Proceedings of Geomorphometry 2009. Zurich, Switzerland, 31 August - 2 September, 2009

Automated Classification of Martian Morphology Using a Terrain Fingerprinting Method

R. Koenders[1], R. C. Lindenbergh[1], T. E. Zegers[2]

[1]Delft University of Technology, P.O.Box 5058, 2600 GB Delft, The Netherlands
Telephone: (+31)610383808
Email: roderikk@gmail.com

[2]Utrecht University, P.O Box 80125, 3508 TC Utrecht, The Netherlands

1. Introduction

The planet Mars has a relatively short human exploration history, while the size of the scientific community studying Mars is also smaller than its Earth equivalent. On the other hand the interest in Mars is large, basically because it is the planet in the solar system most similar to Earth. Several satellites are currently orbiting Mars, and transmit data back in unprecedented detail. In fact, the Martian surface is mapped at up to 5 times higher resolution than the bottom of the ocean here on Earth.

The scientific community studying Mars has already made great discoveries concerning, for example, the variability of the surface (Bibring, 2005), and the presence of water. To learn more about the history of the surface and about the planet as a whole, data generated by different satellite missions will have to be combined. Processing such large, multi-attribute datasets at a global Martian scale requires efficient automated classification methods.

The use of automated classification in combination with geomorphometric data has only recently been possible on Mars with the creation of the global Mars Orbiter Laser Altimeter (MOLA) digital elevation model (DEM) (Smith et al. 2003), as obtained between 1997-2001 by the Mars Global Surveyor. (Bue and Stepinski, 2006) demonstrated the potential of classifying global MOLA DEM data and concluded that similar methodology could be applied on other data sets like the ~60m spatial resolution DEM, as currently under construction from High Resolution Stereo Camera (HRSC) images collected by ESA's Mars Express (Gwinner, 2007).

On Earth, morphological classification has been used for numerous specific applications (Guzetti and Reichenbach, 1994; Hosokawa and Hoshi, 2001). Also only relatively recent it was demonstrated that attributes like gradient and roughness, as derived from elevation data, can be used to construct a multi-attribute feature vector, that, possibly in combination with other data, like intensity or multi-spectral data, can be consecutively applied in land surface and vegetation classification procedures (e.g. Antonorakis et al., 2008; Bork and Su, 2007; Chust et al., 2008).

Even though the use of automated classification on Martian datasets has great potential, it is not yet being used as intensively by the scientific community studying Mars. The research presented in this abstract therefore formalises the methodology presented by Bue and Stepinski (2006) as the Terrain Fingerprinting Method (TFM) in Section 2. We have applied the TFM to several areas on Mars based on the MOLA DEM, which has a maximum spatial resolution of 400 meters per pixel; HRSC DEM, which has a maximum resolution of 50 meters per pixel; and a combination of the MOLA DEM with data from the Mars Express mineralogical spectrometer (OMEGA). The present abstract focuses on an analysis of the combination of OMEGA and MOLA DEM data as presented in Section 3.

2. The Terrain Fingerprinting Method

The Terrain Fingerprinting Method consists of 5 steps that closely follow the steps a geoscientist takes when he or she analyses a terrain.

Step 1: Defining the research question

The first step requires the scientist to define the research area and the type of terrain to be analysed. This is exactly how a scientist starts researching a terrain on Earth; first selects a region to analyse and then he decides what research questions need to be answered about that region – or vice versa.

Step 2: Choosing the attributes

The next step requires the expert to describe how the terrain would be analysed if it was analysed by hand. For example, to categorise a certain region into different geological units a geoscientist would look at terrain attributes including slope, elevation, terrain roughness, and rock colour to determine where the geological units' boundaries lie.

Step 3: Converting unprocessed data

The next step is to translate the attributes found in the previous step into a computer readable format. This involves converting unprocessed satellite data of the region to be studied to data read by GIS software, and finally deriving the required attributes; for example creating slopes from elevation data.

Step 4: Clustering data

In the supervised terrain classification, this step would involve manually classifying the terrain with GIS software. This step is replaced by an automated classification method. This research uses a combination of a partitive clustering technique (Self-organising Maps, Kohonen, 2001) and a hierarchical clustering technique (Ward clustering, Ward, 1963) to create a fingerprint of the terrain analysed automatically, see section 2.1.

Step 5: Analysing and validating results

The final step in the TFM is also very similar to the final step in the manual process. It includes validating, analysing, and interpreting the classification made by the automated classification.

2.1. Automated Classification of Landforms

Classification schemes can generally be divided into two categories: hierarchical and partitive clustering. In hierarchical clustering, each data point can be seen as being on the end of a twig on a tree, which is part of a branch, which connects to the tree. The more two twigs are set apart, the more dissimilar the two points of data are.

This way of classifying is very useful for smaller datasets. However, as each data point corresponds to a twig, the classification tree, and therefore the storage and processing power, grows with every point of data.

On the other hand, partitive clustering does not look at each individual data point; it tries instead to look at the dataset as a whole to find clusters. To continue the analogy, partitive clustering tends to ignore the twigs of the tree and only look at the branches. One of the disadvantages of this approach is that it is more difficult to distinguish between clusters that are closer together.

When classifying large areas of terrain and/or terrain at a high resolution, the number of data points grows very fast. To still be able to classify these terrains, while avoiding the disadvantages of the single methods, a combination of the partitive and hierarchical clustering techniques can be applied. Vesanto (2001) has shown the feasibility of this technique.

The present research uses self-organising maps (Kohonen, 2001) to create an initial mapping of the original data to a lower number of proto-vectors. These proto-vectors are then clustered using so-called Ward clustering (Ward, 1963) to bring them down to 20 classes.

3. Results

To design and validate the Terrain Fingerprinting Method we have applied it to several different use-cases. This section describes each of the TFM steps for one of the use-cases.

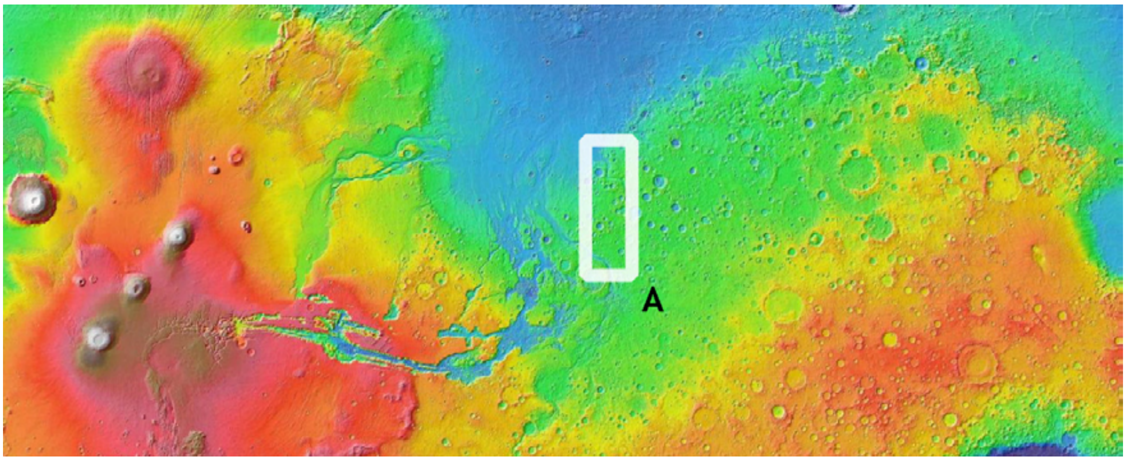

Figure 1. The area analysed in this research.

R*esearch question*: The research question for area A is whether the area contains terrain that is on one hand safe to land on for a rover such as ExoMars, whilst on the other hand the terrain is also interesting from a scientific point of view. In this case safe is defined using the characteristics given in Table 1; and being interesting is based solely on whether there are indicators for phyllosilicates[1] in the OMEGA data.

This particular area—around the Mawrth Vallis—was chosen because it is known for its high phyllosilicates content and its possible suitability as a landing site for NASA's Mars Science Laboratory (Michalski, 2008).

Characteristic	Lower Bound	Upper Bound
Elevation [m]	-	-1000
Slope [°]	0	2

Table 1. Primary safety characteristics for ExoMars.

A*ttributes*: To classify the terrain by hand into safe, unsafe, and interesting units, it was agreed that there were 5 relevant attributes:

1 Phyllosilicates are considered to be an indicator of the past presence of liquid water; ExoMars scientific and exploration goals are to find traces of past and/or present life; since life as we know it requires, among other things, the presence of liquid water, we assume that evidence of phyllosilicates will indicate an interesting region.

1. Elevation: to distinguish between low and high;
2. Slope: to identify crater walls and cliffs;
3. Filled Difference[2]: to identify craters;
4. Filled Slope[3]: to distinguish crater walls and cliffs elsewhere; and
5. Phyllosilicates: to identify interesting terrain, based on Pelkey (2007).

U*nprocessed data*: The classification is based on a set of OMEGA images that covered the region, all of which had approximately the same resolution of 1000 meters per pixel.

The file was read into `GRASS` (GRASS Development Team, 2008) and the attributes were generated using the following `GRASS` routines:

1. Elevation: -
2. Slopes: `r.slope.aspect`
3. Filled Difference: `r.terraflow` and `r.math`
4. Filled Slope: `r.terraflow` and `r.slope.aspect`
5. Accumulation: `r.terraflow`

C*lustering data*: Figure 2a shows a close up of part of the elevation data on which the previously discussed attributes were based. The final classification result is shown in Figure 2b. At the bottom, each colour can be seen to correspond with a class, and the classes are grouped according to their relative distances and the clustering algorithm.

A*nalysis*: The groups indicated in the legend of Figure 2b can also be visually identified in the mapping of the classification: group 1 can be identified on the higher, southern terrain; group 2 fills up the spaces between group 1 and the craters of the said highlands; group 3 can be identified as craters; group 4 is located in the northern lowlands; and group 5 is found in the area between the lowlands and the highlands.

In order to verify these claims, Table 2 summarises the attribute means for each group. The first row shows the means for the attributes over the entire area of study, this can be used to compare the values of the other groupings.

The final column in this table shows the amount of phyllosilicates in the group. According to Pelkey (2007) only levels above 0.02 indicate the presence of this mineral group on that location. None of the groups in Table 2 show this level, though the values in the table represent a mean over an area. When compared to the mean of all groupings, groups 4 and 5 show above average phyllosilicate levels.

It can be concluded from the data given in Table 2 that groups 4 and groups 5 potentially identify terrain types that are both safe and interesting; with the latter being more interesting as it has a higher phyllosilicates content.

---

2 Filled difference is an attribute generated by using a fill algorithm to make lakes out of all the craters and blocked channels, finally the `filled difference' attribute is the difference between this map of lakes and the real elevation map; thus generating domes where craters are located.

3 The filled slope attribute is generated by creating a slope map from the `map of lakes' created in the filled difference attribute procedure.

Groups	Area [%]	Elevati on [m]	Fill. Diff. [m]	Fill. Slope [°]	Slope [°]	Phyllo. [-]
All	100	-2507	233	0.83	2.08	0.00383
1	24	-2008	21	2.52	3.43	0.00368
2	38	-2058	135	0.14	1.21	0.00127
3	13	-3216	1221	0	3.42	0.00193
4	12	-3736	94	0.2	1.01	0.00775
5	8	-2727	19	0.73	1.25	0.00871

Table 2. Attribute means for different groups in area A.

4. Conclusions

Bue and Stepinski (2006) used MOLA data to create an automated classification of Martian terrain. The present research has looked at how their methodology can be formalised to appeal to the broader planetary science community, and how it can be applied to other types of data.

As can be deduced from the process description above, the steps taken for the TFM are almost identical to those taken for a terrain classification done by hand. The fingerprint produced in step 4 of the TFM is the key difference. The terrain fingerprint produced can for instance be used to:

- perform an initial terrain classification on an area to quickly identify the primary terrain classes, which can be used as input for a manual classification;
- locate terrains elsewhere on the surface that have the same fingerprint and could therefore be similar terrains;
- use a classification made for one area and apply it to a different area to quickly, and consistently classify this new area with the same classes as the original area; and
- translate a classification made manually by an expert to a computer readable terrain fingerprint and apply it consistently to other areas.

The most challenging TFM step is the one where the list of attributes is created. During manual terrain classification a geoscientist combines many attributes, including interpretations from previous terrains. More research is required to turn these more `interpretive' attributes into attributes that the computer can understand.

Another property of TFM—and of analysing datasets of Mars in general—is that combining the different datasets into a single frame of reference requires many processing tools and steps. Scaling the analysis to cover larger swathes of terrain will therefore require investigations on how to optimally load the unprocessed data.

Another often-heard criticism about the automated classification method is that the results are not similar enough to how a geoscientist would interpret the terrain. One way to circumvent this criticism is to use a manual terrain classification as the basis for the fingerprint.

In the course of the present research we have identified several items for further research. Due to the lack of ground-truth on Mars it will be important to validate TFM on Earth with terrestrial geoscientists.

It is expected that using geostatistical methods considering both cross-correlation between attributes, and spatial correlation within one attribute, will further optimise the distinction between different terrain types. Moreover the use of geostatistical methods provides a framework to combine datasets with different spatial point densities and/or individual point qualities.

Figure 2. *a)* MOLA elevation data of area A; *b)* Classification made with TFM.

References

Antonarakis, A.S. and Richards, K.S. and Brasington, J.,2008. *Object-based land cover classification using airborne LiDAR*, Remote Sensing of Environment, 112(6), pp. 2988-2998.

Bibring, J.P. and Langevin, Y. and Gendrin, A. and Gondet, B. and Poulet, F. and Berthe, M. and Soufflot, A. and Arvidson, R. and Mangold, N. and Mustard, J. et al, 2005. *Mars Surface Diversity as Revealed by the OMEGA/Mars Express Observations*, Science, Vol 307, 1576-1581.

Bork, E.W. and Su, J.G.,2007. *Integrating LIDAR data and multispectral imagery for enhanced classification of rangeland vegetation: A meta analysis,* Remote Sensing of Environment, 111(1), pp. 11-24.

Bue, B.D. and Stepinski, T.F., 2006. *Automated classification of landforms on Mars*, Computers and Geosciences, Vol 32.

Chust G, Galparsoro I, Borja A, Franco J, Uriarte A (2008) *Coastal and estuarine habitat mapping, using LIDAR height and intensity and multi-spectral imagery*. Estuarine Coastal and Shelf Science 78:633–643.

GRASS Development Team, 2008. *Geographic Resources Analysis Support System (GRASS) Software*, Version 6.3.0. http://grass.osgeo.org .

Guzetti, F. and Reichenbach, P., 1994. *Towards a definition of topographic regions for Italy*, Geomorphology, Vol. 11.

Gwinner, K. et al., 2007. *Global mapping of Mars by systematic derivation of Mars Express HRSC high-resolution digital elevation models and orthoimages*, ISPRS WG IV/7 Extraterrestrial Mapping Workshop, Houston, Texas.

Hosokawa, M. and Hoshi, T., 2001. *Landform classification method using self-organizing map and its application to earthquake damage evaluation*, IEEE 2001 International Geoscience and Remote Sensing Symposium, IGARSS'01.

Kohonen, T., 2001. *Self-organizing Maps*, Springer.

Michalski, J. and Bibring, J. et al, 2008. *Mineral Mapping of High Priority Landing Sites for MSL and Beyond Using Mars Express OMEGA and HRSC Data*. American Geophysical Union, Fall Meeting 2008, abstract P33B-1463

Pelkey, SM and Mustard, JF et al, 2007. *CRISM multispectral summary products: Parameterizing mineral diversity on Mars from reflectance*. J. Geophys. Res, Vol 112.

Smith, D.E., M.T. Zuber, G.A. Neumann, E.A. Guinness, and S. Slavney, 2003. *Mars Global Surveyor Laser Altimeter Mission Experiment Gridded Data Record*, MGS-M-MOLA-5-MEGDR-L3-V1.0, NASA Planetary Data System.

Vesanto, J. and Alhoniemi, E., 200. *Clustering of the self-organizing map*, IEEE Transactions on Neural Networks, Vol. 11.

Ward Jr, J.H., 1963. *Hierarchical grouping to optimize an objective function*, Journal of the American Statistical Association.