QuaSI: Quantile Sparse Image Prior for Spatio-Temporal Denoising of Retinal OCT Data

Franziska Schirrmacher[1,*], Thomas Köhler[1,*], Lennart Husvogt[1],
James G. Fujimoto[2], Joachim Hornegger[1], and Andreas K. Maier[1]

[1] Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
{franziska.schirrmacher, thomas.koehler}@fau.de
[2] Department of Electrical Engineering & Computer Science and Research Laboratory
of Electronics, Massachusetts Institute of Technology
https://www5.cs.fau.de/en/research/software/idaa/
[*] These authors contributed equally to this work.

Abstract. Optical coherence tomography (OCT) enables high-resolution and non-invasive 3D imaging of the human retina but is inherently impaired by speckle noise. This paper introduces a spatio-temporal denoising algorithm for OCT data on a B-scan level using a novel quantile sparse image (QuaSI) prior. To remove speckle noise while preserving image structures of diagnostic relevance, we implement our QuaSI prior via median filter regularization coupled with a Huber data fidelity model in a variational approach. For efficient energy minimization, we develop an alternating direction method of multipliers (ADMM) scheme using a linearization of median filtering. Our spatio-temporal method can handle both, denoising of single B-scans and temporally consecutive B-scans, to gain volumetric OCT data with enhanced signal-to-noise ratio. Our algorithm based on 4 B-scans only achieved comparable performance to averaging 13 B-scans and outperformed other current denoising methods.

1 Introduction

Since its invention in 1991, optical coherence tomography (OCT) [2] has become a standard imaging technique within clinical workflows in ophthalmology. OCT enables non-invasive 3D imaging of retinal layers with spatial resolutions in a micrometer range. These properties contributed to its wide distribution for diagnosis and disease monitoring as well as for computer-aided diagnostics [8]. Besides these merits, OCT suffers from a low signal-to-noise ratio (SNR) due to speckle noise caused by photon interference. Hardware-based techniques, e. g. frequency compounding [12], are able to reduce speckle noise but increase the complexity of the system. Thus, image-based post-processing algorithms to improve the reliability of retinal OCT data are attractive. However, many popular denoising methods such as BM3D [3] have been mainly designed for natural images and can handle non-Gaussian noise only to a certain extend. Furthermore, noise reduction in OCT is a sensitive issue as the preservation of tiny morphological structures is a mandatory requirement. *Spatial* or *single-image* methods perform

noise reduction by filtering single B-scans. Spatial filters suited for OCT are hybrid median filter, Lee filter, Wiener filter, or wavelet thresholding [11]. Popular global methods include non-linear diffusion [14], variational formulations [4], or structure-adaptive Bayesian estimation [15]. More recently, sparse coding [5] using high SNR scans for dictionary learning to denoise single low SNR scans has been proposed. The spatial methods have in common that noise reduction is limited as they utilize single B-scans only. *Temporal* or *multi-image* methods exploit several B-scans acquired sequentially from the same location or nearby positions. A simple method often implemented for commercial systems is averaging of multiple registered B-scans. Recent approaches are wavelet multi-image denoising [9] or matrix completion [1] that have been customized to speckle noise reduction and outperform simple averaging. However, temporal methods require longer acquisition times to gain multiple B-scans and hence increase patient discomfort.

This paper introduces a new *spatio-temporal* OCT denoising algorithm. Our contribution is two-fold: 1) We propose denoising via energy minimization based on the novel class of quantile sparse image (QuaSI) priors. To deal with speckle noise while preserving morphological structures, we regularize with the popular median filter as a special instance of our QuaSI prior. 2) We develop an alternating direction method of multipliers (ADMM) scheme for optimization with the non-linear median filter. Our method can handle both, denoising of single and multiple registered B-scans. Source code of our method is available on our webpage.

2   Proposed Spatio-Temporal Denoising Algorithm

2.1   Noise Model and Energy Minimization Formulation

Our method aims at denoising volumetric OCT data, where a single volume is represented as a stack of B-scans $\tilde{\boldsymbol{G}} \in \mathbb{R}^{L \times N_x \times N_y}$. We denote the l-th B-scan of size $N_x \times N_y$ in vector notation as $\tilde{\boldsymbol{g}}_l \in \mathbb{R}^N$ with $l \in [1, L]$ and $N = N_x N_y$. Each noisy B-scan $\tilde{\boldsymbol{g}}_l$ in a given volume is related to the respective noise-free scan $\tilde{\boldsymbol{f}}_l$ according to multiplicative speckle noise. Following related denoising methods [15,4], we model speckle noise in a logarithmic measurement range according to $\boldsymbol{f}_l = \boldsymbol{g}_l + \boldsymbol{n}_l$, where $\boldsymbol{f}_l = \log(\tilde{\boldsymbol{f}}_l)$, $\boldsymbol{g} = \log(\tilde{\boldsymbol{g}}_l)$ and $\boldsymbol{n}_l \in \mathbb{R}^N$ is additive noise.

Let $\boldsymbol{g}^{(1)}, \ldots, \boldsymbol{g}^{(K)}$ be a set of B-scans that are captured from the same location and registered to each other. We estimate a denoised B-scan $\hat{\boldsymbol{f}}$ according to:

$$\hat{\boldsymbol{f}} = \underset{\boldsymbol{f}}{\arg\min}\, \mathcal{L}(\boldsymbol{f}) = \underset{\boldsymbol{f}}{\arg\min} \sum_{k=1}^{K} \rho\big(\boldsymbol{f} - \boldsymbol{g}^{(k)}\big) + \mu \|\nabla \boldsymbol{f}\|_1 + \lambda R_{\text{QuaSI}}(\boldsymbol{f}). \qquad (1)$$

In (1), the first term denotes the fidelity of $\boldsymbol{f}$ w.r.t. K observed noisy B-scans $\boldsymbol{g}^{(k)}$, $k = 1, \ldots, K$ according to the loss function $\rho : \mathbb{R}^N \to \mathbb{R}_0^+$. The second term is the anisotropic total variation (TV) to regularize the image gradient $\nabla \boldsymbol{f} = (\nabla_x \boldsymbol{f}, \nabla_y \boldsymbol{f})^\top$ with the weight $\mu \geq 0$. The third term denotes regularization according to our proposed QuaSI prior with the weight $\lambda \geq 0$, see Section 2.2.

To define the data fidelity term in (1), we propose to use outlier-insensitive loss functions. In this paper, we use the Huber loss [10] for $\rho(\cdot)$ to model the

image formation in retinal OCT. It is worth noting that the Huber loss can tackle outliers related to non-Gaussian noise while being convex and easy to optimize.

2.2   Quantile Sparse Image Prior

The novel class of priors that we propose is based on quantile filtering. We denote a quantile filter as $\tilde{\boldsymbol{f}} = Q(\boldsymbol{f})$, where $\tilde{f}_i = \text{quantile}_{\mathcal{N}(i)}(f_i, p)$ determines the p-quantile with $p \in [0, 1]$ within the local neighborhood $\mathcal{N}(i)$ centered at the i-th pixel in $\boldsymbol{f}$. Our prior is defined as fixed point under the quantile filter according to:

$$R_{\text{QuaSI}}(\boldsymbol{f}) = ||\boldsymbol{f} - Q(\boldsymbol{f})||_1 \,. \tag{2}$$

In general, our prior is non-convex. Similar forms have also become popular in other inverse problems, e. g. the dark channel prior [7] or regularization by denoising priors [13]. This general model facilitates regularization by various types of order statistics or parameters such as minimum or maximum, first or third quantile, weighted median, etc. To customize (2) for OCT denoising in this paper, we use the $p = 0.5$ quantile that is equivalent to the median. This implements a regularization by the popular median filter within our framework. We found that this facilitates structure-preserving denoising and handles non-Gaussian noise as important prerequisites for the desired application.

2.3   Alternating Direction Method of Multiplier Optimization

The optimization of (1) involves two non-smooth regularization terms related to the TV and our proposed QuaSI prior. For efficient minimization with the non-smooth terms, we adopt ADMM optimization [6]. To this end, we introduce the auxiliary variables $\boldsymbol{u}$ and $\boldsymbol{v}$ and re-formulate (1) via the augmented Lagrangian:

$$\begin{aligned}
\mathcal{L}_{\text{AL}}(\boldsymbol{f}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{b}_u, \boldsymbol{b}_v) = \sum_{k=1}^{K} \rho\big(\boldsymbol{f} - \boldsymbol{g}^{(k)}\big) + \frac{\alpha}{2}\|\boldsymbol{u} - \boldsymbol{f} + Q(\boldsymbol{f}) - \boldsymbol{b}_u\|_2^2 \\
+ \lambda\|\boldsymbol{u}\|_1 + \frac{\beta}{2}\|\boldsymbol{v} - \nabla\boldsymbol{f} - \boldsymbol{b}_v\|_2^2 + \mu\|\boldsymbol{v}\|_1,
\end{aligned} \tag{3}$$

where $\boldsymbol{b}_u$ and $\boldsymbol{b}_v$ denote Bregman variables, and $\alpha > 0$ and $\beta > 0$ are Lagrangian multiplier to enforce the constraints $\boldsymbol{u} = \boldsymbol{f} - Q(\boldsymbol{f})$ and $\boldsymbol{v} = \nabla\boldsymbol{f}$. We iteratively optimize (3) by alternating minimization w.r.t. the individual parameters.

   Notice that direct optimization of (3) is not tractable due to the non-linearity of the quantile operator $Q(\boldsymbol{f})$ and the quantile operator is first linearized as $Q(\boldsymbol{f}) = \boldsymbol{Q}\boldsymbol{f}$ [7]. If $Q(\boldsymbol{f})$ denotes median filtering, we assemble $\boldsymbol{Q}$ element-wise as a binary matrix according to $Q_{ij} = 1 \Leftrightarrow z = j$, where $j = \arg\text{median}_{z \in \mathcal{N}(i)} f_z$ denotes the position of the median in the neighborhood $\mathcal{N}(i)$ centered at the i-th pixel. This construction fulfills $Q(\boldsymbol{f}') = \boldsymbol{Q}\boldsymbol{f}'$ for $\boldsymbol{f}' = \boldsymbol{f}$, while otherwise we use $\boldsymbol{Q}$ as an approximation of the median filter. In the proposed ADMM scheme, we gradually update $\boldsymbol{Q}$ and assemble the linearization from the intermediate image $\boldsymbol{f}^t$ estimated at the previous iteration. Given this linearization, we minimize (3)

Algorithm 1 Spatio-temporal denoising algorithm using ADMM optimization

---

Initialize $\boldsymbol{u}^1 = \boldsymbol{v}^1 = \boldsymbol{0}$, $\boldsymbol{b}_u^1 = \boldsymbol{b}_v^1 = \boldsymbol{0}$ and $\boldsymbol{f}^1 = \text{mean}(\boldsymbol{g}^{(1)}, \dots, \boldsymbol{g}^{(K)})$
for $t = 1, \dots, T_{\text{outer}}$ **do**
    Assemble $\boldsymbol{Q}$ from the intermediate image $\boldsymbol{f}^t$
    for $i = 1, \dots, T_{\text{inner}}$ **do**
        Update the intermediate image $\boldsymbol{f}^{t+1}$ using CG iterations for (4)
        Update the auxiliary variables $\boldsymbol{u}^{t+1}$ and $\boldsymbol{v}^{t+1}$ using (5) - (6)
        Update the Bregman variables $\boldsymbol{b}_u^{t+1}$ and $\boldsymbol{b}_v^{t+1}$ using (7) - (8)
    end for
end for

---

w.r.t. the denoised image $\boldsymbol{f}$. To handle the Huber loss, this is done by iteratively re-weighted least squares (IRLS). This leads to the linear system:

$$\left[2 \sum_{k=1}^{K} \boldsymbol{W}^{(k)} + \alpha \left(\boldsymbol{I} - \boldsymbol{Q} \right)^{\top} \left(\boldsymbol{I} - \boldsymbol{Q} \right) + \beta \nabla^{\top} \nabla \right] \boldsymbol{f}^{t+1}$$
$$= 2 \sum_{k=1}^{K} \boldsymbol{W}^{(k)} \boldsymbol{g}^{(k)} + \alpha \left(\boldsymbol{I} - \boldsymbol{Q} \right)^{\top} \left(\boldsymbol{u} - \boldsymbol{b}_u \right) + \beta \nabla^{\top} (\boldsymbol{v} - \boldsymbol{b}_v), \tag{4}$$

where $\boldsymbol{W}^{(k)}$ denotes a diagonal weight matrix derived from the Huber loss. For IRLS, the weights are computed based on the intermediate image $\boldsymbol{f}^t$ according to $W_{ii}^{(k)} = \rho'(f_i - g_i^{(k)})/(f_i^t - g_i^{(k)})$, where $\rho'(z)$ is the derivative of the Huber loss [10]. The linear system in (4) is solved by conjugate gradient (CG) iterations.

The minimization of (3) w.r.t. to the auxiliary variables is separable and performed element-wise. We obtain closed-form solutions using shrinkage operations:

$$u_i^{t+1} = \text{shrink}\left([\boldsymbol{f}^{t+1} - \boldsymbol{Q}\boldsymbol{f}^{t+1} + \boldsymbol{b}_u^t]_i, \lambda/\alpha \right), \tag{5}$$
$$v_i^{t+1} = \text{shrink}\left([\nabla \boldsymbol{f}^{t+1} + \boldsymbol{b}_v^t]_i, \mu/\beta \right), \tag{6}$$

where $\text{shrink}(z, \gamma) = \text{sign}(z) \max(z - \gamma, 0)$ denotes soft-thresholding associated with the L_1 norm [6]. Finally, the Bregman variables are updated according to:

$$\boldsymbol{b}_u^{t+1} = \boldsymbol{b}_u^t + (\boldsymbol{f}^{t+1} - \boldsymbol{Q}\boldsymbol{f}^{t+1} - \boldsymbol{u}^{t+1}), \tag{7}$$
$$\boldsymbol{b}_v^{t+1} = \boldsymbol{b}_v^t + (\nabla \boldsymbol{f}^{t+1} - \boldsymbol{v}^{t+1}). \tag{8}$$

Algorithm 1 summarizes ADMM using $\boldsymbol{u}^1 = \boldsymbol{v}^1 = \boldsymbol{0}$, $\boldsymbol{b}_u^1 = \boldsymbol{b}_v^1 = \boldsymbol{0}$, and the average of the input B-scans as an initial guess $\boldsymbol{f}^1$. We found experimentally that for the convergence to a stationary point and the speed-up of ADMM it is sufficient to update the linearization only after a couple of iterations. For this reason, we use T_{inner} iterations to update $\boldsymbol{f}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{b}_u$, and $\boldsymbol{b}_v$ and T_{outer} iterations to update $\boldsymbol{Q}$.

3 Experiments and Results

We compare our method to the well known BM3D [3] as well as current OCT noise reduction algorithms, namely Bayesian estimation denoising (BED) [15],

averaging (AVG) of registered B-scans, and wavelet multi-frame denoising (WMF) [9]. As BM3D and BED handle single B-scans only, we apply these methods to the outcome of AVG for fair comparisons. WMF is a pure temporal approach and requires at least two registered B-scans. The parameters of our method were set to $\mu = 0.075 \cdot K$, $\lambda = 5.0 \cdot K$, $\alpha = 100.0 \cdot K$, $\beta = 1.5 \cdot K$, $T_{\text{outer}} = 20$ and $T_{\text{inner}} = 2$ for K B-scans and 3×3 median filtering to setup the QuaSI prior.

Pig Eye Data. To study the behavior of our algorithm quantitatively, we conducted experiments on the publicly available pig eye dataset provided by Mayer et al. [9]. The dataset was captured ex-vivo by scanning a pig eye with a Spectralis HRA & OCT and comprises subsets of 35 eye positions with 13 B-scans each. Following [9], a gold standard B-scan was computed by averaging the 455 scans that have already been registered to each other. We applied the competing denoising methods on subsets of K registered B-scans with $K \in [1, 13]$. Fig. 1 compares our method with and without the proposed QuaSI prior to simple averaging of consecutive B-scans for different numbers of input images. Notice that spatio-temporal denoising substantially enhanced noise reduction compared to averaging. We also found that regularization by our QuaSI prior in combination with the TV prior further enhanced noise reduction compared to TV denoising. Fig. 2 depicts the mean peak-signal-to-noise ratio (PSNR) and structural similarity (SSIM) relative to the gold standard. Here, our spatio-temporal method based on 4 B-scans only was comparable to averaging of 13 B-scans. Compared to TV denoising, the QuaSI prior enhanced the mean PSNR and SSIM by 0.9 dB and 0.03, respectively. The proposed method also outperformed BM3D, BED and WMF in terms of both measures.

Clinical Data. We also investigate denoising on clinical data which were acquired using a prototype ultrahigh-speed swept-source OCT system with 1050 nm wavelength and a sampling rate of 400,000 A-scans per second [2]. Each B-scan was acquired five times in direct succession and the B-scans were registered towards the central one using cross-correlation. We use OCT data from 14 human subjects with two volumes each. The data covers proliferative and non-proliferative diabetic retinopathy, early age-related macular degeneration and one healthy subject. The field size is 3×3 mm with 500 A-scans by 500 B-scans. Our experiments were conducted on the central B-scan of each volume. To quantify noise reduction in the absence of a gold standard, we use the mean-to-standard-deviation ratio $\text{MSR} = \mu_f / \sigma_f$ and the contrast-to-noise ratio $\text{CNR} = |\mu_f - \mu_b| / (0.5(\sigma_f^2 + \sigma_b^2))^{0.5}$ [5,11,15], where μ_i and σ_i, $i \in \{f, b\}$ denote the mean and standard deviation of the intensity in foreground ($i = f$) and background ($i = b$) image regions, respectively. Both measures were determined for five foreground regions and one background region that were manually selected for each B-scan, see Fig. 3a. Fig. 4 depicts the mean MSR and CNR for different numbers of input images. Here, BM3D and our method achieved the best denoising performance in terms of both measures. In particular, the combination of the TV and QuaSI priors consistently outperformed the competing methods. Fig. 3 compares our approach and two competing methods on one example dataset. WMF enabled structure-preserving

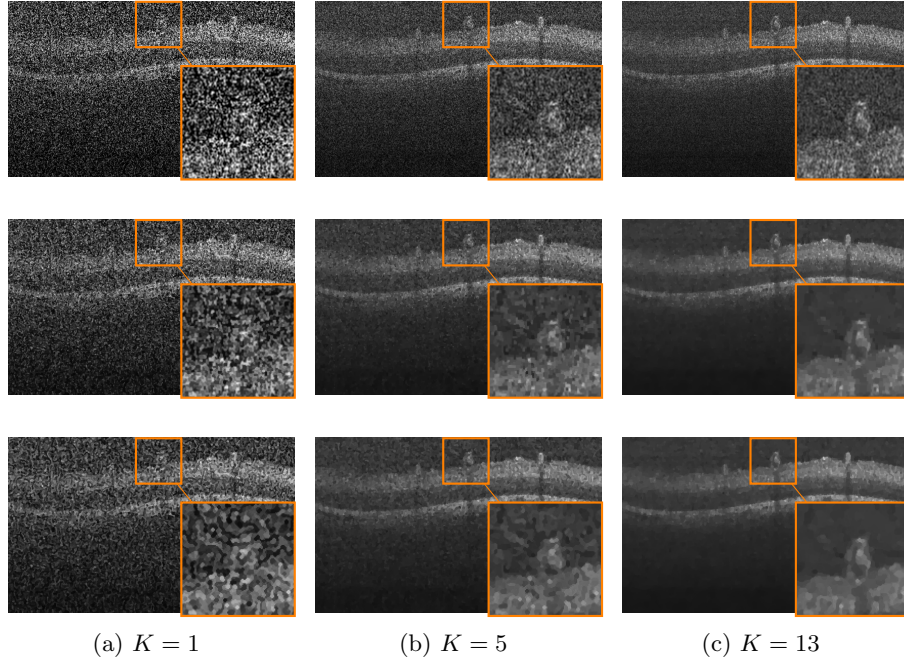(a) $K = 1$ (b) $K = 5$ (c) $K = 13$

Fig. 1: Comparison of simple averaging of consecutive B-scans (top row) to the proposed spatio-temporal denoising with TV regularization only (second row) and TV + QuaSI regularization (third row) for different numbers of input images.
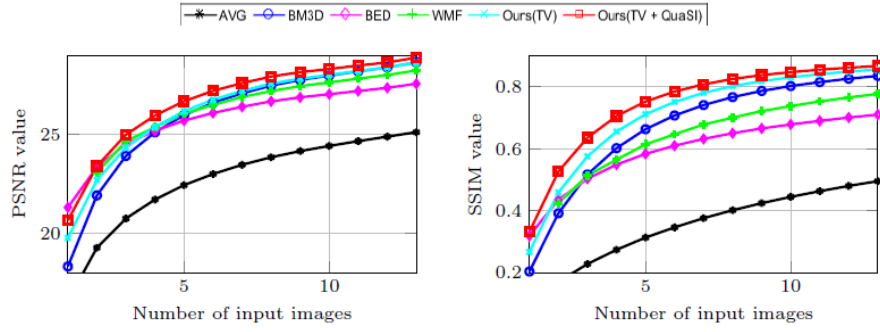

Fig. 2: PSNR and SSIM of different denoising methods averaged over 35 pig eye positions for different numbers of input images.

denoising but suffered from noise in homogeneous areas resulting in lower MSR and CNR measures. BM3D enables a strong noise reduction but suffered from streak artifacts, see the magnified image regions. The proposed method achieved a decent tradeoff between noise reduction and structure preservation.

(a) Noisy image (MSR: 2.68, CNR: 2.47)   (b) BM3D [3] (MSR: 4.61, CNR: 4.85)

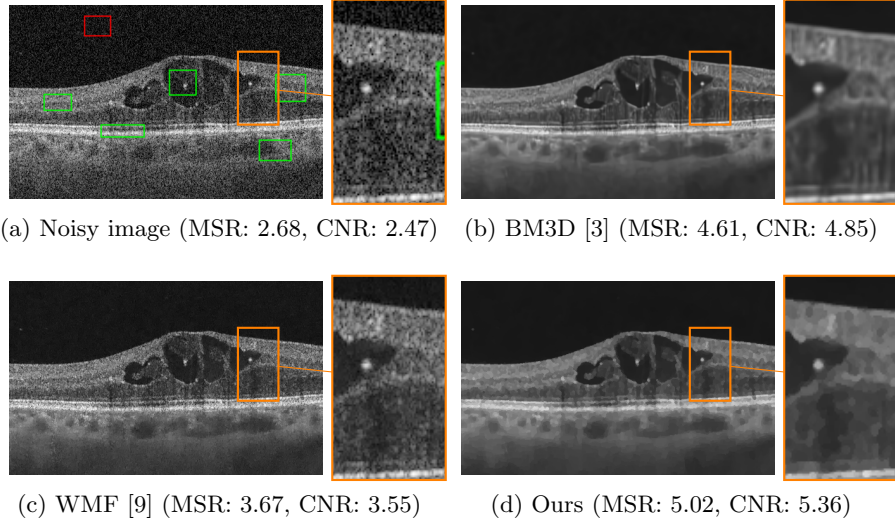(c) WMF [9] (MSR: 3.67, CNR: 3.55)   (d) Ours (MSR: 5.02, CNR: 5.36)

Fig. 3: Denoising on our clinical dataset using $K = 5$ B-scans from a 46 years old male diabetic retinopathy patient. (a) Noisy image with manually selected background (red) and foreground regions (green) to determine MSR and CNR. (b) - (d) BED [15], WMF [9] and our proposed method.
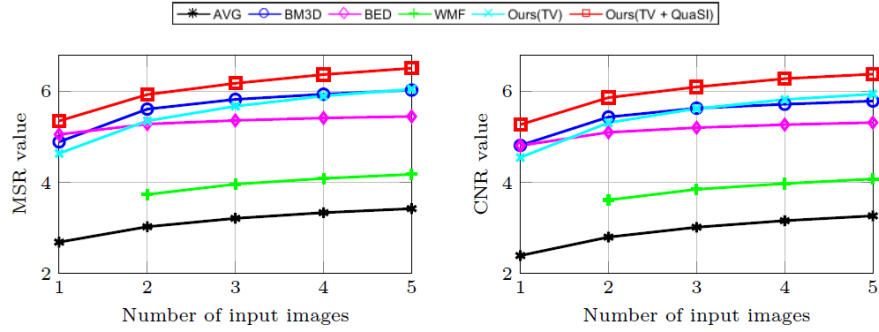

Fig. 4: MSR and CNR measures averaged over 14 subjects to quantify noise reduction on our clinical dataset for different numbers of input images.

4   Conclusion

This paper proposed a spatio-temporal denoising algorithm for OCT data. To effectively reduce speckle noise and preserve morphological structures, we introduced the class of QuaSI priors. We implemented this model via median filter regularization and devolved an ADMM scheme for efficient numerical optimization. Our method can denoise single or multiple registered B-scans. Compared to simple B-scan averaging and state-of-the-art single-image methods, our algorithm

is more effective in reducing speckle noise. In contrast to pure temporal methods, e. g. [9], we can adjust the number of B-scans as a tradeoff between denoising performance and acquisition time. In our future work, we study the impact of our algorithm to common OCT image analysis tasks.

References

1. Cheng, J., Duan, L., Wong, D.W.K., Tao, D., Akiba, M., Liu, J.: Speckle Reduction in Optical Coherence Tomography by Image Registration and Matrix Completion. Proc. MICCAI 2014 8673, 162–169 (2014)
2. Choi, W., Potsaid, B., Jayaraman, V., Baumann, B., Grulkowski, I., Liu, J.J., Lu, C.D., Cable, A.E., Huang, D., Duker, J.S., Fujimoto, J.G.: Phase-sensitive swept-source optical coherence tomography imaging of the human retina with a vertical cavity surface-emitting laser light source. Opt Lett 38(3), 338 (2013)
3. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. IEEE Trans Image Process 16(8), 145–149 (2007)
4. Duan, J., Lu, W., Tench, C., Gottlob, I., Proudlock, F., Samani, N.N., Bai, L.: Denoising optical coherence tomography using second order total generalized variation decomposition. Biomed Signal Process Control 24, 120–127 (2016)
5. Fang, L., Li, S., Nie, Q., Izatt, J.A., Toth, C.A., Farsiu, S.: Sparsity based denoising of spectral domain optical coherence tomography images. Biomed. Opt. Express 3(5), 927–942 (2012)
6. Goldstein, T., Osher, S.: The split bregman method for l1-regularized problems. SIAM J Imaging Sci 2(2), 323–343 (2009)
7. Jinshan Pan, Deqing Sun, Hanspeter Pfister, Ming-Hsuan Yang: Blind image deblurring using dark channel prior. Proc. CVPR 2016 pp. 1628–1636 (2016)
8. Köhler, T., Bock, R., Hornegger, J., Michelson, G.: Computer-Aided Diagnostics and Pattern Recognition: Automated Glaucoma Detection. In: Michelson, G. (ed.) Teleophthalmology in Preventive Medicine, pp. 93–104. Springer (2015)
9. Mayer, M.A., Borsdorf, A., Wagner, M., Hornegger, J., Mardin, C.Y., Tornow, R.P.: Wavelet denoising of multiframe optical coherence tomography data. Biomed Opt Express 3(3), 572 (2012)
10. Ochs, P., Dosovitskiy, A., Brox, T., Pock, T.: On Iteratively Reweighted Algorithms for Nonsmooth Nonconvex Optimization in Computer Vision. SIAM J Imaging Sci 8(1), 331–372 (2015)
11. Ozcan, A., Bilenca, A., Desjardins, A.E., Bouma, B.E., Tearney, G.J.: Speckle reduction in optical coherence tomography images using digital filtering. Journal of the Optical Society of America A 24(7), 1901 (2007)
12. Pircher, M., Gotzinger, E., Leitgeb, R., Fercher, A.F., Hitzenberger, C.K.: Speckle reduction in optical coherence tomography by frequency compounding. J Biomed Opt 8(3), 565 (2003)
13. Romano, Y., Elad, M., Milanfar, P.: The Little Engine that Could: Regularization by Denoising (RED). arXiv preprint arXiv:1611.02862 (2016)
14. Salinas, H., Fernandez, D.: Comparison of PDE-Based Nonlinear Diffusion Approaches for Image Enhancement and Denoising in Optical Coherence Tomography. IEEE Trans Med Imaging 26(6), 761–771 (2007)
15. Wong, A., Mishra, A., Bizheva, K., Clausi, D.a.: General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery. Opt Express 18(8), 8338–8352 (2010)