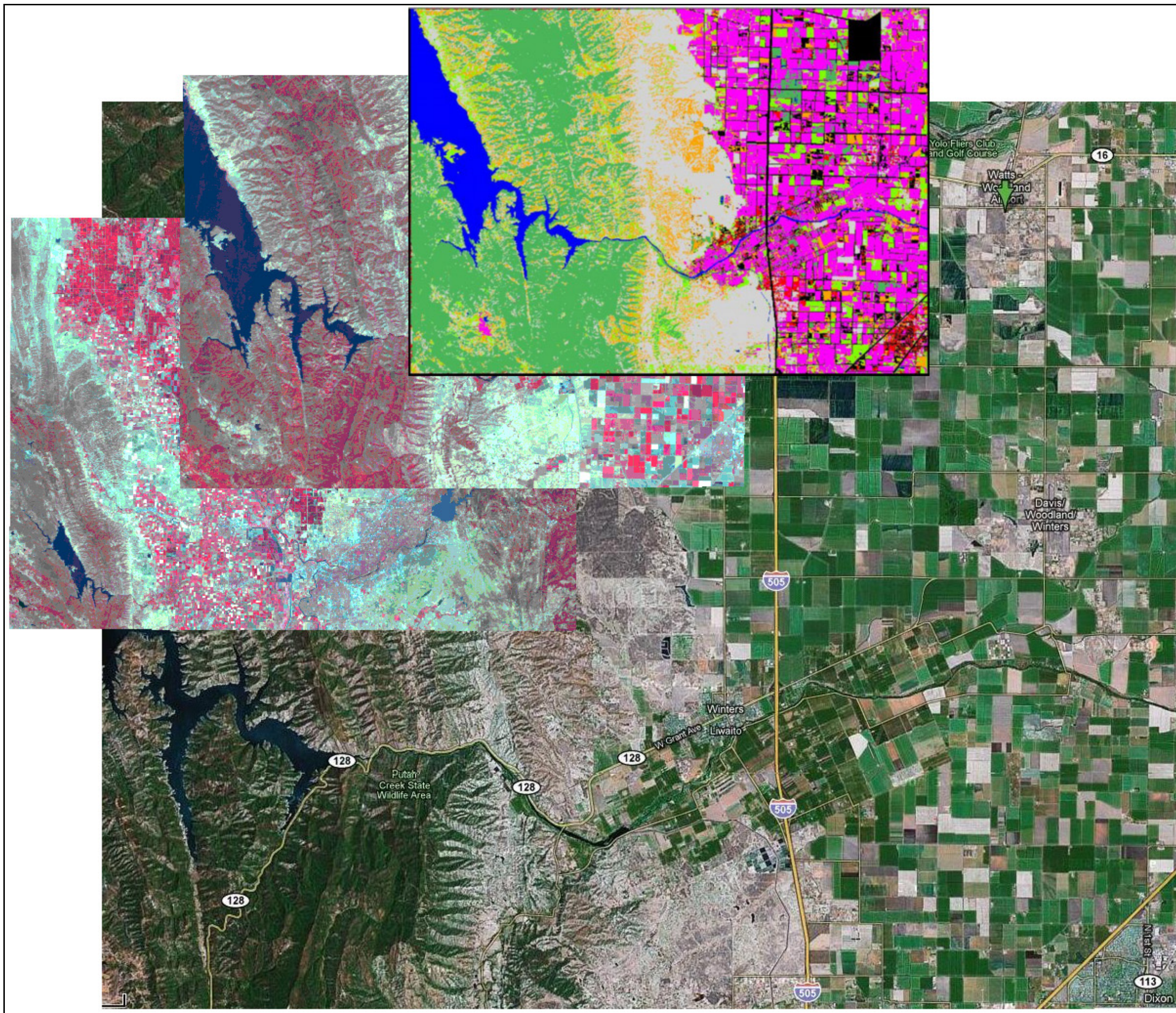


Unsupervised Classification of Spectrally Landsat TM Data

Completed by Valentina Constantinescu valcon34@pacbell.net

and Rohitesh Richardrohitesh70@hotmail.com

Table of Contents

1. Project Overview

The main objective of this project is to classify the vegetation, man-made structures, and miscellaneous objects from the Satellite Image of Dixon-Winters-Esparto-Woodland (Project Area), by using the software, ERDAS Imagine 2010.
The ERDAS Image software performs the classification of an image for identification of terrestrial features based on the spectral analysis. For classification of the Project Area the multispectral data was used for categorization of terrestrial features in specific land covers. There are two ways to classify pixels into different categories: supervised and unsupervised. The classification of unsupervised data through ERDAS Image helped in identifying the terrestrial features in the project Area. The spectral pattern present within the data for each pixel was used as the numerical basis for categorization.

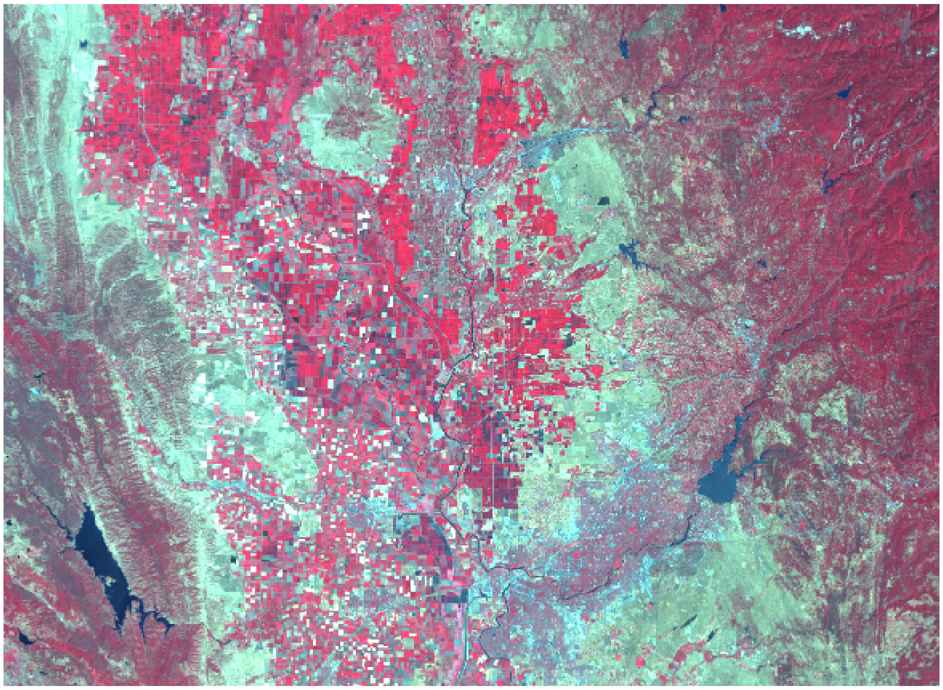
The report below will describe the methods used, issues encountered, and results obtained for this project.

2. Data Sources

For this project, I used a Landsat Thematic Mapper (TM) image covering a large portion of Northern California Area. The characteristics of the TM image I used are summarized at the end of this section.
After acquiring the image from Prof. Jennings, I selected a subset image that included only the area of interest (Dixon-Winters-Esparto-Woodland).
The image characteristics' information is listed below.

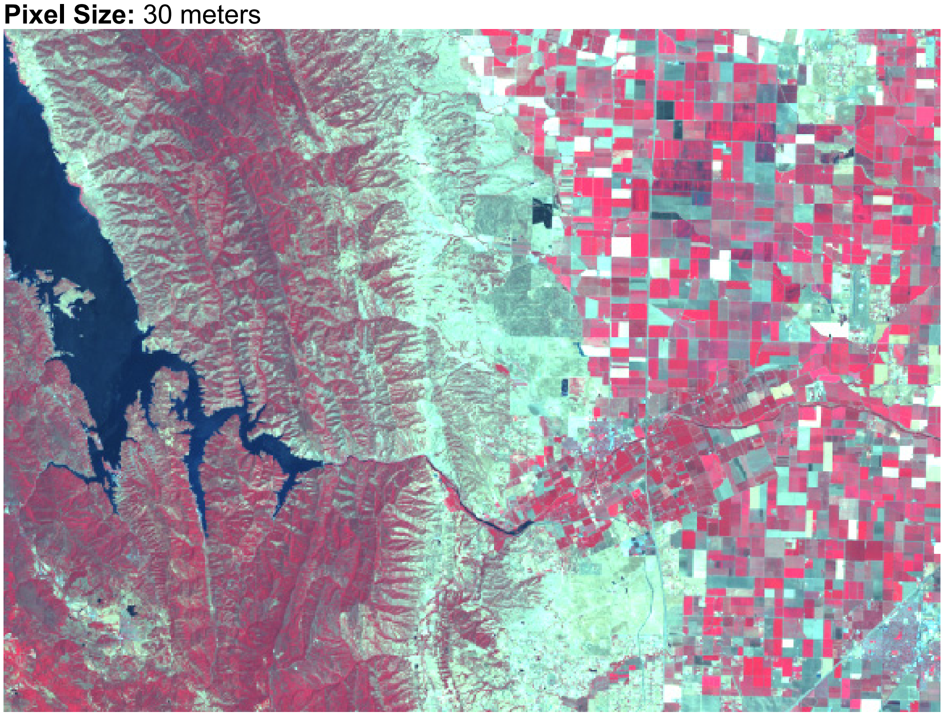
Original Image Characteristics

Satellite: Landsat **System:** Thematic Mapper (TM)
Last Modified: October 15, 2006
Number of Layers: 6
Pixel Size: 30 meters



Subset Information

Location: Dixon-Winters-Esparto-Woodland **Area:** Approx. 400 sq. mi.
Last Modified: April 11, 2010
Number of Layers: 6
Pixel Size: 30 meters



3. Method Used

Unsupervised Classification

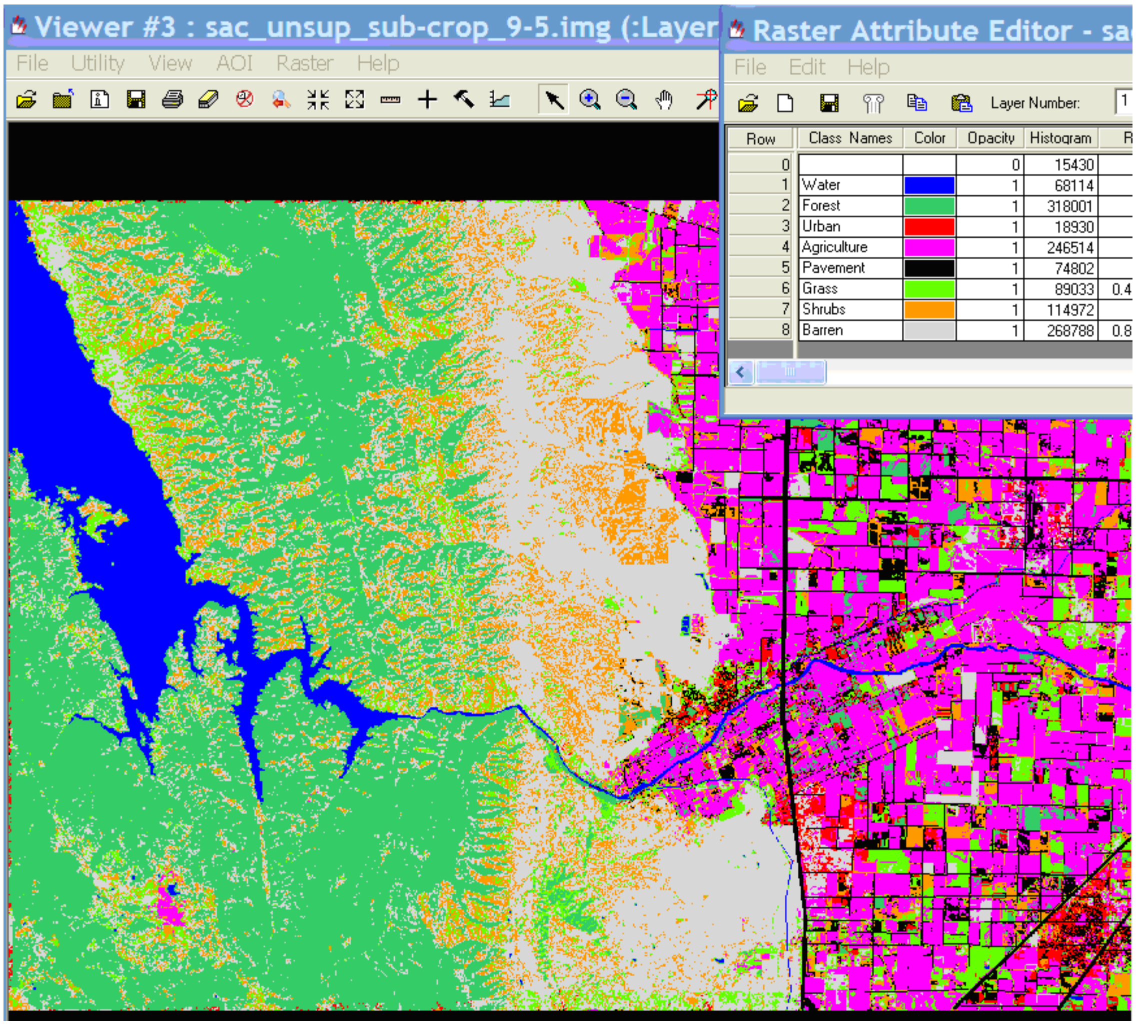
In a multispectral image, each pixel has a spectral signature determined by the reflectance of that pixel in each of the spectral bands. Multispectral classification is an information extraction process that analyzes the spectral signatures and then assigns pixels to classes based on similar signatures (Sabins 283). For example, all of the pixels which represent an area of forested land on a TM image should have roughly the same spectral signature. Classification procedures attempt to group together such similar pixels so that a GIS layer can be generated with each land cover type represented by a different class. The detail of the classes depends on the spectral and spatial resolution characteristics of the imaging system. Landsat TM imagery is usually good for creating a general land cover classification map.

Unsupervised classification is a method in which the computer searches for natural groupings of similar pixels called clusters (Jensen 231). Fewer clusters exist, more pixels within each cluster exist and will vary in terms of spectral signature, and vice versa. In ERDAS unsupervised classification is performed using an algorithm called the Iterative Self-Organizing Data Analysis Technique (ISODATA). Using this algorithm, the analyst input the number of clusters desired and a confidence threshold. The computer will then build clusters iteratively, meaning that with each new iteration, the clusters become more and more refined. The iterations stop when the confidence level (or a maximum number of iterations specified by the user) is reached (Jensen 238). For example, if the user wants 30 clusters at 95% confidence, the computer will iteratively build the clusters until it is 95% confident has attained the best distribution of pixels into 30 clusters.

After the clusters are built, the analyst must select the land cover classes (water, forest, etc.), then assign each cluster to the appropriate class. For this step, it is important user has a good knowledge of the region being mapped, since he or she must decide what land cover the pixels of each cluster represent. Once all clusters have been assigned to a class, the image of clusters can be recoded into a GIS layer which displays each land cover class with a different color.

Once the spectral enhancements were completed on my image, I performed an unsupervised classification (using ISODATA) with 15 classes respectively and a 98% confidence threshold. I set the maximum number of iterations at 15. The final result was an image with 15 groups of pixels each represented by a different color. I was able to highlight each class one at a time and then determine which of the classes it belonged to by interpreting the original multispectral image. Then I changed the cluster color to an appropriate one, for example, I made the water clusters blue. Finally, the image was recoded into the map shown below. The following table lists the land cover classes I was able to distinguish.

Class	Land Cover	Color	Pixel Value Before Recode	Pixel Value After Recode
1	Water	Blue	1	1
2	Forest	Dark Green	2,3	2
3	Urban	Red	4	3
4	Agriculture	Magenta	5-8, 12, 15	4
5	Pavement	Brown	9	5
6	Grass	Green	10, 13	6
7	Shrub	Orange	11	7
8	Barren	Grey	14	8



Note: "Class 0" just refers to unclassified pixels, of which there were none

4. Issues Encountered

4.1 <u>Choosing the appropriate number of classes</u>

For this step, it is important the user has a good knowledge of the region being mapped, since he or she must decide what land cover the pixels of each class represent. Choosing too many classes posed the problem of too many classes be recoded during the later part of classification process.
Whatever, choosing less than optimum classes for classification will result in too many pixels being misclassified and also in a heavily recode process.

Based on the Google Earth Image visual inspection of the project area, I realized that there are not less than 7 classes. So, I decided to go for classification with 15 classes. This gave me the expected seven classes and additional "mixture classes" making the recode process for misclassified pixels easier.

4.2 <u>The time of year when the original satellite image was taken</u>

To make a land cover classification map the time when the satellite image was taken is very important. The vegetation and diverse types of vegetation give a different spectral signatures over the year and also some crops as alfa-alfa would appear similar to grass, especially if they were very low to the ground, the orchard would be misclassified as forest or shrubs (depending how mature are the tree) and the bare field would be misclassified as barren.

4.3 <u>Separating the classes. Recode process</u>

The classes easier to separate one from another were **the forest, barren area, and water features.**

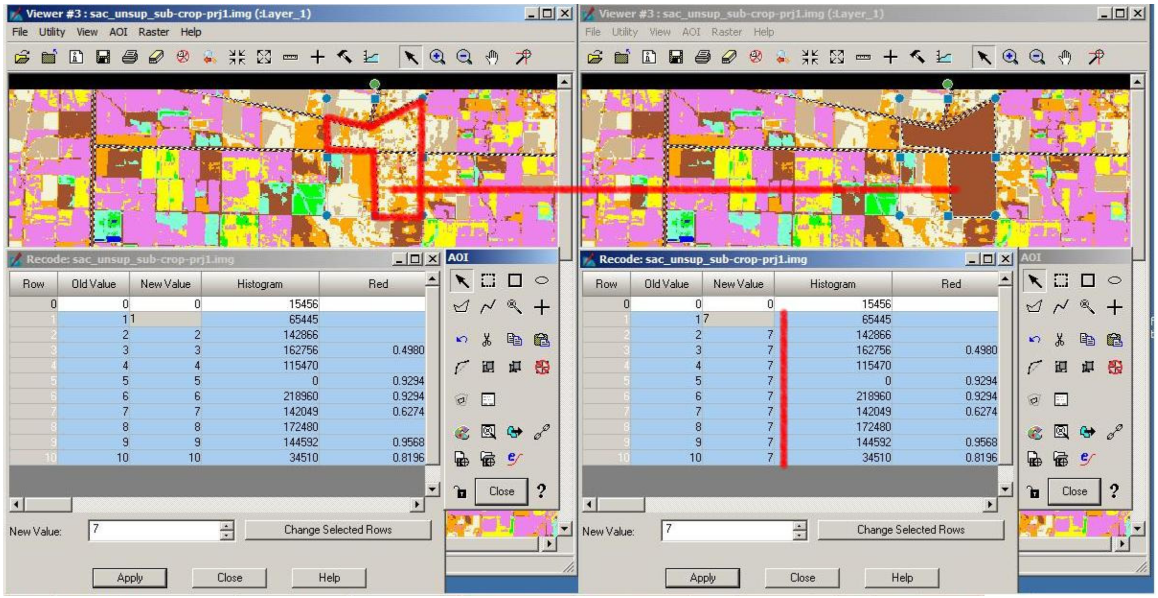
As I expected, **the urban areas** were classified as a mixture of agriculture-grass-forest-pavement, with some patches of barren due to features such as parks, green area, parking lots, and bare fields.

This occurred because some urban features have similar spectral signatures as agriculture/grass/forest and vice versa.

By using the recode process, I recoded those specific pixels emphasizing the predominant desired class.

The airports and new house development areas are classified as barren areas, although they are speckled with pixels classified as urban, which are obviously misclassified.

When I recoded those areas, I decided to recode using the “all-in-one” class. I selected the whole area and assumed same value to all pixels. The result looks like a “patch” cover.   I consider this was not an good idea.



However, I think the spectrally enhanced layers did a pretty good job of separating these classes, especially apparent in the fact that you can see the main roads (linear dark features) in the agricultural areas.

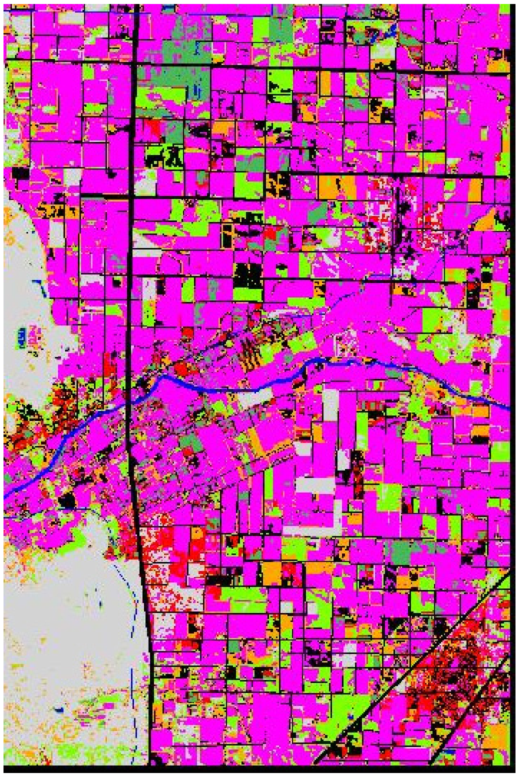
Anyway, I emphasized the highway and the large road recoding them and assigned them to a specific class named “pavement”

The biggest problem with the classification was the issue of crops. The image had a large agricultural area. Since there were areas with different type of crops, it was too difficult to create separate classes called "crops" or “orchards” without creating a lot of misclassified pixels in other areas.

Crops in the image were grouped with forest, shrubs, or grass and the bare fields were classified as barren. It is understandable, though, that some crops would appear similar to grass, especially if they were still very low to the ground, the orchard would be classified as forest or shrubs (depending how mature are the tree) and the bare field would be misclassified as barren.

Also all crops rotation process makes difficult a correct classification in specific classes (based on crop type) for such a large area. So, I decided to define a class named Agriculture which included all type of crops and covered all the agriculture designated area.

In the recode process I decided to keep the lot boundaries (I used the AOI Tools to draw lines and recode them as pavement) and to keep some misclassified lots only to emphasize the large variety of the crops.



5. Conclusions

This project has helped me learn a lot about several aspects of digital image processing. The results met my objective of finding out whether or not spectrally enhanced layers of data could be used to generate a good land cover classification map of the Project Area. As explained above, I feel the unsupervised classification using the enhanced data was a success.

Although this project may not make any significant contributions to digital remote sensing research or a specific real world problem, the experience will definitely be valuable to me for the future GIS and remote sensing projects I might be involved with.

6. References

Lillesand, Kiefer, Chipman. *Remote Sensing and Image Interpretation*. 6th edition

Corner, Brian. *Principal Components Analysis*

Sabins, Floyd F. *Remote Sensing: Principles and Interpretation.* 3rd edition.