Pose Estimation and Map Building with a PMD-Camera for Robot Navigation

A. Prusak, O. Melnychuk, H. Roth

Institute of Automatic Control Engineering, University of Siegen, 57068 Siegen, Germany

I. Schiller, R. Koch

Institute of Computer Science, Christian-Albrechts-University of Kiel, 24098 Kiel, Germany

Abstract: In this paper we describe a joint approach for robot navigation with collision avoidance, pose estimation and map building with a 2.5D PMD (Photonic Mixer Device)-Camera combined with a high-resolution spherical camera. The cameras are mounted at the front of the robot with a certain inclination angle. The navigation and map building consists of two steps: When entering new terrain the robot first scans the surrounding. Simultaneously a 3D-panorama is generated from the PMD-images. In the second step the robot moves along the predefined path, using the PMD-camera for collision avoidance and a combined Structure-from-Motion (SfM) and model-tracking approach for self-localization. The computed poses of the robot are simultaneously used for map building with new measurements from the PMD-camera.

Keywords: robot navigation, self localization, pose estimation, Structure-from-motion, SLAM, 3D-map

1 Introduction

For the realization of dynamic tasks in environment and time, autonomous mobile systems need sequential and fast knowledge of the spatial surroundings. Only on the basis of the fast technical 3D viewing can these systems adapt autonomously and intelligently to unknown surroundings in real-time. Such fast 3D environment capturing concerns in particular self localization and scene modeling. Pose estimation (self-localization) and environment modeling are the base for numerous applications in computer vision and robotics. For an exact pose and an exact environment model, the 2D-CCD-camera should be used as well as the PMD-camera [13]. With this combination of cameras 2D-gray-scale images and 3D images, containing depth data of the environment are available. From this 2D-3D image stream the pose of the moving PMD-camera, respectively the moving robot, can be determined continuously and in real-time. In parallel to that, relevant objects inside the snapshots of the PMD camera can be analyzed. These objects are converted into a suitable format and merged into a so-called virtual environment map, for navigation.

2 Previous Work

2.1 Structure-from-motion (SfM) with PMD-Camera

The PMD-Camera has recently gained attention in the computer vision research. Combining a high resolution camera with a 3D time-of-flight sensor is based in the work of Prasad et al. [1]. The combined approach of using a PMD-camera and a standard 2D CCD-camera for pose estimation with 3D-Range-Data [4] showed that the use of a 3D-Range-Camera can overcome many of the restrictions standard SfM approaches suffer from, like scale ambiguity and need for lateral movement. In [6] a model of a scene is generated in an offline phase. This model is used to support the pose estimation in the online phase. The approach presented here differs from [6] in that the 3D-model is computed online during the capture phase of the robot.

2.2 Map Building and Navigation

Map building with a mobile robots is regarded one of the fundamental problems in mobile robotics. Typical solutions rely upon the integration of the information coming from robot sensors with the a priory knowledge of the surrounding environment. Most previous work used depth data, obtained from 3D laser scans or from odometry. In the literature, the mobile robot mapping problem is often referred to as the simultaneous localization and mapping problem (SLAM) [7]. In the approach of [8] the position tracking of a robot is achieved by comparing the actual laser information of the robot with the precomputed information about the environment. This is done by storing the angle histograms obtained from laser range scans recorded at various locations in the environment. By using the odometry data of the robot and maximizing the correlation between the stored histograms and the laser range scans recorded by the robot while moving, the position and orientation of the robot can be estimated. A similar method is used in [9,10], where hill-climbing is applied to match local maps constructed from sonar information with the global occupancy grid map. In [12] the trimmed iterative closest point algorithm is applied in order to estimate the planar displacement of the robot by matching dense two-dimensional range scans.

3 Mobile Vehicle Tom3D [14-16]

A mobile vehicle should be able to navigate and explore autonomous within a building. Wheel encoders are used for the velocity control and for rough pose estimation. For accurate pose estimation of the mobile vehicle a combined SfM-/modeltracking approach is used. Obstacle recognition and driveway recognition can easily be done with the same images. At the University of Siegen the so-called mobile robot Tom3D (Tele Operated Machine with 3D PMD-Camera) was developed (Fig. 1). It is build with a differential-drive, which has great advantages compared to the Ackermann drive. The mobile vehicle is equipped with an Embedded PC for image processing and communication, as well as a microcontroller C167 for velocity and path control. The access of the operator to the mobile vehicle is possible via Wireless-LAN, Access-Point, LAN and Internet, or is realized directly as an Ad-hoc-Wireless-LAN connection.

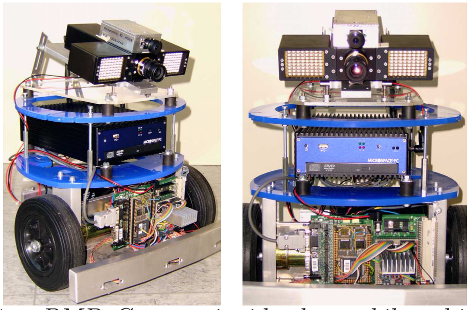

Figure 1 Mobile Robot Tom3D, equipped with 2D-CCD-camera, PMD-Camera, Embedded PC and Microcontroller.

3.1 PMD-Camera inside the mobile vehicle

The PMD camera has an ambiguity interval of 7.5 meters. In order not to exceed the maximum range, the PMD camera must be fixed with a suitable inclination angle on the mobile vehicle. Furthermore the PMD camera must be able to see the driveway in particular in the front area of the mobile vehicle. Altogether the optimum inclination angle is depending on the non-ambiguous interval, optical fov, height of assembly of the PMD camera and the location of the visible front area of the mobile vehicle.

3.2 Obstacle and driveway detection

At first it must be distinguished between so-called negative and positive obstacles. Negative obstacles are deepenings, holes, depressions, etc. within the driveway. However, positive obstacles show rises, knoll, chattels, objects, etc. which lie on the driveway.

Due to of the inclination angle, the PMD camera captures obstacles and driveway of the visible front area of the mobile vehicle in one video image. Therefore it can not be distinguished between obstacles and driveway (Fig. 2) easily. A separation of driveway, negative and positive obstacles must be done: Thereto the PMD image is rotated around its horizontal axis with above described inclination angle. Now the driveway within the PMD image is reduced to a horizontal balk. This balk is always located at the same place within the PMD image, if the mobile vehicle drives through a planar area, i.e. pitch angle and roll angle of the PMD camera respectively the mobile vehicle are equal to zero. If the Tom3D navigates in an outdoor area its pitch and roll angles are unequal to zero and will be measured by the inclination sensor.

3.3 Obstacle avoidance and navigation

The mobile vehicle should avoid positive and negative obstacles. Thereto the above processed PMD image (rotation with inclination angle and if necessary additional rotation with pitch angle around PMD image's horizontal axis as well as

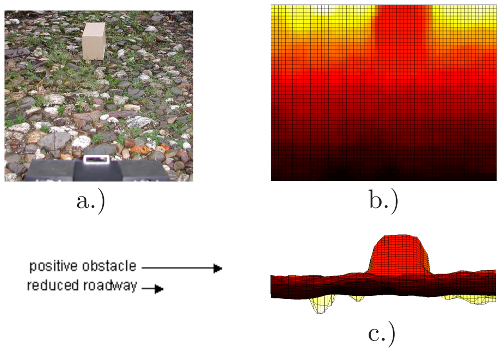

Figure 2 Fast obstacle and driveway recognition:
a.) robot view of the outdoor driveway with a obstacle,
b.) it's PMD video image and
c.) PMD video image after rotation operation with PMD-camera's inclination angle

rotation with roll angle around PMD image's optical central axis) will be divided into eight sectors. Within every sector the high level robot control will determine the respective minimum, followed by the calculation which sector owns the smallest distance. This sector number and its distance value are used as a linguistic input variable for the Fuzzy Logic controller. After this so-called Fuzzification follows the inferencing, in which the manoeuvres for the obstacle avoidance are formulated in linguistic form. Finally within the Defuzzification the results of inferencing are converted in concrete motor control signals for the mobile vehicle.

PMD-images contain a significant amount of noise. A complete control loop therefore consists of PMD-image capturing, image filtering (median filter), obstacle and driveway recognition, obstacle avoidance and engine control. With this continuously working loop, a navigation respectively exploration without environment map and without path control can be realized. In this state the mobile vehicle can only be used for uncontrolled driving or hazard driving. While the mobile vehicle passes through the environment, it is guided by obstacles. If an obstacle is detected a defined direction change is done.

4 Self localization with Structure-from-Motion (SfM)

In this section a method for self-localization with a SfM algorithm is described. Our approach for online SfM on a mobile platform uses the available information of a 2D-CCD camera and a 3D-PMD camera. SfM approaches use 2D features in an image stream which are tracked over sequences. In this work we start the tracking by assuming that nothing is known about the current environment. The tracking itself consists of two steps. Environment scanning with 3D-panorama creation and visual tracking with 3D-panorama.

4.1 Mapping between PMD-Image and Fisheye-Image

By calibrating the two cameras we aquire the relative transform of the two cameras. The calibration is done using the method described in [19]. The PMD-image, with the estimated external and internal camera parameters, is mapped to the 2D spherical image. This results in a pixel to pixel mapping between spherical 2D- and perspective 3D-camera. In fig. 3 the mapping results are demonstrated. On the left a depth-map from the PMD-camera is visualized and in the center the PMD-

image is mapped to the spherical image. To construct the 3D-panorama a set of 2D-features $x_{i,0} i = 1, ..., n$ is detected in the 2D-image and a set of 3D-points X_i is generated from the depth-information in the 3D-image. The 2D interest point detector and the KLT-Tracker, which are used to track the 2D interest points in subsequent images, are similar to [2]. From these 2D-3D-correspondences an initial pose is estimated. However the field-of-view (fov) covered by the PMD-camera is small compared to the fov of the spherical camera. This results in only very few 2D-points in the fov of the PMD-camera and very few 2D-3D-correspondences which leads to unstable pose estimation.
Our solution to this problem is to generate a 3D-depth-panorama from many PMD-images. Therefore the robot rotates around it's axis to the left and right when entering new terrain. The detected 2D-intensity corners $x_{i,0}$ in the spherical image are tracked in following images $x_{i,t}$. The homography $x_{i,0}H = x_{i,t}$ between subsequent images is estimated numerically from the 2D-correspondences, using Levenberg-Marquardt. Obviously when estimating homographies the translation of the camera center is neglected and the assumption is made that the camera rotates around it's center.

Figure 3 PMD-Image, PMD-Image projected to spherical image, spherical image with overlaid depth-panorama

With the estimated homography between the spherical images a 3D-panorama is generated. Therefore all pixels from the PMD-images are projected into a cylindric camera with the same center. An occupation map is computed in the cylindric camera and if more than one measurement is available per pixel the mean is used. From the occupation map a 3D-triangle mesh is generated with the projection parameters of the spherical camera. This 3D-model can now be used to support the visual tracking and be seen as an extension of the depth-image of the PMD-camera. The extended depth-map of the 3D-panorama is shown in fig. 3 on the right, superimposed in the spherical image of the CCD-camera. Fig. 4 shows the estimated 3D-model on the left with the initial camera pose.

4.2 Pose-Estimation with 3D-Panorama

The 3D-panorama model is now used in a combined model-tracking/SfM approach. The robot now drives along the planned path in the unknown terrain. The last estimated pose of the homography estimation is used as the starting pose for further pose estimation. The generated 3D-model is rendered in an offscreen buffer with this initial pose. In the spherical image 2D-interest points are detected and 3D-points are generated from the 2D-interest points in the region of the rendered depth-map. This results in a set of 2D-3D correspondences which are used for pose estimation. In subsequent images the 2D-interest points are tracked and the pose is estimated. Additionally 2D-points which are not in the region of the rendered depth map are triangulated and new 2D-3D-correspondences are generated.

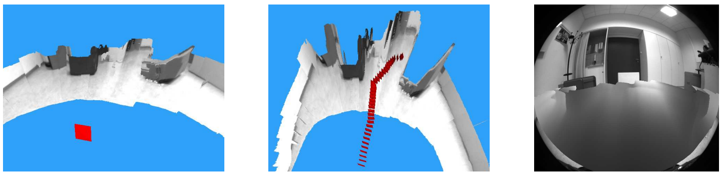

Figure 4 3D-model estimated from homography and initial camera,3D-panorama with full estimated robot path, spherical image with overlaid depth-map from 3D-panorama at frame 127.

In the center of figure 4 the 3D-model is displayed with the estimated camera poses as a path through the 3D-model. The spherical image of the CCD-camera at frame 127 of the tracking is demonstrated together with the rendered depth-map superimposed in the right image. The calculated robot poses from the combined SfM and model-tracking are subsequently used in the map-building and navigation of the Tom3D.

5 Map Building

Pose estimation and map building can not be viewed as two separable tasks, due to the fact that mapping includes both, estimating the position of the robot relative to a map and generating a map using the sensory input and the estimated robot poses. The 3D-points from the 3D-panorama are used, together with the estimated poses, as an input data for the map building. First of all the algorithm converts the PMD-images from spherical into cartesian coordinates.

Due to noise in the PMD-images and the approximation during the creation of the 3D-panorama an error in the estimated poses of the SfM-algorithm may be present. In oder to correct the inaccuracy of the poses, the PMD-images are matched with Trimmed Iterative Closest Point (TrICP) Algorithm. The structure of TrICP is similar to conventional Iterative Closest Point Algorithm (ICP) but it is more applicable to erroneous measurements and shape defects. The main idea is to use the least trimmed squares approach in all phases of the operation. The pose from SfM is used as an initial condition for the iterations. The map is built incrementally by adding the new parts with evaluated pose.

6 Results

The algorithm for map building is implemented on a Tom3D mobile robot equipped with a PMD-camera 3k-S with a resolution of 64×48 pixels. The PMD camera is mounted on the robot with an inclination angle of 38 degrees. The approximation of the movement of the robot in the first phase (turning to left and right) as a pure rotation introduces some errors in the creation of the 3D-panorama. Initially the mean error in z-direction between the rendered panorama and the PMD-depth-map is < 20cm at an object distance of 2m. The pose estimation on the generated 3D-panorama runs in real-time (appr. 7 fps) and is well suited for real time pose estimation and map-building. However the pose estimation using the 3D-panorama lacks absolute accuracy due to noise in the PMD-images and incompleteness of the panorama at occlusions. The map is stored as a 2.5D-grid model with a resolution of 5mm. The map size is 1500×400 elements. Figure 5 shows an incremental building of the environment map step by step. The completed grid based map is capable to be used for the navigation of the mobile robot for obstacles avoidance.

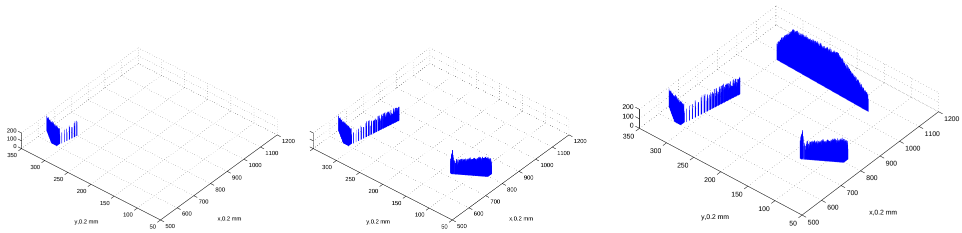

Figure 5 Incremental building of a 3D-map while moving the robot with camera.

7 Conclusion

We presented a combined approach for online mobile pose estimation and map-building with a PMD-camera. By combining a spherical CCD camera with a PMD-camera we exploit the advantages of the big fov of the spherical camera for the feature tracking and pose estimation and the absolute scale of the PMD-camera. The creation of the 3D-panorama by homography estimation extends the small fov of the PMD-camera and initially more 2D-3D-correspondences can be generated which results in a more stable pose estimation. The estimated camera poses can directly be used for online map-building. The map was built in to stages: first the robot pose is estimated using the TrICP Algorithm. The pose data from SfM is used as initialization for the iterations. In the second step the map is build using the incremental mapping approach. The resulting algorithm is capable to calculate 3D maps in real-time (up to 7 fps) and handles noise in PMD data.

Acknowledgments

This work was supported by the German Research Foundation (DFG) as part of the “3DPoseMap” [17] project (KO-2044/3-1 and RO-2384/1-1). Thanks to the DFG [18] for initiating and financing this project.

References and Notes

1 Prasad, T.D.A. and Hartmann, K. and Wolfgang, W. and Ghobadi, S.E. and Sluiter, A. (2006) 'First steps in enhancing 3D vision technique using 2D/3D sensors', University of Siegen,*Computer Vision Winter Workshop 2006, Telc, Czech Republic*, Czech Society for Cybernetics and Informatics, pages 82-86

2 J. Shi and C. Tomasi (1994), 'Good features to track', pp 593-600 , *Conference on Computer Vision and Pattern Recognition*

3 R. Koch and K. Koeser and B. Streckel and J.-F. Evers-Senne (2005), 'Markerless Image-based 3D Tracking for Real-time Augmented Reality Applications', *WIAMIS 2005*,Montreux, Switzerland

4 Streckel, B. and Bartczak, B. and Koch, R. and Kolb, A. (2007), 'Supporting Structure From Motion with a 3D-Range-Camera', *Scandinavian Conf. Image Analysis (SCIA)*

5 Lindner, M.; Kolb, A. (2006), 'Lateral and Depth Calibration of PMD-Distance Sensors', *International Symposium on Visual Computing (ISVC06), Advances in Visual Computing*, Lake Tahoe, Nevada, Nov 6-8, 2006

6 K. Koeser and B. Bartczak and R. Koch (2006), 'Drift-free Pose Estimation with Hemispherical Cameras', *Proceedings of CVMP 2006*, London

7 J.A. Castellnos, J.M.M. Montiel, J. Neira, and J.D. Tardos (1999), 'The SPmap: A probabilistic framework for simultaneous localization and map building.' *IEEE Transactions on robotics and Automation*, 15(5):948-953, 1999

8 G. Weiss, C. Wetzler, and E. von Puttkamer. (1994). 'Keeping track of position and orientation of moving indoor systems by correlation of range finder scans.' In *Proc. of the International conference on Intelligent Robots and Systems*, pages 595-601, 1994.

9 B. Yamauchi. (1996) 'Mobile robot localization in dynamic environments using dead reckoning and evidence grids.' In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 1996.

10 A.C. Schultz, W. Adams, and B. Yamauchi.(1999) 'Integrating exploration, localization, navigation and planning with a common representation.' *Auton. Robots,* 6(3):293-308, 1999.

11 D. Chetverikov, D. Svirko, D. Stepanov, and P. Krsek (2002) 'The trimmed iterative closest point algorithm.' *16th International conference on pattern recognition. ICPR 2002. Quebec* vol 3:545-548, 2002

12 J. Minguez, F. Lamiraux, L. Montesano.(2005) 'Metric-Based Scan Matching Algorithms for Mobile Robot Displacement Estimation' *Pr. of the 2005 IEEE International Conference on Robotics and Automation*, Barcelona, Spain, April 2005

13 Lange, Robert: 3D time-of-flight distance measurement with custom solid-state image sensors in CMOS, CCD technology. University of Siegen, Dissertation, 2000.

14 Ruangpayoongsak, N.; Roth, H.; Schwarte, R: 3D artificial landmark design and recognition for mobile robots. Mechatronics 2006. 4th IFAC Symposium on Mechatronic Systems. Preprints, Heidelberg, Germany, 12-14 Sept. 2006.

15 Alexander Prusak, Hubert Roth, Rudolf Schwarte: Application of 3D-PMD Video Cameras for Tasks in the Autonomous Mobile Robotics. 16th IFAC World Congress, Prague, Czech Republic, from July 4 to July 8, 2005.

16 Niramon Ruangpayoongsak, Hubert Roth, Rudolf Schwarte: Position Calibration of a Mobile Robot Based on 3D Vision. 16th IFAC World Congress, Prague, Czech Republic, from July 4 to July 8, 2005

17 DFG-Project "Dyn3D" (Dynamic 3D-view) with the Subproject "3DPoseMap" (3D pose estimation and 3D mapping with help of PMD-Camera), www.zess.uni-siegen.de/cms/front_content.php?idcat=127

18 DFG, Deutsche Forschungsgemeinschaft (German Research Foundation), www.dfg.de

19 Beder, C. and Koch, R. Calibration of focal length and 3D pose based on the reflectance and depth image of a planar object. Int. J. on Intell. Systems and Techn. and App., Workshop on Dynamic 3D Imaging in Conjunction with DAGM 2007, Heidelberg.