2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Map Matching Scheme for Position Estimation of Planetary Explorer in Natural Terrain

Takashi Kubota, Klaus G.Moesl, Ichiro Nakatani, *Member, IEEE*

Abstract— **This paper describes a position identification method to precisely estimate the relative movement onboard of an autonomous explorer in such natural terrain as Moon or Mars. This paper proposes a new map matching scheme that registers the visual information of consecutive robot positions and expresses the robot displacement in a common coordinate system. Local elevation maps of the environment are therefore retrieved from the 3D perception of a stereo camera. The proposed scheme consists of three main steps: feature point extraction from map data, matching of features points using triangle configurations and a voting procedure, and direct calculation of the robot displacement in six degrees of freedom using quaternions. The effectiveness of the proposed method is confirmed by computer simulations with synthetic terrain and experiments with stereo perception.**

I. Introduction

Intelligent mobile explorer to perform scientific observation and explore the surface of distant planets has attracted a lot of attention of the space exploration community for a long time. Unmanned explorers have been deployed on the moon and on Mars. Various kinds of missions for exploring other planets, moons, and small bodies including asteroids and comets by rovers have been proposed or are currently in operation. In any case, rovers will act as a mini-field geologist by providing access to samples for in-situ analyses, performing experiments, and exploring the extraterrestrial body[1][2].

Envisioned missions have to limit communication contacts between rover and ground control, because of constraints on communication opportunities, resources, and cost. In between these communication windows, the rover must operate autonomously. It fully has to rely on its position estimation and navigation system which will be realized in the form of an interconnection of dead reckoning devices and a vision system to keep all errors low. The harsh and challenging terrain over which the rover will traverse tends to seriously degrade the dead reckoning state estimate. Obstacle climbing, numerous maneuvers including rover turns, and wheel slippage in soft soil act on the degradation.

A lander relay for the rover's position estimation is only feasible for operations around the landing side. Traverses exceeding several hundreds of meters and kilometers will have to rely on relative position estimation based on integrated internal measurements including odometry as well as on external movement estimate by vision[3]. The latter has the advantage that its error does not grow over time. Stereo cameras and laser rangefinder are the two main types of sensors used for rovers to perceive the surrounding environment and thus to localize and navigate the rover.

The authors propose a new algorithm to precisely calculate the relative movement onboard of the rover which exploits stereo vision but generally is independent from the type of 3D perception.

Key idea for having reliable position information is to determine the change in position and attitude of the rover between consecutive image pairs, acquired as the rover moves by using a "tracking" of some characteristic feature points in both images. At each moving step, the current position and orientation of the rover is estimated by matching the current map with the previous one. Maps are represented by Digital Elevation Maps (DEMs) of the environment and feature points are defined as local maxima on these DEMs.

A new method for feature point extraction searches the DEMs on different sub-scales for local maxima. It is a Multi-Scale Peak Detection (MSPD) method which uses a parent-child relationship to establish a connection between the sub-scales and detect prominent peaks. Performance and reliability of the MSPD method are evaluated with a Seed Wandering Method (SWM).

The main registration algorithm organizes the extracted feature points in sets of bare triangles, merely characterized through their 3D coordinates on each elevation map. If two triangles are considered congruent, votes will be given for a successful matching on the point-to-point level of the triangles.

After the verification of the prospective matching points using the number of votes, the movement relative to the preceding position can be retrieved by trigonometric and matrix calculus. The authors could show successful registration of DEMs with an overlap of only 25% if there were prominent, common feature points on both of them.

T.Kubota is with Japan Aerospace Exploration Agency, 3-1-1, Yoshinodai, Sagamihara, 2298510 JAPAN (phone: 81-42-759-8305; fax: 81-42-759-8305; e-mail: kubota@ nnl.isas.jaxa.jp).

K.G.Mosl was with Japan Aerospace Exploration Agency, 3-1-1, Yoshinodai, Sagamihara, 2298510 JAPAN.

I.Nakatani is with Japan Aerospace Exploration Agency, 3-1-1, Yoshinodai, Sagamihara, 2298510 JAPAN.

II. Related Works

Some research groups have focused on the terrain mapping and terrain matching issue. The iterative closest point (ICP) algorithm[4] including its modifications and enhancements is the most popular algorithm. It provides shape registration for geometric data represented in point sets, line segment sets, curves, surfaces, and triangle sets. This algorithm always converges to a local minimum of the error function which might not be the correct global minimum. There is the problem to define an appropriate set of initial states that guarantees a correct global minimum.

This also includes the "Trimmed ICP" algorithm as proposed by Chetverikov et al.[5]. The ICP-CSM algorithm is an enhancement of the ICP algorithm. Using virtual movements of a matchpoint, it verifies correspondences between the points of two data [6]. Zhang gives an extensive survey in iterative methods[7] and exploits the same idea as Besl and McKay[4]. It uses the quaternion technique to compute the motion of the sets of reasonably paired 3D points [4][7][8].

Krotkov discusses a method that extracts features from a given digital elevation map, likely to be identifiable in rover-acquired images. Both, stereo and laser rangefinder, were tested with the stereo process being able to execute more rapidly[9].

A method of matching unequally spaced height maps is described by Gennery[10]. It is a registration method between images from a camera orbiting the celestial body and a ground-based stereo vision or laser rangefinder system. Thompson et al. performed vision-based localization by extracting features from maps and matching these features with features of the current two-dimensional view of the rover. Configurations of views play a decisive role within the matching process and a prominence value is introduced, representing specific combinations of feature properties[11]. Kweon and Kanade combined feature matching, iconic matching, and inertial navigation data with a course-to-fine pixel-based gradient search to register elevation maps.

A prior existing DEM was needed for estimating the vehicle's position in the DEM and reducing the error accumulation during the rover's motion[12]. Kamgar-Parsi et al. used feature-based matching of contours of constant range to register ocean floor images with three degrees of freedom[13]. Sutherland as well as Betke and Gurvits address the localization problem using distinguishable landmarks in the environment. They show that for a given error in angle measurement, the localization error varies depending on the configuration of the landmarks[14][15].

The algorithms for using stereo cameras on the motion estimate of the prevailing Mars Exploration Rovers were originally developed by Matthies. Following his work, Olson et al. did some minor variations and modifications to improve the robustness and accuracy of the method. The key idea is to determine the change in position and attitude of the rover between consecutive image pairs acquired as the rover moves. Maximum likelihood estimation tracks the feature points between the images[16][17][18].

As mentioned beforehand, most of the iterative procedures request a pre-registration of the two data sets. This is not necessary for the herein proposed algorithm. Furthermore, iterative procedures might converge towards an incorrect local maximum and great care needs to be taken on this regard which usually involves some additional procedures with demerits on the original idea of the algorithms [4][5][6][7][12][16][17][18]. Landmark-based registration methods[14][15] require an absolute correct identification and order knowledge of the reference landmarks. This often asks for additional information from texture and vegetation which limits the operational independence of the method. If there are exceptional configurations of landmarks including straight line asymmetric or non-linear ones, ambiguity within the landmark order will be the consequence and further measures have to be taken.

III. Map Matching

This paper proposes a new map matching scheme that can register the visual information of consecutive robot positions and expresses the robot displacement in a common coordinate system. Local elevation maps are therefore retrieved from the 3D perception of a stereo camera. The proposed scheme consists of three main steps: feature point extraction from map data, matching of features points using triangle configurations and a voting procedure, and direct calculation of the robot displacement as shown in Fig.1.

A. Triangle Matching

After loading two DEM sets and extracting the feature points from both of them, triangles are built from each set of feature points. Triangles are a good trade-off between geometric complexity and data aggregation. They have a distinct determination in 3D regarding their vertices and data processing is much more effective than using points or mere lines for matching.

Figure 2 represents two triangles that might be candidates for a successful matching. Triangle *A* comprises of the three feature points *a1*, *a2*, and *a3*. Triangle *B* is made up of feature points *b1*, *b2*, and *b3*, respectively. These two triplets of feature points are identical to some of the top feature points of both DEM sets.

Candidate matches are identified on the similarity of the triangle side lengths. The triangles need to be congruent. Comparing all triangles of set *A* with all triangles of set *B* in respect of congruence may be a down to earth attempt for finding correct matches but is a reliable means to achieve the aim. Here, we propose to form target configurations for the triangle matching process. They help in cutting down the combinatorics and the calculation time.

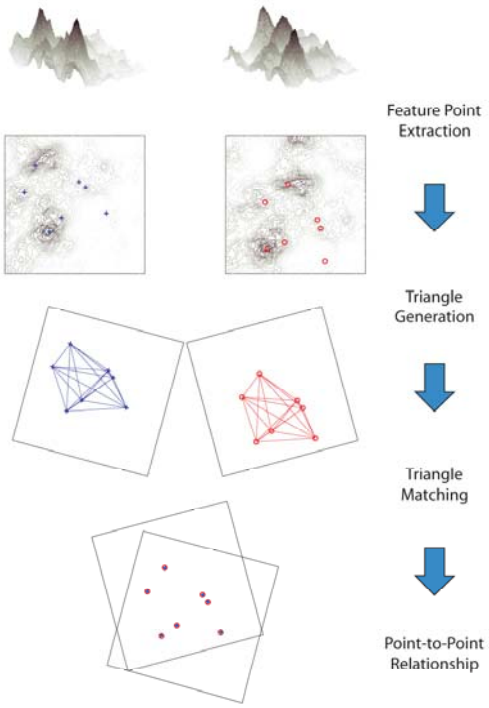

Fig.1. Map Matching Scheme for Position Identification

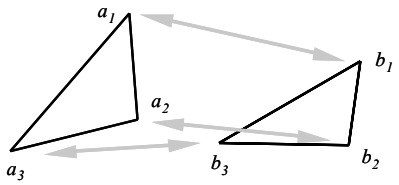

Fig.2. Generation of Triangles of DEM Set *A* and DEM Set *B*

The establishment of target configurations includes a sorting of the built triangles according to one characteristic as seen in Fig.3. The length of the shortest or the longest side, or the triangle surface area might be appropriate. Here, the side length of the shortest triangle side is chosen as sorting criteria. This causes benefit in two aspects: On the one hand, the calculation cost is reduced. On the other hand, a helpful filter may be installed for very small, hardly distinguishable triangles or triangles with a very short side. They have a higher probability of mismatching than bigger triangles due to the limited DEM resolution.

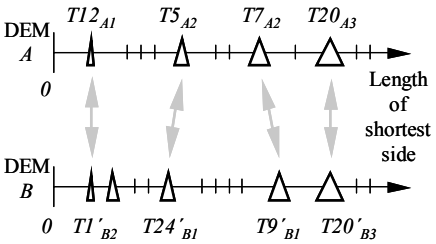

Fig.3. Correspondences of Sorted Triangles

Figure 3 shows the sorted triangles in ascending order of the shortest side. The variables *Txx* and *Tyy'* indicate triangles as built from the list of feature points of set *A* and set *B*, respectively. Their footnotes flag the shortest side length of each triangle that is responsible for the triangle's ranking. Thus, triangle *T12A1* for instance is rated as having the overall shortest side among all the triangle sides of set *A*. On the basis of the preceding sorting, the search scope (among set *B* triangles) for matches with an *A*-set triangle can be limited to the close-up range of the shortest side length of the *A*-set triangle. By doing so, candidate matches of e.g. *T12A1* will be found around *T1'B2*, whereas an attempted match with triangle *T20'B3* does not make any sense. The algorithm allows ambiguous correspondences between several triangles of set *A* and set *B* as this only is the first step of finding the real corresponding matches.

Of course, all three triangle sides of a candidate match need to be congruent. The judging process therefore is structured in three steps: shortest, middle, and longest triangle sides. Eq.(1) gives an examples of judging relations for congruence (cf. Fig.2).

$$\begin{aligned} \left(\overline{a_1a_2} - \overline{b_1b_2}\right)^2 &\le \varepsilon_{12} \\ \left(\overline{a_2a_3} - \overline{b_2b_3}\right)^2 &\le \varepsilon_{23} \\ \left(\overline{a_3a_1} - \overline{b_3b_1}\right)^2 &\le \varepsilon_{31} \end{aligned} \qquad (1)$$

Recalling the example of Fig.2, the Euclidean distance between two points forming a triangle side length calculates according to Eq.(2).

$$\overline{a_ia_j} = \sqrt{\left(x_{A,i} - x_{A,j}\right)^2 + \left(y_{A,i} - y_{A,j}\right)^2 + \left(z_{A,i} - z_{A,j}\right)^2} \qquad (2)$$

The variables ε_{12}, ε_{23}, and ε_{31} are threshold values that allow the consent of matching between an *A*-side and a *B*-side triangle. Ideally, all three threshold values should be equal zero as postulated in Eq.(3) and Eq.(4). However, the strict demand of Eq.(4) is out of reach due to restrictions of the DEM resolution and sensor noise in practice.

$$\varepsilon_{12} = \varepsilon_{23} = \varepsilon_{31} = \varepsilon = \text{const.} \qquad (3)$$

$$\varepsilon = 0 \qquad (4)$$

During our studies all three threshold values have been kept uniform according to Eq.(3).The value ε needs to be heuristically adjusted for best matches. It is dependant on terrain alternation and roughness as well as the kind and direction of motion. Generally it does not exert a very strong influence on the matching results as long as it is reasonably chosen and there are enough common feature points for matching. Mere translational displacements between the two DEMs can be managed with very low ε-values, whereas rotational displacements require a little bit wider tolerances

B. Voting for True Match

As mentioned before there may be some ambiguous and repeated triangle correspondences approved by the matching condition according to Eq.(1), if they have similar side lengths. The arrows in Fig.3 consequently only mark some "candidate matches". Therefore, a verification of matching triangles / feature points using a voting process forms the second part of the map matching algorithm. If all three side lengths fulfill the matching criterion, two triangles are supposed to be congruent and votes will be given for the matching vertices. This is done as illustrated by Eq.(5) and involves some basic combinatorics.

$$\begin{array}{llll} a_0 \leftrightarrow b_1, & a_2 \leftrightarrow b_2 \\ a_0 \leftrightarrow b_2, & a_2 \leftrightarrow b_1 \\ a_2 \leftrightarrow b_2, & a_3 \leftrightarrow b_3 \\ a_2 \leftrightarrow b_3, & a_3 \leftrightarrow b_2 \\ a_3 \leftrightarrow b_3, & a_0 \leftrightarrow b_1 \\ a_3 \leftrightarrow b_1, & a_0 \leftrightarrow b_3 \end{array} \qquad (5)$$

There are four possibilities of point-to-point correspondences for a pair of matching triangle side lengths, but only two of them are true point matches. The voting process gives votes for all four combinations, though. By doing this procedure for all three triangle sides, votes are given twice to the real matching points and only once to "error matches". Furthermore, giving votes for all the vertices combinatorics of all candidate matches between triangles of set A and set B, the real matching feature points distinguish themselves from mismatches. They have a repeatedly positive occurrence in several matching triangles.

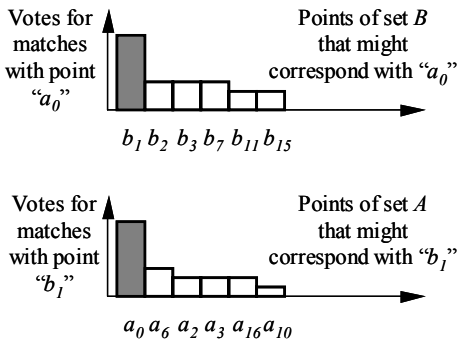

Fig.4. Voting for Matching Triangle Vertices

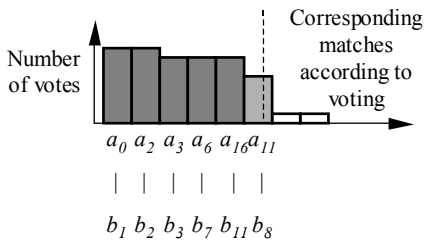

Fig.5. Corresponding Feature Points of DEM A and DEM B

This procedure allows two points of view: First, a separate observation of set A and set B, as visualized in Fig.4. Actually, this point of view asks for giving the four, above mentioned votes twice because of the separated accounting of votes in two sets of feature points. Second, a combined observation of all features points correlations. This may be done in the form of a matrix where the horizontal dimension e.g. refers to DEM A and the vertical dimension to DEM B feature points, which is a very compact representation of the same process as illustrated in Fig.4 and may furthermore be a convenient and efficient model for code implementation. After all candidate matches got votes for possible feature point correspondences, results of real matching feature points turn out to be similar to Fig.4 and can finally be summarized according to Fig.5. Thus, candidate matches in relation to the triangle matching of Eq.(1) are verified and mismatches are separated.

Calculation of the translational and rotational registration value (six degrees of freedom) is performed using quaternions and triplets of point-to-point correspondences of Fig.5. If three degrees of freedom are only requested within the registration task, doublets of point-to-point correspondences will be sufficient for an equivalent calculation. Feature point (a_{11}, b_8) is a special case that may be filtered out of the final number of matches as it got observably less votes than the other positive matches. Indeed, it is a traceable, positive match as unveiled by recalculation but it might also indicate a mismatch if the operator-defined matching parameter ε allows wide tolerances.

IV. Feature Point Extraction

Feature-based methods for localization hold the potential for avoiding many problems inherent with other approaches. Features can be extracted independently from sensor data and then be matched symbolically. Regarding the process of matching, no accurate rendition of the sensed data is needed here. Feature points with the best characteristic traits get the highest ranking as they are most likely to match. Two algorithms for feature point extraction are proposed and both are adapted to surface representation by DEM. A reference or ground plane is discretized into a regular spaced grid and elevation is the vertical distance above or below this reference. Finally, the availability of the best feature points may be taken for granted within the actual process of triangle matching.

A. Multi-Scale Peak Detection (MSPD)

Each elevation map is evenly segmented into patches of several scales, which help in evaluating the feature points of the map. "Scale" denotes the level of resolution under which an elevation map of fixed size is observed.

The MSPD feature point extraction starts with considering the whole DEM as one single element. Each iteration / scale concretizes the resolution of the map by factor 2 till the

highest scale (highest resolution of original, single DEM cells) is reached. Multi-scale cells request an elevation averaging in relation to the DEM cells belonging to them. The actual verification of peaks is done by jumping from scale cell to scale cell and checking whether it is a local maximum to its neighbors. If so, the current scale cell will be recorded on a general list of all multi-scale peaks.

Figure 6 develops the parent-child relationship between peaks of multiple scales. The parameter G indicates the generations of scales one peak possesses. Several peaks of different scales get detected during the verification process of local maxima. These might be big flat hills without real summit, smaller pointed rocks, or any other possible appearance among natural terrain features. Preferable feature points for matching should have high number generations in order to be distinctive and prominent.

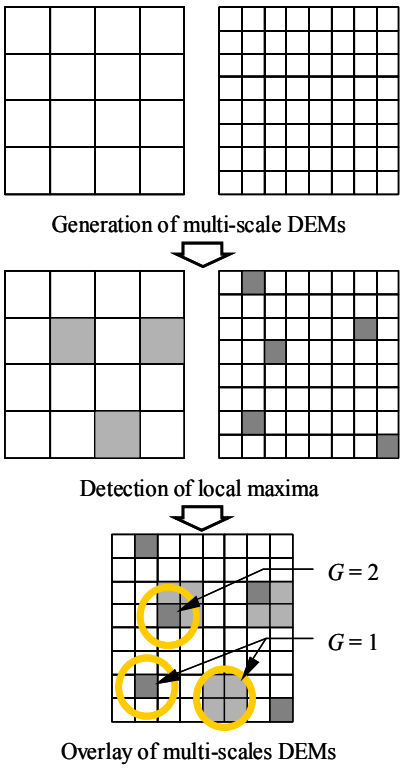

Fig.6. Operational Sequence of MSPD Method

B. *Seed Wandering Method (SWM)*

Seed wandering is a universal procedure for the detection of local maxima. The utilization of seeds accommodates the discrete character of the DEM grid as each seed can easily and directly be assigned to one specific grid cell. Furthermore, a seed is able to wander on the DEM, from one grid cell to the next in correlation to a function routine. The SWM algorithm operates self-contained and does not need any special preparation of the DEM data. It is possible to define the number of seeds to be spread over the DEM. After the seeds got disseminated, the first one starts to wander. By doing so, it compares the elevation values of its neighbors and moves onto the grid cell with the highest elevation. If no higher neighbor can be found, the seed reached its local maximum where it stays. Finally, the next seed starts to wander until it arrives at its local maximum as well. Central idea is that the number of seeds associated to one grid cell is a quality measure for feature points. The more seeds a local maximum gets, the bigger its catchment area was. Thus, it typifies a characteristic feature point, an aspirant for the matching procedure.

V. Experimetal Results

The matching algorithm runs within an amount of time which is a function of the number of extracted feature points. The more points there are extracted for matching, the longer the overall process takes. Using seed wandering (SWM) for the extraction, the operator is free to define any adequate number of desired feature points that will be processed by the registration. If the MSPD method is used, the number of feature points is predefined as a parameter of the MSPD scales, the height discretization of the DEM, and the terrain itself. Thus, it could happen that there is no comfortable number of features regarding a short calculation time using the multi-scale method.

Two successive DEMs should not be too far from each other that there are enough common feature points on both of them. There is a smooth transition from excellent matching to complete mismatching in relation to terrain characteristics, displacement / overlap of the DEMs, stereo and DEM quality, usage of various filters within the whole process, and the definition of the matching parameter ε itself. Removing statistical outliners from stereo perception is considered a pre-processing step, has been implemented as such, and will therefore not be addressed herein.

The matching parameter ε is determined by experiments (cp. Eq.(1) and Eq.(3)). We investigated its influence on matching / mismatching in the range of [0.01 10] and [0.1 10] [pixel] for synthetic and real terrain, respectively. Generally, the usage of all investigated values ε was uncritical as the algorithm runs stable and self-contained with its two steps of triangle matching and verification by voting. However, there might be some optimum sub-ranges of ε in relation to the rover movements (pure translation or combined translation and rotation) and the characteristics of the current terrain.

A. *Matching of Synthetic Terrain*

Various types of synthetic DEM were generated for the verification of the matching method. The raw data is based on 256 x 256[pixel] grids with 0.05 [m] spacing. The two DEMs for matching are derived from this global data sets and laid out on smaller 128 x 128 [pixel] grids. All terrain types with translations up to 50[%] in each grid direction (x, y) and

rotations up to 40[deg] around the surface normal (z) were tested. The algorithm runs stable as long as there are enough common feature points. Figure 7 and Fig.8 show one of cases of matching with rotation.

B. *Matching of Terrain Data from Stereo Perception*

Currently, the experimental stereo system is based on a pre-calibrated, two-lens setup generating 3D information in the form of point clouds. Elevation maps are built by projecting those points onto the horizontal DEM grid of the rover environment. Occlusions are interpolated and the related ridge lines are smoothed for a sound performance of the feature extraction algorithms. DEM resolution is set to 0.04 [m] and the overall grid size is 100 x 100 [pixel] here.

The effectiveness of the proposed new matching procedures on the real 3D data of the stereo system could be shown. Translational displacements can be registered well with a lot of common feature points as seen in Fig.9.

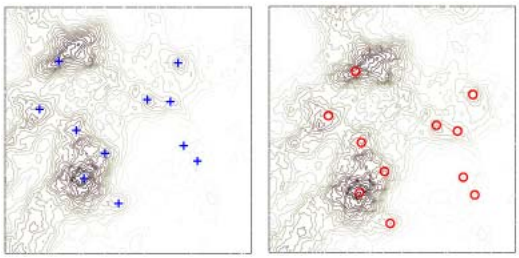

Fig.7. Common Feature Points with Congruent Triangles

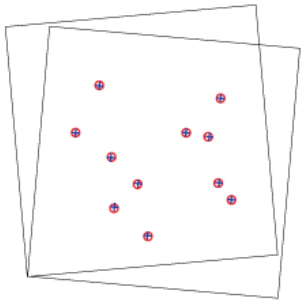

Fig.8. Matched Feature Points with a DEM. (Rotation of 10 [deg])

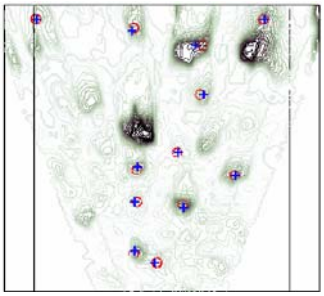

Fig.9. Matched DEMs from Stereo Perception with 0.4[m] Displacement

VI. Conclusion

This paper has presented a new algorithm that exploits stereo vision and elevation maps for an extern sensor feedback on motion. Terrain maps from views of successive locations are matched by the proposed registration algorithm, which may be incorporated into the central navigation procedure. The experimental results showed very good performance on synthetic terrain data and could also perform well on elevation maps from real world stereo.

References

[1] T.Kubota, Y.Kunii, Y.Kuroda, T.Yoshimitsu, T.Okada, Mn.Kato, " Lunar Robotics Exploration by Cooperation with Lander and Micro Rovers," 6th IAA Int. Conf. on Low-Cost Planetary Missions, pp.189-194, 2005.

[2] The Rover Team, "The Pathfinder Microrover," Journal of Geophysical Research, Vol.102, No.E2, pp.3989-4001, 1997.

[3] M.Maurette, "Mars Rover Autonomous Navigation," Autonomous Robots, Vol.14, No.2-3, pp.199-208, 2003.

[4] P.J.Besl and N.D.McKay, "A Method for Registration of 3-D Shapes," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.14, No.2, pp.239-256, 1992.

[5] D.Chetverikov, D.Svirko, D.Stepanov, and P.Krsek, "The Trimmed Iterative Closest Point Algorithm," Proc. of the 16th Int. Conf. on Pattern Recognition, Vol.3, pp.545-548, 2002.

[6] E.Guest, E.Berry, R.A.Baldock, et al., "Robust Point Correspondence Applied to Two-and Three-Dimensional Image Registration," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.23, No.2, pp.165-179, 2001.

[7] Z.Zhang, Iterative Point Matching For Registration of Free-Form Curves and Surfaces, Institut National de Recherche en Informatique et en Automatique, 1992.

[8] Z.Zhang, "A stereovision system for a planetary rover: calibration, correlation, registration, and fusion," Machine Vision and Application, Vol.10, No.1, pp.27-34, 1997.

[9] E.Krotkov, Position Estimation and Autonomous Travel by Mobile Robots in Natural Terrain, Kent Forum Book, 1997.

[10] D.B.Gennery, "Visual Terrain Matching for a Mars Rover," Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.483-491, 1989.

[11] W.B.Thompson, T.C.Henderson, T.L.Colvin, et al., Vision-Based Localization, DARPA98, pp.491-498, 1998.

[12] I.S.Kweon and T.Kanade, "High-Resolution Terrain Map from Multiple Sensor Data," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.14, No.2, pp.278-292, 1992.

[13] B.Kamgar-Parsi, J.L.Jones, and A.Rosenfeld "Registration of Multiple Overlapping Range Images: Scenes Without Distinctive Features," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.13, No.9, pp.857-871, 1991.

[14] K.T.Sutherland, "Ordering Landmarks in a View," ARPA94, pp.(II)1101-1105, 1994.

[15] M. Betke and K. Gurvits, "Mobile Robot Localization Using Landmarks," IEEE Transactions on Robotics and Automation, Vol.13, No. 2, pp.251-263, 1997.

[16] C.F.Olson,"Mobile Robot Self-Localization by Iconic Matching of Range Maps," Proc. of the 8th Int. Conf. on Advanced Robotics, pp.447-452, 1997.

[17] C.F.Olson, L.H.Matthias, M.Schoppers, M.W. Maimone,"Rover Navigation Using Stereo Ego-Motion,"Robotics and Autonomous Systems, Vol.43, pp.215-229, 2003.

[18] R.Li, K.Di, L.H.Matthies et al.,"Rover Localization and Landing Site Mapping Technology for the 2003 Mars Exploration Rover Mission,"Journal of Photogrammetric Engineering and Remote Sensing, Vol.70, No.1, 2003.