Wavelet denoising of multiframe optical coherence tomography data

Markus A. Mayer,[1,2] **Anja Borsdorf,**[1] **Martin Wagner,**[3] **Joachim Hornegger,**[1,2] **Christian Y. Mardin,**[4] **and Ralf P. Tornow**[4]

[1] *Pattern Recognition Lab, Martensstrasse 3, 91058 Erlangen, Germany*
[2] *Erlangen Graduate School in Advanced Optical Technologies (SAOT), Paul-Gordan Str. 6, 91052 Erlangen, Germany*
[3] *University of Heidelberg Medical Center, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany*
[4] *Department of Ophthalmology, Schwabachanlage 6, 91054 Erlangen, Germany*
markus.mayer@informatik.uni-erlangen.de

Abstract: We introduce a novel speckle noise reduction algorithm for OCT images. Contrary to present approaches, the algorithm does not rely on simple averaging of multiple image frames or denoising on the final averaged image. Instead it uses wavelet decompositions of the single frames for a local noise and structure estimation. Based on this analysis, the wavelet detail coefficients are weighted, averaged and reconstructed. At a signal-to-noise gain at about 100% we observe only a minor sharpness decrease, as measured by a full-width-half-maximum reduction of 10.5%. While a similar signal-to-noise gain would require averaging of 29 frames, we achieve this result using only 8 frames as input to the algorithm. A possible application of the proposed algorithm is preprocessing in retinal structure segmentation algorithms, to allow a better differentiation between real tissue information and unwanted speckle noise.

OCIS codes: (100.0100) Image processing; (100.2980) Image enhancement; (100.7410) Wavelets(110.4500); Optical coherence tomography.

References and links

1. D. Huang, E. A. Swanson, C. P. Lin, J. Schuman, W. Stinson, W. Chang, M. Hee, T. Flotte, K. Gregory, C. Puliafito, and et. al., "Optical coherence tomography," Science **254**(5035), 1178–1181 (1991).
2. A. F. Fercher, C. K. Hitzenberger, W. Drexler, G. Kamp, and H. Sattmann, "In vivo optical coherence tomography," Amer. J. Ophthalmology. **116**(1), 113–114 (1993).
3. V. Guedes, J. S. Schuman, E. Hertzmark, G. Wollstein, A. Correnti, R. Mancini, D. Lederer, S. Voskanian, L. Velazquez, H. M. Pakter, T. Pedut-Kloizman, J. G. Fujimoto, and C. Mattox, "Optical coherence tomography measurement of macular and nerve fiber layer thickness in normal and glaucomatous human eyes," Ophthalmology **110**(1), 177–189 (2003).
4. Y. Chen, L. N. Vuong, J. Liu, J. Ho, V. J. Srinivasan, I. Gorczynska, A. J. Witkin, J. S. Duker, J. Schuman, and J. G. Fujimoto, "Three-dimensional ultrahigh resolution optical coherence tomography imaging of age-related macular degeneration," Opt. Express **17**(5), 4046–4060 (2009).
5. F. K. Horn, C. Y. Mardin, R. Laemmer, D. Baleanu, A. Juenemann, F. E. Kruse, and R. P. Tornow, "Correlation between local glaucomatous visual field defects and loss of nerve fiber layer thickness measured with scanning laser polarimetry and spectral domain optical coherence tomograpy," Invest. Ophthalmol. Vis. Sci. **50**(5), 1971–1977 (2009).
6. I. Krebs, S. Hagen, W. Brannath, P. Haas, I. Womastek, G. de Salvo, S. Ansari-Shahrezaei, and S. Binder, "Repeatability and reproducibility of retinal thickness measurements by optical coherence tomography in age-related macular degeneration," Ophthalmology **117**(8), 1577–1584 (2010).

7. M. K. Garvin, M. D. Abramoff, R. Kardon, S. R. Russell, X. Wu, and M. Sonka, "Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search," IEEE Transactions on Medical Imaging **27**(10), 1495 – 1505 (2008).
8. D. Baleanu, R. P. Tornow, F. K. Horn, R. Laemmer, C. W. Roessler, A. G. Juenemann, F. E. Kruse, and C. Y. Mardin, "Retinal nerve fiber layer thickness in normals measured by spectral domain OCT," J. Glaucoma **19**(7), 475–482 (2009).
9. S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," Opt. Express **18**(18), 19,413–19,428 (2010).
10. A. Buades, B. Coll, and J. M. Morel, "On image denoising methods," Technical Note, CMLA (Centre de Mathematiques et de Leurs Applications) (2004).
11. J. M. Schmitt and A. Knüttel, "Model of optical coherence tomography of heterogeneous tissue," J. Opt. Soc. Am. A **14**(6), 1231–1242 (1997).
12. J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," Journal of Biomedical Optics **4**(1), 95–105 (1999).
13. B. Karamata, K. Hassler, M. Laubscher, and T. Lasser, "Speckle statistics in optical coherence tomography," J. Opt. Soc. Am. A **22**(4), 593–596 (2005).
14. N. M. Grzywacz, J. de Juan, C. Ferrone, D. Giannini, D. Huang, G. Koch, V. Russo, O. Tan, and C. Bruni, "Statistics of optical coherence tomography data from human retina," IEEE Transactions on Medical Imaging, **29**(6), 1224 –1237 (2010).
15. P. Lee, W. Gao, and X. Zhang, "Speckle properties of the logarithmically transformed signal in optical coherence tomography," J. Opt. Soc. Am. A **28**(4), 517–522 (2011).
16. M. Bashkansky and J. Reintjes, "Statistics and reduction of speckle in optical coherence tomography," Opt. Lett. **25**(8), 545–547 (2000).
17. N. Iftimia, B. E. Bouma, and G. J. Tearney, "Speckle reduction in optical coherence tomography by path length encoded angular compounding," J. Biomed. Opt. **8**(2), 260–263 (2003).
18. L. Ramrath, G. Moreno, H. Mueller, T. Bonin, G. Huettmann, and A. Schweikard, "Towards multi-directional OCT for speckle noise reduction," in *Proceedings of the 11th international conference on Medical Image Computing and Computer-Assisted Intervention - Part I*, MICCAI '08, pp. 815–823 (Springer-Verlag, Berlin, Heidelberg, 2008).
19. M. Hughes, M. Spring, and A. Podoleanu, "Speckle noise reduction in optical coherence tomography of paint layers," Applied Optics **49**(1), 99–107 (2010).
20. A. E. Desjardins, B. J. Vakoc, G. J. Tearney, and B. E. Bouma, "Speckle reduction in OCT using massively-parallel detection and frequency-domain ranging," Opt. Express **14**(11), 4736–4745 (2006).
21. M. Pircher, E. Gotzinger, R. Leitgeb, A. F. Fercher, and C. K. Hitzenberger, "Speckle reduction in optical coherence tomography by frequency compounding." Journal of Biomedical Optics **8**(3), 565–569 (2003).
22. B. Sander, M. Larsen, L. Thrane, J. Hougaard, and T. M. Jrgensen, "Enhanced optical coherence tomography imaging by multiple scan averaging," British Journal of Ophthalmology **89**(2), 207–212 (2005).
23. T. M. Jrgensen, J. Thomadsen, U. Christensen, W. Soliman, and B. Sander, "Enhancing the signal-to-noise ratio in ophthalmic optical coherence tomography by image registration - method and clinical examples," J. Biomed. Opt. **12**(4), 041,208 (2007).
24. E. Götzinger, M. Pircher, B. Baumann, T. Schmoll, H. Sattmann, R. A. Leitgeb, and C. K. Hitzenberger, "Speckle noise reduction in high speed polarization sensitive spectral domain optical coherence tomography," Opt. Express **19**(15), 14,568–14,585 (2011).
25. R. D. Ferguson, D. Hammer, L. A. Paunescu, S. Beaton, and J. S. Schuman, "Tracking optical coherence tomography," Opt. Lett. **29**(18), 2139–2141 (2004).
26. M. E. B. Jadwiga Rogowska, "Image processing techniques for noise removal, enhancement and segmentation of cartilage OCT images." Physics in Medicine and Biology **47**(4), 641–655 (2002).
27. D. L. Marks, T. S. Ralston, and S. A. Boppart, "Speckle reduction by I-divergence regularization in optical coherence tomography," J. Opt. Soc. Am. A **22**(11), 2366–2371 (2005).
28. 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**(8), 8338–8352 (2010).
29. H. M. Salinas and D. C. Fernandez, "Comparison of PDE-based nonlinear diffusion approaches for image enhancement and denoising in optical coherence tomography," IEEE Transactions on Medical Imaging **26**(6), 761–771 (2007).
30. R. Bernardes, C. Maduro, P. Serranho, A. Araújo, S. Barbeiro, and J. Cunha-Vaz, "Improved adaptive complex diffusion despeckling filter," Opt. Express **18**(23), 24,048–24,059 (2010).
31. P. Puvanathasan and K. Bizheva, "Interval type-II fuzzy anisotropic diffusion algorithm for speckle noise reduction in optical coherence tomography images," Opt. Express **17**(2), 733–746 (2009).
32. D. C. Adler, T. H. Ko, and J. G. Fujimoto, "Speckle reduction in optical coherence tomography images by use of a spatially adaptive wavelet filter," Optics Letters **29**(24), 2878–2880 (2004).
33. S. Chitchian, M. A. Fiddy, and N. M. Fried, "Denoising during optical coherence tomography of the prostate

nerves via wavelet shrinkage using dual-tree complex wavelet transform." Journal of Biomedical Optics **14**(1), 014,031 (2009).
34. 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**(2), 1024–1032 (2010).
35. A. Ozcan, A. Bilenca, A. E. Desjardins, B. E. Bouma, and G. J. Tearney, "Speckle reduction in optical coherence tomography images using digital filtering," J. Opt. Soc. Am. A **24**(7), 1901–1910 (2007).
36. W. Guo and F. Huang, "A Local Mutual Information Guided Denoising Technique and Its Application to Self-calibrated Partially Parallel Imaging," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2008*, D. Metaxas, L. Axel, G. Fichtinger, and G. Szkely, eds., vol. 5242 of *Lecture Notes in Computer Science*, pp. 939–947 (Springer Berlin / Heidelberg, 2008).
37. N. Azzabou and N. Paragios, "Spatio-temporal speckle reduction in ultrasound sequences." Medical Image Computing and Computer-Assisted Intervention **11**(1), 951–958 (2008).
38. A. Borsdorf, R. Raupach, and J. Hornegger, "Multiple CT-reconstructions for locally adaptive anisotropic wavelet denoising," International Journal of Computer Assisted Radiology and Surgery **2**(5), 255–264 (2008).
39. A. Borsdorf, R. Raupach, T. Flohr, and J. Hornegger, "Wavelet based noise reduction in CT-Images using correlation analysis," IEEE Transactions on Medical Imaging **27**(12), 1685–1703 (2008).
40. G. Nason and B. Silverman, "The stationary wavelet transform and some statistical applications," Lecture Notes in Statistics **103**, 281–299 (1995).
41. I. Selesnick, R. Baraniuk, and N. Kingsbury, "The dual-tree complex wavelet transform," Signal Processing Magazine, IEEE **22**(6), 123 – 151 (2005).
42. D. Koozekanani, K. Boyer, and C. Roberts, "Retinal thickness measurements from optical coherence tomography using a Markov boundary model," Medical Imaging, IEEE Transactions on **20**(9), 900–916 (2001).
43. M. Mujat, R. Chan, B. Cense, B. Park, C. Joo, T. Akkin, T. Chen, and J. de Boer, "Retinal nerve fiber layer thickness map determined from optical coherence tomography images," Optics Express **13**(23), 9480–9491 (2005).
44. M. Baroni, P. Fortunato, and A. L. Torre, "Towards quantitative analysis of retinal features in optical coherence tomography," Medical Engineering & Physics **29**(4), 432–441 (2007).

1. Introduction

Optical Coherence Tomography (OCT) [1] has become a well-established modality for depth resolved imaging of translucent tissues. Ophthalmology has particularly benefited from the inventions and improvements made to OCT systems. In vivo imaging of the retina [2] is established in daily clinical practise and measurements performed on these images aid the diagnosis of pathologies by physicians [3, 4, 5, 6]. An example for such measurements is determining the thickness of single retinal layers [7, 8, 9].

To speed up the diagnosis and to avoid observer differences, measurements have been automated by software methods. Regardless of the structure to be observed or automatically segmented, e.g. retinal layers, drusens, or the optic nerve head, one common component in modern OCT systems is the speckle noise suppression, since the OCT images are corrupted by unwanted speckle noise. This noise is generated by the interference of multiple scattered photons and complete loss of photons in the tissue. Suppression methods enhance the tissue structure for better visualization and prevention of automated segmentation errors. The development of these methods is challenging, as speckle noise differs in its properties from the zero mean, isotropic Gaussian noise, which is usually assumed to be present in natural scenes taken with a CCD-sensor. Such a noise model is well understood in its behaviour and removal process [10].

The properties of OCT intensities and speckle behaviour were investigated in [11, 12, 13]. Speckle is not pure noise. It also transports image information. Thus, we differentiate between speckle and speckle noise, where the later is the pure unwanted corruption to an ideal OCT signal. As the speckle is image information, its pattern does not change if no physical parameter of the imaging system is altered. The speckle and thus the noise are spatially correlated [14]. Their distribution properties change depending on the intensity scale space in which the images are viewed, on the position in the image, and on the scattering properties of the imaged tissue [15].

State-of-the-art speckle suppresion methods can be roughly categorized in frame averaging

methods and digital denoising algorithms. The first category can be further split into systems that alter parameters of the imaging system in between the multiple recordings or rely on the imaged subject itself to change the speckle pattern, like for example due to movement of the eye when imaging the retina.

Parameters that are changed in OCT systems to decorrelate speckle of multiple recorded frames are the incident angle of the laserbeam [16, 17, 18, 19], the recording angle of the backreflected light [20] or the frequency of the laser beam [21]. These methods can also be utilized when the imaged object is static, as for example in the imaging of paint layers as performed by Hughes et al. [19]. In the case of imaging of the eye, the position of the object of interest changes constantly due to: a) small motions caused by respiration and heart beat; b) movements of the eye itself, like for example saccadic motion. With the development of fast frequency domain OCT systems, multiple B-Scan recordings at the same or nearly the same positions became possible. To compensate for small motions that might have happened, weighted averaging schemes were proposed [22] or the multiple frames were registered by cross correlation [23, 24]. An additional eye tracking hardware can compensate for transversal motion before the scans take place [25].

A multitude of well known digital denoising methods have been adapted for OCT images. Simple average and median filters have been utilized in [26]. Marks et al. [27] formulated an I-divergence regularization approach for despeckling. A new method based on bayesian estimations in large neighborhoods was formulated by Wong et al. [28]. Diffusion filters have proven to be well suited for OCT speckle suppression. Salinas et al. [29] proposed the use of complex diffusion. This approach was further improved by Bernardes et al. [30]. Puvanathasan et al. [31] included edge information and a noise estimate in their formulation of a nonlinear isotropic diffusion denoising. Another prominent technique uses wavelet thresholding or comparable methods. Wavelet methods have the advantage of performing denoising on multiple resolutions, which is desireable when dealing with neighborhood correlated noise. The proposed formulations of wavelet denoising algorithms for OCT images range from spatially adaptive wavelet filters [32] and the use of modern wavelet decompositions, like the dual tree complex wavelet transformation [33], to curvelets transformations [34].

An extensive comparison of standard digital denoising methods has been performed by Ozcan et al. [35]. Among others, wavelet thresholding with shift invariant wavelets yielded the best results. In addition to comparing various filters, Ozcan et al. also adressed the question of multiframe data and investigated if a denoising of the single frames before averaging has advantages over denoising the averaged frame. They conlude that, in terms of quantitative evaluation metrics, no significant difference could be observed. The computation time however is much lower when only the averaged frame is processed. Ozcan et al. processed the single frames indenpendently before averaging.

We propose a new method for performing denoising of multiple frame data, were the single frames are processed before averaging, but information from all frames is utilized. By comparing a single frame to the other frames in the dataset, we are able to differentiate more precisly between tissue structure and unwanted noise. We formulated this idea in the framework of wavelet denoising, as wavelet denoising is widely acknowledged to provide good results for OCT images. The concept, however, is not limited to wavelet denoising. It may easily be incorporated in other denoising techniques.

Multiframe denoising where single frames are denoised by utilizing a multitude of frames are known from other imaging modalities. Guo and Huang [36] proposed a method for magnetic resonance image denoising, where a low resolution reference frame steers the denoising of a second noise image with higher resolution. Azzabou and Paragios [37] denoised ultrasound sequences by adding temporal filtering within the sequence, assuming that the image content

does not vary to a large extend from frame to frame. As the acquisition process in OCT is much faster than in ultrasound and the scanning location may be further stabilized by the use of an eye tracker or software motion-artifact correction, we can assume no motion at all. The algorithm in our work is inspired by Borsdorf et al. [38, 39]. They reconstructed multiple images from distinct selections out of a set of computed tomography (CT) projections. By computing correlations in between the reconstructions, a wavelet denoising approach could be enhanced. Splitting projection data for multiple reconstructions in CT leads to a higher noise level on the single frames. Thus, only a two frame denoising was proposed. With OCT we can in theory aquire a very large number of frames with equivalent noise level. We make use of this property. Our algorithm expects at least 2 input frames, but more input frames are favored.

The remainder of the paper is structured as follows: In section 2 we describe our dataset, the proposed method for denoising multiple frame data, and the evaluation process for gaining quantitative results. Results and a comparison to standard single frame algorithms, namely median filtering and wavelet thresholding, are presented in section 3. Our work is summarized in section 4. Here conclusions are also drawn and suggestions for future work are given.

2. Method

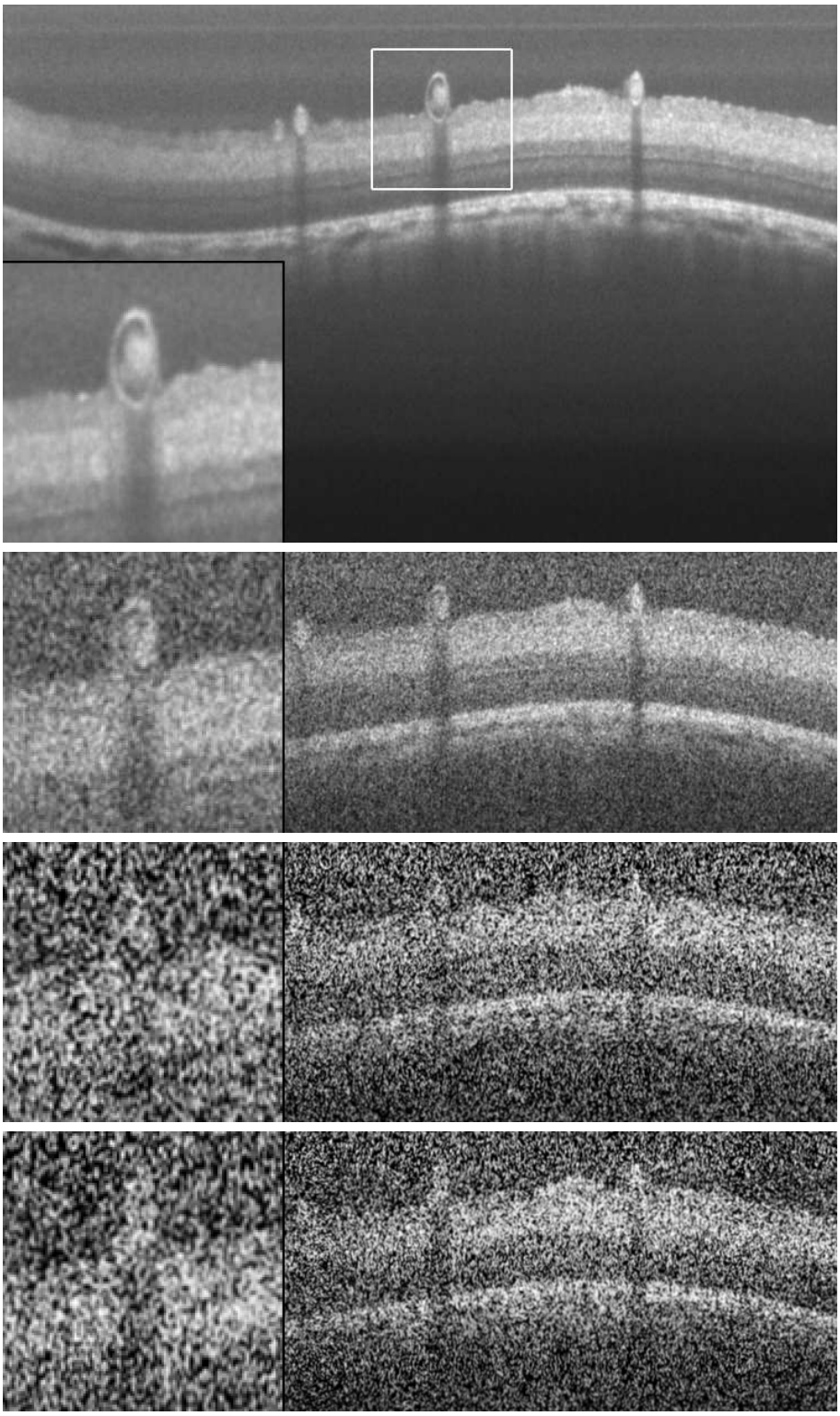
2.1. Data

A dataset for the quantitative evaluation of our method was acquired by scanning a pigs eye ex-vivo with a Spectralis HRA & OCT (Heidelberg Engineering) in high speed mode with 768 A-scans. The pixel spacing of the resulting images is $3.87\mu m$ in the axial direction and $14\mu m$ in the transversal direction. The resolution of the system in the axial direction is $7\mu m$. The eye tracker of the Spectralis was switched off during image acquisition.

The eye was placed in front of the OCT device. It was rotated around the center of its lens. Sets of 13 frames each were recorded at 35 eye positions, corresponding to a complete $0.384mm$ shift in the transversal direction. Speckle noise can be assumed to be uncorrelated in between the scans from the varying positions. The opacity of the eye was increased as the eye lost humidity during transportation. Thus, the image quality was decreased by a lower signal-to-noise ratio. This effect, however, is not unwanted in our evaluation. Our goal is to deal with and improve images with very low signal-to-noise levels as are often observed in daily clinical practise, particularly in the scans of elderly persons with cataract, glaucoma or age-related macula degeneration.

A gold standard image for the evaluation that should contain as little speckle noise as possible was created by averaging all 455 recorded frames. Therefore, in a first step the 35 image sets corresponding to the acquisition at the 35 fixed positions were averaged. Although a large amount of the speckle noise within these sets is correlated, an improvement of the signal-to-noise ratio was observed, as the uncorrelated part of the noise is reduced. Next, the 35 averaged images were registered. First, a manual registration was performed. Afterwards, these manual registrations were automatically optimized by minimizing the sum of squared distances (SSD) between the averaged images. Only rigid transformations, that is translation and rotation, were considered. A powell optimizer was utilized for the optimization. The rotation and translation parameters are applied to all images before further processing. Figure 1 shows examples of: the gold standard, an average of 8 random frames from the dataset, as well as examples of two single frames.

In addition to the pigs eye, the fundus of a normal human subject was recorded to demonstrate the applicability of our algorithm to real word situations. In this case the eye motion was compensated by the built-in eye tracker of the Spectralis. Due to eye motion during the scanning process the single frames of this human eye fundus data set are assumed to have uncorrelated speckle patterns.



(a)

(b)

(c)

(d)

Fig. 1. Example images from the pig eye dataset. (a) Gold standard image, created by averaging 455 frames. An exemplary region is magnified, indicated by the white rectangle. To allow a better visual comparison, this region will also be magnified on all subsequent images shown in this work. (b) Average of 8 randomly selected frames. For display reasons, this image and the remaining images in this work are cropped. (c, d) Single frames from the averaged dataset in (b).

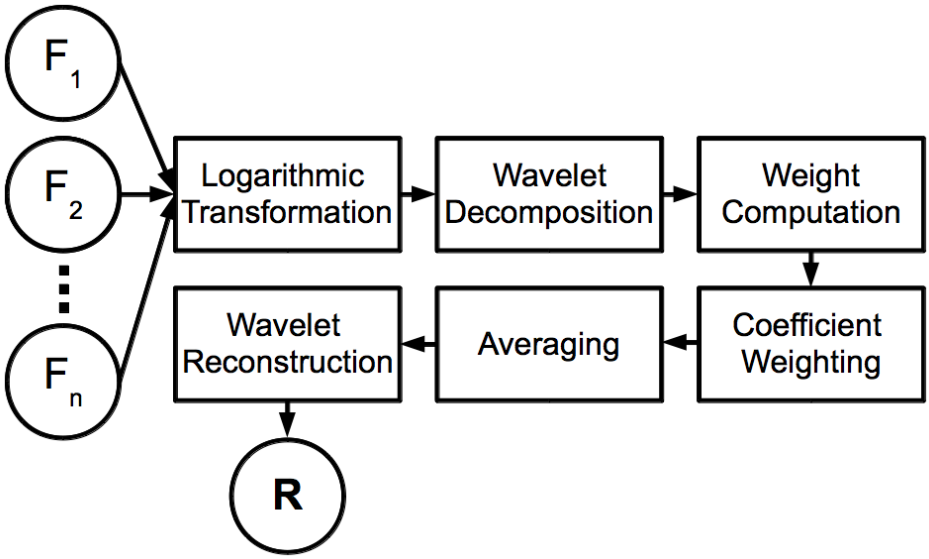


Fig. 2. Flow diagram of the proposed multiframe denoising algorithm. The single images F_i are scaled with a logarithmic transformation and wavelet decomposed. Weights are estimated from the wavelet coefficients and applied to the detail coefficients. Averaging in the wavelet domain and a wavelet reconstruction yields the final image R.

2.2. Algorithm

A complete overview of the algorithm is shown in figure 2. The processing steps are logarithmic scaling, wavelet decomposition, weight estimation, weight application, averaging and wavelet reconstruction. The steps will be explained in detail below. The data input to our algorithm are images F_i, with i ranging from 1 to N. N is the number of images. The denominations "image" and "frame" are treated as equivalent in this work, where the latter better expresses the idea of multiple captures with the same content.

Logarithmic transformation: A logarithmic transformation is applied to the intensities of the input frames. This common practise allows the assumption of a near additive noise model in the logarithmic scale space. An ideal image S is corrupted with noise N_i. N_i differs and is assumed to be uncorrelated with other single frame acquisitions.

$$F_i = S + N_i \tag{1}$$

Furthermore, the standard deviation of the noise $\sigma_i(x)$ at a position x is assumed to be approximately the same for each image as the tissue properties do not change during the acquisition process:

$$\sigma_i(x) \approx \sigma_j(x) \tag{2}$$

where j is also a frame number and $i \neq j$.

Wavelet decomposition: The single frames are decomposed by a wavelet transformation with a maximum decomposition level L. This decomposition yields approximation coefficients A_i^l and detail coefficients $W_{i,D}^l$, where l is the decomposition level and D the direction (horizontal, vertical, diagonal) of the detail coefficients. We use two different wavelet transformations

and compare them to each other. The first is the discrete stationary wavelet transformation [40] with Haar wavelets (DSWT). No downsampling is applied in between the wavelet decomposition levels. The second wavelet transformation is the dual tree complex wavelet transformation (DTCWT) [41]. In both wavelet decompositions, we store the approximation coefficients for each level l, as they are used later on in the algorithm.

Coefficient weighting, averaging, and wavelet reconstruction: The denoising of the images is performed by weighting the detail coefficients at position x with a weight $G^l_{i,D}$

$$\tilde{W}^l_{i,D}(x) = G^l_{i,D}(x) \cdot W^l_{i,D}(x) \tag{3}$$

where $\tilde{W}^l_D$ are the weighted detail coefficients. After weighting, the detail and approximation coefficients are averaged:

$$W^l_D(x) = \frac{1}{N} \sum_{i=1}^{N} \tilde{W}^l_{i,D}(x) \tag{4}$$

$$A^{l_{max}}(x) = \frac{1}{N} \sum_{i=1}^{N} A_i^{l_{max}}(x) \tag{5}$$

The coarsest wavelet decomposition level is denoted by l_{max}. The inverse wavelet transformation of the averaged coefficients yields the final result. The remainder of this subsection presents the proposed weights $G^l_{i,D}$, which exploit the existence of multiple frame data.

Weight computation: Weights for the standard wavelet soft thresholding can be formulated as:

$$G^l_D(x) = \begin{cases} \frac{W^l_D(x) - \mathrm{sgn}(W^l_D(x))\tau}{W^l_D(x)}, & \left|W^l_D(x)\right| > \tau \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

The omission of the frame number i indicates the application of the weights on the averaged data. The thresholding parameter is τ. We compare our method to standard wavelet soft and hard thresholding on the averaged frames. We also compare it to median filtering as a representative of non-wavelet based denoising methods. The latter was chosen as it has been succesfully applied in OCT retinal layer segmentation preprocessing [42, 43, 44]. To allow for a fair comparison, we used a squared window for the median filter, as the wavelet based methods do not prefer a certain direction in advance. It must, however, be noted that in real applications it may be advantageous to choose rectangular shaped windows.

Two different weights are proposed: A significance and a correlation weight. A combination of these two weights is also proposed. The significance weight provides a local noise estimation calculated on the detail coefficients. Contrary to the algorithm proposed by Borsdorf et al. [39] for CT denoising, by having an expected number of input frames larger than 2 we can perform our weight computation on single detail coefficients. The mean squared distance $\sigma_{S,i,D}$ of the detail coefficients of one image to the others is computed at each level l and position x:

$$\sigma^l_{S,i,D}{}^2(x) = \frac{1}{N-1} \sum_{j=1 \wedge j \neq i}^{N} \left(W^l_{i,D}(x) - W^l_{j,D}(x) \right)^2 \tag{7}$$

In the case of the DTCWT, $\sigma_{S,i,D}$ is calculated on the absolute values of the coefficients. This measurement is motivated by the assumption that if a coefficient at a certain position differs to a large extend from the coefficients at the same position on the other frames it is most likely

corrupted by noise. The significance weights G_{sig} which lower the influence of such coefficients on the final result are then calculated by

$$G^l_{sig,i,D}(x) = \begin{cases} 1, & \left|W^l_{i,D}(x)\right| \geq k\sigma^l_{S,i,D}(x) \\ \theta_i(W^l_{1,D}(x),...,W^l_{m,D}(x)), & otherwise \end{cases} \tag{8}$$

where the parameter k controls the amount of noise reduction. As θ we choose

$$\theta_i(w_1,...,w_m) = \frac{1}{m-1}\sum_{j=0\wedge j\neq i}^{m-1}\left|1-\frac{w_i}{w_j}\right| \tag{9}$$

θ_i is scaled to the interval $[0;1]$ after calculation for a single coefficient.

The correlation weight G_{corr} is calculated on the approximation coefficients and provides information on whether a structure is present in the current frame and position or not. The computation we propose is derived from the work of Borsdorf et al. [39], but modified and adapted to the larger input data sets OCT offers. It is motivated by the fact that if edge information is present in a position of a single frame the correlation to the other images calculated in a small neighborhood around this position is higher than in homogeneous regions. If noise has degraded the edge information on one single frame, the correlation will also be lower. So for each approximation coefficient the median of the correlation to each other image within a small neighborhood is calculated:

$$G^l_{corr,i}(x) = \mathrm{med}_{i\neq j}\left(\frac{1}{2}\,\mathrm{Corr}(V^l_i(x),V^l_j(x))+1\right)^p \tag{10}$$

where V^l_i is the vector of all approximation coefficients in a neighborhood (5×5 pixels) around the position x in decomposition layer l of frame i. In the case of the DTCWT, again the absolute values of the coefficients are used for the calculation. *Corr* is the Pearsons correlation coefficient. p is the parameter that controls the amount of noise reduction. The weight $G_{corr,i}$ is applied to the detail coefficients of all 3 directions.

The significance and the correlation weight can be combined. One solution for the combination of the weights is modifying the parameter p in the calculation of the correlation weights with the significance weights:

$$G^l_{comb,i}(x) = G^l_{corr,i}(x) \quad with \quad p = \hat{p}\cdot(1-G^l_{sig,i,D}(x))^2+1 \tag{11}$$

where $\hat{p}$ is a fixed parameter.

2.3. Evaluation

The goal of developing denoising algorithms in general is to reduce the amount of noise without changing the image information. To quantify our results, we measure both the reduction of noise with the signal-to-noise ratio gain (SNR_{gain}) and the integrity of the image information by the integrity of the edges. As we want to avoid edge bluring, the full-with-half-maximum reduction (FWHM_{red}) at certain edges is calculated.

Noise reduction: In image processing, two common definitions for the signal-to-noise ratio (SNR) can be found:

$$\mathrm{SNR}_1 = \frac{\sigma(S)}{\sigma(N)}, \quad or \quad \mathrm{SNR}_2 = \frac{\mu(S)}{\sigma(N)} \tag{12}$$

where $\sigma(S)$ and $\sigma(N)$ denote the standard deviations of the ideal signal and the noise respectively. $\mu(S)$ is the mean value of S. In our evaluation, we consider the improvement of the SNR

of the filtered image compared to a simple averaging. Both of the definitions in equation 12 lead to the same formula:

$$\mathrm{SNR}_{gain} = \frac{\mathrm{SNR}_1(F_f)}{\mathrm{SNR}_1(F_a)} - 1 = \frac{\mathrm{SNR}_2(F_f)}{\mathrm{SNR}_2(F_a)} - 1 = \frac{\sigma(N_a)}{\sigma(N_f)} - 1 \tag{13}$$

where F_f and F_a are the filtered and the averaged image respectively and N_f and N_a are the noise of these images. The real, unknown noise signal is estimated by the difference of each image to the gold-standard image F_g.

$$N_a \approx F_a - F_g, \qquad N_f \approx F_f - F_g \tag{14}$$

The SNR-gain is measured at different regions of interest (ROI) and averaged. The 6 ROIs with approximately to homogeneous intensities are shown as red squares in figure 3.

Edge integrity: The full-width-half-maximum (FWHM) is a measurement for the sharpness of the edge. We examine horizontal edges within a ROI, where the columns are assumed to be perpendicular to the edge, i.e. the values of the columns represent the edge profile. The columns within the ROI are registered by minimizing the mean squared distance to the horizontal average of all columns. The values in the direction of the edge are then summed up to reduce the influence of noise. A logistic sigmoid function $\Gamma(x)$ is fitted to the resulting values with

$$\Gamma(x) = q_1 + \frac{q_2}{1 + \exp\left(-\frac{x+q_3}{q_4}\right)} \tag{15}$$

where q_1 to q_4 are the parameters that are optimized by a nonlinear regression using the Matlab (Mathworks, Inc.) *nlinfit* method. The resulting function is called the edge response function. Its slope and therefore its sharpness is described by its derivative $\Gamma'(x)$ which is given by

$$\Gamma'(x) = \frac{q_2}{q_4} \frac{\exp\left(-\frac{x+q_3}{q_4}\right)}{\left(1 + \exp\left(-\frac{x+q_3}{q_4}\right)\right)^2} \tag{16}$$

The FWHM measures the width of $\Gamma'(x)$ at the half of the maximum ($\Gamma'(-q_3)/2$).

$$\mathrm{FWHM} = \left| q_4 \ln\left(\frac{-2d + 1 - \sqrt{1-4d}}{-2d + 1 + \sqrt{1-4d}}\right)\right|, \quad \text{with} \quad d = \frac{\Gamma'(-q_3)q_4}{q_2} \tag{17}$$

The smaller the FWHM-value the sharper is the edge. To measure the blurring of the edge we calculate the ratio FWHM_{red} between the FWHM in the filtered image and the averaged image generated with the same frames.

$$\mathrm{FWHM}_{red} = \frac{\mathrm{FWHM}(F_f)}{\mathrm{FWHM}(F_a)} - 1 \tag{18}$$

The FWHM_{red} is measured at edges with different contrasts. The 3 edge ROIs are marked in figure 3 as blue rectangles. The values for the different edges are averaged to get one sharpness reduction value for each image.

All quantitative numbers were computed using randomly generated sets of images. At most one frame was randomly selected from each of the possible 35 positions (13 frames per position) of the pig eye data set. To compensate for the random choice, 10 random sets were used for each measurement and the results were averaged. As the parameter choice allows setting the

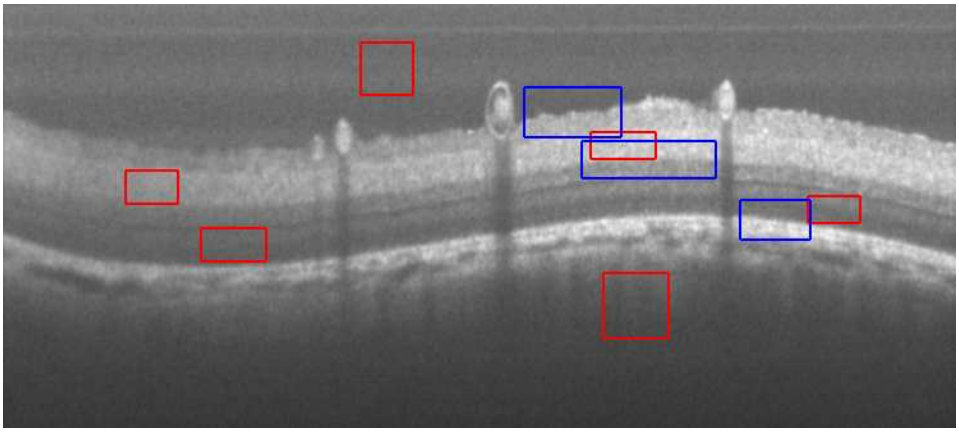


Fig. 3. Gold standard image with the regions of interest marked for the two evaluation metrics. Red rectangle regions are used in the SNR_{gain} measurement. Blue rectangle regions are used in the $FWHM_{red}$ measurements.

strength of the denoising behaviour, we choose comparable results with a SNR_{gain} of around 100% as exemplary quantitative numbers and visual examples.

The implementation of the algorithm was done in Matlab (Mathworks, Inc., Natick, USA). On a Macbook Pro with a 2.66 Intel Core Duo processor and 4GB of memory, a 496 times 394 image set with 8 frames takes 42s to denoise, using the combination of significance and correlation weight. The algorithm was not optimized for speed. However, algorithm speed is not within the scope of this work.

3. Results and Discussion

To assess the performance of our proposed algorithm, we evaluate it quantitatively and qualitatively. We first perform a quantitative evaluation with the metrics presented in section 2.3. The algorithm behaviour with varying parameters is discussed. Afterwards the performance is quantitatively compared to state-of-the-art methods, namely simple frame averaging, median filtering, and wavelet soft and hard thresholding. We conclude the evaluation with a visual inspection of the results on both the pig eye data and the human eye data.

3.1. Parameter Behaviour

The behaviour of the algorithm can be adjusted by the following parameters: The choice of the wavelet, the weight computation method, the parameter k in the significance weight, the parameter p in the correlation weight, and the number of wavelet levels. The number of input frames also has to be taken into account. Due to this large number of parameters, an evaluation of all possible combinations is not feasible. We restrict ourself to parameter combinations that have proven to provide good results in preliminary tests and vary only few parameters to provide an understanding of the algorithm behaviour.

First, we compare the 3 different weight computation methods. The results are computed using 8 frames and 5 wavelet decomposition levels. The quantitative results are computed for both proposed wavelets and are shown in figure 4. The FWHM_{red} is plotted against the SNR_{gain}. An ideal filter would completely leave the edge sharpness intact, which corresponds to a FWHM_{red} of 0%, while providing high SNR_{gain} values.

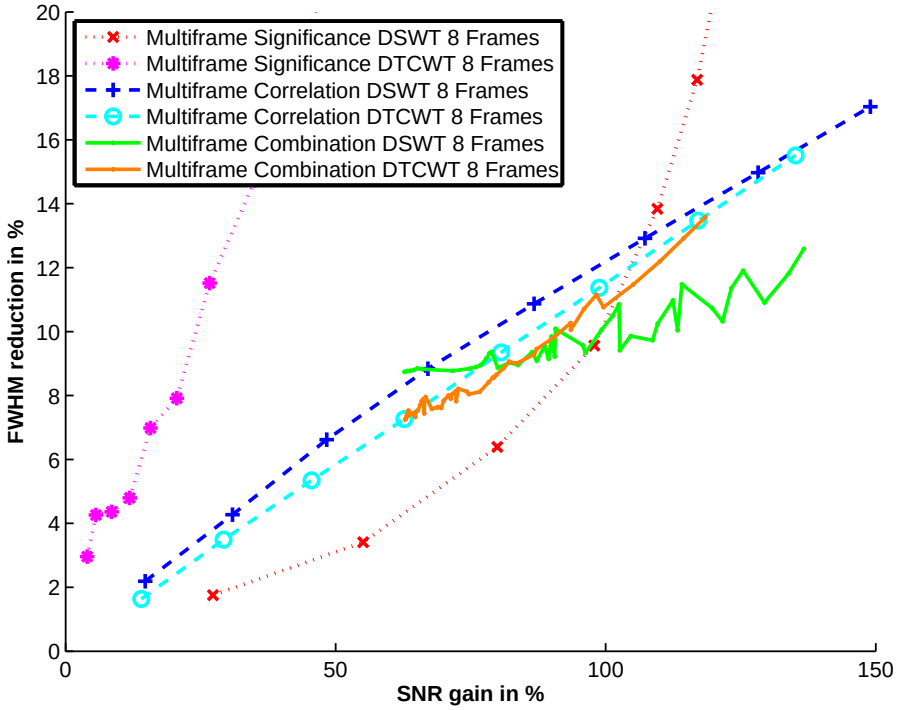


Fig. 4. Sharpness reduction as measured by the full-width-half-maximum reduction (FWHM_{red}) plotted against noise reduction as measured by the signal-to-noise gain (SNR_{gain}) for **8 frames** and 5 wavelet levels using the wavelet multiframe noise reduction method. Results are shown for: Significance weight (parameter k varied) and discrete stationary wavelet transform with Haar wavelets (DSWT); significance weight (parameter k varied) and dual tree complex wavelet transformation (DTCWT); correlation weight (parameter p varied) and DSWT; correlation weight (parameter p varied) and DTCWT; combination of significance and correlation weight (parameter k and p^* varied) and DSWT; combination of significance and correlation weight (parameter k and p^* varied) and DTCWT.

We vary the parameter k in 0.1 steps from 1 to 2.4 in the computation of the significance weight results. Using the DSWT, we achieve 109% SNR_{gain} with a sharpness loss of 13.8%. The DTCWT performed worse. Here a comparable SNR_{gain} of 106.0% lead to a sharpness loss of 40.7%. This discrepancy can be explained by the length of the support of the two wavelet transformations. The significance weight is computed for each detail coefficient position independently without any influence of the neighborhood. However, the value of the detail coefficient is influenced by a larger neighborhood in the spatial domain when the DTCWT is applied compared to when Haar wavelets are used. Thus, noisy coefficients near edges have a higher probabilty of influencing the appearance of the edge. For low SNR_{gain} values, the significance weight with the DSWT delivered the best quantitative results.

In the case of the correlation weight, the support of the wavelets plays only a minor role, as this weight is by itself computed in a window on the approximation coefficients. Thus, the correlation weight, where the parameter p is varied from 0.25 to 2 in 0.25 steps, shows comparable results for both wavelets. With increasing SNR_{gain}, the FWHM_{red} rises almost lineary, yielding an 11.3% FWHM_{red} at 98.8% SNR_{gain} for the DTCWT.

If a high SNR_{gain} is desired, the combination of the significance and the correlation weight provides the best quantitative evaluation results. All combinations of k ranging from 0.5 to 1.5 in 0.1 steps with p ranging from 0.25 to 2.5 in 0.25 steps were tested. The $FWHM_{red}$ is lower than the correlation weight for all SNR_{gain} values, with the DSWT delivering better edge preservation at high noise reduction rates. This effect is most likely also related to the larger support of the DTCWT.

One important question is the behaviour of the algorithm when different numbers of input frames are used. We, therefore, also computed results on 4 frames. The same parameters were used as with 8 frames. Figure 5 shows the amount of noise reduction plotted against edge degradation. When comparing the plots to the 8-frame results in figure 4, two differences can be observed. First, the significance weight performs best only on a decreased range of small SNR_{gain} values. Second, the DSWT outperforms the DTCWT more clearly when both weights are combined. Note, that the SNR_{gain} and $FWHM_{red}$ values in figure 4 and figure 5 do not represent abolute numbers (see section 2.3), but are computed relative to the average of the respective number of frames, so only the *behaviour* of the curves allow an appropriate comparison.

The number of wavelet levels has only a minor influence on the result. With fewer than 5 decomposition levels, the $FWHM_{red}$ slightly increases. We therefore used 5 decomposition levels troughout the results computation. We chose two values shown in table 1 to exemplify this behaviour. Furthermore, table 1 presents the quantitative evaluation results for a SNR_{gain} of roughly 100%.

To sum up, we propose the usage of the significane weight for small amounts of noise reduction, for example for visualization purposes. For heavy noise suppression, as is required, for example, in segmentation preprocessing, a combination of significance and correlation weight is feasible. Suprisingly, although the DTCWT is the more advanced wavelet and delivers good results with single frame denoising, as we will see in the next section, the DSWT with Haar wavelets is better suited to the multiframe method due to its shorter support.

3.2. Comparison to state-of-the-art

A comparison of the proposed multiframe method to averaging, median filtering, and standard wavelet thresholding methods is shown in figure 6. We chose soft thresholding for the DSWT and hard thresholding for the DTCWT, as these outperformed the respective competitor in the preliminary tests. The $FWHM_{red}$ of the frame averaging method was computed in relation to the average of 8 frames (see section 2.3). We averaged up to 40 frames. Exemplary numbers of averages are marked in figure 6.

The traditional DSWT clearly exhibits the worst performance, followed by median filtering with window sizes of 3×3 and 5×5. The quantitative evaluation results of the DTCWT used on a 8-frames average are better than frame averaging of up to 18 frames. If high computational speed is required in preprocessing, this method should be considered. The multiframe method, however, delivers better edge preservation in any case. The amount of edge sharpness loss compared to averaging is small. Averaging 29 frames leads to an SNR_{gain} of 100.0% and a $FWHM_{red}$ of 8.3%. A comparable SNR_{gain} achieved with only 8 frames and the multiframe method yields a 10.5% $FWHM_{red}$.

3.3. Visual inspection

It is known that good quantitative denoising performance does not necesssary lead to a visually pleasing image. Especially when using wavelets, artifacts of the denoising algorithm may disturb the image appearance. We therefore chose example image results from the pig eye dataset, of which quantitative evaluation results were given in table 1, for a visual inspection. The images are shown in figure 7. The parameters of the respective algorithms were adjusted

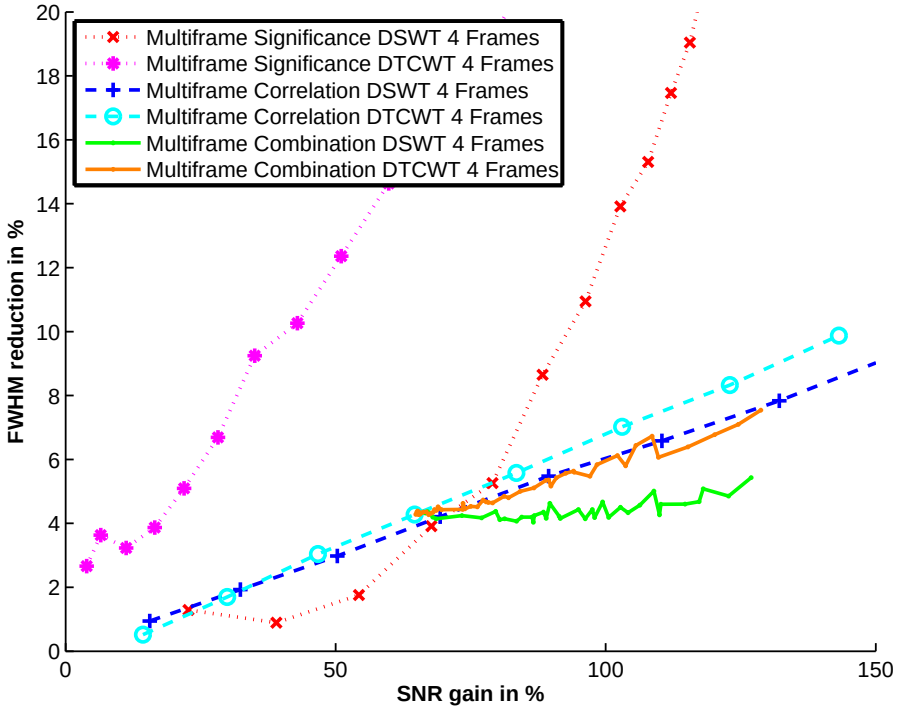


Fig. 5. Sharpness reduction as measured by the full-width-half-maximum reduction ($FWHM_{red}$) plotted against noise reduction as measured by the signal-to-noise gain (SNR_{gain}) for **4 frames** and 5 wavelet levels using the wavelet multiframe noise reduction method. Results are shown for: Significance weight (parameter k varied) and discrete stationary wavelet transform with Haar wavelets (DSWT); significance weight (parameter k varied) and dual tree complex wavelet transformation (DTCWT); correlation weight (parameter p varied) and DSWT; correlation weight (parameter p varied) and DTCWT; combination of significance and correlation weight (parameter k and p^* varied) and DSWT; combination of significance and correlation weight (parameter k and p^* varied) and DTCWT.

so that the SNR_{gain} of all denoised images was roughly about 100%. At this high denoising level, median filtering does remove details and weak edges are blurred (see figure 7b). The standard wavelet hard thresholding with the DTCWT, as shown in figure 7c, is even more unpleasant to the observer, as the image already contains strong denoising artifacts. Together with the quantitative results shown in section 3.2 we can conclude that the single frame denoising with wavelet thresholding is only feasible for low amounts of noise reduction. The use of the multiframe method and the significance weight alone also reaches the limit of its applicability at an SNR_{gain} of 100%, as derived from the quantitative results. Very small wavelet artifacts are visible. Compared to the average image in figure 7a, the speckle reduction, however, is clearly observeable, as the large speckle grains in the average image became much finer and evenly distributed. This behaviour remains intact in the wavelet multiframe results with the combination weight shown in figure 7e. The noise reduction is very strong, while the blood vessel wall shown in the magnification is still perfectly intact, better than in result of the significance weight. In addition, no wavelet artifacts can be observed, even at this high level of noise reduction.

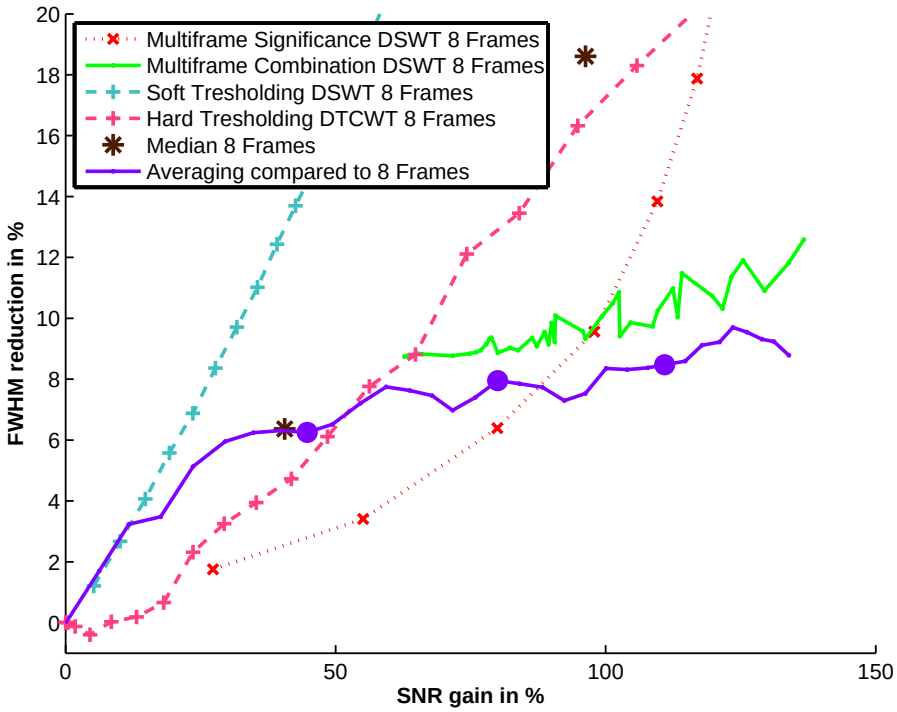


Fig. 6. Sharpness reduction as measured by the full-width-half-maximum reduction (FWHM_{red}) plotted against noise reduction as measured by the signal-to-noise gain (SNR_{gain}) for **8 frames** and 5 wavelet levels. Results are shown for: Wavelet soft thresholding with Haar wavelets and the discrete stationary wavelet transform (DSWT) (threshold varied); wavelet hard thresholding with dual tree complex wavelet transformation (DTCWT) (threshold varied); wavelet multiframe denoising with significance weight (parameter k varied); wavelet multiframe denoising with a combination of significance and correlation weight (parameter k and p^* varied); median filtering (window size of 3×3 and 5×5 pixels). Additionaly, the results averaging of up to 40 frames are shown were an average of 8 frames holds as a reference. The results of averaging 16, 24, and 32 frames are marked with dots.

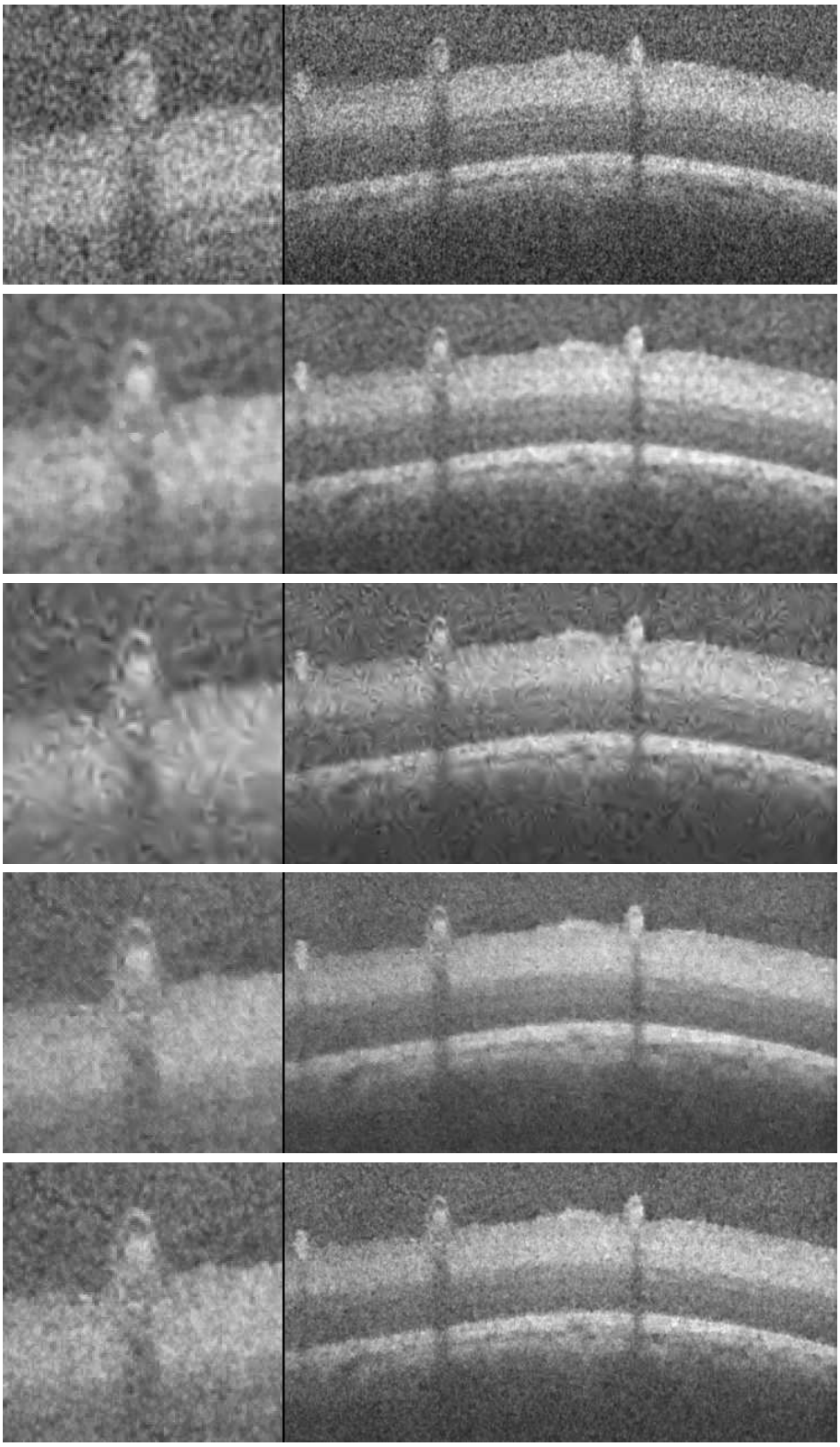


Fig. 7. Result examples from the pig eye dataset, generated from the same 8 randomly selected frames: (a) Average. (b - e) Denoising results. Parameters for the methods were adjusted so that the SNR_{gain} was roughly 100%. (b) Median filtering with a window size of 5×5 on the average of 8 frames. (c) Wavelet hard thresholding with dual tree wavelet on the average of the 8 frames. (d) Wavelet multiframe denoising using the significance weight with Haar wavelets. (e) Wavelets multiframe denoising using the combined significance and correlation weight with Haar wavelets.

Table 1. Quantitative evaluation results of the different denoising methods. The parameters were adjusted so that the signal-to-noise-ratio gain (SNR_{gain}) was about 100%, with the exception of the wavelet soft thresholding using the Haar wavelet, where a SNR_{gain} of 100 can not be achieved due to artifact generation.

Method	**Wavelet**	**Levels**	**# Frames**	$\mathbf{SNR}_{gain}[\%]$	$\mathbf{FWHM}_{red}[\%]$
Averaging	-	-	8	0	0
Averaging	-	-	29	100	8.3
Median Filtering	-	-	8	96.3	18.6
Soft Thresholding	Haar	5	8	75.1	28.7
Hard Thresholding	Dual Tree	5	8	105.8	18.3
Median Filtering	-	-	4	103.6	16.7
Soft Thresholding	Haar	5	4	75.3	22.3
Hard Thresholding	Dual Tree	5	4	101.0	12.2
Significance Weight	Haar	5	8	109.6	13.8
Significance Weight	Dual Tree	5	8	106.0	40.7
Correlation Weight	Haar	5	8	107.2	12.9
Correlation Weight	Dual Tree	5	8	98.8	11.3
Weight Combination	Haar	5	8	101.2	10.5
Weight Combination	Dual Tree	5	8	99.6	10.7
Weight Combination	Haar	3	8	102.4	13.71
Weight Combination	Haar	4	8	99.9	11.9
Weight Combination	Haar	5	4	99.4	4.6
Weight Combination	Dual Tree	5	4	102.2	6.1

Applied on in-vivo data, the observations stay the same, as shown in figure 8. The parameters for the algorithms were the same as the ones in the quantitative evaluation, where a SNR_{gain} of roughly 100% was achieved for 4 frames (see table 1). Again, the median filter removes details and the single frame wavelet thresholding produces a large amount of artifacts. The multiframe method significantly lowers the noise in the background and in homogeneous regions inside the retinal layers, while preserving all edges and details of the retinal structure.

4. Summary and Conclusion

We presented a denoising method, that uses the single captures as input instead of the average of multiple frames, as is common in OCT processing today. The single frames are wavelet transformed, and on the transformed data weights are computed. We proposed two different weights. A significance weight, that determines if noise is locally present and a correlation weight, that determines if structure is present within a local neighborhood. A combination of these weights is also possible. The wavelet detail coefficients are scaled with the weights, averaged and transformed back.

A quantitative evaluation showes that the proposed method is capable of suppressing noise better than median filtering or single frame wavelet denoising on the averaged data. The amount of noise reduction can be adjusted with parameters. A signal-to-noise gain of 101.2% leads to a sharpness reduction measured by full-width-half-maximum reduction of 10.5%. This is only slightly larger than a full-width-half-maximum reduction of 8.3% when averaging 29 frames instead of 8, where a comparable signal-to-noise gain of 100% is achieved. A visual inspection of denoised images from a pig eye ex-vivo and in-vivo human retina scans shows that the

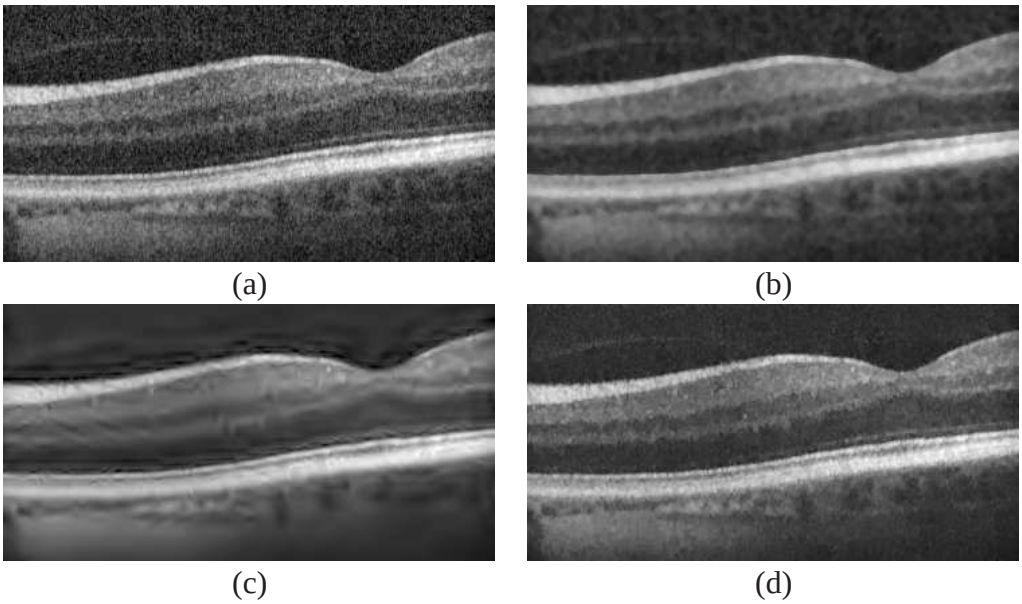


Fig. 8. Result examples from the human eye fundus dataset: (a) Average of 4 frames. (b - d) Denoising results. Parameters for the methods were adjusted such that the SNR_{gain} was roughly 100%. (b) Median filtering on the average of 4 frames. (c) Wavelet hard thresholding with dual tree wavelet on the average of 4 frames. (d) Wavelets multiframe denoising using the combined significance and correlation weight with Haar wavelets.

method reduces noise effectively without degrading structure or generating denoising artifacts. At the shown noise reduction level, this could not be achieved with single frame denoising methods.

The proposed algorithm is applicable in OCT imaging in a clinical setting where the acquisition of a large number of frames for averaging is not possible, for example in elderly patients or patients with severe eye diseases, but where a high quality result is desired. The main application area from our point of view is image preprocessing, for example in the segmentation of retinal layers, as the noise on the images is effectively suppressed without degrading structural information severly. Other applications or automated computations that use multiframe OCT data as input may also benefit from the ideas presented in this work.

Acknowledgement

The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) in the framework of the German excellence initiative.

We thank Ivan W. Selesnick for sharing the Matlab code for the dual tree complex wavelet transformation on his website. In addition, we thank Shahab Chitchian for the fruitful discussions on denoising OCT images.