International Congress Series 1268 (2004) 527–532

www.ics-elsevier.com

Augmented reality provision in robotically assisted minimally invasive surgery

D.A. Wang*, F. Bello, A. Darzi

Department of Surgical Oncology and Technology, Imperial College London, 10th Floor Q.E.Q.M. Wing, St. Mary's Hospital Praed Street, London W2 1NY, UK

Abstract. Since the widespread introduction of minimally invasive surgery, extensive evidence has demonstrated its advantages over open surgery. Although there are clear benefits, such surgery creates an unnatural environment where the surgeon has lost orientation, the eye–hand–target axis and visual depth perception. Computer-enhanced surgical technology has been proposed to overcome some of these drawbacks. Master–slave tele-manipulator systems enhance manipulation capabilities and increase performance precision, but the field of view presented to the surgeon is still restricted. The aim of our research is to improve visualization during robotically assisted minimally invasive surgery by providing augmented reality facilities that will allow the surgeon to view on demand information relevant to the operation. This paper presents a technique for the calibration of the endoscopic cameras and the slave manipulators of the da Vinci™ surgical system. Calibrating the slave manipulators will enable us to address the problem of updating the virtual objects overlaid on the video frames by tracking tool–tissue interactions and the resulting dynamic changes within the scene. Initial work regarding the registration of computer-generated graphics with images obtained from the da Vinci™ surgical system is also described.

Keywords: Augmented reality; Robotic surgery; Enhanced visualization

1. Introduction

The design and development of robotic minimally invasive surgical systems was motivated by the need to overcome the limitations of manual laparoscopic surgery. Although the problems of reduced dexterity of the surgeon and unnatural hand–eye coordination during laparoscopic surgery have been tackled by robotic surgical systems, the field of view transmitted to the surgeon remains restricted. The desire for improved visualization during robotic minimally invasive surgical procedures has motivated research into enhancing the visual field using computer-generated graphics. This combination of simulated objects with the real time view of the scene is known as augmented reality (AR). The provision of AR facilities will allow the surgeon to view information, which might otherwise be unavailable, overlaid onto the view of the operating scene in real time. Our aim is to produce AR facilities

* Corresponding author. Tel.: +44-20-7886-2130; fax: +44-20-7886-1810.
E-mail address: dorothy.a.wang@imperial.ac.uk (D.A. Wang).

doi:10.1016/j.ics.2004.03.057

specifically for the da Vinci™ surgical system, the leading tele-manipulator system manufactured by Intuitive Surgical® [1].

AR systems such as [2–4] provide supplementary information to the surgeon by allowing pre-operative patient data to be reviewed while the operating field is being observed. With respect to the da Vinci™ system, it is anticipated that simulated images will be combined with the stereo video and presented to the surgeon at the console.

A key step towards the provision of AR facilities for the da Vinci™ system is accurate calibration of the stereo endoscope. Recent work in this area has been based on the calibration algorithms proposed by Tsai [5] and Zhang [6]. The method used by Hattori et al. [7] employed a POLARIS location sensor to track markers on the endoscope and the calibration pattern. As in Ref. [8], we made use of the real time kinematic data provided by the da Vinci™'s Application Programming Interface (API) software to avoid the need for a separate tracking device to measure the pose of the endoscope and to incorporate the positional data of the slave manipulators in an additional calibration step designed to facilitate the tracking and updating of the virtual objects.

2. da Vinci™ robotic surgical system

The da Vinci™ surgical system consists of a master console and a patient-side cart. There are three robotic arms on the patient-side cart. One arm holds a stereo endoscopic camera while the other two hold surgical instruments corresponding to the surgeon's left and right hand. The combination of a robotic arm and its attached tool is known as a slave manipulator. Seated at the master console, the surgeon uses the controls to manipulate the surgical instruments. The endoscopic view of the operating scene is transmitted to a binocular display set within the master console. Creating AR facilities for the da Vinci™ system will provide the surgeon with the ability to review the patient's pre-operative data without turning away from the operating scene.

3. Calibration

The calibration procedure implemented in this project is divided into two parts: camera calibration and slave calibration. When viewing a real world scene via a camera, 3D objects are projected onto a 2D display. In order to augment a 2D video scene with computer-generated graphics, the transformation between various coordinate systems must be known (see Fig. 1). This is to ensure that the graphics can be correctly positioned and orientated within the scene.

Camera calibration is an important procedure that extracts camera and 3D metric scene information from 2D images. The intrinsic parameters (the transformation between image and camera coordinate frames) and the extrinsic parameters (the pose with respect to the world coordinate frame) of a camera are determined from calibration. The most common techniques are photogrammetric calibration and self-calibration. In photogrammetric calibration, images of a calibration object consisting of two or three orthogonal planes and whose geometry is accurately known are captured. Self-calibration does not require a calibration object. The camera is moved arbitrarily around a static scene and the intrinsic and extrinsic parameters are calculated through point correspondences between images.

In addition, a slave calibration technique was deemed advantageous. This would enable us to address the issue of updating the virtual objects by tracking tool–tissue interactions

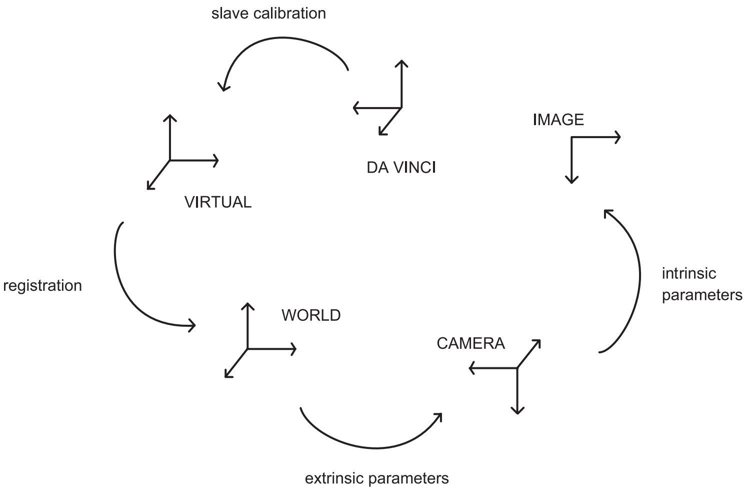

Fig. 1. The transformations between the various coordinate systems present within the environment must be determined.

and the dynamic changes within the scene. For each of the three robotic arms, we can access articulate joint data through da Vinci™'s API and utilize the real time kinematic data. The data is given with respect to the da Vinci™'s own coordinate system. Therefore in order to relate scene objects (measured with respect to a world coordinate system) with the da Vinci™ instruments, the transformation between world and da Vinci™ coordinate systems must be calculated. This is the aim of our slave calibration technique.

3.1. Camera calibration

An offline photogrammetric-based calibration technique was chosen for this project. The method is based on a four-step procedure [9] using a simple black and white checkerboard pattern as a calibration target. The effective grid pattern is a 5×5 array of black and white squares of size 9.4×9.4 mm, giving 36 intersection points. Images of the pattern in at least five different orientations were captured from both left and right video channels of the endoscopic camera (see Fig. 2).

The images were loaded into a MATLAB Toolbox implementation of the calibration algorithm by Bouguet [10]. Neither the position of the object nor its motion need to be

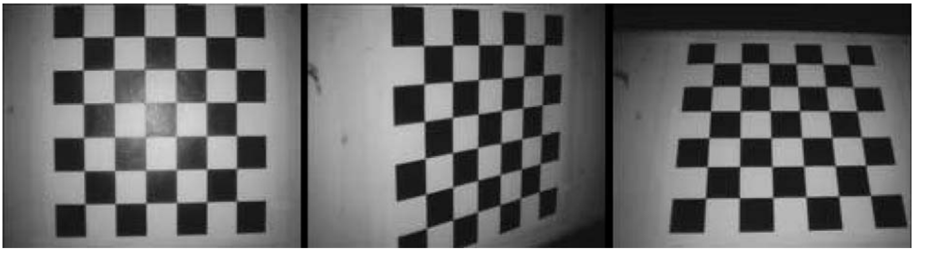

Fig. 2. Images of the calibration grid taken from one of the cameras of the da Vinci™ endoscope.

Fig. 3. Example of images used to select a tool joint for video-based slave calibration.

known. However, one corner of the calibration grid must be defined as the origin of the world coordinate system. Since the extrinsic parameters are relative to the world reference frame, the da Vinci™ endoscopic camera must be calibrated each time AR facilities are required. This technique was chosen due to its simplicity and robustness.

3.2. Slave calibration

For each orientation of the calibration grid, the tip of a closed slave instrument was used to select several (at least four) of the intersection points on the grid. The articulate joint data was gathered for each point and the position of the tip of the instrument relative to the da Vinci™ frame at each point was calculated. As the world location of the probed grid points were known in advance, the transformation matrix between the two coordinate systems could be computed using a well-known algorithm.

A video-based slave calibration procedure was also carried out. Images of the instrument in several different orientations and positions were taken and a pre-specified joint was selected in each of the images. Examples of images used for this calibration procedure are shown in Fig. 3. The results of the camera calibration were used to find the world coordinates of the joint in each position. Then the transformation between these world coordinates and their corresponding da Vinci™ coordinates was computed.

4. Registration

The task of registering two images involves mapping points in one image to their corresponding points in the second image. Our goal is to align the 3D reconstruction produced from pre-operative CT/MRI scans with the intra-operative video stream transmitted by the da Vinci™ endoscopic camera.

Algorithms developed for registration can either be termed rigid or non-rigid, depending on the nature of the transformations considered. Rigid registration performs the mapping using rotations and translations only. In contrast, local deformations are also taken into account during non-rigid registration.

Techniques requiring fiducials are not feasible for applications in minimally invasive surgery. Therefore we aim to produce a non-rigid, intensity-based registration algorithm.

4.1. Method

As a first step towards creating real-time overlays for our application, initial exploratory work has begun using a point-based (rigid) registration technique. A 3D MRI reconstruction of a heart phantom model, manufactured by *Limbs and Things* [11], was superimposed onto its counterpart in a single frame taken from one channel of the da Vinci™ video stream.

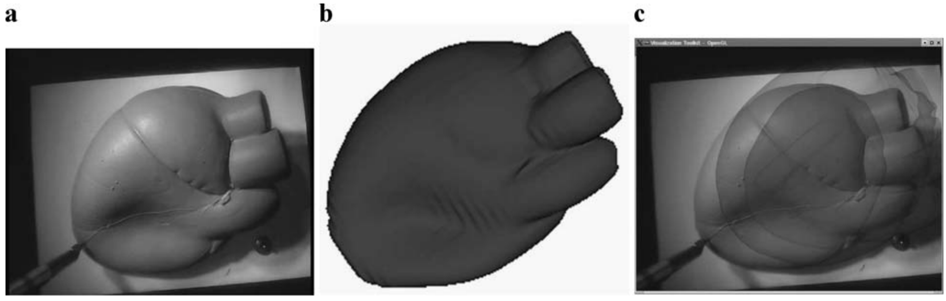

Fig. 4. (a) Image taken from one channel of da Vinci™. (b) 3D reconstruction of the heart phantom. (c) 3D virtual heart superimposed onto its counterpart in the 2D video frame.

Points in the 2D image were selected and their world coordinates were computed using the results of the camera calibration. The corresponding points on the virtual phantom were then selected and the transformation from one point set to the other was computed with software developed by [12]. Combining this transformation with the ones determined from our camera calibration an overlay of the virtual on the real world was generated.

5. Results

Stereo calibration of the da Vinci™ endoscopic camera has provided results that are accurate to within 3 mm. However, our point-selecting slave calibration procedure produced errors as large as 11 mm. With the exception of a few outliers, the video-based slave calibration method generated promising results. Errors using this technique were shown to be lower than those achieved in the point-selecting method.

Initial results of the point-based registration algorithm reveal RMS errors of approximately 14 mm. It is believed that these large errors are caused by the difficulty in accurately selecting corresponding points. An overlay produced using the transformations computed with the registration algorithm can be seen in Fig. 4.

6. Discussion

The aim of our research is to improve visualization during robotically assisted minimally invasive surgery by providing augmented reality facilities. This will allow the surgeon to view on demand information relevant to the operation and make use of such information to define and guide him through the operative volume.

In this paper we have presented a simple offline calibration procedure involving both the stereoscopic camera and the slave manipulators. The results of the calibration will relate the various coordinates systems that are present within the environment. Although large errors were present in the slave calibration results, a new improved version of the system's API is expected to reduce these significantly. Further tests will be carried out with the video-based slave calibration method to determine its suitability over the point-based technique. Through slave calibration, we will be able to determine the image coordinates of known points within the da Vinci™ frame, and vice versa. Data regarding the pose of the endoscopic camera can be obtained from the system's API and therefore, using the results of slave calibration, any changes in camera view may be taken into account and used to update the virtual scene. Also, by being able to locate the position of the tip of the slave instrument, we envisage that it will enable us to track the interactions between instruments and tissue in order to reflect any dynamic changes caused by the instruments on the virtual objects overlaid on the surgical scene.

Our progress in producing an appropriate registration algorithm has also been explained. The current overlays generated would not be beneficial to the surgeon operating. However, we believe that the misalignment of the modalities is caused by an accumulation of errors during the different stages required to create the corresponding point sets to be registered. It is known that a slight deviation in pixel selection causes a large difference in the computed world coordinates. Therefore point selection from the 2D images must be extremely accurate. We are currently working on reducing the errors through the use of a more careful point selection and possibly a more accurate and suitable phantom. It should be noted that our present work in this area is exploratory and not the ultimate goal. We intend to produce a non-rigid, intensity-based algorithm. However, as a small step towards this aim it was considered constructive to begin with a simple point-based registration and to determine the visual accuracy in generating overlays with such an algorithm.

Work towards a fully integrated AR system will continue and we shall explore the application of such system to define a safe operative volume that will constrain the movement of the slave manipulators to avoid them from penetrating high-risk areas.

Acknowledgements

This work is funded by a grant from the BUPA Foundation. We also acknowledge the help and advice received from Intuitive Surgical® in the use of their API software.

References

[1] http://www.intuitivesurgical.com/.
[2] W. Lorensen, R. Kilkinis, Enhancing reality in the operating room, Proc. IEEE Conf. Visual. (1993) 410–415.
[3] F. Devernay, F. Mourgues, È. Coste-Manière, Towards endoscopic augmented reality for robotically assisted minimally invasive cardiac surgery, Proc. MIAR (2001) 16–20.
[4] M. Blackwell, et al., An image overlay system for medical data visualization, Med. Image Anal. 4 (2000) 67–72.
[5] R.Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, IEEE J. Robotics and Automation 3 (4) (1987) 323–344.
[6] Z. Zhang, A flexible new technique for camera calibration, PAMI 22 (11) (2000) 1330–1334.
[7] A. Hattori, et al., A robotic surgery system (da Vinci) with image guided function—system architecture and cholecystectomy application, MMVR (2003).
[8] F. Mourgues, È. Coste-Manière, Flexible calibration of actuated stereoscopic endoscope for overlay in robot assisted surgery, MICCAI (2002).
[9] J. Heikkilä, O. Silvén, A four-step camera calibration procedure with implicit image correction, CVPR (1997) 1106–1112.
[10] J.-Y. Bouguet, http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.
[11] Limbs and Things, http://www.limbsandthings.com/.
[12] D. Rueckert, J. Schnabel, http://www.doc.ic.ac.uk/~dr/.