

### **3. Image Classification**

Image classification is an important part of the remote sensing, image analysis and pattern recognition. In some instances, the classification itself may be the object of the analysis. The image classification therefore forms an important tool for examination of the digital images.

The term classifier refers loosely to a computer program that implements a specific procedure for image classification. The analyst must select a classification method that will best accomplish a specific task. At present, it is not possible to state which classifier is best for all situation as the characteristic of each image and the circumstances for each study vary so greatly. Therefore, it is essential that each analyst understand the alternative strategies for image classification so that he or she may be prepared to select the most appropriate classifier for the task in hand.

At present, there is different image classification procedures used for different purposes by various researchers. These techniques are distinguished in two main ways as:

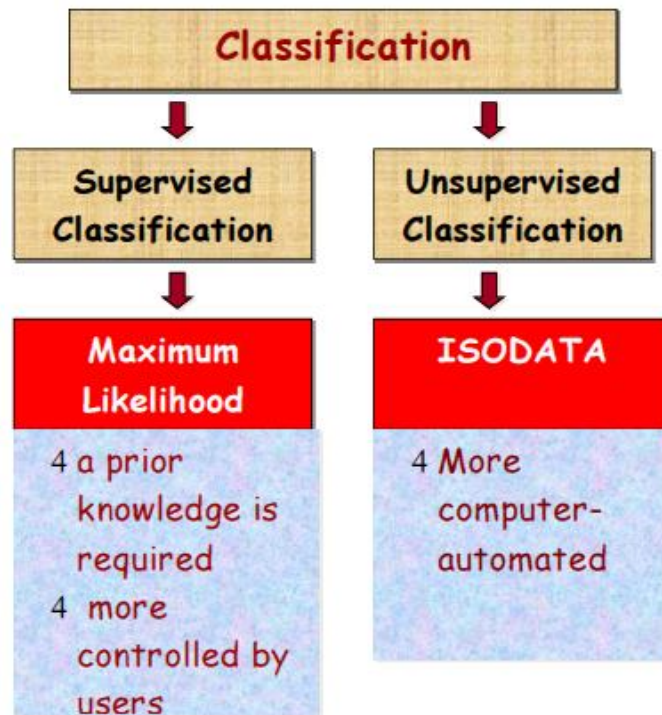
1. Supervised classification

- Analyst identifies representative training sets for each informational class
- Algorithm generates decision boundaries

2. Unsupervised classification

- Algorithm identifies clusters in data
- Analyst labels clusters.

Additionally, supervised classification has different sub classification methods which are named as parallel piped, maximum likelihood, minimum distances and Fisher classifier methods. These methods are named as Hard Classifier.



## Maximum likelihood classification algorithm

- The maximum likelihood decision rule is based on probability.
- It assigns each pixel having pattern measurements or features  $X$  to the class  $i$  whose units are most probable or likely to have given rise to feature vector  $X$ .
- In other words, the probability of a pixel belonging to each of a predefined set of  $m$  classes is calculated, and the pixel is then assigned to the class for which the probability is the highest.
- The maximum likelihood decision rule is one of the most widely used supervised classification algorithms.

The maximum likelihood procedure assumes that the training data statistics for each class in each band are normally distributed (Gaussian).

## **Unsupervised classification (Clustering)**

- ☐ Only some parameters are required to specify from the user to begin this process.
- ☐ Then the computer uses these parameters to uncover statistical patterns that are inherent in the data.
- ☐ Spectral classes do not necessarily correspond to any meaning characteristics of ground objects.
- ☐ After classification, the users must attach the actual meaning to the resulting classes.

## **Iterative Self-Organizing Data Analysis Technique (ISODATA)**

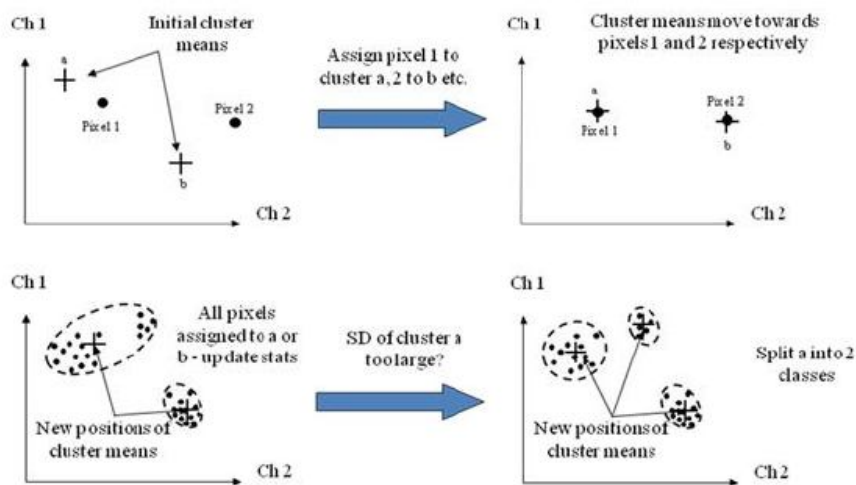
- ☐ ISODATA is self-organizing because it requires relatively little human input.
- ☐ Don't need to know the number of clusters.
- ☐ Algorithm splits and merges clusters.
- ☐ User defines threshold values for parameters.
- ☐ Computer runs algorithm through many iterations until threshold is reached.

### **Unsupervised classification: ISODATA algorithm**

1. Start with (user-defined number) randomly located clusters (mean values).
2. Assign each pixel to nearest cluster (minimum distance).
3. Re-calculate cluster means and standard deviations.
4. If distance between two clusters  $<$  some threshold, merge them.

5. If standard deviation in any one dimension  $>$  some threshold, split into two clusters.
6. Delete clusters with small number of pixels.
7. Re-assign pixels, re-calculate cluster statistics etc. until changes of clusters  $<$  some fixed threshold.

### Original Land sat Image Classified Image



### Drawbacks of ISODATA

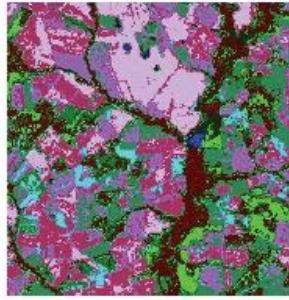
- ☐ May be time consuming if data is much unstructured.
- ☐ Algorithm can spiral out of control leaving only one class.

### Advantages of ISODATA

- ☐ Don't need to know much about the data beforehand.
- ☐ Little user effort required.
- ☐ ISODATA is very effective at identifying spectral clusters in data.



Original Land sat Image



Classified Image

Due to their digital format, the results of digital image classification provide distinct advantages that may make some amount of error tolerable.

These advantages allow the results to be

1. Readily provided in hardcopy map form;
2. Compiled in tabular form to provide area, perimeter, and proximity information (such as edge relationships) for each class; and
- 3 Entered into a geographical information system for subsequent merging and joint analysis with other spatially formatted data.