Degradation analysis of craters on Mars

Magda Sofia Carvoeiras Rocha
magdarocha@ist.utl.pt

Instituto Superior Técnico, Lisboa, Portugal

November 2015

1. Abstract

The rapid growth of image acquisition technology led to an enormous increase on the amount of planetary images taken by satellites. The possibility of analyzing the content of these images was a turning point on planetary research. Impact craters are among the most studied geomorphologic planetary features because they yield information about past geological processes and are a powerful tool for estimating the ages of planetary surfaces. Furthermore studying these structures helps to uncover information that can drive future robotic or even human exploration to the Red Planet. To overcome years of visual analysis, this paper presents a method for the automatically classification of the degree of preservation of Mars craters based on the state of its rim. Texture features were extracted and the performance of the SVM (Support Vector Machine) classifier was tested in craters of different sizes and shapes located on two distinct regions, Mars Lunae Palus and Syrtis Major Quadrangles. The proposed method led to classification accuracies of ~80%.

2. Introduction

Remote Sensing technology consists on taking long distance images of the surface of the Earth and by extracting information from the pixels combined with supervised classification techniques revealed to be a powerful tool for scientists gather image information, without the physical present of technicians in the field. Rapidly this concept was expanded to Mars.

The Red Planet geological structures density, patterns and morphology have always flickered interest among planetary scientists which pursue answers for the dynamics of the planet.

Nowadays there are many techniques to estimate the relative age of the surfaces as crater density and superposition phenomena. Crater classification which provides information about the degradation rate of those same craters says that the more modified a crater, the older it is.

Degradation factors as new crater impacts, landslides and ejecta deposition, lava flooding, floor fracturing, crater infilling, glacial accumulation, winds and tectonic activity contribute to the erosion and degradation of these geological structures.

A fresh looking crater with a low degradation state has been less subject to the erosion factors and for this reason has sharp rims and has low chances of having suffered subsequent impacts from a meteor (it will naturally show a lower degradation index in any built scale). As the age of the crater increases we normally see a change on its morphology, the rim begins to be more rounded and at advanced stages can become hardly noticeable or can be completely erased. The development of an automatic method to quantify the degradation state of impact craters through the detailed analysis of their rims is currently very well welcome by the Planetary Science community, as it could be a new key to new discoveries into the geological history of Mars and also of other planetary bodies.

The work takes as starting point the use of an algorithm [1] developed to thoroughly extract the outline of the crater and by means of Dynamic programming. This has allowed a correct definition of the border of the crater and the extraction of physical statistics related to the crater as its radius and center location.

The second step involves a proper extraction of the features around that same area adjusting the region of the extraction to the size of the crater. Those features measured the data distribution by means of histograms and percentiles vectors.

The last stage involved the construction of a SVM classifier that uses the features extracted from the images into the learning/training process of the classifier, those were previously assign to a label (ground-truth) related to their degree of conservation with the help of knowledge's of investigator Pedro Pina.

3. Background

3.1 Support Vector Machine

Support Vector Machines (SVMs) [2] are used to perform image classification, by mapping input feature vectors onto the underlying image class labels.

The operation of the binary SVM algorithm is based on finding the hyperplane that gives the largest minimum distance to the N training examples: $(x_1, y_1), \ldots, (x_N, y_N)$ where $x_i \in R_m$ is a m-dimensional feature vector representing the i^{th}training sample, and $y_i \in \{-1, 1\}$ is the class label of x_i. The optimal separating hyperplane maximizes is represented with the following expression:

$$w^T x + b = 0,\ \text{w} \in R^m,\ b \text{ is a scalar} \quad \textbf{(1)}$$

The parameter pair (w, b) corresponding to the optimal hyperplane is the solution to the following optimization problem:

Minimize $L(w)$:

$$L(w) = \frac{1}{2}\|w\|^2 \quad \textbf{(2)}$$

Subject to:

$$y_i = (w^T x_i + b) \geq 1, i = 1, \ldots, N. \quad \textbf{(3)}$$

When linear SVM does not retrieve good results it is used the concept of soft margin.

When linear SVM does not achieve good results, a nonlinear SVM is used, the basic idea is to map each feature vector *x* by nonlinearly mapping *φ(x)* to a higher dimensional space in which the optimal hyperplane is found. The Kernel function, $K(x_i, x_j)$ achieves this mapping by computing the inner product of vectors $\varphi(x_i)$ and $\varphi(x_j)$

$$K(x_i, x_j) = < \varphi(x_i), \varphi(x_j) >. \quad \textbf{(4)}$$

4. Methodology

The implementation phases to characterize the crater's rim into the tree states of conservation: Preserved (type '0'), Modified (type '1') and Degraded (type '2') were:

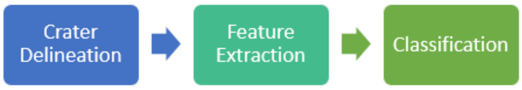

Figure 1: The tree main implementation steps

4.1 Crater delineation

The first step was to detect the crater outline so the entirely rim region could be analyzed. The algorithm used [1] will be briefly summarized. In the first stage, the algorithm detects intensity transactions in the image of the crater and finally links all the edges in order to design the contour. The algorithm proposed uses an Edge Map e(x)∈[0,1], measures the variation of directional intensity on each point and assigns the value e(x)=0 to a pixel when there is a strong intensity variation in the neighborhood of x in a orthogonal direction to the crater contour. The value e(x)=1, is assigned to a pixel if the image is constant in that direction, which means that x is not a part of the contour.

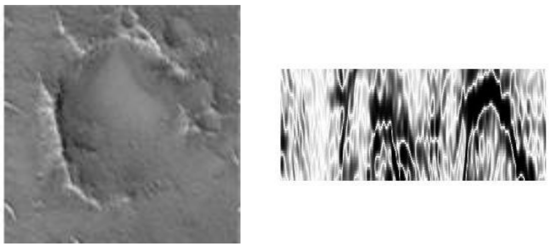

Figure 2: Crater on the left and edge map on the right

The final stage involves the computation of a closed contour, x(s) which minimizes the energy functional:

$$E(x) = \int e(x(s))\, ds + \mathrm{E}_{\mathrm{int}}. \quad \textbf{(5)}$$

The parameter $E_{int}(x)$ measures the deviations of the crater contour, x(s), to a circle and *s is* the arc length parameter of the curve. To simplify the problem the crater was defined in polar coordinates.

The Dynamic Programming algorithm delineated the contour of the craters using as input the Edge Map and minimizing the energy functional. It allowed to properly identify the crater in the input image.

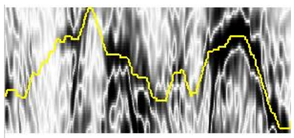

Figure 3: Edge map and optimal contour (yellow line) in polar coordinates related to the crater in Figure 2

4.2 Feature Extraction

The rim was the feature considered more distinctive to characterize the degree of preservation of the crater and the decision about what kind of features would lead to a better classification results was made based on a bunch of factors.

Features based on surface geometry, texture and color were analyzed and revealed to not be ideal for the study possibilities analyzed**.**

Surface geometry information could be useful if 3D images taken by the Rovers were used for distinguishing for example different terrains slopes near the crater's rim.

Color somehow is limited for the surface of Mars and would not make sense because the images are in gray scale.

Texture features revealed to be the ideal to proper reflect the state of the rim. For that reason, to measure the texture of the craters rim, gradient information was extracted across the images. The gradient of an image has two kinds of information, one is the magnitude and the other is the direction of the gradient. Magnitude gives the information of how rapidly the image is changing and direction of the gradient tells which image direction is changing more rapidly.

The rim was divided into 8 sections, the features were extracted for each sector and then the classification was obtained for each individual sector.

The idea of dividing the crater into sections and analyzing them as a combo of the behavior on each one seemed the best approach to overcome some ambiguous behavior on the global appearance of the rim. For example if in a preserved crater a small part of its contour vanishes or if it shows a considerable amount of shadows the crater obviously continues being a preserved crater.

To select which pixels belong to the section ring the Euclidian distance of all image points to all points of the contour. If the condition in equation (6) was true the point belonged to the rim section.

$$R\_min < d_{min}(P) < R\_max \quad \textbf{(6)}$$

Then each point that had verified the condition was assigned to a label depending on their position in the frame.

The Figure 4 shows the division implemented by the algorithm as well as the location of each section.

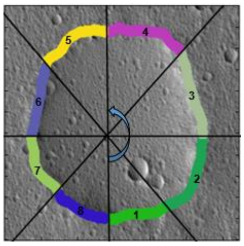

Figure 4: Section identification (by numbers and colors)

Two type of gradient features were used in the classifier: Gradient histograms (3 different types) for each section and percentiles for the gradient values (two types) calculated in the sections.

For each section 3 types of gradient histogram were extracted:

•Histogram of Oriented Gradient (HOG), proposed by Dalal and Triggs [3] for Human detection task.

$$h_i = \sum_{x \in s^k} |\mathrm{g(x)}| \times b_i(\varphi(\mathrm{x})) \quad \textbf{(7)}$$

$$b_i(\varphi) = \begin{cases} 1 \; if \; \varphi \in i^{th} \text{ bin} \\ 0, otherwise \end{cases}, \quad \textbf{(8)}$$

•Gradient Phase Histogram proposed by Schiele and Crowley in 2000 [4].

$$h_i = \sum_{x \in s^k} b_i(\varphi(\mathrm{x})) \quad \textbf{(9)}$$

$$b_i(\varphi) = \begin{cases} 1 \; if \; \varphi \in i^{th} \text{ bin} \\ 0, otherwise \end{cases}. \quad \textbf{(10)}$$

•Gradient Magnitude Histogram also proposed by Schiele and Crowley in 2000 [4].

$$h_i = \sum_{x \in s^k} b_i(|\mathrm{g(x)}|), \quad \textbf{(11)}$$

$$b_i(|g|) = \begin{cases} 1 \; if \; |g| \in i^{th} \text{ bin} \\ 0, otherwise \end{cases}. \quad \textbf{(12)}$$

The second type of feature is the percentiles to the magnitude and direction of the gradient values in each section given the percentage vector. This is a very useful statistic measure to assess the numerical data in a way that lets us know what is the maximum gradient value for each section given the percentile. Percentiles analyzed were the 5th, 25th, 50th, 75th and 95th.

$$P_k = \frac{K(n+1)}{100} \; th \; observation, \; n = \; observations \quad \textbf{(14)}$$

4.3 SVM classifier implementation

(SVM) of the type One-Versus-All was implemented for the multi-class classification of the data since a ternary label system was previously defined as the best way of aggregate the craters into the tree categories.
The ground-truth label defined was ternary, l∈[1,2,3] so the technique used perform pair-wise comparisons between the three classes, creating three binary classifiers that distinguish one class for the other two. The classifier which generates the highest value for its decision function is selected as the winner and the corresponding class label is assigned to the data.
SVMTRAIN function was used to train the classifier :

$$c = \sum_{i=1} \alpha_i \, K(s_i, x) + b. \quad (15)$$

if c ≥ 0, x is classified as a member of the first group, otherwise it is classified as a member of the second group. The three retrieved models by the function SVMTRAIN are used in the second stage of the classifier where using as input the matrix of test instance the LIBSVM function SVMPREDICT estimate the probability of each instance of the test matrix belonging to each of the tree model representing the three possible classes to be assigned.
To avoid overfitting Cross Validation was implemented as rotation model using 59/109 craters (Dataset1/Dataset2) for training and only one for testing.
The parameter optimization was achieved using Nested Cross Validation where the inner loop is used to perform the tuning of the parameters while the outer loop of the Cross Validation is used to compute an estimate of the model error.

5. Experimental Results

Two distinct sets from Mars were selected to develop and test the methodology (Figure 5) Set 1 consists of a younger surface region of plain characteristics while Set 2 consists of a much more irregular surface. Each set is covered by one single HiRISE image (ESP_011491_2090 for Set 1 and ESP_025555_1940 for Set 2).

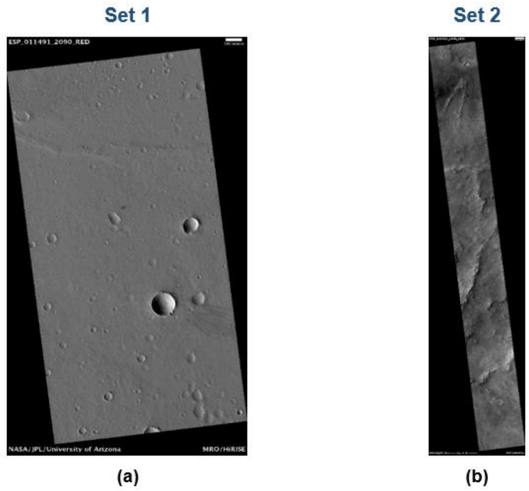

Figure 5: Overview of the two analyzed regions: (a) Recent Impacts- Lunae Palus quadrangle, Scale:25 cm/pixel; (b) Layered Ejecta near Syrtis Major quadrangle, Scale :50cm /pixel

In Table 1 is the details for each selected site, namely the location on Mars according to longitude and latitude in decimal degrees, the image resolution as well as the total number of craters selected in each site with the respective dimensions.

Site	Camera	Image	Lat	Long	Resolution	Craters	Diameter(m)		
#		id	º	º	m/pixel	Nb	min	max	avg
1	HiRISE	ESP_011491_2090	29,57N	371,51E	0.25	60	9	722	67
2	HiRISE	ESP_025555_1940	13.94S	69.60E	0.50	110	10	585	165

Table 1: HiRISE images and craters Datasets

The algorithm was tested over two DataSets of crater images sampled from the two region images in Figure 5. The first one (named DataSet1) is constituted by a total of 60 crater images (480 sections).
The scale of the DataSet 1 images has twice the resolution of the one of DataSet 2, has we can see on Table 3.1. For DataSet 1 the largest crater image has the dimension of 3561x3485 pixels, while the smallest measures 62x61 pixels with an average of 386x404pixels. For DataSet2 the largest image has the dimension of 1673x1769 pixels, while the smallest measure 26x27 pixels, with an average of 235x249 pixels. This DataSet contains 26 images with a size inferior to the size of the smallest crater in DataSet1.
For DataSet1 the overall accuracy while using histograms as features was ~80% for the three types of histogram. Table 2 shows the results for the three classifiers using the histograms as features.

The SVM classifier that used HOG has not reached satisfying results.

SVM classifiers statistics using histograms as feature vectors-Data Set1											
Features type 1 : Gradient Magnitude Histograms				Features type 2 : Gradient Phase Histograms				Features type 3: HOG			
Actual class \ Predicted class	0	1	2	Actual class \ Predicted class	0	1	2	Actual class \ Predicted class	0	1	2
0	7	22	0	0	1	28	0	0	0	29	0
1	0	311	6	1	0	311	6	1	0	311	6
2	0	60	74	2	0	55	79	2	0	60	74
Final Accuracy≈ 81,6%				Final Accuracy≈ 81,5%				Final Accuracy≈ 80,0%			

Table 2: Statistic Results for Data Set1 using histogram as feature vectors

For the Features Type 1 the learning pattern for 3 preservation states was:

- For crater of type '0' the 5th bin histogram(which represents stronger gradients when compared with the lower range bins)must contain a significate amount of points ~20%
- Craters of type '1' have a differentiated pattern many times similar to the histogram pattern of the other two type craters. This is the main reason for the classifier have not reach better results.
- Craters of type '2' have 90% or more of the points in the first bin and almost the rest in the second.

For the Features Type 2 the learning pattern for the 3 preservation stages were:

- For crater of type '0' the histogram as 3 major peaks well differentiated between the three, the one in the middle with more impact than the other two.
- Craters of type '1' once again has a modified pattern similar to crater type'0' and crater type '2'responsible. This is the main reason for the mis-classification achieved in this DataSet.

- Craters of type '2' have histograms with samples relatively uniform distributed for all the bins

For DataSet1 the results while using percentiles as features are represented in the confusion matrixes (Table 3)

SVM classifiers statistics using gradient percentiles as feature vectors-DataSet1							
Features type 4: Percentiles vectors of the magnitude gradient sections				Features type 5: Percentiles vectors of the phase gradient section			
Actual class \ Predicted class	0	1	2	Actual class \ Predicted class	0	1	2
0	0	29	0	0	0	29	0
1	0	315	2	1	0	316	1
2	0	111	23	2	0	88	46
Final Accuracy≈ 70%				Final Accuracy≈ 75,4%			

Table 3 Statistic Results for Data Set1 using data percentiles as feature vectors

The major difficulty while using this two types of features is again the variety of patterns that type '1' crater exhibit.

In DataSet2 it can be concluded that for this Data Set the results were considerably less satisfying. Since the algorithm was dealing with images much smaller and with considerable less definition than those in DataSet1 the feature histogram were compromised.
In some situations the image has reasonable dimensions but the crater on it does not. Because of that only two or three sections could be extracted leading to completely nulls histograms for the other sections In other situations the dimension of the images is too small but the crater image itself occupied almost the entire image, so the algorithm could hardly distinguish the 8 sections in a correct way. In this case, the points do not represent well the sections and the lack of definition of the image prevents the possibility of distinguishing the contour. The results obtained for all the features was ~26%, the classifier was not capable of learning any feature pattern.
The 51 smaller images were removed from DataSet2 and the whole algorithm tested again.
The results increased substantially.
The overall results for the histogram features were approximately 70% as it´s possible to see in Table 4. The results fell in about 10% when compared with the results of DataSet1.

SVM classifiers statistics using histograms as feature vectors-DataSet2									
	Features type 1 : Gradient Magnitude Histograms			Features type 2 : Gradient Phase Histograms			Features type 3: HOG		
Actual class \ Predicted class	0	1	2	0	1	2	0	1	2
0	15	42	0	2	55	0	0	57	0
1	0	279	29	0	278	30	0	280	28
2	0	67	40	0	49	58	0	51	56
	Final Accuracy≈ 70%			Final Accuracy≈ 71%			Final Accuracy≈ 71%		

Table 4: Statistic Results for part of Data Set2 using histogram as feature vectors

This situation could be explained by the fact that this DataSet has samples with a larger ambiguity. For some sections and also for entire craters it becomes harder to assign them into an obvious GT. Ambiguous GT lead to ambiguous results. The increase of the error in GT is the explanation for the decrease in the results.

The percentiles of gradient magnitude and phase had an accuracy of approximately 60% (Table 5) revealing to be in concordance with the histogram results as occurred for DataSet1.

SVM classifiers statistics using gradient percentiles as feature vectors-DataSet2						
	Features type 4: Percentiles vectors of the magnitude gradient sections			Features type 5: Percentiles vectors of the phase gradient section		
Actual class \ Predicted class	0	1	2	0	1	2
0	0	57	0	0	57	0
1	0	273	20	0	270	23
2	0	83	26	0	75	32
	Final Accuracy≈ 63%			Final Accuracy≈ 64%		

Table 5: Statistic Results for part of Data Set2 using data percentiles as feature vectors

6. Conclusions

None of the few methods implemented until now used the rim to classify craters conservation state, so this project can be considered pioneer and useful for future investigations.

The results obtained in crater classification are good enough to be applied in crater images with certain characteristics. Images with some kind of superposition phenomena or affected by shadows due to the Sun position affect negatively the performance of the classifier. In the first case the algorithm actually extracts features that do not belong to the image being classified and in the second case shadows are perceived as texture features and also distort the learning phase of the Support Vector Machine Classifier.

The image resolution is also an important aspect, if we consider small images that are highly noisy it leads to the extraction of very few features that fail because of the redundancy of its information and also by the small number of samples that do not give enough information to be considered representative of a pattern of rim´s behavior. Craters with radius dimensions bellow 16 pixels fit this category.

Images that do not exhibit this kind of behavior achieve satisfying classification results, the classifier was able to learn the tree types of rim. The occurred miss-classifications were easy explained because of the appearance of the rim in the modified category that in some cases presented similarities with the other two preservation stages. Future improvements like combining more than one type of geological crater feature would be a possible approach to the construction of more reliable classifiers.

The application of methods that can detect and remove shadows and the subdivision of the crater into more sections so that the generalization in rims behavior decreases could also optimize the results.

As we know classifiers generally perform poorly on imbalanced datasets like happened in this study. A possible future improve would be using some approaches that could minimize the problem. The main reason for the SVM algorithm to be sensitive to class imbalance is that the soft margin assigns the same cost for both positive and negative misclassifications in the penalty term. Using DEC (Different Error Cost) would cause the separating hyperplane to be skewed towards the minority class, which would finally yield a suboptimal model. Another approach that would make SVM less sensitive to outliers and noisy samples is a technique called Fuzzy SVMs. Therefore, it made sense to apply these methods and compare the performances results.

7. References

[1] Kamarudin, N. D., Ghani, K. A., Mustapha, M., Ismail, A., & Daud, N. G. N. (2014). An Overview of Crater Analyses, Tests and Various Methods of Crater Detection Algorithm. Frontiers in Environmental Engineering.

[2] Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.

[3] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886-893). IEEE

[4] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints.International journal of computer vision, 60(2), 91-110