THE USE OF PATTERN THEORY IN SPECTRAL ANALYSIS OF SURFACE UNITS

Guoqi He, Carlé M. Pieters, John F. Mustard; Department of Geological Sciences, Brown University, Providence, RI 02912

Introduction: Although evaluating the complexity of physical processes that shape planetary surfaces is often beyond the capability of any individual approach, remote sensing techniques are used quite commonly and successfully for the compositional analysis and mapping of remote surface units, and determining their stratigraphic relations (1,2,3,4,5). Image processing tools play a key role in the successful analysis and interpretation (2,5). With the strong physical basis for diagnostic spectral properties of rocks and minerals as background, multispectral information for small areas or separate groups of pixels can often be interpreted directly. However the geologic context and study of all spectral variability pixel by pixel for a whole region is quite impractical and requires tools that can both manipulate large image-cube arrays of data while applying analysis criteria that can be accurately interpreted. Spectral mixing analysis (5,6) is one such approach, but it lacks the ability to distinguish the subtle but important details of spectra variability related to composition. Specific information about composition and physical processes active on the surface are generally deduced from the results. This study tries to use pattern theory to bring together the best of both methodologies and to incorporate the spatial information at the beginning of the analysis. The basic approach here is to define a mathematical expression via pattern theory by which the image data set can be partitioned into geochemical and/or geological meaningful surface components before subsequent more detailed analysis. Its distinction from classification is that rather than a statistic analysis of similarity, this approach uses applied knowledge of spectral properties of surface materials as constraints, and explicitly incorporates spatial coherency as a parameter.

Formulation: Our basic approach is to define a vector **v** which contains fundamental information that we are interested in analyzing in the image. This can be any number of specific properties, and the simplest set is the spectrum at each pixel. For this approach, the vector is constituted by 3 basic elements: 1) the Generator (**G** (4)) which is defined here as the average albedo of each pixel, 2) spectral information (**S**), and 3) a measure of spatial coherency (**B**). Thus the mathematical expressions is of the form **v(G,S,B),** where **S**, and **B** are both Bond Values (7). Bond value is simply the degree to which the parameters are related to the generator directly or indirectly and here includes both spatial and spectral values. The purpose of this formulation is to have the **v()** subset the original or concrete image into mutually exclusive subsets in which relevant compositional and/or information related to process are manageable, with which more detailed mathematical models can then be applied and tested. **B** can be multidimensional vector fields as well, in which case, it can be separated into both bond values, and bond directions (e.g. the bond value describes the lateral variations magnitude, i.e. texture, and the bond direction describes the sign of the variation). Digital elevation models lend themselves well to this formulation.

Discussion of v(G,S,B): The construction of **v(G,S,B)** for a given analytical objective is simply a vector. If we use each of its components to separate the original data cloud into equivalence classes (similarity groups), then the intersection of all the equivalence classes for this given vector will result in a subset of the image. This subset then contains all pixels that satisfy **v(G,S,B).** When the vector has exactly the same dimension as the image data set, the exercise becomes the study of image spectra. Whereas if we select a few of these vectors and partition each individual pixel into their fractions, it becomes a classical spectral mixing analysis. However, due to limited spectral contrast and mimickry (4), these fraction may not be accurate, nor fully model the compositional variability. If we choose **v(G,S,B)** to partition the image into nontrivial compositional similarity groups, the fraction of the partitioned group is maximized. The similarity group partitioning can also be achieved using multiple vectors as boundaries. The mixing analysis of each partitioned similarity group and/or their union will then have a more precise relationships to the mixing model. Thus the **v(G,S,B)** expression allows investigators to partition, and/or select image endmember from the Image partition with properties that can be interpreted directly for their study objectives.

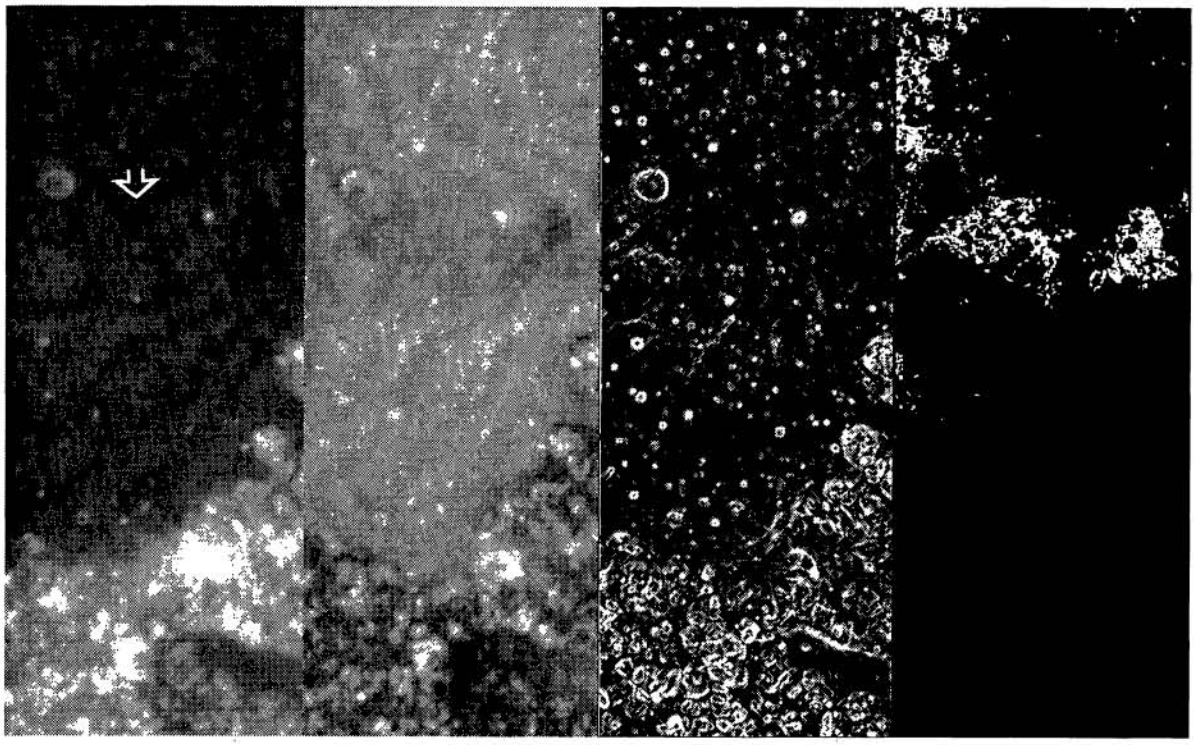
Regular Structures: It is the configuration of Generators - the combinatory principle - that is at the heart of *General Pattern Theory*(7). The typical emphasis of multispectral data analysis, however, is on data clouds bounded by the configurations, or the "distribution" of the "generators". This difference may account for a lack of pattern theory application in most image analyses. In reality, there are very few geological boundaries that are sharp enough (the mixing, rather, is the norm) to warrant the precise mathematical description of regular structures (i.e. boundaries) commonly found in pattern theory.

A Case Study at Mare Crisium: We have begun to test the formulation described above for lunar applications with Clementine data of Mare Crisium. Three major basalt groups have been previously recognized at Mare Crisium using a variety of remote sensing and lunar sample data (3). To identify these units using a spectral mixing analysis we would need at least 6 endmembers to represent all these three basalt groups (fresh & mature for blue, intermediate, and red basalts). In addition we need to have at least 2 endmembers for highland materials (fresh & mature). This is beyond what the five filter Clementine UV/VIS data can accommodate in terms of both the

Pattern Theory and Spectral Analysis. He, Pieters, and Mustard.

number of endmembers and their precision. In this example the reddest mare components are separated from a strip of data (orbit 145, between 15 to 20 degree North,) which includes the Lunar 24 landing site. The image and spectral calibration for each frame are performed using the procedures described on the WWW page www.planetary.brown.edu:/clementine/calibrations.html. Additional image processing required includes merging long and short exposures, filter to filter registration, mosaicing, and Gaussian filtering of 415 and 750 filters to match their spatial resolution. The evaluation of the **v(G,S,B)** expression in this strip includes an initial assignment of 750 albedo as the generator(Figure 1a), the 415/750 ratio (Figure 1b) is used as the initial spectral information coding, and sum of the differences between the generator and its 4 neighbors for all the filters normalized by the generator as the Bond value (Figure 1c).

Results: The resulting combination subset of the above **v(G,S,B)** (with values 9-10.8%, 0.63-0.65, 3-8% respectively) is shown in Figure 1d. The Luna 24 type of red mare data cloud is effectively isolated in this approach. This successful isolation of one data set makes further detailed analysis of more complex combinations possible. Based on this exercise we think it is also possible to separate the Image into a subset of maturation groups, highland and mare mixing boundary, mare-mare mixing boundary, and deposits with distinctive textures.



1a 1b 1c 1d

Figure 1. a) Clementine 750 nm image (used as generators, **G**), of southern Mare Crisium orbit 145 from 15 to 20 degree North. The approximate location for Lunar 24 is indicated with an arrow. b) 415nm/750nm ratio image, (**S**). c) Normalized sum of the differences between the generator and its 4 neighbors for all 5 UVVIS filters to depict the "texture" at pixel level (**B**). d) The resultant red mare similarity group partition using **v**(9-10.8% (750nm reflectance), 0.63-0.65 (415/750 ratio), 3-8% (normalized sum of the difference as bond value)) as limiting boundaries.

References: 1) Pieters et al. (1994), Science 266, 1844-1848; 2) Pieters & Englert (1993) *Remote Geochemical Analysis: Elemental and Mineralogical Composition,*,Cambridge University press; 3). Head et al. (1977), *Mare Crisium*, Pergamon Press 43-74; 4) Sabol et al. (1992) JGR 97, 2659-2672; 5) Adams et al. (1993) *Remote Geochemical Analysis: Elemental and Mineralogical Composition,* Cambridge University press, Ch. 7; 6) Mustard & Head (1996), JGR, (submitted); 7) Grenander, (1993), *General Pattern Theory*, Oxford Science Publications