Wide Baseline Mapping for Mars Rovers

Kaichang Di and Man Peng

Abstract

Wide-baseline mapping technology has been applied in NASA's Mars Exploration Rover (MER) mission to improve the accuracy of mapping of far-range terrain from Rover stereo images. As a basic research topic in photogrammetry and computer vision, it is desirable to perform a comprehensive investigation on the accuracy and automation of wide-baseline mapping. This paper presents the results of a systematic accuracy analysis of wide-baseline mapping through theoretical derivation and Monte Carlo simulation. Automated bundle adjustment and 3D wide-baseline mapping techniques are developed and tested using wide-baseline images acquired by the Spirit Rover. Experimental results demonstrate the effectiveness of the proposed techniques and verify that the mapping capability of rover stereo images can be extended from tens of meters (hard-baseline) to hundreds of meters using wide-baseline mapping techniques.

Introduction

In an unmanned planetary rover mission, 3D mapping of the surrounding terrain is of fundamental importance for safe rover navigation and for scientific investigation of geological features. During the 2003 Mars Exploration Rover (MER) mission, Navcam (navigation camera), and Pancam (panoramic camera) rover images have been extensively used for topographic mapping of both the Spirit and Opportunity landing sites; these high-accuracy maps have greatly supported rover traverse planning and scientific investigation in mission operations (Li *et al.*, 2005; Di *et al.*, 2008). Both Navcam and Pancam stereo cameras are mounted on the same camera bar atop a rover mast. This bar can be rotated 360° in azimuth and ±90° in elevation, enabling the acquisition of full or partial panoramas by Navcam and Pancam camera systems. Navcam is an engineering stereo camera used for navigation purposes (Maki *et al.*, 2003) while Pancam is a high-resolution, multispectral stereo panoramic imager used for scientific investigation of the morphology, topography, and geology of the two MER landing sites (Bell *et al.*, 2003). Pancam has a longer baseline and a longer focal length than Navcam, making it more effective in mapping medium to far objects.

Due to the limited length of a hard-baseline, the expected measurement error from hard-baseline stereo images is less than 1 m within a range of 27 m for Navcam and 55 m for Pancam (Di and Li, 2007). This baseline length usually satisfies the requirements of day-to-day, short-term planning. However, for long-term planning and investigations of distant geological features, a mapping capability of up to hundreds of meters is desirable, making wide-baseline mapping technology necessary. A wide-baseline stereo pair is formed by two or more images taken at two separate rover positions, resulting in a "soft" baseline that is much wider than the "hard" baseline designated on the camera bar (Figure 1). During Mars Exploration Rover mission operations, wide-baseline stereo images have been used to map a number of far-range major features including Husband Hill, McCool Hill, and Von Braun at the Spirit landing site as well as Endurance and Victoria Craters at the Opportunity landing site (Li *et al.*, 2005; Di and Li, 2007; Chen, 2008). These wide-baselines were designed empirically using the parallax equation for "normal case" stereo as a reference and the wide-baseline mapping process involved quite a few manual interactions, for example, manual selection of cross-site (wide-baseline) tie points for bundle adjustment and interactive image matching to obtain terrain details (Di and Li, 2007; Chen, 2008).

Photogrammetric bundle adjustment is a key technique for achieving high-precision topographic products and has been widely used in MER mission for 3D mapping as well as rover localization (Li *et al.*, 2004 and 2005; Di and Li, 2007; Di *et al.*, 2008). Bundle adjustment of an image network formed by relevant Pancam and Navcam stereo images generates high-precision exterior orientation parameters of the images as well as ground positions of tie points. In particular, bundle adjustment of multi-site images ensures that the mapping products generated from bundle-adjusted images are of high accuracy and geometrically seamless.

Image matching is one of the core problems in computer vision and digital photogrammetry. Traditionally, interest point detectors such as the Harris operator (Harris and Stephens, 1988) and the Förstner operator (Förstner and Gulch, 1987) have been used widely in stereo-image matching. In the last decade, wide-baseline matching has been an active area of research in the computer vision community. A variety of invariant detectors and descriptors have been proposed for wide-baseline matching, object recognition, and image retrieval purposes including affine invariant regions (Tuytelaars and Van Gool, 2000), MSER (maximally stable extremal regions) (Matas *et al.*, 2002), Harris-Laplace and Harris-Affine interest-point detectors (Mikolajczyk and Schmid, 2004), SIFT (Scale Invariant Feature Transform) (Lowe 1999 and 2004), and SURF (Speeded-Up Robust Features) (Bay *et al.*, 2008). These detectors and their

State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Applications, Chinese Academy Sciences, P. O. Box 9718, Datun Road, Chaoyang District, Beijing 100101, P.R. China (kcdi@irsa.ac.cn).

Photogrammetric Engineering & Remote Sensing
Vol. 77, No. 6, June 2011, pp. 609–618.

0099-1112/11/7706–0609/$3.00/0

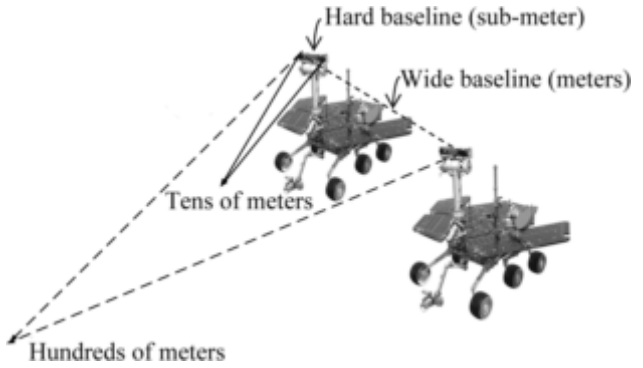

Figure 1. Conceptual illustration of wide-baseline mapping.

descriptors were tested and applied mainly in structured scenes of indoor or outdoor environments on Earth. Olson *et al.* (2003) addressing the issues of wide-baseline stereo vision for Mars rovers, developed an algorithm consisting of the following four steps: (a) motion refinement through Förstner interest-point extraction, hierarchical matching, and Levenberg-Marquardt optimization for motion estimation, (b) image rectification (epipolar resampling), (c) disparity calculation with maximum-likelihood image matching, and (d) triangulation for determination of the 3D positions of matched image pixels. This algorithm has been tested for wide-baseline mapping using images of Mars-like terrain on Earth (Olson *et al.*, 2007) and Mars rover images (Olson and Abi-Rached, 2009). Invariant detectors could be applied to Mars rover images to address the challenges of wide-baseline matching. With the current availability of wide-baseline stereo images from the Spirit and Opportunity rovers, it would be more interesting to test the wide-baseline mapping algorithms using these actual Mars surface images.

Overall, it is highly desirable to perform a systematic study of wide-baseline mapping accuracy and optimal configuration; it is also desirable to develop fully automated techniques for wide-baseline bundle adjustment and dense matching for Mars rovers. This paper presents accuracy analysis results of wide-baseline mapping through theoretical analysis and Monte Carlo simulation. An automated bundle-adjustment technique is developed, and a TIN controlled matching method is applied for wide-baseline mapping. The techniques are tested using wide-baseline images acquired by the Spirit Rover.

It should be noted that the research team and its affiliated institute is not part of the MER operations, and the developed method has not been used in MER operations as well. The purpose of this work is to solve the basic problem of wide-baseline mapping in the research field of photogrammetry and computer vision and test the developed methods using publicly available rover data sets.

Geometric Configuration and Accuracy Analysis for Wide-Baseline Mapping

Theoretical Accuracy Analysis

In order to investigate mapping capability under different wide-baseline configurations, we first performed a theoretical analysis of the possible attainable accuracy. For this research, we used the Pancam parameters to analyze the attainable accuracy of wide-baseline mapping; Navcam parameters were used only to estimate the baseline error, which is equivalent to the rover localization error. The geometric parameters of the MER Pancam and Navcam are listed in Table 1 (Bell *et al.*, 2003; Maki *et al.*, 2003). To simplify the derivation, we assumed that the wide-baseline stereo pair was acquired in a "normal case" configuration, i.e., where the two camera axes are parallel to each other and perpendicular to the baseline (Di and Li, 2007).

TABLE 1. CAMERA PARAMETERS OF MER PANCAM AND NAVCAM

	Pancam	Navcam
Stereo base	30 cm	20 cm
Focal length	43 mm	14.67 mm
Image dimension	1024 × 1024 pixels	1024 × 1024 pixels
Pixel size	$12\,\mu m$	$12\,\mu m$
Field of view	$16.8^\circ \times 16.8^\circ$	$45^\circ \times 45^\circ$

Since measurement error in the range direction is always larger than that in other two directions (Di and Li, 2007), in the following discussion we have used range error to represent mapping accuracy. In the "normal case" stereo image configuration, the range Y can be calculated using the parallax equation:

$$Y = \frac{B}{p} f \tag{1}$$

where f is the focal length, B is the baseline, and p is the stereo parallax of a specific point. It is easy to understand from this equation that the mapping accuracy of hard-baseline stereo depends only on the measured parallax p (because f and B are fixed for Navcam and Pancam, respectively). For wide-baseline stereo, the baseline is not fixed (as it is with a hard-baseline) and therefore the baseline error will affect the mapping accuracy. Through error-propagation derivation (Equation 1), range error can be calculated as:

$$\sigma_Y^2 = \left(\frac{Y}{B}\right)^2 \sigma_B^2 + \left(\frac{Y^2}{Bf}\right)^2 \sigma_p^2 \tag{2}$$

where σ_Y is the standard error of range Y, σ_B is the standard error of the wide-baseline, and σ_p is the parallax measurement error. In digital photogrammetry, σ_p is determined by the accuracy of image matching. According to theoretical analysis, pixel-level image matching (correlation) can reach an accuracy of about one-third pixel (Zhang and Zhang, 1997). In practice, the pixel measurement error can be superior to one-third pixel for image pairs that belong to the same site (hard-baseline) if a sub-pixel matching algorithm is applied. If the image pairs belong to different sites (wide-baseline), the pixel measurement errors may be larger than one-third pixel due to a large difference in viewing angles. Since there are image pairs at both one site (hard-baseline) and two sites in wide-baseline mapping, generally we used one-third pixel as the pixel measurement error in our analysis to represent the overall error.

Baseline error can be depicted by rover localization with the traverse length of baseline B. During MER mission operations, Navcam images were used for localization purpose at most rover locations. Therefore, in this research we have used the Navcam parameters to estimate the baseline error. For simplicity, we assumed there is only one landmark in the center of two sites (the two ends of the wide-baseline). In this ideal case, rover localization can be performed simply by a translation of Site 1 with respect to Site 0 based on the difference between the locations of the landmark measured from the two sites. Thus, the wide-baseline can be calculated as:

$$B = y_0 + y_1 \tag{3}$$

where y_0 and y_1 are coordinates (ranges) of the landmark with respect to Site 0 and Site 1, respectively. Through simple error propagation, the error of the baseline is the sum of the two range errors:

$$\sigma_B^2 = \sigma_{y_0}^2 + \sigma_{y_1}^2. \tag{4}$$

Assuming that the landmark is in the center of these two sites (i.e., $y_0=y_1=B/2$), the range error for the Navcam hard-baseline is calculated from Equation 2 as:

$$\sigma_{y_0}^2 = \sigma_{y_1}^2 = \left(\frac{B^2\sigma_p}{4b_{Navcam}f_{Navcam}}\right)^2 \tag{5}$$

where b_{Navcam} is the Navcam hard-baseline and f_{Navcam} is the Navcam focal length. Note that the first term of Equation 2 was not included here because the fixed Navcam hard-baseline (with no baseline error) was used for localization. Then, Equation 4 can be represented as:

$$\sigma_B^2 = \frac{B^4}{8b_{Navcam}^2 f_{Navcam}^2}\sigma_p^2. \tag{6}$$

Substituting σ_B^2 into Equation 2, the wide-baseline range error is finally represented as:

$$\sigma_Y^2 = \frac{Y^2B^2}{8b_{Navcam}^2 f_{Navcam}^2}\sigma_p^2 + \left(\frac{Y^2}{Bf_{Pancam}}\right)^2\sigma_p^2. \tag{7}$$

It should be noted that in this research, we use Pancam parameters to evaluate mapping accuracy, and use Navcam parameters to estimate the baseline error through rover localization. Therefore, in Equation 7, Navcam parameters b_{Navcam} and f_{Navcam} are used in the baseline error term (the first term), while Pancam parameter f_{Pancam} is used in the parallax measurement error term (the second term).

Based on Equation 7, Figure 2 was generated to show the range measurement errors of targets at different distances from the camera (100 to 700 m) under different wide-baselines. It can be seen that mapping errors do not change monotonically with respect to the baselines; there exists an optimal baseline for a given distance to a target. Table 2 shows the optimal baselines for targets at different distances.

TABLE 2. OPTIMAL BASELINES FOR TARGETS AT DIFFERENT DISTANCES BASED ON THEORETICAL ANALYSIS

Target distance (m)	100	200	300	400	500	600	700
Optimal baseline (m)	4.39	6.21	7.60	8.78	9.82	10.76	11.62
Range error (m)	0.299	0.847	1.556	2.395	3.348	4.401	5.546

As the width of the baseline increases, the area of overlap between the wide-baseline stereo pair becomes smaller. In photogrammetric engineering, the general rule is that end lap should be at least 60 percent. Through simple calculations, we can check to see if the optimal baselines satisfy the needs of end lap. Figure 3 is an illustration of the end lap of a wide-baseline stereo pair. Using similar triangles in Figure 3, we can easily obtain:

$$\frac{Y}{f} = \frac{\Delta X + B}{a} \tag{8}$$

and

$$\Delta x = \Delta X\frac{f}{Y} \tag{9}$$

where a is the image size and ΔX and Δx are the projections of the area of overlap in object space and image space, respectively. Thus, the end lap is:

$$\frac{\Delta x}{a} = 1 - \frac{Bf}{aY}. \tag{10}$$

For Pancam, the value of f/a is 3.499. To insure an end lap of 60 percent or better for Pancam, we should have:

$$0.6 \le 1 - 3.499B/Y < 1 \tag{11}$$

therefore, the valid range of B is:

$$0 < B \le 0.114Y \tag{12}$$

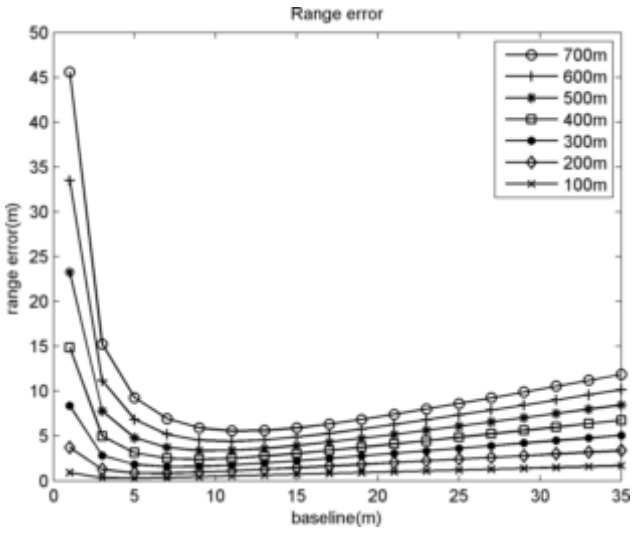

Figure 2. Range measurement error of different Pancam wide-baselines.

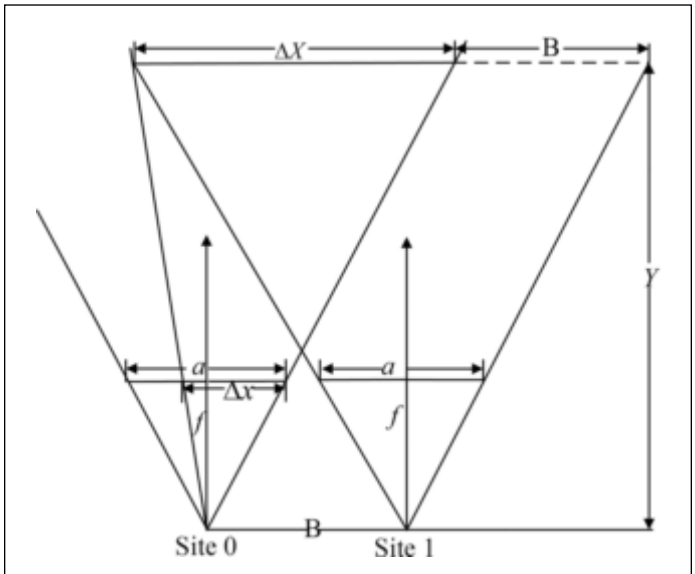

Figure 3. End lap for wide-baseline stereo.

TABLE 3. MAXIMUM BASELINE OF DIFFERENT RANGE DISTANCES

Target distance (m)	100	200	300	400	500	600	700
Maximum baseline (m)	11.4	22.8	34.2	45.6	57.0	68.4	79.8

The maximum baselines for targets at different distances are shown in Table 3. When comparing Tables 2 and 3, it can be seen that the optimal baselines are always less than the corresponding maximum baselines. This verifies the assumption that theoretically derived optimal baselines can be used in practical implementations of wide-baseline mapping.

In the above theoretical analysis, pixel measurement (parallax) error and baseline error were considered for wide-baseline mapping accuracy analysis using the "normal case" stereo configuration. Results from this analysis can be used as a general guideline for wide-baseline design. A more comprehensive analysis, one based on simulation that incorporates additional factors such as angular errors and multi-baseline intersection, is described below.

Monte Carlo Simulation

In the above theoretical derivation, only a single wide-baseline stereo pair was considered. The single stereo pair is formed by linking a single image taken at the first site (Site 0) with a single image taken at the second site (Site 1). In practice, a wide-baseline stereo pair can be acquired at both of the wide-baseline sites (Site 0 and Site 1). Thus, there are three possible stereo modes in wide-baseline mapping: one single image at Site 0 and another single image at Site 1 (single + single), one hard-baseline stereo pair at Site 0 and one single image at Site 1 (stereo + single), or one hard-baseline stereo pair at Site 0 and another hard-baseline stereo pair at Site 1 (stereo + stereo). These three configurations are illustrated in Figure 4. In addition, in the theoretical derivation the baseline error was estimated without taking into consideration the orientation (azimuth, tilt, and swing) error between Site 0 and Site 1. To incorporate these factors for a comprehensive wide-baseline accuracy analysis, we employ the Monte Carlo simulation method. In this simulation process, the known Pancam hard-baseline and focal length along with the fixed relationship between the left and right Pancam cameras were used. As in the theoretical analysis, the pixel measurement error was set to one-third pixel, and the baseline error was estimated using Equation 6. Based on experimental findings, after bundle adjustment the azimuthal inconsistency between the adjacent sites is usually within one pixel, the angular error was set to $\tan^{-1}(1 \text{ pixel}/f)$ for azimuth using the known Pancam pixel size and focal length.

The steps of the Monte Carlo simulation for wide-baseline mapping accuracy analysis are:

1. Generate the exterior orientation parameters (EOPs) for a Pancam hard-baseline stereo of Site 0 of the wide-baseline.
2. Suppose the wide-baseline and the hard-baseline are parallel to a site frame, whose X-axis points to local east, Z-axis points up in the local normal direction, and Y-axis is defined to form a right-handed system. Generate the EOPs for a Pancam hard-baseline stereo of Site 1 by adding the given wide-baseline to the X positions of the Pancam and keeping all other EOPs (Y, Z, azimuth, swing, tilt) the same as those of Site 0.
3. Given the 3D coordinates of a target that is with a specific distance to the center of the wide-baseline, calculate the image coordinates in the wide-baseline images with collinearity equations using the EOPs generated in Steps 1 and 2 above (so far, all the EOPs and 3D and 2D coordinates are "true" values without any error involved).
4. Generate 10,000 Gaussian random numbers with 0 mean and a standard deviation that is equal to the standard error of baseline σ_B as calculated in Equation 6, and add these random errors to the X positions of the EOPs at Site 1; generate 10,000 Gaussian random numbers with 0 mean and a standard deviation of one-third pixel, and add them to the x image coordinates at Site 1; generate 10,000 Gaussian random numbers with 0 mean and standard deviation of $\tan^{-1}(1 \text{ pixel}/f)$, and add these angular errors to the azimuth angle of the EOPs at Site 1; now the EOPs of Site 1 have included baseline and angular errors, and the 2D coordinates in Site 1 have included pixel measurement errors.
5. Calculate the 3D coordinates (10,000 times) of the target by space intersection using the EOPs and image coordinates.
6. Calculate the difference between the calculated and true 3D coordinates of the target, and obtain the RMS error from the 10,000 runs.

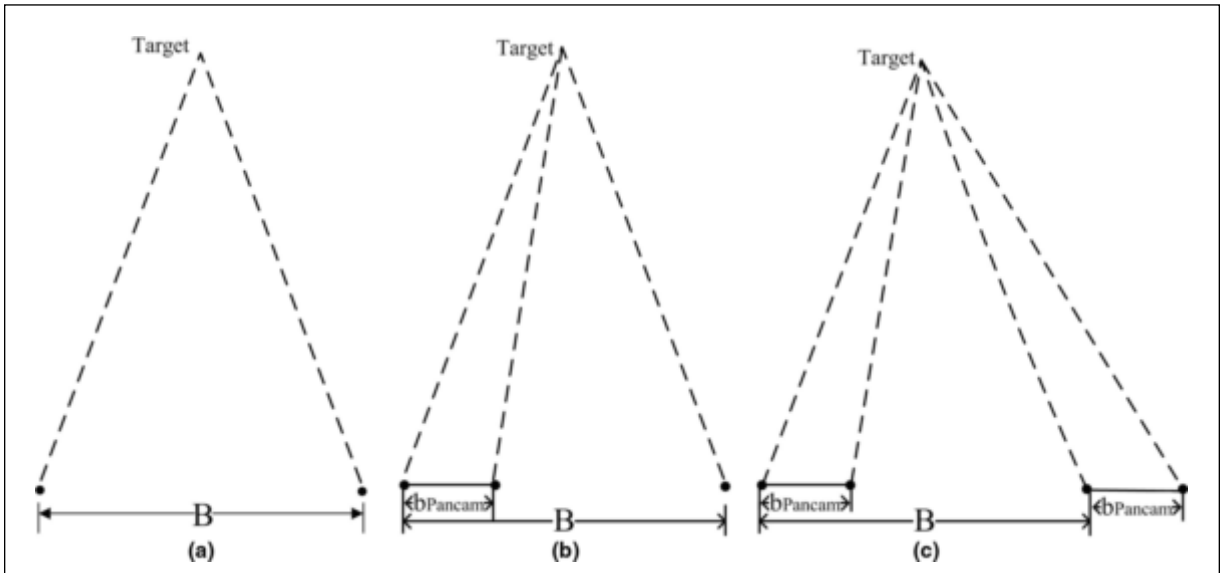

Figure 4. Stereo modes of wide-baselines: (a) "single + single," (b) "stereo + single," and (c) "stereo + stereo."

The above simulation is performed for each combination of baseline and target distance. A complete accuracy analysis is fulfilled by using all the combinations of baseline (1 m to 25 m) and target distance (100 m to 700 m) for all the three stereo modes using Pancam parameters (along with Navcam parameters for baseline error estimation). For stereo modes "stereo + single" and "stereo + stereo", random pixel measurement errors are also added to the *x* coordinates of the right images at Site 0. In addition, in order to verify the error analysis results from theoretical derivation, we simulated the "single + single" stereo mode with pixel measurement error and baseline error without considering angular error, which is the same setting as used in the theoretical analysis. The simulation results were found to be the same as the results of the theoretical analysis.

Figures 5a, 5b, and 5c show mapping error versus the baseline under the "single + single," "stereo + single," and "stereo + stereo" modes, respectively. Table 4 summarizes the optimal baselines and range errors for targets of different distances from the cameras. We can observe from the figures and Table 4 that: (a) there is always an optimal baseline at which the mapping error achieves the lowest value, (b) the farther the distance to the target, the longer the optimal baseline, (c) the mapping accuracy can be improved by multiple-baseline intersection with "stereo + single" and "stereo + stereo" modes, and (d) as multiple-baseline intersection can improve mapping accuracy with redundant observations, the optimal baseline tends to become shorter when employing multiple-baseline intersection. The above results can be used in detailed designs for wide-baseline mapping. Whenever possible, multiple-baseline intersection using "stereo + single" or "stereo + stereo" modes are recommended.

Automated Bundle Adjustment for Wide-baseline Mapping

The relative position and orientation of the wide-baseline images acquired at the two ends of the wide-baseline, which are initially determined by onboard dead-reckoning, may be of poor accuracy due to wheel slippage and IMU drift. To reach the highest theoretical accuracy in wide-baseline mapping, usually a bundle adjustment is necessary to fix the relative position and orientation of the two end sites. The key to the success of wide-baseline bundle adjustment is to select a sufficient number of tie points to link the wide-baseline images to form an image network. The wide-baseline (cross-site) tie points used for this bundle adjustment were chosen manually during MER mission operations. For this paper, we have developed an algorithm for automatic wide-baseline tie-point selection so that the bundle adjustment process is also automated.

Automatic Wide-baseline Tie Point Selection

Tie point selection between wide-baseline images is more challenging than tie point selection between hard-baseline images. As can be seen in Figure 6, the challenge is that close-range and medium-range objects look significantly different in the separate wide-baseline images due to a large difference in viewing angles; in addition, far-range objects lack visible texture. To tackle these challenges, we use a combination of Förstner operator and SURF operator.

Förstner operator (Förstner and Gulch, 1987) tends to provide more unique interest points in natural outdoor scenes with high accuracy and reliability under conditions of similar viewing angle and illumination. Matching of the extracted Förstner interest points is performed using cross correlation followed by least-squares matching. From this process, a sufficient number of interest points can be extracted and matched for medium- and far-range terrain, with only a few interest points being matched for close-range terrain. SURF (Bay *et al.*, 2008) is a scale- and rotation-invariant detector and descriptor, which provides similar results to SIFT though at a faster speed (**http://www.chrisevansdev.com/computer-vision-opensurf.html**). We used

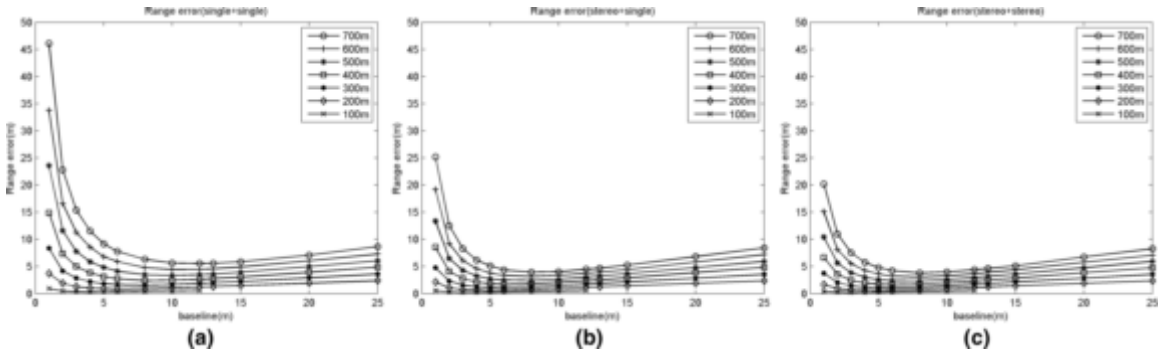

Figure 5. Range errors with different wide-baselines under the three different stereo modes: (a) "single + single," (b) "stereo + single," and (c) "stereo + stereo."

TABLE 4. OPTIMAL BASELINES AND RANGE ERRORS FOR TARGETS AT DIFFERENT DISTANCES FROM THE CAMERAS

Target distance(m)		100	200	300	400	500	600	700
Single + Single	Optimal baseline(m)	4	6	8	8	10	10	12
	Range error(m)	0.304	0.842	1.561	2.425	3.383	4.441	5.577
Stereo + Single	Optimal baseline(m)	3	5	6	6	7	8	9
	Range error(m)	0.217	0.608	1.104	1.719	2.361	3.182	3.953
Stereo + Stereo	Optimal baseline(m)	3	5	6	6	7	8	9
	Range error(m)	0.201	0.586	1.065	1.646	2.307	3.023	3.802

the SURF-64 descriptor and applied a distance-ratio measure (the ratio of Euclidean distance of the closest neighbor to that of the second-closest neighbor) for matching. As a result, SURF provided a sufficient number of matched points for close- and medium-range terrain, but only a few for far-range terrain. By combining the matched points from both the Förstner and the SURF matching processes, we were able to obtain a sufficient number of matched points for close-, medium- and far-range terrain scenarios.

Figure 6a shows the results of feature-point matching of a wide-baseline stereo pair taken by the Spirit rover. Matched points are marked as black crosses and linked by white lines. Those lines that have significant differences in orientation and length in comparison with neighboring lines indicate mismatches. In order to eliminate these mismatches, we computed the fundamental matrix between the wide-baseline image pair using a RANSAC (Fischler and Bolles, 1981) procedure (Kovesi, 2000). Under this iterative procedure, any mismatches are automatically eliminated as outliers. Figure 6b shows the results of feature-point matching after outlier elimination. In this particular pair, 93 mismatches were eliminated from the 222 matched points (Figure 6a), resulting in 129 correctly matched points (Figure 6b). Consequently, wide-baseline tie points were selected from these corrected matched points. By drawing a grid (e.g., 3×3) on the image and selecting the one tie point having the highest correlation coefficient (or distance ratio) in each patch, we can obtain evenly distributed wide-baseline tie points that are good for the subsequent bundle adjustment.

We tested the automatic wide-baseline tie point selection process using eight pairs of Pancam wide-baseline images taken by the Spirit rover at sites APFI (on Sols 774 and 775) and APGB (on Sols 776 to 778) for mapping McCool

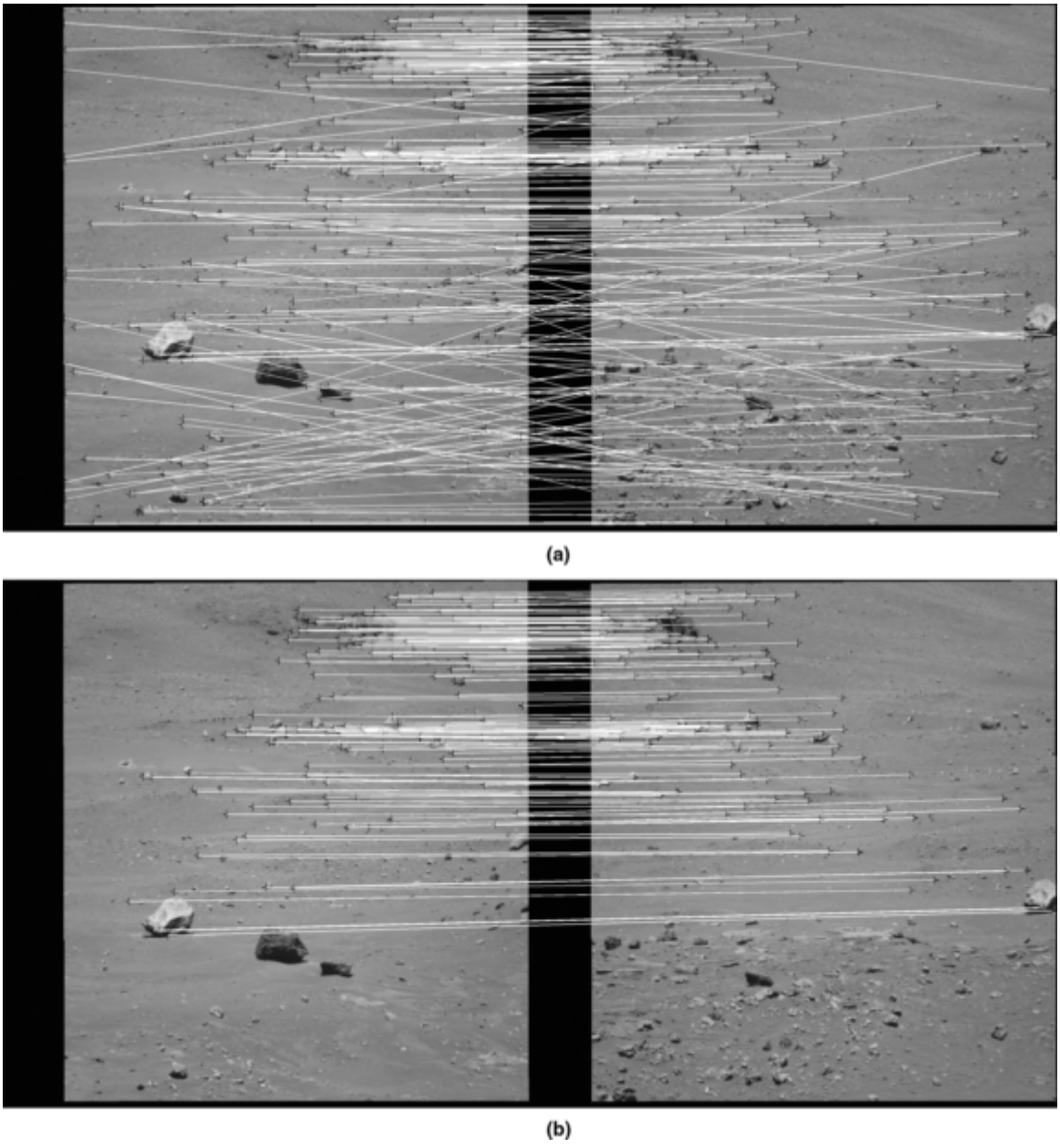

Figure 6. Matched feature points (a) before, and (b) after outlier elimination.

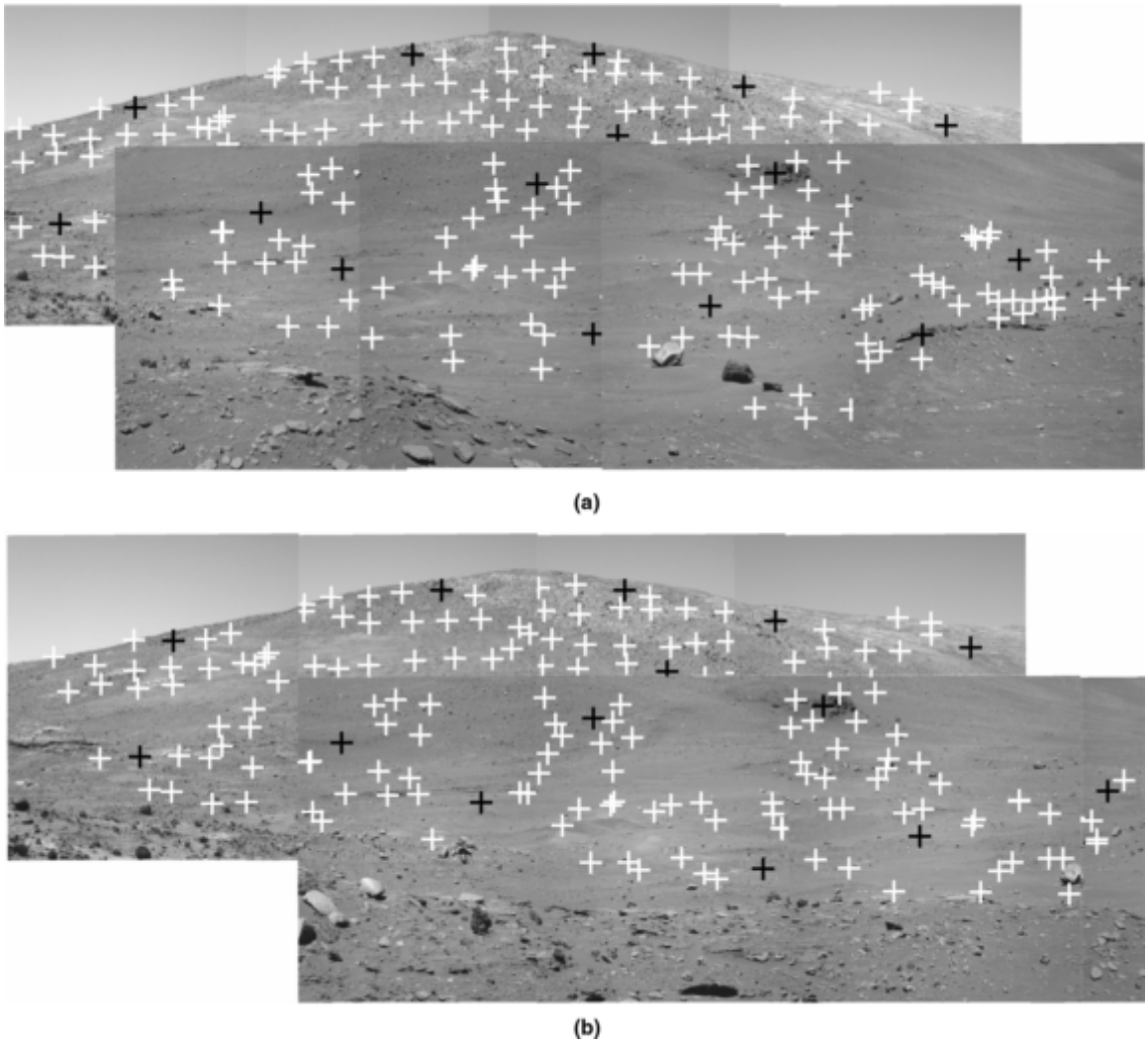

Figure 7. Automatically selected tie points overlaid on image mosaics at sites: (a) APFI, and (b) APGB.

Hill at a range of up to 400m. The length of this wide-baseline was about 8 m. The "stereo + stereo" mode was used, thus we had eight hard-baseline stereo pairs at each site. Image data for epipolar-resampled images (FFL files) was downloaded from the MER Analyst's Notebook website (**http://an.rsl.wustl.edu/mer/mera/mera.htm**) and used in this experiment. Figure 7 shows the automatically selected tie points (white and black crosses) overlaid on the two image mosaics. Overall there were 218 wide-baseline tie points, out of which 15 (black crosses) were used as checkpoints in the subsequent bundle adjustment.

Automated Bundle Adjustment of Wide-baseline Images

Using the collinearity equation as the basic model, bundle adjustment was used to adjust the EOPs of the images and the 3D coordinates of the tie points so that their optimal least-squares values were achieved. The image orientation parameters were extracted in the headers of the FFL files and converted from the CAHV model to EOPs of the conventional photogrammetry model (Di and Li, 2004); they were defined in the Landing Site Local (LSL) frame (Li *et al.*, 2004; Di *et al.*, 2008) and used as the initial values of the unknowns in the bundle adjustment. They were refined through the iterative solutions of the bundle adjustment.

There are six unknown EOPs for each wide-baseline image in the bundle adjustment. In the "stereo + single" and "stereo + stereo" modes, hard-baseline images are also involved. In order to reduce the correlation between the EOPs of the hard-baseline images, the fixed relationship between the left and right images of the hard-baseline stereo pairs was applied as constraint equations in the bundle adjustment to improve the accuracy and reliability of the adjustment computation (Di *et al.*, 2004). Since there is no absolute ground control on the Martian surface, the bundle adjustment is basically a free net adjustment in which the normal matrix is rank deficient. We used the Singular Value Decomposition technique to solve the normal equation in which the minimum norm principle was applied along with the principle of least squares.

The mapping accuracy of wide-baseline images is reflected in the accuracy of the bundle adjustment of the two sites. Since no absolute ground control is available, it is impossible to evaluate the accuracy using ground truth. Instead, the bundle adjustment accuracy was estimated by checking consistencies between positions in the 3D ground space and corresponding positions in 2D image space. From sites APFI and APGB, we selected 15 cross-site tie points (appearing in multiple wide-baseline stereo pairs) as

checkpoints, while another 203 tie points were used in the bundle adjustment. First, the 3D position of each check point was calculated from one wide-baseline stereo pair and then from another wide-baseline stereo pair; any difference was obtained by calculating the 3D distance. The 3D accuracy was then depicted by the average of this difference for all the check points in object space. Second, the 3D position of a check point, as calculated from one wide-baseline stereo pair, was projected onto another wide-baseline stereo pair. The projected image point was compared to the actual measured point and the difference was obtained by calculating the 2D distance. The 2D accuracy was then depicted by the average of the differences for all the checkpoints in image space. The maximal distance from the wide-baseline sites to the checkpoints is up to 400 m. Results were 3D accuracy of 0.563 m and 2D accuracy of 0.714 pixel after bundle adjustment. With a 3D accuracy of 1.118 m and a 2D accuracy of 2.952 pixels before bundle adjustment, it is clear that bundle adjustment has improved accuracy in both image and object space. This bundle adjustment accuracy represents the mapping accuracy of the wide-baseline images.

We also tested the automated bundle adjustment using the Pancam wide-baseline images taken by the Spirit rover at sites AVLF (on Sol 1348) and AVMA (on Sol 1350) for mapping Von Braun Hill at a range up to 120 m. The wide-baseline was about 8 m. The "stereo + single" mode was employed; we selected ten hard-baseline stereo pairs at site AVLF and ten single images at site AVMA. A total of 238 wide-baseline tie points were selected automatically, out of which 18 were used as checkpoints. The maximal distance from the sites to the checkpoints is up to 110 m. Before bundle adjustment, 3D accuracy was 1.681 m, and 2D accuracy was 2.515 pixels; after adjustment, 3D accuracy was 0.623 m, and 2D accuracy was 0.842 pixel. Again, results show that the bundle adjustment improves mapping accuracy, and the method for automatic wide-baseline tie point selection is effective.

Dense Matching and Digital Elevation Model (DEM) Generation

The bundle adjustment provides improved EOPs of the images as well as high-precision 3D ground positions for all the tie points. The 3D positions of all matched feature points (out of which the tie points were selected) were calculated by space intersection. However, their number was still sparse for reconstruction of the details of the terrain. Based on these matched feature points and the refined EOPs, we applied a dense matching of the wide-baseline images. A triangulated irregular network (TIN)-controlled dense image matching method was developed by Li *et al.* (2007) for the dense matching of hard-baseline images. We adopted this strategy and imposed the epipolar geometry of wide-baseline images. First, a TIN was constructed using the matched feature points of the wide-baseline images at Site 0. Second, triangles and the epipolar geometry were used to predict the approximate location of the homologous point in the wide-baseline images at Site 1. Third, image matching was performed in a small search range around the predicted location along the epipolar line using correlation coefficients. Fourth, least-squares matching was applied to all the matched dense points to refine the matched locations to a sub-pixel level. Finally, 3D coordinates of the matched points were calculated using the refined EOPs of the wide-baseline images. In our mapping of McCool Hill, we used a grid of 3×3 pixels for dense matching. Figure 8 shows the results of dense matching of a wide-baseline stereo pair. Overall, we obtained 7,185 matched points (including the previously matched feature points) in this pair. In order to ensure the quality of the DEM, an outlier detection procedure was performed on the dense 3D points using a local-surface-fitting technique. The outliers are caused by mismatches in the dense matching process. Any 3D points that are larger than three times the standard deviation of the distances to the local fitted plane were eliminated as outliers; 331 outliers were eliminated in this experiment. The resultant 44,533 3D points were used to generate a 1 m resolution DEM based on Kriging interpolation. Figure 9 shows the DEM of McCool Hill generated from wide-baseline images.

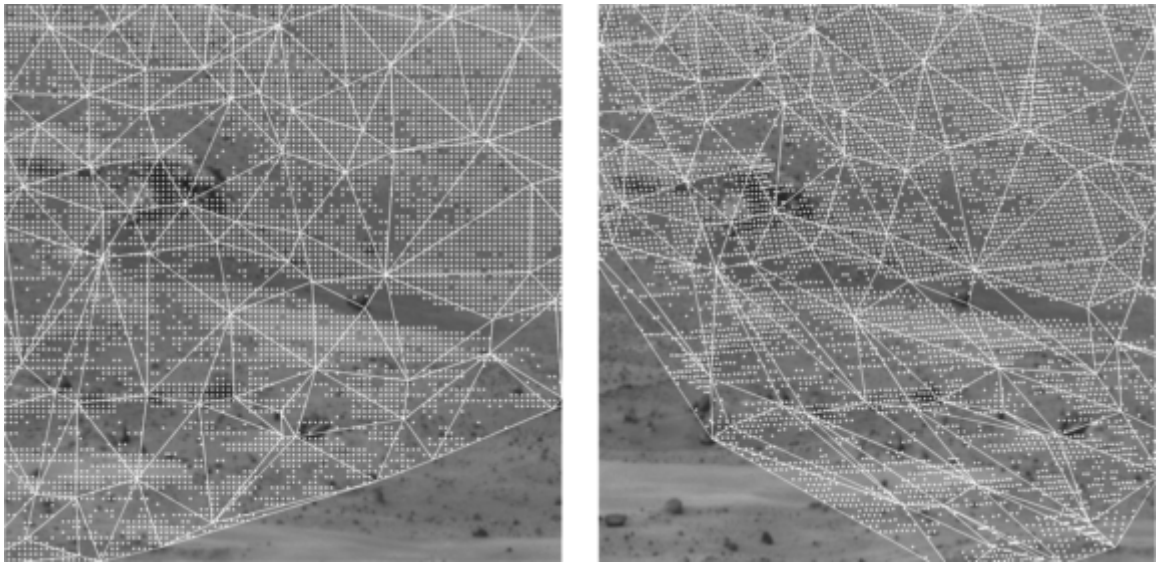

Figure 8. Dense image matching result of a wide-baseline stereo pair.

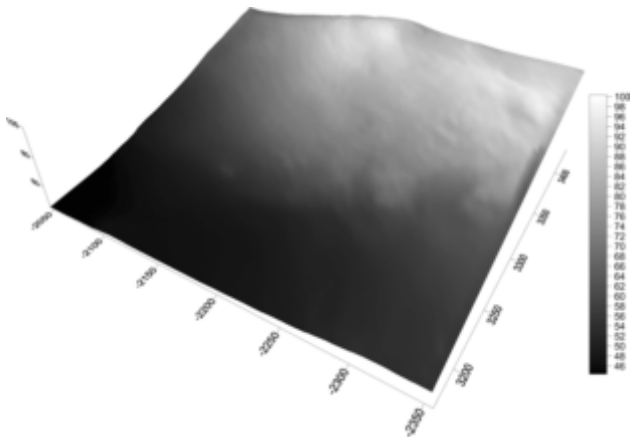

Figure 9. 3D view of the McCool Hill DEM.

Conclusions

This paper has described a systematic accuracy analysis of wide-baseline mapping through theoretical derivation and Monte Carlo simulation. Considering pixel measurement error and baseline error, an analytical equation was derived using the "normal case" configuration. In addition, a more comprehensive analysis was performed using Monte Carlo simulation taking into consideration pixel measurement error, baseline error and angular error along with three separate stereo modes. The results can be used as guidelines in the design of wide-baseline mapping for such variables as optimal baseline and stereo mode.

An automatic method for wide-baseline tie point selection was developed in this research based on a combination of Förstner operator and SURF operator followed by a RANSAC-based outlier elimination procedure. The sufficient number of evenly distributed tie points enabled an automated bundle adjustment. Experiments using wide-baseline images taken by the Spirit rover for mapping McCool Hill and Von Braun Hill demonstrated that the developed automated bundle adjustment is effective. After bundle adjustment, the mapping capability of the rover stereo images can be extended from tens of meters (hard-baseline) to hundreds of meters using wide-baseline mapping techniques. Finally, a TIN-controlled dense image matching method was used to generate dense matched points from wide-baseline images. Consequently, a high-resolution DEM was generated by Kriging interpolation.

The reliability of wide-baseline image matching is an important issue for wide-baseline mapping. The TIN-controlled dense matching may fail in some areas due to large perspective changes and/or poor textures. In the future, we will improve the speed and reliability of dense wide-baseline matching and conduct field tests on Earth, using ground truth to validate the attainable mapping accuracy.

Acknowledgments

Funding of this research by the National High Technology Research and Development Program of China (2009AA12Z310) and National Natural Science Foundation of China (40871202) is acknowledged. The authors appreciate the constructive comments from the reviewers.

References

Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool, 2008. Speeded-up robust features, *Computer Vision and Image Understanding*, 110(3):346–359.

Bell, J.F. III, S.W. Squyres, K.E. Herkenhoff, J.N. Maki, H.M. Arneson, D. Brown, S.A. Collins, A. Dingizian, S.T. Elliot, E.C. Hagerott, A.G. Hayes, M.J. Johnson, J.R. Johnson, J. Joseph, K. Kinch, M.T. Lemmon, R.V. Morris, L. Scherr, M. Schwochert, M.K. Shepard, G.H. Smith, J.N. Sohl-Dickstein, R.J. Sullivan, W.T. Sullivan, and M. Wadsworth, 2003. Mars exploration rover Athena panoramic camera (Pancam) investigation, *Journal of Geophysical Research - Planets*, 108(E12):8063 (doi:10.1029/2003JE002070).

Chen, W., 2008. *Wide Baseline Mapping in Mars Exploration*, M.Sc. thesis, The Ohio State University, Columbus, Ohio, 45 p.

Di, K., and R. Li, 2004. CAHVOR camera model and its photogrammetric conversion for planetary applications, *Journal of Geophysical Research - Planets*, 109(E4):E04004, (doi:10.1029/2003JE002199).

Di, K., F. Xu, and R. Li, 2004. Constrained bundle adjustment of panoramic stereo images for Mars landing-site mapping, *Proceedings of the 4th International Symposium on Mobile Mapping Technology*, 29–31 March, Kunming, China, unpaginated CD-ROM.

Di, K., and R. Li, 2007. Topographic mapping capability analysis of Mars exploration rover 2003 mission imagery, *Proceedings of the 5th International Symposium on Mobile Mapping Technology (MMT 2007)*, 28–31 May, Padua, Italy, unpaginated CD-ROM.

Di, K., F. Xu, J. Wang, S. Agarwal, E. Brodyagina, R. Li, L.H. 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.

Fischler, M.A., and R.C. Bolles, 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Communication of the ACM*, 24(6):381–395.

Förstner, W., and E. Gulch, 1987. A fast operator for detection and precise location of distinct points, corners and centers of circular features, *Proceedings of the ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, 02–04 June, Interlaken, Switzerland, pp. 281–305.

Harris, C., and M. Stephens, 1988. A combined corner and edge detector, *Proceedings of the 4th Alvey Vision Conference*, 31 August - 02 September, Manchester, UK, pp. 147–151.

Kovesi, P.D., 2000. MATLAB and Octave functions for computer vision and image processing, School of Computer Science & Software Engineering, The University of Western Australia, URL: *http://www.csse.uwa.edu.au/~pk/research/matlabfns/* (last date accessed: 17 March 2011).

Li, R., K. Di, L.H. Matthies, R.E. Arvidson, W.M. Folkner, and B.A. Archinal, 2004. Rover localization and landing-site mapping technology for 2003 Mars exploration rover mission, *Photogrammetric Engineering & Remote Sensing*, 70(1):77–90.

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 & Remote Sensing, Special Issue on Mapping Mars*, 71(10):1129–1142.

Li, R., K. Di, A.B. Howard, L.H. Matthies, J. Wang, and S. Agarwal, 2007. Rock modeling and matching for autonomous long-range Mars rover localization, *Journal of Field Robotics*, 24(3):187–203.

Lowe, D.G., 1999. Object recognition from local scale-invariant features, *Proceedings of the International Conference on Computer Vision*, 20–25 September, Corfu, Greece, pp. 1150–1157.

Maki, J.N., J.F. Bell, III, K.E. Herkenhoff, S.W. Squyres, A. Kiely, M. Klimesh, M. Schwochert, T. Litwin, R. Willson, A. Johnson, M. Maimone, E. Baumgartner, A. Collins, M. Wadsworth,

S.T. Elliot, A. Dingizian, D. Brown, E.C. Hagerott, L. Scherr, R. Deen, D. Alexander, and J. Lorre, 2003. Mars exploration rover engineering cameras, *Journal of Geophysical Research - Planets*, 108(E12):8071, (doi:10.1029/2003JE002077).

Matas, J., O. Chum, M. Urban, and T. Pajdla, 2002. Robust wide-baseline stereo from maximally stable extremal regions, *Proceedings of the 13th British Machine Vision Conference (BMVC2002)*, 02–05 September, Cardiff, Wales, UK, pp. 384–394.

Mikolajczyk, K., and C. Schmid, 2004. Scale and affine invariant interest point detectors, *International Journal of Computer Vision*, 60(1):63–86.

Olson, C.F., H. Abi-Rached, M. Ye, and J.P. Hendrich, 2003.Wide-baseline stereo vision for Mars rovers, *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 27–31 October, Las Vegas, Nevada, pp. 1302–1307.

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.

Olson, C.F., and H. Abi-Rached, 2010. Wide-baseline stereo vision for terrain mapping, *Machine Vision and Applications*, 21(5):713–725.

Strecha, C., T. Tuytelaars, and L. Van Gool, 2003. Dense matching of multiple wide baseline, *Proceedings of the 9th IEEE International Conference on Computer Vision*, 14–17 October, Nice, France, pp. 1194–1201.

Schmid, C., R. Mohr, and C. Bauckhage, 2000. Evaluation of interest point detectors, *International Journal of Computer Vision*, 37(1):151–172.

Tuytelaars, T., and L. Van Gool, 2000. Wide-baseline stereo matching based on local, affinely invariant regions, *Proceedings of the 11th British Machine Vision Conference (BMVC)*, 11–14 September, Bristol, UK, pp. 412–422.

Zhang Z., and J. Zhang, 1997. *Digital Photogrammetry*, Press of the Wuhan Technical University of Surveying and Mapping, Wuhan, China, 296 p.

(Received 13 July 2010; accepted 24 September 2010; final version 07 February 2011)