# The estimation of spatial positions by using an omnidirectional camera system

*Maximilian Muffert*<sup>1</sup>, Jan Siegemund<sup>1</sup>, Wolfgang Förstner<sup>1</sup> <sup>1</sup>Intitute of Geodesy and Geoinformation, University of Bonn

#### Abstract

With an omnidirectional camera system, it is possible to take 360°views of the surrounding area at each camera position. These systems are used particularly in robotic applications, in autonomous navigation and supervision technology for ego-motion estimation. In addition to the visual capture of the environment itself, we can compute the parameters of orientation and position from image sequences, i.e. we get three parameters of position and three of orientation (yaw rate, pitch and roll angle) at each time of acquisition. The aim of the presented project is to investigate the quality of the spatial trajectory of a mobile survey vehicle from the recorded image sequences. In this paper, we explain the required photogrammetric background and show the advantages of omnidirectional camera systems for this task. We present the first results on our test set and discuss alternative applications for omnidirectional cameras.

#### Keywords

Omnidirectional camera systems, ego-motion estimation, spatial trajectories, relative orientation, transformations

## INTRODUCTION

The use of multi-sensor systems increases the accuracy and the controllability of position estimation tasks. The multi-sensor systems are often an integration of GPS/INS-systems [GOODALL, C., 2009] or a combination of GPS with gyroscopes and odometers [EICHHORN, A., 2005]. They are often used in precision farming [KUHLMANN, H., SIEMES, M., 2007] or machine control [RETSCHER, G., 2002]. In these applications camera-systems are rarely used, although the parameters of position and orientation can be derived from image sequences and be used for the comparison of accuracy and controllability. Especially in robotic applications ccd-camera systems are used successfully [MOURAGNON, E. et al., 2006].

There are some important advantages using camera systems which should be mentioned: In contrast to GPS, camera systems are more independent and more adaptable because they are as applicable in outdoor as in indoor environments. Issues concerning GPS shadowing or multipath effects are absent. Furthermore, complete signal blockages may happen in kinematic GPS-applications. A new initialisation has to be done, which may take up to five minutes for reasons of ambiguity fixing. During this time, no precise position can be determined. Another disadvantage of GPS which will not arise in the use of camera-systems.

The major advantage of using camera-systems is to get a high level of information at each time of acquisition. A standard acquisition rate between 20Hz and 30Hz [DAVISON, A. J. et al., 2007] and a high image resolution may be given. In addition to the visual capture of the environment itself, we can compute the parameters of orientation and position from image sequences at each time of acquisition with only one sensor. We can interpret the driven trajectory visually. This gives a high level of control and measurement errors can be detected easier.

As a disadvantage it should mentioned that the absolute scale factor of a estimated trajectory can not be defined directly. To obtain absolute position information due to a reference system, we need additional measurements, for example pass-point information. Furthermore, the recorded images need to offer a certain degree of texture which then allows a successful detection of corresponding points in successive images. The high level of information per time ration results in huge data files (~ GB). Consequently a high computer performance is necessary for fast image processing. Table 1 summarizes the above called advantages and disadvantages:

Advantages	Disadvantages	
Suitable for indoor and outdoor environments	No direct estimation of the absolute scale factor	
No shadowing or multipath effects	Prominent structures are needed	
Get a high level of information at each acquisition time	Huge data files	
Estimation of position and orientation with only one sensor	High computer performance is needed	
High level of control on account of admission of the		
environment		
High frame rate		

Table 1: Advantages and Disadvantages of using camera systems:

Considering the mentioned advantages, in the present project the spatial trajectory of a mobile survey vehicle should be estimated using an omnidirectional camera system. This task is also called *ego-motion estimation*. Furthermore, an assessment of the achievable accuracy should be taken. We discuss the required photogrammetric background in section 2 which is essential for ego-motion estimation. Using omnidirectional camera systems, it is possible to capture 360° views of the environment at each camera position. Therefore, especially the orientation can be estimated in a more stable way. In our presented project we use the omnidirectional camera system *Ladybug3* [www.ptgrey.com]. This system consists of a conglomeration of six separate cameras which point in different directions to form the complete omnidirectional camera systems (and especially of the *Ladybug3*) are introduced in section 3. On account of the six separate cameras of the *Ladybug3*, some challenges arise. In contrast to [GLUCKMAN, J., NAYAR, K., 1998], we cannot use the generated omnidirectional image for ego-motion estimation. The concept we developed to overcome this challenge is the key paragraph of the paper (section 4). Finally, we give the first results which are achieved with the new developed concept.

# **ESTIMATION OF SPATIAL POSITIONS USING CAMERA SYSTEMS**

The set of images in chronological order captured by a camera in a definite period of time is called an image sequence. For now, we consider a monocular camera system which is mounted on a mobile platform. The reference point of the camera is the projection centre O. The goal is to determine the motion of the projection centres from the image sequences. We can estimate three parameters of position and three of orientation (yaw rate, pitch and roll angle) at each time of acquisition (Figure 2.1, Left).



Figure 2.1: Left: Ego-motion of a monocular camera; Right: Schematic relationship between the camera system and the image system

To determine the ego-motion of a camera, we define a camera model which is based on the mathematical model of [McGLONE, C., et al., 2004]. We construct the projection ray connecting the object point X, image point of x' and the projection centre O geometrically (Figure 2.1, right). We model the relationship between the object point X and the image point x'. For an easier notation, we represent all parameters in homogeneous coordinates:

$$\mathbf{x}' = \begin{bmatrix} K & \mathbf{0}_3 \end{bmatrix} \mathbf{M}_t \mathbf{X}, \text{ with } \mathbf{0}_3 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}^T$$
 (2.1)

The calibration matrix K defines the coordinate transformation from a spatial camera system into a 2 dimensional (ccd-) image system. This transformation is called the *interior orientation*. This means we estimate the relationship between the projection centre of the camera and the coordinate axes of the image plane (Figure 2.1, right). Concerning a camera with an affine sensor, the interior orientation is defined by five parameters (principle point  $(x_h, y_h)$ , principle distance (c), skew factor (s), scale difference of the coordinate axes (m)). Concerning general cameras, non-linear geometric distortions have to be taken into account. Therefore, the relation between image system and camera system is established in two steps: First, non-linear geometric distortions are corrected and subsequently we achieve *rectified images*. Second, the calibration matrix K is nominated using the previously mentioned parameters:

$$K = \begin{bmatrix} c & c & x_h \\ 0 & c & (1+m) & y_h \\ 0 & 0 & 1 \end{bmatrix}, M_t = \begin{bmatrix} R_t & T_t \\ 0_3^T & 1 \end{bmatrix}$$
(2.2)

For the estimation of the interior orientation comprising a modelling of the non-linear distortion, [ABRAHAM, S., HAU, T., 1997] developed an automatic method. The motion matrix  $M_t$  with its rotation matrix  $R_t$  and the translation vector  $T_t$  describe the spatial orientation and position of the projection centre O at the time t in relation to a reference system. However, we are only interested in the relative motion matrix  $M_t^{t+1}$  between two successive camera positions (defined by the projection centres) which is called the *relative orientation* (Figure 2.2).



#### Figure 2.2: The transformation between the projection centres O' and O'' (relative orientation)

Due to the geometrical exposition in Figure 2.2 the relative motion  $M_t^{t+1}$  can be expressed as:

$$M_t^{t+1} = \left(\begin{array}{c} M_t \end{array}\right)^{-1} M_{t+1} \tag{2.3}$$

[LÄBE T., FÖRSTNER W., 2006] introduce an automatic procedure to estimate the relative orientation from digital rectified image sequences. For each time of acquisition, a best estimate of the relative motion parameters is achieved. We obtain a relative trajectory in a local coordinate system which we define as a *photogrammetric model*. Therefore, an absolute relationship to a reference system is not given. An absolute orientation of the relative trajectory is only possible via a spatial similarity transformation achieved by pass-point information. In this project, we are interested in the relative orientation only and neglect this limitation at first.

The estimation of the relative trajectory with a monocular camera system can even be computed in real-time. Real-time algorithms are particularly used in robotic applications and by autonomous navigation ([DAVISON, A. J. et al., 2007], [MOURAGNON, E. et al., 2006]).

### **OMNIDIRECTIONAL CAMERA SYSTEMS**

## 1.1 Applications and Configurations of OCS

In section 2 we described methods to estimate the ego-motion of a traditional camera system which has a small field of view (about  $45^{\circ}$ ). By contrast, an omnidirectional camera system is used in our present project. Although, the purpose of ego-motion estimation remains. In this section, we present the main applications and configurations of omnidirectional camera systems. Furthermore, the used system *Ladybug3* is introduced.

Omnidirectional camera systems usually generate images of almost the complete sphere seen from one particular point of view. On account of a 360° view of the environment, these systems are widely used in robotic applications, in autonomous navigations and in supervision technology ([YAGI, Y., et al., 2005], [YACHIDA, M., 1998], [GLUCKMAN, J., NAYAR, K., 1998]). The advantages in comparison to usual monocular camera systems are obvious: The generated panoramic image shows a larger field of view (up to 360°) than a usual image. A higher level of information can be used to estimate the ego-motion of the camera. Occlusions and glancing intersections can be handled easier than in the case of a limited field of view. In the following three well known configurations of omnidirectional camera systems and their characteristics are described.

1. Extreme fish eye lenses.

Extreme fish eye lenses are used to take images with a field of view up to 180°. The complete sphere cannot be shown in a single image. Besides the images show large aberrations.

2. Catadioptric systems.

By a catadioptric system, a reflective surface is mounted in lengthening of the object lens. The environment is projected via the reflective surface to the camera. The recording image is therefore the mirror image of the environment. [ZIVKOVIC Z., BOOIJ, O., 2005] for example give a detailed description of the configuration.

3. Mosaic-based camera systems.

Mosaic-based camera systems are made of a conglomeration of several monocular cameras which are put up in different recording directions to form one complete system. Due to the overlapping regions in the single images, an omnidirectional image can be generated.

# 1.2 The Ladybug3

The omnidirectional camera system *Ladybug3* is a mosaic-based camera system which was developed by the company *Point Grey Research*. Five cameras are positioned in a horizontal ring and one camera on the top which points vertically in the sky (Figure 3.1). The single unrectified image sequences from each camera, as well as the stitched omnidirectional image sequence may be captured. Table 3.2 summarizes the main characteristics of the camera which are taken from [POINT GREY RESEARCH, 2008].



Figure 3.1:	The Ladybug3	and the orientation	of the cameras [POIN	T GREY RESEARCH, 2008]
-------------	--------------	---------------------	----------------------	------------------------

Imaging Sensor Type	Six Sony progressive scan colour CCDs	
	(five in a horizontal ring, one on top)	
Maximum Resolution	1616(H) x1232(V) (each sensor)	
Field of view	>80% of full sphere	
Maximum frame rate	16FPS JPEG compressed	
	6.5 FPS uncompressed	
Dimensions 134mm x 141mm		
Mass	2,416kg	

Table 3.2: Main characteristics of the camera:

The use of the *Ladybug3* for ego-motion estimation comes along with some challenges: In the methods mentioned in section 2, we assume that a unique projection centre exists. But since the *Ladybug3* comprises six different projection centres, we cannot use these methods directly. The corresponding effect is shown in the generated omnidirectional image (Figure 3.3): In the overlapping areas, we recognise explicit ambiguities. Caused by the absence of a unique, common projection centre, in the overlapping areas aberrations occur during the stitching process. Therefore, for our applications the generated omnidirectional images are not usable. This is the reason why we decide to treat each camera individually. Consequently, we get image sequences of six monocular camera systems. The (relative) spatial ego-motion of the complete system should be estimated from those six image sequences. The newly developed concept is described in the following section.



Figure 3.3: Each singe unrectified image, the stitched omnidirectional image and aberrations

# **DEVELOPED CONCEPT**

For the solution of the described challenge we have decided on a two-step solution. First, a high precision calibration of each camera is done using the method from [ABRAHAM, S., HAU T., 1997]. With this the interior orientation is available and we can process rectified image sequences. We estimate the six photogrammetric models for each camera using the procedure of [LÄBE T., FÖRSTNER W., 2006] introduced in section 2. We use the Matlab based implementation of [LÄBE T., FÖRSTNER W., 2006] called *Aurelo*. For each time of acquisition t(0,...T) and for each camera c (c=1,..6) the motion matrix  ${}^{c}M_{t}$  comprising the cameras orientation  ${}^{c}R_{t}$  and position  ${}^{c}T_{t}$  (Figure 4.1, left) is obtained. The local coordinate system (of the photogrammetric model) is defined by the projective centre that belongs to the first image of the sequence. As a further result, we obtain the accuracies of the estimated motion parameters represented by the covariance matrices  ${}^{c}\Sigma_{MM}(t)$ .

$${}^{c}M_{t} = \begin{bmatrix} {}^{c}R_{t} & {}^{c}T_{t} \\ 0_{3} & 1 \end{bmatrix}, \qquad {}^{c}\Sigma_{MM}(t)$$

$$(4.1)$$

As in the equation 2.3 we compute the relative motion matrices expressed as follows:

$${}^{c}M_{t}^{t+1} \stackrel{\text{def}}{=} {}^{c}\Delta M_{t} = \left( {}^{c}M_{t} \right)^{-1} {}^{c}M_{t+1}$$

$$(4.2)$$

We receive six different trajectories which differ pairwise by a spatial similarity transformation (Figure 4.1).



Figure 4.1: Schematic process of the developed concept: relative orientations of the cameras (left), relative trajectories (middle), transformed trajectories (right)



Figure 4.2: Geometric relationship between the observations and the unknown parameters

The intention of the second step is to transform the estimated spatial trajectories in a common reference system Z and to estimate an optimum spatial trajectory in relation to a certain reference point from this. We define the reference system Z in the centre of the camera system (Figure 3.1, right). Figure 4.2 shows the geometric relationship between the observations  ${}^{c}\Delta M_{t}$  and the unknown parameters of the spatial similarity transformations  ${}^{z}_{c}M$  (with rotation  ${}^{z}_{c}R$ , translation  ${}^{z}_{c}T$  and scale factor  ${}^{z}_{c}\lambda$ ) in the reference system Z as well as the unknown single reference trajectory  ${}^{z}\Delta M_{t}$ . The relationship between the desired reference trajectory and each of the six observed camera trajectories is given by:

$${}^{c}\Delta M_{t} = \left({}^{z}_{c}M\right)^{-1} {}^{z}\Delta M_{t} \left({}^{z}_{c}M\right)$$

$$\tag{4.3}$$

In more detail, using equation 4.3 we get:

$$\begin{bmatrix} {}^{c}\Delta R_{t} & {}^{c}\Delta T_{t} \\ 0_{3} & 1 \end{bmatrix} = \begin{bmatrix} {}^{z}\lambda_{c}^{z}R & {}^{z}T \\ 0_{3} & 1 \end{bmatrix}^{-1} \begin{bmatrix} {}^{z}\Delta R_{t} & {}^{z}\Delta T_{t} \\ 0_{3} & 1 \end{bmatrix} \begin{bmatrix} {}^{z}\lambda_{c}^{z}R & {}^{z}T \\ 0_{3} & 1 \end{bmatrix}$$
(4.4)

From equation 4.4 we get the non-linear relation of the observed parameters of rotation  ${}^{c}\Delta R_{t}$  and translation  ${}^{c}\Delta T_{t}$  to the unknown parameters as:

$${}^{c}\Delta R_{t} = {}^{z}_{c}R^{T} {}^{z}\Delta R_{t} {}^{z}_{c}R \tag{4.5}$$

$${}^{c}\Delta T_{t} = \frac{1}{\frac{z}{c\lambda}} {}^{z}_{c} R^{T} \left[ {}^{z}\Delta R_{t} {}^{z}_{c} T + {}^{z}\Delta T_{t} - {}^{z}_{c} T \right]$$

$$(4.6)$$

With a linearisation of the equations 4.5 and 4.6, we can set up a Gauss-Markov model. We assume the relative position of the six cameras to be rigid over time and thus the transformations  $({}^{z}_{c}M)$  are kept constant for every time. We get a first approximation for the six transformations  $({}^{z}_{c}M)$  from the calibration file delivered by the manufacturer. With the use of this information and the relationship 4.3, we also obtain approximate values for the unknown trajectory of the reference system. The stochastic model is designed to be based on the covariance matrices of equation 4.1 using error propagation. We now solve for the unknown parameters using the well known mathematics concepts of adjustment calculus [NIEMEIER, W., 2002].

With the present concept, some advantages arise in contrast to monocular systems. At each time of acquisition, we estimate from the combined information of the six camera systems an optimum (relative) position and orientation. This offers a high level of controllability and stability. Also in case

of failure of a camera, the remaining systems may still deliver enough information for a stable position and orientation estimation.

## **TEST SET AND FIRST RESULTS**

We evaluate the developed concept concerning reliability and accuracy in an experimental set up. Since we just offer first results, we concentrate on the accuracy of the parameters of orientation. The experimental set-up is as follows: The camera is mounted on a mobile survey vehicle. The image sequences are captured while the vehicle moves along a slightly curved path.



Figure 5.1: The relative trajectories of the cameras (left), transformed trajectories (right)

Figure 5.1 (left) shows the relative trajectories of the six single cameras determined using the software *Aurelo*. Each single trajectory is modelled in the local camera system relative to the first image of the sequence. Since these trajectories model the relative motion without reference to an absolute world frame, we abstain from scaling the axes in the plots. If we now transfer the trajectories to the reference system Z using the approximated transformations, we obtain the six approximations of the unknown reference trajectory shown in Figure 5.1 (right).

To make a first statement regarding the accuracy of the orientation parameters of the approximated reference trajectories, we refer to the standard deviation of the yaw rate (Figure 5.2). The single cameras deliver different accuracies. Since due to the alignment of the cameras, different admission conditions are given. Because of static image regions in the image sequences (for example, from the mobile survey vehicle) or monotonous regions, as for example blue sky, the estimation of the relative orientation can be affected (cameras 1 and 4). Cameras 2 and 5 deliver very good results. Maximum standard deviations of 0.04 gon and 0.07 gon are given. From experience these results can be reached with the software *Aurelo* under good conditions.

The accuracy of the orientations becomes more inaccurate during the time of acquisition. This can be explained by the fact that the datum is only defined in the first admission time. Like a unilaterally connected polygon, the inaccuracy grows with advancing distance. This problem is to be mastered only with pass-point information.



**Figure 5.2:** The standard deviation of the yaw-rates of the cameras

At this early stage of the project, a best estimate of the reference trajectory and the spatial similarity transformations is not achieved yet. But due to the approximated values, we can compute the average reference trajectory (Figure 5.3).



Figure 5.3: The average reference trajectory

# CONCLUSION

In this paper, we discuss how to estimate the ego motion of a mobile platform using camera systems. We concentrated on omnidirectional camera systems consisting of a conglomeration of single cameras. We depicted the challenges which arose due to the multiple projection centres. As the key contribution we presented an approach to deal with these challenges and to estimate the trajectory of the camera system as well as the spatial relation between the single cameras of the conglomeration. We showed the first results of our approach using the *Ladybug3* of Point Grey and demonstrated the applicability especially for the estimation of orientation.

In the future, our concept should be tested with larger image sequences. We will consider how to optimize our two-step solution to a one-step solution. Furthermore, we will propose the use of the *Ladybug3* in combination with GPS for real-time applications.

## ACKNOWLEDGEMENTS

The authors would like to thank all involved institutes and persons for the support. Special thanks to Thomas Läbe who has supported us by the evaluation with *Aurelo*.

#### REFERENCES

ABRAHAM, S., HAU, T., 1997: *Towards Autonomous High-Precision Calibration of Digital Cameras*. Proceedings of SPIE Annual Meeting, p. 82-93, San Diego 1997.

DAVISON, A. J., REID I. D., MOLTON N., STASSE O., 2007: *MonoSLAM: Real-Time Single Camera SLAM*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 29, Issue 6, p. 1052-1067, June 2007.

EICHHORN, A., 2005: *Ein Beitrag zur Identifikation von dynamischen Strukturmodellen mit Methoden der adaptiven KALMAN-Filterung.* Deutsche Geodätische Kommission, Verlag der Bayerischen Akademie der Wissenschaften, Reihe C, Heft 585, München 2005.

GLUCKMAN, J., NAYAR, K., 1998: *Ego-Motion and Omnidirectional Cameras*. Sixth International Conference on Computer Vision, p. 999-1005, Bombay 1998.

GOODALL, C., 2009: Improving Usability of Low-Cost INS/GPS Navigation System using Intelligent Techniques. The University of Calgary, UCGE Report, number 20276, Calgary 2009.

KUHLMANN, H., SIEMES, M., 2007: Nutzung eines Multi-Sensor-System mit Kalman-Filter für die Bestimmung von Längsbewegungen einer Landmaschine in der Präzisionslandwirtschaft. Ingenieurvermessung 07, Beiträge zum 15. Internationalen Ingenieurvermessungskurs Graz 2007, p. 369-382, Graz 2007.

LÄBE, T., FÖRSTNER, W., 2006: *Automatic relative orientation of images*. Proceedings of the 5th Turkish-German Joint Geodetic Days, Berlin 2007.

McGLONE, C., MIKHAIL, E., BETHEL, J. (eds)., 2004: *Manual of Photogrammetry*. 5th edn. Bethesda, MD, USA: American Society of Photogrammetry and Remote Sensing 2004.

MOURAGNON, E., LHUILLIER M., DHOME, M., DEKEYSER, F. and SAYD, P., 2006: *Monocular Vision Based SLAM for Mobile Robot*. Proceedings of the 18<sup>th</sup> International Conference on Pattern Recognition, Volume 3, p. 1027-1031, Hong Kong 2006.

NIEMEIER, W., 2002: Ausgleichsrechnung: eine Einführung für Studierende und Praktiker des Vermessungs- und Geoinformationswesens. De Gruyter, Berlin, 2002.

NISTER, D., 2004: *An efficient solution to the five-point relative pose problem.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 26, Issue 6, p. 756-770, June 2004.

POINT GREY RESEARCH, 2008: *Ladybug3:Technical Reference Manual*. Version 1.1, Richmond BC, December 2008.

RETSCHER, G., 2002: *Multi-Sensor Systems for Machine Guidance and Control*. XXII International Congress FIG 2002, American Society for Photogrammetry and Remote Sensing, Washington DC 2002.

YACHIDA, M., 1998: *Omnidirectional Sensing and Combined Multiple Sensing*. Proceedings of the 1998 Workshop on Computer Vision for Virtual Reality Based Human Communication, Washington DC, January 1998.

YAGI, Y., IMAI, K., TSUJI, K., YACHIDA, M., 2005: *Iconic memory-based omnidirectional route panorama navigation*. Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 27, Issue 1, p. 78 – 87, January 2005.

ZIVKOVIC Z., BOOIJ O., 2005: *How did we built our hyperbolic mirror omnidirectional camera –practical issues and basic geometry*. IAS technical report (IAS-UVA-05-04), Informatics Institute, University of Amsterdam, Amsterdam 2005.

#### Links:

http://www.ptgrey.com, last accessed on January 08, 2010

Contact: BSc Maximilian Muffert, Institute of Geodesy and Geoinformation, University of Bonn Nussallee 17, D-53115 Bonn, Germany Email: <u>mmuffert@uni-bonn.de</u>