ALGORITHMIC SOLUTION AND SIMULATION RESULTS FOR VISION-BASED AUTONOMOUS MODE OF A PLANETARY ROVER*

ABSTRACT

A vision based navigation (VBN) system is chosen as a basic tool to provide autonomous operations of a planetary rover during space missions. That means that the rover equipped with a stereo vision system and perhaps a laser ranging device shall be able to maintain a high level of autonomy under various illumination conditions and with little a priori information about the underlying scene. As it is specified in the LEDA Moon exploration project currently under focus by the European Space Agency ESA, during the autonomous mode the rover should perform the following operations: on-board absolute localization, elevation model (DEM) generation, obstacle detection and relative localization, global path planning and execution.

Focus of this paper is to simulate some of the path planning and path execution steps. That is done with the help of a laboratory terrain mockup and a high precision camera mounting device. The following operations are performed on the basis of stereo image sequence: 3D scene reconstruction, risk map generation, local path planning, update cameras position during the motion on the basis of landmarks tracking. It is shown that standalone tracking using automatically identified landmarks is robust enough to give navigation data for further stereoscopic reconstruction of the surrounding terrain. Iterative tracking and reconstruction leads to a complete description of the rover path and its surrounding with an accuracy high enough to meet the specifications for unmanned space exploration.

1 INTRODUCTION

Significant progress in outdoor autonomous vehicle development has been made during the last decade. The implementation of the recent technological achievements to planetary rovers is a challenging task outlined in the LEDA Moon exploration project [1]. Three roving control modes must be supported by the Vision-Based Navigation (VBN) system according to LEDA mission objectives:

1. teleoperating mode of ground based piloting;
2. medium autonomy mode or autonomous piloting and ground based navigation;
3. full autonomy mode or autonomous navigation;

This paper outlines the concept of fully autonomous mode for the rover equipped by an active / passive imagery sensors setup.

In the second section a closed-loop algorithmic solution suitable for on-board implementation is described. Section 3 presents simulation results of rover operations with the help of a Lunar terrain mockup.

2 CLOSED LOOP SOLUTION FOR AUTONOMOUS NAVIGATION SCENARIO

2.1 SPECIFIC CONDITIONS

The Rover's VBN system must overcome a set of specific conditions and requirements which are present on a planetary surface where immediate human intervention is in fact impossible. These are:

1. Automatic initial calibration of the vision sensors.

* The work described here is being supported by the Austrian Science Foundation (FWF), Grants M00265-MAT and S7003-MAT, and JOANNEUM RESEARCH.

2. Absence of accurate reference points or landmarks for precise self calibration.
3. Low angle illumination conditions.
4. Low angle viewing conditions.
5. No a priori information about underlying terrain.

On the other hand a relatively slow motion of the rover allows to move in a stop/thinking mode which simplifies the necessary on-board operations.

2.2 STEREO VISION SYSTEM ARRANGEMENT.

Experiences of several rover teams have proved that parallel geometry of stereo cameras is difficult to arrange and maintain. For example, an autonomous walker "Dante" based on stereo vision and designed for Antarctic applications is described In [2]. A special platform has been designed to adjust and maintain parallel geometry for three stereo cameras. Evidently, the parallel stereo geometry can be easily distorted after landing and due to the day/night time temperature variations.

Therefore a 3D stereo reconstruction approach based on arbitrary stereo geometry looks more preferable and reliable. Consequently, the necessary calibration procedure for the stereo system is composed from two steps:

- intrinsic parameters (focal length, principal distance, lens distortion) calibration. It is performed on-ground and considered unchanged (or recalculated depending upon the known temperature conditions) during the rover's operations.
- extrinsic parameters (cameras position and orientation) calibration. These parameters must be updated with respect to a given coordinate system while the rover is moving.

2.3 NAVIGATIONAL STEPS

We propose the following approach to accomplish the rover's autonomous mode:

I. Initialisation phase. This step is performed only once to initialise the rover's operational units. These are:
    - Initial calibration of the imagery sensors (laser, stereo cameras) and measuring units (wheel odometer / accelerometer / inclinometer sensors, gyroscope, star tracker).
    - Self-localisation of the rover position either with respect to the orbiter map of the surface or with respect to the lander.

II. Operational phase. This step consists of the following operations implemented in the cycle:
    - Stop-thinking mode.
        1. Stereo image acquisition;
        2. DEM reconstruction;
        3. Risk map generation;
        4. Local path generation;
    - Path execution mode.
        5. Consecutive image acquisition;
        6. Landmarks tracking;
        7. Update rover position (calibration update);
        8. Reflex obstacle detection;

2.3.1 Initial localisation of the rover position

The accuracy of initial localisation for the lander on the planetary surface depends upon the accuracy of the available surface map taken from orbit. An expected size of a landing area for the LEDA mission is equal to 6 km x 1.5 km with 1.5 m ground resolution. After the rover deployment a local map with respect to the lander (its position will specify the beginning of the local coordinate system) can be used on-board for further rover operations. The rover position and orientation shall be constantly controlled on-board in the local coordinate system. That can be done with the help of the vision system, as described in Section 2.3.7

2.3.2 Initial calibration of the sensors

The initial calibration of all measuring units must be done after the deployment of the rover on the surface. The accuracy of the initial calibration for the stereo cameras depends on the accuracy of the 2D coordinates (in the images) and the accuracy of the 3D coordinates (on the scene) for the reference points. According to the evaluation made in [3], a very accurate relative orientation between the cameras can be accomplished by utilising results from stereo matching.

At this stage the initialisation phase is finished and the rover is ready to perform autonomous operations in cycle.

2.3.3 3D DEM reconstruction

A 3D stereo reconstruction algorithm based on the arbitrary geometry of the vision system has been developed for the purpose of autonomous lander navigation during descent. The underlying approach for stereo matching is entitled Hierarchical Feature Vector Matching (HFVM) [5]. The algorithm is based upon the idea of creating a feature vector for each pixel (surrounding) and comparing these features in image pyramids from low resolution level downward to fine resolution level. Each pixel of the resulting disparity map defines a displacement (parallax) between corresponding points on the left and right images of the stereo pair.

The stereo Locus method is used for 3D coordinates reconstruction from the disparity map [6]. Key idea of the method is to find an intersection of a vertical line in the object space with the surface. The output of the Locus 3D reconstruction are height values in a regular grid projected on the horizontal plane in the given world coordinate system.

2.3.4 Risk map generation

A risk map is generated on the basis of the Locus DEM. The basic idea is to select steep slopes and elevated areas from the DEM and to mark them as hazardous and unsuitable for the rover motion by comparing to some predefined safe threshold. The algorithm evaluates the minimal and maximal heights on the DEM in a local window and then calculates the local inclination. The algorithm has $O(\mathrm{N})$ calculations per pixel (where N is the local window size in the optimised version).

2.3.5 Local path planning

Local path planning on the basis of DEM and risk map is the next step in the processing chain. There are several algorithms to build up a safe local path [7]. Our idea is to consider the points located outside of hazardous areas as the nodes of a directed graph. The path is generated from a predefined start point and should reach a specified target point on the local map in such a way to minimise the length of a sub-graph (Dijkstra algorithm [8]).

The detailed description of the algorithm can be found in [9]. The start point of the path is considered as the graph origin. The length of the graph edges are positive values (weights) defined on the basis of DEM values. As usual, the length of a particular path which joins any two given vertices of the graph is defined as the sum of the edge weights composing the path.

The number of operations for searching the shortest path from the start point to the possible destination points is equal to $O(N^2)$, were *N* is the image size.

2.3.6 Natural feature (landmark) tracking

A landmark tracking algorithm [4] is used to follow upon the tracks of homologue points (Interest Points) in two subsequent images. Corresponding displacements between the Interest Points are then used as data base for calibration update.

To extract the Interest Points from the origin image a derivative of the Moravec operator is used. In fact it is not possible to make even coarse prediction of the motion between two consecutive images (frames) taken while the rover is moving. Therefore, we propose to use a hierarchical approach to identify correspondent Interest Points on the subsequent frames. As a first iteration, the disparities (optical flow) of all points on the image are calculated in coarse resolution. This can be done with the help of HFVM down to a certain image pyramid level. During the second iteration, the coarse disparities are used to calculate the correspondences only for the interest points with high resolution.

2.3.7 Calibration update

The essence of the calibration update method (i.e. identifying an instant camera position and pointing parameters along the rover path) is the following. Let us consider two consecutive images (frames N and N+1, Figure 1) taken from the rover during motion. Let us assume, the 3D coordinates of the Interest Points on the frame N are calculated from the DEM. Having 3D coordinates of Interest Points acquired from frame N and their 2D coordinates from tracking between frame N and frame N+1, a calibration can be performed to obtain position and pointing parameters of frame N+1. The calibration method which keeps fixed intrinsic camera parameters and gains only an update for extrinsic camera parameters has been described in [10]. It takes the particular conditions (3D coordinates noise, flat terrain, large number of points) into consideration.

2.3.8 Reflex obstacle detection

Any unexpected obstacle whose size is above a given threshold as well as unknown areas (shadowing, specula reflection) have to be detected during the rover motion at a minimum distance equal to two times the stopping distance of the rover.

The fact that a bright stripe from the laser is easily recognised on the image field can be used to detect obstacles in the vicinity of the rover. A well-known approach is to analyse the shape of the laser stripe on the images taken during the motion. High curvy portions of the trace can be detected as parts of hazardous areas on the underlying scene.

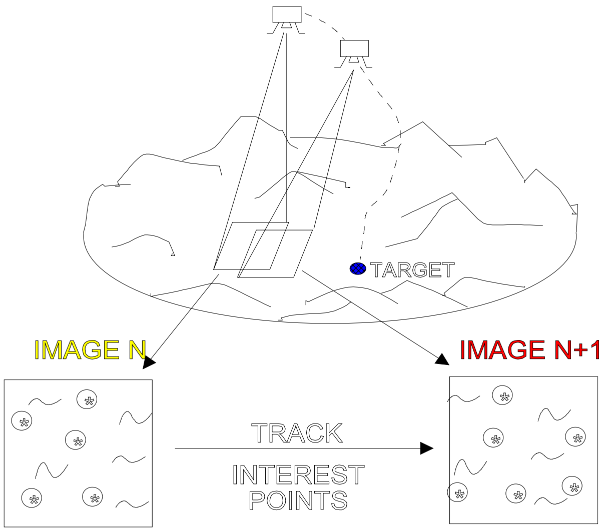

Figure 1: Principle of calibration update using landmark tracking (Example: Spacecraft landing)

2.4 SOFTWARE ORGANISATION, WORKSHARE BETWEEN ON-BOARD AND ON-GROUND PROCESSING

The question which part of the necessary calculations to provide autonomous navigation have to be performed on-board is mostly the question of the on-board computational resources available. Generally, a software should be designed as a combination of separate algorithmic solutions (software blocks) with the minimum possible input-output intermediate data exchange.

It is especially important to separate those software blocks that will need interactive data from the rover sensors as input. All necessary processing shall be sorted out between on-board \ on-ground parts after on-board computational resources are clarified. It is evident that software blocks which need data from the rover sensors as interactive input information are more preferable to be put on-board. The computational complexity of the algorithms to be developed can be considered as a starting point in a trade-off regarding necessary on-board computational power.

3 SIMULATION RESULTS AND ILLUSTRATIONS

This section describes data and processing results collected during the sessions simulating rover operations on the Moon surface. The hardware used for simulation (CamRobII) includes an accurate robot holding two cameras that can be moved with 7 degrees of freedom within a 2m x 1m x 1.5m wide volume. The motion is performed by 7 step engines. CamRobII is controlled by software running on a SPARC workstation. A software interface enables an operator to move the camera on interactive-command basis to capture images and to store video data together with the camera position and orientation data.

A model of the Lunar terrain was placed in a 1.6 m x 2 m bed mockup. For the rover motion simulation session both cameras have been placed in the lowest position and directed forward and slightly downwards. The positions of the cameras are detected with respect to the world coordinate system. The correspondence between CamRobII and world coordinate systems is given by a transformation matrix. Initial coordinates for the very first position of the left stereo camera and the stereo basis for the stereo pairs in the sequence are taken as known. The relative orientation of the stereo configuration based on stereo correspondences was performed using a fixed baseline between the cameras to obtain a certain scale factor. The following sequence of the operations has been performed with the help of the system described above:

1. Stereo sequence image acquisition.
2. Stereo matching and DEM generation.
3. Risk map generation and local path planning.

4. Landmarks tracking and calibration update.

3.1 *INPUT IMAGE ACQUISITION*

A long sequence of the stereo images (40 pairs) was taken. The left image frames with odd image indices are taken by the left camera whereas even frames are regarded to the right camera. Both stereo cameras are set close to the mockup surface and directed slightly downward (15-20 degrees) to obtain a convergent low viewing angle perspective stereo pair of the mockup terrain. After the first pair is acquired both cameras are moved one step forward (15-20 mm) to catch the next stereo pair and so on (Figure 2). The whole camera path is straight for the case presented on the illustrations.

3.2 *DEM GENERATION*

A general elevation model of the mockup terrain (*ortho* DEM) is generated on the basis of each fifth stereo pair (called as *basic stereo pairs*) from the sequence. Intermediate left image frames (e.g. taken by the left camera) between the subsequent *basic stereo pairs* are used for the tracking and calibration update for the left camera. The frequency of the *basic stereo pairs* is defined by the necessary overlap between the reconstructed *ortho* DEM (e.g. > 70%) to keep 3D coordinates known on the underlying surface. The *ortho* DEMs calculated from the subsequent *basic stereo pairs* are merged to generate the entire *ortho* DEM for the underlying terrain. The stereo pairs are matched automatically, the Locus method is applied for the 3D reconstruction. DEM resolution in x and y is selected 1 mm.

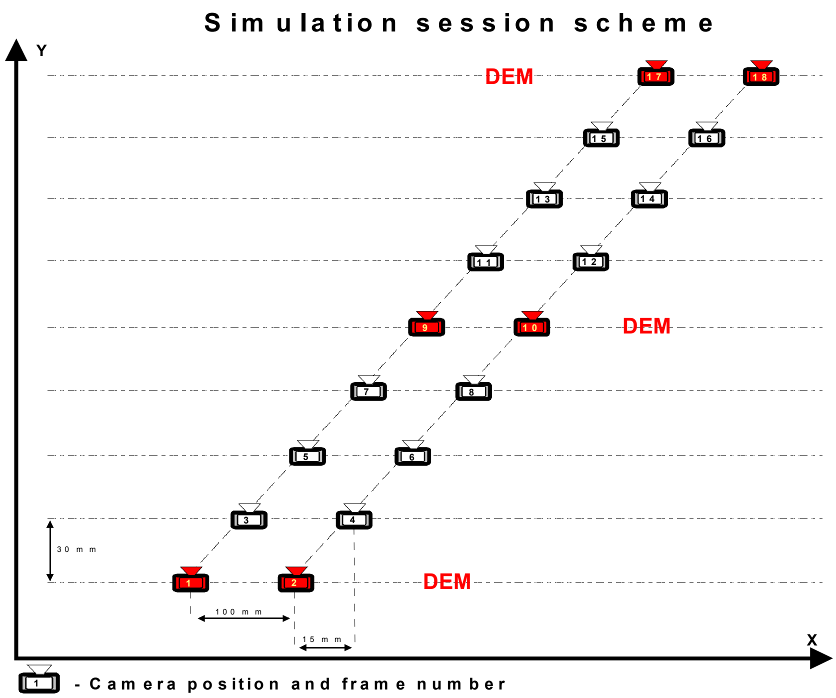

Figure 2: Every fifth stereo pair is used for the reconstruction of a new DEM

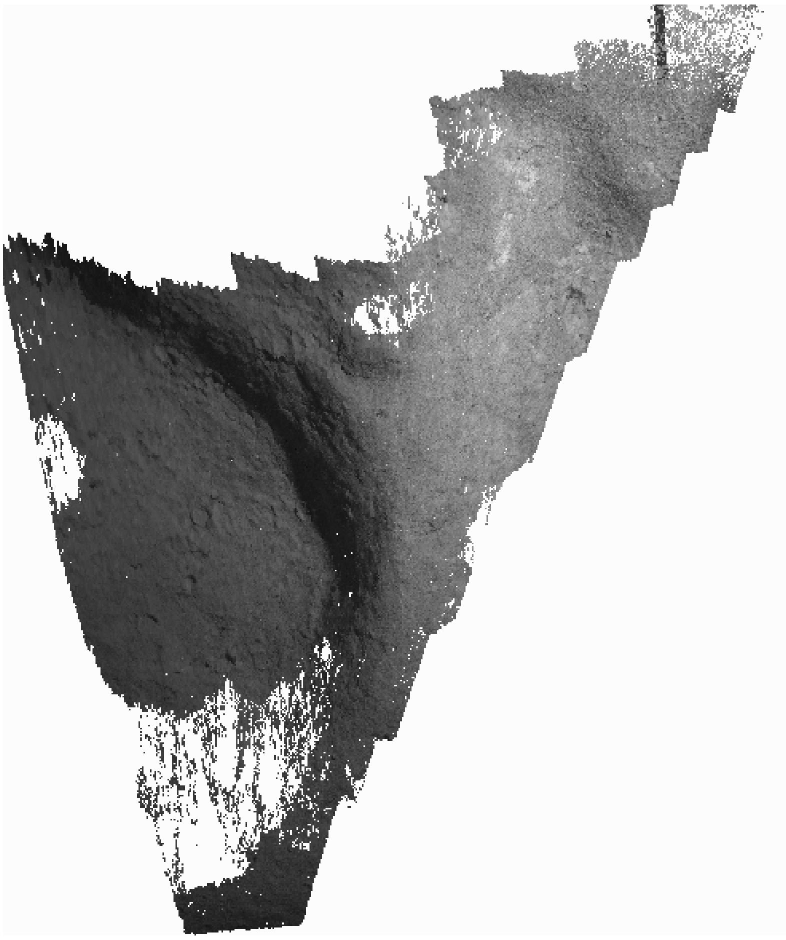

Figure 3 Ortho image merged from ten stereo configurations

Figure 3 depicts the merging result of the ten *ortho* images calculated from the ten subsequent *basic stereo pairs*. Occluded and undefined areas are marked here as white.

3.3 PATH PLANNING

Path planning was done independently from the tracking simulation to demonstrate the robustness of the proposed approach for the Moon like terrain. A local path has been generated on the basis of the once reconstructed DEM.

The generated local path is shown on Figure 4. Hazardous areas unsuitable for the rover motion are marked on the DEM as black. The start and destination points for the rover path are specified by an operator. A safe path is generated automatically within safe areas on the basis of the DEM slopes.

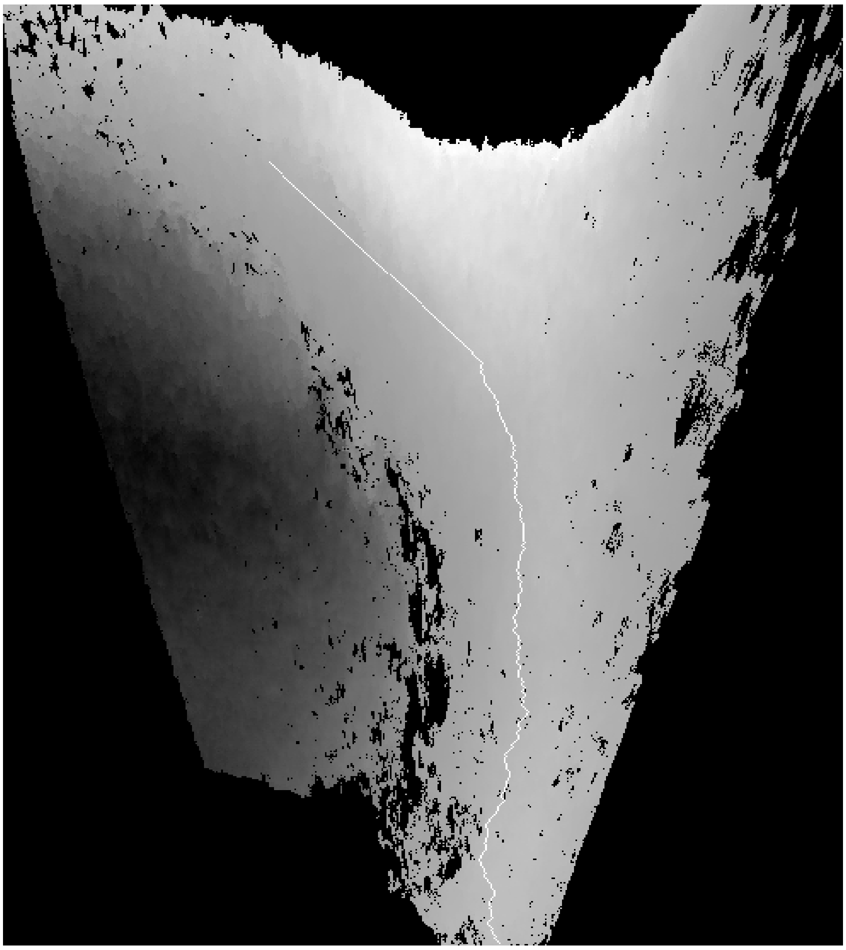

Figure 4. Local path put on the DEM. Unknown and hazardous areas are marked black.

3.4 TRACKING AND CALIBRATION UPDATE

The goal of the landmarks (Interest Points) tracking is to calculate actual displacements between the CamRobII positions (calibration update) on the basis of tracking information for each subsequent image frame. The calibration update results have been compared with the actual CamRobII coordinates.

The image sequence used for the tracking and calibration update is composed with the 10 *basic stereo pairs* (20 frames) and 4 intermediate left image frames between each of them (40 frames). A relative calibration procedure based on stereo matching [3] is used to calculate the coordinates of the right camera for each *basic stereo pair*. A calibration update procedure [4] based on the Interest Points (landmarks) tracking is used to maintain the coordinates of the 4 intermediate frames known. Each forth intermediate frame composes the left image for the next *basic stereo pair* starting therefore the next calibration loop.

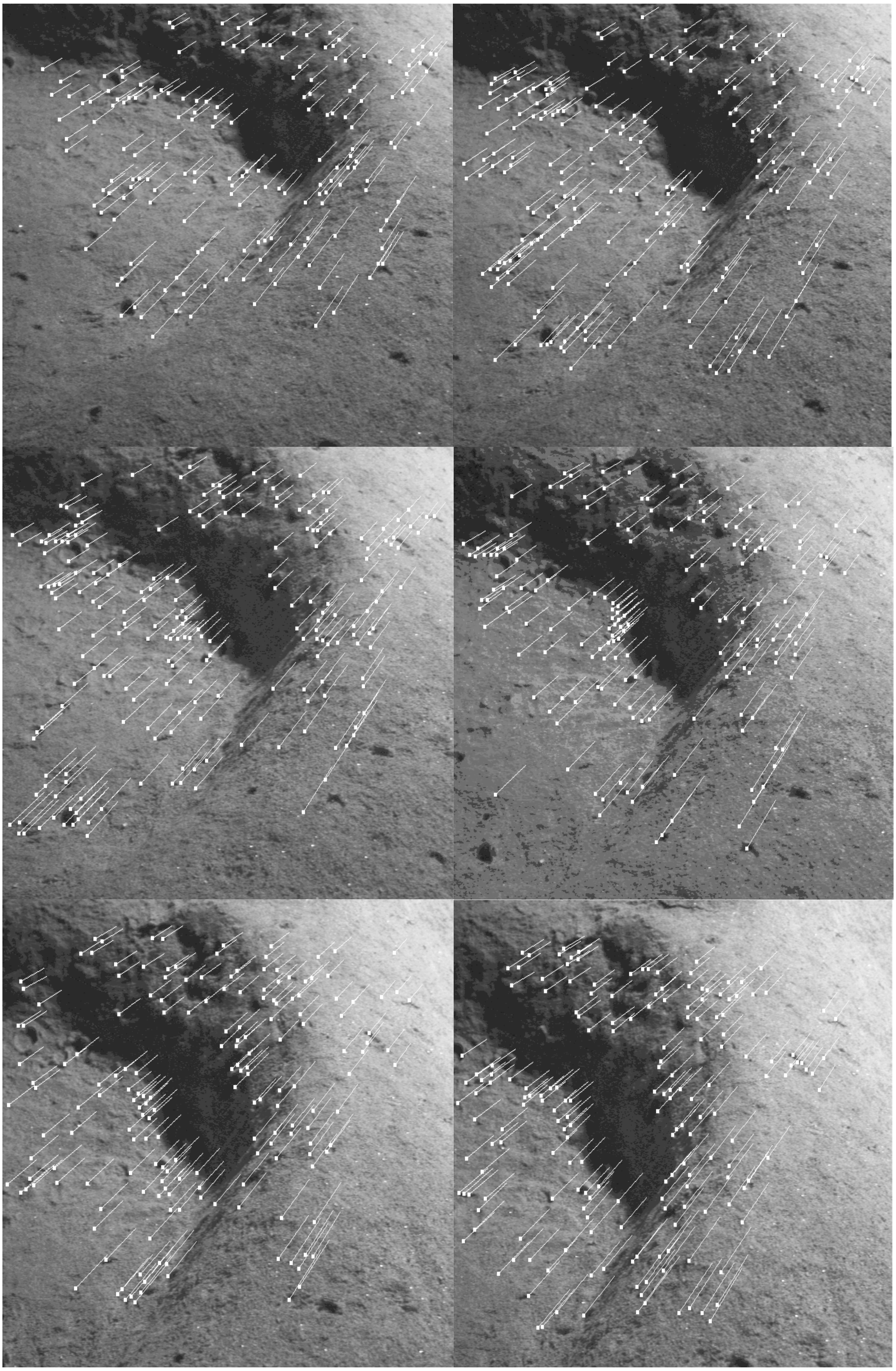

Figure 5. Six consecutive image frames and landmark tracking paths.

The actual CamRobII trajectory used for the stereo sequence acquisition is a straight line. An offset between subsequent odd and even frames is 15 mm in ***X*** direction, and is 30 mm in ***Y*** direction (Figure 2). The stereo basis (SB) for the *basic stereo pairs* is equal to 97 mm. An example of the tracking paths between the corresponding Interest Points on the 6 subsequent even frames are shown on Figure 5. The major parameters during Interest Points tracking are the number of landmarks and their back projection error as a measure of the calibration consistency. The most inconsistent Interest Points are rejected and are not used for calibration. Experiments showed that the optimum number of landmarks necessary for the reliable calibration is at about 200 points, however, a smaller set of landmarks (>200) still leads to robust tracking.

Figure 6 displays the trajectory calculated on the basis of tracking and calibration update for the cameras. The coordinates of the positions which composed the trajectory are presented in Table 1. They show that discrepancy between CamRobII coordinates and tracking positions have not been accumulated along the path. The fact that Y offset values are always above 30 mm is explained by the uncertainty in the scale factor chosen with the estimated stereo baseline.

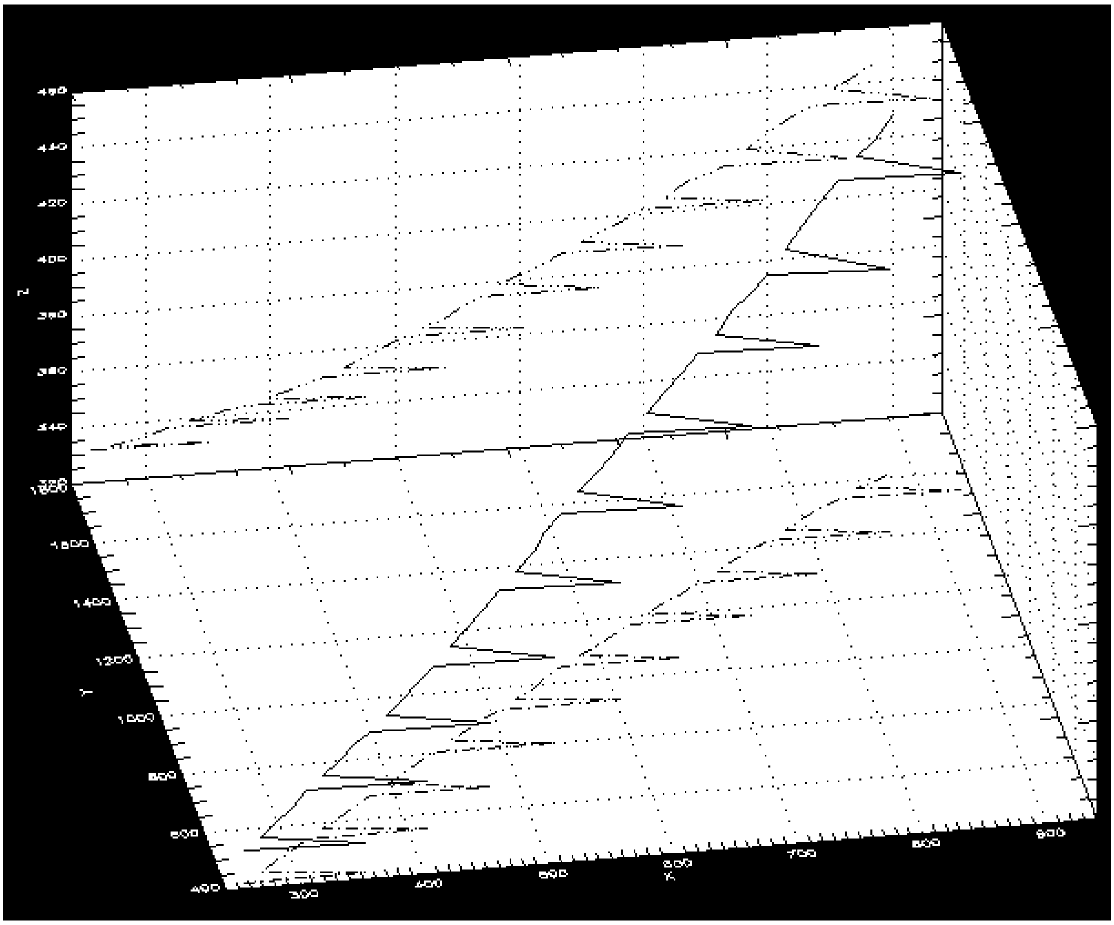

Figure 6. Camera trajectory as calculated on the basis of landmarks tracking.

4 CONCLUSIONS

A closed-loop vision-based algorithmic solution for autonomous rover navigation mode is presented here. The solution integrates the following algorithmic components into the calculating chain:

- Consecutive image stereo data acquisition.
- 3D Digital Elevation Model reconstruction of a terrain.
- Risk map generation.
- Local path planning.
- Landmarks tacking.
- Update rover position (calibration update).

Frame	X	Y	Z	X offset	Y offset	SB
1	256.165	417.157	331.636			
2	353.161	417.710	330.965			96.996
3	271.647	449.668	332.519	15.482	31.957	
5	286.810	482.161	333.775	15.163	32.493	
7	302.299	514.559	335.101	15.488	32.398	
9	317.487	547.343	337.435	15.187	32.784	
10	414.485	547.211	336.911			96.998
11	333.371	579.563	339.031	15.884	32.220	
13	349.014	612.294	339.976	15.643	32.730	
15	364.387	644.058	341.940	15.373	31.764	
17	380.025	676.150	342.851	15.638	32.091	
18	477.022	675.418	342.566			96.996
19	396.148	708.317	344.676	16.123	32.167	
21	411.807	740.070	346.625	15.658	31.753	
23	427.618	772.065	348.582	15.811	31.995	
25	443.073	803.346	350.698	15.454	31.281	
26	540.062	801.939	350.407			96.989
27	458.708	834.988	353.727	15.635	31.642	
29	474.679	866.187	356.595	15.970	31.199	
31	490.766	897.497	359.578	16.087	31.309	
33	506.651	928.671	362.590	15.884	31.174	
34	603.628	926.575	362.290			96.976
35	523.134	959.598	365.552	16.483	30.927	
37	539.555	990.485	368.521	16.420	30.887	
39	551.078	1022.09	371.993	11.523	31.608	
41	567.620	1053.04	374.986	16.542	30.944	
42	664.586	1050.54	374.519			96.966
43	584.219	1084.25	378.625	16.599	31.210	
45	600.645	1115.20	381.423	16.426	30.948	
47	616.792	1146.29	384.380	16.146	31.092	
49	633.345	1177.35	388.005	16.553	31.058	
50	730.297	1174.32	387.529			96.951
51	649.787	1208.06	391.497	16.442	30.710	
53	666.145	1238.78	395.024	16.357	30.724	
55	682.960	1269.23	398.441	16.814	30.444	
57	698.848	1300.17	401.788	15.888	30.939	
58	795.777	1296.49	401.459			96.929
59	716.479	1330.16	404.807	17.631	29.990	
61	733.025	1361.22	409.718	16.545	31.056	
63	749.612	1391.04	411.801	16.587	29.816	
65	766.314	1420.93	414.673	16.701	29.892	
66	863.218	1416.68	413.975			96.904
67	783.056	1451.25	419.980	16.742	30.320	
69	799.582	1481.52	425.081	16.525	30.267	
71	816.465	1511.60	429.738	16.882	30.083	
73	833.774	1541.02	433.723	17.309	29.420	
74	930.657	1536.30	433.065			96.8825
75	850.856	1570.43	438.901	17.082	29.410	
77	867.135	1589.88	441.959	16.278	19.455	
79	884.723	1619.29	446.964	17.588	29.406	

Table 1. Calibration update results on the basis of landmarks tracking. SB: stereo basis calculated.

The algorithmic steps have been simulated with the help of an accurate robot (CamRobII, seven degrees of freedom) placed above a Lunar terrain mockup. Emphasis was put on the critical onboard calculations in providing fully autonomous and robust rover operations.

The following statements can be drawn on the basis of simulation sessions:

- 3D autonomous rover navigation on a Moon-like terrain is feasible.
- The accuracy and robustness of each algorithmic component of the VBN system was shown on a long (50 frames) consecutive image sequence.
- Incredibly long sequence of stable calibration update results without outside intervention demonstrated high performance of the algorithms involved. The results are very important for the accurate local path execution by the rover.
- Assessments for the necessary computation efforts showed that the existing algorithms can be used for on-board implementation. On-ground software can duplicate on-board calculations at a higher level of accuracy to create a virtual reality environment of the planetary surface.

Reflex obstacle detection is the remaining and not simulated aspect in the presented solution. The authors hope on an opportunity to integrate reflex obstacle detection algorithm during real rover operation sessions.

REFERENCES

[1] "LEDA Assessment Study Final report", LEDA-RP-95-02, Revision 0, 1995, June 12.

[2] Ross, B. „A practical Stereo Vision System“, CVPR 93, June 15-18, 1993.

[3] Ulm, M. and Paar. G. "Relative Camera Calibration from Stereo Disparities." In Proc. 3rd ISPRS Conference on Optical 3-D Measurement Techniques, Vienna, October 2-4 1995..

.[4] Paar, G. et al. "Natural Feature Tracking for Autonomous Navigation." In Proc. 28th International Dedicated ISATA Conference on Robotics, Motion and Maching Vision, Stuttgart, October 1995.

[5] Paar,G. and Pölzleitner, W. "Robust Disparity Estimation in Terrain Modeling for Spacecraft Navigation". In Proc. 11th ICPR, 1992.

[6] Bauer, A. and Paar, G. "Stereo Reconstruction from Dense Disparity Maps Using the Locus Method." In Gruen, A. and Kahmen, H., editors, Proc. 2nd Conference on Optical 3-D Measurement Techniques, pp. 460-466, Zurich, Switzerland, October 4-7 1993. ETH Zurich, Wichmann Verlag.

[7] Proy, C., et al. "Improving autonomy of Marsokhod 96". 44th Cong. of the Int. Astronautical Federation. *16-22 Oct. 1993. Graz, Austria.*

[8] Dijkstra, E.W. (1959). A note on two problems in connection with graphs. Numer. Math.*, 1, 269-271.*

[9]. Kolesnik, M. “Vision and Navigation of Marsokhod Rover”. ACCV´95, 5-8 Dec. 1995, pp.III-772 - III-777.

[10] Pölzleitner, W. and Ulm, M. "Robust dynamic 3D motion estimation using landmarks". In Optical Tools for Manufacturing and Advanced Automation. Videometrics II, 1993.