Analysis of Thermal Infrared Multispectral Scanner Data and Correlation to the Geology of Van Winkle Mountain, San Bernardino County, California
Adam Atkinson-Lewis
May 5, 2004

Abstract

Thermal Infrared Multispectral Scanner (TIMS) data from the Providence Mountains in the Mojave Desert of California are examined in the context of local surface geology using ESRI’s ArcGIS software. The correlation between spectral data and lithology is examined, and a search for a basic way to categorize this relationship is conducted. Raw six channel TIMS data is brought into the mapping environment and separated into its component lithological areas. The TIMS raster data is examined for each lithology, and a histogram for each is created. These histograms are then compared to those of three unknown sample sites, and a loose categorization of the unknown lithology is made.

Figure 1- Base map showing location of study within California.
Introduction

This project analyzed the ability to use a Geographic Information System (GIS) to analyze the surface cover of the Van Winkle Mountain area of the Mojave Desert in San Bernardino County, California (see Figure 1). The use of ESRI’s ArcGIS ArcView in this context was explored as relates to the software’s abilities to analyze high-resolution Thermal Infrared Multispectral Scanner (TIMS) data. A local geologic map, which included the Providence Mountains to the north-northeast, was used in conjunction with 1-meter resolution digital orthophoto quadrangles (DOQs) to assist in this investigation.

The geological map, originally in paper form, was digitized to give a general depiction of the local geology, broken up into major rock unit types and compared to the georeferenced six-band TIMS data to determine a correlation between the data types, if any. The software was used to answer the question of whether raw TIMS data can be easily utilized to determine geologic groundcover.

The analysis started with georeferenced DOQs, onto which the TIMS data, in TIFF format, was referenced using anchor point and rotation methods to achieve the best alignment possible. Each channel was then lined up with the preliminary georeferenced TIMS band and then stacked to make an independent 6-band raster that was then used for visual analysis. The geological map, originally on paper, was scanned, georeferenced with the TIMS data and then generalized into geologic units and made into a feature class for comparison.

The statistical assessment of the data involved in the analysis was based off of the grouping of TIMS data into the areas associated with the various rock formations present along the Providence 2 flight line. A histogram was then produced for each rock type raster, and the median raster value found. These values, grouped as an intended indicator for rock type, were then applied to sampled data outside of the analyzed area in effort to accurately predict the surface rock exposure.

Background

This investigation relies heavily upon the interactions of varying rock types with thermal infrared light. Every mineral, and therefore rock type, emits, reflects, or transmits a different infrared signature. Continuous transmission spectra of many minerals have been taken and catalogued using laboratory-based infrared spectroscopes, as seen in Figure 2. The TIMS data used in this project was collected from reflected infrared energy.
Thermal infrared data relies on the molecular structure of the minerals within the surface exposure, and how infrared light interacts with the dipoles of each bond or bonded group in the molecule. Different ion mass, crystal structure, and bond strength are the primary factors which affect the infrared spectrum of a particular mineral or rock; thus the spectral properties of a rock exposure are strongly tied to the composition and crystallographic system of the exposure. A full absorption spectrum is collected to categorically define mineral and rock samples; a remotely-sensed TIMS data set such as the one described in the Data section is limited by atmospheric absorption windows and flight speed, and can only collect data along separate channels, in place of the continuous spectrum collected in a laboratory environment.

### Data

The particular TIMS data used in this analysis, the Providence Mountain flight lines, site #9128 in the Mojave Desert, California, was flown on 09 September, 1991 by a C-130 aircraft under jurisdiction of a NASA Ames Research Center as part of the NASA C-130 Earth Resources Program. The TIMS system used a dispersive grating and a six-element mercury cadmium telluride detector, producing six channels with a detectable range of 8.2µm to 12.2µm. Flown at 11,500 feet, this data set has a resolution of approximately 25 feet, or 7.6 meters (resolutions are specified at 10,000 feet of altitude).

The six frequency channel ranges have been selected based upon an atmospheric infrared absorption window between 8 and 14µm (see Figure 3). This range is exceptionally appropriate, as the most notable absorption peak of silicates lies between 8 and 12µm, and decreases uniformly within this range depending upon specific silicate systems (cyclo-,

<table>
<thead>
<tr>
<th>Channel</th>
<th>Wavelength (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.2-8.6</td>
</tr>
<tr>
<td>2</td>
<td>8.6-9.0</td>
</tr>
<tr>
<td>3</td>
<td>9.0-9.4</td>
</tr>
<tr>
<td>4</td>
<td>9.4-10.2</td>
</tr>
<tr>
<td>5</td>
<td>10.2-11.2</td>
</tr>
<tr>
<td>6</td>
<td>11.2-12.2</td>
</tr>
</tbody>
</table>

This spectrum range also allows for the presence of desert varnish of moderate thickness on the sample, as peak features of the underlying mineralogy are still present despite this layer. The relatively high resolution also allows for the sensing of small exposures and outcrops along the flight lines. Also, the data was collected on a clear day in a dry environment, so there is minimal distortion due to water vapor.

This data, however, was particularly difficult to georeference, due to a non-north-south flight path, as well as lateral distortion and lack of yaw correction due to flight path and the angle at which the data was captured relative to the ground. Random noise, as well as consistent noise in channel 6 due to a faulty component within the instrument, has been cited among instrument parameters, and constitutes a small data discrepancy which does not affect the analyses undergone in this project. Channel 6 was not used in this project.

Figure 4- Georeferenced strip of prov2_1.tif, band 1 of the Providence 2 flight line.
analysis, as channels 1, 3, and 5 have been cited as the most geologically significant channels in this location. A referenced band of TIMS data can be seen in Figure 4.

All of the data sets were without metadata, as they were created specifically for this project. The DOQs, off of which everything else was referenced, came in Transverse Mercator projection, with a 500,000 meter false easting and a 0 meter false northing, using the GCS North American Datum. Although not labeled as such, the data can be assumed to be in Universal Transverse Mercator (UTM) projection in zone 11N.

Data Preprocessing and Preparation

The first step to creating a workable map space was the importation of the four separate DOQ files. These files came as single-band 8-bit TIFF files with header (*.hdr) files that were automatically placed in the correct...

Figure 5- Result of merging four separate digital orthophoto quadrangles.
projection and location by ArcMap. These four DOQs were merged into a single raster file using the following raster calculator equation:

\[
\text{merged\_doq} = \\
\text{Merge([DOQ\_1],[DOQ\_2],[DOQ\_3],[DOQ\_4])}
\]

where “merged\_doq” is the output raster file (names of individual DOQs differ from actual names). The merged DOQ image in figure 5 was produced.

The next step in preparing the data for analysis was the georeferencing of the TIMS data. The data came in two flight lines: Providence 2 and 3. Both flight line data sets were georeferenced and turned into a 6-band raster, but only the Providence 2 flight line was used in analysis, so focus will remain on this data set only. The TIFF files were imported into ArcMap as completely unreferenced files. Using the Georeferencing Tool in the ArcMap environment, the first channel of TIMS data was fit to the screen, which was focusing on the merged DOQ layer. Once common features were found after a careful examination of the vastly different-sized raster files, the TIMS file was rotated slightly and several anchor points were created between the two layers. These points were primarily focused on road and river intersections, as these were the easiest features to identify on both datasets (see Figure 6). These anchor points were used in a second-order transformation to place the band 1 TIMS data for the flight line in the correct position, closely matching the DOQ in the area of Van Winkle Mountain. While not extremely inaccurate, other parts of the map, particularly the northern extent, were not as well matched. While the second-order transformation provided the most local accuracy with the anchor points used, ideally an affine transformation would have been applied, as linear distortion is the only type of inaccuracy implied in the error and parameter documentation for the instrument and flight.

Once the first band of TIMS data was georeferenced, the *.aux files for the remaining bands were substituted with a renamed copy of the prov2_1.aux file. This file substitution, in the Windows environment rather than the ArcCatalog environment, caused ArcGIS to render TIMS bands 2-6 in the correct location, projection, and datum.

The next step in preparing the data for analysis was the creation of a digital version of the geological map of the area. The first step of this process was to scan in the paper map. The image files were then brought into Adobe Photoshop to remove all color information (the original map was in black and white) and increase the contrast to more closely render the black and white nature of the information. An added advantage of this step that became useful in rendering the map image raster was that it created a raster with RGB
values that were identical in each channel. Once this step was done, and the map files were saved in TIFF format, they were brought into the ArcMap environment and appropriately georeferenced in a manner identical to the way in which prov2_1.tif was georeferenced, as described above (see Figure 7).

Once appropriately located, the geologic map data was converted into a shapefile. To do this, a new polygon shapefile was created in ArcCatalog, and the projection information from the merged DOQ file was imported and used for the new shapefile. This file was then imported into ArcMap. This shapefile was then placed above the geology map image, and Editor was used to create the necessary polygons.

![Figure 7- Georeferenced scanned image of geologic map.](image)
To create these polygons, editing was started, and the appropriate layer was selected. The sketch tool was used to create a primary polygon outlining all of Van Winkle Mountain. This polygon was later broken into different polygons using the Cut Polygon Features tool, which requires that the user select the polygon to be cut, and is not as useful as the Auto-Complete Polygon tool, which has been inactivated in this version of ArcView, yet still remains in the Task menu on the Editor toolbar. Polygons with shared edges, which could not be cut due to shape issues with the software, were created using the Stream cursor in place of the Sketch cursor, which allows for exact tracing of the sides of one or several (selected) polygons in the creation of a new polygon. In the making of this map shapefile, all volcanic suite rock types were grouped together for ease of analysis (referred to as VOLC). A further study into this TIMS classification technique could try not making the same grouping. See Figure 8 for a view of the completed geology shapefile.

Figure 8- Completed geology shapefile digitized from the information on the geologic map
Once creation of the polygons was complete, the attribute table for the data was also edited, designating the separate rock units with a text field, designating each geology type. The creation of a new data field in the attribute table for the geology map shapefile took place in ArcCatalog, under the Fields tab in the Properties window of the shapefile. Editing was saved frequently as necessary. See Attribute table in Figure 9 for results.

In order to later make Raster Calculator calculations in ArcMap using the geology map, this shapefile was converted into a raster using the “Convert>Features to Raster” tool in the Spatial Analyst extension toolbar. The new raster file’s cell values were assigned according to geology.

A boundary shapefile was also created to delineate the region of analysis in this project, using ArcMap’s Editor feature in a manner identical to the creation of the geologic map shapefile. For sample analysis to follow, three more shapefiles were created in ArcCatalog and defined with the same projection as the DOQ file. These files were brought into ArcMap and edited to create three rectangles equal in area that reside outside of the analysis boundary polygon (see Figure 10).

Additionally, many other data sets were brought into the ArcMap environment and examined (e.g. elevation datasets, land cover information) that were not used in the final analysis that required *.e00 conversion, etc., which have not been touched upon in this report.
ArcGIS Data Analysis

Once appropriately positioned, projected, displayed, and formatted, the data were analyzed, taking advantage of specific ArcGIS functions. The goal of the analytical steps taken in ArcMap was to derive statistical data from the raster bands of each specific rock unit within the analysis area and to characterize the data so as to be easily compared to data from areas of unknown geology.

The first step in this process was to separate the TIMS data by rock type. To start, the geology raster was reclassified to separate the various geology types. To do this, the geology raster was reclassified six times—one for each geology type present in the analysis area—so that the target rock unit would have a cell value of 1, while all other units would be given a cell value of zero. Once done for each layer, the actual categorization of TIMS data by rock type could proceed.

To isolate TIMS data by geology, the following Raster Calculator calculation was used:

\[ ka_1 = [\text{prov2}_1].tif \times [\text{ka_reclass}] \]

where \( ka_1 \) is the new unit-specific TIMS raster, and \( \text{ka_reclass} \) is the 1/0
cell value raster where exposures of Ka have a value of 1. This was done with the Analysis Mask set to the Boundary layer which contained the analysis area. The above calculation was done three times per geology, once each for bands 1, 3, and 5. The resulting layers are shown in Figure 11.

The resulting eighteen layers each then had statistics calculated for them. This was also done using the Spatial Analyst extension of ArcMap,

![Figure 11- TIMS data for each lithology is separated into different raster files.](image-url)
using the Zonal Statistics tool. To do this, the Zonal Statistics tool was selected, and the layer in question was chosen (by default; defaults used) as both the Zone Dataset and the Value Raster. This calculated statistics for the entire geological exposure area within the analysis area for each lithology. These statistics were then outputted to *.dbf files, and imported into Microsoft Excel to be analyzed and have charts made, which will be discussed later in the report.

Since the ArcGIS part of the analysis of the control data was complete, the sample areas were the next to be examined. A similar Raster Calculator process to the one above was used in this procedure, but the sample area shapefile was used as the Analysis Mask, and a raster of each band was created that solely encompassed each sample area. The zonal statistics for each of these nine layers were then calculated, and brought into Excel for analysis.

Analysis, Interpretation, and Discussion

The analysis of the statistical data in Excel began with the creation of x-y scatter plots of the VALUE and COUNT columns of the ArcMap dbase IV files (*.dbf), effectively creating a histogram of each of the data sets. From the statistical data, the mode value of each TIMS data band was also plotted onto the histograms and labeled. See Figure 12 for the spectral response of each lithology. These graphs can be compared to the graphs of each of the sample areas, as shown in Figure 13.

To more closely analyze the relationship between the histograms of each lithology and each sample site, the mean difference between sample area mode value and lithology mode value for each band of each sample was calculated (Figure14). For histograms with more than one mode cell value, the mean of the two was used. The lower the mean difference, the closer a sample spectrum is to a lithology spectrum, and therefore, the more likely a sample area contains that spectrum’s associated lithology.

From this analysis method, it seems that Sample 1 is most likely Qoa; Sample 2, Qa; Sample 3, VOLC. It is important to note that the mode values for all three bands of Qa and Qoa are especially close, and it can be said that, using this technique, it is rather hard to distinguish between the two using TIMS data.

When compared with Miller’s geologic map, Sample 1 was taken from an area of roughly equal amounts of Qa and Qoa, Sample 2 taken from an area of Qoa, and Sample 3 taken from an area of Qoa that contains alluvial runoff from an area of VOLC. The spectral results are loosely comparable, especially considering the type of rock classified. Qa is an unconsolidated alluvium, Qoa a partially consolidated alluvium, and VOLC a grouping of several types of volcanic lithologies: basalts, tuffs, breccias, etc. Alluvial ground cover, especially unconsolidated, is likely to produce an inconsistent infrared spectrum, dependent on the lithologies upstream of the alluvial deposits. They are likely to have highly varying mineral contents. The volcanic group would most likely constitute much of the alluvial sediment in the Sample 3.
Figure 12- Spectral data of the various geology types within the analysis area.
Conclusions

This project has yielded some useful interpretations of TIMS data and its relation to surface geology in a GIS context. While inconclusive due to the lithologies in the sample area, and possibly limited by the sample and analysis area sizes, there exists a correlation between the TIMS data and lithology as tested. While providing a rough indication of geology, the analysis of raw TIMS data in this project was not comprehensive enough to yield accurate and reliable results.

The methods used, while providing a basic useable output, could be modified to increase the reliability of the determination of geology from TIMS data. The first aspect of this study to change would be to avoid grouping all of
the volcanic rock types together, as they are highly variable in texture and crystallography, and thus thermal infrared emissions. Also, a larger analysis area would be needed to provide a better thermal profile for each lithology, as well as a focus on a more highly lithified area, as a granite is more likely to yield a consistent thermal profile over a given area than is an unconsolidated and variable alluvium.

This examination of the correlation between remote-sensed TIMS data and geology is only a cursory look at what promises to be a very powerful and useful data gathering tool and relationship.

References

7 Donald B. Allen, personal communication. 2004 Apr.

Geologic Map data from:
Figure 15- Complete map showing all information relevant to the analysis in this project.