

A Lunar Terrain Auto Recognition Algorithm with Adaptive Threshold by Gushing and Immersion*

Jiaheng Chen^{1, a}, Zhankai Li^{†1, b} and Xiaolin Tian^{1, 2, c}

¹Faculty of Information Technology, Macau University of Science and Technology, Macao 999078;

²Lunar and Planetary Science Laboratory/Space Science Institute, Macau University of Science and Technology, Macao 999078.

^achenjiaheng1992@126.com, ^bjanlzk@126.com[†], ^cxltian@must.edu.mo

Keywords: Lunar terrain, gushing and immersion, DEM and CCD image, mare and highland.

Abstract. As we know, the recognition of lunar terrain is the beginning of the lunar research. Therefore, an accurate and effective automatic recognition of lunar terrain is an integral part of lunar research. This paper has proposed an improved recognition algorithm for lunar terrain based on gushing and immersion. First of all, we identifying the lunar mare area and the highland area with four features used in k-means clustering on CCD image and DEM data. Next, the boundary of lunar mare and lunar highland will be extracted and calculate the average height on DEM image which had match to the CCD image. Then, we use the average value of boundary area as the threshold of gushing algorithm. Finally, the gushing and immersion algorithm will start at the minimal value of lunar mare area. The result will show the classification of lunar terrain. For the testing, the SI and Crisium had been chosen as two typical areas of lunar terrain. This improved algorithm can distinguish the lunar mare and highland automatically. The algorithm shows a better accuracy than other relevant algorithm.

1. Introduction

The moon is the earth nearest natural satellite. There is no denying that it is important to start the research of universe from the moon [1]. Lunar terrain is the surface of the moon which has rugged topography description like the terrain on earth. There are many various morphological features, distribution pattern and the course of development in it [2]. In order to study the geology and evolution history of the moon, the feature of topography and geomorphology on it should be extracted and analyzed for the research. Due to the differences of properties, material composition and formation years, it is obvious that the moon can also be classified to lunar mare and lunar highland [3]. During the research, we can found that the lunar mare and lunar highland also have plenty of different features in CCD and DEM images.

Gushing and immersion algorithm had been widely used on earth images. This technique usually used on the model of river and ocean. The idea of this algorithm is designed as the principle of the flow water. In image processing, the eight neighbor pixel of the starting point will be detected one by one and each neighbor of it will be saved as a new starting point. Based on the iteration of the process above, the image will be gradually marked according to the gray value. According to the principle of earth river identification, we can distinguish the lunar mare and lunar highland by the same principle.

2. An improved recognition algorithm for lunar terrain based on gushing and immersion

This section has been divided into three parts: the flowchart of the lunar terrain auto recognition algorithm, the rough classification based on feature extraction and clustering, gushing and immersion algorithm use the average value of boundary area as the threshold.

2.1 The flowchart of the new algorithm

The flowchart has been shown in the Fig. 1. The whole algorithm contains two sections. In section one; the algorithm determines the size of the initial block according to the accuracy of the input image

[8-9]. Then on the basis of the clustering result, direct output the classification results of block which has the high reliability classification. In section two, the algorithm will get an adaptive threshold from the average value of boundary area for gushing and immersion algorithm. The start point of gushing algorithm is the minimal value of the DEM data in lunar mare candidate.

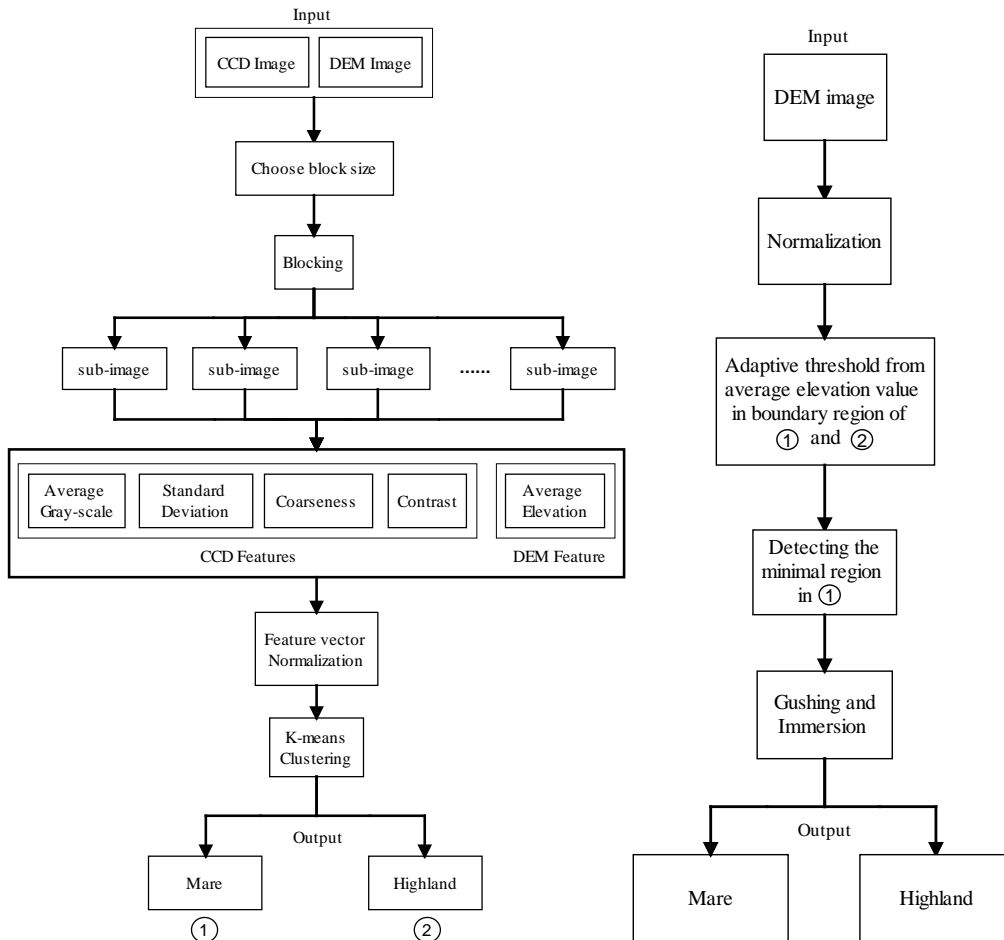


Fig 1 The flow chart of improved algorithm proposed

2.2 Section One: Rough Classification Based On Features

Some of significant features we had summarized according to the characteristics of lunar terrain. And in this part, four features of CCD image and one feature of DEM data used for clustering in this paper will be discussed [8]. They are the Average Gray-scale, the Standard Deviation of gray-scale, the Coarseness and the Contrast in CCD image, and the Average Elevation in DEM data.

2.2.1 The Average Gray-scale

In the lunar terrain CCD image, what is the most obvious difference between mare and highland? It's undoubted that the mare is much darker than the highland. Because of the main components of lunar mare are basalts, and the highland are Light-colored plagioclases [3], the reflectivity of these two main components is different. So, we inevitably choose the Average Gray-scale to describe this difference. Formula to calculate the Average Gray-scale has been shown below:

$$F_{AGS} = \frac{\sum_{i=1}^m \sum_{j=1}^n G(i,j)}{m \times n} \quad (1)$$

Where $G(i, j)$ is the gray-level of each pixel. The m is the length of image and n is the width of image.

2.2.2 The Standard Deviation

Standard Deviation of gray-scale is the most effective way in lunar terrain recognition. It is a measure of the degree of dispersion of a set of data's average value. When we get a large standard deviation, this means that there is a large difference between most value and its average, so the degree of dispersion is high. And a small standard deviation value means that there is a small difference

between most value and average value. It is obvious that the same average does not mean the same standard deviation. Therefore, we can use standard deviation to distinguish two different terrain areas which have the same average gray-scale. Formula to calculate the Standard Deviation has been shown below:

$$F_{SD} = \sqrt{\frac{\sum_{i=1}^m \sum_{j=1}^n (G(i,j) - \bar{G})^2}{m \times n}} \quad (2)$$

$$\bar{G} = F_{AGS} \quad (3)$$

2.2.3 Coarseness

Coarseness is a value, reflecting the granularity of the texture, which is the most basic texture feature [7]. When the only difference between two texture patterns is primitives, the one with larger size of primitive model gives more rough feeling. The calculation method is as follows [4]:

Step 1: Calculate the average gray-scale of $2^k \times 2^k$ pixels in active window.

In formula: $k = 0, 1, 2, 3, 4, 5$; $g(i, j)$ is gray-value which located in pixel (i, j) .

$$A_k(x, y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} g(i, j) / 2^{2k} \quad (4)$$

Step 2: For each point, take differences between pairs of averages corresponding to pairs of non-overlapping neighborhoods just on opposite sides of the point in both horizontal and vertical orientations.

$$E_{k,h} = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \quad (5)$$

$$E_{k,y} = |A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1})| \quad (6)$$

$E_{k,h}$ is the differences in the horizontal case, $E_{k,y}$ is the differences in the vertical case.

According to each pixel points, we can pick the best size which gives the highest output value. In the other words, we use k , the max value of E to configure the best size S_{best}

$$S_{best}(x, y) = 2^k \quad (7)$$

$$E_k = E_{max} = \max(E_1, E_2, E_3, \dots, E_h) \quad (8)$$

Step 3: Finally, F_{crs} can be calculated using the average value of S_{best} of the image, m and n is the width and height of the picture

$$F_{crs} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n S_{best}(i, j) \quad (9)$$

2.2.4 Contrast

Contrast refers brightness level difference between the brightest white in bright area and the darkest black in dark area in an image. The higher the range of pixel difference, the higher value of contrast you will get. Its size is decided by following factors: gray dynamic range, the polarization degree of the histogram in black and white, the sharpness of the edge and the cycle of the repetitive patterns. Generally, contrast usually refers to the first two factors. In this paper, we use the contrast which from Tamura [4]. The calculation as shown in formula:

$$F_{con} = \frac{\sigma}{\alpha_4^{1/4}} \quad (10)$$

It can be easily seen that the variance σ^2 or standard deviation σ about the mean of the gray-levels probability distribution is more preferable as contrast. As is well known, α_4 represents the peak value of image gray value, defined by

$$\alpha_4 = u_4 / \sigma^4 \quad (11)$$

And u_4 is the fourth moment about the mean, the $\sigma = F_{SD}$.

2.2.5 Average Elevation in DEM

In this paper, DEM data recorded surface elevation of each position on the moon. According to the common situation, the elevations of lunar mare always lower than it on lunar highland. Therefore, this

feature in DEM can provide a help to analysis lunar terrain. Here we use average elevation value as the fifth feature to help clustering. The formula has shown below:

$$H_{AGS} = \frac{\sum_{i=1}^m \sum_{j=1}^n D(i,j)}{m \times n} \quad (12)$$

2.2.6 K-means clustering

K-means method is used as the clustering rule in this paper. Given a set of observations (x_1, x_2, \dots, x_n) , where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k ($\leq n$) sets $S = \{S_1, S_2, \dots, S_k\}$ so as to minimize the within-cluster sum of squares (WCSS) (sum of distance functions of each point in the cluster to the K center). In other words, its objective is to find [5]:

$$\arg \min_S \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2 \quad (13)$$

where μ_i is the mean of points in S_i .

2.3 Section Two: Gushing and Immersion

The idea of this algorithm is designed as the principle of the flow water. In image processing, the eight neighbor pixel of the starting point will be detected one by one and each neighbor of it will be saved as a new starting point. Based on the iteration of the process above, the image will be gradually marked according to the gray value. The diffusion principle of eight neighborhoods has shown below in Fig. 2.

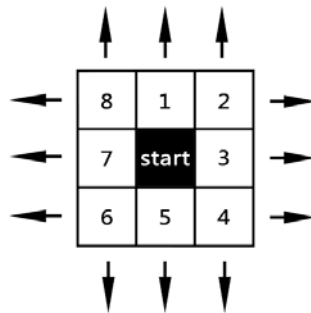


Fig 2 The principle of eight neighborhoods.

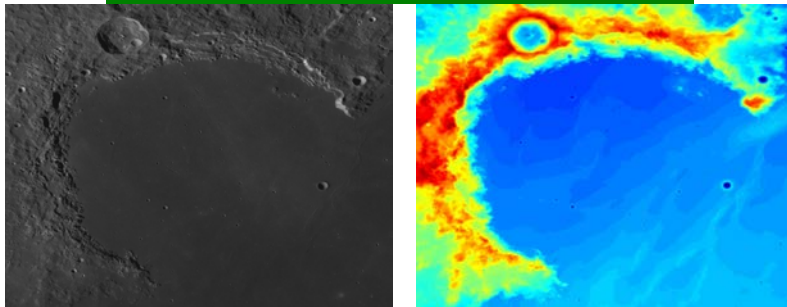
In this paper, the start point of gushing algorithm is the minimal value of the DEM data in lunar mare candidate. These start point are extracted from lunar mare areas based on rough classification. Therefore, the start point won't appear on the highland areas.

Next, the boundary of lunar mare and lunar highland will be extracted and calculate the average height on DEM image which had match to the CCD image. And each block of these two categories terrain from Section one will determine whether its four neighboring blocks are belonging to the category of each other. Then, this algorithm will produce an adaptive threshold water level from the average value of boundary area for gushing and immersion algorithm.

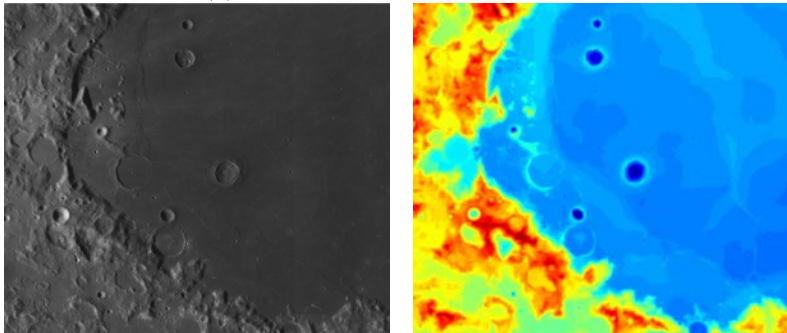
3. Testing result

There are two testing areas used in this paper: 'SI' area and 'Crisium' area which from NASA and show in Fig 3. These CCD and DEM data had already matched of two areas. They are selected manually in order to compare with the previous algorithm [6-7]. The test DEM data is 500 meters per point and the test CCD data is 500 meters per pixel. First, we set the block size is 20×20 pixels. Second, due to the problem of DEM data format, we need to normalize the DEM data at first, and set the range of minimal value is lower than 0.1. And the test results are shown in Fig 4 and Fig 5.

Name	Longitude	Latitude
SI	26°W~39°W	40°N~50°N
Crisium	47.5°E~60°E	9.3°N~20°N



(a) The CCD and DEM of 'SI'



(b) The CCD and DEM of 'Crisium'

Fig 3 The testing areas

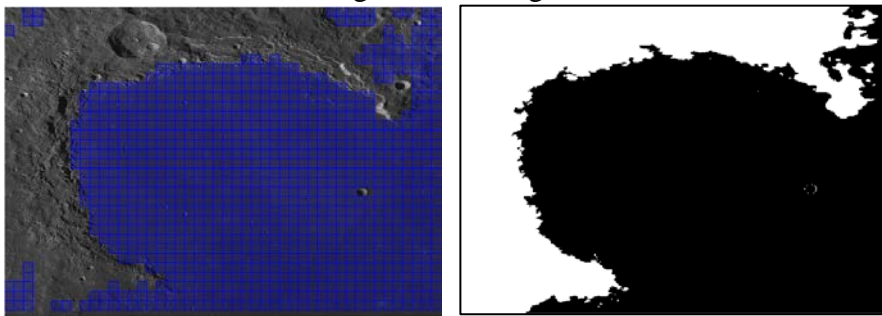


Fig 4 The test results of 'SI'

And comparing with the area of artificial interpretation data that from USGS, the recognition rate can be calculated.

Mare Recognition rate	Highland Recognition rate
99.2331	96.4489

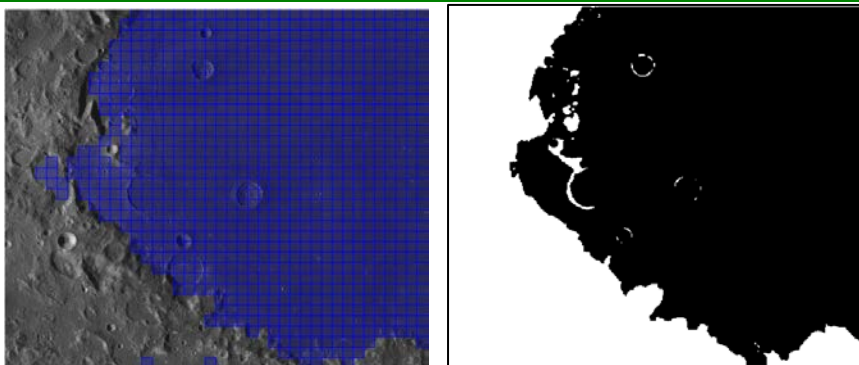


Fig 5 The test results of 'Crisium'

Table 3 The accuracy of Crisium area

Mare Recognition rate	Highland Recognition rate
96.8715	96.8391

4. Summary

From the testing results, we can get some conclusion:

There is a very high rate on recognizing mare and highland in ‘IS’ and ‘Crisium’ area. In order to evaluate the result and make sure the high accuracy is not an accident, so we had calculated the recognition rate of mare and highland. It shows that the value of mare recognition rate and highland recognition rate for ‘SI’ area is 99.2331% and 96.4489%, the value of mare recognition rate and highland recognition rate for ‘Crisium’ area is 96.8715% and 96.0025%. Also the recognition rate of ‘SI’ area is higher than the value given by other previous algorithms [6-7].

However, This algorithm also has some deficiencies.

(1) Error in recognition can be found easily in the area where the elevation value has no obvious differences.

(2) Can not be accurately identified in high-resolution images. Because of there are some gulches existed in the boundary area of highland and mare, if we use gusing and immersion algorithm, it will flow into the gulch. So, some areas of highland will be error identification.

And we will try to solve those issues of the algorithm used in this paper and improve it in the future.

*This work is supported by the Science and Technology Development Fund of Macao (No. 059/2013/A2).

Reference

- [1]. Ziyuan OuYang, Introduction to lunar science (in Chinese). Beijing: China Aerospace Press, 2005.
- [2]. Ziyuan Ouyang. Lunar geology [J]. Advances in Earth Sciences. 1994, 9(2):80-81.
- [3]. Wenzhe Fa, Theoretical Modeling and Parameters Inversion for Microwave Remote Sensing of Lunar Surface (in Chinese), Ph.D. Thesis. Shanghai: School of Information Science, Fudan University, 2008.
- [4]. Tamura, H., Mori, S., & Yamawaki, T. (1978). Textural features corresponding to visual perception. Systems, Man and Cybernetics, IEEE Transactions on, 8(6), 460-473.
- [5]. JA Hartigan, MA Wong, Algorithm AS 136: A k-means clustering algorithm, Applied statistics, 1979-JSTOR.
- [6]. Chenzhi Wang, et al, A Method for Identifying the Lunar Morphology Based on Texture from DEMs (in Chinese), Vol.17, No.1. Jan., 2015.
- [7]. Xie T L, Jiang H K, Wang J L, et al. A new recognition algorithm of the lunar mare area basing on the DEM con-contrast [C]. International Conference on Advanced Materials and Engineering Structural Technology, Qingdao, China, April 25-26, 2015.
- [8]. Zhankai Li. An improved recognition algorithm for lunar terrain based on CCD image. ICMCT 2016. p 1581-1587.
- [9]. Jiang H K, Tian X L and Xu A A. A new segmentation algorithm for lunar surface terrain based on CCD image [J]. Research in Astronomy and Astrophysics. 2014, 15(9): 1-10.