Image Classification II

Supervised Classification

- Using pixels of known classes to identify pixels of unknown classes
- Advantages
  - Generates information classes
  - Self-assessment using training sites
  - Training sites are reusable
- Disadvantages
  - Information classes may not match spectral classes
  - Signature homogeneity of information classes varies
  - Signature uniformity of a class may vary
  - Difficulty and cost of selecting training sites
  - Training sites may not encompass unique spectral classes

Supervised Classification Procedures

- Determines a classification scheme
- Selects training sites on image
- Generates class signatures
- Evaluates class signatures
- Assigns pixels to classes using a classifier

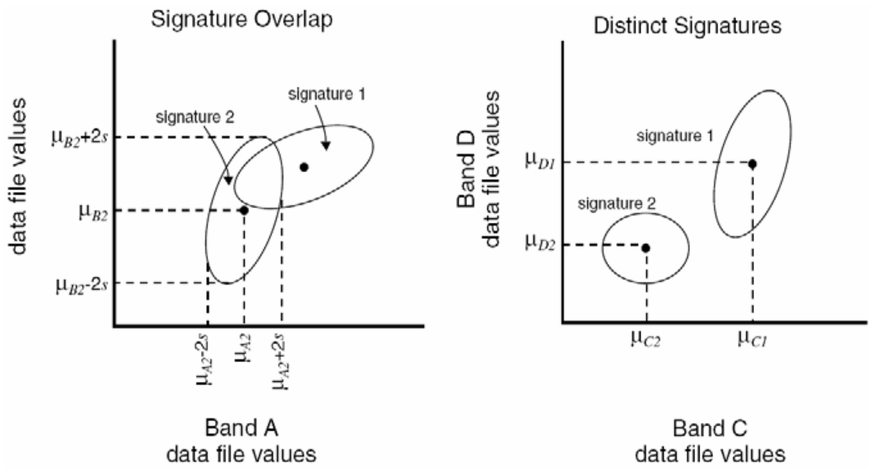
Training Site Selection

- Number of pixels (at least 100 per class)
- Individual training sites should not be too big (10 to 40 pixels per site)
- Sites should be dispersed throughout the image
- Uniform and heterogeneous sites

Signature Evaluation

- Alarm (i.e., preview using parallelepiped classifier)
- Ellipse (mean & stdv)
- Contingency matrix (based on pixels within training sites)
- Separability
  - Euclidean Distance
  - Divergence
  - Transformed Divergence (0 – 2000, > 1700)
  - Jefferies-Matusita Distance (0 – 1414)
- Statistics & histograms
  - Small variations preferred

Ellipse Evaluation



Contingency Matrix

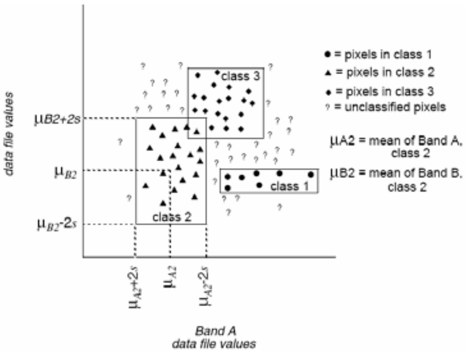
```
ERROR MATRIX
-------------

                    Reference Data
                    --------------
Classified
      Data           Veg        Nonveg     Row Total
----------    ----------    ----------    ----------
       Veg          2395             0          2395
    Nonveg             1          1279          1280

Column Total        2396          1279          3675
```

Classifiers

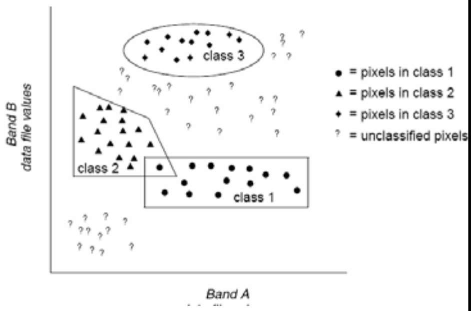
- Nonparametric (faster then parametric classifiers)
  - Parallelepiped
  - Feature space
- Parametric
  - Minimum distance (Euclidean spectral distance): least accurate, most efficient
  - Mahalanobis distance (Euclidean distance + covariance - normal distribution of DN is assumed).
  - Maximum likelihood (Bayesian prob. - normal distribution is assumed): most accurate, least efficient

Nonparametric

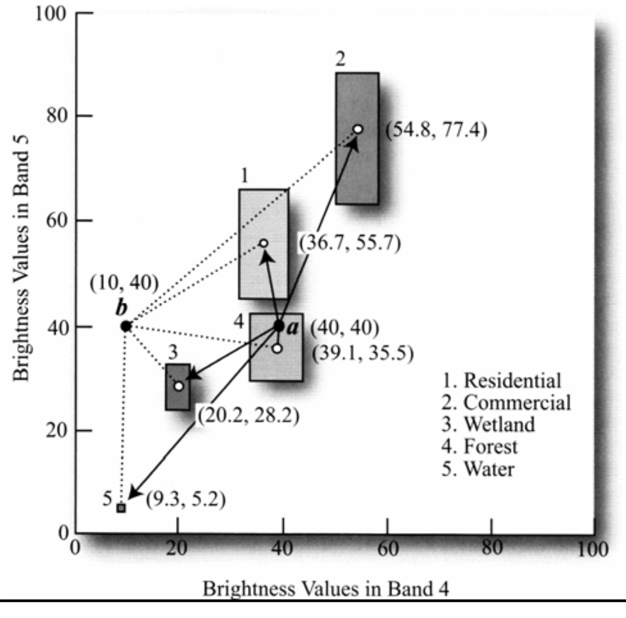
Parallelepiped

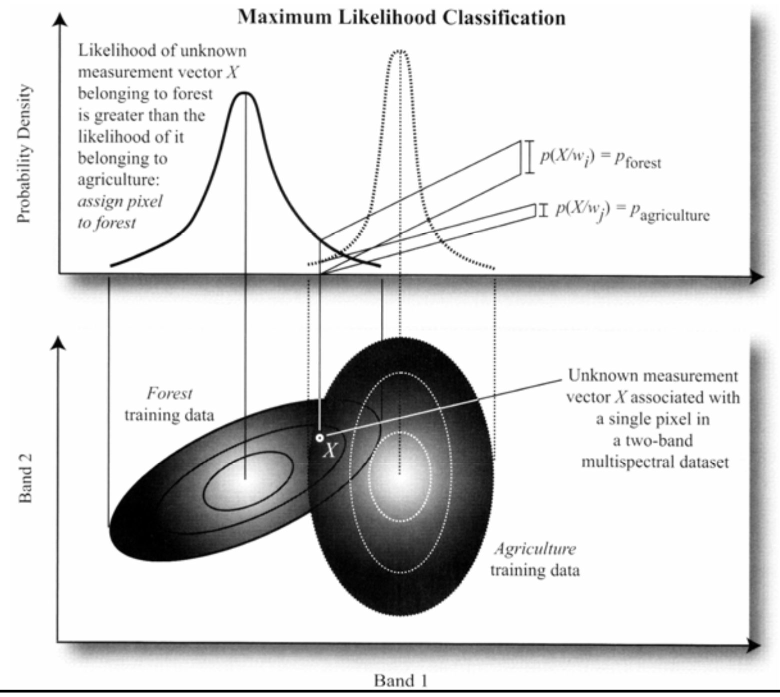


Feature space



Minimum Distance Classifiers





Maximum Likelihood Classification

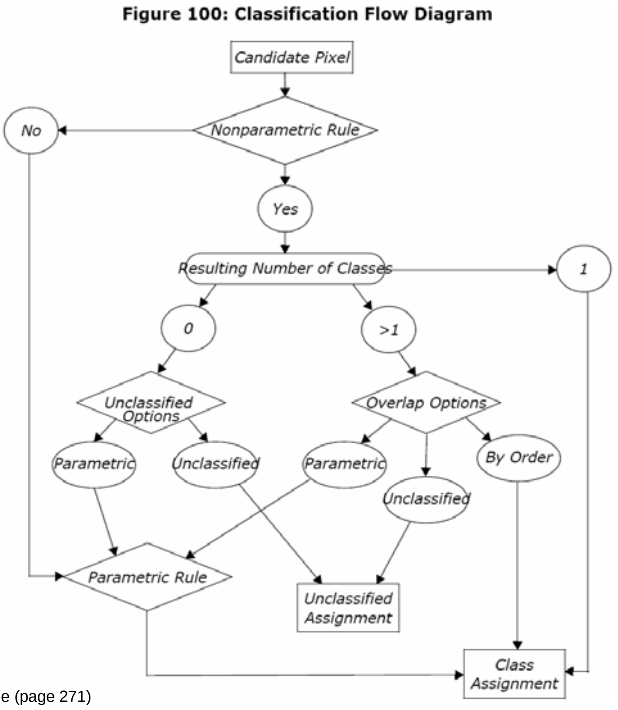


Figure 100: Classification Flow Diagram

ERDAS Imagine Field Guide (page 271)

Classification Terminology

- Features/Feature space/Dimensionality
  - Spectral bands, **textures**, indices, **ancillary GIS layers**…
- Classification schemes
  - Taxonomically correct definitions of classes organized according to logical criteria (e.g., 1976 USGS Anderson's classification)
- Signatures
  - Information classes and spectral classes
- Training sites (fields)
  - Areas for extracting signatures of information classes
- Classifier
  - A procedure to assign pixels to classes based on pixel features and class signatures

---

What are land-use and land cover?

Land-use

- Human activity on, and intention for, the land

Land-cover

- The biophysical characteristics of the landscape

Texture Analysis

- Texture
  - Apparent roughness or smoothness of an image region (Campbell 2002)
  - Frequency of tonal change on an image (Lillesand and Kiefer 2004)
  - Natural scenes containing semi-repetitive arrangements of pixels (Pratt 1991)
- Texture analysis
  - Feature-based (first-order, second-order statistics)
  - Model-based
  - Structural

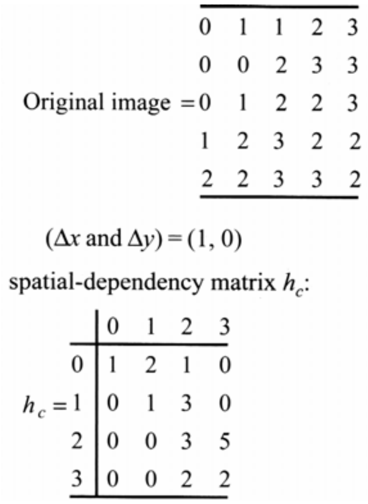
First-Order Texture Statistics

- Stats based on a moving window
  - Mean
  - Standard Deviation
  - Range
  - Entropy

Second-order Texture Statistics

- Stats based on paired pixels
  - Variogram
  - Fourier Analysis
  - Gray-Level Co-occurrence Matrices

GLCM

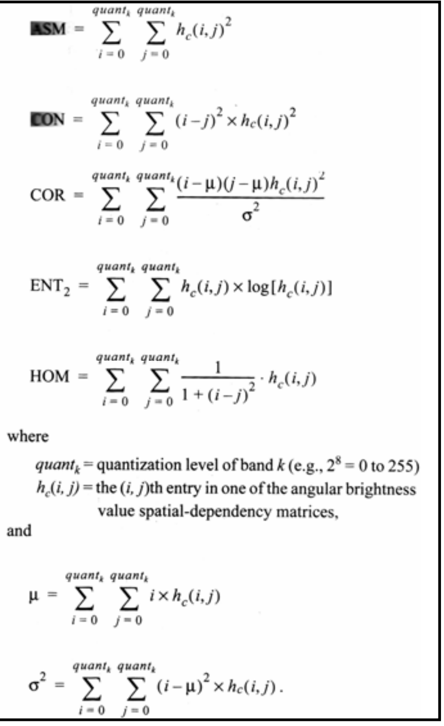


$$\text{Original image} = \begin{array}{ccccc} 0 & 1 & 1 & 2 & 3 \\ 0 & 0 & 2 & 3 & 3 \\ 0 & 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 2 & 2 \\ 2 & 2 & 3 & 3 & 2 \end{array}$$

$(\Delta x \text{ and } \Delta y) = (1, 0)$

spatial-dependency matrix h_c:

$$h_c = \begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 3 & 0 \\ 2 & 0 & 0 & 3 & 5 \\ 3 & 0 & 0 & 2 & 2 \end{array}$$



$$\text{ASM} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i,j)^2$$

$$\text{CON} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} (i-j)^2 \times h_c(i,j)^2$$

$$\text{COR} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} \frac{(i-\mu)(j-\mu)h_c(i,j)^2}{\sigma^2}$$

$$\text{ENT}_2 = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i,j) \times \log[h_c(i,j)]$$

$$\text{HOM} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} \frac{1}{1+(i-j)^2} \cdot h_c(i,j)$$

where

$quant_k$ = quantization level of band k (e.g., $2^8 = 0$ to 255)

$h_c(i, j)$ = the (i, j)th entry in one of the angular brightness value spatial-dependency matrices,

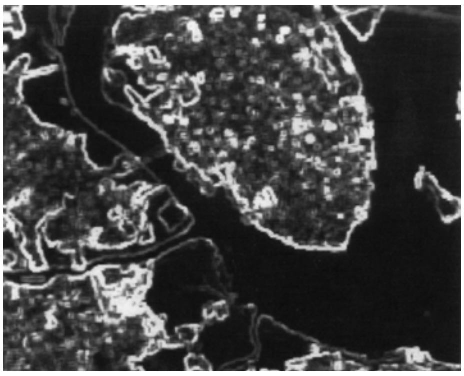
and

$$\mu = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} i \times h_c(i,j)$$

$$\sigma^2 = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} (i-\mu)^2 \times h_c(i,j)\,.$$

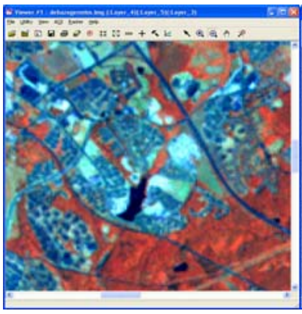
Textural Classification

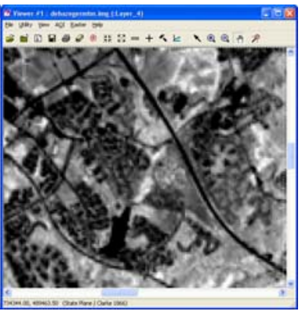
TM Band 4 (NIR)

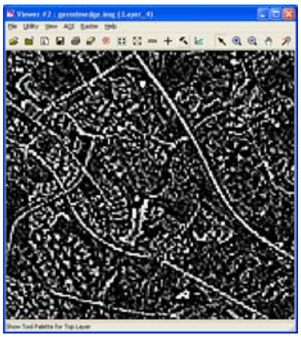


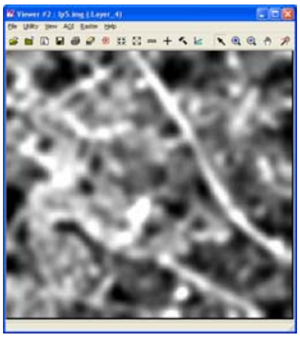
STDV Texture on Band 4

Texture Information









Ancillary Data

- Increase feature space dimensionality so that information classes can be more easily separated
  - For example, elevation and vegetation distribution
- (In)compatibility
  - Physical (data format, resolution, etc.)
  - Logical (do the values used to define the feature space make sense?)

Ancillary Data (cont.)

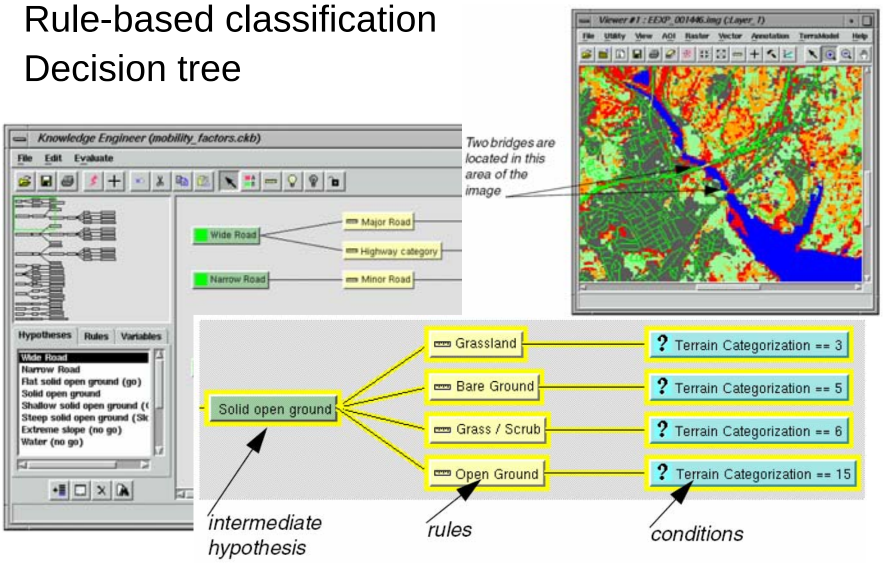
- Stratification
- a priori probability in maximum likelihood classifier
- Contextual classification
- Post-classification sorting (rule-based pixel class adjustment)

Classification Using Ancillary Data

1. Correct for slope and aspect effects and then do classification
2. Classify with aspect data masked using elevation and slope criteria

Contextual Classification

- Rule-based classification
- Decision tree



Final Remarks

- No classification method is inherently superior to any other.
- The process guideline varies among images
- In general, one should generate 10 ~ 15 spectral classes for each intended information class in unsupervised classification (e.g., 20 ~ 30 spectral cls for 2 info cls)
- When determining info class in supervised classification, one should also consider their spectral heterogeneity (e.g., agricultural might include fallow and vegetated fields)

Mixels

- Pure & composite signatures
- Where do mixels occur?
- Are mixels good or bad?

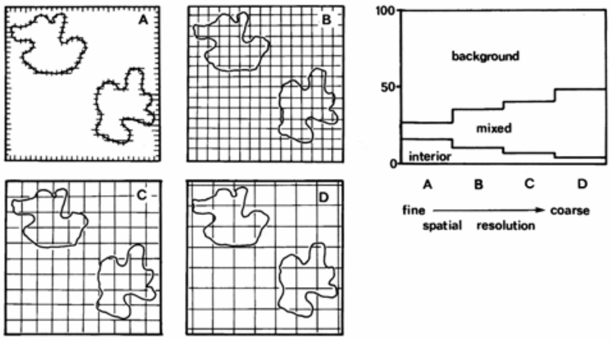
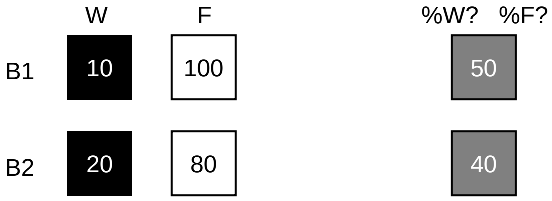


FIGURE 9.7. Influence of spatial resolution on proportions of mixed pixels.

Spectral Unmixing

Pure end members



	W	F	%W? %F?
B1	10	100	50
B2	20	80	40

A_w x 10 + A_f x 100 = 50

A_w x 20 + A_f x 80 = 40

End members:
Tree, soil, water, grass, shadow …

Hard & Fuzzy Classification Schemes

- Hard
  - Mutually exclusive classes
  - Exhaustive
  - Hierarchical
- Fuzzy
  - Fuzzy (e.g., upland forest, forested wetland, water)

Fuzzy Classification

1. Bayes' Theorem and Maximum Likelihood Classification
2. Fuzzy Signature Development
3. Soft Classifier (evaluation of probability that a pixel is a member of a class)
4. Hardeners (forcing decision of class membership)

Fuzzy Signatures

- Training sites (homogeneous vs. fuzzy)
- Fuzzy partition matrix

	Class#1	Class#2	Class#3	…
Site#1	0.7	0.2	0.1	0
Site#2	0.2	0.2	0.4	0…
Site#3	0.5	…		
…	…			