

Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

Optical coherence tomography retinal image reconstruction via nonlocal weighted sparse representation

Ashkan Abbasi
Amirhassan Monadjemi
Leyuan Fang
Hossein Rabbani

SPIE.

Ashkan Abbasi, Amirhassan Monadjemi, Leyuan Fang, Hossein Rabbani, "Optical coherence tomography retinal image reconstruction via nonlocal weighted sparse representation," *J. Biomed. Opt.* **23**(3), 036011 (2018), doi: 10.1117/1.JBO.23.3.036011.

Optical coherence tomography retinal image reconstruction via nonlocal weighted sparse representation

Ashkan Abbasi,^a Amirhassan Monadjemi,^{a,*} Leyuan Fang,^{b,*} and Hossein Rabbani^c

^aUniversity of Isfahan, Department of Artificial Intelligence, Faculty of Computer Engineering, Isfahan, Iran

^bHunan University, College of Electrical and Information Engineering, Changsha, China

^cIsfahan University of Medical Sciences, School of Advanced Technologies in Medicine, Medical Image and Signal Processing Research Center, Department of Biomedical Engineering, Isfahan, Iran

Abstract. We present a nonlocal weighted sparse representation (NWSR) method for reconstruction of retinal optical coherence tomography (OCT) images. To reconstruct a high signal-to-noise ratio and high-resolution OCT images, utilization of efficient denoising and interpolation algorithms are necessary, especially when the original data were subsampled during acquisition. However, the OCT images suffer from the presence of a high level of noise, which makes the estimation of sparse representations a difficult task. Thus, the proposed NWSR method merges sparse representations of multiple similar noisy and denoised patches to better estimate a sparse representation for each patch. First, the sparse representation of each patch is independently computed over an overcomplete dictionary, and then a nonlocal weighted sparse coefficient is computed by averaging representations of similar patches. Since the sparsity can reveal relevant information from noisy patches, combining noisy and denoised patches' representations is beneficial to obtain a more robust estimate of the unknown sparse representation. The denoised patches are obtained by applying an off-the-shelf image denoising method and our method provides an efficient way to exploit information from noisy and denoised patches' representations. The experimental results on denoising and interpolation of spectral domain OCT images demonstrated the effectiveness of the proposed NWSR method over existing state-of-the-art methods.

© 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JBO.23.3.036011](https://doi.org/10.1117/1.JBO.23.3.036011)]

Keywords: optical coherence tomography; image reconstruction; sparse representation; denoising; interpolation; superresolution.

Paper 180013R received Jan. 12, 2018; accepted for publication Mar. 6, 2018; published online Mar. 24, 2018.

1 Introduction

Optical coherence tomography (OCT) is a noninvasive imaging modality, which can provide a cross-sectional view (tomography) of the tissue structures and has been widely used to identify and monitor various ophthalmology diseases.¹ However, due to interferometry nature of OCT imaging, the OCT images are inevitably affected by noise. Also, to avoid motion artifacts from the fixation eye movements, the OCT clinical images are often captured at lower than nominal sampling rates.^{2–6} Therefore, effective denoising and interpolation algorithms are necessary for automated or even manual OCT image analysis.

Interpolation and denoising are two of the most well-known problems in image processing,⁷ and many methods based on various image models have been proposed for OCT image reconstruction.^{8–19} In general, these methods are mostly developed for OCT denoising. Although they can be used with efficient interpolation algorithms, designing a more unified algorithm that is capable of denoising, and interpolation is more appealing.^{5,6,20}

Classic interpolation algorithms produce unsatisfying results when the low-resolution image is noisy and they cannot recover features that are missed in the input noisy image itself.^{4–6,20} Multiframe interpolation methods seems more plausible since

they benefit from more information provided by multiple images.² However, they need motion estimation from images, which is prone to error even for noise-free images.^{8,10,13,14,21,22}

Example-based learning interpolation can exploit the information from a dataset of paired training images. The early approaches are based on finding the nearest neighbors in a large dataset consisting of low-resolution (LR) and high-resolution (HR) patch pairs.²¹ Despite promising results that have been presented, this approach is computationally intensive and the search for nearest neighbors may be negatively affected by noise. These shortcomings can be mitigated using sparse models.^{23–27} Thus, several researchers have used sparse representation and closely related compressed sensing methods to reconstruct OCT data.^{2–6,28–30}

Recently, Fang et al.^{5,6} have tried to denoise and interpolate spectral domain OCT (SDOCT) images by proposing a fast coupled dictionary learning (DL) approach. Basically, this approach is based on the idea of shared sparse representations between LR and HR patches.^{24,31} However, learning coupled dictionaries in presence of noise can mislead estimating supports or coefficient values in sparse representation stage.^{20,32} Therefore, sophisticated regularizers are needed with the cost of higher computational complexity.²⁰ An alternative way is based on using sparse representation of the degraded patch itself. Several researchers have exploited this idea, along with

*Address all correspondence to: Amirhassan Monadjemi, E-mail: Monadjemi@eng.ui.ac.ir; Leyuan Fang, E-mail: fangleiyuan@gmail.com

nonlocal redundancy in natural images.^{25,33–36} To interpolate a natural image with sparse nonlocal models, learning one model for each subspace through clustering has been shown promising results,^{25,33,36} but clustering highly noisy OCT image patches is not easy and inconsistency between LR and HR clusters negatively affects the reconstruction quality.⁶ To tackle with heavier and more realistic noise, two interpolation methods based on the well-known block-matching and 3D filtering (BM3D) sparse representation model are recently introduced.^{37,38} Although the reconstruction results are excellent, these methods use internal BM3D dictionaries that are learned from the corrupted image itself. Therefore, it would be hard to recover details that are lost in the LR image itself.

In this paper, we propose a nonlocal weighted sparse representation (NWSR) to denoise or interpolate a single cross-sectional SDOCT image (also called B-scan). We learn an overcomplete dictionary from an external dataset of SDOCT images and use the learned dictionary in our reconstruction algorithm to enhance the spatial resolution of an input image. Two cues have been used to better estimate sparse representation of an unknown HR patch from noisy LR patches. First, inspired by sparse nonlocal models,^{33–35} we incorporate self-similarity information to obtain a representation for each patch by averaging sparse representations of nonlocal similar patches over an overcomplete dictionary. The sparse representation of each patch is computed independently, and then a weighted sparse coefficient is computed, where weights are determined by the similarity between patches.³⁹ Second, promising results have been shown recently by combining information that are extracted from noisy image and the result of an off-the-shelf denoising method.⁴⁰ Sparse representation can identify relevant information from noisy image patches while suppressing noise. Therefore, we combine noisy and denoised patches' representations to calculate nonlocal weighed sparse coefficients. This results in a more robust estimate of the HR patch representation, which is later used for denoising or interpolation of SDOCT images.

The rest of the paper is organized as follows. In Sec. 2, we briefly review the sparse representation-based denoising model and interpolation model. In Sec. 3, we introduce the proposed NWSR method for the denoising and interpolation of OCT images. Experimental results on clinical OCT data are shown in Sec. 4. The conclusion and future works are presented in Sec. 5.

2 Related Works

2.1 Sparse Representation-Based Denoising

The sparse representation assumes that most signals of interest can be estimated using a linear combination of a few (sparse) elements from a set of basis functions (atoms) that is called a dictionary.²³ Ideally, in the sparse model, corruptions such as noise cannot be well represented sparsely, thus finding the sparse representation of a corrupted signal over a suitable dictionary helps to restore it. To exploit the sparse model, a corrupted image \mathbf{Y} is divided into overlapping small patches. Each patch $Y_i \in \mathbb{R}^{p_1 \times p_2}$ is converted to a column vector $y_i \in \mathbb{R}^{n=p_1 p_2}$ by lexicographic ordering.⁴¹ To avoid numerical instabilities of computing sparse representation, the mean intensity value of each patch is often subtracted from its pixel values.³⁵ Then, the vector y_i is represented sparsely by a coefficient vector $\alpha_i \in \mathbb{R}^k$ over a dictionary matrix $D \in \mathbb{R}^{n \times k}$ by solving the following problem:

$$\min_{\alpha_i} \|y_i - D\alpha_i\|_2^2, \quad \text{s.t. } \|\alpha_i\|_0 < t, \quad (1)$$

where $\|\cdot\|_0$ is the L_0 pseudonorm that counts the number of non-zero elements (support) in the coefficient vector α_i . The constant t is the sparsity level. Assuming that we have an overcomplete dictionary D with $k \gg n$; A pursuit algorithms such as orthogonal matching pursuit (OMP)²³ can be used to approximately solve Eq. (1) such that $x_i \approx D\alpha_i$. After adding the removed mean intensity to each estimate $D\alpha_i$, it could be interpreted as a restored version of Y_i . The dictionary D can be constructed²⁰ or learned from a dataset of image patches; however, it is more desirable to learn the dictionary, because it leads to a more compact one which speeds up the computation of sparse representations.^{24,25,41}

2.2 Sparse Representation-Based Interpolation

An instance of interpolation problem can be formulated as follows: denote an original HR image as $\mathbf{X} \in \mathbb{R}^{r \times c}$, the down-sample operator as S , and the subsampled LR image as $\mathbf{Z} = \mathbf{S}\mathbf{X} \in \mathbb{R}^{(rs) \times (cs)}$. Given a subsampled observed image \mathbf{Z} , the goal of any image interpolation method is to estimate \mathbf{X} by $\hat{\mathbf{Z}}$ such that $\hat{\mathbf{Z}} \sim \mathbf{X}$. To solve this problem with a patch-based sparse model, we need a way to estimate the latent sparse representation of an HR patch from the observed LR patch.

Two main strategies have been used in different methods for this estimation: (1) estimating the sparse representation of an HR patch via the remaining pixels, i.e., only the observed pixels in the given LR patch and (2) estimating the sparse representation of an HR patch by assuming that the sparse representation of the LR patch and its corresponding HR one are similar over a couple of dictionaries, i.e., (D_L, D_H) where D_L is the LR dictionary and D_H is the HR dictionary. The second was pioneered by Yang et al.^{24,31} However, as mentioned earlier, learning coupled dictionaries in the presence of noise are not easy.³²

The first strategy needs one dictionary. At first, rows of the dictionary D that are corresponding to the missing pixels in a given patch y_i are removed by applying a suitable matrix on the dictionary $D'_i = U_i D$. The columns of D'_i are usually normalized to unit one. Then, the sparse representation α_i of an image patch z_i can be found over this new dictionary by solving

$$\min_{\alpha_i} \|z'_i - D'_i \alpha_i\|_2^2 \quad \text{s.t. } \|\alpha_i\|_0 < t, \quad (2)$$

where $z'_i = U_i z_i$ is a vector corresponding to the observed pixels of the patch z_i , and the HR patch x_i is reconstructed by $x_i \sim \hat{x}_i = D\alpha_i$.

3 Proposed NWSR Method for OCT Image Denoising and Interpolation

Given an observed SDOCT image \mathbf{Y} , we want to estimate a denoised HR image \mathbf{X} based on \mathbf{Y} . When \mathbf{Y} and \mathbf{X} have the same spatial resolution, i.e., $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{r \times c}$ the problem is denoising. Following,^{5,6} a spatially subsampled noisy SDOCT image has fewer columns and we can formulate it as $\mathbf{Z} = \mathbf{S}\mathbf{Y} \in \mathbb{R}^{r \times (cs)}$ where S is the column downsampler. Thus, to infer the contents of the denoised HR image \mathbf{X} from the noisy image \mathbf{Y} or the noisy downsampled image \mathbf{Z} ; a dictionary and an image reconstruction algorithm are required.

3.1 Dictionary Learning

To learn a dictionary, we need high signal-to-noise ratio (SNR) SDOCT images that are acquired in works.^{5,16} In the training part of their dataset, there are 10 high-SNR-high-resolution (HH) images. We use these images and extract N overlapping patches of size $p_1 \times p_2$ pixels. For each patch $X_i \in \mathbb{R}^{p_1 \times p_2}$, as described in Sec. 2.1, we convert it to a vector $x_i \in \mathbb{R}^{n=p_1 p_2}$ and remove the mean of its pixels intensity. Then, we solve the following DL problem to learn an overcomplete dictionary $D \in \mathbb{R}^{n \times K}$:

$$\min_{D, \alpha_i} \|\alpha_i\|_0 \quad \text{s.t.} \quad \sum_{i=1}^N \|D\alpha_i - x_i\|_2 \leq \epsilon. \quad (3)$$

In Eq. (3), the representation error ϵ is set to be $\sigma g \sqrt{p_1 \times p_2}$ for all patches,⁴² where σ indicates the level of noise in the training data, and g is a predefined constant, which is called noise gain. We used efficient implementation of K-SVD and batch-OMP algorithm, which is specifically optimized for calculating sparse representations of large sets of signals over a dictionary.⁴³

3.2 Image Reconstruction

Supposing that a noisy LR image \mathbf{Z} is fed to the algorithm as the input image, we densely extract overlapping patches of size $p_1 \times p_2$ from the input image. Then, each patch is processed with our reconstruction algorithm that mainly comprises three steps: (1) for each patch, find similar patches with the help of corresponding patches that are extracted from a denoised version of an input image, (2) compute nonlocal weighted sparse coefficients over the learned dictionary, and (3) reconstruct the whole image by aggregating overlapping estimates. The schematic diagram of our algorithm is shown in Fig. 1 and each step will be described in more detail in the following subsections.

3.2.1 Finding similar patches

Given a patch from the observed image \mathbf{Z} , we want to find similar patches for it. To compare patches, the Euclidean distance can be used; but this distance on the noisy patches is prone to overfitting the noise and return irrelevant results.⁴⁴ Therefore, we use an off-the-shelf denoising algorithm to reduce the amount of noise in \mathbf{Z} . Though any denoising algorithms can be used, we apply an implementation of the nonlocal mean (NLM) filtering, which is provided by Ref. 5.

Concretely, to find K similar patches for a given noisy patch z_i , $i = 1, \dots, M$; we use the corresponding denoised patch \hat{z}_i extracted from the denoised image $\hat{\mathbf{Z}}$. Then, we extract L patches around it, which are in the same row and retrieve the indices of K most similar patches (including the patch itself). The distance is computed as

$$d_{i,j} = (U_i \hat{z}_i - U_i \hat{z}_j)^T (U_i \hat{z}_i - U_i \hat{z}_j), \quad \text{where } j \in R_i. \quad (4)$$

In Eq. (4), the matrix U_i extracts the observed pixels from \hat{z}_i and \hat{z}_j , respectively. The set R_i denotes the indices of patches that have missing values in the same locations as the given patch \hat{z}_i in the neighborhood of size L . Note that, if the patches have different locations of missing values, the sparse representation stage will try to capture information that are relevant for reconstructing those values. Therefore, it may contribute wrongly for the computation of the sparse representation of the current patch in process.

3.2.2 Computing nonlocal weighted sparse coefficients

Given K similar patches for each degraded patch z_i , we combine information from them to obtain a better estimation of sparse representation of the corresponding HR patch. First, we independently compute sparse representation of the patch z_i and its similar denoised patches using Eq. (2). Let us indicate the set containing sparse representations of K patches for the i 'th

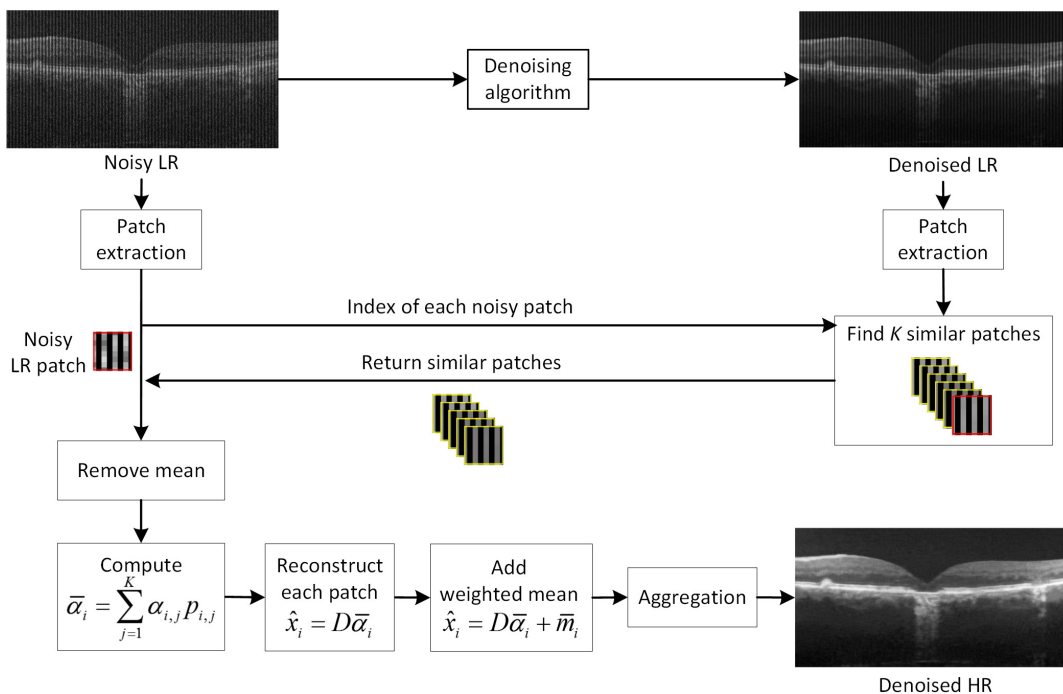


Fig. 1 A schematic diagram of the proposed NWSR method.

patch as $S_i = \{\alpha_{i,1}, \dots, \alpha_{i,K}\}$. Next, a nonlocal weighted sparse coefficient is computed using

$$\bar{\alpha}_i = \sum_{j=1}^K \alpha_{i,j} p_{i,j}. \quad (5)$$

In this equation, $p_{i,j}$ is the probability associated with the sparse representation of the j 'th patch in the set S_i . We assign the probabilities by considering the degree of similarity between the given patch and its similar ones.³⁹ To this end, we use the distance between patches $d_{i,j}$ that are computed in the previous section [Eq. (4)] to define a weight for each patch

$$w_{i,j} = e^{-\frac{d_{i,j}}{2h^2}}. \quad (6)$$

Obviously, the maximum weight is assigned to the sparse representation of the given patch (z_i). The parameter h is a constant that controls the amount of deviation from the i 'th patch. If it is set to a very small value, the patches must be very similar to the i 'th patch to have a significant weight. Equation (7) converts the weights into probabilities and ensures that the probabilities sum to one

$$p_{i,j} = \frac{w_{i,j}}{\sum_{j=1}^K w_{i,j}}. \quad (7)$$

After computing nonlocal weighted sparse coefficients $\bar{\alpha}_i$ for each patch z_i , we can denoise and interpolate it by $D\bar{\alpha}_i$ (Fig. 1).

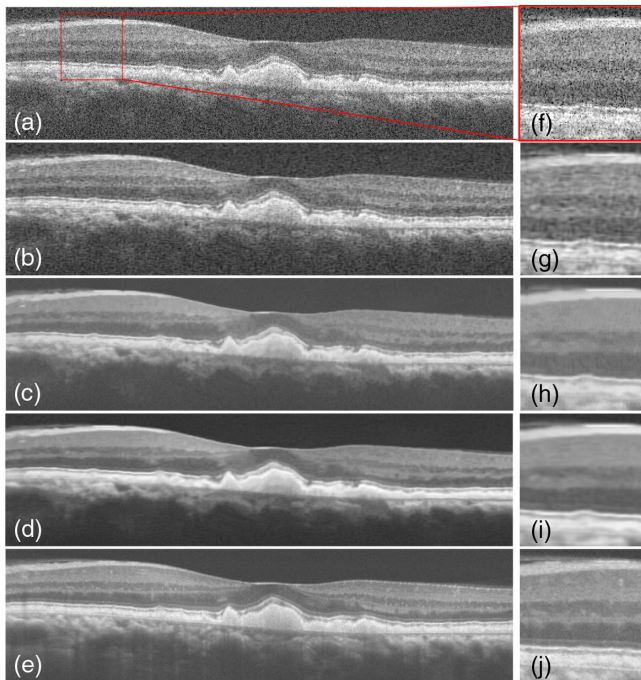


Fig. 2 The effect of the proposed NWSR method for denoising a retinal OCT image. The second column shows a magnified region. (a), (f) Original noisy image; (b), (g) output without using any similar patches for computing sparse representation of each patch; (c), (h) output of the off-the-shelf NLM denoising; (d), (i) output of the proposed NWSR method using five similar patches for each patch to compute its sparse representation; and (e), (j) registered and averaged images. Note that the denoising method [(c) and (h)] produces artifacts and the image is still noisy. The figure is better seen by zooming on a computer screen.

The effect of the proposed NWSR method for image reconstruction is shown in Fig. 2. In the next step, we restore the mean of the patches that were removed during sparse representations, and then merge the patches to reconstruct a whole denoised HR image.

3.2.3 Reconstructing the whole image

Let us indicate the estimation of the i 'th patch with $\hat{x}_i = D\bar{\alpha}_i$. This estimation lacks the mean of intensities for each patch as it was removed in the sparse representation stage. One way to restore the mean is to use the mean of the remaining pixels in each patch.⁴² A better way is to use the mean of the corresponding denoised patch.⁵ As we exploit sparse representations of multiple patches to obtain nonlocal weighted sparse coefficients for each patch, it is natural to expect that the mean can also be recovered in a similar weighted fashion. We compute the mean intensity for the i 'th patch by

$$\bar{m}_i = \sum_j m_j p_j. \quad (8)$$

Then, the mean intensity is added to the estimation by

$$\hat{x}_i = D\bar{\alpha}_i + \bar{m}_i. \quad (9)$$

The whole procedure of restoring each patch is briefly described in Fig. 3. After restoring each patch, we have multiple estimates for each pixel. A common method to form a complete image from a patch-based processing method is to average these estimates by⁴¹

$$\hat{x} = W \left(\sum_{i=1}^M R_i^T \hat{x}_i \right) \quad \text{where, } W = \left(\sum_{i=1}^M R_i^T R_i \right)^{-1}. \quad (10)$$

In Eq. (10), $R_i^T \hat{x}_i$ places the i 'th estimated patch in the appropriate place. The diagonal matrix W contains the weights associated with each pixel based on the number of overlapping patches that reconstruct it. The vector \hat{x} is the recovered image.

4 Experimental Results

In this section, we present some experimental results of the proposed NWSR method for both interpolation and denoising problems. The datasets, parameters, and quantitative image reconstruction metrics are explained. Then, the method is compared quantitatively and qualitatively with other competing methods. The source code of our proposed NWSR method will be released on the website (<https://github.com/ashkan-abbasi66>).

4.1 Datasets

We use datasets that were originally introduced by Ref. 5, and it is available in Ref. 45. The images were all captured by Bioptigen SDOCT imaging systems (Durham, North Carolina). The first dataset was acquired from the central foveal images of 28 eyes of 28 subjects with and without nonneovascular age macular degeneration. For each subject, an HH image was provided by registration of azimuthally repeated B-scan images from the fovea.⁵ The dataset was divided randomly into a training set of 10 subjects and a test set of 18 subjects. The second dataset consists of really subsampled images with 450 columns per image. This dataset was acquired with the same imaging device from 13 human subjects.

Patch Restoration algorithm via NWSR

Input: $\{z_{i,1}, z_{i,2}, \dots, z_{i,K}\}$, a set consist of the i th patch ($z_{i,1} = z_i$) and its similar ones, D , the dictionary.

1. Remove missing values from the patches by $z'_{i,j} = U_i z_{i,j}$, $j = 1, \dots, K$.
2. Remove the mean pixel value of each patch from it.
3. Remove the dictionary rows that are corresponding to the missing values by $D'_i = U_i D$.
4. Compute the sparse representations of the patches ($\{z'_{i,1}, z'_{i,2}, \dots, z'_{i,K}\}$) independently by equation (2).
5. Compute the probabilities by equation (7).
6. Compute the non-local weighted sparse coefficient by equation (5).
7. Restore the weighted means by equation (10).

Output: \hat{x}_i , the restored patch.

Fig. 3 The procedure of restoring a patch via NWSR.

4.2 Quantitative Metrics

To assess objective performance of the methods, we use quantitative metrics that are commonly used in medical image reconstruction. We adopt the peak signal-to-noise-ratio (PSNR), mean-to-standard-deviation ratio (MSR),⁴⁶ contrast-to-noise-ratio (CNR),⁴⁷ and equivalent number of looks (ENL).⁴⁸

The PSNR is a widely used metric that globally measures the intensity difference between the processed and the reference image. The other mentioned metrics do not require the reference image, and their computation requires selecting regions of interest from the images. We compute the MSR by averaging mean to standard deviation (SD) ratio on foreground regions of an image (e.g., red box #2-#6 in Fig. 4). The CNR measures the contrast between foreground regions and background noise by taking into account not only foreground regions but also background region (e.g., red box #1 in Fig. 4). The ENL involves computing mean and SD of background region. Therefore, it evaluates smoothness in background regions. Large ENL indicates a stronger noise smoothing in the corresponding region.⁴⁸

4.3 Compared Methods

We compare the proposed NWSR method quantitatively and qualitatively with other competing methods from the literature. The comparison methods for the OCT image denoising include: K-SVD denoising algorithm,⁴¹ BM3D,³⁴ patch group prior-based denoising (PGPD),⁴⁹ two-dimensional sparsity-based simultaneous denoising and interpolation (2-D-SBSDI),⁵ and BM4D.⁵⁰

For the OCT image interpolation problem, the comparison methods include bicubic, BM3D³⁴ + bicubic, single image scale-up using sparse representation by Zeyde et al.⁵¹ (Zeyde), 2-D-SBSDI,⁵ and BM4D⁵⁰ + bicubic. The BM3D/BM4D + bicubic method is a combination of the BM3D/BM4D denoising approach and the bicubic interpolation approach.

4.4 Algorithm Parameters

We set most of the parameters of the proposed method NWSR parameters based on our experiments. The selected parameters kept unchanged for all images in both interpolation and denoising experiments. We extract patches of size $p_1 \times p_2 = 8 \times 8$ pixels for DL and sparse representation. The smaller patch sizes have negative effects on the performance of the

proposed method and the bigger ones increase the computational cost. We will evaluate the effects of the patch size on the performance of the proposed method in Sec. 4.7. For each patch, we extract $L = 30$ patches in the same row and retrieve $K = 6$ most similar ones (including the patch itself). Although a larger search window might further enhance the reconstruction performance, it has more computational cost. In Sec. 4.8, we will discuss the effect of the number of retrieved similar patches (K) on the performance of the proposed NWSR method. The constant h in the exponential weight function [Eq. (6)] is set to 80. The sparsity level in the image reconstruction algorithm [Eq. (2)] is set to 2. To set the representation error for DL, we estimate the noise level from the training HH images by employing the algorithm published by Ref. 52. The obtained value is $\sigma = 4.6$. The noise gain g is set to 1.65. The procedure terminates in 20 iterations. Adding more iterations hardly results in any improvement. The number of dictionary atoms is set to 128. Further increasing the number of atoms will slightly improve the performance of reconstruction with the expense of higher computation time. We initialize the DL algorithm by randomly selecting patches from the training set. Better initialization methods can be used to improve the algorithm's performance,¹⁸ but we still use random initialization due to its simplicity. The patch size for NLM filtering is set to 6×6 pixels. Increasing the NLM filtering patch size results in over smoothing of textured regions. Instead of removing all of the noise using an NLM filtering with bigger patch size, we reduce the amount of noise using this off-the-shelf algorithm. Then, the remaining noise will be removed by sparse representation. All parameters involved in the compared methods were optimally assigned or chosen as described in the reference papers.

4.5 Results for OCT Image Denoising

Figure 4 demonstrates a visual comparison between the proposed NWSR method and compared methods for denoising of a real retinal OCT test image. As can be seen, the K-SVD, BM3D, and PGPD significantly suppress the noise while introducing visual artifacts. The SBSDI further reduces the artifacts, but it results in a slightly noisy reconstruction. BM4D can greatly reduce the noise, and it can better preserve layer structures due to using multiple images for denoising, but the result exhibits visible artifacts. The proposed NWSR method can effectively reduce noise while preserving many structures, compared to the HH image. Average quantitative results (over

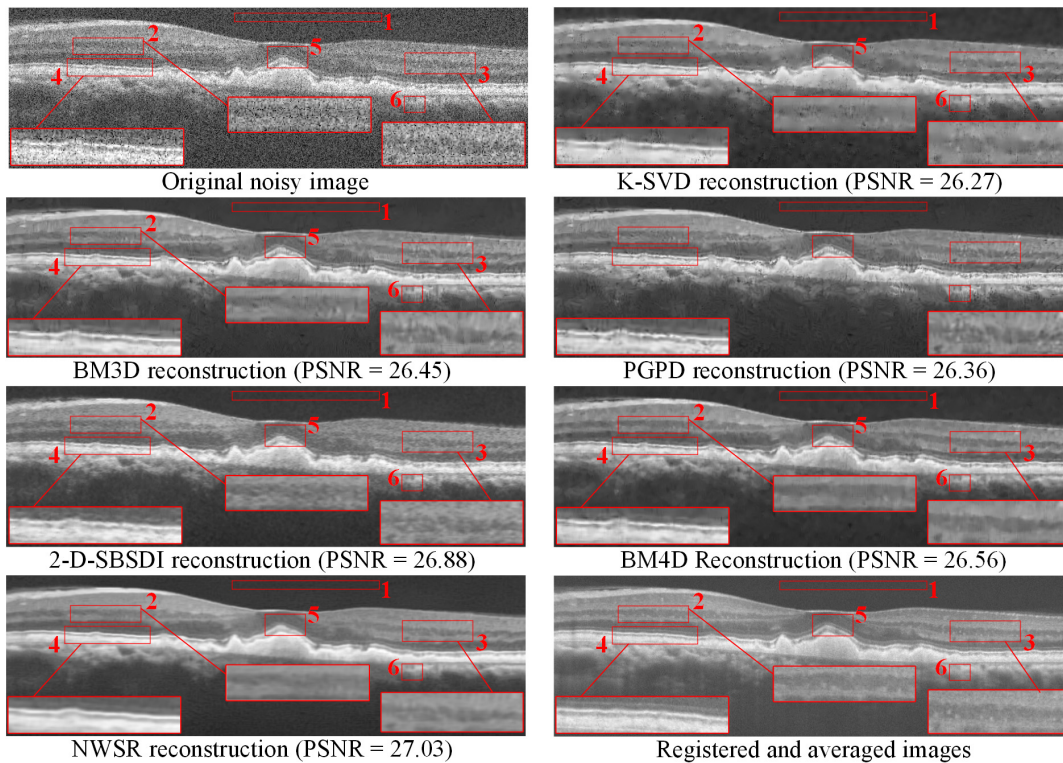


Fig. 4 Visual comparison of the denoised images by KSVD,⁴¹ BM3D,³⁴ PGPD,⁴⁹ 2-D-SBSDI,⁵ BM4D,⁵⁰ and the proposed NWSR method.

18 foveal images) of all the test methods are reported in Table 1. According to this table, the quantitative results of the proposed NWSR method are superior to those of the compared methods in terms of the four quantitative metrics (i.e., MSR, CNR, ENL, and PSNR). The experiments were conducted on a laptop

with an Intel® Core i7-4702MQ processor and 16 GB of RAM. On average, denoising an OCT test image of size 450 × 900 pixels using the proposed NWSR method takes about 108 s, implemented in MATLAB®. We did not optimize the code for speed, thus there is potential to reduce the average running time.

Table 1 Mean and SD of the MSR, CNR, ENL, and PSNR results for reconstructing 18 foveal images by the test methods.

Metric\method		K-SVD ⁴¹	BM3D ³⁴	PGPD ⁴⁹	2-D-SBSDI ⁵	BM4D ⁵⁰	NWSR
MSR	Mean	7.53	6.86	7.17	8.2	7.07	8.67
	SD	1.26	0.96	1.2	1.09	0.81	1.15
	<i>p</i> value	1.91E-06*	9.33E-12*	2.63E-09*	1.28E-06*	2.26E-10*	
CNR	Mean	3.39	3.21	3.22	3.43	3.31	3.45
	SD	0.52	0.45	0.48	0.45	0.44	0.47
	<i>p</i> value	1.74E-01	6.04E-07*	1.70E-05*	3.51E-02*	1.25E-06*	
ENL	Mean	776.24	1228.17	995.42	1539.14	1038.3	1552.68
	SD	150.69	511.56	314.31	305.97	262.91	264.07
	<i>p</i> value	6.53E-12*	2.85E-03*	8.85E-09*	7.31E-01	1.84E-07*	
PSNR	Mean	27.00	27.03	26.97	27.53	27.57	27.79
	SD	2.41	2.42	2.44	2.28	2.49	2.41
	<i>p</i> value	1.26E-04*	1.04E-05*	6.73E-05*	4.28E-06*	2.95E-03*	

Note: Best results in the mean values are shown in bold.
 **p* < 0.05, the metrics for each test method are considered statistically significant.

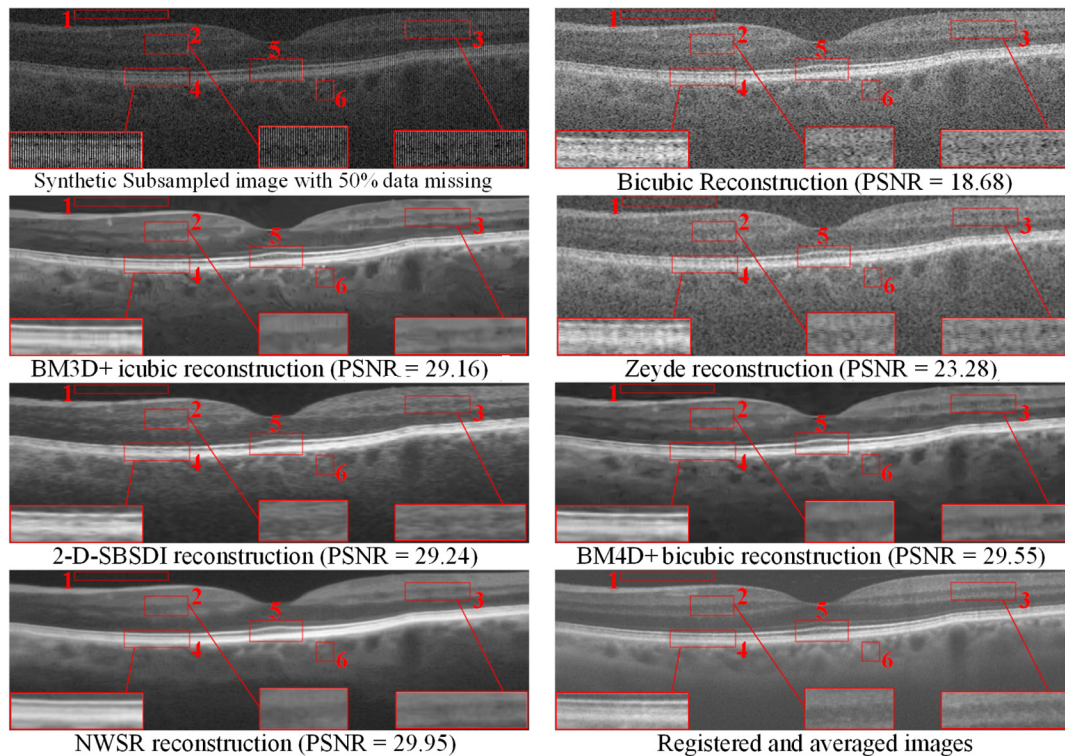


Fig. 5 Visual comparison of the reconstruction results of a synthetic subsampled retinal OCT test image (with 50% data missing) using bicubic, BM3D³⁴ + bicubic, Zeyde,⁵¹ 2-D-SBSDI,⁵ BM4D⁵⁰ + bicubic, and the proposed NWSR method. The figure is better seen by zooming on a computer screen.

4.6 Results for OCT Image Interpolation

Figures 5 and 6 show a visual comparison between the proposed NWSR method and compared methods for reconstruction of two real retinal OCT test images from two different datasets. Figure 5 shows a synthetic subsampled image (with 50% data missing), its visually reconstructed results, and its corresponding HH image. Figure 6 shows a real subsampled image (with 50% data missing) and its visually reconstructed results obtained from all the test methods. As expected, the results of bicubic and Zeyde are noisy, as we can see in Figs. 5 and 6. The BM3D + bicubic and BM4D + bicubic can significantly reduce the noise, but their results are suffered from visible artifacts. The BM4D + bicubic results in less artifacts than BM3D + bicubic and better preserves structures. Similar to denoising experiment, the SBSDI method reduces the artifacts, but it results in a slightly noisy reconstruction. The visual evaluation of the proposed NWSR method indicates good performance of the method in preserving textures and suppressing noise. This can be validated by the average quantitative results that are reported in Tables 2 and 3. The average quantitative results over 18 synthetic subsampled images from the first dataset are reported in Table 2. The quantitative results over 39 real subsampled images are reported in Table 3. As can be observed, the proposed NWSR method provides the best performance, except that for the mean of the ENL for reconstructing the images from the second dataset (Table 3). For this dataset, the best mean of the ENL is achieved by the 2-D-SBSDI method due to the fact that it results in a more aggressive background smoothing. This is also reflected in lower values of other metrics (i.e., MSR and CNR). On average, interpolating an OCT test

image with 50% data missing using the proposed NWSR method takes about 85 s. The implementation was the same as mentioned in Sec. 4.5.

4.7 Effects of Patch Size

To evaluate the effect of patch size on the performance, we vary this parameter in a certain range and set the other parameters to the fixed values described in Sec. 4.4. We present the experiment of reconstructing 18 synthetic subsampled images (with 50% data missing) described in Sec. 4.6. Figure 7 shows the behavior of PSNR versus different patch sizes. It can be seen that smaller patch sizes negatively affect the performance due to the limited spatial information. As the patch size increases, the method significantly suppresses noise but may result in oversmoothing and the loss of image detail. Consequently, we use the patch size of $p_1 \times p_2 = 8 \times 8$ pixels for all the experiments in our paper.

4.8 Effects of Number of Similar Patches to Compute Nonlocal Weighted Sparse Coefficients

Because our proposed NWSR method relies on sparse representations of the K most similar patches, we study here the impact of different values of K on the performance. We reconstruct 18 synthetic subsampled images (with 50% data missing) described in Sec. 4.6 with different values for K while keeping all other parameters unchanged. Figure 8 plots the average PSNR over the test images as a function of K . It can be seen that merging only a few representations yields benefits. The performance gets better as K increases. However, it might result in oversmoothing

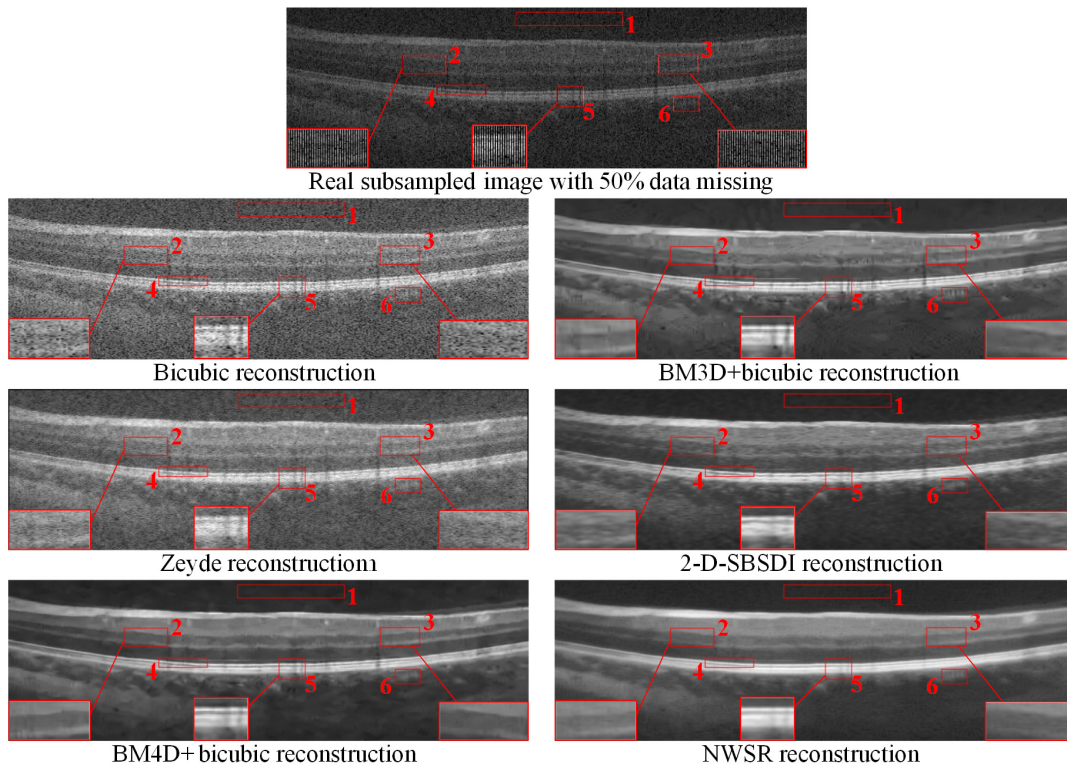


Fig. 6 Visual comparison of the reconstruction results of a real subsampled retinal OCT test image (with 50% data missing) using bicubic, BM3D³⁴ + bicubic, Zeyde,⁵¹ 2-D-SBSDI,⁵ BM4D⁵⁰ + bicubic, and the proposed NWSR method. The figure is better seen by zooming on a computer screen.

Table 2 Mean and SD of the MSR, CNR, ENL, and PSNR results for reconstructing 18 foveal images (with 50% data missing) by the test methods.

Metric\method		Bicubic	BM3D ³⁴ + bicubic	Zeyde ⁵¹	2-D-SBSDI ⁵	BM4D ⁵⁰ + bicubic	NWSR
MSR	Mean	3.44	7.49	5.08	8.75	7.6	10.01
	SD	0.29	1.14	0.43	1.13	0.95	1.39
	<i>p</i> value	1.88E-13*	6.31E-11*	5.08E-12*	1.56E-08*	2.18E-10*	
CNR	Mean	1.75	3.36	2.55	3.57	3.45	3.84
	SD	0.29	0.5	0.4	0.5	0.49	0.56
	<i>p</i> value	1.43E-14*	1.40E-10*	4.21E-13*	3.85E-11*	1.64E-10*	
ENL	Mean	8.62	1345.01	28.69	1885.33	1096.31	1664.19
	SD	1.12	543.74	3.79	545.09	342.14	305.71
	<i>p</i> value	4.72E-14*	6.43E-04*	5.22E-14*	7.58E-03*	6.21E-07*	
PSNR	Mean	18.54	27.26	22.72	27.64	27.76	27.96
	SD	0.48	2.44	0.91	2.27	2.52	2.29
	<i>p</i> value	1.31E-13*	8.76E-06*	2.00E-11*	3.90E-09*	3.14E-02*	

Note: Best results in the mean values are shown in bold.
 **p* < 0.05, the metrics for each test method are considered statistically significant.

Table 3 Mean and SD of the MSR, CNR, and ENL results for reconstructing 39 foveal images (with 50% data missing) by the test methods.

Method	MSR			CNR			ENL		
	Mean	SD	<i>p</i> value	Mean	SD	<i>p</i> value	Mean	SD	<i>p</i> value
Bicubic	3.34	0.17	1.28E-23*	1.67	0.25	3.54E-33*	8.81	0.44	1.49E-33*
BM3D ³⁴ + bicubic	6.69	0.91	3.04E-19*	3.08	0.38	2.84E-25*	1857.20	901.89	1.56E-01
Zeyde ⁵¹	4.86	0.26	7.71E-20*	2.42	0.33	7.58E-29*	29.19	1.40	2.19E-33*
2-D-SBSDI ⁵	7.74	1.34	4.66E-14*	3.25	0.39	7.80E-26*	1914.60	442.58	5.21E-06*
BM4D ⁵⁰ + bicubic	6.86	0.98	3.59E-17*	3.16	0.38	6.33E-22*	1304.11	471.65	1.08E-05*
NWSR	8.62	1.47		3.50	0.43		1675.00	243.26	

Note: Best results in the mean values are shown in bold.
 **p* < 0.05, the metrics for each test method are considered statistically significant.

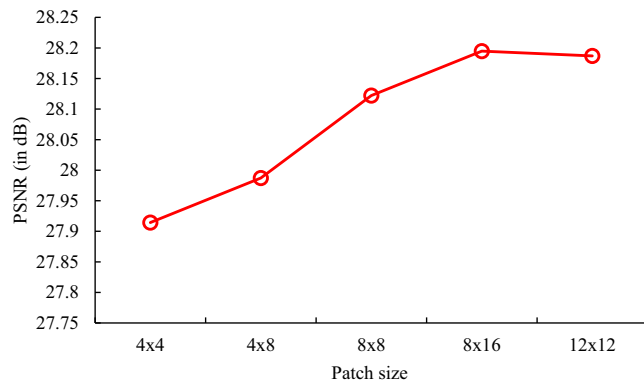


Fig. 7 Effects of the patch size on the performance of the proposed NWSR method. For each patch size, mean of the PSNR results for reconstructing 18 foveal images (with 50% data missing) is reported.

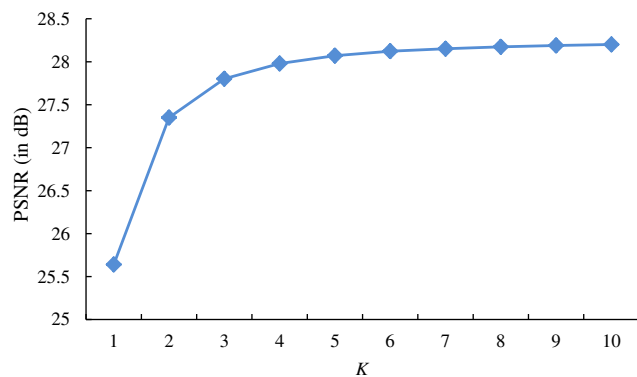


Fig. 8 Effects of number of similar patches (*K*) on the performance of the proposed NWSR method. For each value of *K*, mean of the PSNR results for reconstructing 18 foveal images (with 50% data missing) is reported.

and the loss of image detail. Therefore, we use *K* = 6 for all the experiments in our paper.

5 Conclusion

In this paper, we have presented a competitive method (named NWSR) capable of denoising and interpolating a cross-sectional

SDOCT image. The proposed NWSR method relies on sparse representations of multiple noisy and denoised LR patches to compute a more robust estimate of the HR patch representation. The experiments showed that merging only a few representations yields better image reconstruction results. The experimental results on two datasets of real retinal SDOCT images show the effectiveness of the proposed NWSR method over several leading state-of-the-art image reconstruction methods. The proposed NWSR method is useful for OCT image quality improvement, and it might be able to improve the performance of segmentation algorithms.^{18,53} However, there are limitations to this work that need to be addressed in further research efforts. First, although DL from HH images could provide a way to estimate the subspace of clean images,^{5,16} the HH images are currently not available for other applications of OCT images (e.g., intravascular OCT images of carotid arteries¹¹). Second, the proposed NWSR method is a single-image reconstruction method. Despite its benefits, there is potential to further improve the reconstruction quality by extending the proposed NWSR method to multiframe reconstruction. This is because there are high degrees of spatial correlations between nearby OCT images. Third, the proposed NWSR method uses fixed patch size. In order to efficiently capture OCT image structure, a more effective method might be offered based on adaptive patch size or shape adaptive patches.⁵⁴⁻⁵⁶ Forth, in future study, we would like to investigate the incorporation of high-frequency information into the reconstruction.²⁴ Fifth, although here we only considered the task of retinal OCT image reconstruction, the proposed method can also be applied to reconstruction of other noise-corrupted medical images.^{2,20}

Disclosures

The authors have no relevant financial interests in this article and no potential conflicts of interest to disclose.

Acknowledgments

This work was supported in part by the National Natural Science Fund of China for Distinguished Young Scholars under Grant 61325007 and by the National Natural Science Foundation of China under Grants 61501180 and 61771192.

References

1. M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Rev. Biomed. Eng.* **3**, 169–208 (2010).
2. M. D. Robinson et al., "Novel applications of super-resolution in medical imaging," in *Super-Resolution Imaging*, P. Milanfar, Ed., pp. 384–412, CRC Press, Boca Raton, Florida (2010).
3. M. Young et al., "Real-time high-speed volumetric imaging using compressive sampling optical coherence tomography," *Biomed. Opt. Express* **2**(9), 2690–2697 (2011).
4. A. Boroomand et al., "Multi-penalty conditional random field approach to super-resolved reconstruction of optical coherence tomography images," *Biomed. Opt. Express* **4**(10), 2032–2050 (2013).
5. L. Fang et al., "Fast acquisition and reconstruction of optical coherence tomography images via sparse representation," *IEEE Trans. Med. Imaging* **32**(11), 2034–2049 (2013).
6. L. Fang et al., "Segmentation based sparse reconstruction of optical coherence tomography images," *IEEE Trans. Med. Imaging* **36**(2), 407–421 (2017).
7. P. Milanfar, "A tour of modern image filtering: new insights and methods, both practical and theoretical," *IEEE Signal Process. Mag.* **30**(1), 106–128 (2013).
8. S. Chitchian et al., "Retinal optical coherence tomography image enhancement via shrinkage denoising using double-density dual-tree complex wavelet transform," *J. Biomed. Opt.* **17**, 116004 (2012).
9. Y. Du et al., "Speckle reduction in optical coherence tomography images based on wave atoms," *J. Biomed. Opt.* **19**, 056007 (2014).
10. L. Bian et al., "Multiframe denoising of high-speed optical coherence tomography data using interframe and intraframe priors," *J. Biomed. Opt.* **20**, 036006 (2015).
11. J. Xu et al., "Wavelet domain compounding for speckle reduction in optical coherence tomography," *J. Biomed. Opt.* **18**, 096002 (2013).
12. Z. Amini and H. Rabbani, "Optical coherence tomography image denoising using Gaussianization transform," *J. Biomed. Opt.* **22**(8), 086011 (2017).
13. H. Zhang et al., "Speckle reduction in optical coherence tomography by two-step image registration," *J. Biomed. Opt.* **20**, 036013 (2015).
14. D. Alonso-Caneiro, S. A. Read, and M. J. Collins, "Speckle reduction in optical coherence tomography imaging by affine-motion image registration," *J. Biomed. Opt.* **16**, 116026 (2011).
15. H. M. Salinas and D. C. Fernandez, "Comparison of PDE-Based non-linear diffusion approaches for image enhancement and denoising in optical coherence tomography," *IEEE Trans. Med. Imaging* **26**(6), 761–771 (2007).
16. L. Fang et al., "Sparsity based denoising of spectral domain optical coherence tomography images," *Biomed. Opt. Express* **3**(5), 927–942 (2012).
17. H. Rabbani, M. Sonka, and M. D. Abramoff, "Optical coherence tomography noise reduction using anisotropic local bivariate Gaussian mixture prior in 3D complex wavelet domain," *Int. J. Biomed. Imaging* **2013**, 1–23 (2013).
18. R. Kafieh, H. Rabbani, and I. Selesnick, "Three dimensional data-driven multi scale atomic representation of optical coherence tomography," *IEEE Trans. Med. Imaging* **34**(5), 1042–1062 (2015).
19. Z. Amini and H. Rabbani, "Statistical modeling of retinal optical coherence tomography," *IEEE Trans. Med. Imaging* **35**(6), 1544–1554 (2016).
20. D.-H. Trinh et al., "Novel example-based method for super-resolution and denoising of medical images," *IEEE Trans. Image Process.* **23**(4), 1882–1895 (2014).
21. W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *Int. J. Comput. Vision* **40**(1), 25–47 (2000).
22. M. A. Mayer et al., "Wavelet denoising of multiframe optical coherence tomography data," *Biomed. Opt. Express* **3**(3), 572–589 (2012).
23. A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.* **51**(1), 34–81 (2009).
24. J. Yang et al., "Image super-resolution via sparse representation," *IEEE Trans. Image Process.* **19**(11), 2861–2873 (2010).
25. W. Dong et al., "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.* **20**(7), 1838–1857 (2011).
26. J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.* **17**(1), 53–69 (2008).
27. J. Mairal et al., "Discriminative learned dictionaries for local image analysis," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8 (2008).
28. X. Liu and J. U. Kang, "Compressive SD-OCT: the application of compressed sensing in spectral domain optical coherence tomography," *Opt. Express* **18**(21), 22010–22019 (2010).
29. D. Xu et al., "Modified compressive sensing optical coherence tomography with noise reduction," *Opt. Lett.* **37**(20), 4209–4211 (2012).
30. E. Lebed et al., "Rapid volumetric OCT image acquisition using compressive sampling," *Opt. Express* **18**(20), 21003–21012 (2010).
31. J. Yang et al., "Coupled dictionary training for image super-resolution," *IEEE Trans. Image Process.* **21**(8), 3467–3478 (2012).
32. S. Wang et al., "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 2216–2223 (2012).
33. W. Dong et al., "Nonlocally centralized sparse representation for image restoration," *IEEE Trans. Image Process.* **22**(4), 1620–1630 (2013).
34. K. Dabov et al., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
35. J. Mairal et al., "Non-local sparse models for image restoration," in *IEEE Int. Conf. on Computer Vision*, pp. 2272–2279 (2009).
36. S. Yang et al., "Multitask dictionary learning and sparse representation based single-image super-resolution reconstruction," *Neurocomputing* **74**(17), 3193–3203 (2011).
37. K. Egiazarian and V. Katkovnik, "Single image super-resolution via BM3D sparse coding," in *European Signal Processing Conf.*, pp. 2849–2853 (2015).
38. C. Cruz et al., "Single image super-resolution based on Wiener filter in similarity domain," *IEEE Trans. Image Process.* **27**(3), 1376–1389 (2018).
39. A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 60–65 (2005).
40. A. Singh, F. Porikli, and N. Ahuja, "Super-resolving noisy images," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 2846–2853 (2014).
41. M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.* **15**(12), 3736–3745 (2006).
42. M. Aharon, M. Aharon, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.* **54**(11), 4311–4322 (2006).
43. R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," Technical Report, Technion, Department of Computer Science, Haifa, Israel (2008).
44. I. Mosseri, M. Zontak, and M. Irani, "Combining the power of internal and external denoising," in *IEEE Int. Conf. on Computational Photography*, pp. 1–9 (2013).
45. Fang et al., "Software and datasets used for fast acquisition and reconstruction of optical coherence tomography images via sparse representation," 2013, http://people.duke.edu/~sf59/Fang_TMI_2013.htm (01 January 2018).
46. G. Cincotti, G. Loi, and M. Pappalardo, "Ultrasound medical images with wavelet packets," *IEEE Trans. Med. Imaging* **20**(8), 764–771 (2001).
47. P. Bao and L. Zhang, "Noise reduction for magnetic resonance images via adaptive multiscale products thresholding," *IEEE Trans. Med. Imaging* **22**(9), 1089–1099 (2003).
48. A. Pizurica et al., "Multiresolution denoising for optical coherence tomography: a review and evaluation," *Curr. Med. Imaging Rev.* **4**(4), 270–284 (2008).
49. J. Xu et al., "Patch group based nonlocal self-similarity prior learning for image denoising," in *IEEE Int. Conf. on Computer Vision*, pp. 244–252 (2015).
50. M. Maggioni et al., "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.* **22**(1), 119–133 (2013).

51. R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," *Lect. Notes Comput. Sci.* **6920**, 711–730 (2012).
52. A. Foi, "Noise estimation and removal in MR imaging: the variance-stabilization approach," in *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, pp. 1809–1814 (2011).
53. L. Fang et al., "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomed. Opt. Express* **8**(5), 2732–2744 (2017).
54. H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.* **16**(2), 349–366 (2007).
55. L. Fang et al., "Hyperspectral image classification via multiple-feature-based adaptive sparse representation," *IEEE Trans. Instrum. Meas.* **66**(7), 1646–1657 (2017).
56. L. Fang, H. Zhuo, and S. Li, "Super-resolution of hyperspectral image via superpixel-based sparse representation," *Neurocomputing* **273**, 171–177 (2018).

Ashkan Abbasi received his BS degree in software engineering in 2011 from Hamedan Branch, Islamic Azad University, Hamedan, Iran. He received his MS degree in artificial intelligence from Isfahan University, Isfahan, Iran, in 2013. He is currently working toward his PhD in Isfahan University. From September 2017 to February 2018, he is a visiting student at the Vision and Image Processing Laboratory of Hunan University in Changsha, China. His research interests are medical image analysis and machine learning.

Amirhassan Monadjemi received his BS and MSc degrees in computer engineering from Isfahan University of Technology in 1990 and Shiraz University in 1994, respectively. He received his PhD in

computer engineering, pattern recognition, and image processing, from the University of Bristol, Bristol, England, in 2004. He is now working as an associate professor at the Department of Artificial Intelligence, Faculty of Computer Engineering, Isfahan University, Isfahan, Iran. His research interests include artificial intelligence, image processing, and neural networks.

Leyuan Fang received his PhD from the College of Electrical and Information Engineering, Hunan University, Changsha, China, in 2015. From September 2011 to September 2012, he was a visiting PhD student with the Department of Ophthalmology, Duke University, Durham, North Carolina, USA, supported by the China Scholarship Council. From August 2016 to 2017, he was a postdoc researcher at the Department of Biomedical Engineering, Duke University, Durham, USA. Since January 2017, he has been an associate professor at the College of Electrical and Information Engineering, Hunan University. His research interests include sparse representation and deep learning in medical image processing.

Hossein Rabbani received his MS and PhD degrees in bioelectrical engineering in 2002 and 2008, respectively, from Amirkabir University of Technology (Tehran Polytechnic). In 2007, he was at Queen's University, as a visiting researcher; in 2011 at the University of Iowa as a postdoctoral research scholar; and from 2013 to 2014 at Duke University Eye Center as a postdoctoral fellow. He is currently an associate professor in the Biomedical Engineering Department and Medical Image and Signal Processing Research Center of Isfahan University of Medical Sciences. His current research interests include medical image analysis and modeling, statistical (multidimensional) signal processing, sparse transforms, and image/video restoration.