

Calibration and 3D Measurement from Martian Terrain Images

Maarten Vergauwen, Marc Pollefeys* and Luc Van Gool
ESAT-PSI, K.U.Leuven,
Kard. Mercierlaan 94, B-3001 Heverlee, Belgium
Phone: +32-16-32.10.64, Fax: +32-16-32.17.23
vergauwe | pollefeys | vangool@esat.kuleuven.ac.be

Abstract

In this paper a new on-Mars calibration procedure of a planetary lander vision system is presented. It is based on recent developments in computer vision. The calibration is retrieved from the images of the Mars terrain (the same images from which the 3D measurements of the terrain are obtained). The procedure is based on the relations between multiple views of the same scene. The main advantage of this calibration procedure is that it allows to retrieve the calibration on Mars without any additional requirements on the system. Once the system has been calibrated, the same images can be used to estimate a digital elevation map of the environment around the lander. This means that the terrain based calibration causes no overhead on transmission. 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. A new algorithm was developed to efficiently extract this from the disparity maps. This algorithm can generate a regular digital elevation map at any desired resolution (interpolating if necessary) and easily takes occlusions into account.

1 Introduction

The work described in this paper was performed in the scope of the ROBUST¹ project of the European Space Agency (ESA). In this project an end-to-end system is developed for a planetary exploration mission.

1.1 Hardware

The ROBUST system consists of three important parts.

* Postdoctoral Fellow of the Fund for Scientific Research - Flanders (Belgium)(F.W.O. - Vlaanderen).

¹The ROBUST consortium consists of the Belgian companies SAS and OptiDrive, the K.U.Leuven departments PMA and ESAT-PSI and the German companies DLR and vH&S.

1. the planetary rover: the Nanokhod, a small and simple rover, designed to carry instruments in the immediate surroundings of a lander. It is equipped with a tether cable, providing the rover with power and data connection to the lander which allows a very high ratio instrument-mass/rover-mass [11]. Figure 1 on the left shows an image of the Nanokhod.
2. The Planetary Lander which contains the Imaging Head, an On-Board computer and the Control System for both Nanokhod and Imaging Head. The right image of figure 1 shows one of the cameras.
3. The On Ground Control System

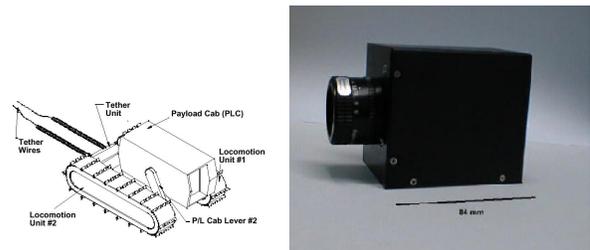


Figure 1: left: the Nanokhod; right: one of the space-approved cameras

The Imaging Head is both used for recording images from which a reconstruction of the planetary terrain is computed and for controlling the motion of the rover. It consists of a stereo head, mounted on a unit which allows for pan and tilt motions and 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.

1.2 Utilization

A typical utilization scenario will deploy the Imaging Head as soon as possible after the landing of the planetary system. 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. Since the cameras have a limited field of view (23x23 degrees) the entire environment is not recorded at once but it is segmented into rings according to the tilt angle and each ring is divided into segments according to the pan angle of the Imaging Head (see figure 2). The outermost boundary of the recorded terrain lies at twenty meters from the camera. For each of the segments a stereo image pair is recorded and sent down. The values of the actual pan and tilt angles can be read out from the encoders of the pan-tilt motors and are sent down together with the corresponding images.

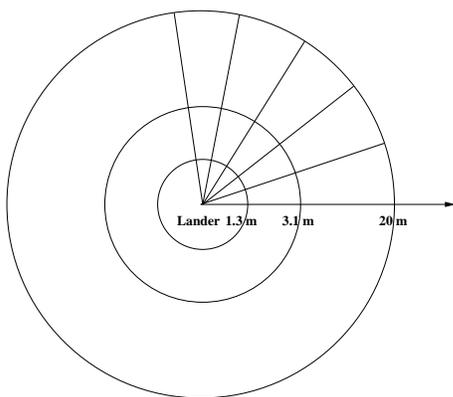


Figure 2: Segmentation of the terrain into segments

In this paper new techniques are described which perform the calibration and the reconstruction tasks.

2 Calibration

Every planetary mission is a high-risk operation. 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. If the camera housing were not so rigidly built and the camera intrinsics were likely to change during launch or landing, Algorithms exist that can retrieve these intrinsic parameters from images too [14, 9].

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 [12].

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. One needs to be absolutely sure of the pose of these objects for the calibration to have any meaningful result. 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. During the design of the Imaging Head, robustness was a very important issue and therefore the number of moving items was minimized. Therefore, no zooming and focusing system was added to the cameras which thus have a constant focal length. 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 3. Once this relative calibration is performed, 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 4 deals with this problem.

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 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).

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 3 makes this clear.

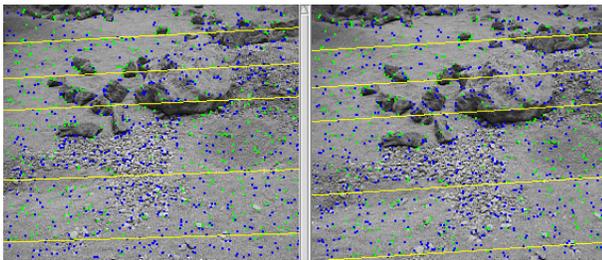


Figure 3: Epipolar geometry of an image pair

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. The result consists of points or *corners* in the images determin-

ing 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 data 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 does not change. Figure 4 makes this clear. Stereo images of differ-

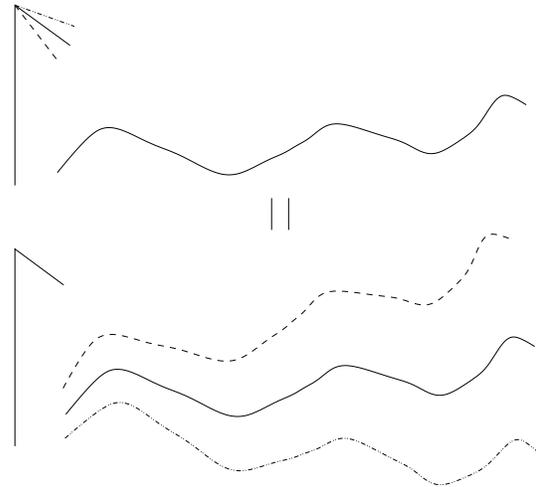


Figure 4: Combining different segments

ent rings are obtained by tilting the Imaging Head. However, one could imagine the camera to be kept steady and the terrain to be tilted. This would result in the same stereo images. That’s why the possible correspondences of the different rings and segments can be combined to compute the epipolar geometry more accurately.

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.

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. [6]. 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. If there were some change in the actual value of the baseline, the consequences of fixing it to the (wrong) value of the specs are not harsh because all measurements are done with the same measurement system. This means that the reconstruction of the terrain in 3D will show some deviation from the real terrain but since the localization of the rover is done using the same imaging system, this is not a problem. Only during the path-planning phase some absolute measurements are needed to make sure the rover is not commanded to climb too steep a slope or to overcome too high an obstacle. An error of these measurements of a few percent is not a problem.

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.

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.

4.1 The Imaging Head frame

For sake of clarity we define a virtual “**Imaging Head frame**” in “the middle” of the two cameras. This means that the relative translation and rotation between the left camera and the Imaging Head frame is equal to the translation and rotation between the Imaging Head frame and the right camera.

$$\begin{aligned} R(l, ih) &= R(ih, r) \\ t(l, ih) &= t(ih, r) \end{aligned}$$

This can easily be calculated from the relative transformation between the two cameras, computed in Section 3, using the “matrix square root” of the rotation matrix.

4.2 From Imaging Head to Lander

Calibrating the relative transformation between the Imaging Head frame and the lander is much more complicated because it 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.

4.3 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. This setup is shown in figure 5. The area shaded from bottom left to top right is visible in the first view. The area shaded

from bottom right to top left is visible in the second. The overlapping area is of course visible in both views.

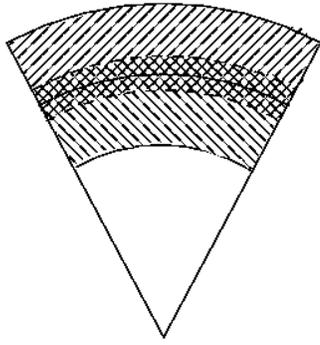


Figure 5: Symbolic representation of the setup for the computation of a relative transformation for a tilt motion

Corresponding features in the images of the first image pair can be found as explained in Section 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 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 setup is shown in figure 6.

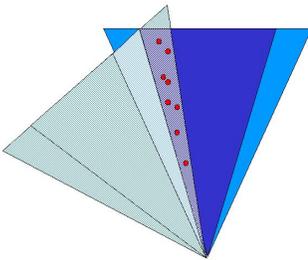


Figure 6: Symbolic representation of the setup for the computation of a relative transformation for a pan motion

It is clear that in this case there are almost no features that are present in all 4 images of the two views. Due to the

vergence 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. These features are represented by a dot in image 6. 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 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. Using the relative transformation between camera and IH from Section 4.1, we have found back the relative transformation between the two stereo views.

4.4 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. In this section we will explain how to use these relative transformations to compute the pan and tilt axes.

Rotation axis from relative transformation First we will explain how a rotation axis can be computed linearly if the relative transformation is known between two frames between which there has been a rotation around this axis. The axis can be represented by its direction and one point on the axis.

If we call $T = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$ the relative transformation between the two frames, than any homogeneous point $p = (x \ y \ z \ w)^T$ on the tilt axis satisfies the equation

$$\begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

because points on the axis do not change under a rotation around this axis.

Finding the direction of the axis is easily done by solving the equation for $w = 0$. Finding a point on the axis is also easy by fixing one coordinate and computing the other two.

Tilt 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. Because the relative transformation of a set of such rotations has been computed, one can calculate the axis of rotation linearly for every transformation as explained before. A least

squares fit yields a good first approximation of the tilt-axis in the IH frame.

Pan 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. This means we can rotate each IH frame back to a theoretic tilt-angle of zero. All relative transformations with a pan motion yield a pan axis and a least squares fit yields a good approximation of the pan-axis in the IH frame for a tilt angle of zero.

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.

5 Results

The calibration algorithms have been tested on artificial data. A planar scene was constructed and pairs of images were generated with a visualization toolkit.

First the relative calibration between the two cameras was computed. The original relative transformation was the following.

$$\begin{pmatrix} 0.9848078 & 0.0 & -0.1736482 & 0.4980973 \\ 0.0 & 1.0 & 0.0 & 0.0 \\ 0.1736482 & 0.0 & 0.9848078 & 0.04357787 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{pmatrix}$$

During calibration data of 9 image pairs was combined and 2591 corners were matched to calculate the relative transformation. Since the scene was planar, calibration would

not have been possible without this combination. The resulting relative transformation was

$$\begin{pmatrix} 0.984812 & 0.000039 & -0.1736424 & 0.498163 \\ -0.000027 & 1.000000 & 0.000072 & -0.000265 \\ 0.173624 & -0.000066 & 0.984812 & 0.042813 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{pmatrix}$$

The pan and tilt axes were calibrated from the same data too. The original pan axis coincided with the Z axis and the tilt axis with the X axis. If we represent both axes as a direction and a point on the axis, the original axes could be represented as follows.

$$\begin{aligned} D_{pan} &= (0.0 \ 0.0 \ 1.0) \\ P_{pan} &= (0.0 \ 0.0 \ 0.0) \end{aligned}$$

$$\begin{aligned} D_{tilt} &= (1.0 \ 0.0 \ 0.0) \\ P_{tilt} &= (0.0 \ 0.0 \ 0.0) \end{aligned}$$

The same 9 image pairs were used to calibrate the axes and the following results were found.

$$\begin{aligned} D_{pan} &= (0.00187 \ 0.01062 \ 0.99993) \\ P_{pan} &= (-0.00388 \ 0.00113 \ -0.00011) \end{aligned}$$

$$\begin{aligned} D_{tilt} &= (0.99999 \ 0.00207 \ -0.00122) \\ P_{tilt} &= (0.0001 \ -0.00516 \ 0.00143) \end{aligned}$$

6 3D Terrain modeling

After the calibration of the IH is performed, the process of generating a 3D model or models of the planetary terrain can commence. This modeling is vital 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). In the following sections we will describe the different steps that are performed to obtain such a DEM.

7 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 [13, 10]. 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 [5]. 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 constraint in the dynamic programming approach.

The algorithm was adapted to yield sub-pixel accuracy by employing a quadratic fit of the disparities. This is shown in figure 7.

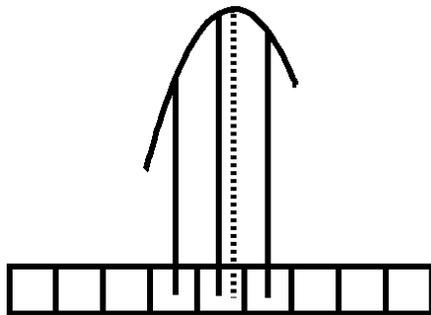


Figure 7: The maximum of the parabola, fitted through three neighboring values of the correlation function yields a correspondence with sub-pixel accuracy.

The stereo algorithm computes the correlation function between a window around a certain pixel in the first image and the windows around all pixels of the same line in the second image. The pixel for which the maximum of this function -subject to certain constraints- is computed is selected as the corresponding pixel of the pixel in the first image. One can fit a parabola through this maximum and the values of the correlation function evaluated in the two neighboring pixels. The maximum of this parabola yields

a sub-pixel correspondence of the pixel in the first image.

The resulting disparity map of the left image of an image pair of an artificial scene can be seen on the right in figure 8. The scene contained a flat surface with a couple of textured cubes. The left image of the image pair can be seen on the left in figure 8.

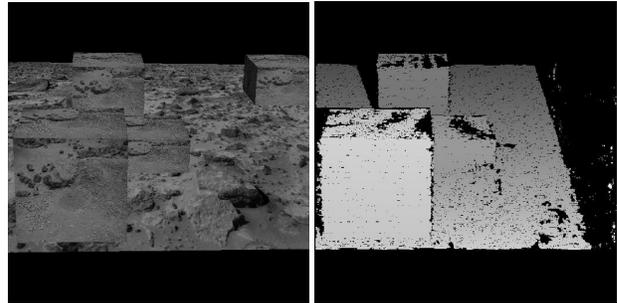


Figure 8: Original left image and resulting disparity map of the left image of an image pair of an artificial scene.

8 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 consist of two steps.

- For each stereo image pair the disparity images are used to construct depth images. These are images with the same size as the original images. The value of each pixel corresponds to its depth: the distance of the point to the camera.
- A limited amount of points of each depth image is reconstructed in 3D. These points form the DEM.

The problem of this scheme is 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.

8.1 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 9.

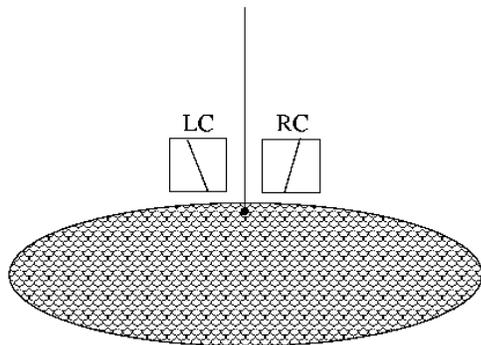


Figure 9: Setup of the DEM generation

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 10 illustrates the algorithm that is used to determine the height of the terrain on that line.

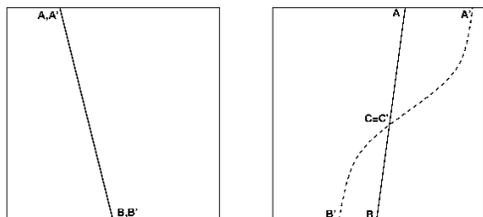


Figure 10: DEM generation in detail

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 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 $A-B$. Some remarks can be made about this procedure.

- It is possible to detect occluded regions easily. This is the case for cells that are not visible in both stereo images. The height value of these cells can not be computed and these cells get a certain predefined value in the DEM which marks them as unseen.
- This particular scheme 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. This is not a problem because the transformation between the views can easily be computed since the calibration has been calculated.

9 Results

A digital elevation map of an artificial scene was constructed using the techniques described in the previous section. A disparity map of one of the pairs of this scene was shown in figure 8. Figure 11 shows the resulting DEM. The left image shows a orthographic top view of the DEM, illustrating the regularity of the grid in 3D. The right image of figure 11 shows a perspective view of the DEM.

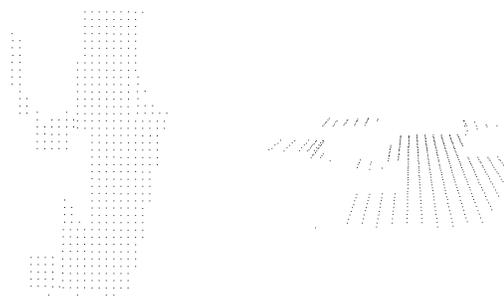


Figure 11: Two views of the resulting DEM, generated from the disparity map of figure 8

The DEM can be used for path-planning purposes but is not suited for visualization. To this end we can construct a triangulated mesh model (TMM) from the DEM. The DEM is triangulated and texture is mapped onto the triangles. Figure 12 shows the result. The reconstructed terrain clearly

shows some holes due to occlusions of the terrain by the large cubes in front.

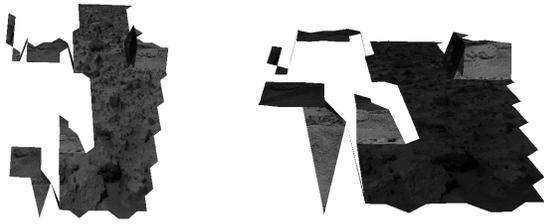


Figure 12: Triangulated mesh model with texture, obtained from the DEM of figure 11

Acknowledgments

We acknowledge support from the Belgian IUAP4/24 'IMechS' project.

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] J. Knight and I. Reid, "Self-calibration of a stereo-rig in a planar scene by data combination", In *Proceedings International Conference on Pattern Recognition (ICPR 2000)*, pp. 411-414, Barcelona, 2000.
- [5] R. Koch, "Automatische Oberflächenmodellierung 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.
- [6] S. Maybank, "Theory of reconstruction from image motion", *Springer Verlag*, 1992.
- [7] J. Mundy and A. Zisserman, "Machine Vision", In J.L. Mundy, A. Zisserman, and D. Forsyth (eds.), *Applications of Invariance in Computer Vision*, Lecture Notes in Computer Science, Vol. 825, Springer-Verlag, 1994.
- [8] M. Pollefeys, R. Koch, M. Vergauwen and L. Van Gool, "Metric 3D Surface Reconstruction from Uncalibrated Image Sequences", In *Proceedings SMILE Workshop (post-ECCV'98)*, LNCS 1506, pp.138-153, Springer-Verlag, 1998.
- [9] M. Pollefeys, R. Koch and L. Van Gool. "Self-Calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters", *International Journal of Computer Vision*.
- [10] 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.
- [11] 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.
- [12] 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
- [13] C. Loop and Z. Zhang. "Computing Rectifying Homographies for Stereo Vision". *IEEE Conf. Computer Vision and Pattern Recognition (CVPR'99)*, Colorado, June 1999.
- [14] A. Zisserman, P. Beardsley and I. Reid, "Metric calibration of a stereo rig", *Proceedings IEEE Workshop on Representation of Visual Scenes*, Cambridge, pp. 93-100, 1995.