# Autonomous Ultrasound Scanning for Intraoperative Tumour Localisation and Diagnosis

L. Zhang<sup>1</sup>, M. Ye<sup>1</sup>, C. Chan<sup>1</sup>, G.-Z. Yang<sup>1</sup>

<sup>1</sup>The Hamlyn Centre for Robotic Surgery, Imperial College London, UK lin.zhang11@imperial.ac.uk

## INTRODUCTION

Intraoperative ultrasound (iUS) has become a popular tool in robot-assisted laparoscopic surgery to facilitate in situ pathology localisation. However, manually controlled ultrasound scanning can cause significant cognitive loads to surgeons, due to the need for maintaining optimal scanning orientation and consistent contact with the tissue, as well as covering a wide area for tumour detection and surveillance. In [1], an iUS probe has been controlled by an auxiliary robotic arm to provide guidance to the surgeon for tumour resection. Although an autonomous tumour dissection framework using the iUS has been proposed in [2] on planar tissue surfaces, a 6-DoF autonomous scanning approach that adapts to arbitrary tissue surfaces has yet to be presented. This paper proposes an automatic iUS scanning framework using the da Vinci<sup>®</sup> Research Kit (dVRK). Our framework consists of trajectory planning based on 3D surface reconstruction, an automatic scanning via a 6-DoF visual servoing guidance, tumour segmentation and reconstruction using the ultrasound images.

#### MATERIALS AND METHODS

This paper made use of a da Vinci<sup>®</sup> Patient Side Manipulator (PSM) via the dVRK [3] that provides kinematic control of the robot. A stereo laparoscope system provides SD (720x576) video streaming for both left and right channels at 25 Hz. A UST-533 linear array ultrasound probe (Aloka Medical, Japan) has been mounted in a custom-made clip that can be picked-up by the robot using Cadiere forceps. A marker (KeyDot<sup>®</sup>, Key Surgical Inc, USA) was attached on the probe to assist visual tracking and pose estimation. A customised PVA cryogel kidney phantom was utilised with similar elastic and chromatic properties to human tissue.

The tissue surface is recovered from a disparity map using a stereo matching method [4]. The region for scanning is chosen by the user, and the scanning trajectory is planned adaptively to the 3D surface at the beginning of the task. To calculate the local poses (along the trajectory), which place the probe perpendicular to the surface, a local coordinate frame {S} on each surface point is defined using surface gradients and normals. To transform the trajectory from frame {S} to frame {M} (the marker frame), a matrix  $H_M^S$  is used and defined based on the calibrated transducer's position. These local poses also enforce that the centre of the transducer is aligned with the surface point.

For every step of visual control, a desired marker pose in frame {C} is calculated as:  $H_C^{M*} = H_C^{S*} \cdot H_S^M$ , where  $H_C^{S*}$ 

is current desired transformation from frame {C} to {S}. As shown in Fig. 1, a desired robot command  $H_B^{E*}$  is calculated based on the current and desired marker pose,  $H_C^M$  and  $H_C^{M*}$ , in a camera frame {C} via:  $H_B^{E*} = H_B^M \cdot H_E^M \cdot H_C^M \cdot H_M^E$ . The transformation  $H_E^M$  between end-effector frame {E} and marker frame {M} is a constant which can be either measured or calibrated using a standard hand-eye calibration method. For every image in a video sequence, the marker is detected using a circular-grid detection method followed by a pose estimation step [5] that calculates the marker's pose  $H_C^M$  in frame {C}. In order to improve the marker detection rate, we have included a tracking component based on pyramidal optical flow [6] to track the circular grid along time.

To segment ultrasound images that contain tumours, speckle noises are removed by applying a Butterworth second order filter followed by an active contour [7] segmentation given a manually chosen seed region. The boundary and holes of the segmented tumour region are smoothed and filled by a set of morphological operations. The boundary is represented as a number of points where each point has correspondence across ultrasound images. In order to reconstruct a 3D model of the tumour using the ultrasound images, we need to transform the segmented boundary points from a 2D image coordinate frame to the 3D camera frame  $\{C\}$ . To this end, an ultrasound coordinate frame {U} is defined in which each pixel of the ultrasound image can be represented as a 3D point  $p^{U}$ . For N segmented boundary points  $p^{U}(i)$  are then transformed to frame {C} via :  $p^{C}(i) = H_{C}^{M} \cdot H_{M}^{U}$ .  $p^{U}(i), i \in [1, N]$ , where  $H_{M}^{U}$  can be found via calibration and measurement. A 3D model of the tumour can then be created by connecting all the corresponding boundary points.



**Fig. 1** Coordinate frames defined in this paper: surface frame  $\{S\}$ , marker frame  $\{M\}$ , end-effector frame  $\{E\}$ , robot base frame  $\{B\}$ , ultrasound image frame  $\{U\}$  and camera frame  $\{C\}$ .



**Fig. 2** A 3D visual comparison between the CT ground truth (a) and reconstructed models of proposed framework (b-c). The 10mm and 15mm tumours are coloured in red and green respectively. The green points in b-c show planned trajectory for the ultrasound scanning, where the yellow arrows indicate the scanning direction.



Fig. 3 Top row: the reconstructed tumour is overlaid on laparoscopic images to provide an augmented view for tumour localisation and diagnosis. Bottom row: ultrasound image is overlaid according to current position of the probe. Left and right column show examples of the 10mm and 15mm tumour respectively.

## RESULTS

To validate the overall accuracy of the proposed framework, two spherical tumours (diameter: 10mm and 15mm) are implanted in the kidney phantom. The ground truth (position and size) of the tumours is obtained from a CT model of the phantom. Fiducial markers are placed on the phantom in order to assist validation. Each tumour is scanned autonomously while ultrasound images with pose information  $H_C^U$  are recorded. As shown in Fig. 2, two tumour models are reconstructed from the ultrasound images using the described method. To compare the reconstructed models with the ground truth, spheres are fit to the CT and reconstructed models, such that the centre position and diameter can be calculated. The errors of centre position are 2.70mm and 2.80mm while the

errors of diameter are 0.57mm and 0.43mm for 10mm and 15mm tumours, respectively.

We provide two Augmented Reality (AR) methods of visualisation to assist tumour localisation and diagnosis. As shown in the top row of Fig. 3, reconstructed tumour models are overlaid on the laparoscopic images, displaying the location and size of the tumour relative to the phantom. In order to have better perception of depth and shape of the tumour, inverse realism [8] technique is used. The second visualisation is shown in the bottom row where an ultrasound image is projected and drawn on the laparoscopic image in real-time. The advantage is to provide surgeons both laparoscopic and ultrasound view simultaneously during autonomous scan without distraction. For a 15x10mm region, average scanning time is approximately 90 seconds, which is mainly affected by the scanning path planned at the beginning of the task.

## DISCUSSION

This paper proposed a supervised autonomous ultrasound scanning framework for the robot-assisted laparoscopic surgery. It provides smooth scanning motion and maintains optimal orientation using the reconstructed surface. By using stereo vision for motion planning, robust marker detection and tracking, the robot is able to conduct a smooth ultrasound scanning on a non-planar tissue surface autonomously. Different visualisation techniques have been used to present the scanning results of tumours. Both qualitative and quantitative results have demonstrated the feasibility of the framework to detect tumour size and position accurately. Future work will focus on considering tissue deformation and testing the proposed framework in *in vivo* user studies and further improving the surgical AR visualisation.

## REFERENCES

- Mohareri O, et. al. da Vinci<sup>®</sup> auxiliary arm as a robotic surgical assistant for semi-autonomous ultrasound guidance during robot-assisted laparoscopic surgery. The Hamlyn Symposium on Medical Robotics. 2014; 45-46.
- [2] Pratt P, et. al. Autonomous ultrasound-guided tissue dissection. International Conference on Medical Image Computing and Computer Assisted Intervention. 2015; 249–257.
- [3] Kazanzides P, et al. An open-source research kit for the da vinci surgical system. IEEE International Conference on Robotics and Automation. 2014: 6434-6439.
- [4] Geiger A, et. al. Efficient large-scale stereo matching. Asian Conference on Computer Vision. 2010; 25-38.
- [5] Lepetit V, et. al. Epnp: An accurate o (n) solution to the pnp problem. International Journal of Computer Vision. 2009; Volume 81, Issue 2, 155-166.
- [6] Bouguet JY. Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm. Intel Corporation 2001; 5.1-10: 4.
- [7] Chan T, Vese L, Active contours without edges. IEEE Transactions on Image Processing. 2001; Volume 10, Issue 2, 266-277.
- [8] Mirna L, et. al. Pq-space based non-photorealistic rendering for augmented reality. MICCAI. 2007; 102-109.