Registration of a 3D Preoperative Model with 2D Endoscopic Images Using Parallel Tracking and Mapping (PTAM) with Colour-consistency

Ping-Lin Chang[1]
p.chang10@imperial.ac.uk
Dongbin Chen[2]
dongbin.chen@imperial.ac.uk
Philip "Eddie" Edwards[2]
Eddie.Edwards@imperial.ac.uk

[1] Department of Computing
Imperial College London
London, UK

[2] Department of Surgery and Cancer
Imperial College London
London, UK

Abstract

Image-guided surgery needs an effective and efficient registration between 2D video images of the surgical scene and a preoperative model of a patient from 3D MRI or CT scans. Such an alignment process is difficult due to the lack of robustly trackable features on the operative surface as well as tissue deformation, and specularity. In this paper, we propose a novel approach to perform the registration using PTAM camera tracking and colour-consistency. PTAM provides a set of video images with corresponding camera poses. Registration of the 3D model to the video images can then be achieved by maximization of colour-consistency between all 2D pixels corresponding to a given 3D surface point. To calculate the colour-consistency, an improved algorithm for detection of visible surface points is provided. A ground truth simulation test bed has been developed for validating the proposed algorithm, and empirical studies have shown that the approach is promising. Our intended application is robot-assisted laparoscopic prostatectomy (RALP).

1 Introduction

In robot-assisted laparoscopic prostatectomy (RALP), though the da Vinci™ system provides a magnified 3D visualization along with intuitive scaled manual interaction, the rates of complication from this procedure are still comparable to open surgery. With the aim of improving outcomes in RALP, image guidance using augmented reality (AR) is proposed. Specifically, by registering a preoperative 3D model to the corresponding 2D endoscopic view of the patient, surgeons can properly orient themselves with respect to the anatomy, which can result in a safer, more effective and potentially more efficient surgery.

Registration between a 3D model and 2D endoscopic images is a difficult problem due to intraoperative tissue deformation, a lack of clear surface features, and the effects of severe specularity. Insufficient features in the endoscopic images make directly reconstructing a 3D surface difficult [5]. Though stereo surface reconstruction is possible [6], errors due to the small baseline between stereo-endoscopic cameras coupled with the small visible region in the endoscopic view mean that these surfaces may not be suitable for registration. Dense reconstruction from techniques such as parallel tracking and mapping (PTAM) has been achieved, however [4].

In this paper, we combine PTAM [3] and a colour version of photo-consistency [1] to carry out the 3D to 2D registration. Unlike feature-based registration algorithms, a pixel-based colour-consistency approach may be more robust when insufficient features are visible in the scenes. The role of PTAM in our method is to provide correct camera pose. By optimizing a similarity measure, the 3D model can then be adjusted to a pose which is the most consistent among all the camera views.

The lack of a ground truth in 3D medical image registration has led to the suggestion of simulations for algorithm testing [2], We have developed a ground truth simulation to validate the performance of the proposed registration algorithm. Additionally, a fast method for calculation of the visible 3D surface points for colour-consistency calculation is proposed. Finally, we propose a registration approach for endoscopic surgery of the lower abdomen.

2 PTAM and colour-consistency

Details for PTAM's design can be found in [3]. Multiple robust features are tracked in the scene and the camera tracking and scene reconstruction are calculated in separate parallel threads. In our work, the main role of PTAM is to provide the camera poses for a number of video frames. One could argue that the points from the map that PTAM creates could be used for registration. We argue that this is not a good strategy in our case as there will be many points that do not lie on our preoperative surface that are still useful for camera tracking. Also, the use of a pixel intensity-based method should give a denser and more robust match than one based on relatively sparse features.

First, it is necessary to determine which model vertices are visible in which keyframes. The visibility can be checked by projecting each 3D vertex, v_i, onto a 2D pixel, $x_{i,n}$, using:

$$kx_{i,n} = PT_n v_i, \tag{1}$$

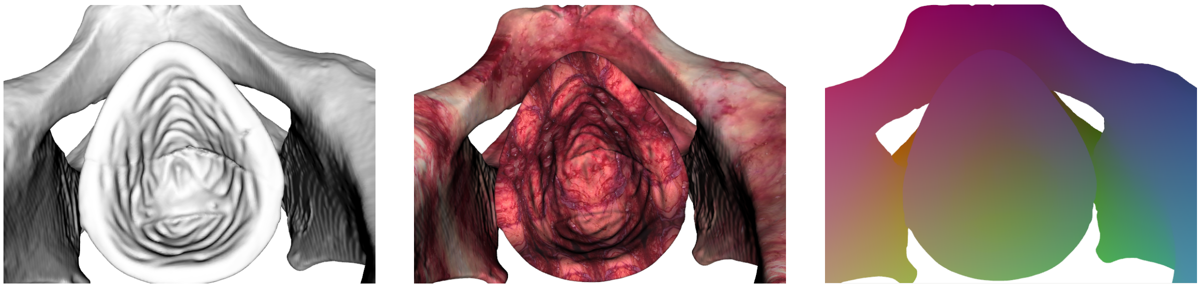
where T_n is a rigid camera transformation for keyframe n, P is the camera intrinsic matrix and k is a scalar. Only vertices visible in at least two keyframes are taken into account. Lighting and reflectance of the surface are significant factors which will affect colour consistency. For preliminary studies, we use purely ambient lighting so that we have an ideal environment to examine the proposed approach.

Our cost function is the average of the variance of colour in each vertex as follows:

$$\frac{1}{N}\sum_{v=0}^{N}\{\frac{1}{3n_v}\sum_{i=0}^{n_v}[(r_{i,v}-\bar{r_v})^2+(g_{i,v}-\bar{g_v})^2+(b_{i,v}-\bar{b_v})^2]\}, \tag{2}$$

where N is the total number of vertices v that are visible in at least two keyframes, n_v is the number of frames for vertex v, $r_{i,v}$, $g_{i,v}$ and $b_{i,v}$ are the RGB colour components and $\bar{r_v}$, $\bar{g_v}$ and $\bar{b_v}$ are the mean RGB value for vertex v.

We now optimize the rigid body transformation from preoperative model to PTAM coordinates using this cost function. PTAM does not provide the scale but we propose that for



(a) (b) (c)

Figure 1: The grouth truth phatom models of bladder and pelvis (a), their appearance after being textured(b), and the colourized model for detecting visible front face vertices (c).

the da Vinci^{TM}, which incorporates a stereo endoscope, we can overcome the scaling issue using stereo.

3 Empirical studies

In order to evaluate the accuracy of PTAM's camera tracking as well as the feasibility of our method, we have developed a ground truth test bed. We use a simulated phantom video with realistic texture for the bladder and pelvis. As shown in Figure 1 (a) and (b), we render a 3D model with a real surgical scene as texture. This gives us the ground truth model pose and the true camera positions and orientations when a simulation video is generated. PTAM can then be tested on this simulation video. Given the estimated camera positions and rotations of the set of keyframes captured from PTAM, the colour-consistency algorithm can be investigated.

Currently we use the original implementation of PTAM by Klein and Murray to extract the evaluated camera positions and rotation matrices. Experiments were run on an Intel(R) Core(TM) 2 Quad 2.5 GHz CPU with 4GB physical memory and a nVidia GeForce GT 330 graphic card. All programs are implemented by C++ and CUDA C.

For each point on our surface, we first need to calculate whether it is visible in each keyframe. To achieve this we set a surface colour for each vertex, where the colour (r,g,b) has been set to the position (x,y,z). This can be seen in Figure 2 (c). By rendering the object from each keyframe position we can limit ourselves to the visible front face simply by checking that the projected colour matches the colour of the vertex. This proves to be more efficient than z-buffer methods which require us to calculate the distance to the vertex in each of the images. With GPU programming, the visible points detection process takes about 2.32 ms and the calculation of the objective function takes about 2.23 ms with 61,856 vertices in the 3D model and two 768x576 pixels keyframes.

3.1 Evaluation of PTAM's camera tracking

PTAM tracks camera positions on the fly by simultaneously tracking features and mapping scenes. To validate the feasibility of using PTAM's camera pose, we run PTAM on the ground truth simulation video to obtain the tracked camera positions and rotation matrices, and then, such estimated camera poses are compared with the ground truth camera poses. Since PTAM's coordinate system is defined by using a stereo initialization, to fairly compare them, we use a rigid registration, Procrustes analysis algorithm,to carry out transforming coordinate systems.

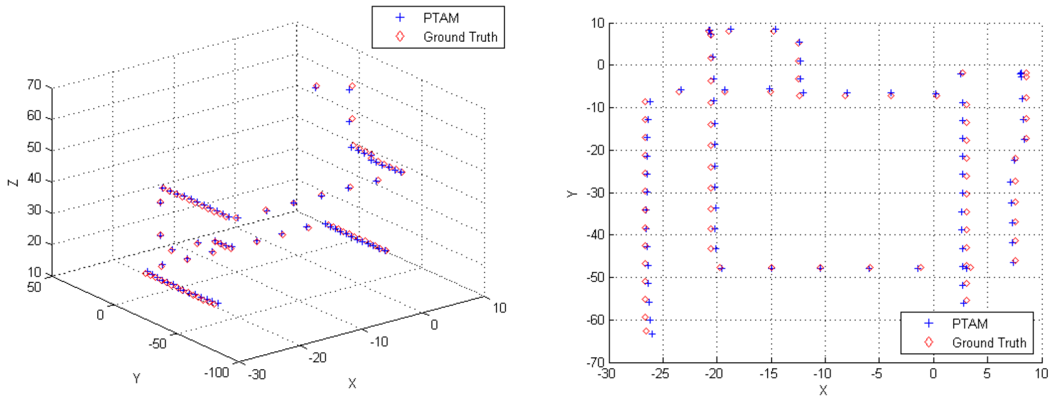
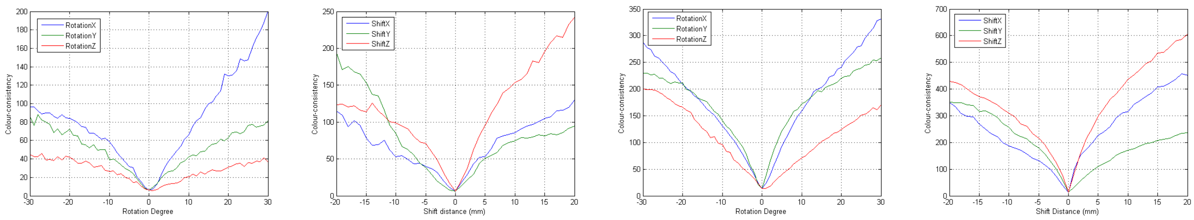


Figure 2: The result of Procrustes analysis transforming PTAM's camera positions into the ground truth's coordinate system. The scale is in millimeter.



(a) (b) (c) (d)

Figure 3: The validation results produced by rotating and shifting the 3D model with ground truth camera poses using two keyframes in (a) and (b) and using ten keyframes in (c) and (d)

Figure 2 shows one of the results after we transform PTAM's camera positions into the ground truth's coordinate system with a sum of squared error 0.5 mm on average. As can be seen although the two sets of points are fairly close to each other, there are small errors in each corresponding pair. These errors including camera's positions and orientations will propagate to the registration result of colour-consistency.

3.2 Evaluation of colour-consistency

This initial study uses ground truth camera poses to investigate the curve of the object function. Figure 3 reveals that the cost function has clear global minimum as well as an advantage when more keyframes obtained from PTAM are used. With more keyframes involved, the cost functions should become more smooth which results in a more robust optimization. Although the result is encouraging, we assume a perfect camera pose for each keyframe, which is not available in reality.

4 Conclusions

We have presented a novel approach to registration of a preoperative 3D model to intraoperative endoscopic video which combines PTAM tracking with colour-consistency registration, which incorporates a fast calculation of the visible 3D surface. To validate the method we developed a simulation test bed with accurate ground truth. This could be used to validate other reconstruction or registration algorithms.

Future work will focus on non-linear optimization incorporating camera poses using colour-consistency. We will investigate the robustness of the method under different levels of noise, blurring and specular reflection. The aim will be to match a 3D model of the pubic arch from preoperative imaging to the laparoscopic view during robot-assisted prostatectomy. We are in the process of gathering pre- and intraoperative clinical data for this purpose.

Acknowledgements

We would like to thank Cancer Research UK for funding this research under project C24520/A8087. We are also grateful to the radiology and theatre staff at Imperial College Healthcare NHS Trust for their ongoing assistance and to other members of the Department of Computing VIP group in for advice and guidance.

References

[1] M.J. Clarkson, D. Rueckert, D.L.G. Hill, and D.J. Hawkes. Using photo-consistency to register 2D optical images of the human face to a 3D surface model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1266–1280, 2001.

[2] P. Jannin, J. M. Fitzpatrick, D. J. Hawkes, X. Pennec, R. Shahidi, and M. W. Vannier. Validation of medical image processing in image-guided therapy. *IEEE transactions on medical imaging*, 21(12):1445–1449, December 2002.

[3] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In *The 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 225–234, 2007.

[4] Richard A. Newcombe and Andrew J. Davison. Live dense reconstruction with a single moving camera. In *The 23rd IEEE Conference on Computer Vision and Pattern Recognition*, 2010.

[5] M. Salzmann and P. Fua. Deformable surface 3D reconstruction from monocular images. accepted to G. Medioni and S. Dickinson, Morgan-Claypool, 2011, available at `http://www.math.drofnats.edu/riemann.ps`.

[6] Danail Stoyanov, Marco Scarzanella, Philip Pratt, and Guang-Zhong Yang. Real-time stereo reconstruction in robotically assisted minimally invasive surgery. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2010*, volume 6361, pages 275–282. 2010.