Satellite Imagery Classification

Lillesand, Chipman & Kiefer, 2015.

February 26, 2026

Outline

Plan

Spectral Pattern Recognition

- **The Core Objective:** Automatically categorize pixels into land cover classes or "themes".
- **The "Individualist" Approach:**
  - Classification is performed **pixel-by-pixel**.
  - Based on spectral reflectance or emissivity signatures.
  - Crucial Note: It ignores the pixel's neighbors or surroundings.
- **Extending the Model:**
  - **Polarization:** Radar imagery.
  - **Temporal:** Multitemporal sequences.
  - **Bidirectional:** Multi-angle imagery (e.g., MISR).

Spatial Pattern Recognition

- **The "Contextual" Approach:** Categorizes pixels based on their relationship with surrounding pixels.
- **Key Variables Considered:**
  - Texture, proximity, size, and shape.
  - Directionality, repetition, and context.
- **Human Synthesis:** Attempts to replicate the visual interpretation process of a human analyst.
- **Complexity:** More computationally intensive than spectral methods.
- **Hybrid Models (OBIA):** Object-Based Image Analysis combines both spectral and spatial logic.

Supervised Classification

The analyst acts as the teacher for the algorithm.

1. **Training Stage:** The analyst identifies representative sample sites of known cover types (**Training Areas**).
2. **Interpretation Key:** These areas are used to compile numerical descriptors of spectral attributes.
3. **Classification:** The computer compares every unknown pixel to the key.
4. **Labeling:** The pixel is assigned to the category it "looks most like" based on statistical strategies.

Unsupervised and Hybrid Classification

Unsupervised Classification

- **Data-First:** Pixels are grouped into "natural" clusters based on spectral similarity.
- **Post-Classification Identity:** The analyst labels the clusters *after* the grouping is done using ground reference data.

Hybrid & Specialized Procedures

- **Hybrid:** Combines supervised and unsupervised steps for better efficiency.
- **Advanced Topics:** Neural networks, mixed pixel analysis, and specialized hyperspectral procedures.

"There is no single 'right' manner... approach depends on data nature and intended application."

Plan

Data Input and Spectral Patterns

- **Sensor Agnostic:** Procedures apply to airborne multispectral data and satellite platforms (Landsat, SPOT, WorldView-2).
- **Multichannel Framework:** Analysis typically involves multiple spectral bands, for example:
  - Visible: Blue, Green, Red
  - Infrared: Near-Infrared (NIR), Thermal Infrared
- **Digital Numbers (DNs):** The sensor measures scene radiance recorded as DNs for each pixel across all bands.
- **Basis for Classification:** If spectral response patterns are sufficiently distinct for different terrain features, they form a “signature” for automated categorization.

The Three-Step Procedure

Supervised classification follows a structured three-stage workflow:

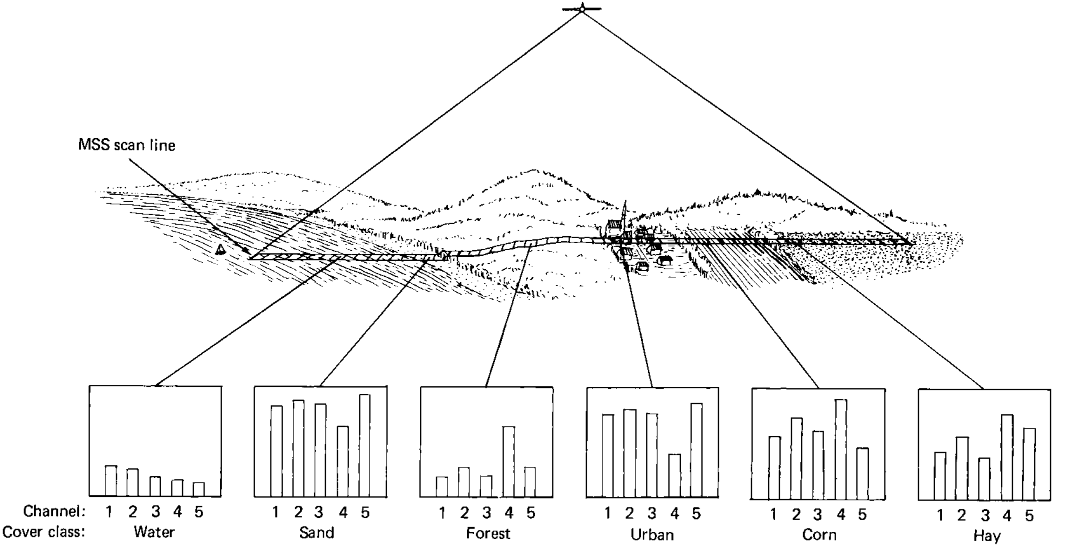
1. **Training Stage:** The analyst identifies representative "training areas" to develop numerical descriptions of the spectral attributes for each land cover type.
2. **Classification Stage:** Each pixel is compared to the training data and categorized into the class it most closely resembles.
3. **Output Stage:** Results are compiled into final products for end-users.

Pixels that do not sufficiently match any training set are labeled as ***"unknown".***

Classification Stage and Output Products

- **The "Heart" of the Process:** The classification stage uses computer-based **decision rules** to evaluate spectral patterns.
- **Versatile Output Formats:**
  - **Thematic Maps:** Visual representations of land cover.
  - **Statistical Tables:** Summary data for each class.
  - **Digital Data Files:** Classified data can be exported directly into a **Geographic Information System (GIS)**.
- **Strategic Integration:** Classification output often becomes GIS input, enabling complex spatial modeling and decision-making.

Multispectral Scan Line Profile



Basic steps

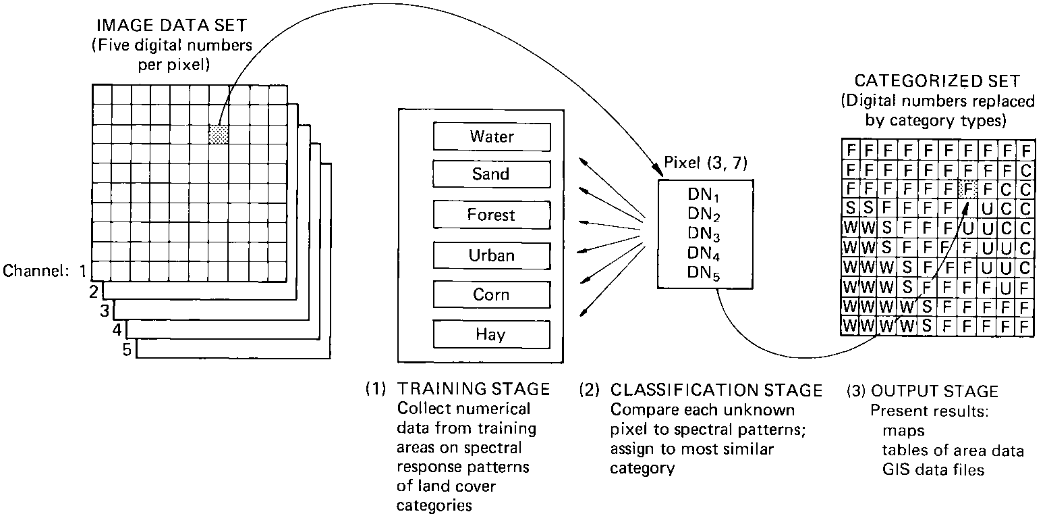


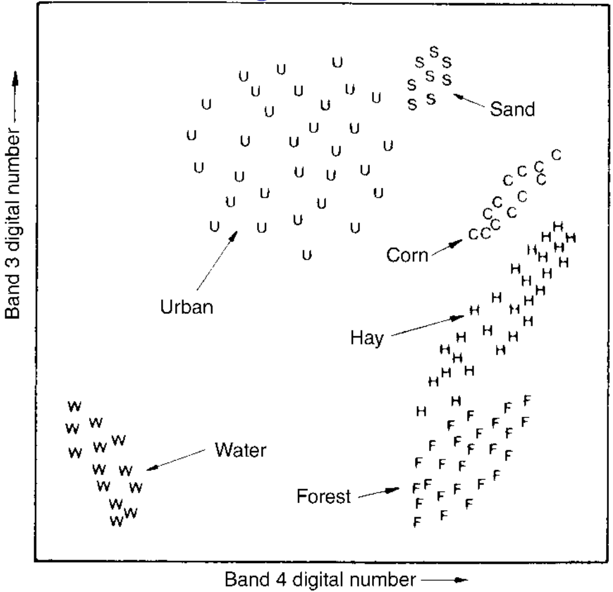
IMAGE DATA SET
(Five digital numbers per pixel)
Channel: 1
2
3
4
5
Water
Sand
Forest
Urban
Corn
Hay
Pixel (3, 7)
DN1
DN2
DN3
DN4
DN5
CATEGORIZED SET
(Digital numbers replaced by category types)
(1) TRAINING STAGE
Collect numerical data from training areas on spectral response patterns of land cover categories
(2) CLASSIFICATION STAGE
Compare each unknown pixel to spectral patterns; assign to most similar category
(3) OUTPUT STAGE
Present results:
maps
tables of area data
GIS data files

Plan

The Concept of Measurement Space

- **Measurement Vectors:** Each pixel is represented as a point in a multi-dimensional space.
- **Scatter Plots:** For a 2-band subset, DNs from Band 3 (y-axis) and Band 4 (x-axis) define a pixel's coordinates.
- **Spectral Clusters:** Training data forms "clouds of points" rather than single values.
- **Natural Variability:** These clouds illustrate the centralizing tendency and inherent variance of specific land cover classes.

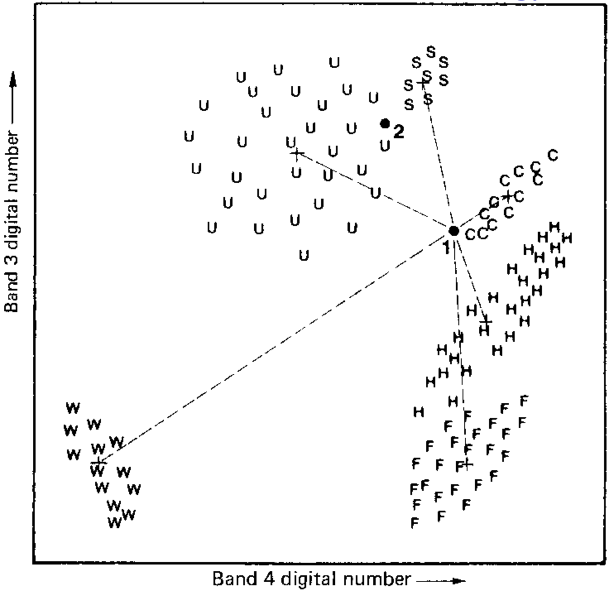
Pixel observations to scatter diagram



Minimum-Distance-to-Means Classifier

- **Mechanism:**
  1. Calculate the **mean vector** (average spectral value) for each category.
  2. Compute the Euclidean distance from an unknown pixel to each mean.
  3. Assign pixel to the "closest" class.
- **Pros:** Mathematically simple and computationally efficient.
- **Cons:** Insensitive to different degrees of **variance**.
- Limitation: A pixel may be mathematically closer to a "tight" cluster (like sand) even if it logically belongs to a "wide" cluster (like urban).

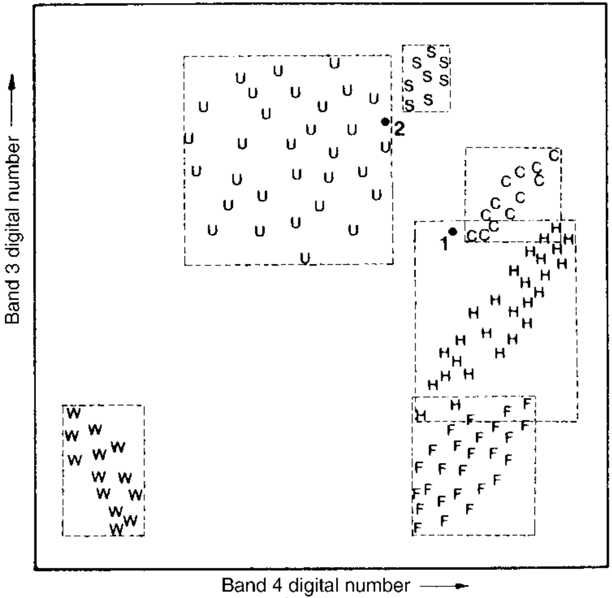
Minimum distance to means classification strategy



Parallelepiped Classifier

- **Mechanism:** Uses the range (min/max) of DNs in each band to create “decision regions” (rectangles in 2D, parallelepipeds in n-D).
- **Sensitivity:** Accounts for category variance by sizing the boxes differently.
- **The Overlap Problem:**
  - If categories overlap, pixels are labeled “not sure”.
  - Fails to account for **covariance** (slanted clouds where bands vary together).
- **Performance:** Very fast, but often poor for highly correlated spectral response patterns.

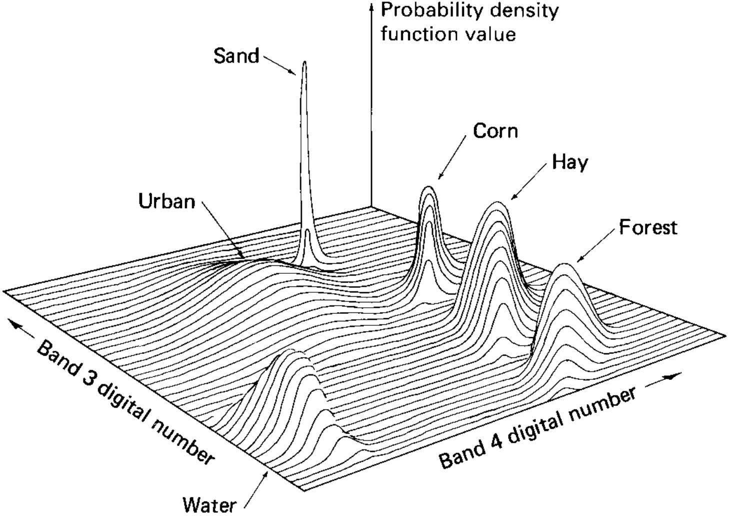
Parallelepiped classification strategy



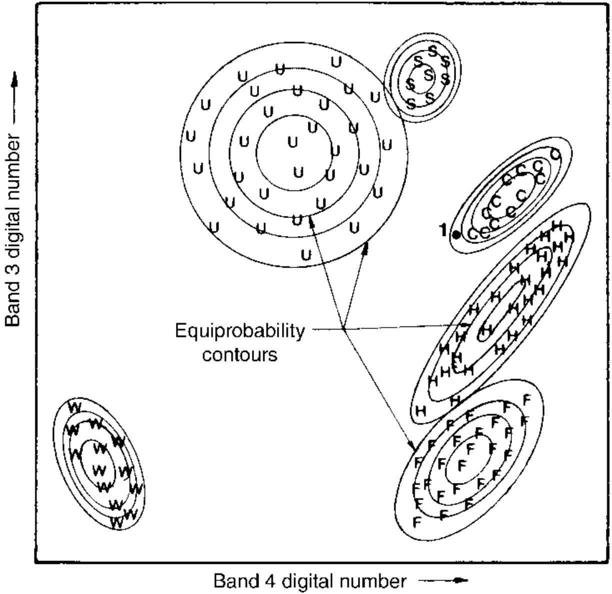
Gaussian Maximum Likelihood Classifier (GMLC)

- **The Probabilistic Leap:** Evaluates both variance and covariance by assuming a **Gaussian (Normal)** distribution.
- **Parameters:** Completely described by the mean vector μ and the covariance matrix Σ.
- **Mechanism:** Calculates the statistical probability of a pixel belonging to each class; assigns it to the "most likely" one.
- **Equi-probability Contours:** Delineates ellipsoidal decision regions that follow the "slant" of the data.

Probability density by a maximum likelihood classifier.



Equiprobability contours by a maximum likelihood classifier.



Bayesian Classifier & Cost Functions

The Bayesian approach refines GMLC by adding two weighting factors:

1. **A Priori Probability:** The anticipated likelihood of a class occurring in the scene (e.g., urban is more likely than rare sand).
2. **Cost of Misclassification:** A weight applied to minimize the "damage" of specific errors.

The Ideal Goal

To minimize the overall "cost" of errors, resulting in a theoretically optimum classification.

Efficiency vs. Accuracy

- **The Drawback of GMLC:** Large number of computations, especially with many bands.
- **Dimensionality Reduction:** Using Principal Component transformations to speed up the process.
- **Decision Tree (Layered) Classifiers:**
  - A "multi-step" approach.
  - Separate simple classes (like water) first using basic thresholds.
  - Reserve complex GMLC logic only for "ambiguous" overlapping classes.

Plan

The Analyst's Critical Role

- **Manual vs. Automated:** While classification is automated, training is a manual process requiring "art and science".
- **The Objective:** To define the location, size, and orientation of the "clouds of points" (spectral signatures) for each class.
- **Requirements:**
  - Close interaction with image data.
  - Substantial reference data.
  - Thorough knowledge of the geographic area.
- Success Factor: The quality of training determines the value of the entire classification effort.

Information Classes vs. Spectral Classes

- **Information Class:** The desired category (e.g., "Water" or "Agriculture").
- **Spectral Class:** The distinct spectral subgroups within an information class.
- **The "Representative" Rule:**
  - One information class often requires **multiple** spectral classes.
  - **Example:** "Water" might include clear water, turbid water, and deep water (each requiring its own training set).
  - **Complexity:** Agriculture may vary by planting date, soil moisture, and crop variety.

Methods of Delineation

Manual Polygons:

- Analyst draws boundaries around known areas.
- **Edge Avoidance:** Avoid pixels on boundaries or "rough" areas to ensure purity.

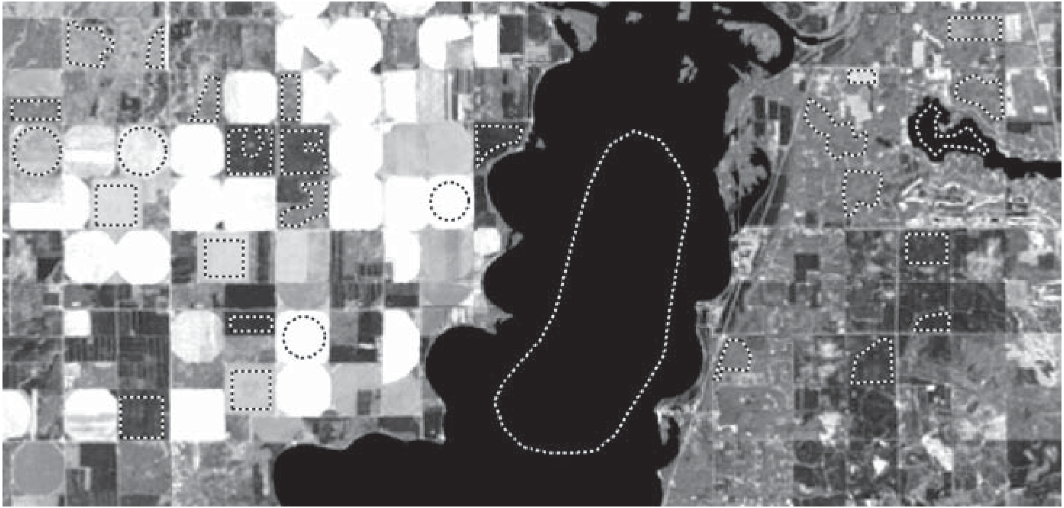
Seed Pixel Approach:

- Analyst selects a representative "seed".
- Algorithm grows the area by including contiguous pixels with similar spectral traits.

Statistical Requirements

- **Sample Size (n = number of bands):**
  - Theoretical minimum: $n + 1$ pixels.
  - **Practical Rule:** Use $10n$ to $100n$ pixels per training set.
- **Spatial Dispersion:**
  - It is better to have many small sites (e.g., 20 sites of 40 pixels) than one giant site (1 site of 800 pixels).
  - Dispersion captures scene-wide variations.
- **Mathematical Basis:** These pixels generate the **mean vector** and **covariance matrix** for the classifier.

Training area polygons



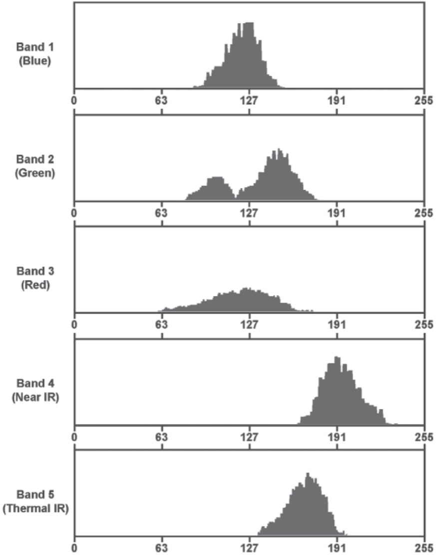
The Refinement Process

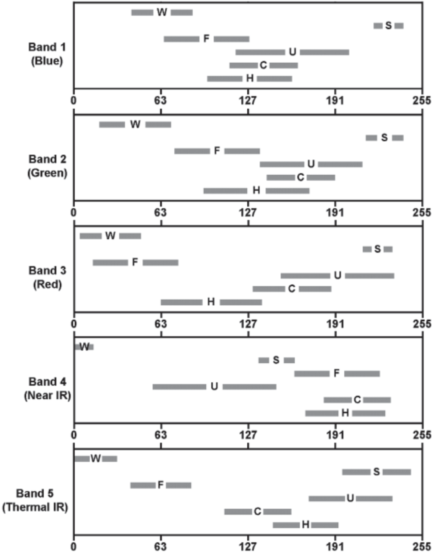
- **Statistical Testing:** Ensure data is **unimodal** and Gaussian (normally distributed).
- **Cleaning Data:**
  - Delete "edge pixels" or outliers (e.g., bare soil found within a crop field).
  - Split bimodal distributions into two distinct spectral classes.
- **Spectral Separability:** Evaluate if training sets for different classes overlap too much.
- **Final Goal:** A clean, non-redundant set of "interpretation keys" for the computer to follow.

Graphical Analysis: Histograms

- **Visual Normality Check:** Essential for Maximum Likelihood classifiers to ensure data follows a Gaussian distribution.
- **Identifying Subclasses:**
  - A **bimodal distribution** (two peaks) suggests the training site contains two distinct subclasses.
  - *Example:* "Hay" might actually be two different varieties or have different illumination conditions.
- **Action:** Split bimodal classes into separate spectral categories to improve overall accuracy.

Visualization of training plot data

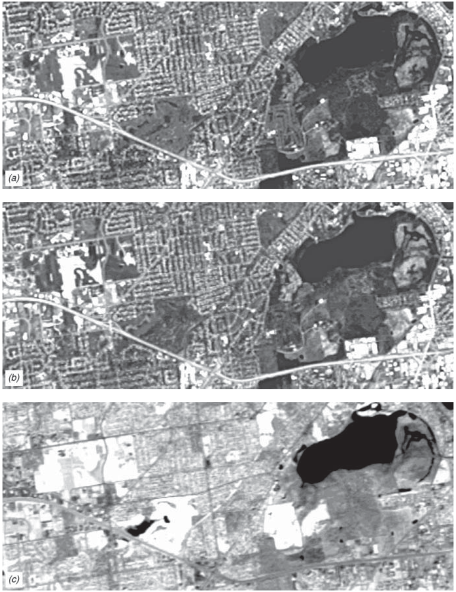




Coincident Spectral Plots

- **Comparison Tool:** Unlike histograms, these allow for direct comparison between different category types.
- **Visualization:** Displays mean response and variance (± 2 standard deviations) for each band.
- **Key Insight:** Reveals spectral overlap between classes (e.g., Hay vs. Corn).
- **Band Selection:** Helps identify which bands offer the best "reversals" or gaps for class separation.

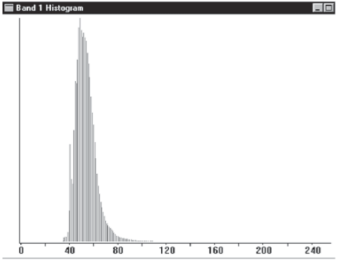
Multispectral Images

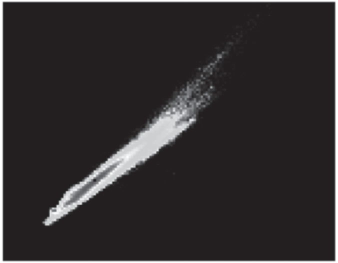


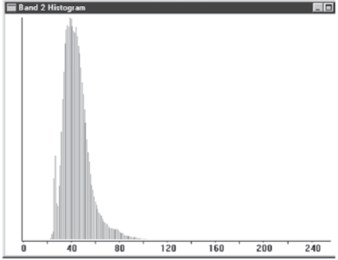
Scatter Diagrams (Measurement Space)

- **Multi-Band View:** Captures relationships that single-band histograms miss.
- **Correlation Analysis:**
  - **High Correlation:** Near-linear "cloud of points" (e.g., Visible Green vs. Red). Harder to separate classes.
  - **Low Correlation:** Expanded "measurement space" (e.g., Red vs. Near-IR). Much easier to separate land cover types.
- **Practical Application:** Often, just two well-chosen, low-correlation bands are enough for generalized classification.

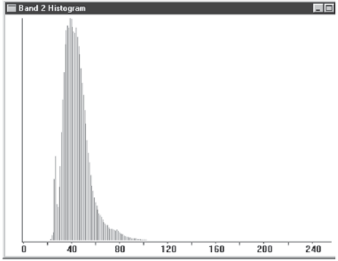
Histograms and two-dimensional scatter diagram

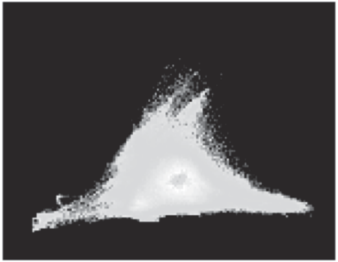


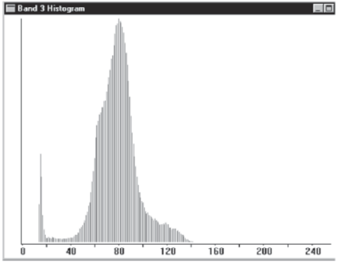




Histograms and two-dimensional scatter diagram







Quantitative Separability Measures

- **Statistical Distance:** Computing a "matrix of divergence" between all class pairs.
- **Transformed Divergence (TD):** A covariance-weighted distance.
  - Values > 1500 generally indicate good separation.
  - Values < 1500 suggest problematic spectral overlap.
- **Jeffries-Matusita (JM) Distance:** Similar to TD but with a fixed scale (maximum 1414).
- **Diagnosis:** If overlap occurs between *different* information classes (e.g., Hay and Corn), refinement is needed.

Divergence Matrix

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.

Self-Classification of Training Data

- **The “Preliminary Test”:** Classifying only the training pixels to see if the computer assigns them correctly.
- **The Error Matrix:** Displays percentages of “as expected” vs. “misclassified” training pixels.
- The Warning: Do NOT confuse this with final map accuracy.
  - Training areas are “pure” examples; the rest of the scene is usually much “messier”.
  - High training accuracy does not guarantee high scene-wide accuracy.

Interactive Preliminary Classification

- **Real-Time Feedback:** Using efficient algorithms (like parallelepiped) to highlight areas the current statistics would capture.
- **Visual Approximation:** Highlighting "classified" pixels in color over the raw imagery.
- **Representative Subsets:** Classifying a small, diverse section of the image first to verify the logic before committing to the full scene.
- **Iteration:** "Trial and error" testing of alternative deletions and poolings of training classes.

Tactics for Difficult Overlaps

When spectral classes refuse to separate, the analyst has several choices:

1. **Merger/Aggregation:** Combine specific classes into broader categories (e.g., Merge “Birch” and “Aspen” into “Deciduous”).
2. **Deletion:** Eliminate rare problem classes to preserve the accuracy of extensive, similar classes.
3. **Trial and Error:** Adjust sample sizes or re-evaluate the identity of training sites.

“If classes are inherently similar, no amount of retraining will make them separable!”

Multi-Image and Ancillary Data

- **The Multi-Image Problem:** Atmospheric and sun-angle changes mean "signatures" from one date rarely work on another.
- **Radiometric Normalization:** Matching secondary images to a "base" image to allow signature extension.
- **The Hybrid Future:**
  - Incorporating **GIS data** (elevation, soil type).
  - Using **Multitemporal** data (seasonal changes).
  - **Spatial Pattern Recognition** to mimic human visual interpretation.