Denoising methods for improving automatic segmentation
in OCT images of human eye

A. STANKIEWICZ¹, T. MARCINIAK¹*, A. DĄBROWSKI¹,
M. STOPA²,³, P. RAKOWICZ²,³, and E. MARCINIAK²

¹Division of Signal Processing and Electronic Systems, Poznan University of Technology, 24 Jana Pawla II St., 60-965 Poznan, Poland
²Clinical Eye Unit and Pediatric Ophthalmology Service, Heliodor Swiecicki University Hospital, Poznan University of Medical Sciences,
16/18 Grunwaldzka St., 60-780 Poznan, Poland
³Department of Optometry and Biology of Visual System, Poznan University of Medical Sciences,
5D Rokietnicka St., 60-806 Poznan, Poland

Abstract. This paper presents analysis of selected noise reduction methods used in optical coherence tomography (OCT) retina images (the so-called B-scans). The tested algorithms include median and averaging filtering, anisotropic diffusion, soft wavelet thresholding, and multiframe wavelet thresholding. Precision of the denoising process was evaluated based on the results of automated retina layers segmentation, since this stage (vital for ophthalmic diagnosis) is strongly dependent on the image quality. Experiments were conducted with a set of 3D low quality scans obtained from 10 healthy patients and 10 patients with vitreoretinal pathologies. Influence of each method on the automatic image segmentation for both groups of patients is thoroughly described. Manual annotations of investigated retina layers provided by ophthalmology experts served as reference data for evaluation of the segmentation algorithm.

Key words: optical coherence tomography (OCT), image denoising, image segmentation, anisotropic diffusion, wavelet thresholding.

1. Introduction

The newest measurement technologies provide automatic visualization and analysis of pathologic tissues. This process can be further extended through application of advanced algorithms for analysis of medical images [1]. The detailed measurements and proper representation of thickness, volume and placement of the examined structures make it easier for the specialists to chose the proper treatment course [2].

Among noninvasive techniques for soft tissue measurement used in ophthalmology is the spectral domain optical coherence tomography (SD-OCT) [3]. This technology is based on illuminating the tissue with a stream of infrared light. The light reflected from inner structures of the eye, is spectrally analyzed to receive data, representing layers of the retina. A set of single scans across the retina (called A-scans) is assembled into a cross-section (i.e. the B-scan) illustrating the layered morphological structure of the retina. A series of B-scans creates a three-dimensional visualization (3D OCT scan) of the retina.

Figure 1 presents a 3D OCT scan annotated with the most commonly identified retina layers: inner limiting membrane (ILM), nerve fiber layer (NFL), ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer (ONL), inner segments of photoreceptors (IS), outer segments of photoreceptors (OS), and retinal pigment epithelium (RPE) [4].

Fig. 1. Example of 3D OCT scan of human retina with annotated most important layers

2. Characteristics of OCT interfaces

The inspection of biometric properties of an eye using the OCT technology is based on segmentation and identification of the most important segments (such as retina layers and borders of the optic disc) and on the analysis of their depth at various
points. This information is required to perform the detailed and proper diagnosis that can point out possible treatment courses. Fig. 2 presents a general scheme of the OCT analysis procedure.

Using modern OCT devices it is possible to acquire even 70,000 A-scans per second [5]. This means performing a 3D scan, of 141 B-scans with the resolution of $640 \times 385$ points each, in around 0.8 second. The fast measurement assures no artifacts caused by involuntary movements of the eyeball. Some of the newest devices employ also a motion correction technology (MCT) in order to minimize this problem [6].

The B-scans constituting the 3D OCT scan can be very noisy, what is the main cause of errors during the automatic segmentation of layers. The research on the nature of noise in the OCT images [7] proves, that it is not an entirely random noise as it contains some specific information. It is called the speckle noise. It depends on:

- the placement in the image
- intensity scale of the image
- reflecting properties of the examined tissue.

It is not possible to automatically assess its level. Due to the above features, elimination of noise from the OCT images is a very difficult task, though necessary for the proper tissue analysis.

Fig. 3 illustrates an example of the B-scan acquired through a 3D OCT examination. This image has a low quality due to the noise. It includes pathological changes caused by the vitreomacular traction syndrome (tearing of the retina in the fovea region due to anomalous vitreous detachment). This pathology causes incorrect segmentation of retina layers performed with the use of the automatic procedure presented in Section 3.

Methods for reducing the speckle noise can be divided into two groups: methods based on averaging of a series of images and algorithms designed for denoising single frames (B-scans).

The first group of methods requires multiple scans of the same area that are next averaged [8, 9]. Since the noise present in every cross-section has various spatial distributions (as opposed to the examined tissue) this is an efficient approach to this noise reduction. The more images used for the averaging, the better is the noise minimization and the tissue structure enhancement.

Unfortunately, this technique significantly prolongs the measurement time. Indeed, for the above example, the standard number of frames is 32, but using MCT even 80, what means the time acquisition of 24 and 62 seconds, respectively for the earlier discussed 3D scan. This is not acceptable for a patient. The prolonged acquisition would also cause shift and rotation between subsequent B-scans in the 3D dataset due to unexpected movement of the patient/eyeball, problems with maintaining vision focus and blinking. These problems do not affect scan acquisition under a second, therefore this method can be used only during the acquisition of a single B-scan (one Line Scan Pattern) through the center of macula.

Methods of the second group used for denoising the B-scans in a volumetric set involve:

- averaging and median filtering [10]
- local Bayesian estimation [12]
- diffusion filtering, including nonlinear anisotropic filtering [13–15]
- wavelet thresholding (e.g. spatially adaptive filtering [16], dual tree complex wavelet transformation [17], curvelets transformation [18]).

The results of research conducted by Ozcan et al. [19] suggest superiority of methods based on wavelet transformations in comparison to other techniques. The published findings were derived from data gathered mainly using images of animal tissues (pigs, rats, mice) [8, 15], human skin [20, 21], and human healthy retina [9]. In many cases the noise reduction methods were analyzed on nonmedical and synthetic images [11, 13, 15].

Furthermore, little information can be found in the literature about analysis of speckle noise on OCT images of pathological changes of the retina. One of the reports that used images from 30 patients with 15 various retina pathologies was presented by Abbirame et al. [22].
Recent literature includes also papers describing application of the wavelet thresholding method for a 3D set of human skin images [20, 21]. This method utilizes the information from neighboring frames to minimize the effect of blurring and emphasize the details in the image.

3. Denoising of 3D OCT scans

In this article we analyze the influence of selected denoising methods on the retina layers segmentation procedure. The tests include 3D sets of OCT images for both healthy and pathological retinas. The inspected pathology is the vitreomacular traction syndrome. The group of investigated algorithms is presented in Fig. 4 and described below.

3.1. Averaging filtering. The basic, commonly used method of image denoising is averaging filtering (further referred to as AVG). It employs the two dimensional convolution between an investigated image and a previously defined filter. Such filter is usually defined as a square matrix with odd numbers of rows and columns (e.g. 3×3), although rectangular filters, such as 3×19, can also be found in the literature [23]. In our analysis we tested four variants of such filters, namely: 3×3, 5×5, 7×7, and 9×9.

3.2. Median filtering. Another commonly used nonlinear technique of image denoising is median filtering (further referred to as MED). Its advantage in preprocessing of OCT examinations is the ability to preserve the edges and tissue features, and to reduce the influence of the speckle noise. During the experiment we evaluated four sizes of this type of filter: 3×3, 5×5, 7×7, and 9×9.

3.3. Anisotropic diffusion filtering. Anisotropic diffusion (AD) is an efficient noise reduction method. For an input image this method, proposed by Perona and Malik [24], defines the denoised image as

\[
\hat{I} = c(x, y)\Delta I, \quad (1)
\]

where \(\Delta\) represents the Laplace operator, and \(c(x, y)\) describes the diffusion coefficient dependent on the position \((x, y)\) in the image space, according to the function

\[
c(x, y) = e^{-\frac{|\nabla|^2}{2\kappa^2}}, \quad (2)
\]

where \(\nabla\) defines gradient, \(\kappa\) is the denoising parameter, being a positive real value that is related to the noise level and to the expected edge preservation in the image. For edge detection we use the image gradient calculated with the Prewitt operator, since it is robust to noise [14].

Through choosing the lower value of the diffusion coefficient we avoid the blurring of edges, while bigger value allows for smoothing of areas between the edges. Thanks to this, the lines and structures in the image, important for interpretation, are preserved. This technique is useful for reducing the speckle noise in OCT images.

3.4. Wavelet thresholding. The wavelet thresholding (WT) method provides good results in denoising OCT images. This is mainly due to the fact that the noise is evenly distributed between wavelet coefficients, while the majority of the informative content is concentrated among the coefficients with high magnitude. By selecting a proper threshold value (which might be a difficult task) we are able to reduce the noise maintaining characteristic features of the image [25].

In this algorithm, a single B-scan \(I_t\), represented in the logarithmic scale, is decomposed with the wavelet transform of the maximum decomposition level \(L\). Thereby receiving approximation coefficients \(A_t^L\) and detail coefficients \(W_t^L\), where \(D\) describes direction (horizontal or vertical) of image filtering. During the experiments we used the soft thresholding method with the Haar wavelet, namely the discrete stationary wavelet transform (DSWT) [26].

The denoising operations consist in reducing the detail coefficients for a position \(x\) in the image, based on the weight \(G_t^L\)

\[
W_t^L(x) = G_t^L(x) \cdot W_t^L(x), \quad (3)
\]

where each weight \(G_t^L\) is calculated for a manually selected threshold \(\tau\) according to the following equation

\[
G_t^L(x) = \begin{cases} 
\frac{|W_t^L(x)| - \text{sgn}(W_t^L(x))\tau}{|W_t^L(x)|} & \text{for } |W_t^L(x)| > \tau \\
0 & \text{otherwise}
\end{cases} \quad (4)
\]

The last step of this algorithm requires performing the inverse wavelet transform. Fig. 5a presents a general scheme of the described algorithm.

3.5. Multiframe wavelet thresholding. The multiframe wavelet thresholding method (MWT) uses a set of frames \(i \in (1, N)\), where \(N\) defines the number of the processed B-scans. It assumes that the noise in image \(i\) is uncorrelated with respect to noises in other images (B-scans), and that its standard deviation \(\sigma_i(x)\) at position \(x\) is the same as in the other images.
Authors of this method [8] propose calculating the weight for the detail coefficients in a way that allows for estimation of the local noise. This weight is called the significance weight $G_{\text{sig},i,D}(x)$ and is calculated as follows

$$G_{\text{sig},i,D}(x) = \begin{cases} 1, & \text{for } |W_{i,D}^L(x)| \geq k \sigma_{S,L,D}(x) \\ \theta_i(W_{i,D}^L(x), ..., W_{K,D}^L(x)), & \text{otherwise} \end{cases},$$

where $\sigma_{S,L,D}(x)$ defines the mean squared distance between the detail coefficients in individual images, parameter $k$ describes the noise reduction level, and $\theta_i$ is the normalized parameter calculated as in (7):

$$\theta_i(w_1, ..., w_N) = \frac{1}{N-1} \sum_{j=1,j \neq i}^N \left(1 - \frac{w_i}{w_j}\right).$$

After obtaining new detail and approximation coefficients with the above weights for all images, the coefficients are averaged and the inverse transform is calculated.

This algorithm was developed for processing a set of frames of the same examined tissue area, as is in the case of multiple acquisition of a single B-scan across the fovea. Bearing in mind, that the distance between the subsequent B-scans in the 3D examination is about 50 μm, a little change in the tissue structure is observed in the neighboring cross-sections. Based on this, we propose to use 3 subsequent frames in the 3D OCT set as input images for this method. A general scheme of this approach is illustrated in Fig. 5b.

### 4. Experiments

#### 4.1. Methodology.

As was mentioned earlier, we measured effectiveness of the denoising methods by analysis of influence on the image segmentation accuracy. There are two reasons for this experiment. First, for OCT images a reference image (an ideal image without noise) does not exist, thus, it is difficult to calculate accuracy of denoising algorithms directly. Second, noise in OCT images is causing errors in the segmentation of the retina layers. The segmentation procedure is in turn a key step in defining the morphological structure of the retina during the diagnosis. Visualization and measurement of the retina layers thickness is the base line for the retina analysis.

Current research concerning segmentation of medical images indicates methods based on the graph theory as the most accurate approach [27]. The algorithm selected for this study (reported by Chiu et al. [23]) treats a single OCT B-scan as a graph, in which every pixel is a node. For the created graph, with previously calculated weights, we use the shortest path
to find the line representing the border between two neighboring layers in the image. In the conducted experiment we used a modified version of this algorithm [28, 29]. The modified version takes continuity of tissues into account across all cross-sections and the signal quality in each column of the image.

The 3D OCT scans used for testing were acquired with Avanti RTvue device (Optovue Inc., Freemont, USA [5]) from 10 patients with vitreomacular traction (VMT) pathology and 10 healthy volunteers. Average age for each group was 71 and 39 years, respectively. Every patient had a full 3D OCT examination of macular region of size $7 \times 7 \times 2$ mm. The examined volume is represented by $141 \times 640 \times 385$ data points (141 B-scans with $640 \times 385$ resolution). Next, the acquired scans were manually segmented by experts from the Clinical Eye Unit at the Heliodor Święcicki Medical University Hospital in Poznań to annotate 7 retina layers (further used as reference data for evaluating investigated algorithms).

In the next step, each image was denoised with the earlier described methods. For every method various parameter values were tested. Table 1 presents a list of the tested parameters and their values, with the best results obtained for values marked with gray shading. Also, as a part of subjective evaluation of the denoising algorithms, the ophthalmology experts were given an opportunity to select the best parameter values. It is worth mentioning that they have intuitively selected bigger values then those resulting from the tests (see Fig. 6).

Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>AVG</th>
<th>MED</th>
<th>AD</th>
<th>DSWT</th>
<th>MWT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mask size</td>
<td>$3 \times 3$</td>
<td>$3 \times 3$</td>
<td>$3 \times 3$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td>Value 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mask size</td>
<td>$5 \times 5$</td>
<td>$5 \times 5$</td>
<td>$5$</td>
<td>$10$</td>
<td>$1$</td>
</tr>
<tr>
<td>Value 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mask size</td>
<td>$7 \times 7$</td>
<td>$7 \times 7$</td>
<td>$10$</td>
<td>$30$</td>
<td>$10$</td>
</tr>
<tr>
<td>Value 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mask size</td>
<td>$9 \times 9$</td>
<td>$9 \times 9$</td>
<td>$20$</td>
<td>$100$</td>
<td>$100$</td>
</tr>
<tr>
<td>Value 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2. Comparison of denoising algorithms. Examples of B-scans obtained after application of each investigated denoising method are illustrated in Figs. 6 and 7. Fig. 6A presents an original B-scan, while an example of employing AVG and MED filtering is illustrated in Fig. 6B and 6C respectively. Fig. 6D, 6E and 6F illustrate the results of AD, DSWT and MWT algorithms.

The original single cross-section (A) has a visibly high noise content (a grainy structure). By comparing it with images (B) and (C) we can notice that these methods cause blurring of the image and the noise is still present. Method (E) also leads to blurring of the image, although the regions of individual tissues are smoothed. Additionally, for bigger threshold values rectangular shaped artifacts appear in the image. The greatest unification and smoothing of tissues can be detected in method (D). Unfortunately, this algorithm reduces the line of the posterior vitreous cortex present in the upper left part of
the image. This line is best revealed by method (F), thanks to which, it is possible to maintain most informative content about the pathology distribution. Method (F) also provides a good quality image with low noise content and visible retina tissue areas separation.

Next, each 3D scan was subjected to automatic image segmentation based on the graph theory. Verification of effectiveness of the implemented methods was based on calculation of the peak signal to noise ratio (PSNR):

$$\text{PSNR} = 10 \log_{10} \frac{\text{MAX}^2}{\text{MSE}}, \text{[dB]}$$

where MAX defines a maximum possible value of the annotation range (in our case it is equal to the height of the image), and MSE is a mean squared error between the automatic and manual segmentations. Due to divergence in annotating layer borders by experts and by computer, the difference between them lower than 5 pixels was classified as negligible (i.e., error equal to zero), what would give PSNR value over 45 dB.

### 4.3. Analysis of segmentation accuracy

During the conducted experiment we used an image segmentation algorithm implemented in the Matlab/Simulink R2014b environment [30] for annotating 7 borders between the retina layers: ILM, NFL/GCL, IPL/INL, INL/OPL, OPL/ONL, IS/OS, and RPE/Choroid using the procedure described in Section 4.1.

Figure 7 presents the segmentation results after each investigated denoising method. As can be seen, method (F) provides best accuracy for segmenting pathological tissue. Although still some inadequacies can be found, especially in places of shading caused by blood vessels or fluids (e.g. for IS/OS border).

The OCT device used for acquiring the image provides automatic segmentation of only 4 retina layers (i.e. ILM, IPL/INL, IS/OS and RPE/Choroid). Thus, comparison of the proposed method to this device is not applicable.

Results of automatic segmentation were calculated separately for images of healthy volunteers and eyes with VMT pathology. Their comparisons for each evaluated method illustrate Figs. 8–12. It is clearly visible that smaller filter masks for both averaging and median filtering provide better results, while bigger masks tend to blur the image. It can be also inferred, that lower values of $\tau$ parameter in the AD method and threshold $\lambda$ for the WSDT method guarantee better performance.

![Illustration of manually segmented layers (A) and results of automatic analysis after noise reduction with selected methods: (B) average filtering (3 × 3 mask), (C) median filtering (3 × 3 mask), (D) anisotropic diffusion ($\kappa = 20$), (E) soft wavelet thresholding ($\lambda = 10$), (F) multiframe wavelet thresholding ($k = 1$). Data is presented in a section of the size 220 × 220 pixels cropped from the center of the image in Fig. 6](image)

![PSNR of retina layers segmentation for healthy and pathological eyes after averaging filtering](image)
Tables 2 and 3 contain PSNR results obtained for the best parameter values for each denoising method. It is worth mentioning, that for borders between hyper-reflective and dark regions (i.e. ILM, IS/OS, and RPE/Choroid) the segmentation results for both groups of patients have higher scores, regardless of the denoising method.

Additionally, a severe VMT pathology, that affected one of the examined patients, has greatly lowered the calculated PSNR value for this group. For this patient the best results in each method gave a score in the range of 36.80 to 38.14 dB, while mean results for other patients in this group were in the range of 43.78 to 44.42 dB. For all patients with the VMT pathology the erroneous segmentation occurred in the area of pathology, and the biggest error was 49 px. This error may be caused by layers irregularities in the area of pathology, as the segmentation algorithm assumes smoothness of the layers borders. This continuity characteristic is expected by the experts and is noticeable in their manual annotations.

### Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>AVG</th>
<th>MED</th>
<th>AD</th>
<th>DSWT</th>
<th>MWT</th>
</tr>
</thead>
<tbody>
<tr>
<td>All layers</td>
<td>37.40</td>
<td>37.40</td>
<td>37.75</td>
<td>37.70</td>
<td>37.88</td>
</tr>
<tr>
<td>ILM</td>
<td>49.28</td>
<td>49.34</td>
<td>48.54</td>
<td>48.56</td>
<td>49.23</td>
</tr>
<tr>
<td>NFL/GCL</td>
<td>38.35</td>
<td>38.24</td>
<td>40.12</td>
<td>40.14</td>
<td>39.11</td>
</tr>
<tr>
<td>IPL/INL</td>
<td>35.56</td>
<td>35.50</td>
<td>35.84</td>
<td>35.80</td>
<td>35.94</td>
</tr>
<tr>
<td>INL/OPL</td>
<td>33.82</td>
<td>33.79</td>
<td>33.90</td>
<td>33.85</td>
<td>34.17</td>
</tr>
<tr>
<td>OPL/ONL</td>
<td>34.87</td>
<td>34.98</td>
<td>35.02</td>
<td>34.91</td>
<td>35.45</td>
</tr>
<tr>
<td>IS/OS</td>
<td>41.64</td>
<td>41.84</td>
<td>42.35</td>
<td>42.38</td>
<td>42.24</td>
</tr>
<tr>
<td>RPE/Choroid</td>
<td>45.67</td>
<td>45.53</td>
<td>46.52</td>
<td>46.47</td>
<td>46.20</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Method</th>
<th>AVG</th>
<th>MED</th>
<th>AD</th>
<th>DSWT</th>
<th>MWT</th>
</tr>
</thead>
<tbody>
<tr>
<td>All layers</td>
<td>43.58</td>
<td>43.29</td>
<td>44.60</td>
<td>44.56</td>
<td>43.15</td>
</tr>
<tr>
<td>ILM</td>
<td>55.66</td>
<td>55.42</td>
<td>55.45</td>
<td>55.55</td>
<td>54.49</td>
</tr>
<tr>
<td>NFL/GCL</td>
<td>40.09</td>
<td>39.95</td>
<td>42.74</td>
<td>42.72</td>
<td>40.84</td>
</tr>
<tr>
<td>IPL/INL</td>
<td>44.31</td>
<td>43.68</td>
<td>44.26</td>
<td>44.16</td>
<td>42.62</td>
</tr>
<tr>
<td>INL/OPL</td>
<td>43.44</td>
<td>42.54</td>
<td>41.79</td>
<td>41.68</td>
<td>40.58</td>
</tr>
<tr>
<td>OPL/ONL</td>
<td>44.62</td>
<td>43.95</td>
<td>44.13</td>
<td>44.04</td>
<td>42.31</td>
</tr>
<tr>
<td>IS/OS</td>
<td>43.08</td>
<td>43.44</td>
<td>48.76</td>
<td>48.68</td>
<td>46.14</td>
</tr>
<tr>
<td>RPE/Choroid</td>
<td>43.91</td>
<td>44.01</td>
<td>45.11</td>
<td>45.34</td>
<td>45.51</td>
</tr>
</tbody>
</table>

### 5. Conclusions

Manufacturers of OCT devices constantly try to overcome the problem of low quality of the acquired images. The denoising methods such as anisotropic diffusion and wavelet thresholding allow for better retina segmentation for both tested groups of patients. Additionally, in case of the VMT pathology we were able to improve accuracy by using the multiframe wavelet thresholding algorithm. We have observed that this approach did not provide significant improvement for images of healthy retinas.

Our experiments confirmed that the proposed method of the multiframe wavelet thresholding improved segmentation accuracy for OCT images of pathological tissues, although change of the noise reduction parameter did not influence the segmentation process.
For healthy data also soft wavelet thresholding of single B-scans and anisotropic diffusion methods are worth exploring, although for these methods it is worth setting the parameter value as small as possible.

The proposed solution can be applied to volumetric datasets to aid during the diagnostic procedure, since precise segmentation allows for early detection of morphological changes and conduction of the thorough assessment of the pathology stage. Thanks to that, it could be possible to select and perform the proper therapy treatment.

Acknowledgements. This work was prepared within the PRELUDIUM CADOCT Project number 2014/15/N/ST6/00710 founded by the National Science Centre in Poland.

REFERENCES


