Lehrstuhl für Bildverarbeitung
Institute of Imaging & Computer Vision

Camera Calibration for Fish-Eye Lenses in Endoscopy with an Application to 3D Reconstruction

Thomas Stehle and Daniel Truhn and Til Aach and Christian Trautwein and Jens Tischendorf

Institute of Imaging and Computer Vision
RWTH Aachen University, 52056 Aachen, Germany
tel: +49 241 80 27860, fax: +49 241 80 22200
web: www.lfb.rwth-aachen.de

in: Proceedings IEEE International Symposium on Biomedical Imaging (ISBI). See also BIBTEX entry below.

BIBTEX:

```
@inproceedings{STE07a,
   author        = {Thomas Stehle and Daniel Truhn and Til Aach and
    Christian Trautwein and Jens Tischendorf},
   title         = {Camera Calibration for Fish-Eye Lenses in Endoscopy
       with an Application to {3D} Reconstruction},
   booktitle     = {Proceedings IEEE International Symposium on Biomedical Imaging (ISBI)},
   publisher     = {IEEE},
   address         = {Washington, D.C.},
   month         = {April 12-15},
   year         = {2007},
   pages         = {1176--1179 (CD-ROM, ISBN: 1-4244-0672-2)}}
```


document created on: April 18, 2007
created from file: isbi-2007.tex
cover page automatically created with CoverPage.sty
(available at your favourite CTAN mirror)

RWTH AACHEN
RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

CAMERA CALIBRATION FOR FISH-EYE LENSES IN ENDOSCOPY WITH AN APPLICATION TO 3D RECONSTRUCTION

Thomas Stehle, Daniel Truhn, Til Aach

RWTH Aachen University
Institute of Imaging and Computer Vision
52056 Aachen, Germany

Christian Trautwein, Jens Tischendorf

University Hospital Aachen
Medical Clinic III
Pauwelsstr. 30, 52074 Aachen, Germany

ABSTRACT

Image analysis tasks such as 3D reconstruction from endoscopic images require compensation of geometric distortions introduced by the lens system. Appropriate camera calibration is thus necessary. Commonly used calibration algorithms rely on the well-known pinhole camera model, extended by parametric terms for radial distortions. In this paper, we demonstrate that these models are not appropriate if very strong distortions occur as is the case for endoscopic fish-eye lenses. As an alternative, we analyze a generic calibration algorithm published recently by Kannala and Brandt, which is based on more general projection equations. We show qualitatively and quantitatively that this algorithm is well suited to deal with significant distortions especially in the image's rim regions. Furthermore, we demonstrate how images of a colon phantom that were corrected in such a manner can be used to obtain a 3D reconstruction.

Index Terms— biomedical image processing, camera calibration, endoscopy, fish-eye lens, 3D reconstruction

1. INTRODUCTION

Cancer of the colon is the fourth most common type of cancer and the second leading cause of cancer death in the United States of America. More than 135,000 new cases are diagnosed and over 56,000 people die from colorectal cancer each year.

Endoscopy is a widely used technique in preventive medical checkup and therapy of colorectal cancer. Often, suspect polyps can be visually classified as being benign or malignant using features like texture [1] and vascularization [2]. If the cancer is already widely spread and a simple polypectomy is not possible anymore, part of the colon must be removed. During surgery planning, it is beneficial to know the exact position of the cancerous area in the patient's colon. One way to provide this information is to build up a three-dimensional (3D) model of the patient's colon, which can support navigation during the intervention.

Strong geometric distortions caused by fish-eye lenses, however, hamper 3D reconstruction because 3D reconstruction from consecutive endoscopic images is very sensitive to geometric distortions as it strongly relies on epipolar geometry and triangulation [3, 4].

Many camera calibration algorithms [5, 6, 7] — also used in endoscopy [8, 9] — are based on the classical pinhole camera model, which describes distortion free image formation using the intercept theorems. To compensate for radial and tangential lens distortion, the model is extended [10, 11, 12] with appropriate parametric correction terms. This model is well suited for image formation in case of narrow angle or even wide angle lenses, but not in case of fish-eye lenses, which are frequently used in endoscopy. As an example, we demonstrate that calibration with Bouguet's algorithm [7] provides sufficiently accurate results only in the central image region, but not at the image's rim where geometric fish-eye distortion is strongest. For 3D reconstruction, the data in the image's rim regions can therefore not be used, thus narrowing the effective field of view.

We analyze a new calibration algorithm published recently by Kannala and Brandt [13], which is based on more general projection equations. We show quantitatively as well as qualitatively that this algorithm provides accurate results in terms of mean reprojection error and stability even in the image's rim regions. Finally, we show the reconstruction of a phantom obtained from complete endoscopic images, which were geometrically compensated using this calibration.

2. IMAGE FORMATION

Geometric image formation consists of two steps: The first step is a mapping of object coordinates given in the world coordinate system to the camera coordinate system using the extrinsic camera parameters. This step is identical for Kannala's and Bouguet's approaches. In the second step, where both approaches differ, object points in camera coordinates are mapped to image coordinates.

2.1. Mapping World Coordinates to Camera Coordinates

The transformation from the world coordinate system (denoted by subscript w) to the camera coordinate system (denoted by subscript c) can be described as a rotation and translation as both coordinate systems are orthogonal and equally scaled. This transform can be written in homogeneous coordinates as

$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} = \begin{pmatrix} \mathcal{R} & \vec{t} \\ \vec{0} & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \tag{1}$$

where $\mathcal{R}$ is a 3×3 rotation matrix and $\vec{t}$ is a translation vector.

2.2. Mapping Camera Coordinates to Image Coordinates

The pinhole camera model describes the radial part by

$$r(\theta) = f \tan\theta \tag{2}$$

where r is the distance from the principal point, f is the focal length and θ is the angle between the incoming ray and the optical axis. One reason for the limited accuracy of the pinhole model is that Eq. (2) diverges at $\theta = \pi/2$. Empirically, fish-eye lenses obey one of the following equations [13],

$$\begin{array}{lllllll} r(\theta) & = & 2f\tan(\theta/2) & & r(\theta) & = & 2f\sin(\theta/2) \\ r(\theta) & = & f\theta & & r(\theta) & = & f\sin(\theta) \end{array} \tag{3}$$

which are known as the stereographic, equidistance, equisolid angle and orthogonal projection equations, respectively. Rather than restricting the approach to one of the equations in (3), Kannala remains generic through expanding r in a Taylor series describing all of the projections above — including the pinhole model — as follows:

$$r(\theta) = \kappa_1\theta + \kappa_2\theta^3 + \kappa_3\theta^5 + \kappa_4\theta^7 + \ldots \tag{4}$$

where the κ_i are to be estimated by the calibration procedure. Thus, the pinhole model is merely a special case of Kannala's model.

3. 3D RECONSTRUCTION

In order to reconstruct the 3D position of the image points, it is necessary to estimate the relative camera movement and orientation between two image frames. Therefore, we estimate the fundamental matrix, which encodes this information among others. After that, we obtain each point's 3D position using triangulation.

3.1. Estimation of the Fundamental Matrix

As a first step to estimate the fundamental matrix, we need to find at least eight point correspondences between the two images. We obtain these correspondences using Lowe's scale invariant feature transform (SIFT) [14]. SIFT usually yields a large number of correspondences containing only few wrong correspondences.

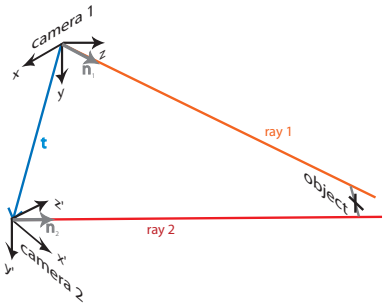


Fig. 1. Triangulation to recover 3D coordinates. Due to inaccuracies, the two rays do not intersect.

We use Hartley's normalized eight point algorithm [15] to compute a first estimate of the fundamental matrix $\mathcal{F}$. Due to the epipolar geometry, the following condition holds for each point correspondence ($\vec{u}_i$,$\vec{u}'_i$)

$$\vec{u}_i^T \mathcal{F} \vec{u}'_i = 0. \tag{5}$$

Rewriting equation (5) as scalar product and using many point correspondences one can build up a system of linear equations

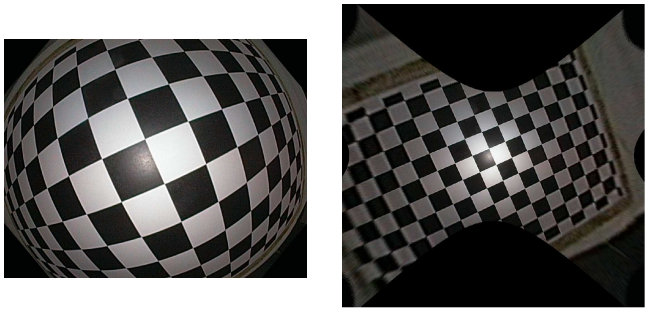
$$\mathcal{U}\vec{F} = \vec{0} \tag{6}$$

where each line of the matrix $\mathcal{U}$ consists of a combination of the coordinates of corresponding points $\vec{u}_i$ and $\vec{u}'_i$. $\vec{F}$ consists of the elements of $\mathcal{F}$. As solving this equation directly, using singular value decomposition, is numerically unstable, Hartley's proposition to first normalize the coordinates of the point correspondences leads to better results.

Wrong correspondences are removed using random sample consensus (RANSAC) [16]. Adapting RANSAC to our problem leads to the following steps:

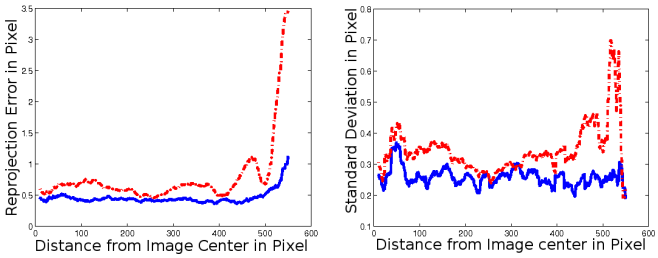
1. Randomly select eight point correspondences and estimate the fundamental matrix.
2. Calculate for remaining point correspondences their distances from the epipolar lines. Point correspondences with sufficiently small distances are called inliers. The remaining point correspondences are called outliers.
3. If the number of inliers is sufficiently large remove all outliers from the set of point correspondences and start again at 1. Otherwise, do not remove the outliers and start again at 1.
4. Repeat until the set of point correspondences converges or a maximum number of iterations is reached.

Finally, use all remaining point correspondences to estimate the fundamental matrix. The relative translation and orientation can now be easily computed [3].



(a) (b)

Fig. 2. (a) Distorted calibration image as captured with an endoscope with fish-eye lens. (b) Result of geometric correction using Kannala's model.



(a) (b)

Fig. 3. Comparison of the reprojection error of Bouguet's (dashed line) and Kannala's (solid line) calibration procedure. (a) Mean reprojection errors. (b) Respective standard deviations.

3.2. Triangulation

Once the orientation of the two cameras with respect to each other is known, it is straightforward to reconstruct the 3D coordinates of those points which can be observed in both images. However, remaining inaccuracies in the determination of point correspondences and of the position and orientation of the two camera frames will have the result that the two rays do not intersect with each other. In that case, we assumed the real position of the object to lie at the location which minimizes the sum of the squared distances to the two rays.

Fig. 1 depicts the geometry of the two camera coordinate systems and illustrates the described problem.

4. EXPERIMENTS

We acquired 23 images of a calibration chessboard with an Olympus "CF H-180 AL" state of the art HDTV video endoscope. Initial guesses of the positions of the chessboard corners were provided and refined to subpixel accuracy using Harris' corner detector [17]. The camera parameters for both camera models were estimated using Bouguet's and Kannala's reference implementations of their respective calibration procedures.

For 3D reconstruction, we designed a colon phantom [4]. The phantom consisted of a tube with a colored chessboard pattern on the inside, which made it possible to identify the absolute position of each corner unambiguously. We inserted the endoscope into the phantom and acquired a video sequence, from which two images were selected as input for the 3D reconstruction.

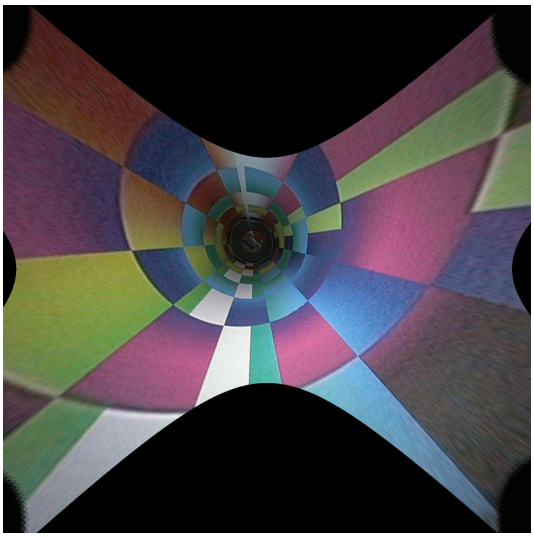
5. RESULTS

One of the calibration images without any corrections is shown in Fig. 2a. The result of geometric correction based on the model by Kannala and Brandt is depicted in Fig. 2b, where evidently only minor distortions remain. Note that the chessboard pattern is now clearly recognizable also at the chessbord border although these image parts are hardly visible in the distorted original. The black regions in the corrected image plane correspond to regions outside of the source image, where no information is available. Fig. 3a gives a quantitative comparison of the reprojection errors of Bouguet's (dashed line) and Kannala's (solid line) algorithms. The errors of the approach by Kannala and Brandt increase only slightly towards the image border, while the errors of the standard model increase sharply up to 8 pixels. The standard deviation shown in Fig. 3b is also lower in case of Kannala's approach.

Fig. 4 gives an example of the complete processing of our colon phantom. Fig. 4a shows the result of geometric distortion correction based on Kannala and Brandt's model. Fig. 4b provides an impression of the reconstructed phantom, which was generated as described above. Note that the phantom itself is not perfectly circular as can be seen on the right hand side of Fig. 4b.

6. DISCUSSION

We have analyzed image formation, camera calibration and distortion correction based on the pinhole model extended by distortion parameters and based on the recently published generic model of Kannala and Brandt. Both camera models were calibrated with images obtained from a state of the art HDTV endoscope using the same point correspondences. The reprojection error as well as its respective standard deviation of Kannala's model were consistently lower than the ones achieved with Bouguet's calibration algorithm. The difference between both models is especially significant at the image's rim regions, where the pinhole-model based algorithm fails, whereas the generic one provides results sufficiently accurate for 3D reconstruction even from data near the image borders. One such reconstruction example - obtained from a colon phantom - was also shown.



(a)

(b)

Fig. 4. Example of distortion correction of our colon phantom using Kannala's algorithm (a) and subsequent 3D reconstruction (b).

7. REFERENCES

[1] S. Kudo, S. Hirota, T. Nakajima, S. Hosobe, H. Kusaka, T. Kobayashi, M. Himori, and A. Yagyuu, "Colorectal tumours and pit pattern.," *J Clin Pathol*, vol. 47, no. 10, pp. 880–885, Oct 1994.

[2] S. Tanaka, S. Oka, M. Hirata, S. Yoshida, I. Kaneko, and K. Chayama, "Pit pattern diagnosis for colorectal neoplasia using narrow band imaging magnification," *Digestive Endoscopy*, vol. 18(s1), pp. 52–56, 2006.

[3] R. Hartley and A. Zisserman, *Multiple View Geometry*, Cambridge University Press, Falls Church, 2000.

[4] D. Truhn, "3D Reconstruction of Endoscopic Images," M.S. thesis, RWTH Aachen University and Imperial College London, 2006.

[5] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," pp. 221–244, 1992.

[6] J. Heikkilä and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *CVPR '97*, Washington, DC, USA, 1997, p. 1106, IEEE Computer Society.

[7] J.-Y. Bouguet, "Visual methods for three-dimensional modeling," *PhD Thesis*, 1999.

[8] C. Wengert, M. Reeff, P. C. Cattin, and G. Szkely, "Fully automatic endoscope calibration for intraoperative use," in *Bildverarbeitung für die Medizin*. March 2006, pp. 419–23, Springer.

[9] S. Rupp, C. Winter, and T. Wittenberg, "Camera calibration from fiberscopic views with accuracy evalution," in *Bildverarbeitung für die Medizin*. March 2006, pp. 424–428, Springer.

[10] C. C. Slama, *Manual of Photogrammetry, Fourth Edition*, American Society of Photogrammetry, Falls Church, 1980.

[11] E. Trucco and A. Verri, *Introductory Techniques for 3D Computer Vision*, Prentice-Hall, Upper Saddle River, 1998.

[12] D. A. Forsyth and J. Ponce, *Computer Vision. A Modern Approach*, Prentice-Hall, Upper Saddle River, 2003.

[13] J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1335–1340, 2006.

[14] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, 2003.

[15] R. Hartley, "In defense of the eight-point algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 133–135, 1997.

[16] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[17] C. Harris and M. Stephens, "A combined corner and edge detector," *Proceedings of The Fourth Alvey Vision Conference*, pp. 147–151, 1988.