Optics Communications xxx (xxxx) xxx

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

# Axial and horizontal registration guided speckle suppression in single-line HD mode for retinal optical coherence tomography images

Ming Liu [a,1], Xinjian Chen [a,b,1], Bo Wang [a,*]

[a] School of Electronic and Information Engineering, Soochow University, 215006, Suzhou, China
[b] State Key Laboratory of Radiation Medicine and Protection, Soochow University, 215123, Suzhou, China

ARTICLE INFO

*Keywords:*
Optical coherence tomography
Speckle suppression
Image registration
Eye movement

ABSTRACT

We propose an image registration method to suppress the speckle noise in retinal optical coherence tomography(OCT) adapted to single-line HD mode hardware implementation. In single-line HD mode, multiple cross-section images at the same straight scan line are obtained repeatedly and rapidly for averaging the speckle noise. Not like static objects, cross-section retinal images are often blurred by direct averaging due to unintentional eye/head motions in axial and horizontal directions. An image registration method which combines regularized dynamic programming algorithm with hill-climbing algorithm is demonstrated to correct axial and lateral motion shifts efficiently. The unaligned images are discarded automatically using a criterion based on probability densities of image gray levels. Finally, an adaptive gamma correction method is adopted to adjust the contrast of averaged OCT image for better retinal feature visualization. The experimental results of the algorithm based on our lab-built OCT setup show the improvement of image quality. The method is automatic and rapid which is potential to be applied in the OCT imaging with micro-motion targets.

## 1. Introduction

Optical coherence tomography(OCT) as an *in vivo*, high resolution and high speed imaging modality, has been widely applied to show cross-sectional structure profiles for biological tissue [1,2]. Due to the transparency of human eyes, OCT is regarded as the ground truth in the field of ophthalmology, especially in the assistant diagnosis of retina diseases, such as glaucoma, macular hole, age-related macular degeneration, *etc* [3–5].

High quality OCT images with suppressed speckle noise are pursued for detailed observation of retinal layer structure in the ophthalmology clinical application [6–8]. The speckle noise is inherent and caused by the interference of light waves scattered by different scattering particles existing in the biological tissue [9]. Some hardware-based optical system design methods are proposed for OCT speckle reduction, which depend on acquiring different uncorrelated speckle patterns for averaging or compressing the sample to create speckle by angular compounding [10,11], frequency compounding [12], spatial compounding [13] and strain compounding [14]. In adaptive optics OCT(AO-OCT), wavefront sensing and wavefront correction devices are employed which can set the focus at the selected layer to reduce the imaging speckle as well as the limited depth of focus [15]. However, these approaches need complicated modifications and additional data acquisition, which are hard to be implemented directly in clinical ophthalmology OCT scanners. Regarded as post-processing approaches, software-based methods rely on filtering or reconstructing the obtained cross-section image(also called B-scan). The methods based on spatial filtering aim to smooth every single B-scan to filter out the influences of speckle noise, such as Wiener filtering [16], bilateral filtering [17], anisotropic diffusion filtering [18]. But these methods may have over-fitting and over-smoothing problems. In addition, methods based on sparse transform like wavelet [19,20], curvelet [21] and dictionary learning [22,23] have achieved good performance in noise reduction but also have some problems, such as insufficient image feature representation, difficult threshold selection, and time-consuming dictionary learning. Moreover, statistical model-based methods are also common for image denoising [24,25]. But the existing statistical models are mostly univariate and may not consider all the key information like neighboring pixels into the model. The low-rank decomposition methods [26,27] are effective in OCT images denoising as well but it relies on local similarity. In the last few years, deep learning networks for despeckling OCT images are also introduced due to their high performance in many computer vision tasks. Deep learning networks treat the denoising problem as a pixel-regression task and try to explore the mapping relationship between the origin and denoisd image pair. The OCT slices from any imaging mode can be despeckled after the inference of the trained model and the inference time is relatively

* Corresponding author.
*E-mail address:* wangbo@suda.edu.cn (B. Wang).
[1] These authors contributed equally.

https://doi.org/10.1016/j.optcom.2021.126807
Received 22 September 2020; Received in revised form 11 January 2021; Accepted 18 January 2021
Available online xxxx
0030-4018/© 2021 Elsevier B.V. All rights reserved.

Please cite this article as: M. Liu, X. Chen and B. Wang, Axial and horizontal registration guided speckle suppression in single-line HD mode for retinal optical coherence tomography images, Optics Communications (2021) 126807, https://doi.org/10.1016/j.optcom.2021.126807.

low. The structural information is paid more attention to because of the special design of the loss function and training module. Mao et al. [28] proposed deep convolutional encoder–decoder networks with symmetric skip connections but it needs long-term dependency. Zhang et al. [29] assume that OCT speckle noise is additive noise and proposed a network structure with residual blocks. Abbasi et al. [30] proposed a multi-input fully-convolutional networks (MIFCN) to further improve the performance. Ma et al. [31] proposed a cGAN based network that focused on the edge information to a greater extent. Guo et al. [32] proposed an end-to-end structure-aware noise reduction generative adversarial network (SNR-GAN) which achieved subtle structural preservation. However, all the deep learning studies require a large amount of OCT data for training which would limit the potential applications. Furthermore, the deep learning methods take images after averaging as ground truth and the result of the methods can only be close to the ground truth but cannot exceed.

Besides the independent hardware-based and software-based methods, there is a mixed way to generate the HD OCT image in many OCT applications. A specific mode (often called as single-line HD mode) is designed to obtain B-scans of the same retinal location by quickly repeating the scan multiple times along only one straight scan line. Considering the laser safety limit and to avoid large eye/head movements, the total scan time is often limited to less than 2.5s. In the limited time, 20~50 B-scans are recorded according to the pixel number and exposure time setting. Then such B-scans are averaged to generate a B-scan with higher quality by suppressing random speckle noises in each B-scan.

However, the OCT image shifts at ten micrometer scale caused by eye/head movement can still be observed in many cases which obstruct the direct averaging. Many works were demonstrated to correct the motion artifacts before multiple B-scans averaging. Jorgensen et al. [33] reported an algorithm based on the regularized dynamic programming with an iterative framework, which is the first to correct both axial and lateral displacements. But given that the framework needs to align multiple B-scans in one iteration, it cannot correct large lateral displacements with the consideration of computational time. Alonso-Caneiro et al. [34] proposed a registration method that relies on the affine-motion scheme. This method involves image transformations of higher degrees but the registration accuracy needs to be improved especially in local areas. Zhang et al. [35] combined the above two methods and reported a two-step image registration method to reduce the speckles in OCT images. This method improves the registration accuracy to a large extent and avoids the discontinuities in the retinal layers. But in the local step, the regularized dynamic programming algorithm is only able to correct the axial shifts, and the lateral displacements may still exist. Also, the local registration is easily affected by the global step and it may cause mistakes in the margin. It should be noted that image registration has been proposed as eye-motion correction methods in many OCT applications, which is independent of hardware implementation [36–39]. Especially, image registration like the strip-based registration method is a standard practice in AO-OCT to produce the cell scale resolution OCT images [40,41]. These approaches are aimed at correcting the axial and lateral misalignment of 3D OCT volume data, thus achieving good visualization of the cone mosaic pattern in a small field of view (FOV) [42,43]. The 3D OCT volume data is acquired by raster scan mode at a high A-scan rate, and the B-scans are obtained at different locations. However, in this paper, the input data is acquired by single-line scan mode at a low A-scan rate with large FOV, which means the B-scans should be obtained at the same location, and the eye movement will become apparent.

To correct the apparent axial and lateral shifts rapidly and robustly, in this paper we proposed a modified registration method using hill-climbing algorithm and dynamic programming algorithm. Dynamic programming is known for its efficient computation to find the shortest path in a global problem which is used to correct the axial shift firstly. Hill-climbing algorithm can accurately obtain the optimal position especially for the large registration range. The combination of these two methods can register full alignment with high speed and high accuracy. In addition, if the eye movement/rotation exceeds the registration range threshold, the images may not be aligned well and will blur the edges in the averaging. An algorithm called de-ghosting is proposed to remove the images to ensure the sharpness of boundaries of retina layers. In the de-ghosting process, the blurred shadow areas are recognized by the brightness of the difference image which is obtained by subtracting the unaligned image with the baseline image. An adaptive gamma correction with weighting distribution(AGCWD) method is applied to further improve the contrast of the image. The overall algorithm is automatic and has a good generalization ability for correcting the motion-induced image shifts in single-line HD mode imaging of clinical OCT applications. Compared to other methods (statistical model, low-rank decomposition, deep learning *etc*), our method tries to solve the denoising task from the image itself. So it does not need to consider such problems like insufficient image feature representation, retinal boundary information, lots of training data, and so on. As a light-weighted, fast, robust image processing algorithm, it does not consume much hardware resources, thus having great clinical application potential.

## 2. Methods

### 2.1. Single-line HD mode

The retinal OCT devices usually provide several different scan modes for particular purposes. The scan patterns can be divided into 3 types by the shape: single line, multiple lines, rectangle area. At the same exposure duration, in the single line and multiple lines mode, the scan lines are often repeated to fulfill the total scan time for cross-section images with higher quality.

In the single-line mode, the position and direction of the scan line can be tuned to cross the specific zone guided by the fundus preview image. The scan lines often repeat 20~50 times according to the exposure duration and numbers of A-scans. The averaging of multiple B-scans at the same position can produce high qualified images. This mode is often called as HD mode and benefits to observe suspicious lesions.

In the multiple lines mode, the scan pattern is a number of lines equally spaced in the horizontal, vertical or radial direction. It decreases to repeat 3~5 times at the same location with the increased scan line numbers.

In the area mode, the raster scan often covers a rectangular area containing more than 100 B-scan slices to produce a 3D image. It cannot repeat all the scans because of the time limit. The quality of OCT images cannot be improved by images averaging, but mainly depend on the digital image processing algorithm like filtering, deep learning *etc*.

Each scan mode of retinal OCT devices leads to different clinical applications. In this paper, we demonstrate a way to produce a high-qualified OCT image in single-line HD mode by averaging the denoised B-scans at the same scan position.

### 2.2. Image averaging method

In single-line HD mode, we reconstruct and denoise the B-scans from rawdata like other modes. But an additional method is further proposed to obtain the despeckled image with better visualization of retinal features by averaging B-scans rapidly and accurately. The overall flowchart is shown in Fig. 1.

In the following, we will present the procedures of the image averaging method:

1, determine a template image as a baseline to align the rest images. A good baseline can accelerate processing speed and accuracy. A template image which is most similar to all the B-scans is determined by calculating the cross-correlation coefficient of the B-scans. The template image is used in both axial and lateral shift correction.

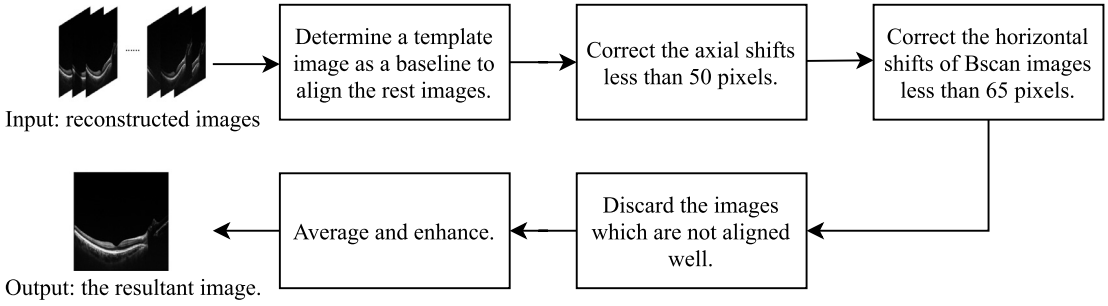


**Fig. 1.** Flowchart of proposed algorithm. Intermediate results are shown in Figs. 3, 4, 5.

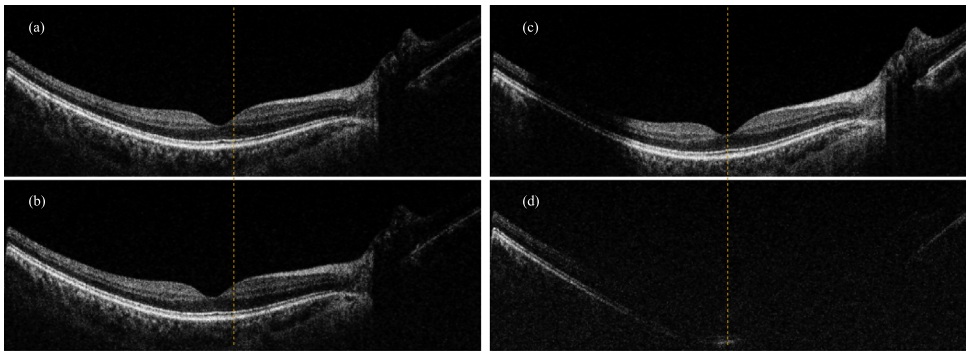


**Fig. 2.** Examples of images in one set of OCT data. (a) The template image selected in the data set. (b) An image with slight eye movement during shooting compared to (a). (c) An image with the retina blocked by eyelashes. (d) An image of missing retinal structure due to blinking.

2, correct the vertical shifts. The vertical shifts are corrected using dynamic programming for axial motion correction. The graph-based method takes neighboring shifts into account while finding the optimal relative shifts of each A-scan, thus making the aligned image very smooth. After that, the horizontal shifts still exist and blur the edge of layers.

3, correct the horizontal shifts of B-scans. This step is the focus of the proposed algorithm. The challenges of horizontal shift correction are the accuracy and speed of finding the peak of the evaluation values in a large range. The hill-climbing algorithm is able to find the optimal shifts within short time by comparing adjacent values in the hill peak.

4, discard the few unaligned images. A registration range threshold is set to −32~32 pixels. To register the images whose motion shifts exceed the range is more difficult and consumes much more calculating time. So a criterion is designed to judge whether the image has the shadow different from the template image. This method ensures the sharpness of retinal layers in the resultant image after averaging;

5, average the registered images. In the averaged image, the areas with lower reflection intensity may be slightly darker by averaging the random noises while the choroid layers are too bright. The retinal layers require careful observation to distinguish, especially in the inner retina area. So the averaged image is further enhanced using adaptive gamma correction with weighting distribution(AGCWD) method to make the layers more discernible.

#### 2.2.1. Determine the template B-scan

Obviously, for a set of number $n$ OCT B-scans obtained in single HD mode, a template image with minimal motion shifts and all retinal structure features is beneficial to accelerate the alignment process. In our lab-built OCT, $n$ can be tuned to 20~50 at different scan frequencies. A "reference image" is made by averaging all the B-scans, and the single B-scan which is most similar to the reference image is selected from the B-scans as the template image. Cross-correlation coefficient is a criterion of similarity. The larger cross-correlation coefficient means more similarity in the two images.

Next, the images lack of retinal structure information should not be selected. Due to eye movement, the retinal structure will shift slightly in most captured images. Compared with Fig. 2(a), the eye moved a little when Fig. 2(b) was imaging. In a few images, the retinal structure may be blocked by eyelashes as shown in Fig. 2(c), or be greatly deformed by squeezing eye muscles involuntarily, or even not be captured with eyes closed as shown in Fig. 2(d). Such B-scans should be excluded before registration for higher efficiency of our algorithm. So 80%(adjustable) B-scans that have higher cross-correlation coefficient to the template image are found. Then the selected images will be registered one by one with the template image.

#### 2.2.2. Vertical alignment: regularized dynamic programming

In this step, each A-scan of the image is aligned by calculating the axial shifts optimally between the current image and the template image. Because of the influences of speckle noise and the vessels on the retina, the shifts leading to the maximum cross-correlation coefficient are not always optimal for correction. To solve this problem, we utilize a graph-based method using regularized dynamic programming algorithm to find the shortest path [33]. Then a continuous and smooth curve of the shifts caused by eye movements is described and the corresponding axial displacements can be compensated well. Fig. 3 shows the averaging results of two images before and after axial alignment. The image in the rectangular box is a magnified view. We can clearly observe that axial shifts are registered well and the vertical displacements are eliminated. The structure between layers is not as blurred as the image above. But by inspecting the sides of the macular fovea, the horizontal shifts (what the short arrow denotes) still obviously exist which are needed to be eliminated.

#### 2.2.3. Horizontal alignment: hill-climbing

The most key step of the method is the alignment of lateral motion in B-scans. Actually, the problem of horizontal registration is the maximum search in the range of lateral shifts. The current image will be shifted by −32 to 32 pixels to the right. The horizontal registration range we set is 65 pixels(about 780um in retina), when the shift value

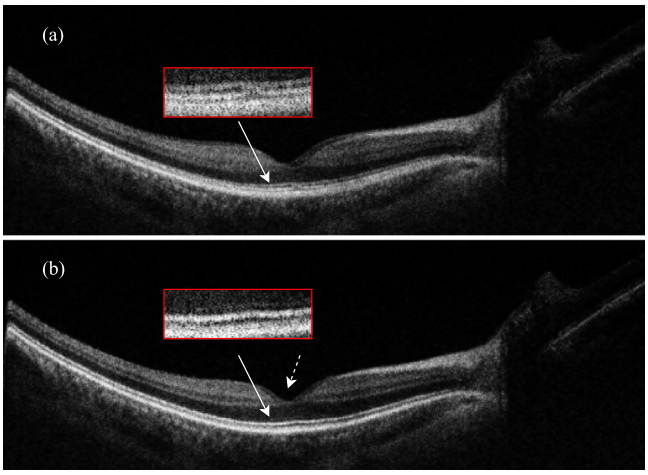


**Fig. 3.** (a) Average of the two unaligned images shown in Figs. 2(a) and 2(b). (b) Average of the two vertically aligned images. The image in the rectangular box is a magnified view of the area pointed by the long arrow. The displacements in the vertical direction are eliminated, but the horizontal shifts (what the dashed arrow denotes) still exist.

is negative, the image is shifted to the left. Vertical shifts need to be aligned after every horizontal shift using the method above to calculate a cross-correlation coefficient accurately. The figure of merit in each lateral shift is shown in Fig. 4(a). As can be seen from the figure, the curve presents a unimodal characteristic in general and the algorithm needs to find the index of the maximum value. Considering the computational time in real applications, it is not efficient to calculate the coefficients of each position to find the maximum value in the range. Hill-climbing method is the typical approach to this case. At first, the initial direction and step length are assigned. After each step, the image is horizontally shifted and vertically aligned, and the cross-correlation coefficient is calculated at the same time. If the evaluation value of the current position is greater than the evaluation value of the previous position, continue to move the step length in the same direction, otherwise reverse the moving direction and reduce the step length until the peak position is found. In Fig. 4(a) we noticed the strong oscillation at the foot of the hill, but the ideal similarity measurement curve should be a parabola. To avoid falling into the local maximum, the initial step length of the hill-climbing method is increased to 16 pixels. Fig. 4(c) shows the average of the two images after horizontal and vertical alignment. Compared with the average of the two vertically aligned images shown in Fig. 4(b), the edges of the macular fovea(white arrows) are very sharp and the retinal layers are also preserved well. The contrast is improved compared to the template image in Fig. 2(a).

We have also tried binary search method to quickly find the optimal shift of the lateral alignment. Take half of the two ends of the registration range and compare these three evaluation values. Pick the largest and second-largest positions and continue to take half until the peak position is found. Compared with hill-climbing method, binary search does not require reentry on both sides of the peak, thus improving the efficiency of the algorithm. But as mentioned above, the farther away from the peak, the more likely to be mistakes. So the binary search method may find a wrong peak position. For example, as shown in Fig. 4(a), the value of −32 shift pixels is greater than the value of 32 shift pixels, and the binary search method will find the peak location in [−32,0] pixel range while the proper peak location is in [0,32] pixel range. To avoid this, we choose the hill-climbing method in this paper.

#### 2.2.4. De-ghost averaging method and enhancement

The algorithm can register the images below the range threshold. But there may be some B-scans out of the range, which are not registered well. For these images with large eye movements, they may show a double/ghost image after averaging. To make the edges sharper, a criterion called de-ghosting is designed to remove the ghost image.

The registered image and the template image are smoothed using median filter at first. Then a difference image is obtained by subtracting the two images. We use the probability density of each gray level in a digital image to judge whether the difference image has a ghost. The probability density function ($\rho$) can be approximated by

$$\rho(l) = n(l)/L \tag{1}$$

where $n(l)$ is the number of pixels at gray level $l$ and $L = M * N$ is the total number of pixels in the image. The value of $l$ is from 0 to 255. The cumulative distribution function ($\eta$) is based on $\rho$, and is formulated as:

$$\eta(l) = \sum_{k=0}^{l} \rho(k). \tag{2}$$

Set the two cumulative distribution function thresholds $t1, t2$. Find the first gray level $v1$ whose $\eta$ is closest to $t1$, and find the second gray level $v2$ whose $\eta$ value is closest to the second parameter $t2$. It means that the probability of pixels with gray levels less than or equal to $v1, v2$ in the difference image is $t1, t2$, respectively. If the image is not aligned well, the unaligned region(where the ghost exists) will be shown in the difference image and the values of pixels in this region will be larger than those in other regions. As values of most pixels in the difference image are small, we choose $0.9 < t1 < 1$, $0.9 < t2 < t1$ to recognize the unaligned area. $V_d$ is defined to judge the probability of ghosting and it is calculated as follows:

$$V_d = v1 - v2. \tag{3}$$

The larger the $V_d$ is, the higher probability of the ghost image will be. If $V_d$ is greater than the threshold $V_t$ we set, we believe that ghosting will occur, and the current image is removed without participating in subsequent averaging as a result. Then we take advantage of the remaining images and a resultant image with high quality is obtained. The flowchart of the de-ghosting algorithm is shown in Fig. 5(e).

The average of the registered images after horizontal and vertical alignment without using the de-ghosting algorithm is shown in Fig. 5(a) and the average of the registered images using the de-ghosting algorithm is shown in Fig. 5(b). The image in the rectangular box is a magnified view of the area denoted by the white arrow. (c) and (d) are the enlarged views of the circled area respectively. It can be observed that something like a shadow appears near the boundary of the retina where the double arrow points. The shadow is what we call ghost. In the resultant image using the de-ghosting algorithm, the ghost is removed and the edges are sharper than the averaged image without using our method. Although the image quality greatly improves after averaging, the areas with lower reflection intensity may be slightly darker by averaging the random noises while the choroid layers are too bright. The distinction of layers in retina is not obvious and the contrast is not satisfying as well.

To make the retinal layers more discernible, adaptive gamma correction with weighting distribution(AGCWD) method is applied for further contrast enhancement [44]. This method utilizes a progressive increment of the original trend to adaptively modify the intensity of the image and effectively avoid the over-enhancement of gamma correction. The result is shown in Fig. 6(a). Fig. 6(b) shows the magnified image of the rectangular box in (a). The contrast between layers is enhanced and we can clearly distinguish nine layers of the retina.

## 3. Experiments and results

### 3.1. Experimental data acquisition

To evaluate the performance of the proposed method, normal and pathological SD-OCT data with only macula and with both macula

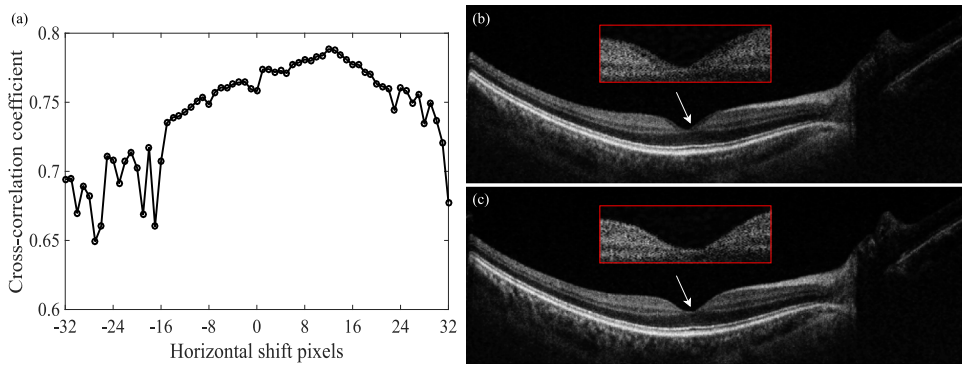


**Fig. 4.** (a) The cross-correlation coefficients of each lateral shift. (b) Average of the two vertically aligned images. (c) Average of the two images after horizontal and vertical alignment. The image in the rectangular box is a magnified view. It can be observed that the two images are registered perfectly.

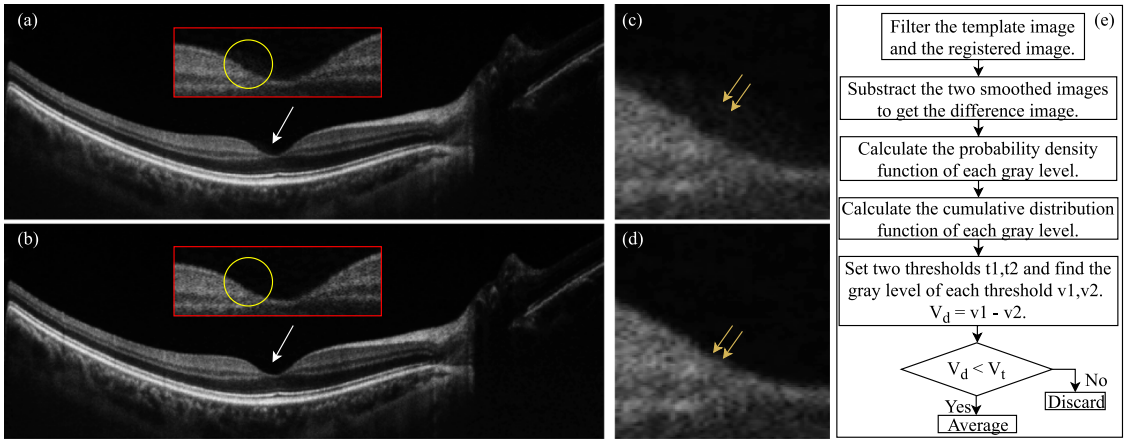


**Fig. 5.** (a) Average of the registered images after horizontal and vertical alignment without using the de-ghosting algorithm. (b) Average of the registered images using the de-ghosting algorithm. The image in the rectangular box is a magnified view of the area denoted by the white arrow. The circled area is the boundary of the retina. (c) and (d) are the enlarged views of the circled area respectively. The area where the double arrow points is the boundary of the retina. There is a shadow next to the boundary in (c), while the edge in (d) is sharper. (e) Flowchart of the de-ghosting averaging algorithm.

and optic nerve head(ONH) were acquired by a classical spectral-domain OCT system. A wideband superluminescent diode with center wavelength of 840 nm and 3 dB bandwidth of 45 nm is used as light source. The scan length is 1000 pixels corresponding to 6 mm with only macula or 12 mm with both macula and ONH. The axial resolution and lateral resolution are 5 μm and 12 μm in our lab-built OCT scanner respectively. The codes were implemented in C++ and tested on a PC with Intel i7-7700 CPU@3.60 GHz and 8 GB of RAM, where eight cores were utilized.

### 3.2. Evaluation metrics

To quantitatively assess the image quality after the method we proposed, signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), peak signal-to-noise ratio (PSNR), structural similarity (SSIM) and edge preservation index(EPI) are used as the evaluation metrics. The metrics are introduced as follows.

SNR is a measure of the noise level in the image [45], defined as:

$$\mathrm{SNR} = \mu_b/\sigma_b \tag{4}$$

where $\mu_b$ and $\sigma_b$ are respectively the mean and the standard deviation of the intensity in the background region.

CNR is a suitable criterion that reflects the contrast between signal and noise regions in the image. It is defined as:

$$\mathrm{CNR} = (1/R)\sum_{r=1}^{R}\left(\frac{\mu_r - \mu_b}{\sqrt{\sigma_r^2 + \sigma_b^2}}\right) \tag{5}$$

where $\mu_r$ and $\sigma_r$ are the mean and the standard deviation respectively of the intensity from a set of $R$ different ROIs that were manually selected in the image.

PSNR is a global quality criterion to estimate noise level, which is computed as [20]:

$$\mathrm{PSNR} = 20 \cdot log_{10}\frac{Ir\_max}{\sqrt{\frac{1}{H}\sum_{h=1}^{H}(Ir\_h - Io\_h)^2}} \tag{6}$$

where $Ir\_h$ represents the intensity of the $h$th pixel in the resultant image after our method, $Io\_h$ is the same $h$th pixel of the template image, $H$ is the total number of the pixels, and $Ir\_max$ is the maximum intensity value of the resultant image.

SSIM is also a global criterion that reflects the structural similarity of the denoised image and the origin image [46], defined as:

$$\mathrm{SSIM} = \frac{(2\mu_{Ir}\mu_{Io} + c1)(2\sigma_{IrIo} + c2)}{(\mu_{Ir}^2 + \mu_{Io}^2 + c1)(\sigma_{Ir}^2 + \sigma_{Io}^2 + c2)} \tag{7}$$

where $\mu_{Ir}$, $\mu_{Io}$ and $\sigma_{Ir}$, $\sigma_{Io}$ are the mean and standard deviation of the resultant image and the template image respectively, $\sigma_{IrIo}$ is their covariance, $c1$ and $c2$ are used to avoid dividing by zero.

EPI is the quantitative metric which indicates the ability of maintaining details of edge in the denoised image [31]. Due to the layered structure of the retina, we only focus on EPI in the longitudinal direction and it is calculated as:

$$\mathrm{EPI} = \frac{\sum_i\sum_j |Ir(i+1,j) - Ir(i,j)|}{\sum_i\sum_j |Io(i+1,j) - Io(i,j)|} \tag{8}$$

where $Ir$ and $Io$ represent the resultant image and the template image, while $i$ and $j$ are coordinates in the longitudinal and lateral direction

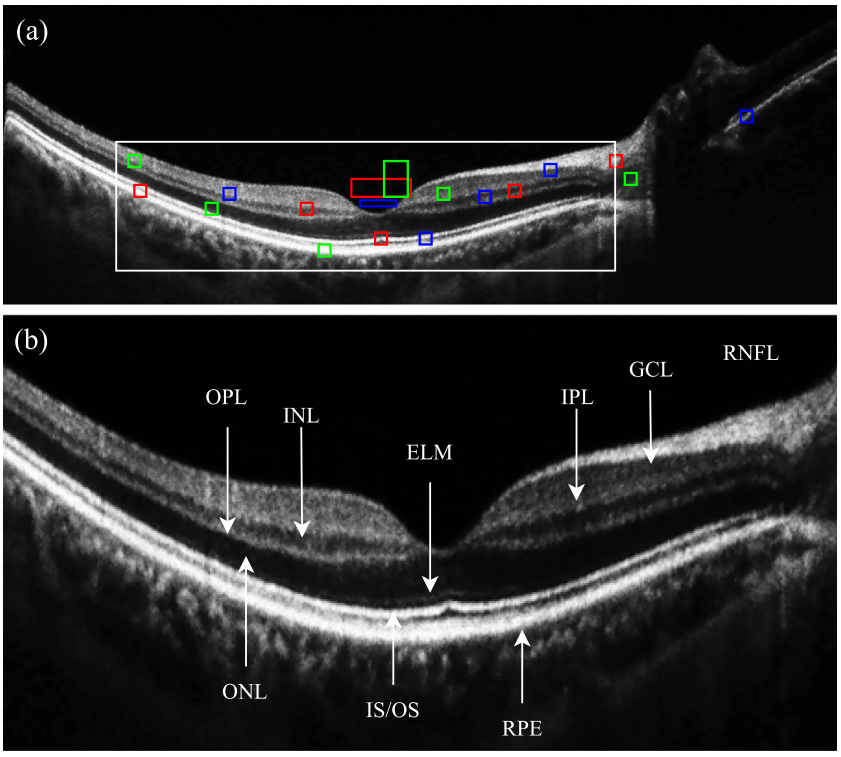


**Fig. 6.** (a) The resultant image in one set of 20 B-scans obtained in single-line HD mode after the proposed method. (b) The enlarged region from the area in the rectangular box in (a). The edges of retinal layerings are preserved well. NFL: nerve fiber layer; GCL: ganglion cell layer; IPL: inner plexiform layer; INL: inner nuclear layer; OPL: outer plexiform layer; ONL: outer nuclear layer; ELM: external limiting membrane; IS/OS: inner and outer segment junctions of the photo-receptor; and RPE: retinal pigmented epithelium.

in the image. In our experiments, we select three bands with height of 7 pixels [31] centered at three boundaries (ILM, INL and RPE) respectively as shown in Fig. 8(a).

### 3.3. Performance of the proposed method

SNR and CNR are typical indicators for judging image quality. To assess the performance of each processing step and the performance of different numbers of B-scans in each set generally, these two evaluation metrics are sufficient. We manually selected 3 groups of ROIs denoted as red, green, blue rectangles which covered all retinal layers, macula and ONH. Each group has five small rectangles with 20 × 20 pixels in the retinal region and one large rectangle in the noisy background region as shown in Fig. 6(a).

A set of 20 B-scans obtained in single-line HD mode is used to test the SNR and CNR performance of each step, shown in Fig. 7(a)(b). The numbers on the $x$-axis indicate the processing steps in our algorithm. It can be seen that SNR and CNR are improved in each step. The vertical alignment makes a slight improvement in SNR, but a significant change (155% improvement) is made after the horizontal adjustment in step2. The CNR is improved by 5% after step1 but basically has no growth in step2. After de-ghosting averaging in step3, the SNR is roughly doubled and CNR is improved from 2 to 2.8. The AGCWD method may reduce SNR from 4.5 to 4 in some regions, partly because the method may also enhance the background noise, resulting in an increase of standard deviation in the background part. But it has great effects on CNR (18% improvement) to the image. The reason for the large rise of ROI2 in step4 is that there are more brighter areas in ROI2, and the enhancement of these areas is more obvious. With respect to the average of two images before registration, it improved the SNR, CNR by around 300% and 65%.

To test the performance in SNR and CNR as functions of the number of averaged images, one set of 50 B-scans is employed. The exposure time of imaging 50 B-scans in a set is the same as imaging 20 B-scans. It can be seen from Fig. 7(c) that SNR increases with the number of averaged images($N$). And the curve is similar to the square root of N and exceeds after 5 images are averaged. Fig. 7(d) presents the CNR improvement. When the number of averaged images is around 15, the CNR improvement can achieve a factor of around 9. The curve becomes flat after 15 images are averaged. This is also why the number of B-scans in single-line HD mode of our lab-built OCT is set to 20.

### 3.4. Comparison to different registration methods

Three sets of data were randomly selected to compare the proposed method with state-of-the-art approaches for image registration guided speckle suppression in OCT, including affine model [34], affine + dynamic programming(DP) [35], DP + binary search(BS) and DP + hill-climbing(HC). Fig. 8 shows the registration guided despeckling results of three sets of 20 OCT retinal images(data 1 in the first row shows the normal retinal image with macula and ONH, data 2 shows the normal retinal image with macula only, data 3 shows the pathological retinal image with macula and ONH). The average performance metrics for all test data are listed in Table 1. The results of the proposed method are higher than the methods compared in SNR, PSNR, SSIM and EPI, which is compliant with the better visual quality.

In the affine method [34], although the calculation time is shorter, the images are not well registered in details. Therefore, a blurred average image is calculated shown in Fig. 8(b). Especially, the thin layer above the RPE complex, known as external limiting membrane(ELM) cannot be viewed clearly. The most evaluation metrics are relatively lower except SNR and CNR because of more averaging images.

After the global and local registration [35], the edges in the averaged image are sharper. But regularized dynamic programming method is only able to correct the axial shifts, the local lateral shifts still exist. Also, after the global alignment using affine model, the margin of the image will be missing, thus causing mistakes in local registration shown in the second row of Fig. 8(c). The evaluation values are higher than

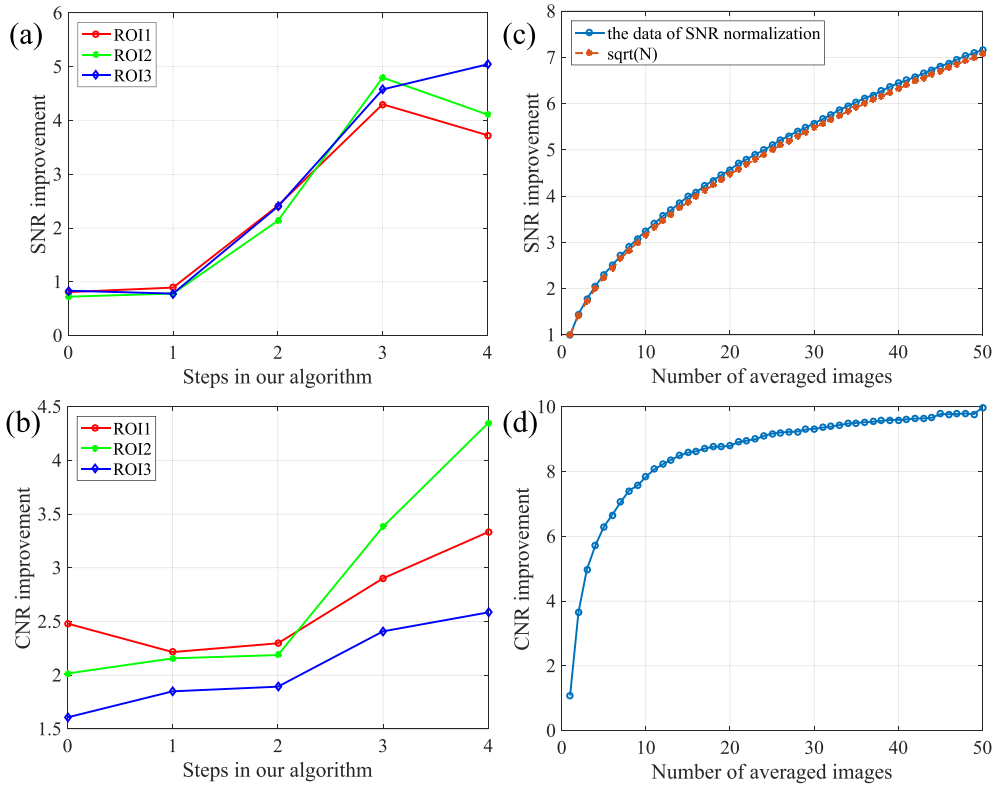


**Fig. 7.** (a)(b) Improvements in signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) as functions of the steps in our algorithm. 0: Average of two images before registration. 1: Average of two images after vertical alignment. 2: Average of two images after horizontal and vertical alignment. 3: Average of multiple images after the de-ghosting method. 4: Enhanced image after the AGCWD method. (c)(d) Improvements in signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) as functions of the number of averaged images.

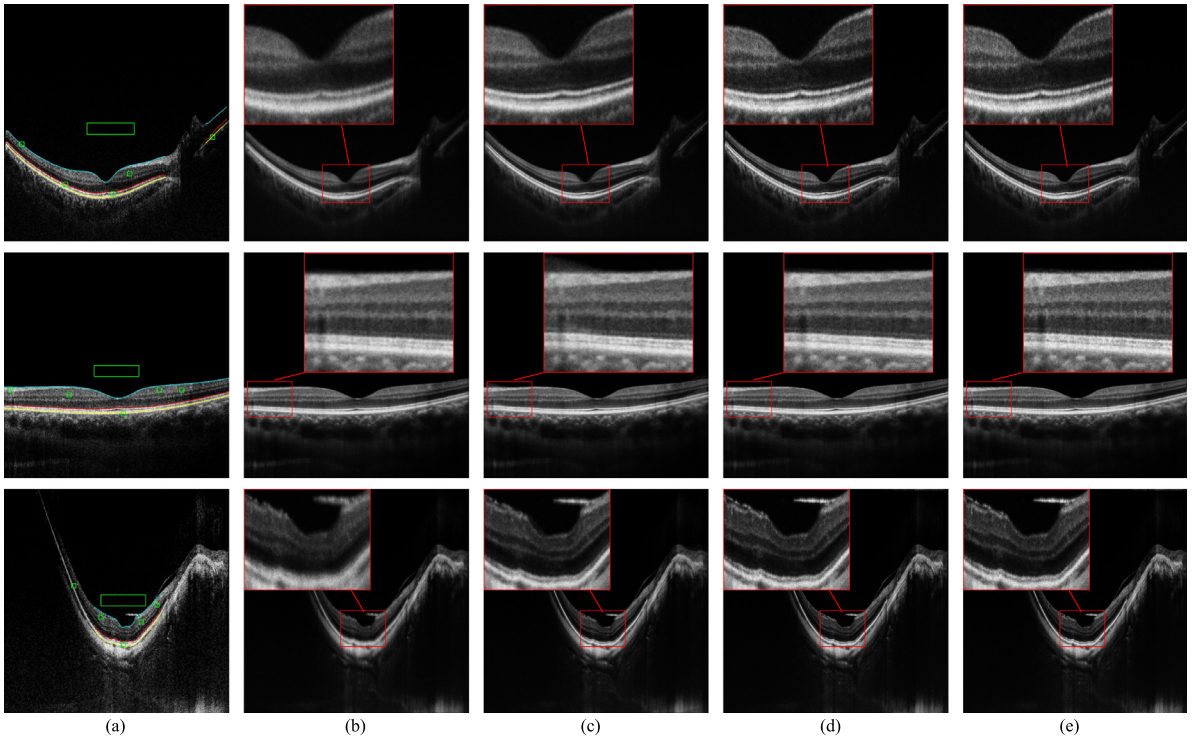


**Fig. 8.** The results of three sets of 20 OCT images. (a) The template image of each set. (b) Results using affine model. (c) Results using affine + DP. (d) Results using DP + BS. (e) Results using DP + HC(proposed). In (b), the images are not well registered in details, thus making the boundaries blurred. In (c), the mistakes of local registration are shown in the second row. In (d)(e), it is difficult to see the differences between these images at a glance. But with careful observation, the boundaries are slightly sharper in (e). The proposed method achieved the best human visual perception.

**Table 1**
Evaluation metrics in average for different methods.

| Methods | Time (s) | SNR | CNR | PSNR (dB) | SSIM | EPI |
|---|---|---|---|---|---|---|
| Affine [34] | **0.41** | 2.56 | **4.29** | 24.96 | 0.914 | 0.754 |
| Affine + DP [35] | 1.77 | 2.55 | 3.76 | 25.28 | 0.921 | 0.881 |
| DP + BS | 8.71 | 2.51 | 3.37 | 25.43 | 0.926 | 0.926 |
| DP + HC(Proposed) | 14.9 | **2.57** | 3.35 | **25.46** | **0.927** | **0.942** |

the results of the affine method but not the highest in PSNR, SSIM and EPI. The sharpness of edges and boundaries is not satisfying.

In Fig. 8(d)(e), it is difficult to see the differences in the resultant images of the DP + BS method and DP + HC method unless carefully inspected. But in facts, as mentioned in 2.2.3, there are fewer mistakes in DP + HC method, the edges are sharper and the metrics of DP + HC are slightly higher than the metrics of DP + BS in the performance of edge-preserving and structural similarity. We remove several images with large eye movement which exceeds the registration range we set, thus resulting in the lower metrics in CNR. But edges of the image become sharper and details like vessels or smaller lesions can be better observed in the resultant image of our method. In conclusion, combining both subjective and objective evaluation criteria, the proposed DP + HC method can obtain best results among the methods compared.

## 4. Conclusions and discussions

In this work, we propose a registration guided speckle suppression method in single-line HD mode for ophthalmology OCT images. The method well preserved the detailed structure information and achieved further enhanced retinal layers visualization in the resultant image. The method has a good generalization ability for different retinal OCT data and can be applied in commercial scanners. Also, it can be extended to registration and enhancement in other OCT data, such as dentistry or dermatology which suffer the unintentional movement of teeth or the vibration of skin. However, the images with deformed retina or with large rotation are still a big challenge, which are discarded in this algorithm. The current code is based on CPU computation which may be accelerated based on GPU computation in the next.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding

This work was supported in part by the National Key R&D Program of China under Grant 2018YFA0701700, the National Basic Research Program of China (973 Program) under Grant 2014CB748600, and in part by the National Natural Science Foundation of China (NSFC) under Grant 61622114.

## References

[1] D. Huang, et al., Optical coherence tomography, Science 254 (5035) (1991) 1178.
[2] M. Wojtkowski, High-speed optical coherence tomography: Basics and applications, Appl. Opt. 49 (16) (2010) D30–61.
[3] M.R. Hee, J. Izatt, E. Swanson, D. Huang, J.G. Fujimoto, Optical coherence tomography of the human retina, Arch Ophthalmol. 113 (3) (1995) 325–332.
[4] G.J. Jaffe, J. Caprioli, Optical coherence tomography to detect and manage retinal disease and glaucoma, Am. J. Ophthalmol. 137 (1) (2004) 156–169.
[5] J.G. Fujimoto, W. Drexler, J.S. Schuman, C.K. Hitzenberger, Optical coherence tomography (oct) in ophthalmology: Introduction, Opt. Express 17 (5) (2009) 3978–3979.
[6] M. Samieinasab, Z. Amini, H. Rabbani, Multivariate statistical modeling of retinal optical coherence tomography, IEEE Trans. Med. Imaging PP (99) (2020) 1–1.
[7] B. Qiu, Z. Huang, X. Liu, X. Meng, Y. You, G. Liu, K. Yang, A. Maier, Q. Ren, Y. Lu, Noise reduction in optical coherence tomography images using a deep neural network with perceptually-sensitive loss function, Biomed. Opt. Express 11 (2) (2020) 817–830.
[8] M. Shamouilian, I. Selesnick, Total variation denoising for optical coherence tomography, in: 2019 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), 2019, pp. 1–5.
[9] J.M. Schmitt, S.H. Xiang, K.M. Yung, Speckle in optical coherence tomography, J. Biomed. Opt. 4 (1) (1999) 95.
[10] N. Iftimia, B.E. Bouma, S. Yun, G. Tearney, Analysis of speckle reduction in optical coherence tomography by path length encoded angular compounding, in: Biomedical Topical Meeting, Optical Society of America, 2004, p. FH30.
[11] H. Wang, A.M. Rollins, Speckle reduction in optical coherence tomography using angular compounding by b-scan doppler-shift encoding, J. Biomed. Opt. 14 (3) (2009) 030512.
[12] M. Pircher, E. Gotzinger, R. Leitgeb, A.F. Fercher, C.K. Hitzenberger, Speckle reduction in optical coherence tomography by frequency compounding, J. Biomed. Opt. 8 (3) (2003) 565.
[13] R.J. Zawadzki, B. Cense, Y. Zhang, S.S. Choi, J.S. Werner, Ultrahigh-resolution optical coherence tomography with monochromatic and chromatic aberration correction, Opt. Express 16 (11) (2008) 8126–8143.
[14] B.F. Kennedy, T.R. Hillman, A. Curatolo, D.D. Sampson, Speckle reduction in optical coherence tomography by strain compounding, Opt. Lett. 35 (14) (2010) 2445–2447.
[15] M. Pircher, R.J. Zawadzki, Review of adaptive optics oct (ao-oct): principles and applications for retinal imaging, Biomed. Opt. Express 8 (5) (2017) 2536–2562.
[16] A. Ozcan, A. Bilenca, A.E. Desjardins, B.E. Bouma, G.J. Tearney, Speckle reduction in optical coherence tomography images using digital filtering, J. Opt. Soc. Amer. A 24 (7) (2007) 1901–1910.
[17] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: International Conference on Computer Vision, 1998.
[18] H.M. Salinas, D.C. Fernandez, Comparison of pde-based nonlinear diffusion approaches for image enhancement and denoising in optical coherence tomography, IEEE Trans. Med. Imaging 26 (6) (2007) p.761–771.
[19] M.A. Mayer, A. Borsdorf, M. Wagner, J. Hornegger, C.Y. Mardin, R.P. Tornow, Wavelet denoising of multiframe optical coherence tomography data, Biomed. Opt. Express 3 (3) (2012) 572–589.
[20] L. Fang, S. Li, R.P. McNabb, Q. Nie, A.N. Kuo, C.A. Toth, J.A. Izatt, S. Farsiu, Fast acquisition and reconstruction of optical coherence tomography images via sparse representation, IEEE Trans. Med. Imaging 32 (11) (2013) 2034–2049.
[21] Z. Jian, L. Yu, B. Rao, B.J. Tromberg, Z. Chen, Three-dimensional speckle suppression in optical coherence tomography based on the curvelet transform - escholarship, Opt. Express 18 (2) (2010) 1024–1032.
[22] M. Esmaeili, A.M. Dehnavi, F. Hajizadeh, H. Rabbani, Three-dimensional curvelet-based dictionary learning for speckle noise removal of optical coherence tomography, Biomed. Opt. Express 11 (2) (2020) 586–608.
[23] L. Fang, S. Li, D. Cunefare, S. Farsiu, Segmentation based sparse reconstruction of optical coherence tomography images, IEEE Trans. Med. Imaging 36 (2) (2017) 407–421.
[24] D.A. Jesus, D.R. Iskander, Assessment of corneal properties based on statistical modeling of oct speckle, Biomed. Opt. Express 8 (1) (2017) 162–176.
[25] M. Li, R. Idoughi, B. Choudhury, W. Heidrich, Statistical model for oct image denoising, Biomed. Opt. Express 8 (9) (2017) 3903–3917.
[26] I. Kopriva, F. Shi, X. Chen, Enhanced low-rank + sparsity decomposition for speckle reduction in optical coherence tomography, J. Biomed. Opt. 21 (7) (2016) 1–9.
[27] J. Cheng, D. Tao, Y. Quan, D.W.K. Wong, G.C.M. Cheung, M. Akiba, J. Liu, Speckle reduction in 3d optical coherence tomography of retina by A-scan reconstruction, IEEE Trans. Med. Imaging 35 (10) (2016) 2270–2279.
[28] X. Mao, C. Shen, Y. Yang, Image denoising using very deep fully convolutional encoder-decoder networks with symmetric skip connections, CoRR, abs/1603.09056, 2016.
[29] K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising, IEEE Trans. Image Process. 26 (7) (2017) 3142–3155.
[30] A. Abbasi, A. Monadjemi, L. Fang, H. Rabbani, Y. Zhang, Three-dimensional optical coherence tomography image denoising through multi-input fully-convolutional networks, 108, 2019, pp. 1–8.
[31] Y. Ma, X. Chen, W. Zhu, X. Cheng, D. Xiang, F. Shi, Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN, Biomed. Opt. Express 9 (2018) 5129.
[32] Y. Guo, K. Wang, S. Yang, Y. Wang, P. Gao, G. Xie, C. Lv, B. Lv, Structure-Aware noise reduction generative adversarial network for optical coherence tomography image, in: H. Fu, M.K. Garvin, T. MacGillivray, Y. Xu, Y. Zheng (Eds.), Ophthalmic Medical Image Analysis, Springer International Publishing, pp. 9–17.

[33] T.M. Jorgensen, J. Thomadsen, U. Christensen, W. Soliman, B. Sander, Enhancing the signal-to-noise ratio in ophthalmic optical coherence tomography by image registration—method and clinical examples, J. Biomed. Opt. 12 (4) (2007) 041208.
[34] D. Alonso-Caneiro, S.A. Read, M.J. Collins, Speckle reduction in optical coherence tomography imaging by affine-motion image registration, J. Biomed. Opt. 16 (11) (2011) p.117027.1–117027.5.
[35] H. Zhang, Z. Li, X. Wang, X. Zhang, Speckle reduction in optical coherence tomography by two-step image registration, J. Biomed. Opt. 20 (3) (2015) 036013.
[36] M. Nasiriavanaki, S. Adabi, Z. Turani, E. Fatemizadeh, A. Clayton, Optical coherence tomography technology and quality improvement methods for optical coherence tomography images of skin: A short review, Biomed. Eng. Comput. Biol. 8 (2017).
[37] M. Azimipour, J.V. Migacz, R.J. Zawadzki, J.S. Werner, R.S. Jonnal, Functional retinal imaging using adaptive optics swept-source OCT at 1.6 MHz, Optica 6 (3) (2019) 300–303.
[38] M. Azimipour, R. Jonnal, J. Werner, R. Zawadzki, Coextensive synchronized slo-oct with adaptive optics for human retinal imaging, Opt. Lett. 44 (2019) 4219.
[39] A. Camino, P. Zang, A. Athwal, S. Ni, Y. Jian, Sensorless adaptive-optics optical coherence tomographic angiography, Biomed. Opt. Express 11 (7) (2020) 3952–3967.
[40] A. Baghaie, Z. Yu, R.M. D'Souza, Involuntary eye motion correction in retinal optical coherence tomography: Hardware or software solution?, Med. Image Anal. 37 (2017) 129–145.
[41] M.F. Shirazi, E. Brunner, M. Laslandes, A. Pollreisz, C.K. Hitzenberger, M. Pircher, Visualizing human photoreceptor and retinal pigment epithelium cell mosaics in a single volume scan over an extended field of view with adaptive optics optical coherence tomography, Biomed. Opt. Express 11 (8) (2020) 4520–4535.
[42] S.B. Stevenson, A. Roorda, Correcting for miniature eye movements in high-resolution scanning laser ophthalmoscopy, in: F. Manns, P.G. Söderberg, A. Ho, B.E. Stuck, M.B. M.D (Eds.), Ophthalmic Technologies XV, 5688, SPIE, International Society for Optics and Photonics, 2005, pp. 145–151.
[43] H.C. Hendargo, R. Estrada, S.J. Chiu, C. Tomasi, S. Farsiu, J.A. Izatt, Automated non-rigid registration and mosaicing for robust imaging of distinct retinal capillary beds using speckle variance optical coherence tomography, Biomed. Opt. Express 4 (6) (2013) 803–821.
[44] S.C. Huang, F.C. Cheng, Y.S. Chiu, Efficient contrast enhancement using adaptive gamma correction with weighting distribution, IEEE Trans. Image Process. A Publication of the IEEE Signal Processing Society 22 (3) (2013) 1032–1041.
[45] Y. Watanabe, H. Hasegawa, S. Maeno, Angular high-speed massively parallel detection spectral-domain optical coherence tomography for speckle reduction, J. Biomed. Opt. 16 (6) (2011) p.060504.1–060504.3.
[46] Zhou Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.