Image Classification

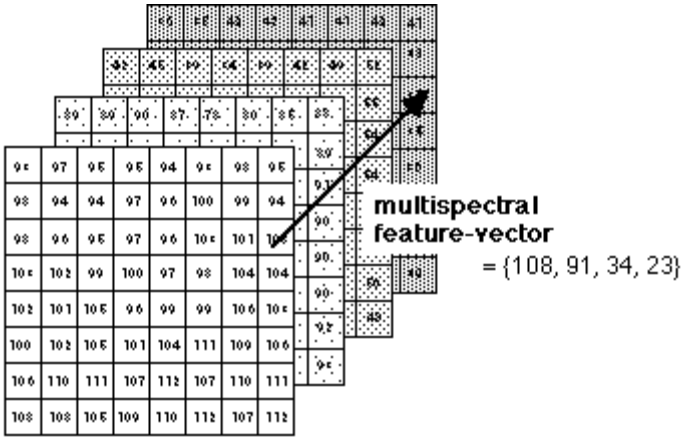
References
1. R. A. Schowengerdt, 1997. *Remote Sensing models and methods for image processing*, 2nd ed., Academic Press, Chap. 9.
2. Lillesand and Kiefer, 1994, *Remote Sensing and Image Interpretation*, 3rd ed., Wiley, Chap. 7.7
2. http://www.watleo.uwaterloo.ca/~piwowar/geog376/ImageAnalysis/Classification/Classification.html
3. http://www.esf.edu/forest/supervisedClass.html

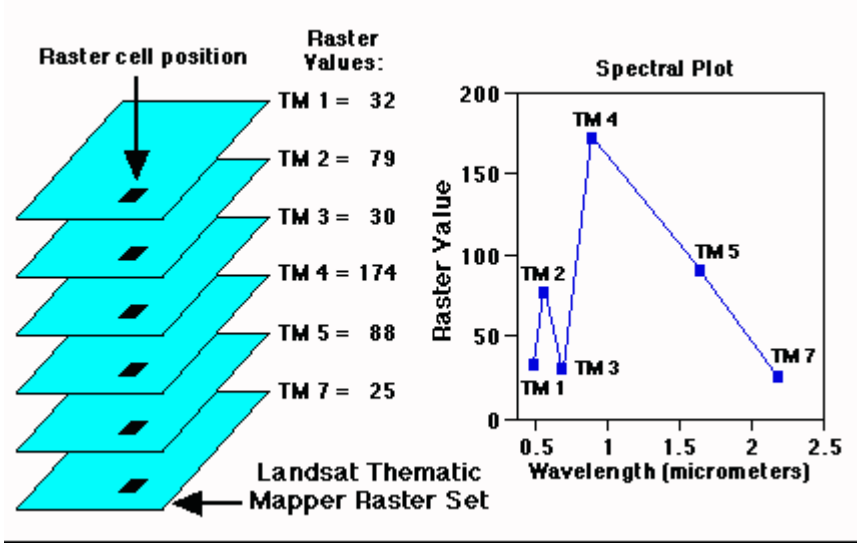
Classification

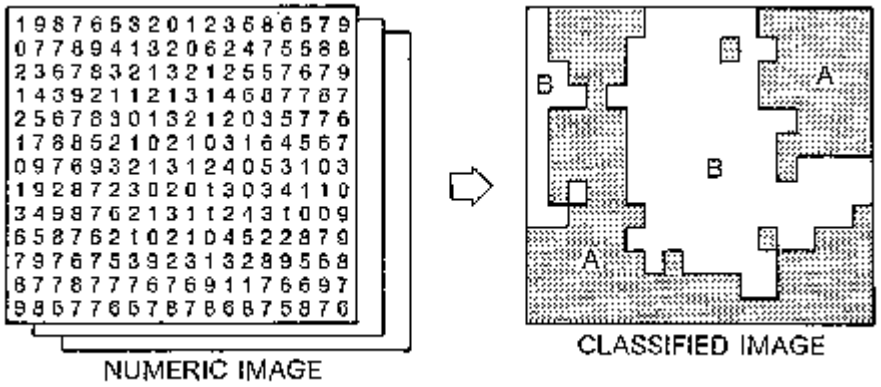
- Definition
  - The process of reducing images to information classes. Classification divides the spectral or spatial **feature space** into several classes based on a decision rule.

- General Procedures
  - **Feature Extraction** : Transform the multispectral image by a spatial or spectral transform to a feature image (optional). Ex) selection of bands, filtering, PCA.
  - **Training** : Extract the pixels to be used for training the classifier to recognize certain categories, or classes. Determine the *discriminant functions* in the feature space. *Supervised* or *unsupervised*
  - **Labeling** : Apply the discriminant functions to the entire feature image and label all pixels. The output consists of one label for each pixel.

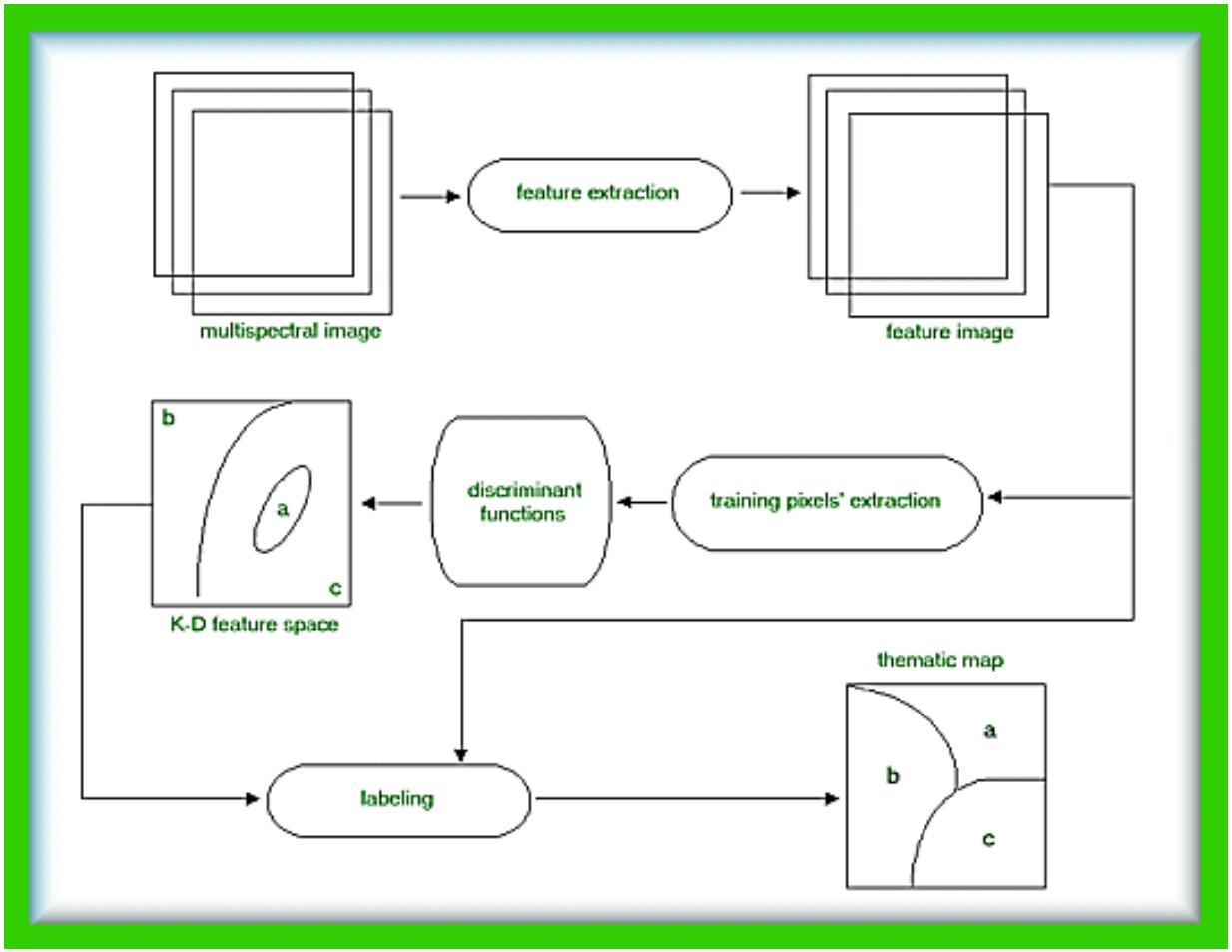
Classification – Feature Space







Classification Procedures



Classification Methods

By the use of **Feature Space**:

- Spectral pattern recognition
- Spatial pattern recognition
- Temporal pattern recognition
- Spatio-spectral pattern recognition

By the use of **Training Method**:

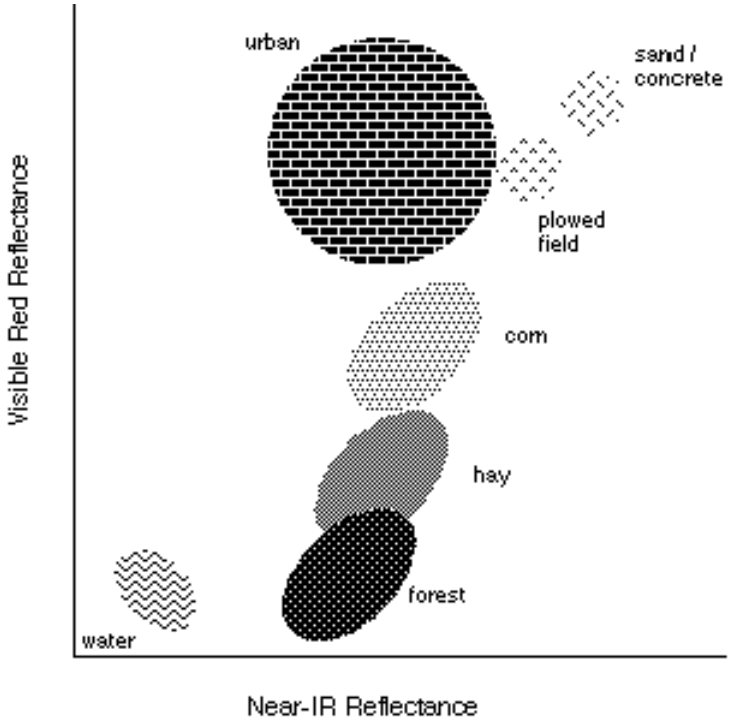
- Supervised Training
- Unsupervised Clustering
- Hybrid (Supervised/Unsupervised) Classification

by the use of **Labeling Method** (classifier):

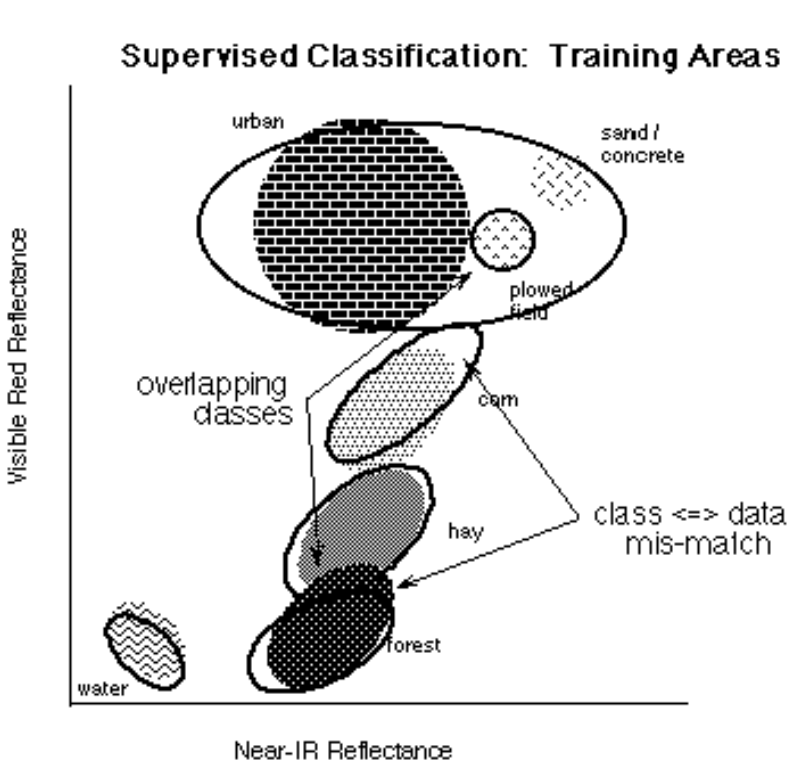
- Non-Parametric: do not use statistics
  - Level-Slice Classifier
  - Parallelepiped Classifier
  - Histogram Estimation Classifier
  - Nearest Neighbors Classifier
  - Artificial Neural Network Classifier
- Parametric: use mean, covariance
  - Nearest Mean Classifier (Minimum Distance Classifier)
  - Maximum Likelihood Classifier

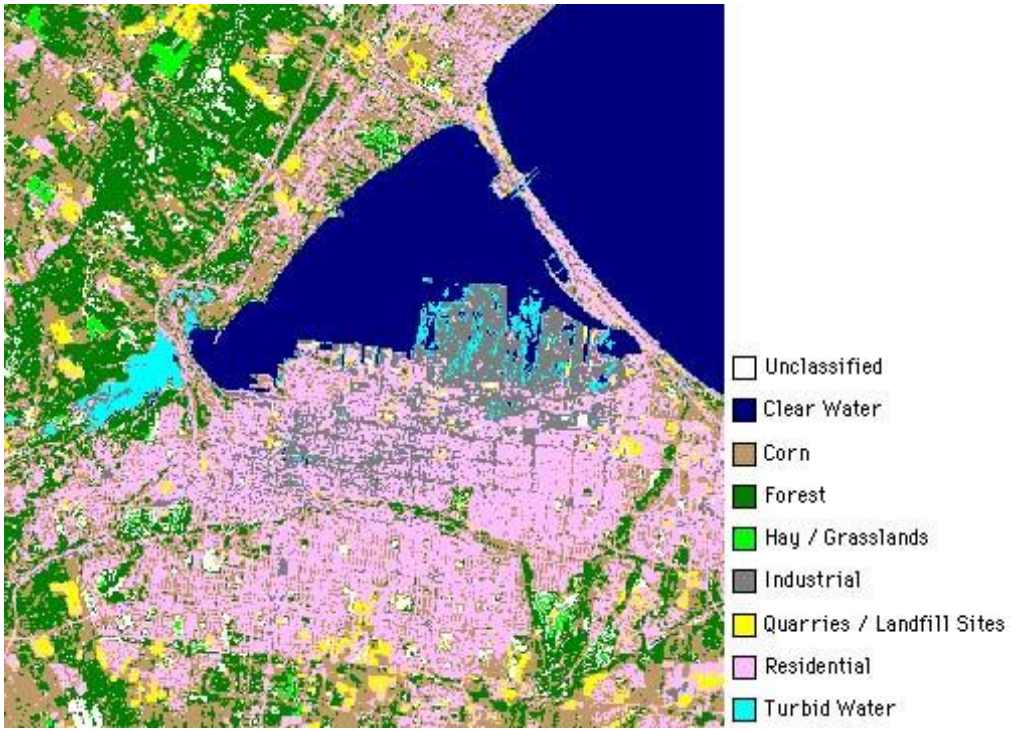
Supervised Classification

- The training area should be a homogeneous sample of the respective class, but at the same time include the range of variability for the class
- More than one training area per class is often used.



Example of Supervised Classification





Unsupervised Classification

- the process of automatically segmenting an image into spectral classes based on natural groupings found in the data

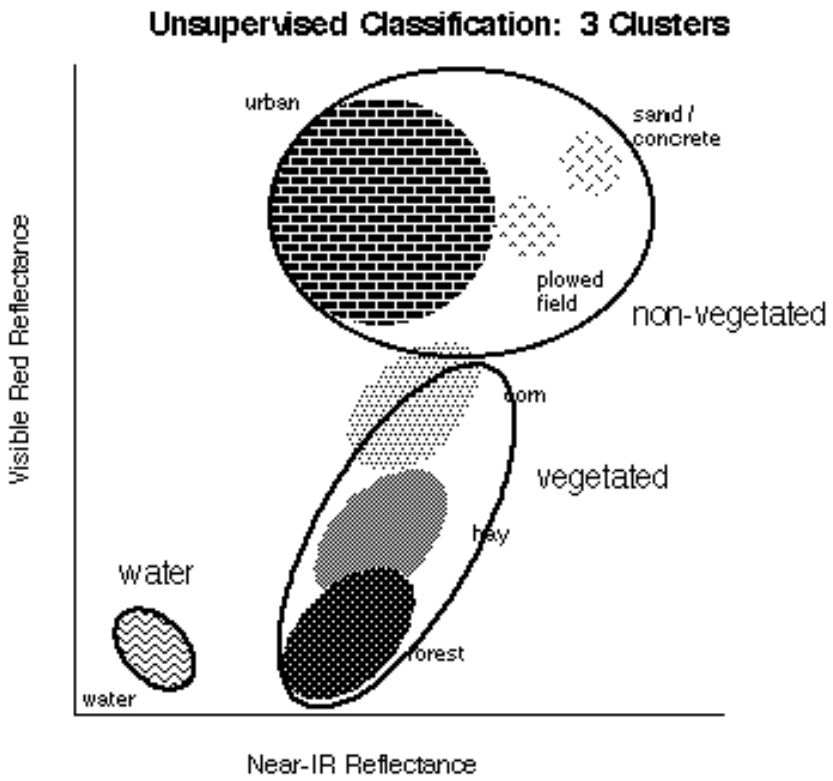
Procedure

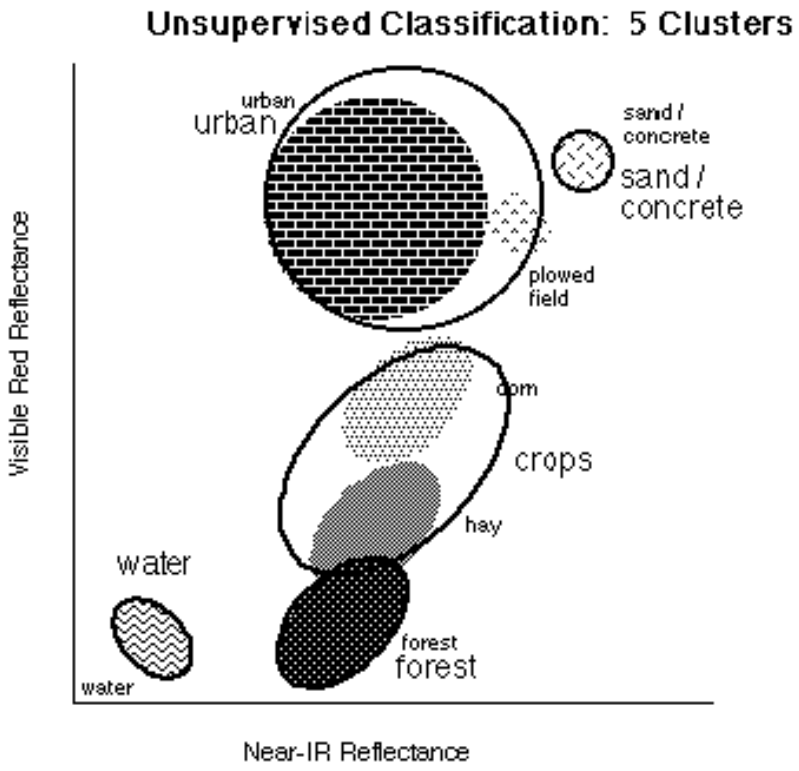
- Classify the image
- Identify clusters (Clustering)
- Accuracy assessment

Clustering

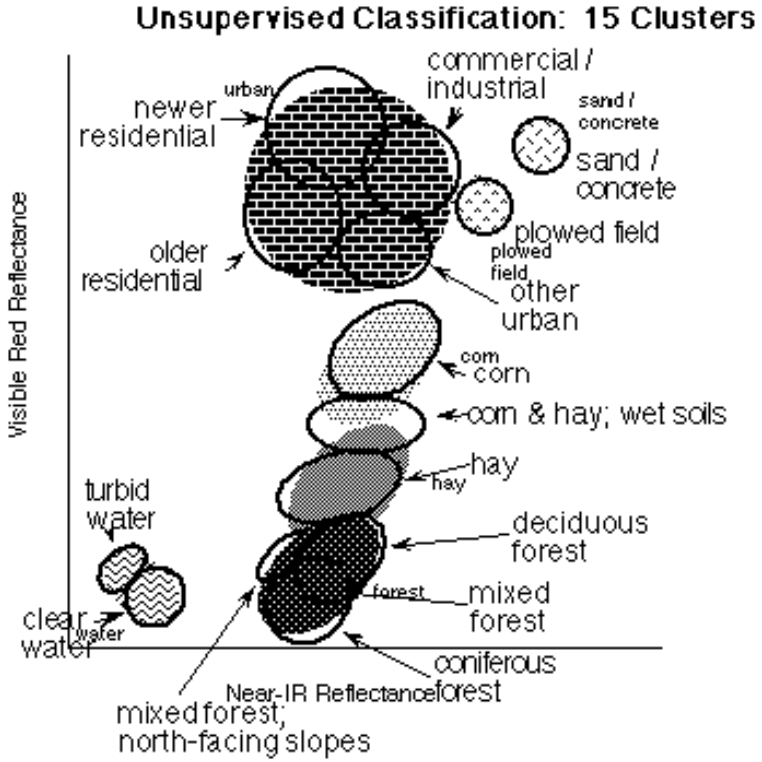
- Sequential Clustering
- K-means Clustering
- ISODATA (Iterative Self Organizing Data) Clustering

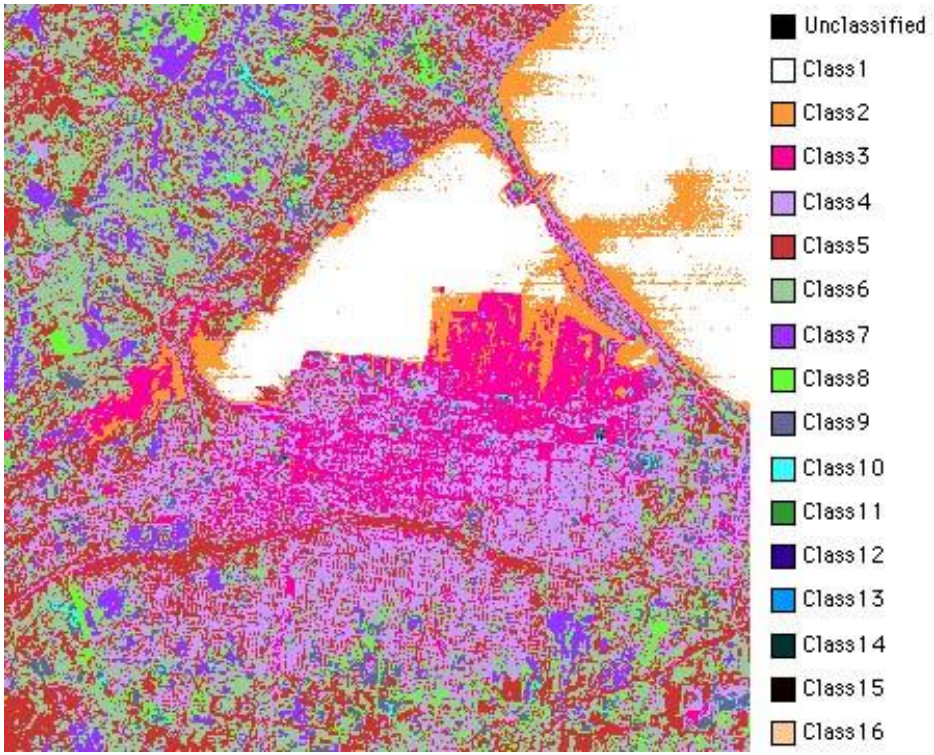
Example of Unsupervised Classification





Example of Unsupervised Classification -continued





Supervised vs. Unsupervised Classification

Supervised

- pre-defined classes
- serious classification errors detectable
- defined classes may not match natural classes
- classes based on information categories
- selected training data may be inadequate
- a priori class training is time-consuming and tedious
- only pre-defined classes will be found

Unsupervised

- unknown classes
- no classification errors
- natural classes may not match desired classes
- classes based on spectral properties
- derived clusters may be unidentifiable
- a posteriori cluster identification is time-consuming and tedious
- unexpected categories may be revealed

Nearest Mean Classifier (Minimum Distance Classifier)

Advantages:

- mathematically simple
- computationally efficient

Disadvantages:

- insensitive to different degrees of variance in the data (point 2)

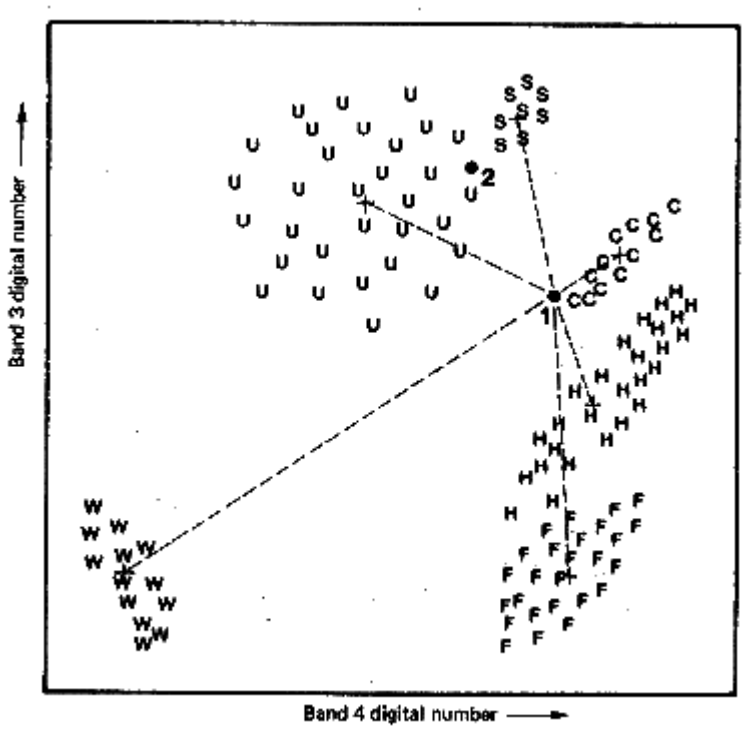


Figure 7.39 Minimum distance to means classification strategy.

Level-Slice Classifier (Parallelepiped Classifier)

- *Rectangular (parallelepipeds* in multidimension) decision range

- Advantages:
  - mathematically simple
  - computationally efficient
  - sensitive to different degrees of variance in the data
- Disadvantages:
  - problems occur in regions of overlap
  - does not account for inter-band covariance (point 1)

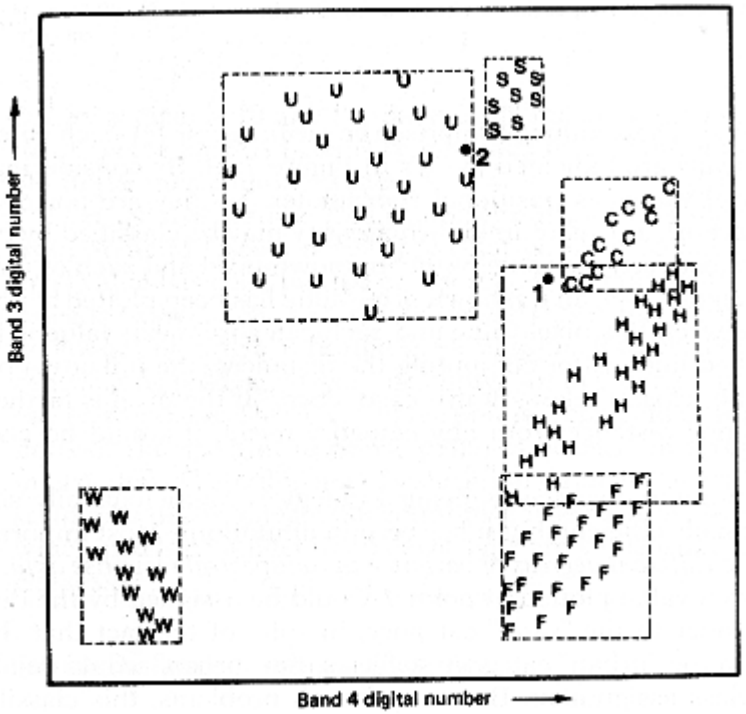


Figure 7.40 Parallelepiped classification strategy.

Maximum Likelihood Classifier

Assmption of Normality

Mean Vector, Covariance Matrix

Probability Density functions

Advantages:

- accounts for covariance between bands
- generally produces the most accurate classifications

Disadvantages:

- requires an assumption of normality in the training data
- mathematically complex
- computationally slow

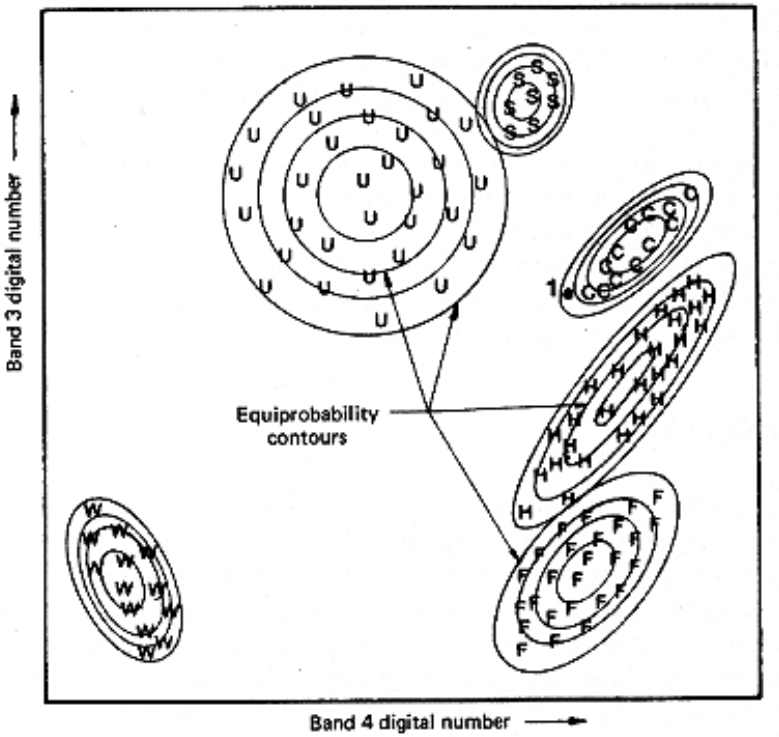
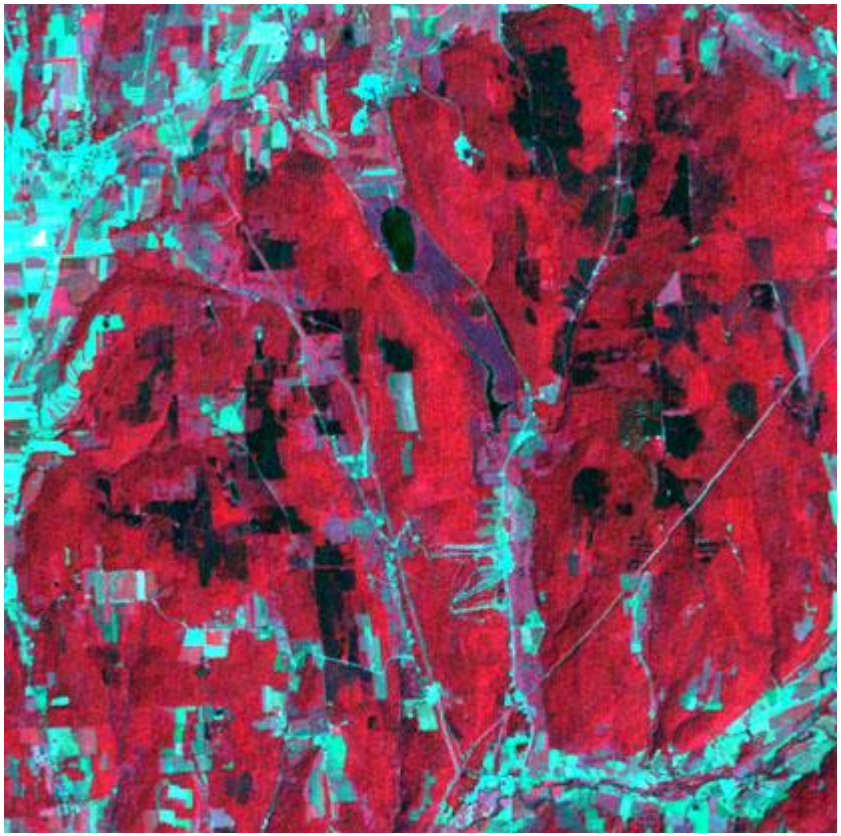
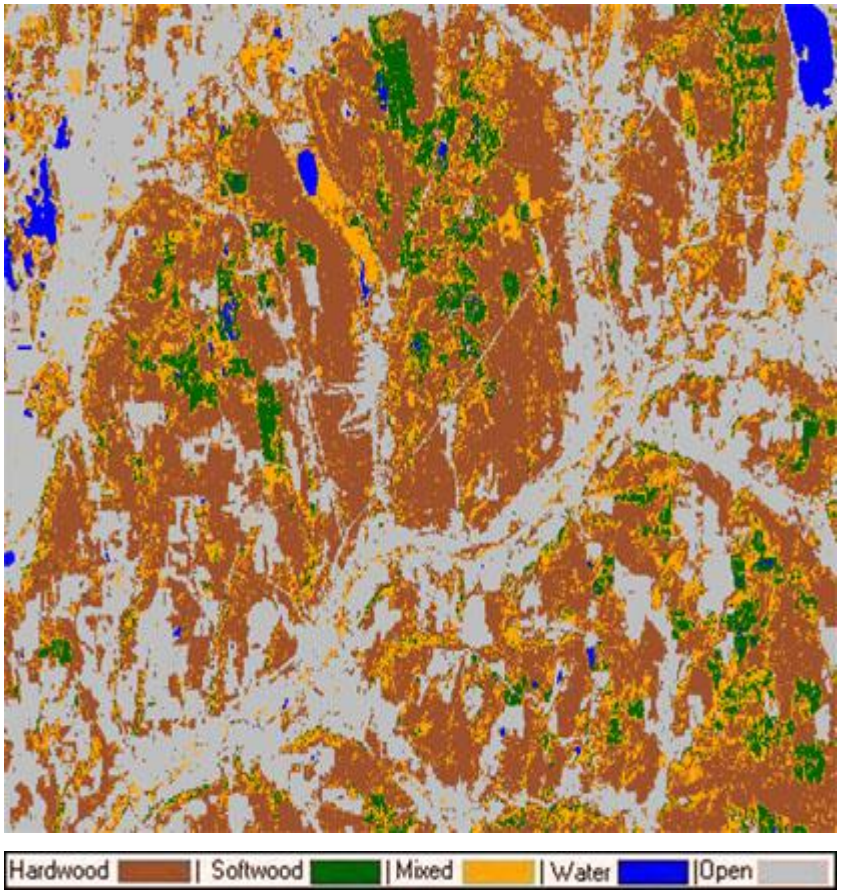


Figure 7.43 Equiprobability contours defined by a maximum likelihood classifier.

Example: Forest Type Classification

http://www.esf.edu/forest/supervisedClass.html





Landsat ETM+, Central New York, 1999/07/28

Classification Error Matrix

	Hardwood	Softwood	Mixed	Water	Open	User's Accuracy
Hardwood	41	3	11	0	1	0.73
Softwood	3	8	14	0	0	0.32
Mixed	4	2	24	0	0	0.80
Water	2	0	3	11	0	0.69
Open	0	0	0	0	8	1.00
Producer's Accuracy	0.82	0.62	0.46	1.00	0.89	
	Overall Accuracy:		0.68			

- The relationship between known reference data (ground truth) and the corresponding results of an automated classification.
- One of the most common means of expressing classification accuracy (also called confusion matrix or contingency table).
- ***Overall Accuracy*** = (Total number of correctly classified pixels)/(Total number of reference pixels).
- ***Producer's Accuracy*** = (Number of correctly classified pixels in each category)/(Number of training set pixels used for that category). This figure indicates how well training set pixels of the given cover type are classified.
- ***User's Accuracy*** = (Number of correctly classified pixels in each category)/(Number of pixels classified in that category). This figure is a measure of commission error and indicates the probability that a pixel classified into a given category actually represents that category on the ground.