A machine vision based autonomous navigation system for Lunar rover: the model and key technique

Hui Li
Planetary Research Institute, China University of Geosciences, Wuhan, China and Key Laboratory of Watershed Ecology and Geographical Environment Monitoring, NASG, China, and
Cheng Zhong
Three Gorges Research Center for Geo-hazard, Ministry of Education China, University of Geosciences, Wuhan, China

Abstract
Purpose – This study aims to find a feasible precise navigation model for the planed Lunar rover. Autonomous navigation is one of the most important missions in the Chinese Lunar exploration project. Machine vision is expected to be a promising option for this mission because of the dramatic development of an image processing technique. However, existing attempts are often subject to low accuracy and errors accumulation.
Design/methodology/approach – In this paper, a novel autonomous navigation model was developed, based on the rigid geometric and photogrammetric theory, including stereo perception, relative positioning and absolute adjustment. The first step was planned to detect accurate three-dimensional (3D) surroundings around the rover by matching stereo-paired images; the second was used to decide the local location and orientation changes of the rover by matching adjacent images; and the third was adopted to find the rover's location in the whole scene by matching ground image with satellite image. Among them, the SURF algorithm that had been commonly believed as the best algorithm for matching images was adopted to find matched images.
Findings – Experiments indicated that the accurate 3D scene, relative positioning and absolute adjustment were easily generated and illustrated with the matching results. More importantly, the proposed algorithm is able to match images with great differences in illumination, scale and observation angle. All experiments and findings in this study proved that the proposed method could be an alternative navigation model for the planed Lunar rover.
Originality/value – With the matching results, an accurate 3D scene, relative positioning and absolute adjustment of rover can be easily generated. The whole test proves that the proposed method could be a feasible navigation model for the planed Lunar rover.

Keywords Machine vision, Image processing, Navigation

Paper type Research paper

1. Introduction

To conduct scientific exploration on the Lunar surface, Lunar rover of the Chinese Lunar exploration project must have the ability to execute tasks in an unstructured environment (Enright *et al.*, 2012). To guarantee that the rover can move smoothly in the different and unstructured surroundings, its navigation system must have a high degree of autonomy and the capabilities of high-accuracy and immediate positioning (Ning, and Fang, 2009).

Lunar rover navigation techniques mainly include absolute positioning and relative positioning. Relative positioning could achieve high accuracy of position and heading in a short time, but the errors probably accumulate over time (which may lead to divergence). Inertial navigation is a typical relative positioning technique, in which accelerometers and gyroscopes are used to track the position and orientation of an object relative to a known starting point, orientation and velocity. Inertial measurement units (IMUs) typically contain three orthogonal rate-gyroscopes and three orthogonal accelerometers, measuring angular velocity and linear acceleration, respectively. IMU is subject to errors accumulating for long distance navigation (Hinueber, 2011). Similarly, dead reckoning may provide the best available information on position, but it is subject to significant errors due to many factors, as both speed and direction must be accurately known at all instants for a position to be determined accurately (Murphy, 2011).

In absolute positioning techniques, such as satellite navigation systems, position and heading errors are bounded and do not accumulate over time, and the output is discrete. However, it is hard to find accurate references in the unknown surroundings on Lunar surface; thus, some commonly used

The current issue and full text archive of this journal is available on Emerald Insight at: **www.emeraldinsight.com/0260-2288.htm**

Sensor Review
36/4 (2016) 377–385

[DOI 10.1108/SR-01-2016-0001]

The present study is supported by Natural Science Foundation of China (41102209, 41102210), Natural Science Foundation of Hubei Province (2010CDB04105), and supported by the Key Laboratory of Watershed Ecology and Geographical Environment Monitoring, NASG.

Received 1 January 2016
Revised 29 April 2016
Accepted 19 May 2016

navigation methods on the Earth are not useful. Although the accuracy and speed of the typical satellite navigation systems, such as the global positioning system, Beidou, are satisfactory, unfortunately, there is no satellite navigation in the Lunar yet (Gibbons, 2008). If we use the radio navigation, the rover control may fail because the speed and accuracy of this method cannot meet the requirement of an autonomous navigation. Also, the Lunar magnetic field is very weak, so magnetic sensor-based methods are ineffective.

Celestial navigation uses angular measurements (sights) between celestial bodies and the visible horizon to locate one's position on the globe, on land as well as at sea. The earliest researchers (Krotkov *et al.*, 1994; Volpe, 1999; Benjamin, 2001) carried out celestial navigation by the altitude difference method through observing the sun, Earth and fixed stars. Recent researchers use vector observations-based quaternion estimation (QUEST) to get the rover heading angle (Trebi-Ollennu *et al.*, 2001; Ali *et al.*, 2005; Yue *et al.*, 2005, 2006). Ashitey proposed an absolute heading detection method for the field-integrated design and operation (FIDO) rover (Ning and Fang, 2006). Ning established a position and attitude determination method based on celestial observations (Ning and Fang, 2009). Ning proposed a Lunar rover kinematics model-based augmented unscented particle filter (ASUPF) as a new autonomous celestial navigation method for dealing with systematic errors and measurement noise (Benjamin, 2001). Pei *et al.* (2009) proposed a strapdown inertial navigation and celestial-navigation-based integrated method for Lunar rovers. In summary, the celestial navigation methods do not combine celestial positioning with orientation and cannot get the absolute heading and location information in real time.

Recently, a machine-vision-based Lunar rover navigation has become a promising option because of the dramatic development of an image processing technique. Visual navigation could provide autonomous positioning by perceiving the environment the system navigates and tracking visual clues, such as objects or landmarks in neighbor images. Furthermore, the Lunar surface is rather rough and complicated, as craters and rocks randomly are distributed all over the surface (Morita *et al.*, 2006). Machine vision is able to detect those obstacles in a short time, thus improving the navigation accuracy and efficiency dramatically. The JPL used a vision system in planetary rover navigation in their Mars exploration (Matthies *et al.*, 1997; Tunstel *et al.*, 2002; Maki *et al.*, 2005). Matthies introduced a stereo vision system for Robby (Matthies *et al.*, 1995), which captured images by two cameras with a 25 cm baseline and then calculated the height of objects with a disparity map. This rover could autonomously move about 100 m in a dry riverbed. "Mars Pathfinder" was equipped with two multi-spectral cameras with a 15 cm baseline, for mapping and positioning (Matthies *et al.*, 1995). To avoid an obstacle and plan a path, the Mars exploration rover (MER) generated a three-dimensional (3D) topographic map with stereo vision (Deen and Lorre, 2005; Wright *et al.*, 2005; Cheng *et al.*, 2006). Actual work indicated that the navigating error was less than 1 cm within 1 m. In Ralter, a sample of Lunar rover, two charge-coupled device (CCD) cameras were used as stereo vision sensors (Alexander *et al.*, 2005; Litwin and Maki, 2005; Biesiadecki *et al.*, 2005). In the system, images were subdivided and partially resampled to improve the matching speed. Ralter could move 1,021 m autonomously in a rough scoria area. To make full use of the field of view, carnegie Mellon University (CMU) developed a stereo vision system with a vertical baseline for highly mobile multi-wheeled vehicle (HMMWV) (Backes *et al.*, 2005). The European Space Agency (ESA) also developed a vision navigation system for planetary detection which consisted of two cameras with a free yaw and pitch angle. Experiments indicated that the accuracy of relative calibration and stereo matching could meet the requirements of Lunar rover. The centre national d'études spatiales (CNES) started the autonomous rover plan for Lunar and Mars exploration (Biesiadecki *et al.*, 2007; Leger *et al.*, 2005; Baumgartner *et al.*, 2005). Its vision system passed the test of autonomous navigation in the simulated Mars surface with various illuminations. In the institut national de recherche en informatique et en automatique (INRIA) and the French autonomous planetary rover program (VAP rover), a novel method for checking accuracy of matching images, an iteration algorithm to reduce computation amount and a weak calibration for 3D reconstruction were put forward to realize immediate positioning (Faugeras *et al.*, 1993). Experiments indicated that the speed of image processing was dramatically improved in this system, without losing image quality.

Although many innovations on camera calibration, stereo matching and 3D reconstruction have been presented by institutes and researchers, most of them are subject to below problems. First, affected by the distance from the camera to an object, the baseline between two cameras, and the resolution of an image, the accuracy of detecting far targets is often low; second, due to the great change of the rover's posture, scale and illumination, matching images at different location is probably very difficult and inefficient; third, the relative positioning process will lead to accumulation of errors.

In this paper, a machine-vision-based autonomous navigation model for the Lunar rover is put forward. First, stereo perception was formed by matching stereo-paired images to detect accurate 3D surroundings around the rover. Second, relative positioning was done by matching neighbor images to decide the local location and orientation changes of the rover. Finally, absolute adjustment was carried out to find the rover's location in a whole scene by matching the ground image with satellite image. The speeded up robust features (SURF) that had been commonly believed to be the best algorithm in the field was adopted to find matched images. With the matching results, accurate 3D scene, relative positioning and absolute adjustment were easily generated and illustrated. With the matching strategy, the model is expected to realize a fast and accurate autonomous navigation for Lunar rover.

2. Methods

The machine-vision-based autonomous navigation model (Figure 1) consists of three parts: the stereo perception model, the relative positioning model and the absolute adjustment model.

The stereo perception model takes as an input a stereo pair and as outputs arrays of the three coordinates X, Y and Z of

Figure 1 The model of autonomous navigation

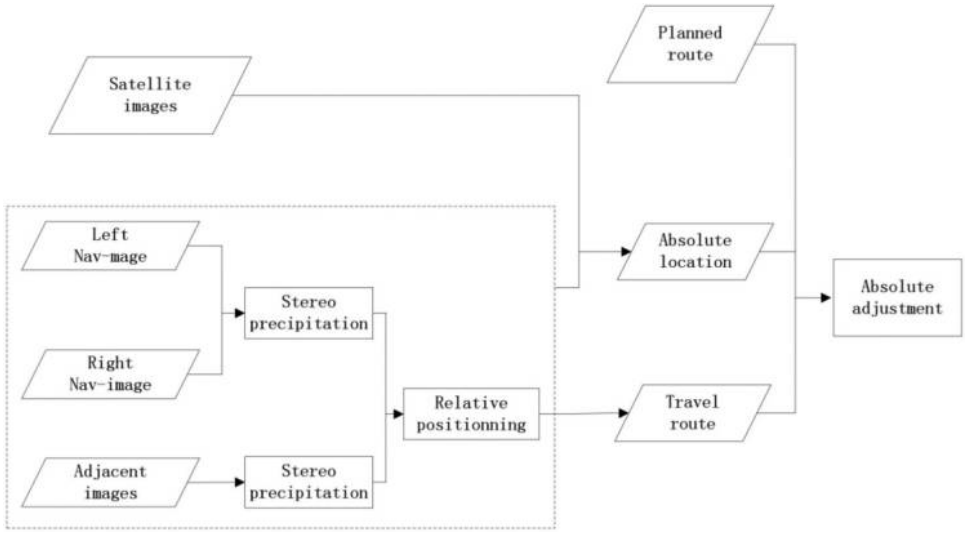


image pixels belonging to references in the camera frame. The points measured by stereo pairs are accumulated into a terrain map centered at the vehicle which provides the local view for objects detection.

The main task of the relative positioning model, namely, finding the change of location and posture between neighbor stereo pairs, includes tracking of features and calculation of parameters, for finding the change of location and posture between neighbor stereo pairs.

The absolute adjustment is used to adjust the accumulated errors from relative positioning, by matching the local images with images from outer equipments. The matching may be very difficult, as the view angle of cameras on the Lunar rover is different from that on the satellites.

2.1 Stereo perception

2.1.1 Stereo perception

Stereo perception is used to detect 3D surroundings around the rover. The two images from neighbor cameras which should cover each other more than 60 per cent are called stereo pairs. Generally, the inner parameters of two cameras are the same and are calibrated before use. Figure 2 shows the principle of Stereo Perception, where the center of camera lens is called the optical center, the line connecting two optical centers is the baseline B, the distance from the optical center to the image is the focal length f and the line through the optical center and perpendicular to the image plane is called the ray axis. The two ray axes are usually parallel to simplify the computation.

Given the projections of an object $W(X, Y, Z)$ are (x_1, y_1) and (x_2, y_2) in the two images, respectively. From Figure 2, it is easy to find the simple geometry relationship as:

$$x_1 = f\frac{X}{Z}, y_1 = y_2 = f\frac{Y}{Z}, x_2 = f\frac{X-B}{Z}. \quad (1)$$

Then, the coordinates of the object W could be solved from equation (1):

$$X = \frac{B \cdot x_1}{D}, Y = \frac{B \cdot y_1}{D}, Z = \frac{B \cdot f}{D} \quad (2)$$

Here, $D = x_1 - x_2$, is called the parallax of stereo vision. The accuracy of solving object's coordinates with equation (2) depends on the length of the baseline. According to equation (2), the parallax is figured out by matching conjugate projections of the object W in two images.

2.1.2 Epipolar geometry

To speed up the process of matching stereo images, the epipolar geometry is adopted, as shown in Figure 3. Here, an unrotated image is produced by placing a virtual image plane in front of each camera's projection center to produce, to simplify the projection geometry. In Figure 3, O_L and O_R represent cameras' projection centers, X represents the point

Figure 2 The principle of stereo perception

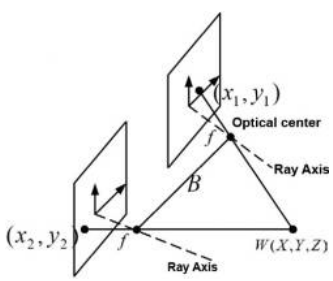
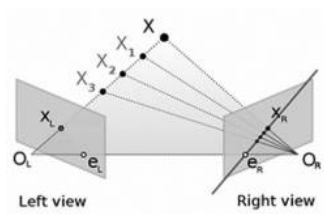


Figure 3 The epipolar geometry



of interest in both cameras, points x_L and x_R indicate the projections of point X onto image planes. The points e_L and e_R are called epipoles or epipolar points, and the plane formed by points X, O_L and O_R form is called the epipolar plane. The intersects of epipolar plane and camera's image plane are called epipolar lines.

In the calibrated environment, we capture epipolar geometric constraint in an algebraic representation known as the *essential* matrix. In the uncalibrated environment, it is captured in the *fundamental* matrix. With the two views, the two-camera coordinate systems are related by a rotation R and a translation T:

$$W_1 = RW_2 + T \tag{3}$$

Taking the vector product with T, followed by the scalar product with W_1, we obtain:

$$W_1 \cdot (T \wedge RW_2) = 0 \tag{4}$$

This can also be written as:

$$W_1^T E W_2 = 0 \tag{5}$$

where:

$$E = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \cdot R$$

is the essential matrix, and $T = (t_x, t_y, t_z)^{\mathrm{T}}$. Equation (5) is the algebraic representation of epipolar geometry for known calibration, and the essential matrix relates corresponding image points expressed in the camera coordinate system.

According to equations (3) to (5), if the relative translation and rotation between two cameras are known, the corresponding epipolar geometry leads to two important observations:

- If the projection point x_L is known, the epipolar line $e_R - x_R$ becomes known and the projection of X must be in the line. Namely, the conjugate of a point could be observed in a known epipolar line. This epipolar constraint could be used to judge whether the two points are the projections of a 3D point.
- If the conjugate points x_L and x_R are known, their projection lines are also known. If the two image points are the projections of a 3D point X, their projection lines must intersect at X.

Matching the neighbor image is critical to build the Epipolar geometry. After that, a 3D object could easily be reconstructed just by searching conjugate features in corresponding epipolar lines.

2.2 Relative positioning

Relative positioning could be applied by tracking evident landmarks in two or more stereo pairs. Then, the movement of cameras could be figured out. The process consists of tracking landmark and calculating transformation parameters. In this study, a 3D transformation of the rover could be figured out, as the 3D coordinates of landmarks have been provided by the stereo perception. It should be noticed that sufficient overlap between stereo pairs are required to detect as many landmarks as possible.

2.2.1 Tracking landmarks

Landmark tracking means matching landmarks in different stereo pairs, consisting of:

- estimating the change of location and posture between two moments, with the speedometer and the attitude measurement tool;
- estimating the 3D and 2D coordinates of the landmarks at the current moment; and
- finding landmarks around the estimating coordinates in the current image. In this process, it is hard to reduce the search work by using epipolar geometry.

It should be noticed that the overlap between stereo images depending on the relative speed between the rover and the shooting is often less than that used in relative positioning.

2.2.2 Solving transformation parameters

As cameras are fixed on the rover, the transformation relationship between them at different moments could reflect the movement of the rover. Given the rover is a rigid body, when conjugated landmarks are matched in different stereo pairs, the transformation parameters could be calculated as:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = R \cdot \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + \begin{bmatrix} t_x \\ t_x \\ t_z \end{bmatrix} \tag{6}$$

where:

$$R = \begin{bmatrix} \cos\phi\cos\kappa & -\cos\phi\sin\kappa & -\sin\phi \\ \cos\omega\sin\kappa - \sin\omega\sin\phi\cos\kappa & \cos\omega\cos\kappa + \sin\omega\sin\phi\sin\kappa & -\sin\phi\sin\omega \\ \sin\omega\sin\kappa - \cos\omega\sin\phi\cos\kappa & \sin\omega\cos\kappa - \cos\omega\sin\phi\sin\kappa & \cos\omega\cos\phi \end{bmatrix}$$

Here, φ, θ, ω indicate the pitching, roll and drift angles, respectively, (X_c,Y_c,Z_c); (X_p,Y_p,Z_p) indicate the 3D coordinates of the landmark at current and previous time, respectively; and (t_x, t_y, t_z) indicate the translation between images at two moments.

2.2.3 Optimizing strategy

To reduce the influence of random errors in the matching process, RANSAC is adopted to find the best subset of matched pairs. It includes:

1 randomly selecting a small subset of matched landmarks and solving the transformation parameters with least square rule;
2 assessing the accuracy and judging whether it is satisfied; and
3 repeating Steps 1 and 2 until finding the subset that has the best accuracy in the solving process.

The relative positioning is expected to be able to handle great transformations between stereo pairs, in case the rover has to move quickly at some time. The method should also be competent with continuous navigation especially in tough area.

2.3 Absolute adjustment

Several methods have been presented for absolute adjustment. Matching the local terrain with global terrain used in celestial navigation is complex and unreliable. Comparatively, matching the image captured by the rover with the satellite image is believed to be easier and more accurate. However, the matching should be very hard due to the great difference in the observation angle and resolution between both the images. In the MER project, images were matched by manual work, which needed considerable time and label resources. In this study, the SURF algorithm is adopted to test whether the images could be matched automatically.

Then, the absolute location of the rover could be figured out by calculating the transformation parameters between the two images. After that, the rover's movement would be checked and adjusted to the planned routes. The 2D affine transformation as equation (7) is adopted to depict the relationship between images:

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_l \\ y_l \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{7}$$

Here, (x_s, y_s) and (x_l, y_l) are coordinates in the two images, respectively, and (t_x,t_y) and θ indicate the translation and the rotation angle, respectively.

2.4 Matching images with SURF

Matching images is one of the most fundamental and critical task for stereo precipitation, relative positioning and absolute adjustment. Thus, the speed, accuracy and reliability of the matching method lay great effects on the performance of the whole visual navigation system. Furthermore, the matching algorithm is expected to handle images with a great difference in the observation angle, rotation angle, scale and illumination, as the rover should be competent with complex and rough Lunar surfaces.

A wide variety of methods have already been proposed in the literatures, such as Harris (Stephens and Harris, 1988), the smallest univalue segment assimilating nucleus (SUSAN) (Smith and Brady, 1997), difference of Gaussian (DOG) (Lowe, 2004), Hessian-Laplace (Mikolajczyk and Schmid, 2004), scale-invariant feature transform (SIFT) (Lowe, 1999) and SURF (Bay *et al.*, 2008). Among them, SURF has been proved as one of the most robust, distinctive and efficient methods for matching images (Bay *et al.*, 2006). It could keep invariant to any change of the rotation, scale and illumination, and produce enough conjugate pairs whether images are full of features.

2.4.1 SURF detector

SURF feature detection runs on an integral image. Given an input image I and a point (x, y), the integral image I_Σ is calculated by summing the values between the point and the origin. Formally, this can be defined by the formula:

$$I_\Sigma(x) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i,j) \tag{8}$$

The SURF detector is based on the determinant of the Laplacian of Gaussians matrix. An approximation to t is using box filter representations to replace the Gaussian kernels, which can be represented as:

$$H_{approx}(x,\sigma) = \begin{bmatrix} D_{xx}(x,\sigma) & D_{xy}(x,\sigma) \\ D_{xy}(x,\sigma) & D_{yy}(x,\sigma) \end{bmatrix} \tag{9}$$

Here, D_{xx} (x,σ) means the convolution of the second-order box filter at scaleσ derivative with the image at point (x, y) and similarly for D_{xy}, D_{yy}. The determinant of the approximated Gaussians is:

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \tag{10}$$

Interest point detection operates on a scale space. In SIFT, the scale space is implemented as an image pyramid where the input image is iteratively convolved with the Gaussian kernel and repeatedly sub-sampled (reduced in size). In SURF, the scale space can be created by applying kernels of increasing size to the original image.

The task of localizing the scale and rotation invariant interest points in the image can be divided into three steps. First, the response values that are below the predetermined threshold are removed. Then, a non-maximal suppression is performed to find a set of candidate points. Finally, feature points are interpolated to find the location in both the space and scale to sub-pixel accuracy.

2.4.2 SURF descriptor

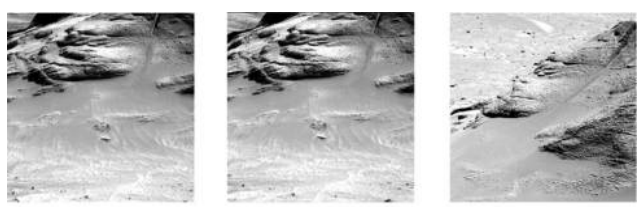
To generate the descriptor, first, a dominate orientation is needed, and then the Haar wavelet response is calculated. To determine the orientation, Haar wavelet responses are calculated for a set pixel within a radius of 6 σ of the detected point, where σ refers to the scale. A circle segment covering an angle of $3/\pi$ is rotated around the origin. At each position, the x and y responses within the segment are summed to form a new vector. The longest vector lends its orientation to the interest point.

To get SURF descriptor, we construct a square window around the interest point which contains the pixels around the feature point of size 20 σ, and then orient the square along the dominant orientation. The descriptor window is divided into 4×4 regular subregions. Within each of these subregions, Haar wavelets are calculated. If we refer to the x and y wavelet responses by d_x and d_y, respectively, we get:

$$Desc\_squre = V\left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|\right) \tag{11}$$

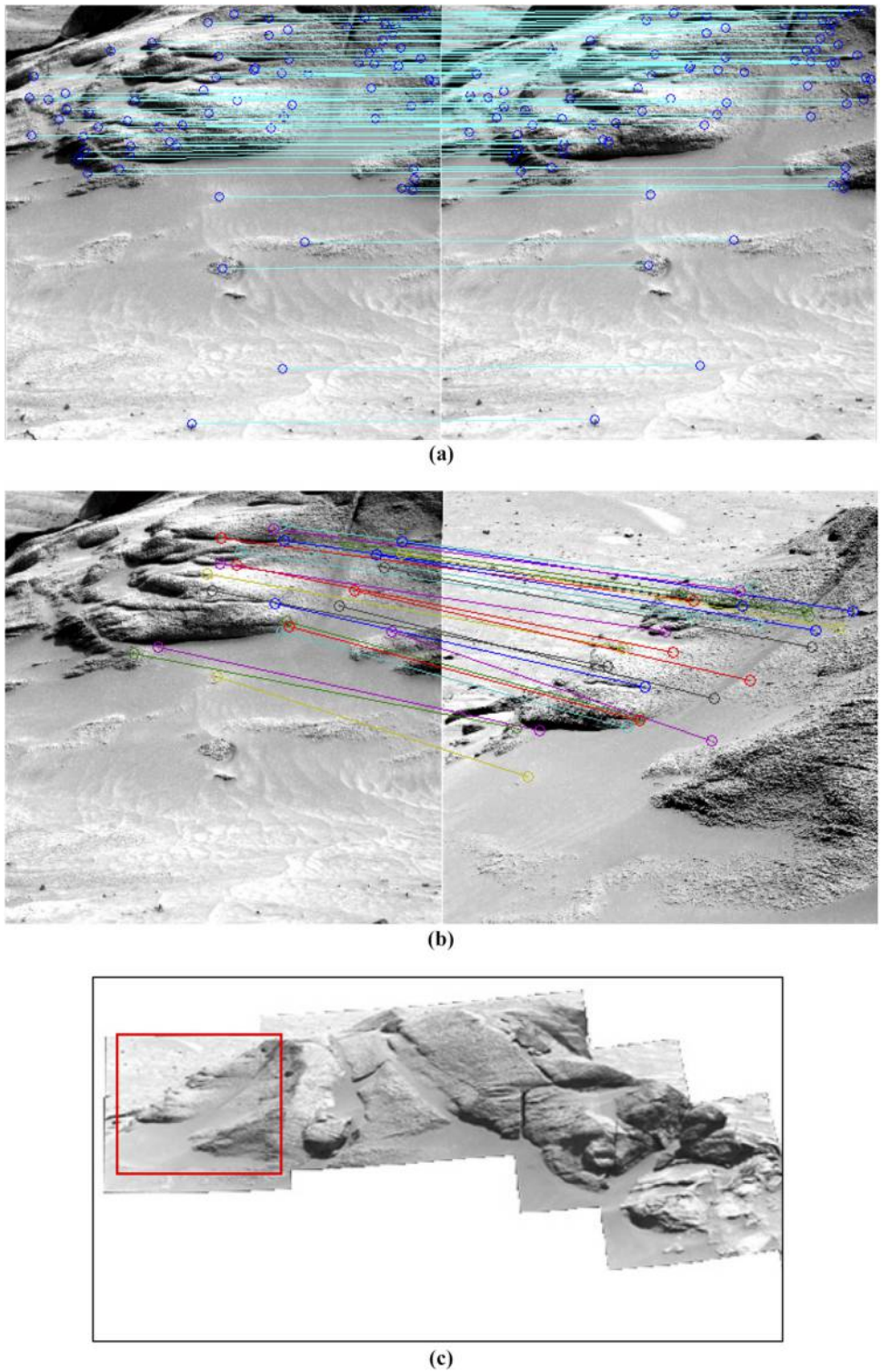
It is a four-dimensional vector for all 4×4 subregions.

Figure 4 Mars ground images



(a) (b) (c)

Figure 5 Matching results



Notes: Matched images are shown in one picture, with lines connecting conjugated features (little circles). Two per cent matches are shown in Figure 5(a) and (b) to display matching results clearly; (a) Matching paired images; (b) matching images captured at different locations; (c) merging matched images

2.4.3 Feature correspondences

Feature correspondence uses the k-d tree nearest neighbor search algorithm base on Euclidean distance. For SURF descriptors, it uses squared distances for the comparison to avoid computing square roots. The squared Euclidean distance is defined as:

$$d(desc1, desc2) = \sqrt{\sum_{i=1}^{64} |desc1(i) - desc2(i)|^2} \quad (12)$$

A homograhy transform is assumed to exist between images to simplify the computation process.

2.4.4 Homograhy matrix estimation

Conjugate points $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$ depicted with a homography matrix in the two images have a relationship as presented in equation (13):

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = H \times \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \quad (13)$$

Then, a consensus model is estimated with RANSAC to minimize matrix error. The iterative process first chooses a minimum subset randomly to fit the transformation model, and then tests all other data with the fitted model. If a point fits the estimated model very well, it would be considered as a hypothetical inlier. At last, the final model could be re-estimated and evaluated with all hypothetical inliers.

3. Experiments and discussions

Images from the Mars' rover are used in the experiments, due to lacking images from Lunar rover. We assume that the rovers and images are similar to that on Lunar surface, so that the experiment results will be useful in the future Lunar rover exploration.

Ground images captured by the opportunity and spirit are downloaded from http://an.rsl.wustl.edu/mer/mera. The image size is 1,024 × 1,024. A pair of images (Figure 4(a) and (b)) and a left image at an adjacent position (Figure 4(c)) are used in our experiment.

Figure 4(a) and (b) similar in all aspects, including illumination, scale and so on. Although matching paired images is often not difficult, its accuracy and efficiency must meet some especial requirements of the Lunar rover, such as immediate response to various terrains. Great differences in illumination, scale and observation angle are found between Figure 4(a) and (c). Especially, the difference in observation angles between them leads to different geometric and gray features between the two images. In this case, it becomes challenging to match them with the most register algorithms.

Figure 5(a) shows the result of matching pared images, which is finished within 0.06 s. Through carefully checking, we assure the error rate is lower than 0.01 per cent and could easily be removed by the RANSAC. From the matching result, the stereo visual, 3D models, slope map and other products could be produced and then used to assist avoiding obstacles and planning routes.

Figure 5(b) shows the matching result between images captured at different locations. Checking the connections between features in the two images carefully, it is found that all matches are correct and accurate and not affected by the great differences in the observation angle. Comparatively, a test with ENVI 5.0 image registration module shows no correct matched features in the images. The results prove that the proposed method has great advantages over typical commercial software on this. With the matching result, geometric relationship between different locations where camera is, such as the rotation, translation and scale parameters, could be figured out. Finally, the actual route consisting of all locations the rover ever stayed, could be mapped.

Figure 5(c) shows the result of merging several matched navigation images, which illustrates a hill's profile. The red rectangle embraced the test images in Figure 4. The result indicates that the proposed method is able to handle images with different observation angle, as their combinations could depict whole targets without any distortion.

Figure 6 shows another example of matching rover's ground images with the proposed method. Figure 6(a) shows that matching images with similar illumination and observation angle is very easy. In Figure 6(b), the rock embraced by a red circle shows very different sizes, colors and coordinates in two images, which is common in the Lunar rover 's photo album. The great number of connections between conjugated features indicates that the proposed method could handle this situation easily. Experiments indicate that the proposed method is able to meet the basic requirements of an autonomous navigation system, e.g. high accuracy and immediate response.

Figure 7 shows an example of autonomous navigation, where a red line indicates the planned route, and the red star represents a rover. The absolute adjustment module is able to judge whether the rover moved along the planned route at

Figure 6 Another example

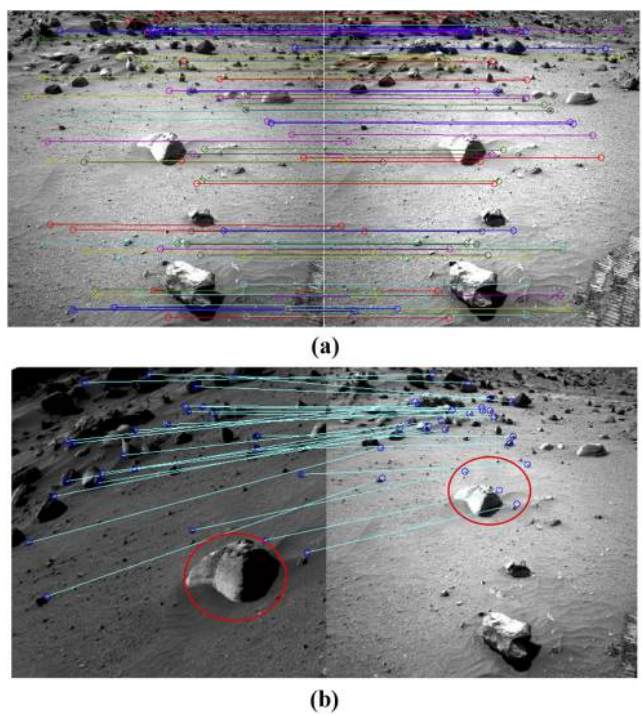
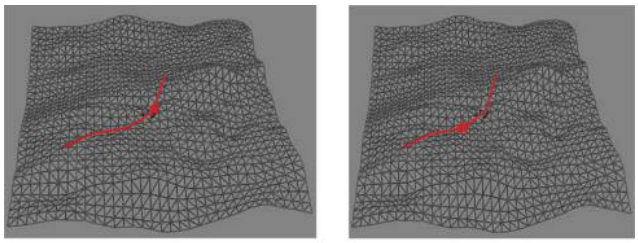


Figure 7 An example of the autonomous navigation



each step by matching local images with satellite images, as shown in Figure 7(a) and (b). The experiment shows that all steps, including the stereo perception, relative positioning and absolute adjustment, could complete in one second. This is thought to be able to satisfy the requirement of immediate response.

4. Conclusions

Machine vision is expected to be a promising option for the Lunar rover autonomous navigation in the Chinese Lunar exploration project. In this paper, a novel autonomous navigation model was developed, based on rigid geometric and photogrammetric theory and solutions. First, the stereo perception was formed by matching stereo-paired images to detect accurate 3D surroundings around the rover. Second, relative positioning was done by matching neighbor images to decide the local location and orientation changes of the rover. Finally, the absolute adjustment was carried out to find the rover's location in the whole scene by matching ground image with satellite image. The SURF that had been commonly believed as the best algorithm in the field was adopted to find matched images. Experiments prove that our algorithm is able to match images with great differences in illumination, scale and observation angle. Through careful visual checking, the error rate in the matching process is lower than 0.01 per cent and could be easily removed by the RANSAC. With the matching results, the actual route consisting of all locations the rover ever stayed, could be mapped, and then the rover's movement could be adjusted by the absolute adjustment module. All experiments and findings in this study proved that the proposed method could be an alternative navigation model for the planed Lunar rover.

References

Alexander, D., Zamani, P., Deen, R., Andres, P., Mortensen, H. (2005), "Automated generation of image products for mars exploration rover mission tactical operations", *Proceedings of IEEE Interntional Conference Systems, Man and Cybernetics, The Big Island, HI, No.* 1, pp. 923-929.

Ali, K.S., Vanelli, C.A., Biesiadecki, J.J., Maimone, M.W., Cheng, Y., San Martin, A.M. and Alexander, J.W. (2005), "Attitude and position estimation on the Mars exploration rovers", *Proceedings of the IEEE Systems, Man and Cybernetics Society, The Big Island, HI*, pp. 20-27.

Backes, P., Calderon, A., Robinson, M., Bajracharya, M. and Helmick, D. (2005), "Automated rover positioning and instrument placement", *Proceedings of IEEE Aerospace Conference, Big Sky, MT*, pp. 1-12.

Baumgartner, E., Bonitz, R., Melko, J., Shiraishi, L. and Leger, P. (2005), "The Mars exploration rover instrument positioning system", *Proceedings of IEEE Aerospace Conference, Big Sky, Nontana*, pp. 1-19.

Bay, H., Ess, A., Tuytelaars, T. and Gool, L.V. (2008), "Speeded-Up Robust Features (SURF)", *Computer Vision and Image Understanding*, Vol. 110 No. 3, pp. 346-359.

Bay, H., Tuytelaars, T. and Van Gool, L. (2006), "SURF: speeded up robust features", in Leonardis, A., Bischof, H., Pinz, A. (Eds), *ECCV 2006: LNCS*, Vol. 3951, Springer, Heidelberg, pp. 404-417.

Benjamin, P.M. (2001), *Celestial Navigation on the Surface of Mars*, Naval Academy, Annapolis, MD.

Biesiadecki, J., Baumgartner, E., Bonitz, R.G., Cooper, B.K., Hartman, F.R., Leger, P.C., Maimone, M.W., Maxwell, S.A., Treibi-Ollenu, A., Tunstel, E.W. and Wright, J.R. (2005), "Mars exploration rover surface operations: driving opportunity at Meridiani Planum", *Proceedings of IEEE International Conference on Systems, Man and Cybernetics, The Big Island, HI, No.* 2, pp. 1823-1830.

Biesiadecki, J., Leger, C. and Maimone, M. (2007), "Tradeoffs Between directed and autonomous driving on the Mars exploration Rovers", *International Journal of Robotics Research*, Vol. 26 No. 1, pp. 91-104.

Cheng, Y., Maimone, M. and Matthies, L. (2006), "Visual odometry on the Mars exploration rovers", *IEEE Robotics and Automation Magazine*, Vol. 13 No. 2, pp. 54-62.

Deen, R. and Lorre, J. (2005), "Seeing in three dimensions: correlation and triangulation of Mars exploration rover imagery", *Proceedings of IEEE International Conference on Systems, Man and Cybernetics, The Big Island, HI, No.* 1, pp. 911-916.

Enright, J., Barfoot, T., Soto, M. (2012), "Star tracking for planetary rovers", *IEEE Aerospace Conference, Big Sky, MT*, pp. 1-13.

Faugeras, O., Hotz, B., Mathieu, H., Vieville, T., Zhang, Z., Pascal, F., Theron, E., Moll, L., Berry, G., Vuillemin, J., Bertin, P. and Proy, C. (1993), "Real time correlation-based stereo: algorithm, implementations and applications", INRIA Technical Report, pp. 1-49.

Gibbons, G. (2008), "China GNSS 101: compass in the rearview mirror", *Inside GNSS*, January/February 2008, pp. 62-63.

Hinueber, E.V. (2011), "Design of an unaided aircraft attitude reference system with medium accurate gyroscopes for higher performance attitude requirements", *Inertial Sensors and Systems – Symposium Gyro Technology, Karlsruhe/Germany (iMAR Navigation/DGON)*.

Krotkov, E., Hebert, M., Bufa, M., Cozman, F. and Robert, L. (1994), "Stereo driving and position estimation for autonomous planetary rovers", *Proceedings of the IARP Workshop on Robotics in Space, Montreal.*

Leger, P., Deen, R. and Bonitz, R. (2005), "Remote image analysis for mars exploration rover mobility and manipulation operations", *Proceedings of IEEE Internationa'l Conference on Systems, Man and Cybernetics, Hawaii*, No. 1, pp. 917-922.

Litwin, T. and Maki, J. (2005), "Imaging services flight software on the Mars exploration rovers", *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, *HI*, No. 1, pp. 895-902.

Lowe, D.G. (1999), "Object recognition from local scale-invariant features", *Proceedings of the International Conference on Computer Vision*, *Kerkyra*, Vol. 2, p. 1150.

Lowe, D.G. (2004), "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, Vol. 60 No. 2, pp. 91-110.

Maki, J., Litwin, Schwochert, T. and Herkenhoff, M.K. (2005), "Operation and performance of the Mars exploration rover imaging system on the martian surface", *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, *HI*, No. 1, pp. 930-936.

Matthies, L., Balch, T. and Wilcox, B. (1997), "Fast optical hazard detection for planetary rovers using multiple spot laser triangulation", *Proceedings of IEEE International Conference on Robotics and Automation*, *Albuquerque, NM*, pp. 859-866.

Matthies, L., Gat, E., Harrision, R., Wilcox, B., Volpe, R. and Litwin, T. (1995), "Mars microrover navigation: performance evaluation and enhancement", *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, *Pittsburgh, PA*, pp. 433-440.

Mikolajczyk, K. and Schmid, C. (2004), "Scale and affine invariant interest point detectors", *International Journal of Computer Vision*, Vol. 60 No. 1, pp. 63-86.

Morita, H., Hild, M., Miura, J. and Shirai, Y. (2006), "Panoramic view-based navigation in outdoor environments based on support vector learning", *Proceedings of IEEE International Conference on Intelligent Robots and Systems (IROS)*, *Beijing*, pp. 2302-2307.

Murphy, C. (2011), Believable Dead Reckoning for Networked Games, Game Engine Gems 2, Lengyel, Eric. AK Peters, pp. 308-326.

Ning, X. and Fang, J. (2009), "A new autonomous celestial navigation method for the lunar rover", *Robotics and Autonomous Systems*, Vol. 57 No. 1, pp. 48-54.

Ning, X.L. and Fang, J.C. (2006), "Position and pose estimation by celestial observation for lunar rovers", *Journal of Beijing University of Aeronautics and Astronautics*, Vol. 32 No. 7, pp. 756-759.

Pei, F.J., Ju, H.H. and Cui, P.Y. (2009), "A long-range autonomous navigation method for lunar rovers", *High Technology Letters*, Vol. 19 No. 10, pp. 1072-1077.

Smith, S.M. and Brady, J.M. (1997), "SUSAN-A new approach to low level image processing", *International Journal of Computer Vision*, Vol. 23 No. 1, pp. 45-78.

Stephens, C. and Harris, C. (1988), "A combined corner and edge detection", *Proceedings of The Fourth Alvey Vision Conference*, *University of Manchester*, pp. 147-151.

Trebi-Ollennu, A., Huntsberger, T., Cheng, Y., Baumgartner, E.T., Kennedy, B. and Schenker, P. (2001), "Design and analysis of a sun sensor for planetary rover absolute heading detection", *IEEE Transactions on Robotics and Automation*, Vol. 17 No. 6, pp. 939-947.

Tunstel, E., Huntsberger, T., Aghazarian, H., Backers, P., Baumgartner, E., Cheng, Y., Garrett, M., Kennedy, B., Leger, C., Magnone, L., Norris, J., Powell, M., Trebi-Ollennu, A. and Schenker, P. (2002), "FIDO Rover field trials as rehearsal for the NASA 2003 Mars exploration rovers mission", *Proceedings of 5th Biannual World Automation Congress*, *Orlando, Florida*, No. 14, pp. 9-13.

Volpe, R. (1999), "Mars rover navigation results using sun sensor heading determination", *Proceedings of the IEEE/RSJ International Conference on Intelligent Robot and Systems*, *Kyongju*, pp. 460-467.

Wright, J., Trebi-Ollennu, A.F., Hartman, F., Cooper, B., Maxwell, S., Yen, J. and Morrison, J. (2005), "Terrain modelling for In-situ activity planning and rehearsal for the Mars exploration rovers", *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, *HI*, No. 2, pp. 1372-1377.

Yue, F.Z., Cui, P.Y., Cui, H.T. and Ju, H.H. (2005), "Earth sensor and accelerometer based autonomous heading detection algorithm research of lunar rover", *Journal of Astronautics*, Vol. 26 No. 5, pp. 553-557.

Yue, F.Z., Cui, P.Y., Cui, H.T. and Ju, H.H. (2006), "Algorithm research on lunar rover autonomous heading detection", *Acta Aeronautica et Astronautica Sinica*, Vol. 27 No. 3, pp. 501-504.

Corresponding author

Cheng Zhong can be contacted at: dr.zhong.c@gmail.com