Patch-based denoising method using low-rank technique and targeted database for optical coherence tomography image

Xiaoming Liu
Zhou Yang
Jia Wang
Jun Liu
Kai Zhang
Wei Hu

Patch-based denoising method using low-rank technique and targeted database for optical coherence tomography image

Xiaoming Liu,^{a,b,*} Zhou Yang,^{a,b} Jia Wang,^{a,b} Jun Liu,^{a,b} Kai Zhang,^{a,b} and Wei Hu^{a,b}

^aWuhan University of Science and Technology, College of Computer Science and Technology, Wuhan, China

^bHubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan, China

Abstract. Image denoising is a crucial step before performing segmentation or feature extraction on an image, which affects the final result in image processing. In recent years, utilizing the self-similarity characteristics of the images, many patch-based image denoising methods have been proposed, but most of them, named the internal denoising methods, utilized the noisy image only where the performances are constrained by the limited information they used. We proposed a patch-based method, which uses a low-rank technique and targeted database, to denoise the optical coherence tomography (OCT) image. When selecting the similar patches for the noisy patch, our method combined internal and external denoising, utilizing the other images relevant to the noisy image, in which our targeted database is made up of these two kinds of images and is an improvement compared with the previous methods. Next, we leverage the low-rank technique to denoise the group matrix consisting of the noisy patch and the corresponding similar patches, for the fact that a clean image can be seen as a low-rank matrix and rank of the noisy image is much larger than the clean image. After the first-step denoising is accomplished, we take advantage of Gabor transform, which considered the layer characteristic of the OCT retinal images, to construct a noisy image before the second step. Experimental results demonstrate that our method compares favorably with the existing state-of-the-art methods. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.4.1.014002](https://doi.org/10.1117/1.JMI.4.1.014002)]

Keywords: low rank; image denoising; Gabor transform; optical coherence tomography.

Paper 16097RR received Jun. 2, 2016; accepted for publication Jan. 16, 2017; published online Feb. 1, 2017.

1 Introduction

Optical coherence tomography (OCT), one of the new tomography techniques, provides high-resolution, cross-sectional tomographic image of internal structures of specimens. The sub-surface images at near-microscopic resolution can be obtained using this technique, for the reason that OCT is based on light rather than on sound. The wavelength of light is much shorter and its speed is much faster than sound waves. OCT is based on low-coherence interferometry, and the technique has advanced greatly in recent decades. The first generation of OCT is time domain OCT (TD-OCT),¹ which can acquire at ~400 axial scans per second with an axial resolution of 8 to 10 μm in tissue. Spectral domain OCT (SD-OCT) has been developed since 2001. Current commercial SD-OCT devices are 40 to 110 times faster than TD-OCT, with an axial resolution of 5 to 7 μm in tissue. Recently, swept source-OCT (SS-OCT) technique has also been developed, and SS-OCT devices can achieve more than 100,000 A scans per second with an axial resolution of about 3 μm in tissue.

OCT can realize a direct imaging of biological tissue in real time without the preparation of the sample, and it does not involve ionizing radiation, which does not harm human health. It has experienced rapid development in various fields, especially in biomedicine applications, such as ophthalmic imaging, cardiovascular imaging, gastrointestinal imaging, and pulmonary imaging.^{1–5} In conventional excisional biopsy, the images

obtained in the stomach are of low quality, whereas the endoscopic OCT can clearly delineate internal tissue microstructure in gastrointestinal tissue.³ Cardiovascular OCT is an inserted high-resolution imaging application, which deploys stents *in vivo* for imaging coronary arteries via OCT technique.⁴ Combining OCT and autofluorescence, Pahlevaninezhad et al.⁵ presented an imaging of pulmonary nodules *in vivo*, which can locate and characterize pulmonary nodules precisely. Ophthalmology is the first medical application of OCT technique,¹ and OCT has revolutionized the clinical practice of ophthalmology. In this paper, we focus on the OCT images obtained from the retina.

However, the OCT image is usually contaminated by noises of different sources, which mainly include source instability,⁶ tissue heterogeneity,⁷ specular reflection,⁸ and speckle noise.⁹ Due to its coherent detection nature, the OCT image is usually severely degraded by the speckle noise in the acquisition, which decreases the contrast of the image and obscures the meaningful features in the image. Specular reflection happens when the local sample surface is glossy and perpendicular to the incident beam, which results in the vertical stripe noise in OCT. Many researchers have studied the recovering of the clean image from its noisy version. In this paper, the aim of our work is for attenuation of the speckle noise.

Numerous image denoising methods have been proposed for natural images in the past decades, which are coarsely classified into three categories: spatial-domain method, transform-domain method, and compound method. Spatial-domain

*Address all correspondence to: Xiaoming Liu, E-mail: liuxiaoming@wust.edu.cn; lxmspace@gmail.com

method denoises the noisy image pixel wisely by weighting and averaging the neighbors of each pixel in which the weights depend on the similarity, such as the nonlocal mean filters.¹⁰ Transform-domain method projects the noisy image on some representation basis and generates the corresponding sparse coefficients, where image information is mostly focused in the few largest ones. So, noise is removed by shrinking the smaller coefficients, such as the wavelet transform.¹¹ Compound method takes advantage of spatial correlation and sparse representation to suppress noise, such as the block matching and 3-D filtering.¹² Among the aforementioned methods, patch-based image denoising¹³ methods have attracted much attention. Its basic steps are as follows: first, extract a small patch from a noisy image; second, pick out a set of similar patches and apply some functions to those patches for obtaining an estimated patch of the clean patch; third, pick a small patch again and repeat the second step; at last, a large number of estimated patches are acquired and the estimated image of clean image is reconstructed by accumulating all the estimated patches. It needs to decide where the similar patches are selected. There are two sources of them: the noisy image itself, which is known as internal denoising,¹⁴ and external database consisting of some images relevant to the noisy image, which is known as external denoising.¹⁵ Through the experiments, Glasner et al.¹⁶ demonstrates the self-similarity characteristic of a natural image that patches tend to recur many times inside the image at a different location. However, there exists a shortcoming that not all the patches recur in an image, or the frequency of the patches recurring cannot reach the effect of image denoising. Levin and Nadler¹⁷ reveals the rare patches effect and regard it as the inherent bounds of internal denoising. External denoising has the ability to overcome the shortcoming and break the performance bounds. Luo et al.¹⁸ proposed a data-dependent denoising procedure to restore noisy image and leverages the similarity of the external database, showing the superiority of the new algorithm over existing methods. A learning-based approach using a neural network, combining denoising results from an internal method and an external method, was proposed by Burger et al.¹⁹

Low-rank technique,²⁰ which is formulated as a minimization problem where the cost function measures the fit between a given matrix (the original data) and an approximating matrix (the optimization data), is subject to a constraint that the approximating matrix has reduced rank. The fact that a clean image has a low-rank matrix and rank of the noisy image is much larger than the clean image is in-line with the above introduction, which can be utilized in image denoising. Li et al.²¹ proposed an image denoising model using a low-rank dictionary and sparse representation, and achieved effective denoising performance. Guo et al.²² proposed a computationally simple denoising algorithm, exploiting the optimal energy compaction property of singular value decomposition (SVD) to lead a low-rank approximation of similar patch groups using the nonlocal self-similarity and the low-rank approximation. However, most of the aforementioned methods are used for noise reduction of the natural image, whereas less is used for the OCT image. Fang et al.²³ presented a multiscale sparsity-based tomographic denoising method, which uses a sparse representation dictionary from the OCT images with high signal-to-noise ratio (SNR) and utilizes the dictionary to denoise the OCT images with low SNR. Thapa et al.²⁴ learned a new efficient nonlinear dictionary on the OCT images for sparsity-based image denoising, which add nonlinear functions to the conventional discrete cosine

transform atoms. Baghaie et al.²⁵ proposed a sparse and low-rank decomposition-based method for suppressing noise in retinal OCT images, which can achieve simultaneous decomposition and batch alignment. Coupled with the curvelet transform's nearly optimal sparse representation, a method for attenuating speckle noise in OCT images is presented by Jian et al.,²⁶ which preserves many subtle features. Adler et al.²⁷ proposed a spatially adaptive 2-D wavelet filter to reduce noise in time-domain and Fourier-domain OCT images, which attenuate noise power in the wavelet domain without significant loss of image sharpness.

In this paper, we propose a patch-based denoising method using the low-rank technique and targeted database for OCT image, which get better denoising effect and performance compared with the existing methods. The contributions of our work include:

1. Different from the other general noise reduction algorithms, the low-rank technique, as one of the emerging and promising algorithms, is adopted and taken as an important step throughout the entire denoising process, which is comparable with other methods in the noise reduction effect.
2. Because patch-based denoising method is a process from local to whole, to ensure good effectiveness of the overall noise reduction, we must make some improvement in local, which take sources of the similar patches into account. In this paper, we optimize the targeted database by combining internal denoising and external denoising, namely taking the noisy image itself and some images similar with the noisy image in structure as sources of the similar patches, which guarantee sufficient number of similar patches can be found for each noisy patch.
3. In our method, the noise reduction of an image is divided into two major steps and the first step is basically the same as the second step except for the input image. Utilizing Gabor transform, a new noisy image is constructed by adding some missing details to layers in the first-step denoised image, which consider the layer characteristics of an OCT image and keep layer boundaries of the final result clear.
4. The choice of patch size has a significant impact on the final noise reduction effect for patch-based denoising method, where square patch is usually utilized by researchers but the square patch is less flexible for different noise levels and the content of noisy image. In this paper, the rectangular patch is investigated, which has higher flexibility and adaptability.

The remainder of the paper is structured as follows. In Sec. 2, we present our algorithm framework and introduce the detailed process of the proposed method. In Sec. 3, we show our experiment results and compare with other methods. We conclude this paper in Sec. 4.

2 Proposed Method

2.1 Method Framework

Our proposed method used the low-rank technique and optimized the targeted database on the basis of patch-based denoising

Fig. 1 Flowchart of the proposed denoising method for OCT image.

method; meanwhile, a further noise reduction step is incorporated into the whole procedure. Figure 1 shows the flowchart of our method framework, which includes the following steps:

1. Load data: A noisy image and targeted database are loaded. The targeted database consists of some OCT images, and we can find the similar patches in these images for each noisy patch.
2. Find similar patches: For every noisy patch in the noisy image, we will select some most similar patches that compose a patch group together with the noisy patch.
3. Patch denoising: With low-rank criterion, SVD is applied to the patch group for removing noise. To get better results, the estimated patch corresponding to the noisy patch is obtained by summing all the patches in the denoised group and doing mean processing.
4. Aggregation: Because each pixel may be covered by several patches, we can reconstruct the first estimated image by weighted average. For every pixel in the noisy image, we should find the corresponding denoised pixels appearing in different estimated patches and weight those pixels.
5. Gabor transform: Any noise reduction algorithm will make image more or less lose partial key information, so is our method. To deal with this problem, and considering the layer characteristics of the OCT image, texture of the first estimate would be located by Gabor transform before second-step denoising.
6. Construct a new noisy image: A new noisy image is generated by adding some missing details back into

the first estimated image that is regarded as the input image for the second step.

7. Further denoising and get final estimate: Repeat steps 2 and 3 until all the noisy patches are denoised. Next, step 4 is performed, and the final estimated image is obtained.

2.2 Load Data

In this paper, the image data used in our experiments are separated into two parts: the noisy image and targeted database. Targeted database consists of the noisy image itself and other relevant external images similar to the former in structure. Figure 2(a) shows a noisy image. From Figs. 2(b)–2(d) several relevant images in the targeted database are shown. Comparing Fig. 2(a) with those three images, we can find that they are similar in layer information, location of layer information, number of layers, and range of gray value. Consequently, sufficient similar patches can be extracted for most noisy patches with the introduced external images, which can improve the denoising effect.

2.3 Find Similar Patches

In this section, we will find some most similar patches for each noisy patch in the noisy image. Because lots of patch comparison are needed, the similarity computation cost needs to be low. Euclidean distance is such a method in the spatial domain, which calculates the difference between two patches and is defined by Ref. 12

$$D(y_j, y_c) = \|y_j - y_c\|_2^2, \quad (1)$$

Fig. 2 Targeted database. (a) The noisy image. (b)–(d) Several relevant images in targeted database.

where $\| \cdot \|_2^2$ denotes the Euclidean distance, and y_j is the j 'th noisy patch with y_c as a corresponding candidate patch. As a matter of fact, the Eq. (1) represents difference of gray value between two patches, which means that the more similar y_j is to y_c , the smaller $D(y_j, y_c)$ is.

In the process of finding similar patches, there are two aspects that require attention. In the first place, transition interval from a noisy patch to next noisy patch ought to set, which should not be larger than the size of patch, otherwise, partial pixels are not covered by patches and the final estimate is distorted. In the next place, when an external image in targeted database is utilized to find similar patches, the Euclidean distances between all patches and the noisy patch are calculated. As described in Sec. 2.2, the noisy image is similar with the relevant image about location of layer information, accordingly a search window is constructed to restrict the region of finding similar patches in each relevant image, which reduce the running time.

In this paper, finding similar patches consists of two steps: first, Eq. (1) is applied to calculate all the Euclidean distances between the noisy patch and the candidate patches in the search window in a relevant image in the targeted database, respectively. The candidate patches are sorted with the corresponding distances in ascending order, and the first n candidate patches is selected and stored. Then repeat above operation in next relevant image until all the relevant images in targeted database are processed. Thus, N possible similar patches (formed by many groups of n candidate patches) are obtained. Second, the possible similar patches are sorted with the corresponding distances in ascending order, and the first m possible patches are selected as similar patches. The noisy patch and m patches are stored as a patch group matrix by treating each patch as a column of the group matrix. The patch group matrix P_j (consisting of $m + 1$ columns) corresponding to j 'th noisy patch is formed by

$$P_j = \{y_j, y_1, \dots, y_m\}. \quad (2)$$

Due to each patch in P_j contains noise, it can be represented as follows:

$$P_j = Q_j + N_j, \quad (3)$$

where Q_j denotes the noise-free group matrix, and N_j denotes the noise matrix.

2.4 Patch Denoising

Because similar patches for the noisy patch have been found in Sec. 2.3, we will apply functions to the group matrix and estimate noise-free version Q_j from noisy version P_j in this part. Ideally, the estimate $\hat{Q}$ should satisfy²⁸

$$\|P_j - \hat{Q}\|_F^2 = \sigma^2, \quad (4)$$

where $\| \cdot \|_F$ represents the Frobenius norm defined as the square root of the sum of the absolute squares of its elements, and σ is the standard deviation of noise. In this case, N_j is completely removed.

From Secs. 2.2 and 2.3, the noisy patch and the corresponding similar patches are basically the same, so they have a high correlation after denoising, which means the rank of Q_j is low. Thus, low-rank approximation is used to estimate Q_j by solving the following optimization problem:²²

$$\hat{Q}_j = \operatorname{argmin}_{Z_j} P_j - Z_{jF}^2, \quad \operatorname{rank}(Z_j) = k, \quad (5)$$

where $\operatorname{rank}(\cdot)$ denotes the rank of matrix. k can be calculated using Eq. (9). And we can solve the above equation by SVD, which is demonstrated by Ref. 18.

With SVD, P_j can be represented as follows:

$$P_j = U\Sigma V^T, \quad (6)$$

where Σ is a diagonal matrix and U, V are two orthonormal matrix. And the solution of Eq. (5) is defined as Ref. 29

$$\hat{Q}_j = P_{j,k} = U\Sigma_k V^T, \quad (7)$$

where Σ_k is obtained from the matrix Σ by keeping the first (largest) k singular values and setting the remaining diagonal elements to zeros.

$$\Sigma_k = \operatorname{diag}(s_1, \dots, s_k, 0, \dots, 0). \quad (8)$$

Next, we need to determine the value of k , which is defined by

$$\operatorname{sum}_k(\Sigma) \geq \eta \cdot \operatorname{sum}(\Sigma) > \operatorname{sum}_{k-1}(\Sigma), \quad (9)$$

where $\operatorname{sum}(\cdot)$ is the sum of the diagonal elements in Σ , $\operatorname{sum}_k(\cdot)$ is the sum of the first k diagonal elements in Σ . η is a tunable parameter (empirically set as 0.9). By the aforementioned formula, the denoised group matrix is obtained, which means each patch in group matrix is denoised. Our method performs denoising for the whole group matrix, rather than for a single patch in group matrix. In other words, $\hat{Q}_j$ is the ideal denoised result for the group matrix P_j , whereas the first column of $\hat{Q}_j$ may contain a small amount of noise and does not represent the ideal denoised patch for the noisy patch [from Eqs. (2), (6), and (7), the first column of $\hat{Q}_j$ only represents the intermediate result of the noisy patch]. So the estimated patch corresponding to the noisy patch is obtained by doing a mean processing

$$q_j = \operatorname{mean}(\hat{Q}_j), \quad (10)$$

where $\operatorname{mean}(\cdot)$ is a process of average on matrix by row and q_j , as the estimate of y_j , is a column vector.

2.5 Aggregation

Repeating the steps described in Secs. 2.3 and 2.4, all the estimated patches corresponding to the noisy patches in the noisy image are obtained. And there are overlapping parts between different patches, which mean that some noisy pixels in noisy patch are estimated for many times. As a result, these estimates of each pixel in the noisy image need to be aggregated, by performing a weighted averaging method with a certain noise reduction effect, to reconstruct the first estimated image. In general, the weighted averaging is assigning the uniform weights to all the estimates of each pixel, which may lead to an over-smoothed result. In this paper, for the j 'th group matrix $\hat{Q}_j$ [defined in Eq. (7)], the weighted formula is defined depending on the rank k and columns $m + 1$ (a noisy patch and its m similar patches) of each group matrix as Ref. 22

$$w_j = \begin{cases} 1 - \frac{k}{m+1}, & k < m+1, \\ \frac{1}{m+1}, & k = m+1. \end{cases} \quad (11)$$

In this paper, i is index of the pixel in the noisy image, and j is index of the noisy patch in the noisy image. If the group matrix is of full rank ($k = m + 1$), any correlation does not exist among patches and the uniform weights are used. If $k < m + 1$, there exists correlation in the group matrix. The smaller k is, the higher the group matrix is correlated and the better estimate of noisy patch using the low-rank technique is. The estimate for the i 'th pixel in the noisy image is expressed as follows:

$$\hat{x}_i = \frac{1}{W} \sum_{j \in \Gamma(x_i)} w_j \hat{x}_{i,j}, \quad (12)$$

where W is the sum of the used weights

$$W = \sum_{j \in \Gamma(x_i)} w_j, \quad (13)$$

where $\Gamma(x_i)$ represents the index set of all group matrices containing the pixel x_i , which is described as

$$\Gamma(x_i) = \{j | x_i \in Q_j, j = 1, \dots, C\}, \quad (14)$$

where x_i represents the i 'th pixel in the noisy image. $\hat{x}_{i,j}$ denotes the estimate of the pixel x_i in the j 'th denoised patch q_j . And C is the number of all group matrices. The first estimated image is reconstructed by the estimates of all pixels in the noisy image.

Figure 3(a) shows a noisy image. Figure 3(b) shows the first estimated image. Comparing the two images, we can see that most of noise in Fig. 3(a) is removed, but a little noise exists and partial detail is lost. So a further denoising is necessary.

2.6 Gabor Transform

Before the second-step denoising, it is necessary to construct a new noisy image. In general, the first-step denoised result is the input image of the second-step denoising. However, it is possible that image will lose some key detail after the first step. As shown in Fig. 3(b), the boundaries between layers become blurred; moreover, the contrast of background region and interest region is reduced. Thickness of retinal layers can reflect structural features of the eye tissue where pathological eyes are different with normal eyes about layer thickness, which is

Fig. 3 The original image and the denoised result. (a) The noisy image. (b) The first estimated image.

compared by Sayanagi et al.³⁰ We can determine the location of the boundaries and layers between the boundaries by Gabor transform in this section and add some missing detail to those regions in the next section.

Gabor transform³¹ is a special case of the short-time Fourier transform, which is used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time. The input image is first multiplied by a Gaussian function, which can be regarded as a window function, and the output image is transformed with a Fourier transform to derive the time-frequency analysis. So, Gabor transform can be used to extract relevant features at different scales and different directions and be used in texture recognition,³² which has achieved good results. In this section, Gabor transform is applied to each pixel for calculating its response values in different directions.

Figure 4(a) shows the structure of retinal layers. Retina is in the back of eye, which is composed of multilayer structure of thin and transparent organization. It is considered to be a

Fig. 4 Structure of retinal layers and some response images after applying Gabor transform. (a) Structure of retinal layers, (b) the response image after applying Gabor transform to Fig. 3(b) in horizontal direction, (c) in 30 deg direction, (d) in vertical direction, (e) the maximum response image after applying Gabor transform to Fig. 3(b), and (f) the location of layers between the boundaries in Fig. 3(b).

hierarchical structure between retinal nerve fiber layer and retinal pigment epithelium (RPE).³³ Inner limiting membrane (ILM) is the boundary of white area and black area at the most top of the retinal image. RPE is the brightest boundary since it has the highest reflectivity in the retinal image. The layers between ILM and RPE are ganglion cell layer, inner plexiform layer, inner nuclear layer, outer plexiform layer, outer nuclear layer, inner photosensitive layer, and outer photosensitive layer. It is clear that those boundaries are approximately horizontal. In other words, every boundary line has a small amplitude oscillation near the horizontal direction, which is set as θ (commonly equal to 45 deg).

All pixels' response values in a direction construct the response image along the direction. Figures 4(b)–4(d) show the response image after applying Gabor transform to Fig. 3(b) in horizontal direction, 30 deg direction, and vertical direction, respectively, which calculate response value for each pixel along the three directions. Because boundaries in OCT image are approximately horizontal, Fig. 4(b) can display some edges. From Figs. 4(b)–4(d), the response values of each pixel along different directions are different with each other. For each pixel in the image, we calculate all the response values along all the directions and maximum of the values is regarded as the maximum response value of the pixel. All the maximum response values can construct a maximum response image. Figure 4(e) shows the maximum response image after applying Gabor transform, which proves the accuracy of determining location of boundaries using Gabor transform, but cannot determine the location of layers.

On the basis of Fig. 4(e), determining the location of layers is a process of locating pixel where its direction corresponding to the maximum response value is at an angle smaller than θ to the horizontal direction. If a pixel meets this condition, its pixel value is set by 1; otherwise, its pixel value is set by 0. It leads to a binary image as shown in Fig. 4(f). Figure 4(f) also shows the location of layers represented by central white region, which is consistent with Fig. 3(b) about contour of layers. As for other white region and black region, they mean uninterested region with less layers information or without layers information, while it may seem a bit messy due to a little noise existing in Fig. 3(b). In fact, this does not affect our final result. Constructing a new noisy image in next section just add some missing details to pixels in location of layers, which can lead to much less noise in the new noisy image than the original noisy image. Comparing Fig. 3(a) with Fig. 3(b), most noise in Fig. 3(a) is removed, which prove the efficiency of denoising, and second-step denoising will achieve a better result.

2.7 Construct a New Noisy Image

In this section, a new noisy image is constructed by add some details eliminated in first step to pixels in location of layers.

$$y' = \hat{y}_1 + \mu(y - \hat{y}_1) \cdot \text{indexlayers}, \quad (15)$$

where $\hat{y}_1$ is the first estimated image, y is the original noisy image, and y' is the new noisy image. Location of layers is described as indexlayers , which represents Fig. 4(f). μ is a tunable parameter. When $\mu \rightarrow 0$, $y' \rightarrow \hat{y}_1$, which is usually adopted by other papers but leads to loss of image information. On the contrary, when $\mu \rightarrow 1$, pixels in location of layers are not denoised, and the second-step estimate is almost the same as the first estimate. In Ref. 22, Guo et al. make the improvement

Fig. 5 The new noisy image.

and set $\mu = 0.5$, which is a trade-off between 0 and 1. But in Guo's method, the missing details are added back into the whole image instead of layers, which cause all the background regions to be contaminated by noise again. However, in our method, we only add the missing details back into layer structures, which make the denoising effect on the background regions better, and can obtain higher contrast between background and foreground in the final denoised image. Figure 5 shows the new noisy image.

Generally, most denoising methods will remove some useful high-frequency signal when reducing high-frequency noises. In our method, the first estimated image can also miss some useful signal, but we can selectively add the missed structural information back to the first estimate, by setting the value of μ and using the Gabor based mask (the index_layers), when constructing a new noisy image for the second-step denoising. In the first step, the more the useful high-frequency signal in the noisy image is removed, the greater the value of μ is set.

2.8 Further Denoising and Get Final Estimate

After constructing a new noisy image, the second-step denoising can go on and the remaining sections are the same as the first step (as shown in the flowchart of Fig. 1). Figure 6 shows the final estimated image, which get better effects compared with the first estimate about edge preserving and contrast.

3 Results

In this section, we will show the experimental results of the proposed method and compared with other state-of-the-art methods. All the methods are implemented in MATLAB 2015a, executed on a desktop computer with an Intel(R) Core(TM) i5-4590M CPU (3.30 GHz) and 8 GB of RAM.

3.1 Comparison Methods

There are many advanced algorithms to achieve great achievements and progress in image denoising, including BM3D,¹² BM3D-SAPCA,³⁴ LPG-PCA,³⁵ NLM,¹⁰ UINTA,^{36,37} and TVSB,³⁸ they are chosen for comparison with our proposed method. Because the first four methods are internal denoising methods, for a fair comparison, we rewrite these methods into a denoising algorithm compatible with external denoising so that finding patch can be performed in the same targeted databases as ours. In addition to this, these methods are all divided

Fig. 6 The final estimated image.

into two steps where the first estimate in the first step is taken as the input of the second step, which exist a slight difference with our methods. Ultimately, the new four methods are external BM3D, external BM3D-SAPCA, external LPG-PCA, and external NLM, respectively. The last two state-of-the-art methods we compared are UINTA^{36,37} and TVSB.³⁸ We give a brief introduction of the aforementioned methods as follows:

- (1) BM3D: The block matching and 3-D filtering, first utilizing self-similarity and sparsity, groups similar patches into 3-D arrays to be dealt by a collaborative filtering. However, image patches containing edges singularities or textures cannot be sparsely represented using this method that may lead to visual artifacts.
- (2) BM3D-SAPCA: An improved BM3D filter, which exploits adaptive-shape patches and principle component analysis (PCA), achieves better denoising results than BM3D but is more time-consuming.
- (3) LPG-PCA: An adaptive image denoising scheme, using PCA and local pixel grouping, uses block matching to group the pixels with similar local structure and shrinks PCA transformation coefficients using the linear minimum mean square-error estimation technique.
- (4) NLM: The nonlocal mean filter, estimating each pixel by a nonlocal averaging of all the pixels in the image and with a weighting for a pixel based on the Euclidean distance, uses the structural redundancy and is considered as an extension of the bilateral filter.
- (5) UINTA: UINTA is an unsupervised information theoretic-based adaptive filtering method for image restoration. It estimates the intensity value of a pixel from its neighborhoods by decreasing the joint entropy between them. The probability used in entropy calculation is obtained through multivariate Parzen window estimation.
- (6) TVSB: It is based on Rudin–Osher–Fatemi model for total variation denoising, a general split Bregman framework is proposed to solve the involved L_1 -regularized optimization problem.

To show the advantage of our method over internal denoising, we will also compare the proposed method to internal PM (internal denoising version of the proposed method), which just enable the targeted database only to consist of the noisy image itself in our method. In addition, our proposed method is compared with PMNG (the proposed method when no Gabor transform is performed), which take the first estimate of the noisy image as the input of the second-step denoising.

3.2 Assessment Criteria

Evaluating the effectiveness of a method both calls for qualitative analysis and quantitative comparison. Specific assessment criteria are required for quantitative comparison where SNR is the most common choice in image denoising. SNR calculates the ratio between effective information and noise information

in an image, which need a clean image. But there is no standard noise-free image in the experiment, so SNR is not feasible for this paper. And, we can calculate the relevant standards of the local areas in a denoised image, which include two criteria: mean-to-standard-deviation ratio (MSR) and contrast-to-noise ratio (CNR). These two criteria are both proportional to the medical image quality. The larger values of them are, the higher quality of image is.

The metrics MSR³⁹ and CNR⁴⁰ are defined as follows:

$$\text{MSR} = \frac{\mu_f}{\sigma_f}, \quad (16)$$

$$\text{CNR} = \frac{|\mu_f - \mu_b|}{\sqrt{0.5(\sigma_f^2 + \sigma_b^2)}}, \quad (17)$$

where μ_f and μ_b are the mean of the foreground region (as #2 to 6 red boxes shown in Fig. 7) and background region (as #1 red boxes shown in Fig. 7), respectively. σ_f and σ_b are the standard deviation of the foreground region and background region, respectively.

For Eqs. (16) and (17), a clean image can be divided into foreground region and background region, and the grayscale value in each region tends to be uniform. But, for the noisy image, the grayscale values in regions tend to change rapidly, which make the standard deviations in the regions become greater. In general, foreground region contains structure information and mainly consists of a series of pixels with relatively large grayscale value in a clean image, whereas grayscale values in foreground region become smaller when the clean image is contaminated by noise. So the mean in noisy foreground region gets smaller than that in clean foreground region. For clean background region, its noisy version is the result that values of some pixels in the clean region become larger, which means that the mean in the noisy region is larger than in the clean region. In general, the mean in the foreground region is much larger than in the background region. This paragraph can be summarized into the following derivation process:

$$R_c \rightarrow R_n \Rightarrow \sigma_f \uparrow, \sigma_b \uparrow, \mu_f \downarrow, \mu_b \uparrow \Rightarrow \text{MSR} \downarrow, \text{CNR} \downarrow, \quad (18)$$

where $R_c \rightarrow R_n$ represents the process from clean region (foreground region or background region) to noisy region. $\uparrow$ represents the increase, and $\downarrow$ represents the decrease.

3.3 Local Denoising Effect and Quantitative Comparison

To evaluate the effectiveness of the proposed method, we compared its performance with eight well-known denoising approaches: external BM3D, external BM3D-SAPCA, external LPG-PCA, external NLM, internal PM, TVSB, UINTA, and PMNG. In these experiments, the parameters are set to the default values. Figure 7 shows the local denoising effect using different methods. Six regions, which are taken as objects for criteria calculation, are selected and outlined with red rectangle, one of which is background region marked with red number 1, and five of which are foreground regions marked from red number two to six. To observe the local denoising effect more clearly, three regions are magnified two times, which are connected to the new sized regions, respectively, by red line.

It is found that denoising effect of our proposed method is better than other four methods. BM3D uses the fixed 3-D basis

Fig. 7 (a) The original noisy image. The denoising results using (b) external BM3D, (c) external BM3D-SAPCA, (d) external LPG-PCA, (e) external NLM, (f) internal PM, (g) PMNG, (h) TVSB, (i) UINTA, and (j) our proposed method.

function that is less adapted to the edges and textures and leads to some artifacts as shown in Fig. 7(b). Figures 7(c) and 7(d) show the denoising results using BM3D-SAPCA and LPG-PCA methods, respectively. NLM is a nonlocal averaging method, which means that too much averaging can lead to detail loss in image as shown in Fig. 7(e). Comparing Fig. 7(e) with Fig. 7(j), contrast of the denoising result using our method between background and foreground is higher, which is the result that the lost details in the first step are added to the new noisy image by Gabor transform. Figure 7(f) is inferior to Fig. 7(j) in noise attenuation because not sufficient numbers of similar patches are found for many noisy patches using internal denoising. Figure 7(j) is better than Fig. 7(g) in preserving image details and layers information, resulting from the operation of Gabor transform. Figure 7(h) is the result of TVSB, which is also inferior to Fig. 7(j), the reasons may lie in the fact that TVSB is a pixelwise method (while our method, like NLM, is a patch-based method), and it cannot utilize the structure information embedded in a patch. In addition, it also ignores the self-similarity characteristic in an image. It has the advantage in speed as shown in Sec. 3.5. Figure 7(i) shows the result of UINTA. UINTA is also a patch-based method, and it can tune the bandwidth used in Parzen windowing for density estimation. It only uses the information in the noisy image itself without referring to the targeted database, and the denoising quality is lower than our result.

Usually, data calculated on a test image does not certainly indicate whether this method is effective or not, but we test data from multiple images. Quantitatively, MSR and CNR on six regions (as #1 to 6 red boxes shown in Fig. 7) from 20 noisy OCT images are measured at first. Next, we averaged the MSR and CNR values obtained from #2 to 6 red boxes shown in Fig. 7 in each image. Finally, we calculate the mean and the standard deviation of these averaged MSR and CNR results across all the test images in Table 1. From the table, the means of CNR and MSR calculated by our proposed method are largest, which demonstrate that our method achieves better noise reduction effect than other methods.

3.4 Comparison Under Different Patch Sizes

An important parameter in our experiment is patch size, which seriously affects noise reduction ability of our method. The determination of patch size is subject to a lot of factors, such as the size of noisy image, the level of noise, the structure characteristic of image, and so on. Next, we will determine the optimal patch size by experimental results. Figure 8 shows the denoising results using our proposed method under five different patch sizes. Obviously, the denoising results can be roughly divided into two categories, one of which has general noise reduction performance consisting of Figs. 8(b) and 8(c), and the other has better noise reduction performance consisting of

Table 1 Mean and standard deviation of the MSR and CNR results for 20 noisy OCT images using external BM3D, external BM3D-SAPCA, external LPG-PCA, external NLM, internal PM, PMNG, TVSB, UINTA, and the proposed method.

	Mean (CNR)	Standard deviation (CNR)	Mean (MSR)	Standard deviation (MSR)
Original	0.9320	0.3690	2.7404	0.2343
eBM3D	1.5088	0.5178	4.6390	0.3141
eBM3D-SAPCA	1.0653	0.4068	2.9931	0.2076
eLPG-PCA	1.0238	0.3964	2.9365	0.2159
eNLM	1.6833	0.5185	5.7581	0.5550
iPM	1.6453	0.5764	5.7259	0.4533
PMNG	1.7365	0.5854	6.0469	0.5362
TVSB	1.7159	0.5997	5.9057	0.5894
UINTA	1.6229	0.4880	5.1190	0.4342
Our method	1.8173	0.6003	6.4262	0.5748

Note: The best results in the table are shown in bold.

Figs. 8(d)–8(f). Six regions are circled with red rectangle, which is similar to the previous work. It is difficult to distinguish which is best in noise reduction among Figs. 8(d)–8(f) only through visual effect, due to a very small difference existing among them. The optimal patch size is decided by calculating the relevant criteria values under different patch sizes and making a comparison.

Here, we have to make a quantitative comparison similar to Table 1. MSR and CNR on six regions (as #1 to 6 red boxes shown in Fig. 8) from 20 noisy OCT images are measured at first. Next, we averaged the MSR and CNR values obtained from #2 to 6 red boxes shown in Fig. 7 in each image. Finally, we calculate the mean and the standard deviation of these averaged MSR and CNR results across all the test images in Table 2. From the table, MSR and CNR get maximum values under 7×9 patch size, which is taken as the optimal patch size in our method.

3.5 Overall Visual Comparison and Running Time Comparison

In this section, a global comparison is presented, which is helpful for us to understand the noise reduction effect of each method directly visually. The facts show that our method outperforms the other several methods. Figure 9 shows the denoising results using different methods.

Running times for denoising 20 noisy OCT images are measured. We calculate the mean of these running times across all the test images in Table 3. From the table, we can see that the mean running time of the proposed method is longer than most of other methods, but this is acceptable since our method has better robustness according to Table 1 and achieved better denoising effect (as shown in Fig. 9). Our method is currently implemented in MATLAB, and the running time can be greatly reduced when implemented with C or C++ language.

Fig. 8 (a) The original noisy image. The denoising results using our proposed method under (b) 5×7 , (c) 6×6 , (d) 7×9 , (e) 8×8 , (f) 9×11 patch sizes, respectively.**Table 2** Mean and standard deviation of the MSR and CNR results for twenty noisy OCT images using our proposed method under five patch sizes: 5×7 , 6×6 , 7×9 , 8×8 , and 9×11 .

	Mean (CNR)	Standard deviation (CNR)	Mean (MSR)	Standard deviation (MSR)
Original	1.1803	0.2604	2.9296	0.2101
5×7	2.0096	0.3771	5.2106	0.5553
6×6	2.0222	0.3717	5.2725	0.5993
7×9	2.4498	0.4520	6.3256	0.9662
8×8	2.3071	0.4358	5.9804	0.8459
9×11	2.3723	0.4339	6.1910	0.8676

Note: The best results in the table are shown in bold.

Fig. 9 (a) The original noisy image, (b) the denoising result using external BM3D, (c) external BM3D-SAPCA, (d) external LPG-PCA, (e) external NLM, (f) internal PM, (g) PMNG, (h) TVSB, (i) UINTA, and (j) our proposed method.

The computation time for UINTA is the longest even it is implemented in C++,⁴¹ for each pixel, since it is an iterative version on the NLM, the iteration is time-consuming. TVSB spend the shortest time.

Table 3 Mean of running time for denoising 20 noisy OCT images using external BM3D, external BM3D-SAPCA, external LPG-PCA, external NLM, internal PM, TVSB, UINTA, and the proposed method.

Method	Mean (time) (s)
eBM3D	140
eBM3D-SAPCA	162
eLPG-PCA	163
eNLM	123
iPM	242
TVSB	107
UINTA	459
Our method	346

4 Conclusion

In this paper, we proposed a patch-based method for image denoising. Taking source of similar patches into consideration, a combination of internal denoising and external denoising is adopted when searching for similar patches. To denoise each patch, the low-rank technique is employed, and that is equivalent to performing SVD for patch groups by preserving only a few largest singular values due to the optimal energy compaction property of SVD. Creating a new noisy image before the second-step denoising takes advantage of Gabor transform to locate texture, which consider directional distribution characteristics of layer information in OCT image. The performance is improved when rectangular patch is selected instead of the usual square patch. Experiment demonstrates that our proposed method compares favorably with the existing state-of-the-art methods.

Disclosures

The authors, Zhou Yang, Jia Wang, and Wei Hu, have no financial disclosures.

Acknowledgments

Xiaoming Liu was partially supported by the National Natural Science Foundation of China (Nos. 61403287, 61572381, 61273303), China Postdoctoral Science Foundation

(No. 2014M552039) and the Natural Science Foundation of Hubei Province (No. 2014CFB288). Jun Liu was supported by the National Natural Science Foundation of China (No. 31201121). Kai Zhang was supported by the National Natural Science Foundation of China (No. 61472293).

References

1. D. Huang et al., "Optical coherence tomography," *Science* **254**, 1178–1181 (1991).
2. A. F. Fercher, "Optical coherence tomography," *J. Biomed. Opt.* **1**, 157–173 (1996).
3. A. F. Fercher et al., "Optical coherence tomography-principles and applications," *Rep. Prog. Phys.* **66**, 239–303 (2003).
4. H. G. Bezerra et al., "Intracoronary optical coherence tomography: a comprehensive review: clinical and research applications," *JACC: Cardiovasc. Interventions* **2**, 1035–1046 (2009).
5. H. Pahlevaninezhad et al., "Endoscopic Doppler optical coherence tomography and autofluorescence imaging of peripheral pulmonary nodules and vasculature," *Biomed. Opt. Express* **6**, 4191–4199 (2015).
6. A. I. Kholodnykh et al., "Precision of measurement of tissue optical properties with optical coherence tomography," *Appl. Opt.* **42**, 3027–3037 (2003).
7. J. Schmitt and A. Knüttel, "Model of optical coherence tomography of heterogeneous tissue," *J. Opt. Soc. Am. A* **14**, 1231–1242 (1997).
8. Y. Zhang, *Optical Coherence Tomography Guided Laser-Cochleostomy*, KIT Scientific Publishing, Karlsruhe, Germany (2015).
9. J. M. Schmitt, S. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *J. Biomed. Opt.* **4**, 95–105 (1999).
10. A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.* **4**, 490–530 (2005).
11. M. Antonini et al., "Image coding using wavelet transform," *IEEE Trans. Image Process.* **1**, 205–220 (1992).
12. K. Dabov et al., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**, 2080–2095 (2007).
13. C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Trans. Image Process.* **15**, 2866–2878 (2006).
14. M. Zontak and M. Irani, "Internal statistics of a single natural image," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 977–984 (2011).
15. I. Mosseri, M. Zontak, and M. Irani, "Combining the power of internal and external denoising," in *IEEE Int. Conf. on Computational Photography (ICCP '13)*, pp. 1–9 (2013).
16. D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *IEEE 12th Int. Conf. on Computer Vision*, pp. 349–356 (2009).
17. A. Levin and B. Nadler, "Natural image denoising: optimality and inherent bounds," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 2833–2840 (2011).
18. E. Luo, S. H. Chan, and T. Q. Nguyen, "Adaptive image denoising by targeted databases," *IEEE Trans. Image Process.* **24**, 2167–2181 (2015).
19. H. C. Burger, C. Schuler, and S. Harmeling, "Learning how to combine internal and external denoising methods," in *German Conf. on Pattern Recognition*, Springer, pp. 121–130 (2013).
20. N. Srebro and T. Jaakkola, "Weighted low-rank approximations," in *Int. Conf. on Machine Learning (ICML '03)*, pp. 720–727 (2003).
21. T. Li et al., "Image denoising using low-rank dictionary and sparse representation," in *Tenth Int. Conf. on Computational Intelligence and Security (CIS '14)*, pp. 228–232 (2014).
22. Q. Guo et al., "An efficient SVD-based method for image denoising," *IEEE Trans. Circuits Syst. Video Technol.* **26**, 868–880 (2016).
23. L. Fang et al., "Sparsity based denoising of spectral domain optical coherence tomography images," *Biomed. Opt. Express* **3**, 927–942 (2012).
24. D. Thapa, K. Raahemifar, and V. Lakshminarayanan, "A new efficient dictionary and its implementation on retinal images," in *19th Int. Conf. on Digital Signal Processing (DSP '14)*, pp. 841–846 (2014).
25. A. Baghaie, R. M. D'souza, and Z. Yu, "Sparse and low rank decomposition based batch image alignment for speckle reduction of retinal OCT images," in *IEEE 12th Int. Symp. on Biomedical Imaging (ISBI '15)*, pp. 226–230 (2015).
26. Z. Jian et al., "Three-dimensional speckle suppression in optical coherence tomography based on the curvelet transform," *Opt. Express* **18**, 1024–1032 (2010).
27. D. C. Adler, T. H. Ko, and J. G. Fujimoto, "Speckle reduction in optical coherence tomography images by use of a spatially adaptive wavelet filter," *Opt. Lett.* **29**, 2878–2880 (2004).
28. B. Carpentieri, I. S. Duff, and L. Giraud, "Sparse pattern selection strategies for robust Frobenius-norm minimization preconditioners in electromagnetism," *Numer. Linear Algebra Appl.* **7**, 667–685 (2000).
29. C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika* **1**, 211–218 (1936).
30. K. Sayanagi, S. Sharma, and P. K. Kaiser, "Comparison of retinal thickness measurements between three-dimensional and radial scans on spectral-domain optical coherence tomography," *Am. J. Ophthalmol.* **148**, 431–438 (2009).
31. E. Sejdić, I. Djurović, and J. Jiang, "Time—frequency feature representation using energy concentration: an overview of recent advances," *Digital Signal Process.* **19**, 153–183 (2009).
32. J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. Acoust. Speech Signal Process.* **36**, 1169–1179 (1988).
33. M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Rev. Biomed. Eng.* **3**, 169–208 (2010).
34. K. Dabov et al., "BM3D image denoising with shape-adaptive principal component analysis," in *Signal Processing with Adaptive Sparse Structured Representations (SPARS'09)* (2009).
35. L. Zhang et al., "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognit.* **43**, 1531–1549 (2010).
36. S. P. Awate and R. T. Whitaker, "Higher-order image statistics for unsupervised, information-theoretic, adaptive, image filtering," in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 44–51 (2005).
37. S. P. Awate and R. T. Whitaker, "Unsupervised, information-theoretic, adaptive image filtering for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.* **28**, 364–376 (2006).
38. T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imag. Sci.* **2**, 323–343 (2009).
39. G. Cincotti, G. Loi, and M. Pappalardo, "Frequency decomposition and compounding of ultrasound medical images with wavelet packets," *IEEE Trans. Med. Imaging* **20**, 764–771 (2001).
40. P. Bao and L. Zhang, "Noise reduction for magnetic resonance images via adaptive multiscale products thresholding," *IEEE Trans. Med. Imaging* **22**, 1089–1099 (2003).
41. Kitware Inc., "Patch based denoising image filter," ITK Insight Segmentation and Registration Toolkit, 2016, https://itk.org/Doxygen42/html/group__ITKDenosing.html (2 June 2016).

Xiaoming Liu received his PhD from the College of Computer Science and Technology at Zhejiang University, China, in 2007. He has been a visiting scholar in the Image Display, Enhancement, and Analysis (IDEA) Group at the University of North Carolina at Chapel Hill (UNC). He is a professor with the College of Computer Science and Technology, Wuhan University of Science and Technology, China. His research interests include medical image processing, pattern recognition, and machine learning.

Zhou Yang is now a postgraduate student in the College of Computer Science and Technology at Wuhan University of Science and Technology in China. His research interests mainly include medical image processing and pattern recognition.

Jia Wang is currently a postgraduate student in the College of Computer Science and Technology at Wuhan University of Science and Technology in China. His research interests mainly include medical image processing and machine learning.

Jun Liu received his PhD from the Wuhan University of Science and Technology, China, in 2012. He has been a visiting scholar at the

University of Michigan at Ann Arbor in 2016. He is an associate professor at the College of Computer Science and Technology, Wuhan University of Science and Technology, China. His research interests include ultrasound image analysis, medical image processing, and pattern recognition.

Kai Zhang received the PhD from Huazhong University of Science and Technology in 2008. He received his postdoctoral training at Peking University from 2008 to 2010. Currently, he is a professor in the School of Computer Science and Technology at Wuhan

University of Science and Technology. His research interests include intelligent computing and combinatorial optimization.

Wei Hu received his master's and PhD degrees in computer science from Zhejiang University in 2005 and 2008, respectively. He was an assistant professor at Zhejiang University from 2008–2010. He is currently associate professor at Wuhan University of Science and Technology. His current research interests include medical image processing, computer architecture, and embedded systems.