

# Robotically Surgical Vessel Localization Using Robust Hybrid Video Motion Magnification

Wenkang Fan<sup>✉</sup>, Zhuohui Zheng, Wankang Zeng, Yinran Chen, Hui-Qing Zeng, Hong Shi,  
and Xiongbiao Luo<sup>✉</sup>, *Senior Member, IEEE*

**Abstract**—Vessel and neurovascular bundle localization plays an essential role in endoscopic and robotic surgery. It still remains challenging to spare vessels and neurovascular bundles to avoid inadvertent injury due to limited visual and tactile perception of surgeons. This work assumes that surgeons have great difficulty in intuitively perceiving small pulsatile motion of vessels and neurovascular bundles from complex surgical field provided by endoscopic videos, and proposes a new surgical video pulsatile motion magnification method to help surgeons easily and precisely recognize vessels or neurovascular bundles by their visual systems. The new method consists of robust hybrid temporal filtering and deeply learned spatial decomposition. The proposed hybrid temporal filtering can significantly magnify pulsatile motion more consistent with reality and simultaneously keep non-pulsating regions in magnified videos almost identical to original videos, and learning-based spatial decomposition can reduce noise and ring artifacts in magnified videos. We evaluate our method on surgical videos acquired from robotic prostatectomy, with the experimental results showing that our method essentially outperforms current motion magnification approaches. In particular, visual quality and quantitative assessment of our method are certainly better than these methods.

**Index Terms**—Surgical robotics; laparoscopy, computer vision for medical robotics, vessel localization, motion magnification, hybrid temporal filtering.

## I. INTRODUCTION

ENDOSCOPIC and robotic surgery are widely performed in the operating room because of their fast recovery and

Manuscript received October 15, 2020; accepted January 25, 2021. Date of publication February 12, 2021; date of current version March 2, 2021. This letter was recommended for publication by Associate Editor L. Fichera and Editor P. Valdastri upon evaluation of the reviewers' comments. This work was supported in part by the National Nature Science Foundation of China under Grant 61971367, in part by the Fujian Provincial Technology Innovation Joint Funds under Grant 2019Y9091, in part by the Fujian Provincial Natural Science Foundation under Grant 2020J01004, in part by the National Natural Science Foundation of China under Grant 62001403, in part by the Fundamental Research Funds for the Central Universities China under Grant 20720200093, and in part by the Natural Science Foundation of Fujian Province of China under Grant 2020J05003. (*Corresponding author: Wenkang Fan*)

Wenkang Fan, Zhuohui Zheng, Wankang Zeng, Yinran Chen, and Xiongbiao Luo are with the Department of Computer Science, Xiamen University, Xiamen 361005, China (e-mail: 23020181154270@stu.xmu.edu.cn; 23020181154261@stu.xmu.edu.cn; 23020181154252@stu.xmu.edu.cn; yinran\_chen@xmu.edu.cn; xiongbiao.luo@gmail.com).

Hui-Qing Zeng is with the Zhongshan Hospital, Xiamen University, Xiamen 361004, China (e-mail: luowan@outlook.com).

Hong Shi is with the Fujian Cancer Hospital & Fujian Medical University Cancer Hospital, Xiamen, Fujian 361005, China (e-mail: endoshihong@hotmail.com).

This letter has supplementary downloadable material available at <https://doi.org/10.1109/LRA.2021.3058906>, provided by the authors.

Digital Object Identifier 10.1109/LRA.2021.3058906

less complications. Endoscopes integrated with video cameras are usually used to provide surgeons with surgical visualization to manipulate various surgical tools. Sparing vessels and neurovascular bundles is essential to reduce surgical risks during various surgical procedures. Unfortunately, it is difficult to intuitively identify on-site vessels and neurovascular bundles from endoscopic videos due to limited visual and tactile perception of surgeons.

Numerous vessel and neurovascular bundle identification and localization methods are generally classified into of three main categories: (1) preoperative imaging, (2) interventional imaging, and (3) endoscopic imaging. While some preoperative imaging methods [1] and interventional imaging methods such as ultrasound [2], [3] and florescent [4], [5] can locate vessels or bundles, they are inconvenient and inaccurate in clinical practices [6]. More recently, video motion magnification as endoscopic imaging method is interesting and effective [7], [8] that can enhance small pulsatile motions to recognize pulsating arteries and veins [9]. This method directly magnifies small motion in endoscopic video to enable surgeons to easily and conveniently locate vessels with naked eyes without additional configurations in the operating room.

## A. Related Work

Currently, Eulerian approach has become the main method of video motion magnification which contain three main steps: spatial decomposition, temporal filtering, and motion reconstruction. Wu *et al.* [10] used the Laplacian pyramid to decompose video frames and directly magnify intensity-based motion, while Wadhwa *et al.* [11], [12] employed the Riesz pyramid to decompose video frames and obtain phase-based motion. More interestingly, Oh *et al.* [13] decomposed video frames using convolutional neural networks to learn motion representation that can greatly reduce noise and ring artifacts. These Eulerian approaches [10]–[13] use bandpass filtering to extract frequency components of interest. Unfortunately, bandpass filtering results in losing original surgical video characteristics since it only preserves sine-wave components. To address this problem, Janatka *et al.* [14] introduced third-order Gaussian filtering into accelerated magnification [15]. Such a filter can highlight cardio-physiological features and maintain the original pulsation characteristics but it will introduce additional noise and artifacts.

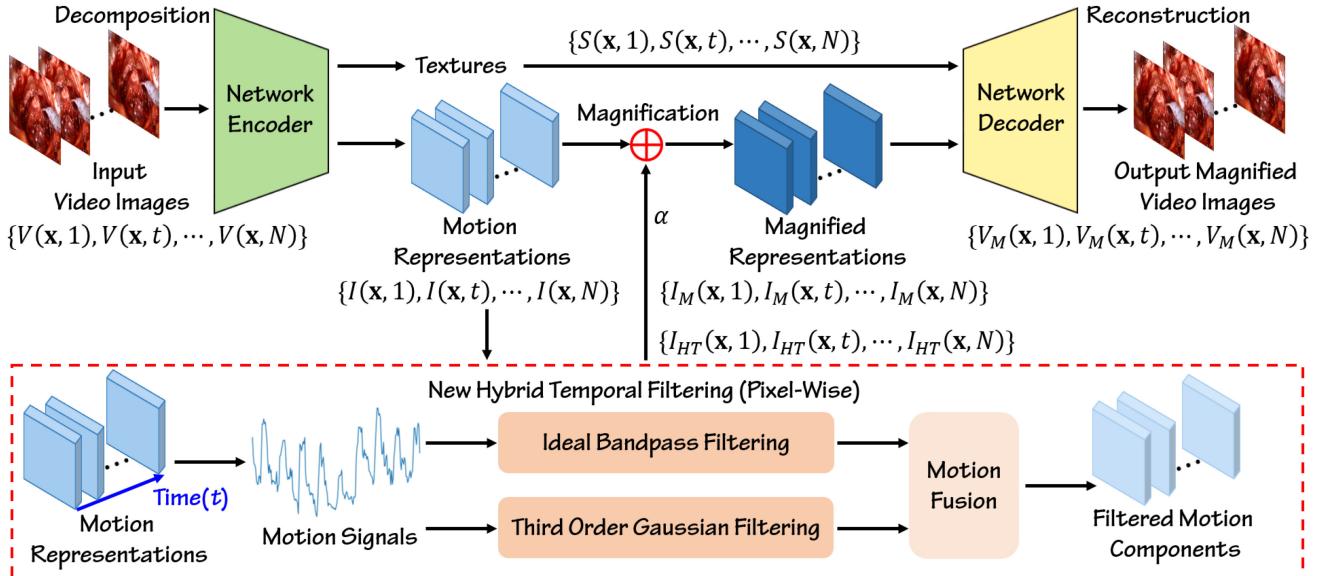


Fig. 1. New hybrid temporal filtering with deeply learned decomposition and reconstruction for pulsatile motion magnification.

### B. Contribution

This work proposes a new motion magnification strategy that employs hybrid temporal filtering and deeply learned spatial decomposition to precisely enhance or magnify small pulsatile motion without noise and ring artifacts in surgical videos, enabling surgeons to conveniently and intuitively locate vessels with naked eyes without additional configurations. The main contributions of this work are clarified as follows:

First, We propose a robust hybrid temporal filtering that fuses the advantages of bandpass filtering and third-order Gaussian filtering. The hybrid temporal filtering can essentially magnify pulsatile motion more consistent with reality, but also keep non-pulsating regions unmagnified or unchanged in the magnified endoscopic videos. Next, Compared to current methods, our new strategy to integrate hybrid temporal filtering with deeply learned spatial decomposition can greatly reduce noise and ring artifacts while significantly magnifying pulsatile motion. Additionally, this work also presents a thorough comparison of Eulerian motion magnification methods for endoscopic videos to precisely and intuitively identify vessels or neurovascular bundles during robotic surgery.

## II. APPROACHES

This section details our new motion magnification method that contains three steps: deeply learned spatial decomposition, hybrid temporal filtering, and motion reconstruction. Fig. 1 illustrates the pipeline of our proposed method.

### A. Deeply Learned Spatial Decomposition

Let  $\mathcal{V} = \{V(x, 1), V(x, t), \dots, V(x, N)\}_{t=1}^N$  ( $N$  is the number of frames) be a set of video sequences. Each frame  $V(x, t)$  is spatially decomposed into two feature maps of texture  $\mathcal{S} = \{S(x, 1), S(x, t), \dots, S(x, N)\}_{t=1}^N$  and shape  $\mathcal{I} =$

$\{I(x, 1), I(x, t), \dots, I(x, N)\}_{t=1}^N$  by the encoder [13]:

$$\text{Decompose}(V(x, t)) = [S(x, t), I(x, t)], \quad (1)$$

where texture  $S(x, t)$  indicates the intensity and shape  $I(x, t)$  represents the motion information on image  $V(x, t)$  in accordance with the work [13].

### B. New Hybrid Temporal Filtering

After obtaining  $\mathcal{I} = \{I(x, 1), I(x, t), \dots, I(x, N)\}_{t=1}^N$ , we propose hybrid temporal filtering to extract a set of interested motion-frequency components from these representations.

Our proposed filtering first uses bandpass filtering and third-order Gaussian filtering to extract any interesting components from the motion signal  $I(x, t)$ , obtaining two interesting components  $I_{BP}(x, t)$  and  $I_{TOG}(x, t)$ , respectively. After that, the proposed hybrid filtering fuses the two extracted components to obtain new interested components  $I_{HT}(x, t)$ .

Bandpass filtering extracts the interested component  $I_{BP}(x, t)$  from  $I(x, t)$  by using the discrete short-time Fourier transform (STFT) to deal with finite frames at one time:

$$\text{STFT}\{\mathcal{I}(\tau, \omega)\} = \sum_{t=-\infty}^{\infty} I(x, t) \omega(t - \tau) e^{-j\omega t}, \quad (2)$$

where  $\omega(\tau)$  is the window function. STFT is usually performed by the fast Fourier transform (FFT)  $\mathcal{I}(\tau, \omega)$ . The spectrum  $\mathcal{I}(\tau, \omega)$  contains the low- and high-frequency components corresponding to the pulsatile motion and lung-breathing/heart-beating motion in surgical videos, respectively. We preserve the high-frequency component and obtain the filtered component  $I_{BP}(x, t)$  by the inverse FFT.

Janatka *et al.* [14] believe that the third-order components of  $I(x, t)$  are useful characteristics to represent the pulsatile motion. Therefore, we also use third-order Gaussian filtering to

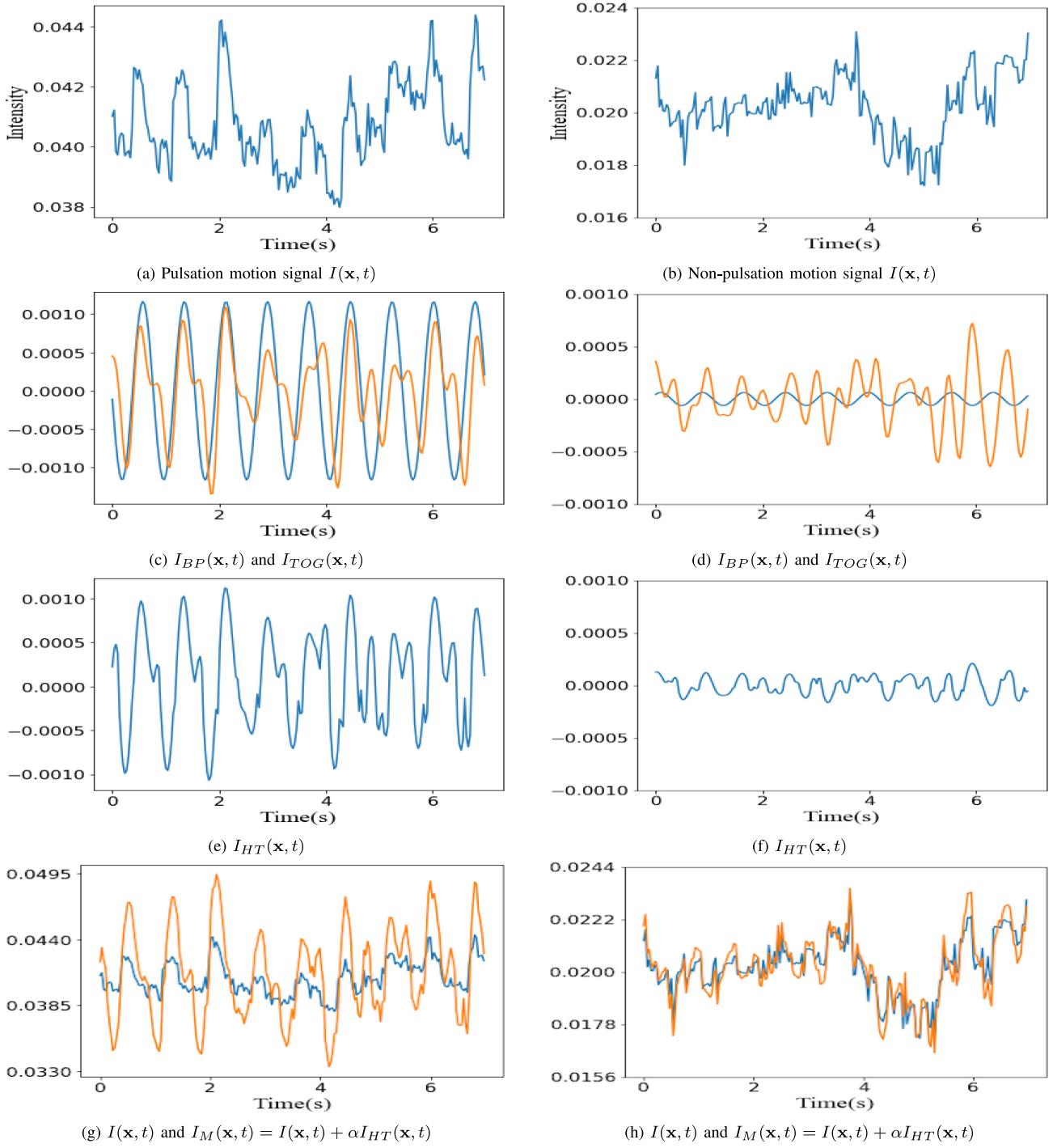


Fig. 2. Pulsation (1st column) and non-pulsation (2nd column) motion signals are extracted by bandpass filtering ((c)-(d) blue), third-order Gaussian filtering ((c)-(d) orange), and our hybrid temporal filtering ((e) and (f)). (g) and (h) show the input (blue) and magnified (orange) motion signals based on our hybrid temporal filtering. The magnification factor  $\alpha$  is set to 5 here.

extract the third-order component  $I_{TOG}(\mathbf{x}, t)$

$$I_{TOG}(\mathbf{x}, t) = I(\mathbf{x}, t) \otimes \frac{\partial^3 G_\sigma(t)}{\partial t^3}, \quad (3)$$

where  $\otimes$  is the convolution operator and  $G_\sigma(t)$  is the Gaussian kernel with the standard deviation  $\sigma = r/4w\sqrt{2}$  computed by [16], where  $r$  is the frame rate of the video and  $w$  is the interested motion frequency. In order for the convolution

kernel to contain the entire third-order Gaussian function, the window size should be set to more than  $r/w$ . Note that the filter horizontally flips the kernel and sets the window center in the crest of the third-order Gaussian function to make the trend and initial phase of the component consistent with the original signal, respectively.

Fig. 2(a), (b) show the pulsation and non-pulsation signals. Fig. 2(c), (d) demonstrate the advantages and disadvantages of

the bandpass and third-order Gaussian filters. Bandpass filtering can effectively extract the interested component and introduce little noise at the non-pulsating position since the size of the extracted component depends on the similarity between the motion signal and sinusoidal wave at the interested frequency, which also results in loss of real characteristics of the original signal. The third-order Gaussian filtering can maintain real characteristics of the original signal at the pulsatile position but introduce noise at the non-pulsating position since this filtering only perceives limited range of frames.

Based on the filtered pulsation and non-pulsation motion analysis above, we can fuse the results of the two filtering and obtain  $\tilde{I}_{HT}(\mathbf{x}, t)$  in the way that the component  $I_{BP}(\mathbf{x}, t)$  are used to weight or constrain the component  $I_{TOG}(\mathbf{x}, t)$

$$\tilde{I}_{HT}(\mathbf{x}, t) = |I_{BP}(\mathbf{x}, t)|I_{TOG}(\mathbf{x}, t), \quad (4)$$

which possibly leads to an over-large amplitude of the fused component at the pulsatile position. Fortunately, the amplitudes of the two components extracted by the two filtering are almost the same at the pulsatile position (Fig. 2). Therefore, a hybrid temporal filtering naturally integrating the bandpass and third-order Gaussian filtering can be established to obtain a new fused motion component  $I_{HT}(\mathbf{x}, t)$

$$I_{HT}(\mathbf{x}, t) = \text{sgn}(I_{TOG}(\mathbf{x}, t))\sqrt{\tilde{I}_{HT}(\mathbf{x}, t)}. \quad (5)$$

Our hybrid temporal filtering aims to keep the consistent characteristics between the unmagnified and magnified pulsatile motion while reducing noise and artifacts. The new interested component  $I_{HT}(\mathbf{x}, t)$  actually fused the advantages of  $I_{BP}(\mathbf{x}, t)$  and  $I_{TOG}(\mathbf{x}, t)$ . Fig. 2(e)(f) illustrates that our filtering can maintain the original pulsation characteristics at the pulsatile motion position while suppressing noise and artifacts at the non-pulsating position.

We use our proposed hybrid temporal filtering to process all the motion representations and obtain new motion components  $\mathcal{I}_{HT} = \{I_{HT}(\mathbf{x}, 1), I_{HT}(\mathbf{x}, 2), \dots, I_{HT}(\mathbf{x}, N)\}_{t=1}^N$  that are magnified and reconstructed in the following section.

### C. Deeply Learned Motion Reconstruction

We magnify the new motion component  $I_{HT}(\mathbf{x}, t)$  extracted above and obtain the magnified motion  $I_M(\mathbf{x}, t)$  by

$$I_M(\mathbf{x}, t) = I(\mathbf{x}, t) + \alpha I_{HT}(\mathbf{x}, t), \quad (6)$$

where  $\alpha$  is the magnification factor experimentally determined. Fig. 2(g)(h) further demonstrate that our filtering can certainly magnify the pulsation signal with the same characteristics as the original signal and unmagnify non-pulsating signals while reducing noise in magnified surgical videos.

Similar to the encoder decomposition, we use the network decoder to reconstruct texture  $S(\mathbf{x}, t)$  and magnified motion  $I_M(\mathbf{x}, t)$  and obtain the magnified frame  $V_M(\mathbf{x}, t)$

$$V_M(\mathbf{x}, t) = \text{Reconstruct}[S(\mathbf{x}, t), I_M(\mathbf{x}, t)]. \quad (7)$$

Finally, we obtain the magnified video  $\mathcal{V}_M = \{V_M(\mathbf{x}, t)\}_{t=1}^N$ .

### III. EXPERIMENTAL SETTINGS

Surgical videos were collected by the da Vinci Si surgical system during robotic-assisted laparoscopic radical prostatectomy. All the data were acquired under a protocol approved by the Research Ethics board of Western University, Canada. We built a video acquisition system that can obtain video  $1920 \times 1080$  sequences at a speed of 30 frames per second. We evaluated all the different motion magnification methods by seven robotic prostatectomy videos. The architecture and parameters of the network encoder and decoder were used the same as the previous work [13]. The training data were natural images from the previous work [13]. We selected the natural images where the target with the enlarged texture was similar to the background, making the training images more consistent to clinical endoscopic images. The threshold of the interested pulsatile motion frequency in this study was determined by observing the intensity signal at the obviously pulsatile motion position in surgical videos.

We visually compare different spatial motion decomposition and temporal filtering methods while using two objective metrics of the structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) [17] to quantitatively evaluate these methods. It is generally reasonable to calculate SSIM and PSNR between the original video and magnified video for quantitative assessment since it is somewhat impossible to generate clinical ground truth data for the validation of endoscopic video motion magnification. We also believe that SSIM can describe the consistency of the pulsatile motion between the magnified and original videos because it mainly measures the structural similarity between images while PSNR can indicate the amount of additional noise introduced after the magnification.

### IV. RESULTS

*1) Magnification Factor:* We experimentally determine the magnification factor  $\alpha$ . We use the same  $\alpha$  to compare our hybrid temporal filtering to bandpass filtering and third-order Gaussian filtering since the amplitude of the interested motion components extracted by the three filtering are almost same at the pulsatile motion position (Fig. 2). We tested many different values of  $\alpha$  and found that the learning-based method using the factor 10 provides good pulsatile motion magnification. Then, we adjusted the factors of the other two decomposition methods in order to generate the (almost) same amplitude of the pulsatile motion as the learning-based, and eventually set the magnification factor to be 25 and 8 for the intensity- and phase-based methods, respectively (Fig. 4).

*2) Decomposition Comparison:* Figs. 3 and 4 investigate the three different motion decomposition methods combined with our hybrid temporal filtering using the fixed magnification factors. The intensity-based method produces much noise and distortion because it directly magnifies the intensity map. Although the phase-based method outperforms the intensity-based one, it still generates much ring artifacts because the phase map obtained by the complex steerable pyramid decomposition can be disordered during magnification. The learning-base method

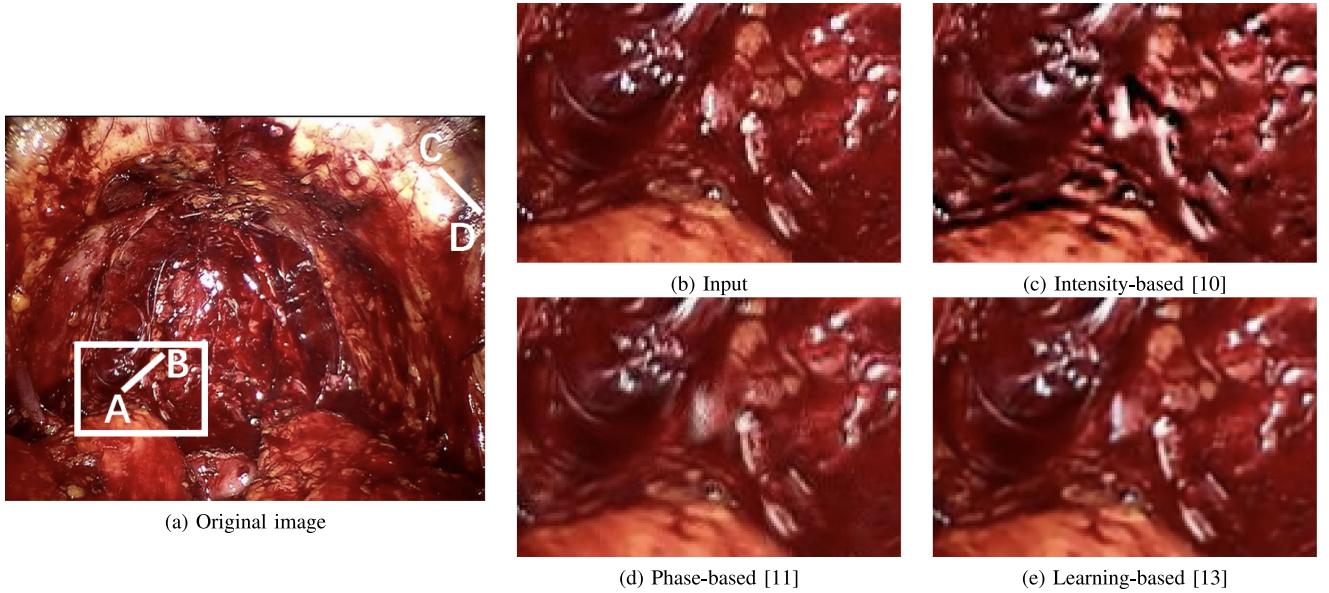


Fig. 3. Comparison of region AB (white square) in the image (a) magnified by the three different spatial decomposition methods integrated with our hybrid temporal filtering. The learning-based decomposition method shows the best performance.

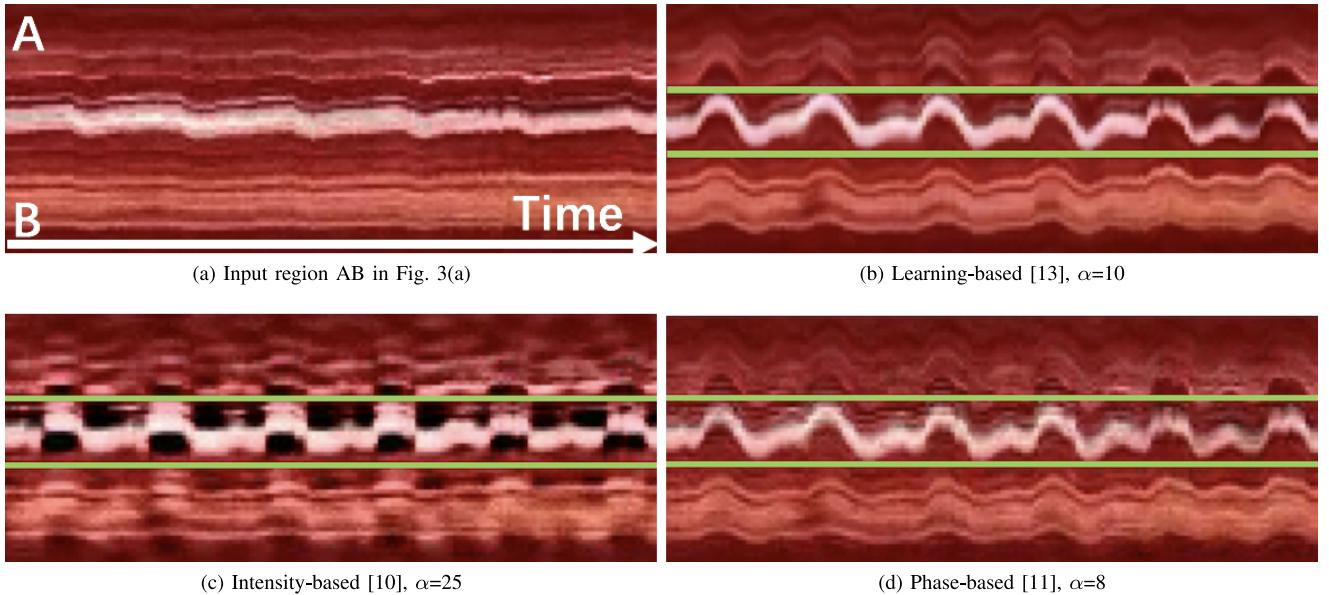


Fig. 4. Experimentally determined the magnification factors by visually comparing the time expansion diagrams generated from three different motion decompositions integrated with hybrid temporal filtering for video motion magnification. The magnification factors used at the same pulsation position (green lines) show the consistent performance.

generally introduces less noise and ring artifacts which can ensure the intensity and texture informations of the magnified video are consistent with the original.

*3) Magnification Comparison:* Fig. 5 displays the magnification results of using the three filters integrated with the learning-based decomposition method. All the three filters can magnify the pulsatile motion but bandpass filtering cannot preserve the original features of pulsation at the pulsation regions, because it only extract and magnify the sinusoidal wave of motion signals. While both third-order Gaussian and hybrid temporal filtering methods can highlight cardio-physiological features, leading to

the magnified video more consistent with reality according to the analysis of pulsatile motion of [14]. More interestingly, the bandpass and hybrid temporal filters introduce little noise on the non-pulsating regions while third-order Gaussian filtering generates additional noise.

Table I summarizes the average SSIM and PSNR coefficients of all seven videos of the nine motion magnification methods. Table I demonstrates that our method  $M_9(0.908, 28.92)$  of hybrid temporal filtering with deeply learned decomposition works much better than the other eight methods, while both the learning-based motion decomposition and hybrid temporal

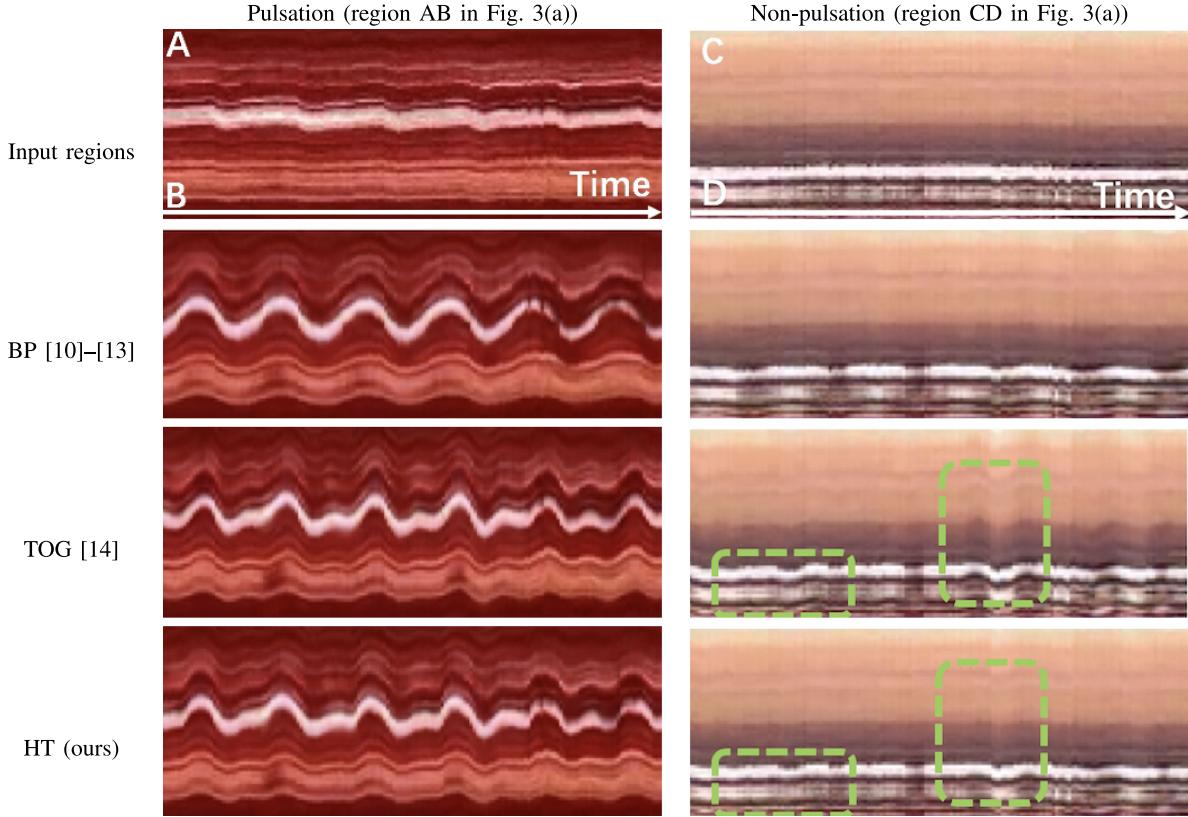


Fig. 5. Visual comparison of time expansion diagrams of pulsation (*left*) and non-pulsation (*right*) regions in the original and magnified videos obtained by the three different temporal filtering strategies integrated with the learning-based decomposition and reconstruction method. Rows BP, TOG, and HT illustrate the magnified results corresponding to bandpass filtering [10]–[13], third-order Gaussian filtering [14], and our proposed hybrid temporal filtering, respectively. The green squares on the diagrams indicate and compare whether noise or artifacts is significantly reduced. Our new hybrid temporal filtering generally outperforms other filters when combining with the learning-based method.

TABLE I

TWO COMMONLY USED IMAGE QUALITY METRICS OF SSIM AND PSNR FOR QUANTITATIVE ASSESSMENT OF THE VIDEO MOTION MAGNIFICATION METHODS THAT COMBINE THE THREE MOTION DECOMPOSITION APPROACHES WITH THE THREE TEMPORAL FILTERS:  $M_i$  INDICATES THE  $i$ -TH METHOD, E.G.,  $M_1$  IS THE MOTION MAGNIFICATION METHOD OF A COMBINATION OF BP [10]–[13] AND INTENSITY-BASED ( $\alpha = 25$ ) [10]

$M_i$ (SSIM, PSNR)	Intensity-based ( $\alpha = 25$ ) [10]	Phase-based ( $\alpha = 8$ ) [12]	Learning-based ( $\alpha = 10$ ) [13]	Average
BP [10]–[13]	$M_1(0.753, 21.91)$	$M_2(0.793, 25.65)$	$M_3(0.875, 27.20)$	(0.807, 24.92)
TOG [14]	$M_4(0.740, 22.87)$	$M_5(0.831, 27.31)$	$M_6(0.857, 27.00)$	(0.809, 25.73)
HT (ours)	$M_7(0.804, 24.26)$	$M_8(0.840, 27.52)$	$M_9(0.908, 28.92)$	(0.851, 26.90)
Average	(0.766, 23.01)	(0.821, 26.83)	(0.880, 27.71)	--

filtering outperform the other two decomposition and filtering methods, respectively.

4) *Subjective Assessment*: Six surgeons manually and subjectively evaluated all the magnified endoscopic videos. All of them generally agreed that our method can effectively assist them to identify the location of vessels and neurovascular bundles in endoscopic prostatectomy videos. Additionally, the visual quality of the magnified videos obtained by our method was better than the results of others through their intuitive perception.

## V. DISCUSSION

This work aims to magnify endoscopic video pulsatile motion to precisely localize vessels and neurovascular bundles during robotic surgery. We analyze the properties of pulsatile motion corresponding to various temporal filters used in current motion

magnification methods and explore a new video motion magnification strategy with robust hybrid temporal filtering and deeply learned spatial decomposition. Our new hybrid temporal filtering fuses the advantages of the bandpass and third-order Gaussian filters which can highlight cardio-physiological features and introduce little noise and artifacts, and we establish an accurate and robust video motion magnification framework that employs the proposed hybrid temporal filtering with deeply learned spatial decomposition. Such a new video motion magnification strategy much outperforms the other eight methods.

Our method still suffer from some potential limitations. First, the magnification factor  $\alpha$  was subjectively determined to provide good and intuitive visual quality of the experimental results. Selecting the optimal magnification factor is a challenging problem for different patients in clinical practices. Second, our

new motion magnification approach still introduces noise in the magnified videos since the learning-based spatial decomposition method was trained by natural video sequences different from complex surgical video images. Additionally, the computational efficiency of our proposed filtering and magnification methods implemented on Matlab does not meet real-time requirement of clinical applications, and can be improved by GPU and code optimization techniques.

In summary, this work presents a new robust video motion magnification method to precisely locate vessels and neurovascular bundles in robotic surgery. Such a method successfully combines hybrid temporal filtering with deeply learned spatial decomposition, outperforming other magnification methods.

#### ACKNOWLEDGMENT

The authors would like to thank the assistance of Drs. Yan-Ping Du, Xiao Cheng, Suyu Chen, and Jianhua Chen for manual evaluation of the magnified video motion results.

#### REFERENCES

- [1] T. D. McClure *et al.*, “Use of MR imaging to determine preservation of the neurovascular bundles at robotic-assisted laparoscopic prostatectomy,” *Radiology*, vol. 262, no. 3, pp. 873–883, 2012.
- [2] A. Tarutis, A. C. Timmermans, P. C. Wouters, M. Kacprowicz, G. M. van Dam, and V. Ntziachristos, “Optoacoustic imaging of human vasculature: Feasibility by using a handheld probe,” *Radiology*, vol. 281, no. 1, pp. 256–263, 2016.
- [3] G. Marquis-Gravel *et al.*, “Ultrasound guidance versus anatomical landmark approach for femoral artery access in coronary angiography: A randomized controlled trial and a meta analysis,” *J. Interventional Cardiol.*, vol. 31, no. 4, pp. 496–503, 2018.
- [4] T. B. Manny, M. Patel, and A. K. Hemal, “Fluorescence-enhanced robotic radical prostatectomy using real-time lymphangiography and tissue marking with percutaneous injection of unconjugated indocyanine green: The initial clinical experience in 50 patients,” *Eur. Urol.*, vol. 65, no. 6, pp. 1162–1168, 2014.
- [5] R. M. Schols, N. J. Connell, and L. P. S. Stassen, “Near-infrared fluorescence imaging for real-time intraoperative anatomical guidance in minimally invasive surgery: A systematic review of the literature,” *World J. Surg.*, vol. 39, no. 5, pp. 1069–1079, 2015.
- [6] A. N. Sridhar *et al.*, “Image-guided robotic interventions for prostate cancer,” *Nat. Rev. Urol.*, vol. 10, no. 8, pp. 452–462, 2013.
- [7] F. Adams *et al.*, “Algorithm-based motion magnification for video processing in urological laparoscopy,” *J. Endourol.*, vol. 31, no. 6, pp. 583–587, 2017.
- [8] D. Alinovi, G. Ferrari, F. Pisani, and R. Raheli, “Respiratory rate monitoring by video processing using local motion magnification,” in *Proc. Eur. Signal Process. Conf.*, 2018, pp. 1794–1798.
- [9] A. J. McLeod, J. S. Baxter, S. de Ribaupierre, and T. M. Peters, “Motion magnification for endoscopic surgery,” in *Soc. Photo-Opt. Instrum. Engineers Med. Imag.*, 2014, Art. no. 90360C.
- [10] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, “Eulerian video magnification for revealing subtle changes in the world,” *ACM Trans. Graph.*, vol. 31, no. 4, pp. 1–8, 2012.
- [11] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, “Phase-based video motion processing,” *ACM Trans. Graph.*, vol. 32, no. 4, pp. 1–10, 2013.
- [12] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, “Riesz pyramids for fast phase-based video magnification,” in *Proc. IEEE Int. Conf. Comput. Photography*, 2014, pp. 1–10.
- [13] T.-H. Oh *et al.*, “Learning-based video motion magnification,” in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 633–648.
- [14] M. Janatka, A. Sridhar, J. Kelly, and D. Stoyanov, “Higher order of motion magnification for vessel localisation in surgical video,” in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2018, pp. 307–314.
- [15] Y. Zhang, S. L. Pintea, and J. C. Van Gemert, “Video acceleration magnification,” in *Proc. Conf. Comput. Pattern Recognit.*, 2017, pp. 529–537.
- [16] K. Mikolajczyk and C. Schmid, “Indexing based on scale invariant interest points,” in *Proc. Int. Conf. Comput. Vis.*, 2001, pp. 525–531.
- [17] A. Hore and D. Ziou, “Image quality metrics: PSNR vs. SSIM,” in *Proc. 20th Int. Conf. Pattern Recognit.*, 2010, pp. 2366–2369.