Shearlet Transform: a Good Candidate for Compressed Sensing in Optical Coherence Tomography

Lesley-Ann Duflot[1,2], Alexandre Krupa[1], Brahim Tamadazte[2], and Nicolas Andreff[2]

Abstract— This paper deals with the development of a fast and smart acquisition technique of Optical Coherence Tomography (OCT) data that has the capability to reconstruct missing data of OCT image. The main objective is to reduce the acquisition time (i.e., increase the frame rate) of an OCT-scan system by choosing a trajectory that covers entirely the image but that does not take all the measurements. The reconstruction of the missing data is achieved by applying an updated Fast Iterative Soft-Thresholding Algorithm (FISTA) on a sparse representation of the image. Several sparse representations have been tested and the shearlets-based approach seems to outperform the other ones (e.g. wavelets, curvelets and by applying bilinear interpolation). The targeted application is a fast OCT imaging solution allowing an efficient compensation of the artefacts induced by the patient physiological motions for diagnostic purpose through optical biopsies (3D micrometric resolution optical images). The obtained results seem promising in terms of low time processing and improvement of the quality of the reconstructed image compared to the traditional sparse acquisition.

I. INTRODUCTION

Optical Coherence Tomography is an emerging technology for performing high-resolution, cross-sectional, tomographic imaging. OCT was first developed in 1991 by Huang and his colleagues in James Fujimoto's laboratory at the Massachusetts Institute of Technology (MIT) as an innovative technology for non-invasive imaging of biological tissues and was tested in vitro for viewing the retina and the coronary artery [1]. The first clinical *in-vivo* tests in the human optic disc and macula were performed in 1993 [2]. In subsequent years, this technology was improved and the first device was commercialized in 1996. Currently, OCT systems are widely used in ophthalmology for treating macular diseases and glaucoma, reaching micrometer resolution and millimeter penetration depth.

For these reasons, OCT is an established medical imaging technique used in various medical and mini- and non-invasive surgical applications (e.g., dermatology, cardiology, etc.). This growing interest for the use of OCT in the medical fields as diagnostic means has led research activities to evolve OCT devices to more compact systems able to access within very limited areas (millimetric size). However, the main drawback of OCT systems is the inappropriate acquisition time (especially for OCT volume i.e., 3D optical biopsy) that is too long. Few researches have tried to enhance the acquisition time, with compressed sensing (CS) approaches [3], [4].

In fact, the Compressed Sensing paradigm can meet this challenge. It was introduced by Candès, D. Donoho, J. Romberg and T. Tao [5] who demonstrate that it is possible to overcome the limitations imposed by the Shannon sampling theorem, long considered as a cornerstone in information theory. This means that any signal or image can be reconstructed with few loss when there exist a basis Ψ in which the signal is very sparse and a basis Φ in which the sampling is performed that are mutually incoherent. As mentioned above, *tomography* in particular is an area where it is very important to build a protocol for rapid and efficient acquisition.

This work is part of an ambitious project dealing with the development of a microrobotic-based nasal endoscopy system embedding a fiber-based OCT system[1]. The medical aim is the *in-situ* characterization (optical biopsy) of the olfactory mucosa in order to prevent an eventual smell deficiency. In this case, CS could provide an interesting way to perform repetitive and rapid optical biopsies of the olfactory mucosa using a fiber-based OCT system. Therefore, this paper focuses on the design of a CS approach in order to improve the OCT system frame rate and consequently to reduce the image artefacts especially due to the long acquisition time. More precisely, it deals with OCT compressed sensing using an update optimization algorithm based on Fast Iterative Shrinkage Algorithm (FISTA) [6] on different image transforms: wavelets, curvelets, shearlets.

In the remainder of this paper, Section II recalls the compressed sensing basics. The OCT image formation and acquisition are discussed in Section III and the proposed adaptation of the FISTA algorithm for CS is presented in Section IV. The comparison between the different image transforms with our updated FISTA and a simple bilinear interpolation using OCT image of organic and non-organic samples are presented in Section V.

II. OCT IMAGING SYSTEM

The working principle of an OCT is close to ultrasound imaging (US): it is based on the measure of the echo time delay and the intensity of the back-scattered or reflected light from the sample [1]. The time of flight of the light-wave

[1]L. A. Duflot and A. Krupa are with the Inria Rennes-Bretagne Atlantique, Campus universitaire de Beaulieu, Rennes 35042, France lesley-ann.duflot@inria.fr - alexandre.krupa@inria.fr

[2]L. A. Duflot, B. Tamadazte and N. Andreff are with FEMTO-ST, AS2M, Univ. Bourgogne Franche-Comté, Univ. de Franche-Comté/CNRS/ENSMM, 25000 Besançon, France brahim.tamadazte@femto-st.fr – nicolas.andreff@femto-st.fr

provides information about the position of the scattering site in the sample, while the intensity of the back-scattered light depends on the optical properties of the tissue. Since light travels faster than sound , direct measurement of the echo time as in US systems is not viable. In OCT systems this measure is indirectly obtained by using low-coherence interferometry and implementing a simple Michelson interferometer. In our experimental test, we use a Fourier-Domain OCT system (see Fig. 1) Testo III from Thorlabs[2].
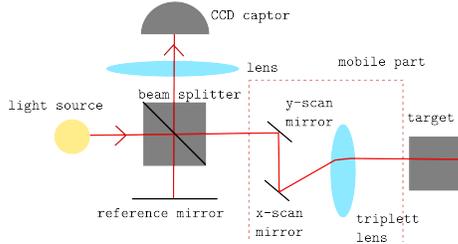

Fig. 1.   Structure of a Fourier-Domain OCT system.

III. COMPRESSED SENSING BASICS

Shannon and Nyquist's theorem verifies that under uniform sampling, it is necessary to acquire data at a frequency twice the bandwidth. Actually, this condition remains valid from a theoretical point of view but is often violated in practice. CS is a perfect example of this reality as demonstrated by Candès and Romberg [5], [7]. Thereby, it becomes possible to reconstruct a signal (respectively an image) with few loss when there exist a basis Ψ in which the signal (respectively an image) is very sparse and a basis Φ in which the sampling is performed that are mutually incoherent. It results that a signal of size n can be reconstructed from m measurements ($m << n$) if the following conditions are respected:

- *Sparsity*: there exists a basis (i.e., the representation basis or the sparse basis) Ψ, where the signal is defined as k-sparse, i.e. contains only k non-zero coefficients;
- *Incoherence*: there exists a low coherence μ between the basis Ψ and the basis Φ where the signal is acquired (i.e., the sensing initial basis).

With $\mu(\Phi, \Psi) = \sqrt{n}\Big(\max_{1<k,j<n}\big(|\Phi_j, \Psi_k|\big)\Big)$ where, n is the length of the two bases Φ and Ψ, Φ_j is the j_{th} component of Φ and Ψ_k is the k_{th} component of Ψ. This means that the maximal correlation between two signal values, $\mathbf{v}$ defined in Φ basis and $\mathbf{v}$' chosen in Ψ basis, must be low. Indeed, in case of a 2D image, one pixel in the sparse basis Ψ can be obtained by several pixels in the sensing basis Φ. Let us consider the following example:

- $\mathbf{x} \in \mathbb{R}^n$, the signal to be reconstructed;
- $\mathbf{y} \in \mathbb{R}^m$, with $m << n$, the measured signal;
- $\Phi \in \mathbb{R}^{m\times n}$, represents the sensing matrix, such as $\mathbf{y} = \Phi\mathbf{x}$;
- $\Psi \in \mathbb{R}^{n\times n}$, defines the sparse matrix, such as $\mathbf{x} = \Psi\mathbf{s}$, where $\mathbf{s}$ is a k-sparse vector.

[2]www.thorlabs.com

Now, the CS issue can be considered as an optimization problem which consists of the minimization of the ℓ1-norm of $\mathbf{x}$ with respect to the equalities $\mathbf{y} = \Phi\mathbf{x}$ and $\mathbf{x} = \Psi\mathbf{s}$. Generally, in case of a 2D image signal, $\mathbf{s}$ is composed by a majority of *zeros* in the sparse basis. Thereby, our problem can be written as

$$\min \|\mathbf{s}\|_1 \text{ subject to } \mathbf{y} = \Phi\Psi\mathbf{s} \qquad (1)$$

It can be underlined that, based on the literature, the maximum rate of signal recovery is ensured when $m \geq 4k$ [7].

IV. UPDATED FAST ITERATIVE SHRINKAGE-THRESHOLDING ALGORITHM

Fast Iterative Soft-Thresholding Algorithm (FISTA) is usually used for image denoising [6] and inpainting [8]. The goal is to resolve a general linear equation: $\mathbf{A}\mathbf{x} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$ are known and $\mathbf{x} \in \mathbb{R}^n$ is unknown. Due to the fact that it is an under-determined large scale problem ($m << n$), the use of standard matrix decomposition techniques (such as LU, SQ or SVD) is impossible. The sparsity is the only way to find a solution. The Fast Iterative Soft-Thresholding Algorithm (FISTA) can solve any regularized problem of type $\min_{\mathbf{x}}(f(\mathbf{x}) + g(\mathbf{x}))$.

Thus, FISTA needs to define the gradient ∇f of the f function, a proximal operator $prox_g$ of the function g and an initialisation x_0 of the solution, often chosen as a vector of 0. For instance, if the problem is a basic least square problem, $f(\mathbf{x})$ is taken as $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ and $g(\mathbf{x})$ is taken as 0.

Data: $prox_g$, ∇f,$\mathbf{x}_0$
Result: $\mathbf{x}$
$\mathbf{y}_1 = \mathbf{x}_0, t_1 = 1$;
while $\mathbf{x}_{k+1} - \mathbf{x}_k > \epsilon$ **do**
  $\quad \mathbf{x}_k = prox_g(\mathbf{y}_k - t_k\nabla f(\mathbf{y}_k))$;
  $\quad t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$;
  $\quad \mathbf{y}_{k+1} = \mathbf{x}_k + \frac{t_k-1}{t_{k+1}}(\mathbf{x}_k - \mathbf{x}_{k-1})$;
end

Algorithm 1: FISTA

In our case, let us consider $\mathbf{b}$ the measurements (pixel intensities) vector and $\mathbf{A}$ the matrix representing the coordinates of the $\mathbf{b}$ values ($\mathbf{A}$ taken in the Φ basis presented in the Section III), $\mathbb{W}$ the transformation from the measurements basis towards a basis where signal will be sparse (the Ψ basis presented in the Section III). The ℓ1-minimisation problem is defined as

$$\mathbf{x}_{l1} = \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + l\|\mathbb{W}\mathbf{x}\|_1 \qquad (2)$$

with $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ represents the fidelity of $\mathbf{x}$ to the acquired data $\mathbf{b}$ and $\|\mathbb{W}\mathbf{x}\|_1$ represents the sparsity of $\mathbf{x}$ in a certain basis Ψ, l is a simple trade-off between these two constraints. According to the ISTA formulation $f(\mathbf{x})$ is taken as $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ and $g(\mathbf{x})$ is taken as $\|\mathbb{W}\mathbf{x}\|_1$. Traditionally, [3], $prox_g$ and ∇f are taken as:

- $prox_g = T_a(\mathbf{x}) = sgn(\mathbf{x}_i) \times Max(|\mathbf{x}_i| - a, 0)$, a soft-thresholding function with a parameter a.
- $\nabla f(\mathbf{x}) = (\mathbf{A}\mathtt{W}^{-1})^T(\mathbf{A}\mathtt{W}^{-1}\mathbf{x} - \mathbf{b})$ where $\mathtt{W}^{-1}$ is the inverse of $\mathtt{W}$.

However in our case, in order to speed up the algorithm and to test sparse bases other than wavelets, more easily ($\mathtt{W}^{-1}$ is not computable as a simple matrix for curvelets and shearlets, what prevents us to compute ∇f in the previous form) a change of $prox_g$ and ∇f has been done:

- $prox_g = \mathtt{W}^{-1}T_a\mathtt{W}(\mathbf{x})$
- $\nabla f(\mathbf{x}) = \mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$

Thus, the main-step of the usual algorithm is replaces by :

$$\mathbf{x}_k = \mathtt{W}^{-1}T_a\mathtt{W}(\mathbf{y}_k - t_k\mathbf{A}^T(\mathbf{A}\mathbf{y}_k - \mathbf{b})) \qquad (3)$$

In order to test the proposed updated FISTA algorithm (u-FISTA), we opted for the well-known sparse representation bases Ψ: wavelets [9], curvelets [10] and shearlets [11] respectively computed in matlab with Wavelab 850[3], Ridgelet and Curvelet first generation Toolbox[4] and Shearlab[5]. The soft-thresholding parameter has been optimized for each case.

V. VALIDATION RESULTS

A. Score measurements

The work described in this paper was tested using different OCT images. In order to judge quantitatively the updated FISTA algorithm, using the different image sparse basis representations, various scores were considered to compute the similarity degree between the reference OCT image $\mathbf{I_r}$ and the obtained image $\mathbf{I_c}$ (of size L) after a compressed sensing process, this using, for instance, the Peak Signal Noise Ratio (PSNR):

$$PSNR = 10 * log_{10}(\frac{255^2}{MSE}) \quad \text{with} \quad MSE = \frac{(\mathbf{I}_c - \mathbf{I}_r)^2}{L}. \qquad (4)$$

Comparing two images $\mathbf{I_c}$ and $\mathbf{I_r}$ only with their signal does not always give a good visualization of the morphological proximity between the two images. Therefore the Structural SIMilarity (SSIM) [12], value between -1 and 1, of the image is often a good way to give a score to an image reconstruction.

$$SSIM = [ss(\mathbf{I}_c, \mathbf{I}_r)].[ls(\mathbf{I}_c.\mathbf{I}_r].[cs(\mathbf{I}_c.\mathbf{I}_r)] \qquad (5)$$

with ss, ls and cs, the structure, the luminance and the contrast similarity of two images, respectively. Thus, (5) can be reformulated as

$$SSIM = \frac{(2\mu_{\mathbf{I}_c}\mu_{\mathbf{I}_r} + c_1)(2\sigma_{\mathbf{I}_c}\sigma_{\mathbf{I}_r} + c_2)(2cov_{\mathbf{I}_c\mathbf{I}_r} + c3)}{(\mu_{\mathbf{I}_c}^2 + \mu_{\mathbf{I}_r}^2 + c_1)(\sigma_{\mathbf{I}_c}^2 + \sigma_{\mathbf{I}_r}^2 + c_2)(\sigma_{\mathbf{I}_c}\sigma_{\mathbf{I}_r} + c_3)} \qquad (6)$$

with $\mu_\mathbf{I}$ the mean of $\mathbf{I}$, σ_x^2 the variance of $\mathbf{I}$, $cov_{\mathbf{I}_1\mathbf{I}_2}$ the covariance of $\mathbf{I}_1$ and $\mathbf{I}_2$, c_1, c_2 and c_3 some variables to stabilize the division.

[3]http://www-stat.stanford.edu/ wavelab/
[4]http://www.mathworks.com/matlabcentral/fileexchange/31559
[5]http://www.shearlab.org/

B. Comparison between the different reconstruction frameworks:

Comparison between u-FISTA with several sparse bases and bilinear interpolation were performed by reconstructing six 256x256 pixels OCT images representing a *grape*, a *kiwi*, a *banana*, a *chicken* and two *mice epithelium system* samples (Fig. 2). To evaluate our reconstruction framework, a mask was applied to the OCT image, in order to simulate a partial acquisition of the samples, then the reconstruction was done. The reconstruction quality was measured by the two indicators that are the PSNR and SSIM presented previously.
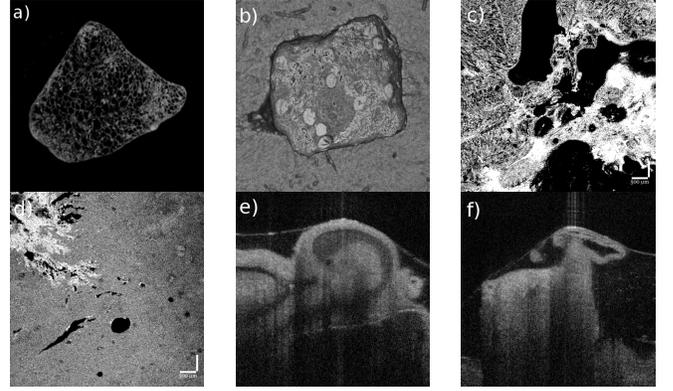

Fig. 2.   (a) grape (b) kiwi (c) banana (d) chicken (e & f) mice epithelium system

The compressed sensing theory advices to use a random acquisition mask. However, this choice is not always relevant in practice. A better solution for OCT application is to used a mask corresponding to a spiral acquisition trajectory (Fig.4 (b)) in order to take the most data as possible in a large coverage area inside an envelope ball with a minimum of time.

Figure 3 shows the mean of the PSNR and the mean of the SSIM of the different reconstructions (i.e., bilinear interpolation or u-FISTA with wavelets, curvelets or shearlets) with a random mask (dashed line) or with a spiral mask (solid line) for different percent of acquired samples. A spiral mask seems to decrease the quality of u-FISTA with curvelets, but with the two other sparse bases, the performances of the spiral mask are at least the same or even better. Thus, the use of a spiral mask will accelerate the acquisition of the data but will not decrease the quality of the reconstruction compared to a random mask commonly advised in the CS literature.

Figure 4 gives a representative example of the reconstruction for the spiral mask. First, the reconstruction with bilinear interpolation involves some structural artefacts (Fig.4(c)) and is good only when the percent of acquired samples is high. Concerning the reconstruction with u-FISTA, the choice of the reconstruction basis affects a lot the aspect of the image. With the wavelets, some important details of the image are missing in the reconstruction (Fig.4 (d)) where some rounds is
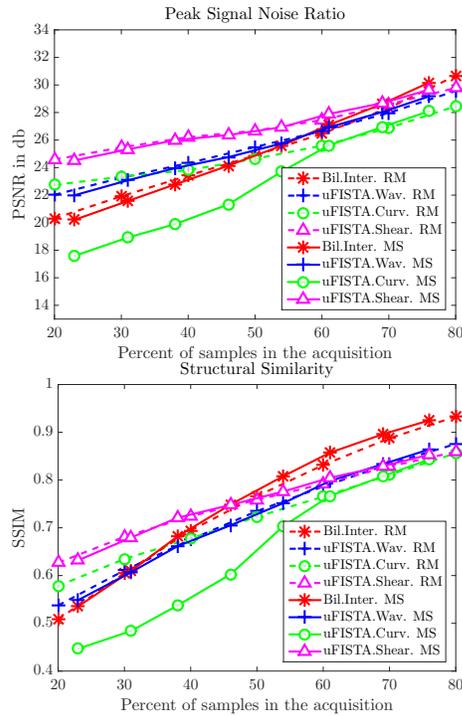
Fig. 3. (left) PSNR; (right) SSIM - for different percent of samples and for different type of mask (RM: Random Mask, MS: Mask Spiral)
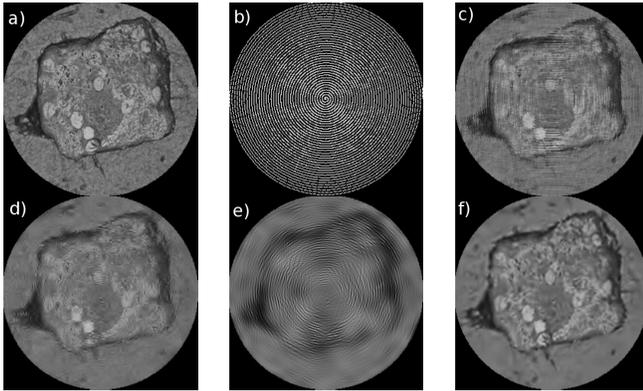

Fig. 4. (a) Reference; (b) Mask with 31% of samples; Reconstruction: (c) Bilinear Interpolation (d) FISTA with Wavelet (e) FISTA with Curvelet (f) FISTA with Shearlet.

the center of the image are not clearly visible. With curvelets, u-FISTA gives good reconstruction with a random mask (as is it showed in Fig. 3) but it is particularly not adapted to a spiral mask (Fig.4 (e)). That is probably due to the fact that curvelets transformation is computed using some rotations which disturbs the transfer to the curvelet basis. With the shearlets, whether it be with random or with spiral mask (Fig.4 (f)), all the information is kept and there are no artefacts. Figure 3 shows that the shearlets-based approach is not only good when the number of sample is low but also efficient with the use of the spiral mask.

VI. CONCLUSION & FUTURE WORKS

In this paper was presented a preliminary study of the feasibility of using Compressive Sensing for OCT images.

By comparing several bases where the signal is sparse with our reconstruction algorithm based on u-FISTA, the good performance of the shearlets basis was experimentally demonstrated. Moreover this paper points out that using a spiral trajectory, more adapted to OCT acquisition framework, to acquire the measurements does not degrade the quality of the reconstruction of the entire image. Even more, u-FISTA combination with shearlets basis is well appropriate for spiral scanning in OCT imaging since it is the best candidate for compressed sensing on this modality.

These results was obtained on pre-acquired image in which a spiral-mask was applied. Further work will be to test it with a image obtained with a OCT scanning where the laser beam will follow a real spiral trajectory, what is not possible yet with our current hardware setup. Another and not trivial advantage of our method is the reduction of physiological artefact by faster acquisition of the image. Therefore, future test of our method on a sample that exhibits physiological motion could be promising. The algorithm of reconstruction based on u-FISTA was chosen due to its simplicity and its efficiency but it will also be interesting to implement the Augmented Lagrangian Multiplier method and to compare it with u-FISTA. Another interesting improvement could be an adaptation of the method in order to allow on-line reconstruction of the image before finishing the whole acquisition of the sample.

VII. ACKNOWLEDGEMENTS

REFERENCES

[1] Huang, Swanson et al. Optical coherence tomography. Science, vol.254, n5035, p. 1178-1181. (1991)
[2] Swanson, Izatt et al, In vivo retinal imaging by optical coherence tomography. Optics letters, 18(21), 1864-1866 (1993)
[3] Lebed et al, Rapid volumetric OCT image acquisition using compressive sampling, Optics express, Vol 18, p.21003–21012 (2010)
[4] Xu, Huanget et al, Real-time compressive sensing spectral domain optical coherence tomography. Optics letters, 39(1), 76-79. (2014)
[5] Candes, Romberg and Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,Information Theory, IEEE Transactions on, Vol.52, p.489-509 (2006)
[6] Beck, Amir and Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM Journal on Imaging Sciences , Vol.2, p.183-202 (2009)
[7] Candes and Wakin, An Introduction To Compressive Sampling,Signal Processing Magazine, IEEE, Vol.25, p.21-30 (2008)
[8] Yu, Cao et al., Fast Wavelet Thresholding Algorithms for Face Image Inpainting, in GreenCom (2013)
[9] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, (2008)
[10] Ma and Plonka, A review of curvelets and recent applications, IEEE Signal Processing Magazine, Vol.27, p.118-133 (2010)
[11] G. Kutyniok and D. Labate, Introduction to shearlets, in *Shearlets*. Springer (2012).
[12] Wang, Bovik et al., Image quality assessment: from error visibility to structural similarity, Image Processing, IEEE Transactions on, Vol.13, p.600-612 (2004)