

# Crater Extraction and Classification System for Lunar Images

Rie Honda and Ryushi Azuma

( Department of Mathematics and Information Science, Kochi University )

## Abstract

A large number of images of planetary surfaces have recently been obtained by satellite remote sensing. The surfaces of solid planets are covered with a tremendous number of craters of various shape and size that reflect the geological history and rheology of the planetary surfaces. However, the extraction of craters and the following analysis remains difficult because it requires a great deal of human work. In the present study we examined the crater detection and categorization process in view of data mining from a large-scale scientific image database. The investigated system consists of an image database and modules for the detection and categorization of craters. The detection modules are based on an image processing method including binarization, circular object detection using the Hough transformation and the Genetic Algorithm. Extracted crater candidates are normalized concerned with size and intensity and accumulated in the secondary database. These images are categorized through learning by self-organizing mapping. We show here preliminary results for both extraction and categorization based on the above-described framework.

## 1 Introduction

A large amount of remote-sensing image data has recently been obtained by satellites due to the improvement of sensors and telemetry systems. These conditions have led scientists to request the construction of a large-scale scientific image database, to extract the characteristics of these images, and to construct a secondary database without manual operation. In addition, there is a demand for both theoretical and practical investigations into knowledge discovery and data mining system for these data.

Based on their interests in knowledge discovery and data-mining systems for scientific images, Fayyad et. al. (1996) developed the Sky Image Cataloging and Analysis tool (SKICAT) for the second Palomar Observatory Sky Survey. Smyth et al. (1996) and Burl et al.(1998) have reported a discovery system of venusian volcanoes from Synthetic Aperture Radar images taken by the spacecraft Magellan. These studies have been aimed at the automatic extraction of objects of interest and classification based on a machine-learning method. Furthermore, Burl et al.(1999) have recently reported an

automated feature-detection system for planetary images named Diamond Eye.

In this study we consider a knowledge discovery support system for the lunar multi-spectral images taken by the U.S. spacecraft Clementine (Nozette et. al., 1994), particularly focusing on crater detection and categorization problems.

A more through description of the crater analysis of planetary surfaces is given in chapter 2, and a system overview as well as details of the method of extraction of crater candidates and their categorization using self-organizing mapping is described in chapter 3. This chapter gives the results of the implementation and related experiments. Chapter 4 presents a general discussion and a description of the problems to be investigated in future studies.

## 2 Backgrounds

It is widely known that solid planet surfaces are covered with a tremendous number of craters of various sizes and shapes (Melosh, 1988). Most of these craters have formed as a result of meteoroid impacts. Assuming a size-frequency distribution of meteoroids and its time variations, we can estimate the relative formation age of the local area of a solid planet surface by counting and sizing craters.

Furthermore the craters also have a variety of shapes, including bowl-floored craters, flat-floored craters, flat-floored craters with central peaks, and degraded craters, reflecting the surface rheology and subsurface structure. Additionally, the craters are considered to be windows into the interiors of planetary bodies. Central peaks in particular are composed of deeper material that has been lifted up after crater excavation.

Despite the importance of crater data and the tremendous amount of data available, crater analysis remains dependent on human vision and manual operation.

In 1994, the U. S. Spacecraft Clementine obtained lunar global images by means of UV-VIS sensors. These data are archived in 88 volumes of CD-ROMs and the U. S. Geological Survey has composed a mosaic images called lunar digital image model (LDIM, 15 volumes of CD-ROMs) from the 750-nm band data.

By using the LDIM, we examine here the automatic extraction and categorization of craters as a prototype of an image data-mining system for planetary imagery. We believe that this concept will be easily extended to the other features such as volcanoes, grooves, and lineaments.

### 3 Method

#### 3.1 Overview

An overview of the image-mining system for craters is illustrated in Figure 1. The system includes image database, preprocessing by filtering and binarization, circular object detection by a genetic algorithm (GA) and Hough transform(HT), and categorization using self-organizing map (SOM). Additional leaning and clustering utilizes the feature map previously learned by SOM.

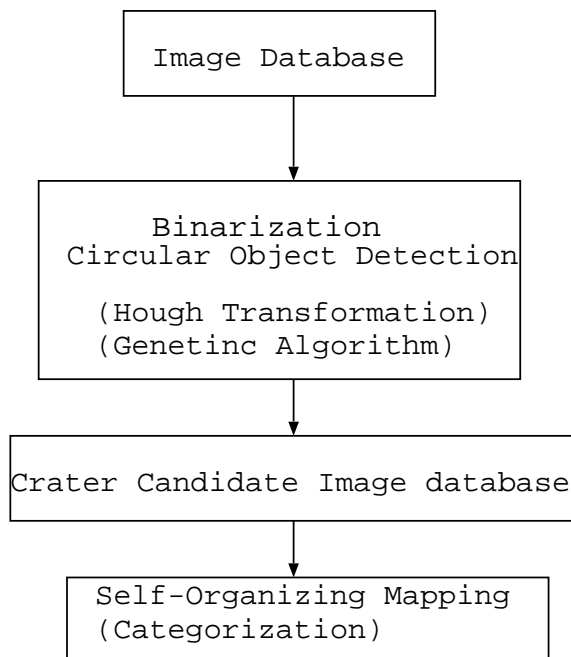


Figure 1: Overview of crater detection and categorizing support system.

The image database is composed of a Clementine LDIM consisting of a set of images whose height and width are equal to a 7 degree in latitude. Thus the image database is composed by tuple as follows:

{Image name, Reticle longitudes, Reticle latitudes }.

LDIM are grayscale images, most of which are taken when the spacecraft and the sun are situated in the same direction in relation to the lunar surface. Images around 60 degree in latitude are taken at a solar altitude of approximately 30 degree and are thus appropriate for terrain analysis.

Since we recognize craters based on spatial intensity changes, we preprocess the images by operating an edge detection filter (Robert filter). The threshold for binarization is determined statistically from the intensity histogram, and the images are transformed into

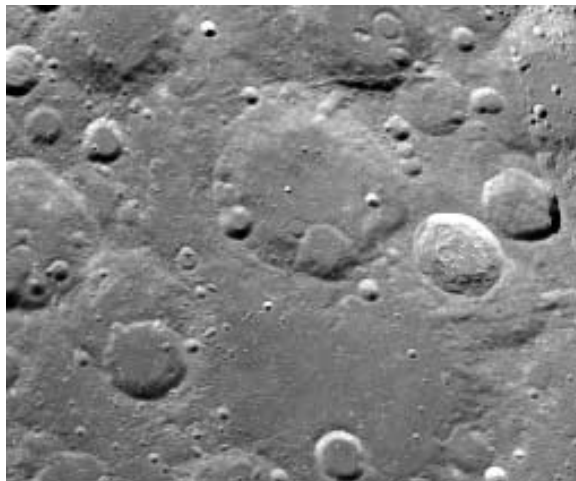


Figure 2: Original lunar image.

binary data.

### 3.2 Circular object detection

After the binarization, the crater extraction is considered in terms of circular signal detection by a pattern-matching method. We examined the following two approaches for comparison.

#### 3.2.1 Combinational Hough transform

The Hough transform is a commonly used method for extracting geometrically simple parametric figures from binary images. The signals on images are projected into the parameter space of the figure. In this case, the parameters are the center of the circle and its radius. Watanabe and Shibata(1990) have reported combinational Hough Transform (CHT) that uses pyramid images and a pair of signals in a restricted region to simplify the projection to a parameter space, showing that CBT can significantly reduce computational time and improve the solution accuracy. We therefore implemented this method for crater detection.

Figure 2 shows the original lunar sample image, and Figures 3 and 4 show the craters detected by CHT for the lower and upper pyramid image, respectively. These figures show that clear craters are detected relatively well. We also examined noise reduction techniques such as thinning, expansion, and shrinking, and the removal of isolated signal related to this process, and found that noise reduction process strongly affects the computational time and accuracy of the solution. Thus it is crucial to produce "good" binary images in this method.

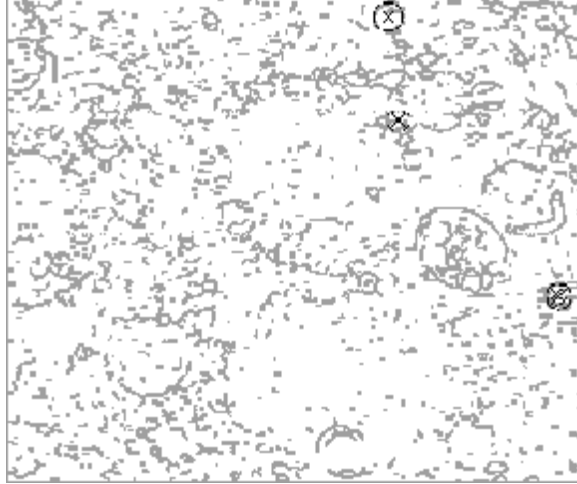


Figure 3: The results of circle detection by CHT for the lower pyramid image. Crosses represent the center of the detected craters, and solid circles represent the extracted crater rims.

### 3.2.2 Genetic Algorithm

A genetic algorithm is commonly used to obtain the single solution in the optimizing problem. In order to implement this algorithm for circular object detection, we set the gene as follows:

$$\{x_i, y_i, r_i\},$$

where  $(x_i, y_i)$  is the center of the circle represented by i-th gene and  $r_i$  is the radius of the circle. The adaptation for the i-th gene,  $g_i$ , is calculated by drawing the circle of the i-th gene on the binary image and checking the overlap:

$$g_i = n_i/N_i$$

where  $n_i$  is the number of black pixels on the circle of i-th gene, and  $N_i$  is the total number of pixels on the circle.

In order to avoid random noise being incorporated into the solution, we modified  $g_i$  as follows:

$$g'_i = g_i - g_{i,r=fr_i},$$

where  $0 < f < 1.0$ , typically  $f=0.3$ .

Since it is possible to have many craters in a single image, we have included a process to erase the detected circles from the original images, and the same process is continued iteratively.

Figure 5 shows thus the detected craters from Figure 2. This figure shows the GA is less able to filter noises than CHT, and that, at present, more false signals are detected. However, it appears that

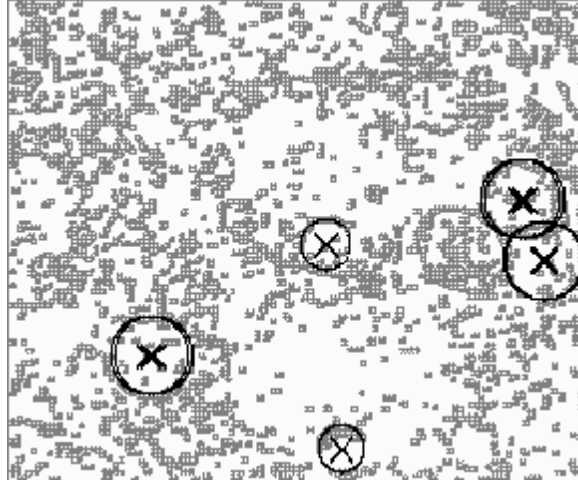


Figure 4: The same as in Figure 3 except we show here the results for the upper pyramid image.

the GA can be also applied to very noisy images for which the CBT cannot be used. We believe that the optimization of parameters for GA in the crater detection problem should be a subject of future research.

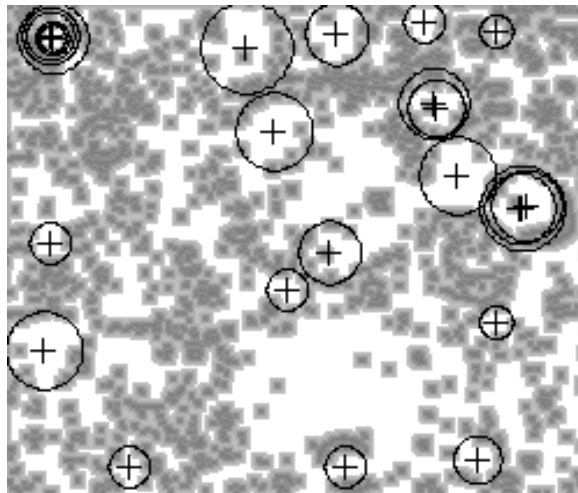


Figure 5: The same as in Figure 3, except we show here the result of genetic algorithm.

### 3.2.3 Selection of definite candidates from the GA and CHT results

The crater candidates extracted by CHT and GA contain plural solutions for a single solution. Thus we consider here the following method for grouping the candidates into definite solutions.

1. All solutions detected in one image are listed by the following tuple,

$$t_i = (x_i, y_i, r_i, m_i, \text{imagename}_i),$$

where  $m_i$  is the matching ratio.

2. The lists are sorted in the order of the values of the matching ratios.
3. The  $i$ -th tuple is projected into the parameter space composed of  $(x, y, r)$ , the process described below selects the definite candidates :

```

k = 0;
for i = 1, i < im, ++ i
    j = 0;
    while (sqrt((x_i - x_j)^2 + (y_i - y_j)^2 + (r_i - r_j)^2) > threshold and j <= k)
        j = j + 1;
    end of while
    if (j = k)
        t_k = t_i; k = k + 1;
    else
        t_j = t_i;
    endif
endif

```

where the value of  $im$  is the number of candidates,  $k$  is the number of the definite candidates, and  $\{t_i | 0 \leq i \leq k\}$  is the list of definite candidates.

### 3.3 Categorization of craters by SOM

In order to categorize the craters, we cut out the crater images of their diameter width and height from original 'grayscale' images.

We used Kohonen's self-organizing mapping (Kohonen, 1992), as this method has previously been implemented for clustering of weather (cloud) images by GMS 5 (Katayama and Konishi, 1999), and it has been demonstrated that it can successfully categorize the image characteristics. Kohonen's self-organizing mapping uses a two-dimensional feature map whose cells have feature vectors. The distance between the input vector and each cell's feature vector is calculated and the input vector is placed into the closest cell. At the same time, the feature vector in the cells adjacent to the closest cell is modified so that it moves closer to the input vector. As a result of this iterative learning process, we obtain a feature map that reflects variations in the input vectors and at the same time adequate clustering of the input vectors.

We considered three types of input vectors for crater analysis.

- Image vectors normalized with respect to intensity and size. The normalized intensity is represented by

$$p'(i, j) = \{p(i, j) - p_{mean}\} / \sigma,$$

where  $p(i, j)$  is the intensity of the  $i$ -th and  $j$ -th pixel for each crater image, and  $p_{mean}$  and  $\sigma$  are the average and the standard deviation, respectively, in the window.

- An intensity histogram calculated from the normalized image.
- A discrete cosine transform (DCT) component calculated from the normalized image.

We examined the first two types for learning by SOM in the present work, since DCT requires additional execution time for preprocessing and thus is considered to be inferior to data mining from a large quantity of scientific data.

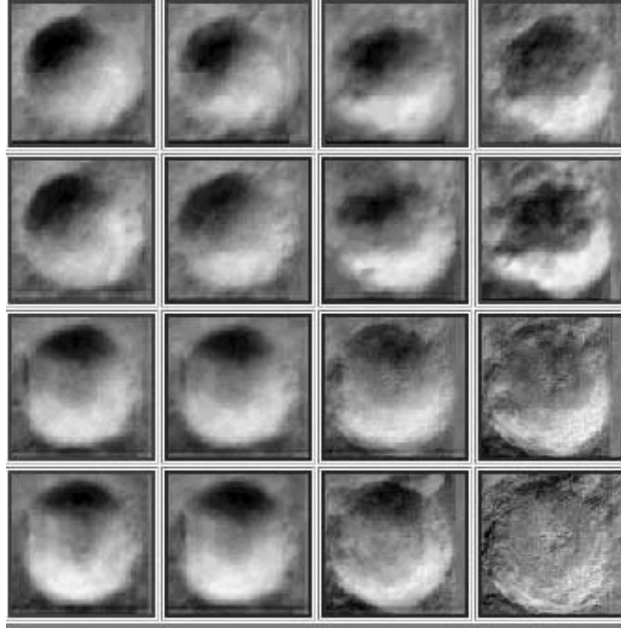


Figure 6: Feature map obtained by SOM using normalized image vectors for 37 craters extracted from two LDIM images.

As preliminary experiments we processed two LDIM images. The craters are selected both visually and manually since neither the algorithm of GA or CHT is robust enough for extremely noisy lunar images. We selected 37 craters, including simple bowl-shaped craters, degraded craters, and flat floored craters. Each image is normalized to the size of 100-pixel square, and the intensities are normalized to an average of 128 and a standard deviation of 40.

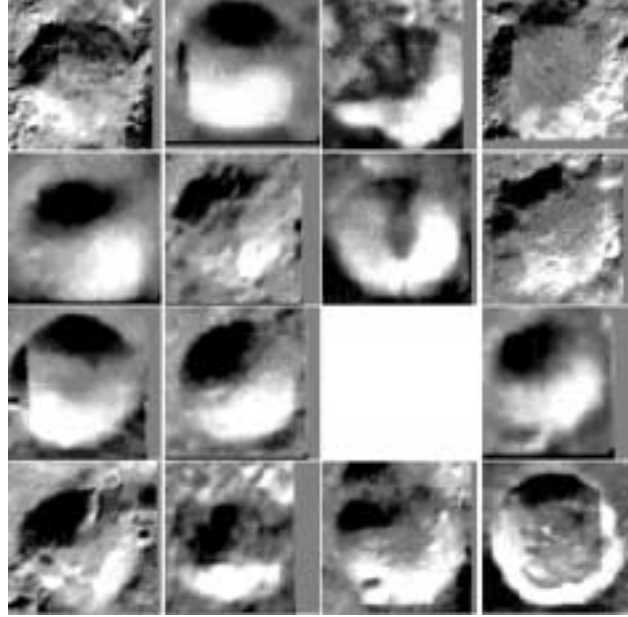


Figure 7: The same as in Figure 6 except for the use of the intensity histogram for input vectors and for taking the cut images that minimize the distance in each cell as representative images.

The SOM block size is taken to be  $4 \times 4$ , the number of iteration is 5000, and the neighborhood learning distance is 1.

Figure 6 shows thus obtained feature map taking the normalized image vector as the input. The simple bowl-shaped craters are clustered into the upper left cells, and the flat-floored craters are clustered in the lower left corner cells. This result indicates that normalized image vectors can successfully distinguish the variations in crater shape that are recognized in human vision.

Figure 7 shows the feature map of the representative images after learning by SOM when the intensity histogram vectors are taken as the input vectors. This figure shows no significant meaning in the distribution of cut images in the feature map and thus a failure to learn the crater patterns from input vectors.

We have not fully considered DCT components as input vectors, but this method is considered to be inappropriate for data mining because the time required for pre-processing by DCT is considered to be comparable to that spent by the SOM. The above-described results indicate that normalized image vectors classify crater images more effectively than histogram vector concerned with characteristics recognized by human vision.

Furthermore, we have extended the above-described experiments to 170 craters located in the higher middle latitudes ranging from 56 degrees to 63 degrees N. All the craters were extracted both visually

and manually, and normalized image vectors were taken as the input data.

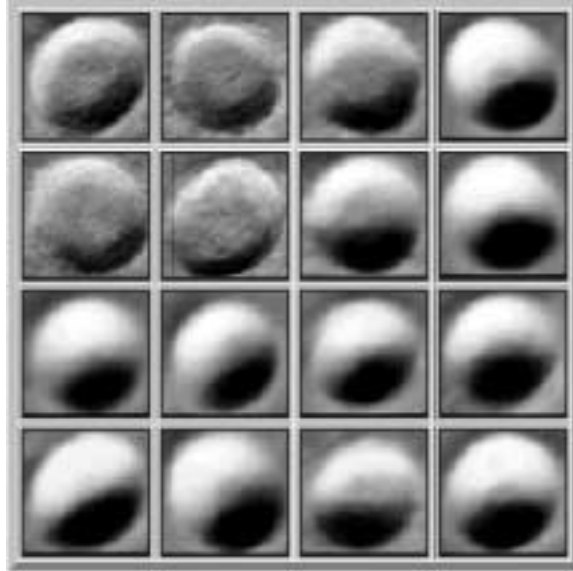


Figure 8: The same as in Figure 6, except that the learning included 170 craters in the northern hemisphere (ranging from 58 degrees N to 63 degrees N) of the moon.

Figure 8 shows the feature map obtained by SOM after 5000 iterations. The result indicate that a similar patterns in the feature map as those observed previously, including a clustering of bowl-shaped craters and flat-floored craters in the corner cells on the diagonal line.

By visually classifying the input images in each of the cells, we were able to group the cells into four super-categories as indicated by Figure 9 . The number of the cells in each category ranges from two to ten. Table 1 presents a semantical description of these four categories. Categories 1 and 2 are composed of flat-floored craters, and 50 % of category 1 and 15 % of category 2 are central-peaked craters, respectively. Category 3, the intermediate between categories 2 and 4, consists of cone-shaped craters. Category 4, which accounts for 72 % of all the craters, consists of bowl-shaped craters.

The similarity of the results from two independent data sets (Figures 6 and 8) suggests that the normalized image vectors are efficient as inputs for crater categorization problems. We have two other problems in the normalization: sun direction and solar altitude, both of which are known as the imaging conditions in most cases. We consider the sun direction problem to be resolved by the rotation of images before the cut out, and the use of a map with

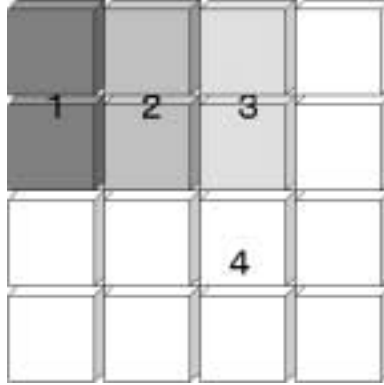


Figure 9: The super-categories in the feature map in figure 8. The semantical description is given in table 1.

more cells and grouping is effective with regard to the solar altitude problem.

Table 1 Semantical description of the crater categories.

Category	Number	Description	Percentage of central-peaked craters
1	19	Flat-floored.	50
2	13	Flat-floored.	15
3	16	Cone-shaped.	–
4	122	Bowl shaped.	–

## 4 Concluding Remarks

We examined the crater detection and categorization process in view of data mining from a large-scale scientific image database. We examined both the genetic algorithm and the combinational Hough transform as crater detection modules. Presently, CBT appears to be more robust for highly noisy lunar images, however, the optimization of various parameters for both methods should be considered in future studies.

We examined self-organizing mapping in relation to the categorization of craters, by taking normalized image vectors, intensity histograms, and DCT component, as input vectors, respectively. The experimental results show that normalized image vectors produce the most meaningful feature maps with recognizable simple bowl-shaped craters, cone-shaped craters, flat-floored craters and flat-floored craters with a central peak.

## 5 Acknowledgements.

The authors are grateful to Prof. O. Konishi for offering us the prototype of the SOM program and for our fruitful discussions. We would also like to thank to Dr. Y. Iijima for discussion, H. Yokogawa, S. Yamanaka, and T. Tsunemitsu for helps with the image processing. This research is supported by grants-in-aid for intensive research (A) (2) (Project 1130215) from the Ministry of Education, Science, and Culture of Japan.

## References

- Burl, M. C. et. al. (1998), Learning to Recognize Volcanoes on Venus Machine Learning ,30, 165–195.
- Burl, M. C. et al.(1999), Mining for Image Content, Systems, Cybernetics, and Informatics / Information Systems: Analysis and Synthesis, (Orlando, FL)
- Fayyad, U. M. et. al. (1996). Automating the Analysis and Cataloging of Sky Survey. *Advances in Knowledge Discovery and Data Mining*, (pp. 471–493). AAAI Press.
- Katayama, K. and O. Konishi.(1999). Construction of Satellite Image Databases for supporting Knowledge Discovery (in Japanese). *Information Processing Society of Japan -Transactions on Database*, Vol.40, No.SIG5(TOD2), 69–78.
- Kohonen, T. (1997). *Self-Organizing Maps*, 2nd eds., Springer.
- Melosh, H. J.(1988). *Impact Cratering*, Oxford University Press.
- Nozette, S. et. al. (1994). The Clementine Mission to the Moon: Scientific overview. *Science*, 266 , 1835–1839.
- Smyth, P. et. al. (1996). Modeling subjective uncertainty in image annotation . *Advances in Knowledge Discovery and Data Mining*, (pp. 517-539). AAAI Press.