MULTISPECTRAL CLASSIFICATION

DEFINITIONS
MULTISPECTRAL CLASSIFICATION

- **Thematic maps**
  
  Informational description of a given area
  
  *Symbolic labels at each pixel replace numeric DNs in images*

- **Themes**
  
  For example: soil, water, vegetation
  
  *Desire direct relation between themes and geophysical variables, but a strong connection does not always exist*

  Also commonly called *classes*

- **Physical Models**
  
  Explain some within-class variation in “signatures,” e.g. due to topographic shading

  Remaining variation attributed to

  - natural variation in signatures
  - spectral mixing within GIFOV
MULTISPECTRAL CLASSIFICATION

THE CLASSIFICATION PROCESS

Feature Extraction

- Example features

  original multispectral bands

  subset of bands

  derived features

    - principal components
    - vegetation indexes
    - spatial properties

- Good features help to suppress or correct for known physical sources of variation

  For example, multispectral ratios can reduce topographic shading
MULTISPECTRAL CLASSIFICATION

Data flow in classification

scene

atmosphere → sensor → multispectral image

feature extraction

feature image

training

determine discriminant functions

extract training pixels

classifier

K-D feature space

labeling

thematic map

Dr. Robert A. Schowengerdt, schowengerdt@ece.arizona.edu, 520-621-2706 (voice), -8076 (fax)
MULTISPECTRAL CLASSIFICATION

- Feature space

Because of class variability, the “signature” is actually a statistical distribution of feature vectors.

Successful classification requires separated distributions, i.e. minimal overlap.

Example with 2-D feature space, e.g. two bands of a multispectral image:

![Diagram showing high and low separability in feature space](image-url)
MULTISPECTRAL CLASSIFICATION

The Importance of Image Scale and Resolution

• Different classes require different sensor GIFOVs for identification

  For example, man-made categories (streets, buildings, etc) require GIFOVs of 10m or less

  Sea surface temperature, or global NDVI, can be measured sufficiently with larger GIFOVs, such as 1km

• Themes are often organized into a hierarchy

  Higher levels in the hierarchy generally require higher resolution (smaller GIFOVs)

  Different classes and hierarchies can be defined for different applications
## MULTISPECTRAL CLASSIFICATION

### Anderson land-use and land-cover labeling scheme for Levels I and II

<table>
<thead>
<tr>
<th>Level I</th>
<th>Level II</th>
</tr>
</thead>
</table>
| 1 urban/built-up land | 11 residential  
12 commercial and services  
13 industrial  
14 transportation, communication, and utilities | 15 industrial and commercial complexes  
16 mixed urban/built-up land  
17 other urban/built-up land |
| 2 agricultural land | 21 cropland and pasture  
22 orchards, groves, vineyards, nurseries and ornamental horticultural areas | 23 confined feeding operations  
24 other agricultural land |
| 3 rangeland | 31 herbaceous rangeland  
32 shrub and brush rangeland | 23 mixed rangeland |
| 4 forest land | 41 deciduous forest land  
42 evergreen forest land | 43 mixed forest land |
| 5 water | 51 streams and canals  
52 lakes | 53 reservoirs  
54 bays and estuaries |
| 6 wetland | 61 forested wetland  
62 non-forested wetland | |
| 7 barren land | 71 dry salt flats  
72 beaches  
73 sandy areas other than beaches  
74 bare exposed rock | 75 strip mines, quarries, and gravel pits  
76 transitional areas  
77 mixed barren land |
| 8 tundra | 81 shrub and brush tundra  
82 herbaceous tundra  
83 bare ground tundra | 84 wet tundra  
85 mixed tundra |
| 9 perennial snow or ice | 91 perennial snowfields | 92 glaciers |
MULTISPECTRAL CLASSIFICATION

*Anderson land-use and land-cover labeling scheme for Level III*

<table>
<thead>
<tr>
<th>Level II</th>
<th>Level III</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 residential</td>
<td>111 single family units</td>
</tr>
<tr>
<td></td>
<td>112 multifamily units</td>
</tr>
<tr>
<td></td>
<td>113 group quarters</td>
</tr>
<tr>
<td></td>
<td>114 residential hotels</td>
</tr>
<tr>
<td></td>
<td>115 mobile home parks</td>
</tr>
<tr>
<td></td>
<td>116 transient lodgings</td>
</tr>
<tr>
<td></td>
<td>117 other</td>
</tr>
</tbody>
</table>
MULTISPECTRAL CLASSIFICATION

Hard Versus Soft Classification

- Labeling of pixels accomplished by partitioning the feature space
- **Hard** classification results in one class/pixel
- **Soft** classification results in multiple classes/pixel, with associated likelihoods for each class:
  - "probabilities" parametric statistical classifiers
  - "membership grades" fuzzy classifiers
  - "fractions" mixture analysis
- **Soft** classification is more descriptive of reality
  - Accomodates within- and between-class variation
  - Accomodates class mixing within GIFOV
MULTISPECTRAL CLASSIFICATION

Visualization of hard versus soft classification

multispectral image  \( \rightarrow \)  \( K \)-dimensional feature space  \( \rightarrow \)  thematic map

hard classification

soft classification
MULTISPECTRAL CLASSIFICATION

Supervised and Unsupervised Training

- **Must** "teach" the classifier algorithm to recognize the classes of interest

- **Use** sample pixels to represent the classes ("prototypes," "exemplars," or "training samples")

- **Training**

  *derive features from sample pixels to determine decision boundaries between classes, according to the selected classification criteria*
MULTISPECTRAL CLASSIFICATION

Supervised Training

• Training samples are labeled

  Ground truth

  Interpretation of higher resolution imagery (e.g. aerial photography)

  Existing maps

• Identifying training sites (pixel samples) for each class can be laborious

  Desire class-homogeneous sites, without mixture among classes

  Also desire representation of full within-class variability

  Therefore, often need more than one site/class

  Some classes may have small spatial extent, e.g. “asphalt” roads
MULTISPECTRAL CLASSIFICATION

- Separability analysis

Useful for pre-classification analysis of supervised class signatures

Desire some measure of interclass similarity (or dissimilarity)

3 similarity measures based on class means

Dr. Robert A. Schowengerdt, schowengerdt@ece.arizona.edu, 520-621-2706 (voice), -8076 (fax)
### MULTISPECTRAL CLASSIFICATION

**Similarity measures using class means only or class means and variances/covariances**

<table>
<thead>
<tr>
<th>name</th>
<th>formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>city block</td>
<td>[ L_1 =</td>
</tr>
<tr>
<td>Euclidean</td>
<td>[ L_2 = |\mu_a - \mu_b| = \left[ (\mu_a - \mu_b)^T (\mu_a - \mu_b) \right]^{1/2} ] [ = \left[ \sum_{k=1}^{K} (m_{ak} - m_{bk})^2 \right]^{1/2} ]</td>
</tr>
<tr>
<td>angular</td>
<td>[ ANG = \arccos \left( \frac{\mu_a^T \mu_b}{|\mu_a| |\mu_b|} \right) ]</td>
</tr>
<tr>
<td>normalized city block</td>
<td>[ NL_1 = \sum_{k=1}^{K} \frac{</td>
</tr>
<tr>
<td>Mahalanobis</td>
<td>[ MH = \left[ (\mu_a - \mu_b)^T \left( \frac{C_a + C_b}{2} \right)^{-1} (\mu_a - \mu_b) \right]^{1/2} ]</td>
</tr>
</tbody>
</table>
Separability analysis can be used to find the best subset of features for given training data

- Calculate average separability among classes, using all possible combinations of \( k = 2 \) features, 3 features, etc. (\( k \leq K \))
- Choose the \( k \) features (out of \( K \) total features) that have highest separability for the given classes
- Can reduce computation load
MULTISPECTRAL CLASSIFICATION

Unsupervised Training

- **Training samples are unlabeled**
  
  Sites need not be homogeneous, can use entire image if desired

- **Algorithm used to find “natural” groupings of similar data, i.e. “clusters”**
  
  Resulting clusters are labeled by analyst using any available information (as in supervised training)

  Clusters seldom are true concentrations of data, but just optimally-partitioned regions within feature space

- **Data-driven, whereas supervised training is analyst-driven**
  
  Clustering can be helpful in finding homogeneous image areas for supervised training sites
MULTISPECTRAL CLASSIFICATION

• **K-means clustering algorithm**

  **Iterative procedure**

  Adjusts cluster partitioning such that mean-squared distance of training samples from cluster means is minimized

  **Example convergence criteria**

  • **net mean migration** < threshold

  • \[ \Delta \mu(i) = \sum_{k=1}^{K} |\mu_k(i) - \mu_k(i-1)| \quad (Eq. 9.1) \]

  • **number of pixels that change labels between iterations** < threshold
MULTISPECTRAL CLASSIFICATION

ISODATA algorithm

- adds automated merging of similar clusters and splitting of heterogenous clusters

K-Means Clustering

- Specify $K$ initial cluster mean estimates ("seeds")
  
  Note: $K$ is not the number of features

- Partition feature space according to nearest-mean rule

- Calculate revised cluster mean estimates from partitioned data

- Repeat 2 and 3 until convergence criterion is met
MULTISPECTRAL CLASSIFICATION

Schematic of ISODATA behavior

Specify K initial cluster mean estimates ("seeds") and partition feature space

Calculate revised cluster mean estimates from partitioned data

after iteration 2

after iteration 3

mean vector migration
MULTISPECTRAL CLASSIFICATION

net mean migration versus iteration number

![Graph showing net mean migration versus iteration number]
MULTISPECTRAL CLASSIFICATION

Image example

- Marana, AZ, TM data, bands 3 and 4

Color IR composite

- dark soil
- light soil
- crop
MULTISPECTRAL CLASSIFICATION

unsupervised cluster maps for various values of $K$

$K = 2$

4

5

6

7
feature space cluster mean locations corresponding to Fig. 9-8

At least 6 clusters needed to “capture” vegetation class because it is a numerically small population
MULTISPECTRAL CLASSIFICATION

residual magnitude image error maps corresponding to Fig. 9-8

Error is greatly reduced for vegetation with 6 or more clusters
MULTISPECTRAL CLASSIFICATION

total residual magnitude error in cluster map

Error drops by 25% in going from 5 to 6 clusters
MULTISPECTRAL CLASSIFICATION

Hybrid Supervised/Unsupervised Training

• Combines the benefits of both types of training

Hybrid Training

• Cluster the image into a large number of cluster classes (say 50 or more)
  Insures that classes are separable

• Using ancillary data, supervise the labeling of the clusters
  Many will be merged into single, labeled classes
  Result is a labeled map of fewer classes

• Accept labeled map as final, or do supervised classification, using labeled training samples from cluster map
MULTISPECTRAL CLASSIFICATION

NONPARAMETRIC CLASSIFICATION
MULTISPECTRAL CLASSIFICATION

Level-Slice Classifier

- No assumed statistical distribution for classes, i.e. nonparametric
- Robust to variation in class signatures
- Performance cannot be predicted or supported by statistical theory
- Simple and efficient
MULTISPECTRAL CLASSIFICATION

- Includes outlier “unclassified” label

**Level-Slice Classifier**
- Define $L$ “boxes” in K-D feature space, one for each class at class mean ± $k$ times class standard deviation
- Some software allows interactive specification of box boundaries in feature space
- If pixel vector lies within one box, assign that label to pixel
- If pixel vector lies in two or more overlapping boxes, use “tie-breaking” scheme
  - nearest-mean rule

![Feature-space decision boundaries](image)
MULTISPECTRAL CLASSIFICATION

Artificial Neural Network (ANN) Classifier

- **Classifier is a connected structure of simple decision elements (processing nodes)**

- **Classification capability contained in connecting weights between nodes**

- **Train with an iterative, optimization algorithm**

  For given training sample input vector, calculate output classification vector
  
  - supervised training

  Attempt to minimize mean-squared error in output classification vector (desire 1 at class node of interest, 0 for all other class nodes) over all training data

- **Three-layer network can form second-order (nonlinear) decision boundaries**
MULTISPECTRAL CLASSIFICATION

Basic 3-layer ANN

- Input nodes ($i$)
- Hidden layer nodes ($j$)
- Output nodes ($k$)

- Input pattern $p_i$
- Output pattern $o_k$

- Weights $w_{ji}$
- Weights $w_{kj}$
MULTISPECTRAL CLASSIFICATION

• Processing Nodes

  Sums all the weighted inputs and passes the sum through a soft decision function (threshold)

Components of a processing node

(typically, the decision function is a sigmoid function)

\[ f(S) = \frac{1}{1 + e^{-S}} \]  
(Eq. 9-4)
**MULTISPECTRAL CLASSIFICATION**

**Sigmoid function**

The threshold is usually “soft,” as above, but a “hard” threshold performs similarly.

At each **hidden layer node** \( j \), the input signals are weighted and summed, and then passed through the decision function

\[
S_j = \sum_i w_{ji} p_i
\]

(Eq. 9-2)

\[
h_j = f(S_j) = f\left(\sum_i w_{ji} p_i\right)
\]

At each **output layer node** \( k \), the signals from the hidden layer nodes are weighted and summed, and then passed through the decision function

\[
S_k = \sum_j w_{kj} h_j
\]

(Eq. 9-3)

\[
o_k = f(S_k) = f\left(\sum_j w_{kj} h_j\right)
\]

\[
= f\left(\sum_j w_{kj}\left[\sum_i w_{ji} p_i\right]\right)
\]
MULTISPECTRAL CLASSIFICATION

- **Back-Propagation Algorithm**

  *Training algorithm to find “optimal” weights for the given training data*

  *Member of iterative, gradient-descent family of algorithms*

  **Steps**

  1. select training pixels for each class and specify desired output vector $d_k$ for class $k$ (typically $d_{m=k} = 0.9$ and $d_{m\neq k} = 0.1$)
  2. initialize weights as random numbers between 0 and 1 (typically small $\approx 0$)
  3. specify one training cycle:
     - after each training pixel (*sequential*)
     - after all training pixels in each class or
     - after all training pixels in all classes (*batch*)
  4. propagate training data forward through net, one pixel at a time
  5. after each training cycle, calculate the output $o_k$ and the error relative to the desired output $d_k$ (the 1/2 factor is a mathematical convenience)
MULTISPECTRAL CLASSIFICATION

$$\frac{\|e\|^2}{2} = \frac{1}{2} \sum_{p=1}^{P} \sum_{k} (d_k - o_k)^2 \quad (Eq. 9-5)$$

- 6. adjust weight \(w_{kj}\) by,

$$\Delta w_{kj} = LR \sum_{p=1}^{P} (d_k - o_k) \left[ \frac{df(S)}{dS} \right]_{S_{kj}} h_j \quad (Eq. 9-6)$$

where \(LR\) is the Learning Rate (controls speed of convergence)

- 7. adjust \(w_{ji}\) by

$$\Delta w_{ji} = LR \sum_{p=1}^{P} \left\{ \frac{df(S)}{dS} \left| S_{ji} \right\{ \left[ (d_k - o_k) \frac{df(S)}{dS} \right]_{S_{kj}} w_{kj} \right\} \right\} \quad (Eq. 9-7)$$

- 8. do another training cycle (steps 4 through 7) until \(\epsilon < \) threshold for all patterns (classes)
MULTISPECTRAL CLASSIFICATION

Nonparametric Classification Examples

- Level-slice

Example level-slice classification

thematic map  DN feature space

DN4  DN3

crop
light soil
dark soil
MULTISPECTRAL CLASSIFICATION

- Artificial neural network

Convergence of the back-propagation algorithm
MULTISPECTRAL CLASSIFICATION

Hard classification maps at three stages of iteration
MULTISPECTRAL CLASSIFICATION

Final ANN hard classification map

thematic map

DN feature space

crop

light soil

dark soil
MULTISPECTRAL CLASSIFICATION

Soft classification maps at four iterations

All classes “build” over iterations

Small crop class takes longer to “build” than other classes
MULTISPECTRAL CLASSIFICATION

PARAMETRIC CLASSIFICATION
MULTISPECTRAL CLASSIFICATION

Estimation of Model Parameters

- Classifier depends on parameters of assumed class statistical distribution

- The class conditional probabilities, \( p(f|i) \), for a feature vector \( f \), are estimated from the training data for each class \( i \)

  \( p(f|i) \) has unit area (unit volume in K-D)

- The a priori probabilities, \( p(j) \), are estimated by the total percent coverage of each class

  Usually set equal for all classes, i.e. \( \frac{1}{\text{# classes}} \) (unbiased)

- To perform a classification of non-training pixels, we need the a posteriori probabilities, \( p(i|f) \)
MULTISPECTRAL CLASSIFICATION

• **Bayes Rule**

\[ p(i|f) = \frac{p(f|i)p(i)}{p(f)} \quad (\text{Eq. 9-11}) \]

\[ p(f) = \sum_{i} p(f|i)p(i) \quad (\text{Eq. 9-12}) \]

- is the total probability of feature vector \( f \), and is independent of class.

• **Bayes decision rule**

If \( p(i|f) > p(j|f) \), for all \( j \neq i \), assign pixel to class \( i \)

Intuitive behavior: assign pixel to class with largest *a posteriori* probability
MULTISPECTRAL CLASSIFICATION

Discriminant Functions

- Restate Bayes decision rule

\[ D_i(f) > D_j(f), \text{ for all } j \neq i, \text{ assign pixel to class } i \]

where the discriminant function for class \( i \), \( D_i(f) \), is given by,

\[ D_i(f) = p(i|f)p(f) = p(f|i)p(i). \quad (\text{Eq. } 9-13) \]

- Any monotonic function of \( p(i|f)p(f) \), also works as a discriminant function

For example

\[ D_i(f) = Ap(i|f)p(f) + B = Ap(f|i)p(i) + B \quad (\text{Eq. } 9-14) \]

or

\[ D_i(f) = \ln[p(i|f)p(f)] = \ln[p(f|i)p(i)] \quad (\text{Eq. } 9-15) \]

The log transformation in Eq. 9-15 is advantageous for normal distributions
MULTISPECTRAL CLASSIFICATION

The Normal Distribution Model

- **Common statistical assumption**

  *Note, Bayes decision rule does not require normal distributions*

  Little research or data on normality of training pixels in remote sensing data

- **Assuming normal (Gaussian) distribution in K-D**

\[
N(DN; \mu, C) = \frac{1}{|C|^{1/2}(2\pi)^{K/2}} e^{-(DN - \mu) C^{-1} (DN - \mu)/2} \quad \text{(Eq. 4-18)}
\]

and Eq. 9-15,

\[
D_i(f) = \ln[p(i)] - \frac{1}{2} \left\{ K \ln[2\pi] + \ln|C_i| + (f - \mu_i)^T C_i^{-1} (f - \mu_i) \right\} \quad \text{(Eq. 9-16)}
\]

Only the last term needs to be recalculated at each pixel
MULTISPECTRAL CLASSIFICATION

**Decision boundaries (Gaussian PDF)**

- located where the probabilities of classes are equal
- quadratics in 2-D and hyperquadrics in K-D
- The decision rule is equivalent to setting, for two classes \(a\) and \(b\),

\[
\ln[p(a|f)p(f)] = \ln[p(b|f)p(f)] \quad \text{or} \quad p(a|f) = p(b|f) \quad \text{(Eq. 9-18)}
\]
MULTISPECTRAL CLASSIFICATION

**Maximum-likelihood classifier**

- **Minimizes total misclassification error**,

\[ \varepsilon_{\text{total}} = \sum_{i \neq j} \varepsilon(i|j) \]

where \( \varepsilon(i|j) \) is the error that a pixel belonging to class \( j \) is mislabeled as class \( i \)

**Quadratic decision boundaries in 2-D feature space**
MULTISPECTRAL CLASSIFICATION

**Maximum-Likelihood Classifier**

- Estimate mean vector $m_i$ and covariance matrix $C_i$ for each class from the DN mean and covariance of training data.

- Calculate discriminant functions at each image pixel (Eq. 9-16).

- For each unlabeled pixel, assign the label of the class with the largest discriminant function (i.e. the largest a posteriori probability).
The Nearest-Mean Classifier

- Special case of the maximum-likelihood classifier
- Also called minimum-Euclidean distance classifier
- Assumes

Class covariance matrices are equal and diagonal

\[
c_0 = \begin{bmatrix} c_0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & c_0 \end{bmatrix} \quad (Eq. 9-23)
\]

A priori probabilities are equal
MULTISPECTRAL CLASSIFICATION

- **Discriminant function then becomes**

\[
D_i(f) = A - \frac{(f - \mu_i)^T(f - \mu_i)}{2c_0}
\]  
(Eq. 9-24)

which is (almost) the Euclidean distance of a pixel vector \( f \) from a class mean vector \( \mu_i \).

This discriminant function will be a maximum when the distance is a minimum (because of negative sign on the second term).

So, maximizing the discriminant function minimizes the Euclidean distance.
MULTISPECTRAL CLASSIFICATION

Parametric Classification Examples

- Nearest-mean

Nearest-mean classification in image and feature space
MULTISPECTRAL CLASSIFICATION

- Maximum-likelihood

Maximum-likelihood classification in image and feature space, without and with probability threshold
MULTISPECTRAL CLASSIFICATION

Modeling of class distributions with Gaussians

- Two-class example (Lake Anna, Virginia, MSS image)

Image with training sites
MULTISPECTRAL CLASSIFICATION

Gaussian models compared to class distributions

![Diagrams showing Gaussian models compared to class distributions.](image-url)
MULTISPECTRAL CLASSIFICATION

Classifications with one training site/class

nearest-mean

maximum-likelihood
MULTISPECTRAL CLASSIFICATION

Effect of a priori probabilities on model fit

![Diagram showing the effect of a priori probabilities on model fit.](image-url)
MULTISPECTRAL CLASSIFICATION

Model fit with 2 training sites for water
MULTISPECTRAL CLASSIFICATION

Final maximum-likelihood classification
MULTISPECTRAL CLASSIFICATION

SPATIAL-SPECTRAL SEGMENTATION
MULTISPECTRAL CLASSIFICATION

Spatial-Spectral Model

- neighboring pixels are likely to belong to same spectral class
- Goal is to group “connected” pixels into spatial segments
- May be thought of as “spatial clustering,” using a spectral similarity criterion

Region Growing

- Types of “connectedness”
  
  4-connected
  
  - only vertical and horizontal neighbors
  
  8-connected
  
  - all 8 neighbors, including diagonal neighbors
MULTISPECTRAL CLASSIFICATION

Connectedness definitions

4-connected

8-connected

3 x 3 pixel neighborhood

possible connected region
MULTISPECTRAL CLASSIFICATION

4-connected region growing algorithm

- DN difference threshold $t$ between a pixel and its neighbors
- Proceed top-to-bottom, left-to-right across image with 3-pixel window

Pixel and label windows

- Apply series of rules to test for connectivity at each pixel
### MULTISPECTRAL CLASSIFICATION

#### Rules for segmentation

<table>
<thead>
<tr>
<th>case</th>
<th>if</th>
<th>then</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. start new region</td>
<td>$</td>
<td>DN - DN_u</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>DN - DN_l</td>
</tr>
<tr>
<td>2. merge with upper</td>
<td>$</td>
<td>DN - DN_u</td>
</tr>
<tr>
<td>region</td>
<td>$</td>
<td>DN - DN_l</td>
</tr>
<tr>
<td>3. merge with left</td>
<td>$</td>
<td>DN - DN_u</td>
</tr>
<tr>
<td>region</td>
<td>$</td>
<td>DN - DN_l</td>
</tr>
<tr>
<td>4. merge with upper and</td>
<td>$</td>
<td>DN - DN_u</td>
</tr>
<tr>
<td>left region</td>
<td>$</td>
<td>DN - DN_l</td>
</tr>
<tr>
<td></td>
<td>$L_l = L_u$</td>
<td></td>
</tr>
<tr>
<td>5. relabel and merge</td>
<td>$</td>
<td>DN - DN_u</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>DN - DN_l</td>
</tr>
<tr>
<td></td>
<td>$L_l \neq L_u$</td>
<td></td>
</tr>
</tbody>
</table>
MULTISPECTRAL CLASSIFICATION

• **Steps**

1. **Label first row of pixels**
   - assign \( L = 0 \) at first pixel (upper left corner of image)
   - use modified Cases 1 and 3 for remaining pixels in first row

<table>
<thead>
<tr>
<th>modified case</th>
<th>if</th>
<th>then</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. start new region</td>
<td>(</td>
<td>DN - DN_l</td>
</tr>
<tr>
<td>3. merge with left region</td>
<td>(</td>
<td>DN - DN_l</td>
</tr>
</tbody>
</table>

2. **Label first pixel, second row**
   - use modified Cases 1 and 3

<table>
<thead>
<tr>
<th>modified case</th>
<th>if</th>
<th>then</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. start new region</td>
<td>(</td>
<td>DN - DN_u</td>
</tr>
<tr>
<td>3. merge with upper region</td>
<td>(</td>
<td>DN - DN_u</td>
</tr>
</tbody>
</table>

3. **Label remaining pixels in second row using all Cases 1 - 5 in Table 9-6.**
MULTISPECTRAL CLASSIFICATION

4. Repeat steps 2 and 3 for remaining rows until finished

• Optional second pass can be done to merge similar regions
  decide on merging two connected regions using threshold on the DN variance of the resulting combined region

• Final pass required to resequence labels
MULTISPECTRAL CLASSIFICATION

Signature maps for two thresholds

**TM3**

**TM4**

Larger $t$ results in more aggregation and, therefore, fewer, but larger, regions

$t = 2$

$t = 5$
MULTISPECTRAL CLASSIFICATION

- Convergence properties

Speed of convergence (as a function of threshold $t$) depends on image structure

Error maps between signature maps and image

Number of regions and DN error versus threshold

Dr. Robert A. Schowengerdt, schowengerdt@ece.arizona.edu, 520-621-2706 (voice), -8076 (fax)
MULTISPECTRAL CLASSIFICATION

• **Effect on multispectral feature space**

  *Spatial segmentation causes “coalescing” of DN vectors into “clusters”*

  • Connected object regions in an image tend to consist of several or more pixels with similar spectral vectors

  *Can be a useful precursor to spectral clustering*

  • Disconnected regions with similar spectral vectors will have different labels from segmentation; such regions will be given the same label by clustering

  • Adds spatial connectedness information to spectral similarity information
MULTISPECTRAL CLASSIFICATION

Example using TM (Marana, AZ) scattergram

original image  segmented image (t = 5)
MULTISPECTRAL CLASSIFICATION

SUBPIXEL CLASSIFICATION
MULTISPECTRAL CLASSIFICATION

Spatial-Spectral Mixing

- Spatial integration within the sensor $GIFOV$

\[GIFOV = 1 \times 1 \quad 3 \times 3 \quad 7 \times 7 \quad 13 \times 13\]
MULTISPECTRAL CLASSIFICATION

• Weighted spatial integration within net sensor spatial response

scene  detector GIFOV  effective net GIFOV
MULTISPECTRAL CLASSIFICATION

- Simple geometric model for spatial-spectral mixing

One GIFOV:

Class a: 65% area, spectrum $E_a$

Class b: 20% area, spectrum $E_b$

Class c: 15% area, spectrum $E_c$

Total spectrum at pixel: $DN = 0.65E_a + 0.20E_b + 0.15E_c$
**MULTISPECTRAL CLASSIFICATION**

- **Linear mixing model** based on proportions of the spectral signatures within each pixel

Assumes the spectrum at each pixel can be represented by a linear combination of pure “end-member” spectra

- In vector-matrix form, for K bands and N end-members:

\[
DN_{ij} = Ef_{ij}
\]

where

- \(DN_{ij}\) is the \(K \times 1\) hyperspectral vector at pixel \(ij\)
- \(E\) is the \(K \times N\) end member spectrum matrix. Each column is the spectrum for one end-member
- \(f_{ij}\) is the \(N \times 1\) fraction vector of each end-member for pixel \(ij\)

- To find \(f\) at each pixel, we must invert the above equation.
MULTISPECTRAL CLASSIFICATION

- Since $K \gg N$, inversion is a least-squares problem to find the estimated fraction vector $\hat{f}_{ij}$ that minimizes the total error:

$$\min [\varepsilon^T_{ij} \varepsilon_{ij}] = (DN_{ij} - \hat{E}f_{ij})^T (DN_{ij} - \hat{E}f_{ij}) \quad (Eq. 9-34)$$

- Use pseudoinverse solution, at each pixel:

$$\hat{f}_{ij} = (E^T E)^{-1} E^T \cdot DN_{ij} \quad (Eq. 9-33)$$

- Additional contraints:

$$f_n \geq 0, \quad \sum_{n=1}^{N} f_n = 1 \quad (Eq. 9-26 \text{ and } 9-27)$$
MULTISPECTRAL CLASSIFICATION

- Convex hull

*Defined by outer boundary connecting end-members*

*Three possible endmember sets in 2-D*
MULTISPECTRAL CLASSIFICATION

• Ways to find the end-members

  Laboratory or field reflectance spectra. **Image data must be calibrated to reflectance.**

  Image pixels modeled as mixtures of library reflectance spectra. **Image data must be calibrated to reflectance.**

  Automated techniques based on image transforms, e.g. PCT

  K-dimensional interactive visualization tools
MULTISPECTRAL CLASSIFICATION

Unmixing example

- Two bands of TM image with three classes
- 2-D multispectral vector at each pixel,

\[ DN_{ij} = Ef_{ij} \quad (Eq. 9-28) \]

or,

\[
\begin{bmatrix}
DN_3 \\
DN_4
\end{bmatrix} =
\begin{bmatrix}
E_{dksoil3} & E_{ltsoil3} & E_{crop3} \\
E_{dksoil4} & E_{ltsoil4} & E_{crop4}
\end{bmatrix}
\begin{bmatrix}
f_{dksoil} \\
f_{ltsoil} \\
f_{crop}
\end{bmatrix} \quad (Eq. 9-29)
\]

which is underdetermined.
MULTISPECTRAL CLASSIFICATION

- **Incorporate full coverage constraint at each pixel:**

  \[ 1 = f_{\text{darksoil}} + f_{\text{lightsoil}} + f_{\text{crop}} \]  
  \(\text{(Eq. 9-30)}\)

- **Augmented mixing equation**

  \[
  \begin{bmatrix}
  DN_3 \\
  DN_4 \\
  1
  \end{bmatrix}
  =
  \begin{bmatrix}
  E_{\text{dksoil}} & E_{\text{ltsoil}} & E_{\text{crop}} \\
  E_{\text{dksoil}} & E_{\text{ltsoil}} & E_{\text{crop}} \\
  1 & 1 & 1
  \end{bmatrix}
  \begin{bmatrix}
  f_{\text{dksoil}} \\
  f_{\text{ltsoil}} \\
  f_{\text{crop}}
  \end{bmatrix}
  \]  
  \(\text{(Eq. 9-31)}\)

- **Solve exactly for class fractions**

  \[
  \begin{bmatrix}
  f_{\text{dksoil}} \\
  f_{\text{ltsoil}} \\
  f_{\text{crop}}
  \end{bmatrix}
  =
  \begin{bmatrix}
  E_{\text{dksoil}} & E_{\text{ltsoil}} & E_{\text{crop}} \\
  E_{\text{dksoil}} & E_{\text{ltsoil}} & E_{\text{crop}} \\
  1 & 1 & 1
  \end{bmatrix}^{-1}
  \begin{bmatrix}
  DN_3 \\
  DN_4 \\
  1
  \end{bmatrix}
  \]  
  \(9-32\)
MULTISPECTRAL CLASSIFICATION

Class fraction maps produced for two different sets of end-members

- **data-defined endmembers**
- **“virtual” endmembers**

**dark soil**

**light soil**

**crop**
MULTISPECTRAL CLASSIFICATION

End-member values for example

<table>
<thead>
<tr>
<th>endmember type</th>
<th>band</th>
<th>dark soil</th>
<th>light soil</th>
<th>crop</th>
</tr>
</thead>
<tbody>
<tr>
<td>data-defined</td>
<td>3</td>
<td>18</td>
<td>84</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>14</td>
<td>72</td>
<td>84</td>
</tr>
<tr>
<td>virtual</td>
<td>3</td>
<td>15</td>
<td>93</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6</td>
<td>77</td>
<td>90</td>
</tr>
</tbody>
</table>
MULTISPECTRAL CLASSIFICATION

Alternate interpretations of mixing analysis

- **Class fractions represent spectral mixing**
  
  “intimate” mixing caused by lack of one-to-one correspondance between class labels and spectral signatures

- **Class fractions represent spatial mixing within GIFOV and spatial response function**
  
  Independent of spectral similarity of classes

- **Other mixing indicators**
  
  output node values in a ANN classification
  
  *a posteriori* probabilities in a maximum-likelihood classification
MULTISPECTRAL CLASSIFICATION

Classification Comparisons

- Soft image maps

- TM CIR (bands 4, 3, 2)
- ANN output
- mixing fractions
MULTISPECTRAL CLASSIFICATION

- **Hard image and feature maps**

![Example images and graphs related to multispectral classification.](image-url)