Exercise #7a - Classification

• 7.1 Introduction to Different Classification Methods

• 7.2 Supervised Classification

Note: Make sure you read this exercise. Otherwise you might not do the steps correctly. Also, remember that if you want more information on how a part of Imagine works, you can try the Help button.

Objective: In Exercise 7 you will learn about two different types of classification: Supervised and Unsupervised classification. In Supervised classification you will define a set of training areas which will be used in a Maximum Likelihood Classification. The Unsupervised classification will be an ISODATA clustering.

7.1 Introduction to Different Classification Methods

In image classification, the aim is to categorize all pixels in an image, for example, into land cover classes (e.g., forest, water, agriculture...). Two classification procedures commonly used that we will do in Lab #7 are Supervised Classification and Unsupervised Classification.

Supervised classification first requires input from the analyst about how to assign land cover classes to the pixels. The input from the analyst is often called "training" data.

Unsupervised classification first clusters the data using statistical methods, and then the analyst assigns land cover classes to the clusters.

In this lab, pretend you are working as a Remote Sensing analyst at the National Mapping Agency. Your boss comes in, and asks you to make a vegetation map over the part of Carpathian Mts. area. She gives you the satellite data, some forest stand maps and some aerial photographs. She asks if you can try to map the following classes:

- Water
- Agriculture
- Coniferous Forest
- Deciduous Forest
- Clouds
- Bare Soil

The data

The image you will be looking at:

- File name: tm070717.img
Path/Row: A subset from 185/026  
Sensor: Landsat 5 TM  
Pixel size: 30m x 30m  
Acquisition Date: July 17, 2007  
Bands: 6 band file - TM bands 1, 2, 3, 4, 5, 7  
Correction level: Geometrically corrected UTM, Zone 34  
Geographic area: Bohorodehany - Nadvirna, Ivano-Frankivska district

7.2 Supervised Classification

Open image tm070717.img. Use the Band combination TM band 4, TM band 3, TM band 2 (RGB).

Before you get going with the rest of the exercise, there are some areas within this scene that I have labeled with their cover types. Bring up the annotation file test.ovr. This is for your information, so you can see what some different cover types look like. Use the inquire cursor to see what the pixel values are (You'll either have to a) use Arrange Layers, and change the order of the two files in order to use the Inquire Cursor, or b) you can open a 2nd viewer with only the satellite image in it, and Geolink it to the viewer with the annotation). Change band combinations to the most common ones used for viewing: a) TM bands 3, 2, 1 in RGB; b) TM bands 4, 3, 2 in RGB; c) TM bands 4, 5, 3 in RGB)

Also keep in mind the following illustrations to guide you as to what land cover types the DN values might refer to. Keep in mind that the DNs are never exactly the same in two different images for the "same" land cover type. A pine forest in one image will have different DNs from a pine forest in another image, or even another area. There are many factors that influence the DNs: atmosphere, the season the image was taken, the site conditions, the crown closure - just to name a few. However, the relative patterns with the DNs may be similar.

Training set selection

Lillesand and Kiefer's textbook (Page 544) says that "Whereas the actual classification of multispectral image data is a highly automated process, assembling the training data needed for classification is anything but automatic. In many ways, the training effort required in supervised classification is both an art and a science."

"Training areas" are collected by drawing a boundary around areas that are representative of the cover-types you want to map in the image. In supervised classification, these training areas are used to "tell" the computer program what the classes you want to classify should look like spectrally.

Your training data should be from independent data sources such as airphoto interpretation or field sampling. These data are often called "ground truth". You also need an independent
source of data to check the results of your classification (also called ground truth), and these cannot be the same sites as you used in your training areas.

When you are making a classification, you should first decide on the classes (and their definitions) you can map in your image, considering the data you have (for example, the field data and satellite data). These classes you want to classify are called "information classes".

Several training areas are collected for each information class, in order to get a good representation of that class. Spectral signatures for each information class are calculated using the DNs from the training areas.

Open the Signature Editor by going to the Classifier Icon on the Main Icon Panel, then Signature Editor. This opens up the tool that will contain your Signature Set. The three main buttons you will be using are defined below.

You will be collecting training areas (all the areas assembled together is called a "training set"). The ground truth data you will use is a combination of aerial photos, the stand maps you used earlier in the photo-interpretation exercise, some other map data, and also a bit of your own visual interpretation of the satellite imagery.

Training areas should be homogenous (meaning that the appearance due to the DNs of the pixels, should look quite the same). There are no rules for a minimum or maximum size of a training area. Certainly the training area should not be so large that it includes non-representative pixels. Neither should it be so small that you do not get several pixels to
represent your class. As a rule of thumb, I try to draw a training area that contains anywhere from 5 to 30 pixels. But, as said before, there are no real size limits.

Lillesand and Kiefer's textbook offers a few guidelines on training area selection: 1) it is not uncommon to acquire 100 or more training areas when classifying an image; 2) in theory, the minimum total number of pixels in a training set is the number of spectral bands + 1, in practice the number of pixels in a training set is a minimum of from 10 x number of bands, depending on the classes one wants to classify in the scene.

When you draw your training areas, you will want to be zoomed in to do this work. You can either zoom in and out of the image in your Viewer, or you can create a Magnifying window in which to work. Do this by going to the View/Create Magnifier on your Viewer Menu. Do whichever you feel is easier.

Go to the Viewer menu item AOI/Tools. This opens the AOI tool palette. AOI stands for Area Of Interest. You will use this to draw the training areas. There are a couple different ways to draw the training areas just now. The first one we will use is Manual Delineation; Choose the Create Polygon AOI tool.

Bare Soil training areas using Polygon AOIs

For the information class Bare Soil, and knowing what you know about how bare soil should look spectrally (refer also to “Typical Spectral Response Patterns”), draw an AOI polygon in a bare soil area. Because you've just created this AOI, it is still "selected". Then in the Signature Editor, click on the Create New Signature from AOI button. This takes the selected AOI and adds it into the signature set. Under the Signature name, change this to Bare. Change the color for the Bare Soil classes to bright pink.

Water training areas using Region Growing
Now you'll try another way of collecting a training area, but this time for the Water class. In Imagine it's called "Region Growing", but in other literature it's also known as "Seeding". Go to a body of water in the image. Draw a quite large AOI (using the Create Polygon AOI tool) within this body of water - you don't need to draw it too carefully. Then go to the Region Grow Properties button on the AOI Tool palette (see above graphic).

Pressing the Region Grow Properties button brings up a dialog in which you will be setting some constraints to tell the Region Grow how many pixels and how much spectral variation you want to have. Change the geographic constraint distance to be 30 pixels (so you don't create too large a polygon), and the spectral Euclidean distance to be 3 (this is in DN. The choice of these numbers depends on the number of training areas you might choose, and the broadness or narrowness of the spectral signature of the information class you are choosing a training area for). Under Options, make sure Include Island Polygon is turned on. (Update Region Mean is also turned on, but it doesn't matter for us just now.) Close the options and Close the Region Growing Properties dialog.

With the Region Grow, you are going to pick a single pixel within the AOI you just drew. Region Grow then selects other pixels which are spectrally similar. Go back to the AOI Tool Palette and choose the Region Grow AOI, and left click it on the class you are targeting (in this case Water) within the previous large AOI that you drew.

This makes a new AOI (and keeps it selected), with at most 30 pixels that are all within 3 DNs of each other in each band. Add this AOI to your Signature Editor, and label it Water. Change the color of the class.

**Saving your work!**

Once you have done this, make sure you save your Signature File now, and sometimes throughout the exercise. In the Signature Editor, go to File/SaveAs, and call it
Sigl.sig. Until you have finished collecting all the training areas, you will save this file as Sigl.sig.

Also save your AOI file now, and sometimes throughout the collection of training areas. To do this, in the Viewer where you are collecting the AOIs, go to File/Save/AOI Layer As, and save this as Training.aoi.

Continue, and collect 5 training areas each for bare soil and water.

---

Tip: If you do choose to draw the training polygons in the Magnifying window, you will have to close the magnifying window after you draw the polygon, and then choose “save”. Otherwise, it will not save your polygons! If you are trying to draw an AOI in your Viewer, and it just doesn’t work, it probably means that your AOI layer is not the Top Layer. Go into View/Arrange Layers, and move the AOI layer to the top.

---

Viewing your training areas with Image Alarm

There is a tool in the Signature Editor that does a quick evaluation of other pixels in your image that a training area is spectrally similar to. You can use this function to see whether the training area you have chosen is approximately a reasonable one or not. To do this, first fit the image to window. Then, select the training area you want to test, by selecting the row of that area. For example, start with your first Bare Soil training area. Make sure the little > sign is pointing to this signature (See graphic below). Then in the Signature Editor, go to View/Image Alarm. This brings up another dialog box. Here you can accept the defaults. (The viewing of overlap areas you can do if you choose several training areas at once, to see where the signatures might overlap.) The areas that this training area might classify is highlighted in whatever color you have for that class in the Signature Editor.

Try using Image Alarm to check some of your classes and see how this works. You can use it again, anytime during the exercise. It should highlight the pixels which are most spectrally similar in the image and give you an idea whether your training set is good or not. If you have an Image Alarm which highlights the whole image, this tells you that this training area is too broad and too general spectrally, and will look like many other classes in the image. If the Image Alarm doesn't highlight any areas, then the training set may have too narrow a spectral range.
If you do get such a problem signature in this part of the exercise (one that highlights almost everything in the scene), you can delete that area by clicking on the row, then right clicking and choosing Delete Selection. Then add a new training area to replace the one you deleted.

**Bog training areas and Importing an Arc/Info coverage**

For Bogs, you will use the 1:100,000 scale "Blue" landuse map (Bla kartan) to identify bog areas. The map sheet name you will use is called "y_21jno". First, you will need to import the map sheet, which is an Arc/Info vector coverage. Go to the Import button on the Main Menu Icon Panel. Under Type choose Arc_Interchange to Coverage, and under Media choose File. For the input file you should have file y_21jno.e00, and the output file (call the output y_21jno) should be directed into your home directory (N: drive). That's because it's important to not just copy or move around this Arc/Info file y_21jno. Arc/Info files require a special copying, and will not work if they are just copied using a regular Windows copy). The .eOO extension of the input file is the Export version of an Arc/Info file which you can copy freely. (See the graphic below for how the Import dialog should look approximately.)

Choose OK on the Import/Export dialog box. Then a dialog box appears, in which you specify that you want to import a COVERAGE. Say OK. Now it should start importing the coverage. Close the Import/Export dialog box.
In the Viewer with the satellite data, Go to File/Open/Vector Layer, and choose the file y_21jno. Under the tab Vector Options (just like the Raster Options tab), make sure the Clear Display is not checked on, and make sure the Use Symbology is Checked on, and the Option says that it is set to the file y_21jno.evs. The .evs file is the color scheme for y_21jno, and has been pre-set. In this map, there are three different bog types: the light brown color is wet bogs, or "bla myr", the darker brown color is dryer bogs or "brun myr", and the orange color is forest covered bogs (Skogkladdmyr). Water is blue, Urban is pink, Forest is green, and Open Land is yellow in the Blue Map.

Look at View/Arrange layers. You can see that you now have three layers in your Viewer. AiTange it so that the image data is the top layer, and the vector file is the second layer, and the AOI is the third layer. Make sure to hit the Apply button after you Arrange Layers. Use the Swipe function to identify training areas in the satellite imagery that will be good for the two bog types wet bogs and dryer bogs. Create AOIs either with Region Growing or with a Polygon AOI. Remember to Arrange Layers and raise your AOI layer to the top in order to add an AOL. This means that as you try to find and draw AOIs, you may have to do a bit of Arranging the Layers back and forth.

Also consider that if you use Region Growing, you may have to change the limits we set earlier for water training areas. Notice that there is much more variability in the bog areas. This means that you may have to try to collect training areas that represent some of this variability. Label these training areas Wetbog1...Wetbog5 (or however many you collect), and Drybog1...Drybog5.

Save your signature file (Sigl.sig).

**Regenerating Forest**
Regenerating Forest is defined here as stands recently cut, and we'll select stands with volumes of 1 to 15 m³/ha. Identify these areas from the stand maps (Stands 71, 135, 170, 195 or 46 for example). Choose 5 training areas for regrowth using Polygon AOIs and call them Regrowth1,... Regrowth5.

Save your signature file (Sigl.sig)

**Forest training area selection**

For the forest classes, you will use the SPOT 5 image with 5m/pixel spatial resolution as ground truth. Here the forest areas will be defined as stands with a volume of approximately 30 m³/ha and upwards. It is known that there is both Norway Spruce and Beech in the mountainous area. One of the stands is quite small, and hard to identify pixels that would make a good training area.

So, in this case, we don't have enough ground truth to collect a good representation of spruce vegetation. Therefore, we're going to have to tell the boss that we don't have enough ground data to classify Spruce as it's own class, but will instead decide to classify both Pine and Spruce together as a "Coniferous" class.

Collect one training area, where the pixels are homogenous (like below). In the Signature Editor, for the Signature Name, call this Coniferous1.
Now you will collect training areas for other Coniferous forest areas. When you look at the image, you can see there is a wide variety of spectral patterns. This can be due to many reasons, including the species mixture or the density of the forest there. It is a good idea to try to get a representative variety of the spectral variability that you see. Take these training areas using Polygon AOIs. Call them Coniferous2...up to however many you take.

Continue taking Coniferous forest training areas, but identify stands that might have mixes of Pine and Spruce (and not so much Deciduous). These training areas should also be called Coniferous and have a number after the name (e.g., Coniferous7, if it's your 7th Coniferous training area.). Now take a few training areas in more sparse Coniferous areas (less than 75 m3/ha but greater than 30 m3/ha).

Don’t forget to save your signature file (Sig1.sig). For the practice of using aerial photographs, use the black and white photos to identify two more training areas for the Coniferous information class. You will have to use what you learned in the previous photo interpretation labs here, to try to identify Coniferous forest using only black and white photos. Color-IR photos would be more ideal, because you would then see more of a difference between Coniferous and Deciduous trees. However, we would need a light table up in the computer lab to look at the negatives that we have. If you happen to choose a wrong area, it will show up later, and you can delete it then if it does.

We will also try to classify Deciduous forest. Using the stand maps, look into stands that have almost completely Deciduous volume as the dominant factor. This is a very small area, but for the sake of this exercise, try to take about 3 training areas here in this area. This is not a good
habit (taking training areas all from the same spot), but we will analyze the results later, and this could be instructive.

When you are done collecting all of these training areas, your Sigl.sig file should look approximately something like below (But NOT Exactly! Because you chose your own different training areas!)

Running the Supervised Classification

You will now try the first run of a Supervised Classification. This area is quite small, and the classification goes quickly. It is most likely that you will need to adjust your training areas, perhaps adding or deleting signatures. In Part II, you will learn more about evaluating those signatures, and run the classification again. But for now, it's good to get a rough idea of how good your current training data might be.

In the Signature Editor, go to Classify/Supervised. This brings up a dialog in which you will give a new output file name. Call this Superl.img. You will not output a distance file. You will use the Parametric Rule of Maximum Likelihood for this Supervised classification. Press OK.
The Maximum Likelihood Classification is a process in which the probability of an unclassified pixel belonging to a user-defined class is calculated. The unclassified pixel is then assigned the class of the "most likely" (most highly statistically probable or similar) class. The Maximum Likelihood classifier basically takes each individual pixel within the satellite imagery, compares the DNs of that unclassified pixel to all of the training sets, and uses a probability function to assign the pixel to the pre-defined class that it is most like, or with the highest probability.

After the process is done, open a new Viewer, and open the Superl.img file. Using the Inquire Cursor in the Superl.img file, you can look at which training areas where used to label each pixel (You have to widen the Inquire Cursor window, as shown in the graphic).

#1 You should just look in general at your result and also the original satellite data. In Part II of this exercise we will really evaluate the result with field data. For now, just look at it visually. You can also use the aerial photos to look and compare some areas. How did your classification turn out? Do you have some confusion between classes? Which classes?

You are done with Part I of Exercise #7. Write down your answer to the question for Part II. You will need it then. You do not need to send any answers today.

Make sure that you save the following files to your own N: drive!!!