

Enhanced Lunar Topographic Mapping Using Multiple Stereo Images Taken by Yutu-2 Rover with Changing Illumination Conditions

Wenhui Wan, Jia Wang, Kaichang Di, Jian Li, Zhaoqin Liu, Peng Man, Yexin Wang, Tianyi Yu, Chuankai Liu, and Lichun Li

Abstract

In a planetary-rover exploration mission, stereovision-based 3D reconstruction has been widely applied to topographic mapping of the planetary surface using stereo cameras onboard the rover. In this study, we propose an enhanced topographic mapping method based on multiple stereo images taken at the same rover location with changing illumination conditions. Key steps of the method include dense matching of stereo images, 3D point-cloud generation, point-cloud co-registration, and fusion. The final point cloud has more complete coverage and more details of the terrain than that conventionally generated from a single stereo pair. The effectiveness of the proposed method is verified by experiments using the Yutu-2 rover, in which two data sets were acquired by the navigation cameras at two locations and under changing illumination conditions. This method, which does not involve complex operations, has great potential for application in planetary-rover and lander missions.

Introduction

In a planetary-rover exploration mission, topographic mapping plays a crucial role to support science and engineering operations. It provides fundamental terrain information for topographic analysis, obstacle detection, path planning, rover navigation, and so on (Li *et al.* 2005; Alexander *et al.* 2006; Olson *et al.* 2007; Di *et al.* 2020). Thus, the performance of topographic mapping largely determines the safety of the rover's travel and its effectiveness in fulfilling scientific exploration tasks.

Due to their low power consumption, rich information acquisition, and reliable performance, stereo cameras have been used as the primary payload for topographic mapping in rover missions. The captured stereo images are usually transmitted from the rover to the ground-control station and inputted to the data-processing system for image preprocessing, 3D reconstruction, visual localization, and further scientific analysis. This off-line processing methodology has been successfully applied in Mars and lunar rover missions, such as the Mars Exploration Rover missions and Mars Science Laboratory mission (Li *et al.* 2005; Alexander *et al.* 2006; Matthies *et al.* 2007; Di *et al.* 2008; Stein *et al.* 2019) and Chang'e-3 and Chang'e-4 rover missions (Z. Liu *et al.* 2015, 2020; Di *et al.* 2020). During such a rover mission, topographic mapping is performed routinely with stereo images captured at the

waypoints along the rover traverse. The topographic mapping products are used not only in obstacle detection and path planning but also in measurement and designation of science targets for detailed in situ exploration.

The rover topographic mapping techniques used in previous missions are basically based on stereovision methods (Matthies *et al.* 2007). Image matching is the core technique in stereovision and has been extensively researched in the fields of photogrammetry and computer vision. Image-matching methods can be broadly classified as sparse or dense matching. Sparse matching methods match distinguished features (e.g., interest points; Gruen 1985; Förstner and Gülch 1987; Harris and Stephens 1988) to provide high-precision but sparse correspondence between the stereo images. Dense matching methods provide a pixel-by-pixel disparity map for dense 3D point-cloud generation (Scharstein and Szeliski 2002; Zhu *et al.* 2007), and have been applied in Mars and lunar rover exploration missions successfully (Alexander *et al.* 2006; Matthies *et al.* 2007; Di *et al.* 2008; Z. Liu *et al.* 2015, 2020). Recently, Semi-Global Matching and related optimized methods (Hirschmuller 2008; Michael *et al.* 2013; Besse *et al.* 2014) have become popular the vision-based 3D mapping applications because of their improved performance. Without requiring a manually designed model for image-feature extraction, deep learning-based dense matching methods have been proposed, using a large amount of proper training data (Luo *et al.* 2016; J.-R. Chang and Chen 2018). To overcome the limitation of the hard baseline of the stereo camera onboard a planetary rover, wide-baseline methods with image acquisition at two or more waypoints were developed for topographic mapping with higher accuracy for farther ranges (Olson *et al.* 2003; Di and Peng 2011; Wu *et al.* 2011).

Rover imagery of the lunar surface is often of poor texture, caused by the featureless surface itself or the severe influence of changes in illumination conditions due to the lack of atmosphere. This poses great challenges to the precision and completeness of topographic mapping products from stereo pairs of images. Recently, some illumination-invariant matching methods (Xu *et al.* 2016; Y.-J. Chang and Ho 2017) have been developed for urban or indoor scenes. These methods may not be directly applicable in matching of lunar rover images with different illumination conditions, though. Shading information associated with different illumination conditions has been exploited to produce lunar topographic products from orbital images using a photometric stereo technique (W. Liu *et al.* 2018; W. Liu and Wu 2020). The model and techniques are

Wenhui Wan, Kaichang Di, Zhaoqin Liu, Peng Man, and Yexin Wang are with the State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China (dikc@radi.ac.cn).

Jia Wang, Jian Li, Tianyi Yu, Chuankai Liu, and Lichun Li are with Beijing Aerospace Control Center, Beijing, China.

Contributed by Norbert Pfeifer, August 7, 2020 (sent for review January 21, 2021; reviewed by Wai Chung Liu, Cristina Re).

Photogrammetric Engineering & Remote Sensing
Vol. 87, No. 8, August 2021, pp. 567–576.
0099-1112/21/567–576

© 2021 American Society for Photogrammetry
and Remote Sensing
doi: 10.14358/PERS.87.8.567

also not directly applicable to lunar rover images, which are close-range perspective views.

During surface operations in a lunar rover mission—e.g., the *Chang'e-4* mission—the rover often stops and stays at the same location for quite a long time, for in situ exploration of the science targets or to hibernate in the extremely cold lunar night. If the rover took stereo images at the same location but at different times with different illumination conditions, would the multiple stereo images be helpful in producing better topographic products than a single stereo pair? This is the question that inspired us to begin this research.

This article proposes a novel and easy-to-implement method of rover-based lunar topographic mapping using multiple stereo images taken at the same location but with changing illumination conditions. The effectiveness of the proposed method is verified using stereo images acquired by the Yutu-2 rover. The enhanced lunar topographic mapping results, which have higher precision and more details, can be used in missions to support rover exploration planning and decision making. The proposed method is particularly useful when we want to get a higher-quality 3D model of the exploration area and science target but the rover cannot get closer to the target because of traversing constraints. It is also directly applicable to a lunar sample-return mission—e.g., the planned *Chang'e-5* mission—in which the stereo cameras are fixed on the lander, to provide enhanced mapping of the sampling areas.

The rest of the article is organized as follows. The next section elaborates the methodology and describes the unique image data sets acquired by the Yutu-2 rover for verification of the proposed topographic mapping method. In the section after that, the experimental results and detailed performance analysis based on the rover images are reported. Conclusions and future work are discussed in the last section.

Methodology and Data

The main idea on which the method is proposed is that multiple stereo images taken at the same location and the same attitude but with different illumination conditions (i.e., different solar azimuth and elevation angles) can provide more information about the terrain than any single stereo pair in the data set. As we know, it is usually not possible to reconstruct the 3D terrain in shadow areas from one stereo pair through image matching. Multiple stereo images with different illumination conditions can solve this problem because shadow areas in one stereo pair may become illuminated in other stereo pairs. For a terrain patch not in shadow area, i.e., visible in all stereo images, the image-matching results of the same image pixel in the left images will be slightly different through subpixel matching (e.g., the least-squares matching method), because of the microstructural changes in visual appearance of the same patch in different stereo pairs. Consequently, the 3D coordinates of the same point in different stereo pairs will be different. Taking these 3D points as the repeated observations of the same physical point and merging them based on the measurement uncertainties will result in a 3D point with higher precision. Our proposed method aims to take advantage of multiple stereo images with different illumination conditions to enhance the capability of rover-based lunar topographic mapping.

The proposed lunar topographic mapping method, using stereo images captured in different illumination conditions, consists of three major steps: dense 3D point-cloud generation from each stereo pair, co-registration of multiple 3D point clouds, and fusion of surface point clouds. As shown in Figure 1, each pair of stereo images is inputted individually for dense matching and generation of 3D point clouds. Then, co-registration of multiple point clouds is performed through a global Iterative Closest Point (ICP) algorithm under the principle of 2D and 3D weighted-distance minimization.

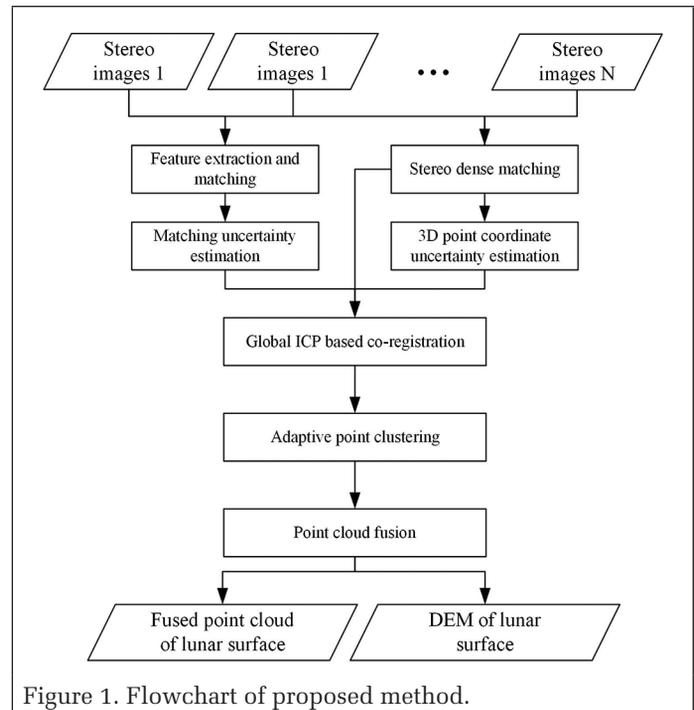


Figure 1. Flowchart of proposed method.

Theoretically, if the positions and attitudes of the multiple stereo pairs are exactly the same, this co-registration step can be skipped. But in practice, co-registration is necessary due to slight inconsistencies among the actual orientations of the stereo images caused by the control and measurement uncertainties of rover mast drives. Later, the co-registered 3D point clouds are fused into a merged surface point set with adaptive point clustering based on position consistency. Finally, the topographic products are provided in two forms: the fused point-cloud data, and the digital elevation model (DEM) generated by regular grid interpolation. The details of the proposed method follow.

Multiple Stereo-Image Acquisition by Yutu-2 Rover at the Same Location

China's *Chang'e-4* probe successfully landed on the far side of the moon in the Von Kármán crater on 3 January 2019, and the rover Yutu-2 was released from the lander and started surface exploration on the same day (Di *et al.* 2019). The Yutu-2 rover carries three pairs of stereo cameras: the navigation camera (Navcam), the hazard-avoidance camera (Hazcam), and the panoramic camera (Pancam; Z. Liu *et al.* 2020). The Navcam and Hazcam are engineering sensors designed for navigation and obstacle avoidance, and have been used to acquire images at every waypoint. The Pancam is one of the science payloads and has been used at certain waypoints for long-term planning of scientific exploration. The Hazcam is fixed on the front of the rover with a downward angle of 30°. The Navcam and Pancam are mounted on the same camera bar atop the rover mast and can acquire stereo images pointing to different azimuth and elevation angles by rotations through the yaw and pitch drives. Controlled by the deployment drive, the camera mast is deployed to vertical position when working in a lunar day and stowed to horizontal position when hibernating in the lunar night. Before the *Chang'e-4* probe was launched, the interior orientation parameters and distortion-model parameters of the cameras were calibrated precisely in the calibration field containing evenly distributed 3D control points. The installation parameters of each camera and mast joints related to the rover body were also calibrated by cube prism-based measurements. The exterior orientation parameters of Navcam and Pancam images related to the rover body coordinate system can be calculated with the parameters of mast installation and yaw/pitch rotations.

In this study, in order to verify the proposed enhanced topographic mapping method, we designed the data-acquisition strategy and implemented it on the eighth, ninth, and 10th lunar day at the ground-control station, Beijing Aerospace Control Center. Specifically, three pairs of Navcam images were captured at waypoint LE01002 in the morning of the 10th lunar day. The captured stereo images and their detail views are shown in Figure 2.

The second data set contains three pairs of Navcam stereo images captured at waypoint LE00803 in the afternoon of the eighth lunar day and one pair in the morning of the ninth lunar day. The images are shown in Figure 3. The solar azimuth angles and elevation angles of the two image data sets in the local-site coordinate system are listed in Table 1.

As can be seen in the images of the two data sets, the areas of the shadowed image patches became larger with the decrease of the solar elevation angle. We note also that image pair 2d was captured with the opposite solar azimuth angle from image pairs 2a to 2c. Therefore, the shadow areas in image pairs 2a to 2c were imaged effectively through image pair

2d, which should be particularly helpful in reconstructing the lunar surface more completely.

Dense Matching of Stereo Images and 3D Point-Cloud Generation

In our proposed method, dense matching is adopted for each pair of stereo images captured by the rover in order to obtain the 3D point cloud. Before the dense matching, it is necessary to perform image enhancement to improve the image quality, which is helpful for getting more matched points. We use methods of grayscale stretch and histogram specification to enlarge the image contrast of a single image and decrease the histogram differences between left and right images. Then the pinhole-based distortion model and calibrated relative exterior orientation parameters of the Navcam are adopted for epipolar rectification (Barnard and Fischler 1982).

The pinhole-based distortion correction model for Navcam imagery is

$$\begin{cases} x'_d = x'_s + x'_s(k_1r^2 + k_2r^4 + k_3r^6 + \alpha) + \beta y'_s + p_1(r^2 + 2x'_s y'_s) + 2p_2 x'_s y'_s \\ y'_d = y'_s + y'_s(k_1r^2 + k_2r^4 + k_3r^6) + 2p_1 x'_s y'_s + p_2(r^2 + 2y'^2_s) \end{cases} \quad (1)$$

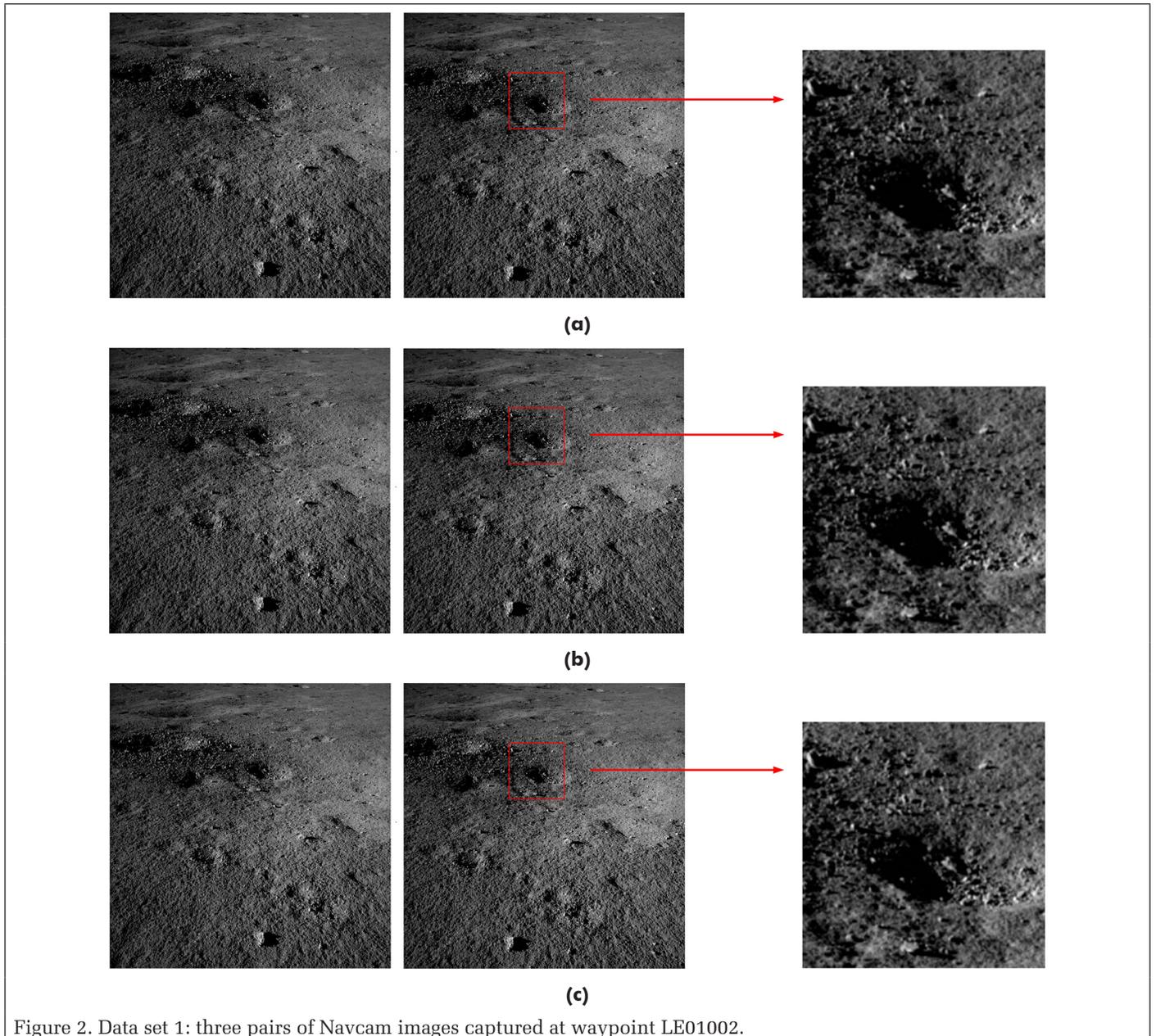


Figure 2. Data set 1: three pairs of Navcam images captured at waypoint LE01002.

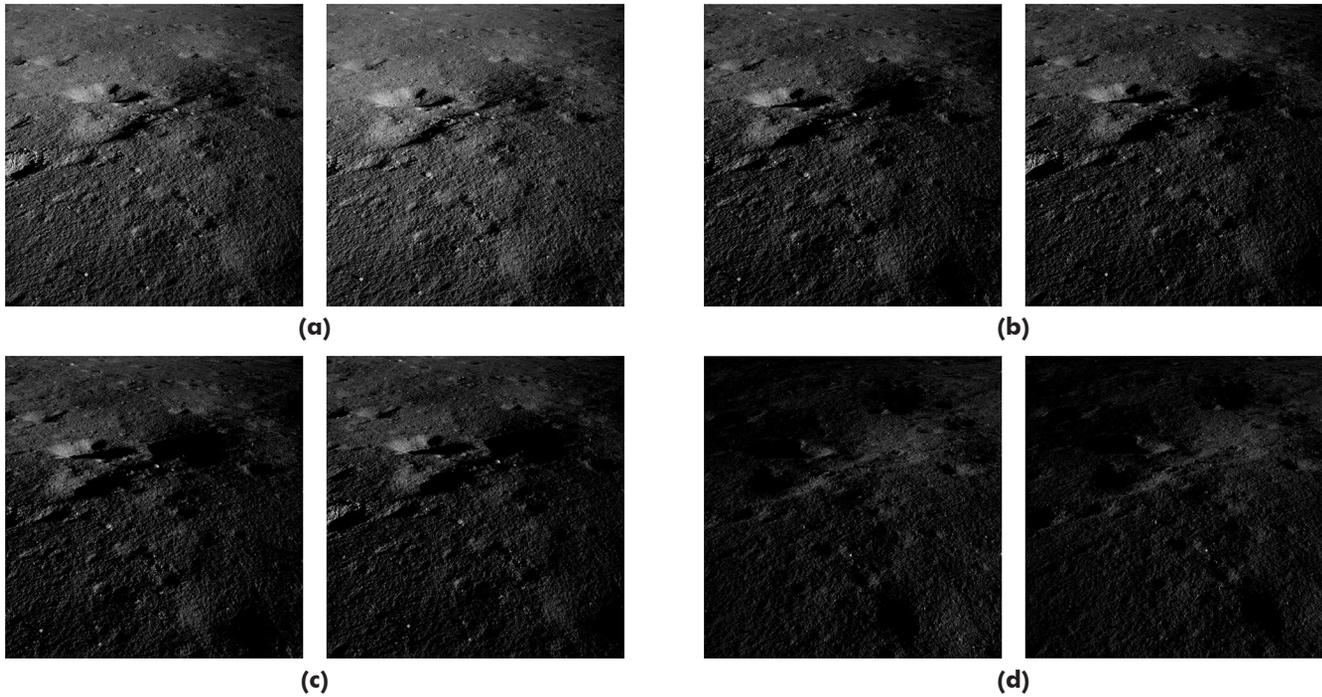


Figure 3. Data set 2: four pairs of Navcam images captured at waypoint LE00803.

where (x'_s, y'_s) and (x'_d, y'_d) are the coordinates of the original point and the corrected point in the principal point-centered image-plane coordinate system in millimeters; k_1, k_2, k_3 are the radial distortion coefficients; p_1, p_2 are the tangential distortion coefficients; r^2 is a coefficient calculated by $r^2 = x'_s + y'_s{}^2$; α is the non-square scale factor of the charge-coupled device; β is the non-orthogonal distortion coefficient of the charge-coupled device, and the point coordinates (x', y') in the image-plane coordinate system in millimeters are calculated by

$$\begin{cases} x' = (x - x_0) \cdot p_s \\ y' = (y - y_0) \cdot p_s \end{cases} \quad (2)$$

where (x, y) are the image coordinates in pixels, with the origin centered at the top left corner of the image; (x_0, y_0) are the principal-point coordinates in pixels; and p_s is the pixel size in millimeters.

Semi-Global Matching (Hirschmuller 2008) is applied to the generation of the disparity map by scanning the image pixel with the minimum in energy function along the x-axis of the epipolar rectified images. The energy function is represented as

$$E(D) = \sum_{\mathbf{p}} \left(C(\mathbf{p}, D_{\mathbf{p}}) + \sum_{\mathbf{q} \in N_p} P_1 T[|D_{\mathbf{p}} - D_{\mathbf{q}}| = 1] + \sum_{\mathbf{q} \in N_p} P_2 T[|D_{\mathbf{p}} - D_{\mathbf{q}}| > 1] \right) \quad (3)$$

where $E(D)$ is the sum of all pixel matching costs for the disparities of image D ; $C(\mathbf{p}, D_{\mathbf{p}})$ is the matching cost of pixel \mathbf{p} with disparity $D_{\mathbf{p}}$; \mathbf{q} is the pixel set in neighborhood N_p of \mathbf{p} ; and P_1 and P_2 are the penalty coefficients for the pixels with disparity changes in \mathbf{q} of, respectively, one pixel and more than one pixel. Considering that the grayscale distributions of the same physical point in the left and right images are still slightly different, due to different observation angles and different responses of the left and right cameras, the ADCensus algorithm is used for calculating the matching cost (Mei *et al.* 2011). It is expressed as

Table 1. Solar azimuth angles and elevation angles of the two stereo-image data sets.

Stereo Image No.	Solar Elevation Angle (°)	Solar Azimuth Angle (°)	Acquisition Time (UTC+8:00)
1a	18.3	67.9	2019-09-24 09:36:02
1b	19.2	66.9	2019-09-24 11:53:28
1c	20.0	65.9	2019-09-24 14:19:27
2a	16.2	-71.8	2019-08-06 17:45:04
2b	11.4	-77.8	2019-08-06 20:07:27
2c	9.9	-78.7	2019-08-07 09:54:51
2d	7.8	80.4	2019-08-24 14:19:22

$$C_{\text{ADCensus}}(\mathbf{p}, D_{\mathbf{p}}) = \rho(C_{\text{Census}}(\mathbf{p}, D_{\mathbf{p}}), \lambda_{\text{Census}}) + \rho(C_{\text{AD}}(\mathbf{p}, D_{\mathbf{p}}), \lambda_{\text{AD}}) \quad (4)$$

where C_{Census} and C_{AD} are the cost values from the census transform and absolute differences, respectively; λ is the pre-set coefficient of cost estimation; and $\rho(c, \lambda)$ is a function on variable c represented as

$$\rho(c, \lambda) = 1 - \exp\left(-\frac{c}{\lambda}\right) \quad (5)$$

The census cost C_{Census} is computed as the Hamming distance of the two bit strings, which are formed from sequential judgments whether the grayscale value of the centered pixel is higher or lower than that of each neighboring pixel. Then the disparity map is obtained by means of minimizing $E(D)$ based on dynamic programming. Note that the disparity map only contains the disparity values in the horizontal direction of the image, because of epipolar rectification. Finally, least-squares matching (Gruen 1985) is applied to refine the disparity map to subpixel values. After dense matching, the 3D coordinates of matched points are calculated by space intersection with the exterior orientation parameters of the stereo images. As a result, 3D point-cloud data are generated for each stereo pair of images.

Co-registration of Multiple Point Clouds

Due to the control and measurement uncertainties of the rover's mast drives that deploy the mast or rotate the camera bar, inconsistencies may exist among the 3D point clouds generated from the stereo Navcam images, even if they were acquired at the same waypoint with the same nominal pointing directions. Thus, point-cloud co-registration is performed to obtain the relative parameters of rotation and translation. ICP is a useful method for point-cloud registration by minimizing the mean square distance of two point clouds without feature point matching (Besl and McKay 1992). However, because of flat terrain in some exploration areas, the registration solution may fall into a local minimum. Considering that the images captured under changing illumination conditions still have a certain grayscale similarity, 2D feature points are extracted and matched among different stereo images, so as to constrain the global ICP registration.

To obtain the matched points robustly, the scale-invariant feature transform (Lowe 2004) is used to extract and match the feature points in the left images of stereo pairs. Then the corresponding points in the right images are matched by the least-square matching method. Outlier detection is performed to eliminate incorrectly matched points. Finally, the left image is divided into evenly spaced grids, and the point in each grid with the highest correlation coefficient is chosen to get evenly distributed matched points within and among stereo pairs. Afterwards, the geometric constraints between stereo pairs, constructed by the feature matched points, will be incorporated into the global ICP solution.

In our global ICP solution, we also consider the uncertainties of the 3D points by calculating weighted 3D distances. According to the principle of photogrammetry (Di and Li 2007), the 3D coordinate uncertainties of object point P can be estimated by

$$\begin{cases} \sigma_x = \sqrt{\left(\frac{Z_p^2}{Bf}\right)^2 \left(\frac{x}{f}\right)^2 \sigma_d^2 + \left(\frac{Z_p}{f}\right)^2 \sigma_x^2} \\ \sigma_y = \sqrt{\left(\frac{Z_p^2}{Bf}\right)^2 \left(\frac{y}{f}\right)^2 \sigma_d^2 + \left(\frac{Z_p}{f}\right)^2 \sigma_y^2} \\ \sigma_z = \frac{Z_p}{Bf} \sigma_d \end{cases} \quad (6)$$

where $(\sigma_x, \sigma_y, \sigma_z)$ are the 3D errors (standard deviations) of point P ; B and f are the baseline and focal length, respectively, of the stereo cameras; Z_p is the z-axis (perpendicular to the image plane) coordinate of P in the camera reference coordinate system; (x, y) are the 2D coordinates of the corresponding image point p ; (σ_x, σ_y) are the 2D measurement errors of p ; and σ_d is the error of horizontal disparity of p , represented by the stereo-matching error of p and calculated as (Zhang and Zhang 1997)

$$\sigma_d = \sqrt{\frac{2(1-\rho^2)}{N} \cdot \left(\frac{\sigma_g^2}{\sigma_g'^2}\right)} \quad (7)$$

where ρ is the correlation coefficient of the matched points, N is the number of pixels in the matching window, σ_g is the grayscale variance in the matching window, and σ_g' is the variance of first-order grayscale difference. By combining Equations 6 and 7, the 3D coordinate uncertainties of the point cloud are estimated for the following co-registration.

In our global ICP solution, the optimization principle is built on minimized 2D and 3D weighted distance, which is expressed as

$$L(R, T) = a \cdot D_{2D}(R, T) + b \cdot D_{3D}(R, T) \quad (8)$$

where R and T are the rotation and translation parameters between the reference point cloud P_R and the point cloud to be registered P_C ; a and b are the preset weight coefficients; $D_{2D}(R, T)$ is the 2D mean weighted distance of the matched feature points, expressed as

$$D_{2D}(R, T) = \frac{1}{M} \sum_{i=1}^M m_i^R - F_p(m_i^C, R, T) \quad (9)$$

and $D_{3D}(R, T)$ is the mean weighted 3D distance, which is expressed as

$$D_{3D}(R, T) = \frac{1}{N} \sum_{i=1}^N \omega_i [p_i^R - (R \cdot p_i^C - T)] \quad (10)$$

where M is the number of feature matched points; m_i^R and m_i^C are the i^{th} matched points in the left images of the reference stereo frame I_R and the co-registered frame I_C ; F_p is the projection function that transforms the coordinates in I_C to in I_R ; N is the pair number of the closest points between P_R and P_C ; and ω_i is the distance weight between the corresponding 3D point p_i^R and p_i^C , which is

$$\omega_i = \frac{2}{\sigma_i^R + \sigma_i^P} \quad (11)$$

where σ_i^R and σ_i^P are the errors of points p_i^R and p_i^C obtained by Equations 6 and 7. By minimizing Equation 8, the rotation and translation parameters are solved to complete the co-registration.

Fusion of Multiple Point Clouds

The co-registered point clouds are considered as the multiple observations of the lunar surface in different illumination conditions. The point clouds generated from the stereo images are then fused to obtain a density point cloud to better represent the lunar surface. In the data-fusion process, each point from all of the point clouds of single stereo pairs is inputted to the surface point set along with the point uncertainty. Point clustering is performed for all points one by one. As a result, each cluster contains a seed point having the lowest uncertainty and other points associated to the seed point as repeated observations. Then by merging the points in each cluster to one point using the weighted average, the final fused surface point set is obtained for enhanced topographic mapping. The workflow of multiple point-cloud fusion is shown in Figure 4.

Since the stereo matched image points are the original source of the 3D points, the estimation of position consistency is performed in the image space. All 3D points of the generated point clouds are projected to the left image space of the stereo pair which has the most matched points, to obtain the corresponding 2D coordinates. All the points are sorted by the 2D matched uncertainties in increasing order to obtain the ordered point set, which is defined as

$$\mathbf{P}_0 = \{p_i(i, l, X_{2D}, X_{3D}, \sigma_{2D}, \sigma_{3D}, C_{ID}) \mid i \in (1, m), l \in (1, n)\} \quad (12)$$

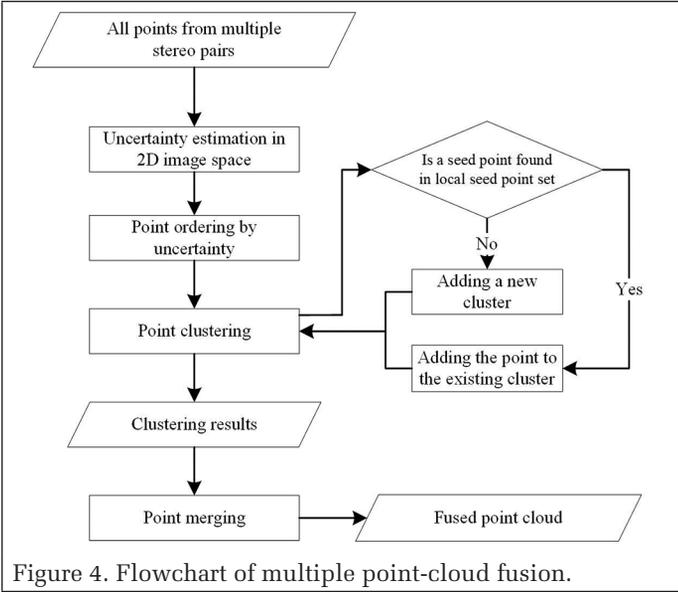


Figure 4. Flowchart of multiple point-cloud fusion.

where i is the point index in the point set; l is the point-cloud index; X_{2D} and X_{3D} are the 2D and 3D coordinates, respectively; σ_{2D} and σ_{3D} are the 2D and 3D uncertainties of the point calculated by Equations 7 and 6; m is the total number of points; n is the number of point clouds; and C_{ID} is the cluster index, which is set to i by default. Then each point p_i in P_o is moved by the ascending order of the uncertainties to the initially empty surface point set P_s by the uncertainty clustering. In the process of point clustering, p_i 's corresponding local seed point set from P_s , which is defined as $P_N = \{p_j | j \in (1, k)\}$, is obtained by searching in the neighboring area centered at X_{2D}^i of P_s with the radius of σ_{2D}^i . If p_i and p_j have the minimal weighted distance L , which is calculated by

$$L = \sigma_{2D}^i + X_{2D}^i - X_{2D}^j \quad (13)$$

and $l_i \neq l_j$, p_i is added to P_s by setting C_{ID}^i to C_{ID}^j as a new member of the cluster containing p_j . Otherwise, p_i is used to construct a new cluster as the seed point. Due to the small search radius σ_{2D}^i (typically 0.2 pixel), the points in one cluster are close to each other, which ensures that the clustered points are repeated observations of the same physical point from multiple stereo pairs. After each point is clustered in P_o iteratively, the optimally clustered point set P_s is obtained. Then each point cluster containing more than one point is merged to an optimally estimated point p_f . The 3D coordinates X_{3D}^f of the merged point can be obtained by

$$X_{3D}^f = \sum_{h=1}^N w_h \cdot X_{3D}^h \quad (14)$$

where X_{3D}^h is the 3D coordinates of cluster point; N is the point number of the cluster; and w_h is the weight coefficient, calculated by

$$w_h = \frac{1}{\sigma_{3D}^h} \bigg/ \sum_{c=1}^N \left(\frac{1}{\sigma_{3D}^c} \right) \quad (15)$$

According to measurement-error theory, the weighted average of the repeated 3D point measurements will have a higher precision than a single measurement. By processing all clusters and adding all the new points, the fused surface point set is obtained. This point cloud is optimal in that it has a higher

density and higher precision than the point cloud generated from a single stereo pair.

Experimental Results

Topographic Mapping Results

The experiments were performed using the two data sets already described. First, each image pair was used to generate a point cloud. Then the fused point sets of the lunar surface were obtained by applying the methods of point-cloud co-registration and fusion. Figure 5 shows the generated point clouds of the first data set and the detailed view of the point cloud corresponding to the image areas shown in Figure 2, with Figure 5d and 5e showing the fused surface point sets that have denser points than the single point cloud produced by one stereo pair of images. Figure 6 shows the reconstruction results of the second data set. The point numbers of the single stereo point clouds and the fused clouds can be found in Table 2.

From Figures 5 and 6 and Table 2, we can observe that in both data sets, the fused clouds have many more points than the point clouds generated from single stereo images, providing enhanced topographic mapping capability. Note that image pair 2d was captured with the opposite solar azimuth angle to the other three stereo pairs in data set 2, and thus the shadow areas in those pairs were imaged effectively through image pair 2d. This is particularly helpful for reconstructing the lunar surface with more complete coverage, as shown in Figure 6e.

DEM Evaluation

To further evaluate the performance of the proposed method, we applied statistical analysis of the DEMs generated in the experiments. For convenience of comparison, two test areas were selected in each data set, at close and far range from

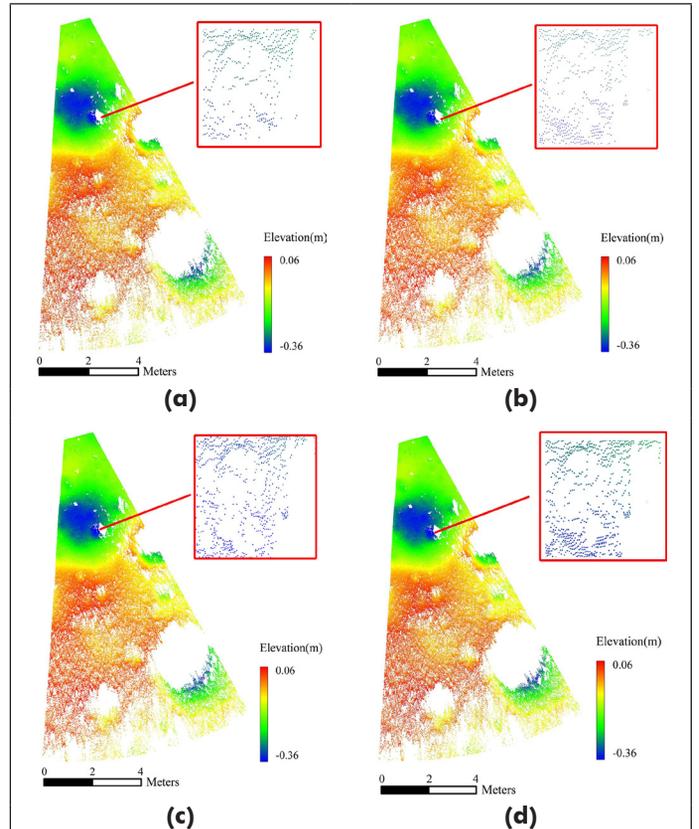


Figure 5. The dense point clouds of the first data set: (a to c) Navcam image pairs 1a, 1b, and 1c, respectively; (d) the fused point set of the surface.

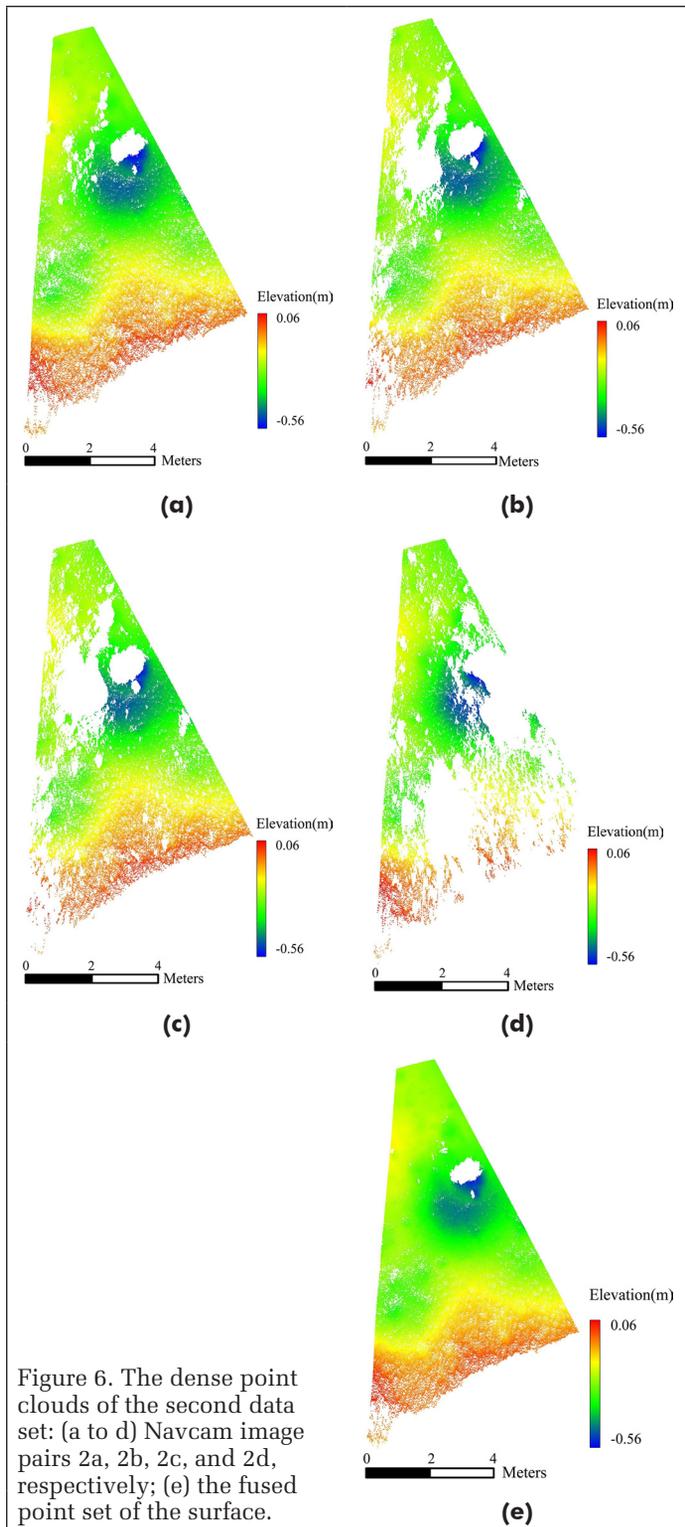


Figure 6. The dense point clouds of the second data set: (a to d) Navcam image pairs 2a, 2b, 2c, and 2d, respectively; (e) the fused point set of the surface.

Table 2. Point numbers of the point clouds.

Data Set	Point Cloud	Number of Points
1	1a	616 428
	1b	629 842
	1c	644 015
	Fused result	824 815
2	2a	634 784
	2b	401 470
	2c	331 660
	2d	266 062
	Fused result	910 187

the rover. Regular grid DEMs, interpolated with the points in selected areas, were used to perform the quality analysis using the statistical measures of sharpness and mean gradient. In evaluating the first data set, two areas A and B were selected to produce the DEMs from the fused point set. For comparative analysis, the points in the corresponding areas from the point cloud of image pair 1c—which had the most points among image pairs of the data set—were also used to generate single-pair DEMs. The two selected areas and the DEMs are shown in Figure 7. In order to preserve the details of the reconstruction results, the grid spacings of the DEMs generated in Areas A and B were set to 1 and 2 mm, respectively, to facilitate detailed comparison. To investigate the details of the topographic mapping results, we obtained the profiles of the DEMs along Profile A in Figure 7b and Profile B in Figure 7d. Figure 8 shows the red elevation profiles of the DEMs from the fused points and blue single-pair DEM profiles.

Similarly, we selected Areas C and D from the results of the second data set for DEM-based evaluation. The point cloud of image pair 2a was chosen as the data source of the single-pair DEM. Figure 9 shows the DEMs in Areas C and D. Figure 10 shows Profiles C and D. As shown in Figures 8 and 10, more details can be seen in the DEMs generated from fused point clouds.

Statistical measures of mean gradient and sharpness, Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE), and Natural Image Quality Evaluator (NIQE) were used to evaluate the produced DEMs. The sharpness depicts the mean local terrain relief, which is calculated as the distance-weighted average of height differences of a grid to its eight neighboring grids in the DEM (Wang *et al.* 2004). BRISQUE and NIQE (Mittal *et al.* 2012, 2013) were originally used to evaluate image quality; DEMs with more details and less noise will achieve better results (lower scores). As shown in Table 3, compared with the best results of a single image pair, the DEMs from fused points have better quality, demonstrating the effectiveness of the proposed method in lunar topographic mapping with higher quality and more details.

Conclusion

In this study, a novel method for enhanced lunar topographic mapping is proposed using multiple rover stereo images acquired at one location with changing illumination conditions. The point clouds generated from each pair of stereo images were fused into a global optimized lunar surface point set. To eliminate errors in the measurements of rover mast drives, point-cloud registration was performed using the global ICP method with constraints from 2D matched features to enhance the consistency among point clouds. Each point was inputted for clustering based on the position consistency, to find repeated surface observations. The 3D coordinates were estimated by all 3D points in the cluster, with the position uncertainties. By processing all clusters of repeated observations and adding the new observations, a fused surface point cloud with larger spatial coverage (higher point density) and lower positioning uncertainty (higher precision) can be obtained.

To validate the effectiveness of the proposed method, we acquired two image data sets using the Yutu-2 rover's Navcam at waypoints LE01002 and LE00803 under changing illumination conditions. Experimental results show that the topographic mapping products (point clouds and DEMs) generated by the proposed method have higher quality than those conventionally derived from single stereo images in terms of revealing more details of the lunar surface. Furthermore, the spatial coverage of the products is improved by elimination of shadow areas with the help of changing illuminations. Without involving complex engineering operations, the proposed method provides a novel solution to enhanced

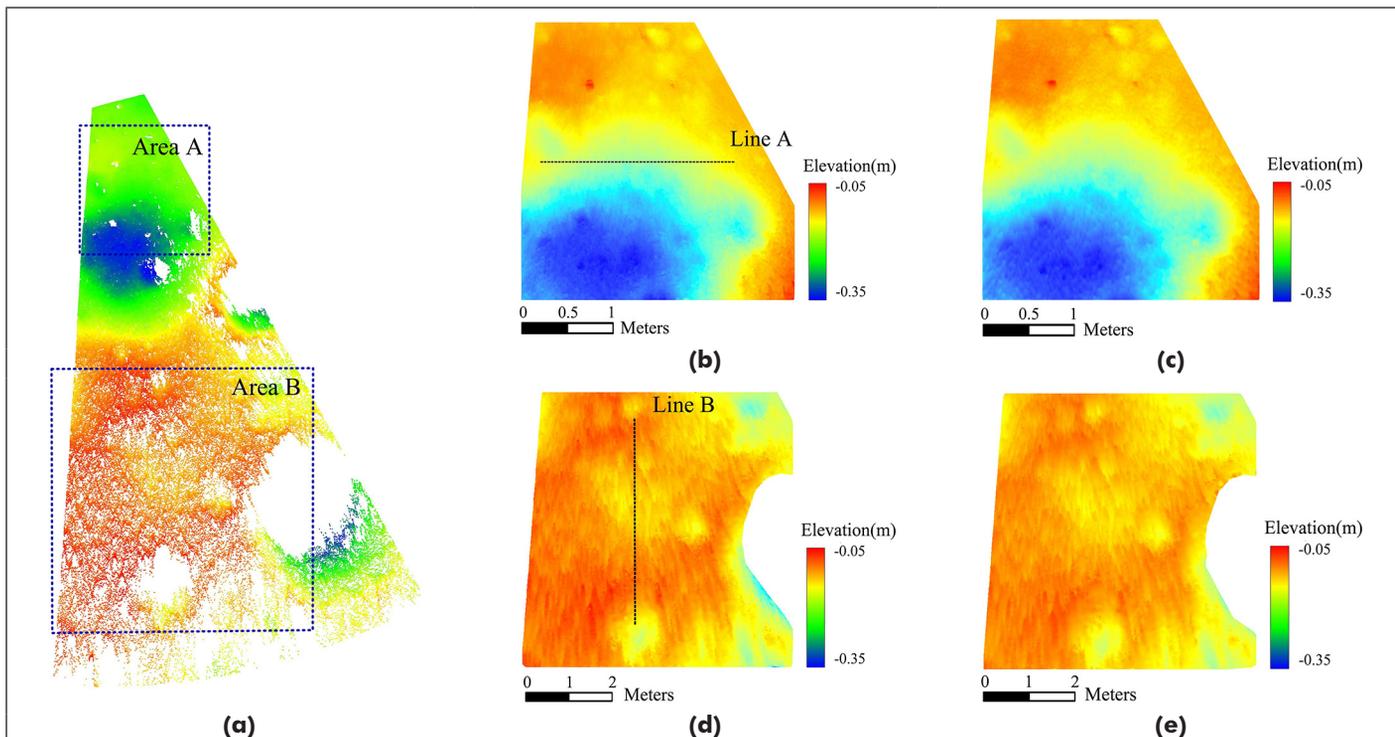


Figure 7. The selected areas and generated digital elevation models of the first data set: (a) test Areas A and B; (b, d) digital elevation models generated from the fused point sets; (c, e) digital elevation models from the point cloud of image pair 1c.

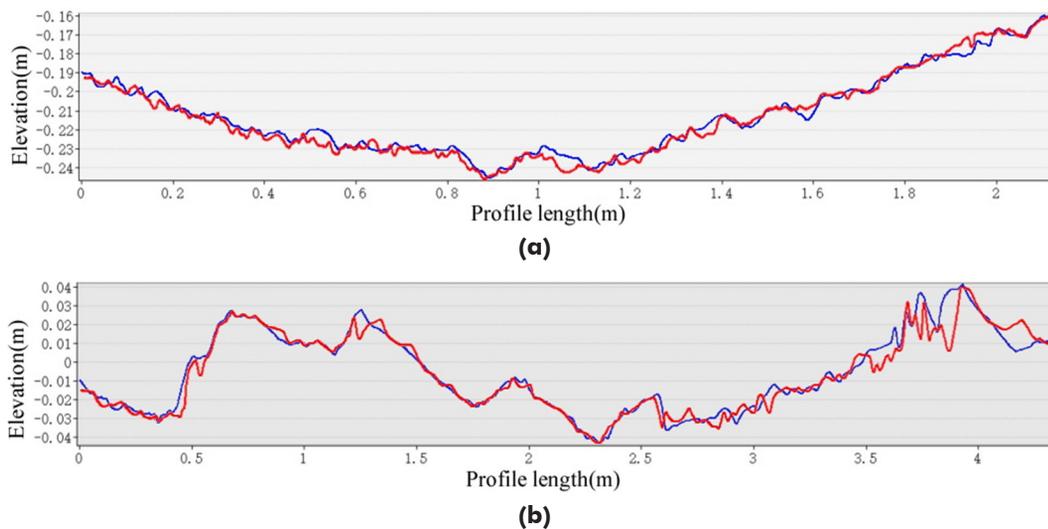


Figure 8. Elevation Profiles A (a) and B (b). Red lines are from fused digital elevation models and blue lines are from single-pair digital elevation models.

Table 3. Quality evaluation results of the digital elevation models.

Area	Source	Mean Gradient	Sharpness	BRISQUE	NIQE
A	Fused points	0.28	0.25	46.51	11.29
	Single pair	0.23	0.19	46.82	11.66
	Points of brute-force merging	0.26	0.21	47.46	11.46
B	Fused points	0.23	0.20	48.49	8.86
	Single pair	0.19	0.17	51.51	10.22
	Points of brute-force merging	0.20	0.18	49.02	8.94
C	Fused points	0.15	0.26	45.75	10.87
	Single pair	0.11	0.19	47.62	10.94
	Points of brute-force merging	0.12	0.21	47.78	10.90
D	Fused points	0.27	0.24	48.23	10.43
	Single pair	0.19	0.17	48.33	10.47
	Points of brute-force merging	0.22	0.19	48.40	10.43

BRISQUE = Blind/Referenceless Image Spatial Quality Evaluator; NIQE = Natural Image Quality Evaluator.

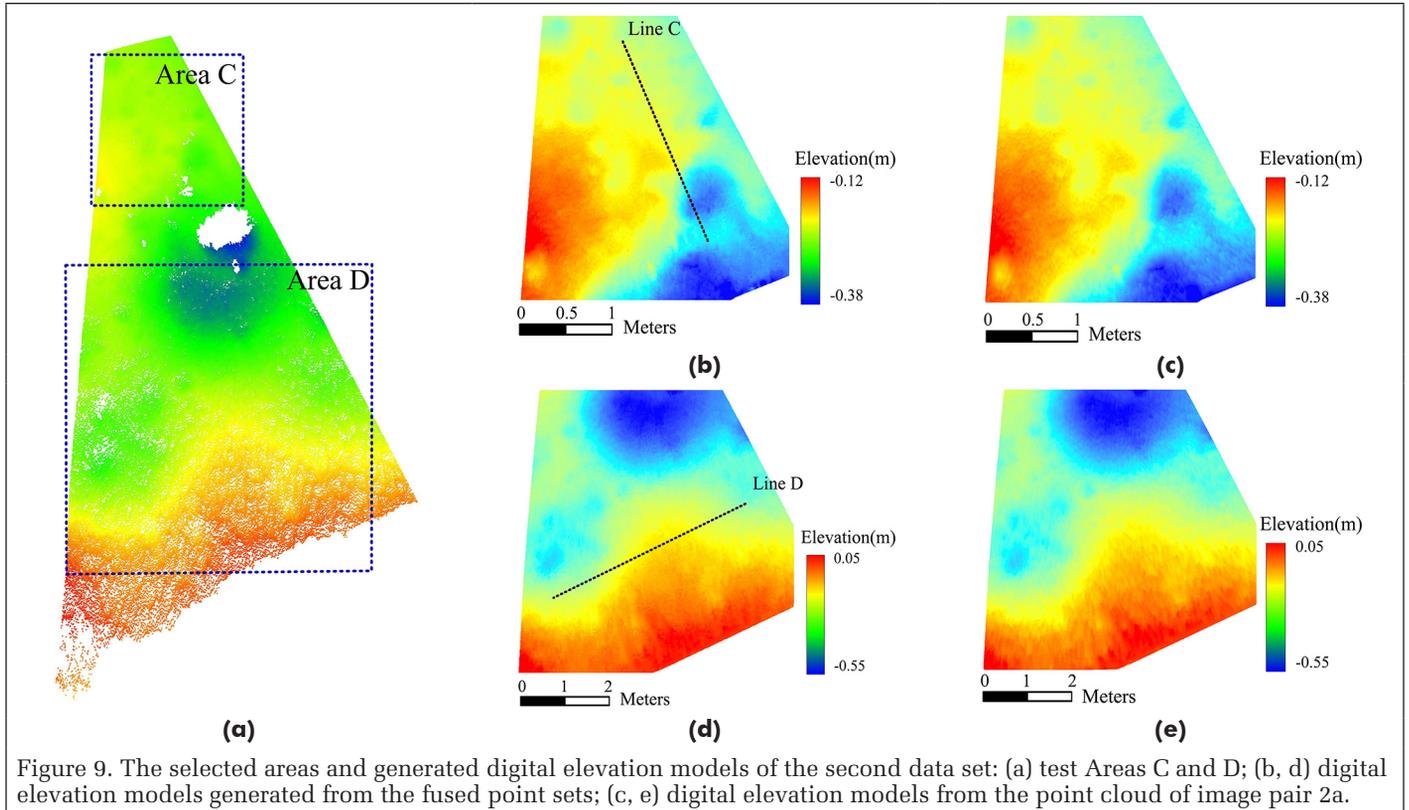


Figure 9. The selected areas and generated digital elevation models of the second data set: (a) test Areas C and D; (b, d) digital elevation models generated from the fused point sets; (c, e) digital elevation models from the point cloud of image pair 2a.

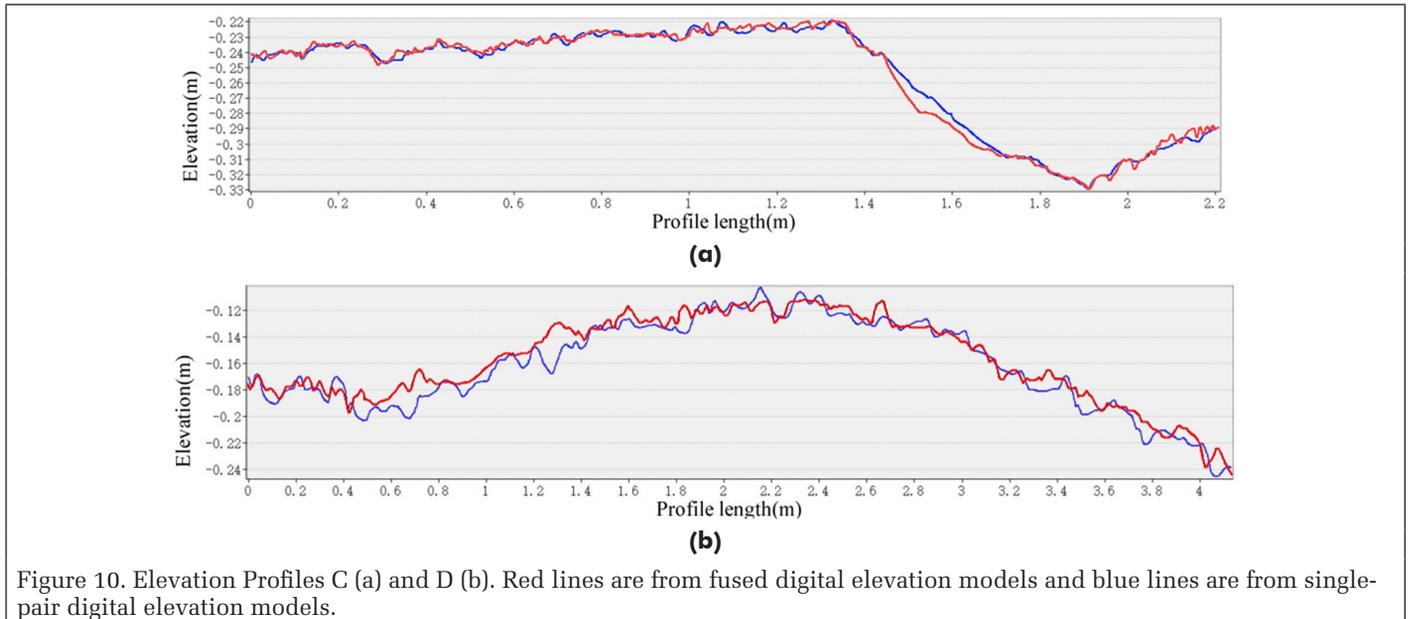


Figure 10. Elevation Profiles C (a) and D (b). Red lines are from fused digital elevation models and blue lines are from single-pair digital elevation models.

lunar topographic mapping, which has great potential to be applied in planetary rover and lander missions. In future work, shape-from-shading techniques could be combined with the proposed stereo mapping method to further enhance the topographic mapping capability using these kinds of image data sets. Furthermore, photometric properties of the surface materials could be studied based on these unique data sets with changing illumination conditions, after appropriate photometric calibration.

Acknowledgments

The *Chang'e-4* mission was carried out by the Chinese Lunar Exploration Program. This work was supported by the National Key Research and Development Program of China

(grant 2018YFB1305004) and the National Natural Science Foundation of China (grants 41941003 and 41771488).

References

- Alexander, D. A., R. G. Deen, P. M. Andres, P. Zamani, H. B. Mortensen, A. C. Chen, M. K. Cayanan, J. R. Hall, V. S. Klochko, O. Pariser, C. L. Stanley, C. K. Thompson and G. M. Yagi. 2006. Processing of Mars Exploration Rover imagery for science and operations planning. *Journal of Geophysical Research: Planets* 111(E2):E02S02.
- Barnard, S. T. and M. A. Fischler. 1982. Computational stereo. *ACM Computing Surveys* 14(4):553–572.
- Besl, P. J. and N. D. McKay. 1992. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2):239–256.

- Besse, F., C. Rother, A. Fitzgibbon and J. Kautz. 2014. PMBP: PatchMatch belief propagation for correspondence field estimation. *International Journal of Computer Vision* 110(1):2–13.
- Chang, J.-R. and Y.-A. Chen. 2018. Pyramid stereo matching network. Pages 5410–5418 in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, held in City, St./Country, DD–DD Month YYYY. Edited by J. Editor. City, St./Country: Publisher.
- Chang, Y.-J. and Y.-S. Ho. 2017. Pixel-based adaptive normalized cross correlation for illumination invariant stereo matching. Pages 124–129 in *Proceedings of IS&T International Symposium on Electronic Imaging, Stereoscopic Displays and Applications, XXVIII*.
- Di, K. and R. Li. 2007. Topographic mapping capability analysis of Mars Exploration Rover 2003 mission imagery. Pages 28–31 in *Proceedings of the 5th International Symposium on Mobile Mapping Technology (MMT 2007)*.
- Di, K., Z. Liu, B. Liu, W. Wan, M. Peng, Y. Wang, S. Gou, Z. Yue, X. Xin, M. Jia and S. Niu. 2019. Chang'e-4 lander localization based on multi-source data. *Journal of Remote Sensing* 23(1):177–184.
- Di, K., Z. Liu, W. Wan, M. Peng, B. Liu, Y. Wang, S. Gou and Z. Yue. 2020. Geospatial technologies for Chang'e-3 and Chang'e-4 lunar rover missions. *Geo-spatial Information Science* 23(1):87–97.
- Di, K. and M. Peng. 2011. Wide baseline mapping for Mars rovers. *Photogrammetric Engineering and Remote Sensing* 77(6):609–618.
- Di, K., F. Xu, J. Wang, S. Agarwal, E. Brodyagina, R. Li and L. Matthies. 2008. Photogrammetric processing of rover imagery of the 2003 Mars Exploration Rover mission. *ISPRS Journal of Photogrammetry and Remote Sensing* 63(2):181–201.
- Förstner, W. and E. A. Gülch. 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. Pages PP–PP in *Proceedings of the ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*.
- Gruen, A. W. 1985. Adaptive least squares correlation: A powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography* 14(3):175–187.
- Harris, C. and M. Stephens. 1988. A combined corner and edge detector. Pages PP–PP in *Proceedings of the 4th Alvey Vision Conference*.
- Hirschmuller, H. 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(2):328–341.
- Li, R., S. W. Squyres, R. E. Arvidson, B. A. Archinal, J. Bell, Y. Cheng, L. Crumpler, D. J. Des Marais, K. Di, T. A. Ely, M. Golombek, E. Graat, J. Grant, J. Guinn, A. Johnson, R. Greeley, R. L. Kirk, M. Maimone, L. H. Matthies, M. Malin, T. Parker, M. Sims, L. A. Soderblom, S. Thompson, J. Wang, P. Whelley and F. Xu. 2005. Initial results of rover localization and topographic mapping for the 2003 Mars Exploration Rover mission. *Photogrammetric Engineering and Remote Sensing* 71(10):1129–1142.
- Liu, W. C. and B. Wu. 2020. An integrated photogrammetric and photoclinometric approach for illumination-invariant pixel-resolution 3D mapping of the lunar surface. *ISPRS Journal of Photogrammetry and Remote Sensing* 159:153–168.
- Liu, W. C., B. Wu and C. Wohler. 2018. Effects of illumination differences on photometric stereo shape-and-albedo-from-shading for precision lunar surface reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing* 136:58–72.
- Liu, Z. Q., K. C. Di, M. Peng, W. H. Wan, B. Liu, L. C. Li, T. Y. Yu, B. F. Wang, J. L. Zhou and H. M. Chen. 2015. High precision landing site mapping and rover localization for Chang'e-3 mission. *Science China Physics, Mechanics & Astronomy* 58(1):1–11.
- Liu, Z., K. Di, J. Li, J. Xie, X. Cui, L. Xi, W. Wan, M. Peng, B. Liu, Y. Wang, S. Gou, Z. Yue, T. Yu, L. Li, J. Wang, C. Liu, X. Xin, M. Jia, Z. Bo, J. Liu, R. Wang, S. Niu, K. Zhang, Y. You, B. Liu and J. Liu. 2020. Landing site topographic mapping and rover localization for Chang'e-4 mission. *Science China Information Sciences* 63(4):140901.
- Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2):91–110.
- Luo, W., A. G. Schwing and R. Urtasun. 2016. Efficient deep learning for stereo matching. Pages 5695–5703 in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, held in Las Vegas, Nev. 27–30 June 2016. [https://doi: 10.1109/CVPR.2016.614](https://doi.org/10.1109/CVPR.2016.614).
- Matthies, L., M. Maimone, A. Johnson, Y. Cheng, R. Willson, C. Villalpando, S. Goldberg, A. Huertas, A. Stein and A. Angelova. 2007. Computer vision on Mars. *International Journal of Computer Vision* 75(1):67–92.
- Mei, X., X. Sun, M. Zhou, S. Jiao, H. Wang and X. Zhang. 2011. On building an accurate stereo matching system on graphics hardware. Pages PP–PP in 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), held in Barcelona, Spain 6–13 November 2011. [https://doi: 10.1109/ICCVW.2011.6130280](https://doi.org/10.1109/ICCVW.2011.6130280).
- Michael, M., J. Salmen, J. Stallkamp and M. Schlipf. 2013. Real-time stereo vision: Optimizing Semi-Global Matching. Pages 1197–1202 in 2013 IEEE Intelligent Vehicles Symposium (IV), held in Gold Coast, QLD, Australia 23–26 June 2013. [https://doi: 10.1109/IVS.2013.6629629](https://doi.org/10.1109/IVS.2013.6629629).
- Mittal, A., A. K. Moorthy and A. C. Bovik. 2012. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing* 21(12):4695–4708.
- Mittal, A., R. Soundararajan and A. C. Bovik. 2013. Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters* 20(3):209–212.
- Olson, C. F., H. Abi-Rached, M. Ye and J. P. Hendrich. 2003. Wide-baseline stereo vision for Mars rovers. Pages 1302–1307 in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, held in Las Vegas, Nev. 27–31 October 2003. [https://doi: 10.1109/IROS.2003.1248825](https://doi.org/10.1109/IROS.2003.1248825).
- Olson, C. F., L. H. Matthies, J. R. Wright, R. Li and K. Di. 2007. Visual terrain mapping for Mars exploration. *Computer Vision and Image Understanding* 105(1):73–85.
- Scharstein, D. and R. Szeliski. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47(1–3):7–42.
- Stein, T. C., R. E. Arvidson and F. Zhou. 2019. PDS analyst's notebook for MSL and MER. Page PP in *Ninth International Conference on Mars*.
- Wang, H., W. Zhong, J. Wang and D. Xia. 2004. Research of measurement for digital image definition. *Journal of Image and Graphics* 9(7):581–586.
- Wu, B., Y. Zhang and Q. Zhu. 2011. A triangulation-based hierarchical image matching method for wide-baseline images. *Photogrammetric Engineering and Remote Sensing* 77(7):695–708.
- Xu, J., Q. Yang, J. Tang and Z. Feng. 2016. Linear time illumination invariant stereo matching. *International Journal of Computer Vision* 119(2):179–193.
- Zhang, Z. and J. Zhang. 1997. *Digital Photogrammetry*. Wuhan, China: Press of the Wuhan Technical University of Surveying and Mapping.
- Zhu, Q., B. Wu and Y. Tian. 2007. Propagation strategies for stereo image matching based on the dynamic triangle constraint. *ISPRS Journal of Photogrammetry and Remote Sensing* 62(4):295–308.