Classification of remotely sensed images

N. Dudeni, P. Debba

CSIR, Logistics and Quantitative Methods, CSIR Built Environment

Outline

OVERVIEW OF REMOTE SENSING

- Useful tool in observing the earth
- Widely applicable in natural resource management, earth systems science, etc.
- Sensors record electromagnetic radiation (EMR) from earth's surface
- Each part of EM spectrum results in a 2-dimensional image

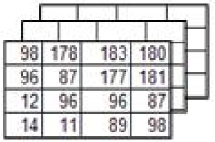


Figure: example of a remotely-sensed image

N. Dudeni, P. Debba

OVERVIEW OF REMOTE SENSING

- Useful tool in observing the earth
- Widely applicable in natural resource management, earth systems science, etc.
- Sensors record electromagnetic radiation (EMR) from earth's surface
- Each part of EM spectrum results in a 2-dimensional image

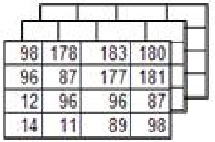


Figure: example of a remotely-sensed image

OVERVIEW OF REMOTE SENSING

- Useful tool in observing the earth
- Widely applicable in natural resource management, earth systems science, etc.
- Sensors record electromagnetic radiation (EMR) from earth's surface
- Each part of EM spectrum results in a 2-dimensional image

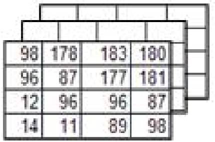


Figure: example of a remotely-sensed image

OVERVIEW OF REMOTE SENSING

- Useful tool in observing the earth
- Widely applicable in natural resource management, earth systems science, etc.
- Sensors record electromagnetic radiation (EMR) from earth's surface
- Each part of EM spectrum results in a 2-dimensional image

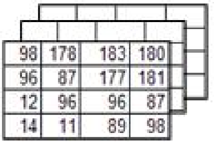


Figure: example of a remotely-sensed image

OVERVIEW OF REMOTE SENSING

- Useful tool in observing the earth
- Widely applicable in natural resource management, earth systems science, etc.
- Sensors record electromagnetic radiation (EMR) from earth's surface
- Each part of EM spectrum results in a 2-dimensional image

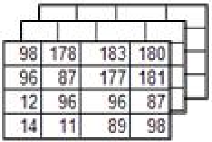


Figure: example of a remotely-sensed image

OVERVIEW OF REMOTE SENSING (cont. . .)

- Statistical techniques are required for processing and classification of images
- Methods of multivariate analysis (discriminant, cluster, and principal component analyses)

Classification of remotely sensed images

N. Dudeni, P. Debba

Introduction to Remote Sensing

Introduction to Image Classification

Objective of the study

Classification algorithms by group

Unsupervised algorithms

Supervised classification algorithms

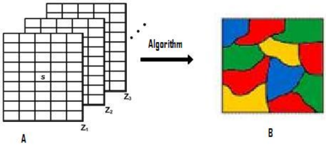
Spatial Classification: ICM algorithm

- Statistical techniques are required for processing and classification of images
- Methods of multivariate analysis (discriminant, cluster, and principal component analyses)

OVERVIEW OF REMOTE SENSING (cont. . .)

- Statistical techniques are required for processing and classification of images
- Methods of multivariate analysis (discriminant, cluster, and principal component analyses)

OVERVIEW OF IMAGE CLASSIFICATION

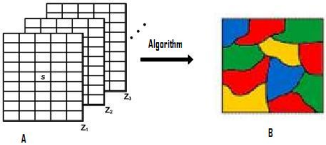


Suppose that:

- Data are located on a lattice D with locations $s = (u, v)'$
- $Z(s) = (Z_1(s), Z_2(s), \ldots, Z_p(s))$ is the data vector for a pixel located at s
- $\theta(s)$ is an unknown ground class to which pixel s belongs

Objective is to classify the pixel at location s to the one of the k clusters

OVERVIEW OF IMAGE CLASSIFICATION

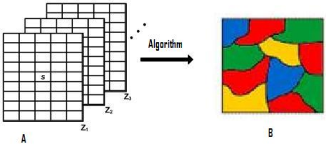


Suppose that:

- Data are located on a lattice D with locations $s = (u, v)'$
- $Z(s) = (Z_1(s), Z_2(s), \ldots, Z_p(s))$ is the data vector for a pixel located at s
- $\theta(s)$ is an unknown ground class to which pixel s belongs

Objective is to classify the pixel at location s to the one of the k clusters

OVERVIEW OF IMAGE CLASSIFICATION

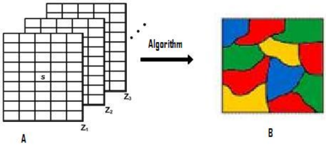


Suppose that:

- Data are located on a lattice D with locations $s = (u, v)'$
- $Z(s) = (Z_1(s), Z_2(s), \ldots, Z_p(s))$ is the data vector for a pixel located at s
- $\theta(s)$ is an unknown ground class to which pixel s belongs

Objective is to classify the pixel at location s to the one of the k clusters

OVERVIEW OF IMAGE CLASSIFICATION



Suppose that:

- Data are located on a lattice D with locations $s = (u, v)'$
- $Z(s) = (Z_1(s), Z_2(s), \ldots, Z_p(s))$ is the data vector for a pixel located at s
- $\theta(s)$ is an unknown ground class to which pixel s belongs

Objective is to classify the pixel at location s to the one of the k clusters

OVERVIEW OF IMAGE CLASSIFICATION (cont. . .)

CLASSES OF IMAGE CLASSIFICATION

- Classification algorithms are broadly divided into supervised and unsupervised
- Spatial classification algorithms
  - Spatial classification techniques seek to identify and group pixels that are geographically 'close' and are 'similar' w.r.t spectral values

OVERVIEW OF IMAGE CLASSIFICATION (cont. . .)

CLASSES OF IMAGE CLASSIFICATION

- Classification algorithms are broadly divided into supervised and unsupervised
- Spatial classification algorithms
    - Spatial classification techniques seek to identify and group pixels that are geographically 'close' and are 'similar' w.r.t spectral values

OVERVIEW OF IMAGE CLASSIFICATION (cont. . .)

CLASSES OF IMAGE CLASSIFICATION

- Classification algorithms are broadly divided into supervised and unsupervised
- Spatial classification algorithms
  - Spatial classification techniques seek to identify and group pixels that are geographically 'close' and are 'similar' w.r.t spectral values

OVERVIEW OF IMAGE CLASSIFICATION (cont. . .)

CLASSES OF IMAGE CLASSIFICATION

- Classification algorithms are broadly divided into supervised and unsupervised
- Spatial classification algorithms
  - Spatial classification techniques seek to identify and group pixels that are geographically 'close' and are 'similar' w.r.t spectral values

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Unsupervised vs Supervised classification

Unsupervised

- Classifies pixels purely based on spectral properties
- No prior knowledge of the actual groupings or structures in the image
- k groups of classes to be identified in the image are to specified
- To determine the identity of natural groupings, reference data are used

Supervised

- Classification is supervised based on knowledge of actual surface types
- k training classes are identified prior to classification
- These classes are used to train the algorithm to recognize similar pixels for each class
- Each image pixel is then compared to each training class

Background and Objective of the study

- Examine several classification algorithms (unsupervised, supervised, and SPATIAL)
- Demonstrate the application of these algorithms to remotely sensed images
- Identify which is the “best” method for image classification
- long-term goals include developing new classification and segmentation algorithm(s)

Data or image used for classification

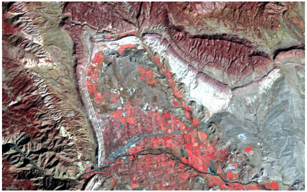


Figure: Image used in the classification

- Bright red areas representing high infra-red corresponding to healthy vegetation
- Slight red areas represent coniferous trees
- Geological and urban classes are available on the image

Data or image used for classification (cont. . .

- (Software data) Landsat data from Canon City, Colorado in the USA
- Ground truth data with **three** training samples
- The training sample had **811** ground truth data points
- Data had **6 bands**,which were all used for classification
- The **500** points used for training the supervised algorithms
- The other **311** points used for validation of classification algorithm

Data or image used for classification (cont. . .

- (Software data) Landsat data from Canon City, Colorado in the USA
- Ground truth data with **three** training samples
- The training sample had **811** ground truth data points
- Data had **6 bands**,which were all used for classification
- The **500** points used for training the supervised algorithms
- The other **311** points used for validation of classification algorithm

Data or image used for classification (cont. . .

- (Software data) Landsat data from Canon City, Colorado in the USA
- Ground truth data with **three** training samples
- The training sample had **811** ground truth data points
- Data had **6 bands**,which were all used for classification
- The **500** points used for training the supervised algorithms
- The other **311** points used for validation of classification algorithm

Data or image used for classification (cont. . .

- (Software data) Landsat data from Canon City, Colorado in the USA
- Ground truth data with **three** training samples
- The training sample had **811** ground truth data points
- Data had **6 bands**,which were all used for classification
- The **500** points used for training the supervised algorithms
- The other **311** points used for validation of classification algorithm

Data or image used for classification (cont. . .

- (Software data) Landsat data from Canon City, Colorado in the USA
- Ground truth data with **three** training samples
- The training sample had **811** ground truth data points
- Data had **6 bands**,which were all used for classification
- The **500** points used for training the supervised algorithms
- The other **311** points used for validation of classification algorithm

Data or image used for classification (cont. . .

- (Software data) Landsat data from Canon City, Colorado in the USA
- Ground truth data with **three** training samples
- The training sample had **811** ground truth data points
- Data had **6 bands**,which were all used for classification
- The **500** points used for training the supervised algorithms
- The other **311** points used for validation of classification algorithm

Data or image used for classification (cont. . .

- (Software data) Landsat data from Canon City, Colorado in the USA
- Ground truth data with **three** training samples
- The training sample had **811** ground truth data points
- Data had **6 bands**,which were all used for classification
- The **500** points used for training the supervised algorithms
- The other **311** points used for validation of classification algorithm

Classification algorithms by group(cont. . .)

Unsupervised classification algorithms

- K-means

Supervised classification algorithms

- Maximum likelihood

Spatial classification algorithms

- Iterated conditional modes (ICM) algorithm

Classification algorithms by group(cont. . .)

Unsupervised classification algorithms

- K-means

Supervised classification algorithms

- Maximum likelihood

Spatial classification algorithms

- Iterated conditional modes (ICM) algorithm

Classification algorithms by group(cont. . .)

Unsupervised classification algorithms

- K-means

Supervised classification algorithms

- Maximum likelihood

Spatial classification algorithms

- Iterated conditional modes (ICM) algorithm

Classification algorithms by group(cont. . .)

Unsupervised classification algorithms

- K-means

Supervised classification algorithms

- Maximum likelihood

Spatial classification algorithms

- Iterated conditional modes (ICM) algorithm

Unsupervised Classification: The K-means algorithm

- Iterative procedure that assigns each pixel $Z(s)$ to the class that has the closest centroid
- Closeness is usually with respect to the Euclidean distance between the class centroid and the pixel
- The objective function of the k-means seeks to minimize the $SS_{distances}$ defined by:

$$E = \sum_{i=1}^{k} \sum_{Z(s) \in C_i} |Z(s) - \mu_i|^2 \qquad (1)$$

where μ_i is the mean of the class that pixel $Z(s)$ is assigned to

Unsupervised Classification: The K-means algorithm

- Iterative procedure that assigns each pixel $Z(s)$ to the class that has the closest centroid
- Closeness is usually with respect to the Euclidean distance between the class centroid and the pixel
- The objective function of the k-means seeks to minimize the $SS_{distances}$ defined by:

$$E = \sum_{i=1}^{k} \sum_{Z(s) \in C_i} |Z(s) - \mu_i|^2 \qquad (1)$$

where μ_i is the mean of the class that pixel $Z(s)$ is assigned to

Unsupervised Classification: The K-means algorithm

- Iterative procedure that assigns each pixel $Z(s)$ to the class that has the closest centroid
- Closeness is usually with respect to the Euclidean distance between the class centroid and the pixel
- The objective function of the k-means seeks to minimize the $SS_{distances}$ defined by:

$$E = \sum_{i=1}^{k} \sum_{Z(s)\in C_i} |Z(s) - \mu_i|^2 \qquad (1)$$

where μ_i is the mean of the class that pixel $Z(s)$ is assigned to

Unsupervised Classification: The K-means algorithm

- Iterative procedure that assigns each pixel $Z(s)$ to the class that has the closest centroid
- Closeness is usually with respect to the Euclidean distance between the class centroid and the pixel
- The objective function of the k-means seeks to minimize the $SS_{distances}$ defined by:

$$E = \sum_{i=1}^{k} \sum_{Z(s) \in C_i} |Z(s) - \mu_i|^2 \quad (1)$$

where μ_i is the mean of the class that pixel $Z(s)$ is assigned to

Unsupervised Classification: The K-means algorithm

- Iterative procedure that assigns each pixel $Z(s)$ to the class that has the closest centroid
- Closeness is usually with respect to the Euclidean distance between the class centroid and the pixel
- The objective function of the k-means seeks to minimize the $SS_{distances}$ defined by:

$$E = \sum_{i=1}^{k} \sum_{Z(s) \in C_i} |Z(s) - \mu_i|^2 \qquad (1)$$

where μ_i is the mean of the class that pixel $Z(s)$ is assigned to

Unsupervised Classification: The K-means algorithm

- Iterative procedure that assigns each pixel $Z(s)$ to the class that has the closest centroid
- Closeness is usually with respect to the Euclidean distance between the class centroid and the pixel
- The objective function of the k-means seeks to minimize the $SS_{distances}$ defined by:

$$E = \sum_{i=1}^{k} \sum_{Z(s) \in C_i} |Z(s) - \mu_i|^2 \qquad (1)$$

where μ_i is the mean of the class that pixel $Z(s)$ is assigned to

Unsupervised Classification: The K-means algorithm(cont. . .)

- Minimizing the E ($SS_{distances}$) is equivalent to minimizing the Mean Squared Error (MSE)

$$MSE = \frac{E}{(N-k)z} \tag{2}$$

where N is the number of pixels
k indicates the number of classes
z is the number of spectral bands

Unsupervised Classification: The K-means algorithm(cont. . .)

- Minimizing the E ($SS_{distances}$) is equivalent to minimizing the Mean Squared Error (MSE)

$$MSE = \frac{E}{(N-k)z} \tag{2}$$

where N is the number of pixels
k indicates the number of classes
z is the number of spectral bands

Unsupervised Classification: The K-means algorithm(cont. . .)

- Minimizing the E ($SS_{distances}$) is equivalent to minimizing the Mean Squared Error (MSE)

$$MSE = \frac{E}{(N-k)z} \tag{2}$$

where N is the number of pixels
k indicates the number of classes
z is the number of spectral bands

Unsupervised Classification: The K-means algorithm(cont. . .)

- Minimizing the E ($SS_{distances}$) is equivalent to minimizing the Mean Squared Error (MSE)

$$MSE = \frac{E}{(N-k)z} \tag{2}$$

where N is the number of pixels
k indicates the number of classes
z is the number of spectral bands

Unsupervised Classification: The K-means algorithm(cont. . .)

- Minimizing the E ($SS_{distances}$) is equivalent to minimizing the Mean Squared Error (MSE)

$$MSE = \frac{E}{(N-k)z} \quad (2)$$

where N is the number of pixels
k indicates the number of classes
z is the number of spectral bands

Unsupervised Classification: The K-means algorithm(cont. . .)

- Minimizing the E ($SS_{distances}$) is equivalent to minimizing the Mean Squared Error (MSE)

$$MSE = \frac{E}{(N-k)z} \tag{2}$$

where N is the number of pixels
k indicates the number of classes
z is the number of spectral bands

Unsupervised Classification: The K-means algorithm(cont. . .)

- Minimizing the E ($SS_{distances}$) is equivalent to minimizing the Mean Squared Error (MSE)

$$MSE = \frac{E}{(N-k)z} \tag{2}$$

where N is the number of pixels
k indicates the number of classes
z is the number of spectral bands

Unsupervised Classification: The K-means algorithm (cont. . .)

- K-means works best for images with clusters that are spherical in shape and have some variance
- Sensitive to initial centroids representing initial class means
- Not robust to outliers

Unsupervised Classification: The K-means algorithm (cont. . .)

- K-means works best for images with clusters that are spherical in shape and have some variance
- Sensitive to initial centroids representing initial class means
- Not robust to outliers

Unsupervised Classification: The K-means algorithm (cont. . .)

- K-means works best for images with clusters that are spherical in shape and have some variance
- Sensitive to initial centroids representing initial class means
- Not robust to outliers

Unsupervised Classification: The K-means algorithm (cont. . .)

- K-means works best for images with clusters that are spherical in shape and have some variance
- Sensitive to initial centroids representing initial class means
- Not robust to outliers

Supervised Classification: The Maximum Likelihood algorithm

- Parametric method
- The distribution pattern of each category can be described by the mean vector and the covariance matrix
- The pixel is assigned to the class for which the probability is the highest
- For data with multiple bands, an n-dimensional multivariate normal density function is computed using the following function:

Supervised Classification: The Maximum Likelihood algorithm

- Parametric method
- The distribution pattern of each category can be described by the mean vector and the covariance matrix
- The pixel is assigned to the class for which the probability is the highest
- For data with multiple bands, an n-dimensional multivariate normal density function is computed using the following function:

Supervised Classification: The Maximum Likelihood algorithm

- Parametric method
- The distribution pattern of each category can be described by the mean vector and the covariance matrix
- The pixel is assigned to the class for which the probability is the highest
- For data with multiple bands, an n-dimensional multivariate normal density function is computed using the following function:

Supervised Classification: The Maximum Likelihood algorithm

- Parametric method
- The distribution pattern of each category can be described by the mean vector and the covariance matrix
- The pixel is assigned to the class for which the probability is the highest
- For data with multiple bands, an n-dimensional multivariate normal density function is computed using the following function:

Supervised Classification: The Maximum Likelihood algorithm

- Parametric method
- The distribution pattern of each category can be described by the mean vector and the covariance matrix
- The pixel is assigned to the class for which the probability is the highest
- For data with multiple bands, an n-dimensional multivariate normal density function is computed using the following function:

$$P(Z(s)|\theta) = \frac{1}{2\pi^{\frac{1}{2}}} |V_i|^{\frac{1}{2}} \exp\left[\frac{1}{2}(Z(s) - \mu_i)^T V_i^{-1}(Z(s) - \mu_i)\right] \tag{3}$$

where:
$|V_i|$ is the determinant of a covariance matrix,
V_i^{-1} is the inverse of a covariance matrix, and
$(Z(s) - \mu_i)^T$ is the transpose of the vector of $(Z(s) - \mu_i)$
the mean vector (μ_i) and the covariance matrix V_i for each class are estimated from the training data

Supervised Classification: The Maximum Likelihood algorithm (cont. . .)

$$P(Z(s)|\theta) = \frac{1}{2\pi^{\frac{1}{2}}} |V_i|^{\frac{1}{2}} \exp\left[\frac{1}{2}(Z(s) - \mu_i)^T V_i^{-1}(Z(s) - \mu_i)\right] \tag{3}$$

where:
$|V_i|$ is the determinant of a covariance matrix,
V_i^{-1} is the inverse of a covariance matrix, and
$(Z(s) - \mu_i)^T$ is the transpose of the vector of $(Z(s) - \mu_i)$
the mean vector (μ_i) and the covariance matrix V_i for each class are estimated from the training data

Supervised Classification: The Maximum Likelihood algorithm (cont. . .)

- This classification is based on ordinary discriminant analysis and Baye's classification rule
- Suppose we have an image with pixel location s, and $Z(s) = Z_1(s), \ldots, Z_p(s)$ representing the pixel values at each location
- Given $\theta(s)$, the RSID (Z(s))are distributed according to a conditional density function $f(.|\theta(s))$, which does not depend on s
- Let $\{\pi(k) : k = 1, 2, \ldots, K\}$ be a prior probability *dist*$^{\underline{n}}$ of $\theta(s)$
- Based on *realiz*$^{\underline{n}}$ $z(s)$ of $Z(s)$ Baye's *class*$^{\underline{n}}$ rule declares that $\hat{\theta}(s) = k^*$, where k^* maximizes the posterior distribution

$$p(k|z(s)) = \frac{f(z(s)|\pi(k))}{\sum_{i=1}^{k} f(z(s)|\pi(l))} \tag{4}$$

- The denominator remains constant in k and the numerator gets maximized while the probability of assigning pixels to correct classes gets maximized

Supervised Classification: The Maximum Likelihood algorithm (cont. . .)

$$p(k|z(s)) = \frac{f(z(s)|\pi(k))}{\sum_{i=1}^{k} f(z(s)|\pi(l))} \tag{5}$$

- If the distribution of pixels forming a category from the training data with mean μ_k, is normal and has identical covariance matrices, then the this results in ordinary discriminant analysis
- Maximum likelihood approach assumes that data vectors in neighboring pixels are independent (ignores spatial dependencies)
- Neighboring pixels may not be marginally independent even if they are independent conditional on ground class
- Spatial dependencies can result in biased estimates of Σ, and hence increased classification errors

Supervised Classification: The Maximum Likelihood algorithm (cont. . .)

$$p(k|z(s)) = \frac{f(z(s)|\pi(k))}{\sum_{i=1}^{k} f(z(s)|\pi(l))} \quad (5)$$

- If the distribution of pixels forming a category from the training data with mean μ_k, is normal and has identical covariance matrices, then the this results in ordinary discriminant analysis
- Maximum likelihood approach assumes that data vectors in neighboring pixels are independent (ignores spatial dependencies)
- Neighboring pixels may not be marginally independent even if they are independent conditional on ground class
- Spatial dependencies can result in biased estimates of Σ, and hence increased classification errors

Iterated Conditional Modes (ICM) Algorithm

Assume

- θ is the unknown ground class
- given θ, $Z(s_1), Z(s_2), \ldots, Z(s_n)$ are conditionally independent, and
- $Z(s_i)$ has the same conditional density function given by the following:

$$g(Z(s_i)|\theta(s_i)) \tag{6}$$

depends only on $\theta(s_i)$

Iterated Conditional Modes (ICM) Algorithm (cont. . .)

- Let $\theta_t(s)$ denote the classification at pixel location s for the i^{th} iterate.
- For the initial classification $t = 0$ $\hat{\theta}_0(s_i)$ can be obtained from ordinary discriminant analysis (that is, ML classification) or from k-means classification.
- For each pixel s at time $t = 1, 2, \ldots, k$ assign $\theta_t(s) = k*$, where $k*$ maximizes the following probability function:

$$P(\theta_t(s) = k|\{\hat{\theta}_{t-1}(u) : u \neq s\}; Z)$$
$$= c^{-1} g(Z(s)|\hat{\theta}_t(s) = k) \cdot \pi(\hat{\theta}_t(s) = k|\{\hat{\theta}_{t-1}(u) : u \in N_s\})$$

Iterated Conditional Modes (ICM) Algorithm (cont. . .)

- $g(Z(s)|\hat{\theta}_t(s) = k)$ takes care of spectral information
- c^{-1} is the appropriate normalizing constant
- π is a prior distribution
- N_s is the neighborhood structure (pixel contiguity)
- $\pi(\hat{\theta}_t(s) = k|\{\hat{\theta}_{t-1}(u) : u \in N_s\}$ takes care of spatial info of pixels

$$min_k\{(Z(s) - {\mu_k}^t)^T(Z(s) - {\mu_k}^t) - \beta\nu^t {N_s}^t(k)\} \tag{7}$$

Accuracy of the classification was assessed in terms of

- Confusion matrices
- Kappa coefficients
  - measures the degree to which two judges (classification method vs ground truth) concur in respect with classifying N pixels into k mutually exclusive categories

$$k = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \tag{8}$$

Accuracy of the classification was assessed in terms of

- Confusion matrices
- Kappa coefficients
  - measures the degree to which two judges (classification method vs ground truth) concur in respect with classifying N pixels into k mutually exclusive categories

$$k = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \tag{8}$$

Accuracy of the classification was assessed in terms of

- Confusion matrices
- Kappa coefficients
  - measures the degree to which two judges (classification method vs ground truth) concur in respect with classifying N pixels into k mutually exclusive categories

$$k = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \tag{8}$$

K-means classification results

K-means classification accuracy

- Overall Accuracy $= (288/311)92.6\%$
- Kappa Coefficient $= 0.8854$

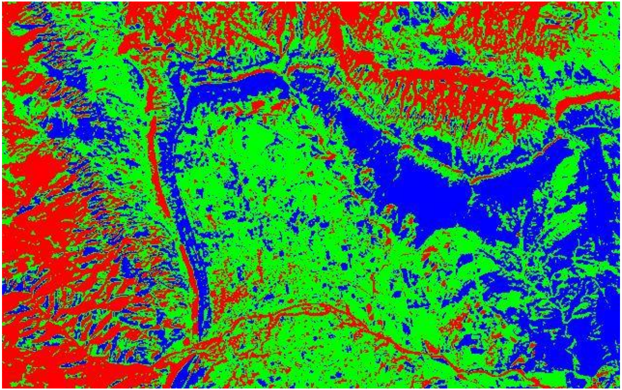


Figure: K-means classified image

Maximum likelihood classification

Maximum likelihood classification accuracy

- Overall Accuracy $= (311/311) \quad 100\%$
- Kappa Coefficient $= 1.0000$

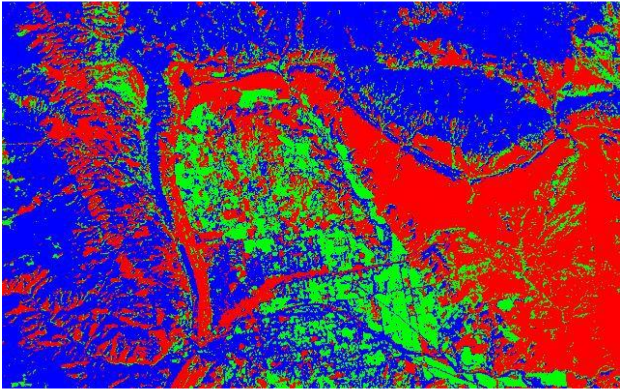


Figure: Maximum likelihood classified image

Iterated conditional modes (ICM) classification

ICM classification accuracy

- Overall Accuracy $= (269/311) 86.5\%$
- Kappa Coefficient $= 0.7958$

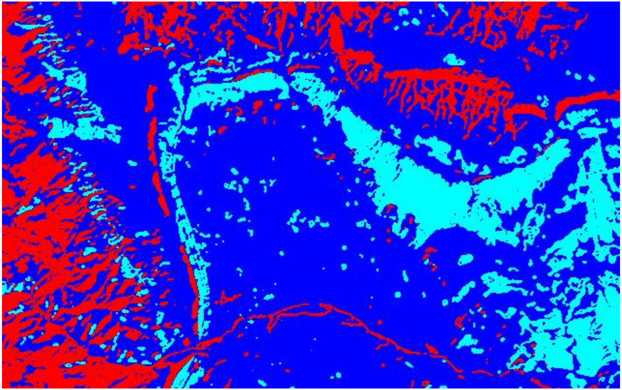


Figure: ICM classified image

Comparison between the classification algorithms

Classification method	Classification accuracy	Kappa Coefficient
Supervised classification		
Maximum Likelihood	(311/311) 100%	1
Unsupervised classification		
K-means	(288/311) 92.6%	0.885
Spatial Classification		
Iterated Conditional Modes	(269/311) 86.5%	0.796

- The accuracy of the maximum likelihood method is high because it is supervised and that the ground truth only contains three regions of interest. There are, however, pixels that are unclassified but do not fall in the validation set and therefore does not decrease the accuracy of the classification
- The K-means and the ICM did not seem to work well as they did perfom well in terms of validation accuracy

The End

THANK YOU