LA502 Special Studies Remote Sensing

Image Classification_2

Dr. Ragab Khalil
Department of Landscape Architecture
Faculty of Environmental Design
King AbdulAziz University
Room 103

Supervised classification

Overview

- Image classification techniques
- Supervised classification
- Classification stage
- Training stage

Classification techniques

Digital image classification techniques can generally be classified into two types:

- Unsupervised classification techniques,
- Supervised classification techniques.

Last week the Unsupervised classification was covered.

Supervised Classification

Supervised classification techniques require the image analyst to define the classification categories and identify a representative samples of pixels to the computer.

The computer then assign all of the remaining pixels to one of the predefined classes on the basis of the similarities between the digital number in the training pixels and the digital numbers in all other pixels.

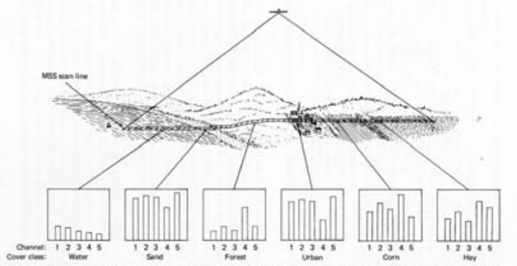


Figure 7.37 Selected multispectral scanner measurements made along one scan line. Sensor covers the following spectral bands: 1, blue; 2, green; 3, red; 4, near infrared; 5, thermal infrared.

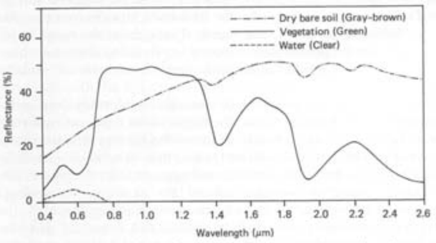


Figure 1.10 Typical spectral reflectance curves for vegetation, soil, and water. (Adapted from Swain and Davis, 1978.)

Supervised Classification steps

1. The analyst defines a classification scheme,
2. *Training stage:*
   - The analyst identifies pixels known to fall in each class for the computer,
3. *Classification stage:*
   - The computer put each image pixel into a class based on the multispectral data ranges of the training pixels, and then
4. *Output stage:*
   - The computer generates a classified image.

Classification Steps

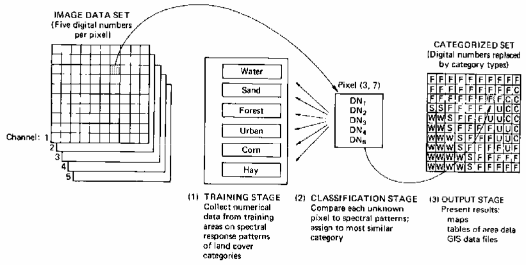


Figure 7.38 Basic steps in supervised classification.

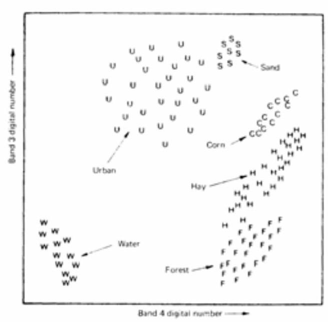


Figure 7.39 Pixel observations from selected training sites plotted on scatter diagram.

Classifiers

One of several possible classification strategies, or *classifiers*, will then be applied to assign the remaining pixels into one of the predefined classes. Three classifiers will considered here:

- The minimum-distance-to-means classifier,
- The parallelepiped classifier, and
- The Gaussian maximum likelihood classifier.

Minimum-distance-to-means classifier

- Calculates mean of the spectral values for the training set in each band and for each category
- Measures the distance from a pixel of unknown identify to the mean of each category
- Assigns the pixel to the category with the shortest distance
- Assigns a pixel as "unknown" if the pixel is beyond the distances defined by the analyst

C

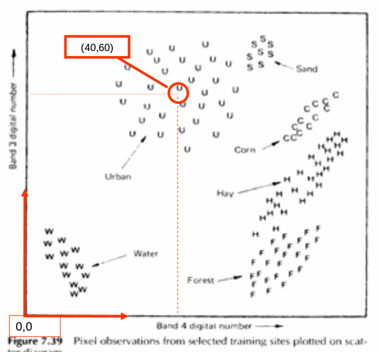


Figure 7.39 Pixel observations from selected training sites plotted on scatter diagram.

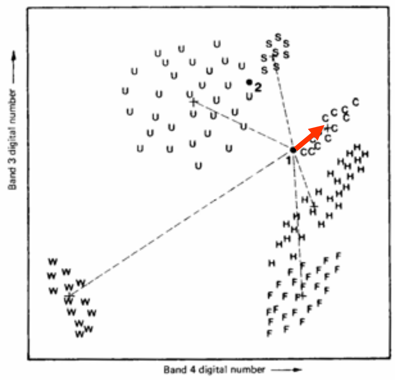


Figure 7.40 Minimum distance to means classification strategy.

Minimum-distance-to-means classifier

- Advantage
  computationally simple and fast
- Disadvantage
  insensitive to differences in variance among categories

Parallelepiped Classifier

- Forms a decision region by the maximum and minimum values of the training set in each band and for each category
- Assigns a pixel to the category where the pixel falls in
- Assigns a pixel as "unknown" if it falls outside of all regions

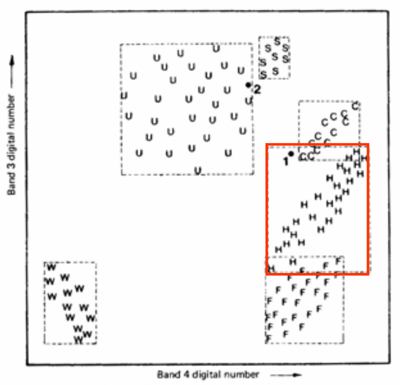


Figure 7.41 Parallelepiped classification strategy.

Parallelepiped Classifier ..

- Advantage
  - computationally simple and fast
  - takes differences in variance into account
- Disadvantage
  - performs poorly when the regions overlap because of high correlation between categories (high covariance)

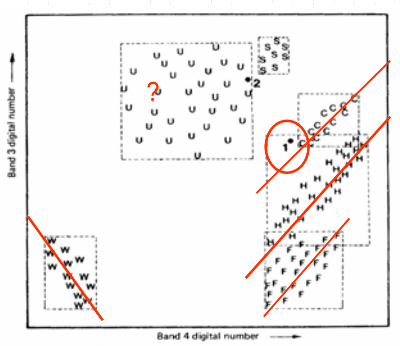


Figure 7.41 Parallelepiped classification strategy.

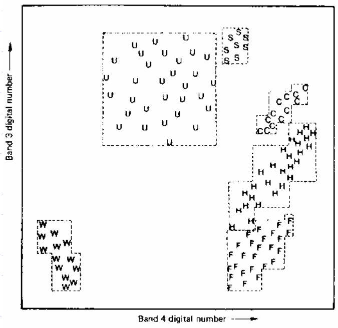


Figure 7.42 Parallelepiped classification strategy employing stepped decision region boundaries.

Gaussian Maximum likelihood Classifier

- Assumes the (probability density function) distribution of the training set is normal
- Describes the membership of a pixel in a category by probability terms
- The probability is computed based on probability density function for each category

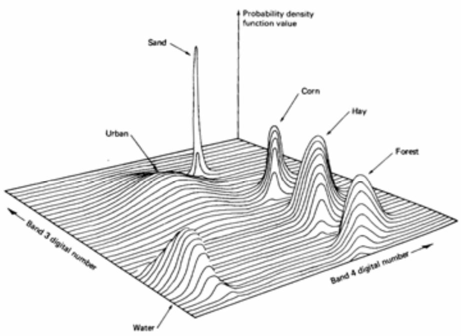


Figure 7.43 Probability density functions defined by a maximum-likelihood classifier.

Gaussian Maximum likelihood Classifier ..

- A pixel may occur in several categories but with different probabilities
- Assign a pixel to the category with the highest probability

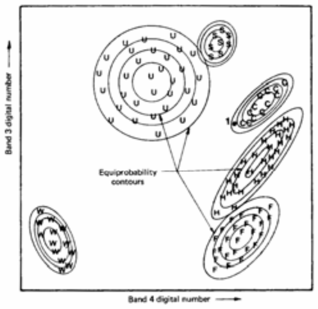


Figure 7.44 Equiprobability contours defined by a maximum-likelihood classifier.

Gaussian Maximum likelihood Classifier ..

- Advantage
  takes into account the distance, variance, and covariance
- Disadvantage
  computationally intensive

Training

- Collect a set of statistics that describe the spectral response pattern for each land cover type to be classified
- Select several spectral classes representative of each land cover category
- Avoid pixels between land cover types

Training sites

Training sites are geographic areas within which all of the pixels are believed to belong to a single spectral class, i.e. uniform areas of known land cover types.

Training sites

•ERDAS IMAGINE uses *areas of interest* to enclose training sites.

•An area of interest is a polygon enclosing one or more pixels representing a land cover class in the proposed classification scheme.

Training sites

- The more general name for an area of interest used in this manner is a *viewing window*.
- All of the pixels within a viewing window should be representative of a particular spectral class.

Dr. Ragab Khalil KAAU - FED – LA502: RS 28/46

Training sites

Because all of the pixels in a viewing window represent a particular spectral class, viewing windows should avoid edge areas so that transition zones are not included in the class.

Dr. Ragab Khalil KAAU - FED – LA502: RS 29/46

Training ..

- A minimum of n+1 pixels must be selected (n=number of bands)
- More pixels will improve statistical representation, 10n or 100n are common
- Spatially dispersed training areas throughout the scene better represent the variation of the cover types

Dr. Ragab Khalil KAAU - FED – LA502: RS 30/46

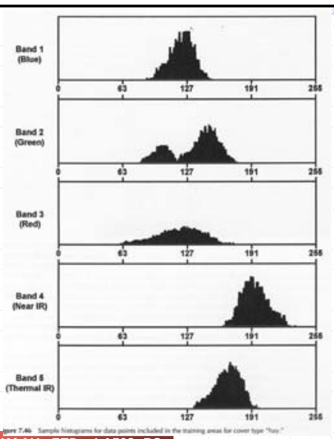
Training Set Refinement

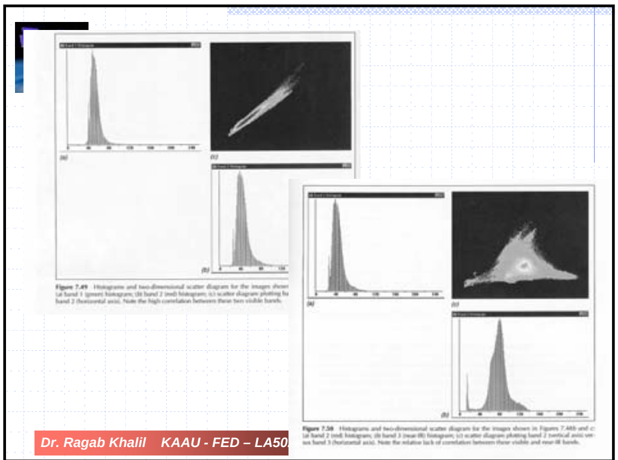
- Graphic representation
- Quantitative expression
- Self-classification

Training Set Refinement ..

Graphic representation

- It is necessary to display histograms of training sets to check for normality and purity
- Coincident spectral plot with 2 std dev from the mean is useful to check for category overlap
- 2-D scatter gram is also useful for refinement





Training Set Refinement ..

Quantitative expression

- Euclidian
- Divergence
- Transformed divergence (maximum 2000)
- Jefferies- Matusita (maximum 1414)

higher values indicate greater separability

TABLE 7.1 Portion of a Divergence Matrix Used to Evaluate Pairwise Training Class Spectral Separability

Spectral Class[a]	W1	W2	W3	C1	C2	C3	C4	H1	H2 · · ·
W1	0								
W2	1185	0							
W3	1410	680	0						
C1	1997	2000	1910	0					
C2	1953	1890	1874	860	0				
C3	1980	1953	1930	1340	1353	0			
C4	1992	1997	2000	1700	1810	1749	0		
H1	2000	1839	1911	1410	1123	860	1712	0	
H2	1995	1967	1935	1563	1602	1197	1621	721	0
⋮	⋮								

[a] W, water; C, corn; H, hay.

Training Set Refinement ..

Training set self-classification

- Interactive preliminary classification
  - use simple and fast classifier to classify the entire scene
- Representative sub-scene classification

Interactive Preliminary Classification

During this process, the analyst is able to add or remove pixels from the preliminary training classes to immediately see the results of the modified classification.

This interactive process easily identifies individual pixels that either improve or degrade the quality of the classification, resulting in final spectral classes that produce good results.

Representative subscene classification

Representative subscene classification is similar to interactive preliminary classification except that the classification is performed with the type of classifier that the analyst plans to use in the final classification step and only part of the entire image is classified.

Using part of the full image allows the analyst to concentrate on areas where the cover types are already well known and makes the process less time consuming.

Digital Image Processing

Final Comments on the Training Stage

Final Comments

Image classifiers are designed for efficiency (speed) *and* accuracy.

The training stage, however, must be conducted in a manner that produces maximum accuracy. Introducing shortcuts at the training stage is likely to produce poor results in the final classification.

Final Comments

The most difficult part of the training process is not the development of spectral classes for distinctly different land cover types such as water, forest and agriculture.

Instead, the problems arise in developing spectral classes that will assign pixels in transition zones and areas of mixed cover types to the appropriate land cover classes.

Final Comments

By definition, transition zones are areas containing elements of the cover types that are more homogenous on either side of the transition zone.

As a result, spectral signatures often have to be developed separately for transition zones and the analyst needs to develop a consistent classification scheme that will either assign these areas to one or another primary cover type or to a transitional cover type.

Final Comments

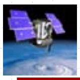
Refining spectral classes is often a trial-and-error process where the analyst adds or removes pixels from a developing training class to test the effect that the modification has on the ability of the classifier to produce acceptable results.

Sometimes it is necessary to remove rarely-occurring land cover type from the classification scheme in order to avoid misclassifying pixels that belong to more common land cover types.

Final Comments

The training stage may also make it apparent that some of the originally-proposed land cover classes will need to be combined into more general classes because the spectral responses of the proposed classes are indistinguishable.

For example, the data may not make it possible to distinguish individual tree species even though they are adequate to classify trees into more general categories such as evergreen, deciduous and mixed forests.



Final Comments

Sometimes, the inherent spectral responses of similar cover types may make it impossible to separate them using a single multiband image.

In these cases, it may be necessary to acquire additional data in the form of field data or images acquired on other dates or with other sensors, or to use other classification techniques such as visual interpretation of the digital image or of other images including higher resolution digital images or aerial photography.