Flexible Calibration of Actuated Stereoscopic Endoscope for Overlay in Robot Assisted Surgery

Fabien Mourgues and Ève Coste-Manière

Chir Team,
www.inria.fr/chir
Inria, 2004 route des lucioles,
06902 Sophia-Antipolis Cedex, France
{fabien.mourgues,eve.coste-maniere}@inria.fr

Abstract. Robotic assistance have greatly benefited the operative gesture in mini-invasive surgery. Nevertheless, the surgeon is still suffering from the restricted vision of the operating field through the endoscope. We thus propose to augment the endoscopic images with preoperative data.
This paper focuses on the use of a priori information to initialise the overlay by a precise calibration of the actuated stereoscopic endoscope. The flexibility of the proposed method avoids any additional tracking system in the operating room and can be applied to other augmented reality systems. We present quantitative experimental calibration results with the da Vinci^{TM} surgical system, as well as the use of these results to initialise the overlay of endoscopic images of a plastic heart with a coronary artery model.

1 Introduction

Mini-invasive surgery presents many advantages for the patient: the incisions and the trauma are reduced, the post-operative recovery time is limited. It takes now advantages of robot assistance: the endoscope and the instruments are controlled by the surgeon through a master console. The comfort and the precision of the gesture are improved. In the da Vinci robot system, a stereoscopic endoscope provides the surgeon with an immersive 3D vision of the operating field thus providing good depth perception. However it still suffers from the classical constraints of the mini-invasive surgery: the keyhole vision is confined and greatly increases the difficulty of locating landmarks necessary for the operation. For example in the coronary artery bypass interventions, two difficulties are present: distinguishing the target artery among the grooves of the other arteries and locating the position of the stenosis on the latter. During the preoperative planning the surgeon visualises this information on the 3D+t model of the coronary tree built from angiography acquisitions [1] but the transposition in the endoscopic view of the operating field is not easy.

T. Dohi and R. Kikinis (Eds.): MICCAI 2002, LNCS 2488, pp. 25–34, 2002.
© Springer-Verlag Berlin Heidelberg 2002

Thus, the enhancement of the surgeon's vision with pre or intraoperative information has become an important research challenge. It has been addressed in classical laparoscopic procedures [2] by restoring the surgeon's natural point of view, or in neurosurgery [3], by mapping endoscopic images on a preoperative brain surface. Some results of superposition in the laparoscopic view are given in [4] but no study concerns the augmentation of the 3D vision of the operating field in robot assisted surgery.

In coronary artery bypass interventions, the superposition of cartographic information (labelled arteries, direction to the target) requires new research efforts as described in [5]. The operating conditions - left lung collapse, CO_2 insufflation in the rib cage to provide space for the intervention, pericardium opening, organs motion - increase the difficulty of the registration with preoperative acquisitions (angiography, CT-scan).

We propose to address the problem in two steps:

1. use a priori information by registering the intraoperative endoscopic view and **the patient model in the preoperative situation**. This step needs:

- the calibration of the actuated endoscope (cameras parameters and position of the camera frames with respect to the actuated arm) to localise the latter in the robot frame.
- the **external** rigid registration of the patient in the operating room with respect to the preoperative acquisitions and the registration of the robot with respect to the patient. This step is solved using fiducial markers on the patient skin, segmented in CT-scan images and pointed by the robot arm [6].

This first step should give a good initialisation of the intraoperative overlay.

2. use of intraoperative acquisition. The heart deformations, displacements and motions must be considered to improve the coherency of the augmentation of the endoscopic images. The stereoscopic endoscope simplify the extraction of 3D information [7] with no need of structured light [2] and we investigate the use of appropriate landmarks for our cardiac application. The simultaneous use of other intraoperative acquisitions modalities (eg. ultrasound) could help the registration.

This paper mainly addresses the first step of the overall intraoperative registration problem: the calibration of the actuated stereoscopic endoscope by a new flexible method compatible with the operating room constraints. This method can be applied to any Augmented Reality (AR) system with cameras attached to a sensor localised with respect to a fixed reference. It presents some experimental calibration results and finally illustrates the first use of these results to initialise an endoscopic overlay.

2 Related Works and Contributions

Our endoscopic overlay system can be decomposed as a classical video-see through AR system: the two endoscopic cameras delivering the view of the real

scene on two display screens (for a 3D vision of the operating field) and a localisation system. The orientation and position of a frame attached to the rigid endoscope is continuously computed from the articular coordinates of the robot. The problem of calibration can be divided into two parts: the determination of the two camera models and the estimation of the orientation and position of the cameras with respect to the localised frame (figure 1). The camera calibration is now well studied [8,9,10] and applied to the endoscopes [11,4]. On the other hand, the calibration of see-through display tracked with magnetic or optical technologies have been addressed for AR applications. In [12] a video see-through system is calibrated using a 3D calibration object. The object and the Head-Mounted Display are localised with an optical tracker and the calibration is performed with Tsai's method [8]. The same method is used in [4] for a laparoscope. In [13] the calibration pattern is reduced to a single 3D point fixed to the tracker transmitter box. The method is active and adapted to the stereo case.

In our specific problem, the localisation of the endoscope is given by the robot arm. We developed a calibration method based on a planar grid [10] which avoids any additional tracking system to localise the pattern. This pattern, in an *arbitrary* position, is observed with the actuated endoscope from unknown points of view. The position of the grid with respect to the robot is an output of the algorithm. The calibration protocol is flexible, in one step (camera parameters including radial distortion, endoscope tip - camera transformation) and easily usable in the operating room.

3 Material and Methods

After a description of the endoscope model and the protocol, the two parts of the calibration process are developed: the first one consists of the determination of the geometry of the two cameras: intrinsic parameters and radial distortion, rigid transform between the two camera frames, by extending to the stereo case the method proposed in [10]. The second one consists of computing and refining the transformation D_{off} (see figure 1) between the camera frames and the endoscope tip.

3.1 Actuated Endoscope Model

The stereoscopic endoscope is set on a actuated arm. Its position D_{er_i} with respect to the fix robot-base is deduced from the articular coordinates. The figure 1 illustrates the frames of the robot-base, the actuated endoscopic arm and the two cameras. Each of the two cameras is described by the classical pinhole model [14]. The intrinsic parameters matrices are noted:

$$A^{k=0,1} = \begin{bmatrix} f_u^k & \gamma^k & u_0^k \\ 0 & f_v^k & v_0^k \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

By introducing the rigid transform D_{01} between the two camera frames, the projection matrices can be written in the first camera frame as:

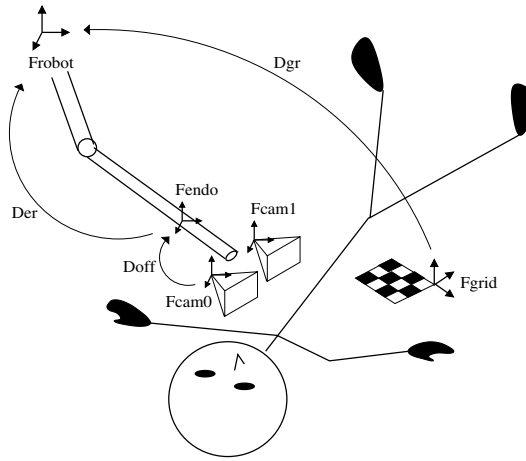


Fig. 1. Definition of the frames.

$$P^0 = A^0\,[I_3|0_3]\ and\ P^1 = A^1\,[R_{01}|T_{01}] \tag{2}$$

We consider the two first terms k_1^k and k_2^k of the radial lens distortion. Finally the endoscope model is completed by the rigid transform D_{off} between the right camera frame and the endoscope tip given by the articular coordinates.

3.2 Protocol

Before the introduction of the endoscope in its trocar, the focus is adjusted and a grid is set in front of the endoscope in an arbitrary position with respect to the robot and the patient. The actuated endoscope is moved around the grid and a few views are acquired. The corners of the grid are extracted and the algorithm is run. The overall process only takes a few minutes.

Fig. 2. Left and right view of the calibration grid.

3.3 Determination of the Parameters of the Two Cameras

This first part of the method consists of the determination of the geometry of the two cameras: intrinsic parameters and radial distortion, rigid transform between the two camera frames.

Intrinsic parameters and pose estimation. The intrinsic parameters are estimated assuming the cameras are free of distortion. As shown in [10], they are derived from the properties of the collineations between the planar grid and the projected points. With at least three views, all the five intrinsics parameters may be estimated in the generic case. For each point of view, the pose D_i of the right camera with respect to the grid is determined from the intrinsic parameters and the properties of the collineations. Then, an estimate D_{01} of the transformation between the two camera frames is derived.

Refinement. The intrinsic parameters, the poses D_i (rotation vector r_i, translation T_i) of the right camera with respect to the grid and D_{01} (rotation vector r_{01}, translation T_{01}) are refined by globally minimising the distance from the projection of the points M_j of the grid to the detected ones $m_{ij}^{0,1}$ in all available stereo pairs. We also introduce the distortion coefficients k_1^k and k_2^k, initialised at 0. This non-linear criterion is minimised using the Levenberg-Marquardt algorithm [15]:

$$\sum_i \sum_j \begin{array}{l} \|m_{ij}^0 - m(f_u^0, f_v^0, \gamma^0, u_0^0, v_0^0, k_1^0, k_2^0, r_i, T_i, M_j)\|^2 \\ + \|m_{ij}^1 - m(f_u^1, f_v^1, \gamma^1, u_0^1, v_0^1, k_1^1, k_2^1, r_i, T_i, r_{01}, T_{01}, M_j)\|^2 \end{array} \quad (3)$$

Such bundle adjustment technique is commonly used in photogrammetry (see [16] for a recent review).

3.4 Determination of the Cameras - Endoscope Tip Transform

This second part of the method can be applied to any AR system with cameras attached to a sensor localised with respect to a fixed reference (optic or magnetic tracker). The absolute position of the calibration grid is simultaneously determined.

Use of the endoscope displacements. As illustrated in figure 1, the absolute position D_{gr} of the calibration grid is unknown. We overcome the difficulty by using the displacement matrices between two positions: the displacement of the endoscope with respect to the robot and the displacement of the cameras with respect to the grid. Between the positions i and $i+1$ of the endoscope, we can write:

$$D_{m_{i\to i+1}} = D_{i+1} * D_i^{-1} = D_{off}^{-1} D_{er_{i+1}}^{-1} D_{er_i} D_{off} \quad (4)$$

id est:

$$D_{er_{i+1}} D_{off} D_{m_{i\to i+1}} = D_{er_i} D_{off} \quad (5)$$

The matrices can be parameterised using quaternions q [17] and vectors of translation, leading to the following system:

$$q_{er_{i+1}} * q_{off} * q_{m_{i\to i+1}} = q_{er_i} * q_{off} \quad (6)$$

$$\left(R_{er_{i+1}} - R_{er_i}\right) T_{off} = T_{er_i} - T_{er_{i+1}} - R_{er_{i+1}} R_{off} T_{m_{i\to i+1}} \quad (7)$$

By expanding the equation (6) for n displacements of the endoscope, we have for the rotation:

$$Aq_{off} = 0 \tag{8}$$

This last equation is solved using a singular value decomposition of the $4n \times 4$ matrix A (see [18] fro a recent review and analysis) . Then, the expansion of the equation (7) , gives for the translation:

$$BT_{off} = C \tag{9}$$

where B is a $3n \times 3$ matrix and C a vector of dimension $3n$. The problem is solved from:

$$T_{off} = \left(B^t B\right)^{-1} B^t C \tag{10}$$

The grid-robot transformation is then deduced from any endoscope position:

$$D_{gr} = D_{er_i} * D_{off} * D_i \tag{11}$$

Refinement. The position of the grid D_{gr} and the cameras - endoscope tip transform D_{off} are refined by minimising, using the Levenberg-Marquardt algorithm [15], the following criterion. The residual error is a quantitative indication of the performance of the superposition of the points in the images and the virtual view generated using the endoscope model and the state of the actuated endoscope.

$$\begin{array}{r} \sum_i \sum_j \quad \|m^0_{ij} - m^0_i(r_{off}, T_{off}, r_{gr}, T_{gr}, M_j)\|^2 \\ + \|m^1_{ij} - m^1_i(r_{off}, T_{off}, r_{gr}, T_{gr}, M_j)\|^2 \end{array} \tag{12}$$

4 Experiments and Results

The da Vinci™ surgical system was used for experiments with the cardiac surgery team at Hôpital Européen Georges Pompidou, Paris, France. An acquisition system has been developed in the team to synchronously acquire the two video channels of the endoscope and the state of the robot with dedicated APIs written by Intuitive Surgical. The PC-based system serves this data through a network from the operating room of the hospital. A very high speed network allows us to collect the data and work from our lab in Sophia-Antipolis. A surgical robotics simulator previously developed [6] (see also our companion paper) is connected through the network and computes in real time the two virtual views corresponding to the endoscope position by using the calibration results.

Endoscope Calibration. We acquired several data sets and verified the stability of the result. For each data set of n endoscopic positions, we applied the algorithm on all the $n - 1$ endoscopic positions subsets.

After optimisation of the parameters of the two cameras, the reprojection error was less than half a pixel (eq. 3). We also computed the RMS error of the

distance from the 2D points to the epipolar lines which measures the quality of the calibration. The residual error was around one pixel.

The transformation between the cameras and the tip of the endoscope was estimated as explained earlier. The residual error computed using the calibration results and the forward kinematics of the robot (eq. 12) was of the order of 4 pixels on the data subsets. We obtained equivalent reprojection error on the position of the endoscope non used in the calibration procedure. We hope to improve this error by using a calibration procedure that would correct the measured position of the endoscopic arm.

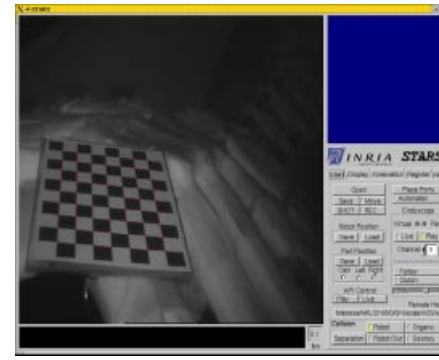
3D reconstruction of the operating field. As described in [5] we investigate the use of surfacic information to refine the overall registration problem. The experimental setup is composed of a plastic phantom composed of a heart and rib cage. After calibration of the actuated endoscope, the cameras parameters are used to reconstruct in 3D by stereoscopy the surface of the plastic heart. We use the cameras to endoscope transformation and the forward kinematics of the robot to fuse the surface patches corresponding to different positions of the endoscope for the fixed robot frame. Figure 3(c) depicts preliminary results of 3D reconstruction of the heart surface. These results are displayed with segmented surface of the heart from preoperative CT scan. Animal experiments are currently carried out to better validate the proposed approach and assess its adaptability.

Overlay in the endoscope. The the rib cage of the phantom is registered to preoperative CT data using a fiducial markers based pointing mechanism. On the other hand, a 3D model of the coronary arteries had been obtained from angiographic projections [1]. After calibration, the endoscope is inserted in its trocar and placed in operating position. We use the calibration results to model the cameras, find the cameras to robot transformation and perform the first augmentation of the endoscopic images with the preoperative model of the arteries. Figure 3(d) illustrates this overlay. We have observed a misalignment which can be attributed to the external registration error, the limited precision of the robot kinematics and the precision of the endoscope calibration. The different sources of error are being further investigated through a more discriminating experimental setup that we are currently building.

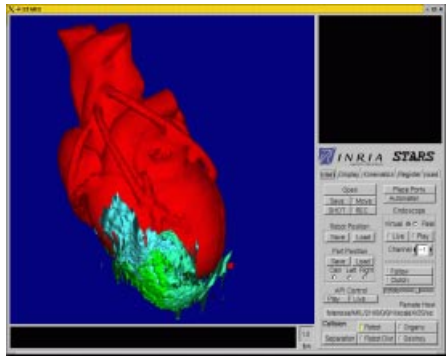
5 Discussion and Future Trends

This work addresses the problem of augmenting endoscopic 3D vision in robot assisted surgery. It focuses on the use of a priori information to initialise the overlay by a precise calibration of the actuated stereoscopic endoscope. The proposed method avoids any additional tracking system in the operating room. Its flexibility is compatible with the constraints of the operating theatre and it can be applied to a variety AR systems.

(a) experimental setup



(b) calibration pattern



(c) surface patches of the plastic heart reconstructed from the endoscopic images and the preoperative heart model segmented in CT-scan.

(d) superposition of the preoperative model of the coronary arteries in the endoscopic view of the plastic heart.

Fig. 3. Experimental results

3D reconstructions of the operating field performed in an experimental setup composed of a plastic heart inside a rib cage illustrate the quality of the results. The experiments were performed with Professor Alain Carpentier's cardiac surgery team and the Intuitive Surgical team at the Hôpital Européen Georges Pompidou, Paris, France. The error measured on a planar calibration pattern observed from an unknown point of view in an operative configuration is around 4 pixels. We investigate the reduction of this error by more precisely calibrating the mechanic of the actuated endoscope. This improvement should have consequences on the use of the direct robot kinematics to initialise the overlay and we are building an experimental setup to quantify the error of the registration string from the preoperative acquisitions to the intraoperative overlay in the endoscope.

Meanwhile we address, the second aspect of the registration problem: the use of intraoperative information. We investigate, on in vivo images of animals, the way to refine the displacement of the endoscope deduced from the forward robot kinematics, to deal with the dynamic aspect and to constrain the registration by using eligible landmarks for our cardiac application.

Acknowledgements

Thierry Vieville and Olivier Faugeras are gratefully acknowledged for some fruitful ideas during the development of this work. We also thank Cyril Coirier from Intuitive Surgical France and Prof. Alain Carpentier and the cardiac surgery team at the Hôpital Européen Georges Pompidou, Paris.

References

1. Mourgues, F., Devernay, F., Malandain, G., Coste-Manière, E.: 3d+t modeling of coronary artery tree from standard non simultaneous angiograms. In: Proceedings MICCAI. Volume 2208 of Lecture Notes in Computer Science., Springer (2001) 1320–1322
2. Fuchs, H., Livingston, M., Raskar, R., Colucci, D., Keller, K., State, A., Crawford, J., Rademacher, P., Drake, S., A.Meyer: Augmented reality visualization for laparoscopic surgery. In: Proc. of MICCAI'98. Volume 1496 of LNCS., Springer (1998) 934–943
3. Dey, D., Slomka, P., Gobbi, D., Peters, T.: Mixed reality merging of endoscopic images and 3-D surfaces. In: Proceedings MICCAI. Volume 1935 of Lecture Notes in Computer Science., Springer (2000) 796–803
4. Buck, S.D., Cleynenbreugel, J.V., Geys, I., Koninckx, T., Koninck, P.R., Suetens, P.: A system to support laparoscopic surgery by augmented reality visualization. In: Proceedings MICCAI. Volume 2208 of Lecture Notes in Computer Science., Springer (2001) 691–698
5. Devernay, F., Mourgues, F., Coste-Manière, E.: Towards endoscopic augmented reality for robotically assisted minimally invasive cardiac surgery. In IEEE, ed.: Proceedings of Medical Imaging and Augmented Reality. (2001) 16–20
6. Coste-Manière, È., Adhami, L., Severac-Bastide, R., Lobontiu, Adrian Salisbury, J.K.J., Boissonnat, J.D., Swarup, N., Guthart, G., Mousseaux, É., Carpentier, A.: Optimized port placement for the totally endoscopic coronary artery bypass grafting using the da Vinci robotic system. In Russ, D., Singh, S., eds.: Lecture Notes in Control and Information Sciences, Experimental Robotics VII. Volume 271., Springer (2001)
7. Mourgues, F., Devernay, F., Coste-Manière, E.: 3D reconstruction of the operating field for image overlay in 3D-endoscopic surgery. In: Proceedings of International Symposium on Augmented Reality. (2001)
8. Tsai, R.Y.: 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 **3** (1987) 323–344
9. Maybank, S.J., Faugeras, O.D.: A theory of self-calibration of a moving camera. The International Journal of Computer Vision **8** (1992) 123–152

10. Zhang, Z.: Flexible camera calibration by viewing a plane from unknown orientations. In: Proceedings of the 7th International Conference on Computer Vision, Kerkyra, Greece, IEEE Computer Society, IEEE Computer Society Press (1999) 666–673
11. Khadem, R., Bax, M.R., Johnson, J.A., Wilkinson, E.P., Shahidi, R.: Endoscope calibration and accuracy testing for 3d/2d image registration. In: Proceedings MICCAI. Volume 2208 of Lecture Notes in Computer Science., Springer (2001) 1361–1362
12. Sauer, F., Khamene, A., Bascle, B., Rubino, G.: A head-mounted display system for augmented reality image guidance: Towards clinical evaluation for iMRI-guided neurosurgery. In: Proceedings MICCAI. Volume 2208 of Lecture Notes in Computer Science., Springer (2001) 707–716
13. Genc, Y., Sauer, F., Wenzel, F., Tuceryan, M., Navab, N.: Optical see-through HMD calibration: A stereo method validated with a video see-through system. In: Proceedings of International Symposium on Augmented Reality. (2000)
14. Faugeras, O.: Three-Dimensional Computer Vision. MIT Press (1993)
15. More, J.: The levenberg-marquardt algorithm, implementation and theory. In Watson, G.A., ed.: Numerical Analysis. Lecture Notes in Mathematics 630. Springer-Verlag (1977)
16. Triggs, B.: Optimal estimation of matching constraints. In Koch, R., Gool, L.V., eds.: Workshop on 3D Structure from Multiple Images of Large-scale Environments SMILE'98. Lecture Notes in Computer Science (1998)
17. Pervin, E., Webb, J.: Quaternions in computer vision and robotics. In: Proceedings of the International Conference on Computer Vision and Pattern Recognition, Washington, IEEE (1983) 382–383
18. Papadopoulo, T., Lourakis, M.: Estimating the jacobian of the singular value decomposition: Theory and applications. Research Report 3961, INRIA Sophia-Antipolis (2000)