Hybrid method for both calibration and registration of an endoscope with an active optical tracker

Gaëtan Marti[a], Vincent Bettschart[b], Jean-Sébastien Billiard[c], Charles Baur[a]

[a] *VRAI Group, LSRO, IPR, Swiss Federal Institute of Technology (EPFL), Switzerland*
[b] *Division of General Surgery, University Hospital of Sherbrooke (CHUS), Quebec, Canada*
[c] *Department of Radiology, University Hospital of Sherbrooke (CHUS), Quebec, Canada*

Abstract

In this paper, we present a hybrid method for calibration of an endoscope and its registration with an active optical tracker. Practically, both operations are done simultaneously by moving an active optical marker in the field of view of the two devices. By segmenting image data, the LEDs composing the marker are extracted and the transformation matrix between the two referentials (homography) is calculated. By reformulating the calibration problem, registration and calibration parameters are extracted from the homography. As camera calibration and registration is an indispensable step for augmented reality or image guided applications, this technique can easily be used in the operating field because it is fast, accurate and reliable. We currently are using this technique with an augmented reality system for laparoscopic procedures.

Keywords: calibration, registration, augmented reality, image-guided surgery

1. Introduction

In one and a half decade, laparoscopic surgery has gained wide acceptance in both the surgical community and the general population. The obvious advantages of a minimal aggression, including the preservation of the abdominal wall integrity, have given to laparoscopic approaches a widespread use. Operations such as cholecystectomy, appendicetomy, adrenalectomy, colectomy are performed mostly laparoscopically.

However some intrinsic limitations of the technique are not been overcame. The perception of the depth is lost, as well as the tactile information. The degree of freedom of the laparoscopic tools is much less than with a hand in open surgical field. Therefore, dissection during laparoscopic surgery is often somewhat hazardous and takes the longest part of the operation.

Many operations, like liver or pancreatic resections, cannot yet be habitually performed by laparoscopy. In order to be able to deal safely with such highly vascularized organs, there is a need for a positioning system, allowing the surgeon to know where the relevant structures are, sparing the dissection time. This implies the development of augmented reality tools specifically designed for laparoscopic surgery.

Camera calibration is an indispensable step for augmented reality or image guided applications. Among the different calibration techniques, the photometric calibration is current the most used in the medical field [1-6]. It consists of observing a calibration object, like a chessboard or a planar grid, whose geometry in the 3-D space is known with a good precision. This technique uses some snapshot representing the calibration object at different poses and extracts object's features in the image, to finally correlate them with the 3-D model of the reference object. A calibration algorithm then extracts the homography matrix and the camera intrinsic parameters (focal lengths, optical centre, skew, and optionally radial distortions parameters), as well as extrinsic parameters (rotation and translation of the calibration object relatively to its model coordinates).

Registration of a medical camera relatively to an optical tracker is also an important step for image guided applications. On way to perform this operation consist on finding some common cues visible simultaneously by the two devices.

In this paper, we present a technique using an active optical marker as the calibration and registration object. By reformulating the calibration problem, it is so possible to express the extrinsic parameters as the registration matrix. Calibration and registration are done simultaneously by moving an active marker both visible by the optical tracker and the endoscope. This technique can easily be deployed in the operating field because it is fast, accurate and reliable.

2. Methods

Different optical trackers like *easyTrack* [7] or *Polaris*® [8] are using active markers composed of luminescent diodes (LEDs) that are emitting in the near infrared (~ 850 nm). Although, this wavelength is not visible by human eyes, most classic camera sensors are sensitive to the near infrared and can detect without difficulties LEDs' flashes (figure 1). Active markers, taken at different positions, can thus be used as calibration objects during the procedure.

The calibration procedure consists of placing an emitting marker on the endoscope and another one in its field of view. Both markers should be visible by the optical tracking device but only the free one has to be seen by the camera. When moving this marker, the tracker records the 3-D position of its LEDs. Endoscope images are simultaneously acquired and the corresponding 2-D positions of the LEDs are extracted via segmentation. A resampling of the data is usually necessary due to the acquisition rate difference of the two devices.

The first calibration step finds the linear transformation from the tracker coordinates (X_i, Y_i, Z_i) to the endoscope image coordinates (u_i, v_i). Using a homogeneous 3 x 4 matrix representation for matrix A the following equation can be written:

$$\underbrace{\begin{pmatrix} u_i \\ v_i \\ 1 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}}_{A} \cdot \underbrace{\begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}}_{X} \tag{1}$$

The matrix **A**, also called homography, has 11 degrees of freedom (3 rotations, 3 translations, 5 intrinsic parameters) and has an arbitrary scale factor involved, so one of the entries can be set to 1 without loss of generality (i.e. $a_{xy} = a_{xy} / a_{34}$).

The direct linear transformation, initially developed by Y. Abdel-Aziz [9] is used to solve this system. Assume that the camera is pinhole without radial distortions, the correspondence between the image points and the tracker points is the following projection:

$$\begin{cases} u_i = \dfrac{a_{11}X_i + a_{12}Y_i + a_{13}Z_i + a_{14}}{a_{31}X_i + a_{32}Y_i + a_{33}Z_i + a_{34}} \\ v_i = \dfrac{a_{21}X_i + a_{22}Y_i + a_{23}Z_i + a_{24}}{a_{31}X_i + a_{32}Y_i + a_{33}Z_i + a_{34}} \end{cases} \tag{2}$$

For N corresponding points (N ≥ 6), the equation 2 can be rewritten:

$$\underbrace{\begin{pmatrix} -X_1 & -Y_1 & -Z_1 & -1 & 0 & 0 & 0 & 0 & u_1X_1 & u_1Y_1 & u_1Z_1 & u_1 \\ 0 & 0 & 0 & 0 & -X_1 & -Y_1 & -Z_1 & -1 & v_1X_1 & v_1Y_1 & v_1Z_1 & v_1 \\ & & & & & & \vdots & & & & & \\ -X_N & -Y_N & -Z_N & -1 & 0 & 0 & 0 & 0 & u_NX_N & u_NY_N & u_NZ_N & u_N \\ 0 & 0 & 0 & 0 & -X_N & -Y_N & -Z_N & -1 & v_NX_N & v_NY_N & v_NZ_N & v_N \end{pmatrix}}_{L} \cdot \underbrace{\begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{33} \\ a_{34} \end{pmatrix}}_{a} = 0 \tag{3}$$

Where **L** is a 2N x 9 matrix and matrix **a** is the matrix **A** rewritten as a column vector. The well known solution of this over constrained homogenous system is the right singular vector of **L** associated with the smallest singular value (or equivalently, the eigenvector of $\mathbf{L}^{\mathrm{T}}\mathbf{L}$ associated with the smallest eigenvalue).

This approximate linear solution for **L** is used as the starting point for a non-linear minimisation of the difference between the measured and the projected points. This optimization will take into account the radial distortions and the camera real model:

$$\min_A \sum_{i=1}^{N} \left\| \underbrace{\begin{pmatrix} u_i \\ v_i \\ 1 \end{pmatrix}}_{measure} - \underbrace{A \cdot \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}}_{projection} \right\|^2 \tag{4}$$

Extraction of intrinsic and extrinsic parameters from **L** is performed using either Faugeras method [10] or QR decomposition [11] depending on the number of points available. For example, using the first method, we have:

$$A = \underbrace{\begin{pmatrix} -f_x & 0 & u_c \\ 0 & -f_y & v_c \\ 0 & 0 & 1 \end{pmatrix}}_{I} \cdot \underbrace{\begin{pmatrix} \vec{r}_1^{\,t} & \\ \vec{r}_2^{\,t} & \vec{t} \\ \vec{r}_3^{\,t} & \end{pmatrix}}_{E} = I \cdot E = \begin{pmatrix} -f_x \vec{r}_1^{\,t} + u_c \vec{r}_3^{\,t} & -f_x t_x + u_c t_z \\ -f_y \vec{r}_2^{\,t} + v_c \vec{r}_3^{\,t} & -f_y t_y + v_c t_z \\ \vec{r}_3^{\,t} & t_z \end{pmatrix} \qquad (5)$$

I is the intrinsic matrix, f_x and f_y are the focal length in the horizontal and vertical directions, (u_0, v_0) is the optical centre. Note that this technique assumes that the skew (i_{12}) is equal to 0. **E** is the extrinsic matrix (rotation and translation). Parameters can easily be extracted from the Faugeras decomposition (equation 5). Other decomposition method, like absolute conics, can alternatively be used to extract the intrinsic and extrinsic parameters.

Depending on the optics distortions, radial correction may be necessary. The reader is referred to [5] for a more elaborated discussion about this topic.

The resulting extrinsic parameters are expressed in the referential of the tracker. The final step consists of formulating this transformation in the referential of the marker fixed on the endoscope. This final transformation is the searched registration. For augmented reality applications, the resulting intrinsic parameters can optionally be used to set the virtual camera and viewport settings.

3. Results

We currently are developing an augmented reality system for laparoscopic procedures. Our system allows displaying preoperative 3D models or real-time ultrasound data of the liver in the endoscopic image during minimal invasive interventions. The proposed method is currently used to calibrate and register a 10 mm endoscope with an optical tracker easyTrack 500 [7].

The segmentation of endoscopic images is based on a growing region method (figure 1b). LED's blobs are sorted depending on their aspect ratio, size, mean intensity and neighbour pixels variance intensity attributes which are statistically pertinent. The LED position used to compute the homography is the barycentre of the blob. To improve detection reliability and speed, the initial homography can be guessed with the first 4 points correspondences by mean of either the described method or a 2D-3D pose estimation technique [12]. This estimate gives a cue of the approximate position of the LEDs in the images and allows improving the search success and speed.

In our system, about 20 LEDs positions per second can be processed with this technique and calibration starts with about 150 samples acquired in the full field of view of the camera. Some of these samples contain badly segmented centres and should be filtered before homography calculation. First, the homography is estimated with the whole samples using equation 3. Then, the distance between the measured and the projected points indicates if the sample is an outlier. If so, it is removed and the homography is estimated again.

Calibration accuracy is closed to conventional grids methods but is faster and do not require an extra calibration object in the operating field. Results has been compared with the gold-standard "Camera Calibration Toolbox for Matlab" [6] and do not differ for more than one percent. Total time for data acquisition, camera calibration and registration is less than one minute using a modern laptop. This method is about twice faster than the system we previously have developed and which needs a calibration grid.

We currently are validating our technique with different kind of endoscopes, as well as, medical microscopes. For example, figure 2 presents a test of augmented reality we have performed during a kidney transplant on living donor. The 3D preoperative model was at bootstrap manually and rigidly registered. The endoscope allows keeping the rigid registration correct while moving. We also plan to quantitatively compare our algorithm with the computer vision gold standard calibration algorithms for early 2004 and to adapt a non-rigid registration algorithm on the augmented reality application.

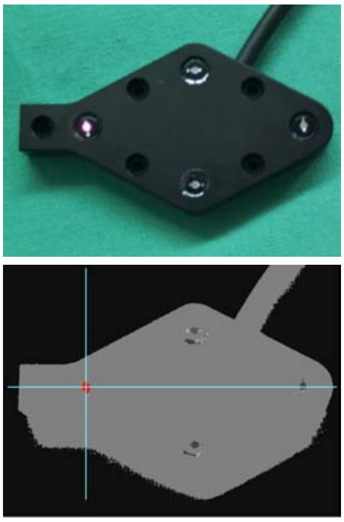

Figure 1: (a) Marker and (b) Segmented marker. Example of illuminated LED (left) and corresponding region-based segmentation.

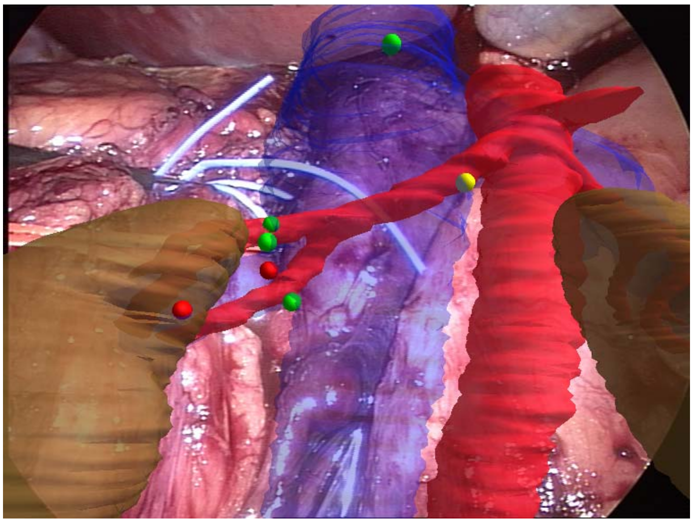

Figure 2: Proposed technique used in an augmented reality application for kidney transplant on living donors. The initial 3D model rigid registration is done manually and is kept while moving the endoscope.

4. Conclusion

We present a new method for both calibration and registration of camera images with an active optical tracker. By exploiting the fact that markers' LEDs are both visible by the endoscope and the optical tracker, the calibration problem can be reformulated to give simultaneously the registration of the camera relatively to the tracker and the intrinsic parameters of the camera. Compared to existing techniques using calibration objects, the proposed approach is ergonomic because medical markers are designed for sterilization and the LEDs can easily be located in the images. This new combined method will also improve time saving during clinical use because technical preparation with this technique is short.

5. References

[1] R. Tsai, An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, 1986, pp. 364–374.

[2] R. Tsai, A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses, IEEE Journal of Robotics and Automation, Vol. RA–3, No. 4, August 1987, pp. 323–344.

[3] B. Horn, Tsai's camera calibration method revisited, Copyright 2000, http://www.ai.mit.edu/people/bkph/papers/tsaiexplain.pdf.

[4] Z. Zhang, Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. International Conference on Computer Vision *(ICCV'99)*, Corfu, Greece, September 1999, pages 666-673.

[5] Z. Zhang, A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.

[6] J.-Y. Bouguet, Camera Calibration Toolbox for Matlab, last updated 2004, http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

[7] *easyTrack 200* and *500*, products of Atracsys sàrl, www.atracsys.com

[8] *Polaris*®, product of Northern Digital - NDI, www.ndi.ca

[9] Y. Abdel-Aziz, H. Karara, Direct linear transformation into object space coordinates in close-range photogrammetry, Proc. Symposium on Close-Range Photogrammetry, 1971, Urbana, Illinois, p. 1-18.

[10] O. Faugeras, G. Toscani, Camera calibration for 3D computer vision, Proc. International Workshop on Industrial Applications of Machine Vision and Machine Intelligence, Silken, Japan, 1987, p. 240-247.

[11] QR Decomposition, http://mathworld.wolfram.com/QRDecomposition.html

[12] S. Price, Solving for Perspective Inversion, last updated 1996, http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MARBLE/high/pia/solving.htm