Camera Calibration from Fiberscopic Views with Accuracy Evaluation

Stephan Rupp[1], Christian Winter[1,2] and Thomas Wittenberg[1]

[1]Fraunhofer-Institute IIS, Erlangen
[2]Chair of Information Technologies, University of Erlangen-Nuremberg, Erlangen
Email: ruppsn@iis.fraunhofer.de

Abstract. In many medical applications glass-fiber endoscopes are used to acquire images from complex hollows for diagnostic and interventional purposes. For a complete exploration and understanding of an antrum, a 3-D model of the scene might be desirable or yet necessary. In order to perform a 3-D reconstruction, in most cases a calibration of the imaging system is required. In this paper we apply an adaptive raster reduction algorithm and an interpolation technique in order to eliminate the typical comb structure of fiberscopes enabling the appliance of common camera calibration techniques. Moreover we present a model of the fiberscopic transmission line in order to synthetize fiberscopic images from rigid endoscope images enabling an evaluation of the calibration results with respect to the comb structure removal approaches.

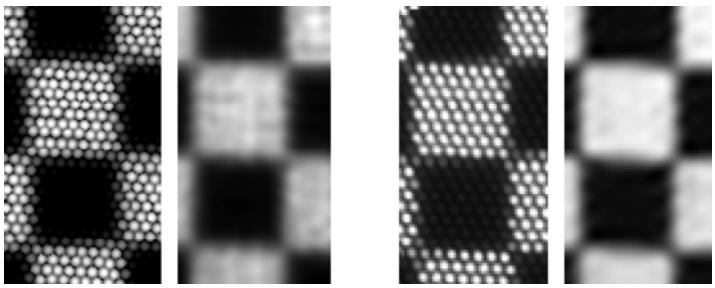
1 Problem Statement

Modern techniques for medical diagnostic and therapy in keyhole-surgery scenarios make use of flexible endoscopes. Their characteristic bendable image conductor consists of a very limited number of coated fibers which lead to the so called *comb structure*. When performing 3-D reconstruction from images, a calibration of the optical system is necessary which is usually performed by observing a special calibration object and extracting the image coordinates of projected calibration marks with high accuracy. When utilizing fiberscopes, the image conductor introduces a disturbing comb structure to the images that anticipate a robust reference point extraction that is required for a calibration.

2 Related Work and Contribution

Camera calibration has been studied intensivley in the past years, starting in the photogrammetry community and more recently in computer vision including accuracy evaluations for off-the-shelf cameras [1, 2, 3]. However, practically no camera calibration from fiberscopic views exists in literature nor any publications covering the impact of the comb structure removal on the calibration result.

Fig. 1. Zoomed image regions exhibit the performance of the rendered approximation: A synthesized fiberscopic view (a) in comparison to a real fiberscopic image (c) along with the filtered versions (b, d) after removing the comb structure.



In this work we enable the appliance of common camera calibration techniques to fiberscopes by applying an adaptive reconstruction algorithm as proposed by Winter et al. [4] and an interpolation scheme for eliminating the typical comb structure. Moreover, we enable an evaluation of the calibration by modelling the fiberscopic transmission channel in order to synthesize fiberscopic images from rigid endoscopes' images.

Our contribution provides the opportunity to quantify the influence of fiberscopic images on the calibration process and make an estimation regarding the over-all performance of a 3-D reconstruction computed from fiberscopic images.

3 Methods

When using fiberscopes in combination with cameras, a crude sub-sampling is made on the tip of the fiberscope, while an over-sampling takes place where the fiber-transmitted image is digitized (fig. 1(c)). This reveals a problem when trying to quantizise the performance of a calibration. Since we can only access the resampled image at the proximal end of the transmission line, we have no possibility to relate it to the unaffected image in front of the fiberscope's tip – in other words, we do not have any access to a *ground truth* and therefore no possibility to rate the performance of the calibration.

To approach this problem, we model the transmission and synthesize images originating from rigid endoscopes. By calibrating the rigid endoscope with its own images followed by a calibration of the same system with synthesized and filtered images, we gain access to a performance measurement of the calibration process.

3.1 Modelling the Image Conductor

The image transmission by a bundle of glass fibers is modelled in a five step process: At first, a *white image* of a glass fiber endoscope is needed in order to extract the fiber centers from the image. This is done by applying adaptive

thresholding technique along with a distance transform and a morphologic filtering. Once the fiber centers have been located, a delaunay triangulation is performed providing a regular triangle mesh. From the triangulation, the dual graph, the voronoi net, is taken to approximate the cladding of the real glass fibers. With this tesselation artifical glass fiber images can be synthesized for an abitrary image $\mathbf{I}$. Let $I(x, y)$ denote the pixel intensity at (x, y) in the arbitrary image and $\mathbf{S}$ the synthesized, fiberscopic image.

Since a single fiber only transmits one intensity information, all the intensities of the pixels within a voronoi cell $\mathcal{C}$ are collected and related to the fiber's capacity $I_{\text{cap}}(\mathcal{C})$, thus yielding the intensity information $I(\mathcal{C})$ that is transmitted:

$$I(\mathcal{C}) = \frac{1}{I_{\text{cap}}(\mathcal{C})} \sum_{(x,y)\in\mathcal{C}} I(x, y) \; . \tag{1}$$

Here, the fiber's capacity is defined as the maximum intensity that the cell $\mathcal{C}$ can transmit, that is the maximum intensity $I_{\max}$ cumulated over all the pixels within the cell:

$$I_{\text{cap}}(\mathcal{C}) = \sum_{(x,y)\in\mathcal{C}} I_{\max} \; . \tag{2}$$

According to the far field intensity distribution of a real index glass fiber [5], the intensity $I(\mathcal{C})$ associated to the cell $\mathcal{C}$ is weightened with an adaptive 2-D gaussian distribution, positioned at the fiber's center (c_x, c_y):

$$S_{\mathcal{C}}(x, y) = I(\mathcal{C}) \exp(-\frac{(x - c_x)^2 + (y - c_y)^2}{\sigma^2}) \quad \forall\, (x, y) \in \mathcal{C} \tag{3}$$

This procedure is performed for every voronoi cell yielding an artifical fiberscopic image $\mathbf{S}$ of the image $\mathbf{I}$ exhibiting the typical honeycomb structure (fig. 1(a)).

When dealing with endoscopic images, a global shading effect is observable being due to the non-uniform illumination source at the endoscope's tip [6]. This shading is modelled to be Gaussian across the whole image.

3.2 Comb Structure Removal

To eliminate the comb structure an adaptive filtering approach is applied. Based on the spectrum of the image, the Nyquist-Shannon Sampling Theorem is applied to obtain optimal parameters for the generation of a band-rejecting mask [4]. Additionally, we compare the results with a second, non-adaptive interpolation method. This makes use of the delaunay mesh containing the fibers' centers and performs an interpolation based on homogenous barycentric coordinates:

$$I(\mathbf{p}) = b_0\, I(\mathbf{c}_0) + b_1\, I(\mathbf{c}_1) + b_2\, I(\mathbf{c}_2) \quad \text{with} \quad b_0 + b_1 + b_2 = 1 \tag{4}$$

and

$$b_i = \frac{A_{\mathbf{c}_j,\mathbf{p},\mathbf{c}_k}}{A_{\mathbf{c}_0,\mathbf{c}_1,\mathbf{c}_2}} = \frac{||(\mathbf{c}_j - \mathbf{p}) \times (\mathbf{c}_k - \mathbf{p})||}{||(\mathbf{c}_1 - \mathbf{c}_0) \times (\mathbf{c}_2 - \mathbf{c}_0)||} \qquad \begin{matrix} i, j, k = 0 \ldots 2 \\ i \neq j, i \neq k, j \neq k \end{matrix} \tag{5}$$

The intensity $I(\mathbf{p})$ at a pixel $\mathbf{p} = (x, y)$ is calculated from the weighted sum of the intensities at the fiber centers $I(\mathbf{c}_0), I(\mathbf{c}_1), I(\mathbf{c}_2)$ of the circumscribed triangle. The *barycentric coordinate* b_i is defined by the ratio of the area of a subtriangle $(\mathbf{c}_j, \mathbf{p}, \mathbf{c}_k)$ and the area of the circumscribed triangle $(\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2)$.

3.3 Calibration

The characteristics of the imaging system are determined by the well-known calibration technique by Zhang [1] modelling the relationship between the 2-D pixel coordinates and 3-D world coordinates by a projection matrix $\tilde{\mathbf{P}}$, which maps points from the projection space $\mathcal{P}^3$ to the projective plane $\mathcal{P}^2$:

$$\tilde{\mathbf{P}} = \lambda_{\mathrm{w}} \underbrace{\begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} \mathbf{R}_{\mathrm{w}} & \mathbf{t}_{\mathrm{w}} \\ \mathbf{0}_3^T & 1 \end{pmatrix}}_{\mathbf{D}_{\mathrm{w}}}$$

The 3×3 matrix $\mathbf{A}$, whose four entries are called *intrinsic parameters* model the imaging process whereas the 4×4 displacement matrix $\mathbf{D}_{\mathrm{w}}$ describes the external orientation of the camera (*extrinsic parameters*).

4 Experiments And Results

For an accuracy evaluation, we calibrate several rigid endoscopes of different diameters and manufacturers. For any given endoscope we record $F = 10$ images of a 8-by-6 checkerboard ($N = 48$ calibration marks) from different directions and determine the *back-projection error* being defined by:

$$\epsilon = \sum_{i=1}^{N} \sum_{j=1}^{F} \epsilon_{ij}, \quad \epsilon_{ij} = \left\| \begin{pmatrix} u_{ij} \\ v_{ij} \\ 1 \end{pmatrix} - \tilde{\mathbf{P}} \cdot \begin{pmatrix} X_{ij} \\ Y_{ij} \\ Z_{ij} \\ 1 \end{pmatrix} \right\|_2 \tag{6}$$

The back-projection error of a single calibration feature ϵ_{ij} is defined by the Euclidean distance between its initially extracted image coordinates $(u_{ij}, v_{ij})'$ and the corresponding 3-D world coordinates $(X_{ij}, Y_{ij}, Z_{ij})'$ being projected back to the image plane with the calibrated projection matrix $\tilde{\mathbf{P}}$. The sum of all these errors yields an appropriate measure for the quality of the calibration [1].

In a further step, we synthesize fiberscopic images from the acquired images and apply the adaptive filtering method as well as the barycentric interpolation for the removal of the comb structure that is followed by a calibration. The results for an exemplary rigid endoscope with a diameter of $10.0mm$ and a viewing angle of $70°$ are depicted in table 1.

The results exhibit that a camera calibration from fiberscopic views is basically feasible and that the back-projection error is in principle four to five times as high as in the non-fiberscopic case as expected due to the corresponding subsampling factor between the fibers' and sensor resolution. With respect to the

Table 1. The performance of the back-projection error for a calibration from synthesized, fiberscopic images. The unit of the error is in pixel whereas the bracketed values determine the error related to a standardized image dimension of 640x480 pixel.

Image Number	Rigid Endoscope	Adaptive Filtering	Barycentric Interpolation
1	0.4197 (0.2623)	1.1703 (0.7313)	1.6954 (0.9578)
2	0.4788 (0.2993)	1.5511 (0.9694)	1.2076 (0.7547)
3	0.4347 (0.2712)	1.4938 (0.9337)	1.2326 (0.7703)
4	0.3160 (0.1975)	1.2232 (0.7645)	1.4122 (0.8826)
5	0.4906 (0.3066)	1.6628 (1.0394)	1.4434 (0.9025)
6	0.4358 (0.2723)	1.2456 (0.7785)	1.3522 (0.8345)
7	0.3904 (0.2401)	1.3318 (0.8324)	1.3437 (0.8398)
8	0.3514 (0.2197)	1.2200 (0.7625)	3.5252 (2.2032)
9	0.3270 (0.2044)	1.3219 (0.8262)	1.5217 (0.9512)
10	0.4128 (0.2580)	1.4484 (0.9053)	2.3997 (1.4998)
Mean error	0.4058 (0.2536)	1.3669 (0.8543)	1.6954 (1.0597)

algorithm the less complex interpolation method performs slightly worse than the adaptive filtering approach caused by the algorithm for the extraction of the calibration marks.

5 Discussion and Conclusion

In the proposed work, we model the transmission channel of fiberscopes in order to synthesize fiberscopic images from images obtained with rigid endoscope. We apply this model along with comb-structure removal algorithms in order to facilitate and evaluate the camera calibration from fiberscopic views. First results show that the calibration of fiberscopes is feasible and - as expected - the results perform slightly worse compared to the calibration of a rigid endoscope.

This insight of well functional calibration environments hopefully helps to push many medical applications requiring the flexibility of fiberscopes.

Acknowledgement. This work was supported by the German Research Foundation (DFG) in the context of Graduiertenkolleg 244 and SFB 603 (TP A7)

References

1. Zhang Z. A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 2000;22(11):1330–1334.
2. Heikkilä J. Geometric Camera Calibration Using Circular Control Points. IEEE Transactions on Pattern Analysis and Machine Intelligence 2000;22(10):1066 – 1077.
3. Tsai RY. A Versatile Camera Calibration Technique for High-Accuaricy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. IEEE Transactions on Robotics and Automation 1987;4:323 – 344.
4. Winter C, et al. Automatic adaptive removal of fiberscopic comb-structure by spectral masking. Computer Assisted Radiology and Surgery 2005;.
5. Young M. Optics And Lasers. Springer, Heidelberg; 2000.
6. Relling J, et al. Technische Endoskopie. Expert, Malmsheim; 2001.