Calibration, Terrain Reconstruction and Path Planning for a Planetary Exploration System

Maarten Vergauwen[1], Marc Pollefeys[1], Ronny Moreas[2], Fuyi Xu[2],
Gianfranco Visentin[3], Luc Van Gool[1] and Hendrik Van Brussel[2].

1. ESAT-PSI, K.U.Leuven, Kasteelpark Arenberg 10, B-3001 Heverlee, Belgium
2. PMA, K.U.Leuven, Celestijnenlaan 300 B, B-3001 Heverlee, Belgium
3. ESA-ESTEC, Keplerlaan 1, Noordwijk, The Netherlands

Keywords planetary exploration, calibration, 3D reconstruction, path planning.

Abstract

In this paper, three important preparation tasks of a planetary exploration system are described, namely calibration, terrain reconstruction and path planning. The calibration is retrieved from the images of the planetary terrain (the same images from which the 3D measurements of the terrain are obtained). Once the system has been calibrated, the same images can be used to estimate a digital elevation map of the environment around the lander. These images are first processed pairwise using a stereo algorithm yielding sub-pixel disparity maps. In order to perform the path planning a digital elevation map is required. An algorithm is described which can generate a regular digital elevation map at any desired resolution (interpolating if necessary) and easily takes occlusions into account. The results of the calibration and reconstruction are then used by the path planning module which is capable of computing the optimal path between points of interest of the terrain.

1 Introduction

The work described in this paper has been performed in the scope of the ROBUST[1] project of the European Space Agency. In this project an end-to-end control system has been developed for a planetary exploration system. In order to reduce costs and improve efficiency the system was designed in such a way that as much preparation as possible is done on the ground control system. This preparation includes calibration of the stereo camera system, reconstruction of the planetary terrain and computation of an ideal path for the rover to traverse.

[1]The official name for this project is PSPE. The consortium consists of the Belgian companies SAS and OptiDrive, the German companies DLR and vH&S and the electrical and mechanical department of the Belgian university K.U.Leuven.

1.1 System Overview

The system that has been designed for the ROBUST project consists of three important parts.

- A Planetary Rover.
  The **Nanokhod** [8], a small and robust rover with a very high payload–mass versus total–mass ratio was selected. Due to its design the Nanokhod can traverse a wide variety of surfaces. It can hold up to four instruments in its payload cab and is operated from the lander.

- A Planetary Lander.
  The lander holds an on–board Lander Control System and an Imaging Head (IH). The latter is a stereo head mounted on a pan–tilt unit which is approximately 1.5 meter high. The two cameras of the stereo head are space approved 1024x1024 CCD cameras. The stereo head has a baseline of 0.5 meter. The images, taken by the digital scanners on the stereo head, are used for calibration, reconstruction of the terrain and localization of the rover. The latter makes use of four Light Emitting Diodes on the rover payload cab.

- An On Ground Control System.
  This system, consisting of one or more workstations, performs preparations, simulation, verification and playback tasks. The calibration, terrain reconstruction and path planning modules that are described in this paper are executed on this On Ground Control System.

1.2 Typical utilization scenario

A typical utilization scenario will deploy the Imaging Head as soon as possible after the landing of the planetary lander. Because of the strain on the parts during launch and landing, the Imaging Head needs to be recalibrated. To accomplish this, it takes images of the terrain which are sent to earth where the calibration is performed using these images. From the same images a 3D reconstruction of the terrain is then computed. Based on this reconstruction, scientists can indicate interesting points for the rover to visit. A path planning algorithm will then compute the optimal path. This path is simulated and verified and then uploaded to the lander where it is executed by the Nanokhod.

2 Calibration

During launch and landing, the lander and its contents are subject to extreme forces. The mechanical properties of the Imaging Head are likely to have been affected by mechanical and thermal effects. For high accuracy equipment, such as the Imaging Head, a small change in these mechanical properties results in large degradation of the results, unless the new properties can be estimated. The cameras themselves are built so that the intrinsic parameters during the mission can be assumed identical to the parameters obtained through calibration on ground.

2.1 Using markers ?

Traditional calibration algorithms rely on known calibration objects with well-defined optical characteristics in the scene. If cameras take images of these artificial objects, the pose of the cameras can be computed, yielding the extrinsic (mechanical) calibration of the cameras [9].

There are two reasons why this scheme is not suited in our case where the Imaging Head is deployed on a distant planet. First there is the problem of where to place the calibration objects. It is of course impossible to add objects to the terrain, so one has to think of placing calibration markers on the lander itself. A typical lander consist of a cocoon which opens after landing, comparable to an opening flower. The markers could be applied to the opening "petals". However, one is never sure of the exact position of these petals which makes the markers much harder to use. Even if one did dispose of accurate markers on the lander, a second problem arises. To maximize robustness, no zooming and focusing system was added to the cameras. Since the accuracy of the stereo matching decreases with the square of the distance, the cameras are focussed on infinity to gain as much accuracy in the far regions as possible. As a consequence, the images of near regions are blurred. Since the markers would be on the lander, images of the markers would always be blurred, reducing the accuracy of the calibration up to the point where the markers are useless.

It is clear that standard calibration algorithms can not be used in our system. A new strategy had to be developed that only uses images of the terrain to calibrate the Imaging Head.

2.2 Strategy

The calibration procedure that was implemented for the ROBUST project is able to calibrate the Imaging Head using images of the terrain only. This means that the images which are sent down from the planet to earth to reconstruct the terrain, can also be used for calibrating the Imaging Head. Therefore, the terrain based calibration causes no overhead on transmission.

The calibration of the extrinsic (mechanical) properties of the Imaging Head is split into two parts which are executed consecutively. First the relative transformation between the two cameras is computed. This is explained in Section 2.3. Once this relative calibration is computed, a procedure can be performed which computes the relative transformations between the cameras and the lander. This boils down to computing the pan and tilt axes of the pan-tilt unit. Section 2.4 deals with this problem.

2.3 Relative calibration

The relative transformation between the two cameras of the Imaging Head can be computed from images of the terrain only. The algorithm to do this uses the concept of the essential matrix. This matrix represents the epipolar geometry between two views, including the internal parameters of the cameras as extra information. We make use of the fact that the relative transformation between the cameras does not change when the the different segments of the terrain are recorded, which allows for different measurements of the epipolar geometry to be combined to yield one accurate solution.

If the essential matrix between the two views is computed, the relative transformation (position and orientation) between the two cameras can be calculated up to the baseline (i.e. the distance between the two cameras).

2.3.1 Computing epipolar geometry

The first step in obtaining the relative calibration is the computation of the epipolar geometry of the stereo head. The epipolar geometry constraint limits the search for the correspondence of a point in one image to points on a line in the second image. Figure 1 makes this clear.

If one wants to find back the epipolar geometry between two images automatically, a filter, called the "Harris Corner Detector" [3] is applied to the images first.

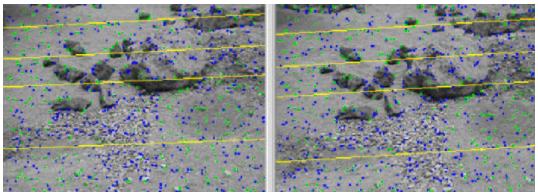

Figure 1: Epipolar geometry of an image pair

The result consists of points or *corners* in the images determining where the image intensity changes significantly in two orthogonal directions. Next, the corners are matched automatically between pairs of images using cross correlation. This process yields a set of possible matches which is typically contaminated with an important number of wrong matches or outliers. Therefore a robust matching scheme, called RANSAC[2], is used to compute and update epipolar geometry and matches iteratively.

In the case of the ROBUST Imaging Head the images of the different segments of the terrain can be combined to compute the epipolar geometry much more robustly because the relative transformation between the cameras and therefore the epipolar geometry does not change.

It is even the case that a specific degenerate case for the computation of the epipolar geometry is solved by the combination scheme we described. Computing the epipolar geometry of a pair of images of a **planar** scene is impossible from correspondences only. If the planetary terrain is planar or close to it, computing the epipolar geometry for one pair of images becomes an ill-posed problem. By combining correspondences from different segments, this problem is solved.

2.3.2 Computing relative transformation

Once the epipolar geometry is computed in the form of the fundamental matrix F, the relative transformation between the two cameras of the Imaging Head can be calculated. First the essential matrix is constructed. This is easily done since

$$E = K^T F K$$

with K the 3x3 matrix with the intrinsic calibration of the cameras. To derive the relative translation and rotation from the essential matrix, we refer to the work of Maybank et al. [5]. There, it is explained, based on Lie group theory, that the relative rotation R can be computed as

$$R = U B V^T$$

with U and V the matrices containing the left and right singular vectors of E (i.e. $E = U \Sigma V^T$), and B a matrix representing a rotation around the optical camera axis over $\frac{\pi}{2}$ or $-\frac{\pi}{2}$. The direction of the relative translation is easily computed since the epipole (the projection of the second camera center in the first image) is known from the fundamental matrix. It is now clear that there is one parameter we can not calibrate: the actual value of the baseline. We can however assume that this value will not deviate much from the mechanical specs. This assumption is valid since it is unlikely that the distance between the cameras, which are fixed on the tilt axis, will change much.

The computed values for R and t are used as an initialization for a non-linear Levenberg-Marquardt minimization which finds back the values of R and t that minimize sum of all distances between points and their corresponding epipolar lines. The result is a very accurate calibration of the relative transformation between the two images.

2.4 Pan-tilt calibration

Computing the relative transformation between the two cameras is an important part of the calibration but it does not suffice. For rover localization and generation of terrain reconstructions, the transformations between the cameras and the Imaging Head and between the Imaging Head and the lander need to be known as well.

2.4.1 From Imaging Head to Lander

Calibrating the relative transformation between the Imaging Head frame and the lander implies calibration of the pan and tilt axes. It is clear that this transformation depends on the actual angle of rotation around both the pan and tilt axis. From the world's point of view, the motion of the IH can be described as a rotation around the pan axis followed by a rotation around the tilt axis. The pan axis is never altered but the orientation of the tilt axis depends on the pan angle.

If we look from the point of view of the IH however, it is the tilt axis that never changes and the orientation of the pan axis depends on the tilt angle.

The latter view will be employed because it fits very well in the philosophy where one derives the entire chain of calibration transformations from the cameras, which are the only measurement device, to the lander.

When we speak of THE pan and tilt axis, we mean the pan and tilt axis for a pan and tilt angle of 0.

2.4.2 Relative transformations between views

To calibrate the pan and tilt axes, stereo images of the same ring and the same segment are used respectively. Especially the overlap between consecutive stereo images is important in the strategy.

Tilt For the calibration of the tilt axis, a stereo image of the outer ring of a certain segment is recorded. The IH is commanded to execute a tilt motion and to record a stereo image of the second ring. One has to make sure that there is sufficient overlap between the two image-pairs. .

Corresponding features in the images of the first image pair can be found as explained in Section 2.3.1. Because we know the relative transformation between the two cameras, we can reconstruct the features in 3D. The same is done in the second image pair. Because of the overlap, some of the features will be visible in both image pairs. We can find correspondences between these features by running the matching algorithm of Section 2.3.1 on the two images of the left or the right camera. The corresponding features allow us to align the reconstruction of the second pair with the reconstruction of the first pair. This yields the relative transformation between the first and second IH frame.

Pan For the pan axis, the computation of the relative transformation between two views is slightly different. The IH records a stereo pair, is commanded to execute a pan motion and records a second stereo pair.

In this case there are almost no features that are present in all 4 images of the two views. Due to the verging of the two cameras however, there are some features that can be seen in one stereo view and in one image of the other image pair. Again we find back corresponding features between the left image of the first pair and the right image of the second pair with the algorithm of Section 2.3.1. Because the features are visible in both images of the first stereo view, they can be reconstructed in 3D. They are also visible in one image of the second view, so one can apply a pose-estimation of the camera of the second pair in which the features are visible, yielding the pose of this camera in the frame of the first view and thus the relative transformation for the pan motion.

2.4.3 Actual calibration of pan and tilt axes

The previous section provides us with a set of relative transformations between IH frames. Part of these come from tilt motions, the other part from pan motions. If one knows the relative transformation between two poses of the IH and one assumes a pure rotation between the two poses, the rotation axis can be computed linearly.

First the tilt axis will be calibrated. For this we will use the relative transformation between views where the motion between the poses was a pure tilt rotation.

From the set of relative transformations between two IH frames where the relative motion is a pure pan rotation, the pan axis can be computed linearly as well. However, from the point of the view of the cameras, the pose of the pan axis changes according to the tilt angle. This means one first has to "undo" the influence of the tilt rotation before one can use the relative transformation to compute the pan axis. This is not a problem because a good approximation of the tilt axis has been computed in the previous step.

Iterative procedure During the acquisition of the data one tries not to change the pan angle if a pure tilt rotation is executed and vice versa. In any real system however, there will be deviations from the desired angles. This means that the computation of the tilt axis will not be correct because the linear algorithm computes the real rotation axis, which is not the tilt axis if there is an -even small- pan component. But there is a solution to this problem. In the second step a good approximation of the pan axis was found, so if we account for the small deviations of pan with the current computed value of the pan axis, we can recompute the tilt axis more accurately. This in turn allows us to update the pan axis etc. We can repeat this iterative procedure until the solution for the axes has converged. In reality three iterations have proven to be sufficient.

3 3D Terrain modeling

After calibration of the Imaging Head is performed, the process of generating a 3D model or models of the planetary terrain can commence. This modeling is vital for to accomplish the goal of planetary exploration. Its input are all images of the terrain and the calibration of the Imaging Head. The output of the terrain modeling can have different forms but the most important is the Digital Elevation Map (DEM). Due to the known setup of the system and the results of the calibration, a specific algorithm has been developed to perform the 3D terrain modeling. If such information were unknown, other techniques, as described in [7] can still retreive 3D information but these will not be discussed here.

3.1 Generation of disparity maps

On each pair of images recorded by the Imaging Head, a stereo algorithm is applied to compute the disparity maps from the left image to the right and from the right image to the left. Disparity maps are an elegant way to describe correspondences between two images if the images are **rectified** first. The process of rectification re-maps the image pair to standard geometry with the epipolar lines coinciding with the image scan lines [10, 6]. The correspondence search is then reduced to a matching of the image points along each image scan-line. The result (the disparity maps) is an image where the value of each pixel corresponds with the number of pixels one has to move

to left or right to find the corresponding pixel in the other image.

In addition to the epipolar geometry other constraints like preserving the order of neighboring pixels, bidirectional uniqueness of the match and detection of occlusions can be exploited.

The dense correspondence scheme we employ to construct the disparity maps is the one described in [4]. This scheme is based on the dynamic programming scheme of Cox [1]. It operates on rectified image pairs and incorporates the above mentioned constraints. The matcher searches at each pixel in the left image for the maximum normalized cross correlation in the right image by shifting a small measurement window along the corresponding scan line. Matching ambiguities are resolved by exploiting the ordering constrain in the dynamic programming approach.

The algorithm was adapted to yield sub-pixel accuracy by employing a quadratic fit of the disparities.

3.2 Digital Elevation Maps

A digital elevation map or DEM can be seen as a collection of points in a top view of the 3D terrain where each point has its own height or elevation. Classical approaches to generate DEMs from disparity maps or depth maps suffer from the problem that the resulting DEM has no regular form in 3D. A possible solution to this is to triangulate the resulting points which forms a 3D surface. Intersecting this surface with vertical lines of a regular grid yields a regular DEM. However, in this strategy no optimal use is made of the information present in the disparity maps. The algorithm we devised does use this information. It is explained in the following section.

3.2.1 Generating a regular DEM

The algorithm proposed for generating regular DEMs in the ROBUST project fills in a top view image of the terrain completely, i.e. a height value can be computed for every pixel in the top view image, except for pixels that are not visible in the IH because of occlusions. These occlusions are found in a very simple way. The principle of the algorithm is illustrated in figure 2.

The terrain is divided into cells: the pixels of the DEM. For each cell the stereo pair image is selected in which the cell would be visible if it had a height of zero. A vertical line is drawn and the projection of this line in the left and right disparity image of the stereo pair is computed. Figure 3 illustrates the algorithm that is used to determine the height of the terrain on that line.

The two images are the disparity images of the left and right stereo image respectively. The solid line A-B is the projection of the vertical line in both disparity images. Now imagine placing a light where the left camera

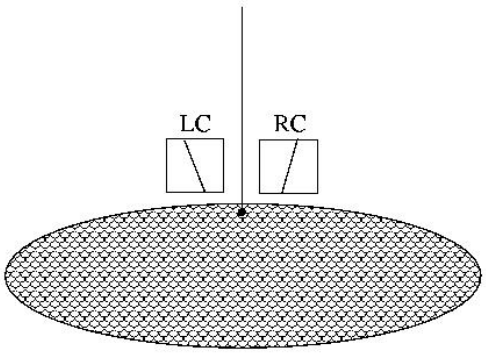

Figure 2: Setup of the DEM generation

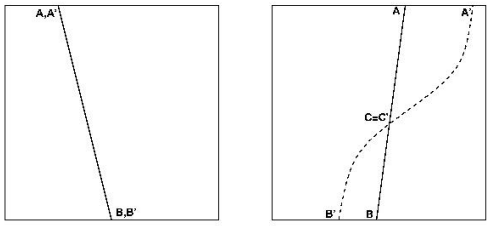

Figure 3: DEM generation in detail

is. This light shines on the vertical line which throws a shadow on the terrain. In the left image this shadow of course has the same projection as the line itself. In the right image however this is not the case. The projection of the shadow in this image is the smooth curve from A' to B'. The part of this curve from A' to C' is the *real* part of the shadow (i.e. it would be visible on the terrain). The part from C' to B' can be seen as the *virtual* part of the shadow, coming from the part of the vertical line below the surface of the terrain. This shadow–curve can be computed using the disparity in the left disparity image of every pixel of the projected line A–B. The intersection point C of the vertical line and the terrain can then be found as the point where the shadow A'–B' intersects the line A–B.

Occluded regions are easily detected since in this case no intersection point C exists. The height value of occluded cells can not be computed and these cells get a certain value in the DEM which marks them as unseen. This particular scheme also makes it possible to generate regular digital elevation maps at any desired resolution, interpolating automatically if needed. For the parts of the terrain close to the boundary of a ring, different parts of the vertical line will be projected in different stereo views. Therefore it is possible that data of two different stereo views has to be combined.

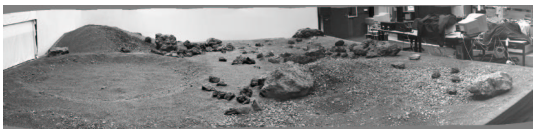

Figure 4: Panoramic image of a planetary testbed

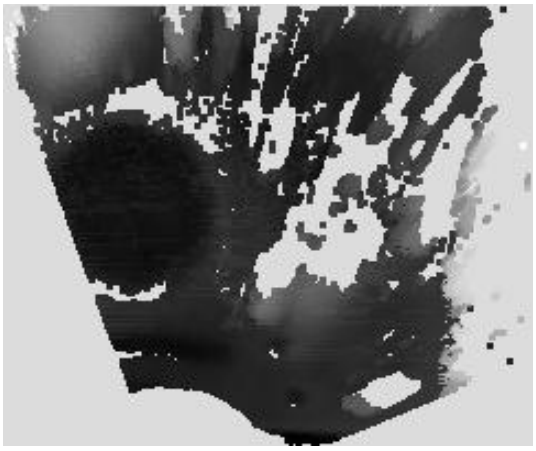

Figure 5: Resulting DEM of the testbed

3.3 Results

Figure 4 shows a panoramic image of the planetary testbed at ESTEC in Noordwijk (The Netherlands) on which the system was tested before and during the ASTRA 2000 workshop.

In figure 5 the resulting digital elevation map (DEM) is shown. There are quite some (white) gaps in the map that are due to occlusions.

The resulting DEM can be triangulated and textured to yield a 3D reconstruction of the the terrain. Figure 6 shows some views of this reconstruction.

4 Path Planning

Once the planetary terrain has been reconstructed, scientists can indicate sites of interest (Points Of Reference or PORs) which the Nanokhod should visit. Paths have to be computed between successive PORs. This is where the automatic Path Planner plays its role. Given a terrain map, an initial rover position and heading and a desired rover position and heading, the Path Planner will find a path, using an A* framework, which takes the rover from the initial state to the goal state and which is optimal in terms of a series of parameters, like energy consumption, risk of tip-over, use of tether etc.

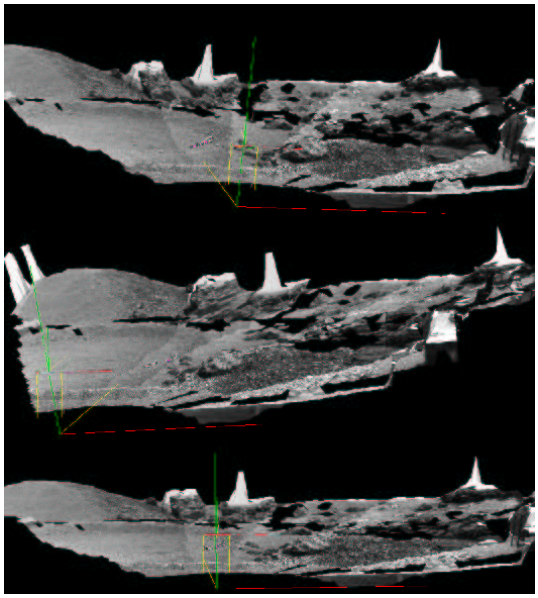

Figure 6: Resulting views of the reconstruction

4.1 Travel Cost Map

The Travel Cost Map (TCM) provides a measure for the cost of traversal based on metrics inherent to the terrain. In the current implementation, a simple metric based on the gradient of the Digital Elevation Map (DEM) is used. Another metric used characterizes the uncertainty of the terrain data, the farther from the lander camera the higher the uncertainty. Areas occluded by rocks also have high uncertainty. Soil characteristics are taken into account by the ROG Map.

4.2 The Hierarchical Approach

The rover can move according to a set of available operators (also called rover movements), which take the rover from one position and heading (this pair is also known as a state) to another position/heading. Each operator has an associated cost. The main term of this cost is computed from the above mentioned TCM. Given that A* is computationally very complex, finding a path in a reasonably large terrain, using complex operators for the rover movements, can take a long time. Moreover, given that our path execution will certainly be less than perfect, it is not desirable to find a path which is optimal, whilst closely surrounded by a difficult area. Considering both these facts has led to the choice of a hierarchical approach to the path planning problem.

4.2.1 Finding the Corridor

At the first stage, a traverse is planned between the start and goal states using A* covering the whole ter-

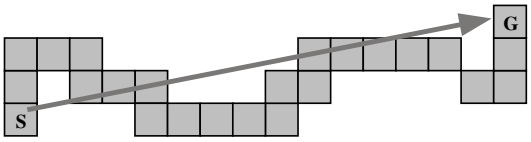

Figure 7: Corridor

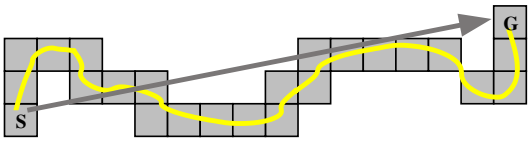

Figure 8: Refined path

rain but with reduced resolution, the cells being somewhat larger than the size of the Nanokhod so that the rover can manoeuvre comfortable within the corridor. A low–resolution TCM is used for this. The transition operators are simple forward, backward, left and right, allowing to apply a highly optimized and fast version of A*. The result is a corridor (see figure 7) in which the rover may safely move. The cells belonging to the corridor are marked in a Restrain Grid which is used in the second stage of the algorithm. The user has the ability to mark additional cells to enlarge the search space.

4.2.2 Refinement of the Path

At the second stage the path is refined using the high–resolution TCM. By restricting the search to cells marked in the Restrain Grid constructed in the previous stage more complex operators and full available resolution can be used within reasonable time constraints.

The rover is abstracted as a directed point, representing its position and heading. The representation of the operators to take the rover from one state to another is kept very general, i.e. a rotation followed by a translation. As such, the set of operator is very customizable. The cost of applying an operator is determined by using a number of cost evaluation points. The cost is calculated as a weighted sum of the costs at these points, evaluated at the resulting rover pose. The evaluation points represent parts of the 'virtual' rover during an after the completion of the corresponding move. The position of the evaluation points are calculated based on the rover dimensions, the parameters of the rover movement, the desired safety margin and the resolution of the TCM.

4.3 Segmentation of the Path

The result of the hierarchical A* algorithm is a high-resolution path (see figure 8), represented by an ordered list of rover poses, bringing the rover from its start pose to the desired destination. This representation must be converted to a Path Segment List (PSL) which can be executed by the rover. The PSL is sequence of straight path segments associated with a rover action. An iterative procedure is used to approximate the high-resolution path taking into account the traversability along the path.

5 Conclusion

In this paper, three important preparation tasks of a planetary exploration system have been described. The calibration, terrain reconstruction and path planning modules are based on recent developments in computer vision and robotics and have been tested on both artificial data and on a planetary testbed.

Acknowledgments

We acknowledge support from the Belgian IUAP4/24 'IMechS' project. We also wish to thank all partners of the ROBUST project for the collaboration.

References

[1] I. Cox, S. Hingorani and S. Rao, "A Maximum Likelihood Stereo Algorithm", In *Computer Vision and Image Understanding*, Vol. 63, No. 3, May 1996.

[2] M. Fischler and R. Bolles: "RANdom SAmpling Consensus: a paradigm for model fitting with application to image analysis and automated cartography", In *Commun. Assoc. Comp. Mach.*, 24:381-95, 1981.

[3] C. Harris and M. Stephens: "A combined corner and edge detector", In *Fourth Alvey Vision Conference*, pp. 147-151, 1988.

[4] R. Koch, "Automatische Oberflachenmodellierung starrer dreidimensionaler Objekte aus stereoskopischen Rundum-Ansichten", *PhD thesis*, University of Hannover, Germany, 1996 also published as *Fortschritte-Berichte VDI*, Reihe 10, Nr.499, VDI Verlag, 1997.

[5] S. Maybank, "Theory of reconstruction from image motion", *Springer Verlag*, 1992.

[6] M. Pollefeys, R. Koch and L. Van Gool, "A simple and efficient rectification method for general motion", *Proc.ICCV'99 (international Conference on Computer Vision)*, pp.496-501, Corfu (Greece), 1999.

[7] M. Pollefeys, R. Koch, M. Vergauwen and L. Van Gool, "Metric 3D Surface Reconstruction from Uncalibrated Image Sequences", *Proc. SMILE Workshop (post-ECCV'98)*, LNCS 1506, pp.138-153, Springer-Verlag, 1998.

[8] R. Rieder, H. Wanke, H. v. Hoerner. e.a., "Nanokhod, a miniature deployment device with instrumentation for chemical, mineralogical and geological analysis of planetary surfaces, for use in connection with fixed planetary surface stations", In *Lunar and Planetary Science*, XXVI,pp. 1261-1262, 1995.

[9] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses." *IEEE Journal of Robotics and Automation*, 3(4): pp. 324-344, 1987

[10] C. Loop and Z. Zhang. "Computing Rectifying Homographies for Stereo Vision". IEEE Conf. Computer Vision and Pattern Recognition (CVPR'99), Colorado, June 1999.