Auto Localization and Segmentation of Occluded Vessels in Robot-Assisted Partial Nephrectomy

Alborz Amir-Khalili[1], Jean-Marc Peyrat[2], Julien Abinahed[2], Osama Al-Alao[3], Abdulla Al-Ansari[2,3], Ghassan Hamarneh[4], and Rafeef Abugharbieh[1]

[1] BiSICL, University of British Columbia, Vancouver, Canada
[2] Qatar Robotic Surgery Centre, Qatar Science & Technology Park, Doha, Qatar
[3] Urology Department, Hamad General Hospital, Hamad Medical Corporation, Qatar
[4] Medical Image Analysis Lab, Simon Fraser University, Burnaby, Canada

Abstract. Hilar dissection is an important and delicate stage in partial nephrectomy during which surgeons remove connective tissue surrounding renal vasculature. Potentially serious complications arise when vessels occluded by fat are missed in the endoscopic view and are not appropriately clamped. To aid in vessel discovery, we propose an automatic method to localize and label occluded vasculature. Our segmentation technique is adapted from phase-based video magnification, in which we measure subtle motion from periodic changes in local phase information albeit for labeling rather than magnification. We measure local phase through spatial decomposition of each frame of the endoscopic video using complex wavelet pairs. We then assign segmentation labels based on identifying responses of regions exhibiting temporal local phase changes matching the heart rate frequency. Our method is evaluated with a retrospective study of eight real robot-assisted partial nephrectomies demonstrating utility for surgical guidance that could potentially reduce operation times and complication rates.

1 Introduction

Approximately 30,000 new cases of kidney cancer, generally renal cell carcinoma, are detected each year in the U.S. alone. Kidney resection, also known as a nephrectomy, is the only known effective treatment for this type of localized cancer [1]. Robot-assisted partial nephrectomy (RAPN) refers to nephron-sparing techniques performed with surgical robots in which only the cancerous cells are excised while the kidney is reconstructed to retain functionality.

The intraoperative aspect of RAPN procedures can be organized into five stages [2]: 1) Bowel mobilization; 2) Hilar dissection and control; 3) Identification and demarcation of tumor margins; 4) Resection of tumor; and 5) Reconstruction of the kidney (renorrhaphy). Hilar dissection stands out as a daunting stage that requires significant expertise since improper clamping due to overlooked accessory renal vessels can cause significant bleeding during resection [3].

Hilar dissection is a delicate procedure during which the surgeon dissects through the Gerota's fascia and removes the connective tissue that surrounds

the renal artery (RA) and vein (RV). This task is complex due to the variability in patient vasculature and the amount of perinephric fat which surrounds the kidney. Access to the hilum grants the surgeon control over the flow of blood into and out of the kidney, which is very important as warm ischemia is required during the excision of the tumor to minimize internal hemorrhaging.

In some cases, accessory vessels that branch off from the RA or the abdominal aorta are accidentally missed as they lie hidden behind a thick layer of perinephric fat. In one study of 200 laparoscopic partial nephrectomy cases by world leading surgeons, seven incidents of intraoperative bleeding were reported as a result of inadequate hilar control, two of which were directly caused by missed accessory vessels [4]. Although the number of incidents is low, other studies have observed the existence of accessory vessels in up to 35% of patients [5,6]. If the surgeon's level of experience is limited, the incidence of bleeding may be much higher. The implications are many, aside from obvious complications that would arise from internal hemorrhaging, bleeding may also jeopardize the surgical outcome by occluding the surgeon's view as the tumor is being resected.

Surgeons often make use of preoperative medical images in identifying troublesome accessory vessels [7]. Even with a detailed scan and segmented pre-op plan, surgeons are still burdened with the task of mentally transferring this abstraction onto the surgical site during the operation. Reducing the difficulty of navigation has been attempted by state-of-the-art approaches that used multimodal registration to align the preoperative surgical map of the vessels onto the surgeon's endoscopic view [8]. Such techniques have to date not been extensively validated in clinical practice, possibly because they require very delicate selection of parameters, use of invasive fiducials, or are computationally complex to the extent that the algorithms cannot perform in real-time. Recent methods favor the use of hardware solutions such as near infrared flourescence imaging [9] or algorithmic methods that only use color intensity information from the endoscope for enhancing RAPN by highlighting vasculature based on perfusion models [10]. Hardware solutions are not widely accessible as they are cost restrictive and both methods fail to identify vessels that are hidden under a layer of fat.

Our work is motivated by the need for an automated guidance system that can work in parallel to the techniques mentioned above to reduce the complications and the time required to perform hilar dissection by assisting the surgeon in localizing hidden accessory vessels. Our proposed system aims at highlighting occluded vessels by analyzing the complementary temporal motion characteristics of the scene as acquired by the endoscope. Our method is inspired by video magnification techniques developed for natural scenes [11,12], where an Eulerian approach to analyzing flow within a video sequence can be used to magnify periodic motions that are nearly invisible to the human eye. An extension of [11] was very recently implemented in the context of robot-assisted surgery [13]. In our case, we adapted the phased-based video magnification [12] to detect and label subtle motion patterns instead of magnifying them. Our method is evaluated with a retrospective study of eight RAPN cases to show its potential utility for surgeons.

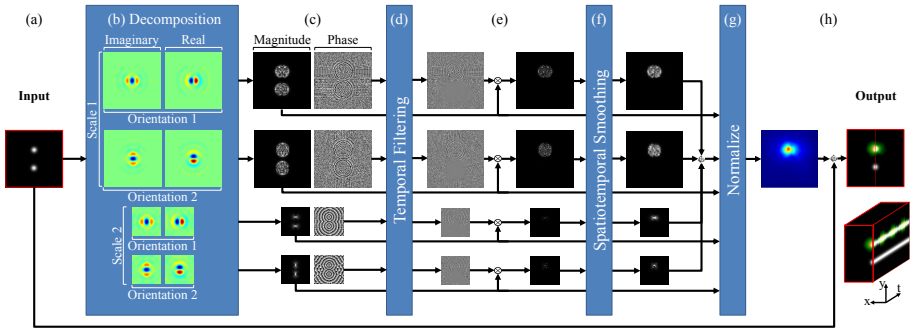

Fig. 1: Overview of our proposed method: (a) First frame of a synthetic input video composed of one circle that pulsates (top) and another that remains stationary (bottom). (b) Steerable filter bank with illustrated impulse responses are used to decompose the information inside each frame into (c) magnitude and local phase at different scales and orientations. (d) The phase information of all frames of the video is temporally filtered using an ideal bandpass filter centered on the frequency of the pulsating circle. (f) A spatiotemporal median filter is applied to (e) the magnitude weighted response of the filtered phases to remove phase noise. (g) The filtered results are then combined and (h) added back to the input as an overlay. A spatiotemporal cross section of the video illustrates four pulsations across 30 frames of the synthetic video.

2 Methods

The goal is to highlight occluded vasculature near the renal hilum. These regions in the video exhibit periodic pulsatile motion within a narrow temporal passband centered around the heart rate. This sub-pixel motion is faintly visible on the surface of the occluding adipose tissue. We can relate the position p of such a region at time t to its original position P, at time $t = 0$, with the displacement vector $u(P, t)$ such that $p = u(P, t) + P$. By analyzing the vector u with respect to time, for a dense set of regions starting at every pixels in the first frame of acquisition, we can label regions that exhibit the desired behaviour.

Dense tracking of pixels in an image through time is computationally expensive, more so when the motions occur at the subpixel level [14]. To overcome this limitation, our proposed segmentation method (Fig. 1) relies on Eulerian motion processing techniques to label regions that pulsate like a vessel.

The Eulerian motion magnification work [11] relies on the first-order approximation of changes in intensities to estimate motion. Analysis in [11], demonstrated that this approximation is susceptible to noise at high spatial frequencies, especially at spatial locations where the curvature of change in intensity is high. A recent study demonstrated that a second order approximation of the change in intensity using the Hessian matrix is less susceptible to errors and provides a metric for attenuating noise locally [13]. Even with a second order approximation, these gradient based methods are prone to error in salient regions. We have chosen to estimate this local motion from the change in instantaneous phase

of complex sinusoid decomposition of each image using the more recent phase-based motion processing [12] technique as it has been proven to be more robust to high frequency noise and, by extension, non-Lambertian specular highlights that are abundant in endoscopic video.

Our extension to their method can be described in the same 1D intuitive manner proposed by *Wadhwa et al* [12], without loss of generality, as follows. Given a frame of video (Fig. 1a), represented as an intensity function $f(p)$ that maps an intensity value to a given particle at position p, is decomposed into spatial sub-bands (Fig. 1b)

$$f(p) = f(P + u(P,t)) = \sum_{\omega=-\infty}^{\infty} A_\omega e^{i\omega(P+u(P,t))} \tag{1}$$

with each sub-band representing a complex sinusoid $S_\omega(p,t) = A_\omega e^{i\omega(P+u(P,t))}$ at spatial frequency ω, the local phase (Fig. 1c) is defined as $\omega(P + u(P,t))$.

The motion vectors $u(P,t)$ can be extracted from a temporal sequence of local phase measurements using a DC balanced bandpass filter (Fig. 1d); the filter response of the temporal bandpass filter is denoted by B_ω. If the passband is wide enough, we can compute the displacement vector entirely at each sub-band such that $B_\omega(p,t) = \omega u(P,t)$. If the passband of the filter is tuned to the typical heart rate of a patient, we can isolate components of the local motion that are synchronous with the heart rate and vascular pulsation.

We adapt our method to generate fuzzy segmentation labels from the computed local motion. Local phase changes, B_ω, are first attenuated in regions where the magnitude response A_ω of the spatial sub-band is weak by computing the product between the bandpassed phases and the normalized magnitude of the spatial filter response vectors to obtain $Q_\omega = \hat{A}_\omega B_\omega$ (Fig. 1e). Local phase measurements are wrapped between the interval $(-\pi, \pi]$ so Q_ω will contain impulse noise. We remove noise from the product Q_ω using a spatiotemporal median filter (Fig. 1f). This denoised product $\tilde{Q}_\omega$ is finally incorporated into a voting scheme (Fig. 1g) using magnitude weighted averaging to obtain fuzzy labels

$$H(p,t) = \frac{\sum_{\omega=-\infty}^{\infty} \lfloor \frac{A_\omega \tilde{Q}_\omega}{2\pi\omega} - T \rfloor}{\sum_{\omega=-\infty}^{\infty} A_\omega + \epsilon}, \tag{2}$$

where T is an optional noise compensation term representing the maximum response that can be generated from noise alone and ϵ is a small number to avoid division by zero. The resulting sequence of fuzzy labels H may be displayed as an overlay or separately to highlight this pulsatile motion (Fig. 1h).

3 Results

Video sequences from eight real RAPN interventions were used for validation. All endoscopic video data were acquired by a da Vinci Si surgical system (Intuitive Surgical, California, USA). HD (1080i) videos were resized to 480×270 pixels.

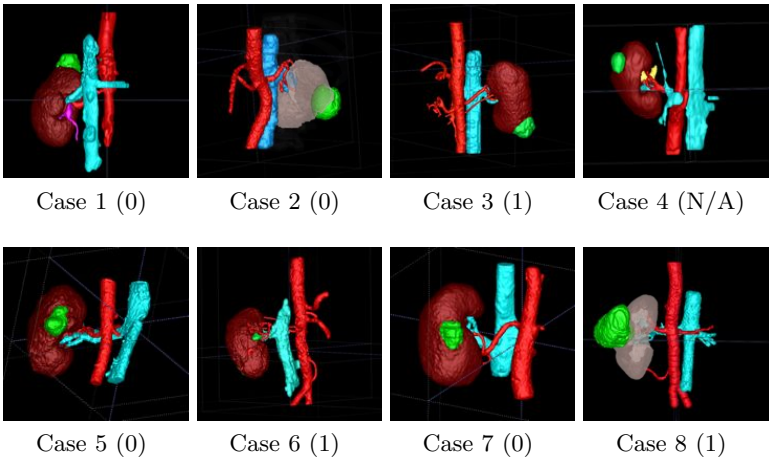

Fig. 2: Manual segmentation of each RAPN case, number of accessory vessels in parenthesis, showing kidney (brown), tumor/cyst (green), veins (cyan), and arteries (red).

Sixteen complex steerable filter pairs were used (four orientations at four scales) with one filter per octave using publicly available code [15]. The passband of the temporal filter was set between 60 to 120 beats per minute and the noise threshold T was set to zero. Average runtime of our unoptimized MATLAB code to process a four second clip (120 frames) was 65 seconds.

In order to provide a framework for validation, we compare the segmentations obtained through our guidance system against manually localized vasculature. To achieve this, we segmented the kidney, tumor/cyst, inferior vena cava, abdominal aorta, RA, RV, and accessory vessels (Fig. 2) using a semi-automatic segmentation algorithm [16]. The resulting meshes were manually aligned onto the first frame of each endoscopic scene (Fig. 3a) using a rigid transformation. Anatomical landmarks such as the contour of the kidney or visible parts of the vessels were used to guide the registration process. Our best estimate of the ground truth is presented in Fig. 3b. Small observable discrepancies between the aligned model and the endoscopic view are attributed to non-rigid deformations of the organs and vasculature caused by deformation during insufflation, retraction, or the mobilization of organs during the dissection.

In our experiments, it was observed that although venous and arterial structures pulsate at the same frequency, their pulsations are not in-phase. The motion of the inferior vena cava and RV typically succeeds that of the RA and abdominal aorta by an average of six frames. The results in Fig. 3 illustrate this phenomenon in two frames of the segmented video. Compared to the reference in Fig. 3b, the motions highlighted in Fig. 3c correspond to the cyan structures (venous) and Fig. 3d corresponds to the red structures (arterial).

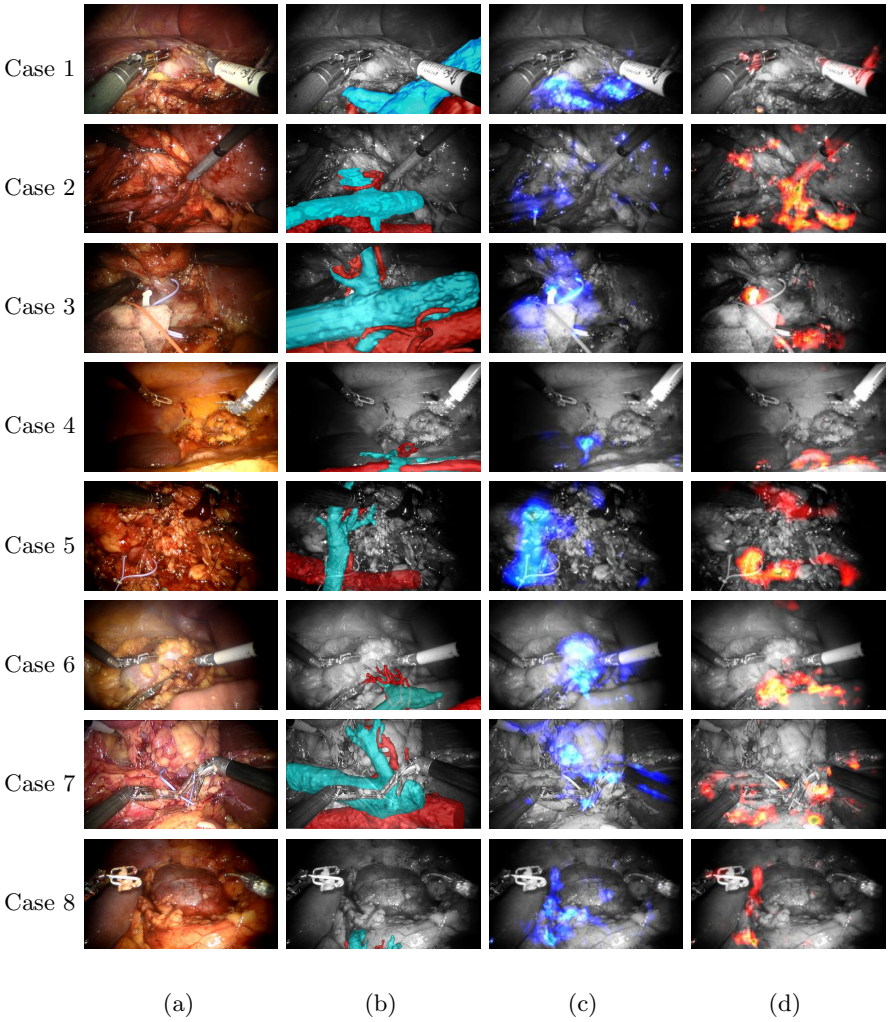

Fig. 3: Exemplar video frames with the proposed automatic localization of veins and arteries in the scene. (a) The first frame of the sequence, (b) manually localized venous (cyan) and arterial (red) structures, and segmentation of (c) veins and (d) arteries.

From the results we can further observe that, in Case 1, the small RA is correctly identified at the hilum. The mislabeling of the RA in Case 2 is attributed to retraction by the surgical instrument. The small accessory RA to the left of RV is also identified in Case 3. In Case 4, the suprarenal vein is misaligned due to mobilization of the spleen. Case 5 illustrates that retraction has shifted the abdominal aorta up. Branching of RA is detected on both sides of RV in Case 6 and the pulsation of heavy vascular region has casued the tumor to pulsate in

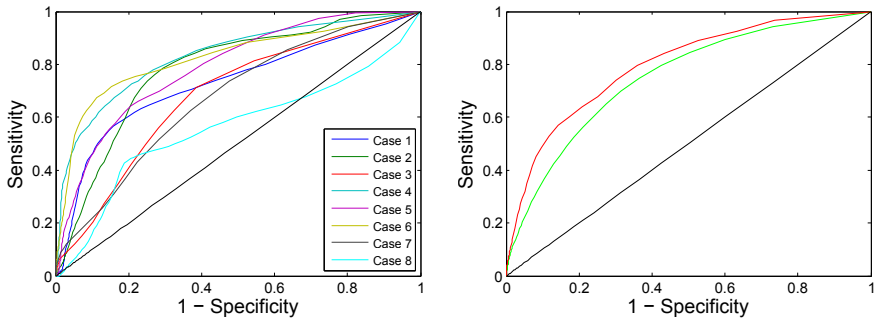

Fig. 4: Left: ROC of all cases. Right: median (red) and mean (green) of all ROC. Mean area under all ROC is 0.76 with a standard deviation of 0.08.

the centre of the frame. False positives are observed in Case 7 due to the motion of the tools in the scene. Ideally, surgical instruments should remain motionless during the acquisition of the video. The last case posed a big challenge as the vasculature is heavily occluded by the bowel and many false positives are detected in fluid filled cavities to the left of the cyst.

To validate our labels quantitatively, the segmentations were binarized (at a fixed threshold throughout the sequence) and combined across the frames of the video. This resulting binarized image was then compared to the binarized version of the reference manual segmentation in Fig. 3b, combining all vasculature into a single mask. Fig. 4 illustrates the segmentation performance of all cases, at different threshold values, via their receiver operating characteristics (ROC).

4 Conclusion

We have described a novel method for localizing and labelling regions in endoscopic video that contain occluded vessels. Our method extends Eulerian phase-based video motion processing techniques to detect and label small motions that are barely visible on the surface of the perinephric fat. To the best of our knowledge, we are the first to attempt the challenging task of localizing occluded vasculature in endoscopic video without the use of additional hardware or preoretative scans. We validated our novel method qualitatively in a retrospective *in vivo* study to verify its application in a clinical setting. Conservative quantitative validation of our method demonstrates that it is suitable for integration alongside existing techniques (as an additional cue) that use other visible features such as color, shape and texture. In the future, we plan to extend this method by developing an adaptive estimation for noise and optimizing the code to operate in real-time. We are actively exploring the use of obtained segmentations as an additional data term for guiding automatic non-rigid registration of preoperative surgical models with endoscopic video in the context of RAPN. Finally, we aim to assess its applicability to other surgical interventions.

Acknowledgement The authors would like to thank Mr. Masoud Nosrati, Mr. Jeremy Kawahara, and Mr. Ivan Figueroa for their assistance with the data acquisition. This publication was made possible by NPRP Grant #4-161-2-056 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. Drucker, B.J.: Renal cell carcinoma: current status and future prospects. Cancer treatment reviews **31**(7) (2005) 536–545
2. Gill, I.S., et al.: Laparoscopic partial nephrectomy for renal tumor: Duplicating open surgical techniques. The Journal of Urology **167**(2, Part 1) (2002) 469 – 476
3. Singh, I.: Robot-assisted laparoscopic partial nephrectomy: Current review of the technique and literature. Journal of minimal access surgery **5**(4) (2009) 87
4. Ramani, A.P., Desai, M.M., Steinberg, A.P., Ng, C.S., Abreu, S.C., Kaouk, J.H., Finelli, A., Novick, A.C., Gill, I.S.: Complications of laparoscopic partial nephrectomy in 200 cases. The Journal of urology **173**(1) (2005) 42–47
5. Urban, B.A., Ratner, L.E., Fishman, E.K.: Three-dimensional volume-rendered CT angiography of the renal arteries and veins: Normal anatomy, variants, and clinical applications. RadioGraphics **21**(2) (2001) 373–386
6. Sampaio, F., Passos, M.: Renal arteries: anatomic study for surgical and radiological practice. Surgical and Radiologic Anatomy **14**(2) (1992) 113–117
7. Mottrie, A., De Naeyer, G., Schatteman, P., et al.: Impact of the learning curve on perioperative outcomes in patients who underwent robotic partial nephrectomy for parenchymal renal tumours. European Urology **58**(1) (2010) 127–133
8. Teber, D., Guven, S., Simpfendörfer, T., Baumhauer, M., Güven, E.O., Yencilek, F., Gözen, A.S., Rassweiler, J.: Augmented reality: a new tool to improve surgical accuracy during laparoscopic partial nephrectomy? Preliminary in vitro and in vivo results. European Urology **56**(2) (2009) 332–338
9. Tobis, S., Knopf, J., Silvers, C., Yao, J., et al.: Near infrared fluorescence imaging with robotic assisted laparoscopic partial nephrectomy: initial clinical experience for renal cortical tumors. The Journal of Urology **186**(1) (2011) 47–52
10. Crane, N.J., Gillern, S.M., Tajkarimi, K., Levin, I.W., Pinto, P.A., et al.: Visual enhancement of laparoscopic partial nephrectomy with 3-charge coupled device camera: assessing intraoperative tissue perfusion and vascular anatomy by visible hemoglobin spectral response. The Journal of Urology **184**(4) (2010) 1279–1285
11. Wu, H.Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., Freeman, W.T.: Eulerian video magnification for revealing subtle changes in the world. ACM Transactions on Graphics **31**(4) (2012) 65
12. Wadhwa, N., Rubinstein, M., Durand, F., Freeman, W.T.: Phase-based video motion processing. ACM Transactions on Graphics **32**(4) (2013) 80
13. McLeod, A.J., Baxter, J.S., de Ribaupierre, S., Peters, T.M.: Motion magnification for endoscopic surgery. In: SPIE: Medical Imaging. Volume 9036. (2014) 9036–11
14. Liu, C.: Beyond pixels: exploring new representations and applications for motion analysis. PhD thesis, Massachusetts Institute of Technology (2009)
15. Portilla, J., Simoncelli, E.P.: A parametric texture model based on joint statistics of complex wavelet coefficients. IJCV **40**(1) (2000) 49–70
16. Yushkevich, P.A., Piven, J., Cody Hazlett, H., Gimpel Smith, R., Ho, S., Gee, J.C., Gerig, G.: User-guided 3D active contour segmentation of anatomical structures. Neuroimage **31**(3) (2006) 1116–1128