Detecting Impact Craters in Planetary Images Using Machine Learning

T. F. Stepinski[1], Wei Ding,[2] R. Vilalta[3]
[1]Dept. of Geography, Univ. of Cincinnati, OH 45221, USA. [2]Dept. of Computer Science, Univ. of Massachusetts Boston, 100 Morrissey Blvd. Boston, MA 02125-339, USA. [3]Dept. of Computer Science, University of Houston, 4800 Calhoun Rd., Houston, TX 77204, USA.

ABSTRACT

Prompted by crater counts as the only available tool for measuring remotely the relative ages of geologic formations on planets, advances in remote sensing have produced a very large database of high resolution planetary images, opening up an opportunity to survey much more numerous small craters improving the spatial and temporal resolution of stratigraphy. Automating the process of crater detection is key to generate comprehensive surveys of smaller craters. Here we discuss two supervised machine learning techniques for crater detection algorithms (CDA): identification of craters from digital elevation models (also known as range images), and identification of craters from panchromatic images. We present applications of both techniques and demonstrate how such automated analysis has produced new knowledge about planet Mars.

INTRODUCTION

Impact craters are structures formed by collisions of meteoroids with planetary surfaces. They are common features on all hard-surface bodies in the Solar System, but are most abundant on bodies such as the Moon, Mercury, or Mars where they can accumulate over geologically long times due to slow surface erosion rates. The importance of impact craters stems from the wealth of information that detailed analysis of their distributions and morphology can bring forth. In particular, in the absence of in situ measurements, crater counting is the only technique for establishing relative chronology of different planetary surfaces (Wise & Minkowski, 1980) (Tanaka, 1986). Simply put, heavily cratered surfaces are relatively older than less cratered surfaces. Thus, surveying impact craters is one of the most fundamental tools of planetary geology science (Hartmann, Martian cratering VI: Crater count isochrons and evidence for recent volcanism from Mars Global Surveyor, 1999) (Hartmann & Neukum, 2001).

Presently, all such surveys are done manually via visual inspection of images. Manually compiled databases of craters are either spatially comprehensive, but restricted to only the largest craters (Barlow, 1988) (Rodionova, et al., 2000) (Andersson & Whitaker, 1982) (Kozlova, Michael, Rodinova, & Shevchenko, 2001), or size comprehensive, but limited to only narrowly defined geographical locations. Using spatially comprehensive catalogs of only the largest craters allows for establishing relative chronology on large spatial scale and coarse temporal resolution. This is because large craters are rare, so their counts must be collected from spatially extended regions in order to accumulate a sufficient number of samples for accurate statistics (cumulative distribution of crater sizes is well approximated by a power law with index equal to -2). A finer spatial resolution of the stratigraphy can only be obtained from statistics of much more numerous smaller craters. Compiling global or regional catalogs of small craters, however, would be a very laborious process, ill-suited for the standard technique of manual visual detection.

Advances in gathering planetary data by space probes has resulted in a deluge of high resolution images that show craters as small as 100 m in diameter, and can be combined into mosaics covering entire surfaces of planets such as Mars, the Moon, and soon the Mercury. It is now clear that, if left to manual surveys, the fraction of cataloged craters to the craters actually present in the available and forthcoming imagery data will continue to drop precipitously. Progress in measuring surface relative chronology with increasing spatial and temporal accuracies can only be achieved by automating the process of crater surveying. Because of the importance of craters to the field of planetary science, there have been numerous attempts to develop a "crater detection algorithm" (CDA). Despite a large body of work, practitioners of planetary science continue to count craters manually, resulting in a lack of progress (relative to available data) in improving the surface chronology. This is because most approaches to CDA are restricted to demonstration that a particular algorithm achieves high accuracy on a particular image or set of images containing relatively simple "textbook" craters, whereas practitioners of planetary science are looking for a robust algorithm having a decent performance on all possible surfaces. In reality, craters are rarely simply circular on a relatively uniform background; craters appearance in an image depends on their level of degradation, on their internal morphologies, on 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. Thus the construction of a robust and practical CDA stands as a significant challenge to the scientific community.

Because of a large variety of crater forms, as well as diversity of backgrounds, we contend that a CDA must be based on machine learning principles, in order to be robust enough for actual application. An objective of this chapter is to describe our research on machine learning applied to robust crater detection. We start by reviewing the existing literature on CDA, emphasizing the difference between semi-automatic approaches, based purely on image processing which identify circular features in an image leaving actual decision of whether a structure is indeed a crater to an analyst, and fully automatic approaches, based on a combination of image processing and machine learning which make their own decisions about structures being craters. We then proceed to describe two different families of machine-based CDAs, one constructed for identification of craters from digital elevation models (also known as range images), and another constructed for identification of craters from panchromatic images. We present applications of both CDAs and demonstrate how such automated analysis led to new knowledge about planet Mars. We finish with enumerating existing challenges and indicating future research directions.

APPROACHES TO AUTO--DETECTION OF CRATERS

Because of the importance of craters for planetary science, the literature on crater detections algorithms is extensive (Salamuniccar & Loncaric, GT-57633 catalogue of martian impact craters developed for evaluation of crater detection algorithms, 2008) (Salamuniccar & Loncaric, 2008). Since 2008, there has been a further increase in publications on auto-cataloging of craters; (Salamaniccar & Loncaric, 2010) provides extensive references for those newer works. From the data source point of view, the CDAs can be divided into those that detect craters from images (most often panchromatic images), and those that detect craters from digital elevation models (DEMs). DEM is a raster-type dataset that stores the value of elevation in each cell.

Image-based crater-detection approaches could be divided into those that dispense with machine learning and those that exploit it. The first class of methods rely exclusively on pattern recognition techniques to identify crater rims having circular or elliptical features in an image (for example, (Barata, Alves, Saraiva,

& Pina, 2004) (Cheng, Johnson, Matthies, & Olson, 2002) (Honda, Iijima, & Konishi, 2002) (Kim, Muller, S., J., & Neukum, 2005) (Leroy, Medioni, & Matthies, 2001) (Salamaniccar & Loncaric, 2010) (Salamuniccar & Loncaric, 2008) (Salamuniccar & Loncaric, 2008). The general idea of such methods is to first preprocess an image to enhance the edges of the crater rims, and then to detect the craters using variants of the Hough Transform (Hough V, 1962), genetic algorithms (Honda, Iijima, & Konishi, 2002), or the radial consistency algorithm (Earl, Chicarro, Koeberl, Marchetti, & Milnes, 2005) that identifies regions of rotational symmetry.

The second class of methods (for example, (Burl, Stough, Colwell, Bierhaus, Merline, & Chapman, 2001) (Plesko, Brumby, Asphaug, Chamberlain, & Engel., 2005) (Vinogradova, Burl, & Mjolsness, 2002) (Wetzler, Honda, Enke, Merline, Chapman, & Burl, 2005) utilize machine learning to facilitate crater detection. In a learning phase, the training set of images containing craters labeled by domain experts is fed to an algorithm. In the detection phase, the previously trained algorithm detects craters in a new, unlabeled set of images (Burl, Stough, Colwell, Bierhaus, Merline, & Chapman, 2001); (Vinogradova, Burl, & Mjolsness, 2002) used a continuously scalable template-model technique to achieve detection. (Wetzler, Honda, Enke, Merline, Chapman, & Burl, 2005) tested a number of machine learning algorithms and reported that support vector machines achieve the best rate of crater detection. (Plesko, Brumby, Asphaug, Chamberlain, & Engel., 2005) used a genetic programming to generate a population of random-detection algorithms whose performance is iteratively improved using a training set as selection criteria.

Image-based CDAs must employ multistep algorithms to combat inherent limitations of imagery data (see previous section). (Honda, Iijima, & Konishi, 2002) first clusters the set of images from which craters are to be detected with respect to image quality and apply separate, optimized detection algorithm to each cluster. (Kim, Muller, S., J., & Neukum, 2005) verified detected craters by template matching and employed neural networks to remove false detections. In spite of such measures, image-based crater-detection algorithms had only limited success. When applied to imagery data, the machine learning-based CDA algorithms work well for small craters (Plesko, Brumby, Asphaug, Chamberlain, & Engel., 2005), our papers) and/or for relatively simple terrain, but their efficiency drops in proportion to the complexity of the terrain (Vinogradova, Burl, & Mjolsness, 2002). Methods that don't employ machine learning work well in the limited context of an autonomous spacecraft navigation system (Cheng, Johnson, Matthies, & Olson, 2002) because of the relative simplicity of asteroid surfaces.

For several planets (Mars, the Moon, and soon Mercury) high resolution DEMs with near global coverage are available. DEMs are much more fundamental descriptors of planetary surfaces than images. They are suitable for a quantitative geomorphic analysis and are well suited for automated identification of craters. Some authors (Salamaniccar & Loncaric, 2010) (Salamuniccar & Loncaric, 2008) (Salamuniccar & Loncaric, 2008) (Stepinski, Mendenhall, & Bue, Robust automated identification of martian impact craters, 2007) identify craters in a DEM in a manner similar to the identification of craters in an image – through rim detection. Alternatively, (Stepinski, Mendenhall, & Bue., 2009) fully utilize tree-dimensional character of the DEM data and identify craters from DEM as round-shaped depression of certain depth. Overall, it is preferable to detect craters from DEMs than from images. However, DEMs are still limited in availability and resolution, so detecting craters from images is still necessary.

TOWARD ROBUST DETECTION OF CRATERS

Crater detection algorithms design issues

Crater detection may be an interesting computer science challenge, but ultimately any CDA algorithm must address the needs of the end user – a planetary scientist that needs to survey craters over some extent of planetary surface. Most needed are CDAs capable of global surveys of small craters; such algorithms are required to identify up to millions of craters from large collections of images (or DEMs) in a robust

manner and with minimal involvement from an analyst. Because of the sheer size of the task, a practical algorithm must rely on machine learning. Algorithms based exclusively on pattern recognition are not discriminating enough to have sufficient accuracy and human examination of the results, and are out of question for tasks involving thousands to millions of craters. The overall design architecture for a robust CDA consists of three components: (1) identification of crater candidates (pattern recognition task), (2) binary classification of crater candidates into craters and non-craters (machine learning task), and (3) an ability to adjust a classifier to a new type of surface in order to maintain its performance while minimizing the cost of adjustment to an analyst.

The design of some CDAs combines the first two components; crater candidates are not calculated, instead a decision about a block of image being a crater is done directly on the basis of its pattern. In our experience, such approach results in either high computational cost or inferior accuracy. For example, a viable CDA design may be based on a combination of texture features (Papageorgiou, Oren, & Poggio, 1998) and boosting algorithm (Viola & Jones, 2004). This is an example of a design that concentrates on machine learning at the expense of image processing. In such a CDA, implemented by (Rodionova, et al., 2000), an exhaustive search generates blindly "crater candidates" consisting of square image blocks of all possible sizes centered on all possible locations in an image. Each block is classified as either containing a crater or not by boosting a classifier on the basis of its grayscale texture. Such design is capable of yielding a CDA of sufficient accuracy, however, at a large computational cost. This is because a classifier needs to evaluate a very large number of image blocks, the overwhelming majority of which contain no craters or only fragments of craters. An opposite design approach, one that concentrates on image processing at the expense of machine learning, was introduced by (Urbach & Stepinski, 2009). They detect craters efficiently taking advantage of the fact that photographic imprint of a crater contains crescent-like highlight and shadow regions. Their CDA utilizes methods of mathematical morphology (Serra, 1982) to design scale-invariant and rotation-invariant shape filters for identification of crescent like regions. Because a single application of shape filter to an entire image identifies all craters irrespective of their sizes and orientation, such CDA is very efficient. The lack of a machine learning component in deciding which shapes to include in the filters, however, results in poor crater identification accuracy. Our point of view is that separating identification of crater candidates (for example, by means of mathematical morphology with relatively relaxed shape criteria) with machine learning (for example a boosting algorithm applied only to crater candidates and not to all conceivable image blocks) offers the best possible design that optimizes both performance and accuracy.

Identification of craters from topography

For planet Mars and the Moon, global–scale datasets of topography exist in the form of DEMs. Mars was the first planet for which such dataset existed, so much of our effort focuses on auto-surveying craters on Mars, however the methodology can be applied without systemic changes to the Moon; the only change necessary is the training set, as the Martian and lunar craters have somewhat different character and shape.
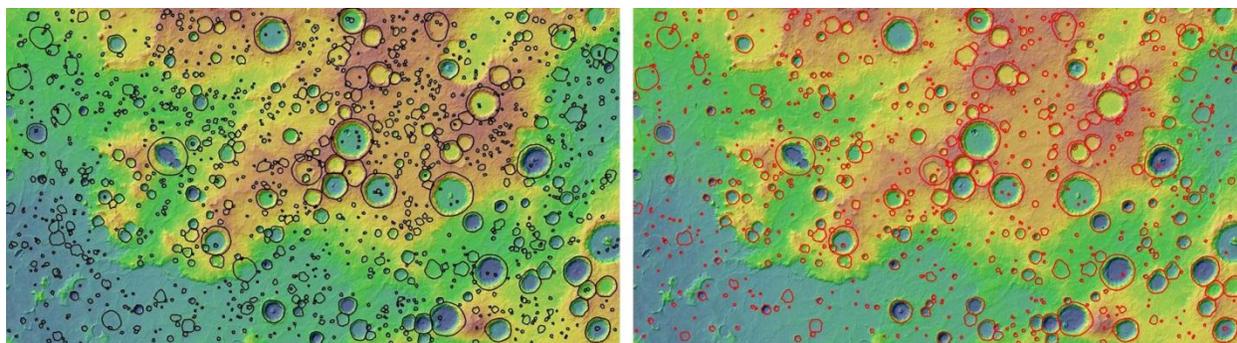
Our algorithm for automatic identification of craters from Martian topographic data (Urbach & Stepinski, 2009) is designed to have separate modules for identification of crater candidates and for the final classification. Such modular architecture separates the two major challenges facing all feature-finding algorithms, completeness (minimization of false negative detections) and accuracy (minimization of false positive detections). In the present context, false negatives are craters not identified by the algorithm, and false positives are non-crater features identified by the algorithm as craters. The module for identification of crater candidates is designed to address the completeness issue; its input is the topographic dataset (hereafter referred to as a *site* to underscore its geographical meaning), and its output is a list of crater candidates containing, as its subset, the true craters. The module for final classification of crater candidates is designed to address the accuracy issue. This is why it is based on machine learning; its input

is the list of crater candidates, and its output is the same list but with labels indicating whether a given depression is a crater or not

The core concept behind the "candidates" module is that, in the topographic data, craters are depressions. Thus, the role of the candidate-finding module is to identify all topographic depressions in a given dataset. The challenge is to identify all depressions including superposed depressions and irregular depressions. In principle, depressions in topographic dataset could be identified by the ''flooding" algorithm (O'Callagnan & Mark, 1984). The flooding algorithm identifies depressions in the DEM by raising elevation of pixels within them to the level of the lowest pour point. However, (Stepinski, Mendenhall, & Bue, 2007) have pointed out that flooding algorithms cannot be utilized as an accurate depression-finder in actual Martian landscapes because superposed depressions, common in such landscapes, are identified as a single depression by the flooding algorithm.
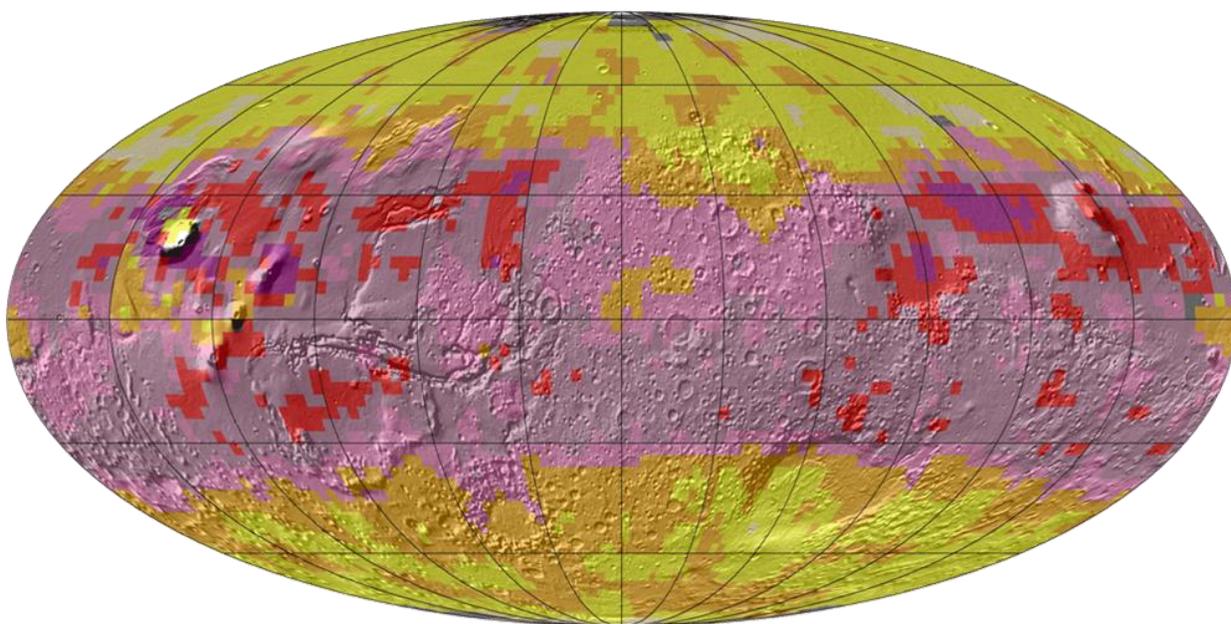
In our algorithm, the problem of nested depressions is addressed by first identifying only the smallest depressions, then proceeding to identification of successively larger depressions in subsequent steps. In order to separate depressions of different scales, we introduce a function that transforms the original landscape (as given by the DEM) into an artificial landscape optimized for identification of depressions having a given length scale. This transformation calculates the degree to which elevation gradients anchored to pixels in the given-scale neighborhood of the focal pixel are aligned to point toward that pixel (for details see (Urbach & Stepinski, 2009). Such transformation preserves depressions of a given scale, however, features having smaller scale are smoothed out, and features having larger scale bigger are suppressed. Topographic depressions in the transformed landscape are identified as upward-concave regions. Concativity is determined by calculating the discrete second derivative in principal directions; the single-connected regions of upward-concave pixels mark individual topographic depressions of a given scale. This depression-finding procedure is repeated using a transformation with increasingly larger spatial scale in order to identify increasingly larger depressions. All identified depressions are crater candidates; for each of them we calculate five features: diameter, depth, depth/diameter ratio, and two features describing a planar shape of the depression (elongation and lumpiness).

The machine learning module uses the C4.5 decision tree algorithm (Quinlan, 1993) as implemented in the software package WEKA (Witten & Frank, 2005). Because a crater's morphology depends on its size, we have acquired three different training sets for three different sizes of craters (large, medium, and small). With the resolution of the grid being about 0.5 km, the small craters are those having diameter of about 5 km, medium are those having diameters between 5 and 10 km, and large are those having diameter larger than 10 km. The training sets are constructed iteratively. First a relatively small number of crater candidates is hand-labeled by a human expert and a classifier is built based on this initial training set. This classifier is then applied to all crater candidates in a given site and the results are visually reviewed and corrected if necessary. The corrected results constitute a new, much larger training set. This procedure is repeated for a number of sites. The final three training sets contain 5970, 1010, and 431 labeled examples, respectively. The final pruned decision tree constructed on the basis of 5970 examples and pertaining to identification of smallest craters has 59 nodes and 30 leaves; its expected accuracy is 96.2%. The final pruned decision tree constructed on the basis of 1010 and pertaining to identification of medium-size craters has 25 nodes and 13 leaves; its expected accuracy is 90.1%. Finally, the decision tree constructed on the basis of 431 examples and pertaining to large craters has 15 nodes and 8 leaves; its expected accuracy is 89%. Accuracies are measured using 10-fold cross-validation (Kohavi, 1995). In order to visualize the process of crater detection, Figure 1 shows a small portion of the entire Martian surface with (Left) crater candidates indicated by black outlines and (Right) craters indicated by red outlines.

Figure 1. Visualizing crater detection from topographic data. Martian surface topography is illustrated by gradients of colors from blue (low elevations) to red (high elevations). (Left) Black outlines indicate depressions (crater candidates). (Right) Red outlines indicate craters.

The methodology described above was used to conduct automatic survey of craters over the entire surface of planet Mars (Stepinski & Urbach, 2009). The result of this survey is a catalog of 75,919 craters (ranging in size from 1.36 km to 347 km.) listing coordinates of the center of each crater, its diameter (D) and depth estimate (d). This survey constitutes a major progress in the study of Martian surface. A ``carpet coverage" of Mars surface by 75,919 craters with estimated depths makes, for the first time, possible construction of planet-wide maps showing geographical distribution of depths for craters of different sizes. These maps led to an independent confirmation of the notion that ice exists just beneath the surface at the higher latitudes of the planet but not in its equatorial regions. Fig. 2 shows a map of crater relative depths (d/D) for a subset of our auto survey restricted to craters in the 5-10 km size range. It shows that the craters are deeper at the equator than at higher latitude regions. Because craters of the same size should be equally deep, this result indicate that craters at high latitudes experiences some process subsequent to their formation that led to their shallow nature. The best interpretation is that these craters were emplaced in a ground rich in ice, a substance that undergoes a shape relaxation on geologic times. The craters located near the equator preserved their original depths indicating lack of ice there.

Figure 2. Map of craters relative depths created using attributes of 75,919 craters automatically detected by our crater detection algorithm. Color gradient from red to yellow indicates deep to shallow craters.
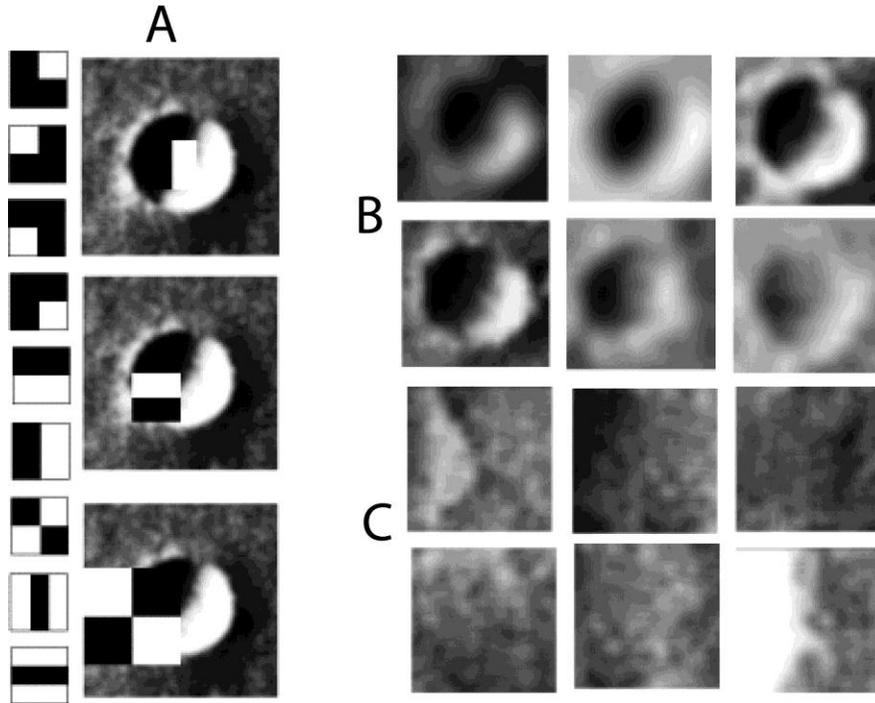
Identification of craters from images

Presently, topographic data is too coarse for identification of craters that are smaller than about 1 km in diameter. However, those small craters are most useful for improvements in surface chronology. Auto-surveys of craters from topography offers unique capabilities for estimating crater depths and may lead to knowledge discovery (see previous sub-section), but for improved accuracy of surface chronology, where only size and not the depth of the crater is required, image-based auto-surveys are necessary. In planetary context, high resolution images are panchromatic (grayscale), so the task is to find craters in grayscale images. Our approach to image-based crater detection is consistent with our philosophy that the process works best if divided into two parts: identification of crater candidates, and finding true craters from amongst the candidates. In application to image-based crater detection, these modules are based on different principles when compared to topography-based detection (discussed above). This is because craters in images look different than craters in 3-D terrains.

The core concept behind the "candidates" module is that, in the imagery data, a small crater appears as a pair of semi-circular of crescent-like highlight and shadow regions. Thus, (Urbach & Stepinski, 2009) proposed identifying small craters utilizing tools of mathematical morphology (Serra, 1982). Rotation and shape invariant filters can be constructed for identification of crescent-like regions in an image. 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, craters are not the only features on planetary surfaces that may have such a double crescent imprint in an image. Thus, a mathematical morphology-based algorithm is not a very accurate stand-alone crater detection algorithm, but it is a very efficient crater candidate finding algorithm. The efficiency of identifying crater candidates in an image using our algorithm becomes clear if one considers alternatives. The only alternative is to use exhaustive search of sub-windows of different sizes and locations within an image. There are many orders of magnitude more such sub-windows than actual craters; mathematical morphology identifies number of crater candidates that is about the same order of magnitude as the number of craters.

The output of the crater candidate detection module is a list of shapes (each containing a pair of crescent-like shadow and highlight regions). This list constitutes the input to a machine learning module. The first design choice is to select image features as discriminants between craters and other surface objects on the list of candidates. Unlike the case of topographic data, no clear, physics-based features are available, because image is not a physical reality of the surface but rather its projection into 2-D space under given illumination. We encode images of crater candidates in terms of texture features (Ding, et al., 2011). First, each candidate (irregular fragment of an image) is embedded in a square image block, centered on the location of the candidate and having a dimension twice the diameter of the candidate. Second, texture features are extracted from the image block using a simple geometrical technique for texture feature extraction first proposed in (Papageorgiou, Oren, & Poggio, 1998), and popularized by (Viola & Jones, 2004) in the context of face recognition. Figure 2A illustrates the construction of such features; only one of possible mask sizes and one mask location are shown. Such texture features are broadly utilized for object detection and have proven to work well for crater detection (Martins, Pina, Marques, & Silveira., 2008). Overall, we represent each crater candidate by 1089 texture features; see (Ding, et al., 2011) for details. A large number of features restricts our choices of a learning algorithm; for example, decision trees or support vector machines would be ineffective in such context. We need to select automatically only those features that are most useful in discriminating between craters and non-craters. We use a variant (Viola & Jones, 2004) of AdaBoost algorithm (Freund & Schapire, 1995) that simultaneously selects the best features and trains the classifier.

We evaluated our approach to image-based crater detection using high resolution (12.5 meters/pixel) images of Mars taken by the High Resolution Stereo Camera (HRSC) on-board the Mars Express

spacecraft. In this site an analyst has manually cataloged 1937 craters having diameters larger than 16 pixels in size but smaller than 400 pixels. These craters are deemed "detectable" by our method and serve as the ground truth.
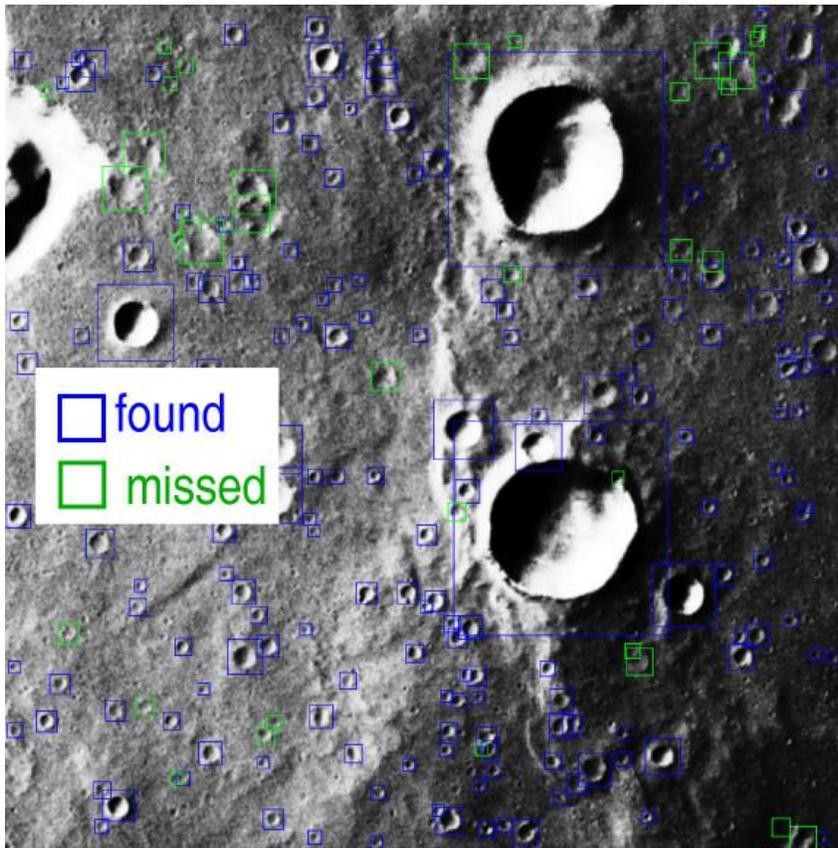

Figure 3. (A) The concept of texture features. Different shape masks used to calculate texture features are shown on the left. The value of a particular mask-feature is obtained by placing the mask over a selected part of an image block and subtracting the sum of grayscale values of the pixels covered by black sectors of the mask from the sum of grayscale values of the pixels covered by white sectors of the mask. (B) Positive examples of craters. (C) Negative examples of craters.

The deep valley passes through the middle of the image introducing heterogeneity of the terrain. In order to account for this heterogeneity in the evaluation of our algorithm, we divided the image into three sections (West, Central, and East). Our experimental setup corresponds to the likely use of an algorithm by a planetary scientist who is expected to have a small image annotated (ground truth), but wants to find the craters in a different, larger image. Note that an alternative (not likely to be of practical importance in planetary research) is to have a training set scattered across the large image. Applying the crater candidate finding module to the test site results in the identification of 14,004 crater candidates. We have chosen 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. Figure 3B and 3C shows few of the positive and negative examples. In the training phase of the classification module the AdaBoost algorithm ranks the features by their importance in distinguishing between craters and non-craters and establishes the minimum number of features (about 100) required number for classification phase. The selected features focus on detecting a boundary between the shadow and the highlight portions of a crater.

We use the trained classifier to detect the craters in the entire image. Overall the detection rate is 81% compared with 64% for an algorithm based on mathematical morphology alone. Thus application of supervised learning improved the performance of the classifier by 17%. Moreover, the detection rate in the East section of an image (from where training set was sampled) is 87%. However, detection rate in

West and central sections, which did not contribute to the training set) is about 79%. Figure 4 shows a fragment of the entire image with found and missed craters indicated by different color outlines. This is because the West and Central regions contain crater candidates having character unaccounted for in the training set. We could remedy this drop in performance by selecting candidates from the West and 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 (see above). Future research should incorporate techniques of transfer learning and active learning 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.

Figure 4. Fragment of the grayscale image used to evaluate our image-based CDA. Found and missed craters are shown as indicated.

Solutions and Recommendations

Automatic crater detection is very likely to become the first popular tool for a planetary scientist that employs supervised learning. The technology has still some way to go before delivering a robust and usable implementation of a crater detection algorithm. In many ways the problem is similar to face detecting algorithms but the technology may take longer to develop because of lower demand and, consequently, fewer resources devoted to the research. The most promising approach to automate crater detection is to find them from topographic data. We expect that the first truly robust crater detector will be based on the high resolution DEMs which are in the process of being obtained for the Moon and planet

Mercury. No new topography data gathering mission to the planet mars is planned, so we are likely to be stuck with currently available medium resolution DEMs, although higher resolution DEMs may be constructed locally from stereo pair images. Detecting craters from images will need to be done for planets where high resolution topography data is not available.

Irrespectively of the data type used, we advocate a two-step design for crater detection algorithm. Such design separates the process of finding crater candidates from the process of labeling the candidates. Both steps require further research, although it is the first step (finding the candidates) that can benefit most from further research. Existing supervised learning techniques appear sufficient for the second step of CDA, however further research is needed for identification of the best set of features. More work is necessary on how to assess the accuracy of CDAs. There are problems with the ground truth as well as with the process of matching finds with the ground truth. Unlike the case of human faces, what is and what is not a crater is sometimes debatable, and experts may differ in their opinions. Thus, in the crater detection problem, there is no such thing as a completely objective ground truth. Moreover, the ground truth not only consists of the presence/absence of the crater but also of the size of the crater. Measuring the size of the crater is not as straightforward as it may appear and there is no guarantee that expert collected measurements are always accurate. Thus comparing craters identified by a CDA with ground truth requires a certain degree of flexibility with respect to position and size; finding the best, standard way to do it remains a challenge.

FUTURE RESEARCH DIRECTIONS

The most important challenge in designing a robust and usable CDA is how to maintain high accuracy throughout changing images (or DEMs) without significant additional training. The character of planetary surface changes with location and so does the appearance of craters embedded in those surfaces. Planetary scientist expects to obtain a CDA and use it "out of the box" without training or with minimal training. A CDA that requires extensive training will not be accepted by the planetary community. A solution to this problem is to collect an extensive training set that reflects many known types of surfaces. However, there is a great variety of planetary surfaces and not all of them could be anticipated by CDA designers. This is why future research must concentrate on incorporating elements of active learning, semi-supervised learning, and learning transfer methods to offer adaptation of CDA to different datasets. First attempts (Ding W. S., 2010) (Ding, et al., 2011) on incorporating transfer learning to crater detection show some promise but much more work is necessary.

CONCLUSION

Crater survey is an important task in the field of planetary science, which until now has been performed manually. With increasing amount of available images and other planetary data, this task is ripe for automation via machine learning techniques. In this chapter we have given an overview of the field, pointing the readers to the existing literature and presenting our own contribution to designing and implementing crater detection algorithms. Our approach is underpinned by the philosophy that an efficient and accurate CDA needs to be divided into a step that finds crater candidates, and a step that narrows the field of crater candidates to just craters. Within this flexible framework we demonstrated two different algorithms, one for detection of craters from topographic data, and another for detection of craters from images. Although the two algorithms share our design philosophy, they differ very much in particularities. We have also demonstrated, using Mars as an example, how obtaining a global catalog craters can lead to knowledge discovery.

The algorithms presented here give reasonably accurate surveys of craters, but can benefit from better training sets. We suggest that the most important challenge of CDA is to incorporate elements of transfer learning and/or machine learning to allow for efficient addition of training samples as the need arises.

REFERENCES

Andersson, L. B., & Whitaker, B. A. (1982). NASA catalogue of lunar nomenclature. *NASA Reference Publication* , 1097.

Barata, T., Alves, E. I., Saraiva, J., & Pina, P. (2004). Automatic Recognition of Impact Craters on the Surface of Mars. *ICIAR, 2*, pp. 489-496.

Barlow, N. G. (1988). Crater size-distributions and a revised martian relative. *Icarus , 75 (2)*, 285-305.

Burl, M. C., Stough, T., Colwell, W., Bierhaus, E. B., Merline, W. J., & Chapman, C. (2001). Automated detection of craters and other geological features. *Int. Symp. Artif. Intell. Robot. and Autom. Space.* Montreal.

Cheng, Y., Johnson, A. E., Matthies, L. H., & Olson, C. F. (2002). Optical landmark detection for spacecraft navigation. *the 13th Annual AAS/AIAA Space Flight Mechanics Meeting*, (pp. 1785-1803). Puerto Rico.

Ding, W. S. (2010). Automatic detection of craters in planetary images: An embedded framework using feature selection and boosting. *the 19th ACM International Conference on Information and Knowledge Management.* Toronto, Canada.

Ding, W., Stepinski, T., Mu, Y., Bandeira, L., Vilalta, R., Wu, Y., et al. (2011). Sub-Kilometer Crater Discovery with Boosting and Transfer Learning. *ACM Transactions on Intelligent Systems and Technology* .

Earl, J., Chicarro, A. F., Koeberl, C., Marchetti, P. G., & Milnes, M. (2005). Automatic Recognition of Crater-like Structures in Terrestrial and Planetary Images. *36th Annual Lunar and Planetary Science Conference.* League City.

Freund, Y., & Schapire, R. (1995). A decision-theoretic generalization of on-line learning and an application to boosting. *Computational Learning Theory: Eurocolt.* , 23-37.

Hartmann, W. K. (1999). Martian cratering VI: Crater count isochrons and evidence for recent volcanism from Mars Global Surveyor. *Meteoritics & Planetary Science , 34* (2), 166-177.

Hartmann, W. K., & Neukum, G. (2001). Cratering chronology and evolution of Mars. *Chronology and Evolution of Mars* , 165-194.

Honda, R., Iijima, Y., & Konishi, O. (2002). Mining of topographic feature from heterogeneous imagery and its application to lunar craters. *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project* .

Hough V, P. C. (1962). United States Patent.

Kim, J., Muller, J.-P., S., V. G., J., M., & Neukum, G. (2005). Automated crater detection: a new tool for Mars cartography and chronology. *Photogrammetric Engineering and Remote Sensing , 71*, 1205-1217.

Kozlova, E. A., Michael, G. G., Rodinova, J. F., & Shevchenko, V. V. (2001). Compilation and preliminary analysis of a catalogue. *Lunar and Planetary Science , XXXII*, 1231.

Leroy, B., Medioni, G., & Matthies, E. J. (2001). Crater detection for autonomous landing on asteroids. *Image and Vision Computing , 19*, 787-792.

Martins, R., Pina, P., Marques, J. S., & Silveira., M. (2008). Crater detection by a boosting approach. *IEEE Geoscience and Remote Sensing Letters , 6* (1), 127-131.

O'Callagnan, J., & Mark, D. (1984). The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process , 28*, 328-344.

Papageorgiou, C., Oren, M., & Poggio, T. (1998). A general framework for object detection. *Sixth International Conference on Computer Vision*, (pp. 555-562).

Plesko, C., Brumby, S., Asphaug, E., Chamberlain, D., & Engel., T. (2005). Automatic crater counts on Mars. *unar and Planetary Science XXXV.*

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning.* San Francisco: Morgan Kaufmann Publishers.

Rodionova, F. J., Dekchtyareva, K. I., Khramchikhin, A., Michael, G. G., Ajukov, S. V., Pugacheva, S. G., et al. (2000). Morphological catalogue of the craters of Mars. *ESA-ESTEC* .

Salamaniccar, G., & Loncaric, S. (2010). Method for crater detection from martian digital topography data using gradient value/orientation, morphometry, vote analysis, slip tuning, and calibration. *IEEE Transactions on Geoscience and Remote* .

Salamuniccar, G., & Loncaric, S. (2008). GT-57633 catalogue of martian impact craters developed for evaluation of crater detection algorithms. *Planetary and Space Science* .

Salamuniccar, G., & Loncaric, S. (2008). Open framework for objective evaluation of crater detection algorithms with first test-field subsytem based on MOLA data. *Advances in Space Research , 42*, 6-19.

Serra, J. (1982). *Image Analysis and Mathematical Morphology.* Academic Press.

Stepinski, T. F., & Urbach, E. R. (2009). The First Automatic Survey of Impact Craters on Mars: Global Maps of Depth/Diameter Ratio. *40th Lunar and Planetary Science Conference, (Lunar and Planetary Science XL).* The Woodlands.

Stepinski, T. F., Mendenhall, M. P., & Bue, B. D. (2007). Robust automated identification of martian impact craters. *Lunar and Planetary Science XXXVIII.* Lunar and Planetary Institute.

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

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

Urbach, E. R., & Stepinski, T. F. (2009). Automatic detection of sub-km craters in high resolution planetary images. *Planetary and Space Science , 57*, 880-887.

Vinogradova, T., Burl, M., & Mjolsness, E. (2002). Training of a crater detection algorithm for Mars crater imagery. *IEEE Aerospace Conference Proceedings*, *7*, pp. 3201-3211.

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision , 57*, 147-154.

Wetzler, P., Honda, R., Enke, B., Merline, W., Chapman, C., & Burl, M. (2005). Learning to detect small impact craters. *Seventh IEEE Workshops on Application of Computer Vision.*, *1*, pp. 178-184. Breckenridge.

Wise, U. D., & Minkowski, G. (1980). *Dating methodology of small, homogeneous crater populations applied to the tempe-utopia trough region on Mars.* Goddard Space Flight. Greenbelt: NASA.

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques, 2nd Edition.* San Francisco: Morgan Kaufmann.

KEYWORD: Crater Detection, Supervised Learning, Machine Learning, Mars