arXiv:1312.1931v1 [cs.CV] 6 Dec 2013

Denoising High Speed Optical Coherence Tomography Data Using Inter-frame and Intra-frame Priors

Liheng Bian, Jinli Suo, Feng Chen and Qionghai Dai

Department of Automation, Tsinghua University, Beijing, 100084, China

lihengbian@gmail.com

Abstract: Optical coherence tomography (OCT) is an important interferometric diagonal technique which provides cross-sectional views of the subsurface microstructure of biological tissue. However, the imaging quality of high-speed OCT is limited due to the large speckle noise. To address this problem, this paper proposes a multi-frame algorithmic method to denoise OCT volume images. Mathematically, both inter-frame and intra frame priors are unified into an optimization model: low rank constraint and sparsity are imposed on registered frames and sparsity of image gradient. We use a convex algorithm based on Augmented Lagrangian Method (ALM) method to solve the model. The results reveal that our approach outperforms state-of-the-arts methods in terms of both speckle noise suppression and crucial detail reservation.

© 2019 Optical Society of America

OCIS codes: (000.0000) General.

References and links

1. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito *et al.*, "Optical coherence tomography," Science **254**, 1178–1181 (1991).
2. J. M. Schmitt, "Optical coherence tomography (oct): a review," Selected Topics in Quantum Electronics, IEEE Journal of **5**, 1205–1215 (1999).
3. J. Welzel, "Optical coherence tomography in dermatology: a review," Skin Research and Technology **7**, 1–9 (2001).
4. J. M. Schmitt, S. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," Journal of Biomedical Optics **4**, 95–105 (1999).
5. A. Cameron, D. Lui, A. Boroomand, J. Glaister, A. Wong, and K. Bizheva, "Stochastic speckle noise compensation in optical coherence tomography using non-stationary spline-based speckle noise modelling," Biomedical optics express **4**, 1769 (2013).
6. M. Wojtkowski, V. Srinivasan, T. Ko, J. Fujimoto, A. Kowalczyk, and J. Duker, "Ultrahigh-resolution, high-speed, fourier domain optical coherence tomography and methods for dispersion compensation," Optics Express **12**, 2404–2422 (2004).
7. M. A. Choma, M. V. Sarunic, C. Yang, and J. A. Izatt, "Sensitivity advantage of swept source and fourier domain optical coherence tomography," Optics Express **11**, 2183–2189 (2003).
8. M. A. Mayer, A. Borsdorf, M. Wagner, J. Hornegger, C. Y. Mardin, R. P. Tornow *et al.*, "Wavelet denoising of multiframe optical coherence tomography data," Biomedical optics express **3**, 572 (2012).
9. J. Schmitt and A. Knüttel, "Model of optical coherence tomography of heterogeneous tissue," JOSA A **14**, 1231–1242 (1997).
10. M. Bashkansky and J. Reintjes, "Statistics and reduction of speckle in optical coherence tomography," Optics Letters **25**, 545–547 (2000).
11. B. Karamata, K. Hassler, M. Laubscher, and T. Lasser, "Speckle statistics in optical coherence tomography," JOSA A **22**, 593–596 (2005).
12. A. Ozcan, A. Bilenca, A. E. Desjardins, B. E. Bouma, and G. J. Tearney, "Speckle reduction in optical coherence tomography images using digital filtering," JOSA A **24**, 1901–1910 (2007).

13. Y. Yue, M. M. Croitoru, A. Bidani, J. B. Zwischenberger, and J. W. Clark Jr, "Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images," Medical Imaging, IEEE Transactions on **25**, 297–311 (2006).
14. F. Zhang, Y. M. Yoo, L. M. Koh, and Y. Kim, "Nonlinear diffusion in laplacian pyramid domain for ultrasonic speckle reduction," Medical Imaging, IEEE Transactions on **26**, 200–211 (2007).
15. M. Gargesha, M. W. Jenkins, A. M. Rollins, and D. L. Wilson, "Denoising and 4d visualization of oct images," Optics express **16**, 12313–12333 (2008).
16. J. Xie, Y. Jiang, H.-T. Tsui, and P.-A. Heng, "Boundary enhancement and speckle reduction for ultrasound images via salient structure extraction," Biomedical Engineering, IEEE Transactions on **53**, 2300–2309 (2006).
17. D. L. Marks, T. S. Ralston, and S. A. Boppart, "Speckle reduction by i-divergence regularization in optical coherence tomography," JOSA A **22**, 2366–2371 (2005).
18. A. Wong, A. Mishra, K. Bizheva, and D. A. Clausi, "General bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery," Opt. Express **18**, 8338–8352 (2010).
19. L. Fang, S. Li, Q. Nie, J. A. Izatt, C. A. Toth, and S. Farsiu, "Sparsity based denoising of spectral domain optical coherence tomography images," Biomedical optics express **3**, 927–942 (2012).
20. N. Iftimia, B. E. Bouma, and G. J. Tearney, "Speckle reduction in optical coherence tomography by path length encoded angular compounding," Journal of biomedical optics **8**, 260–263 (2003).
21. L. Ramrath, G. Moreno, H. Mueller, T. Bonin, G. Huettmann, and A. Schweikard, "Towards multi-directional oct for speckle noise reduction," in "Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008," (Springer, 2008), pp. 815–823.
22. M. Hughes, M. Spring, and A. Podoleanu, "Speckle noise reduction in optical coherence tomography of paint layers," Applied optics **49**, 99–107 (2010).
23. A. Desjardins, B. Vakoc, G. Tearney, and B. Bouma, "Speckle reduction in oct using massively-parallel detection and frequency-domain ranging," Optics express **14**, 4736 (2006).
24. A. Desjardins, B. Vakoc, W. Oh, S. Motaghiannezam, G. Tearney, and B. Bouma, "Angle-resolved optical coherence tomography with sequential angular selectivity for speckle reduction," Optics express **15**, 6200 (2007).
25. M. Pircher, E. Go, R. Leitgeb, A. F. Fercher, C. K. Hitzenberger *et al.*, "Speckle reduction in optical coherence tomography by frequency compounding," Journal of Biomedical Optics **8**, 565–569 (2003).
26. Z. Jian, L. Yu, B. Rao, B. J. Tromberg, and Z. Chen, "Three-dimensional speckle suppression in optical coherence tomography based on the curvelet transform," Optics express **18**, 1024 (2010).
27. Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," arXiv preprint arXiv:1009.5055 (2010).
28. A. Leigh, A. Wong, D. A. Clausi, and P. Fieguth, "Comprehensive analysis on the effects of noise estimation strategies on image noise artifact suppression performance," in "Multimedia (ISM), 2011 IEEE International Symposium on," (IEEE, 2011), pp. 97–104.
29. T. S. Cho, N. Joshi, C. L. Zitnick, S. B. Kang, R. Szeliski, and W. T. Freeman, "A content-aware image prior," in "Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on," (IEEE, 2010), pp. 169–176.
30. T. S. Cho, C. L. Zitnick, N. Joshi, S. B. Kang, R. Szeliski, and W. T. Freeman, "Image restoration by matching gradient distributions," Pattern Analysis and Machine Intelligence, IEEE Transactions on **34**, 683–694 (2012).
31. F. Heide, M. Rouf, M. B. Hullin, B. Labitzke, W. Heidrich, and A. Kolb, "High-quality computational imaging through simple lenses," ACM Trans. Graph.(Proc. SIGGRAPH 2013) (2013).
32. Y. Deng, Q. Dai, and Z. Zhang, "An overview of computational sparse models and their applications in artificial intelligence," in "Artificial Intelligence, Evolutionary Computing and Metaheuristics," (Springer, 2013), pp. 345–369.
33. H. Lee, A. Battle, R. Raina, and A. Ng, "Efficient sparse coding algorithms," in "Advances in neural information processing systems," (2006), pp. 801–808.
34. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," Image Processing, IEEE Transactions on **13**, 600–612 (2004).
35. R. Bernardes, C. Maduro, P. Serranho, A. Araújo, S. Barbeiro, and J. Cunha-Vaz, "Improved adaptive complex diffusion despeckling filter," (2010).
36. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," Image Processing, IEEE Transactions on **16**, 2080–2095 (2007).
37. M. A. Mayer, A. Borsdorf, M. Wagner, J. Hornegger, C. Y. Mardin, R. P. Tornow *et al.*, "Image denoising algorithms archive," `http://www5.cs.fau.de/en/research/software/idaa/` (2012).
38. M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "A nonlocal transform-domain filter for volumetric data denoising and reconstruction," (2007).

1. Introduction

Optical coherence tomography (OCT), which dates back to 1999[1], provides cross-sectional views of the subsurface microstructure of biological tissues[2][3] and has developed into a

widely used diagonal technique in medical field due to its non-invasive diagonal manner. Different from common CCD imaging, OCT is an interferometric technique and typically employs near-infrared laser light to penetrate into the scattering medium before capturing the backscattered optical waves for the final imagery[4]. The signal acquisition speed of OCT systems has been greatly improved with the development of high speed sensors and tunable lasers at MHz scanning rates, which allows real-time imaging of vivo tissues [5]. Furthermore, with the development of ultrahigh resolution OCT (UHROCT) [6] and Fourier domain OCT (FDOCT) [7], it is feasible to visualize biological tissue at cellular level up to depths of nearly 1mm below the tissue with higher sensitivity and image quality. Thus, OCT technique can capture morphological details such as individual tissue layers, clusters of specialized cells, small blood vessels, and so on. Specifically, ophthalmology has particularly benefited from OCT with which it can image retina to aid the diagnosis of eye diseases [8].

One of the important challenges limiting high-speed OCT's development is its unsatisfying image quality caused by speckle noise. Due to the coherence of optical waves, speckle noise arises under limited spatial-frequency bandwidth of the interference signals [4]. The generation mechanism of OCT determines that the properties of speckle noise are related not only to laser source but also to tissue's structural properties, and thus results in non-stationary speckle noise over the entire image. There has been some work investigating the properties of speckle noise in OCT images [4][9][10][11], showing that speckle noise carries both image information of tissue structures and a single noise component. Due to the requirement of high precision in medical diagnosis, it is greatly vital to remove speckle noise from OCT images for enhancing image quality in order to recognize diseases correctly.

Large efforts has been invested in denoising OCT images and various approaches are reported. Here we classify these methods into two categories, namely *single-frame methods* and *multi-frame methods*. Single-frame methods often assume a prior parametric or non-parametric model for signal and noise, and then remove noise or use statistic methods to remove noise probabilistically from a single image. Aydogan et al.[12] apply various digital filters to denoise OCT images and the results indicate that nonorthogonal wavelet filters together with enhanced Lee and adaptive Wiener filters can significantly reduce speckle noise. Based on the wavelet filter, Yong et al.[13] utilize the iterative edge enhancement feature of nonlinear diffusion to improve denoising results. Similarly, [14] uses nonlinear diffusion in Laplacian pyramid domain to filter ultrasonic images. Kovesi Nw filtering technique and Laplacian Pyramid Nonlinear Diffusion (LPND) are unified together in [15] to remove both shot noise and speckle noise from OCT images. Concerning on image contents, Jun et al.[16] propose a salient structure extraction algorithm combining an adaptive speckle suspension term to enhance ultrasound images while restraining speckle noise. The work in [17] uses a regularization method to minimize Csiszar's I-divergence measurement, which would extrapolate additional details to improve visual effects. Alexander et al.[18] and Andrew et al.[5] use statistic Bayesian least square model to reduce OCT speckle noise in logarithmic space. Different from the above filtering or statistic methods, based on sparse coding, the work in [19] learns an over complete dictionary from high-SNR images, then utilizes this dictionary to reconstruct low-SNR OCT images and result in significant noise suppression. In all, making use of the intrinsic redundancy of a single-frame OCT image can help noise removal to a large extent.

Benefiting from the development of high speed OCT imaging systems, researchers gradually begin to investigate using multiple OCT frames to attenuate speckle noise, namely multi-frame methods. Technically these methods could be classified into two categories as well, namely hardware methods and algorithmic methods. The general idea of hardware methods is changing parameters of OCT imaging systems to decorrelate speckle noise in multiple frames from the ground truth, and then get a noise-free OCT image by directly averaging these frames after im-

age registration. For example, [10][20][21][22] alter the angle of incident light when shooting different frames, [23][24] change the detection angle of backreflected light, while [25] changes the frequency of the laser beam in the imaging progress. The main disadvantage of hardware methods is the complex data acquisition procedure, which would greatly increase the design complexity of OCT imaging systems[5]. There are several recent algorithmic multi-frame denoising methods for OCT images making use of the redundancy of latent sharp OCT image in frequency domain. The work in [26] performs 3D curvelet transform to the volume data, and determine a threshold for coefficient attenuation before inverse 3D curvelet transform. Similarly, Markus et al.[8] use the same idea but replacing the 3D curvelet transform with wavelet transform. In spite of the promising performance, these denoising algorithms run the risk of losing crucial details by truncating the wavelet coefficients.

In this study, we propose an algorithmic multi-frame optimization method making use of both inter-frame and intra-frame priors to denoise OCT volume images. Firstly, registering the images in the cubic OCT data results in the low-rank property of the data at the image count dimension. Secondly, by applying the sparse gradient prior of natural images to OCT data, we force the gradient of OCT images being still sparse. Base on the above two priors, combining the restraints of image formation and noise variance, we build a preliminary non-convex optimization model to minimize the rank of OCT volume data and the sparsity of its gradient. To get the closed-form global optimal solution, some mathematical transforms and approximations are made to make the model convex, and a numeric algorithm based on the augmented lagrangian multiplier method [27] is provided by this work. In summary, the contributions of this study are mainly in three aspects as follows:

- Introduce both inter-frame and intra-frame priors of OCT data, which effectively promise separation between signal and speckle noise.
- Use non-parametric and non-stationary noise restraint, which simplifies the optimization model and raises the precision.
- Derive a convex optimization algorithm to solve the mathematic model, which ensures that the problem converges to a global optimum.

The remainder of this paper is organized as follows: In Sec. 2, two preprocessing operations including frame registration and noise estimation are introduced firstly, followed by model construction, algorithm derivation and explanation of several evaluation methods we use to judge recovered quality. Then we apply our method to actual OCT images in Sec. 3, and compare our approach with several other classical methods in recovered image quality and algorithm runtime. Convergence analysis and model parameter analysis are conducted in this section as well. Finally, we conclude this paper with some conclusions and discussions in Sec. 4.

2. Method

2.1. Frame registration and noise estimation

There tend to be slight misalignment among OCT image sequences in vivo imaging due to the life dynamics or other system factors[8], though the optical coherence tomography facilities are usually of high shot speed. Therefore, we should firstly register the frames before further operations. Here we adopt the registration method in [8], where a powell optimizer is utilized for minimizing the sum of squared distances (SSD) among multiple registered images. Specifically, the approach applies translation and rotation movements to warp the pixels describing the same tissue point in different frames to the same image coordinate.

Another preprocessing operation is estimating standard deviation of the image noise. Considering the satisfying performance of median absolute deviation (MAD) method for noise

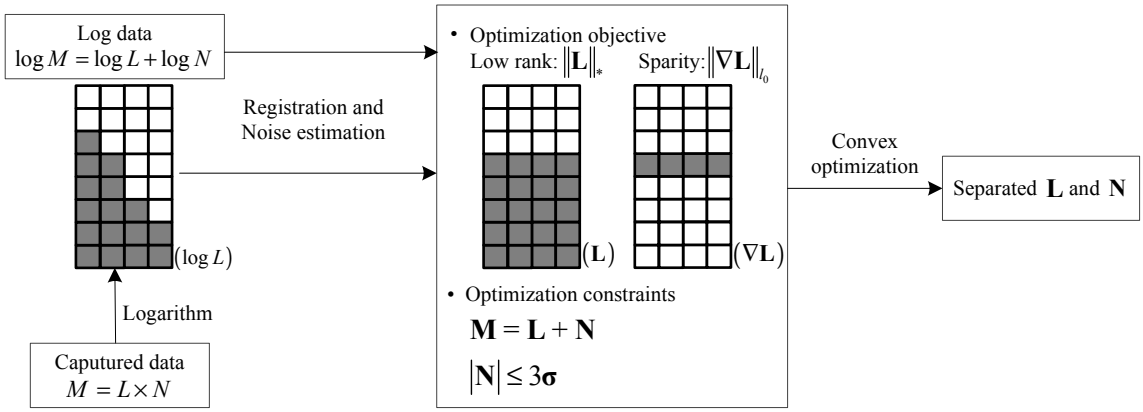


Fig. 1. **Framework of the proposed approach.** After logarithm, each frame is represented by a single column, and there is slight misalignment compared to other frames, as ($\log L$) shows. Then by noise estimation and frame registration, we get not only the optimization constraints but also the optimization objective, namely $\mathbf{L}$'s low rank and the sparsity of $\bigtriangledown\mathbf{L}$, as shown by ($\mathbf{L}$) and ($\bigtriangledown\mathbf{L}$). Finally, the model is iteratively solved by a convex optimization algorithm, thus $\mathbf{L}$ is separated from $\mathbf{N}$.

estimation[28], we utilize the MAD noise estimation method in [5]. Because of the scattering of laser source, the speckle noise in OCT images is multiplicative[9][10], and can be described as

$$M(s) = L(s) \times N(s), \tag{1}$$

where $M(s)$ represents the captured data at location s, and $L(s)$ and $N(s)$ denote the ground truth data and measurement noise at the same location, respectively. To convert the relationship between F and N from multiplication to addition, we take logarithm to both sides of Eq.(1) and get

$$\log M(s) = \log L(s) + \log N(s). \tag{2}$$

In the following sections, we assume that all the variables have been log transformed.

Because of similar tissue properties and light directions in the neighborhood, we assume uniform noise parameters within a small patch. Therefore, the MAD within a small neighborhood $\mathbb{N}_1$ of pixel s is firstly computed in the log space as

$$\hat{\sigma}(s) = 1.4826 \mathscr{M}_{\mathbb{N}_1}\left(|\log M(s_i) - \mathscr{M}_{\mathbb{N}_1}(\log M)|\right) \tag{3}$$

where $\mathscr{M}$ stands for median operation and s_i is the ith neighboring point of s. To make the estimated deviation more precise, we choose a larger region $\mathbb{N}_2$ around s, calculate $\hat{\sigma}$ of each small neighborhood within this region, and treat the mode of these $\hat{\sigma}$ as the preliminary noise deviation estimation of s as

$$\overline{\sigma}(s) = mode_{\mathbb{N}_2}(\hat{\sigma}). \tag{4}$$

Empirical studies in [5] tell that the noise estimation performs best when the neighborhood sizes of $\mathbb{N}_1$ and $\mathbb{N}_2$ are 9×9 and 15×15, respectively. Finally, to force smoothness of noise parameters among adjacent pixels, a cubic spline fitting progress is conducted to amend $\overline{\sigma}$ and get the final standard deviation estimation of OCT noise.

2.2. Modeling

Suppose that there are k frames of OCT data and the resolution of each frame is $m \times n$ pixels, we denote temporally registered noisy OCT images, their latent sharp version, and the measurement noise as $\mathbf{M}$, $\mathbf{L}$ and $\mathbf{N}$, respectively (all in logarithm space). Mathematically, the image

formation equation can be written as

$$\mathbf{M} = \mathbf{L} + \mathbf{N}. \tag{5}$$

Representing each frame with a column vector, the dimensions of $\mathbf{M}$, $\mathbf{L}$, and $\mathbf{N}$ are all $(m \times n) \times k$. After registration, theoretically the entries at one specific row of $\mathbf{L}$ should be exactly the same, just as Fig.1 shows. Therefore, in mathematics we can treat $\mathbf{L}$ as a low rank matrix and use its nuclear norm $\mathbf{L}$ as $||\mathbf{L}||_*$ as inter-frame priors in denoising.

According to statistical studies[29][30], the adjacent pixels in natural images are of similar intensities and thus the distribution of image gradient centers around zero at large probability, as shown in Fig.1. In the task of natural image restoration such as denoising, sparse gradient has been widely used and effective prior. Although captured via a different imaging mechanism from usual CCD imaging, the OCT tissue images still follow similar statistics and we also impose gradient sparsity of the latent OCT images as intra-frame prior. Specifically, we minimize its l_0 norm $||\nabla\mathbf{L}||_{l_0}$ which counts the number of non-zero entries. In implementation, we use matrix multiplication for gradient calculation because image gradient is usually calculated by a high frequency filtering. Adopting the same representation in [31], we replace $||\nabla\mathbf{L}||_{l_0}$ as $\sum_{a=1}^{2}||\mathbf{H}_a\mathbf{L}||_{l_0}$, where $\mathbf{H}_1$ and $\mathbf{H}_2$ are gradient operators defined as the diagonal matrices of corresponding h_1 and h_2, i.e., [-1,1] and [-1;1].

As mentioned before, speckle noise includes both the image information and a single noise component whose mean value is zero. What we concern most is the removal of the noise part. So in the following, we treat the first information component of speckle as a part of image groundtruth and concentrate on attenuating the zero-mean noise component. According to the three sigma rule which indicates that nearly all (99.73%) of the instances of a random variable lie within 3 times standard deviation of its mean, we can formulate the noise constraint as

$$|\mathbf{N}| \leq 3\boldsymbol{\sigma} \tag{6}$$

where $\boldsymbol{\sigma}$ is the standard deviation matrix whose dimension is $(m \times n) \times k$ as well. By introducing a nonnegative matrix variable $\boldsymbol{\varepsilon}$, we can replace the above inequality with an equality as

$$\mathbf{N} \odot \mathbf{N} - 9\boldsymbol{\sigma} \odot \boldsymbol{\sigma} + \boldsymbol{\varepsilon} = \mathbf{0} \tag{7}$$

where $\odot$ is the component-wise product that for any two matrices $\mathbf{X}$ and $\mathbf{Y}$, i.e., $(\mathbf{X} \odot \mathbf{Y})_{ij} = \mathbf{X}_{ij}\mathbf{Y}_{ij}$.

Based on the above notations, the optimization model can be defined as

$$\begin{aligned} \{\mathbf{L}^*, \mathbf{N}^*\} = \arg\min \quad & ||\mathbf{L}||_* + \lambda \sum_{a=1}^{2} ||\mathbf{H}_a\mathbf{L}||_{l_0} \\ s.t. \quad & \mathbf{M} = \mathbf{L} + \mathbf{N}, \\ & \mathbf{N} \odot \mathbf{N} - 9\boldsymbol{\sigma} \odot \boldsymbol{\sigma} + \boldsymbol{\varepsilon} = \mathbf{0} \end{aligned} \tag{8}$$

where λ is a positive weighting parameter balancing different terms.

2.3. Derivation

The above model defined in (8) is obviously non-convex, so we firstly conduct some transformations to convexify it and resort to convex optimization method for the final solution. As shown in [32][33], replacing l_0 norm with l_1 norm is one typical convexification transformation, and we replace $||\mathbf{H}_a\mathbf{L}||_{l_0}$ with $||\mathbf{H}_a\mathbf{L}||_{l_1}$, where $||\cdot||_{l_1}$ denotes the sum of the absolute values of matrix entries. Further, it is hard to get its closed form solution of the convexified objective

funcion due to the high nonlinearity. To decrease the nonlinearity, we replace the variables, whose nuclear norm or l_1 norm needs minimization, with introduced auxiliary variables $\mathbf{S}_1$ and $\mathbf{S}_2$. Besides, we pack $\sum_{a=1}^{2}||\mathbf{H}_a\mathbf{L}||_1$ as $||\mathbf{PL}||_1$ for computation simplicity, where $\mathbf{P}=[\mathbf{H}_1;\mathbf{H}_2]$. Now we can rewrite the model as

$$\begin{aligned} min. \quad & ||\mathbf{S}_1||_* + \lambda||\mathbf{S}_2||_1 \\ s.t. \quad & \mathbf{G}_1 = \mathbf{S}_1 - \mathbf{L}, \\ & \mathbf{G}_2 = \mathbf{S}_2 - \mathbf{PL}, \\ & \mathbf{G}_3 = \mathbf{M} - \mathbf{L} + \mathbf{N}, \\ & \mathbf{G}_4 = \mathbf{N} \odot \mathbf{N} - 9\boldsymbol{\sigma} \odot \boldsymbol{\sigma} + \boldsymbol{\varepsilon} \end{aligned} \tag{9}$$

where $\mathbf{G}_{1...4}$ are supposed to be $\mathbf{0}$ in theory.

The augmented Lagrangian of the formulation in Eq.(9) is derived as

$$f = ||\mathbf{S}_1||_* + \lambda||\mathbf{S}_2||_1 + \sum_{j=1}^{4}\left(<\mathbf{Y}_j, \mathbf{G}_j> + \frac{\theta}{2}||\mathbf{G}_j||_F^2\right) \tag{10}$$

where $<\cdot,\cdot>$ denotes the inner product, $\mathbf{Y}$ defines the Lagrangian Multipliers (in matrix form) and θ is the parameter balancing the equation constraints. According to the lagrangian formulation, we derive the updating rules for $\mathbf{S}, \mathbf{L}, \mathbf{N}, \varepsilon$ and $\mathbf{Y}$ respectively as follows according to the augmented lagrange multiplier (ALM) algorithm in [27].

Optimize S. By removing all the items irrelevant to $\mathbf{S}_1$ in f, we can get

$$\begin{aligned} f(\mathbf{S}_1) &= ||\mathbf{S}_1||_* + <\mathbf{Y}_1, \mathbf{G}_1> + \frac{\theta}{2}||\mathbf{G}_1||_F^2 \\ &= ||\mathbf{S}_1||_* + \frac{\theta}{2}||\mathbf{G}_1 + \theta^{-1}\mathbf{Y}_1||_F^2 \\ &= ||\mathbf{S}_1||_* + \frac{\theta}{2}||\mathbf{S}_1 - (\mathbf{L}^{(k)} - \theta^{-1}\mathbf{Y}_1^{(k)})||_F^2. \end{aligned}$$

According to ALM algorithm, we can get the updating rule of $\mathbf{S}_1$ as

$$\mathbf{S}_1^{(k+1)} = U s_{\theta^{-1}}(S_{temp}) V^T, \tag{11}$$

where $US_{temp}V^T$ is the SVD of $\mathbf{L}^{(k)} - \theta^{-1}\mathbf{Y}_1^{(k)}$, and

$$s_{\theta^{-1}}(x) = \begin{cases} x - \theta^{-1}, & x > \theta^{-1} \\ x + \theta^{-1}, & x < -\theta^{-1} \\ 0, & others. \end{cases}$$

Similarly, keeping items related to $\mathbf{S}_2$ in f yields

$$\begin{aligned} f(\mathbf{S}_2) &= \lambda||\mathbf{S}_2||_1 + \frac{\theta}{2}||\mathbf{S}_2 - \mathbf{PL}^{(k)} + \theta^{-1}\mathbf{Y}_2^{(k)}||_F^2 \\ &= \lambda\left(||\mathbf{S}_2||_1 + \frac{\theta}{2\lambda}||\mathbf{S}_2 - (\mathbf{PL}^{(k)} - \theta^{-1}\mathbf{Y}_2^{(k)})||_F^2\right), \end{aligned}$$

so we can get the updating rule of $\mathbf{S}_2$

$$\mathbf{S}_2^{(k+1)} = s_{\frac{\lambda}{\theta}}(\mathbf{PL}^{(k)} - \theta^{-1}\mathbf{Y}_2^{(k)}). \tag{12}$$

Optimize L. Keeping only the items related to $\mathbf{L}$, f is simplified into

$$\begin{aligned} f(\mathbf{L}) &= \sum_{j=1}^{3}\left(<\mathbf{Y}_j,\mathbf{G}_j>+\frac{\theta}{2}||\mathbf{G}_j||_F^2\right) \\ &= \frac{\theta}{2}||\mathbf{S}_1-\mathbf{L}+\theta^{-1}\mathbf{Y}_1||_F^2+\frac{\theta}{2}||\mathbf{S}_2-\mathbf{PL}+\theta^{-1}\mathbf{Y}_2||_F^2+\frac{\theta}{2}||\mathbf{M}-\mathbf{L}-\mathbf{N}+\theta^{-1}\mathbf{Y}_3||_F^2 \end{aligned}$$

The partial derivative of $f(\mathbf{L})$ to $\mathbf{L}$ is

$$\begin{aligned} \frac{\partial f(\mathbf{L})}{\partial \mathbf{L}} &= \theta(\mathbf{L}-\mathbf{S}_1^{(k)}-\theta^{-1}\mathbf{Y}_1^{(k)})+\theta[\mathbf{P}^T\mathbf{PL}-\mathbf{P}^T(\mathbf{S}_2^{(k)}+\theta^{-1}\mathbf{Y}_2^{(k)})] \\ &+ \theta(\mathbf{L}-\mathbf{M}+\mathbf{N}^{(k)}-\theta^{-1}\mathbf{Y}_3^{(k)}). \end{aligned}$$

Obviously, it's hard to compute the value of $\mathbf{L}$ satisfying $\frac{\partial f(\mathbf{L})}{\partial \mathbf{L}}=0$. So we resort to gradient descent method to update $\mathbf{L}$ approximatively:

$$\mathbf{L}^{(k+1)}=\mathbf{L}^{(k)}-\Delta\times\frac{\partial f(\mathbf{L})}{\partial \mathbf{L}}|_{\mathbf{L}=\mathbf{L}^{(k)}} \tag{13}$$

Here Δ is the descent step parameter.

Optimize N and $\boldsymbol{\varepsilon}$. Similar to updating $\mathbf{L}$, the closed form solution to $\frac{\partial f(\mathbf{N})}{\partial \mathbf{N}}=0$ is not easy to get either. Therefore, we here take derivative with respect to the optimization variables and adopt gradient descent approach to update $\mathbf{N}$ iteratively as follows.

$$\mathbf{N}^{(k+1)}=\mathbf{N}^{(k)}-\Delta\times\frac{\partial f(\mathbf{N})}{\partial \mathbf{N}}|_{\mathbf{N}=\mathbf{N}^{(k)}}, \tag{14}$$

with the partial derivative $\frac{\partial f(\mathbf{N})}{\partial \mathbf{N}}$ defined as

$$\frac{\partial f(\mathbf{N})}{\partial \mathbf{N}} = \theta\pi(\mathbf{N}-\mathbf{M}+\mathbf{L}^{(k)}-\theta^{-1}\mathbf{Y}_3^{(k)})+\theta[\mathbf{N}\odot\mathbf{N}-9\boldsymbol{\sigma}\odot\boldsymbol{\sigma}+\boldsymbol{\varepsilon}^{(k)}+\theta^{-1}\mathbf{Y}_4^{(k)}]\odot 2\mathbf{N}.$$

.

The derivative of f with respect to ε is

$$\frac{\partial f(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}}=\theta(\boldsymbol{\varepsilon}+\mathbf{N}^{(k)}\odot\mathbf{N}^{(k)}-9\boldsymbol{\sigma}\odot\boldsymbol{\sigma}+\theta^{-1}\mathbf{Y}_4^{(k)})$$

and under the nonnegative assumption on $\boldsymbol{\varepsilon}$, we can get its updating rule as

$$\boldsymbol{\varepsilon}^{(k+1)}=\max\left(9\boldsymbol{\sigma}\odot\boldsymbol{\sigma}-\mathbf{N}^{(k)}\odot\mathbf{N}^{(k)}-\theta^{-1}\mathbf{Y}_4^{(k)},\ \mathbf{0}\right). \tag{15}$$

For more clarity, the entire algorithm based on the above derivations is summarized as follows (please see the attached derivation material for detailed derivation).

2.4. Evaluation criteria

To qualify the performance of our approach compared to other methods, we utilize two kinds of the most common image quality criteria, namely peak signal-to-noise ratio (PSNR) and structure similarity (SSIM) [34] as metrics to evaluate recovered images.

Algorithm 1: Algorithm for OCT image Denoising

Input : Capturing data $\mathbf{M}$, Noise standard deviation $\boldsymbol{\sigma}$
Output : Denoised frames $\mathbf{L}$ and separated noise $\mathbf{N}$.

$\mathbf{L}^{(0)} = \overline{\mathbf{M}}$, $\mathbf{N}^{(0)} = \mathbf{M} - \mathbf{L}^{(0)}$, $\boldsymbol{\varepsilon}^{(0)} = \mathbf{0}$; $\mathbf{Y}_{1\cdots4} = \mathbf{0}$;
while *not converged* **do**
  Update $\mathbf{L}^{(k+1)}$ according to (13);
  Update $\mathbf{S}_{\{1,2\}}^{(k+1)}$ according to (11) and (12);
  Update $\mathbf{N}^{(k+1)}$ according to (14);
  Update slack variables $\boldsymbol{\varepsilon}^{(k+1)}$ according to (15);
  $\mathbf{Y}_{\{1\ldots4\}}^{(k+1)} = \mathbf{Y}_{\{1\ldots4\}}^{(k)} + \theta^{-1}\mathbf{G}_{\{1\ldots4\}}$;
  $\theta^{(k+1)} = \min\left(\rho\theta^{(k)}, \theta_{max}\right)$;
  $k := k+1$;
end

PSNR analysis: Peak signal-to-noise ratio has traditionally been used to assess the image quality of image $\mathbf{L}_{m\times n}$ regarding to the groundtruth image $\mathbf{I}_{m\times n}$, which is calculated as

$$\begin{aligned} PSNR &= 10\times\log_{10}\left(\frac{MAX^2}{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[\mathbf{L}(i,j)-\mathbf{I}(i,j)]^2}\right) \\ &= 10\times\log_{10}\left(\frac{mn\times MAX^2}{\sum_{i=1}^{m}\sum_{j=1}^{n}[\mathbf{L}(i,j)-\mathbf{I}(i,j)]^2}\right) \end{aligned} \tag{16}$$

where $MAX = 2^b - 1$ is the maximum possible value of image pixels represented by using b bits. For example, if the pixels are stored using 8 bits, then $MAX = 2^8 - 1 = 255$. From the equation we can see that PSNR mainly describes the value difference of pixels between two images, and would be smaller with greater difference. Thus we expect that the restored image better owns great PSNR. In general, typical values of PSNR are between 20 dB and 40 dB, provided that the bit depth is 8.

SSIM analysis: Structure similarity is proposed in [34] designed to extract structural information from images and assess structural similarities between two images. The criterion method firstly select two sets of patches from $\mathbf{L}$ and $\mathbf{I}$, and then calculate SSIM between two corresponding patches p_l and p_i as

$$SSIM(p_l, p_i) = \frac{(2\mu_l\mu_i + c_1)(2\sigma_{li} + c_2)}{(\mu_l^2 + \mu_i^2 + c_1)(\sigma_l^2 + \sigma_i^2 + c_2)} \tag{17}$$

where μ_l and μ_i are respectively the average pixel values of patch p_l and p_i, while σ_l and σ_i are the patches' variances. σ_{li} is the covariance between p_l and p_i. c_1 and c_2 are calculated as $c_1 = (k_1 MAX)^2$, $c_2 = (k_2 MAX)^2$, where k_1 and k_2 are two variables set by users whose default values are 0.01 and 0.03, respectively. The final SSIM value between two images is the average of all the patches' SSIM values.

SSIM concerns on images' structural similarity and thus improves the assessment ability of traditional methods such as PSNR, which only reveals single pixels' difference between two images. SSIM ranges from 0 to 1, and is greater when two images are more similar to each other structurally.

3. Experimental Results

In this section, we implement our denoising method on OCT noisy images, and compare our denosing results with that of several classical methods including complex diffusion method [35], Bayesian method [18], BM3D [36] and multi-frame wavelet OCT denoising method [8]. Besides, runtime and convergence analysis together with study on algorithm performance due to different values of the model parameter λ are also conducted.

3.1. Data and parameters

We firstly illustrate the source of OCT data and the values of algorithm parameters we use for our experiments in this subsection. For OCT data, we use the public pigeye OCT dataset published in [37], which is also used as experiment source data in [8]. The dataset is acquired by using a Spectralis HRA & OCT (Heidelberg Engineering) to scan a pig eye in high speed mode with 768 A-scans. There are totally 455 images in the dataset including 35 sets, and 13 frames are included in each set sharing the same imaging position. There exists slight misalignment between two different sets for different imaging positions, resulting that speckle noise is decoupled between the scans.

To assess recovered images' quality, we need a noise-free benchmark image for comparison. However, due to the high speed of imagery, the original frames' signal-to-noise ratio is very low. So we consider using the averaged image of all the 455 frames which have been registered in advance for image quality assessment, which is same as the technique in [8].

There are several important parameters in our algorithm whose values should be declared ahead of time as well. As shown in Alg.1, the algorithm parameters include λ, $\theta^{(0)}$, ρ, θ_{max}, and Δ. λ is the model parameter to decide the regulation relationship between low-rank objective $||\mathbf{L}||_*$ and sparsity objective $||\nabla\mathbf{L}||_1$, as shown in Eq.(8), and is set to be 0.2 for the best algorithm performance. $\theta^{(0)}$, ρ, and θ_{max} are algorithm parameters to determine the varying value of iterative θ. Here we set $\theta^{(0)} = 1e-2$, $\rho = 1.6$, and $\theta_{max} = 1e1$. Δ is the descent step parameter for $\mathbf{L}$ and $\mathbf{N}$ in the iteration process. Considering different value characters of $\mathbf{L}$ and $\mathbf{N}$, we choose different descent steps for them as $\Delta_{\mathbf{L}} = 5e-3$ and $\Delta_{\mathbf{N}} = 1e-2$.

3.2. Comparison to other methods

In this section, we run our algorithm and the other four kinds of previous methods on the OCT dataset, and compare algorithms' performance in both recovered image quality and algorithm runtime. What should be noticed is that the first two methods, namely complex diffusion method and Bayesian method are both single-frame methods. So we take the averaged image of registered frames as algorithm input. Besides, for the first three methods, the standard deviation is assumed to be space-invariant. To tackle with this problem, we use the maximum value of deviation matrix estimated by MAD method described in Sec.(2.1) as the input noise standard deviation. Utilizing the above techniques, we run Matlab implementations of the overall five methods on an Intel Xeon 2.27 GHz CPU workstation with 64 bit Windows 7 system. By random selection, the serial numbers of input sequential frames including 8 frames range from '35_6' to '35_13'. Recovered images are shown in Fig.(2), where only one of eight images is presented for each method. Numerical image assessments including PSNR and SSIM of every method , together with running time are recorded in Tab. 1.

Referring to Fig.2, we can see that the recovered images of anisotropic diffusion method and Bayesian method still contains undesired noise, which affects image quality seriously. BM3D has a better performance than the above two methods, but there exists some artifacts similar to 'watermarks', and too many details are attenuated, which makes the image relatively blurred. Multi-frame wavelet method and our method are both superior to the first three approaches. Comparing the results of these two methods, we can see that in smooth regions, the wavelet

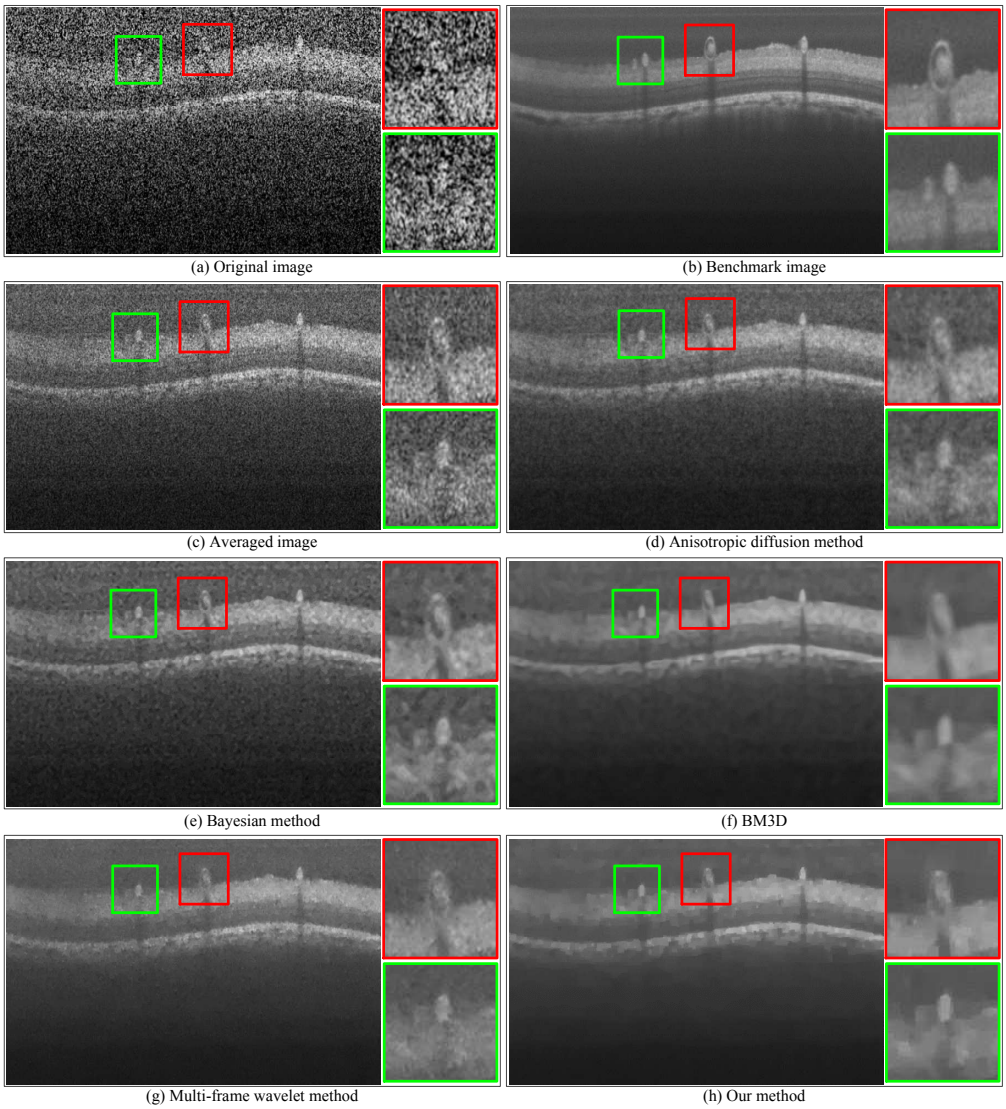


Fig. 2. **Comparison with state-of-the-arts.** (a) is the original image in log space, while (b) is the averaged image of 455 registered frames. (c) is the averaged image of the input 8 frames, and (d)-(g) are results of four comparison methods. The result of our method is shown in (h). The two clipped patches on the right of every subfigure are closeups of the regions of interest.

method leaves out more noise than our method, while in textured regions the result of our method maintains higher color contrast which would improve visual effects. Stronger comparison is presented in the closeups. For example, in the green rectangle clip region, wavelet method nearly blurs the information of the left white spot, while our method still contains grey value changes which would offer some important information for diagnosis.

More numerical assessments are shown in Tab.1. We can see that our method could increase noisy image's PSNR from 17.19 to 31.74, and increase SSIM from around 0.13 to 0.91. Regarding of PSNR and SSIM, our method is superior to all the other methods. Considering the only

Table 1. Quantitative comparison of different denoising methods

	Metric	Input	Average	Diffusion	Bayesian	BM3D	Wavelet	Ours
Entire Image	PSNR(dB)	17.19	24.56	29.14	28.38	31.32	30.75	**31.74**
	SSIM	0.1272	0.4515	0.7359	0.7043	0.8982	0.8650	**0.9121**
	Running time(s)	—	—	79	33	49	477	519
Red Clip	PSNR(dB)	15.03	22.02	26.60	26.07	28.24	27.85	**28.92**
	SSIM	0.0631	0.2921	0.6504	0.6346	**0.8134**	0.7378	0.8124
Green Clip	PSNR(dB)	15.13	21.91	26.14	25.35	27.97	27.83	**28.75**
	SSIM	0.0604	0.2547	0.6094	0.5749	0.7901	0.7232	**0.8033**

two kinds of multi-frame methods, namely the multi-frame wavelet method and our method, it can be concluded that our method obtains higher PSNR and SSIM respectively by 1 and 0.1 than wavelet method on average. In addition, one of the single-frame methods, namely BM3D, behaves very well regarding of the two numerical criteria, which is better than the other methods except for our approach. This is not surprised considering that the averaged image of 8 frames is utilized as the input of the three single-frame denoising methods.

Runtime results are also shown in Tab.1, from which we can see that three kinds of single-frame methods own the same runtime level under 100s, while two multi-frame methods need around 500s to recover all the 8 frames. What should be noticed is that for the three single-frame methods, the showing items are only the runtime for recovering one single frame. If we replace the averaged image input with eight frames in turn and ignore the decreased image quality, then for each of the three single-frame methods, the algorithm runtime would be around eight times of the present interval. In this view, we can conclude that the diffusion method is the most time-consuming, followed by our method, the wavelet method and BM3D. The Bayesian method owns the minimum runtime.

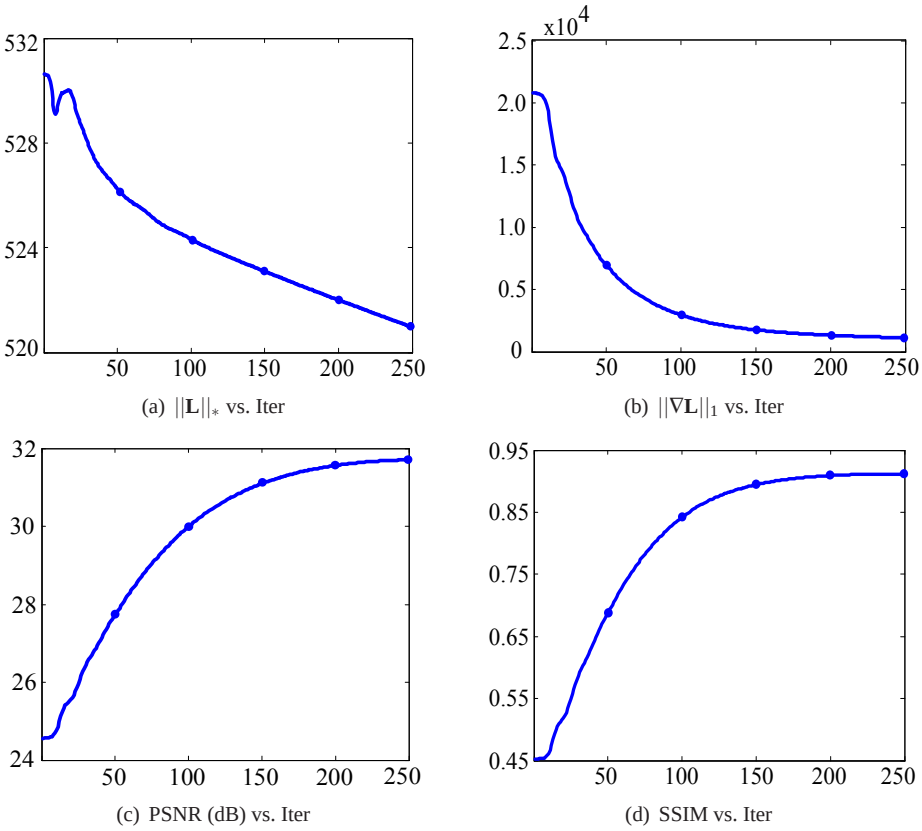
3.3. Convergence analysis

The convergence properties of using l_1 norm to replace l_0 norm and using ALM algorithm framework to solve l_1 problem have been proven in a lot of work[27][32][33]. Here we don't give extra theoretical analysis about algorithm convergence. In stead, we study some objective variables' evolution that we are interested in during the iteration, which would reveal more characters of our optimization framework and would show convergence properties from the numerical point of view.

What we concern most is the evolutions of the two objective items in our model and the recovered images' quality. So according to the above analysis, we record all the values of $||\mathbf{L}||_*$, $||\nabla\mathbf{L}||_1$ and recovered image's two quality criteria values, namely PSNR and SSIM, in every iteration. The results are shown in Fig.3, from which we can see that values of two objectives both descend in the process, while values of two criteria both increase. Specifically, $||\nabla\mathbf{L}||_1$, PSNR and SSIM all approach to stable values from around 200th iteration. For $||\mathbf{L}||_*$, though the values fluctuate at beginning, it descends steadily in the following process. Comparing value's decrease proportion of $||\mathbf{L}||_*$ with that of $||\nabla\mathbf{L}||_1$, we can see that $||\nabla\mathbf{L}||_1$ has a more obvious attenuation.

3.4. Model parameter behavior

In the model represented in Eqs.(8), the only model parameter that we can adjust is λ, which controls the regulation relationship between the low rank item $||\mathbf{L}||_*$ and the sparsity item $||\nabla\mathbf{L}||_1$. To be more specific, the regulation of low rank item will become loose compared to that of the sparsity item as λ increases. To shed light on the impact caused by value changes of λ on optimization progress and final results, the value of λ is set to range from 0.05 to 0.5, and



(a) $||\mathbf{L}||_*$ vs. Iter (b) $||\nabla\mathbf{L}||_1$ vs. Iter

(c) PSNR (dB) vs. Iter (d) SSIM vs. Iter

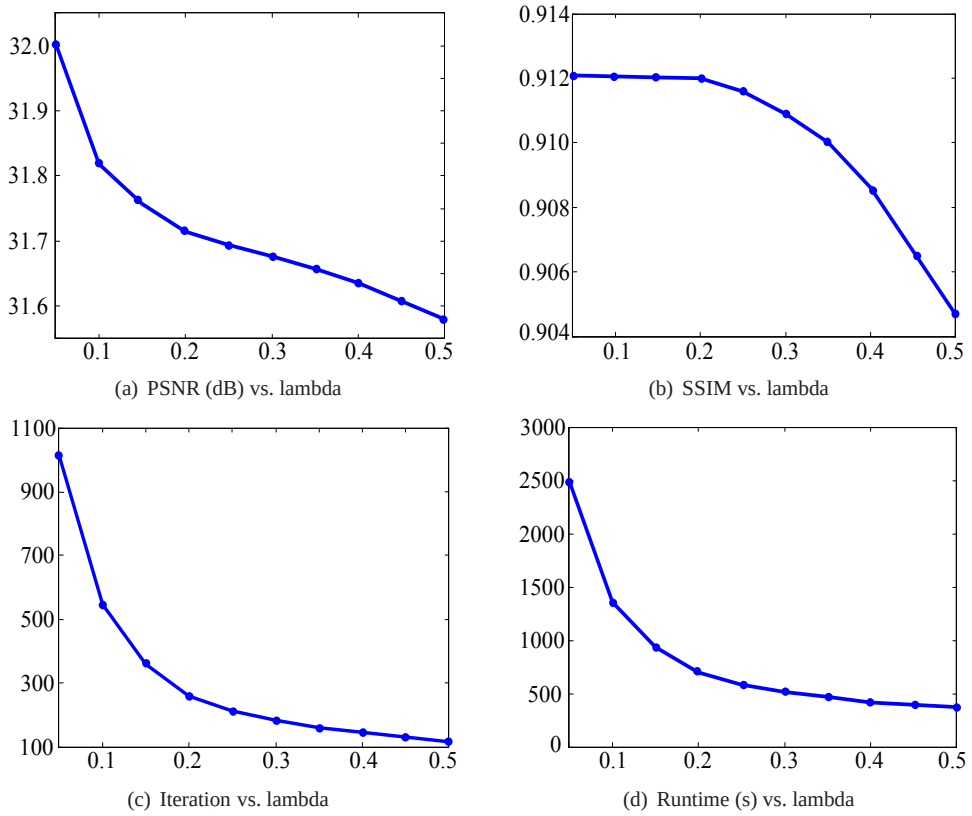
Fig. 3. **Numerical analysis on the evolution of objective terms and quantitative evaluations.** (a) represents the iterative changes of $||\mathbf{L}||_*$ standing for the rank of $\mathbf{L}$. The evolution of image gradients' sparsity $||\nabla\mathbf{L}||_1$ is shown in (b). (c) and (d) respectively present value changes of PSNR and SSIM in iteration process.

we record iterations and runtime of our algorithm, together with PSNR and SSIM of recovered images for each value of λ. The results are shown in Fig.4.

From the figure we can see that it will take more iterations and time to converge for our algorithm as λ increases, while the quality of recovered images becomes better. Specifically, iteration and runtime share the same variation trend, and both increase acceleratedly as λ decreases. SSIM doesn't change much until λ is greater than 0.2, while PSNR owns fatest descent rate when λ is smaller than 0.2. So how to choose an appropriate value for λ is related to the tradeoff between image quality and algorithm runtime. If one concerns more about the algorithm's running speed, then λ should be set to be greater. In the contrary, λ should be set smaller if there are stronger requirements for recovered images' quality.

4. Conclusion and Discussion

In this study, we propose a multi-frame OCT denoising method utilizing two priors including an inter-frame prior and an intra-frame prior. Specifically, the inter-frame prior means low rank of registered OCT frames and the intra-frame prior is sparsity of image gradient. In our optimization framework we derived for attenuating speckle noise, logarithmic transformation and image registration are firstly done to the original captured OCT frames. Then we use nuclear



(a) PSNR (dB) vs. lambda

(b) SSIM vs. lambda

(c) Iteration vs. lambda

(d) Runtime (s) vs. lambda

Fig. 4. **Numerical analysis on algorithm performance under different values of λ.** (a) and (b) respectively show PSNR and SSIM of recovered images with respect to that λ ranges from 0.05 to 0.5, while (c) and (d) record corresponding iterations and algorithm runtime.

norm to regulate the rank of image matrix in which each frame is aligned to one column, and use l_0 norm to represent the sparsity regulation of corresponding image gradient matrix. Together with image formation in log space and restriction of noise range, we build a nonconvex model for multiple OCT frames containing two objective items and two constraints. To make the model solvable and get a global optimal solution, the replacement of l_0 norm with l_1 norm and some other approximations are made to the model to transform the non-convex model to a convex one. Next, an iterative optimization framework based on ALM algorithm is derived to solve the model. Thus, by applying the algorithm to the transformed convex model, we could separate speckle noise from ideal signals and get the final recovered images.

We run our algorithm as well as another four kinds of previous methods on the pigeye OCT volume dataset available on the internet, and compare the performance of our method with that of them. The four kinds of methods are respectively complex diffusion method, Bayesian method , BM3D and multi-frame wavelet OCT denoising method. Besides, the comparison is conducted regarding of three criteria that includes PSNR, SSIM of recovered images and algorithm runtime. The results reveal that our method is superior to the other four approaches for that our method could effectively attenuate speckle noise while preserving more important image details. For numerical assessment, our approach could improve OCT image's quality by increasing image's PSNR from 17.19 to 31.74 and image's SSIM from 0.1272 to 0.9121,

which are better than the other four methods. What's more, influence on algorithm performance caused by the regulation model parameter λ is also studied and we could conclude that though PSNR and SSIM both increase with the increase of λ, corresponding iteration and runtime also increase logarithmically. This would guide one to choose appropriate value for λ according to different weigh between image quality and runtime.

In a word, the proposed approach could recover high quality noise-free OCT images from a little number of original captured frames, which would do a lot of help to occasions requiring tiny image details, such as medical diagnosis and pathological analysis. The method briefly benefits from the development of high speed OCT imaging systems, which assures that we could capture a set of OCT frames in little time and that misalignment among serial frames is slight. Besides, due to the non-parametric and non-stationary noise model we derived, our approach could be widely applied to other kinds of multi-frame image denoising tasks with different noise models.