

Automatic Patient Registration for Port Placement in Minimally Invasive Endoscopic Surgery

Marco Feuerstein^{1,2}, Stephen M. Wildhirt², Robert Bauernschmitt², and Nassir Navab¹

¹ Computer Aided Medical Procedures (CAMP) Group, TU Munich, Germany

² Department of Cardiothoracic Surgery, German Heart Center Munich, Germany

Abstract. Optimal port placement is a delicate issue in minimally invasive endoscopic surgery, particularly in robotically assisted surgery. A good choice of the instruments' and endoscope's ports can avoid time-consuming consecutive new port placement. We present a novel method to intuitively and precisely plan the port placement. The patient is registered to its pre-operative CT by just moving the endoscope around fiducials, which are attached to the patient's thorax and are visible in its CT. Their 3D positions are automatically reconstructed. Without prior time-consuming segmentation, the pre-operative CT volume is directly rendered with respect to the endoscope or instruments. This enables the simulation of a camera flight through the patient's interior along the instruments' axes to easily validate possible ports.

1 Introduction

Ideal port placement is one of the key issues in minimally invasive endoscopic surgery, particularly in robotically assisted surgery. The optimal choice of the instruments' ports provides full access to the whole operation region as well as adequate surgeon dexterity. This can avoid time-consuming new port placement, which is a strain on every patient.

In the current clinical work flow, the surgical staff selects all ports by palpation of external anatomic landmarks, primarily based on their previous experience. However, if these external landmarks do not correspond to the individual internal anatomy of each patient, a misplacement of ports can occur. Several methods have been proposed to improve and automate the optimal placement of ports [1–4]. They all have two major disadvantages: 1) They rely on the time-consuming manual or semi-automatic segmentation of pre-operative imaging data from CT or MRI, which is essential to reconstruct models of any involved anatomy, e.g. ribs, heart, and soft tissue. These 3D models are used to automatically compute the port locations. 2) They lack a practical and accurate way to transfer the planned port locations to the operating room, which is achieved by registering the patient to the pre-operative data.

In any case, the patient registration process is based on matching anatomical or artificial landmarks, which are visible on both the patient and its CT data.

Adhami and Coste-Maniere use the end effectors of the da Vinci telemanipulator to point to fiducials, which are attached to the patient [1]. Due to their shape and intensity, the fiducials can be segmented automatically in the CT data. Intra-operatively, they move the robot arm’s end effector to every single fiducial in order to get its position in the robot coordinate frame. This is a time-consuming and unnatural task. Similarly, Selha et al use the sensor of an additional electromagnetic tracking system [3] as a pointing device. However, they base their registration on anatomical landmarks. Both electromagnetic tracking and the use of anatomical landmarks introduce an inherent imprecision when determining corresponding landmarks.

We propose a fast, practical, and easy method to register the CT data to the patient. Spherical CT visible self-adhesive fiducials are stuck on the patient’s skin. They are segmented automatically in its CT data. Intra-operatively, instead of pointing to the fiducials, we only move the tracked endoscope around the fiducials and acquire a set of images from differing, but arbitrary poses. To simplify the acquisition process, not all fiducials need to be seen by the camera in a single image. By automatically detecting the fiducials in these images, we reconstruct their 3D positions in the tracking (=world) coordinate frame. Point based registration methods enable us to match them with the CT data. For port placement, the surgical staff simply moves the tracked instruments or endoscope to the positions where it wishes to place their corresponding ports. A virtual camera is placed on top of the instruments’ end effectors or the endoscope’s camera center. It is able to simulate a flight through the patient’s interior by rendering the CT volume as it would be seen by the endoscope. In this natural way, optimal port placements can easily be identified without prior segmentation of patient’s anatomy or any tedious pointing device. Our method is applicable to any tracked endoscope, no matter whether it is tracked by an optical tracking system, a mechanical one such as da Vinci, or any other tracking system.

In order to reconstruct the 3D positions of the fiducials, the endoscope’s pose and intrinsic parameters need to be determined. This is achieved by a one-time hand-eye calibration, as described in section 2. In section 3, we present our algorithms for 3D reconstruction and patient registration. Further details on the provision of volume rendering for port placement can be found in section 4. Our conducted experiments on a thorax phantom are described in section 5. We conclude in section 6 with an evaluation of our presented methods and a short outlook on future research.

2 Calibration of the Endoscope

For our application, the endoscope camera is rigidly attached to a sensor, e.g. to a marker target seen by an optical tracking system or an actuated robot arm as for da Vinci. The main purpose of calibrating the endoscope is to model the transformation of a 3D world point onto the camera’s 2D image plane, so the projection of a fiducial onto the endoscope’s image can be reproduced mathematically.

In detail, a point X_w in the world frame is first transformed into the sensor frame by wT_s , from where it is transformed into the camera frame by sT_c , and finally mapped onto the image plane by the camera's calibration matrix K . The transformation wT_s can be directly received from the tracking system. The rigid transformation sT_c from sensor to camera coordinate frame and the intrinsic camera parameters stored in K need to be computed once. Additionally, the rather large radial and tangential lens distortion of endoscopes needs to be corrected for.

To compute all unknowns, a classical hand-eye calibration approach is taken [5–7]. Therefore, a flat checkerboard pattern is placed arbitrarily. The tracked camera performs a series of n motions. At the pause of each motion, the camera acquires an image of the pattern and the pose of the attached sensor is recorded. Having at least two motions (rotations around distinguished axes) or three poses, respectively, the offset sT_c along with the camera's intrinsic parameters and distortion coefficients can be computed as follows:

First, the intrinsics, distortion coefficients, and camera poses in the pattern coordinate frame ($pT_c^1 \dots pT_c^m$: the transformations from pattern to camera coordinate frame) are computed using the gold standard algorithms for camera calibration [8, 9].

Second, the rigid offset between sensor and camera is computed. All transformations involved during a single motion from pose i to pose j can be seen on figure 1(a). The camera motions can be easily computed from previous results. Analogous, the sensor motions can be received from the recorded poses. To compute sT_c , the following so-called hand-eye equation needs to be solved:

$$\forall i = 1 \dots n, j = 1 \dots n, i \neq j : sT_c {}^i T_s^j = {}^i T_c^j sT_c \quad (1)$$

This can be achieved by decomposing the involved matrices, as described by Tsai/Lenz and others [5–7].

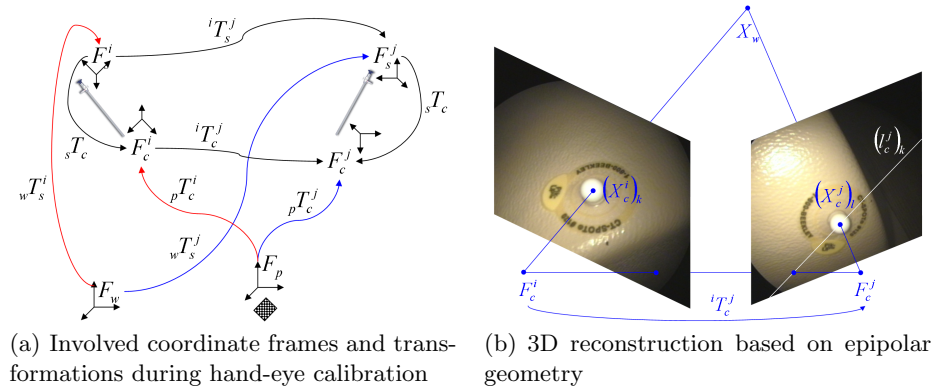


Fig. 1. The principles of hand-eye calibration and epipolar geometry

3 Automatic 3D Reconstruction for Patient Registration

For patient registration, three essential steps are required: 1) All fiducials must be segmented in the CT volume to determine the positions of their centroids. 2) Their positions in the tracking coordinate frame need to be reconstructed using the images, which are acquired by the calibrated endoscope camera and show the fiducials. 3) The resulting point sets need to be matched in order to register the patient to its CT data set.

The automatic segmentation of the fiducials in the CT volume can be achieved by using standard image processing techniques based on thresholding, filling, morphology, and subtraction [10, 11]. The centroids of all segmented fiducials can be computed very precisely by weighing their associated voxel intensities and incorporating partial volume effects.

For finding the 3D positions of the fiducials in the tracking coordinate frame, two iterations are performed for each image i containing an arbitrary number m of fiducials. First, the 2D positions $x_1^i \dots x_m^i$ of all visible fiducials are extracted automatically after undistortion of the image. Similar techniques as for the segmentation of the CT data are used, which also incorporate edge detection and color information of the fiducials and patient's skin [11]. Second, the properties of epipolar geometry are applied to reconstruct their 3D positions [9, 12], as illustrated on figure 1(b).

Next, all 2D point pairs corresponding to the same 3D point are used to optimally reconstruct the 3D point. For all 2D points, their associated projection rays $r_1 \dots r_s$ are constructed, which intersect the camera center $C_r = {}_c t_w^i$ and the point's projection onto the image plane $P_r = {}_c R_w^i (X_c^i)_k + {}_c t_w^i$, where ${}_c R_w^i = ({}_w R_c^i)^T$ and ${}_c t_w^i = -({}_w R_c^i)^T {}_w t_c^i$. They can be represented using the camera center C_r as starting point and a directional unit vector d_r :

$$r_r = C_r + \lambda_r d_r = C_r + \lambda_r \frac{P_r - C_r}{\|P_r - C_r\|} \quad (2)$$

The associated midpoint X_w can be computed, which is closest in average to all s rays. Therefore, following overdetermined system of linear equations has to be minimized:

$$\sum_{r=1}^s \|C_r + \lambda_r d_r - X_w\|^2 \quad (3)$$

As stated by Sturm et al, this linear least squares problem may be solved using the Pseudo-inverse [13]. Finally, these results can be further improved by using the Levenberg-Marquardt iteration to minimize following equation:

$$\sum_{r=1}^s \left\| K \begin{bmatrix} {}_s R_c | {}_s t_c \end{bmatrix} \begin{bmatrix} ({}_w R_s)_r & ({}_w t_s)_r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ 1 \end{bmatrix} - \begin{bmatrix} x_r \\ 1 \end{bmatrix} \right\|^2 \quad (4)$$

After the reconstruction of all 3D points from their associated 2D points, they need to be matched with the points segmented in the CT data set. Therefore, the correct point correspondences need to be identified and the transformation from

the CT coordinate frame into the world coordinate frame, where the patient is registered in, needs to be computed. This can be done by a distance-weighted graph matching approach along with a point based registration algorithm [14, 15]. Finally, the patient’s CT volume is registered in the same coordinate frame as the patient.

4 Volume Rendered Port Placement

Once the CT volume is registered to the patient, it can be visualized with respect to any tracked instruments, overlaid onto the real images of the endoscope, displayed simultaneously with the real endoscopic images [16], or even be used to extend the endoscopic images to improve the surgeon’s orientation. In our case, the CT volume is directly rendered as it would be seen by virtual cameras put in front of the tracked instruments or by the real endoscope camera. This can be used for placing the instruments and endoscope or their corresponding ports, respectively, in an optimal way. By virtually moving the camera in and out the volume along the instruments’ or endoscope’s main axes the surgeon can intuitively verify, whether the current poses of the instruments and the endoscope are ideal to reach the whole operating region.

For rendering the volume, pre-defined transfer functions are offered to assign specific colors and opacities to certain image intensities, which can be modified interactively to the surgeon’s needs. This makes it easy to realistically visualize only the anatomy, which is essential for the success of the intervention, e.g. for cardiac surgery the bones, aorta, heart, and other main arteries such as the left and right arteria mammaria interna, which were contrasted for CT. To provide a fast and though detailed visualization during port placement in real time, our volume renderer is using the graphic card’s GPU (graphical processing unit) to perform the computations for 3D texture mapping.

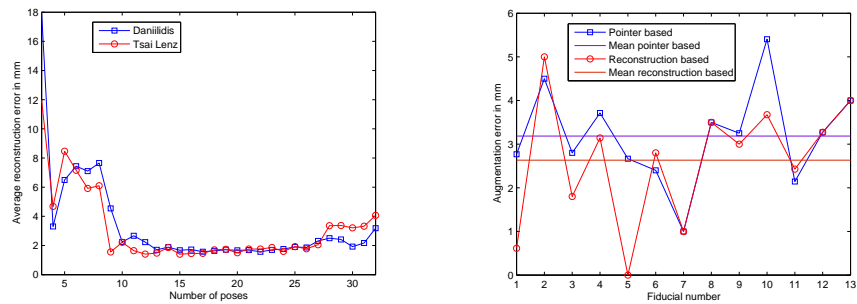
Utilizing this approach of direct volume rendering without prior segmentation and generation of polygonal 3D models of the patient’s anatomy saves a noticeable amount of time for planning the ports and leaves the control to the surgical staff during port placement.

5 Experimental Results

For our experiments we used a 30 degrees laparoscope tracked by an optical tracking system, which has a root mean square error of 0.53 millimeters for the viewing axis and 0.32 millimeters for the other axes when tracking retroreflective markers. We implemented two classical hand-eye calibration methods by Tsai/Lenz and Daniilidis [5, 6]. Tsai and Lenz combine two QR decompositions to determine the translation and rotation, whereas Daniilidis uses dual quaternions and a single singular value decomposition.

The intrinsic and extrinsic camera parameters were estimated from 32 frames. For hand-eye calibration, 3 to 32 endoscope poses and all possible motions between them were used to estimate the transformation from sensor to camera,

resulting in 30 transformation matrices. To validate these matrices, the positions of 9 retroreflective spherical markers were reconstructed from 6 endoscopic images. These reconstructions were compared to the measurements of the optical tracking system. The average distance of the reconstructed points to the measurements of the tracking system was computed for each transformation matrix. As visualized in figure 2(a), a typical hand-eye calibration incorporating 10 to 25 poses gave errors between 1.4 and 2 millimeters. The described hand-eye calibration is done off-line, so above results remain valid for a long period of time and only need to be verified every now and then.



(a) Average reconstruction error for retroreflective markers depending on the hand-eye calibration method and number of poses (b) The augmentation error for all 13 fiducials measured directly in the video

Fig. 2. Experimental reconstruction and augmentation error

To determine the augmentation error during port placement, 13 CT visible spherical fiducials with a diameter of 4 mm were attached to a plastic thorax phantom containing a heart model. After a CT scan and segmentation of all fiducials, the phantom was placed arbitrarily. The 3D positions of 4 fiducials were reconstructed automatically by moving the tracked endoscope around them, using 3 to 4 images from differing poses for each fiducial. The other 9 fiducials were just used later for validating the augmentation from many different viewing directions, so in practice they are not needed. Next, the CT-to-tracking transformation of the 4 fiducials was computed.

Having the sensor-to-camera and CT-to-tracking transformations as well as intrinsic camera parameters and distortion coefficients, the endoscopic images can be undistorted and the CT volume can be augmented on them. To verify the augmentation, the distances of all 13 fiducials from the real images to a semi-transparent augmentation in an orthogonal view were measured. An average error of 2.6 mm could be determined. This is fully sufficient for a precise port placement. We also compared our automatic 3D reconstruction method to a pointer based approach. Therefore, a pointing device tracked by the optical tracking system was used to record the positions of the 4 fiducials. Again, the

CT-to-tracking transformation was computed and used for the augmentation. With this method, we only achieved an average error of 3.2 mm, i.e. our method almost systematically performs better than the pointer based one. The comparison is visualized in figure 2(b).

A port placement application was implemented offering three visualization modes, as displayed on figure 3. In the first mode, the undistorted real endoscopic image is displayed. The second mode additionally augments the volume on the phantom in a half-transparent mode, so the accuracy of the overlay can be verified by the surgeon. In a third purely virtual mode the surgeon can switch the endoscope optics from 30 degrees to 0 degrees and move the camera in and out the volume along the endoscope’s main axis to validate a possible port. The augmentation of a $512 \times 512 \times 444$ CT volume and undistortion of the camera frames with a resolution of 800×600 pixels was achieved in real time (15 fps).

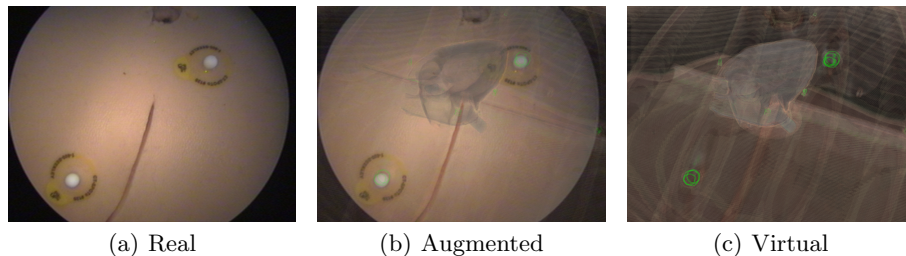


Fig. 3. 3 visualization modes for the same endoscope pose: 3(a): Real camera image, 3(b): Transparent augmented view outlining fiducials, ribs, and heart (the virtual green contours correctly match the white fiducials in the video image), 3(c): Purely virtual view, which can be used for port placement to move the camera in and out

6 Conclusion

In this work we addressed and solved two substantial problems of current approaches dealing with the improvement and automation of port placement: Time-consuming segmentation of patient’s anatomy and inadequate patient registration. Moreover, our technique not only supports the surgeon during port placement, it also enhances the endoscopic images by undistortion. Besides the tracking system used to determine the endoscope’s pose no further tracking system is needed. We reckon that our technique can supplement the current clinical work flow easily, because we keep it simple and still leave the control to the surgeon during port placement. Apart from the pre-operative attachment of 4 fiducials to the patient and a short and intuitive intra-operative patient registration procedure we do not alter the conventional clinical work flow. Our method can be applied to any minimally invasive endoscopic procedure provided that pre-operative patient data is available.

This method is more precise than the usual method of pointing the robot or hand-held endoscope/instrument to each fiducial. It also fits more smoothly into the surgical work flow, as it only requires the surgical staff to move the endoscope camera over the patient's body. Our work does not address organ deformations and motions caused by the insufflation of carbon dioxide and respiratory as well as cardiovascular effects. In this sense, the system only provides an approximative result and relies on the surgeon's expertise for further considerations of possible deformations.

Acknowledgements. This work was supported in part by the German Research Foundation (DFG) within the Collaborative Research Centre SFB 453 on "High-Fidelity Telepresence and Teleaction". We like to thank A.R.T. GmbH for providing the tracking system.

References

1. Adhami, L., Coste-Manière, È.: A versatile system for computer integrated mini-invasive robotic surgery. In: MICCAI '02, Japan (2002) 272–281
2. Chiu, A.M., Dey, D., Drangova, M., Boyd, W.D., Peters, T.M.: 3-d image guidance for minimally invasive robotic coronary artery bypass. *HSF* **3** (2000) 224–231
3. Selha, S.D., Dupont, P.E., Howe, R.D., Torchiana, D.F.: Dexterity optimization by port placement in robot-assisted minimally invasive surgery. In: Proceedings of SPIE TTT VIII. Volume 4570. (2002) 97–104
4. Traub, J., Feuerstein, M., Bauer, M., Schirmbeck, E.U., Najafi, H., Bauernschmitt, R., Klinker, G.: Augmented reality for port placement and navigation in robotically assisted minimally invasive cardiovascular surgery. In: CARS 2004, Chicago, USA (2004) 735–740
5. Tsai, R., Lenz, R.: Real time versatile robotics hand/eye calibration using 3d machine vision. In: CRA88. (1988) 554–561
6. Daniilidis, K.: Hand-eye calibration using dual quaternions. *The International Journal of Robotics Research* **18** (1999) 286–298
7. Mourgues, F., Coste-Manière, È.: Flexible calibration of actuated stereoscopic endoscope for overlay in robot assisted surgery. In: MICCAI, Japan (2002) 25–34
8. Bouguet, J.Y.: Camera calibration toolbox for Matlab. <http://www.vision.caltech.edu/bouguetj/calib.doc/> (2004)
9. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Second edn. Cambridge University Press, ISBN: 0521540518 (2004)
10. Wang, M.Y., Maurer, C. R., J., Fitzpatrick, J.M., Maciunas, R.J.: An automatic technique for finding and localizing externally attached markers in ct and mr volume images of the head. *IEEE Trans Bio Med Eng* **43** (1996) 627–637
11. Nicolau, S., Garcia, A., Pennec, X., Soler, L., Ayache, N.: Augmented reality guided radio-frequency tumor ablation. *CAVW* **16** (2005)
12. Trucco, E., Verri, A.: Introductory Techniques for 3-D Computer Vision. Prentice Hall PTR (1998)
13. Sturm, P., Ramalingam, S., Lodha, S.K.: On calibration, structure-from-motion and multi-view geometry for panoramic camera models. In: Panoramic Photogrammetry Workshop, Berlin, Germany (2005)
14. Gold, S., Rangarajan, A.: A graduated assignment algorithm for graph matching. *IEEE Trans. Pattern Anal. Mach. Intell.* **18** (1996) 377–388
15. Walker, M.W., Shao, L., Volz, R.A.: Estimating 3-d location parameters using dual number quaternions. *CVGIP: Image Underst.* **54** (1991) 358–367
16. Shahidi, R., et al: Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *IEEE Trans. Med. Imaging* **21** (2002) 1524–1535