Available online at www.sciencedirect.com

SciVerse ScienceDirect

Advances in Space Research 49 (2012) 64–74

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

Detection of sub-kilometer craters in high resolution planetary images using shape and texture features

Lourenço Bandeira [a,*], Wei Ding [b], Tomasz F. Stepinski [c]

[a] Centre for Natural Resources and the Environment, Instituto Superior Tecnico, 1049-001 Lisboa, Portugal
[b] Department of Computer Science, College of Science and Mathematics, University of Massachusetts Boston, 100 Morrissey Blvd., Boston, MA 02125-3393, USA
[c] Department of Geography, University of Cincinnati, Cincinnati, OH 45221-0131, USA

Received 10 December 2010; received in revised form 4 August 2011; accepted 22 August 2011
Available online 28 August 2011

Abstract

Counting craters is a paramount tool of planetary analysis because it provides relative dating of planetary surfaces. Dating surfaces with high spatial resolution requires counting a very large number of small, sub-kilometer size craters. Exhaustive manual surveys of such craters over extensive regions are impractical, sparking interest in designing crater detection algorithms (CDAs). As a part of our effort to design a CDA, which is robust and practical for planetary research analysis, we propose a crater detection approach that utilizes both shape and texture features to identify efficiently sub-kilometer craters in high resolution panchromatic images. First, a mathematical morphology-based shape analysis is used to identify regions in an image that may contain craters; only those regions – crater candidates – are the subject of further processing. Second, image texture features in combination with the boosting ensemble supervised learning algorithm are used to accurately classify previously identified candidates into craters and non-craters. The design of the proposed CDA is described and its performance is evaluated using a high resolution image of Mars for which sub-kilometer craters have been manually identified. The overall detection rate of the proposed CDA is 81%, the branching factor is 0.14, and the overall quality factor is 72%. This performance is a significant improvement over the previous CDA based exclusively on the shape features. The combination of performance level and computational efficiency offered by this CDA makes it attractive for practical application.

Keywords: Automatic crater detection; Pattern recognition; Craters; Mars

1. Introduction

Craters are topographic features on planetary surfaces resulting from impacts of meteoroids. They are found on all hard-surface bodies in the solar system but are most abundant on bodies such as the Moon or Mars where they can accumulate due to slow surface erosion rates. Importance of craters stems from their utility to provide relative chronology of different planetary surfaces (Wise and Minkowski, 1980; Tanaka, 1986); simply put, heavily cratered surfaces are relatively older than less cratered surfaces. Because statistics of crater sizes form the basis for geologic stratigraphy, crater counting is a routine activity in planetary science (Hartmann, 1999; Hartmann and Neukum, 2001). Presently, all crater surveys are done by means of visual inspection of images resulting in databases which are either spatially comprehensive but restricted to only the largest craters (Barlow, 1988; Rodionova et al., 2000; Andersson and Whitaker, 1982; Kozlova et al., 2001), or size comprehensive but limited to very specific geographical location. The size distribution of craters can always be approximated by the power-law (Tanaka, 1986); large craters are rare and small craters are abundant. Counts of large craters must be collected from spatially extended regions in order to accumulate sufficient number of samples for accurate statistics. Thus, geologic stratigraphy based on

* Corresponding author.
E-mail addresses: lpcbandeira@ist.utl.pt (L. Bandeira), ding@cs.umb.edu (W. Ding), stepintz@uc.edu (T.F. Stepinski).

0273-1177/$36.00 © 2011 COSPAR. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.asr.2011.08.021

manually collected databases has coarse spatial resolution. Finer spatial resolution of the stratigraphy can only be obtained from statistics of smaller craters, and the only viable means to obtain spatially comprehensive databases of smaller craters is through automating the process of crater detection.

The literature on crater detections algorithms (CDAs) is extensive, (Salamuniccar et al., 2011) tabulates 77 past publications on various CDA approaches however, no algorithm capable of becoming a standard tool for practitioners of planetary science has emerged. Crater counts continue to be done via visual inspection of images even as data sets of high resolution images keep on increasing. The lack of robust and practical CDA, that could be utilized as an "off-the-shelf" application, is due to the fact that real (as oppose to idealized) craters are difficult objects to detect in an image as they lack specific features that can reliably discriminate them from other objects (or collection of objects) also present on planetary surfaces. Craters appearance in an image depends on their level of degradation, on their internal morphologies (presence of central peaks, peak rings, central pits, wall terraces, etc.) and the degree of overlapping with other craters, on image quality (illumination angle, surface properties, atmospheric state), and on their sizes that may differ by orders of magnitude.

One challenge of a CDA is to find a set of features that best discriminates between craters and non-crater objects. Such set of features will depend on the type of surface, illumination properties, and the sizes and shapes of craters. Existing algorithms focus on detection of larger craters located on simple surfaces; such focus dictates a specific choice of image features. However, an actual need is to detect smaller, sub-kilometer size craters, which are not necessary located at smooth surfaces, and are present in an image that may have variable level of illumination. Even smaller craters, having sizes below 100 m cannot be utilized for surface dating because they erode on relatively fast time scale and/or are results of secondary impacts. Thus, in this paper we refer to craters as "small" if they have sub-kilometer size; note that such "small" craters may contain hundreds or even thousands of pixels in a high resolution image having resolution of the order of 1–10 m/pixel. An algorithm designed to detect sub-kilometer craters should be build around different set of image features than an algorithm designed to detect large craters located at a simple terrain. The need to detect small craters stems from their shear number; much less abundant, larger craters can be detected via visual inspection with substantial but not overwhelming effort, but vast quantity of small craters all but requires an automated detection approach.

Another challenge of a CDA is to utilize an optimal detection technique. Existing approaches to detection technique can be divided into two general categories: unsupervised (fully autonomous) and supervised (which require an input from a domain expert). Detailed discussion of how these two approaches are utilized in crater detection algorithms can be found in Bue and Stepinski (2007) and Stepinski et al. (2009). Unsupervised approaches rely on pattern recognition techniques to identify crater rims in an image as circular objects. Supervised approaches use machine learning techniques to train a classifier that is subsequently used to separate craters from other objects found in an image. Note that in both approaches craters are detected by a process of narrowing a set of their potential candidates. In an unsupervised approach the narrowing is achieved through thresholding a value of parameter that measures how well a candidate fits a circle. In a supervised approach the narrowing is achieved by thresholding a probability of positive detection by a classifier. Thus, a good CDA requires a well chosen set of discriminative features, an efficient method for finding all possible crater candidates, and a set of criteria that can accurately separate craters from non-craters in the set of candidates.

The purpose of this paper is to contribute toward development of a CDA especially designed toward surveying sub-kilometers size craters in high resolution planetary images. Our key insight is an observation that discriminative features utilized to obtain a set of crater candidates do not have to be the same as discriminative features utilized to extract craters from the set of candidates. Indeed, we propose to use shape features for rapid identification of a set of viable crater candidates and to use texture features for accurate classification of the candidates.

Recently it was proposed (Urbach and Stepinski, 2009) that small crater candidates can be conveniently and efficiently detected in an image because each of them contains a pair of characteristic shapes. Planets on which craters are preserved have very tenuous atmospheres, so all images are taken in direct sunlight. Under such illumination conditions, small craters, which are bowl-like topographic depressions, yield photographic imprints containing semicircular of crescent-like highlight and shadow regions. Note that larger craters, which are not bowl-shaped, show no such imprints. A detection method based on this idea utilizes mathematical morphology (Matheron, 1975; Serra, 1982) to design scale and rotation-invariant shape filters for identification of such regions. Because a single application of shape filter to an image identifies all crater candidates irrespective of their sizes (within a limit) and orientation, the shape-based method is very efficient and thus well-suited for detecting small crater candidates in large images. However, using a single pair of shapes to distinguish between true craters and other objects present in the set of candidates yields less than optimal classification results (Urbach and Stepinski, 2009). Thus, the shape-based algorithm (Urbach and Stepinski, 2009) is not the best complete crater detector, but it is a very efficient crater candidate detector.

Another recent study (Martins et al., 2009) has demonstrated that simple image texture features (Papageorgiou et al., 1998) are effective in discriminating between craters and non-crater objects when coupled to a boosting algorithm (Viola and Jones, 2004) that combines a number of most discriminative single-feature weak classifiers into a

sole, strong classifier. The boosting approach, first proposed in the context of face recognition, is effective because its learning strategy adopts well to a large variety found in images of craters. On the other hand, such CDA has no build-in strategy for finding crater candidates; the classifier is applied to an exhaustive set of image fragments making it inefficient for detecting small craters in large images. Thus, the texture-based algorithm is an accurate but inefficient crater detector. It's efficiency can be increased significantly by providing a compact set of crater candidates, which would make performing an exhaustive search unnecessary.

In this paper we present a CDA, oriented toward detection of sub-kilometer craters that uses shape features (Urbach and Stepinski, 2009) to detect efficiently a set of feasible crater candidates and utilizes texture features to classify accurately the candidates into craters and non-craters. Thus, our approach utilizes the best characteristics of the two methods. We demonstrate that such hybrid algorithm significantly outperforms an algorithm (Urbach and Stepinski, 2009) based on shape features alone – the only other algorithm designed for efficient detection of sub-kilometer craters. This work is a continuation of early tests where we verified the effectiveness of this method (Bandeira et al., 2010).

The rest of the paper is organized as follows. In Section 2 we describe our shape-based methodology for finding crater candidates and in Section 3 we describe our texture-based method for classification of the candidates into craters and non-craters. In Section 4 we introduce our test site, represented by a panchromatic, high resolution image of cratered terrain on Mars, and present results of calculations aimed at assessing the performance of our method. Conclusions and direction of future research are given in Section 5.

2. Finding crater candidates

The first phase of our sub-kilometer CDA is devoted to an efficient identification of a set of viable crater candidates. This set is to be submitted (in the second phase of our CDA) to a classifier for separation of true craters from non-crater objects. Having a dedicated algorithm for finding crater candidates improves an overall efficiency of crater detection. In principle, a pool of candidates could be created by an exhaustive search of an entire image. In such an approach candidates are generated by extracting image blocks of different sizes centered at each pixel in an image. This leads to an unnecessarily large number of image block extraction, and to a wasteful amount of evaluations at the classification phase. There exist several methods for more efficient identification of crater candidates in an image, including Bandeira et al. (2007), Salamuniccar and Loncaric (2010). The recent method (Urbach and Stepinski, 2009) based on shape features identifies sub-kilometer crater candidates in a particularly efficient manner; we utilize this method in our CDA.

Algorithm 1. Finding crater candidates using shape features

input: High resolution panchromatic image of planetary surface.
output: The list containing crater candidates (ID number of a candidate, x and y coordinates of candidate center, estimate of candidate diameter.)

1 Invert the image;
2 **for** image = original **to** inverted **do**
3 Extraction of highlight (shadow) shapes;
4 Removal of oversize shapes using median filter;
5 Removal of shapes that lack sufficient contrast using power filter (Young and Evans, 2003);
6 Removal of undersize shapes using area filter;
7 Apply shape filters to identify highlight and shadow shapes having geometries consistent with being parts of craters.
8 **end**
8 Match identified highlights and shadows shapes into crater candidates. Calculate x and y coordinates and estimate diameters of crater candidates.

The key insight of Urbach and Stepinski (2009) is an observation that the craters (crater candidates) could be recognized in an image as a pair of crescent-like highlight and shadow regions or shapes; hence we refer to their method as based on shape features. Algorithm 1 identifies high level steps needed to translate that insight into actual calculation. For detailed description of the shape-based algorithm we refer the reader to the original work (Urbach and Stepinski, 2009); here we provide only a short summary of this technique. The input is a panchromatic image which contains many highlight and shadow shapes. The algorithm identifies and processes highlight and shadow shapes simultaneously using inverted image to identify and process the shadow shapes (line 3 in the Algorithm 1). The goal of the chain of processes (lines 4 to 7 in the Algorithm 1) is to eliminate all shapes that are not indicative of craters. The shape processing utilizes extensively attribute filters (Breen and Jones, 1996); these filters use criteria evaluated using numerical attributes of shapes to decide whether to eliminate or keep any given shape. The efficiency of the shape-based approach stems from translational, rotational, and scale (only shape filters) invariance of attributes used in the filters. Thus, a single application of any filter to an image removes all undesired shapes. In the final step (line 9 in the Algorithm 1), the remaining highlight and shadow shapes are matched to each other to identify crater candidates; for each candidate coordinates of its center and the estimate of its diameter are calculated. The shape-based method was originally proposed (Urbach and Stepinski, 2009) as a complete CDA,

however, the set of crater candidates identified by Algorithm 1 has a high number of false positive detections. In Urbach and Stepinski (2009) the number of false positives was reduced by classifying the output of Algorithm 1 using additional features pertaining to relationships between a pair of (highlight and shadow) shapes constituting each crater candidate. We submit that subjecting the output of Algorithm 1 to a classifier based on texture features rather than additional shape features results in a more accurate CDA.

3. Detection of craters

Craters are detected from amongst the set of crater candidates (the output of Algorithm 1) on the basis of image texture features using a variation of the method proposed in Martins et al. (2009). The method has two components. First, an overcomplete set of texture features is constructed to represent an image block containing each crater candidate. Second, a simple and efficient classifier/feature selector is built using the AdaBoost learning algorithm to select features that discriminate best between craters and non crater objects, and to assign probabilities of being a crater to each candidate. Although a version of this algorithm is available at Open Computer Vision library (http://opencv.willowgarage.com/wiki/FaceDetection), we decided to implement our own version in Matlab, easier to integrate with previous work.

3.1. Image preprocessing

In a preprocessing step an image is subjected to histogram equalization (Press, 2007) transformation. The purpose of such preprocessing is two-fold: first, it increases contrast and make detection of craters easier; second, it makes many different images, each having different level of illumination and contrast, look similar thus minimizing the need for establishing separate training set for each individual image. Subsequently, we apply a median filter (with a 3×3 kernel) to smooth the noise introduced by the equalization.

3.2. Texture features

For each crater candidate a square image block, centered on the location of the candidate and having a dimension twice the diameter of the candidate, is extracted from an overall image representing the entire scene. Examples of such image blocks, representing candidates having different sizes, are shown in Fig. 1. Image blocks stemming from all identified crater candidates must be encoded in terms of texture features. Texture feature extraction algorithms can be grouped into statistical, geometrical, model based, and signal processing based (Tuceryan and Jain, 1998). From the variety of possible algorithms we have chosen a simple geometrical method of texture features extraction first proposed in Papageorgiou et al. (1998), and popularized in the context of face recognition by Viola and Jones (2004). Such texture features are broadly utilized for object detection (for example, face detection) and have proven to work well for crater detection (Martins et al., 2009). Numerical values of the features are calculated using a family of binary masks. Each mask is a rectangular region subdivided into two kinds of sectors (referred to as black and white sectors for specificity). The value of a particular mask-feature is obtained by placing the mask over a selected part of an image block (representing a crater candidate) and subtracting the sum of grayscale values of the pixels covered by black sectors of the mask from the sum of grayscale values of the pixels covered by white sectors of the mask. By changing the size of the mask, an arrangement of its sectors, and its placement within an image block, a large set of mask-features is assembled. This set provides a rich, overcomplete representation of crater candidates and forms a basis for their classification into two classes of objects: true craters and non-crater objects. An overcompleted representation of objects is desirable because it helps to accommodate intra-class variability without compromising inter-class discriminability.

Because we aim at detecting round craters characterized by large scale shadows and highlights we decided to use only square mask-features in order to significantly reduce the overall number of features. Note that in the context of face detection, and even in the previous application of

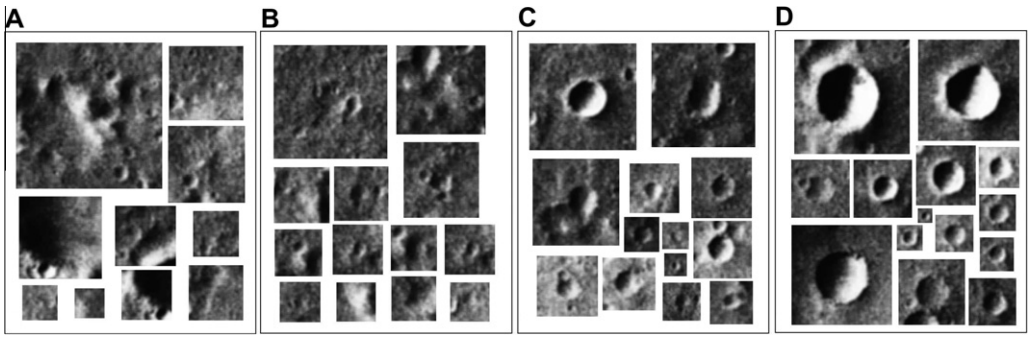

Fig. 1. Examples of image blocks representing crater candidates C; each block is twice the size of the candidate. Classification of the candidates (see Section 3) assigns them probabilities, $H(C)$, of being a crater; (A) $0 < H(C) \leqslant 0.4$, (B) $0.4 < H(C) \leqslant 0.6$, (C) $0.6 < H(C) \leqslant 0.8$, (D) $0.8 < H(C) \leqslant 1.0$.

Fig. 2. Nine types of masks used for extraction of texture features.

mask-features to crater detection (Martins et al., 2009), a larger set of rectangular mask-features was used. In the case of face detection rectangular features are justified by existence of rectangular face components (nose, eyes), but in the context of crater detection rectangular features are not necessary. In our implementation of texture features, each crater candidate image block is resized (using a bilinear interpolation method) from its original size to a standardized size of 48×48 pixels. We use mask-features in four different sizes, 12×12, 24×24, 36×36 or 48×48 pixels, respectively. In each size nine different patterns of black and white sectors are used (see Fig. 1). These patterns are chosen to detect edges of the crater as well as the shadows/highlight line across the crater (see Fig. 1 for examples of image blocks containing craters). Finally, each mask is applied in multiple spots located all across an image block and separated from each other by a third of mask size. Overall, we represent each crater candidate (its corresponding image block) by 1089 features. Each of the nine mask pattern types (see Fig. 2) results in 121 features; 100 for 12×12 pixels size masks, 16 for 24×24 pixels size masks, 4 for 36×36 pixels size masks, and 1 for 48×48 pixels size masks. The features are linearly normalized to the range of $(-1, 1)$ to eliminate dependence on size of the mask. Efficient calculation of all features is achieved through so-called "integral image representation" (Viola and Jones, 2004).

3.3. Feature selection and classification

Given a set of 1089 texture-based features and a training set containing craters and non-crater objects, a number of machine learning approaches (for example, decision tree, support vector machines, maximum likelihood, or neural networks) could be used to derive a binary (crater vs. non-crater) classifier. However, the purpose of having an overcompleted set of features is to use a classifier that, during its training phase, automatically selects only those features that are most useful in discriminating between crater and non-craters. Note that whether any particular feature is discriminative or not is not known a priori and feature selection is indeed a part of the learning process. Following Martins et al. (2009) we use a variant (Viola and Jones, 2004) of AdaBoost algorithm (Freund and Schapire, 1997) that simultaneously selects the best features and trains the classifier.

In general, the AdaBoost selects a limited subset from a large pool of simple, rather poorly performing learning algorithms (called weak classifiers) and combines the selected algorithms into a strong, well-performing classifier. The variant of the AdaBoost algorithm used in this paper identifies weak classifiers with texture features. For each of our 1089 texture features a simple classifier in the form of a single node decision tree (a decision stump) is constructed

$$h(C; f, p, \theta) = \begin{cases} 1 & \text{if } p \cdot f(C) \geqslant p \cdot \theta \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where C is an image block representing a crater candidate and $f(C)$ is the numerical value of a specific texture feature applied to C. The weak, feature-specific classifier $h(C)$ is parameterized by a threshold θ and a polarity (direction of inequality) $p = \{1, -1\}$ and yields values of 1 (crater) or 0 (non-crater). Values of the parameters are set so the minimum number of crater candidates in the training set is misclassified. Each weak classifier by itself is expected to perform poorly because no single texture-based feature has sufficiently high discriminative power. However, different features reflect different aspects of discrimination between craters and non-craters, so the combination of many weak classifiers is expected to perform well.

Algorithm 2. Building the strong, texture-based classifier through boosting

input: Training set of crater candidates $(C_1, y_1), \ldots, (C_n, y_n)$, where C_i is the ith crater candidate image block and $y_i = 0, 1$ for craters and non-craters, respectively.
output: The final strong classifier that assigns each C_i a probability of being a true crater.

1 Initialize weights $w_{1,i} = (1/2m)$, $(1/2l)$ for $y_i = 0, 1$, where m and l are number of carters and non-craters, respectively in the training set;
2 **for** $t = 1$ **to** T, where T is a desired number of most discriminative weak classifiers **do**
3 Normalize the weights, $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}$;
4 Select the weak classifier that minimizes the weighted error $\epsilon_t = \min_{f,p,\theta} \sum_i w_{t,i} |h(C_i, f, p, \theta) - y_i|$;
5 Define $h_t(C) = h(C, f_t, p_t, \theta_t)$, where f_t, p_t, and θ_t are the minimizers of ϵ_t;
6 Update the weights, $w_{t+1,i} = w_{t,i} \beta^{1-e_i}$, where $e_i = 0$ if C_i is classified correctly, and $e_i = 1$ if C_i is classified incorrectly; $\beta_t = \epsilon_t / (1 - \epsilon_t)$.
7 **end**
8 The final strong classifier is given by

$$H(C) = \begin{cases} \text{crater} & \text{if } \frac{\sum_{t=1}^{T} \alpha_t \cdot h_t(C)}{\sum_{t=1}^{T} \alpha_t} > \mu \\ \text{non-crater} & \text{otherwise} \end{cases}$$

where $\alpha_t = \log(1/\beta_t)$ and μ is a threshold probability.

Algorithm 2 summarizes a variant of the AdaBoost procedure we utilized for building the strong classifier. The algorithm requires a training set of n image blocks depicting both craters and non-craters. It selects T most

discriminative weak classifiers from the total available pool of 1089 such classifiers. In order to achieve this goal the algorithm goes through T "rounds" of weak classifier selection (the main loop spanning the lines 2–7 in the Algorithm 2). In each round the weak classifier that minimizes the weighted error (line 4 in the Algorithm 2) is selected. The algorithm also assigns an importance, α_t to the selected weak classifier; providing that $\epsilon_t \leqslant 0.5$, the α_t gets larger as ϵ_t gets smaller. All training examples are assigned the weights (line 1 of the Algorithm 2). These weights are recalculated at the end of each round (line 6 in the Algorithm 2) in such a way as to *decrease* the weights of training examples correctly identified in the present round thus forcing the algorithm to focus on other, "harder" examples in the next round. Once a desired number of rounds are performed, the strong classifier is constructed from the T selected weak classifiers and their importance factors (line 8 in the Algorithm 2). The strong classifier, $H(C)$, depends on the user-defined value of probability threshold $\mu \leqslant 1$. Choosing a large value of μ minimizes the value of false detection percentage (FDP) but results in a relatively small value of true detection percentage (TDP), whereas choosing a smaller value of μ results in increased values of both TDP and FDP.

3.4. Post-processing

Our shape-based methodology allows for multiple candidates, classified as "craters," to correspond to a single crater. For this, we use elimination formulas consistent with earlier work on morphology based algorithm (Bandeira et al., 2010; Ding et al., 2011), where crater-labeled candidates fulfilling the following criteria are taken to correspond to a single crater and grouped together:

$$\frac{|d_i - d_j|}{max(d_i, d_j)} \leqslant \alpha \quad \text{and} \quad \frac{dist(p_i, p_j)}{max(d_i, d_j)} \leqslant \gamma \tag{2}$$

here, candidates i and $j \neq i$ have diameters (positions) $d_i(p_i)$ and $d_j(p_j)$, respectively. We have determined experimentally the best choices for free parameters $\alpha = \gamma = 0.5$. The candidate having the highest value of $H(C)$ within the group is selected as a positive "identification" of a given crater; the remaining candidates within that group are discarded without becoming false negatives.

4. Test site results

4.1. Dataset and experimental setup

In order to test the performance of our CDA we use a portion of the nadir panchromatic, 12.5 m/pixel image (h0905) of Mars taken by the High Resolution Stereo Camera (HRSC) on-board the Mars Express spacecraft. The selected site is centered on Nanedi Valles – a prominent geologic feature; it extends ~40 km (north-south) by ~59 km (east–west) and covers ~2360 km^2 of heavily cratered terrain. In this site an analyst has manually cataloged 3050 craters having diameters between 40 and 6600 m. The site and its cataloged craters are shown in Fig. 3, available for download at http://cerena.ist.utl.pt/lpcbandeira/downloads.html, also with the detection results obtained. The Nanedi Valles passes through the middle of the site

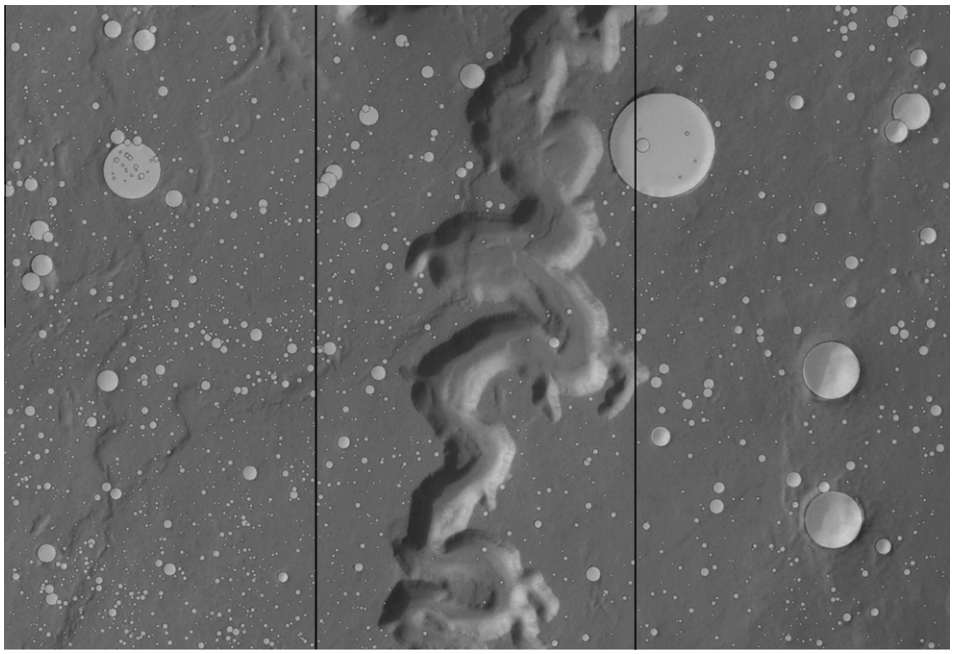

Fig. 3. Image of test site located in the region of Nanedi Valles on Mars, between longitudes of −47°.66′E/ − 48°.68′E, and latitudes of 7°.28′N/7°.95′N (from the top-left corner: X-Offset = 1500, Y-Offset = 36000, Width = 4700, Height = 3200). The 3050 craters manually identified by an analyst are shown as lighter gray circles with sizes proportional to the diameters of the craters. Vertical lines indicate boundaries between West, Central, and East sections.

introducing heterogeneity of the terrain. In order to account for this heterogeneity in the evaluation of our CDA, three sections, labeled West, Central (containing Nanedi Valles) and East, have been defined (see Fig. 3). We recognize that our algorithm may not be able to detect smallest craters identify by an analyst. Thus, our goal is not to compare the result of detection to the entire manual catalog, but rather to determine the smallest size of a crater that can be reliably detected by the algorithm. This size is determined by the desire detection rate; we aim at detection rate of about 80%.

Applying Algorithm 1 to our test site results in identification of 14004 crater candidates. We have chosen $n = 633$ candidates, all located in the northern half of the east section of the test site, to constitute the training set; 211 of them are true craters and 422 are non-crater objects. The false examples were randomly picked out from amongst the candidates not overlaying any crater in the ground truth catalog. Note that the training set contains only 7% of the craters in the entire site and is sampled from approximately 20% of the area of the site. Thus, our setup corresponds to a likely use of the CDA, wherein a user wants to train the CDA on a small image and use the CDA to find craters in a larger image. In the training phase we have applied the Algorithm 2 to the training set using all features ($T = 1089$). This is because we have no a priori knowledge about a number of features that are sufficient for adequate performance. However, the AdaBoost algorithm ranks the features by their importance factor α_t providing a tool for establishing the minimum required number of features for future application. Fig. 4 shows the features selected by the algorithm in the first six rounds (top six features). The selected features (except for the feature # 5) appear to focus on detecting a boundary between the shadow and the highlight portions of a crater. Note that the features were selected on the basis of the entire training set, but are shown on Fig. 4 over a single, particularly regular crater. Thus, we cannot fully interpret the selection logic without examining the entire training set.

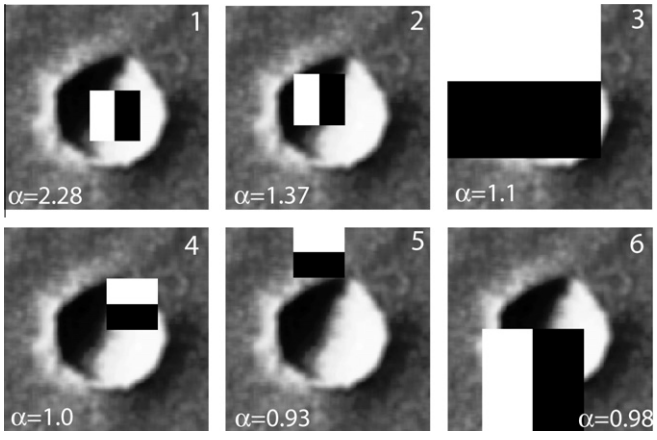

Fig. 4. The top six mask-features selected with AdaBoost using the training set as described in the text. All six mask-features are shown on top of the same gray scale image block showing a well-defined crater. A crater is a depression in the surface and appears in an image as a pair of shadow and highlight semi-circular shapes. An illumination is from the north-west.

4.2. Performance evaluation methodology

By running a series of calculations with progressively larger lower limit on the size of allowable craters we have determined that only craters larger than 16 pixels (200 m in the h0905 image) can be reliably identified in our test site (detection rate of about 80%). Thus, our algorithm, which requires a certain number of pixels to determine a shape and the texture of an image block, cannot detect craters as small as the human eye can. Nevertheless, the lower limit of 200 m is quite sufficient from practical point of view (see Section 1). In addition, the overall size of an image puts an upper limit on the size of identifiable crater because of paucity of larger craters for inclusion in the training set. We have determined that only craters smaller than 400 pixels (5000 m in the h0905 image) can be reliably identified in our test site. Again, this is not a serious limitation because such craters are very rare (in an area considered here) and can be easily counted manually. Thus, in our test image, we compare the results of our CDA to an identifiable subset of the craters cataloged by an analyst that consists of the craters in the range of 16 pixels < diameter < 400 pixels. There are 1937 such craters; hereafter we refer to them as the ground truth.

To evaluate the performance of our method we use quality factors (Shufelt, 1999): true detection percentage $D = 100 \cdot TP/(TP + FN)$, branching factor $B = FP/TP$, and quality percentage $Q = 100 \cdot TP/(TP + FP + FN)$. Here, TP stands for the number of true positive detections (detected craters that are actual craters), FP stands for the number of false positive detections (detected craters that are not), and FN stands for the number of false negative "detections" (non-detection of real craters). For completeness, TN is the number of true negatives (non-craters correctly identified as such). D can be treated as a measure of crater-detection performance (larger is better), B as a measure of delineation performance (smaller is better), and Q as an overall measure of algorithm performance (larger is better).

Determining "detection" of craters is nontrivial as the craters constituting the ground truth catalog and the crater candidates labeled by our classifier are not expected to correspond exactly to each other. We use the following criteria to match a labeled candidate to a crater in the ground truth catalog (gt):

$$\frac{|d_{gt} - d_i|}{max(d_{gt}, d_i)} \leqslant \beta, \quad \frac{dist(p_{gt}, p_i)}{max(d_{gt}, d_i)} \leqslant 1, \quad dist(p_{gt}, p_i) \leqslant \delta \qquad (3)$$

Values of β or δ determine desired strictness of the match between a candidate and the ground truth crater; we have determined experimentally their best values, $\beta = 0.5$ and $\delta = 26$ pixels.

4.3. Results

With the values of TP, FP and FN calculated according to a description given above we proceed to establish a minimum number of features required to yield a satisfactory classifier. We calculated receiver operating characteristic (ROC) curves (Fawcett, 2004) corresponding to crater detection classifiers, $H(C)$, indexed by a number of features they use; classifiers using 10, 100, 150, 200, and 1089 top features were evaluated. The ROC curve is a parametric curve that compares two characteristics of the classifier, true positive rate $TPR = TP/(TP + FN)$ (the same as D but expressed as a fraction) and false positive rate $FPR = FP/(FP + TP)$ as the probability threshold μ changes. It depicts changes of relative trade-off between benefit (TPR) and cost (FPR) with μ. Fig. 5 shows the inverted ROC curves for considered classifiers. The curves for 100, 150, 200, and 1089 features are identical on the scale of Fig. 5 and are represented by a single curve labeled $T = 1089$. We use an inverted ROC curve (TPR on the x-axis and FPR on the y-axis) in order to help the prospective user (likely to be a planetary scientist) to chose a desired value of TPR and thus the probability threshold μ. Note that in order to make Fig. 5 more easy to view we did not indicate the changes of μ along the curves but this data is available.

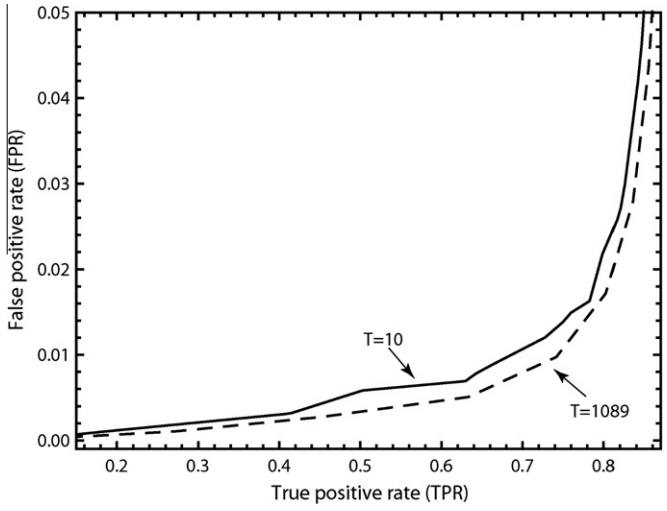

Fig. 5. Inverted receiver operating characteristic (ROC) curves for the strong crater detection classifier $H(C)$ constructed using $T = 10$ top features (solid line) and 100, 150, 200, and 1089 features (dashed line).

Comparison of the ROC curves depicted on Fig. 5 indicates that a classifier built with 100 top features is sufficient; inclusion of additional features does not improve noticeably detection performance over any range of values of μ. A classifier built with as little as 10 top features performs well but worse than the one built with 100 top features. Therefore we conclude that it is sufficient to construct a classifier using only 100 (out of 1089) top features to achieve optimal performance; if speed of the classification is an important factor, a classifier with only 10 top features can also be considered. After deciding on the number of features used to built a classifier, the next step is to select a desired value of TPR which fixes a value of the probability threshold μ. It is important to note that a variability of μ along the curve depends on the classifier. For example, for the $T = 10$ curve the range is from $\mu = 0.87$ at the low-TPR/low-FPR end of the curve to $\mu = 0.4$ at the high-TPR/high-FPR end of the curve, whereas for the $T = 1089$ curve the range is between $\mu = 0.65$ and $\mu = 0.43$.

For a more in-depth evaluation of our CDA, and to compare the performance of our hybrid CDA with a performance of a CDA based on shape features alone we chose a classifier built using 100 top features. Aiming at TPR of about 0.8 we choose $\mu = 0.525$ as read from the ROC curve annotated with the values of μ. Table 1 lists values of quality factors D, Q, and B for the entire test site and, separately, for its three sections (West, Central, and East). For comparison, it also lists the values of the same quality factors calculated for the CDA (Urbach and Stepinski, 2009) based exclusively on shape features. The lower section of the table shows improvement in the values of performance metrics calculated as {value(shape/texture)-value(shape)}/value(shape).

Over the entire site our CDA achieves detection rate $D = 80.7\%$ with branching factor $B = 0.136$ for the overall

Table 1
Performance of our shape/texture-based CDA vs. the performance of the CDA based exclusively on shape features. D stands for true detection percentage; Q for quality percentage and B for branching factor.

Type	Entire test site		West section		Central section		East section	
	Shape/texture	Shape only	Shape/texture	Shape only	Shape/texture	Shape only	Shape/texture	Shape only
Quality factors								
D (%)	80.7	63.53	79.1	59.37	78.3	61.52	86.9	75.48
B	0.136	0.248	0.115	0.259	0.243	0.247	0.0923	0.229
Q (%)	72.3	54.9	72.5	51.5	65.8	53.4	80.5	64.3
Type	All regions		West region		Central region		East region	
Improvement in classification performance								
Improv. D (%)	27.0		33.2		27.3		15.1	
Improv. B (%)	45.0		55.6		1.6		59.7	
Improv. Q (%)	31.7		40.8		23.2		25.2	

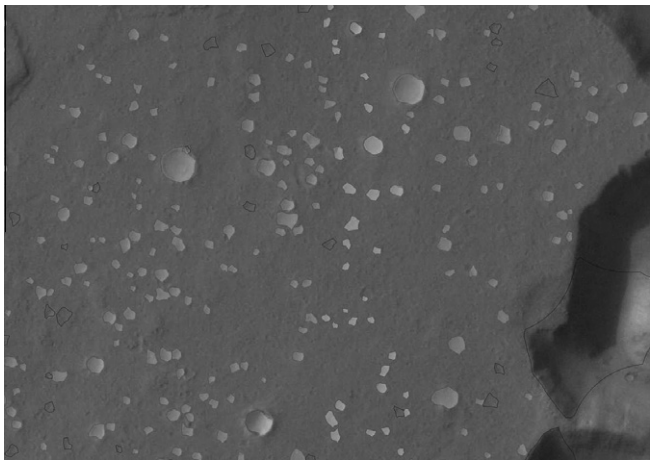

Fig. 6. Detail of the test site showing the craters detected by our algorithm. Shapes in gray are TP and outlined shapes are FP.

quality rate of $Q = 72.3\%$. This is a good performance for the CDA devoted to finding sub-kilometer craters; this level of performance is acceptable for practical planetary research. For illustrative purposes, in Fig. 6 we show some of these detections on a small portion of the test site. Note that our CDA performs best in the East section of an image which contains the training set. The performance drops somewhat in the West section of the image, which does not contribute samples to the training set, but is characterized by a terrain similar to that found in the East section. The performance drops further in the Central section of the image, which also does not contribute samples to the training set, but its terrain has markedly different character from the rest of the image. Because the Central region contains crater candidates having character unaccounted for in the training set the number of false positives and false negatives in this region is higher than in the rest of an image. We could remedy this drop in performance by selecting candidates from the Central region to the training set. However, this would be contrary to our overall approach to testing carter detection methods in accordance with how they are expected to be used by planetary scientists. In planetary application, the user would like to utilize already available catalogs of craters as the training set. Such catalogs are the results of exhaustive surveys over small sites rather than casual surveys over large sites. Thus, the CDA should be able to generalize the knowledge from a small site to the much larger region. If the larger region is inhomogeneous with respect to appearance of the craters, the performance of the CDA is expected to be spatially variable – exactly what we have found out in our test site. See below, in the conclusions section, for how we propose to address this issue in future research.

Table 1 clearly demonstrates superior performance of our new CDA based on a combination of shape features (to identify the crater candidates) and texture features (to classify the candidates) over the CDA (described in Urbach and Stepinski, 2009) based solely on shape features. The last section of Table 1 shows the improvement in classification performance that are gained by incorporating texture into the CDA. All quality factors are improved by double digits percentages, but the branching factor has improved the most (except for the Central section of the image) indicating that texture features significantly reduced the number of false positive detections.

A similar, possibly slightly higher, accuracy could be achieved by a classifier using only the texture features and relying on exhaustive search to generate blindly crater candidates. Because the candidate set, as identified by Algorithm 1, includes vast majority of identifiable craters but not all of them, texture-based algorithm with exhaustive search is expected to yield slightly higher accuracy. However, an algorithm based on exhaustive search is vastly slower. For image having size $imgWidth \times imgHeight$ pixels, the number of square sub-windows having size $subSize \times subSize$ is

$$n(subSize) = \left(\frac{imgWidth - subSize}{shift}\right)\left(\frac{imgHight - subSize}{shift}\right) \quad (4)$$

where $shift$ is the number of pixels by which the window is moved to its new position; we assume that $shift$ is the same in horizontal and vertical directions. For our test site $imgWidth = 4700$ and $imgHeight = 3200$. For compatibility with our calculations we assume that sub-windows are generated starting from $subSize = 32$ pixels and increase in size by the factor of 1.5 until they reach $subSize = 546$ pixels. We further assume that $shift = 0.03subSize$. Under such conservative assumptions, the classifier relying on exhaustive search needs to classify over 28 millions sub-windows as craters candidates. This is over 2000 times more than 14004 evaluation performed by our proposed method. Using a 3 GHz PC the computational cost of classifying 14004 candidates is 84 s using a classifier based on the 100 top features. There is an overhead of 866 s for the Algorithm 1 to identify the candidates in the test site. The algorithm based on exhaustive search has also some small overhead associated with generation of sub-windows, but its major computational cost is the 47 h that are needed to classify all 28 millions of "candidates." Considering that our test site constitutes only 1/44 of the size of the entire HRSC image h0905, and that hundreds of such images are ready to be utilized, the two orders of magnitude speed gain of our approach over the exhaustive search approach makes the large scale identification of small craters practical. Using faster computers and more advanced versions of AdaBoost classifier (Lienhart et al., 2003) would significantly reduce classification time, however, our method would maintain a built-in advantage of having to evaluate orders of magnitude smaller set of candidates in exchange for a small overhead associated with their identification using Algorithm 1. In parallel, we are exploring on how to adapt to new surfaces with the training done in previous regions using techniques of transfer learning (Ding et al., 2011).

5. Conclusions

In this paper we presented a new, hybrid algorithm for detection of sub-kilometer craters in high resolution panchromatic image. The presented algorithm combine the efficiency of shape analysis in identifying crater candidates with precision of texture features/AdaBoost classifier in distinguishing between craters and non-craters. The application of the new CDA to a test image confirms our design assumptions; the new algorithm identifies sub-kilometers craters with much better precision than an algorithm based on shape features alone. It is also much faster than an algorithm based on texture features alone that employs an exhaustive search. The training requirements are not overwhelming (only 7% of the candidates are utilized for training) and the performance of the new algorithm is sufficiently high ($D = 81\%$, $B = 0.14$, $Q = 72\%$) for its utilization in research-oriented crater counting application. Future research, aimed at further improvement of performance and robustness of the algorithm, will focus on several shortcomings of the present algorithm and lessons learned in the process of its development.

First, the weakest link (from the accuracy point of view) of the present algorithm is the shape-based identification of crater candidates. Of the 1937 identifiable ground truth craters, 201 are not covered by any crater candidate thus becoming false negatives by "default." With all ground truth craters covered by candidates the detection rate of our algorithm could be improved to up to 91% using the same training set. Thus, further research should concentrate on improvements to the shape-based identification of the candidates. One way to increase the number of candidates is to relax the shape filter criteria. However, since the utility of the shape-based identification of the candidates is its computational efficiency, this efficiency should not be sacrificed by over-relaxed shape criteria.

Second, we have demonstrated on our test image, that, in the realistic scenario of acquiring samples for the training set from spatially limited region, the performance of the CDA will depend on the similarity of the crater candidates to which it is applied to the candidates constituting the training set. This is a basic property of supervised learning not limited to our specific classifier. Future research should incorporate techniques of transfer learning (see, for example, Dai et al., 2007) in order to allow the user to modify (with minimum necessary effort) the training set to take into account changing character of craters at different images. The focus of such effort should be on intelligent selection of new samples that exemplify differences between existing training set and the character of new candidates.

Finally, our texture-based features are not rotationally invariant. Applying our CDA with its present training set to an image taken at different solar azimuth angle requires either rotating the new image, or rotating the image blocks constituting the training set. The method, in its present form cannot be applied to a mosaic of images taken at different solar azimuth angles. Future research should address development of rotationally invariant texture features.

Acknowledgments

This work was supported by NASA under Grant NNX11AC13G.

References

Andersson, L.B., Whitaker, B.A. Nasa catalogue of lunar nomenclature, in: NASA Reference Publication 1097, 1982.

Bandeira, L., Saraiva, Jose, Pina, Pedro Impact crater recognition on Mars based on a probability volume created by template matching. IEEE T. Geosci. Remote 45 (12), 4008–4015, 2007.

Bandeira, L., Ding, W., Stepinski, T.F. Automatic detection of sub-km craters using shape and texture information. Lunar Planetary Sci. XLI (CD-ROM), Abs. 1144, 2010.

Barlow, N.G. Crater size-frequency distributions and a revised Martian relative chronology. Icarus 75 (2), 285–305, 1988.

Breen, E.J., Jones, R. Attribute openings, thinnings and granulometries. Comput. Vis. Image Und. 64, 604–614, 1996.

Bue, B.D., Stepinski, T.F. Machine detection of Martian impact craters from digital topography data. IEEE T. Geosci. Remote 45 (1), 265–274, 2007.

Dai, W., Yang, Q., Xue, G.-R., Yu, Y. Boosting for transfer learning, in: ICML '07: Proceedings of the 24th International Conference on Machine learning, pp. 193–200, 2007.

Ding, W., Stepinski, T., Mu, Y., Bandeira, L., Vilalta, R., Wu, Y., Lu, Z., Cao, T., Wu, X. Sub-kilometer crater discovery with boosting and transfer learning. ACM Trans. Intell. Syst. Technol. 2 (4), 39, 2011.

Fawcett, T. Roc graphs: Notes and practical considerations for researchers. Technical report, Tech Report HPL-2003-4, HP Laboratories, 2004.

Freund, Y., Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55 (1), 119–139, 1997.

Hartmann, W.K. Martian cratering vi. crater count isochrons and evidence for recent volcanism from Mars global surveyor. Meteor. Planet. Sci. 34, 167–177, 1999.

Hartmann, W.K., Neukum, G. Cratering chronology and evolution of Mars, in: Altwegg, K., Ehrenfreund, P., Geiss, J., Huebner, W.F. (Eds.), Composition and Origin of Cometary Materials. Kluwer Academic Publishers, The Netherlands, pp. 165–194, 2001.

Kozlova, E.A., Michael, G.G., Rodinova, J.F., Shevchenko, V.V. Compilation and preliminary analysis of a catalogue of craters of mercury, in: Lunar and Planetary Science XXXII (CD-ROM), Abs. 1231, 2001.

Lienhart, R., Kuranov, A., Pisarevsky, V. Empirical analysis of detection cascades of boosted classifiers for rapid object detection, in: DAGM'03, 25th Pattern Recognition Symposium, Madgeburg, Germany, pp. 297–304, 2003.

Martins, R., Pina, P., Marques, J.S., Silveira, M. Crater detection by a boosting approach. IEEE Geosci. Remote Sensing Lett. 6 (1), 127–131, 2009.

Matheron, G. Random Sets and Integral Geometry. John Wiley, 1975.

Papageorgiou, C.P., Oren, M., Poggio, T. A general framework for object detection, in: Sixth International Conference on Computer Vision, Bombay, India, pp. 555–562, 1998.

Press, W.K. Digital Image Processing: PIKS Scientific Inside. Wiley Interscience, New York, 2007.

Rodionova, F.J., Dekchtyareva, K.I., Khramchikhin, A.A., Michael, G.G., Ajukov, S.V., Pugacheva, S.G., Shevchenko, V.V. Morphological catalogue of the craters of Mars, in: ESA-ESTEC, 2000.

Salamuniccar, G., Loncaric, S. Method for crater detection from Martian digital topography data using gradient value/orientation, morphometry, vote analysis, slip tuning, and calibration. IEEE T. Geosci. Remote 48 (5), 2317–2329, 10.1109/TGRS.2009.2037750, 2010.

Salamuniccar, G., Loncaric, S., Pina, P., Bandeira, L., Saraiva, J. MA130301GT catalogue of Martian impact craters and advanced evaluation of crater detection algorithms using diverse topography and image datasets. Planet. Space Sci. 59 (1), 111–131, 2011.

Serra, J. Image Analysis and Mathematical Morphology, 2nd ed Academic Press, New York, vol. 1, 1982.

Shufelt, J.A. Performance evaluation and analysis of monocular building extraction from aerial imagery. IEEE Trans. Pattern Anal. Mach. Intell. 21 (4), 311–326, 1999.

Stepinski, T.F., Mendenhall, M.P., Bue, B.D. Machine cataloging of impact craters on Mars. Icarus 203 (1), 77–87, 2009.

Tanaka, K.L. The stratigraphy of Mars. J. Geophys. Res. 91 (B13), 139–158, 1986.

Tuceryan, M., Jain, A.K. Texture analysis, The Handbook of Pattern Recognition and Computer Vision, 2nd ed., pp. 207–248, 1998.

Urbach, Erik R., Stepinski, Tomasz F. Automatic detection of sub-km craters in high resolution planetary images. Planet. Space Sci. 57 (7), 880–887, 2009.

Viola, P., Jones, M. Robust real-time face detection. Int. J. Comput. Vis. 57 (2), 137–154, 2004.

Wise, D.U., Minkowski, G. Dating methodology of small, homogeneous crater populations applied to the tempe-utopia trough region on Mars, in: NASA Tech. Memo 81 776, 122-124, NASA Goddard Space Flight Center, Greenbelt, MD, 1980.

Young, N., Evans, A.N. Psychovisually tuned attribute operators for pre-processing digital video. IEEE Proc. Vis. Image Signal Process. 150 (5), 277–286, 2003.