LA502 Special Studies Remote Sensing

Image Classification _1

Dr. Ragab Khalil
Department of Landscape Architecture
Faculty of Environmental Design
King AbdulAziz University
Room 103

Overview

- Image classification techniques
- Unsupervised Classification
- K-Means Algorithms
- ISODATA

Digital Image interpretation

While visual image interpretation techniques rely on shape, size, pattern, tone, texture, shadows, and association.

Digital image interpretation relies mainly on *color*, i.e. on comparisons of digital numbers found in different bands in different parts of an image.

Objective

The objective of digital image classification procedures is to categorize the pixels in an image into land cover classes.

The output is a *thematic* image with a limited number of feature classes as opposed to a continuous image with varying shades of gray or varying colors representing a continuous range of spectral reflectances.

Spectral signature

The range of digital numbers in different bands for particular features is known as a *spectral pattern* or *spectral signature*.

Pattern in this sense does not have a spatial component. A *spectral pattern* can be composed of adjacent pixels or widely separated pixels.

Classification techniques

Digital image classification techniques can generally be classified into two types:

- Unsupervised classification techniques,
- Supervised classification techniques.

Often, however, these types are used sequentially or iteratively.

Unsupervised Classification

Unsupervised image classification techniques rely on the computer to classify spectrally-similar pixels into classes

Supervised Classification

Supervised classification techniques require the image analyst to define the classification categories and identify a representative samples of pixels to the computer.

The computer then assign all of the remaining pixels to one of the predefined classes on the basis of the similarities between the digital number in the training pixels and the digital numbers in all other pixels.

Comparison

The procedural steps are reversed in these two classification methodologies:

- In *supervised* classification, the analyst defines land cover types and then develops spectral classes that can be used by the computer to identify those pixels that are members of each class.
- In *unsupervised* classification, the computer develops spectral classes and then the analyst associates the spectral classes with land cover types.

Digital Image Processing

Unsupervised Classification

Computer's job

Unsupervised classification differs from supervised classification in that:

- The image analyst does not design the classification scheme nor develop training classes, and
- The computer uses algorithms that aggregate similar pixels into classes based on their similarity with each other and their dissimilarity to the remaining pixels rather than their likely land cover types.

Analyst's job

With unsupervised classification, the land cover types associated with each class are initially unknown and the computer produces no information to aid in their identity.

It is the image analyst's job to associate the classes defined by the computer with the land cover types in the image that these classes represent.

Advantages of Unsupervised classification

Unsupervised classification has two significant advantages over supervised classification:

- The computer can assign pixels to spectrally-distinct classes which an analyst might not recognize as existing, and
- The computer can identify a much larger number of spectrally-distinct classes than an analyst might consider to exist.

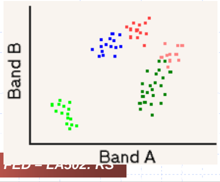
Advantages of Unsupervised classification

Even if an analyst recognizes that distinct subclasses exist, unsupervised classification techniques allow the analyst to avoid developing spectral classes for each unique class and subclass.

Instead, the computer creates a large number of distinct classes and then the analyst can combine them into final classes as deemed appropriate.

Spectral classes

The methods used in both supervised and unsupervised classification require the assignment of individual pixels into a finite number of spectral classes. Each of these spectral classes is presumed to represent a unique land cover type.



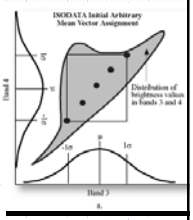
Digital Image Processing

K-Means Algorithms

k-means algorithms

Standard k-means algorithms require the analyst to set the number of clusters to be defined.

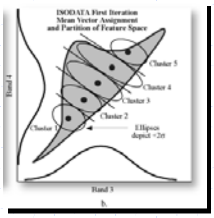
The computer then selects one pixel to initially represent each class. These *seed* pixels are arbitrarily scattered throughout the multidimensional image space defined by the digital numbers in each available band for each pixel. In other words, the seed pixels have large differences in most of their digital numbers.



k-means algorithms

After selecting seed pixels, the computer uses the mean vector of each seed pixel to assign the remaining pixels to clusters around the nearest mean vector.

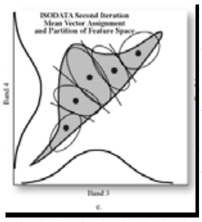
New mean vectors are then calculated using all of the pixels in each of these new clusters. All of the pixels in the image are then reassigned to the nearest of these new mean vectors.



k-means algorithms

The process of calculating new mean vectors and reassigning pixels to the nearest mean vector is repeated until only a limited number of pixels need to be shifted to other classes because there is little movement of the mean vectors.

The threshold percentage of pixels reassigned is another variable controlled by the image analyst, i.e. the inputs to the computer include a percentage of reassigned pixels below which the process stops.

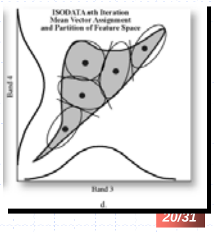


k-means algorithms

A final input provided by the analyst is the number of iterations allowed.

The computer stops calculating new mean vectors and reassigning pixels when:

- Fewer pixels than the input threshold are being reassigned, or
- The maximum number of iterations has been completed.



ISODATA

ERDAS IMAGINE provides a variant of the k-means approach known as *Iterative Self-Organizing Data Analysis* or *ISODATA*.

The ISODATA algorithm is similar to other k-means algorithms, but a significant difference is that the ISODATA algorithm allows the computer to determine the final number of clusters while this value is set by the image analyst in other k-means applications.

ISODATA

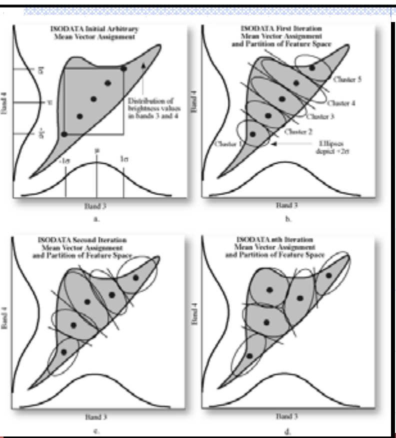
The number of classes can change because the computer is allowed to merge spectrally-similar preliminary clusters (i.e. clusters whose mean vectors are nearby) and to split clusters whose standard deviation within any single band is larger than a predefined threshold.

ISODATA

If splitting a cluster with a large standard deviation produces clusters that are smaller than an analyst-specified threshold, the new clusters are simply eliminated and their constituent pixels are reassigned to the remaining cluster whose mean vector is nearest.

As stated earlier, the process then repeats until either few pixels are being reassigned or a maximum number of iterations has been completed.

ISODATA



Digital Image Processing

Final Comments on Unsupervised Classification

Final Comments

Unsupervised classification procedures are designed to identify spectrally-similar classes of pixels.

They are incapable, however, of associating these classes with landcover classes. This associative process can be as difficult as the process of developing and refining the spectral classes that are used in supervised classifications.

Final Comments

There are three possible relationships between spectral classes and land cover classes:

- A one-to-one relationship,
- A many-to-one relationship, and
- A one-to-many relationship.

Many-to-one and one-to-one relationships are more common and are relatively easy to deal with. One-to-many relationships, though, are especially problematic.

Final Comments

A one-to-one relationship exists when each of the spectral classes represents a distinct landcover class.

If this type of relationship exists, the analyst only needs to recognize the relationships and assign appropriate class names.

Final Comments

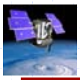
In a many-to-one relationship, two or more spectral classes are logically grouped to define a single landcover class.

For example, an unsupervised classification might produce distinct spectral classes that the analyst recognizes as deep clear water, slightly turbid lakes, and shallow muddy ponds. These can conveniently be assigned to a water landcover class unless the analyst is especially interested in the differences between these water features.

Final Comments

The analyst's job is more difficult, however, if one to many relationships exist.

For example, the analyst may wish to produce a classification that separates deciduous and evergreen forest types in a forestry application. If the computer generates three spectral classes which the analyst recognizes as deciduous, evergreen and mixed forest, these spectral classes don't provide any method to achieve the analyst's objective.



Summery

- Image classification techniques
- Unsupervised Classification
- K-Means Algorithms
- ISODATA