

ERDAS Field Guide[™]

Fifth Edition, Revised and Expanded

ERDAS[®], Inc.
Atlanta, Georgia

Copyright © 1982 - 1999 by ERDAS[®], Inc. All rights reserved.

First Edition published 1990. Second Edition published 1991.
Third Edition reprinted 1995. Fourth Edition printed 1997.
Fifth Edition printed 1999.

Printed in the United States of America.

ERDAS proprietary - copying and disclosure prohibited without express permission from ERDAS, Inc.

ERDAS Worldwide Headquarters
2801 Buford Highway, NE
Atlanta, Georgia 30329-2137 USA
www.erdas.com
Toll Free: 877/GO ERDAS (463-7327)
Phone: 404/248-9000
Fax: 404/248-9400
User Support: 404/248-9777

ERDAS in Europe, Africa, Middle East, Asia/Pacific Rim
Phone: 011 44 1223 881 774
Fax: 011 44 1223 880 160

All Other Worldwide Inquiries
Phone: 404/248-9000
Fax: 404/248-9400

The information in this document is subject to change without notice.

Acknowledgments

The *ERDAS Field Guide*[™] was originally researched, written, edited, and designed by Chris Smith and Nicki Brown of ERDAS, Inc. The Second Edition was produced by Chris Smith, Nicki Brown, Nancy Pyden, and Dana Wormer of ERDAS, Inc., with assistance from Diana Margaret and Susanne Strater. The Third Edition was written and edited by Chris Smith, Nancy Pyden, and Pam Cole of ERDAS, Inc. The fourth edition was written and edited by Stacey Schrader and Russ Pouncey of ERDAS, Inc. The fifth edition was written and edited by Russ Pouncey, Kris Swanson, and Kathy Hart of ERDAS, Inc. Many, many thanks go to Brad Skelton, ERDAS Engineering Director, and the ERDAS Software Engineers for their significant contributions to this and previous editions. Without them this manual would not have been possible. Thanks also to Derrold Holcomb for lending his expertise on the Enhancement chapter. Many others at ERDAS provided valuable comments and suggestions in an extensive review process.

A special thanks to those industry experts who took time out of their hectic schedules to review previous editions of the *ERDAS Field Guide*. Of these "external" reviewers, Russell G. Congalton, D. Cunningham, Thomas Hack, Michael E. Hodgson, David McKinsey, and D. Way deserve recognition for their contributions to previous editions.

Copyrights

Copyright 1999 ERDAS, Inc. All rights reserved. ERDAS, ERDAS, Inc., and ERDAS IMAGINE are registered trademarks; CellArray, IMAGINE Developers' Toolkit, IMAGINE Expert Classifier, IMAGINE IFSAR DEM, IMAGINE NITF, IMAGINE OrthoBASE, IMAGINE OrthoMAX, IMAGINE OrthoRadar, IMAGINE Radar Interpreter, IMAGINE Radar Mapping Suite, IMAGINE StereoSAR DEM, and IMAGINE Vector are trademarks. Other brand and product names are the properties of their respective owners. 11/99 Part No. 27.

Table of Contents

Table of Contents	iii
List of Figures	xv
List of Tables	xxi
Preface	xxv
Introduction	xxv
Conventions Used in this Book	xxv
CHAPTER 1	
Raster Data	
Introduction	1
Image Data	1
Bands	2
Coordinate Systems	3
Remote Sensing	4
Absorption/Reflection Spectra	5
Resolution	14
Spectral Resolution	14
Spatial Resolution	15
Radiometric Resolution	16
Temporal Resolution	17
Data Correction	17
Line Dropout	18
Striping	18
Data Storage	18
Storage Formats	18
Storage Media	21
Calculating Disk Space	23
ERDAS IMAGINE Format (.img)	24
Image File Organization	27
Consistent Naming Convention	27
Keeping Track of Image Files	28
Geocoded Data	28
Using Image Data in GIS	29
Subsetting and Mosaicking	29
Enhancement	30
Multispectral Classification	30
Editing Raster Data	30
Editing Continuous (Athematic) Data	31
Interpolation Techniques	31

Table of Contents

CHAPTER 2

Vector Layers

Introduction	35
Coordinates	37
Vector Layers	37
Topology	37
Vector Files	38
Attribute Information	39
Displaying Vector Data	40
Symbolization	41
Vector Data Sources	42
Digitizing	43
Tablet Digitizing	43
Screen Digitizing	44
Imported Vector Data	45
Raster to Vector Conversion	45
Other Vector Data Types	46
Shapefile Vector Format	46
SDE	46
SDTS	47

CHAPTER 3

Raster and Vector Data Sources

Introduction	49
Importing and Exporting	49
Raster Data	49
Raster Data Sources	53
Annotation Data	54
Generic Binary Data	55
Vector Data	55
Satellite Data	56
Satellite System	57
Satellite Characteristics	57
IKONOS	58
IRS	59
Landsat 1-5	60
Landsat 7	64
NLAPS	65
NOAA Polar Orbiter Data	65
OrbView3	67
SeaWiFS	67
SPOT	68
SPOT4	70
Radar Data	70
Advantages of Using Radar Data	70
Radar Sensors	71
Speckle Noise	73

Applications for Radar Data	74
Current Radar Sensors	74
Future Radar Sensors	79
Image Data from Aircraft	79
AIRSAR	79
AVIRIS	80
DAEDALUS TMS	80
Image Data from Scanning	81
Photogrammetric Scanners	81
Desktop Scanners	81
Aerial Photography	82
DOQs	82
ADRG Data	82
ARC System	83
ADRG File Format	83
.OVR (overview)	83
.IMG (scanned image data)	85
.Lxx (legend data)	85
ADRG File Naming Convention	87
ADRI Data	88
.OVR (overview)	89
.IMG (scanned image data)	89
ADRI File Naming Convention	89
Raster Product Format	90
CIB	91
CADRG	92
Topographic Data	92
DEM	93
DTED	94
Using Topographic Data	94
GPS Data	94
Introduction	94
Satellite Position	95
Differential Correction	95
Applications of GPS Data	96
Ordering Raster Data	97
Addresses to Contact	98
Raster Data from Other Software Vendors	100
ERDAS Ver. 7.X	100
GRID and GRID Stacks	101
JFIF (JPEG)	101
MrSID	102
SDTS	102
SUN Raster	102
TIFF	103
GeoTIFF	104
Vector Data from Other Software Vendors	105
ARCGEN	106

Table of Contents

AutoCAD (DXF)	106
DLG	107
ETAK	108
IGES	108
TIGER	109

CHAPTER 4

Image Display

Introduction	111
Display Memory Size	111
Pixel	112
Colors	112
Colormap and Colorcells	113
Display Types	114
8-bit PseudoColor	115
24-bit DirectColor	115
24-bit TrueColor	116
PC Displays	117
Displaying Raster Layers	118
Continuous Raster Layers	118
Thematic Raster Layers	123
Using the Viewer	126
Pyramid Layers	127
Dithering	129
Viewing Layers	130
Viewing Multiple Layers	131
Linking Viewers	132
Zoom and Roam	133
Geographic Information	134
Enhancing Continuous Raster Layers	134
Creating New Image Files	135

CHAPTER 5

Enhancement

Introduction	137
Display vs. File Enhancement	137
Spatial Modeling Enhancements	138
Correcting Data	140
Radiometric Correction: Visible/Infrared Imagery	141
Atmospheric Effects	142
Geometric Correction	143
Radiometric Enhancement	143
Contrast Stretching	144
Histogram Equalization	149
Histogram Matching	153
Brightness Inversion	153
Spatial Enhancement	154
Convolution Filtering	155
Crisp	159

Resolution Merge	160
Adaptive Filter	162
Spectral Enhancement	163
Principal Components Analysis	164
Decorrelation Stretch	168
Tasseled Cap	168
RGB to IHS	169
IHS to RGB	172
Indices	173
Hyperspectral Image Processing.....	176
Normalize	177
IAR Reflectance	177
Log Residuals	178
Rescale	178
Processing Sequence	179
Spectrum Average	179
Signal to Noise	180
Mean per Pixel	180
Profile Tools	180
Wavelength Axis	182
Spectral Library	182
Classification	182
System Requirements	183
Fourier Analysis	183
FFT	185
Fourier Magnitude	185
IFFT	188
Filtering	189
Windows	191
Fourier Noise Removal	194
Homomorphic Filtering	195
Radar Imagery Enhancement.....	196
Speckle Noise	197
Edge Detection	204
Texture	208
Radiometric Correction: Radar Imagery	210
Slant-to-Ground Range Correction	212
Merging Radar with VIS/IR Imagery	213

CHAPTER 6

Classification

Introduction	217
The Classification Process	217
Training	217
Signatures	218
Decision Rule	218
Classification Tips	219
Classification Scheme	219
Iterative Classification	220

Table of Contents

Supervised vs. Unsupervised Training	220
Classifying Enhanced Data	220
Dimensionality	220
Supervised Training	221
Training Samples and Feature Space Objects	221
Selecting Training Samples	222
Evaluating Training Samples	224
Selecting Feature Space Objects	224
Unsupervised Training	227
ISODATA Clustering	227
RGB Clustering	232
Signature Files	234
Evaluating Signatures	235
Alarm	236
Ellipse	236
Contingency Matrix	237
Separability	238
Signature Manipulation	241
Classification Decision Rules	242
Nonparametric Rules	242
Parametric Rules	243
Parallelepiped	244
Feature Space	246
Minimum Distance	247
Mahalanobis Distance	249
Maximum Likelihood/Bayesian	250
Fuzzy Methodology	251
Fuzzy Classification	251
Fuzzy Convolution	252
Expert Classification	253
Knowledge Engineer	253
Knowledge Classifier	255
Evaluating Classification	256
Thresholding	256
Accuracy Assessment	258
Output File	260

CHAPTER 7

Photogrammetric Concepts

Introduction	261
What is Photogrammetry?	261
Types of Photographs and Images	262
Why use Photogrammetry?	262
Photogrammetry vs. Conventional Geometric Correction	263
Single Frame Orthorectification vs. Block Triangulation	263
Image and Data Acquisition	265
Photogrammetric Scanners	266

Desktop Scanners	267
Scanning Resolutions	267
Coordinate Systems	269
Terrestrial Photography	271
Interior Orientation	272
Principal Point and Focal Length	273
Fiducial Marks	273
Lens Distortion	274
Exterior Orientation.	276
The Collinearity Equation	277
Photogrammetric Solutions.	279
Space Resection	279
Space Forward Intersection	280
Bundle Block Adjustment	281
Least Squares Adjustment	284
Self-calibrating Bundle Adjustment	286
Automatic Gross Error Detection	287
GCPs	287
GCP Requirements	288
Processing Multiple Strips of Imagery	289
Tie Points	290
Automatic Tie Point Collection	291
Image Matching Techniques	292
Area Based Matching	292
Feature Based Matching	294
Relation Based Matching	295
Image Pyramid	295
Satellite Photogrammetry	296
SPOT Interior Orientation	298
SPOT Exterior Orientation	299
Collinearity Equations and Satellite Block Triangulation	302
Orthorectification	303

CHAPTER 8
Radar Concepts

Introduction	307
IMAGINE OrthoRadar Theory	307
Algorithm Description	309
IMAGINE StereoSAR DEM Theory	315
Introduction	315
Input	316
Subset	318
Despeckle	318
Degrade	319
Register	319
Constrain	320
Match	320
Degrade	325

Table of Contents

Height	325
IMAGINE IFSAR DEM Theory.....	326
Introduction	326
Electromagnetic Wave Background	326
The Interferometric Model	329
Image Registration	334
Phase Noise Reduction	336
Phase Flattening	337
Phase Unwrapping	338
Conclusions	342
CHAPTER 9	
Rectification	
Introduction	343
Orthorectification	344
When to Rectify	344
When to Georeference Only	345
Disadvantages of Rectification	346
Rectification Steps	346
Ground Control Points	346
GCPs in ERDAS IMAGINE	347
Entering GCPs	347
GCP Prediction and Matching	348
Polynomial Transformation.....	349
Linear Transformations	350
Nonlinear Transformations	352
Effects of Order	354
Minimum Number of GCPs	359
Rubber Sheeting	360
Triangle-Based Finite Element Analysis	360
Triangulation	360
Triangle-based rectification	361
Linear transformation	361
Nonlinear transformation	361
Check Point Analysis	362
RMS Error	362
Residuals and RMS Error Per GCP	362
Total RMS Error	363
Error Contribution by Point	363
Tolerance of RMS Error	364
Evaluating RMS Error	364
Resampling Methods	365
Rectifying to Lat/Lon	366
Nearest Neighbor	366
Bilinear Interpolation	367
Cubic Convolution	370
Map-to-Map Coordinate Conversions	373

CHAPTER 10
Terrain Analysis

Introduction 375

Topographic Data 376

Slope Images 376

Aspect Images 379

Shaded Relief 381

Topographic Normalization 382

 Lambertian Reflectance Model 383

 Non-Lambertian Model 384

CHAPTER 11
Geographic Information Systems

Introduction 385

Data Input 387

Continuous Layers 388

Thematic Layers 389

 Statistics 390

Vector Layers 391

Attributes 391

 Raster Attributes 391

 Vector Attributes 393

Analysis 394

 ERDAS IMAGINE Analysis Tools 394

 Analysis Procedures 395

Proximity Analysis 396

Contiguity Analysis 397

Neighborhood Analysis 398

Recoding 400

Overlaying 401

Indexing 402

Matrix Analysis 404

Modeling 404

Graphical Modeling 405

 Model Maker Functions 408

 Objects 409

 Data Types 410

 Output Parameters 410

 Using Attributes in Models 411

Script Modeling 412

 Statements 413

 Data Types 414

 Variables 415

Table of Contents

Vector Analysis	415
Editing Vector Layers	415
Constructing Topology (Coverages Only)	416
Building and Cleaning Coverages	417
CHAPTER 12	
Cartography	
Introduction	419
Types of Maps	419
Thematic Maps	421
Annotation	422
Scale	423
Legends	427
Neatlines, Tick Marks, and Grid Lines	427
Symbols	428
Labels and Descriptive Text	429
Typography and Lettering	430
Map Projections	432
Properties of Map Projections	433
Projection Types	434
Geographical and Planar Coordinates	437
Available Map Projections	438
Choosing a Map Projection	443
Spheroids	444
Map Composition	447
Map Accuracy	448
CHAPTER 13	
Hardcopy Output	
Introduction	451
Printing Maps	451
Scale and Resolution	452
Map Scaling Examples	453
Mechanics of Printing	455
Halftone Printing	455
Continuous Tone Printing	455
Contrast and Color Tables	456
RGB to CMY Conversion	456
Appendix A	
Math Topics	
Introduction	459
Summation	459
Statistics	459



Histogram	459
Bin Functions	460
Mean	462
Normal Distribution	463
Variance	464
Standard Deviation	465
Parameters	466
Covariance	466
Covariance Matrix	467
Dimensionality of Data	467
Measurement Vector	467
Mean Vector	468
Feature Space	469
Feature Space Images	470
n-Dimensional Histogram	470
Spectral Distance	471
Polynomials	471
Order	471
Transformation Matrix	472
Matrix Algebra	473
Matrix Notation	473
Matrix Multiplication	473
Transposition	474

Appendix B Map Projections

Introduction	477
USGS Projections	477
Alaska Conformal	480
Albers Conical Equal Area	481
Azimuthal Equidistant	483
Behrmann	486
Bonne	487
Cassini	489
Eckert I	490
Eckert II	492
Eckert III	493
Eckert IV	495
Eckert V	496
Eckert VI	498
EOSAT SOM	499
Equidistant Conic	500
Equidistant Cylindrical	502
Equirectangular (Plate Carrée)	503
Gall Stereographic	505
Gauss Kruger	506
General Vertical Near-side Perspective	507
Geographic (Lat/Lon)	508
Gnomonic	510
Hammer	511

Table of Contents

Interrupted Goode Homolosine	513
Interrupted Mollweide	514
Lambert Azimuthal Equal Area	515
Lambert Conformal Conic	518
Loximuthal	521
Mercator	523
Miller Cylindrical	525
Modified Transverse Mercator	528
Mollweide	529
New Zealand Map Grid	530
Oblated Equal Area	531
Oblique Mercator (Hotine)	532
Orthographic	535
Plate Carrée	537
Polar Stereographic	539
Polyconic	542
Quartic Authalic	544
Robinson	545
RSO	547
Sinusoidal	548
Space Oblique Mercator	550
Space Oblique Mercator (Formats A & B)	552
State Plane	553
Stereographic	563
Stereographic (Extended)	565
Transverse Mercator	566
Two Point Equidistant	567
UTM	569
Van der Grinten I	571
Wagner IV	574
Wagner VII	575
Winkel I	577
External Projections	578
Bipolar Oblique Conic Conformal	580
Cassini-Soldner	581
Laborde Oblique Mercator	582
Minimum Error Conformal	582
Modified Polyconic	583
Modified Stereographic	584
Mollweide Equal Area	585
Rectified Skew Orthomorphic	586
Robinson Pseudocylindrical	586
Southern Orientated Gauss Conformal	587
Swiss Cylindrical	587
Winkel's Tripel	587
Glossary	589
Bibliography	639
Index	653

List of Figures

Figure 1-1	Pixels and Bands in a Raster Image	2
Figure 1-2	Typical File Coordinates	4
Figure 1-3	Electromagnetic Spectrum	5
Figure 1-4	Sun Illumination Spectral Irradiance at the Earth's Surface	7
Figure 1-5	Factors Affecting Radiation	8
Figure 1-6	Reflectance Spectra	10
Figure 1-7	Laboratory Spectra of Clay Minerals in the Infrared Region	12
Figure 1-8	IFOV	16
Figure 1-9	Brightness Values	17
Figure 1-10	Landsat TM—Band 2 (Four Types of Resolution)	17
Figure 1-11	Band Interleaved by Line (BIL)	19
Figure 1-12	Band Sequential (BSQ)	20
Figure 1-13	Image Files Store Raster Layers	24
Figure 1-14	Example of a Thematic Raster Layer	25
Figure 1-15	Examples of Continuous Raster Layers	26
Figure 2-1	Vector Elements	35
Figure 2-2	Vertices	36
Figure 2-3	Workspace Structure	39
Figure 2-4	Attribute CellArray	40
Figure 2-5	Symbolization Example	42
Figure 2-6	Digitizing Tablet	43
Figure 2-7	Raster Format Converted to Vector Format	45
Figure 3-1	Multispectral Imagery Comparison	58
Figure 3-2	Landsat MSS vs. Landsat TM	63
Figure 3-3	SPOT Panchromatic vs. SPOT XS	69
Figure 3-4	SLAR Radar	71
Figure 3-5	Received Radar Signal	72
Figure 3-6	Radar Reflection from Different Sources and Distances	72
Figure 3-7	ADRG Overview File Displayed in a Viewer	84
Figure 3-8	Subset Area with Overlapping ZDRs	85
Figure 3-9	Seamless Nine Image DR	88
Figure 3-10	ADRI Overview File Displayed in a Viewer	89
Figure 3-11	Arc/second Format	93
Figure 3-12	Common Uses of GPS Data	97
Figure 4-1	Example of One Seat with One Display and Two Screens	111
Figure 4-2	Transforming Data File Values to a Colorcell Value	115
Figure 4-3	Transforming Data File Values to a Colorcell Value	116
Figure 4-4	Transforming Data File Values to Screen Values	117
Figure 4-5	Contrast Stretch and Colorcell Values	120
Figure 4-6	Stretching by Min/Max vs. Standard Deviation	121
Figure 4-7	Continuous Raster Layer Display Process	122

List of Figures

Figure 4-8	Thematic Raster Layer Display Process	125
Figure 4-9	Pyramid Layers	128
Figure 4-10	Example of Dithering	129
Figure 4-11	Example of Color Patches	130
Figure 4-12	Linked Viewers	133
Figure 5-1	Histograms of Radiometrically Enhanced Data	144
Figure 5-2	Graph of a Lookup Table	145
Figure 5-3	Enhancement with Lookup Tables	145
Figure 5-4	Nonlinear Radiometric Enhancement	146
Figure 5-5	Piecewise Linear Contrast Stretch	147
Figure 5-6	Contrast Stretch Using Lookup Tables, and Effect on Histogram	149
Figure 5-7	Histogram Equalization	150
Figure 5-8	Histogram Equalization Example	151
Figure 5-9	Equalized Histogram	152
Figure 5-10	Histogram Matching	153
Figure 5-11	Spatial Frequencies	154
Figure 5-12	Applying a Convolution Kernel	156
Figure 5-13	Output Values for Convolution Kernel	157
Figure 5-14	Local Luminance Intercept	163
Figure 5-15	Two Band Scatterplot	164
Figure 5-16	First Principal Component	165
Figure 5-17	Range of First Principal Component	165
Figure 5-18	Second Principal Component	166
Figure 5-19	Intensity, Hue, and Saturation Color Coordinate System	170
Figure 5-20	Hyperspectral Data Axes	176
Figure 5-21	Rescale Graphical User Interface (GUI)	178
Figure 5-22	Spectrum Average GUI	179
Figure 5-23	Spectral Profile	180
Figure 5-24	Two-Dimensional Spatial Profile	181
Figure 5-25	Three-Dimensional Spatial Profile	181
Figure 5-26	Surface Profile	182
Figure 5-27	One-Dimensional Fourier Analysis	184
Figure 5-28	Example of Fourier Magnitude	186
Figure 5-29	The Padding Technique	188
Figure 5-30	Comparison of Direct and Fourier Domain Processing	190
Figure 5-31	An Ideal Cross Section	192
Figure 5-32	High-Pass Filtering Using the Ideal Window	192
Figure 5-33	Filtering Using the Bartlett Window	193
Figure 5-34	Filtering Using the Butterworth Window	193
Figure 5-35	Homomorphic Filtering Process	196
Figure 5-36	Effects of Mean and Median Filters	199
Figure 5-37	Regions of Local Region Filter	200
Figure 5-38	One-dimensional, Continuous Edge, and Line Models	204
Figure 5-39	A Noisy Edge Superimposed on an Ideal Edge	205

Figure 5-40	Edge and Line Derivatives	205
Figure 5-41	Adjust Brightness Function	211
Figure 5-42	Range Lines vs. Lines of Constant Range	212
Figure 5-43	Slant-to-Ground Range Correction	212
Figure 6-1	Example of a Feature Space Image	224
Figure 6-2	Process for Defining a Feature Space Object	226
Figure 6-3	ISODATA Arbitrary Clusters	229
Figure 6-4	ISODATA First Pass	229
Figure 6-5	ISODATA Second Pass	230
Figure 6-6	RGB Clustering.	233
Figure 6-7	Ellipse Evaluation of Signatures	237
Figure 6-8	Classification Flow Diagram.	243
Figure 6-9	Parallelepiped Classification Using Two Standard Deviations as Limits.	244
Figure 6-10	Parallelepiped Corners Compared to the Signature Ellipse	246
Figure 6-11	Feature Space Classification	246
Figure 6-12	Minimum Spectral Distance	248
Figure 6-13	Knowledge Engineer Editing Window	253
Figure 6-14	Example of a Decision Tree Branch	254
Figure 6-15	Split Rule Decision Tree Branch.	254
Figure 6-16	Knowledge Classifier Classes of Interest	255
Figure 6-17	Histogram of a Distance Image.	256
Figure 6-18	Interactive Thresholding Tips.	257
Figure 7-1	Exposure Stations Along a Flight Path.	265
Figure 7-2	A Regular Rectangular Block of Aerial Photos	266
Figure 7-3	Pixel Coordinates and Image Coordinates	269
Figure 7-4	Image Space Coordinate System	270
Figure 7-5	Terrestrial Photography	271
Figure 7-6	Internal Geometry.	272
Figure 7-7	Pixel Coordinate System vs. Image Space Coordinate System.	273
Figure 7-8	Radial vs. Tangential Lens Distortion.	275
Figure 7-9	Elements of Exterior Orientation	276
Figure 7-10	Space Forward Intersection	280
Figure 7-11	Photogrammetric Configuration	282
Figure 7-12	GCP Configuration	289
Figure 7-13	GCPs in a Block of Images	289
Figure 7-14	Point Distribution for Triangulation	290
Figure 7-15	Tie Points in a Block.	291
Figure 7-16	Image Pyramid for Matching at Coarse to Full Resolution	295
Figure 7-17	Perspective Centers of SPOT Scan Lines	296
Figure 7-18	Image Coordinates in a Satellite Scene	297
Figure 7-19	Interior Orientation of a SPOT Scene	298
Figure 7-20	Inclination of a Satellite Stereo-Scene (View from North to South)	300
Figure 7-21	Velocity Vector and Orientation Angle of a Single Scene	301
Figure 7-22	Ideal Point Distribution Over a Satellite Scene for Triangulation	302

List of Figures

Figure 7-23	Orthorectification	303
Figure 7-24	Digital Orthophoto—Finding Gray Values	304
Figure 8-1	Doppler Cone	312
Figure 8-2	Sparse Mapping and Output Grids	314
Figure 8-3	IMAGINE StereoSAR DEM Process Flow	315
Figure 8-4	SAR Image Intersection	316
Figure 8-5	UL Corner of the Reference Image	321
Figure 8-6	UL Corner of the Match Image	321
Figure 8-7	Image Pyramid	322
Figure 8-8	Electromagnetic Wave	327
Figure 8-9	Variation of Electric Field in Time	327
Figure 8-10	Effect of Time and Distance on Energy	328
Figure 8-11	Geometric Model for an Interferometric SAR System	329
Figure 8-12	Differential Collection Geometry	333
Figure 8-13	Interferometric Phase Image without Filtering	336
Figure 8-14	Interferometric Phase Image with Filtering	337
Figure 8-15	Interferometric Phase Image without Phase Flattening	338
Figure 8-16	Electromagnetic Wave Traveling through Space	338
Figure 8-17	One-dimensional Continuous vs. Wrapped Phase Function	339
Figure 8-18	Sequence of Unwrapped Phase Images	340
Figure 8-19	Wrapped vs. Unwrapped Phase Images	341
Figure 9-1	Polynomial Curve vs. GCPs	350
Figure 9-2	Linear Transformations	351
Figure 9-3	Nonlinear Transformations	352
Figure 9-4	Transformation Example—1st-Order	355
Figure 9-5	Transformation Example—2nd GCP Changed	356
Figure 9-6	Transformation Example—2nd-Order	357
Figure 9-7	Transformation Example—4th GCP Added	357
Figure 9-8	Transformation Example—3rd-Order	358
Figure 9-9	Transformation Example—Effect of a 3rd-Order Transformation	358
Figure 9-10	Triangle Network	360
Figure 9-11	Residuals and RMS Error Per Point	363
Figure 9-12	RMS Error Tolerance	364
Figure 9-13	Resampling	365
Figure 9-14	Nearest Neighbor	367
Figure 9-15	Bilinear Interpolation	368
Figure 9-16	Linear Interpolation	368
Figure 9-17	Cubic Convolution	371
Figure 10-1	Regularly Spaced Terrain Data Points	376
Figure 10-2	3 × 3 Window Calculates the Slope at Each Pixel	377
Figure 10-3	Slope Calculation Example	378
Figure 10-4	3 × 3 Window Calculates the Aspect at Each Pixel	380
Figure 10-5	Aspect Calculation Example	381
Figure 10-6	Shaded Relief	382

Figure 11-1	Data Input	387
Figure 11-2	Raster Attributes for Inlandc.img	392
Figure 11-3	Vector Attributes CellArray	394
Figure 11-4	Proximity Analysis	396
Figure 11-5	Contiguity Analysis	397
Figure 11-6	Using a Mask	399
Figure 11-7	Sum Option of Neighborhood Analysis (Image Interpreter)	400
Figure 11-8	Overlay	402
Figure 11-9	Indexing	403
Figure 11-10	Graphical Model for Sensitivity Analysis	406
Figure 11-11	Graphical Model Structure	407
Figure 11-12	Modeling Objects	410
Figure 11-13	Graphical and Script Models For Tasseled Cap Transformation	413
Figure 11-14	Layer Errors	418
Figure 12-1	Sample Scale Bars	424
Figure 12-2	Sample Legend	427
Figure 12-3	Sample Neatline, Tick Marks, and Grid Lines	428
Figure 12-4	Sample Symbols	429
Figure 12-5	Sample Sans Serif and Serif Typefaces with Various Styles Applied	431
Figure 12-6	Good Lettering vs. Bad Lettering	432
Figure 12-7	Projection Types	434
Figure 12-8	Tangent and Secant Cones	436
Figure 12-9	Tangent and Secant Cylinders	436
Figure 12-10	Ellipse	444
Figure 13-1	Layout for a Book Map and a Paneled Map	452
Figure 13-2	Sample Map Composition	453
Figure A-1	Histogram	460
Figure A-2	Normal Distribution	463
Figure A-3	Measurement Vector	468
Figure A-4	Mean Vector	468
Figure A-5	Two Band Plot	469
Figure A-6	Two-band Scatterplot	470
Figure B-1	Albers Conical Equal Area Projection	483
Figure B-2	Polar Aspect of the Azimuthal Equidistant Projection	485
Figure B-3	Behrmann Cylindrical Equal-Area Projection	487
Figure B-4	Bonne Projection	488
Figure B-5	Cassini Projection	490
Figure B-6	Eckert I Projection	491
Figure B-7	Eckert II Projection	493
Figure B-8	Eckert III Projection	494
Figure B-9	Eckert IV Projection	496
Figure B-10	Eckert V Projection	497
Figure B-11	Eckert VI Projection	499
Figure B-12	Equidistant Conic Projection	502

List of Figures

Figure B-13	Equirectangular Projection	504
Figure B-14	Geographic Projection	509
Figure B-15	Hammer Projection	512
Figure B-16	Interrupted Goode Homolosine Projection	514
Figure B-17	Interrupted Mollweide Projection	515
Figure B-18	Lambert Azimuthal Equal Area Projection	517
Figure B-19	Lambert Conformal Conic Projection	520
Figure B-20	Loximuthal Projection.	522
Figure B-21	Mercator Projection	525
Figure B-22	Miller Cylindrical Projection	527
Figure B-23	Mollweide Projection	530
Figure B-24	Oblique Mercator Projection	534
Figure B-25	Orthographic Projection	537
Figure B-26	Plate Carrée Projection.	538
Figure B-27	Polar Stereographic Projection and its Geometric Construction	541
Figure B-28	Polyconic Projection of North America.	543
Figure B-29	Quartic Authalic Projection.	545
Figure B-30	Robinson Projection	546
Figure B-31	Sinusoidal Projection	549
Figure B-32	Space Oblique Mercator Projection	551
Figure B-33	Zones of the State Plane Coordinate System	554
Figure B-34	Stereographic Projection	565
Figure B-35	Two Point Equidistant Projection	569
Figure B-36	Zones of the Universal Transverse Mercator Grid in the United States	570
Figure B-37	Van der Grinten I Projection	573
Figure B-38	Wagner IV Projection	575
Figure B-39	Wagner VII Projection	576
Figure B-40	Winkel I Projection	578
Figure B-41	Winkel's Tripel Projection.	588

List of Tables

Table 1-1	Bandwidths Used in Remote Sensing	13
Table 2-1	Description of File Types	38
Table 3-1	Raster Data Formats	50
Table 3-2	Annotation Data Formats	54
Table 3-3	Vector Data Formats	55
Table 3-4	IKONOS Bands and Frequencies	59
Table 3-5	LISS-III Bands and Frequencies	59
Table 3-6	Panchromatic Band and Frequency	60
Table 3-7	WiFS Bands and Frequencies	60
Table 3-8	Landsat 7 Characteristics	64
Table 3-9	Orb View-3 Bands and Frequencies	67
Table 3-10	SeaWiFS Bands and Frequencies	68
Table 3-11	SPOT 4 Bands and Frequencies	70
Table 3-12	Commonly Used Bands for Radar Imaging	73
Table 3-13	Current Radar Sensors	74
Table 3-14	JERS-1 Bands and Frequencies	76
Table 3-15	RADARSAT Beam Mode Resolution	77
Table 3-16	SIR-C/X-SAR Bands and Frequencies	79
Table 3-17	DAEDALUS TMS Bands and Frequencies	80
Table 3-18	ARC System Chart Types	86
Table 3-19	Legend Files for the ARC System Chart Types	86
Table 3-20	Common Raster Data Products	97
Table 3-21	File Types Created by Screendump	103
Table 3-22	The Most Common TIFF Format Elements	104
Table 3-23	Conversion of DXF Entries	107
Table 3-24	Conversion of IGES Entities	109
Table 4-1	Colorcell Example	113
Table 4-2	Commonly Used RGB Colors	124
Table 4-3	Overview of Zoom Ratio	134
Table 5-1	Description of Modeling Functions Available for Enhancement	139
Table 5-2	Theoretical Coefficient of Variation Values	201
Table 5-3	Parameters for Sigma Filter	202
Table 5-4	Pre-Classification Sequence	202
Table 6-1	Training Sample Comparison	224
Table 6-2	Feature Space Signatures	227
Table 6-3	ISODATA Clustering	231
Table 6-4	RGB Clustering	234
Table 6-5	Parallelepiped Decision Rule	245
Table 6-6	Feature Space Decision Rule	247
Table 6-7	Minimum Distance Decision Rule	249
Table 6-8	Mahalanobis Decision Rule	250

List of Tables

Table 6-9	Maximum Likelihood/Bayesian Decision Rule	251
Table 7-1	Scanning Resolutions	268
Table 8-1	SAR Parameters Required for Orthorectification	307
Table 8-2	STD_LP_HD Correlator	323
Table 9-1	Number of GCPs per Order of Transformation	359
Table 9-2	Nearest Neighbor Resampling	367
Table 9-3	Bilinear Interpolation Resampling	370
Table 9-4	Cubic Convolution Resampling	372
Table 11-1	Example of a Recoded Land Cover Layer	401
Table 11-2	Model Maker Functions	408
Table 11-3	Attribute Information for parks.img	411
Table 11-4	General Editing Operations and Supporting Feature Types	416
Table 11-5	Comparison of Building and Cleaning Coverages	417
Table 12-1	Types of Maps	420
Table 12-2	Common Map Scales	424
Table 12-3	Pixels per Inch	425
Table 12-4	Acres and Hectares per Pixel	426
Table 12-5	Map Projections	441
Table 12-6	Projection Parameters	442
Table 12-7	Spheroids for use with ERDAS IMAGINE	445
Table B-1	Alaska Conformal Summary	480
Table B-2	Albers Conical Equal Area Summary	481
Table B-3	Azimuthal Equidistant Summary	483
Table B-4	Behrmann Summary	486
Table B-5	Bonne Summary	487
Table B-6	Cassini Summary	489
Table B-7	Eckert I Summary	490
Table B-8	Eckert II Summary	492
Table B-9	Eckert III Summary	493
Table B-10	Eckert IV Summary	495
Table B-11	Eckert V Summary	496
Table B-12	Eckert VI Summary	498
Table B-13	Equidistant Conic Summary	500
Table B-14	Equiarectangular (Plate Carrée) Summary	503
Table B-15	Gall Stereographic Summary	505
Table B-16	General Vertical Near-side Perspective Summary	507
Table B-17	Gnomonic Summary	510
Table B-18	Hammer Summary	511
Table B-19	Interrupted Goode Homolosine	513
Table B-20	Lambert Azimuthal Equal Area Summary	515
Table B-21	Lambert Conformal Conic Summary	518
Table B-22	Loximuthal Summary	521
Table B-23	Mercator Summary	523
Table B-24	Miller Cylindrical Summary	525

Table B-25	Modified Transverse Mercator Summary	528
Table B-26	Mollweide Summary	529
Table B-27	New Zealand Map Grid Summary	530
Table B-28	Oblique Mercator (Hotine) Summary	532
Table B-29	Orthographic Summary	535
Table B-30	Polar Stereographic Summary	539
Table B-31	Polyconic Summary	542
Table B-32	Quartic Authalic	544
Table B-33	Robinson Summary	545
Table B-34	RSO Summary	547
Table B-35	Sinusoidal Summary	548
Table B-36	Space Oblique Mercator Summary	550
Table B-37	NAD27 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States	554
Table B-38	NAD83 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States	559
Table B-39	Stereographic Summary	563
Table B-40	Transverse Mercator Summary	566
Table B-41	Two Point Equidistant Summary	568
Table B-42	UTM Zones, Central Meridians, and Longitude Ranges	571
Table B-43	Van der Grinten I Summary	572
Table B-44	Wagner IV Summary	574
Table B-45	Wagner VII	575
Table B-46	Winkel I Summary	577
Table B-47	Bipolar Oblique Conic Conformal Summary	580
Table B-48	Cassini-Soldner Summary	581
Table B-49	Modified Polyconic Summary	583
Table B-50	Modified Stereographic Summary	584
Table B-51	Mollweide Equal Area Summary	585
Table B-52	Robinson Pseudocylindrical Summary	586
Table B-53	Winkel's Tripel Summary	587

List of Tables

Preface

Introduction

The purpose of the *ERDAS Field Guide*[™] is to provide background information on why one might use particular geographic information system (GIS) and image processing functions and how the software is manipulating the data, rather than what buttons to push to actually perform those functions. This book is also aimed at a diverse audience: from those who are new to geoprocessing to those savvy users who have been in this industry for years. For the novice, the *ERDAS Field Guide* provides a brief history of the field, an extensive glossary of terms, and notes about applications for the different processes described. For the experienced user, the *ERDAS Field Guide* includes the formulas and algorithms that are used in the code, so that he or she can see exactly how each operation works.

Although the *ERDAS Field Guide* is primarily a reference to basic image processing and GIS concepts, it is geared toward ERDAS IMAGINE[®] users and the functions within ERDAS IMAGINE software, such as GIS analysis, image processing, cartography and map projections, graphics display hardware, statistics, and remote sensing. However, in some cases, processes and functions are described that may not be in the current version of the software, but planned for a future release. There may also be functions described that are not available on your system, due to the actual package that you are using.

The enthusiasm with which the first four editions of the *ERDAS Field Guide* were received has been extremely gratifying, both to the authors and to ERDAS[®] as a whole. First conceived as a helpful manual for ERDAS users, the *ERDAS Field Guide* is now being used as a textbook, lab manual, and training guide throughout the world.

The *ERDAS Field Guide* will continue to expand and improve to keep pace with the profession. Suggestions and ideas for future editions are always welcome, and should be addressed to the Technical Writing division of Engineering at ERDAS[®], Inc., in Atlanta, Georgia.

Conventions Used in this Book

The following paragraphs are used throughout the *ERDAS Field Guide* and other ERDAS IMAGINE documentation.



These paragraphs contain strong warnings or important tips.



These paragraphs direct you to the ERDAS IMAGINE software function that accomplishes the described task.



These paragraphs lead you to other chapters in the ERDAS Field Guide or other manuals for additional information.

NOTE: Notes give additional instruction

Preface

CHAPTER 1

Raster Data

Introduction

The ERDAS IMAGINE system incorporates the functions of both image processing and GIS. These functions include importing, viewing, altering, and analyzing raster and vector data sets.

This chapter is an introduction to raster data, including:

- remote sensing
- data storage formats
- different types of resolution
- radiometric correction
- geocoded data
- raster data in GIS



See "CHAPTER 2: Vector Layers" for more information on vector data.

Image Data

In general terms, an image is a digital picture or representation of an object. Remotely sensed image data are digital representations of the Earth. Image data are stored in data files, also called image files, on magnetic tapes, computer disks, or other media. The data consist only of numbers. These representations form images when they are displayed on a screen or are output to hardcopy.

Each number in an image file is a data file value. Data file values are sometimes referred to as pixels. The term pixel is abbreviated from picture element. A pixel is the smallest part of a picture (the area being scanned) with a single value. The data file value is the measured brightness value of the pixel at a specific wavelength.

Raster image data are laid out in a grid similar to the squares on a checkerboard. Each cell of the grid is represented by a pixel, also known as a grid cell.

In remotely sensed image data, each pixel represents an area of the Earth at a specific location. The data file value assigned to that pixel is the record of reflected radiation or emitted heat from the Earth's surface at that location.

Data file values may also represent elevation, as in digital elevation models (DEMs).

NOTE: DEMs are not remotely sensed image data, but are currently being produced from stereo points in radar imagery.



The terms *pixel* and *data file value* are not interchangeable in ERDAS IMAGINE. *Pixel* is used as a broad term with many meanings, one of which is *data file value*. One *pixel* in a file may consist of many *data file values*. When an image is displayed or printed, other types of values are represented by a *pixel*.



See "CHAPTER 4: Image Display" for more information on how images are displayed.

Bands

Image data may include several bands of information. Each band is a set of data file values for a specific portion of the electromagnetic spectrum of reflected light or emitted heat (red, green, blue, near-infrared, infrared, thermal, etc.) or some other user-defined information created by combining or enhancing the original bands, or creating new bands from other sources.

ERDAS IMAGINE programs can handle an unlimited number of bands of image data in a single file.

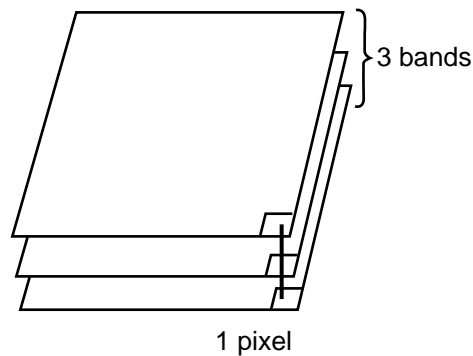


Figure 1-1 Pixels and Bands in a Raster Image



See "CHAPTER 5: Enhancement" for more information on combining or enhancing bands of data.

Bands vs. Layers

In ERDAS IMAGINE, bands of data are occasionally referred to as layers. Once a band is imported into a GIS, it becomes a layer of information which can be processed in various ways. Additional layers can be created and added to the image file (.img extension) in ERDAS IMAGINE, such as layers created by combining existing layers. Read more about image files on page 24.

Layers vs. Viewer Layers

The Viewer permits several images to be layered, in which case each image (including a multiband image) may be a layer.

Numerical Types

The range and the type of numbers used in a raster layer determine how the layer is displayed and processed. For example, a layer of elevation data with values ranging from -51.257 to 553.401 would be treated differently from a layer using only two values to show land and water.

The data file values in raster layers generally fall into these categories:

- Nominal data file values are simply categorized and named. The actual value used for each category has no inherent meaning—it is simply a class value. An example of a nominal raster layer would be a thematic layer showing tree species.
- Ordinal data are similar to nominal data, except that the file values put the classes in a rank or order. For example, a layer with classes numbered and named 1 - Good, 2 - Moderate, and 3 - Poor is an ordinal system.
- Interval data file values have an order, but the intervals between the values are also meaningful. Interval data measure some characteristic, such as elevation or degrees Fahrenheit, which does not necessarily have an absolute zero. (The difference between two values in interval data is meaningful.)
- Ratio data measure a condition that has a natural zero, such as electromagnetic radiation (as in most remotely sensed data), rainfall, or slope.

Nominal and ordinal data lend themselves to applications in which categories, or themes, are used. Therefore, these layers are sometimes called categorical or thematic.

Likewise, interval and ratio layers are more likely to measure a condition, causing the file values to represent continuous gradations across the layer. Such layers are called continuous.

Coordinate Systems

The location of a pixel in a file or on a displayed or printed image is expressed using a coordinate system. In two-dimensional coordinate systems, locations are organized in a grid of columns and rows. Each location on the grid is expressed as a pair of coordinates known as X and Y. The X coordinate specifies the column of the grid, and the Y coordinate specifies the row. Image data organized into such a grid are known as raster data.

There are two basic coordinate systems used in ERDAS IMAGINE:

- file coordinates—indicate the location of a pixel within the image (data file)
- map coordinates—indicate the location of a pixel in a map

File Coordinates

File coordinates refer to the location of the pixels within the image (data) file. File coordinates for the pixel in the upper left corner of the image always begin at 0,0.

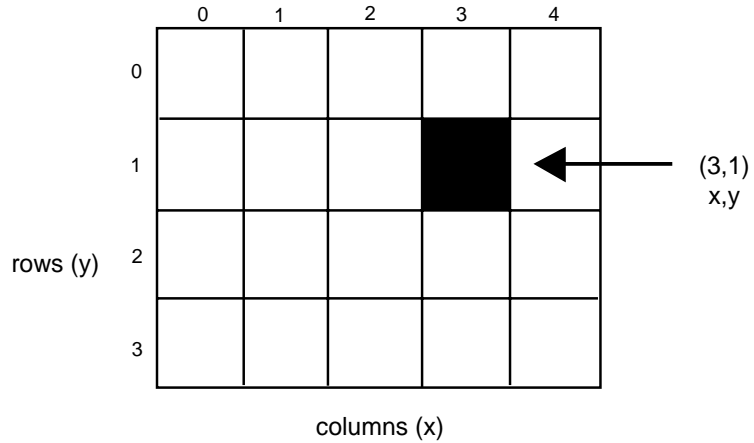


Figure 1-2 Typical File Coordinates

Map Coordinates

Map coordinates may be expressed in one of a number of map coordinate or projection systems. The type of map coordinates used by a data file depends on the method used to create the file (remote sensing, scanning an existing map, etc.). In ERDAS IMAGINE, a data file can be converted from one map coordinate system to another.



For more information on map coordinates and projection systems, see "CHAPTER 12: Cartography" or "Appendix B: Map Projections." See "CHAPTER 9: Rectification" for more information on changing the map coordinate system of a data file.

Remote Sensing

Remote sensing is the acquisition of data about an object or scene by a sensor that is far from the object (Colwell 1983). Aerial photography, satellite imagery, and radar are all forms of remotely sensed data.

Usually, remotely sensed data refer to data of the Earth collected from sensors on satellites or aircraft. Most of the images used as input to the ERDAS IMAGINE system are remotely sensed. However, you are not limited to remotely sensed data.



This section is a brief introduction to remote sensing. There are many books available for more detailed information, including Colwell 1983, Swain and Davis 1978, and Slater 1980 (see "Bibliography").

Electromagnetic Radiation Spectrum

The sensors on remote sensing platforms usually record electromagnetic radiation. Electromagnetic radiation (EMR) is energy transmitted through space in the form of electric and magnetic waves (Star and Estes 1990). Remote sensors are made up of detectors that record specific wavelengths of the electromagnetic spectrum. The electromagnetic spectrum is the range of electromagnetic radiation extending from cosmic waves to radio waves (Jensen 1996).

All types of land cover (rock types, water bodies, etc.) absorb a portion of the electromagnetic spectrum, giving a distinguishable signature of electromagnetic radiation. Armed with the knowledge of which wavelengths are absorbed by certain features and the intensity of the reflectance, you can analyze a remotely sensed image and make fairly accurate assumptions about the scene. Figure 1-3 illustrates the electromagnetic spectrum (Suits 1983; Star and Estes 1990).

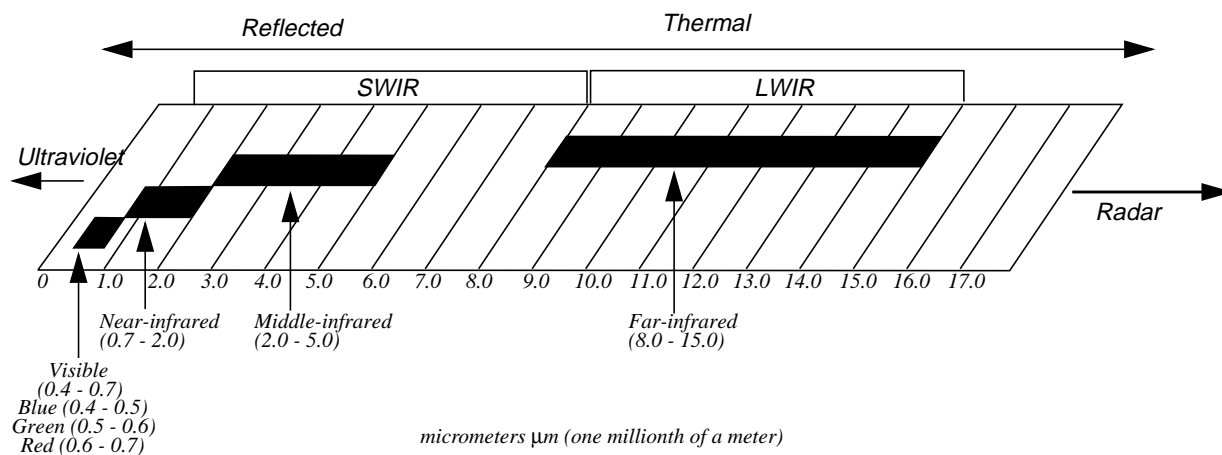


Figure 1-3 Electromagnetic Spectrum

SWIR and LWIR

The near-infrared and middle-infrared regions of the electromagnetic spectrum are sometimes referred to as the short wave infrared region (SWIR). This is to distinguish this area from the thermal or far infrared region, which is often referred to as the long wave infrared region (LWIR). The SWIR is characterized by reflected radiation whereas the LWIR is characterized by emitted radiation.

Absorption/Reflection Spectra

When radiation interacts with matter, some wavelengths are absorbed and others are reflected. To enhance features in image data, it is necessary to understand how vegetation, soils, water, and other land covers reflect and absorb radiation. The study of the absorption and reflection of EMR waves is called spectroscopy.

Spectroscopy

Most commercial sensors, with the exception of imaging radar sensors, are passive solar imaging sensors. Passive solar imaging sensors can only receive radiation waves; they cannot transmit radiation. (Imaging radar sensors are active sensors that emit a burst of microwave radiation and receive the backscattered radiation.)

The use of passive solar imaging sensors to characterize or identify a material of interest is based on the principles of spectroscopy. Therefore, to fully utilize a visible/infrared (VIS/IR) multispectral data set and properly apply enhancement algorithms, it is necessary to understand these basic principles. Spectroscopy reveals the:

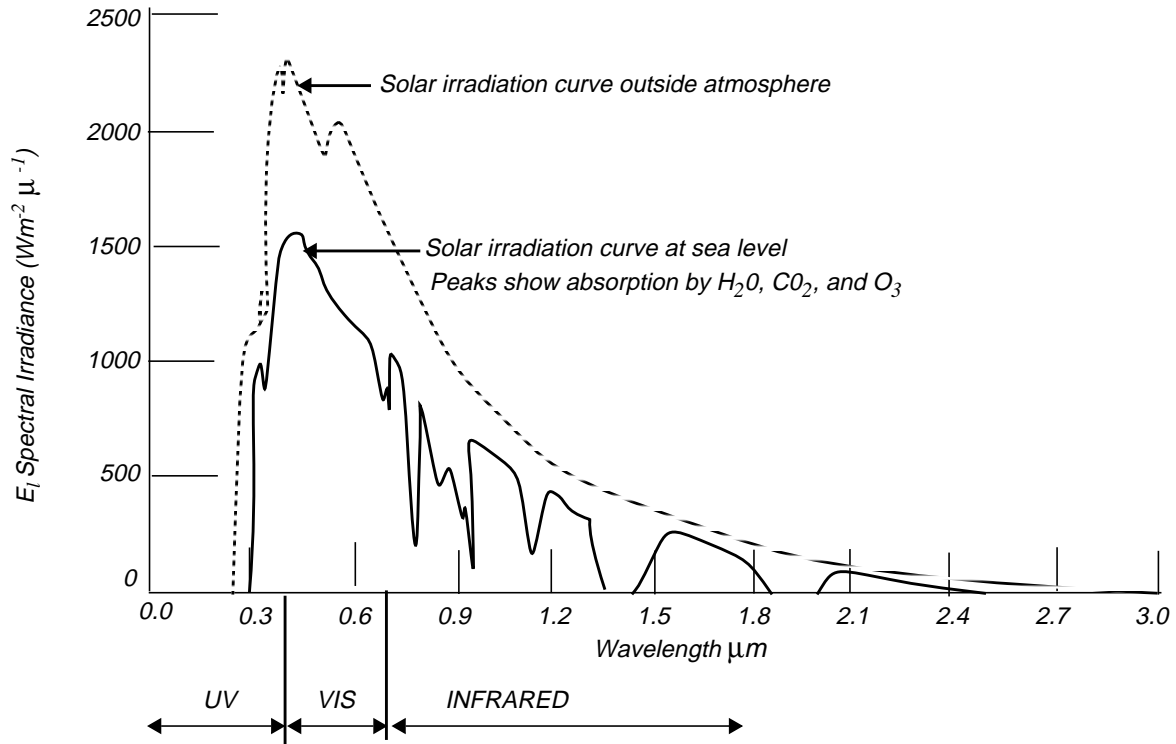
- absorption spectra—the EMR wavelengths that are absorbed by specific materials of interest
- reflection spectra—the EMR wavelengths that are reflected by specific materials of interest

Absorption Spectra

Absorption is based on the molecular bonds in the (surface) material. Which wavelengths are absorbed depends upon the chemical composition and crystalline structure of the material. For pure compounds, these absorption bands are so specific that the SWIR region is often called an infrared fingerprint.

Atmospheric Absorption

In remote sensing, the sun is the radiation source for passive sensors. However, the sun does not emit the same amount of radiation at all wavelengths. Figure 1-4 shows the solar irradiation curve, which is far from linear.

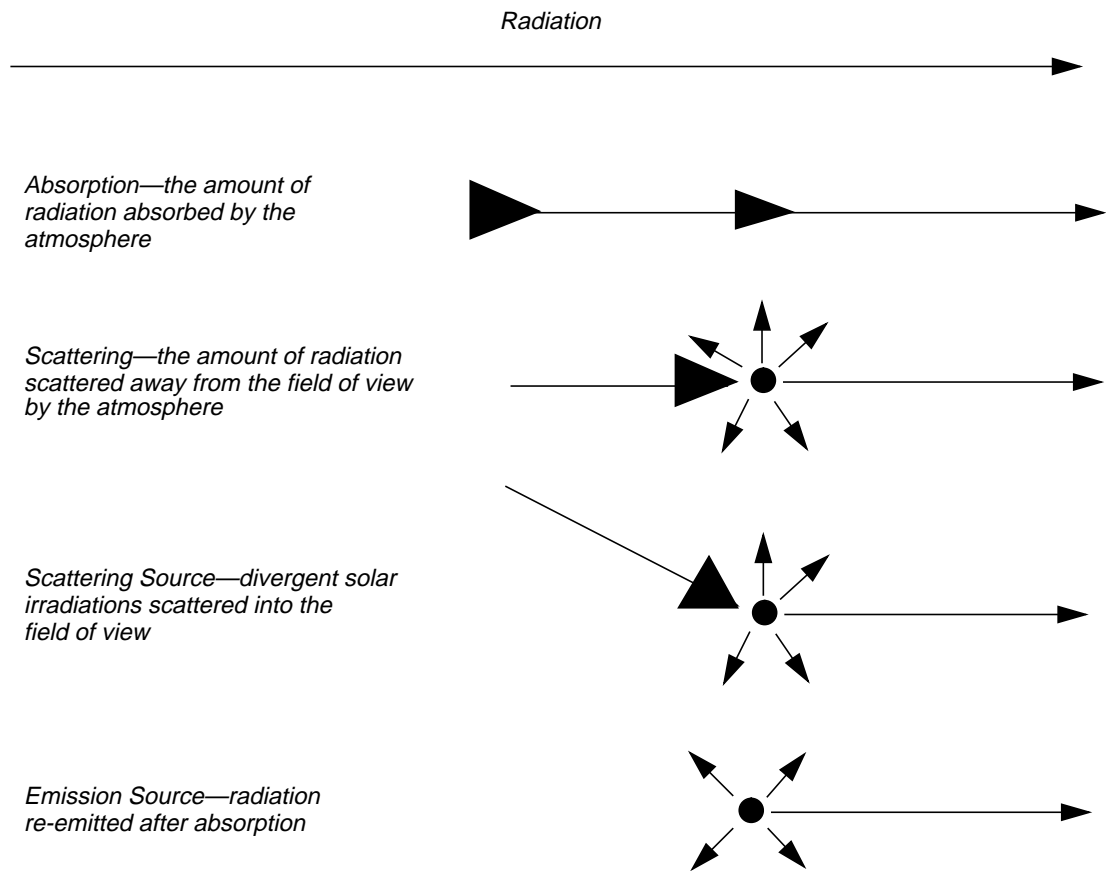


Source: Modified from Chahine, et al 1983

Figure 1-4 Sun Illumination Spectral Irradiance at the Earth's Surface

Solar radiation must travel through the Earth's atmosphere before it reaches the Earth's surface. As it travels through the atmosphere, radiation is affected by four phenomena (Elachi 1987):

- absorption—the amount of radiation absorbed by the atmosphere
- scattering—the amount of radiation scattered away from the field of view by the atmosphere
- scattering source—divergent solar irradiation scattered into the field of view
- emission source—radiation re-emitted after absorption



Source: Elachi 1987

Figure 1-5 Factors Affecting Radiation

Absorption is not a linear phenomena—it is logarithmic with concentration (Flaschka 1969). In addition, the concentration of atmospheric gases, especially water vapor, is variable. The other major gases of importance are carbon dioxide (CO₂) and ozone (O₃), which can vary considerably around urban areas. Thus, the extent of atmospheric absorbance varies with humidity, elevation, proximity to (or downwind of) urban smog, and other factors.

Scattering is modeled as Rayleigh scattering with a commonly used algorithm that accounts for the scattering of short wavelength energy by the gas molecules in the atmosphere (Pratt 1991)—for example, ozone. Scattering is variable with both wavelength and atmospheric aerosols. Aerosols differ regionally (ocean vs. desert) and daily (for example, Los Angeles smog has different concentrations daily).

Scattering source and emission source may account for only 5% of the variance. These factors are minor, but they must be considered for accurate calculation. After interaction with the target material, the reflected radiation must travel back through the atmosphere and be subjected to these phenomena a second time to arrive at the satellite.

The mathematical models that attempt to quantify the total atmospheric effect on the solar illumination are called radiative transfer equations. Some of the most commonly used are Lowtran (Kneizys 1988) and Modtran (Berk 1989).



See "CHAPTER 5: Enhancement" for more information on atmospheric modeling.

Reflectance Spectra

After rigorously defining the incident radiation (solar irradiation at target), it is possible to study the interaction of the radiation with the target material. When an electromagnetic wave (solar illumination in this case) strikes a target surface, three interactions are possible (Elachi 1987):

- reflection
- transmission
- scattering

It is the reflected radiation, generally modeled as bidirectional reflectance (Clark 1984), that is measured by the remote sensor.

Remotely sensed data are made up of reflectance values. The resulting reflectance values translate into discrete digital numbers (or values) recorded by the sensing device. These gray scale values fit within a certain bit range (such as 0 to 255, which is 8-bit data) depending on the characteristics of the sensor.

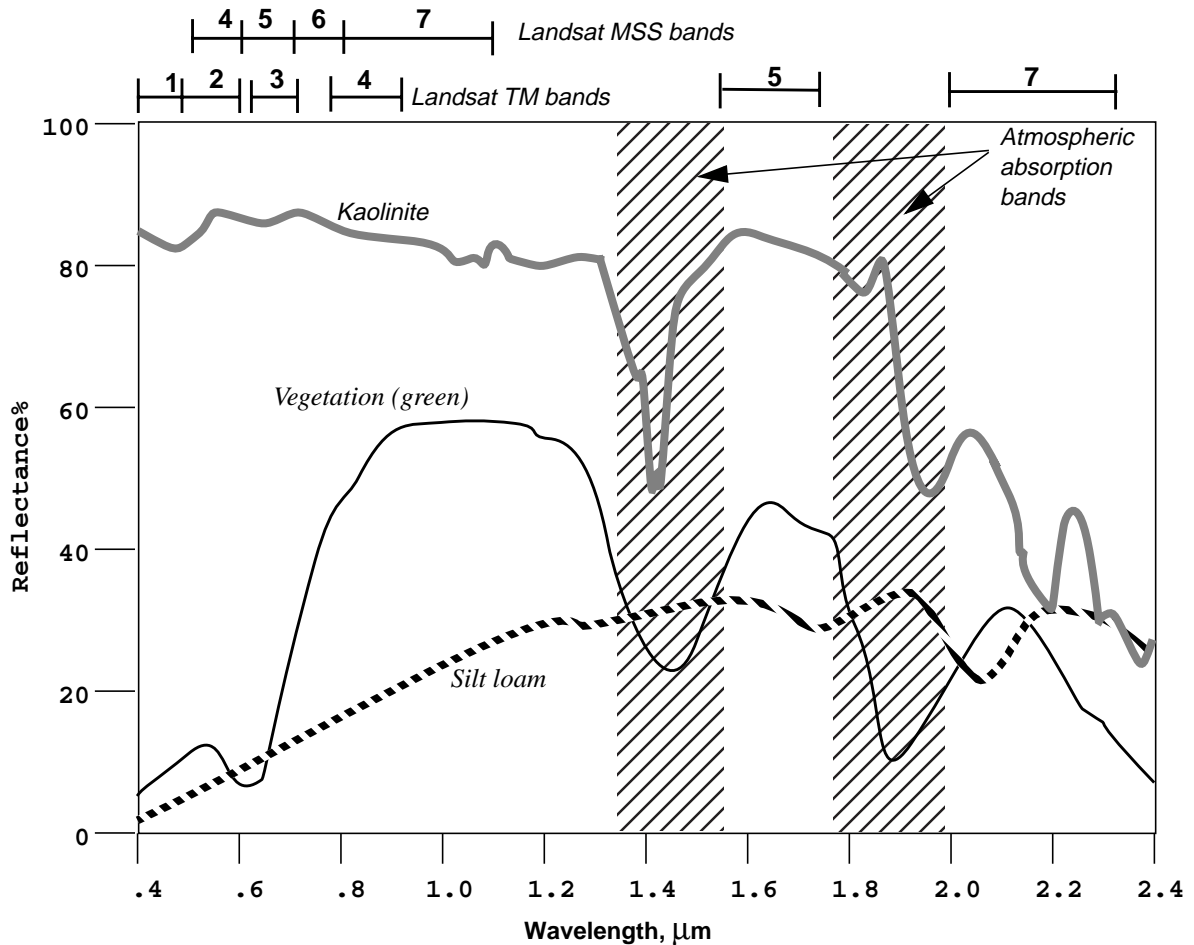
Each satellite sensor detector is designed to record a specific portion of the electromagnetic spectrum. For example, Landsat Thematic Mapper (TM) band 1 records the 0.45 to 0.52 μm portion of the spectrum and is designed for water body penetration, making it useful for coastal water mapping. It is also useful for soil/vegetation discriminations, forest type mapping, and cultural features identification (Lillesand and Kiefer 1987).

The characteristics of each sensor provide the first level of constraints on how to approach the task of enhancing specific features, such as vegetation or urban areas. Therefore, when choosing an enhancement technique, one should pay close attention to the characteristics of the land cover types within the constraints imposed by the individual sensors.

The use of VIS/IR imagery for target discrimination, whether the target is mineral, vegetation, man-made, or even the atmosphere itself, is based on the reflectance spectrum of the material of interest (see Figure 1-6). Every material has a characteristic spectrum based on the chemical composition of the material. When sunlight (the illumination source for VIS/IR imagery) strikes a target, certain wavelengths are absorbed by the chemical bonds; the rest are reflected back to the sensor. It is, in fact, the wavelengths that are *not* returned to the sensor that provide information about the imaged area.

Raster Data

Specific wavelengths are also absorbed by gases in the atmosphere (H_2O vapor, CO_2 , O_2 , etc.). If the atmosphere absorbs a large percentage of the radiation, it becomes difficult or impossible to use that particular wavelength(s) to study the Earth. For the present Landsat and Systeme Pour l'observation de la Terre (SPOT) sensors, only the water vapor bands are considered strong enough to exclude the use of their spectral absorption region. Figure 1-6 shows how Landsat TM bands 5 and 7 were carefully placed to avoid these regions. Absorption by other atmospheric gases was not extensive enough to eliminate the use of the spectral region for present day broad band sensors.



Source: Modified from Fraser 1986; Crist 1986; Sabins 1987

Figure 1-6 Reflectance Spectra

NOTE: This chart is for comparison purposes only. It is not meant to show actual values. The spectra are offset to better display the lines.

An inspection of the spectra reveals the theoretical basis of some of the indices in the ERDAS IMAGINE Image Interpreter. Consider the vegetation index TM4/TM3. It is readily apparent that for vegetation this value could be very large. For soils, the value could be much smaller, and for clay minerals, the value could be near zero. Conversely, when the clay ratio TM5/TM7 is considered, the opposite applies.

Hyperspectral Data

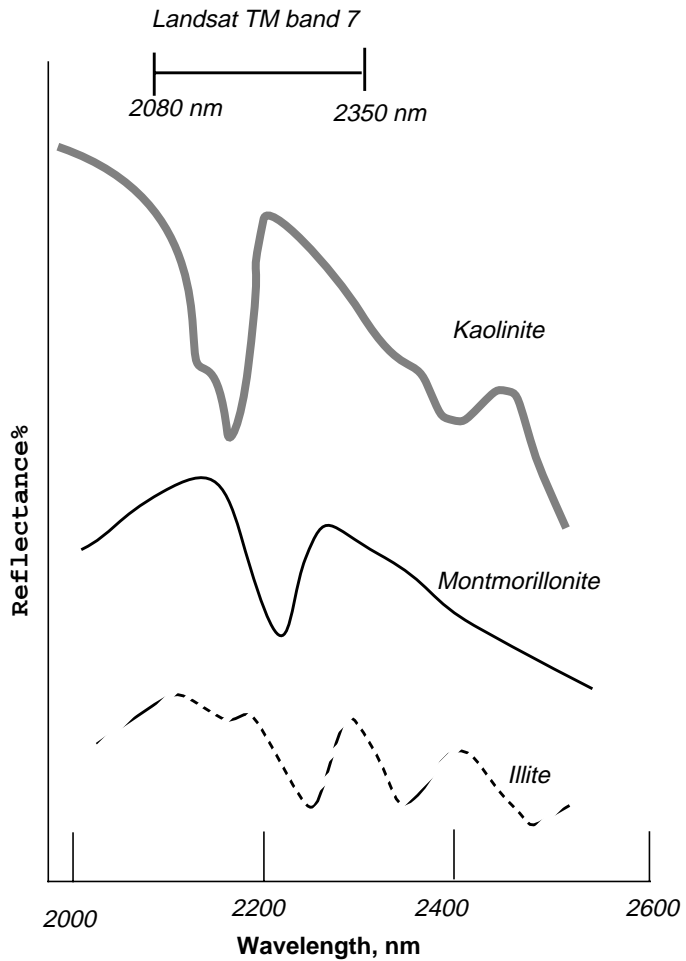
As remote sensing moves toward the use of more and narrower bands (for example, AVIRIS with 224 bands only 10 nm wide), absorption by specific atmospheric gases must be considered. These multiband sensors are called hyperspectral sensors. As more and more of the incident radiation is absorbed by the atmosphere, the digital number (DN) values of that band get lower, eventually becoming useless—unless one is studying the atmosphere. Someone wanting to measure the atmospheric content of a specific gas could utilize the bands of specific absorption.

NOTE: Hyperspectral bands are generally measured in nanometers (nm).

Figure 1-6 shows the spectral bandwidths of the channels for the Landsat sensors plotted above the absorption spectra of some common natural materials (kaolin clay, silty loam soil, and green vegetation). Note that while the spectra are continuous, the Landsat channels are segmented or discontinuous. We can still use the spectra in interpreting the Landsat data. For example, a Normalized Difference Vegetation Index (NDVI) ratio for the three would be very different and, therefore, could be used to discriminate between the three materials. Similarly, the ratio TM5/TM7 is commonly used to measure the concentration of clay minerals. Evaluation of the spectra shows why.

Figure 1-7 shows detail of the absorption spectra of three clay minerals. Because of the wide bandpass (2080 to 2350 nm) of TM band 7, it is not possible to discern between these three minerals with the Landsat sensor. As mentioned, the AVIRIS hyperspectral sensor has a large number of approximately 10 nm wide bands. With the proper selection of band ratios, mineral identification becomes possible. With this data set, it would be possible to discriminate between these three clay minerals, again using band ratios. For example, a color composite image prepared from RGB = 2160nm/2190nm, 2220nm/2250nm, 2350nm/2488nm could produce a color-coded clay mineral image-map.

The commercial airborne multispectral scanners are used in a similar fashion. The Airborne Imaging Spectrometer from the Geophysical & Environmental Research Corp. (GER) has 79 bands in the UV, visible, SWIR, and thermal-infrared regions. The Airborne Multispectral Scanner Mk2 by Geoscan Pty, Ltd., has up to 52 bands in the visible, SWIR, and thermal-infrared regions. To properly utilize these hyperspectral sensors, you must understand the phenomenon involved and have some idea of the target materials being sought.



Source: Modified from Sabins 1987

Figure 1-7 Laboratory Spectra of Clay Minerals in the Infrared Region

NOTE: Spectra are offset vertically for clarity.



The characteristics of Landsat, AVIRIS, and other data types are discussed in "CHAPTER 3: Raster and Vector Data Sources". See page 176 of "CHAPTER 5: Enhancement" for more information on the NDVI ratio.

Imaging Radar Data

Radar remote sensors can be broken into two broad categories: passive and active. The passive sensors record the very low intensity, microwave radiation naturally emitted by the Earth. Because of the very low intensity, these images have low spatial resolution (i.e., large pixel size).

It is the active sensors, termed imaging radar, that are introducing a new generation of satellite imagery to remote sensing. To produce an image, these satellites emit a directed beam of microwave energy at the target, and then collect the backscattered (reflected) radiation from the target scene. Because they must emit a powerful burst of energy, these satellites require large solar collectors and storage batteries. For this reason, they cannot operate continuously; some satellites are limited to 10 minutes of operation per hour.

The microwave energy emitted by an active radar sensor is coherent and defined by a narrow bandwidth. The following table summarizes the bandwidths used in remote sensing.

Table 1-1 Bandwidths Used in Remote Sensing

Band Designation*	Wavelength (λ), cm	Frequency (ν), GHz (10^9 cycles \cdot sec$^{-1}$)
Ka (0.86 cm)	0.8 to 1.1	40.0 to 26.5
K	1.1 to 1.7	26.5 to 18.0
Ku	1.7 to 2.4	18.0 to 12.5
X (3.0 cm, 3.2 cm)	2.4 to 3.8	12.5 to 8.0
C	3.8 to 7.5	8.0 to 4.0
S	7.5 to 15.0	4.0 to 2.0
L (23.5 cm, 25.0 cm)	15.0 to 30.0	2.0 to 1.0
P	30.0 to 100.0	1.0 to 0.3

*Wavelengths commonly used in imaging radars are shown in parentheses.

A key element of a radar sensor is the antenna. For a given position in space, the resolution of the resultant image is a function of the antenna size. This is termed a real-aperture radar (RAR). At some point, it becomes impossible to make a large enough antenna to create the desired spatial resolution. To get around this problem, processing techniques have been developed which combine the signals received by the sensor as it travels over the target. Thus, the antenna is perceived to be as long as the sensor path during backscatter reception. This is termed a synthetic aperture and the sensor a synthetic aperture radar (SAR).

The received signal is termed a phase history or echo hologram. It contains a time history of the radar signal over all the targets in the scene, and is itself a low resolution RAR image. In order to produce a high resolution image, this phase history is processed through a hardware/software system called an SAR processor. The SAR processor software requires operator input parameters, such as information about the sensor flight path and the radar sensor's characteristics, to process the raw signal data into an image. These input parameters depend on the desired result or intended application of the output imagery.

One of the most valuable advantages of imaging radar is that it creates images from its own energy source and therefore is not dependent on sunlight. Thus one can record uniform imagery any time of the day or night. In addition, the microwave frequencies at which imaging radars operate are largely unaffected by the atmosphere. This allows image collection through cloud cover or rain storms. However, the backscattered signal can be affected. Radar images collected during heavy rainfall are often seriously attenuated, which decreases the signal-to-noise ratio (SNR). In addition, the atmosphere does cause perturbations in the signal phase, which decreases resolution of output products, such as the SAR image or generated DEMs.

Resolution

Resolution is a broad term commonly used to describe:

- the number of pixels you can display on a display device, or
- the area on the ground that a pixel represents in an image file.

These broad definitions are inadequate when describing remotely sensed data. Four distinct types of resolution must be considered:

- spectral—the specific wavelength intervals that a sensor can record
- spatial—the area on the ground represented by each pixel
- radiometric—the number of possible data file values in each band (indicated by the number of bits into which the recorded energy is divided)
- temporal—how often a sensor obtains imagery of a particular area

These four domains contain separate information that can be extracted from the raw data.

Spectral Resolution

Spectral resolution refers to the specific wavelength intervals in the electromagnetic spectrum that a sensor can record (Simonett 1983). For example, band 1 of the Landsat TM sensor records energy between 0.45 and 0.52 μm in the visible part of the spectrum.

Wide intervals in the electromagnetic spectrum are referred to as coarse spectral resolution, and narrow intervals are referred to as fine spectral resolution. For example, the SPOT panchromatic sensor is considered to have coarse spectral resolution because it records EMR between 0.51 and 0.73 μm . On the other hand, band 3 of the Landsat TM sensor has fine spectral resolution because it records EMR between 0.63 and 0.69 μm (Jensen 1996).

NOTE: The spectral resolution does not indicate how many levels the signal is broken into.

Spatial Resolution

Spatial resolution is a measure of the smallest object that can be resolved by the sensor, or the area on the ground represented by each pixel (Simonett 1983). The finer the resolution, the lower the number. For instance, a spatial resolution of 79 meters is coarser than a spatial resolution of 10 meters.

Scale

The terms large-scale imagery and small-scale imagery often refer to spatial resolution. Scale is the ratio of distance on a map as related to the true distance on the ground (Star and Estes 1990).

Large-scale in remote sensing refers to imagery in which each pixel represents a small area on the ground, such as SPOT data, with a spatial resolution of 10 m or 20 m. Small scale refers to imagery in which each pixel represents a large area on the ground, such as Advanced Very High Resolution Radiometer (AVHRR) data, with a spatial resolution of 1.1 km.

This terminology is derived from the fraction used to represent the scale of the map, such as 1:50,000. Small-scale imagery is represented by a small fraction (one over a very large number). Large-scale imagery is represented by a larger fraction (one over a smaller number). Generally, anything smaller than 1:250,000 is considered small-scale imagery.

NOTE: Scale and spatial resolution are not always the same thing. An image always has the same spatial resolution, but it can be presented at different scales (Simonett 1983).

Instantaneous Field of View

Spatial resolution is also described as the instantaneous field of view (IFOV) of the sensor, although the IFOV is not always the same as the area represented by each pixel. The IFOV is a measure of the area viewed by a single detector in a given instant in time (Star and Estes 1990). For example, Landsat MSS data have an IFOV of 79×79 meters, but there is an overlap of 11.5 meters in each pass of the scanner, so the actual area represented by each pixel is 56.5×79 meters (usually rounded to 57×79 meters).

Even though the IFOV is not the same as the spatial resolution, it is important to know the number of pixels into which the total field of view for the image is broken. Objects smaller than the stated pixel size may still be detectable in the image if they contrast with the background, such as roads, drainage patterns, etc.

On the other hand, objects the same size as the stated pixel size (or larger) may not be detectable if there are brighter or more dominant objects nearby. In Figure 1-8, a house sits in the middle of four pixels. If the house has a reflectance similar to its surroundings, the data file values for each of these pixels reflect the area around the house, not the house itself, since the house does not dominate any one of the four pixels. However, if the house has a significantly different reflectance than its surroundings, it may still be detectable.

Raster Data

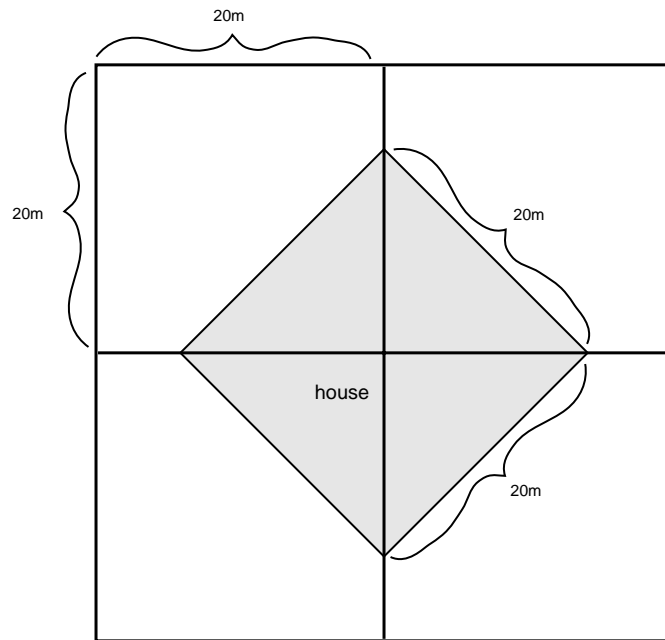


Figure 1-8 IFOV

Radiometric Resolution

Radiometric resolution refers to the dynamic range, or number of possible data file values in each band. This is referred to by the number of bits into which the recorded energy is divided.

For instance, in 8-bit data, the data file values range from 0 to 255 for each pixel, but in 7-bit data, the data file values for each pixel range from 0 to 128.

In Figure 1-9, 8-bit and 7-bit data are illustrated. The sensor measures the EMR in its range. The total intensity of the energy from 0 to the maximum amount the sensor measures is broken down into 256 brightness values for 8-bit data, and 128 brightness values for 7-bit data.

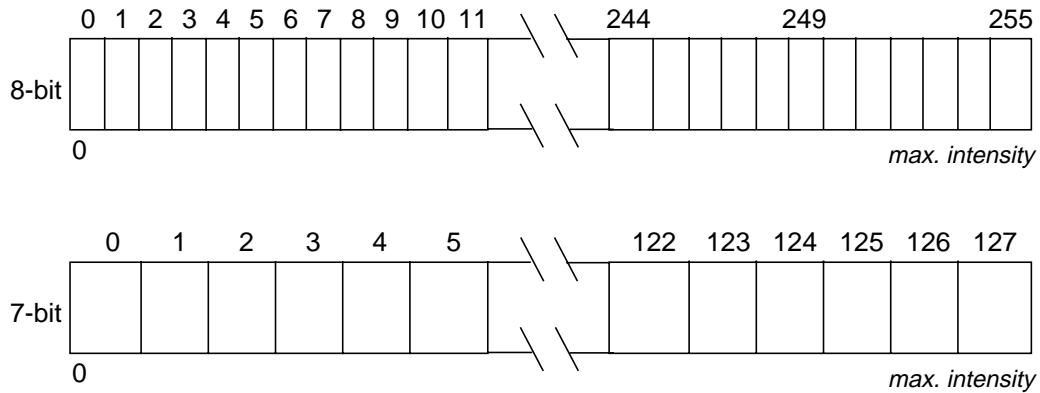


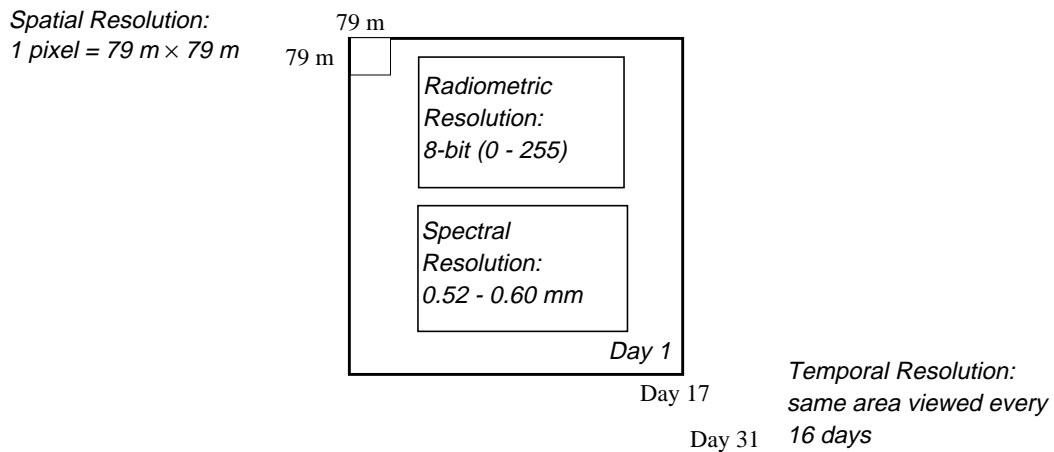
Figure 1-9 Brightness Values

Temporal Resolution

Temporal resolution refers to how often a sensor obtains imagery of a particular area. For example, the Landsat satellite can view the same area of the globe once every 16 days. SPOT, on the other hand, can revisit the same area every three days.

NOTE: Temporal resolution is an important factor to consider in change detection studies.

Figure 1-10 illustrates all four types of resolution:



Source: EOSAT

Figure 1-10 Landsat TM—Band 2 (Four Types of Resolution)

Data Correction

There are several types of errors that can be manifested in remotely sensed data. Among these are line dropout and striping. These errors can be corrected to an extent in GIS by radiometric and geometric correction functions.

NOTE: Radiometric errors are usually already corrected in data from EOSAT or SPOT.



See "CHAPTER 5: Enhancement" for more information on radiometric and geometric correction.

Line Dropout

Line dropout occurs when a detector either completely fails to function or becomes temporarily saturated during a scan (like the effect of a camera flash on a human retina). The result is a line or partial line of data with higher data file values, creating a horizontal streak until the detector(s) recovers, if it recovers.

Line dropout is usually corrected by replacing the bad line with a line of estimated data file values. The estimated line is based on the lines above and below it.



You can correct line dropout using the 5×5 Median Filter from the Radar Speckle Suppression function. The Convolution and Focal Analysis functions in the ERDAS IMAGINE Image Interpreter also corrects for line dropout.

Striping

Striping or banding occurs if a detector goes out of adjustment—that is, it provides readings consistently greater than or less than the other detectors for the same band over the same ground cover.



Use ERDAS IMAGINE Image Interpreter or ERDAS IMAGINE Spatial Modeler for implementing algorithms to eliminate striping. The ERDAS IMAGINE Spatial Modeler editing capabilities allow you to adapt the algorithms to best address the data.

Data Storage

Image data can be stored on a variety of media—tapes, CD-ROMs, or floppy diskettes, for example—but *how* the data are stored (e.g., structure) is more important than on what they are stored.

All computer data are in binary format. The basic unit of binary data is a bit. A bit can have two possible values—0 and 1, or “off” and “on” respectively. A set of bits, however, can have many more values, depending on the number of bits used. The number of values that can be expressed by a set of bits is 2 to the power of the number of bits used.

A byte is 8 bits of data. Generally, file size and disk space are referred to by number of bytes. For example, a PC may have 640 kilobytes (1,024 bytes = 1 kilobyte) of RAM (random access memory), or a file may need 55,698 bytes of disk space. A megabyte (Mb) is about one million bytes. A gigabyte (Gb) is about one billion bytes.

Storage Formats

Image data can be arranged in several ways on a tape or other media. The most common storage formats are:

- BIL (band interleaved by line)
- BSQ (band sequential)
- BIP (band interleaved by pixel)

For a single band of data, all formats (BIL, BIP, and BSQ) are identical, as long as the data are not blocked.



Blocked data are discussed under "Storage Media" on page 21.

BIL

In BIL (band interleaved by line) format, each record in the file contains a scan line (row) of data for one band (Slater 1980). All bands of data for a given line are stored consecutively within the file as shown in Figure 1-11.

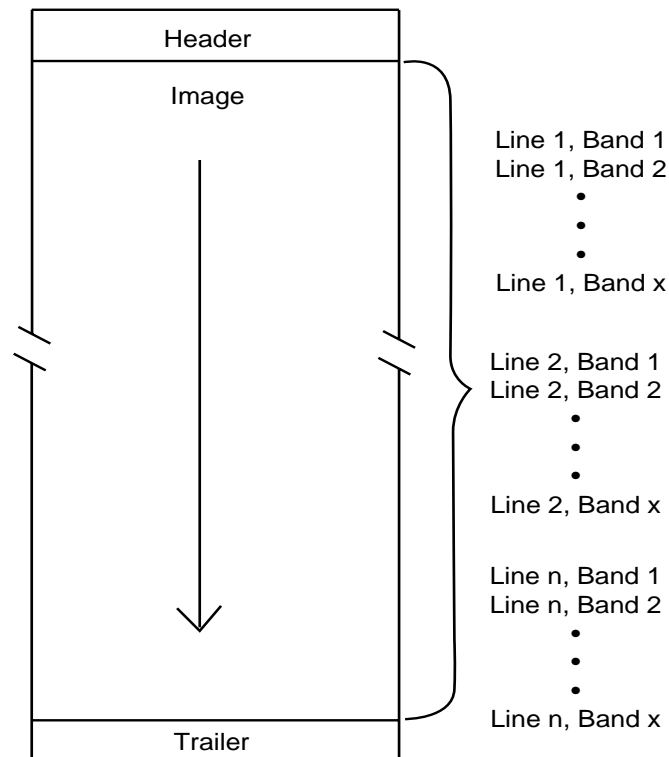


Figure 1-11 Band Interleaved by Line (BIL)

NOTE: Although a header and trailer file are shown in this diagram, not all BIL data contain header and trailer files.

BSQ

In BSQ (band sequential) format, each entire band is stored consecutively in the same file (Slater 1980). This format is advantageous, in that:

- one band can be read and viewed easily, and
- multiple bands can be easily loaded in any order.

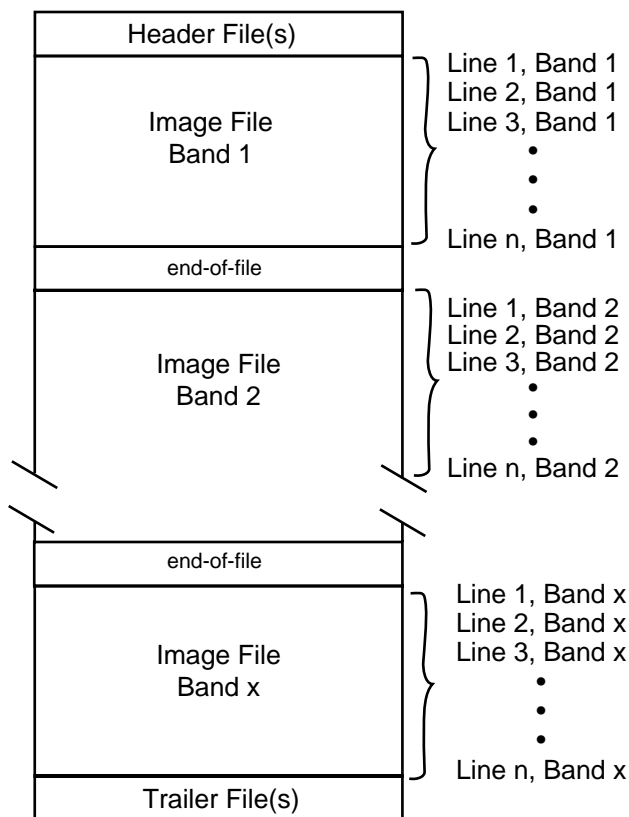


Figure 1-12 Band Sequential (BSQ)

Landsat TM data are stored in a type of BSQ format known as fast format. Fast format data have the following characteristics:

- Files are not split between tapes. If a band starts on the first tape, it ends on the first tape.
- An end-of-file (EOF) marker follows each band.
- An end-of-volume marker marks the end of each volume (tape). An end-of-volume marker consists of three end-of-file markers.
- There is one header file per tape.
- There are no header records preceding the image data.
- Regular products (not geocoded) are normally unblocked. Geocoded products are normally blocked (EOSAT).

ERDAS IMAGINE imports all of the header and image file information.



See "Geocoded Data" on page 28 for more information on geocoded data.

BIP

In BIP (band interleaved by pixel) format, the values for each band are ordered within a given pixel. The pixels are arranged sequentially on the tape (Slater 1980). The sequence for BIP format is:

Pixel 1, Band 1

Pixel 1, Band 2

Pixel 1, Band 3

•

•

•

Pixel 2, Band 1

Pixel 2, Band 2

Pixel 2, Band 3

•

•

•

Storage Media

Today, most raster data are available on a variety of storage media to meet the needs of users, depending on the system hardware and devices available. When ordering data, it is sometimes possible to select the type of media preferred. The most common forms of storage media are discussed in the following section:

- 9-track tape
- 4 mm tape
- 8 mm tape
- 1/4" cartridge tape
- CD-ROM/optical disk

Other types of storage media are:

- floppy disk (3.5" or 5.25")
- film, photograph, or paper
- videotape

Tape

The data on a tape can be divided into logical records and physical records. A record is the basic storage unit on a tape.

Raster Data

- A logical record is a series of bytes that form a unit. For example, all the data for one line of an image may form a logical record.
- A physical record is a consecutive series of bytes on a magnetic tape, followed by a gap, or blank space, on the tape.

Blocked Data

For reasons of efficiency, data can be blocked to fit more on a tape. Blocked data are sequenced so that there are more logical records in each physical record. The number of logical records in each physical record is the blocking factor. For instance, a record may contain 28,000 bytes, but only 4000 columns due to a blocking factor of 7.

Tape Contents

Tapes are available in a variety of sizes and storage capacities. To obtain information about the data on a particular tape, read the tape label or box, or read the header file. Often, there is limited information on the outside of the tape. Therefore, it may be necessary to read the header files on each tape for specific information, such as:

- number of tapes that hold the data set
- number of columns (in pixels)
- number of rows (in pixels)
- data storage format—BIL, BSQ, BIP
- pixel depth—4-bit, 8-bit, 10-bit, 12-bit, 16-bit
- number of bands
- blocking factor
- number of header files and header records

4 mm Tapes

The 4 mm tape is a relative newcomer in the world of GIS. This tape is a mere 2" × 1.75" in size, but it can hold up to 2 Gb of data. This petite cassette offers an obvious shipping and storage advantage because of its size.

8 mm Tapes

The 8 mm tape offers the advantage of storing vast amounts of data. Tapes are available in 5 and 10 Gb storage capacities (although some tape drives cannot handle the 10 Gb size). The 8 mm tape is a 2.5" × 4" cassette, which makes it easy to ship and handle.

1/4" Cartridge Tapes

This tape format falls between the 8 mm and 9-track in physical size and storage capacity. The tape is approximately 4" × 6" in size and stores up to 150 Mb of data.

9-Track Tapes

A 9-track tape is an older format that was the standard for two decades. It is a large circular tape approximately 10" in diameter. It requires a 9-track tape drive as a peripheral device for retrieving data. The size and storage capability make 9-track less convenient than 8 mm or 1/4" tapes. However, 9-track tapes are still widely used.

A single 9-track tape may be referred to as a volume. The complete set of tapes that contains one image is referred to as a volume set.

The storage format of a 9-track tape in binary format is described by the number of bits per inch, bpi, on the tape. The tapes most commonly used have either 1600 or 6250 bpi. The number of bits per inch on a tape is also referred to as the tape density. Depending on the length of the tape, 9-tracks can store between 120-150 Mb of data.

CD-ROM

Data such as ADRG and Digital Line Graphs (DLG) are most often available on CD-ROM, although many types of data can be requested in CD-ROM format. A CD-ROM is an optical read-only storage device which can be read with a CD player. CD-ROMs offer the advantage of storing large amounts of data in a small, compact device. Up to 644 Mb can be stored on a CD-ROM. Also, since this device is read-only, it protects the data from accidentally being overwritten, erased, or changed from its original integrity. This is the most stable of the current media storage types and data stored on CD-ROM are expected to last for decades without degradation.

Calculating Disk Space

To calculate the amount of disk space a raster file requires on an ERDAS IMAGINE system, use the following formula:

$$[(x \times y \times b) \times n] \times 1.2 = \text{output file size}$$

Where:

y = rows

x = columns

b = number of bytes per pixel

n = number of bands

1.2 adds 10% to the file size for pyramid layers (starting with the second) and 10% for miscellaneous adjustments, such as histograms, lookup tables, etc.

NOTE: This output file size is approximate.

For example, to load a 3 band, 16-bit file with 500 rows and 500 columns, about 2,100,000 bytes of disk space is needed.

$$[(500 \times 500) \times 2] \times 3 \times 1.4 = 2,100,000 \quad \text{or } 2.1 \text{ Mb}$$

Raster Data

Bytes Per Pixel

The number of bytes per pixel is listed below:

4-bit data: .5

8-bit data: 1.0

16-bit data: 2.0

NOTE: On the PC, disk space is shown in bytes. On the workstation, disk space is shown as kilobytes (1,024 bytes).

ERDAS IMAGINE Format (.img)

In ERDAS IMAGINE, file name extensions identify the file type. When data are imported into ERDAS IMAGINE, they are converted to the ERDAS IMAGINE file format and stored in image files. ERDAS IMAGINE image files (.img) can contain two types of raster layers:

- thematic
- continuous

An image file can store a combination of thematic and continuous layers, or just one type.

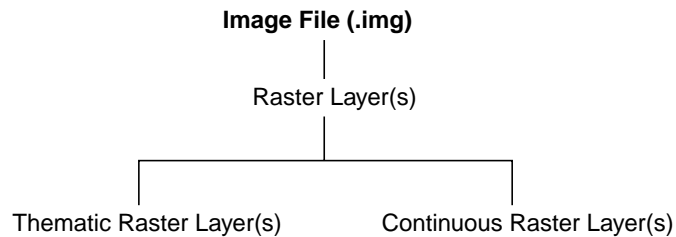


Figure 1-13 Image Files Store Raster Layers

ERDAS Version 7.5 Users

For Version 7.5 users, when importing a GIS file from Version 7.5, it becomes an image file with one thematic raster layer. When importing a LAN file, each band becomes a continuous raster layer within an image file.

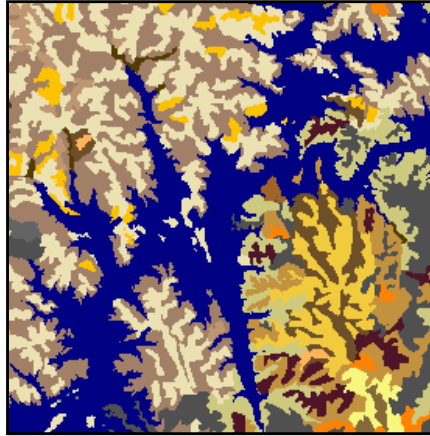
Thematic Raster Layer

Thematic data are raster layers that contain qualitative, categorical information about an area. A thematic layer is contained within an image file. Thematic layers lend themselves to applications in which categories or themes are used. Thematic raster layers are used to represent data measured on a nominal or ordinal scale, such as:

- soils
- land use
- land cover
- roads

- hydrology

NOTE: Thematic raster layers are displayed as pseudo color layers.



soils

Figure 1-14 Example of a Thematic Raster Layer



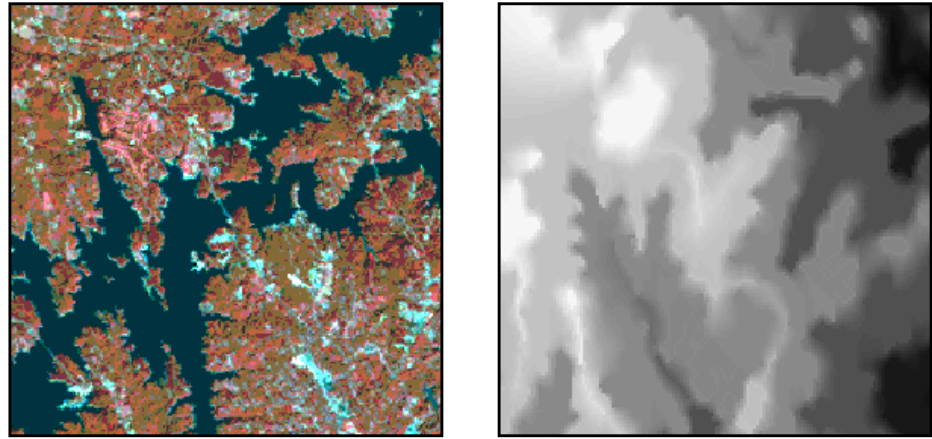
See "CHAPTER 4: Image Display" for information on displaying thematic raster layers.

Continuous Raster Layer

Continuous data are raster layers that contain quantitative (measuring a characteristic on an interval or ratio scale) and related, continuous values. Continuous raster layers can be multiband (e.g., Landsat TM data) or single band (e.g., SPOT panchromatic data). The following types of data are examples of continuous raster layers:

- Landsat
- SPOT
- digitized (scanned) aerial photograph
- DEM
- slope
- temperature

NOTE: Continuous raster layers can be displayed as either a gray scale raster layer or a true color raster layer.



Landsat TM

DEM

Figure 1-15 Examples of Continuous Raster Layers

Tiled Data

Data in the .img format are tiled data. Tiled data are stored in tiles that can be set to any size.



The default tile size for image files is 64 × 64 pixels.

Image File Contents

The image files contain the following additional information about the data:

- the data file values
- statistics
- lookup tables
- map coordinates
- map projection



This additional information can be viewed using the Image Information function located on the Viewer's tool bar.

Statistics

In ERDAS IMAGINE, the file statistics are generated from the data file values in the layer and incorporated into the image file. This statistical information is used to create many program defaults, and helps you make processing decisions.

Pyramid Layers

Sometimes a large image takes longer than normal to display in the Viewer. The pyramid layer option enables you to display large images faster. Pyramid layers are image layers which are successively reduced by the power of 2 and resampled.



The Pyramid Layer option is available in the Image Information function located on the Viewer's tool bar and, from the Import function.



See "CHAPTER 4: Image Display" for more information on pyramid layers. See the On-Line Help for detailed information on ERDAS IMAGINE file formats.

Image File Organization

Data are easy to locate if the data files are well organized. Well organized files also make data more accessible to anyone who uses the system. Using consistent naming conventions and the ERDAS IMAGINE Image Catalog helps keep image files well organized and accessible.

Consistent Naming Convention

Many processes create an output file, and every time a file is created, it is necessary to assign a file name. The name that is used can either cause confusion about the process that has taken place, or it can clarify and give direction. For example, if the name of the output file is image.img, it is difficult to determine the contents of the file. On the other hand, if a standard nomenclature is developed in which the file name refers to a process or contents of the file, it is possible to determine the progress of a project and contents of a file by examining the directory.

Develop a naming convention that is based on the contents of the file. This helps everyone involved know what the file contains. For example, in a project to create a map composition for Lake Lanier, a directory for the files may look similar to the one below:

```
lanierTM.img
lanierSPOT.img
lanierSymbols.ovr
lanierlegends.map.ovr
lanierScalebars.map.ovr
lanier.map
lanier.plt
lanier.gcc
lanierUTM.img
```

From this listing, one can make some educated guesses about the contents of each file based on naming conventions used. For example, lanierTM.img is probably a Landsat TM scene of Lake Lanier. The file lanier.map is probably a map composition that has map frames with lanierTM.img and lanierSPOT.img data in them. The file lanierUTM.img was probably created when lanierTM.img was rectified to a UTM map projection.

Keeping Track of Image Files

Using a database to store information about images enables you to track image files (.img) without having to know the name or location of the file. The database can be queried for specific parameters (e.g., size, type, map projection) and the database returns a list of image files that match the search criteria. This file information helps to quickly determine which image(s) to use, where it is located, and its ancillary data. An image database is especially helpful when there are many image files and even many on-going projects. For example, you could use the database to search for all of the image files of Georgia that have a UTM map projection.



Use the ERDAS IMAGINE Image Catalog to track and store information for image files (.img) that are imported and created in ERDAS IMAGINE.

NOTE: All information in the Image Catalog database, except archive information, is extracted from the image file header. Therefore, if this information is modified in the Image Information utility, it is necessary to recatalog the image in order to update the information in the Image Catalog database.

ERDAS IMAGINE Image Catalog

The ERDAS IMAGINE Image Catalog database is designed to serve as a library and information management system for image files (.img) that are imported and created in ERDAS IMAGINE. The information for the image files is displayed in the Image Catalog CellArray™. This CellArray enables you to view all of the ancillary data for the image files in the database. When records are queried based on specific criteria, the image files that match the criteria are highlighted in the CellArray. It is also possible to graphically view the coverage of the selected image files on a map in a canvas window.

When it is necessary to store some data on a tape, the ERDAS IMAGINE Image Catalog database enables you to archive image files to external devices. The Image Catalog CellArray shows which tape the image file is stored on, and the file can be easily retrieved from the tape device to a designated disk directory. The archived image files are copies of the files on disk—nothing is removed from the disk. Once the file is archived, it can be removed from the disk, if you like.

Geocoded Data

Geocoding, also known as georeferencing, is the geographical registration or coding of the pixels in an image. Geocoded data are images that have been rectified to a particular map projection and pixel size.

Raw, remotely-sensed image data are gathered by a sensor on a platform, such as an aircraft or satellite. In this raw form, the image data are not referenced to a map projection. Rectification is the process of projecting the data onto a plane and making them conform to a map projection system.

It is possible to geocode raw image data with the ERDAS IMAGINE rectification tools. Geocoded data are also available from Space Imaging EOSAT and SPOT.



See "Appendix B: Map Projections" for detailed information on the different projections available. See "CHAPTER 9: Rectification" for information on geocoding raw imagery with ERDAS IMAGINE.

Using Image Data in GIS

ERDAS IMAGINE provides many tools designed to extract the necessary information from the images in a database. The following chapters in this book describe many of these processes.

This section briefly describes some basic image file techniques that may be useful for any application.

Subsetting and Mosaicking

Within ERDAS IMAGINE, there are options available to make additional image files from those acquired from EOSAT, SPOT, etc. These options involve combining files, mosaicking, and subsetting.

ERDAS IMAGINE programs allow image data with an unlimited number of bands, but the most common satellite data types—Landsat and SPOT—have seven or fewer bands. Image files can be created with more than seven bands.

It may be useful to combine data from two different dates into one file. This is called multitemporal imagery. For example, a user may want to combine Landsat TM from one date with TM data from a later date, then perform a classification based on the combined data. This is particularly useful for change detection studies.

You can also incorporate elevation data into an existing image file as another band, or create new bands through various enhancement techniques.



To combine two or more image files, each file must be georeferenced to the same coordinate system, or to each other. See "CHAPTER 9: Rectification" for information on georeferencing images.

Subset

Subsetting refers to breaking out a portion of a large file into one or more smaller files.

Often, image files contain areas much larger than a particular study area. In these cases, it is helpful to reduce the size of the image file to include only the area of interest (AOI). This not only eliminates the extraneous data in the file, but it speeds up processing due to the smaller amount of data to process. This can be important when dealing with multiband data.



The ERDAS IMAGINE Import option often lets you define a subset area of an image to preview or import. You can also use the Subset option from ERDAS IMAGINE Image Interpreter to define a subset area.

Raster Data

Mosaic

On the other hand, the study area in which you are interested may span several image files. In this case, it is necessary to combine the images to create one large file. This is called mosaicking.



To create a mosaicked image, use the Mosaic Images option from the Data Preparation menu. All of the images to be mosaicked must be georeferenced to the same coordinate system.

Enhancement

Image enhancement is the process of making an image more interpretable for a particular application (Faust 1989). Enhancement can make important features of raw, remotely sensed data and aerial photographs more interpretable to the human eye. Enhancement techniques are often used instead of classification for extracting useful information from images.

There are many enhancement techniques available. They range in complexity from a simple contrast stretch, where the original data file values are stretched to fit the range of the display device, to principal components analysis, where the number of image file bands can be reduced and new bands created to account for the most variance in the data.



See "CHAPTER 5: Enhancement" for more information on enhancement techniques.

Multispectral Classification

Image data are often used to create thematic files through multispectral classification. This entails using spectral pattern recognition to identify groups of pixels that represent a common characteristic of the scene, such as soil type or vegetation.



See "CHAPTER 6: Classification" for a detailed explanation of classification procedures.

Editing Raster Data

ERDAS IMAGINE provides raster editing tools for editing the data values of thematic and continuous raster data. This is primarily a correction mechanism that enables you to correct bad data values which produce noise, such as spikes and holes in imagery. The raster editing functions can be applied to the entire image or a user-selected area of interest (AOI).

With raster editing, data values in thematic data can also be recoded according to class. Recoding is a function that reassigns data values to a region or to an entire class of pixels.



See "CHAPTER 11: Geographic Information Systems" for information about recoding data. See "CHAPTER 5: Enhancement" for information about reducing data noise using spatial filtering.

The ERDAS IMAGINE raster editing functions allow the use of focal and global spatial modeling functions for computing the values to replace noisy pixels or areas in continuous or thematic data.

Focal operations are filters that calculate the replacement value based on a window (3×3 , 5×5 , etc.), and replace the pixel of interest with the replacement value. Therefore this function affects one pixel at a time, and the number of surrounding pixels that influence the value is determined by the size of the moving window.

Global operations calculate the replacement value for an entire area rather than affecting one pixel at a time. These functions, specifically the Majority option, are more applicable to thematic data.



See the ERDAS IMAGINE On-Line Help for information about using and selecting AOIs.



The raster editing tools are available in the Viewer.

Editing Continuous (Athematic) Data

Editing DEMs

DEMs occasionally contain spurious pixels or bad data. These spikes, holes, and other noises caused by automatic DEM extraction can be corrected by editing the raster data values and replacing them with meaningful values. This discussion of raster editing focuses on DEM editing.



The ERDAS IMAGINE Raster Editing functionality was originally designed to edit DEMs, but it can also be used with images of other continuous data sources, such as radar, SPOT, Landsat, and digitized photographs.

When editing continuous raster data, you can modify or replace original pixel values with the following:

- a constant value—enter a known constant value for areas such as lakes.
- the average of the buffering pixels—replace the original pixel value with the average of the pixels in a specified buffer area around the AOI. This is used where the constant values of the AOI are not known, but the area is flat or homogeneous with little variation (for example, a lake).
- the original data value plus a constant value—add a negative constant value to the original data values to compensate for the height of trees and other vertical features in the DEM. This technique is commonly used in forested areas.
- spatial filtering—filter data values to eliminate noise such as spikes or holes in the data.
- interpolation techniques (discussed below).

Interpolation Techniques

While the previously listed raster editing techniques are perfectly suitable for some applications, the following interpolation techniques provide the best methods for raster editing:

- 2-D polynomial—surface approximation
- multisurface functions—with least squares prediction
- distance weighting

Each pixel's data value is interpolated from the reference points in the data file. These interpolation techniques are described below:

2-D Polynomial

This interpolation technique provides faster interpolation calculations than distance weighting and multisurface functions. The following equation is used:

$$V = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + \dots$$

Where:

V = data value (elevation value for DEM)
 a = polynomial coefficients
 x = x coordinate
 y = y coordinate

Multisurface Functions

The multisurface technique provides the most accurate results for editing DEMs that have been created through automatic extraction. The following equation is used:

$$V = \sum W_i Q_i$$

Where:

V = output data value (elevation value for DEM)
 W_i = coefficients which are derived by the least squares method
 Q_i = distance-related kernels which are actually interpretable as continuous single value surfaces

Source: Wang 1990

Distance Weighting

The weighting function determines how the output data values are interpolated from a set of reference data points. For each pixel, the values of all reference points are weighted by a value corresponding with the distance between each point and the pixel.

The weighting function used in ERDAS IMAGINE is:

$$W = \left(\frac{S}{D} - 1 \right)^2$$

Where:

S = normalization factor

D = distance from output data point and reference point

The value for any given pixel is calculated by taking the sum of weighting factors for all reference points multiplied by the data values of those points, and dividing by the sum of the weighting factors:

$$V = \frac{\sum_{i=1}^n W_i \times V_i}{\sum_{i=1}^n W_i}$$

Where:

V = output data value (elevation value for DEM)

i = i^{th} reference point

W_i = weighting factor of point i

V_i =data value of point i

n = number of reference points

Source: Wang 1990

Raster Data

CHAPTER 2

Vector Layers

Introduction

ERDAS IMAGINE is designed to integrate two data types, raster and vector, into one system. While the previous chapter explored the characteristics of raster data, this chapter is focused on vector data. The vector data structure in ERDAS IMAGINE is based on the ArcInfo data model (developed by ESRI, Inc.). This chapter describes vector data, attribute information, and symbolization.



You do not need ArcInfo software or an ArcInfo license to use the vector capabilities in ERDAS IMAGINE. Since the ArcInfo data model is used in ERDAS IMAGINE, you can use ArcInfo coverages directly without importing them.



See "CHAPTER 11: Geographic Information Systems" for information on editing vector layers and using vector data in a GIS.

Vector data consist of:

- points
- lines
- polygons

Each is illustrated in Figure 2-1.

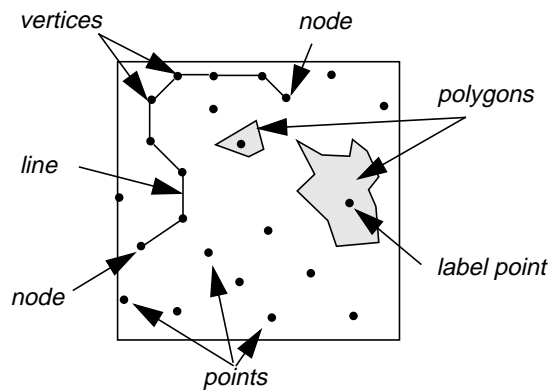


Figure 2-1 Vector Elements

Points

A point is represented by a single x, y coordinate pair. Points can represent the location of a geographic feature or a point that has no area, such as a mountain peak. Label points are also used to identify polygons (see Figure 2-2).

Lines

A line (polyline) is a set of line segments and represents a linear geographic feature, such as a river, road, or utility line. Lines can also represent nongeographical boundaries, such as voting districts, school zones, contour lines, etc.

Polygons

A polygon is a closed line or closed set of lines defining a homogeneous area, such as soil type, land use, or water body. Polygons can also be used to represent nongeographical features, such as wildlife habitats, state borders, commercial districts, etc. Polygons also contain label points that identify the polygon. The label point links the polygon to its attributes.

Vertex

The points that define a line are vertices. A vertex is a point that defines an element, such as the endpoint of a line segment or a location in a polygon where the line segment defining the polygon changes direction. The ending points of a line are called nodes. Each line has two nodes: a from-node and a to-node. The from-node is the first vertex in a line. The to-node is the last vertex in a line. Lines join other lines only at nodes. A series of lines in which the from-node of the first line joins the to-node of the last line is a polygon.

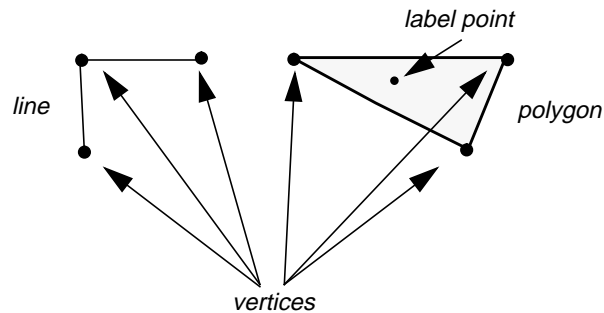


Figure 2-2 Vertices

In Figure 2-2, the line and the polygon are each defined by three vertices.

Coordinates

Vector data are expressed by the coordinates of vertices. The vertices that define each element are referenced with x, y, or Cartesian, coordinates. In some instances, those coordinates may be inches [as in some computer-aided design (CAD) applications], but often the coordinates are map coordinates, such as State Plane, Universal Transverse Mercator (UTM), or Lambert Conformal Conic. Vector data digitized from an ungeoreferenced image are expressed in file coordinates.

Tics

Vector layers are referenced to coordinates or a map projection system using tic files that contain geographic control points for the layer. Every vector layer must have a tic file. Tics are not topologically linked to other features in the layer and do not have descriptive data associated with them.

Vector Layers

Although it is possible to have points, lines, and polygons in a single layer, a layer typically consists of one type of feature. It is possible to have one vector layer for streams (lines) and another layer for parcels (polygons). A vector layer is defined as a set of features where each feature has a location (defined by coordinates and topological pointers to other features) and, possibly attributes (defined as a set of named items or variables) (ESRI 1989). Vector layers contain both the vector features (points, lines, polygons) and the attribute information.

Usually, vector layers are also divided by the type of information they represent. This enables the user to isolate data into themes, similar to the themes used in raster layers. Political districts and soil types would probably be in separate layers, even though both are represented with polygons. If the project requires that the coincidence of features in two or more layers be studied, the user can overlay them or create a new layer.



See "CHAPTER 11: Geographic Information Systems" for more information about analyzing vector layers.

Topology

The spatial relationships between features in a vector layer are defined using topology. In topological vector data, a mathematical procedure is used to define connections between features, identify adjacent polygons, and define a feature as a set of other features (e.g., a polygon is made of connecting lines) (ESRI 1990).

Topology is not automatically created when a vector layer is created. It must be added later using specific functions. Topology must be updated after a layer is edited also.



"Digitizing" on page 43 describes how topology is created for a new or edited vector layer.

Vector Layers

Vector Files

As mentioned above, the ERDAS IMAGINE vector structure is based on the ArcInfo data model used for ARC coverages. This georelational data model is actually a set of files using the computer's operating system for file management and input/output. An ERDAS IMAGINE vector layer is stored in subdirectories on the disk. Vector data are represented by a set of logical tables of information, stored as files within the subdirectory. These files may serve the following purposes:

- define features
- provide feature attributes
- cross-reference feature definition files
- provide descriptive information for the coverage as a whole

A workspace is a location that contains one or more vector layers. Workspaces provide a convenient means for organizing layers into related groups. They also provide a place for the storage of tabular data not directly tied to a particular layer. Each workspace is completely independent. It is possible to have an unlimited number of workspaces and an unlimited number of vector layers in a workspace. Table 2-1 summarizes the types of files that are used to make up vector layers.

Table 2-1 Description of File Types

File Type	File	Description
Feature Definition Files	ARC	Line coordinates and topology
	CNT	Polygon centroid coordinates
	LAB	Label point coordinates and topology
	TIC	Tic coordinates
Feature Attribute Files	AAT	Line (arc) attribute table
	PAT	Polygon or point attribute table
Feature Cross-Reference File	PAL	Polygon/line/node cross-reference file
Layer Description Files	BND	Coordinate extremes
	LOG	Layer history file
	PRJ	Coordinate definition file
	TOL	Layer tolerance file

Figure 2-3 illustrates how a typical vector workspace is set up (ESRI 1992).

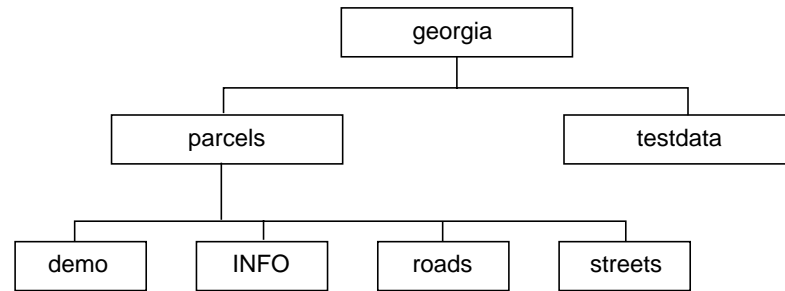


Figure 2-3 Workspace Structure



Because vector layers are stored in directories rather than in simple files, you **MUST** use the utilities provided in ERDAS IMAGINE to copy and rename them. A utility is also provided to update path names that are no longer correct due to the use of regular system commands on vector layers.



See the ESRI documentation for more detailed information about the different vector files.

Attribute Information

Along with points, lines, and polygons, a vector layer can have a wealth of associated descriptive, or attribute, information associated with it. Attribute information is displayed in CellArrays. This is the same information that is stored in the INFO database of ArcInfo. Some attributes are automatically generated when the layer is created. Custom fields can be added to each attribute table. Attribute fields can contain numerical or character data.

The attributes for a roads layer may look similar to the example in Figure 2-4. You can select features in the layer based on the attribute information. Likewise, when a row is selected in the attribute CellArray, that feature is highlighted in the Viewer.

Attributes for /usr/imagine/data/imagineqa/arcdata/usa25m/rds_25m

Record	fnode#	tnode#	lpoly#	rpoly#	length	rds_25m#	rds_25m-id	inter_rte1	inter_rte2	inter_rte3	us_rte1	us_rte2	us_rte3	state_rte1	state_rte2	state_fips	state_name	sub_region	
37	37	41	2	2	3568.972	37	322	90	0	0	0	0	0	0	0	0	53	Washington	Pacific
38	36	37	2	2	10992.120	38	326	90	0	0	0	0	0	0	0	0	53	Washington	Pacific
39	43	42	2	2	47.197	39	328	90	0	0	0	0	0	0	0	0	53	Washington	Pacific
40	42	44	3	3	11975.447	40	328	90	0	0	0	0	0	0	0	0	16	Idaho	Mtn
41	41	43	2	2	15350.767	41	332	90	0	0	0	0	0	0	0	0	53	Washington	Pacific
42	44	45	3	3	8430.396	42	343	90	0	0	0	0	0	0	0	0	16	Idaho	Mtn
43	46	40	2	2	12548.521	43	344	90	0	0	395	0	0	0	0	0	53	Washington	Pacific
44	47	34	4	4	23374.289	44	352	15	0	0	91	0	0	0	0	0	30	Montana	Mtn
45	45	48	3	3	5405.501	45	354	90	0	0	0	0	0	0	0	0	16	Idaho	Mtn
46	49	36	2	2	15864.804	46	366	5	0	0	12	0	0	0	0	0	53	Washington	Pacific
47	32	50	2	2	38922.578	47	368	90	0	0	0	0	0	0	0	0	53	Washington	Pacific
48	46	51	3	3	8227.392	48	371	90	0	0	0	0	0	0	0	0	16	Idaho	Mtn

Figure 2-4 Attribute CellArray

Using Imported Attribute Data

When external data types are imported into ERDAS IMAGINE, only the required attribute information is imported into the attribute tables (AAT and PAT files) of the new vector layer. The rest of the attribute information is written to one of the following INFO files:

- <layer name>.ACODE—arc attribute information
- <layer name>.PCODE—polygon attribute information
- <layer name>.XCODE—point attribute information

To utilize all of this attribute information, the INFO files can be merged into the PAT and AAT files. Once this attribute information has been merged, it can be viewed in CellArrays and edited as desired. This new information can then be exported back to its original format.

The complete path of the file must be specified when establishing an INFO file name in a Viewer application, such as exporting attributes or merging attributes, as shown in the following example:

```
/georgia/parcels/info!arc!parcels.pcode
```



Use the Attributes option in the Viewer to view and manipulate vector attribute data, including merging and exporting. (The Raster Attribute Editor is for raster attributes only and cannot be used to edit vector attributes.)



See the ERDAS IMAGINE On-Line Help for more information about using CellArrays.

Displaying Vector Data

Vector data are displayed in Viewers, as are other data types in ERDAS IMAGINE. You can display a single vector layer, overlay several layers in one Viewer, or display a vector layer(s) over a raster layer(s).

In layers that contain more than one feature (a combination of points, lines, and polygons), you can select which features to display. For example, if you are studying parcels, you may want to display only the polygons in a layer that also contains street centerlines (lines).

Color Schemes

Vector data are usually assigned class values in the same manner as the pixels in a thematic raster file. These class values correspond to different colors on the display screen. As with a pseudo color image, you can assign a color scheme for displaying the vector classes.



See "CHAPTER 4: Image Display" for a thorough discussion of how images are displayed.

Symbolization

Vector layers can be displayed with symbolization, meaning that the attributes can be used to determine how points, lines, and polygons are rendered. Points, lines, polygons, and nodes are symbolized using styles and symbols similar to annotation. For example, if a point layer represents cities and towns, the appropriate symbol could be used at each point based on the population of that area.

Points

Point symbolization options include symbol, size, and color. The symbols available are the same symbols available for annotation.

Lines

Lines can be symbolized with varying line patterns, composition, width, and color. The line styles available are the same as those available for annotation.

Polygons

Polygons can be symbolized as lines or as filled polygons. Polygons symbolized as lines can have varying line styles (see "Lines" above). For filled polygons, either a solid fill color or a repeated symbol can be selected. When symbols are used, you select the symbol to use, the symbol size, symbol color, background color, and the x- and y-separation between symbols. Figure 2-5 illustrates a pattern fill.

Vector Layers

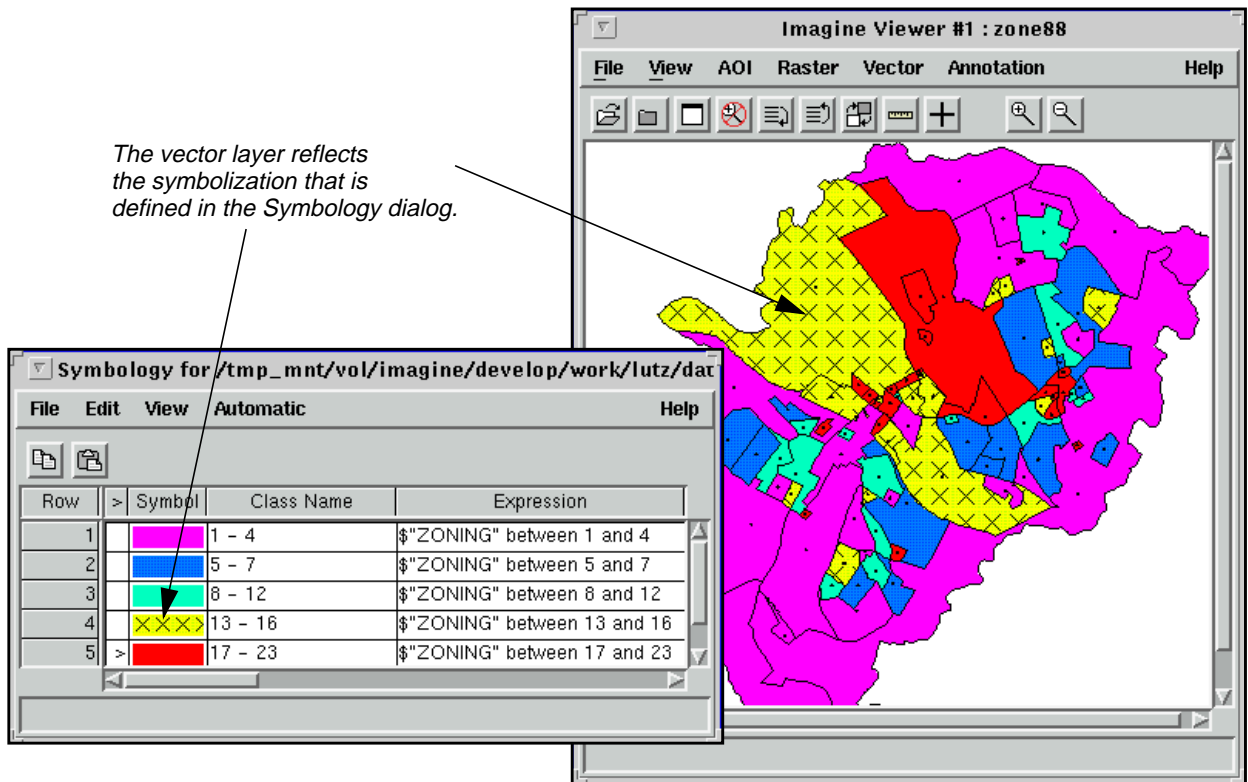


Figure 2-5 Symbolization Example



See the ERDAS IMAGINE Tour Guides or On-Line Help for information about selecting features and using CellArrays.

Vector Data Sources

Vector data are created by:

- tablet digitizing—maps, photographs, or other hardcopy data can be digitized using a digitizing tablet
- screen digitizing—create new vector layers by using the mouse to digitize on the screen
- using other software packages—many external vector data types can be converted to ERDAS IMAGINE vector layers
- converting raster layers—raster layers can be converted to vector layers

Each of these options is discussed in a separate section.

Digitizing

In the broadest sense, digitizing refers to any process that converts nondigital data into numbers. However, in ERDAS IMAGINE, the digitizing of vectors refers to the creation of vector data from hardcopy materials or raster images that are traced using a digitizer keypad on a digitizing tablet or a mouse on a displayed image.

Any image not already in digital format must be digitized before it can be read by the computer and incorporated into the database. Most Landsat, SPOT, or other satellite data are already in digital format upon receipt, so it is not necessary to digitize them. However, you may also have maps, photographs, or other nondigital data that contain information you want to incorporate into the study. Or, you may want to extract certain features from a digital image to include in a vector layer. Tablet digitizing and screen digitizing enable you to digitize certain features of a map or photograph, such as roads, bodies of water, voting districts, and so forth.

Tablet Digitizing

Tablet digitizing involves the use of a digitizing tablet to transfer nondigital data such as maps or photographs to vector format. The digitizing tablet contains an internal electronic grid that transmits data to ERDAS IMAGINE on cue from a digitizer keypad operated by you.

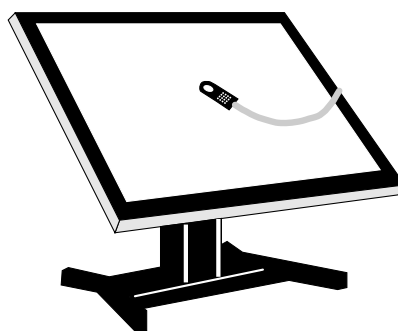


Figure 2-6 Digitizing Tablet

Digitizer Setup

The map or photograph to be digitized is secured on the tablet, and a coordinate system is established with a setup procedure.

Digitizer Operation

The handheld digitizer keypad features a small window with a crosshair and keypad buttons. Position the intersection of the crosshair directly over the point to be digitized. Depending on the type of equipment and the program being used, one of the input buttons is pushed to tell the system which function to perform, such as:

- digitize a point (i.e., transmit map coordinate data),
- connect a point to previous points,
- assign a particular value to the point or polygon, or
- measure the distance between points, etc.

Move the puck along the desired polygon boundaries or lines, digitizing points at appropriate intervals (where lines curve or change direction), until all the points are collected.



Newly created vector layers do not contain topological data. You must create topology using the Build or Clean options. This is discussed further in "CHAPTER 11: Geographic Information Systems"

Digitizing Modes

There are two modes used in digitizing:

- point mode—one point is generated each time a keypad button is pressed
- stream mode—points are generated continuously at specified intervals, while the puck is in proximity to the surface of the digitizing tablet



You can create a new vector layer from the Viewer. Select the Tablet Input function from the Viewer to use a digitizing tablet to enter new information into that layer.

Measurement

The digitizing tablet can also be used to measure both linear and areal distances on a map or photograph. The digitizer puck is used to outline the areas to measure. You can measure:

- lengths and angles by drawing a line
- perimeters and areas using a polygonal, rectangular, or elliptical shape
- positions by specifying a particular point

Measurements can be saved to a file, printed, and copied. These operations can also be performed with screen digitizing.



Select the Measure function from the Viewer or click on the Ruler tool in the Viewer tool bar to enable tablet or screen measurement.

Screen Digitizing

In screen digitizing, vector data are drawn with a mouse in the Viewer using the displayed image as a reference. These data are then written to a vector layer.

Screen digitizing is used for the same purposes as tablet digitizing, such as:

- digitizing roads, bodies of water, political boundaries
- selecting training samples for input to the classification programs
- outlining an area of interest for any number of applications



Create a new vector layer from the Viewer.

Imported Vector Data

Many types of vector data from other software packages can be incorporated into the ERDAS IMAGINE system. These data formats include:

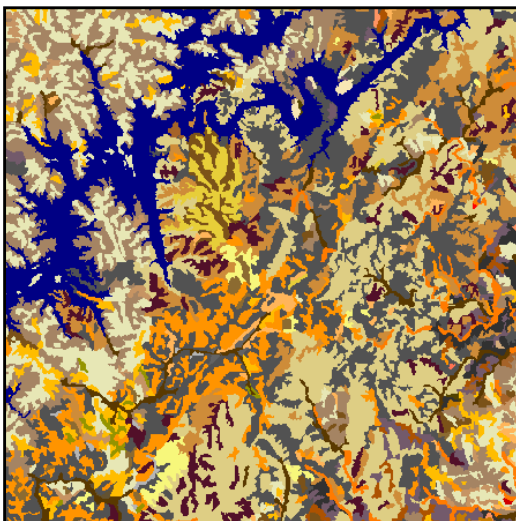
- ArcInfo GENERATE format files from ESRI, Inc.
- ArcInfo INTERCHANGE files from ESRI, Inc.
- ArcView Shapefiles from ESRI, Inc.
- Digital Line Graphs (DLG) from U.S.G.S.
- Digital Exchange Files (DXF) from Autodesk, Inc.
- ETAK MapBase files from ETAK, Inc.
- Initial Graphics Exchange Standard (IGES) files
- Intergraph Design (DGN) files from Intergraph
- Spatial Data Transfer Standard (SDTS) vector files
- Topologically Integrated Geographic Encoding and Referencing System (TIGER) files from the U.S. Census Bureau
- Vector Product Format (VPF) files from the Defense Mapping Agency



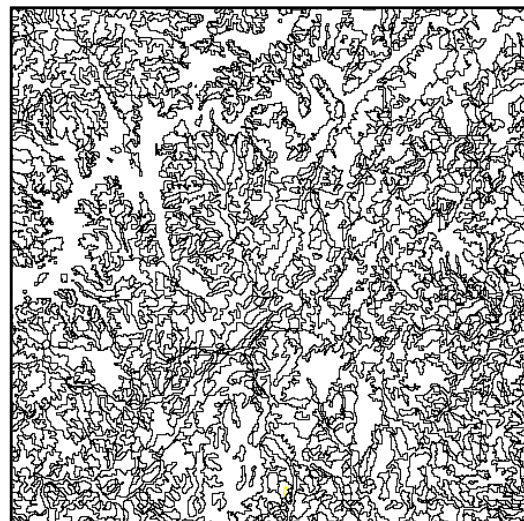
See "CHAPTER 3: Raster and Vector Data Sources" for more information on these data.

Raster to Vector Conversion

A raster layer can be converted to a vector layer and used as another layer in a vector database. The following diagram illustrates a thematic file in raster format that has been converted to vector format.



Raster soils layer



Soils layer converted to vector polygon layer

Figure 2-7 Raster Format Converted to Vector Format

Vector Layers

Most commonly, thematic raster data rather than continuous data are converted to vector format, since converting continuous layers may create more vector features than are practical or even manageable.



Convert vector data to raster data, and vice versa, using IMAGINE Vector™.

Other Vector Data Types

While this chapter has focused mainly on the ArcInfo coverage format, there are other types of vector formats that you can use in ERDAS IMAGINE. The two primary types are:

- shapefile
- Spatial Database Engine (SDE)

Shapefile Vector Format

The shapefile vector format was designed by ESRI. You can now use shapefile format (extension .shp) in ERDAS IMAGINE. You can now:

- display shapefiles
- create shapefiles
- edit shapefiles
- attribute shapefiles
- symbolize shapefiles
- print shapefiles

The shapefile contains spatial data, such as boundary information.

SDE

Like the shapefile format, the Spatial Database Engine (SDE) is a vector format designed by ESRI. The data layers are stored in a relational database management system (RDBMS) such as Oracle, or SQL Server. Some of the features of SDE include:

- storage of large, untiled spatial layers for fast retrieval
- powerful and flexible query capabilities using the SQL where clause
- operation in a client-server environment
- multiuser access to the data

ERDAS IMAGINE has the capability to act as a client to access SDE vector layers stored in a database. To do this, it uses a wizard interface to connect ERDAS IMAGINE to a SDE database, and selects one of the vector layers. Additionally, it can join business tables with the vector layer, and generate a subset of features by imposing attribute constraints (e.g., SQL where clause).

The definition of the vector layer as extracted from a SDE database is stored in a <layername>.sdv file, and can be loaded as a regular ERDAS IMAGINE data file. ERDAS IMAGINE supports the SDE projection systems. Currently, ERDAS IMAGINE's SDE capability is read-only. For example, features can be queried and AOIs can be created, but not edited.

SDTS

SDTS stands for Spatial Data Transfer Standard. SDTS is used to transfer spatial data between computer systems. Such data includes attribute, georeferencing, data quality report, data dictionary, and supporting metadata.

According to the USGS, the

implementation of SDTS is of significant interest to users and producers of digital spatial data because of the potential for increased access to and sharing of spatial data, the reduction of information loss in data exchange, the elimination of the duplication of data acquisition, and the increase in the quality and integrity of spatial data. (USGS 1999)

The components of SDTS are broken down into six parts. The first three parts are related, but independent, and are concerned with the transfer of spatial data. The last three parts provide definitions for rules and formats for applying SDTS to the exchange of data. The parts of SDTS are as follows:

- Part 1—Logical Specifications
- Part 2—Spatial Features
- Part 3—ISO 8211 Encoding
- Part 4—Topological Vector Profile
- Part 5—Raster Profile
- Part 6—Point Profile

Vector Layers

CHAPTER 3

Raster and Vector Data Sources

Introduction

This chapter is an introduction to the most common raster and vector data types that can be used with the ERDAS IMAGINE software package. The raster data types covered include:

- visible/infrared satellite data
- radar imagery
- airborne sensor data
- scanned or digitized maps and photographs
- digital terrain models (DTMs)

The vector data types covered include:

- ArcInfo GENERATE format files
 - AutoCAD Digital Exchange Files (DXF)
 - United States Geological Survey (USGS) Digital Line Graphs (DLG)
 - MapBase digital street network files (ETAK)
 - U.S. Department of Commerce Initial Graphics Exchange Standard files (IGES)
 - U.S. Census Bureau Topologically Integrated Geographic Encoding and Referencing System files (TIGER)
-

Importing and Exporting

Raster Data

There is an abundance of data available for use in GIS today. In addition to satellite and airborne imagery, raster data sources include digital x-rays, sonar, microscopic imagery, video digitized data, and many other sources.

Because of the wide variety of data formats, ERDAS IMAGINE provides two options for importing data:

- import for specific formats
- generic import for general formats

Import

Table 3-1 lists some of the raster data formats that can be imported to exported from, directly read from, and directly written to ERDAS IMAGINE.

Raster and Vector Data Sources



There is a distinct difference between import and direct read. Import means that the data is converted from its original format into another format (e.g. IMG, TIFF, or GRID Stack), which can be read directly by ERDAS IMAGINE. Direct read formats are those formats which the Viewer and many of its associated tools can read immediately without any conversion process.

NOTE: Annotation and Vector data formats are listed separately.

Table 3-1 Raster Data Formats

Data Type	Import	Export	Direct Read	Direct Write
ADRG	•			
ADRI	•			
ARCGEN	•	•		
Arc Coverage			•	•
ArcInfo & Space Imaging BIL, BIP, BSQ	•	•	•	
Arc Interchange	•	•		
ASCII	•			
AVHRR (NOAA)	•			
AVHRR (Dundee Format)	•			
AVHRR (Sharp)	•			
BIL, BIP, BSQ ^a (Generic Binary)	•	•	• ^b	
CADRG (Compressed ADRG)	•		•	
CIB (Controlled Image Base)	•		•	
DAEDALUS	•			
USGS DEM	•	•		
DOQ	•	•		
DOQ (JPEG)	•			
DTED	•		•	
ER Mapper			•	

Table 3-1 Raster Data Formats (Continued)

Data Type	Import	Export	Direct Read	Direct Write
ERS (I-PAF CEOS)	•			
ERS (Conae-PAF CEOS)	•			
ERS (Tel Aviv-PAF CEOS)	•			
ERS (D-PAF CEOS)	•			
ERS (UK-PAF CEOS)	•			
FIT			•	
Generic Binary (BIL, BIP, BSQ) ^a	•	•	• ^b	
GeoTIFF	•	•	•	•
GIS (Erdas 7.x)	•	•	•	
GRASS	•	•		
GRID	•	•	•	
GRID Stack	•	•	•	•
GRID Stack 7.x	•	•	•	
GRD (Surfer: ASCII/Binary)	•	•		
IRS-1C (EOSAT Fast Format C)	•			
IRS-1C (EUROMAP Fast Format C)	•			
JFIF (JPEG)	•	•	•	
Landsat-7 Fast-L7A ACRES	•			
Landsat-7 Fast-L7A EROS	•			
LAN (Erdas 7.x)	•	•	•	
MrSID			•	
MSS Landsat	•			
NLAPS Data Format (NDF)	•			
PCX	•	•	•	
RADARSAT (Vancouver CEOS)	•			

Table 3-1 Raster Data Formats (Continued)

Data Type	Import	Export	Direct Read	Direct Write
RADARSAT (Acres CEOS)	•			
RADARSAT (West Freugh CEOS)	•			
Raster Product Format	•		•	
SDE			•	
SDTS	•	•		
SeaWiFS L1B and L2A (OrbView)	•			
Shapefile	•	•	•	•
SPOT	•			
SPOT CCRS	•			
SPOT (GeoSpot)	•			
SPOT SICORP MetroView	•			
SUN Raster	•	•		
TIFF	•	•	•	•
TM Landsat Acres Fast Format	•			
TM Landsat Acres Standard Format	•			
TM Landsat EOSAT Fast Format	•			
TM Landsat EOSAT Standard Format	•			
TM Landsat ESA Fast Format	•			
TM Landsat ESA Standard Format	•			
TM Landsat IRS Fast Format	•			
TM Landsat IRS Standard Format	•			
TM Landsat-7 Fast-L7A ACRES	•			
TM Landsat-7 Fast-L7A EROS	•			

Table 3-1 Raster Data Formats (Continued)

Data Type	Import	Export	Direct Read	Direct Write
TM Landsat Radarsat Fast Format	●			
TM Landsat Radarsat Standard Format	●			

a. See "Generic Binary Data" on page 55.

b Direct read of generic binary data requires an accompanying header file in the ESRI ArcInfo, Space Imaging, or ERDAS IMAGINE formats.

The import function converts raster data to the ERDAS IMAGINE file format (.img), or other formats directly writable by ERDAS IMAGINE. The import function imports the data file values that make up the raster image, as well as the ephemeris or additional data inherent to the data structure. For example, when the user imports Landsat data, ERDAS IMAGINE also imports the georeferencing data for the image.



Raster data formats cannot be exported as vector data formats unless they are converted with the Vector utilities.



Each direct function is programmed specifically for that type of data and cannot be used to import other data types.

Raster Data Sources

NITFS

NITFS stands for the National Imagery Transmission Format Standard. NITFS is designed to pack numerous image compositions with complete annotation, text attachments, and imagery-associated metadata.

According to Jordan and Beck,

NITFS is an unclassified format that is based on ISO/IEC 12087-5, Basic Image Interchange Format (BIIF). The NITFS implementation of BIIF is documented in U.S. Military Standard 2500B, establishing a standard data format for digital imagery and imagery-related products.

NITFS was first introduced in 1990 and was for use by the government and intelligence agencies. NITFS is now the standard for military organizations as well as commercial industries.

Jordan and Beck list the following attributes of NITF files:

- provide a common basis for storage and digital interchange of images and associated data among existing and future systems

Raster and Vector Data Sources

- support interoperability by simultaneously providing a data format for shared access applications while also serving as a standard message format for dissemination of images and associated data (text, symbols, labels) via digital communications
- require minimal preprocessing and post-processing of transmitted data
- support variable image sizes and resolution
- minimize formatting overhead, particularly for those users transmitting only a small amount of data or with limited bandwidth
- provide universal features and functions without requiring commonality of hardware or proprietary software.

Moreover, NITF files support the following:

- multiple images
- annotation on images
- ASCII text files to accompany imagery and annotation
- metadata to go with imagery, annotation and text

The process of translating NITFS files is a cross-translation process. One system's internal representation for the files and their associated data is processed and put into the NITF format. The receiving system reformats the NITF file, and converts it for the receiving systems internal representation of the files and associated data.

In ERDAS IMAGINE, the IMAGINE NITF™ software accepts such information and assembles it into one file in the standard NITF format.

Source: Jordan and Beck 1999

Annotation Data

Annotation data can also be imported directly. Table 3-2 lists the Annotation formats.

There is a distinct difference between import and direct read. Import means that the data is converted from its original format into another format (e.g. IMG, TIFF, or GRID Stack), which can be read directly by ERDAS IMAGINE. Direct read formats are those formats which the Viewer and many of its associated tools can read immediately without any conversion process.

Table 3-2 Annotation Data Formats

Data Type	Import	Export	Direct Read	Direct Write
ANT (Erdas 7.x)	•	•		
ASCII To Point Annotation	•			
DXF To Annotation	•			

Generic Binary Data

The Generic Binary import option is a flexible program which enables the user to define the data structure for ERDAS IMAGINE. This program allows the import of BIL, BIP, and BSQ data that are stored in left to right, top to bottom row order. Data formats from unsigned 1-bit up to 64-bit floating point can be imported. This program imports only the data file values—it does not import ephemeris data, such as georeferencing information. However, this ephemeris data can be viewed using the Data View option (from the Utility menu or the Import dialog).

Complex data cannot be imported using this program; however, they can be imported as two real images and then combined into one complex image using the Spatial Modeler.



You cannot import tiled or compressed data using the Generic Binary import option.

Vector Data

Vector layers can be created within ERDAS IMAGINE by digitizing points, lines, and polygons using a digitizing tablet or the computer screen. Several vector data types, which are available from a variety of government agencies and private companies, can also be imported. Table 3-3 lists some of the vector data formats that can be imported to, and exported from, ERDAS IMAGINE:



There is a distinct difference between import and direct read. Import means that the data is converted from its original format into another format (e.g. IMG, TIFF, or GRID Stack), which can be read directly by ERDAS IMAGINE. Direct read formats are those formats which the Viewer and many of its associated tools can read immediately without any conversion process.

Table 3-3 Vector Data Formats

Data Type	Import	Export	Direct Read	Direct Write
ARCGEN	●	●		
Arc Interchange	●	●		
Arc_Interchange to Coverage	●			
Arc_Interchange to Grid	●			
ASCII To Point Coverage	●			
DFAD	●	●		
DGN (Intergraph IGDS)	●			
DIG Files (Erdas 7.x)	●			
DLG	●	●		
DXF to Annotation	●			

Table 3-3 Vector Data Formats (Continued)

Data Type	Import	Export	Direct Read	Direct Write
DXF to Coverage	•			
ETAK	•			
IGDS (Intergraph .dgn File)	•			
IGES	•	•		
MIF/MID (MapInfo) to Coverage	•			
SDE			•	
SDTS	•	•		
Shapefile			•	•
Terramodel			•	
TIGER	•	•		
VPF	•	•		

Once imported, the vector data are automatically converted to ERDAS IMAGINE vector layers.



These vector formats are discussed in more detail in "Vector Data from Other Software Vendors" on page 105. See "CHAPTER 2: Vector Layers" for more information on ERDAS IMAGINE vector layers.



Import and export vector data with the Import/Export function. You can also convert vector layers to raster format, and vice versa, with the IMAGINE Vector utilities.

Satellite Data

There are several data acquisition options available including photography, aerial sensors, and sophisticated satellite scanners. However, a satellite system offers these advantages:

- Digital data gathered by a satellite sensor can be transmitted over radio or microwave communications links and stored on magnetic tapes, so they are easily processed and analyzed by a computer.
- Many satellites orbit the Earth, so the same area can be covered on a regular basis for change detection.
- Once the satellite is launched, the cost for data acquisition is less than that for aircraft data.

- Satellites have very stable geometry, meaning that there is less chance for distortion or skew in the final image.



Use the Import/Export function to import a variety of satellite data.

Satellite System

A satellite system is composed of a scanner with sensors and a satellite platform. The sensors are made up of detectors.

- The scanner is the entire data acquisition system, such as the Landsat TM scanner or the SPOT panchromatic scanner (Lillesand and Kiefer 1987). It includes the sensor and the detectors.
- A sensor is a device that gathers energy, converts it to a signal, and presents it in a form suitable for obtaining information about the environment (Colwell 1983).
- A detector is the device in a sensor system that records electromagnetic radiation. For example, in the sensor system on the Landsat TM scanner there are 16 detectors for each wavelength band (except band 6, which has 4 detectors).

In a satellite system, the total width of the area on the ground covered by the scanner is called the swath width, or width of the total field of view (FOV). FOV differs from IFOV in that the IFOV is a measure of the field of view of each detector. The FOV is a measure of the field of view of all the detectors combined.

Satellite Characteristics

The U. S. Landsat and the French SPOT satellites are two important data acquisition satellites. These systems provide the majority of remotely-sensed digital images in use today. The Landsat and SPOT satellites have several characteristics in common:

- They have sun-synchronous orbits, meaning that they rotate around the Earth at the same rate as the Earth rotates on its axis, so data are always collected at the same local time of day over the same region.
- They both record electromagnetic radiation in one or more bands. Multiband data are referred to as multispectral imagery. Single band, or monochrome, imagery is called panchromatic.
- Both scanners can produce nadir views. Nadir is the area on the ground directly beneath the scanner's detectors.

NOTE: The current SPOT system has the ability to collect off-nadir stereo imagery.

Image Data Comparison

Figure 3-1 shows a comparison of the electromagnetic spectrum recorded by Landsat TM, Landsat MSS, SPOT, and National Oceanic and Atmospheric Administration (NOAA) AVHRR data. These data are described in detail in the following sections.

Raster and Vector Data Sources

	Landsat MSS (1,2,3,4)	Landsat TM (4, 5)	SPOT XS	SPOT Pan	NOAA AVHRR ¹
.5		Band 1			
.6	Band 1	Band 2	Band 1	Band 1	Band 1
.7	Band 2	Band 3	Band 2		
.8	Band 3	Band 4	Band 3		
.9	Band 4				Band 2
1.0					
1.1					
1.2					
1.3					
1.4					
1.5					
1.6		Band 5			
1.7					
1.8					
1.9					
2.0					
2.1		Band 7			
2.2					
2.3					
2.4					
2.5					
2.6					
3.0					
3.5					Band 3
4.0					
5.0					
6.0					
7.0					
8.0					
9.0					
10.0					
11.0		Band 6			Band 4
12.0					Band 5
13.0					

¹ NOAA AVHRR band 5 is not on the NOAA 10 satellite, but is on NOAA 11.

Figure 3-1 Multispectral Imagery Comparison

IKONOS

The IKONOS satellite was launched in September of 1999 by the Athena II rocket.

The resolution of the panchromatic sensor is 1 m. The resolution of the multispectral scanner is 4 m. The swath width is 13 km at nadir. The accuracy with out ground control is 12 m horizontally, and 10 m vertically; with ground control it is 2 m horizontally, and 3 m vertically.

IKONOS orbits at an altitude of 423 miles, or 681 kilometers. The revisit time is 2.9 days at 1 m resolution, and 1.5 days at 1.5 m resolution

Table 3-4 IKONOS Bands and Frequencies

Bands	Frequencies
1, Blue	0.45 to 0.52 μm
2, Green	0.52 to 0.60 μm
3, Red	0.63 to 0.69 μm
4, Near IR	0.76 to 0.90 μm
Pan	0.45 to 0.90 μm

Source: *Space Imaging 1999*

IRS

IRS-1C

The IRS-1C sensor was launched in December of 1995.

The repeat coverage of IRS-1C is every 24 days. The sensor has a 744 km swath width.

The IRS-1C satellite has three sensors on board with which to capture images of the Earth. Those sensors are as follows:

LISS-III

LISS-III has a spatial resolution of 23 m, with the exception of the SW Infrared band, which is 70 m. Bands 2, 3, and 4 have a swath width of 142 kilometers; band 5 has a swath width of 148 km. Repeat coverage occurs every 24 days at the Equator.

Table 3-5 LISS-III Bands and Frequencies

Bands	Frequencies
Blue	---
Green	0.52 to 0.59 μm
Red	0.62 to 0.68 μm
Near Infrared	0.77 to 0.86 μm
SW Infrared	1.55 to 1.70 μm

Raster and Vector Data Sources

Panchromatic sensor

The panchromatic sensor has 5.8 m spatial resolution, as well as stereo capability. Its swath width is 70 m. Repeat coverage is every 24 days at the Equator. The revisit time is every five days, with 28° off-nadir viewing.

Table 3-6 Panchromatic Band and Frequency

Band	Frequency
Pan	0.5 to 0.75 μm

Wide Field Sensor (WiFS)

- WiFS has a 188 m spatial resolution, and repeat coverage every five days at the Equator. The swath width is 774 km.

Table 3-7 WiFS Bands and Frequencies

Bands	Frequencies
Red	0.62 to 0.68 μm
NIR	0.77 to 0.86 μm

Source: Space Imaging 1999

IRS-1D

IRS-1D was launched in September of 1997. It collects imagery at a spatial resolution of 5.8 m. IRS-1D's sensors were copied for IRS-1C, which was launched in December 1995.

Imagery collected by IRS-1D is distributed in black and white format. The panchromatic imagery "reveals objects on the Earth's surface (such) as transportation networks, large ships, parks and opens space, and built-up urban areas" (Space Imaging 1999). This information can be used to classify land cover in applications such as urban planning and agriculture. The Space Imaging facility located in Norman, Oklahoma has been obtaining IRS-1D data since 1997.



For band and wavelength data on IRS-1D, see "IRS-1C" on page 59.

Source: Space Imaging 1998

Landsat 1-5

In 1972, the National Aeronautics and Space Administration (NASA) initiated the first civilian program specializing in the acquisition of remotely sensed digital satellite data. The first system was called ERTS (Earth Resources Technology Satellites), and later renamed to Landsat. There have been several Landsat satellites launched since 1972. Landsats 1, 2, and 3 are no longer operating, but Landsats 4 and 5 are still in orbit gathering data.

Landsats 1, 2, and 3 gathered Multispectral Scanner (MSS) data and Landsats 4 and 5 collect MSS and TM data. MSS and TM are discussed in more detail in the following sections.

NOTE: Landsat data are available through the Earth Observation Satellite Company (EOSAT) or the Earth Resources Observation Systems (EROS) Data Center. See "Ordering Raster Data" on page 97 for more information.

MSS

The MSS from Landsats 4 and 5 has a swath width of approximately 185×170 km from a height of approximately 900 km for Landsats 1, 2, and 3, and 705 km for Landsats 4 and 5. MSS data are widely used for general geologic studies as well as vegetation inventories.

The spatial resolution of MSS data is 56×79 m, with a 79×79 m IFOV. A typical scene contains approximately 2340 rows and 3240 columns. The radiometric resolution is 6-bit, but it is stored as 8-bit (Lillesand and Kiefer 1987).

Detectors record electromagnetic radiation (EMR) in four bands:

- Bands 1 and 2 are in the visible portion of the spectrum and are useful in detecting cultural features, such as roads. These bands also show detail in water.
- Bands 3 and 4 are in the near-infrared portion of the spectrum and can be used in land/water and vegetation discrimination.

1 =Green, 0.50-0.60 μm

This band scans the region between the blue and red chlorophyll absorption bands. It corresponds to the green reflectance of healthy vegetation, and it is also useful for mapping water bodies.

2 =Red, 0.60-0.70 μm

This is the red chlorophyll absorption band of healthy green vegetation and represents one of the most important bands for vegetation discrimination. It is also useful for determining soil boundary and geological boundary delineations and cultural features.

3 =Reflective infrared, 0.70-0.80 μm

This band is especially responsive to the amount of vegetation biomass present in a scene. It is useful for crop identification and emphasizes soil/crop and land/water contrasts.

4 =Reflective infrared, 0.80-1.10 μm

This band is useful for vegetation surveys and for penetrating haze (Jensen 1996).

TM

The TM scanner is a multispectral scanning system much like the MSS, except that the TM sensor records reflected/emitted electromagnetic energy from the visible, reflective-infrared, middle-infrared, and thermal-infrared regions of the spectrum. TM has higher spatial, spectral, and radiometric resolution than MSS.

Raster and Vector Data Sources

TM has a swath width of approximately 185 km from a height of approximately 705 km. It is useful for vegetation type and health determination, soil moisture, snow and cloud differentiation, rock type discrimination, etc.

The spatial resolution of TM is 28.5×28.5 m for all bands except the thermal (band 6), which has a spatial resolution of 120×120 m. The larger pixel size of this band is necessary for adequate signal strength. However, the thermal band is resampled to 28.5×28.5 m to match the other bands. The radiometric resolution is 8-bit, meaning that each pixel has a possible range of data values from 0 to 255.

Detectors record EMR in seven bands:

- Bands 1, 2, and 3 are in the visible portion of the spectrum and are useful in detecting cultural features such as roads. These bands also show detail in water.
- Bands 4, 5, and 7 are in the reflective-infrared portion of the spectrum and can be used in land/water discrimination.
- Band 6 is in the thermal portion of the spectrum and is used for thermal mapping (Jensen 1996; Lillesand and Kiefer 1987).

1 =Blue, 0.45-0.52 μm

This band is useful for mapping coastal water areas, differentiating between soil and vegetation, forest type mapping, and detecting cultural features.

2 =Green, 0.52-0.60 μm

This band corresponds to the green reflectance of healthy vegetation. Also useful for cultural feature identification.

3 =Red, 0.63-0.69 μm

This band is useful for discriminating between many plant species. It is also useful for determining soil boundary and geological boundary delineations as well as cultural features.

4 =Reflective-infrared, 0.76-0.90 μm

This band is especially responsive to the amount of vegetation biomass present in a scene. It is useful for crop identification and emphasizes soil/crop and land/water contrasts.

5 =Mid-infrared, 1.55-1.75 μm

This band is sensitive to the amount of water in plants, which is useful in crop drought studies and in plant health analyses. This is also one of the few bands that can be used to discriminate between clouds, snow, and ice.

6 =Thermal-infrared, 10.40-12.50 μm

This band is useful for vegetation and crop stress detection, heat intensity, insecticide applications, and for locating thermal pollution. It can also be used to locate geothermal activity.

7 =Mid-infrared, 2.08-2.35 μm

This band is important for the discrimination of geologic rock type and soil boundaries, as well as soil and vegetation moisture content.

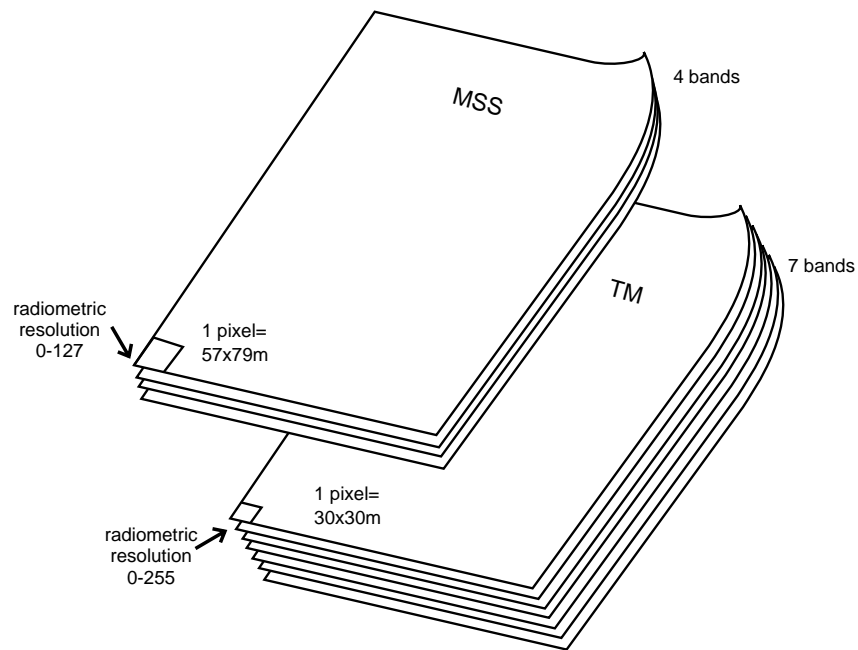


Figure 3-2 Landsat MSS vs. Landsat TM

Band Combinations for Displaying TM Data

Different combinations of the TM bands can be displayed to create different composite effects. The following combinations are commonly used to display images:

NOTE: The order of the bands corresponds to the Red, Green, and Blue (RGB) color guns of the monitor.

- Bands 3,2,1 create a true color composite. True color means that objects look as they would to the naked eye—similar to a color photograph.
- Bands 4,3,2 create a false color composite. False color composites appear similar to an infrared photograph where objects do not have the same colors or contrasts as they would naturally. For instance, in an infrared image, vegetation appears red, water appears navy or black, etc.
- Bands 5,4,2 create a pseudo color composite. (A thematic image is also a pseudo color image.) In pseudo color, the colors do *not* reflect the features in natural colors. For instance, roads may be red, water yellow, and vegetation blue.

Different color schemes can be used to bring out or enhance the features under study. These are by no means all of the useful combinations of these seven bands. The bands to be used are determined by the particular application.



See "CHAPTER 4: Image Display" for more information on how images are displayed, "CHAPTER 5: Enhancement" for more information on how images can be enhanced, and "Ordering Raster Data" on page 97 for information on types of Landsat data available.

Raster and Vector Data Sources

Landsat 7

The Landsat 7 satellite, launched in 1999, uses Enhanced Thematic Mapper Plus (ETM+) to observe the Earth. The capabilities new to Landsat 7 include the following:

- 15m spatial resolution panchromatic band
- 5% radiometric calibration with full aperture
- 60m spatial resolution thermal IR channel

The primary receiving station for Landsat 7 data is located in Sioux Falls, South Dakota at the USGS EROS Data Center (EDC). ETM+ data is transmitted using X-band direct downlink at a rate of 150 Mbps. Landsat 7 is capable of capturing scenes without cloud obstruction, and the receiving stations can obtain this data in real time using the X-band. Stations located around the globe, however, are only able to receive data for the portion of the ETM+ ground track where the satellite can be seen by the receiving station.

Landsat 7 Data Types

One type of data available from Landsat 7 is browse data. Browse data is “a lower resolution image for determining image location, quality and information content.” The other type of data is metadata, which is “descriptive information on the image.” This information is available via the internet within 24 hours of being received by the primary ground station. Moreover, EDC processes the data to Level 0r. This data has been corrected for scan direction and band alignment errors only. Level 1G data, which is corrected, is also available.

Landsat 7 Specifications

Information about the spectral range and ground resolution of the bands of the Landsat 7 satellite is provided in the following table:

Table 3-8 Landsat 7 Characteristics

Band Number	Spectral Range in microns	Ground Resolution (m)
1	.45 to .515	30
2	.525 to .605	30
3	.63 to .690	30
4	.75 to .90	30
5	1.55 to 1.75	30
6	10.40 to 12.5	60
7	2.09 to 2.35	30
Pan	.52 to .90	15

Landsat 7 has a swath width of 185 kilometers. The repeat coverage interval is 16 days, or 233 orbits. The satellite orbits the Earth at 705 kilometers.

Source: NASA 1998

NLAPS

The National Landsat Archive Production System (NLAPS) is the Landsat processing system used by EROS. The NLAPS system is able to “produce systematically-corrected, and terrain corrected products. . .” (USGS 1999).

Landsat data received from satellites is generated into TM corrected data using the NLAPS by:

- correcting and validating the mirror scan and payload correction data
- providing for image framing by generating a series of scene center parameters
- synchronizing telemetry data with video data
- estimating linear motion deviation of scan mirror/scan line corrections
- generating benchmark correction matrices for specified map projections
- producing along- and across-scan high-frequency line matrices

According to the USGS, the products provided by NLAPS include the following:

- image data and the metadata describing the image
- processing procedure, which contains information describing the process by which the image data were produced
- DEM data and the metadata describing them (available only with terrain corrected products)



For information about the Landsat data processed by NLAPS, see "Landsat 1-5" on page 60 and "Landsat 7" on page 64.

Source: USGS 1999

NOAA Polar Orbiter Data

NOAA has sponsored several polar orbiting satellites to collect data of the Earth. These satellites were originally designed for meteorological applications, but the data gathered have been used in many fields—from agronomy to oceanography (Needham 1986).

The first of these satellites to be launched was the TIROS-N in 1978. Since the TIROS-N, five additional NOAA satellites have been launched. Of these, the last three are still in orbit gathering data.

AVHRR

The NOAA AVHRR data are small-scale data and often cover an entire country. The swath width is 2700 km and the satellites orbit at a height of approximately 833 km (Kidwell 1988; Needham 1986).

The AVHRR system allows for direct transmission in real-time of data called High Resolution Picture Transmission (HRPT). It also allows for about ten minutes of data to be recorded over any portion of the world on two recorders on board the satellite. These recorded data are called Local Area Coverage (LAC). LAC and HRPT have identical formats; the only difference is that HRPT are transmitted directly and LAC are recorded.

Raster and Vector Data Sources

There are three basic formats for AVHRR data which can be imported into ERDAS IMAGINE:

- LAC—data recorded on board the sensor with a spatial resolution of approximately 1.1×1.1 km,
- HRPT—direct transmission of AVHRR data in real-time with the same resolution as LAC, and
- GAC—data produced from LAC data by using only 1 out of every 3 scan lines. GAC data have a spatial resolution of approximately 4×4 km.

AVHRR data are available in 10-bit packed and 16-bit unpacked format. The term packed refers to the way in which the data are written to the tape. Packed data are compressed to fit more data on each tape (Kidwell 1988).

AVHRR images are useful for snow cover mapping, flood monitoring, vegetation mapping, regional soil moisture analysis, wildfire fuel mapping, fire detection, dust and sandstorm monitoring, and various geologic applications (Lillesand and Kiefer 1987). The entire globe can be viewed in 14.5 days. There may be four or five bands, depending on when the data were acquired.

1 =Visible, 0.58-0.68 μm

This band corresponds to the green reflectance of healthy vegetation and is important for vegetation discrimination.

2 =Near-infrared, 0.725-1.10 μm

This band is especially responsive to the amount of vegetation biomass present in a scene. It is useful for crop identification and emphasizes soil/crop and land/water contrasts.

3 =Thermal-infrared, 3.55-3.93 μm

This is a thermal band that can be used for snow and ice discrimination. It is also useful for detecting fires.

4 =Thermal-infrared, 10.50-11.50 μm (NOAA-6, 8, 10), 10.30-11.30 μm (NOAA-7, 9, 11)

This band is useful for vegetation and crop stress detection. It can also be used to locate geothermal activity.

5 =Thermal-infrared, 10.50-11.50 μm (NOAA-6, 8, 10), 11.50-12.50 μm (NOAA-7, 9, 11)

See band 4.

AVHRR data have a radiometric resolution of 10-bits, meaning that each pixel has a possible data file value between 0 and 1023. AVHRR scenes may contain one band, a combination of bands, or all bands. All bands are referred to as a full set, and selected bands are referred to as an extract.



See "Ordering Raster Data" on page 97 for information on the types of NOAA data available.



Use the Import/Export function to import AVHRR data.

OrbView3

OrbView is a high-resolution satellite scheduled for launch by ORBIMAGE in the year 2000.

The OrbView-3 satellite will provide both 1 m panchromatic imagery and 4 m multispectral imagery. "One-meter imagery will enable the viewing of houses, automobiles and aircraft, and will make it possible to create highly precise digital maps and three-dimensional fly-through scenes. Four-meter multispectral imagery will provide color and infrared information to further characterize cities, rural areas and undeveloped land from space" (ORBIMAGE 1999). Specific applications include telecommunications and utilities, agriculture and forestry.

OrbView-3's swath width is 8 km, with an image area of 64 km². The revisit time is less than 3 days. OrbView-3 orbits the Earth at an altitude of 470 km.

Table 3-9 Orb View-3 Bands and Frequencies

Bands	Frequencies
1	450 to 520 nm
2	520 to 600 nm
3	625 to 695 nm
4	760 to 900 nm
Pan	450 to 900 nm

Source: ORBVIEW 1999

SeaWiFS

The Sea-viewing Wide Field-of-View Sensor (SeaWiFS) instrument is on-board the SeaStar spacecraft, which was launched in 1997. The SeaStar spacecraft's orbit is circular, at an altitude of 705 km. The satellite uses an attitude control system (ACS), which maintains orbit, as well as performs solar and lunar calibration maneuvers. The ACS also provides attitude information within one SeaWiFS pixel.

The SeaWiFS instrument is made up of an optical scanner and an electronics module. The swath width is 2,801 km LAC/HRPT (958.3 degrees) and 1,502 km GAC (45 degrees). The spatial resolution is 1.1 km LAC and 4.5 km GAC. The revisit time is one day.

Table 3-10 SeaWiFS Bands and Frequencies

Band	Wavelength
1	402 to 422 nm
2	433 to 453 nm
3	480 to 500 nm
4	500 to 520 nm
5	545 to 565 nm
6	660 to 680 nm
7	745 to 785 nm
8	845 to 885 nm

Source: NASA 1999

SPOT

The first SPOT satellite, developed by the French Centre National d'Etudes Spatiales (CNES), was launched in early 1986. The second SPOT satellite was launched in 1990 and the third was launched in 1993. The sensors operate in two modes, multispectral and panchromatic. SPOT is commonly referred to as a pushbroom scanner meaning that all scanning parts are fixed, and scanning is accomplished by the forward motion of the scanner. SPOT pushes 3000/6000 sensors along its orbit. This is different from Landsat which scans with 16 detectors perpendicular to its orbit.

The SPOT satellite can observe the same area on the globe once every 26 days. The SPOT scanner normally produces nadir views, but it does have off-nadir viewing capability. Off-nadir refers to any point that is not directly beneath the detectors, but off to an angle. Using this off-nadir capability, one area on the Earth can be viewed as often as every 3 days.

This off-nadir viewing can be programmed from the ground control station, and is quite useful for collecting data in a region not directly in the path of the scanner or in the event of a natural or man-made disaster, where timeliness of data acquisition is crucial. It is also very useful in collecting stereo data from which elevation data can be extracted.

The width of the swath observed varies between 60 km for nadir viewing and 80 km for off-nadir viewing at a height of 832 km (Jensen 1996).

Panchromatic

SPOT Panchromatic (meaning sensitive to all visible colors) has 10 × 10 m spatial resolution, contains 1 band—0.51 to 0.73 μm—and is similar to a black and white photograph. It has a radiometric resolution of 8 bits (Jensen 1996).

XS

SPOT XS, or multispectral, has 20 × 20 m spatial resolution, 8-bit radiometric resolution, and contains 3 bands (Jensen 1996).

1 =Green, 0.50-0.59 μm

This band corresponds to the green reflectance of healthy vegetation.

2 =Red, 0.61-0.68 μm

This band is useful for discriminating between plant species. It is also useful for soil boundary and geological boundary delineations.

3 =Reflective infrared, 0.79-0.89 μm

This band is especially responsive to the amount of vegetation biomass present in a scene. It is useful for crop identification and emphasizes soil/crop and land/water contrasts.

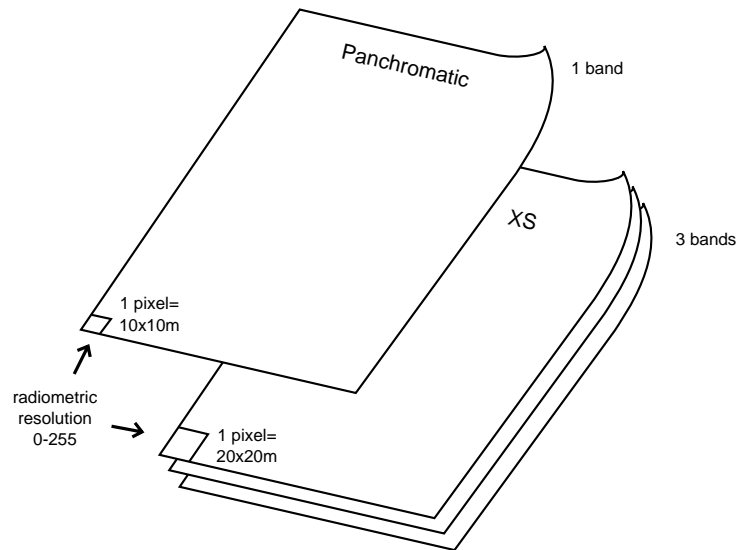


Figure 3-3 SPOT Panchromatic vs. SPOT XS



See "Ordering Raster Data" on page 97 for information on the types of SPOT data available.

Stereoscopic Pairs

Two observations can be made by the panchromatic scanner on successive days, so that the two images are acquired at angles on either side of the vertical, resulting in stereoscopic imagery. Stereoscopic imagery can also be achieved by using one vertical scene and one off-nadir scene. This type of imagery can be used to produce a single image, or topographic and planimetric maps (Jensen 1996).

Topographic maps indicate elevation. Planimetric maps correctly represent horizontal distances between objects (Star and Estes 1990).



See "Topographic Data" on page 92 and "CHAPTER 10: Terrain Analysis" for more information about topographic data and how SPOT stereopairs and aerial photographs can be used to create elevation data and orthographic images.

Raster and Vector Data Sources

SPOT4

The SPOT4 satellite was launched in 1998. SPOT 4 carries High Resolution Visible Infrared (HR VIR) instruments that obtain information in the visible and near-infrared spectral bands.

The SPOT 4 satellite orbits the Earth at 822 km at the Equator. The SPOT 4 satellite has two sensors on board: a multispectral sensor, and a panchromatic sensor. The multispectral scanner has a pixel size of 20 × 20 m, and a swath width of 60 km. The panchromatic scanner has a pixel size of 10 × 10 m, and a swath width of 60 km.

Table 3-11 SPOT 4 Bands and Frequencies

Band	Frequencies
1	0.50 to 0.59 μm
2	0.61 to 0.68 μm
3	0.79 to 0.89 μm
4	1.58 to 1.75 μm
Pan	0.61 to 0.68 μm

Source: SPOT Image 1998, 1999

Radar Data

Simply put, radar data are produced when:

- a radar transmitter emits a beam of micro or millimeter waves,
- the waves reflect from the surfaces they strike, and
- the backscattered radiation is detected by the radar system's receiving antenna, which is tuned to the frequency of the transmitted waves.

The resultant radar data can be used to produce radar images.



While there is a specific importer for data from RADARSAT and others, most types of radar image data can be imported into ERDAS IMAGINE with the Generic import option of Import/Export. The Generic SAR Node of the IMAGINE Radar Mapping Suite™ can be used to create or edit the radar ephemeris.

A radar system can be airborne, spaceborne, or ground-based. Airborne radar systems have typically been mounted on civilian and military aircraft, but in 1978, the radar satellite Seasat-1 was launched. The radar data from that mission and subsequent spaceborne radar systems have been a valuable addition to the data available for use in GIS processing. Researchers are finding that a combination of the characteristics of radar data and visible/infrared data is providing a more complete picture of the Earth. In the last decade, the importance and applications of radar have grown rapidly.

Advantages of Using Radar Data

Radar data have several advantages over other types of remotely sensed imagery:

- Radar microwaves can penetrate the atmosphere day or night under virtually all weather conditions, providing data even in the presence of haze, light rain, snow, clouds, or smoke.
- Under certain circumstances, radar can partially penetrate arid and hyperarid surfaces, revealing subsurface features of the Earth.
- Although radar does not penetrate standing water, it can reflect the surface action of oceans, lakes, and other bodies of water. Surface eddies, swells, and waves are greatly affected by the bottom features of the water body, and a careful study of surface action can provide accurate details about the bottom features.

Radar Sensors

Radar images are generated by two different types of sensors:

- SLAR (Side-looking Airborne Radar)—uses an antenna which is fixed below an aircraft and pointed to the side to transmit and receive the radar signal. (See Figure 3-4.)
- SAR—uses a side-looking, fixed antenna to create a synthetic aperture. SAR sensors are mounted on satellites and the NASA Space Shuttle. The sensor transmits and receives as it is moving. The signals received over a time interval are combined to create the image.

Both SLAR and SAR systems use side-looking geometry. Figure 3-4 shows a representation of an airborne SLAR system. Figure 3-5 shows a graph of the data received from the radiation transmitted in Figure 3-4. Notice how the data correspond to the terrain in Figure 3-4. These data can be used to produce a radar image of the target area. A target is any object or feature that is the subject of the radar scan.

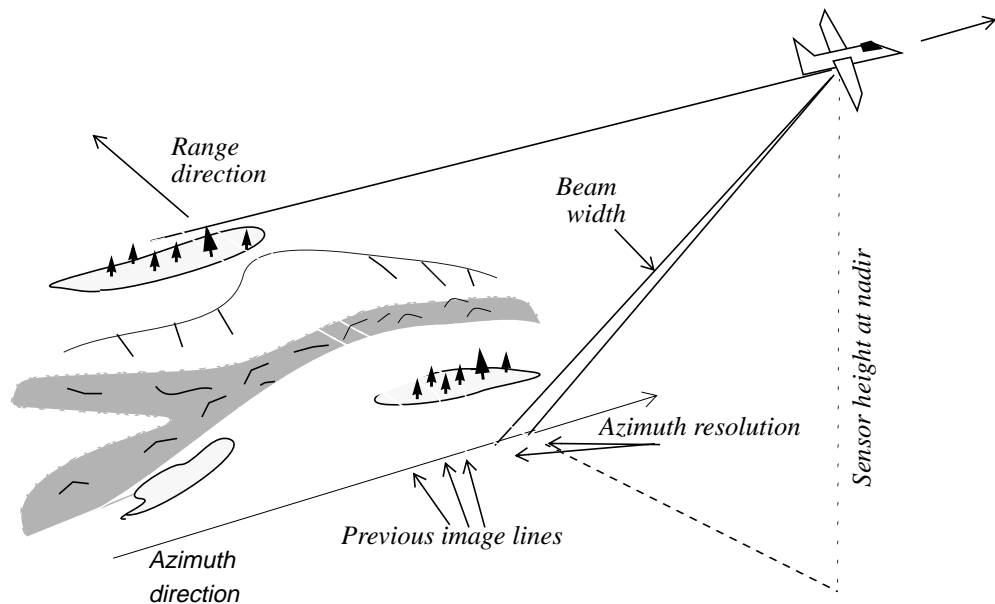


Figure 3-4 SLAR Radar

Source: Lillesand and Kiefer 1987

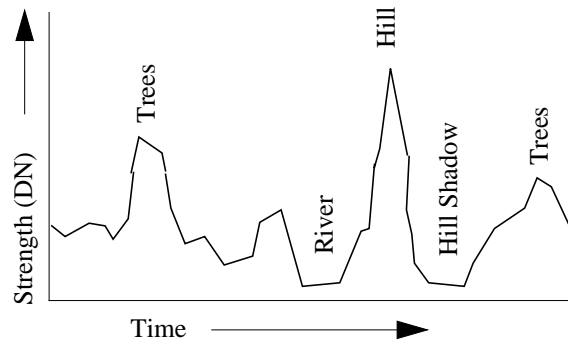


Figure 3-5 Received Radar Signal

Active and Passive Sensors

An active radar sensor gives off a burst of coherent radiation that reflects from the target, unlike a passive microwave sensor which simply receives the low-level radiation naturally emitted by targets.

Like the coherent light from a laser, the waves emitted by active sensors travel in phase and interact minimally on their way to the target area. After interaction with the target area, these waves are no longer in phase. This is due to the different distances they travel from different targets, or single versus multiple bounce scattering.

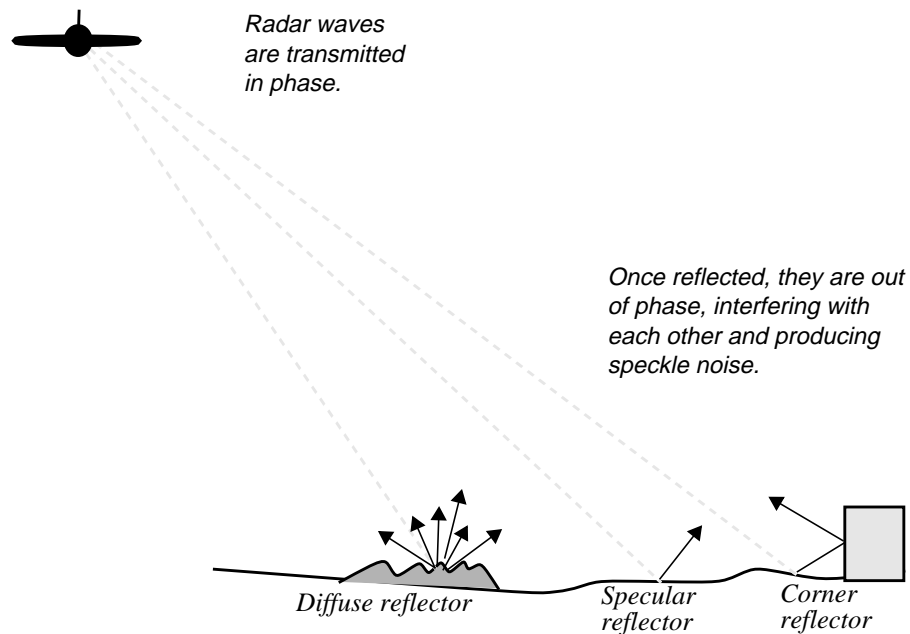


Figure 3-6 Radar Reflection from Different Sources and Distances

Source: Lillesand and Kiefer 1987

Currently, these bands are commonly used for radar imaging systems:

Table 3-12 Commonly Used Bands for Radar Imaging

Band	Frequency Range	Wavelength Range	Radar System
X	5.20-10.90 GHZ	5.77-2.75 cm	USGS SLAR
C	3.9-6.2 GHZ	3.8-7.6 cm	ERS-1, Fuyo 1
L	0.39-1.55 GHZ	76.9-19.3 cm	SIR-A,B, Almaz
P	0.225-0.391 GHZ	40.0-76.9 cm	AIRSAR



More information about these radar systems is given later in this chapter.

Radar bands were named arbitrarily when radar was first developed by the military. The letter designations have no special meaning.

NOTE: The C band overlaps the X band. Wavelength ranges may vary slightly between sensors.

Speckle Noise

Once out of phase, the radar waves can interfere constructively or destructively to produce light and dark pixels known as speckle noise. Speckle noise in radar data must be reduced before the data can be utilized. However, the radar image processing programs used to reduce speckle noise also produce changes to the image. This consideration, combined with the fact that different applications and sensor outputs necessitate different speckle removal models, has lead ERDAS to offer several speckle reduction algorithms.



When processing radar data, the order in which the image processing programs are implemented is crucial. This is especially true when considering the removal of speckle noise. Since any image processing done before removal of the speckle results in the noise being incorporated into and degrading the image, do not rectify, correct to ground range, or in any way resample the pixel values before removing speckle noise. A rotation using nearest neighbor might be permissible.

The IMAGINE Radar Interpreter™ allows you to:

- import radar data into the GIS as a stand-alone source or as an additional layer with other imagery sources
- remove speckle noise
- enhance edges
- perform texture analysis
- perform radiometric and slant-to-ground range correction

IMAGINE OrthoRadar™ allows you to orthorectify radar imagery.

The IMAGINE IFSAR DEM™ module allows you to generate DEMs from SAR data using interferometric techniques.

Raster and Vector Data Sources

The IMAGINE StereoSAR DEM™ module allows you to generate DEMs from SAR data using stereoscopic techniques.



See "CHAPTER 5: Enhancement" and "CHAPTER 8: Radar Concepts" for more information on radar imagery enhancement.

Applications for Radar Data

Radar data can be used independently in GIS applications or combined with other satellite data, such as Landsat, SPOT, or AVHRR. Possible GIS applications for radar data include:

- Geology—radar's ability to partially penetrate land cover and sensitivity to micro relief makes radar data useful in geologic mapping, mineral exploration, and archaeology.
- Classification—a radar scene can be merged with visible/infrared data as an additional layer(s) in vegetation classification for timber mapping, crop monitoring, etc.
- Glaciology—the ability to provide imagery of ocean and ice phenomena makes radar an important tool for monitoring climatic change through polar ice variation.
- Oceanography—radar is used for wind and wave measurement, sea-state and weather forecasting, and monitoring ocean circulation, tides, and polar oceans.
- Hydrology—radar data are proving useful for measuring soil moisture content and mapping snow distribution and water content.
- Ship monitoring—the ability to provide day/night all-weather imaging, as well as detect ships and associated wakes, makes radar a tool that can be used for ship navigation through frozen ocean areas such as the Arctic or North Atlantic Passage.
- Offshore oil activities—radar data are used to provide ice updates for offshore drilling rigs, determining weather and sea conditions for drilling and installation operations, and detecting oil spills.
- Pollution monitoring—radar can detect oil on the surface of water and can be used to track the spread of an oil spill.

Current Radar Sensors

Table 3-13 gives a brief description of currently available radar sensors. This is not a complete list of such sensors, but it does represent the ones most useful for GIS applications.

Table 3-13 Current Radar Sensors

	ERS-1, 2	JERS-1	SIR-A, B	SIR-C	RADARSAT	Almaz-1
Availability	operational	defunct	1981, 1984	1994	operational	1991-1992
Resolution	12.5 m	18 m	25 m	25 m	10-100 m	15 m
Revisit Time	35 days	44 days	NA	NA	3 days	NA
Scene Area	100X100 km	75X100 km	30X60 km	variable swath	50X50 to 500X500 km	40X100 km
Bands	C band	L band	L band	L,C,X bands	C band	C band

Almaz-1

Almaz was launched by the Soviet Union in 1987. The SAR operated with a single frequency SAR, which was attached to a spacecraft. Almaz-1 provided optically-processed data. The Almaz mission was largely kept secret

Almaz-1 was launched in 1991, and provides S-band information. It also includes a "single polarization SAR as well as a sounding radiometric scanner (RMS) system and several infrared bands" (Atlantis Scientific 1997).

The swath width of Almaz-1 is 20-45 km, the range resolution is 15-30 m, and the azimuth resolution is 15 m.

Source: NASA 1996; Atlantis Scientific 1997

ERS-1

ERS-1, a radar satellite, was launched by ESA in July of 1991. One of its primary instruments is the Along-Track Scanning Radiometer (ATSR). The ATSR monitors changes in vegetation of the Earth's surface.

The instruments aboard ERS-1 include: SAR Image Mode, SAR Wave Mode, Wind Scatterometer, Radar Altimeter, and Along Track Scanning Radiometer-1 (European Space Agency 1997).

ERS-1 receiving stations are located all over the world, in countries such as Sweden, Norway, and Canada.

Some of the information that is obtained from the ERS-1 (as well as ERS-2, to follow) includes:

- maps of the surface of the Earth through clouds
- physical ocean features and atmospheric phenomena
- maps and ice patterns of polar regions
- database information for use in modeling
- surface elevation changes

According to ESA,

...ERS-1 provides both global and regional views of the Earth, regardless of cloud coverage and sunlight conditions. An operational near-real-time capability for data acquisition, processing and dissemination, offering global data sets within three hours of observation, has allowed the development of time-critical applications particularly in weather, marine and ice forecasting, which are of great importance for many industrial activities (European Space Agency 1996).

Source: European Space Agency 1996

Raster and Vector Data Sources

ERS-2

ERS-2, a radar satellite, was launched by ESA in April of 1995. It has an instrument called GOME, which stands for Global Ozone Monitoring Experiment. This instrument is designed to evaluate atmospheric chemistry. ERS-2, like ERS-1 makes use of the ATSR.

The instruments aboard ERS-2 include: SAR Image Mode, SAR Wave Mode, Wind Scatterometer, Radar Altimeter, Along Track Scanning Radiometer-2, and the Global Ozone Monitoring Experiment.

ERS-2 receiving stations are located all over the world. Facilities that process and archive ERS-2 data are also located around the globe.

One of the benefits of the ERS-2 satellite is that, along with ERS-1, it can provide data from the exact same type of synthetic aperture radar (SAR).

ERS-2 provides many different types of information. See "ERS-1" for some of the most common types. Data obtained from ERS-2 used in conjunction with that from ERS-1 enables you to perform interferometric tasks. Using the data from the two sensors, DEMs can be created.



For information about ERDAS IMAGINE's interferometric software, IMAGINE IFSAR DEM, see "IMAGINE IFSAR DEM Theory" on page 315.

Source: European Space Agency 1996

JERS-1

JERS stands for Japanese Earth Resources Satellite. The JERS-1 satellite was launched in February of 1992, with an SAR instrument and a 4-band optical sensor aboard. The SAR sensor's ground resolution is 18 m, and the optical sensor's ground resolution is roughly 18 m across-track and 24 m along-track. The revisit time of the satellite is every 44 days. The satellite travels at an altitude of 568 km, at an inclination of 97.67°.

Table 3-14 JERS-1 Bands and Frequencies

Bands	Frequencies
1	.55 to .60 μm
2	.63 to .69 μm
3	.76 to .86 μm
4 ¹	.76 to .86 μm

¹ *Viewing 15.3° forward*

JERS-1 data comes in two different formats: European and Worldwide. The European data format consists mainly of coverage for Europe and Antarctica. The Worldwide data format has images that were acquired from stations around the globe. According to NASA, “a reduction in transmitter power has limited the use of JERS-1 data” (NASA 1996).

Source: Eurimage 1999; NASA 1996

RADARSAT

RADARSAT satellites carry SARs, which are capable of transmitting signals that can be received through clouds and during nighttime hours. RADARSAT satellites have multiple imaging modes for collecting data, which include Fine, Standard, Wide, ScanSAR Narrow, ScanSAR Wide, Extended (H), and Extended (L). The resolution and swath width varies with each one of these modes, but in general, Fine offers the best resolution: 8 m.

Table 3-15 RADARSAT Beam Mode Resolution

Beam Mode	Resolution
Fine Beam Mode	8 m
Standard Beam Mode	25 m
Wide Beam Mode	30 m
ScanSAR Narrow Beam Mode	50 m
ScanSAR Wide Beam Mode	100 m
Extended High Beam Mode	25 m
Low Beam Mode	35 m

The types of RADARSAT image products include: Single Data, Single Look Complex, Path Image, Path Image Plus, Map Image, Precision Map Image, and Orthorectified. You can obtain this data in forms ranging from CD-ROM to print.

The RADARSAT satellite uses a single frequency, C-band. The altitude of the satellite is 496 miles, or 798 km. The satellite is able to image the entire Earth, and its path is repeated every 24 days. The swath width is 500 km. Daily coverage is available of the Arctic, and any area of Canada can be obtained within three days.

Source: RADARSAT 1999; Space Imaging 1999

SIR-A

SIR stands for Spaceborne Imaging Radar. SIR-A was launched and began collecting data in 1981. The SIR-A mission built on the Seasat SAR mission that preceded it by increasing the incidence angle with which it captured images. The primary goal of the SIR-A mission was to collect geological information. This information did not have as pronounced a layover effect as previous imagery.

Raster and Vector Data Sources

An important achievement of SIR-A data is that it is capable of penetrating surfaces to obtain information. For example, NASA says that the L-band capability of SIR-A enabled the discovery of dry river beds in the Sahara Desert.

SIR-1 uses L-band, has a swath width of 50 km, a range resolution of 40 m, and an azimuth resolution of 40 m (Atlantis Scientific 1997).



For information on the ERDAS IMAGINE software that reduces layover effect, IMAGINE OrthoRadar, see "IMAGINE OrthoRadar Theory" on page 307.

Source: NASA 1995, 1996; Atlantis Scientific 1997

SIR-B

SIR-B was launched and began collecting data in 1984. SIR-B improved over SIR-A by using an articulating antenna. This antenna allowed the incidence angle to range between 15 and 60 degrees. This enabled the mapping of surface features using "multiple-incidence angle backscatter signatures" (NASA 1996).

SIR-B uses L-band, has a swath width of 10-60 km, a range resolution of 60-10 m, and an azimuth resolution of 25 m (Atlantis Scientific 1997).

Source: NASA 1995, 1996; Atlantis Scientific 1997

SIR-C

SIR-C is part of a radar system, SIR-C/X-SAR, which was launched in 1994. The system is able to ". . .measure, from space, the radar signature of the surface at three different wavelengths, and to make measurements for different polarizations at two of those wavelengths" (NASA 1997). Moreover, it can supply ". . .images of the magnitude of radar backscatter for four polarization combinations" (NASA 1995).

The data provided by SIR-C/X-SAR allows measurement of the following:

- vegetation type, extent, and deforestation
- soil moisture content
- ocean dynamics, wave and surface wind speeds and directions
- volcanism and tectonic activity
- soil erosion and desertification

The antenna of the system is composed of three antennas: one at L-band, one at C-band, and one at X-band. The antenna was assembled by the Jet Propulsion Laboratory. The acquisition of data at three different wavelengths makes SIR-C/X-SAR data very useful. The SIR-C and X-SAR do not have to be operated together: they can also be operated independent of one another.

SIR-C/X-SAR data come in resolutions from 10 to 200 m. The swath width of the sensor varies from 15 to 90 km, which depends on the direction the antenna is pointing. The system orbits the Earth at 225 km above the surface.

Table 3-16 SIR-C/X-SAR Bands and Frequencies

Bands	Frequencies
L-Band	0.235 m
C-Band	0.058 m
X-Band	0.031 m

Source: NASA 1995, 1997

Future Radar Sensors

Several radar satellites are planned for launch within the next several years, but only a few programs will be successful. Following are two scheduled programs which are known to be highly achievable.

Light SAR

NASA and Jet Propulsion Laboratories (JPL) are currently designing a radar satellite called Light SAR. Present plans are for this to be a multipolar sensor operating at L-band.

Radarsat-2

The Canadian Space Agency is working on the follow-on system to Radarsat 1. Present plans are to include multipolar, C-band imagery.

Image Data from Aircraft

Image data can also be acquired from multispectral scanners or radar sensors aboard aircraft, as well as satellites. This is useful if there is not time to wait for the next satellite to pass over a particular area, or if it is necessary to achieve a specific spatial or spectral resolution that cannot be attained with satellite sensors.

For example, this type of data can be beneficial in the event of a natural or man-made disaster, because there is more control over when and where the data are gathered.

Two common types of airborne image data are:

- Airborne Synthetic Aperture Radar (AIRSAR)
- Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)

AIRSAR

AIRSAR is an experimental airborne radar sensor developed by JPL, Pasadena, California, under a contract with NASA. AIRSAR data have been available since 1983.

This sensor collects data at three frequencies:

- C-band
- L-band
- P-band

Because this sensor measures at three different wavelengths, different scales of surface roughness are obtained. The AIRSAR sensor has an IFOV of 10 m and a swath width of 12 km.

Raster and Vector Data Sources

AIRSAR data have been used in many applications such as measuring snow wetness, classifying vegetation, and estimating soil moisture.

NOTE: These data are distributed in a compressed format. They must be decompressed before loading with an algorithm available from JPL. See "Addresses to Contact" on page 98 for contact information.

AVIRIS

The AVIRIS was also developed by JPL under a contract with NASA. AVIRIS data have been available since 1987.

This sensor produces multispectral data that have 224 narrow bands. These bands are 10 nm wide and cover the spectral range of .4 - 2.4 μm . The swath width is 11 km, and the spatial resolution is 20 m. This sensor is flown at an altitude of approximately 20 km. The data are recorded at 10-bit radiometric resolution.

DAEDALUS TMS

DAEDALUS is a thematic mapper simulator (TMS), which simulates the characteristics, such as spatial and radiometric, of the TM sensor on Landsat spacecraft.

The DAEDALUS TMS orbits at 65,000 feet, and has a ground resolution of 25 meters. The total scan angle is 43 degrees, and the swath width is 15.6 km. DAEDALUS TMS is flown aboard the NASA ER-2 aircraft.

The DAEDALUS TMS spectral bands are as follows:

Table 3-17 DAEDALUS TMS Bands and Frequencies

DAEDALUS Channel	TM Band	Wavelength
1	A	0.42 to 0.45 μm
2	1	0.45 to 0.52 μm
3	2	0.52 to 0.60 μm
4	B	0.60 to 0.62 μm
5	3	0.63 to 0.69 μm
6	C	0.69 to 0.75 μm
7	4	0.76 to 0.90 μm
8	D	0.91 to 1.05 μm
9	5	1.55 to 1.75 μm
10	7	2.08 to 2.35 μm
11	6	8.5 to 14.0 low gain μm
12	6	8.5 to 14.0 high gain μm

Source: NASA 1995

Image Data from Scanning

Hardcopy maps and photographs can be incorporated into the ERDAS IMAGINE environment through the use of a scanning device to transfer them into a digital (raster) format.

In scanning, the map, photograph, transparency, or other object to be scanned is typically placed on a flat surface, and the scanner scans across the object to record the image. The image is then transferred from analog to digital data.

There are many commonly used scanners for GIS and other desktop applications, such as Eikonix (Eikonix Corp., Huntsville, Alabama) or Vexcel (Vexcel Imaging Corp., Boulder, Colorado). Many scanners produce a Tagged Image File Format (TIFF) file, which can be used directly by ERDAS IMAGINE.



Use the Import/Export function to import scanned data.

Eikonix data can be obtained in the ERDAS IMAGINE .img format using the XSCAN™ Tool by Ektron and then imported directly into ERDAS IMAGINE.

Photogrammetric Scanners

There are photogrammetric high quality scanners and desktop scanners. Photogrammetric quality scanners are special devices capable of high image quality and excellent positional accuracy. Use of this type of scanner results in geometric accuracies similar to traditional analog and analytical photogrammetric instruments. These scanners are necessary for digital photogrammetric applications that have high accuracy requirements.

These units usually scan only film because film is superior to paper, both in terms of image detail and geometry. These units usually have a Root Mean Square Error (RMSE) positional accuracy of 4 microns or less, and are capable of scanning at a maximum resolution of 5 to 10 microns.

The required pixel resolution varies depending on the application. Aerial triangulation and feature collection applications often scan in the 10 to 15 micron range. Orthophoto applications often use 15- to 30-micron pixels. Color film is less sharp than panchromatic, therefore color ortho applications often use 20- to 40-micron pixels.

Desktop Scanners

Desktop scanners are general purpose devices. They lack the image detail and geometric accuracy of photogrammetric quality units, but they are much less expensive. When using a desktop scanner, you should make sure that the active area is at least 9 × 9 inches (i.e., A3-type scanners), enabling you to capture the entire photo frame.

Desktop scanners are appropriate for less rigorous uses, such as digital photogrammetry in support of GIS or remote sensing applications. Calibrating these units improves geometric accuracy, but the results are still inferior to photogrammetric units. The image correlation techniques that are necessary for automatic tie point collection and elevation extraction are often sensitive to scan quality. Therefore, errors can be introduced into the photogrammetric solution that are attributable to scanning errors.

Raster and Vector Data Sources

Aerial Photography

Aerial photographs such as NAPP photos are most widely used data sources in photogrammetry. They can not be utilized in softcopy or digital photogrammetric applications until scanned. The standard dimensions of the aerial photos are 9×9 inches or 230×230 mm. The ground area covered by the photo depends on the scale. The scanning resolution determines the digital image file size and pixel size.

For example, for a 1:40,000 scale standard block of white aerial photos scanned at 25 microns (1016 dots per inch), the ground pixel size is 1×1 m². The resulting file size is about 85 MB. It is not recommended to scan a photo with a scanning resolution less than 5 microns or larger than 5080 dpi.

DOQs

DOQ stands for digital orthophoto quadrangle. USGS defines a DOQ as a computer-generated image of an aerial photo, which has been orthorectified to give it map coordinates. DOQs can provide accurate map measurements.

The format of the DOQ is a grayscale image that covers 3.75 minutes of latitude by 3.75 minutes of longitude. DOQs use the North American Datum of 1983, and the Universal Transverse Mercator projection. Each pixel of a DOQ represents a square meter. 3.75-minute quarter quadrangles have a 1:12,000 scale. 7.5-minute quadrangles have a 1:24,000 scale. Some DOQs are available in color-infrared, which is especially useful for vegetation monitoring.

DOQs can be used in land use and planning, management of natural resources, environmental impact assessments, and watershed analysis, among other applications. A DOQ can also be used as “a cartographic base on which to overlay any number of associated thematic layers for displaying, generating, and modifying planimetric data or associated data files” (USGS 1999).

According to the USGS:

DOQ production begins with an aerial photo and requires four elements: (1) at least three ground positions that can be identified within the photo; (2) camera calibration specifications, such as focal length; (3) a digital elevation model (DEM) of the area covered by the photo; (4) and a high-resolution digital image of the photo, produced by scanning. The photo is processed pixel by pixel to produce an image with features in true geographic positions (USGS 1999).

Source: USGS 1999

ADRG Data

ADRG (ARC Digitized Raster Graphic) data come from the National Imagery and Mapping Agency (NIMA), which was formerly known as the Defense Mapping Agency (DMA). ADRG data are primarily used for military purposes by defense contractors. The data are in 128×128 pixel tiled, 8-bit format stored on CD-ROM. ADRG data provide large amounts of hardcopy graphic data without having to store and maintain the actual hardcopy graphics.

ADRG data consist of digital copies of NIMA hardcopy graphics transformed into the ARC system and accompanied by ASCII encoded support files. These digital copies are produced by scanning each hardcopy graphic into three images: red, green, and blue. The data are scanned at a nominal collection interval of 100 microns (254 lines per inch). When these images are combined, they provide a 3-band digital representation of the original hardcopy graphic.

ARC System

The ARC system (Equal Arc-Second Raster Chart/Map) provides a rectangular coordinate and projection system at any scale for the Earth's ellipsoid, based on the World Geodetic System 1984 (WGS 84). The ARC System divides the surface of the ellipsoid into 18 latitudinal bands called zones. Zones 1 - 9 cover the Northern hemisphere and zones 10 - 18 cover the Southern hemisphere. Zone 9 is the North Polar region. Zone 18 is the South Polar region.

Distribution Rectangles

For distribution, ADRG are divided into geographic data sets called Distribution Rectangles (DRs). A DR may include data from one or more source charts or maps. The boundary of a DR is a geographic rectangle that typically coincides with chart and map neatlines.

Zone Distribution Rectangles

Each DR is divided into Zone Distribution Rectangles (ZDRs). There is one ZDR for each ARC System zone covering any part of the DR. The ZDR contains all the DR data that fall within that zone's limits. ZDRs typically overlap by 1,024 rows of pixels, which allows for easier mosaicking. Each ZDR is stored on the CD-ROM as a single raster image file (.IMG). Included in each image file are all raster data for a DR from a single ARC System zone, and padding pixels needed to fulfill format requirements. The padding pixels are black and have a zero value.



The padding pixels are not imported by ERDAS IMAGINE, nor are they counted when figuring the pixel height and width of each image.

ADRG File Format

Each CD-ROM contains up to eight different file types which make up the ADRG format. ERDAS IMAGINE imports three types of ADRG data files:

- .OVR (Overview)
- .IMG (Image)
- .Lxx (Legend or marginalia data)

NOTE: Compressed ADRG (CADRG) is a different format, which may be imported or read directly.



The ADRG .IMG and .OVR file formats are different from the ERDAS IMAGINE .img and .ovr file formats.

.OVR (overview)

The overview file contains a 16:1 reduced resolution image of the whole DR. There is an overview file for each DR on a CD-ROM.

Raster and Vector Data Sources

Importing ADRG Subsets

Since DRs can be rather large, it may be beneficial to import a subset of the DR data for the application. ERDAS IMAGINE enables you to define a subset of the data from the preview image (see Figure 3-8).



You can import from only one ZDR at a time. If a subset covers multiple ZDRs, they must be imported separately and mosaicked with the Mosaic option.

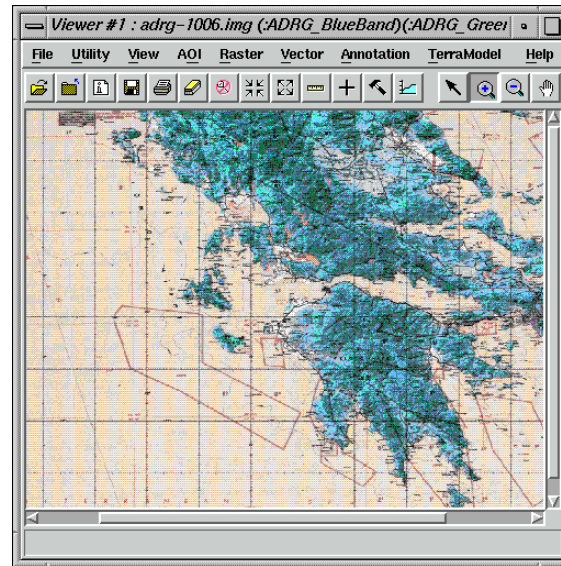


Figure 3-7 ADRG Overview File Displayed in a Viewer

The white rectangle in Figure 3-8 represents the DR. The subset area in this illustration would have to be imported as three files: one for each zone in the DR.

Notice how the ZDRs overlap. Therefore, the .IMG files for Zones 2 and 4 would also be included in the subset area.

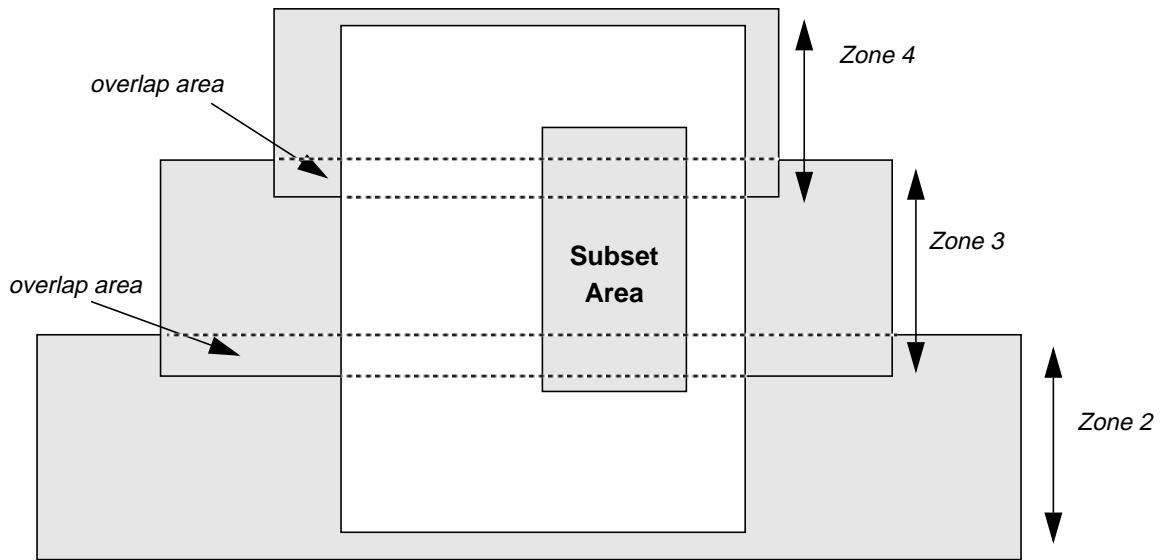


Figure 3-8 Subset Area with Overlapping ZDRs

.IMG (scanned image data)

The .IMG files are the data files containing the actual scanned hardcopy graphic(s). Each .IMG file contains one ZDR plus padding pixels. The Import function converts the .IMG data files on the CD-ROM to the ERDAS IMAGINE file format (.img). The image file can then be displayed in a Viewer.

.Lxx (legend data)

Legend files contain a variety of diagrams and accompanying information. This is information that typically appears in the margin or legend of the source graphic.



This information can be imported into ERDAS IMAGINE and viewed. It can also be added to a map composition with the ERDAS IMAGINE Map Composer.

Each legend file contains information based on one of these diagram types:

- Index (IN)—shows the approximate geographical position of the graphic and its relationship to other graphics in the region.
- Elevation/Depth Tint (EL)—depicts the colors or tints using a multicolored graphic that represent different elevations or depth bands on the printed map or chart.
- Slope (SL)—represents the percent and degree of slope appearing in slope bands.
- Boundary (BN)—depicts the geopolitical boundaries included on the map or chart.
- Accuracy (HA, VA, AC)—depicts the horizontal and vertical accuracies of selected map or chart areas. AC represents a combined horizontal and vertical accuracy diagram.
- Geographic Reference (GE)—depicts the positioning information as referenced to the World Geographic Reference System.

Raster and Vector Data Sources

- Grid Reference (GR)—depicts specific information needed for positional determination with reference to a particular grid system.
- Glossary (GL)—gives brief lists of foreign geographical names appearing on the map or chart with their English-language equivalents.
- Landmark Feature Symbols (LS)—depict navigationally-prominent entities.

ARC System Charts

The ADRG data on each CD-ROM are based on one of these chart types from the ARC system:

Table 3-18 ARC System Chart Types

ARC System Chart Type	Scale
GNC (Global Navigation Chart)	1:5,000,000
JNC-A (Jet Navigation Chart - Air)	1:3,000,000
JNC (Jet Navigation Chart)	1:2,000,000
ONC (Operational Navigation Chart)	1:1,000,000
TPC (Tactical Pilot Chart)	1:500,000
JOG-A (Joint Operations Graphic - Air)	1:250,000
JOG-G (Joint Operations Graphic - Ground)	1:250,000
JOG-C (Joint Operations Graphic - Combined)	1:250,000
JOG-R (Joint Operations Graphic - Radar)	1:250,000
ATC (Series 200 Air Target Chart)	1:200,000
TLM (Topographic Line Map)	1:50,000

Each ARC System chart type has certain legend files associated with the image(s) on the CD-ROM. The legend files associated with each chart type are checked in Table 3-19.

Table 3-19 Legend Files for the ARC System Chart Types

ARC System Chart	IN	EL	SL	BN	VA	HA	AC	GE	GR	GL	LS
GNC	•	•									
JNC / JNC-A	•	•			•		•	•			
ONC	•	•						•		•	•
TPC	•	•						•	•	•	•
JOG-A	•	•			•	•	•	•	•	•	
JOG-G / JOG-C	•				•	•	•	•	•	•	

Table 3-19 Legend Files for the ARC System Chart Types (Continued)

ARC System Chart	IN	EL	SL	BN	VA	HA	AC	GE	GR	GL	LS
JOG-R	•				•	•	•		•	•	
ATC	•				•	•	•		•		
TLM	•	•	•	•					•	•	

ADRG File Naming Convention

The ADRG file naming convention is based on a series of codes: sscddzz

- ss = the chart series code (see the table of ARC System charts)
- cc = the country code
- dd = the DR number on the CD-ROM (01-99). DRs are numbered beginning with 01 for the northwesternmost DR and increasing sequentially west to east, then north to south.
- zz = the zone rectangle number (01-18)

For example, in the ADRG filename JNUR0101.IMG:

- JN = Jet Navigation. This ADRG file is taken from a Jet Navigation chart.
- UR = Europe. The data is coverage of a European continent.
- 01 = This is the first DR on the CD-ROM, providing coverage of the northwestern edge of the image area.
- 01 = This is the first zone rectangle of the DR.
- .IMG = This file contains the actual scanned image data for a ZDR.



You may change this name when the file is imported into ERDAS IMAGINE. If you do not specify a file name, ERDAS IMAGINE uses the ADRG file name for the image.

Legend File Names

Legend file names include a code to designate the type of diagram information contained in the file (see the previous legend file description). For example, the file JNUR01IN.L01 means:

- JN = Jet Navigation. This ADRG file is taken from a Jet Navigation chart.
- UR = Europe. The data is coverage of a European continent.
- 01 = This is the first DR on the CD-ROM, providing coverage of the northwestern edge of the image area.
- IN = This indicates that this file is an index diagram from the original hardcopy graphic.
- .L01 = This legend file contains information for the source graphic 01. The source graphics in each DR are numbered beginning with 01 for the northwesternmost source graphic, increasing sequentially west to east, then north to south. Source directories and their files include this number code within their names.

Raster and Vector Data Sources



For more detailed information on ADRG file naming conventions, see the *National Imagery and Mapping Agency Product Specifications for ARC Digitized Raster Graphics (ADRG)*, published by the NIMA Aerospace Center.

ADRI Data

ADRI (ARC Digital Raster Imagery), like ADRG data, are also from the NIMA and are currently available only to Department of Defense contractors. The data are in 128 × 128 tiled, 8-bit format, stored on 8 mm tape in band sequential format.

ADRI consists of SPOT panchromatic satellite imagery transformed into the ARC system and accompanied by ASCII encoded support files.

Like ADRG, ADRI data are stored in the ARC system in DRs. Each DR consists of all or part of one or more images mosaicked to meet the ARC bounding rectangle, which encloses a 1 degree by 1 degree geographic area. (See Figure 3-9.) Source images are orthorectified to mean sea level using NIMA Level I Digital Terrain Elevation Data (DTED) or equivalent data (Air Force Intelligence Support Agency 1991).



See the previous section on ADRG data for more information on the ARC system. See more about DTED data on page 94.

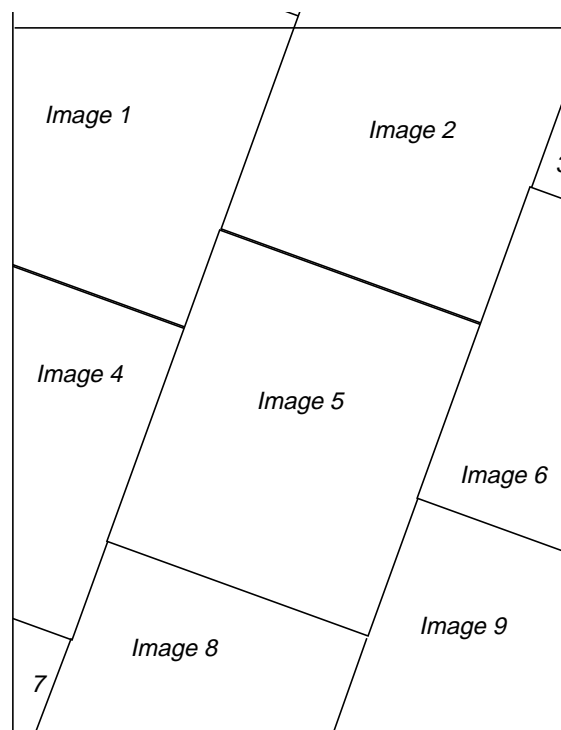


Figure 3-9 Seamless Nine Image DR

In ADRI data, each DR contains only one ZDR. Each ZDR is stored as a single raster image file, with no overlapping areas.

There are six different file types that make up the ADRI format: two types of data files, three types of header files, and a color test patch file. ERDAS IMAGINE imports two types of ADRI data files:

- .OVR (Overview)
- .IMG (Image)



The ADRI .IMG and .OVR file formats are different from the ERDAS IMAGINE .img and .ovr file formats.

.OVR (overview)

The overview file (.OVR) contains a 16:1 reduced resolution image of the whole DR. There is an overview file for each DR on a tape. The .OVR images show the mosaicking from the source images and the dates when the source images were collected. (See Figure 3-10.) This does not appear on the ZDR image.

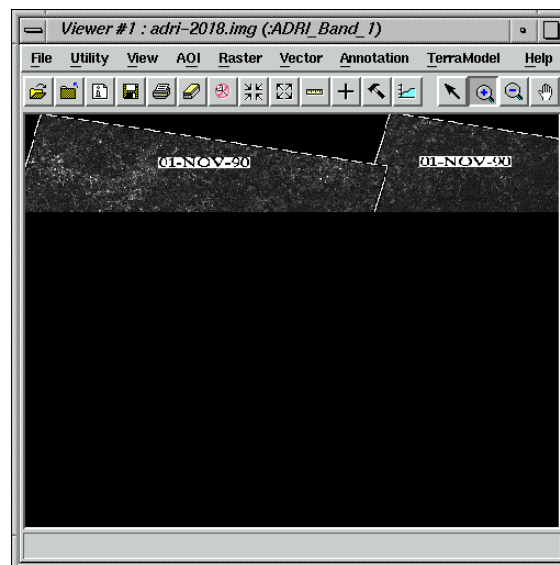


Figure 3-10 ADRI Overview File Displayed in a Viewer

.IMG (scanned image data)

The .IMG files contain the actual mosaicked images. Each .IMG file contains one ZDR plus any padding pixels needed to fit the ARC boundaries. Padding pixels are black and have a zero data value. The ERDAS IMAGINE Import function converts the .IMG data files to the ERDAS IMAGINE file format (.img). The image file can then be displayed in a Viewer. Padding pixels are not imported, nor are they counted in image height or width.

ADRI File Naming Convention

The ADRI file naming convention is based on a series of codes: sscddzz

- ss = the image source code:

Raster and Vector Data Sources

- SP (SPOT panchromatic)
- SX (SPOT multispectral) (not currently available)
- TM (Landsat Thematic Mapper) (not currently available)
- cc = the country code
- dd = the DR number on the tape (01-99). DRs are numbered beginning with 01 for the northwesternmost DR and increasing sequentially west to east, then north to south.
- zz = the zone rectangle number (01-18)

For example, in the ADRI filename SPUR0101.IMG:

- SP = SPOT 10 m panchromatic image
- UR = Europe. The data is coverage of a European continent.
- 01 = This is the first Distribution Rectangle on the CD-ROM, providing coverage of the northwestern edge of the image area.
- 01 = This is the first zone rectangle of the Distribution Rectangle.
- .IMG = This file contains the actual scanned image data for a ZDR.



You may change this name when the file is imported into ERDAS IMAGINE. If you do not specify a file name, ERDAS IMAGINE uses the ADRI file name for the image.

Raster Product Format

The Raster Product Format (RPF), from NIMA, is primarily used for military purposes by defense contractors. RPF data are organized in 1536×1536 frames, with an internal tile size of 256×256 pixels. RPF data are stored in an 8-bit format, with or without a pseudocolor lookup table, on CD-ROM.

RPF Data are projected to the ARC system, based on the World Geodetic System 1984 (WGS 84). The ARC System divides the surface of the ellipsoid into 18 latitudinal bands called zones. Zones 1-9 cover the Northern hemisphere and zones A-J cover the Southern hemisphere. Zone 9 is the North Polar region. Zone J is the South Polar region

Polar data is projected to the Azimuthal Equidistant projection. In nonpolar zones, data is in the Equirectangular projection, which is proportional to latitude and longitude. ERDAS IMAGINE V8.4 introduces the option to use either Equirectangular or Geographic coordinates for nonpolar RPF data. The aspect ratio of projected RPF data is nearly 1; frames appear to be square, and measurement is possible. Unprojected RPFs seldom have an aspect of ratio of 1, but may be easier to combine with other data in Geographic coordinates.

Two military products are currently based upon the general RPF specification:

- Controlled Image Base (CIB)
- Compressed ADRG (CADRG)

RPF employs Vector Quantization (VQ) to compress the frames. A vector is a 4×4 tile of 8-bit pixel values. VQ evaluates all of the vectors within the image, and reduces each vector into a single 12-bit lookup value. Since only 4096 unique vector values are possible, VQ is lossy, but the space savings are substantial. Most of the processing effort of VQ is incurred in the compression stage, permitting fast decompression by the users of the data in the field.

RPF data are stored on CD-ROM, with the following structure:

- The root of the CD-ROM contains an RPF directory. This RPF directory is often referred to as the root of the product.
- The RPF directory contains a table-of-contents file, named A.TOC, which describes the location of all of the frames in the product, and
- The RPF directory contains one or more subdirectories containing RPF frame files. RPF frame file names typically encode the map zone and location of the frame within the map series.
- Overview images may appear at various points in the directory tree. Overview images illustrate the location of a set of frames with respect to political and geographic boundaries. Overview images typically have an .OVx file extension, such as .OVR or .OV1.

All RPF frames, overview images, and table-of-contents files are physically formatted within an NITF message. Since an RPF image is broken up into several NITF messages, ERDAS IMAGINE treats RPF and NITF as distinct formats.

Loading RPF Data

RPF frames may be imported or read directly. The direct read feature, introduced in ERDAS IMAGINE V8.4, is generally preferable since multiple frames with the same resolution can be read as a single image. Import may still be desirable if you wish to examine the metadata provided by a specific frame. ERDAS IMAGINE V8.4 supplies four image types related to RPF:

- RPF Product—combines the entire contents of an RPF CD, excluding overview images, as a single image, provided all frames are within the same ARC map zone and resolution. The RPF directory at the root of the CD-ROM is the image to be loaded.
- RPF Cell—combines all of the frames within a given subdirectory, provided they all have the same resolution and reside within the same ARC map zone. The directory containing frames is the image to be read as an RPF cell.
- RPF Frame—reads a single frame file.
- RPF Overview—reads a single overview frame file.

CIB

CIB is grayscale imagery produced from rectified imagery and physically formatted as a compressed RPF. CIB offers a compression ratio of 8:1 over its predecessor, ADRI. CIB is often based upon SPOT panchromatic data or reformatted ADRI data, but can be produced from other sources of imagery.

Raster and Vector Data Sources

CADRG

CADRG data consist of digital copies of NIMA hardcopy graphics transformed into the ARC system. The data are scanned at a nominal collection interval of 150 microns. The resulting image is 8-bit pseudocolor, which is physically formatted as a compressed RPF.

CADRG is a successor to ADRG, Compressed Aeronautical Chart (CAC), and Compressed Raster Graphics (CRG). CADRG offers a compression ratio of 55:1 over ADRG, due to the coarser collection interval, VQ compression, and the encoding as 8-bit pseudocolor, instead of 24-bit truecolor.

Topographic Data

Satellite data can also be used to create elevation, or topographic data through the use of stereoscopic pairs, as discussed above under SPOT. Radar sensor data can also be a source of topographic information, as discussed in "Chapter 8: Terrain Analysis." However, most available elevation data are created with stereo photography and topographic maps.

ERDAS IMAGINE software can load and use:

- USGS DEMs
- DTED

Arc/second Format

Most elevation data are in arc/second format. Arc/second refers to data in the Latitude/Longitude (Lat/Lon) coordinate system. The data are not rectangular, but follow the arc of the Earth's latitudinal and longitudinal lines.

Each degree of latitude and longitude is made up of 60 minutes. Each minute is made up of 60 seconds. Arc/second data are often referred to by the number of seconds in each pixel. For example, 3 arc/second data have pixels which are 3 × 3 seconds in size. The actual area represented by each pixel is a function of its latitude. Figure 3-11 illustrates a 1° × 1° area of the Earth.

A row of data file values from a DEM or DTED file is called a profile. The profiles of DEM and DTED run south to north, that is, the first pixel of the record is the southernmost pixel.

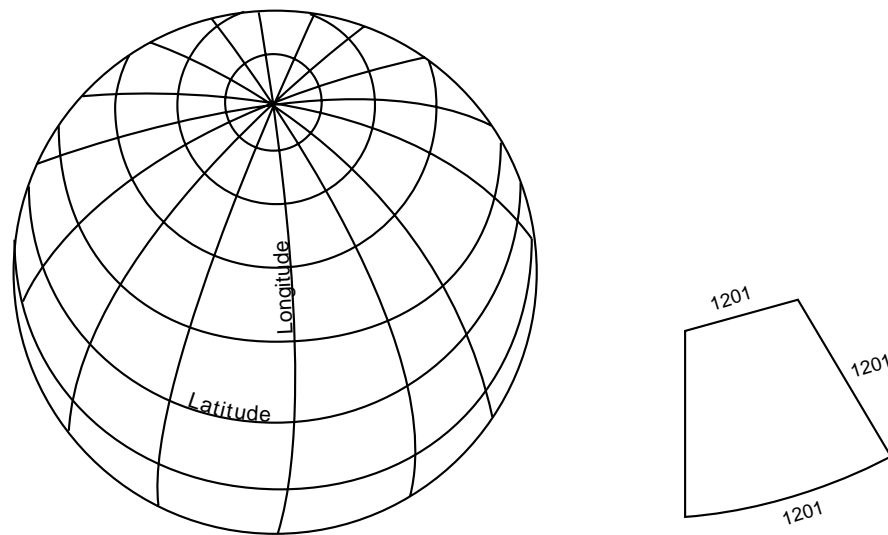


Figure 3-11 Arc/second Format

In Figure 3-11, there are 1201 pixels in the first row and 1201 pixels in the last row, but the area represented by each pixel increases in size from the top of the file to the bottom of the file. The extracted section in the example above has been exaggerated to illustrate this point.

Arc/second data used in conjunction with other image data, such as TM or SPOT, must be rectified or projected onto a planar coordinate system such as UTM.

DEM

DEMs are digital elevation model data. DEM was originally a term reserved for elevation data provided by the USGS, but it is now used to describe any digital elevation data.

DEMs can be:

- purchased from USGS (for US areas only)
- created from stereopairs (derived from satellite data or aerial photographs)



See "CHAPTER 10: Terrain Analysis" for more information on using DEMs. See "Ordering Raster Data" on page 97 for information on ordering DEMs.

USGS DEMs

There are two types of DEMs that are most commonly available from USGS:

- 1:24,000 scale, also called 7.5-minute DEM, is usually referenced to the UTM coordinate system. It has a spatial resolution of 30×30 m.
- 1:250,000 scale is available only in Arc/second format.

Both types have a 16-bit range of elevation values, meaning each pixel can have a possible elevation of -32,768 to 32,767.

Raster and Vector Data Sources

DEM data are stored in ASCII format. The data file values in ASCII format are stored as ASCII characters rather than as zeros and ones like the data file values in binary data.

DEM data files from USGS are initially oriented so that North is on the right side of the image instead of at the top. ERDAS IMAGINE rotates the data 90° counterclockwise as part of the Import process so that coordinates read with any ERDAS IMAGINE program are correct.

DTED

DTED data are produced by the National Imagery and Mapping Agency (NIMA) and are available only to US government agencies and their contractors. DTED data are distributed on 9-track tapes and on CD-ROM.

There are two types of DTED data available:

- DTED 1—a 1° × 1° area of coverage
- DTED 2—a 1° × 1° or less area of coverage

Both are in Arc/second format and are distributed in cells. A cell is a 1° × 1° area of coverage. Both have a 16-bit range of elevation values.

Like DEMs, DTED data files are also oriented so that North is on the right side of the image instead of at the top. ERDAS IMAGINE rotates the data 90° counterclockwise as part of the Import process so that coordinates read with any ERDAS IMAGINE program are correct.

Using Topographic Data

Topographic data have many uses in a GIS. For example, topographic data can be used in conjunction with other data to:

- calculate the shortest and most navigable path over a mountain range
- assess the visibility from various lookout points or along roads
- simulate travel through a landscape
- determine rates of snow melt
- orthorectify satellite or airborne images
- create aspect and slope layers
- provide ancillary data from image classification



See "CHAPTER 10: Terrain Analysis" for more information about using topographic and elevation data.

GPS Data

Introduction

Global Positioning System (GPS) data has been in existence since the launch of the first satellite in the US Navigation System with Time and Ranging (NAVSTAR) system on February 22, 1978, and the availability of a full constellation of satellites since 1994. Initially, the system was available to US military personnel only, but from 1993 onwards the system started to be used (in a degraded mode) by the general public. There is also a Russian GPS system called GLONASS with similar capabilities.

The US NAVSTAR GPS consists of a constellation of 24 satellites orbiting the Earth, broadcasting data that allows a GPS receiver to calculate its spatial position.

Satellite Position

Positions are determined through the traditional ranging technique. The satellites orbit the Earth (at an altitude of 20,200 km) in such a manner that several are always visible at any location on the Earth's surface. A GPS receiver with line of site to a GPS satellite can determine how long the signal broadcast by the satellite has taken to reach its location, and therefore can determine the distance to the satellite. Thus, if the GPS receiver can see three or more satellites and determine the distance to each, the GPS receiver can calculate its own position based on the known positions of the satellites (i.e., the intersection of the spheres of distance from the satellite locations). Theoretically, only three satellites should be required to find the 3D position of the receiver, but various inaccuracies (largely based on the quality of the clock within the GPS receiver that is used to time the arrival of the signal) mean that at least four satellites are generally required to determine a three-dimensional (3D) x, y, z position.

The explanation above is an over-simplification of the technique used, but does show the concept behind the use of the GPS system for determining position. The accuracy of that position is affected by several factors, including the number of satellites that can be seen by a receiver, but especially for commercial users by Selective Availability. Each satellite actually sends two signals at different frequencies. One is for civilian use and one for military use. The signal used for commercial receivers has an error introduced to it called Selective Availability. Selective Availability introduces a positional inaccuracy of up to 100m to commercial GPS receivers. This is mainly intended to limit the use of highly accurate GPS positioning to hostile users, but the errors can be ameliorated through various techniques, such as keeping the GPS receiver stationary; thereby allowing it to average out the errors, or through more advanced techniques discussed in the following sections.

Differential Correction

Differential Correction (or Differential GPS - DGPS) can be used to remove the majority of the effects of Selective Availability. The technique works by using a second GPS unit (or base station) that is stationary at a precisely known position. As this GPS knows where it actually is, it can compare this location with the position it calculates from GPS satellites at any particular time and calculate an error vector for that time (i.e., the distance and direction that the GPS reading is in error from the real position). A log of such error vectors can then be compared with GPS readings taken from the first, mobile unit (the field unit that is actually taking GPS location readings of features). Under the assumption that the field unit had line of site to the same GPS satellites to acquire its position as the base station, each field-read position (with an appropriate time stamp) can be compared to the error vector for that time and the position corrected using the inverse of the vector. This is generally performed using specialist differential correction software.

Real Time Differential GPS (RDGPS) takes this technique one step further by having the base station communicate the error vector via radio to the field unit in real time. The field unit can then automatically update its own location in real time. The main disadvantage of this technique is that the range that a GPS base station can broadcast over is generally limited, thereby restricting the range the mobile unit can be used away from the base station. One of the biggest uses of this technique is for ocean navigation in coastal areas, where base stations have been set up along coastlines and around ports so that the GPS systems on board ships can get accurate real time positional information to help in shallow-water navigation.

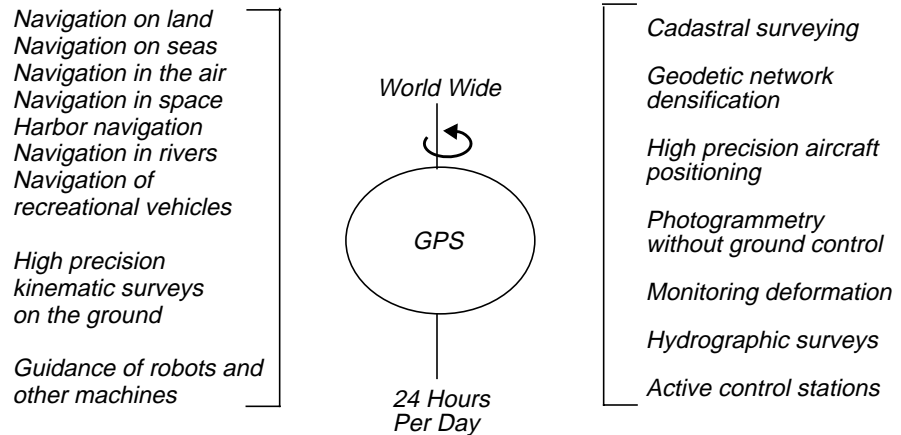
Raster and Vector Data Sources

Applications of GPS Data

GPS data finds many uses in remote sensing and GIS applications, such as:

- Collection of ground truth data, even spectral properties of real-world conditions at known geographic positions, for use in image classification and validation. The user in the field identifies a homogeneous area of identifiable land cover or use on the ground and records its location using the GPS receiver. These locations can then be plotted over an image to either train a supervised classifier or to test the validity of a classification.
- Moving map applications take the concept of relating the GPS positional information to your geographic data layers one step further by having the GPS position displayed in real time over the geographical data layers. Thus you take a computer out into the field and connect the GPS receiver to the computer, usually via the serial port. Remote sensing and GIS data layers are then displayed on the computer and the positional signal from the GPS receiver is plotted on top of them.
- GPS receivers can be used for the collection of positional information for known point features on the ground. If these can be identified in an image, the positional data can be used as Ground Control Points (GCPs) for geocorrecting the imagery to a map projection system. If the imagery is of high resolution, this generally requires differential correction of the positional data.
- DGPS data can be used to directly capture GIS data and survey data for direct use in a GIS or CAD system. In this regard the GPS receiver can be compared to using a digitizing tablet to collect data, but instead of pointing and clicking at features on a paper document, you are pointing and clicking on the real features to capture the information.
- Precision agriculture uses GPS extensively in conjunction with Variable Rate Technology (VRT). VRT relies on the use of a VRT controller box connected to a GPS and the pumping mechanism for a tank full of fertilizers/pesticides/seeds/water/etc. A digital polygon map (often derived from remotely sensed data) in the controller specifies a predefined amount to dispense for each polygonal region. As the tractor pulls the tank around the field the GPS logs the position that is compared to the map position in memory. The correct amount is then dispensed at that location. The aim of this process is to maximize yields without causing any environmental damage.
- GPS is often used in conjunction with airborne surveys. The aircraft, as well as carrying a camera or scanner, has on board one or more GPS receivers tied to an inertial navigation system. As each frame is exposed precise information is captured (or calculated in post processing) on the x, y, z and roll, pitch, yaw of the aircraft. Each image in the aerial survey block thus has initial exterior orientation parameters which therefore minimizes the need for control in a block triangulation process.

Figure 3-12 shows some additional uses for GPS coordinates.



Source: Leick 1990

Figure 3-12 Common Uses of GPS Data

Ordering Raster Data

Table 3-20 describes the different Landsat, SPOT, AVHRR, and DEM products that can be ordered. Information in this chart does not reflect all the products that are available, but only the most common types that can be imported into ERDAS IMAGINE.

Table 3-20 Common Raster Data Products

Data Type	Ground Covered	Pixel Size	# of Bands	Format	Available Geocoded
Landsat TM Full Scene	185 × 170 km	28.5 m	7	Fast (BSQ)	●
Landsat TM Quarter Scene	92.5 × 80 km	28.5 m	7	Fast (BSQ)	●
Landsat MSS Full Scene	185 × 170 km	79 × 56 m	4	BSQ, BIL	
SPOT	60 × 60 km	10 m and 20 m	1 - 3	BIL	●
NOAA AVHRR (LAC)	2700 × 2700 km	1.1 km	1-5	10-bit packed or unpacked	
NOAA AVHRR (GAC)	4000 × 4000 km	4 km	1-5	10-bit packed or unpacked	
USGS DEM 1:24,000	7.5' × 7.5'	30 m	1	ASCII	● (UTM)
USGS DEM 1:250,000	1° × 1°	3" × 3"	1	ASCII	

Raster and Vector Data Sources

Addresses to Contact

For more information about these and related products, contact the following agencies:

- Landsat MSS, TM, and ETM data:
Space Imaging
12076 Grant Street
Thornton, CO 80241 USA
Telephone (US): 800/425-2997
Telephone: 303/254-2000
Fax: 303/254-2215
Internet: www.spaceimage.com
- SPOT data:
SPOT Image Corporation
1897 Preston White Dr.
Reston, VA 22091-4368 USA
Telephone: 703-620-2200
Fax: 703-648-1813
Internet: www.spot.com
- NOAA AVHRR data:
Office of Satellite Operations
NOAA/National Environment Satellite, Data, and Information Service
World Weather Building, Room 100
Washington, DC 20233 USA
Internet: www.nesdis.noaa.gov
- AVHRR Dundee Format
NERC Satellite Station
University of Dundee
Dundee, Scotland DD1 4HN
Telephone: 44/1382-34-4409
Fax: 44/1382-202-575
Internet: www.sat.dundee.ac.uk
- Cartographic data including, maps, airphotos, space images, DEMs, planimetric data, and related information from federal, state, and private agencies:
National Mapping Division
U.S. Geological Survey, National Center
12201 Sunrise Valley Drive
Reston, VA 20192 USA
Telephone: 703/648-4000
Internet: mapping.usgs.gov
- ADRG data (available only to defense contractors):
NIMA (National Imagery and Mapping Agency)
ATTN: PMSC
Combat Support Center
Washington, DC 20315-0010 USA
- ADRI data (available only to defense contractors):
Rome Laboratory/IRRP
Image Products Branch
Griffiss AFB, NY 13440-5700 USA

- Landsat data:
Customer Services
U.S. Geological Survey
EROS Data Center
47914 252nd Street
Sioux Falls, SD 57198 USA
Telephone: 800/252-4547
Fax: 605/594-6589
Internet: edcwww.cr.usgs.gov/eros-home.html
- ERS-1 radar data:
RADARSAT International, Inc.
265 Carling Ave., Suite 204
Ottawa, Ontario
Canada K1S 2E1
Telephone: 613-238-6413
Fax: 613-238-5425
Internet: www.rsi.ca
- JERS-1 (Fuyo 1) radar data:
National Space Development Agency of Japan (NASDA)
Earth Observation Research Center
Roppongi - First Bldg., 1-9-9
Roppongi, Minato-ku
Tokyo 106-0032, Japan
Telephone: 81/3-3224-7040
Fax: 81/3-3224-7051
Internet: yyy.tksc.nasda.go.jp
- SIR-A, B, C radar data:
Jet Propulsion Laboratories
California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109-8099 USA
Telephone: 818/354-4321
Internet: www.jpl.nasa.gov
- RADARSAT data:
RADARSAT International, Inc.
265 Carling Ave., Suite 204
Ottawa, Ontario
Canada K1S 2E1
Telephone: 613-238-5424
Fax: 613-238-5425
Internet: www.rsi.ca
- Almaz radar data:
NPO Mashinostroenia
Scientific Engineering Center "Almaz"
33 Gagarin Street
Reutov, 143952, Russia
Telephone: 7/095-538-3018
Fax: 7/095-302-2001
E:mail: NPO@mashstroy.msk.su

Raster and Vector Data Sources

- U.S. Government RADARSAT sales:
Joel Porter
Lockheed Martin Astronautics
M/S: DC4001
12999 Deer Creek Canyon Rd.
Littleton, CO 80127
Telephone: 303-977-3233
Fax: 303-971-9827
E:mail: joel.porter@den.mmc.com

Raster Data from Other Software Vendors

ERDAS IMAGINE also enables you to import data created by other software vendors. This way, if another type of digital data system is currently in use, or if data is received from another system, it easily converts to the ERDAS IMAGINE file format for use in ERDAS IMAGINE.

Data from other vendors may come in that specific vendor's format, or in a standard format which can be used by several vendors. The Import and/or Direct Read function handles these raster data types from other software systems:

- ERDAS Ver. 7.X
- GRID and GRID Stacks
- JFIF (JPEG)
- MrSID
- SDTS
- Sun Raster
- TIFF and GeoTIFF



Other data types might be imported using the Generic Binary import option.

Vector to Raster Conversion

Vector data can also be a source of raster data by converting it to raster format.



Convert a vector layer to a raster layer, or vice versa, by using IMAGINE Vector.

ERDAS Ver. 7.X

The ERDAS Ver. 7.X series was the predecessor of ERDAS IMAGINE software. The two basic types of ERDAS Ver. 7.X data files are indicated by the file name extensions:

- .LAN—a multiband continuous image file (the name is derived from the Landsat satellite)
- .GIS—a single-band thematic data file in which pixels are divided into discrete categories (the name is derived from geographic information system)

.LAN and .GIS image files are stored in the same format. The image data are arranged in a BIL format and can be 4-bit, 8-bit, or 16-bit. The ERDAS Ver. 7.X file structure includes:

- a header record at the beginning of the file
- the data file values
- a statistics or trailer file



When you import a .GIS file, it becomes an image file with one thematic raster layer. When you import a .LAN file, each band becomes a continuous raster layer within the image file.

GRID and GRID Stacks

GRID is a raster geoprocessing program distributed by Environmental Systems Research Institute, Inc. (ESRI) in Redlands, California. GRID is designed to complement the vector data model system, ArcInfo is a well-known vector GIS that is also distributed by ESRI. The name GRID is taken from the raster data format of presenting information in a grid of cells.

The data format for GRID is a compressed tiled raster data structure. Like ArcInfo Coverages, a GRID is stored as a set of files in a directory, including files to keep the attributes of the GRID.

Each GRID represents a single layer of continuous or thematic imagery, but it is also possible to combine GRIDs files into a multilayer image. A GRID Stack (.stk) file names multiple GRIDs to be treated as a multilayer image. Starting with ArcInfo version 7.0, ESRI introduced the STK format, referred to in ERDAS software as GRID Stack 7.x, which contains multiple GRIDs. The GRID Stack 7.x format keeps attribute tables for the entire stack in a separate directory, in a manner similar to that of GRIDs and Coverages.

JFIF (JPEG)

JPEG is a set of compression techniques established by the Joint Photographic Experts Group (JPEG). The most commonly used form of JPEG involves Discrete Cosine Transformation (DCT), thresholding, followed by Huffman encoding. Since the output image is not exactly the same as the input image, this form of JPEG is considered to be lossy. JPEG can compresses monochrome imagery, but achieves compression ratios of 20:1 or higher with color (RGB) imagery, by taking advantage of the fact that the data being compressed is a visible image. The integrity of the source image is preserved by focussing its compression on aspects of the image that are less noticeable to the human eye. JPEG cannot be used on thematic imagery, due to the change in pixel values.

There is a lossless form of JPEG compression that uses DCT followed by nonlossy encoding, but it is not frequently used since it only yields an approximate compression ratio of 2:1. ERDAS IMAGINE only handles the lossy form of JPEG.

While JPEG compression is used by other file formats, including TIFF, the JPEG File Interchange Format (JFIF) is a standard file format used to store JPEG-compressed imagery.

The ISO JPEG committee is currently working on a new enhancement to the JPEG standard known as JPEG 2000, which will incorporate wavelet compression techniques and more flexibility in JPEG compression.

Raster and Vector Data Sources

MrSID

Multiresolution Seamless Image Database (MrSID, pronounced Mister Sid) is a wavelet transform-based compression algorithm designed by LizardTech, Inc. in Seattle, Washington (<http://www.lizardtech.com>). The novel developments in MrSID include a memory efficient implementation and automatic inclusion of pyramid layers in every data set, both of which make MrSID well-suited to provide efficient storage and retrieval of very large digital images.

The underlying wavelet-based compression methodology used in MrSID yields high compression ratios while satisfying stringent image quality requirements. The compression technique used in MrSID is lossy (i.e., the compression-decompression process does not reproduce the source data pixel-for-pixel). Lossy compression is not appropriate for thematic imagery, but is essential for large continuous images since it allows much higher compression ratios than lossless methods (e.g., the Lempel-Ziv-Welch, LZW, algorithm used in the GIF and TIFF image formats). At standard compression ratios, MrSID encoded imagery is visually lossless. On typical remotely sensed imagery, lossless methods provide compression ratios of perhaps 2:1, whereas MrSID provides excellent image quality at compression ratios of 30:1 or more.

SDTS

The Spatial Data Transfer Standard (SDTS) was developed by the USGS to promote and facilitate the transfer of georeferenced data and its associated metadata between dissimilar computer systems without loss of fidelity. To achieve these goals, SDTS uses a flexible, self-describing method of encoding data, which has enough structure to permit interoperability.

For metadata, SDTS requires a number of statements regarding data accuracy. In addition to the standard metadata, the producer may supply detailed attribute data correlated to any image feature.

SDTS Profiles

The SDTS standard is organized into profiles. Profiles identify a restricted subset of the standard needed to solve a certain problem domain. Two subsets of interest to ERDAS IMAGINE users are:

- Topological Vector Profile (TVP), which covers attributed vector data. This is imported via the SDTS (Vector) title.
- SDTS Raster Profile and Extensions (SRPE), which covers gridded raster data. This is imported as SDTS Raster.

For more information on SDTS, consult the SDTS web page at <http://mcmcweb.er.usgs.gov/sdts>.

SUN Raster

A SUN Raster file is an image captured from a monitor display. In addition to GIS, SUN Raster files can be used in desktop publishing applications or any application where a screen capture would be useful.

There are two basic ways to create a SUN Raster file on a SUN workstation:

- use the OpenWindows Snapshot application
- use the UNIX screendump command

Both methods read the contents of a frame buffer and write the display data to a user-specified file. Depending on the display hardware and options chosen, screendump can create any of the file types listed in Table 3-21.

Table 3-21 File Types Created by Screendump

File Type	Available Compression
1-bit black and white	None, RLE (run-length encoded)
8-bit color paletted (256 colors)	None, RLE
24-bit RGB true color	None, RLE
32-bit RGB true color	None, RLE

The data are stored in BIP format.

TIFF

TIFF was developed by Aldus Corp. (Seattle, Washington) in 1986 in conjunction with major scanner vendors who needed an easily portable file format for raster image data. Today, the TIFF format is a widely supported format used in video, fax transmission, medical imaging, satellite imaging, document storage and retrieval, and desktop publishing applications. In addition, the GeoTIFF extensions permit TIFF files to be geocoded.

The TIFF format’s main appeal is its flexibility. It handles black and white line images, as well as gray scale and color images, which can be easily transported between different operating systems and computers.

TIFF File Formats

TIFF’s great flexibility can also cause occasional problems in compatibility. This is because TIFF is really a family of file formats that is comprised of a variety of elements within the format.

Table 3-22 shows key Baseline TIFF format elements and the values for those elements supported by ERDAS IMAGINE.



Any TIFF file that contains an unsupported value for one of these elements may not be compatible with ERDAS IMAGINE.

Table 3-22 The Most Common TIFF Format Elements

Byte Order	Intel (LSB/MSB)
	Motorola (MSB/LSB)
Image Type	Black and white
	Gray scale
	Inverted gray scale
	Color palette
	RGB (3-band)
Configuration	BIP
	BSQ
Bits Per Plane^a	1 ^b , 2 ^b , 4, 8, 16 ^c , 32 ^c , 64 ^c
Compression^d	None
	CCITT G3 (B&W only)
	CCITT G4 (B&W only)
	Packbits
	LZW ^e
	LZW with horizontal differencing ^e

^a All bands must contain the same number of bits (i.e., 4,4,4 or 8,8,8). Multiband data with bit depths differing per band cannot be imported into ERDAS IMAGINE.

^b Must be imported and exported as 4-bit data.

^c Direct read/write only.

^d Compression supported on import and direct read/write only.

^e LZW is governed by patents and is not supported by the basic version of ERDAS IMAGINE.

GeoTIFF

According to the GeoTIFF Format Specification, Revision 1.0, "The GeoTIFF spec defines a set of TIFF tags provided to describe all 'Cartographic' information associated with TIFF imagery that originates from satellite imaging systems, scanned aerial photography, scanned maps, digital elevation models, or as a result of geographic analysis" (Ritter and Ruth 1999).

The GeoTIFF format separates cartographic information into two parts: georeferencing and geocoding.

Georeferencing

Georeferencing is the process of linking the raster space of an image to a model space (i.e., a map system). Raster space defines how the coordinate system grid lines are placed relative to the centers of the pixels of the image. In ERDA IMAGINE, the grid lines of the coordinate system always intersect at the center of a pixel. GeoTIFF allows the raster space to be defined either as having grid lines intersecting at the centers of the pixels (PixelIsPoint) or as having grid lines intersecting at the upper left corner of the pixels (PixelIsArea). ERDAS IMAGINE converts the georeferencing values for PixelIsArea images so that they conform to its raster space definition.

GeoTIFF allows georeferencing via a scale and an offset, a full affine transformation, or a set of tie points. ERDAS IMAGINE currently ignores GeoTIFF georeferencing in the form of multiple tie points.

Geocoding

Geocoding is the process of linking coordinates in model space to the Earth's surface. Geocoding allows for the specification of projection, datum, ellipsoid, etc. ERDAS IMAGINE can interpret GeoTIFF geocoding so that latitude and longitude of the GeoTIFF images map coordinates can be determined. This interpretation also allows the GeoTIFF image to be reprojected.

In GeoTIFF, the units of the map coordinates are obtained from the geocoding, not from the georeferencing. In addition, GeoTIFF defines a set of standard projected coordinate systems. The use of a standard projected coordinate system in GeoTIFF constrains the units that can be used with that standard system. Therefore, if the units used with a projection in ERDAS IMAGINE are not equal to the implied units of an equivalent GeoTIFF geocoding, ERDAS IMAGINE transforms the georeferencing to conform to the implied units so that the standard projected coordinate system code can be used. The alternative (preserving the georeferencing as is and producing a nonstandard projected coordinate system) is regarded as less interoperable.

**Vector Data from
Other Software
Vendors**

It is possible to directly import several common vector formats into ERDAS IMAGINE. These files become vector layers when imported. These data can then be used for the analyses and, in most cases, exported back to their original format (if desired).

Although data can be converted from one type to another by importing a file into ERDAS IMAGINE and then exporting the ERDAS IMAGINE file into another format, the import and export routines were designed to work together. For example, if you have information in AutoCAD that you would like to use in the GIS, you can import a Drawing Interchange File (DXF) into ERDAS IMAGINE, do the analysis, and then export the data back to DXF format.

In most cases, attribute data are also imported into ERDAS IMAGINE. Each of the following sections lists the types of attribute data that are imported.



Use Import/Export to import vector data from other software vendors into ERDAS IMAGINE vector layers. These routines are based on ArcInfo data conversion routines.

Raster and Vector Data Sources



See "CHAPTER 2: Vector Layers" for more information on ERDAS IMAGINE vector layers. See "CHAPTER 11: Geographic Information Systems" for more information about using vector data in a GIS.

ARCGEN

ARCGEN files are ASCII files created with the ArcInfo UNGENERATE command. The import ARCGEN program is used to import features to a new layer. Topology is not created or maintained, therefore the coverage must be built or cleaned after it is imported into ERDAS IMAGINE.



ARCGEN files must be properly prepared before they are imported into ERDAS IMAGINE. If there is a syntax error in the data file, the import process may not work. If this happens, you must kill the process, correct the data file, and then try importing again.



See the ArcInfo documentation for more information about these files.

AutoCAD (DXF)

AutoCAD is a vector software package distributed by Autodesk, Inc. (Sausalito, California). AutoCAD is a computer-aided design program that enables the user to draw two- and three-dimensional models. This software is frequently used in architecture, engineering, urban planning, and many other applications.

AutoCAD DXF is the standard interchange format used by most CAD systems. The AutoCAD program DXFOUT creates a DXF file that can be converted to an ERDAS IMAGINE vector layer. AutoCAD files can also be output to IGES format using the AutoCAD program IGESOUT.



See "IGES" on page 108 for more information about IGES files.

DXF files can be converted in the ASCII or binary format. The binary format is an optional format for AutoCAD Releases 10 and 11. It is structured just like the ASCII format, only the data are in binary format.

DXF files are composed of a series of related layers. Each layer contains one or more drawing elements or entities. An entity is a drawing element that can be placed into an AutoCAD drawing with a single command. When converted to an ERDAS IMAGINE vector layer, each entity becomes a single feature. Table 3-23 describes how various DXF entities are converted to ERDAS IMAGINE.

Table 3-23 Conversion of DXF Entries

DXF Entity	ERDAS IMAGINE Feature	Comments
Line 3DLine	Line	These entities become two point lines. The initial Z value of 3D entities is stored.
Trace Solid 3DFace	Line	These entities become four or five point lines. The initial Z value of 3D entities is stored.
Circle Arc	Line	These entities form lines. Circles are composed of 361 points—one vertex for each degree. The first and last point is at the same location.
Polyline	Line	These entities can be grouped to form a single line having many vertices.
Point Shape	Point	These entities become point features in a layer.

The ERDAS IMAGINE import process also imports line and point attribute data (if they exist) and creates an INFO directory with the appropriate ACODE (arc attributes) and XCODE (point attributes) files. If an imported DXF file is exported back to DXF format, this information is also exported.



Refer to an AutoCAD manual for more information about the format of DXF files.

DLG

DLGs are furnished by the U.S. Geological Survey and provide planimetric base map information, such as transportation, hydrography, contours, and public land survey boundaries. DLG files are available for the following USGS map series:

- 7.5- and 15-minute topographic quadrangles
- 1:100,000-scale quadrangles
- 1:2,000,000-scale national atlas maps

DLGs are topological files that contain nodes, lines, and areas (similar to the points, lines, and polygons in an ERDAS IMAGINE vector layer). DLGs also store attribute information in the form of major and minor code pairs. Code pairs are encoded in two integer fields, each containing six digits. The major code describes the class of the feature (road, stream, etc.) and the minor code stores more specific information about the feature.

Raster and Vector Data Sources

DLGs can be imported in standard format (144 bytes per record) and optional format (80 bytes per record). You can export to DLG-3 optional format. Most DLGs are in the Universal Transverse Mercator (UTM) map projection. However, the 1:2,000,000 scale series is in geographic coordinates.

The ERDAS IMAGINE import process also imports point, line, and polygon attribute data (if they exist) and creates an INFO directory with the appropriate ACODE, PCODE (polygon attributes), and XCODE files. If an imported DLG file is exported back to DLG format, this information is also exported.



To maintain the topology of a vector layer created from a DLG file, you must Build or Clean it. See "CHAPTER 11: Geographic Information Systems" on page 385 for information on this process.

ETAK

ETAK's MapBase is an ASCII digital street centerline map product available from ETAK, Inc. (Menlo Park, California). ETAK files are similar in content to the Dual Independent Map Encoding (DIME) format used by the U.S. Census Bureau. Each record represents a single linear feature with address and political, census, and ZIP code boundary information. ETAK has also included road class designations and, in some areas, major landmark features.

There are four possible types of ETAK features:

- DIME or D types—if the feature type is D, a line is created along with a corresponding ACODE (arc attribute) record. The coordinates are stored in Lat/Lon decimal degrees.
- Alternate address or A types—each record contains an alternate address record for a line. These records are written to the attribute file, and are useful for building address coverages.
- Shape features or S types—shape records are used to add vertices to the lines. The coordinates for these features are in Lat/Lon decimal degrees.
- Landmark or L types—if the feature type is L and you opt to output a landmark layer, then a point feature is created along with an associated PCODE record.



ERDAS IMAGINE vector data cannot be exported to ETAK format.

IGES

IGES files are often used to transfer CAD data between systems. IGES Version 3.0 format, published by the U.S. Department of Commerce, is in uncompressed ASCII format only.

IGES files can be produced in AutoCAD using the IGESOUT command. The following IGES entities can be converted:

Table 3-24 Conversion of IGES Entities

IGES Entity	ERDAS IMAGINE Feature
IGES Entity 100 (Circular Arc Entities)	Lines
IGES Entity 106 (Copious Data Entities)	Lines
IGES Entity 106 (Line Entities)	Lines
IGES Entity 116 (Point Entities)	Points

The ERDAS IMAGINE import process also imports line and point attribute data (if they exist) and creates an INFO directory with the appropriate ACODE and XCODE files. If an imported IGES file is exported back to IGES format, this information is also exported.

TIGER

TIGER files are line network products of the U.S. Census Bureau. The Census Bureau is using the TIGER system to create and maintain a digital cartographic database that covers the United States, Puerto Rico, Guam, the Virgin Islands, American Samoa, and the Trust Territories of the Pacific.

TIGER/Line is the line network product of the TIGER system. The cartographic base is taken from Geographic Base File/Dual Independent Map Encoding (GBF/DIME), where available, and from the USGS 1:100,000-scale national map series, SPOT imagery, and a variety of other sources in all other areas, in order to have continuous coverage for the entire United States. In addition to line segments, TIGER files contain census geographic codes and, in metropolitan areas, address ranges for the left and right sides of each segment. TIGER files are available in ASCII format on both CD-ROM and tape media. All released versions after April 1989 are supported.

There is a great deal of attribute information provided with TIGER/Line files. Line and point attribute information can be converted into ERDAS IMAGINE format. The ERDAS IMAGINE import process creates an INFO directory with the appropriate ACODE and XCODE files. If an imported TIGER file is exported back to TIGER format, this information is also exported.

TIGER attributes include the following:

- Version numbers—TIGER/Line file version number.
- Permanent record numbers—each line segment is assigned a permanent record number that is maintained throughout all versions of TIGER/Line files.
- Source codes—each line and landmark point feature is assigned a code to specify the original source.
- Census feature class codes—line segments representing physical features are coded based on the USGS classification codes in DLG-3 files.
- Street attributes—includes street address information for selected urban areas.

Raster and Vector Data Sources

- Legal and statistical area attributes—legal areas include states, counties, townships, towns, incorporated cities, Indian reservations, and national parks. Statistical areas are areas used during the census-taking, where legal areas are not adequate for reporting statistics.
- Political boundaries—the election precincts or voting districts may contain a variety of areas, including wards, legislative districts, and election districts.
- Landmarks—landmark area and point features include schools, military installations, airports, hospitals, mountain peaks, campgrounds, rivers, and lakes.



TIGER files for major metropolitan areas outside of the United States (e.g., Puerto Rico, Guam) do not have address ranges.

Disk Space Requirements

TIGER/Line files are partitioned into counties ranging in size from less than a megabyte to almost 120 megabytes. The average size is approximately 10 megabytes. To determine the amount of disk space required to convert a set of TIGER/Line files, use this rule: the size of the converted layers is approximately the same size as the files used in the conversion. The amount of additional scratch space needed depends on the largest file and whether it needs to be sorted. The amount usually required is about double the size of the file being sorted.



The information presented in this section, “Vector Data from Other Software Vendors”, was obtained from the Data Conversion and the 6.0 ARC Command References manuals, both published by ESRI, Inc., 1992.

CHAPTER 4

Image Display

Introduction

This section defines some important terms that are relevant to image display. Most of the terminology and definitions used in this chapter are based on the X Window System (Massachusetts Institute of Technology) terminology. This may differ from other systems, such as Microsoft Windows NT.

A seat is a combination of an X-server and a host workstation.

- A host workstation consists of a CPU, keyboard, mouse, and a display.
- A display may consist of multiple screens. These screens work together, making it possible to move the mouse from one screen to the next.
- The display hardware contains the memory that is used to produce the image. This hardware determines which types of displays are available (e.g., true color or pseudo color) and the pixel depth (e.g., 8-bit or 24-bit).

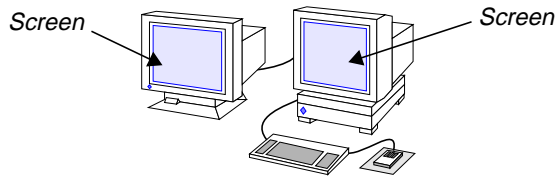


Figure 4-1 Example of One Seat with One Display and Two Screens

Display Memory Size

The size of memory varies for different displays. It is expressed in terms of:

- display resolution, which is expressed as the horizontal and vertical dimensions of memory—the number of pixels that can be viewed on the display screen. Some typical display resolutions are 1152×900 , 1280×1024 , and 1024×780 . For the PC, typical resolutions are 640×480 , 800×600 , 1024×768 , and 1280×1024 , and
- the number of bits for each pixel or pixel depth, as explained below.

Bits for Image Plane

A bit is a binary digit, meaning a number that can have two possible values—0 and 1, or off and on. A set of bits, however, can have many more values, depending upon the number of bits used. The number of values that can be expressed by a set of bits is 2 to the power of the number of bits used. For example, the number of values that can be expressed by 3 bits is $2^3 = 8$.

Image Display

Displays are referred to in terms of a number of bits, such as 8-bit or 24-bit. These bits are used to determine the number of possible brightness values. For example, in a 24-bit display, 24 bits per pixel breaks down to eight bits for each of the three color guns per pixel. The number of possible values that can be expressed by eight bits is 2^8 , or 256. Therefore, on a 24-bit display, each color gun of a pixel can have any one of 256 possible brightness values, expressed by the range of values 0 to 255.

The combination of the three color guns, each with 256 possible brightness values, yields 256^3 , (or 2^{24} , for the 24-bit image display), or 16,777,216 possible colors for each pixel on a 24-bit display. If the display being used is not 24-bit, the example above calculates the number of possible brightness values and colors that can be displayed.

Pixel

The term pixel is abbreviated from picture element. As an element, a pixel is the smallest part of a digital picture (image). Raster image data are divided by a grid, in which each cell of the grid is represented by a pixel. A pixel is also called a grid cell.

Pixel is a broad term that is used for both:

- the data file value(s) for one data unit in an image (file pixels), or
- one grid location on a display or printout (display pixels).

Usually, one pixel in a file corresponds to one pixel in a display or printout. However, an image can be magnified or reduced so that one file pixel no longer corresponds to one pixel in the display or printout. For example, if an image is displayed with a magnification factor of 2, then one file pixel takes up 4 (2×2) grid cells on the display screen.

To display an image, a file pixel that consists of one or more numbers must be transformed into a display pixel with properties that can be seen, such as brightness and color. Whereas the file pixel has values that are relevant to data (such as wavelength of reflected light), the displayed pixel must have a particular color or gray level that represents these data file values.

Colors

Human perception of color comes from the relative amounts of red, green, and blue light that are measured by the cones (sensors) in the eye. Red, green, and blue light can be added together to produce a wide variety of colors—a wider variety than can be formed from the combinations of any three other colors. Red, green, and blue are therefore the additive primary colors.

A nearly infinite number of shades can be produced when red, green, and blue light are combined. On a display, different colors (combinations of red, green, and blue) allow you to perceive changes across an image. Color displays that are available today yield 2^{24} , or 16,777,216 colors. Each color has a possible 256 different values (2^8).

Color Guns

On a display, color guns direct electron beams that fall on red, green, and blue phosphors. The phosphors glow at certain frequencies to produce different colors. Color monitors are often called RGB monitors, referring to the primary colors.

The red, green, and blue phosphors on the picture tube appear as tiny colored dots on the display screen. The human eye integrates these dots together, and combinations of red, green, and blue are perceived. Each pixel is represented by an equal number of red, green, and blue phosphors.

Brightness Values

Brightness values (or intensity values) are the quantities of each primary color to be output to each displayed pixel. When an image is displayed, brightness values are calculated for all three color guns, for every pixel.

All of the colors that can be output to a display can be expressed with three brightness values—one for each color gun.

Colormap and Colorcells

A color on the screen is created by a combination of red, green, and blue values, where each of these components is represented as an 8-bit value. Therefore, 24 bits are needed to represent a color. Since many systems have only an 8-bit display, a colormap is used to translate the 8-bit value into a color. A colormap is an ordered set of colorcells, which is used to perform a function on a set of input values. To display or print an image, the colormap translates data file values in memory into brightness values for each color gun. Colormaps are not limited to 8-bit displays.

Colormap vs. Lookup Table

The colormap is a function of the display hardware, whereas a lookup table is a function of ERDAS IMAGINE. When a contrast adjustment is performed on an image in ERDAS IMAGINE, lookup tables are used. However, if the auto-update function is being used to view the adjustments in near real-time, then the colormap is being used to map the image through the lookup table. This process allows the colors on the screen to be updated in near real-time. This chapter explains how the colormap is used to display imagery.

Colorcells

There is a colorcell in the colormap for each data file value. The red, green, and blue values assigned to the colorcell control the brightness of the color guns for the displayed pixel (Nye 1990). The number of colorcells in a colormap is determined by the number of bits in the display (e.g., 8-bit, 24-bit).

For example, if a pixel with a data file value of 40 was assigned a display value (colorcell value) of 24, then this pixel uses the brightness values for the 24th colorcell in the colormap. In the colormap below (Table 4-1), this pixel is displayed as blue.

Table 4-1 Colorcell Example

Colorcell Index	Red	Green	Blue
1	255	0	0
2	0	170	90
3	0	0	255
24	0	0	255

Image Display

The colormap is controlled by the X Windows system. There are 256 colorcells in a colormap with an 8-bit display. This means that 256 colors can be displayed simultaneously on the display. With a 24-bit display, there are 256 colorcells for each color: red, green, and blue. This offers $256 \times 256 \times 256$, or 16,777,216 different colors.

When an application requests a color, the server specifies which colorcell contains that color and returns the color. Colorcells can be read-only or read/write.

Read-only Colorcells

The color assigned to a read-only colorcell can be shared by other application windows, but it cannot be changed once it is set. To change the color of a pixel on the display, it would not be possible to change the color for the corresponding colorcell. Instead, the pixel value would have to be changed and the image redisplayed. For this reason, it is not possible to use auto-update operations in ERDAS IMAGINE with read-only colorcells.

Read/Write Colorcells

The color assigned to a read/write colorcell can be changed, but it cannot be shared by other application windows. An application can easily change the color of displayed pixels by changing the color for the colorcell that corresponds to the pixel value. This allows applications to use auto update operations. However, this colorcell cannot be shared by other application windows, and all of the colorcells in the colormap could quickly be utilized.

Changeable Colormaps

Some colormaps can have both read-only and read/write colorcells. This type of colormap allows applications to utilize the type of colorcell that would be most preferred.

Display Types

The possible range of different colors is determined by the display type. ERDAS IMAGINE supports the following types of displays:

- 8-bit PseudoColor
- 15-bit HiColor (for Windows NT)
- 24-bit DirectColor
- 24-bit TrueColor

The above display types are explained in more detail below.



A display may offer more than one visual type and pixel depth. See the ERDAS IMAGINE V8.4 Installation Guide for more information on specific display hardware.

32-bit Displays

A 32-bit display is a combination of an 8-bit PseudoColor and 24-bit DirectColor, or TrueColor display. Whether or not it is DirectColor or TrueColor depends on the display hardware.

8-bit PseudoColor

An 8-bit PseudoColor display has a colormap with 256 colorcells. Each cell has a red, green, and blue brightness value, giving 256 combinations of red, green, and blue. The data file value for the pixel is transformed into a colorcell value. The brightness values for the colorcell that is specified by this colorcell value are used to define the color to be displayed.

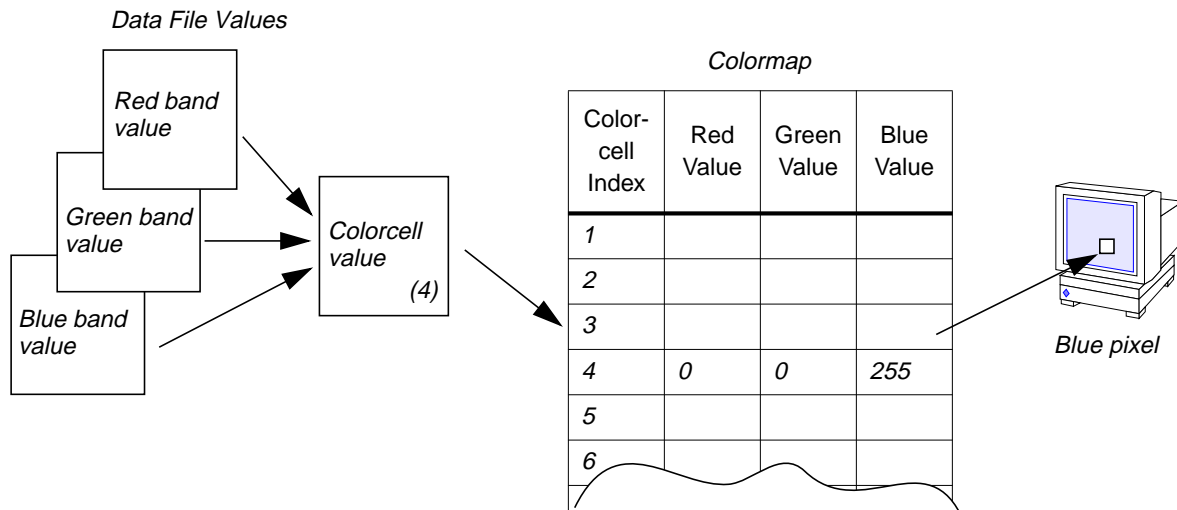


Figure 4-2 Transforming Data File Values to a Colorcell Value

In Figure 4-2, data file values for a pixel of three continuous raster layers (bands) is transformed to a colorcell value. Since the colorcell value is four, the pixel is displayed with the brightness values of the fourth colorcell (blue).

This display grants a small number of colors to ERDAS IMAGINE. It works well with thematic raster layers containing less than 200 colors and with gray scale continuous raster layers. For image files with three continuous raster layers (bands), the colors are severely limited because, under ideal conditions, 256 colors are available on an 8-bit display, while 8-bit, 3-band image files can contain over 16,000,000 different colors.

Auto Update

An 8-bit PseudoColor display has read-only and read/write colorcells, allowing ERDAS IMAGINE to perform near real-time color modifications using Auto Update and Auto Apply options.

24-bit DirectColor

A 24-bit DirectColor display enables you to view up to three bands of data at one time, creating displayed pixels that represent the relationships between the bands by their colors. Since this is a 24-bit display, it offers up to 256 shades of red, 256 shades of green, and 256 shades of blue, which is approximately 16 million different colors (256^3). The data file values for each band are transformed into colorcell values. The colorcell that is specified by these values is used to define the color to be displayed.

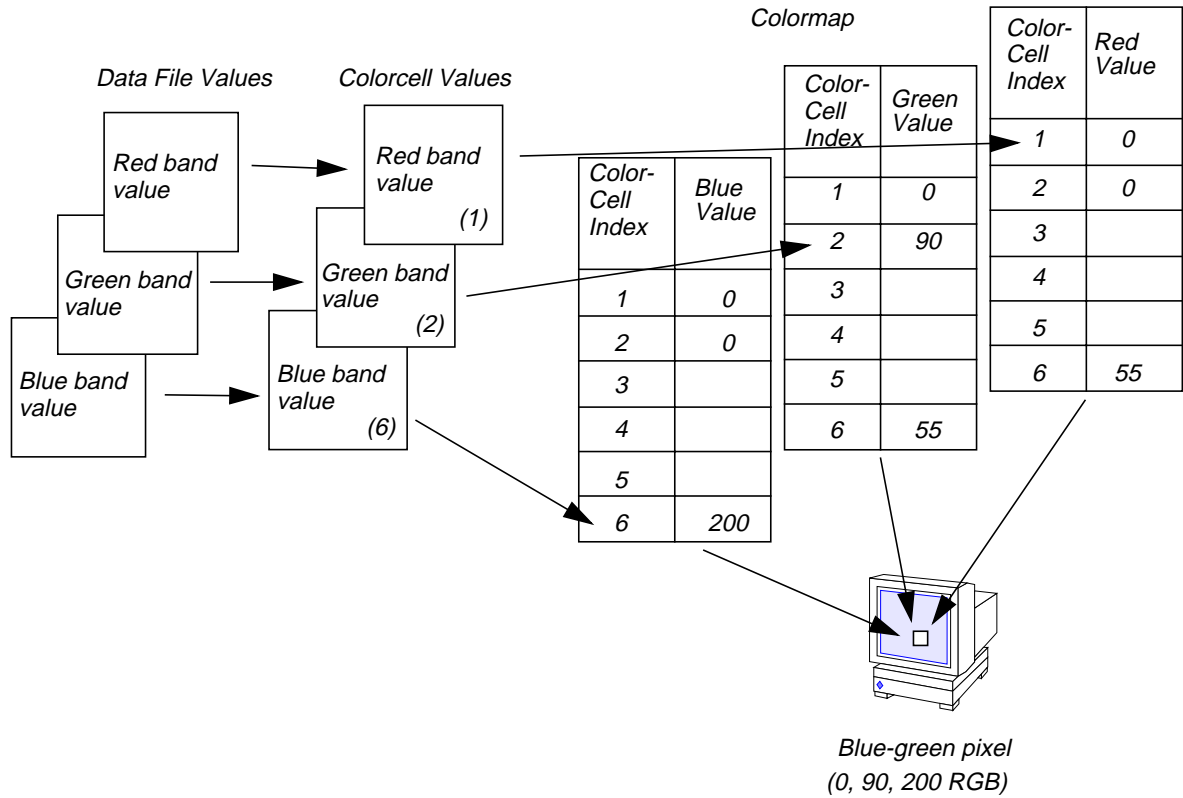


Figure 4-3 Transforming Data File Values to a Colorcell Value

In Figure 4-3, data file values for a pixel of three continuous raster layers (bands) are transformed to separate colorcell values for each band. Since the colorcell value is 1 for the red band, 2 for the green band, and 6 for the blue band, the RGB brightness values are 0, 90, 200. This displays the pixel as a blue-green color.

This type of display grants a very large number of colors to ERDAS IMAGINE and it works well with all types of data.

Auto Update

A 24-bit DirectColor display has read-only and read/write colorcells, allowing ERDAS IMAGINE to perform real-time color modifications using the Auto Update and Auto Apply options.

24-bit TrueColor

A 24-bit TrueColor display enables you to view up to three continuous raster layers (bands) of data at one time, creating displayed pixels that represent the relationships between the bands by their colors. The data file values for the pixels are transformed into screen values and the colors are based on these values. Therefore, the color for the pixel is calculated without querying the server and the colormap. The colormap for a 24-bit TrueColor display is not available for ERDAS IMAGINE applications. Once a color is assigned to a screen value, it cannot be changed, but the color can be shared by other applications.

The screen values are used as the brightness values for the red, green, and blue color guns. Since this is a 24-bit display, it offers 256 shades of red, 256 shades of green, and 256 shades of blue, which is approximately 16 million different colors (256^3).

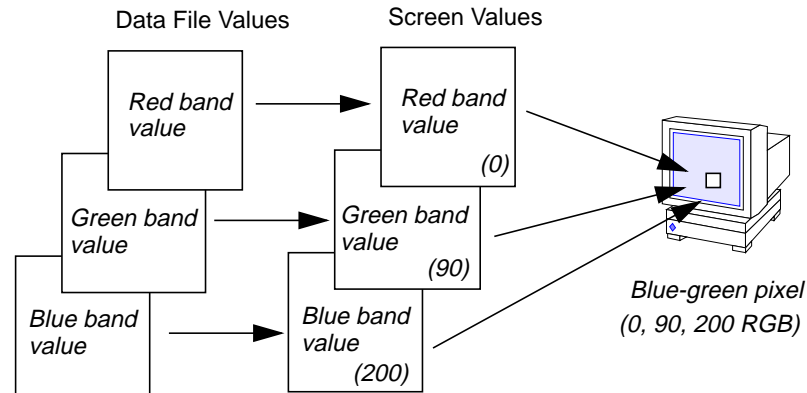


Figure 4-4 Transforming Data File Values to Screen Values

In Figure 4-4, data file values for a pixel of three continuous raster layers (bands) are transformed to separate screen values for each band. Since the screen value is 0 for the red band, 90 for the green band, and 200 for the blue band, the RGB brightness values are 0, 90, and 200. This displays the pixel as a blue-green color.

Auto Update

The 24-bit TrueColor display does not use the colormap in ERDAS IMAGINE, and thus does not provide ERDAS IMAGINE with any real-time color changing capability. Each time a color is changed, the screen values must be calculated and the image must be redrawn.

Color Quality

The 24-bit TrueColor visual provides the best color quality possible with standard equipment. There is no color degradation under any circumstances with this display.

PC Displays

ERDAS IMAGINE for Microsoft Windows NT supports the following visual type and pixel depths:

- 8-bit PseudoColor
- 15-bit HiColor
- 24-bit TrueColor

8-bit PseudoColor

An 8-bit PseudoColor display for the PC uses the same type of colormap as the X Windows 8-bit PseudoColor display, except that each colorcell has a range of 0 to 63 on most video display adapters, instead of 0 to 255. Therefore, each colorcell has a red, green, and blue brightness value, giving 64 different combinations of red, green, and blue. The colormap, however, is the same as the X Windows 8-bit PseudoColor display. It has 256 colorcells allowing 256 different colors to be displayed simultaneously.

15-bit HiColor

A 15-bit HiColor display for the PC assigns colors the same way as the X Windows 24-bit TrueColor display, except that it offers 32 shades of red, 32 shades of green, and 32 shades of blue, for a total of 32,768 possible color combinations. Some video display adapters allocate 6 bits to the green color gun, allowing 64,000 colors. These adapters use a 16-bit color scheme.

24-bit TrueColor

A 24-bit TrueColor display for the PC assigns colors the same way as the X Windows 24-bit TrueColor display.

Displaying Raster Layers

Image files (.img) are raster files in the ERDAS IMAGINE format. There are two types of raster layers:

- continuous
- thematic

Thematic raster layers require a different display process than continuous raster layers. This section explains how each raster layer type is displayed.

Continuous Raster Layers

An image file (.img) can contain several continuous raster layers; therefore, each pixel can have multiple data file values. When displaying an image file with continuous raster layers, it is possible to assign which layers (bands) are to be displayed with each of the three color guns. The data file values in each layer are input to the assigned color gun. The most useful color assignments are those that allow for an easy interpretation of the displayed image. For example:

- A natural-color image approximates the colors that would appear to a human observer of the scene.
- A color-infrared image shows the scene as it would appear on color-infrared film, which is familiar to many analysts.

Band assignments are often expressed in R,G,B order. For example, the assignment 4,2,1 means that band 4 is assigned to red, band 2 to green, and band 1 to blue. Below are some widely used band to color gun assignments (Faust 1989):

- Landsat TM—natural color: 3,2,1
This is natural color because band 3 is red and is assigned to the red color gun, band 2 is green and is assigned to the green color gun, and band 1 is blue and is assigned to the blue color gun.

- Landsat TM—color-infrared: 4,3,2
This is infrared because band 4 = infrared.
- SPOT Multispectral—color-infrared: 3,2,1
This is infrared because band 3 = infrared.

Contrast Table

When an image is displayed, ERDAS IMAGINE automatically creates a contrast table for continuous raster layers. The red, green, and blue brightness values for each band are stored in this table.

Since the data file values in continuous raster layers are quantitative and related, the brightness values in the colormap are also quantitative and related. The screen pixels represent the relationships between the values of the file pixels by their colors. For example, a screen pixel that is bright red has a high brightness value in the red color gun, and a high data file value in the layer assigned to red, relative to other data file values in that layer.

The brightness values often differ from the data file values, but they usually remain in the same order of lowest to highest. Some meaningful relationships between the values are usually maintained.

Contrast Stretch

Different displays have different ranges of possible brightness values. The range of most displays is 0 to 255 for each color gun.

Since the data file values in a continuous raster layer often represent raw data (such as elevation or an amount of reflected light), the range of data file values is often not the same as the range of brightness values of the display. Therefore, a contrast stretch is usually performed, which stretches the range of the values to fit the range of the display.

For example, Figure 4-5 shows a layer that has data file values from 30 to 40. When these values are used as brightness values, the contrast of the displayed image is poor. A contrast stretch simply stretches the range between the lower and higher data file values, so that the contrast of the displayed image is higher—that is, lower data file values are displayed with the lowest brightness values, and higher data file values are displayed with the highest brightness values.

The colormap stretches the range of colorcell values from 30 to 40, to the range 0 to 255. Because the output values are incremented at regular intervals, this stretch is a linear contrast stretch. (The numbers in Figure 4-5 are approximations and do not show an exact linear relationship.)

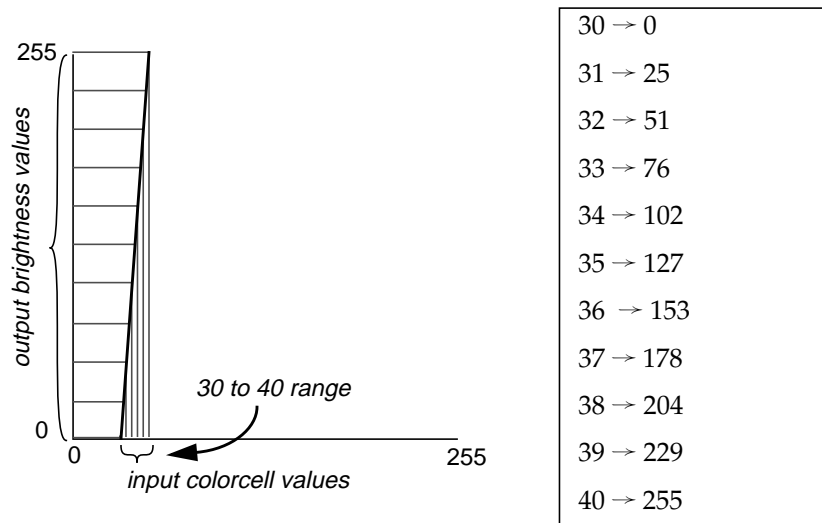


Figure 4-5 Contrast Stretch and Colorcell Values



See "CHAPTER 5: Enhancement" for more information about contrast stretching. Contrast stretching is performed the same way for display purposes as it is for permanent image enhancement.



A two standard deviation linear contrast stretch is applied to stretch pixel values of all image files from 0 to 255 before they are displayed in the Viewer, unless a saved contrast stretch exists (the file is not changed). This often improves the initial appearance of the data in the Viewer.

Statistics Files

To perform a contrast stretch, certain statistics are necessary, such as the mean and the standard deviation of the data file values in each layer.



Use the Image Information utility to create and view statistics for a raster layer.

Usually, not all of the data file values are used in the contrast stretch calculations. The minimum and maximum data file values of each band are often too extreme to produce good results. When the minimum and maximum are extreme in relation to the rest of the data, then the majority of data file values are not stretched across a very wide range, and the displayed image has low contrast.

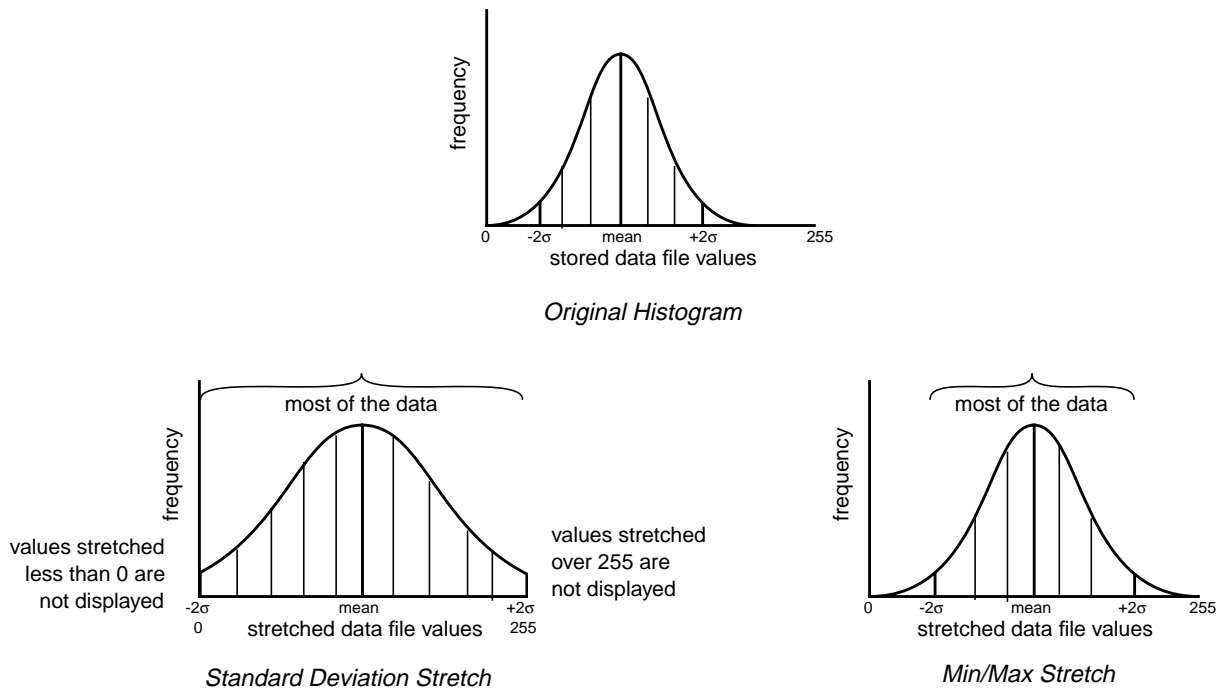


Figure 4-6 Stretching by Min/Max vs. Standard Deviation

The mean and standard deviation of the data file values for each band are used to locate the majority of the data file values. The number of standard deviations above and below the mean can be entered, which determines the range of data used in the stretch.



See "Appendix A: Math Topics" for more information on mean and standard deviation.



Use the Contrast Tools dialog, which is accessible from the Lookup Table Modification dialog, to enter the number of standard deviations to be used in the contrast stretch.

24-bit DirectColor and TrueColor Displays

Figure 4-7 illustrates the general process of displaying three continuous raster layers on a 24-bit DirectColor display. The process is similar on a TrueColor display except that the colormap is not used.

Image Display

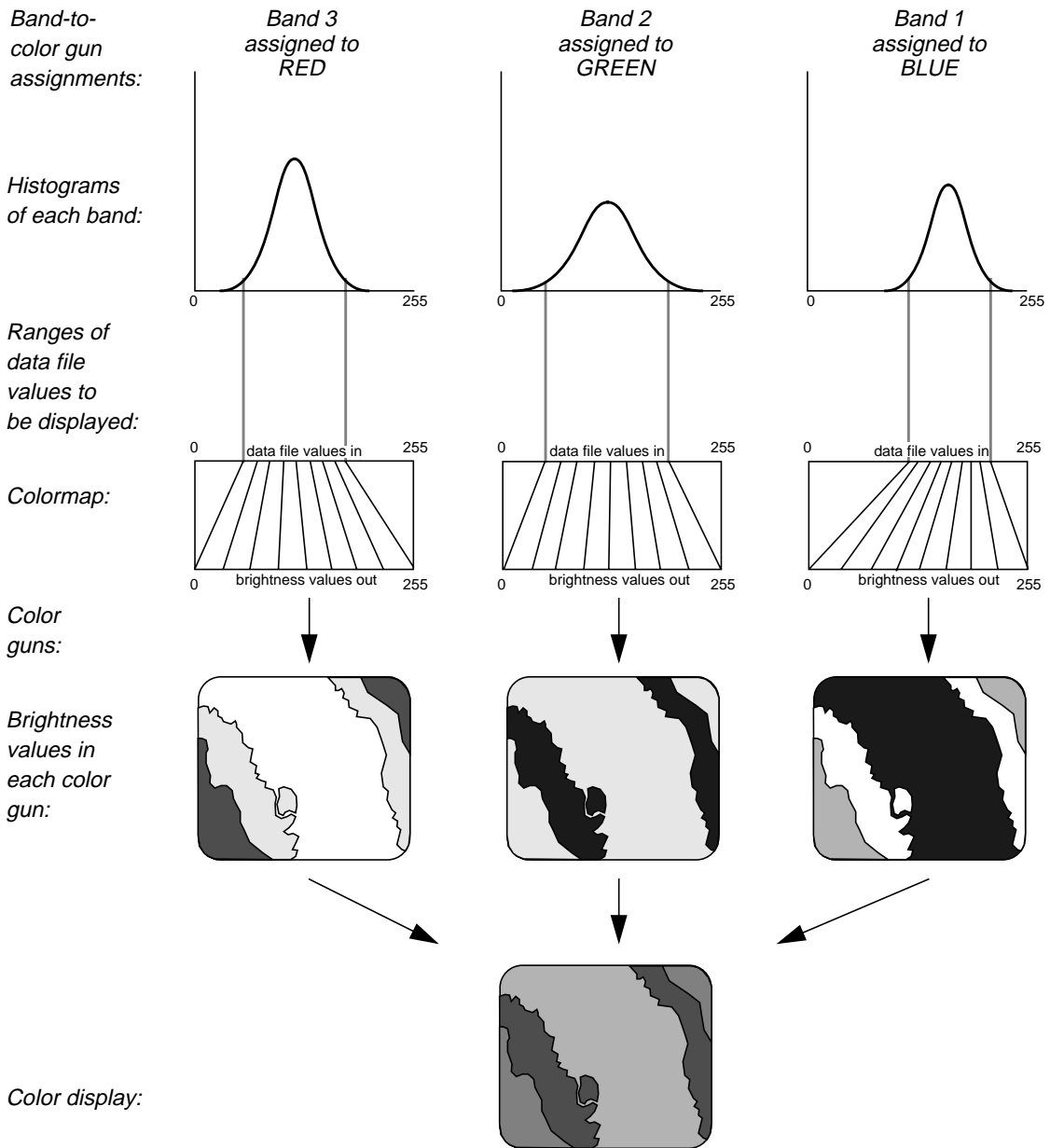


Figure 4-7 Continuous Raster Layer Display Process

8-bit PseudoColor Display

When displaying continuous raster layers on an 8-bit PseudoColor display, the data file values from the red, green, and blue bands are combined and transformed to a colorcell value in the colormap. This colorcell then provides the red, green, and blue brightness values. Since there are only 256 colors available, a continuous raster layer looks different when it is displayed in an 8-bit display than a 24-bit display that offers 16 million different colors. However, the Viewer performs dithering with the available colors in the colormap to let a smaller set of colors appear to be a larger set of colors.



See "Dithering" on page 129 for more information.

Thematic Raster Layers

A thematic raster layer generally contains pixels that have been classified, or put into distinct categories. Each data file value is a class value, which is simply a number for a particular category. A thematic raster layer is stored in an image (.img) file. Only one data file value—the class value—is stored for each pixel.

Since these class values are not necessarily related, the gradations that are possible in true color mode are not usually useful in pseudo color. The class system gives the thematic layer a discrete look, in which each class can have its own color.

Color Table

When a thematic raster layer is displayed, ERDAS IMAGINE automatically creates a color table. The red, green, and blue brightness values for each class are stored in this table.

RGB Colors

Individual color schemes can be created by combining red, green, and blue in different combinations, and assigning colors to the classes of a thematic layer.

Colors can be expressed numerically, as the brightness values for each color gun. Brightness values of a display generally range from 0 to 255, however, ERDAS IMAGINE translates the values from 0 to 1. The maximum brightness value for the display device is scaled to 1. The colors listed in Table 4-2 are based on the range that is used to assign brightness values in ERDAS IMAGINE.

Table 4-2 contains only a partial listing of commonly used colors. Over 16 million colors are possible on a 24-bit display.

Table 4-2 Commonly Used RGB Colors

Color	Red	Green	Blue
Red	1	0	0
Red-Orange	1	.392	0
Orange	.608	.588	0
Yellow	1	1	0
Yellow-Green	.490	1	0
Green	0	1	0
Cyan	0	1	1
Blue	0	0	1
Blue-Violet	.392	0	.471
Violet	.588	0	.588
Black	0	0	0
White	1	1	1
Gray	.498	.498	.498
Brown	.373	.227	0

NOTE: Black is the absence of all color (0,0,0) and white is created from the highest values of all three colors (1,1,1). To lighten a color, increase all three brightness values. To darken a color, decrease all three brightness values.



Use the Raster Attribute Editor to create your own color scheme.

24-bit DirectColor and TrueColor Displays

Figure 4-8 illustrates the general process of displaying thematic raster layers on a 24-bit DirectColor display. The process is similar on a TrueColor display except that the colormap is not used.



Display a thematic raster layer from the Viewer.

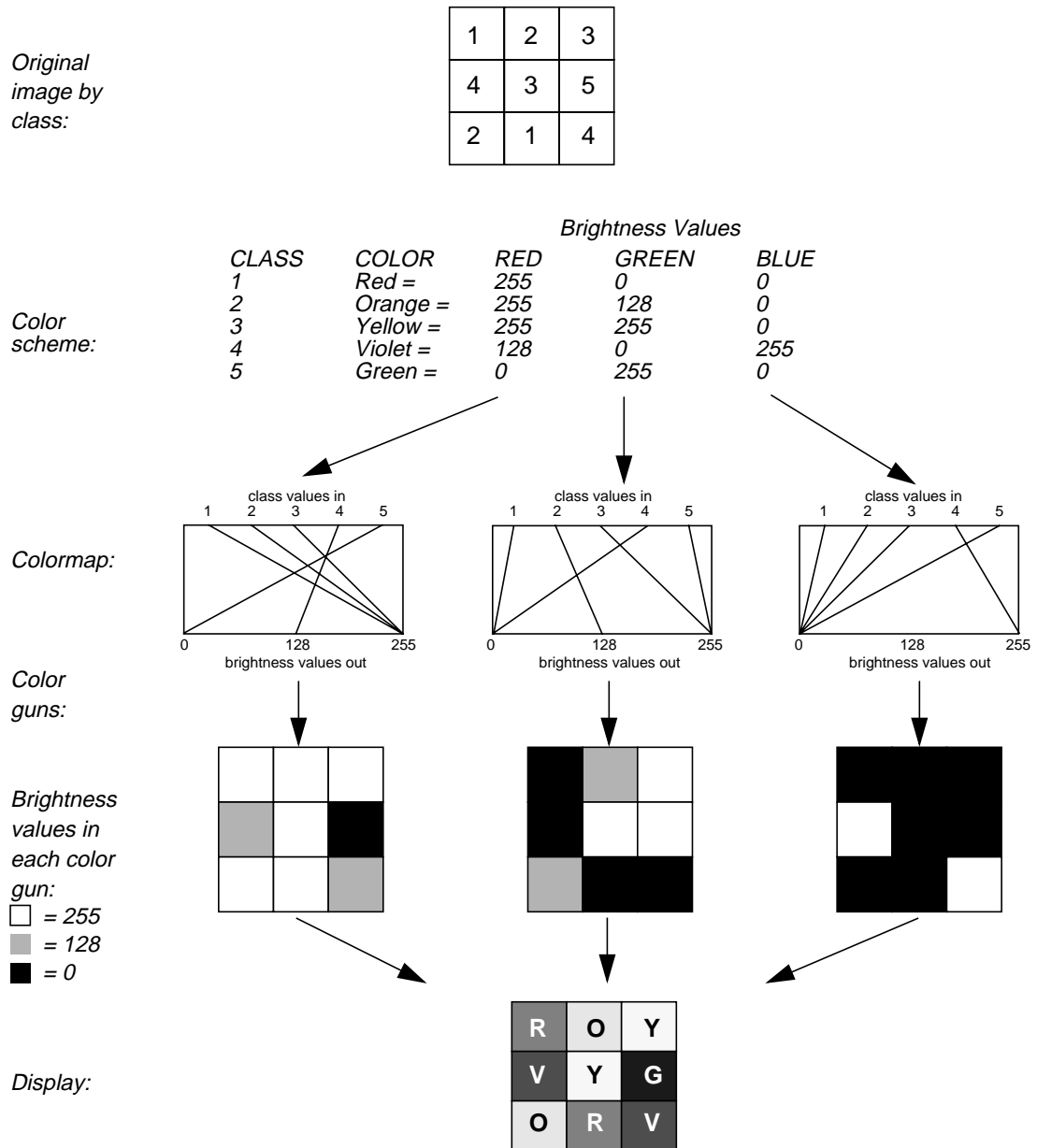


Figure 4-8 Thematic Raster Layer Display Process

8-bit PseudoColor Display

The colormap is a limited resource that is shared among all of the applications that are running concurrently. Because of the limited resources, ERDAS IMAGINE does not typically have access to the entire colormap.

Using the Viewer

The Viewer is a window for displaying raster, vector, and annotation layers. You can open as many Viewer windows as their window manager supports.

NOTE: The more Viewers that are opened simultaneously, the more RAM memory is necessary.

The Viewer not only makes digital images visible quickly, but it can also be used as a tool for image processing and raster GIS modeling. The uses of the Viewer are listed briefly in this section, and described in greater detail in other chapters of the *ERDAS Field Guide*.

Colormap

ERDAS IMAGINE does not use the entire colormap because there are other applications that also need to use it, including the window manager, terminal windows, Arc View, or a clock. Therefore, there are some limitations to the number of colors that the Viewer can display simultaneously, and flickering may occur as well.

Color Flickering

If an application requests a new color that does not exist in the colormap, the server assigns that color to an empty colorcell. However, if there are not any available colorcells and the application requires a private colorcell, then a private colormap is created for the application window. Since this is a private colormap, when the cursor is moved out of the window, the server uses the main colormap and the brightness values assigned to the colorcells. Therefore, the colors in the private colormap are not applied and the screen flickers. Once the cursor is moved into the application window, the correct colors are applied for that window.

Resampling

When a raster layer(s) is displayed, the file pixels may be resampled for display on the screen. Resampling is used to calculate pixel values when one raster grid must be fitted to another. In this case, the raster grid defined by the file must be fit to the grid of screen pixels in the Viewer.

All Viewer operations are file-based. So, any time an image is resampled in the Viewer, the Viewer uses the file as its source. If the raster layer is magnified or reduced, the Viewer refits the file grid to the new screen grid.

The resampling methods available are:

- Nearest Neighbor—uses the value of the closest pixel to assign to the output pixel value.
- Bilinear Interpolation—uses the data file values of four pixels in a 2×2 window to calculate an output value with a bilinear function.
- Cubic Convolution—uses the data file values of 16 pixels in a 4×4 window to calculate an output value with a cubic function.



These are discussed in detail in "CHAPTER 9: Rectification".



The default resampling method is Nearest Neighbor.

Preference Editor

The Preference Editor enables you to set parameters for the Viewer that affect the way the Viewer operates.



See the ERDAS IMAGINE On-Line Help for the Preference Editor for information on how to set preferences for the Viewer.

Pyramid Layers

Sometimes a large image file may take a long time to display in the Viewer or to be resampled by an application. The Pyramid Layer option enables you to display large images faster and allows certain applications to rapidly access the resampled data. Pyramid layers are image layers which are copies of the original layer successively reduced by the power of 2 and then resampled. If the raster layer is thematic, then it is resampled using the Nearest Neighbor method. If the raster layer is continuous, it is resampled by a method that is similar to Cubic Convolution. The data file values for sixteen pixels in a 4 × 4 window are used to calculate an output data file value with a filter function.



See "CHAPTER 9: Rectification" for more information on Nearest Neighbor.

The number of pyramid layers created depends on the size of the original image. A larger image produces more pyramid layers. When the Create Pyramid Layer option is selected, ERDAS IMAGINE automatically creates successively reduced layers until the final pyramid layer can be contained in one block. The default block size is 64 × 64 pixels.



See "CHAPTER 1: Raster Data" for information on block size.

Pyramid layers are added as additional layers in the image file. However, these layers cannot be accessed for display. The file size is increased by approximately one-third when pyramid layers are created. The actual increase in file size can be determined by multiplying the layer size by this formula:

$$\sum_{i=0}^n \frac{1}{4^i}$$

Where:

n = number of pyramid layers

NOTE: This equation is applicable to all types of pyramid layers: internal and external.

Image Display

Pyramid layers do not appear as layers which can be processed: they are for viewing purposes only. Therefore, they do not appear as layers in other parts of the ERDAS IMAGINE system (e.g., the Arrange Layers dialog).



The Image Files (General) section of the Preference Editor contains a preference for the Initial Pyramid Layer Number. By default, the value is set to 2. This means that the first pyramid layer generated is discarded. In Figure 4-9 below, the 2K × 2K layer is discarded. If you wish to keep that layer, then set the Initial Pyramid Layer Number to 1.



Pyramid layers can be deleted through the Image Information utility. However, when pyramid layers are deleted, they are not deleted from the image file; therefore, the image file size does not change, but ERDAS IMAGINE utilizes this file space, if necessary. Pyramid layers are deleted from viewing and resampling access only - that is, they can no longer be viewed or used in an application.

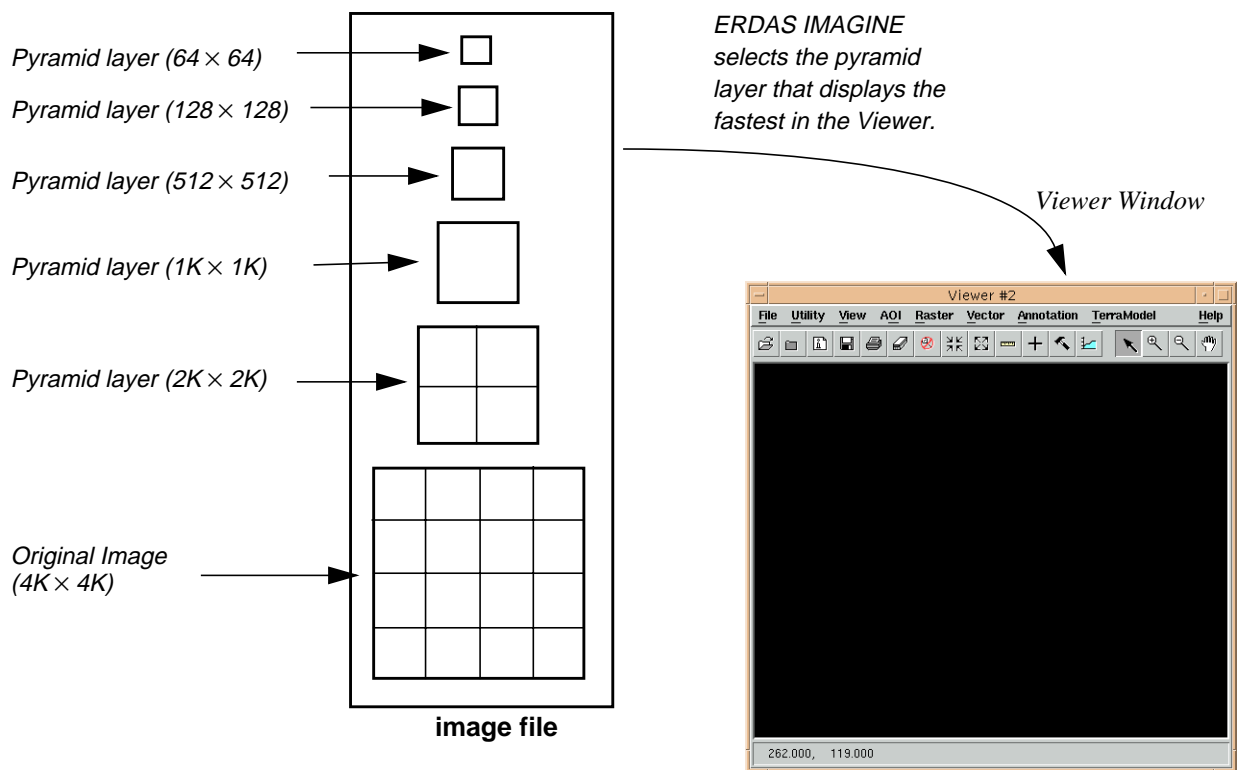


Figure 4-9 Pyramid Layers

For example, a file that is 4K × 4K pixels could take a long time to display when the image is fit to the Viewer. The Compute Pyramid Layers option creates additional layers successively reduced from 4K × 4K, to 2K × 2K, 1K × 1K, 512 × 512, 128 × 128, down to 64 × 64. ERDAS IMAGINE then selects the pyramid layer size most appropriate for display in the Viewer window when the image is displayed.



The Compute Pyramid Layers option is available from Import and the Image Information utility.



For more information about the .img format, see "CHAPTER 1: Raster Data" and the On-Line Help.

External Pyramid Layers

Pyramid layers can be either internal or external. If you choose external pyramid layers, they are stored with the same name in the same directory as the image with which they are associated, but with the .rrd extension. For example, an image named tm_image1.img has external pyramid layers contained in a file named tm_image1.rrd.

The extension .rrd stands for reduced resolution data set. You can delete the external pyramid layers associated with an image by accessing the Image Information dialog. Unlike internal pyramid layers, external pyramid layers do not affect the size of the associated image.

Dithering

A display is capable of viewing only a limited number of colors simultaneously. For example, an 8-bit display has a colormap with 256 colorcells, therefore, a maximum of 256 colors can be displayed at the same time. If some colors are being used for auto update color adjustment while other colors are still being used for other imagery, the color quality degrades.

Dithering lets a smaller set of colors appear to be a larger set of colors. If the desired display color is not available, a dithering algorithm mixes available colors to provide something that looks like the desired color.

For a simple example, assume the system can display only two colors: black and white, and you want to display gray. This can be accomplished by alternating the display of black and white pixels.

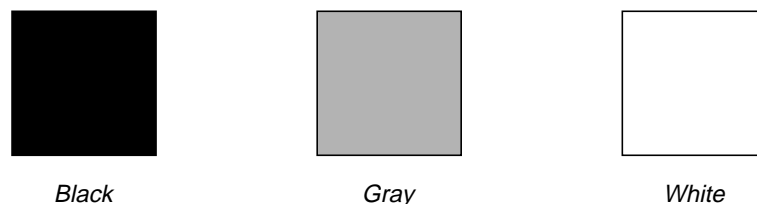


Figure 4-10 Example of Dithering

Image Display

In Figure 4-10, dithering is used between a black pixel and a white pixel to obtain a gray pixel.

The colors that the Viewer dithers between are similar to each other, and are dithered on the pixel level. Using similar colors and dithering on the pixel level makes the image appear smooth.



Dithering allows multiple images to be displayed in different Viewers without refreshing the currently displayed image(s) each time a new image is displayed.

Color Patches

When the Viewer performs dithering, it uses patches of 2×2 pixels. If the desired color has an exact match, then all of the values in the patch match it. If the desired color is halfway between two of the usable colors, the patch contains two pixels of each of the surrounding usable colors. If it is $3/4$ of the way between two usable colors, the patch contains 3 pixels of the color it is closest to, and 1 pixel of the color that is second closest. Figure 4-11 shows what the color patches would look like if the usable colors were black and white and the desired color was gray.

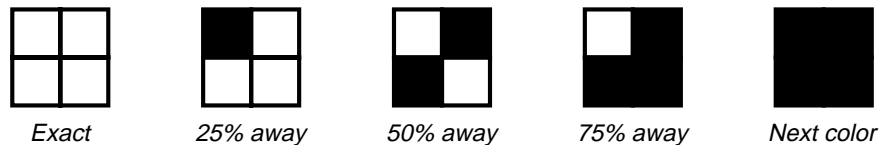


Figure 4-11 Example of Color Patches

If the desired color is not an even multiple of $1/4$ of the way between two allowable colors, it is rounded to the nearest $1/4$. The Viewer separately dithers the red, green, and blue components of a desired color.

Color Artifacts

Since the Viewer requires 2×2 pixel patches to represent a color, and actual images typically have a different color for each pixel, artifacts may appear in an image that has been dithered. Usually, the difference in color resolution is insignificant, because adjacent pixels are normally similar to each other. Similarity between adjacent pixels usually smooths out artifacts that appear.

Viewing Layers

The Viewer displays layers as one of the following types of view layers:

- annotation
- vector
- pseudo color
- gray scale
- true color

Annotation View Layer

When an annotation layer (xxx.ovr) is displayed in the Viewer, it is displayed as an annotation view layer.

Vector View Layer

A Vector layer is displayed in the Viewer as a vector view layer.

Pseudo Color View Layer

When a raster layer is displayed as a pseudo color layer in the Viewer, the colormap uses the RGB brightness values for the one layer in the RGB table. This is most appropriate for thematic layers. If the layer is a continuous raster layer, the layer would initially appear gray, since there are not any values in the RGB table.

Gray Scale View Layer

When a raster layer is displayed as a gray scale layer in the Viewer, the colormap uses the brightness values in the contrast table for one layer. This layer is then displayed in all three color guns, producing a gray scale image. A continuous raster layer may be displayed as a gray scale view layer.

True Color View Layer

Continuous raster layers should be displayed as true color layers in the Viewer. The colormap uses the RGB brightness values for three layers in the contrast table: one for each color gun to display the set of layers.

Viewing Multiple Layers

It is possible to view as many layers of all types (with the exception of vector layers, which have a limit of 10) at one time in a single Viewer.

To overlay multiple layers in one Viewer, they must all be referenced to the same map coordinate system. The layers are positioned geographically within the window, and resampled to the same scale as previously displayed layers. Therefore, raster layers in one Viewer can have different cell sizes.

When multiple layers are magnified or reduced, raster layers are resampled from the file to fit to the new scale.



Display multiple layers from the Viewer. Be sure to turn off the Clear Display checkbox when you open subsequent layers.

Overlapping Layers

When layers overlap, the order in which the layers are opened is very important. The last layer that is opened always appears to be on top of the previously opened layers.

In a raster layer, it is possible to make values of zero transparent in the Viewer, meaning that they have no opacity. Thus, if a raster layer with zeros is displayed over other layers, the areas with zero values allow the underlying layers to show through.

Image Display

Opacity is a measure of how opaque, or solid, a color is displayed in a raster layer. Opacity is a component of the color scheme of categorical data displayed in pseudo color.

- 100% opacity means that a color is completely opaque, and cannot be seen through.
- 50% opacity lets some color show, and lets some of the underlying layers show through. The effect is like looking at the underlying layers through a colored fog.
- 0% opacity allows underlying layers to show completely.



By manipulating opacity, you can compare two or more layers of raster data that are displayed in a Viewer. Opacity can be set at any value in the range of 0% to 100%. Use the Arrange Layers dialog to restack layers in a Viewer so that they overlap in a different order, if needed.

Non-Overlapping Layers

Multiple layers that are opened in the same Viewer do not have to overlap. Layers that cover distinct geographic areas can be opened in the same Viewer. The layers are automatically positioned in the Viewer window according to their map coordinates, and are positioned relative to one another geographically. The map coordinate systems for the layers must be the same.

Linking Viewers

Linking Viewers is appropriate when two Viewers cover the same geographic area (at least partially), and are referenced to the same map units. When two Viewers are linked:

- Either the same geographic point is displayed in the centers of both Viewers, or a box shows where one view fits inside the other.
- Scrolling one Viewer affects the other.
- You can manipulate the zoom ratio of one Viewer from another.
- Any inquire cursors in one Viewer appear in the other, for multiple-Viewer pixel inquiry.
- The auto-zoom is enabled, if the Viewers have the same zoom ratio and nearly the same window size.

It is often helpful to display a wide view of a scene in one Viewer, and then a close-up of a particular area in another Viewer. When two such Viewers are linked, a box opens in the wide view window to show where the close-up view lies.

Any image that is displayed at a magnification (higher zoom ratio) of another image in a linked Viewer is represented in the other Viewer by a box. If several Viewers are linked together, there may be multiple boxes in that Viewer.

Figure 4-12 shows how one view fits inside the other linked Viewer. The link box shows the extent of the larger-scale view.

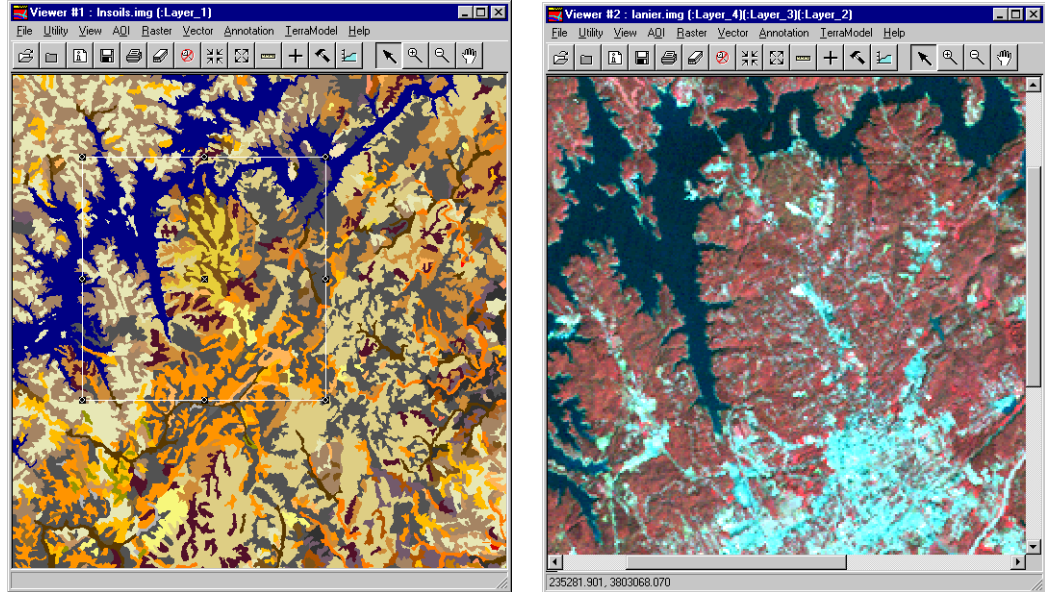


Figure 4-12 Linked Viewers

Zoom and Roam

Zooming enlarges an image on the display. When an image is zoomed, it can be roamed (scrolled) so that the desired portion of the image appears on the display screen. Any image that does not fit entirely in the Viewer can be roamed and/or zoomed. Roaming and zooming have no effect on how the image is stored in the file.

The zoom ratio describes the size of the image on the screen in terms of the number of file pixels used to store the image. It is the ratio of the number of screen pixels in the X or Y dimension to the number that are used to display the corresponding file pixels.

A zoom ratio greater than 1 is a magnification, which makes the image features appear larger in the Viewer. A zoom ratio less than 1 is a reduction, which makes the image features appear smaller in the Viewer.

Table 4-3 Overview of Zoom Ratio

A zoom ratio of 1 means...	each file pixel is displayed with 1 screen pixel in the Viewer.
A zoom ratio of 2 means...	each file pixel is displayed with a block of 2 × 2 screen pixels. Effectively, the image is displayed at 200%.
A zoom ratio of 0.5 means...	each block of 2 × 2 file pixels is displayed with 1 screen pixel. Effectively, the image is displayed at 50%.

NOTE: ERDAS IMAGINE allows floating point zoom ratios, so that images can be zoomed at virtually any scale (i.e., continuous fractional zoom). Resampling is necessary whenever an image is displayed with a new pixel grid. The resampling method used when an image is zoomed is the same one used when the image is displayed, as specified in the Open Raster Layer dialog. The default resampling method is Nearest Neighbor.



Zoom the data in the Viewer via the Viewer menu bar, the Viewer tool bar, or the Quick View right-button menu.

Geographic Information

To prepare to run many programs, it may be necessary to determine the data file coordinates, map coordinates, or data file values for a particular pixel or a group of pixels. By displaying the image in the Viewer and then selecting the pixel(s) of interest, important information about the pixel(s) can be viewed.



The Quick View right-button menu gives you options to view information about a specific pixel. Use the Raster Attribute Editor to access information about classes in a thematic layer.



See "CHAPTER 11: Geographic Information Systems" for information about attribute data.

Enhancing Continuous Raster Layers

Working with the brightness values in the colormap is useful for image enhancement. Often, a trial and error approach is needed to produce an image that has the right contrast and highlights the right features. By using the tools in the Viewer, it is possible to quickly view the effects of different enhancement techniques, undo enhancements that are not helpful, and then save the best results to disk.



Use the Raster options from the Viewer to enhance continuous raster layers.



See "CHAPTER 5: Enhancement" for more information on enhancing continuous raster layers.

Creating New Image Files

It is easy to create a new image file (.img) from the layer(s) displayed in the Viewer. The new image file contains three continuous raster layers (RGB), regardless of how many layers are currently displayed. The Image Information utility must be used to create statistics for the new image file before the file is enhanced.

Annotation layers can be converted to raster format, and written to an image file. Or, vector data can be gridded into an image, overwriting the values of the pixels in the image plane, and incorporated into the same band as the image.



Use the Viewer to .img function to create a new image file from the currently displayed raster layers.

Image Display

CHAPTER 5

Enhancement

Introduction

Image enhancement is the process of making an image more interpretable for a particular application (Faust 1989). Enhancement makes important features of raw, remotely sensed data more interpretable to the human eye. Enhancement techniques are often used instead of classification techniques for feature extraction—studying and locating areas and objects on the ground and deriving useful information from images.

The techniques to be used in image enhancement depend upon:

- Your data—the different bands of Landsat, SPOT, and other imaging sensors are selected to detect certain features. You must know the parameters of the bands being used before performing any enhancement. (See "CHAPTER 1: Raster Data" for more details.)
- Your objective—for example, sharpening an image to identify features that can be used for training samples requires a different set of enhancement techniques than reducing the number of bands in the study. You must have a clear idea of the final product desired before enhancement is performed.
- Your expectations—what you think you are going to find.
- Your background—your experience in performing enhancement.

This chapter discusses the following enhancement techniques available with ERDAS IMAGINE:

- Data correction—radiometric and geometric correction
- Radiometric enhancement—enhancing images based on the values of individual pixels
- Spatial enhancement—enhancing images based on the values of individual and neighboring pixels
- Spectral enhancement—enhancing images by transforming the values of each pixel on a multiband basis
- Hyperspectral image processing—an extension of the techniques used for multispectral data sets
- Fourier analysis—techniques for eliminating periodic noise in imagery
- Radar imagery enhancement—techniques specifically designed for enhancing radar imagery



See "Bibliography" to find current literature that provides a more detailed discussion of image processing enhancement techniques.

Display vs. File Enhancement

With ERDAS IMAGINE, image enhancement may be performed:

Enhancement

- temporarily, upon the image that is displayed in the Viewer (by manipulating the function and display memories), or
- permanently, upon the image data in the data file.

Enhancing a displayed image is much faster than enhancing an image on disk. If one is looking for certain visual effects, it may be beneficial to perform some trial and error enhancement techniques on the display. Then, when the desired results are obtained, the values that are stored in the display device memory can be used to make the same changes to the data file.



For more information about displayed images and the memory of the display device, see "CHAPTER 4: Image Display".

Spatial Modeling Enhancements

Two types of models for enhancement can be created in ERDAS IMAGINE:

- Graphical models—use Model Maker (Spatial Modeler) to easily, and with great flexibility, construct models that can be used to enhance the data.
- Script models—for even greater flexibility, use the Spatial Modeler Language (SML) to construct models in script form. SML enables you to write scripts which can be written, edited, and run from the Spatial Modeler component or directly from the command line. You can edit models created with Model Maker using SML or Model Maker.

Although a graphical model and a script model look different, they produce the same results when applied.

Image Interpreter

ERDAS IMAGINE supplies many algorithms constructed as models, which are ready to be applied with user-input parameters at the touch of a button. These graphical models, created with Model Maker, are listed as menu functions in the Image Interpreter. These functions are mentioned throughout this chapter. Just remember, these are modeling functions which can be edited and adapted as needed with Model Maker or the SML.



See "CHAPTER 11: Geographic Information Systems" for more information on Raster Modeling.

The modeling functions available for enhancement in Image Interpreter are briefly described in Table 5-1.

Table 5-1 Description of Modeling Functions Available for Enhancement

Function	Description
SPATIAL ENHANCEMENT	These functions enhance the image using the values of individual and surrounding pixels.
Convolution	Uses a matrix to average small sets of pixels across an image.
Non-directional Edge	Averages the results from two orthogonal 1st derivative edge detectors.
Focal Analysis	Enables you to perform one of several analyses on class values in an image file using a process similar to convolution filtering.
Texture	Defines texture as a quantitative characteristic in an image.
Adaptive Filter	Varies the contrast stretch for each pixel depending upon the DN values in the surrounding moving window.
Statistical Filter	Produces the pixel output DN by averaging pixels within a moving window that fall within a statistically defined range.
Resolution Merge	Merges imagery of differing spatial resolutions.
Crisp	Sharpens the overall scene luminance without distorting the thematic content of the image.
RADIOMETRIC ENHANCEMENT	These functions enhance the image using the values of individual pixels within each band.
LUT (Lookup Table) Stretch	Creates an output image that contains the data values as modified by a lookup table.
Histogram Equalization	Redistributes pixel values with a nonlinear contrast stretch so that there are approximately the same number of pixels with each value within a range.
Histogram Match	Mathematically determines a lookup table that converts the histogram of one image to resemble the histogram of another.
Brightness Inversion	Allows both linear and nonlinear reversal of the image intensity range.
Haze Reduction*	Dehazes Landsat 4 and 5 TM data and panchromatic data.
Noise Reduction*	Removes noise using an adaptive filter.
Destripe TM Data	Removes striping from a raw TM4 or TM5 data file.
SPECTRAL ENHANCEMENT	These functions enhance the image by transforming the values of each pixel on a multiband basis.
Principal Components	Compresses redundant data values into fewer bands, which are often more interpretable than the source data.
Inverse Principal Components	Performs an inverse principal components analysis.

Table 5-1 Description of Modeling Functions Available for Enhancement

Function	Description
Decorrelation Stretch	Applies a contrast stretch to the principal components of an image.
Tasseled Cap	Rotates the data structure axes to optimize data viewing for vegetation studies.
RGB to IHS	Transforms red, green, blue values to intensity, hue, saturation values.
IHS to RGB	Transforms intensity, hue, saturation values to red, green, blue values.
Indices	Performs band ratios that are commonly used in mineral and vegetation studies.
Natural Color	Simulates natural color for TM data.
FOURIER ANALYSIS	These functions enhance the image by applying a Fourier Transform to the data. NOTE: These functions are currently view only—no manipulation is allowed.
Fourier Transform*	Enables you to utilize a highly efficient version of the Discrete Fourier Transform (DFT).
Fourier Transform Editor*	Enables you to edit Fourier images using many interactive tools and filters.
Inverse Fourier Transform*	Computes the inverse two-dimensional Fast Fourier Transform (FFT) of the spectrum stored.
Fourier Magnitude*	Converts the Fourier Transform image into the more familiar Fourier Magnitude image.
Periodic Noise Removal*	Automatically removes striping and other periodic noise from images.
Homomorphic Filter*	Enhances imagery using an illumination/reflectance model.

* Indicates functions that are not graphical models.

NOTE: There are other Image Interpreter functions that do not necessarily apply to image enhancement.

Correcting Data

Each generation of sensors shows improved data acquisition and image quality over previous generations. However, some anomalies still exist that are inherent to certain sensors and can be corrected by applying mathematical formulas derived from the distortions (Lillesand and Kiefer 1979). In addition, the natural distortion that results from the curvature and rotation of the Earth in relation to the sensor platform produces distortions in the image data, which can also be corrected.

Radiometric Correction

Generally, there are two types of data correction: radiometric and geometric. Radiometric correction addresses variations in the pixel intensities (DNs) that are not caused by the object or scene being scanned. These variations include:

- differing sensitivities or malfunctioning of the detectors
- topographic effects
- atmospheric effects

Geometric Correction

Geometric correction addresses errors in the relative positions of pixels. These errors are induced by:

- sensor viewing geometry
- terrain variations



Because of the differences in radiometric and geometric correction between traditional, passively detected visible/infrared imagery and actively acquired radar imagery, the two are discussed separately. See “Radar Imagery Enhancement” on page 196.

**Radiometric Correction:
Visible/Infrared Imagery**
Striping

Striping or banding occurs if a detector goes out of adjustment—that is, it provides readings consistently greater than or less than the other detectors for the same band over the same ground cover.

Some Landsat 1, 2, and 3 data have striping every sixth line, because of improper calibration of some of the 24 detectors that were used by the MSS. The stripes are not constant data values, nor is there a constant error factor or bias. The differing response of the errant detector is a complex function of the data value sensed.

This problem has been largely eliminated in the newer sensors. Various algorithms have been advanced in current literature to help correct this problem in the older data. Among these algorithms are simple along-line convolution, high-pass filtering, and forward and reverse principal component transformations (Crippen 1989).

Data from airborne multispectral or hyperspectral imaging scanners also shows a pronounced striping pattern due to varying offsets in the multielement detectors. This effect can be further exacerbated by unfavorable sun angle. These artifacts can be minimized by correcting each scan line to a scene-derived average (Kruse 1988).



Use the Image Interpreter or the Spatial Modeler to implement algorithms to eliminate striping. The Spatial Modeler editing capabilities allow you to adapt the algorithms to best address the data. The IMAGINE Radar Interpreter Adjust Brightness function also corrects some of these problems.

Line Dropout

Another common remote sensing device error is line dropout. Line dropout occurs when a detector either completely fails to function, or becomes temporarily saturated during a scan (like the effect of a camera flash on the retina). The result is a line or partial line of data with higher data file values, creating a horizontal streak until the detector(s) recovers, if it recovers.

Line dropout is usually corrected by replacing the bad line with a line of estimated data file values, which is based on the lines above and below it.

Atmospheric Effects

The effects of the atmosphere upon remotely-sensed data are not considered errors, since they are part of the signal received by the sensing device (Bernstein 1983). However, it is often important to remove atmospheric effects, especially for scene matching and change detection analysis.

Over the past 30 years, a number of algorithms have been developed to correct for variations in atmospheric transmission. Four categories are mentioned here:

- dark pixel subtraction
- radiance to reflectance conversion
- linear regressions
- atmospheric modeling



Use the Spatial Modeler to construct the algorithms for these operations.

Dark Pixel Subtraction

The dark pixel subtraction technique assumes that the pixel of lowest DN in each band should really be zero, and hence its radiometric value (DN) is the result of atmosphere-induced additive errors. These assumptions are very tenuous and recent work indicates that this method may actually degrade rather than improve the data (Crane 1971; Chavez et al 1977).

Radiance to Reflectance Conversion

Radiance to reflectance conversion requires knowledge of the true ground reflectance of at least two targets in the image. These can come from either at-site reflectance measurements, or they can be taken from a reflectance table for standard materials. The latter approach involves assumptions about the targets in the image.

Linear Regressions

A number of methods using linear regressions have been tried. These techniques use bispectral plots and assume that the position of any pixel along that plot is strictly a result of illumination. The slope then equals the relative reflectivities for the two spectral bands. At an illumination of zero, the regression plots should pass through the bispectral origin. Offsets from this represent the additive extraneous components, due to atmosphere effects (Crippen 1987).

Atmospheric Modeling

Atmospheric modeling is computationally complex and requires either assumptions or inputs concerning the atmosphere at the time of imaging. The atmospheric model used to define the computations is frequently Lowtran or Modtran (Kneizys et al 1988). This model requires inputs such as atmospheric profile (e.g., pressure, temperature, water vapor, ozone), aerosol type, elevation, solar zenith angle, and sensor viewing angle.

Accurate atmospheric modeling is essential in preprocessing hyperspectral data sets where bandwidths are typically 10 nm or less. These narrow bandwidth corrections can then be combined to simulate the much wider bandwidths of Landsat or SPOT sensors (Richter 1990).

Geometric Correction

As previously noted, geometric correction is applied to raw sensor data to correct errors of perspective due to the Earth's curvature and sensor motion. Today, some of these errors are commonly removed at the sensor's data processing center. In the past, some data from Landsat MSS 1, 2, and 3 were not corrected before distribution.

Many visible/infrared sensors are not nadir-viewing; they look to the side. For some applications, such as stereo viewing or DEM generation, this is an advantage. For other applications, it is a complicating factor.

In addition, even a nadir-viewing sensor is viewing only the scene center at true nadir. Other pixels, especially those on the view periphery, are viewed off-nadir. For scenes covering very large geographic areas (such as AVHRR), this can be a significant problem.

This and other factors, such as Earth curvature, result in geometric imperfections in the sensor image. Terrain variations have the same distorting effect, but on a smaller (pixel-by-pixel) scale. These factors can be addressed by rectifying the image to a map.



See "CHAPTER 9: Rectification" for more information on geometric correction using rectification and "CHAPTER 7: Photogrammetric Concepts" for more information on orthorectification.

A more rigorous geometric correction utilizes a DEM and sensor position information to correct these distortions. This is orthorectification.

Radiometric Enhancement

Radiometric enhancement deals with the individual values of the pixels in the image. It differs from spatial enhancement (discussed on page 154), which takes into account the values of neighboring pixels.

Depending on the points and the bands in which they appear, radiometric enhancements that are applied to one band may not be appropriate for other bands. Therefore, the radiometric enhancement of a multiband image can usually be considered as a series of independent, single-band enhancements (Faust 1989).

Radiometric enhancement usually does not bring out the contrast of every pixel in an image. Contrast can be lost between some pixels, while gained on others.

(*j* and *k* are reference points)

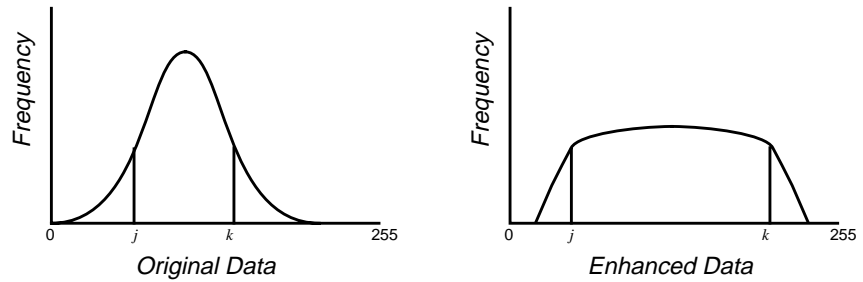


Figure 5-1 Histograms of Radiometrically Enhanced Data

In Figure 5-1, the range between *j* and *k* in the histogram of the original data is about one third of the total range of the data. When the same data are radiometrically enhanced, the range between *j* and *k* can be widened. Therefore, the pixels between *j* and *k* gain contrast—it is easier to distinguish different brightness values in these pixels.

However, the pixels outside the range between *j* and *k* are more grouped together than in the original histogram to compensate for the stretch between *j* and *k*. Contrast among these pixels is lost.

Contrast Stretching

When radiometric enhancements are performed on the display device, the transformation of data file values into brightness values is illustrated by the graph of a lookup table.

For example, Figure 5-2 shows the graph of a lookup table that increases the contrast of data file values in the middle range of the input data (the range within the brackets). Note that the input range within the bracket is narrow, but the output brightness values for the same pixels are stretched over a wider range. This process is called contrast stretching.

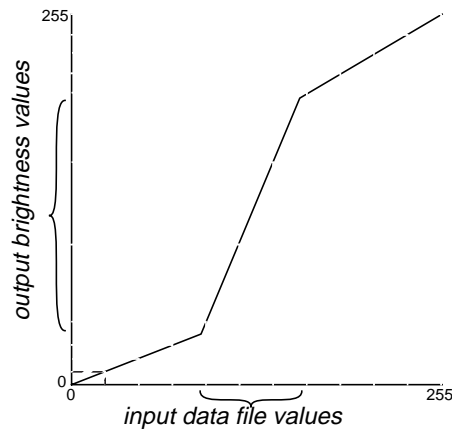


Figure 5-2 Graph of a Lookup Table

Notice that the graph line with the steepest (highest) slope brings out the most contrast by stretching output values farther apart.

Linear and Nonlinear

The terms linear and nonlinear, when describing types of spectral enhancement, refer to the function that is applied to the data to perform the enhancement. A piecewise linear stretch uses a polyline function to increase contrast to varying degrees over different ranges of the data, as in Figure 5-3.

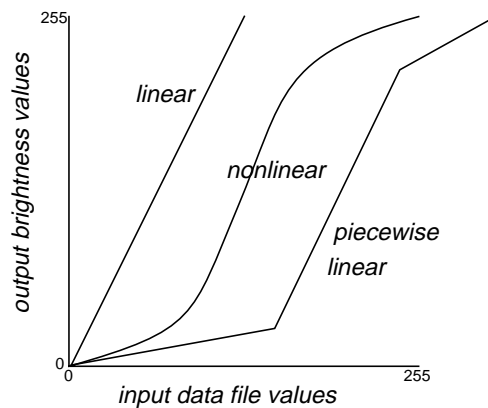


Figure 5-3 Enhancement with Lookup Tables

Linear Contrast Stretch

A linear contrast stretch is a simple way to improve the visible contrast of an image. It is often necessary to contrast-stretch raw image data, so that they can be seen on the display.

In most raw data, the data file values fall within a narrow range—usually a range much narrower than the display device is capable of displaying. That range can be expanded to utilize the total range of the display device (usually 0 to 255).



A two standard deviation linear contrast stretch is automatically applied to images displayed in the Viewer.

Nonlinear Contrast Stretch

A nonlinear spectral enhancement can be used to gradually increase or decrease contrast over a range, instead of applying the same amount of contrast (slope) across the entire image. Usually, nonlinear enhancements bring out the contrast in one range while decreasing the contrast in other ranges. The graph of the function in Figure 5-4 shows one example.

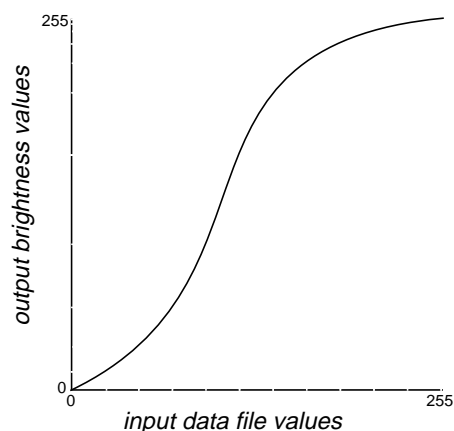


Figure 5-4 Nonlinear Radiometric Enhancement

Piecewise Linear Contrast Stretch

A piecewise linear contrast stretch allows for the enhancement of a specific portion of data by dividing the lookup table into three sections: low, middle, and high. It enables you to create a number of straight line segments that can simulate a curve. You can enhance the contrast or brightness of any section in a single color gun at a time. This technique is very useful for enhancing image areas in shadow or other areas of low contrast.



In ERDAS IMAGINE, the Piecewise Linear Contrast function is set up so that there are always pixels in each data file value from 0 to 255. You can manipulate the percentage of pixels in a particular range, but you cannot eliminate a range of data file values.

A piecewise linear contrast stretch normally follows two rules:

- 1) The data values are continuous; there can be no break in the values between High, Middle, and Low. Range specifications adjust in relation to any changes to maintain the data value range.
- 2) The data values specified can go only in an upward, increasing direction, as shown in Figure 5-5.

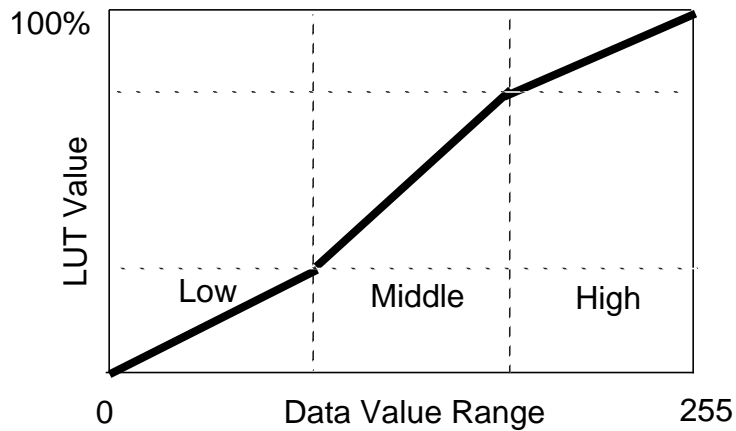


Figure 5-5 Piecewise Linear Contrast Stretch

The contrast value for each range represents the percent of the available output range that particular range occupies. The brightness value for each range represents the middle of the total range of brightness values occupied by that range. Since rules 1 and 2 above are enforced, as the contrast and brightness values are changed, they may affect the contrast and brightness of other ranges. For example, if the contrast of the low range increases, it forces the contrast of the middle to decrease.

Contrast Stretch on the Display

Usually, a contrast stretch is performed on the display device only, so that the data file values are not changed. Lookup tables are created that convert the range of data file values to the maximum range of the display device. You can then edit and save the contrast stretch values and lookup tables as part of the raster data image file. These values are loaded into the Viewer as the default display values the next time the image is displayed.



In ERDAS IMAGINE, you can permanently change the data file values to the lookup table values. Use the Image Interpreter LUT Stretch function to create an .img output file with the same data values as the displayed contrast stretched image.



See "CHAPTER 1: Raster Data" for more information on the data contained in image files.

The statistics in the image file contain the mean, standard deviation, and other statistics on each band of data. The mean and standard deviation are used to determine the range of data file values to be translated into brightness values or new data file values. You can specify the number of standard deviations from the mean that are to be used in the contrast stretch. Usually the data file values that are two standard deviations above and below the mean are used. If the data have a normal distribution, then this range represents approximately 95 percent of the data.

The mean and standard deviation are used instead of the minimum and maximum data file values because the minimum and maximum data file values are usually not representative of most of the data. A notable exception occurs when the feature being sought is in shadow. The shadow pixels are usually at the low extreme of the data file values, outside the range of two standard deviations from the mean.



The use of these statistics in contrast stretching is discussed and illustrated in "CHAPTER 4: Image Display". Statistical terms are discussed in "Appendix A: Math Topics".

Varying the Contrast Stretch

There are variations of the contrast stretch that can be used to change the contrast of values over a specific range, or by a specific amount. By manipulating the lookup tables as in Figure 5-6, the maximum contrast in the features of an image can be brought out.

Figure 5-6 shows how the contrast stretch manipulates the histogram of the data, increasing contrast in some areas and decreasing it in others. This is also a good example of a piecewise linear contrast stretch, which is created by adding breakpoints to the histogram.

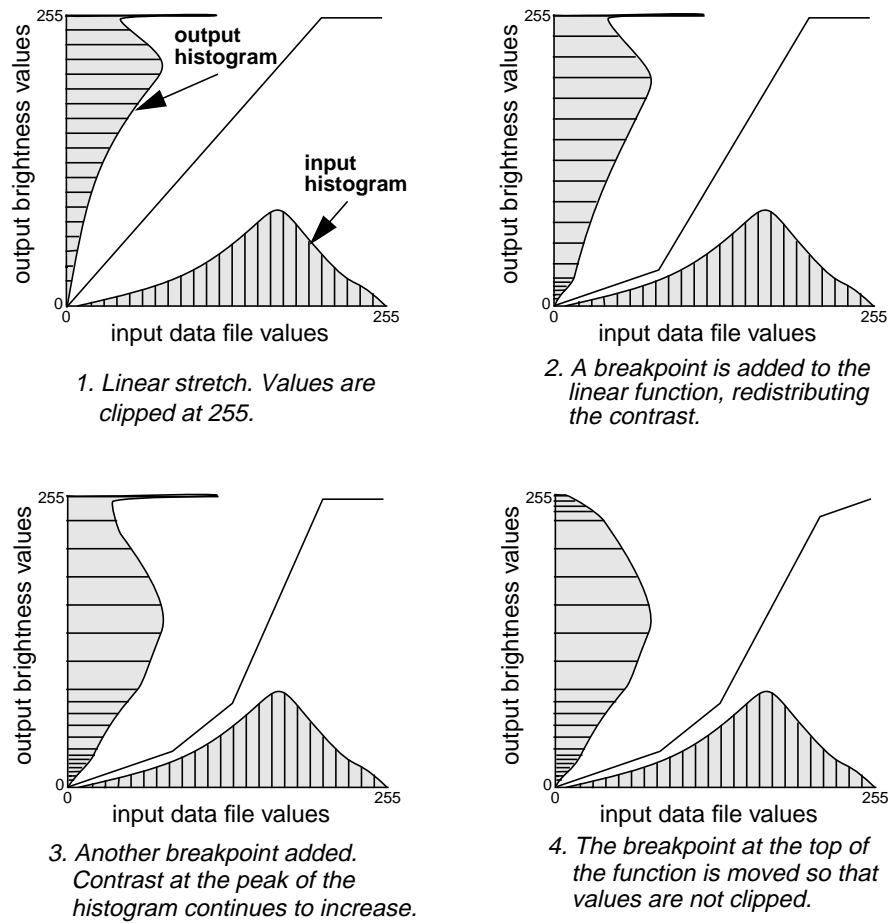


Figure 5-6 Contrast Stretch Using Lookup Tables, and Effect on Histogram

Histogram Equalization

Histogram equalization is a nonlinear stretch that redistributes pixel values so that there is approximately the same number of pixels with each value within a range. The result approximates a flat histogram. Therefore, contrast is increased at the peaks of the histogram and lessened at the tails.

Histogram equalization can also separate pixels into distinct groups if there are few output values over a wide range. This can have the visual effect of a crude classification.

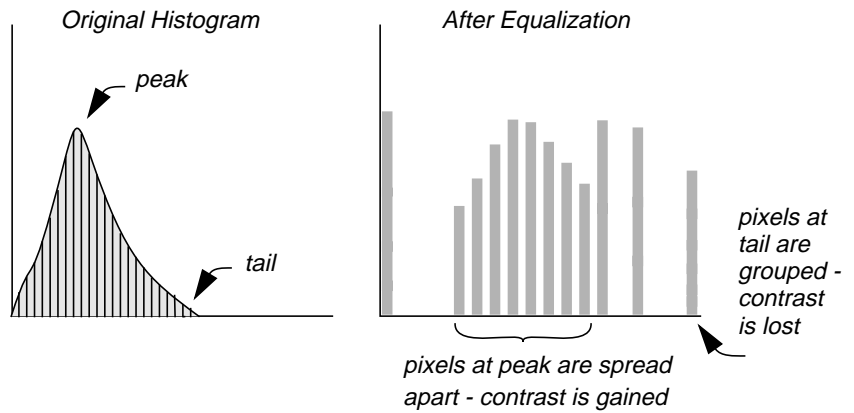


Figure 5-7 Histogram Equalization

To perform a histogram equalization, the pixel values of an image (either data file values or brightness values) are reassigned to a certain number of bins, which are simply numbered sets of pixels. The pixels are then given new values, based upon the bins to which they are assigned.

The following parameters are entered:

- N - the number of bins to which pixel values can be assigned. If there are many bins or many pixels with the same value(s), some bins may be empty.
- M - the maximum of the range of the output values. The range of the output values is from 0 to M .

The total number of pixels is divided by the number of bins, equaling the number of pixels per bin, as shown in the following equation:

$$A = \frac{T}{N}$$

Where:

- N = the number of bins
- T = the total number of pixels in the image
- A = the equalized number of pixels per bin

The pixels of each input value are assigned to bins, so that the number of pixels in each bin is as close to A as possible. Consider Figure 5-8:

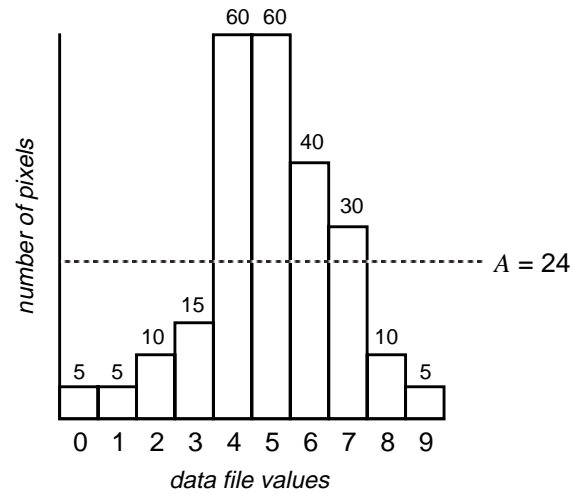


Figure 5-8 Histogram Equalization Example

There are 240 pixels represented by this histogram. To equalize this histogram to 10 bins, there would be:

$$240 \text{ pixels} / 10 \text{ bins} = 24 \text{ pixels per bin} = A$$

To assign pixels to bins, the following equation is used:

$$B_i = \text{int} \left[\frac{\left(\sum_{k=1}^{i-1} H_k \right) + \frac{H_i}{2}}{A} \right]$$

Where:

- A = equalized number of pixels per bin (see above)
- H_i = the number of values with the value i (histogram)
- int = integer function (truncating real numbers to integer)
- B_i = bin number for pixels with value i

Source: Modified from Gonzalez and Wintz 1977

The 10 bins are rescaled to the range 0 to M . In this example, $M = 9$, because the input values ranged from 0 to 9, so that the equalized histogram can be compared to the original. The output histogram of this equalized image looks like Figure 5-9:

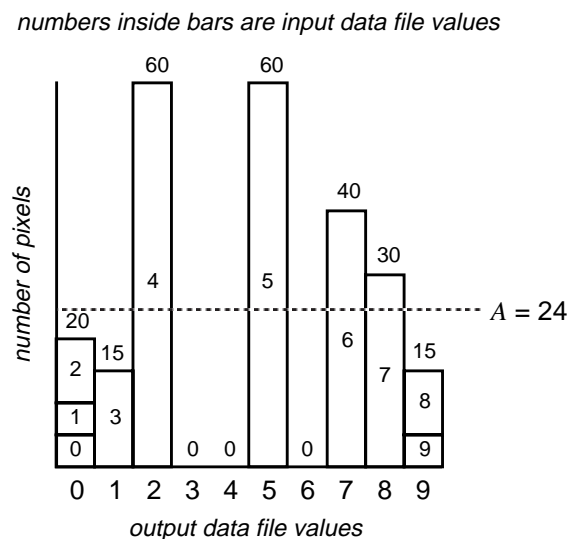


Figure 5-9 Equalized Histogram

Effect on Contrast

By comparing the original histogram of the example data with the one above, you can see that the enhanced image gains contrast in the peaks of the original histogram. For example, the input range of 3 to 7 is stretched to the range 1 to 8. However, data values at the tails of the original histogram are grouped together. Input values 0 through 2 all have the output value of 0. So, contrast among the tail pixels, which usually make up the darkest and brightest regions of the input image, is lost.

The resulting histogram is not exactly flat, since the pixels can rarely be grouped together into bins with an equal number of pixels. Sets of pixels with the same value are never split up to form equal bins.

Level Slice

A level slice is similar to a histogram equalization in that it divides the data into equal amounts. A level slice on a true color display creates a stair-stepped lookup table. The effect on the data is that input file values are grouped together at regular intervals into a discrete number of levels, each with one output brightness value.

To perform a true color level slice, you must specify a range for the output brightness values and a number of output levels. The lookup table is then stair-stepped so that there is an equal number of input pixels in each of the output levels.

Histogram Matching

Histogram matching is the process of determining a lookup table that converts the histogram of one image to resemble the histogram of another. Histogram matching is useful for matching data of the same or adjacent scenes that were scanned on separate days, or are slightly different because of sun angle or atmospheric effects. This is especially useful for mosaicking or change detection.

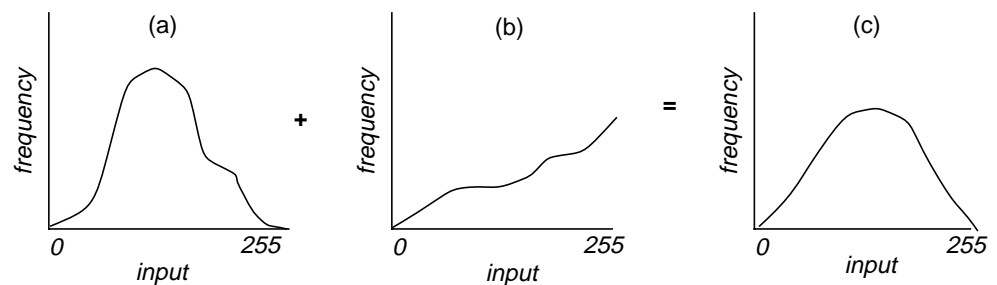
To achieve good results with histogram matching, the two input images should have similar characteristics:

- The general shape of the histogram curves should be similar.
- Relative dark and light features in the image should be the same.
- For some applications, the spatial resolution of the data should be the same.
- The relative distributions of land covers should be about the same, even when matching scenes that are not of the same area. If one image has clouds and the other does not, then the clouds should be removed before matching the histograms. This can be done using the AOI function. The AOI function is available from the Viewer menu bar.



In ERDAS IMAGINE, histogram matching is performed band-to-band (e.g., band 2 of one image is matched to band 2 of the other image).

To match the histograms, a lookup table is mathematically derived, which serves as a function for converting one histogram to the other, as illustrated in Figure 5-10.



Source histogram (a), mapped through the lookup table (b), approximates model histogram (c).

Figure 5-10 Histogram Matching

Brightness Inversion

The brightness inversion functions produce images that have the opposite contrast of the original image. Dark detail becomes light, and light detail becomes dark. This can also be used to invert a negative image that has been scanned to produce a positive image.

Brightness inversion has two options: inverse and reverse. Both options convert the input data range (commonly 0 - 255) to 0 - 1.0. A min-max remapping is used to simultaneously stretch the image and handle any input bit format. The output image is in floating point format, so a min-max stretch is used to convert the output image into 8-bit format.

Enhancement

Inverse is useful for emphasizing detail that would otherwise be lost in the darkness of the low DN pixels. This function applies the following algorithm:

$$DN_{out} = 1.0 \text{ if } 0.0 < DN_{in} < 0.1$$
$$DN_{out} = \frac{0.1}{DN_{in}} \text{ if } 0.1 < DN < 1$$

Reverse is a linear function that simply reverses the DN values:

$$DN_{out} = 1.0 - DN_{in}$$

Source: Pratt 1986

Spatial Enhancement

While radiometric enhancements operate on each pixel individually, spatial enhancement modifies pixel values based on the values of surrounding pixels. Spatial enhancement deals largely with spatial frequency, which is the difference between the highest and lowest values of a contiguous set of pixels. Jensen (1986) defines spatial frequency as “the number of changes in brightness value per unit distance for any particular part of an image.”

Consider the examples in Figure 5-11:

- zero spatial frequency—a flat image, in which every pixel has the same value
- low spatial frequency—an image consisting of a smoothly varying gray scale
- highest spatial frequency—an image consisting of a checkerboard of black and white pixels

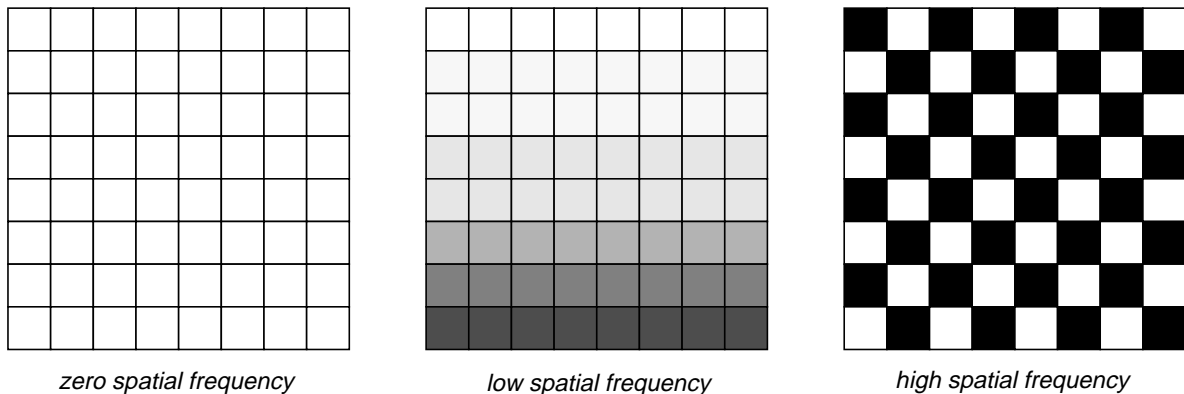


Figure 5-11 Spatial Frequencies

This section contains a brief description of the following:

- Convolution, Crisp, and Adaptive filtering
- Resolution merging



See "Radar Imagery Enhancement" on page 196 for a discussion of Edge Detection and Texture Analysis. These spatial enhancement techniques can be applied to any type of data.

Convolution Filtering

Convolution filtering is the process of averaging small sets of pixels across an image. Convolution filtering is used to change the spatial frequency characteristics of an image (Jensen 1996).

A convolution kernel is a matrix of numbers that is used to average the value of each pixel with the values of surrounding pixels in a particular way. The numbers in the matrix serve to weight this average toward particular pixels. These numbers are often called coefficients, because they are used as such in the mathematical equations.



In ERDAS IMAGINE, there are four ways you can apply convolution filtering to an image:

- 1) The kernel filtering option in the Viewer
- 2) The Convolution function in Image Interpreter
- 3) The IMAGINE Radar Interpreter Edge Enhancement function
- 4) The Convolution function in Model Maker

Filtering is a broad term, which refers to the altering of spatial or spectral features for image enhancement (Jensen 1996). Convolution filtering is one method of spatial filtering. Some texts may use the terms synonymously.

Convolution Example

To understand how one pixel is convolved, imagine that the convolution kernel is overlaid on the data file values of the image (in one band), so that the pixel to be convolved is in the center of the window.

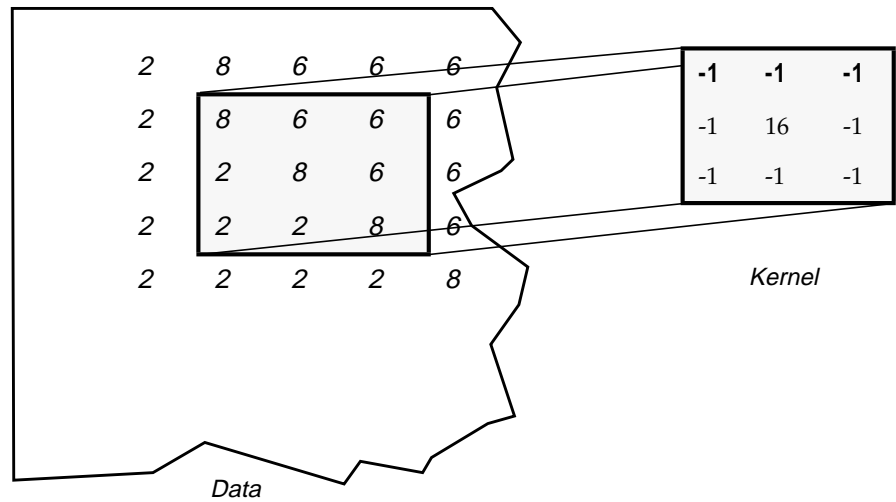


Figure 5-12 Applying a Convolution Kernel

Figure 5-12 shows a 3 × 3 convolution kernel being applied to the pixel in the third column, third row of the sample data (the pixel that corresponds to the center of the kernel).

To compute the output value for this pixel, each value in the convolution kernel is multiplied by the image pixel value that corresponds to it. These products are summed, and the total is divided by the sum of the values in the kernel, as shown here:

$$\begin{aligned}
 & \text{integer } [\\
 & (-1 \times 8) + (-1 \times 6) + (-1 \times 6) + \\
 & (-1 \times 2) + (16 \times 8) + (-1 \times 6) + \\
 & (-1 \times 2) + (-1 \times 2) + (-1 \times 8) : (-1 + -1 + -1 + -1 + 16 + -1 + -1 + -1 + -1)] \\
 & = \text{int} [(128-40) / (16-8)] \\
 & = \text{int} (88 / 8) = \text{int} (11) = 11
 \end{aligned}$$

When the 2 × 2 set of pixels near the center of this 5 × 5 image is convolved, the output values are:

	1	2	3	4	5
1	2	8	6	6	6
2	2	11	5	6	6
3	2	0	11	6	6
4	2	2	2	8	6
5	2	2	2	2	8

Figure 5-13 Output Values for Convolution Kernel

The kernel used in this example is a high frequency kernel, as explained below. It is important to note that the relatively lower values become lower, and the higher values become higher, thus increasing the spatial frequency of the image.

Convolution Formula

The following formula is used to derive an output data file value for the pixel being convolved (in the center):

$$V = \left[\frac{\sum_{i=1}^q \left(\sum_{j=1}^q f_{ij} d_{ij} \right)}{F} \right]$$

Where:

f_{ij} = the coefficient of a convolution kernel at position i,j (in the kernel)

d_{ij} = the data value of the pixel that corresponds to f_{ij}

q = the dimension of the kernel, assuming a square kernel (if $q=3$, the kernel is 3×3)

F = either the sum of the coefficients of the kernel, or 1 if the sum of coefficients is zero

V = the output pixel value

In cases where V is less than 0, V is clipped to 0.

Source: Modified from Jensen 1996; Schowengerdt 1983

The sum of the coefficients (F) is used as the denominator of the equation above, so that the output values are in relatively the same range as the input values. Since F cannot equal zero (division by zero is not defined), F is set to 1 if the sum is zero.

Zero-Sum Kernels

Zero-sum kernels are kernels in which the sum of all coefficients in the kernel equals zero. When a zero-sum kernel is used, then the sum of the coefficients is not used in the convolution equation, as above. In this case, no division is performed ($F = 1$), since division by zero is not defined.

This generally causes the output values to be:

- zero in areas where all input values are equal (no edges)
- low in areas of low spatial frequency
- extreme in areas of high spatial frequency (high values become much higher, low values become much lower)

Therefore, a zero-sum kernel is an edge detector, which usually smooths out or zeros out areas of low spatial frequency and creates a sharp contrast where spatial frequency is high, which is at the edges between homogeneous (homogeneity is low spatial frequency) groups of pixels. The resulting image often consists of only edges and zeros.

Zero-sum kernels can be biased to detect edges in a particular direction. For example, this 3×3 kernel is biased to the south (Jensen 1996).

-1	-1	-1
1	-2	1
1	1	1



See the section on "Edge Detection" on page 204 for more detailed information.

High-Frequency Kernels

A high-frequency kernel, or high-pass kernel, has the effect of increasing spatial frequency.

High-frequency kernels serve as edge enhancers, since they bring out the edges between homogeneous groups of pixels. Unlike edge detectors (such as zero-sum kernels), they highlight edges and do not necessarily eliminate other features.

-1	-1	-1
-1	16	-1
-1	-1	-1

When this kernel is used on a set of pixels in which a relatively low value is surrounded by higher values, like this...

BEFORE			AFTER		
204	200	197	204	200	197
201	106	209	201	9	209
198	200	210	198	200	210

...the low value gets lower. Inversely, when the kernel is used on a set of pixels in which a relatively high value is surrounded by lower values...

BEFORE			AFTER		
64	60	57	64	60	57
61	125	69	61	187	69
58	60	70	58	60	70

...the high value becomes higher. In either case, spatial frequency is increased by this kernel.

Low-Frequency Kernels

Below is an example of a low-frequency kernel, or low-pass kernel, which decreases spatial frequency.

1	1	1
1	1	1
1	1	1

This kernel simply averages the values of the pixels, causing them to be more homogeneous. The resulting image looks either more smooth or more blurred.



For information on applying filters to thematic layers, see "CHAPTER 11: Geographic Information Systems".

Crisp

The Crisp filter sharpens the overall scene luminance without distorting the interband variance content of the image. This is a useful enhancement if the image is blurred due to atmospheric haze, rapid sensor motion, or a broad point spread function of the sensor.

The algorithm used for this function is:

- 1) Calculate principal components of multiband input image.
- 2) Convolve PC-1 with summary filter.
- 3) Retransform to RGB space.

Source: ERDAS (Faust 1993)

The logic of the algorithm is that the first principal component (PC-1) of an image is assumed to contain the overall scene luminance. The other PCs represent intra-scene variance. Thus, you can sharpen only PC-1 and then reverse the principal components calculation to reconstruct the original image. Luminance is sharpened, but variance is retained.

Resolution Merge

The resolution of a specific sensor can refer to radiometric, spatial, spectral, or temporal resolution.



See "CHAPTER 1: Raster Data" for a full description of resolution types.

Landsat TM sensors have seven bands with a spatial resolution of 28.5 m. SPOT panchromatic has one broad band with very good spatial resolution—10 m. Combining these two images to yield a seven-band data set with 10 m resolution provides the best characteristics of both sensors.

A number of models have been suggested to achieve this image merge. Welch and Ehlers (1987) used forward-reverse RGB to IHS transforms, replacing I (from transformed TM data) with the SPOT panchromatic image. However, this technique is limited to three bands (R,G,B).

Chavez (1991), among others, uses the forward-reverse principal components transforms with the SPOT image, replacing PC-1.

In the above two techniques, it is assumed that the intensity component (PC-1 or I) is spectrally equivalent to the SPOT panchromatic image, and that all the spectral information is contained in the other PCs or in H and S. Since SPOT data do not cover the full spectral range that TM data do, this assumption does not strictly hold. It is unacceptable to resample the thermal band (TM6) based on the visible (SPOT panchromatic) image.

Another technique (Schowengerdt 1980) combines a high frequency image derived from the high spatial resolution data (i.e., SPOT panchromatic) additively with the high spectral resolution Landsat TM image.

The Resolution Merge function has two different options for resampling low spatial resolution data to a higher spatial resolution while retaining spectral information:

- forward-reverse principal components transform
- multiplicative

Principal Components Merge

Because a major goal of this merge is to retain the spectral information of the six TM bands (1-5, 7), this algorithm is mathematically rigorous. It is assumed that:

- PC-1 contains only overall scene luminance; all interband variation is contained in the other 5 PCs, and
- scene luminance in the SWIR bands is identical to visible scene luminance.

With the above assumptions, the forward transform into PCs is made. PC-1 is removed and its numerical range (min to max) is determined. The high spatial resolution image is then remapped so that its histogram shape is kept constant, but it is in the same numerical range as PC-1. It is then substituted for PC-1 and the reverse transform is applied. This remapping is done so that the mathematics of the reverse transform do not distort the thematic information (Welch and Ehlers 1987).

Multiplicative

The second technique in the Image Interpreter uses a simple multiplicative algorithm:

$$(DN_{TMI}) (DN_{SPOT}) = DN_{new TMI}$$

The algorithm is derived from the four component technique of Crippen (Crippen 1989). In this paper, it is argued that of the four possible arithmetic methods to incorporate an intensity image into a chromatic image (addition, subtraction, division, and multiplication), only multiplication is unlikely to distort the color.

However, in his study Crippen first removed the intensity component via band ratios, spectral indices, or PC transform. The algorithm shown above operates on the original image. The result is an increased presence of the intensity component. For many applications, this is desirable. People involved in urban or suburban studies, city planning, and utilities routing often want roads and cultural features (which tend toward high reflection) to be pronounced in the image.

Brovoy Transform

In the Brovov Transform method, three bands are used according to the following formula:

$$[DN_{B1} / (DN_{B1} + DN_{B2} + DN_{B3})] \times [DN_{high\ res.\ image}] = DN_{B1_new}$$

$$[DN_{B2} / (DN_{B1} + DN_{B2} + DN_{B3})] \times [DN_{high\ res.\ image}] = DN_{B2_new}$$

$$[DN_{B3} / (DN_{B1} + DN_{B2} + DN_{B3})] \times [DN_{high\ res.\ image}] = DN_{B3_new}$$

Where:

B = band

The Brovov Transform was developed to visually increase contrast in the low and high ends of an image's histogram (i.e., to provide contrast in shadows, water and high reflectance areas such as urban features). Consequently, the Brovov Transform should not be used if preserving the original scene radiometry is important. However, it is good for producing RGB images with a higher degree of contrast in the low and high ends of the image histogram and for producing visually appealing images.

Since the Brovey Transform is intended to produce RGB images, only three bands at a time should be merged from the input multispectral scene, such as bands 3, 2, 1 from a SPOT or Landsat TM image or 4, 3, 2 from a Landsat TM image. The resulting merged image should then be displayed with bands 1, 2, 3 to RGB.

Adaptive Filter

Contrast enhancement (image stretching) is a widely applicable standard image processing technique. However, even adjustable stretches like the piecewise linear stretch act on the scene globally. There are many circumstances where this is not the optimum approach. For example, coastal studies where much of the water detail is spread through a very low DN range and the land detail is spread through a much higher DN range would be such a circumstance. In these cases, a filter that adapts the stretch to the region of interest (the area within the moving window) would produce a better enhancement. Adaptive filters attempt to achieve this (Fahnestock and Schowengerdt 1983, Peli and Lim 1982, Schwartz 1977).



ERDAS IMAGINE supplies two adaptive filters with user-adjustable parameters. The Adaptive Filter function in Image Interpreter can be applied to undegraded images, such as SPOT, Landsat, and digitized photographs. The Image Enhancement function in IMAGINE Radar Interpreter is better for degraded or difficult images.

Scenes to be adaptively filtered can be divided into three broad and overlapping categories:

- Undegraded—these scenes have good and uniform illumination overall. Given a choice, these are the scenes one would prefer to obtain from imagery sources such as Space Imaging or SPOT.
- Low luminance—these scenes have an overall or regional less than optimum intensity. An underexposed photograph (scanned) or shadowed areas would be in this category. These scenes need an increase in both contrast and overall scene luminance.
- High luminance—these scenes are characterized by overall excessively high DN values. Examples of such circumstances would be an over-exposed (scanned) photograph or a scene with a light cloud cover or haze. These scenes need a decrease in luminance and an increase in contrast.

No single filter with fixed parameters can address this wide variety of conditions. In addition, multiband images may require different parameters for each band. Without the use of adaptive filters, the different bands would have to be separated into one-band files, enhanced, and then recombined.

For this function, the image is separated into high and low frequency component images. The low frequency image is considered to be overall scene luminance. These two component parts are then recombined in various relative amounts using multipliers derived from LUTs. These LUTs are driven by the overall scene luminance:

$$DN_{out} = K(DN_{Hi}) + DN_{LL}$$

Where:

K = user-selected contrast multiplier

H_i = high luminance (derives from the LUT)

LL = local luminance (derives from the LUT)

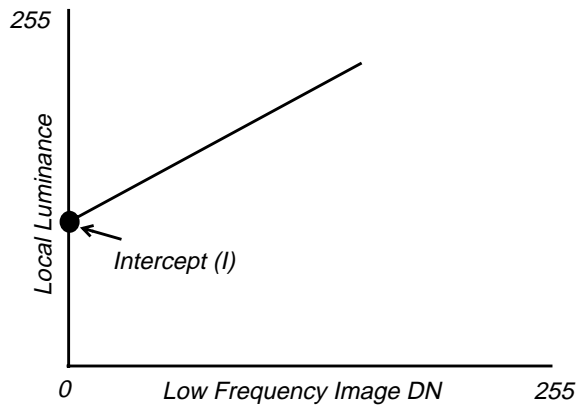


Figure 5-14 Local Luminance Intercept

Figure 5-14 shows the local luminance intercept, which is the output luminance value that an input luminance value of 0 would be assigned.

Spectral Enhancement

The enhancement techniques that follow require more than one band of data. They can be used to:

- compress bands of data that are similar
- extract new bands of data that are more interpretable to the eye
- apply mathematical transforms and algorithms
- display a wider variety of information in the three available color guns (R,G,B)



In this documentation, some examples are illustrated with two-dimensional graphs. However, you are not limited to two-dimensional (two-band) data. ERDAS IMAGINE programs allow an unlimited number of bands to be used.

Keep in mind that processing such data sets can require a large amount of computer swap space. In practice, the principles outlined below apply to any number of bands.



Some of these enhancements can be used to prepare data for classification. However, this is a risky practice unless you are very familiar with your data and the changes that you are making to it. Anytime you alter values, you risk losing some information.

Principal Components Analysis

Principal components analysis (PCA) is often used as a method of data compression. It allows redundant data to be compacted into fewer bands—that is, the dimensionality of the data is reduced. The bands of PCA data are noncorrelated and independent, and are often more interpretable than the source data (Jensen 1996; Faust 1989).

The process is easily explained graphically with an example of data in two bands. Below is an example of a two-band scatterplot, which shows the relationships of data file values in two bands. The values of one band are plotted against those of the other. If both bands have normal distributions, an ellipse shape results.



Scatterplots and normal distributions are discussed in "Appendix A: Math Topics."

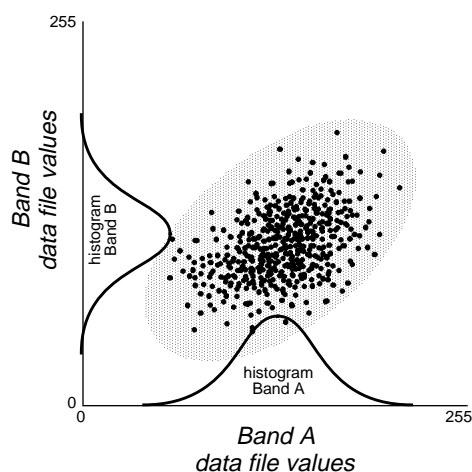


Figure 5-15 Two Band Scatterplot

Ellipse Diagram

In an n -dimensional histogram, an ellipse (2 dimensions), ellipsoid (3 dimensions), or hyperellipsoid (more than 3 dimensions) is formed if the distributions of each input band are normal or near normal. (The term ellipse is used for general purposes here.)

To perform PCA, the axes of the spectral space are rotated, changing the coordinates of each pixel in spectral space, as well as the data file values. The new axes are parallel to the axes of the ellipse.

First Principal Component

The length and direction of the widest transect of the ellipse are calculated using matrix algebra in a process explained below. The transect, which corresponds to the major (longest) axis of the ellipse, is called the first principal component of the data. The direction of the first principal component is the first eigenvector, and its length is the first eigenvalue (Taylor 1977).

A new axis of the spectral space is defined by this first principal component. The points in the scatterplot are now given new coordinates, which correspond to this new axis. Since, in spectral space, the coordinates of the points are the data file values, new data file values are derived from this process. These values are stored in the first principal component band of a new data file.

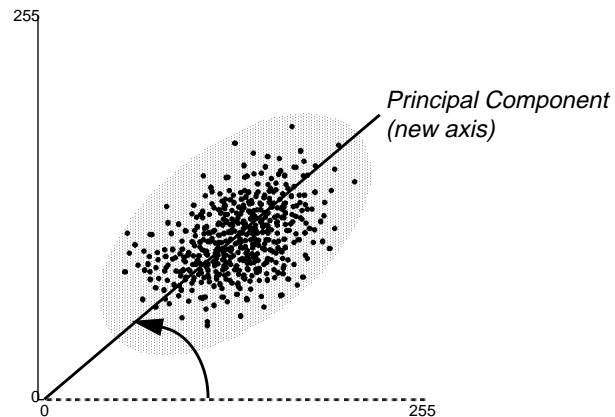


Figure 5-16 First Principal Component

The first principal component shows the direction and length of the widest transect of the ellipse. Therefore, as an axis in spectral space, it measures the highest variation within the data. In Figure 5-17 it is easy to see that the first eigenvalue is always greater than the ranges of the input bands, just as the hypotenuse of a right triangle must always be longer than the legs.

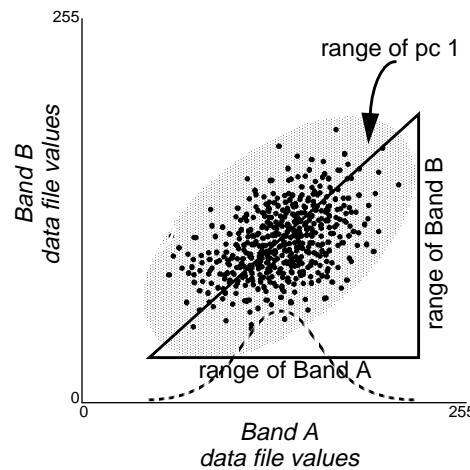


Figure 5-17 Range of First Principal Component

Successive Principal Components

The second principal component is the widest transect of the ellipse that is orthogonal (perpendicular) to the first principal component. As such, the second principal component describes the largest amount of variance in the data that is not already described by the first principal component (Taylor 1977). In a two-dimensional analysis, the second principal component corresponds to the minor axis of the ellipse.

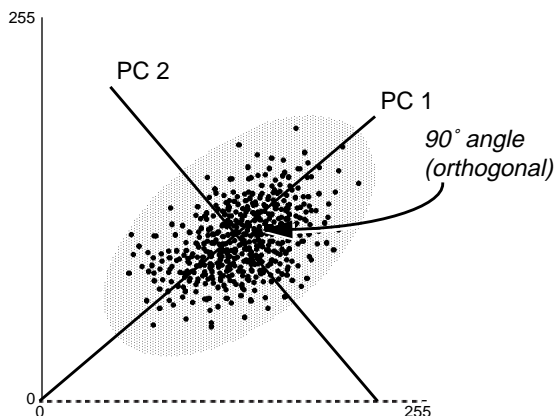


Figure 5-18 Second Principal Component

In n dimensions, there are n principal components. Each successive principal component:

- is the widest transect of the ellipse that is orthogonal to the previous components in the n -dimensional space of the scatterplot (Faust 1989), and
- accounts for a decreasing amount of the variation in the data which is not already accounted for by previous principal components (Taylor 1977).

Although there are n output bands in a PCA, the first few bands account for a high proportion of the variance in the data—in some cases, almost 100%. Therefore, PCA is useful for compressing data into fewer bands.

In other applications, useful information can be gathered from the principal component bands with the least variance. These bands can show subtle details in the image that were obscured by higher contrast in the original image. These bands may also show regular noise in the data (for example, the striping in old MSS data) (Faust 1989).

Computing Principal Components

To compute a principal components transformation, a linear transformation is performed on the data—meaning that the coordinates of each pixel in spectral space (the original data file values) are recomputed using a linear equation. The result of the transformation is that the axes in n -dimensional spectral space are shifted and rotated to be relative to the axes of the ellipse.

To perform the linear transformation, the eigenvectors and eigenvalues of the n principal components must be mathematically derived from the covariance matrix, as shown in the following equation:

$$V = \begin{bmatrix} v_1 & 0 & 0 & \dots & 0 \\ 0 & v_2 & 0 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & v_n \end{bmatrix}$$

$$E \text{ Cov } E^T = V$$

Where:

Cov = the covariance matrix

E = the matrix of eigenvectors

T = the transposition function

V = a diagonal matrix of eigenvalues, in which all nondiagonal elements are zeros

V is computed so that its nonzero elements are ordered from greatest to least, so that $v_1 > v_2 > v_3 \dots > v_n$.

Source: Faust 1989



A full explanation of this computation can be found in Gonzalez and Wintz 1977.

The matrix V is the covariance matrix of the output principal component file. The zeros represent the covariance between bands (there is none), and the eigenvalues are the variance values for each band. Because the eigenvalues are ordered from v_1 to v_n , the first eigenvalue is the largest and represents the most variance in the data.

Each column of the resulting eigenvector matrix, E , describes a unit-length vector in spectral space, which shows the direction of the principal component (the ellipse axis). The numbers are used as coefficients in the following equation, to transform the original data file values into the principal component values.

$$P_e = \sum_{k=1}^n d_k E_{ke}$$

Enhancement

Where:

e = the number of the principal component (first, second)

P_e = the output principal component value for principal component band e

k = a particular input band

n = the total number of bands

d_k = an input data file value in band k

E = the eigenvector matrix, such that E_{ke} = the element of that matrix at row k , column e

Source: Modified from Gonzalez and Wintz 1977

Decorrelation Stretch

The purpose of a contrast stretch is to:

- alter the distribution of the image DN values within the 0 - 255 range of the display device, and
- utilize the full range of values in a linear fashion.

The decorrelation stretch stretches the principal components of an image, not to the original image.

A principal components transform converts a multiband image into a set of mutually orthogonal images portraying inter-band variance. Depending on the DN ranges and the variance of the individual input bands, these new images (PCs) occupy only a portion of the possible 0 - 255 data range.

Each PC is separately stretched to fully utilize the data range. The new stretched PC composite image is then retransformed to the original data areas.

Either the original PCs or the stretched PCs may be saved as a permanent image file for viewing after the stretch.

NOTE: Storage of PCs as floating point, single precision is probably appropriate in this case.

Tasseled Cap

The different bands in a multispectral image can be visualized as defining an N-dimensional space where N is the number of bands. Each pixel, positioned according to its DN value in each band, lies within the N-dimensional space. This pixel distribution is determined by the absorption/reflection spectra of the imaged material. This clustering of the pixels is termed the data structure (Crist & Kauth 1986).



See "CHAPTER 1: Raster Data" for more information on absorption/reflection spectra. See the discussion on "Principal Components Analysis" on page 164.

The data structure can be considered a multidimensional hyperellipsoid. The principal axes of this data structure are not necessarily aligned with the axes of the data space (defined as the bands of the input image). They are more directly related to the absorption spectra. For viewing purposes, it is advantageous to rotate the N-dimensional space such that one or two of the data structure axes are aligned with the Viewer X and Y axes. In particular, you could view the axes that are largest for the data structure produced by the absorption peaks of special interest for the application.

For example, a geologist and a botanist are interested in different absorption features. They would want to view different data structures and therefore, different data structure axes. Both would benefit from viewing the data in a way that would maximize visibility of the data structure of interest.

The Tasseled Cap transformation offers a way to optimize data viewing for vegetation studies. Research has produced three data structure axes that define the vegetation information content (Crist et al 1986, Crist & Kauth 1986):

- Brightness—a weighted sum of all bands, defined in the direction of the principal variation in soil reflectance.
- Greenness—orthogonal to brightness, a contrast between the near-infrared and visible bands. Strongly related to the amount of green vegetation in the scene.
- Wetness—relates to canopy and soil moisture (Lillesand and Kiefer 1987).

A simple calculation (linear combination) then rotates the data space to present any of these axes to you.

These rotations are sensor-dependent, but once defined for a particular sensor (say Landsat 4 TM), the same rotation works for any scene taken by that sensor. The increased dimensionality (number of bands) of TM vs. MSS allowed Crist *et al* to define three additional axes, termed Haze, Fifth, and Sixth. Lavreau (1991) has used this haze parameter to devise an algorithm to dehaze Landsat imagery.

The Tasseled Cap algorithm implemented in the Image Interpreter provides the correct coefficient for MSS, TM4, and TM5 imagery. For TM4, the calculations are:

$$\text{Brightness} = .3037(TM1) + .2793(TM2) + .4743(TM3) + .5585(TM4) + .5082(TM5) + .1863(TM7)$$

$$\text{Greenness} = -.2848(TM1) - .2435(TM2) - .5436(TM3) + .7243(TM4) + .0840(TM5) - .1800(TM7)$$

$$\text{Wetness} = .1509(TM1) + .1973(TM2) + .3279(TM3) + .3406(TM4) - .7112(TM5) - .4572(TM7)$$

$$\text{Haze} = .8832(TM1) - .0819(TM2) - .4580(TM3) - .0032(TM4) - .0563(TM5) + .0130(TM7)$$

Source: Modified from Crist et al 1986, Jensen 1996

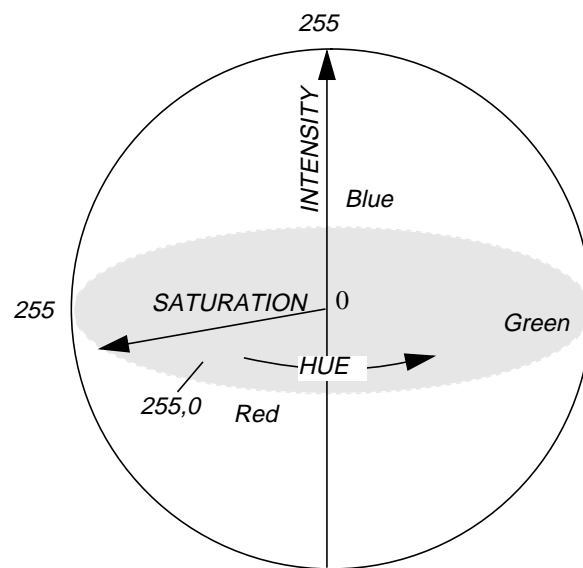
RGB to IHS

The color monitors used for image display on image processing systems have three color guns. These correspond to red, green, and blue (R,G,B), the additive primary colors. When displaying three bands of a multiband data set, the viewed image is said to be in R,G,B space.

Enhancement

However, it is possible to define an alternate color space that uses intensity (I), hue (H), and saturation (S) as the three positioned parameters (in lieu of R,G, and B). This system is advantageous in that it presents colors more nearly as perceived by the human eye.

- Intensity is the overall brightness of the scene (like PC-1) and varies from 0 (black) to 1 (white).
- Saturation represents the purity of color and also varies linearly from 0 to 1.
- Hue is representative of the color or dominant wavelength of the pixel. It varies from 0 at the red midpoint through green and blue back to the red midpoint at 360. It is a circular dimension (see Figure 5-19). In Figure 5-19, 0 to 255 is the selected range; it could be defined as any data range. However, hue must vary from 0 to 360 to define the entire sphere (Buchanan 1979).



Source: Buchanan 1979

Figure 5-19 Intensity, Hue, and Saturation Color Coordinate System



To use the RGB to IHS transform, use the RGB to IHS function from Image Interpreter.

The algorithm used in the Image Interpreter RGB to IHS transform is (Conrac 1980):

$$R = \frac{M - r}{M - m}$$

$$G = \frac{M - g}{M - m}$$

$$B = \frac{M - b}{M - m}$$

Where:

R, G, B are each in the range of 0 to 1.0.

r, g, b are each in the range of 0 to 1.0.

$M =$ largest value, $r, g, \text{ or } b$

$m =$ least value, $r, g, \text{ or } b$

NOTE: At least one of the $R, G, \text{ or } B$ values is 0, corresponding to the color with the largest value, and at least one of the $R, G, \text{ or } B$ values is 1, corresponding to the color with the least value.

The equation for calculating intensity in the range of 0 to 1.0 is:

$$I = \frac{M + m}{2}$$

The equations for calculating saturation in the range of 0 to 1.0 are:

$$\text{If } M = m, S = 0$$

$$\text{If } I \leq 0.5, S = \frac{M - m}{M + m}$$

$$\text{If } I > 0.5, S = \frac{M - m}{2 - M - m}$$

The equations for calculating hue in the range of 0 to 360 are:

$$\text{If } M = m, H = 0$$

$$\text{If } R = M, H = 60(2 + b - g)$$

$$\text{If } G = M, H = 60(4 + r - b)$$

$$\text{If } B = M, H = 60(6 + g - r)$$

Where:

R, G, B are each in the range of 0 to 1.0.

$M =$ largest value, $R, G,$ or B

$m =$ least value, $R, G,$ or B

IHS to RGB

The family of IHS to RGB is intended as a complement to the standard RGB to IHS transform.

In the IHS to RGB algorithm, a min-max stretch is applied to either intensity (I), saturation (S), or both, so that they more fully utilize the 0 to 1 value range. The values for hue (H), a circular dimension, are 0 to 360. However, depending on the dynamic range of the DN values of the input image, it is possible that I or S or both occupy only a part of the 0 to 1 range. In this model, a min-max stretch is applied to either I, S, or both, so that they more fully utilize the 0 to 1 value range. After stretching, the full IHS image is retransformed back to the original RGB space. As the parameter Hue is not modified, it largely defines what we perceive as color, and the resultant image looks very much like the input image.

It is not essential that the input parameters (IHS) to this transform be derived from an RGB to IHS transform. You could define I and/or S as other parameters, set Hue at 0 to 360, and then transform to RGB space. This is a method of color coding other data sets.

In another approach (Daily 1983), H and I are replaced by low- and high-frequency radar imagery. You can also replace I with radar intensity before the IHS to RGB transform (Holcomb 1993). Chavez evaluates the use of the IHS to RGB transform to resolution merge Landsat TM with SPOT panchromatic imagery (Chavez 1991).

NOTE: Use the Spatial Modeler for this analysis.



See the previous section on RGB to IHS transform for more information.

The algorithm used by ERDAS IMAGINE for the IHS to RGB function is (Conrac 1980):

Given: H in the range of 0 to 360; I and S in the range of 0 to 1.0

If $I \leq 0.5$, $M = I(1 + S)$

If $I > 0.5$, $M = I + S - I(S)$

*$m = 2 * I - M$*

The equations for calculating R in the range of 0 to 1.0 are:

$$\text{If } H < 60, R = m + (M - m) \left(\frac{H}{60} \right)$$

$$\text{If } 60 \leq H < 180, R = M$$

$$\text{If } 180 \leq H < 240, R = m + (M - m) \left(\frac{240 - H}{60} \right)$$

$$\text{If } 240 \leq H \leq 360, R = m$$

The equations for calculating G in the range of 0 to 1.0 are:

$$\text{If } H < 120, G = m$$

$$\text{If } 120 \leq H < 180, G = m + (M - m) \left(\frac{H - 120}{60} \right)$$

$$\text{If } 180 \leq H < 300, G = M$$

$$\text{If } 300 \leq H \leq 360, G = m + (M - m) \left(\frac{360 - H}{60} \right)$$

Equations for calculating B in the range of 0 to 1.0:

$$\text{If } H < 60, B = M$$

$$\text{If } 60 \leq H < 120, B = m + (M - m) \left(\frac{120 - H}{60} \right)$$

$$\text{If } 120 \leq H < 240, B = m$$

$$\text{If } 240 \leq H < 300, B = m + (M - m) \left(\frac{H - 240}{60} \right)$$

$$\text{If } 300 \leq H \leq 360, B = M$$

Indices

Indices are used to create output images by mathematically combining the DN values of different bands. These may be simplistic:

$$(Band X - Band Y)$$

or more complex:

$$\frac{Band X - Band Y}{Band X + Band Y}$$

In many instances, these indices are ratios of band DN values:

$$\frac{Band X}{Band Y}$$

These ratio images are derived from the absorption/reflection spectra of the material of interest. The absorption is based on the molecular bonds in the (surface) material. Thus, the ratio often gives information on the chemical composition of the target.



See "CHAPTER 1: Raster Data" for more information on the absorption/reflection spectra.

Applications

- Indices are used extensively in mineral exploration and vegetation analysis to bring out small differences between various rock types and vegetation classes. In many cases, judiciously chosen indices can highlight and enhance differences that cannot be observed in the display of the original color bands.
- Indices can also be used to minimize shadow effects in satellite and aircraft multispectral images. Black and white images of individual indices or a color combination of three ratios may be generated.
- Certain combinations of TM ratios are routinely used by geologists for interpretation of Landsat imagery for mineral type. For example: Red 5/7, Green 5/4, Blue 3/1.

Integer Scaling Considerations

The output images obtained by applying indices are generally created in floating point to preserve all numerical precision. If there are two bands, A and B, then:

$$\text{ratio} = A/B$$

If $A \gg B$ (much greater than), then a normal integer scaling would be sufficient. If $A > B$ and A is never much greater than B, scaling might be a problem in that the data range might only go from 1 to 2 or from 1 to 3. In this case, integer scaling would give very little contrast.

For cases in which $A < B$ or $A \ll B$, integer scaling would always truncate to 0. All fractional data would be lost. A multiplication constant factor would also not be very effective in seeing the data contrast between 0 and 1, which may very well be a substantial part of the data image. One approach to handling the entire ratio range is to actually process the function:

$$\text{ratio} = \text{atan}(A/B)$$

This would give a better representation for $A/B < 1$ as well as for $A/B > 1$ (Faust 1992).

Index Examples

The following are examples of indices that have been preprogrammed in the Image Interpreter in ERDAS IMAGINE:

- IR/R (infrared/red)
- SQRT (IR/R)
- Vegetation Index = $IR-R$

- Normalized Difference Vegetation Index (NDVI) =

$$\frac{IR - R}{IR + R}$$

- Transformed NDVI (TNDVI) =

$$\sqrt{\frac{IR - R}{IR + R} + 0.5}$$

- Iron Oxide = TM 3/1
- Clay Minerals = TM 5/7
- Ferrous Minerals = TM 5/4
- Mineral Composite = TM 5/7, 5/4, 3/1
- Hydrothermal Composite = TM 5/7, 3/1, 4/3

Source: Modified from Sabins 1987; Jensen 1996; Tucker 1979

The following table shows the infrared (IR) and red (R) band for some common sensors (Tucker 1979, Jensen 1996):

Sensor	IR Band	R Band
Landsat MSS	7	5
SPOT XS	3	2
Landsat TM	4	3
NOAA AVHRR	2	1

Image Algebra

Image algebra is a general term used to describe operations that combine the pixels of two or more raster layers in mathematical combinations. For example, the calculation:

(infrared band) - (red band)

$$DN_{ir} - DN_{red}$$

yields a simple, yet very useful, measure of the presence of vegetation. At the other extreme is the Tasseled Cap calculation (described in the following pages), which uses a more complicated mathematical combination of as many as six bands to define vegetation.

Band ratios, such as:

$$\frac{TM5}{TM7} = \text{clay minerals}$$

are also commonly used. These are derived from the absorption spectra of the material of interest. The numerator is a baseline of background absorption and the denominator is an absorption peak.



See "CHAPTER 1: Raster Data" for more information on absorption/reflection spectra.

NDVI is a combination of addition, subtraction, and division:

$$NDVI = \frac{IR - R}{IR + R}$$

Hyperspectral Image Processing

Hyperspectral image processing is, in many respects, simply an extension of the techniques used for multispectral data sets; indeed, there is no set number of bands beyond which a data set is hyperspectral. Thus, many of the techniques or algorithms currently used for multispectral data sets are logically applicable, regardless of the number of bands in the data set (see the discussion of Figure 1-7 on page 12 of this manual). What is of relevance in evaluating these data sets is not the number of bands *per se*, but the spectral bandwidth of the bands (channels). As the bandwidths get smaller, it becomes possible to view the data set as an absorption spectrum rather than a collection of discontinuous bands. Analysis of the data in this fashion is termed imaging spectrometry.

A hyperspectral image data set is recognized as a three-dimensional pixel array. As in a traditional raster image, the x-axis is the column indicator and the y-axis is the row indicator. The z-axis is the band number or, more correctly, the wavelength of that band (channel). A hyperspectral image can be visualized as shown in Figure 5-20.

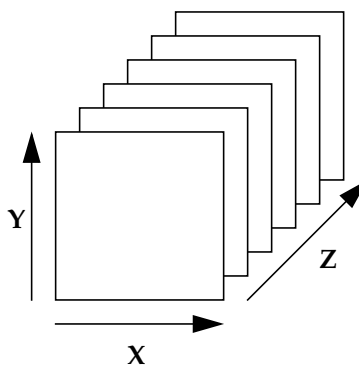


Figure 5-20 Hyperspectral Data Axes

A data set with narrow contiguous bands can be plotted as a continuous spectrum and compared to a library of known spectra using full profile spectral pattern fitting algorithms. A serious complication in using this approach is assuring that all spectra are corrected to the same background.

At present, it is possible to obtain spectral libraries of common materials. The JPL and USGS mineral spectra libraries are included in ERDAS IMAGINE. These are laboratory-measured reflectance spectra of reference minerals, often of high purity and defined particle size. The spectrometer is commonly purged with pure nitrogen to avoid absorbance by atmospheric gases. Conversely, the remote sensor records an image after the sunlight has passed through the atmosphere (twice) with variable and unknown amounts of water vapor, CO₂. (This atmospheric absorbance curve is shown in Figure 1-4.) The unknown atmospheric absorbances superimposed upon the Earth's surface reflectances makes comparison to laboratory spectra or spectra taken with a different atmosphere inexact. Indeed, it has been shown that atmospheric composition can vary within a single scene. This complicates the use of spectral signatures even within one scene. Atmospheric absorption and scattering is discussed on pages 6 through 10 of this manual.

A number of approaches have been advanced to help compensate for this atmospheric contamination of the spectra. These are introduced briefly on page 142 of this manual for the general case. Two specific techniques, Internal Average Relative Reflectance (IARR) and Log Residuals, are implemented in ERDAS IMAGINE. These have the advantage of not requiring auxiliary input information; the correction parameters are scene-derived. The disadvantage is that they produce relative reflectances (i.e., they can be compared to reference spectra in a semi-quantitative manner only).

Normalize

Pixel albedo is affected by sensor look angle and local topographic effects. For airborne sensors, this look angle effect can be large across a scene. It is less pronounced for satellite sensors. Some scanners look to both sides of the aircraft. For these data sets, the average scene luminance between the two half-scenes can be large. To help minimize these effects, an equal area normalization algorithm can be applied (Zamudio and Atkinson 1990). This calculation shifts each (pixel) spectrum to the same overall average brightness. This enhancement must be used with a consideration of whether this assumption is valid for the scene. For an image that contains two (or more) distinctly different regions (e.g., half ocean and half forest), this may not be a valid assumption. Correctly applied, this normalization algorithm helps remove albedo variations and topographic effects.

IAR Reflectance

As discussed above, it is desired to convert the spectra recorded by the sensor into a form that can be compared to known reference spectra. This technique calculates a relative reflectance by dividing each spectrum (pixel) by the scene average spectrum (Kruse 1988). The algorithm is based on the assumption that this scene average spectrum is largely composed of the atmospheric contribution and that the atmosphere is uniform across the scene. However, these assumptions are not always valid. In particular, the average spectrum could contain absorption features related to target materials of interest. The algorithm could then overcompensate for (i.e., remove) these absorbance features. The average spectrum should be visually inspected to check for this possibility. Properly applied, this technique can remove the majority of atmospheric effects.

Enhancement

Log Residuals

The Log Residuals technique was originally described by Green and Craig (1985), but has been variously modified by researchers. The version implemented here is similar to the approach of Lyon (1987). The algorithm can be conceptualized as:

$$\text{Output Spectrum} = (\text{input spectrum}) - (\text{average spectrum}) - (\text{pixel brightness}) + (\text{image brightness})$$

All parameters in the above equation are in logarithmic space, hence the name.

This algorithm corrects the image for atmospheric absorption, systemic instrumental variation, and illuminance differences between pixels.

Rescale

Many hyperspectral scanners record the data in a format larger than 8-bit. In addition, many of the calculations used to correct the data are performed with a floating point format to preserve precision. At some point, it is advantageous to compress the data back into an 8-bit range for effective storage and/or display. However, when rescaling data to be used for imaging spectrometry analysis, it is necessary to consider all data values within the data cube, not just within the layer of interest. This algorithm is designed to maintain the 3-dimensional integrity of the data values. Any bit format can be input. The output image is always 8-bit.

When rescaling a data cube, a decision must be made as to which bands to include in the rescaling. Clearly, a bad band (i.e., a low S/N layer) should be excluded. Some sensors image in different regions of the electromagnetic (EM) spectrum (e.g., reflective and thermal infrared or long- and short-wave reflective infrared). When rescaling these data sets, it may be appropriate to rescale each EM region separately. These can be input using the Select Layer option in the Viewer.

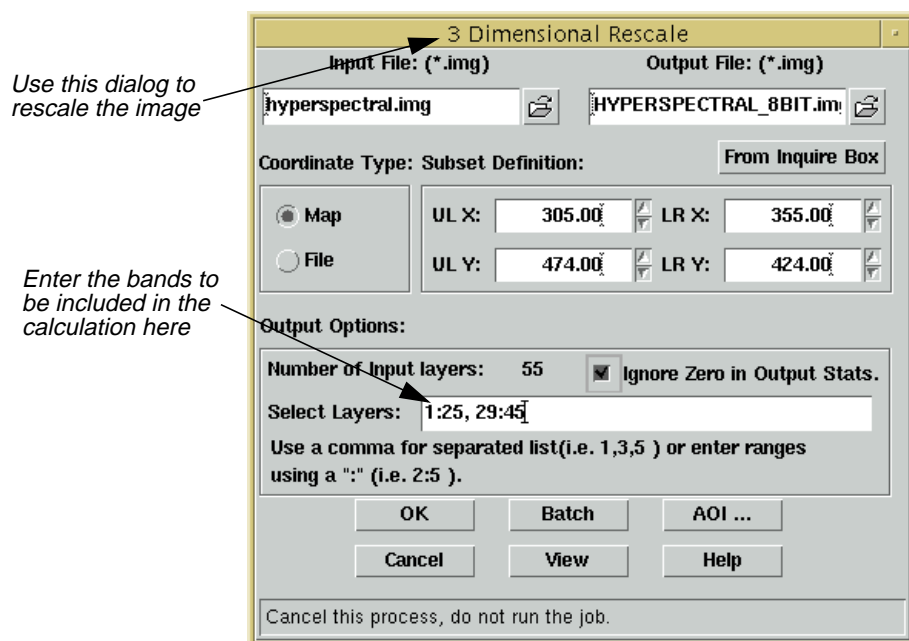


Figure 5-21 Rescale Graphical User Interface (GUI)

NOTE: Bands 26 through 28, and 46 through 55 have been deleted from the calculation. The deleted bands are still rescaled, but they are not factored into the rescale calculation.

Processing Sequence

The above (and other) processing steps are utilized to convert the raw image into a form that is easier to interpret. This interpretation often involves comparing the imagery, either visually or automatically, to laboratory spectra or other known end-member spectra. At present there is no widely accepted standard processing sequence to achieve this, although some have been advanced in the scientific literature (Zamudio and Atkinson 1990; Kruse 1988; Green and Craig 1985; Lyon 1987). Two common processing sequences have been programmed as single automatic enhancements, as follows:

- Automatic Relative Reflectance—Implements the following algorithms: Normalize, IAR Reflectance, and Rescale.
- Automatic Log Residuals—Implements the following algorithms: Normalize, Log Residuals, and Rescale.

Spectrum Average

In some instances, it may be desirable to average together several pixels. This is mentioned above under "IAR Reflectance" on page 177 as a test for applicability. In preparing reference spectra for classification, or to save in the Spectral Library, an average spectrum may be more representative than a single pixel. Note that to implement this function it is necessary to define which pixels to average using the AOI tools. This enables you to average any set of pixels that is defined; the pixels do not need to be contiguous and there is no limit on the number of pixels averaged. Note that the output from this program is a single pixel with the same number of input bands as the original image.

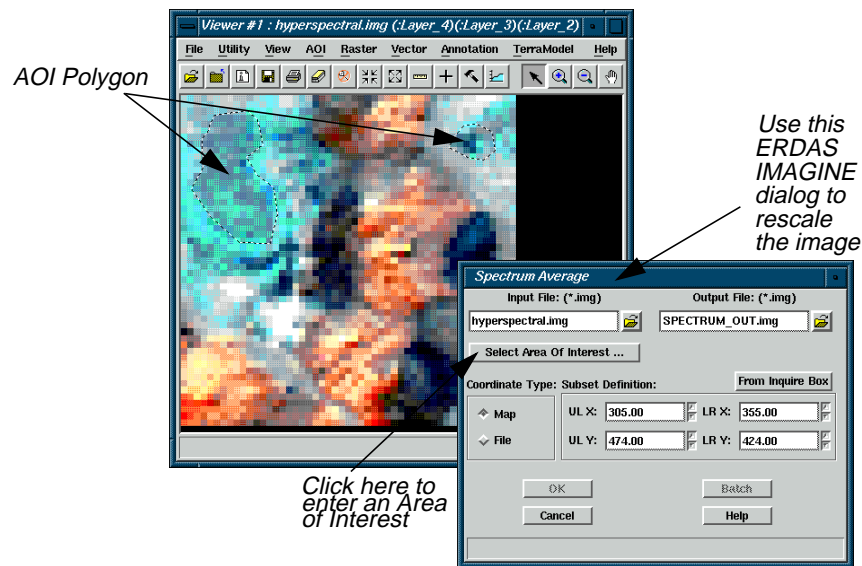


Figure 5-22 Spectrum Average GUI

Enhancement

Signal to Noise

The signal-to-noise (S/N) ratio is commonly used to evaluate the usefulness or validity of a particular band. In this implementation, S/N is defined as Mean/Std.Dev. in a 3×3 moving window. After running this function on a data set, each layer in the output image should be visually inspected to evaluate suitability for inclusion into the analysis. Layers deemed unacceptable can be excluded from the processing by using the Select Layers option of the various Graphical User Interfaces (GUIs). This can be used as a sensor evaluation tool.

Mean per Pixel

This algorithm outputs a single band, regardless of the number of input bands. By visually inspecting this output image, it is possible to see if particular pixels are outside the norm. While this does not mean that these pixels are incorrect, they should be evaluated in this context. For example, a CCD detector could have several sites (pixels) that are dead or have an anomalous response, these would be revealed in the mean-per-pixel image. This can be used as a sensor evaluation tool.

Profile Tools

To aid in visualizing this three-dimensional data cube, three basic tools have been designed:

- Spectral Profile—a display that plots the reflectance spectrum of a designated pixel, as shown in Figure 5-23.

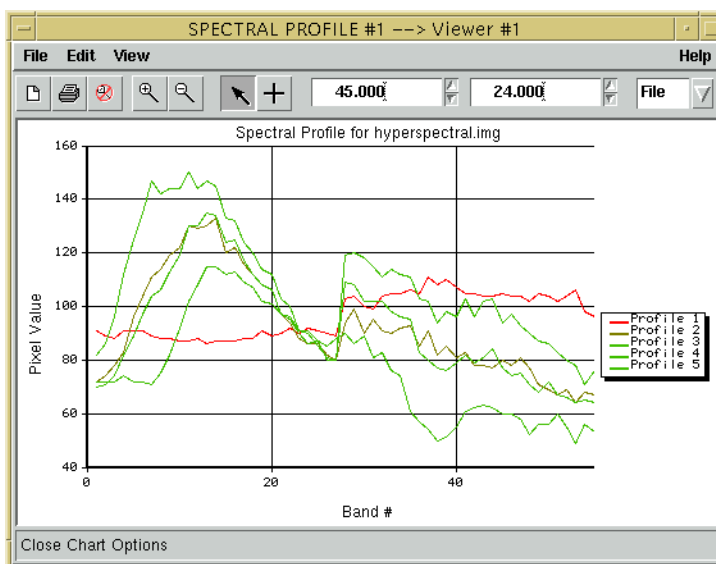


Figure 5-23 Spectral Profile

- Spatial Profile—a display that plots spectral information along a user-defined polyline. The data can be displayed two-dimensionally for a single band, as in Figure 5-24.

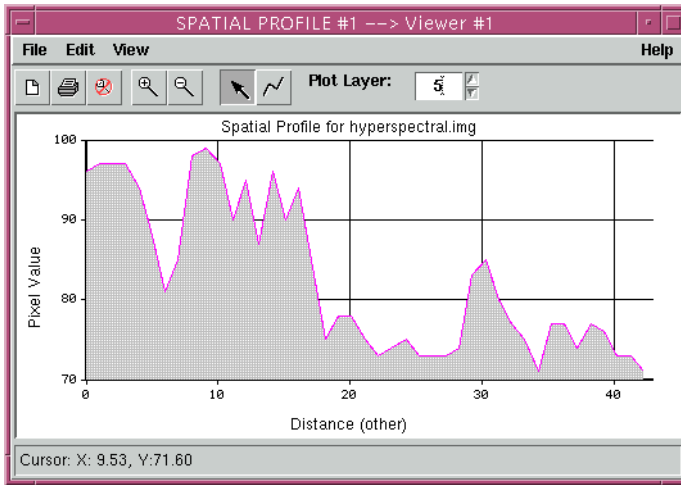


Figure 5-24 Two-Dimensional Spatial Profile

The data can also be displayed three-dimensionally for multiple bands, as in Figure 5-25.

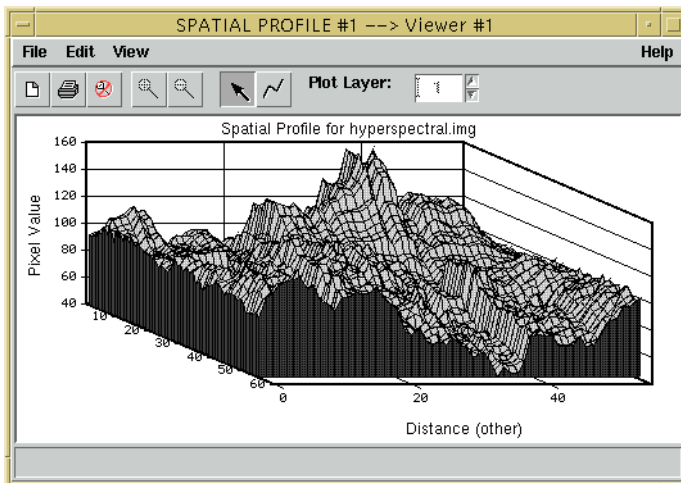


Figure 5-25 Three-Dimensional Spatial Profile

- **Surface Profile**—a display that allows you to designate an x,y area and view any selected layer, z.

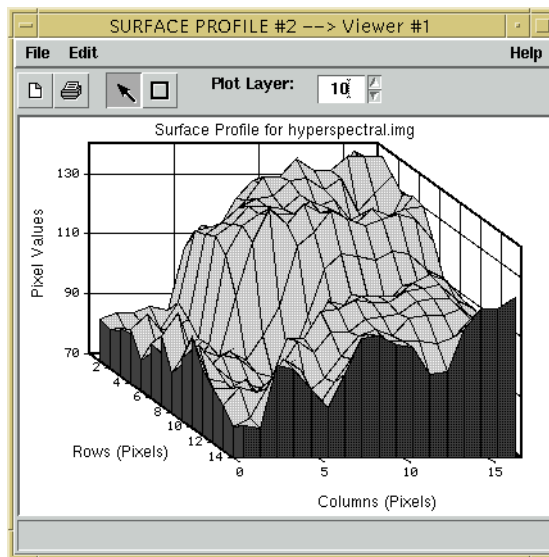


Figure 5-26 Surface Profile

Wavelength Axis

Data tapes containing hyperspectral imagery commonly designate the bands as a simple numerical sequence. When plotted using the profile tools, this yields an x-axis labeled as 1,2,3,4, etc. Elsewhere on the tape or in the accompanying documentation is a file that lists the center frequency and width of each band. This information should be linked to the image intensity values for accurate analysis or comparison to other spectra, such as the Spectra Libraries.

Spectral Library

As discussed on page 177, two spectral libraries are presently included in the software package (JPL and USGS). In addition, it is possible to extract spectra (pixels) from a data set or prepare average spectra from an image and save these in a user-derived spectral library. This library can then be used for visual comparison with other image spectra, or it can be used as input signatures in a classification.

Classification

The advent of data sets with very large numbers of bands has pressed the limits of the traditional classifiers such as Isodata, Maximum Likelihood, and Minimum Distance, but has not obviated their usefulness. Much research has been directed toward the use of Artificial Neural Networks (ANN) to more fully utilize the information content of hyperspectral images (Merenyi, Taranik, Monor, and Farrand 1996). To date, however, these advanced techniques have proven to be only marginally better at a considerable cost in complexity and computation. For certain applications, both Maximum Likelihood (Benediktsson, Swain, Ersoy, and Hong 1990) and Minimum Distance (Merenyi, Taranik, Monor, and Farrand 1996) have proven to be appropriate. "CHAPTER 6: Classification" contains a detailed discussion of these classification techniques.

A second category of classification techniques utilizes the imaging spectroscopy model for approaching hyperspectral data sets. This approach requires a library of possible end-member materials. These can be from laboratory measurements using a scanning spectrometer and reference standards (Clark, Gallagher, and Swayze 1990). The JPL and USGS libraries are compiled this way. The reference spectra (signatures) can also be scene-derived from either the scene under study or another similar scene (Adams, Smith, and Gillespie 1989).

System Requirements

Because of the large number of bands, a hyperspectral data set can be surprisingly large. For example, an AVIRIS scene is only 512×614 pixels in dimension, which seems small. However, when multiplied by 224 bands (channels) and 16 bits, it requires over 140 megabytes of data storage space. To process this scene requires corresponding large swap and temp space. In practice, it has been found that a 48 Mb memory board and 100 Mb of swap space is a minimum requirement for efficient processing. Temporary file space requirements depend upon the process being run.

Fourier Analysis

Image enhancement techniques can be divided into two basic categories: point and neighborhood. Point techniques enhance the pixel based only on its value, with no concern for the values of neighboring pixels. These techniques include contrast stretches (nonadaptive), classification, and level slices. Neighborhood techniques enhance a pixel based on the values of surrounding pixels. As a result, these techniques require the processing of a possibly large number of pixels for each output pixel. The most common way of implementing these enhancements is via a moving window convolution. However, as the size of the moving window increases, the number of requisite calculations becomes enormous. An enhancement that requires a convolution operation in the spatial domain can be implemented as a simple multiplication in frequency space—a much faster calculation.

In ERDAS IMAGINE, the FFT is used to convert a raster image from the spatial (normal) domain into a frequency domain image. The FFT calculation converts the image into a series of two-dimensional sine waves of various frequencies. The Fourier image itself cannot be easily viewed, but the magnitude of the image can be calculated, which can then be displayed either in the Viewer or in the FFT Editor. Analysts can edit the Fourier image to reduce noise or remove periodic features, such as striping. Once the Fourier image is edited, it is then transformed back into the spatial domain by using an IFFT. The result is an enhanced version of the original image.

This section focuses on the Fourier editing techniques available in the FFT Editor. Some rules and guidelines for using these tools are presented in this document. Also included are some examples of techniques that generally work for specific applications, such as striping.

NOTE: You may also want to refer to the works cited at the end of this section for more information.

The basic premise behind a Fourier transform is that any one-dimensional function, $f(x)$ (which might be a row of pixels), can be represented by a Fourier series consisting of some combination of sine and cosine terms and their associated coefficients. For example, a line of pixels with a high spatial frequency gray scale pattern might be represented in terms of a single coefficient multiplied by a $\sin(x)$ function. High spatial frequencies are those that represent frequent gray scale changes in a short pixel distance. Low spatial frequencies represent infrequent gray scale changes that occur gradually over a relatively large number of pixel distances. A more complicated function, $f(x)$, might have to be represented by many sine and cosine terms with their associated coefficients.

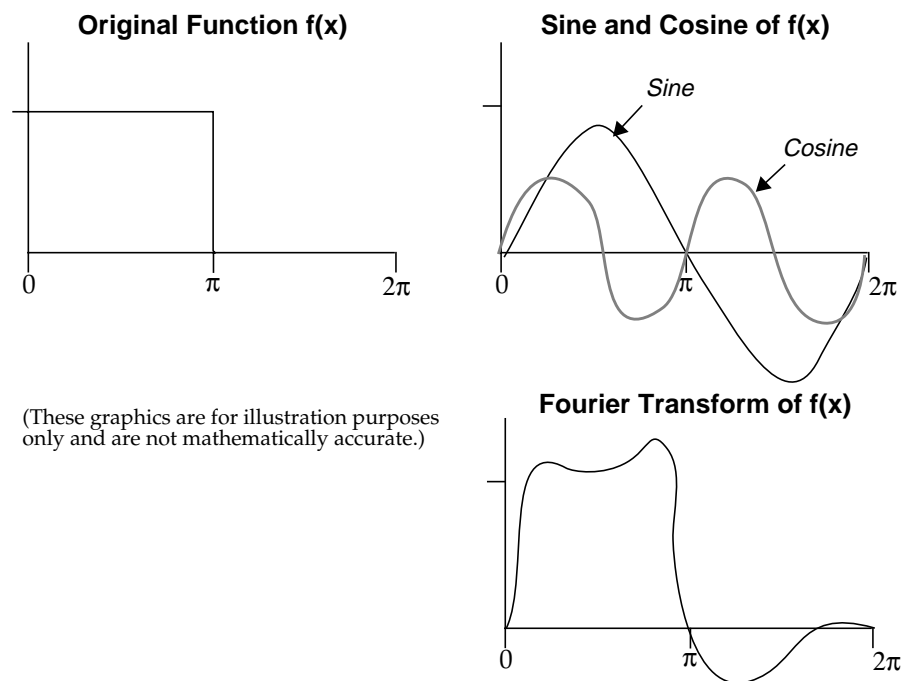


Figure 5-27 One-Dimensional Fourier Analysis

Figure 5-27 shows how a function $f(x)$ can be represented as a linear combination of sine and cosine. The Fourier transform of that same function is also shown.

A Fourier transform is a linear transformation that allows calculation of the coefficients necessary for the sine and cosine terms to adequately represent the image. This theory is used extensively in electronics and signal processing, where electrical signals are continuous and not discrete. Therefore, DFT has been developed. Because of the computational load in calculating the values for all the sine and cosine terms along with the coefficient multiplications, a highly efficient version of the DFT was developed and called the FFT.

To handle images which consist of many one-dimensional rows of pixels, a two-dimensional FFT has been devised that incrementally uses one-dimensional FFTs in each direction and then combines the result. These images are symmetrical about the origin.

Applications

Fourier transformations are typically used for the removal of noise such as striping, spots, or vibration in imagery by identifying periodicities (areas of high spatial frequency). Fourier editing can be used to remove regular errors in data such as those caused by sensor anomalies (e.g., striping). This analysis technique can also be used across bands as another form of pattern/feature recognition.

FFT

The FFT calculation is:

$$F(u, v) \leftarrow \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) e^{-j2\pi ux/M - j2\pi vy/N}]$$

Where:

- M = the number of pixels horizontally
- N = the number of pixels vertically
- u, v = spatial frequency variables
- $e \approx 2.71828$, the natural logarithm base
- $j \equiv$ the imaginary component of a complex number

The number of pixels horizontally and vertically must each be a power of two. If the dimensions of the input image are not a power of two, they are padded up to the next highest power of two. There is more information about this later in this section.

Source: Modified from Oppenheim 1975; Press 1988



Images computed by this algorithm are saved with an .fft file extension.



You should run a Fourier Magnitude transform on an .fft file before viewing it in the Viewer. The FFT Editor automatically displays the magnitude without further processing.

Fourier Magnitude

The raster image generated by the FFT calculation is not an optimum image for viewing or editing. Each pixel of a Fourier image is a complex number (i.e., it has two components: real and imaginary). For display as a single image, these components are combined in a root-sum of squares operation. Also, since the dynamic range of Fourier spectra vastly exceeds the range of a typical display device, the Fourier Magnitude calculation involves a logarithmic function.

Finally, a Fourier image is symmetric about the origin ($u, v = 0, 0$). If the origin is plotted at the upper left corner, the symmetry is more difficult to see than if the origin is at the center of the image. Therefore, in the Fourier magnitude image, the origin is shifted to the center of the raster array.

In this transformation, each .fft layer is processed twice. First, the maximum magnitude, $|X|_{max}$ is computed. Then, the following computation is performed for each FFT element magnitude x :

Enhancement

$$y(x) = 255.0 \ln \left[\left(\frac{|x|}{|x|_{max}} \right) (e - 1) + 1 \right]$$

Where:

- x = input FFT element
- y = the normalized log magnitude of the FFT element
- $|x|_{max}$ = the maximum magnitude
- $e \approx 2.71828$, the natural logarithm base
- $| |$ = the magnitude operator

This function was chosen so that y would be proportional to the logarithm of a linear function of x , with $y(0)=0$ and $y(|x|_{max}) = 255$.

Source: ERDAS

In Figure 5-28, Image A is one band of a badly striped Landsat TM scene. Image B is the Fourier Magnitude image derived from the Landsat image.

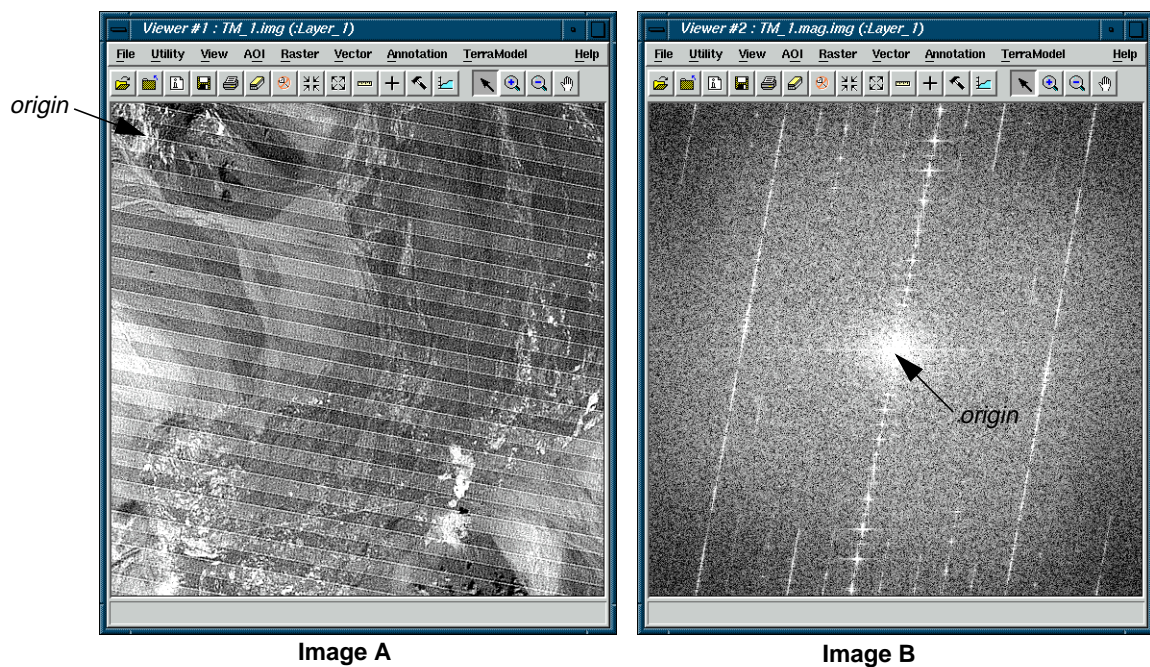


Figure 5-28 Example of Fourier Magnitude

Note that, although Image A has been transformed into Image B, these raster images are very different symmetrically. The origin of Image A is at $(x,y) = (0,0)$ in the upper left corner. In Image B, the origin $(u,v) = (0,0)$ is in the center of the raster. The low frequencies are plotted near this origin while the higher frequencies are plotted further out. Generally, the majority of the information in an image is in the low frequencies. This is indicated by the bright area at the center (origin) of the Fourier image.

It is important to realize that a position in a Fourier image, designated as (u,v) , does not always represent the same frequency, because it depends on the size of the input raster image. A large spatial domain image contains components of lower frequency than a small spatial domain image. As mentioned, these lower frequencies are plotted nearer to the center $(u,v = 0,0)$ of the Fourier image. Note that the units of spatial frequency are inverse length, e.g., m^{-1} .

The sampling increments in the spatial and frequency domain are related by:

$$\Delta u = \frac{1}{M\Delta x}$$

$$\Delta v = \frac{1}{N\Delta y}$$

Where:

M = horizontal image size in pixels

N = vertical image size in pixels

Δx = pixel size

Δy = pixel size

For example, converting a 512×512 Landsat TM image (pixel size = 28.5m) into a Fourier image:

$$\Delta u = \Delta v = \frac{1}{512 \times 28.5} = 6.85 \times 10^{-5} m^{-1}$$

u or v	Frequency
0	0
1	$6.85 \times 10^{-5} \times M^{-1}$
2	$13.7 \times 10^{-5} \times M^{-1}$

If the Landsat TM image was 1024×1024 :

$$\Delta u = \Delta v = \frac{1}{1024 \times 28.5} = 3.42 \times 10^{-5} m^{-1}$$

u or v	Frequency
0	0
1	3.42×10^{-5}
2	6.85×10^{-5}

Enhancement

So, as noted above, the frequency represented by a (u,v) position depends on the size of the input image.

For the above calculation, the sample images are 512×512 and 1024×1024 (powers of two). These were selected because the FFT calculation requires that the height and width of the input image be a power of two (although the image need not be square). In practice, input images usually do not meet this criterion. Three possible solutions are available in ERDAS IMAGINE:

- Subset the image.
- Pad the image—the input raster is increased in size to the next power of two by imbedding it in a field of the mean value of the entire input image.
- Resample the image so that its height and width are powers of two.

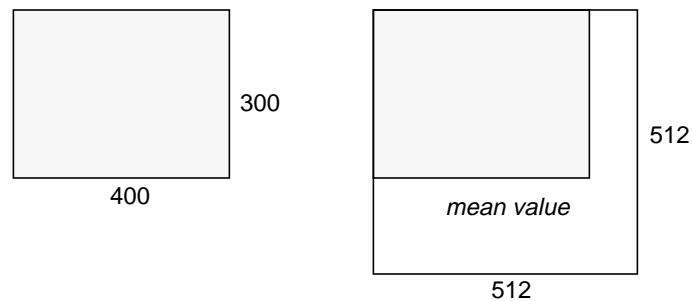


Figure 5-29 The Padding Technique

The padding technique is automatically performed by the FFT program. It produces a minimum of artifacts in the output Fourier image. If the image is subset using a power of two (i.e., 64×64 , 128×128 , 64×128), no padding is used.

IFFT

The IFFT computes the inverse two-dimensional FFT of the spectrum stored.

- The input file must be in the compressed .fft format described earlier (i.e., output from the FFT or FFT Editor).
- If the original image was padded by the FFT program, the padding is automatically removed by IFFT.
- This program creates (and deletes, upon normal termination) a temporary file large enough to contain one entire band of .fft data. The specific expression calculated by this program is:

$$f(x, y) \leftarrow \frac{1}{N_1 N_2} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} [F(u, v) e^{j2\pi ux/M + j2\pi vy/N}]$$

$$0 \leq x \leq M-1, 0 \leq y \leq N-1$$

Where:

M = the number of pixels horizontally
 N = the number of pixels vertically
 u, v = spatial frequency variables
 $e \approx 2.71828$, the natural logarithm base

Source: Modified from Oppenheim and Press 1988



Images computed by this algorithm are saved with an `.iffimg` file extension by default.

Filtering

Operations performed in the frequency (Fourier) domain can be visualized in the context of the familiar convolution function. The mathematical basis of this interrelationship is the convolution theorem, which states that a convolution operation in the spatial domain is equivalent to a multiplication operation in the frequency domain:

$$g(x,y) = h(x,y) * f(x,y) \equiv G(u,v) = H(u,v) \times F(u,v)$$

Where:

$f(x,y)$ = input image
 $h(x,y)$ = position invariant operation (convolution kernel)
 $g(x,y)$ = output image
 G, F, H = Fourier transforms of g, f, h

The names high-pass, low-pass, high-frequency indicate that these convolution functions derive from the frequency domain.

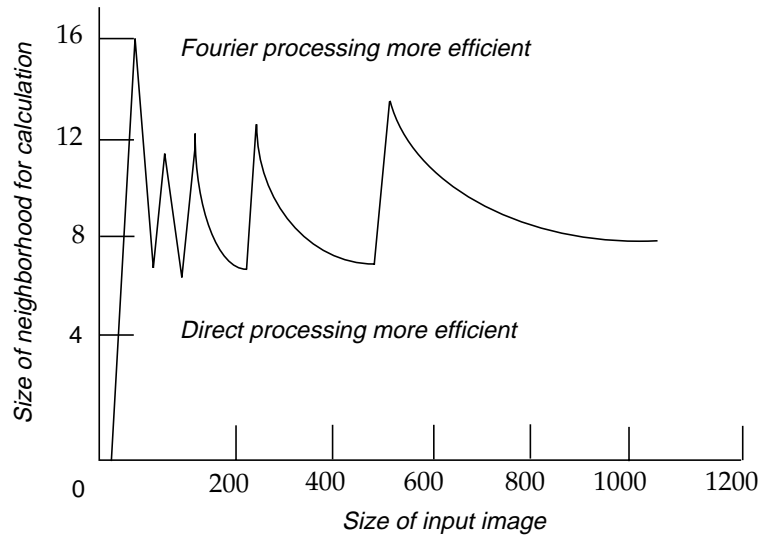
Low-Pass Filtering

The simplest example of this relationship is the low-pass kernel. The name, low-pass kernel, is derived from a filter that would pass low frequencies and block (filter out) high frequencies. In practice, this is easily achieved in the spatial domain by the $M = N = 3$ kernel:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Obviously, as the size of the image and, particularly, the size of the low-pass kernel increases, the calculation becomes more time-consuming. Depending on the size of the input image and the size of the kernel, it can be faster to generate a low-pass image via Fourier processing.

Figure 5-30 compares Direct and Fourier domain processing for finite area convolution.



Source: Pratt 1991

Figure 5-30 Comparison of Direct and Fourier Domain Processing

In the Fourier domain, the low-pass operation is implemented by attenuating the pixels' frequencies that satisfy:

$$u^2 + v^2 > D_0^2$$

D_0 is often called the cutoff frequency.

As mentioned, the low-pass information is concentrated toward the origin of the Fourier image. Thus, a smaller radius (r) has the same effect as a larger N (where N is the size of a kernel) in a spatial domain low-pass convolution.

As was pointed out earlier, the frequency represented by a particular u, v (or r) position depends on the size of the input image. Thus, a low-pass operation of $r = 20$ is equivalent to a spatial low-pass of various kernel sizes, depending on the size of the input image.

For example:

Image Size	Fourier Low-Pass $r =$	Convolution Low-Pass $N =$
64 × 64	50	3
	30	3.5
	20	5
	10	9
	5	14
128 × 128	20	13
	10	22
256 × 256	20	25
	10	42

This table shows that using a window on a 64 × 64 Fourier image with a radius of 50 as the cutoff is the same as using a 3 × 3 low-pass kernel on a 64 × 64 spatial domain image.

High-Pass Filtering

Just as images can be smoothed (blurred) by attenuating the high-frequency components of an image using low-pass filters, images can be sharpened and edge-enhanced by attenuating the low-frequency components using high-pass filters. In the Fourier domain, the high-pass operation is implemented by attenuating the pixels' frequencies that satisfy:

$$u^2 + v^2 < D_0^2$$

Windows

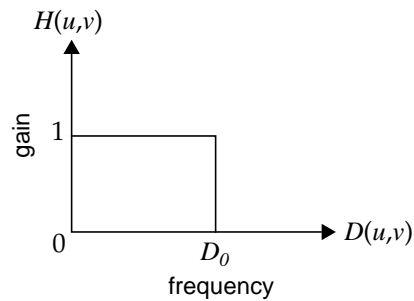
The attenuation discussed above can be done in many different ways. In ERDAS IMAGINE Fourier processing, five window functions are provided to achieve different types of attenuation:

- Ideal
- Bartlett (triangular)
- Butterworth
- Gaussian
- Hanning (cosine)

Each of these windows must be defined when a frequency domain process is used. This application is perhaps easiest understood in the context of the high-pass and low-pass filter operations. Each window is discussed in more detail:

Ideal

The simplest low-pass filtering is accomplished using the ideal window, so named because its cutoff point is absolute. Note that in Figure 5-31 the cross section is ideal.

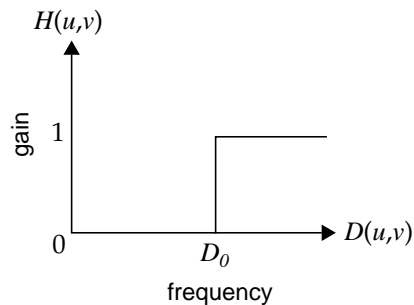


$$H(u,v) = 1 \text{ if } D(u,v) \leq D_0$$
$$0 \text{ if } D(u,v) > D_0$$

Figure 5-31 An Ideal Cross Section

All frequencies inside a circle of a radius D_0 are retained completely (passed), and all frequencies outside the radius are completely attenuated. The point D_0 is termed the cutoff frequency.

High-pass filtering using the ideal window looks like the following illustration:



$$H(u,v) = 0 \text{ if } D(u,v) \leq D_0$$
$$1 \text{ if } D(u,v) > D_0$$

Figure 5-32 High-Pass Filtering Using the Ideal Window

All frequencies inside a circle of a radius D_0 are completely attenuated, and all frequencies outside the radius are retained completely (passed).

A major disadvantage of the ideal filter is that it can cause ringing artifacts, particularly if the radius (r) is small. The smoother functions (e.g., Butterworth and Hanning) minimize this effect.

Bartlett

Filtering using the Bartlett window is a triangular function, as shown in the following low- and high-pass cross sections:

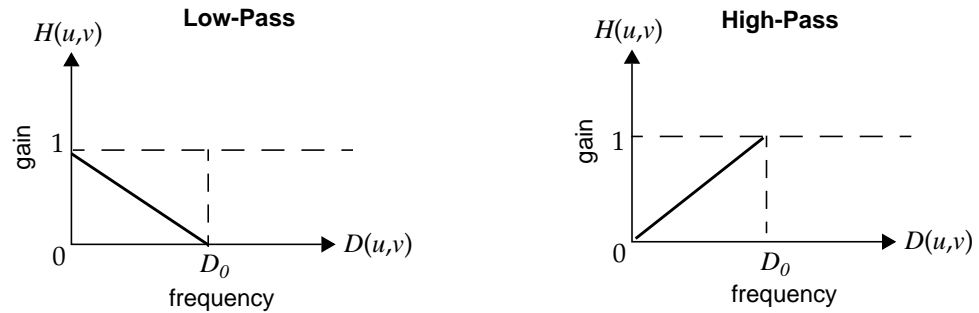


Figure 5-33 Filtering Using the Bartlett Window

Butterworth, Gaussian, and Hanning

The Butterworth, Gaussian, and Hanning windows are all smooth and greatly reduce the effect of ringing. The differences between them are minor and are of interest mainly to experts. For most normal types of Fourier image enhancement, they are essentially interchangeable.

The Butterworth window reduces the ringing effect because it does not contain abrupt changes in value or slope. The following low- and high-pass cross sections illustrate this:

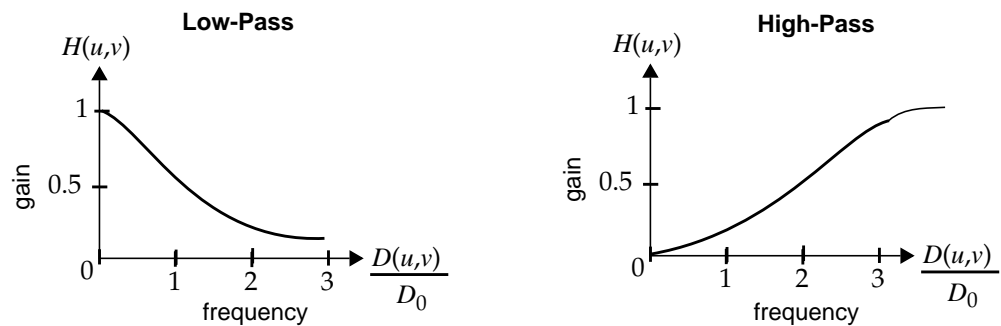


Figure 5-34 Filtering Using the Butterworth Window

The equation for the low-pass Butterworth window is:

$$H(u,v) = \frac{1}{1 + [(D(u, v))/D_0]^{2n}}$$

NOTE: The Butterworth window approaches its window center gain asymptotically.

Enhancement

The equation for the Gaussian low-pass window is:

$$H(u,v) = e^{-\left(\frac{x}{D_0}\right)^2}$$

The equation for the Hanning low-pass window is:

$$H(u,v) = \frac{1}{2} \left(1 + \cos\left(\frac{\pi x}{2D_0}\right) \right)$$

for $0 \leq x \leq 2D_0$

0 otherwise.

Fourier Noise Removal

Occasionally, images are corrupted by noise that is periodic in nature. An example of this is the scan lines that are present in some TM images. When these images are transformed into Fourier space, the periodic line pattern becomes a radial line. The Fourier Analysis functions provide two main tools for reducing noise in images:

- Editing
- Automatic removal of periodic noise

Editing

In practice, it has been found that radial lines centered at the Fourier origin ($u,v = 0,0$) are best removed using back-to-back wedges centered at $(0,0)$. It is possible to remove these lines using very narrow wedges with the Ideal window. However, the sudden transitions resulting from zeroing-out sections of a Fourier image causes a ringing of the image when it is transformed back into the spatial domain. This effect can be lessened by using a less abrupt window, such as Butterworth.

Other types of noise can produce artifacts, such as lines not centered at $u,v = 0,0$ or circular spots in the Fourier image. These can be removed using the tools provided in the FFT Editor. As these artifacts are always symmetrical in the Fourier magnitude image, editing tools operate on both components simultaneously. The FFT Editor contains tools that enable you to attenuate a circular or rectangular region anywhere on the image.

Automatic Periodic Noise Removal

The use of the FFT Editor, as described above, enables you to selectively and accurately remove periodic noise from any image. However, operator interaction and a bit of trial and error are required. The automatic periodic noise removal algorithm has been devised to address images degraded uniformly by striping or other periodic anomalies. Use of this algorithm requires a minimum of input from you.

The image is first divided into 128×128 pixel blocks. The Fourier Transform of each block is calculated and the log-magnitudes of each FFT block are averaged. The averaging removes all frequency domain quantities except those that are present in each block (i.e., some sort of periodic interference). The average power spectrum is then used as a filter to adjust the FFT of the entire image. When the IFFT is performed, the result is an image that should have any periodic noise eliminated or significantly reduced. This method is partially based on the algorithms outlined in Cannon *et al* 1983 and Srinivasan *et al* 1988.



Select the Periodic Noise Removal option from Image Interpreter to use this function.

Homomorphic Filtering

Homomorphic filtering is based upon the principle that an image may be modeled as the product of illumination and reflectance components:

$$I(x,y) = i(x,y) \times r(x,y)$$

Where:

$I(x,y)$ = image intensity (DN) at pixel x,y

$i(x,y)$ = illumination of pixel x,y

$r(x,y)$ = reflectance at pixel x,y

The illumination image is a function of lighting conditions and shadows. The reflectance image is a function of the object being imaged. A log function can be used to separate the two components (i and r) of the image:

$$\ln I(x,y) = \ln i(x,y) + \ln r(x,y)$$

This transforms the image from multiplicative to additive superposition. With the two component images separated, any linear operation can be performed. In this application, the image is now transformed into Fourier space. Because the illumination component usually dominates the low frequencies, while the reflectance component dominates the higher frequencies, the image may be effectively manipulated in the Fourier domain.

By using a filter on the Fourier image, which increases the high-frequency components, the reflectance image (related to the target material) may be enhanced, while the illumination image (related to the scene illumination) is de-emphasized.



Select the Homomorphic Filter option from Image Interpreter to use this function.

By applying an IFFT followed by an exponential function, the enhanced image is returned to the normal spatial domain. The flow chart in Figure 5-35 summarizes the homomorphic filtering process in ERDAS IMAGINE.

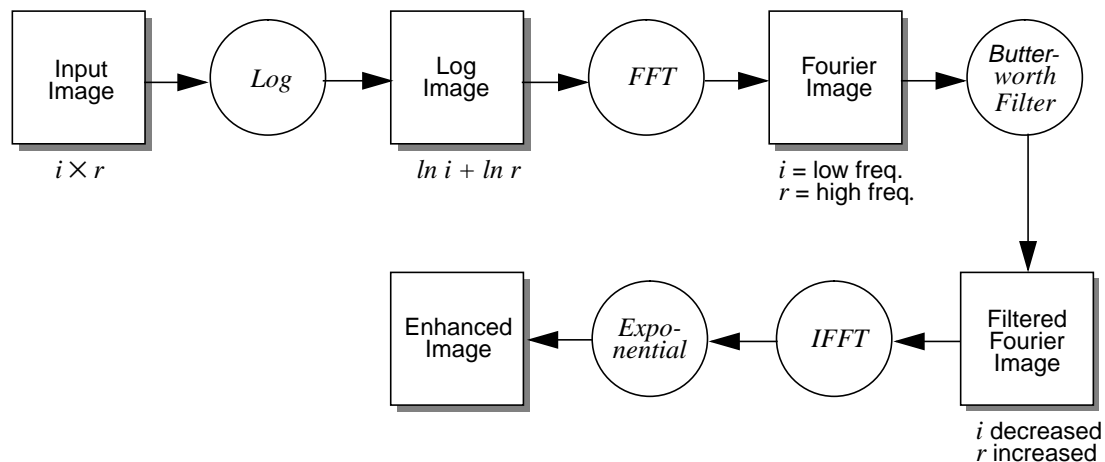


Figure 5-35 Homomorphic Filtering Process

As mentioned earlier, if an input image is not a power of two, the ERDAS IMAGINE Fourier analysis software automatically pads the image to the next largest size to make it a power of two. For manual editing, this causes no problems. However, in automatic processing, such as the homomorphic filter, the artifacts induced by the padding may have a deleterious effect on the output image. For this reason, it is recommended that images that are not a power of two be subset before being used in an automatic process.



A detailed description of the theory behind Fourier series and Fourier transforms is given in Gonzales and Wintz (1977). See also Oppenheim (1975) and Press (1988).

Radar Imagery Enhancement

The nature of the surface phenomena involved in radar imaging is inherently different from that of visible/infrared (VIS/IR) images. When VIS/IR radiation strikes a surface it is either absorbed, reflected, or transmitted. The absorption is based on the molecular bonds in the (surface) material. Thus, this imagery provides information on the chemical composition of the target.

When radar microwaves strike a surface, they are reflected according to the physical and electrical properties of the surface, rather than the chemical composition. The strength of radar return is affected by slope, roughness, and vegetation cover. The conductivity of a target area is related to the porosity of the soil and its water content. Consequently, radar and VIS/IR data are complementary; they provide different information about the target area. An image in which these two data types are intelligently combined can present much more information than either image by itself.



See "CHAPTER 1: Raster Data" and "CHAPTER 3: Raster and Vector Data Sources" for more information on radar data.

This section describes enhancement techniques that are particularly useful for radar imagery. While these techniques can be applied to other types of image data, this discussion focuses on the special requirements of radar imagery enhancement. The IMAGINE Radar Interpreter provides a sophisticated set of image processing tools designed specifically for use with radar imagery. This section describes the functions of the IMAGINE Radar Interpreter.



For information on the Radar Image Enhancement function, see the section on “Radiometric Enhancement” on page 143.

Speckle Noise

Speckle noise is commonly observed in radar (microwave or millimeter wave) sensing systems, although it may appear in any type of remotely sensed image utilizing coherent radiation. An active radar sensor gives off a burst of coherent radiation that reflects from the target, unlike a passive microwave sensor that simply receives the low-level radiation naturally emitted by targets.

Like the light from a laser, the waves emitted by active sensors travel in phase and interact minimally on their way to the target area. After interaction with the target area, these waves are no longer in phase. This is because of the different distances they travel from targets, or single versus multiple bounce scattering.

Once out of phase, radar waves can interact to produce light and dark pixels known as speckle noise. Speckle noise must be reduced before the data can be effectively utilized. However, the image processing programs used to reduce speckle noise produce changes in the image.



Because any image processing done before removal of the speckle results in the noise being incorporated into and degrading the image, you should not rectify, correct to ground range, or in any way resample, enhance, or classify the pixel values before removing speckle noise. Functions using Nearest Neighbor are technically permissible, but not advisable.

Since different applications and different sensors necessitate different speckle removal models, IMAGINE Radar Interpreter includes several speckle reduction algorithms:

- Mean filter
- Median filter
- Lee-Sigma filter
- Local Region filter
- Lee filter
- Frost filter
- Gamma-MAP filter

NOTE: Speckle noise in radar images cannot be completely removed. However, it can be reduced significantly.

These filters are described in the following sections:

Mean Filter

The Mean filter is a simple calculation. The pixel of interest (center of window) is replaced by the arithmetic average of all values within the window. This filter does not remove the aberrant (speckle) value; it averages it into the data.

In theory, a bright and a dark pixel within the same window would cancel each other out. This consideration would argue in favor of a large window size (e.g., 7×7). However, averaging results in a loss of detail, which argues for a small window size.

In general, this is the least satisfactory method of speckle reduction. It is useful for applications where loss of resolution is not a problem.

Median Filter

A better way to reduce speckle, but still simplistic, is the Median filter. This filter operates by arranging all DN values in sequential order within the window that you define. The pixel of interest is replaced by the value in the center of this distribution. A Median filter is useful for removing pulse or spike noise. Pulse functions of less than one-half of the moving window width are suppressed or eliminated. In addition, step functions or ramp functions are retained.

The effect of Mean and Median filters on various signals is shown (for one dimension) in Figure 5-36.

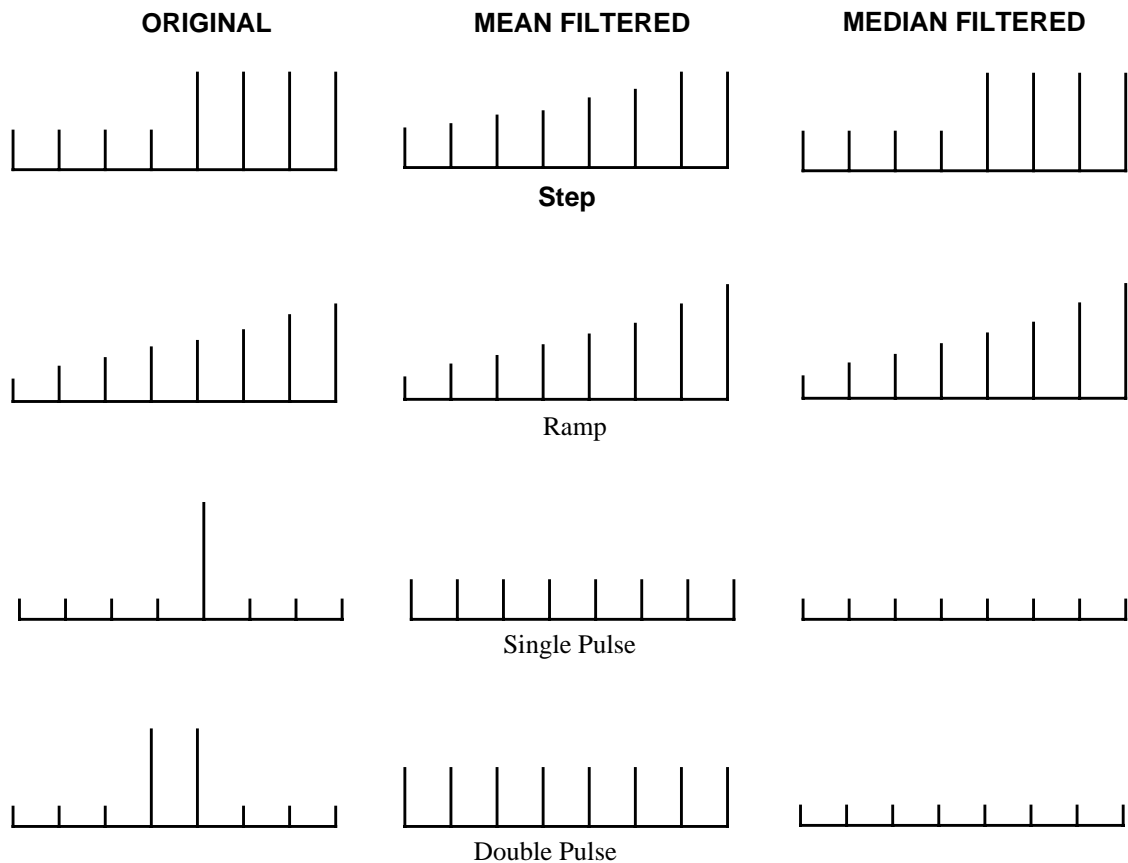


Figure 5-36 Effects of Mean and Median Filters

The Median filter is useful for noise suppression in any image. It does not affect step or ramp functions; it is an edge preserving filter (Pratt 1991). It is also applicable in removing pulse function noise, which results from the inherent pulsing of microwaves. An example of the application of the Median filter is the removal of dead-detector striping, as found in Landsat 4 TM data (Crippen 1989).

Local Region Filter

The Local Region filter divides the moving window into eight regions based on angular position (North, South, East, West, NW, NE, SW, and SE). Figure 5-37 shows a 5×5 moving window and the regions of the Local Region filter.

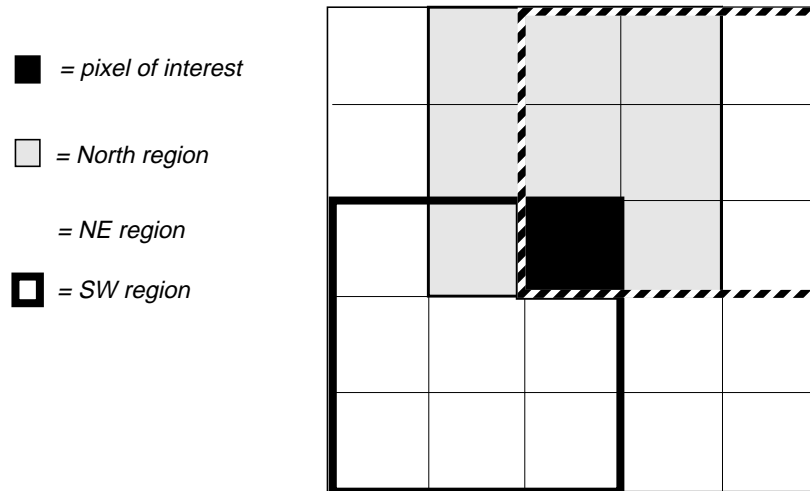


Figure 5-37 Regions of Local Region Filter

For each region, the variance is calculated as follows:

$$Variance = \frac{\sum(DN_{x,y} - Mean)^2}{n - 1}$$

Source: Nagao 1979

The algorithm compares the variance values of the regions surrounding the pixel of interest. The pixel of interest is replaced by the mean of all DN values within the region with the lowest variance (i.e., the most uniform region). A region with low variance is assumed to have pixels minimally affected by wave interference, yet very similar to the pixel of interest. A region of low variance is probably such for several surrounding pixels.

The result is that the output image is composed of numerous uniform areas, the size of which is determined by the moving window size. In practice, this filter can be utilized sequentially 2 or 3 times, increasing the window size. The resultant output image is an appropriate input to a classification application.

Sigma and Lee Filters

The Sigma and Lee filters utilize the statistical distribution of the DN values within the moving window to estimate what the pixel of interest should be.

Speckle in imaging radar can be mathematically modeled as multiplicative noise with a mean of 1. The standard deviation of the noise can be mathematically defined as:

$$Standard\ Deviation \Rightarrow \frac{\sqrt{VARIANCE}}{MEAN} = Coefficient\ Of\ Variation \Rightarrow sigma (\sigma)$$

The coefficient of variation, as a scene-derived parameter, is used as an input parameter in the Sigma and Lee filters. It is also useful in evaluating and modifying VIS/IR data for input to a 4-band composite image, or in preparing a 3-band ratio color composite (Crippen 1989).

It can be assumed that imaging radar data noise follows a Gaussian distribution. This would yield a theoretical value for Standard Deviation (SD) of .52 for 1-look radar data and SD = .26 for 4-look radar data.

Table 5-2 gives theoretical coefficient of variation values for various look-average radar scenes:

Table 5-2 Theoretical Coefficient of Variation Values

# of Looks (scenes)	Coef. of Variation Value
1	.52
2	.37
3	.30
4	.26
6	.21
8	.18

The Lee filters are based on the assumption that the mean and variance of the pixel of interest are equal to the local mean and variance of all pixels within the moving window you select.

The actual calculation used for the Lee filter is:

$$DN_{out} = [Mean] + K[DN_{in} - Mean]$$

Where:

Mean = average of pixels in a moving window

$$K = \frac{Var(x)}{[Mean]^2 \sigma^2 + Var(x)}$$

The variance of x [Var (x)] is defined as:

$$Var(x) = \left(\frac{[Variance\ within\ window] + [Mean\ within\ window]^2}{[Sigma]^2 + 1} \right) - [Mean\ within\ window]^2$$

Source: Lee 1981

The Sigma filter is based on the probability of a Gaussian distribution. It is assumed that 95.5% of random samples are within a 2 standard deviation (2 sigma) range. This noise suppression filter replaces the pixel of interest with the average of all DN values within the moving window that fall within the designated range.

As with all the radar speckle filters, you must specify a moving window size. The center pixel of the moving window is the pixel of interest.

As with the Statistics filter, a coefficient of variation specific to the data set must be input. Finally, you must specify how many standard deviations to use (2, 1, or 0.5) to define the accepted range.

The statistical filters (Sigma and Statistics) are logically applicable to any data set for preprocessing. Any sensor system has various sources of noise, resulting in a few erratic pixels. In VIS/IR imagery, most natural scenes are found to follow a normal distribution of DN values, thus filtering at 2 standard deviations should remove this noise. This is particularly true of experimental sensor systems that frequently have significant noise problems.

These speckle filters can be used iteratively. You must view and evaluate the resultant image after each pass (the data histogram is useful for this), and then decide if another pass is appropriate and what parameters to use on the next pass. For example, three passes of the Sigma filter with the following parameters are very effective when used with any type of data:

Table 5-3 Parameters for Sigma Filter

Pass	Sigma Value	Sigma Multiplier	Window Size
1	0.26	0.5	3 × 3
2	0.26	1	5 × 5
3	0.26	2	7 × 7

Similarly, there is no reason why successive passes must be of the same filter. The following sequence is useful prior to a classification:

Table 5-4 Pre-Classification Sequence

Filter	Pass	Sigma Value	Sigma Multiplier	Window Size
Lee	1	0.26	NA	3 × 3
Lee	2	0.26	NA	5 × 5
Local Region	3	NA	NA	5 × 5 or 7 × 7



With all speckle reduction filters there is a payoff between noise reduction and loss of resolution. Each data set and each application have a different acceptable balance between these two factors. The ERDAS IMAGINE filters have been designed to be versatile and gentle in reducing noise (and resolution).

Frost Filter

The Frost filter is a minimum mean square error algorithm that adapts to the local statistics of the image. The local statistics serve as weighting parameters for the impulse response of the filter (moving window). This algorithm assumes that noise is multiplicative with stationary statistics.

The formula used is:

$$DN = \sum_{n \times n} K \alpha e^{-\alpha|t|}$$

Where:

K=normalization constant

\bar{I} =local mean

σ =local variance

$\bar{\sigma}$ =image coefficient of variation value

$|t| = |X - X_0| + |Y - Y_0|$

n=moving window size

And

$$\alpha = (4/n\bar{\sigma}^2)(\sigma^2/\bar{I}^2)$$

Source: Lopes, Nezry, Touzi, Laur 1990

Gamma-MAP Filter

The Maximum A Posteriori (MAP) filter attempts to estimate the original pixel DN, which is assumed to lie between the local average and the degraded (actual) pixel DN. MAP logic maximizes the a posteriori probability density function with respect to the original image.

Many speckle reduction filters (e.g., Lee, Lee-Sigma, Frost) assume a Gaussian distribution for the speckle noise. Recent work has shown this to be an invalid assumption. Natural vegetated areas have been shown to be more properly modeled as having a Gamma distributed cross section. This algorithm incorporates this assumption. The exact formula used is the cubic equation:

$$\hat{I}^3 - \bar{I}^2 + \sigma(\hat{I} - DN) = 0$$

Enhancement

Where:

\hat{I} =sought value

\bar{I} =local mean

DN=input value

σ =original image variance

Source: Frost, Stiles, Shanmugan, Holtzman 1982

Edge Detection

Edge and line detection are important operations in digital image processing. For example, geologists are often interested in mapping lineaments, which may be fault lines or bedding structures. For this purpose, edge and line detection are major enhancement techniques.

In selecting an algorithm, it is first necessary to understand the nature of what is being enhanced. Edge detection could imply amplifying an edge, a line, or a spot (see Figure 5-38).

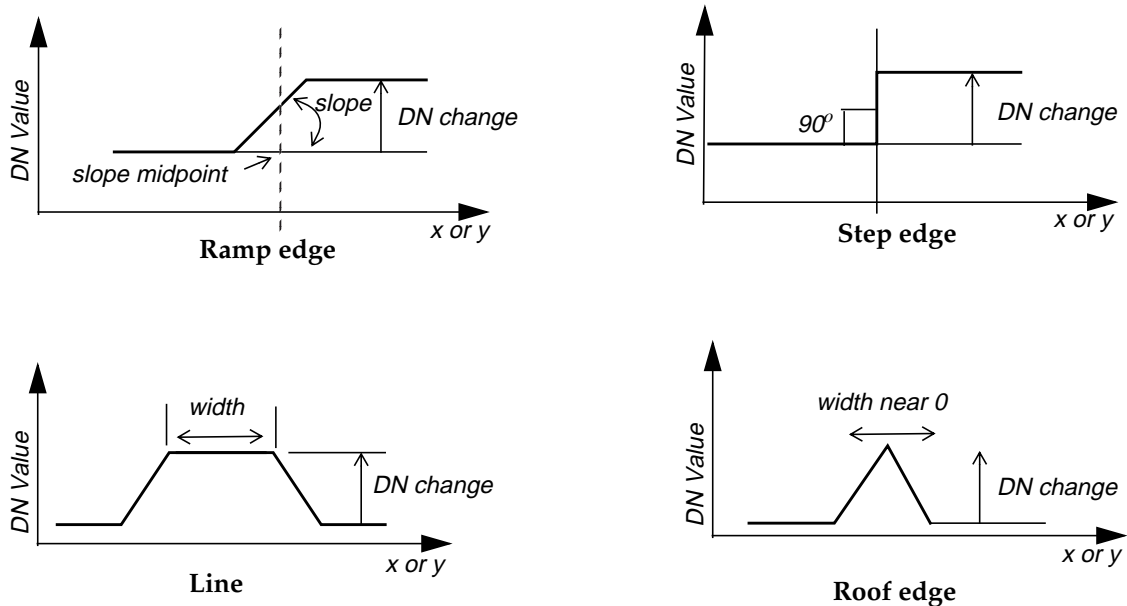


Figure 5-38 One-dimensional, Continuous Edge, and Line Models

- Ramp edge—an edge modeled as a ramp, increasing in DN value from a low to a high level, or vice versa. Distinguished by DN change, slope, and slope midpoint.
- Step edge—a ramp edge with a slope angle of 90 degrees.
- Line—a region bounded on each end by an edge; width must be less than the moving window size.
- Roof edge—a line with a width near zero.

The models in Figure 5-38 represent ideal theoretical edges. However, real data values vary to produce a more distorted edge due to sensor noise or vibration. (see Figure 5-39). There are no perfect edges in raster data, hence the need for edge detection algorithms.

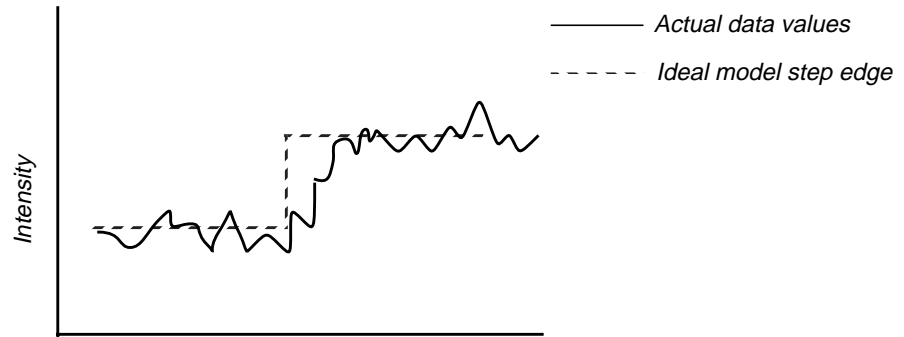


Figure 5-39 A Noisy Edge Superimposed on an Ideal Edge

Edge detection algorithms can be broken down into 1st-order derivative and 2nd-order derivative operations. Figure 5-40 shows ideal one-dimensional edge and line intensity curves with the associated 1st-order and 2nd-order derivatives.

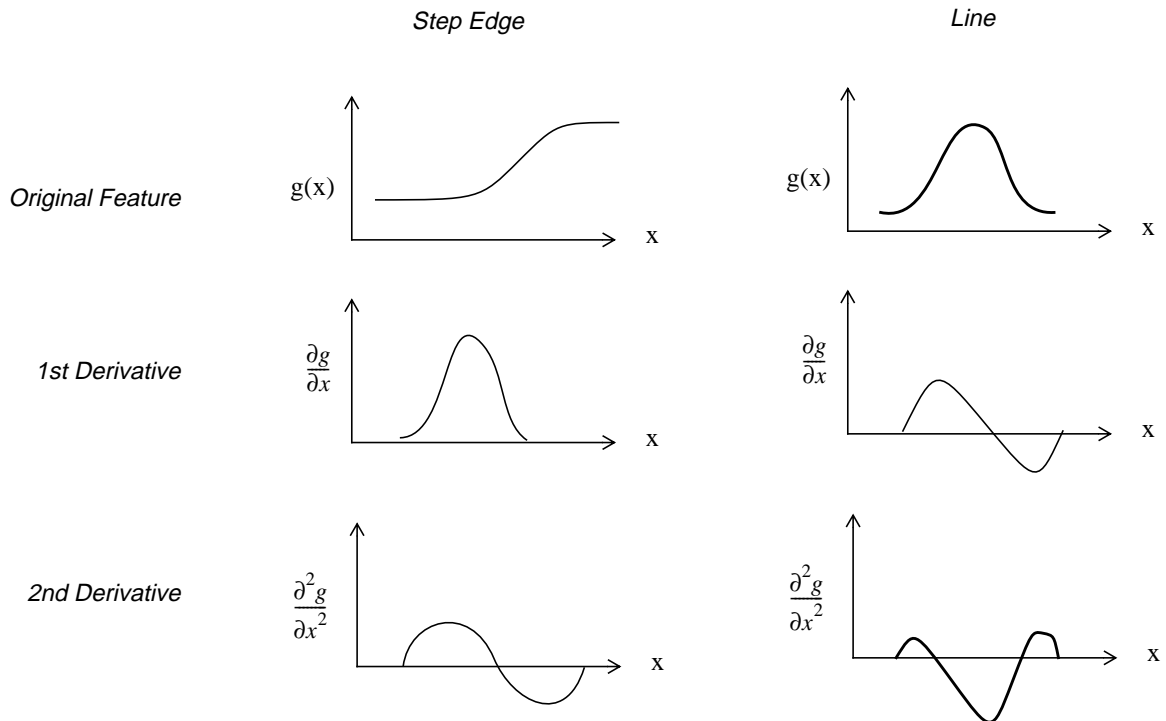


Figure 5-40 Edge and Line Derivatives

The 1st-order derivative kernel(s) derives from the simple Prewitt kernel:

$$\frac{\partial}{\partial x} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \frac{\partial}{\partial y} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

The 2nd-order derivative kernel(s) derives from Laplacian operators:

$$\frac{\partial^2}{\partial x^2} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \quad \text{and} \quad \frac{\partial^2}{\partial y^2} = \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$$

1st-Order Derivatives (Prewitt)

The IMAGINE Radar Interpreter utilizes sets of template matching operators. These operators approximate to the eight possible compass orientations (North, South, East, West, Northeast, Northwest, Southeast, Southwest). The compass names indicate the slope direction creating maximum response. (Gradient kernels with zero weighting, i.e., the sum of the kernel coefficient is zero, have no output in uniform regions.) The detected edge is orthogonal to the gradient direction.

To avoid positional shift, all operating windows are odd number arrays, with the center pixel being the pixel of interest. Extension of the 3 × 3 impulse response arrays to a larger size is not clear cut—different authors suggest different lines of rationale. For example, it may be advantageous to extend the 3-level (Prewitt 1970) to:

$$\begin{bmatrix} 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 & -1 \end{bmatrix}$$

or the following might be beneficial:

$$\begin{bmatrix} 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 4 & 2 & 0 & -2 & -4 \\ 4 & 2 & 0 & -2 & -4 \\ 4 & 2 & 0 & -2 & -4 \\ 4 & 2 & 0 & -2 & -4 \\ 4 & 2 & 0 & -2 & -4 \end{bmatrix}$$

Larger template arrays provide a greater noise immunity, but are computationally more demanding.

Zero-Sum Filters

A common type of edge detection kernel is a zero-sum filter. For this type of filter, the coefficients are designed to add up to zero. Following are examples of two zero-sum filters:

$$\text{Sobel} = \begin{matrix} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \\ \text{horizontal} & \text{vertical} \end{matrix}$$

$$\text{Prewitt} = \begin{matrix} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \\ \text{horizontal} & \text{vertical} \end{matrix}$$



Prior to edge enhancement, you should reduce speckle noise by using the IMAGINE Radar Interpreter Speckle Suppression function.

2nd-Order Derivatives (Laplacian Operators)

The second category of edge enhancers is 2nd-order derivative or Laplacian operators. These are best for line (or spot) detection as distinct from ramp edges. ERDAS IMAGINE Radar Interpreter offers two such arrays:

Unweighted line:

$$\begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

Weighted line:

$$\begin{bmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{bmatrix}$$

Source: Pratt 1991



Some researchers have found that a combination of 1st- and 2nd-order derivative images produces the best output. See Eberlein and Weszka (1975) for information about subtracting the 2nd-order derivative (Laplacian) image from the 1st-order derivative image (gradient).

Texture

According to Pratt (1991), "Many portions of images of natural scenes are devoid of sharp edges over large areas. In these areas the scene can often be characterized as exhibiting a consistent structure analogous to the texture of cloth. Image texture measurements can be used to segment an image and classify its segments."

As an enhancement, texture is particularly applicable to radar data, although it may be applied to any type of data with varying results. For example, it has been shown (Blom *et al* 1982) that a three-layer variance image using 15×15 , 31×31 , and 61×61 windows can be combined into a three-color RGB (red, green, blue) image that is useful for geologic discrimination. The same could apply to a vegetation classification.

You could also prepare a three-color image using three different functions operating through the same (or different) size moving window(s). However, each data set and application would need different moving window sizes and/or texture measures to maximize the discrimination.

Radar Texture Analysis

While texture analysis has been useful in the enhancement of VIS/IR image data, it is showing even greater applicability to radar imagery. In part, this stems from the nature of the imaging process itself.

The interaction of the radar waves with the surface of interest is dominated by reflection involving the surface roughness at the wavelength scale. In VIS/IR imaging, the phenomena involved is absorption at the molecular level. Also, as we know from array-type antennae, radar is especially sensitive to regularity that is a multiple of its wavelength. This provides for a more precise method for quantifying the character of texture in a radar return.

The ability to use radar data to detect texture and provide topographic information about an image is a major advantage over other types of imagery where texture is not a quantitative characteristic.

The texture transforms can be used in several ways to enhance the use of radar imagery. Adding the radar intensity image as an additional layer in a (vegetation) classification is fairly straightforward and may be useful. However, the proper texture image (function and window size) can greatly increase the discrimination. Using known test sites, one can experiment to discern which texture image best aids the classification. For example, the texture image could then be added as an additional layer to the TM bands.



As radar data come into wider use, other mathematical texture definitions may prove useful and will be added to the IMAGINE Radar Interpreter. In practice, you interactively decide which algorithm and window size is best for your data and application.

Texture Analysis Algorithms

While texture has typically been a qualitative measure, it can be enhanced with mathematical algorithms. Many algorithms appear in literature for specific applications (Haralick 1979; Irons 1981).

The algorithms incorporated into ERDAS IMAGINE are those which are applicable in a wide variety of situations and are not computationally over-demanding. This later point becomes critical as the moving window size increases. Research has shown that very large moving windows are often needed for proper enhancement. For example, Blom (Blom *et al* 1982) uses up to a 61×61 window.

Four algorithms are currently utilized for texture enhancement in ERDAS IMAGINE:

- mean Euclidean distance (1st-order)
- variance (2nd-order)
- skewness (3rd-order)
- kurtosis (4th-order)

Mean Euclidean Distance

These algorithms are shown below (Irons and Petersen 1981):

$$\text{Mean Euclidean Distance} = \frac{\sum[\sum_{\lambda}(x_{c\lambda} - x_{ij\lambda})^2]}{n - 1}^{\frac{1}{2}}$$

Where:

$x_{ij\lambda}$ = DN value for spectral band λ and pixel (i,j) of a multispectral image
 $x_{c\lambda}$ = DN value for spectral band λ of a window's center pixel
 n = number of pixels in a window

Variance

NOTE: When the Variance option is chosen, the program computes the standard deviation using the moving window. To compute the actual variance, you need to square the results of the program.

$$\text{Variance} = \frac{\sum(x_{ij} - M)^2}{n - 1}$$

Where:

x_{ij} = DN value of pixel (i,j)
 n = number of pixels in a window
 M = Mean of the moving window, where:

$$Mean = \frac{\sum x_{ij}}{n}$$

Skewness

$$Skew = \frac{|\sum (x_{ij} - M)^3|}{(n - 1)(V)^{\frac{3}{2}}}$$

Where:

- x_{ij} = DN value of pixel (ij)
- n = number of pixels in a window
- M = Mean of the moving window (see above)
- V = Variance (see above)

Kurtosis

$$Kurtosis = \frac{\sum (x_{ij} - M)^4}{(n - 1)(V)^2}$$

Where:

- x_{ij} = DN value of pixel (ij)
- n = number of pixels in a window
- M = Mean of the moving window (see above)
- V = Variance (see above)



Texture analysis is available from the Texture function in Image Interpreter and from the IMAGINE Radar Interpreter Texture Analysis function.

Radiometric Correction: Radar Imagery

The raw radar image frequently contains radiometric errors due to:

- imperfections in the transmit and receive pattern of the radar antenna
- errors due to the coherent pulse (i.e., speckle)
- the inherently stronger signal from a near range (closest to the sensor flight path) than a far range (farthest from the sensor flight path) target

Many imaging radar systems use a single antenna that transmits the coherent radar burst and receives the return echo. However, no antenna is perfect; it may have various lobes, dead spots, and imperfections. This causes the received signal to be slightly distorted radiometrically. In addition, range fall-off causes far range targets to be darker (less return signal).

These two problems can be addressed by adjusting the average brightness of each range line to a constant—usually the average overall scene brightness (Chavez 1986). This requires that each line of constant range be long enough to reasonably approximate the overall scene brightness (see Figure 5-41). This approach is generic; it is not specific to any particular radar sensor.



The Adjust Brightness function in ERDAS IMAGINE works by correcting each range line average. For this to be a valid approach, the number of data values must be large enough to provide good average values. Be careful not to use too small an image. This depends upon the character of the scene itself.

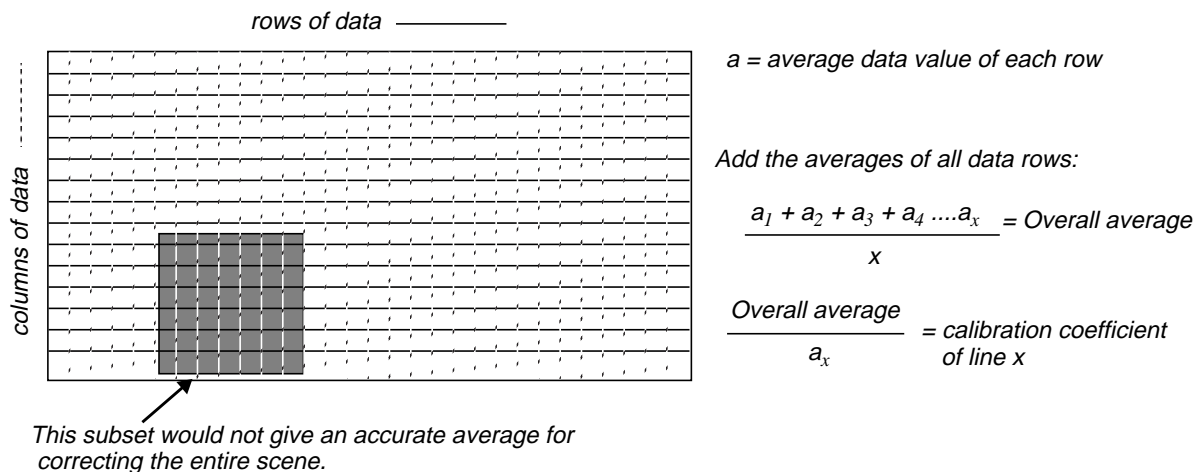


Figure 5-41 Adjust Brightness Function

Range Lines/Lines of Constant Range

Lines of constant range are not the same thing as range lines:

- Range lines—lines that are perpendicular to the flight of the sensor
- Lines of constant range—lines that are parallel to the flight of the sensor
- Range direction—same as range lines

Because radiometric errors are a function of the imaging geometry, the image must be correctly oriented during the correction process. For the algorithm to correctly address the data set, you must tell ERDAS IMAGINE whether the lines of constant range are in columns or rows in the displayed image.

Figure 5-42 shows the lines of constant range in columns, parallel to the sides of the display screen:

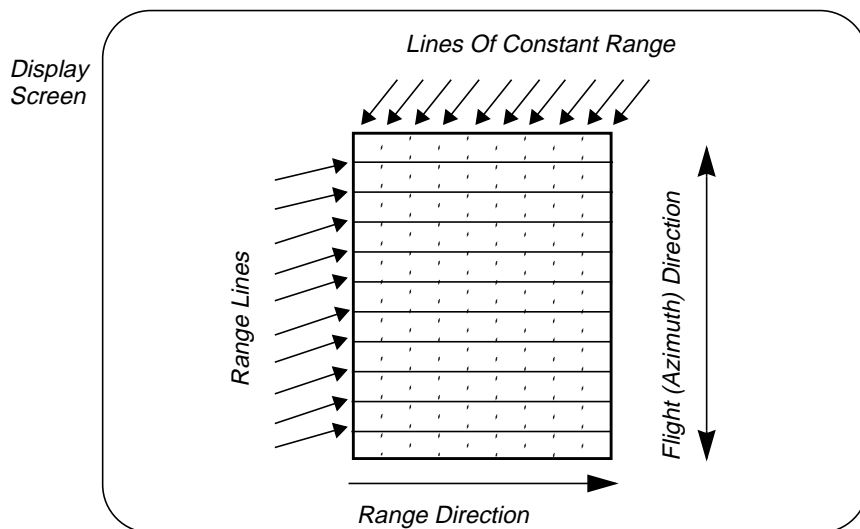


Figure 5-42 Range Lines vs. Lines of Constant Range

Slant-to-Ground Range Correction

Radar images also require slant-to-ground range correction, which is similar in concept to orthorectifying a VIS/IR image. By design, an imaging radar is always side-looking. In practice, the depression angle is usually 75° at most. In operation, the radar sensor determines the range (distance to) each target, as shown in Figure 5-43.

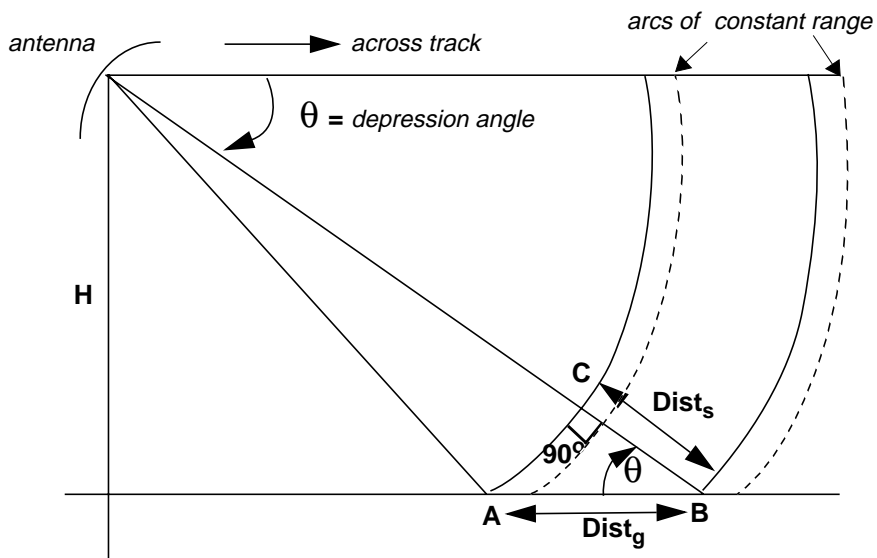


Figure 5-43 Slant-to-Ground Range Correction

Assuming that angle ACB is a right angle, you can approximate:

$$Dist_s \approx (Dist_g)(\cos \theta)$$

Where:

$Dist_s$ = slant range distance

$Dist_g$ = ground range distance

$$\cos \theta = \frac{Dist_s}{Dist_g}$$

Source: Leberl 1990

This has the effect of compressing the near range areas more than the far range areas. For many applications, this may not be important. However, to geocode the scene or to register radar to infrared or visible imagery, the scene must be corrected to a ground range format. To do this, the following parameters relating to the imaging geometry are needed:

- Depression angle (θ)—angular distance between sensor horizon and scene center
- Sensor height (H)—elevation of sensor (in meters) above its nadir point
- Beam width—angular distance between near range and far range for entire scene
- Pixel size (in meters)—range input image cell size



This information is usually found in the header file of data. Use the Data View option to view this information. If it is not contained in the header file, you must obtain this information from the data supplier.

Once the scene is range-format corrected, pixel size can be changed for coregistration with other data sets.

Merging Radar with VIS/IR Imagery

As aforementioned, the phenomena involved in radar imaging is quite different from that in VIS/IR imaging. Because these two sensor types give different information about the same target (chemical vs. physical), they are complementary data sets. If the two images are correctly combined, the resultant image conveys both chemical and physical information and could prove more useful than either image alone.

The methods for merging radar and VIS/IR data are still experimental and open for exploration. The following methods are suggested for experimentation:

- Codisplaying in a Viewer
- RGB to IHS transforms

- Principal components transform
- Multiplicative



The ultimate goal of enhancement is not mathematical or logical purity; it is feature extraction. There are currently no rules to suggest which options yield the best results for a particular application; you must experiment. The option that proves to be most useful depends upon the data sets (both radar and VIS/IR), your experience, and your final objective.

Codisplaying

The simplest and most frequently used method of combining radar with VIS/IR imagery is codisplaying on an RGB color monitor. In this technique, the radar image is displayed with one (typically the red) gun, while the green and blue guns display VIS/IR bands or band ratios. This technique follows from no logical model and does not truly merge the two data sets.



Use the Viewer with the Clear Display option disabled for this type of merge. Select the color guns to display the different layers.

RGB to IHS Transforms

Another common technique uses the RGB to IHS transforms. In this technique, an RGB color composite of bands (or band derivatives, such as ratios) is transformed into IHS color space. The intensity component is replaced by the radar image, and the scene is reverse transformed. This technique integrally merges the two data types.



For more information, see "RGB to IHS" on page 169.

Principal Components Transform

A similar image merge involves utilizing the PC transformation of the VIS/IR image. With this transform, more than three components can be used. These are converted to a series of principal components. The first PC, PC-1, is generally accepted to correlate with overall scene brightness. This value is replaced by the radar image and the reverse transform is applied.



For more information, see "Principal Components Analysis" on page 164.

Multiplicative

A final method to consider is the multiplicative technique. This requires several chromatic components and a multiplicative component, which is assigned to the image intensity. In practice, the chromatic components are usually band ratios or PCs; the radar image is input multiplicatively as intensity (Holcomb 1993).

The two sensor merge models using transforms to integrate the two data sets (PC and RGB to IHS) are based on the assumption that the radar intensity correlates with the intensity that the transform derives from the data inputs. However, the logic of mathematically merging radar with VIS/IR data sets is inherently different from the logic of the SPOT/TM merges (as discussed in "Resolution Merge" on page 160). It cannot be assumed that the radar intensity is a surrogate for, or equivalent to, the VIS/IR intensity. The acceptability of this assumption depends on the specific case.

For example, Landsat TM imagery is often used to aid in mineral exploration. A common display for this purpose is RGB = TM5/TM7, TM5/TM4, TM3/TM1; the logic being that if all three ratios are high, the sites suited for mineral exploration are bright overall. If the target area is accompanied by silicification, which results in an area of dense angular rock, this should be the case. However, if the alteration zone is basaltic rock to kaolinite/alunite, then the radar return could be weaker than the surrounding rock. In this case, radar would not correlate with high 5/7, 5/4, 3/1 intensity and the substitution would not produce the desired results (Holcomb 1993).

Enhancement

CHAPTER 6

Classification

Introduction

Multispectral classification is the process of sorting pixels into a finite number of individual classes, or categories of data, based on their data file values. If a pixel satisfies a certain set of criteria, the pixel is assigned to the class that corresponds to that criteria. This process is also referred to as image segmentation.

Depending on the type of information you want to extract from the original data, classes may be associated with known features on the ground or may simply represent areas that look different to the computer. An example of a classified image is a land cover map, showing vegetation, bare land, pasture, urban, etc.

The Classification Process

Pattern Recognition

Pattern recognition is the science—and art—of finding meaningful patterns in data, which can be extracted through classification. By spatially and spectrally enhancing an image, pattern recognition can be performed with the human eye; the human brain automatically sorts certain textures and colors into categories.

In a computer system, spectral pattern recognition can be more scientific. Statistics are derived from the spectral characteristics of all pixels in an image. Then, the pixels are sorted based on mathematical criteria. The classification process breaks down into two parts: training and classifying (using a decision rule).

Training

First, the computer system must be trained to recognize patterns in the data. Training is the process of defining the criteria by which these patterns are recognized (Hord 1982). Training can be performed with either a supervised or an unsupervised method, as explained below.

Supervised Training

Supervised training is closely controlled by the analyst. In this process, you select pixels that represent patterns or land cover features that you recognize, or that you can identify with help from other sources, such as aerial photos, ground truth data, or maps. Knowledge of the data, and of the classes desired, is required before classification.

By identifying patterns, you can instruct the computer system to identify pixels with similar characteristics. If the classification is accurate, the resulting classes represent the categories within the data that you originally identified.

Classification

Unsupervised Training

Unsupervised training is more computer-automated. It enables you to specify some parameters that the computer uses to uncover statistical patterns that are inherent in the data. These patterns do not necessarily correspond to directly meaningful characteristics of the scene, such as contiguous, easily recognized areas of a particular soil type or land use. They are simply clusters of pixels with similar spectral characteristics. In some cases, it may be more important to identify groups of pixels with similar spectral characteristics than it is to sort pixels into recognizable categories.

Unsupervised training is dependent upon the data itself for the definition of classes. This method is usually used when less is known about the data before classification. It is then the analyst's responsibility, after classification, to attach meaning to the resulting classes (Jensen 1996). Unsupervised classification is useful only if the classes can be appropriately interpreted.

Signatures

The result of training is a set of signatures that defines a training sample or cluster. Each signature corresponds to a class, and is used with a decision rule (explained below) to assign the pixels in the image file to a class. Signatures in ERDAS IMAGINE can be parametric or nonparametric.

A parametric signature is based on statistical parameters (e.g., mean and covariance matrix) of the pixels that are in the training sample or cluster. Supervised and unsupervised training can generate parametric signatures. A set of parametric signatures can be used to train a statistically-based classifier (e.g., maximum likelihood) to define the classes.

A nonparametric signature is not based on statistics, but on discrete objects (polygons or rectangles) in a feature space image. These feature space objects are used to define the boundaries for the classes. A nonparametric classifier uses a set of nonparametric signatures to assign pixels to a class based on their location either inside or outside the area in the feature space image. Supervised training is used to generate nonparametric signatures (Kloer 1994).

ERDAS IMAGINE enables you to generate statistics for a nonparametric signature. This function allows a feature space object to be used to create a parametric signature from the image being classified. However, since a parametric classifier requires a normal distribution of data, the only feature space object for which this would be mathematically valid would be an ellipse (Kloer 1994).

When both parametric and nonparametric signatures are used to classify an image, you are more able to analyze and visualize the class definitions than either type of signature provides independently (Kloer 1994).



See "Appendix A: Math Topics" for information on feature space images and how they are created.

Decision Rule

After the signatures are defined, the pixels of the image are sorted into classes based on the signatures by use of a classification decision rule. The decision rule is a mathematical algorithm that, using data contained in the signature, performs the actual sorting of pixels into distinct class values.

Parametric Decision Rule

A parametric decision rule is trained by the parametric signatures. These signatures are defined by the mean vector and covariance matrix for the data file values of the pixels in the signatures. When a parametric decision rule is used, every pixel is assigned to a class since the parametric decision space is continuous (Kloer 1994).

Nonparametric Decision Rule

A nonparametric decision rule is not based on statistics; therefore, it is independent of the properties of the data. If a pixel is located within the boundary of a nonparametric signature, then this decision rule assigns the pixel to the signature's class. Basically, a nonparametric decision rule determines whether or not the pixel is located inside of nonparametric signature boundary.

Classification Tips
Classification Scheme

Usually, classification is performed with a set of target classes in mind. Such a set is called a classification scheme (or classification system). The purpose of such a scheme is to provide a framework for organizing and categorizing the information that can be extracted from the data (Jensen 1983). The proper classification scheme includes classes that are both important to the study and discernible from the data on hand. Most schemes have a hierarchical structure, which can describe a study area in several levels of detail.

A number of classification schemes have been developed by specialists who have inventoried a geographic region. Some references for professionally-developed schemes are listed below:

- Anderson, J.R., et al. 1976. "A Land Use and Land Cover Classification System for Use with Remote Sensor Data." *U.S. Geological Survey Professional Paper 964*.
- Cowardin, Lewis M., et al. 1979. *Classification of Wetlands and Deepwater Habitats of the United States*. Washington, D.C.: U.S. Fish and Wildlife Service.
- Florida Topographic Bureau, Thematic Mapping Section. 1985. *Florida Land Use, Cover and Forms Classification System*. Florida Department of Transportation, Procedure No. 550-010-001-a.
- Michigan Land Use Classification and Reference Committee. 1975. *Michigan Land Cover/Use Classification System*. Lansing, Michigan: State of Michigan Office of Land Use.

Other states or government agencies may also have specialized land use/cover studies.

It is recommended that the classification process is begun by defining a classification scheme for the application, using previously developed schemes, like those above, as a general framework.

Classification

Iterative Classification

A process is iterative when it repeats an action. The objective of the ERDAS IMAGINE system is to enable you to iteratively create and refine signatures and classified image files to arrive at a desired final classification. The ERDAS IMAGINE classification utilities are tools to be used as needed, not a numbered list of steps that must always be followed in order.

The total classification can be achieved with either the supervised or unsupervised methods, or a combination of both. Some examples are below:

- Signatures created from both supervised and unsupervised training can be merged and appended together.
- Signature evaluation tools can be used to indicate which signatures are spectrally similar. This helps to determine which signatures should be merged or deleted. These tools also help define optimum band combinations for classification. Using the optimum band combination may reduce the time required to run a classification process.
- Since classifications (supervised or unsupervised) can be based on a particular area of interest (either defined in a raster layer or an .aoi layer), signatures and classifications can be generated from previous classification results.

Supervised vs. Unsupervised Training

In supervised training, it is important to have a set of desired classes in mind, and then create the appropriate signatures from the data. You must also have some way of recognizing pixels that represent the classes that you want to extract.

Supervised classification is usually appropriate when you want to identify relatively few classes, when you have selected training sites that can be verified with ground truth data, or when you can identify distinct, homogeneous regions that represent each class.

On the other hand, if you want the classes to be determined by spectral distinctions that are inherent in the data so that you can define the classes later, then the application is better suited to unsupervised training. Unsupervised training enables you to define many classes easily, and identify classes that are not in contiguous, easily recognized regions.

NOTE: Supervised classification also includes using a set of classes that is generated from an unsupervised classification. Using a combination of supervised and unsupervised classification may yield optimum results, especially with large data sets (e.g., multiple Landsat scenes). For example, unsupervised classification may be useful for generating a basic set of classes, then supervised classification can be used for further definition of the classes.

Classifying Enhanced Data

For many specialized applications, classifying data that have been merged, spectrally merged or enhanced—with principal components, image algebra, or other transformations—can produce very specific and meaningful results. However, without understanding the data and the enhancements used, it is recommended that only the original, remotely-sensed data be classified.

Dimensionality

Dimensionality refers to the number of layers being classified. For example, a data file with 3 layers is said to be 3-dimensional, since 3-dimensional feature space is plotted to analyze the data.



Feature space and dimensionality are discussed in "Appendix A: Math Topics".

Adding Dimensions

Using programs in ERDAS IMAGINE, you can add layers to existing image files. Therefore, you can incorporate data (called ancillary data) other than remotely-sensed data into the classification. Using ancillary data enables you to incorporate variables into the classification from, for example, vector layers, previously classified data, or elevation data. The data file values of the ancillary data become an additional feature of each pixel, thus influencing the classification (Jensen 1996).

Limiting Dimensions

Although ERDAS IMAGINE allows an unlimited number of layers of data to be used for one classification, it is usually wise to reduce the dimensionality of the data as much as possible. Often, certain layers of data are redundant or extraneous to the task at hand. Unnecessary data take up valuable disk space, and cause the computer system to perform more arduous calculations, which slows down processing.



Use the Signature Editor to evaluate separability to calculate the best subset of layer combinations. Use the Image Interpreter functions to merge or subset layers. Use the Image Information tool (on the Viewer's tool bar) to delete a layer(s).

Supervised Training

Supervised training requires *a priori* (already known) information about the data, such as:

- What type of classes need to be extracted? Soil type? Land use? Vegetation?
- What classes are most likely to be present in the data? That is, which types of land cover, soil, or vegetation (or whatever) are represented by the data?

In supervised training, you rely on your own pattern recognition skills and *a priori* knowledge of the data to help the system determine the statistical criteria (signatures) for data classification.

To select reliable samples, you should know some information—either spatial or spectral—about the pixels that you want to classify.

The location of a specific characteristic, such as a land cover type, may be known through ground truthing. Ground truthing refers to the acquisition of knowledge about the study area from field work, analysis of aerial photography, personal experience, etc. Ground truth data are considered to be the most accurate (true) data available about the area of study. They should be collected at the same time as the remotely sensed data, so that the data correspond as much as possible (Star and Estes 1990). However, some ground data may not be very accurate due to a number of errors and inaccuracies.

Training Samples and Feature Space Objects

Training samples (also called samples) are sets of pixels that represent what is recognized as a discernible pattern, or potential class. The system calculates statistics from the sample pixels to create a parametric signature for the class.

Classification

The following terms are sometimes used interchangeably in reference to training samples. For clarity, they are used in this documentation as follows:

- Training sample, or sample, is a set of pixels selected to represent a potential class. The data file values for these pixels are used to generate a parametric signature.
- Training field, or training site, is the geographical AOI in the image represented by the pixels in a sample. Usually, it is previously identified with the use of ground truth data.

Feature space objects are user-defined AOIs in a feature space image. The feature space signature is based on these objects.

Selecting Training Samples

It is important that training samples be representative of the class that you are trying to identify. This does not necessarily mean that they must contain a large number of pixels or be dispersed across a wide region of the data. The selection of training samples depends largely upon your knowledge of the data, of the study area, and of the classes that you want to extract.

ERDAS IMAGINE enables you to identify training samples using one or more of the following methods:

- using a vector layer
- defining a polygon in the image
- identifying a training sample of contiguous pixels with similar spectral characteristics
- identifying a training sample of contiguous pixels within a certain area, with or without similar spectral characteristics
- using a class from a thematic raster layer from an image file of the same area (i.e., the result of an unsupervised classification)

Digitized Polygon

Training samples can be identified by their geographical location (training sites, using maps, ground truth data). The locations of the training sites can be digitized from maps with the ERDAS IMAGINE Vector or AOI tools. Polygons representing these areas are then stored as vector layers. The vector layers can then be used as input to the AOI tools and used as training samples to create signatures.



Use the Vector and AOI tools to digitize training samples from a map. Use the Signature Editor to create signatures from training samples that are identified with digitized polygons.

User-defined Polygon

Using your pattern recognition skills (with or without supplemental ground truth information), you can identify samples by examining a displayed image of the data and drawing a polygon around the training site(s) of interest. For example, if it is known that oak trees reflect certain frequencies of green and infrared light according to ground truth data, you may be able to base your sample selections on the data (taking atmospheric conditions, sun angle, time, date, and other variations into account). The area within the polygon(s) would be used to create a signature.



Use the AOI tools to define the polygon(s) to be used as the training sample. Use the Signature Editor to create signatures from training samples that are identified with the polygons.

Identify Seed Pixel

With the Seed Properties dialog and AOI tools, the cursor (crosshair) can be used to identify a single pixel (seed pixel) that is representative of the training sample. This seed pixel is used as a model pixel, against which the pixels that are contiguous to it are compared based on parameters specified by you.

When one or more of the contiguous pixels is accepted, the mean of the sample is calculated from the accepted pixels. Then, the pixels contiguous to the sample are compared in the same way. This process repeats until no pixels that are contiguous to the sample satisfy the spectral parameters. In effect, the sample grows outward from the model pixel with each iteration. These homogenous pixels are converted from individual raster pixels to a polygon and used as an AOI layer.



Select the Seed Properties option in the Viewer to identify training samples with a seed pixel.

Seed Pixel Method with Spatial Limits

The training sample identified with the seed pixel method can be limited to a particular region by defining the geographic distance and area.



Vector layers (polygons or lines) can be displayed as the top layer in the Viewer, and the boundaries can then be used as an AOI for training samples defined under Seed Properties.

Thematic Raster Layer

A training sample can be defined by using class values from a thematic raster layer (see Table 6-1). The data file values in the training sample are used to create a signature. The training sample can be defined by as many class values as desired.

NOTE: The thematic raster layer must have the same coordinate system as the image file being classified.

Table 6-1 Training Sample Comparison

Method	Advantages	Disadvantages
Digitized Polygon	precise map coordinates, represents known ground information	may overestimate class variance, time-consuming
User-defined Polygon	high degree of user control	may overestimate class variance, time-consuming
Seed Pixel	auto-assisted, less time	may underestimate class variance
Thematic Raster Layer	allows iterative classifying	must have previously defined thematic layer

Evaluating Training Samples

Selecting training samples is often an iterative process. To generate signatures that accurately represent the classes to be identified, you may have to repeatedly select training samples, evaluate the signatures that are generated from the samples, and then either take new samples or manipulate the signatures as necessary. Signature manipulation may involve merging, deleting, or appending from one file to another. It is also possible to perform a classification using the known signatures, then mask out areas that are not classified to use in gathering more signatures.



See "Evaluating Signatures" on page 235 for methods of determining the accuracy of the signatures created from your training samples.

Selecting Feature Space Objects

The ERDAS IMAGINE Feature Space tools enable you to interactively define feature space objects (AOIs) in the feature space image(s). A feature space image is simply a graph of the data file values of one band of data against the values of another band (often called a scatterplot). In ERDAS IMAGINE, a feature space image has the same data structure as a raster image; therefore, feature space images can be used with other ERDAS IMAGINE utilities, including zoom, color level slicing, virtual roam, Spatial Modeler, and Map Composer.

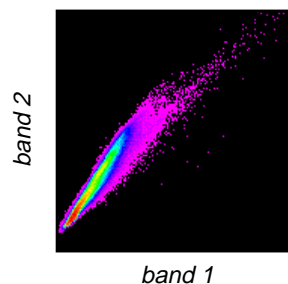


Figure 6-1 Example of a Feature Space Image

The transformation of a multilayer raster image into a feature space image is done by mapping the input pixel values to a position in the feature space image. This transformation defines only the pixel position in the feature space image. It does not define the pixel's value. The pixel values in the feature space image can be the accumulated frequency, which is calculated when the feature space image is defined. The pixel values can also be provided by a thematic raster layer of the same geometry as the source multilayer image. Mapping a thematic layer into a feature space image can be useful for evaluating the validity of the parametric and nonparametric decision boundaries of a classification (Kloer 1994).



When you display a feature space image file (.fsp.img) in a Viewer, the colors reflect the density of points for both bands. The bright tones represent a high density and the dark tones represent a low density.

Create Nonparametric Signature

You can define a feature space object (AOI) in the feature space image and use it directly as a nonparametric signature. Since the Viewers for the feature space image and the image being classified are both linked to the ERDAS IMAGINE Signature Editor, it is possible to mask AOIs from the image being classified to the feature space image, and vice versa. You can also directly link a cursor in the image Viewer to the feature space Viewer. These functions help determine a location for the AOI in the feature space image.

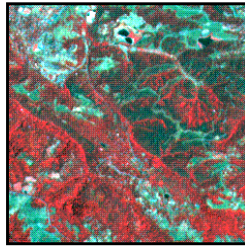
A single feature space image, but multiple AOIs, can be used to define the signature. This signature is taken within the feature space image, not the image being classified. The pixels in the image that correspond to the data file values in the signature (i.e., feature space object) are assigned to that class.

One fundamental difference between using the feature space image to define a training sample and the other traditional methods is that it is a nonparametric signature. The decisions made in the classification process have no dependency on the statistics of the pixels. This helps improve classification accuracies for specific nonnormal classes, such as urban and exposed rock (Faust, et al 1991).

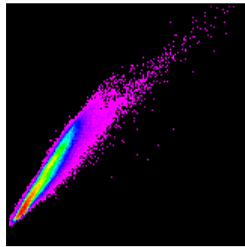


See "Appendix A: Math Topics" for information on feature space images.

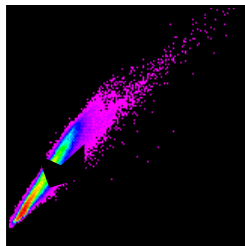
Classification



Display the image file to be classified in a Viewer (layers 3, 2, 1).



Create feature space image from the image file being classified (layer 1 vs. layer 2).



Draw an AOI (feature space object around the desired area in the feature space image. Once you have a desired AOI, it can be used as a signature.

A decision rule is used to analyze each pixel in the image file being classified, and the pixels with the corresponding data file values are assigned to the feature space class.

Figure 6-2 Process for Defining a Feature Space Object

Evaluate Feature Space Signatures

Using the Feature Space tools, it is also possible to use a feature space signature to generate a mask. Once it is defined as a mask, the pixels under the mask are identified in the image file and highlighted in the Viewer. The image displayed in the Viewer must be the image from which the feature space image was created. This process helps you to visually analyze the correlations between various spectral bands to determine which combination of bands brings out the desired features in the image.

You can have as many feature space images with different band combinations as desired. Any polygon or rectangle in these feature space images can be used as a nonparametric signature. However, only one feature space image can be used per signature. The polygons in the feature space image can be easily modified and/or masked until the desired regions of the image have been identified.



Use the Feature Space tools in the Signature Editor to create a feature space image and mask the signature. Use the AOI tools to draw polygons.

Table 6-2 Feature Space Signatures

Advantages	Disadvantages
Provide an accurate way to classify a class with a nonnormal distribution (e.g., residential and urban).	The classification decision process allows overlap and unclassified pixels.
Certain features may be more visually identifiable in a feature space image.	The feature space image may be difficult to interpret.
The classification decision process is fast.	

Unsupervised Training

Unsupervised training requires only minimal initial input from you. However, you have the task of interpreting the classes that are created by the unsupervised training algorithm.

Unsupervised training is also called clustering, because it is based on the natural groupings of pixels in image data when they are plotted in feature space. According to the specified parameters, these groups can later be merged, disregarded, otherwise manipulated, or used as the basis of a signature.



Feature space is explained in "Appendix A: Math Topics".

Clusters

Clusters are defined with a clustering algorithm, which often uses all or many of the pixels in the input data file for its analysis. The clustering algorithm has no regard for the contiguity of the pixels that define each cluster.

- The Iterative Self-Organizing Data Analysis Technique (ISODATA) (Tou and Gonzalez 1974) clustering method uses spectral distance as in the sequential method, but iteratively classifies the pixels, redefines the criteria for each class, and classifies again, so that the spectral distance patterns in the data gradually emerge.
- The RGB clustering method is more specialized than the ISODATA method. It applies to three-band, 8-bit data. RGB clustering plots pixels in three-dimensional feature space, and divides that space into sections that are used to define clusters.

Each of these methods is explained below, along with its advantages and disadvantages.



Some of the statistics terms used in this section are explained in "Appendix A: Math Topics".

ISODATA Clustering

ISODATA is iterative in that it repeatedly performs an entire classification (outputting a thematic raster layer) and recalculates statistics. Self-Organizing refers to the way in which it locates clusters with minimum user input.

Classification

The ISODATA method uses minimum spectral distance to assign a cluster for each candidate pixel. The process begins with a specified number of arbitrary cluster means or the means of existing signatures, and then it processes repetitively, so that those means shift to the means of the clusters in the data.

Because the ISODATA method is iterative, it is not biased to the top of the data file, as are the one-pass clustering algorithms.



Use the Unsupervised Classification utility in the Signature Editor to perform ISODATA clustering.

ISODATA Clustering Parameters

To perform ISODATA clustering, you specify:

- N - the maximum number of clusters to be considered. Since each cluster is the basis for a class, this number becomes the maximum number of classes to be formed. The ISODATA process begins by determining N arbitrary cluster means. Some clusters with too few pixels can be eliminated, leaving less than N clusters.
- T - a convergence threshold, which is the maximum percentage of pixels whose class values are allowed to be unchanged between iterations.
- M - the maximum number of iterations to be performed.

Initial Cluster Means

On the first iteration of the ISODATA algorithm, the means of N clusters can be arbitrarily determined. After each iteration, a new mean for each cluster is calculated, based on the actual spectral locations of the pixels in the cluster, instead of the initial arbitrary calculation. Then, these new means are used for defining clusters in the next iteration. The process continues until there is little change between iterations (Swain 1973).

The initial cluster means are distributed in feature space along a vector that runs between the point at spectral coordinates $(\mu_1-\sigma_1, \mu_2-\sigma_2, \mu_3-\sigma_3, \dots, \mu_n-\sigma_n)$ and the coordinates $(\mu_1+\sigma_1, \mu_2+\sigma_2, \mu_3+\sigma_3, \dots, \mu_n+\sigma_n)$. Such a vector in two dimensions is illustrated in Figure 6-3. The initial cluster means are evenly distributed between $(\mu_A-\sigma_A, \mu_B-\sigma_B)$ and $(\mu_A+\sigma_A, \mu_B+\sigma_B)$.

5 arbitrary cluster means in two-dimensional spectral space

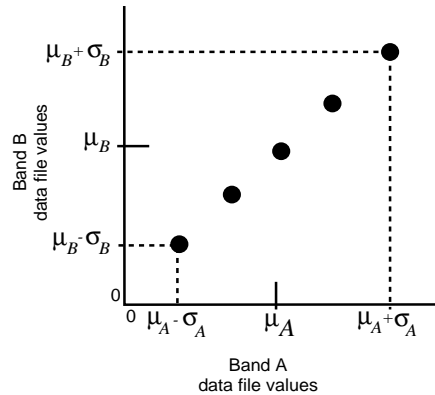


Figure 6-3 ISODATA Arbitrary Clusters

Pixel Analysis

Pixels are analyzed beginning with the upper left corner of the image and going left to right, block by block.

The spectral distance between the candidate pixel and each cluster mean is calculated. The pixel is assigned to the cluster whose mean is the closest. The ISODATA function creates an output image file with a thematic raster layer and/or a signature file (.sig) as a result of the clustering. At the end of each iteration, an image file exists that shows the assignments of the pixels to the clusters.

Considering the regular, arbitrary assignment of the initial cluster means, the first iteration of the ISODATA algorithm always gives results similar to those in Figure 6-4.

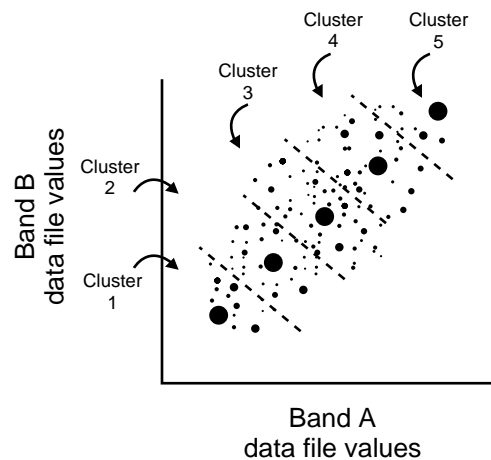


Figure 6-4 ISODATA First Pass

Classification

For the second iteration, the means of all clusters are recalculated, causing them to shift in feature space. The entire process is repeated—each candidate pixel is compared to the new cluster means and assigned to the closest cluster mean.

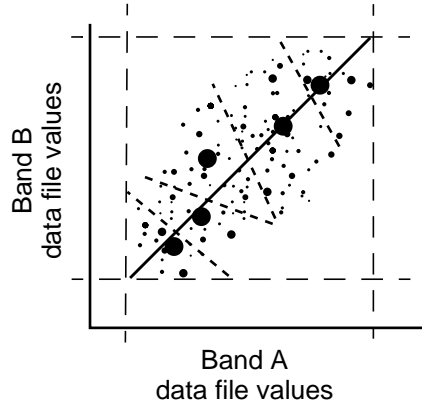


Figure 6-5 ISODATA Second Pass

Percentage Unchanged

After each iteration, the normalized percentage of pixels whose assignments are unchanged since the last iteration is displayed in the dialog. When this number reaches T (the convergence threshold), the program terminates.

It is possible for the percentage of unchanged pixels to never converge or reach T (the convergence threshold). Therefore, it may be beneficial to monitor the percentage, or specify a reasonable maximum number of iterations, M , so that the program does not run indefinitely.

Table 6-3 ISODATA Clustering

Advantages	Disadvantages
Because it is iterative, clustering is not geographically biased to the top or bottom pixels of the data file.	The clustering process is time-consuming, because it can repeat many times.
This algorithm is highly successful at finding the spectral clusters that are inherent in the data. It does not matter where the initial arbitrary cluster means are located, as long as enough iterations are allowed.	Does not account for pixel spatial homogeneity.
A preliminary thematic raster layer is created, which gives results similar to using a minimum distance classifier (as explained below) on the signatures that are created. This thematic raster layer can be used for analyzing and manipulating the signatures before actual classification takes place.	

Principal Component Method

Whereas clustering creates signatures depending on pixels' spectral reflectance by adding pixels together, the principal component method actually subtracts pixels. Principal Components Analysis (PCA) is a method of data compression. With it, you can eliminate data that is redundant by compacting it into fewer bands.

The resulting bands are noncorrelated and independent. You may find these bands more interpretable than the source data. PCA can be performed on up to 256 bands with ERDAS IMAGINE. As a type of spectral enhancement, you are required to specify the number of components you want output from the original data.

Recommended Decision Rule

Although the ISODATA algorithm is the most similar to the minimum distance decision rule, the signatures can produce good results with any type of classification. Therefore, no particular decision rule is recommended over others.

In most cases, the signatures created by ISODATA are merged, deleted, or appended to other signature sets. The image file created by ISODATA is the same as the image file that is created by a minimum distance classification, except for the nonconvergent pixels (100-T% of the pixels).



Use the Merge and Delete options in the Signature Editor to manipulate signatures.



Use the Unsupervised Classification utility in the Signature Editor to perform ISODATA clustering, generate signatures, and classify the resulting signatures.

Classification

RGB Clustering



The RGB Clustering and Advanced RGB Clustering functions in Image Interpreter create a thematic raster layer. However, no signature file is created and no other classification decision rule is used. In practice, RGB Clustering differs greatly from the other clustering methods, but it does employ a clustering algorithm.

RGB clustering is a simple classification and data compression technique for three bands of data. It is a fast and simple algorithm that quickly compresses a three-band image into a single band pseudocolor image, without necessarily classifying any particular features.

The algorithm plots all pixels in 3-dimensional feature space and then partitions this space into clusters on a grid. In the more simplistic version of this function, each of these clusters becomes a class in the output thematic raster layer.

The advanced version requires that a minimum threshold on the clusters be set so that only clusters at least as large as the threshold become output classes. This allows for more color variation in the output file. Pixels that do not fall into any of the remaining clusters are assigned to the cluster with the smallest city-block distance from the pixel. In this case, the city-block distance is calculated as the sum of the distances in the red, green, and blue directions in 3-dimensional space.

Along each axis of the three-dimensional scatterplot, each input histogram is scaled so that the partitions divide the histograms between specified limits—either a specified number of standard deviations above and below the mean, or between the minimum and maximum data values for each band.

The default number of divisions per band is listed below:

- Red is divided into 7 sections (32 for advanced version)
- Green is divided into 6 sections (32 for advanced version)
- Blue is divided into 6 sections (32 for advanced version)

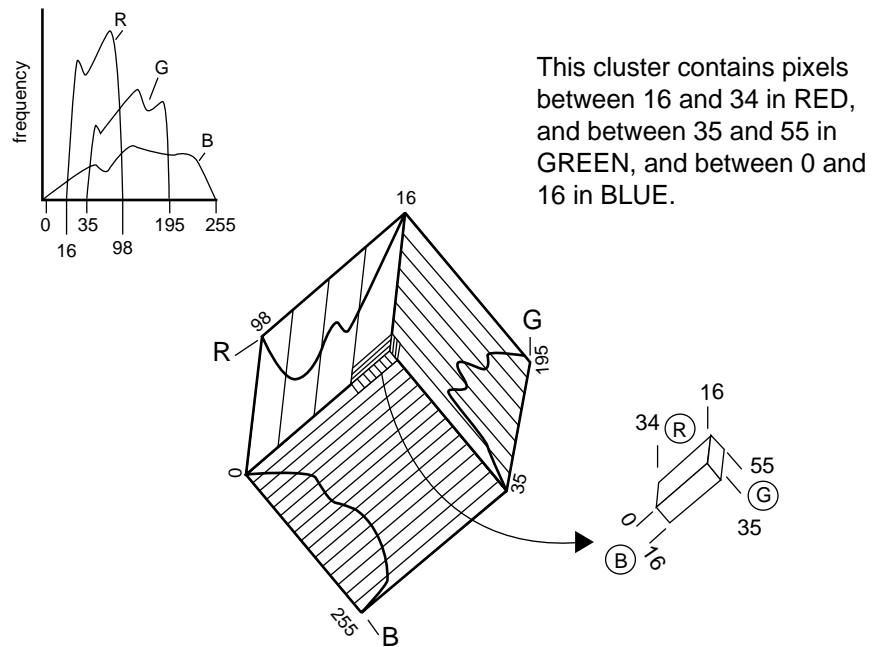


Figure 6-6 RGB Clustering

Partitioning Parameters

It is necessary to specify the number of R, G, and B sections in each dimension of the 3-dimensional scatterplot. The number of sections should vary according to the histograms of each band. Broad histograms should be divided into more sections, and narrow histograms should be divided into fewer sections (see Figure 6-6).



It is possible to interactively change these parameters in the RGB Clustering function in the Image Interpreter. The number of classes is calculated based on the current parameters, and it displays on the command screen.

Table 6-4 RGB Clustering

Advantages	Disadvantages
The fastest classification method. It is designed to provide a fast, simple classification for applications that do not require specific classes.	Exactly three bands must be input, which is not suitable for all applications.
Not biased to the top or bottom of the data file. The order in which the pixels are examined does not influence the outcome.	Does not always create thematic classes that can be analyzed for informational purposes.
(Advanced version only) A highly interactive function, allowing an iterative adjustment of the parameters until the number of clusters and the thresholds are satisfactory for analysis.	

Tips

Some starting values that usually produce good results with the simple RGB clustering are:

R = 7
G = 6
B = 6

which results in $7 \times 6 \times 6 = 252$ classes.

To decrease the number of output colors/classes or to darken the output, decrease these values.

For the Advanced RGB clustering function, start with higher values for R, G, and B. Adjust by raising the threshold parameter and/or decreasing the R, G, and B parameter values until the desired number of output classes is obtained.

Signature Files

A signature is a set of data that defines a training sample, feature space object (AOI), or cluster. The signature is used in a classification process. Each classification decision rule (algorithm) requires some signature attributes as input—these are stored in the signature file (.sig). Signatures in ERDAS IMAGINE can be parametric or nonparametric.

The following attributes are standard for all signatures (parametric and nonparametric):

- name—identifies the signature and is used as the class name in the output thematic raster layer. The default signature name is Class <number>.
- color—the color for the signature and the color for the class in the output thematic raster layer. This color is also used with other signature visualization functions, such as alarms, masking, ellipses, etc.

- value—the output class value for the signature. The output class value does not necessarily need to be the class number of the signature. This value should be a positive integer.
- order—the order to process the signatures for order-dependent processes, such as signature alarms and parallelepiped classifications.
- parallelepiped limits—the limits used in the parallelepiped classification.

Parametric Signature

A parametric signature is based on statistical parameters (e.g., mean and covariance matrix) of the pixels that are in the training sample or cluster. A parametric signature includes the following attributes in addition to the standard attributes for signatures:

- the number of bands in the input image (as processed by the training program)
- the minimum and maximum data file value in each band for each sample or cluster (minimum vector and maximum vector)
- the mean data file value in each band for each sample or cluster (mean vector)
- the covariance matrix for each sample or cluster
- the number of pixels in the sample or cluster

Nonparametric Signature

A nonparametric signature is based on an AOI that you define in the feature space image for the image file being classified. A nonparametric classifier uses a set of nonparametric signatures to assign pixels to a class based on their location, either inside or outside the area in the feature space image.



The format of the .sig file is described in the On-Line Help. Information on these statistics can be found in "Appendix A: Math Topics".

Evaluating Signatures

Once signatures are created, they can be evaluated, deleted, renamed, and merged with signatures from other files. Merging signatures enables you to perform complex classifications with signatures that are derived from more than one training method (supervised and/or unsupervised, parametric and/or nonparametric).



Use the Signature Editor to view the contents of each signature, manipulate signatures, and perform your own mathematical tests on the statistics.

Using Signature Data

There are tests to perform that can help determine whether the signature data are a true representation of the pixels to be classified for each class. You can evaluate signatures that were created either from supervised or unsupervised training. The evaluation methods in ERDAS IMAGINE include:

Classification

- Alarm—using your own pattern recognition ability, you view the estimated classified area for a signature (using the parallelepiped decision rule) against a display of the original image.
- Ellipse—view ellipse diagrams and scatterplots of data file values for every pair of bands.
- Contingency matrix—do a quick classification of the pixels in a set of training samples to see what percentage of the sample pixels are actually classified as expected. These percentages are presented in a contingency matrix. This method is for supervised training only, for which polygons of the training samples exist.
- Divergence—measure the divergence (statistical distance) between signatures and determine band subsets that maximize the classification.
- Statistics and histograms—analyze statistics and histograms of the signatures to make evaluations and comparisons.

NOTE: If the signature is nonparametric (i.e., a feature space signature), you can use only the alarm evaluation method.

After analyzing the signatures, it would be beneficial to merge or delete them, eliminate redundant bands from the data, add new bands of data, or perform any other operations to improve the classification.

Alarm

The alarm evaluation enables you to compare an estimated classification of one or more signatures against the original data, as it appears in the Viewer. According to the parallelepiped decision rule, the pixels that fit the classification criteria are highlighted in the displayed image. You also have the option to indicate an overlap by having it appear in a different color.

With this test, you can use your own pattern recognition skills, or some ground truth data, to determine the accuracy of a signature.



Use the Signature Alarm utility in the Signature Editor to perform n-dimensional alarms on the image in the Viewer, using the parallelepiped decision rule. The alarm utility creates a functional layer, and the Viewer allows you to toggle between the image layer and the functional layer.

Ellipse

In this evaluation, ellipses of concentration are calculated with the means and standard deviations stored in the signature file. It is also possible to generate parallelepiped rectangles, means, and labels.

In this evaluation, the mean and the standard deviation of every signature are used to represent the ellipse in 2-dimensional feature space. The ellipse is displayed in a feature space image.



Ellipses are explained and illustrated in "Appendix A: Math Topics" under the discussion of Scatterplots.

When the ellipses in the feature space image show extensive overlap, then the spectral characteristics of the pixels represented by the signatures cannot be distinguished in the two bands that are graphed. In the best case, there is no overlap. Some overlap, however, is expected.

Figure 6-7 shows how ellipses are plotted and how they can overlap. The first graph shows how the ellipses are plotted based on the range of 2 standard deviations from the mean. This range can be altered, changing the ellipse plots. Analyzing the plots with differing numbers of standard deviations is useful for determining the limits of a parallelepiped classification.

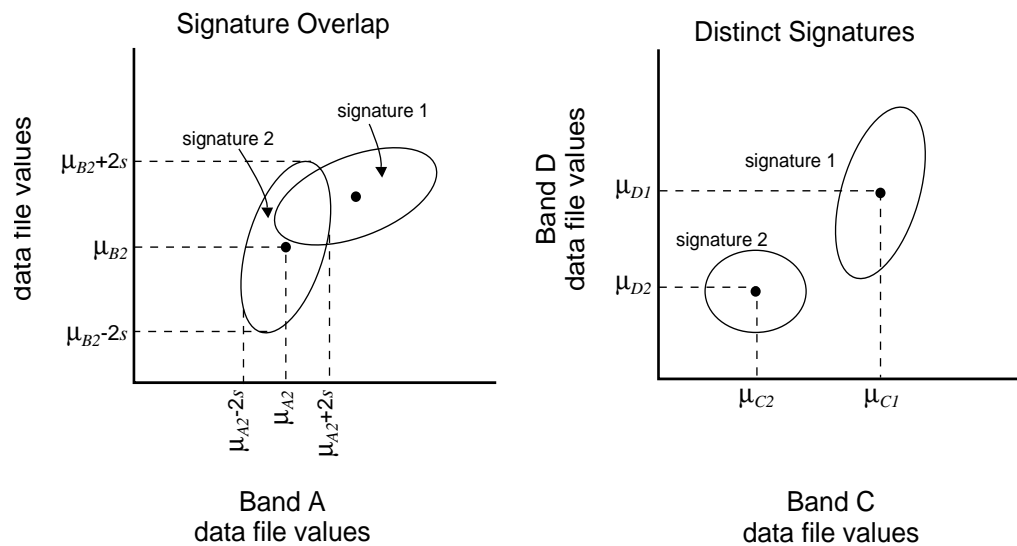


Figure 6-7 Ellipse Evaluation of Signatures

By analyzing the ellipse graphs for all band pairs, you can determine which signatures and which bands provide accurate classification results.



Use the Signature Editor to create a feature space image and to view an ellipse(s) of signature data.

Contingency Matrix

NOTE: This evaluation classifies all of the pixels in the selected AOIs and compares the results to the pixels of a training sample.

The pixels of each training sample are not always so homogeneous that every pixel in a sample is actually classified to its corresponding class. Each sample pixel only weights the statistics that determine the classes. However, if the signature statistics for each sample are distinct from those of the other samples, then a high percentage of each sample's pixels is classified as expected.

Classification

In this evaluation, a quick classification of the sample pixels is performed using the minimum distance, maximum likelihood, or Mahalanobis distance decision rule. Then, a contingency matrix is presented, which contains the number and percentages of pixels that are classified as expected.



Use the Signature Editor to perform the contingency matrix evaluation.

Separability

Signature separability is a statistical measure of distance between two signatures. Separability can be calculated for any combination of bands that is used in the classification, enabling you to rule out any bands that are not useful in the results of the classification.

For the distance (Euclidean) evaluation, the spectral distance between the mean vectors of each pair of signatures is computed. If the spectral distance between two samples is not significant for any pair of bands, then they may not be distinct enough to produce a successful classification.

The spectral distance is also the basis of the minimum distance classification (as explained below). Therefore, computing the distances between signatures helps you predict the results of a minimum distance classification.



Use the Signature Editor to compute signature separability and distance and automatically generate the report.

The formulas used to calculate separability are related to the maximum likelihood decision rule. Therefore, evaluating signature separability helps you predict the results of a maximum likelihood classification. The maximum likelihood decision rule is explained below.

There are three options for calculating the separability. All of these formulas take into account the covariances of the signatures in the bands being compared, as well as the mean vectors of the signatures.



Refer to "Appendix A: Math Topics" for information on the mean vector and covariance matrix.

Divergence

The formula for computing Divergence (D_{ij}) is as follows:

$$D_{ij} = \frac{1}{2} \text{tr}((C_i - C_j)(C_i^{-1} - C_j^{-1})) + \frac{1}{2} \text{tr}((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T)$$

Where:

- i and j = the two signatures (classes) being compared
- C_i = the covariance matrix of signature i
- μ_i = the mean vector of signature i
- tr= the trace function (matrix algebra)
- T = the transposition function

Source: Swain and Davis 1978

Transformed Divergence

The formula for computing Transformed Divergence (TD) is as follows:

$$D_{ij} = \frac{1}{2} \text{tr}((C_i - C_j)(C_i^{-1} - C_j^{-1})) + \frac{1}{2} \text{tr}((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T)$$

$$TD_{ij} = 2000 \left(1 - \exp\left(\frac{-D_{ij}}{8}\right) \right)$$

Where:

- i and j = the two signatures (classes) being compared
- C_i = the covariance matrix of signature i
- μ_i = the mean vector of signature i
- tr= the trace function (matrix algebra)
- T = the transposition function

Source: Swain and Davis 1978

According to Jensen, the transformed divergence “gives an exponentially decreasing weight to increasing distances between the classes.” The scale of the divergence values can range from 0 to 2,000. Interpreting your results after applying transformed divergence requires you to analyze those numerical divergence values. As a general rule, if the result is greater than 1,900, then the classes can be separated. Between 1,700 and 1,900, the separation is fairly good. Below 1,700, the separation is poor (Jensen 1996).

Jeffries-Matusita Distance

The formula for computing Jeffries-Matusita Distance (JM) is as follows:

$$\alpha = \frac{1}{8} (\mu_i - \mu_j)^T \left(\frac{C_i + C_j}{2} \right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{|(C_i + C_j)/2|}{\sqrt{|C_i| \times |C_j|}} \right)$$

Classification

$$JM_{ij} = \sqrt{2(1 - e^{-\alpha})}$$

Where:

- i and j = the two signatures (classes) being compared
- C_i = the covariance matrix of signature i
- μ_i = the mean vector of signature i
- \ln = the natural logarithm function
- $|C_i|$ = the determinant of C_i (matrix algebra)

Source: Swain and Davis 1978

According to Jensen, "The JM distance has a saturating behavior with increasing class separation like transformed divergence. However, it is not as computationally efficient as transformed divergence" (Jensen 1996).

Separability

Both transformed divergence and Jeffries-Matusita distance have upper and lower bounds. If the calculated divergence is equal to the appropriate upper bound, then the signatures can be said to be totally separable in the bands being studied. A calculated divergence of zero means that the signatures are inseparable.

- TD is between 0 and 2000.
- JM is between 0 and 1414.

A separability listing is a report of the computed divergence for every class pair and one band combination. The listing contains every divergence value for the bands studied for every possible pair of signatures.

The separability listing also contains the average divergence and the minimum divergence for the band set. These numbers can be compared to other separability listings (for other band combinations), to determine which set of bands is the most useful for classification.

Weight Factors

As with the Bayesian classifier (explained below with maximum likelihood), weight factors may be specified for each signature. These weight factors are based on *a priori* probabilities that any given pixel is assigned to each class. For example, if you know that twice as many pixels should be assigned to Class A as to Class B, then Class A should receive a weight factor that is twice that of Class B.

NOTE: The weight factors do not influence the divergence equations (for TD or JM), but they do influence the report of the best average and best minimum separability.

The weight factors for each signature are used to compute a weighted divergence with the following calculation:

$$W_{ij} = \frac{\sum_{i=1}^{c-1} \left(\sum_{j=i+1}^c f_i f_j U_{ij} \right)}{\frac{1}{2} \left[\left(\sum_{i=1}^c f_i \right)^2 - \sum_{i=1}^c f_i^2 \right]}$$

Where:

i and j = the two signatures (classes) being compared

U_{ij} = the unweighted divergence between i and j

W_{ij} = the weighted divergence between i and j

c = the number of signatures (classes)

f_i = the weight factor for signature i

Probability of Error

The Jeffries-Matusita distance is related to the pairwise probability of error, which is the probability that a pixel assigned to class i is actually in class j . Within a range, this probability can be estimated according to the expression below:

$$\frac{1}{16} (2 - JM_{ij}^2)^2 \leq P_e \leq 1 - \frac{1}{2} \left(1 + \frac{1}{2} JM_{ij}^2 \right)$$

Where:

i and j = the signatures (classes) being compared

JM_{ij} = the Jeffries-Matusita distance between i and j

P_e = the probability that a pixel is misclassified from i to j

Source: Swain and Davis 1978

Signature Manipulation

In many cases, training must be repeated several times before the desired signatures are produced. Signatures can be gathered from different sources—different training samples, feature space images, and different clustering programs—all using different techniques. After each signature file is evaluated, you may merge, delete, or create new signatures. The desired signatures can finally be moved to one signature file to be used in the classification.

The following operations upon signatures and signature files are possible with ERDAS IMAGINE:

- View the contents of the signature statistics
- View histograms of the samples or clusters that were used to derive the signatures
- Delete unwanted signatures

Classification

- Merge signatures together, so that they form one larger class when classified
- Append signatures from other files. You can combine signatures that are derived from different training methods for use in one classification.



Use the Signature Editor to view statistics and histogram listings and to delete, merge, append, and rename signatures within a signature file.

Classification Decision Rules

Once a set of reliable signatures has been created and evaluated, the next step is to perform a classification of the data. Each pixel is analyzed independently. The measurement vector for each pixel is compared to each signature, according to a decision rule, or algorithm. Pixels that pass the criteria that are established by the decision rule are then assigned to the class for that signature. ERDAS IMAGINE enables you to classify the data both parametrically with statistical representation, and nonparametrically as objects in feature space. Figure 6-8 shows the flow of an image pixel through the classification decision making process in ERDAS IMAGINE (Kloer 1994).

If a nonparametric rule is not set, then the pixel is classified using only the parametric rule. All of the parametric signatures are tested. If a nonparametric rule is set, the pixel is tested against all of the signatures with nonparametric definitions. This rule results in the following conditions:

- If the nonparametric test results in one unique class, the pixel is assigned to that class.
- If the nonparametric test results in zero classes (i.e., the pixel lies outside all the nonparametric decision boundaries), then the unclassified rule is applied. With this rule, the pixel is either classified by the parametric rule or left unclassified.
- If the pixel falls into more than one class as a result of the nonparametric test, the overlap rule is applied. With this rule, the pixel is either classified by the parametric rule, processing order, or left unclassified.

Nonparametric Rules

ERDAS IMAGINE provides these decision rules for nonparametric signatures:

- parallelepiped
- feature space

Unclassified Options

ERDAS IMAGINE provides these options if the pixel is not classified by the nonparametric rule:

- parametric rule
- unclassified

Overlap Options

ERDAS IMAGINE provides these options if the pixel falls into more than one feature space object:

- parametric rule
- by order
- unclassified

Parametric Rules

ERDAS IMAGINE provides these commonly-used decision rules for parametric signatures:

- minimum distance
- Mahalanobis distance
- maximum likelihood (with Bayesian variation)

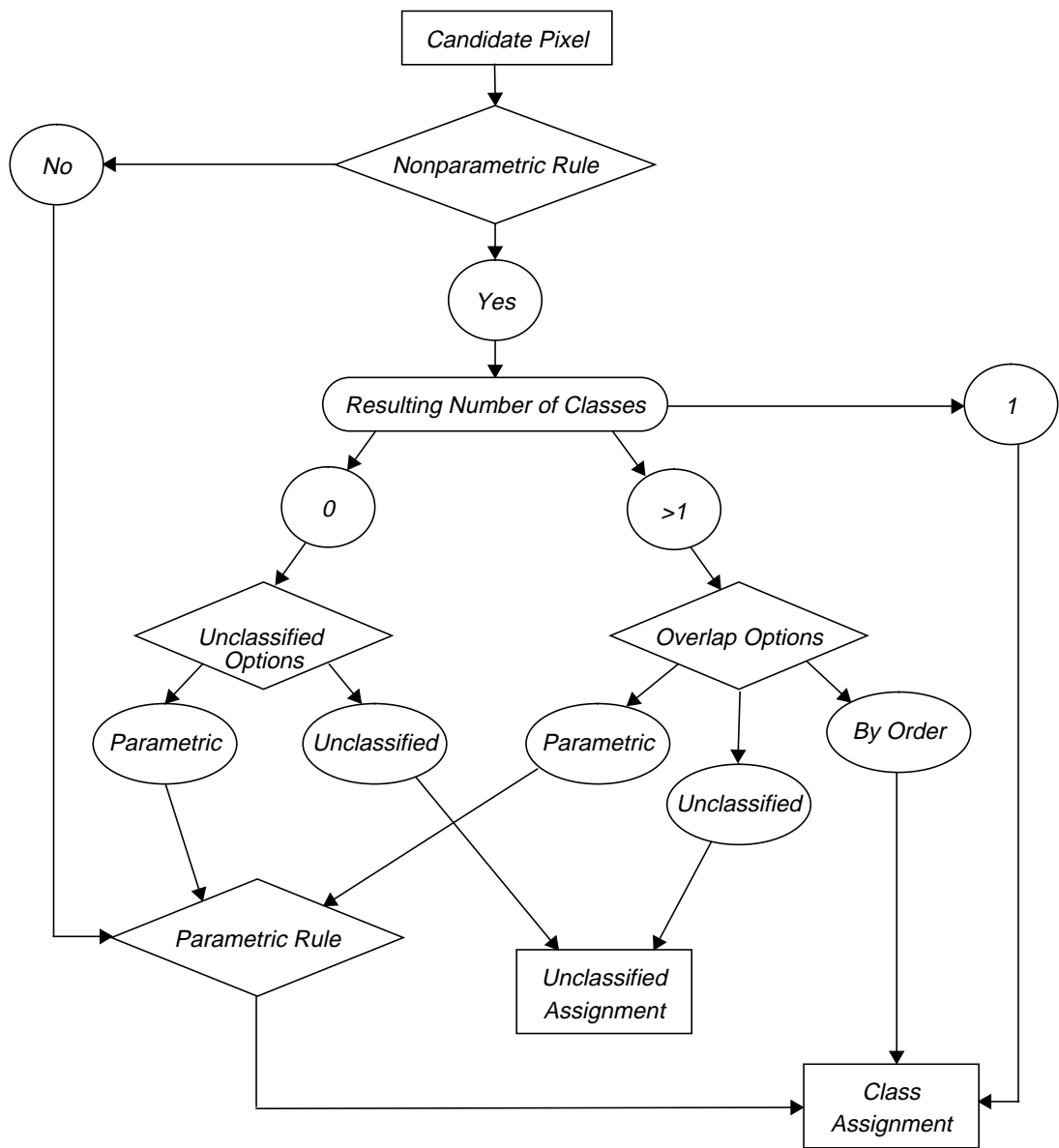


Figure 6-8 Classification Flow Diagram

Classification

Parallelepiped

In the parallelepiped decision rule, the data file values of the candidate pixel are compared to upper and lower limits. These limits can be either:

- the minimum and maximum data file values of each band in the signature,
- the mean of each band, plus and minus a number of standard deviations, or
- any limits that you specify, based on your knowledge of the data and signatures. This knowledge may come from the signature evaluation techniques discussed above.



These limits can be set using the Parallelepiped Limits utility in the Signature Editor.

There are high and low limits for every signature in every band. When a pixel's data file values are between the limits for every band in a signature, then the pixel is assigned to that signature's class. Figure 6-9 is a two-dimensional example of a parallelepiped classification.

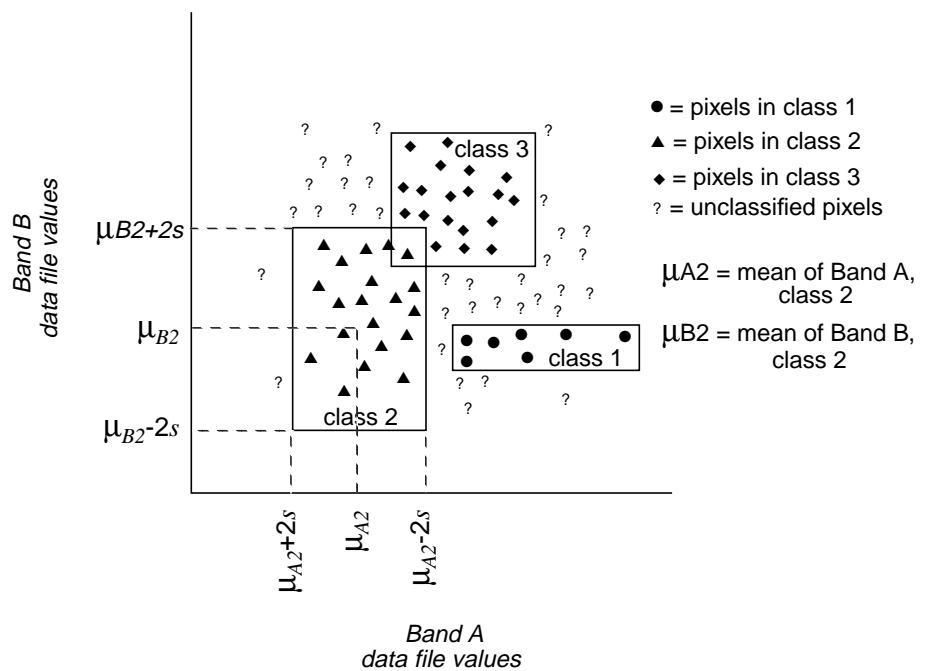


Figure 6-9 Parallelepiped Classification Using Two Standard Deviations as Limits

The large rectangles in Figure 6-9 are called parallelepipeds. They are the regions within the limits for each signature.

Overlap Region

In cases where a pixel may fall into the overlap region of two or more parallelepipeds, you must define how the pixel can be classified.

- The pixel can be classified by the order of the signatures. If one of the signatures is first and the other signature is fourth, the pixel is assigned to the first signature's class. This order can be set in the Signature Editor.
- The pixel can be classified by the defined parametric decision rule. The pixel is tested against the overlapping signatures only. If neither of these signatures is parametric, then the pixel is left unclassified. If only one of the signatures is parametric, then the pixel is automatically assigned to that signature's class.
- The pixel can be left unclassified.

Regions Outside of the Boundaries

If the pixel does not fall into one of the parallelepipeds, then you must define how the pixel can be classified.

- The pixel can be classified by the defined parametric decision rule. The pixel is tested against all of the parametric signatures. If none of the signatures is parametric, then the pixel is left unclassified.
- The pixel can be left unclassified.



Use the Supervised Classification utility in the Signature Editor to perform a parallelepiped classification.

Table 6-5 Parallelepiped Decision Rule

Advantages	Disadvantages
Fast and simple, since the data file values are compared to limits that remain constant for each band in each signature.	Since parallelepipeds have corners, pixels that are actually quite far, spectrally, from the mean of the signature may be classified. An example of this is shown in Figure 6-10.
Often useful for a first-pass, broad classification, this decision rule quickly narrows down the number of possible classes to which each pixel can be assigned before the more time-consuming calculations are made, thus cutting processing time (e.g., minimum distance, Mahalanobis distance, or maximum likelihood).	
Not dependent on normal distributions.	

Classification

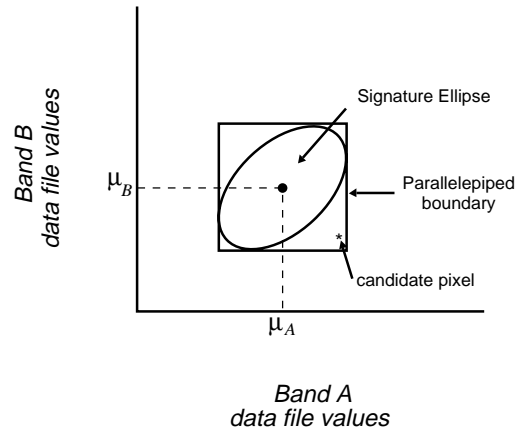


Figure 6-10 Parallelepiped Corners Compared to the Signature Ellipse

Feature Space

The feature space decision rule determines whether or not a candidate pixel lies within the nonparametric signature in the feature space image. When a pixel's data file values are in the feature space signature, then the pixel is assigned to that signature's class. Figure 6-11 is a two-dimensional example of a feature space classification. The polygons in this figure are AOIs used to define the feature space signatures.

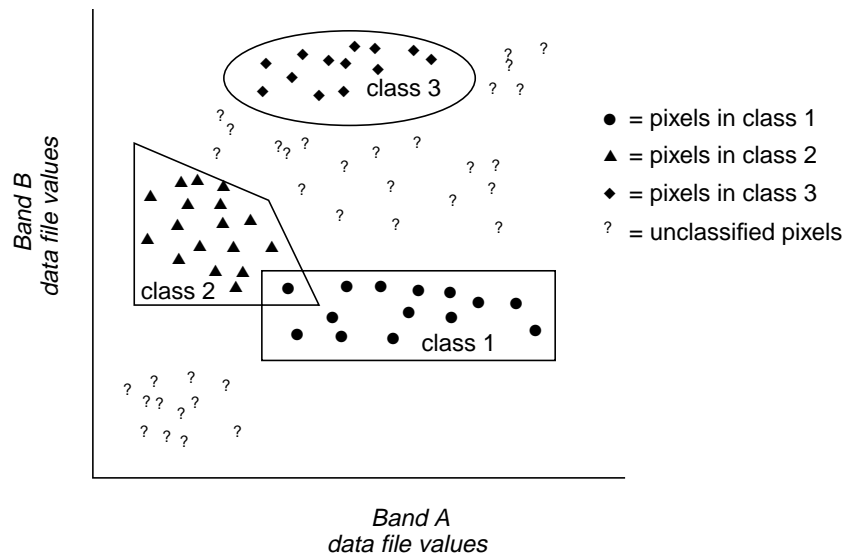


Figure 6-11 Feature Space Classification

Overlap Region

In cases where a pixel may fall into the overlap region of two or more AOIs, you must define how the pixel can be classified.

- The pixel can be classified by the order of the feature space signatures. If one of the signatures is first and the other signature is fourth, the pixel is assigned to the first signature’s class. This order can be set in the Signature Editor.
- The pixel can be classified by the defined parametric decision rule. The pixel is tested against the overlapping signatures only. If neither of these feature space signatures is parametric, then the pixel is left unclassified. If only one of the signatures is parametric, then the pixel is assigned automatically to that signature’s class.
- The pixel can be left unclassified.

Regions Outside of the AOIs

If the pixel does not fall into one of the AOIs for the feature space signatures, then you must define how the pixel can be classified.

- The pixel can be classified by the defined parametric decision rule. The pixel is tested against all of the parametric signatures. If none of the signatures is parametric, then the pixel is left unclassified.
- The pixel can be left unclassified.

Table 6-6 Feature Space Decision Rule

Advantages	Disadvantages
Often useful for a first-pass, broad classification.	The feature space decision rule allows overlap and unclassified pixels.
Provides an accurate way to classify a class with a nonnormal distribution (e.g., residential and urban).	The feature space image may be difficult to interpret.
Certain features may be more visually identifiable, which can help discriminate between classes that are spectrally similar and hard to differentiate with parametric information.	
The feature space method is fast.	



Use the Decision Rules utility in the Signature Editor to perform a feature space classification.

Minimum Distance

The minimum distance decision rule (also called spectral distance) calculates the spectral distance between the measurement vector for the candidate pixel and the mean vector for each signature.

Classification

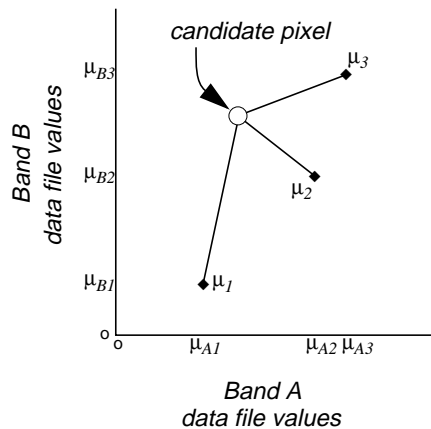


Figure 6-12 Minimum Spectral Distance

In Figure 6-12, spectral distance is illustrated by the lines from the candidate pixel to the means of the three signatures. The candidate pixel is assigned to the class with the closest mean.

The equation for classifying by spectral distance is based on the equation for Euclidean distance:

$$SD_{xyc} = \sqrt{\sum_{i=1}^n (\mu_{ci} - X_{xyi})^2}$$

Where:

n = number of bands (dimensions)

i = a particular band

c = a particular class

X_{xyi} = data file value of pixel x,y in band i

μ_{ci} = mean of data file values in band i for the sample for class c

SD_{xyc} = spectral distance from pixel x,y to the mean of class c

Source: Swain and Davis 1978

When spectral distance is computed for all possible values of c (all possible classes), the class of the candidate pixel is assigned to the class for which SD is the lowest.

Table 6-7 Minimum Distance Decision Rule

Advantages	Disadvantages
Since every pixel is spectrally closer to either one sample mean or another, there are no unclassified pixels.	Pixels that should be unclassified (i.e., they are not spectrally close to the mean of any sample, within limits that are reasonable to you) become classified. However, this problem is alleviated by thresholding out the pixels that are farthest from the means of their classes. (See the discussion of Thresholding on page 256.)
The fastest decision rule to compute, except for parallelepiped.	Does not consider class variability. For example, a class like an urban land cover class is made up of pixels with a high variance, which may tend to be farther from the mean of the signature. Using this decision rule, outlying urban pixels may be improperly classified. Inversely, a class with less variance, like water, may tend to overclassify (that is, classify more pixels than are appropriate to the class), because the pixels that belong to the class are usually spectrally closer to their mean than those of other classes to their means.

Mahalanobis Distance



The Mahalanobis distance algorithm assumes that the histograms of the bands have normal distributions. If this is not the case, you may have better results with the parallelepiped or minimum distance decision rule, or by performing a first-pass parallelepiped classification.

Mahalanobis distance is similar to minimum distance, except that the covariance matrix is used in the equation. Variance and covariance are figured in so that clusters that are highly varied lead to similarly varied classes, and vice versa. For example, when classifying urban areas—typically a class whose pixels vary widely—correctly classified pixels may be farther from the mean than those of a class for water, which is usually not a highly varied class (Swain and Davis 1978).

The equation for the Mahalanobis distance classifier is as follows:

$$D = (X - M_c)^T (Cov_c^{-1}) (X - M_c)$$

Where:

- D= Mahalanobis distance
- c= a particular class
- X= the measurement vector of the candidate pixel
- M_c= the mean vector of the signature of class c
- Cov_c= the covariance matrix of the pixels in the signature of class c
- Cov_c⁻¹= inverse of Cov_c
- T= transposition function

The pixel is assigned to the class, c, for which D is the lowest.

Table 6-8 Mahalanobis Decision Rule

Advantages	Disadvantages
Takes the variability of classes into account, unlike minimum distance or parallelepiped.	Tends to overclassify signatures with relatively large values in the covariance matrix. If there is a large dispersion of the pixels in a cluster or training sample, then the covariance matrix of that signature contains large values.
May be more useful than minimum distance in cases where statistical criteria (as expressed in the covariance matrix) must be taken into account, but the weighting factors that are available with the maximum likelihood/Bayesian option are not needed.	Slower to compute than parallelepiped or minimum distance.
	Mahalanobis distance is parametric, meaning that it relies heavily on a normal distribution of the data in each input band.

Maximum Likelihood/Bayesian



The maximum likelihood algorithm assumes that the histograms of the bands of data have normal distributions. If this is not the case, you may have better results with the parallelepiped or minimum distance decision rule, or by performing a first-pass parallelepiped classification.

The maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes, and that the input bands have normal distributions.

Bayesian Classifier

If you have *a priori* knowledge that the probabilities are not equal for all classes, you can specify weight factors for particular classes. This variation of the maximum likelihood decision rule is known as the Bayesian decision rule (Hord 1982). Unless you have *a priori* knowledge of the probabilities, it is recommended that they not be specified. In this case, these weights default to 1.0 in the equation.

The equation for the maximum likelihood/Bayesian classifier is as follows:

$$D = \ln(a_c) - [0.5 \ln(|Cov_c|)] - [0.5 (X-M_c)^T (Cov_c^{-1}) (X-M_c)]$$

Where:

- D = weighted distance (likelihood)
- c = a particular class
- X = the measurement vector of the candidate pixel
- M_c = the mean vector of the sample of class c
- a_c = percent probability that any candidate pixel is a member of class c (defaults to 1.0, or is entered from *a priori* knowledge)
- Cov_c = the covariance matrix of the pixels in the sample of class c
- $|Cov_c|$ = determinant of Cov_c (matrix algebra)
- Cov_c^{-1} = inverse of Cov_c (matrix algebra)
- \ln = natural logarithm function
- T = transposition function (matrix algebra)

The inverse and determinant of a matrix, along with the difference and transposition of vectors, would be explained in a textbook of matrix algebra.

The pixel is assigned to the class, c , for which D is the lowest.

Table 6-9 Maximum Likelihood/Bayesian Decision Rule

Advantages	Disadvantages
The most accurate of the classifiers in the ERDAS IMAGINE system (if the input samples/clusters have a normal distribution), because it takes the most variables into consideration.	An extensive equation that takes a long time to compute. The computation time increases with the number of input bands.
Takes the variability of classes into account by using the covariance matrix, as does Mahalanobis distance.	Maximum likelihood is parametric, meaning that it relies heavily on a normal distribution of the data in each input band.
	Tends to overclassify signatures with relatively large values in the covariance matrix. If there is a large dispersion of the pixels in a cluster or training sample, then the covariance matrix of that signature contains large values.

Fuzzy Methodology

Fuzzy Classification

The Fuzzy Classification method takes into account that there are pixels of mixed make-up, that is, a pixel cannot be definitively assigned to one category. Jensen notes that, “Clearly, there needs to be a way to make the classification algorithms more sensitive to the imprecise (fuzzy) nature of the real world” (Jensen 1996).

Fuzzy classification is designed to help you work with data that may not fall into exactly one category or another. Fuzzy classification works using a membership function, wherein a pixel’s value is determined by whether it is closer to one class than another. A fuzzy classification does not have definite boundaries, and each pixel can belong to several different classes (Jensen 1996).

Classification

Like traditional classification, fuzzy classification still uses training, but the biggest difference is that “it is also possible to obtain information on the various constituent classes found in a mixed pixel. . .” (Jensen 1996). Jensen goes on to explain that the process of collecting training sites in a fuzzy classification is not as strict as a traditional classification. In the fuzzy method, the training sites do not have to have pixels that are exactly the same.

Once you have a fuzzy classification, the fuzzy convolution utility allows you to perform a moving window convolution on a fuzzy classification with multiple output class assignments. Using the multilayer classification and distance file, the convolution creates a new single class output file by computing a total weighted distance for all classes in the window.

Fuzzy Convolution

The Fuzzy Convolution operation creates a single classification layer by calculating the total weighted inverse distance of all the classes in a window of pixels. Then, it assigns the center pixel in the class with the largest total inverse distance summed over the entire set of fuzzy classification layers.

This has the effect of creating a context-based classification to reduce the speckle or salt and pepper in the classification. Classes with a very small distance value remain unchanged while classes with higher distance values may change to a neighboring value if there is a sufficient number of neighboring pixels with class values and small corresponding distance values. The following equation is used in the calculation:

$$T[k] = \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} \sum_{l=0}^{n-1} \frac{w_{ij}}{D_{ijl}[k]}$$

Where:

i = row index of window

j = column index of window

s = size of window (3, 5, or 7)

l = layer index of fuzzy set

n = number of fuzzy layers used

W = weight table for window

k = class value

$D[k]$ = distance file value for class k

$T[k]$ = total weighted distance of window for class k

The center pixel is assigned the class with the maximum $T[k]$.

Expert Classification

Expert classification can be performed using the IMAGINE Expert Classifier™. The expert classification software provides a rules-based approach to multispectral image classification, post-classification refinement, and GIS modeling. In essence, an expert classification system is a hierarchy of rules, or a decision tree, that describes the conditions under which a set of low level constituent information gets abstracted into a set of high level informational classes. The constituent information consists of user-defined variables and includes raster imagery, vector coverages, spatial models, external programs, and simple scalars.

A rule is a conditional statement, or list of conditional statements, about the variable's data values and/or attributes that determine an informational component or hypotheses. Multiple rules and hypotheses can be linked together into a hierarchy that ultimately describes a final set of target informational classes or terminal hypotheses. Confidence values associated with each condition are also combined to provide a confidence image corresponding to the final output classified image.

The IMAGINE Expert Classifier is composed of two parts: the Knowledge Engineer and the Knowledge Classifier. The Knowledge Engineer provides the interface for an expert with first-hand knowledge of the data and the application to identify the variables, rules, and output classes of interest and create the hierarchical decision tree. The Knowledge Classifier provides an interface for a nonexpert to apply the knowledge base and create the output classification.

Knowledge Engineer

With the Knowledge Engineer, you can open knowledge bases, which are presented as decision trees in editing windows.

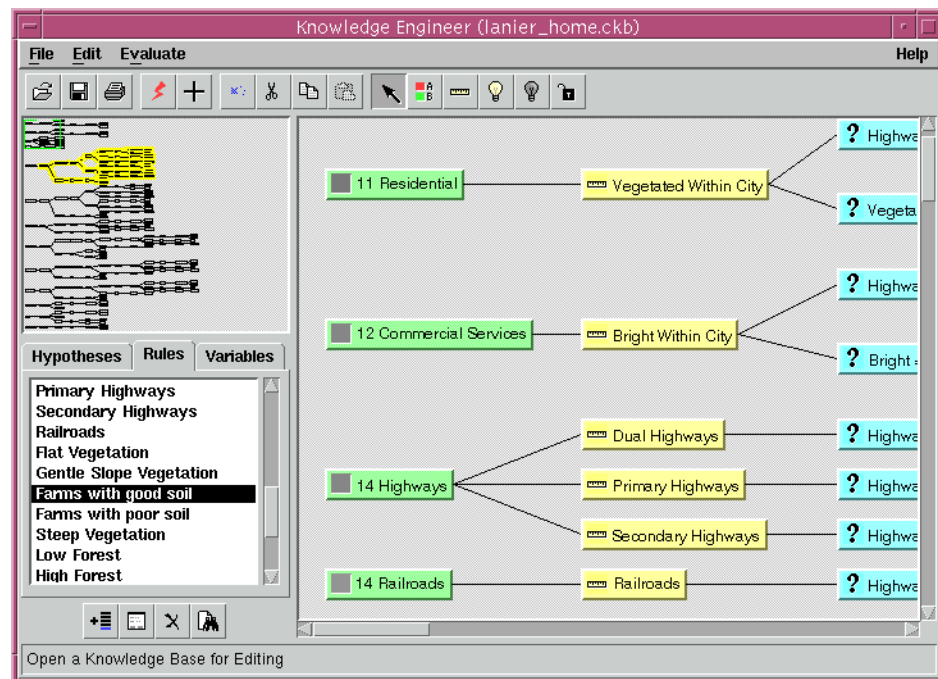


Figure 6-13 Knowledge Engineer Editing Window

Classification

In Figure 6-13, the upper left corner of the editing window is an overview of the entire decision tree with a green box indicating the position within the knowledge base of the currently displayed portion of the decision tree. This box can be dragged to change the view of the decision tree graphic in the display window on the right. The branch containing the currently selected hypotheses, rule, or condition is highlighted in the overview.

The decision tree grows in depth when the hypothesis of one rule is referred to by a condition of another rule. The terminal hypotheses of the decision tree represent the final classes of interest. Intermediate hypotheses may also be flagged as being a class of interest. This may occur when there is an association between classes.

Figure 6-14 represents a single branch of a decision tree depicting a hypothesis, its rule, and conditions.

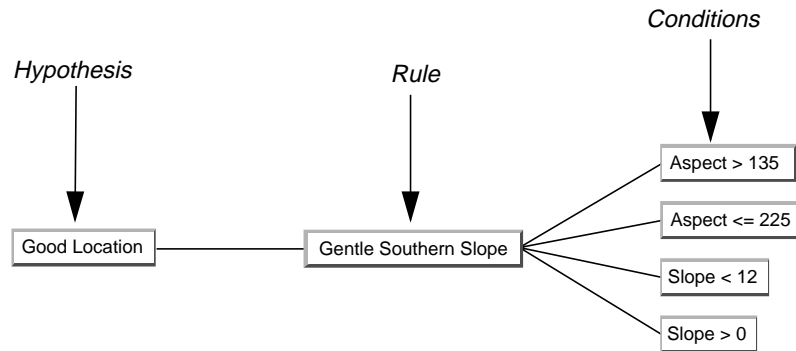


Figure 6-14 Example of a Decision Tree Branch

In this example, the rule, which is Gentle Southern Slope, determines the hypothesis, Good Location. The rule has four conditions depicted on the right side, all of which must be satisfied for the rule to be true.

However, the rule may be split if either Southern or Gentle slope defines the Good Location hypothesis. While both conditions must still be true to fire a rule, only one rule must be true to satisfy the hypothesis.

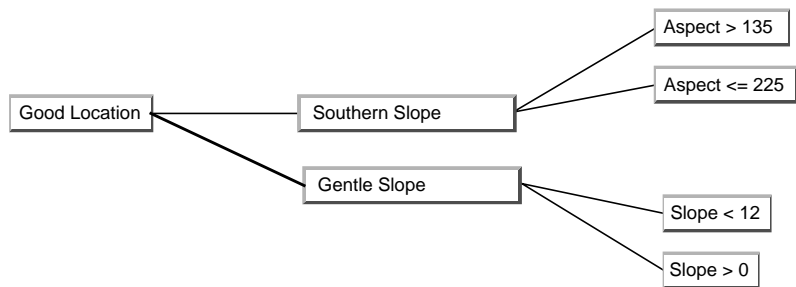


Figure 6-15 Split Rule Decision Tree Branch

Variable Editor

The Knowledge Engineer also makes use of a Variable Editor when classifying images. The Variable editor provides for the definition of the variable objects to be used in the rules conditions.

The two types of variables are raster and scalar. Raster variables may be defined by imagery, feature layers (including vector layers), graphic spatial models, or by running other programs. Scalar variables may be defined with an explicit value, or defined as the output from a model or external program.

Evaluating the Output of the Knowledge Engineer

The task of creating a useful, well-constructed knowledge base requires numerous iterations of trial, evaluation, and refinement. To facilitate this process, two options are provided. First, you can use the Test Classification to produce a test classification using the current knowledge base. Second, you can use the Classification Pathway Cursor to evaluate the results. This tool allows you to move a crosshair over the image in a Viewer to establish a confidence level for areas in the image.

Knowledge Classifier

The Knowledge Classifier is composed of two parts: an application with a user interface, and a command line executable. The user interface application allows you to input a limited set of parameters to control the use of the knowledge base. The user interface is designed as a wizard to lead you through pages of input parameters.

After selecting a knowledge base, you are prompted to select classes. The following is an example classes dialog:

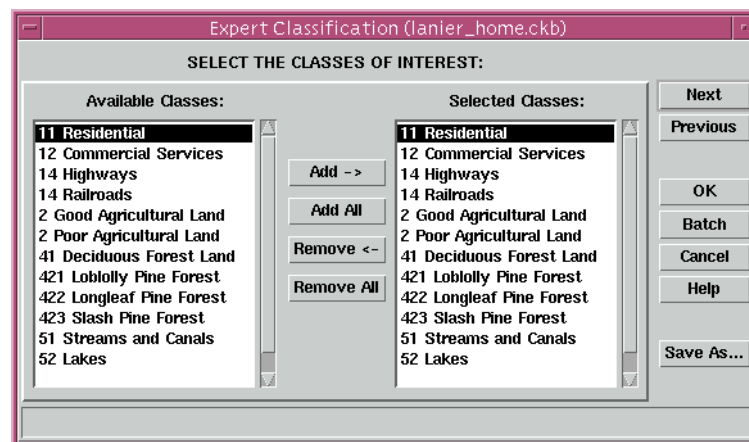


Figure 6-16 Knowledge Classifier Classes of Interest

After you select the input data for classification, the classification output options, output files, output area, output cell size, and output map projection, the Knowledge Classifier process can begin. An inference engine then evaluates all hypotheses at each location (calculating variable values, if required), and assigns the hypothesis with the highest confidence. The output of the Knowledge Classifier is a thematic image, and optionally, a confidence image.

Evaluating Classification

After a classification is performed, these methods are available for testing the accuracy of the classification:

- Thresholding—Use a probability image file to screen out misclassified pixels.
- Accuracy Assessment—Compare the classification to ground truth or other data.

Thresholding

Thresholding is the process of identifying the pixels in a classified image that are the most likely to be classified incorrectly. These pixels are put into another class (usually class 0). These pixels are identified statistically, based upon the distance measures that were used in the classification decision rule.

Distance File

When a minimum distance, Mahalanobis distance, or maximum likelihood classification is performed, a distance image file can be produced in addition to the output thematic raster layer. A distance image file is a one-band, 32-bit offset continuous raster layer in which each data file value represents the result of a spectral distance equation, depending upon the decision rule used.

- In a minimum distance classification, each distance value is the Euclidean spectral distance between the measurement vector of the pixel and the mean vector of the pixel's class.
- In a Mahalanobis distance or maximum likelihood classification, the distance value is the Mahalanobis distance between the measurement vector of the pixel and the mean vector of the pixel's class.

The brighter pixels (with the higher distance file values) are spectrally farther from the signature means for the classes to which they re assigned. They are more likely to be misclassified.

The darker pixels are spectrally nearer, and more likely to be classified correctly. If supervised training was used, the darkest pixels are usually the training samples.

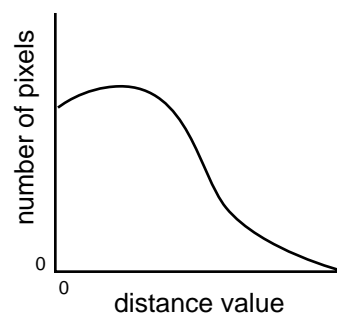


Figure 6-17 Histogram of a Distance Image

Figure 6-17 shows how the histogram of the distance image usually appears. This distribution is called a chi-square distribution, as opposed to a normal distribution, which is a symmetrical bell curve.

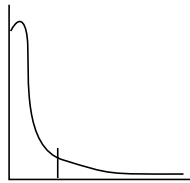
Threshold

The pixels that are the most likely to be misclassified have the higher distance file values at the tail of this histogram. At some point that you define—either mathematically or visually—the tail of this histogram is cut off. The cutoff point is the threshold.

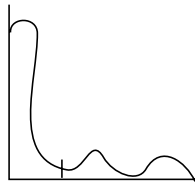
To determine the threshold:

- interactively change the threshold with the mouse, when a distance histogram is displayed while using the threshold function. This option enables you to select a chi-square value by selecting the cut-off value in the distance histogram, or
- input a chi-square parameter or distance measurement, so that the threshold can be calculated statistically.

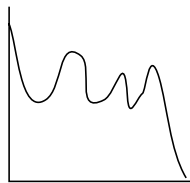
In both cases, thresholding has the effect of cutting the tail off of the histogram of the distance image file, representing the pixels with the highest distance values.



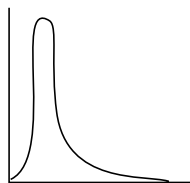
Smooth chi-square shape—try to find the breakpoint where the curve becomes more horizontal, and cut off the tail.



Minor mode(s) (peaks) in the curve probably indicate that the class picked up other features that were not represented in the signature. You probably want to threshold these features out.



Not a good class. The signature for this class probably represented a polymodal (multi-peaked) distribution.



Peak of the curve is shifted from 0. Indicates that the signature mean is off-center from the pixels it represents. You may need to take a new signature and reclassify.

Figure 6-18 Interactive Thresholding Tips

Classification

Figure 6-18 shows some example distance histograms. With each example is an explanation of what the curve might mean, and how to threshold it.

Chi-square Statistics

If the minimum distance classifier is used, then the threshold is simply a certain spectral distance. However, if Mahalanobis or maximum likelihood are used, then chi-square statistics are used to compare probabilities (Swain and Davis 1978).

When statistics are used to calculate the threshold, the threshold is more clearly defined as follows:

T is the distance value at which $C\%$ of the pixels in a class have a distance value greater than or equal to T .

Where:

T = the threshold for a class

$C\%$ = the percentage of pixels that are believed to be misclassified, known as the confidence level

T is related to the distance values by means of chi-square statistics. The value X^2 (chi-squared) is used in the equation. X^2 is a function of:

- the number of bands of data used—known in chi-square statistics as the number of degrees of freedom
- the confidence level

When classifying an image in ERDAS IMAGINE, the classified image automatically has the degrees of freedom (i.e., number of bands) used for the classification. The chi-square table is built into the threshold application.

NOTE: In this application of chi-square statistics, the value of X^2 is an approximation. Chi-square statistics are generally applied to independent variables (having no covariance), which is not usually true of image data.



A further discussion of chi-square statistics can be found in a statistics text.



Use the Classification Threshold utility to perform the thresholding.

Accuracy Assessment

Accuracy assessment is a general term for comparing the classification to geographical data that are assumed to be true, in order to determine the accuracy of the classification process. Usually, the assumed-true data are derived from ground truth data.

It is usually not practical to ground truth or otherwise test every pixel of a classified image. Therefore, a set of reference pixels is usually used. Reference pixels are points on the classified image for which actual data are (or will be) known. The reference pixels are randomly selected (Congalton 1991).

NOTE: You can use the ERDAS IMAGINE Accuracy Assessment utility to perform an accuracy assessment for any thematic layer. This layer does not have to be classified by ERDAS IMAGINE (e.g., you can run an accuracy assessment on a thematic layer that was classified in ERDAS Version 7.5 and imported into ERDAS IMAGINE).

Random Reference Pixels

When reference pixels are selected by the analyst, it is often tempting to select the same pixels for testing the classification that were used in the training samples. This biases the test, since the training samples are the basis of the classification. By allowing the reference pixels to be selected at random, the possibility of bias is lessened or eliminated (Congalton 1991).

The number of reference pixels is an important factor in determining the accuracy of the classification. It has been shown that more than 250 reference pixels are needed to estimate the mean accuracy of a class to within plus or minus five percent (Congalton 1991).

ERDAS IMAGINE uses a square window to select the reference pixels. The size of the window can be defined by you. Three different types of distribution are offered for selecting the random pixels:

- random—no rules are used
- stratified random—the number of points is stratified to the distribution of thematic layer classes
- equalized random—each class has an equal number of random points



Use the Accuracy Assessment utility to generate random reference points.

Accuracy Assessment CellArray

An Accuracy Assessment CellArray is created to compare the classified image with reference data. This CellArray is simply a list of class values for the pixels in the classified image file and the class values for the corresponding reference pixels. The class values for the reference pixels are input by you. The CellArray data reside in an image file.



Use the Accuracy Assessment CellArray to enter reference pixels for the class values.

Error Reports

From the Accuracy Assessment CellArray, two kinds of reports can be derived.

- The error matrix simply compares the reference points to the classified points in a $c \times c$ matrix, where c is the number of classes (including class 0).

Classification

- The accuracy report calculates statistics of the percentages of accuracy, based upon the results of the error matrix.

When interpreting the reports, it is important to observe the percentage of correctly classified pixels and to determine the nature of errors of the producer and yourself.



Use the Accuracy Assessment utility to generate the error matrix and accuracy reports.

Kappa Coefficient

The Kappa coefficient expresses the proportionate reduction in error generated by a classification process compared with the error of a completely random classification. For example, a value of .82 implies that the classification process is avoiding 82 percent of the errors that a completely random classification generates (Congalton 1991).



For more information on the Kappa coefficient, see a statistics manual.

Output File

When classifying an image file, the output file is an image file with a thematic raster layer. This file automatically contains the following data:

- class values
- class names
- color table
- statistics
- histogram

The image file also contains any signature attributes that were selected in the ERDAS IMAGINE Supervised Classification utility.



The class names, values, and colors can be set with the Signature Editor or the Raster Attribute Editor.

CHAPTER 7

Photogrammetric Concepts

Introduction

What is Photogrammetry?

Photogrammetry is the "art, science and technology of obtaining reliable information about physical objects and the environment through the process of recording, measuring and interpreting photographic images and patterns of electromagnetic radiant imagery and other phenomena" (ASP 1980).

Photogrammetry was invented in 1851 by Laussedat, and has continued to develop over the last 140 years. Over time, the development of photogrammetry has passed through the phases of plane table photogrammetry, analog photogrammetry, analytical photogrammetry, and has now entered the phase of digital photogrammetry (Konecny 1994).

The traditional, and largest, application of photogrammetry is to extract topographic information (e.g., topographic maps) from aerial images. However, photogrammetric techniques have also been applied to process satellite images and close range images in order to acquire topographic or nontopographic information of photographed objects.

Prior to the invention of the airplane, photographs taken on the ground were used to extract the relationships between objects using geometric principles. This was during the phase of plane table photogrammetry.

In analog photogrammetry, starting with stereomeasurement in 1901, optical or mechanical instruments were used to reconstruct three-dimensional geometry from two overlapping photographs. The main product during this phase was topographic maps.

In analytical photogrammetry, the computer replaces some expensive optical and mechanical components. The resulting devices were analog/digital hybrids. Analytical aerotriangulation, analytical plotters, and orthophoto projectors were the main developments during this phase. Outputs of analytical photogrammetry can be topographic maps, but can also be digital products, such as digital maps and DEMs.

Digital photogrammetry is photogrammetry as applied to digital images that are stored and processed on a computer. Digital images can be scanned from photographs or can be directly captured by digital cameras. Many photogrammetric tasks can be highly automated in digital photogrammetry (e.g., automatic DEM extraction and digital orthophoto generation). Digital photogrammetry is sometimes called softcopy photogrammetry. The output products are in digital form, such as digital maps, DEMs, and digital orthophotos saved on computer storage media. Therefore, they can be easily stored, managed, and applied by you. With the development of digital photogrammetry, photogrammetric techniques are more closely integrated into remote sensing and GIS.

Photogrammetric Concepts

Digital photogrammetric systems employ sophisticated software to automate the tasks associated with conventional photogrammetry, thereby minimizing the extent of manual interaction required to perform photogrammetric operations. IMAGINE OrthoBASE™ is such a photogrammetric system.

Photogrammetry can be used to measure and interpret information from hardcopy photographs or images. Sometimes the process of measuring information from photography and satellite imagery is considered metric photogrammetry, such as creating DEMs. Interpreting information from photography and imagery is considered interpretative photogrammetry, such as identifying and discriminating between various tree types as represented on a photograph or image (Wolf 1983).

Types of Photographs and Images

The types of photographs and images that can be processed within IMAGINE OrthoBASE include aerial, terrestrial, close range, and oblique. Aerial or vertical (near vertical) photographs and images are taken from a high vantage point above the Earth's surface. The camera axis of aerial or vertical photography is commonly directed vertically (or near vertically) down. Aerial photographs and images are commonly used for topographic and planimetric mapping projects. Aerial photographs and images are commonly captured from an aircraft or satellite.

Terrestrial or ground-based photographs and images are taken with the camera stationed on or close to the Earth's surface. Terrestrial and close range photographs and images are commonly used for applications involved with archeology, geomorphology, civil engineering, architecture, industry, etc.

Oblique photographs and images are similar to aerial photographs and images, except the camera axis is intentionally inclined at an angle with the vertical. Oblique photographs and images are commonly used for reconnaissance and corridor mapping applications.

Digital photogrammetric systems use digitized photographs or digital images as the primary source of input. Digital imagery can be obtained from various sources. These include:

- Digitizing existing hardcopy photographs
- Using digital cameras to record imagery
- Using sensors on board satellites such as Landsat and SPOT to record imagery



This document uses the term imagery in reference to photography and imagery obtained from various sources. This includes aerial and terrestrial photography, digital and video camera imagery, 35 mm photography, medium to large format photography, scanned photography, and satellite imagery.

Why use Photogrammetry?

As stated in the previous section, raw aerial photography and satellite imagery have large geometric distortion that is caused by various systematic and nonsystematic factors. The photogrammetric modeling based on collinearity equations eliminates these errors most efficiently, and creates the most reliable orthoimages from the raw imagery. It is unique in terms of considering the image-forming geometry, utilizing information between overlapping images, and explicitly dealing with the third dimension: elevation.

In addition to orthoimages, photogrammetry can also provide other geographic information such as a DEM, topographic features, and line maps reliably and efficiently. In essence, photogrammetry produces accurate and precise geographic information from a wide range of photographs and images. Any measurement taken on a photogrammetrically processed photograph or image reflects a measurement taken on the ground. Rather than constantly go to the field to measure distances, areas, angles, and point positions on the Earth's surface, photogrammetric tools allow for the accurate collection of information from imagery. Photogrammetric approaches for collecting geographic information save time and money, and maintain the highest accuracies.

**Photogrammetry vs.
Conventional Geometric
Correction**

Conventional techniques of geometric correction such as polynomial transformation are based on general functions not directly related to the specific distortion or error sources. They have been successful in the field of remote sensing and GIS applications, especially when dealing with low resolution and narrow field of view satellite imagery such as Landsat and SPOT data (Yang 1997). General functions have the advantage of simplicity. They can provide a reasonable geometric modeling alternative when little is known about the geometric nature of the image data.

However, conventional techniques generally process the images one at a time. They cannot provide an integrated solution for multiple images or photographs simultaneously and efficiently. It is very difficult, if not impossible, for conventional techniques to achieve a reasonable accuracy without a great number of GCPs when dealing with large-scale imagery, images having severe systematic and/or nonsystematic errors, and images covering rough terrain. Misalignment is more likely to occur when mosaicking separately rectified images. This misalignment could result in inaccurate geographic information being collected from the rectified images. Furthermore, it is impossible for a conventional technique to create a three-dimensional stereo model or to extract the elevation information from two overlapping images. There is no way for conventional techniques to accurately derive geometric information about the sensor that captured the imagery.

Photogrammetric techniques overcome all the problems mentioned above by using least squares bundle block adjustment. This solution is integrated and accurate.

IMAGINE OrthoBASE can process hundreds of images or photographs with very few GCPs, while at the same time eliminating the misalignment problem associated with creating image mosaics. In short, less time, less money, less manual effort, but more geographic fidelity can be realized using the photogrammetric solution.

**Single Frame
Orthorectification vs.
Block Triangulation**

Single frame orthorectification techniques orthorectify one image at a time using a technique known as space resection. In this respect, a minimum of three GCPs is required for each image. For example, in order to orthorectify 50 aerial photographs, a minimum of 150 GCPs is required. This includes manually identifying and measuring each GCP for each image individually. Once the GCPs are measured, space resection techniques compute the camera/sensor position and orientation as it existed at the time of data capture. This information, along with a DEM, is used to account for the negative impacts associated with geometric errors. Additional variables associated with systematic error are not considered.

Single frame orthorectification techniques do not utilize the internal relationship between adjacent images in a block to minimize and distribute the errors commonly associated with GCPs, image measurements, DEMs, and camera/sensor information. Therefore, during the mosaic procedure, misalignment between adjacent images is common since error has not been minimized and distributed throughout the block.

Aerial or block triangulation is the process of establishing a mathematical relationship between the images contained in a project, the camera or sensor model, and the ground. The information resulting from aerial triangulation is required as input for the orthorectification, DEM, and stereopair creation processes. The term aerial triangulation is commonly used when processing aerial photography and imagery. The term block triangulation, or simply triangulation, is used when processing satellite imagery. The techniques differ slightly as a function of the type of imagery being processed.

Classic aerial triangulation using optical-mechanical analog and analytical stereo plotters is primarily used for the collection of GCPs using a technique known as control point extension. Since the cost of collecting GCPs is very large, photogrammetric techniques are accepted as the ideal approach for collecting GCPs over large areas using photography rather than conventional ground surveying techniques. Control point extension involves the manual photo measurement of ground points appearing on overlapping images. These ground points are commonly referred to as tie points. Once the points are measured, the ground coordinates associated with the tie points can be determined using photogrammetric techniques employed by analog or analytical stereo plotters. These points are then referred to as ground control points (GCPs).

With the advent of digital photogrammetry, classic aerial triangulation has been extended to provide greater functionality. IMAGINE OrthoBASE uses a mathematical technique known as bundle block adjustment for aerial triangulation. Bundle block adjustment provides three primary functions:

- To determine the position and orientation for each image in a project as they existed at the time of photographic or image exposure. The resulting parameters are referred to as exterior orientation parameters. In order to estimate the exterior orientation parameters, a minimum of three GCPs is required for the entire block, regardless of how many images are contained within the project.
- To determine the ground coordinates of any tie points manually or automatically measured on the overlap areas of multiple images. The highly precise ground point determination of tie points is useful for generating control points from imagery in lieu of ground surveying techniques. Additionally, if a large number of ground points is generated, then a DEM can be interpolated using the Create Surface tool in ERDAS IMAGINE.
- To minimize and distribute the errors associated with the imagery, image measurements, GCPs, and so forth. The bundle block adjustment processes information from an entire block of imagery in one simultaneous solution (i.e., a bundle) using statistical techniques (i.e., adjustment component) to automatically identify, distribute, and remove error.

Because the images are processed in one step, the misalignment issues associated with creating mosaics are resolved.

Image and Data Acquisition

During photographic or image collection, overlapping images are exposed along a direction of flight. Most photogrammetric applications involve the use of overlapping images. In using more than one image, the geometry associated with the camera/sensor, image, and ground can be defined to greater accuracies and precision.

During the collection of imagery, each point in the flight path at which the camera exposes the film, or the sensor captures the imagery, is called an exposure station (see Figure 7-1).

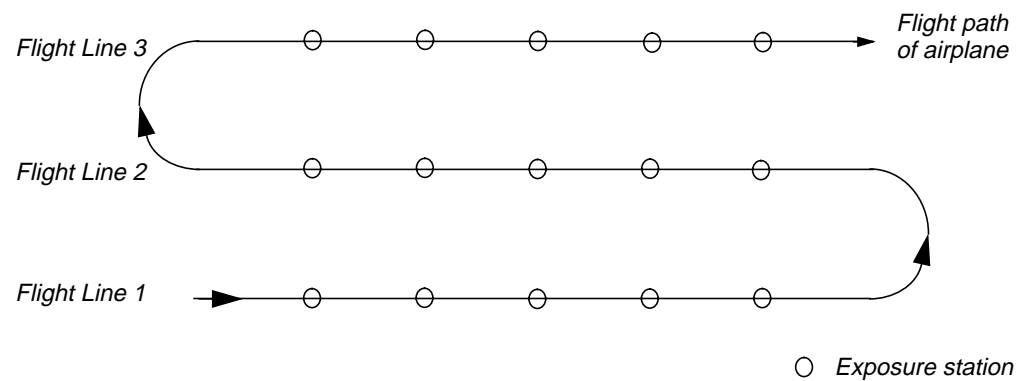


Figure 7-1 Exposure Stations Along a Flight Path

Each photograph or image that is exposed has a corresponding image scale associated with it. The image scale expresses the average ratio between a distance in the image and the same distance on the ground. It is computed as focal length divided by the flying height above the mean ground elevation. For example, with a flying height of 1000 m and a focal length of 15 cm, the image scale (SI) would be 1:6667.

NOTE: The flying height above ground is used, versus the altitude above sea level.

A strip of photographs consists of images captured along a flight line, normally with an overlap of 60%. All photos in the strip are assumed to be taken at approximately the same flying height and with a constant distance between exposure stations. Camera tilt relative to the vertical is assumed to be minimal.

The photographs from several flight paths can be combined to form a block of photographs. A block of photographs consists of a number of parallel strips, normally with a sidelap of 20-30%. Block triangulation techniques are used to transform all of the images in a block and ground points into a homologous coordinate system.

A regular block of photos is a rectangular block in which the number of photos in each strip is the same. Figure 7-2 shows a block of 5×2 photographs.

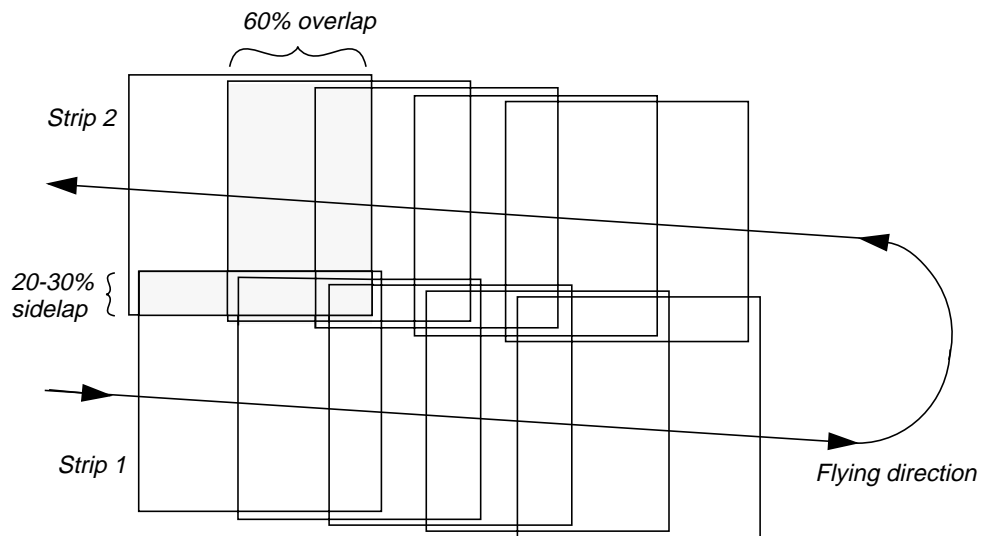


Figure 7-2 A Regular Rectangular Block of Aerial Photos

Photogrammetric Scanners

Photogrammetric quality scanners are special devices capable of high image quality and excellent positional accuracy. Use of this type of scanner results in geometric accuracies similar to traditional analog and analytical photogrammetric instruments. These scanners are necessary for digital photogrammetric applications that have high accuracy requirements.

These units usually scan only film because film is superior to paper, both in terms of image detail and geometry. These units usually have a Root Mean Square Error (RMSE) positional accuracy of 4 microns or less, and are capable of scanning at a maximum resolution of 5 to 10 microns (5 microns is equivalent to approximately 5,000 pixels per inch).

The required pixel resolution varies depending on the application. Aerial triangulation and feature collection applications often scan in the 10- to 15-micron range. Orthophoto applications often use 15- to 30-micron pixels. Color film is less sharp than panchromatic, therefore color ortho applications often use 20- to 40-micron pixels.

Desktop Scanners

Desktop scanners are general purpose devices. They lack the image detail and geometric accuracy of photogrammetric quality units, but they are much less expensive. When using a desktop scanner, you should make sure that the active area is at least 9 X 9 inches (i.e., A3 type scanners), enabling you to capture the entire photo frame.

Desktop scanners are appropriate for less rigorous uses, such as digital photogrammetry in support of GIS or remote sensing applications. Calibrating these units improves geometric accuracy, but the results are still inferior to photogrammetric units. The image correlation techniques that are necessary for automatic tie point collection and elevation extraction are often sensitive to scan quality. Therefore, errors can be introduced into the photogrammetric solution that are attributable to scanning errors. IMAGINE OrthoBASE accounts for systematic errors attributed to scanning errors.

Scanning Resolutions

One of the primary factors contributing to the overall accuracy of block triangulation and orthorectification is the resolution of the imagery being used. Image resolution is commonly determined by the scanning resolution (if film photography is being used), or by the pixel resolution of the sensor. In order to optimize the attainable accuracy of a solution, the scanning resolution must be considered. The appropriate scanning resolution is determined by balancing the accuracy requirements versus the size of the mapping project and the time required to process the project. Table 7-1, "Scanning Resolutions," on page 268 lists the scanning resolutions associated with various scales of photography and image file size.

Table 7-1 Scanning Resolutions

	12 microns (2117 dpi¹)	16 microns (1588 dpi)	25 microns (1016 dpi)	50 microns (508 dpi)	85 microns (300 dpi)
Photo Scale 1 to	Ground Coverage (meters)	Ground Coverage (meters)	Ground Coverage (meters)	Ground Coverage (meters)	Ground Coverage (meters)
1800	0.0216	0.0288	0.045	0.09	0.153
2400	0.0288	0.0384	0.06	0.12	0.204
3000	0.036	0.048	0.075	0.15	0.255
3600	0.0432	0.0576	0.09	0.18	0.306
4200	0.0504	0.0672	0.105	0.21	0.357
4800	0.0576	0.0768	0.12	0.24	0.408
5400	0.0648	0.0864	0.135	0.27	0.459
6000	0.072	0.096	0.15	0.3	0.51
6600	0.0792	0.1056	0.165	0.33	0.561
7200	0.0864	0.1152	0.18	0.36	0.612
7800	0.0936	0.1248	0.195	0.39	0.663
8400	0.1008	0.1344	0.21	0.42	0.714
9000	0.108	0.144	0.225	0.45	0.765
9600	0.1152	0.1536	0.24	0.48	0.816
10800	0.1296	0.1728	0.27	0.54	0.918
12000	0.144	0.192	0.3	0.6	1.02
15000	0.18	0.24	0.375	0.75	1.275
18000	0.216	0.288	0.45	0.9	1.53
24000	0.288	0.384	0.6	1.2	2.04
30000	0.36	0.48	0.75	1.5	2.55
40000	0.48	0.64	1	2	3.4
50000	0.6	0.8	1.25	2.5	4.25
60000	0.72	0.96	1.5	3	5.1
B/W File Size (MB)	363	204	84	21	7
Color File Size (MB)	1089	612	252	63	21

¹ dots per inch

The ground coverage column refers to the ground coverage per pixel. Thus, a 1:40000 scale photograph scanned at 25 microns [1016 dots per inch (dpi)] has a ground coverage per pixel of $1 \text{ m} \times 1 \text{ m}$. The resulting file size is approximately 85 MB, assuming a square 9×9 inch photograph.

Coordinate Systems

Conceptually, photogrammetry involves establishing the relationship between the camera or sensor used to capture imagery, the imagery itself, and the ground. In order to understand and define this relationship, each of the three variables associated with the relationship must be defined with respect to a coordinate space and coordinate system.

Pixel Coordinate System

The file coordinates of a digital image are defined in a pixel coordinate system. A pixel coordinate system is usually a coordinate system with its origin in the upper-left corner of the image, the x-axis pointing to the right, the y-axis pointing downward, and the unit in pixels, as shown by axis c and r in Figure 7-3. These file coordinates (c, r) can also be thought of as the pixel column and row number. This coordinate system is referenced as pixel coordinates (c, r) in this chapter.

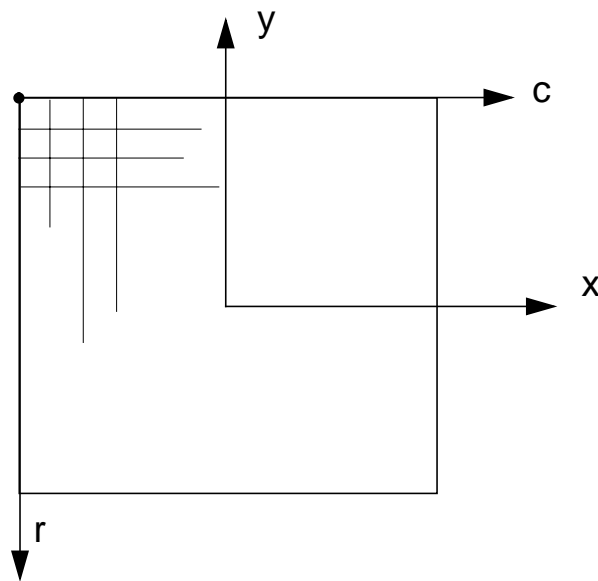


Figure 7-3 Pixel Coordinates and Image Coordinates

Image Coordinate System

An image coordinate system or an image plane coordinate system is usually defined as a two-dimensional coordinate system occurring on the image plane with its origin at the image center, normally at the principal point or at the intersection of the fiducial marks as illustrated by axis x and y in Figure 7-3. Image coordinates are used to describe positions on the film plane. Image coordinate units are usually millimeters or microns. This coordinate system is referenced as image coordinates (x, y) in this chapter.

Image Space Coordinate System

An image space coordinate system (Figure 7-4) is identical to image coordinates, except that it adds a third axis (z). The origin of the image space coordinate system is defined at the perspective center S as shown in Figure 7-4. Its x -axis and y -axis are parallel to the x -axis and y -axis in the image plane coordinate system. The z -axis is the optical axis, therefore the z value of an image point in the image space coordinate system is usually equal to $-f$ (focal length). Image space coordinates are used to describe positions inside the camera and usually use units in millimeters or microns. This coordinate system is referenced as image space coordinates (x, y, z) in this chapter.

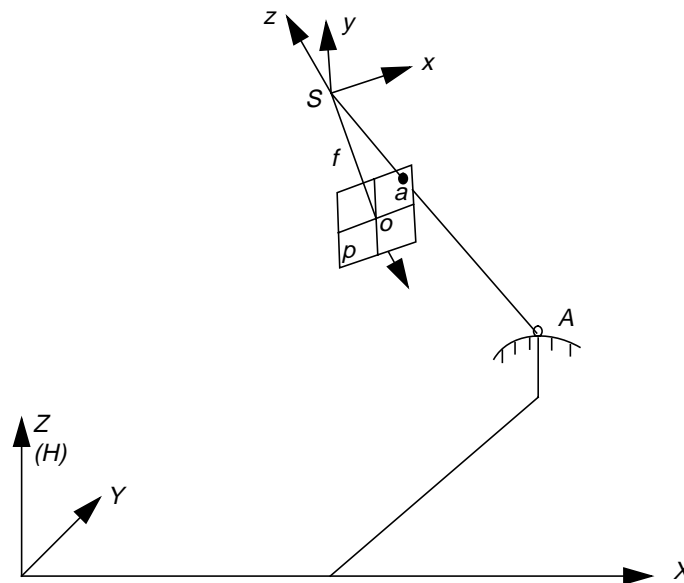


Figure 7-4 Image Space Coordinate System

Ground Coordinate System

A ground coordinate system is usually defined as a three-dimensional coordinate system that utilizes a known map projection. Ground coordinates (X, Y, Z) are usually expressed in feet or meters. The Z value is elevation above mean sea level for a given vertical datum. This coordinate system is referenced as ground coordinates (X, Y, Z) in this chapter.

Geocentric and Topocentric Coordinate System

Most photogrammetric applications account for the Earth's curvature in their calculations. This is done by adding a correction value or by computing geometry in a coordinate system which includes curvature. Two such systems are geocentric and topocentric coordinates.

A geocentric coordinate system has its origin at the center of the Earth ellipsoid. The Z -axis equals the rotational axis of the Earth, and the X -axis passes through the Greenwich meridian. The Y -axis is perpendicular to both the Z -axis and X -axis, so as to create a three-dimensional coordinate system that follows the right hand rule.

A topocentric coordinate system has its origin at the center of the image projected on the Earth ellipsoid. The three perpendicular coordinate axes are defined on a tangential plane at this center point. The plane is called the reference plane or the local datum. The x-axis is oriented eastward, the y-axis northward, and the z-axis is vertical to the reference plane (up).

For simplicity of presentation, the remainder of this chapter does not explicitly reference geocentric or topocentric coordinates. Basic photogrammetric principles can be presented without adding this additional level of complexity.

Terrestrial Photography

Photogrammetric applications associated with terrestrial or ground-based images utilize slightly different image and ground space coordinate systems. Figure 7-5 illustrates the two coordinate systems associated with image space and ground space.

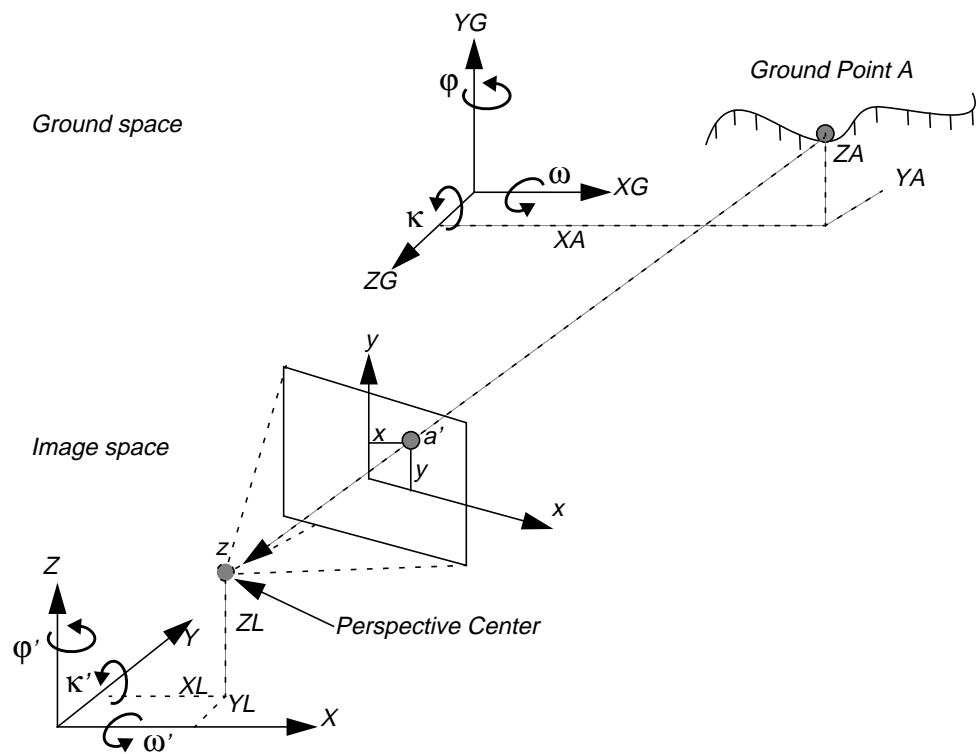


Figure 7-5 Terrestrial Photography

The image and ground space coordinate systems are right-handed coordinate systems. Most terrestrial applications use a ground space coordinate system that was defined using a localized Cartesian coordinate system.

Photogrammetric Concepts

The image space coordinate system directs the z-axis toward the imaged object and the y-axis directed North up. The image x-axis is similar to that used in aerial applications. The X_L , Y_L , and Z_L coordinates define the position of the perspective center as it existed at the time of image capture. The ground coordinates of ground point A (X_A , Y_A , and Z_A) are defined within the ground space coordinate system (X_G , Y_G , and Z_G). With this definition, the rotation angles ω , ϕ , and κ are still defined as in the aerial photography conventions. In IMAGINE OrthoBASE, you can also use the ground (X , Y , Z) coordinate system to directly define GCPs. Thus, GCPs do not need to be transformed. Then the definition of rotation angles ω' , ϕ' , and κ' are different, as shown in Figure 7-5.

Interior Orientation

Interior orientation defines the internal geometry of a camera or sensor as it existed at the time of data capture. The variables associated with image space are defined during the process of interior orientation. Interior orientation is primarily used to transform the image pixel coordinate system or other image coordinate measurement system to the image space coordinate system.

Figure 7-6 illustrates the variables associated with the internal geometry of an image captured from an aerial camera, where o represents the principal point and a represents an image point.

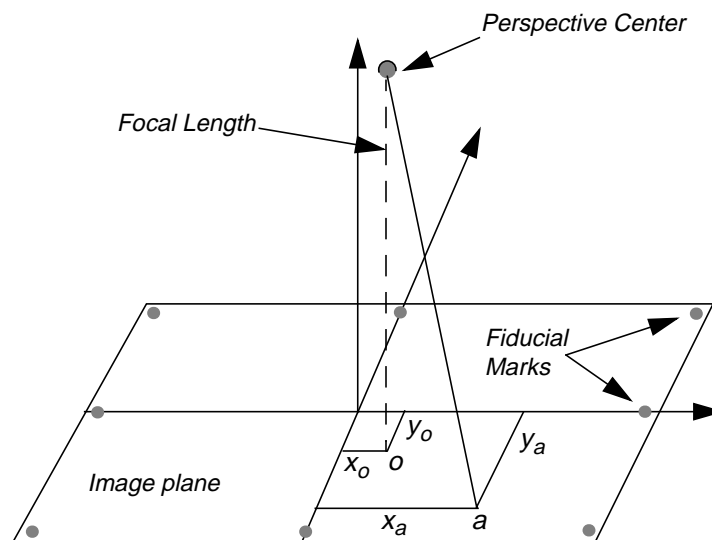


Figure 7-6 Internal Geometry

The internal geometry of a camera is defined by specifying the following variables:

- Principal point
- Focal length
- Fiducial marks
- Lens distortion

Principal Point and Focal Length

The principal point is mathematically defined as the intersection of the perpendicular line through the perspective center of the image plane. The length from the principal point to the perspective center is called the focal length (Wang 1990).

The image plane is commonly referred to as the focal plane. For wide-angle aerial cameras, the focal length is approximately 152 mm, or 6 inches. For some digital cameras, the focal length is 28 mm. Prior to conducting photogrammetric projects, the focal length of a metric camera is accurately determined or calibrated in a laboratory environment.

This mathematical definition is the basis of triangulation, but difficult to determine optically. The optical definition of principal point is the image position where the optical axis intersects the image plane. In the laboratory, this is calibrated in two forms: principal point of autocollimation and principal point of symmetry, which can be seen from the camera calibration report. Most applications prefer to use the principal point of symmetry since it can best compensate for the lens distortion.

Fiducial Marks

As stated previously, one of the steps associated with interior orientation involves determining the image position of the principal point for each image in the project. Therefore, the image positions of the fiducial marks are measured on the image, and subsequently compared to the calibrated coordinates of each fiducial mark.

Since the image space coordinate system has not yet been defined for each image, the measured image coordinates of the fiducial marks are referenced to a pixel or file coordinate system. The pixel coordinate system has an x coordinate (column) and a y coordinate (row). The origin of the pixel coordinate system is the upper left corner of the image having a row and column value of 0 and 0, respectively. Figure 7-7 illustrates the difference between the pixel coordinate system and the image space coordinate system.

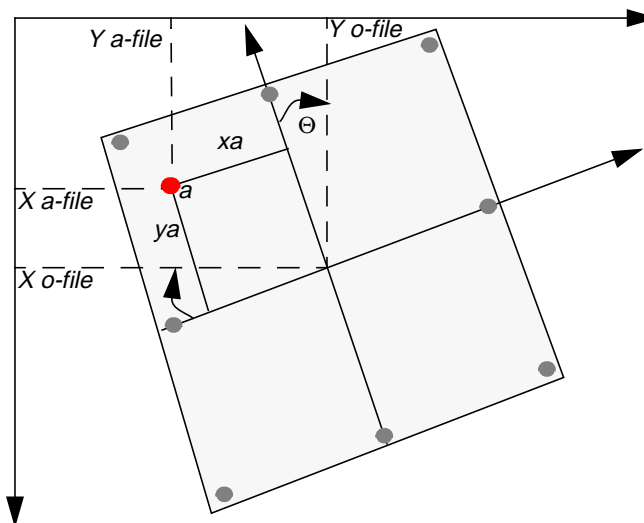


Figure 7-7 Pixel Coordinate System vs. Image Space Coordinate System

Photogrammetric Concepts

Using a two-dimensional affine transformation, the relationship between the pixel coordinate system and the image space coordinate system is defined. The following two-dimensional affine transformation equations can be used to determine the coefficients required to transform pixel coordinate measurements to the image coordinates:

$$x = a_1 + a_2X + a_3Y$$

$$y = b_1 + b_2X + b_3Y$$

The x and y image coordinates associated with the calibrated fiducial marks and the X and Y pixel coordinates of the measured fiducial marks are used to determine six affine transformation coefficients. The resulting six coefficients can then be used to transform each set of row (y) and column (x) pixel coordinates to image coordinates.

The quality of the two-dimensional affine transformation is represented using a root mean square (RMS) error. The RMS error represents the degree of correspondence between the calibrated fiducial mark coordinates and their respective measured image coordinate values. Large RMS errors indicate poor correspondence. This can be attributed to film deformation, poor scanning quality, out-of-date calibration information, or image mismeasurement.

The affine transformation also defines the translation between the origin of the pixel coordinate system and the image coordinate system (x_{o-file} and y_{o-file}). Additionally, the affine transformation takes into consideration rotation of the image coordinate system by considering angle Θ (see Figure 7-7). A scanned image of an aerial photograph is normally rotated due to the scanning procedure.

The degree of variation between the x - and y -axis is referred to as nonorthogonality. The two-dimensional affine transformation also considers the extent of nonorthogonality. The scale difference between the x -axis and the y -axis is also considered using the affine transformation.

Lens Distortion

Lens distortion deteriorates the positional accuracy of image points located on the image plane. Two types of radial lens distortion exist: radial and tangential lens distortion. Lens distortion occurs when light rays passing through the lens are bent, thereby changing directions and intersecting the image plane at positions deviant from the norm. Figure 7-8 illustrates the difference between radial and tangential lens distortion.

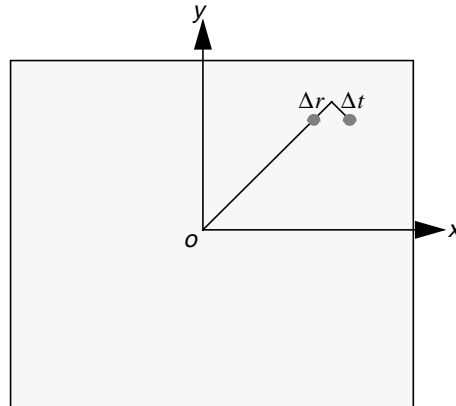


Figure 7-8 Radial vs. Tangential Lens Distortion

Radial lens distortion causes imaged points to be distorted along radial lines from the principal point o . The effect of radial lens distortion is represented as Δr . Radial lens distortion is also commonly referred to as symmetric lens distortion. Tangential lens distortion occurs at right angles to the radial lines from the principal point. The effect of tangential lens distortion is represented as Δt . Since tangential lens distortion is much smaller in magnitude than radial lens distortion, it is considered negligible.

The effects of lens distortion are commonly determined in a laboratory during the camera calibration procedure.

The effects of radial lens distortion throughout an image can be approximated using a polynomial. The following polynomial is used to determine coefficients associated with radial lens distortion:

$$\Delta r = k_0 r + k_1 r^3 + k_2 r^5$$

Δr represents the radial distortion along a radial distance r from the principal point (Wolf 1983). In most camera calibration reports, the lens distortion value is provided as a function of radial distance from the principal point or field angle. IMAGINE OrthoBASE accommodates radial lens distortion parameters in both scenarios.

Three coefficients (k_0 , k_1 , and k_2) are computed using statistical techniques. Once the coefficients are computed, each measurement taken on an image is corrected for radial lens distortion.

Exterior Orientation

Exterior orientation defines the position and angular orientation associated with an image. The variables defining the position and orientation of an image are referred to as the elements of exterior orientation. The elements of exterior orientation define the characteristics associated with an image at the time of exposure or capture. The positional elements of exterior orientation include X_0 , Y_0 , and Z_0 . They define the position of the perspective center (O) with respect to the ground space coordinate system (X , Y , and Z). Z_0 is commonly referred to as the height of the camera above sea level, which is commonly defined by a datum.

The angular or rotational elements of exterior orientation describe the relationship between the ground space coordinate system (X , Y , and Z) and the image space coordinate system (x , y , and z). Three rotation angles are commonly used to define angular orientation. They are omega (ω), phi (ϕ), and kappa (κ). Figure 7-9 illustrates the elements of exterior orientation.

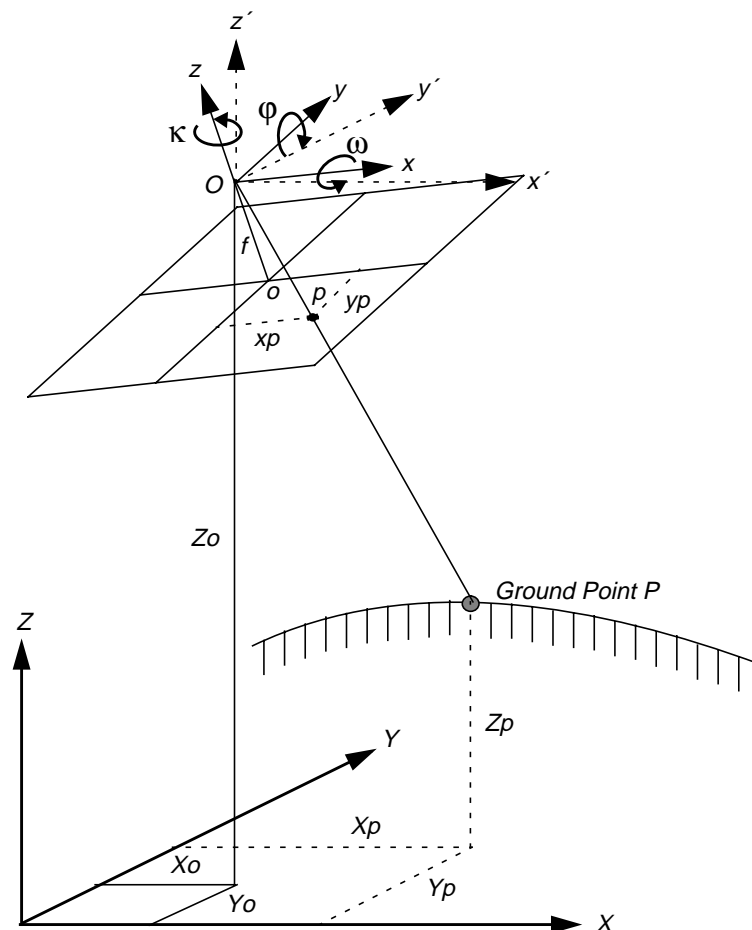


Figure 7-9 Elements of Exterior Orientation

Omega is a rotation about the photographic x-axis, phi is a rotation about the photographic y-axis, and kappa is a rotation about the photographic z-axis, which are defined as being positive if they are counterclockwise when viewed from the positive end of their respective axis. Different conventions are used to define the order and direction of the three rotation angles (Wang 1990). The ISPRS recommends the use of the ω , ϕ , κ convention. The photographic z-axis is equivalent to the optical axis (focal length). The x' , y' , and z' coordinates are parallel to the ground space coordinate system.

Using the three rotation angles, the relationship between the image space coordinate system (x , y , and z) and ground space coordinate system (X , Y , and Z or x' , y' , and z') can be determined. A 3 X 3 matrix defining the relationship between the two systems is used. This is referred to as the orientation or rotation matrix, M . The rotation matrix can be defined as follows:

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

The rotation matrix is derived by applying a sequential rotation of omega about the x-axis, phi about the y-axis, and kappa about the z-axis.

The Collinearity Equation

The following section defines the relationship between the camera/sensor, the image, and the ground. Most photogrammetric tools utilize the following formulations in one form or another.

With reference to Figure 7-9, an image vector a can be defined as the vector from the exposure station O to the image point p . A ground space or object space vector A can be defined as the vector from the exposure station O to the ground point P . The image vector and ground vector are collinear, inferring that a line extending from the exposure station to the image point and to the ground is linear.

The image vector and ground vector are only collinear if one is a scalar multiple of the other. Therefore, the following statement can be made:

$$a = kA$$

Where k is a scalar multiple. The image and ground vectors must be within the same coordinate system. Therefore, image vector a is comprised of the following components:

$$a = \begin{bmatrix} xp - xo \\ yp - yo \\ -f \end{bmatrix}$$

Where xo and yo represent the image coordinates of the principal point.

Photogrammetric Concepts

Similarly, the ground vector can be formulated as follows:

$$A = \begin{bmatrix} X_p - X_o \\ Y_p - Y_o \\ Z_p - Z_o \end{bmatrix}$$

In order for the image and ground vectors to be within the same coordinate system, the ground vector must be multiplied by the rotation matrix M . The following equation can be formulated:

$$a = kMA$$

Where:

$$\begin{bmatrix} xp - x_o \\ yp - y_o \\ -f \end{bmatrix} = kM \begin{bmatrix} X_p - X_o \\ Y_p - Y_o \\ Z_p - Z_o \end{bmatrix}$$

The above equation defines the relationship between the perspective center of the camera/sensor exposure station and ground point P appearing on an image with an image point location of p . This equation forms the basis of the collinearity condition that is used in most photogrammetric operations. The collinearity condition specifies that the exposure station, ground point, and its corresponding image point location must all lie along a straight line, thereby being collinear. Two equations comprise the collinearity condition.

$$x_p - x_o = -f \left[\frac{m_{11}(X_p - X_{o_1}) + m_{12}(Y_p - Y_{o_1}) + m_{13}(Z_p - Z_{o_1})}{m_{31}(X_p - X_{o_1}) + m_{32}(Y_p - Y_{o_1}) + m_{33}(Z_p - Z_{o_1})} \right]$$

$$y_p - y_o = -f \left[\frac{m_{21}(X_p - X_{o_1}) + m_{22}(Y_p - Y_{o_1}) + m_{23}(Z_p - Z_{o_1})}{m_{31}(X_p - X_{o_1}) + m_{32}(Y_p - Y_{o_1}) + m_{33}(Z_p - Z_{o_1})} \right]$$

One set of equations can be formulated for each ground point appearing on an image. The collinearity condition is commonly used to define the relationship between the camera/sensor, the image, and the ground.

Photogrammetric Solutions

As stated previously, digital photogrammetry is used for many applications, ranging from orthorectification, automated elevation extraction, stereopair creation, feature collection, highly accurate point determination, and control point extension.

For any of the aforementioned tasks to be undertaken, a relationship between the camera/sensor, the image(s) in a project, and the ground must be defined. The following variables are used to define the relationship:

- Exterior orientation parameters for each image
- Interior orientation parameters for each image
- Accurate representation of the ground

Well-known obstacles in photogrammetry include defining the interior and exterior orientation parameters for each image in a project using a minimum number of GCPs. Due to the costs and labor intensive procedures associated with collecting ground control, most photogrammetric applications do not have an abundant number of GCPs. Additionally, the exterior orientation parameters associated with an image are normally unknown.

Depending on the input data provided, photogrammetric techniques such as space resection, space forward intersection, and bundle block adjustment are used to define the variables required to perform orthorectification, automated DEM extraction, stereopair creation, highly accurate point determination, and control point extension.

Space Resection

Space resection is a technique that is commonly used to determine the exterior orientation parameters associated with one image or many images based on known GCPs. Space resection is based on the collinearity condition. Space resection using the collinearity condition specifies that, for any image, the exposure station, the ground point, and its corresponding image point must lie along a straight line.

If a minimum number of three GCPs is known in the X, Y, and Z direction, space resection techniques can be used to determine the six exterior orientation parameters associated with an image. Space resection assumes that camera information is available.

Space resection is commonly used to perform single frame orthorectification, where one image is processed at a time. If multiple images are being used, space resection techniques require that a minimum of three GCPs be located on each image being processed.

Using the collinearity condition, the positions of the exterior orientation parameters are computed. Light rays originating from at least three GCPs intersect through the image plane through the image positions of the GCPs and resect at the perspective center of the camera or sensor. Using least squares adjustment techniques, the most probable positions of exterior orientation can be computed. Space resection techniques can be applied to one image or multiple images.

Photogrammetric Concepts

Space Forward Intersection

Space forward intersection is a technique that is commonly used to determine the ground coordinates X , Y , and Z of points that appear in the overlapping areas of two or more images based on known interior orientation and known exterior orientation parameters. The collinearity condition is enforced, stating that the corresponding light rays from the two exposure stations pass through the corresponding image points on the two images and intersect at the same ground point. Figure 7-10 illustrates the concept associated with space forward intersection.

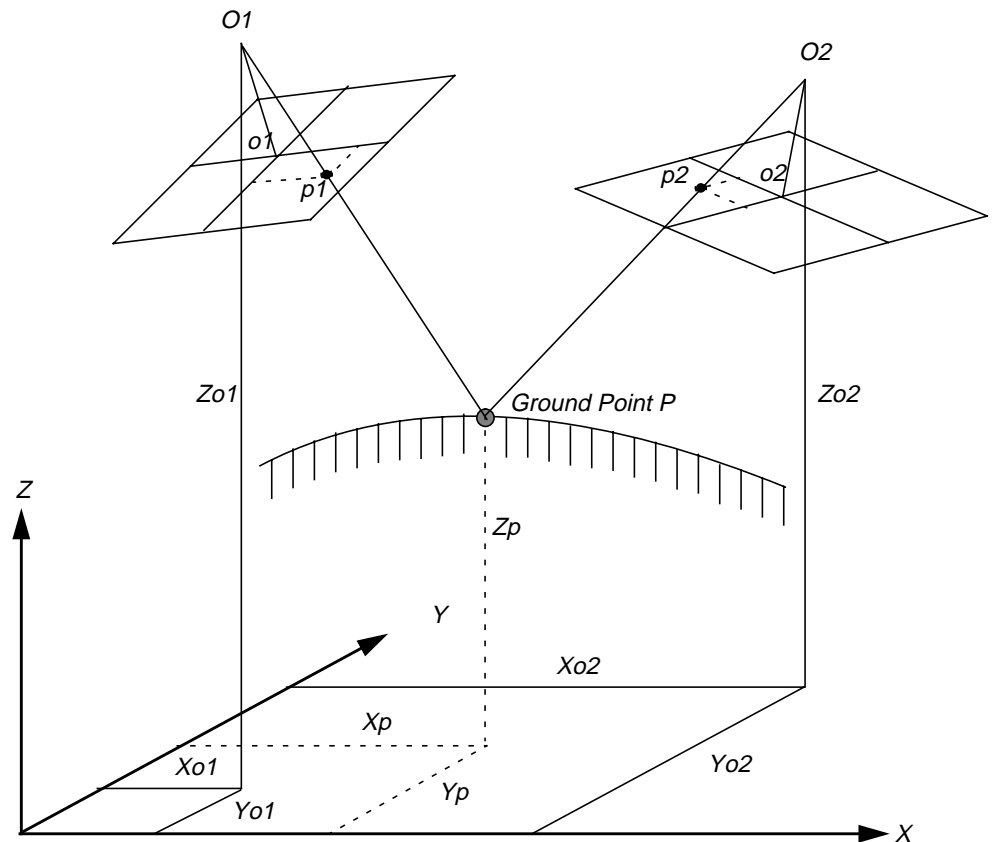


Figure 7-10 Space Forward Intersection

Space forward intersection techniques assume that the exterior orientation parameters associated with the images are known. Using the collinearity equations, the exterior orientation parameters along with the image coordinate measurements of point p on Image 1 and Image 2 are input to compute the X_p , Y_p , and Z_p coordinates of ground point p .

Space forward intersection techniques can be used for applications associated with collecting GCPs, cadastral mapping using airborne surveying techniques, and highly accurate point determination.

Bundle Block Adjustment

For mapping projects having more than two images, the use of space intersection and space resection techniques is limited. This can be attributed to the lack of information required to perform these tasks. For example, it is fairly uncommon for the exterior orientation parameters to be highly accurate for each photograph or image in a project, since these values are generated photogrammetrically. Airborne GPS and INS techniques normally provide initial approximations to exterior orientation, but the final values for these parameters must be adjusted to attain higher accuracies.

Similarly, rarely are there enough accurate GCPs for a project of 30 or more images to perform space resection (i.e., a minimum of 90 is required). In the case that there are enough GCPs, the time required to identify and measure all of the points would be costly.

The costs associated with block triangulation and orthorectification are largely dependent on the number of GCPs used. To minimize the costs of a mapping project, fewer GCPs are collected and used. To ensure that high accuracies are attained, an approach known as bundle block adjustment is used.

A bundle block adjustment is best defined by examining the individual words in the term. A bundled solution is computed including the exterior orientation parameters of each image in a block and the X, Y, and Z coordinates of tie points and adjusted GCPs. A block of images contained in a project is simultaneously processed in one solution. A statistical technique known as least squares adjustment is used to estimate the bundled solution for the entire block while also minimizing and distributing error.

Block triangulation is the process of defining the mathematical relationship between the images contained within a block, the camera or sensor model, and the ground. Once the relationship has been defined, accurate imagery and information concerning the Earth's surface can be created.

When processing frame camera, digital camera, videography, and nonmetric camera imagery, block triangulation is commonly referred to as aerial triangulation (AT). When processing imagery collected with a pushbroom sensor, block triangulation is commonly referred to as triangulation.

There are several models for block triangulation. The common models used in photogrammetry are block triangulation with the strip method, the independent model method, and the bundle method. Among them, the bundle block adjustment is the most rigorous of the above methods, considering the minimization and distribution of errors. Bundle block adjustment uses the collinearity condition as the basis for formulating the relationship between image space and ground space. IMAGINE OrthoBASE uses bundle block adjustment techniques.

In order to understand the concepts associated with bundle block adjustment, an example comprising two images with three GCPs with X, Y, and Z coordinates that are known is used. Additionally, six tie points are available. Figure 7-11 illustrates the photogrammetric configuration.

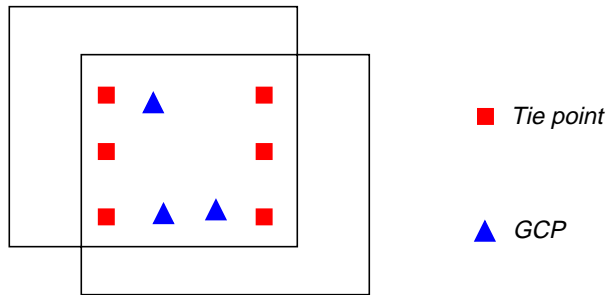


Figure 7-11 Photogrammetric Configuration

Forming the Collinearity Equations

For each measured GCP, there are two corresponding image coordinates (x and y). Thus, two collinearity equations can be formulated to represent the relationship between the ground point and the corresponding image measurements. In the context of bundle block adjustment, these equations are known as observation equations.

If a GCP has been measured on the overlapping areas of two images, four equations can be written: two for image measurements on the left image comprising the pair and two for the image measurements made on the right image comprising the pair. Thus, GCP A measured on the overlap areas of image left and image right has four collinearity formulas.

$$x_{a_1} - x_o = -f \left[\frac{m_{11}(X_A - X_{o_1}) + m_{12}(Y_A - Y_{o_1}) + m_{13}(Z_A - Z_{o_1})}{m_{31}(X_A - X_{o_1}) + m_{32}(Y_A - Y_{o_1}) + m_{33}(Z_A - Z_{o_1})} \right]$$

$$y_{a_1} - y_o = -f \left[\frac{m_{21}(X_A - X_{o_1}) + m_{22}(Y_A - Y_{o_1}) + m_{23}(Z_A - Z_{o_1})}{m_{31}(X_A - X_{o_1}) + m_{32}(Y_A - Y_{o_1}) + m_{33}(Z_A - Z_{o_1})} \right]$$

$$x_{a_2} - x_o = -f \left[\frac{m'_{11}(X_A - X_{o_2}) + m'_{12}(Y_A - Y_{o_2}) + m'_{13}(Z_A - Z_{o_2})}{m'_{31}(X_A - X_{o_2}) + m'_{32}(Y_A - Y_{o_2}) + m'_{33}(Z_A - Z_{o_2})} \right]$$

$$y_{a_2} - y_o = -f \left[\frac{m'_{21}(X_A - X_{o_2}) + m'_{22}(Y_A - Y_{o_2}) + m'_{23}(Z_A - Z_{o_2})}{m'_{31}(X_A - X_{o_2}) + m'_{32}(Y_A - Y_{o_2}) + m'_{33}(Z_A - Z_{o_2})} \right]$$

One image measurement of GCP A on Image 1:

$$x_{a_1}, y_{a_1}$$

One image measurement of GCP A on Image 2:

$$x_{a_2}, y_{a_2}$$

Positional elements of exterior orientation on Image 1:

$$X_{o_1}, Y_{o_1}, Z_{o_1}$$

Positional elements of exterior orientation on Image 2:

$$X_{o_2}, Y_{o_2}, Z_{o_2}$$

If three GCPs have been measured on the overlap areas of two images, twelve equations can be formulated, which includes four equations for each GCP (refer to Figure 7-11).

Additionally, if six tie points have been measured on the overlap areas of the two images, twenty-four equations can be formulated, which includes four for each tie point. This is a total of 36 observation equations (refer to Figure 7-11).

Photogrammetric Concepts

The previous example has the following unknowns:

- Six exterior orientation parameters for the left image (i.e., X , Y , Z , ω , ϕ , κ)
- Six exterior orientation parameters for the right image (i.e., X , Y , Z , ω , ϕ and κ), and
- X , Y , and Z coordinates of the tie points. Thus, for six tie points, this includes eighteen unknowns (six tie points times three X , Y , Z coordinates).

The total number of unknowns is 30. The overall quality of a bundle block adjustment is largely a function of the quality and redundancy in the input data. In this scenario, the redundancy in the project can be computed by subtracting the number of unknowns, 30, by the number of knowns, 36. The resulting redundancy is six. This term is commonly referred to as the degrees of freedom in a solution.

Once each observation equation is formulated, the collinearity condition can be solved using an approach referred to as least squares adjustment.

Least Squares Adjustment

Least squares adjustment is a statistical technique that is used to estimate the unknown parameters associated with a solution while also minimizing error within the solution. With respect to block triangulation, least squares adjustment techniques are used to:

- Estimate or adjust the values associated with exterior orientation
- Estimate the X , Y , and Z coordinates associated with tie points
- Estimate or adjust the values associated with interior orientation
- Minimize and distribute data error through the network of observations

Data error is attributed to the inaccuracy associated with the input GCP coordinates, measured tie point and GCP image positions, camera information, and systematic errors.

The least squares approach requires iterative processing until a solution is attained. A solution is obtained when the residuals associated with the input data are minimized.

The least squares approach involves determining the corrections to the unknown parameters based on the criteria of minimizing input measurement residuals. The residuals are derived from the difference between the measured (i.e., user input) and computed value for any particular measurement in a project. In the block triangulation process, a functional model can be formed based upon the collinearity equations.

The functional model refers to the specification of an equation that can be used to relate measurements to parameters. In the context of photogrammetry, measurements include the image locations of GCPs and GCP coordinates, while the exterior orientations of all the images are important parameters estimated by the block triangulation process.

The residuals, which are minimized, include the image coordinates of the GCPs and tie points along with the known ground coordinates of the GCPs. A simplified version of the least squares condition can be broken down into a formulation as follows:

$$V = AX - L, \text{ including a weight matrix } P$$

Where:

- V = the matrix containing the image coordinate residuals
- A = the matrix containing the partial derivatives with respect to the unknown parameters, including exterior orientation, interior orientation, XYZ tie point, and GCP coordinates
- X = the matrix containing the corrections to the unknown parameters
- L = the matrix containing the input observations (i.e., image coordinates and GCP coordinates)

The components of the least squares condition are directly related to the functional model based on collinearity equations. The A matrix is formed by differentiating the functional model, which is based on collinearity equations, with respect to the unknown parameters such as exterior orientation, etc. The L matrix is formed by subtracting the initial results obtained from the functional model with newly estimated results determined from a new iteration of processing. The X matrix contains the corrections to the unknown exterior orientation parameters. The X matrix is calculated in the following manner:

$$X = (A^t P A)^{-1} A^t P L$$

Where:

- X = the matrix containing the corrections to the unknown parameters
- A = the matrix containing the partial derivatives with respect to the unknown parameters
- t = the matrix transposed
- P = the matrix containing the weights of the observations
- L = the matrix containing the observations

Once a least squares iteration of processing is completed, the corrections to the unknown parameters are added to the initial estimates. For example, if initial approximations to exterior orientation are provided from Airborne GPS and INS information, the estimated corrections computed from the least squares adjustment are added to the initial value to compute the updated exterior orientation values. This iterative process of least squares adjustment continues until the corrections to the unknown parameters are less than a user-specified threshold (commonly referred to as a convergence value).

The V residual matrix is computed at the end of each iteration of processing. Once an iteration is completed, the new estimates for the unknown parameters are used to recompute the input observations such as the image coordinate values. The difference between the initial measurements and the new estimates is obtained to provide the residuals. Residuals provide preliminary indications of the accuracy of a solution. The residual values indicate the degree to which a particular observation (input) fits with the functional model. For example, the image residuals have the capability of reflecting GCP collection in the field. After each successive iteration of processing, the residuals become smaller until they are satisfactorily minimized.

Once the least squares adjustment is completed, the block triangulation results include:

- Final exterior orientation parameters of each image in a block and their accuracy
- Final interior orientation parameters of each image in a block and their accuracy
- X , Y , and Z tie point coordinates and their accuracy
- Adjusted GCP coordinates and their residuals
- Image coordinate residuals

The results from the block triangulation are then used as the primary input for the following tasks:

- Stereo pair creation
- Feature collection
- Highly accurate point determination
- DEM extraction
- Orthorectification

Self-calibrating Bundle Adjustment

Normally, there are more or less systematic errors related to the imaging and processing system, such as lens distortion, film distortion, atmosphere refraction, scanner errors, etc. These errors reduce the accuracy of triangulation results, especially in dealing with large-scale imagery and high accuracy triangulation. There are several ways to reduce the influences of the systematic errors, like *a posteriori*-compensation, test-field calibration, and the most common approach: self-calibration (Konecny 1984; Wang 1990).

The self-calibrating methods use additional parameters in the triangulation process to eliminate the systematic errors. How well it works depends on many factors such as the strength of the block (overlap amount, crossing flight lines), the GCP and tie point distribution and amount, the size of systematic errors versus random errors, the significance of the additional parameters, the correlation between additional parameters, and other unknowns.

There was intensive research and development for additional parameter models in photogrammetry in the 70s and the 80s, and many research results are available (e.g., Bauer 1972; Brown 1975; Ebner 1976; Grün 1978; Jacobsen 1980 and 1982; Li 1985; Wang 1988, Stojić, *et al.* 1998). Based on these scientific reports, IMAGINE OrthoBASE provides four groups of additional parameters for you to choose for different triangulation circumstances. In addition, IMAGINE OrthoBASE allows the interior orientation parameters to be analytically calibrated within its self-calibrating bundle block adjustment capability.

Automatic Gross Error Detection

Normal random errors are subject to statistical normal distribution. In contrast, gross errors refer to errors that are large and are not subject to normal distribution. The gross errors among the input data for triangulation can lead to unreliable results. Research during the 80s in the photogrammetric community resulted in significant achievements in automatic gross error detection in the triangulation process (e.g., Kubik 1982; Li 1983 and 1985; Jacobsen 1984; El-Hakin 1984; Wang 1988).

Methods for gross error detection began with residual checking using data-snooping and were later extended to robust estimation (Wang 1990). The most common robust estimation method is the iteration with selective weight functions. Based on the scientific research results from the photogrammetric community, IMAGINE OrthoBASE offers two robust error detection methods within the triangulation process.

It is worth mentioning that the effect of the automatic error detection depends not only on the mathematical model, but also depends on the redundancy in the block. Therefore, more tie points in more overlap areas contribute better gross error detection. In addition, inaccurate GCPs can distribute their errors to correct tie points, therefore the ground and image coordinates of GCPs should have better accuracy than tie points when comparing them within the same scale space.

GCPs

The instrumental component of establishing an accurate relationship between the images in a project, the camera/sensor, and the ground is GCPs. GCPs are identifiable features located on the Earth's surface that have known ground coordinates in X, Y, and Z. A full GCP has X, Y, and Z (elevation of the point) coordinates associated with it. Horizontal control only specifies the X, Y, while vertical control only specifies the Z. The following features on the Earth's surface are commonly used as GCPs:

- Intersection of roads
- Utility infrastructure (e.g., fire hydrants and manhole covers)
- Intersection of agricultural plots of land
- Survey benchmarks

Depending on the type of mapping project, GCPs can be collected from the following sources:

- Theodolite survey (millimeter to centimeter accuracy)
- Total station survey (millimeter to centimeter accuracy)
- Ground GPS (centimeter to meter accuracy)
- Planimetric and topographic maps (accuracy varies as a function of map scale, approximate accuracy between several meters to 40 meters or more)
- Digital orthorectified images (X and Y coordinates can be collected to an accuracy dependent on the resolution of the orthorectified image)
- DEMs (for the collection of vertical GCPs having Z coordinates associated with them, where accuracy is dependent on the resolution of the DEM and the accuracy of the input DEM)

Photogrammetric Concepts

When imagery or photography is exposed, GCPs are recorded and subsequently displayed on the photography or imagery. During GCP measurement in IMAGINE OrthoBASE, the image positions of GCPs appearing on an image or on the overlap areas of the images are collected.

It is highly recommended that a greater number of GCPs be available than are actually used in the block triangulation. Additional GCPs can be used as check points to independently verify the overall quality and accuracy of the block triangulation solution. A check point analysis compares the photogrammetrically computed ground coordinates of the check points to the original values. The result of the analysis is an RMSE that defines the degree of correspondence between the computed values and the original values. Lower RMSE values indicate better results.

GCP Requirements

The minimum GCP requirements for an accurate mapping project vary with respect to the size of the project. With respect to establishing a relationship between image space and ground space, the theoretical minimum number of GCPs is two GCPs having X, Y, and Z coordinates and one GCP having a Z coordinate associated with it. This is a total of seven observations.

In establishing the mathematical relationship between image space and object space, seven parameters defining the relationship must be determined. The seven parameters include a scale factor (describing the scale difference between image space and ground space); X, Y, Z (defining the positional differences between image space and object space); and three rotation angles (omega, phi, and kappa) that define the rotational relationship between image space and ground space.

In order to compute a unique solution, at least seven known parameters must be available. In using the two X, Y, Z GCPs and one vertical (Z) GCP, the relationship can be defined. However, to increase the accuracy of a mapping project, using more GCPs is highly recommended.

The following descriptions are provided for various projects:

Processing One Image

If processing one image for the purpose of orthorectification (i.e., a single frame orthorectification), the minimum number of GCPs required is three. Each GCP must have an X, Y, and Z coordinate associated with it. The GCPs should be evenly distributed to ensure that the camera/sensor is accurately modeled.

Processing a Strip of Images

If processing a strip of adjacent images, two GCPs for every third image is recommended. To increase the quality of orthorectification, measuring three GCPs at the corner edges of a strip is advantageous. Thus, during block triangulation a stronger geometry can be enforced in areas where there is less redundancy such as the corner edges of a strip or a block.

Figure 7-12 illustrates the GCP configuration for a strip of images having 60% overlap. The triangles represent the GCPs. Thus, the image positions of the GCPs are measured on the overlap areas of the imagery.

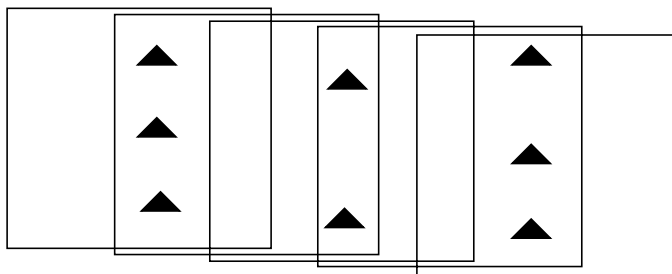


Figure 7-12 GCP Configuration

Processing Multiple Strips of Imagery

Figure 7-13 depicts the standard GCP configuration for a block of images, comprising four strips of images, each containing eight overlapping images.

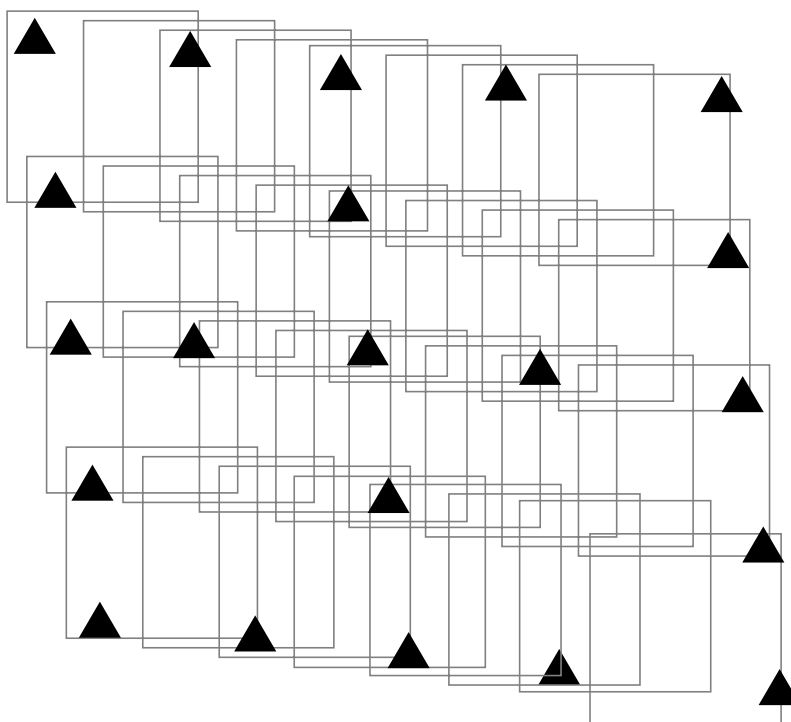


Figure 7-13 GCPs in a Block of Images

In this case, the GCPs form a strong geometric network of observations. As a general rule, it is advantageous to have at least one GCP on every third image of a block. Additionally, whenever possible, locate GCPs that lie on multiple images, around the outside edges of a block, and at certain distances from one another within the block.

Tie Points

A tie point is a point that has ground coordinates that are not known, but is visually recognizable in the overlap area between two or more images. The corresponding image positions of tie points appearing on the overlap areas of multiple images is identified and measured. Ground coordinates for tie points are computed during block triangulation. Tie points can be measured both manually and automatically.

Tie points should be visually well-defined in all images. Ideally, they should show good contrast in two directions, like the corner of a building or a road intersection. Tie points should also be well distributed over the area of the block. Typically, nine tie points in each image are adequate for block triangulation. Figure 7-14 depicts the placement of tie points.

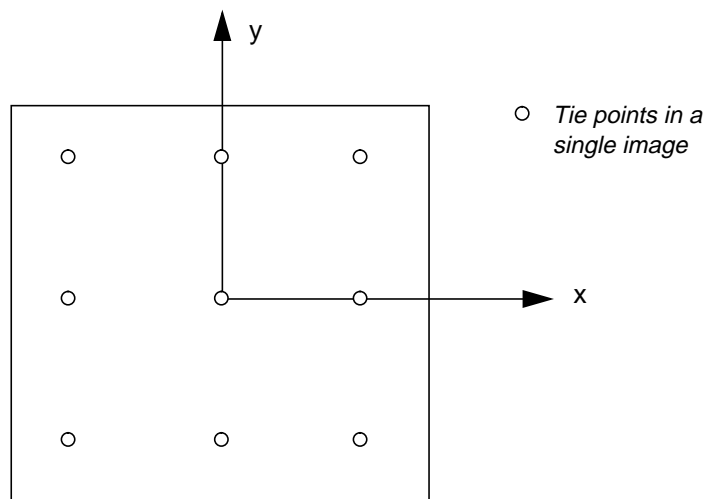


Figure 7-14 Point Distribution for Triangulation

In a block of images with 60% overlap and 25-30% sidelap, nine points are sufficient to tie together the block as well as individual strips (see Figure 7-15).

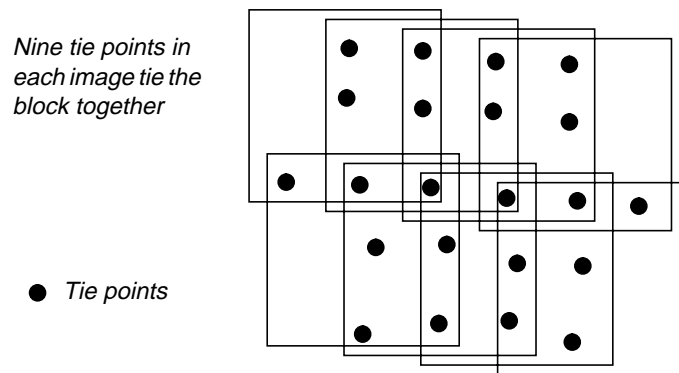


Figure 7-15 Tie Points in a Block

Automatic Tie Point Collection

Selecting and measuring tie points is very time-consuming and costly. Therefore, in recent years, one of the major focal points of research and development in photogrammetry has concentrated on the automated triangulation where the automatic tie point collection is the main issue.

The other part of the automated triangulation is the automatic control point identification, which is still unsolved due to the complication of the scenario. There are several valuable research results available for automated triangulation (e.g., Agouris and Schenk 1996; Heipke 1996; Krzystek 1998; Mayr 1995; Schenk 1997; Tang 1997; Tsingas 1995; Wang 1998).

After investigating the advantages and the weaknesses of the existing methods, IMAGINE OrthoBASE was designed to incorporate an advanced method for automatic tie point collection. It is designed to work with a variety of digital images such as aerial images, satellite images, digital camera images, and close range images. It also supports the processing of multiple strips including adjacent, diagonal, and cross-strips.

Automatic tie point collection within IMAGINE OrthoBASE successfully performs the following tasks:

- Automatic block configuration. Based on the initial input requirements, IMAGINE OrthoBASE automatically detects the relationship of the block with respect to image adjacency.
- Automatic tie point extraction. The feature point extraction algorithms are used here to extract the candidates of tie points.
- Point transfer. Feature points appearing on multiple images are automatically matched and identified.
- Gross error detection. Erroneous points are automatically identified and removed from the solution.
- Tie point selection. The intended number of tie points defined by you is automatically selected as the final number of tie points.

The image matching strategies incorporated in IMAGINE OrthoBASE for automatic tie point collection include the coarse-to-fine matching; feature-based matching with geometrical and topological constraints, which is simplified from the structural matching algorithm (Wang 1998); and the least square matching for the high accuracy of tie points.

Image Matching Techniques

Image matching refers to the automatic identification and measurement of corresponding image points that are located on the overlapping area of multiple images. The various image matching methods can be divided into three categories including:

- Area based matching
- Feature based matching
- Relation based matching

Area Based Matching

Area based matching is also called signal based matching. This method determines the correspondence between two image areas according to the similarity of their gray level values. The cross correlation and least squares correlation techniques are well-known methods for area based matching.

Correlation Windows

Area based matching uses correlation windows. These windows consist of a local neighborhood of pixels. One example of correlation windows is square neighborhoods (e.g., 3×3 , 5×5 , 7×7 pixels). In practice, the windows vary in shape and dimension based on the matching technique. Area correlation uses the characteristics of these windows to match ground feature locations in one image to ground features on the other.

A reference window is the source window on the first image, which remains at a constant location. Its dimensions are usually square in size (e.g., 3×3 , 5×5 , etc.). Search windows are candidate windows on the second image that are evaluated relative to the reference window. During correlation, many different search windows are examined until a location is found that best matches the reference window.

Correlation Calculations

Two correlation calculations are described below: cross correlation and least squares correlation. Most area based matching calculations, including these methods, normalize the correlation windows. Therefore, it is not necessary to balance the contrast or brightness prior to running correlation. Cross correlation is more robust in that it requires a less accurate *a priori* position than least squares. However, its precision is limited to one pixel. Least squares correlation can achieve precision levels of one-tenth of a pixel, but requires an *a priori* position that is accurate to about two pixels. In practice, cross correlation is often followed by least squares for high accuracy.

Cross Correlation

Cross correlation computes the correlation coefficient of the gray values between the template window and the search window according to the following equation:

$$\rho = \frac{\sum_{i,j} [g_1(c_1, r_1) - \bar{g}_1][g_2(c_2, r_2) - \bar{g}_2]}{\sqrt{\sum_{i,j} [g_1(c_1, r_1) - \bar{g}_1]^2 \sum_{i,j} [g_2(c_2, r_2) - \bar{g}_2]^2}}$$

with

$$\bar{g}_1 = \frac{1}{n} \sum_{i,j} g_1(c_1, r_1) \quad \bar{g}_2 = \frac{1}{n} \sum_{i,j} g_2(c_2, r_2)$$

Where:

- ρ = the correlation coefficient
- $g(c,r)$ = the gray value of the pixel (c,r)
- c_1, r_1 = the pixel coordinates on the left image
- c_2, r_2 = the pixel coordinates on the right image
- n = the total number of pixels in the window
- i, j = pixel index into the correlation window

When using the area based cross correlation, it is necessary to have a good initial position for the two correlation windows. If the exterior orientation parameters of the images being matched are known, a good initial position can be determined. Also, if the contrast in the windows is very poor, the correlation can fail.

Least Squares Correlation

Least squares correlation uses the least squares estimation to derive parameters that best fit a search window to a reference window. This technique has been investigated thoroughly in photogrammetry (Ackermann 1983; Grün and Baltsavias 1988; Helava 1988). It accounts for both gray scale and geometric differences, making it especially useful when ground features on one image look somewhat different on the other image (differences which occur when the surface terrain is quite steep or when the viewing angles are quite different).

Least squares correlation is iterative. The parameters calculated during the initial pass are used in the calculation of the second pass and so on, until an optimum solution is determined. Least squares matching can result in high positional accuracy (about 0.1 pixels). However, it is sensitive to initial approximations. The initial coordinates for the search window prior to correlation must be accurate to about two pixels or better.

When least squares correlation fits a search window to the reference window, both radiometric (pixel gray values) and geometric (location, size, and shape of the search window) transformations are calculated.

Photogrammetric Concepts

For example, suppose the change in gray values between two correlation windows is represented as a linear relationship. Also assume that the change in the window's geometry is represented by an affine transformation.

$$g_2(c_2, r_2) = h_0 + h_1 g_1(c_1, r_1)$$

$$c_2 = a_0 + a_1 c_1 + a_2 r_1$$

$$r_2 = b_0 + b_1 c_1 + b_2 r_1$$

Where:

c_1, r_1 = the pixel coordinate in the reference window

c_2, r_2 = the pixel coordinate in the search window

$g_1(c_1, r_1)$ = the gray value of pixel (c_1, r_1)

$g_2(c_2, r_2)$ = the gray value of pixel (c_2, r_2)

h_0, h_1 = linear gray value transformation parameters

a_0, a_1, a_2 = affine geometric transformation parameters

b_0, b_1, b_2 = affine geometric transformation parameters

Based on this assumption, the error equation for each pixel is derived, as shown in the following equation:

$$v = (a_1 + a_2 c_1 + a_3 r_1) g_c + (b_1 + b_2 c_1 + b_3 r_1) g_r - h_1 - h_2 g_1(c_1, r_1) + \Delta g$$

with $\Delta g = g_2(c_2, r_2) - g_1(c_1, r_1)$

Where:

g_c and g_r are the gradients of $g_2(c_2, r_2)$.

Feature Based Matching

Feature based matching determines the correspondence between two image features. Most feature based techniques match extracted point features (this is called feature point matching), as opposed to other features, such as lines or complex objects. The feature points are also commonly referred to as interest points. Poor contrast areas can be avoided with feature based matching.

In order to implement feature based matching, the image features must initially be extracted. There are several well-known operators for feature point extraction. Examples include the Moravec Operator, the Dreschler Operator, and the Förstner Operator (Förstner and Gülch 1987; Lü 1988).

After the features are extracted, the attributes of the features are compared between two images. The feature pair having the attributes with the best fit is recognized as a match. IMAGINE OrthoBASE utilizes the Förstner interest operator to extract feature points.

Relation Based Matching Relation based matching is also called structural matching (Vosselman and Haala 1992; Wang 1994 and 1995). This kind of matching technique uses the image features and the relationship between the features. With relation based matching, the corresponding image structures can be recognized automatically, without any *a priori* information. However, the process is time-consuming since it deals with varying types of information. Relation based matching can also be applied for the automatic recognition of control points.

Image Pyramid Because of the large amount of image data, the image pyramid is usually adopted during the image matching techniques to reduce the computation time and to increase the matching reliability. The pyramid is a data structure consisting of the same image represented several times, at a decreasing spatial resolution each time. Each level of the pyramid contains the image at a particular resolution.

The matching process is performed at each level of resolution. The search is first performed at the lowest resolution level and subsequently at each higher level of resolution. Figure 7-16 shows a four-level image pyramid.

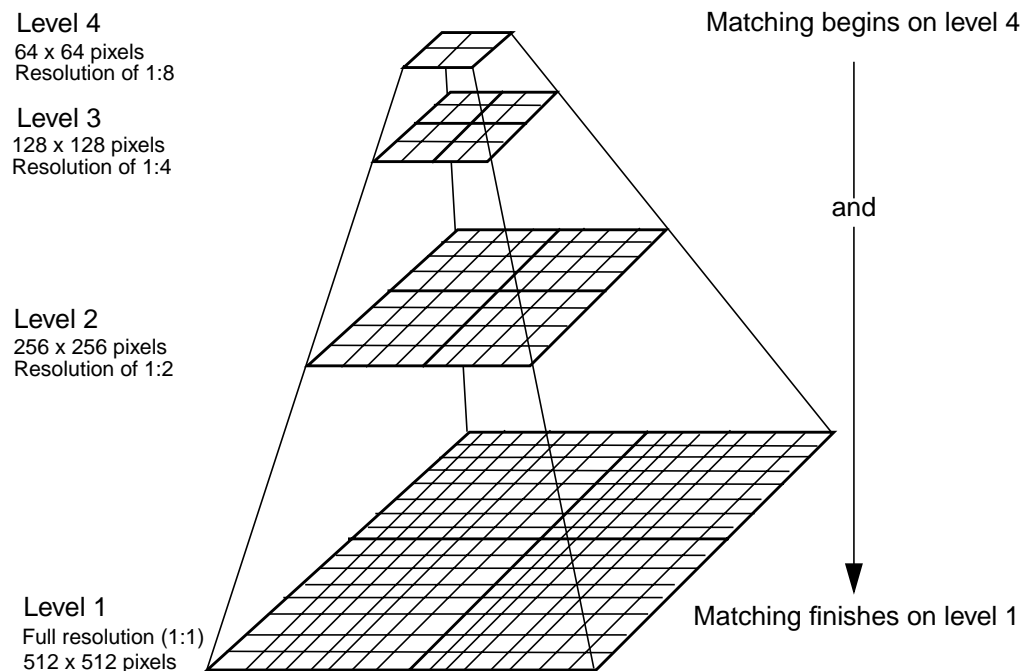


Figure 7-16 Image Pyramid for Matching at Coarse to Full Resolution

There are different resampling methods available for generating an image pyramid. Theoretical and practical investigations show that the resampling methods based on the Gaussian filter, which are approximated by a binomial filter, have the superior properties concerning preserving the image contents and reducing the computation time (Wang 1994). Therefore, IMAGINE OrthoBASE uses this kind of pyramid layer instead of those currently available under ERDAS IMAGINE, which are overwritten automatically by IMAGINE OrthoBASE.

Satellite Photogrammetry

Satellite photogrammetry has slight variations compared to photogrammetric applications associated with aerial frame cameras. This document makes reference to the SPOT and IRS-1C satellites. The SPOT satellite provides 10-meter panchromatic imagery and 20-meter multispectral imagery (four multispectral bands of information).

The SPOT satellite carries two high resolution visible (HRV) sensors, each of which is a pushbroom scanner that takes a sequence of line images while the satellite circles the Earth. The focal length of the camera optic is 1084 mm, which is very large relative to the length of the camera (78 mm). The field of view is 4.1 degrees. The satellite orbit is circular, North-South and South-North, about 830 km above the Earth, and sun-synchronous. A sun-synchronous orbit is one in which the orbital rotation is the same rate as the Earth's rotation.

The Indian Remote Sensing (IRS-1C) satellite utilizes a pushbroom sensor consisting of three individual CCDs. The ground resolution of the imagery ranges between 5 to 6 meters. The focal length of the optic is approximately 982 mm. The pixel size of the CCD is 7 microns. The images captured from the three CCDs are processed independently or merged into one image and system corrected to account for the systematic error associated with the sensor.

Both the SPOT and IRS-1C satellites collect imagery by scanning along a line. This line is referred to as the scan line. For each line scanned within the SPOT and IRS-1C sensors, there is a unique perspective center and a unique set of rotation angles. The location of the perspective center relative to the line scanner is constant for each line (interior orientation and focal length). Since the motion of the satellite is smooth and practically linear over the length of a scene, the perspective centers of all scan lines of a scene are assumed to lie along a smooth line. Figure 7-17 illustrates the scanning technique.

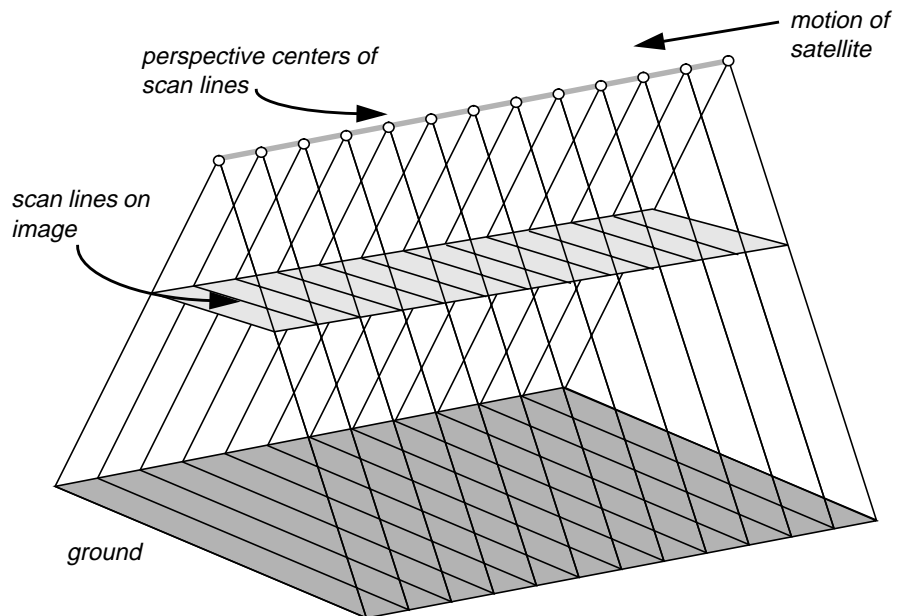


Figure 7-17 Perspective Centers of SPOT Scan Lines

The satellite exposure station is defined as the perspective center in ground coordinates for the center scan line. The image captured by the satellite is called a scene. For example, a SPOT Pan 1A scene is composed of 6000 lines. For SPOT Pan 1A imagery, each of these lines consists of 6000 pixels. Each line is exposed for 1.5 milliseconds, so it takes 9 seconds to scan the entire scene. (A scene from SPOT XS 1A is composed of only 3000 lines and 3000 columns and has 20-meter pixels, while Pan has 10-meter pixels.)

NOTE: The following section addresses only the 10 meter SPOT Pan scenario.

A pixel in the SPOT image records the light detected by one of the 6000 light sensitive elements in the camera. Each pixel is defined by file coordinates (column and row numbers). The physical dimension of a single, light-sensitive element is 13×13 microns. This is the pixel size in image coordinates. The center of the scene is the center pixel of the center scan line. It is the origin of the image coordinate system. Figure 7-18 depicts image coordinates in a satellite scene:

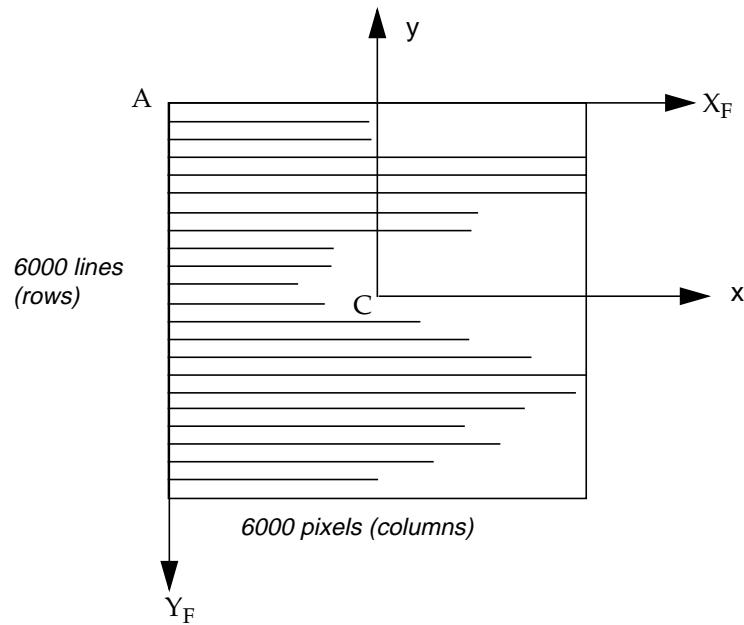


Figure 7-18 Image Coordinates in a Satellite Scene

Where:

- A = origin of file coordinates
- A- X_F , A- Y_F = file coordinate axes
- C = origin of image coordinates (center of scene)
- C-x, C-y = image coordinate axes

Photogrammetric Concepts

SPOT Interior Orientation

Figure 7-19 shows the interior orientation of a satellite scene. The transformation between file coordinates and image coordinates is constant.

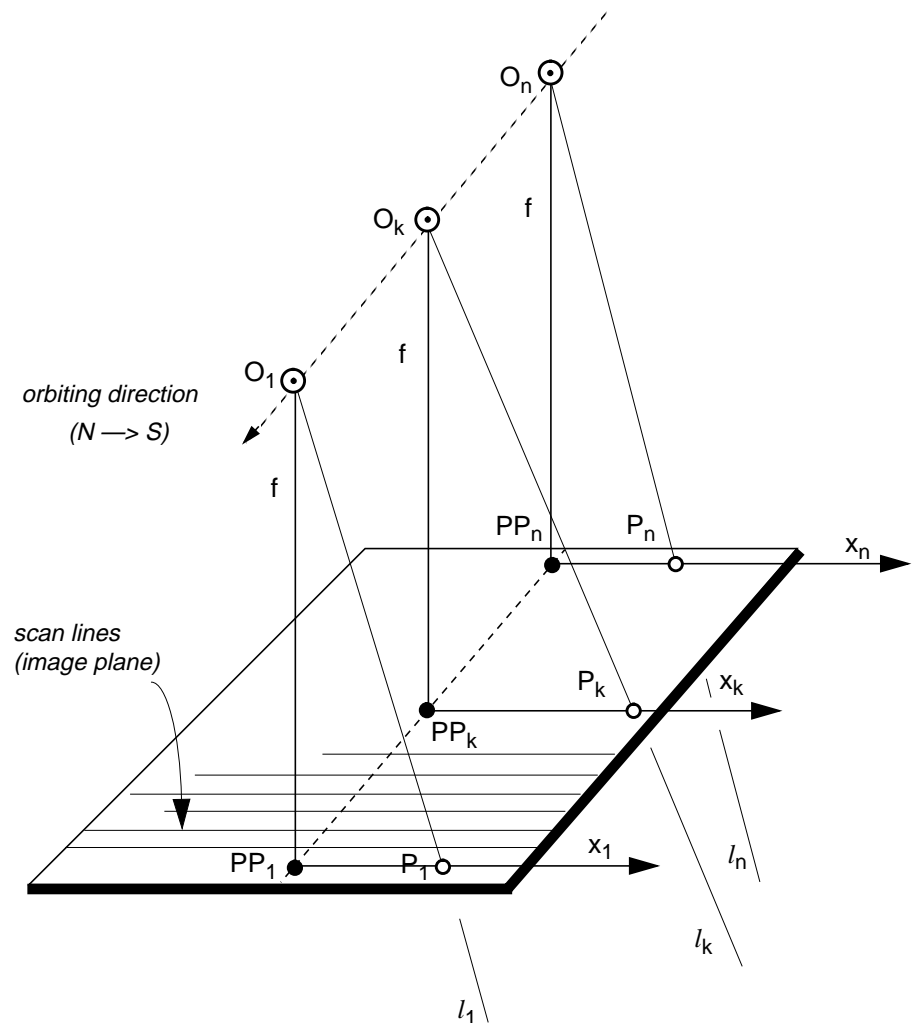


Figure 7-19 Interior Orientation of a SPOT Scene

For each scan line, a separate bundle of light rays is defined, where:

- P_k = image point
- x_k = x value of image coordinates for scan line k
- f = focal length of the camera
- O_k = perspective center for scan line k, aligned along the orbit
- PP_k = principal point for scan line k
- l_k = light rays for scan line, bundled at perspective center O_k

SPOT Exterior Orientation

SPOT satellite geometry is stable and the sensor parameters, such as focal length, are well-known. However, the triangulation of SPOT scenes is somewhat unstable because of the narrow, almost parallel bundles of light rays.

Ephemeris data for the orbit are available in the header file of SPOT scenes. They give the satellite's position in three-dimensional, geocentric coordinates at 60-second increments. The velocity vector and some rotational velocities relating to the attitude of the camera are given, as well as the exact time of the center scan line of the scene. The header of the data file of a SPOT scene contains ephemeris data, which provides information about the recording of the data and the satellite orbit.

Ephemeris data that can be used in satellite triangulation include:

- Position of the satellite in geocentric coordinates (with the origin at the center of the Earth) to the nearest second
- Velocity vector, which is the direction of the satellite's travel
- Attitude changes of the camera
- Time of exposure (exact) of the center scan line of the scene

The geocentric coordinates included with the ephemeris data are converted to a local ground system for use in triangulation. The center of a satellite scene is interpolated from the header data.

Light rays in a bundle defined by the SPOT sensor are almost parallel, lessening the importance of the satellite's position. Instead, the inclination angles (incidence angles) of the cameras on board the satellite become the critical data.

The scanner can produce a nadir view. Nadir is the point directly below the camera. SPOT has off-nadir viewing capability. Off-nadir refers to any point that is not directly beneath the satellite, but is off to an angle (i.e., East or West of the nadir).

A stereo scene is achieved when two images of the same area are acquired on different days from different orbits, one taken East of the other. For this to occur, there must be significant differences in the inclination angles.

Inclination is the angle between a vertical on the ground at the center of the scene and a light ray from the exposure station. This angle defines the degree of off-nadir viewing when the scene was recorded. The cameras can be tilted in increments of a minimum of 0.6 to a maximum of 27 degrees to the East (negative inclination) or West (positive inclination). Figure 7-20 illustrates the inclination.

Photogrammetric Concepts

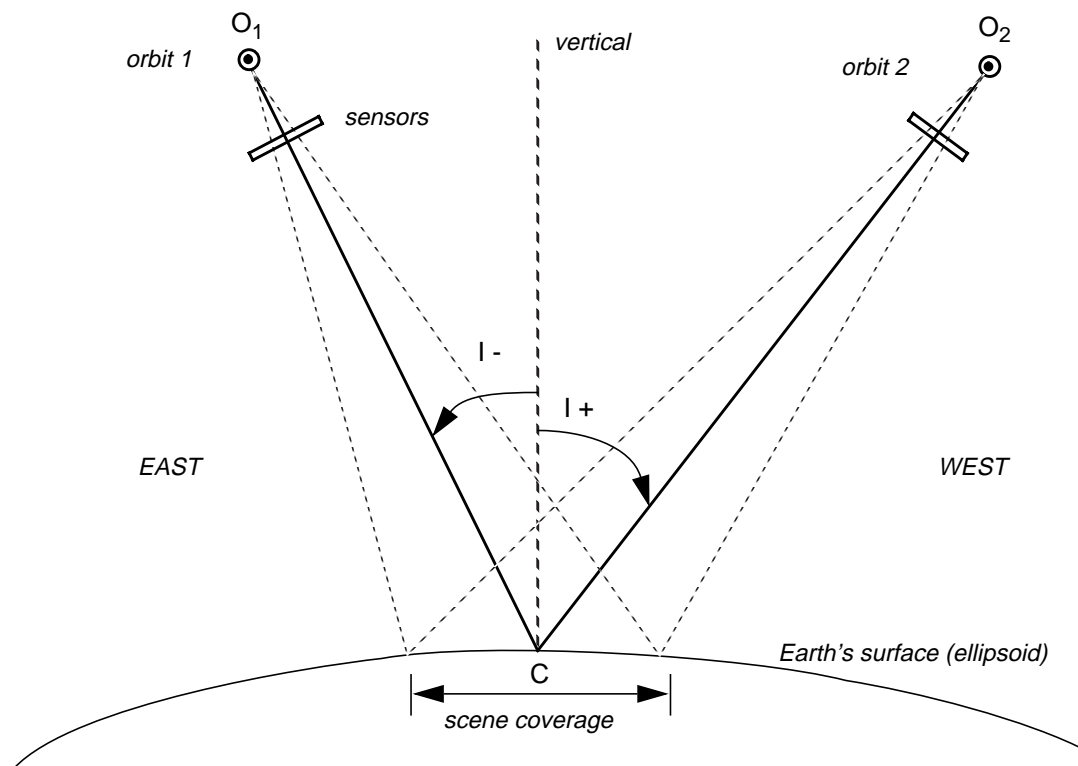


Figure 7-20 Inclination of a Satellite Stereo-Scene (View from North to South)

Where:

- C = center of the scene
- I- = eastward inclination
- I+ = westward inclination
- O₁, O₂ = exposure stations (perspective centers of imagery)

The orientation angle of a satellite scene is the angle between a perpendicular to the center scan line and the North direction. The spatial motion of the satellite is described by the velocity vector. The real motion of the satellite above the ground is further distorted by the Earth's rotation.

The velocity vector of a satellite is the satellite's velocity if measured as a vector through a point on the spheroid. It provides a technique to represent the satellite's speed as if the imaged area were flat instead of being a curved surface (see Figure 7-21).

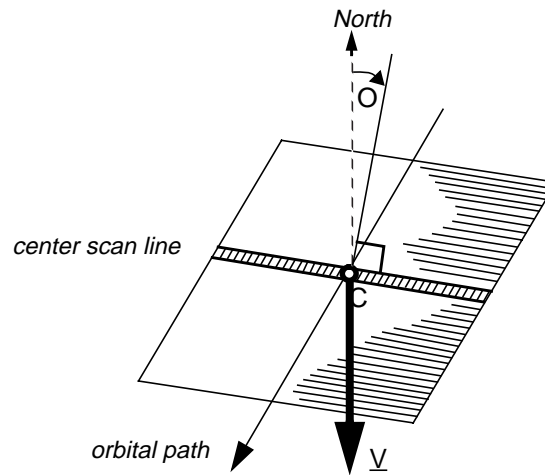


Figure 7-21 Velocity Vector and Orientation Angle of a Single Scene

Where:

- O = orientation angle
- C = center of the scene
- \underline{V} = velocity vector

Satellite block triangulation provides a model for calculating the spatial relationship between a satellite sensor and the ground coordinate system for each line of data. This relationship is expressed as the exterior orientation, which consists of

- the perspective center of the center scan line (i.e., X, Y, and Z),
- the change of perspective centers along the orbit,
- the three rotations of the center scan line (i.e., omega, phi, and kappa), and
- the changes of angles along the orbit.

In addition to fitting the bundle of light rays to the known points, satellite block triangulation also accounts for the motion of the satellite by determining the relationship of the perspective centers and rotation angles of the scan lines. It is assumed that the satellite travels in a smooth motion as a scene is being scanned. Therefore, once the exterior orientation of the center scan line is determined, the exterior orientation of any other scan line is calculated based on the distance of that scan line from the center, and the changes of the perspective center location and rotation angles.

Bundle adjustment for triangulating a satellite scene is similar to the bundle adjustment used for aerial images. A least squares adjustment is used to derive a set of parameters that comes the closest to fitting the control points to their known ground coordinates, and to intersecting tie points.

Photogrammetric Concepts

Collinearity Equations and Satellite Block Triangulation

The resulting parameters of satellite bundle adjustment are:

- Ground coordinates of the perspective center of the center scan line
- Rotation angles for the center scan line
- Coefficients, from which the perspective center and rotation angles of all other scan lines are calculated
- Ground coordinates of all tie points

Modified collinearity equations are used to compute the exterior orientation parameters associated with the respective scan lines in the satellite scenes. Each scan line has a unique perspective center and individual rotation angles. When the satellite moves from one scan line to the next, these parameters change. Due to the smooth motion of the satellite in orbit, the changes are small and can be modeled by low order polynomial functions.

Control for Satellite Block Triangulation

Both GCPs and tie points can be used for satellite block triangulation of a stereo scene. For triangulating a single scene, only GCPs are used. In this case, space resection techniques are used to compute the exterior orientation parameters associated with the satellite as they existed at the time of image capture. A minimum of six GCPs is necessary. Ten or more GCPs are recommended to obtain a good triangulation result.

The best locations for GCPs in the scene are shown below in Figure 7-22.

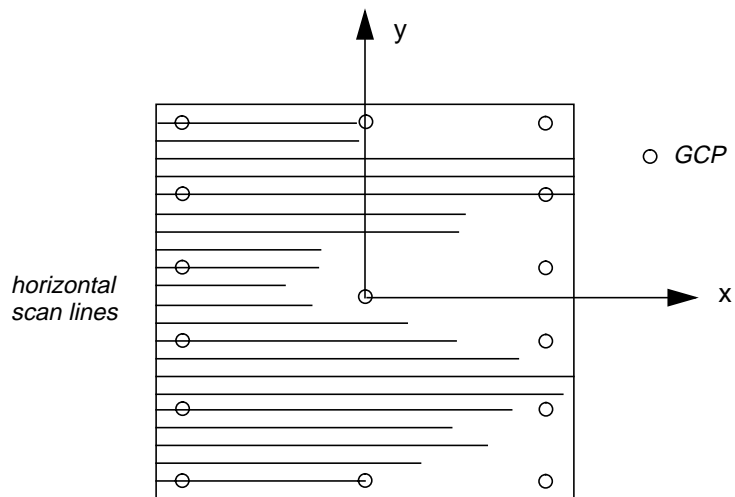


Figure 7-22 Ideal Point Distribution Over a Satellite Scene for Triangulation

Orthorectification

As stated previously, orthorectification is the process of removing geometric errors inherent within photography and imagery. The variables contributing to geometric errors include, but are not limited to:

- Camera and sensor orientation
- Systematic error associated with the camera or sensor
- Topographic relief displacement
- Earth curvature

By performing block triangulation or single frame resection, the parameters associated with camera and sensor orientation are defined. Utilizing least squares adjustment techniques during block triangulation minimizes the errors associated with camera or sensor instability. Additionally, the use of self-calibrating bundle adjustment (SCBA) techniques along with Additional Parameter (AP) modeling accounts for the systematic errors associated with camera interior geometry. The effects of the Earth's curvature are significant if a large photo block or satellite imagery is involved. They are accounted for during the block triangulation procedure by setting the relevant option. The effects of topographic relief displacement are accounted for by utilizing a DEM during the orthorectification procedure.

The orthorectification process takes the raw digital imagery and applies a DEM and triangulation results to create an orthorectified image. Once an orthorectified image is created, each pixel within the image possesses geometric fidelity. Thus, measurements taken off an orthorectified image represent the corresponding measurements as if they were taken on the Earth's surface (see Figure 7-23).

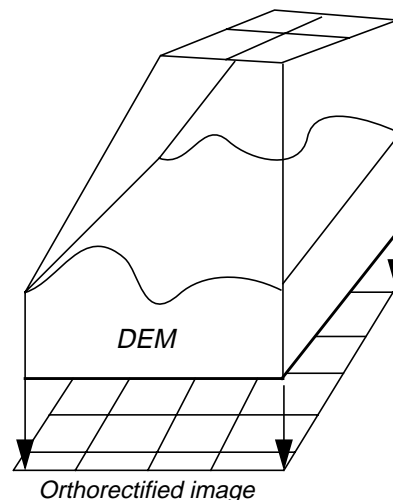


Figure 7-23 Orthorectification

Photogrammetric Concepts

An image or photograph with an orthographic projection is one for which every point looks as if an observer were looking straight down at it, along a line of sight that is orthogonal (perpendicular) to the Earth. The resulting orthorectified image is known as a digital orthoimage (see Figure 7-24).

Relief displacement is corrected by taking each pixel of a DEM and finding the equivalent position in the satellite or aerial image. A brightness value is determined for this location based on resampling of the surrounding pixels. The brightness value, elevation, and exterior orientation information are used to calculate the equivalent location in the orthoimage file.

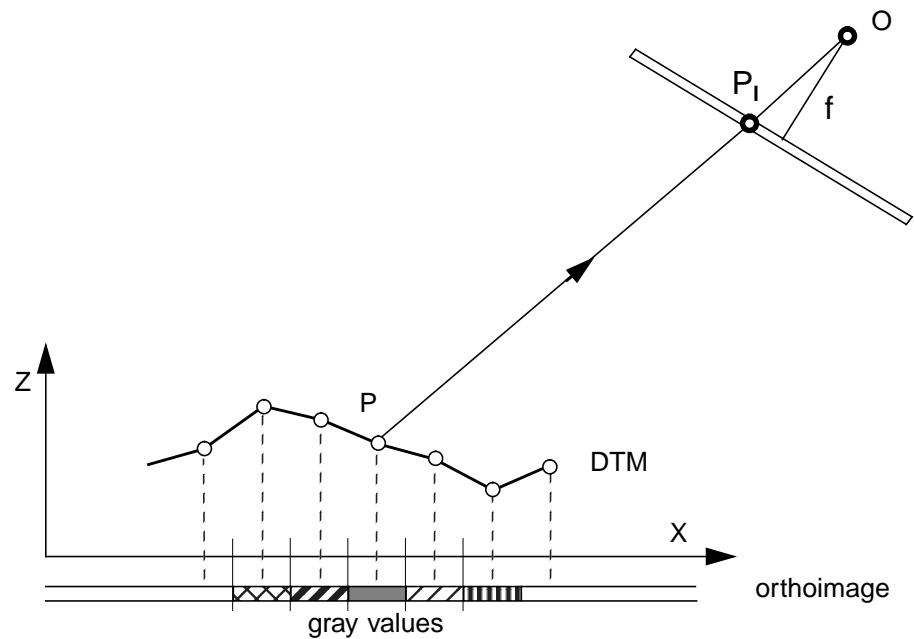


Figure 7-24 Digital Orthophoto—Finding Gray Values

Where:

- P = ground point
- P_1 = image point
- O = perspective center (origin)
- X,Z = ground coordinates (in DTM file)
- f = focal length

In contrast to conventional rectification techniques, orthorectification relies on the digital elevation data, unless the terrain is flat. Various sources of elevation data exist, such as the USGS DEM and a DEM automatically created from stereo image pairs. They are subject to data uncertainty, due in part to the generalization or imperfections in the creation process. The quality of the digital orthoimage is significantly affected by this uncertainty. For different image data, different accuracy levels of DEMs are required to limit the uncertainty-related errors within a controlled limit. While the near-vertical viewing SPOT scene can use very coarse DEMs, images with large incidence angles need better elevation data such as USGS level-1 DEMs. For aerial photographs with a scale larger than 1:60000, elevation data accurate to 1 meter is recommended. The 1 meter accuracy reflects the accuracy of the Z coordinates in the DEM, not the DEM resolution or posting.



Detailed discussion of DEM requirements for orthorectification can be found in Yang and Williams (1997). See "Bibliography".

Resampling methods used are nearest neighbor, bilinear interpolation, and cubic convolution. Generally, when the cell sizes of orthoimage pixels are selected, they should be similar or larger than the cell sizes of the original image. For example, if the image was scanned at 25 microns (1016 dpi) producing an image of 9K X 9K pixels, one pixel would represent 0.025 mm on the image. Assuming that the image scale of this photo is 1:40000, then the cell size on the ground is about 1 m. For the orthoimage, it is appropriate to choose a pixel spacing of 1 m or larger. Choosing a smaller pixel size oversamples the original image.



For information, see the scanning resolutions table, Table 7-1, on page 268.

For SPOT Pan images, a cell size of 10 meters is appropriate. Any further enlargement from the original scene to the orthophoto does not improve the image detail. For IRS-1C images, a cell size of 6 meters is appropriate.

CHAPTER 8

Radar Concepts

Introduction

Radar images are quite different from other remotely sensed imagery you might use with ERDAS IMAGINE software. For example, radar images may have speckle noise. Radar images, do, however, contain a great deal of information. ERDAS IMAGINE has many radar packages, including IMAGINE Radar Interpreter, IMAGINE OrthoRadar, IMAGINE StereoSAR DEM, IMAGINE IFSAR DEM, and the Generic SAR Node with which you can analyze your radar imagery. You have already learned about the various methods of speckle suppression—those are IMAGINE Radar Interpreter functions.

This chapter tells you about the advanced radar processing packages that ERDAS IMAGINE has to offer. The following sections go into detail about the geometry and functionality of those modules of the IMAGINE Radar Mapping Suite.

IMAGINE OrthoRadar Theory

Parameters Required for Orthorectification

SAR image orthorectification requires certain information about the sensor and the SAR image. Different sensors (RADARSAT, ERS, etc.) express these parameters in different ways and in different units. To simplify the design of our SAR tools and easily support future sensors, all SAR images and sensors are described using our Generic SAR model. The sensor-specific parameters are converted to a Generic SAR model on import.

The following table lists the parameters of the Generic SAR model and their units. These parameters can be viewed in the SAR Parameters tab on the main Generic SAR Model Properties (IMAGINE OrthoRadar) dialog.

Table 8-1 SAR Parameters Required for Orthorectification

Parameter	Description	Units
sensor	The sensor that produced the image (RADARSAT, ERS, etc.)	
coord_sys	Coordinate system for ephemeris (I = inertial, F = fixed body or Earth rotating)	
year	Year of data collection	
month	Month of data collection	
day	Day of data collection	
doy	GMT day of year of data collection	
num_samples	Number of samples in each line in the image (= number of samples in range)	

Table 8-1 SAR Parameters Required for Orthorectification (Continued)

Parameter	Description	Units
num_lines	Number of lines in the image (= number of lines in azimuth)	
image_start_time	Image start time in seconds of day	sec
first_pt_secs_of_day	Time of first ephemeris point provided in seconds of day. This parameter is updated during orbit adjustment	sec
first_pt_org	The original time of the first ephemeris point provided in seconds of day	
time_interval	Time interval between ephemeris points	sec
time_interval_org	Time interval between ephemeris points. This parameter is updated during orbit adjustment	sec
image_duration	Slow time duration of image	sec
image_end_time	Same as image_start_time + image_duration	sec
semimajor	Semimajor axis of Earth model used during SAR processing	m
semiminor	Semiminor axis of Earth model used during SAR processing	m
target_height	Assumed height of scene above Earth model used during SAR processing	m
look_side	Side to which sensor is pointed (left or right). Sometimes called sensor clock angle where -90 deg is left-looking and 90 deg is right-looking	deg
wavelength	Wavelength of sensor	m
sampling_rate	Range sampling rate	Hz
range_pix_spacing	Slant or ground range pixel spacing	m
near_slant_range	Slant range to near range pixel	m
num_pos_pts	Number of ephemeris points provided	
projection	Slant or ground range projection	
gnd2slt_coeffs[6]	Coefficients used in polynomial transform from ground to slant range. Used when the image is in ground range	
time_dir_pixels	Time direction in the pixel (range) direction	
time_dir_lines	Time direction in the line (azimuth) direction	
rsx, rsy, rsz	Array of spacecraft positions (x, y, z) in an Earth Fixed Body coordinate system	m
vsx, vsy, vsz	Array of spacecraft velocities (x, y, z) in an Earth Fixed Body coordinate system	m/sec
orbitState	Flag indicating if the orbit has been adjusted	

Table 8-1 SAR Parameters Required for Orthorectification (Continued)

Parameter	Description	Units
rs_coeffs[9]	Coefficients used to model the sensor orbit positions as a function of time	
vs_coeffs[9]	Coefficients used to model the sensor orbit velocities as a function of time	
sub_unity_subset	Flag indicating that the entire image is present	
sub_range_start	Indicates the starting range sample of the current raster relative to the original image	
sub_range_end	Indicates the ending range sample of the current raster relative to the original image	
sub_range_degrade	Indicates the range sample degrade factor of the current raster relative to the original image	
sub_range_num_samples	Indicates the range number of samples of the current raster	
sub_azimuth_start	Indicates the starting azimuth line of the current raster relative to the original image	
sub_azimuth_end	Indicates the ending azimuth line of the current raster relative to the original image	
sub_azimuth_degrade	Indicates the azimuth degrade factor of the current raster relative to the original image	
sub_azimuth_num_lines	Indicates the azimuth number of lines	

Algorithm Description

Overview

The orthorectification process consists of several steps:

- ephemeris modeling and refinement (if GCPs are provided)
- sparse mapping grid generation
- output formation (including terrain corrections)

Each of these steps is described in detail in the following sections.

Ephemeris Coordinate System

The positions and velocities of the spacecraft are internally assumed to be in an Earth Fixed Body coordinate system. If the ephemeris are provided in inertial coordinate system, IMAGINE OrthoRadar converts them from inertial to Earth Fixed Body coordinates.

The Earth Fixed Body coordinate system is an Earth-centered Cartesian coordinate system that rotates with the Earth. The x-axis radiates from the center of the Earth through the 0 longitude point on the equator. The z-axis radiates from the center of the Earth through the geographic North Pole. The y-axis completes the right-handed Cartesian coordinate system.

Ephemeris Modeling

The platform ephemeris is described by three or more platform locations and velocities. To predict the platform position and velocity at some time (t):

$$R_{s,x} = a_1 + a_2 t + a_3 t^2$$

$$R_{s,y} = b_1 + b_2 t + b_3 t^2$$

$$R_{s,z} = c_1 + c_2 t + c_3 t^2$$

$$V_{s,x} = d_1 + d_2 t + d_3 t^2$$

$$V_{s,y} = e_1 + e_2 t + e_3 t^2$$

$$V_{s,z} = f_1 + f_2 t + f_3 t^2$$

Where R_s is the sensor position and V_s is the sensor velocity:

$$R_s = [R_{s,x} \ R_{s,y} \ R_{s,z}]^T$$

$$V_s = [V_{s,x} \ V_{s,y} \ V_{s,z}]^T$$

To determine the model coefficients $\{a_i, b_i, c_i\}$ and $\{d_i, e_i, f_i\}$, first do some preprocessing. Select the best three consecutive data points prior to fitting (if more than three points are available). The best three data points must span the entire image in time. If more than one set of three data points spans the image, then select the set of three that has a center time closest to the center time of the image.

Once a set of three consecutive data points is found, model the ephemeris with an exact solution.

Form matrix A :

$$A = \begin{bmatrix} 1.0 & t_1 & t_1^2 \\ 1.0 & t_2 & t_2^2 \\ 1.0 & t_3 & t_3^2 \end{bmatrix}$$

Where t_1 , t_2 , and t_3 are the times associated with each platform position. Select t such that $t = 0.0$ corresponds to the time of the second position point. Form vector b :

$$b = [R_{s,x}(1) \ R_{s,x}(2) \ R_{s,x}(3)]^T$$

Where $R_{s,x}(i)$ is the x-coordinate of the i-th platform position ($i=1:3$). We wish to solve $Ax = b$ where x is:

$$x = [a_1 \ a_2 \ a_3]^T$$

To do so, use LU decomposition. The process is repeated for: $R_{s,y}$, $R_{s,z}$, $V_{s,x}$, $V_{s,y}$, and $V_{s,z}$

SAR Imaging Model

Before discussing the ephemeris adjustment, it is important to understand how to get from a pixel in the SAR image (as specified by a range line and range pixel) to a target position on the Earth [specified in Earth Centered System (ECS) coordinates or x, y, z]. This process is used throughout the ephemeris adjustment and the orthorectification itself.

For each range line and range pixel in the SAR image, the corresponding target location (\mathbf{R}_t) is determined. The target location can be described as (lat, lon, elev) or (x, y, z) in ECS. The target can either lie on a smooth Earth ellipsoid or on a smooth Earth ellipsoid plus an elevation model.

In either case, the location of \mathbf{R}_t is determined by finding the intersection of the Doppler cone, range sphere, and Earth model. In order to do this, first find the Doppler centroid and slant range for a given SAR image pixel.

Let i = range pixel and j = range line.

Time

Time $T(j)$ is thus:

$$T(j) = T(0) + \frac{j-1}{N_a-1} * t_{dur}$$

Where $T(0)$ is the image start time, N_a is the number of range lines, and t_{dur} is the image duration time.

Doppler Centroid

The computation of the Doppler centroid f_d to use with the SAR imaging model depends on how the data was processed. If the data was deskewed, this value is always 0. If the data is skewed, then this value may be a nonzero constant or may vary with i .

Slant Range

The computation of the slant range to the pixel i depends on the projection of the image. If the data is in a slant range projection, then the computation of slant range is straightforward:

$$R_{sl}(i) = r_{sl} + (i-1) * \Delta r_{sr}$$

Where $\mathbf{R}_{sl}(i)$ is the slant range to pixel i , r_{sl} is the near slant range, and Δr_{sr} is the slant range pixel spacing.

If the projection is a ground range projection, then this computation is potentially more complicated and depends on how the data was originally projected into a ground range projection by the SAR processor.

Intersection of Doppler Cone, Range Sphere, and Earth Model

To find the location of the target \mathbf{R}_t corresponding to a given range pixel and range line, the intersection of the Doppler cone, range sphere, and Earth model must be found. For an ellipsoid, these may be described as follows:

$$f_D = \frac{2}{\lambda R_{sl}} (\bar{\mathbf{R}}_s - \bar{\mathbf{R}}_t) \cdot (\bar{\mathbf{V}}_s - \bar{\mathbf{V}}_t), \quad f_D > 0 \text{ for forward squint}$$

$$R_{sl} = | \mathbf{R}_s - \mathbf{R}_t |$$

$$\frac{R_s(x)^2 + R_s(y)^2}{(R_e + h_{\text{targ}})^2} + \frac{R_s(z)^2}{(R_m + h_{\text{targ}})^2} = 1$$

Where \mathbf{R}_s and \mathbf{V}_s are the platform position and velocity respectively, \mathbf{V}_t is the target velocity ($= \mathbf{0}$, in this coordinate system), R_e is the Earth semimajor axis, and R_m is the Earth semiminor axis. The platform position and velocity vectors \mathbf{R}_s and \mathbf{V}_s can be found as a function of time $\mathbf{T}(j)$ using the ephemeris equations developed previously

Figure 8.1 Graphically illustrates the solution for the target location given the sensor ephemeris, doppler cone, range sphere, and flat Earth model.

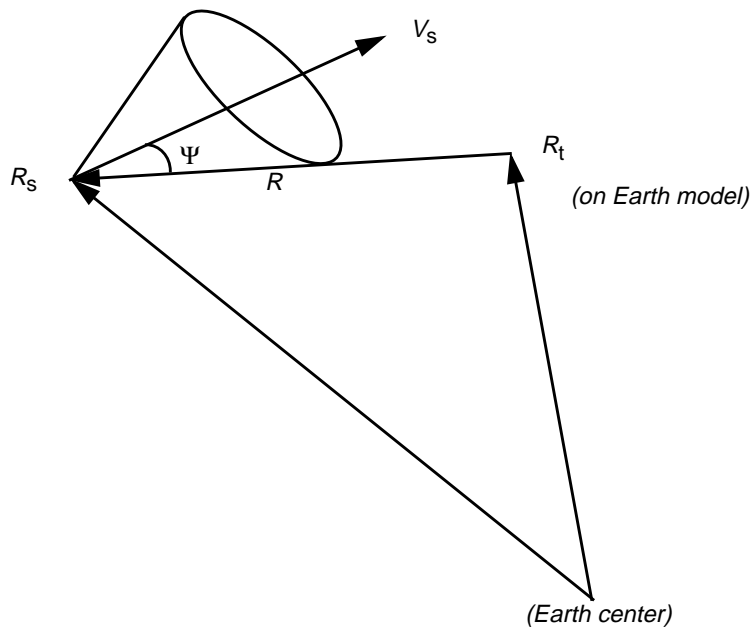


Figure 8-1 Doppler Cone

Ephemeris Adjustment

There are three possible adjustments that can be made: along track, cross track, and radial. In IMAGINE OrthoRadar, the along track adjustment is performed separately. The cross track and radial adjustments are made simultaneously. These adjustments are made using residuals associated with GCPs. Each GCP has a map coordinate (such as lat, lon) and an elevation. Also, an SAR image range line and range pixel must be given. The SAR image range line and range pixel are converted to \mathbf{R}_t using the method described previously (substituting $h_{\text{targ}} = \text{elevation of GCP above ellipsoid used in SAR processing}$).

The along track adjustment is computed first, followed by the cross track and radial adjustments. The two adjustment steps are then repeated.

AOrthorectification

The ultimate goal in orthorectification is to determine, for a given target location on the ground, the associated range line and range pixel from the input SAR image, including the effects of terrain.

To do this, there are several steps. First, take the target location and locate the associated range line and range pixel from the input SAR image assuming smooth terrain. This places you in approximately the correct range line. Next, look up the elevation at the target from the input DEM. The elevation, in combination with the known slant range to the target, is used to determine the correct range pixel. The data can now be interpolated from the input SAR image.

Sparse Mapping Grid

Select a block size \mathbf{M} . For every \mathbf{M} th range line and \mathbf{M} th range pixel, compute \mathbf{R}_t on a smooth ellipsoid (using the SAR Earth model), and save these values in an array. Smaller \mathbf{M} implies less distortion between grid points.

Regardless of \mathbf{M} and the total number of samples and lines in the input SAR image, always compute \mathbf{R}_t at the end of every line and for the very last line. The spacing between points in the sparse mapping grid is regular except at the far edges of the grid.

Output Formation

For each point in the output grid, there is an associated \mathbf{R}_t . This target should fall on the surface of the Earth model used for SAR processing, thus a conversion is made between the Earth model used for the output grid and the Earth model used during SAR processing.

The process of orthorectification starts with a location on the ground. The line and pixel location of the pixel to this map location can be determined from the map location and the sparse mapping grid. The value at this pixel location is then assigned to the map location. Figure 8-4 illustrates this process.

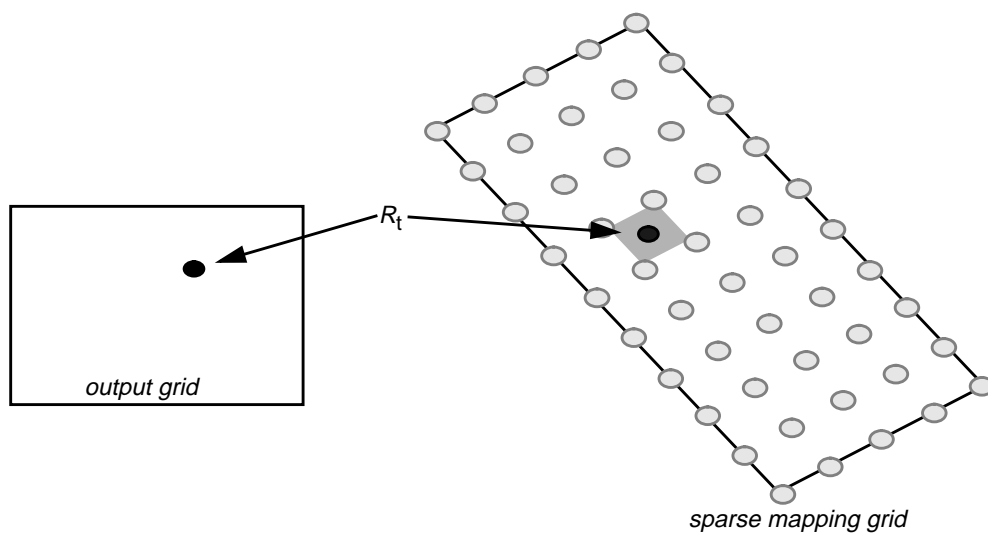


Figure 8-2 Sparse Mapping and Output Grids

**IMAGINE StereoSAR
DEM Theory**

Introduction

This chapter details the theory that supports IMAGINE StereoSAR DEM processing.

To understand the way IMAGINE StereoSAR DEM works to create DEMs, it is first helpful to look at the process from beginning to end. Figure 8-3 shows a stylized process for basic operation of the IMAGINE StereoSAR DEM module.

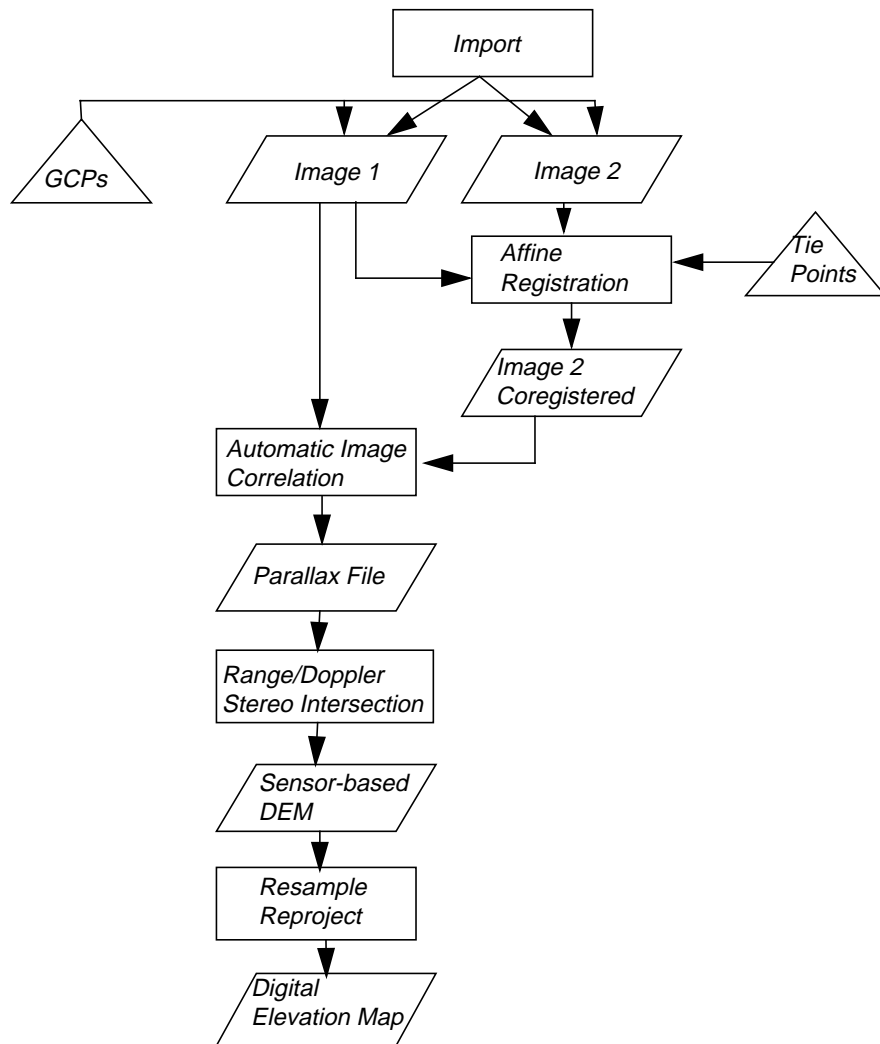


Figure 8-3 IMAGINE StereoSAR DEM Process Flow

The following discussion includes algorithm descriptions as well as discussion of various processing options and selection of processing parameters.

Input

There are many elements to consider in the Input step. These include beam mode selection, importing files, orbit correction, and ephemeris data.

Beam Mode Selection

Final accuracy and precision of the DEM produced by the IMAGINE StereoSAR DEM module is predicated on two separate calculation sequences. These are the automatic image correlation and the sensor position/triangulation calculations. These two calculation sequences are joined in the final step: Height.

The two initial calculation sequences have disparate beam mode demands. Automatic correlation works best with images acquired with as little angular divergence as possible. This is because different imaging angles produce different-looking images, and the automatic correlator is looking for image similarity. The requirement of image similarity is the same reason images acquired at different times can be hard to correlate. For example, images taken of agricultural areas during different seasons can be extremely different and, therefore, difficult or impossible for the automatic correlator to process successfully.

Conversely, the triangulation calculation is most accurate when there is a large intersection angle between the two images (see Figure 8-4). This results in images that are truly different due to geometric distortion. The ERDAS IMAGINE automatic image correlator has proven sufficiently robust to match images with significant distortion if the proper correlator parameters are used.

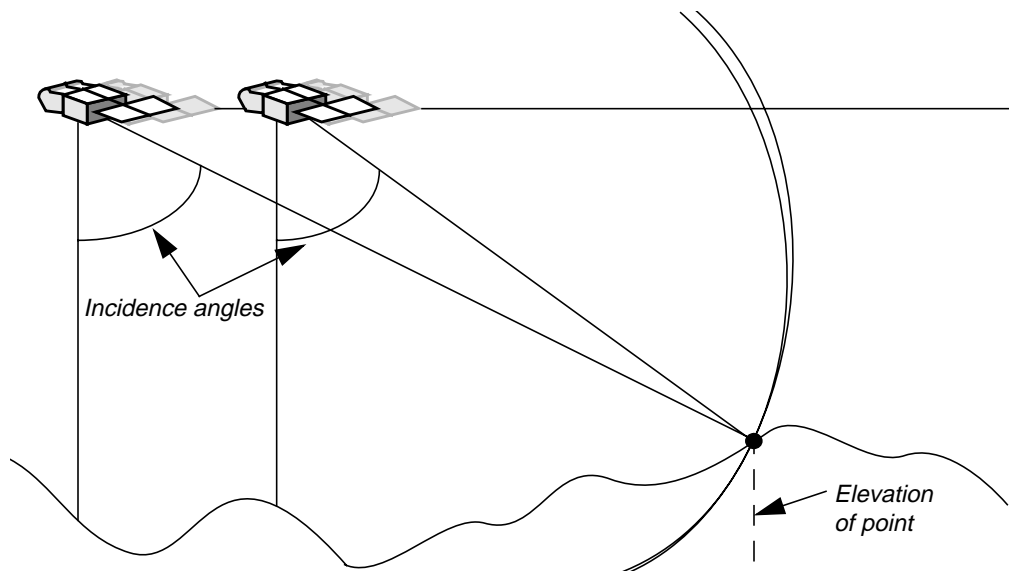


Figure 8-4 SAR Image Intersection

NOTE: IMAGINE StereoSAR DEM has built-in checks that assure the sensor associated with the Reference image is closer to the imaged area than the sensor associated with the Match image.

A third factor, cost effectiveness, must also often be evaluated. First, select either Fine or Standard Beam modes. Fine Beam images with a pixel size of six meters would seem, at first glance, to offer a much better DEM than Standard Beam with 12.5-meter pixels. However, a Fine Beam image covers only one-fourth the area of a Standard Beam image and produces a DEM only minimally better.

Various Standard Beam combinations, such as an S3/S6 or an S3/S7, cover a larger area per scene, but only for the overlap area which might be only three-quarters of the scene area. Testing at both ERDAS and RADARSAT has indicated that a stereopair consisting of a Wide Beam mode 2 image and a Standard Beam mode 7 image produces the most cost-effective DEM at a resolution consistent with the resolution of the instrument and the technique.

Import

The imagery required for the IMAGINE StereoSAR DEM module can be imported using the ERDAS IMAGINE radar-specific importers for either RADARSAT or ESA (ERS-1, ERS-2). These importers automatically extract data from the image header files and store it in an Hfa file attached to the image. In addition, they abstract key parameters necessary for sensor modeling and attach these to the image as a Generic SAR Node Hfa file. Other radar imagery (e.g., SIR-C) can be imported using the Generic Binary Importer. The Generic SAR Node can then be used to attach the Generic SAR Node Hfa file.

Orbit Correction

Extensive testing of both the IMAGINE OrthoRadar and IMAGINE StereoSAR DEM modules has indicated that the ephemeris data from the RADARSAT and the ESA radar satellites is very accurate (see appended accuracy reports). However, the accuracy does vary with each image, and there is no *a priori* way to determine the accuracy of a particular data set.

The modules of the IMAGINE Radar Mapping Suite: IMAGINE OrthoRadar, IMAGINE StereoSAR DEM, and IMAGINE IFSAR DEM, allow for correction of the sensor model using GCPs. Since the supplied orbit ephemeris is very accurate, orbit correction should only be attempted if you have very good GCPs. In practice, it has been found that GCPs from 1:24 000 scale maps or a handheld GPS are the minimum acceptable accuracy. In some instances, a single accurate GCP has been found to result in a significant increase in accuracy.

As with image warping, a uniform distribution of GCPs results in a better overall result and a lower RMS error. Again, accurate GCPs are an essential requirement. If your GCPs are questionable, you are probably better off not using them. Similarly, the GCP must be recognizable in the radar imagery to within plus or minus one to two pixels. Road intersections, reservoir dams, airports, or similar man-made features are usually best. Lacking one very accurate and locatable GCP, it would be best to utilize several good GCPs dispersed throughout the image as would be done for a rectification.

Ellipsoid vs. Geoid Heights

The IMAGINE Radar Mapping Suite is based on the World Geodetic System (WGS) 84 Earth ellipsoid. The sensor model uses this ellipsoid for the sensor geometry. For maximum accuracy, all GCPs used to refine the sensor model for all IMAGINE Radar Mapping Suite modules (IMAGINE OrthoRadar, IMAGINE StereoSAR DEM, or IMAGINE IFSAR DEM) should be converted to this ellipsoid in all three dimensions: latitude, longitude, and elevation.

Note that, while ERDAS IMAGINE reprojection converts latitude and longitude to UTM WGS 84 for many input projections, it does not modify the elevation values. To do this, it is necessary to determine the elevation offset between WGS 84 and the datum of the input GCPs. For some input datums this can be accomplished using the Web site: www.ngs.noaa.gov/GEOID/geoid.html. This offset must then be added to, or subtracted from, the input GCP. Many handheld GPS units can be set to output in WGS 84 coordinates.

One elegant feature of the IMAGINE StereoSAR DEM module is that orbit refinement using GCPs can be applied at any time in the process flow without losing the processing work to that stage. The stereopair can even be processed all the way through to a final DEM and then you can go back and refine the orbit. This refined orbit is transferred through all the intermediate files (Subset, Despeckle, etc.). Only the final step Height would need to be rerun using the new refined orbit model.

The ephemeris normally received with RADARSAT, or ERS-1 and ERS-2 imagery is based on an extrapolation of the sensor orbit from previous positions. If the satellite received an orbit correction command, this effect might not be reflected in the previous position extrapolation. The receiving stations for both satellites also do ephemeris calculations that include post image acquisition sensor positions. These are generally more accurate. They are not, unfortunately, easy to acquire and attach to the imagery.

Refined Ephemeris

For information about Refined Ephemeris, see "IMAGINE IFSAR DEM Theory" on page 326.

Subset

Use of the Subset option is straightforward. It is not necessary that the two subsets define exactly the same area: an approximation is acceptable. This option is normally used in two circumstances. First, it can be used to define a small subset for testing correlation parameters prior to running a full scene. Also, it would be used to constrain the two input images to only the overlap area. Constraining the input images is useful for saving data space, but is not necessary for functioning of IMAGINE StereoSAR DEM — it is purely optional.

Despeckle

The functions to despeckle the images prior to automatic correlation are optional. The rationale for despeckling at this time is twofold. One, image speckle noise is not correlated between the two images: it is randomly distributed in both. Thus, it only serves to confuse the automatic correlation calculation. Presence of speckle noise could contribute to false positives during the correlation process.

Secondly, as discussed under Beam Mode Selection, the two images the software is trying to match are different due to viewing geometry differences. The slight low pass character of the despeckle algorithm may actually move both images toward a more uniform appearance, which aids automatic correlation.

Functionally, the despeckling algorithms presented here are identical to those available in the IMAGINE Radar Interpreter. In practice, a 3×3 or 5×5 kernel has been found to work acceptably. Note that all ERDAS IMAGINE speckle reduction algorithms allow the kernel to be tuned to the image being processed via the Coefficient of Variation. Calculation of this parameter is accessed through the IMAGINE Radar Interpreter Speckle Suppression interface.



See the ERDAS IMAGINE Tour Guides for the IMAGINE Radar Interpreter tour guide.

Degrade

The Degrade option offered at this step in the processing is commonly used for two purposes. If the input imagery is Single Look Complex (SLC), the pixels are not square (this is shown as the Range and Azimuth pixel spacing sizes). It may be desirable at this time to adjust the Y scale factor to produce pixels that are more square. This is purely an option; the software accurately processes undegraded SLC imagery.

Secondly, if data space or processing time is limited, it may be useful to reduce the overall size of the image file while still processing the full images. Under those circumstances, a reduction of two or three in both X and Y might be appropriate. Note that the processing flow recommended for maximum accuracy processes the full resolution scenes and correlates for every pixel. Degrade is used subsequent to Match to lower DEM variance (LE90) and increase pixel size to approximately the desired output posting.

Rescale

This operation converts the input imagery bit format, commonly unsigned 16-bit, to unsigned 8-bit using a two standard deviations stretch. This is done to reduce the overall data file sizes. Testing has not shown any advantage to retaining the original 16-bit format, and use of this option is routinely recommended.

Register

Register is the first of the Process Steps (other than Input) that must be done. This operation serves two important functions, and proper user input at this processing level affects the speed of subsequent processing, and may affect the accuracy of the final output DEM.

The registration operation uses an affine transformation to rotate the Match image so that it more closely aligns with the Reference image. The purpose is to adjust the images so that the elevation-induced pixel offset (parallax) is mostly in the range (x-axis) direction (i.e., the images are nearly epipolar). Doing this greatly reduces the required size of the search window in the Match step.

One output of this step is the minimum and maximum parallax offsets, in pixels, in both the x- and y-axis directions. These values must be recorded by the operator and are used in the Match step to tune the IMAGINE StereoSAR DEM correlator parameter file (.ssc). These values are critical to this tuning operation and, therefore, must be correctly extracted from the Register step.

Two basic guidelines define the selection process for the tie points used for the registration. First, as with any image-to-image registration, a better result is obtained if the tie points are uniformly distributed throughout the images. Second, since you want the calculation to output the minimum and maximum parallax offsets in both the x- and y-axis directions, the tie points selected must be those that have the minimum and maximum parallax.

In practice, the following procedure has been found successful. First, select a fairly uniform grid of about eight tie points that defines the lowest elevation within the image. Coastlines, river flood plains, roads, and agricultural fields commonly meet this criteria. Use of the Solve Geometric Model icon on the StereoSAR Registration Tool should yield values in the -5 to +5 range at this time. Next, identify and select three or four of the highest elevations within the image. After selecting each tie point, click the Solve Geometric Model icon and note the effect of each tie point on the minimum and maximum parallax values. When you feel you have quantified these values, write them down and apply the resultant transform to the image.

Constrain

This option is intended to allow you to define areas where it is not necessary to search the entire search window area. A region of lakes would be such an area. This reduces processing time and also minimizes the likelihood of finding false positives. This option is not implemented at present.

Match

An essential component, and the major time-saver of the IMAGINE StereoSAR DEM software is automatic image correlation.

In automatic image correlation, a small subset (image chip) of the Reference image termed the template (see Figure 8-5), is compared to various regions of the Match image's search area (Figure 8-6) to find the best Match point. The center pixel of the template is then said to be correlated with the center pixel of the Match region. The software then proceeds to the next pixel of interest, which becomes the center pixel of the new template.

Figure 8-5 shows the upper left (UL) corner of the Reference image. An 11×11 pixel template is shown centered on the pixel of interest: $X = 8, Y = 8$.

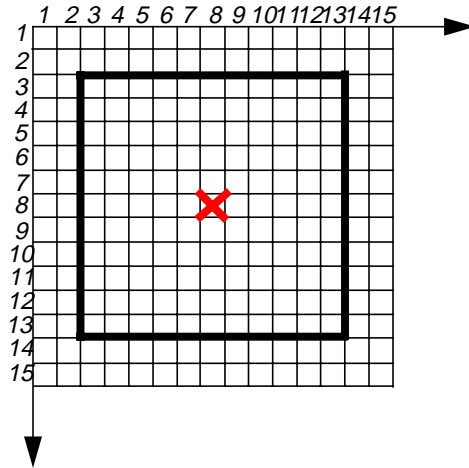


Figure 8-5 UL Corner of the Reference Image

Figure 8-6 shows the UL corner of the Match image. The 11×11 pixel template is shown centered on the initial estimated correlation pixel $X = 8, Y = 8$. The 15×7 pixel search area is shown in a dashed line. Since most of the parallax shift is in the range direction (x-axis), the search area should always be a rectangle to minimize search time.

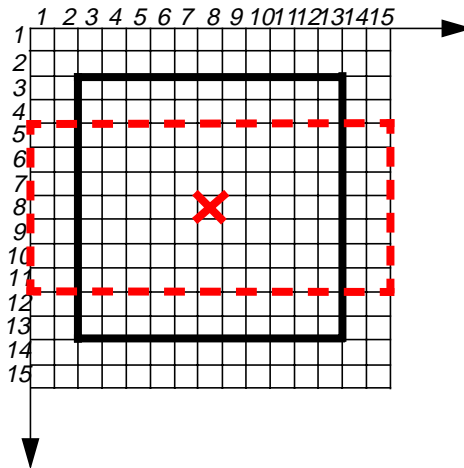


Figure 8-6 UL Corner of the Match Image

The ERDAS IMAGINE automatic image correlator works on the hierarchical pyramid technique. This means that the image is successively reduced in resolution to provide a coregistered set of images of increasing pixel size (see Figure 8-7). The automatic correlation software starts at the top of the resolution pyramid with the lowest resolution image being processed first. The results of this process are filtered and interpolated before being passed to the next highest resolution layer as the initial estimated correlation point. From this estimated point, the search is performed on this higher resolution layer.

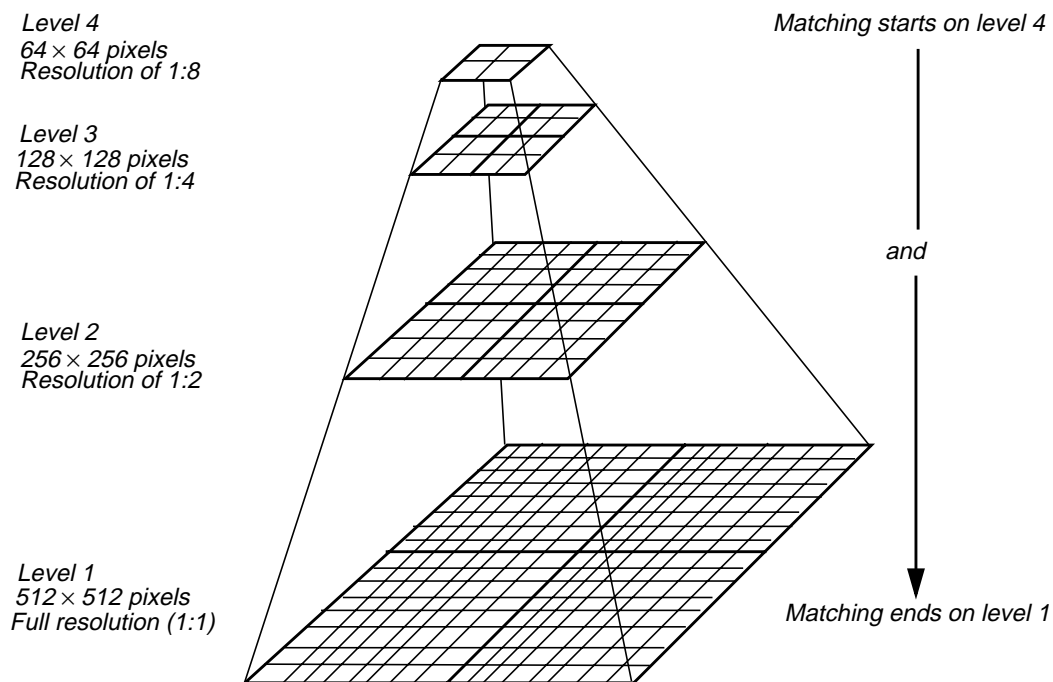


Figure 8-7 Image Pyramid

Template Size

The size of the template directly affects computation time: a larger image chip takes more time. However, too small of a template could contain insufficient image detail to allow accurate matching. A balance must be struck between these two competing criteria, and is somewhat image-dependent. A suitable template for a suburban area with roads, fields, and other features could be much smaller than the required template for a vast region of uniform ground cover. Because of viewing geometry-induced differences in the Reference and Match images, the template from the Reference image is never identical to any area of the Match image. The template must be large enough to minimize this effect.

The IMAGINE StereoSAR DEM correlator parameters shown in Table 8-2 are for the library file Std_LP_HD.ssc. These parameters are appropriate for a RADARSAT Standard Beam mode (Std) stereopair with low parallax (LP) and high density of detail (HD). The low parallax parameters are appropriate for images of low to moderate topography. The high density of detail (HD) parameters are appropriate for the suburban area discussed above.

Table 8-2 STD_LP_HD Correlator

Level	Average	Size X	Size Y	Search -X	Search +X	Search -Y	Search +Y
1	1	20	20	2	2	1	1
2	2	60	60	3	4	1	1
3	3	90	90	8	20	2	3
4	4	120	120	10	30	2	5
5	5	180	180	20	60	2	8
6	6	220	220	25	70	3	10

Level	Step X	Step Y	Threshold	Value	Vector X	Vector Y	Applied
1	2	2	0.30000	0.00000	0.00000	0.00000	0
2	8	8	0.20000	0.00000	0.00000	0.00000	0
3	20	20	0.20000	0.00000	0.00000	0.00000	0
4	50	50	0.20000	0.00000	0.00000	0.00000	0
5	65	65	0.20000	0.00000	0.00000	0.00000	0
6	80	80	0.10000	0.00000	0.00000	0.00000	0

Note that the size of the template (Size X and Size Y) increases as you go up the resolution pyramid. This size is the effective size if it were on the bottom of the pyramid (i.e., the full resolution image). Since they are actually on reduced resolution levels of the pyramid, they are functionally smaller. Thus, the 220 × 220 template on Level 6 is actually only 36 × 36 during the actual search. By stating the template size relative to the full resolution image, it is easy to display a box of approximate size on the input image to evaluate the amount of detail available to the correlator, and thus optimize the template sizes.

Search Area

Considerable computer time is expended in searching the Match image for the exact Match point. Thus, this search area should be minimized. (In addition, searching too large of an area increases the possibility of a false match.) For this reason, the software first requires that the two images be registered. This gives the software a rough idea of where the Match point might be. In stereo DEM generation, you are looking for the offset of a point in the Match image from its corresponding point in the Reference image (parallax). The minimum and maximum displacement is quantified in the Register step and is used to restrain the search area.

In Figure 8-6, the search area is defined by four parameters: $-X$, $+X$, $-Y$, and $+Y$. Most of the displacement in radar imagery is a function of the look angle and is in the range or x-axis direction. Thus, the search area is always a rectangle emphasizing the x-axis. Because the total search area (and, therefore, the total time) is X times Y , it is important to keep these values to a minimum. Careful use of the IMAGINE StereoSAR DEM software at the Register step easily achieves this.

Step Size

Because a radar stereopair typically contains millions of pixels, it is not desirable to correlate every pixel at every level of the hierarchical pyramid, nor is this even necessary to achieve an accurate result. The density at which the automatic correlator is to operate at each level in the resolution pyramid is determined by the step size (posting). The approach used is to keep posting tighter (smaller step size) as the correlator works down the resolution pyramid. For maximum accuracy, it is recommended to correlate every pixel at the full resolution level. This result is then compressed by the Degrade step to the desired DEM cell size.

Threshold

The degree of similarity between the Reference template and each possible Match region within the search area must be quantified by a mathematical metric. IMAGINE StereoSAR DEM uses the widely accepted normalized correlation coefficient. The range of possible values extends from -1 to $+1$, with $+1$ being an identical match. The algorithm uses the maximum value within the search area as the correlation point.

The threshold in Table 8-2 is the minimum numerical value of the normalized correlation coefficient that is accepted as a correlation point. If no value within the entire search area attains this minimum, there is not a Match point for that level of the resolution pyramid. In this case, the initial estimated position, passed from the previous level of the resolution pyramid, is retained as the Match point.

Correlator Library

To aid both the novice and the expert in rapidly selecting and refining an IMAGINE StereoSAR DEM correlator parameter file for a specific image pair, a library of tested parameter files has been assembled and is included with the software. These files are labeled using the following syntax: (RADARSAT Beam mode)_(Magnitude of Parallax)_(Density of Detail).

RADARSAT Beam Mode

Correlator parameter files are available for both Standard (Std) and Fine (Fine) Beam modes. An essential difference between these two categories is that, with the Fine Beam mode, more pixels are required (i.e., a larger template) to contain the same number of image features than with a Standard Beam image of the same area.

Magnitude of Parallax

The magnitude of the parallax is divided into high parallax (_HP) and low parallax (_LP) options. This determination is based upon the elevation changes and slopes within the images and is somewhat subjective. This parameter determines the size of the search area.

Density of Detail

The level of detail within each template is divided into high density (_HD) and low density (_LD) options. The density of detail for a suburban area with roads, fields, and other features would be much higher than the density of detail for a vast region of uniform ground cover. This parameter, in conjunction with beam mode, determines the required template sizes.

Quick Tests

It is often advantageous to quickly produce a low resolution DEM to verify that the automatic image correlator is optimum before correlating on every pixel to produce the final DEM.

For this purpose, a Quick Test (_QT) correlator parameter file has been provided for each of the full resolution correlator parameter files in the .ssc library. These correlators process the image only through resolution pyramid Level 3. Processing time up to this level has been found to be acceptably fast, and testing has shown that if the image is successfully processed to this level, the correlator parameter file is probably appropriate.

Evaluation of the parallax files produced by the Quick Test correlators and subsequent modification of the correlator parameter file is discussed in "IMAGINE StereoSAR DEM Application" in the *IMAGINE Radar Mapping Suite Tour Guide*.

Degrade

The second Degrade step compresses the final parallax image file (Level 1). While not strictly necessary, it is logical and has proven advantageous to reduce the pixel size at this time to approximately the intended posting of the final output DEM. Doing so at this time decreases the variance (LE90) of the final DEM through averaging.

Height

This step combines the information from the above processing steps to derive surface elevations. The sensor models of the two input images are combined to derive the stereo intersection geometry. The parallax values for each pixel are processed through this geometric relationship to derive a DEM in sensor (pixel) coordinates.

Comprehensive testing of the IMAGINE StereoSAR DEM module has indicated that, with reasonable data sets and careful work, the output DEM falls between DTED Level I and DTED Level II. This corresponds to between USGS 30-meter and USGS 90-meter DEMs. Thus, an output pixel size of 40 to 50 meters is consistent with this expected precision.

The final step is to resample and reproject this sensor DEM in to the desired final output DEM. The entire ERDAS IMAGINE reprojection package is accessed within the IMAGINE StereoSAR DEM module.

IMAGINE IFSAR DEM Theory

Introduction

Terrain height extraction is one of the most important applications for SAR images. There are two basic techniques for extracting height from SAR images: stereo and interferometry. Stereo height extraction is much like the optical process and is discussed in "IMAGINE StereoSAR DEM Theory" on page 315. The subject of this section is SAR interferometry (IFSAR).

Height extraction from IFSAR takes advantage of one of the unique qualities of SAR images: distance information from the sensor to the ground is recorded for every pixel in the SAR image. Unlike optical and IR images, which contain only the intensity of the energy received to the sensor, SAR images contain distance information in the form of phase. This distance is simply the number of wavelengths of the source radiation from the sensor to a given point on the ground. SAR sensors can record this information because, unlike optical and IR sensors, their radiation source is active and coherent.

Unfortunately, this distance phase information in a single SAR image is mixed with phase noise from the ground and other effects. For this reason, it is impossible to extract just the distance phase from the total phase in a single SAR image. However, if two SAR images are available that cover the same area from slightly different vantage points, the phase of one can be subtracted from the phase of the other to produce the distance difference of the two SAR images (hence the term interferometry). This is because the other phase effects for the two images are approximately equal and cancel out each other when subtracted. What is left is a measure of the distance difference from one image to the other. From this difference and the orbit information, the height of every pixel can be calculated.

This chapter covers basic concepts and processing steps needed to extract terrain height from a pair of interferometric SAR images.

Electromagnetic Wave Background

In order to understand the SAR interferometric process, one must have a general understanding of electromagnetic waves and how they propagate. An electromagnetic wave is a changing electric field that produces a changing magnetic field that produces a changing electric field, and so on. As this process repeats, energy is propagated through empty space at the speed of light.

Figure 8-8 gives a description of the type of electromagnetic wave that we are interested in. In this diagram, **E** indicates the electric field and **H** represents the magnetic field. The directions of **E** and **H** are mutually perpendicular everywhere. In a uniform plane, wave **E** and **H** lie in a plane and have the same value everywhere in that plane.

A wave of this type with both \mathbf{E} and \mathbf{H} transverse to the direction of propagation is called a Transverse ElectroMagnetic (TEM) wave. If the electric field \mathbf{E} has only a component in the y direction and the magnetic field \mathbf{H} has only a component in the z direction, then the wave is said to be polarized in the y direction (vertically polarized). Polarization is generally defined as the direction of the electric field component with the understanding that the magnetic field is perpendicular to it.

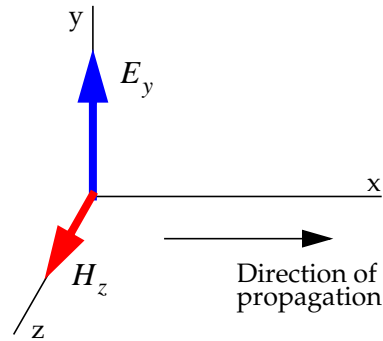


Figure 8-8 Electromagnetic Wave

The electromagnetic wave described above is the type that is sent and received by an SAR. The SAR, like most equipment that uses electromagnetic waves, is only sensitive to the electric field component of the wave; therefore, we restrict our discussion to it. The electric field of the wave has two main properties that we must understand in order to understand SAR and interferometry. These are the magnitude and phase of the wave. Figure 8-9 shows that the electric field varies with time.

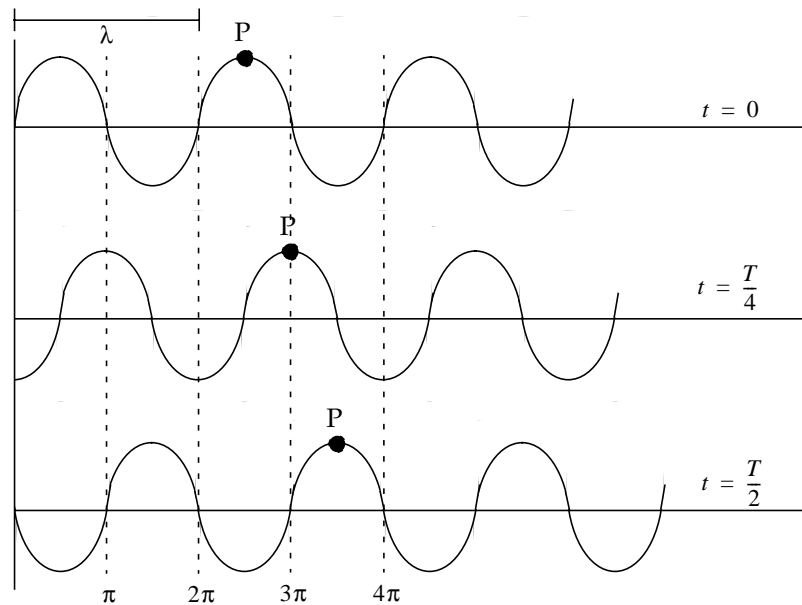


Figure 8-9 Variation of Electric Field in Time

The figure shows how the wave phase varies with time at three different moments. In the figure λ is the wavelength and T is the time required for the wave to travel one full wavelength. P is a point of constant phase and moves to the right as time progresses. The wave has a specific phase value at any given moment in time and at a specific point along its direction of travel. The wave can be expressed in the form of Equation 8-1.

Equation 8-1

$$E_y = \cos(\omega t + \beta x) \quad \text{Where}$$
$$w = \frac{2\pi}{T} \quad \text{and}$$
$$\beta = \frac{2\pi}{\lambda}$$

Equation 8-1 is expressed in Cartesian coordinates and assumes that the maximum magnitude of E_y is unity. It is more useful to express this equation in exponential form and include a maximum term as in Equation 8-2.

Equation 8-2

$$E_y = E_0 \cdot e^{j(\omega t \pm \beta x)}$$

So far we have described the definition and behavior of the electromagnetic wave phase as a function of time and distance. It is also important to understand how the strength or magnitude behaves with time and distance from the transmitter. As the wave moves away from the transmitter, its total energy stays the same but is spread over a larger distance. This means that the energy at any one point (or its energy density) decreases with time and distance as shown in Figure 8-10.

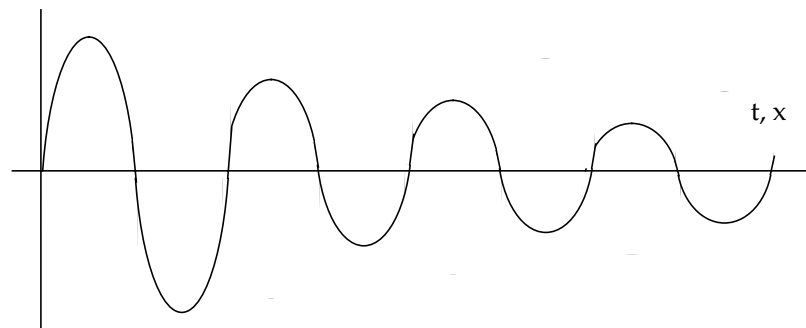


Figure 8-10 Effect of Time and Distance on Energy

The magnitude of the wave decreases exponentially as the distance from the transmitter increases. Equation 8-2 represents the general form of the electromagnetic wave that we are interested in for SAR and IFSAR applications. Later, we further simplify this expression given certain restrictions of an SAR sensor.

The Interferometric Model

Most uses of SAR imagery involve a display of the magnitude of the image reflectivity and discard the phase when the complex image is magnitude-detected. The phase of an image pixel representing a single scatterer is deterministic; however, the phase of an image pixel represents multiple scatterers (in the same resolution cell), and is made up of both a deterministic and nondeterministic, statistical part. For this reason, pixel phase in a single SAR image is generally not useful. However, with proper selection of an imaging geometry, two SAR images can be collected that have nearly identical nondeterministic phase components. These two SAR images can be subtracted, leaving only a useful deterministic phase difference of the two images.

Figure 8-11 provides the basic geometric model for an interferometric SAR system.

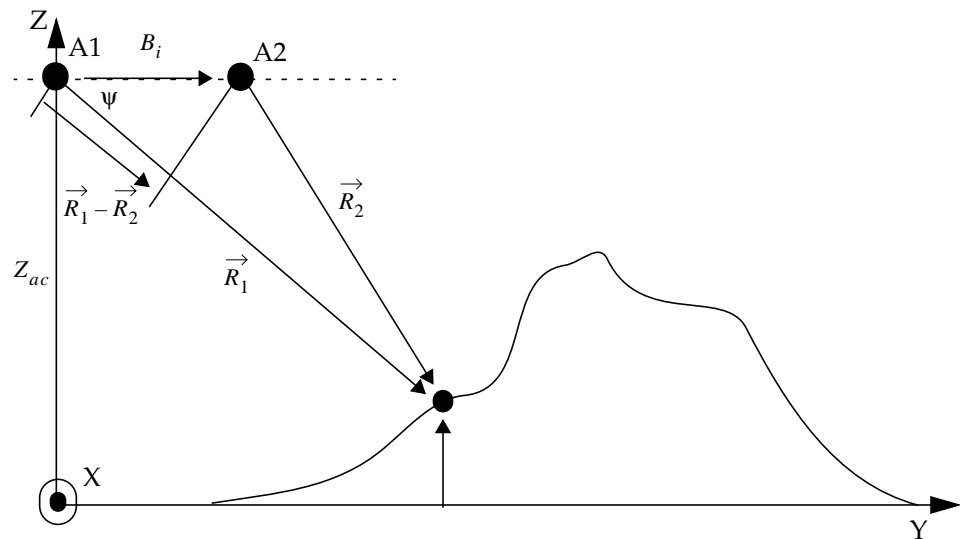


Figure 8-11 Geometric Model for an Interferometric SAR System

Where

- A1 = antenna 1
- A2 = antenna 2
- B_i = baseline
- \vec{R}_1 = vector from antenna 1 to point of interest
- \vec{R}_2 = vector from antenna 2 to point of interest
- Ψ = angle between R_1 and baseline vectors (depression angle)
- Z_{ac} = antenna 1 height

A rigid baseline B_i separates two antennas, A1 and A2. This separation causes the two antennas to illuminate the scene at slightly different depression angles relative to the baseline. Here, ψ is the nominal depression angle from A1 to the scatterer relative to the baseline. The model assumes that the platform travels at constant velocity in the X direction while the baseline remains parallel to the Y axis at a constant height Z_{ac} above the XY plane.

The electromagnetic wave Equation 8-2 describes the signal data collected by each antenna. The two sets of signal data differ primarily because of the small differences in the data collection geometry. Complex images are generated from the signal data received by each antenna.

As stated earlier, the phase of an image pixel represents the phase of multiple scatters in the same resolution cell and consists of both deterministic and unknown random components. A data collection for SAR interferometry adheres to special conditions to ensure that the random component of the phase is nearly identical in the two images. The deterministic phase in a single image is due to the two-way propagation path between the associated antenna and the target.

From our previously derived equation for an electromagnetic wave, and assuming the standard SAR configuration in which the perpendicular distance from the SAR to the target does not change, we can write the complex quantities representing a corresponding pair of image pixels, P_1 and P_2 , from image 1 and image 2 as Equation 8-3 and Equation 8-4.

Equation 8-3

$$P_1 = a_1 \cdot e^{j(\theta_1 + \Phi_1)}$$

and

Equation 8-4

$$P_2 = a_2 \cdot e^{j(\theta_2 + \Phi_2)}$$

The quantities a_1 and a_2 represent the magnitudes of each image pixel. Generally, these magnitudes are approximately equal. The quantities θ_1 and θ_2 are the random components of pixel phase. They represent the vector summations of returns from all unresolved scatterers within the resolution cell and include contributions from receiver noise. With proper system design and collection geometry, they are nearly equal. The quantities Φ_1 and Φ_2 are the deterministic contribution to the phase of the image pixel. The desired function of the interferometer is to provide a measure of the phase difference, $\Phi_1 - \Phi_2$.

Next, we must relate the phase value to the distance vector from each antenna to the point of interest. This is done by recognizing that phase and the wavelength of the electromagnetic wave represent distance in number of wavelengths. Equation 8-5 relates phase to distance and wavelength.

Equation 8-5

$$\Phi_i = \frac{4\pi R_i}{\lambda}$$

Multiplication of one image and the complex conjugate of the second image on a pixel-by-pixel basis yields the phase difference between corresponding pixels in the two images. This complex product produces the interferogram I with

Equation 8-6

$$I = P_1 \cdot P_2'$$

Where ' denotes the complex conjugate operation. With θ_1 and θ_2 nearly equal and a_1 and a_2 nearly equal, the two images differ primarily in how the slight difference in collection depression angles affects Φ_1 and Φ_2 . Ideally then, each pixel in the interferogram has the form

Equation 8-7

$$I = a^2 \cdot e^{-j\left(\frac{4\pi}{\lambda}(R_1 - R_2)\right)} = a^2 \cdot e^{j\phi_{12}}$$

using $a_1 = a_2 = a$. The amplitude a^2 of the interferogram corresponds to image intensity. The phase ϕ_{12} of the interferogram becomes

Equation 8-8

$$\phi_{12} = \frac{4\pi(R_2 - R_1)}{\lambda}$$

which is the quantity used to derive the depression angle to the point of interest relative to the baseline and, eventually, information about the scatterer height relative to the XY plane. Using the following approximation allows us to arrive at an equation relating the interferogram phase to the nominal depression angle.

Equation 8-9

$$R_2 - R_1 \approx B_i \cos(\psi)$$

Equation 8-10

$$\phi_{12} \approx \frac{4\pi B_i \cos(\psi)}{\lambda}$$

In Equation 8-9 and Equation 8-10, ψ is the nominal depression angle from the center of the baseline to the scatterer relative to the baseline. No phase difference indicates that $\psi = 90$ degrees and the scatterer is in the plane through the center of and orthogonal to the baseline. The interferometric phase involves many radians of phase for scatterers at other depression angles since the range difference $R_2 - R_1$ is many wavelengths. In practice, however, an interferometric system does not measure the total pixel phase difference. Rather, it measures only the phase difference that remains after subtracting all full 2π intervals present (module- 2π).

To estimate the actual depression angle to a particular scatterer, the interferometer must measure the total pixel phase difference of many cycles. This information is available, for instance, by unwrapping the raw interferometric phase measurements beginning at a known scene location. Phase unwrapping is discussed in further detail in "Phase Unwrapping" on page 338.

Because of the ambiguity imposed by the wrapped phase problem, it is necessary to seek the relative depression angle and relative height among scatterers within a scene rather than their absolute depression angle and height. The differential of Equation 8-10 with respect to ψ provides this relative measure. This differential is

Equation 8-11

$$\Delta\phi_T = -\left(\frac{4\pi B_i}{\lambda}\right)\sin(\psi)\Delta\psi$$

or

Equation 8-12

$$\Delta\psi = -\left[\frac{\lambda}{4\pi B_i \sin(\psi)}\right]\Delta\phi_{12}$$

This result indicates that two pixels in the interferogram that differ in phase by ϕ_{12} represent scatterers differing in depression angle by $\Delta\psi$. Figure 8-12 shows the differential collection geometry.

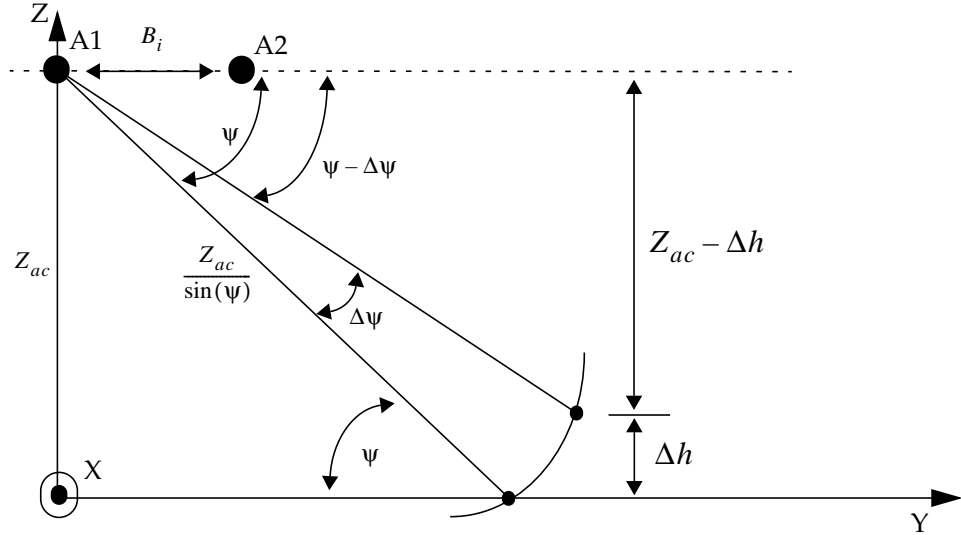


Figure 8-12 Differential Collection Geometry

From this geometry, a change $\Delta\psi$ in depression angle is related to a change Δh in height (at the same range from mid-baseline) by Equation 8-13.

Equation 8-13

$$Z_{ac} - \Delta h = \frac{Z_{ac} \sin(\psi - \Delta\psi)}{\sin(\psi)}$$

Using a useful small-angle approximation to Equation 8-13 and substituting Equation 8-12 into Equation 8-13 provides the result Equation 8-14 for Δh .

Equation 8-14

$$\Delta h \approx Z_{ac} \cot(\psi) \Delta\psi$$

$$\Delta h = - \left[\frac{\lambda Z_{ac} \cot(\psi)}{4\pi B_i \sin(\psi)} \right] \Delta\phi_{12}$$

Note that, because we are calculating differential height, we need at least one known height value in order to calculate absolute height. This translates into a need for at least one GCP in order to calculate absolute heights from the IMAGINE IFSAR DEM process.

In this section, we have derived the mathematical model needed to calculate height from interferometric phase information. In order to put this model into practice, there are several important processes that must be performed. These processes are image registration, phase noise reduction, phase flattening, and phase unwrapping. These processes are discussed in the following sections.

Image Registration

In the discussion of the interferometric model of the last section, we assumed that the pixels had been identified in each image that contained the phase information for the scatterer of interest. Aligning the images from the two antennas is the purpose of the image registration step. For interferometric systems that employ two antennas attached by a fixed boom and collect data simultaneously, this registration is simple and deterministic. Given the collection geometry, the registration can be calculated without referring to the data. For repeat pass systems, the registration is not quite so simple. Since the collection geometry cannot be precisely known, we must use the data to help us achieve image registration.

The registration process for repeat pass interferometric systems is generally broken into two steps: pixel and sub-pixel registration. Pixel registration involves using the magnitude (visible) part of each image to remove the image misregistration down to around a pixel. This means that, after pixel registration, the two images are registered to within one or two pixels of each other in both the range and azimuth directions.

Pixel registration is best accomplished using a standard window correlator to compare the magnitudes of the two images over a specified window. You usually specify a starting point in the two images, a window size, and a search range for the correlator to search over. The process identifies the pixel offset that produces the highest match between the two images, and therefore the best interferogram. One offset is enough to pixel register the two images.

Pixel registration, in general, produces a reasonable interferogram, but not the best possible. This is because of the nature of the phase function for each of the images. In order to form an image from the original signal data collected for each image, it is required that the phase functions in range and azimuth be Nyquist sampled.

Nyquist sampling simply means that the original continuous function can be reconstructed from the sampled data. This means that, while the magnitude resolution is limited to the pixel sizes (often less than that), the phase function can be reconstructed to much higher resolutions. Because it is the phase functions that ultimately provide the height information, it is important to register them as closely as possible. This fine registration of the phase functions is the goal of the sub-pixel registration step.

Sub-pixel registration is achieved by starting at the pixel registration offset and searching over upsampled versions of the phase functions for the best possible interferogram. When this best interferogram is found, the sub-pixel offset has been identified. In order to accomplish this, we must construct higher resolution phase functions from the data. In general this is done using the relation from signal processing theory shown in Equation 8-15.

Equation 8-15

$$i(r + \Delta r, a + \Delta a) = \zeta^{-1}[I(u, v) \cdot e^{-j(u\Delta r + v\Delta a)}]$$

Where

- r = range independent variable
- a = azimuth independent variable
- $i(r, a)$ = interferogram in spacial domain
- $I(u, v)$ = interferogram in frequency domain
- Δr = sub-pixel range offset (i.e., 0.25)
- Δa = sub-pixel azimuth offset (i.e., 0.75)
- ζ^{-1} = inverse Fourier transform

Applying this relation directly requires two-dimensional (2D) Fourier transforms and inverse Fourier transforms for each window tested. This is impractical given the computing requirements of Fourier transforms. Fortunately, we can achieve the upsampled phase functions we need using 2D sinc interpolation, which involves convolving a 2D sinc function of a given size over our search region. Equation 8-16 defines the sinc function for one dimension.

Equation 8-16

$$\frac{\sin(n\pi)}{n\pi}$$

Using sinc interpolation is a fast and efficient method of reconstructing parts of the phase functions which are at sub-pixel locations.

In general, one sub-pixel offset is not enough to sub-pixel register two SAR images over the entire collection. Unlike the pixel registration, sub-pixel registration is dependent on the pixel location, especially the range location. For this reason, it is important to generate a sub-pixel offset function that varies with range position. Two sub-pixel offsets, one at the near range and one at the far range, are enough to generate this function. This sub-pixel register function provides the weights for the sinc interpolator needed to register one image to the other during the formation of the interferogram.

Phase Noise Reduction

We mentioned in "The Interferometric Model" on page 329 that it is necessary to unwrap the phase of the interferogram before it can be used to calculate heights. From a practical and implementational point of view, the phase unwrapping step is the most difficult. We discuss phase unwrapping more in "Phase Unwrapping" on page 338.

Before unwrapping, we can do a few things to the data that make the phase unwrapping easier. The first of these is to reduce the noise in the interferometric phase function. Phase noise is introduced by radar system noise, image misregistration, and speckle effects caused by the complex nature of the imagery. Reducing this noise is done by applying a coherent average filter of a given window size over the entire interferogram. This filter is similar to the more familiar averaging filter, except that it operates on the complex function instead of just the magnitudes. The form of this filter is given in Equation 8-17.

Equation 8-17

$$\hat{i}(r, a) = \frac{\sum_{i=0}^N \sum_{j=0}^M RE[i(r+i, a+j)] + jImg[(r+i, a+j)]}{M+N}$$

Figure 8-13 shows an interferometric phase image without filtering; Figure 8-14 shows the same phase image with filtering.

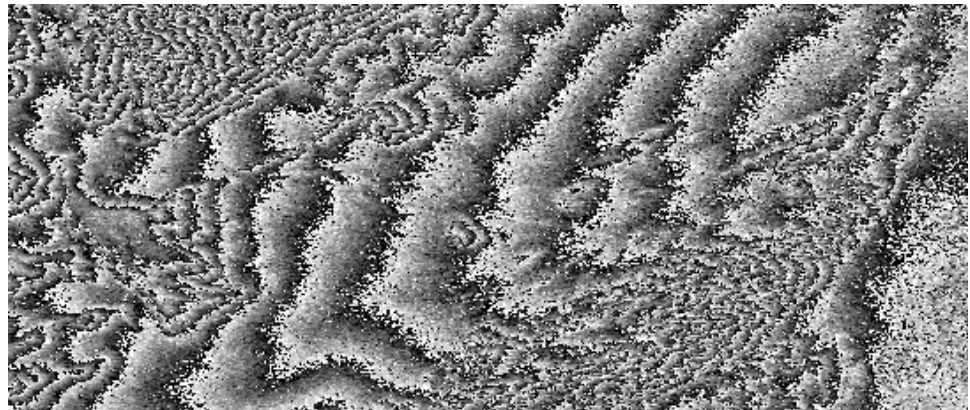


Figure 8-13 Interferometric Phase Image without Filtering

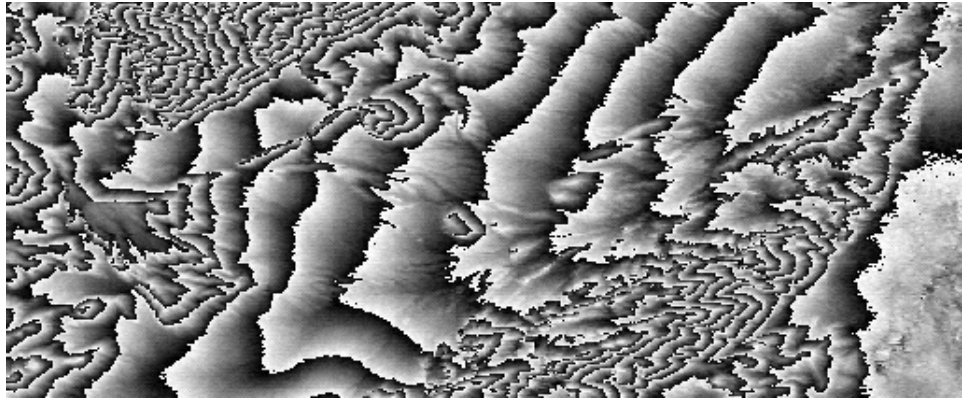


Figure 8-14 Interferometric Phase Image with Filtering

The sharp ridges that look like contour lines in Figure 8-13 and Figure 8-14 show where the phase functions wrap. The goal of the phase unwrapping step is to make this one continuous function. This is discussed in greater detail in "Phase Unwrapping" on page 338. Notice how the filtered image of Figure 8-14 is much cleaner than that of Figure 8-13. This filtering makes the phase unwrapping much easier.

Phase Flattening

The phase function of Figure 8-14 is fairly well behaved and is ready to be unwrapped. There are relatively few wrap lines and they are distinct. Notice in the areas where the elevation is changing more rapidly (mountain regions) the frequency of the wrapping increases. In general, the higher the wrapping frequency, the more difficult the area is to unwrap. Once the wrapping frequency exceeds the spacial sampling of the phase image, information is lost. An important technique in reducing this wrapping frequency is phase flattening.

Phase flattening involves removing high frequency phase wrapping caused by the collection geometry. This high frequency wrapping is mainly in the range direction, and is because of the range separation of the antennas during the collection. Recall that it is this range separation that gives the phase difference and therefore the height information. The phase function of Figure 8-14 has already had phase flattening applied to it. Figure 8-15 shows this same phase function without phase flattening applied.

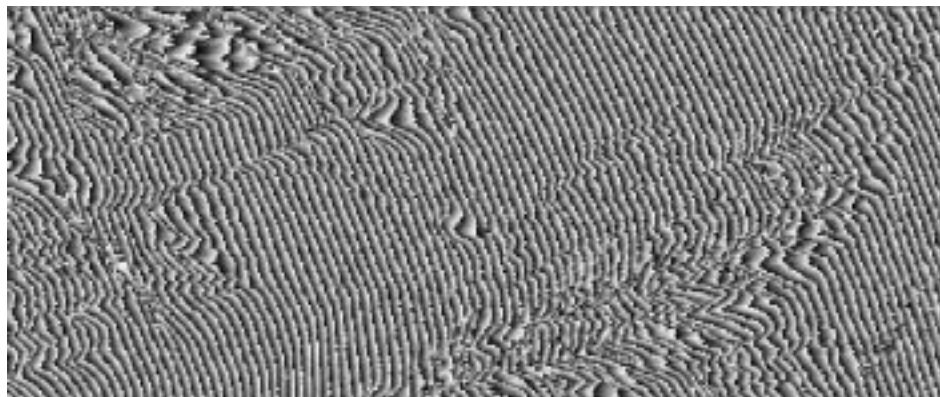


Figure 8-15 Interferometric Phase Image without Phase Flattening

Phase flattening is achieved by removing the phase function that would result if the imaging area was flat from the actual phase function recorded in the interferogram. It is possible, using the equations derived in "The Interferometric Model" on page 329, to calculate this flat Earth phase function and subtract it from the data phase function.

It should be obvious that the phase function in Figure 8-14 is easier to unwrap than the phase function of Figure 8-15.

Phase Unwrapping

We stated in "The Interferometric Model" on page 329 that we must unwrap the interferometric phase before we can use it to calculate height values. In "Phase Noise Reduction" on page 336 and "Phase Flattening" on page 337, we develop methods of making the phase unwrapping job easier. This section further defines the phase unwrapping problem and describes how to solve it.

As an electromagnetic wave travels through space, it cycles through its maximum and minimum phase values many times as shown in Figure 8-16.

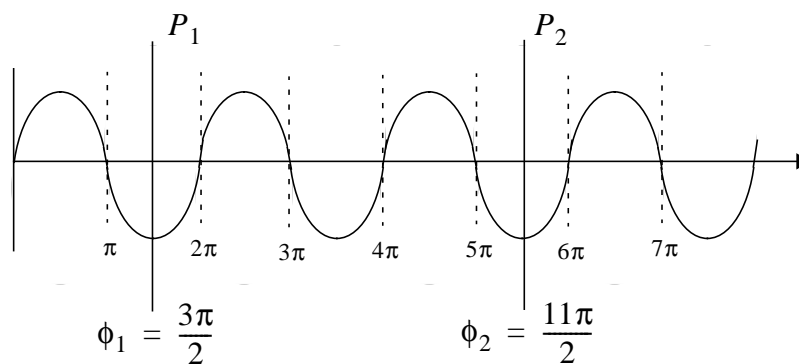


Figure 8-16 Electromagnetic Wave Traveling through Space

The phase difference between points P_1 and P_2 is given by Equation 8-18.

Equation 8-18

$$\phi_2 - \phi_1 = \frac{11\pi}{2} - \frac{3\pi}{2} = 4\pi$$

Recall from Equation 8-8 that finding the phase difference at two points is the key to extracting height from interferometric phase. Unfortunately, an interferometric system does not measure the total pixel phase difference. Rather, it measures only the phase difference that remains after subtracting all full 2π intervals present (module- 2π). This results in the following value for the phase difference of Equation 8-18.

Equation 8-19

$$\phi_2 - \phi_1 = (\text{mod}2\pi)\left(\frac{11\pi}{2}\right) - (\text{mod}2\pi)\left(\frac{3\pi}{2}\right) = \frac{3\pi}{2} - \frac{3\pi}{2} = 0$$

Figure 8-17 further illustrates the difference between a one-dimensional continuous and wrapped phase function. Notice that when the phase value of the continuous function reaches 2π , the wrapped phase function returns to 0 and continues from there. The job of the phase unwrapping is to take a wrapped phase function and reconstruct the continuous function from it.

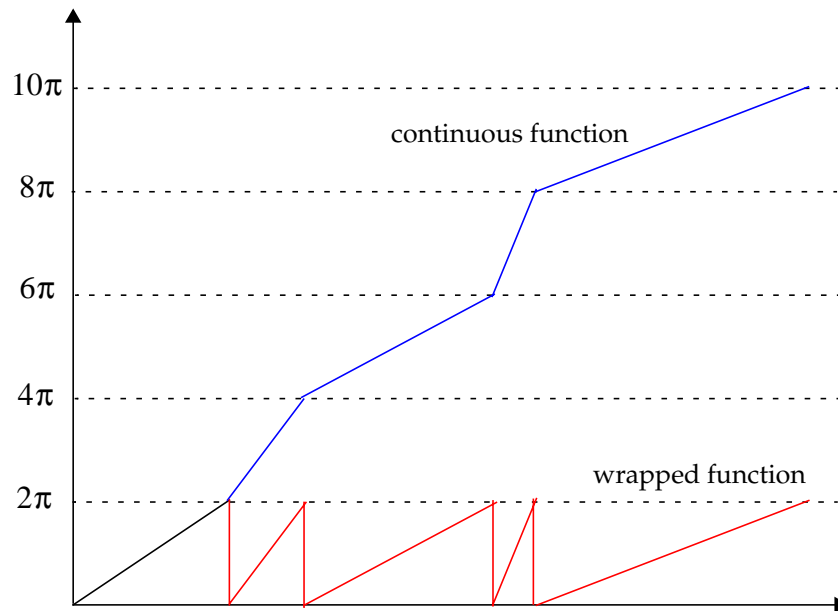


Figure 8-17 One-dimensional Continuous vs. Wrapped Phase Function

There has been much research and many different methods derived for unwrapping the 2D phase function of an interferometric SAR phase image. A detailed discussion of all or any one of these methods is beyond the scope of this chapter. The most successful approaches employ algorithms which unwrap the easy or good areas first and then move on to more difficult areas. Good areas are regions in which the phase function is relatively flat and the correlation is high. This prevents errors in the tough areas from corrupting good regions. Figure 8-18 shows a sequence of unwrapped phase images for the phase function of Figure 8-14.

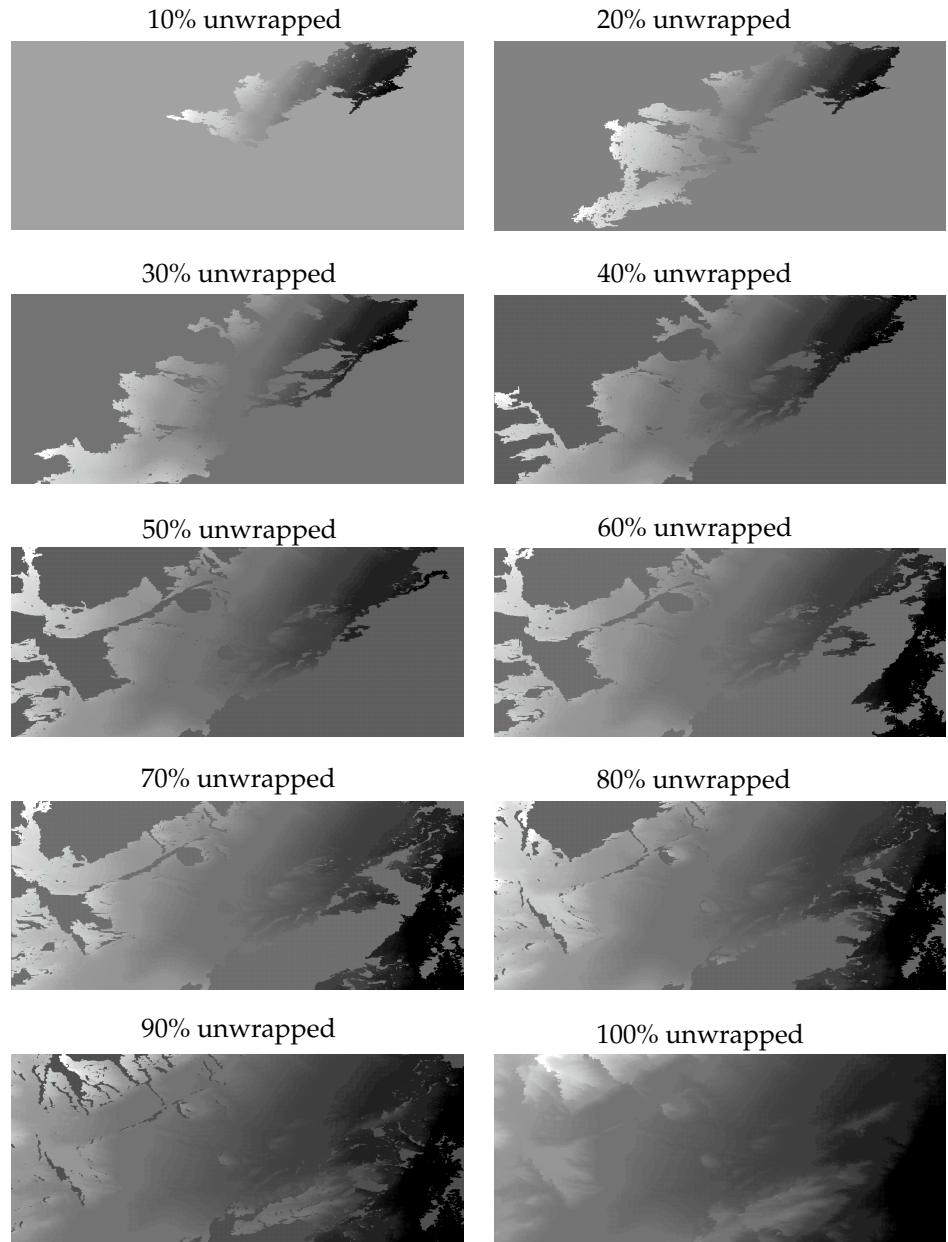
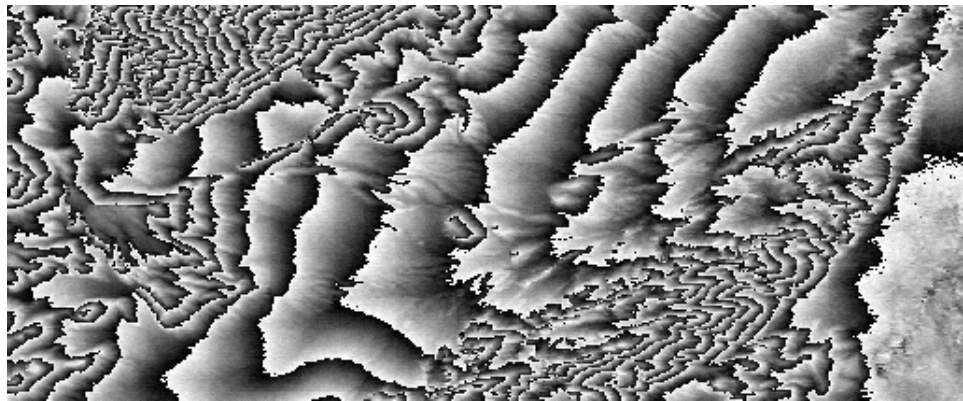
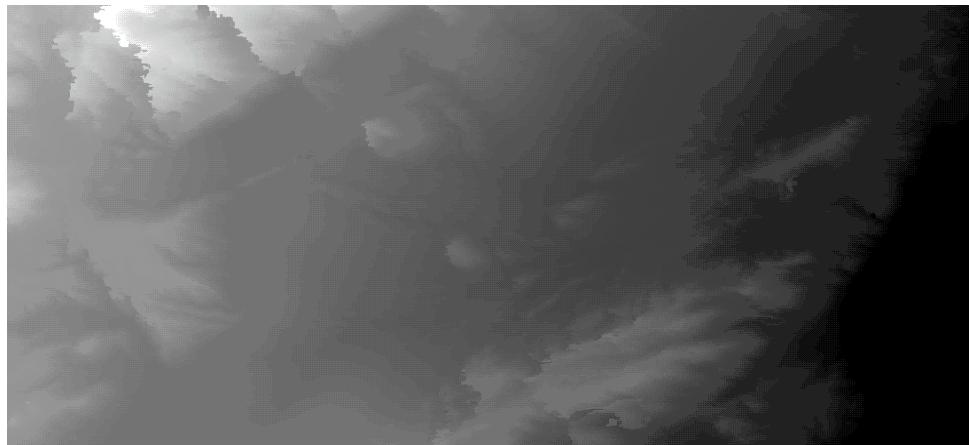


Figure 8-18 Sequence of Unwrapped Phase Images

Figure 8-19 shows the wrapped phase compared to the unwrapped phase image.



Wrapped phase image



Unwrapped phase image

Figure 8-19 Wrapped vs. Unwrapped Phase Images

The unwrapped phase values can now be combined with the collection position information to calculate height values for each pixel in the interferogram.

Radar Concepts

Conclusions

SAR interferometry uses the unique properties of SAR images to extract height information from SAR interferometric image pairs. Given a good image pair and good information about the collection geometry, IMAGINE IFSAR DEM can produce very high quality results. The best IMAGINE IFSAR DEM results are acquired with dual antenna systems that collect both images at once. It is also possible to do IFSAR processing on repeat pass systems. These systems have the advantage of only requiring one antenna, and therefore are cheaper to build. However, the quality of repeat pass IFSAR is very sensitive to the collection conditions because of the fact that the images were not collected at the same time. Weather and terrain changes that occur between the collection of the two images can greatly degrade the coherence of the image pair. This reduction in coherence makes each part of the IMAGINE IFSAR DEM process more difficult.

CHAPTER 9

Rectification

Introduction

Raw, remotely sensed image data gathered by a satellite or aircraft are representations of the irregular surface of the Earth. Even images of seemingly flat areas are distorted by both the curvature of the Earth and the sensor being used. This chapter covers the processes of geometrically correcting an image so that it can be represented on a planar surface, conform to other images, and have the integrity of a map.

A map projection system is any system designed to represent the surface of a sphere or spheroid (such as the Earth) on a plane. There are a number of different map projection methods. Since flattening a sphere to a plane causes distortions to the surface, each map projection system compromises accuracy between certain properties, such as conservation of distance, angle, or area. For example, in equal area map projections, a circle of a specified diameter drawn at any location on the map represents the same total area. This is useful for comparing land use area, density, and many other applications. However, to maintain equal area, the shapes, angles, and scale in parts of the map may be distorted (Jensen 1996).

There are a number of map coordinate systems for determining location on an image. These coordinate systems conform to a grid, and are expressed as X,Y (column, row) pairs of numbers. Each map projection system is associated with a map coordinate system.

Rectification is the process of transforming the data from one grid system into another grid system using a geometric transformation. While polynomial transformation and triangle-based methods are described in this chapter, discussion about various rectification techniques can be found in Yang (1997). Since the pixels of the new grid may not align with the pixels of the original grid, the pixels must be resampled. Resampling is the process of extrapolating data values for the pixels on the new grid from the values of the source pixels.

Registration

In many cases, images of one area that are collected from different sources must be used together. To be able to compare separate images pixel by pixel, the pixel grids of each image must conform to the other images in the data base. The tools for rectifying image data are used to transform disparate images to the same coordinate system. Registration is the process of making an image conform to another image. A map coordinate system is not necessarily involved. For example, if image A is not rectified and it is being used with image B, then image B must be registered to image A so that they conform to each other. In this example, image A is not rectified to a particular map projection, so there is no need to rectify image B to a map projection.

Georeferencing

Georeferencing refers to the process of assigning map coordinates to image data. The image data may already be projected onto the desired plane, but not yet referenced to the proper coordinate system. Rectification, by definition, involves georeferencing, since all map projection systems are associated with map coordinates. Image-to-image registration involves georeferencing only if the reference image is already georeferenced. Georeferencing, by itself, involves changing only the map coordinate information in the image file. The grid of the image does not change.

Geocoded data are images that have been rectified to a particular map projection and pixel size, and usually have had radiometric corrections applied. It is possible to purchase image data that is already geocoded. Geocoded data should be rectified only if they must conform to a different projection system or be registered to other rectified data.

Latitude/Longitude

Lat/Lon is a spherical coordinate system that is not associated with a map projection. Lat/Lon expresses locations in the terms of a spheroid, not a plane. Therefore, an image is not usually rectified to Lat/Lon, although it is possible to convert images to Lat/Lon, and some tips for doing so are included in this chapter.



You can view map projection information for a particular file using the Image Information utility. Image Information allows you to modify map information that is incorrect. However, you cannot rectify data using Image Information. You must use the Rectification tools described in this chapter.



The properties of map projections and of particular map projection systems are discussed in "CHAPTER 12: Cartography" and "Appendix B: Map Projections."

Orthorectification

Orthorectification is a form of rectification that corrects for terrain displacement and can be used if there is a DEM of the study area. It is based on collinearity equations, which can be derived by using 3D GCPs. In relatively flat areas, orthorectification is not necessary, but in mountainous areas (or on aerial photographs of buildings), where a high degree of accuracy is required, orthorectification is recommended.



See "CHAPTER 7: Photogrammetric Concepts" for more information on orthorectification.

When to Rectify

Rectification is necessary in cases where the pixel grid of the image must be changed to fit a map projection system or a reference image. There are several reasons for rectifying image data:

- comparing pixels scene to scene in applications, such as change detection or thermal inertia mapping (day and night comparison)
- developing GIS data bases for GIS modeling

- identifying training samples according to map coordinates prior to classification
- creating accurate scaled photomaps
- overlaying an image with vector data, such as ArcInfo
- comparing images that are originally at different scales
- extracting accurate distance and area measurements
- mosaicking images
- performing any other analyses requiring precise geographic locations

Before rectifying the data, you must determine the appropriate coordinate system for the data base. To select the optimum map projection and coordinate system, the primary use for the data base must be considered.

If you are doing a government project, the projection may be predetermined. A commonly used projection in the United States government is State Plane. Use an equal area projection for thematic or distribution maps and conformal or equal area projections for presentation maps. Before selecting a map projection, consider the following:

- How large or small an area is mapped? Different projections are intended for different size areas.
- Where on the globe is the study area? Polar regions and equatorial regions require different projections for maximum accuracy.
- What is the extent of the study area? Circular, north-south, east-west, and oblique areas may all require different projection systems (ESRI 1992).

When to Georeference Only

Rectification is not necessary if there is no distortion in the image. For example, if an image file is produced by scanning or digitizing a paper map that is in the desired projection system, then that image is already planar and does not require rectification unless there is some skew or rotation of the image. Scanning and digitizing produce images that are planar, but do not contain any map coordinate information. These images need only to be georeferenced, which is a much simpler process than rectification. In many cases, the image header can simply be updated with new map coordinate information. This involves redefining:

- the map coordinate of the upper left corner of the image
- the cell size (the area represented by each pixel)

This information is usually the same for each layer of an image file, although it could be different. For example, the cell size of band 6 of Landsat TM data is different than the cell size of the other bands.



Use the Image Information utility to modify image file header information that is incorrect.

Rectification

Disadvantages of Rectification

During rectification, the data file values of rectified pixels must be resampled to fit into a new grid of pixel rows and columns. Although some of the algorithms for calculating these values are highly reliable, some spectral integrity of the data can be lost during rectification. If map coordinates or map units are not needed in the application, then it may be wiser not to rectify the image. An unrectified image is more spectrally correct than a rectified image.

Classification

Some analysts recommend classification before rectification, since the classification is then based on the original data values. Another benefit is that a thematic file has only one band to rectify instead of the multiple bands of a continuous file. On the other hand, it may be beneficial to rectify the data first, especially when using GPS data for the GCPs. Since these data are very accurate, the classification may be more accurate if the new coordinates help to locate better training samples.

Thematic Files

Nearest neighbor is the only appropriate resampling method for thematic files, which may be a drawback in some applications. The available resampling methods are discussed in detail later in this chapter.

Rectification Steps

NOTE: Registration and rectification involve similar sets of procedures. Throughout this documentation, many references to rectification also apply to image-to-image registration.

Usually, rectification is the conversion of data file coordinates to some other grid and coordinate system, called a reference system. Rectifying or registering image data on disk involves the following general steps, regardless of the application:

1. Locate GCPs.
2. Compute and test a transformation.
3. Create an output image file with the new coordinate information in the header. The pixels must be resampled to conform to the new grid.

Images can be rectified on the display (in a Viewer) or on the disk. Display rectification is temporary, but disk rectification is permanent, because a new file is created. Disk rectification involves:

- rearranging the pixels of the image onto a new grid, which conforms to a plane in the new map projection and coordinate system
- inserting new information to the header of the file, such as the upper left corner map coordinates and the area represented by each pixel

Ground Control Points

GCPs are specific pixels in an image for which the output map coordinates (or other output coordinates) are known. GCPs consist of two X,Y pairs of coordinates:

- source coordinates—usually data file coordinates in the image being rectified
- reference coordinates—the coordinates of the map or reference image to which the source image is being registered

The term map coordinates is sometimes used loosely to apply to reference coordinates and rectified coordinates. These coordinates are not limited to map coordinates. For example, in image-to-image registration, map coordinates are not necessary.

GCPs in ERDAS IMAGINE

Any ERDAS IMAGINE image can have one GCP set associated with it. The GCP set is stored in the image file along with the raster layers. If a GCP set exists for the top file that is displayed in the Viewer, then those GCPs can be displayed when the GCP Tool is opened.

In the CellArray of GCP data that displays in the GCP Tool, one column shows the point ID of each GCP. The point ID is a name given to GCPs in separate files that represent the same geographic location. Such GCPs are called corresponding GCPs.

A default point ID string is provided (such as GCP #1), but you can enter your own unique ID strings to set up corresponding GCPs as needed. Even though only one set of GCPs is associated with an image file, one GCP set can include GCPs for a number of rectifications by changing the point IDs for different groups of corresponding GCPs.

Entering GCPs

Accurate GCPs are essential for an accurate rectification. From the GCPs, the rectified coordinates for all other points in the image are extrapolated. Select many GCPs throughout the scene. The more dispersed the GCPs are, the more reliable the rectification is. GCPs for large-scale imagery might include the intersection of two roads, airport runways, utility corridors, towers, or buildings. For small-scale imagery, larger features such as urban areas or geologic features may be used. Landmarks that can vary (e.g., the edges of lakes or other water bodies, vegetation, etc.) should not be used.

The source and reference coordinates of the GCPs can be entered in the following ways:

- They may be known *a priori*, and entered at the keyboard.
- Use the mouse to select a pixel from an image in the Viewer. With both the source and destination Viewers open, enter source coordinates and reference coordinates for image-to-image registration.
- Use a digitizing tablet to register an image to a hardcopy map.



Information on the use and setup of a digitizing tablet is discussed in "CHAPTER 2: Vector Layers."

Digitizing Tablet Option

If GCPs are digitized from a hardcopy map and a digitizing tablet, accurate base maps must be collected. You should try to match the resolution of the imagery with the scale and projection of the source map. For example, 1:24,000 scale USGS quadrangles make good base maps for rectifying Landsat TM and SPOT imagery. Avoid using maps over 1:250,000, if possible. Coarser maps (i.e., 1:250,000) are more suitable for imagery of lower resolution (i.e., AVHRR) and finer base maps (i.e., 1:24,000) are more suitable for imagery of finer resolution (i.e., Landsat and SPOT).

Mouse Option

When entering GCPs with the mouse, you should try to match coarser resolution imagery to finer resolution imagery (i.e., Landsat TM to SPOT), and avoid stretching resolution spans greater than a cubic convolution radius (a 4×4 area). In other words, you should not try to match Landsat MSS to SPOT or Landsat TM to an aerial photograph.

How GCPs are Stored

GCPs entered with the mouse are stored in the image file, and those entered at the keyboard or digitized using a digitizing tablet are stored in a separate file with the extension .gcc.

Automated GCP prediction enables you to pick a GCP in either coordinate system and automatically locate that point in the other coordinate system based on the current transformation parameters.

Automated GCP matching is a step beyond GCP prediction. For image-to-image rectification, a GCP selected in one image is precisely matched to its counterpart in the other image using the spectral characteristics of the data and the geometric transformation. GCP matching enables you to fine tune a rectification for highly accurate results.

Both of these methods require an existing transformation which consists of a set of coefficients used to convert the coordinates from one system to another.

GCP Prediction

GCP prediction is a useful technique to help determine if enough GCPs have been gathered. After selecting several GCPs, select a point in either the source or the destination image, then use GCP prediction to locate the corresponding GCP on the other image (map). This point is determined based on the current transformation derived from existing GCPs. Examine the automatically generated point and see how accurate it is. If it is within an acceptable range of accuracy, then there may be enough GCPs to perform an accurate rectification (depending upon how evenly dispersed the GCPs are). If the automatically generated point is not accurate, then more GCPs should be gathered before rectifying the image.

GCP prediction can also be used when applying an existing transformation to another image in a data set. This saves time in selecting another set of GCPs by hand. Once the GCPs are automatically selected, those that do not meet an acceptable level of error can be edited.

GCP Matching

In GCP matching, you can select which layers from the source and destination images to use. Since the matching process is based on the reflectance values, select layers that have similar spectral wavelengths, such as two visible bands or two infrared bands. You can perform histogram matching to ensure that there is no offset between the images. You can also select the radius from the predicted GCP from which the matching operation searches for a spectrally similar pixels. The search window can be any odd size between 5×5 and 21×21 .



Histogram matching is discussed in "CHAPTER 5: Enhancement".

A correlation threshold is used to accept or discard points. The correlation ranges from -1.000 to +1.000. The threshold is an absolute value threshold ranging from 0.000 to 1.000. A value of 0.000 indicates a bad match and a value of 1.000 indicates an exact match. Values above 0.8000 or 0.9000 are recommended. If a match cannot be made because the absolute value of the correlation is less than the threshold, you have the option to discard points.

Polynomial Transformation

Polynomial equations are used to convert source file coordinates to rectified map coordinates. Depending upon the distortion in the imagery, the number of GCPs used, and their locations relative to one another, complex polynomial equations may be required to express the needed transformation. The degree of complexity of the polynomial is expressed as the order of the polynomial. The order is simply the highest exponent used in the polynomial.

The order of transformation is the order of the polynomial used in the transformation. ERDAS IMAGINE allows 1st- through nth-order transformations. Usually, 1st-order or 2nd-order transformations are used.



You can specify the order of the transformation you want to use in the Transform Editor.



A discussion of polynomials and order is included in "Appendix A: Math Topics".

Transformation Matrix

A transformation matrix is computed from the GCPs. The matrix consists of coefficients that are used in polynomial equations to convert the coordinates. The size of the matrix depends upon the order of transformation. The goal in calculating the coefficients of the transformation matrix is to derive the polynomial equations for which there is the least possible amount of error when they are used to transform the reference coordinates of the GCPs into the source coordinates. It is not always possible to derive coefficients that produce no error. For example, in Figure 9-1, GCPs are plotted on a graph and compared to the curve that is expressed by a polynomial.

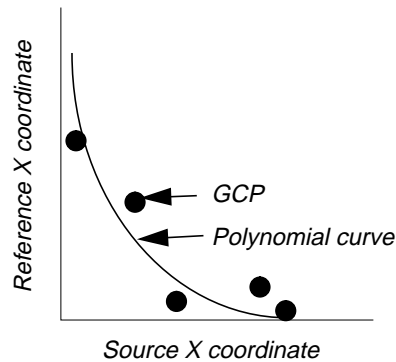


Figure 9-1 Polynomial Curve vs. GCPs

Every GCP influences the coefficients, even if there is not a perfect fit of each GCP to the polynomial that the coefficients represent. The distance between the GCP reference coordinate and the curve is called RMS error, which is discussed later in this chapter. The least squares regression method is used to calculate the transformation matrix from the GCPs. This common method is discussed in statistics textbooks.

Linear Transformations

A 1st-order transformation is a linear transformation. It can change:

- location in X and/or Y
- scale in X and/or Y
- skew in X and/or Y
- rotation

First-order transformations can be used to project raw imagery to a planar map projection, convert a planar map projection to another planar map projection, and when rectifying relatively small image areas. You can perform simple linear transformations to an image displayed in a Viewer or to the transformation matrix itself. Linear transformations may be required before collecting GCPs on the displayed image. You can reorient skewed Landsat TM data, rotate scanned quad sheets according to the angle of declination stated in the legend, and rotate descending data so that north is up.

A 1st-order transformation can also be used for data that are already projected onto a plane. For example, SPOT and Landsat Level 1B data are already transformed to a plane, but may not be rectified to the desired map projection. When doing this type of rectification, it is not advisable to increase the order of transformation if at first a high RMS error occurs. Examine other factors first, such as the GCP source and distribution, and look for systematic errors.

ERDAS IMAGINE provides the following options for 1st-order transformations:

- scale
- offset
- rotate
- reflect

Scale

Scale is the same as the zoom option in the Viewer, except that you can specify different scaling factors for X and Y.



If you are scaling an image in the Viewer, the zoom option undoes any changes to the scale that you do, and vice versa.

Offset

Offset moves the image by a user-specified number of pixels in the X and Y directions. For rotation, you can specify any positive or negative number of degrees for clockwise and counterclockwise rotation. Rotation occurs around the center pixel of the image.

Reflection

Reflection options enable you to perform the following operations:

- left to right reflection
- top to bottom reflection
- left to right and top to bottom reflection (equal to a 180° rotation)



Linear adjustments are available from the Viewer or from the Transform Editor. You can perform linear transformations in the Viewer and then load that transformation to the Transform Editor, or you can perform the linear transformations directly on the transformation matrix.

Figure 9-2 illustrates how the data are changed in linear transformations.

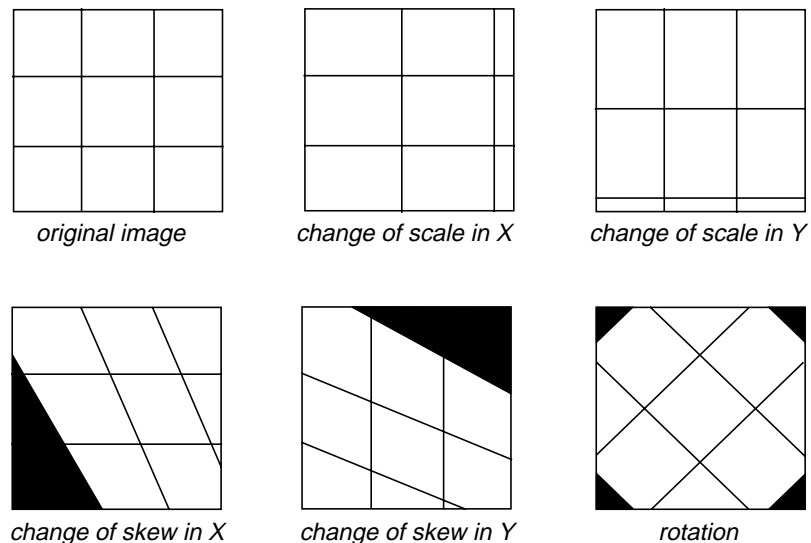


Figure 9-2 Linear Transformations

Rectification

The transformation matrix for a 1st-order transformation consists of six coefficients—three for each coordinate (X and Y).

$$\begin{array}{ccc} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \end{array}$$

Coefficients are used in a 1st-order polynomial as follows:

$$x_o = a_0 + a_1x + a_2y$$

$$y_o = b_0 + b_1x + b_2y$$

Where:

x and y are source coordinates (input)
 x_o and y_o are rectified coordinates (output)
the coefficients of the transformation matrix are as above



The position of the coefficients in the matrix and the assignment of the coefficients in the polynomial is an ERDAS IMAGINE convention. Other representations of a 1st-order transformation matrix may take a different form.

Nonlinear Transformations

Transformations of the 2nd-order or higher are nonlinear transformations. These transformations can correct nonlinear distortions. The process of correcting nonlinear distortions is also known as rubber sheeting. Figure 9-3 illustrates the effects of some nonlinear transformations.

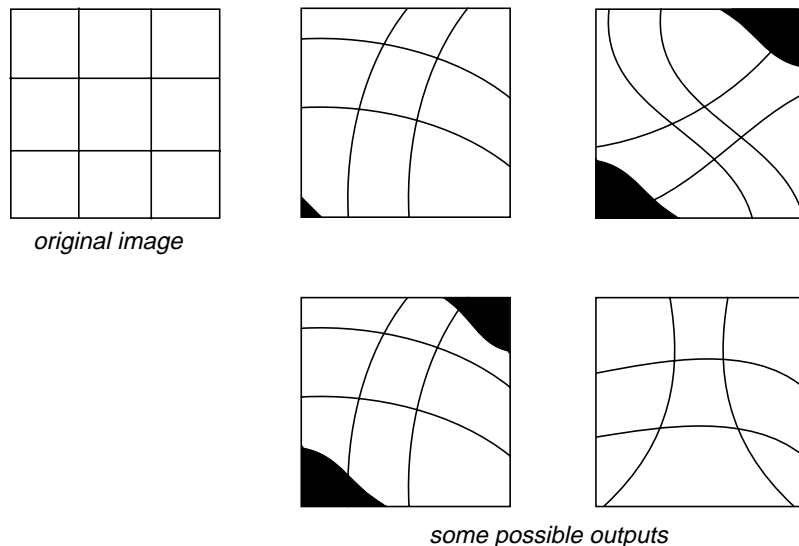


Figure 9-3 Nonlinear Transformations

Second-order transformations can be used to convert Lat/Lon data to a planar projection, for data covering a large area (to account for the Earth's curvature), and with distorted data (for example, due to camera lens distortion). Third-order transformations are used with distorted aerial photographs, on scans of warped maps and with radar imagery. Fourth-order transformations can be used on very distorted aerial photographs.

The transformation matrix for a transformation of order t contains this number of coefficients:

$$2 \sum_{i=1}^{t+1} i$$

It is multiplied by two for the two sets of coefficients—one set for X, one for Y.

An easier way to arrive at the same number is:

$$(t + 1) \times (t + 2)$$

Clearly, the size of the transformation matrix increases with the order of the transformation.

Higher Order Polynomials

The polynomial equations for a t -order transformation take this form:

$$x_o = \left(\begin{array}{c} t \\ \Sigma \\ i = o \end{array} \right) \left(\begin{array}{c} i \\ \Sigma \\ j = o \end{array} \right) \quad a_k \times x^{i-j} \times y^j$$

$$y_o = \left(\begin{array}{c} t \\ \Sigma \\ i = o \end{array} \right) \left(\begin{array}{c} i \\ \Sigma \\ j = o \end{array} \right) \quad b_k \times x^{i-j} \times y^j$$

Rectification

Where:

t is the order of the polynomial

a_k and b_k are coefficients

the subscript k in a_k and b_k is determined by:

$$k = \frac{i \cdot i + j}{2} + j$$

An example of 3rd-order transformation equations for X and Y, using numbers, is:

$$x_o = 5 + 4x - 6y + 10x^2 - 5xy + 1y^2 + 3x^3 + 7x^2y - 11xy^2 + 4y^3$$

$$y_o = 13 + 12x + 4y + 1x^2 - 21xy + 11y^2 - 1x^3 + 2x^2y + 5xy^2 + 12y^3$$

These equations use a total of 20 coefficients, or

$$(3 + 1) \times (3 + 2)$$

Effects of Order

The computation and output of a higher-order polynomial equation are more complex than that of a lower-order polynomial equation. Therefore, higher-order polynomials are used to perform more complicated image rectifications. To understand the effects of different orders of transformation in image rectification, it is helpful to see the output of various orders of polynomials.

The following example uses only one coordinate (X), instead of two (X,Y), which are used in the polynomials for rectification. This enables you to draw two-dimensional graphs that illustrate the way that higher orders of transformation affect the output image.

NOTE: Because only the X coordinate is used in these examples, the number of GCPs used is less than the number required to actually perform the different orders of transformation.

Coefficients like those presented in this example would generally be calculated by the least squares regression method. Suppose GCPs are entered with these X coordinates:

Source X Coordinate (input)	Reference X Coordinate (output)
1d	17
2	9
3	1

These GCPs allow a 1st-order transformation of the X coordinates, which is satisfied by this equation (the coefficients are in parentheses):

$$x_r = (25) + (-8)x_i$$

Where:

x_r = the reference X coordinate

x_i = the source X coordinate

This equation takes on the same format as the equation of a line ($y = mx + b$). In mathematical terms, a 1st-order polynomial is linear. Therefore, a 1st-order transformation is also known as a linear transformation. This equation is graphed in Figure 9-4.

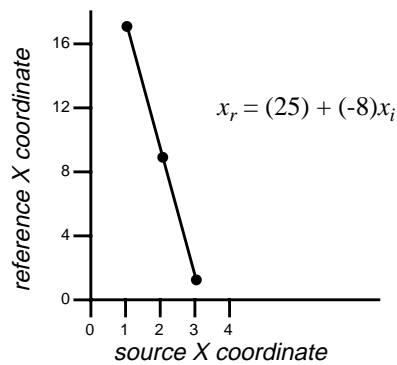


Figure 9-4 Transformation Example—1st-Order

Rectification

However, what if the second GCP were changed as follows?

Source X Coordinate (input)	Reference X Coordinate (output)
1	17
2	7
3	1

These points are plotted against each other in Figure 9-5.

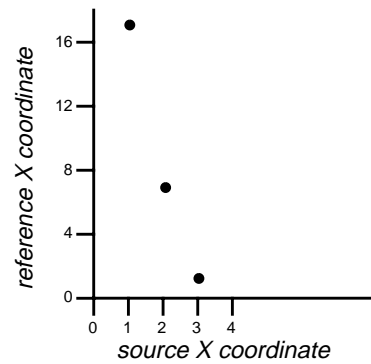


Figure 9-5 Transformation Example—2nd GCP Changed

A line cannot connect these points, which illustrates that they cannot be expressed by a 1st-order polynomial, like the one above. In this case, a 2nd-order polynomial equation expresses these points:

$$x_r = (31) + (-16)x_i + (2)x_i^2$$

Polynomials of the 2nd-order or higher are nonlinear. The graph of this curve is drawn in Figure 9-6.

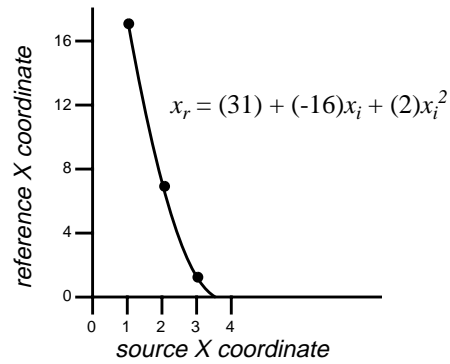


Figure 9-6 Transformation Example—2nd-Order

What if one more GCP were added to the list?

Source X Coordinate (input)	Reference X Coordinate (output)
1	17
2	7
3	1
4	5

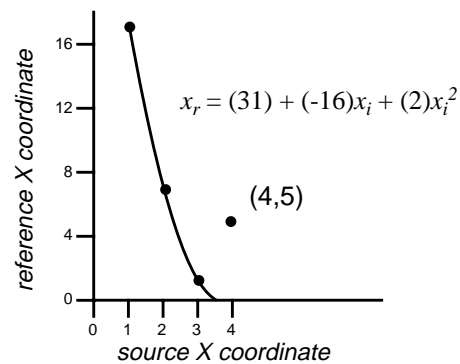


Figure 9-7 Transformation Example—4th GCP Added

As illustrated in Figure 9-7, this fourth GCP does not fit on the curve of the 2nd-order polynomial equation. To ensure that all of the GCPs fit, the order of the transformation could be increased to 3rd-order. The equation and graph in Figure 9-8 could then result.

Rectification

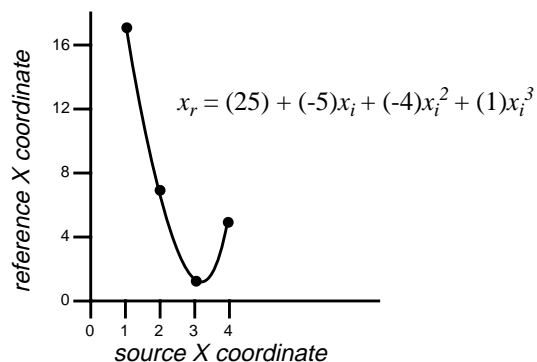


Figure 9-8 Transformation Example—3rd-Order

Figure 9-8 illustrates a 3rd-order transformation. However, this equation may be unnecessarily complex. Performing a coordinate transformation with this equation may cause unwanted distortions in the output image for the sake of a perfect fit for all the GCPs. In this example, a 3rd-order transformation probably would be too high, because the output pixels would be arranged in a different order than the input pixels, in the X direction.

Source X Coordinate (input)	Reference X Coordinate (output)
1	$x_o(1) = 17$
2	$x_o(2) = 7$
3	$x_o(3) = 1$
4	$x_o(4) = 5$

$$x_o(1) > x_o(2) > x_o(4) > x_o(3)$$

$$17 > 7 > 5 > 1$$

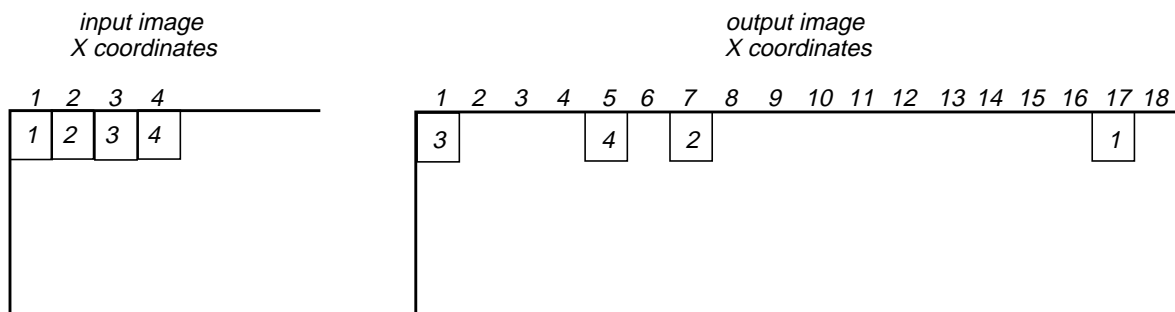


Figure 9-9 Transformation Example—Effect of a 3rd-Order Transformation

In this case, a higher order of transformation would probably not produce the desired results.

Minimum Number of GCPs

Higher orders of transformation can be used to correct more complicated types of distortion. However, to use a higher order of transformation, more GCPs are needed. For instance, three points define a plane. Therefore, to perform a 1st-order transformation, which is expressed by the equation of a plane, at least three GCPs are needed. Similarly, the equation used in a 2nd-order transformation is the equation of a paraboloid. Six points are required to define a paraboloid. Therefore, at least six GCPs are required to perform a 2nd-order transformation. The minimum number of points required to perform a transformation of order t equals:

$$\frac{((t + 1)(t + 2))}{2}$$

Use more than the minimum number of GCPs whenever possible. Although it is possible to get a perfect fit, it is rare, no matter how many GCPs are used.

For 1st- through 10th-order transformations, the minimum number of GCPs required to perform a transformation is listed in the following table:

Table 9-1 Number of GCPs per Order of Transformation

Order of Transformation	Minimum GCPs Required
1	3
2	6
3	10
4	15
5	21
6	28
7	36
8	45
9	55
10	66



For the best rectification results, you should always use more than the minimum number of GCPs, and they should be well-distributed.

Rubber Sheeting

Triangle-Based Finite Element Analysis

The finite element analysis is a powerful tool for solving complicated computation problems which can be approached by small simpler pieces. It has been widely used as a local interpolation technique in geographic applications. For image rectification, the known control points can be triangulated into many triangles. Each triangle has three control points as its vertices. Then, the polynomial transformation can be used to establish mathematical relationships between source and destination systems for each triangle. Because the transformation exactly passes through each control point and is not in a uniform manner, finite element analysis is also called rubber sheeting. It can also be called as the triangle-based rectification because the transformation and resampling for image rectification are performed on a triangle-by-triangle basis.

This triangle-based technique should be used when other rectification methods such as polynomial transformation and photogrammetric modeling cannot produce acceptable results.

Triangulation

To perform the triangle-based rectification, it is necessary to triangulate the control points into a mesh of triangles. Watson (1994) summarily listed four kinds of triangulation, including the arbitrary, optimal, Greedy and Delaunay triangulation. Of the four kinds, the Delaunay triangulation is most widely used and is adopted because of the smaller angle variations of the resulting triangles.

The Delaunay triangulation can be constructed by the empty circumcircle criterion. The circumcircle formed from three points of any triangle does not have any other point inside. The triangles defined this way are the most equiangular possible.

Figure 9-10 shows an example of the triangle network formed by 13 control points.

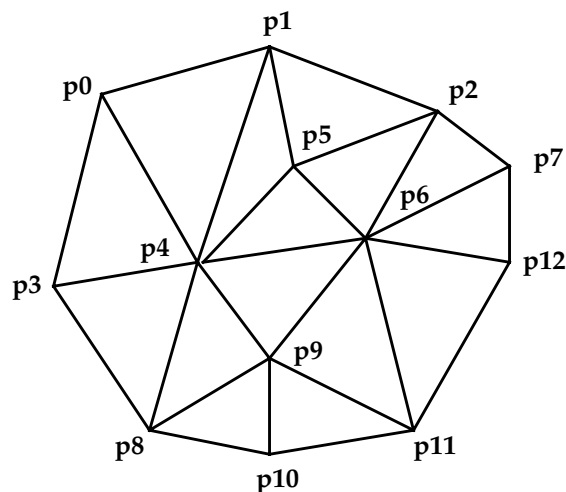


Figure 9-10 Triangle Network

Triangle-based rectification

Once the triangle mesh has been generated and the spatial order of the control points is available, the geometric rectification can be done on a triangle-by-triangle basis. This triangle-based method is appealing because it breaks the entire region into smaller subsets. If the geometric problem of the entire region is very complicated, the geometry of each subset can be much simpler and modeled through simple transformation.

For each triangle, the polynomials can be used as the general transformation form between source and destination systems.

Linear transformation

The easiest and fastest is the linear transformation with the first order polynomials:

$$\begin{cases} x_o = a_0 + a_1x + a_2y \\ y_o = b_0 + b_1x + b_2y \end{cases}$$

There is no need for extra information because there are three known conditions in each triangle and three unknown coefficients for each polynomial.

Nonlinear transformation

Even though the linear transformation is easy and fast, it has one disadvantage. The transitions between triangles are not always smooth. This phenomenon is obvious when shaded relief or contour lines are derived from the DEM which is generated by the linear rubber sheeting. It is caused by incorporating the slope change of the control data at the triangle edges and vertices. In order to distribute the slope change smoothly across triangles, the nonlinear transformation with polynomial order larger than one is used by considering the gradient information.

The fifth order or quintic polynomial transformation is chosen here as the nonlinear rubber sheeting technique in this dissertation. It is a smooth function. The transformation function and its first order partial derivative are continuous. It is not difficult to construct (Akima 1978). The formulation is simply as the following:

$$\begin{cases} x_o = \sum_{i=0}^5 \sum_{j=0}^i a_k \cdot x^{i-j} \cdot y^j \\ y_o = \sum_{i=0}^5 \sum_{j=0}^i b_k \cdot x^{i-j} \cdot y^j \end{cases}$$

which has 21 coefficients for each polynomial to be determined. For solving these unknowns, 21 conditions should be available. For each vertex of the triangle, one point value is given, and two first order and three second order partial derivatives can be easily derived by establishing a second order polynomial using vertices in the neighborhood of the vertex. Then the total 18 conditions are ready to be used. Three more conditions can be obtained by assuming that the normal partial derivative on each edge of the triangle is a cubic polynomial, which means that the sum of the polynomial items beyond the third order in the normal partial derivative has a value zero.

Rectification

Check Point Analysis

It should be emphasized that the independent check point analysis is critical for determining the accuracy of rubber sheeting modeling. For an exact modeling method like rubber sheeting, the ground control points, which are used in the modeling process, do not have much geometric residuals remaining. To evaluate the geometric transformation between source and destination coordinate systems, the accuracy assessment using independent check points is recommended.

RMS Error

RMS error is the distance between the input (source) location of a GCP and the retransformed location for the same GCP. In other words, it is the difference between the desired output coordinate for a GCP and the actual output coordinate for the same point, when the point is transformed with the geometric transformation.

RMS error is calculated with a distance equation:

$$\text{RMS error} = \sqrt{(x_r - x_i)^2 + (y_r - y_i)^2}$$

Where:

x_i and y_i are the input source coordinates

x_r and y_r are the retransformed coordinates

RMS error is expressed as a distance in the source coordinate system. If data file coordinates are the source coordinates, then the RMS error is a distance in pixel widths. For example, an RMS error of 2 means that the reference pixel is 2 pixels away from the retransformed pixel.

Residuals and RMS Error Per GCP

The GCP Tool contains columns for the X and Y residuals. Residuals are the distances between the source and retransformed coordinates in one direction. They are shown for each GCP. The X residual is the distance between the source X coordinate and the retransformed X coordinate. The Y residual is the distance between the source Y coordinate and the retransformed Y coordinate.

If the GCPs are consistently off in either the X or the Y direction, more points should be added in that direction. This is a common problem in off-nadir data.

RMS Error Per GCP

The RMS error of each point is reported to help you evaluate the GCPs. This is calculated with a distance formula:

$$R_i = \sqrt{XR_i^2 + YR_i^2}$$

Where:

R_i = the RMS error for GCP_{*i*}

XR_i = the X residual for GCP_{*i*}

YR_i = the Y residual for GCP_{*i*}

Figure 9-11 illustrates the relationship between the residuals and the RMS error per point.

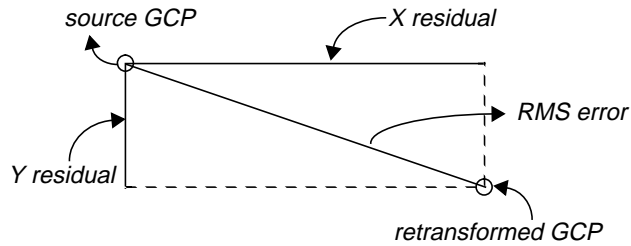


Figure 9-11 Residuals and RMS Error Per Point

Total RMS Error

From the residuals, the following calculations are made to determine the total RMS error, the X RMS error, and the Y RMS error:

$$R_x = \sqrt{\frac{1}{n} \sum_{i=1}^n XR_i^2}$$

$$R_y = \sqrt{\frac{1}{n} \sum_{i=1}^n YR_i^2}$$

$$T = \sqrt{R_x^2 + R_y^2} \quad \text{or} \quad \sqrt{\frac{1}{n} \sum_{i=1}^n XR_i^2 + YR_i^2}$$

Where:

R_x = X RMS error

R_y = Y RMS error

T = total RMS error

n = the number of GCPs

i = GCP number

XR_i = the X residual for GCP_{*i*}

YR_i = the Y residual for GCP_{*i*}

Error Contribution by Point

A normalized value representing each point's RMS error in relation to the total RMS error is also reported. This value is listed in the Contribution column of the GCP Tool.

Rectification

$$E_i = \frac{R_i}{T}$$

Where:

E_i = error contribution of GCP_{*i*}

R_i = the RMS error for GCP_{*i*}

T = total RMS error

Tolerance of RMS Error

In most cases, it is advantageous to tolerate a certain amount of error rather than take a more complex transformation. The amount of RMS error that is tolerated can be thought of as a window around each source coordinate, inside which a retransformed coordinate is considered to be correct (that is, close enough to use). For example, if the RMS error tolerance is 2, then the retransformed pixel can be 2 pixels away from the source pixel and still be considered accurate.

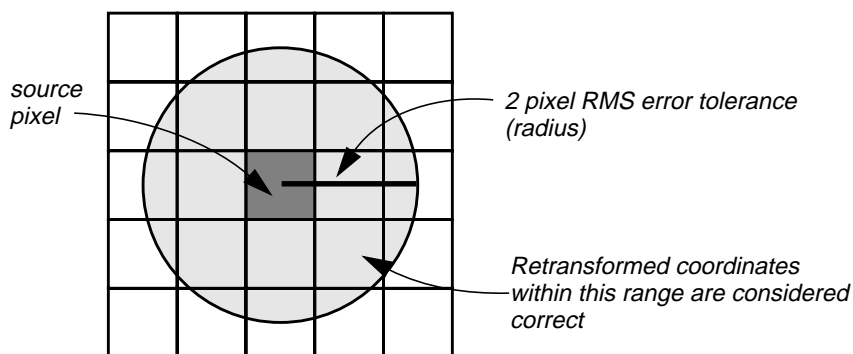


Figure 9-12 RMS Error Tolerance

Acceptable RMS error is determined by the end use of the data base, the type of data being used, and the accuracy of the GCPs and ancillary data being used. For example, GCPs acquired from GPS should have an accuracy of about 10 m, but GCPs from 1:24,000-scale maps should have an accuracy of about 20 m.

It is important to remember that RMS error is reported in pixels. Therefore, if you are rectifying Landsat TM data and want the rectification to be accurate to within 30 meters, the RMS error should not exceed 1.00. Acceptable accuracy depends on the image area and the particular project.

Evaluating RMS Error

To determine the order of polynomial transformation, you can assess the relative distortion in going from image to map or map to map. One should start with a 1st-order transformation unless it is known that it does not work. It is possible to repeatedly compute transformation matrices until an acceptable RMS error is reached.



Most rectifications are either 1st-order or 2nd-order. The danger of using higher order rectifications is that the more complicated the equation for the transformation, the less regular and predictable the results are. To fit all of the GCPs, there may be very high distortion in the image.

After each computation of a transformation and RMS error, there are four options:

- Throw out the GCP with the highest RMS error, assuming that this GCP is the least accurate. Another transformation can then be computed from the remaining GCPs. A closer fit should be possible. However, if this is the only GCP in a particular region of the image, it may cause greater error to remove it.
- Tolerate a higher amount of RMS error.
- Increase the complexity of transformation, creating more complex geometric alterations in the image. A transformation can then be computed that can accommodate the GCPs with less error.
- Select only the points for which you have the most confidence.

Resampling Methods

The next step in the rectification/registration process is to create the output file. Since the grid of pixels in the source image rarely matches the grid for the reference image, the pixels are resampled so that new data file values for the output file can be calculated.

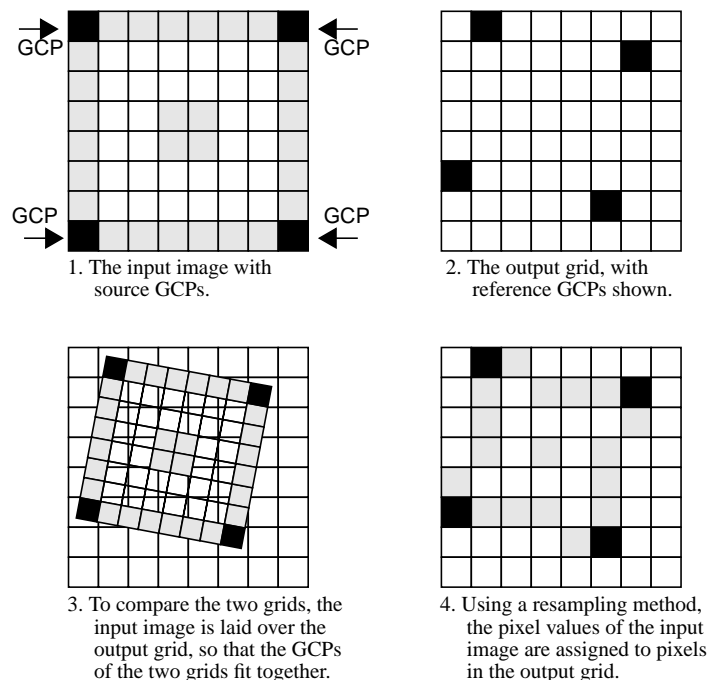


Figure 9-13 Resampling

The following resampling methods are supported in ERDAS IMAGINE:

Rectification

- Nearest neighbor—uses the value of the closest pixel to assign to the output pixel value.
- Bilinear interpolation—uses the data file values of four pixels in a 2×2 window to calculate an output value with a bilinear function.
- Cubic convolution—uses the data file values of sixteen pixels in a 4×4 window to calculate an output value with a cubic function.

In all methods, the number of rows and columns of pixels in the output is calculated from the dimensions of the output map, which is determined by the geometric transformation and the cell size. The output corners (upper left and lower right) of the output file can be specified. The default values are calculated so that the entire source file is resampled to the destination file.

If an image to image rectification is being performed, it may be beneficial to specify the output corners relative to the reference file system, so that the images are coregistered. In this case, the upper left X and upper left Y coordinate are 0,0 and not the defaults.



*If the output units are pixels, then the origin of the image is the upper left corner.
Otherwise, the origin is the lower left corner.*

Rectifying to Lat/Lon

You can specify the nominal cell size if the output coordinate system is Lat/Lon. The output cell size for a geographic projection (i.e., Lat/Lon) is always in angular units of decimal degrees. However, if you want the cell to be a specific size in meters, you can enter meters and calculate the equivalent size in decimal degrees. For example, if you want the output file cell size to be 30×30 meters, then the program would calculate what this size would be in decimal degrees and automatically update the output cell size. Since the transformation between angular (decimal degrees) and nominal (meters) measurements varies across the image, the transformation is based on the center of the output file.



Enter the nominal cell size in the Nominal Cell Size dialog.

Nearest Neighbor

To determine an output pixel's nearest neighbor, the rectified coordinates (x_o, y_o) of the pixel are retransformed back to the source coordinate system using the inverse of the transformation. The retransformed coordinates (x_r, y_r) are used in bilinear interpolation and cubic convolution as well. The pixel that is closest to the retransformed coordinates (x_r, y_r) is the nearest neighbor. The data file value(s) for that pixel become the data file value(s) of the pixel in the output image.

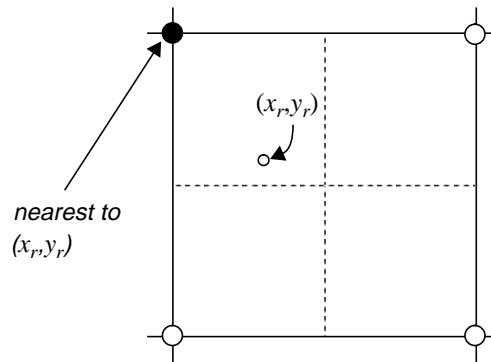


Figure 9-14 Nearest Neighbor

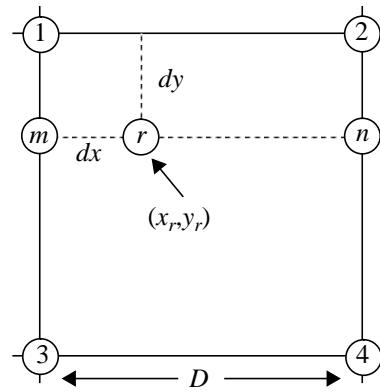
Table 9-2 Nearest Neighbor Resampling

Advantages	Disadvantages
Transfers original data values without averaging them as the other methods do; therefore, the extremes and subtleties of the data values are not lost. This is an important consideration when discriminating between vegetation types, locating an edge associated with a lineament, or determining different levels of turbidity or temperatures in a lake (Jensen 1996).	When this method is used to resample from a larger to a smaller grid size, there is usually a stair stepped effect around diagonal lines and curves.
Suitable for use before classification.	Data values may be dropped, while other values may be duplicated.
The easiest of the three methods to compute and the fastest to use.	Using on linear thematic data (e.g., roads, streams) may result in breaks or gaps in a network of linear data.
Appropriate for thematic files, which can have data file values based on a qualitative (nominal or ordinal) system or a quantitative (interval or ratio) system. The averaging that is performed with bilinear interpolation and cubic convolution is not suited to a qualitative class value system.	

Bilinear Interpolation

In bilinear interpolation, the data file value of the rectified pixel is based upon the distances between the retransformed coordinate location (x_r, y_r) and the four closest pixels in the input (source) image (see Figure 9-15). In this example, the neighbor pixels are numbered 1, 2, 3, and 4. Given the data file values of these four pixels on a grid, the task is to calculate a data file value for $r (V_r)$.

Rectification



r is the location of the retransformed coordinate

Figure 9-15 Bilinear Interpolation

To calculate V_r , first V_m and V_n are considered. By interpolating V_m and V_n , you can perform linear interpolation, which is a simple process to illustrate. If the data file values are plotted in a graph relative to their distances from one another, then a visual linear interpolation is apparent. The data file value of m (V_m) is a function of the change in the data file value between pixels 3 and 1 (that is, $V_3 - V_1$).

Calculating a data file value as a function of spatial distance between two pixels

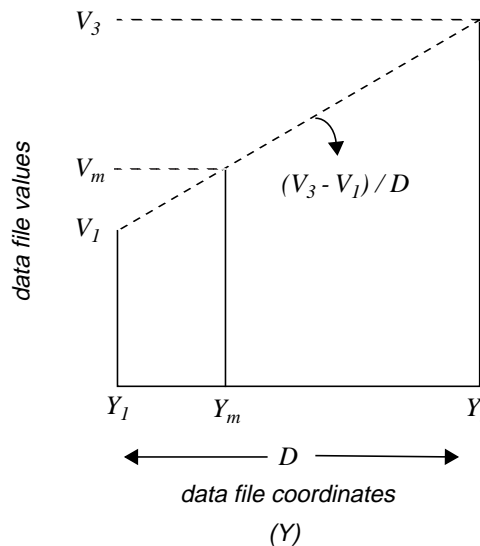


Figure 9-16 Linear Interpolation

The equation for calculating V_m from V_1 and V_3 is:

$$V_m = \left[\frac{V_3 - V_1}{D} \right] \times dy + V_1$$

Where:

Y_i = the Y coordinate for pixel i

V_i = the data file value for pixel i

dy = the distance between Y_l and Y_m in the source coordinate system

D = the distance between Y_l and Y_3 in the source coordinate system

If one considers that $(V_3 - V_1 / D)$ is the slope of the line in the graph above, then this equation translates to the equation of a line in $y = mx + b$ form.

Similarly, the equation for calculating the data file value for n (V_n) in the pixel grid is:

$$V_n = \left[\frac{V_4 - V_2}{D} \right] \times dy + V_2$$

From V_n and V_m , the data file value for r , which is at the retransformed coordinate location (x_r, y_r) , can be calculated in the same manner:

$$V_r = \left[\frac{V_n - V_m}{D} \right] \times dx + V_m$$

The following is attained by plugging in the equations for V_m and V_n to this final equation for V_r :

$$V_r = \frac{\left[\frac{V_4 - V_2}{D} \times dy + V_2 \right] - \left[\frac{V_3 - V_1}{D} \times dy + V_1 \right]}{D} \times dx + \left[\frac{V_3 - V_1}{D} \times dy + V_1 \right]$$

$$V_r = \frac{V_1(D - dx)(D - dy) + V_2(dx)(D - dy) + V_3(D - dx)(dy) + V_4(dx)(dy)}{D_2}$$

In most cases $D = 1$, since data file coordinates are used as the source coordinates and data file coordinates increment by 1.

Some equations for bilinear interpolation express the output data file value as:

$$V_r = \sum w_i V_i$$

Rectification

Where:

w_i is a weighting factor

The equation above could be expressed in a similar format, in which the calculation of w_i is apparent:

$$V_r = \sum_{i=1}^4 \frac{(D - \Delta x_i)(D - \Delta y_i)}{D^2} \times V_i$$

Where:

Δx_i = the change in the X direction between (x_r, y_r) and the data file coordinate of pixel i

Δy_i = the change in the Y direction between (x_r, y_r) and the data file coordinate of pixel i

V_i = the data file value for pixel i

D = the distance between pixels (in X or Y) in the source coordinate system

For each of the four pixels, the data file value is weighted more if the pixel is closer to (x_r, y_r) .

Table 9-3 Bilinear Interpolation Resampling

Advantages	Disadvantages
Results in output images that are smoother, without the stair-stepped effect that is possible with nearest neighbor.	Since pixels are averaged, bilinear interpolation has the effect of a low-frequency convolution. Edges are smoothed, and some extremes of the data file values are lost.
More spatially accurate than nearest neighbor.	
This method is often used when changing the cell size of the data, such as in SPOT/TM merges within the 2×2 resampling matrix limit.	



See "CHAPTER 5: Enhancement" for more about convolution filtering.

Cubic Convolution

Cubic convolution is similar to bilinear interpolation, except that:

- a set of 16 pixels, in a 4×4 array, are averaged to determine the output data file value, and
- an approximation of a cubic function, rather than a linear function, is applied to those 16 input values.

To identify the 16 pixels in relation to the retransformed coordinate (x_r, y_r) , the pixel (i, j) is used, such that:

$$i = \text{int}(x_r)$$

$$j = \text{int}(y_r)$$

This assumes that (x_r, y_r) is expressed in data file coordinates (pixels). The pixels around (i, j) make up a 4×4 grid of input pixels, as illustrated in Figure 9-17.

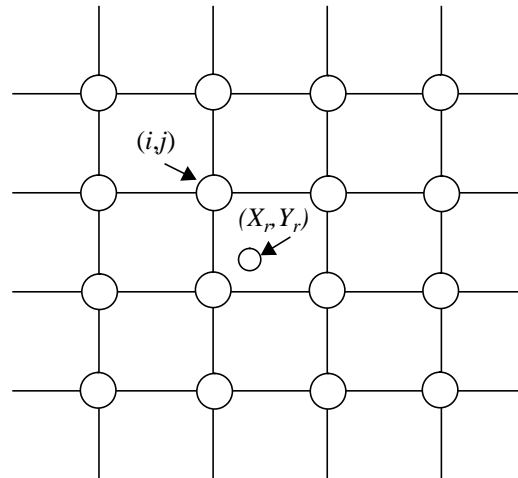


Figure 9-17 Cubic Convolution

Since a cubic, rather than a linear, function is used to weight the 16 input pixels, the pixels farther from (x_r, y_r) have exponentially less weight than those closer to (x_r, y_r) .

Several versions of the cubic convolution equation are used in the field. Different equations have different effects upon the output data file values. Some convolutions may have more of the effect of a low-frequency filter (like bilinear interpolation), serving to average and smooth the values. Others may tend to sharpen the image, like a high-frequency filter. The cubic convolution used in ERDAS IMAGINE is a compromise between low-frequency and high-frequency. The general effect of the cubic convolution depends upon the data.

The formula used in ERDAS IMAGINE is:

$$V_r = \sum_{n=1}^4 V(i-1, j+n-2) \times f(d(i-1, j+n-2) + 1)$$

$$+ V(i, j+n-2) \times f(d(i, j+n-2))$$

$$+ V(i+1, j+n-2) \times f(d(i+1, j+n-2) - 1)$$

Rectification

$$+ V(i + 2, j + n - 2) \times f(d(i + 2, j + n - 2) - 2)$$

Where:

$$i = \text{int}(x_r)$$

$$j = \text{int}(y_r)$$

$d(i, j)$ = the distance between a pixel with coordinates (i, j) and (x_r, y_r)

$V(i, j)$ = the data file value of pixel (i, j)

V_r = the output data file value

$a = -0.5$ (a constant which differs in other applications of cubic convolution)

$f(x)$ = the following function:

$$f(x) = \begin{cases} (a + 2)|x|^3 - (a + 3)|x|^2 + 1 & \text{if } |x| < 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a & \text{if } 1 < |x| < 2 \\ 0 & \text{otherwise} \end{cases}$$

Source: Atkinson 1985

In most cases, a value for a of -0.5 tends to produce output layers with a mean and standard deviation closer to that of the original data (Atkinson 1985).

Table 9-4 Cubic Convolution Resampling

Advantages	Disadvantages
Uses 4×4 resampling. In most cases, the mean and standard deviation of the output pixels match the mean and standard deviation of the input pixels more closely than any other resampling method.	Data values may be altered.
The effect of the cubic curve weighting can both sharpen the image and smooth out noise (Atkinson 1985). The actual effects depend upon the data being used.	The most computationally intensive resampling method, and is therefore the slowest.
This method is recommended when you are dramatically changing the cell size of the data, such as in TM/aerial photo merges (i.e., matches the 4×4 window more closely than the 2×2 window).	

**Map-to-Map
Coordinate
Conversions**

There are many instances when you may need to change a map that is already registered to a planar projection to another projection. Some examples of when this is required are as follows (ESRI 1992):

- When combining two maps with different projection characteristics.
- When the projection used for the files in the data base does not produce the desired properties of a map.
- When it is necessary to combine data from more than one zone of a projection, such as UTM or State Plane.

A change in the projection is a geometric change—distances, areas, and scale are represented differently. Therefore, the conversion process requires that pixels be resampled.

Resampling causes some of the spectral integrity of the data to be lost (see the disadvantages of the resampling methods explained previously). So, it is not usually wise to resample data that have already been resampled if the accuracy of data file values is important to the application. If the original unrectified data are available, it is usually wiser to rectify that data to a second map projection system than to lose a generation by converting rectified data and resampling it a second time.

Conversion Process

To convert the map coordinate system of any georeferenced image, ERDAS IMAGINE provides a shortcut to the rectification process. In this procedure, GCPs are generated automatically along the intersections of a grid that you specify. The program calculates the reference coordinates for the GCPs with the appropriate conversion formula and a transformation that can be used in the regular rectification process.

Vector Data

Converting the map coordinates of vector data is much easier than converting raster data. Since vector data are stored by the coordinates of nodes, each coordinate is simply converted using the appropriate conversion formula. There are no coordinates between nodes to extrapolate.

Rectification

CHAPTER 10

Terrain Analysis

Introduction

Terrain analysis involves the processing and graphic simulation of elevation data. Terrain analysis software functions usually work with topographic data (also called terrain data or elevation data), in which an elevation (or Z value) is recorded at each X,Y location. However, terrain analysis functions are not restricted to topographic data. Any series of values, such as population densities, ground water pressure values, magnetic and gravity measurements, and chemical concentrations, may be used.

Topographic data are essential for studies of trafficability, route design, nonpoint source pollution, intervisibility, siting of recreation areas, etc. (Welch 1990). Especially useful are products derived from topographic data. These include:

- slope images—illustrates changes in elevation over distance. Slope images are usually color-coded according to the steepness of the terrain at each pixel.
- aspect images—illustrates the prevailing direction that the slope faces at each pixel.
- shaded relief images—illustrates variations in terrain by differentiating areas that would be illuminated or shadowed by a light source simulating the sun.

Topographic data and its derivative products have many applications, including:

- calculating the shortest and most navigable path over a mountain range for constructing a road or routing a transmission line
- determining rates of snow melt based on variations in sun shadow, which is influenced by slope, aspect, and elevation

Terrain data are often used as a component in complex GIS modeling or classification routines. They can, for example, be a key to identifying wildlife habitats that are associated with specific elevations. Slope and aspect images are often an important factor in assessing the suitability of a site for a proposed use. Terrain data can also be used for vegetation classification based on species that are terrain-sensitive (e.g., Alpine vegetation).



Although this chapter mainly discusses the use of topographic data, the ERDAS IMAGINE terrain analysis functions can be used on data types other than topographic data.



See "CHAPTER 11: Geographic Information Systems" for more information about GIS modeling.

Topographic Data

Topographic data are usually expressed as a series of points with X,Y, and Z values. When topographic data are collected in the field, they are surveyed at a series of points including the extreme high and low points of the terrain along features of interest that define the topography such as streams and ridge lines, and at various points in between.

DEM and DTED are expressed as regularly spaced points. To create DEM and DTED files, a regular grid is overlaid on the topographic contours. Elevations are read at each grid intersection point, as shown in Figure 10-1.

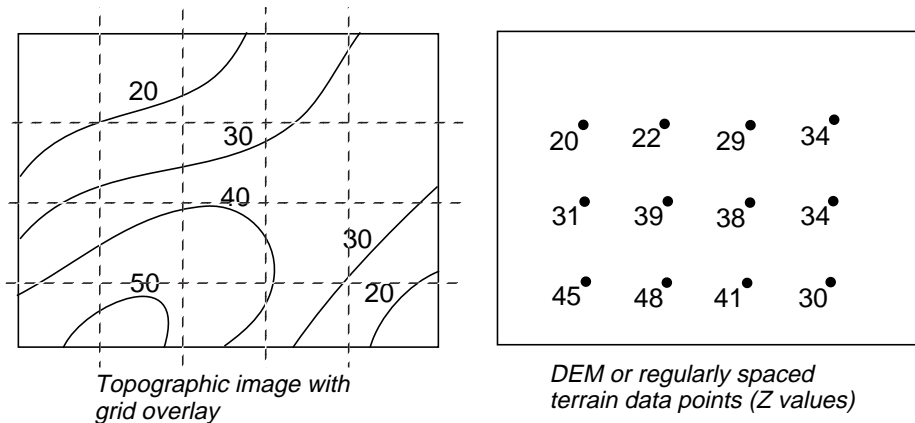


Figure 10-1 Regularly Spaced Terrain Data Points

Elevation data are derived from ground surveys and through manual photogrammetric methods. Elevation points can also be generated through digital orthographic methods.



See "CHAPTER 3: Raster and Vector Data Sources" for more details on DEM and DTED data. See "CHAPTER 7: Photogrammetric Concepts" for more information on the digital orthographic process.



To make topographic data usable in ERDAS IMAGINE, they must be represented as a surface, or DEM. A DEM is a one-band image file where the value of each pixel is a specific elevation value. A gray scale is used to differentiate variations in terrain.



DEMs can be edited with the Raster Editing capabilities of ERDAS IMAGINE. See "CHAPTER 1: Raster Data" for more information.

Slope Images

Slope is expressed as the change in elevation over a certain distance. In this case, the certain distance is the size of the pixel. Slope is most often expressed as a percentage, but can also be calculated in degrees.



Use the Slope function in Image Interpreter to generate a slope image.

In ERDAS IMAGINE, the relationship between percentage and degree expressions of slope is as follows:

- a 45° angle is considered a 100% slope
- a 90° angle is considered a 200% slope
- slopes less than 45° fall within the 1 - 100% range
- slopes between 45° and 90° are expressed as 100 - 200% slopes

A 3×3 pixel window is used to calculate the slope at each pixel. For a pixel at location X, Y , the elevations around it are used to calculate the slope as shown in Figure 10-2. In Figure 10-2, each pixel has a ground resolution of 30×30 meters.

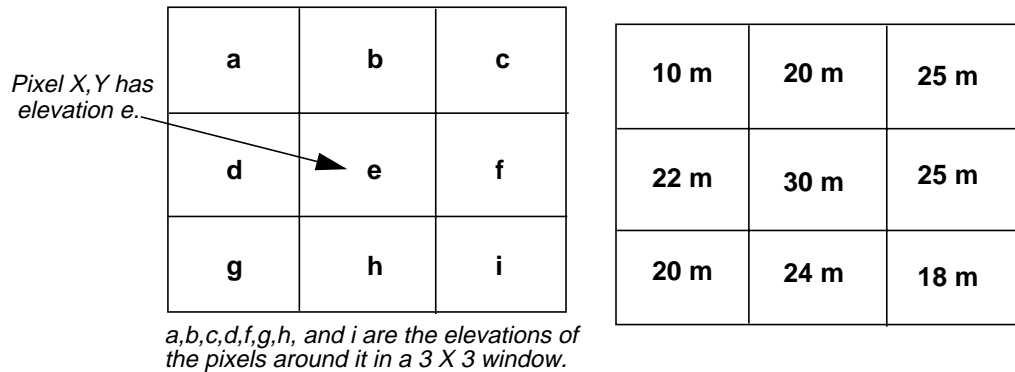


Figure 10-2 3×3 Window Calculates the Slope at Each Pixel

First, the average elevation changes per unit of distance in the x and y direction (Δx and Δy) are calculated as:

$$\Delta x_1 = c - a$$

$$\Delta x_2 = f - d$$

$$\Delta x_3 = i - g$$

$$\Delta y_1 = a - g$$

$$\Delta y_2 = b - h$$

$$\Delta y_3 = c - i$$

$$\Delta x = (\Delta x_1 + \Delta x_2 + \Delta x_3) / 3 \times x_s$$

$$\Delta y = (\Delta y_1 + \Delta y_2 + \Delta y_3) / 3 \times y_s$$

Terrain Analysis

Where:

$a...i$ = elevation values of pixels in a 3×3 window, as shown above

x_s = x pixel size = 30 meters

y_s = y pixel size = 30 meters

The slope at pixel x,y is calculated as:

$$s = \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2}}{2} \qquad s = 0.0967$$

$$\text{if } s \leq 1 \quad \text{percent slope} = s \times 100$$

$$\text{if } s > 1 \quad \text{percent slope} = 200 - \frac{100}{s}$$

$$\text{slope in degrees} = \tan^{-1}(s) \times \frac{180}{\pi}$$

Example

Slope images are often used in road planning. For example, if the Department of Transportation specifies a maximum of 15% slope on any road, it would be possible to recode all slope values that are greater than 15% as unsuitable for road building.

A hypothetical example is given in Figure 10-3, which shows how the slope is calculated for a single pixel.

10 m	20 m	25 m
22 m		25 m
20 m	24 m	18 m

*The pixel for which slope is being calculated is shaded.
The elevations of the neighboring pixels are given in meters.*

Figure 10-3 Slope Calculation Example

So, for the hypothetical example:

$$\Delta x_1 = 25 - 10 = 15$$

$$\Delta y_1 = 10 - 20 = -10$$

$$\Delta x_2 = 25 - 22 = 3$$

$$\Delta y_2 = 20 - 24 = -4$$

$$\Delta x_3 = 18 - 20 = -2$$

$$\Delta y_3 = 25 - 18 = 7$$

$$\Delta x = \frac{15 + 3 - 2}{30 \times 3} = 0.177$$

$$\Delta y = \frac{-10 - 4 + 7}{30 \times 3} = -0.078$$

For the example, the slope is:

$$\text{slope in degrees} = \tan^{-1}(s) \times \frac{180}{\pi} = \tan^{-1}(0.0967) \times 57.30 = 5.54$$

$$\text{percent slope} = 0.0967 \times 100 = 9.67\%$$

Aspect Images

An aspect image is an image file that is gray scale coded according to the prevailing direction of the slope at each pixel. Aspect is expressed in degrees from north, clockwise, from 0 to 360. Due north is 0 degrees. A value of 90 degrees is due east, 180 degrees is due south, and 270 degrees is due west. A value of 361 degrees is used to identify flat surfaces such as water bodies.



Use the *Aspect* function in *Image Interpreter* to generate an aspect image.

As with slope calculations, aspect uses a 3×3 window around each pixel to calculate the prevailing direction it faces. For pixel x,y with the following elevation values around it, the average changes in elevation in both x and y directions are calculated first. Each pixel is 30×30 meters in the following example:

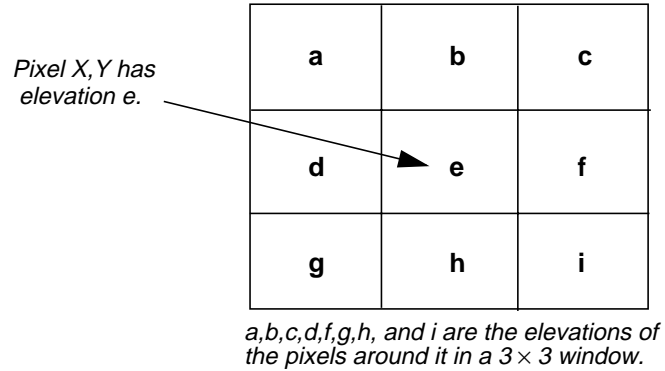


Figure 10-4 3×3 Window Calculates the Aspect at Each Pixel

$$\Delta x_1 = c - a$$

$$\Delta y_1 = a - g$$

$$\Delta x_2 = f - d$$

$$\Delta y_2 = b - h$$

$$\Delta x_3 = i - g$$

$$\Delta y_3 = c - i$$

Where:

$a...i$ = elevation values of pixels in a 3×3 window as shown above

$$\Delta x = (\Delta x_1 + \Delta x_2 + \Delta x_3) / 3$$

$$\Delta y = (\Delta y_1 + \Delta y_2 + \Delta y_3) / 3$$

If $\Delta x = 0$ and $\Delta y = 0$, then the aspect is flat (coded to 361 degrees). Otherwise, θ is calculated as:

$$\theta = \tan^{-1}\left(\frac{\Delta x}{\Delta y}\right)$$



Note that θ is calculated in radians

then aspect is $180 + \theta$ (in degrees).

Example

Aspect files are used in many of the same applications as slope files. In transportation planning, for example, north facing slopes are often avoided. Especially in northern climates, these would be exposed to the most severe weather and would hold snow and ice the longest. It would be possible to recode all pixels with north facing aspects as undesirable for road building.

A hypothetical example is given in Figure 10-5, which shows how the aspect is calculated for a single pixel.

10 m	20 m	25 m
22 m		25 m
20 m	24 m	18 m

*The pixel for which slope is being calculated is shaded.
The elevations of the neighboring pixels are given in meters.*

Figure 10-5 Aspect Calculation Example

$$\Delta x = \frac{15 + 3 - 2}{3} = 5.33 \qquad \Delta y = \frac{-10 - 4 + 7}{3} = -2.33$$

$$\theta = \tan^{-1}\left(\frac{5.33}{-2.33}\right) = 1.98$$

$$1.98 \text{ radians} = 113.6 \text{ degrees}$$

$$\text{aspect} = 180 + 113.6 = 293.6 \text{ degrees}$$

Shaded Relief

A shaded relief image provides an illustration of variations in elevation. Based on a user-specified position of the sun, areas that would be in sunlight are highlighted and areas that would be in shadow are shaded. Shaded relief images are generated from an elevation surface, alone or in combination with an image file draped over the terrain.

It is important to note that the relief program identifies shadowed areas—i.e., those that are not in direct sun. It does not calculate the shadow that is cast by topographic features onto the surrounding surface.

Terrain Analysis

For example, a high mountain with sunlight coming from the northwest would be symbolized as follows in shaded relief. Only the portions of the mountain that would be in shadow from a northwest light would be shaded. The software would not simulate a shadow that the mountain would cast on the southeast side.

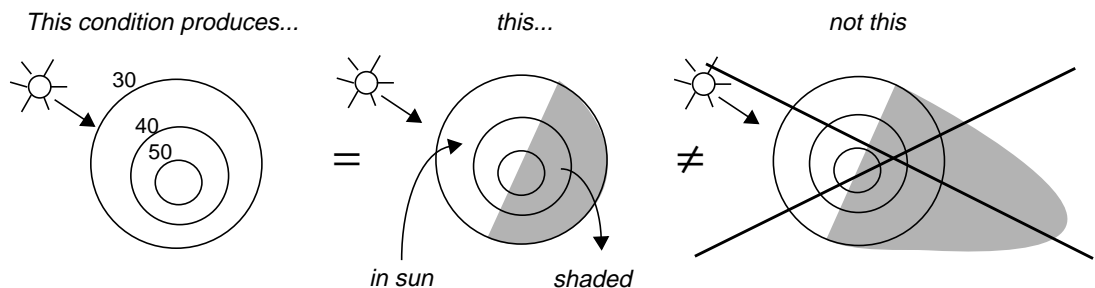


Figure 10-6 Shaded Relief

Shaded relief images are an effective graphic tool. They can also be used in analysis, e.g., snow melt over an area spanned by an elevation surface. A series of relief images can be generated to simulate the movement of the sun over the landscape. Snow melt rates can then be estimated for each pixel based on the amount of time it spends in sun or shadow. Shaded relief images can also be used to enhance subtle detail in gray scale images such as aeromagnetic, radar, gravity maps, etc.



Use the Shaded Relief function in Image Interpreter to generate a relief image.

In calculating relief, the software compares the user-specified sun position and angle with the angle each pixel faces. Each pixel is assigned a value between -1 and +1 to indicate the amount of light reflectance at that pixel.

- Negative numbers and zero values represent shadowed areas.
- Positive numbers represent sunny areas, with +1 assigned to the areas of highest reflectance.

The reflectance values are then applied to the original pixel values to get the final result. All negative values are set to 0 or to the minimum light level specified by you. These indicate shadowed areas. Light reflectance in sunny areas falls within a range of values depending on whether the pixel is directly facing the sun or not. (In the example above, pixels facing northwest would be the brightest. Pixels facing north-northwest and west-northwest would not be quite as bright.)

In a relief file, which is a DEM that shows surface relief, the surface reflectance values are multiplied by the color lookup values for the image file.

Topographic Normalization

Digital imagery from mountainous regions often contains a radiometric distortion known as topographic effect. Topographic effect results from the differences in illumination due to the angle of the sun and the angle of the terrain. This causes a variation in the image brightness values. Topographic effect is a combination of:

- incident illumination—the orientation of the surface with respect to the rays of the sun
- exitance angle—the amount of reflected energy as a function of the slope angle
- surface cover characteristics—rugged terrain with high mountains or steep slopes (Hodgson and Shelley 1993)

One way to reduce topographic effect in digital imagery is by applying transformations based on the Lambertian or Non-Lambertian reflectance models. These models normalize the imagery, which makes it appear as if it were a flat surface.



The Topographic Normalize function in Image Interpreter uses a Lambertian Reflectance model to normalize topographic effect in VIS/IR imagery.

When using the Topographic Normalization model, the following information is needed:

- solar elevation and azimuth at time of image acquisition
- DEM file
- original imagery file (after atmospheric corrections)

Lambertian Reflectance Model

The Lambertian Reflectance model assumes that the surface reflects incident solar energy uniformly in all directions, and that variations in reflectance are due to the amount of incident radiation.

The following equation produces normalized brightness values (Colby 1991; Smith *et al* 1980):

$$BV_{normal} \lambda = BV_{observed} \lambda / \cos i$$

Where:

$BV_{normal} \lambda$ = normalized brightness values
 $BV_{observed} \lambda$ = observed brightness values
 $\cos i$ = cosine of the incidence angle

Incidence Angle

The incidence angle is defined from:

$$\cos i = \cos (90 - \theta_s) \cos \theta_n + \sin (90 - \theta_s) \sin \theta_n \cos (\phi_s - \phi_n)$$

Where:

i = the angle between the solar rays and the normal to the surface
 θ_s = the elevation of the sun
 ϕ_s = the azimuth of the sun
 θ_n = the slope of each surface element
 ϕ_n = the aspect of each surface element

Terrain Analysis

If the surface has a slope of 0 degrees, then aspect is undefined and i is simply $90 - \theta_s$.

Non-Lambertian Model

Minnaert (1961) proposed that the observed surface does not reflect incident solar energy uniformly in all directions. Instead, he formulated the Non-Lambertian model, which takes into account variations in the terrain. This model, although more computationally demanding than the Lambertian model, may present more accurate results.

In a Non-Lambertian Reflectance model, the following equation is used to normalize the brightness values in the image (Colby 1991; Smith *et al* 1980):

$$BV_{normal \lambda} = (BV_{observed \lambda} \cos e) / (\cos^k i \cos^k e)$$

Where:

$BV_{normal \lambda}$ = normalized brightness values

$BV_{observed \lambda}$ = observed brightness values

$\cos i$ = cosine of the incidence angle

$\cos e$ = cosine of the exitance angle, or slope angle

k = the empirically derived Minnaert constant

Minnaert Constant

The Minnaert constant (k) may be found by regressing a set of observed brightness values from the remotely sensed imagery with known slope and aspect values, provided that all the observations in this set are the same type of land cover. The k value is the slope of the regression line (Hodgson and Shelley 1993):

$$\log (BV_{observed \lambda} \cos e) = \log BV_{normal \lambda} + k \log (\cos i \cos e)$$



Use the Spatial Modeler to create a model based on the Non-Lambertian model.

NOTE: The Non-Lambertian model does not detect surfaces that are shadowed by intervening topographic features between each pixel and the sun. For these areas, a line-of-sight algorithm can identify such shadowed pixels.

CHAPTER 11

Geographic Information Systems

Introduction

The dawning of GIS can legitimately be traced back to the beginning of the human race. The earliest known map dates back to 2500 B.C., but there were probably maps before that time. Since then, humans have been continually improving the methods of conveying spatial information. The mid-eighteenth century brought the use of map overlays to show troop movements in the Revolutionary War. This could be considered an early GIS. The first British census in 1825 led to the science of demography, another application for GIS. During the 1800s, many different cartographers and scientists were all discovering the power of overlays to convey multiple levels of information about an area (Star and Estes).

Frederick Law Olmstead has long been considered the father of Landscape Architecture for his pioneering work in the early 20th century. Many of the methods Olmstead used in Landscape Architecture also involved the use of hand-drawn overlays. This type of analysis was beginning to be used for a much wider range of applications, such as change detection, urban planning, and resource management (Rado 1992).

The first system to be called a GIS was the Canadian Geographic Information System, developed in 1962 by Roger Tomlinson of the Canada Land Inventory. Unlike earlier systems that were developed for a specific application, this system was designed to store digitized map data and land-based attributes in an easily accessible format for all of Canada. This system is still in operation today (Parent and Church 1987).

In 1969, Ian McHarg's influential work, *Design with Nature*, was published. This work on land suitability/capability analysis (SCA), a system designed to analyze many data layers to produce a plan map, discussed the use of overlays of spatially referenced data layers for resource planning and management (Star and Estes 1990).

The era of modern GIS really started in the 1970s, as analysts began to program computers to automate some of the manual processes. Software companies like ESRI and ERDAS developed software packages that could input, display, and manipulate geographic data to create new layers of information. The steady advances in features and power of the hardware over the last ten years—and the decrease in hardware costs—have made GIS technology accessible to a wide range of users. The growth rate of the GIS industry in the last several years has exceeded even the most optimistic projections.

Today, a GIS is a unique system designed to input, store, retrieve, manipulate, and analyze layers of geographic data to produce interpretable information. A GIS should also be able to create reports and maps (Marble 1990). The GIS database may include computer images, hardcopy maps, statistical data, or any other data that is needed in a study. Although the term GIS is commonly used to describe software packages, a true GIS includes knowledgeable staff, a training program, budgets, marketing, hardware, data, and software (Walker and Miller 1990). GIS technology can be used in almost any geography-related discipline, from Landscape Architecture to natural resource management to transportation routing.

The central purpose of a GIS is to turn geographic data into useful information—the answers to real-life questions—questions such as:

- How can we monitor the influence of global climatic changes on the Earth's resources?
- How should political districts be redrawn in a growing metropolitan area?
- Where is the best place for a shopping center that is most convenient to shoppers and least harmful to the local ecology?
- What areas should be protected to ensure the survival of endangered species?
- How can communities be better prepared to face natural disasters, such as earthquakes, tornadoes, hurricanes, and floods?

Information vs. Data

Information, as opposed to data, is independently meaningful. It is relevant to a particular problem or question:

- “The land cover at coordinate N875250, E757261 has a data file value 8,” is data.
- “Land cover with a value of 8 are on slopes too steep for development,” is information.

You can input data into a GIS and output information. The information you wish to derive determines the type of data that must be input. For example, if you are looking for a suitable refuge for bald eagles, zip code data is probably not needed, while land cover data may be useful.

For this reason, the first step in any GIS project is usually an assessment of the scope and goals of the study. Once the project is defined, you can begin the process of building the database. Although software and data are commercially available, a custom database must be created for the particular project and study area. The database must be designed to meet the needs of the organization and objectives. ERDAS IMAGINE provides tools required to build and manipulate a GIS database.

Successful GIS implementation typically includes two major steps:

- data input
- analysis

Data input involves collecting the necessary data layers into a GIS database. In the analysis phase, these data layers are combined and manipulated in order to create new layers and to extract meaningful information from them. This chapter discusses these steps in detail.

Data Input

Acquiring the appropriate data for a project involves creating a database of layers that encompasses the study area. A database created with ERDAS IMAGINE can consist of:

- continuous layers (satellite imagery, aerial photographs, elevation data, etc.)
- thematic layers (land use, vegetation, hydrology, soils, slope, etc.)
- vector layers (streets, utility and communication lines, parcels, etc.)
- statistics (frequency of an occurrence, population demographics, etc.)
- attribute data (characteristics of roads, land, imagery, etc.)

The ERDAS IMAGINE software package employs a hierarchical, object-oriented architecture that utilizes both raster imagery and topological vector data. Raster images are stored in image files, and vector layers are coverages or shapefiles based on the ESRI ArcInfo and ArcView data models. The seamless integration of these two types of data enables you to reap the benefits of both data formats in one system.

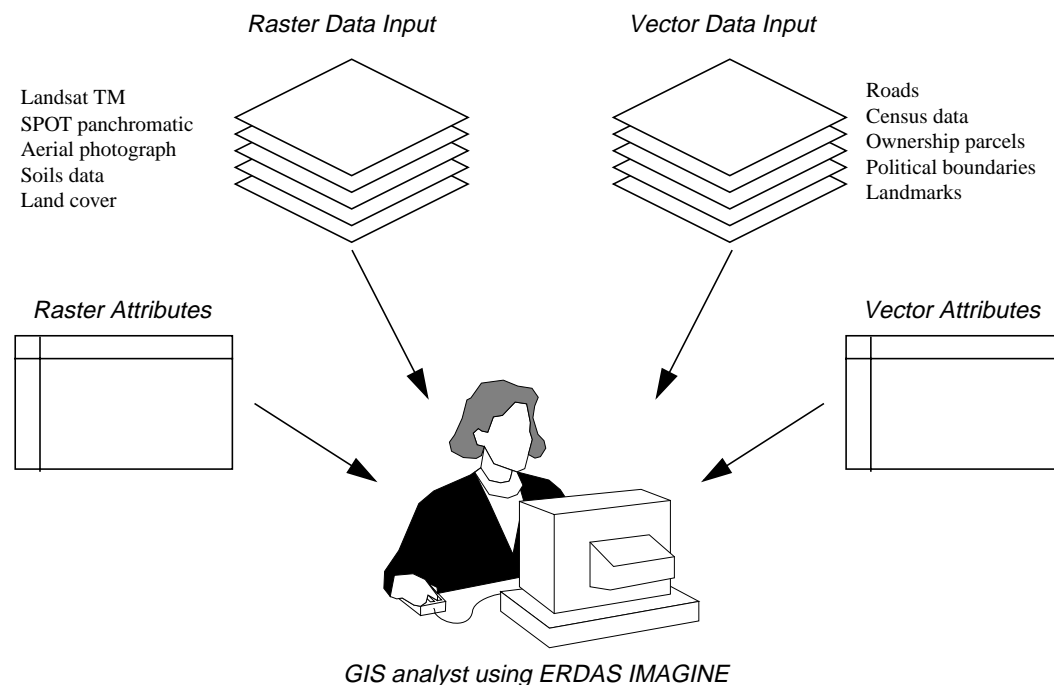


Figure 11-1 Data Input

Raster data might be more appropriate in the following applications:

- site selection
- natural resource management
- petroleum exploration
- mission planning
- change detection

On the other hand, vector data may be better suited for these applications:

- urban planning
- tax assessment and planning
- traffic engineering
- facilities management

The advantage of an integrated raster and vector system such as ERDAS IMAGINE is that one data structure does not have to be chosen over the other. Both data formats can be used and the functions of both types of systems can be accessed. Depending upon the project, only raster or vector data may be needed, but most applications benefit from using both.

Themes and Layers

A database usually consists of files with data of the same geographical area, with each file containing different types of information. For example, a database for the city recreation department might include files of all the parks in the area. These files might depict park boundaries, county and municipal boundaries, vegetation types, soil types, drainage basins, slope, roads, etc. Each of these files contains different information—each is a different theme. The concept of themes has evolved from early GISs, in which transparent overlays were created for each theme and combined (overlaid) in different ways to derive new information.

A single theme may require more than a simple raster or vector file to fully describe it. In addition to the image, there may be attribute data that describe the information, a color scheme, or meaningful annotation for the image. The full collection of data that describe a certain theme is called a layer.

Depending upon the goals of a project, it may be helpful to combine several themes into one layer. For example, if you want to propose a new park site, you might create one layer that shows roads, land cover, land ownership, slope, etc., and indicate through the use of colors and/or annotation which areas would be best for the new site. This one layer would then include many separate themes. Much of GIS analysis is concerned with combining individual themes into one or more layers that answer the questions driving the analysis. This chapter explores these analysis techniques.

Continuous Layers

Continuous raster layers are quantitative (measuring a characteristic) and have related, continuous values. Continuous raster layers can be multiband (e.g., Landsat TM) or single band (e.g., SPOT panchromatic).

Satellite images, aerial photographs, elevation data, scanned maps, and other continuous raster layers can be incorporated into a database and provide a wealth of information that is not available in thematic layers or vector layers. In fact, these layers often form the foundation of the database. Extremely accurate base maps can be created from rectified satellite images or aerial photographs. Then, all other layers that are added to the database can be registered to this base map.

Once used only for image processing, continuous data are now being incorporated into GIS databases and used in combination with thematic data to influence processing algorithms or as backdrop imagery on which to display the results of analyses. Current satellite data and aerial photographs are also effective in updating outdated vector data. The vectors can be overlaid on the raster backdrop and updated dynamically to reflect new or changed features, such as roads, utility lines, or land use. This chapter explores the many uses of continuous data in a GIS.



See "CHAPTER 1: Raster Data" for more information on continuous data.

Thematic Layers

Thematic data are typically represented as single layers of information stored as image files and containing discrete classes. Classes are simply categories of pixels which represent the same condition. An example of a thematic layer is a vegetation classification with discrete classes representing coniferous forest, deciduous forest, wetlands, agriculture, urban, etc.

A thematic layer is sometimes called a variable, because it represents one of many characteristics about the study area. Since thematic layers usually have only one band, they are usually displayed in pseudo color mode, where particular colors are often assigned to help visualize the information. For example, blues are usually used for water features, greens for healthy vegetation, etc.



See "CHAPTER 4: Image Display" for information on pseudo color display.

Class Numbering Systems

As opposed to the data file values of continuous raster layers, which are generally multiband and statistically related, the data file values of thematic raster layers can have a nominal, ordinal, interval, or ratio relationship (Star and Estes 1990).

- Nominal classes represent categories with no particular order. Usually, these are characteristics that are not associated with quantities (e.g., soil type or political area).
- Ordinal classes are those that have a sequence, such as poor, good, better, and best. An ordinal class numbering system is often created from a nominal system, in which classes have been ranked by some criteria. In the case of the recreation department database used in the previous example, the final layer may rank the proposed park sites according to their overall suitability.
- Interval classes also have a natural sequence, but the distance between each value is meaningful as well. This numbering system might be used for temperature data.
- Ratio classes differ from interval classes only in that ratio classes have a natural zero point, such as rainfall amounts.

The variable being analyzed, and the way that it contributes to the final product, determines the class numbering system used in the thematic layers. Layers that have one numbering system can easily be recoded to a new system. This is discussed in detail under "Recoding" on page 400.

Classification

Thematic layers can be generated from remotely sensed data (e.g., Landsat TM, SPOT) by using the ERDAS IMAGINE Image Interpreter, Classification, and Spatial Modeler tools. A frequent and popular application is the creation of land cover classification schemes through the use of both supervised (user-assisted) and unsupervised (automatic) pattern-recognition algorithms contained within ERDAS IMAGINE. The output is a single thematic layer that represents specific classes based on the approach selected.



See "CHAPTER 6: Classification" for more information.

Vector Data Converted to Raster Format

Vector layers can be converted to raster format if the raster format is more appropriate for an application. Typical vector layers, such as communication lines, streams, boundaries, and other linear features, can easily be converted to raster format within ERDAS IMAGINE for further analysis. Spatial Modeler automatically converts vector layers to raster for processing.



Use the Vector Utilities menu from the Vector icon in the ERDAS IMAGINE icon panel to convert vector layers to raster format, or use the vector layers directly in Spatial Modeler.



Other sources of raster data are discussed in "CHAPTER 3: Raster and Vector Data Sources".

Statistics

Both continuous and thematic layers include statistical information. Thematic layers contain the following information:

- a histogram of the data values, which is the total number of pixels in each class
- a list of class names that correspond to class values
- a list of class values
- a color table, stored as brightness values in red, green, and blue, which make up the colors of each class when the layer is displayed

For thematic data, these statistics are called attributes and may be accompanied by many other types of information, as described in "Attributes".



Use the Image Information option on the Viewer's tool bar to generate or update statistics for image files.



See "CHAPTER 1: Raster Data" for more information about the statistics stored with continuous layers.

Vector Layers

The vector layers used in ERDAS IMAGINE are based on the ArcInfo data model and consist of points, lines, and polygons. These layers are topologically complete, meaning that the spatial relationships between features are maintained. Vector layers can be used to represent transportation routes, utility corridors, communication lines, tax parcels, school zones, voting districts, landmarks, population density, etc. Vector layers can be analyzed independently or in combination with continuous and thematic raster layers.

Starting with ERDAS IMAGINE V8.4, vector layers may also be shapefiles based on the ArcView data model.

Vector data can be acquired from several private and governmental agencies. Vector data can also be created in ERDAS IMAGINE by digitizing on the screen, using a digitizing tablet, or converting other data types to vector format.



See "CHAPTER 2: Vector Layers" for more information on the characteristics of vector data.

Attributes

Text and numerical data that are associated with the classes of a thematic layer or the features in a vector layer are called attributes. This information can take the form of character strings, integer numbers, or floating point numbers. Attributes work much like the data that are handled by database management software. You may define fields, which are categories of information about each class. A record is the set of all attribute data for one class. Each record is like an index card, containing information about one class or feature in a file of many index cards, which contain similar information for the other classes or features.

Attribute information for raster layers is stored in the image file. Vector attribute information is stored in either an INFO file, dbf file, or SDE database. In both cases, there are fields that are automatically generated by the software, but more fields can be added as needed to fully describe the data. Both are viewed in CellArrays, which allow you to display and manipulate the information. However, raster and vector attributes are handled slightly differently, so a separate section on each follows.

Raster Attributes

In ERDAS IMAGINE, raster attributes for image files are accessible from the Raster Attribute Editor. The Raster Attribute Editor contains a CellArray, which is similar to a table or spreadsheet that not only displays the information, but also includes options for importing, exporting, copying, editing, and other operations.

Figure 11-2 shows the attributes for a land cover classification layer.

Row	Color	Red	Green	Blue	Opacity	Class_Names	Histogram
0		0	0	0	0	BG	0
1		0	0.784314	0	1	DECIDUOUS	71228
2		0	0.568627	0	1	DECIDUOUS/PINE MIXED	80848
3		0	0.392157	0	1	PINE	24911
4		0	0	0.784314	1	WATER	23070
5		0.509804	0.415686	0	1	AGRICULTURE	26986
6		0.941176	0.823529	0.196078	1	BARE GROUND	7400
7		0	1	0.666667	1	GRASS	12518

Figure 11-2 Raster Attributes for Inlandc.img

Most thematic layers contain the following attribute fields:

- Class Name
- Class Value
- Color table (red, green, and blue values)
- Opacity percentage
- Histogram (number of pixels in the file that belong to the class)

As many additional attribute fields as needed can be defined for each class.



See "CHAPTER 6: Classification" for more information about the attribute information that is automatically generated when new thematic layers are created in the classification process.

Viewing Raster Attributes

Simply viewing attribute information can be a valuable analysis tool. Depending on the type of information associated with the layers of a database, processing may be further refined by comparing the attributes of several files. When both the raster layer and its associated attribute information are displayed, you can select features in one using the other. For example, to locate the class name associated with a particular area in a displayed image, simply click in that area with the mouse and the associated row is highlighted in the Raster Attribute Editor.

Attribute information is accessible in several places throughout ERDAS IMAGINE. In some cases it is read-only and in other cases it is a fully functioning editor, allowing the information to be modified.

Manipulating Raster Attributes

The applications for manipulating attributes are as varied as the applications for GIS. The attribute information in a database depends on the goals of the project. Some of the attribute editing capabilities in ERDAS IMAGINE include:

- import/export ASCII information to and from other software packages, such as spreadsheets and word processors
- cut, copy, and paste individual cells, rows, or columns to and from the same Raster Attribute Editor or among several Raster Attribute Editors
- generate reports that include all or a subset of the information in the Raster Attribute Editor
- use formulas to populate cells
- directly edit cells by entering in new information

The Raster Attribute Editor in ERDAS IMAGINE also includes a color cell column, so that class (object) colors can be viewed or changed. In addition to direct manipulation, attributes can be changed by other programs. For example, some of the Image Interpreter functions calculate statistics that are automatically added to the Raster Attribute Editor. Models that read and/or modify attribute information can also be written.



See "CHAPTER 5: Enhancement" for more information on the Image Interpreter. There is more information on GIS modeling, starting on page 405.

Vector Attributes

Vector attributes are stored in the Vector Attributes CellArrays. You can simply view attributes or use them to:

- select features in a vector layer for further processing
- determine how vectors are symbolized
- label features

Figure 11-3 shows the attributes for a roads layer.

The screenshot shows a window titled "Attributes for /usr/imagene/data/imageneqa/vector/arcinfo/germ/roads". The window contains a table with the following data:

Record	FNODE#	TNODE#	LPOLY#	RPOLY#	LENGTH	ROADS#	ROADS-ID	CLASS
1	3	11	0	0	3288.524	1	1	1
2	12	4	0	0	3288.993	2	27	1
3	11	16	0	0	578.119	3	2	1
4	17	12	0	0	863.370	4	26	1
5	16	20	0	0	457.817	5	3	1
6	17	21	0	0	390.772	6	184	1
7	20	24	0	0	479.687	7	4	1
8	25	21	0	0	484.224	8	24	1
9	24	28	0	0	639.895	9	5	1
10	30	25	0	0	723.476	10	23	1
11	28	42	0	0	3170.395	11	6	1
12	43	30	0	0	3119.037	12	22	1
13	42	46	0	0	646.384	13	7	1
14	48	43	0	0	568.821	14	21	1
15	46	50	0	0	406.756	15	8	1
16	51	48	0	0	476.809	16	20	1
17	50	54	0	0	506.585	17	9	1

Figure 11-3 Vector Attributes CellArray



See "CHAPTER 2: Vector Layers" for more information about vector attributes.

Analysis

ERDAS IMAGINE Analysis Tools

In ERDAS IMAGINE, GIS analysis functions and algorithms are accessible through three main tools:

- script models created with SML
- graphical models created with Model Maker
- prepackaged functions in Image Interpreter

Spatial Modeler Language

SML is the basis for all ERDAS IMAGINE GIS functions. It is a modeling language that enables you to create script (text) models for a variety of applications. Models may be used to create custom algorithms that best suit your data and objectives.

Model Maker

Model Maker is essentially SML linked to a graphical interface. This enables you to create graphical models using a palette of easy-to-use tools. Graphical models can be run, edited, saved in libraries, or converted to script form and edited further, using SML.

NOTE: References to the Spatial Modeler in this chapter mean that the named procedure can be accomplished using both Model Maker and SML.

Image Interpreter

The Image Interpreter houses a set of common functions that were all created using either Model Maker or SML. They have been given a dialog interface to match the other processes in ERDAS IMAGINE. In most cases, these processes can be run from a single dialog. However, the actual models are also provided with the software to enable customized processing.

Many of the functions described in the following sections can be accomplished using any of these tools. Model Maker is also easy to use and utilizes many of the same steps that would be performed when drawing a flow chart of an analysis. SML is intended for more advanced analyses, and has been designed using natural language commands and simple syntax rules. Some applications may require a combination of these tools.

Customizing ERDAS IMAGINE Tools

ERDAS Macro Language (EML) enables you to create and add new and/or customized dialogs. If new capabilities are needed, they can be created with the IMAGINE Developers' Toolkit™. Using these tools, a GIS that is completely customized to a specific application and its preferences can be created.



See the ERDAS IMAGINE On-Line Help for more information about EML and the IMAGINE Developers' Toolkit.

Analysis Procedures

Once the database (layers and attribute data) is assembled, the layers can be analyzed and new information extracted. Some information can be extracted simply by looking at the layers and visually comparing them to other layers. However, new information can be retrieved by combining and comparing layers using the following procedures:

- Proximity analysis—the process of categorizing and evaluating pixels based on their distances from other pixels in a specified class or classes.
- Contiguity analysis—enables you to identify regions of pixels in the same class and to filter out small regions.
- Neighborhood analysis—any image processing technique that takes surrounding pixels into consideration, such as convolution filtering and scanning. This is similar to the convolution filtering performed on continuous data. Several types of analyses can be performed, such as boundary, density, mean, sum, etc.
- Recoding—enables you to assign new class values to all or a subset of the classes in a layer.

- **Overlaying**—creates a new file with either the maximum or minimum value of the input layers.
- **Indexing**—adds the values of the input layers.
- **Matrix analysis**—outputs the coincidence values of the input layers.
- **Graphical modeling**—enables you to combine data layers in an unlimited number of ways. For example, an output layer created from modeling can represent the desired combination of themes from many input layers.
- **Script modeling**—offers all of the capabilities of graphical modeling with the ability to perform more complex functions, such as conditional looping.

Using an Area of Interest

Any of these functions can be performed on a single layer or multiple layers. You can also select a particular AOI that is defined in a separate file (AOI layer, thematic raster layer, or vector layer) or an AOI that is selected immediately preceding the operation by entering specific coordinates or by selecting the area in a Viewer.

Proximity Analysis

Many applications require some measurement of distance or proximity. For example, a real estate developer would be concerned with the distance between a potential site for a shopping center and an interchange to a major highway.

Proximity analysis determines which pixels of a layer are located at specified distances from pixels in a certain class or classes. A new thematic layer (image file) is created, which is categorized by the distance of each pixel from specified classes of the input layer. This new file then becomes a new layer of the database and provides a buffer zone around the specified class(es). In further analysis, it may be beneficial to weight other factors, based on whether they fall inside or outside the buffer zone.

Figure 11-4 shows a layer containing lakes and streams and the resulting layer after a proximity analysis is run to create a buffer zone around all of the water features.

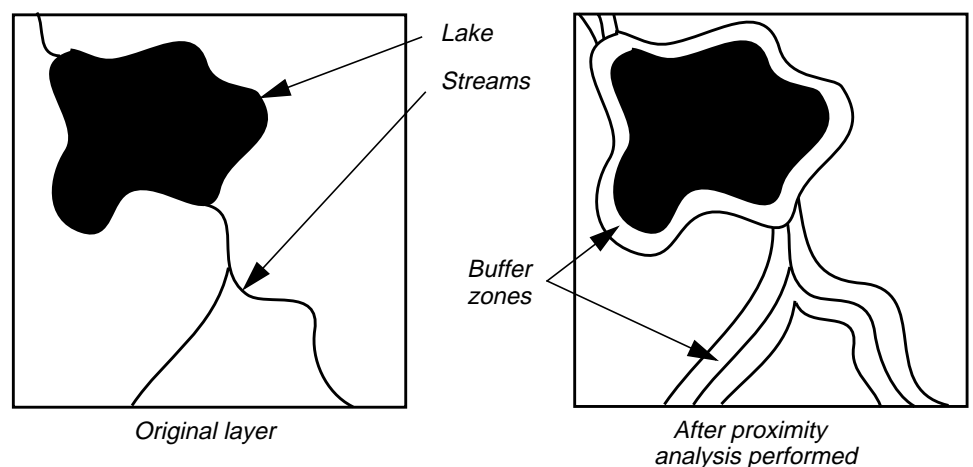


Figure 11-4 Proximity Analysis



Use the Search (GIS Analysis) function in Image Interpreter or Spatial Modeler to perform a proximity analysis.

Contiguity Analysis

Contiguity analysis is concerned with the ways in which pixels of a class are grouped together. Groups of contiguous pixels in the same class, called raster regions, or clumps, can be identified by their sizes and manipulated. One application of this tool would be an analysis for locating helicopter landing zones that require at least 250 contiguous pixels at 10-meter resolution.

Contiguity analysis can be used to:

- divide a large class into separate raster regions, or
- eliminate raster regions that are too small to be considered for an application.

Filtering Clumps

In cases where very small clumps are not useful, they can be filtered out according to their sizes. This is sometimes referred to as eliminating the salt and pepper effects, or sieving. In Figure 11-5, all of the small clumps in the original (clumped) layer are eliminated.

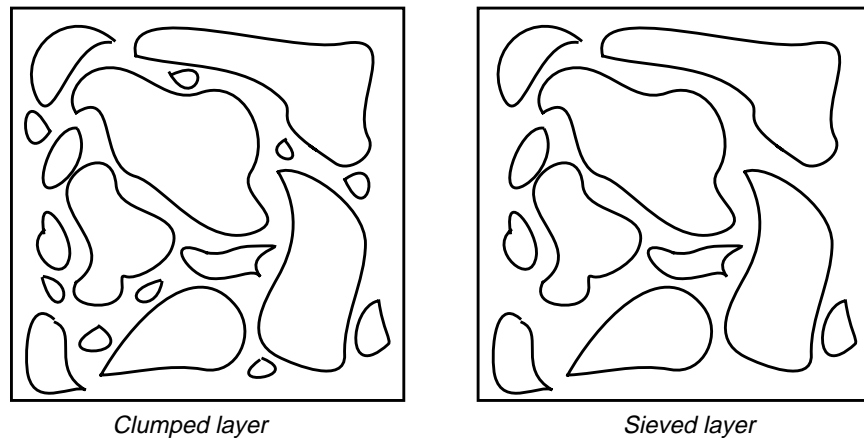


Figure 11-5 Contiguity Analysis



Use the Clump and Sieve (GIS Analysis) function in Image Interpreter or Spatial Modeler to perform contiguity analysis.

Neighborhood Analysis

With a process similar to the convolution filtering of continuous raster layers, thematic raster layers can also be filtered. The GIS filtering process is sometimes referred to as scanning, but is not to be confused with data capture via a digital camera. Neighborhood analysis is based on local or neighborhood characteristics of the data (Star and Estes 1990).

Every pixel is analyzed spatially, according to the pixels that surround it. The number and the location of the surrounding pixels is determined by a scanning window, which is defined by you. These operations are known as focal operations. The scanning window can be of any size in SML. In Model Maker, it has the following constraints:

- circular, with a maximum diameter of 512 pixels
- doughnut-shaped, with a maximum outer radius of 256
- rectangular, up to 512×512 pixels, with the option to mask-out certain pixels



Use the Neighborhood (GIS Analysis) function in Image Interpreter or Spatial Modeler to perform neighborhood analysis. The scanning window used in Image Interpreter can be 3×3 , 5×5 , or 7×7 . The scanning window in Model Maker is defined by you and can be up to 512×512 . The scanning window in SML can be of any size.

Defining Scan Area

You may define the area of the file to be scanned. The scanning window moves only through this area as the analysis is performed. The area may be defined in one or all of the following ways:

- Specify a rectangular portion of the file to scan. The output layer contains only the specified area.
- Specify an area that is defined by an existing AOI layer, an annotation overlay, or a vector layer. The area(s) within the polygon are scanned, and the other areas remain the same. The output layer is the same size as the input layer or the selected rectangular portion.
- Specify a class or classes in another thematic layer to be used as a mask. The pixels in the scanned layer that correspond to the pixels of the selected class or classes in the mask layer are scanned, while the other pixels remain the same.

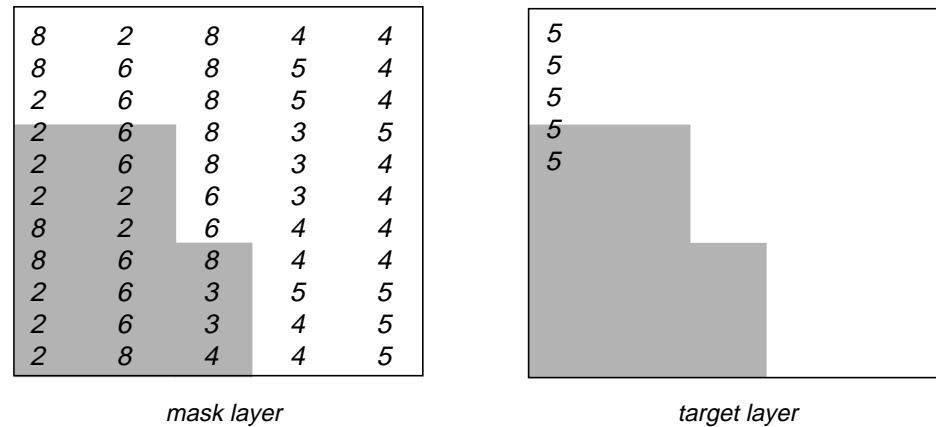


Figure 11-6 Using a Mask

In Figure 11-6, class 2 in the mask layer was selected for the mask. Only the corresponding (shaded) pixels in the target layer are scanned—the other values remain unchanged.

Neighborhood analysis creates a new thematic layer. There are several types of analysis that can be performed upon each window of pixels, as described below:

- **Boundary**—detects boundaries between classes. The output layer contains only boundary pixels. This is useful for creating boundary or edge lines from classes, such as a land/water interface.
- **Density**—outputs the number of pixels that have the same class value as the center (analyzed) pixel. This is also a measure of homogeneity (sameness), based upon the analyzed pixel. This is often useful in assessing vegetation crown closure.
- **Diversity**—outputs the number of class values that are present within the window. Diversity is also a measure of heterogeneity (difference).
- **Majority**—outputs the class value that represents the majority of the class values in the window. The value is defined by you. This option operates like a low-frequency filter to clean up a salt and pepper layer.
- **Maximum**—outputs the greatest class value within the window. This can be used to emphasize classes with the higher class values or to eliminate linear features or boundaries.
- **Mean**—averages the class values. If class values represent quantitative data, then this option can work like a convolution filter. This is mostly used on ordinal or interval data.
- **Median**—outputs the statistical median of the class values in the window. This option may be useful if class values represent quantitative data.
- **Minimum**—outputs the least or smallest class value within the window. The value is defined by you. This can be used to emphasize classes with the low class values.

- **Minority**—outputs the least common of the class values that are within the window. This option can be used to identify the least common classes. It can also be used to highlight disconnected linear features.
- **Rank**—outputs the number of pixels in the scan window whose value is less than the center pixel.
- **Standard deviation**—outputs the standard deviation of class values in the window.
- **Sum**—totals the class values. In a file where class values are ranked, totaling enables you to further rank pixels based on their proximity to high-ranking pixels.

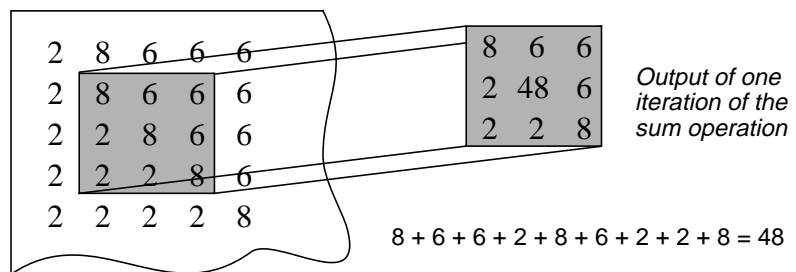


Figure 11-7 Sum Option of Neighborhood Analysis (Image Interpreter)

In Figure 11-7, the Sum option of Neighborhood (Image Interpreter) is applied to a 3×3 window of pixels in the input layer. In the output layer, the analyzed pixel is given a value based on the total of all of the pixels in the window.



The analyzed pixel is always the center pixel of the scanning window. In this example, only the pixel in the third column and third row of the file is summed.

Recoding

Class values can be recoded to new values. Recoding involves the assignment of new values to one or more classes. Recoding is used to:

- reduce the number of classes
- combine classes
- assign different class values to existing classes

When an ordinal, ratio, or interval class numbering system is used, recoding can be used to assign classes to appropriate values. Recoding is often performed to make later steps easier. For example, in creating a model that outputs good, better, and best areas, it may be beneficial to recode the input layers so all of the best classes have the highest class values.

In the following example (Table 11-1), a land cover layer is recoded so that the most environmentally sensitive areas (Riparian and Wetlands) have higher class values.

Table 11-1 Example of a Recoded Land Cover Layer

Value	New Value	Class Name
0	0	Background
1	4	Riparian
2	1	Grassland and Scrub
3	1	Chaparral
4	4	Wetlands
5	1	Emergent Vegetation
6	1	Water



Use the Recode (GIS Analysis) function in Image Interpreter or Spatial Modeler to recode layers.

Overlaying

Thematic data layers can be overlaid to create a composite layer. The output layer contains either the minimum or the maximum class values of the input layers. For example, if an area was in class 5 in one layer, and in class 3 in another, and the maximum class value dominated, then the same area would be coded to class 5 in the output layer, as shown in Figure 11-8.

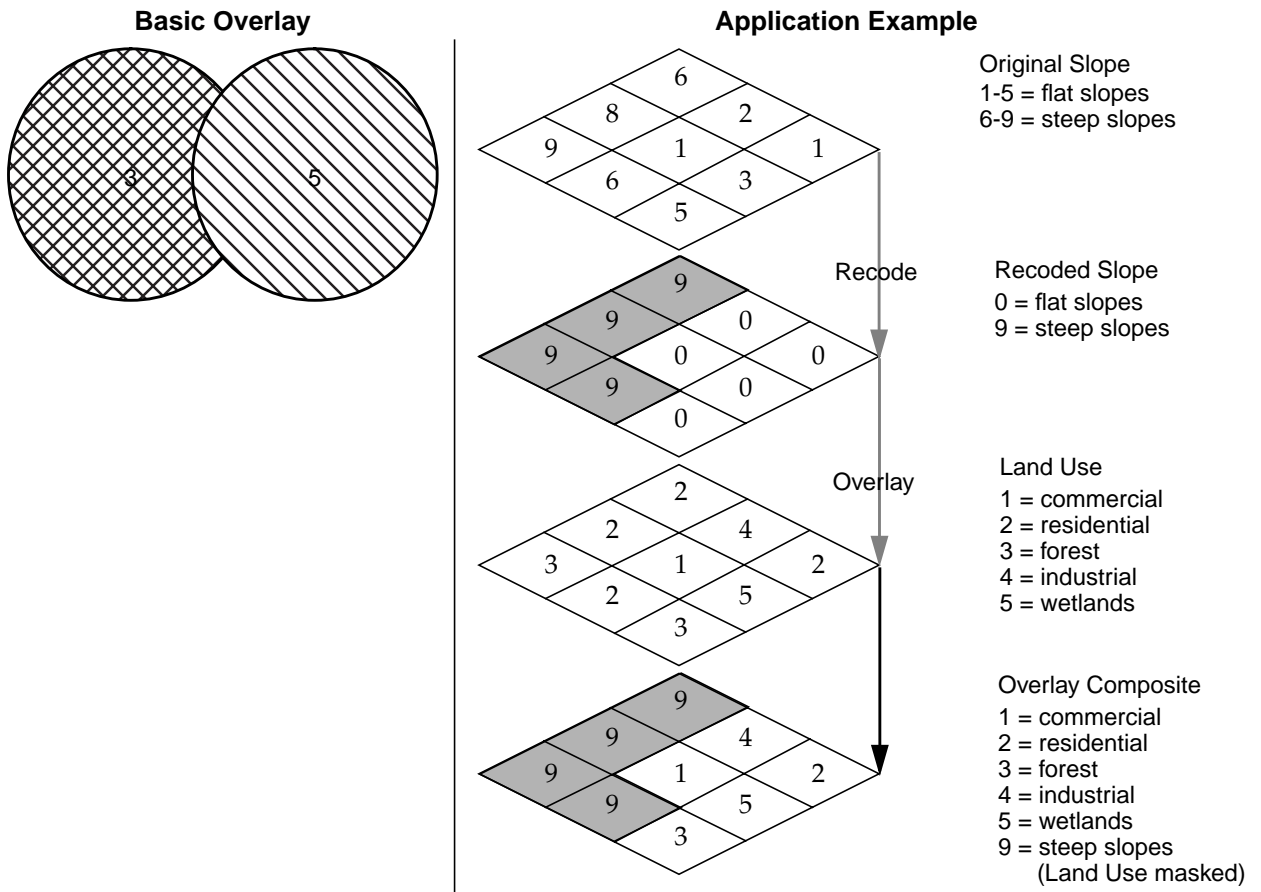


Figure 11-8 Overlay

The application example in Figure 11-8 shows the result of combining two layers—slope and land use. The slope layer is first recoded to combine all steep slopes into one value. When overlaid with the land use layer, the highest data file values (the steep slopes) dominate in the output layer.



Use the *Overlay (GIS Analysis)* function in *Image Interpreter* or *Spatial Modeler* to overlay layers.

Indexing

Thematic layers can be indexed (added) to create a composite layer. The output layer contains the sums of the input layer values. For example, the intersection of class 3 in one layer and class 5 in another would result in class 8 in the output layer, as shown in Figure 11-9.

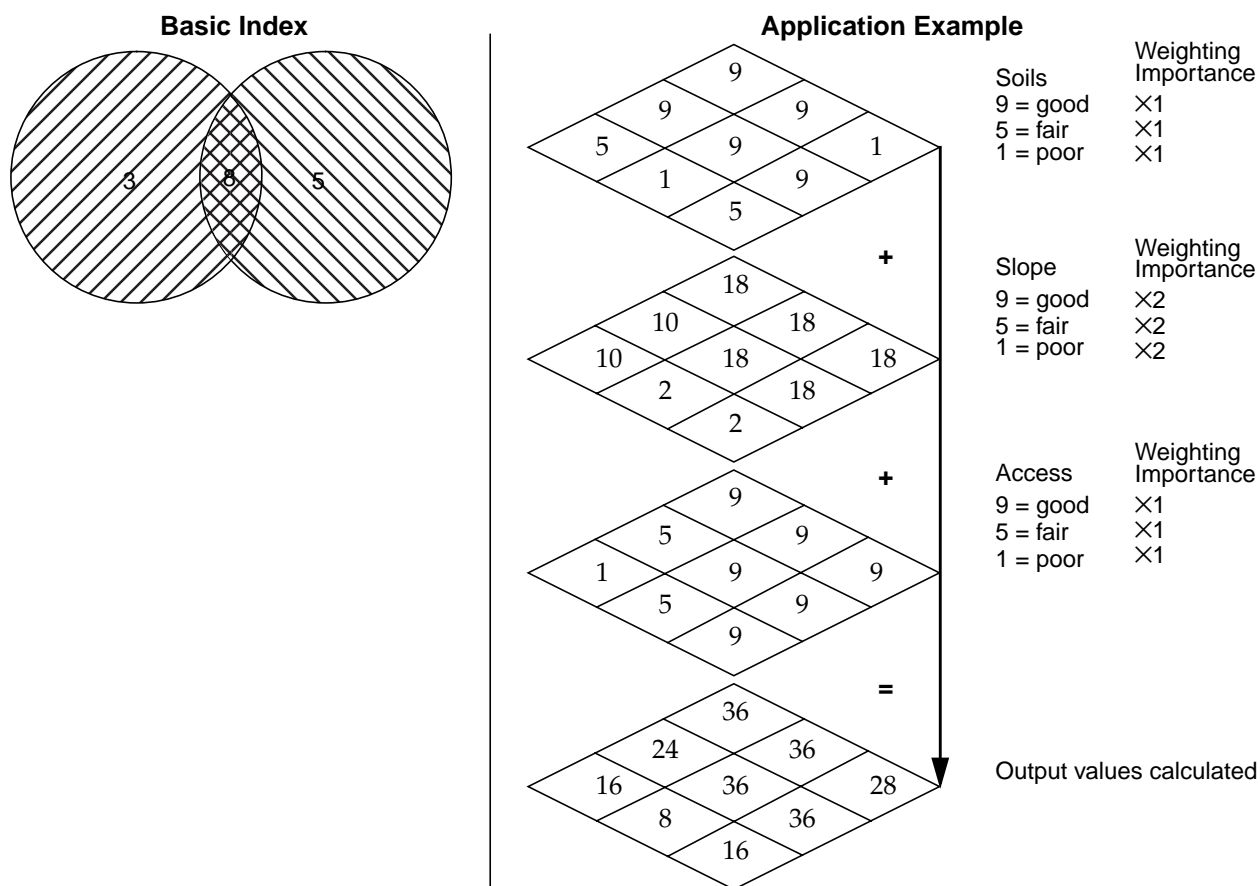


Figure 11-9 Indexing

The application example in Figure 11-9 shows the result of indexing. In this example, you want to develop a new subdivision, and the most likely sites are where there is the best combination (highest value) of good soils, good slope, and good access. Because good slope is a more critical factor to you than good soils or good access, a weighting factor is applied to the slope layer. A weighting factor has the effect of multiplying all input values by some constant. In this example, slope is given a weight of 2.



Use the *Index (GIS Analysis)* function in the *Image Interpreter* or *Spatial Modeler* to index layers.

Matrix Analysis

Matrix analysis produces a thematic layer that contains a separate class for every coincidence of classes in two layers. The output is best described with a matrix diagram.

		input layer 2 data values (columns)					
		0	1	2	3	4	5
input layer 1 data values (rows)	0	0	0	0	0	0	0
	1	0	1	2	3	4	5
	2	0	6	7	8	9	10
	3	0	11	12	13	14	15

In this diagram, the classes of the two input layers represent the rows and columns of the matrix. The output classes are assigned according to the coincidence of any two input classes.



All combinations of 0 and any other class are coded to 0, because 0 is usually the background class, representing an area that is not being studied.

Unlike overlaying or indexing, the resulting class values of a matrix operation are unique for each coincidence of two input class values. In this example, the output class value at column 1, row 3 is 11, and the output class at column 3, row 1 is 3. If these files were indexed (summed) instead of matrixed, both combinations would be coded to class 4.



Use the Matrix (GIS Analysis) function in Image Interpreter or Spatial Modeler to matrix layers.

Modeling

Modeling is a powerful and flexible analysis tool. Modeling is the process of creating new layers from combining or operating upon existing layers. Modeling enables you to create a small set of layers—perhaps even a single layer—which, at a glance, contains many types of information about the study area.

For example, if you want to find the best areas for a bird sanctuary, taking into account vegetation, availability of water, climate, and distance from highly developed areas, you would create a thematic layer for each of these criteria. Then, each of these layers would be input to a model. The modeling process would create one thematic layer, showing only the best areas for the sanctuary.

The set of procedures that define the criteria is called a model. In ERDAS IMAGINE, models can be created graphically and resemble a flow chart of steps, or they can be created using a script language. Although these two types of models look different, they are essentially the same—input files are defined, functions and/or operators are specified, and outputs are defined. The model is run and a new output layer(s) is created. Models can utilize analysis functions that have been previously defined, or new functions can be created by you.



Use the Model Maker function in Spatial Modeler to create graphical models and SML to create script models.

Data Layers

In modeling, the concept of layers is especially important. Before computers were used for modeling, the most widely used approach was to overlay registered maps on paper or transparencies, with each map corresponding to a separate theme. Today, digital files replace these hardcopy layers and allow much more flexibility for recoloring, recoding, and reproducing geographical information (Steinitz, Parker, and Jordan 1976).

In a model, the corresponding pixels at the same coordinates in all input layers are addressed as if they were physically overlaid like hardcopy maps.

Graphical Modeling

Graphical modeling enables you to draw models using a palette of tools that defines inputs, functions, and outputs. This type of modeling is very similar to drawing flowcharts, in that you identify a logical flow of steps needed to perform the desired action. Through the extensive functions and operators available in the ERDAS IMAGINE graphical modeling program, you can analyze many layers of data in very few steps without creating intermediate files that occupy extra disk space. Modeling is performed using a graphical editor that eliminates the need to learn a programming language. Complex models can be developed easily and then quickly edited and re-run on another data set.



Use the Model Maker function in Spatial Modeler to create graphical models.

Image Processing and GIS

In ERDAS IMAGINE, the traditional GIS functions (e.g., neighborhood analysis, proximity analysis, recode, overlay, index, etc.) can be performed in models, as well as image processing functions. Both thematic and continuous layers can be input into models that accomplish many objectives at once.

For example, suppose there is a need to assess the environmental sensitivity of an area for development. An output layer can be created that ranks most to least sensitive regions based on several factors, such as slope, land cover, and floodplain. To visualize the location of these areas, the output thematic layer can be overlaid onto a high resolution, continuous raster layer (e.g., SPOT panchromatic) that has had a convolution filter applied. All of this can be accomplished in a single model (as shown in Figure 11-10).

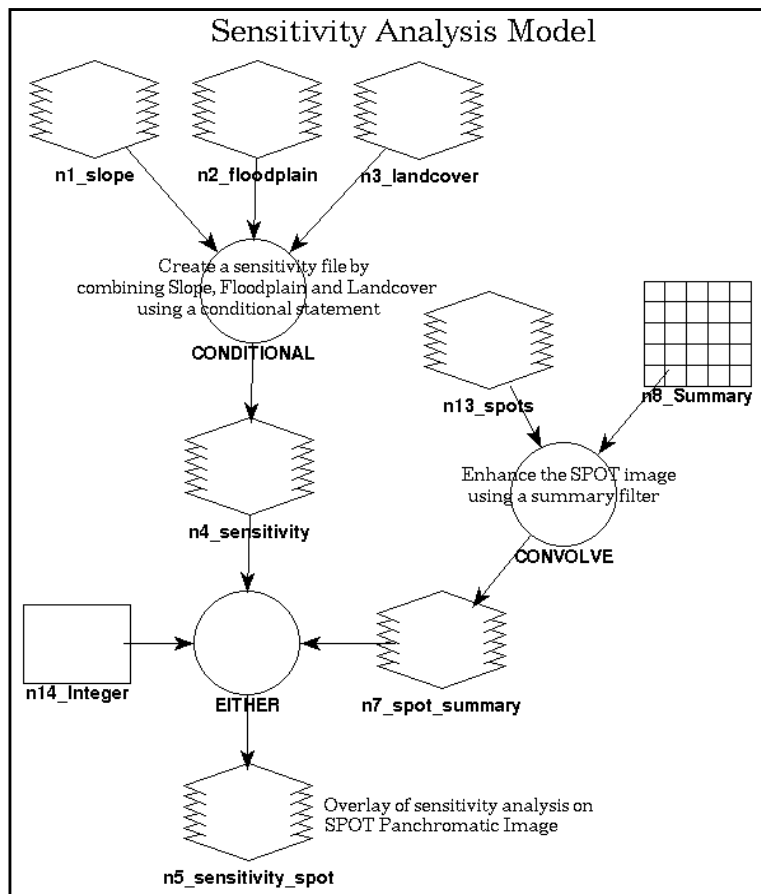


Figure 11-10 Graphical Model for Sensitivity Analysis



See the ERDAS IMAGINE Tour Guides™ manual for step-by-step instructions on creating the environmental sensitivity model in Figure 11-10. Descriptions of all of the graphical models delivered with ERDAS IMAGINE are available in the On-Line Help.

Model Structure

A model created with Model Maker is essentially a flow chart that defines:

- the input image(s), vector(s), matrix(ces), table(s), and scalar(s) to be analyzed
- calculations, functions, or operations to be performed on the input data

- the output image(s), matrix(ces), table(s), and scalars to be created

The graphical models created in Model Maker all have the same basic structure: input, function, output. The number of inputs, functions, and outputs can vary, but the overall form remains constant. All components must be connected to one another before the model can be executed. The model on the left in Figure 11-11 is the most basic form. The model on the right is more complex, but it retains the same input/function/output flow.

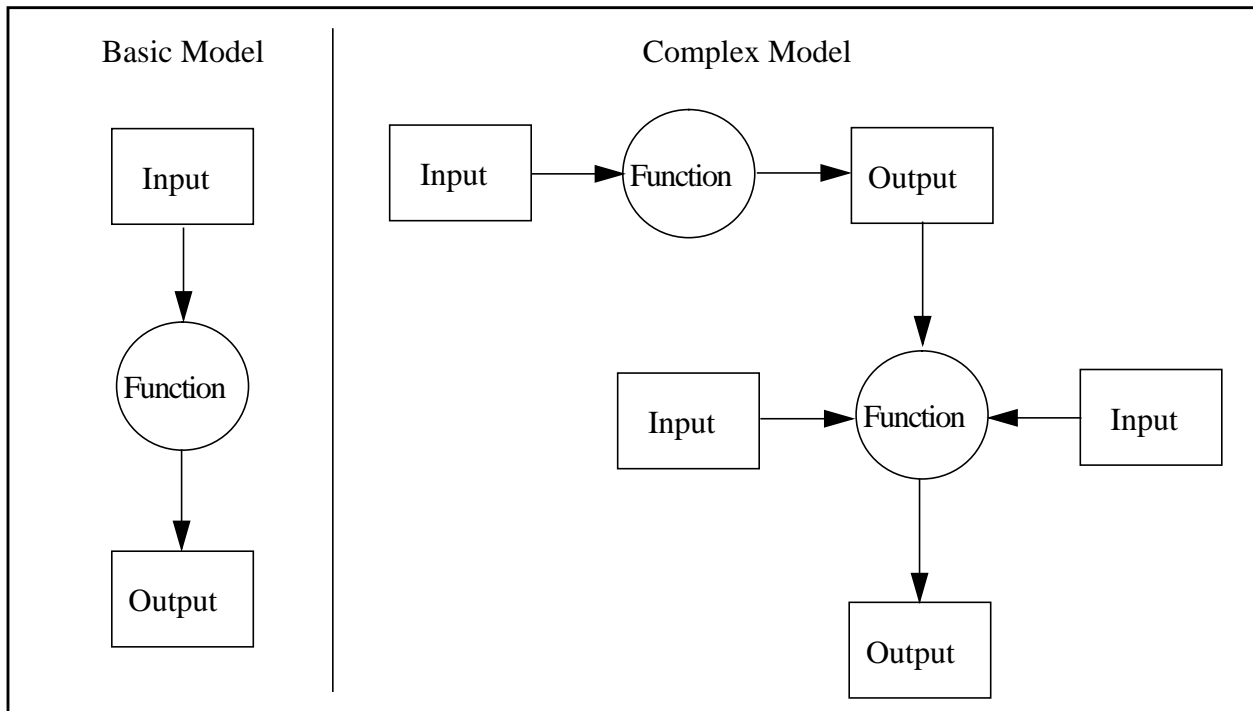


Figure 11-11 Graphical Model Structure

Graphical models are stored in ASCII files with the .gmd extension. There are several sample graphical models delivered with ERDAS IMAGINE that can be used as is or edited for more customized processing.



See the On-Line Help for instructions on editing existing models.

Model Maker Functions

The functions available in Model Maker are divided into the following categories:

Table 11-2 Model Maker Functions

Category	Description
Analysis	Includes convolution filtering, histogram matching, contrast stretch, principal components, and more.
Arithmetic	Perform basic arithmetic functions including addition, subtraction, multiplication, division, factorial, and modulus.
Bitwise	Use bitwise and, or, exclusive or, and not.
Boolean	Perform logical functions including and, or, and not.
Color	Manipulate colors to and from RGB (red, green, blue) and IHS (intensity, hue, saturation).
Conditional	Run logical tests using conditional statements and either...if...or...otherwise.
Data Generation	Create raster layers from map coordinates, column numbers, or row numbers. Create a matrix or table from a list of scalars.
Descriptor	Read attribute information and map a raster through an attribute column.
Distance	Perform distance functions, including proximity analysis.
Exponential	Use exponential operators, including natural and common logarithmic, power, and square root.
Focal (Scan)	Perform neighborhood analysis functions, including boundary, density, diversity, majority, mean, minority, rank, standard deviation, sum, and others.
Focal Use Opts	Constraints on which pixel values to include in calculations for the Focal (Scan) function.
Focal Apply Opts	Constraints on which pixel values to apply the results of calculations for the Focal (Scan) function.
Global	Analyze an entire layer and output one value, such as diversity, maximum, mean, minimum, standard deviation, sum, and more.
Matrix	Multiply, divide, and transpose matrices, as well as convert a matrix to a table and vice versa.
Other	Includes over 20 miscellaneous functions for data type conversion, various tests, and other utilities.
Relational	Includes equality, inequality, greater than, less than, greater than or equal, less than or equal, and others.
Size	Measure cell X and Y size, layer width and height, number of rows and columns, etc.
Stack Statistics	Perform operations over a stack of layers including diversity, majority, max, mean, median, min, minority, standard deviation, and sum.
Statistical	Includes density, diversity, majority, mean, rank, standard deviation, and more.
String	Manipulate character strings.

Table 11-2 Model Maker Functions (Continued)

Category	Description
Surface	Calculate aspect and degree/percent slope and produce shaded relief.
Trigonometric	Use common trigonometric functions, including sine/arcsine, cosine/arccosine, tangent/arctangent, hyperbolic arcsine, arccosine, cosine, sine, and tangent.
Zonal	Perform zonal operations including summary, diversity, majority, max, mean, min, range, and standard deviation.



These functions are also available for script modeling.



See the ERDAS IMAGINE Tour Guides and the On-Line SML manual for complete instructions on using Model Maker, and more detailed information about the available functions and operators.

Objects

Within Model Maker, an object is an input to or output from a function. The five basic object types used in Model Maker are:

- raster
- vector
- matrix
- table
- scalar

Raster

A raster object is a single layer or multilayer array of pixel data. Rasters are typically used to specify and manipulate data from image files.

Vector

Vector data in either a vector coverage, shapefile, or annotation layer can be read directly into the Model Maker, converted from vector to raster, then processed similarly to raster data; Model Maker cannot write to coverages, or shapefiles or annotation layers.

Matrix

A matrix object is a set of numbers arranged in a two-dimensional array. A matrix has a fixed number of rows and columns. Matrices may be used to store convolution kernels or the neighborhood definition used in neighborhood functions. They can also be used to store covariance matrices, eigenvector matrices, or matrices of linear combination coefficients.

Table

A table object is a series of numeric values, colors, or character strings. A table has one column and a fixed number of rows. Tables are typically used to store columns from the Raster Attribute Editor or a list of values that pertains to the individual layers of a set of layers. For example, a table with four rows could be used to store the maximum value from each layer of a four layer image file. A table may consist of up to 32,767 rows. Information in the table can be attributes, calculated (e.g., histograms), or defined by you.

Scalar

A scalar object is a single numeric value, color, or character string. Scalars are often used as weighting factors.

The graphics used in Model Maker to represent each of these objects are shown in Figure 11-12.

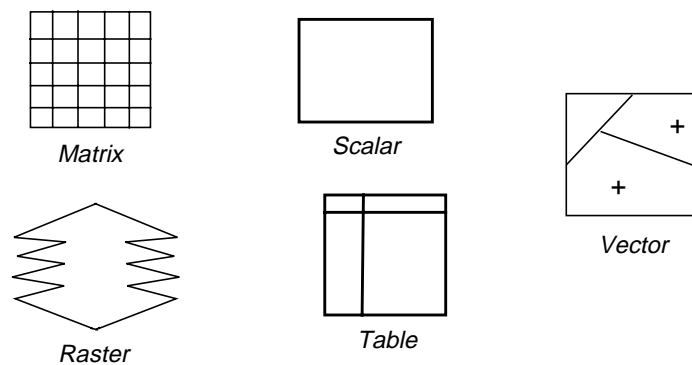


Figure 11-12 Modeling Objects

Data Types

The five object types described above may be any of the following data types:

- Binary—either 0 (false) or 1 (true)
- Integer—integer values from -2,147,483,648 to 2,147,483,648 (signed 32-bit integer)
- Float—floating point data (double precision)
- String—a character string (for table objects only)

Input and output data types do not have to be the same. Using SML, you can change the data type of input files before they are processed.

Output Parameters

Since it is possible to have several inputs in one model, you can optionally define the working window and the pixel cell size of the output data along with the output map projection.

Working Window

Raster layers of differing areas can be input into one model. However, the image area, or working window, must be specified in order to use it in the model calculations. Either of the following options can be selected:

- Union—the model operates on the union of all input rasters. (This is the default.)
- Intersection—the model uses only the area of the rasters that is common to all input rasters.

Pixel Cell Size

Input rasters may also be of differing resolution (pixel size), so you must select the output cell size as either:

- Minimum—the minimum cell size of the input layers is used (this is the default setting).
- Maximum—the maximum cell size of the input layers is used.
- Other—specify a new cell size.

Map Projection

The output map projection defaults to be the same as the first input, or projection may be selected to be the same as a chosen input. The output projection may also be selected from a projection library.

Using Attributes in Models

With the criteria function in Model Maker, attribute data can be used to determine output values. The criteria function simplifies the process of creating a conditional statement. The criteria function can be used to build a table of conditions that must be satisfied to output a particular row value for an attribute (or cell value) associated with the selected raster.

The inputs to a criteria function are rasters or vectors. The columns of the criteria table represent either attributes associated with a raster layer or the layer itself, if the cell values are of direct interest. Criteria which must be met for each output column are entered in a cell in that column (e.g., >5). Multiple sets of criteria may be entered in multiple rows. The output raster contains the first row number of a set of criteria that were met for a raster cell.

Example

For example, consider the sample thematic layer, parks.img, that contains the following attribute information:

Table 11-3 Attribute Information for parks.img

Class Name	Histogram	Acres	Path Condition	Turf Condition	Car Spaces
Grant Park	2456	403.45	Fair	Good	127
Piedmont Park	5167	547.88	Good	Fair	94
Candler Park	763	128.90	Excellent	Excellent	65
Springdale Park	548	46.33	None	Excellent	0

A simple model could create one output layer that shows only the parks in need of repairs. The following logic would therefore be coded into the model:

“If Turf Condition is not Good or Excellent, and if Path Condition is not Good or Excellent, then the output class value is 1. Otherwise, the output class value is 2.”

More than one input layer can also be used. For example, a model could be created, using the input layers parks.img and soils.img, that shows the soil types for parks with either fair or poor turf condition. Attributes can be used from every input file.

The following is a slightly more complex example:

If you have a land cover file and you want to create a file of pine forests larger than 10 acres, the criteria function could be used to output values only for areas that satisfy the conditions of being both pine forest and larger than 10 acres. The output file would have two classes: pine forests larger than 10 acres and background. If you want the output file to show varying sizes of pine forest, you would simply add more conditions to the criteria table.

Comparisons of attributes can also be combined with mathematical and logical functions on the class values of the input file(s). With these capabilities, highly complex models can be created.



See the ERDAS IMAGINE Tour Guides or the On-Line Help for specific instructions on using the criteria function.

Script Modeling

SML is a script language used internally by Model Maker to execute the operations specified in the graphical models that are created. SML can also be used to directly write to models you create. It includes all of the functions available in Model Maker, plus:

- conditional branching and looping
- the ability to use complex data types

Graphical models created with Model Maker can be output to a script file (text only) in SML. These scripts can then be edited with a text editor using SML syntax and rerun or saved in a library. Script models can also be written from scratch in the text editor. They are stored in ASCII .mdl files.



The Text Editor is available from the Tools menu located on the ERDAS IMAGINE menu bar and from the Model Librarian (Spatial Modeler).

In Figure 11-13, both the graphical and script models are shown for a tasseled cap transformation. Notice how even the annotation on the graphical model is included in the automatically generated script model. Generating script models from graphical models may aid in learning SML.

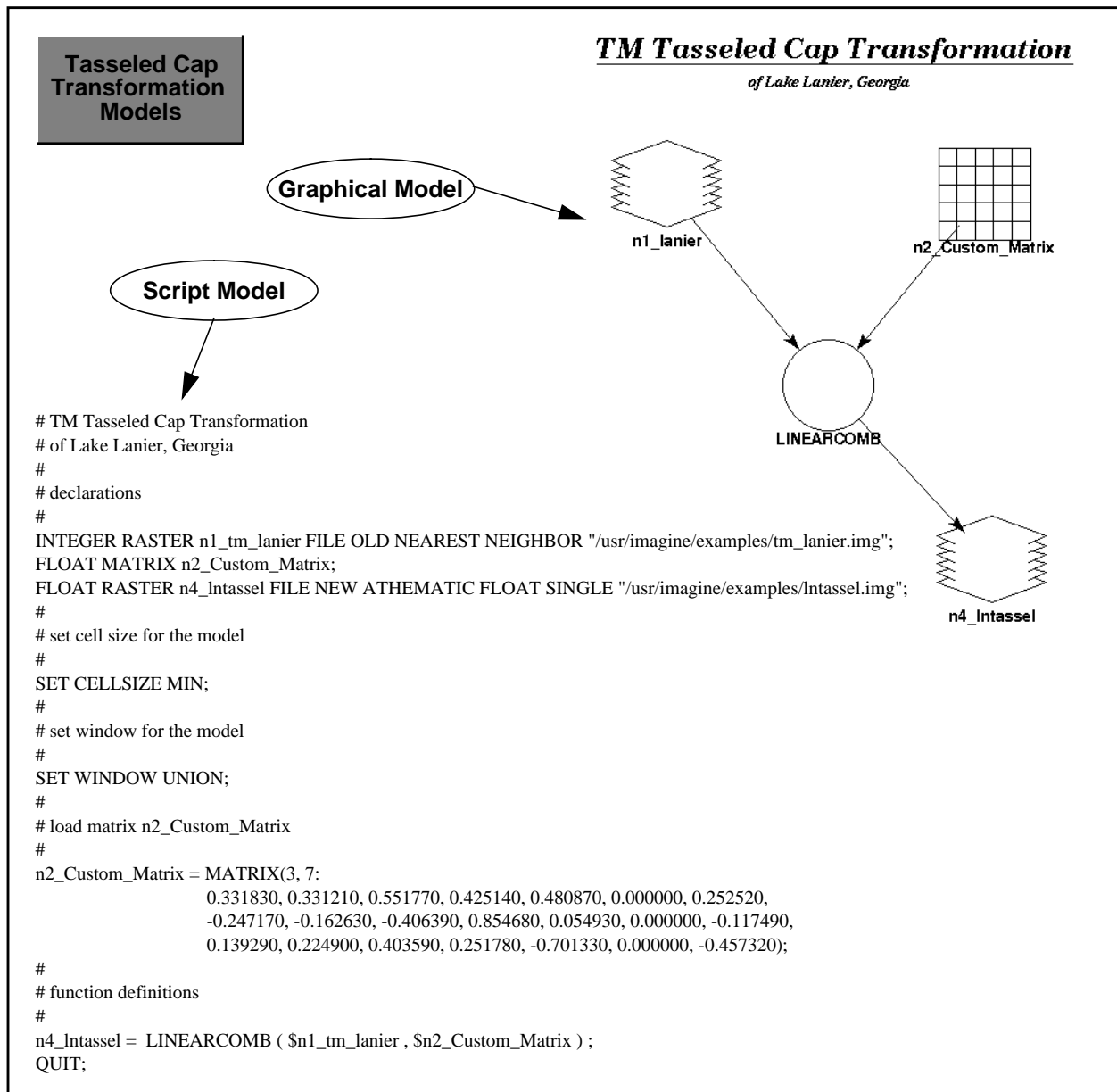


Figure 11-13 Graphical and Script Models For Tasseled Cap Transformation



Convert graphical models to scripts using Model Maker. Open existing script models from the Model Librarian (Spatial Modeler).

Statements

A script model consists primarily of one or more statements. Each statement falls into one of the following categories:

- Declaration—defines objects to be manipulated within the model

- Assignment—assigns a value to an object
- Show and View—enables you to see and interpret results from the model
- Set—defines the scope of the model or establishes default values used by the Modeler
- Macro Definition—defines substitution text associated with a macro name
- Quit—ends execution of the model

SML also includes flow control structures so that you can utilize conditional branching and looping in the models and statement block structures, which cause a set of statements to be executed as a group.

Declaration Example

In the script model in Figure 11-13, the following lines form the declaration portion of the model:

```
INTEGER RASTER n1_tm_lanier FILE OLD NEAREST NEIGHBOR
"/usr/imagene/examples/tm_lanier.img";

FLOAT MATRIX n2_Custom_Matrix;

FLOAT RASTER n4_lntassel FILE NEW ATHEMATIC FLOAT SINGLE
"/usr/imagene/examples/lntassel.img";
```

Set Example

The following set statements are used:

```
SET CELLSIZE MIN;

SET WINDOW UNION;
```

Assignment Example

The following assignment statements are used:

```
n2_Custom_Matrix = MATRIX(3, 7:
0.331830, 0.331210, 0.551770, 0.425140, 0.480870,
0.000000, 0.252520,
-0.247170, -0.162630, -0.406390, 0.854680, 0.054930,
0.000000, -0.117490,
0.139290, 0.224900, 0.403590, 0.251780, -0.701330,
0.000000, -0.457320);

n4_lntassel = LINEARCOMB ( $n1_tm_lanier ,
$n2_Custom_Matrix ) ;
```

Data Types

In addition to the data types utilized by Graphical Modeling, script model objects can store data in the following data types:

- Complex—complex data (double precision)

- Color—three floating point numbers in the range of 0.0 to 1.0, representing intensity of red, green, and blue

Variables

Variables are objects in the Modeler that have been associated with a name using a declaration statement. The declaration statement defines the data type and object type of the variable. The declaration may also associate a raster variable with certain layers of an image file or a table variable with an attribute table. Assignment statements are used to set or change the value of a variable.



For script model syntax rules, descriptions of all available functions and operators, and sample models, see the On-Line SML manual.

Vector Analysis

Most of the operations discussed in the previous pages of this chapter focus on raster data. However, in a complete GIS database, both raster and vector layers are present. One of the most common applications involving the combination of raster and vector data is the updating of vector layers using current raster imagery as a backdrop for vector editing. For example, if a vector database is more than one or two years old, then there are probably errors due to changes in the area (new roads, moved roads, new development, etc.). When displaying existing vector layers over a raster layer, you can dynamically update the vector layer by digitizing new or changed features on the screen.

Vector layers can also be used to indicate an AOI for further processing. Assume you want to run a site suitability model on only areas designated for commercial development in the zoning ordinances. By selecting these zones in a vector polygon layer, you could restrict the model to only those areas in the raster input files.

Vector layers can also be used as inputs to models. Updated or new attributes may also be written to vector layers in models.

Editing Vector Layers

Editable features are polygons (as lines), lines, label points, and nodes. There can be multiple features selected with a mixture of any and all feature types. Editing operations and commands can be performed on multiple or single selections. In addition to the basic editing operations (e.g., cut, paste, copy, delete), you can also perform the following operations on the line features in multiple or single selections:

- spline—smooths or generalizes all currently selected lines using a specified grain tolerance
- generalize—weeds vertices from selected lines using a specified tolerance
- split/unsplit—makes two lines from one by adding a node or joins two lines by removing a node
- densify—adds vertices to selected lines at a tolerance you specify
- reshape (for single lines only)—enables you to move the vertices of a line

Reshaping (adding, deleting, or moving a vertex or node) can be done on a single selected line. Table 11-4 details general editing operations and the feature types that support each of those operations.

Table 11-4 General Editing Operations and Supporting Feature Types

	Add	Delete	Move	Reshape
Points	yes	yes	yes	no
Lines	yes	yes	yes	yes
Polygons	yes	yes	yes	no
Nodes	yes	yes	yes	no

The Undo utility may be applied to any edits. The software stores all edits in sequential order, so that continually pressing Undo reverses the editing.



For more information on vectors, see "CHAPTER 3: Raster and Vector Data Sources" on page 49.

Constructing Topology (Coverages Only)

Either the Build or Clean option can be used to construct topology. To create spatial relationships between features in a vector layer, it is necessary to create topology. After a vector layer is edited, the topology must be constructed to maintain the topological relationships between features. When topology is constructed, each feature is assigned an internal number. These numbers are then used to determine line connectivity and polygon contiguity. Once calculated, these values are recorded and stored in that layer's associated attribute table.



You must also reconstruct the topology of vector layers imported into ERDAS IMAGINE.

When topology is constructed, feature attribute tables are created with several automatically created fields. Different fields are stored for the different types of layers. The automatically generated fields for a line layer are:

- FNODE#—the internal node number for the beginning of a line (from-node)
- TNODE#—the internal number for the end of a line (to-node)
- LPOLY#—the internal number for the polygon to the left of the line (zero for layers containing only lines and no polygons)
- RPOLY#—the internal number for the polygon to the right of the line (zero for layers containing only lines and no polygons)
- LENGTH—length of each line, measured in layer units
- Cover#—internal line number (values assigned by ERDAS IMAGINE)
- Cover-ID—user-ID (values modified by you)

The automatically generated fields for a point or polygon layer are:

- AREA—area of each polygon, measured in layer units (zero for layers containing only points and no polygons)
- PERIMETER—length of each polygon boundary, measured in layer units (zero for layers containing only points and no polygons)
- Cover#—internal polygon number (values assigned by ERDAS IMAGINE)
- Cover-ID—user-ID (values modified by you)

Building and Cleaning Coverages

The Build option processes points, lines, and polygons, but the Clean option processes only lines and polygons. Build recognizes only existing intersections (nodes), whereas Clean creates intersections (nodes) wherever lines cross one another. The differences in these two options are summarized in Table 11-5 (ESRI 1990).

Table 11-5 Comparison of Building and Cleaning Coverages

Capabilities	Build	Clean
Processes:		
Polygons	Yes	Yes
Lines	Yes	Yes
Points	Yes	No
Numbers features	Yes	Yes
Calculates spatial measurements	Yes	Yes
Creates intersections	No	Yes
Processing speed	Faster	Slower

Errors

Constructing topology also helps to identify errors in the layer. Some of the common errors found are:

- Lines with less than two nodes
- Polygons that are not closed
- Polygons that have no label point or too many label points
- User-IDs that are not unique

Constructing topology can identify the errors mentioned above. When topology is constructed, line intersections are created, the lines that make up each polygon are identified, and a label point is associated with each polygon. Until topology is constructed, no polygons exist and lines that cross each other are not connected at a node, because there is no intersection.



Construct topology using the Vector Utilities menu from the Vector icon in the ERDAS IMAGINE icon panel.



You should not build or clean a layer that is displayed in a Viewer, nor should you try to display a layer that is being built or cleaned.

When the Build or Clean options are used to construct the topology of a vector layer, potential node errors are marked with special symbols. These symbols are listed below (ESRI 1990).

◇ Pseudo nodes, drawn with a diamond symbol, occur where a single line connects with itself (an island) or where only two lines intersect. Pseudo nodes do not necessarily indicate an error or a problem. Acceptable pseudo nodes may represent an island (a spatial pseudo node) or the point where a road changes from pavement to gravel (an attribute pseudo node).

□ A dangling node, represented by a square symbol, refers to the unconstructed node of a dangling line. Every line begins and ends at a node point. So if a line does not close properly, or was digitized past an intersection, it registers as a dangling node. In some cases, a dangling node may be acceptable. For example, in a street centerline map, cul-de-sacs are often represented by dangling nodes.

In polygon layers there may be label errors—usually no label point for a polygon, or more than one label point for a polygon. In the latter case, two or more points may have been mistakenly digitized for a polygon, or it may be that a line does not intersect another line, resulting in an open polygon.

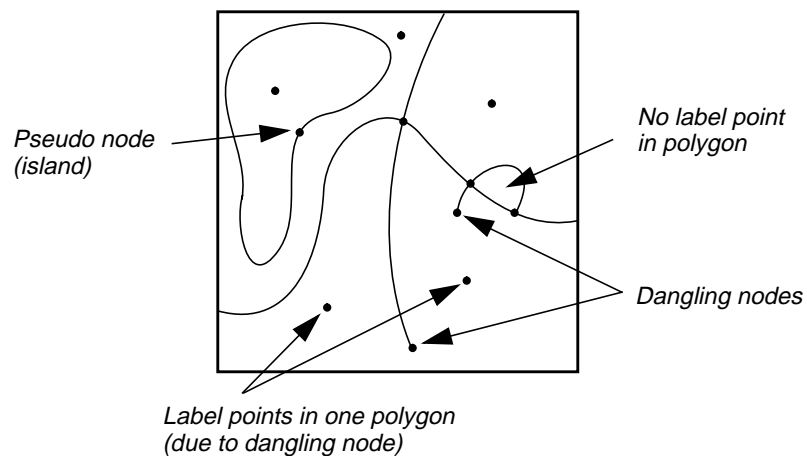


Figure 11-14 Layer Errors

Errors detected in a layer can be corrected by changing the tolerances set for that layer and building or cleaning again, or by editing the layer manually, then running Build or Clean.



Refer to the ERDAS IMAGINE Tour Guides manual for step-by-step instructions on editing vector layers.

CHAPTER 12

Cartography

Introduction

Maps and mapping are the subject of the art and science known as cartography—creating two-dimensional representations of our three-dimensional Earth. These representations were once hand-drawn with paper and pen. But now, map production is largely automated—and the final output is not always paper. The capabilities of a computer system are invaluable to map users, who often need to know much more about an area than can be reproduced on paper, no matter how large that piece of paper is or how small the annotation is. Maps stored on a computer can be queried, analyzed, and updated quickly.

As the veteran GIS and image processing authority, Roger F. Tomlinson, said: “Mapped and related statistical data do form the greatest storehouse of knowledge about the condition of the living space of mankind.” With this thought in mind, it only makes sense that maps be created as accurately as possible and be as accessible as possible.

In the past, map making was carried out by mapping agencies who took the analyst’s (be they surveyors, photogrammetrists, or draftsmen) information and created a map to illustrate that information. But today, in many cases, the analyst is the cartographer and can design his maps to best suit the data and the end user.

This chapter defines some basic cartographic terms and explains how maps are created within the ERDAS IMAGINE environment.



Use the Map Composer to create hardcopy and softcopy maps and presentation graphics.



This chapter concentrates on the production of digital maps. See "CHAPTER 13: Hardcopy Output" for information about printing hardcopy maps.

Types of Maps

A map is a graphic representation of spatial relationships on the Earth or other planets. Maps can take on many forms and sizes, depending on the intended use of the map. Maps no longer refer only to hardcopy output. In this manual, the maps discussed begin as digital files and may be printed later as desired.

Some of the different types of maps are defined in Table 12-1, “Types of Maps,” on page 420.

Table 12-1 Types of Maps

Map	Purpose
Aspect	A map that shows the prevailing direction that a slope faces at each pixel. Aspect maps are often color-coded to show the eight major compass directions, or any of 360 degrees.
Base	A map portraying background reference information onto which other information is placed. Base maps usually show the location and extent of natural Earth surface features and permanent human-made objects. Raster imagery, orthophotos, and orthoimages are often used as base maps.
Bathymetric	A map portraying the shape of a water body or reservoir using isobaths (depth contours).
Cadastral	A map showing the boundaries of the subdivisions of land for purposes of describing and recording ownership or taxation.
Choropleth	A map portraying properties of a surface using area symbols. Area symbols usually represent categorized classes of the mapped phenomenon.
Composite	A map on which the combined information from different thematic maps is presented.
Contour	A map in which lines are used to connect points of equal elevation. Lines are often spaced in increments of ten or twenty feet or meters.
Derivative	A map created by altering, combining, or analyzing other maps.
Index	A reference map that outlines the mapped area, identifies all of the component maps for the area if several map sheets are required, and identifies all adjacent map sheets.
Inset	A map that is an enlargement of some congested area of a smaller scale map, and that is usually placed on the same sheet with the smaller scale main map.
Isarithmic	A map that uses isarithms (lines connecting points of the same value for any of the characteristics used in the representation of surfaces) to represent a statistical surface. Also called an isometric map.
Isopleth	A map on which isopleths (lines representing quantities that cannot exist at a point, such as population density) are used to represent some selected quantity.
Morphometric	A map representing morphological features of the Earth's surface.
Outline	A map showing the limits of a specific set of mapping entities, such as counties, NTS quads, etc. Outline maps usually contain a very small number of details over the desired boundaries with their descriptive codes.
Planimetric	A map showing only the horizontal position of geographic objects, without topographic features or elevation contours.
Relief	Any map that appears to be, or is, three-dimensional. Also called a shaded relief map.
Slope	A map that shows changes in elevation over distance. Slope maps are usually color-coded according to the steepness of the terrain at each pixel.

Map	Purpose
Thematic	A map illustrating the class characterizations of a particular spatial variable (e.g., soils, land cover, hydrology, etc.)
Topographic	A map depicting terrain relief.
Viewshed	A map showing only those areas visible (or invisible) from a specified point(s). Also called a line-of-sight map or a visibility map.



In ERDAS IMAGINE, maps are stored as a map file with a .map extension.

Thematic Maps

Thematic maps comprise a large portion of the maps that many organizations create. For this reason, this map type is explored in more detail.

Thematic maps may be subdivided into two groups:

- qualitative
- quantitative

A qualitative map shows the spatial distribution or location of a kind of nominal data. For example, a map showing corn fields in the United States would be a qualitative map. It would not show how much corn is produced in each location, or production relative to the other areas.

A quantitative map displays the spatial aspects of numerical data. A map showing corn production (volume) in each area would be a quantitative map. Quantitative maps show ordinal (less than/greater than) and interval/ratio (difference) scale data (Dent 1985).



You can create thematic data layers from continuous data (aerial photography and satellite images) using the ERDAS IMAGINE classification capabilities. See "CHAPTER 6: Classification" for more information.

Base Information

Thematic maps should include a base of information so that the reader can easily relate the thematic data to the real world. This base may be as simple as an outline of counties, states, or countries, to something more complex, such as an aerial photograph or satellite image. In the past, it was difficult and expensive to produce maps that included both thematic and continuous data, but technological advances have made this easy.

For example, in a thematic map showing flood plains in the Mississippi River valley, you could overlay the thematic data onto a line coverage of state borders or a satellite image of the area. The satellite image can provide more detail about the areas bordering the flood plains. This may be valuable information when planning emergency response and resource management efforts for the area. Satellite images can also provide very current information about an area, and can assist you in assessing the accuracy of a thematic image.



In ERDAS IMAGINE, you can include multiple layers in a single map composition. See “Map Composition” on page 447 for more information about creating maps.

Color Selection

The colors used in thematic maps may or may not have anything to do with the class or category of information shown. Cartographers usually try to use a color scheme that highlights the primary purpose of the map. The map reader’s perception of colors also plays an important role. Most people are more sensitive to red, followed by green, yellow, blue, and purple. Although color selection is left entirely up to the map designer, some guidelines have been established (Robinson and Sale 1969).

- When mapping interval or ordinal data, the higher ranks and greater amounts are generally represented by darker colors.
- Use blues for water.
- When mapping elevation data, start with blues for water, greens in the lowlands, ranging up through yellows and browns to reds in the higher elevations. This progression should not be used for series other than elevation.
- In temperature mapping, use red, orange, and yellow for warm temperatures and blue, green, and gray for cool temperatures.
- In land cover mapping, use yellows and tans for dryness and sparse vegetation and greens for lush vegetation.
- Use browns for land forms.



Use the Raster Attributes option in the Viewer to select and modify class colors.

Annotation

A map is more than just an image(s) on a background. Since a map is a form of communication, it must convey information that may not be obvious by looking at the image. Therefore, maps usually contain several annotation elements to explain the map. Annotation is any explanatory material that accompanies a map to denote graphical features on the map. This annotation may take the form of:

- scale bars
- legends
- neatlines, tick marks, and grid lines
- symbols (north arrows, etc.)
- labels (rivers, mountains, cities, etc.)
- descriptive text (title, copyright, credits, production notes, etc.)

The annotation listed above is made up of single elements. The basic annotation elements in ERDAS IMAGINE include:

- rectangles (including squares)
- ellipses (including circles)

- polygons and polylines
- text

These elements can be used to create more complex annotation, such as legends, scale bars, etc. These annotation components are actually groups of the basic elements and can be ungrouped and edited like any other graphic. You can also create your own groups to form symbols that are not in the ERDAS IMAGINE symbol library. (Symbols are discussed in more detail under "Symbols" on page 428.)



Create annotation using the Annotation tool palette in the Viewer or in a map composition.

How Annotation is Stored

An annotation layer is a set of annotation elements that is drawn in a Viewer or Map Composer window and stored in a file. Annotation that is created in a Viewer window is stored in a separate file from the other data in the Viewer. These annotation files are called overlay files (.ovr extension). Map annotation that is created in a Map Composer window is also stored in an .ovr file, which is named after the map composition. For example, the annotation for a file called lanier.map would be lanier.map.ovr.

Scale

Map scale is a statement that relates distance on a map to distance on the Earth's surface. It is perhaps the most important information on a map, since the level of detail and map accuracy are both factors of the map scale. Scale is directly related to the map extent, or the area of the Earth's surface to be mapped. If a relatively small area is to be mapped, such as a neighborhood or subdivision, then the scale can be larger. If a large area is to be mapped, such as an entire continent, the scale must be smaller. Generally, the smaller the scale, the less detailed the map can be. As a rule, anything smaller than 1:250,000 is considered small-scale.

Scale can be reported in several ways, including:

- representative fraction
- verbal statement
- scale bar

Representative Fraction

Map scale is often noted as a simple ratio or fraction called a representative fraction. A map in which one inch on the map equals 24,000 inches on the ground could be described as having a scale of 1:24,000 or 1/24,000. The units on both sides of the ratio must be the same.

Verbal Statement

A verbal statement of scale describes the distance on the map to the distance on the ground. A verbal statement describing a scale of 1:1,000,000 is approximately 1 inch to 16 miles. The units on the map and on the ground do not have to be the same in a verbal statement. One-inch and 6-inch maps of the British Ordnance Survey are often referred to by this method (1 inch to 1 mile, 6 inches to 1 mile) (Robinson and Sale 1969).

Scale Bars

A scale bar is a graphic annotation element that describes map scale. It shows the distance on paper that represents a geographical distance on the map. Maps often include more than one scale bar to indicate various measurement systems, such as kilometers and miles.

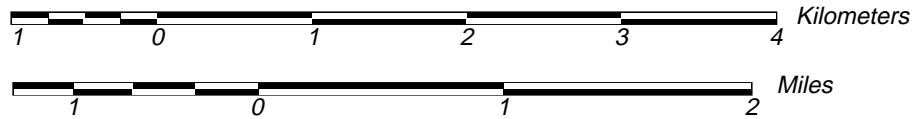


Figure 12-1 Sample Scale Bars



Use the Scale Bar tool in the Annotation tool palette to automatically create representative fractions and scale bars. Use the Text tool to create a verbal statement.

Common Map Scales

You can create maps with an unlimited number of scales, however, there are some commonly used scales. Table 12-2 lists these scales and their equivalents (Robinson and Sale 1969).

Table 12-2 Common Map Scales

Map Scale	1/40 inch represents	1 inch represents	1 centimeter represents	1 mile is represented by	1 kilometer is represented by
1:2,000	4.200 ft	56.000 yd	20.000 m	31.680 in	50.00 cm
1:5,000	10.425 ft	139.000 yd	50.000 m	12.670 in	20.00 cm
1:10,000	6.952 yd	0.158 mi	0.100 km	6.340 in	10.00 cm
1:15,840	11.000 yd	0.250 mi	0.156 km	4.000 in	6.25 cm
1:20,000	13.904 yd	0.316 mi	0.200 km	3.170 in	5.00 cm
1:24,000	16.676 yd	0.379 mi	0.240 km	2.640 in	4.17 cm
1:25,000	17.380 yd	0.395 mi	0.250 km	2.530 in	4.00 cm
1:31,680	22.000 yd	0.500 mi	0.317 km	2.000 in	3.16 cm
1:50,000	34.716 yd	0.789 mi	0.500 km	1.270 in	2.00 cm
1:62,500	43.384 yd	0.986 mi	0.625 km	1.014 in	1.60 cm
1:63,360	0.025 mi	1.000 mi	0.634 km	1.000 in	1.58 cm
1:75,000	0.030 mi	1.180 mi	0.750 km	0.845 in	1.33 cm
1:80,000	0.032 mi	1.260 mi	0.800 km	0.792 in	1.25 cm
1:100,000	0.040 mi	1.580 mi	1.000 km	0.634 in	1.00 cm

Table 12-2 Common Map Scales (Continued)

Map Scale	1/40 inch represents	1 inch represents	1 centimeter represents	1 mile is represented by	1 kilometer is represented by
1:125,000	0.050 mi	1.970 mi	1.250 km	0.507 in	8.00 mm
1:250,000	0.099 mi	3.950 mi	2.500 km	0.253 in	4.00 mm
1:500,000	0.197 mi	7.890 mi	5.000 km	0.127 in	2.00 mm
1:1,000,000	0.395 mi	15.780 mi	10.000 km	0.063 in	1.00 mm

Table 12-3 shows the number of pixels per inch for selected scales and pixel sizes.

Table 12-3 Pixels per Inch

Pixel Size (m)	SCALE							
	1"=100' 1:1200	1"=200' 1:2400	1"=500' 1:6000	1"=1000' 1:12000	1"=1500' 1:18000	1"=2000' 1:24000	1"=4167' 1:50000	1"=1 mile 1:63360
1	30.49	60.96	152.40	304.80	457.20	609.60	1270.00	1609.35
2	15.24	30.48	76.20	152.40	228.60	304.80	635.00	804.67
2.5	12.13	24.38	60.96	121.92	182.88	243.84	508.00	643.74
5	6.10	12.19	30.48	60.96	91.44	121.92	254.00	321.87
10	3.05	6.10	15.24	30.48	45.72	60.96	127.00	160.93
15	2.03	4.06	10.16	20.32	30.48	40.64	84.67	107.29
20	1.52	3.05	7.62	15.24	22.86	30.48	63.50	80.47
25	1.22	2.44	6.10	12.19	18.29	24.38	50.80	64.37
30	1.02	2.03	5.08	10.16	15.240	20.32	42.33	53.64
35	.87	1.74	4.35	8.71	13.08	17.42	36.29	45.98
40	.76	1.52	3.81	7.62	11.43	15.24	31.75	40.23
45	.68	1.35	3.39	6.77	10.16	13.55	28.22	35.76
50	.61	1.22	3.05	6.10	9.14	12.19	25.40	32.19
75	.41	.81	2.03	4.06	6.10	8.13	16.93	21.46
100	.30	.61	1.52	3.05	4.57	6.10	12.70	16.09
150	.20	.41	1.02	2.03	3.05	4.06	8.47	10.73
200	.15	.30	.76	1.52	2.29	3.05	6.35	8.05
250	.12	.24	.61	1.22	1.83	2.44	5.08	6.44
300	.10	.30	.51	1.02	1.52	2.03	4.23	5.36
350	.09	.17	.44	.87	1.31	1.74	3.63	4.60
400	.08	.15	.38	.76	1.14	1.52	3.18	4.02
450	.07	.14	.34	.68	1.02	1.35	2.82	3.58

Table 12-3 Pixels per Inch (Continued)

Pixel Size (m)	SCALE							
	1"=100' 1:1200	1"=200' 1:2400	1"=500' 1:6000	1"=1000' 1:12000	1"=1500' 1:18000	1"=2000' 1:24000	1"=4167' 1:50000	1"=1 mile 1:63360
500	.06	.12	.30	.61	.91	1.22	2.54	3.22
600	.05	.10	.25	.51	.76	1.02	2.12	2.69
700	.04	.09	.22	.44	.65	.87	1.81	2.30
800	.04	.08	.19	.38	.57	.76	1.59	2.01
900	.03	.07	.17	.34	.51	.68	1.41	1.79
1000	.03	.06	.15	.30	.46	.61	1.27	1.61

Courtesy of D. Cunningham and D. Way, The Ohio State University.
 Table 12-4 lists the number of acres and hectares per pixel for various pixel sizes.

Table 12-4 Acres and Hectares per Pixel

Pixel Size (m)	Acres	Hectares
1	0.0002	0.0001
2	0.0010	0.0004
2.5	0.0015	0.0006
5	0.0062	0.0025
10	0.0247	0.0100
15	0.0556	0.0225
20	0.0988	0.0400
25	0.1544	0.0625
30	0.2224	0.0900
35	0.3027	0.1225
40	0.3954	0.1600
45	0.5004	0.2025
50	0.6178	0.2500
75	1.3900	0.5625
100	2.4710	1.0000
150	5.5598	2.2500
200	9.8842	4.0000
250	15.4440	6.2500
300	22.2394	9.0000
350	30.2703	12.2500
400	39.5367	16.0000

Table 12-4 Acres and Hectares per Pixel (Continued)

Pixel Size (m)	Acres	Hectares
450	50.0386	20.2500
500	61.7761	25.0000
600	88.9576	36.0000
700	121.0812	49.0000
800	158.1468	64.0000
900	200.1546	81.0000
1000	247.1044	100.0000

Courtesy of D. Cunningham and D. Way, The Ohio State University.

Legends

A legend is a key to the colors, symbols, and line styles that are used in a map. Legends are especially useful for maps of categorical data displayed in pseudo color, where each color represents a different feature or category. A legend can also be created for a single layer of continuous data, displayed in gray scale. Legends are likewise used to describe all unknown or unique symbols utilized. Symbols in legends should appear exactly the same size and color as they appear on the map (Robinson and Sale 1969).

Legend





	pasture
	forest
	swamp
	developed

Figure 12-2 Sample Legend



Use the Legend tool in the Annotation tool palette to automatically create color legends. Symbol legends are not created automatically, but can be created manually.

Neatlines, Tick Marks, and Grid Lines

Neatlines, tick marks, and grid lines serve to provide a georeferencing system for map detail and are based on the map projection of the image shown.

- A neatline is a rectangular border around the image area of a map. It differs from the map border in that the border usually encloses the entire map, not just the image area.
- Tick marks are small lines along the edge of the image area or neatline that indicate regular intervals of distance.

- Grid lines are intersecting lines that indicate regular intervals of distance, based on a coordinate system. Usually, they are an extension of tick marks. It is often helpful to place grid lines over the image area of a map. This is becoming less common on thematic maps, but is really up to the map designer. If the grid lines help readers understand the content of the map, they should be used.

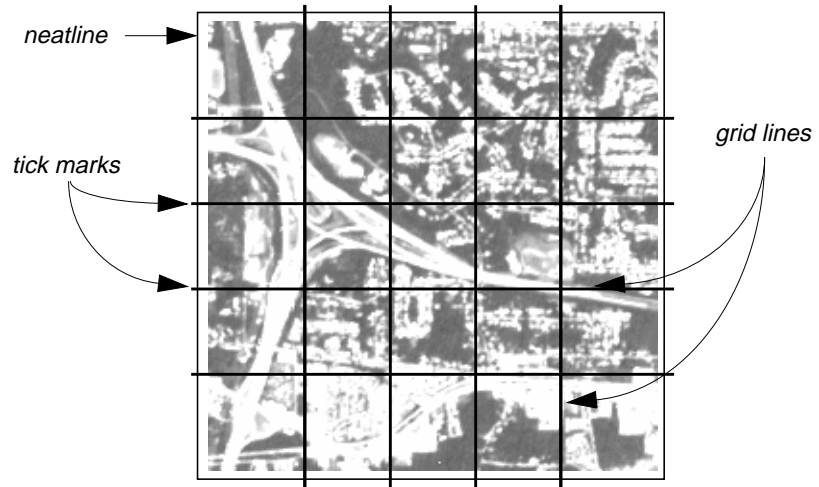


Figure 12-3 Sample Neatline, Tick Marks, and Grid Lines

Grid lines may also be referred to as a graticule.



Graticules are discussed in more detail in “Map Projections” on page 432.



Use the Grid/Tick tool in the Annotation tool palette to create neatlines, tick marks, and grid lines. Tick marks and grid lines can also be created over images displayed in a Viewer. See the On-Line Help for instructions.

Symbols

Since maps are a greatly reduced version of the real-world, objects cannot be depicted in their true shape or size. Therefore, a set of symbols is devised to represent real-world objects. There are two major classes of symbols:

- replicative
- abstract

Replicative symbols are designed to look like their real-world counterparts; they represent tangible objects, such as coastlines, trees, railroads, and houses. Abstract symbols usually take the form of geometric shapes, such as circles, squares, and triangles. They are traditionally used to represent amounts that vary from place to place, such as population density, amount of rainfall, etc. (Dent 1985).

Both replicative and abstract symbols are composed of one or more of the following annotation elements:

- point
- line
- area

Symbol Types

These basic elements can be combined to create three different types of replicative symbols:

- plan—formed after the basic outline of the object it represents. For example, the symbol for a house might be a square, because most houses are rectangular.
- profile—formed like the profile of an object. Profile symbols generally represent vertical objects, such as trees, windmills, oil wells, etc.
- function—formed after the activity that a symbol represents. For example, on a map of a state park, a symbol of a tent would indicate the location of a camping area.

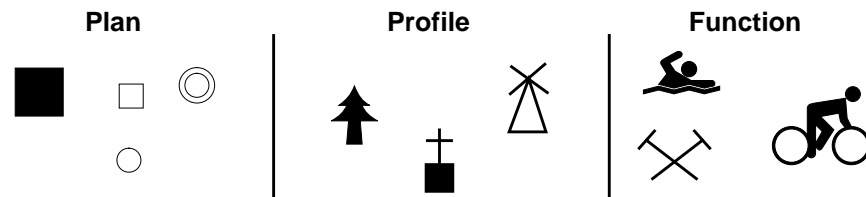


Figure 12-4 Sample Symbols

Symbols can have different sizes, colors, and patterns to indicate different meanings within a map. The use of size, color, and pattern generally shows qualitative or quantitative differences among areas marked. For example, if a circle is used to show cities and towns, larger circles would be used to show areas with higher population. A specific color could be used to indicate county seats. Since symbols are not drawn to scale, their placement is crucial to effective communication.



Use the Symbol tool in the Annotation tool palette and the symbol library to place symbols in maps.

Labels and Descriptive Text

Place names and other labels convey important information to the reader about the features on the map. Any features that help orient the reader or are important to the content of the map should be labeled. Descriptive text on a map can include the map title and subtitle, copyright information, captions, credits, production notes, or other explanatory material.

Title

The map title usually draws attention by virtue of its size. It focuses the reader's attention on the primary purpose of the map. The title may be omitted, however, if captions are provided outside of the image area (Dent 1985).

Credits

Map credits (or source information) can include the data source and acquisition date, accuracy information, and other details that are required or helpful to readers. For example, if you include data that you do not own in a map, you must give credit to the owner.



Use the Text tool in the Annotation tool palette to add labels and descriptive text to maps.

Typography and Lettering

The choice of type fonts and styles and how names are lettered can make the difference between a clear and attractive map and a jumble of imagery and text. As with many other aspects of map design, this is a very subjective area and many organizations already have guidelines to use. This section is intended as an introduction to the concepts involved and to convey traditional guidelines, where available.

If your organization does not have a set of guidelines for the appearance of maps and you plan to produce many in the future, it would be beneficial to develop a style guide specifically for mapping. This ensures that all of the maps produced follow the same conventions, regardless of who actually makes the map.



ERDAS IMAGINE enables you to make map templates to facilitate the development of map standards within your organization.

Type Styles

Type style refers to the appearance of the text and may include font, size, and style (bold, italic, underline, etc.). Although the type styles used in maps are purely a matter of the designer's taste, the following techniques help to make maps more legible (Robinson and Sale 1969; Dent 1985).

- Do not use too many different typefaces in a single map. Generally, one or two styles are enough when also using the variations of those type faces (e.g., bold, italic, underline, etc.). When using two typefaces, use a serif and a sans serif, rather than two different serif fonts or two different sans serif fonts [e.g., Sans (sans serif) and Roman (serif) could be used together in one map].
- Avoid ornate text styles because they can be difficult to read.
- Exercise caution in using very thin letters that may not reproduce well. On the other hand, using letters that are too bold may obscure important information in the image.
- Use different sizes of type for showing varying levels of importance. For example, on a map with city and town labels, city names are usually in a larger type size than the town names. Use no more than four to six different type sizes.

- Put more important text in labels, titles, and names in all capital letters and lesser important text in lowercase with initial capitals. This is a matter of personal preference, although names in which the letters must be spread out across a large area are better in all capital letters. (Studies have found that capital letters are more difficult to read, therefore lowercase letters might improve the legibility of the map.)
- In the past, hydrology, landform, and other natural features were labeled in italic. However, this is not strictly adhered to by map makers today, although water features are still nearly always labeled in italic.

Sans Serif	Serif
Sans 10 pt regular	Roman 10 pt regular
<i>Sans 10 pt italic</i>	<i>Roman 10 pt italic</i>
Sans 10 pt bold	Roman 10 pt bold
<i>Sans 10 pt bold italic</i>	<i>Roman 10 pt bold italic</i>
SANS 10 PT ALL CAPS	ROMAN 10 PT ALL CAPS

Figure 12-5 Sample Sans Serif and Serif Typefaces with Various Styles Applied



Use the Styles dialog to adjust the style of text.

Lettering

Lettering refers to the way in which place names and other labels are added to a map. Letter spacing, orientation, and position are the three most important factors in lettering. Here again, there are no set rules for how lettering is to appear. Much is determined by the purpose of the map and the end user. Many organizations have developed their own rules for lettering. Here is a list of guidelines that have been used by cartographers in the past (Robinson and Sale 1969; Dent 1985).

- Names should be either entirely on land or water—not overlapping both.
- Lettering should generally be oriented to match the orientation structure of the map. In large-scale maps this means parallel with the upper and lower edges, and in small-scale maps, this means in line with the parallels of latitude.
- Type should not be curved (i.e., different from preceding bullet) unless it is necessary to do so.
- If lettering must be disoriented, it should never be set in a straight line, but should always have a slight curve.
- Names should be letter spaced (i.e., space between individual letters, or kerning) as little as necessary.
- Where the continuity of names and other map data, such as lines and tones, conflicts with the lettering, the data, but not the names, should be interrupted.
- Lettering should never be upside-down in any respect.

- Lettering that refers to point locations should be placed above or below the point, preferably above and to the right.
- The letters identifying linear features (roads, rivers, railroads, etc.) should not be spaced. The word(s) should be repeated along the feature as often as necessary to facilitate identification. These labels should be placed above the feature and river names should slant in the direction of the river flow (if the label is italic).
- For geographical names, use the native language of the intended map user. For an English-speaking audience, the name Germany should be used, rather than Deutscheland.

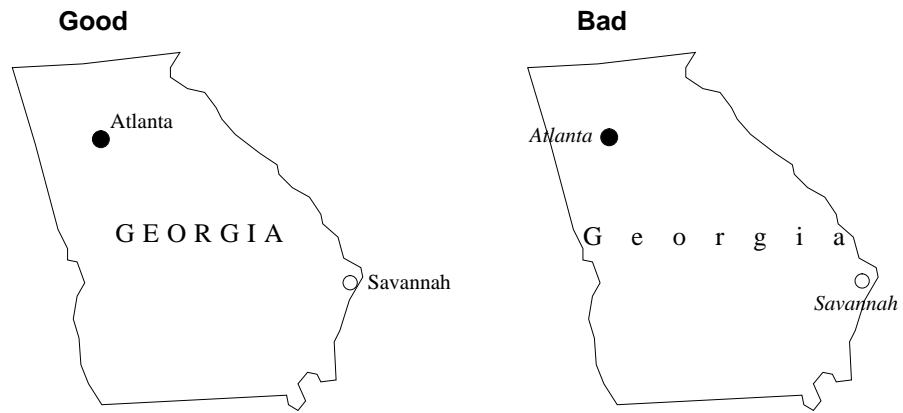


Figure 12-6 Good Lettering vs. Bad Lettering

Text Color

Many cartographers argue that all lettering on a map should be black. However, the map may be well-served by incorporating color into its design. In fact, studies have shown that coding labels by color can improve a reader's ability to find information (Dent 1985).

Map Projections



This section is adapted from "Map Projections for Use with the Geographic Information System" by Lee and Walsh, 1984.

A map projection is the manner in which the spherical surface of the Earth is represented on a flat (two-dimensional) surface. This can be accomplished by direct geometric projection or by a mathematically derived transformation. There are many kinds of projections, but all involve transfer of the distinctive global patterns of parallels of latitude and meridians of longitude onto an easily flattened surface, or developable surface.

The three most common developable surfaces are the cylinder, cone, and plane (Figure 12-7). A plane is already flat, while a cylinder or cone may be cut and laid out flat, without stretching. Thus, map projections may be classified into three general families: cylindrical, conical, and azimuthal or planar.



Map projections are selected in the Projection Chooser. For more information about the Projection Chooser, see the ERDAS IMAGINE On-Line Help.

Properties of Map Projections

Regardless of what type of projection is used, it is inevitable that some error or distortion occurs in transforming a spherical surface into a flat surface. Ideally, a distortion-free map has four valuable properties:

- conformality
- equivalence
- equidistance
- true direction

Each of these properties is explained below. No map projection can be true in all of these properties. Therefore, each projection is devised to be true in selected properties, or most often, a compromise among selected properties. Projections that compromise in this manner are known as compromise projections.

Conformality is the characteristic of true shape, wherein a projection preserves the shape of any small geographical area. This is accomplished by exact transformation of angles around points. One necessary condition is the perpendicular intersection of grid lines as on the globe. The property of conformality is important in maps which are used for analyzing, guiding, or recording motion, as in navigation. A conformal map or projection is one that has the property of true shape.

Equivalence is the characteristic of equal area, meaning that areas on one portion of a map are in scale with areas in any other portion. Preservation of equivalence involves inexact transformation of angles around points and thus, is mutually exclusive with conformality except along one or two selected lines. The property of equivalence is important in maps that are used for comparing density and distribution data, as in populations.

Equidistance is the characteristic of true distance measuring. The scale of distance is constant over the entire map. This property can be fulfilled on any given map from one, or at most two, points in any direction or along certain lines. Equidistance is important in maps that are used for analyzing measurements (i.e., road distances). Typically, reference lines such as the equator or a meridian are chosen to have equidistance and are termed standard parallels or standard meridians.

True direction is characterized by a direction line between two points that crosses reference lines (e.g., meridians) at a constant angle or azimuth. An azimuth is an angle measured clockwise from a meridian, going north to east. The line of constant or equal direction is termed a rhumb line.

The property of constant direction makes it comparatively easy to chart a navigational course. However, on a spherical surface, the shortest surface distance between two points is not a rhumb line, but a great circle, being an arc of a circle whose center is the center of the Earth. Along a great circle, azimuths constantly change (unless the great circle is the equator or a meridian).

Thus, a more desirable property than true direction may be where great circles are represented by straight lines. This characteristic is most important in aviation. Note that all meridians are great circles, but the only parallel that is a great circle is the equator.

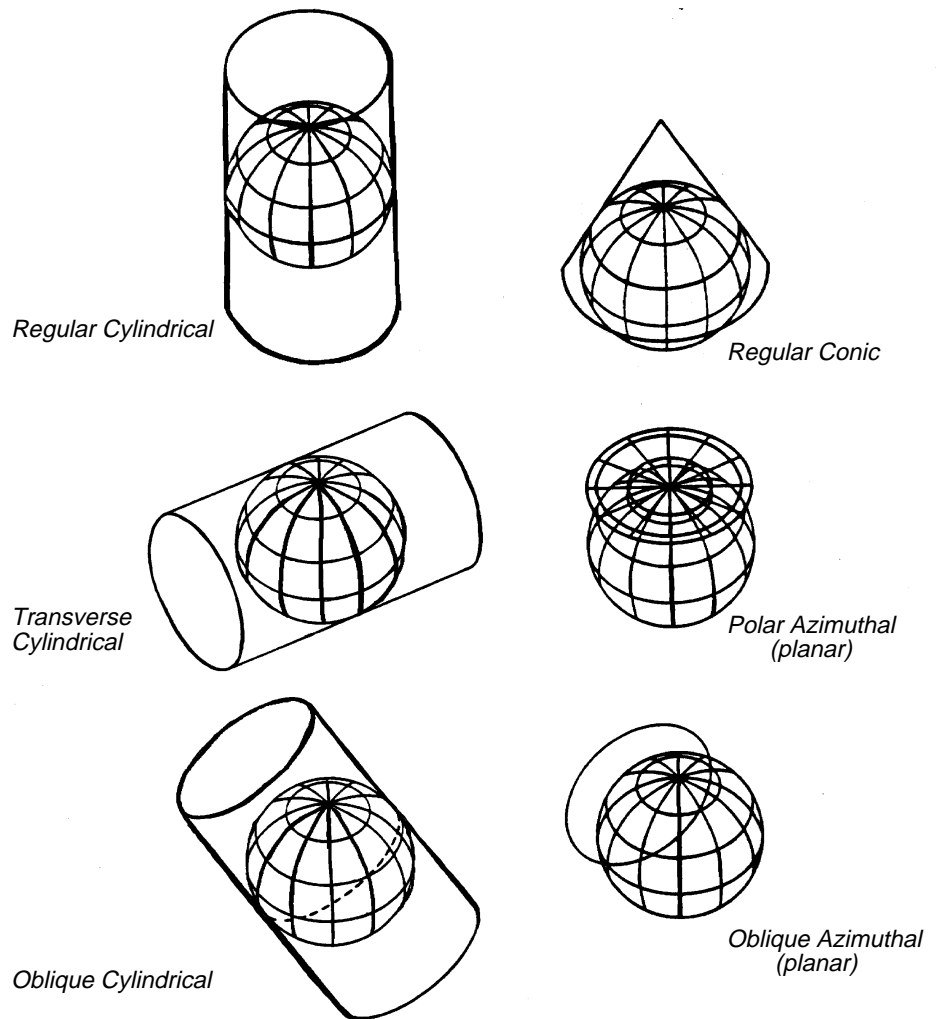


Figure 12-7 Projection Types

Projection Types

Although a great number of projections have been devised, the majority of them are geometric or mathematical variants of the basic direct geometric projection families described below. Choice of the projection to be used depends upon the true property or combination of properties desired for effective cartographic analysis.

Azimuthal Projections

Azimuthal projections, also called planar projections, are accomplished by drawing lines from a given perspective point through the globe onto a tangent plane. This is conceptually equivalent to tracing a shadow of a figure cast by a light source. A tangent plane intersects the global surface at only one point and is perpendicular to a line passing through the center of the sphere. Thus, these projections are symmetrical around a chosen center or central meridian. Choice of the projection center determines the aspect, or orientation, of the projection surface.

Azimuthal projections may be centered:

- on the poles (polar aspect)
- at a point on the equator (equatorial aspect)
- at any other orientation (oblique aspect)

The origin of the projection lines—that is, the perspective point—may also assume various positions. For example, it may be:

- the center of the Earth (gnomonic)
- an infinite distance away (orthographic)
- on the Earth's surface, opposite the projection plane (stereographic)

Conical Projections

Conical projections are accomplished by intersecting, or touching, a cone with the global surface and mathematically projecting lines onto this developable surface.

A tangent cone intersects the global surface to form a circle. Along this line of intersection, the map is error-free and possess equidistance. Usually, this line is a parallel, termed the standard parallel.

Cones may also be secant, and intersect the global surface, forming two circles that possess equidistance. In this case, the cone slices underneath the global surface, between the standard parallels. Note that the use of the word secant, in this instance, is only conceptual and not geometrically accurate. Conceptually, the conical aspect may be polar, equatorial, or oblique. Only polar conical projections are supported in ERDAS IMAGINE.

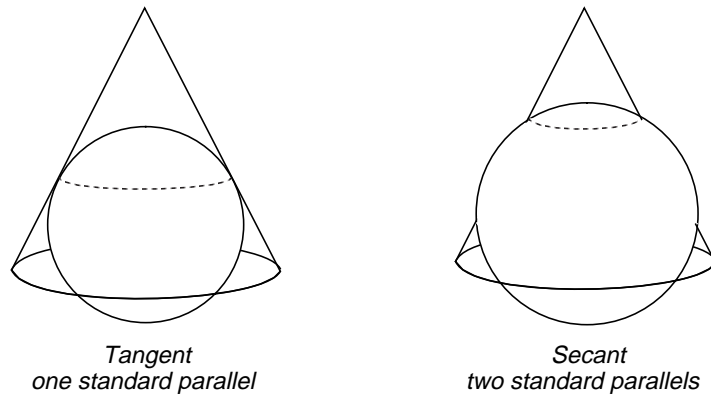


Figure 12-8 Tangent and Secant Cones

Cylindrical Projections

Cylindrical projections are accomplished by intersecting, or touching, a cylinder with the global surface. The surface is mathematically projected onto the cylinder, which is then cut and unrolled.

A tangent cylinder intersects the global surface on only one line to form a circle, as with a tangent cone. This central line of the projection is commonly the equator and possesses equidistance.

If the cylinder is rotated 90 degrees from the vertical (i.e., the long axis becomes horizontal), then the aspect becomes transverse, wherein the central line of the projection becomes a chosen standard meridian as opposed to a standard parallel. A secant cylinder, one slightly less in diameter than the globe, has two lines possessing equidistance.

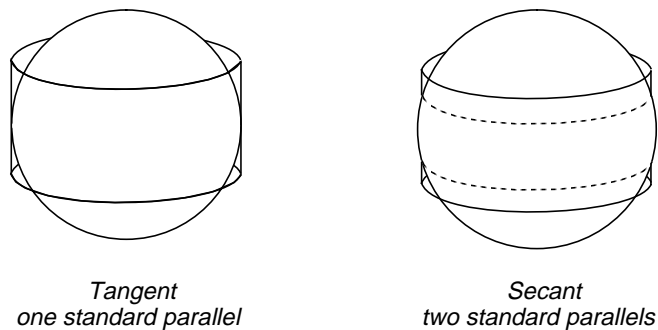


Figure 12-9 Tangent and Secant Cylinders

Perhaps the most famous cylindrical projection is the Mercator, which became the standard navigational map. Mercator possesses true direction and conformality.

Other Projections

The projections discussed so far are projections that are created by projecting from a sphere (the Earth) onto a plane, cone, or cylinder. Many other projections cannot be created so easily.

Modified projections are modified versions of another projection. For example, the Space Oblique Mercator projection is a modification of the Mercator projection. These modifications are made to reduce distortion, often by including additional standard lines or a different pattern of distortion.

Pseudo projections have only some of the characteristics of another class projection. For example, the Sinusoidal is called a pseudocylindrical projection because all lines of latitude are straight and parallel, and all meridians are equally spaced. However, it cannot truly be a cylindrical projection, because all meridians except the central meridian are curved. This results in the Earth appearing oval instead of rectangular (ESRI 1991).

Geographical and Planar Coordinates

Map projections require a point of reference on the Earth's surface. Most often this is the center, or origin, of the projection. This point is defined in two coordinate systems:

- geographical
- planar

Geographical

Geographical, or spherical, coordinates are based on the network of latitude and longitude (Lat/Lon) lines that make up the graticule of the Earth. Within the graticule, lines of longitude are called meridians, which run north/south, with the prime meridian at 0° (Greenwich, England). Meridians are designated as 0° to 180°, east or west of the prime meridian. The 180° meridian (opposite the prime meridian) is the International Dateline.

Lines of latitude are called parallels, which run east/west. Parallels are designated as 0° at the equator to 90° at the poles. The equator is the largest parallel. Latitude and longitude are defined with respect to an origin located at the intersection of the equator and the prime meridian. Lat/Lon coordinates are reported in degrees, minutes, and seconds. Map projections are various arrangements of the Earth's latitude and longitude lines onto a plane.

Planar

Planar, or Cartesian, coordinates are defined by a column and row position on a planar grid (X,Y). The origin of a planar coordinate system is typically located south and west of the origin of the projection. Coordinates increase from 0,0 going east and north. The origin of the projection, being a false origin, is defined by values of false easting and false northing. Grid references always contain an even number of digits, and the first half refers to the easting and the second half the northing.

In practice, this eliminates negative coordinate values and allows locations on a map projection to be defined by positive coordinate pairs. Values of false easting are read first and may be in meters or feet.

Available Map Projections

In ERDAS IMAGINE, map projection information appears in the Projection Chooser, which is used to georeference images and to convert map coordinates from one type of projection to another. The Projection Chooser provides the following projections:

USGS Projections

- Alaska Conformal
- Albers Conical Equal Area
- Azimuthal Equidistant
- Behrmann
- Bonne
- Cassini
- Conic Equidistant
- Eckert I
- Eckert II
- Eckert III
- Eckert IV
- Eckert V
- Eckert VI
- EOSAT SOM
- Equidistant Conic
- Equidistant Cylindrical
- Equiarectangular
- Gall Stereographic
- Gauss Kruger
- General Vertical Near-side Perspective
- Geographic (Lat/Lon)
- Gnomonic
- Hammer
- Interrupted Goode Homolosine
- Interrupted Mollweide
- Lambert Azimuthal Equal-area
- Lambert Conformal Conic
- Loximuthal
- Mercator
- Miller Cylindrical
- Modified Transverse Mercator
- Mollweide

- New Zealand Map Grid
- Oblated Equal Area
- Oblique Mercator (Hotine)
- Orthographic
- Plate Carrée
- Polar Stereographic
- Polyconic
- Quartic Authalic
- Robinson
- RSO
- Sinusoidal
- Space Oblique Mercator
- Space Oblique Mercator (Formats A & B)
- State Plane
- Stereographic
- Stereographic (Extended)
- Transverse Mercator
- Two Point Equidistant
- UTM
- Van Der Grinten I
- Wagner IV
- Wagner VII
- Winkel I
- Winkel II

External Projections

- Albers Equal Area
- Azimuthal Equidistant
- Bipolar Oblique Conic Conformal
- Cassini-Soldner
- Conic Equidistant
- Laborde Oblique Mercator
- Lambert Azimuthal Equal Area
- Lambert Conformal Conic
- Mercator
- Minimum Error Conformal
- Modified Polyconic

- Modified Stereographic
- Mollweide Equal Area
- Oblique Mercator
- Orthographic
- Plate Carrée
- Rectified Skew Orthomorphic
- Regular Polyconic
- Robinson Pseudocylindrical
- Sinusoidal
- Southern Orientated Gauss Conformal
- Stereographic
- Swiss Cylindrical
- Stereographic (Oblique)
- Transverse Mercator
- Universal Transverse Mercator
- Van der Grinten
- Winkel's Tripel

Choice of the projection to be used depends upon the desired major property and the region to be mapped (see Table 12-5, "Map Projections," on page 441). After choosing the desired map projection, several parameters are required for its definition (see Table 12-6, "Projection Parameters," on page 442). These parameters fall into three general classes: (1) definition of the spheroid, (2) definition of the surface viewing window, and (3) definition of scale.

For each map projection, a menu of spheroids displays, along with appropriate prompts that enable you to specify these parameters.

Units

Use the units of measure that are appropriate for the map projection type.

- Lat/Lon coordinates are expressed in decimal degrees. When prompted, you can use the DD function to convert coordinates in degrees, minutes, seconds format to decimal. For example, for 30°51'12":

$$\begin{aligned}\text{dd}(30,51,12) &= 30.85333 \\ -\text{dd}(30,51,12) &= -30.85333\end{aligned}$$

or

$$\mathbf{30:51:12} = 30.85333$$

You can also enter Lat/Lon coordinates in radians.

- State Plane coordinates are expressed in feet or meters.
- All other coordinates are expressed in meters.



Note also that values for longitude west of Greenwich, England, and values for latitude south of the equator are to be entered as negatives.

Table 12-5 Map Projections

Map projection	Construction	Property	Use
0 Geographic	N/A	N/A	Data entry, spherical coordinates
1 Universal Transverse Mercator	Cylinder (see #9)	Conformal	Data entry, plane coordinates
2 State Plane	(see #4, 7,9,20)	Conformal	Data entry, plane coordinates
3 Albers Conical Equal Area	Cone	Equivalent	Middle latitudes, E-W expanses
4 Lambert Conformal Conic	Cone	Conformal True Direction	Middle latitudes, E-W expanses flight (straight great circles)
5 Mercator	Cylinder	Conformal True Direction	Nonpolar regions, navigation (straight rhumb lines)
6 Polar Stereographic	Plane	Conformal	Polar regions
7 Polyconic	Cone	Compromise	N-S expanses
8 Equidistant Conic	Cone	Equidistant	Middle latitudes, E-W expanses
9 Transverse Mercator	Cylinder	Conformal	N-S expanses
10 Stereographic	Plane	Conformal	Hemispheres, continents
11 Lambert Azimuthal Equal Area	Plane	Equivalent	Square or round expanses
12 Azimuthal Equidistant	Plane	Equidistant	Polar regions, radio/seismic work (straight great circles)
13 Gnomonic	Plane	Compromise	Navigation, seismic work (straight great circles)
14 Orthographic	Plane	Compromise	Globes, pictorial
15 General Vertical Near-Side Perspective	Plane	Compromise	Hemispheres or less
16 Sinusoidal	Pseudo-Cylinder	Equivalent	N-S expanses or equatorial regions
17 Equirectangular	Cylinder	Compromise	City maps, computer plotting (simplistic)
18 Miller Cylindrical	Cylinder	Compromise	World maps
19 Van der Grinten I	N/A	Compromise	World maps
20 Oblique Mercator	Cylinder	Conformal	Oblique expanses (e.g., Hawaiian islands), satellite tracking
21 Space Oblique Mercator	Cylinder	Conformal	Mapping of Landsat imagery
22 Modified Transverse Mercator	Cylinder	Conformal	Alaska

Table 12-6 Projection Parameters

Parameter	Projection type (#) ^a																					
	3	4	5	6	7	8 ^b	9	10	11	12	13	14	15	16	17	18	19	20 ^b	21 ^b	22		
Definition of Spheroid																						
Spheroid selections	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Definition of Surface Viewing Window																						
False easting	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
False northing	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Longitude of central meridian	X	X	X		X	X	X							X	X	X	X					
Latitude of origin of projection	X	X			X	X	X											X				
Longitude of center of projection								X	X	X	X	X	X									
Latitude of center of projection								X	X	X	X	X	X									
Latitude of first standard parallel	X	X			X																	
Latitude of second standard parallel	X	X			X																	
Latitude of true scale			X	X																		
Longitude below pole				X																		
Definition of Scale																						
Scale factor at central meridian								X														
Height of perspective point above sphere												X										
Scale factor at center of projection																		X				

a. Numbers are used for reference only and correspond to the numbers used in Table 1. Parameters for definition of map projection types 0-2 are not applicable and are described in the text.

b. Additional parameters required for definition of the map projection are described in the text of Appendix C.

Choosing a Map Projection

Map Projection Uses in a GIS

Selecting a map projection for the GIS database enables you to (Maling 1992):

- decide how to best display the area of interest or illustrate the results of analysis
- register all imagery to a single coordinate system for easier comparisons
- test the accuracy of the information and perform measurements on the data

Deciding Factors

Depending on your applications and the uses for the maps created, one or several map projections may be used. Many factors must be weighed when selecting a projection, including:

- type of map
- special properties that must be preserved
- types of data to be mapped
- map accuracy
- scale

If you are mapping a relatively small area, virtually any map projection is acceptable. In mapping large areas (entire countries, continents, and the world), the choice of map projection becomes more critical. In small areas, the amount of distortion in a particular projection is barely, if at all, noticeable. In large areas, there may be little or no distortion in the center of the map, but distortion increases outward toward the edges of the map.

Guidelines

Since the sixteenth century, there have been three fundamental rules regarding map projection use (Maling 1992):

- if the country to be mapped lies in the tropics, use a cylindrical projection
- if the country to be mapped lies in the temperate latitudes, use a conical projection
- if the map is required to show one of the polar regions, use an azimuthal projection

These rules are no longer held so strongly. There are too many factors to consider in map projection selection for broad generalizations to be effective today. The purpose of a particular map and the merits of the individual projections must be examined before an educated choice can be made. However, there are some guidelines that may help you select a projection (Pearson 1990):

- Statistical data should be displayed using an equal area projection to maintain proper proportions (although shape may be sacrificed).
- Equal area projections are well-suited to thematic data.
- Where shape is important, use a conformal projection.

Spheroids

The previous discussion of direct geometric map projections assumes that the Earth is a sphere, and for many maps this is satisfactory. However, due to rotation of the Earth around its axis, the planet bulges slightly at the equator. This flattening of the sphere makes it an oblate spheroid, which is an ellipse rotated around its shorter axis.

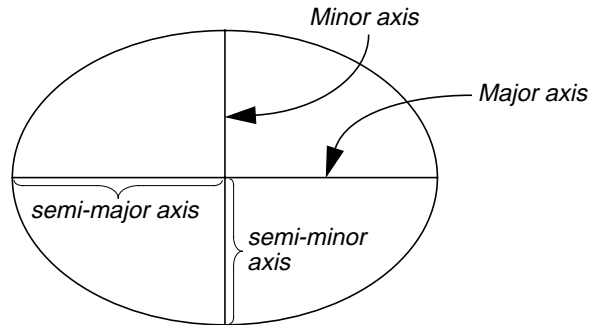


Figure 12-10 Ellipse

An ellipse is defined by its semi-major (long) and semi-minor (short) axes.

The amount of flattening of the Earth is expressed as the ratio:

$$f = (a - b) / a$$

Where:

a = the equatorial radius (semi-major axis)

b = the polar radius (semi-minor axis)

Most map projections use eccentricity (e^2) rather than flattening. The relationship is:

$$e^2 = 2f - f^2$$

The flattening of the Earth is about 1 part in 300, and becomes significant in map accuracy at a scale of 1:100,000 or larger.

Calculation of a map projection requires definition of the spheroid (or ellipsoid) in terms of the length of axes and eccentricity squared (or radius of the reference sphere). Several principal spheroids are in use by one or more countries. Differences are due primarily to calculation of the spheroid for a particular region of the Earth's surface. Only recently have satellite tracking data provided spheroid determinations for the entire Earth. However, these spheroids may not give the best fit for a particular region. In North America, the spheroid in use is the Clarke 1866 for NAD27 and GRS 1980 for NAD83 (State Plane).

If other regions are to be mapped, different projections should be used. Upon choosing a desired projection type, you have the option to choose from the following list of spheroids:

Airy
 Australian National
 Bessel
 Clarke 1866
 Clarke 1880
 Everest
 GRS 1980
 Helmert
 Hough
 International 1909
 Krasovsky
 Mercury 1960
 Modified Airy
 Modified Everest
 Modified Mercury 1968
 New International 1967
 Southeast Asia
 Sphere of Nominal Radius of Earth
 Sphere of Radius 6370977m
 Walbeck
 WGS 66
 WGS 72
 WGS 84



The spheroids listed above are the most commonly used. There are many other spheroids available, and they are listed in the Projection Chooser. These additional spheroids are not documented in this manual. You can use the IMAGINE Developers' Toolkit to add your own map projections and spheroids to ERDAS IMAGINE.

The semi-major and semi-minor axes of all supported spheroids are listed in Table 12-7, as well as the principal uses of these spheroids.

Table 12-7 Spheroids for use with ERDAS IMAGINE

Spheroid	Semi-Major Axis	Semi-Minor Axis	Use
165	6378165.0	6356783.0	Global
Airy (1940)	6377563.0	6356256.91	England
Airy Modified (1849)			Ireland
Australian National (1965)	6378160.0	6356774.719	Australia
Bessel (1841)	6377397.155	6356078.96284	Central Europe, Chile, and Indonesia
Bessell (Namibia)	6377483.865	6356165.383	Namibia
Clarke 1858	6378293.0	6356619.0	Global
Clarke 1866	6378206.4	6356583.8	North America and the Philippines

Table 12-7 Spheroids for use with ERDAS IMAGINE (Continued)

Spheroid	Semi-Major Axis	Semi-Minor Axis	Use
Clarke 1880	6378249.145	6356514.86955	France and Africa
Clarke 1880 IGN	6378249.2	6356515.0	Global
Everest (1830)	6377276.3452	6356075.4133	India, Burma, and Pakistan
Everest (1956)	6377301.243	6356100.2284	India, Nepal
Everest (1969)	6377295.664	6356094.6679	Global
Everest (Malaysia & Singapore)	6377304.063	6356103.038993	Global
Everest (Pakistan)	6377309.613	6356108.570542	Pakistan
Everest (Sabah & Sarawak)	6377298.556	6356097.5503	Brunei, East Malaysia
Fischer (1960)	6378166.0	6356784.2836	Global
Fischer (1968)	6378150.0	6356768.3372	Global
GRS 1980 (Geodetic Reference System)	6378137.0	6356752.31414	Adopted in North America for 1983 Earth-centered coordinate system (satellite)
Hayford	6378388.0	6356911.946128	Global
Helmert	6378200.0	6356818.16962789092	Egypt
Hough	6378270.0	6356794.343479	As International 1909 above, with modification of ellipse axes
IAU 1965	6378160.0	6356775.0	Global
Indonesian 1974	6378160.0	6356774.504086	Global
International 1909 (= Hayford)	6378388.0	6356911.94613	Remaining parts of the world not listed here
IUGG 1967	6378160.0	6356774.516	Hungary
Krasovsky (1940)	6378245.0	6356863.0188	Former Soviet Union and some East European countries
Mercury 1960	6378166.0	6356794.283666	Early satellite, rarely used
Modified Airy	6377341.89	6356036.143	As Airy above, more recent version
Modified Everest	6377304.063	6356103.039	As Everest above, more recent version
Modified Mercury 1968	6378150.0	6356768.337303	As Mercury 1960 above, more recent calculation
Modified Fischer (1960)	6378155.0	6356773.3205	Singapore
New International 1967	6378157.5	6356772.2	As International 1909 below, more recent calculation
SGS 85 (Soviet Geodetic System 1985)	6378136.0	6356751.3016	Soviet Union
South American (1969)	6378160.0	6356774.7192	South America
Southeast Asia	6378155.0	6356773.3205	As named
Sphere	6371000.0	6371000.0	Global

Table 12-7 Spheroids for use with ERDAS IMAGINE (Continued)

Spheroid	Semi-Major Axis	Semi-Minor Axis	Use
Sphere of Nominal Radius of Earth	6370997.0	6370997.0	A perfect sphere
Sphere of Radius 6370997 m	6370997.0	6370997.0	A perfect sphere with the same surface area as the Clarke 1866 spheroid
Walbeck (1819)	6376896.0	6355834.8467	Soviet Union, up to 1910
WGS 60 (World Geodetic System 1960)	6378165.0	6356783.287	Global
WGS 66 (World Geodetic System 1966)	6378145.0	6356759.769356	As WGS 72 above, older version
WGS 72 (World Geodetic System 1972)	6378135.0	6356750.519915	NASA (satellite)
WGS 84 (World Geodetic System 1984)	6378137.0	6356752.31424517929	As WGS 72, more recent calculation

Map Composition

Learning Map Composition

Cartography and map composition may seem like an entirely new discipline to many GIS and image processing analysts—and that is partly true. But, by learning the basics of map design, the results of your analyses can be communicated much more effectively. Map composition is also much easier than in the past when maps were hand drawn. Many GIS analysts may already know more about cartography than they realize, simply because they have access to map-making software. Perhaps the first maps you made were imitations of existing maps, but that is how we learn. This chapter is certainly not a textbook on cartography; it is merely an overview of some of the issues involved in creating cartographically-correct products.

Plan the Map

After your analysis is complete, you can begin map composition. The first step in creating a map is to plan its contents and layout. The following questions may aid in the planning process:

- How is this map going to be used?
- Will the map have a single theme or many?
- Is this a single map, or is it part of a series of similar maps?
- Who is the intended audience? What is the level of their knowledge about the subject matter?
- Will it remain in digital form and be viewed on the computer screen or will it be printed?
- If it is going to be printed, how big will it be? Will it be printed in color or black and white?
- Are there map guidelines already set up by your organization?

The answers to these questions can help to determine the type of information that must go into the composition and the layout of that information. For example, suppose you are going to do a series of maps about global deforestation for presentation to Congress, and you are going to print these maps in color on an inkjet printer. This scenario might lead to the following conclusions:

- A format (layout) should be developed for the series, so that all the maps produced have the same style.
- The colors used should be chosen carefully, since the maps are printed in color.
- Political boundaries might need to be included, since they influence the types of actions that can be taken in each deforested area.
- The typeface size and style to be used for titles, captions, and labels have to be larger than for maps printed on 8.5" × 11.0" sheets. The type styles selected should be the same for all maps.
- Select symbols that are widely recognized, and make sure they are all explained in a legend.
- Cultural features (roads, urban centers, etc.) may be added for locational reference.
- Include a statement about the accuracy of each map, since these maps may be used in very high-level decisions.

Once this information is in hand, you can actually begin sketching the look of the map on a sheet of paper. It is helpful for you to know how you want the map to look before starting the ERDAS IMAGINE Map Composer. Doing so ensures that all of the necessary data layers are available, and makes the composition phase go quickly.



See "CHAPTER 3: Map Composer" in the *ERDAS IMAGINE Tour Guides for step-by-step instructions on creating a map*. Refer to the *On-Line Help* for details about how *Map Composer* works.

Map Accuracy

Maps are often used to influence legislation, promote a cause, or enlighten a particular group before decisions are made. In these cases, especially, map accuracy is of the utmost importance. There are many factors that influence map accuracy: the projection used, scale, base data, generalization, etc. The analyst/cartographer must be aware of these factors before map production begins. The accuracy of the map, in a large part, determines its usefulness. It is usually up to individual organizations to perform accuracy assessment and decide how those findings are reflected in the products they produce. However, several agencies have established guidelines for map makers.

US National Map Accuracy Standard

The United States Bureau of the Budget has developed the US National Map Accuracy Standard in an effort to standardize accuracy reporting on maps. These guidelines are summarized below (Fisher 1991):

- On scales smaller than 1:20,000, not more than 10 percent of points tested should be more than 1/50 inch in horizontal error, where points refer only to points that can be well-defined on the ground.

- On maps with scales larger than 1:20,000, the corresponding error term is 1/30 inch.
- At no more than 10 percent of the elevations tested can contours be in error by more than one half of the contour interval.
- Accuracy should be tested by comparison of actual map data with survey data of higher accuracy (not necessarily with ground truth).
- If maps have been tested and do meet these standards, a statement should be made to that effect in the legend.
- Maps that have been tested but fail to meet the requirements should omit all mention of the standards on the legend.

USGS Land Use and Land Cover Map Guidelines

The USGS has set standards of their own for land use and land cover maps (Fisher 1991):

- The minimum level of accuracy in identifying land use and land cover categories is 85%.
- The several categories shown should have about the same accuracy.
- Accuracy should be maintained between interpreters and times of sensing.

USDA SCS Soils Maps Guidelines

The United States Department of Agriculture (USDA) has set standards for Soil Conservation Service (SCS) soils maps (Fisher 1991):

- Up to 25% of the pedons may be of other soil types than those named if they do not present a major hindrance to land management.
- Up to only 10% of pedons may be of other soil types than those named if they do present a major hindrance to land management.
- No single included soil type may occupy more than 10% of the area of the map unit.

Digitized Hardcopy Maps

Another method of expanding the database is by digitizing existing hardcopy maps. Although this may seem like an easy way to gather more information, care must be taken in pursuing this avenue if it is necessary to maintain a particular level of accuracy. If the hardcopy maps that are digitized are outdated, or were not produced using the same accuracy standards that are currently in use, the digitized map may negatively influence the overall accuracy of the database.

CHAPTER 13

Hardcopy Output

Introduction

Hardcopy output refers to any output of image data to paper. These topics are covered in this chapter:

- printing maps
- the mechanics of printing



For additional information, see "CHAPTER 8: Windows Printing" in the ERDAS IMAGINE V8.4 Installation Guide.

Printing Maps

ERDAS IMAGINE enables you to create and output a variety of types of hardcopy maps, with several referencing features.

Scaled Maps

A scaled map is a georeferenced map that has been projected to a map projection, and is accurately laid-out and referenced to represent distances and locations. A scaled map usually has a legend, that includes a scale, such as 1 inch = 1000 feet. The scale is often expressed as a ratio, like 1:12,000, where 1 inch on the map represents 12,000 inches on the ground.



See "CHAPTER 9: Rectification" for information on rectifying and georeferencing images and "CHAPTER 12: Cartography" for information on creating maps.

Printing Large Maps

Some scaled maps do not fit on the paper that is used by the printer. These methods are used to print and store large maps:

- A book map is laid out like the pages of a book. Each page fits on the paper used by the printer. There is a border, but no tick marks on every page.
- A paneled map is designed to be spliced together into a large paper map; therefore, borders and tick marks appear on the outer edges of the large map.

Hardcopy Output

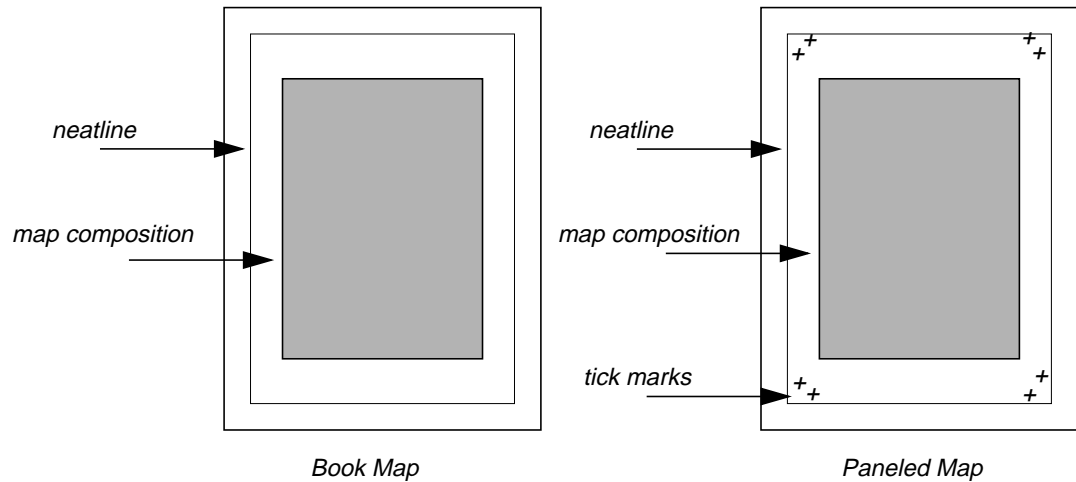


Figure 13-1 Layout for a Book Map and a Paneled Map

Scale and Resolution

The following scales and resolutions are noticeable during the process of creating a map composition and sending the composition to a hardcopy device:

- spatial resolution of the image
- display scale of the map composition
- map scale of the image(s)
- map composition to paper scale
- device resolution

Spatial Resolution

Spatial resolution is the area on the ground represented by each raw image data pixel.

Display Scale

Display scale is the distance on the screen as related to one unit on paper. For example, if the map composition is 24 inches by 36 inches, it would not be possible to view the entire composition on the screen. Therefore, the scale could be set to 1:0.25 so that the entire map composition would be in view.

Map Scale

The map scale is the distance on a map as related to the true distance on the ground, or the area that one pixel represents measured in map units. The map scale is defined when you create an image area in the map composition. One map composition can have multiple image areas set at different scales. These areas may need to be shown at different scales for different applications.

Map Composition to Paper Scale

This scale is the original size of the map composition as related to the desired output size on paper.

Device Resolution

The number of dots that are printed per unit—for example, 300 dots per inch (DPI).



Use the ERDAS IMAGINE Map Composer to define the above scales and resolutions.

Map Scaling Examples

The ERDAS IMAGINE Map Composer enables you to define a map size, as well as the size and scale for the image area within the map composition. The examples in this section focus on the relationship between these factors and the output file created by Map Composer for the specific hardcopy device or file format. Figure 13-2 is the map composition that is used in the examples. This composition was originally created using the ERDAS IMAGINE Map Composer at a size of 22" × 34", and the hardcopy output must be in two different formats.

- It must be output to a PostScript printer on an 8.5" × 11" piece of paper.
- A TIFF file must be created and sent to a film recorder having a 1,000 DPI resolution.

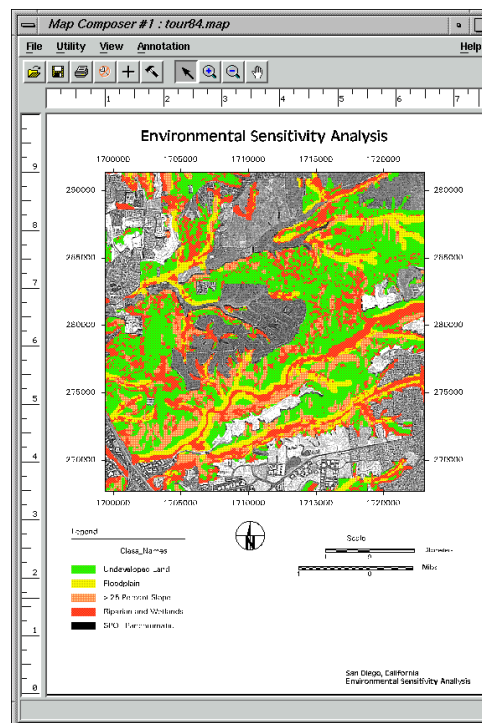


Figure 13-2 Sample Map Composition

Output to PostScript Printer

Since the map was created at 22" × 34", the map composition to paper scale needs to be calculated so that the composition fits on an 8.5" × 11" piece of paper. If this scale is set for a one to one ratio, then the composition is paneled.

Hardcopy Output

To determine the map composition to paper scale factor, it is necessary to calculate the most limiting direction. Since the printable area for the printer is approximately $8.1'' \times 8.6''$, these numbers are used in the calculation.

- $8.1'' / 22'' = 0.36$ (horizontal direction)
- $8.6'' / 34'' = 0.23$ (vertical direction)

The vertical direction is the most limiting; therefore, the map composition to paper scale would be set for 0.23.



If the specified size of the map (width and height) is greater than the printable area for the printer, the output hardcopy map is paneled. See the hardware manual of the hardcopy device for information about the printable area of the device.



Use the Print Map Composition dialog to output a map composition to a PostScript printer.

Output to TIFF

The limiting factor in this example is not page size, but disk space (600 MB total). A three-band image file must be created in order to convert the map composition to .tif file. Due to the three bands and the high resolution, the image file could be very large. The .tif file is output to a film recorder with a 1,000 DPI device resolution.

To determine the number of megabytes for the map composition, the X and Y dimensions need to be calculated:

- $X = 22 \text{ inches} \times 1,000 \text{ dots/inch} = 22,000$
- $Y = 34 \times 1,000 = 34,000$
- $22,000 \times 34,000 \times 3 = 2,244 \text{ MB}$ (multiplied by 3 since there are 3 bands)

Although this appears to be an unmanageable file size, it is possible to reduce the file size with little image degradation. The image file created from the map composition must be less than half to accommodate the .tif file, because the total disk space is only 600 megabytes. Dividing the map composition by three in both X and Y directions ($2,244 \text{ MB} / 3 / 3$) results in approximately a 250 megabyte file. This file size is small enough to process and leaves enough room for the image to TIFF conversion. This division is accomplished by specifying a $1/3$ or 0.333 map composition to paper scale when outputting the map composition to an image file.

Once the image file is created and exported to TIFF format, it can be sent to a film recorder that accepts .tif files. Remember, the file must be enlarged three times to compensate for the reduction during the image file creation.



See the hardware manual of the hardcopy device for information about the DPI device resolution.



Use the ERDAS IMAGINE Print Map Composition dialog to output a map composition to an image file.

Mechanics of Printing

This section describes the mechanics of transferring an image or map composition from a data file to a hardcopy map.

Halftone Printing

Halftoning is the process of converting a continuous tone image into a pattern of dots. A newspaper photograph is a common example of halftoning.

To make a color illustration, halftones in the primary colors (cyan, magenta, and yellow), plus black, are overlaid. The halftone dots of different colors, in close proximity, create the effect of blended colors in much the same way that phosphorescent dots on a color computer monitor combine red, green, and blue to create other colors. By using different patterns of dots, colors can have different intensities. The dots for halftoning are a fixed density—either a dot is there or it is not there.

For scaled maps, each output pixel may contain one or more dot patterns. If a very large image file is being printed onto a small piece of paper, data file pixels are skipped to accommodate the reduction.

Hardcopy Devices

The following hardcopy devices use halftoning to output an image or map composition:

- Tektronix Inkjet Printer
- Tektronix Phaser Printer



See the user's manual for the hardcopy device for more information about halftone printing.

Continuous Tone Printing

Continuous tone printing enables you to output color imagery using the four process colors (cyan, magenta, yellow, and black). By using varying percentages of these colors, it is possible to create a wide range of colors. The printer converts digital data from the host computer into a continuous tone image. The quality of the output picture is similar to a photograph. The output is smoother than halftoning because the dots for continuous tone printing can vary in density.

Hardcopy Output

Example

There are different processes by which continuous tone printers generate a map. One example is a process called thermal dye transfer. The entire image or map composition is loaded into the printer's memory. While the paper moves through the printer, heat is used to transfer the dye from a ribbon, which has the dyes for all of the four process colors, to the paper. The density of the dot depends on the amount of heat applied by the printer to transfer the dye. The amount of heat applied is determined by the brightness values of the input image. This allows the printer to control the amount of dye that is transferred to the paper to create a continuous tone image.

Hardcopy Devices

The following hardcopy device uses continuous toning to output an image or map composition:

- Tektronix Phaser II SD

NOTE: The above printers do not necessarily use the thermal dye transfer process to generate a map.



See the user's manual for the hardcopy device for more information about continuous tone printing.

Contrast and Color Tables

ERDAS IMAGINE contrast and color tables are used for some printing processes, just as they are used in displaying an image. For continuous raster layers, they are loaded from the ERDAS IMAGINE contrast table. For thematic layers, they are loaded from the color table. The translation of data file values to brightness values is performed entirely by the software program.

RGB to CMY Conversion

Colors

Since a printer uses ink instead of light to create a visual image, the primary colors of pigment (cyan, magenta, and yellow) are used in printing, instead of the primary colors of light (red, green, and blue). Cyan, magenta, and yellow can be combined to make black through a subtractive process, whereas the primary colors of light are additive—red, green, and blue combine to make white (Gonzalez and Wintz 1977).

The data file values that are sent to the printer and the contrast and color tables that accompany the data file are all in the RGB color scheme. The RGB brightness values in the contrast and color tables must be converted to cyan, magenta, and yellow (CMY) values.

The RGB primary colors are the opposites of the CMY colors—meaning, for example, that the presence of cyan in a color means an equal lack of red. To convert the values, each RGB brightness value is subtracted from the maximum brightness value to produce the brightness value for the opposite color. The following equation shows this relationship:

$$C = MAX - R$$
$$M = MAX - G$$
$$Y = MAX - B$$

Where:

MAX = the maximum brightness value
 R = red value from lookup table
 G = green value from lookup table
 B = blue value from lookup table
 C = calculated cyan value
 M = calculated magenta value
 Y = calculated yellow value

Black Ink

Although, theoretically, cyan, magenta, and yellow combine to create black ink, the color that results is often a dark, muddy brown. Many printers also use black ink for a truer black.

NOTE: Black ink may not be available on all printers. Consult the user's manual for your printer.

Images often appear darker when printed than they do when displayed on the display device. Therefore, it may be beneficial to improve the contrast and brightness of an image before it is printed.



Use the programs discussed in "CHAPTER 5: Enhancement" to brighten or enhance an image before it is printed.

Appendix A

Math Topics

Introduction

This appendix is a cursory overview of some of the basic mathematical concepts that are applicable to image processing. Its purpose is to educate the novice reader, and to put these formulas and concepts into the context of image processing and remote sensing applications.

Summation

A commonly used notation throughout this and other discussions is the Sigma (Σ), used to denote a summation of values.

For example, the notation

$$\sum_{i=1}^{10} i$$

is the sum of all values of i , ranging from 1 to 10, which equals:

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = 55.$$

Similarly, the value i may be a subscript, which denotes an ordered set of values. For example,

$$\sum_{i=1}^4 Q_i = 3 + 5 + 7 + 2 = 17$$

Where:

$$Q_1 = 3$$

$$Q_2 = 5$$

$$Q_3 = 7$$

$$Q_4 = 2$$

Statistics

Histogram

In ERDAS IMAGINE image data files, each data file value (defined by its row, column, and band) is a variable. ERDAS IMAGINE supports the following data types:

- 1, 2, and 4-bit
- 8, 16, and 32-bit signed
- 8, 16, and 32-bit unsigned
- 32 and 64-bit floating point
- 64 and 128-bit complex floating point

Distribution, as used in statistics, is the set of frequencies with which an event occurs, or that a variable has a particular value.

A histogram is a graph of data frequency or distribution. For a single band of data, the horizontal axis of a histogram is the range of all possible data file values. The vertical axis is the number of pixels that have each data value.

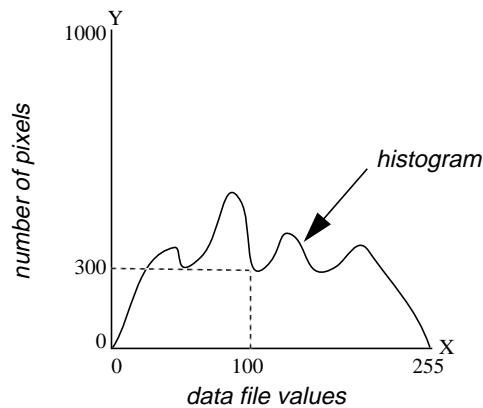


Figure A-1 Histogram

Figure A-1 shows the histogram for a band of data in which Y pixels have data value X. For example, in this graph, 300 pixels (y) have the data file value of 100 (x).

Bin Functions

Bins are used to group ranges of data values together for better manageability. Histograms and other descriptor columns for 1, 2, 4, and 8-bit data are easy to handle since they contain a maximum of 256 rows. However, to have a row in a descriptor table for every possible data value in floating point, complex, and 32-bit integer data would yield an enormous amount of information. Therefore, the bin function is provided to serve as a data reduction tool.

Example of a Bin Function

Suppose you have a floating point data layer with values ranging from 0.0 to 1.0. You could set up a descriptor table of 100 rows, with each row or bin corresponding to a data range of .01 in the layer.

The bins would look like the following:

Bin Number	Data Range
0	$X < 0.01$
1	$0.01 \leq X < 0.02$
2	$0.02 \leq X < 0.03$
.	
.	
.	
98	$0.98 \leq X < 0.99$
99	$0.99 \leq X$

Then, for example, row 23 of the histogram table would contain the number of pixels in the layer whose value fell between .023 and .024.

Types of Bin Functions

The bin function establishes the relationship between data values and rows in the descriptor table. There are four types of bin functions used in ERDAS IMAGINE image layers:

- **DIRECT**—one bin per integer value. Used by default for 1, 2, 4, and 8-bit integer data, but may be used for other data types as well. The direct bin function may include an offset for negative data or data in which the minimum value is greater than zero.

For example, a direct bin with 900 bins and an offset of -601 would look like the following:

Bin Number	Data Range
0	$X \leq -600.5$
1	$-600.5 < X \leq -599.5$
.	
.	
.	
599	$-2.5 < X \leq -1.5$
600	$-1.5 < X \leq -0.5$
601	$-0.5 < X < 0.5$
602	$0.5 \leq X < 1.5$
603	$1.5 \leq X < 2.5$
.	
.	

Bin Number	Data Range
.	
898	$296.5 \leq X < 297.5$
899	$297.5 \leq X$

- LINEAR—establishes a linear mapping between data values and bin numbers, as in our first example, mapping the data range 0.0 to 1.0 to bin numbers 0 to 99.

The bin number is computed by:

$$bin = numbins * (x - min) / (max - min)$$

$$if (bin < 0) bin = 0$$

$$if (bin \geq numbins) bin = numbins - 1$$

Where:

bin = resulting bin number

numbins = number of bins

x = data value

min = lower limit (usually minimum data value)

max = upper limit (usually maximum data value) LOG - establishes a logarithmic mapping between data values and bin numbers.

- LOG—establishes a logarithmic mapping between data values and bin numbers.

The bin number is computed by:

$$bin = numbins * (\ln(1.0 + ((x - min)/(max - min))) / \ln(2.0))$$

$$if (bin < 0) bin = 0$$

$$if (bin \geq numbins) bin = numbins - 1$$

- EXPLICIT—explicitly defines mapping between each bin number and data range.

Mean

The mean (μ) of a set of values is its statistical average, such that, if Q_i represents a set of k values:

$$\mu = \frac{Q_1 + Q_2 + Q_3 + \dots + Q_k}{k}$$

or

$$\mu = \sum_{i=1}^k \frac{Q_i}{k}$$

The mean of data with a normal distribution is the value at the peak of the curve—the point where the distribution balances.

Normal Distribution

Our general ideas about an average, whether it be average age, average test score, or the average amount of spectral reflectance from oak trees in the spring, are made visible in the graph of a normal distribution, or bell curve.

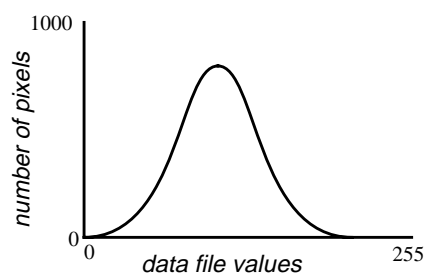


Figure A-2 Normal Distribution

Average usually refers to a central value on a bell curve, although all distributions have averages. In a normal distribution, most values are at or near the middle, as shown by the peak of the bell curve. Values that are more extreme are more rare, as shown by the tails at the ends of the curve.

The Normal Distributions are a family of bell shaped distributions that turn up frequently under certain special circumstances. For example, a normal distribution would occur if you were to compare the bands in a desert image. The bands would be very similar, but would vary slightly.

Each Normal Distribution uses just two parameters, σ and μ , to control the shape and location of the resulting probability graph through the equation:

$$f(x) = \frac{e^{-\left(\frac{x-\mu}{2\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

Where

- x = the quantity's distribution that is being approximated
- π and e = famous mathematical constants

The parameter μ controls how much the bell is shifted horizontally so that its average matches the average of the distribution of x , while σ adjusts the width of the bell to try to encompass the spread of the given distribution. In choosing to approximate a distribution by the nearest of the Normal Distributions, we describe the many values in the bin function of its distribution with just two parameters. It is a significant simplification that can greatly ease the computational burden of many operations, but like all simplifications, it reduces the accuracy of the conclusions we can draw.

The normal distribution is the most widely encountered model for probability. Many natural phenomena can be predicted or estimated according to the law of averages that is implied by the bell curve (Larsen and Marx 1981).

A normal distribution in remotely sensed data is meaningful—it is a sign that some characteristic of an object can be measured by the average amount of electromagnetic radiation that the object reflects. This relationship between the data and a physical scene or object is what makes image processing applicable to various types of land analysis.

The mean and standard deviation are often used by computer programs that process and analyze image data.

Variance

The mean of a set of values locates only the average value—it does not adequately describe the set of values by itself. It is helpful to know how much the data varies from its mean. However, a simple average of the differences between each value and the mean equals zero in every case, by definition of the mean. Therefore, the squares of these differences are averaged so that a meaningful number results (Larsen and Marx 1981).

In theory, the variance is calculated as follows:

$$Var_Q = E\langle(Q - \mu_Q)^2\rangle$$

Where:

E = expected value (weighted average)

2 = squared to make the distance a positive number

In practice, the use of this equation for variance does not usually reflect the exact nature of the values that are used in the equation. These values are usually only samples of a large data set, and therefore, the mean and variance of the entire data set are estimated, not known.

The equation used in practice follows. This is called the minimum variance unbiased estimator of the variance, or the sample variance (notated σ^2).

$$\sigma_Q^2 \approx \frac{\sum_{i=1}^k (Q_i - \mu_Q)^2}{k-1}$$

Where:

i = a particular pixel

k = the number of pixels (the higher the number, the better the approximation)

The theory behind this equation is discussed in chapters on point estimates and sufficient statistics, and covered in most statistics texts.

NOTE: The variance is expressed in units squared (e.g., square inches, square data values, etc.), so it may result in a number that is much higher than any of the original values.

Standard Deviation

Since the variance is expressed in units squared, a more useful value is the square root of the variance, which is expressed in units and can be related back to the original values (Larsen and Marx 1981). The square root of the variance is the standard deviation.

Based on the equation for sample variance (s^2), the sample standard deviation (s_Q) for a set of values Q is computed as follows:

$$s_Q = \sqrt{\frac{\sum_{i=1}^k (Q_i - \mu_Q)^2}{k-1}}$$

In any distribution:

- approximately 68% of the values are within one standard deviation of μ , that is, between $\mu-s$ and $\mu+s$
- more than 1/2 of the values are between $\mu-2s$ and $\mu+2s$
- more than 3/4 of the values are between $\mu-3s$ and $\mu+3s$

An example of a simple application of these rules is seen in the ERDAS IMAGINE Viewer. When 8-bit data are displayed in the Viewer, ERDAS IMAGINE automatically applies a 2 standard deviation stretch that remaps all data file values between $\mu-2s$ and $\mu+2s$ (more than 1/2 of the data) to the range of possible brightness values on the display device.

Standard deviations are used because the lowest and highest data file values may be much farther from the mean than $2s$.



For more information on contrast stretch, see "CHAPTER 5: Enhancement."

Parameters

As described above, the standard deviation describes how a fixed percentage of the data varies from the mean. The mean and standard deviation are known as parameters, which are sufficient to describe a normal curve (Johnston 1980).

When the mean and standard deviation are known, they can be used to estimate other calculations about the data. In computer programs, it is much more convenient to estimate calculations with a mean and standard deviation than it is to repeatedly sample the actual data.

Algorithms that use parameters are parametric. The closer that the distribution of the data resembles a normal curve, the more accurate the parametric estimates of the data are. ERDAS IMAGINE classification algorithms that use signature files (.sig) are parametric, since the mean and standard deviation of each sample or cluster are stored in the file to represent the distribution of the values.

Covariance

In many image processing procedures, the relationships between two bands of data are important. Covariance measures the tendencies of data file values in the same pixel, but in different bands, to vary with each other, in relation to the means of their respective bands. These bands must be linear.

Theoretically speaking, whereas variance is the average square of the differences between values and their mean in one band, covariance is the average product of the differences of corresponding values in two different bands from their respective means. Compare the following equation for covariance to the previous one for variance:

$$Cov_{QR} = E\langle(Q - \mu_Q)(R - \mu_R)\rangle$$

Where:

Q and R = data file values in two bands

E = expected value

In practice, the sample covariance is computed with this equation:

$$C_{QR} \approx \frac{\sum_{i=1}^k (Q_i - \mu_Q)(R_i - \mu_R)}{k}$$

Where:

i = a particular pixel

k = the number of pixels

Like variance, covariance is expressed in units squared.

Covariance Matrix

The covariance matrix is an $n \times n$ matrix that contains all of the variances and covariances within n bands of data. Below is an example of a covariance matrix for four bands of data:

	band A	band B	band C	band D
band A	Var_A	Cov_{BA}	Cov_{CA}	Cov_{DA}
band B	Cov_{AB}	Var_B	Cov_{CB}	Cov_{DB}
band C	Cov_{AC}	Cov_{BC}	Var_C	Cov_{DC}
band D	Cov_{AD}	Cov_{BD}	Cov_{CD}	Var_D

The covariance matrix is symmetrical—for example, $Cov_{AB} = Cov_{BA}$.

The covariance of one band of data with itself is the variance of that band:

$$C_{QQ} = \frac{\sum_{i=1}^k (Q_i - \mu_Q)(Q_i - \mu_Q)}{k - 1} = \frac{\sum_{i=1}^k (Q_i - \mu_Q)^2}{k - 1}$$

Therefore, the diagonal of the covariance matrix consists of the band variances.

The covariance matrix is an organized format for storing variance and covariance information on a computer system, so that it needs to be computed only once. Also, the matrix itself can be used in matrix equations, as in principal components analysis.



See "Matrix Algebra" on page 473 for more information on matrices.

Dimensionality of Data

Spectral Dimensionality is determined by the number of sets of values being used in a process. In image processing, each band of data is a set of values. An image with four bands of data is said to be four-dimensional (Jensen 1996).

NOTE: The letter n is used consistently in this documentation to stand for the number of dimensions (bands) of image data.

Measurement Vector

The measurement vector of a pixel is the set of data file values for one pixel in all n bands. Although image data files are stored band-by-band, it is often necessary to extract the measurement vectors for individual pixels.

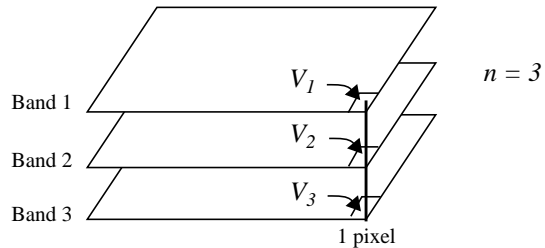


Figure A-3 Measurement Vector

According to Figure A-3,

i = particular band

V_i = the data file value of the pixel in band i , then the measurement vector for this pixel is:

$$\begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix}$$



See "Matrix Algebra" on page 473 for an explanation of vectors.

Mean Vector

When the measurement vectors of several pixels are analyzed, a **mean vector** is often calculated. This is the vector of the means of the data file values in each band. It has n elements.

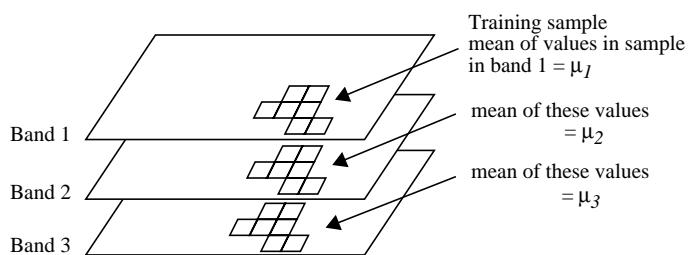


Figure A-4 Mean Vector

According to Figure A-4,

i = a particular band

μ_i = the mean of the data file values of the pixels being studied, in band i ,
then the mean vector for this training sample is:

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$$

Feature Space

Many algorithms in image processing compare the values of two or more bands of data. The programs that perform these functions abstractly plot the data file values of the bands being studied against each other. An example of such a plot in two dimensions (two bands) is illustrated in Figure A-5.

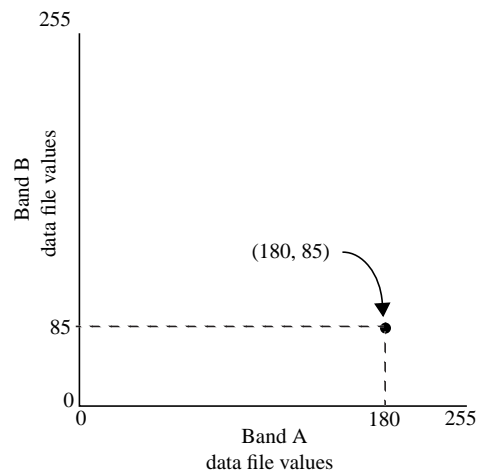


Figure A-5 Two Band Plot

NOTE: If the image is 2-dimensional, the plot doesn't always have to be 2-dimensional.

In Figure A-5, the pixel that is plotted has a measurement vector of:

$$\begin{bmatrix} 180 \\ 85 \end{bmatrix}$$

The graph above implies physical dimensions for the sake of illustration. Actually, these dimensions are based on spectral characteristics represented by the digital image data. As opposed to physical space, the pixel above is plotted in feature space. Feature space is an abstract space that is defined by spectral units, such as an amount of electromagnetic radiation.

Feature Space Images

Several techniques for the processing of multiband data make use of a two-dimensional histogram, or feature space image. This is simply a graph of the data file values of one band of data against the values of another band.

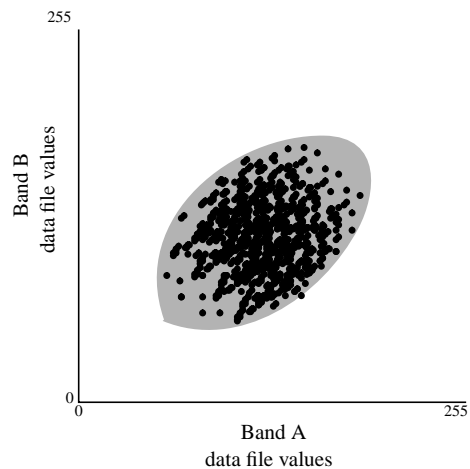


Figure A-6 Two-band Scatterplot

The scatterplot pictured in Figure A-6 can be described as a simplification of a two-dimensional histogram, where the data file values of one band have been plotted against the data file values of another band. This figure shows that when the values in the bands being plotted have jointly normal distributions, the feature space forms an ellipse.

This ellipse is used in several algorithms—specifically, for evaluating training samples for image classification. Also, two-dimensional feature space images with ellipses are helpful to illustrate principal components analysis.



See "CHAPTER 5: Enhancement" for more information on principal components analysis, "CHAPTER 6: Classification" for information on training sample evaluation, and "CHAPTER 9: Rectification" for more information on orders of transformation.

n-Dimensional Histogram

If two-dimensional data can be plotted on a two-dimensional histogram, as above, then n -dimensional data can, abstractly, be plotted on an n -dimensional histogram, defining n -dimensional spectral space.

Each point on an n -dimensional scatterplot has n coordinates in that spectral space — a coordinate for each axis. The n coordinates are the elements of the measurement vector for the corresponding pixel.

In some image enhancement algorithms (most notably, principal components), the points in the scatterplot are replotted, or the spectral space is redefined in such a way that the coordinates are changed, thus transforming the measurement vector of the pixel.

When all data sets (bands) have jointly normal distributions, the scatterplot forms a hyperellipsoid. The prefix “hyper” refers to an abstract geometrical shape, which is defined in more than three dimensions.

NOTE: In this documentation, 2-dimensional examples are used to illustrate concepts that apply to any number of dimensions of data. The 2-dimensional examples are best suited for creating illustrations to be printed.

Spectral Distance

Euclidean Spectral distance is distance in n -dimensional spectral space. It is a number that allows two measurement vectors to be compared for similarity. The spectral distance between two pixels can be calculated as follows:

$$D = \sqrt{\sum_{i=1}^n (d_i - e_i)^2}$$

Where:

- D = spectral distance
- n = number of bands (dimensions)
- i = a particular band
- d_i = data file value of pixel d in band i
- e_i = data file value of pixel e in band i

This is the equation for Euclidean distance—in two dimensions (when $n = 2$), it can be simplified to the Pythagorean Theorem ($c^2 = a^2 + b^2$), or in this case:

$$D^2 = (d_i - e_i)^2 + (d_j - e_j)^2$$

Polynomials

A polynomial is a mathematical expression consisting of variables and coefficients. A coefficient is a constant, which is multiplied by a variable in the expression.

Order

The variables in polynomial expressions can be raised to exponents. The highest exponent in a polynomial determines the order of the polynomial.

A polynomial with one variable, x , takes this form:

$$A + Bx + Cx^2 + Dx^3 + \dots + \Omega x^t$$

Where:

$A, B, C, D \dots \Omega$ = coefficients
 t = the order of the polynomial

NOTE: If one or all of $A, B, C, D \dots$ are 0, then the nature, but not the complexity, of the transformation is changed. Mathematically, Ω cannot be 0.

A polynomial with two variables, x and y , takes this form:

$$x_o = \left(\begin{array}{c} t \\ \Sigma \\ i = o \end{array} \right) \left(\begin{array}{c} i \\ \Sigma \\ j = o \end{array} \right) \quad a_k \times x^{i-j} \times y^j$$

$$y_o = \left(\begin{array}{c} t \\ \Sigma \\ i = o \end{array} \right) \left(\begin{array}{c} i \\ \Sigma \\ j = o \end{array} \right) \quad b_k \times x^{i-j} \times y^j$$

Where:

t is the order of the polynomial

a_k and b_k are coefficients

the subscript k in a_k and b_k is determined by:

$$k = \frac{i \cdot i + j}{2} + j$$

A numerical example of 3rd-order transformation equations for x and y is:

$$x_o = 5 + 4x - 6y + 10x^2 - 5xy + 1y^2 + 3x^3 + 7x^2y - 11xy^2 + 4y^3$$

$$y_o = 13 + 12x + 4y + 1x^2 - 21xy + 1y^2 - 1x^3 + 2x^2y + 5xy^2 + 12y^3$$

Polynomial equations are used in image rectification to transform the coordinates of an input file to the coordinates of another system. The order of the polynomial used in this process is the order of transformation.

Transformation Matrix

In the case of first order image rectification, the variables in the polynomials (x and y) are the source coordinates of a GCP. The coefficients are computed from the GCPs and stored as a transformation matrix.



A detailed discussion of GCPs, orders of transformation, and transformation matrices is included in "CHAPTER 9: Rectification."

Matrix Algebra

A matrix is a set of numbers or values arranged in a rectangular array. If a matrix has i rows and j columns, it is said to be an i by j matrix.

A one-dimensional matrix, having one column (i by 1) is one of many kinds of vectors. For example, the measurement vector of a pixel is an n -element vector of the data file values of the pixel, where n is equal to the number of bands.



See "CHAPTER 5: Enhancement" for information on eigenvectors.

Matrix Notation

Matrices and vectors are usually designated with a single capital letter, such as M . For example:

$$M = \begin{bmatrix} 2.2 & 4.6 \\ 6.1 & 8.3 \\ 10.0 & 12.4 \end{bmatrix}$$

One value in the matrix M would be specified by its position, which is its row and column (in that order) in the matrix. One element of the array (one value) is designated with a lower case letter and its position:

$$m_{3,2} = 12.4$$

With column vectors, it is simpler to use only one number to designate the position:

$$G = \begin{bmatrix} 2.8 \\ 6.5 \\ 10.1 \end{bmatrix}$$

$$G_2 = 6.5$$

Matrix Multiplication

A simple example of the application of matrix multiplication is a 1st-order transformation matrix. The coefficients are stored in a 2×3 matrix:

$$C = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

Then, where:

Math Topics

$$x_o = a_1 + a_2x_i + a_3y_i$$

$$y_o = b_1 + b_2x_i + b_3y_i$$

x_i and y_i = source coordinates

x_o and y_o = rectified coordinates

the coefficients of the transformation matrix are as above

The above could be expressed by a matrix equation:

$$\begin{bmatrix} x_o \\ y_o \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \begin{bmatrix} 1 \\ x_i \\ y_i \end{bmatrix}$$

$R = CS$, or

Where:

S = a matrix of the source coordinates (3 by 1)

C = the transformation matrix (2 by 3)

R = the matrix of rectified coordinates (2 by 1)

The sizes of the matrices are shown above to demonstrate a rule of matrix multiplication. To multiply two matrices, the first matrix must have the same number of columns as the second matrix has rows. For example, if the first matrix is a by b , and the second matrix is m by n , then b must equal m , and the product matrix has the size a by n .

The formula for multiplying two matrices is:

$$(fg)_{ij} = \sum_{k=1}^m f_{ik}g_{kj}$$

for every i from 1 to a

for every j from 1 to n

Where:

i = a row in the product matrix

j = a column in the product matrix

f = an (a by b) matrix

g = an (m by n) matrix (b must equal m)

fg is an a by n matrix.

Transposition

The transposition of a matrix is derived by interchanging its rows and columns. Transposition is denoted by T , as in the following example (Cullen 1972).

$$G = \begin{bmatrix} 2 & 3 \\ 6 & 4 \\ 10 & 12 \end{bmatrix}$$

$$G^T = \begin{bmatrix} 2 & 6 & 10 \\ 3 & 4 & 12 \end{bmatrix}$$



For more information on transposition, see "Computing Principal Components" in "CHAPTER 5: Enhancement" and "Classification Decision Rules" in "CHAPTER 6: Classification."

Appendix B

Map Projections

Introduction

This appendix is an alphabetical listing of the map projections supported in ERDAS IMAGINE. It is divided into two sections:

- USGS projections, beginning on page 477
- External projections, beginning on page 578

The external projections were implemented outside of ERDAS IMAGINE so that you could add to these using the IMAGINE Developers' Toolkit. The projections in each section are presented in alphabetical order.

The information in this appendix is adapted from:

- *Map Projections for Use with the Geographic Information System* (Lee and Walsh 1984)
- *Map Projections—A Working Manual* (Snyder 1987)
- *ArcInfo HELP* (ESRI 1997)

Other sources are noted in the text.



For general information about map projection types, refer to "CHAPTER 12: Cartography".



Rectify an image to a particular map projection using the ERDAS IMAGINE Rectification tools. View, add, or change projection information using the Image Information option.

NOTE: You cannot rectify to a new map projection using the Image Information option. You should change map projection information using Image Information only if you know the information to be incorrect. Use the rectification tools to actually georeference an image to a new map projection system.

USGS Projections

The following USGS map projections are supported in ERDAS IMAGINE and are described in this section:

- Alaska Conformal
- Albers Conical Equal Area
- Azimuthal Equidistant

Map Projections

- Behrmann
- Bonne
- Cassini
- Conic Equidistant
- Eckert I
- Eckert II
- Eckert III
- Eckert IV
- Eckert V
- Eckert VI
- EOSAT SOM
- Equidistant Cylindrical
- Equiarectangular
- Gall Stereographic
- Gauss Kruger
- General Vertical Near-side Perspective
- Geographic (Lat/Lon)
- Gnomonic
- Hammer
- Interrupted Goode Homolosine
- Interrupted Mollweide
- Lambert Azimuthal Equal Area
- Lambert Conformal Conic
- Loximuthal
- Mercator
- Miller Cylindrical
- Modified Transverse Mercator
- Mollweide

- New Zealand Map Grid
- Oblated Equal Area
- Oblique Mercator (Hotine)
- Orthographic
- Plate Carrée
- Polar Stereographic
- Polyconic
- Quartic Authalic
- Robinson
- RSO
- Sinusoidal
- Space Oblique Mercator
- Space Oblique Mercator (Formats A & B)
- State Plane
- Stereographic
- Stereographic (Extended)
- Transverse Mercator
- Two Point Equidistant
- UTM
- Van der Grinten I
- Wagner IV
- Wagner VII
- Winkel I
- Winkel II

Map Projections

Alaska Conformal

Table B-1 Alaska Conformal Summary

Construction	Modified planar
Property	Conformal
Meridians	N/A
Parallels	N/A
Graticule spacing	N/A
Linear scale	The minimum scale factor is 0.997 at roughly 62.5° N, 156° W. Scale increases outward from these coordinates. Most of Alaska and the Aleutian Islands (with the exception of the panhandle) is bounded by a line of true scale. The scale factor for Alaska is from 0.997 to 1.003. That is one quarter the range for a corresponding conic projection (Snyder 1987).
Uses	This projection is useful for mapping the complete state of Alaska on the Clarke 1866 spheroid or NAD27, but not with other datums and spheroids. Distortion increases as distance from Alaska increases.

Use of this projection results in a conformal map of Alaska. It has little scale distortion as compared to other conformal projections. The method of projection is “modified planar. [It is] a sixth-order-equation modification of an oblique Stereographic conformal projection on the Clarke 1866 spheroid. The origin is at 64° N, 152° W” (ESRI 1997).

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Alaska Conformal is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445 in "CHAPTER 12: Cartography".

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Albers Conical Equal Area

Table B-2 Albers Conical Equal Area Summary

Construction	Cone
Property	Equal-area
Meridians	Meridians are straight lines converging on the polar axis, but not at the pole.
Parallels	Parallels are arcs of concentric circles concave toward a pole.
Graticule spacing	Meridian spacing is equal on the standard parallels and decreases toward the poles. Parallel spacing decreases away from the standard parallels and increases between them. Meridians and parallels intersect each other at right angles. The graticule spacing preserves the property of equivalence of area. The graticule is symmetrical.
Linear scale	Linear scale is true on the standard parallels. Maximum scale error is 1.25% on a map of the United States (48 states) with standard parallels of 29.5°N and 45.5°N.
Uses	Used for thematic maps. Used for large countries with an east-west orientation. Maps based on the Albers Conical Equal Area for Alaska use standard parallels 55°N and 65°N; for Hawaii, the standard parallels are 8°N and 18°N. The National Atlas of the United States, United States Base Map (48 states), and the Geologic map of the United States are based on the standard parallels of 29.5°N and 45.5°N.

The Albers Conical Equal Area projection is mathematically based on a cone that is conceptually secant on two parallels. There is no areal deformation. The North or South Pole is represented by an arc. It retains its properties at various scales, and individual sheets can be joined along their edges.

This projection produces very accurate area and distance measurements in the middle latitudes (Figure B-1). Thus, Albers Conical Equal Area is well-suited to countries or continents where north-south depth is about 3/5 the breadth of east-west. When this projection is used for the continental US, the two standard parallels are 29.5° and 45.5° North.

Map Projections

This projection possesses the property of equal-area, and the standard parallels are correct in scale and in every direction. Thus, there is no angular distortion (i.e., meridians intersect parallels at right angles), and conformality exists along the standard parallels. Like other conics, Albers Conical Equal Area has concentric arcs for parallels and equally spaced radii for meridians. Parallels are not equally spaced, but are farthest apart between the standard parallels and closer together on the north and south edges.

Albers Conical Equal Area is the projection exclusively used by the USGS for sectional maps of all 50 states of the US in the National Atlas of 1970.

Prompts

The following prompts display in the Projection Chooser once Albers Conical Equal Area is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Latitude of 1st standard parallel

Latitude of 2nd standard parallel

Enter two values for the desired control lines of the projection (i.e., the standard parallels). Note that the first standard parallel is the southernmost.

Then, define the origin of the map projection in both spherical and rectangular coordinates.

Longitude of central meridian

Latitude of origin of projection

Enter values for longitude of the desired central meridian and latitude of the origin of projection.

False easting at central meridian

False northing at origin

Enter values of false easting and false northing, corresponding to the intersection of the central meridian and the latitude of the origin of projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates from occurring within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

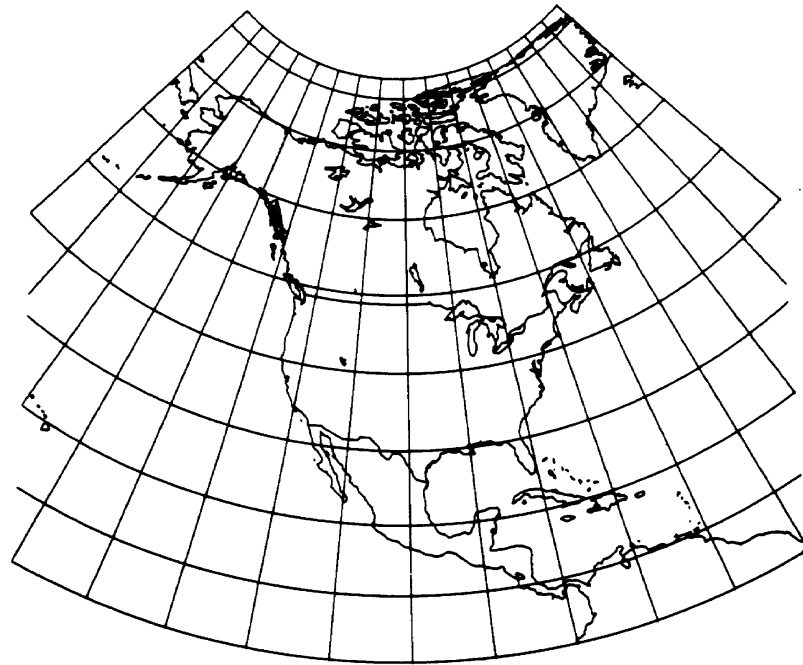


Figure B-1 Albers Conical Equal Area Projection

In Figure B-1, the standard parallels are 20°N and 60°N. Note the change in spacing of the parallels.

Azimuthal Equidistant

Table B-3 Azimuthal Equidistant Summary

Construction	Plane
Property	Equidistant
Meridians	<p>Polar aspect: the meridians are straight lines radiating from the point of tangency.</p> <p>Oblique aspect: the meridians are complex curves concave toward the point of tangency.</p> <p>Equatorial aspect: the meridians are complex curves concave toward a straight central meridian, except the outer meridian of a hemisphere, which is a circle.</p>
Parallels	<p>Polar aspect: the parallels are concentric circles.</p> <p>Oblique aspect: the parallels are complex curves.</p> <p>Equatorial aspect: the parallels are complex curves concave toward the nearest pole; the Equator is straight.</p>

Table B-3 Azimuthal Equidistant Summary (Continued)

Graticule spacing	Polar aspect: the meridian spacing is equal and increases away from the point of tangency. Parallel spacing is equidistant. Angular and area deformation increase away from the point of tangency.
Linear scale	Polar aspect: linear scale is true from the point of tangency along the meridians only. Oblique and equatorial aspects: linear scale is true from the point of tangency. In all aspects, the projection shows distances true to scale when measured between the point of tangency and any other point on the map.
Uses	The Azimuthal Equidistant projection is used for radio and seismic work, as every place in the world is shown at its true distance and direction from the point of tangency. The USGS uses the oblique aspect in the National Atlas and for large-scale mapping of Micronesia. The polar aspect is used as the emblem of the United Nations.

The Azimuthal Equidistant projection is mathematically based on a plane tangent to the Earth. The entire Earth can be represented, but generally less than one hemisphere is portrayed, though the other hemisphere can be portrayed, but is much distorted. It has true direction and true distance scaling from the point of tangency.

This projection is used mostly for polar projections because latitude rings divide meridians at equal intervals with a polar aspect (Figure B-2). Linear scale distortion is moderate and increases toward the periphery. Meridians are equally spaced, and all distances and directions are shown accurately from the central point.

This projection can also be used to center on any point on the Earth (e.g., a city) and distance measurements are true from that central point. Distances are not correct or true along parallels, and the projection is neither equal-area nor conformal. Also, straight lines radiating from the center of this projection represent great circles.

Prompts

The following prompts display in the Projection Chooser if Azimuthal Equidistant is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the center of the map projection in both spherical and rectangular coordinates.

Longitude of center of projection

Latitude of center of projection

Enter values for the longitude and latitude of the desired center of the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

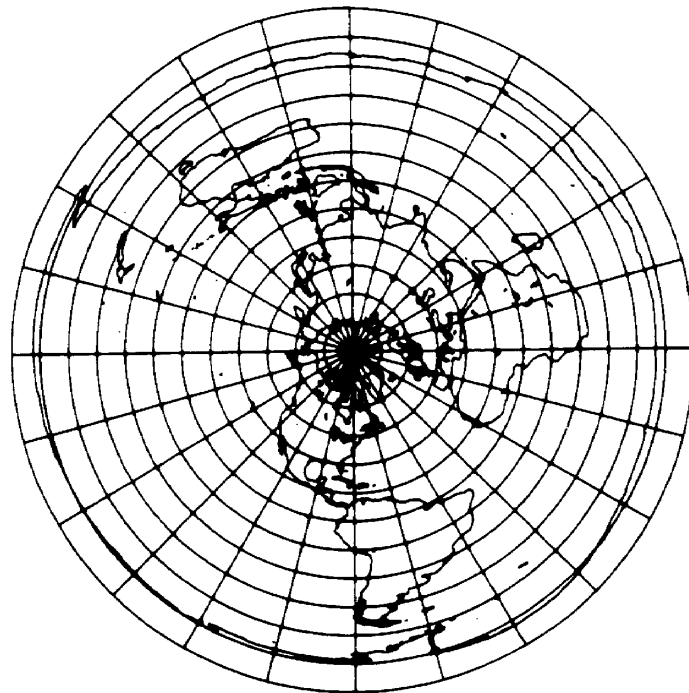


Figure B-2 Polar Aspect of the Azimuthal Equidistant Projection

This projection is commonly used in atlases for polar maps.

Table B-4 Behrmann Summary

Construction	Cylindrical
Property	Equal-area
Meridians	Straight parallel lines that are equally spaced and 0.42 the length of the Equator.
Parallels	Straight lines that are unequally spaced and farthest apart near the Equator, perpendicular to meridians.
Graticule spacing	See Meridians and Parallels. Poles are straight lines the same length as the Equator. Symmetry is present about any meridian or the Equator.
Linear scale	Scale is true along latitudes 30° N and S.
Uses	Used for creating world maps.

With the exception of compression in the horizontal direction and expansion in the vertical direction, the Behrmann projection is the same as the Lambert Cylindrical Equal-area projection. These changes prevent distortion at latitudes 30° N and S instead of at the Equator.

Source: Snyder and Voxland 1989

Prompts

The following prompts display in the Projection Chooser once Behrmann is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

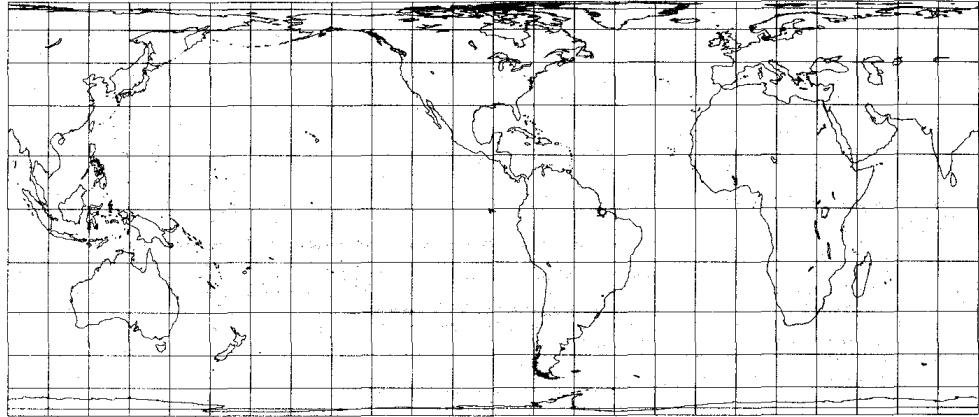
Longitude of central meridian

Enter a value of the longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-3 Behrmann Cylindrical Equal-Area Projection

Bonne

Table B-5 Bonne Summary

Construction	Pseudocone
Property	Equal-area
Meridians	N/A
Parallels	Parallels are concentric arcs that are equally spaced.
Graticule spacing	The central meridian is a linear graticule.
Linear scale	Scale is true along the central meridian and all parallels.
Uses	This projection is best used on maps of continents and small areas. There is some distortion.

The Bonne projection is an equal-area projection. True scale is achievable along the central meridian and all parallels. Although it was used in the 1800s and early 1900s, Bonne was replaced by Lambert Azimuthal Equal Area (see page 515) by the mapping company Rand McNally & Co. and Hammond, Inc.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Bonne is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Map Projections

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Latitude of standard parallel

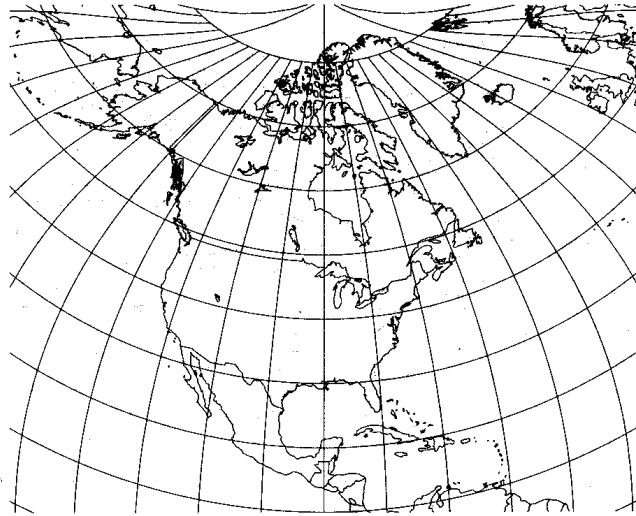
Longitude of central meridian

Enter values of the latitude of standard parallel and the longitude of central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-4 Bonne Projection

Cassini

Table B-6 Cassini Summary

Construction	Cylinder
Property	Compromise
Meridians	N/A
Parallels	N/A
Graticule spacing	Linear graticules are located at the Equator, central meridian, as well as meridians 90° from the central meridian.
Linear scale	With increasing distance from central meridian, scale distortion increases. Scale is true along the central meridian and lines perpendicular to the central meridian.
Uses	Cassini is used for large maps of areas near the central meridian. The extent is 5 degrees to either side of the central meridian.

The Cassini projection is a transverse cylindrical projection and is neither equal-area or conformal. It is best used for areas that are mostly in the north-south direction.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Cassini is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Scale Factor

Enter the scale factor.

Longitude of central meridian

Latitude of origin of projection

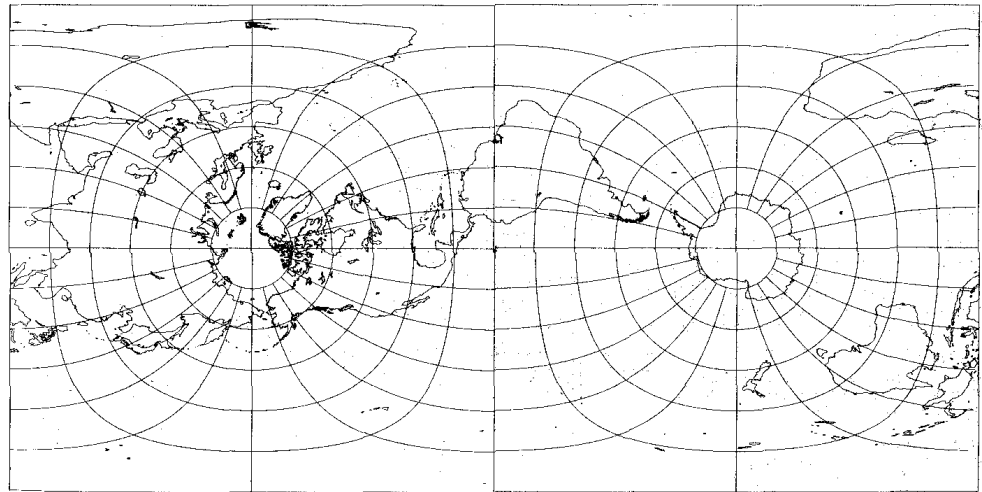
Enter the values for longitude of central meridian and latitude of origin of projection.

False easting

False northing

Map Projections

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-5 Cassini Projection

Eckert I

Table B-7 Eckert I Summary

Construction	Pseudocylinder
Property	Neither conformal nor equal-area
Meridians	Meridians are converging straight lines that are equally spaced and broken at the Equator.
Parallels	Parallels are perpendicular to the central meridian, equally spaced straight parallel lines.
Graticule spacing	See Meridians and Parallels. Poles are lines one half the length of the Equator. Symmetry exists about the central meridian or the Equator.
Linear scale	Scale is true along latitudes 47° 10' N and S. Scale is constant at any latitude (and latitude of opposite sign) and any meridian.
Uses	This projection is used as a novelty to show a straight-line graticule.

A great amount of distortion at the Equator is due to the break at the Equator.

Source: Snyder and Voxland 1989

Prompts

The following prompts display in the Projection Chooser once Eckert I is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

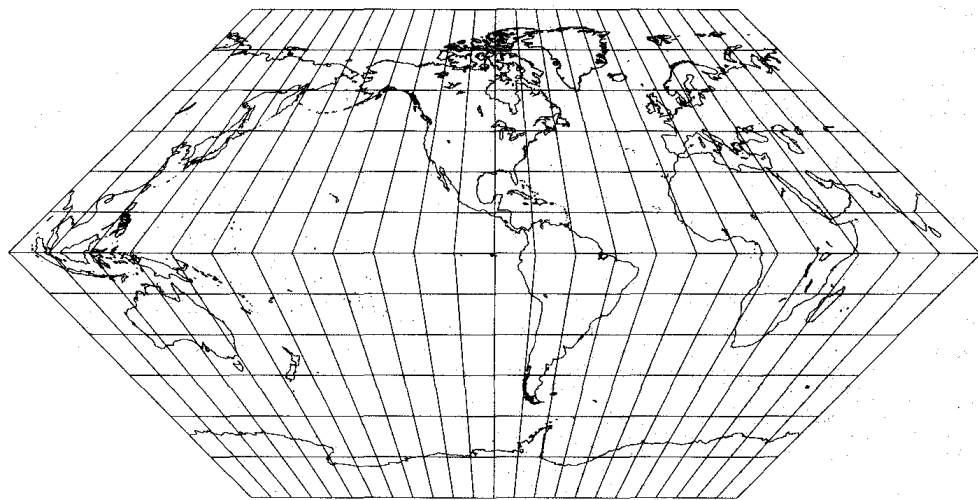
Longitude of central meridian

Enter a value of the longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-6 Eckert I Projection

Eckert II

Table B-8 Eckert II Summary

Construction	Pseudocylinder
Property	Equal-area
Meridians	Meridians are straight lines that are equally spaced and broken at the Equator. Central meridian is one half as long as the Equator.
Parallels	Parallels are straight parallel lines that are unequally spaced. The greatest separation is close to the Equator. Parallels are perpendicular to the central meridian.
Graticule spacing	See Meridians and Parallels. Pole lines are half the length of the Equator. Symmetry exists at the central meridian or the Equator.
Linear scale	Scale is true along altitudes 55° 10' N, and S. Scale is constant along any latitude.
Uses	This projection is used as a novelty to show straight-line equal-area graticule.

The break at the Equator creates a great amount of distortion there. Eckert II is similar to the Eckert I projection. The Eckert I projection has meridians positioned identically to Eckert II, but the Eckert I projection has equidistant parallels.

Source: Snyder and Voxland 1989

Prompts

The following prompts display in the Projection Chooser once Eckert II is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

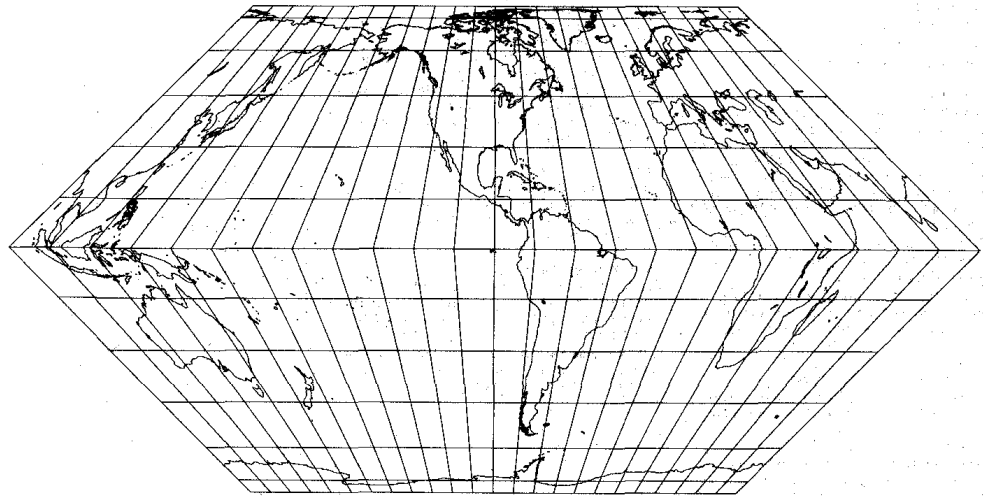
Longitude of central meridian

Enter a value of the longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-7 Eckert II Projection

Eckert III

Table B-9 Eckert III Summary

Construction	Pseudocylinder
Property	Area is not preserved.
Meridians	Meridians are elliptical curves that are equally spaced elliptical curves. The meridians $\pm 180^\circ$ from the central meridian are semicircles. The poles and the central meridian are straight lines one half the length of the Equator.
Parallels	Parallels are equally spaced straight lines.
Graticule spacing	See Meridians and Parallels. Pole lines are half the length of the Equator. Symmetry exists at the central meridian or the Equator.
Linear scale	Scale is correct only along 37° and $55'$ N and S. Features close to poles are compressed in the north-south direction.
Uses	Used for mapping the world.

Source: ESRI 1997

In the Eckert III projection, “no point is free of all scale distortion, but the Equator is free of angular distortion” (Snyder and Voxland 1989).

Map Projections

Prompts

The following prompts display in the Projection Chooser once Eckert III is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

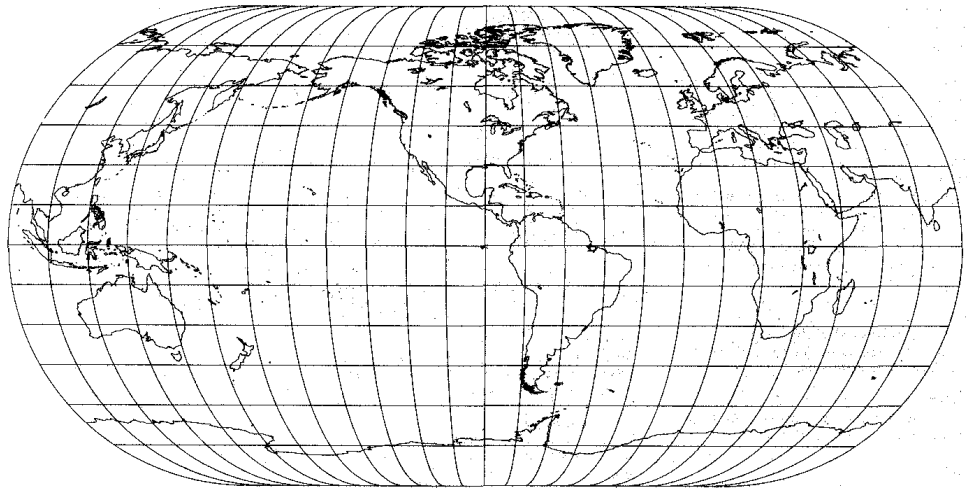
Longitude of central meridian

Enter a value of the longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-8 Eckert III Projection

Eckert IV

Table B-10 Eckert IV Summary

Construction	Pseudocylinder
Property	Equal-area
Meridians	Meridians are elliptical arcs that are equally spaced.
Parallels	Parallels are straight lines that are unequally spaced and closer together at the poles.
Graticule spacing	See Meridians and Parallels. The poles and the central meridian are straight lines one half the length of the Equator.
Linear scale	“Scale is distorted north-south 40 percent along the Equator relative to the east-west dimension. This distortion decreases to zero at 40° 30' N and S and at the central meridian. Scale is correct only along these parallels. Nearer the poles, features are compressed in the north-south direction” (ESRI 1997).
Uses	Use for world maps only.

The Eckert IV projection is best used for thematic maps of the globe. An example of a thematic map is one depicting land cover.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Eckert IV is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

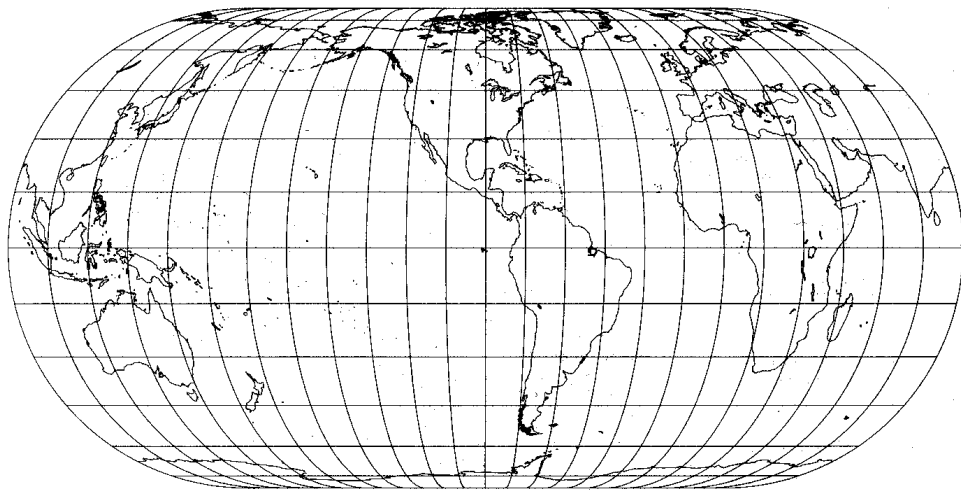
Longitude of central meridian

Enter a value of the longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-9 Eckert IV Projection

Eckert V

Table B-11 Eckert V Summary

Construction	Pseudocylinder
Property	Area is not preserved.
Meridians	Meridians are sinusoidal curves that are equally spaced. The poles and the central meridian are straight lines one half as long as the Equator.
Parallels	Parallels are straight lines that are equally spaced.
Graticule spacing	See Meridians and Parallels.
Linear scale	Scale is correct only along 37° 55' N and S. Features near the poles are compressed in the north-south direction.
Uses	This projection is best used for thematic world maps.

The Eckert V projection is only supported on a sphere. Like Eckert III, no point is free of all scale distortion, but the Equator is free of angular distortion" (Snyder and Voxland 1989).

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Eckert V is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

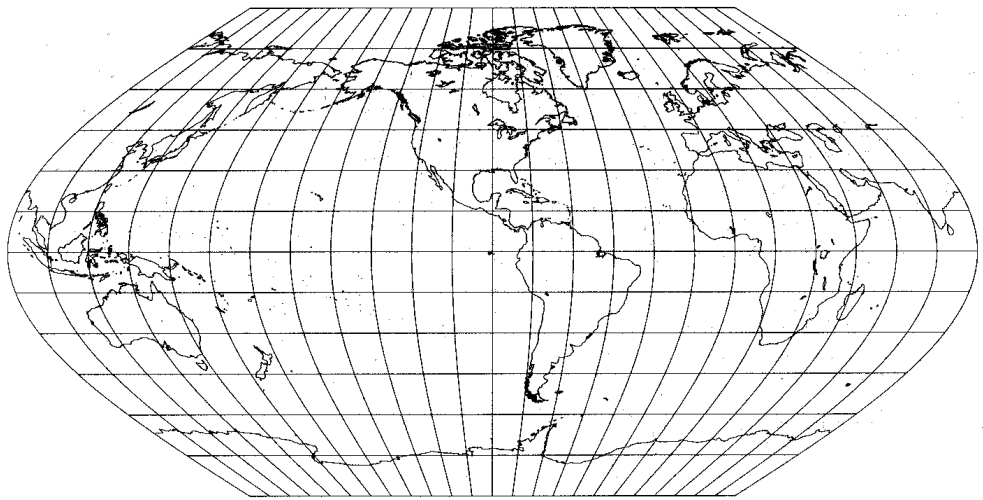
Longitude of central meridian

Enter a value of the longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-10 Eckert V Projection

Eckert VI

Table B-12 Eckert VI Summary

Construction	Pseudocylinder
Property	Equal-area
Meridians	Meridians are sinusoidal curves that are equally spaced.
Parallels	Parallels are unequally spaced straight lines, closer together at the poles.
Graticule spacing	See Meridians and Parallels. The poles and the central meridian are straight lines one half the length of the Equator.
Linear scale	“Scale is distorted north-south 29 percent along the Equator relative to the east-west dimension. This distortion decreases to zero at 49° 16’ N and S at the central meridian. Scale is correct only along these parallels. Nearer the poles, features are compressed in the north-south direction” (ESRI 1997).
Uses	Use for world maps only.

The Eckert VI projection is best used for thematic maps. An example of a thematic map is one depicting land cover.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Eckert VI is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

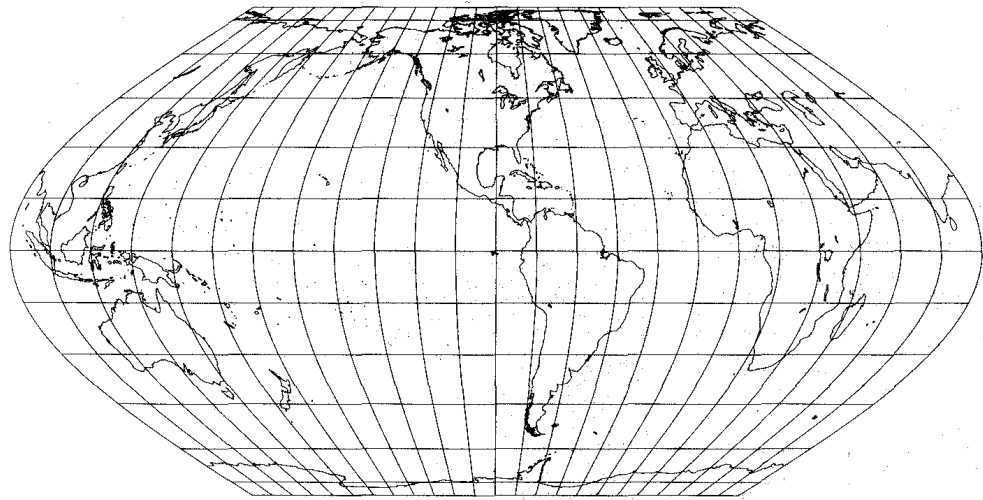
Longitude of central meridian

Enter a value of the longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-11 Eckert VI Projection

EOSAT SOM

The EOSAT SOM projection is similar to the Space Oblique Mercator projection. The main exception to the similarity is that the EOSAT SOM projection's X and Y coordinates are switched.



For information on the Space Oblique Mercator projection, see page 550.

Prompts

The following prompts display in the Projection Chooser once EOSAT SOM is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Landsat vehicle ID (1-5)

Specify whether the data are from Landsat 1, 2, 3, 4, or 5.

Orbital path number (1-251 or 1-233)

Map Projections

For Landsats 1, 2, and 3, the path range is from 1 to 251. For Landsats 4 and 5, the path range is from 1 to 233.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Equidistant Conic

Table B-13 Equidistant Conic Summary

Construction	Cone
Property	Equidistant
Meridians	Meridians are straight lines converging on a polar axis but not at the pole.
Parallels	Parallels are arcs of concentric circles concave toward a pole.
Graticule spacing	Meridian spacing is true on the standard parallels and decreases toward the pole. Parallels are placed at true scale along the meridians. Meridians and parallels intersect each other at right angles. The graticule is symmetrical.
Linear scale	Linear scale is true along all meridians and along the standard parallel or parallels.
Uses	The Equidistant Conic projection is used in atlases for portraying mid-latitude areas. It is good for representing regions with a few degrees of latitude lying on one side of the Equator. It was used in the former Soviet Union for mapping the entire country (ESRI 1992).

With Equidistant Conic (Simple Conic) projections, correct distance is achieved along the line(s) of contact with the cone, and parallels are equidistantly spaced. It can be used with either one (A) or two (B) standard parallels. This projection is neither conformal nor equal-area, but the north-south scale along meridians is correct. The North or South Pole is represented by an arc. Because scale distortion increases with increasing distance from the line(s) of contact, the Equidistant Conic is used mostly for mapping regions predominantly east-west in extent. The USGS uses the Equidistant Conic in an approximate form for a map of Alaska.

Prompts

The following prompts display in the Projection Chooser if Equidistant Conic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the origin of the projection in both spherical and rectangular coordinates.

Longitude of central meridian

Latitude of origin of projection

Enter values for the longitude of the desired central meridian and the latitude of the origin of projection.

False easting

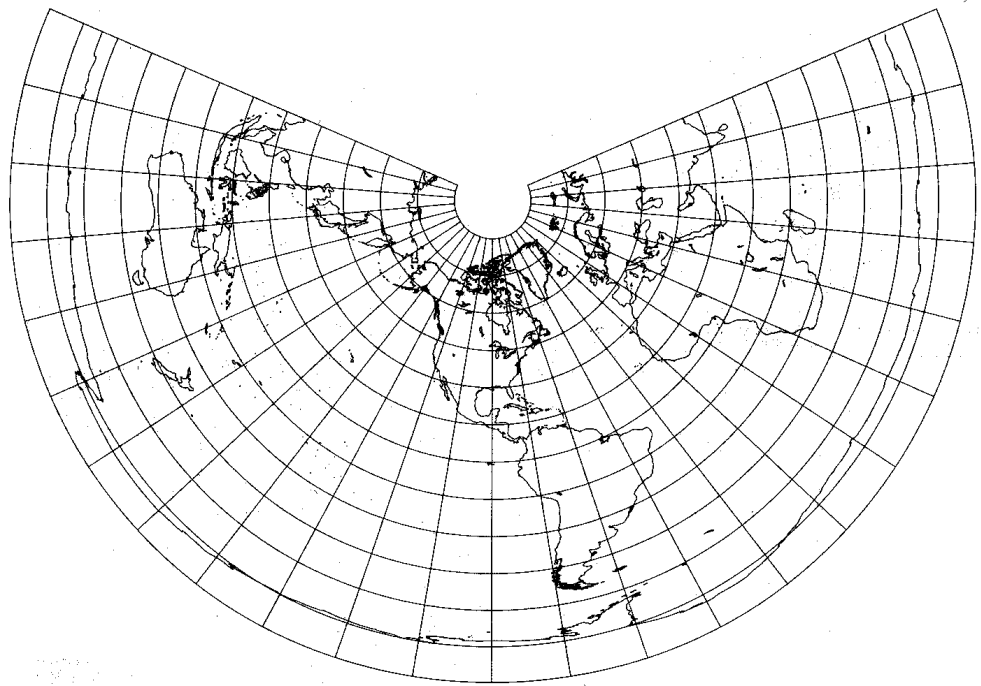
False northing

Enter values of false easting and false northing corresponding to the intersection of the central meridian and the latitude of the origin of projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

One or two standard parallels?

Latitude of standard parallel

Enter one or two values for the desired control line(s) of the projection, i.e., the standard parallel(s). Note that if two standard parallels are used, the first is the southernmost.



Source: Snyder and Voxland 1989

Figure B-12 Equidistant Conic Projection

Equidistant Cylindrical

The Equidistant Cylindrical projection is similar to the Equirectangular projection.



For information on the Equirectangular projection, see page 503.

Prompts

The following prompts display in the Projection Chooser if Equidistant Cylindrical is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Longitude of standard parallel

Latitude of true scale

Enter a value for longitude of the standard parallel and the latitude of true scale.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Equirectangular (Plate Carrée)

Table B-14 Equirectangular (Plate Carrée) Summary

Construction	Cylinder
Property	Compromise
Meridians	All meridians are straight lines.
Parallels	All parallels are straight lines.
Graticule spacing	Equally spaced parallel meridians and latitude lines cross at right angles.
Linear scale	The scale is correct along all meridians and along the standard parallels (ESRI 1992).
Uses	Best used for city maps, or other small areas with map scales small enough to reduce the obvious distortion. Used for simple portrayals of the world or regions with minimal geographic data, such as index maps (ESRI 1992).

Also called Simple Cylindrical, Equirectangular is composed of equally spaced, parallel meridians and latitude lines that cross at right angles on a rectangular map. Each rectangle formed by the grid is equal in area, shape, and size. Equirectangular is not conformal nor equal-area, but it does contain less distortion than the Mercator in polar regions. Scale is true on all meridians and on the central parallel. Directions due north, south, east, and west are true, but all other directions are distorted. The Equator is the standard parallel, true to scale and free of distortion. However, this projection may be centered anywhere.

This projection is valuable for its ease in computer plotting. It is useful for mapping small areas, such as city maps, because of its simplicity. The USGS uses Equirectangular for index maps of the conterminous US with insets of Alaska, Hawaii, and various islands. However, neither scale nor projection is marked to avoid implying that the maps are suitable for normal geographic information.

Map Projections

Prompts

The following prompts display in the Projection Chooser if Equirectangular is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Longitude of central meridian

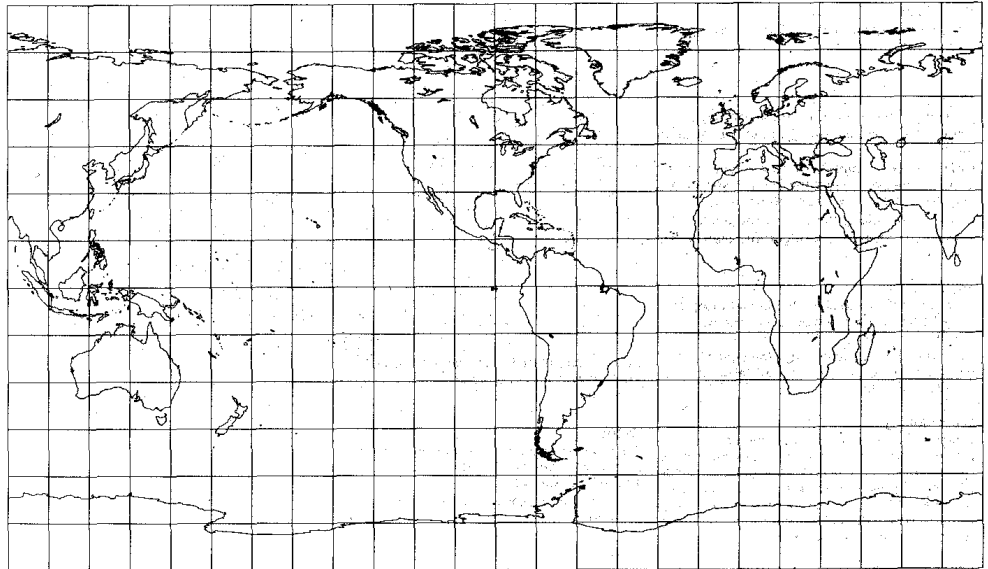
Latitude of true scale

Enter a value for longitude of the desired central meridian to center the projection and the latitude of true scale.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-13 Equirectangular Projection

Gall Stereographic

Table B-15 Gall Stereographic Summary

Construction	Cylinder
Property	Compromise
Meridians	Meridians are straight lines that are equally spaced.
Parallels	Parallels are straight lines that have increased space with distance from the Equator.
Graticule spacing	All meridians and parallels are linear.
Linear scale	“Scale is true in all directions along latitudes 45° N and S. Scale is constant along parallels and is symmetrical around the Equator. Distances are compressed between latitudes 45° N and S, and expanded beyond them” (ESRI 1997).
Uses	Use for world maps only.

The Gall Stereographic projection was created in 1855. The two standard parallels are located at 45° N and 45° S. This projection is used for world maps.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Gall Stereographic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Longitude of central meridian

Enter a value for longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Map Projections

Gauss Kruger

The Gauss Kruger projection is the same as the Transverse Mercator projection, with the exception that Gauss Kruger uses a fixed scale factor of 1. Gauss Kruger is available only in ellipsoidal form.

Many countries such as China and Germany use Gauss Kruger in 3-degree zones instead of 6-degree zones for UTM.



For information on the Transverse Mercator projection, see page 566.

Prompts

The following prompts display in the Projection Chooser once Gauss Kruger is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.

Scale factor

Designate the desired scale factor. This parameter is used to modify scale distortion. A value of one indicates true scale only along the central meridian. It may be desirable to have true scale along two lines equidistant from and parallel to the central meridian, or to lessen scale distortion away from the central meridian. A factor of less than, but close to one is often used.

Longitude of central meridian

Enter a value for the longitude of the desired central meridian to center the projection.

Latitude of origin of projection

Enter the value for the latitude of origin of projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

General Vertical Near-side Perspective

Table B-16 General Vertical Near-side Perspective Summary

Construction	Plane
Property	Compromise
Meridians	The central meridian is a straight line in all aspects. In the polar aspect all meridians are straight. In the equatorial aspect the Equator is straight (ESRI 1992).
Parallels	Parallels on vertical polar aspects are concentric circles. Nearly all other parallels are elliptical arcs, except that certain angles of tilt may cause some parallels to be shown as parabolas or hyperbolas.
Graticule spacing	Polar aspect: parallels are concentric circles that are not evenly spaced. Meridians are evenly spaced and spacing increases from the center of the projection. Equatorial and oblique aspects: parallels are elliptical arcs that are not evenly spaced. Meridians are elliptical arcs that are not evenly spaced, except for the central meridian, which is a straight line.
Linear scale	Radial scale decreases from true scale at the center to zero on the projection edge. The scale perpendicular to the radii decreases, but not as rapidly (ESRI 1992).
Uses	Often used to show the Earth or other planets and satellites as seen from space. Used as an aesthetic presentation, rather than for technical applications (ESRI 1992).

General Vertical Near-side Perspective presents a picture of the Earth as if a photograph were taken at some distance less than infinity. The map user simply identifies area of coverage, distance of view, and angle of view. It is a variation of the General Perspective projection in which the “camera” precisely faces the center of the Earth.

Central meridian and a particular parallel (if shown) are straight lines. Other meridians and parallels are usually arcs of circles or ellipses, but some may be parabolas or hyperbolas. Like all perspective projections, General Vertical Near-side Perspective cannot illustrate the entire globe on one map—it can represent only part of one hemisphere.

Prompts

The following prompts display in the Projection Chooser if General Vertical Near-side Perspective is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Map Projections

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Height of perspective point

Enter a value for the desired height of the perspective point above the sphere in the same units as the radius.

Then, define the center of the map projection in both spherical and rectangular coordinates.

Longitude of center of projection

Latitude of center of projection

Enter values for the longitude and latitude of the desired center of the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Geographic (Lat/Lon)

The Geographic is a spherical coordinate system composed of parallels of latitude (Lat) and meridians of longitude (Lon) (Figure B-14). Both divide the circumference of the Earth into 360 degrees. Degrees are further subdivided into minutes and seconds (60 sec = 1 minute, 60 min = 1 degree).

Because the Earth spins on an axis between the North and South Poles, this allows construction of concentric, parallel circles, with a reference line exactly at the north-south center, termed the Equator. The series of circles north of the Equator is termed north latitudes and runs from 0° latitude (the Equator) to 90° North latitude (the North Pole), and similarly southward. Position in an east-west direction is determined from lines of longitude. These lines are not parallel, and they converge at the poles. However, they intersect lines of latitude perpendicularly.

Unlike the Equator in the latitude system, there is no natural zero meridian. In 1884, it was finally agreed that the meridian of the Royal Observatory in Greenwich, England, would be the prime meridian. Thus, the origin of the geographic coordinate system is the intersection of the Equator and the prime meridian. Note that the 180° meridian is the international date line.

If you choose Geographic from the Projection Chooser, the following prompts display:

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.



Note that in responding to prompts for other projections, values for longitude are negative west of Greenwich and values for latitude are negative south of the Equator.

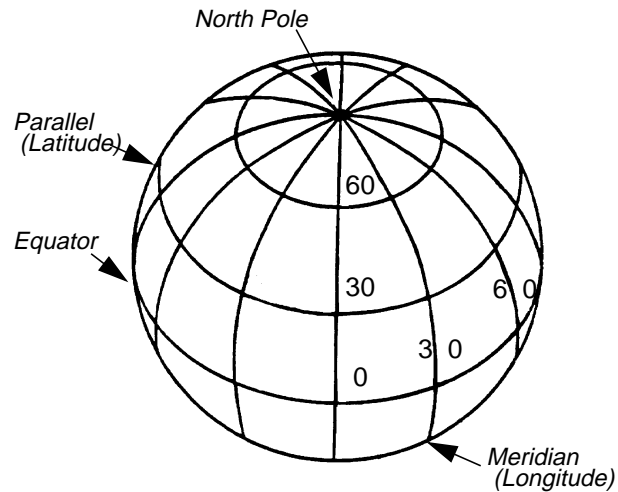


Figure B-14 Geographic Projection

Figure B-14 shows the graticule of meridians and parallels on the global surface.

Gnomonic

Table B-17 Gnomonic Summary

Construction	Plane
Property	Compromise
Meridians	Polar aspect: the meridians are straight lines radiating from the point of tangency. Oblique and equatorial aspects: the meridians are straight lines.
Parallels	Polar aspect: the parallels are concentric circles. Oblique and equatorial aspects: parallels are ellipses, parabolas, or hyperbolas concave toward the poles (except for the Equator, which is straight).
Graticule spacing	Polar aspect: the meridian spacing is equal and increases away from the pole. The parallel spacing increases rapidly from the pole. Oblique and equatorial aspects: the graticule spacing increases very rapidly away from the center of the projection.
Linear scale	Linear scale and angular and areal deformation are extreme, rapidly increasing away from the center of the projection.
Uses	The Gnomonic projection is used in seismic work because seismic waves travel in approximately great circles. It is used with the Mercator projection for navigation.

Gnomonic is a perspective projection that projects onto a tangent plane from a position in the center of the Earth. Because of the close perspective, this projection is limited to less than a hemisphere. However, it is the only projection which shows all great circles as straight lines. With a polar aspect, the latitude intervals increase rapidly from the center outwards.

With an equatorial or oblique aspect, the Equator is straight. Meridians are straight and parallel, while intervals between parallels increase rapidly from the center and parallels are convex to the Equator.

Because great circles are straight, this projection is useful for air and sea navigation. Rhumb lines are curved, which is the opposite of the Mercator projection.

Prompts

The following prompts display in the Projection Chooser if Gnomonic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the center of the map projection in both spherical and rectangular coordinates.

Longitude of center of projection

Latitude of center of projection

Enter values for the longitude and latitude of the desired center of the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Hammer

Table B-18 Hammer Summary

Construction	Modified Azimuth
Property	Equal-area
Meridians	The central meridian is half as long as the Equator and a straight line. Others are curved and concave toward the central meridian and unequally spaced.
Parallels	With the exception of the Equator, all parallels are complex curves that have a concave shape toward the nearest pole.
Graticule spacing	Only the Equator and central meridian are straight lines.
Linear scale	Scale lessens along the Equator and central meridian as proximity to the origin grows.
Uses	Use for world maps only.

The Hammer projection is useful for mapping the world. In particular, the Hammer projection is suited for thematic maps of the world, such as land cover.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Hammer is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Map Projections

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

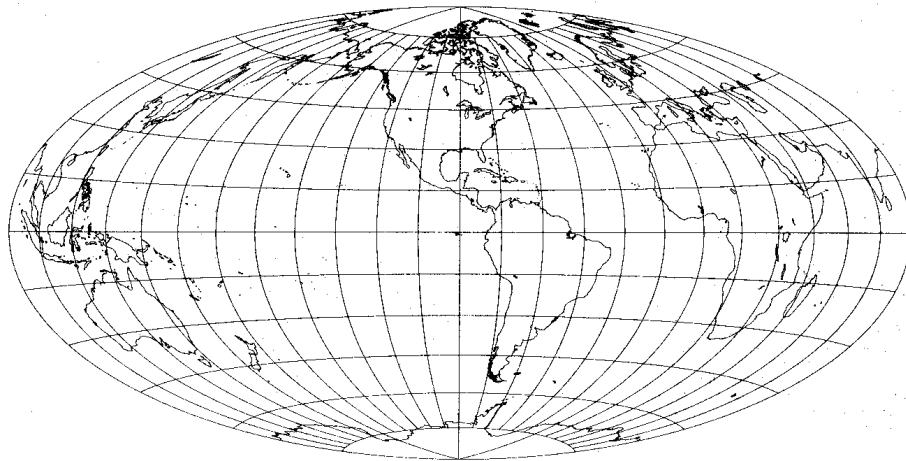
Longitude of central meridian

Enter a value for longitude of the desired central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-15 Hammer Projection

**Interrupted Goode
Homolosine**

Table B-19 Interrupted Goode Homolosine

Construction	Pseudocylindrical
Property	Equal-area
Meridians	“In the interrupted form, there are six central meridians, each a straight line 0.22 as long as the Equator but not crossing the Equator. Other meridians are equally spaced sinusoidal curves between latitudes 40° 44' N and S. and elliptical arcs elsewhere, all concave toward the central meridian. There is a slight bend in meridians at the 40° 44' latitudes” (Snyder and Voxland 1989).
Parallels	Parallels are straight parallel lines, which are perpendicular to the central meridians. Between latitudes 40° 44' N and S, they are equally spaced. Parallels gradually get closer together closer to the poles.
Graticule spacing	See Meridians and Parallels. Poles are points. Symmetry is nonexistent in the interrupted form.
Linear scale	Scale is true at each latitude between 40° 44' N and S along the central meridian within the same latitude range. Scale varies with increased latitudes.
Uses	This projection is useful for world maps.

Source: Snyder and Voxland 1989

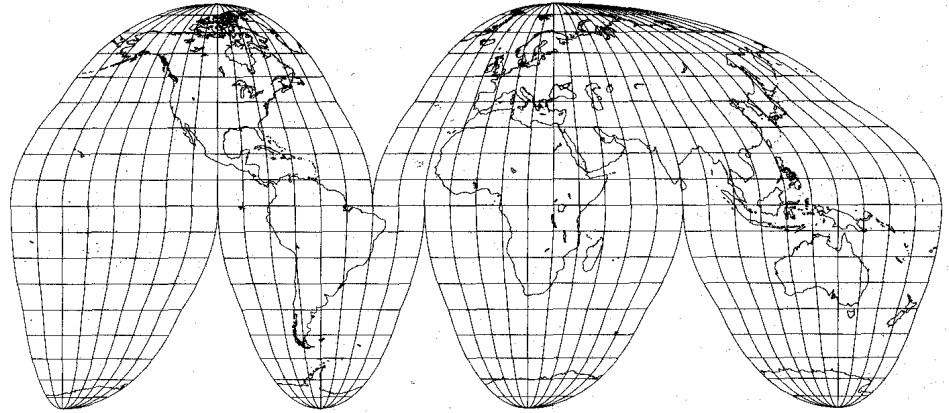
Prompts

The following prompts display in the Projection Chooser once Interrupted Goode Homolosine is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



Source: Snyder and Voxland 1989

Figure B-16 Interrupted Goode Homolosine Projection

Interrupted Mollweide

The interrupted Mollweide projection reduces the distortion of the Mollweide projection. It is interrupted into six regions with fixed parameters for each region.

Source: Snyder and Voxland 1989



For information on the Mollweide projection, see page 529.

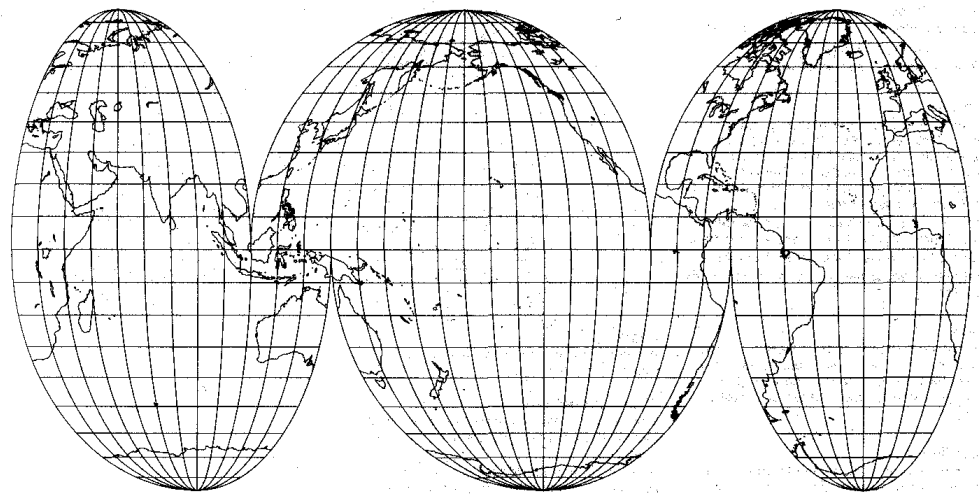
Prompts

The following prompts display in the Projection Chooser once Interrupted Mollweide is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



Source: Snyder and Voxland 1989

Figure B-17 Interrupted Mollweide Projection

Lambert Azimuthal Equal Area

Table B-20 Lambert Azimuthal Equal Area Summary

Construction	Plane
Property	Equal-area
Meridians	Polar aspect: the meridians are straight lines radiating from the point of tangency. Oblique and equatorial aspects: meridians are complex curves concave toward a straight central meridian, except the outer meridian of a hemisphere, which is a circle.
Parallels	Polar aspect: parallels are concentric circles. Oblique and equatorial aspects: the parallels are complex curves. The Equator on the equatorial aspect is a straight line.
Graticule spacing	Polar aspect: the meridian spacing is equal and increases, and the parallel spacing is unequal and decreases toward the periphery of the projection. The graticule spacing, in all aspects, retains the property of equivalence of area.
Linear scale	Linear scale is better than most azimuthals, but not as good as the equidistant. Angular deformation increases toward the periphery of the projection. Scale decreases radially toward the periphery of the map projection. Scale increases perpendicular to the radii toward the periphery.
Uses	The polar aspect is used by the USGS in the National Atlas. The polar, oblique, and equatorial aspects are used by the USGS for the Circum-Pacific Map.

Map Projections

The Lambert Azimuthal Equal Area projection is mathematically based on a plane tangent to the Earth. It is the only projection that can accurately represent both area and true direction from the center of the projection (Figure B-18). This central point can be located anywhere. Concentric circles are closer together toward the edge of the map, and the scale distorts accordingly. This projection is well-suited to square or round land masses. This projection generally represents only one hemisphere.

In the polar aspect, latitude rings decrease their intervals from the center outwards. In the equatorial aspect, parallels are curves flattened in the middle. Meridians are also curved, except for the central meridian, and spacing decreases toward the edges.

Prompts

The following prompts display in the Projection Chooser if Lambert Azimuthal Equal Area is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the center of the map projection in both spherical and rectangular coordinates.

Longitude of center of projection

Latitude of center of projection

Enter values for the longitude and latitude of the desired center of the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

In Figure B-18, three views of the Lambert Azimuthal Equal Area projection are shown: A) Polar aspect, showing one hemisphere; B) Equatorial aspect, frequently used in old atlases for maps of the eastern and western hemispheres; C) Oblique aspect, centered on 40°N.

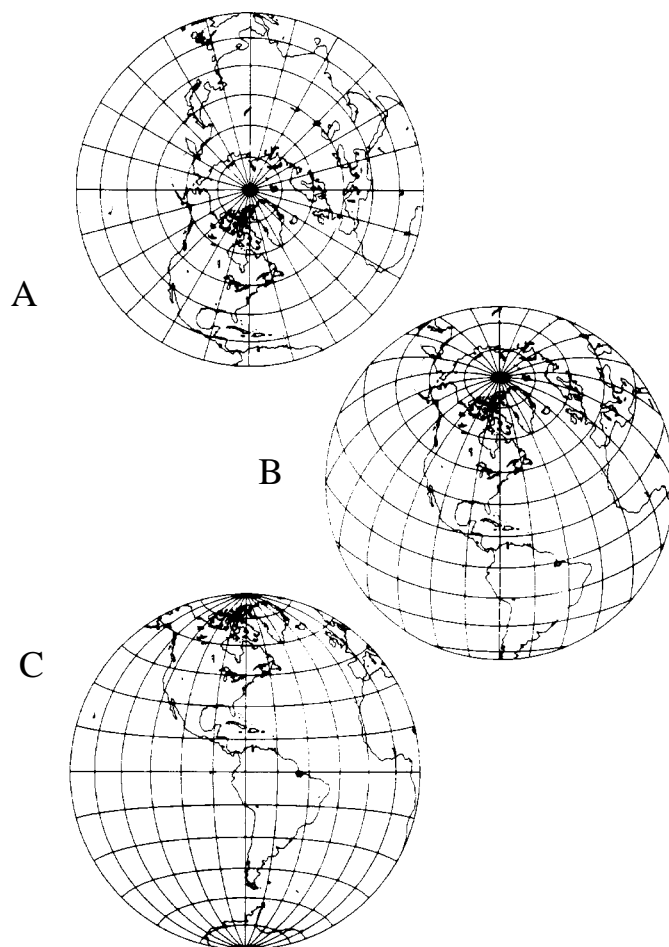


Figure B-18 Lambert Azimuthal Equal Area Projection

Lambert Conformal Conic

Table B-21 Lambert Conformal Conic Summary

Construction	Cone
Property	Conformal
Meridians	Meridians are straight lines converging at a pole.
Parallels	Parallels are arcs of concentric circles concave toward a pole and centered at a pole.
Graticule spacing	Meridian spacing is true on the standard parallels and decreases toward the pole. Parallel spacing increases away from the standard parallels and decreases between them. Meridians and parallels intersect each other at right angles. The graticule spacing retains the property of conformality. The graticule is symmetrical.
Linear scale	Linear scale is true on standard parallels. Maximum scale error is 2.5% on a map of the United States (48 states) with standard parallels at 33°N and 45°N.
Uses	Used for large countries in the mid-latitudes having an east-west orientation. The United States (50 states) Base Map uses standard parallels at 37°N and 65°N. Some of the National Topographic Map Series 7.5-minute and 15-minute quadrangles, and the State Base Map series are constructed on this projection. The latter series uses standard parallels of 33°N and 45°N. Aeronautical charts for Alaska use standard parallels at 55°N and 65°N. The National Atlas of Canada uses standard parallels at 49°N and 77°N.

This projection is very similar to Albers Conical Equal Area, described previously. It is mathematically based on a cone that is tangent at one parallel or, more often, that is conceptually secant on two parallels (Figure B-19). Areal distortion is minimal, but increases away from the standard parallels. North or South Pole is represented by a point—the other pole cannot be shown. Great circle lines are approximately straight. It retains its properties at various scales, and sheets can be joined along their edges. This projection, like Albers, is most valuable in middle latitudes, especially in a country sprawling east to west like the US. The standard parallels for the US are 33° and 45°N.

The major property of this projection is its conformality. At all coordinates, meridians and parallels cross at right angles. The correct angles produce correct shapes. Also, great circles are approximately straight. The conformal property of Lambert Conformal Conic, and the straightness of great circles makes it valuable for landmark flying.

Lambert Conformal Conic is the State Plane coordinate system projection for states of predominant east-west expanse. Since 1962, Lambert Conformal Conic has been used for the International Map of the World between 84°N and 80°S.

In comparison with Albers Conical Equal Area, Lambert Conformal Conic possesses true shape of small areas, whereas Albers possesses equal-area. Unlike Albers, parallels of Lambert Conformal Conic are spaced at increasing intervals the farther north or south they are from the standard parallels.

Prompts

The following prompts display in the Projection Chooser if Lambert Conformal Conic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Latitude of 1st standard parallel

Latitude of 2nd standard parallel

Enter two values for the desired control lines of the projection, i.e., the standard parallels. Note that the first standard parallel is the southernmost.

Then, define the origin of the map projection in both spherical and rectangular coordinates.

Longitude of central meridian

Latitude of origin of projection

Enter values for longitude of the desired central meridian and latitude of the origin of projection.

False easting at central meridian

False northing at origin

Enter values of false easting and false northing corresponding to the intersection of the central meridian, and the latitude of the origin of projection. These values must be in meters. It is often convenient to make them large enough to ensure that there are no negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Figure B-19 Lambert Conformal Conic Projection

In Figure B-19, the standard parallels are 20°N and 60°N. Note the change in spacing of the parallels.

Loximuthal

Table B-22 Loximuthal Summary

Construction	Pseudocylindrical
Property	Neither conformal nor equal-area
Meridians	The “central meridian is a straight line generally over half as long as the Equator, depending on the central latitude. If the central latitude is the Equator, the ration is 0.5; if it is 40° N or S, the ratio is 0.65. Other meridians are equally spaced complex curves intersecting at the poles and concave toward the central meridian” (Snyder and Voxland 1989).
Parallels	Parallels are straight parallel lines that are equally spaced. They are perpendicular to the central meridian.
Graticule spacing	See Meridians and Parallels. Poles are points. Symmetry exists about the central meridian. Symmetry also exists at the Equator if it is designated as the central latitude.
Linear scale	Scale is true along the central meridian. Scale is also constant along any given latitude, but different for the latitude of opposite sign.
Uses	Used for world maps where loxodromes (rhumb lines) are emphasized.

The distortion of the Loximuthal projection is average to pronounced. Distortion is not present at the central latitude on the central meridian. What is most noteworthy about the loximuthal projection is the loxodromes that are “straight, true to scale, and correct in azimuth from the center” (Snyder and Voxland 1989).

Source: Snyder and Voxland 1989

Prompts

The following prompts display in the Projection Chooser if Loximuthal is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Longitude of central meridian

Latitude of central parallel

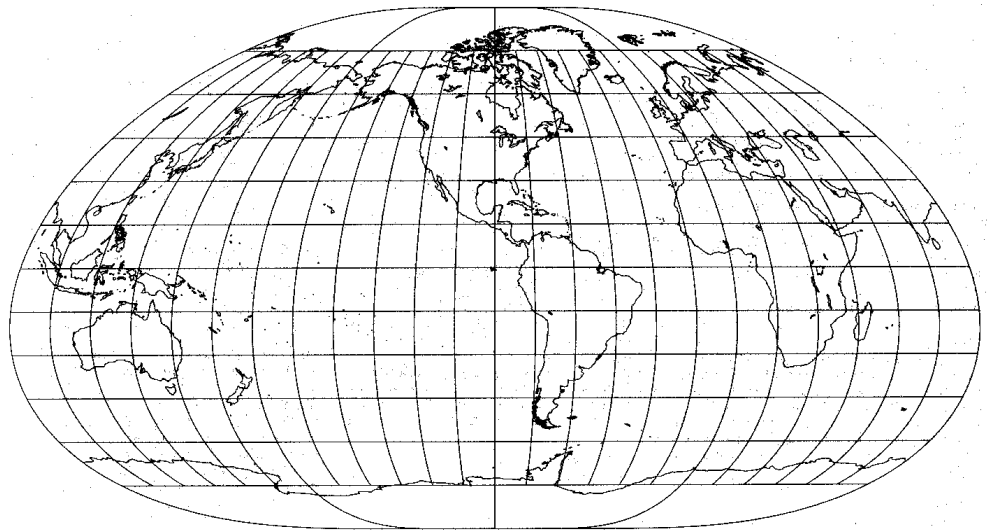
Map Projections

Enter a value for longitude of the desired central meridian to center the projection and the latitude of central parallel.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-20 Loximuthal Projection

Mercator

Table B-23 Mercator Summary

Construction	Cylinder
Property	Conformal
Meridians	Meridians are straight and parallel.
Parallels	Parallels are straight and parallel.
Graticule spacing	Meridian spacing is equal and the parallel spacing increases away from the Equator. The graticule spacing retains the property of conformality. The graticule is symmetrical. Meridians intersect parallels at right angles.
Linear scale	Linear scale is true along the Equator only (line of tangency), or along two parallels equidistant from the Equator (the secant form). Scale can be determined by measuring one degree of latitude, which equals 60 nautical miles, 69 statute miles, or 111 kilometers.
Uses	An excellent projection for equatorial regions. Otherwise, the Mercator is a special-purpose map projection best suited for navigation. Secant constructions are used for large-scale coastal charts. The use of the Mercator map projection as the base for nautical charts is universal. Examples are the charts published by the National Ocean Survey, US Dept. of Commerce.

This famous cylindrical projection was originally designed by Flemish map maker Gerhardus Mercator in 1569 to aid navigation (Figure B-21). Meridians and parallels are straight lines and cross at 90° angles. Angular relationships are preserved. However, to preserve conformality, parallels are placed increasingly farther apart with increasing distance from the Equator. Due to extreme scale distortion in high latitudes, the projection is rarely extended beyond 80°N or S unless the latitude of true scale is other than the Equator. Distance scales are usually furnished for several latitudes.

This projection can be thought of as being mathematically based on a cylinder tangent at the Equator. Any straight line is a constant-azimuth (rhumb) line. Areal enlargement is extreme away from the Equator; poles cannot be represented. Shape is true only within any small area. It is a reasonably accurate projection within a 15° band along the line of tangency.

Rhumb lines, which show constant direction, are straight. For this reason, a Mercator map was very valuable to sea navigators. However, rhumb lines are not the shortest path—great circles are the shortest path. Most great circles appear as long arcs when drawn on a Mercator map.

Map Projections

Prompts

The following prompts display in the Projection Chooser if Mercator is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the origin of the map projection in both spherical and rectangular coordinates.

Longitude of central meridian

Latitude of true scale

Enter values for longitude of the desired central meridian and latitude at which true scale is desired. Selection of a parameter other than the Equator can be useful for making maps in extreme north or south latitudes.

False easting at central meridian

False northing at origin

Enter values of false easting and false northing corresponding to the intersection of the central meridian and the latitude of true scale. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

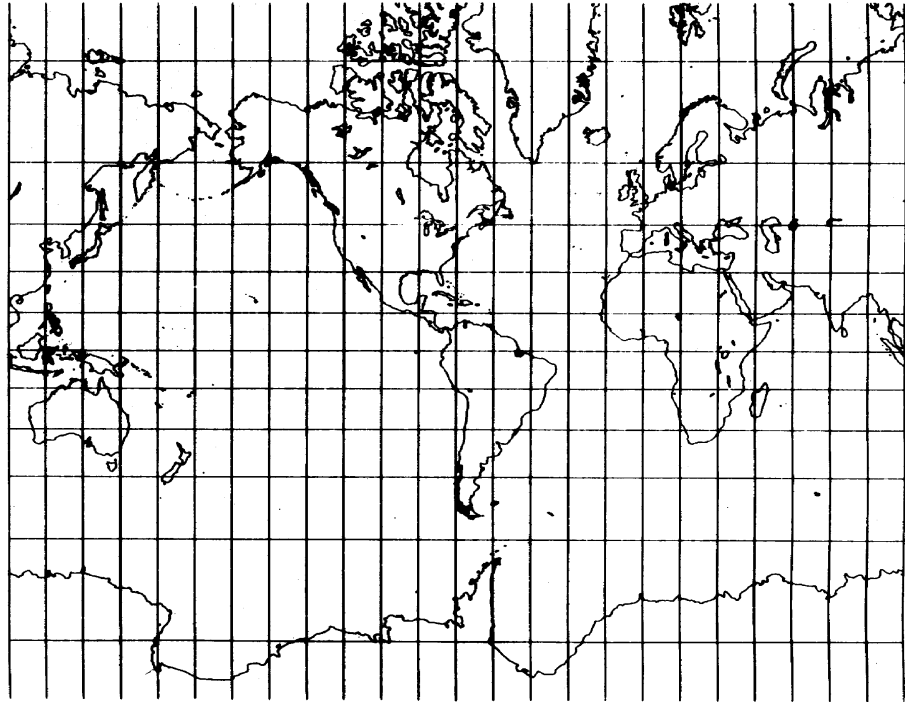


Figure B-21 Mercator Projection

In Figure B-21, all angles are shown correctly, therefore small shapes are true (i.e., the map is conformal). Rhumb lines are straight, which makes it useful for navigation.

Miller Cylindrical

Table B-24 Miller Cylindrical Summary

Construction	Cylinder
Property	Compromise
Meridians	All meridians are straight lines.
Parallels	All parallels are straight lines.
Graticule spacing	Meridians are parallel and equally spaced, the lines of latitude are parallel, and the distance between them increases toward the poles. Both poles are represented as straight lines. Meridians and parallels intersect at right angles (ESRI 1992).
Linear scale	While the standard parallels, or lines, that are true to scale and free of distortion, are at latitudes 45°N and S, only the Equator is standard.
Uses	Useful for world maps.

Map Projections

Miller Cylindrical is a modification of the Mercator projection (Figure B-22). It is similar to the Mercator from the Equator to 45°, but latitude line intervals are modified so that the distance between them increases less rapidly. Thus, beyond 45°, Miller Cylindrical lessens the extreme exaggeration of the Mercator. Miller Cylindrical also includes the poles as straight lines whereas the Mercator does not.

Meridians and parallels are straight lines intersecting at right angles. Meridians are equidistant, while parallels are spaced farther apart the farther they are from the Equator. Miller Cylindrical is not equal-area, equidistant, or conformal. Miller Cylindrical is used for world maps and in several atlases.

Prompts

The following prompts display in the Projection Chooser if Miller Cylindrical is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Longitude of central meridian

Enter a value for the longitude of the desired central meridian to center the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

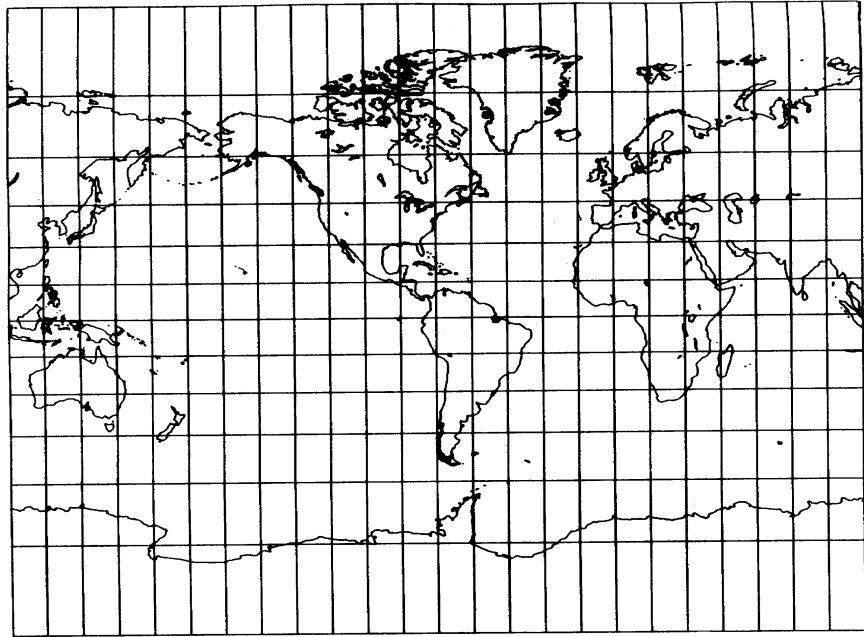


Figure B-22 Miller Cylindrical Projection

This projection resembles the Mercator, but has less distortion in polar regions. Miller Cylindrical is neither conformal nor equal-area.

Map Projections

Modified Transverse Mercator

Table B-25 Modified Transverse Mercator Summary

Construction	Cone
Property	Equidistant
Meridians	On pre-1973 editions of the Alaska Map E, meridians are curved concave toward the center of the projection. On post-1973 editions, the meridians are straight.
Parallels	Parallels are arcs concave to the pole.
Graticule spacing	Meridian spacing is approximately equal and decreases toward the pole. Parallels are approximately equally spaced. The graticule is symmetrical on post-1973 editions of the Alaska Map E.
Linear scale	Linear scale is more nearly correct along the meridians than along the parallels.
Uses	USGS's Alaska Map E at the scale of 1:2,500,000. The Bathymetric Maps Eastern Continental Margin U.S.A., published by the American Association of Petroleum Geologists, uses the straight meridians on its Modified Transverse Mercator and is more equivalent to the Equidistant Conic map projection.

In 1972, the USGS devised a projection specifically for the revision of a 1954 map of Alaska which, like its predecessors, was based on the Polyconic projection. This projection was drawn to a scale of 1:2,000,000 and published at 1:2,500,000 (map "E") and 1:1,584,000 (map "B"). Graphically prepared by adapting coordinates for the UTM projection, it is identified as the Modified Transverse Mercator projection. It resembles the Transverse Mercator in a very limited manner and cannot be considered a cylindrical projection. It resembles the Equidistant Conic projection for the ellipsoid in actual construction. The projection was also used in 1974 for a base map of the Aleutian-Bering Sea Region published at 1:2,500,000 scale.

It is found to be most closely equivalent to the Equidistant Conic for the Clarke 1866 ellipsoid, with the scale along the meridians reduced to 0.9992 of true scale, and the standard parallels at latitude 66.09° and 53.50°N.

Prompts

The following prompts display in the Projection Chooser if Modified Transverse Mercator is selected. Respond to the prompts as described.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Mollweide

Table B-26 Mollweide Summary

Construction	Pseudocylinder
Property	Equal-area
Meridians	Meridians are elliptical arcs that are equally spaced. The exception is the central meridian, which is a straight line.
Parallels	All parallels are straight lines.
Graticule spacing	The Equator and the central meridian are linear graticules.
Linear scale	Scale is accurate along latitudes 40° 44' N and S at the central meridian. Distortion becomes more pronounced farther from the lines, and is severe at the extremes of the projection.
Uses	Use for world maps only.

Carl B. Mollweide designed the projection in 1805. It is an equal-area pseudocylindrical projection. The Mollweide projection is used primarily for thematic maps of the world.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Mollweide is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Longitude of central meridian

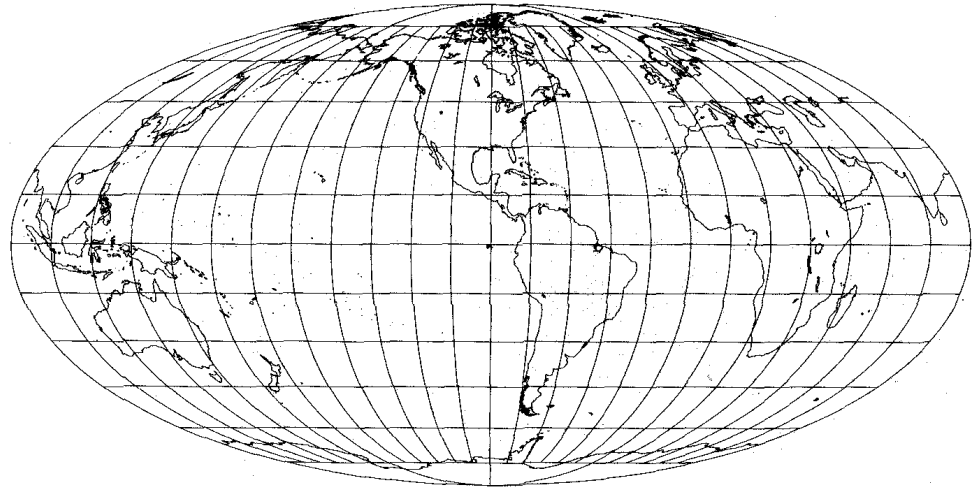
Enter a value for longitude of the desired central meridian.

False easting

False northing

Map Projections

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-23 Mollweide Projection

New Zealand Map Grid

Table B-27 New Zealand Map Grid Summary

Construction	Modified cylindrical
Property	Conformal
Meridians	N/A
Parallels	N/A
Graticule spacing	None
Linear scale	Scale is within 0.02 percent of actual scale for the country of New Zealand.
Uses	This projection is useful only for maps of New Zealand.

This projection is used only for mapping New Zealand.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once New Zealand Map Grid is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

The Spheroid Name defaults to International 1909. The Datum Name defaults to Geodetic Datum 1949. These fields are not editable.

Easting Shift

Northing Shift

The Easting and Northing shifts are reported in meters.

Oblated Equal Area**Prompts**

The following prompts display in the Projection Chooser once Oblated Equal Area is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.

Parameter M

Parameter N

Enter the obliterated equal area oval shape of parameters M and N.

Longitude of center of projection

Latitude of center of projection

Enter the longitude of the center of the projection and the latitude of the center of the projection.

Rotation angle

Enter the obliterated equal area oval rotation angle.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Map Projections

Oblique Mercator (Hotine)

Table B-28 Oblique Mercator (Hotine) Summary

Construction	Cylinder
Property	Conformal
Meridians	Meridians are complex curves concave toward the line of tangency, except each 180th meridian is straight.
Parallels	Parallels are complex curves concave toward the nearest pole.
Graticule spacing	Graticule spacing increases away from the line of tangency and retains the property of conformality.
Linear scale	Linear scale is true along the line of tangency, or along two lines of equidistance from and parallel to the line of tangency.
Uses	Useful for plotting linear configurations that are situated along a line oblique to the Earth's Equator. Examples are: NASA Surveyor Satellite tracking charts, ERTS flight indexes, strip charts for navigation, and the National Geographic Society's maps "West Indies," "Countries of the Caribbean," "Hawaii," and "New Zealand."

Oblique Mercator is a cylindrical, conformal projection that intersects the global surface along a great circle. It is equivalent to a Mercator projection that has been altered by rotating the cylinder so that the central line of the projection is a great circle path instead of the Equator. Shape is true only within any small area. Areal enlargement increases away from the line of tangency. Projection is reasonably accurate within a 15° band along the line of tangency.

The USGS uses the Hotine version of Oblique Mercator. The Hotine version is based on a study of conformal projections published by British geodesist Martin Hotine in 1946-47. Prior to the implementation of the Space Oblique Mercator, the Hotine version was used for mapping Landsat satellite imagery.

Prompts

The following prompts display in the Projection Chooser if Oblique Mercator (Hotine) is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Scale factor at center

Designate the desired scale factor along the central line of the projection. This parameter may be used to modify scale distortion away from this central line. A value of 1.0 indicates true scale only along the central line. A value of less than, but close to one is often used to lessen scale distortion away from the central line.

Latitude of point of origin

False easting

False northing

The center of the projection is defined by rectangular coordinates of false easting and false northing. The origin of rectangular coordinates on this projection occurs at the nearest intersection of the central line with the Earth's Equator. To shift the origin to the intersection of the latitude of the origin entered above and the central line of the projection, compute coordinates of the latter point with zero false eastings and northings, reverse the signs of the coordinates obtained, and use these for false eastings and northings. These values must be in meters.

It is often convenient to add additional values so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Do you want to enter either:

A) Azimuth East of North for central line and the longitude of the point of origin

B) The latitude and longitude of the first and second points defining the central line

These formats differ slightly in definition of the central line of the projection.

Format A

For format A, the additional prompts are:

Azimuth east of north for central line

Longitude of point of origin

Format A defines the central line of the projection by the angle east of north to the desired great circle path and by the latitude and longitude of the point along the great circle path from which the angle is measured. Appropriate values should be entered.

Map Projections

Format B

For format B, the additional prompts are:

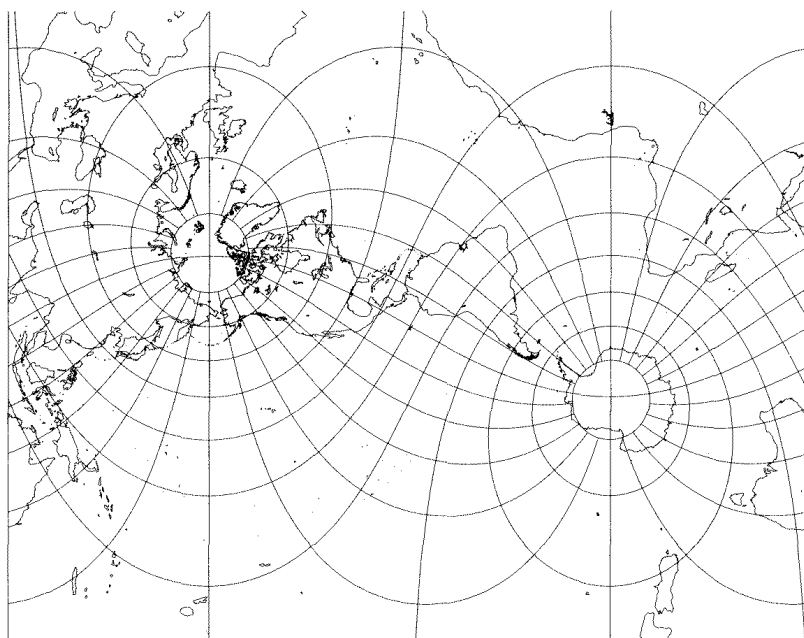
Longitude of 1st point

Latitude of 1st point

Longitude of 2nd point

Latitude of 2nd point

Format B defines the central line of the projection by the latitude of a point on the central line which has the desired scale factor entered previously and by the longitude and latitude of two points along the desired great circle path. Appropriate values should be entered.



Source: Snyder and Voxland 1989

Figure B-24 Oblique Mercator Projection

Orthographic

Table B-29 Orthographic Summary

Construction	Plane
Property	Compromise
Meridians	<p>Polar aspect: the meridians are straight lines radiating from the point of tangency.</p> <p>Oblique aspect: the meridians are ellipses, concave toward the center of the projection.</p> <p>Equatorial aspect: the meridians are ellipses concave toward the straight central meridian.</p>
Parallels	<p>Polar aspect: the parallels are concentric circles.</p> <p>Oblique aspect: the parallels are ellipses concave toward the poles.</p> <p>Equatorial aspect: the parallels are straight and parallel.</p>
Graticule spacing	<p>Polar aspect: meridian spacing is equal and increases, and the parallel decreases from the point of tangency.</p> <p>Oblique and equatorial aspects: the graticule spacing decreases away from the center of the projection.</p>
Linear scale	Scale is true on the parallels in the polar aspect and on all circles centered at the pole of the projection in all aspects. Scale decreases along lines radiating from the center of the projection.
Uses	USGS uses the Orthographic map projection in the National Atlas.

The Orthographic projection is geometrically based on a plane tangent to the Earth, and the point of projection is at infinity (Figure B-25). The Earth appears as it would from outer space. Light rays that cast the projection are parallel and intersect the tangent plane at right angles. This projection is a truly graphic representation of the Earth, and is a projection in which distortion becomes a visual aid. It is the most familiar of the azimuthal map projections. Directions from the center of the projection are true.

This projection is limited to one hemisphere and shrinks those areas toward the periphery. In the polar aspect, latitude ring intervals decrease from the center outwards at a much greater rate than with Lambert Azimuthal. In the equatorial aspect, the central meridian and parallels are straight, with spaces closing up toward the outer edge.

The Orthographic projection seldom appears in atlases. Its utility is more pictorial than technical. Orthographic has been used as a basis for artistic maps by Rand McNally and the USGS.

Map Projections

Prompts

The following prompts display in the Projection Chooser if Orthographic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the center of the map projection in both spherical and rectangular coordinates.

Longitude of center of projection

Latitude of center of projection

Enter values for the longitude and latitude of the desired center of the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Three views of the Orthographic projection are shown in Figure B-25: A) Polar aspect; B) Equatorial aspect; C) Oblique aspect, centered at 40°N and showing the classic globe-like view.

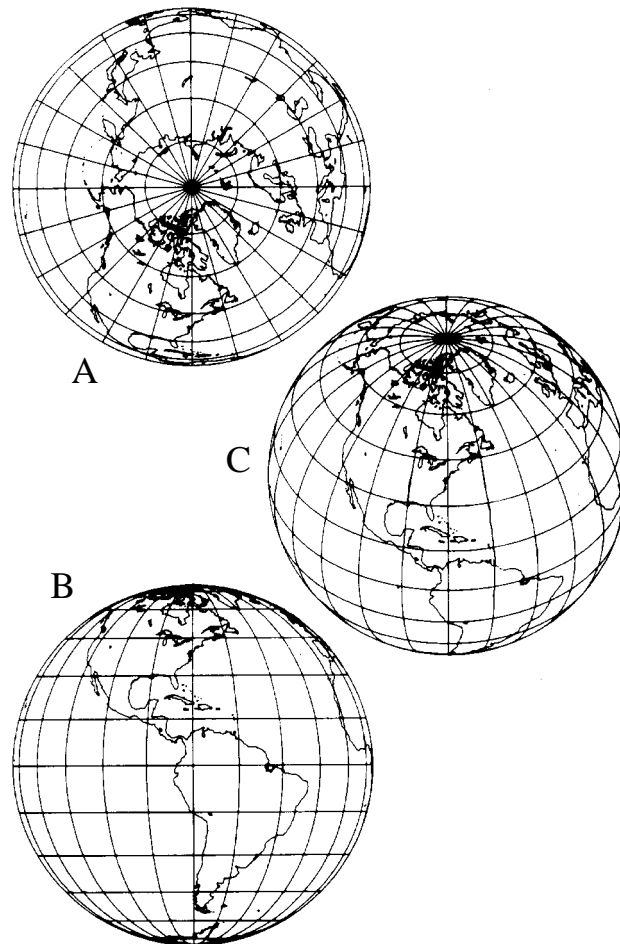


Figure B-25 Orthographic Projection

Plate Carrée

The parameters for the Plate Carrée projection are identical to that of the Equirectangular projection.



For information on the Equirectangular projection, see page 503.

Prompts

The following prompts display in the Projection Chooser if Plate Carrée is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Longitude of central meridian

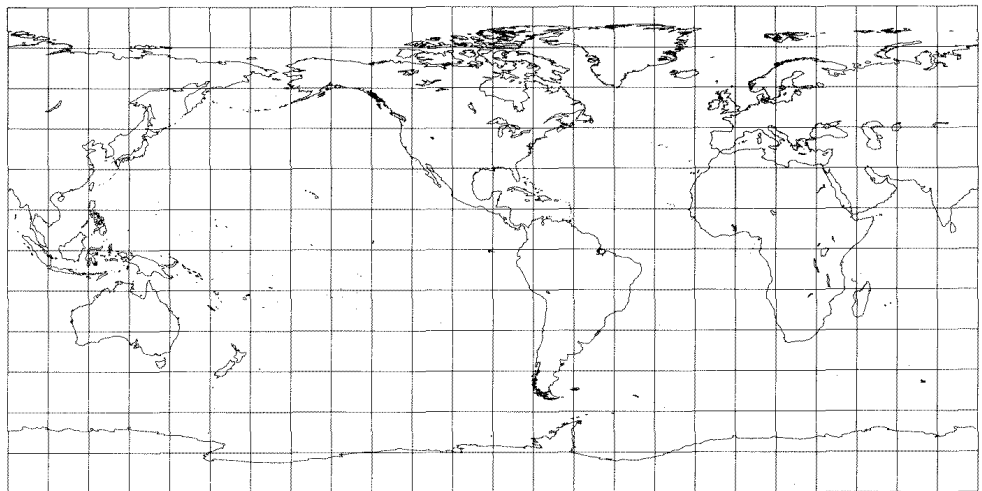
Latitude of true scale

Enter a value for longitude of the desired central meridian to center the projection and the latitude of true scale.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-26 Plate Carrée Projection

Polar Stereographic

Table B-30 Polar Stereographic Summary

Construction	Plane
Property	Conformal
Meridians	Meridians are straight.
Parallels	Parallels are concentric circles.
Graticule spacing	The distance between parallels increases with distance from the central pole.
Linear scale	The scale increases with distance from the center. If a standard parallel is chosen rather than one of the poles, this latitude represents the true scale, and the scale nearer the pole is reduced.
Uses	Polar regions (conformal). In the Universal Polar Stereographic (UPS) system, the scale factor at the pole is made 0.994, thus making the standard parallel (latitude of true scale) approximately 81°07'N or S.

The Polar Stereographic may be used to accommodate all regions not included in the UTM coordinate system, regions north of 84°N and 80°S. This form is called Universal Polar Stereographic (UPS). The projection is equivalent to the polar aspect of the Stereographic projection on a spheroid. The central point is either the North Pole or the South Pole. Of all the polar aspect planar projections, this is the only one that is conformal.

The point of tangency is a single point—either the North Pole or the South Pole. If the plane is secant instead of tangent, the point of global contact is a line of latitude (ESRI 1992).

Polar Stereographic is an azimuthal projection obtained by projecting from the opposite pole (Figure B-27). All of either the northern or southern hemispheres can be shown, but not both. This projection produces a circular map with one of the poles at the center.

Polar Stereographic stretches areas toward the periphery, and scale increases for areas farther from the central pole. Meridians are straight and radiating; parallels are concentric circles. Even though scale and area are not constant with Polar Stereographic, this projection, like all stereographic projections, possesses the property of conformality.

The Astrogeology Center of the Geological Survey at Flagstaff, Arizona, has been using the Polar Stereographic projection for the mapping of polar areas of every planet and satellite for which there is sufficient information.

Map Projections

Prompts

The following prompts display in the Projection Chooser if Polar Stereographic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the origin of the map projection in both spherical and rectangular coordinates. Ellipsoid projections of the polar regions normally use the International 1909 spheroid (ESRI 1992).

Longitude below pole of map

Enter a value for longitude directed straight down below the pole for a north polar aspect, or straight up from the pole for a south polar aspect. This is equivalent to centering the map with a desired meridian.

Latitude of true scale

Enter a value for latitude at which true scale is desired. For secant projections, specify the latitude of true scale as any line of latitude other than 90°N or S. For tangential projections, specify the latitude of true scale as the North Pole, 90 00 00, or the South Pole, -90 00 00 (ESRI 1992).

False easting

False northing

Enter values of false easting and false northing corresponding to the pole. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

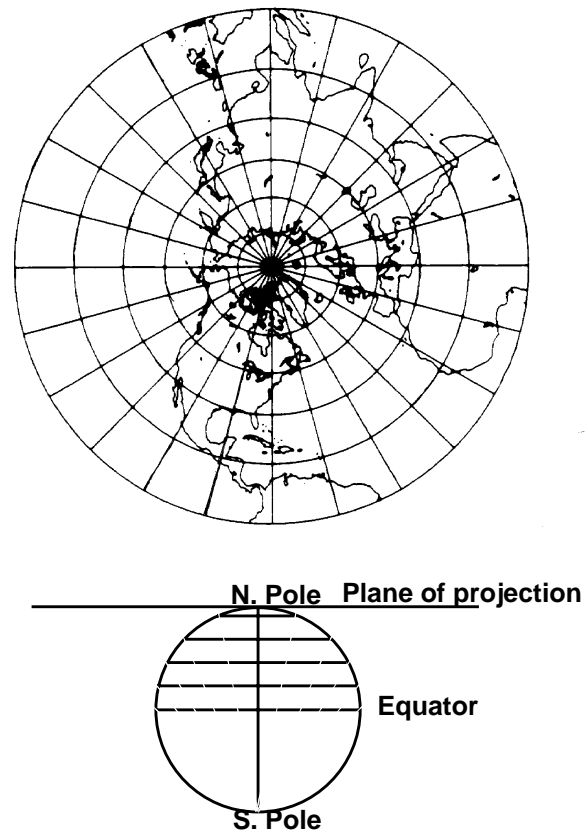


Figure B-27 Polar Stereographic Projection and its Geometric Construction
This projection is conformal and is the most scientific projection for polar regions.

Polyconic

Table B-31 Polyconic Summary

Construction	Cone
Property	Compromise
Meridians	The central meridian is a straight line, but all other meridians are complex curves.
Parallels	Parallels (except the Equator) are nonconcentric circular arcs. The Equator is a straight line.
Graticule spacing	All parallels are arcs of circles, but not concentric. All meridians, except the central meridian, are concave toward the central meridian. Parallels cross the central meridian at equal intervals but get farther apart at the east and west peripheries.
Linear scale	The scale along each parallel and along the central meridian of the projection is accurate. It increases along the meridians as the distance from the central meridian increases (ESRI 1992).
Uses	Used for 7.5-minute and 15-minute topographic USGS quad sheets, from 1886 to about 1957 (ESRI 1992). Used almost exclusively in slightly modified form for large-scale mapping in the United States until the 1950s.

Polyconic was developed in 1820 by Ferdinand Hassler specifically for mapping the eastern coast of the US (Figure B-28). Polyconic projections are made up of an infinite number of conic projections tangent to an infinite number of parallels. These conic projections are placed in relation to a central meridian. Polyconic projections compromise properties such as equal-area and conformality, although the central meridian is held true to scale.

This projection is used mostly for north-south oriented maps. Distortion increases greatly the farther east and west an area is from the central meridian.

Prompts

The following prompts display in the Projection Chooser if Polyconic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the origin of the map projection in both spherical and rectangular coordinates.

Longitude of central meridian

Latitude of origin of projection

Enter values for longitude of the desired central meridian and latitude of the origin of projection.

False easting at central meridian

False northing at origin

Enter values of false easting and false northing corresponding to the intersection of the central meridian and the latitude of the origin of projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

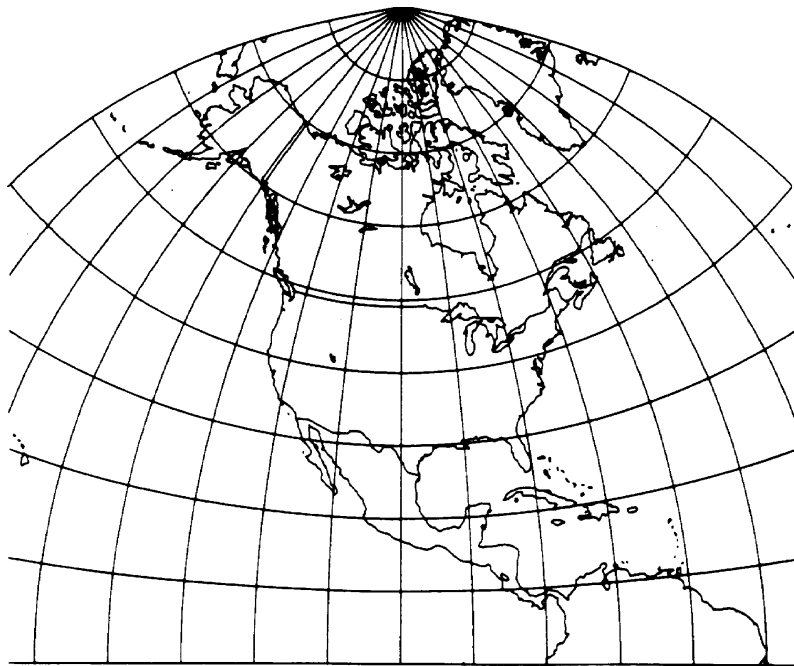


Figure B-28 Polyconic Projection of North America

In Figure B-28, the central meridian is 100°W . This projection is used by the USGS for topographic quadrangle maps.

Map Projections

Quartic Authalic

Table B-32 Quartic Authalic

Construction	Pseudocylindrical
Property	Equal-area
Meridians	The central meridian is a straight line, and is 0.45 as long as the Equator. The other meridians are curves that are equally spaced. They fit a “fourth-order (quartic) equation and concave toward the central meridian” (Snyder and Voxland 1989).
Parallels	Parallels are straight parallel lines that are unequally spaced. The parallels have the greatest distance between in proximity to the Equator. Parallel spacing changes slowly, and parallels are perpendicular to the central meridian.
Graticule spacing	See Meridians and Parallels. Poles are points. Symmetry exists about the central meridian or the Equator.
Linear scale	Scale is accurate along the Equator. Scale is constant along each latitude, and is the same for the latitude of opposite sign.
Uses	The McBryde-Thomas Flat-Polar Quartic projection uses Quartic Authalic as its base (Snyder and Voxland 1989). Used for world maps.

Outer meridians at high latitudes have great distortion. If the Quartic Authalic projection is interrupted, distortion can be reduced.

Source: Snyder and Voxland 1989

Prompts

The following prompts display in the Projection Chooser once Quartic Authalic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

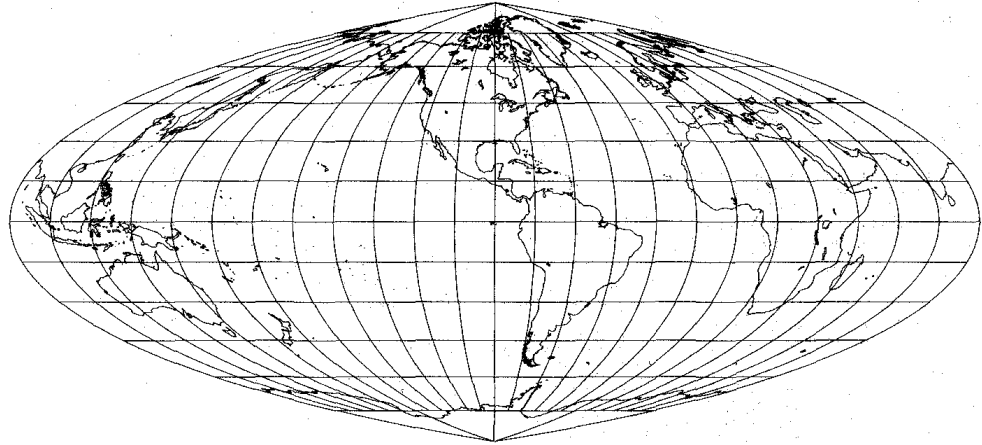
Longitude of central meridian

Enter the value of the longitude of the central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-29 Quartic Authalic Projection

Robinson

Table B-33 Robinson Summary

Construction	Pseudocylinder
Property	Neither conformal nor equal-area
Meridians	Meridians are equally spaced, and concave toward the central meridian, and look like elliptical arcs (ESRI 1997).
Parallels	Parallels are equally spaced straight lines between 38° N and S.
Graticule spacing	The central meridian and all parallels are linear.
Linear scale	Scale is true along latitudes 38° N and S. Scale is constant along any specific latitude, and for the latitude of opposite sign.
Uses	Useful for thematic and common world maps.

According to ESRI, the Robinson “central meridian is a straight line 0.51 times the length of the Equator. Parallels are equally spaced straight lines between 38° N and S; spacing decreases beyond these limits. The poles are 0.53 times the length of the Equator. The projection is based upon tabular coordinates instead of mathematical formulas” (ESRI 1997).

This projection has been used both by Rand McNally and the National Geographic Society.

Source: ESRI 1997

Map Projections

Prompts

The following prompts display in the Projection Chooser once Robinson is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

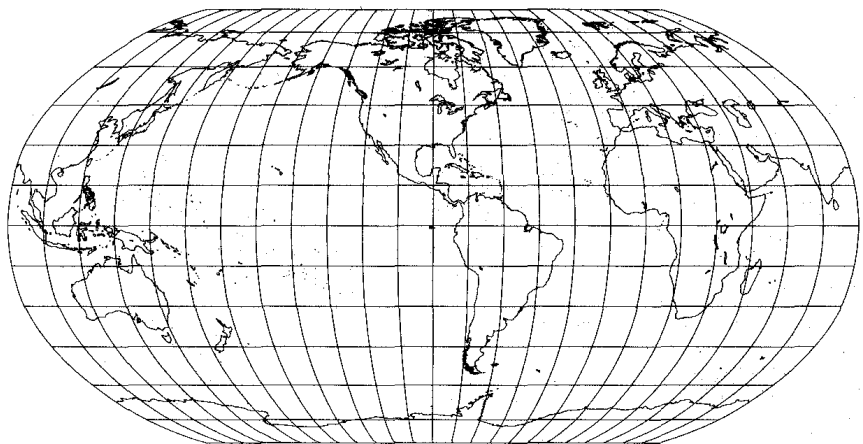
Longitude of central meridian

Enter the value of the longitude of the central meridian.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-30 Robinson Projection

RSO

Table B-34 RSO Summary

Construction	Cylinder
Property	Conformal
Meridians	Two meridians are 180 degrees apart.
Parallels	N/A
Graticule spacing	Graticules are two meridians 180 degrees apart.
Linear scale	"A line of true scale is drawn at an angle to the central meridian" (ESRI 1997).
Uses	This projection should be used to map areas of Brunei and Malaysia.

The acronym RSO stands for Rectified Skewed Orthomorphic. This projection is used to map areas of Brunei and Malaysia, and is each country's national projection.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once RSO is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

RSO Type

Select the RSO Type. You can choose from Borneo or Malaysia.

Sinusoidal

Table B-35 Sinusoidal Summary

Construction	Pseudocylinder
Property	Equal-area
Meridians	Meridians are sinusoidal curves, curved toward a straight central meridian.
Parallels	All parallels are straight, parallel lines.
Graticule spacing	Meridian spacing is equal and decreases toward the poles. Parallel spacing is equal. The graticule spacing retains the property of equivalence of area.
Linear scale	Linear scale is true on the parallels and the central meridian.
Uses	Used as an equal-area projection to portray areas that have a maximum extent in a north-south direction. Used as a world equal-area projection in atlases to show distribution patterns. Used by the USGS as the base for maps showing prospective hydrocarbon provinces of the world, and sedimentary basins of the world.

Sometimes called the Sanson-Flamsteed, Sinusoidal is a projection with some characteristics of a cylindrical projection—often called a pseudocylindrical type. The central meridian is the only straight meridian—all others become sinusoidal curves. All parallels are straight and the correct length. Parallels are also the correct distance from the Equator, which, for a complete world map, is twice as long as the central meridian.

Sinusoidal maps achieve the property of equal-area but not conformality. The Equator and central meridian are distortion free, but distortion becomes pronounced near outer meridians, especially in polar regions.

Interrupting a Sinusoidal world or hemisphere map can lessen distortion. The interrupted Sinusoidal contains less distortion because each interrupted area can be constructed to contain a separate central meridian. Central meridians may be different for the northern and southern hemispheres and may be selected to minimize distortion of continents or oceans.

Sinusoidal is particularly suited for less than world areas, especially those bordering the Equator, such as South America or Africa. Sinusoidal is also used by the USGS as a base map for showing prospective hydrocarbon provinces and sedimentary basins of the world.

Prompts

The following prompts display in the Projection Chooser if Sinusoidal is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

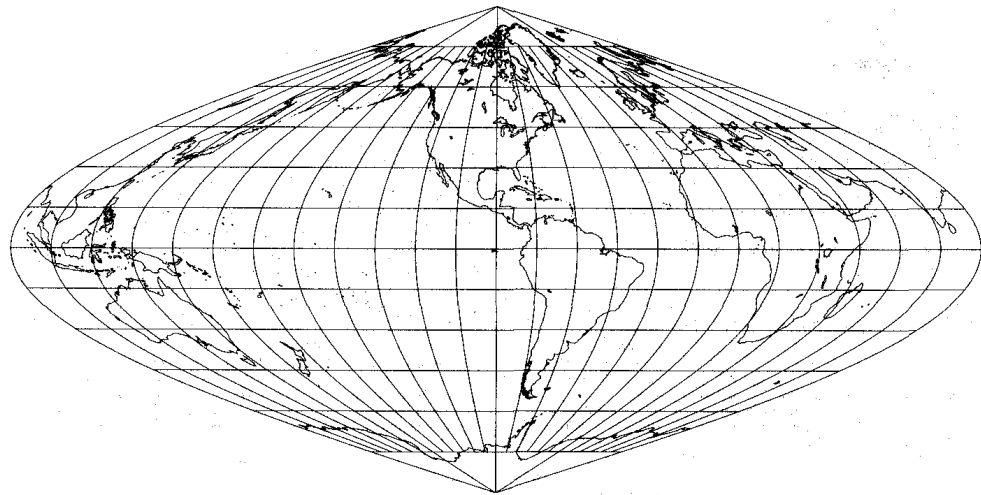
Longitude of central meridian

Enter a value for the longitude of the desired central meridian to center the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-31 Sinusoidal Projection

Space Oblique Mercator

Table B-36 Space Oblique Mercator Summary

Construction	Cylinder
Property	Conformal
Meridians	All meridians are curved lines except for the meridian crossed by the groundtrack at each polar approach.
Parallels	All parallels are curved lines.
Graticule spacing	There are no graticules.
Linear scale	Scale is true along the groundtrack, and varies approximately 0.01% within sensing range (ESRI 1992).
Uses	Used for georectification of, and continuous mapping from, satellite imagery. Standard format for data from Landsats 4 and 5 (ESRI 1992).

The Space Oblique Mercator (SOM) projection is nearly conformal and has little scale distortion within the sensing range of an orbiting mapping satellite such as Landsat. It is the first projection to incorporate the Earth's rotation with respect to the orbiting satellite.

The method of projection used is the modified cylindrical, for which the central line is curved and defined by the groundtrack of the orbit of the satellite. The line of tangency is conceptual and there are no graticules.

The SOM projection is defined by USGS. According to USGS, the X axis passes through the descending node for each daytime scene. The Y axis is perpendicular to the X axis, to form a Cartesian coordinate system. The direction of the X axis in a daytime Landsat scene is in the direction of the satellite motion—south. The Y axis is directed east. For SOM projections used by EOSAT, the axes are switched; the X axis is directed east and the Y axis is directed south.

The SOM projection is specifically designed to minimize distortion within sensing range of a mapping satellite as it orbits the Earth. It can be used for the rectification of, and continuous mapping from, satellite imagery. It is the standard format for data from Landsats 4 and 5. Plots for adjacent paths do not match without transformation (ESRI 1991).

Prompts

The following prompts display in the Projection Chooser if Space Oblique Mercator is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Landsat vehicle ID (1-5)

Specify whether the data are from Landsat 1, 2, 3, 4, or 5.

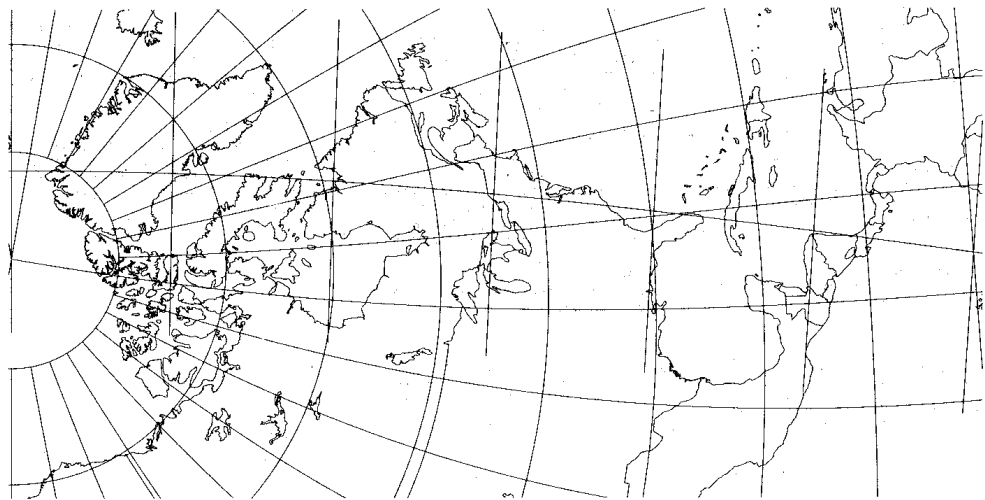
Orbital path number (1-251 or 1-233)

For Landsats 1, 2, and 3, the path range is from 1 to 251. For Landsats 4 and 5, the path range is from 1 to 233.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-32 Space Oblique Mercator Projection

Map Projections

Space Oblique Mercator (Formats A & B)

The Space Oblique Mercator (Formats A&B) projection is similar to the Space Oblique Mercator projection.



For information on the Space Oblique Mercator projection, see "Robinson" on page 550.

Prompts

The following prompts display in the Projection Chooser once Space Oblique Mercator (Formats A & B) is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Format A (Generic Satellite)

Inclination of orbit at ascending node

Period of satellite revolution in minutes

Longitude of ascending orbit at equator

Landsat path flag

If you select Format A of the Space Oblique Mercator projection, you need to supply the information listed above.

Format B (Landsat)

Landsat vehicle ID (1-5)

Specify whether the data are from Landsat 1, 2, 3, 4, or 5.

Path number (1-251 or 1-233)

For Landsats 1, 2, and 3, the path range is from 1 to 251. For Landsats 4 and 5, the path range is from 1 to 233.

State Plane

The State Plane is an X,Y coordinate system (not a map projection); its zones divide the US into over 130 sections, each with its own projection surface and grid network (Figure B-33). With the exception of very narrow states, such as Delaware, New Jersey, and New Hampshire, most states are divided into between two and ten zones. The Lambert Conformal projection is used for zones extending mostly in an east-west direction. The Transverse Mercator projection is used for zones extending mostly in a north-south direction. Alaska, Florida, and New York use either Transverse Mercator or Lambert Conformal for different areas. The Aleutian panhandle of Alaska is prepared on the Oblique Mercator projection.

Zone boundaries follow state and county lines, and, because each zone is small, distortion is less than one in 10,000. Each zone has a centrally located origin and a central meridian that passes through this origin. Two zone numbering systems are currently in use—the USGS code system and the National Ocean Service (NOS) code system (Table B-37 and Table B-38), but other numbering systems exist.

Prompts

The following prompts appear in the Projection Chooser if State Plane is selected. Respond to the prompts as described.

State Plane Zone

Enter either the USGS zone code number as a positive value, or the NOS zone code number as a negative value.

NAD27 or NAD83 or HARN

Either North America Datum 1927 (NAD27), North America Datum 1983 (NAD83), or High Accuracy Reference Network (HARN) may be used to perform the State Plane calculations.

- NAD27 is based on the Clarke 1866 spheroid.
- NAD83 and HARN are based on the GRS 1980 spheroid. Some zone numbers have been changed or deleted from NAD27.

Tables for both NAD27 and NAD83 zone numbers follow (Table B-37 and Table B-38). These tables include both USGS and NOS code systems.



Figure B-33 Zones of the State Plane Coordinate System

The following abbreviations are used in Table B-37 and Table B-38:

- Tr Merc = Transverse Mercator
- Lambert = Lambert Conformal Conic
- Oblique = Oblique Mercator (Hotine)
- Polycon = Polyconic

Table B-37 NAD27 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States

State	Zone Name	Type	Code Number	
			USGS	NOS
Alabama	East	Tr Merc	3101	-101
	West	Tr Merc	3126	-102
Alaska	1	Oblique	6101	-5001
	2	Tr Merc	6126	-5002
	3	Tr Merc	6151	-5003
	4	Tr Merc	6176	-5004
	5	Tr Merc	6201	-5005
	6	Tr Merc	6226	-5006
	7	Tr Merc	6251	-5007
	8	Tr Merc	6276	-5008
	9	Tr Merc	6301	-5009
	10	Lambert	6326	-5010

Table B-37 NAD27 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States (Continued)

State			Code Number	
	Zone Name	Type	USGS	NOS
American Samoa	-----	Lambert	-----	-5302
Arizona	East	Tr Merc	3151	-201
	Central	Tr Merc	3176	-202
	West	Tr Merc	3201	-203
Arkansas	North	Lambert	3226	-301
	South	Lambert	3251	-302
California	I	Lambert	3276	-401
	II	Lambert	3301	-402
	III	Lambert	3326	-403
	IV	Lambert	3351	-404
	V	Lambert	3376	-405
	VI	Lambert	3401	-406
	VII	Lambert	3426	-407
Colorado	North	Lambert	3451	-501
	Central	Lambert	3476	-502
	South	Lambert	3501	-503
Connecticut	-----	Lambert	3526	-600
Delaware	-----	Tr Merc	3551	-700
District of Columbia	Use Maryland or Virginia North			
Florida	East	Tr Merc	3601	-901
	West	Tr Merc	3626	-902
	North	Lambert	3576	-903
Georgia	East	Tr Merc	3651	-1001
	West	Tr Merc	3676	-1002
Guam	-----	Polycon	-----	-5400
Hawaii	1	Tr Merc	5876	-5101
	2	Tr Merc	5901	-5102
	3	Tr Merc	5926	-5103
	4	Tr Merc	5951	-5104
	5	Tr Merc	5976	-5105

Table B-37 NAD27 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States (Continued)

State			Code Number	
	Zone Name	Type	USGS	NOS
Idaho	East	Tr Merc	3701	-1101
	Central	Tr Merc	3726	-1102
	West	Tr Merc	3751	-1103
Illinois	East	Tr Merc	3776	-1201
	West	Tr Merc	3801	-1202
Indiana	East	Tr Merc	3826	-1301
	West	Tr Merc	3851	-1302
Iowa	North	Lambert	3876	-1401
	South	Lambert	3901	-1402
Kansas	North	Lambert	3926	-1501
	South	Lambert	3951	-1502
Kentucky	North	Lambert	3976	-1601
	South	Lambert	4001	-1602
Louisiana	North	Lambert	4026	-1701
	South	Lambert	4051	-1702
	Offshore	Lambert	6426	-1703
Maine	East	Tr Merc	4076	-1801
	West	Tr Merc	4101	-1802
Maryland	-----	Lambert	4126	-1900
Massachusetts	Mainland	Lambert	4151	-2001
	Island	Lambert	4176	-2002
Michigan (Tr Merc)	East	Tr Merc	4201	-2101
	Central	Tr Merc	4226	-2102
	West	Tr Merc	4251	-2103
Michigan (Lambert)	North	Lambert	6351	-2111
	Central	Lambert	6376	-2112
	South	Lambert	6401	-2113
Minnesota	North	Lambert	4276	-2201
	Central	Lambert	4301	-2202
	South	Lambert	4326	-2203
Mississippi	East	Tr Merc	4351	-2301
	West	Tr Merc	4376	-2302

Table B-37 NAD27 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States (Continued)

State			Code Number	
	Zone Name	Type	USGS	NOS
Missouri	East	Tr Merc	4401	-2401
	Central	Tr Merc	4426	-2402
	West	Tr Merc	4451	-2403
Montana	North	Lambert	4476	-2501
	Central	Lambert	4501	-2502
	South	Lambert	4526	-2503
Nebraska	North	Lambert	4551	-2601
	South	Lambert	4576	-2602
Nevada	East	Tr Merc	4601	-2701
	Central	Tr Merc	4626	-2702
	West	Tr Merc	4651	-2703
New Hampshire	-----	Tr Merc	4676	-2800
New Jersey	-----	Tr Merc	4701	-2900
New Mexico	East	Tr Merc	4726	-3001
	Central	Tr Merc	4751	-3002
	West	Tr Merc	4776	-3003
New York	East	Tr Merc	4801	-3101
	Central	Tr Merc	4826	-3102
	West	Tr Merc	4851	-3103
	Long Island	Lambert	4876	-3104
North Carolina	-----	Lambert	4901	-3200
North Dakota	North	Lambert	4926	-3301
	South	Lambert	4951	-3302
Ohio	North	Lambert	4976	-3401
	South	Lambert	5001	-3402
Oklahoma	North	Lambert	5026	-3501
	South	Lambert	5051	-3502
Oregon	North	Lambert	5076	-3601
	South	Lambert	5101	-3602
Pennsylvania	North	Lambert	5126	-3701
	South	Lambert	5151	-3702
Puerto Rico	-----	Lambert	6001	-5201

Table B-37 NAD27 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States (Continued)

State			Code Number	
	Zone Name	Type	USGS	NOS
Rhode Island	-----	Tr Merc	5176	-3800
South Carolina	North	Lambert	5201	-3901
	South	Lambert	5226	-3902
South Dakota	North	Lambert	5251	-4001
	South	Lambert	5276	-4002
St. Croix	-----	Lambert	6051	-5202
Tennessee	-----	Lambert	5301	-4100
Texas	North	Lambert	5326	-4201
	North Central	Lambert	5351	-4202
	Central	Lambert	5376	-4203
	South Central	Lambert	5401	-4204
	South	Lambert	5426	-4205
Utah	North	Lambert	5451	-4301
	Central	Lambert	5476	-4302
	South	Lambert	5501	-4303
Vermont	-----	Tr Merc	5526	-4400
Virginia	North	Lambert	5551	-4501
	South	Lambert	5576	-4502
Virgin Islands	-----	Lambert	6026	-5201
Washington	North	Lambert	5601	-4601
	South	Lambert	5626	-4602
West Virginia	North	Lambert	5651	-4701
	South	Lambert	5676	-4702
Wisconsin	North	Lambert	5701	-4801
	Central	Lambert	5726	-4802
	South	Lambert	5751	-4803
Wyoming	East	Tr Merc	5776	-4901
	East Central	Tr Merc	5801	-4902
	West Central	Tr Merc	5826	-4903
	West	Tr Merc	5851	-4904

Table B-38 NAD83 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States

State	Zone Name	Type	Code Number	
			USGS	NOS
Alabama	East	Tr Merc	3101	-101
	West	Tr Merc	3126	-102
Alaska	1	Oblique	6101	-5001
	2	Tr Merc	6126	-5002
	3	Tr Merc	6151	-5003
	4	Tr Merc	6176	-5004
	5	Tr Merc	6201	-5005
	6	Tr Merc	6226	-5006
	7	Tr Merc	6251	-5007
	8	Tr Merc	6276	-5008
	9	Tr Merc	6301	-5009
	10	Lambert	6326	-5010
Arizona	East	Tr Merc	3151	-201
	Central	Tr Merc	3176	-202
	West	Tr Merc	3201	-203
Arkansas	North	Lambert	3226	-301
	South	Lambert	3251	-302
California	I	Lambert	3276	-401
	II	Lambert	3301	-402
	III	Lambert	3326	-403
	IV	Lambert	3351	-404
	V	Lambert	3376	-405
	VI	Lambert	3401	-406
Colorado	North	Lambert	3451	-501
	Central	Lambert	3476	-502
	South	Lambert	3501	-503
Connecticut	-----	Lambert	3526	-600
Delaware	-----	Tr Merc	3551	-700
District of Columbia	Use Maryland or Virginia North			
Florida	East	Tr Merc	3601	-901
	West	Tr Merc	3626	-902
	North	Lambert	3576	-903

Table B-38 NAD83 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States (Continued)

State	Zone Name	Type	Code Number	
			USGS	NOS
Georgia	East	Tr Merc	3651	-1001
	West	Tr Merc	3676	-1002
Hawaii	1	Tr Merc	5876	-5101
	2	Tr Merc	5901	-5102
	3	Tr Merc	5926	-5103
	4	Tr Merc	5951	-5104
	5	Tr Merc	5976	-5105
Idaho	East	Tr Merc	3701	-1101
	Central	Tr Merc	3726	-1102
	West	Tr Merc	3751	-1103
Illinois	East	Tr Merc	3776	-1201
	West	Tr Merc	3801	-1202
Indiana	East	Tr Merc	3826	-1301
	West	Tr Merc	3851	-1302
Iowa	North	Lambert	3876	-1401
	South	Lambert	3901	-1402
Kansas	North	Lambert	3926	-1501
	South	Lambert	3951	-1502
Kentucky	North	Lambert	3976	-1601
	South	Lambert	4001	-1602
Louisiana	North	Lambert	4026	-1701
	South	Lambert	4051	-1702
	Offshore	Lambert	6426	-1703
Maine	East	Tr Merc	4076	-1801
	West	Tr Merc	4101	-1802
Maryland	-----	Lambert	4126	-1900
Massachusetts	Mainland	Lambert	4151	-2001
	Island	Lambert	4176	-2002
Michigan	North	Lambert	6351	-2111
	Central	Lambert	6376	-2112
	South	Lambert	6401	-2113

Table B-38 NAD83 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States (Continued)

State	Zone Name	Type	Code Number	
			USGS	NOS
Minnesota	North	Lambert	4276	-2201
	Central	Lambert	4301	-2202
	South	Lambert	4326	-2203
Mississippi	East	Tr Merc	4351	-2301
	West	Tr Merc	4376	-2302
Missouri	East	Tr Merc	4401	-2401
	Central	Tr Merc	4426	-2402
	West	Tr Merc	4451	-2403
Montana	-----	Lambert	4476	-2500
Nebraska	-----	Lambert	4551	-2600
Nevada	East	Tr Merc	4601	-2701
	Central	Tr Merc	4626	-2702
	West	Tr Merc	4651	-2703
New Hampshire	-----	Tr Merc	4676	-2800
New Jersey	-----	Tr Merc	4701	-2900
New Mexico	East	Tr Merc	4726	-3001
	Central	Tr Merc	4751	-3002
	West	Tr Merc	4776	-3003
New York	East	Tr Merc	4801	-3101
	Central	Tr Merc	4826	-3102
	West	Tr Merc	4851	-3103
	Long Island	Lambert	4876	-3104
North Carolina	-----	Lambert	4901	-3200
North Dakota	North	Lambert	4926	-3301
	South	Lambert	4951	-3302
Ohio	North	Lambert	4976	-3401
	South	Lambert	5001	-3402
Oklahoma	North	Lambert	5026	-3501
	South	Lambert	5051	-3502
Oregon	North	Lambert	5076	-3601
	South	Lambert	5101	-3602

Table B-38 NAD83 State Plane Coordinate System Zone Numbers, Projection Types, and Zone Code Numbers for the United States (Continued)

State	Zone Name	Type	Code Number	
			USGS	NOS
Pennsylvania	North	Lambert	5126	-3701
	South	Lambert	5151	-3702
Puerto Rico	-----	Lambert	6001	-5201
Rhode Island	-----	Tr Merc	5176	-3800
South Carolina	-----	Lambert	5201	-3900
South Dakota	-----	Lambert	5251	-4001
	South	Lambert	5276	-4002
Tennessee	-----	Lambert	5301	-4100
Texas	North	Lambert	5326	-4201
	North Central	Lambert	5351	-4202
	Central	Lambert	5376	-4203
	South Central	Lambert	5401	-4204
	South	Lambert	5426	-4205
Utah	North	Lambert	5451	-4301
	Central	Lambert	5476	-4302
	South	Lambert	5501	-4303
Vermont	-----	Tr Merc	5526	-4400
Virginia	North	Lambert	5551	-4501
	South	Lambert	5576	-4502
Virgin Islands	-----	Lambert	6026	-5201
Washington	North	Lambert	5601	-4601
	South	Lambert	5626	-4602
West Virginia	North	Lambert	5651	-4701
	South	Lambert	5676	-4702
Wisconsin	North	Lambert	5701	-4801
	Central	Lambert	5726	-4802
	South	Lambert	5751	-4803
Wyoming	East	Tr Merc	5776	-4901
	East Central	Tr Merc	5801	-4902
	West Central	Tr Merc	5826	-4903
	West	Tr Merc	5851	-4904

Stereographic

Table B-39 Stereographic Summary

Construction	Plane
Property	Conformal
Meridians	Polar aspect: the meridians are straight lines radiating from the point of tangency. Oblique and equatorial aspects: the meridians are arcs of circles concave toward a straight central meridian. In the equatorial aspect, the outer meridian of the hemisphere is a circle centered at the projection center.
Parallels	Polar aspect: the parallels are concentric circles. Oblique aspect: the parallels are nonconcentric arcs of circles concave toward one of the poles with one parallel being a straight line. Equatorial aspect: parallels are nonconcentric arcs of circles concave toward the poles; the Equator is straight.
Graticule spacing	The graticule spacing increases away from the center of the projection in all aspects and it retains the property of conformality.
Linear scale	Scale increases toward the periphery of the projection.
Uses	The Stereographic projection is the most widely used azimuthal projection, mainly used for portraying large, continent-sized areas of similar extent in all directions. It is used in geophysics for solving problems in spherical geometry. The polar aspect is used for topographic maps and navigational charts. The American Geographical Society uses this projection as the basis for its "Map of the Arctic." The USGS uses it as the basis for maps of Antarctica.

Stereographic is a perspective projection in which points are projected from a position on the opposite side of the globe onto a plane tangent to the Earth (Figure B-34). All of one hemisphere can easily be shown, but it is impossible to show both hemispheres in their entirety from one center. It is the only azimuthal projection that preserves truth of angles and local shape. Scale increases and parallels become more widely spaced farther from the center.

In the equatorial aspect, all parallels except the Equator are circular arcs. In the polar aspect, latitude rings are spaced farther apart, with increasing distance from the pole.

Map Projections

Prompts

The following prompts display in the Projection Chooser if Stereographic is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Define the center of the map projection in both spherical and rectangular coordinates.

Longitude of center of projection

Latitude of center of projection

Enter values for the longitude and latitude of the desired center of the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

The Stereographic is the only azimuthal projection which is conformal. Figure B-34 shows two views: A) Equatorial aspect, often used in the 16th and 17th centuries for maps of hemispheres; and B) Oblique aspect, centered on 40°N.

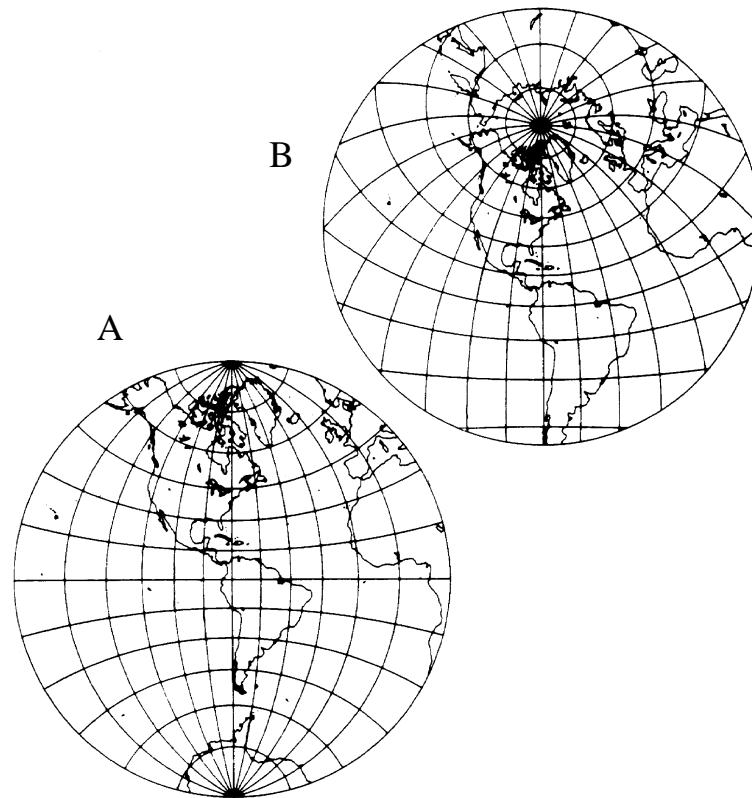


Figure B-34 Stereographic Projection

**Stereographic
(Extended)**

The Stereographic (Extended) projection has the same attributes as the Stereographic projection, with the exception of the ability to define scale factors.



For details about the Stereographic projection, see "Stereographic" on page 563.

Prompts

The following prompts display in the Projection Chooser once Stereographic (Extended) is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Scale factor

Map Projections

Designate the desired scale factor. This parameter is used to modify scale distortion. A value of one indicates true scale only along the central meridian. It may be desirable to have true scale along two lines equidistant from and parallel to the central meridian, or to lessen scale distortion away from the central meridian. A factor of less than, but close to one is often used.

Longitude of origin of projection

Latitude of origin of projection

Enter the values for longitude of origin of projection and latitude of origin of projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Transverse Mercator

Table B-40 Transverse Mercator Summary

Construction	Cylinder
Property	Conformal
Meridians	Meridians are complex curves concave toward a straight central meridian that is tangent to the globe. The straight central meridian intersects the Equator and one meridian at a 90° angle.
Parallels	Parallels are complex curves concave toward the nearest pole; the Equator is straight.
Graticule spacing	Parallels are spaced at their true distances on the straight central meridian. Graticule spacing increases away from the tangent meridian. The graticule retains the property of conformality.
Linear scale	Linear scale is true along the line of tangency, or along two lines equidistant from, and parallel to, the line of tangency.
Uses	Used where the north-south dimension is greater than the east west dimension. Used as the base for the USGS 1:250,000-scale series, and for some of the 7.5-minute and 15-minute quadrangles of the National Topographic Map Series.

Transverse Mercator is similar to the Mercator projection except that the axis of the projection cylinder is rotated 90° from the vertical (polar) axis. The contact line is then a chosen meridian instead of the Equator, and this central meridian runs from pole to pole. It loses the properties of straight meridians and straight parallels of the standard Mercator projection (except for the central meridian, the two meridians 90° away, and the Equator).

Transverse Mercator also loses the straight rhumb lines of the Mercator map, but it is a conformal projection. Scale is true along the central meridian or along two straight lines equidistant from, and parallel to, the central meridian. It cannot be edge-joined in an east-west direction if each sheet has its own central meridian.

In the United States, Transverse Mercator is the projection used in the State Plane coordinate system for states with predominant north-south extent. The entire Earth from 84°N to 80°S is mapped with a system of projections called the Universal Transverse Mercator.

Prompts

The following prompts display in the Projection Chooser if Transverse Mercator is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Scale factor at central meridian

Designate the desired scale factor at the central meridian. This parameter is used to modify scale distortion. A value of one indicates true scale only along the central meridian. It may be desirable to have true scale along two lines equidistant from and parallel to the central meridian, or to lessen scale distortion away from the central meridian. A factor of less than, but close to one is often used.

Finally, define the origin of the map projection in both spherical and rectangular coordinates.

Longitude of central meridian

Latitude of origin of projection

Enter values for longitude of the desired central meridian and latitude of the origin of projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the intersection of the central meridian and the latitude of the origin of projection. These values must be in meters. It is often convenient to make them large enough so that there are no negative coordinates within the region of the map projection. That is, origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Two Point Equidistant

Table B-41 Two Point Equidistant Summary

Construction	Modified planar
Property	Compromise
Meridians	N/A
Parallels	N/A
Graticule spacing	N/A
Linear scale	N/A
Uses	The Two Point Equidistant projection “does not represent great circle paths” (ESRI 1997). There is little distortion if two chosen points are within 45 degrees of each other.

The Two Point Equidistant projection is used to show the distance from “either of two chosen points to any other point on a map” (ESRI 1997). Note that the first point has to be west of the second point. This projection has been used by the National Geographic Society to map areas of Asia.

Source: ESRI 1997

Prompts

The following prompts display in the Projection Chooser once Two Point Equidistant is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

False easting

False northing

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

Longitude of 1st point

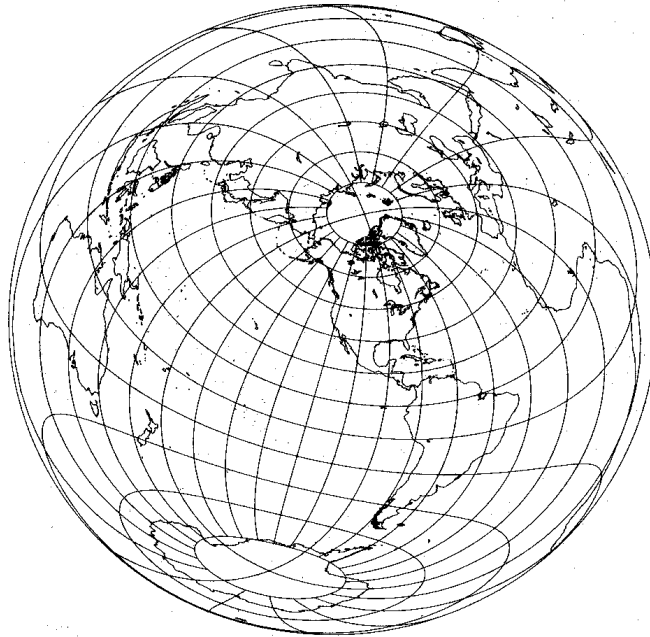
Latitude of 1st point

Enter the longitude and latitude values of the first point.

Longitude of 2nd point

Latitude of 2nd point

Enter the longitude and latitude values of the second point.



Source: Snyder and Voxland 1989

Figure B-35 Two Point Equidistant Projection

UTM

UTM is an international plane (rectangular) coordinate system developed by the US Army that extends around the world from 84°N to 80°S. The world is divided into 60 zones each covering six degrees longitude. Each zone extends three degrees eastward and three degrees westward from its central meridian. Zones are numbered consecutively west to east from the 180° meridian (Figure B-36, Table B-42).

The Transverse Mercator projection is then applied to each UTM zone. Transverse Mercator is a transverse form of the Mercator cylindrical projection. The projection cylinder is rotated 90° from the vertical (polar) axis and can then be placed to intersect at a chosen central meridian. The UTM system specifies the central meridian of each zone. With a separate projection for each UTM zone, a high degree of accuracy is possible (one part in 1000 maximum distortion within each zone).

If the map to be projected extends beyond the border of the UTM zone, the entire map may be projected for any UTM zone specified by you.



See "Space Oblique Mercator (Formats A & B)" on page 566 for more information.

Map Projections

Prompts

The following prompts display in the Projection Chooser if UTM is chosen.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

UTM Zone

North or South

All values in Table B-42 are in full degrees east (E) or west (W) of the Greenwich prime meridian (0).

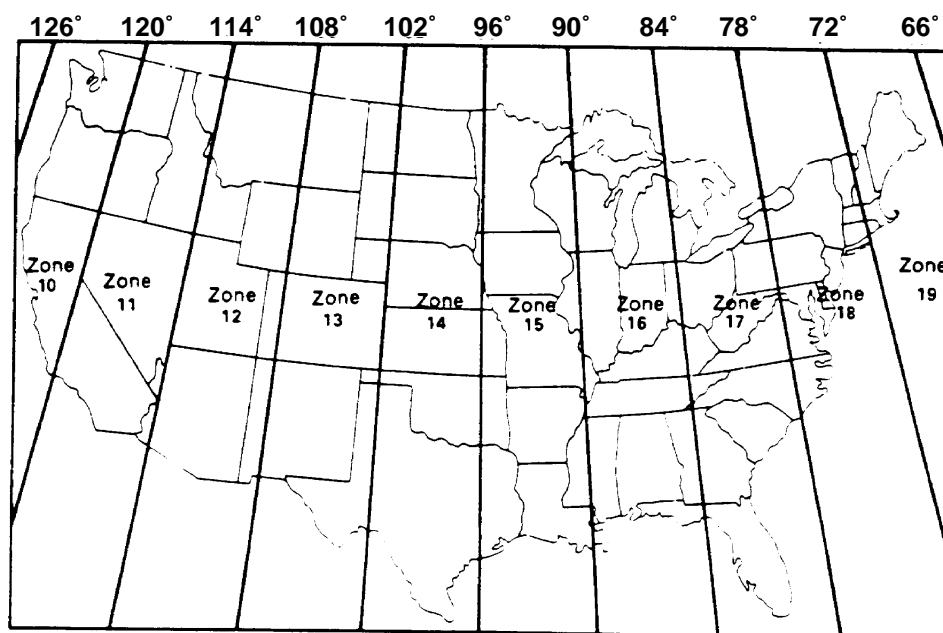


Figure B-36 Zones of the Universal Transverse Mercator Grid in the United States

Table B-42 UTM Zones, Central Meridians, and Longitude Ranges

Zone	Central Meridian	Range	Zone	Central Meridian	Range
1	177W	180W-174W	31	3E	0-6E
2	171W	174W-168W	32	9E	6E-12E
3	165W	168W-162W	33	15E	12E-18E
4	159W	162W-156W	34	21E	18E-24E
5	153W	156W-150W	35	27E	24E-30E
6	147W	150W-144W	36	33E	30E-36E
7	141W	144W-138W	37	39E	36E-42E
8	135W	138W-132W	38	45E	42E-48E
9	129W	132W-126W	39	51E	48E-54E
10	123W	126W-120W	40	57E	54E-60E
11	117W	120W-114W	41	63E	60E-66E
12	111W	114W-108W	42	69E	66E-72E
13	105W	108W-102W	43	75E	72E-78E
14	99W	102W-96W	44	81E	78E-84E
15	93W	96W-90W	45	87E	84E-90E
16	87W	90W-84W	46	93E	90E-96E
17	81W	84W-78W	47	99E	96E-102E
18	75W	78W-72W	48	105E	102E-108E
19	69W	72W-66W	49	111E	108E-114E
20	63W	66W-60W	50	117E	114E-120E
21	57W	60W-54W	51	123E	120E-126E
22	51W	54W-48W	52	129E	126E-132E
23	45W	48W-42W	53	135E	132E-138E
24	39W	42W-36W	54	141E	138E-144E
25	33W	36W-30W	55	147E	144E-150E
26	27W	30W-24W	56	153E	150E-156E
27	21W	24W-18W	57	159E	156E-162E
28	15W	18W-12W	58	165E	162E-168E
29	9W	12W-6W	59	171E	168E-174E
30	3W	6W-0	60	177E	174E-180E

Van der Grinten I

Table B-43 Van der Grinten I Summary

Construction	Miscellaneous
Property	Compromise
Meridians	Meridians are circular arcs concave toward a straight central meridian.
Parallels	Parallels are circular arcs concave toward the poles, except for a straight Equator.
Graticule spacing	Meridian spacing is equal at the Equator. The parallels are spaced farther apart toward the poles. The central meridian and Equator are straight lines. The poles are commonly not represented. The graticule spacing results in a compromise of all properties.
Linear scale	Linear scale is true along the Equator. Scale increases rapidly toward the poles.
Uses	The Van der Grinten projection is used by the National Geographic Society for world maps. Used by the USGS to show distribution of mineral resources on the sea floor.

The Van der Grinten I projection produces a map that is neither conformal nor equal-area (Figure B-37). It compromises all properties, and represents the Earth within a circle.

All lines are curved except the central meridian and the Equator. Parallels are spaced farther apart toward the poles. Meridian spacing is equal at the Equator. Scale is true along the Equator, but increases rapidly toward the poles, which are usually not represented.

Van der Grinten I avoids the excessive stretching of the Mercator and the shape distortion of many of the equal-area projections. It has been used to show distribution of mineral resources on the ocean floor.

Prompts

The following prompts display in the Projection Chooser if Van der Grinten I is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Longitude of central meridian

Enter a value for the longitude of the desired central meridian to center the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.

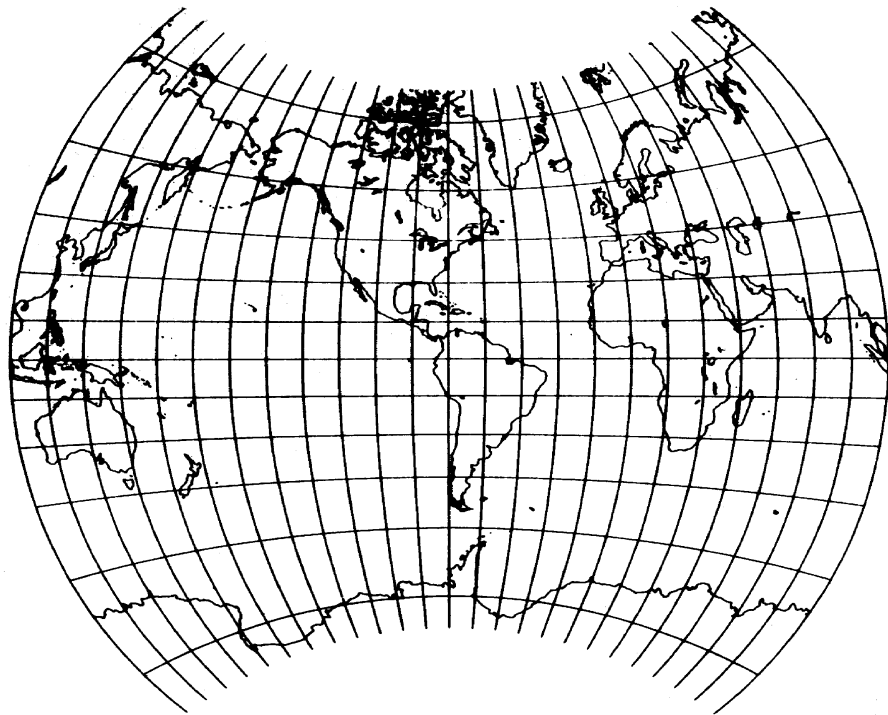


Figure B-37 Van der Grinten I Projection

The Van der Grinten I projection resembles the Mercator, but it is not conformal.

Map Projections

Wagner IV

Table B-44 Wagner IV Summary

Construction	Pseudocylinder
Property	Equal-area
Meridians	The central meridian is a straight line one half as long as the Equator. The other meridians are portions of ellipses that are equally spaced. They are concave towards the central meridian. The meridians at 103° 55' E and W of the central meridian are circular arcs.
Parallels	Parallels are unequally spaced. Parallels have the widest space between them at the Equator, and are perpendicular to the central meridian.
Graticule spacing	See Meridians and Parallels. Poles are lines one half as long as the Equator. Symmetry exists around the central meridian or the Equator.
Linear scale	Scale is accurate along latitudes 42° 59' N and S. Scale is constant along any specific latitude as well as the latitude of opposite sign.
Uses	Useful for world maps.

The Wagner IV Projection has distortion primarily in the polar regions.

Source: Snyder and Voxland 1989

Prompts

The following prompts display in the Projection Chooser if Wagner IV is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

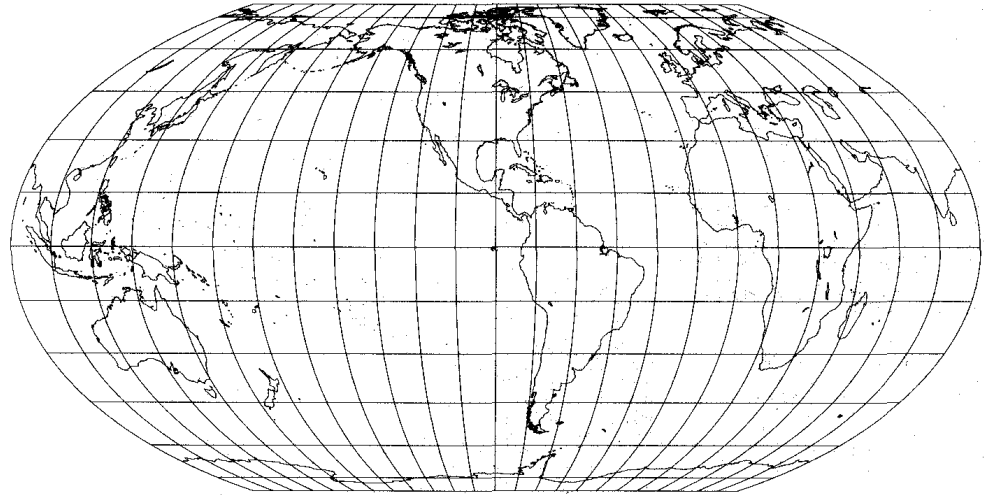
Longitude of central meridian

Enter a value for the longitude of the desired central meridian to center the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-38 Wagner IV Projection

Wagner VII

Table B-45 Wagner VII

Construction	Modified azimuthal
Property	Equal-area
Meridians	Central meridian is straight and half the Equator's length. Other meridians are unequally spaced curves. They are concave toward the central meridian.
Parallels	The Equator is straight; the other parallels are curved. Other parallels are unequally spaced curves, which are concave toward the closest pole.
Graticule spacing	See Meridians and Parallels. Poles are curved lines. Symmetry exists about the central meridian or the Equator.
Linear scale	Scale decreases along the central meridian and the Equator relative to distance from the center of the Wagner VII projection.
Uses	Used for world maps.

Distortion is prevalent in polar areas. The Wagner VII projection is modified based on the Hammer projection. "The poles correspond to the 65th parallels on the Hammer [projection], and meridians are repositioned" (Snyder and Voxland 1989).

Source: Snyder and Voxland 1989

Prompts

The following prompts display in the Projection Chooser if Wagner IV is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

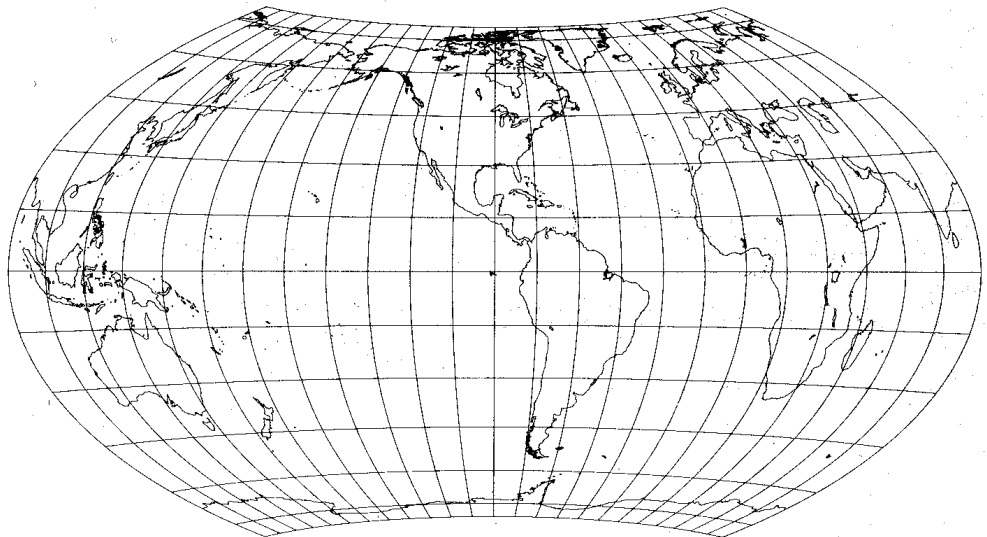
Longitude of central meridian

Enter a value for the longitude of the desired central meridian to center the projection.

False easting

False northing

Enter values of false easting and false northing corresponding to the center of the projection. These values must be in meters. It is often convenient to make them large enough to prevent negative coordinates within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-39 Wagner VII Projection

Winkel I

Table B-46 Winkel I Summary

Construction	Pseudocylinder
Property	Neither conformal nor equal-area
Meridians	Central meridian is a straight line 0.61 the length of the Equator. The other meridians are sinusoidal curves that are equally spaced and concave toward the central meridian.
Parallels	Parallels are equally spaced.
Graticule spacing	See Meridians and Parallels. Pole lines are 0.61 the length of the Equator. Symmetry exists about the central meridian or the Equator.
Linear scale	Scale is true along latitudes 50° 28' N and S. Scale is constant along any given latitude as well as the latitude of the opposite sign.
Uses	Used for world maps.

The Winkel I projection is “not free of distortion at any point” (Snyder and Voxland 1989).

Source: Snyder and Voxland 1989

Prompts

The following prompts display in the Projection Chooser once Winkel I is selected. Respond to the prompts as described.

Spheroid Name

Datum Name

Select the spheroid and datum to use.



The list of available spheroids is located on page 445.

Latitude of standard parallel

Longitude of central meridian

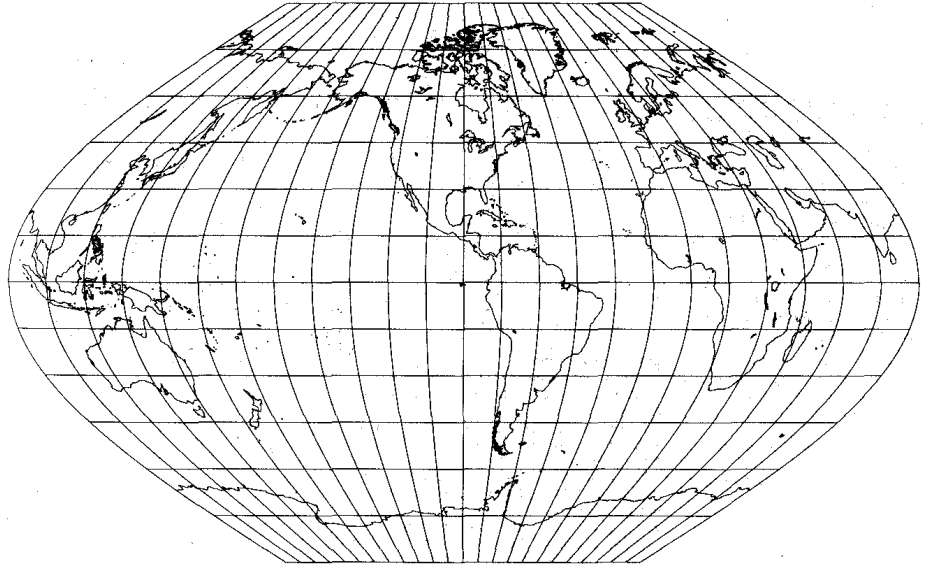
Enter values of the latitude of standard parallel and the longitude of central meridian.

False easting

False northing

Map Projections

Enter values of false easting and false northing corresponding to the desired center of the projection. These values must be in meters. It is often convenient to make them large enough so that no negative coordinates occur within the region of the map projection. That is, the origin of the rectangular coordinate system should fall outside of the map projection to the south and west.



Source: Snyder and Voxland 1989

Figure B-40 Winkel I Projection

External Projections

The following external projections are supported in ERDAS IMAGINE and are described in this section. Some of these projections were discussed in the previous section. Those descriptions are not repeated here. Simply refer to the page number in parentheses for more information.

NOTE: ERDAS IMAGINE does not support datum shifts for these external projections.

- Albers Equal Area (see page 481 for "Albers Conical Equal Area")
- Azimuthal Equidistant (see page 483)
- Bipolar Oblique Conic Conformal
- Cassini-Soldner
- Conic Equidistant (see page 500 for "Equidistant Conic")
- Laborde Oblique Mercator
- Lambert Azimuthal Equal Area (see page 515)
- Lambert Conformal Conic (see page 518)

- Mercator (see page 523)
- Minimum Error Conformal
- Modified Polyconic
- Modified Stereographic
- Mollweide Equal Area (see page 529 for "Mollweide")
- Oblique Mercator (see page 532)
- Orthographic (see page 535)
- Plate Carrée (see page 503)
- Rectified Skew Orthomorphic (see page 547 for "Robinson")
- Regular Polyconic (see page 542)
- Robinson Pseudocylindrical (see page 545 for "Robinson")
- Sinusoidal (see page 548)
- Southern Orientated Gauss Conformal
- Stereographic (see page 563)
- Swiss Cylindrical
- Stereographic (Oblique) (see page 563)
- Transverse Mercator (see page 566)
- Universal Transverse Mercator (see page 569)
- Van der Grinten (see page 571)
- Winkel's Triplet

Map Projections

Bipolar Oblique Conic Conformal

Table B-47 Bipolar Oblique Conic Conformal Summary

Construction	Cone
Property	Conformal
Meridians	Meridians are complex curves concave toward the center of the projection.
Parallels	Parallels are complex curves concave toward the nearest pole.
Graticule spacing	Graticule spacing increases away from the lines of true scale and retains the property of conformality.
Linear scale	Linear scale is true along two lines that do not lie along any meridian or parallel. Scale is compressed between these lines and expanded beyond them. Linear scale is generally good, but there is as much as a 10% error at the edge of the projection as used.
Uses	Used to represent one or both of the American continents. Examples are the Basement map of North America and the Tectonic map of North America.

The Bipolar Oblique Conic Conformal projection was developed by O.M. Miller and William A. Briesemeister in 1941 specifically for mapping North and South America, and maintains conformality for these regions. It is based upon the Lambert Conformal Conic, using two oblique conic projections side-by-side. The two oblique conics are joined with the poles 104° apart. A great circle arc 104° long begins at 20°S and 110°W, cuts through Central America, and terminates at 45°N and approximately 19°59'36"W. The scale of the map is then increased by approximately 3.5%. The origin of the coordinates is made 17°15'N, 73°02'W.



Refer to "Interrupted Goode Homolosine" on page 518 for more information.

Prompts

The following prompts display in the Projection Chooser if Bipolar Oblique Conic Conformal is selected.

Projection Name

Spheroid Type

Datum Name

Cassini-Soldner

Table B-48 Cassini-Soldner Summary

Construction	Cylinder
Property	Compromise
Meridians	Central meridian, each meridian 90° from the central meridian, and the Equator are straight lines. Other meridians are complex curves.
Parallels	Parallels are complex curves.
Graticule spacing	Complex curves for all meridians and parallels, except for the Equator, the central meridian, and each meridian 90° away from the central meridian, all of which are straight.
Linear scale	Scale is true along the central meridian, and along lines perpendicular to the central meridian. Scale is constant but not true along lines parallel to the central meridian on the spherical form, and nearly so for the ellipsoid.
Uses	Used for topographic mapping, formerly in England and currently in a few other countries, such as Denmark, Germany, and Malaysia.

The Cassini projection was devised by C. F. Cassini de Thury in 1745 for the survey of France. Mathematical analysis by J. G. von Soldner in the early 19th century led to more accurate ellipsoidal formulas. Today, it has largely been replaced by the Transverse Mercator projection, although it is still in limited use outside of the United States. It was one of the major topographic mapping projections until the early 20th century.

The spherical form of the projection bears the same relation to the Equidistant Cylindrical, or Plate Carrée, projection that the spherical Transverse Mercator bears to the regular Mercator. Instead of having the straight meridians and parallels of the Equidistant Cylindrical, the Cassini has complex curves for each, except for the Equator, the central meridian, and each meridian 90° away from the central meridian, all of which are straight.

There is no distortion along the central meridian if it is maintained at true scale, which is the usual case. If it is given a reduced scale factor, the lines of true scale are two straight lines on the map, parallel to and equidistant from, the central meridian. There is no distortion along them instead.

The scale is correct along the central meridian, and also along any straight line perpendicular to the central meridian. It gradually increases in a direction parallel to the central meridian as the distance from that meridian increases, but the scale is constant along any straight line on the map that is parallel to the central meridian. Therefore, Cassini-Soldner is more suitable for regions that are predominantly north-south in extent, such as Great Britain, than regions extending in other directions. The projection is neither equal-area nor conformal, but it has a compromise of both features.

Map Projections

The Cassini-Soldner projection was adopted by the Ordnance Survey for the official survey of Great Britain during the second half of the 19th century. A system equivalent to the oblique Cassini-Soldner projection was used in early coordinate transformations for ERTS (now Landsat) satellite imagery, but it was changed to Oblique Mercator (Hotine) in 1978, and to the Space Oblique Mercator in 1982.

Prompts

The following prompts display in the Projection Chooser if Cassini-Soldner is selected.

Projection Name

Spheroid Type

Datum Name

Laborde Oblique Mercator

In 1928, Laborde combined a conformal sphere with a complex-algebra transformation of the Oblique Mercator projection for the topographic mapping of Madagascar. This variation is now known as the Laborde Oblique Mercator. The central line is a great circle arc.



See "Oblique Mercator (Hotine)" on page 528 for more information.

Prompts

The following prompts display in the Projection Chooser if Laborde Oblique Mercator is selected.

Projection Name

Spheroid Type

Datum Name

Minimum Error Conformal

The Minimum Error Conformal projection is the same as the New Zealand Map Grid projection.



For information about the New Zealand Map Grid projection, see page 530.

Modified Polyconic

Table B-49 Modified Polyconic Summary

Construction	Cone
Property	Compromise
Meridians	All meridians are straight.
Parallels	Parallels are circular arcs. The top and bottom parallels of each sheet are nonconcentric circular arcs.
Graticule spacing	The top and bottom parallels of each sheet are nonconcentric circular arcs. The two parallels are spaced from each other according to the true scale along the central meridian, which is slightly reduced.
Linear scale	Scale is true along each parallel and along two meridians, but no parallel is standard.
Uses	Used for the International Map of the World (IMW) series until 1962.

The Modified Polyconic projection was devised by Lallemand of France, and in 1909 it was adopted by the International Map Committee (IMC) in London as the basis for the 1:1,000,000-scale International Map of the World (IMW) series.

The projection differs from the ordinary Polyconic in two principal features: all meridians are straight, and there are two meridians that are made true to scale. Adjacent sheets fit together exactly not only north to south, but east to west. There is still a gap when mosaicking in all directions, in that there is a gap between each diagonal sheet, and either one or the other adjacent sheet.

In 1962, a U.N. conference on the IMW adopted the Lambert Conformal Conic and the Polar Stereographic projections to replace the Modified Polyconic.



See "Polyconic" on page 542 for more information.

Prompts

The following prompts display in the Projection Chooser if Modified Polyconic is selected.

Projection Name

Spheroid Type

Datum Name

Map Projections

Modified Stereographic

Table B-50 Modified Stereographic Summary

Construction	Plane
Property	Conformal
Meridians	All meridians are normally complex curves, although some may be straight under certain conditions.
Parallels	All parallels are complex curves, although some may be straight under certain conditions.
Graticule spacing	The graticule is normally not symmetrical about any axis or point.
Linear scale	Scale is true along irregular lines, but the map is usually designed to minimize scale variation throughout a selected region.
Uses	Used for maps of continents in the Eastern Hemisphere, for the Pacific Ocean, and for maps of Alaska and the 50 United States.

The meridians and parallels of the Modified Stereographic projection are generally curved, and there is usually no symmetry about any point or line. There are limitations to these transformations. Most of them can only be used within a limited range. As the distance from the projection center increases, the meridians, parallels, and shorelines begin to exhibit loops, overlapping, and other undesirable curves. A world map using the GS50 (50-State) projection is almost illegible with meridians and parallels intertwined like wild vines.

Prompts

The following prompts display in the Projection Chooser if Modified Stereographic is selected.

Projection Name

Spheroid Type

Datum Name

Mollweide Equal Area

Table B-51 Mollweide Equal Area Summary

Construction	Pseudocylinder
Property	Equal-area
Meridians	All of the meridians are ellipses. The central meridian is a straight line, and 90° meridians are circular arcs (Pearson 1990).
Parallels	The Equator and parallels are straight lines perpendicular to the central meridian, but they are not equally spaced.
Graticule spacing	Linear graticules include the central meridian and the Equator (ESRI 1992). Meridians are equally spaced along the Equator and along all other parallels. The parallels are straight parallel lines, but they are not equally spaced. The poles are points.
Linear scale	Scale is true along latitudes 40°44'N and S. Distortion increases with distance from these lines and becomes severe at the edges of the projection (ESRI 1992).
Uses	Often used for world maps (Pearson 1990). Suitable for thematic or distribution mapping of the entire world, frequently in interrupted form (ESRI 1992).

The second oldest pseudocylindrical projection that is still in use (after the Sinusoidal) was presented by Carl B. Mollweide (1774 - 1825) of Halle, Germany, in 1805. It is an equal-area projection of the Earth within an ellipse. It has had a profound effect on world map projections in the 20th century, especially as an inspiration for other important projections, such as the Van der Grinten.

The Mollweide is normally used for world maps and occasionally for a very large region, such as the Pacific Ocean. This is because only two points on the Mollweide are completely free of distortion unless the projection is interrupted. These are the points at latitudes 40°44'12"N and S on the central meridian(s).

The world is shown in an ellipse with the Equator, its major axis, twice as long as the central meridian, its minor axis. The meridians 90° east and west of the central meridian form a complete circle. All other meridians are elliptical arcs which, with their opposite numbers on the other side of the central meridian, form complete ellipses that meet at the poles.

Prompts

The following prompts display in the Projection Chooser if Mollweide Equal Area is selected.

Projection Name

Spheroid Type

Datum Name

Map Projections

Rectified Skew Orthomorphic

Martin Hotine (1898 - 1968) called the Oblique Mercator projection the Rectified Skew Orthomorphic projection.



See "Oblique Mercator (Hotine)" on page 528 for more information.

Prompts

The following prompts display in the Projection Chooser if Rectified Skew Orthomorphic is selected.

Projection Name

Spheroid Type

Datum Name

Robinson Pseudocylindrical

Table B-52 Robinson Pseudocylindrical Summary

Construction	Pseudocylinder
Property	Compromise
Meridians	Meridians are elliptical arcs, equally spaced, and concave toward the central meridian.
Parallels	Parallels are straight lines.
Graticule spacing	Parallels are straight lines and are parallel. The individual parallels are evenly divided by the meridians (Pearson 1990).
Linear scale	Generally, scale is made true along latitudes 38°N and S. Scale is constant along any given latitude, and for the latitude of opposite sign (ESRI 1992).
Uses	Developed for use in general and thematic world maps. Used by Rand McNally since the 1960s and by the National Geographic Society since 1988 for general and thematic world maps (ESRI 1992).

The Robinson Pseudocylindrical projection provides a means of showing the entire Earth in an uninterrupted form. The continents appear as units and are in relatively correct size and location. Poles are represented as lines.

Meridians are equally spaced and resemble elliptical arcs, concave toward the central meridian. The central meridian is a straight line 0.51 times the length of the Equator. Parallels are equally spaced straight lines between 38°N and S, and then the spacing decreases beyond these limits. The poles are 0.53 times the length of the Equator. The projection is based upon tabular coordinates instead of mathematical formulas (ESRI 1992).

Prompts

The following prompts display in the Projection Chooser if Robinson Pseudocylindrical is selected.

Projection Name

Spheroid Type

Datum Name

**Southern Orientated
Gauss Conformal**

Southern Orientated Gauss Conformal is another name for the Transverse Mercator projection, after mathematician Friedrich Gauss (1777 - 1855). It is also called the Gauss-Krüger projection.



See "Space Oblique Mercator (Formats A & B)" on page 566 for more information.

Prompts

The following prompts display in the Projection Chooser if Southern Orientated Gauss Conformal is selected.

Projection Name

Spheroid Type

Datum Name

Swiss Cylindrical

The Swiss Cylindrical projection is a cylindrical projection used by the Swiss Landestopographie, which is a form of the Oblique Mercator projection.



For information about the Oblique Mercator projection, see page 532.

Winkel's Tripel**Table B-53 Winkel's Tripel Summary**

Construction	Modified azimuthal
Property	Neither conformal nor equal-area
Meridians	Central meridian is straight. Other meridians are curved and are equally spaced along the Equator, and concave toward the central meridian.
Parallels	Equidistant spacing of parallels. Equator and the poles are straight. Other parallels are curved and concave toward the nearest pole.

Table B-53 Winkel's Tripel Summary (Continued)

Graticule spacing	Symmetry is maintained along the central meridian or the Equator.
Linear scale	Scale is true along the central meridian and constant along the Equator.
Uses	Used for world maps.

Winkel's Tripel was formulated in 1921 by Oswald Winkel of Germany. It is a combined projection that is the arithmetic mean of the Plate Carrée and Aitoff's projection (Maling).

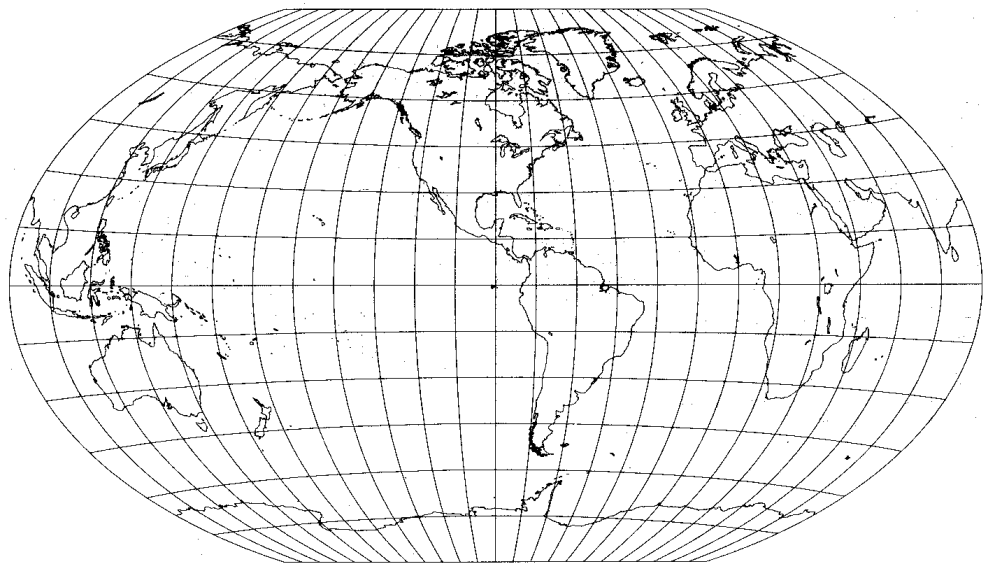
Prompts

The following prompts display in the Projection Chooser if Winkel's Tripel is selected.

Projection Name

Spheroid Type

Datum Name



Source: Snyder and Voxland 1989

Figure B-41 Winkel's Tripel Projection

Glossary

Numerics

2D—two-dimensional.

3D—three-dimensional.

A

absorption spectra—the electromagnetic radiation wavelengths that are absorbed by specific materials of interest.

abstract symbol—an annotation symbol that has a geometric shape, such as a circle, square, or triangle. These symbols often represent amounts that vary from place to place, such as population density, yearly rainfall, etc.

a priori—already or previously known.

accuracy assessment—the comparison of a classification to geographical data that is assumed to be true. Usually, the assumed-true data are derived from ground truthing.

accuracy report—in classification accuracy assessment, a list of the percentages of accuracy, which is computed from the error matrix.

ACS—see **attitude control system**.

active sensors—the solar imaging sensors that both emit and receive radiation.

ADRG—see **ARC Digitized Raster Graphic**.

ADRI—see **ARC Digital Raster Imagery**.

aerial stereopair—two photos taken at adjacent exposure stations.

Airborne Synthetic Aperture Radar—an experimental airborne radar sensor developed by Jet Propulsion Laboratories (JPL), Pasadena, California, under a contract with NASA. AIRSAR data have been available since 1983.

Airborne Visible/Infrared Imaging Spectrometer—(AVIRIS) a sensor developed by JPL (Pasadena, California) under a contract with NASA that produces multispectral data with 224 narrow bands. These bands are 10 nm wide and cover the spectral range of 0.4-2.4 nm. AVIRIS data have been available since 1987.

AIRSAR—see **Airborne Synthetic Aperture Radar**.

alarm—a test of a training sample, usually used before the signature statistics are calculated. An alarm highlights an area on the display that is an approximation of the area that would be classified with a signature. The original data can then be compared to the highlighted area.

Almaz—a Russian radar satellite that completed its mission in 1992.

Along-Track Scanning Radiometer—(ATSR) instrument aboard the European Space Agency's ERS-1 and ERS-2 satellites, which detects changes in the amount of vegetation on the Earth's surface.

American Standard Code for Information Interchange—(ASCII) a "basis of character sets. . .to convey some control codes, space, numbers, most basic punctuation, and unaccented letters a-z and A-Z" (Free On-Line Dictionary of Computing 1999).

analog photogrammetry—optical or mechanical instruments used to reconstruct three-dimensional geometry from two overlapping photographs.

analytical photogrammetry—the computer replaces optical and mechanical components by substituting analog measurement and calculation with mathematical computation.

ancillary data—the data, other than remotely sensed data, that are used to aid in the classification process.

ANN—see **Artificial Neural Networks**.

annotation—the explanatory material accompanying an image or map. In ERDAS IMAGINE, annotation consists of text, lines, polygons, ellipses, rectangles, legends, scale bars, grid lines, tick marks, neatlines, and symbols that denote geographical features.

annotation layer—a set of annotation elements that is drawn in a Viewer or Map Composer window and stored in a file (.ovr extension).

AOI—see **area of interest**.

arc—see **line**.

ARC system (Equal Arc-Second Raster Chart/Map)—a system that provides a rectangular coordinate and projection system at any scale for the Earth's ellipsoid, based on the World Geodetic System 1984 (WGS 84).

ARC Digital Raster Imagery—Defense Mapping Agency (DMA) data that consist of SPOT panchromatic, SPOT multispectral, or Landsat TM satellite imagery transformed into the ARC system and accompanied by ASCII encoded support files. These data are available only to Department of Defense contractors.

ARC Digitized Raster Graphic—data from the Defense Mapping Agency (DMA) that consist of digital copies of DMA hardcopy graphics transformed into the ARC system and accompanied by ASCII encoded support files. These data are primarily used for military purposes by defense contractors.

ARC GENERATE data—vector data created with the ArcInfo UNGENERATE command.

arc/second—a unit of measure that can be applied to data in the Lat/Lon coordinate system. Each pixel represents the distance covered by one second of latitude or longitude. For example, in 3 arc/second data, each pixel represents an area three seconds latitude by three seconds longitude.

area—a measurement of a surface.

- area based matching**—an image matching technique that determines the correspondence between two image areas according to the similarity of their gray level values.
- area of interest** —(AOI) a point, line, or polygon that is selected as a training sample or as the image area to be used in an operation. AOIs can be stored in separate .aoi files.
- Artificial Neural Networks**—(ANN) data classifiers that may process hyperspectral images with a large number of bands.
- ASCII**—see **American Standard Code for Information Interchange**.
- aspect**—the orientation, or the direction that a surface faces, with respect to the directions of the compass: north, south, east, west.
- aspect image**—a thematic raster image that shows the prevailing direction that each pixel faces.
- aspect map**—a map that is color coded according to the prevailing direction of the slope at each pixel.
- ATSR**—see **Along-Track Scanning Radiometer**.
- attitude control system**—(ACS) system used by SeaWiFS instrument to sustain orbit, conduct lunar and solar calibration procedures, and supply attitude information within one SeaWiFS pixel (NASA 1999).
- attribute**—the tabular information associated with a raster or vector layer.
- average**—the statistical mean; the sum of a set of values divided by the number of values in the set.
- AVHRR**—Advanced Very High Resolution Radiometer data. Small-scale imagery produced by an NOAA polar orbiting satellite. It has a spatial resolution of 1.1×1.1 km or 4×4 km.
- AVIRIS**—see **Airborne Visible/Infrared Imaging Spectrometer**.
- azimuth**—an angle measured clockwise from a meridian, going north to east.
- azimuthal projection**—a map projection that is created from projecting the surface of the Earth to the surface of a plane.

-
- B**
- band**—a set of data file values for a specific portion of the electromagnetic spectrum of reflected light or emitted heat (red, green, blue, near-infrared, infrared, thermal, etc.), or some other user-defined information created by combining or enhancing the original bands, or creating new bands from other sources. Sometimes called channel.
- banding**—see **striping**.
- base map**—a map portraying background reference information onto which other information is placed. Base maps usually show the location and extent of natural surface features and permanent human-made features.

Basic Image Interchange Format—(BIIF) the basis for the NITFS format.

batch file—a file that is created in the Batch mode of ERDAS IMAGINE. All steps are recorded for a later run. This file can be edited.

batch mode—a mode of operating ERDAS IMAGINE in which steps are recorded for later use.

bathymetric map—a map portraying the shape of a water body or reservoir using isobaths (depth contours).

Bayesian—a variation of the maximum likelihood classifier, based on the Bayes Law of probability. The Bayesian classifier allows the application of *a priori* weighting factors, representing the probabilities that pixels are assigned to each class.

BIIF—see **Basic Image Interchange Format**.

BIL—band interleaved by line. A form of data storage in which each record in the file contains a scan line (row) of data for one band. All bands of data for a given line are stored consecutively within the file.

bilinear interpolation—a resampling method that uses the data file values of four pixels in a 2×2 window to calculate an output data file value by computing a weighted average of the input data file values with a bilinear function.

bin function—a mathematical function that establishes the relationship between data file values and rows in a descriptor table.

bins—ordered sets of pixels. Pixels are sorted into a specified number of bins. The pixels are then given new values based upon the bins to which they are assigned.

BIP—band interleaved by pixel. A form of data storage in which the values for each band are ordered within a given pixel. The pixels are arranged sequentially on the tape.

bit—a binary digit, meaning a number that can have two possible values 0 and 1, or off and on. A set of bits, however, can have many more values, depending upon the number of bits used. The number of values that can be expressed by a set of bits is 2 to the power of the number of bits used. For example, the number of values that can be expressed by 3 bits is 8 ($2^3 = 8$).

block of photographs—formed by the combined exposures of a flight. The block consists of a number of parallel strips with a sidelap of 20-30%.

blocked—a method of storing data on 9-track tapes so that there are more logical records in each physical record.

blocking factor—the number of logical records in each physical record. For instance, a record may contain 28,000 bytes, but only 4,000 columns due to a blocking factor of 7.

book map—a map laid out like the pages of a book. Each page fits on the paper used by the printer. There are neatlines and tick marks on all sides of every page.

Boolean—logical, based upon, or reducible to a true or false condition.

- border**—on a map, a line that usually encloses the entire map, not just the image area as does a neatline.
- boundary**—a neighborhood analysis technique that is used to detect boundaries between thematic classes.
- bpi**—bits per inch. A measure of data storage density for magnetic tapes.
- breakline**—an elevation polyline in which each vertex has its own X, Y, Z value.
- brightness value**—the quantity of a primary color (red, green, blue) to be output to a pixel on the display device. Also called intensity value, function memory value, pixel value, display value, and screen value.
- BSQ**—band sequential. A data storage format in which each band is contained in a separate file.
- buffer zone**—a specific area around a feature that is isolated for or from further analysis. For example, buffer zones are often generated around streams in site assessment studies so that further analyses exclude these areas that are often unsuitable for development.
- build**—the process of constructing the topology of a vector layer by processing points, lines, and polygons. See **clean**.
- bundle**—the unit of photogrammetric triangulation after each point measured in an image is connected with the perspective center by a straight light ray. There is one bundle of light rays for each image.
- bundle attitude**—defined by a spatial rotation matrix consisting of three angles (κ , ω , ϕ).
- bundle location**—defined by the perspective center, expressed in units of the specified map projection.
- byte**—8 bits of data.

-
- C**
- CAC**—see **Compressed Aeronautical Chart**.
- CAD**—see **computer-aided design**.
- cadastral map**—a map showing the boundaries of the subdivisions of land for purposes of describing and recording ownership or taxation.
- CADRG**—see **Compressed ADRG**.
- calibration certificate/report**—in aerial photography, the manufacturer of the camera specifies the interior orientation in the form of a certificate or report.
- Cartesian**—a coordinate system in which data are organized on a grid and points on the grid are referenced by their X,Y coordinates.
- cartography**—the art and science of creating maps.
- categorical data**—see **thematic data**.
- CCT**—see **computer compatible tape**.

CD-ROM—a read-only storage device read by a CD-ROM player.

cell—1. a $1^\circ \times 1^\circ$ area of coverage. DTED (Digital Terrain Elevation Data) are distributed in cells. 2. a pixel; grid cell.

cell size—the area that one pixel represents, measured in map units. For example, one cell in the image may represent an area $30' \times 30'$ on the ground. Sometimes called pixel size.

center of the scene—the center pixel of the center scan line; the center of a satellite image.

central processing unit—(CPU) “the part of a computer which controls all the other parts. . .the CPU consists of the control unit, the arithmetic and logic unit (ALU) and memory (registers, cache, RAM and ROM) as well as various temporary buffers and other logic” (Free On-Line Dictionary of Computing 1999).

character—a number, letter, or punctuation symbol. One character usually occupies one byte when stored on a computer.

check point—additional ground points used to independently verify the degree of accuracy of a triangulation.

check point analysis—the act of using check points to independently verify the degree of accuracy of a triangulation.

chi-square distribution—a nonsymmetrical data distribution: its curve is characterized by a tail that represents the highest and least frequent data values. In classification thresholding, the tail represents the pixels that are most likely to be classified incorrectly.

choropleth map—a map portraying properties of a surface using area symbols. Area symbols usually represent categorized classes of the mapped phenomenon.

CIB—see **Controlled Image Base**.

city-block distance—the physical or spectral distance that is measured as the sum of distances that are perpendicular to one another.

class—a set of pixels in a GIS file that represents areas that share some condition. Classes are usually formed through classification of a continuous raster layer.

class value—a data file value of a thematic file that identifies a pixel as belonging to a particular class.

classification—the process of assigning the pixels of a continuous raster image to discrete categories.

classification accuracy table—for accuracy assessment, a list of known values of reference pixels, supported by some ground truth or other *a priori* knowledge of the true class, and a list of the classified values of the same pixels, from a classified file to be tested.

classification scheme—(or classification system) a set of target classes. The purpose of such a scheme is to provide a framework for organizing and categorizing the information that can be extracted from the data.

- clean**—the process of constructing the topology of a vector layer by processing lines and polygons. See **build**.
- client**—on a computer on a network, a program that accesses a server utility that is on another machine on the network.
- clump**—a contiguous group of pixels in one class. Also called raster region.
- clustering**—unsupervised training; the process of generating signatures based on the natural groupings of pixels in image data when they are plotted in spectral space.
- clusters**—the natural groupings of pixels when plotted in spectral space.
- CMY**—cyan, magenta, yellow. Primary colors of pigment used by printers, whereas display devices use RGB.
- CNES**—Centre National d'Etudes Spatiales. The corporation was founded in 1961. It provides support for ESA. CNES suggests and executes programs. (CNES 1998)
- coefficient**—one number in a matrix, or a constant in a polynomial expression.
- coefficient of variation**—a scene-derived parameter that is used as input to the Sigma and Local Statistics radar enhancement filters.
- collinearity**—a nonlinear mathematical model that photogrammetric triangulation is based upon. Collinearity equations describe the relationship among image coordinates, ground coordinates, and orientation parameters.
- colorcell**—the location where the data file values are stored in the colormap. The red, green, and blue values assigned to the colorcell control the brightness of the color guns for the displayed pixel.
- color guns**—on a display device, the red, green, and blue phosphors that are illuminated on the picture tube in varying brightnesses to create different colors. On a color printer, color guns are the devices that apply cyan, yellow, magenta, and sometimes black ink to paper.
- colormap**—an ordered set of colorcells, which is used to perform a function on a set of input values.
- color printer**—a printer that prints color or black and white imagery, as well as text. ERDAS IMAGINE supports several color printers.
- color scheme**—a set of lookup tables that assigns red, green, and blue brightness values to classes when a layer is displayed.
- composite map**—a map on which the combined information from different thematic maps is presented.
- Compressed ADRG**—(CADRG) a military data product based upon the general RPF specification.
- Compressed Aeronautical Chart**—(CAC) precursor to CADRG.
- Compressed Raster Graphics**—(CRG) precursor to CADRG.

compromise projection—a map projection that compromises among two or more of the map projection properties of conformality, equivalence, equidistance, and true direction.

computer-aided design—(CAD) computer application used for design and GPS survey.

computer compatible tape—(CCT) a magnetic tape used to transfer and store digital data.

confidence level—the percentage of pixels that are believed to be misclassified.

conformal—a map or map projection that has the property of conformality, or true shape.

conformality—the property of a map projection to represent true shape, wherein a projection preserves the shape of any small geographical area. This is accomplished by exact transformation of angles around points.

conic projection—a map projection that is created from projecting the surface of the Earth to the surface of a cone.

connectivity radius—the distance (in pixels) that pixels can be from one another to be considered contiguous. The connectivity radius is used in connectivity analysis.

contiguity analysis—a study of the ways in which pixels of a class are grouped together spatially. Groups of contiguous pixels in the same class, called raster regions, or clumps, can be identified by their sizes and manipulated.

contingency matrix—a matrix that contains the number and percentages of pixels that were classified as expected.

continuous—a term used to describe raster data layers that contain quantitative and related values. See **continuous data**.

continuous data—a type of raster data that are quantitative (measuring a characteristic) and have related, continuous values, such as remotely sensed images (e.g., Landsat, SPOT, etc.).

contour map—a map in which a series of lines connects points of equal elevation.

contrast stretch—the process of reassigning a range of values to another range, usually according to a linear function. Contrast stretching is often used in displaying continuous raster layers, since the range of data file values is usually much narrower than the range of brightness values on the display device.

control point—a point with known coordinates in the ground coordinate system, expressed in the units of the specified map projection.

Controlled Image Base—(CIB) a military data product based upon the general RPF specification.

convolution filtering—the process of averaging small sets of pixels across an image. Used to change the spatial frequency characteristics of an image.

- convolution kernel**—a matrix of numbers that is used to average the value of each pixel with the values of surrounding pixels in a particular way. The numbers in the matrix serve to weight this average toward particular pixels.
- coordinate system**—a method of expressing location. In two-dimensional coordinate systems, locations are expressed by a column and row, also called x and y.
- correlation threshold**—a value used in rectification to determine whether to accept or discard GCPs. The threshold is an absolute value threshold ranging from 0.000 to 1.000.
- correlation windows**—windows that consist of a local neighborhood of pixels. One example is square neighborhoods (e.g., 3×3 , 5×5 , 7×7 pixels).
- corresponding GCPs**—the GCPs that are located in the same geographic location as the selected GCPs, but are selected in different files.
- covariance**—measures the tendencies of data file values for the same pixel, but in different bands, to vary with each other in relation to the means of their respective bands. These bands must be linear. Covariance is defined as the average product of the differences between the data file values in each band and the mean of each band.
- covariance matrix**—a square matrix that contains all of the variances and covariances within the bands in a data file.
- CPU**— see **central processing unit**.
- credits**—on maps, the text that can include the data source and acquisition date, accuracy information, and other details that are required for or helpful to readers.
- CRG**—see **Compressed Raster Graphics**.
- crisp filter**—a filter used to sharpen the overall scene luminance without distorting the interband variance content of the image.
- cross correlation**—a calculation that computes the correlation coefficient of the gray values between the template window and the search window.
- cubic convolution**—a method of resampling that uses the data file values of sixteen pixels in a 4×4 window to calculate an output data file value with a cubic function.
- current directory**—also called default directory, it is the directory that you are in. It is the default path.
- cylindrical projection**—a map projection that is created from projecting the surface of the Earth to the surface of a cylinder.

D **dangling node**—a line that does not close to form a polygon, or that extends past an intersection.

data—1. in the context of remote sensing, a computer file containing numbers that represent a remotely sensed image, and can be processed to display that image.
2. a collection of numbers, strings, or facts that requires some processing before it is meaningful.

database (one word)—a relational data structure usually used to store tabular information. Examples of popular databases include SYBASE, dBase, Oracle, INFO, etc.

data base (two words)—in ERDAS IMAGINE, a set of continuous and thematic raster layers, vector layers, attribute information, and other kinds of data that represent one area of interest. A data base is usually part of a GIS.

data file—a computer file that contains numbers that represent an image.

data file value—each number in an image file. Also called file value, image file value, DN, brightness value, pixel.

datum—see **reference plane**.

DCT—see **Discrete Cosine Transformation**.

decision rule—an equation or algorithm that is used to classify image data after signatures have been created. The decision rule is used to process the data file values based upon the signature statistics.

decorrelation stretch—a technique used to stretch the principal components of an image, not the original image.

default directory—see **current directory**.

Defense Mapping Agency—(DMA) agency that supplies VPF, ARC digital raster, DRG, ADRG, and DTED files.

degrees of freedom—when chi-square statistics are used in thresholding, the number of bands in the classified file.

DEM—see **digital elevation model**.

densify—the process of adding vertices to selected lines at a user-specified tolerance.

density—1. the number of bits per inch on a magnetic tape. 9-track tapes are commonly stored at 1600 and 6250 bpi. 2. a neighborhood analysis technique that outputs the number of pixels that have the same value as the analyzed pixel in a user-specified window.

derivative map—a map created by altering, combining, or analyzing other maps.

descriptor—see **attribute**.

desktop scanners—general purpose devices that lack the image detail and geometric accuracy of photogrammetric quality units, but are much less expensive.

detector—the device in a sensor system that records electromagnetic radiation.

developable surface—a flat surface, or a surface that can be easily flattened by being cut and unrolled, such as the surface of a cone or a cylinder.

DFT—see **Discrete Fourier Transform**.

DGPS—see **Differential Correction**.

Differential Correction—(DGPS) can be used to remove the majority of the effects of Selective Availability.

digital elevation model—(DEM) continuous raster layers in which data file values represent elevation. DEMs are available from the USGS at 1:24,000 and 1:250,000 scale, and can be produced with terrain analysis programs, IMAGINE IFSAR DEM, IMAGINE OrthoMAX™, and IMAGINE StereoSAR DEM.

Digital Number—(DN) variation in pixel intensity due to composition of what it represents. For example, the DN of water is different from that of land. DN is expressed in a value—typically from 0-255.

digital orthophoto—an aerial photo or satellite scene that has been transformed by the orthogonal projection, yielding a map that is free of most significant geometric distortions.

digital orthophoto quadrangle—(DOQ) a computer-generated image of an aerial photo. (USGS 1999)

digital photogrammetry—photogrammetry as applied to digital images that are stored and processed on a computer. Digital images can be scanned from photographs or can be directly captured by digital cameras.

Digital Line Graph—(DLG) a vector data format created by the USGS.

Digital Terrain Elevation Data—(DTED) data produced by the DMA. DTED data comes in two types, both in Arc/second format: DTED 1—a 1° × 1° area of coverage, and DTED 2—a 1° × 1° or less area of coverage.

digital terrain model—(DTM) a discrete expression of topography in a data array, consisting of a group of planimetric coordinates (X,Y) and the elevations of the ground points and breaklines.

digitized raster graphic—(DRG) a digital replica of DMA hardcopy graphic products. See also **ADRG**.

digitizing—any process that converts nondigital data into numeric data, usually to be stored on a computer. In ERDAS IMAGINE, digitizing refers to the creation of vector data from hardcopy materials or raster images that are traced using a digitizer keypad on a digitizing tablet, or a mouse on a display device.

DIME—see **Dual Independent Map Encoding**.

dimensionality—a term referring to the number of bands being classified. For example, a data file with three bands is said to be three-dimensional, since three-dimensional spectral space is plotted to analyze the data.

directory—an area of a computer disk that is designated to hold a set of files. Usually, directories are arranged in a tree structure, in which directories can also contain many levels of subdirectories.

- Discrete Cosine Transformation**—(DCT) an element of a commonly used form of JPEG, which is a compression technique.
- Discrete Fourier Transform**—(DFT) method of removing striping and other noise from radar images. See also **Fast Fourier Transform**.
- displacement**—the degree of geometric distortion for a point that is not on the nadir line.
- display device**—the computer hardware consisting of a memory board and a monitor. It displays a visible image from a data file or from some user operation.
- display driver**—the ERDAS IMAGINE utility that interfaces between the computer running ERDAS IMAGINE software and the display device.
- display memory**—the subset of image memory that is actually viewed on the display screen.
- display pixel**—one grid location on a display device or printout.
- display resolution**—the number of pixels that can be viewed on the display device monitor, horizontally and vertically (i.e., 512×512 or 1024×1024).
- distance**—see **Euclidean distance**, **spectral distance**.
- distance image file**—a one-band, 16-bit file that can be created in the classification process, in which each data file value represents the result of the distance equation used in the program. Distance image files generally have a chi-square distribution.
- distribution**—the set of frequencies with which an event occurs, or the set of probabilities that a variable has a particular value.
- distribution rectangles**—(DR) the geographic data sets into which ADRG data are divided.
- dithering**—a display technique that is used in ERDAS IMAGINE to allow a smaller set of colors appear to be a larger set of colors.
- divergence**—a statistical measure of distance between two or more signatures. Divergence can be calculated for any combination of bands used in the classification; bands that diminish the results of the classification can be ruled out.
- diversity**—a neighborhood analysis technique that outputs the number of different values within a user-specified window.
- DLG**—see **Digital Line Graph**.
- DMA**—see **Defense Mapping Agency**.
- DN**—see **Digital Number**.
- DOQ**—see **digital orthophoto quadrangle**.
- dot patterns**—the matrices of dots used to represent brightness values on hardcopy maps and images.

dots per inch—(DPI) when referring to the resolution of an output device, such as a printer, the number of dots that are printed per unit—for example, 300 dots per inch.

double precision—a measure of accuracy in which fifteen significant digits can be stored for a coordinate.

downsampling—the skipping of pixels during the display or processing of the scanning process.

DPI—see **dots per inch**.

DR—see **distribution rectangles**.

DTED—see **Digital Terrain Elevation Data**.

DTM—see **digital terrain model**.

Dual Independent Map Encoding—(DIME) a type of ETAK feature wherein a line is created along with a corresponding ACODE (arc attribute) record. The coordinates are stored in Lat/Lon decimal degrees. Each record represents a single linear feature.

DXF—Data Exchange Format. A format for storing vector data in ASCII files, used by AutoCAD software.

dynamic range—see **radiometric resolution**.

E

Earth Observation Satellite Company—(EOSAT) a private company that directs the Landsat satellites and distributes Landsat imagery.

Earth Resources Observation Systems—(EROS) a division of the USGS National Mapping Division. EROS is involved with managing data and creating systems, as well as research. (USGS)

Earth Resources Technology Satellites—(ERTS) in 1972, NASA's first civilian program to acquire remotely sensed digital satellite data, later renamed to Landsat.

EDC—see **EROS Data Center**.

edge detector—a convolution kernel, which is usually a zero-sum kernel, that smooths out or zeros out areas of low spatial frequency and creates a sharp contrast where spatial frequency is high. High spatial frequency is at the edges between homogeneous groups of pixels.

edge enhancer—a high-frequency convolution kernel that brings out the edges between homogeneous groups of pixels. Unlike an edge detector, it only highlights edges, it does not necessarily eliminate other features.

eigenvalue—the length of a principal component that measures the variance of a principal component band. See also **principal components**.

eigenvector—the direction of a principal component represented as coefficients in an eigenvector matrix which is computed from the eigenvalues. See also **principal components**.

electromagnetic—(EM) type of spectrum consisting of different regions such as thermal infrared and long-wave and short-wave reflective.

electromagnetic radiation—(EMR) the energy transmitted through space in the form of electric and magnetic waves.

electromagnetic spectrum—the range of electromagnetic radiation extending from cosmic waves to radio waves, characterized by frequency or wavelength.

element—an entity of vector data, such as a point, line, or polygon.

elevation data—see **terrain data**, **DEM**.

ellipse—a two-dimensional figure that is formed in a two-dimensional scatterplot when both bands plotted have normal distributions. The ellipse is defined by the standard deviations of the input bands. Ellipse plots are often used to test signatures before classification.

EM—see **electromagnetic**.

EML—see **ERDAS Macro Language**.

EMR—see **electromagnetic radiation**.

end-of-file mark—(EOF) usually a half-inch strip of blank tape that signifies the end of a file that is stored on magnetic tape.

end-of-volume mark—(EOV) usually three EOFs marking the end of a tape.

Enhanced Thematic Mapper Plus—(ETM+) the observing instrument on Landsat 7.

enhancement—the process of making an image more interpretable for a particular application. Enhancement can make important features of raw, remotely sensed data more interpretable to the human eye.

entity—an AutoCAD drawing element that can be placed in an AutoCAD drawing with a single command.

Environmental Systems Research Institute—(ESRI) company based in Redlands, California, which produces software such as ArcInfo and ArcView. ESRI has created many data formats, including GRID and GRID Stack.

EOF—see **end-of-file mark**.

EOSAT—see **Earth Observation Satellite Company**.

EOV— see **end-of-volume mark**.

ephemeris data—contained in the header of the data file of a SPOT scene, provides information about the recording of the data and the satellite orbit.

epipolar stereopair—a stereopair without y-parallax.

equal area—see **equivalence**.

- equatorial aspect**—a map projection that is centered around the equator or a point on the equator.
- equidistance**—the property of a map projection to represent true distances from an identified point.
- equivalence**—the property of a map projection to represent all areas in true proportion to one another.
- ERDAS Macro Language**—(EML) computer language that can be used to create custom dialogs in ERDAS IMAGINE, or to edit existing dialogs and functions for your specific application.
- EROS**—see **Earth Resources Observation Systems**.
- EROS Data Center**—(EDC) a division of USGS, located in Sioux Falls, SD, which is the primary receiving center for Landsat 7 data.
- error matrix**—in classification accuracy assessment, a square matrix showing the number of reference pixels that have the same values as the actual classified points.
- ERS-1**—the European Space Agency's (ESA) radar satellite launched in July 1991, currently provides the most comprehensive radar data available. ERS-2 was launched in 1995.
- ERTS**—see **Earth Resources Technology Satellites**.
- ESA**—see **European Space Agency**.
- ESRI**—see **Environmental Systems Research Institute**.
- ETAK MapBase**—an ASCII digital street centerline map product available from ETAK, Inc. (Menlo Park, California).
- ETM+**—see **Enhanced Thematic Mapper Plus**.
- Euclidean distance**—the distance, either in physical or abstract (e.g., spectral) space, that is computed based on the equation of a straight line.
- exposure station**—during image acquisition, each point in the flight path at which the camera exposes the film.
- extend**—the process of moving selected dangling lines up a specified distance so that they intersect existing lines.
- extension**—the three letters after the period in a file name that usually identify the type of file.
- extent**—1. the image area to be displayed in a Viewer. 2. the area of the Earth's surface to be mapped.
- exterior orientation**—all images of a block of aerial photographs in the ground coordinate system are computed during photogrammetric triangulation using a limited number of points with known coordinates. The exterior orientation of an image consists of the exposure station and the camera attitude at this moment.

exterior orientation parameters—the perspective center's ground coordinates in a specified map projection, and three rotation angles around the coordinate axes.

European Space Agency—(ESA) company with two satellites, ERS-1 and ERS-2, that collect radar data. For more information, visit the ESA web site at <http://www.esa.int>.

extract—selected bands of a complete set of NOAA AVHRR data.

F

false color—a color scheme in which features have expected colors. For instance, vegetation is green, water is blue, etc. These are not necessarily the true colors of these features.

false easting—an offset between the y-origin of a map projection and the y-origin of a map. Typically used so that no y-coordinates are negative.

false northing—an offset between the x-origin of a map projection and the x-origin of a map. Typically used so that no x-coordinates are negative.

fast format—a type of BSQ format used by EOSAT to store Landsat TM data.

Fast Fourier Transform—(FFT) a type of Fourier Transform faster than the DFT. Designed to remove noise and periodic features from radar images. It converts a raster image from the spatial domain into a frequency domain image.

feature based matching—an image matching technique that determines the correspondence between two image features.

feature collection—the process of identifying, delineating, and labeling various types of natural and human-made phenomena from remotely-sensed images.

feature extraction—the process of studying and locating areas and objects on the ground and deriving useful information from images.

feature space—an abstract space that is defined by spectral units (such as an amount of electromagnetic radiation).

feature space area of interest—a user-selected area of interest (AOI) that is selected from a feature space image.

feature space image—a graph of the data file values of one band of data against the values of another band (often called a scatterplot).

FFT—see **Fast Fourier Transform**.

fiducial center—the center of an aerial photo.

fiducials—four or eight reference markers fixed on the frame of an aerial metric camera and visible in each exposure. Fiducials are used to compute the transformation from data file to image coordinates.

field—in an attribute database, a category of information about each class or feature, such as Class name and Histogram.

field of view—(FOV) in perspective views, an angle that defines how far the view is generated to each side of the line of sight.

- file coordinates**—the location of a pixel within the file in x,y coordinates. The upper left file coordinate is usually 0,0.
- file pixel**—the data file value for one data unit in an image file.
- file specification or filespec**—the complete file name, including the drive and path, if necessary. If a drive or path is not specified, the file is assumed to be in the current drive and directory.
- filled**—referring to polygons; a filled polygon is solid or has a pattern, but is not transparent. An unfilled polygon is simply a closed vector that outlines the area of the polygon.
- filtering**—the removal of spatial or spectral features for data enhancement. Convolution filtering is one method of spatial filtering. Some texts may use the terms filtering and spatial filtering synonymously.
- flip**—the process of reversing the from-to direction of selected lines or links.
- focal length**—the orthogonal distance from the perspective center to the image plane.
- focal operations**—filters that use a moving window to calculate new values for each pixel in the image based on the values of the surrounding pixels.
- focal plane**—the plane of the film or scanner used in obtaining an aerial photo.
- Fourier analysis**—an image enhancement technique that was derived from signal processing.
- FOV**—see **field of view**.
- from-node**—the first vertex in a line.
- full set**—all bands of a NOAA AVHRR data set.
- function memories**—areas of the display device memory that store the lookup tables, which translate image memory values into brightness values.
- function symbol**—an annotation symbol that represents an activity. For example, on a map of a state park, a symbol of a tent would indicate the location of a camping area.
- Fuyo 1 (JERS-1)**—the Japanese radar satellite launched in February 1992.

G

- GAC**—see **global area coverage**.
- GBF**—see **Geographic Base File**.
- GCP**—see **ground control point**.
- GCP matching**—for image-to-image rectification, a GCP selected in one image is precisely matched to its counterpart in the other image using the spectral characteristics of the data and the transformation matrix.
- GCP prediction**—the process of picking a GCP in either coordinate system and automatically locating that point in the other coordinate system based on the current transformation parameters.

- generalize**—the process of weeding vertices from selected lines using a specified tolerance.
- geocentric coordinate system**—a coordinate system that has its origin at the center of the Earth ellipsoid. The Z_G -axis equals the rotational axis of the Earth, and the X_G -axis passes through the Greenwich meridian. The Y_G -axis is perpendicular to both the Z_G -axis and the X_G -axis, so as to create a three-dimensional coordinate system that follows the right-hand rule.
- geocoded data**—an image(s) that has been rectified to a particular map projection and cell size and has had radiometric corrections applied.
- Geographic Base File**—(GBF) along with DIME, sometimes provides the cartographic base for TIGER/Line files, which cover the US, Puerto Rico, Guam, the Virgin Islands, American Samoa, and the Trust Territories of the Pacific.
- geographic information system**—(GIS) a unique system designed for a particular application that stores, enhances, combines, and analyzes layers of geographic data to produce interpretable information. A GIS may include computer images, hardcopy maps, statistical data, and any other data needed for a study, as well as computer software and human knowledge. GISs are used for solving complex geographic planning and management problems.
- geographical coordinates**—a coordinate system for explaining the surface of the Earth. Geographical coordinates are defined by latitude and by longitude (Lat/Lon), with respect to an origin located at the intersection of the equator and the prime (Greenwich) meridian.
- geometric correction**—the correction of errors of skew, rotation, and perspective in raw, remotely sensed data.
- georeferencing**—the process of assigning map coordinates to image data and resampling the pixels of the image to conform to the map projection grid.
- GeoTIFF**—TIFF files that are geocoded.
- gigabyte**—(Gb) about one billion bytes.
- GIS**—see **geographic information system**.
- GIS file**—a single-band ERDAS Ver. 7.X data file in which pixels are divided into discrete categories.
- global area coverage**—(GAC) a type of NOAA AVHRR data with a spatial resolution of 4×4 km.
- global operations**—functions that calculate a single value for an entire area, rather than for each pixel like focal functions.
- GLobal NAVigation Satellite System**—(GLONASS) a satellite-based navigation system produced by the Russian Space Forces. It provides three-dimensional locations, velocity, and time measurements for both civilian and military applications. GLONASS started its mission in 1993. (Magellan Corporation 1999)
- Global Ozone Monitoring Experiment**—(GOME) instrument aboard ESA's ERS-2 satellite, which studies atmospheric chemistry (European Space Agency 1996).

- Global Positioning System**—(GPS) system used for the collection of GCPs, which uses orbiting satellites to pinpoint precise locations on the Earth's surface.
- GLONASS**—see **GLOBAL NAVIGATION SATELLITE SYSTEM**.
- .gmd file**—the ERDAS IMAGINE graphical model file created with Model Maker (Spatial Modeler).
- gnomonic**—an azimuthal projection obtained from a perspective at the center of the Earth.
- GOME**—see **Global Ozone Monitoring Experiment**.
- GPS**—see **Global Positioning System**.
- graphical modeling**—a technique used to combine data layers in an unlimited number of ways using icons to represent input data, functions, and output data. For example, an output layer created from modeling can represent the desired combination of themes from many input layers.
- graphical model**—a model created with Model Maker (Spatial Modeler). Graphical models are put together like flow charts and are stored in .gmd files.
- Graphical User Interface**—(GUI) the dialogs and menus of ERDAS IMAGINE that enable you to execute commands to analyze your imagery.
- graticule**—the network of parallels of latitude and meridians of longitude applied to the global surface and projected onto maps.
- gray scale**—a color scheme with a gradation of gray tones ranging from black to white.
- great circle**—an arc of a circle for which the center is the center of the Earth. A great circle is the shortest possible surface route between two points on the Earth.
- GRID**—a compressed tiled raster data structure that is stored as a set of files in a directory, including files to keep the attributes of the GRID.
- grid cell**—a pixel.
- grid lines**—intersecting lines that indicate regular intervals of distance based on a coordinate system. Sometimes called a graticule.
- GRID Stack**—multiple GRIDs to be treated as a multilayer image.
- ground control point**—(GCP) specific pixel in image data for which the output map coordinates (or other output coordinates) are known. GCPs are used for computing a transformation matrix, for use in rectifying an image.
- ground coordinate system**—a three-dimensional coordinate system which utilizes a known map projection. Ground coordinates (X,Y,Z) are usually expressed in feet or meters.
- ground truth**—data that are taken from the actual area being studied.
- ground truthing**—the acquisition of knowledge about the study area from field work, analysis of aerial photography, personal experience, etc. Ground truth data are considered to be the most accurate (true) data available about the area of study.

GUI—see **Graphical User Interface**.

H

halftoning—the process of using dots of varying size or arrangements (rather than varying intensity) to form varying degrees of a color.

hardcopy output—any output of digital computer (softcopy) data to paper.

HARN—see **High Accuracy Reference Network**.

header file—a file usually found before the actual image data on tapes or CD-ROMs that contains information about the data, such as number of bands, upper left coordinates, map projection, etc.

header record—the first part of an image file that contains general information about the data in the file, such as the number of columns and rows, number of bands, database coordinates of the upper left corner, and the pixel depth. The contents of header records vary depending on the type of data.

HFA—see **Hierarchal File Architecture System**.

Hierarchal File Architecture System—(HFA) a format that allows different types of information about a file to be stored in a tree-structured fashion. The tree is made of nodes that contain information such as ephemeris data.

High Accuracy Reference Network—(HARN) HARN is based on the GRS 1980 spheroid, and can be used to perform State Plane calculations.

high-frequency kernel—a convolution kernel that increases the spatial frequency of an image. Also called high-pass kernel.

High Resolution Picture Transmission—(HRPT) the direct transmission of AVHRR data in real-time with the same resolution as LAC.

High Resolution Visible Infrared—(HR VIR) a pushbroom scanner on the SPOT 4 satellite, which captures information in the visible and near-infrared bands (SPOT Image 1999).

High Resolution Visible sensor—(HRV) a pushbroom scanner on a SPOT satellite that takes a sequence of line images while the satellite circles the Earth.

histogram—a graph of data distribution, or a chart of the number of pixels that have each possible data file value. For a single band of data, the horizontal axis of a histogram graph is the range of all possible data file values. The vertical axis is a measure of pixels that have each data value.

histogram equalization—the process of redistributing pixel values so that there are approximately the same number of pixels with each value within a range. The result is a nearly flat histogram.

histogram matching—the process of determining a lookup table that converts the histogram of one band of an image or one color gun to resemble another histogram.

horizontal control—the horizontal distribution of GCPs in aerial triangulation (x,y - planimetry).

host workstation—a CPU, keyboard, mouse, and a display.

HRPT—see **High Resolution Picture Transmission**.

HRV—see **High Resolution Visible sensor**.

HR VIR—see **High Resolution Visible Infrared**.

hue—a component of IHS (intensity, hue, saturation) that is representative of the color or dominant wavelength of the pixel. It varies from 0 to 360. Blue = 0 (and 360), magenta = 60, red = 120, yellow = 180, green = 240, and cyan = 300.

hyperspectral sensors—the imaging sensors that record multiple bands of data, such as the AVIRIS with 224 bands.

I

IARR—see **Internal Average Relative Reflectance**.

IFFT—see **Inverse Fast Fourier Transform**.

IFOV—see **instantaneous field of view**.

IGES—see **Initial Graphics Exchange Standard files**.

IHS—intensity, hue, saturation. An alternate color space from RGB (red, green, blue). This system is advantageous in that it presents colors more nearly as perceived by the human eye. See **intensity**, **hue**, and **saturation**.

image—a picture or representation of an object or scene on paper, or a display screen. Remotely sensed images are digital representations of the Earth.

image algebra—any type of algebraic function that is applied to the data file values in one or more bands.

image center—the center of the aerial photo or satellite scene.

image coordinate system—the location of each point in the image is expressed for purposes of photogrammetric triangulation.

image data—digital representations of the Earth that can be used in computer image processing and GIS analyses.

image file—a file containing raster image data. Image files in ERDAS IMAGINE have the extension .img. Image files from the ERDAS Ver. 7.X series software have the extension .LAN or .GIS.

image matching—the automatic acquisition of corresponding image points on the overlapping area of two images.

image memory—the portion of the display device memory that stores data file values (which may be transformed or processed by the software that accesses the display device).

image pair—see **stereopair**.

image processing—the manipulation of digital image data, including (but not limited to) enhancement, classification, and rectification operations.

image pyramid—a data structure consisting of the same image represented several times, at a decreasing spatial resolution each time. Each level of the pyramid contains the image at a particular resolution.

image scale—(SI) expresses the average ratio between a distance in the image and the same distance on the ground. It is computed as focal length divided by the flying height above the mean ground elevation.

image space coordinate system—identical to image coordinates, except that it adds a third axis (z) that is used to describe positions inside the camera. The units are usually in millimeters or microns.

IMC—see **International Map Committee**.

.img file—(also, **image file**) an ERDAS IMAGINE file that stores continuous or thematic raster layers.

IMW—see **International Map of the World**.

inclination—the angle between a vertical on the ground at the center of the scene and a light ray from the exposure station, which defines the degree of off-nadir viewing when the scene was recorded.

indexing—a function applied to thematic layers that adds the data file values of two or more layers together, creating a new output layer. Weighting factors can be applied to one or more layers to add more importance to those layers in the final sum.

index map—a reference map that outlines the mapped area, identifies all of the component maps for the area if several map sheets are required, and identifies all adjacent map sheets.

Indian Remote Sensing Satellite—(IRS) satellites operated by Space Imaging, including IRS-1A, IRS-1B, IRS-1C, and IRS-1D.

indices—the process used to create output images by mathematically combining the DN values of different bands.

information—something that is independently meaningful, as opposed to data, which are not independently meaningful.

Initial Graphics Exchange Standard files—(IGES) files often used to transfer CAD data between systems. IGES Version 3.0 format, published by the U.S. Department of Commerce, is in uncompressed ASCII format only.

initialization—a process that ensures all values in a file or in computer memory are equal until additional information is added or processed to overwrite these values. Usually the initialization value is 0. If initialization is not performed on a data file, there could be random data values in the file.

inset map—a map that is an enlargement of some congested area of a smaller scale map, and that is usually placed on the same sheet with the smaller scale main map.

instantaneous field of view—(IFOV) a measure of the area viewed by a single detector on a scanning system in a given instant in time.

intensity—a component of IHS (intensity, hue, saturation), which is the overall brightness of the scene and varies from 0 (black) to 1 (white).

interferometry—method of subtracting the phase of one SAR image from another to derive height information.

interior orientation—defines the geometry of the sensor that captured a particular image.

Internal Average Relative Reflectance—(IARR) a technique designed to compensate for atmospheric contamination of the spectra.

International Map Committee—(IMC) located in London, the committee responsible for creating the International Map of the World series.

International Map of the World—(IMW) a series of maps produced by the International Map Committee. Maps are in 1:1,000,000 scale.

intersection—the area or set that is common to two or more input areas or sets.

interval data—a type of data in which thematic class values have a natural sequence, and in which the distances between values are meaningful.

Inverse Fast Fourier Transform—(IFFT) used after the Fast Fourier Transform to transform a Fourier image back into the spatial domain. See also **Fast Fourier Transform**.

IR—infrared portion of the electromagnetic spectrum. See also **electromagnetic spectrum**.

IRS—see **Indian Remote Sensing Satellite**.

isarithmic map—a map that uses isarithms (lines connecting points of the same value for any of the characteristics used in the representation of surfaces) to represent a statistical surface. Also called an isometric map.

ISODATA clustering—see **Iterative Self-Organizing Data Analysis Technique**.

island—A single line that connects with itself.

isopleth map—a map on which isopleths (lines representing quantities that cannot exist at a point, such as population density) are used to represent some selected quantity.

iterative—a term used to describe a process in which some operation is performed repeatedly.

Iterative Self-Organizing Data Analysis Technique—(ISODATA clustering) a method of clustering that uses spectral distance as in the sequential method, but iteratively classifies the pixels, redefines the criteria for each class, and classifies again, so that the spectral distance patterns in the data gradually emerge.

J

JERS-1 (Fuyo 1)—the Japanese radar satellite launched in February 1992.

Glossary

K

Jet Propulsion Laboratories—(JPL) “the lead U.S. center for robotic exploration of the solar system.” JPL is managed for NASA by the California Institute of Technology. For more information, see the JPL web site at <http://www.jpl.nasa.gov> (NASA 1999).

JFIF—see **JPEG File Interchange Format**.

join—the process of interactively entering the side lot lines when the front and rear lines have already been established.

Joint Photographic Experts Group—(JPEG) 1. a group responsible for creating a set of compression techniques. 2. Compression techniques are also called JPEG.

JPEG—see **Joint Photographic Experts Group**.

JPEG File Interchange Format—(JFIF) standard file format used to store JPEG-compressed imagery.

JPL—see **Jet Propulsion Laboratories**.

K

Kappa coefficient—a number that expresses the proportionate reduction in error generated by a classification process compared with the error of a completely random classification.

kernel—see **convolution kernel**.

L

label—in annotation, the text that conveys important information to the reader about map features.

label point—a point within a polygon that defines that polygon.

LAC—see **local area coverage**.

.LAN files—multiband ERDAS Ver. 7.X image files (the name originally derived from the Landsat satellite). LAN files usually contain raw or enhanced remotely sensed data.

land cover map—a map of the visible ground features of a scene, such as vegetation, bare land, pasture, urban areas, etc.

Landsat—a series of Earth-orbiting satellites that gather MSS and TM imagery, operated by EOSAT.

large-scale—a description used to represent a map or data file having a large ratio between the area on the map (such as inches or pixels), and the area that is represented (such as feet). In large-scale image data, each pixel represents a small area on the ground, such as SPOT data, with a spatial resolution of 10 or 20 meters.

Lat/Lon—Latitude/Longitude, a map coordinate system.

- layer**—1. a band or channel of data. 2. a single band or set of three bands displayed using the red, green, and blue color guns of the ERDAS IMAGINE Viewer. A layer could be a remotely sensed image, an aerial photograph, an annotation layer, a vector layer, an area of interest layer, etc. 3. a component of a GIS data base that contains all of the data for one theme. A layer consists of a thematic image file, and may also include attributes.
- least squares correlation**—uses the least squares estimation to derive parameters that best fit a search window to a reference window.
- least squares regression**—the method used to calculate the transformation matrix from the GCPs. This method is discussed in statistics textbooks.
- legend**—the reference that lists the colors, symbols, line patterns, shadings, and other annotation that is used on a map, and their meanings. The legend often includes the map's title, scale, origin, and other information.
- lettering**—the manner in which place names and other labels are added to a map, including letter spacing, orientation, and position.
- level 1A (SPOT)**—an image that corresponds to raw sensor data to which only radiometric corrections have been applied.
- level 1B (SPOT)**—an image that has been corrected for the Earth's rotation and to make all pixels 10×10 on the ground. Pixels are resampled from the level 1A sensor data by cubic polynomials.
- level slice**—the process of applying a color scheme by equally dividing the input values (image memory values) into a certain number of bins, and applying the same color to all pixels in each bin. Usually, a ROYGBIV (red, orange, yellow, green, blue, indigo, violet) color scheme is used.
- line**—1. a vector data element consisting of a line (the set of pixels directly between two points), or an unclosed set of lines. 2. a row of pixels in a data file.
- line dropout**—a data error that occurs when a detector in a satellite either completely fails to function or becomes temporarily overloaded during a scan. The result is a line, or partial line of data with incorrect data file values creating a horizontal streak until the detector(s) recovers, if it recovers.
- linear**—a description of a function that can be graphed as a straight line or a series of lines. Linear equations (transformations) can generally be expressed in the form of the equation of a line or plane. Also called 1st-order.
- linear contrast stretch**—an enhancement technique that outputs new values at regular intervals.
- linear transformation**—a 1st-order rectification. A linear transformation can change location in X and/or Y, scale in X and/or Y, skew in X and/or Y, and rotation.
- line of sight**—in perspective views, the point(s) and direction from which the viewer is looking into the image.
- local area coverage**—(LAC) a type of NOAA AVHRR data with a spatial resolution of 1.1×1.1 km.

logical record—a series of bytes that form a unit on a 9-track tape. For example, all the data for one line of an image may form a logical record. One or more logical records make up a physical record on a tape.

long wave infrared region—(LWIR) the thermal or far-infrared region of the electromagnetic spectrum.

lookup table—(LUT) an ordered set of numbers that is used to perform a function on a set of input values. To display or print an image, lookup tables translate data file values into brightness values.

lossy—“a term describing a data compression algorithm which actually reduces the amount of information in the data, rather than just the number of bits used to represent that information” (Free On-Line Dictionary of Computing 1999).

low-frequency kernel—a convolution kernel that decreases spatial frequency. Also called low-pass kernel.

LUT—see **lookup table**.

LWIR—see **long wave infrared region**.

M

Machine Independent Format—(MIF) a format designed to store data in a way that it can be read by a number of different machines.

magnify—the process of displaying one file pixel over a block of display pixels. For example, if the magnification factor is 3, then each file pixel takes up a block of 3×3 display pixels. Magnification differs from zooming in that the magnified image is loaded directly to image memory.

magnitude—an element of an electromagnetic wave. Magnitude of a wave decreases exponentially as the distance from the transmitter increases.

Mahalanobis distance—a classification decision rule that is similar to the minimum distance decision rule, except that a covariance matrix is used in the equation.

majority—a neighborhood analysis technique that outputs the most common value of the data file values in a user-specified window.

MAP—see **Maximum A Posteriori**.

map—a graphic representation of spatial relationships on the Earth or other planets.

map coordinates—a system of expressing locations on the Earth’s surface using a particular map projection, such as UTM, State Plane, or Polyconic.

map frame—an annotation element that indicates where an image is placed in a map composition.

map projection—a method of representing the three-dimensional spherical surface of a planet on a two-dimensional map surface. All map projections involve the transfer of latitude and longitude onto an easily flattened surface.

matrix—a set of numbers arranged in a rectangular array. If a matrix has i rows and j columns, it is said to be an $i \times j$ matrix.

- matrix analysis**—a method of combining two thematic layers in which the output layer contains a separate class for every combination of two input classes.
- matrix object**—in Model Maker (Spatial Modeler), a set of numbers in a two-dimensional array.
- maximum**—a neighborhood analysis technique that outputs the greatest value of the data file values in a user-specified window.
- Maximum A Posteriori**—(MAP) a filter (Gamma-MAP) that is designed to estimate the original DN value of a pixel, which it assumes is between the local average and the degraded DN.
- maximum likelihood**—a classification decision rule based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes, and that the input bands have normal distributions.
- .mdl file**—an ERDAS IMAGINE script model created with the Spatial Modeler Language.
- mean**—1. the statistical average; the sum of a set of values divided by the number of values in the set. 2. a neighborhood analysis technique that outputs the mean value of the data file values in a user-specified window.
- mean vector**—an ordered set of means for a set of variables (bands). For a data file, the mean vector is the set of means for all bands in the file.
- measurement vector**—the set of data file values for one pixel in all bands of a data file.
- median**—1. the central value in a set of data such that an equal number of values are greater than and less than the median. 2. a neighborhood analysis technique that outputs the median value of the data file values in a user-specified window.
- megabyte**—(Mb) about one million bytes.
- memory resident**—a term referring to the occupation of a part of a computer's RAM (random access memory), so that a program is available for use without being loaded into memory from disk.
- mensuration**—the measurement of linear or areal distance.
- meridian**—a line of longitude, going north and south. See **geographical coordinates**.
- MIF**—see **Machine Independent Format**.
- minimum**—a neighborhood analysis technique that outputs the least value of the data file values in a user-specified window.
- minimum distance**—a classification decision rule that calculates the spectral distance between the measurement vector for each candidate pixel and the mean vector for each signature. Also called spectral distance.
- minority**—a neighborhood analysis technique that outputs the least common value of the data file values in a user-specified window.

mode—the most commonly-occurring value in a set of data. In a histogram, the mode is the peak of the curve.

model—in a GIS, the set of expressions, or steps, that defines your criteria and creates an output layer.

modeling—the process of creating new layers from combining or operating upon existing layers. Modeling allows the creation of new classes from existing classes and the creation of a small set of images—perhaps even a single image—which, at a glance, contains many types of information about a scene.

modified projection—a map projection that is a modified version of another projection. For example, the Space Oblique Mercator projection is a modification of the Mercator projection.

monochrome image—an image produced from one band or layer, or contained in one color gun of the display device.

morphometric map—a map representing morphological features of the Earth's surface.

mosaicking—the process of piecing together images side by side, to create a larger image.

MrSID—see **Multiresolution Seamless Image Database**.

MSS—see **multispectral scanner**.

Multiresolution Seamless Image Database—(MrSID) a wavelet transform-based compression algorithm designed by LizardTech, Inc.

multispectral classification—the process of sorting pixels into a finite number of individual classes, or categories of data, based on data file values in multiple bands. See also **classification**.

multispectral imagery—satellite imagery with data recorded in two or more bands.

multispectral scanner—(MSS) Landsat satellite data acquired in four bands with a spatial resolution of 57 × 79 meters.

multitemporal—data from two or more different dates.

N

NAD27—see **North America Datum 1927**.

NAD83—see **North America Datum 1983**.

nadir—the area on the ground directly beneath a scanner's detectors.

nadir line—the average of the left and right edge lines of a Landsat image.

nadir point—the center of the nadir line in vertically viewed imagery.

NASA—see **National Aeronautics and Space Administration**.

National Aeronautics and Space Administration—(NASA) an organization that studies outer space. For more information, visit the NASA web site at <http://www.nasa.gov>.

- National Imagery and Mapping Agency**—(NIMA) formerly DMA. The agency was formed in October of 1996. NIMA supplies current imagery and geospatial data (NIMA 1998).
- National Imagery Transmission Format Standard**—(NITFS) a format designed to package imagery with complete annotation, text attachments, and imagery-associated metadata (Jordan and Beck 1999).
- National Ocean Service**—(NOS) the organization that created a zone numbering system for the State Plane coordinate system. A division of NOAA. For more information, visit the NOS web site at <http://www.nos.noaa.gov>.
- National Oceanic and Atmospheric Administration**—(NOAA) an organization that studies weather, water bodies, and encourages conservation. For more information, visit the NOAA web site at <http://www.noaa.gov>.
- Navigation System with Time and Ranging**—(NAVSTAR) satellite launched in 1978 for collection of GPS data.
- NAVSTAR**—see **Navigation System with Time and Ranging**.
- NDVI**—see **Normalized Difference Vegetation Index**.
- nearest neighbor**—a resampling method in which the output data file value is equal to the input pixel that has coordinates closest to the retransformed coordinates of the output pixel.
- neatline**—a rectangular border printed around a map. On scaled maps, neatlines usually have tick marks that indicate intervals of map coordinates or distance.
- negative inclination**—the sensors are tilted in increments of 0.6° to a maximum of 27° to the east.
- neighborhood analysis**—any image processing technique that takes surrounding pixels into consideration, such as convolution filtering and scanning.
- NIMA**—see **National Imagery and Mapping Agency**.
- 9-track**—CCTs that hold digital data.
- NITFS**—see **National Imagery Transmission Format Standard**.
- NOAA**—see **National Oceanic and Atmospheric Administration**.
- node**—the ending points of a line. See **from-node** and **to-node**.
- nominal data**—a type of data in which classes have no inherent order, and therefore are qualitative.
- nonlinear**—describing a function that cannot be expressed as the graph of a line or in the form of the equation of a line or plane. Nonlinear equations usually contain expressions with exponents. Second-order (2nd-order) or higher-order equations and transformations are nonlinear.
- nonlinear transformation**—a 2nd-order or higher rectification.

nonparametric signature—a signature for classification that is based on polygons or rectangles that are defined in the feature space image for the image file. There is no statistical basis for a nonparametric signature; it is simply an area in a feature space image.

normal—the state of having a normal distribution.

normal distribution—a symmetrical data distribution that can be expressed in terms of the mean and standard deviation of the data. The normal distribution is the most widely encountered model for probability, and is characterized by the bell curve. Also called Gaussian distribution.

normalize—a process that makes an image appear as if it were a flat surface. This technique is used to reduce topographic effect.

Normalized Difference Vegetation Index—(NDVI) the formula for NDVI is $(IR - R) / (IR + R)$, where IR stands for the infrared portion of the electromagnetic spectrum, and R stands for the red portion of the electromagnetic spectrum. NDVI finds areas of vegetation in imagery.

North America Datum 1927—(NAD27) a datum created in 1927 that is based on the Clarke 1866 spheroid. Commonly used in conjunction with the State Plane coordinate system.

North America Datum 1983—(NAD83) a datum created in 1983 that is based on the GRS 1980 spheroid. Commonly used in conjunction with the State Plane coordinate system.

NOS—see **National Ocean Service**.

NPO Mashinostroenia—a company based in Russia that develops satellites, such as Almaz 1-B, for GIS application.

number maps—maps that output actual data file values or brightness values, allowing the analysis of the values of every pixel in a file or on the display screen.

numeric keypad—the set of numeric and/or mathematical operator keys (+, -, etc.) that is usually on the right side of the keyboard.

Nyquist—in image registration, the original continuous function can be reconstructed from the sampled data, and phase function can be reconstructed to much higher resolution.

O

object—in models, an input to or output from a function. See **matrix object**, **raster object**, **scalar object**, **table object**.

oblique aspect—a map projection that is not oriented around a pole or the Equator.

observation—in photogrammetric triangulation, a grouping of the image coordinates for a GCP.

off-nadir—any point that is not directly beneath a scanner's detectors, but off to an angle. The SPOT scanner allows off-nadir viewing.

1:24,000—1:24,000 scale data, also called 7.5-minute DEM, available from USGS. It is usually referenced to the UTM coordinate system and has a spatial resolution of 30×30 meters.

1:250,000—1:250,000 scale DEM data available from USGS. Available only in arc/second format.

opacity—a measure of how opaque, or solid, a color is displayed in a raster layer.

operating system—(OS) the most basic means of communicating with the computer. It manages the storage of information in files and directories, input from devices such as the keyboard and mouse, and output to devices such as the monitor.

orbit—a circular, north-south and south-north path that a satellite travels above the Earth.

order—the complexity of a function, polynomial expression, or curve. In a polynomial expression, the order is simply the highest exponent used in the polynomial. See also **linear**, **nonlinear**.

ordinal data—a type of data that includes discrete lists of classes with an inherent order, such as classes of streams—first order, second order, third order, etc.

orientation angle—the angle between a perpendicular to the center scan line and the North direction in a satellite scene.

orthographic—an azimuthal projection with an infinite perspective.

orthorectification—see **orthorectification**.

orthoimage—see **digital orthophoto**.

orthomap—an image map product produced from orthoimages, or orthoimage mosaics, that is similar to a standard map in that it usually includes additional information, such as map coordinate grids, scale bars, north arrows, and other marginalia.

orthorectification—a form of rectification that corrects for terrain displacement and can be used if a DEM of the study area is available.

OS—see **operating system**.

outline map—a map showing the limits of a specific set of mapping entities such as counties. Outline maps usually contain a very small number of details over the desired boundaries with their descriptive codes.

overlay—1. a function that creates a composite file containing either the minimum or the maximum class values of the input files. Overlay sometimes refers generically to a combination of layers. 2. the process of displaying a classified file over the original image to inspect the classification.

overlay file—an ERDAS IMAGINE annotation file (.ovr extension).

.ovr file—an ERDAS IMAGINE annotation file.

panchromatic imagery—single-band or monochrome satellite imagery.

paneled map—a map designed to be spliced together into a large paper map.

Therefore, neatlines and tick marks appear on the outer edges of the large map.

pairwise mode—an operation mode in rectification that allows the registration of one image to an image in another Viewer, a map on a digitizing tablet, or coordinates entered at the keyboard.

parallax—displacement of a GCP appearing in a stereopair as a function of the position of the sensors at the time of image capture. You can adjust parallax in both the X and the Y direction so that the image point in both images appears in the same image space.

parallel—a line of latitude, going east and west.

parallelepiped—1. a classification decision rule in which the data file values of the candidate pixel are compared to upper and lower limits. 2. the limits of a parallelepiped classification, especially when graphed as rectangles.

parameter—1. any variable that determines the outcome of a function or operation. 2. the mean and standard deviation of data, which are sufficient to describe a normal curve.

parametric signature—a signature that is based on statistical parameters (e.g., mean and covariance matrix) of the pixels that are in the training sample or cluster.

passive sensors—solar imaging sensors that can only receive radiation waves and cannot transmit radiation.

path—the drive, directories, and subdirectories that specify the location of a file.

pattern recognition—the science and art of finding meaningful patterns in data, which can be extracted through classification.

PC—see **principal components**.

PCA—see **principal components analysis**.

perspective center—1. a point in the image coordinate system defined by the x and y coordinates of the principal point and the focal length of the sensor. 2. after triangulation, a point in the ground coordinate system that defines the sensor's position relative to the ground.

perspective projection—the projection of points by straight lines from a given perspective point to an intersection with the plane of projection.

phase—an element of an electromagnetic wave.

phase flattening—In IMAGINE IFSAR DEM, removes the phase function that would result if the imaging area was flat from the actual phase function recorded in the interferogram.

phase unwrapping—In IMAGINE IFSAR DEM, the process of taking a wrapped phase function and reconstructing the continuous function from it.

- photogrammetric quality scanners**—special devices capable of high image quality and excellent positional accuracy. Use of this type of scanner results in geometric accuracies similar to traditional analog and analytical photogrammetric instruments.
- photogrammetry**—the "art, science and technology of obtaining reliable information about physical objects and the environment through the process of recording, measuring and interpreting photographic images and patterns of electromagnetic radiant imagery and other phenomena" (ASP 1980).
- physical record**—a consecutive series of bytes on a 9-track tape followed by a gap, or blank space, on the tape.
- piecewise linear contrast stretch**—a spectral enhancement technique used to enhance a specific portion of data by dividing the lookup table into three sections: low, middle, and high.
- pixel**—abbreviated from picture element; the smallest part of a picture (image).
- pixel coordinate system**—a coordinate system with its origin in the upper left corner of the image, the x-axis pointing to the right, the y-axis pointing downward, and the units in pixels.
- pixel depth**—the number of bits required to store all of the data file values in a file. For example, data with a pixel depth of 8, or 8-bit data, have 256 values ($2^8 = 256$), ranging from 0 to 255.
- pixel size**—the physical dimension of a single light-sensitive element (13×13 microns).
- planar coordinates**—coordinates that are defined by a column and row position on a grid (x,y).
- planar projection**—see **azimuthal projection**.
- plane table photogrammetry**—prior to the invention of the airplane, photographs taken on the ground were used to extract the geometric relationships between objects using the principles of Descriptive Geometry.
- planimetric map**—a map that correctly represents horizontal distances between objects.
- plan symbol**—an annotation symbol that is formed after the basic outline of the object it represents. For example, the symbol for a house might be a square, since most houses are rectangular.
- point**—1. an element consisting of a single (x,y) coordinate pair. Also called grid cell.
2. a vertex of an element. Also called a node.
- point ID**—in rectification, a name given to GCPs in separate files that represent the same geographic location.
- point mode**—a digitizing mode in which one vertex is generated each time a keypad button is pressed.
- polar aspect**—a map projection that is centered around a pole.

- polarization**—the direction of the electric field component with the understanding that the magnetic field is perpendicular to it.
- polygon**—a set of closed line segments defining an area.
- polynomial**—a mathematical expression consisting of variables and coefficients. A coefficient is a constant that is multiplied by a variable in the expression.
- positive inclination**—the sensors are tilted in increments of 0.6° to a maximum of 27° to the west.
- primary colors**—colors from which all other available colors are derived. On a display monitor, the primary colors red, green, and blue are combined to produce all other colors. On a color printer, cyan, yellow, and magenta inks are combined.
- principal components**—(PC) the transects of a scatterplot of two or more bands of data that represent the widest variance and successively smaller amounts of variance that are not already represented. Principal components are orthogonal (perpendicular) to one another. In principal components analysis, the data are transformed so that the principal components become the axes of the scatterplot of the output data.
- principal components analysis**—(PCA) a method of data compression that allows redundant data to be compressed into fewer bands (Jensen 1996; Faust 1989).
- principal component band**—a band of data that is output by principal components analysis. Principal component bands are uncorrelated and nonredundant, since each principal component describes different variance within the original data.
- principal components analysis**—the process of calculating principal components and outputting principal component bands. It allows redundant data to be compacted into fewer bands (i.e., the dimensionality of the data is reduced).
- principal point (X_p , Y_p)**—the point in the image plane onto which the perspective center is projected, located directly beneath the interior orientation.
- printer**—a device that prints text, full color imagery, and/or graphics. See **color printer**, **text printer**.
- profile**—a row of data file values from a DEM or DTED file. The profiles of DEM and DTED run south to north (i.e., the first pixel of the record is the southernmost pixel).
- profile symbol**—an annotation symbol that is formed like the profile of an object. Profile symbols generally represent vertical objects such as trees, windmills, oil wells, etc.
- proximity analysis**—a technique used to determine which pixels of a thematic layer are located at specified distances from pixels in a class or classes. A new layer is created that is classified by the distance of each pixel from specified classes of the input layer.

pseudo color—a method of displaying an image (usually a thematic layer) that allows the classes to have distinct colors. The class values of the single band file are translated through all three function memories that store a color scheme for the image.

pseudo node—a single line that connects with itself (an island), or where only two lines intersect.

pseudo projection—a map projection that has only some of the characteristics of another projection.

pushbroom—a scanner in which all scanning parts are fixed, and scanning is accomplished by the forward motion of the scanner, such as the SPOT scanner.

pyramid layers—image layers which are successively reduced by the power of 2 and resampled. Pyramid layers enable large images to display faster.

Q

quadrangle—1. any of the hardcopy maps distributed by USGS such as the 7.5-minute quadrangle or the 15-minute quadrangle. 2. one quarter of a full Landsat TM scene. Commonly called a quad.

qualitative map—a map that shows the spatial distribution or location of a kind of nominal data. For example, a map showing corn fields in the US would be a qualitative map. It would not show how much corn is produced in each location, or production relative to other areas.

quantitative map—a map that displays the spatial aspects of numerical data. A map showing corn production (volume) in each area would be a quantitative map.

R

radar data—the remotely sensed data that are produced when a radar transmitter emits a beam of micro or millimeter waves, the waves reflect from the surfaces they strike, and the backscattered radiation is detected by the radar system's receiving antenna, which is tuned to the frequency of the transmitted waves.

RADARSAT—a Canadian radar satellite.

radiative transfer equations—the mathematical models that attempt to quantify the total atmospheric effect of solar illumination.

radiometric correction—the correction of variations in data that are not caused by the object or scene being scanned, such as scanner malfunction and atmospheric interference.

radiometric enhancement—an enhancement technique that deals with the individual values of pixels in an image.

radiometric resolution—the dynamic range, or number of possible data file values, in each band. This is referred to by the number of bits into which the recorded energy is divided. See **pixel depth**.

RAM—see **random-access memory**.

random-access memory—(RAM) memory used for applications and data storage on a CPU (Free On-Line Dictionary of Computing 1999).

rank—a neighborhood analysis technique that outputs the number of values in a user-specified window that are less than the analyzed value.

RAR—see **Real-Aperture Radar**.

raster data—data that are organized in a grid of columns and rows. Raster data usually represent a planar graph or geographical area. Raster data in ERDAS IMAGINE are stored in image files.

raster object—in Model Maker (Spatial Modeler), a single raster layer or set of layers.

Raster Product Format—(RPF) Data from NIMA, used primarily for military purposes. Organized in 1536×1536 frames, with an internal tile size of 256×256 pixels.

raster region—a contiguous group of pixels in one GIS class. Also called clump.

ratio data—a data type in which thematic class values have the same properties as interval values, except that ratio values have a natural zero or starting point.

RDBMS—see **relational database management system**.

RDGPS—see **Real Time Differential GPS**.

Real-Aperture Radar—(RAR) a radar sensor that uses its side-looking, fixed antenna to transmit and receive the radar impulse. For a given position in space, the resolution of the resultant image is a function of the antenna size. The signal is processed independently of subsequent return signals.

Real Time Differential GPS—(RDGPS) takes the Differential Correction technique one step further by having the base station communicate the error vector via radio to the field unit in real time.

recoding—the assignment of new values to one or more classes.

record—1. the set of all attribute data for one class of feature. 2. the basic storage unit on a 9-track tape.

rectification—the process of making image data conform to a map projection system. In many cases, the image must also be oriented so that the north direction corresponds to the top of the image.

rectified coordinates—the coordinates of a pixel in a file that has been rectified, which are extrapolated from the GCPs. Ideally, the rectified coordinates for the GCPs are exactly equal to the reference coordinates. Because there is often some error tolerated in the rectification, this is not always the case.

reduce—the process of skipping file pixels when displaying an image so that a larger area can be represented on the display screen. For example, a reduction factor of 3 would cause only the pixel at every third row and column to be displayed, so that each displayed pixel represents a 3×3 block of file pixels.

- reference coordinates**—the coordinates of the map or reference image to which a source (input) image is being registered. GCPs consist of both input coordinates and reference coordinates for each point.
- reference pixels**—in classification accuracy assessment, pixels for which the correct GIS class is known from ground truth or other data. The reference pixels can be selected by you, or randomly selected.
- reference plane**—In a topocentric coordinate system, the tangential plane at the center of the image on the Earth ellipsoid, on which the three perpendicular coordinate axes are defined.
- reference system**—the map coordinate system to which an image is registered.
- reference window**—the source window on the first image of an image pair, which remains at a constant location. See also **correlation windows** and **search windows**.
- reflection spectra**—the electromagnetic radiation wavelengths that are reflected by specific materials of interest.
- registration**—the process of making image data conform to another image. A map coordinate system is not necessarily involved.
- regular block of photos**—a rectangular block in which the number of photos in each strip is the same; this includes a single strip or a single stereopair.
- relational database management system**—(RDBMS) system that stores SDE database layers.
- relation based matching**—an image matching technique that uses the image features and the relation among the features to automatically recognize the corresponding image structures without any *a priori* information.
- relief map**—a map that appears to be or is three-dimensional.
- remote sensing**—the measurement or acquisition of data about an object or scene by a satellite or other instrument above or far from the object. Aerial photography, satellite imagery, and radar are all forms of remote sensing.
- replicative symbol**—an annotation symbol that is designed to look like its real-world counterpart. These symbols are often used to represent trees, railroads, houses, etc.
- representative fraction**—the ratio or fraction used to denote map scale.
- resampling**—the process of extrapolating data file values for the pixels in a new grid when data have been rectified or registered to another image.
- rescaling**—the process of compressing data from one format to another. In ERDAS IMAGINE, this typically means compressing a 16-bit file to an 8-bit file.
- reshape**—the process of redigitizing a portion of a line.

residuals—in rectification, the distances between the source and retransformed coordinates in one direction. In ERDAS IMAGINE, they are shown for each GCP. The X residual is the distance between the source X coordinate and the retransformed X coordinate. The Y residual is the distance between the source Y coordinate and the retransformed Y coordinate.

resolution—a level of precision in data. For specific types of resolution see **display resolution**, **radiometric resolution**, **spatial resolution**, **spectral resolution**, and **temporal resolution**.

resolution merging—the process of sharpening a lower-resolution multiband image by merging it with a higher-resolution monochrome image.

retransformed—in the rectification process, a coordinate in the reference (output) coordinate system that has transformed back into the input coordinate system. The amount of error in the transformation can be determined by computing the difference between the original coordinates and the retransformed coordinates. See **RMS error**.

RGB—red, green, blue. The primary additive colors that are used on most display hardware to display imagery.

RGB clustering—a clustering method for 24-bit data (three 8-bit bands) that plots pixels in three-dimensional spectral space, and divides that space into sections that are used to define clusters. The output color scheme of an RGB-clustered image resembles that of the input file.

rhumb line—a line of true direction that crosses meridians at a constant angle.

right-hand rule—a convention in three-dimensional coordinate systems (X,Y,Z) that determines the location of the positive Z axis. If you place your right hand fingers on the positive X axis and curl your fingers toward the positive Y axis, the direction your thumb is pointing is the positive Z axis direction.

RMS error—the distance between the input (source) location of a GCP and the retransformed location for the same GCP. RMS error is calculated with a distance equation.

RMSE—(Root Mean Square Error) used to measure how well a specific calculated solution fits the original data. For each observation of a phenomena, a variation can be computed between the actual observation and a calculated value. (The method of obtaining a calculated value is application-specific.) Each variation is then squared. The sum of these squared values is divided by the number of observations and then the square root is taken. This is the RMSE value.

roam—the process of moving across a display so that different areas of the image appear on the display screen.

root—the first part of a file name, which usually identifies the file's specific contents.

ROYGBIV—a color scheme ranging through red, orange, yellow, green, blue, indigo, and violet at regular intervals.

RPF—see **Raster Product Format**.

rubber sheeting—the application of a nonlinear rectification (2nd-order or higher).

S

sample—see **training sample**.

SAR—see **Synthetic Aperture Radar**.

saturation—a component of IHS which represents the purity of color and also varies linearly from 0 to 1.

SCA—see **suitability/capability analysis**.

scale—1. the ratio of distance on a map as related to the true distance on the ground.
2. cell size. 3. the processing of values through a lookup table.

scale bar—a graphic annotation element that describes map scale. It shows the distance on paper that represents a geographical distance on the map.

scalar object—in Model Maker (Spatial Modeler), a single numeric value.

scaled map—a georeferenced map that is accurately arranged and referenced to represent distances and locations. A scaled map usually has a legend that includes a scale, such as 1 inch = 1000 feet. The scale is often expressed as a ratio like 1:12,000 where 1 inch on the map equals 12,000 inches on the ground.

scanner—the entire data acquisition system, such as the Landsat TM scanner or the SPOT panchromatic scanner.

scanning—1. the transfer of analog data, such as photographs, maps, or another viewable image, into a digital (raster) format. 2. a process similar to convolution filtering that uses a kernel for specialized neighborhood analyses, such as total, average, minimum, maximum, boundary, and majority.

scatterplot—a graph, usually in two dimensions, in which the data file values of one band are plotted against the data file values of another band.

scene—the image captured by a satellite.

screen coordinates—the location of a pixel on the display screen, beginning with 0,0 in the upper left corner.

screen digitizing—the process of drawing vector graphics on the display screen with a mouse. A displayed image can be used as a reference.

script modeling—the technique of combining data layers in an unlimited number of ways. Script modeling offers all of the capabilities of graphical modeling with the ability to perform more complex functions, such as conditional looping.

script model—a model that is comprised of text only and is created with the SML. Script models are stored in .mdl files.

SCS—see **Soil Conservation Service**.

SD—see **standard deviation**.

SDE—see **Spatial Database Engine**.

SDTS—see **spatial data transfer standard**.

SDTS Raster Profile and Extensions—(SRPE) an SDTS profile that covers gridded raster data.

search radius—in surfacing routines, the distance around each pixel within which the software searches for terrain data points.

search windows—candidate windows on the second image of an image pair that are evaluated relative to the reference window.

seat—a combination of an X-server and a host workstation.

Sea-viewing Wide Field-of-View Sensor—(SeaWiFS) a sensor located on many different satellites such as ORBVIEW's OrbView-2, and NASA's SeaStar.

SeaWiFS—see **Sea-viewing Wide Field-of-View Sensor**.

secant—the intersection of two points or lines. In the case of conic or cylindrical map projections, a secant cone or cylinder intersects the surface of a globe at two circles.

Selective Availability—introduces a positional inaccuracy of up to 100 m to commercial GPS receivers.

sensor—a device that gathers energy, converts it to a digital value, and presents it in a form suitable for obtaining information about the environment.

separability—a statistical measure of distance between two signatures.

separability listing—a report of signature divergence which lists the computed divergence for every class pair and one band combination. The listing contains every divergence value for the bands studied for every possible pair of signatures.

sequential clustering—a method of clustering that analyzes pixels of an image line by line and groups them by spectral distance. Clusters are determined based on relative spectral distance and the number of pixels per cluster.

server—on a computer in a network, a utility that makes some resource or service available to the other machines on the network (such as access to a tape drive).

shaded relief image—a thematic raster image that shows variations in elevation based on a user-specified position of the sun. Areas that would be in sunlight are highlighted and areas that would be in shadow are shaded.

shaded relief map—a map of variations in elevation based on a user-specified position of the sun. Areas that would be in sunlight are highlighted and areas that would be in shadow are shaded.

shapefile—an ESRI vector format that contains spatial data. Shapefiles have the .shp extension.

short wave infrared region—(SWIR) the near-infrared and middle-infrared regions of the electromagnetic spectrum.

SI—see **image scale**.

Side-looking Airborne Radar—(SLAR) a radar sensor that uses an antenna which is fixed below an aircraft and pointed to the side to transmit and receive the radar signal.

signal based matching—see **area based matching**.

Signal-to-Noise ratio—(S/N) in hyperspectral image processing, a ratio used to evaluate the usefulness or validity of a particular band of data.

signature—a set of statistics that defines a training sample or cluster. The signature is used in a classification process. Each signature corresponds to a GIS class that is created from the signatures with a classification decision rule.

skew—a condition in satellite data, caused by the rotation of the Earth eastward, which causes the position of the satellite relative to the Earth to move westward. Therefore, each line of data represents terrain that is slightly west of the data in the previous line.

SLAR—see **Side-looking Airborne Radar**.

slope—the change in elevation over a certain distance. Slope can be reported as a percentage or in degrees.

slope image—a thematic raster image that shows changes in elevation over distance. Slope images are usually color coded to show the steepness of the terrain at each pixel.

slope map—a map that is color coded to show changes in elevation over distance.

small-scale—for a map or data file, having a small ratio between the area of the imagery (such as inches or pixels) and the area that is represented (such as feet). In small-scale image data, each pixel represents a large area on the ground, such as NOAA AVHRR data, with a spatial resolution of 1.1 km.

SML—see **Spatial Modeler Language**.

S/N—see **Signal-to-Noise ratio**.

softcopy photogrammetry—see **digital photogrammetry**.

Soil Conservation Service—(SCS) an organization that produces soil maps (Fisher 1991) with guidelines provided by the USDA.

SOM—see **Space Oblique Mercator**.

source coordinates—in the rectification process, the input coordinates.

Spaceborne Imaging Radar—(SIR-A, SIR-B, and SIR-C) the radar sensors that fly aboard NASA space shuttles. SIR-A flew aboard the 1981 NASA Space Shuttle Columbia. That data and SIR-B data from a later Space Shuttle mission are still valuable sources of radar data. The SIR-C sensor was launched in 1994.

Space Oblique Mercator—(SOM) a projection available in ERDAS IMAGINE that is nearly conformal and has little scale distortion within the sensing range of an orbiting mapping satellite such as Landsat.

spatial data transfer standard—(SDTS) “a robust way of transferring Earth-referenced spatial data between dissimilar computer systems with the potential for no information loss” (USGS 1999).

Spatial Database Engine—(SDE) An ESRI vector format that manages a database theme. SDE allows you to access databases that may contain large amounts of information (ESRI 1996).

spatial enhancement—the process of modifying the values of pixels in an image relative to the pixels that surround them.

spatial frequency—the difference between the highest and lowest values of a contiguous set of pixels.

Spatial Modeler Language—(SML) a script language used internally by Model Maker (Spatial Modeler) to execute the operations specified in the graphical models you create. SML can also be used to write application-specific models.

spatial resolution—a measure of the smallest object that can be resolved by the sensor, or the area on the ground represented by each pixel.

speckle noise—the light and dark pixel noise that appears in radar data.

spectral distance—the distance in spectral space computed as Euclidean distance in n-dimensions, where n is the number of bands.

spectral enhancement—the process of modifying the pixels of an image based on the original values of each pixel, independent of the values of surrounding pixels.

spectral resolution—the specific wavelength intervals in the electromagnetic spectrum that a sensor can record.

spectral space—an abstract space that is defined by spectral units (such as an amount of electromagnetic radiation). The notion of spectral space is used to describe enhancement and classification techniques that compute the spectral distance between n-dimensional vectors, where n is the number of bands in the data.

spectroscopy—the study of the absorption and reflection of electromagnetic radiation (EMR) waves.

spliced map—a map that is printed on separate pages, but intended to be joined together into one large map. Neatlines and tick marks appear only on the pages which make up the outer edges of the whole map.

spline—the process of smoothing or generalizing all currently selected lines using a specified grain tolerance during vector editing.

split—the process of making two lines from one by adding a node.

SPOT—a series of Earth-orbiting satellites operated by the Centre National d’Etudes Spatiales (CNES) of France.

SRPE—see **SDTS Raster Profile and Extensions**.

STA—see **statistics file**.

- standard deviation**—(SD) 1. the square root of the variance of a set of values which is used as a measurement of the spread of the values. 2. a neighborhood analysis technique that outputs the standard deviation of the data file values of a user-specified window.
- standard meridian**—see **standard parallel**.
- standard parallel**—the line of latitude where the surface of a globe conceptually intersects with the surface of the projection cylinder or cone.
- statement**—in script models, properly formatted lines that perform a specific task in a model. Statements fall into the following categories: declaration, assignment, show, view, set, macro definition, and quit.
- statistical clustering**—a clustering method that tests 3×3 sets of pixels for homogeneity, and builds clusters only from the statistics of the homogeneous sets of pixels.
- statistics file**—(STA) an ERDAS IMAGINE Ver. 7.X trailer file for LAN data that contains statistics about the data.
- stereographic**—1. the process of projecting onto a tangent plane from the opposite side of the Earth. 2. the process of acquiring images at angles on either side of the vertical.
- stereopair**—a set of two remotely-sensed images that overlap, providing two views of the terrain in the overlap area.
- stereo-scene**—achieved when two images of the same area are acquired on different days from different orbits, one taken east of the vertical, and the other taken west of the nadir.
- stream mode**—a digitizing mode in which vertices are generated continuously while the digitizer keypad is in proximity to the surface of the digitizing tablet.
- string**—a line of text. A string usually has a fixed length (number of characters).
- strip of photographs**—consists of images captured along a flight-line, normally with an overlap of 60% for stereo coverage. All photos in the strip are assumed to be taken at approximately the same flying height and with a constant distance between exposure stations. Camera tilt relative to the vertical is assumed to be minimal.
- striping**—a data error that occurs if a detector on a scanning system goes out of adjustment—that is, it provides readings consistently greater than or less than the other detectors for the same band over the same ground cover. Also called banding.
- structure based matching**—see **relation based matching**.
- subsetting**—the process of breaking out a portion of a large image file into one or more smaller files.
- suitability/capability analysis**—(SCA) a system designed to analyze many data layers to produce a plan map. Discussed in McHarg's book *Design with Nature* (Star and Estes 1990).

sum—a neighborhood analysis technique that outputs the total of the data file values in a user-specified window.

Sun raster data—imagery captured from a Sun monitor display.

sun-synchronous—a term used to describe Earth-orbiting satellites that rotate around the Earth at the same rate as the Earth rotates on its axis.

supervised training—any method of generating signatures for classification, in which the analyst is directly involved in the pattern recognition process. Usually, supervised training requires the analyst to select training samples from the data that represent patterns to be classified.

surface—a one-band file in which the value of each pixel is a specific elevation value.

swath width—in a satellite system, the total width of the area on the ground covered by the scanner.

SWIR—see **short wave infrared region**.

symbol—an annotation element that consists of other elements (sub-elements). See **plan symbol**, **profile symbol**, and **function symbol**.

symbolization—a method of displaying vector data in which attribute information is used to determine how features are rendered. For example, points indicating cities and towns can appear differently based on the population field stored in the attribute database for each of those areas.

Synthetic Aperture Radar—(SAR) a radar sensor that uses its side-looking, fixed antenna to create a synthetic aperture. SAR sensors are mounted on satellites, aircraft, and the NASA Space Shuttle. The sensor transmits and receives as it is moving. The signals received over a time interval are combined to create the image.

T

table object—in Model Maker (Spatial Modeler), a series of numeric values or character strings.

tablet digitizing—the process of using a digitizing tablet to transfer nondigital data such as maps or photographs to vector format.

Tagged Imaged File Format—see **TIFF data**.

tangent—an intersection at one point or line. In the case of conic or cylindrical map projections, a tangent cone or cylinder intersects the surface of a globe in a circle.

Tasseled Cap transformation—an image enhancement technique that optimizes data viewing for vegetation studies.

TEM—see **transverse electromagnetic wave**.

temporal resolution—the frequency with which a sensor obtains imagery of a particular area.

terrain analysis—the processing and graphic simulation of elevation data.

terrain data—elevation data expressed as a series of x, y, and z values that are either regularly or irregularly spaced.

text printer—a device used to print characters onto paper, usually used for lists, documents, and reports. If a color printer is not necessary or is unavailable, images can be printed using a text printer. Also called a line printer.

thematic data—raster data that are qualitative and categorical. Thematic layers often contain classes of related information, such as land cover, soil type, slope, etc. In ERDAS IMAGINE, thematic data are stored in image files.

thematic layer—see **thematic data**.

thematic map—a map illustrating the class characterizations of a particular spatial variable such as soils, land cover, hydrology, etc.

Thematic Mapper—(TM) Landsat data acquired in seven bands with a spatial resolution of 30 × 30 meters.

thematic mapper simulator—(TMS) an instrument “designed to simulate spectral, spatial, and radiometric characteristics of the Thematic Mapper sensor on the Landsat-4 and 5 spacecraft” (NASA 1995).

theme—a particular type of information, such as soil type or land use, that is represented in a layer.

3D perspective view—a simulated three-dimensional view of terrain.

threshold—a limit, or cutoff point, usually a maximum allowable amount of error in an analysis. In classification, thresholding is the process of identifying a maximum distance between a pixel and the mean of the signature to which it was classified.

tick marks—small lines along the edge of the image area or neatline that indicate regular intervals of distance.

tie point—a point; its ground coordinates are not known, but can be recognized visually in the overlap or sidelap area between two images.

TIFF data—Tagged Image File Format data is a raster file format developed by Aldus, Corp. (Seattle, Washington), in 1986 for the easy transportation of data.

TIGER—see **Topologically Integrated Geographic Encoding and Referencing System**.

tiled data—the storage format of ERDAS IMAGINE image files.

TIN—see **triangulated irregular network**.

TM—see **Thematic Mapper**.

TMS—see **thematic mapper simulator**.

TNDVI—see **Transformed Normalized Distribution Vegetative Index**.

to-node—the last vertex in a line.

topocentric coordinate system—a coordinate system that has its origin at the center of the image on the Earth ellipsoid. The three perpendicular coordinate axes are defined on a tangential plane at this center point. The x-axis is oriented eastward, the y-axis northward, and the z-axis is vertical to the reference plane (up).

topographic—a term indicating elevation.

topographic data—a type of raster data in which pixel values represent elevation.

topographic effect—a distortion found in imagery from mountainous regions that results from the differences in illumination due to the angle of the sun and the angle of the terrain.

topographic map—a map depicting terrain relief.

Topologically Integrated Geographic Encoding and Referencing System—(TIGER) files are line network products of the US Census Bureau.

Topological Vector Profile—(TVP) a profile of SDTS that covers attributed vector data.

topology—a term that defines the spatial relationships between features in a vector layer.

total RMS error—the total root mean square (RMS) error for an entire image. Total RMS error takes into account the RMS error of each GCP.

trailer file—1. an ERDAS IMAGINE Ver. 7.X file with a .TRL extension that accompanies a GIS file and contains information about the GIS classes. 2. a file following the image data on a 9-track tape.

training—the process of defining the criteria by which patterns in image data are recognized for the purpose of classification.

training field—the geographical area represented by the pixels in a training sample. Usually, it is previously identified with the use of ground truth data or aerial photography. Also called training site.

training sample—a set of pixels selected to represent a potential class. Also called sample.

transformation matrix—a set of coefficients that is computed from GCPs, and used in polynomial equations to convert coordinates from one system to another. The size of the matrix depends upon the order of the transformation.

Transformed Normalized Distribution Vegetative Index—(TNDVI) adds 0.5 to the NDVI equation, then takes the square root. Created by Deering *et al* in 1975 (Jensen 1996).

transposition—the interchanging of the rows and columns of a matrix, denoted with T .

transverse aspect—the orientation of a map in which the central line of the projection, which is normally the equator, is rotated 90 degrees so that it follows a meridian.

transverse electromagnetic wave—(TEM) a wave where both E (electric field) and H (magnetic field) are transverse to the direction of propagation.

triangulated irregular network—(TIN) a specific representation of DTMs in which elevation points can occur at irregular intervals.

triangulation—establishes the geometry of the camera or sensor relative to objects on the Earth's surface.

true color—a method of displaying an image (usually from a continuous raster layer) that retains the relationships between data file values and represents multiple bands with separate color guns. The image memory values from each displayed band are translated through the function memory of the corresponding color gun.

true direction—the property of a map projection to represent the direction between two points with a straight rhumb line, which crosses meridians at a constant angle.

TVP—see **Topological Vector Profile**.

U

union—the area or set that is the combination of two or more input areas or sets without repetition.

United States Department of Agriculture—(USDA) an organization regulating the agriculture of the US. For more information, visit the web site www.usda.gov.

United States Geological Survey—(USGS) an organization dealing with biology, geology, mapping, and water. For more information, visit the web site www.usgs.gov.

Universal Polar Stereographic—(UPS) a mapping system used in conjunction with the Polar Stereographic projection that makes the scale factor at the pole 0.994.

Universal Transverse Mercator—(UTM) UTM is an international plane (rectangular) coordinate system developed by the US Army that extends around the world from 84°N to 80°S. The world is divided into 60 zones each covering six degrees longitude. Each zone extends three degrees eastward and three degrees westward from its central meridian. Zones are numbered consecutively west to east from the 180° meridian.

unscaled map—a hardcopy map that is not referenced to any particular scale in which one file pixel is equal to one printed pixel.

unsplit—the process of joining two lines by removing a node.

unsupervised training—a computer-automated method of pattern recognition in which some parameters are specified by the user and are used to uncover statistical patterns that are inherent in the data.

UPS—see **Universal Polar Stereographic**.

USDA—see **United States Department of Agriculture**.

USGS—see **United States Geological Survey**.

UTM—see **Universal Transverse Mercator**.

V

variable—1. a numeric value that is changeable, usually represented with a letter. 2. a thematic layer. 3. one band of a multiband image. 4. in models, objects which have been associated with a name using a declaration statement.

variable rate technology—(VRT) in precision agriculture, used with GPS data. VRT relies on the use of a VRT controller box connected to a GPS and the pumping mechanism for a tank full of fertilizers/pesticides/seeds/water/etc.

variance—the measure of central tendency.

vector—1. a line element. 2. a one-dimensional matrix, having either one row (1 by j), or one column (i by 1). See also **mean vector**, **measurement vector**.

vector data—data that represent physical forms (elements) such as points, lines, and polygons. Only the vertices of vector data are stored, instead of every point that makes up the element. ERDAS IMAGINE vector data are based on the ArcInfo data model and are stored in directories, rather than individual files. See **workspace**.

vector layer—a set of vector features and their associated attributes.

Vector Quantization—(VQ) used to compress frames of RPF data.

velocity vector—the satellite's velocity if measured as a vector through a point on the spheroid.

verbal statement—a statement that describes the distance on the map to the distance on the ground. A verbal statement describing a scale of 1:1,000,000 is approximately 1 inch to 16 miles. The units on the map and on the ground do not have to be the same in a verbal statement.

vertex—a point that defines an element, such as a point where a line changes direction.

vertical control—the vertical distribution of GCPs in aerial triangulation (z - elevation).

vertices—plural of vertex.

viewshed analysis—the calculation of all areas that can be seen from a particular viewing point or path.

viewshed map—a map showing only those areas visible (or invisible) from a specified point(s).

VIS/IR—see **visible/infrared imagery**.

visible/infrared imagery—(VIS/IR) a type of multispectral data set that is based on the reflectance spectrum of the material of interest.

volume—a medium for data storage, such as a magnetic disk or a tape.

volume set—the complete set of tapes that contains one image.

VPF—see **vector product format**.

VQ—see **Vector Quantization**.

VRT—see **variable rate technology**.

-
- W**
- wavelet**—"a waveform that is bounded in both frequency and duration" (Free On-Line Dictionary of Computing 1999).
- weight**—the number of values in a set; particularly, in clustering algorithms, the weight of a cluster is the number of pixels that have been averaged into it.
- weighting factor**—a parameter that increases the importance of an input variable. For example, in GIS indexing, one input layer can be assigned a weighting factor that multiplies the class values in that layer by that factor, causing that layer to have more importance in the output file.
- weighting function**—in surfacing routines, a function applied to elevation values for determining new output values.
- WGS**—see **World Geodetic System**.
- Wide Field Sensor**—(WiFS) sensor aboard IRS-1C with 188m spatial resolution.
- WiFS**—see **Wide Field Sensor**.
- working window**—the image area to be used in a model. This can be set to either the union or intersection of the input layers.
- workspace**—a location that contains one or more vector layers. A workspace is made up of several directories.
- World Geodetic System**—(WGS) a spheroid. Earth ellipsoid with multiple versions including: WGS 66, 72, and 84.
- write ring**—a protection device that allows data to be written to a 9-track tape when the ring is in place, but not when it is removed.

-
- X**
- X residual**—in RMS error reports, the distance between the source X coordinate and the retransformed X coordinate.
- X RMS error**—the root mean square (RMS) error in the X direction.

-
- Y**
- Y residual**—in RMS error reports, the distance between the source Y coordinate and the retransformed Y coordinate.
- Y RMS error**—the root mean square (RMS) error in the Y direction.

-
- Z**
- ZDR**—see **zone distribution rectangles**.
- zero-sum kernel**—a convolution kernel in which the sum of all the coefficients is zero. Zero-sum kernels are usually edge detectors.
- zone distribution rectangles**—(ZDRs) the images into which each distribution DR are divided in ADRG data.

Glossary

Z

zoom—the process of expanding displayed pixels on an image so they can be more closely studied. Zooming is similar to magnification, except that it changes the display only temporarily, leaving image memory the same.

Bibliography

- Ackermann, F., 1983. "High precision digital image correlation." *Proceedings of 39th Photogrammetric Week*, Institute of Photogrammetry, University of Stuttgart, pp.231-243.
- Adams, J.B., Smith, M.O., and Gillespie, A.R. 1989. "Simple Models for Complex Natural Surfaces: A Strategy for the Hyperspectral Era of Remote Sensing." *Proceedings IEEE Intl. Geosciences and Remote Sensing Symposium*. 1:16-21.
- Agouris, P. and Schenk, T., 1996. "Automated Aerotriangulation Using Multiple Image Multipoint Matching." *Photogrammetric Engineering and Remote Sensing*, 62(6): 703-710.
- Akima, H., 1978, A Method for Bivariate Interpolation and Smooth Surface Fitting for Irregularly Distributed Data Points, *ACM Transactions on Mathematical Software* 4(2), pp. 148-159.
- American Society of Photogrammetry. 1980. *Photogrammetric Engineering and Remote Sensing* XLVI:10:1249.
- Atkinson, Paula. 1985. "Preliminary Results of the Effect of Resampling on Thematic Mapper Imagery." *1985 ACSM-ASPRS Fall Convention Technical Papers*. Falls Church, Virginia: American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping.
- Atlantis Scientific, Inc. "Sources of SAR Data: Sources of SAR Data." at http://www.atlsci.com/library/sar_sources.html , 23 March 1997.
- Battrick, Bruce, and Lois Proud, eds. May 1992. *ERS-1 User Handbook*. Noordwijk, The Netherlands: European Space Agency, ESA Publications Division, c/o ESTEC.
- Bauer, H., and Müller, J., 1972. "Height accuracy of blocks and bundle block adjustment with additional parameters." *ISPRS 12th Congress*, Ottawa.
- Benediktsson, J.A., Swain, P.H., Ersoy, O.K., and Hong, D. 1990. "Neural Network Approaches Versus Statistical Methods in Classification of Multisource Remote Sensing Data." *IEEE Transactions on Geoscience and Remote Sensing* 28:4:540-51.
- Berk, A., et al. 1989. *MODTRAN: A Moderate Resolution Model for LOWTRAN 7*. Hanscom Air Force Base, Massachusetts: U.S. Air Force Geophysical Laboratory (AFGL).
- Bernstein, Ralph, et al. 1983. "Image Geometry and Rectification." Chapter 21 in *Manual of Remote Sensing*, edited by Robert N. Colwell. Falls Church, Virginia: American Society of Photogrammetry.

Bibliography

- Billingsley, Fred C., et al. 1983. "Data Processing and Reprocessing." Chapter 17 in *Manual of Remote Sensing*, edited by Robert N. Colwell. Falls Church, Virginia: American Society of Photogrammetry.
- Blom, Ronald G., and Michael Daily. July 1982. "Radar Image Processing for Rock-Type Discrimination." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-20, No. 3.
- Buchanan, M. D. 1979. "Effective Utilization of Color in Multidimensional Data Presentation." *Proceedings of the Society of Photo-Optical Engineers*, Vol. 199: 9-19.
- Burrus, C. S., and T. W. Parks. 1985. *DFT/FFT and Convolution Algorithms: Theory and Implementation*. New York: John Wiley & Sons, Inc.
- Cannon, Michael, Alex Lehar, and Fred Preston, 1983. "Background Pattern Removal by Power Spectral Filtering." *Applied Optics*, Vol. 22, No. 6: 777-779.
- Carter, James R. 1989. "On Defining the Geographic Information System." *Fundamentals of Geographic Information Systems: A Compendium*, edited by William J. Ripple. Bethesda, Maryland: American Society for Photogrammetric Engineering and Remote Sensing and the American Congress on Surveying and Mapping.
- Chahine, Moustafa T., et al. 1983. "Interaction Mechanisms within the Atmosphere." Chapter 5 in *Manual of Remote Sensing*, edited by Robert N. Colwell. Falls Church, Virginia: American Society of Photogrammetry.
- Chavez, Pat S., Jr., et al. 1991. "Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data: Landsat TM and SPOT Panchromatic." *Photogrammetric Engineering & Remote Sensing*, Vol. 57, No. 3: 295-303.
- Chavez, Pat S., Jr., and Graydon L. Berlin. 1986. "Restoration Techniques for SIR-B Digital Radar Images." Paper presented at the Fifth Thematic Conference: Remote Sensing for Exploration Geology, Reno, Nevada.
- Clark, Roger N., and Ted L. Roush. 1984. "Reflectance Spectroscopy: Quantitative Analysis Techniques for Remote Sensing Applications." *Journal of Geophysical Research*, Vol. 89, No. B7: 6329-6340.
- Clark, R.N., Gallagher, A.J., and Swayze, G.A. 1990. "Material Absorption Band Depth Mapping of Imagine Spectrometer Data using a Complete Band Shape Least-Square Fit with Library Reference Spectra." *Proceedings of the Second AVIRIS Conference*. JPL Pub. 90-54.
- CNES. "CNES: CNES: Centre National D'Etudes Spatiales" at <http://sads.cnes.fr/ceos/cdrom-98/ceos1/cnes/gb/lecnes.htm> , 03 December 1998.
- Colby, J. D. 1991. "Topographic Normalization in Rugged Terrain." *Photogrammetric Engineering & Remote Sensing*, Vol. 57, No. 5: 531-537.

- Colwell, Robert N., ed. 1983. *Manual of Remote Sensing*. Falls Church, Virginia: American Society of Photogrammetry.
- Congalton, R. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment*, Vol. 37: 35-46.
- Conrac Corp., Conrac Division. 1980. *Raster Graphics Handbook*. Covina, California: Conrac Corp.
- Crippen, Robert E. July 1989. "Development of Remote Sensing Techniques for the Investigation of Neotectonic Activity, Eastern Transverse Ranges and Vicinity, Southern California." Ph.D. Diss., University of California, Santa Barbara.
- Crippen, Robert E. 1989. "A Simple Spatial Filtering Routine for the Cosmetic Removal of Scan-Line Noise from Landsat TM P-Tape Imagery." *Photogrammetric Engineering & Remote Sensing*, Vol. 55, No. 3: 327-331.
- Crippen, Robert E. 1987. "The Regression Intersection Method of Adjusting Image Data for Band Ratioing." *International Journal of Remote Sensing*, Vol. 8, No. 2: 137-155.
- Crist, E. P., et al. 1986. "Vegetation and Soils Information Contained in Transformed Thematic Mapper Data." *Proceedings of IGARSS' 86 Symposium*, ESA Publications Division, ESA SP-254.
- Crist, E. P., and R. J. Kauth. 1986. "The Tasseled Cap De-Mystified." *Photogrammetric Engineering & Remote Sensing*, Vol. 52, No. 1: 81-86.
- Cullen, Charles G. 1972. *Matrices and Linear Transformations*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Daily, Mike. 1983. "Hue-Saturation-Intensity Split-Spectrum Processing of Seasat Radar Imagery." *Photogrammetric Engineering & Remote Sensing*, Vol. 49, No. 3: 349-355.
- Dangermond, Jack. 1988. "A Review of Digital Data Commonly Available and Some of the Practical Problems of Entering Them into a GIS." *Fundamentals of Geographic Information Systems: A Compendium*, edited by William J. Ripple. Bethesda, Maryland: American Society for Photogrammetric Engineering and Remote Sensing and the American Congress on Surveying and Mapping.
- Defense Mapping Agency Aerospace Center. 1989. *Defense Mapping Agency Product Specifications for ARC Digitized Raster Graphics (ADRG)*. St. Louis, Missouri: DMA Aerospace Center.
- Dent, Borden D. 1985. *Principles of Thematic Map Design*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Duda, Richard O., and Peter E. Hart. 1973. *Pattern Classification and Scene Analysis*. New York: John Wiley & Sons, Inc.

Bibliography

- Eberlein, R. B., and J. S. Weszka. 1975. "Mixtures of Derivative Operators as Edge Detectors." *Computer Graphics and Image Processing*, Vol. 4: 180-183.
- Ebner, H., 1976. "Self-calibrating block adjustment." *Bildmessung und Luftbildwesen*, Vol. 4.
- Elachi, Charles. 1987 *Introduction to the Physics and Techniques of Remote Sensing*. New York: John Wiley & Sons.
- Elachi, Charles. 1992. "Radar Images of the Earth from Space." *Exploring Space*.
- Elachi, Charles. 1987. *Spaceborne Radar Remote Sensing: Applications and Techniques*. New York: IEEE Press.
- Elassal, Atef A., and Vincent M. Caruso. 1983. *USGS Digital Cartographic Data Standards: Digital Elevation Models*. Circular 895-B. Reston, Virginia: U.S. Geological Survey.
- El-Hakin, S.F., 1984. "A step-by-step strategy for gross error detection." *PE&RS*, 1984/6.
- ESRI. 1992. *ARC Command References 6.0*. Redlands, California: ESRI, Inc.
- ESRI. 1992. *Data Conversion: Supported Data Translators*. Redlands, California: ESRI, Inc.
- ESRI. 1992. *Managing Tabular Data*. Redlands, California: ESRI, Inc.
- ESRI. 1992. *Map Projections & Coordinate Management: Concepts and Procedures*. Redlands, California: ESRI, Inc.
- ESRI. 1990. *Understanding GIS: The ArcInfo Method*. Redlands, California: ESRI, Inc.
- ESRI, 1996. *Using ArcView GIS*. Redlands, California: ESRI, Inc.
- ESRI, 1997. ArcInfo. Version 7.2.1. *ArcInfo HELP*. Redlands, California: ESRI, Inc.
- Eurimage. "JERS-1: JERS-1." at http://www.eurimage.com/Products/JERS_1.html , 09 Nov 1998.
- European Space Agency. "ERS-2: A Continuation of the ERS-1 Success: ERS-2: A Continuation of the ERS-1 Success." at <http://esapub.esrin.esa.it/bulletin/bullet83/ducho83.htm> , 27 June 1996.
- European Space Agency. "SAR Mission Planning for ERS-1 and ERS-2: SAR Mission Planning for ERS-1 and ERS-2." at <http://esapub.esrin.esa.it/bulletin/bullet90/b90delia.htm> , 31 July 1997.
- Fahnestock, James D., and Robert A. Schowengerdt. 1983. "Spatially Variant Contrast Enhancement Using Local Range Modification." *Optical Engineering*, Vol. 22, No. 3.
- Faust, Nickolas L. 1989. "Image Enhancement." Volume 20, Supplement 5 of *Encyclopedia of Computer Science and Technology*, edited by Allen Kent and James G. Williams. New York: Marcel Dekker, Inc.

- FGDC, 1997. *Content Standards for Digital Orthoimagery*. Federal Geographic Data Committee, Washington, DC.
- Fisher, P. F. 1991. "Spatial Data Sources and Data Problems." *Geographical Information Systems: Principles and Applications*, edited by David J. Maguire, Michael F. Goodchild, and David W. Rhind. New York: Longman Scientific & Technical.
- Flaschka, H. A. 1969. *Quantitative Analytical Chemistry: Vol 1*. New York: Barnes & Noble, Inc.
- Förstner, W. and Gülch, E., 1987. "A fast operator for detection and precise location of distinct points, corners and centers of circular features." *Proceedings of Intercommission Conf. on Fast Processing of Photogrammetric Data*, 2-4 June, Interlaken, Switzerland, pp. 281-305. Available from Institute of Geodesy and Photogrammetry, ETH Zurich.
- Fraser, S. J., et al. 1986. "Targeting Epithermal Alteration and Gossans in Weathered and Vegetated Terrains Using Aircraft Scanners: Successful Australian Case Histories." Paper presented at the fifth Thematic Conference: Remote Sensing for Exploration Geology, Reno, Nevada.
- Freden, Stanley C., and Frederick Gordon, Jr. 1983. "Landsat Satellites." Chapter 12 in *Manual of Remote Sensing*, edited by Robert N. Colwell. Falls Church, Virginia: American Society of Photogrammetry.
- Free On-Line Dictionary Of Computing. "American Standard Code for Information Interchange from FOLDOC: American Standard Code for Information Interchange." at <http://foldoc.doc.ic.ac.uk/foldoc/> , 24 October 1999.
- Free On-Line Dictionary Of Computing. "central processing unit from FOLDOC: central processing unit." at http://foldoc.doc.ic.ac.uk/foldoc , 24 October 1999.
- Free On-Line Dictionary Of Computing. "lossy from FOLDOC: lossy." at http://foldoc.doc.ic.ac.uk/foldoc , 11 November 1999.
- Free On-Line Dictionary Of Computing. "random-access memory from FOLDOC: random-access memory." at http://foldoc.doc.ic.ac.uk/foldoc , 11 November 1999.
- Free On-Line Dictionary Of Computing. "wavelet from FOLDOC: wavelet." at http://foldoc.doc.ic.ac.uk/foldoc , 11 November 1999.
- Frost, Victor S., Stiles, Josephine A., Shanmugan, K. S., and Holtzman, Julian C. 1982. "A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-4, No. 2, March 1982.
- Geological Remote Sensing Group Newsletter*. May 1992. No. 5. Institute of Hydrology, Wallingford, OX10, United Kingdom.

Bibliography

- Gonzalez, Rafael C., and Paul Wintz. 1977. *Digital Image Processing*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Gonzalez, Rafael C., and Richard E. Woods. 1992. *Digital Image Processing*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Green, A.A. and Craig, M.D. 1985. "Analysis of Aircraft Spectrometer Data with Logarithmic Residuals." *Proceedings of the AIS Data Analysis Workshop*. JPL Pub. 85-41:111-119.
- Grün, A., 1978. "Experiences with self calibrating bundle adjustment." *Proceedings of ACSM-ASP Convention*, Washington.
- Grün, A. and Baltsavias, E.P., 1988. "Geometrically constrained multiphoto matching." *Photogrammetric Engineering and Remote Sensing*, Vol.54-5, pp. 309-312.
- Guptill, Stephen C., ed. 1988. *A Process for Evaluating Geographic Information Systems*. U.S. Geological Survey Open-File Report 88-105.
- Haralick, Robert M. 1979. "Statistical and Structural Approaches to Texture." *Proceedings of the IEEE*, Vol. 67, No. 5: 786-804. Seattle, Washington.
- Heipke, Ch., 1996. "Automation of interior, relative and absolute orientation." *International Archives of Photogrammetry and Remote Sensing*, Vol. 31, Part B3, pp. 297 - 311.
- Helava, U.V., 1988. "Object space least square correlation." *International Archives of Photogrammetry and Remote Sensing*, Vol. 27, Part B3, p. 321.
- Hodgson, Michael E., and Bill M. Shelley. 1993. "Removing the Topographic Effect in Remotely Sensed Imagery." *ERDAS Monitor*, Fall 1993. Contact Dr. Hodgson, Dept. of Geography, University of Colorado, Boulder, CO 80309-0260.
- Holcomb, Derrold W. 1993. "Merging Radar and VIS/IR Imagery." Paper submitted to the 1993 ERIM Conference, Pasadena, California.
- Hord, R. Michael. 1982. *Digital Image Processing of Remotely Sensed Data*. New York: Academic Press.
- Iron, James R., and Gary W. Petersen. 1981. "Texture Transforms of Remote Sensing Data," *Remote Sensing of Environment*, Vol. 11: 359-370.
- Jacobsen, K., 1980. *Vorschläge zur Konzeption und zur Bearbeitung von Bündelblockausgleichungen*. Ph.D. dissertation, wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, No. 102.
- Jacobsen, K., 1982. *Programmgesteuerte Auswahl der zusätzlicher Parameter*. *Bildmessung und Luftbildwesen*, p. 213.
- Jacobsen, K., 1984. "Experiences in blunder detection." *ISPRS 15th Congress*, Rio de Janeiro.

- Jacobsen, K., 1994. "Combined block adjustment with precise differential GPS data." *International Archives of Photogrammetry and Remote Sensing*, Vol. 30, Part B3, p. 422.
- Jensen, John R., et al. 1983. "Urban/Suburban Land Use Analysis." Chapter 30 in *Manual of Remote Sensing*, edited by Robert N. Colwell. Falls Church, Virginia: American Society of Photogrammetry.
- Jensen, John R. 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Johnston, R. J. 1980. *Multivariate Statistical Analysis in Geography*. Essex, England: Longman Group Ltd.
- Jordan, III, Lawrie E., Bruce Q. Rado, and Stephen L. Sperry. 1992. "Meeting the Needs of the GIS and Image Processing Industry in the 1990s." *Photogrammetric Engineering & Remote Sensing*, Vol. 58, No. 8: 1249-1251.
- Jordan, III, Lawrie E. and Larry Beck. 1999. *NITFS—The National Imagery Transmission Format Standard*. Atlanta, Georgia: ERDAS, Inc.
- Keates, J. S. 1973. *Cartographic Design and Production*. London: Longman Group Ltd.
- Kennedy, Michael 1996. "The Global Positioning System and GIS: An Introduction." Chelsea, Michigan: Ann Arbor Press, Inc.
- Kidwell, Katherine B., ed. 1988. *NOAA Polar Orbiter Data (TIROS-N, NOAA-6, NOAA-7, NOAA-8, NOAA-9, and NOAA-10) Users Guide*. Washington, DC: National Oceanic and Atmospheric Administration.
- Kloer, Brian R. 1994. "Hybrid Parametric/Non-parametric Image Classification." Paper presented at the ACSM-ASPRS Annual Convention, April 1994, Reno, Nevada.
- Kneizys, F. X., et al. 1988. *Users Guide to LOWTRAN 7*. Hanscom AFB, Massachusetts: Air Force Geophysics Laboratory.
- Knuth, Donald E. 1987. "Digital Halftones by Dot Diffusion." *ACM Transactions on Graphics*, Vol. 6: 245-273.
- Konecny, G., 1994. "New trends in technology, and their applications: photogrammetry and remote sensing from analog to digital." *Thirteenth United Nations Regional Cartographic Conference for Asia and the Pacific*, Beijing, 9-15 May 1994.
- Konecny, G. and Lehmann, G., 1984. *Photogrammetrie*. Walter de Gruyter Verlag, Berlin.
- Kraus, K., 1984. *Photogrammetrie*. Band II. Dümmlers Verlag, Bonn.
- Kruse, Fred A. 1988. "Use of Airborne Imaging Spectrometer Data to Map Minerals Associated with Hydrothermally Altered Rocks in the Northern Grapevine Mountains, Nevada." *Remote Sensing of the Environment*, Vol. 24: 31-51.

Bibliography

- Krzystek, P., 1998. "On the use of matching techniques for automatic aerial triangulation." *Proceedings of ISPRS commission III conference, 1998*. Columbus, Ohio, USA.
- Kubik, K., 1982. "An error theory for the Danish method." *ISPRS Commission III conference, Helsinki, Finland*.
- Larsen, Richard J., and Morris L. Marx. 1981. *An Introduction to Mathematical Statistics and Its Applications*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Lavreau, J. 1991. "De-Hazing Landsat Thematic Mapper Images." *Photogrammetric Engineering & Remote Sensing*, Vol. 57, No. 10: 1297-1302.
- Leberl, Franz W. 1990. *Radargrammetric Image Processing*. Norwood, Massachusetts: Artech House, Inc.
- Lee, J. E., and J. M. Walsh. 1984. *Map Projections for Use with the Geographic Information System*. U.S. Fish and Wildlife Service, FWS/OBS-84/17.
- Lee, Jong-Sen. 1981. "Speckle Analysis and Smoothing of Synthetic Aperture Radar Images." *Computer Graphics and Image Processing*, Vol. 17:24-32.
- Leick, Alfred. 1990. *GPS Satellite Surveying*. New York, New York: John Wiley & Sons.
- Li, D., 1983. *Ein Verfahren zur Aufdeckung grober Fehler mit Hilfe der a posteriori-Varianzschätzung*. *Bildmessung und Luftbildwesen*. Vol. 5.
- Li, D., 1985. *Theorie und Untersuchung der Trennbarkeit von groben Paßpunktfehlern und systematischen Bildfehlern bei der photogrammetrischen punktbestimmung*. Ph.D. dissertation, Deutsche Geodätische Kommission, Reihe C, No. 324.
- Lillesand, Thomas M., and Ralph W. Kiefer. 1987. *Remote Sensing and Image Interpretation*. New York: John Wiley & Sons, Inc.
- Lopes, A., Nezry, E., Touzi, R., and Laur, H. 1990. "Maximum A Posteriori Speckle Filtering and First Order Texture Models in SAR Images." *International Geoscience and Remote Sensing Symposium (IGARSS)*.
- Lü, Yan, 1988. "Interest operator and fast implementation." *IASPRS* Vol. 27, B2, Kyoto, 1988.
- Lue, Yan and Kurt Novak. 1991. "Recursive Grid - Dynamic Window Matching for Automatic DEM Generation." *1991 ACSM-ASPRS Fall Convention Technical Papers*.
- Lyon, R.J.P. 1987. "Evaluation of AIS-2 Data over Hydrothermally Altered Granitoid Rocks." *Proceedings of the Third AIS Data Analysis Workshop*. JPL Pub. 87-30:107-119.

- Magellan Corporation. "GLONASS: Technology Facts, GLONASS and the GPS+GLONASS Advantage." at <http://www.magellangps.com/geninfo/glonass.htm> , 24 August 1999.
- Maling, D. H. 1992. *Coordinate Systems and Map Projections*. 2nd ed. New York: Pergamon Press.
- Marble, Duane F. 1990. "Geographic Information Systems: An Overview." *Introductory Readings in Geographic Information Systems*, edited by Donna J. Peuquet and Duane F. Marble. Bristol, Pennsylvania: Taylor & Francis, Inc.
- Mayr, W., 1995. "Aspects of automatic aerotriangulation." *Proceedings of 45th Photogrammetric Week*, Wichmann Verlag, Karlsruhe, pp. 225 - 234.
- Mendenhall, William, and Richard L. Scheaffer. 1973. *Mathematical Statistics with Applications*. North Scituate, Massachusetts: Duxbery Press.
- Menon, Sudhakar, Peng Gao, and CiXiang Zhan. 1991. "GRID: A Data Model and Functional Map Algebra for Raster Geo-processing." *GIS/LIS '91 Proceedings*, Vol. 2: 551-561. Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing.
- Merenyi, E., Taranik, J.V., Monor, Tim, and Farrand, W. March 1996. "Quantitative Comparison of Neural Network and Conventional Classifiers for Hyperspectral Imagery." *Proceedings of the Sixth AVIRIS Conference*. JPL Pub.
- Minnaert, J. L., and G. Szeicz. 1961. "The Reciprocity Principle in Lunar Photometry." *Astrophysics Journal*, Vol. 93: 403-410.
- Moffit, Francis H. and Edward M. Mikhail, 1980. *Photogrammetry*. New York: Harper& Row Publishers.
- NASA. "Daedalus TMS Description: Thematic Mapper Simulators (TMS)." at http://geo.arc.nasa.gov/esdstaff/jskiles/top-down/OTTER/OTTER_docs/DAEDALUS.html , 20 December 1995.
- NASA. "Mission Overview: Mission Overview." at <http://southport.jpl.nasa.gov/science/missiono.html> , 11 January 1995.
- NASA. "SAR Document: SAR Development." at http://southport.jpl.nasa.gov/reports/iwgsar/3_SAR_Development.html , 21 October 1996.
- NASA. "SIR-C Description: What is SIR-C/X-SAR." at <http://southport.jpl.nasa.gov/desc/SIRCdesc.html> , 14 January 1997.
- NASA. "Landsat 7 Facts: Landsat 7." at <http://geo.arc.nasa.gov/sge/landsat/landsat.html> , 05 October 1999

Bibliography

- NASA. "SeaWiFS Project—Spacecraft and Sensor Overview: An Overview of SeaWiFS and the SeaStar Spacecraft." at <http://seawifs.gsfc.nasa.gov/SEAWIFS/SEASTAR/SPACECRAFT.html> , 17 August 1999.
- Nagao, Makoto, and Takashi Matsuyama. 1978. "Edge Preserving Smoothing." *Computer Graphics and Image Processing*, Vol. 9: 394-407.
- Needham, Bruce H. 1986. "Availability of Remotely Sensed Data and Information from the U.S. National Oceanic and Atmospheric Administration's Satellite Data Services Division." Chapter 9 in *Satellite Remote Sensing for Resources Development*, edited by Karl-Heinz Szekiela. Gaithersburg, Maryland: Graham & Trotman, Inc.
- Nichols, David, et al. 1983. "Digital Hardware." Chapter 20 in *Manual of Remote Sensing*, edited by Robert N. Colwell. Falls Church, Virginia: American Society of Photogrammetry.
- NIMA. "nimafs.html: The National Imagery and Mapping Agency Fact Sheet." at <http://164.214.2.59/general/factsheets/nimafs.html> , October 1998.
- Oppenheim, Alan V., and Ronald W. Schafer. 1975. *Digital Signal Processing*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- ORBIMAGE. "OrbView-3: High-resolution Imagery in Real-time." at <http://www.orbimage.com/satellite/orbview3/orbview3.html> , 13 July 1999.
- Parent, Phillip, and Richard Church. 1987. "Evolution of Geographic Information Systems as Decision Making Tools." *Fundamentals of Geographic Information Systems: A Compendium*, edited by William J. Ripple. Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping.
- Pearson, Frederick. 1990. *Map Projections: Theory and Applications*. Boca Raton, Florida: CRC Press, Inc.
- Peli, Tamar, and Jae S. Lim. 1982. "Adaptive Filtering for Image Enhancement." *Optical Engineering*, Vol. 21, No. 1.
- Pratt, William K. 1991. *Digital Image Processing*. New York: John Wiley & Sons, Inc.
- Press, William H., et al. 1988. *Numerical Recipes in C*. New York, New York: Cambridge University Press.
- Prewitt, J. M. S. 1970. "Object Enhancement and Extraction." *Picture Processing and Psychopictorics*, edited by B. S. Lipkin, and A. Resenfeld. New York: Academic Press.
- RADARSAT. "RADARSAT: RADARSAT Specifications." at <http://radarsat.space.gc.ca/> , 14 September 1999.

- Rado, Bruce Q. 1992. "An Historical Analysis of GIS." *Mapping Tomorrow's Resources*. Logan, Utah: Utah State University.
- Richter, Rudolf. 1990. "A Fast Atmospheric Correction Algorithm Applied to Landsat TM Images." *International Journal of Remote Sensing*, Vol. 11, No. 1: 159-166.
- Ritter, Dr. Niles and Mike Ruth. "GeoTIFF Spec: GeoTIFF Format Specification." Rev. 1.0. at <http://www.remotesensing.org/geotiff/spec/geotiffhome.html> , 02 March 1999.
- Robinson, Arthur H., and Randall D. Sale. 1969. *Elements of Cartography*. 3rd ed. New York: John Wiley & Sons, Inc.
- Sabins, Floyd F., Jr. 1987. *Remote Sensing Principles and Interpretation*. New York: W. H. Freeman and Co.
- Sader, S. A., and J. C. Winne. 1992. "RGB-NDVI Colour Composites For Visualizing Forest Change Dynamics." *International Journal of Remote Sensing*, Vol. 13, No. 16: 3055-3067.
- Schenk, T., 1997. "Towards automatic aerial triangulation." *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(3):110-121.
- Schowengerdt, Robert A. 1983. *Techniques for Image Processing and Classification in Remote Sensing*. New York: Academic Press.
- Schowengerdt, Robert A. 1980. "Reconstruction of Multispatial, Multispectral Image Data Using Spatial Frequency Content." *Photogrammetric Engineering & Remote Sensing*, Vol. 46, No. 10: 1325-1334.
- Schwartz, A. A., and J. M. Soha. 1977. "Variable Threshold Zonal Filtering." *Applied Optics*, Vol. 16, No. 7.
- Short, Nicholas M. 1982. *The Landsat Tutorial Workbook: Basics of Satellite Remote Sensing*. Washington, DC: National Aeronautics and Space Administration.
- Simonett, David S., et al. 1983. "The Development and Principles of Remote Sensing." Chapter 1 in *Manual of Remote Sensing*, edited by Robert N. Colwell. Falls Church, Virginia: American Society of Photogrammetry.
- Slater, P. N. 1980. *Remote Sensing: Optics and Optical Systems*. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc.
- Smith, J., T. Lin, and K. Ranson. 1980. "The Lambertian Assumption and Landsat Data." *Photogrammetric Engineering & Remote Sensing*, Vol. 46, No. 9: 1183-1189.
- Snyder, John P. 1987. *Map Projections--A Working Manual*. Geological Survey Professional Paper 1532. Washington, DC: United States Government Printing Office.

Bibliography

- Snyder, John P., and Philip M. Voxland. 1989. *An Album of Map Projections*. U.S. Geological Survey Professional Paper 1453. Washington, DC: United States Government Printing Office.
- Space Imaging. "Space Imaging—IKONOS: IKONOS." at <http://www.spaceimage.com/aboutus/satellites/IKONOS/ikonos.html> , 30 September 1999.
- Space Imaging. "IRS-ID Satellite Imagery Available for Sale Worldwide: IRS-ID Satellite Imagery Available for Sale Worldwide." at <http://www.spaceimage.com/newsroom/releases/1998/IRS1Dworldwide.html> , 01 October 1999.
- Space Imaging. "Space Imaging—IRS: IRS (Indian Remote Sensing Satellite)." at <http://www.spaceimage.com/aboutus/satellites/IRS/IRS.html> , 17 September 1999.
- Space Imaging. "Space Imaging—Landsat: LANDSAT TM." at <http://www.spaceimage.com/aboutus/satellites/Landsat/landsat.html> , 17 September 1999.
- Space Imaging. "Space Imaging—RADARSAT: RADARSAT." at <http://www.spaceimage.com/aboutus/satellites/RADARSAT/radarsat.htm> , 17 September 1999.
- SPOT Image. "Page Title: SPOT 4—In Service!" at <http://www.spot.com/spot/home/news/press/Commish.htm> , 09 June 1999.
- SPOT Image. "Spot Image: Technical Data: SPOT System Technical Data." at <http://www.spot.com/spot/home/system/introsat/seltec/seltec.htm> , 12 February 1999.
- Srinivasan, Ram, Michael Cannon, and James White, 1988. "Landsat Destriping Using Power Spectral Filtering." *Optical Engineering*, Vol. 27, No. 11: 939-943.
- Star, Jeffrey, and John Estes. 1990. *Geographic Information Systems: An Introduction*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Steinitz, Carl, Paul Parker, and Lawrie E. Jordan, III. 1976. "Hand Drawn Overlays: Their History and Perspective Uses." *Landscape Architecture*, Vol. 66.
- Stimson, George W. 1983. *Introduction to Airborne Radar*. El Segundo, California: Hughes Aircraft Company.
- Stojić, M., *et al.*, 1998. "The assessment of sediment transport rates by automated digital photogrammetry." *PE&RS*. Vol. 64, No. 5, pp. 387 - 395.
- Suits, Gwynn H. 1983. "The Nature of Electromagnetic Radiation." Chapter 2 in *Manual of Remote Sensing*, edited by Robert N. Colwell. Falls Church, Virginia: American Society of Photogrammetry.

- Swain, Philip H. 1973. *Pattern Recognition: A Basis for Remote Sensing Data Analysis* (LARS Information Note 111572). West Lafayette, Indiana: The Laboratory for Applications of Remote Sensing, Purdue University.
- Swain, Philip H., and Shirley M. Davis. 1978. *Remote Sensing: The Quantitative Approach*. New York: McGraw Hill Book Company.
- Tang, L., Braun, J. and Debitsch R., 1997. "Automatic aerotriangulation - concept, realization and results." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol.52, pp121-131.
- Taylor, Peter J. 1977. *Quantitative Methods in Geography: An Introduction to Spatial Analysis*. Boston, Massachusetts: Houghton Mifflin Company.
- TIFF Developer's Toolkit*. 1990. Seattle, Washington: Aldus Corp.
- Tou, Julius T., and Rafael C. Gonzalez. 1974. *Pattern Recognition Principles*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Tsingas, V., 1995. "Operational use and empirical results of automatic aerial triangulation." *Proceedings of 45th Photogrammetric Week*, Wichmann Verlag, Karlsruhe, pp. 207 - 214.
- Tucker, Compton J. 1979. "Red and Photographic Infrared Linear Combinations for Monitoring Vegetation." *Remote Sensing of Environment*, Vol. 8: 127-150.
- United States Geological Survey. "No web page title: About the EROS Data Center." at http://edcwww.cr.usgs.gov/content_about.html , no date given.
- United States Geological Survey. "LANDSAT_TM: Landsat Thematic Mapper Data." at http://edc.usgs.gov/glis/hyper/guide/landsat_tm , 13 September 1999.
- United States Geological Survey. "Landsat: National Landsat Archive Production System (NLAPS)." at <http://edc.usgs.gov/glis/hyper/guide/nlaps.html> , no date given.
- United States Geological Survey. "USGS TerraServer—Digital Backyard: Digital Orthophoto Quadrangles (DOQ's): Digital Orthophoto Quadrangles." at <http://mapping.usgs.gov/digitalbackyard/doqs.html>, no date given.
- United States Geological Survey. "What is SDTS?: What is SDTS?" at <http://mcmweb.er.usgs.gov/sdts/whatsdts.html> , no date given.
- Vosselman, G. and Haala, N., 1992. "Erkennung topographischer Paßpunkte durch relationale Zuordnung." *Zeitschrift für Photogrammetrie und Fernerkundung*, (6): 170-176.
- Walker, Terri C., and Richard K. Miller. 1990. *Geographic Information Systems: An Assessment of Technology, Applications, and Products*. Madison, Georgia: SEAI Technical Publications.

Bibliography

- Wang, Y., 1988. "A combined adjustment program system for close range photogrammetry." *Journal of Wuhan Technical University of Surveying and Mapping*, Vol. 12, No. 2.
- Wang, Y., 1994. *Strukturzuordnung zur automatischen Oberflächenrekonstruktion*. Ph.D. dissertation, wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, No. 207.
- Wang, Y., 1998. "Principles and applications of structural image matching." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol.53, pp. 154-165.
- Wang, Z., 1990. *Principles of photogrammetry (with Remote Sensing)*. Press of Wuhan Technical University of Surveying and Mapping and Publishing House of Surveying and Mapping, Beijing, China.
- Watson, David, 1994, *Contouring: A Guide to the Analysis and Display of Spatial Data*, Elsevier Science, New York.
- Welch, Roy. 1990. "3-D Terrain Modeling for GIS Applications." *GIS World*, Vol. 3, No. 5.
- Welch, R., and W. Ehlers. 1987. "Merging Multiresolution SPOT HRV and Landsat TM Data." *Photogrammetric Engineering & Remote Sensing*, Vol. 53, No. 3: 301-303.
- Wolberg, George. 1990. *Digital Image Warping*. IEEE Computer Society Press Monograph.
- Wolf, Paul R. 1980. "Definitions of Terms and Symbols used in Photogrammetry." *Manual of Photogrammetry*. Ed. Chester C. Slama. Falls Church, Virginia: American Society of Photogrammetry.
- Wolf, Paul R., 1983. *Elements of Photogrammetry*. New York: McGraw-Hill, Inc.
- Wong, K.W. 1980. "Basic Mathematics of Photogrammetry." *Manual of Photogrammetry*. Ed. Chester C. Slama. Falls Church, Virginia: American Society of Photogrammetry.
- Yang, X., 1997. *Georeferencing CAMS Data: Polynomial Rectification and Beyond*, Dissertation, University of South Carolina, Columbia, SC.
- Yang, X. and D. Williams, 1997. "The Effect of DEM Data Uncertainty on the Quality of Orthoimage Generation." *Proceedings of GIS/LIS 97*, Cincinnati, Ohio.
- Yang, Xinghe, R. Robinson, H. Lin, and A. Zusmanis. 1993. "Digital Ortho Corrections Using Pre-transformation Distortion Adjustment." *1993 ASPRS Technical Papers*. New Orleans. Vol. 3: 425-434.
- Zamudio, J.A. and Atkinson, W.W. 1990. "Analysis of AVIRIS data for Spectral Discrimination of Geologic Materials in the Dolly Varden Mountains." *Proceedings of the Second AVIRIS Conference*. JPL Pub. 90-54:162-66.

Index

Symbols

.OV1 (overview image) 91
.OVR (overview image) 91
.stk (GRID Stack file) 101

Numerics

1:24,000 scale 93
1:250,000 scale 93
2D affine transformation 274
4 mm tape 21, 22
7.5-minute DEM 93
8 mm tape 21, 22
9-track tape 21, 23

A

a priori 221, 347
Absorption 7, 10
 spectra 6, 11
Absorption spectra 174
Accuracy assessment 256, 258
Accuracy report 260
Active sensor 6
Adaptive filter 162
Additional Parameter modeling (AP) 303
ADRG 23, 50, 82
 file naming convention 87
 ordering 98
ADRI 50, 88
 file naming convention 89
 ordering 98
Aerial photography 82
Aerial photos 4, 25, 262
Aerial triangulation (AT) 264, 281
Airborne GPS 281
Airborne imagery 49
Airborne Imaging Spectrometer 11
Airborne Multispectral Scanner Mk2 11

AIRSAR 73, 79
Aitoff 588
Albers Conical Equal Area 481, 518
Almaz 73
Almaz 1-B 74
Almaz-1 74
Along track adjustment 313
Analog photogrammetry 261
Analytical photogrammetry 261
Annotation 54, 126, 131, 422
 element 422
 in script models 412
 layer 423
ANT (Erdas 7.x) (annotation) 54
AP 303
Arc Coverage 50
Arc Interchange (raster) 50
Arc Interchange (vector) 55
ARC system 83, 90
Arc/second format 92
Arc_Interchange to Coverage 55
Arc_Interchange to Grid 55
ARCGEN 50, 55, 106
ArcInfo 49, 50, 101, 106, 345, 387, 391
 coverages 35
 data model 35, 38
 UNGENERATE 106
ArcInfo GENERATE 45
ArcInfo INTERCHANGE 45
ArcView 45
Area based matching 292
Area of interest 30, 153, 396
ASCII 94, 393
ASCII Raster 50
ASCII To Point Annotation (annotation) 54
ASCII To Point Coverage (vector) 55
Aspect 375, 379, 435
 calculating 379
 equatorial 435
 oblique 435

Index

B

- polar 435
- transverse 436
- AT 281
- Atmospheric correction 153
- Atmospheric effect 142
- Atmospheric modeling 143
- Attribute
 - imported 40
 - in model 411
 - information 35, 37, 39
 - raster 391
 - thematic 390, 391
 - vector 391, 393
 - viewing 392
- Auto update 115, 116, 117
- AutoCAD 45, 49, 106
- Automatic image correlation 320
- Automatic tie point collection 290
- Average 463
- AVHRR 15, 50, 57, 65, 143, 175, 347
 - extract 66
 - full set 66
 - ordering 97
- AVHRR (Dundee Format) 50
- AVHRR (NOAA) 50
- AVIRIS 11, 79, 80
- Azimuth 433
- Azimuthal Equidistant 483, 484

B

- Band 2, 61, 62, 66, 68
 - displaying 118
- Banding 18, 141
 - see also striping
- Bartlett window 193
- Basic Image Interchange Format 53
- Bayesian classifier 240
- Beam mode selection 316
- Behrmann 486
- BIL 18, 19, 50

- Bin 150
- Binary format 18
- BIP 18, 21, 50
- Bipolar Oblique Conic Conformal 580
- Bit 18
 - display device 111
 - in files (depth) 16, 18, 22
- Block of images 289
- Block triangulation 264, 281
- Blocking factor 20, 22
- Bonne 487
- Border 427
- Bpi 23
- Brightness inversion 153
- Brightness value 16, 112, 113, 119, 123, 144
- Brovey Transform 161
- BSQ 18, 19, 50
- Buffer zone 396
- Bundle block adjustment 264, 281
- Butterworth window 193
- Byte 18

C

- CADRG (Compressed ADRG) 50
- Canadian Geographic Information System
 - 385
- Cartesian coordinate 37
- Cartography 419
- Cartridge tape 21, 22
- Cassini 489, 581
- Cassini-Soldner 581
- Categorical
 - data 3
- CCD 296
- CD-ROM 18, 21, 23
- Cell 94
- Change detection 17, 29, 56, 344
- Chi-square
 - distribution 256
 - statistics 258

- CIB 91
- CIB (Controlled Image Base) 50
- Class 217, 389
 - name 390, 392
 - value 123, 392
 - numbering systems 389, 400
- Classical aerial triangulation 264
- Classification 30, 74, 200, 217, 390
 - and enhanced data 163, 220
 - and rectified data 346
 - and terrain analysis 375
 - evaluating 256
 - flow chart 243
 - iterative 220, 224, 241
 - scheme 219
- Clump 397
- Clustering 227
 - ISODATA 227
 - RGB 227, 232
- Coefficient 471
 - in convolution 155
 - of variation 201
- Collinearity condition 278
- Collinearity equations 282, 302
- Color gun 112, 119
- Color scheme 41, 390
- Color table 123, 392
 - for printing 456
- Colorcell 113
 - read-only 114
- Color-infrared 118
- Colormap 113, 126
- Complex image 55
- Confidence level 258
- Conformality 433
- Contiguity analysis 395
- Contingency matrix 236, 237, 238
- Continuous
 - data 3
- Continuous data
 - see data
- Contrast stretch 30, 144
 - for display 119, 147, 466
 - linear 145
 - min/max vs. standard deviation 120, 148
 - nonlinear 146
 - piecewise linear 146
- Contrast table 119
- Control point extension 264
- Controlled Image Base 91
- Convergence value 285
- Convolution 18
 - cubic 370
 - filtering 155, 370, 371, 398
 - kernel
 - crisp 159
 - edge detector 158
 - edge enhancer 158
 - gradient 206
 - high frequency 157, 158
 - low frequency 159, 370
 - Prewitt 205
 - zero-sum 158
- Convolution kernel 155
 - high frequency 371
 - low frequency 371
 - Prewitt 206
- Coordinate
 - Cartesian 37, 437
 - conversion 373
 - file 3, 37, 134
 - geographic 437, 508
 - map 3, 4, 37, 343, 346
 - planar 437
 - reference 346
 - retransformed 367
 - source 346
 - spherical 437
- Coordinate system 3, 269
 - ground space 269

Index

D

- image space 269
- Correlation calculations 292
- Correlation coefficient 324
- Correlation threshold 349
- Correlation windows 292
- Correlator library 324
- Covariance 238, 249, 466
 - sample 466
- Covariance matrix 167, 235, 240, 251, 467
- Cross correlation 292

D

- DAEDALUS 50, 80
- DAEDALUS TMS
 - bands/frequencies 80
- Data 386
 - airborne sensor 49
 - ancillary 221
 - categorical 3
 - complex 55
 - compression 164, 232
 - continuous 3, 24, 118, 388
 - displaying 122
 - creating 135
 - elevation 221, 375
 - enhancement 138
 - from aircraft 79
 - geocoded 20, 28, 344
 - gray scale 131
 - hyperspectral 11
 - interval 3
 - nominal 3
 - ordering 97
 - ordinal 3
 - packed 66
 - pseudo color 131
 - radar 49, 70
 - applications 74
 - bands 73
 - merging 213

- raster 3, 126
 - converting to vector 100
 - editing 30
 - formats (BIL, etc.) 22
 - importing and exporting 49
 - in GIS 387
 - sources 49
- ratio 3
- satellite 49
- structure 168
- thematic 3, 24, 123, 223, 390
 - displaying 125
- tiled 26, 55
- topographic 92, 375
 - using 94
- true color 131
- vector 126, 131, 345, 373, 391
 - converting to raster 100, 390
 - copying 39
 - displaying 40
 - editing 415
 - densify 415
 - generalize 415
 - reshape 415
 - spline 415
 - split 415
 - unsplit 415
 - from raster data 42
 - importing 42, 55
 - in GIS 388
 - renaming 39
 - sources 42, 45, 49
 - structure 38
- viewing 130
 - multiple layers 131
 - overlapping layers 131
- Data correction 17, 30, 137, 140
 - geometric 141, 143, 343
 - radiometric 141, 210, 344
- Data file value 1, 30

- display 119, 134, 146
 - in classification 217
- Data storage 18
- Database
 - image 28
- Decision rule 218, 242
 - Bayesian 250
 - feature space 246
 - Mahalanobis distance 249
 - maximum likelihood 250, 256
 - minimum distance 247, 250, 256, 258
 - non-parametric 242
 - parallelepiped 244
 - parametric 243
- Decision tree 254
- Decorrelation stretch 168
- Degrees of freedom 258
- DEM 1, 25, 92, 93, 143, 344, 376
 - editing 31
 - interpolation 31
 - ordering 97
- Density 23
- Density of detail 325
- Descriptive information
 - see attribute information 39
- Desktop Scanners 81
- Desktop scanners 267
- Detector 57, 141
- DFAD 55
- DGN 45
- DGN (Intergraph IGDS) (vector) 55
- Differential collection geometry 333
- Differential correction 95
- DIG Files (Erdas 7.x) (vector) 55
- Digital image 43
- Digital orthophoto 304
- Digital orthophotography 376
- Digital photogrammetry 261
- Digital picture
 - see image 112
- Digital terrain model (DTM) 49
- Digitizing 43, 345
 - GCPs 347
 - operation modes 44
 - point mode 44
 - screen 42, 44
 - stream mode 44
 - tablet 42
- DIME 108
- Dimensionality 164, 220, 221, 470
- Direction of flight 265
- Discrete Cosine Transformation 101
- Disk space 23
- Diskette 18, 21
- Display
 - 32-bit 114
 - DirectColor 114, 115, 121, 124
 - HiColor 114, 117
 - PC 118
 - PseudoColor 114, 115, 117, 122, 125
 - TrueColor 114, 116, 117, 121, 124
- Display device 111, 119, 144, 147, 169
- Display memory 138
- Display resolution 111
- Distance image file 256, 257
- Distribution 460
- Distribution Rectangle 83
- Dithering 129
 - color artifacts 130
 - color patch 130
- Divergence 236
 - signature 238
 - transformed 239
- DLG 23, 45, 49, 55, 107
- Doppler centroid 311
- Doppler cone 312
- DOQ 50
- DOQ (JPEG) 50
- DOQs 82
- DTED 50, 88, 92, 376

Index

E

- DTM 49
- DXF 45, 49, 106
- DXF to Annotation 55
- DXF to Coverage 56
- Dynamic range 16

- E**
- Earth Centered System 311
- Earth Fixed Body coordinate system 309
- Earth model 312
- Eckert I 490
- Eckert II 492
- Eckert III 493
- Eckert IV 495
- Eckert V 496
- Eckert VI 498
- Edge detection 158, 204
- Edge enhancement 158
- Eigenvalue 164
- Eigenvector 164, 167
- Eikonix 81
- Electric field 327
- Electromagnetic radiation 5, 57
- Electromagnetic spectrum 2, 5
 - long wave infrared region 5
 - short wave infrared region 5
- Electromagnetic wave 326, 338
- Elements of exterior orientation 276
- Ellipse 164, 236
- Ellipsoid 318
- Enhancement 30, 63, 137, 217
 - linear 119, 145
 - nonlinear 145
 - on display 126, 134, 138
 - radar data 137, 197
 - radiometric 137, 143, 154
 - spatial 137, 143, 154
 - spectral 137, 163
- Entity (AutoCAD) 107
- EOF (end of file) 20
- EOSAT 17, 29, 61, 98
- EOSAT SOM 499
- EOV (end of volume) 20
- Ephemeris adjustment 313
- Ephemeris coordinate system 309
- Ephemeris data 299
- Ephemeris modeling 310
- equal area
 - see equivalence
- Equidistance 433
- Equidistant Conic 500, 528
- Equidistant Cylindrical 502, 581
- Equirectangular 503
- Equivalence 433
- ER Mapper 50
- ERDAS macro language (EML) 395
- ERDAS Version 7.X 24, 100
- EROS Data Center 99
- Error matrix 259
- ERS (Conae-PAF CEOS) 51
- ERS (D-PAF CEOS) 51
- ERS (I-PAF CEOS) 51
- ERS (Tel Aviv-PAF CEOS) 51
- ERS (UK-PAF CEOS) 51
- ERS-1 73, 74, 75
 - ordering 99
- ERS-2 74, 76
- ESRI 35, 385
- ETAK 45, 49, 56, 108
- Euclidean distance 248, 256, 471
- Expected value 464
- Expert classification 253
- Exposure station 265
- Extent 423
- Exterior orientation 276
 - SPOT 299

- F**
- .fsp.img file 225
- False color 63

- False easting 437
- False northing 437
- Fast format 20
- Fast Fourier Transform 183
- Feature based matching 294
- Feature extraction 137
- Feature point matching 294
- Feature space 469
 - area of interest 222, 225
 - image 224, 470
- Fiducial marks 273
- Field 391
- File
 - .fsp.img 225
 - .gcc 348
 - .GIS 24, 100
 - .gmd 407
 - .img 2, 24, 118
 - .LAN 24, 100
 - .mdl 412
 - .ovr 423
 - .sig 466
 - archiving 28
 - header 22
 - output 23, 27, 365, 366, 412
 - .img 147
 - classification 260
 - pixel 112
 - tic 37
- File coordinate 3
- File name 27
- Film recorder 454
- Filter
 - adaptive 162
 - Fourier image 189
 - Frost 197, 203
 - Gamma-MAP 197, 203
 - high-pass 191
 - homomorphic 195
 - Lee 197, 200
 - Lee-Sigma 197
 - local region 197, 199
 - low-pass 189
 - mean 197, 198
 - median 18, 197, 198, 199
 - periodic noise removal 194
 - Sigma 200
 - zero-sum 207
- Filtering 155
 - see also convolution filtering
- Flattening 337
- Flight path 265
- Focal analysis 18
- Focal length 273
- Focal operation 31, 398
- Focal plane 273
- Fourier analysis 137, 184
- Fourier magnitude 183
 - calculating 185
- Fourier Transform
 - calculation 185
 - Editor
 - window functions 191
 - filtering 189
 - high-pass filtering 191
 - inverse
 - calculating 188
 - low-pass filtering 189
 - neighborhood techniques 183
 - noise removal 194
 - point techniques 183
- Frequency
 - statistical 460
- Frost filter 197, 203
- Function memory 138
- Fuyo 1 73
 - ordering 99
- Fuzzy convolution 252
- Fuzzy methodology 251

Index

G

G

- .gcc file 348
- .GIS file 24, 100
- .gmd file 407
- GAC (Global Area Coverage) 66
- Gamma-MAP filter 197, 203
- Gauss Kruger 506
- Gaussian distribution 202
- Gauss-Krüger 587
- GCP 346, 472
 - corresponding 347
 - digitizing 347
 - matching 348
 - minimum required 359
 - prediction 348
 - selecting 347
- GCP configuration 289
- GCP requirements 288
- GCPs 287
- General Vertical Near-side Perspective 507
- Generic Binary 51
- Geocentric coordinate system 270
- Geocoded data 20
- Geocoding
 - GeoTIFF 105
- Geographic Information System
 - see GIS
- Geoid 318
- Geology 74
- Georeference 344, 451
- Georeferencing
 - GeoTIFF 105
- GeoTIFF 51, 104
 - geocoding 105
 - georeferencing 105
- Gigabyte 18
- GIS 1
 - database 387
 - defined 386
 - history 385
 - GIS (Erdas 7.x) 51
- Glaciology 74
- Global operation 31
- GLONASS 94
- Gnomonic 510
- GOME 76
- GPS data 94
- GPS data applications 96
- GPS satellite position 95
- Gradient kernel 206
- Graphical model 138, 394
 - convert to script 412
 - create 407
- Graphical modeling 396, 405
- GRASS 51
- Graticule 428, 437
- Gray scale 376
- Gray values 304
- Great circle 434
- GRID 51, 100, 101
- GRID (Surfer ASCII/Binary) 51
- Grid cell 1, 112
- Grid line 428
- GRID Stack 51
- GRID Stack7x 51
- GRID Stacks 101
- Ground Control Point
 - see GCP
- Ground coordinate system 270
- Ground space 269
- Ground truth 217, 221, 222, 236
- Ground truth data 96
- Ground-based photographs 262

H

- Halftone 455
- Hammer 511
- Hardcopy 451
- Hardware 111

- HD 323
 - Header
 - file 20, 22
 - record 20
 - HFA file 317
 - Hierarchical pyramid technique 322
 - High density (HD) 323
 - High Resolution Visible sensors (HRV) 296
 - Histogram 144, 232, 390, 392, 459, 460, 470
 - breakpoint 148
 - signature 236, 241
 - Histogram equalization
 - formula 151
 - Histogram match 153
 - Homomorphic filtering 195
 - Host workstation 111
 - Hotine 532
 - HRPT (High Resolution Picture Transmission) 66
 - HRV 296
 - Hue 170
 - Huffman encoding 101
 - Hydrology 74
 - Hyperspectral data 11
 - Hyperspectral image processing 137
- I**
- .img file 2, 24
 - Ideal window 192
 - IFOV (instantaneous field of view) 15
 - IGDS 56
 - IGES 45, 49, 56, 108
 - IHS to RGB 172
 - IKONOS 58
 - bands/frequencies 59
 - Image 1, 112, 137
 - airborne 49
 - complex 55
 - digital 43
 - microscopic 49
 - pseudo color 41
 - radar 49
 - raster 43
 - ungeoreferenced 37
 - Image algebra 175, 220
 - Image Catalog 27, 28
 - Image coordinate system 269
 - Image data 1
 - Image display 2, 111
 - Image file 1, 118, 376
 - statistics 459
 - Image Information 120, 128, 344, 345, 390
 - Image Interpreter 11, 18, 138, 147, 161, 162, 169, 210, 232, 377, 379, 382, 383, 390, 393, 394
 - functions 139
 - Image processing 1
 - Image pyramid 295
 - Image registration 334
 - Image scale (SI) 265, 305
 - Image space 269, 273
 - Image space coordinate system 270
 - IMAGINE Developers' Toolkit 395, 445, 477
 - IMAGINE Expert Classifier 253
 - IMAGINE OrthoRadar algorithm description 309
 - IMAGINE Radar Interpreter 197
 - IMAGINE StereoSAR DEM
 - Constrain
 - theory 320
 - Degrade
 - theory 319, 325
 - Despeckle
 - theory 318
 - Height
 - theory 325
 - Input
 - theory 316
 - Match
 - theory 320
 - Register

Index

J

- theory 319
- Rescale
 - theory 319
- Subset
 - theory 318
- IMAGINE StereoSAR DEM process flow 315
- Import
 - direct 49
 - generic 55
- Incidence angles 299
- Inclination 300
- Inclination angles 299
- Index 173, 396, 402
 - application 174
 - vegetation 11
- Inertial navigation system 96
- INFO 40, 107, 108, 109, 391
 - path 40
 - see also ArcInfo
- Information (vs. data) 386
- INS 281
- Intensity 170
- Interferometric model 329
- Interior orientation 272
 - SPOT 298
- International Dateline 437
- Interpretative photogrammetry 262
- Interrupted Goode Homolosine 513
- Interrupted Mollweide 514
- Interval
 - classes 367, 389
 - data 3
- IRS 59
- IRS-1C 59
- IRS-1C (EOSAT Fast Format C) 51
- IRS-1C (EUROMAP Fast Format C) 51
- IRS-1D 60

J

- Jeffries-Matusita distance 239
- JERS-1 74, 76
 - bands/frequencies 76
 - ordering 99
- JFIF (JPEG) 51, 101
- JPEG File Interchange Format 101

K

- Kappa 277
- Kappa coefficient 260
- Knowledge Classifier 255
- Knowledge Engineer 253
- Kurtosis 209

L

- .LAN file 24, 100
- Laborde Oblique Mercator 582
- LAC (Local Area Coverage) 66
- Lambert Azimuthal Equal Area 516
- Lambert Conformal 553
- Lambert Conformal Conic 518, 580, 583
- Lambertian reflectance model 383
- LAN (Erdas 7.x) 51
- Landsat 10, 17, 25, 43, 51, 57, 137, 143, 162, 550
 - description 57, 60
 - history 60
 - MSS 15, 61, 141, 143, 175
 - ordering 97
 - TM 9, 10, 14, 20, 61, 160, 172, 175, 199, 215, 345, 347
 - displaying 118
- Landsat 7 64
 - characteristics 64
 - data types 64
- Laplacian operator 206, 207
- Latitude/Longitude 92, 344, 437, 508
 - rectifying 366
- Layer 2, 388, 405
- Least squares adjustment 284

Least squares condition 284
Least squares correlation 293
Least squares regression 350, 355
Lee filter 197, 200
Lee-Sigma filter 197
Legend 427
Lens distortion 274
Level 1B data 350
Level slice 152
Light SAR 79
Line 36, 41, 55, 107
Line detection 204
Line dropout 18, 142
Linear regression 142
Linear transformation 350, 355
Lines of constant range 211
LISS-III 59
 bands/frequencies 59
Local region filter 197, 199
Long wave infrared region 5
Lookup table 113, 144
 display 148
Low parallax (LP) 323
Lowtran 9, 143
Loximuthal 521
LP 323

M

.mdl file 412
Magnification 112, 132, 133
Magnitude of parallax 325
Mahalanobis distance 256
Map 419
 accuracy 443, 448
 aspect 420
 base 420
 bathymetric 420
 book 451
 cadastral 420
 choropleth 420

 colors in 422
 composite 420
 composition 447
 contour 420
 credit 430
 derivative 420
 hardcopy 449
 index 420
 inset 420
 isarithmic 420
 isopleth 420
 label 429
 land cover 217
 lettering 431
 morphometric 420
 outline 420
 output to TIFF 454
 paneled 451
 planimetric 69, 420
 printing 451
 continuous tone 455
 with black ink 457
 qualitative 421
 quantitative 421
 relief 420
 scale 443, 452
 scaled 451, 455
 shaded relief 420
 slope 420
 thematic 421
 title 430
 topographic 69, 421
 typography 430
 viewshed 421
Map Composer 419, 453
Map coordinate 3, 4, 346
 conversion 373
Map projection 343, 345, 432, 477
 azimuthal 433, 435, 443
 compromise 433

Index

M

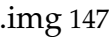
- conformal 443
- conical 433, 435, 443
- cylindrical 433, 436, 443
- equal area 443
- external 439, 578
- gnomonic 435
- modified 437
- orthographic 435
- planar 433, 435
- pseudo 437
- see also specific projection
- selecting 443
- stereographic 435
- types 434
- units 440
- USGS 438, 477
- MapBase
 - see ETAK
- Mapping 419
- Mask 398
- Matrix 473
 - analysis 396, 404
 - contingency 236, 237, 238
 - covariance 167, 235, 240, 251, 467
 - error 259
 - transformation 346, 348, 365, 472
- Matrix algebra
 - and transformation matrix 473
 - multiplication 473
 - notation 473
 - transposition 474
- Maximum likelihood
 - classification decision rule 238
- Mean 120, 148, 236, 462, 464, 466
 - of ISODATA clusters 228
 - vector 240, 468
- Mean Euclidean distance 209
- Mean filter 197, 198
- Measurement 44
- Measurement vector 467, 469, 473
- Median filter 197, 198, 199
- Megabyte 18
- Mercator 437, 523, 526, 532, 566, 572, 581
- Meridian 437
- Metric photogrammetry 262
- Microscopic imagery 49
- Microsoft Windows NT 111, 117
- MIF/MID (MapInfo) to Coverage 56
- Miller Cylindrical 526
- Minimum distance
 - classification decision rule 231, 238
- Minimum Error Conformal 582
- Minnaert constant 384
- Model 405, 406
- Model Maker 138, 394, 405
 - criteria function 411
 - functions 408
 - object 409
 - data type 410
 - matrix 409
 - raster 409
 - scalar 410
 - table 410
 - working window 411
- Modeling 31, 404
 - and image processing 405
 - and terrain analysis 375
 - using conditional statements 411
- Modified Polyconic 583
- Modified Stereographic 584
- Modified Transverse Mercator 528
- Modtran 9, 143
- Mollweide 529
- Mollweide Equal Area 585
- Mosaic 29, 30, 345
- MrSID 51, 102
- MSS Landsat 51
- Multiplicative algorithm 161
- Multispectral imagery 57, 68
- Multitemporal imagery 29

N

Nadir 57, 299
 NASA 60, 71, 79
 NASA ER-2 80
 NASA/JPL 79
 Natural-color 118
 NAVSTAR 94
 Nearest neighbor
 see resample
 Neatline 427
 Negative inclination 299
 Neighborhood analysis 395, 398
 boundary 399
 density 399
 diversity 399
 majority 399
 maximum 399
 mean 399
 median 399
 minimum 399
 minority 400
 rank 400
 standard deviation 400
 sum 400
 New Zealand Map Grid 530
 NITFS 53
 NLAPS 51, 65
 NOAA 65
 Node 36
 dangling 418
 from-node 36
 pseudo 418
 to-node 36
 Noise reduction 336
 Noise removal 194
 Nominal
 classes 367, 389
 data 3
 Non-Lambertian reflectance model 383, 384
 Nonlinear transformation 352, 356, 358

Normal distribution 164, 249, 250, 251, 463, 466,
 470
 Normalized correlation coefficient 324
 Normalized Difference Vegetation Index
 (NDVI) 11, 175, 176

O

.OV1 (overview image) 91
 .OVR (overview image) 91
 .ovr file 423
 Oblated Equal Area 531
 Oblique Mercator 532, 553, 582, 586
 Oblique photographs 262
 Observation equations 282
 Oceanography 74
 Off-nadir 68, 299
 Offset 350
 Oil exploration 74
 Omega 277
 Opacity 132, 392
 Optical disk 21
 Orb View-3
 bands/frequencies 67
 Orbit correction 317
 OrbView3 67
 Order
 of polynomial 471
 of transformation 472
 Ordinal
 classes 367, 389
 data 3
 Orientation 277
 Orientation angle 301
 Orthorectification 94, 143, 212, 344
 Orthographic 535
 Orthorectification 303, 313, 344
 Output file 23, 365, 366, 412
  147
 classification 260
 Output formation 313

Index

P

Overlay 396, 401
Overlay file 423
Overview images 91

P

Panchromatic
 band/frequency 60
Panchromatic imagery 57, 68
Panchromatic sensor 60
Parallax 324
Parallel 437
Parallelepiped
 alarm 236
Parameter 466
Parametric 250, 251, 466
Passive sensor 6
Pattern recognition 217, 236
PCX 51
Periodic noise removal 194
Phase 336, 337, 338, 340
Phase function 339
Phi 277
Photogrammetric configuration 282
Photogrammetric scanners 81, 266
Photogrammetry 261
Photograph 49, 56, 162
 aerial 79
 ordering 98
Pixel 1, 112
 depth 111
 display 112
 file vs. display 112
Pixel coordinate system 269, 273
Plane table photogrammetry 261
Plate Carrée 503, 537, 581, 588
Point 36, 41, 55, 107
 label 36
Point ID 347
Polar Stereographic 539, 583
Pollution monitoring 74
Polyconic 528, 542, 583
Polygon 36, 41, 55, 107, 179, 222, 398, 415
Polynomial 353, 471
Positive inclination 299
Posting 324
PostScript 453
Preference Editor 127
Prewitt kernel 205, 206
Primary color 112, 456
 RGB vs. CMY 456
Principal component band 165
Principal components 30, 160, 161, 164, 168,
 220, 467, 470, 471
 computing 166
Principal point 273
Printer
 PostScript 453
 Tektronix Inkjet 455
 Tektronix Phaser 455
 Tektronix Phaser II SD 456
Probability 241, 250
Processing a strip of images 288
Processing one image 288
Profile 92
Projection
 external
 Bipolar Oblique Conic Conformal
 580
 Cassini-Soldner 581
 Laborde Oblique Mercator 582
 Modified Polyconic 583
 Modified Stereographic 584
 Mollweide Equal Area 585
 Rectified Skew Orthomorphic 586
 Robinson Pseudocylindrical 586
 Southern Orientated Gauss Confor-
 mal 587
 Winkel's Tripel 587
 perspective 563
USGS

- Alaska Conformal 480
 - Albers Conical Equal Area 481
 - Azimuthal Equidistant 483
 - Bonne 487
 - Cassini 489
 - Conic Equidistant 500
 - Eckert IV 495
 - Eckert VI 498
 - Equiarectangular (Plate Carrée) 503
 - Gall Stereographic
 - Gall Stereographic 505
 - General Vertical Nearside Perspective 507
 - Geographic (Lat/Lon) 508
 - Gnomonic 510
 - Hammer 511
 - Lambert Azimuthal Equal Area 515
 - Lambert Conformal Conic 518
 - Mercator 523
 - Miller Cylindrical 525
 - Modified Transverse Mercator 528
 - Mollweide 529
 - Oblique Mercator (Hotine) 532
 - Orthographic 535
 - Polar Stereographic 539
 - Polyconic 542
 - Sinusoidal 548
 - Space Oblique Mercator 550
 - State Plane 553
 - Stereographic 563
 - Transverse Mercator 566
 - Two Point Equidistant 567
 - UTM 569
 - Van der Grinten I 571
 - Projection Chooser 438
 - Proximity analysis 395
 - Pseudo color 63
 - display 389
 - Pseudo color image 41
 - Pushbroom scanner 68
 - Pyramid layer 27, 127
 - Pyramid technique 322
 - Pythagorean Theorem 471
- Q**
- Quartic Authalic 544
 - Quick tests 325
- R**
- Radar 4, 18, 162, 206, 207, 208
 - Radar imagery 49
 - RADARSAT 52, 74, 77
 - beam mode resolution 77
 - ordering 99
 - RADARSAT (Acres CEOS) 52
 - RADARSAT (Vancouver CEOS) 51
 - RADARSAT Beam mode 325
 - Radial lens distortion 274, 275
 - Radiative transfer equation 9
 - RAM 126
 - Range line 211
 - Range sphere 312
 - Raster
 - data 3
 - Raster Attribute Editor 124, 392
 - Raster editing 30
 - Raster image 43
 - Raster Product Format 52, 90
 - Raster region 397
 - Ratio
 - classes 367, 389
 - data 3
 - Rayleigh scattering 8
 - Real time differential GPS 95
 - Recode 30, 395, 400, 405
 - Record 21, 391
 - logical 22
 - physical 22
 - Rectification 28, 343
 - process 346

Index

S

- Rectified Skew Orthomorphic 586
 - Reduction 133
 - Reference coordinate 346
 - Reference pixel 259
 - Reference plane 271
 - Refined ephemeris 318
 - Reflect 350, 351
 - Reflection spectra 6
 - see absorption spectra
 - Registration 343
 - vs. rectification 346
 - Relation based matching 295
 - Remote sensing 343
 - Report
 - generate 393
 - Resample 343, 346, 365
 - Bilinear Interpolation 126, 366, 367
 - Cubic Convolution 126, 366, 370
 - for display 126
 - Nearest Neighbor 126, 366, 370
 - Residuals 362
 - Resolution 14, 267
 - display 111
 - merge 160
 - radiometric 14, 16, 17, 61
 - spatial 15, 17, 153, 452
 - spectral 14, 17
 - temporal 14, 17
 - Resolution merge
 - Brovey Transform 161
 - Multiplicative 161
 - Principal Components Merge 161
 - Retransformed coordinate 367
 - RGB monitor 112
 - RGB to IHS 214
 - Rhumb line 433
 - Right hand rule 270
 - RMS error 274, 348, 350, 362, 364
 - tolerance 364
 - total 363
 - RMSE 266
 - Roam 133
 - Robinson 545
 - Robinson Pseudocylindrical 586
 - Root mean square error 274
 - Root Mean Square Error (RMSE) 266
 - Rotate 350
 - Rotation matrix 277
 - RPF 90
 - RPF Cell 91
 - RPF Frame 91
 - RPF Overview 91
 - RPF Product 91
 - RSO 547
 - Rubber sheeting 352
- S**
- .sig file 466
 - .stk (GRID Stack file) 101
 - Sanson-Flamsteed 548
 - SAR 71
 - SAR image intersection 316
 - SAR imaging model 311
 - Satellite 56
 - imagery 4
 - system 56
 - Satellite photogrammetry 296
 - Satellite scene 297
 - Saturation 170
 - Scale 15, 350, 351, 423, 451
 - display 452
 - equivalents 424
 - large 15
 - map 452
 - paper 452
 - determining 454
 - pixels per inch 425
 - representative fraction 423
 - small 15
 - verbal statement 423

- Scale bar 424
- Scaled map 455
- Scan line 296
- Scanner 57
- Scanners 81
- Scanning 81, 345
- Scanning resolutions 267, 268
- Scanning window 398
- Scattering 7
 - Rayleigh 8
- Scatterplot 164, 232, 470
 - feature space 224
- SCBA 303
- Screendump command 102
- Script model 138, 394
 - data type 414
 - library 412
 - statement 413
- Script modeling 396
- SDE 52
- SDTS 45, 102
- SDTS (raster) 52
- SDTS (vector) 56
- SDTS profiles 102
- SDTS Raster Profile and Extensions 102
- Search area 320, 324
- Seasat-1 70
- Seat 111
- SeaWiFS 67
 - bands/frequencies 68
- SeaWiFS L1B and L2A 52
- Secant 435
- Seed properties 223
- Self-calibrating bundle adjustment (SCBA)
 - 303
- Sensor 57
 - active 6, 72
 - passive 6, 72
 - radar 74
- Separability
 - listing 240
 - signature 238
- Shaded relief 375, 381
 - calculating 382
- Shadow
 - enhancing 146
- Shapefile 52
- Ship monitoring 74
- Short wave infrared region 5
- SI 305
- Sigma filter 200
- Sigma notation 459
- Signal based matching 292
- Signature 218, 221, 224, 234
 - alarm 236
 - append 242
 - contingency matrix 238
 - delete 241
 - divergence 236, 238
 - ellipse 236
 - evaluating 235
 - file 466
 - histogram 236, 241
 - manipulating 224, 241
 - merge 242
 - non-parametric 218, 225, 235
 - parametric 218, 235
 - separability 238, 240
 - statistics 236, 241
 - transformed divergence 239
- Simple Conic 500
- Simple Cylindrical 503
- Single frame orthorectification 263
- Sinusoidal 437, 548
- SIR-A 73, 74, 77
 - ordering 99
- SIR-B 73, 74, 78
 - ordering 99
- SIR-C 74, 78
 - ordering 99

Index

S

- SIR-C/X-SAR 78
 - bands/frequencies 79
- Skew 350
- Skewness 209
- Slant range 311
- Slant-to-ground range correction 212
- SLAR 71, 73
- Slope 375, 376
 - calculating 377
- Softcopy photogrammetry 261
- Source coordinate 346
- Southern Orientated Gauss Conformal 587
- Space forward intersection 280
- Space Oblique Mercator 437, 532, 550
- Space Oblique Mercator (Formats A & B) 552
- Space resection 263, 279
- Sparse mapping grid 313
- Spatial frequency 154
- Spatial Modeler 18, 55, 138, 384, 390, 405
- Spatial Modeler Language 138, 394, 405, 412
- Speckle noise 73, 197
 - removing 197
- Speckle suppression 18
 - local region filter 199
 - mean filter 198
 - median filter 198, 199
 - Sigma filter 200
- Spectral dimensionality 467
- Spectral distance 238, 247, 256, 258, 471
 - in ISODATA clustering 228
- Spectral space 164, 166, 470
- Spectroscopy 5
- Spheroid 444
- SPOT 10, 15, 17, 25, 29, 43, 52, 57, 68, 88, 109, 137, 143, 162, 347
 - ordering 97
 - panchromatic 14, 160, 172
 - XS 68, 175
 - displaying 119
- SPOT (GeoSpot) 52
- SPOT 4
 - bands/frequencies 70
- SPOT CCRS 52
- SPOT exterior orientation 299
- SPOT interior orientation 298
- SPOT SICORP MetroView 52
- SPOT4 70
- SRPE 102
- Standard Beam Mode 323
- Standard deviation 120, 148, 236, 464
 - sample 465
- Standard meridian 433, 436
- Standard parallel 433, 435
- State Plane 440, 444, 518, 553, 567
- Statistics 27, 459
 - signature 236
- Step size 324
- Stereographic 539, 563
- Stereographic (Extended) 565
- Stereoscopic imagery 69
- Striping 18, 166, 199
- Subset 29
- Summation 459
- Sun angle 381
- SUN Raster 52
- Sun Raster 100, 102
- Surface generation
 - weighting function 32
- Swath width 57
- Swiss Cylindrical 587
- Symbol 428
 - abstract 428
 - function 429
 - plan 429
 - profile 429
 - replicative 428
- Symbolization 35
- Symbology 41
- Symmetric lens distortion 275

T

Tangent 435
Tangential lens distortion 274
Tape 18
Tasseled Cap transformation 169, 175
Tektronix
 Inkjet Printer 455
 Phaser II SD 456
 Phaser Printer 455
Template 320
Template size 322
Terramodel 56
Terrestrial photography 262, 271
Texture analysis 208
Thematic
 data 3
Thematic data
 see data
Theme 388
Threshold 257, 324
Thresholding 256
Thresholding (classification) 249
Tick mark 427
Tie point distribution 290
Tie points 264, 290
TIFF 52, 81, 100, 103, 453, 454
TIGER 45, 49, 56, 109
 disk space requirement 110
Time 311
TM Landsat Acres Fast Format 52
TM Landsat Acres Standard Format 52
TM Landsat EOSAT Fast Format 52
TM Landsat EOSAT Standard Format 52
TM Landsat ESA Fast Format 52
TM Landsat ESA Standard Format 52
TM Landsat IRS Fast Format 52
TM Landsat IRS Standard Format 52
TM Landsat Radarsat Fast Format 53
TM Landsat Radarsat Standard Format 53
TM Landsat-7 Fast-L7A EROS 52

Topocentric coordinate system 271
Topographic effect 382
Topological Vector Profile 102
Topology 37, 416
 constructing 416
Total field of view 57
Total RMS error 363
Training 217
 supervised 217, 220
 supervised vs. unsupervised 220
 unsupervised 218, 220, 227
Training field 222
Training sample 221, 224, 259, 345, 470
 defining 222
 evaluating 224
Training site
 see training field
Transformation
 1st-order 350
 linear 350, 355
 matrix 349, 353
 nonlinear 352, 356, 358
 order 349
Transformation matrix 346, 348, 349, 353, 365,
 472
Transposition 251, 474
Transposition function 239, 249, 251
Transverse Mercator 528, 553, 566, 569, 581, 587
True color 63
True direction 433
Two Point Equidistant 567
Type style
 on maps 430

U
Ungeoreferenced image 37
Universal Polar Stereographic 539
Universal Transverse Mercator 528, 567, 569
Unwrapped phase 340
Unwrapping 338

Index

V

USGS DEM 50
UTM
 see Universal Transverse Mercator

V

V residual matrix 286
Van der Grinten I 572, 585
Variable 389
 in models 415
Variable rate technology 96
Variance 209, 249, 464, 465, 466, 467
 sample 464
Vector 473
Vector data
 see data
Vector layer 37
Vector Quantization 91
Vegetation index 11, 174
Velocity vector 301
Vertex 36
Viewer 126, 127, 138, 396
 dithering 130
 linking 132
Volume 23
 set 23
VPF 45, 56
VQ 91

W

Wagner IV 574
Wagner VII 575
Wavelet 101
Weight factor 403
 classification
 separability 240
Weighting function (surfacing) 32
Wide Field Sensor 60
WiFS 60
 bands/frequencies 60
Winkel I 577

Winkel II 578
Winkel's Tripel 587
Workspace 38

X

X matrix 285
X residual 362
X RMS error 363
X Window 111
XSCAN 81

Y

Y residual 362
Y RMS error 363

Z

Zero-sum filter 158, 207
Zone 83, 90
Zone distribution rectangle (ZDR) 83
Zoom 132, 133