Buletinul Ştiinţific al Universităţii Politehnica Timişoara

TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Volume 60(74), Issue 1, 2015

Optical Coherence Tomography Speckle Reduction in the Wavelets Domain

Cristina Stolojescu-Crisan[1] Alexandru Isar[2]

Abstract – This paper proposes a denoising method that associates the Hyperanalytic Wavelet Transform (HWT) with a Maximum A Posteriori (MAP) filter named bishrink. The method is tested on Optical coherence tomography (OCT) images. The experimental results prove that the denoising algorithm can effectively reduce the speckle noise, while preserving the structural and textural features and improves the quality of OCT images.

Keywords: — Denoising, Hyperanalytic Wavelet Transform, Optical coherence tomography, bishrink filter, speckle noise

I. INTRODUCTION

Worldwide, degenerative eye diseases such as macular degeneration, glaucoma, cataract, or retinal detachment are the main causes of blindness [1]. More, retinal diseases are already the most common cause of childhood blindness worldwide [2]. The main microvascular complication of diabetes in the eye is the diabetic retinopathy (DR), which is found in almost 20% of newly diagnosed diabetic people. Age-related macular degeneration (AMD) is another retinal disease discussed and highlighted as a growing concern and it is already the third largest cause of blindness in the world. The annual incidence of Retinal detachment (RD) is estimated at 10/100,000 per year. Globally, 90 eyes are blinded by RD every hour [3].

Optical coherence tomography (OCT) is a non-invasive imaging test that provides high resolution images of retinal structures, helping the early detection, diagnosis and treatment guidance for retinal diseases in their early stages, before vision is affected. The OCT produces cross sectional view of the retina, with an accuracy ranging from 5 to 10 microns [4]. It is analogous to ultrasound imaging, except that it uses light instead of sound [5-6].

One of the main limitations of OCT images is the presence of an unwanted speckle noise, a multiplicative noise that affects small and low-intensity features. Many well known digital denoising methods have been adapted for OCT images, including median filtering [7-8], anisotropic diffusion filters [8-9], or bayesian estimations [10]. Wavelets based denoising methods have the advantage of performing denoising on multiple resolutions. The Dual Tree Complex Wavelet Transform has been used in [11], while the curvelet transform was used in [12].

This paper presents a speckle reduction method in the wavelets domain, that associates the Hyperanalytic Wavelet Transform (HWT) with a Maximum A Posteriori (MAP) filter called bishrink.

The rest of the paper is structured as follows: Section II is dedicated to the theoretical part regarding the proposed denoising method. In Section III, the experimental results obtained for real OCT images are presented, while the last section is dedicated to conclusions.

II. MATERIAL AND METHODS

Images denoising methods can be classified in two distinct categories: methods acting in the spatial domain and the methods acting in the wavelets domain [13]. This paper is focused on the second category. This class of denoising methods has three steps:

1. Computation of a wavelet transform,
2. Detail coefficients filtering, and
3. Computation of the corresponding inverse wavelet transform.

Regarding the first and the last step, there are various wavelet transforms that can be used. One of them is the Discrete Wavelet Transform (DWT). However, it has three main disadvantages: it is not shift invariant, the associated mother wavelets are not symmetric, and its directional selectivity is poor. An alternative to the use of the DWT is the Undecimated Discrete Wavelet Transform (UDWT). The UDWT, also called Stationary Wavelet Transform (SWT), was used in [14]. However, even if the UDWT is translations invariant, its directional selectivity is poor and it is very redundant [13]. The previously stated

[1] Faculty of Electronics and Telecommunications, Communications Dept.
Bd. V. Parvan 2, 300223 Timisoara, Romania, e-mail cristina.stolojescu-crisan@upt.ro
[2] Faculty of Electronics and Telecommunications, Communications Dept.
Bd. V. Parvan 2, 300223 Timisoara, Romania, e-mail alexandru.isar@upt.ro

three disadvantages of the DWT can be diminished using complex wavelet transforms. The interest in complex wavelets may be linked to the development of the dual filter bank [15-16]. The DT-CWT is a quadrature pair of DWT trees and its coefficients may be interpreted as arising from the DWT associated with a quasi-analytic wavelet. The main property of the 2D DT-CWT is the quasi-shift invariance [13]: perfect shift invariance at level 1, and approximately achieved shift invariance beyond this level. In this paper, we will focus on the HWT. The HWT is quite similar to the DT-CWT behavior. However, the DT-CWT requires special mother wavelets, while for the implementation of the HWT classical mother wavelets, such as the ones belonging to the Daubechies family, can be used.

Concerning the second step of wavelets based denoising algorithms, one of the most efficient denoising methods implies the use of maximum a posteriori (MAP) filters. An interesting MAP filter is the bishrink filter.

A. *The Hyperanalytic Wavelet Transform (HWT)*

Being given the real mother wavelets, $\psi(x,y)$, the hypercomplex mother wavelet associated to $\psi(x,y)$ is defined as:

$$\psi_a(x,y)=\psi(x,y)+i\mathcal{H}_x\{\psi(x,y)\}+ \\ +j\mathcal{H}_y\{\psi(x,y)\}+k\mathcal{H}_x\{\mathcal{H}_y\{\psi(x,y)\}\}, \quad (1)$$

where $i^2=j^2=-k^2=-1$, $ij=ji=k$ and $\mathcal{H}$ represents the Hilbert transform [13].

The HWT of an image $f(x,y)$ can be computed as:

$$HWT_f=HWT\{f(x,y)\}=\langle f(x,y),\psi_a(x,y)\rangle. \quad (2)$$

Using (1) and (2), it results:

$$HWT_f=DWT\{f(x,y)\}+iDWT\{\mathcal{H}_x\{f(x,y)\}\}+ \\ +jDWT\{\mathcal{H}_y\{f(x,y)\}\}+kDWT\{\mathcal{H}_y\{\mathcal{H}_x\{f(x,y)\}\}\} \quad (3)$$

In the end we obtain:

$$HWT_f=\langle f_a(x,y),\psi(x,y)\rangle=DWT\{f_a(x,y)\}. \quad (4)$$

The HWT of the image can be obtained using the 2D-DWT of its associated hypercomplex image. The HWT implementation is presented in Fig. 1.

The HWT implementation shown in Fig. 1 uses four trees, each one implementing a 2D-DWT: the first one is applied to the input image, the next two trees are applied to the 1D Hilbert transforms computed across the lines ($\mathcal{H}_x$) or columns ($\mathcal{H}_y$) of the input image, and the last tree is applied to the result obtained by the computation of the two 1D Hilbert transforms on the input image.

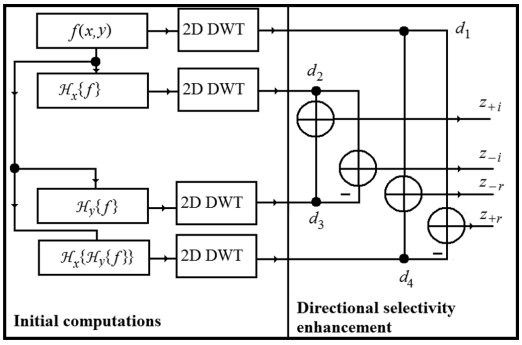


Fig. 1. The 2D HWT implementation architecture.

B. *Bishrink filtering*

The bishrink filter is a MAP filter that takes into account the interscale dependency of wavelet coefficients. Based on the observation $y = w + n$, where n represents the wavelet transform of the noise, n_i, obtained as the logarithm of the speckle $n_i=\log sp$, and w represents the wavelet transform of the useful component corresponding to the input image s, obtained as the logarithm of the noiseless component of the acquired image $s=\log u$. The MAP estimation of w is given by:

$$\hat{w}(y)=\underset{w}{\operatorname{argmax}}\left\{\ln\left(p_n(y-w)\,p_w(w)\right)\right\}, \quad (5)$$

where p_n is the noise probability density function (pdf), when the noise is AWGN (independent), while the a priori distribution of the parameter w, or "prior" $p_w(w)$ contains what is known before making the measurements.

For the construction of the bishrink filter, the noise is assumed to be i.i.d. Gaussian [17], because the HWT is a unitary transform which do not correlate the i.i.d. Gaussian noise [18]:

$$p_{\mathbf{n}}(\mathbf{n})=\frac{1}{2\pi\sigma_n^2}\cdot e^{-\frac{n_1^2+n_2^2}{2\sigma_n^2}},\ \mathbf{n}=[n_1,n_2]. \quad (6)$$

The model of a noiseless image is given using a heavy tailed distribution:

$$p_{\mathbf{w}}(\mathbf{w})=\frac{3}{2\pi\sigma^2}\cdot e^{-\frac{\sqrt{3}}{\sigma}\sqrt{w_1^2+w_2^2}},\ \mathbf{w}=[w_1,w_2]. \quad (7)$$

If we replace these two pdfs in equation (6) we obtain:

$$^{1}\hat{\mathbf{w}}(\mathbf{y}) = \underset{\mathbf{w}}{\operatorname{argmax}} \left\{ \ln \left(\frac{1}{2\pi\sigma_n^2} e^{-\frac{(y_1-w_1)^2+(y_2-w_2)^2}{2\sigma_n^2}} \cdot \frac{3}{2\pi\sigma^2} e^{-\frac{\sqrt{3}}{\sigma}\sqrt{w_1^2+w_2^2}} \right) \right\} \quad (8)$$

After several computations it results:

$$\begin{cases} w_1 = \dfrac{\sigma\sqrt{w_1^2+w_2^2}}{\sigma\sqrt{w_1^2+w_2^2}+\sqrt{3}\sigma_n^2} y_1 \\ w_2 = \dfrac{\sigma\sqrt{w_1^2+w_2^2}}{\sigma\sqrt{w_1^2+w_2^2}+\sqrt{3}\sigma_n^2} y_2 \end{cases} \quad (9)$$

By making the sum $w_1^2 + w_2^2$, the following result is obtained:

$$w_1^2+w_2^2 = \frac{\sigma^2\left(w_1^2+w_2^2\right)}{\left[\sigma\sqrt{w_1^2+w_2^2}+\sqrt{3}\sigma_n^2\right]^2}\left(y_1^2+y_2^2\right)$$

$$\Updownarrow$$

$$w_1^2+w_2^2 = \frac{\left[\sigma\sqrt{w_1^2+w_2^2}+\sqrt{3}\sigma_n^2\right]^2}{\sigma^2} = y_1^2+y_2^2 \quad (10)$$

In the end it results:

$$\sqrt{w_1^2+w_2^2} = \left(\sqrt{y_1^2+y_2^2}-\sqrt{3}\frac{\sigma_n^2}{\sigma}\right)_+ \quad (11)$$

By combining equation (8) and equation (9), we obtain:

$$\begin{cases} ^{1}\hat{w}_1 = \dfrac{\left(\sqrt{y_1^2+y_2^2}-\sqrt{3}\dfrac{\sigma_n^2}{\sigma}\right)_+}{\sqrt{y_1^2+y_2^2}} y_1 \\ ^{1}\hat{w}_2 = \dfrac{\left(\sqrt{y_1^2+y_2^2}-\sqrt{3}\dfrac{\sigma_n^2}{\sigma}\right)_+}{\sqrt{y_1^2+y_2^2}} y_2 \end{cases} \quad (12)$$

Thus, the input-output relation of the bishrink filter is:

$$\hat{w}_1 = \frac{\left(\sqrt{y_1^2+y_2^2}-\dfrac{\sqrt{3}\sigma_n^2}{\sigma}\right)_+}{\sqrt{y_1^2+y_2^2}} y_1 \quad (13)$$

The bishrink filter requires prior knowledge of the noise variance and of the marginal variance of the noise-less image for each wavelet coefficient. For the estimation of the noise variance from the noisy wavelet coefficients, a robust median estimator from the wavelet coefficients finest scale is used [19]:

$$\hat{\sigma}_n^2 = \frac{\text{median}\left(|y_i|\right)}{0.6745}, \quad y_i \in \text{sub-band HH}. \quad (14)$$

The marginal variance of the k^{th} coefficient can be estimated using neighboring coefficients in the region $N(k)$, a squared shaped window centered on this coefficient, with the size of 7×7 [21]. The estimation can be done using the equation:

$$\sigma_y^2 = \sigma^2 + \sigma_n^2, \quad (15)$$

where σ_y^2 represents the marginal variance of the noisy observations y_1 and y_2.

It results:

$$\hat{\sigma} = \sqrt{\left(\hat{\sigma}_y^2 - \hat{\sigma}_n^2\right)_+} \quad (16)$$

For the estimation of the marginal variance of the noisy observations, the following relation is proposed in [17]:

$$\hat{\sigma}_y^2 = \frac{1}{M}\sum_{y_i \in N(k)} y_i^2, \quad (17)$$

where the neighborhood $N(k)$ has the size M.

In order to estimate the local standard deviation of the useful component corresponding to the parent coefficients, $\hat{\sigma}_2$, in a given sub-band, the sub-band is first interpolated by the repetition of each line and column. Then, by applying the relations (16) and (17), the local standard deviation of the useful component corresponding to the child coefficients is obtained:

$$\hat{\sigma} = \frac{\hat{\sigma}_1 + 0.5\cdot\hat{\sigma}_2}{2} \quad (18)$$

The local variance of a pixel also gives some information about the frequency content of the region to which the considered pixel belongs: pixels having low local variances imply a corresponding region with low frequencies, while pixels having high local variances imply a corresponding region containing high frequencies.

The estimation of the noise variance is obtained using the equation:

$$\hat{\sigma}_n^2 = median\left(|y_i|\right), \quad y_i \in \text{ sub-band HH}. \quad (19)$$

The standard deviation of the noiseless coefficients can be estimated as:

$$\hat{\sigma} = \begin{cases} \sqrt{\frac{1}{M} \sum_{y_i \in N(k)} y_i^2 - \hat{\sigma}_n^2}, & if \ \frac{1}{M} \sum_{y_i \in N(k)} y_i^2 - \hat{\sigma}_n^2 > 0 \\ 0, & if \ not \end{cases} \tag{20}$$

where M is the size of the moving window $N(k)$, centered on the kth pixel of the acquired image.

The sensitivity of the bishrink filter with the estimation of the noise standard deviation $\hat{\sigma}_n$ can be computed with the relation:

$$S_{\hat{w}_1}^{\hat{\sigma}_n} = \frac{d\hat{w}_1}{d\hat{\sigma}_n} \cdot \frac{\hat{\sigma}_n}{\hat{w}_1} \tag{21}$$

Using the input-output relation of the bishrink filter in equation (10) we obtain:

$$S_{\hat{w}_1}^{\hat{\sigma}_n} = \begin{cases} \frac{-2\sqrt{3}\hat{\sigma}_n^2}{\hat{\sigma}\sqrt{y_1^2 + y_2^2 - \sqrt{3}\hat{\sigma}_n^2}}, & \text{if } \sqrt{y_1^2 + y_2^2} > \frac{\sqrt{3}\hat{\sigma}_n^2}{\hat{\sigma}} \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

The absolute value of the sensitivity is an increasing function of $\hat{\sigma}_n$. The performance of the bishrink filter decreases with the increase of the noise standard deviation estimation value.

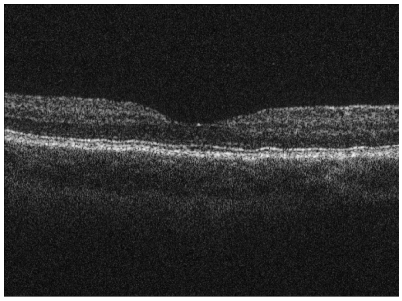
An important parameter of the bishrink filter is the local estimation of the noiseless image marginal variance ($\hat{\sigma}$). The sensitivity of the estimation $\hat{w}_1$ with $\hat{\sigma}$ is given by:

$$S_{\hat{w}_1}^{\hat{\sigma}} = \begin{cases} \frac{\sqrt{3}\hat{\sigma}_n^2}{\hat{\sigma}\sqrt{y_1^2 + y_2^2 - \sqrt{3}\hat{\sigma}_n^2}}, & \text{if } \sqrt{y_1^2 + y_2^2} > \frac{\sqrt{3}\hat{\sigma}_n^2}{\hat{\sigma}} \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

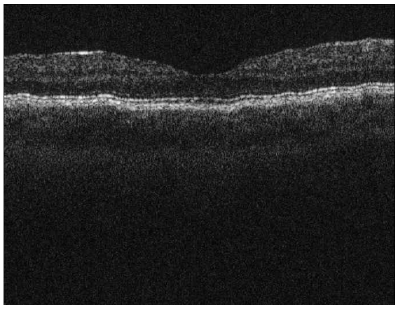
The estimation precision using the bishrink filter decreases with the decreasing of $\hat{\sigma}$.

III. RESULTS

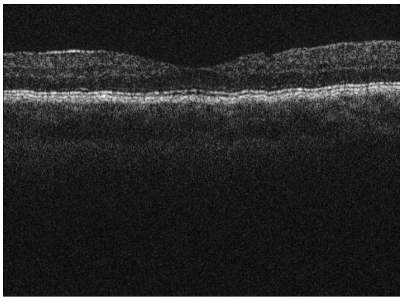
In this section, we test our denoising approach on three OCT images shown in Fig. 2.



a) OCT 1



b) OCT 2



c) OCT 3

Fig. 2. The three OCT images used for testing.

The obtained results are analyzed in terms of the noise variance and the Equivalent Number of Looks (ENL) which quantifies the homogeneity degree of a region. The ENL is defined by the ratio of the squares of pixels mean and variance situated in the considered region. It can be computed as follows:

$$ENL = \left(\frac{mean}{standard\ deviation} \right)^2. \tag{24}$$

The results are shown in Table 1.

Table 1

Images	ENLi	ENLo	σ_{ni}	σ_{no}
OCT 1	5.19	73.94	8.94	0.28
OCT 2	5.56	86.82	8.37	0.33
OCT 3	6.03	100.64	8.386	0.31

In Table 1, ENLi represents the input ENL value, while the ENLo is the value obtained after the denoising procedure. σ_{ni} and σ_{no} are the values of the noise variance before and after the denoising.

The denoising algorithm significantly reduces the noise variance and the ENL output values indicate a good performance of the proposed denoising algorithm.

In Fig. 3 two homogenous regions (before and after denoising) from each test images, are compared.

Based on visual inspection, the proposed denoising method seems to be effective.

IV. CONCLUSIONS

This paper presents an effective wavelets based denoising system for OCT images. Wavelets based denoising methods have the advantage of performing denoising on multiple resolutions which is useful in the case of correlated noise.

The proposed denoising algorithm associates the Hyperanalytic Wavelet Transform and with the bishrink filter. The implementation of the HWT is very simple and flexible, permitting the use of any orthogonal or biorthogonal real mother wavelets for its computation. In this paper we used the Daubechies family of mother wavelets.

The experimental results presented in Table 1 and in Fig. 3 highlight the effectiveness of the proposed algorithm.

Acknowledgement

This work was partially supported by the strategic grant POSDRU/159/1.5/S/137070 (2014) of the Ministry of National Education, Romania, co-financed by the European Social Fund – Investing in People, within the Sectoral Operational Programme Human Resources Development 2007-2013.

REFERENCES

[1] C. Delcourt, "Nutrition and age-related eye diseases: the Alienor (Antioxydants, Lipides Essentiels, Nutrition et maladies OculaiRes) Study", *Journal of Nutrition Health Aging*, 14(10), pp. 854-61, 2010.

[2] D. Yorston, "Retinal Diseases and VISION 2020", *Community Eye Health,* 16(46), pp. 19–20, 2003.

[3] S. Shah, "Blindness and visual impairment due to retinal diseases", *Community Eye Health,* 22(69), pp. 8–9, 2009.

[4] D. Huang, E. Swanson, C. Lin, J. Schuman, W. Stinson, W. Chang, M. Hee, T. Flotte, K. Gregory, C. Puliafito C, J. Fujimoto, "Optical coherence tomography", *Science* 254, pp. 1178-1181, 1991.

[5] M. Born, E. Wolf, *Interference and Diffraction with Partially Coherent Light*, fourth ed., Principles of Optics, Pergamon Press, United Kingdom pp. 491–505, 1970.

[6] Kiernan, W. Mieler, and S. Hariprasad, "Spectral-domain optical coherence tomography: a comparison of modern high-resolution retinal imaging systems", *American Journal of Ophthalmology* 149(1), pp.18–31, 2010.

[7] M. E. B. Jadwiga Rogowska, "Image processing techniques for noise removal, enhancement and segmentation of cartilage OCT images", *Physics in Medicine and Biology,* 47(4), pp. 641–655, 2012.

[8] C. P. Loizu, C. Theofanous, M. Pantziaris, et al., "Despeckle filtering software toolbox for ultrasound imaging of the common carotid artery", *Computer Methods and Programs in Biomedicine*, I14, pp. 109-124, 2014.

[9] P. Puvanathasan and K. Bizheva, "Interval type-II fuzzy anisotropic diffusion algorithm for speckle noise reduction in optical coherence tomography images", *Opt. Express*, 17(2),pp. 733–746, 2009.

[10] A. Wong, A. Mishra, K. Bizheva, and D. A. Clausi, "General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery", *Opt. Expres*s, 18(8), pp. 8338–8352, 2010.

[11] S. Chitchian, M. A. Fiddy, and N. M. Fried, "Denoising during optical coherence tomography of the prostate nerves via wavelet shrinkage using dual-tree complex wavelet transform", *Journal of Biomedical Optics*, 14(1), pp. 14-31, 2009.

[12] Z. Jian, L. Yu, B. Rao, B. J. Tromberg, and Z. Chen, "Three-dimensional speckle suppression in optical coherence tomography based on the curvelet transform," *Optics Express*, 18(2), pp. 1024–1032, 2010.

[13] A. Isar, I. Firoiu, C. Nafornita, S. Moga, "Sonar Images Denoising", in: Nikolai Kolev, N. (Ed.), Sonar Systems. INTECH, Croatia, pp. 173-206, 2011.

[14] S. Foucher, G. B. Benie, J. M. Boucher, "Multiscale MAP Filtering of SAR images", *IEEE Transactions on Image Processing,* 10(1), pp. 49-60, 2001.

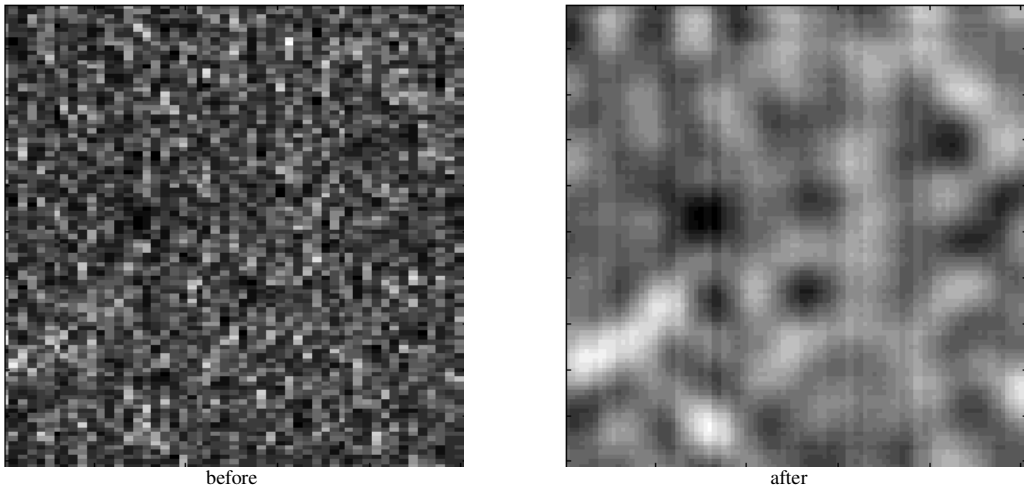
[15] N. Kingsbury, "The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement," Proceedings of EUSIPCO, Rhodes, Greece, 1998, pp. 319–322.

[16] N. Kingsbury, "Complex Wavelets for Shift Invariant Analysis and Filtering of Signals", *Applied and Computational Harmonic Analysis,* vol. 10, pp. 234-253, 2001.

[17] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency", *IEEE Transactions on Signal Processing,* 50(11), pp. 2744-2756, 2002.

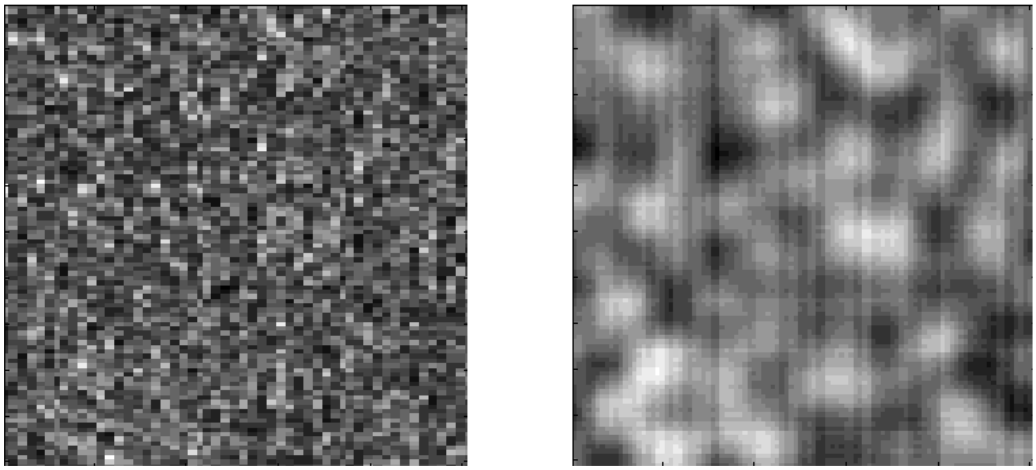
[18] I. Firoiu, C. Nafornita, D. Isar, A. Isar, "Bayesian hyperanalytic denoising of SONAR images", *IEEE Geoscience and Remote Sensing Letters,* 8(6), pp. 1065-1069, 2011.

[19] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage", *Biometrika*. 81(3), pp. 425-455, 1994.



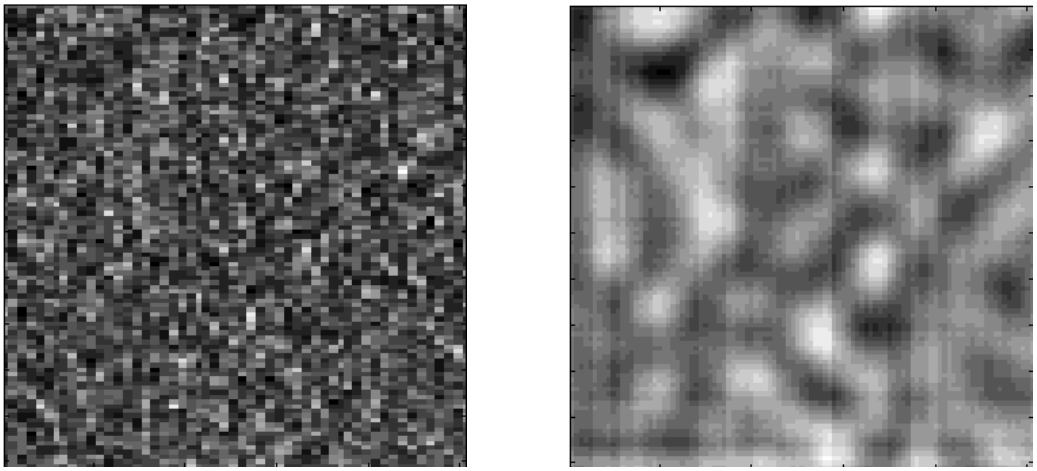
before after

a) OCT 1



before after

b) OCT 2



before after

c) OCT 3

Fig. 3. Results for OCT images in an homogenous region before / after HWT+bishrink.