Andrew Nordquist - @01078209

Digital Image Processing for Bolivian Anthropological Study

6 May, 2008
Project Final Report for EES 5083 – Digital Image Processing Class

Introduction. A professor of Anthropology, Dr. Sonia Alconini, wants to investigate possible ancient village sites in an area of Bolivia that is adjacent to the border with Peru and northeast of the shores of Lake Titicaca. There are no road maps nor other ortho-rectified maps of the study area; this has been borne out by many months of investigation by Dr. Alconini. The area is about 1280 square kilometers in size, and consists of mainly steep mountainous terrain. Although most of the area is a national park, there are some existing villages and towns.

She wants to be able to find prospective sites at which to begin her investigations. Such sites could be evident by the remains of terraces built into the mountainside by the tenants of the ancient villages. The terraces were used for farming in ancient times, and some may still be in use today. A more well-known example is the terracing constructed in the steep terrain by ancient Incas near the Machu Picchu site in Peru.

She will begin field work in June 2008. She already has prepared some digital information that is stored and accessed using ArcGIS.

Acquisition of Imagery. At the beginning of my participation in her investigation, I gathered some sample 1-m and 5-m resolution digital satellite images from some Internet sites, such as Satellite Imaging Corporation (www.satimagingcorp.com). I showed her the difference between 1-m, 5-m, and coarser resolutions when it comes to trying to view objects and features on the ground.

The 1-m images provided sufficient justification to increase her budget for acquiring images. She procured 0.6-m resolution panchromatic (PAN) QuickBird imagery available for about 75% of her study area. She was already in possession of a 5-m PAN Indian Remote Sensing satellite (IRS) scene that contains about 95% of her study area. There is a small gap between the two QuickBird scenes that is less than 500 m wide; otherwise, QuickBird imagery covers the western portion of the study area. The combination of the one IRS PAN and two QuickBird PAN and MUL images provides 100% coverage of the study area.

The QuickBird Imagery was acquired (indirectly) from Digital Globe. Digital Globe would provide only pan-sharpened imagery for the same price that our supplier charged for providing us the “bundle” of separate MUL and PAN imagery. The specification called for a georeferenced, “ortho-ready” format, UTM zone 19 South projection, WGS-84 Datum. Tile files (.TIL) were provided so the imagery could be read into ENVI as an External file format (Mosaic Tiled QuickBird product). Additional files were provided, under the category of “GIS Files”. This folder contains SHP files designated as “Order Shape”, “Product Shape”, “Strip Shape”, and “Tile Shape”.

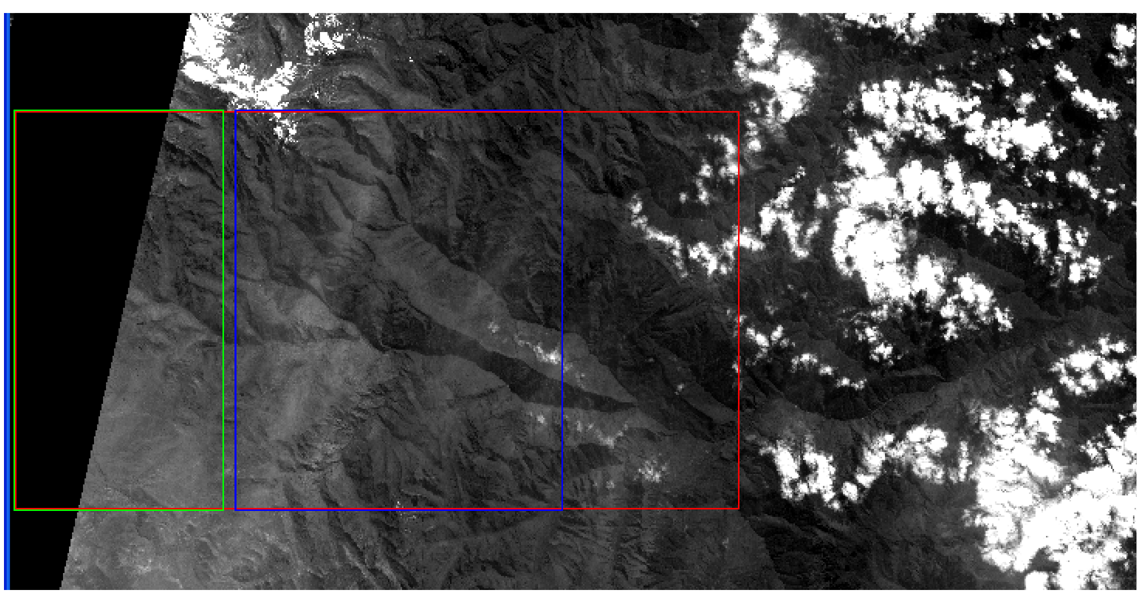
<u>Registration Problem(s).</u> A problem has been identified. The feature registration between the QuickBird (QB) images and the IRS image is not good. All the images are georeferenced, but there is some error between the UTM projection data for each image. The coverage of both QB images was converted to EVF vector files and superposed on the IRS coverage of the study area. Comparing the two images and assuming that the QB images are "true", the registration of the IRS image is off by over 2 km from the QB images.

The difference in meters was estimated from a feature identifiable in both images to the edge of the EVF rectangle using the Cursor Location and Measurement Tool windows in ENVI. By this method, the EVF rectangle overlay for the QB East image is about 2425 m west of where it should be on the IRS image. The EVF rectangle overlay for the QB West image is about 2354 m west of where it should be. Does this mean that the registration is bad for QB and not IRS?

A suggestion has been made to do an image-to-map registration, with the QB images as the base and the IRS image as a warp. Will this solve the problem? Not easily, but perhaps it will. Ground Control Points (GCPs) are still being constructed for the IRS image, and registration will happen soon.

How about just changing the georeferencing for the offending image(s)? I discovered that I could edit the IRS header information. I could change the XY coordinates of the reference point and effectively re-register the image. Is this a valid procedure? Is it possible to use GCPs to find the offset of the reference point, change the reference point and be done with re-registration?

Will ortho-rectifying all three images separately solve the problem? There are rectified images available in Dr. Alconini's ArcGIS database. Would it be possible to perform a similar image-to-map registration, with the map as a base, to correct the problem?



On the previous page is a picture of the three EVFs overlaid on a portion of the IRS image. The EVF for the Study Area was imported from an ArcGIS SHP file. You can see there is good coverage of the Study Area, if all three images are used. Even if the three rectangles need to be moved 2 km east to be correct, there is plenty of data available. However, the further east one goes, the more clouds enter the image. Fortunately, the far eastern portion of the study area is less important, because it is mostly jungle.

It appears there is another registration problem. The rectified images contained in Dr. Alconini's stored ArcGIS data are all consistent with each other, but they appear to be different from the QB images. It seems the ArcGIS features are about 900-1000 m north of the same features in the QB imagery.

I realize that registration is a common problem when one is dealing with information from multiple sources. How do I find out which one is "correct"? Will it take GCPs generated from field work to truly rectify all this data?

Other, minor problems. There is some streaking in the IRS image, possibly some very thin, high clouds over the study area. It could also be sunlight reflecting off wet terrain. All I have of the IRS data is the PAN imagery. Can this whiteness be eliminated, or is it really a problem?

Digital Elevation Model (DEM). I have downloaded the 90-m SRTM DEM image for the Study Area. This data is poor, as it has many "holes" where there were clouds at the time of image acquisition. This data is not acceptable in the current form to use as a DEM. I have written an IDL routine to replace the "bad" data with a constant number.

It turns out that Sonia has had a student in Bolivia "fill in" the spotty SRTM data and create an ArcGIS TIN file of the results. I was able to convert that TIN data into raster data, using the 3D Analyst extension in ArcGIS. I finally used another ArcGIS feature to export the raster data to a TIF file. I chose to create several candidate DEM files with varying horizontal resolution: 200m, 100m, 50m, and 10m. The 10-m resolution data looks very good, and we can use that TIF file as the DEM for the entire study area. Figures 2 and 3 on the next page show the original SRTM data and the results derived from the TIN file.

Figure 2. SRTM DEM data from NASA. EVFs are superposed. Black regions have no DEM data. I considered this data unacceptable in this form.

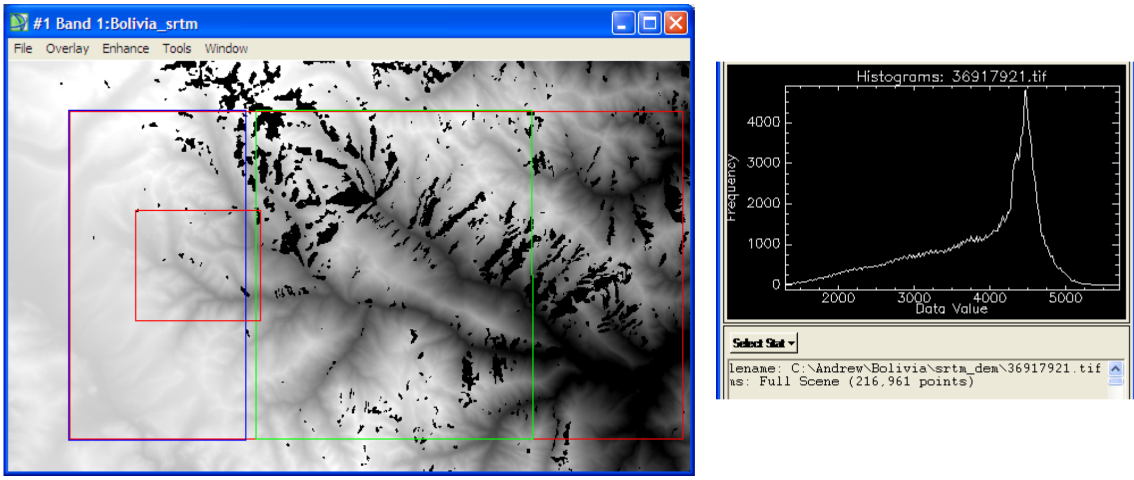
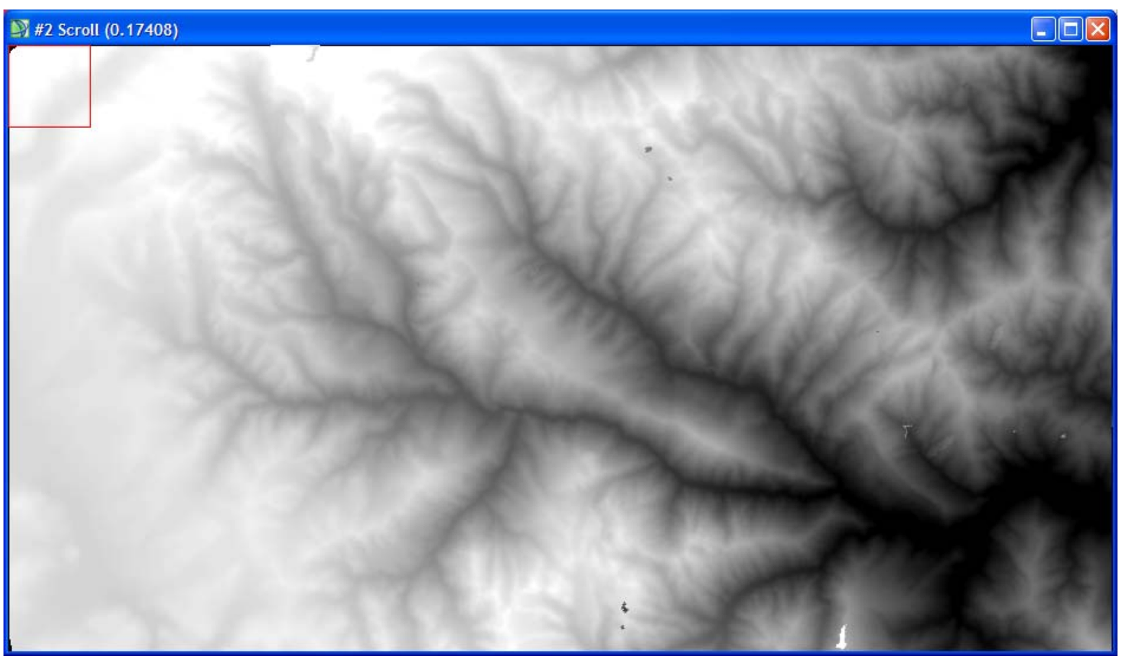


Figure 3. DEM data from the "MergedFernandez" TIN file in the ArcGIS database. There's only a couple of small glitches that can be eliminated with a little ENVI work.



Pan-sharpening. The two QB images are composed of a total of 15 tiles, as supplied. I subjected each of these tiles to the Gram-Schmidt Spectral Sharpening technique. For the most part, the results of the pan sharpening are spectacular. However, the technique produces files that are about 2 GB in size for each tile.

I quickly ran out of space on my hard drive for storage, and procured a 160 GB portable hard drive on which to store the project data. So far, about 60 GB of the 160 GB is used for this project.

I encountered a problem with the results of this technique in four of the fifteen images. The Statistics reported some DN values were ‘way out of range. I repeated the GS sharpening procedure, and ended up with two pan-sharpened versions of the affected tiles. Sometimes I had to combine different bands from each of the different attempts, but I managed to get an entire spectrum of sharpened image for each tile.

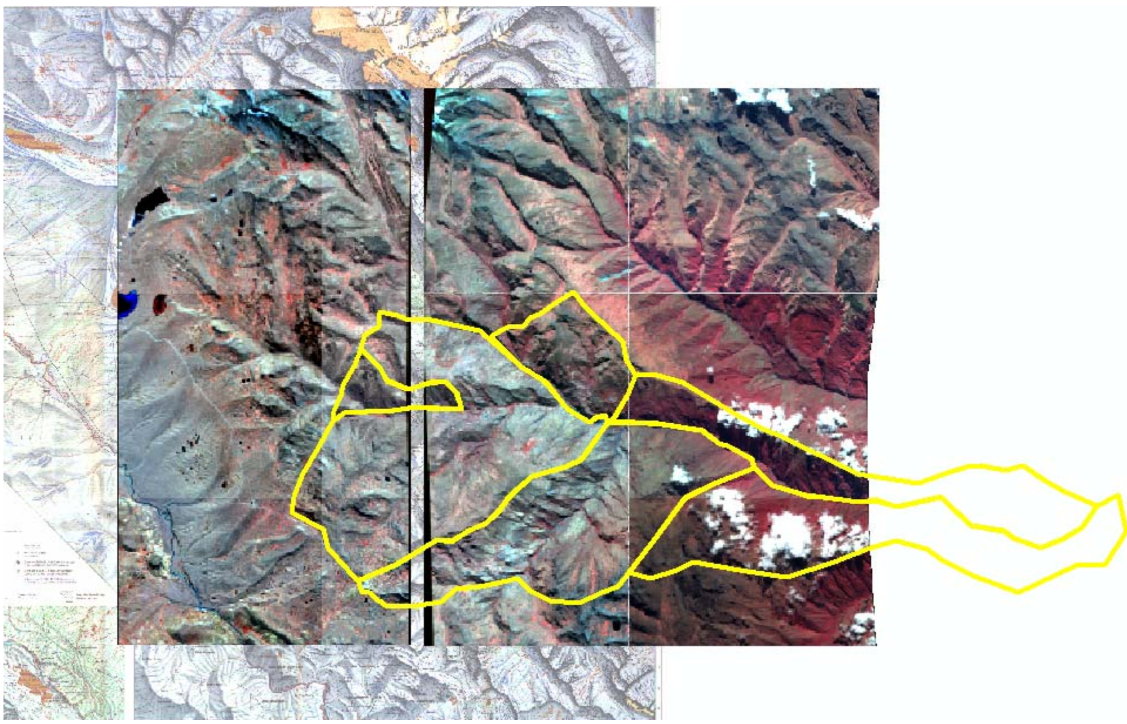


Figure 4. Image from ArcGIS that shows a rectified map in the background, the pan-sharpened QB images, and the six study area subzones. QB images are 4-3-2 RGB.

QuickBird images for Visual Inspection. The pan-sharpened tiles were added to the ArcGIS database. The results of that addition are shown above in Figure 4. This allows Sonia to view the images and make her own decisions and conclusions. I hope that the work I am doing will assist her greatly in performing this task. I would like to find a way to quickly change the view of the images from 4-3-2 RGB to 3-2-1 RGB. I think I can also use the histograms of all 15 tiles to compute a stretch to use for viewing the data. That way, the apparent intensity can be more uniform than it is now.

Classification of Features. I have identified six different classes of land use features and I have started using them to prepare the images for further examination. The six feature classes are:

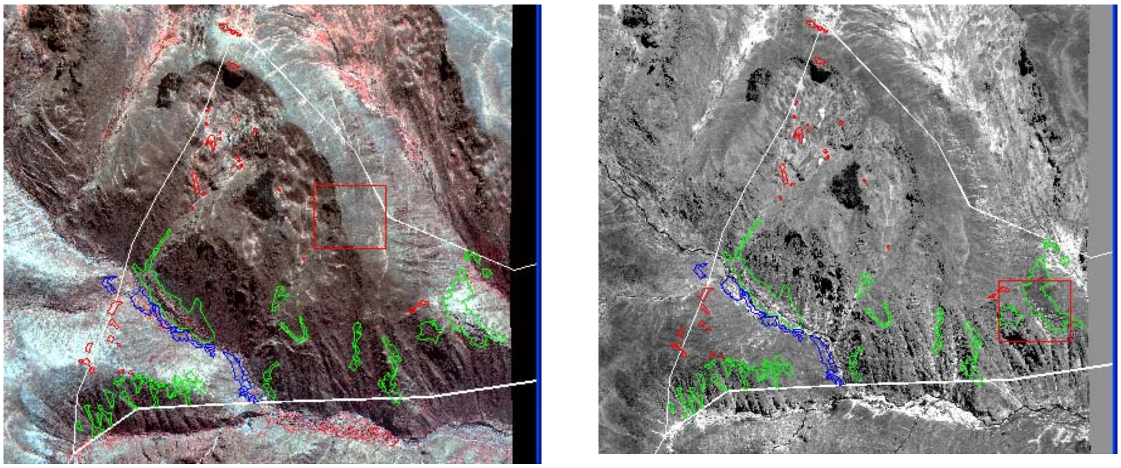
- Ranch
- Abandoned Ranch
- Wide, Gentle Terrace (Terrace A)
- Steep Terrace (Terrace B)
- Abandoned Terrace (Terrace C)
- Other Agricultural use

The ranches are identified by the presence of what appear to be stone walls used to corral herds of grazing animals. They also tend to be higher in elevation than the other classes. The "Other Ag" class tends to be at the bottom of river valleys. A ranch is abandoned if there appear to be no buildings nearby, or there are apparent ruins of buildings nearby. Abandoned terraces either have no water in them, or appear to be in some state of disrepair or non-use. They could also show some damage. There are many examples of what appear to be terraced land brimming with water. Entire mountainsides can be observed showing this feature. It is sometimes difficult to determine what is terraced and what could be more like flooded land than flooded terraces.

The wider, gentler terraces are mixed with the steeper terraces in the same terrain. What kind they are depends on the slope of the land. Steeper terraces are typically 5m or narrower in width, while the gentler terraces are wider than 10m.

Sonia has approved of these classifications and has encouraged me to also look for and identify canals, roads, bridges, and other features that may prove useful to her while in the field. Figure 5 follows, and shows the classifications in one subzone of the study area overlaid on both the pan-sharpened image and an NDVI image.

<u>Figure 5.</u> Results of Chari subzone classification. Exported original ROIs to SHPs and EVFs. Originally had 14 regions. On the left is a 4-3-2 RGB image. On the right is an NDVI image plot. Red = Ranches, Green = Terraces, Blue = Other Ag.



Summary of Progress. This section details progress of the project to date. Many pictures relating to the analysis of the images were created and are being stored in a project log book and a project log file. I plan to stick with this project for as long as Sonia desires my assistance.

1. **Images for Visual Inspection.** All of the QB images can be viewed from ArcGIS. There are 15 separate tiles, each of them are pan-sharpened, 0.6-m resolution multispectral imagery.

2. **Classification.** I have created NDVI and ISODATA classification images from each of the QB East and West images. I have computed a statistics plot of the NDVI in the classes that I've identified below.

3. **Identify Features.** I have created six classes of land use: ranch, abandoned ranch, wide terrace, steep terrace, abandoned terrace, and other agricultural. I have identified areas of each of these classes in one of six subzones of the main study area. I have imported this data into ArcGIS as SHP files.

4. **Digital Elevation Model (DEM).** The SRTM DEM is data is poor, as it has many "holes" where there were clouds at the time of image acquisition. It turns out that Sonia has had a student in Bolivia "fill in" the spotty SRTM data and create a TIN file of the results. I was able to convert that TIN data into raster data, and eventually created a TIF file. I chose to create files with varying horizontal resolution: 200m, 100m, 50m, and 10m. The 10-m resolution data looks very good, and we can use that TIF file as the DEM for the entire study area.

5. **Pan-sharpening.** The pan-sharpening effort is complete. The two QB images are composed of a total of 15 tiles, as supplied. I subjected each of these tiles to the Gram-Schmidt Spectral Sharpening technique. I encountered a problem with the results of this technique, but repeating the GS sharpening procedure allowed collection of an entire pan-sharpened image for the problem tiles.

6. **Prepare to Ortho-Rectify the Image to Make a Map.** I still need to familiarize myself with the rectification procedure.

7. **Perform Ortho-rectification if Possible.** Some of the maps in the ArcGIS database are rectified. I'm uncertain as to the accuracy of this rectification.