



Cluster-based filtering framework for removing speckles with structural protection in OCT images

Shengjian Huang¹ · Chen Tang¹ · Min Xu¹ · Zhenkun Lei²

Received: 10 August 2020 / Accepted: 20 August 2021 / Published online: 16 October 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Generally speaking, it is the essential core of image filtering to keep the texture features better while denoising the image. To some extent, optical coherence tomography retina images have speckle noise, which masks the texture features of the image, and thus causes misjudgment to the doctor's diagnosis. In this paper, we first propose a cluster-based filtering framework for removing speckles with structural protection in OCT images. The overall process can be divided into preprocessing, structure extraction and structure denoising. First, in the preprocessing stage, we propose to use the shearlet (SHT) method for preliminary filtering and combine block search and matching to achieve structure protection. Then in the structure extraction stage, we propose to use the relative total variation algorithm to achieve structure extraction, combined with fuzzy C-means Clustering filters out the background noise to obtain the structure mask of the image. Finally, in the structure denoising stage, we propose a new variational Block matching 3D (BM3D)- L^2 method, and the structure of the image and the noise are described in BM3D space and L^2 space, respectively. By assigning appropriate values to the parameters, image noise can be better eliminated, and the structural texture of the image can be protected. We test the proposed method on seven large noisy OCT images, which include five human retinal OCT images and two mouse optic nerve OCT images. In addition, we also compare it with SHT, BM3D, TV-SHT and TV-BM3D methods, which were proved to be effective in denoising. The performances of these methods are quantitatively evaluated in terms of the signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR) and the averaged equivalent number of looks (ENL) at the aspects of speckle reduction and structure texture protection. Vast experiments show that our proposed method can effectively reduce speckle noise in OCT images, protect important structural information and improve image quality. Here, we believe that our method will improve image segmentation, medical diagnosis, and can use this as training samples to improve the accuracy of machine learning.

1 Introduction

Optical coherence tomography (OCT) is an emerging tomography technology, which is based on the principle of optical interference to observe the tissue information inside the organism. In this process, as a non-contact, non-invasive, non-ionizing imaging technology, OCT has been widely used in biomedical imaging and clinical with its advantages of high safety, fast scanning speed and high resolution diagnostic area [1–3]. However, due to factors, such as the

wavelength of the imaging beam, the mutual interference of scattered light, and the differences between different structures of the image, speckle has become an inherent phenomenon of the OCT measurement system [4]. In addition, speckles can also lead to a decrease in the contrast of each part of the medical image, which brings certain difficulties to the doctor's diagnosis. Therefore, reducing speckle noise in OCT images, enhancing contrast in abnormal areas, and protecting image detail information are objects that researchers have paid close attention to in recent years. To improve the image quality, in the past three decades, many improvements have been proposed in terms of system design and post-processing.

From the perspective of system design [5–10], the noise reduction performance of the OCT image can be directly and effectively achieved, but the premise of achieving this performance is that the OCT system needs to be continuously modified, such as adjusting the light source and adjusting

✉ Chen Tang
tangchen@tju.edu.cn

¹ School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, People's Republic of China

² State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian 116024, People's Republic of China

the scanning angle. As a result, a lot of time is consumed and there will be speckle noise in the obtained OCT image.

Based on the image post-processing, the idea with speckle removal has been investigated in the past few years, and numerous different approaches have been proposed in the past several years. Generally speaking, these algorithms are classified as the spatial domain method and the transform domain method.

In the spatial domain algorithm, numerous and diverse methods have been developed, such as the adaptive median filtering [11], enhanced Lee filtering [12], adaptive Wiener filtering [13], Bayesian least squares estimation [14], the diffusion-based algorithms (PDEs) [15–17] and the nonlocal filters-based algorithms [18–21].

Particularly, the total variation (TV) model which was first proposed by Rudin et al. [22], is of great interest and probably the most successful one because of its simple form and decent edge preserving performance. In addition, many modified TV de-speckling methods [23–26], with different tuning functions, have been proposed gradually. However, the fixed (spatially invariant) regularization parameter used in TV models still limits the improvement of speckle reduction. In fact, the “staircase” effect often appears in TV and PDE de-speckling results. This is because the spatial information on signal and noise cannot be adequately reflected by these methods.

In addition to the spatial domain algorithm, the transform domain algorithm is another important de-speckling strategy. This strategy decomposes the original image across multiple scales with a multiscale transform and reduces the speckle noise with a shrinkage technique to thresh the obtained coefficients. In other words, their de-speckling performance can be enhanced by choosing a proper decomposition basis and applying an effective shrinkage technique. In the transform domain, there are mainly the wavelet transform [27], the adaptive wavelet transform [28], the SHT transform [29], the contourlet transform [30], the curvelet filtering [31] and block matching and 3D (BM3D) transform [32]. Relatively speaking, the methods realize better speckle reduction and more effective structure preservation. However, some unwanted wavelet-related visual artefacts are introduced into the results. In addition, these artefacts often degrade the image quality especially in the textural regions.

Generally speaking, the spatial domain denoising method has the advantages of directness, flexibility, and ease of implementation, while the transform domain denoising method can sparsely represent images in the frequency domain. It can be seen from the above that the two have unique advantages in reducing image speckle and structural protection in the spatial and transform domains, respectively. By combining these algorithms and making full use of the advantages of each algorithm, the speckle removal performance can be further improved. For example, Ma et al.

reduced pseudo Gibbs artifacts using TV-based methods [33, 34]. In addition, Xu et al. studied the SHT-based TV (TV-SHT) method and Huang et al.’s BM3D-based TV method [35, 36]. These methods usually need to select and adjust specific parameters to constantly balance the reduction of artifacts, the suppression of speckle noise, and the retention of structural texture. However, in the case of high speckle, even with these methods, pseudo Gibbs artifacts are still inevitable. Recently, Hossein et al. [45] propose to use the K-means clustering method to cluster images, and then pass three pixel-based de-speckle algorithms, including Lee filter, adaptive Wiener filter and mean filter method for filtering to obtain the good results. Hu et al. [44] proposed the retinex OCT image enhancement algorithm based on the clustering method, which selectively enhances the structural part, and uses a three-dimensional block matching algorithm to filter it, which further reduces the speckle noise and improves the image quality. It is worth mentioning that literature [44] uses FCM to divide the image into background and structure parts. However, this algorithm is based on feature calculation and clustering. When the noise is large, the clustering result of the image will be limited. Therefore, pre-processing operations are required in this paper. In addition, machine learning and deep learning methods [37–39] have provided many new ideas for speckle reduction. To some extent, these algorithms can quickly remove noise through the set network architecture. In this process, the training set plays a very important role, which directly determines the quality of the prediction results. However, the training set of these algorithms usually contains more background noise, so that the final predicted image still has noise, and even the structure information of the image is blurred. To a certain extent, the accuracy and adaptability of these methods will be reduced.

Inspired by this, to eliminate the influence of background large noise on image structure information during the denoising process. In this paper, we first propose a cluster-based filtering framework for removing speckles with structural protection in OCT images. The overall process can be divided into preprocessing, structure extraction and structure denoising. First, in the preprocessing stage, we propose to use the SHT method for preliminary filtering and combine block search and matching to achieve structure protection. Then in the structure extraction stage, we propose to use the relative total variation algorithm to achieve structure extraction, combined with fuzzy C-means Clustering filters out the background noise to obtain the structure mask of the image. Finally, in the structure denoising stage, we propose a new variational BM3D- L^2 method, and the structure of the image and the noise are described in BM3D space and L^2 space, respectively. By assigning appropriate values to the parameters, image noise can be better eliminated, and the structural texture of the image can be protected. With this method, seven large noisy OCT images are successfully

de-speckled, in which human retina OCT images with the different disease information and two mouse normal optic nerve OCT images are selected, respectively, in this paper. In addition, we also use the three quantitative indicators of SNR, CNR and ENL to evaluate the ability of background noise reduction, structural noise removal and details information smoothing protection. In addition, we compared the proposed method with SHT, BM3D, TV-SHT and TV-BM3D methods. The experimental results of seven large noisy OCT images show that the proposed method is superior to the above three methods, and effectively improves the quality of the original OCT noise image.

The remainder of this paper is organized as follows: Sect. 2 reviews the theory of SHT transform, block matching 3D and relative total variation. Besides, we details the proposed method. Section 3 gives the results of different methods and analyzes them. Finally, the conclusion is given in Sect. 4.

2 The description of related methods

2.1 The brief reviews of the related methods

2.1.1 SHT transform

The SHT space is a special synthetic wavelet with strong directional sensitivity that provides the best representation of the image's geometric features in the direction and shape of the image. Recently, this transform has been applied to image denoising.

Given an image u , the forward and backward continuous 2D SHT transform is defined as

$$\text{SH}_\psi u(a, s, t) = \langle u, \psi_{a,s,t} \rangle, \quad (1)$$

$$u = \int_{R^2} \int_{-\infty}^{\infty} \int_0^{\infty} \langle u, \psi_{a,s,t} \rangle \psi_{a,s,t} \frac{da}{a^3} ds dt, \quad (2)$$

where $\psi_{a,s,t} = |\det M_{a,s}|^{-\frac{1}{2}} \psi(M_{a,s}^{-1}x - t)$
 $\left(M_{a,s} = B_s A_s, A_s = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, B_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \right)$.

As shown in the above equation, $\langle u, \psi_{a,s,t} \rangle$ are the SHT transform coefficients of u , the matrix $M_{a,s}$ contains two different operations that can be decomposed into the product of the shear matrix B_s and the anisotropic expansion matrix A_s . The parabolic scaling matrix A_s is an anisotropic dilation, and the geometric shear matrix B_s parameterizes the orientations using the variable s . $\psi_{a,s,t}$ is a set of local functions with scales, directions, and positional parameters of a, s, t , respectively. The SHT space can capture the change of the image, such as the edge and the contour, by the reduction of

the scale parameter a . The frequency-domain shear structure of shear waves is shown in Fig. 1. It is worth mentioning that when the value of a is small, the smaller the cut area; therefore, it can capture the frequency information of the image well.

2.1.2 Block matching 3D (BM3D)

Block matching 3D (BM3D) [32] is one of the hotspot algorithms used in OCT image denoising in recent years. Its main idea is to aggregate similar blocks in the image, and then denoise through three-dimensional wavelet shrinkage. While this algorithm can sparsely express the image, the most important thing is to consider the pixel block of the image, which can retain the detailed information of the image and improve the visual quality of the image. Its flowchart can be seen in Fig. 2.

The algorithm principle of BM3D mainly includes two parts. One part is to use three-dimensional transformation combined with wavelet hard threshold shrinkage coefficient to estimate the initial filtered image, and the other part is based on the result of the above operation, using three-dimensional transformation combined with Wiener filtering to shrinkage coefficient. Finally, the algorithm output the filtered image.

2.1.3 Relative total variation (RTV)

Relative total variation [40] is an extension of the total variation algorithm, which has good structure extraction performance. It combines the total variation of pixels in the window with the inherent change of the entire space (especially for the visually significant area) to make meaningful content in the image as well as the texture edges are adjusted appropriately. Let R and S be the input image and the result image, respectively. The effect of the term $(R_p - S_p)$ is that there is no significant deviation between the input and the result. The objective function is finally expressed as

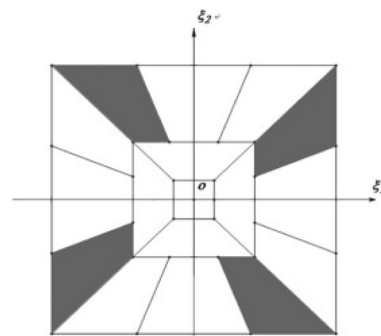


Fig. 1 SHT support tiling in frequency domain

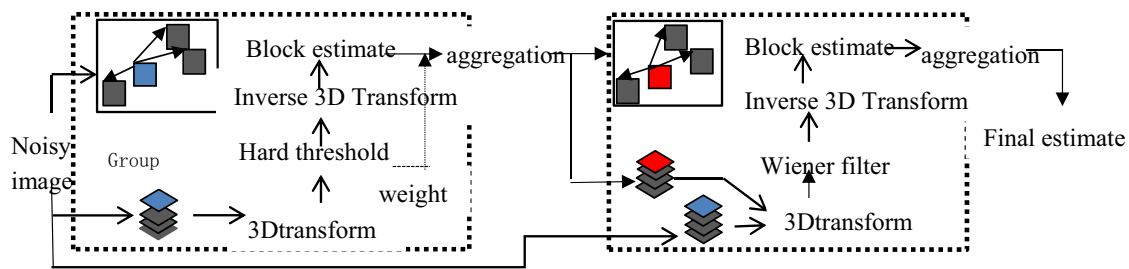


Fig. 2 Block matching 3D (BM3D) operation flowchart

$$\arg \min_s \sum_n (R_n - S_n)^2 + \lambda \cdot \left(\frac{l_x(n)}{\zeta_x(n) + \epsilon} + \frac{l_y(n)}{\zeta_y(n) + \epsilon} \right), \tag{3}$$

where the value of λ typically varies in a small range [0.01,0.03] in practice. The ϵ is a small positive number to avoid division by zero. The general pixelwise windowed total variation measure, written as

$$l_x(n) = \sum_{m \in R(n)} g_{n,m} \cdot |(\partial_x R)_m|, \tag{4}$$

$$l_y(n) = \sum_{m \in R(n)} g_{n,m} \cdot |(\partial_y R)_m|, \tag{5}$$

$$g_{n,m} \propto \exp\left(-\frac{(x_n - x_m)^2 + (y_n - y_m)^2}{2w^2}\right), \tag{6}$$

where m belongs to the rectangular region centered at pixel n . $l_x(n)$ and $l_y(n)$, respectively, are the local weighted total variation of pixel n in the x and y directions. $g_{n,m}$ is a weighting function defined according to spatial affinity, and the w control window size. The specific solution process can be seen in Ref. [40].

2.2 The proposed cluster-based filtering framework

By reviewing the above algorithms, we propose a new three-stage filtering algorithm, including filtering, the structure

extraction and the structure denoising. The implementation steps are described as follows:

2.2.1 Preprocessing

First, a forward SHT transform is performed on the input image $I(x, y)$, to obtain the SHT transform coefficients CSH_I :

$$CSH_I = \text{SHT}(I(x, y)), \tag{7}$$

where $I(x, y)$ represents original input image, CSH_I represents SHT transform coefficients, and $\text{SHT}(\cdot)$ represents SHT transform (Fig. 3).

Then the transform coefficients subject to a hard threshold before an inverse SHT transform is carried out:

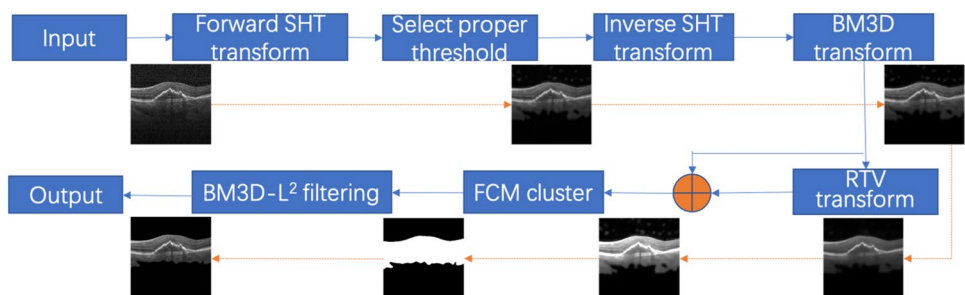
$$CSH_I(i) = \text{shrink}(CSH_I(i)), \quad i = 0, 1, 2, 3, 4, \tag{8}$$

$$\text{shrink}(CSH_I(i)) = \begin{cases} CSH_I; & CSH_I > w \\ 0; & \\ -CSH_I; & -CSH_I < -w \end{cases}, \tag{9}$$

$$w = \text{Tscalars}(s) * \text{sigma} * \text{dst_scalars}(j), \quad s = 0, 1, 2, 3, 4; j = 1, 2, 3, 4, \tag{10}$$

where sigma is the parameter selected according the noise level, $\text{Tscalars}(\cdot)$ is a threshold scalar and its value is set to [thr 5 1 2 4] in this paper. $\text{dst_scalars}(j)$ is the scalar cell array of the noise estimation level.

Fig. 3 Proposed filtering framework



In addition, a backward SHT transform will operate on the SHT coefficients:

$$I_1(x, y) = \text{SHT}^{-1}(\text{CSH}_i(i)), i = 0, 1, 2, 3, 4, \quad (11)$$

where $\text{SHT}^{-1}(\cdot)$ represents inverse SHT transform and $I_1(x, y)$ represents the image obtained by the inverse transform. In addition, the filtering effect of SHT is largely related to the parameter sigma, and the different values of sigma will eventually lead to different performance of image denoising.

Second, it combines block search and matching to achieve structural registration. Using the excellent sparse expression performance of BM3D, image registration is realized through BM3D's block matching, and the structure of the image is preserved. Expressed by the following mathematical formula as

$$I_2(x, y) = B_m(I_1(x, y), 40), \quad (12)$$

where $I_2(x, y)$ is the result after BM3D's block search and matching processing with noise level 40 for $I_1(x, y)$ to greater achieve the structure registration, and $B_m(\cdot)$ represents the BM3D transform.

2.2.2 Structure extraction

First, we use relative total variation for image $I_2(x, y)$ to extract the structure of the image, and obtain the image $I_3(x, y)$. Expressed by the following mathematical formula as

$$I_3 = \arg \min_{I_3} \sum_n (I_2(n) - I_3(n))^2 + \lambda \cdot \left(\frac{I_x(n)}{\zeta_x(n) + \varepsilon} + \frac{I_y(n)}{\zeta_y(n) + \varepsilon} \right), \quad (13)$$

where the value of λ typically varies in a small range [0.01, 0.03] in practice. In this paper, we set it to 0.01. ε is a small positive number to avoid division by zero. The first term of this formula is used as the fidelity term so that the processed image cannot be too deviated. The second term is used as a regular term to capture the structural information of the set window relative to the global change and adjust the balance through the parameter λ . More detailed information can be seen in Sect. 2.1.3 and the specific solution process can refer to Ref. [40].

Then, we superimpose the obtained result with the above structure registration result to highlight the structure of the image and obtain $I_4(x, y)$. It can be expressed as a mathematical formula:

$$I_4(x, y) = I_2(x, y) + I_3(x, y). \quad (14)$$

Second, we cluster the obtained images $I_4(x, y)$ with fuzzy C-means clustering. In this process, we divide it into two categories to facilitate the separation of the structure from the background noise.

The standard FCM algorithm assigns pixels to each category using fuzzy memberships. In this paper, we classify the pixels in image $I_4(x, y)$. Let $X = [x_1, x_2, \dots, x_n]^T$ denote an image $I_4(x, y)$ with N pixels to be partitioned into two clusters. FCM clustering algorithm is an iterative optimization that minimizes the cost function [41]. The energy function of the clustering mask is calculated as

$$\text{Mask} = \sum_{j=1}^N \sum_{i=1}^c \mu_{ij}^m \|x_j - v_i\|^2, \quad (15)$$

where μ_{ij} represents the membership of pixel x_j in the i th cluster, c is the number of cluster, v_i is the i th cluster center; $\|\cdot\|$ is the norm metric; $m \in [1, \infty]$ is the cluster fuzziness. In this paper, we set c and m to 2, respectively. The membership functions and clusters are updated as

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{\frac{2}{m-1}}}, \quad (16)$$

and

$$v_{ij} = \frac{\sum_{j=1}^N \mu_{ij}^m x_j}{\sum_{j=1}^N \mu_{ij}^m}. \quad (17)$$

In addition, we obtained the *Mask* of the structure of the image, as shown in Fig. 10. In addition, then get the structure of the image $G(x, y)$. It can be expressed as a mathematical formula:

$$G(x, y) = \text{Mask} * I(x, y). \quad (18)$$

2.2.3 Structure denoising

Here, we first propose a new variational method to filter the structure image $G(x, y)$:

The energy function of the proposed method can be written as follows:

$$I_s, I_n = \arg \min_{I_s, I_n} \|I_s\|_{\text{BM3D}} + \delta \|I_n\|_{L^2}^2, G = I_s + I_n, \quad (19)$$

where I_s and I_n represent the structure of the image and the noise contained in the image, respectively. $\|\cdot\|_{\text{BM3D}}$ represents block matching 3D space with noise level parameter σ_1 , and $\|\cdot\|_{L^2}^2$ represents L^2 space, which is square integrable space and is the most commonly used and simplest functional space.

By minimizing the energy function theoretically obtain satisfactory filtering results, so as to achieve the purpose of separating structure from noise, and then better remove noise. Furthermore, we numerically solve the designed energy function.

The above energy function can be converted to the following unconstrained model by Lagrangian multiplier [42], and an effective measure to solve the model is to minimize each variable I_s and I_n in the energy function in turn:

$$I_s, I_n = \arg \min_{I_s, I_n} \|I_s\|_{BM3D} + \delta \|I_n\|_{L^2}^2 + \frac{1}{2} \|G - I_s - I_n\|_{L^2}^2. \tag{20}$$

With I_n being fixed, I_s is a solution of

$$I_s = \|G - I_n\|_{BM3D}. \tag{21}$$

With I_s being fixed, I_n is a solution of

$$I_n = \frac{1}{2\delta + 1} (G - I_s), \tag{22}$$

where δ belongs to the noise coefficient, we set it to 2 in this paper.

We summarize the main steps for structure-denoising in algorithm A, as follows:

Algorithm A

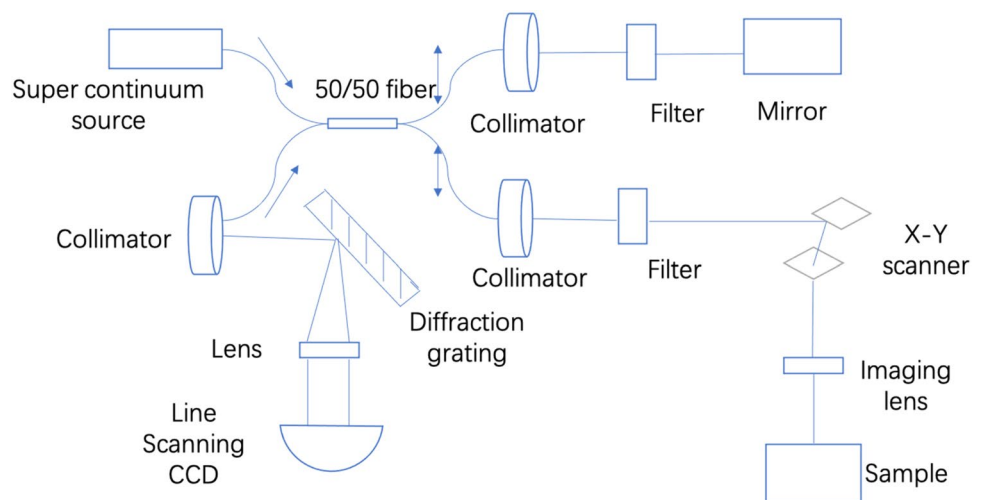
1. Initialization: $I_s = I_n = 0$, G is the initial image
2. Iterations:
 - 2.1 $I_s = \|G - I_n\|_{BM3D}$
 - 2.2 $I_n = \frac{1}{2\delta + 1} (G - I_s)$
 - 2.3 If $n = N_0$, $G = I_s + I_n$
3. Stop test: we stop if $n > N$

In algorithm A, N_0 is the number of iterative updates in the middle, and N is the maximum number of iterations of the update.

3 Experiment and discussion

In this section, we test and verify the performance of our proposed method using seven large noisy OCT images, which include the five human retina images and two mouse optic nerve images. These images are obtained through the traditional SD-OCT system, which has the advantages of high-speed scanning and high accuracy, as shown in Fig. 4. In addition, these human retina images are from a public data set [43] and mouse optic nerve images are from data set [46], as shown in Fig. 5. For the convenience of readers, we named these images as OCT-A, OCT-B, OCT-C,

Fig. 4 Traditional SD-OCT system



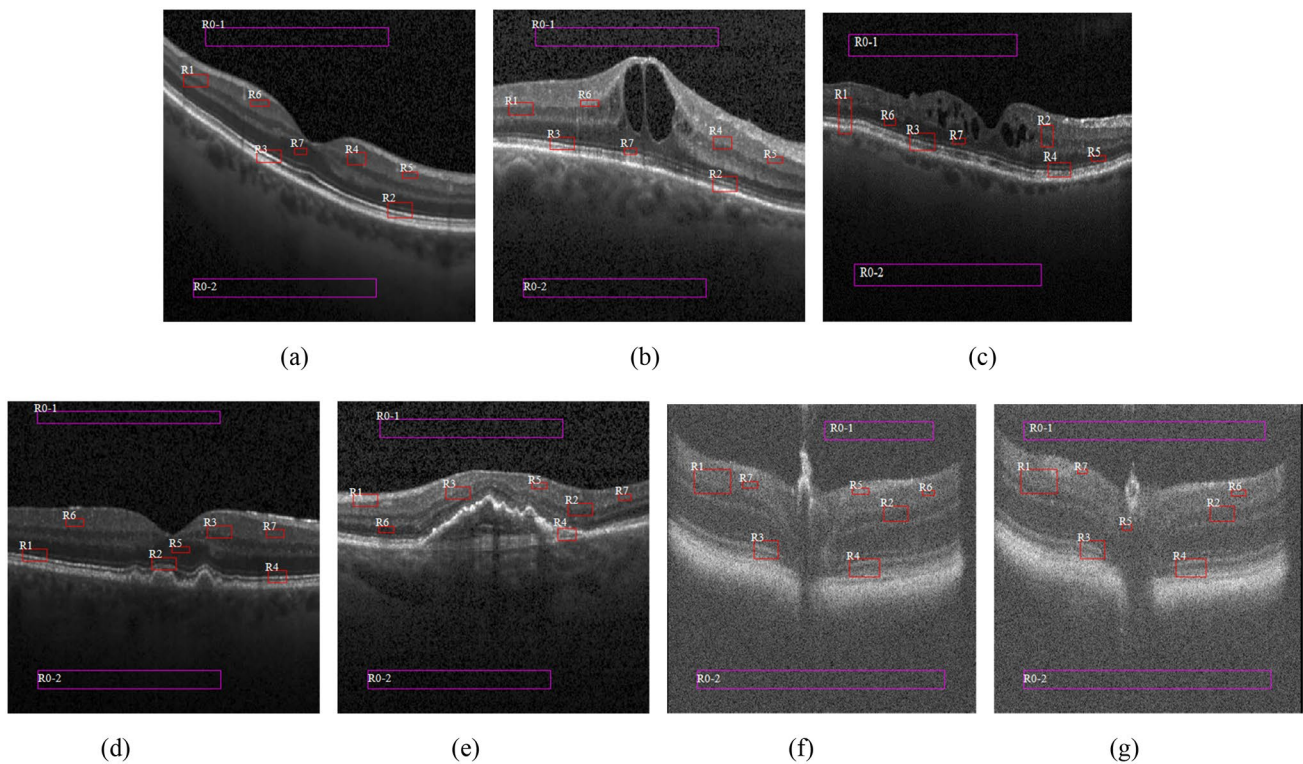


Fig. 5 Original noisy OCT images: **a** OCT-A; **b** OCT-B; **c** OCT-C; **d** OCT-D; **e** OCT-E; **f** OCT-mouse1; **g** OCT-mouse2

OCT-D, OCT-E, OCT-mouse1, OCT-mouse2. According to Fig. 5, it can be seen that the lesion areas of different OCT images are different, and the different areas all contain a certain amount of speckle noise, which greatly affects the structure information of the image and brings trouble to the doctor's diagnosis. It is worth noting that OCT-B–C have a rich porous structure.

All experiments are processed by MATLAB R2016a under the same conditions of a Lenovo ThinkPad E450 computer equipped with 2.2 GHz CPU and 4 GB RAM memory.

3.1 Performance evaluation

For the quantitative evaluation, some aspects are used to quantify the image quality, including the signal-to-noise ratio (SNR), the contrast-to-noise ratio (CNR) and the averaged equivalent number of homogeneous areas (ENL). In Fig. 5, some areas are marked to calculate SNR and CNR, in which R0-1–R0-2 as the ROIs are used to achieve the SNR, R1–R4 as the ROIs are used to achieve the CNR and R5–R7 as the ROIs are used to achieve the ENL. The overall selected region of interest covers the structural information of the image and can better reflect the microscopic mathematical characteristics of the image, so that the image structure can be expressed in detail. The three index equations are, respectively, defined as follows:

$$\text{SNR} = 10 \cdot \log_{10} \left(\frac{\max(I^2)}{\sigma_n^2} \right), \quad (23)$$

$$\text{CNR} = \frac{1}{m} \left(\sum_{i=1}^m \frac{\mu_i - \mu_n}{\sqrt{\sigma_i^2 + \sigma_n^2}} \right), \quad (24)$$

$$\text{ENL} = \frac{1}{H} \left(\sum_{i=1}^H \frac{\mu_i^2}{\sigma_i^2} \right), \quad (25)$$

where μ_n and σ_n^2 represent the mean and variance of the noise region, m and H indicates the number of regions selected in the original noisy image. It can be seen from Eq. (23) that SNR is often used to measure the amount of noise contained in the signal. The higher the value, the less noise the signal contains and the better the image quality. In this paper, it can be used as an indicator of speckle noise reduction capability.

It can be seen from Eq. (24) that CNR is a measure of the ability of image structure to remove noise. The more noise is removed from the structure part, the more the image structure contrast will increase. In addition, as an index for quantizing the noise reduction performance of the structure, the higher the CNR value, the less the noise content of the image structure part, which means the greater the degree

of separation between the image structure and the noise. According to Eq. (25), it can be seen that ENL is used to measure the smoothness of detail information in the image. The larger the value of ENL, the smoother the details of the image and the less noise in the details of the image.

3.2 Experimental results and discussion

Here, we use the selected algorithm to conduct experimental research on the obtained OCT experimental images, including SHT, BM3D, TV-SHT, TV-BM3D and the proposed method. Generally speaking, for a fair comparison, we compare the best results of each method. In this process, the parameters of all algorithms are selected based on better filtering performance, in which the results of the BM3D method depend on thresholding factor of noise level σ , the results of the SHT method depend on the parameter: sigma, the results of the TV-SHT method depend on the parameters: $\mu, \lambda, \eta', \eta''$, and the results of the TV-BM3D method depend on the parameters: $\mu, \alpha, \delta, \beta, N_0, N$.

We show the performance of seven OCT images in numerical forms among Tables 1, 2 and 3. From the tables, it can be seen that our proposed method outperforms all the

previous methods (SHT, BM3D, TV-SHT and TV-BM3D) by achieving highest values in SNR, CNR and ENL. It is worth mentioning that because the background noise is eliminated, the noise variance is 0, and the SNR is expressed by max. According to Tables 1 and 2, by comparing the value of SNR and CNR, it can be seen that the background noise removal and structure denoising ability of TV-BM3D in most OCT images is better than SHT, BM3D and TV-SHT methods. Finally, according to Table 3, by comparing the value of ENL, it can be seen that the smooth structure of BM3D in OCT image is much stronger than SHT, TV-SHT and TV-BM3D methods. We also show the parameter values of different methods for all OCT images in Table 4. Combining the three quantization indexes, we can find that our proposed method has achieved good results in the three aspects of background noise removal, structure denoising and structure smoothing. In addition, from the visual effect point of view, our proposed filtering and structure extraction methods have achieved good results.

Figure 6 shows the results of the OCT images processed by SHT algorithm. From a visual point of view, it can be seen from the result graph that the algorithm can remove background noise to a certain extent, and keep the high

Table 1 SNR results in OCT images

Methods	OCT-A	OCT-B	OCT-C	OCT-D	OCT-E	OCT-mouse1	OCT-mouse2
Original	28.4875	26.7188	30.3994	29.7485	26.4991	24.0597	23.2360
SHT	34.6176	40.5572	40.7131	42.9058	40.4294	44.7045	42.9840
BM3D	34.1465	34.7036	40.0868	40.3351	37.0123	44.1496	41.8459
TV-SHT	34.5276	33.0496	46.6660	45.5074	29.8923	37.1079	34.8761
TV-BM3D	41.6474	48.9775	54.0247	46.0266	42.3412	44.4088	42.0033
Proposed	max	max	max	max	max	max	max

Table 2 CNR results in OCT images

Methods	OCT-A	OCT-B	OCT-C	OCT-D	OCT-E	OCT-mouse1	OCT-mouse2
Original	2.9169	3.3158	2.4612	2.5605	3.4325	1.1668	1.1407
SHT	4.2335	6.4812	5.1884	3.8465	5.9748	3.3216	2.7961
BM3D	3.8126	5.7560	4.5172	3.1684	5.3844	2.9919	2.6245
TV-SHT	4.7077	5.5717	6.1564	4.5206	5.1583	2.5811	2.3439
TV-BM3D	4.7220	7.2749	4.4851	3.3090	6.0058	3.8383	3.4251
Proposed	5.0518	11.3234	5.8258	5.3396	7.1093	9.5959	9.1442

Table 3 ENL results in OCT images

Methods	OCT-A	OCT-B	OCT-C	OCT-D	OCT-E	OCT-mouse1	OCT-mouse2
Original	441.3	123.6270	48.5	161.5	110.8	35.4	31.2
SHT	695.0	238.7960	215.5	1097.4	673.7	3085.6	4981.9
BM3D	2724.3	347.9196	475.1	1265.7	2107.8	3820.9	9128.3
TV-SHT	360.5	228.0433	164.9	705.0	377.1	333.2	405.0
TV-BM3D	1157.3	475.0802	138.9057	1949.2	1668.3	8499.0	4752.2
Proposed	3986.2	356.6745	469.7235	2260.6	3291.9	3260.3	5656.0

Table 4 Parameters with different methods for OCT images

	Methods	Parameters
OCT-A	SHT	$\sigma = 30$
	BM3D	$\sigma = 20$
	TV-SHT	$\mu = 1, \lambda = 1, \eta' = 0.09, \eta'' = 100$
	TV-BM3D	$\mu = 0.63, \alpha = 2.0, \delta = 0.075, \beta = 0.15, N_0 = 5, N = 2$
	Proposed	thr = 34, $\sigma = 0.051015, \alpha = 20, N_0 = 2, N = 3$
OCT-B	SHT	$\sigma = 40$
	BM3D	$\sigma = 50$
	TV-SHT	$\mu = 1, \lambda = 1, \eta' = 0.1, \eta'' = 100$
	TV-BM3D	$\mu = 0.33, \alpha = 1.6, \delta = 0.065, \beta = 0.11, N_0 = 4, N = 2$
	Proposed	thr = 33, $\sigma = 0.071205, \sigma = 20, N_0 = 2, N = 3$
OCT-C	SHT	$\sigma = 30$
	BM3D	$\sigma = 40$
	TV-SHT	$\mu = 1, \lambda = 1, \eta' = 0.1, \eta'' = 100$
	TV-BM3D	$\mu = 0.75, \alpha = 2.6, \delta = 0.085, \beta = 0.22, N_0 = 5, N = 2$
	Proposed	thr = 29, $\sigma = 0.041235, \sigma = 20, N_0 = 2, N = 3$
OCT-D	SHT	$\sigma = 30$
	BM3D	$\sigma = 50$
	TV-SHT	$\mu = 1, \lambda = 1, \eta' = 0.1, \eta'' = 100$
	TV-BM3D	$\mu = 0.63, \alpha = 1.9, \delta = 0.075, \beta = 0.143, N_0 = 4, N = 2$
	Proposed	thr = 34, $\sigma = 0.070125, \sigma = 20, N_0 = 2, N = 3$
OCT-E	SHT	$\sigma = 35$
	BM3D	$\sigma = 60$
	TV-SHT	$\mu = 1, \lambda = 1, \eta' = 0.1, \eta'' = 100$
	TV-BM3D	$\mu = 0.67, \alpha = 2.2, \delta = 0.065, \beta = 0.143, N_0 = 4, N = 2$
	Proposed	thr = 38, $\sigma = 0.063205, \sigma = 20, N_0 = 2, N = 4$
OCT-mouse1	SHT	$\sigma = 25$
	BM3D	$\sigma = 60$
	TV-SHT	$\mu = 1, \lambda = 1, \eta' = 0.1, \eta'' = 100$
	TV-BM3D	$\mu = 0.37, \alpha = 2.7, \delta = 0.075, \beta = 0.203, N_0 = 5, N = 2$
	Proposed	thr = 66, $\sigma = 0.052145, \sigma = 20, N_0 = 2, N = 3$
OCT-mouse2	SHT	$\sigma = 45$
	BM3D	$\sigma = 40$
	TV-SHT	$\mu = 1, \lambda = 1, \eta' = 0.1, \eta'' = 100$
	TV-BM3D	$\mu = 0.39, \alpha = 2.9, \delta = 0.085, \beta = 0.247, N_0 = 5, N = 2$
	Proposed	thr = 78, $\sigma = 0.064205, \sigma = 20, N_0 = 2, N = 4$

frequency part of the image relatively complete. However, when the background noise is too large, part of the low-frequency structure of the image is often missing during the filtering process. In addition, there are some scratches at the same time. Figure 7 shows the results of the OCT images processed by the BM3D algorithm. From a visual point of view, the background of the image is turbid, and still retains a lot of noise. However, the algorithm has better protection ability for the image structure, and can describe the image structure vividly in the smoothing process, making the image structure clearer. Figure 8 shows the results of the OCT images processed by the TV-SHT method. From a visual point of view, the structural part of the image gradually becomes prominent, but the background of some

excessively noisy images is accompanied by stronger noise artifacts and more scratches appear in the structural part. In terms of quantization index, in comparison, due to its effect on noise compensation, the TV-SHT method has relatively low overall values in terms of three indicators. However, with its excellent high-frequency sparse expression, the structure of the image is more complete, despite the presence of scratches. Figure 9 shows the results of the OCT images processed by the TV-BM3D method. It can be seen that under the background of large noise, the effect of this method is cleaner than the overall visual effect of the above method, and the structure of the image can be maintained to a certain extent. At the same time, it can be seen from the quantitative index that the method reduces the background

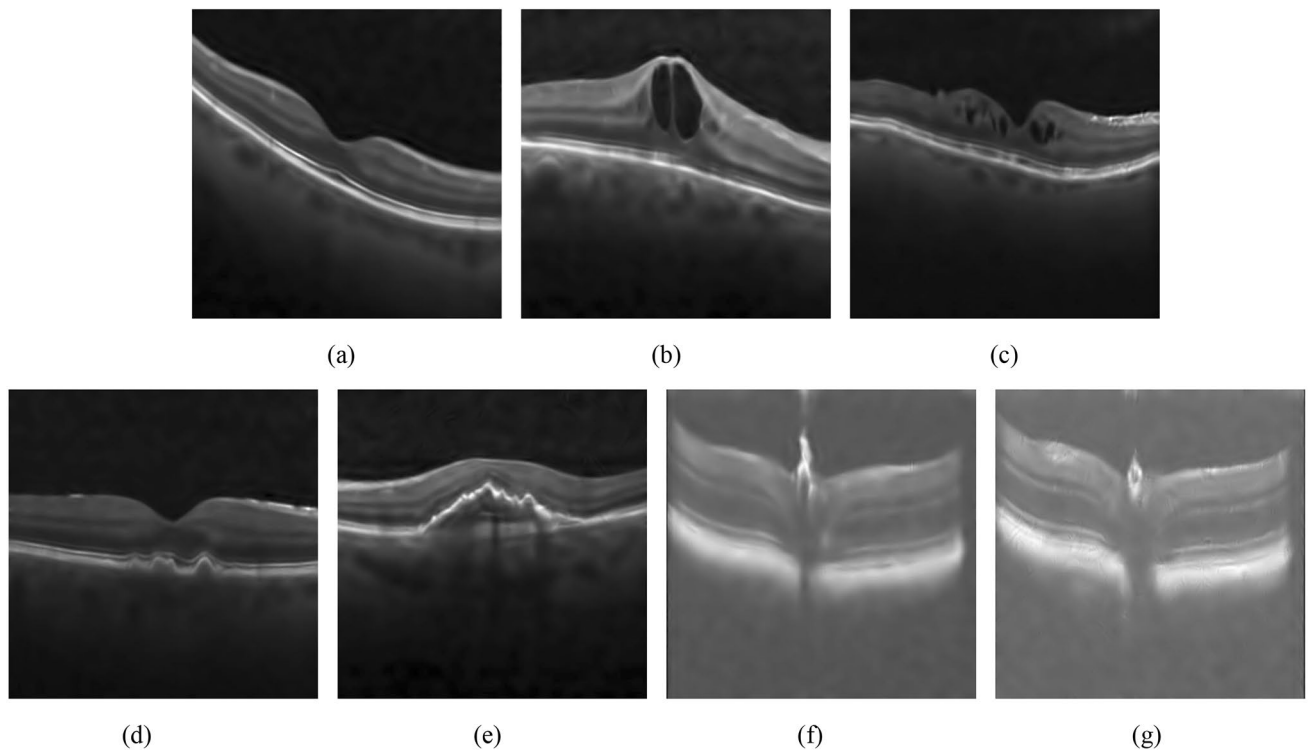


Fig. 6 Filtering results processed by SHT method: **a** OCT-A; **b** OCT-B; **c** OCT-C; **d** OCT-D; **e** OCT-E; **f** OCT-mouse1; **g** OCT-mouse2

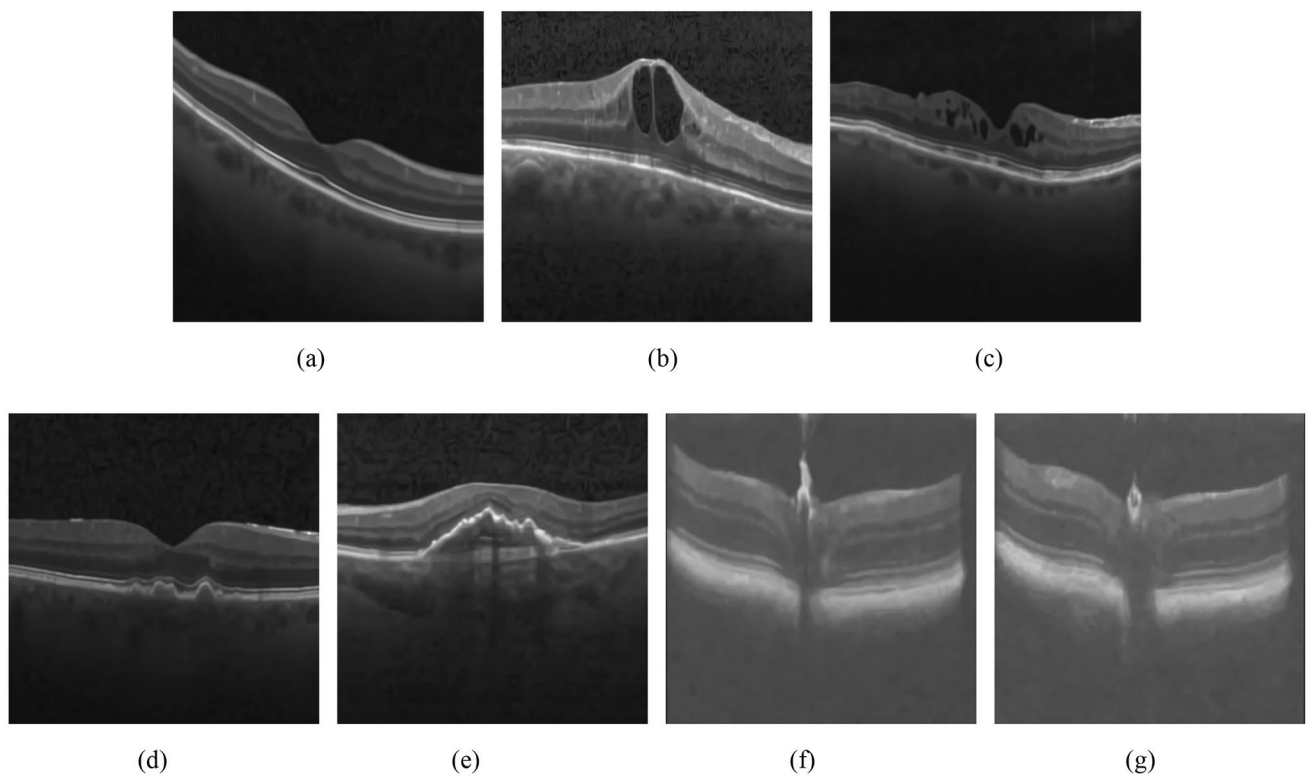


Fig. 7 Filtering results processed by BM3D method: **a** OCT-A; **b** OCT-B; **c** OCT-C; **d** OCT-D; **e** OCT-E; **f** OCT-mouse1; **g** OCT-mouse2

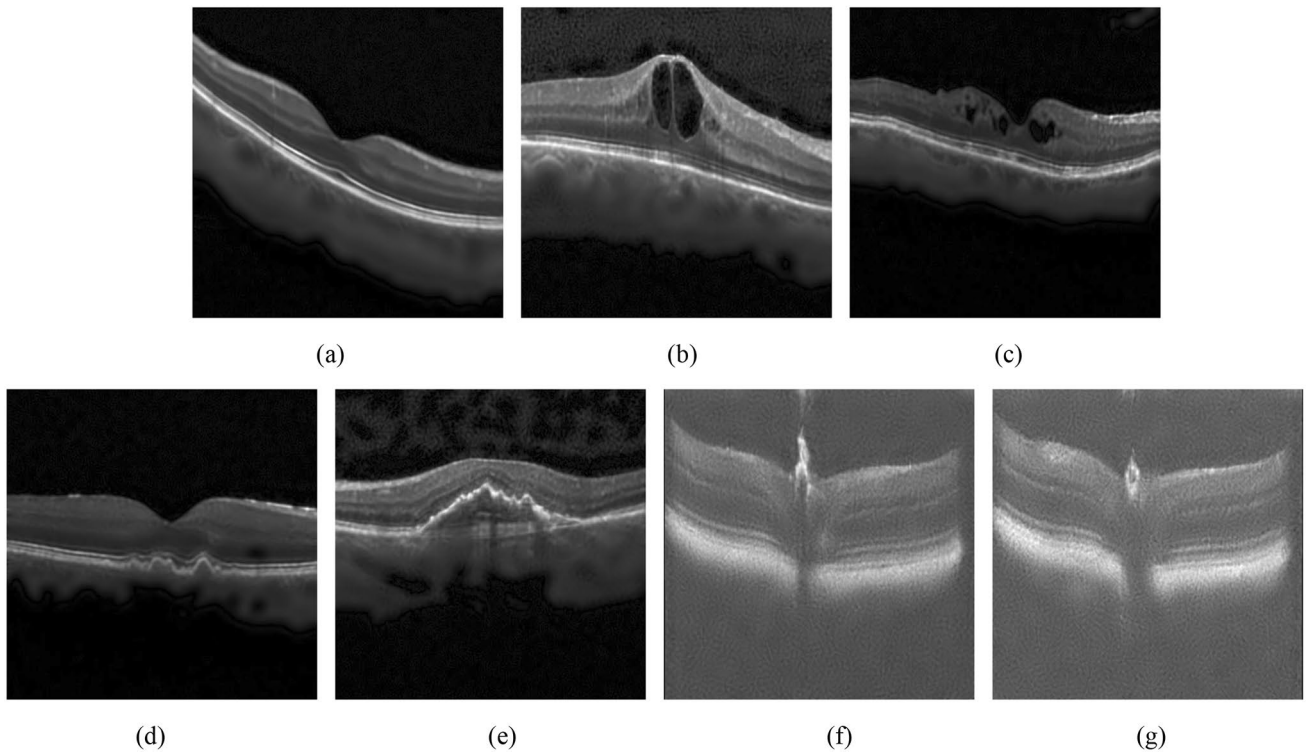


Fig. 8 Filtering results processed by TV-SHT method: **a** OCT-A; **b** OCT-B; **c** OCT-C; **d** OCT-D; **e** OCT-E; **f** OCT-mouse1; **g** OCT-mouse2

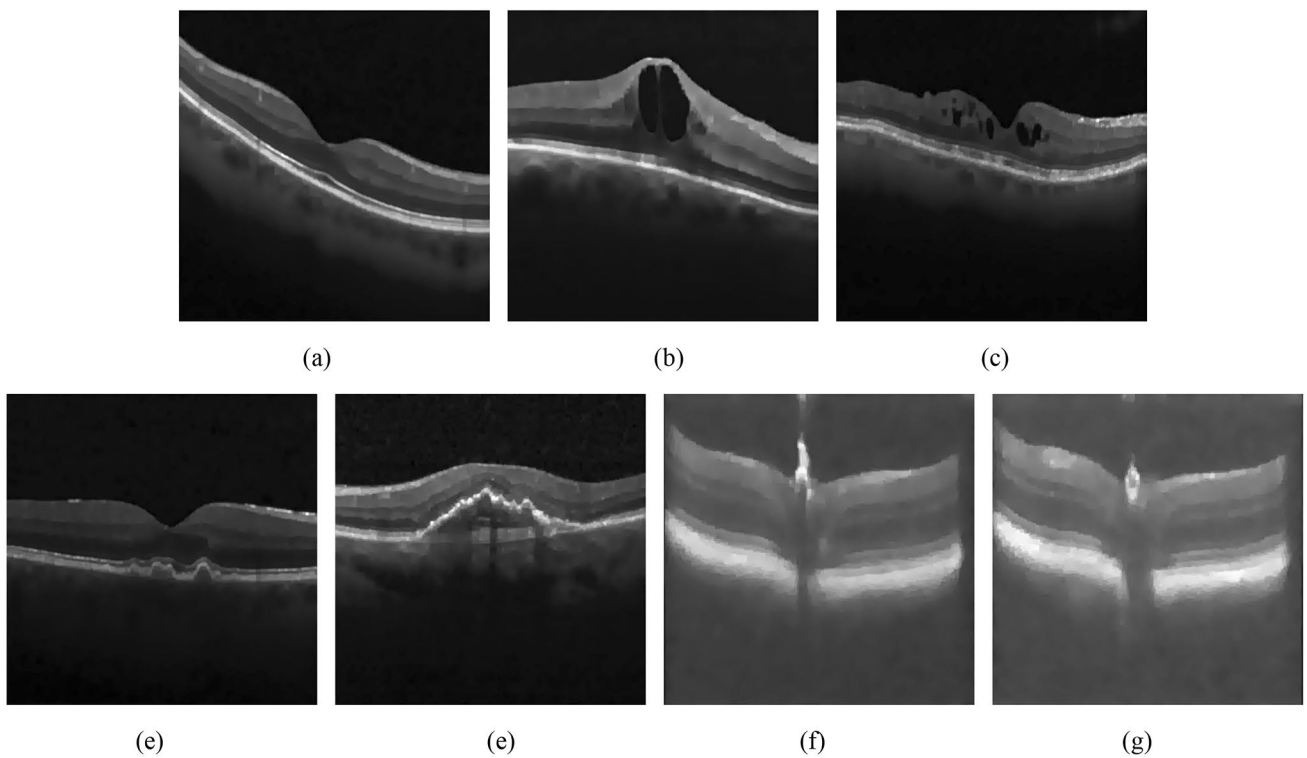


Fig. 9 Filtering results processed by TV-BM3D method: **a** OCT-A; **b** OCT-B; **c** OCT-C; **d** OCT-D; **e** OCT-E; **f** OCT-mouse1; **g** OCT-mouse2

noise better, and the contrast has been improved to a certain extent. However, inevitably, this filtering method leads to the loss of some minute structure information of the image. Especially in the case of large noise and low contrast, the image loss will be more serious.

Figure 10 shows the image structure mask obtained after processing by our proposed method, which clearly shows us the structure of the image. Figure 11 shows the result of using the new BM3D- L^2 variational algorithm on the structure of the image. Since the extracted structure contains some noise, to better preserve the image details, we adopt the idea of BM3D- L^2 . The algorithm describes the structure of the image in BM3D space, and describes the noise in L^2 space. It can be seen that the image structure noise is eliminated, and the overall appearance is very clean, almost close to the noise-free image. According to the quantization index, the proposed algorithm effectively removes background noise and structural noise, and also has a very significant improvement in detail smooth protection, fully demonstrating the superiority of our algorithm. We can get the following conclusions from the experimental results:

1) For the first time, by combining the SHT and BM3D for image preprocessing, the background noise of the image can be effectively filtered, and protect important structural information during this process. Then it is

combined with the relative total variation, enhancing the structural information of the image.

- 2) According to the experimental results, it is found that, respectively, using SHT, BM3D, TV-SHT and TV-BM3D methods to filter the original image cannot completely eliminate the background noise, and the presence of noise in the filtering process will affect the structural information. The fuzzy C-means algorithm effectively combines the ideas of machine learning algorithms for image segmentation, cleverly separates the structural information of the image from the background, and obtains the structural mask of the image, thereby eliminating the effect of structural denoising.
- 3) For the filtering of the structure part, we first propose the BM3D- L^2 variational algorithm, which divides the structure into two parts: image structure information and noise information, and the two sections are described in BM3D space and L^2 space. In this way, the structure can be separated from the noise part. According to the experimental results, it can be found that the proposed algorithm has significant advantages in image structure denoising and detail protection.

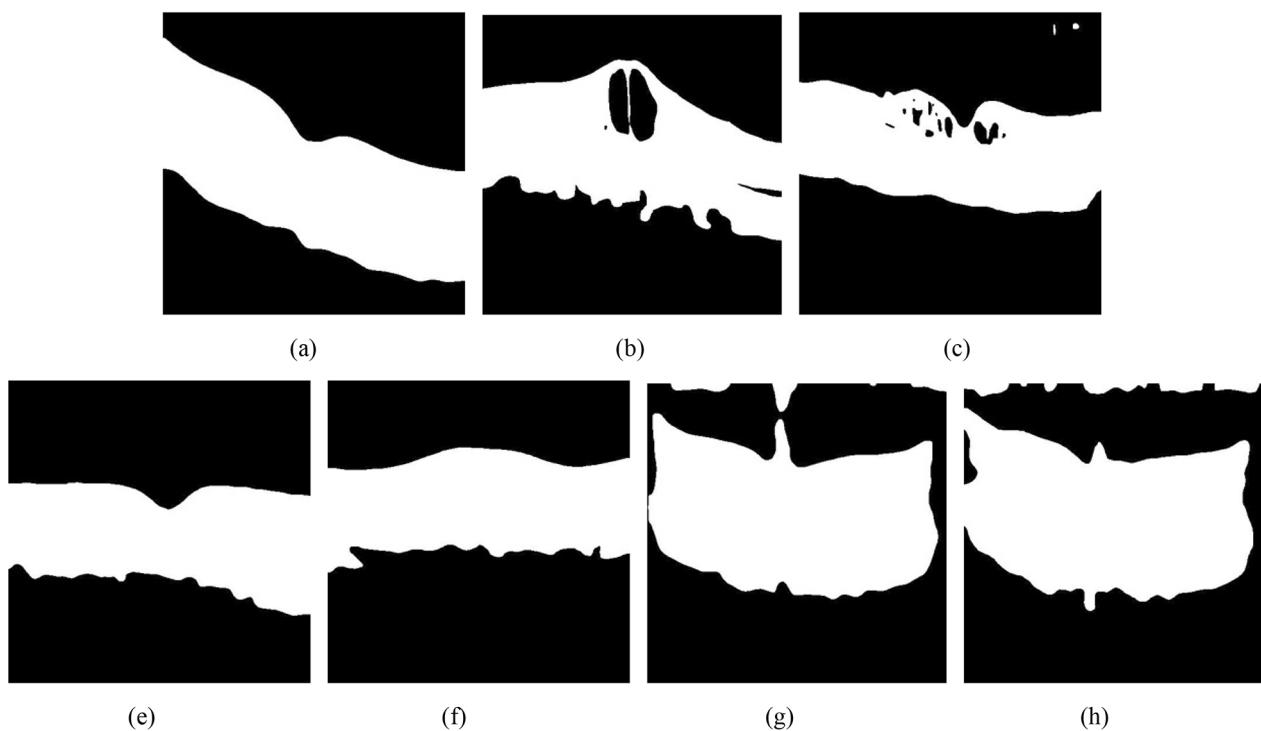


Fig. 10 Structural clustering masks by proposed method: **a** OCT-A; **b** OCT-B; **c** OCT-C; **d** OCT-D; **e** OCT-E; **f** OCT-mouse1; **g** OCT-mouse2

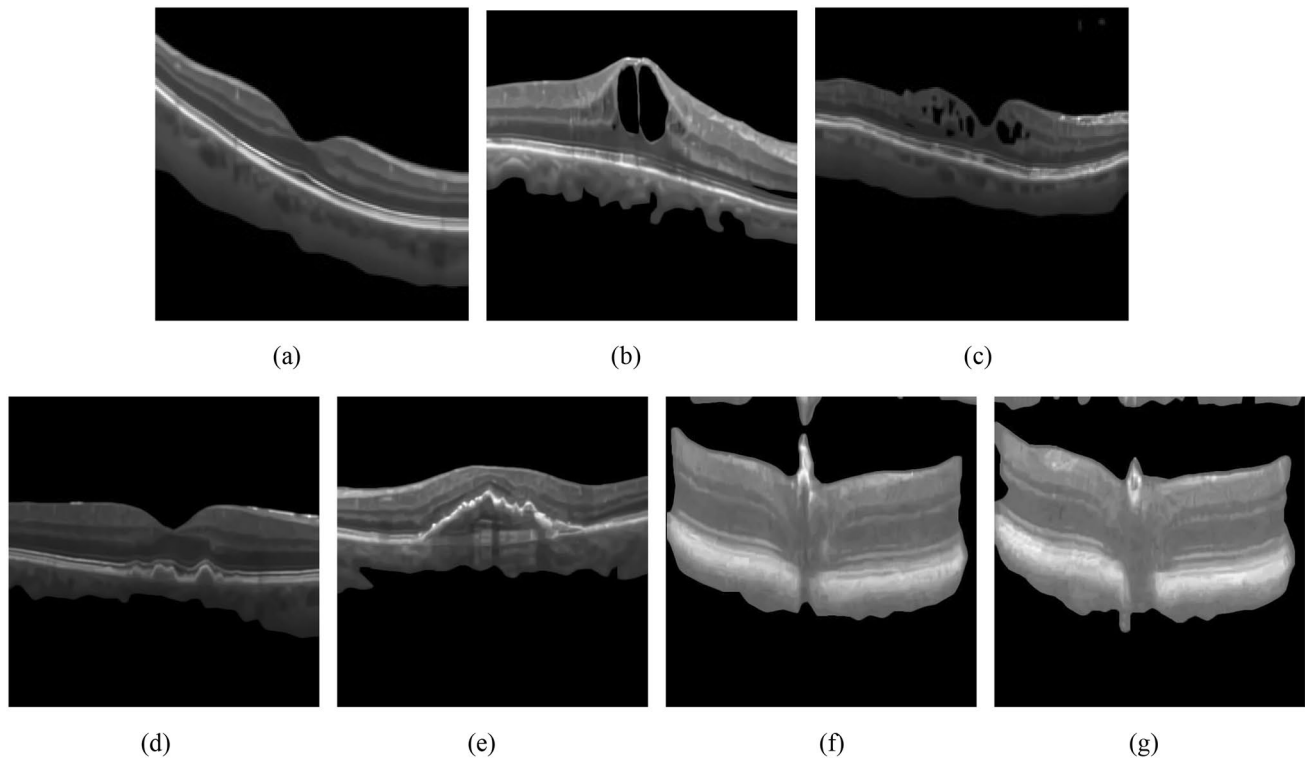


Fig. 11 Filtering results processed by proposed method: **a** OCT-A; **b** OCT-B; **c** OCT-C; **d** OCT-D; **e** OCT-E; **f** OCT-mouse1; **g** OCT-mouse2

4 Conclusion

In this paper, we first construct an image filtering framework based on clustering for large noise OCT images. The framework can be divided into preprocessing, structure extraction and structure denoising. In the preprocessing stage, we use the SHT method to filter out background noise, and combine 3D block matching to denoise the image structure to achieve structural protection. In the structure extraction stage, we use the relative total variation algorithm to extract the structure, and combine the fuzzy C-means clustering algorithm to obtain the image structure mask. In the structure denoising stage, we propose a new variational BM3D- L^2 method, which uses BM3D space and L^2 space to describe the structure and noise of the image. During this process, we can better eliminate the image structure noise and protect the image structure texture by assigning appropriate parameter values. We tested our method on seven large noisy OCT images, and compared it with SHT, BM3D, TV-SHT and TV-BM3D methods. Combining quantitative and qualitative indicators, it can be found that our proposed method is effective, which can reduce the speckle noise of large-noise OCT images while protecting the structural information.

Acknowledgements This work was supported by the National Natural Science Foundation of China (NNSFC) (Grant nos. 11772081, 11972106).

References

1. D. Huang, E.A. Swanson, C.P. Lin et al., *Science* **254**, 1178–1181 (1991)
2. W. Drexler, U. Morgner, R.K. Ghanta et al., *Nat. Med.* **7**, 502–507 (2001)
3. F. Fercher, W. Drexler, C.K. Hitzenberger et al., *Rep. Prog. Phys.* **66**, 239–303 (2003)
4. J.M. Schmitt, S.H. Xiang, K.M. Yung, *J. Biomed. Opt.* **4**, 95–105 (1999)
5. M. Wojtkowski, *Appl. Opt.* **49**, D30–D61 (2010)
6. Y. Matsuo, T. Sakamoto, T. Yamashita et al., *Investig. Ophthalmol. Vis. Sci.* **54**, 7630–7636 (2013)
7. M. Wojtkowski, R. Leitgeb, A. Kowalczyk et al., *J. Biomed. Opt.* **7**, 457–463 (2002)
8. M. Gora, K. Karnowski, M. Szkulmowski et al., *Opt. Express* **17**, 14880–14894 (2009)
9. M. Hughes, M. Spring, A. Podoleanu, *Appl. Opt.* **49**, 99–107 (2010)
10. M. Szkulmowski, I. Gorczynska, D. Szigal et al., *Opt. Express* **20**, 1337–1359 (2012)
11. T. Loupas, W. McDicken, P. Allen, *IEEE Trans. Circuits Syst.* **36**, 129–135 (1989)
12. J.S. Lee, *Comput. Graph. Image Process.* **17**, 24–32 (1981)
13. G. Franceschetti, V. Pascazio, G. Schirrinzi, *J. Opt. Soc. Am. A* **12**, 686–694 (1995)

14. A. Wong, A. Mishra, K. Bizheva et al., *Opt. Express* **18**, 8338–8352 (2010)
15. Y. Yu, S.T. Acton, *IEEE Trans. Image Process.* **11**, 1260–1270 (2002)
16. H.M. Salinas, D.C. Fernández, *IEEE Trans. Med. Imaging* **26**, 761–771 (2007)
17. R. Bernardes, C. Maduro, P. Serranho et al., *Opt. Express* **18**, 24048–24059 (2010)
18. J. Aum, J.H. Kim, J. Jeong, *Appl. Opt.* **54**, D43–D50 (2015)
19. H. Yu, J. Gao, A. Li, *Opt. Lett.* **41**, 994–997 (2016)
20. H. Chen, S. Fu, H. Wang et al., *J. Biomed. Opt.* **23**, 036014 (2018)
21. H. Lv, S. Fu, C. Zhang et al., *Opt. Express* **26**, 11804 (2018)
22. L. Rudin, S. Osher, E. Fatemi, *Phys. D.* **60**, 259–268 (1992)
23. D.L. Donoho, M. Johnstone, *J. Am. Stat. Assoc.* **90**, 1200–1224 (1995)
24. D.L. Marks, T.S. Ralston, S.A. Boppart, *J. Opt. Soc. Am. A* **22**, 2365–2371 (2005)
25. G. Aubert, P. Kornprobst, Springer Science & Business Media (2006)
26. G. Gong, H. Zhang, M. Yao, *Opt. Express* **23**, 24699–24712 (2015)
27. J.X. Deng, Y.M. Liang, *Acta Photonica Sin.* **29**, 2138–2141 (2009)
28. D.C. Adler, T.H. Ko, J.G. Fujimoto, *Opt. Lett.* **29**, 2878–2880 (2004)
29. D. Gupta, R.S. Anand, B. Tyagi, *IET Image Process.* **9**, 107–117 (2015)
30. Q. Guo, F. Dong, S. Sun et al., *IET Image Process.* **7**, 442–450 (2013)
31. Z. Jian, Z. Yu, L. Yu et al., *Opt. Lett.* **34**, 1516–1518 (2009)
32. K. Dabov, A. Foi, V. Katkovnik et al., *IEEE Trans. Image Process.* **16**, 2080–2095 (2007)
33. J. Ma, G. Plonka, *IEEE Trans. Image Process* **16**, 2198–2206 (2007)
34. T. Zeng, X. Li, M. Ng, *Commun. Comput. Phys.* **8**, 976 (2010)
35. M. Xu, C. Tang, M. Chen et al., *Opt. Laser. Eng.* **122**, 265–283 (2019)
36. S. Huang, C. Tang, M. Xu et al., *Appl. Opt.* **58**, 6233–6243 (2019)
37. S. Adabi, E. Rashedi, A. Clayton et al., *J. Biomed. Opt.*, **23**, 016013 (2018)
38. Y. Ma, X. Chen, W. Zhu et al., *Biomed. Opt. Express* **9**, 5129–5146 (2018)
39. K.J. Halupka, B.J. Antony, M.H. Lee et al., *Biomed. Opt. Express* **9**, 6205–6221 (2018)
40. Y. Chen, J. Li, R. Nian et al., in *OCEANS 2017 - Anchorage*, pp. 1–6 (2017).
41. K.S. Chuang, H.L. Tzeng, S. Chen et al., *Comput. Med. Imaging Graph.* **30**(1), 9–15 (2006)
42. J.F. Aujol, A. Chambolle, *Int. J. Comput. Vis.* **63**, 85–104 (2005)
43. D.S. Kermany, M. Goldbaum, W. Cai et al., *Cell* **172**(5), 1122–1131 (2018)
44. Hu. Yibing, C. Tang, Xu. Min et al., *Appl. Opt.* **58**, 9861–9869 (2019)
45. M. Hossein Eybposh, Z. Turani, D. Mehregan et al., *Biomed. Opt. Express* **9**, 6359–6373 (2018)
46. L. Fang, *IEEE Trans. Med. Imaging* **32**(11), 2034–2049 (2013)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.