



# Denoising in optical coherence tomography volumes for improved 3D visualization

LJUBOMIR JOVANOV\*  AND WILFRIED PHILIPS

*imec-IPi, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium*

\*[ljubomir.jovanov@ugent.be](mailto:ljubomir.jovanov@ugent.be)

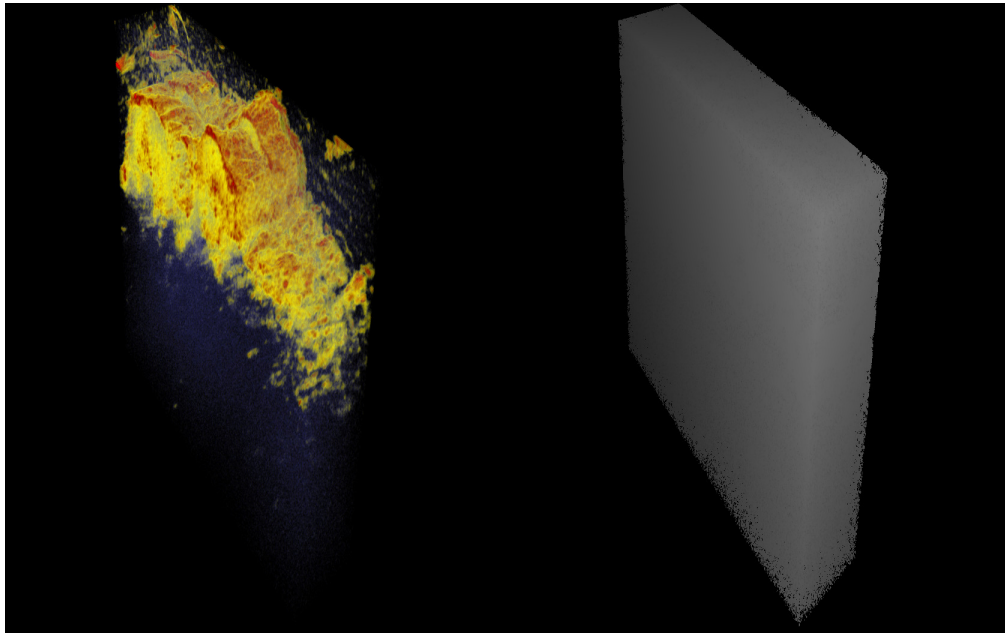
**Abstract:** Optical coherence tomography (OCT) has already become one of the most important diagnostic tools in different fields of medicine, as well as in various industrial applications. The most important characteristic of OCT is its high resolution, both in depth and the transverse direction. Together with the information on the tissue density, OCT offers highly precise information on tissue geometry. However, the detectability of small and low-intensity features in OCT scans is limited by the presence of speckle noise. In this paper we present a new volumetric method for noise removal in OCT volumes, which aims at improving the quality of rendered 3D volumes. In order to remove noise uniformly, while preserving important details, the proposed algorithm simultaneously observes the estimated amounts of noise and the sharpness measure, and iteratively enhances the volume until it reaches the required quality. We evaluate the proposed method using four quality measures as well as visually, by evaluating the visualization of OCT volumes on an auto-stereoscopic 3D screen. The results show that the proposed method outperforms reference methods both visually and in terms of objective measures.

© 2024 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

## 1. Introduction

Optical coherence tomography (OCT) is one of the most recent and the most promising medical imaging techniques. One of the first successful methods for two dimensional OCT imaging of the human eye was presented by Fercher [1]. A significantly improved OCT imaging method, capable of creating cross-sectional images with micrometer resolution was presented in the paper of Huang et al. [2]. The most important applications of OCT are in ophthalmology, cardiology - for diagnosis of coronary artery disease, and for analysis of multilayer structures (e.g. in paintings).

Thanks to its remarkable depth and transversal resolution, OCT enables the creation of high resolution 3D volumetric images, which facilitates precise localization of structures inside the tissue being investigated. By using volumetric rendering methods, the complete 3D structure can be recovered from 2D OCT slices and visualized on a 3D display. Such visualization can be observed as an additional source of information which may improve human ability to localize features inside the volume. However, the effective resolution of OCT scans is limited by speckle noise, which reduces the detectability of small and low intensity features in each individual slice. Moreover, speckle noise degrades the quality of the depth maps created from the OCT volume, and consequently, their visualization on an autostereoscopic display. In particular, here we refer to autostereoscopic displays which take texture and depth as an input. As a practical example, Fig. 1 shows a noisy OCT volume on the left, and the corresponding noisy depth map on the right – the speckle noise has completely masked the features of interest in the volume, resulting in a depth map shaped similarly to the bounding box of the volume. The right image of Fig. 1 also shows that noise affects the texture of the volume, but to a much lesser extent. Although it is possible to obtain some 3D information from a noisy OCT scan by choosing different rendering parameters, often the fine details of the texture will be severely distorted. To address these problems and significantly improve the quality of OCT images and volumes, noise removal techniques can be applied prior to visualization. Numerous OCT denoising methods have been proposed in the last two decades, both in single and multiresolution domains.



**Fig. 1.** Data inputs needed for multiview autostereoscopic display: Noisy texture of the OCT volume (left) and the corresponding noisy depth map rendered using linear opacity values (right). Both the texture image and the depth map are required to visualize the volume on an autostereoscopic display. Speckle noise has completely masked the features of interest in the depth map.

The first group of OCT denoising methods are general despeckling methods, such as the algorithms of Lee and Kuan et al. [3,4], which are based on the use of the second order statistics, optimization of the minimum-mean squared error criterion, and a multiplicative speckle model. In the case of additive noise, the method of Lee [3] is based on the minimum mean-squared error estimator in its simplest form. This method can also be applied to multiplicative noise by employing a statistical optimal linear approximation. The method of Kuan et al. [4] extends the method of Lee [3] to the non-stationary noise case. To develop this extension, Lee derives a minimum mean-squared error (MMSE) estimate as a conditional mean of the noise-free signal, given the noisy estimate. The methods of Lee and Kuan et al. [3,4] are computationally efficient, since both methods are non-iterative and operate in the pixel domain, without applying signal transformations. The main drawback of both methods is that their quality performance lags behind more recent denoising methods.

One of the best performing single resolution denoising methods for OCT images is the Rotating Kernel Transform (RKT) filter by Rogowska et al. [5]. This method performs filtering by using a set of oriented kernels for each pixel, and selects the optimal output. Each of the filtering kernels is line shaped and rotated in small steps around  $360^\circ$ . Although this method removes noise quite efficiently, it does not take noise characteristics of OCT scans into account, but rather performs classical convolution with a set of predefined, differently oriented kernels. This method also introduces angular pattern artifacts due to the use of differently oriented kernels.

One of the first methods which takes into account characteristics of OCT images was presented in the paper of Adler et al. [6]. The spatially adaptive wavelet thresholding applied in this method is based on the method presented by Chang et al. [7]. Unlike earlier wavelet thresholding methods, the method from [7] utilizes a spatially variable threshold for better adaptation to the

spatially changing statistics of images. In this paper, wavelet coefficients are modeled as a random variable obeying a generalized Gaussian distribution. To estimate the unknown parameter of this model, context modeling is used to derive a threshold. This method achieves much better performance than the pixel domain methods reviewed in the previous paragraph. Compared to those, this method has higher computational complexity, due to the wavelet transform and sorting operations.

To further improve the performance of wavelet-based methods, Mayer et al. [8] have employed information from multiple slices for volumetric denoising, based on weighted averaging of wavelet coefficients from neighboring OCT slices. In this method, two different weights are defined: a significance and a correlation weights. Significance weights are related to a local noise variance estimation, calculated using detail coefficients. To calculate correlation weights, the authors rely on the approximation coefficients to provide information on whether a certain structure exists in the current frame or not. Thanks to the use of redundancy present in neighboring slices, this method efficiently removes noise from the slices. As in the previous wavelet-based method, this method has a higher computational complexity. Due to the use of multiple frames, in this case computation complexity is even higher, in addition to greater memory requirements.

Different principles were used in the method of Fang et al. [9], where sparse dictionary learning was employed on reference low noise frames to reduce speckle in noisy OCT volumes. Thanks to the customized scanning ability, a certain number of higher signal-to-noise ratio slices (SNR) is captured, which are subsequently used to learn a sparse representation dictionary using the proposed MSBTD algorithm. Based on the sparse dictionary, denoised slices are obtained by finding a linear combination of dictionary atoms by minimizing a cost function. The cost function consists from  $l_2$  norm of the difference between noisy slice and the linear combination of atoms and  $l_0$  norms of the coefficients of employed dictionary atoms. This method efficiently removes noise, while improving contrast and re-introducing lost details in the slices. The main drawback of this method is even larger computational complexity and the requirement for high SNR dataset for dictionary learning.

The method of Wang et al. [10] et al. performs noise estimation and OCT volume speckle reduction in a sparse domain using adaptive shrinkage. This method combines non-local denoising and thresholding in a sparse domain. While removing noise efficiently, this method is computationally intensive. Similarly, in the method of Li et al. [11], a statistical model was employed for denoising of OCT images. In this method, denoising is formulated as a convex optimization problem. To solve this problem, a new fast iterative algorithm is proposed, based on a linearized alternating direction method of multipliers (ADMM). In the approach of Devalla et al. [12], a multi-scale U-Net network based approach was proposed, based on earlier DRUNET approach from the same authors [13]. The authors propose a deep neural network, based on U-Net, residual learning, dilated convolutions and multi-scale hierarchical feature extraction. Thanks to the U-Net and its skip connections, the network learns both the local and contextual information.

The method proposed by Rico-Jimenez et al. [14] performs denoising based on deep neural network trained on OCT data. This algorithm relies on the self-fusion network presented by Devalla et al. in [12], which utilizes similarity between adjacent slices, to fuse 3 frames and achieve improved PSNR. To further reduce the processing time, the authors only use self-fused frames for training of U-Net. The parameters of the U-Net are optimized using the Adam optimization algorithm on a training set consisting of 9 optic nerve head (ONH) volumes. This method is capable of real-time denoising of OCT volumes. However, methods from [12,14] suffer from slight image-smoothing due to the convolution and rigid registration. Moreover, this method requires robust training data to avoid introduction of blur. Finally, in the paper of Huang et al. [15], the authors propose a groundtruth-free deep learning approach for real-time noise reduction. This method is based on a modified asymmetric convolution-SRResNet (AC-SRResNet), and

relies on an asymmetric convolution block for feature extraction, which improves the accuracy of deep learning. To optimize the network, the authors used Adam optimizer. The performance of this method lags behind the method presented in [10]. In this work, we present an OCT denoising algorithm specifically aimed at improving 3D volume visualization. The algorithm automatically detects regions of interest and locally adapts to the content.

This paper is organized as follows: Section 2 describes the proposed method and Section 3 contains the evaluation of the proposed method. Conclusions and the future work are presented in Section 4.

## 2. Proposed method

In this section we describe the proposed OCT volume denoising algorithm, which exploits redundancy in neighboring slices to perform noise reduction. The main goal of the proposed algorithm is to enable high visual quality of rendered 3D OCT volumes. The motivation for this lies in the fact that 3D displays coupled with 3D input devices like Kinect and Leap Motion sensors [16,17], could give radiologists better insight into tissues, as discussed in the paper of Gargesha et al. [18]. Moreover, when interacting with images using 3D input devices, better visual quality would facilitate high precision manual navigation through the regions of interest. However, the main obstacle to achieving better visibility and better rendered volume quality is *noise*. To improve the quality of rendered OCT volumes and enable interactivity, we perform 3D denoising along the trajectories of volume rendering rays and adapt the intensity of denoising to the local characteristics of the tissue. Our goal is to enable the propagation of rendering rays through regions with lower tissue density, while preserving important features in regions which contain tissues of interest. In this paper we propose a novel volumetric filter which adapts to the local tissue density and local noise standard deviation in order to remove noise uniformly in the whole volume. By applying these denoising principles, the processed volumes exhibit improved objective quality, measured as equivalent number of looks (ENL), edge preservation (EP), and contrast to noise ratio (CNR) as defined in Section 3.

The main contributions of the proposed method are:

- Improvement of 3D volume visualization by performing denoising along the rays used by the volume rendering algorithm
- Locally adaptive noise reduction based on joint optimization of the quality measures such as EP and SNR
- Tunable noise reduction parameters to facilitate volume visualization
- Automatic detection of regions of interest in the image for optimization and local adaptation

We explain the proposed method in detail in the following subsections.

### 2.1. Robustness of the volume rendering algorithms in the presence of noise

In this subsection we concisely describe volume rendering algorithm used to extract the geometry and the texture from OCT scans, because we incorporate its basic principles into the proposed denoising algorithm. The main objective of the proposed volume denoising algorithm is to achieve the highest quality of rendered depth maps possible, for 3D visualization on an auto-stereoscopic (3D) display and to enable more precise and comfortable interaction and manual annotation. Therefore, by incorporating some aspects of volume rendering into the denoising algorithm, we are able to achieve higher quality rendered depth images.

We perform volume rendering using the GPU-assisted algorithm implemented in the VTK software library created by Schroeder et al. [19], based on the volume ray casting method, originally proposed by Kajiya [20]. Series of b-scans are observed as a volume containing tissue

density for each voxel in 3D space. B-scans are brightness scans created by scanning the tissue laterally and then stacking them into a transversal image. In order to achieve better differentiation of tissues in the volume, we define the opacity function  $\alpha(x, y, z)$  and the color function  $c(x, y, z)$ , which map the tissue density into opacity values, and color values respectively.

The color mapping function maps the voxel value into triplets of RGB values, which facilitates differentiation of regions with different tissue density value, as they are coloured differently after applying this function. By applying the opacity function, it is possible to determine the visibility of objects of certain density. The opacity function is especially important for geometry rendering, since by carefully choosing this function's parameters, it is possible to improve the visibility of small features of a certain density in the volume's depth map.

Based on these two values for each voxel, the volume ray casting algorithm reconstructs the texture and the geometry. The geometry of the volume is rendered using tissue densities and the opacity transfer function. We calculate the gradient after applying the opacity function and determine the orientation of the surface based on the value of the gradient. Similarly as in the texture rendering step, we propagate an imaginary ray through the volume, from the front to the back. The surface is detected when the gradient value of the opacity function is larger than the predefined threshold value.

From the position of the high opacity gradient, we can determine the geometry of the front surface, i.e. its depth map. Since the propagation of the ray is stopped once the surface is detected, computational load is significantly reduced, as we do not perform calculations for the whole length of the ray. It is also possible to render multiple layers of front surfaces. Noise affects the volume rendering process twofold; by adding noise to the rendered texture of the volume and by affecting the rendered depth map of the volume. As we can see in Fig. 1, small details in the regions of interest in the texture image are masked by noise, especially in the lower part of the volume. Rendering of the volume depth map is affected more severely than the texture, which can be seen in the left part of Fig. 1.

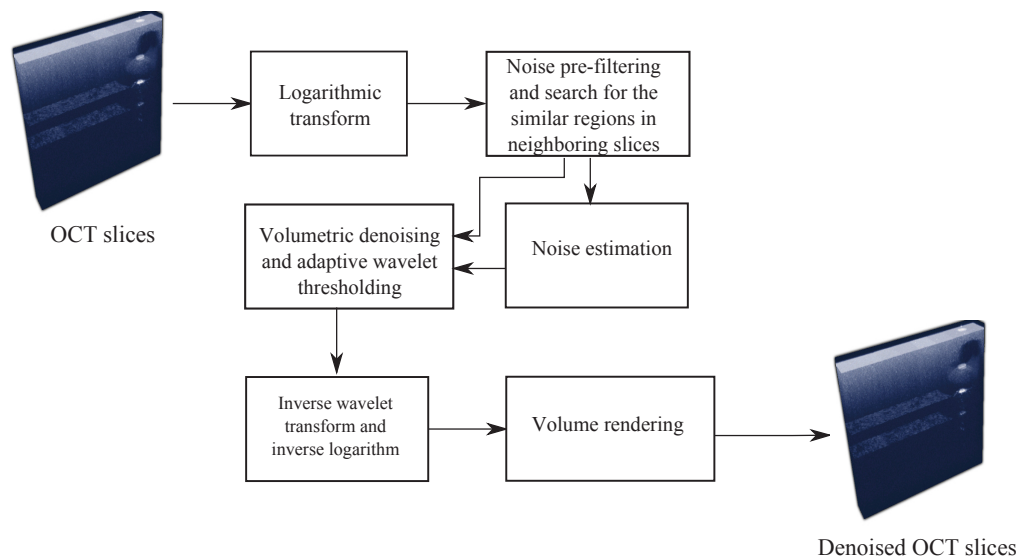
Since noise prevents the propagation of the ray through the tissue, starting at the boundaries of the volume, the resulting depth map has the shape of the bounding box around the tissue. Moreover, the depth rendering step requires calculating of the gradient of volume densities, which results in even higher noise levels. On the other side, integration performed in the texture rendering step slightly reduces the noise level in the rendered texture. With these two findings in mind, in order to improve the quality of the rendering, the proposed denoising algorithm should remove the noise consistently, in order to free the path of the rendering rays, all the way to the surface of interest in the volume.

## 2.2. *Main principles of the proposed algorithm*

Modifying the opacity function only has a limited effect on improving the quality of the texture and the depth; in order to improve the visual quality of the rendered volume, denoising of OCT B-scans becomes necessary. In the ideal case, denoising should be performed along the currently rendered viewing direction, to prevent ray scattering and occlusions due to the noise.

Without the loss of generality and to improve the clarity, here we will consider volume denoising along orthogonal directions, i.e. x,y and z axis. The denoising principles described here can be applied in any direction selected in the visualization software. The proposed denoising algorithm takes into account some specific characteristics of OCT images, such as variable speckle contrast, due to the propagation of the laser beam through the tissue. This results in a more uniform denoising throughout the whole volume. Structures inside the volume cannot be always observed as rigid bodies, so therefore it is not possible to use any global alignment techniques to register neighboring slices. Therefore, it is necessary to follow tissue structures locally through the volume.

The main principles of the proposed method are illustrated in Fig. 2. The first step converts multiplicative noise into approximately white Gaussian noise. In the second step we perform wavelet transformation of the volume. Then, we perform volumetric denoising along the viewing direction. An integral part of this step is the estimation of regions of interest, noise characteristics, and direction of the free space, and lastly noise reduction. Finally, slices are transformed into the base domain and visualized. The main idea of the proposed method is to perform noise suppression in such a way that the visibility of regions of interest is improved both in the 2D slices and in the 3D volume rendering. These two goals are incorporated in the algorithm by optimizing standard OCT image quality measures and integrating volume rendering principles into the denoising process. After applying multi-resolution decomposition, each slice is represented with a low pass image and a number of high pass bands, representing high frequency content. The low pass component image typically contains slowly varying correlated noise, which is difficult to suppress. These artefacts make the detection of background regions more difficult.



**Fig. 2.** Block diagram of the proposed method, with its main components and data flows. Detailed description of the diagram can be found in Section 2.2

In order to perform ROI and background detection, and enable ray propagation, we perform volumetric filtering of the low pass band along the direction of the rendering ray. This filtering is performed in a non-iterative manner with constant variance, as explained in the next section. Analogous to the volume rendering, we start denoising from the front of the volume and continue towards the interior, along the selected direction. In order to improve the rendering quality, we perform classification onto regions of interest and background regions. For each identified region of interest, we calculate the standard quality measures mentioned in the previous section and defined in Section 3, while SNR is calculated in the background.

### 2.2.1. Algorithm description

As previously stated, the main idea of the proposed method is to improve the propagation of rendering rays through the areas with low density and to improve the quality of regions of interest. To achieve this we rely on the redundancy contained in multiple slices. We operate on a buffer containing  $N_B = 7$  slices in the direction which is selected in the visualization software. For the first  $\lfloor \frac{N_B}{2} \rfloor$  slices, the buffer is mirrored.

**Algorithm 1:** The proposed algorithm pseudo-code

---

```

Result: Denoised volume
Fetch buffer of noisy slices;
for each_slice do
  for each_pixel do
    while ( $\sigma_{res} > \sigma_D$  or  $EP > EP_D$ ) do
      Find the best matching patch currently;
      Add the current patch to the input set  $\mathcal{N}_p$ ;
      Estimate  $\hat{x}$  from the current patch set using Equation 4;
      Calculate current values of EP and  $\sigma$ ;
    end
  end
end

```

---

In order to exploit the power of multi-slice methods, we perform a search for the most similar blocks in the neighboring slices, taking into account the tissue type in the neighbourhood of the current location and the offsets of surrounding patches. For the current patch at the location  $(i, j, t)$  we find the most similar patches by minimizing the following functional:

$$\arg \min_{p_{k,l}} \left[ \sum_{k,l \in \mathcal{N}_{i,j}} |p_{i,j} - p_{k,l}| + \alpha \cdot S + \beta \cdot dR \right] \quad (1)$$

where  $p_{i,j}$  is the current patch,  $p_{k,l}$  is the patch from the neighborhood  $\mathcal{N}_{i,j}$ ,  $S$  is the type of tissue to which the patch  $p_{j,k}$  belongs,  $\alpha, \beta$  are empirically determined constants, and  $dR$  is the spatial offset of the patch  $p_{k,l}$  relative to the patch  $p_{i,j}$ . The first factor from the regularization component of the Eq. (1),  $S$ , has a value of 0, when the center of the current patch  $p_{i,j}$  belongs to the same tissue type as the patch from the search region  $p_{k,l}$ , and 1 otherwise. This factor is scaled by a constant  $\alpha$  to additionally reduce mixing, due to the averaging of different classes of patches. The second regularization factor,  $dR$ , penalizes the distance of the patch in the search area from the current patch location, and is calculated as follows:

$$dR = \sum_{l \in \mathcal{N}_{i,j}} \sqrt{(dx(l) - med(dx))^2 + (dy(l) - med(dy))^2} \quad (2)$$

where  $dx$  and  $dy$  are displacements of the most similar patches from the central patch and  $med$  is median of all displacements within the neighbourhood  $\mathcal{N}$ . Once the best matching patch that minimizes Eq. (2) is found, the next best patch is found by minimizing the same equation, and the previously found good matches are removed from the search space. We proceed with the patch selection step until the denoising criterion is met, as illustrated in the pseudo-code in Algorithm 1.

Found patches are further processed, in order to maximally suppress noise and enhance details necessary for visualization and successful diagnosis. Selected patches are enhanced simultaneously with the search. Let us denote noisy sequence after logarithmic transform as:

$$y = x + n \quad (3)$$

where  $x$  denotes the unknown, noise-free volume and  $n$  denotes the noise. Our goal is to reconstruct the unknown noise free volume  $\hat{x}$ , which is as close as possible to the hypothetical noise-free image and satisfies the criteria set in the optimization process. To estimate  $x$ , we apply the following steps. For each newly found patch, we calculate a preliminary result by applying

the following operator in each wavelet band:

$$\hat{x} = F\left(\sum_{l \in \mathcal{P}} w(y_l) y_l\right) \quad (4)$$

where  $y_l$  is the current patch from the set,  $w(y_l)$  is the weight function for the current patch in the set, and  $F(y_l)$  is a function applied to each pixel of the patch depending on its type and segment to which it belongs. New patches are added to this sum until the convergence criterion of filtering is met. Weights  $w(y_l)$  are defined as:

$$w(y_l) = \frac{e^{-\frac{\|y_l - y_c\|^2}{2\sigma_l^2}} \cdot s_{y_l}}{\sum_{n \in \mathcal{P}} e^{-\frac{\|y_n - y_c\|^2}{2\sigma_l^2}} \cdot s_{y_n}} \quad (5)$$

where  $y_c$  is the patch at the central location,  $y_l$  is the patch from the set of the most similar patches  $\mathcal{P}$ , and  $s_{y_l}$  is an indicator variable which switches different processing on, in the case of pixels from different segments. We calculate the norm of difference between the central and the currently processed patch, by pixel-wise subtraction of values and summing their absolute differences. The proposed volumetric filtering is executed until denoising criteria are met. The first criteria that we propose is related to the residual noise. We perform searching for new candidate patches until the noise variance drops under a certain threshold. In each step we calculate the noise variance using the Eq. (6).

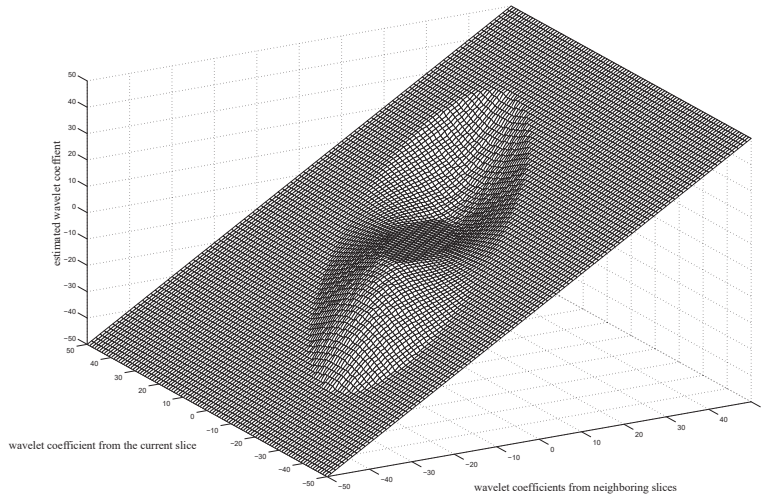
$$\sigma_{y_c}^2 = \sum_{n \in \mathcal{P}} w(y_l)^2 \sigma_{y_l}^2 \quad (6)$$

where  $\sigma_{y_l}^2$  is locally estimated noise variance at the position  $l$ ,  $\sigma_{y_c}^2$  noise variance at the currently denoised position, and  $w(y_l)$  weight of the pixel at the position  $l$ . If the absolute difference between noise variances in two successive iterations becomes smaller than threshold, we stop the filtering.

At the other hand, excessive averaging in the areas with features of interest might lead to excessive blur. In order to prevent this, we calculate the value of the EP, defined in 3.2, and stop filtering if its value drops below the minimum desired value, which is checked by comparing values in successive iterations. Thresholds for both stopping criteria can be set from the GUI of the visualization algorithms to meet the preferences of the user or automatically determined for the best detail visibility. Finally we define a shrinkage function for per-slice spatial filtering, which suppress the remaining noise in the patches:

$$\hat{x} = F(\mathcal{I}_{z-1}, \mathcal{I}_z, y_l, y_{l-1}) \quad (7)$$

Relying on the first, iterative weighted averaging step, noise variance is already significantly reduced, which relaxes requirements for the spatial filtering within one slice. The robustness of the shrinkage operation is additionally improved by introducing oriented spatially registered indicator functions from the previous slice and next slice, in addition to spatial indicators from the current slice. The shrinkage function boosts the high frequency components in order to improve the values of quality indicators. The low pass band of each slice is also filtered temporally using the same operator, but with the constant variance. This way, low frequency artefacts are removed and ray propagation is improved. At each point noise is estimated locally using diagonal wavelet band. Finally, we perform spatial filtering of the residual noise, by applying wavelet shrinkage function shown in Fig. 3. This shrinkage function is obtained by combining two membership functions: one related to the value of the current wavelet coefficient and the other for the evidence that edge is present.



**Fig. 3.** Wavelet shrinkage function for spatial filtering

These two functions are shown in Fig. 4. The main idea of the proposed filtering is to leave the current wavelet coefficient unchanged or boosted, if coefficients in its neighborhood are sufficiently large and the coefficient itself is large. Otherwise if neighboring coefficients have low values, it is highly likely that the current, high coefficient represents noise and that it should be suppressed. The spatial denoising method presented in this paper is based on our earlier method presented in [21]. For the sake of conciseness here we present only main differences compared to [21]. The wavelet coefficient at the spatial position  $k$  is estimated as:

$$\hat{y}_k = \tilde{\gamma}(\mathbf{w}_k, \mathbf{z}_k) \cdot w_k \tag{8}$$

where  $\mathbf{z}_k = \begin{bmatrix} z_k^C & z_k^N \end{bmatrix}^T$  represents a vector consisting of local spatial activity indicators (LAI) at the position  $k$  in the current and neighbouring slices,  $\mathbf{w}_k = \begin{bmatrix} w_k^C & w_k^N \end{bmatrix}^T$  is a vector containing the values of the wavelet coefficients in the current and neighbouring slices at the location  $k$  and  $\tilde{\gamma}$  denotes a wavelet estimator function. Membership functions are dependent on the noise covariance matrix  $C_n$ :

$$\mu_{\tilde{w}}(w_k) = \mu_w(\|C_n^{-\frac{1}{2}} w_k\|) \tag{9}$$

Based on Eq. (9), we obtain the expression for membership functions:

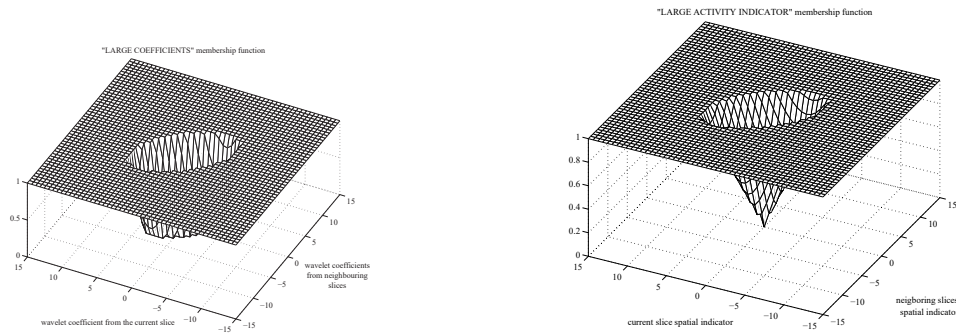
$$\mu_{\tilde{w}}(\mathbf{w}_k) = \mu_w(C_n^{-\frac{1}{2}} \mathbf{w}_k) \quad \text{and} \quad \mu_{\tilde{z}}(\mathbf{z}_k) = \mu_w(C_n^{-\frac{1}{2}} \mathbf{z}_k). \tag{10}$$

In the case of the positive semi-definite matrix  $C_n$ , equation:

$$\|C_n^{-1/2} \mathbf{w}_k\|^2 = \mathbf{w}_k^T C_n^{-1} \mathbf{w}_k = T^2 \tag{11}$$

represents an ellipsoid in a two dimensional space. If the signal at the current location  $\mathbf{w}_k$  corresponds to a position inside the ellipsoid, the signal of interest (SOI) cannot be detected at the location  $k$ . Otherwise, the signal of interest is present at the given location.

The parameter  $T$  from the Eq. (11) determines the sensitivity of the SOI detector. Typical value for the threshold is  $T = 1$ . We calculate the value of the membership function for the current value of the vector variable  $\mathbf{w}_k$ . Similarly as in [21] we calculate the values of signal presence indicators relying on the anisotropic spatial indicators, calculated based on the values of the current and spatially aligned neighboring slices.



(a) Membership function describing prominence of the edge

(b) Membership function describing neighborhood with strong edges

**Fig. 4.** Membership functions

### 3. Experimental results

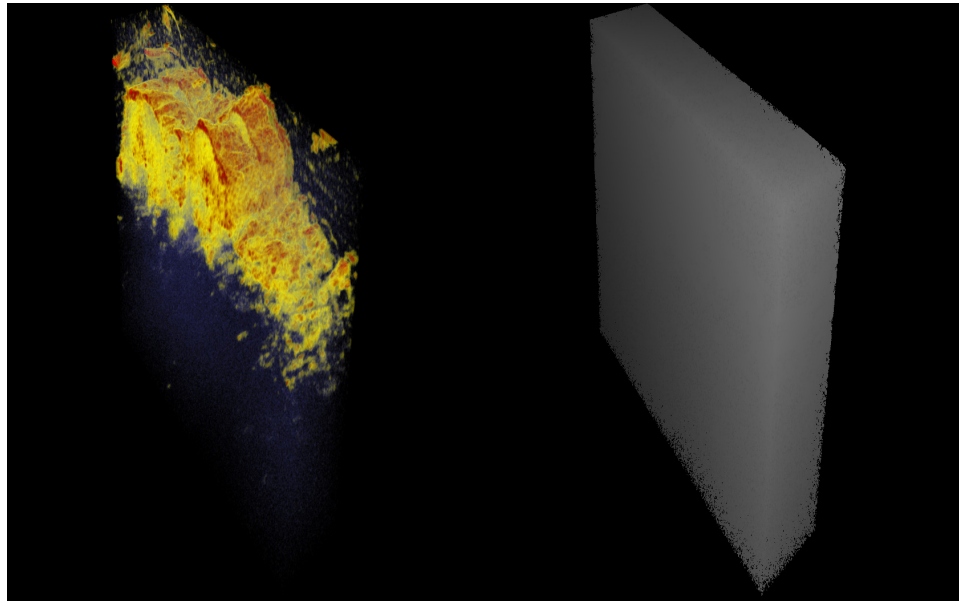
In this section we evaluate the performance of the proposed method on live tissues scans acquired using the SkinDex 300 OCT scanner, manufactured by ISIS Optronics, Mannheim Germany. This OCT scanner belongs to the class of time-domain OCT devices. It is based on a Michelson interferometer, employing four optical arms. During the scanning process, the optical path in the reference arm remains constant, while only the distal tip of a fiber in the sample arm keeps moving. The central wavelength of the source is  $\lambda_0=1300\text{nm}$  and the bandwidth is  $\Delta\lambda=100\text{nm}$ . For further information about this OCT scanner used, we refer the interested reader to the following paper [22]. OCT scans were produced using the following settings: dimensions of each slice were  $904\times 1005\mu\text{m}$  in axial and transverse directions respectively, the distance between slices was  $1.1\mu\text{m}$ , while the resolution of the slice was  $556\times 500$  pixels. Volumes typically contained 126 slices. In the following subsections we will first show visual results represented by 3D renders and B-scans of OCT volumes denoised using the proposed method, followed by quantitative comparisons.

#### 3.1. Visual evaluation of the proposed algorithm

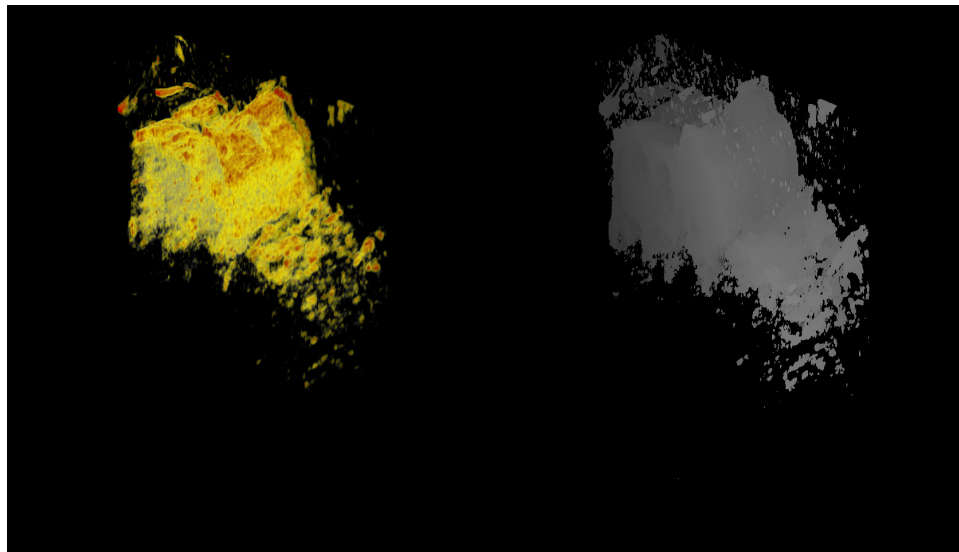
Volume rendering of noisy OCT data requires careful adjustment of rendering parameters in order to achieve acceptable results. However, in practice, it would be desirable to use one universal set of rendering parameters to obtain high rendered volume quality without needing additional adjustments. If we observe volume rendering for a standard set of rendering parameters, as shown in in Fig. 1, we can see that the rendered geometry, as well as the texture, become unusable due to the noise, since it is only possible to see the texture and the front surface of the volume. After applying the proposed volume denoising method, the resulting volume rendering is significantly improved and it becomes possible to distinguish details in the tissue, as can be seen in Fig. 5. Volume denoising allows rendering rays to propagate through the volume, diffuse and reflect at the regions of interest, thereby generating plausible rendering. We have also provided a supplementary material associated with this article at the website of the publisher, which supports the conclusions of our visual experiments (see [Visualization 1](#)).

#### 3.2. Evaluation of the proposed algorithm using quantitative performance measures

The main goal of the proposed method is to achieve the highest quality of rendered 3D geometry. Although a certain estimation of the quality can be obtained visually, it is necessary to quantify these results by using objective quality measures. In this subsection, we quantitatively compare



(a)



(b)

**Fig. 5.** Visual comparison between a noisy OCT volume and the result of the proposed method: a) Noisy texture of the OCT volume (left) and the corresponding noisy depth map rendered using linear opacity values (right), b) Denoised texture of the OCT volume (left) and the corresponding denoised depth map rendered using linear opacity values (right)(see [Visualization 1](#)). The proposed volume denoising method significantly improves visibility of tissue details and the depth map, both required to visualize the volume on an autostereoscopic display.

the performance of the proposed algorithm and recent state-of-the-art multi-frame OCT denoising

algorithms presented in [8] and [9] in terms of well established no-reference quality metrics. For this purpose we employ the following quality metrics:

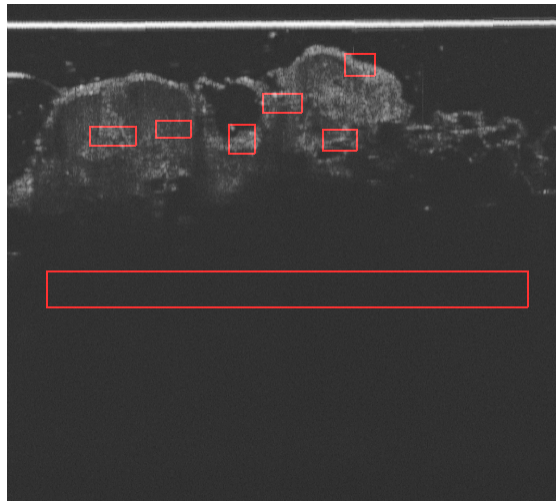
- average signal-to-noise ratio:  $SNR = 10 \log \frac{\max(A^2)}{\sigma^2}$ ,
- contrast to noise ratio:  $CNR = \frac{1}{R} \sum_{r=1}^R \frac{|\mu_r - \mu_b|}{\sqrt{\sigma_r^2 + \sigma_b^2}}$ ,
- equivalent number of looks:  $ENL = \frac{1}{H} \sum_{h=1}^H \frac{\mu_h^2}{\sigma_h^2}$ ,
- edge preservation:  $EP = \frac{\Sigma(\nabla^2 M - \overline{\nabla^2 M})(\nabla^2 A - \overline{\nabla^2 A})}{\sqrt{\Sigma(\nabla^2 M - \overline{\nabla^2 M})^2 \cdot \Sigma(\nabla^2 A - \overline{\nabla^2 A})^2}}$ .

where  $A$  represents the maximal value in the processed image;  $\sigma^2$  represents the variance of the background region;  $\mu_b$  and  $\sigma_b^2$  represent the mean value and the variance of the background region;  $\mu_r$  and  $\sigma_r^2$  represent the mean and the variance of the  $r$ -th region of interest;  $\mu_h$  and  $\sigma_h^2$  represent the mean value and the variance of the  $h$ -th region of interest;  $\nabla^2 M$  and  $\nabla^2 A$  represent the Laplacian operator performed on the original and the filtered image respectively;  $\overline{\nabla^2 M}$  and  $\overline{\nabla^2 A}$  represent the mean value of a  $3 \times 3$  neighbourhood around the pixel location of  $\nabla^2 M$  and  $\nabla^2 A$  respectively.

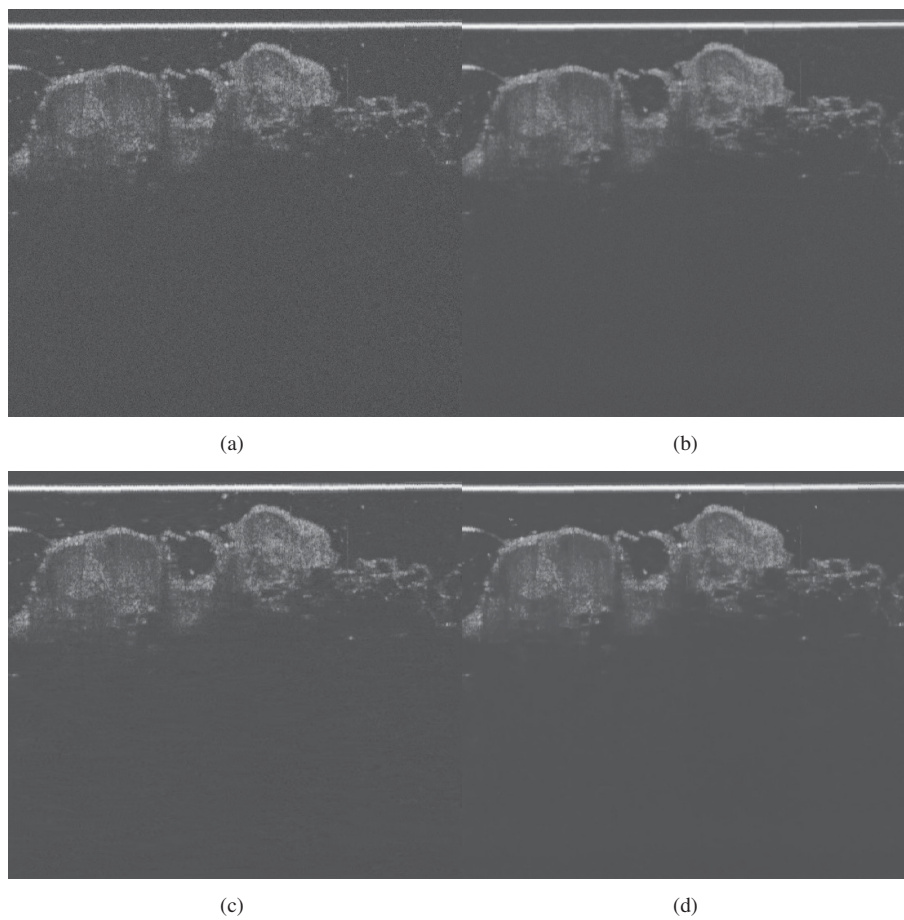
We evaluate the above metrics on a several carefully chosen regions of interest (ROI) which contain details of interest in the tissue, like edges and textured regions shown in Fig. 6. In the ideal case, the selected ROIs should include all types of texture, edges and piece-wise flat regions, representing different types of tissues, to better evaluate the quality of the denoising algorithm. The quality measures defined above, were proposed since noise-free OCT images are not available for direct visual comparison and mean squared error calculation. Each of these quality measures is related to an aspect of the perceived image quality. The global performance of the denoising algorithms is estimated using SNR, since it compares the maximum value of the image and the standard deviation of noise in the background region. CNR characterizes the contrast between the background and the regions of interest. Smoothness of a homogeneous region of interest is characterized using ENL, while EP characterizes the ability of the algorithm to preserve edges.

Results of the qualitative evaluation of the tested method is given in Table 1. The performance of the proposed method compares favourably to the state-of-are-art methods. More precisely, the values of SNR, EP and ENL metrics achieved by the proposed method are the highest among the compared methods, while the value of CNR is slightly lower than the best performing method. On the other hand the best performing method regarding CNR [8] exhibits a loss of fine details and a slightly distorted geometry of the tissue. The algorithm of [9] achieve slightly better preservation of significant details than the proposed method, while at the same time suffering from slight block and low-frequency artefacts. These quantitative results can also be visually confirmed by observing Fig. 7, in the case of the challenging OCT volume containing a large amount of fine details and texture. The proposed method performs equally well on all types of OCT volumes. One of the reasons for this is that we do not employ learning algorithms that might not sufficiently generalize to different types of OCT volumes.

We have conducted statistical analysis testing to determine whether the proposed and the reference methods performed significantly different on the qualitative measures. Due to the non-normal distribution, we have used the non-parametric unpaired Kruskal-Wallis test and



**Fig. 6.** Regions of interest used for the evaluation of the quality measures

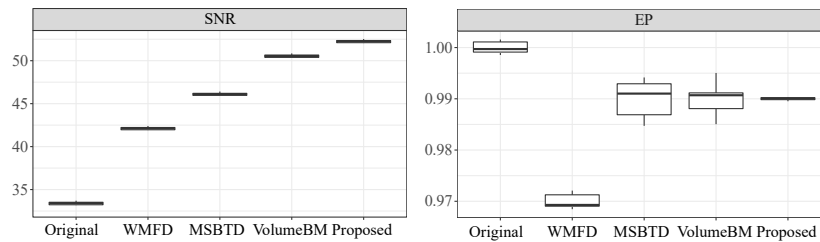


**Fig. 7.** Visual comparison of the results: a.) Noisy slice b.) Result of WMFD [8] c.) Result of MSBTD [9] d.) Result obtained using the proposed method

**Table 1. Quantitative comparison between the proposed and reference methods**

Method	SNR[dB]	CNR[dB]	ENL	EP
Original Noisy	33.34	2.28	333	<b>1</b>
WMFD [8]	42.05	<b>3.54</b>	563	0.97
MSBTD [9]	46.05	2.94	403	0.99
VolumeBM [10]	50.47	2.95	396	0.99
The proposed method	<b>52.18</b>	<b>3.54</b>	<b>605</b>	0.99

conducted pairwise comparisons using Wilcoxon's test with Bonferonni correction. For SNR and EP, there were statistically significant differences between methods including the Original, with  $p < 0.0001$ ,  $p < 0.001$ , respectively. In the case of EP, the proposed method was significantly different from WMFD and the Original, but not from MSBTD and VolumeBM. Sample sizes required to detect significant differences between methods, including the Original, were estimated as 13 for SNR and 13 for EP. Estimates were conducted for the Kruskal-Wallis test, with a target power of 80%, significance level of 0.05, the five groups with the currently measured means, and an assumption of normally distributed samples. Analysis was conducted in R using rstatix version 0.7.0 and sample size estimates were conducted with the *kwsamplesize* function in MultNonParam version 1.3.9. We illustrate the findings of the statistical analysis in Fig. 8, for SNR and EP quality measures. The boxplot shows the consistent results of the proposed methods, with a small sample variance. This also confirms our experimental findings that the proposed method is capable to efficiently denoise a wide variety of different OCT volumes.

**Fig. 8.** Statistics of SNR and EP quality measures

#### 4. Conclusions and future work

In this paper we have presented a new volumetric denoising algorithm. The main application of the proposed method is the improved visualization of volumes on autostereoscopic displays which use texture images and depth as an inputs. We show that the proposed method reduces noise uniformly and therefore significantly improves the quality of the visualization on the autostereoscopic display. Besides visual evaluations, we have also performed comparisons with the reference methods from the literature in terms of objective quality measures and found out that the proposed method compares favorably with the above mentioned methods. In the future, we plan to further optimize the proposed algorithm and perform closer integration with the visualization software.

**Funding.** Interuniversitair Micro-Electronica Centrum VZW.

**Acknowledgments.** The authors would like to thank Asli Kumcu for constructive comments and for proofreading the manuscript.

**Disclosures.** The authors declare no conflicts of interest.

**Data availability.** Data underlying the results presented in this paper are not publicly available at this time, due to privacy reasons, but may be obtained from the authors upon reasonable request.

## References

1. A. F. Fercher, "Ophthalmic interferometry," in *Proceedings of the International Conference on Optics in Life Sciences*, G. von Bally and S. Khanna, eds., 1990, pp. 221–228.
2. D. Huang, E. Swanson, C. Lin, *et al.*, "Optical coherence tomography," *Science* **254**(5035), 1178–1181 (1991).
3. J. S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2**(2), 165–168 (1980).
4. D. T. Kuan, A. A. Sawchuk, T. C. Strand, *et al.*, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-7**(2), 165–177 (1985).
5. J. Rogowska and M. Brezinski, "Evaluation of the adaptive speckle suppression filter for coronary optical coherence tomography imaging," *IEEE Trans. Med. Imaging* **19**(12), 1261–1266 (2000).
6. 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," *Opt. Lett.* **29**(24), 2878–2880 (2004).
7. S. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. on Image Process.* **9**(9), 1522–1531 (2000).
8. M. A. Mayer, A. Borsdorf, M. Wagner, *et al.*, "Wavelet denoising of multiframe optical coherence tomography data," *Biomed. Opt. Express* **3**(3), 572–589 (2012).
9. L. Fang, S. Li, Q. Nie, *et al.*, "Sparsity based denoising of spectral domain optical coherence tomography images," *Biomed. Opt. Express* **3**(5), 927–942 (2012).
10. L. Wang, Z. Meng, X. S. Yao, *et al.*, "Adaptive Speckle Reduction in OCT Volume Data Based on Block-Matching and 3-D Filtering," *IEEE Photonics Technol. Lett.* **24**(20), 1802–1804 (2012).
11. M. Li, R. Idoughi, B. Choudhury, *et al.*, "Statistical model for OCT image denoising," *Biomed. Opt. Express* **8**(9), 3903–3917 (2017).
12. S. K. Devalla, G. Subramanian, T. Pham, *et al.*, "A Deep Learning Approach to Denoise Optical Coherence Tomography Images of the Optic Nerve Head," *Sci. Rep.* **9**(1), 14454 (2019).
13. S. K. Devalla, P. K. Renukanand, B. K. Sreedhar, *et al.*, "DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images," *Biomed. Opt. Express* **9**(7), 3244–3265 (2018).
14. J. Rico-Jimenez, D. Hu, E. Tang, *et al.*, "Real-time OCT image denoising using a self-fusion neural network," *Biomed. Opt. Express* **13**(3), 1398–1409 (2022).
15. Y. Huang, N. Zhang, and Q. Hao, "Real-time noise reduction based on ground truth free deep learning for optical coherence tomography," *Biomed. Opt. Express* **12**(4), 2027–2040 (2021).
16. C. S. Bamji, P. O'Connor, T. Elkhatib, *et al.*, "A 0.13  $\mu\text{m}$  CMOS System-on-Chip for a  $512 \times 424$  Time-of-Flight Image Sensor With Multi-Frequency Photo-Demodulation up to 130 MHz and 2 GS/s ADC," *IEEE J. Solid-State Circuits* **50**(1), 303–319 (2015).
17. A. Sharma, A. Yadav, S. Srivastava, *et al.*, "Analysis of movement and gesture recognition using Leap Motion Controller," *Procedia Comput. Sci.* **132**, 551–556 (2018).
18. M. Gargsha, M. W. Jenkins, A. M. Rollins, *et al.*, "Denoising and 4D visualization of OCT images," *Opt. Express* **16**(16), 12313–12333 (2008).
19. W. Schroeder, K. Martin, and B. Lorensen, "Visualization Toolkit: An Object-Oriented Approach to 3D Graphics," 4th Edition, <https://vtk.org/vtk-textbook/>.
20. J. Kajiya and B. Von Herzen, "Ray tracing volume densities," *SIGGRAPH Comput. Graph.* **18**(3), 165–174 (1984).
21. L. Jovanov, A. Pižurica, and W. Philips, "Fuzzy logic-based approach to wavelet denoising of 3D images produced by time-of-flight cameras," *Opt. Express* **18**(22), 22651–22676 (2010).
22. A. Knüttel, S. Bonev, and W. Knaak, "New method for evaluation of in vivo scattering and refractive index properties obtained with optical coherence tomography," *J. Biomed. Opt.* **9**(2), 265–273 (2004).