MULTI-DIMENSIONAL DENOISING OF REAL-TIME OCT IMAGING DATA

Tyler S. Ralston[⋆]*, Ian Atkinson*[♮]*, Farzad Kamalabadi*[♮]*, and Stephen A. Boppart*[⋆]

University of Illinois at Urbana-Champaign
Department of Electrical and Computer Engineering
[⋆]Beckman Institute for Advanced Science and Technology, [♮]Coordinated Science Laboratory

ABSTRACT

We present a novel scheme for blind suppression of noise from a sequence of optical coherence tomography (OCT) images, such as those collected on a real-time OCT imaging system. In contrast to virtually all existing approaches to OCT denoising, our technique is specifically aimed at collections of images and is able to exploit the correlations among those images. The proposed method approximates the optimal linear denoising operator for log-transformed data using a 2-D discrete wavelet transform (DWT) to decorrelate in space and the discrete Fourier transform (DFT), or an estimated transform, to decorrelate in time. Decorrelated coefficients are then denoised and converted back to the image domain to produce denoised OCT images. Real-time OCT data processed with this technique shows significant reduction in noise.

1. INTRODUCTION

Optical coherence tomography is a near-infrared imaging and microscopy technique capable of micrometer-scale resolutions in biological specimens [1]. The imaging principles behind OCT are similar to those of ultrasound imaging, except OCT relies on backscattered light instead of backscattered sound. OCT is capable of real-time in vivo imaging with multiple frames per second with little image degradation due to motion of the specimen [2, 3].

Despite the high-resolution OCT provides, human interpretation of fine details is often complicated by the presence of noise. Noise in OCT images is primarily speckle, which has prompted the development of various hardware-based techniques, such as polarization diversity, spatial compounding, and frequency compounding. Although such techniques do reduce speckle, each involves hardware modifications that can be expensive and inconvenient to implement. This has led to various post-processing algorithms, such as median filtering, homomorphic Wiener filtering, and multi-resolution wavelet analysis, to suppress noise in OCT data. Wiener filtering utilizes the second-order signal and noise statistics and is useful in removing additive white Gaussian noise, but is not effective at suppressing multiplicative speckle noise. Arsenault addressed this problem by applying the Wiener filter to the logarithm of an image and computing the exponent of the resulting data to obtain the final image [4].

In this paper, we describe a multi-dimensional denoising procedure for post-processing a sequence of OCT images collected from a real-time or three-dimensional OCT imaging system. Unlike existing approaches to suppression of noise in OCT data, our technique is specifically aimed at the multiple images case and is therefore able to leverage the additional information these multiple images provide.

Section 2 discusses multi-dimensional denoising of a sequence of OCT images. Before motivating and describing the proposed denoising method, optimal denoising of an OCT image sequence, and its associated issues, is explored. In Section 3, the proposed technique is applied to a sequence of 50 in vivo OCT images to demonstrate its ability to reduce image noise.

2. MULTI-DIMENSIONAL DENOISING OF AN OCT IMAGE SEQUENCE

Real-time OCT imaging produces a time-series of images, each of which must be estimated to suppress noise. The most basic approach to this problem would be to process each image separately, treating it as independent of the others. A major shortcoming of this type of scheme is that it fails to capitalize on any correlations between the images. A more sophisticated technique would attempt to compute the denoised estimate by accounting not only for the information contained in the individual images, but also for the correlation among them.

Let $Y_n(m_x, m_y)$ denote position (m_x, m_y) of the n^{th} *observed* image from a sequence of N OCT images with spatial dimensions $M_x \times M_y$. The observed OCT signal may be modeled as

$$Y_n(m_x, m_y) = S_n(m_x, m_y) V_n(m_x, m_y) + W_n(m_x, m_y)$$

where $S_n(m_x, m_y)$ is the desired OCT signal that has been corrupted by the multiplicative noise $V_n(m_x, m_y)$ and the additive white Gaussian noise $W_n(m_x, m_y)$. The multiplicative noise term models speckle noise present in OCT data and the additive noise captures the sensor noise.

2.1. Logarithmically Transformed OCT Data

The combination of additive and multiplicative noise in OCT data makes direct processing of the OCT images a challenging task. A common approach to this address this problem is to log-transform the observed data prior to processing. To understand the effects this transformation has on OCT data let

$$X = S_n(m_x, m_y) V_n(m_x, m_y) \qquad Z = W_n(m_x, m_y)$$

for some n, m_x, and m_y, which separates the multiplicative and additive components into X and Z, respectively.

It is well known that the observed OCT signal generated from a polarized light source and free of additive noise obeys a Rayleigh distribution and has a probability density of:

$$p(X) = \frac{X}{4\sigma^2 k^2 A_r^2} e^{-\frac{X^2}{8\sigma^2 k^2 A_r^2}}$$

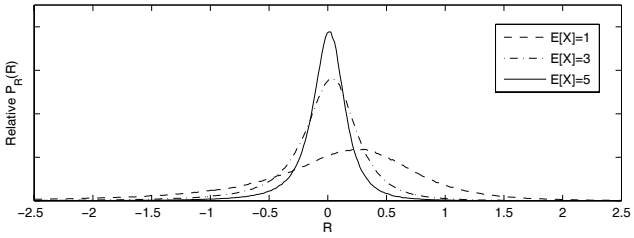

Fig. 1. Numerically computed probability density of the difference between the logarithm of the complete observation and the logarithm of the speckle only observation. Larger values of $E[X]$ correspond to the speckle being more dominate. This translates to the difference distribution being approximated more closely by a zero-mean Gaussian distribution.

where k is a constant factor, A_r is the signal amplitude from the reference arm, and σ is the standard deviation [5]. It can be shown [6] that the probability density of the natural logarithm of this value is

$$p(X_L) = \frac{e^{2X_L}}{4\sigma^2 k^2 A_r^2} exp\left(-\frac{e^{2X_L}}{8\sigma^2 k^2 A_r^2}\right)$$

where $x_L = \log_e(X)$. Though not exact, the distribution of x_L can be reasonably approximated with a Gaussian distribution. Thus, log-transforming an OCT signal (corrupted only by speckle) approximately converts the multiplicative speckle noise into additive Gaussian noise.

If applying a logarithm to an OCT signal converts the multiplicative speckle into approximately Gaussian noise, a natural question to ask is what is the effect of this transformation on the original additive noise. Let R denote the difference between the logarithm of the complete observation and the logarithm of the speckle only observation

$$R = \log_e(X+Z) - \log_e(X)$$

The distribution of this difference signal was found numerically for the case where Z is a standard normal and X is a Rayleigh with various mean values. As can be seen from Figure 1, this difference value is also well approximated by a Gaussian distribution. This, combined with the results for a speckle only observation, implies that $\log_e(X+Z)$ is reasonably well approximated by a signal plus additive Gaussian noise. And, when speckle is the dominate noise, as is often the case, the additive log-domain noise is zero mean. Therefore, we will model the logarithm of the observed data as

$$y_n(m_x, m_y) = s_n(m_x, m_y) + w_n(m_x, m_y)$$

where $y_n(m_x, m_y) = \log_e(Y_n(m_x, m_y))$, $s_n(m_x, m_y)$ is the log of the desired signal, and $w_n(m_x, m_y)$ is zero-mean Gaussian noise that is independent of the signal and captures the total additive noise in the log-domain.

2.2. Optimal Denoising of an OCT Image Sequence

The basic task of denoising is to provide an estimate of the noise-free signal. In the context of the problem at hand, we aim to estimate log-domain signal $s_n(m_x, m_y)$ given the log-domain observation $y_n(m_x, m_y)$.

Collapsing our notation into vector form, let $\mathbf{y}$, $\mathbf{s}$, and $\mathbf{w}$ denote the length-NM_xM_y vectors containing the elements (ordered by space and then time) of the observed signal, desired signal, and additive noise, respectively, in the log-domain. A linear estimate of the signal from the observation is of the general form

$$\hat{\mathbf{s}} = \mathbf{G}\mathbf{y}$$

where $\mathbf{G}$ is an $NM_xM_y \times NM_xM_y$ matrix implementing the estimator.

Recalling that our log-domain noise is modeled as zero-mean and independent of the signal, the minimum mean-square error estimator is the Wiener filter, which has the well known form

$$\mathbf{G}_o = \mathbf{R}_{\mathbf{ss}}\left(\mathbf{R}_{\mathbf{ss}} + \mathbf{R}_{\mathbf{ww}}\right)^{-1}$$

Here $\mathbf{R}_{\mathbf{ss}} = E\left[\mathbf{s}\mathbf{s}^T\right]$ and $\mathbf{R}_{\mathbf{ww}} = E\left[\mathbf{w}\mathbf{w}^T\right]$ are the $NM_xM_y \times NM_xM_y$ signal and noise correlation matrices, respectively. The eigenexpansion of $\mathbf{G}_o$ is

$$\mathbf{G}_o = \mathbf{U}_{\mathbf{ss}}\mathbf{\Lambda}_{\mathbf{G}}\mathbf{U}_{\mathbf{ss}}^H \tag{1}$$

$\mathbf{U}_{\mathbf{ss}}^H$ will optimally decorrelate the signal $\mathbf{s}$ in both time and space and is therefore termed the *temporal-spatial KL transform*.

2.3. Blind Denoising of an OCT Image Sequence

Despite its mathematical optimality, the Wiener filter is generally useful only in theory as its computation requires knowledge of the second-order signal and noise statistics and the inversion of a large matrix ($NM_xM_y \times NM_xM_y$). These facts motivate us to create an alternate denoising technique for a series of OCT images.

From (1), it can be seen that the Wiener filter creates an estimate of the true (log-domain) signal by

1. decorrelating the signal in both space and time
2. weighting the decorrelated coefficients
3. recorrelating the signal in both space and time

Our approach to creating an alternate denoising scheme for a sequence of OCT images to mimic the Wiener filter's behavior of decorrelate, weight, recorrelate, in a manner that does not suffer from the issues associated with the optimal estimator. This type of "approximate Wiener filter" approach has been successfully applied to other multi-dimensional estimation problems [7].

In general, the correlation between any two samples in the sequence of images will depend on both the spatial locations of the samples and the indexes of the images those samples are part of. That is, $E\left[s_{n_1}(m_{x_1}, m_{y_1})\, s_{n_1}(m_{x_1}, m_{y_1})\right]$ is dependent on all six indexes, which makes accurate modeling of this correlation difficult. To simply this problem, we will assume that $\mathbf{R}_{\mathbf{ss}}$ may be separated as

$$\mathbf{R}_{\mathbf{ss}} = \mathbf{R}_{\mathbf{s}_t\mathbf{s}_t} \otimes \mathbf{R}_{\mathbf{s}_s\mathbf{s}_s} \tag{2}$$

where $\mathbf{R}_{\mathbf{s}_t\mathbf{s}_t}$ is the $N \times N$ matrix capturing the temporal correlation of the data and $\mathbf{R}_{\mathbf{s}_s\mathbf{s}_s}$ is the $M \times M$ matrix capturing the spatial correlation (in both dimensions). Although the validity of this assumption will depend on the imaged data, it should be reasonable in most cases since $\mathbf{R}_{\mathbf{s}_t\mathbf{s}_t}$ and $\mathbf{R}_{\mathbf{s}_s\mathbf{s}_s}$ can always be set to the *average* correlation matrix in the corresponding dimension. In addition, this assumption of separability has been successfully assumed in several applications requiring multi-dimensional decorrelation [7, 8]

Eigenexpansion of $\mathbf{R}_{\mathbf{s}_t\mathbf{s}_t}$ and $\mathbf{R}_{\mathbf{s}_s\mathbf{s}_s}$ gives

$$\mathbf{R}_{\mathbf{s}_t\mathbf{s}_t} = \mathbf{U}_{\mathbf{s}_t}\mathbf{\Lambda}_{\mathbf{s}_t}\mathbf{U}_{\mathbf{s}_t}^H$$
$$\mathbf{R}_{\mathbf{s}_s\mathbf{s}_s} = \mathbf{U}_{\mathbf{s}_s}\mathbf{\Lambda}_{\mathbf{s}_s}\mathbf{U}_{\mathbf{s}_s}^H$$

$\mathbf{U}_{\mathbf{s}_t}^H$ and $\mathbf{U}_{\mathbf{s}_s}^H$ are termed the *temporal KL transform* and the *spatial KL transform* as they optimally decorrelate the data in time and space, respectively. Applying a basic property of Kronecker products [9], we can show that

$$\mathbf{U}_{\mathbf{ss}}^H = \mathbf{U}_{\mathbf{s}_t}^H \otimes \mathbf{U}_{\mathbf{s}_s}^H \tag{3}$$

The significance of (3) is that the spatial-temporal KL transform, which optimally decorrelates the signal in both space and time, can be decomposed into the spatial KL transform and the temporal KL transform. This allows us to independently address, and approximate, the spatial and temporal decorrelation.

2.3.1. Approximating the Temporal KL Transform

The role of the temporal KL transform is to temporally decorrelate the signal. The ideal approximation to $\mathbf{U}_{\mathbf{s}_t}^H$ should therefore provide excellent decorrelation of the underlying signal in the time dimension without requiring knowledge, statistical or otherwise, of the that signal.

In many situations, the object being imaged may fit into a classification that makes temporal decorrelation more straight forward. For example, if the real-time imaging is performed on an object that has periodic (e.g. cardiac or respiratory function) or slowly varying motion, then the frequency domain representation of a single pixel (over all images) will often contain a small number of significant coefficients. In such cases, the DFT will provide reasonable temporal decorrelation and may be used to approximate the temporal KL transform.

When the source object cannot be classified in any helpful manner (e.g., non-periodic movement or 3-D imaging), it is unlikely that the DFT will provide adequate temporal decorrelation. In these cases, the temporal KL transform may be estimated from the observed data by computing the average (empirical) temporal correlation matrix and computing its eigendecomposition to yield the appropriate transformation. A description of this procedure is described (for multi-spectral images) in [8].

2.3.2. Approximating the Spatial KL Transform

Just as the temporal KL transform decorrelates the signal in time, the spatial KL transform decorrelates the signal in space. Again, we seek to approximate this transform in a manner that does not rely on signal or noise information of any kind.

It is well known that a wavelet basis forms an approximate KL basis for a wide class of signals, including images with natural features (e.g., edges, uniform regions, etc.). This means that the DWT performs blind signal decorrelation effectively for a large number of signals. For this reason, wavelet bases are often used in applications that rely on signal decorrelation such as compression and estimation. With this in mind, we will approximate the spatial KL transform with a 2-D DWT.

2.3.3. Approximating the Weighting Values

The final approximation we must make in order to have a complete estimation scheme is to the weighting values. The role of these values in both the Wiener filter our proposed approximation is to weight each decorrelated coefficient based on its relative content of signal and noise. To limit noise contributions, a coefficient that contains primarily noise should be assigned a weight close to zero. Conversely, to maximize signal contributions, a coefficient that contains primarily signal should be assigned a weight close to one. This is not unlike the thresholding and shrinkage operators that are ubiquitous in wavelet-based denoising applications [10, 11]. Therefore, we will approximate the weighting coefficients with a threshold or shrinkage operation.

Since we have chosen to employ a 2-D DWT to approximately decorrelate in the signal space, the coefficients we need to weight will exist in a wavelet domain. This enables us to employ any one of the many thresholding and shrinkage techniques developed for wavelet domain coefficients, including both simple coefficient shrinkage [11] and sophisticated adaptive techniques [10].

2.3.4. Overall Denoising Scheme

Having approximated the temporal and spatial KL transforms and the weighting values of the optimal denoising operator, we now have a complete scheme that overcomes the limitations of the Wiener filter and is capable of blind denoising of a sequence of OCT images. The complete technique consists of eight steps:

1. Compute logarithm of image data
2. [Optional] Estimate the temporal KL transform
3. Decorrelate in time using the DFT or estimated temporal KL transform
4. Decorrelate in space using the 2-D DWT
5. Apply thresholding or shrinkage to decorrelated values
6. Recorrelate in space using 2-D inverse DWT
7. Recorrelate in time using the inverse DFT or estimated temporal KL transform
8. Compute exponential to convert back into image data

3. EXAMPLE

Spectral domain OCT [12] data was acquired using system configured as shown in Figure 2 from an in vivo Xenopus laevis (African frog) tadpole, a common developmental biology animal model and one used routinely for demonstrating OCT imaging performance. Imaging was performed along the ventral surface of the tadpole. Real-time motion of the heart was acquired using a 20-mm focal length lens over a 2 mm by 2 mm scan with an image size of 1000 pixels by 1024 pixels at a processed rate of 6 frames per second. A total of 50 frames were captured.

The 50 OCT images were processed with the proposed denoising scheme using a Daubechies length-8 wavelet and the BayesShrink adaptive thresholding rule described in [10]. Since this dataset consists of periodic motion, the DFT was used for temporal decorrelation. For comparison, the denoising was repeated using an estimated temporal KL transform rather than the DFT, no visual difference between the two denoising results could be seen. Representative acquired and denoised images are shown in Figure 3 and a magnified subsection is presented in Figure 4. As can be seen from these images, the proposed scheme effectively suppresses noise in the images without significant signal distortion. Region 2, outlined in white on the acquired image in Figure 3, primarily contains background noise. The standard deviation in this region was reduced from 104.37 in the acquired image, to 24.28 in the denoised image. Note that the significant noise present above the heart is due to strong back-reflections from the heart. Since this does not fit the multiplicative or additive model we do not expect this noise to be effectively removed.

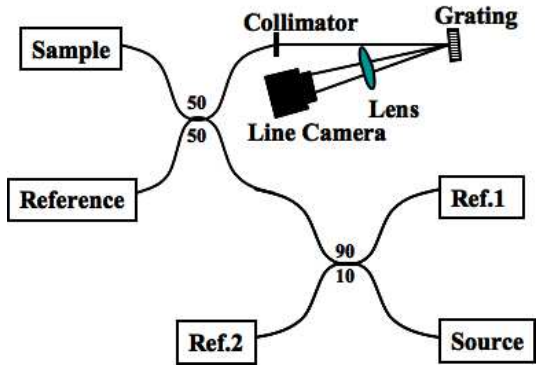

Fig. 2. Spectral domain optical coherence tomography system setup. The system is comprised of a fiber-based interferometer and a free-space spectral detector. A titanium sapphire source with a center wavelength of 800 nm and a bandwidth of 100 nm was used to illuminate the sample. The additional 90/10 splitter allows for reference of the source spectrum.

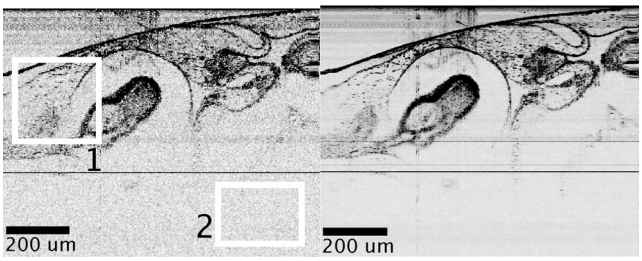

Fig. 3. Representative frame for acquired (left) and denoised (right) data. The white outline region marked 1 in acquired image is the location of the magnified images shown in Figure 4.

4. CONCLUSIONS

Optimal denoising of a sequence of log-transformed OCT images is achieved by the Wiener filter, which requires second-order statistical knowledge of both the true signal and the corrupting noise. Key to the Wiener filter's performance is its ability to optimally decorrelate the true signal in both time and space and to weight these decorrelate values according to their relative signal and noise content.

Blind spatial decorrelation can be achieved using a 2-D DWT in place of the optimal spatial decorrelator. A suitable transform to decorrelate the data temporally can be estimated from the log-domain images. When the sequence of images contains periodic of slowly varying motion the DFT can be used to decorrelate in time, enabling the transform estimation step to be skipped.

Weighting decorrelated coefficients to suppress noise and preserve signal may be accomplished using any of the various wavelet thresholding and shrinkage rules that exist; we found that the BayesShrink adaptive threshold performs well. This rule, combined with the 2-D DWT and estimated temporal decorrelating transform (or DFT when appropriate), yields a complete denoising scheme for a sequence of OCT images.

Although the technique itself is not capable of operating in real-time, applying it to a sequence of 50 images acquired from a real-time OCT imaging system removed an appreciable amount of noise from the images and decreased the standard deviation of the background noise by a factor greater than four.

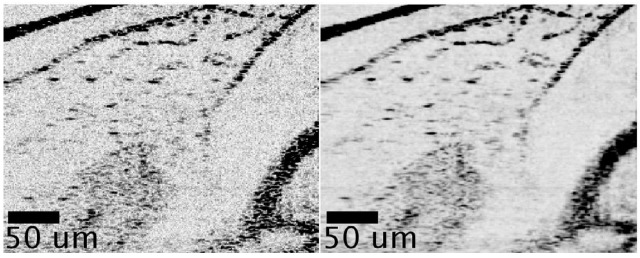

Fig. 4. Magnified view of region 1 of acquired (left) and denoised (right) data.

5. REFERENCES

[1] Joseph M. Schmitt, "Optical coherence tomography (OCT): A review," *IEEE J. Select. Topics Quantum Electron.*, vol. 5, no. 4, pp. 1205–1215, 1999.

[2] S. Bourquin, A. D. Aguirre, I. Hartl, P. Hsuing, T. H. Ko, and J. G. Fujimoto, "Ultrahigh resolution real time OCT imaging using a compact femtosecond Nd:Glass laser and nonlinear fiber," *Optics Express*, vol. 11, no. 24, pp. 3290–3297, 2003.

[3] S. H. Yun, G. J. Tearney, B. E. Bouma, B. H. Park, and J. F. de Boer, "High-speed spectral domain optical coherence tomography at 1.3 μm wavelength," *Optics Express*, vol. 11, no. 26, pp. 3598–3604, 2003.

[4] H. H. Arsenault and G. April, "Properties of speckle integrated with a finite apeture and logarithmically transformed," *Optical Society of America*, vol. 66, pp. 1160–1163, 1976.

[5] Michael Pircher, Erich Götzinger, Rainer Leitgeb, Adolf F. Fercher, and Christoph. K. Hitzenberger, "Speckle reduction in optical coherence tomography by frequency compounding," *Journal of Biomedical Optics*, vol. 8, no. 3, pp. 565–569, July 2003.

[6] Hua Xie, Leland E. Pierce, and Fawwaz T. Ulaby, "Statistical properties of logarithmically transformed speckle," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, no. 3, pp. 721–727, Mar. 2002.

[7] Ian Atkinson, Farzad Kamalabadi, Satish Mohan, and Douglas L. Jones, "Asymptotically-optimal blind multichannel image estimation," *IEEE Trans. Image Processing*, In Press.

[8] B. R. Hunt and Olaf Kübler, "Karhunen-loeve multispectral image restoration, part I: Theory," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 3, pp. 592–600, June 1984.

[9] John W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, no. 9, pp. 772–781, Sept. 1978.

[10] S. Grace Chang, Bin Yu, and Martin Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Processing*, vol. 9, no. 9, pp. 1532–1546, Sept. 2000.

[11] David L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[12] A. F. Fercher, C. K. Hitzenberger, G. Kamp, and S. Y. El-Zaiat, "Measurements of intraocular distances by backscattering spectral interferometry," *Opt. Comm.*, vol. 117, pp. 43–48, 1995.