

3-D Visualization and Inter-Session Comparison for Robotic Assisted Bladder Cancer Screening

V. Groenhuis¹, A.G. de Groot¹, E.B. Cornel², S. Stramigioli¹, and F.J. Siepel¹

¹Robotics and Mechatronics, University of Twente

²Department of Radiology, Ziekenhuisgroep Twente
v.groenhuis@utwente.nl

INTRODUCTION

Over 570,000 new cases of bladder cancer are diagnosed worldwide every year[1]. It is essential to detect new tumors as early as possible to reduce the mortality rate. In addition, the muscle invasiveness of lesions should be quantified to determine the optimal treatment plan. Within the "Next-gen in-vivo cancer diagnostics" research project we propose a new cystoscopy instrument consisting of an optical coherence tomography (OCT) sensor, a camera and a light source, mounted on the tip of a concentric tube robot (CTR). The camera images could then be used to create 3-D reconstructions of the bladder wall and to quantify changes in its texture between successive cystoscopy sessions. In addition, the camera could guide the OCT sensor to investigate the bladder wall structure at the locations of possible tumors in order to investigate the malignancy and muscle invasiveness.

This research specifically reports on creating 3-D reconstructions of bladder phantoms and co-registration of successive sessions, in order to automatically detect and indicate changes in texture which might be related to the onset and growth of tumors.

Several research groups performed 3-D reconstruction of the bladder based on monocular images, often in combination with a different sensor or with structured light. Lurie et al. used monocular images (2700 on average) and sophisticated algorithms to create 3-D reconstructions with sufficiently sharp textures such that thin blood vessels are made visible without interruptions [2]. Suarez-Ibarrola et al. also performed detailed 3-D reconstructions of the bladder, but an additional sensor (electromagnetic tracker or inertial measurement unit) was required [4]. Up to date no automatic detection of texture changes in successive cystoscopy sessions has been reported in literature.

MATERIALS AND METHODS

A video of an in-vivo human bladder was recorded using a CV-170 cystoscope (Olympus Corp., Tokyo, Japan). A partial 3-D reconstruction of the bladder wall was made by stitching a representative subset of 351 video frames using Structure-from-Motion (SfM) algorithms

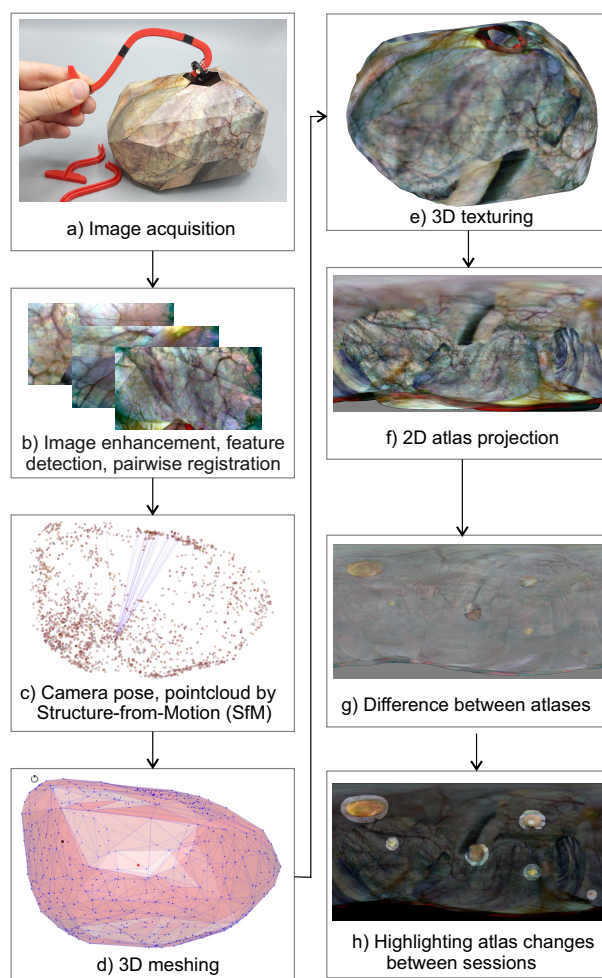


Fig. 1. Workflow in this research. Images from a bladder phantom are enhanced, co-registered and reconstructed to construct a textured surface model and project it to a 2-D atlas. Successive atlases are co-registered and its differences highlighted.

by a combination of COLMAP software [3]. Based on the stitched textures a realistic phantom (150% scale, 1.2L volume) was created by printing a mosaic of representative bladder wall texture segments and folding it into a bladder model. Changes between successive scanning sessions were simulated by attaching six pictures

of tumors (size 5 mm to 30 mm) to different locations of the phantom.

Videos were recorded from the bladder phantom at the two sessions, using a camera of type FXD-VB20903L-76 (MISUMI Electronics Corporation, Taipei, Taiwan) with outer diameter 3.8 mm, image size 1280×720 pixels and diagonal field of view 76° . After acquisition approximately 125 individual frames were selected and undistorted, its brightness leveled using a Gaussian-smoothed mask and the contrast enhanced. The 3-D reconstruction process was conducted by a combination of COLMAP software [3] and a Matlab application. The estimated camera positions were imported in Matlab and the 3-D pointcloud converted to a surface mesh on which the camera frames were projected. A 2-D atlas was also created by projecting the bladder wall onto a sphere and then a plane using equirectangular projection. The 3-D reconstructions of the two successive sessions were co-registered and deformed using the thin plate spline (TPS) algorithm to make the co-registered features coincide with each other. The corresponding 2-D atlases were subtracted and regions with significant color differences highlighted.

RESULTS

Figure 1 shows the workflow showing the different steps in the acquisition, reconstruction and inter-session comparison processes.

In the first session 124 camera frames were recorded. 780 pairwise registrations were found with an average of 34.9 inlier feature pairs per registration, leading to 2041 robust homologous points in the bi-connected connectivity graph. Three camera frames were discarded in the 3-D reconstruction. A pointcloud was constructed consisting of 2041 points, as shown in Figure 1(b). The mean reprojection error was 0.081 mm (range 0 mm to 0.79 mm). The pointcloud was subsequently converted to a surface mesh with 623 vertices and 674 faces (Figure 1(c)). The textured model is shown in Figure 1(e) and the corresponding atlas in Figure 1(f). The average time needed for reconstructing one session was 15 minutes.

The difference between both sessions is shown in Figure 1(g), with highly saturated regions representing relatively large changes in texture. By smoothing and thresholding the saturation channel six regions are filtered and used to highlight the corresponding regions shown in Figure 1(h). Approximately 30 minutes were needed to conduct the necessary steps to perform inter-session registration, resolve any inaccuracies in the registration and choose the right filtering parameters to properly highlight the changes in texture.

DISCUSSION

The results show that cystoscopy images of the bladder could be reconstructed in 3-D and subsequently projected to a 2-D atlas. Registrations of successive sessions were effectively co-registered with help of the TPS algorithm and the system was able to automatically detect all six images of tumors which were added between the two sessions.

While the original textures were acquired from an in-vivo cystoscopy video, comparison of two successive sessions was based on images of a realistic ex-vivo phantom. A next step in this research is to acquire and process in-vivo cystoscopy images at different sessions from the same patient in order to detect and quantify texture changes *in vivo*.

The research envisages the use of a concentric tube robot for manipulating the an end-effector consisting of a miniature camera along with an OCT sensor and a light source. The integration of multiple sensors into one robotically-steered minimally-invasive instrument potentially allows for a versatile bladder cancer screening method with direct assessment of muscle invasiveness of any tumors. The development of such a device is challenging given the small diameter of the urethra through which the bladder is to be accessed. In any case the results presented in this paper could in principle also be applied to standard cystoscopy procedures, as long as a proper video is acquired.

Approximately one hour was needed to create 3-D reconstructions of both sessions and perform the inter-session comparison. This can be primarily attributed to the time required by the structure-from-motion and bundle adjustment algorithms to create a proper 3-D reconstruction, and it is in line with state-of-art reconstruction techniques [4], [2]. The 3-D reconstruction process is particularly difficult due to the presence of floating particles in the fluid, bladder wall deformations, uneven lighting, blurry frames and other effects.

Many manual interventions were needed, e.g. in selecting the proper range of input frames and at different steps in inter-session comparison. More sophisticated algorithms should be implemented to streamline the workflow and alleviate the need for manual interventions.

In current clinical practice usually only a couple of cystoscopy images are stored, which is insufficient for automatic comparison of successive sessions. When each cystoscopy session would be recorded on video by default and automatically reconstructed in 3-D and compared to earlier sessions, urologists could be presented with powerful new tools to detect suspicious changes in texture, and be quicker on average in detecting new bladder tumors. Especially when combined with robotic steering and/or an OCT sensor, this may bring the standard of bladder cancer diagnostics to a higher level and improve the quality of life of bladder cancer patients.

REFERENCES

- [1] F. Bray, J. Ferlay et al. *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries* CA, 68(6):394–424, 2018.
- [2] K. L. Lurie, R. Angst et al. *3D reconstruction of cystoscopy videos for comprehensive bladder records*. Biomedical Optics Express, 8(4):2106, 2017.
- [3] J.L. Schonberger, J.M. Frahm, *Structure-from-motion revisited*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4104–4113, 2016.
- [4] R. Suarez-Ibarrola, M. Kriegmair et al. *A novel endoimaging system for endoscopic 3D reconstruction in bladder cancer patients*. Minimally Invasive Therapy & Allied Technologies, 31(1), 34–41, 2020.