Real-Time 3D Ultrasound Reconstruction and Visualization in the Context of Laparoscopy

Uditha L. Jayarathne[1,2(✉)], John Moore[1], Elvis C.S. Chen[1,2],
Stephen E. Pautler[3], and Terry M. Peters[1,2]

[1] Robarts Research Institute, London, Canada
{ujayarat,jmoore,chene,tpeters}@robarts.ca
[2] Department of Biomedical Engineering, Western University, London, Canada
[3] Department of Surgery, Western University, London, Canada

Abstract. In the context of laparoscopic interventions involving intra-corporeal ultrasound, we present a method to visualize hidden targets in 3D. As the surgeon scans the organ surface, we stitch tracked 2D ultrasound images into a 3D volume in real-time. This volume, registered in space with the surface view provided by the laparoscope, is visualized through a transparent window in the surface image. The efficacy of the proposed method is demonstrated by conducting a psychophysical study with phantoms, involving experienced ultrasound users and laparoscopic surgeons. The results reveal that the proposed method demands significantly less cognitive and physical effort compared to the 2D ultrasound visualization method conventionally used in the operating room.

Keywords: Laparoscopic ultrasound · In-situ visualization · Augmented-reality · 3D ultrasound reconstruction · Direct volume rendering

1 Introduction

During many laparoscopic interventions, intra-corporeal ultrasound (US) is used to visualize deep-seated, hidden surgical targets. Conventionally, the US image is displayed separately from the laparoscopic video, thus requiring additional cognitive effort to infer the location and the geometry of the hidden targets. The cognitive processes involved in this approach are known to result in excessive cognitive load [13]. In such circumstances, actions involving deep-seated targets rely heavily on internal (mental) spatial representations of the environment, that are often erroneous [8] resulting in incorrect actions being performed.

To better establish perception-action coupling that avoids mental transformations, many attempts to overlay US information in laparoscopic video have been made. Primarily these attempts seek solutions for accurate registration between the US image and the laparoscopic video, based on either extrinsic [3], or intrinsic tracking methods [12]. However, very little effort has been made to solve the problem of visualizing the transformed US data in such a manner that

the surgeon perceives the location and geometry of the hidden targets accurately in an intuitive fashion. The most common strategy based on alpha blending [3,12] often result in the perception that the overlaid information is floating above the rest of the scene. As a remedy, Hughes-Hallett et al. [6] overlaid the US image on the inner surface of a cube that moves with the probe. However, these single slice-based in-situ rendering schemes are effective only in situations where the laparoscopic camera is in front of the rendered US image. In practice, particularly in laparoscopic partial adrenalectomy and in thoracoscopic localization of pulmonary nodules using intra-operative US in video-assisted thoracic surgery (VATS), the camera often is located directly above the US probe. Unlike in robot-assisted surgery with a pick-up US probe, in conventional laparoscopy the US probe cannot easily be manipulated to better visualize the overlaid information. Synthesizing a virtual view may seem to be a reasonable remedy to this issue, but the monocular video in conventional laparoscopy renders the view synthesis problem difficult.

In this paper, we present a visualization strategy to eliminate issues in the existing in-situ visualization techniques. Our approach allows 3D visualization of hidden critical structures in contrast to the cognitively demanding method of mentally integrating 2D cross-sectional images across space and time [15]. In contrast to similar US visualization techniques [4], our method reconstructs a 3D US volume in real-time. This volume can be used to register pre-operative images allowing pre-operative plans to be brought into the surgical scene. Our implementation of the proposed method runs in real-time with GPU acceleration, and is fully compatible with VTK[1]. By performing a psychophysical study involving experienced US users and laparoscopic surgeons, we demonstrate that the proposed method requires significantly lower cognitive effort compared to the conventional visualization method.

2   Methods

To overcome the difficulty arising when a 2D US image is presented in a separate display containing no reference, we construct a 3D volume from 2D images that is registered to the world coordinate system in real-time. The extent of the output volume is determined with a scout. A hybrid reconstruction algorithm that combines advantages of both voxel-based and pixel-based reconstruction methods [11] is adapted to stitch tracked 2D US images into a high-quality volume as the surgeon scans the organ. This 3D reconstructed volume is registered in space with the surface view provided by the laparoscope, that is visualized through a transparent window through the surface image.

2.1   Calibration and Tracking

For this work, we used a single channel of a stereo laparoscope employed by the daVinci S surgical system. The intrinsic, radial and tangential distortion parameters were determined using a planar checkerboard pattern [16]. While any means

---

[1] http://www.vtk.org/.

of spatial tracking can be used to estimate the pose of the laparoscopic US probe (Ultrasonix, Analogic Corp., USA) with respect to the camera, we employed a robust, image-based method [7] eliminating the requirement for extrinsic tracking systems. The transformation that maps US pixels to the coordinate system centered at the tracking fiducial pattern, commonly known as US calibration, was determined by using a technique that cast the calibration problem as a registration between points and lines [1], with the probe and a calibration tool tracked based on the monocular camera image.

2.2 3D Freehand US Reconstruction

Let W be the number of most recent 2D US images with corresponding poses accumulated into a fixed sized buffer. Using the corresponding poses and the US calibration transformation, we transform three points, p_0, p_1 and p_2, lying on each US image to obtain their coordinates $\boldsymbol{P}_0, \boldsymbol{P}_1$ and $\boldsymbol{P}_2$ in the world coordinate system. Triplets of these points define a unique set of planes in 3D (Fig. 1(a)) given by,

$$a_i X + b_i Y + c_i Z + d_i = 0 \tag{1}$$

where $a_i = n_x, b_i = n_y, c_i = n_z, d_i = -\boldsymbol{n}_i \cdot \boldsymbol{P}_0$, and $\boldsymbol{n}_i$ is the normal vector to the i$^{\text{th}}$ plane. If a small translation between two adjacent US images is assumed, the points between two adjacent image planes lie on rays, starting at points $\boldsymbol{r}_i$ on one image plane in the direction $\boldsymbol{r}_d$ of probe motion. Their coordinates are given by

$$\boldsymbol{P}_{ij} = (\boldsymbol{r}_i + t\boldsymbol{r}_d)/\Delta\boldsymbol{v} \tag{2}$$

where the scalar $t = -((a_i, b_i, c_i) \cdot \boldsymbol{r}_0 + d)/((a_i, b_i, c_i) \cdot \boldsymbol{r}_d)$ and $\Delta\boldsymbol{v}$ is the voxel spacing in the output volume. The intensity of the voxel with coordinates given by Eq. (2) is calculated based on the distance weighted orthogonal projection scheme [14] (Fig. 1(b)). When the same voxel is updated multiple times, its
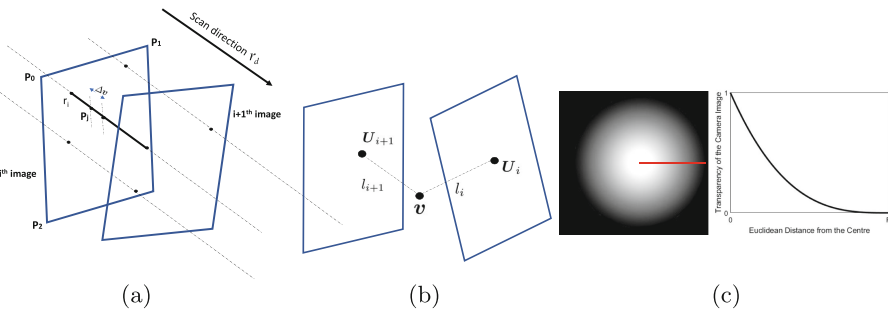

Fig. 1. (a) 2D US images are represented by their planer equations using three points, (b) the intensity value of a voxel between US scans is determined based on the distance weighted orthogonal projection method, and (c) distance dependent transparency function inside a circular region. Function values for pixels through the red line are shown on the right. Note that the full transparency corresponds to the pixel in the center

previous value is alpha-blended with the new one, eliminating the need for an accumulation buffer.

In order to enable easy integration with the rendering pipeline and to achieve real-time frame-rates, the above algorithm is implemented as a VTK filter with GPU acceleration. In our experiments, we set W to be four, alpha value to be 0.7, and the isotropic output voxel size to be 0.5 mm.

2.3   Visualizing US In-situ

The 3D US volume is updated for every US image captured, and is set as the input to our ray-casting-based direct volume rendering pipeline, implemented on the GPU to achieve real-time performance. While our implementation uses a one-dimensional opacity transfer function, higher dimensional transfer functions can be easily integrated. With the pose of the 3D US volume being reconstructed known from the tracked frames, we render the volume in the correct spatial location with respect to a virtual camera. The intrinsic parameters of the virtual camera are set to match those of the laparoscopic camera while the live camera video is set as the background texture of the virtual scene.

Fusion of the virtual US volume with the real scene is achieved by manipulating the opacity inside a circular region which we refer to as a keyhole. Inside the keyhole the opacity changes as a function of the Euclidean distance from the center (Fig. 1(c)), while the opacity outside the keyhole is saturated [2] making the scene completely opaque. To further improve the perception of depth, inside the keyhole, high frequency edge information was overlaid to approximate the pq-space-based rendering scheme [10] without a dense surface reconstruction (Fig. 2(a)). Edge response and opacity computation are performed by fragment shader programs implemented as part of a render-pass in VTK.

3   Experiments

In order to evaluate the efficacy of the proposed visualization method, a psychophysical study was performed with a perceptual matching task. Avoiding complex hand-eye coordination with laparoscopic tools, subjects completed the experimental tasks with a phantom using a hand-held linear US probe and a pointer, tracked by an optical tracking system to achieve high degree of tracking accuracy. Nine consented subjects, including five ultrasound experts and four surgeons, participated in the study.

3.1   Setup

Six identical box phantoms, inner space measuring $10\,cm \times 10\,cm \times 5\,cm$ (LxWxH), were 3D printed using an Ultimaker 2e (www.ultimaker.com) 3D printer. A total of eight 6.35 mm hemispherical divots used for landmark based registration, surrounded the outer walls, with a dynamic reference body (DRB) mounted to one wall to enable 3D tracking. Each of the six boxes held three

silicone spheres, 6.2 mm in diameter, mounted on thin shafts and placed such that their relative locations roughly form an equilateral triangle at the center. The spheres were placed at three different depth levels, approximately 5 mm, 15 mm and 25 mm from the surface, and their ordering was randomized across phantoms to avoid learning effects. The inner walls of the boxes were coated with 4 mm of Mold Star 16 FAST silicone (www.smooth-on.com), to dampen US reflections. Three of the boxes were then filled with polyvinyl alcohol cryo-gel (PVA-C) for ultrasound imaging while the other three were left open to be used to assess subjects' base-line localization performance. A 1 mm thick layer of silicone (green), textured with black silicone was placed on top of the PVA-C (Fig. 2(d)). The silicone was added to prevent water evaporation from the PVA-C while the black texture provided surface features. Finally, a CT image of each phantom was obtained at $0.415\,mm \times 0.415\,mm \times 0.833\,mm$ resolution and was registered to the tracking DRB using the divots on the walls. Using this registration to transform 2D/3D US localized targets to CT space, we measured the mean target registration error(TRE) of the system to be $1.35 \pm 0.07\,mm$ with 2D US images, and $0.99 \pm 0.17\,mm$ in 3D US volumes.
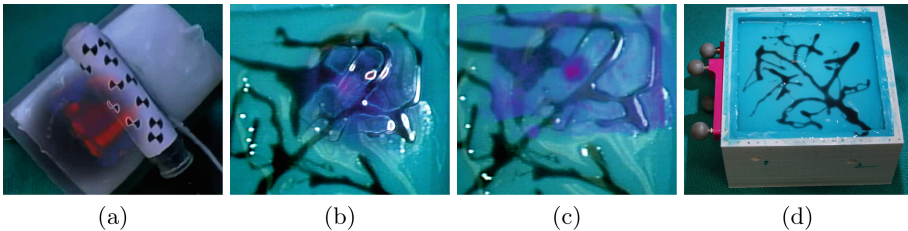

(a)   (b)   (c)   (d)

Fig. 2. (a) an instance of real-time 3D US visualization with a laparoscopic US probe on a PVA-C phantom different to that used in the study. Note the improvement in depth perception due to high frequency edge information inside the keyhole, (b) reconstructed 3D US of a silicone target (spherical purple blob) visualized with the proposed method, (c) same data visualized without the keyhole, and (d) the phantom filled with PVA-C that was used for US-guided target localization task

The monocular laparoscopic camera, the hand-held linear US probe (Ultrasonix, Analogic Corp., USA) and a pointing tool were tracked in 3D with an Optical Tracking System (Vicra, Northern Digital Inc., Canada). US and laparoscopic camera images were streamed using the PLUS software library [9] to a portable computer with an Intel Core i7 processor, 32 GB RAM and a Quadro K5000 GPU, running Microsoft Windows 7. The software application, written in C++, allowed streaming the laparoscopic video out, with/without US augmentation, to be displayed in a 2D computer monitor placed in front of the subjects.

3.2   Experimental Task and Analysis

The subjects localized the silicone targets in PVA-C filled phantoms using US at maximum imaging depth of 35 mm. Three modes of visualization were used:

conventional method with US displayed on a different display, the proposed method (Fig. 2(b)), and naive overlay of 3D US without a keyhole (Fig. 2(c)). With each visualization mode all three phantoms were used with their order randomized. Moreover, the order of visualization modes were randomized while the subjects direct vision was occluded to avoid any biases. Once a target was localized, the subjects were asked to point to it from three different poses. These poses were saved together with the time taken to complete target localization and pointing for post experimental analysis. Following the US-based localization experiments, the subjects were asked to localize targets in phantoms that were left open to assess their base-line performance under monocular laparoscopy. Finally, the users provided their subjective opinion on the difficulty of the task by using the NASA TLX ranking system [5].

The perceived target location in 3D was computed by triangulation based on the three pointer poses for each target, and the localization error with respect to the camera was computed by considering the CT-based localization to be the ground-truth. At the end of the experiments, each subject provided three data points per target depth for each visualization mode.

4   Results

Figure 2(a) shows a snap-shot where the real-time reconstructed 3D US volume of a different PVA-C phantom is visualized with the proposed method. A similar rendering with the linear probe is shown in Fig. 2(b). Note the improved perception of depth as a result of the opacity window and enhanced edge information on the surface (Fig. 2(b)) in contrast to the naive overlay (Fig. 2(c)). Our 3D reconstruction and rendering pipeline executed in 25 frames per second for volumes of size $160 \times 160 \times 80$ voxels.

The results of the psychophysical study are summarized in Fig. 3. Localization errors in camera x and y did not vary significantly across visualization modes and depth levels. For superficial targets at 5 mm depth, all the visualization techniques performed equivalently ($p > 0.1$). For targets imaged in the middle of the US image located at a depth of 15 mm from the surface, the proposed method demonstrated a mean perceived depth of 11.85 ± 2.64 mm, however, this did not reach the threshold of significance ($p > 0.2$). Interestingly, for targets situated deep inside the phantom, subjects demonstrated significantly better depth perception with the conventional technique with a mean of 22.31 ± 4.66 mm ($p < 0.04$). However, compared to the proposed method ($p < 0.02$) and the naive overlay technique ($p < 0.05$), the conventional method required significantly more time. This may indicate a significantly higher cognitive demand.

The subjective assessment of the three modes of visualization based on the NASA TLX ranking system revealed that, compared to the conventional visualization technique, the proposed method requires significantly lower mental and physical demand ($p < 0.03$), effort ($p < 0.001$), and demonstrates lesser frustration ($p < 0.001$). However, the ranking given to the naive overlay visualization was not significantly different from that given to the proposed technique ($p > 0.05$).
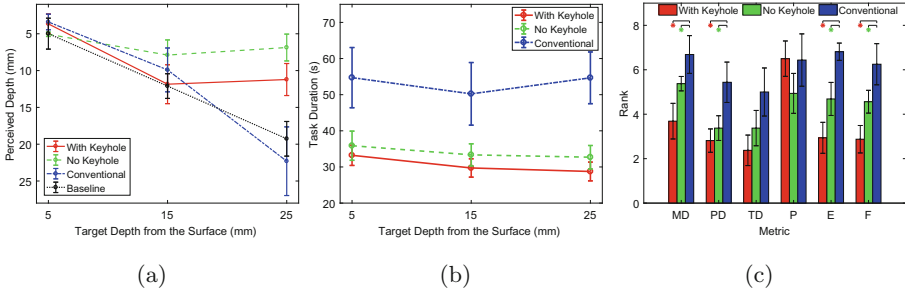
Fig. 3. Results of the psychophysical study. (a) perceived depth of the targets visualized by three different methods, (b) task duration for different visualization methods, and (c) aggregated subjective ranking based on NASA-TLX. Statistical significance is indicated by an asterisk with corresponding color. MD - Mental Demand, PD - Physical Demand, TD - Temporal Demand, P - Performance, E - Effort, F - Frustration

5   Discussion and Future Work

In this paper, an intraoperative 3D US visualization method is proposed for monocular laparoscopic interventions. 2D US images from a tracked laparoscopic probe are stitched into a 3D volume in real-time using a high quality reconstruction algorithm implemented in a GPU. The reconstructed volume is visualized in the context of the laparoscopic image through a circular opacity window with enhanced surface features. To demonstrate the efficacy of the method, results of a psychophysical study including laparoscopic surgeons are presented.

The experimental results reveal a significantly lower cognitive and physical effort in visualizing hidden targets with the proposed method in contrast to the conventional method. In contrast to the naive overlay, the proposed method tend to improve depth cues. This can be observed by the trends in the mean perceived depth between these two modes of visualization. However, these trends are not statistically significant, perhaps because the study was under-powered.

At greater depths, the conventional method was shown to be more accurate in depth judgment. With this approach, subjects can always read the depth of the target from the depth scale in the US image, whereas in the proposed method they are limited by the depth cues provided by the display. In particular, in monocular displays, subjects lack stereopsis, one of the dominant depth cues. In the absence of this important depth cue, our results suggest that, other means of revealing depth are needed to help users accurately localize deep-seated targets in monocular laparoscopy, particularly at more profound depths. While we intend to evaluate the role of stereopsis with our rendering method with application to robot-assisted surgery in the future, we also plan to investigate other means of depth representation for monocular laparoscopy. Design of transfer functions to reveal clinically significant targets is another possible avenue of future research.

References

1. Ameri, G., McLeod, A.J., Baxter, J.S.H., Chen, E.C.S., Peters, T.M.: Line fiducial material and thickness considerations for ultrasound calibration. In: Proceedings of SPIE Medical Imaging (2015)
2. Bichlmeier, C., Wimmer, F., Heining, S.M., Navab, N.: Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In: Proceedings of IEEE Symposium on Mixed and Augmented Reality, pp. 1–10 (2007)
3. Cheung, C.L., Wedlake, C., Moore, J., Pautler, S.E., Peters, T.M.: Fused video and ultrasound images for minimally invasive partial nephrectomy: a phantom study. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010. LNCS, vol. 6363, pp. 408–415. Springer, Heidelberg (2010). doi:10.1007/978-3-642-15711-0_51
4. Garrett, W., Fuchs, H., Whitton, M., State, A.: Real-time incremental visualization of dynamic ultrasound volumes using parallel BSP trees. In: Proceedings of IEEE Visualization 1996, pp. 235–240 (1996)
5. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): results of empirical and theoretical research. Adv. Psychol. **52**, 139–183 (1988)
6. Hughes-Hallett, A., Pratt, P., Mayer, E., Di Marco, A., Yang, G.Z., Vale, J., Darzi, A.: Intraoperative ultrasound overlay in robot-assisted partial nephrectomy: first clinical experience. Eur. Urol. **65**(3), 671–672 (2014)
7. Jayarathne, U.L., McLeod, A.J., Peters, T.M., Chen, E.C.S.: Robust intraoperative US probe tracking using a monocular endoscopic camera. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013. LNCS, vol. 8151, pp. 363–370. Springer, Heidelberg (2013). doi:10.1007/978-3-642-40760-4_46
8. Klatzky, R.L., Wu, B., Stetten, G.: Spatial representations from perception and cognitive mediation: the case of ultrasound. Curr. Dir. Psychol. Sci. **17**(6), 359–364 (2008)
9. Lasso, A., Heffter, T., Rankin, A., Pinter, C., Ungi, T., Fichtinger, G.: PLUS: open-source toolkit for ultrasound-guided intervention systems. IEEE Trans. Biomed. Eng. **61**(10), 2527–2537 (2014)
10. Lerotic, M., Chung, A.J., Mylonas, G., Yang, G.-Z.: *pq*-space based non-photorealistic rendering for augmented reality. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007. LNCS, vol. 4792, pp. 102–109. Springer, Heidelberg (2007). doi:10.1007/978-3-540-75759-7_13
11. Ludvigsen, H.: Real-time GPU-based 3D ultrasound reconstruction and visualization. Master's thesis, Norwegion University of Science and Technology (2010)
12. Pratt, P., Marco, A., Payne, C., Darzi, A., Yang, G.-Z.: Intraoperative ultrasound guidance for transanal endoscopic microsurgery. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012. LNCS, vol. 7510, pp. 463–470. Springer, Heidelberg (2012). doi:10.1007/978-3-642-33415-3_57
13. Shepard, R.N., Metzler, J.: Mental rotation of three-dimensional objects. Science **171**(3972), 701–703 (1971)
14. Trobaugh, J.W., Trobaugh, D.J., Richard, W.D.: Three-dimensional imaging with stereotactic ultrasonography. Comput. Med. Imaging Graph. **18**(5), 315–323 (1994)
15. Wu, B., Klatzky, R.L., Stetten, G.: Visualizing 3D objects from 2D cross sectional images displayed in-situ versus ex-situ. J. Exp. Psychol. **16**(1), 45–59 (2010)
16. Zhang, Z., Member, S.: A flexible new technique for camera calibration. IEEE Trans. PAMI **22**(11), 1330–1334 (2000)