3D-CAMERA BASED NAVIGATION OF A MOBILE ROBOT IN AN AGRICULTURAL ENVIRONMENT

Marc Schulze

Institute of Photogrammetry and Remote Sensing
University of Technology Dresden
01069 Dresden, Germany
marc.schulze@tu-dresden.de

Commission V - Student Consortium

KEY WORDS: 3D-camera, navigation, mobile robot, agricultural environment, voxel data

ABSTRACT:

The paper describes experiments performed in an ongoing research project on field robots accomplished with a simple interface controlled track vehicle equipped with a 204x204 pixel 25 Hz range camera. The goal is to use the 3D-camera data to support the navigation and collision avoidance of the vehicle in agricultural applications. The paper concentrates on the generation of plant canopy density maps and the removal of outliers in a filtering step applying 3D morphology techniques.

1 INTRODUCTION

Mapping their environment, recognizing obstacles and moving around like human beings is a high goal for autonomous mobile robot navigation. The control of a robot with different sensors is a current research field with high potential. One of the most powerful sensors for control and orientation of robots are cameras. This holds especially for modern range cameras acquiring images at a high temporal resolution and getting range information for each pixel simultaneously.
In agricultural applications, a mobile robot may have the task to navigate through rows of plants, for instance in order to spray, water or harvest them. Due to the rough and often slippery ground, wheel encoders are not very helpful in these tasks. Here, a range camera, possibly combined with GPS and compass devices, may be a rather valuable tool to generate a 3D map of the robots environment and to detect open space to navigate. Simultaneously, it may also deliver information on plant structure, growth and harvesting.

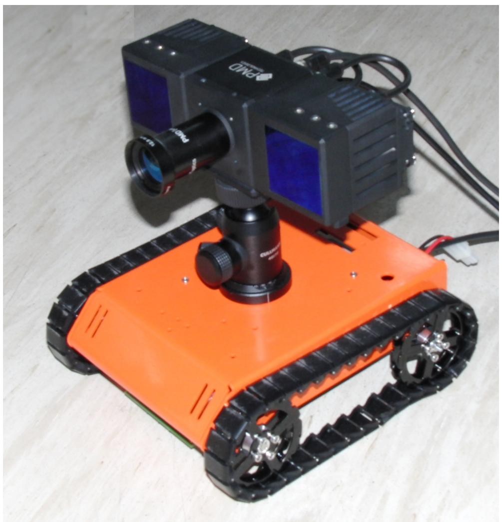


Figure 1: Used mobile robot with PMD CamCube 2.0

Parameter	Value
Camara Type	PMD CamCube 2.0
Measurement Range	0.3 to 7 m
Repeatability (1σ)	<3 mm
Frame Rate (3D)	25 fps
Illumination Wavelength	870 nm
Sensor Size	204**x**204 pixel

Table 1: datasheet PMD CamCube 2.0 (PMD, 2010)

State Of The Art

Agricultural applications with mobile robots and 3D-cameras are still largly unexplored. Unlike for instance in urban environments regular geometries are rare in an agricultural environment (Cole and Newman, 2006) (Weingarten, 2006). This means for example that connection of point clouds will often has to be realized by single points (Wang et al., 2009). The calculated transformation parameters between successive 3D-camera point clouds give information of the current position of the robot. This mapping and localization algorithm is called **S**imultaneous **L**ocalization **A**nd **M**apping or **C**urrent **M**apping and **L**ocalization.
At the end of a chain of transformations the whole high density point cloud gives a 3D-representation of the agricultural environment. Based on this data, it is also possible to determine plant parameters such as height, density and volume (Hosoi and Omasa, 2009). Moreover the mapped point can used to navigate in an autonomous watering process.

2 BASICS OF SENSOR AND DATA

The data used in this study were acquired by a PMD-CamCube 2.0 (Tab. 1) mounted on a mobile robot driving.

Basics

The advantages of 3D-cameras against stereo systems and laser scanners can be seen in the non-sequential data acquisition mode, their high temporal resolution and the fact that intensity and range images are recorded simultaneously.

$$I = \frac{m_1 + m_2 + m_3 + m_4}{4} \tag{1}$$

$$A = \frac{\sqrt{(m_3 - m_1)^2 + (m_4 - m_2)^2}}{2} \quad (2)$$

$$\varphi = \arctan \frac{m_4 - m_2}{m_1 - m_3} \quad (3)$$

$$D = \frac{\lambda_{max}}{2} \cdot \frac{\varphi}{2\pi} \quad (4)$$

On the other hand measurement field is limit by wavelength because of an ambiguity problem. Therefore the maximum visible range is $\lambda/2$. The accuracy of the distance for each pixel depends on the measurement condition (e.g. background illumination) and object properties(e.g. color and surface) (Kahlmann et al., 2006) (Westfeld, 2007). Improvements of the data recording technology and calibration of sensor may increase the quality and accuracy (Kahlmann, 2007) (Westfeld, 2007).
High precise estimated camera constant c, position of principle point (x_h, y_h) and distortion parameters can recover the 3D points accurate by 2D data (Fig. 2).

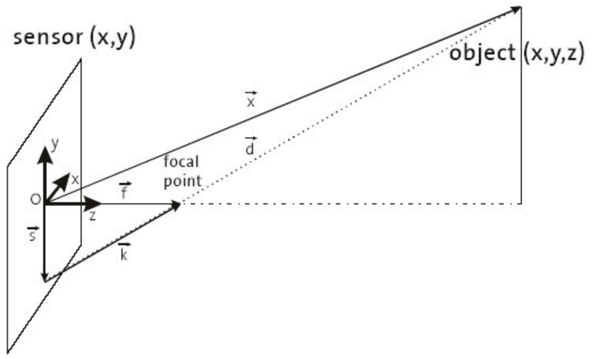


Figure 2: ray course for recovering object point (Kahlmann, 2007)

$$\vec{x} = \vec{s} + D \cdot \|\vec{k}\| = \vec{s} + \vec{d} \quad (5)$$

$$\vec{x} = \begin{pmatrix} x' \\ y' \\ 0 \end{pmatrix} + \begin{pmatrix} -x' \\ -y' \\ f \end{pmatrix} \cdot \frac{D}{\sqrt{x'^2 + y'^2 + f^2}} \quad (6)$$

3 WORKFLOW

This section describes the process of acquiring images for successful transformation of point cloud and pose estimation of mobile robot (Fig. 3). Most popular algorithms for robust and fast feature extraction are **S**cale **I**nvariant **F**eature **T**ransform (SIFT) (Lowe, 1999) and **S**peed-**U**p **R**obust **F**eatures (SURF) (Bay et al., 2008). Both are scale and rotation invariant and help to detect feature points in two following images. This is a fast and efficient method to find corresponding points and determine parameters of transformation. For robust outlier detection of transformation parameters a **RAN**dom **SA**mple **C**onsensus (RANSAC) (Fischler and Bolles, 1981) is used. Valid transformation parameters estimated by RANSAC include the most inliers within a small threshold. After robust outlier test a fine fit named **M**ethod of **L**east **S**quares (MLS) follows. The calculated parameters of transformation represents values to connect local point cloud to global system and pose estimation of mobile robot.
After transforming points into global system the existing space will divide in same sized cubes called volume elements (voxel). Voxel are the 3D opposite of area based picture elements (pixel). The number of voxels depends on size of recording object points and voxel size. It is possible to filter voxel data with morphological operators to increase quality and quantity of data like in images. At the end it is possible to sum up all voxels and calculate the volume of plant canopy in object space.

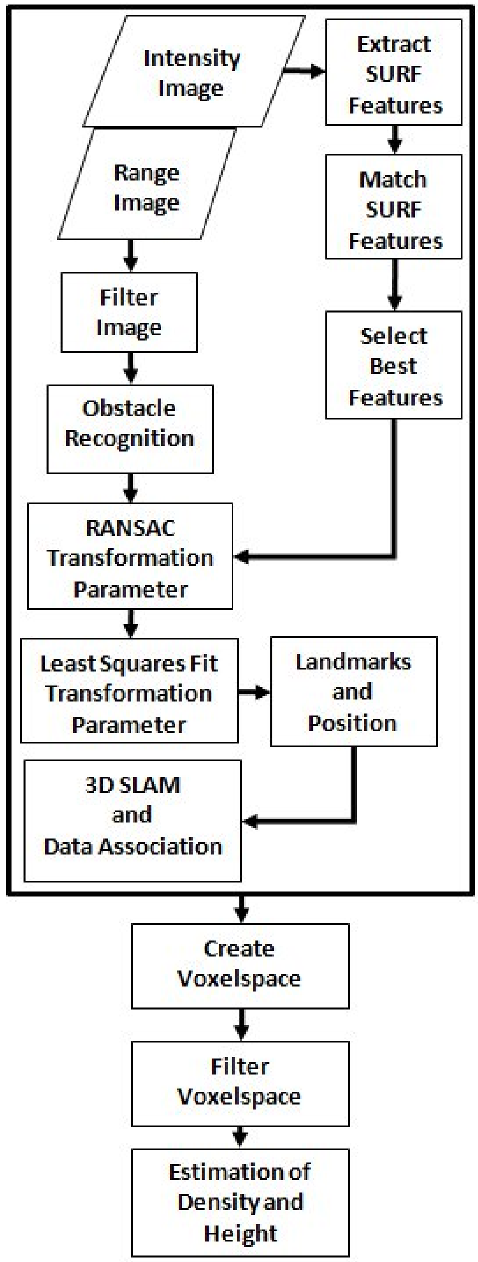


Figure 3: Flowchart of processing (modify of (Wang et al., 2009))

3.1 Intensity Image

The PMD CamCube 2.0 consists of a 2D image sensor (204**x**204 pixel) and two light sources, which emit illumination in near infrared (NIR). To determine the intensity value of each pixel the modulated sinusoidal signal is sampled four times (Chapter 2) and an average value is calculated (Eq. 1). In spite of small resolution of sensor some feature extraction methods deliver best results. That is important to find corresponding points and get transformation parameters.

3.2 SURF Feature Extraction

There is a large number of scale and rotation invariant feature extraction methods to find corresponding point between images.

SIFT (Lowe, 1999) and SURF (Bay et al., 2008) are the most popular and effective methods. The main advantage of these algorithms are to determine feature descriptors that allow quick comparisons with other features. (Wang et al., 2009) compare SIFT and SURF and discover that SURF detect more features in small sensor configuration. To detect some outliers it is useful to choose a threshold for matching the feature points. The matching results give information about reliability of matching. Small values could be an advice for bad or wrong matching. Inside SURF it is useful to choose a high threshold (Fig. 4) with the result of fewer points but better reliability.

Figure 4: Corresponding point with SURF in two following images

3.3 Range Image

Distance measurement of 3D-cameras based on phase shift method (chapter 2). Modulate light in NIR is send with maximum wavelength of 14m. Calculation of distances will be realized by comparison of start and end phase. Results are saved in a 16-bit raw image. Distance values from 0 to 7m were associated with values from 0 to 65536 (16-bit). So the theoretical distance resolution is ca. 0.1mm.

3.4 Range Image Filter

Range data of short integration time are subject of strong noise but allow acquire images while fast movements. Reasons are multi path effects and different behavior of reflection by surfaces (Guðmundsson, 2006), (Gut, 2004). The additional usage of amplitude values can help to minimize noise. Therefore pixels with unusually high or low amplitudes were eliminated. These pixels could be influenced by multi path or some other errors.
Smoothing data with median is another kind of outlier elimination. The advantage of median filter is independence of single outliers.
Avoid obstacles close to the robot is one of the main achievements. Therefore it is important to map objects and analyze their position. At first separating background and close-range objects with static background filter of 1m. Only objects with distances under 1m are important for obstacle recognition.

3.5 Obstacle Recognition

After filtering close-range information and background an algorithm have to find drive line between plants or obstacles. In this situation a constraint is given. On left and right side of image are plants so it is necessary to divide the point cloud into two clusters. A fast algorithm which does that is k-mean filter. It needs only the number of expected clusters and divide point cloud by itself (Fig. 5). k-mean is a fast algorithm for clustering and segmentation but works not always correct. If there are found two clusters a bounding box deliver height and width of obstacle. If one of the two boxes cut a predefined rectangle, it will follow a command to control in another direction.

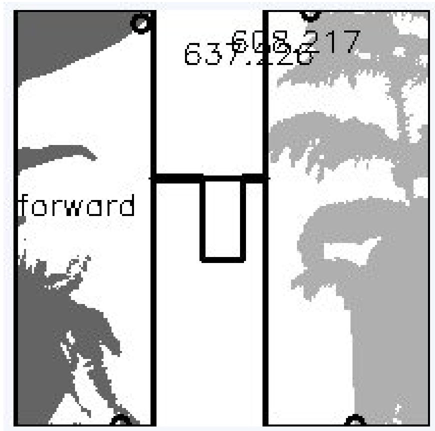


Figure 5: Clustering of point cloud

3.6 Transformation

Base of further 3D data processing is the transformation in a global coordinate system. The origin of global reference system is focal point of first camera position. Other viewpoints will be transformed. The determination of parameters includes two parts. On the one hand a 3-DOF RANSAC is used and on the other hand a 6-DOF least squares fitting. With SURF (chapter 3.2) detected corresponding points in two following images will be expanded to 3D data with information from range image. For a unique solution of 6-DOF equation three points are needed. Usually there are more than three corresponding points. So a graduation between all parameters is use to get the best result.
At first a robust 3-DOF (eq. 7) RANSAC is used to determine translation parameters (X_0, Y_0, Z_0). RANSAC picks randomly one point and calculates the parameters. Parameters that include the most point pairs within a small border were set up as valid translation parameters. Point pairs outside defined border are outliers and have no further influence.
Valid points are used for 6-DOF (eq. 8) fine fit with MLS. MLS needs approximate values for computation. Values for translation are given by RANSAC. Approximation for rotation angles ω,φ,κ were set to zero, because movements especially rotations are small from image to image. Local calculated transformation parameters are sum up and represents global parameters.

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{7}$$

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + \underline{\mathbf{R}} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{8}$$

3.7 Voxel Space

After the transformation into a global reference system it is possible to approximate the density and height of plants. The object space with all 3D points will be estimated by height, width and depth and divided into user-defined voxels. If a voxel do not include points it is probably plant free space. Next step of data processing will be include filter methods of 3D morphological operators (Fig. 6).

4 RESULTS

The developed algorithm is an efficient and fast method to analyze combine data of range and intensity. Created 3D data are

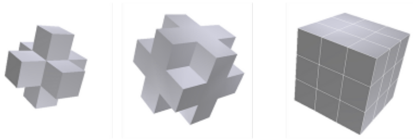


Figure 6: 3D morphologic kernels (Schmidt, 2010)

linked to a global system and so it is possible to analyze whole object space.

4.1 Obstacle Recognition

Amplitude and range image are helpful to analyze the reliability of record data. Outliers with bad amplitudes caused by multi path could be eliminate. The additional background filter of range image supports a higher level of object priority. This combined information are the base of obstacle recognition and robot control. Segmentation with k-mean is a fast method to find clusters of known numbers. K-mean algorithm delivers always a solution, but it is not check for correctness. So it is possible that segmentation and classification are false and also further robot control. The robot will probably crash with obstacle.

4.2 Transformation

The estimation of parameters with RANSAC and MLS represents on the one hand a robust outlier test and on the other hand a fine fit with conclusion of accuracy.
For RANSAC algorithm you have to set iteration number and maximum tolerance to model. For absolute measurement accuracy of CamCube best results were reached with 100 iterations and 1cm tolerance. On average there was found 70 corresponding points and 30 point pairs pass the outlier test. Number of founded feature points depends on contrast and existing intensity variation. After checking corresponding points for validation a fine fit is planned. Approximate values for translation are given by RANSAC and rotation angles were set to zero. Standard deviation of translation is maximum $\sigma X_0 = 1.92mm$, $\sigma Y_0 = 1.95mm$, $\sigma Z_0 = 1.87mm$. So the maximum 3D deviation is $\sigma \vec{X}_{0_{3D}} = 3.3mm$. The standard deviation of rotation could not estimate exactly, because quaternions were used. Quaternions used to avoid singular matrices in MLS. In result rotation angles was small with $< 2°$. Standard deviations are increasing step by step because variance propagation. Therefore following viewpoints own a bigger error ellipse (fig. 7).

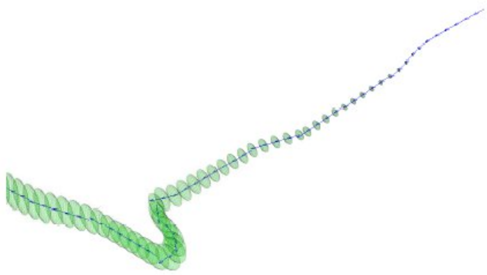
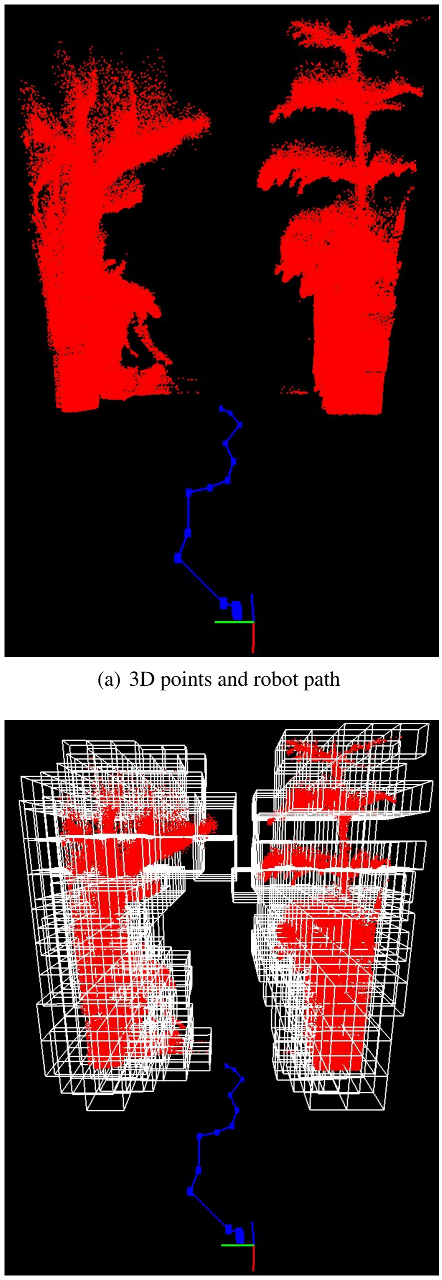


Figure 7: Increasing standard deviation by variance propagation (Cole and Newman, 2006)

4.3 Voxel Space

To divide object space in smaller voxel cells has some advantages for analysis. There are information about position, height and density of plants. This requires to a optimal voxel size. Large voxel cells have the drawback of fewer details of plants and the density approximation will be false. This also have a disadvantage for filtering with morphological operators. If there are too small voxels the time of calculation will be higher.



(a) 3D points and robot path

(b) Voxel space

Figure 8: 3D view of scene

5 CONCLUSIONS AND FUTURE WORK

The descriptive algorithm enables connecting 3D point clouds via 6-DOF transformation and analyzes them. At the moment robot control depending 100% on reliability of k-mean clustering. So if k-mean is wrong the robot probably crashes. So the estimated control has to validate with other controls before. Computation time is also an existing problem. The velocity of robot has to be slow for correct analyzing and controlling. If movements of robot to fast, there is no time to recognize obstacles and avoid collusion. At the beginning of transformation there is a small standard deviation. Caused by variance propagation the transformation and position of robot will be uncertain. A continuously reset with **G**lobal **P**ositioning **S**ystem (GPS) will minimize the error. Next step of work include a detailed voxel analysis.

REFERENCES

Bay, H., Ess, A., Tuytelaars, T. and Gool, L. V., 2008. Speeded-up robust features (surf). Vision and Image Understanding (CVIU) 110(3), pp. 346–359.

Cole, D. and Newman, P., 2006. Using laser range data for 3d slam in outdoor environments. Technical report, Oxford University Robotics Research Group.

Fischler, M. and Bolles, R., 1981. Random sample consensus: A paradigm for model fitting with apphcatlons to image analysis and automated cartography. Communications of the ACM 24, pp. 381–395.

Guðmundsson, S. A., 2006. Robot vision applications using the csem swissranger camera. Technical report, Technical University of Denmark.

Gut, O., 2004. Untersuchungen des 3d-sensors swissranger. Master's thesis, ETH Zurich.

Hosoi, F. and Omasa, K., 2009. Estimating vertical plant area density profile and growth parameters of a wheat canopy at different growth stages using three-dimensional portable lidar imaging. ISPRS Journal of Photogrammetry and Remote Sensing 64, pp. 151–158.

Kahlmann, T., 2007. Range Imaging Metrology: Investigation, Calibration and Development. PhD thesis, ETH Zurich.

Kahlmann, T., Remondino, F. and Ingensand, H., 2006. Calibration for increased accuracy of the range imaging camera swissranger. In: ISPRS Commission V Symposium 'Image Engineering and Vision Metrology'.

Lowe, D., 1999. Object recognition from local scale-invariant features. In: Proceedings of the International Conference on Computer Vision, pp. 1150–1157.

PMD, 2010. Pmd[vision] camcube 2.0. http://www.pmdtec.com/.

Schmidt, A., 2010. Entwicklung von bildanalyseverfahren zur ermittlung von strömungsbeiwerten in waldgebieten auf der basis von voxelrepräsentationen terrestrischer laserscannerdaten. Master's thesis, TU Dresden.

Wang, J. J., Hu, G., Huang, S. and Dissanayake, G., 2009. 3d landmarks extraction from a range imager data for slam. In: Australasian Conference on Robotics and Automation.

Weingarten, J., 2006. Feature-Based 3D SLAM. PhD thesis, Ecole Polytechnique Federale de Lausanne.

Westfeld, P., 2007. Ansätze zur kalibrierung des range-imaging-sensors sr-3000 unter simultaner verwendung von intensitäts- und entferungsbildern. In: Photogrammetrie - Laserscanning - Optische 3D-Messtechnik.