

Uncertainty-Encoded Augmented Reality for Robot-Assisted Partial Nephrectomy: A Phantom Study

Alborz Amir-Khalili¹, Masoud S. Nosrati², Jean-Marc Peyrat³,
Ghassan Hamarneh², and Rafeef Abugharbieh¹

¹ University of British Columbia, Vancouver, Canada

² Simon Fraser University, Burnaby, Canada

³ Qatar Robotic Surgery Centre, Qatar Science and Technology Park, Doha, Qatar

Abstract. In most robot-assisted surgical interventions, multimodal fusion of pre- and intra-operative data is highly valuable, affording the surgeon a more comprehensive understanding of the surgical scene observed through the stereo endoscopic camera. More specifically, in the case of partial nephrectomy, fusing pre-operative segmentations of kidney and tumor with the stereo endoscopic view can guide tumor localization and the identification of resection margins. However, the surgeons are often unable to reliably assess the levels of trust they can bestow on what is overlaid on the screen. In this paper, we present the proof-of-concept of an uncertainty-encoded augmented reality framework and novel visualizations of the uncertainties derived from the pre-operative CT segmentation onto the surgeon’s stereo endoscopic view. To verify its clinical potential, the proposed method is applied to an *ex vivo* lamb kidney. The results are contrasted to different visualization solutions based on crisp segmentation demonstrating that our method provides valuable additional information that can help the surgeon during the resection planning.

1 Introduction

The emergence of robot-assisted interventions using medical robots (e.g. da Vinci Surgical System, Intuitive Surgical, Inc., Sunnyvale, CA, USA), has been shown to increase the accuracy and reduce the operative trauma associated with complex interventions. In partial nephrectomies, for instance, a crucial step is tumor identification during which the surgeon localizes the kidney tumor mass and identifies the resection margins. This step is important to properly plan and speed up the succeeding stage of tumor mass excision during which blood flow can only be safely obstructed for a limited time. More importantly, the accuracy of this step is necessary not only to preserve kidney function by sparing as much healthy tissue as possible, but also to avoid tumor recurrence by resecting all cancerous tissue.

The tumor identification step is usually performed with the help of multi-modal source of information at the surgeon’s disposal: pre-operative scans (typically 3D CT and/or MR) and intra-operative data (2.5D stereo endoscopic data

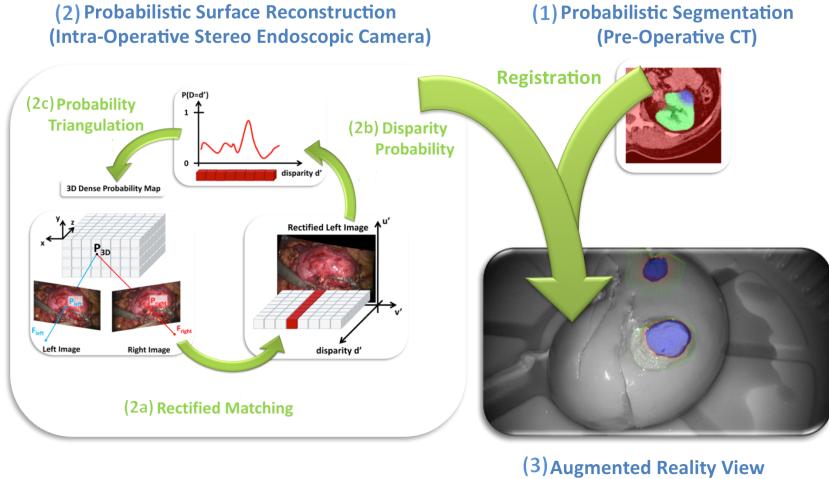


Fig. 1: Our uncertainty-encoded image-guidance framework consists of extracting 1) the probabilistic kidney/tumor boundaries from the CT volume prior to the operation and 2) the corresponding probabilistic surface information from the stereo endoscopic views intra-operatively using 2a) computational stereo matching techniques, 2b) converting matching weights into probability values, and 2c) triangulating the surface probabilities into the same domain as the CT. Finally, 3) we register the pre-operative boundary uncertainties to the stereo endoscope using probabilistic surface reconstruction information and visualize the isoprobability contours onto the surgeon's console.

and, when available, laparoscopic 2D/3D ultrasound). Currently, these rich and complementary sources of information are just displayed on the surgeon's console in a tiled fashion (i.e. side-by-side) or even sometimes on a separate screen of a workstation nearby. These typical display setups require substantial additional effort from the surgeon to piece together a 3D mental map of the surgical scene that integrates all information together in order to localize the tumor and adjacent tissue. Hence, an augmented reality view, in which the endoscopic video stream is overlaid with highlighted kidney and tumor boundaries, can substantially reduce the effort required by the surgeon to achieve accurate and quick tumor excision.

To the best of our knowledge, all current methods rely on the visualization of a crisp segmentation only [1]. This renders the surgeon highly susceptible to the varying levels of confidence in what is overlaid on the screen. Segmentations are hardly ever 100% accurate for many possible reasons: graded decomposition [2], image acquisition artifacts, inter-expert segmentation variability, and fuzzy image segmentation [3,4]. These uncertainties can be important in subsequent analyses and decision-making [2,5].

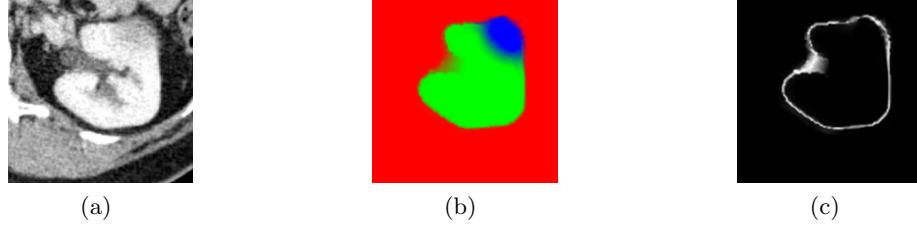


Fig. 2: Probabilistic pre-operative CT segmentation. (a) Original CT. (b) Membership probabilities of kidney (green), tumor (blue), and background (red). (c) Background boundary location probability (0 in black and 1 in white).

In this paper, we propose to provide a visualization of uncertainties at the kidney and tumor boundaries as a visual cue to assist the surgeon in finding the optimal resection strategy. This is similar in concept to what is currently being explored in radiotherapy for brain tumors when extrapolating glioma invasion with variable margins [6]. Our visual cues are derived from shape boundary uncertainties in the probabilistic segmentation of the pre-operative CT. This information is then registered to the endoscopic view as explained in Fig. 1. We apply our method to an *ex vivo* lamb kidney to create an uncertainty-encoded augmented reality view. We compare our results to standard guidance methods that use crisp segmentations and clearly demonstrate the benefits of our method and its utility for resection planning.

2 Methods

We first describe the probabilistic segmentation of the pre-operative CT that provides uncertainties about the boundary localization of kidney and tumor. Secondly, we perform a probabilistic 3D surface reconstruction from stereo endoscopy to which the probabilistic segmentation is directly registered.

2.1 Probabilistic Segmentation of Pre-Operative CT Scans

The probabilistic segmentation of the pre-operative CT is based on the random walker segmentation algorithm [4,7] that generates membership probabilities of three manually seeded regions: background (BG: red), kidney (KD: green), and tumor (TM: blue) (Fig. 2b).

We denote the resulting multi-label probabilistic CT segmentation by:

$$P_{seg}^{CT} : \Omega \subset \mathbb{R}^3 \rightarrow \mathbf{p} \in \mathcal{S}^2 ,$$

where $\mathbf{p} = [p_{BG}, p_{KD}, p_{TM}]$ belongs to the simplex of order 2, and Ω is the spatial domain of the CT. From this multi-label probabilistic segmentation, we

can extract the membership probability map of background P_{BG}^{CT} , kidney P_{KD}^{CT} and tumor P_{TM}^{CT} regions.

We also compute the likelihood $P_{surface}^{CT}$ of the *surface* union of kidney and tumor in the pre-operative CT (Fig. 2c) by combining the membership probabilities of being *inside* the kidney P_{KD}^{CT} and inside the tumor P_{TM}^{CT} as follows:

$$P_{surface}^{CT} = 1 - \frac{|(P_{KD}^{CT} + P_{TM}^{CT}) - 0.5|}{0.5}. \quad (1)$$

2.2 Probabilistic Stereo-Endoscopic Surface Reconstruction

We propose an extension of traditional computational stereo techniques of surface reconstruction from a single crisp surface [8] to a probabilistic representation of surfaces in 3-space.

Dense Matching of Left and Right Stereo Images Using polar rectification [9] with the camera calibration parameters, the 2D dense matching of left and right stereo images is simplified to a 1D matching along parallel epipolar lines in the left and right rectified images. We use the normalized cross correlation (NCC) ratio on greyscale images as a matching similarity metric. This metric has the advantage of being less prone to changes in illumination. In contrast with current state-of-the-art methods, e.g. [10,11,12], instead of computing one set of robust and optimal matches, we retain *all* possible matches with their associated disparity (displacement $d \in \mathbb{Z}$ between matching points along the same horizontal line of the rectified images) and similarity measure ($c \in [-1, 1]$).

Construction of a 3D Probabilistic Voxel Map In order to facilitate the pre-op to intra-op registration detailed in Section 2.3, we first create a 3D probabilistic voxel map in which each voxel stores the probability of being at the surface of the stereo endoscopic scene. To achieve this, we compute the disparity probability values by converting the NCC profile $\mathbf{c} = [c_1, c_2, \dots, c_{N_d}]$ computed previously at every pixel $(u, v) \in \Omega_{2D} \subset \mathbb{R}^2$ in one of the rectified images for different disparities $d \in \mathcal{D} = \{d_1, d_2, \dots, d_{N_d}\}$, where N_d is the total number of disparities. Basically, the NCC profiles are stacked into a 3D correlation map:

$$NCC_{3D}^{stereo} : (u, v, d_i) \in \Omega_{3D} \rightarrow c_i \in [-1, 1] \quad (2)$$

and converted into a 3D probabilistic voxel map using the Gibbs measure as follows:

$$P_{3D}^{stereo}(u, v, d_i) = \frac{\exp(-\beta(\max_d(NCC_{3D}^{stereo}(u, v, d)) - NCC_{3D}^{stereo}(u, v, d_i)))}{W(\beta)}, \quad (3)$$

where $W(\beta) = \sum_d \exp(-\beta(\max_d(NCC_{3D}^{stereo}(u, v, d)) - NCC_{3D}^{stereo}(u, v, d_i)))$ is the partition function, and β is a free parameter.

Finally, the 3D position of each matched pair of points in the stereo views is triangulated with the camera projection matrices to transform P_{3D}^{stereo} into a probabilistic voxel map $P_{surface}^{stereo}$ in real world 3D space:

$$P_{surface}^{stereo} : (x, y, z) \in \Omega_{3D} \rightarrow [p, 1-p] \in \mathcal{S}^1, \quad (4)$$

where $p \in [0, 1]$ is the likelihood of a surface at voxel (x, y, z) in real world 3D space.

2.3 Registration of Stereo Camera and CT

We initialize the registration of the CT to the stereo camera in a semi-automatic manner using manually matched landmarks between the original CT, left and right camera views. In this first step, we use a similarity transformation to model the combination of (1) a rigid transformation to cope with different reference frames between stereo camera and CT acquisitions and (2) a global scaling to cope with ambiguities resulting from possible camera calibration errors. The resulting transformation is then refined with an automatic similarity registration of $P_{surface}^{CT}$ to $P_{surface}^{stereo}$ obtained respectively from (1) and (4). Finally, a non-linear registration step of these two volumes with a B-Spline transformation model is performed to cope with deformations occurring between the pre-operative CT acquisition and the surgical scene. We used `elastix` [13] with the sum of squared differences (SSD) similarity metric for the two last automatic registration steps.

3 Results

3.1 Materials

For validation purposes, we fabricated an *ex vivo* phantom using a lamb kidney and implanted artificial tumors inside it. Different materials (chewing gum and olive pit) were used to emulate low and high contrast kidney-tumor boundaries within the CT. The chewing gum was placed on the surface of the kidney to emulate a partially exophytic tumor/cyst (Fig. 3a) and the olive pit was planted deep inside the kidney (close to the renal pelvis) representing a completely endophytic tumor (Fig. 3c).

A 16 slice Siemens Somatom CT scanner was used to acquire a high resolution CT volume of the phantom. The resulting volume is composed of 130 (0.600 mm thick) transverse slices of 512×512 pixels (0.215 mm pixel spacing). Stereo endoscopy data was captured with a calibrated da Vinci S system at full HD 1080i resolution.

3.2 *Ex vivo* Lamb Kidney Study

The Random Walker segmentation algorithm was applied with manual seeding of each label in the CT volume. The probabilistic labeling corresponding to the

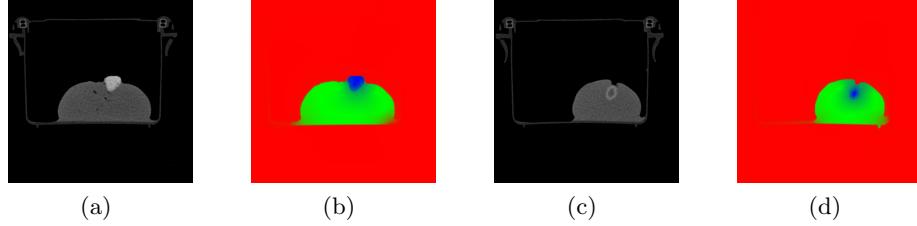


Fig. 3: Transverse slices of CT volume depicting our *ex vivo* lamb kidney phantom with (a) an exophytic and (c) an endophytic artificial tumor. (b) and (d) are probabilistic Random Walker segmentations of (a) and (c), respectively. Tumor labels are colored blue, kidney is colored green, and the background is red.

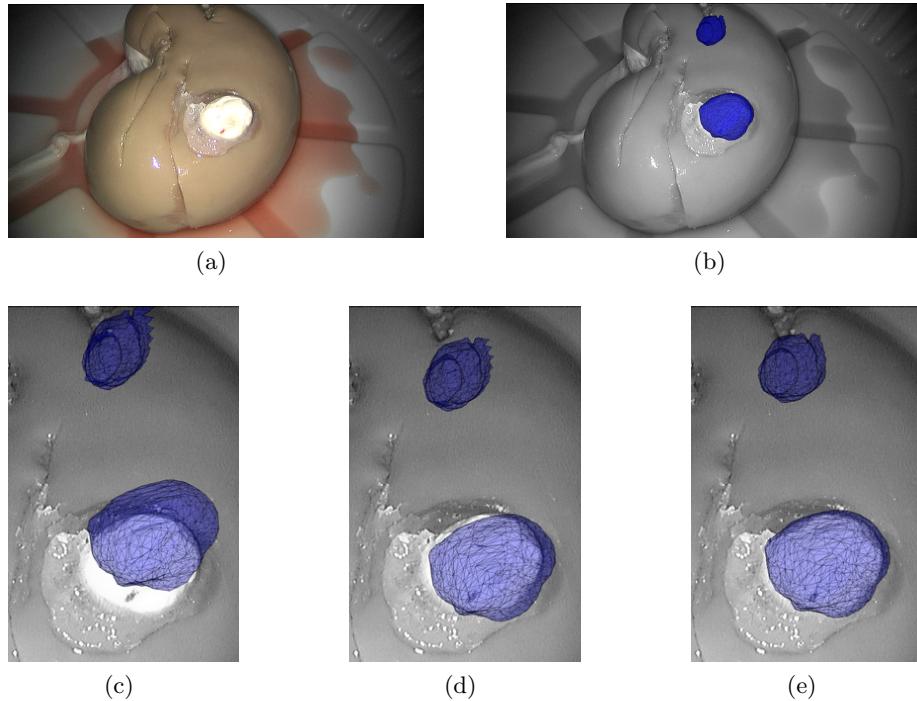


Fig. 4: Results of registration: (a) Original left stereo camera view, (b) final registered crisp mesh of the tumors (blue) projected on top of the image. Close-up views depicting intermediate results of the registration: (c) pose estimation, (d) automatic similarity transformation, and (e) non-rigid registration.

two simulated tumors is illustrated in Fig. 3b and 3d. Note that the diffusion of uncertainty in the endophytic case is more visible compared to the exophytic

tumor; this is a direct result of weaker contrast (CT intensity values: difference in pit/gum composition) at the kidney-tumor boundary. We were careful to keep the distances between the manually placed seeds and the visible boundaries constant to decrease the influence of seed placement on the resulting segmentations.

As illustrated in Fig. 4a, our phantom is quite smooth and lacks unique features on its surface. This results in a largely uncertain reconstruction from our stereo matching algorithm, which in turn causes the registration to be sensitive to the initial pose estimation. Successful registration was achieved after estimating the pose (Fig. 4c) using only four manually selected corresponding landmarks. The outcome of the registration was verified visually (Fig. 4) by projecting the kidney and tumor surfaces on both left and right endoscopy views. A small error in alignment (< 1 mm) is observed in the resulting registration, this is due to the error in reconstruction which is attributed to lack of texture on the phantom.

In order to verify the usefulness of probabilistic boundary visualization, we present four visualization scenarios. In the **first case** (Fig. 4b), we generate a crisp mesh model of the tumor by thresholding the probabilistic segmented CT volume to extract the most probable kidney-tumor boundary. In our **second case**, we project the previously generated mesh onto a 2D plane (normal to the camera) and extract its contour (Fig. 5a). This particular approach does not provide the surgeon with much additional information. Without any visible information (e.g. in the endophytic case) the surgeon's confidence regarding the visualized crisp boundary is, at best, changing isotropically away from the contour (as emulated in Fig. 5b). **Third case**, we isotropically dilate the 3D thresholded volume of the tumors by 1 mm increments and overlay the corresponding projected 2D contours (Fig. 5c). The resulting 2D contours dilate anisotropically as they are influenced by the orientation and shape of the tumor in 3-space. **Fourth case**, we propose thresholding the probabilistic volume at increasingly conservative confidence intervals instead of isotropic dilation to obtain isoprobability contours (Fig. 5d). In this case, we are essentially guiding the dilation of resection boundaries using the underlying uncertainty information extracted during the probabilistic segmentation of the CT. These results are consistent with our initial observation that the diffusion of uncertainties are greater in the endophytic case (pit/gum difference).

We presented the four cases to expert urology surgeons. The general consensus was that the information presented in the fourth case (Fig. 5d) is promising. A valid critique was made regarding the number of contours being overlaid on the endoscopy view: it obstructs the kidney more than the simple crisp solution (Fig. 5a). In order to address this problem, we present a complimentary visualization scenario in which uncertainties are projected onto a single crisp contour. We accomplish this by computing the minimum distance between the most probable contour and the most conservative one at every location of the most probable contour (distance from inner-most to outer-most contours in Fig. 5d). A lower distance implies a higher confidence in the boundary localization as it indicates a sharper edge in the probability map. We then transform these distances into a relative color map and use it to color-code the crisp contour (Fig. 6).

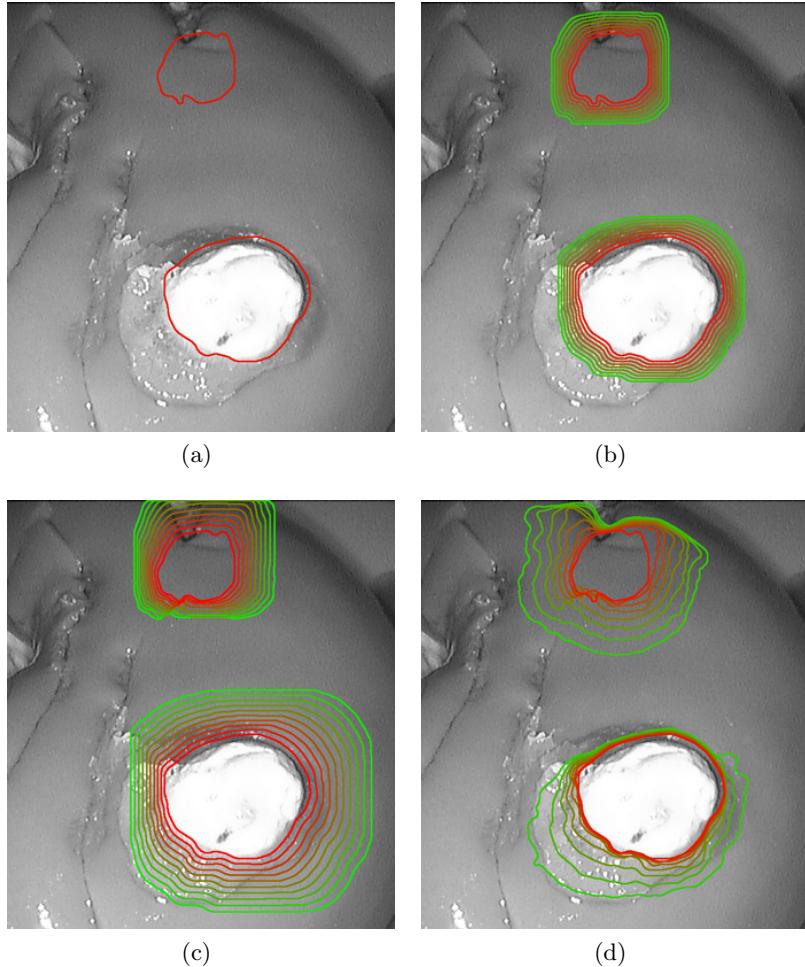


Fig. 5: Augmented reality view of *Ex vivo* lamb kidney with endophytic and exophytic artificial tumors showing different visualization scenarios: (a) crisp contour of projected mesh, (b) isotropic 2D diffusion of the crisp contour, (c) 2D projections of the crisp mesh dilated in 3D by 1 mm increments, (d) 2D projections of 3D isoprobabilities from 0.5 to 0.15. Contours range from the most probable boundary (red) to the most conservative boundary (green).

This final visualization scenario does not only provide the most probable tumor boundary localization, but also provide information about its local confidence. This visualization can guide the surgeon to quickly identify the best (most confident) place to start the resection. During the resection, the surgeon can always opt for the fourth case to see exactly how the uncertainty is diffused spatially.

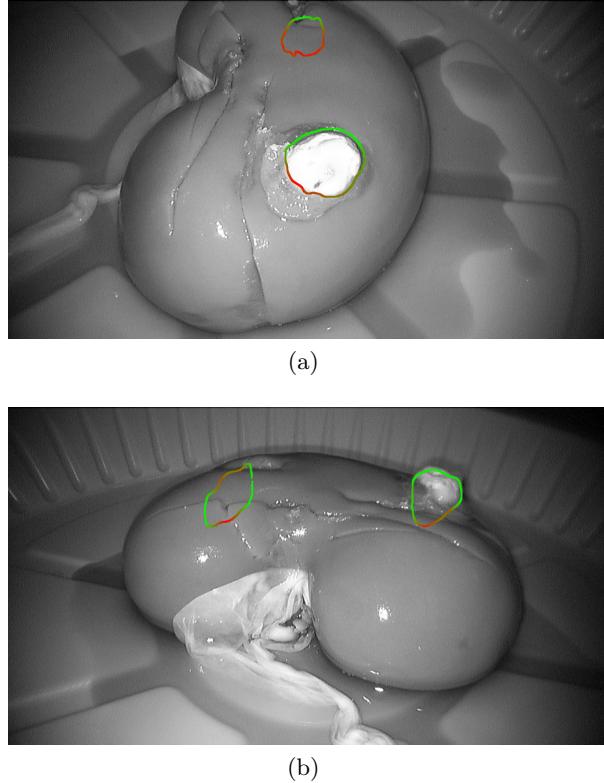


Fig. 6: (a) Top and (b) side views of *ex vivo* lamb kidney augmented with uncertainty-driven tumor boundary localization. Uncertainty is encoded into the tumor boundary ranging from certain (green) to uncertain (red).

4 Conclusion

We proposed a framework that enables extraction and registration of probabilistic data from two complimentary sources of information available in robot-assisted surgical interventions. Our approach provides the confidence in the resulting augmented information which can help the surgeon during the localization of excision margins before resection.

The novel visualization we presented is a proof of concept. The next step is to validate our experiments on clinical data and more realistic *ex vivo* phantoms with agar-based tumors of varying intensities, shapes and sizes [14]. We plan to conduct in-depth summative usability tests in addition to more formative usability tests to fully validate the integration of our uncertainty encoded visualization techniques into the clinical workflow. In the near future we aim to automate the initialization (pose estimation) steps and facilitate real-time operation of this framework. Although in this paper we presented uncertainty encoding from pre-

operative CT, we will be taking advantage of other intra-operative sources of uncertainty to improve the confidence at the localized boundary while new data is acquired during the resection.

Acknowledgement The authors would like to thank Dr. Abdulla Al Ansari and Dr. Osama Al-Alao for their expert opinion and assistance with the data acquisition. This publication was made possible by NPRP Grant from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

1. P. Pratt, E. Mayer, J. Vale, D. Cohen, E. Edwards, A. Darzi, and G.-Z. Yang, “An effective visualisation and registration system for image-guided robotic partial nephrectomy,” *Journal of Robotic Surgery*, pp. 1–9, 2012.
2. J. Udupa and G. Grevera, “Go digital, go fuzzy,” *Pattern Recognition Letters*, vol. 23, no. 6, pp. 743–754, 2002.
3. Y. Zhang, M. Brady, and S. Smith, “Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm,” *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, 2001.
4. L. Grady, “Random walks for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1768–1783, 2006.
5. S. Warfield, K. Zou, and W. Wells, “Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation,” *IEEE Transactions on Medical Imaging*, vol. 23, no. 7, pp. 903–921, 2004.
6. E. Konukoglu, O. Clatz, P. Bondiau, H. Delingette, and N. Ayache, “Extrapolating glioma invasion margin in brain magnetic resonance images: suggesting new irradiation margins,” *Medical Image Analysis*, vol. 14, no. 2, pp. 111–125, 2010.
7. A. K. Sinop and L. Grady, “A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm,” in *ICCV*, pp. 1–8, 2007.
8. R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge Univ Press, 2nd ed., 2004.
9. M. Pollefeys, R. Koch, and L. Van Gool, “A simple and efficient rectification method for general motion,” in *ICCV*, pp. 496–501, 1999.
10. D. Stoyanov, M. V. Scarzanella, P. Pratt, and G.-Z. Yang, “Real-time stereo reconstruction in robotically assisted minimally invasive surgery,” in *MICCAI*, pp. 275–282, 2010.
11. S. Bernhardt, J. Abi-Nahed, and R. Abugharbieh, “Robust dense endoscopic stereo reconstruction for minimally invasive surgery,” *MICCAI workshop on Medical Computer Vision (MCV)*, pp. 198–207, 2012.
12. S. Röhl, S. Bodenstedt, S. Suwelack, H. Kenngott, B. Müller-Stich, R. Dillmann, and S. Speidel, “Real-time surface reconstruction from stereo endoscopic images for intraoperative registration,” in *Proc. SPIE*, vol. 7964, p. 796414, 2011.
13. S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim, “Elastix: a toolbox for intensity based medical image registration,” *IEEE Transactions on Medical Imaging*, vol. 29, pp. 196–205, 2010.
14. J. S. Huber, Q. Peng, and W. W. Moses, “Multi-modality phantom development,” *IEEE Transactions on Nuclear Science*, vol. 56, no. 5, pp. 2722–2727, 2009.