GST 101: Introduction to Geospatial Technology Lab Series

Lab 6: Understanding Remote Sensing and Aerial Photography

Document Version: 2013-07-30

Organization: Del Mar College
Author: Richard Smith

The development of this document is funded by the Department of Labor (DOL) Trade Adjustment Assistance Community College and Career Training (TAACCCT) Grant No. TC-22525-11-60-A-48; The National Information Security, Geospatial Technologies Consortium (NISGTC) is an entity of Collin College of Texas, Bellevue College of Washington, Bunker Hill Community College of Massachusetts, Del Mar College of Texas, Moraine Valley Community College of Illinois, Rio Salado College of Arizona, and Salt Lake Community College of Utah.

Contents

Introduction

This lab is part of a series of lab exercises designed through a grant initiative by the National Information, Security & Geospatial Technologies Consortium (NISGTC), funded by the United States Department of Labor in partnership with the Department of Education under the Trade Adjustment Assistance Community College and Career Training Grant Program (TAACCCT).

In this lab, students will learn how to use MultiSpec, a freeware multispectral image data analysis system created at the Purdue Research Foundation. MultiSpec provides the ability to analyze and classify imagery data, among other tasks. This lab covers four tutorial exercises supplied by the MultiSpec team, and provides an introduction to the software package.

Your instructor may require that you provide screen captures and/or exported files. Please check with your instructor for the requirements specific to your class.

This lab includes the following tasks:

1. Display and Inspection of Image Data
2. Image Enhancement
3. Unsupervised Classification (Cluster Analysis)
4. Supervised Classification

Objective: Apply the Basics of MultiSpec for Image Analysis

Image analysis is one of the largest uses of remote sensing imagery, especially with imagery that has recorded wavelengths beyond the visible spectrum. MultiSpec is a software package that allows you to conduct analysis on imagery beyond the visible spectrum and to analyze patterns in the images.

In this lab, you will complete four tutorials provided by the MultiSpec team. Exercises 1 and 2 teach you the basics of the software, and how to load and enhance imagery. Exercises 3 and 4 teach you how to analyze imagery so the computer can then interpret and classify the objects found on the image.

Lab Settings

Required Virtual Machines and Applications

Windows Machine User Account	Train
Windows Machine User Password	Train1ng$

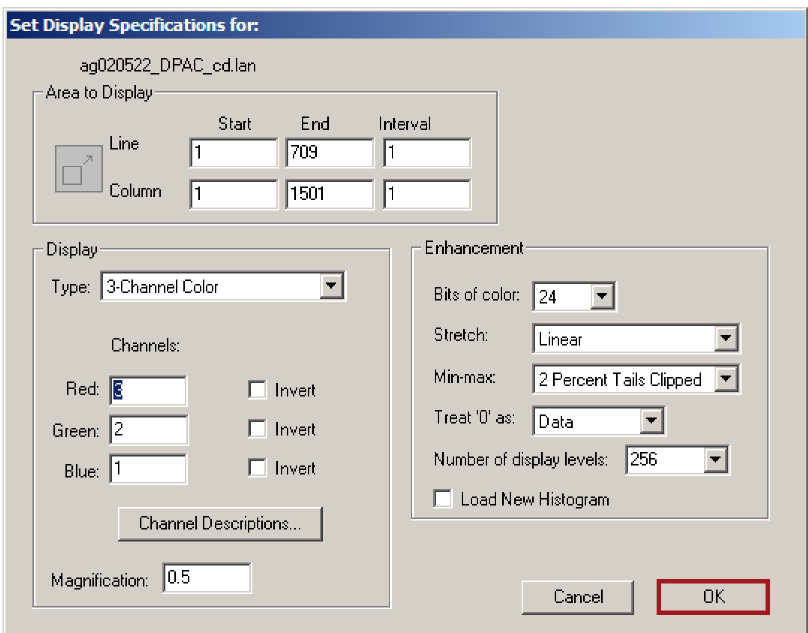
1 MultiSpec Exercise 1: Display and Inspection of Image Data

Requirements: MultiSpec application and image titled "ag020522_DPAC_cd.lan".

In this exercise, you will display an aircraft image of a portion of the Davis Purdue Agriculture Center in Randolph County, Indiana and view the data in several ways using MultiSpec.

1. Start MultiSpec using the icon on the desktop or from MultiSpec in the Startup Menu.
2. From the ***File*** menu choose ***Open Image. . .***. A dialog box will open to allow one to select the data file one wishes to use.
3. Select **ag020522_DPAC_cd.lan** in the Lab 6->Data folder and ***Open,*** or simply double-click on ag020522_DPAC_cd.lan

This is a segment (709 lines x 1501 columns of pixels) of a 3-channel image of DPAC collected on May 22, 2002. Next, a dialog box will appear to allow one to choose among various options for the image display.



Note that by default, the area designated for display is the whole scene and the ***3-Channel Color*** Display Type is selected. The default settings call for the Red screen color to be derived from band 3, Green from band 2 and Blue from band 1. These particular choices will cause the screen image to be in a 3-color format approximating color Infrared film.

4. Select **OK.**

5. This step may not occur for all situations. If the data histogram has not previously been calculated and stored (in a .sta file), another dialog box will be presented allowing the choice of regions to be histogrammed, so that the channel data values can be properly assigned to screen colors. The default options built into this dialog box are satisfactory.

   Select **OK** to begin the histogramming.

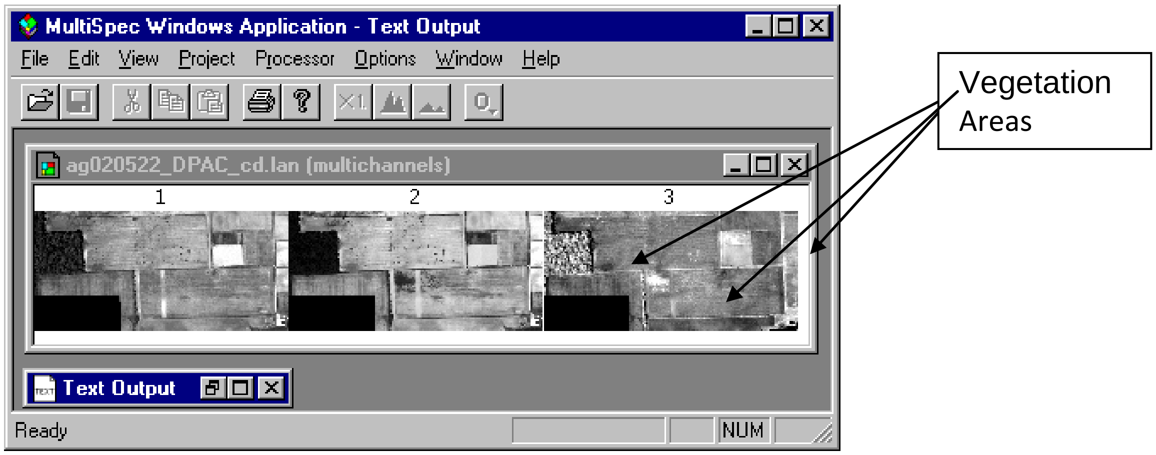
After the histograms of all of the channels have been computed, the information will be saved to a file named “ag020522_DPAC_cd.sta” so that they will not have to be re-computed when needed again.

Note that if a .sta file already exists with the default name, a dialog box will be presented allowing you to overwrite the existing .sta file or save to a different location.

The image of the data will now appear. Notice that just above the image window, in the toolbar there are two small boxes with large and small “mountains.” These are image zooming buttons allowing one to zoom in (large mountain) or out (small mountain) from the current image scale.

6. Experiment with the zoom functions.
   a. Just to the left of the image zooming buttons is another button, which shows ***X 1*** in grayed form. This button allows one to go to X1 magnification directly. The current zoom magnification is displayed along the bottom of the MultiSpec application window in the box labeled “Zoom=.”
   b. Some other options are to hold the *Ctrl* key down while zooming to change the zoom step factor to 0.1 instead of 1. In other words, the zoom factor will change from 1.0 to 1.1 to 1.2 etc. instead of 1, 2, 3, etc. (Note that one uses the ‘Option’ key on the Macintosh version to do this.)
   c. You can make a selection within the image by clicking in the image window and dragging to select a rectangle. If a selected area exists in the image, any zooming will be centered on the selected area if possible. Clear selection using the “Delete” key on your keyboard.

7. Next, view a side-by-side channel display for data quality inspection. From the ***Processor*** menu, select ***Display Image…*** to bring up the display dialog box. Then select **Display Type** “Side-by-Side Channels,” from the **Type** dropdown, and select **OK** to display all three channels in the image side by side.

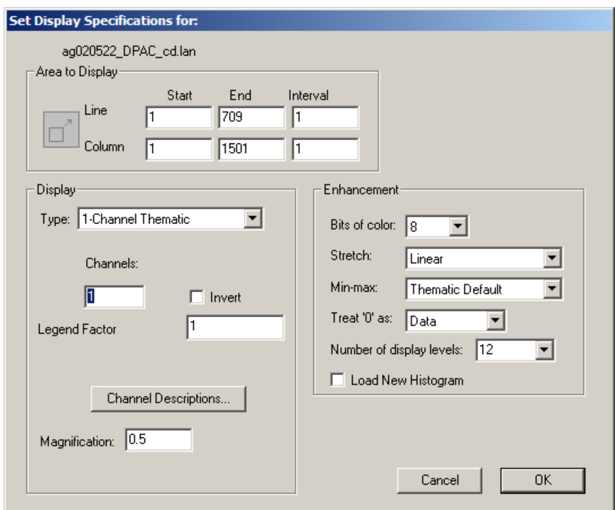


The above image window will be displayed (after zooming out) which shows all three channels displayed side-by-side. Note that the vegetation areas in channel 3 are brighter than the same areas in channels 1 and 2.

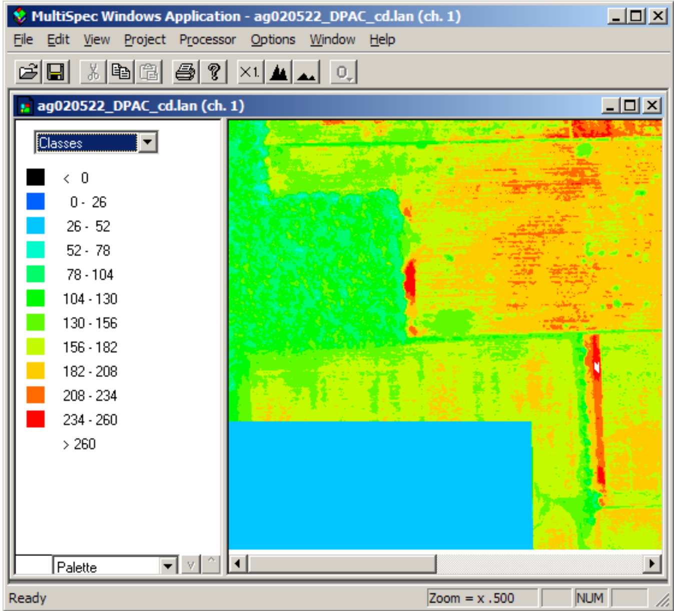
The side-by-side channel display is a good way to verify that the channels are registered correctly. In other words, the same location in the image is at the same pixel location in all channels.

8. To do this, **select an area** in one channel near a field intersection. This same selected area will be drawn in all of the channels. One can then verify that the selected area is at the same location in each channel.
9. Redisplay the 3-channel image with channels 3, 2 and 1 as Red, Green and Blue. **Coordinates View.** One can also display a “coordinates view” along the top of the image to present the cursor (mouse) location and selected areas in the image. To do this, select the **View->Coordinates View** menu item in the Windows version of MultiSpec [or select the **Window->Show Coordinate View** menu item in the Macintosh version]

10. If map coordinate information exists for the image, one can display the coordinates as map units. Use the popup menu on the left side of the coordinate view to select the map units. The area of the selection can be displayed as the number of pixels or in units of acres, hectares, etc. using the dropdown menu to the left of "Scale." The scale of the image will also be displayed.
11. We can also display the cursor coordinates as latitude-longitude. To do this, select **Edit–>Image Map Parameters…** and set the **UTM Zone** to **16** and the **Datum** to **WGS 84** and select **OK**. Note that the format that this image file is stored in (ERDAS *.lan) does not allow for storage of this information. Other image file formats such as GeoTIFF do.
12. Now, change the display again to the 1-channel Thematic display option (see the screenshot below). The 1-channel Thematic display type is useful to display "product" type images such as MODIS NDVI or any of the other many MODIS products.



The data values are grouped into the desired number of levels and a legend is displayed to the left of the image indicating which palette colors are associated with each range of data (see the screenshot below). One can also enter a factor to multiply the data values displayed in the legend to reflect the actual measurement unit. Sometimes the data value may be the measurement value times 100 or 1000. One can use the Min/Max User Specified dialog box item to set the min and max values for the range of data to be displayed. Black is the default color for data values less than the minimum and white is the default color for values greater than the maximum. (Note: This feature can be considered as a supervised 1-channel levels classifier.)



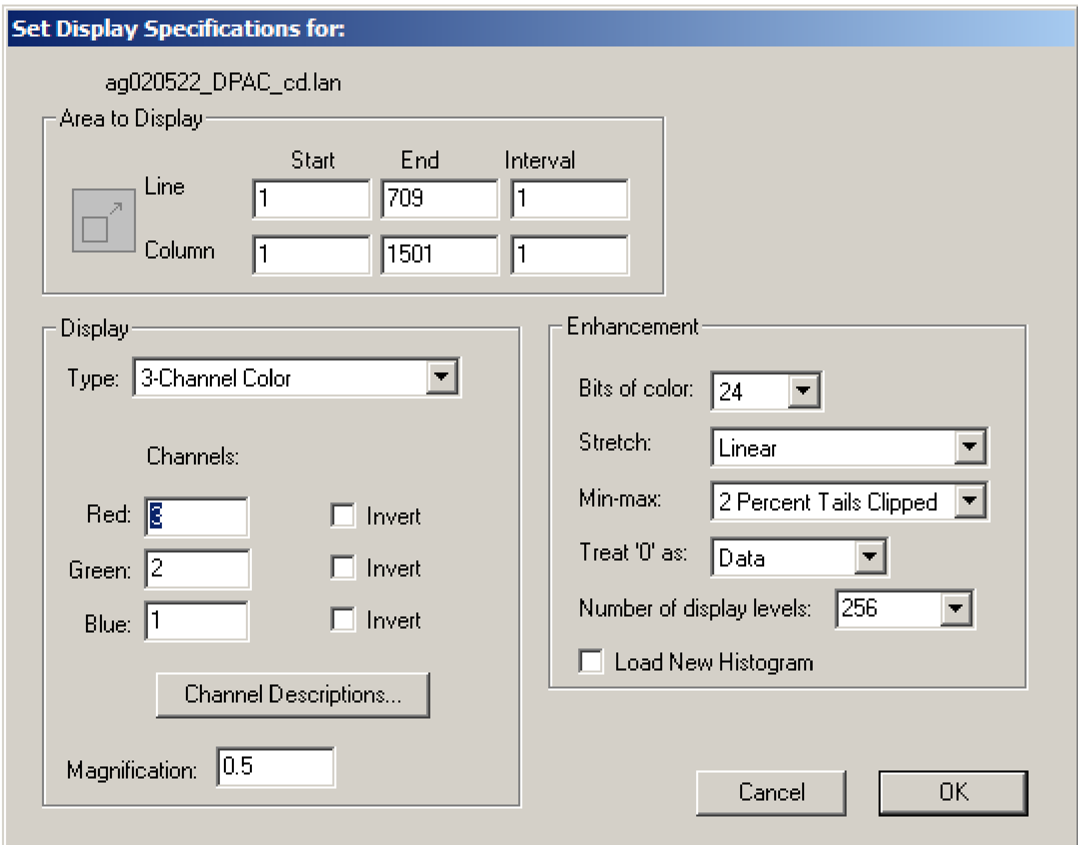
You can use the Reformat - Change Image File Format processor to create thematic images based on what is displayed in the image window.

2 MultiSpec Exercise 2: Image Enhancement

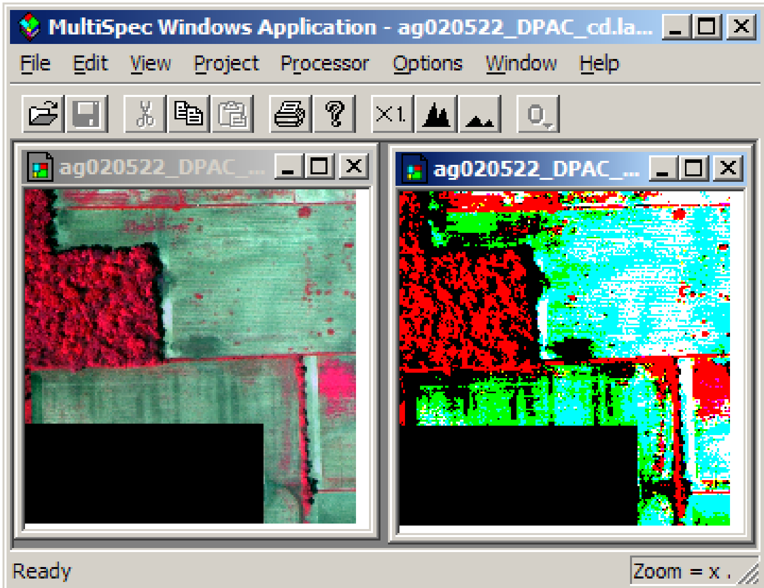
Requirements: MultiSpec application and image titled "ag020522_DPAC_cd.lan".

Open the image if it is not already displayed in a multispectral image window following the guidelines given in Exercise 1 (Display and Inspection of Image Data with MultiSpec).

One can control the enhancement of the image in the multispectral image window by setting five different options in the **Enhancement** section of the **Display Specifications** dialog box including Bits of color, Stretch, Min-maxes, Treat '0' as, and Display levels per channel.

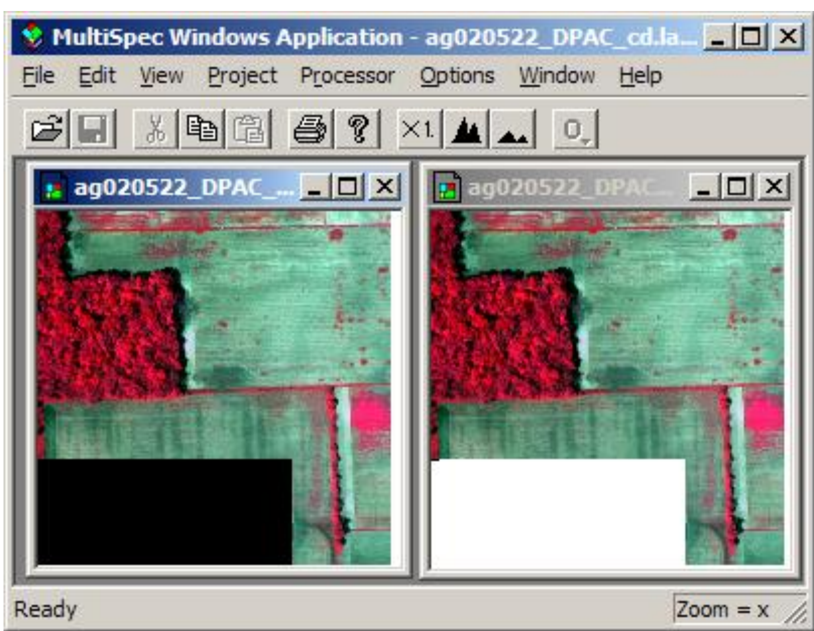


1. The **Bits of color** default is 24 and the **Display levels per channel** default is 256 for all monitors as the current maximum number of colors possible. In the Display Specifications window, select the 8 bits of color option and notice the change in your image.



The figure above illustrates 256 display levels per channel (millions of colors) on the left and 2 display levels per channel (8 colors) on the right.

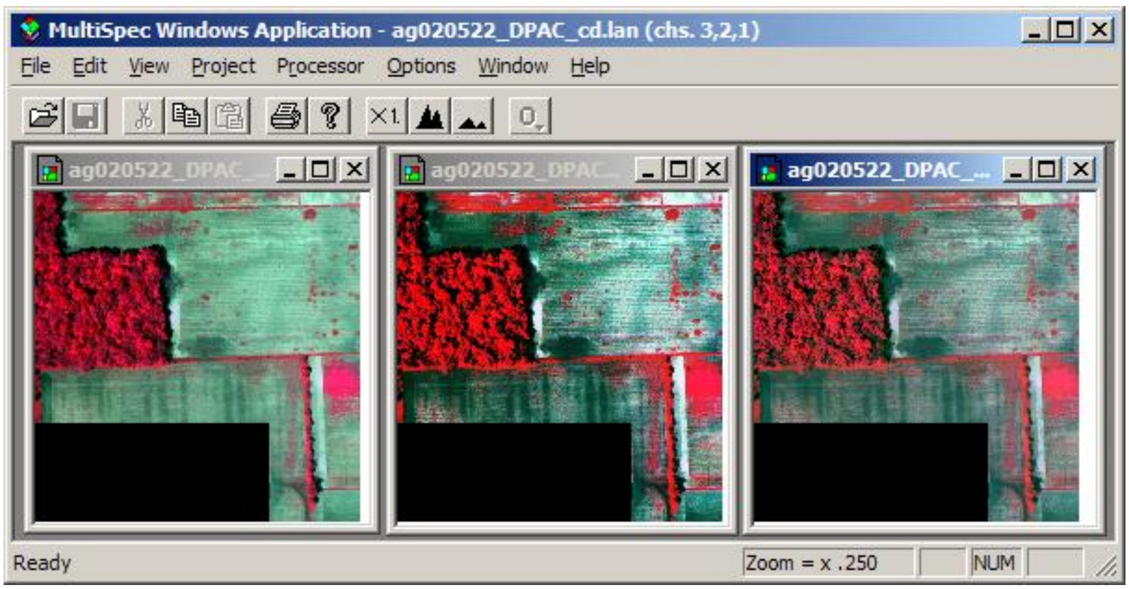
2. The **Treat '0' as** data setting causes 0 values to be displayed as black. However, if 0 actually represents background or 'no data,' one may want to select the background option to cause 0's in all channels to be displayed as white. Switch the **Treat '0' as** option to **Black** and view the image. Then switch it to **White** and note the difference. The figure below illustrates 0's treated as black on the left and 0's as white on the right.



The Stretch and Min-maxes are usually the options used to enhance different parts of the image. They control the process by which each possible data value in the image data is assigned to all possible display levels.

There are three choices for Stretch: Linear, Equal Area, and Gaussian. In the case of Linear Stretch, gray scale intervals are equally spaced over the data range, while the Equal Area Stretch choice causes them to be set so that each interval represents about the same number of pixels. Though nonlinear, the equal area stretch will provide maximum contrast. The Gaussian selection causes the distribution of the number of pixels assigned to the gray scale intervals to represent a Gaussian curve.

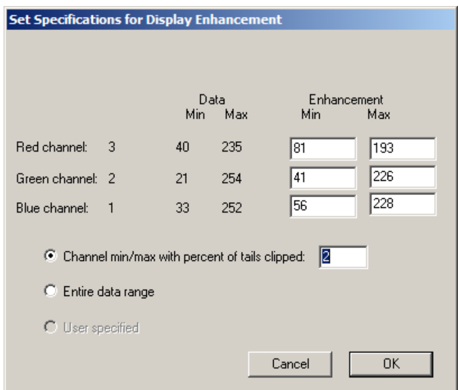
3. Change the Stretch (in the Display Specifications window) to **Equal Area** and **Gaussian** to see how the image changes. Additionally, if one holds the 'Option' key down (Mac version) or the 'Right' Mouse button down (Windows version) before you select the enhancement popup menu with the left button, you can change the number of standard deviations that the data will be fit to for the Gaussian selection. The default is 2.3.



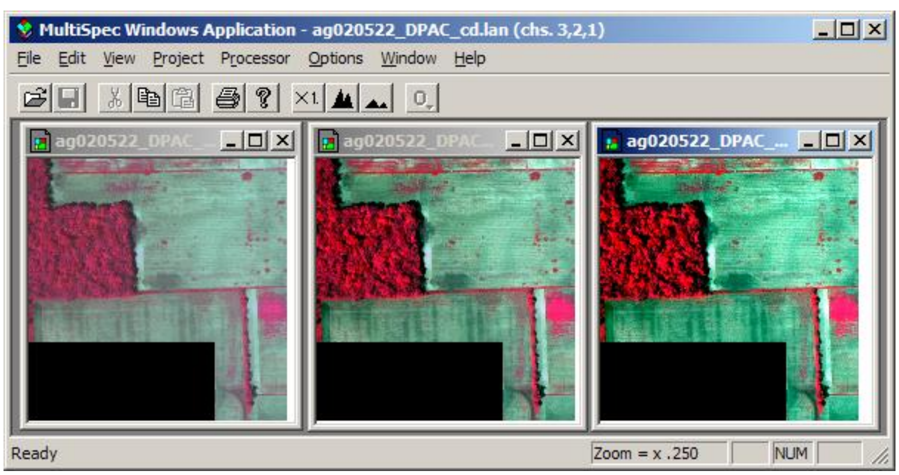
Linear, equal area and Gaussian stretches are illustrated in figure above left to right.

4. The Min-maxes option (also in the Display Specifications window) allows one to select the beginning and end data values of the image histogram to be used for assigning the pixels in the gray scale intervals defined by the Stretch option. The "Entire Range" choice for this option would cause lowest data value possible in the image, 0 for 8-bit unsigned data, to be the first data value displayed for lowest display value (black) and 255 to be the last data value displayed for the highest display value (white). However, if the actual range of the data is only 50 to 150, then the data will only be represented by grays not black to white; there will not be much contrast in the image.

5. The Clip 2% of Tails choice will cause the selected begin and end range of data values for a given channel to represent those data values in which 2 percent of them in the histogram are outside of the selected range. The intent of this choice is to reduce the chance that a small number of extreme outlier data values in the image will unduly influence the display enhancement. This choice usually results in a display of the data that has better contrast over all.
6. Selection of the **User Specified…**choice presents a dialog box (illustrated below) to allow one to set the percent of tails clipped to be something other than 2%. You can also set your own min-max values to stretch the gray levels across. The actual data min and max values computed from the histogram are included in the dialog box. Choose two different percentages, clip the tails using those values, and see how the image is altered. Additionally, choose Entire Range for the Max-min and see how that changes the image.



The entire range, 2% clip and 4% clip is illustrated below from left to right.



3 MultiSpec Exercise 3: Unsupervised Classification (Cluster Analysis)

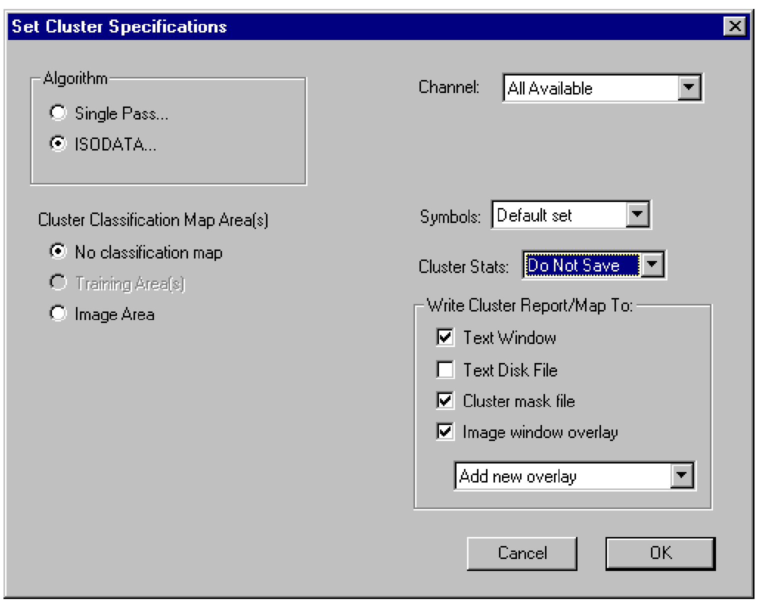
Requirements: MultiSpec application and image titled "ag020522_DPAC_cd.lan".

Two Clustering algorithms are available in MultiSpec. They are useful in grouping similar pixels in the image into clusters or categories. One algorithm implemented is a simple one-pass type. The second is an iterative type called ISODATA. We will use the ISODATA algorithm for this exercise.

To start this exercise, be sure that the "ag020522_DPAC_cd.lan" image that was used in exercises 1 and 2 is open. Also clear any selections in the image window by striking the "Delete Key."

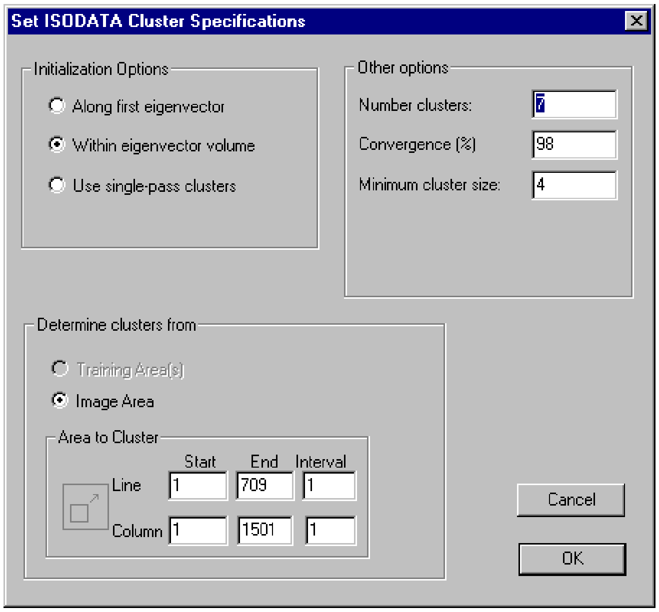
A cluster analysis will be run using the image file represented by the active (top-most) multispectral image window.

1. From the ***Processor*** menu, select ***Cluster…*** to bring up the Cluster Specifications dialog box. Select "**Do Not Save**" in the Cluster Stats: dropdown menu. Select "**Cluster mask file**" and select "**Image window overlay**" under the "Write Cluster Report/Map To" group.



These will give us a cluster map as a thematic image disk file and displayed as an overlay on the multispectral image window.

2. Next, select "ISODATA…" This will cause the ISODATA Specifications dialog box to be displayed as shown below. Enter 7 for the number of clusters, 100 for the convergence percentage and set the line and column intervals to 1 (if needed) and then select **OK**.



3. You are now back to the Cluster Specifications dialog box. Select **OK** to close this dialog box and start the clustering operation.
4. You will be prompted to enter a name for the cluster map disk file and where to save the file. Just use the defaults by selecting **OK** in the Save Cluster Map dialog box.

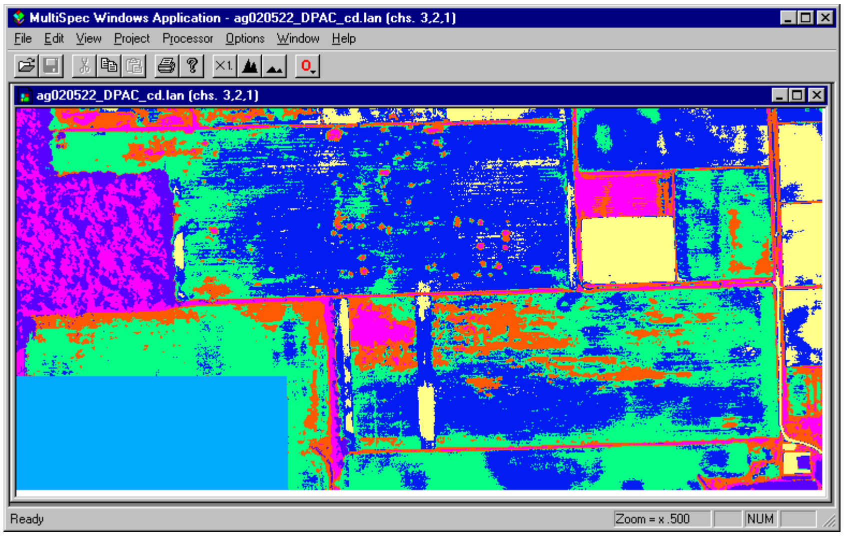
A cluster map will now be created with around 7 classes in an unsupervised manner. You will notice the colors change in the image window as the pixels are sorted into cluster classes during each iteration. After the final iteration, a thematic image file with a map of the cluster classes will be saved to disk.

The text output for the cluster operation will be at the end of text output window. The information includes the mean values for each of the channels for each cluster for both the initial condition and the ending condition. If the map information is available for the image, the final area for each cluster is listed in the units specified in the coordinate view for the image window.

Usually the convergence is set for a little less than 100 so that the process does not take too long to complete. We used 100 in this example so that you have a chance to watch the pixels change cluster classes which illustrates the nature of the ISODATA algorithm.

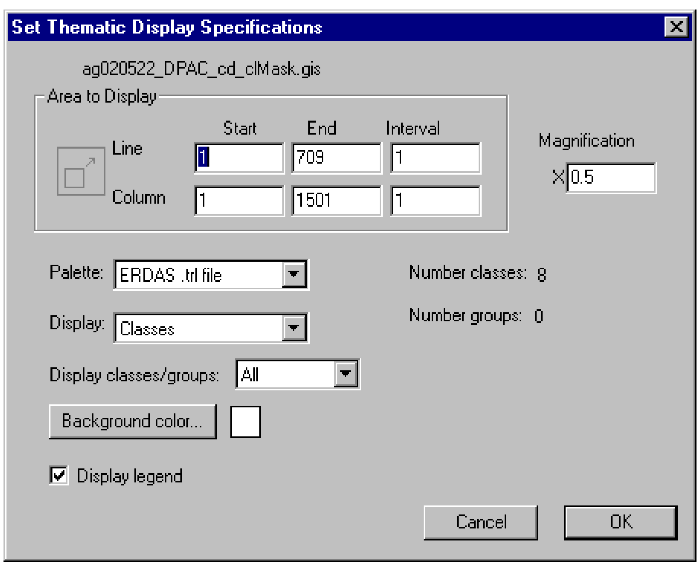
The cluster map overlay on the multispectral image window will look similar to the following screenshot.

5. You can turn the overlay on and off by using the “Red O” popup menu button in the toolbar to the right of the “small mountain” zoom button.



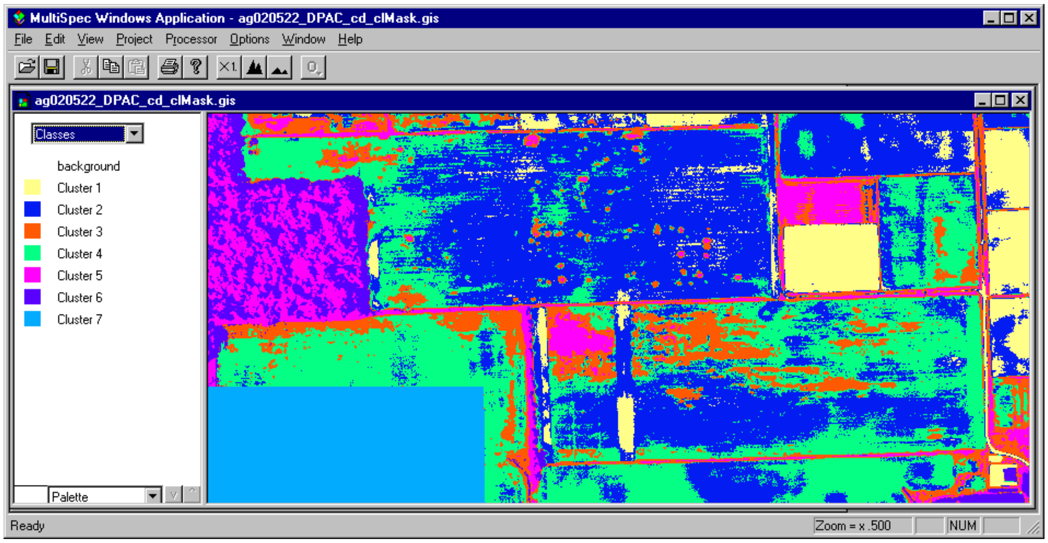
6. Now, we will open the cluster map thematic image file. This will be the same image as is shown in the overlay above but you will have more control over the cluster classes.
7. From the ***File*** menu, select ***Open Image…*** to bring up the open image dialog box. You may have to change the “Files of type” popup to All Files or Thematic Files. Then select “ag020522_DPAC_cd_clMask.gis” and then select **OK**.

The Thematic Display Specifications dialog box shown below will be displayed.



8. The default settings are fine; select **OK** in the Display Thematic Image dialog box. This opens a thematic image window.

9. Observe that the cluster class legend is on the left in the thematic image window below.



10. Compare the Cluster map with the original image window. Note clusters that represent trees, sparse vegetation, thicker vegetation, different soil colors and shadows.
11. There are several things one can do to evaluate the image. Move the cursor over a color chip, hold the shift key down (cursor will change to an open eye) and click the left mouse button down and up to cause the colors for that class to blink off and on (alternate between white and the color). If one holds down both the shift and ctrl keys and then click the left mouse button down and up, then all of the other classes will blink off and on. These procedures are helpful in understanding the extent of the classes in the image.
12. You can also change the class color by double clicking on the color chip.
13. Change the cluster class names by double-clicking on the name to the right of the color chip to match what the clusters represent. (The list of interpreted cluster names that I come up with are: Light soil, Medium colored soil, Sparse vegetation (weeds), Dark soil, Dense vegetation (trees, wheat, weeds), Shadowed trees, Image blank.
14. Next, group the cluster classes together in information groups by selecting Classes/Groups in the dropdown menu above the legend. Then drag the cluster classes into similar information group categories. Again, one can double-click on the group name to change the name. For example, change the appearance of the thematic image to represent Bare Soil, Vegetation and Non-image informational classes. The dropdown in the legend also allows one to display the original cluster (spectral) classes.

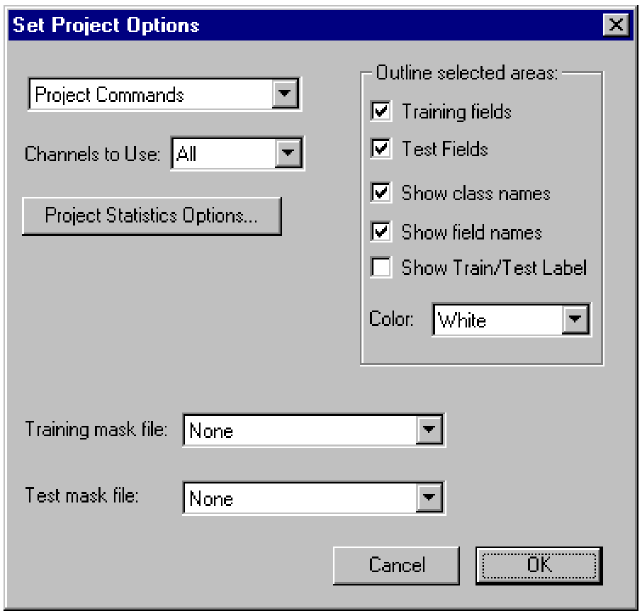
4 MultiSpec Exercise 4: Supervised Classification

Requirements: MultiSpec application and image titled "ag020522_DPAC_cd.lan".

One can also do a supervised classification by selecting training areas for specified classes from known areas. Open the image if it is not already displayed in a multispectral image window following the guidelines given in Exercise 1.

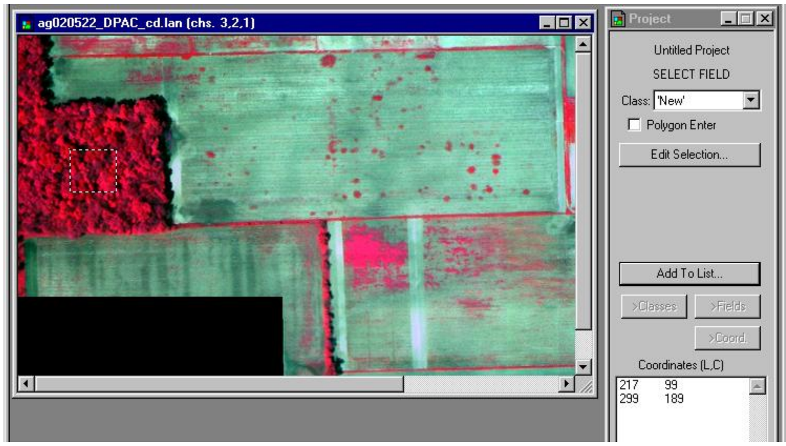
4.1 Select Training Fields

1. If a project window is open, click on the Window tab in the main menu. **If** a file entitled Project is listed, double-click in the upper left to close the current project and project window. One may have been created during the cluster analysis.



2. Now, we will select the training fields. From the **Processor** tab in the main menu select **Statistics** and then **OK** in the **Set Project Options** dialog box (The default settings for this exercise are satisfactory).
   A new window labeled Project will appear on the screen that will be used in a moment (if the window covers the entire screen, click the Restore Down button).

3. Now, let's select some training fields. To select training fields for each class, simply "drag" a rectangular area on the image (or, with polygon option selected, click on the corners of the desired polygon), as is shown in the screenshot below.

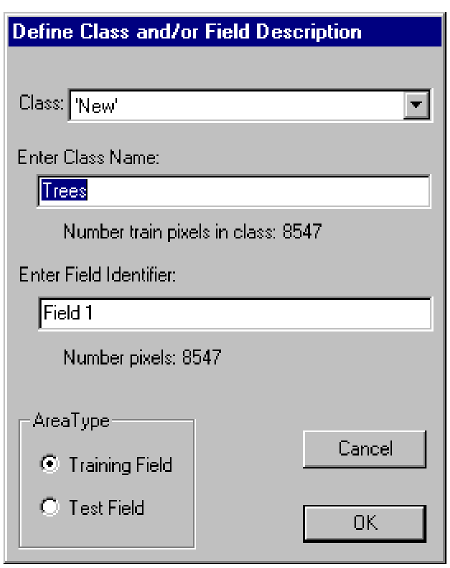


The screenshot shows our selection of the Tree training field in the image window.

Caution: A mistake that is made many times is to select training areas very near edges of a field. One should stay away from the edges by a couple of pixels to reduce the chance of edge affects.

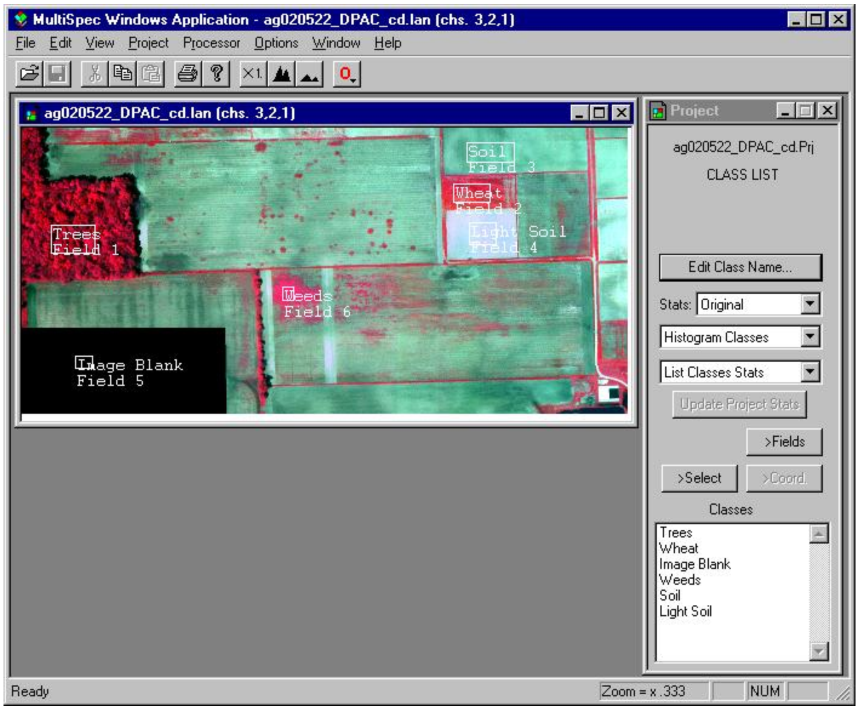
In the ***Project dialog box***, the coordinates (row and column numbers) of the upper left corner and the lower right corner of the selected area will appear.

4. Select the ***Add To List*** button. A dialog box will appear like the one shown below. Type ***Trees*** into the Class Name box and then select ***OK***.



Note that you can designate the selected area as a training or test area.

5. Since there is to be only one training field for this class, we are ready to select the training for the second training class. Drag across the second training field in the ***Image*** window shown below for **Wheat**.
    a. Select the ***Add To List*** button in the ***Project*** window.
    b. Select the training areas for the rest of the six classes as shown below – Image Blank, Weeds, Soil, and Light soil.



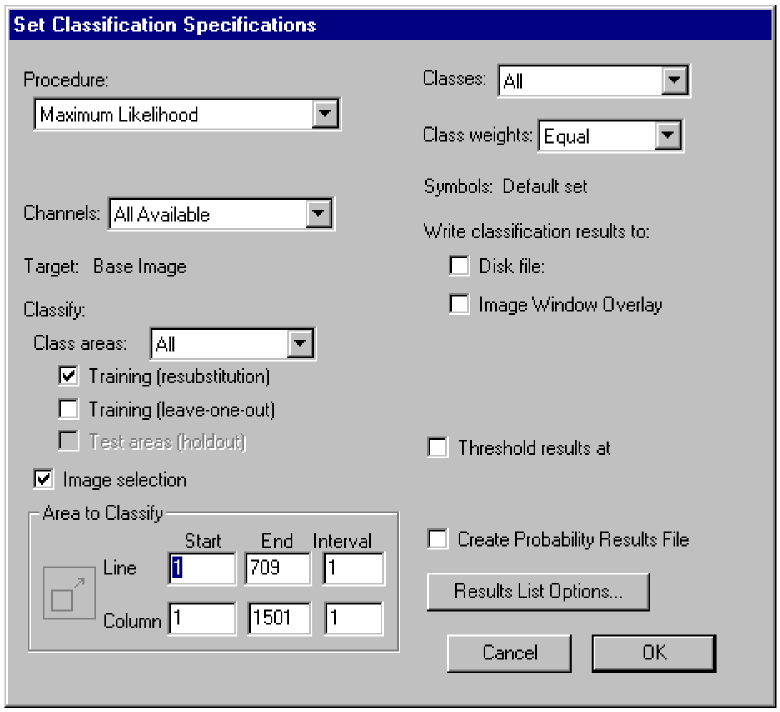
The Project Window can have four different modes - the Select training field mode (>Select button), the Class list mode (>Classes button), the Field list mode (>Fields button), and the Coordinate list mode (>Coord. button).

6. If you wish to delete a class, select the class in the class list and then select **Cut Class** in the Edit tab on the main menu. You can do the same for deleting a specific field.
7. One can also use polygonal type fields to define training classes. To do this, select the **Polygon Enter** checkbox in the Project Window when in Select mode. Click in the image window to define each corner of the polygon. Double click on the last point. To turn the polygon type selection off, just select the **Polygon Enter** checkbox to deselect it. This is especially useful to select odd-shaped classes.

Note that the clustering step described in Exercise 3 can be used to determine how many classes might be separable in a given data set and where to define training areas such that the spectral characteristics of the pixels are similar.

4.2 Classification

1. From the **Processor** tab in the main menu, select ***Classify…***. In the Set Classification Specifications dialog box which appears, select the √ next to **Image Selection** to de-select it since, during this pass, it is desired to classify only the training fields in order to obtain an initial estimate of the quality of the class definition and training.



Note that under ***Write classification results to:*** One can also select the ***Disk File*** button causing a disk file version of the results to be written. Since we have no need to keep this file in this case, leave this button unselected.

Since the other default options are satisfactory, select **OK** and then **OK** to the Update Project Statistics? dialog box to begin the classification.

The classification will be complete momentarily.

2. From the **Window** tab in the main menu select **Text Output**, to bring the text window forward and make it active, since it contains the classification results. The "TRAINING CLASS PERFORMANCE (Resubstitution Method)" table tabulates how the pixels of each field and class were classified. See example table below. There should be nearly 100% accuracy on the training fields. If the Reference Accuracy is particularly low (say less than 50%) for a class, then the training pixels for that class should be reexamined and new training pixels selected.

Classification of Training Fields

TRAINING CLASS PERFORMANCE (Resubstitution Method)

Project Class Name	Reference Class Number	Accuracy+ (%)	Number Samples	1 Trees	2 Wheat	3 Image Blank	4 Weeds	5 Soil	6 Light Soil
				Number of Samples in Class					
Trees	1	98.4	8249	8117	126	0	6	0	0
Wheat	2	97.5	4277	102	4170	0	5	0	0
Image Blank	3	100.0	1485	0	0	1485	0	0	0
Weeds	4	97.9	1102	2	21	0	1079	0	0
Soil	5	100.0	6069	0	0	0	0	6069	0
Light Soil	6	100.0	3933	0	1	0	0	0	3932
TOTAL			25115	8221	4318	1485	1090	6069	3932
Reliability Accuracy (%)*				98.7	96.6	100.0	99.0	100.0	100.0

OVERALL CLASS PERFORMANCE (24852 / 25115) = 99.0%
Kappa Statistic (X100) = 98.6%. Kappa Variance = 0.000001.

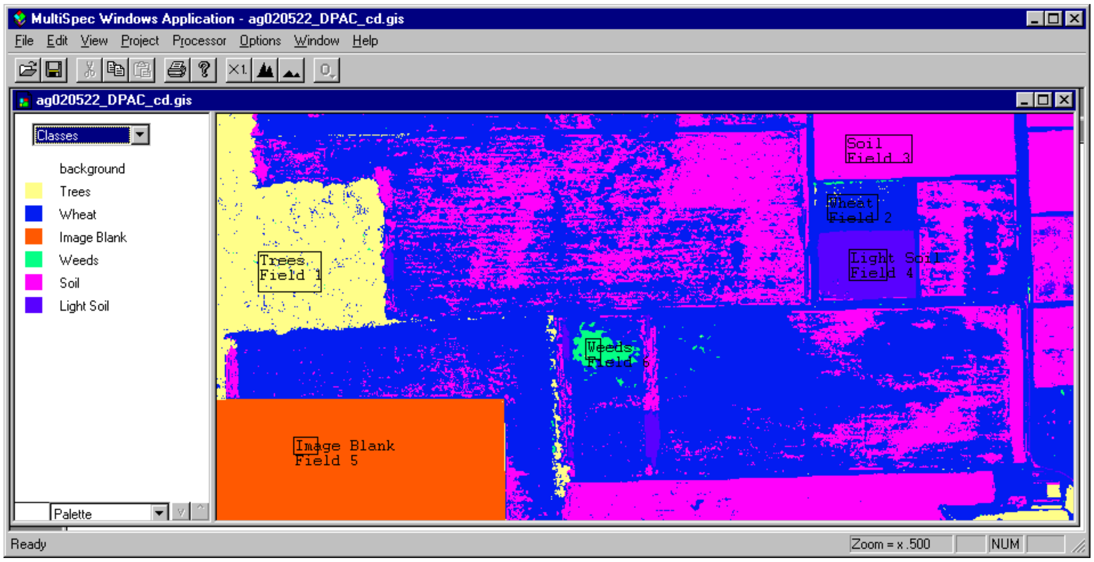
+ (100 - percent omission error); also called producer's accuracy.
* (100 - percent commission error); also called user's accuracy.

3. Assuming satisfactory results, we are ready to classify the whole area. From the ***Processor*** tab in the main menu choose ***Classify***…
    a. Under **Areas to Classify** de-select **Training (resubstitution)**.
    b. Then, select **Image selection**. Make sure that the entire area of the image is to be classified (lines 1-709 and columns 1-1501). Select the square button, if activated, to the left of lines and columns, to force all lines and columns in the image to be used.
    c. Also select **Disk File** under **Write classification results to:** so that a disk file for later use will be created.
    d. One can also select **Image Window Overlay** to cause the classification to be displayed as an overlay on the multispectral image window if you wish to.
    e. Also select the **Create Probability Results File** checkbox so that a classification probability map will be saved to a disk file.
    f. Then select **OK**.
    g. Select **Save** twice in the dialog boxes that follow regarding a file name for the results. We will use the default names and locations for the output classification file and for the probability map file.
    h. As soon as the classification is complete, one will see a summary of the results displayed in the text window.

4. You can save the project using the **File–>Save Project** menu item. You will be presented with a dialog box to enter the name (or use the default name). The training and test areas that you selected will be saved. You can open this file up at a later time if you wish to continue your analyses.

4.3 View Classification Map

1. Now, open the classified image named **ag020522_DPAC_cd.gis**. It should appear similar to the image below (it may have different colors). This is a Thematic type image.

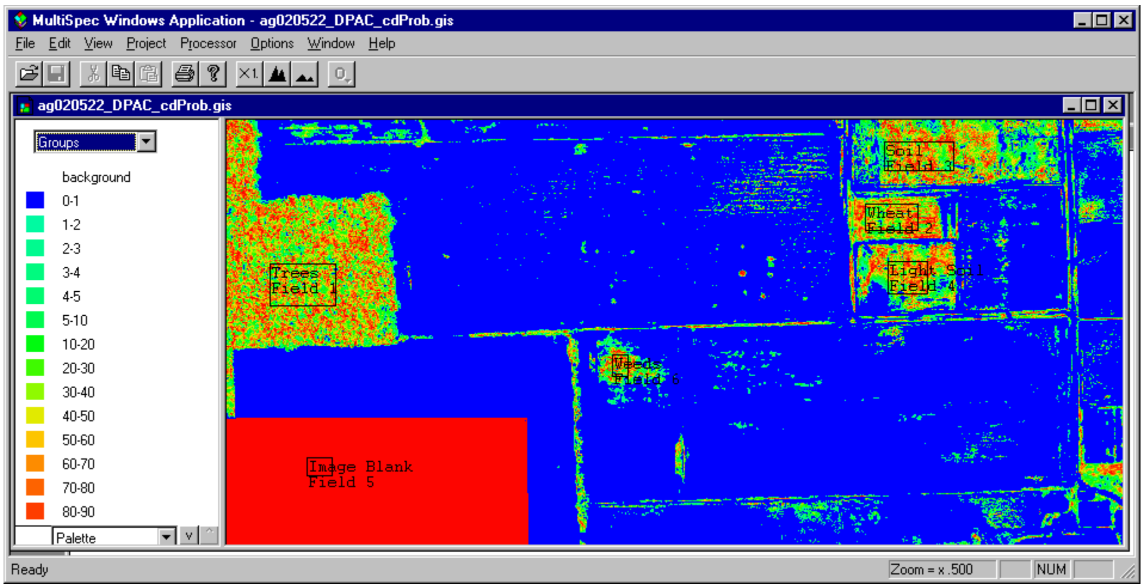


2. After displaying the classified image “ag020522_DPAC_cd.gis”, from the ***Project*** tab in the main menu, select ***Add as Associated Image*** to cause the training field outlines to be drawn on the image. Change the field outline color to black using the **Processor->Statistics…** menu item and selecting **Black** in the **Color** dropdown.

4.4 Classification Probability Map.

Next, we will view the classification probability map to evaluate which portions of the image have lower and higher probabilities of being classified correctly. There may be other classes in the image that our training fields do not adequately represent.

1. Open the probability map named ***ag020522_DPAC_cdProb.gis***. It should appear similar to the image below after associating it to the Project. This is a Thematic type image.



Yellow to red colors represent a high probability of being correct. These pixels are very close to our training pixels for the classified class. Dark green to blues represent a low probability of being correct. These pixels are very far from the training pixels for all of the classes. The classified image though may still represent the area well enough for our purposes. For our classification, we could probably do a better job of separating out the sparse weeds from the wheat. There are probably more than the six classes that we selected than can be separated successfully.

Conclusion

In this lab, you have learned the basics of MultiSpec, a freeware multispectral image data analysis system. MultiSpec is an excellent example of many free-to-use programs available for imagery analysis and GIS-related tasks.

There are many other operations that one can do with MultiSpec including several Reformatting processes. See other exercises or the MultiSpec Introduction at the MultiSpec web site for more information. The MultiSpec web site is: http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/. Or contact Larry Biehl at biehl@purdue.edu with questions.

Discussion Questions

1. With what other objects in imagery could you use supervised classification? Describe two use-cases.

2. Even though the imagery used in these tutorials contained recordings of wavelengths beyond the visible spectrum, why do you think each range of hyperspectral information was mapped to the colors red, green, and blue?

3. Photointerpretation is the process of converting images to information through human interpretation. Based on the results from exercises 3 and 4 in this lab, do you think that computers can interpret as well as humans?

