

Remote Sensing

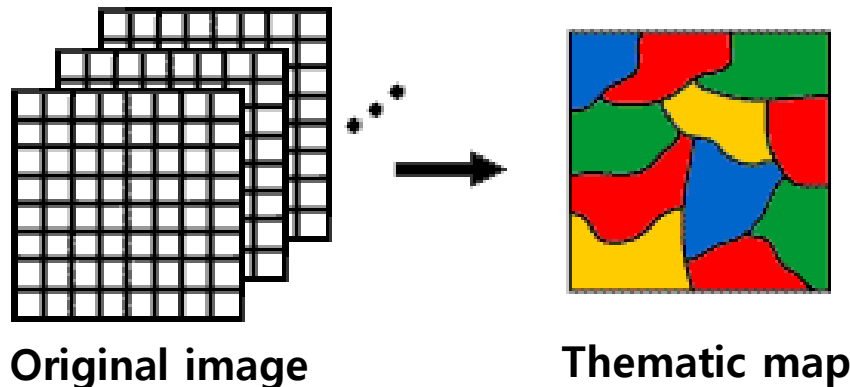
Ch. 4 이미지 해석(Interpretation) 및 분석(Analysis) (Part 3 of 3)

4.7 이미지 분류(Classification) 및 분석(Analysis)

4.8 데이터 통합(Integration) 및 분석(Analysis)

4.7 Image Classification and Analysis

- **분석가(human analyst)**들은 이미지의 대상들을 분류할 때, 다양한 속성 또는 관심 있는 토지 피복 클래스를 나타내는 동질의 픽셀 그룹을 식별하기 위해 **시각적 해석의 요소를** 사용합니다.
- **디지털 이미지 분류기(Digital image classification)**는 하나 이상의 스펙트럼 대역의 디지털 숫자로 표시되는 스펙트럼 정보를 사용하며, 이러한 스펙트럼 정보에 기초하여 **각각의 픽셀에 대한 분류**를 한다. 이러한 유형의 분류를 **스펙트럼 패턴 인식(spectral pattern recognition)**이라고 한다.
- 두 경우 모두, **목표는 특정 클래스 또는 테마** (예를 들어 물, 침엽수림(coniferous forest), 낙엽수림(deciduous forest), 옥수수, 밀, 등)로 이미지의 **모든 픽셀을 할당**하는 것입니다. 생성된 분류된 이미지는 **특정 주제에 속하는 각각의 픽셀의 모자이크**로 구성되고, 본질적으로 **원래의 이미지의 주제도(a thematic "map")**가 된다.

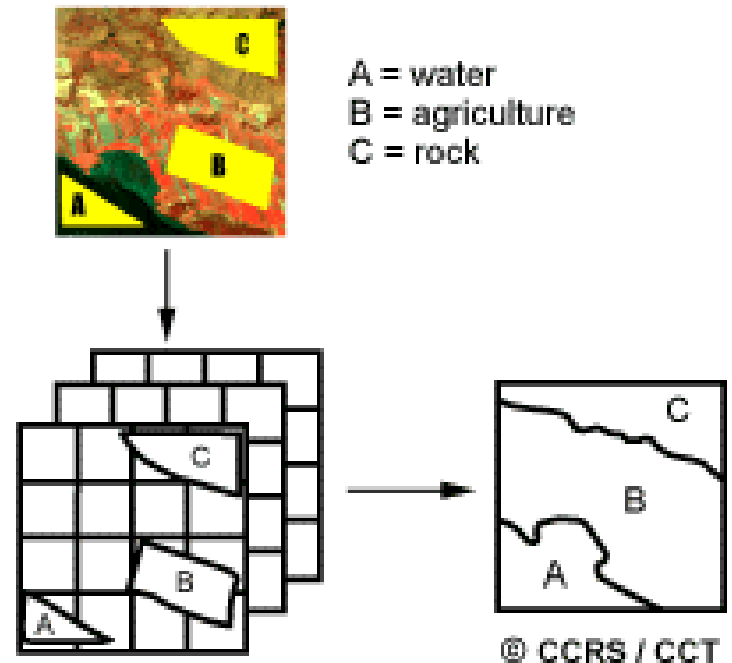


4.7 Image Classification and Analysis

- 클래스에 대해 이야기 할 때, **정보 클래스(information classes)** 와 **스펙트럼 클래스(spectral classes)**를 구분해야 합니다.
 - ✓ **정보 클래스(Information classes)** → 작물의 종류, 숲의 종류 나 수종, 지질학적 단위 또는 암석 종류, 등과 같은 분석가가 실제로 이미지에서 식별하고자 하는 **관심 의 범주(categories of interest)**.
 - ✓ **스펙트럼 클래스(Spectral classes)** → 데이터의 각각의 스펙트럼 채널 안에서의 밝기 값이 **균일한(또는 거의 유사한) 픽셀 그룹 (group of pixels that are uniform (or near-similar))**
- **분류의 목적**은 데이터의 스펙트럼 클래스를 관심이 있는 **정보 클래스에 매치하는 것**이다. 이러한 클래스의 두 가지 유형 간에 **간단한 일대일 매치는 흔하지 않다**.
 - ✓ 오히려 **독특한 스펙트럼 클래스**가 나타날 수 있으나, 이것이 반드시 분석가에 특정 사용 또는 관심의 **정보 클래스와 일치하지는 않습니다**.
 - ✓ 오히려, **폭 넓은 정보 클래스** (예를 들면 숲)가 독특한 스펙트럼 변화를 갖는 **다수의 스펙트럼 서브 클래스를 포함** 할 수 있다. 숲의 예를 들면, 스펙트럼 서브 클래스로는 연령, 종 및 밀도의 변화이거나, 그림자나 조도 변화의 결과에 기인 할 수 있다
- 서로 다른 스펙트럼 클래스의 **유용성** 및 대응되는 **유용한 정보 클래스에 대한 결정은 해석가의 직무(analyst's job)**이다.

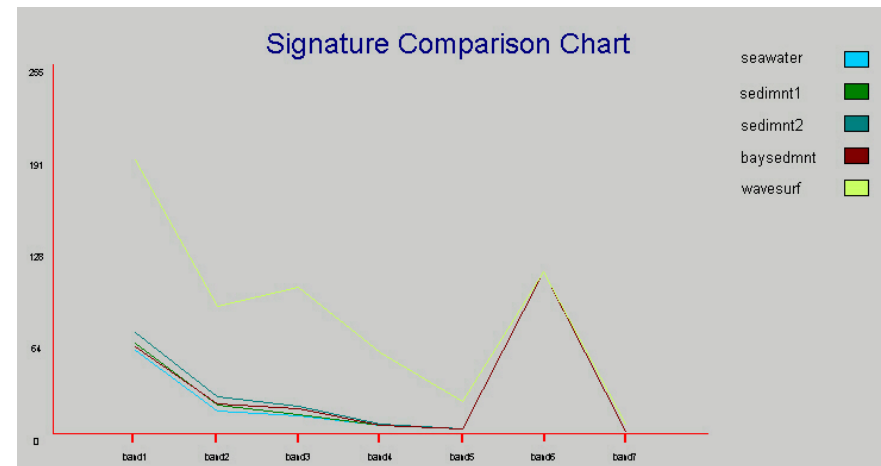
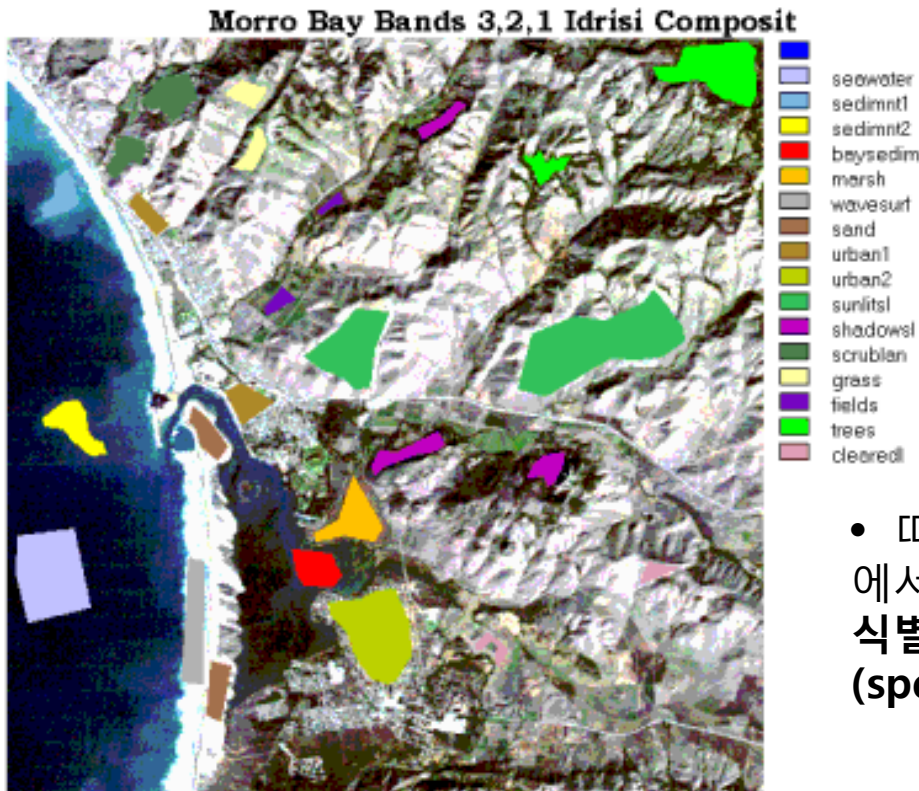
4.7 Image Classification and Analysis

- 일반적인 **분류 절차(classification procedures)**는 사용되는 방법에 따라 두 가지로 분류 될 수 있다 :
 - (1) **감독 분류(supervised classification)**
 - (2) **무감독 분류(unsupervised classification)**.
- **감독 분류(supervised classification)**에서 **분석가**는 이미지에서 관심 있는 표면 커버 타입들(**정보 클래스(information classes)**)의 **대표적인 균일한 샘플들을 선정한다**. 이 샘플들은 **훈련영역(training areas)**이라고 합니다.
- 적절한 훈련영역의 선택은 **분석가의** 이미지에 존재하는 실제 **표면 커버 타입에 대한 지식**과 **지리적 지역에 대한 친숙함**을 기반으로 합니다. 따라서, **분석가(analyst)**는 **특정 클래스 집합의 항목화(categorization)**를 "**감독(supervising)**"합니다.
- 이러한 영역을 구성하는 픽셀의 모든 **스펙트럼 밴드들에서의 숫자 정보**는 각 클래스에 대한 **스펙트럼적으로 유사한 영역을 인식** 할 수 있도록 **컴퓨터를 훈련(train)**하는데 사용됩니다.



4.7 Image Classification and Analysis

- 컴퓨터는 각 훈련 클래스에 대한 수치적인 "**특징(signatures)**"을 결정하기 위해 특별한 프로그램이나 알고리즘을 사용합니다. 컴퓨터가 각각의 클래스에 대한 특징을 결정하면, 이미지의 각 화소는 이러한 특징과 비교하고 디지털적으로 가장 근접하게 "**유사한**" 클래스로 분류된다



- 따라서, 감독 분류(supervised classification)에서는 먼저 정보 클래스(information classes)를 식별한 다음, 그들을 나타내는 스펙트럼 클래스(spectral classe)를 결정하기 위하여 사용된다.

정보 클래스(Information classes)

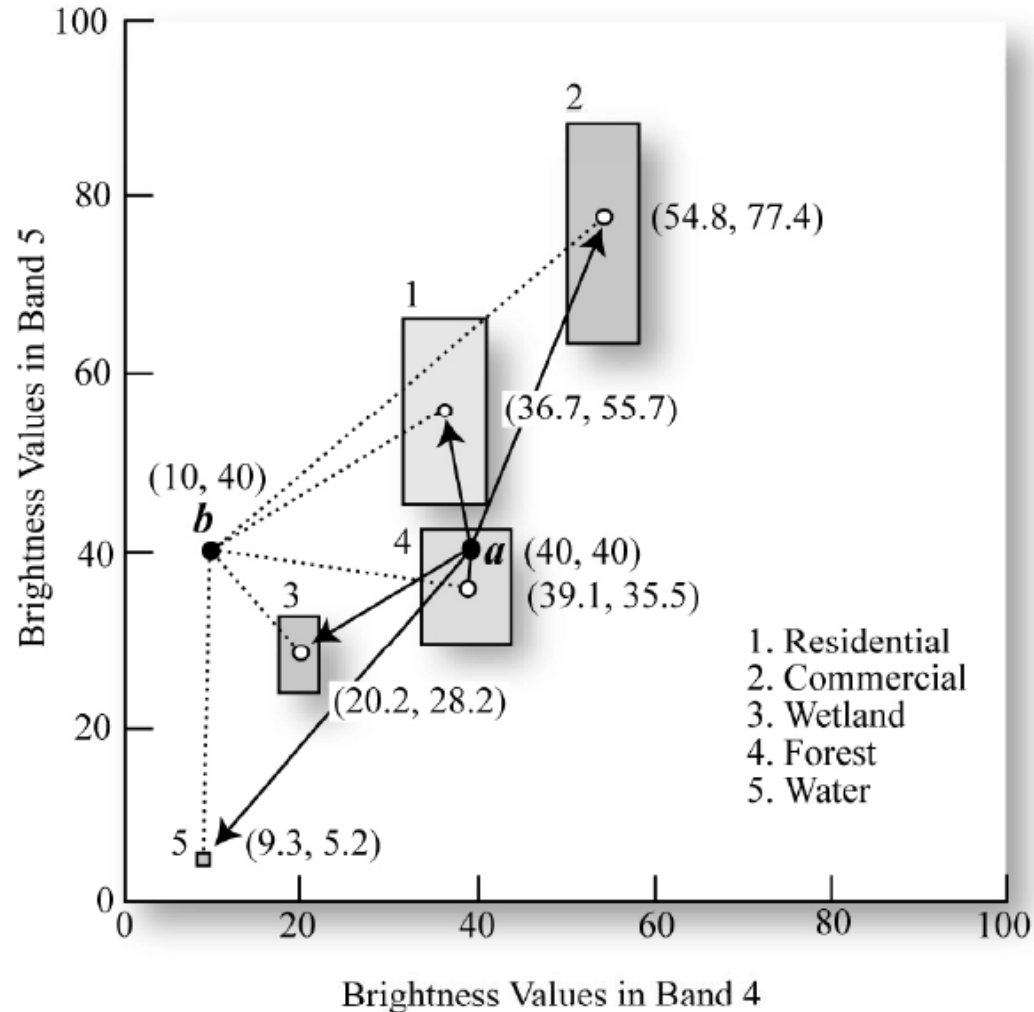


스펙트럼 클래스(Spectral classes)

4.7 Image Classification and Analysis

평행육면체(Parallelepiped) 분류 알고리즘

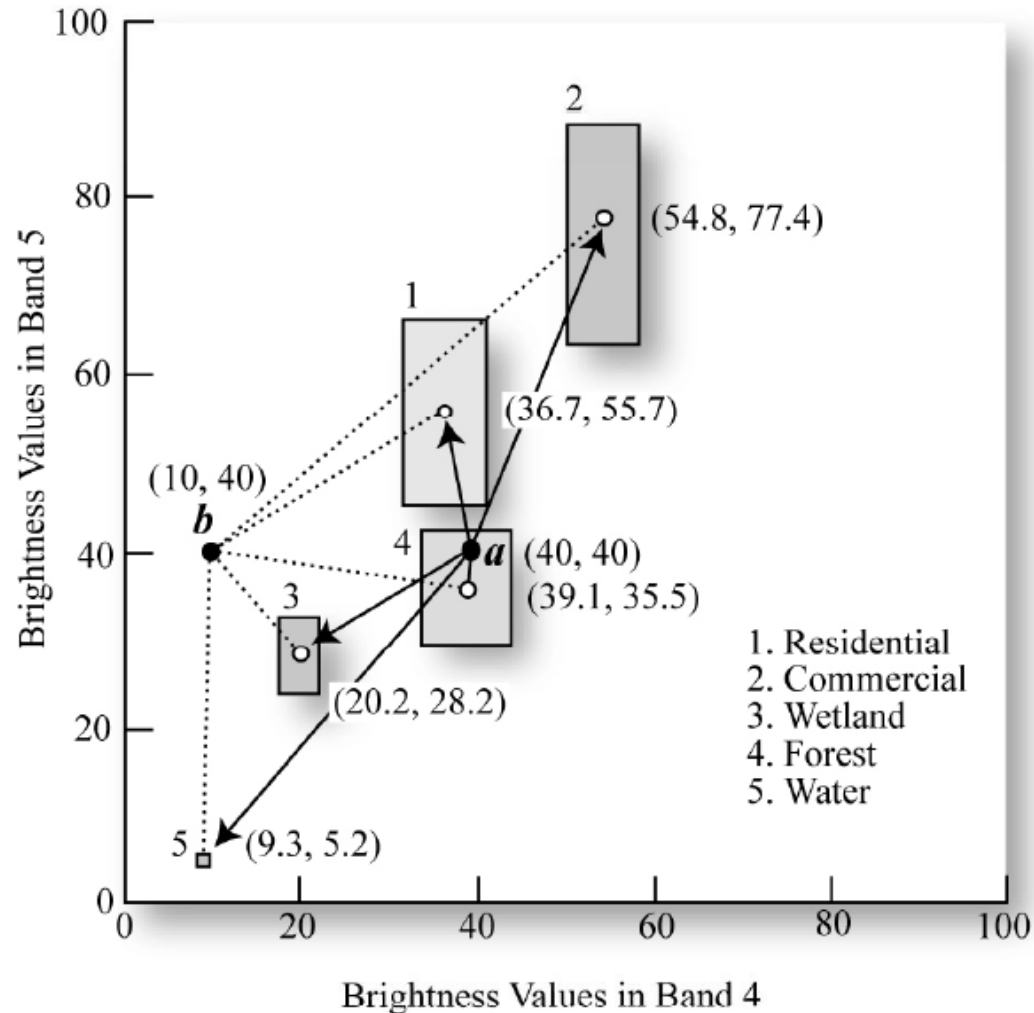
- 각 클래스에 대한 **최소 및 최대 DN**s를 결정하고, 이미지를 **분류**하기 위한 **임계값**으로 사용.
- **장점** : 훈련과 사용이 간단하며, 빠른 연산.
- **단점** : 평행육면체들 사이의 간격에 놓인 픽셀들이 분류될 수 없다. 평행육면체가 겹치는 영역의 픽셀들이 분류될 수 없다.



4.7 Image Classification and Analysis

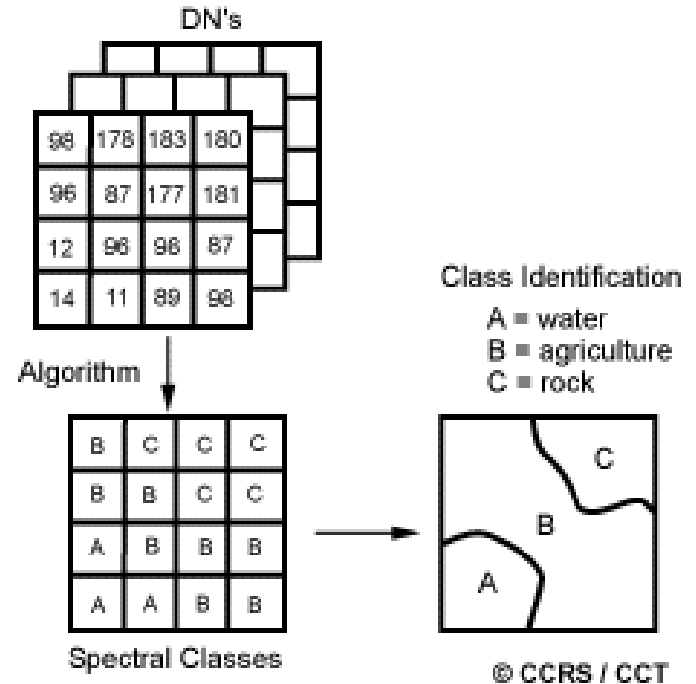
최소거리(Minimum Distance) 분류 알고리즘

- 각 클래스의 **무게중심(centroid)**이 각 클래스에 대한 밴드의 **평균값**을 계산하여 결정한다.
- 각각의 이미지 픽셀에 대하여, 각 **무게중심까지의 N차원 거리**를 계산하고 **가장 가까운 무게중심**을 갖는 클래스로 결정한다.
- **장점** : 수학적으로 간단하고, 효율적인 계산
- **단점** : 스펙트럼 응답 자료의 분산의 다른 정도에 민감하지 않음.



4.7 Image Classification and Analysis

- 본질적으로 **무감독 분류(Unsupervised classification)**는 감독 분류(supervised classification) 과정의 반대이다. 우선 **데이터의 수치 정보에 기초하여 스펙트럼 클래스를 그룹화**하고, 해석가가 **정보 클래스로 매치**시킨다 (가능하다면).
- **클러스터링 알고리즘(clustering algorithms)**이라 불리는 프로그램이 데이터의 **자연적(통계적) 그룹핑** 또는 **구조를 결정**하는 데 사용된다. 보통, 분석가가 데이터 내에서 **얼마나 많은 그룹** 또는 클러스터를 찾는 지를 지정합니다. 원하는 클래스의 개수를 지정하는 것 외에도, 분석가는 **클러스터들 간의 이격 거리와 각 클러스터 내의 분산**에 관련된 **파라미터를 지정**할 수도 있다.



- 이러한 반복적 클러스터링 처리의 최종 결과는, 해석가의 입장에서는 **일부 클러스터들의 결합**, 또는 **더 세분화해야 할 클러스터들**을 초래할 수도 있다 - 이들 각각은 클러스터링 알고리즘의 또 다른 적용을 필요로 하게 된다. 따라서, 무감독 분류가 **인간의 완전한 개입이 없는 것은 아니다**. 그러나, 감독 분류에서와 같이 클래스가 미리 정해진 설정에서 시작되지는 않습니다.

4.7 Image Classification and Analysis

반복적 자기-조직화 자료분석 기술 (Iterative Self-Organizing Data Analysis

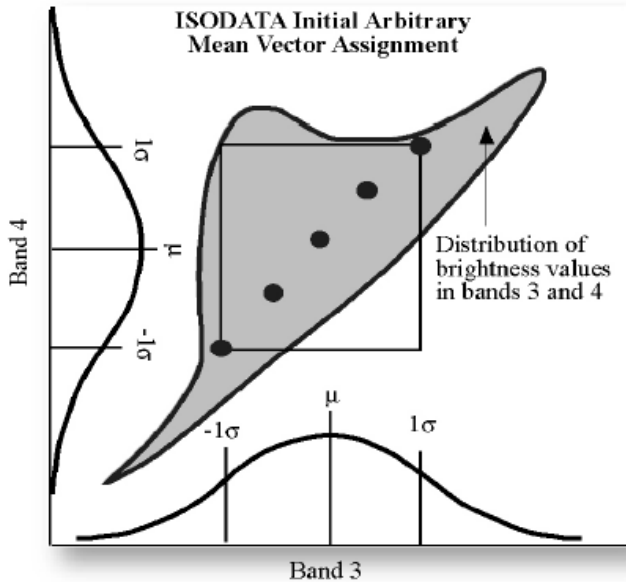
Technique : ISODATA) : 체험적(heuristic) 과정의 포괄적인 세트가 반복적 분류 알고리즘으로 통합 된 것.

ISODATA 알고리즘은 다음을 포함한다.

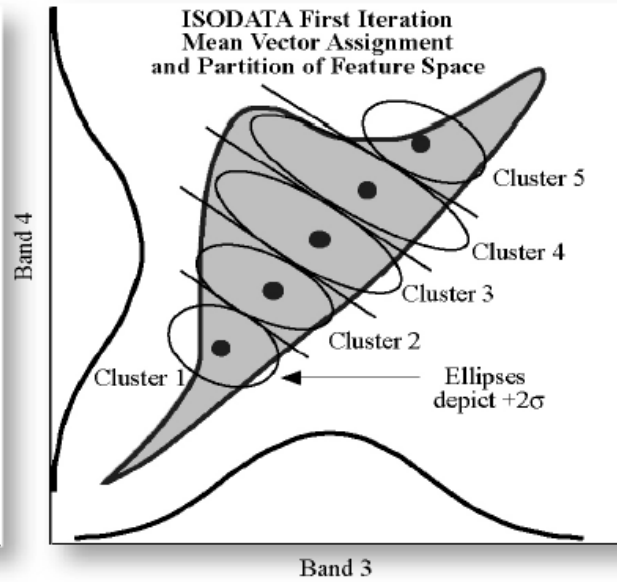
- a) 군집(cluster)들 간의 거리가 사용자가 명시한 임계값 보다 작으면 병합
- b) 하나의 군집(cluster)을 둘로 나누기 위한 규칙

ISODATA 는 사람의 입력이 비교적 작아 자기-조직화(self-organizing)이긴 하지만, 일반적으로 해석가로부터 다음과 같은 일련의 파라미터들의 명시가 요구된다 :

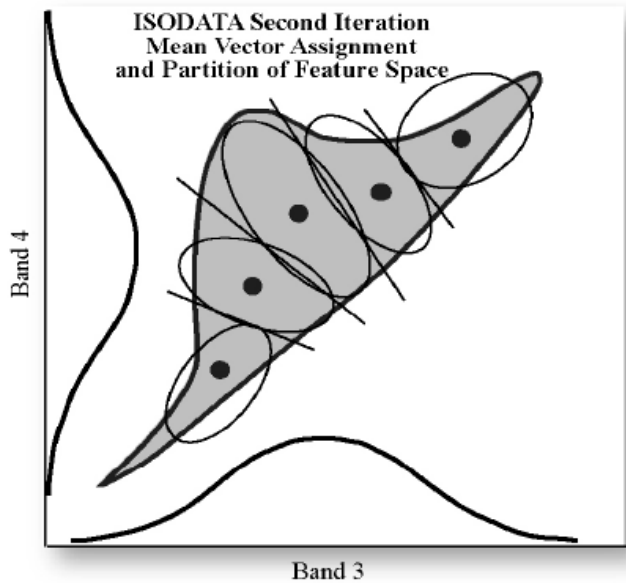
- $N_{\text{MIN_EX}}$: 군집당 최소 표본 수
- N_D : 대략적으로 요구되는 군집 개수
- σ : 분리가 요구되는 최대 분산 파라미터
- D_{MERGE} : 결합이 요구되는 최대 분리 거리
- N_{MERGE} : 결합될 수 있는 최대 군집 수



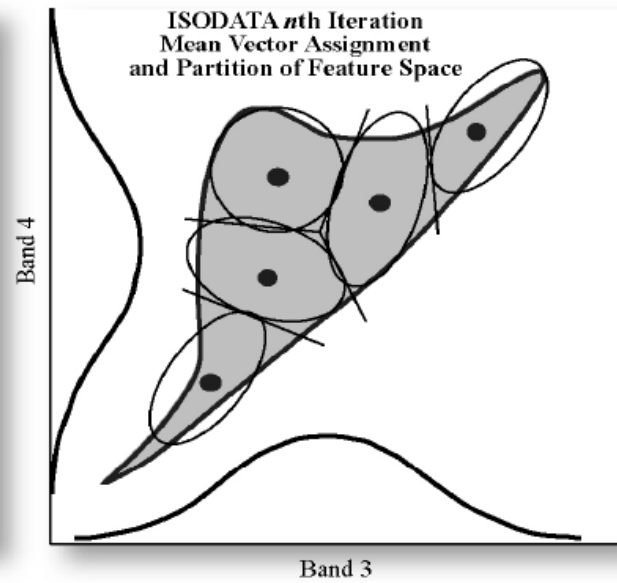
a.



b.



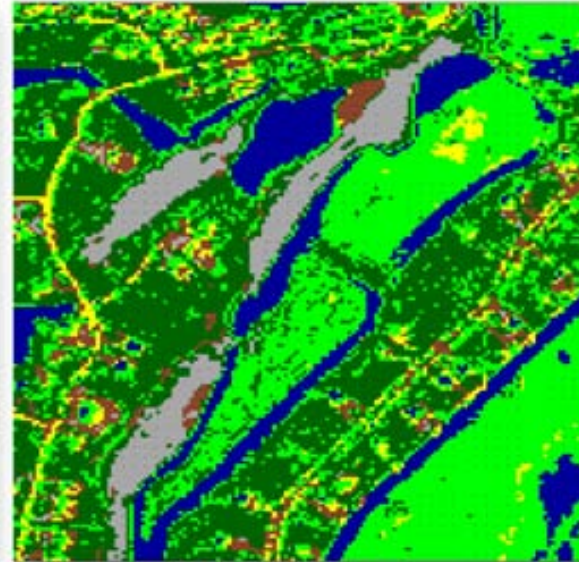
c.



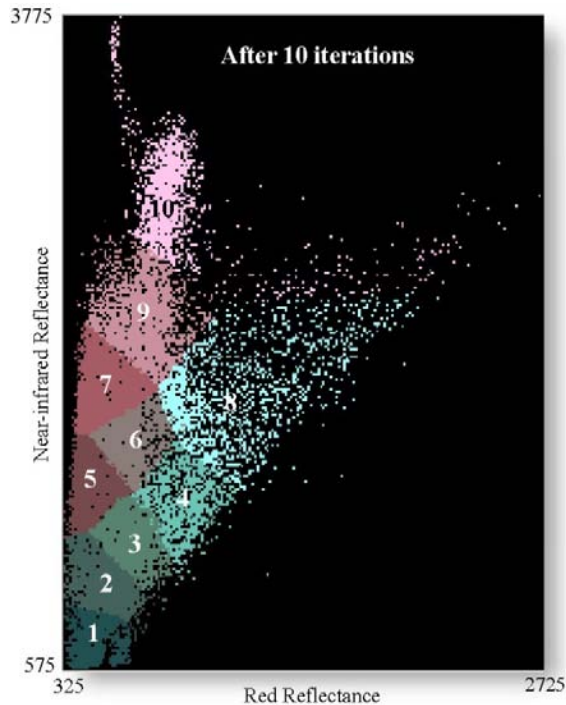
d.

Information Classes Derived from an ISODATA Unsupervised Classification Using 10 Iterations and 10 Mean Vectors of an Area Near North Inlet, SC

A color composite of multi-spectral data.



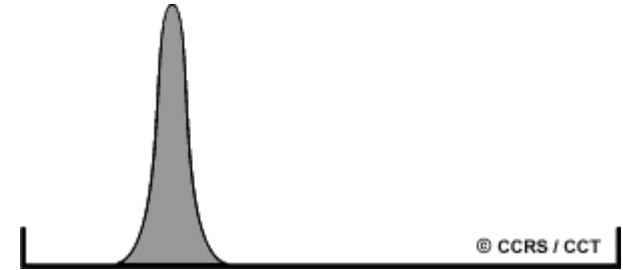
Classification map derived from 10 ISODATA clusters.



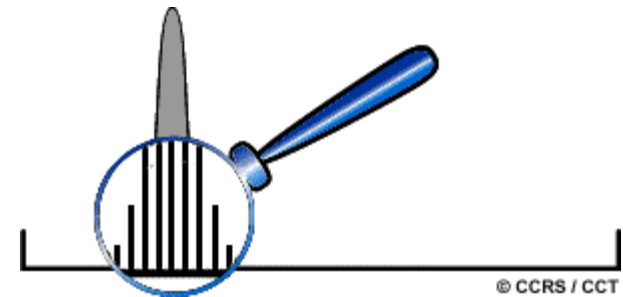
Class	Legend
1. Water	Blue
2. Wetland	Light Green
3. Wetland	Medium Green
4. Roof/asphalt	Yellow
5. Forest	Dark Green
6. Wetland	Light Green
7. Forest	Dark Green
8. Bare soil	Red
9. Forest	Dark Green
10. Fairway	Grey

4.7 Image Classification and Analysis

QUIZ 당신은 위성 영상에 대한 분류를 수행하려고 합니다. 그러나 그것의 히스토그램을 검사해보니, 유용한 데이터의 범위가 매우 좁은 것을 알 수 있었습니다. **분류를 시도하기 이전에**, 당신은 **선형 대비 스트레칭(linear contrast stretch)**으로 이미지 향상을 하겠습니까?



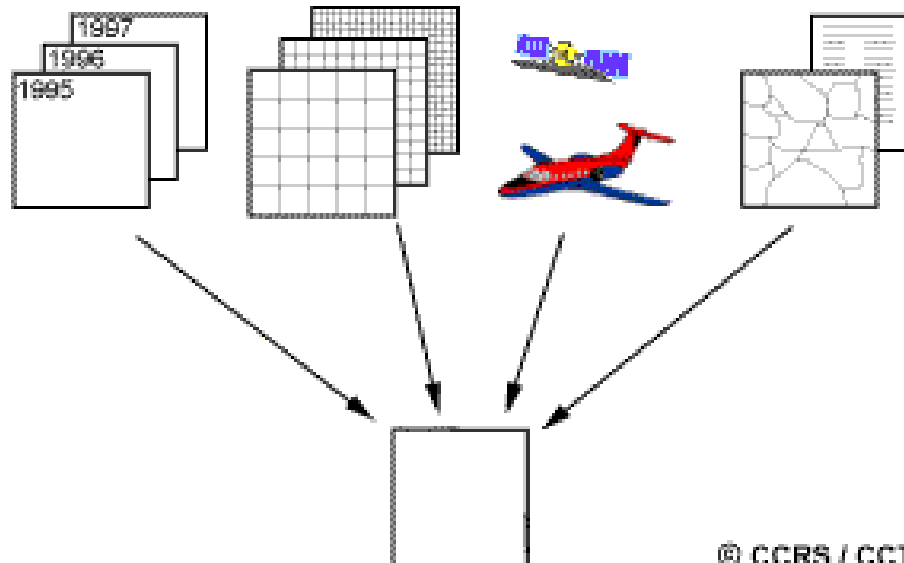
ANS 이미지의 '**향상(enhancement)**'은 오직 시각적으로 그 내용을 인식하고 분석하기 위해서 수행됩니다. 향상(enhancement)은 분류 알고리즘에 관한 한, 아무것도 유용한 정보를 추가하지 않습니다. 이를 다른 시각으로 살펴보면: 두 개의 픽셀이 하나의 디지털 단위(one digital unit)의 밝기 값의 차이가 있는 경우, 눈으로 그 미묘한 차이를 발견하는 것은 매우 어렵습니다. 그러나 컴퓨터에서 그 차이는 그냥 100 배 더 큰 것처럼 '**분명한**' 차이로 인식됩니다.



이미지의 향상된 버전은 '**훈련(training)**' 영역을 (눈을 사용한) 선택하는 데 도움이 될 수 있습니다, 하지만 당신은 여전히 향상되지 않은 버전에서 분류를 수행하게 됩니다.

4.8 자료 통합(Integration) 및 분석(Analysis)

- 원격 탐사 데이터 소스가 항공 사진이었던 **아날로그 원격 탐사의 초기에**, 서로 다른 소스에서 데이터의 **통합**을 역량은 **제한적**이었다. 오늘날, 센서의 다양한 배열에서 **대부분의 데이터가 디지털 형식으로** 이용할 수 있어, **데이터 통합(data integration)**은 **해석 및 분석에 사용되는 일반적인 방법**이다. 데이터 통합은 근본적으로 더 나은 또는 더 많은 정보를 추출하기 위한 노력의 일환으로 **여러 소스에서 데이터 결합(combining)하거나 병합(merging)**을 포함한다. 여기에는 **다중-시간(multi-temporal)**, **다중-해상도(multi-resolution)**, **다중-센서(multi-sensor)**, 또는 **다중-데이터 타입(multi-data type)**의 데이터를 포함 할 수 있다.



4.8 Data Integration and Analysis

- 서로 다른 시간에 수집된 이미지는 변화된 지역을 식별하기 위해 통합됩니다. **다중 시간 변화 감지(Multi-temporal change detection)**는 이미지 뺄셈(image subtraction), 또는 **다중 분류 비교(multiple classification comparisons)**나 **통합된 다중-시간 데이터 세트(integrated multi-temporal data sets)**를 사용한 분류와 같은 더 복잡한 방법에 의해서 달성 될 수 있다.

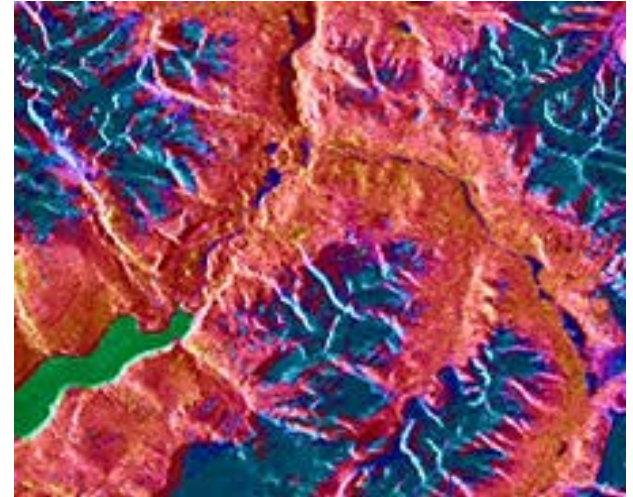
- **다중 해상도 데이터 병합(Multi-resolution data merging)**은 다양한 응용에서 유용하다. **높은 공간 해상도 데이터의 낮은 해상도 데이터와의 병합**은 이미지의 **공간적 세부사항을 매우 선명하게** 하며, **대상들의 식별을 향상시킬 수** 있습니다.

- **10m 전색성(panchromatic) 데이터가 20m 멀티 스펙트럼 데이터와 쉽게 병합** 될 수 있는 **SPOT 데이터가 이 방법에 적합**하다. 또한, **전색성 데이터는 향상된 공간 해상도를** 제공하는 반면에 **멀티 스펙트럼 데이터는 좋은 스펙트럼 해상도를** 유지하는 역할을 한다.



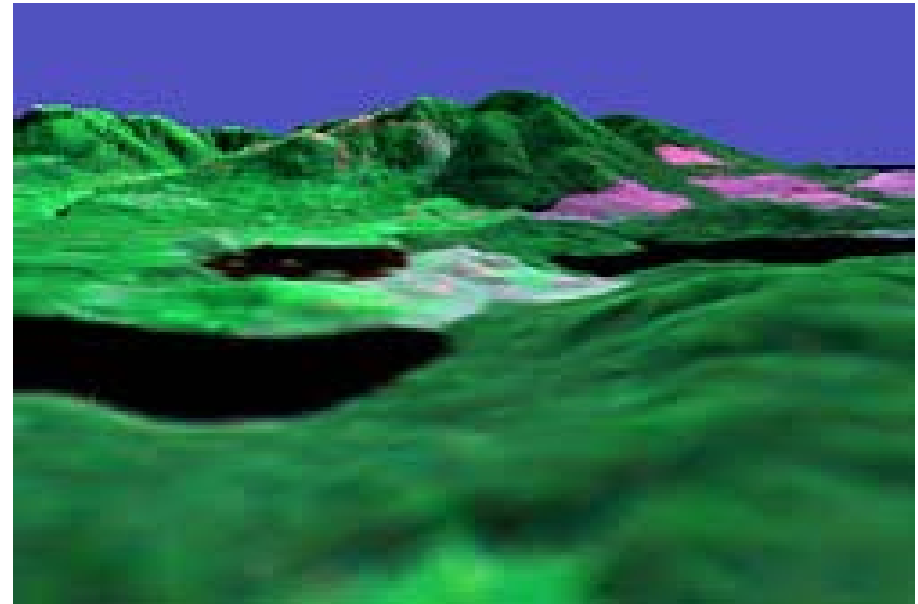
4.8 Data Integration and Analysis

- 다른 센서의 데이터도 또한 **다중-센서 데이터 융합 (multi-sensor data fusion)**의 개념으로 병합 할 수 있습니다.
- 이 기술의 좋은 예는 **멀티 스펙트럼 광학 데이터(multi-spectral optical data)**와 **레이더 이미지(radar imagery)**의 조합입니다. 지표의 이 두 가지 스펙트럼 표현은 **상호 보완적인 정보를 제공** 할 수 있습니다.
 - ✓ **광학 자료(optical data)**는, **지표 커버 유형**을 구별하는 데 유용한 **상세한 스펙트럼 정보**를 제공
 - ✓ **레이더 이미지(radar imagery)**는 이미지의 **구조적인 세부사항(structural detail)**을 강조한다



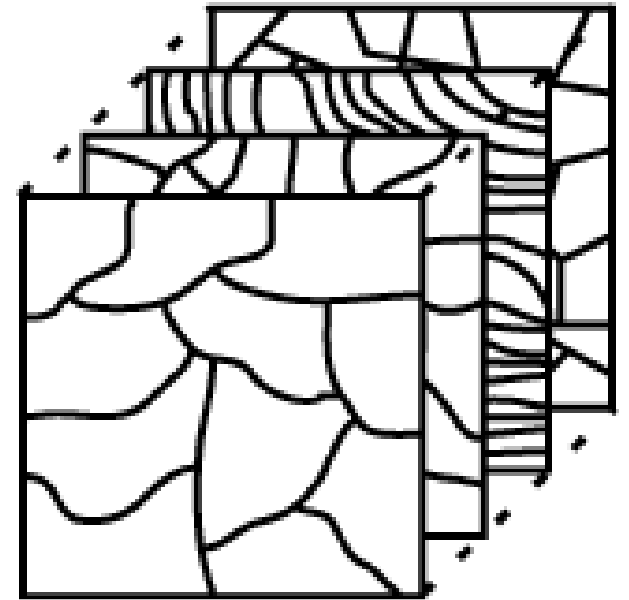
4.8 Data Integration and Analysis

- **멀티 센서 데이터 통합(multi-sensor data integration)**의 응용은 일반적으로 데이터가 서로 또는 **공통의 지리적 좌표 시스템** 또는 **기반 지도** 중의 하나에 **기하학적으로 등록** 할 것이 요구된다. 이것은 또한 **다른 보조(ancillary)(추가) 데이터** 소스가 **원격 탐사 데이터와 통합** 될 수 있게 해준다..
- 예를 들어, **디지털 고도(Digital Elevation)** 또는 **디지털 지형 모델(Digital Terrain Models) (DEM/DTM)** 이라는 디지털 형태 고도 데이터는, 다양한 목적을 위해서 **원격 감지 데이터와 결합** 될 수 있다. **DEM/DTM** 은 지형 및 경사 변동으로 인한 영향이 보정 될 수 있게 하여, 잠재적으로 얻어진 분류 결과의 정확성을 증가 시켜, 이미지 분류에 유용 할 수 있다.
- DEM/DTM은 영상화된 지역의 시각화를 향상하기 위해서, **원격 탐사 이미지를 고도 데이터(elevation data)에 올려 놓음**으로 해서(draping remote sensing imagery), **3차원 입체 투시도 (three-dimensional perspective views)**를 생성하는 데 유용합니다.



4.8 Data Integration and Analysis

- 다른 유형의 그리고 서로 다른 소스에서의 데이터를 결합하는 것이 데이터 통합 및 분석의 최종 결과이다. 모든 데이터 소스가 기하학적으로 공통 지리적 기반에 등록 된 디지털 환경에서는 정보 추출을 위한 가능성이 매우 넓어진다. 이것은 디지털 지리 정보 시스템 (Geographical Information System : GIS) 데이터베이스에서의 분석을 위한 개념이다.
- 공간적으로 참조(referenced spatially)될 수 있는 모든 데이터 소스가 이러한 환경의 유형에서 사용될 수 있다. DEM/DTM은 이러한 데이터 종류의 한 예입니다. 다른 예로는 응용에 따라 토양 타입, 토지 피복 클래스, 숲 종, 도로 네트워크, 그리고 많은 다른 것들의 디지털 지도가 있을 수 있습니다.



© CCRS / CCT

- 지도 형식의 원격 탐사 자료의 분류의 결과는 GIS 에서 기존지도 데이터를 업데이트하기 위한 또 다른 데이터 소스로 사용될 수 있다. 본질적으로, 단일 데이터 소스를 사용하여 단독으로 할 때 보다도, 다양한 데이터를 함께 분석(analyzing diverse data sets together)하여 시너지 방식으로 더 나은 그리고 보다 정확한 정보를 추출 할 수 있게 된다.