Image Classification

Present by:
Dr.Weerakaset Suanpaga
D.Eng(RS&GIS)

http://pirun.ku.ac.th/~fengwks/rs/

6.1 Concept of Classification

- Objectives of Classification
- Advantages of Multi-Spectral data for Classification
- Variation of Multi-Spectra Data
- Segmentation in Feature Domain
- Supervised and Un-Supervised Calssification
- Land Cover and Land Use
- Existing Land Cover Class

Objectives of Classification

- **To create Maps such as Landuse Map, Forest Map, Crop Map, Shrimp pond Map, Mangrove Map, etc.**
- **Carry out quantitative interpretation using mathematical /statistical modeling.**
- **To assign corresponding class to groups with homogeneous characteristics, with the aim of discriminating multiple objects from each other within the image.**
- **The level is called class. Classification will be executed on the base of spectrally defined features, such as density, texture etc. in the feature space. It can be said that classification divides the feature space into several classes based on a decision rule.**
- Classes are for such as Land use, Land Cover, Crop Type, Forest Types, and etc.

RS Image Classification

- Multi-Spectral Data Classification
  - Assumption - Different surface materials have defferent sepectral reflectance
  - K-dimensional vector (K:number of band)
  - divide K-dimensional feature space into few regions (classes)

Concept of Classification of Remote Sensing

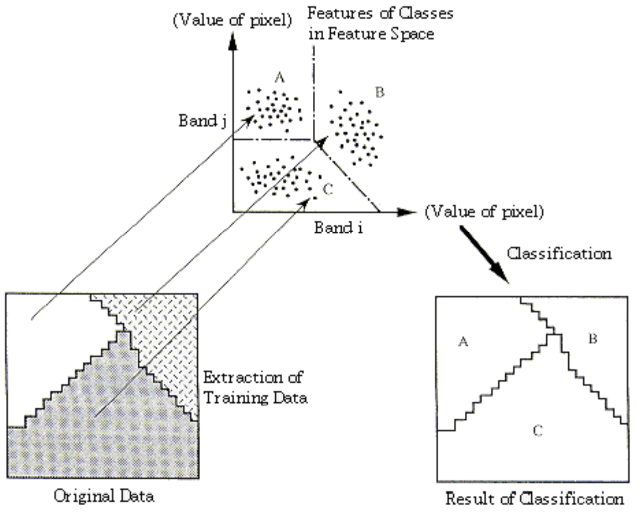
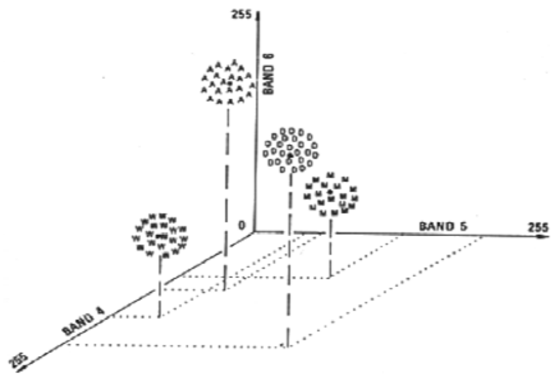


Figure 11.1.1 Concept of Classification of Remotely Sensed Data

5

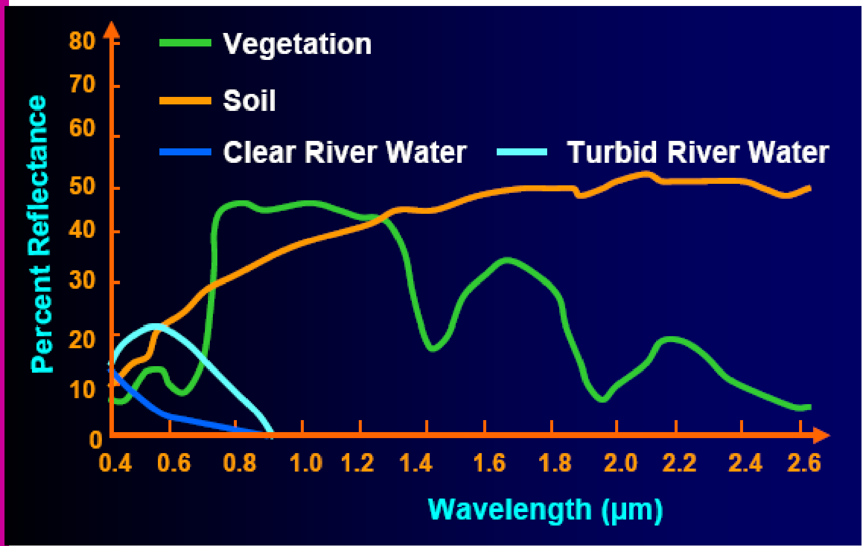
Segmentation in Feature Domain

- In general, the separation of all classes requires more than two spectral bands. Because the clusters occur in K-dimensions.



6

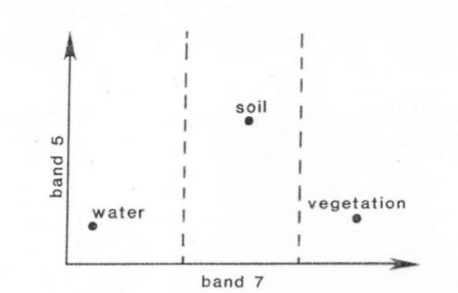
Spectral Reflectance



7

Multi-spectral Classification

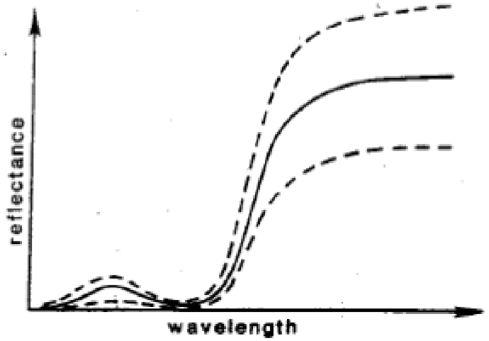
- The spectral signature is a K-dimensional vector whose coordinates are the measured radiance in each spectral band. If every pixel from each land cover has same radiance with in the class, only 1 band (IR) would be enough for classification for the case of water, soil and vegetation below.



8

Variation of Multispectral data

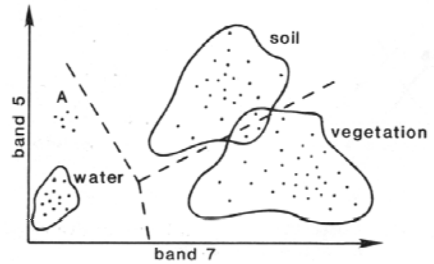
- In reality, the spectral radiance of a given surface material is not characterized by a single, deterministic curve, but by a family of curves with a range of variability.



9

Segmentation in Multi-dimensional feature space

- Thus, it is very common to find big overlaps among distributions in one band information.
- By combining other bands, we can improve the accuracy of classification, which is a segmentation in a multi-dimensional feature space.



10

Supervised and Un-Supervised Classification

- Supervised Classification
  - Classify each pixel into a pre-established class.
  - Population statistics of each class is to be identified by training areas.
  - Each pixel will be classified into a class which has similar (nearest) property with the pixel.
- Un-supervised Classification
  - Analyze inherent structure of the data
  - Unconstrained by external knowledge about area
  - When knowledge about the area is not enough
- Combination
  - Un-Supervised Classification -> Ground Truth -> Supervised Classification

11

Unsupervised

- **Clustering algorithm**
- **Objective and statistically valid**
- **May not be meaningful**
- **Class identification required**

12

Supervised

- Uses training areas
- Classes will be meaningful
- Classes may not be statistically valid

Supervised method

- Density Slicing
- Box classifiers
- Nearest neighbour
- Maximum likelihood
- End member analysis

Density Slicing

- Simplest, easiest to implement
- Uses only one band
- Prone to ambiguity

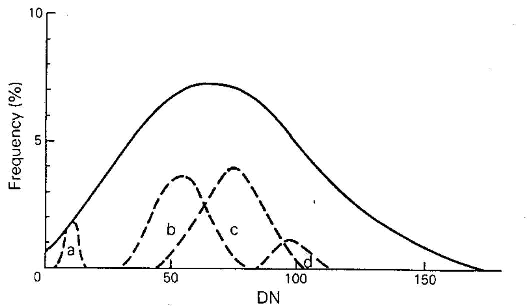
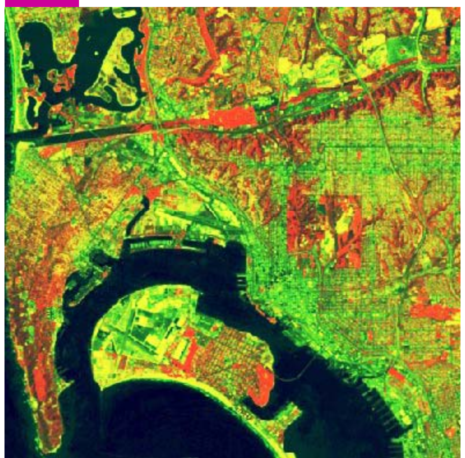
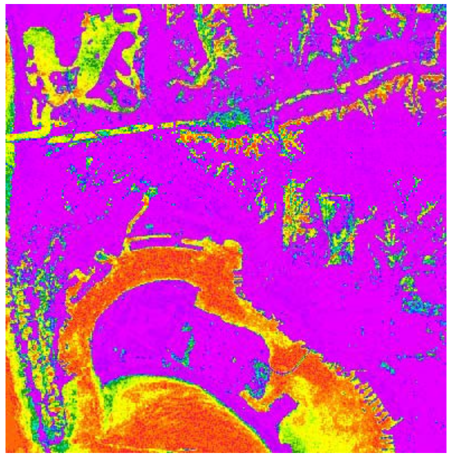


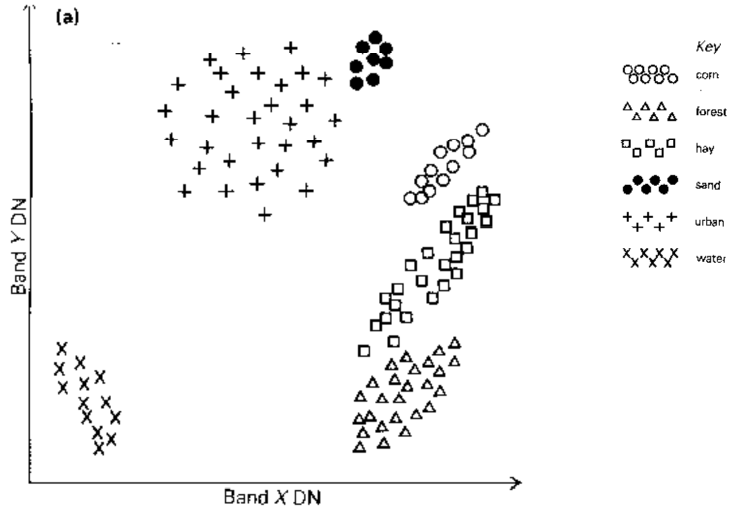
Figure 5.37 The histogram of one band for a whole scene conceals within it smaller histograms for particular classes of surface. In this case (a)–(d) are hypothetical classes.

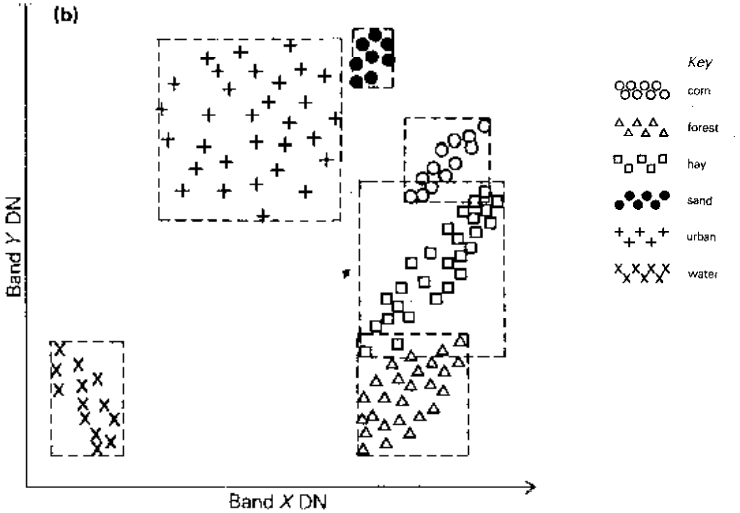




Box Classifiers

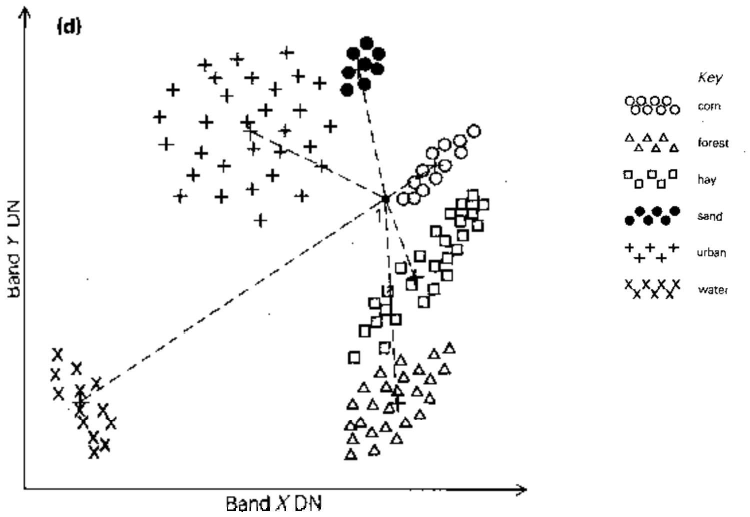
- Multi-band density slicing
- Defines a spectral "volume" for each class
- Reduces ambiguity
- Boundary solutions are arbitrary





Nearest Neighbour

- Defines a typical pixel for each class
- Assigns pixels on the basis of spectral distance
- Can separate diverse classes
- Boundary problems remain unresolved



Maximum Likelihood

- Most Popular methods
- **Defines a typical pixel for each class**
- **Calculates the probability that each pixel in the image belongs to that class**
- **Maps classes on the basis of confidence levels**
- **Boundary problems resolved**

Bayes Theory

- feature x --- for example, the gray level of each pixel
  - p(x | i) : probability density function in class i
  - p(i) : a priori probabilities
  - p(i | x) : a posteriori probabilities
- Bayes Rule
  - p(i | x) : p(x | i) p(i) / p(x)
    - ✓ If we observed feature x, what is the probability to be class i ?
  - p(x) = Σ p (x | i) p(i)
- Bayes Dicision Rule
  - one dimensional, two-class classification problem
  - a pixel belongs to class 1 if p(x|1)p(1) > p(x|2)p(2)
  - a pixel belongs to class 2 if p(x|2)p(2) > p(x|1)p(1)

Bayes Decision Rule

Forest		Agriculture	
0.6		0.4	
0.3	0.7	0.9	0.1
f1	f2	f1	f2
0.18	0.42	0.36	.04

p(forest) = 0.6
p(Agr) = 0.4
P(f1|Forest)=0.3
P(f2|Forest)=0.7
P(f1|Agr)=0.9
P(f2|Agr)=0.1

p(f1|Forest) *p(Forest)= 0.3 x 0.6 = 0.18
p(f2|Forest) *p(Forest)= 0.7 x 0.6 = 0.42
p(f1|Agr) *p(Agr)= 0.9 x 0.4 = 0.36
p(f2|Agr) *p(Agr)= 0.1 x 0.4 = 0.04

f1		f2	
0.54		0.46	
0.18	0.36	0.42	0.04
Forest	Agri	Forest	Ag
0.33	0.67	0.91	0.09

p(Forest|f1)=p(f1|Forest)*p(Forest) / p(f1) = 0.18 / 0.54 = 0.33
p(Agr|f1)=p(f1|Agr)*p(Agr) / p(f1) = 0.36 / 0.54 = 0.67
p(Forest|f2)=p(f2|Forest)*p(Forest) / p(f2) = 0.42 / 0.46 = 0.91
p(Agr|f2)=p(f2|Agr)*p(Agr) / p(f2) = 0.04 / 0.46 = 0.09

Discriminant Function

- The Bayes Dicision Rule is restated as
  - a pixel belongs to class 1 if D1(x) > D2(x)
  - a pixel belongs to class 2 if D2(x) > D1(x)
- where Di is called discriminant function and is given by
  - Di(x) = p(x | i) p(i)
  - However P(i) is unknown, we assume p(i)=p(j)

Assumption of Normal Distribution

- If the class probability distributions are normal

$$p(x \mid i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right)$$

$$\mu_i = \text{mean of } x \text{ for class } i$$

$$\sigma_i^2 = \text{variance of } x \text{ for class } i$$

- Bayes optimal discriminant function for class i is then

$$D_i(x) = \ln[p(x \mid i)\, p(i)]$$

$$= \ln[p(i)] - \frac{1}{2}\ln[2\pi] - \frac{1}{2}\ln[\sigma_i^2] - \frac{(x-\mu_i)^2}{2\sigma_i^2}$$

- p(i) is unknown. Assumption of p(i) = p(j),

$$D_i(x) = -\frac{1}{2}\ln[\sigma_i^2] - \frac{(x-\mu_i)^2}{2\sigma_i^2}$$

Extension to K Dimension

$$p(x \mid i) = \frac{1}{(2\pi)^{K/2}|\Sigma_i|^{1/2}} \exp\left[-\frac{1}{2}(X-\mu_i)^t \Sigma_i^{-1}(X-\mu_i)\right]$$

$$D_i(x) = \ln[p(i)] - \frac{K}{2}\ln[2\pi] - \frac{1}{2}\ln[|\Sigma_i|] - \frac{1}{2}(X-\mu_i)^t\Sigma_i^{-1}(X-\mu_i)$$

$$D_i(x) = -\frac{1}{2}\ln[|\Sigma_i|] - \frac{1}{2}(X-\mu_i)^t\Sigma_i^{-1}(X-\mu_i)$$

X : vector of image data (K dimension) $\quad X = [x_1, x_2, ..., x_k]$

μ_i: mean vector for class i $\quad \mu_i = [m_1, m_2, ..., m_k]$

Σ_i: variance-covariance matrix for class i

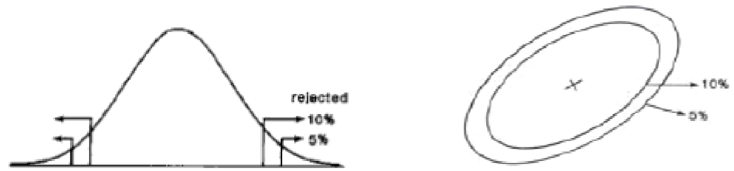
$$\Sigma_i = \begin{bmatrix} \sigma_{11} & \sigma_{11} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{bmatrix}$$

$|\Sigma_i|$: determinant of Σ_i

Σ_i^{-1}: inverse matrix of Σ_i

Thresholding

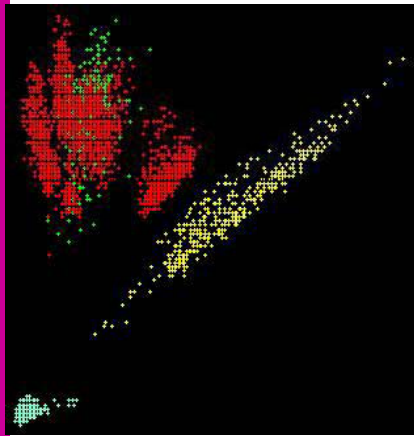
- Eliminate pixels which have low posteriori probability

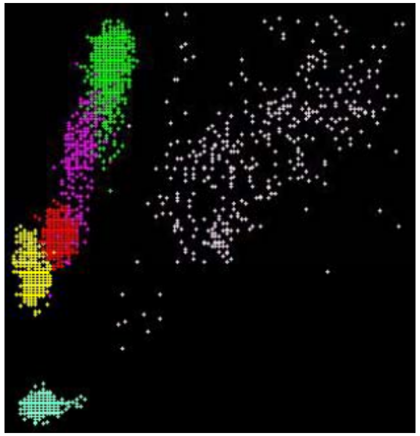


(a) ONE DIMENSION

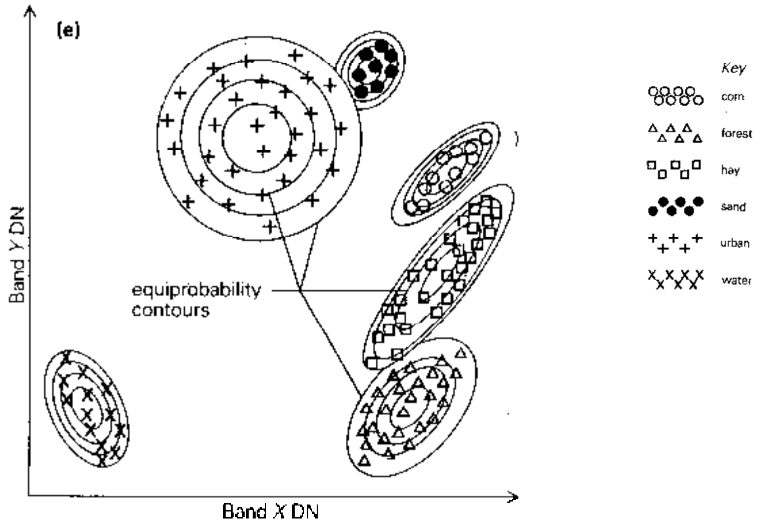
(b) TWO DIMENSIONS

Actual Distribution

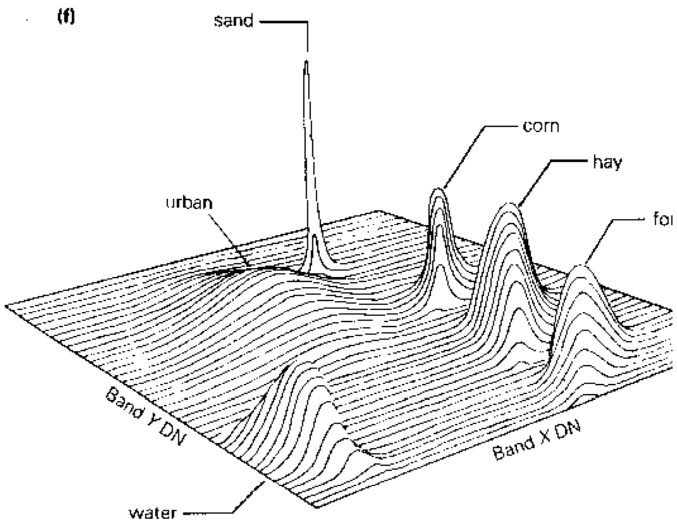




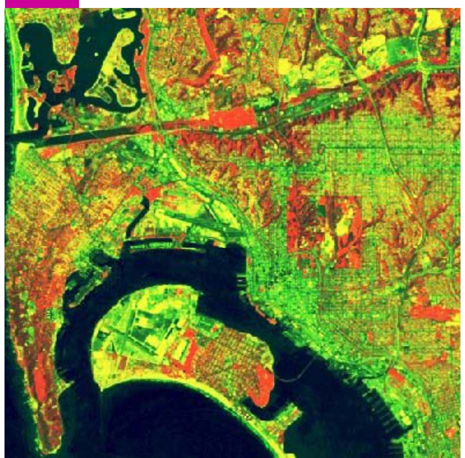
29

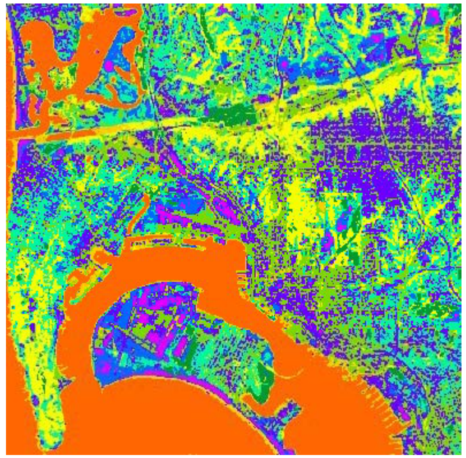


30



31





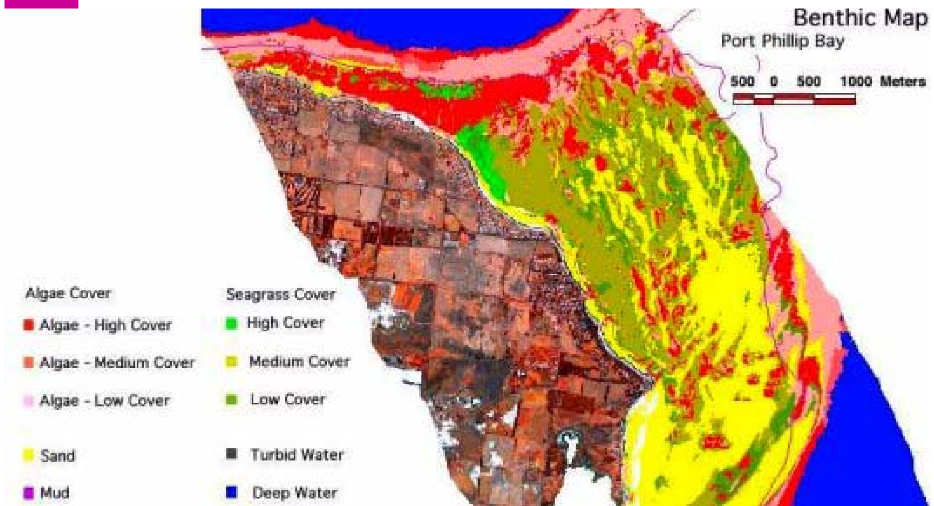
32

End Member Analysis

- Still experimental
- Uses a library of known spectral curves to match the observed curve
- Must have N+1 bands to avoid ambiguity
- Limited by data requirements

33

34



35

6.4 Unsupervised Classification

- To determine the inherent structure of the data, unconstrained by

external knowledge about the area.

- To produce clusters automatically, which consists of pixels with similar spectral signature
  - Hierarchical Clustering
    - ✓ Evaluate distance between clusters
    - ✓ Merge a pair of clusters which have the minimum distance.
    - ✓ Members are not reallocated to different clusters
- Non-Hierarchical Clustering
  - K-mean, ISODATA method
  - Reallocation of members
  - Merge and Division of clusters

36

Hierarchical clustering

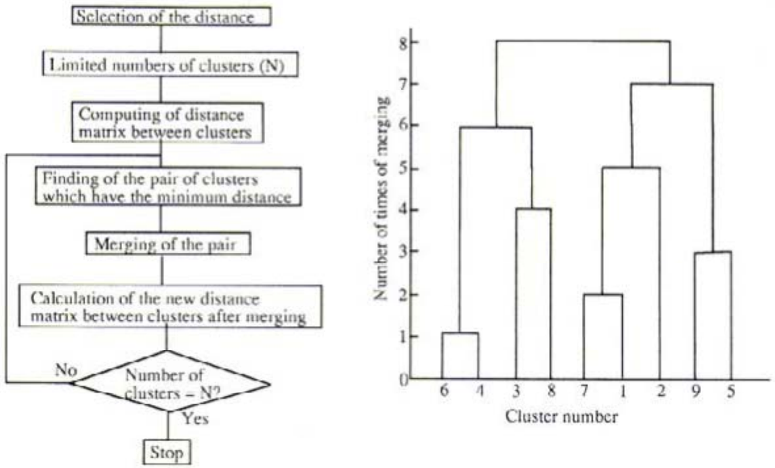


Figure 11.3.1 Flow and an example of hierarchical clustering

ISODATA method

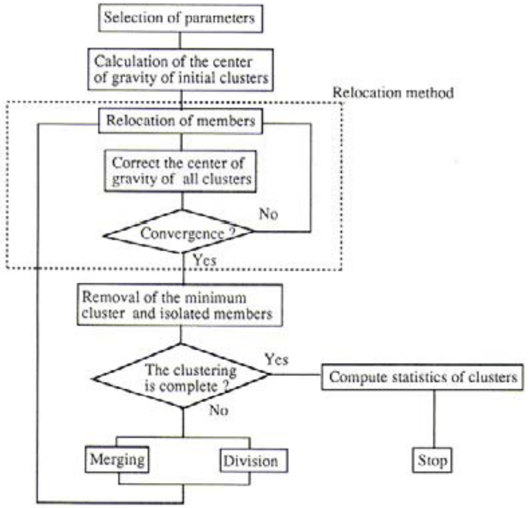
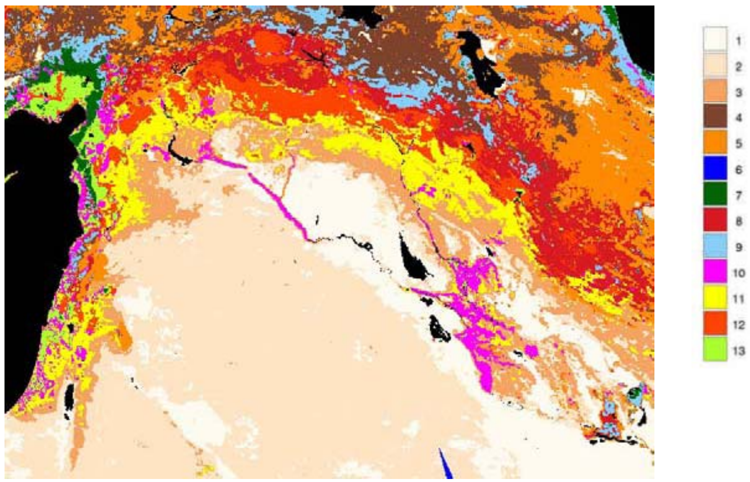


Figure 11.3.2 Flow of non-hierarchical clustering (ISODATA method)

ISODATA Unsupervised Classification example



6.5 Accuracy Assessment

- **Accuracy assessments determine the quality of the information derived from remotely sensed data (Congalton and Green, 1999).**
- **Accuracy assessment is important to produce reliable maps.**
- **Assessments can be either qualitative or quantitative. In qualitative assessments, we determine if a map "looks right" by comparing what we see in the imagery with what we see on the ground.**
- **However quantitative assessments attempt to identify and measure remote sensing-based map error. In such assessments, we compare map data with reference or ground truth data.**

Reference/Ground truth data collection

- Usually we divide ground truth into two.
  - 50% is used for supervised classification training
  - 50% is used for accuracy assessment
- Aerial photographs
- Other Maps
- Ground based data is assumed to be 100% correct in accuracy assessments, hence it's very important that the data is collected carefully. It should be collected consistently with vigilant quality control.

41

Common quantitative error assessments

- Error Matrix or Confusion Matrix – assesses accuracy for each class as well as for the whole image; this includes errors of inclusion and errors of exclusion
- We must accept some level of error as a trade off for the cost savings of remotely sensed data (Congalton and Kass, 1999)

42

Confusion Matrix

Classified

Reference (Ground Truth)

	Urban	Crop	Range	Water	Forest	Barren	Total	PA	EO
Urban	150	21	9	7	17	30	**234**	64.1%	35.9%
Crop	0	730	93	14	115	21	**973**	75.0%	25.0%
Range	33	121	320	23	54	43	**594**	53.9%	46.1%
Water	3	18	11	83	8	3	**126**	65.9%	34.1%
Forest	23	81	12	4	350	13	**483**	72.5%	27.5%
Barren	39	8	15	3	11	115	**191**	60.2%	39.8%
Total	**248**	**979**	**460**	**134**	**555**	**225**	**1748**		
CA	60.5%	74.6%	69.6%	61.9%	63.1%	51.1%			
EC	39.5%	25.4%	30.4%	38.1%	36.9%	48.9%			

Total Pixel		2601
Correct Pixel	150+730+320+83+350+11	1748
Overall Accuracy	=1748/2601	67.2%

PA	Producers Accuracy
CA(UA)	Consumer's (User's) Accuracy
EO	Error of Omission = 100%−PA
EC	Error of Comission = 100%−CA

43

END

Thank you for Attention

44