Fully Automatic Endoscope Calibration for Intraoperative Use

Christian Wengert, Mireille Reeff, Philippe C. Cattin
and Gábor Székely

Computer Vision Laboratory, ETH Zurich, 8092 Zurich, Switzerland
{wengert,reeff,cattin,szekely}@vision.ee.ethz.ch

Abstract. As of today endoscopes have been only used as a keyhole to look inside the human body. Our goal is to enhance the endoscope to a full imaging device providing better quantitative and qualitative data. Possible applications for such an enhanced endoscope are referencing, navigation and 3D visualization during endoscopic surgery. To obtain accurate results, a reliable and fully automatic calibration method for the endoscopic camera has been developed which can be used within the operating room (OR). Special care has been taken to ensure robustness against inevitable distortions and inhomogeneous illumination.

1 Introduction

Real-time quantitative measurements and 3D visualization during endoscopic surgery can provide the clinician with valuable additional information. Such an enhanced endoscope allows for endoscopic navigation and referencing, updating preoperative models, visual augmentation, and finally better depth perception for the surgeon, as structures on the border of the image are no longer distorted. For all of these applications it is essential to perform a pre- or even intraoperative camera calibration to obtain accurate and reliable results. This is more challenging inside the OR as under usual laboratory conditions, as the sterilization should not be affected and the required time has to be minimal. The main difficulties are the inhomogeneous illumination and the severe lens and projective distortion due to the wide-angle optics. Depending on the kind of surgery, endoscopes are either used in liquid or in air which has also to be taken into account during calibration, as the intrinsic parameters such as the focal length vary in different mediums.

2 State of the art

Several generic camera calibration methods have been reported in the literature, but one of the most commonly used is the implementation of [1] using the camera model of Heikkilä [2]. It is based on extracting the corners of a chessboard and matching these 2D image points to the corresponding 3D points of the chessboard. Although being straightforward, it needs some experience and time to

get reasonable calibration results, especially with highly distorted endoscopic optics. Other authors [3, 4] studied endoscope calibration, but there was neither an improvement on the accuracy compared to [2] nor did they explicitly address the usage inside the OR. [5] used a "center dot" to determine the orientation of the dot grid. The center dot detection is based on the relative radius (classification in large dots vs. small dots), which could be problematic in the presence of severe distortions.

3 Methods

Several aspects of performing camera calibration in the OR are crucial and have been largely ignored in the literature. Preserving the sterility, the ease of use, inhomogeneous illumination conditions and the generally very severe lens distortion need to be investigated carefully and taken properly into consideration to guarantee a successful application in the OR. A typical image is illustrated in Figure 1, where the inhomogeneous illumination and the lens distortion can be seen. Instead of the commonly used chessboard we designed a new pattern consisting of a grid of dots and two special marks for identifying the local coordinate system. The grid is made of aluminum and the pattern is oxidized onto its surface, thus withstanding any sterilization method. Additionally, we built a sterilizable mechanical system to position the endoscope rigidly over the calibration plate for tremor-free images (only needed for interlaced cameras). The procedure for calibrating the system is straightforward: The surgeon needs to take at least 2 images, although 15 images are recommended, of the calibration grid from different viewpoints. Except the endoscope's light source, no additional illumination is required. Once the images are acquired, the calibration is performed fully automatically. The images are processed, the points extracted and matched against a virtual representation of the grid points to establish the necessary 2D/3D correspondences, which allow to calibrate the camera using the camera model from Heikkilä.

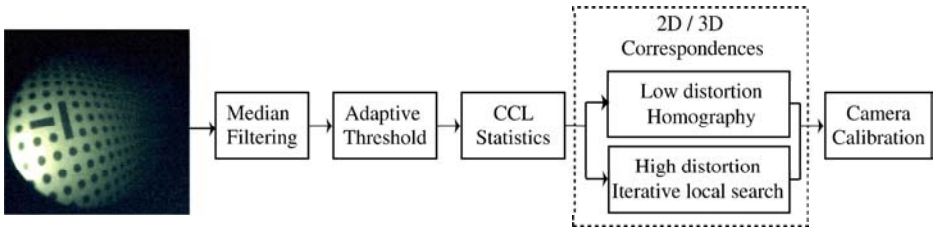
3.1 Hardware

The hardware used for testing our framework consists of two different optics: A nearly distortion-free 10 mm/25° and a highly distorted 3.8 mm/12° laparoscope. Both were used together with a progressive scan IEEE1394 color camera with a resolution of 640x480 pixels and a 2/3" CMOS chip. A C-Mount adapter allows to connect the optics to the camera. A 250W medical light source for endoscopes was used for illuminating the grid.

3.2 Algorithm

After denoising the image with a 3×3-median filter, we use a low-pass filter to determine a local threshold for each pixel. This allows to cope with the varying illumination and to extract a maximum number of dots. From the binary image,

Fig. 1. Algorithm outline, starting from the raw image (note the two orientation marks on the left)



all blobs are extracted using connected component labelling. Each blob is then characterized with statistical and geometric properties such as its ellipticity, solidity, area, and the center of gravity, serving as the grid point location. The solidity is defined as the ratio of a blob's area and the area of its convex hull, whereas the ellipticity is the ratio of its major and minor axis length.

In a first stage all unusable blobs with a solidity less than 0.80 and the very small blobs, caused by image noise, are eliminated. Next, the two special marks are recognized, as their ellipticity is larger than 2, whereas normal dots have an ellipticity close to 1. The area difference of the two marks allows us to discriminate between the larger (origin of the local coordinate system) and the smaller mark. The orientation of the special marks gives us two directions to start searching for adjacent points. By assuming that the local lens distortion is small from one grid point to the other, we adapt the search direction and radius after every matched grid point. The new search radius is defined by $r_{n+1} = 1.2 \times \|\vec{x}_n - \vec{x}_{n-1}\|$ and the new search direction for finding the next point is defined as $\vec{d}_{n+1} = (\vec{x}_n - \vec{x}_{n-1})/\|\vec{x}_n - \vec{x}_{n-1}\|$ where $\vec{x}_n$ are the image coordinates of the grid point at step n. If a point $\vec{x}_{n+1}$ is within a distance of r_{n+1} and the angle between $\vec{x}_{n+1} - \vec{x}_n$ and $\vec{d}_{n+1}$ is smaller than 25°, this point is matched. Otherwise the search procedure in this direction is stopped. For low distortion optics, the correspondence establishment can be greatly speeded up once 4 non-collinear points have been matched. The image coordinates of these matched points can then be used to calculate a projective transformation between the points and the their grid coordinates. The missing correspondences can then be very efficiently established using this transformation.

Once all correspondences have been found, we compute a final projective transformation between all matched point pairs. The residual error allows then to eliminate images where the correspondence establishment failed. Even for severely distorted images, the error of a wrongly matched image is at least one order of magnitude higher than for a correctly matched one. Finally, the camera's intrinsic parameters are computed from these matches using Heikkilä's model. In order to refine the calibration we eliminate points having a backprojection error of more than one pixel and recalibrate with the new set of points.

4 Results

We performed 20 calibrations for each endoscope with 15 images each and reached a steady mean backprojection error of 0.25 pixels with a standard deviation of 0.1 pixels for the 10 mm optics. The 3.8 mm optic with the high lens distortion performs slightly inferior with a backprojection error of 0.35 ± 0.1 pixels. Using an additional external light source, the registration error could be slightly improved to 0.2 ± 0.1 pixels for the 10 mm laparoscope.

The influence of a different medium other than air has also been investigated. As expected, the intrinsic parameters change. However, the backprojection error for the camera calibration in water were in the same range as those in air.

The processing time for extracting the points and establishing the correspondences is less than one second per image. Thus, the actual time needed for calibrating mainly depends on the time required to grab the calibration images. With a progressive scan camera, the images can be made dynamically and the surgeon only has to take several shots of the calibration grid requiering less than 1 minute. In case of an interlaced camera, the endoscope has to be static for every shot and repositioned afterwards, which generally takes around 3 minutes.

In comparison, the 3.8 mm laparoscope has been calibrated by hand using a chessboard and the standard calibration toolbox [1]. The best backprojection errors where more than 0.5 pixels and one calibration took roughly one hour, as all the corner points had to be extracted manually.

4.1 External tracking

If an external tracking device is used to report the camera's position and orientation, we need to compute the transformation between the world coordinate system of the tracking device and the camera's internal coordinate system lying on the image plane. We attached an active infrared marker with a diameter of 50mm to the endoscope which has been tracked using an EasyTrack 500 from Atracsys LLC. In order to find the relationship between the tracker, resp. the marker and the camera, we propose an add-on to our calibration framework which allows to establish the missing relationship in a second step based on the camera calibration. The latter allows to compute the camera pose in the grid's local coordinate system which is defined by the two special marks. We can establish an equation of the type $A \cdot X = X \cdot B$ where X is the transformation between the marker and the camera, A is the movement of the marker relative to the tracking device and B is the camera motion relative to the calibration grid frame. A similar problem exists in the robotics community, where the calibration between a robot's end-effector and a camera attached to it is known as *Hand-Eye calibration* [6, 7, 8, 9].

The Hand-Eye calibration was performed using [6] and the error has been measured by backprojecting the 3D grid points computed in the tracker coordinate system. In our setup the marker was positioned 150 mm from the tip of the endoscope. This resulted in a backprojection error of 3 ± 1 pixels. The backprojection error depends from the distance between the tool center point (TCP)

and the marker and the size of the marker itself. As has been shown in [10], the rotational error of the tracker measurements defined by the marker size has a much bigger influence on the backprojection error compared to the translational error. In addition, the error linearly depends on the distance of the marker to the TCP.

5 Conclusion

We propose a fast, reliable and completely automatic endoscope calibration system for in-OR use, which can also be used in different mediums other than air. This allows to remove the lens distortion in order to quantitatively analyze and process endoscopic images. The precision we achieved is better and the process is much faster than doing it manually. With our method, the endoscope can be calibrated before or even during the surgery and neither an engineer nor a technician needs to be present, thus allowing to use an "enhanced"endoscope inside the OR. The system has been tested under lab conditions and inside the OR showing equally good performance.

Acknowledgment. This research has been supported by the NCCR CO-ME of the Swiss National Science Foundation.

References

1. Bouguet JY. Camera Calibration Toolbox for Matlab.
Internet: www.vision.caltech.edu/bouguetj, last visit: 08.12.2005. Available from: www.vision.caltech.edu/bouguetj/calib_doc/
2. Heikkilä J, Silven O. A Four-step Camera Calibration Procedure with Implicit Image Correction. In: Conference on Computer Vision and Pattern Recognition. IEEE Computer Society; 1997. p. 1106.
3. Zhang ZY. Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. In: International Conference on Computer Vision; 1999. p. 666–673.
4. Zhang C, et al. Nonlinear distortion correction in endoscopic video images. In: International Conference on Image Processing; 2000. p. 439–442.
5. Shahidi R, et al. Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. IEEE Trans Med Imaging 2002;21(12).
6. Tsai RY, Lenz RK. A New Technique for Fully Autonomous and Efficient 3D Robotics Hand Eye Calibration. IEEE Transactions on Robotics and Automation 1989.
7. Daniilidis K. Hand-Eye Calibration Using Dual Quaternions. International Journal of Robotics Research 1999;18(3):286–298.
8. Chou JCK, et al. Quaternions Approach to Solve the Kinematic Equation of Rotation of a Sensor-Mounted Robotic Manipulator. In: IEEE International Conference on Robotics and Proceedings Automation. vol. 2; 1988. p. 656–662.
9. Park FC, et al. Robot Sensor Calibration: Solving AX=XB on the Euclidean group. IEEE Transactions on Robotics and Automation 1994;10(5):717–721.
10. Bianchi G, Wengert C, et al. Camera-Marker Alignment Framework and Comparison with Hand-Eye Calibration for Augmented Reality Applications. In: IEEE and ACM International Symposium on Mixed and Augmented Reality; 2005.