Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 13-16 August 2006

# STEREO VISION BASED MOTION ESTIMATION FOR LUNAR ROVER NAVIGATION

**PING-YUAN CUI, FU-ZHAN YUE, HU-TAO CUI**

Deep Space Exploration Research Center, Harbin Institute of Technology, Harbin 150080, China
E-MAIL: cui@astro.hit.edu.cn, yuefzh@hit.edu.cn

**Abstract:**

**Accurate navigation autonomously on uneven terrain is a very important technique for lunar rover to execute long-range exploration on lunar surface. This paper presents a stereo vision based algorithm for lunar rover autonomous navigation that can enable for precision localization. Our techniques mainly consist of image processing and motion estimation. To improve the performance of motion estimation, robust linear motion estimation is executed to reject the outliers of image processing and estimate the motion initially, then, Levenberg-Marquardt nonlinear estimation is executed to estimate the motion precisely. The result of our autonomous navigation algorithm is an estimation of attitude and position, which can be passed directly to path planning and motion control system.**

**Keywords:**

**Stereo vision; Notion estimation; Lunar rover**

## 1. Introduction

In lunar exploration mission, it is very critical to determine the position of the lunar rover with a high degree of accuracy both for exploration achievement of scientific goals and for safe navigation. Localization of lunar rover when moving on lunar uneven terrain has been a very active research topic in the planetary exploration and robotics communities. With the development of deep space exploration, numerous localization schemes and techniques have be developed in the literature[1-7], but, for the special condition of lunar environment, some position estimation techniques cannot work on the Moon. Use of the popular Global Position System (GPS) is not feasible, since its signals are not readily detectable from the lunar surface and there is no plan of building GPS for the moon in the near future. Sonar and magnetic sensors cannot work for lunar rover[2]. Dead reckoning sensors including IMU and odometry are suffered from error accumulation, and cannot used for rover long range navigation.

Stereo vision based motion estimation for lunar rover navigation is an autonomously localization technique, and stereo vision can also provide more environment information which can be used for rover obstacle avoidance and lunar DEM construction. With the development of computer processing technology, the precision of image processing can reach a sub-pixel level, and time of image processing can be accepted in real time application. Therefore, the technique of stereo vision based motion estimation can be used on the lunar rover and improve its performance of autonomous navigation.

In this paper, we describe a stereo vision image based motion estimation algorithm, which can be used on board autonomously. The localization scheme mainly has two stages: image processing and motion estimation. In the motion estimation stage, to reject the outliers and improve the precision, we use a robust motion estimation algorithm followed with a nonlinear estimation method. The output of our motion estimation is the attitude and position parameter.

## 2. Images processing

The first step in stereo vision based motion estimation is sequence images processing, which includes feature detection, stereo matching and feature tracking[8,9].

### 2.1. Feature detection

Features that corresponding to high curvature points and being helpful to estimate motion parameter are detected in successive stereo pairs[10]. In our application, Harris corner detector is used to extract the interest point. To recover the feature position up to sub-pixel, we improve the performance of Harris operator by weighting the pixels and normalizing the gravity-center of pixels.

### 2.2. Stereo matching (1)

Stereo matching can estimate the 3D position of the landmarks that feature correspond. We use Zhang's correspondence method[11] to achieve feature matching, The

1-4244-0060-0/06/$20.00 ©2006 IEEE

procedure uses a correlation-based search algorithm firstly to forms a candidate match set. The correlation technique selects the consistent matches by a given threshold which the correlation score must be higher than, inevitably, a feature point in the left image may be paired several points in the right image, and vice versa. Relaxation technique is then used to reorganize the candidate matches by propagating continuity and uniqueness constraints, then the problem of disambiguating matching is resolved by minimizing the total strength of the match (SM). Triangulation using the known relative position between the cameras is then used to determine the position of the landmark with respect to the camera frame. The output of stereo matching is a 3D position set of the features correspond $\{L_i^b \mid i = 1, \cdots, n\}$.

## 2.3. Feature tracking

After the motion of lunar rover, the features change more or less in image plane. The procedure to locate the features detected before motion in the camera view after motion is called feature tracking. Feature tracking is some similar to stereo matching, only the difference between them is that superposition degree, for motion uncertain of lunar rover, the superposition degree of images before and after motion is larger than that of stereo images. In machine vision literature, Feature tracking can be split in to two groups of algorithms: correlation based methods and optical flow based methods[12]. Correlation based method can also be appropriate when the motion of feature in the image is expected to be large, so we use correlation-based method, the search window is larger than that of stereo matching in order to achieve high tracking performance.

## 2.4. Stereo matching (2)

A second stereo matching step is performed to estimate the 3D positions of landmarks that features correspond with respect to the stereo images after motion. This step is same as the previous stereo matching procedure, and triangulation is performed to estimate the position of landmarks in the camera frame after motion. The output of the second stereo matching is a 3D position set of the features correspond $\{L_i^a \mid i = 1, \cdots, n\}$.

The image processing procedure is a very important part of motion estimation, it's processing rapidity and precision can directly affect the performance of motion estimate. The output of image processing is the 3D position set of matching features $\{(L_i^b, L_i^a) \mid i = 1, \cdots, n\}$.

# 3. Motion estimation

Given the noisy landmarks positions from stereo data, we can estimate the motion between the stereo views. Assume that the motion between two stereo views can be described by a rigid transformation, the motion estimate model can be described as follows:

$$L_i^a = R_j L_i^b + T_j + e_j \tag{1}$$

Where $R_j$ and $T_j$ denote the rotation and translation between pre-move and post-move stereo views respectively, $e_j$ combines the errors in the observed positions of the landmarks at both locations.

Currently, image processing technique has reached a more perfect condition, however, it is hard to avoid errors. The errors mainly lie in two ways, first feature localization errors result from feature detection stage; second the outliers result from matching process. To improve the performance of lunar rover motion estimate, Errors described above must be eliminated effectively.

In our algorithm, motion estimation is divided into two stage process. First an initial estimate of the motion is computed using a orthogonalizing method, this algorithm is applied multiple times using different sets of features to eliminate feature matching and tracking outliers and determine a robust LMedS estimate of motion. Since precise estimate of motion is needed for lunar rover long range traverse, so the result of robust estimate is then used as input to a more accurate nonlinear algorithm that solves for the motion parameters directly.

## 3.1. Robust motion estimation

Robust motion estimation serves two purposes: to detect and reject the outliers and to provide an initial estimate of $R_j$ and $T_j$. The robust motion algorithm uses the orthogonalizing method to provide a closed form solution of motion parameter, and uses Least Median of Squares method to detect and reject the outliers.

## 3.2. Orthogonalizing method

The orthogonalizing method is a special algorithm for motion estimation which is used only in the condition of three pairs of non-collinear matching features. Selecting three elements in set of $\{(L_i^b, L_i^a) \mid i = 1, \cdots, n\}$, and let:

$$v_1^b = L_{n2}^b - L_{n1}^b \qquad v_2^b = L_{n3}^b - L_{n1}^b$$
$$v_1^a = L_{n2}^a - L_{n1}^a \qquad v_2^a = L_{n3}^a - L_{n1}^a$$

Then unit orthogonal basis can be formed as follows:

$$r_1^b = \frac{v_1^b}{\left\| v_1^b \right\|} \quad r_2^b = \frac{v_2^b - (v_2^b, r_1^b) r_1^b}{\left\| v_2^b - (v_2^b, r_1^b) r_1^b \right\|} \quad r_3^b = r_1^b \times r_2^b$$

$$r_1^a = \frac{v_1^a}{\left\| v_1^a \right\|} \quad r_2^a = \frac{v_2^a - (v_2^a, r_1^a) r_1^a}{\left\| v_2^a - (v_2^a, r_1^a) r_1^a \right\|} \quad r_3^a = r_1^a \times r_2^a$$

Where $(v_2^b, r_1^b)$ denotes the inner product of two vectors, thus it can be seen, the matrixes $M = [r_1^b, r_2^b, r_3^b]$ and $N = [r_1^a, r_2^a, r_3^a]$ both are unit orthogonal matrixes, and:

$$N = R_j M \tag{2}$$

Therefore:

$$R_j = N M^T \tag{3}$$

Using the solved rotation matrix $R_j$, the translation $T_j$ can be calculated as:

$$T_j = \frac{1}{3} \sum_{i=1}^{3} (L_i^a - R_j L_i^b) \tag{4}$$

### 3.3. LMedS motion estimation

Least Median of Squares (LMedS) is a effective optimization algorithm, which is also robust when outliers and colored noise exist. Theoretically speaking, with half matching data influenced by outliers, LMedS can also hold good and provide perfect estimate parameter.

Randomly select three pairs of matching features $(L_{n1}^b, L_{n1}^a)$, $(L_{n2}^b, L_{n2}^a)$, $(L_{n3}^b, L_{n3}^a)$, if they are collinear, select again until the features are non-collinear. By using orthogonalizing method described above, we can calculate the motion parameter $R_j$ and $T_j$, substituting them in equation (1), and the image plane error is described:

$$\Gamma_{(i,j)} = \left\| L_i^a - (R_j L_i^b + T_j) \right\|^2 \tag{5}$$

Compute its median:

$$\Gamma_j = med(\Gamma_{(i,j)}) \tag{6}$$

Repeating the process above $m$ times, we can have $m$ median value, calculate the least one:

$$\Gamma_{\min} = \min_{1 \le j \le m} (\Gamma_j) \tag{7}$$

Then the estimated motion parameter corresponding to the least median is the final output of robust motion estimation $\widetilde{R}_j$ and $\widetilde{T}_j$. To compensate for the deficiency of LMedS efficiency and to reject the outliers, the robust standard deviation estimate is given by[13]:

$$\sigma_s = 1.4826 \left( 1 + \frac{5}{n-3} \right) \sqrt{\Gamma_{\min}} \tag{8}$$

Based on $\sigma_s$, the matching pairs whose square errors are greater than $(2.5\sigma_s)^2$ are considered as outliers and should be eliminated from the set, therefore, the 3D position set of matching features changes into $\{(L_i^b, L_i^a) \mid i = 1, \cdots, \widetilde{n}\}$.

The sampling number $m$ of subsets depends on the probability $\xi$ of a sample free of outliers being selected and the percentage of outlier feature matching in the 3D position set of matching features, which can be written as:

$$m = \frac{\ln(1-\xi)}{\ln(1-(1-\varepsilon)^3)} \tag{9}$$

If set $\xi = 0.999$, $\varepsilon = 20\%$, it results in the selection of $m = 10$. Accordingly, the number $n$ of non-collinear features matched correctly in stereo image pair should satisfy $n(n-1)(n-2)/3! > m$.

For calculation of each sampling feature pair can be executed independently, this algorithm can be speedup by parallel computation.

### 3.4. Nonlinear motion estimation

Robust motion estimation can eliminate the outliers and provide an initial estimate of motion parameter, but it cannot eliminate feature localization errors perfectly, then cannot provide precise estimate of lunar motion. In condition of good initial estimate from robust motion estimation stage, nonlinear algorithm can solves for motion parameters directly and accurately. There exists many nonlinear motion estimation algorithms in the vision literature, for highly accurate motion parameter estimate, we use the Levenberg-Marquardt algorithm.

Levenberg-Marquardt nonlinear motion estimation attempts to minimize the image plane error function to estimate the motion parameter. Before we express the error function, we need to detail the motion parameters over which the minimization will take place. As described above, the motion between pre-move and post-move views is presented as a translation $T$ and a rotation $R$. To simplify the parameter estimation process and avoid singularity, we represent the rotation with a unit quaternion $q = [q_0, q_1, q_2, q_3]$. If we describe $T = [T_x, T_y, T_z]^T$, together the unit quaternion and translation $T$ , we comprise the parameter state vector $a$ [12]:

$$a = [q_0, q_1, q_2, q_3, T_x, T_y, T_z]^T \tag{10}$$

Based on the description above, estimate of motion parameter $a$ is equal to minimizing the error function for all the feature matching pairs in $\{(L_i^b, L_i^a) \mid i = 1, \cdots, \widetilde{n}\}$, the error function is:

$$\Gamma(a) = \sum_{i=1}^{\widetilde{n}} \left\| L_i^a - f_i(a) \right\|^2 \tag{11}$$

Where $f_i(a) = R(a)L_i^b + T$, which represents the projection of the 3D position in pre-move frame into the post-move frame with $a$ given.

To estimate the motion parameters, we minimize the error function (11) using the Levenberg-Marquardt algorithm for nonlinear minimization, this algorithm finds the minimum of $\Gamma(a)$ by iteratively solving[14]:

$$(A + \lambda I)\delta a = b \tag{12}$$

Where:

$$A = \sum_{i=1}^{\widetilde{n}} \left( \frac{\partial}{\partial a} f_i(a) \right)^T \left( \frac{\partial}{\partial a} f_i(a) \right) \tag{13}$$

$$b = -\sum_{i=1}^{\widetilde{n}} \left( \frac{\partial}{\partial a} f_i(a) \right)^T (L_i^a - f_i(a)) \tag{14}$$

And $\lambda$ is a scalar whose value is changed at each iteration. After each iteration, the estimate of $a$ is updated to $a + \delta a$ .When the value of $\lambda$ is smaller, the convergence performance of Levenberg- Marquardt is incline to that of Gauss-Newton algorithm, when $\lambda$ is larger, Levenberg-Marquardt algorithm is similar to steepest descent. With the dynamic adjust of $\lambda$ , Levenberg-Marquardt algorithm improves the performance of motion estimation.

In this paper, $\frac{\partial}{\partial a} f_i(a)$ is a $3 \times 7$ matrix, as:

$$\frac{\partial}{\partial a} f_i(a) = \left[ \frac{\partial}{\partial q} f_i(a) \quad \frac{\partial}{\partial T} f_i(a) \right] = \left[ \frac{\partial (R(q) L_i^b)}{\partial q} \quad I_{3\times3} \right] \tag{15}$$

The initial value $a_0$ Levenberg-Marquardt needed comes from the result $\widetilde{R}_j$ and $\widetilde{T}_j$ of robust motion estimate. The flow chart of Levenberg-Marquardt algorithm for motion estimation is shown in Figure 3

For the rotation parameter being represented by a unit quaternion, we enforce the constraint by setting $q \leftarrow (q + \delta q)/\|q + \delta q\|$ during the update of the parameter vector $a$ .

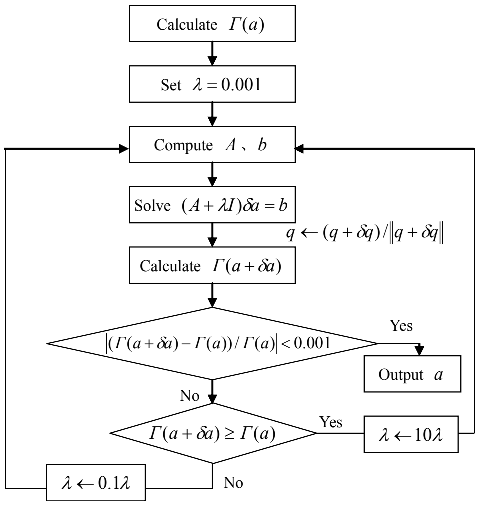
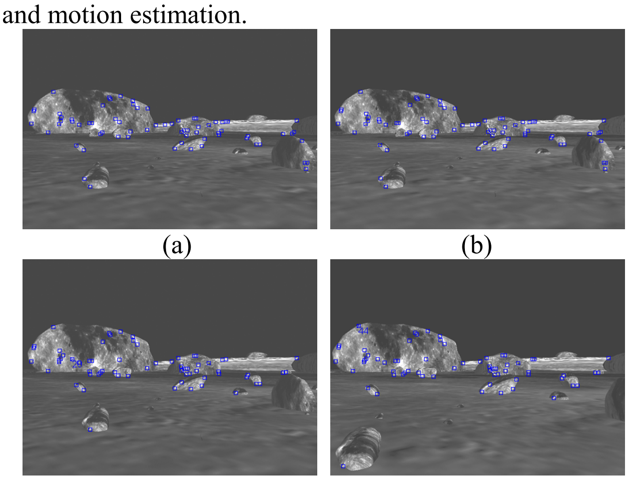


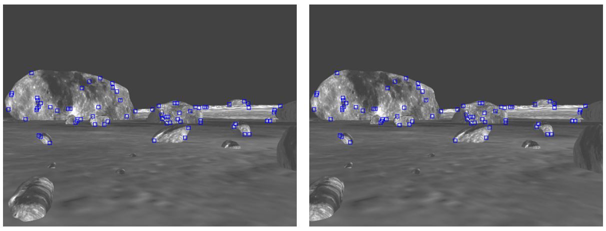

Figure 1. Flow chart of Levenberg-Marquardt estimate

## 4. Simulation

To verify our motion estimation algorithm presented in this paper on lunar environment, we build a simulator for lunar rover simulation of stereo vision motion estimation by using virtual reality modeling tool of openinventor. The simulator can simulate real motion of lunar rover on lunar uneven terrain, and provide a sequence stereo vision images, based on this, we testify the algorithm of feature detection, stereo matching, feature tracking and also the algorithm of two stage motion estimation, in our application, we select the left image of stereo vision as the primary one to execute feature tracking, see figure 1 2 3 for the result of processing and motion estimation.



(a) (b) (c) (d)



(e) (f)

Figure 2. (a) and (b) left and right image of stereo matching (1) respectively. (c) and (d) feature tracking in pre-move and post-move left image respectively. (e) and (f) left and right image of stereo matching (2) respectively.

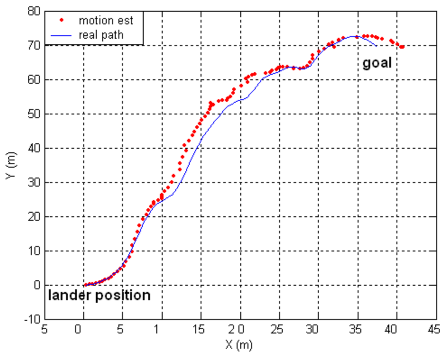



Figure 3. Position estimation result of lunar rover

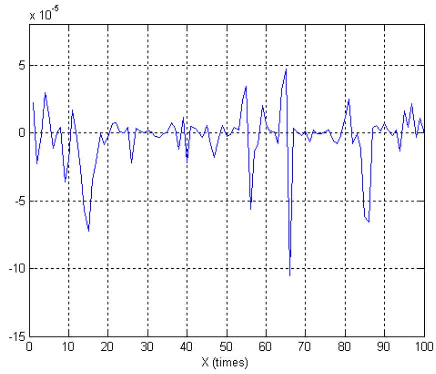



Figure 4. Standard deviation of attitude parameter $q_0$

## 5. Conclusions

We have presented a stereo vision based motion estimation technique for lunar rover autonomous navigation. Shown as the simulation, stereo vision based motion estimation can decrease errors in rover motion to a level and can satisfy the lunar rover for long range navigation. More research work would be carried out in the future is to analyze the lunar special environment, based on which we can improve the precision of image processing, and to consider the camera inner parameter fluctuation resulting from difference in temperature of moon surface, therefore, to pertinently improve the performance of lunar rover stereo vision based motion estimation. To imply this technique to lunar exploration, more work will be done about field test of the autonomous navigation scheme, and more detail consideration will be included.

## References

[1] Lacroix S, Mallet A, "Integration of concurrent localization algorithms for a planetary rover", In 6th International Symposium on Artificial Intelligence, Robotics and Automation in Space, Canada, pp. 380-387, 2001.

[2] Yasutaka Fuke, Eric Krotkov, "Dead Reckoning for a Lunar Rover on Uneven Terrain", Proceedings of the 1996 IEEE International Conference on Robotics and Automation, pp. 411-416, April. 1996.

[3] Nister, D. Naroditsky, O. Bergen, J, "Visual odometry", Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 652-659, July. 2004.

[4] Maarten Vergauwen, Marc Pollefeys, Luc J. Van Gool, "A stereo vision system for support of planetary surface exploration" Proceedings of the Second International Workshop on Computer Vision Systems, pp. 298-312, 2001.

[5] Yang Cheng and Mark Maimone and Larry Matthies, "Visual Odometry on the Mars Exploration Rovers", 2005 IEEE Conference on Systems, Man and Cybernetics, pp. 903-910, Oct. 2005.

[6] D.M. Helmick, Y. Cheng, S.I. Roumeliotis, D. Clouse, L. Matthies, "Path following using visual odometry for a mars rover in high-slip environments", In IEEE Aerospace Conference, pp. 772-789, March. 2004.

[7] Landzettel, K. Steinmetz, B.M. Brunner, B. Arbter, K.; Pollefeys, M.; Vergauwen, M.; Moreas, R.; Xu, V.; Steinicke, L.; Fontaine, B, "A micro-rover navigation and control system for autonomous planetary exploration", The International Journal of the Robotics Society of Japan: Advanced Robotics, Special Issue Service Robots in Space, pp. 285-314, 2005.

[8] Clark F. Olson, Larry H. Matthies, Marcel Schoppers, and Mark W. Maimone, "Rover Navigation Using Stereo Ego-motion", Robotics and Autonomous Systems, pp. 215-229, 2003.

[9] Clark F. Olson, Larry H. Matthies, Marcel Schoppers, Mark W. Maimone, "Stereo Ego-motion Improvements for Robust Rover Navigation", In Proceedings of the IEEE International Conference on Robotics and Automation, pp. 1099-1104, 2001.

[10] Krotkov E, Hebert M, Buffa M, F. Cozman, Robert L, "Stereo driving and position estimation for autonomous planetary rovers", International Advanced Robotics Programme Workshop on Robotics in Space, Canada: Montreal, pp. 263-268, 1994.

[11] Z. Zhang, R. Deriche, O. Faugeras, and Q. Luong, "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry", Artificial Intelligence, Vol 78, pp 87-119, 1995.

[12] Andrew E. Johnson and Larry H. Matthies, "Precise Image-Based Motion Estimation for Autonomous Small Body Exploration", 5th International Symposium on Artificial Intelligence, Robotics and Automation in Space, pp. 627-634, June. 1999.

[13] Ma Song de, Zhang Zheng you, Computer Vision, Science Press. Beijing, 1998.

[14] Zhang Guang cheng, Huang Shi bao, Hou Ze hua, Optimal calculation methods, Chengdu Science and Technology University Press, Chengdu, 1989.