Universal digital filtering for denoising volumetric retinal OCT and OCT angiography in 3D shearlet domain

JIANLONG YANG1,* , YAN HU2, LIYANG FANG1, JUN CHENG1, AND JIANG LIU1,2

1Cixi Institute of Biomedical Engineering, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, China
2Department of Computer Science and Engineering, Southern University of Science and Technology, China
*Corresponding author: yangjianlong@nimte.ac.cn

Retinal optical coherence tomography (OCT) and OCT angiography (OCTA) suffer from the degeneration of image quality due to speckle noise and bulk-motion noise, respectively. Because the cross-sectional retina has distinctive features in OCT and OCTA B-scans, existing digital filters that can denoise OCT efficiently are unable to handle the bulk-motion noise in OCTA. In this Letter, we propose a universal digital filtering approach that is capable of minimizing both types of noise. Considering the retinal capillaries in OCTA are hard to differentiate in B-scans while having distinct curvilinear structures in 3D volumes, we decompose the volumetric OCT and OCTA data with 3D shearlets thus efficiently separate the retinal tissues and vessels with the noise in this transform domain. Compared with wavelets and curvelets, the shearlets provide better representation of the layer edges in OCT and the vasculature in OCTA. Qualitative and quantitative results show the proposed method outperforms the state-of-the-art OCT and OCTA denoising methods. Besides, the superiority of 3D denoising is demonstrated by comparing the 3D shearlet filtering with its 2D counterpart.

© 2019 Optical Society of America

OCIS codes: (110.4500) Optical coherence tomography; (030.4280) Noise in imaging systems; (110.4155) Multiframe image processing; (170.4470) Ophthalmology.

http://dx.doi.org/10.1364/ao.XX.XXXXXX

Optical coherence tomography (OCT) is a non-invasive cross-sectional 3D imaging modality that has been widely used in the studies and clinical diagnosis of various diseases in ophthalmology [1]. In recent years, its functional extension, OCT angiography (OCTA) has been becoming popular because it can acquire capillary vasculature without injecting contrast agents such as fluorescent dyes [2].

OCT optical systems use partially coherence lasers as the light source, so the unavoidable spike noise is a major constrain of the imaging quality [3]. On the other hand, OCT calculates the temporal decorrelation of several repeated OCT scans, which boosts the difference between vasculature and the surrounding tissue. However, both the blood flow in vessels and the involuntary eye movement of human lead to high decorrelation. The latter causes the bulk-motion noise in OCTA images [4].

Minimizing the speckle noise in OCT while preserving the edge and texture information of retinal layers is very challenging. Numerous methods and algorithms have been developed for this task [5–9]. For retinal OCTA, most of the vessels and capillaries distribute on the plane perpendicular to the incident probe light, which makes the bulk-motion noise difficult to remove in the cross-sectional (B-scan) OCTA images. Frequency compounding has been used in denosing the bulk motion but sacrifices the axial resolution [10]. Jia et al. further used mean subtraction to minimize the bulk motion noise [11].

However, none of the methods above can handle both the speckle and bulk-motion noise simultaneously. As demonstrated in Fig 1, we employed the state-of-the-art denoising algorithm, block matching & 4D collaborative filtering (BM4D) [8], and the wavelet based singular value decomposition (K-

Fig. 1. Comparison of different methods in denoising OCT and OCTA, including the BM4D, K-SVD, and the proposed 3D shearlet filtering.
Fig. 2. Workflow of the proposed 3D shearlet filtering for OCT and OCTA denoising. OCTA data is employed here for demonstration. Firstly, the OCTA volume is decomposed into a set of 3D shearlet coefficients, each of them represents different features of the original data. We then apply the hard threshold to all the coefficients. Finally, all of the thresholded coefficients are used to reconstruct the denoised OCTA volume.

SVD method [9] to denoise OCT and OCTA. The results show they could suppress the speckle noise in OCT B-scan but lead to blur and the loss of capillaries in OCTA B-scan and en face OCTA. In this Letter, we propose to denoise both OCT and OCTA with digital filtering in shearlet domain. Shearlet is an anisotropic extension of wavelet [12], which has been mathematically proven to be more efficient than wavelet and curvelet in representing highly anisotropic features like edges and curvilinear structures [13]. Instead of using 2D shearlets to decompose single B-scan, where the vessels and bulk-motion noise have similar point-like structures, we propose to leverage the curvilinear information of retinal vasculature in adjacent B-scans (volumetric data) by using 3D shearlet decomposition. In the 3D shearlet domain, the representations of different shearlets are capable of separating the retinal tissues and vessels with the noise efficiently. So we could minimize the noise by simply applying hard thresholding. Note that more sophisticated techniques, such as adaptive thresholding and total variation, have been widely adapted in denoising tasks [7, 13]. We employ the hard thresholding by following the principle of Occam’s razor, which implies using the simplest approach will justify the robustness of the proposed method to the most extent.}

Figure 2 is the workflow of the proposed 3D shearlet filtering for OCT and OCTA denoising. OCTA data is employed here for demonstration. The OCTA volume is decomposed into a set of 3D shearlet coefficients, each of them represents different features of the original data. For a specific B-scan which locates at the green line position, the retinal vessels and the bulk-motion noise are represented diversely by different shearlets, e.g., more noise features are emerged in $c_1$ and $c_{28}$ than those in $c_{14}$. We then apply the hard threshold to all the coefficients. As demonstrated in Fig. 2, Because $c_1$ and $c_{28}$ has more noise features, so less information is left after the thresholding. Finally, all of the thresholded coefficients are used to reconstruct the denoised OCTA volume.

Briefly, the shearlet transform is mapping a multivariate signal $f \in L^2(\mathbb{R}^2)$ to a set of coefficients $c$ of a generating function $\psi$.

$$f \rightarrow S\Phi f = (f, \psi)$$

The generating function $\psi$, which includes parabolic scaling variable $a > 0$ for altering resolutions, shearing variable $s \in \mathbb{R}$ for altering directions, and translation variable $t \in \mathbb{R}^2$ for altering positions, can be written as

$$\psi_{a,s,t} = a^{3/4} \psi(S_a \begin{bmatrix} 0 & a^{1/2} \\ a & 0 \end{bmatrix} \cdot -t)$$

where $S_a$ is the shearing matrix in the form of

$$S_a = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}.$$  \hspace{1cm} (3)

We can see the shearlet transform is quite similar to the wavelet transform except it can handle anisotropic scaling and shearing. However, the direct numerical implementation of the shearlet transform is difficult. Because of the directional bias problem, the so-called cone-adapted shearlet system needs to be introduced [14]. It partitions the Fourier-domain into four cones including two horizontal and two vertical high-pass region and a squared low-pass region. A scaling function $\phi$ can cover the squared region. Two new generating functions $\phi_h$ and $\phi_v$ are associated to the horizontal and vertical cones, respectively. Then the cone-adapted shearlet system $S\Phi_{\psi_h,\psi_v} = \Phi(\phi) \cup \Psi_h(\psi_h) \cup \Psi_v(\psi_v)$ can be written as

$$\Phi(\phi) = \{ \phi_t = \phi(\cdot - t) : t \in \mathbb{R}^2 \}$$

$$\Psi_h(\psi_h) = \{ \psi_h = a^{-3/4} \psi_h(A_h S^{-1}_a \cdot -t) : a \in (0,1], |s| \leq 1 + a^{1/2}, t \in \mathbb{R}^2 \}$$

$$\Psi_v(\psi_v) = \{ \psi_v = a^{-3/4} \psi_v(A_v S^{-1}_a \cdot -t) : a \in (0,1], |s| \leq 1 + a^{1/2}, t \in \mathbb{R}^2 \}$$

where $A_h = diag(a, a^{1/2})$ is the scaling matrix for the horizontal cone. $A_v = diag(a^{1/2}, a)$ is the scaling matrix for the vertical cone. The cone-adapted shearlet system can be directly digitalized by introducing a sampling factor $c = (c_1, c_2) \in \mathbb{R}_+^2$ in the translation index.

The 3D digital shearlet filter is the product of two 2D digital shearlet filters in the frequency domain [15]:

$$\hat{\psi}_{t_k}^{3D}(\xi) = \hat{\phi}_j^{2D}(\xi_1, \xi_2) \hat{\psi}_k^{2D}(\xi_1, \xi_3)$$

where $j$ and $k$ are the discrete scale and shearing parameters, respectively. $\xi_{1,2,3}$ is the 3D coordinates in Fourier domain. Thus
the 3D shearlet decomposition of a signal $f \in L^2(\mathbb{Z}^3)$ can be defined as

$$DSH_{j,k,m}(f) = \langle \psi_{j,k} \ast f \rangle(m),$$

(6)

where $m$ is the discrete translation parameter.

The proposed method was realized in MATLAB R2018a. The 3D shearlet decomposition and reconstruction of the volumetric OCT/OCTA data are implemented using the open-source shearLab 3D code [16]. A total of 99 shearlets are used in the representation of edges, curvilinear structures, and texture in OCT/OCTA. We optimized the hard thresholding level via visual comparison of the denoised images. Using a personal workstation with Intel Xeon E5-2695 CPU, 128GB RAM, and Nvidia GTX 1080 Ti 12GB GPU, the processing time for a single OCT or OCTA volume is $\sim 70$ s without GPU acceleration and $\sim 19$ s with GPU acceleration. It also should be mentioned that the 3D shearlet filtering also has time-consumption advantage for the denoising tasks. Under the same hardware configuration, the processing time using the K-SVD and BM4D for the same OCT/OCTA volume are $\sim 47$ s, and $\sim 332$ s, respectively.

We employed TOPCON DRI OCT-1 ATLANTIS for OCT acquisition and ZEISS CIRRUS OCT with AngioPlex module for collecting OCTA data. Each of the volumetric OCT data covers $6 \times 6 \text{ mm}^2$ field-of-view (FOV) corresponding to 256 B-frames. Each B-frame contains 512 A-lines and 992 pixels along the depth direction. The OCTA data has 1024 pixels along the depth direction. The human study protocol was approved by the Institutional Review Board of Cixi Institute of Biomedical Engineering, Chinese Academy of Sciences, and followed the tenets of the Declaration of Helsinki. Twenty healthy subjects (age $25.1 \pm 8.5$ years) were imaged by the machines and scan protocols mentioned above. Besides, the high-definition (HD) scan mode of the TOPCON machine, which averages 96 repeated scans at the same B-scan position, was employed as the denoised results using spatial averaging.

The minimization of the OCTA bulk-motion noise using the 3D shearlet filtering is compared with two widely-used methods, median subtraction [11] and pixel averaging [10]. The pixel averaging is used to mimic the performance of the frequency compounding (used in the SSADA algorithm for removing the bulk motion). Because the frequency compounding is implemented on OCTA raw data, which is unavailable in the commercial systems. An Gaussian kernel of [6 1] is employed for the depth-direction averaging. Figure 3 shows the qualitative comparison of a representative OCTA B-frame. The intensity profiles at the green line position are shown in Fig. 3(A). In the zoomed view on the left side, which corresponds to the foveal avascular zone, we can see the 3D shearlet filtering achieves a flat background. The median subtraction method has a lower background value due to the subtraction operation, but still has a larger variation on the signal. As for the zoom-in view on the right side, we can see the 3D shearlet filtering achieves the highest signal strength. The mean-subtraction method is capable of reducing background but has a side-effect of weakening the signal strength.

We further evaluated the performance of OCTA in 3D volume. For clear visualization of the results, we segmented the OCTA volume into three en face projected slabs including superficial vascular plexus (SVP), intermediate capillary plexus (ICP), and deep capillary plexus (DCP) as shown in Fig. 4. We followed their definitions in [11]. We can see the proposed 3D shearlet filtering provides the best signal-to-noise ratio (SNR) for all the retinal slabs at different depths. Note that, for enhancing and denoising the curvilinear vasculature, various digital filters such as Frangi filter and Gabor filter [17] could be efficient in the en face plane. But they are unable to handle the volumetric OCTA data.

For evaluating the denoising performance of the proposed 3D shearlet filtering on OCT data, we also included 2D shearlet filtering in the comparison, which has been shown to be efficient in OCT denoising [18] very recently. Figure 5 demonstrates the denoised OCT B-frame using the 2D shearlet (B) and the 3D shearlet proposed in this work (C). Figure 5(A) is a HD scan image from TOPCON OCT system using the principle of spatial averaging. We can see the 3D shearlet filtering provides better SNR and resolution than its 2D counterpart, which confirms that using the additional information provided by adjacent B-frames in OCT volume could benefit the digital denoising performance [6]. In combination with the top row in Fig. 1, we can further analyze the OCT denoising performance. An important criterion is the preservation of the external limiting membrane (ELM) layer. We can see the middle part of the ELM is well preserved in all of the denoised image while its right part is barely seen in the K-SVD denoised image. For the left part of the ELM layer is only can be clearly seen in the denoised images using spatial averaging and 3D shearlet filtering. A similar situation happens to another important criterion, which is the separation among the photoreceptor layers and retinal pigment epithelium. Clear separation may only be observed in the denoised image using the spatial averaging and 3D shearlet filtering. As for the image SNR, the BM4D algorithm and the 3D shearlet filtering not only minimize the speckle noise but also significantly enhances the SNR compared with other approaches including the spatial averaging. However, the enhancement also leads to the loss of texture information, which is severer in the BM4D result.

The quantitative evaluation of the OCT denoising per-
which can be calculated as $\text{PSNR} = 10 \log_{10} \frac{1}{mn \sum \frac{(I_m - I_n)^2}{I_m^2}}$, where $m$ and $n$ are the column and row numbers of the images, respectively. $I_{\text{max}}$ is the maximum pixel value of the denoised image. (3) Structure similarity index (SSIM), which is defined as $\text{SSIM}(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$, where $\mu_x$, $\mu_y$, $\sigma_x$, $\sigma_y$, and $\sigma_{xy}$ are the local means, standard deviations, and cross-covariance for images $x$, $y$. $C_1 = (0.01T)^2$ and $C_2 = (0.03T)^2$ with $T$ is the maximum value allowed for the data. The results is shown in Table 1. In consistency with the qualitative comparison discussed above, the 3D shearlet filtering outperforms other OCT denoising methods on all three metrics.

In summary, we have demonstrated the capacity of the 3D shearlet filtering in denoising both the volumetric OCT and OCTA data. Comparing with other denoising methods, it is efficient and fast thus has broad prospects in OCT/OCTA-related biomedical applications. Further improvement of this method could be the incorporation with the strategies that have been widely used in denoising tasks, such as local or sub-band adaptive filtering and total variation regularization.

Funding Information. Ningbo 3315 Innovation team grant; Zhejiang Provincial Natural Science Foundation (LQ19H180001); Ningbo Public Welfare Science and Technology Project (2018C50049).

Disclosures. The authors declare no conflicts of interest.

Table 1. Quantitative comparison of the performance of the OCT denoising methods

| Methods     | MSE  | PSNR | SSIM |
|-------------|------|------|------|
| Original    | 0.29 | 16.32| 0.31 |
| K-SVD       | 0.11 | 19.15| 0.63 |
| BM4D        | 0.07 | 21.46| 0.75 |
| Shearlet 2D | 0.10 | 20.08| 0.70 |
| Shearlet 3D | 0.06 | 21.81| 0.76 |

formance could be conducted, under the assumption that the spatial averaging denoised image is noise-free, namely, the so-called ground truth in computer vision and image processing fields. This approach has been adopted in previous works [6, 18]. We employed the HD scan image in Fig. 5(A) as the ground truth image $I_{gt}$. It was compared with the original or denoised images $I$. We utilized three quantitative metrics in the evaluation. (1) Mean square error (MSE). It is defined as $\text{MSE} = \frac{\Sigma (I - I_{gt})^2}{\Sigma |I_{gt}|}$. (2) Peak signal to noise ratio (PSNR),

Fig. 4. Denoising comparison of the superficial, intermediate, and deep retinal slabs segmented from a OCTA volume. SVP: superficial vascular plexus. ICP: intermediate capillary plexus. DCP: deep capillary plexus.

Fig. 5. Denoising comparison using (A) spatial averaging, (B) 2D shearlet filtering, and (C) 3D shearlet filtering.
nicka, C. G. Owen, and S. A. Barman, Comput. methods programs biomedicine 108, 407 (2012).

18. M. Xu, C. Tang, M. Chen, Y. Qiu, and Z. Lei, Opt. Lasers Eng. 122, 265 (2019).