COMPARISON OF RECONSTRUCTION ALGORITHMS IN COMPRESSED SENSING APPLIED TO BIOLOGICAL IMAGING

Yoann Le Montagner[1,2], *Elsa Angelini*[2], *Jean-Christophe Olivo-Marin*[1]

[1]Institut Pasteur, Unité d'Analyse d'Images Quantitative, France
[2]Institut Telecom, Telecom ParisTech, CNRS LTCI, France

ABSTRACT

In this paper, we propose a short presentation of the *compressed sensing* imaging framework, along with a review of recent applications in the biomedical imaging field. One of the critical issue that used to hinder the application of compressed sensing in a bioimaging context is the computational cost of the underlying image reconstruction process. However, some recently published algorithms manage to overcome this difficulty, leading to acceptable reconstruction computational times. We illustrate with simulations on biological images of fluorescence microscopy a comparison of three reconstruction algorithms, evaluating data fidelity and computational efficiency.

Index Terms— Compressed sensing, sampling pattern, convex optimization, Fourier transform.

1. COMPRESSED SENSING: A NEW SAMPLING THEORY

1.1. Generalities

Compressed sensing (CS) is a new sampling theory, introduced in [4, 6]. According to this theory, structured signals (e.g. spatially sparse signals) can be acquired with a sampling rate far smaller than the Nyquist rate without any loss of quality. Instead of acquiring a large amount of data and then compress it, the CS proposes to sample the signal in a space with dense information and to reconstruct the signal via an optimization process. The underlying hypothesis made on the signal is that there exists a basis $\Psi = (\psi_i)_i$ (which has to be known *a priori*) in which the signal of interest $x \in \mathbb{R}^N$ is sparse:

$$x = \sum_{i=0}^{N-1} \alpha_i \psi_i \quad \text{with} \quad \sum_{i=0}^{N-1} \|\alpha_i\|_{l_0} \ll N \qquad (1)$$

In practice, CS has generated a great interest since eq. (1) holds for a large variety of signals. For instance, it is well known that biomedical images are sparse in the wavelet domain or in the Heaviside basis (which is equivalent to say that their derivatives are sparse in the canonical basis of $\mathbb{R}^N$).

Under the assumption that eq. (1) holds for the signal of interest x, it has been proved in [4] that, with a high probability, x can be retrieved from a small number of samples y on a measurement basis $\Phi = (\varphi_i)_i$, assuming that Φ and Ψ are incoherent. The signal is then recovered as the solution of the following convex minimization problem (where the norm l_1 replaces the norm l_0 in the sparsity constraint):

$$\min_{x \in \mathbb{R}^N} \|\Psi x\|_{l_1} \quad \text{s.t. } \Phi x = y \qquad (2)$$

Another interest of the CS is that this sampling technique is somehow intrinsically robust to noise. Indeed, since eq. (1) does not hold for the noisy component of a signal, this component is filtered out during the reconstruction process. It has been demonstrated in [5] that such filtering could be enforced by relaxing the constraint $\Phi x = y$ with the following optimization problem:

$$\min_{x \in \mathbb{R}^N} \|\Psi x\|_{l_1} \quad \text{s.t. } \|\Phi x - y\|_{l_2} \leq \epsilon \qquad (3)$$

where ϵ is tuned according to the noise nevel in the measurements y. Using eq. (3), we can see compressed sensing as a denoising technique.

In the special case of piecewise smooth images, it can be advantageous to substitute the l_1 norm of the transformed signal $\|\Psi x\|_{l_1}$ in eq. (2) and (3) with the total variation TV semi-norm of the signal, expressed as $\|x\|_{\text{TV}} = \sum_{p,q} \sqrt{\partial_1 x(p,q)^2 + \partial_2 x(p,q)^2}$ for a 2D image x. The TV semi-norm enforces sparsity for the image gradient and is generally faster to compute than $\|\Psi x\|_{l_1}$.

1.2. Sampling in the Fourier basis

For natural images (which have a sparse gradient or which are sparse in a wavelet basis), a common choice is to use Fourier basis vectors to design Φ with a particular sampling pattern in the Fourier domain (e.g. random or star-shape sampling). For a given level of undersampling, the patterns specified by Φ will greatly influence the quality of the CS reconstruction, as illustrated in fig. 1. In this figure, we see that the random-Gaussian sampling[1] performs better (both in terms of image quality and convergence speed[2]) than the random-uniform sampling, although the lack of samples in the high frequencies leads to some blurring near sharp edges.

The mathematical results presented in [4] suggests that the Fourier sampling pattern should be chosen uniformly random. However, a structured sampling pattern made of 22 radial lines in the Fourier space was also used to illustrate the reconstruction capabilities of CS. With such pattern, the low frequencies are over-represented compared to the high frequencies, unlike the uniformly random sampling hypothesis.

Recently, non-uniform Fourier sampling patterns was also studied in [19, 10, 8]), exploiting a higher sampling rate for low frequencies. A careful study of the properties of such patterns in terms of reconstruction quality and noise robustness remains open.

---

[1]By random-Gaussian sampling we mean that the probability of sampling the coefficient corresponding to the spatial frequency $k \in \mathbb{R}^2$ in the Fourier space is given by $\exp\left(-\left(\frac{\|k\|_2}{\rho}\right)^2\right)$ where the parameter ρ is tuned according to the defined sampling rate.

[2]These simulations were performed with Matlab running on a Linux workstation with a 2.5 GHz Core 2 Duo CPU and 6 GB of memory; all the image pixels are valued between 0 and 1.
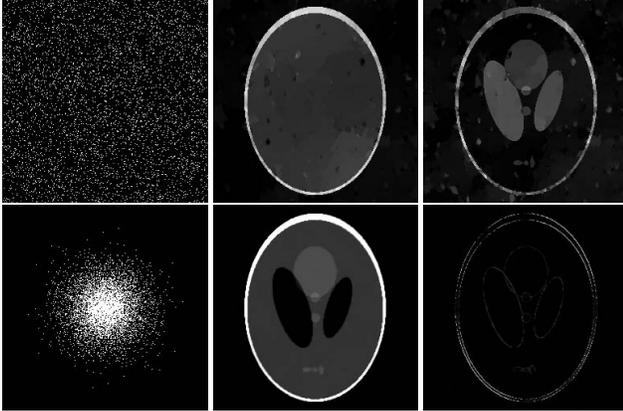
	Uniform (top)	Gaussian (bottom)
RMS of the residue	8.3×10^{-2}	3.0×10^{-2}
#iterations	1240	432
Execution time	37 sec.	14 sec.

Fig. 1. Shepp-Logan image with an additive Gaussian noise with standard deviation $\sigma = 1.5 \times 10^{-2}$ reconstructed using two different random sampling patterns: from a uniform distribution and from a Gaussian distribution. Reconstructions used the NESTA algorithm and TV regularization, using only 7% of the Fourier coefficients. Left: sampling pattern; middle: result; right: residue between the result and the original image.

2. CS APPLICATIONS IN BIOMEDICAL IMAGING

Over the past few years, several works have been carried out to apply the CS framework to various biomedical imaging techniques. Practical interest of CS imaging started in the context of magnetic resonance imaging (MRI) which performs signal acquisition directly in the Fourier domain, also called the k-space (see [10]). CS sampling enables significant reduction of acquisition time, for both static (see [14]) and dynamic (see [7]) screening. Image acquisition acceleration factors between 3 and 6 were reported with CS-MRI and has lead to the exploitation of novel imaging applications such as 3D imaging of the vocal tract during sustained sound production in [8]. CS has also been combined with parallel imaging to further reduce the acquisition time, as in [9].

Tomography-based imaging can also exploit CS sampling prior to the reconstruction process. For example, in [16], the authors examined the possibility of reducing the number of acquisition angles while preserving image resolution in the context of photo-acoustic tomography. There are also some great similarities between the family of maximum likelihood estimators for tomographic image reconstruction studied in [17] and the optimization problems involved in CS imaging.

For biological imaging, studies have been carried out in [12, 11] to apply the CS framework on fluorescence optical microscopy, as a denoising tool. In practice, optical Fourier measures could be used with a CS sampling scheme to reduce the exposure time of the biological samples and the associated photobleaching effect, while preserving a high signal-to-noise ratio. However, handling the missing phase information remains a challenging issue, which was only addressed via the temporal interpolation of partial phase information in [12].

Another application considers image reconstruction in digital holographic imaging with signal recovered from very few random measurements (see [13]).

Practical use of CS remains hindered by the computational cost of the optimization process for the image reconstruction. Several optimization algorithms have been proposed for CS reconstruction, and we review three important ones, used in biomedical CS applications.

3. CS RECONSTRUCTION ALGORITHMS

3.1. General overview

CS imaging relies on the ability to efficiently solve a convex minimization problem, either pb. (2) or (3), usually with the TV seminorm rather than the l_1 norm on wavelet coefficients. From a mathematical point of view, these problems are quite challenging, for two main reasons:

1. the objective function, either $\|\cdot\|_{l_1}$ or $\|\cdot\|_{\mathrm{TV}}$, is not smooth;

2. the dimension of the solution space $\mathbb{R}^N$ is very large for imaging applications (N can be greater than 10^6).

Former approaches for solving these problems consisted in recasting them as linear or second cone order programs (see [5, 3]), and then solving them using general purposes iterative optimization framework such as interior-point methods (see [2]). However, the computation of a solution through these methods often requires to invert a linear system of equations whose size is at least equal to N, which make them very time-consuming in high dimensions. Moreover, such general-purpose approach acts as a black-box solver, not taking into consideration the algebraic properties of Φ and Ψ, such as the fact that there often exists a fast algorithm (fast Fourier or wavelet transform) to project the signal x. Therefore, new heuristics that take advantage of these properties have emerged within the last few years, making the CS optimization problems more tractable and practical in clinical setting. We now review three popular optimization algorithms, used in biomedical applications[3]. Our results are presented in sec. 3.5.

3.2. NESTA

In [1], Becker *et al.* introduced the NESTA specialized algorithm to efficiently solve CS reconstruction problems. This algorithm is based on Nesterov's work on minimizing non-smooth functions (see [15]).

NESTA, which can solve both the l_1 and TV problems, is an iterative algorithm that produces a decreasing sequence of iterates converging to the solution. At each step, the new guessed solution is expressed as linear combination of two terms:

- first, a term that makes the iterate evolve in the opposite direction to the objective function gradient at the current point,

- second, a term that somehow keeps track of the previous gradient directions. It has been proved in [15] that this additional term helps to improve the convergence properties of the algorithm.

---

[3]Matlab implementations of those algorithms are provided by their authors on the following websites:

- NESTA: http://www.acm.caltech.edu/~nesta/

- RecPF: http://www.caam.rice.edu/~optimization/L1/RecPF/

- SPGL1: http://www.cs.ubc.ca/labs/scl/spgl1/

Both terms are defined as a solution of a quadratic optimization sub-problem, whose solution can be expressed analytically. Moreover, if the property $\Phi\Phi^* = \text{Id}$ holds (i.e. the rows of Φ form an orthonormal family), all matrix inverts that appear in the analytical solutions of the quadratic subproblems can be simplified, enabling very fast computation of the solution. The algorithmic complexity of each NESTA iteration is $\mathcal{O}(N + \mathcal{C}_\Phi)$, where $\mathcal{C}_\Phi$ is the complexity of applying Φ or Φ^*. When Φ is a subset of the rows of the Fourier transform matrix, $\Phi\Phi^* = \text{Id}$ holds, and $\mathcal{C}_\Phi = \mathcal{O}(N\log(N))$, and the computational cost of each NESTA step becomes reasonable.

3.3. RecPF

In [20], Yang *et al.* introduce the RecPF algorithm that is able to solve the CS reconstruction problems (either l_1 or TV) in the special case where Φ is a subsample of the Fourier transform[4].

More precisely, this algorithm solves the following problem:

$$\min_{x \in \mathbb{R}^N} \|x\|_{\text{TV}} + \frac{\lambda}{2} \|\Phi x - y\|_{l_2}^2 \tag{4}$$

Both (3) and (4) are equivalent, although the parameter ϵ in (3) can be closely related to the noise level observed in the signal, while λ is more abstract. The key idea of RecPF is to recast (4) into the following minimization problem:

$$\min_{x,w} \sum_{i=0}^{N-1} \left(\|w_i\|_{l_2} + \frac{\beta}{2} \|w_i - D_i x\|_{l_2}^2 \right) + \frac{\lambda}{2} \|\Phi x - y\|_{l_2}^2$$
$$\text{s.t. } \begin{cases} x \in \mathbb{R}^N \\ \forall i \quad w_i \in \mathbb{R}^2 \end{cases} \tag{5}$$

where each D_i represents the $2 \times N$ matrix that computes the two spatial derivatives of x at pixel i. One can note that (4) and (5) are equivalent when $\beta \to \infty$. The algorithm consists then in minimizing eq. (5) in w for a fixed x, then in x for the previous optimal w, and so on:

- minimization in w can be performed in $\mathcal{O}(N)$ since (5) is separable in w;
- minimization in x is a quadratic problem whose solution can be computed in $\mathcal{O}(N\log(N))$ thanks to the special properties of the Fourier transform towards the convolution product.

Overall, each RecPF iteration has an algorithmic complexity equal to $\mathcal{O}(N\log(N))$.

3.4. SPGL1

In [18], Van den Berg *et al.* introduce the SPGL1 algorithm for solving the following optimization problem:

$$\min_{x \in \mathbb{R}^N} \|x\|_{l_1} \text{ s.t. } \|\Phi x - y\|_{l_2} \le \epsilon \tag{6}$$

This algorithm does not require any specific algebraic property on Φ, provided that its product against a vector can be computed efficiently. The adaption of SPGL1 to solve pb. (3) is straightforward when the same property holds for Ψ. However, SPGL1 does not work for TV minimization.

The key idea presented in [18] is to study the following optimization problem:

$$\min_{x \in \mathbb{R}^N} \|\Phi x - y\|_{l_2} \text{ s.t. } \|x\|_{l_1} \le \tau \tag{7}$$

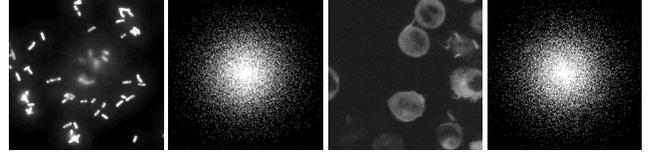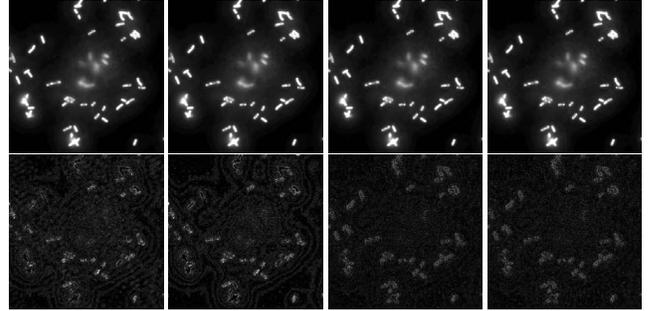[4]RecPF can also be adapted to be used with the discrete cosine transform

Fig. 2. Input images, with corresponding random-Gaussian sampled masks. Left: shigella image (425×425 pixels); right: lymphocytes T image (400×400 pixels).

	TV **minimization**		l_1 **minimization**	
	NESTA	**RecPF**	**NESTA**	**SPGL1**
RMS	7.3×10^{-3}	6.9×10^{-3}	9.5×10^{-3}	9.8×10^{-3}
#iter.	132	27	63	514
Time	19 sec.	2 sec.	20 sec.	89 sec.

Fig. 3. Simulation results for the Shigella image. Top: reconstructed images; bottom: residue between the reconstructed and the input images.

By solving several instances of this problem for different values of the parameter τ, the solution of pb. (6) can be recovered. Pb. (7) is then solved through an iterative gradient descent with backprojection on the convex set of admissible solutions. SPGL1 is reported to be particularly efficient for the following reasons:

- use of an efficient algorithm for computing the projection on the convex set $\{x \text{ s.t. } \|x\|_{l_1} \le \tau\}$;
- prediction of the evolution of successive values of τ via the exploitation of the properties of the Pareto curve $\epsilon = \phi(\tau)$, where $\phi(\tau)$ is the solution of (7), to reduce the number of (7) instances needed.

3.5. Comparison of reconstruction algorithms

We tested the three algorithms by simulations on two microscopic images presented in fig. 2. Images were reconstructed by NESTA and RecPF through TV minimization, and by NESTA and SPGL1 through l_1 minimization (using Daubechies 4 wavelet as the sparsifying basis Ψ). For all experiments, we used the same random-Gaussian sampling pattern measuring 15% of the coefficients. Results obtained with these four minimization schemes are presented in figs. 3 and 4.

We can first notice that there are large differences between the algorithm execution times: for the same input, RecPF can be roughly 10 times faster than NESTA, which is itself about 10 times faster than SPGL1. This observation about SPGL1 is somehow coherent with the comparison presented in [1], where the authors noticed that this algorithm can perform very fast computations (even faster than
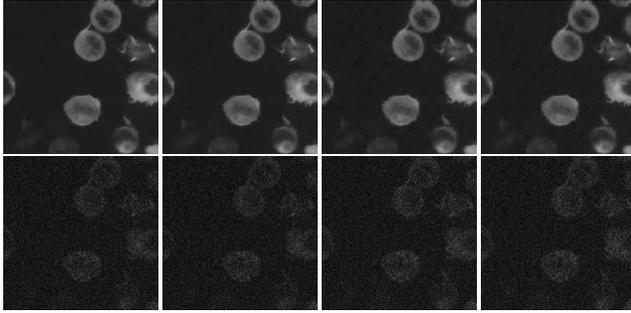
	TV **minimization**		l_1 **minimization**	
	NESTA	**RecPF**	**NESTA**	**SPGL1**
RMS	2.5×10^{-2}	2.5×10^{-2}	2.6×10^{-2}	2.6×10^{-2}
#iter.	144	21	63	2382
Time	14 sec.	1 sec.	16 sec.	560 sec.

Fig. 4. Simulation results for the lymphocytes T image.

NESTA), but has varying computational times depending on the input signal, whereas NESTA computation time seems to be very stable. This instability might make SPGL1 unsuitable for practical CS imaging. The fact that RecPF is faster than NESTA is not sufficient to disqualify this algorithm, at least for three reasons:

- NESTA is more general than RecPF, as RecPF is limited to Fourier sampling;

- Value of the Lagrange parameter λ in RecPF can be difficult to tune, whereas the ϵ parameter in NESTA can be related to *a priori* knowledge on the noise level that affects the input image (see [1]);

- the RecPF Matlab toolbox uses some C MEX functions, whereas the NESTA toolbox does not, which can introduce a bias in the computational time.

In terms of reconstruction quality, all the algorithms produce overall similar results (see figs. 3 and 4). The denoising effect of the CS can be observed on the residual images, which display Poisson distributions, meaning that we observe higher residual values (i.e. noise) in higher intensity areas. Looking more closely at the residual patterns, we noticed that the NESTA algorithm with TV minimization exhibits some unsuitable low-frequency textural artefacts (see fig. 5). These artefacts correspond to the few low-frequency coefficients not sampled by the CS, and which were not reconstructed correctly in the image. The specific sensitivity of the NESTA-TV reconstruction to this missing sampled information remains unclear.

4. CONCLUSION

In this paper, we have proposed a short review of recent CS imaging applications for biomedical applications. In the special case of Fourier-based sampling, our experiments confirmed that Gaussian random sampling is recommended over uniform random sampling; but this remains to be studied more thoroughly from a theoretical point of view.

We also compared the NESTA, RecPF and SPGL1 optimization algorithms for biomedical image reconstruction with Fourier-based sampling. Experiments have illustrated the denoising capabilities of the CS, while revealing highly variable computational times and residue qualities: the TV semi-norm associated with RecPF seems to lead to faster reconstructions and less artefacts in the residue, for Fourier-based CS imaging.
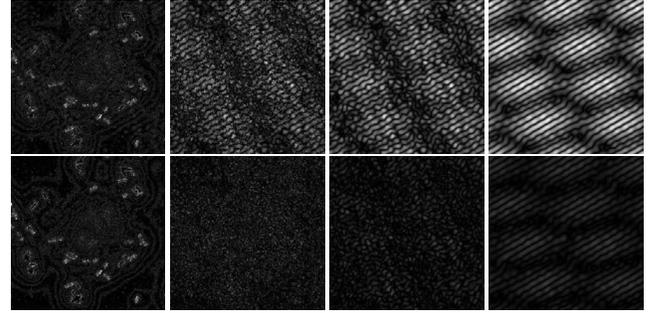

Fig. 5. Residual images obtained for Shigella through TV minimization with NESTA (top row) and RecPF (bottom row), and three low-pass filtered versions of these images, with decreasing cut-off frequency. Columns used the same colormap.

5. REFERENCES

[1] S. Becker, J. Bobin, and E. Candès. NESTA: A fast and accurate first-order method for sparse recovery. Technical report, Caltech, 2009.

[2] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[3] E. Candès and J. Romberg. L1-MAGIC: Recovery of sparse signals via convex programming. Technical report, Caltech, 2005.

[4] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52:489–509, 2006.

[5] E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inacurrate measurements. *Communications on Pure and Applied Mathematics*, 59:1207–1223, 2006.

[6] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52:1289–1306, 2006.

[7] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye. k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI. *Magnetic Resonance in Medicine*, 61:103–116, 2009.

[8] Y.-C. Kim, S. S. Narayanan, and K. S. Nayak. Accelerated three-dimensional upper airway MRI using compressed sensing. *Magnetic Resonance in Medicine*, 61:1434–1440, 2009.

[9] D. Liang, B. Liu, J. Wang, and L. Ying. Accelerating SENSE using compressed sensing. *Magnetic Resonance in Medicine*, 62:1574–1584, 2009.

[10] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58:1182–1195, 2007.

[11] M. Marim, E. Angelini, and J.-C. Olivo-Marin. A compressed sensing approach for biological microscopic image processing. In *ISBI*, pages 1374–1377. IEEE, 2009.

[12] M. Marim, E. Angelini, and J.-C. Olivo-Marin. Compressed sensing in microscopy with random projections in the Fourier domain. In *ICIP*, pages 2121–2124. IEEE, 2009.

[13] M. Marim, M. Atlan, E. Angelini, and J.-C. Olivo-Marin. Compressed sensing with off-axis frequency-shifting holography. *Optics Letters*, 35:871–873, 2010.

[14] M. Murphy, K. Keutzer, S. Vasanawala, and M. Lustig. Clinically feasible reconstruction time for L1-SPIRiT parallel imaging and compressed sensing MRI. In *ISMRM*, 2010.

[15] Y. Nesterov. Gradient methods for minimizing composite objective function. Technical report, Université Catholique de Louvain, 2007.

[16] J. Provost and F. Lesage. The application of compressed sensing for photo-acoustic tomography. *IEEE Transactions on Medical Imaging*, 28:585–594, 2009.

[17] J. W. Stayman and J. A. Fessler. Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction. *IEEE Transactions on Medical Imaging*, 19:601–615, 2000.

[18] E. van den Berg and M. P. Friedlander. Probing the Pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 31:890–912, 2008.

[19] Z. Wang. *New sampling and detection approaches for compressed sensing and their application to ultra wideband communications*. PhD thesis, University of Delaware, January 2010.

[20] J. Yang, Y. Zhang, and W. Yin. A fast TVL1-L2 minimization algorithm for signal reconstruction from partial Fourier data. Technical report, Rice University, 2008.