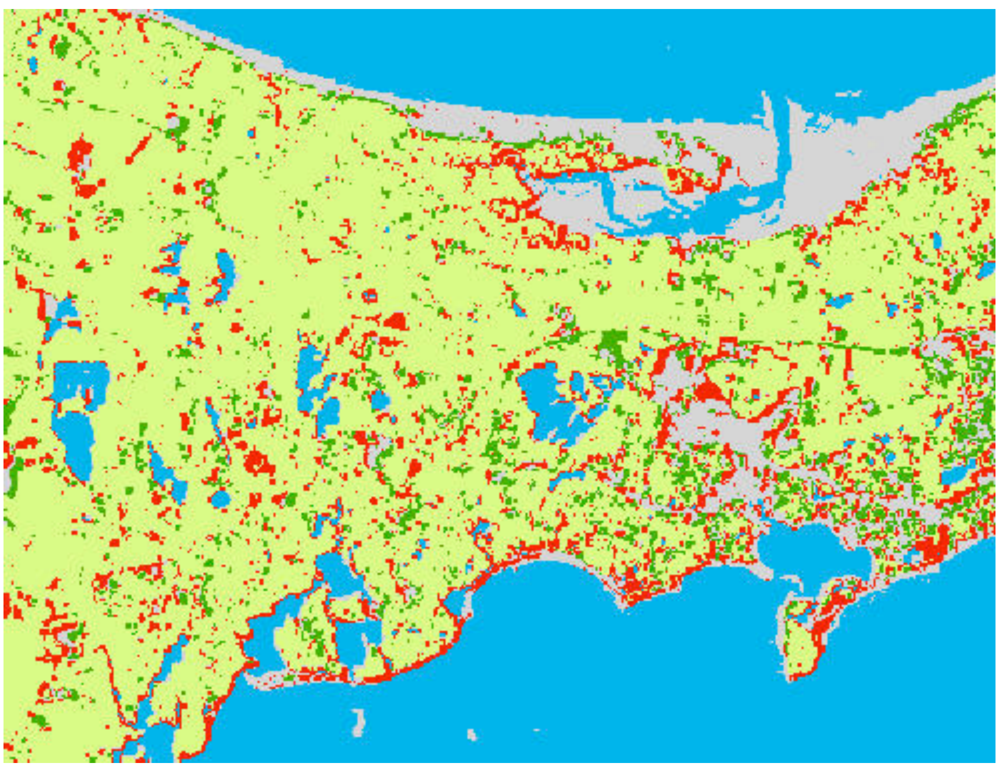
Classifying Multispectral Images

Based on an exercise by Paul Cote, Graduate School of Design, Harvard University

Background

The purpose of this exercise is to gain some hands-on experience with the fundamentals of image classification. Multiband images of the Earth's surface are a very important source of information about land cover and land use. Because satellites beam back information every day, this imagery can be a terrific source of both current and historic information. Subtracting the historic from the current can provide an estimate of change in the landscape, provided one can be sure that the classification of the images yields consistent results.



Classified image showing changes in areas that were (if we trust our classification) cleared (red) and places that were (evidently) reforested (dark green) between ~1977 and 1999.

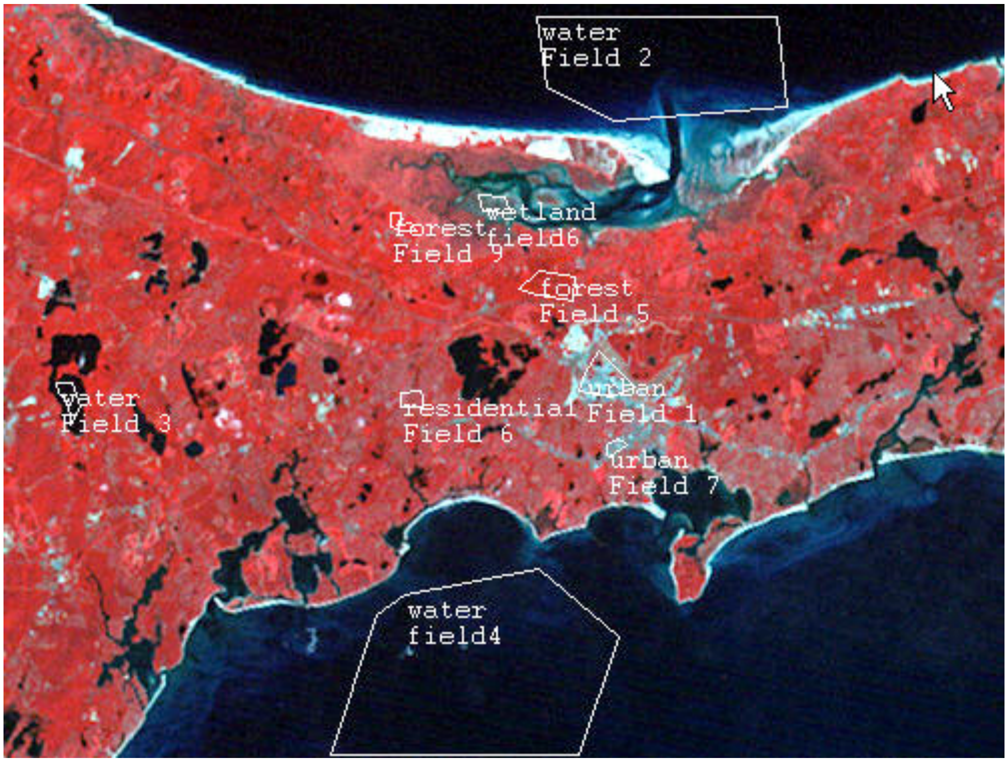
The exercise includes Landsat images from 1977 and 1999. We also have GIS land use polygon data from 1971 and 1999 as a reference to compare with the images. The images below provide examples of the data available in the exercise, and more.

1977 Landsat image of Cape Cod.

1999 Landsat image of Cape Cod.

1999 Landsat image around the town of Barnstable, MA.

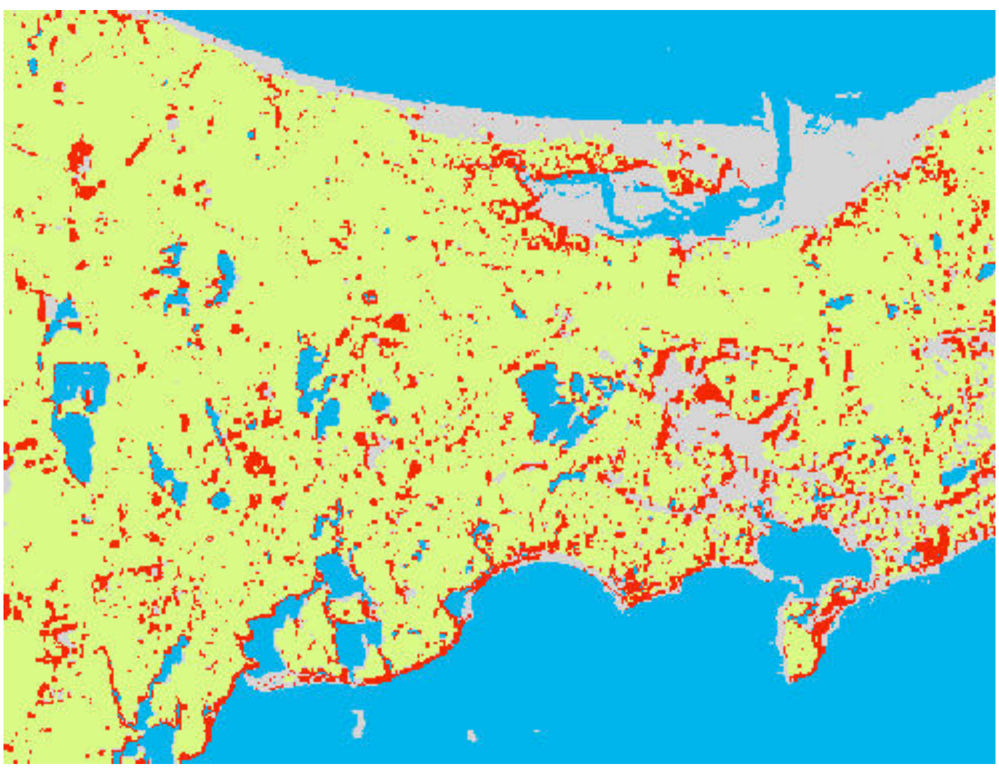
1999 GIS land use classification around the town of Barnstable, MA
are overlain on the Landsat image.



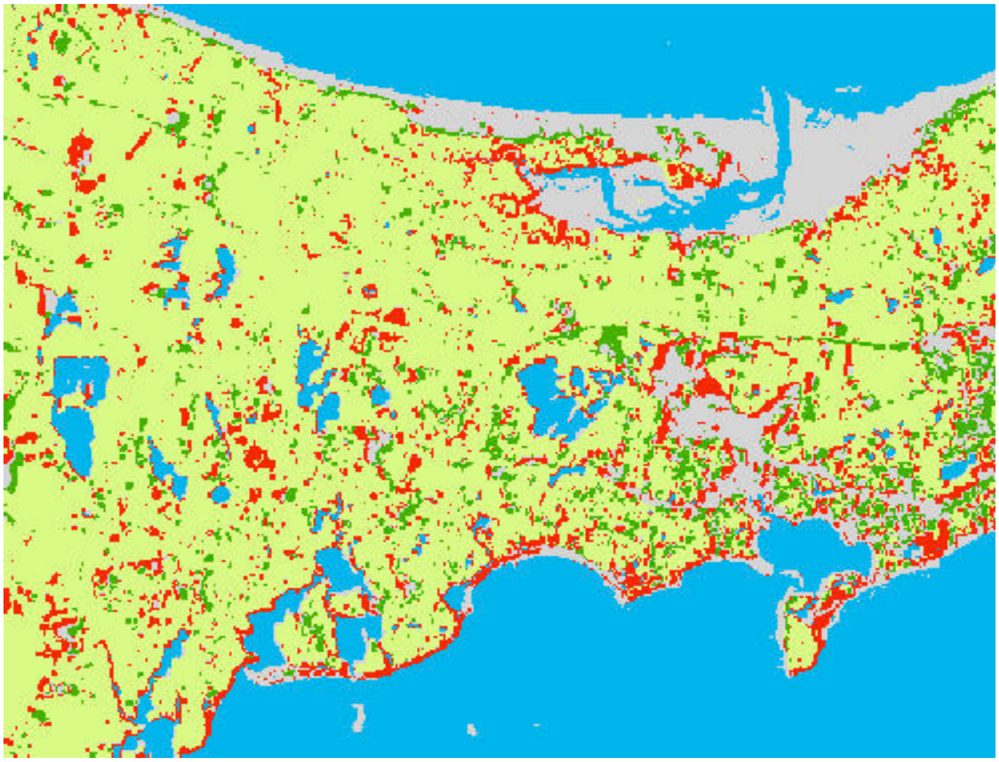
Sample training polygons created for the Barnstable, MA region.

Classified 1977 image in the Barnstable region where
blue = water/wetland, green = forested, gray = non-forested

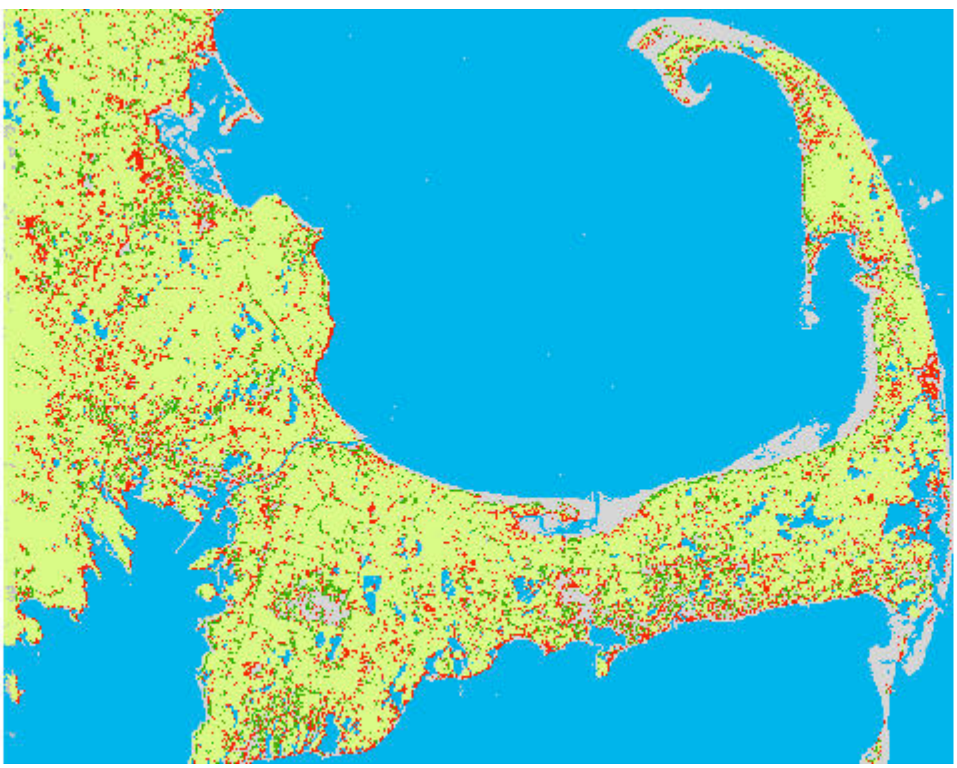
Classified 1999 image in Barnstable, MA region where
blue = water/wetland, green = forested, gray = non-forested



Areas in red are where forest was lost between 1977 and 1999 near Barnstable



Forest gained (dark green) and lost (red) between 1977 and 1999 near Barnstable.



Forest gained (dark green) and lost (red) for entire image region between 1977 and 1999.

Remotely-sensed data are very unstructured, meaning that they are, in most cases, simply a record of relative reflectance (sometimes emittance) of particular wavelengths of radiation within the electromagnetic spectrum. Hence there is usually a lot of interpretive work that must go into extracting useful information from these images. There are many software programs that can assist in this task. This exercise has been written for a very good, very compact and *free* image processing package called MultiSpec, from Purdue University (http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/).

In this exercise, we will investigate changes in land use on Cape Cod, Massachusetts. The setting is interesting because of the sensitive landscapes which were under considerable development pressures in the latter part of the 20th century. This pressure led to the establishment of the Cape Cod Commission (http://www.capecodcommission.org) in 1990 and the Cape Cod National Seashore (http://www.nps.gov/caco) in 1961.

Introducing the Data

The images are from the Global Land Cover Facility at the University of Maryland (http://glcf.umiacs.umd.edu), and represent multispectral, multiband output from the MultiSpectral Scanner (MSS) and the Enhanced Thematic Mapper (ETM). The MSS was a sensor aboard the first Landsat satellite that was active in the 1970s, and the ETM is part of the Landsat 7 satellite mission. The Landsat Technical Documentation at http://glcf.umiacs.umd.edu/library/guide/techguide_landsat.pdf provides more information.

It is common practice to use the red, green, and near infrared bands to differentiate vegetation and development, so we have chosen the appropriate image bands to include in our tif files (which can only hold three bands).

Part of the problem with classifying an image is making training samples based on some understanding of land cover in the area of interest. To aid in this process the exercise includes a land use GIS shapefile for the town of Barnstable, MA, from Massachusetts State Office of Geographic and Environmental Information (MassGIS; http://www.state.ma.us/mgis).

Opening and Exploring the Data in ArcGIS

ArcGIS will provide a useful reference for the satellite imagery before classifying it using MultiSpec. It can also be invaluable in further analyzing the results of the image classification process.

1. Copy the **capecod** folder in **\Data\Image_Processing** to your student folder.
2. Open the ArcMap Document, **map1**.mxd.
3. Explore the images **1999_etm_1234.tif** and **1977_mss_123.tif**, as well as the land use data **landuse 1971** and **Land Use 1999**. Make sure to expand the land use data so as to see all of the various classes.
4. Zoom into the **Land Use 1999**. Use the Identify tool to click around on the land use polygons as well as the pixels in the **1999_etm_1234** image in order to get a feel for both land use categories, as well as the spectral signatures adjacent to them.
5. Keep ArcMap open.

Performing an *Unsupervised* Classification in MultiSpec

…where the software will categorize the pixels in the data solely on the basis of image statistics, without any training samples or prior knowledge (i.e., supervision) of the area.

1. Double-click on the MultiSpec desktop icon to open the MultiSpec application.

2. In MultiSpec, use **File →Open Image**, and navigate to your **capecod** folder in your student account so that you can open the **1999_etm_234** image. Accept all the defaults and then click OK.
3. Zoom in and out using the "mountain" buttons on the tool bar at the top of the MultiSpec window.
4. From the **Processor** menu, choose **Cluster.**
5. Click **Single Pass**, and accept the defaults in the single-pass options dialog. Then in the Cluster dialog, click **Image Area**. If your image area parameters do not start at 1, click the little "Full Image Area Icon" at the left side of the little image area section of this window.
6. Set **Write Cluster Report/Map** to **Text Disk File** (therefore unclick **Text Window**)
7. Then click OK to begin the clustering. Save to the same file prefix.
8. Do not be alarmed by the **Project** window. Just move it to the side.
9. Once the clustering is finished you can open the result as an image. You will have to set the **Files of Type** pull-down to **All Files** in order to find it. The file will have a **.clu** suffix (**1999_etm.234.clu**)

Unsupervised Clustering Results

Clustering is a good way to start an image classification because there are classes of pixels that are decipherable because of their distinct signatures. Some of these signatures will be readily identifiable as land cover classes (e.g., water). But the clustering algorithm has not done such a good job distinguishing certain other classes, such as wetlands or residential.

Fine Tuning Cluster Classes Based on GIS Data

1. We will change the color of the clusters on the image to make them easier to associate with land-cover characteristics.
2. Notice the **Project Window** that showed up on the right side of your screen when you did the clustering. It has some buttons on it called >**Classes** and >**Fields**

3. Put the project menu in **Classes** mode.
4. It will display a list of the classes that were found by the clustering. These correspond with the classes in your clustered image.
5. Rename classes to the categories were below, which were derived by a close examination of the land use polygons in ArcMap. Choose an entry from the list and click **Edit Class Name.** Change …

   - Cluster 1 to Water-Based Recreation1
   - Cluster 2 to Low-Density Residential
   - Cluster 3 to Water-Based Recreation2
   - Cluster 4 to Open Land
   - Cluster 5 to Water
   - Cluster 6 to Saltwater Wetland
   - Cluster 7 to Forest

These categories are, of course, not perfect. Why??

Performing a *Supervised* Classification in MultiSpec

…where the software will categorize the pixels in the image as guided by prior knowledge of the location and identity of land cover types in the region. This knowledge comes from field work, study of aerial photographs, or other independent sources of information. In this case, the main source of information is from the land use polygons already provided for us in ArcMap.

Now, we will add our own input to help the classifier distinguish features that we feel are important.

1. Make your original image the active window.
2. Set your **Project Menu** to **>Classes**
3. Now choose **Classify** from the **Processor** menu. In the **Set Classification Specifications** dialog, make sure to uncheck the box that says **"Write Classification Results”** to **“Image Windows Overlay"** and to check the one that

says **"Disk File"**, and for **ERDAS.GIS** (there are is no ESRI ArcMap option). The rest of the defaults should be fine. Save the result to the same file prefix.

4. Once the classification is finished you can open the result as an image. You will again have to set the **Files of Type** pull-down to **All Files** in order to find it. The file will have a **.gis** suffix (**1999_etm.234.gis**)
5. Look at the image. You can also go to your **Text Output** window (bring it to the front by selecting it from the Window menu) and look at the statistical summary of your classification.
6. In order to see more information about the signatures that MultiSpec has identified from your reclassification, go to the project window, push the **"Classes"** button and choose **"List Stats..."** from the pull-down menu. Additional statistics for each class will be listed at the bottom in the **Text Output** window.
7. Save your classified image by going to **File →"Save Image to GEOTIFF As"** and naming it **"Supervised_1999.tif"**
8. Save your MultiSpec project by going to **File →"Save Project As"** and naming it **"Supervised_1999"** (and save your Text Output too, to **Text_1999).**

Open the Classification Result in ArcGIS

Now that you have extracted some new information from your satellite image, you may want to bring it into ArcGIS to use it with other datasets, including the results of other image classifications. MultiSpec makes this fairly easy, as it allowed you to save the image as a **geotiff,** meaning that is already *georeferenced*, and thus should automatically line up with your ArcGIS layers.

1. Click the **Add Data** button in ArcMap.
2. Connect to your data folder.
3. Select **Supervised_1999.tif**
4. Build pyramids

5. Once the new image is brought into ArcMap, it should align with the other data. Explore the fruits of your labors!

Classifying the 1977 Image

Create a new MultiSpec project with the 1977 image, **1977_mss_234.tif**

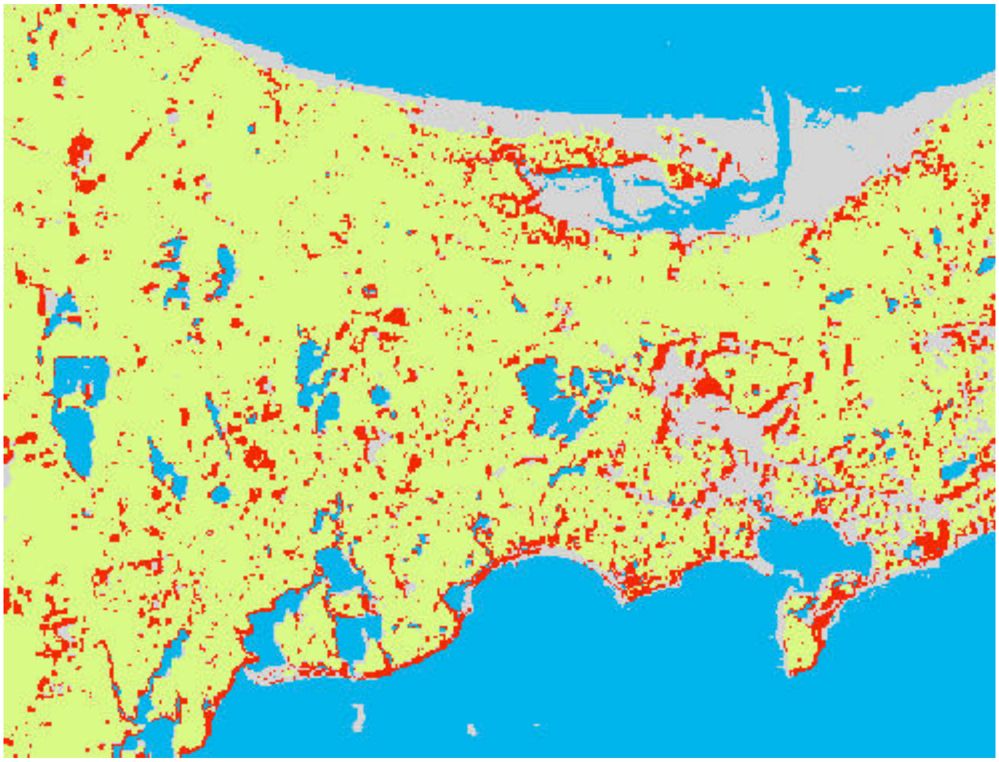
1. Repeat the steps for both unsupervised and supervised classifications on this image. Change…

    - Cluster 1 to Open Land
    - Cluster 2 to Forest
    - Cluster 3 to Water
    - Cluster 4 to Saltwater Wetland
    - Cluster 5 to Forest2
    - Cluster 6 to Low Density Residential

2. Saved the supervised classification to a geotiff named **supervised_1977.tif**

Bring **supervised_1977.tif** into your ArcMap view, zoom into the Barnstable region and note the differences in the pixel classifications between this 1977 image and the 1999 classification (**supervised_1999.tif**).

Here we are just eyeballing the data to note changes in land use over small areas. But in a much more extensive image processing exercise we would systematically extract the differences between the 1977 and the 1999 images. This is beyond the scope of our IDES workshop, but as an example, ArcGIS has a function in its Spatial Analyst extension called **combine**. For each cell in the output grid, **combine** assigns a new value representing the combination of or change between values for that location in both the input and output grids. So in our analysis, one of the zones in the combined grid would be for those cells that were classified forest in 1977, but were then found to be unforested in 1999, likely to due to residential or industrial development.



Areas in red are where forest was lost between 1977 and 1999 near Barnstable, resulting from a systematic comparison between image classification for both years.

The End! Congratulations!