Noise Reduction of Swept-Source Optical Coherence Tomography via Compressed Sensing

Volume 10, Number 1, February 2018

Site Luo
Qiang Guo
Hui Zhao
Xin An
Liang Zhou
Huikai Xie
Jianyu Tang
Xiao Wang
Hongwei Chen
Li Huo
Noise Reduction of Swept-Source Optical Coherence Tomography via Compressed Sensing

Site Luo, Qiang Guo, Hui Zhao, Xin An, Liang Zhou, Huikai Xie, Jianyu Tang, Xiao Wang, Hongwei Chen, and Li Huo

1Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
2Foshan Optomedic Technologies Company, Ltd., Foshan, China
3Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou 100864, China
4University of Florida, Gainesville, FL 32611 USA
5Wuxi WiO Technologies Company, Ltd., Wuxi, China

DOI:10.1109/JPHOT.2017.2787662
1943-0655 © 2017 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Abstract: In this paper, we investigate noise reduction in swept-source optical coherence tomography (OCT) using compressed sensing (CS). Multiple scan averaging is a classical method used to enhance the quality of OCT images by reducing the noise of a system. However, the conventional averaging method requires a repetitive scan at the same location and thus reduces the imaging speed. In this paper, the sparsity property of an OCT A-scan is utilized, and one full A-scan OCT image can be reconstructed from a portion of the acquired data during one sweep period using CS. Thus, multiple OCT A-scans can be reconstructed from a single sweep. The average A-scans yield a better quality than the single A-scan obtained from the whole data acquired during a sweep period. We demonstrate that the average of five reconstructed A-scans from a single sweep using CS offers an image quality and depth resolution similar to those obtained by averaging three sequential A-scans from three sweeps using the conventional averaging method. This proposed method can shorten the time required to perform repetitive scans and thus improve the imaging speed.

Index Terms: Swept-source optical coherence tomography, compressed sensing, noise reduction, multiple scan averaging.

1. Introduction
Optical coherence tomography (OCT) is a rapidly developing imaging technology in recent decades [1], [2] that allows one to obtain a sequence of cross-sectional images with high resolution of 1 to 10 μm. Currently, several types of OCT techniques with faster speed and higher resolutions have been developed [3]–[5]. Compared with time-domain OCT (TD-OCT), spectral domain optical coherence tomography (SD-OCT) and swept-source optical coherence tomography (SS-OCT) appear to have superior sensitivity and higher imaging speeds [6]–[8]. Because of the interferometric nature of OCT, noise reduction has long been a focus of OCT research. In the past few years, various denoising methods have been proposed to improve the quality of OCT images, such as filtering,
wavelet analysis, and speckle statistics [9]–[12]. Multiple B-scan averaging is a classical method used to enhance the quality of OCT images by reducing the speckle noise; this method is also known to efficiently reduce the noise of a system [13]–[16]. However, the conventional averaging method requires a repetitive B-scan at the same location and thus decreases the imaging speed.

In recent years, a novel technique called compressed sensing (CS) has been applied to SD-OCT that allows OCT image acquisition using a photodetector array with fewer pixels [17], achieves real-time high-speed OCT imaging [18]–[21], and obtains volumetric OCT images quickly [22]–[24]. The CS method has been demonstrated to be an effective means to improve the signal-to-noise ratio (SNR) because it exploits the signal sparsity [25]. Fang et al. presented a novel hybrid approach to learn a sparse representation dictionary for each of these high-SNR images in a SD-OCT system and employ these dictionaries to denoise the neighboring low-SNR B-scans [26]. Xu et al. demonstrated a modified CS method to achieve better image quality [27]. However, the aforementioned denoising methods based on CS are all focused on SD-OCT. Currently an investigation of noise reduction in SS-OCT using a CS method has not been found in the literature.

In this paper, we demonstrate a method using CS to enhance the image quality of SS-OCT based on A-scans. In our SS-OCT, the sparsity property of an OCT A-scan is utilized, and each A-scan can be fully reconstructed from a portion of the data acquired during a sweep period using CS. Thus, multiple reconstructed A-scans can be obtained in each sweep period. The average of these scans has a better visual quality than the single A-scan obtained from the whole data acquired during a sweep period. In our experiment, it is found that averaging five reconstructed A-scans from a single sweep using CS offers a similar image quality compared to that of averaging three sequential A-scans from three sweeps using the conventional averaging method, and no degradation of the depth resolution is observed. This proposed method can reduce the number of the repetitive scans and thus improve the imaging speed. One outstanding advantage for this CS-based SS-OCT method is that it is only based on software and requires no hardware modification. To the authors’ best knowledge, this is the first time CS has been used to improve the SNR of SS-OCT.

2. Methods

2.1 The Signal-To-Noise Ratio of Swept-Source Optical Coherence Tomography

In a SS-OCT system, the optical signals reflected back from the reference arm and the sample arm are combined and then directed onto a photodetector. Tomographic images with the depth information of the sample can be reconstructed by processing the interference signal from the photodetector using discrete Fourier transform (DFT). According to the Wiener-Khintchine theorem, power spectrum of a signal is the amplitude of the Fourier transform of its autocorrelation. The SNR of the SS-OCT is proportional to the number of evenly spaced wavenumber samples per A-scan (denoted as) and is given by [2], [6], [7]:

\[
\text{SNR} = \frac{|F_s(z)|^2}{\langle F_s^2 \rangle} = \frac{N_s \langle i_s^2 \rangle}{2 \langle i_n^2 \rangle}
\]  

(1)

where \( z \) is the depth, and \( i_s, i_n \) are the signal’s photocurrent and the noise’s photocurrent, respectively, and \( F_s \) and \( F_n \) are the Fourier transform intensities of \( i_s \) and \( i_n \) at depth of \( z \), respectively.

2.2 Compressed Sensing

For OCT, the signals from A-scans are sparse, i.e., a relatively small number of data points with sufficiently large coefficients capture most of the depth information. Let \( x \) denote the OCT image data points from an A-scan of size \( N \times 1 \) that is a sparse vector with at most \( K \) nonzero entries \( (K \ll N) \). In an OCT system, the relationship between the A-scan image data \( x \) and the acquired data \( y \) of size \( N \times 1 \) is given by (2):

\[
x = Fy
\]  

(2)
where $F \in \mathbb{R}^{N \times N}$ is a Fourier matrix. According to CS theory, the vector $x$ can be reconstructed exactly from a subvector of $y$, denoted here as $y_u$, with a length of $M$ ($M < N$), and $M = O(K \cdot \log(N/K))$ as shown in (3). The corresponding Fourier matrix should be a submatrix of $F$ (denoted here as $F_u$) [28].

$$y_u = F_u^{-1}x$$

(3)

The goal of the reconstruction process is to seek approximate solutions to the problem posed in (3):

$$\min_x \frac{1}{2} \|y_u - F_u^{-1}x\|^2 + \tau \|x\|_1$$

(4)

where $\tau$ is a nonnegative parameter, $\|y_u - F_u^{-1}x\|_2$ denotes the Euclidean norm of $y_u - F_u^{-1}x$, and $\|x\|_1 = \sum_i |x_i|$ is the $\ell_1$ norm of $x$. Finally, an estimated A-scan vector $x_e$ is obtained [29]–[32].

Due to the sparsity property of an OCT A-scan, it can be reconstructed reliably from partial data acquired during a sweep period using CS [28], [32]. A whole set of A-scan data are separated into five data groups, and every data group can reconstruct reliably the original signal and random noise. The random noise reconstructed by the CS algorithm can be reduced through averaging multiple B-scan images.

To evaluate the OCT image quality, the average image of 10 repetitive OCT B-scan images was considered as the noiseless reference image [21]. The image processing flow diagram is shown in Fig. 1. Ten B-scan data sets ($1000 \times 1024$) before DFT are collected, which are repetitive scans at the same location. Every data set can be constructed to form a B-scan image with DFT (named 1024DFT image). The average image of the ten 1024DFT images is considered as the ‘noiseless’ image, as shown in the first image of Fig. 1. The averaging images of sequential three and two 1024DFT images are also obtained with the same method and shown in Fig. 1(a) and (b). The single 1024DFT image is shown in Fig. 1(c). The first 512 data of one A-scan of the original B-scan data set is extracted out to be a new data set ($1000 \times 512$). The reconstructed image (named 512DFT image) of the new data set is shown in Fig. 1(d). It should have worse image quality than 1024DFT image because of the lower amount of data and signal information. However, the new data set can be expanded to the size of $1000 \times 1024$ using the CS algorithm (named 512CS1024 image), as shown in Fig. 1(e). Five new data sets are extracted from the 1024 data points of a single A-scan image, which are the data points 1-512, 129-640, 257-768, 384-895, and 513-1024, respectively. Then five 512CS1024 images are reconstructed with the GPSR_BB CS algorithm reported by Figueiredo, M [33]. The five data sets all consist of 512 data points. The 512 data length is a balanced choice to achieve an appropriate image quality. The averaging image of the five 512CS1024 images is shown in Fig. 1(f). And the corresponding images are shown in Figs. 3 and 4 for imaging two samples.

For a two-dimensional (2-D) B-scan image, we use the peak signal-to-noise ratio (PSNR) to evaluate the image quality. The PSNR is expressed as [26]:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2$$

(5)

$$PSNR = 10 \cdot \log_{10} \left( \frac{\text{MAX}^2}{MSE} \right) = 20 \cdot \log_{10} \left( \frac{\text{MAX}}{\sqrt{MSE}} \right)$$

(6)

Where $MSE$ is the mean square error of two images, $m$ and $n$ is the image’s size, $I$ is the standard original image, $K$ is the noisy image, and MAX is the maximum pixel value. Because the OCT images data are 8 bits, the MAX is 255. As shown in Fig. 1, the PSNRs between the six types of images and the “noiseless” image are defined as PSNR-a~PSNR-f.
3. Experimental Setup and Results

The light source was a MEMS-based wavelength swept source (HSL-20-100-B). The data acquisition card (ATS9350, Alarztec Technologies) served as a dual channel device. The swept source generated a clock signal to control the sampling process. Each acquired B-scan image consisted of 1000 A-scans, and each A-scan was composed of 1024 evenly spaced wavenumber samples. If only the first 512 sample points were used to reconstruct an A-scan image by using DFT, then the SNR would dramatically decrease. However, we can use the CS method described in Section 2.2 to generate an A-scan of length 1024 from the first 512 sample points, resulting in an enhancement of the SNR.

The first sample in the experiment was two coverslips, each 160 $\mu m$ in thickness, with an air gap of approximately 15 $\mu m$ between them. Fig. 2 shows an A-scan obtained from all 1024 sample points during a sweep period using DFT (named 1024DFT), an A-scan obtained from the first 512 sample points with zero-filling to 1024 (named 512DFT) and an A-scan of length 1024 reconstructed from the first 512 sample points using CS (named 512CS1024). The other two A-scan lines is the average of three 1024DFT line (named 3Aver1024DFT) and the average of five 512CS1024 line (named 5Aver512CS1024). Fig. 2(a) shows a comparison of the five A-scan curves at the same location (corresponding to the black line, the blue line, the red line, the green line and the yellow line). The horizontal axis denotes the index of depth at A-scan pixels, and the vertical axis denotes...
Fig. 2. (a) Composite A-scan curves of 1024DFT, 512DFT, 512CS1024, 3Aver1024DFT and 5Aver512CS1024 at the same location. (b) The enlarged image of the region containing four reflecting surfaces. (c) The enlarged image of the region containing noise.

Fig. 3. The image result of two coverslips with six methods. (a) The average of three 1024DFT images. (b) The average image of two 1024DFT images. (c) The 1024DFT image. (d) The 512DFT image. (e) The 512CS1024 image. (f) The average image of five 512CS1024 images.

the normalized power. The A-scan curve shows four reflecting surfaces corresponding to the top and bottom surfaces of the two coverslips. Fig. 2(b) shows a magnified view of the curves of the four reflecting surface regions, and the 1024DFT curve shows a higher resolution than the 512DFT curve and the 512CS1024 curve; the 3Aver1024DFT curve and 5Aver512CS1024 curve appear to have similar resolution to the 1024DFT curve. Fig. 2(c) shows a zoom-in view of the noise region, indicating that the SNR of the A-scan obtained from 512 sample points using DFT is lower than the SNR of the other four. However, the reconstructed A-scan using the CS method has approximately the same SNR as the A-scan obtained from treatment of the all 1024 sample points using DFT.
The noise intensity of 3Aver1024 DFT curve and 5Aver512CS1024 curve are approximately equal to each other, and both of them show lower fluctuation than that of the 1024DFT curve, which is the result of multiple-images-averaging effect.

The image results of two coverslips are shown in Fig. 3, and the six images correspond to the six types of image processing methods illustrated in Fig. 1. The 512DFT image (Fig. 3(d)) shows the worst image quality in the all six images. The 512CS1024 image and the average image of five 512CS1024 images decrease the speckle noise, i.e., the CS method improved the image quality. The air gap region cannot be distinguished at the 512DFT image (Fig. 3(d)), whereas the 512CS1024 image and the average image of five 512CS1024 images (Fig. 3(e) and (f)) show the air gap clearly. A similar trend is also observed from the single 1024DFT image and the averaging images of two and three 1024DFT images; the image quality becomes better with a larger averaging number. Moreover, Fig. 3(a) and (f) appear to have similar image quality based on visual inspection; this evaluation can be supported by the PSNR value.

To evaluate the image quality quantitatively, the PSNRs among those images and the 'noiseless' image are calculated, as listed in Table 1. The PSNR-d, which is the lowest value of the all six cases, results from the smallest data set. The value of PSNR-d, PSNR-e and SNR-f suggest that CS method is capable of improving image quality in SS-OCT technique. However, the 512CS1024 image has lower SNR compared to the 1024DFT image. The PSNR of averaging five 512CS1024 images is approximately the same as that of averaging three 1024DFT images. The light beam is scanned by a galvanometer. Thus the straight lines of the interfaces in the coverslip stack become fan shaped [34].

### Table 1
The PSNR Values of the Two-Coverslips Sample

<table>
<thead>
<tr>
<th></th>
<th>PSNR-a</th>
<th>PSNR-b</th>
<th>PSNR-c</th>
<th>PSNR-d</th>
<th>PSNR-e</th>
<th>PSNR-f</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50.61</td>
<td>49.73</td>
<td>47.84</td>
<td>46.80</td>
<td>47.44</td>
<td>50.75</td>
</tr>
</tbody>
</table>

The noise intensity of 3Aver1024 DFT curve and 5Aver512CS1024 curve are approximately equal to each other, and both of them show lower fluctuation than that of the 1024DFT curve, which is the result of multiple-images-averaging effect.
TABLE 2
The PSNR Values of the Small Intestine Sample

<table>
<thead>
<tr>
<th>PSNR-a</th>
<th>PSNR-b</th>
<th>PSNR-c</th>
<th>PSNR-d</th>
<th>PSNR-e</th>
<th>PSNR-f</th>
</tr>
</thead>
<tbody>
<tr>
<td>43.29</td>
<td>41.93</td>
<td>39.51</td>
<td>37.86</td>
<td>38.98</td>
<td>43.39</td>
</tr>
</tbody>
</table>

Fig. 5. (a) The A-scan of six image processing methods. (b) The enlarged image of the region containing noise.

We also analyzed a biological sample (small intestine of pig); the results are shown in Fig. 4. Similarly, we calculated the respective PSNRs, as listed in Table 2. The two coverslips sample has only four reflecting surfaces, which yields a higher SNR. In contrast, the small intestine sample is a biological tissue with high absorption, resulting in lower PSNR values.

A similar pattern can be observed from the imaging results of both samples (Tables 1 and 2). Upon visual inspection, it is obvious that the 1024DFT image and 512CS1024 image show higher image quality when compared to the 512DFT image. The average image of three 1024DFT images appear to have nearly the same visual quality as the average image of five 512CS1024 images and are much better than that of the single 1024DFT image. The PSNR values shown in Tables 1 and 2 further validate the above observation. Fig. 5 shows a comparison of the A-scans using the six image processing methods and the noise range located at upside of surface (20∼90 pixels). The 512DFT curve presents the highest noise level, which is approximately 6 dB more than that of the 3Aver1024DFT curve and 5Aver512CS1024 curve. The 2Aver1024 line, 1024DFT line and 512CS1024 line are located in the middle range. The qualitative relationship of the six curves greatly matches that of Table 2.

4. Conclusions
In this study, we demonstrated a method that utilizes a CS algorithm to improve the SNR of SS-OCT images. A two-layer coverslip sample and a pig’s small intestine sample were imaged to test the CS algorithm. The experiments showed that the 512CS1024 image and 1024DFT image achieve nearly the same resolution and noise level. The experimental results also suggested that the image quality of averaging five 512CS1024 images reconstructed from a single B-scan 1024DFT image is much
better than that of the original 1024DFT image and is equivalent to that of averaging three 1024DFT images. In other words, the number of multiple repetitive scans can be reduced by one-third, i.e., the image speed can be improved by 2 times.

As we know, the CS algorithm is based on sparsity property, and a mirror reflecting surface can be imaged just by a few data points. In our study, the high-quality images of the one-coverslip sample and two-coverslips sample can also be reconstructed by 256 data points, and even 128 data points, of every A-scan using the CS method. However, this CS algorithm has poor ability to reconstruct biological tissue. Usually, a CS algorithm shows a worse performance for a non-sparse sample when compared to a sparse laminar sample, such as a coverslip, thin film, retina, etc. We are attempting to develop a modified CS algorithm to obtain a similar image quality of a 1024DFT image with 256 (or even 128) data in biological tissue. In this manner, we can achieve higher imaging speed or a higher SNR.

References


