Calibration of focal length and 3d pose based on the reflectance and depth image of a planar object

Christian Beder and Reinhard Koch

Computer Science Department
Kiel University
Germany
{beder,rk}@mip.informatik.uni-kiel.de

Abstract: The estimation of position and orientation of a PMD camera in a global reference frame is required by many measurement applications based on such systems. PMD cameras produce a depth as well as a reflectance image of low resolution compared to standard optical cameras, so that calibration of the cameras based on the reflecance image alone is difficult.

We will present a novel approach for calibrating the focal length and 3d pose of a PMD camera based on the depth and reflectance image of a planar checkerboard pattern. By integrating both sources of information higher accuracies can be achieved. Furthermore, one single image is sufficient for calibrating the focal length as well as the 3d pose from a planar reference object. This is because the depth measurements are orthogonal to the lateral intensity measurements and provide direct metric information.

Keywords: calibration, pose estimation, PMD camera

1 Introduction

Recently real-time active 3d range cameras based on time-of-flight technology like the Photonic Mixer Device (PMD) have become available. Calibration and 3d pose estimation of such systems is an important sub-task for many measurement applications.

Furthermore, PMD range cameras are used in conjunction with other measurement devices. For instance the accuracy of stereo systems and PMD cameras is compared in [2] and a fused surface reconstruction is proposed in [1], which requires a reliable relative calibration of the stereo and the PMD camera. Also in [9] a fusion of PMD and stereo is proposed. The relative orientation of a PMD and an optical camera is also required in [12]. We will provide a simple and easy method for establishing such relative orientation based on images of a planar checkerboard pattern.

Calibration of PMD cameras is described in [8] and [10], who both use photogrammetric methods described for instance in [13] and [3] on the reflectance images to estimate focal length and 3d pose. The depth image is not used for pose

estimation, because the scope of those works is also the depth calibration of the cameras. It is pointed out, that the poor quality of the low resolution intensity images makes precise localization very difficult (cf. also [11]). In contrast to those works we assume the depth to be calibrated, so that it can be used for estimating 3d pose and focal length, thereby improving the achievable accuracy.

We will present a novel approach for calibrating focal length and 3d pose based on both the reflectance and the depth image provided by a PMD camera. In contrast to optical cameras only one single image is required for calibrating focal length and 3d pose, as the depth information is orthogonal to the lateral information contained in the intensity image. Due to the high redundancy of the depth measurements, highly accurate results should be achievable.

In section 2 the geometric camera model will be described. This model is used in section 3 to estimate focal length and 3d pose from one single reflectance and depth image of a planar checkerboard calibration pattern. Finally we will present some experiments in section 5.

2 Camera model

The geometric camera model assumed in the following is as follows. A 3d point is projected according to the homogeneous equation (cf. [7], p.142)

$$\mathbf{x}_i \propto \mathsf{K}R(\boldsymbol{X}_i - \boldsymbol{C}) \tag{1}$$

where $\boldsymbol{C}$ is the unknown position of the camera, R is the unknown rotation of the camera and the intrinsic parameters are contained in the calibration matrix

$$\mathsf{K} = \begin{pmatrix} c & 0 & x_h \\ 0 & c & y_h \\ 0 & 0 & 1 \end{pmatrix} \tag{2}$$

We assume the principal point (x_h, y_h) to be known (usually in the center of the image), the skew to be zero and the aspect ratio to be one. Hence, only the focal length c is introduced as an unknown parameter here.

In contrast to classical optical cameras we assume a depth image to be given, so that for each pixel $\mathbf{x}_i$ also the corresponding depth λ_i is known. Hence, for each pixel the corresponding 3d point $\boldsymbol{X}_i$ can be computed as follows

$$\boldsymbol{X}_i = \lambda_i \frac{R^\mathsf{T}\mathsf{K}^{-1}\mathbf{x}_i}{\sqrt{\mathbf{x}_i^\mathsf{T}\mathsf{K}^{-\mathsf{T}}\mathsf{K}^{-1}\mathbf{x}_i}} + \boldsymbol{C} \tag{3}$$

Assuming without loss of generality the planar reference object to be located at the $Z = 0$ plane, one can derive one constraint per pixel by requiring $X_z = 0$. Using equation (3) this is expressible as

$$\lambda_i \boldsymbol{r}_z^\mathsf{T}\mathsf{K}^{-1}\mathbf{x}_i + C_z\sqrt{\mathbf{x}_i^\mathsf{T}\mathsf{K}^{-\mathsf{T}}\mathsf{K}^{-1}\mathbf{x}_i} = 0 \tag{4}$$

or equivalent

$$\lambda_i = -\frac{C_z\sqrt{\mathbf{x}_i^\mathsf{T}\mathsf{K}^{-\mathsf{T}}\mathsf{K}^{-1}\mathbf{x}_i}}{\boldsymbol{r}_z^\mathsf{T}\mathsf{K}^{-1}\mathbf{x}_i} \tag{5}$$

Obviously a planar reference object is not sufficient for calibration using depth alone. Hence the reflectance image has to be used. Therefore the intensities are assumed to be equal to the known intensities of the reference object

$$I_{ref}(\boldsymbol{X}_i) = I_{ref}\left(\lambda_i \frac{R^\mathsf{T}\mathsf{K}^{-1}\mathbf{x}_i}{\sqrt{\mathbf{x}_i^\mathsf{T}\mathsf{K}^{-\mathsf{T}}\mathsf{K}^{-1}\mathbf{x}_i}} + \boldsymbol{C}\right) = I(\boldsymbol{x}_i) \tag{6}$$

or by substituting λ_i

$$I(\boldsymbol{x}_i) = I_{ref}\left(\boldsymbol{C} - \frac{C_z R^\mathsf{T}\mathsf{K}^{-1}\mathbf{x}_i}{\boldsymbol{r}_z^\mathsf{T}\mathsf{K}^{-1}\mathbf{x}_i}\right) \tag{7}$$

where $I(\boldsymbol{x}_i)$ denotes the reflectance image and I_{ref} denotes the reference image (here a smoothed checkerboard pattern).

In the following we will show, how equation (5) and (7) can be used to estimate the unknown focal length and pose of the camera.

3 Optimization

Now we will show, how the focal length and the pose can be determined from the depth and the reflectance image. Therefore we assume the focal length $c^{(0)}$ as well as rotation $R^{(0)}$ and position $\boldsymbol{C}^{(0)}$ to be approximately known. We will postpone the discussion of obtaining initial values to the next section and first show how to estimate those quantities starting from approximate initial values.

Representing the rotation matrix using its Taylor expansion (cf. [5], p.53)

$$R^{(\nu+1)} \approx R^{(\nu)} + \begin{pmatrix} 0 & -\kappa & \phi \\ \kappa & 0 & -\omega \\ -\phi & \omega & 0 \end{pmatrix} \tag{8}$$

and collecting the unknown parameters into a vector

$$\boldsymbol{p} = \begin{pmatrix} c & \omega & \phi & \kappa & C_x & C_y & C_z \end{pmatrix}^\mathsf{T} \tag{9}$$

one can synthesize the depth image from the parameters using equation (5) as

$$\lambda_i = f_{i1}(\boldsymbol{p}) \approx f_{i1}\left(\boldsymbol{p}^{(\nu)}\right) + \left.\frac{\partial f_{i1}}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}^{(\nu)}} \Delta\boldsymbol{p} \tag{10}$$

and synthesize the reflectance image from the parameters using equation (7) as

$$I(\boldsymbol{x}_i) = f_{i2}(\boldsymbol{p}) \approx f_{i2}\left(\boldsymbol{p}^{(\nu)}\right) + \left.\frac{\partial f_{i2}}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}^{(\nu)}} \Delta\boldsymbol{p} \tag{11}$$

Note, that in this formulation no derivative of any observed noisy low-resolution depth or reflectance image is required for the Taylor expansion.

Denoting the Jacobians with

$$\mathsf{A}_{i1} = \left.\frac{\partial f_{i1}}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}^{(\nu)}} \qquad \mathsf{A}_{i2} = \left.\frac{\partial f_{i2}}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}^{(\nu)}} \tag{12}$$

and

$$\Delta \boldsymbol{l}_{i1} = \lambda_i - f_{i1}\left(\boldsymbol{p}^{(\nu)}\right) \qquad \Delta \boldsymbol{l}_{i2} = I(\boldsymbol{x}_i) - f_{i2}\left(\boldsymbol{p}^{(\nu)}\right) \tag{13}$$

one obtains the parameter covariance as (cf. [5], p.87)

$$\mathsf{C}_{\boldsymbol{pp}}^{(\nu+1)} = \left(\sum_i^N \frac{1}{\sigma_1^2}\mathsf{A}_{i1}^\mathsf{T}\mathsf{A}_{i1} + \frac{1}{\sigma_2^2}\mathsf{A}_{i2}^\mathsf{T}\mathsf{A}_{i2}\right)^{-1} \tag{14}$$

From this the parameter update is computed as

$$\Delta \boldsymbol{p} = \mathsf{C}_{\boldsymbol{pp}}\left(\sum_i^N \frac{1}{\sigma_1^2}\mathsf{A}_{i1}^\mathsf{T}\Delta \boldsymbol{l}_{i1} + \frac{1}{\sigma_2^2}\mathsf{A}_{i2}^\mathsf{T}\Delta \boldsymbol{l}_{i2}\right) \tag{15}$$

yielding the improved parameter vector

$$\boldsymbol{p}^{(\nu+1)} = \boldsymbol{p}^{(\nu)} + \Delta \boldsymbol{p} \tag{16}$$

The residuals in the depth image are given by

$$\Omega_1 = \sum_i^N ||\mathsf{A}_{i1}\Delta \boldsymbol{p} - \Delta \boldsymbol{l}_{i1}||^2 \tag{17}$$

and the residuals in the reflectance image are given by

$$\Omega_2 = \sum_i^N ||\mathsf{A}_{i2}\Delta \boldsymbol{p} - \Delta \boldsymbol{l}_{i2}||^2 \tag{18}$$

so that the variance factors can be updated according to (cf. [5], p.91)

$$\left(\sigma_1^{(\nu+1)}\right)^2 = \left(\sigma_1^{(\nu)}\right)^2 \frac{\Omega_1}{N-7} \qquad \text{and} \qquad \left(\sigma_2^{(\nu+1)}\right)^2 = \left(\sigma_2^{(\nu)}\right)^2 \frac{\Omega_2}{N-7} \tag{19}$$

Starting with initial variance factors $\sigma_1^{(0)} = \sigma_2^{(0)} = 1$ this process should be iterated until convergence. We chose as convergence criterion, that the update is smaller than 1% of the expected accuacy, i.e. $\Delta \boldsymbol{p}\mathsf{C}_{\boldsymbol{pp}}^{-1}\Delta \boldsymbol{p} < 0.01$.

Up to now we have shown, how optimal estimates for the parameters can be obtained by stepwise improvement. In the following we will discuss, how non-optimal initial values can be determined from the input data.

4 Initial values

The iterative method described in the previous sections has to be initialized with approximate initial values. To obtain those initial values, the checkerboard pattern has to be identified in the intensity image. This can either be done manually or using the techniques implemented in [4]. We chose to manually select the four corners of the checkerboard pattern.

After identifying the corners of the checkerboard pattern in the image, we computed the average depth $\overline{\lambda}$ of this region from the depth image. Knowing the diagonal of the checkerboard in image space d_i as well as in object space d_o an approximate focal length can be computed as $c^{(0)} = \overline{\lambda}\frac{d_i}{d_o}$.

Using this approximate focal length and the known 2d-3d correspondences on the planar checkerboard, standard backward intersection techniques (cf. [6],p.786) can be used to obtain an initial pose of the camera.

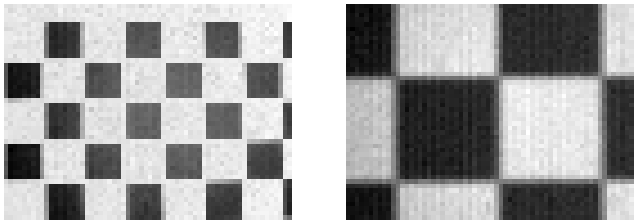

Figure 1 Two typical reflectance images of our PMD camera. The two images constitute the two extrema of a zoom sequence exploiting the full zoom range (see text).

$\sigma_c[pel]$	$\sigma_\omega[\,^\circ]$	$\sigma_\phi[\,^\circ]$	$\sigma_\kappa[\,^\circ]$	$\sigma_X[mm]$	$\sigma_Y[mm]$	$\sigma_Z[mm]$
0.2081	0.1197	0.0694	0.0601	3.9994	6.8511	0.3785

Table 1 Computed standard deviations for a fronto-parallel image of a checkerboard pattern at approximately $3m$ distance with computed focal length $c = 184.3$, i.e. an opening angle of approximately $20^\circ \times 15^\circ$, corresponding to the reflectance image depicted on the left hand side of figure 1.

5 Results

We applied the presented method for the task of calibrating a PMD camera using the setup depicted in figure 2. In figure 1 two exemplary reflectance images of the checkerboard calibration pattern having a resolution of 64×48 pixels are shown.

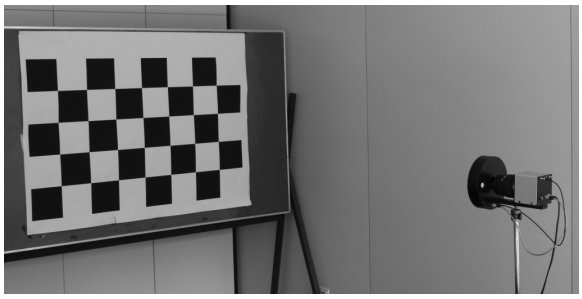

Figure 2 The setup used for calibration of the PMD camera.

A calibration from a fronto-parallel checkerboard pattern at approximately $3m$ distance with opening angle of approximately $20^\circ \times 15^\circ$ corresponding to the reflectance image depicted on the left hand side of figure 1 yielded the accuracies presented in table 1. It can be seen, that the computed values are quite competitive and that the accuracy in direction of the depth are one order of magnitute more accurate. This high accuracy results from the high redundancy of the estimation and the orthogonality of reflectance and depth information. However, the result depends on the assumption, that there is no systematic error contained in the depth measurements, so that an accurate depth calibration of the PMD camera (cf. [10]) is crucial.

In table 2 the computed correlations between the parameters are depicted. It can be seen, that rotation and lateral motion are highly correlated due to the narrow opening angle of the PMD camera. This effect is even worse if the PMD camera is zoomed, i.e. the opening angle is narrowed. To quantify this effect we took a zoom sequence ranging over the full possible range starting at the image depicted on the left hand side of figure 1 having a computed focal length of $184.3pel$ and ending at the image depicted on the right hand side of figure 1 having a computed focal

	c	ω	ϕ	κ	X	Y	Z
c	1.0000	-0.0847	-0.1005	-0.0441	-0.1021	0.0915	-0.2118
ω	-0.0847	1.0000	0.0360	0.0291	0.0358	-0.9985	0.2668
ϕ	-0.1005	0.0360	1.0000	0.0327	0.9960	-0.0362	0.0220
κ	-0.0441	0.0291	0.0327	1.0000	0.0222	-0.0314	0.0114
X	-0.1021	0.0358	0.9960	0.0222	1.0000	-0.0360	0.0257
Y	0.0915	-0.9985	-0.0362	-0.0314	-0.0360	1.0000	-0.2708
Z	-0.2118	0.2668	0.0220	0.0114	0.0257	-0.2708	1.0000

Table 2 Computed correlations between the parameters corresponding to the accuracy values shown in table 1

length of 477.9*pel*. The resulting correlation between corresponding rotation and translation parameters is depicted in figure 3. As expected the correlation further increases with decreasing opening angle.

Next we moved the calibration pattern in front of the PMD camera in order to assess the performance of the calibration. The empirical mean of the computed focal lengths for different views of the calibration pattern was $\mu_c = 183.9$ with an empirical standard deviation of $\sigma_c = 4.63$. The deviation of the empirical standard deviation and the computed standard deviations can be explained by systematic errors in the depth measurement as well as by the correlation of the focal length with the pose, especially its Z-component.

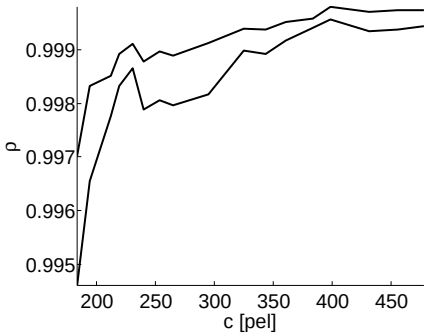

Figure 3 Correlation ρ between rotation and translation against estimated focal length c in pixels for the zoom sequence.

Simultaneously we took pictures of the calibration pattern with a standard 1024×786 camera rigidly coupled with the PMD camera and determined its poses for reference using a standard bundle adjustment. In order to eliminate the effect of correlation between focal length and pose we re-estimated the pose parameters for the PMD images with a fixed focal length of $\mu_c = 183.9$. The resulting poses for the PMD cameras and the corresponding optical cameras are shown in figure 4. Note, that the poses of the PMD cameras are estimated individually while the poses of the optical cameras are obtained using an overall adjustment.

	mean	standard deviation
X	27.8285 cm	9.7977 cm
Y	47.4395 cm	6.0589 cm
Z	74.3781 cm	6.5851 cm

Table 3 Estimated relative translation parameters corresponding to the setup shown in figure 4 together with their standard deviations.

To quantify those results we computed the relative pose of the optical camera with respect to the PMD camera. Because both cameras were rigidly coupled, this should not change over the sequence. The relative translation together with its standard deviations is presented in table 3. Although the standard deviations of the translation are up to 10*cm*, the significance of this result is limited. The reason for this is the large correlation between rotation and translation parameters resulting

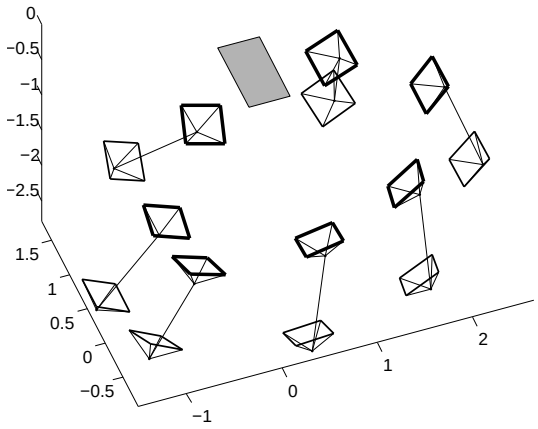

Figure 4 Estimated poses of the PMD cameras (bold) together with corresponding poses of the optical cameras. The position of the calibration pattern is shown in gray.

from the narrow opening angle of the PMD camera. Taking those correlations into account, a precise relative orientation between two cameras for the task of object reconstruction is possible (cf. [2] and [1]).

6 Conclusion

We presented a simple and accurate method for the calibration of focal length and 3d pose based on a reflectance and a depth image of a planar calibration pattern, which enables us to locate a PMD camera in a global coordinate system, and allows for the relative orientation between the PMD camera and a standard optical camera.

It has been shown, that the major limitation results from the narrow opening angle of current PMD cameras. In particular the low resolution of the resulting images is not the limiting factor, because depth and intensity measurements complement each other and allow highly accurate estimations. This is in contrast to the view of [11].

The presented method is a per-image calibration, which estimates all parameters for each image seperately. While this approach is simple and easily applicable, future work will include estimating the calibration parameters from multiple images and including classical optical cameras with larger resolution and more important larger opening angles into the adjustment in order to obtain more accurate results and reduce the correlations between the parameters.

Another interesting approach, which has to be exploited in the future, would be the joint distance, focal length and pose calibration of a PMD camera based on such a multi-camera adjustment.

Acknowledgements

The PMD camera used in the experiments is courtesy of the MultiCam project at ZESS, University of Siegen, Germany. This work was supported by the German Research Foundation (DFG) as part of the 3DPoseMap project (KO-2044/3-1).

References and Notes

1 Christian Beder, Bogumil Bartczak, and Reinhard Koch. A combined approach for estimating patchlets from PMD depth images and stereo intensity images. In F.A. Hamprecht, C. Schnörr, and B. Jähne, editors, *Proceedings of the DAGM 2007*, number 4713 in LNCS, pages 11–20. Springer, 2007.

2 Christian Beder, Bogumil Bartczak, and Reinhard Koch. A comparison of PMD-cameras and stereo-vision for the task of surface reconstruction using patchlets. In *IEEE/ISPRS BenCOS Workshop 2007*, 2007.

3 J.Y. Bouguet. *Visual methods for three-dimensional modelling*. PhD thesis, 1999.

4 Intel Corporation. Opencv: Open source computer vision library. http://www.intel.com/technology/computing/opencv/index.htm.

5 Wolfgang Förstner and Bernhard Wrobel. Mathematical concepts in photogrammetry. In J.C.McGlone, E.M.Mikhail, and J.Bethel, editors, *Manual of Photogrammetry*, pages 15–180. ASPRS, 2004.

6 Wolfgang Förstner, Bernhard Wrobel, Fidel Paderes, Robert Craig, Clive Fraser, and John Dolloff. Analytical photogrammetric operations. In J.C.McGlone, E.M.Mikhail, and J.Bethel, editors, *Manual of Photogrammetry*, pages 763–948. ASPRS, 2004.

7 R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

8 T. Kahlmann, F. Remondino, and H. Ingensand. Calibration for increased accuracy of the range imaging camera SwissrangerTM. In *IEVM06*, 2006.

9 K.D. Kuhnert and M. Stommel. Fusion of stereo-camera and PMD-camera data for real-time suited precise 3d environment reconstruction. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, October 2006.

10 M. Lindner and A. Kolb. Lateral and depth calibration of PMD-distance sensors. In *International Symposium on Visual Computing (ISVC06)*, volume 2, pages 524–533. Springer, 2006.

11 T.D.A. Prasad, K. Hartmann, W. Wolfgang, S.E. Ghobadi, and A. Sluiter. First steps in enhancing 3d vision technique using 2d/3d sensors. In V. Chum, O.Franc, editor, *Computer Vision Winter Workshop 2006*, pages 82–86, University of Siegen, 2006. Czech Society for Cybernetics and Informatics.

12 B Streckel, B. Bartczak, R. Koch, and A. Kolb. Supporting structure from motion with a 3d-range-camera. In *Scandinavian Conference on Image Analysis (SCIA07)*, June 2007.

13 Zhengyou Zhang. Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. In *Proceedings of the International Conference on Computer Vision*, pages 666–673, Corfu, Greece, 1999.