A compressive sensing based transmissive single-pixel camera

Filipe Magalhães[a,b], Mehrdad Abolbashari[c,d], Faramarz Farahi[c,d], Francisco M. Araújo[a], Miguel V. Correia[a,b*]

[a] INESC Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal;
[b] Universidade do Porto, Faculdade de Engenharia, Departamento de Engenharia Electrotécnica e de Computadores, Rua Dr. Roberto Frias, s/n 4200-465 Porto, Portugal;
[c] Center for Optoelectronics and Optical Communications, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
[d] Department of Physics and Optical Science, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

ABSTRACT

Compressive sensing (CS) has recently emerged and is now a subject of increasing research and discussion, undergoing significant advances at an incredible pace.

The novel theory of CS provides a fundamentally new approach to data acquisition which overcomes the common wisdom of information theory, specifically that provided by the Shannon-Nyquist sampling theorem. Perhaps surprisingly, it predicts that certain signals or images can be accurately, and sometimes even exactly, recovered from what was previously believed to be highly incomplete measurements (information).

As the requirements of many applications nowadays often exceed the capabilities of traditional imaging architectures, there has been an increasing deal of interest in so-called computational imaging (CI). CI systems are hybrid imagers in which computation assumes a central role in the image formation process.

Therefore, in the light of CS theory, we present a transmissive single-pixel camera that integrates a liquid crystal display (LCD) as an incoherent random coding device, yielding CS-typical compressed observations, since the beginning of the image acquisition process.

This camera has been incorporated into an optical microscope and the obtained results can be exploited towards the development of compressive-sensing-based cameras for pixel-level adaptive gain imaging or fluorescence microscopy.

Keywords: Compressive Sensing, Single-pixel Camera, Computational Imaging, ℓ_1-norm, convex optimization

1. INTRODUCTION

It is clear that the Nyquist-Shannon sampling theorem has been a fundamental rule of signal processing for many years and can be found in nearly all signal acquisition protocols, being extensively used from consumer video and audio electronics to medical imaging devices or communication systems. Basically, it states that a band-limited input signal can be recovered without distortion if it is sampled at a rate of at least twice the highest frequency component of interest within the signal. For some signals, such as images that are not naturally band limited, the sampling rate is dictated not by the Nyquist-Shannon theorem but by the desired temporal or spatial resolution. However, it is common in such systems to use an anti-aliasing low-pass filter to band limit the signal before sampling it, and so the Nyquist-Shannon theorem plays an implicit role [1].

In the last few years, an alternative theory has emerged, showing that super-resolved signals and images can be reconstructed from far fewer data or measurements than what is usually considered necessary. This is the main concept of compressive sensing (CS), also known as compressed sensing, compressive sampling and sparse sampling. In fact,

* Further author information: (Send correspondence to Filipe Magalhães)
e-mail: filipe.magalhaes@inescporto.pt, Telephone: +351 22 209 42 42

“the theory was so revolutionary when it was created a few years ago that an early paper outlining it was initially rejected on the basis that its claims appeared impossible to substantiate [2].”

CS relies on the empirical observation that many types of signals or images can be well approximated by a sparse expansion in terms of a suitable basis, that is, by only a small number of non-zero coefficients. This is the key aspect of many lossy compression techniques such as JPEG and MP3, where compression is achieved by simply storing only the largest basis coefficients.

In CS, since the number of samples taken is smaller than the number of coefficients in the full image or signal, converting the information back to the intended domain would involve solving an underdetermined matrix equation. Thus, there would be a huge number of candidate solutions and, as a result, we must find a strategy to select the “best” solution.

Different approaches to recover information from incomplete data sets have existed for several decades. One of its earliest applications was related with reflection seismology, in which a sparse reflection function (indicating meaningful changes between surface layers) was sought from band limited data [1, 3, 4]. It was, however, very recently, that the field has gained increasing attention, when Emmanuel J. Candès, Justin Romberg and Terence Tao [5], discovered that it was possible to reconstruct Magnetic Resonance Imaging (MRI) data from what appeared to be highly incomplete data sets in face of the Nyquist-Shannon criterion (see Fig. 1). Following Candès *et al.* work, this decoding or reconstruction problem can be seen as an optimization problem and be efficiently solved using the ℓ_1-norm [6].

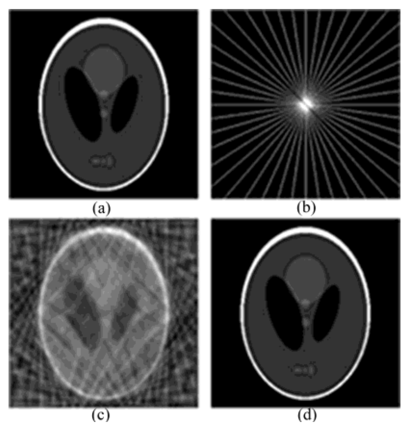

Figure 1 - Example of a simple recovery problem. (a) The Logan–Shepp phantom test image. (b) Sampling domain in the frequency plane; Fourier coefficients are sampled along 22 approximately radial lines. (c) Minimum energy reconstruction obtained by setting unobserved Fourier coefficients to zero. (d) Compressive sensing based reconstruction. This reconstruction is an exact replica of the image in (a) [5].

As a result, CS has become a kind of revolutionary research topic that draws from diverse fields, such as mathematics, engineering, signal processing, probability and statistics, convex optimization, random matrix theory and computer science.

Undergoing significant advances, CS has proved to be far reaching and has enabled several applications in many fields, such as: distributed source coding in sensor networks [7, 8], coding, analog–digital (A/D) conversion, remote wireless sensing [1, 9] and inverse problems, such as those presented by MRI [10].

The current paper is organized as follows: after a brief introduction (section 1), some mathematical background essential to the understanding of CS is shown (section 2); then, in section 3, CS is presented along with some of its principal properties and it is explained why ℓ_1-norm is such a good option for compressive sensing; next, in section 4 the innovative transmissive single-pixel camera is presented; afterwards, experimental results obtained with this single-pixel camera are shown and discussed; in the end, the main conclusions of this work are exposed.

2. COMPRESSIVE SENSING BACKGROUND

In order to become possible, CS is built upon two principles: sparsity, related with the signals of interest, and incoherence, related with the sensing modality.

2.1 K-sparse and compressible signals

Let's consider a real-valued, finite-length, one dimensional, discrete-time signal x, which can be viewed as a $N \times 1$ column vector in $\Re^{\mathrm{N}}$ with elements $x[n]$, with $n = 1,2,\ldots,\mathrm{N}$. Any signal in $\Re^{\mathrm{N}}$ can be represented in terms of a basis of $N \times 1$ vectors $\{\psi_i\}_{i=1}^{\mathrm{N}}$. For simplicity, let's assume that the basis is orthonormal. Using the $N \times N$ basis matrix $\Psi = \{\psi_1, \psi_2, \ldots, \psi_{\mathrm{N}}\}$ with the vectors $\{\psi_i\}$ as columns, a signal x can be expressed as:

$$x = \sum_{i=1}^{\mathrm{N}} s_i \psi_i \text{ or } x = \Psi s, \tag{1}$$

where s is the $N \times 1$ column vector of weighting coefficients $s_i = \langle x, \psi_i \rangle = \psi_i^T x$. s and x are equivalent representations of the signal with x in time or space domain and s in Ψ domain [11].

The signal x is K-sparse if it is a linear combination of only K basis vectors, which means that only K of the s_i coefficients in Eq. (1) are non-zero, while the remaining $(N-K)$ coefficients are zero. In addition, the signal x is compressible if the representation in Eq. (1) has just a few large coefficients and many small coefficients, setting the basis of transform coding. Therefore, we can say that a signal x is sparse in the Ψ domain if the coefficient sequence is supported on a small set, and compressible if the sequence is concentrated near a small set.

In face of the typical data acquisition paradigm, huge amounts of data are collected only to be in large part discarded at the compression stage to facilitate storage and transmission. Imagine, for example, a digital camera that captures images with millions of pixels but eventually encodes the image in just a few hundred kilobytes. Clearly, this is a tremendously wasteful process and suffers from three principal drawbacks. First, the initial number of samples N may be large, even if the desired K is small. Second, the set of all N transform coefficients $\{s_i\}$ must be computed even though all but K of them will be discarded. Third, there is an overhead that is introduced by the encoding of the large coefficients locations [1, 11].

2.2 Recovering K-sparse signals

Following the work presented in [12], Candès and Tao developed a refined version of the *Uniform Uncertainty Principle* (UUP) [13], which has proved to be essential to the study of the general robustness of CS. This key notion was then named *Restricted Isometry Property* (RIP) and can be defined as follows:

For each integer $K = 1,2,\ldots$, define the isometry constant δ_K of a measurement matrix A as the smallest number such that

$$(1-\delta_K)\|x\|_{\ell_2}^2 \le \|Ax\|_{\ell_2}^2 \le (1+\delta_K)\|x\|_{\ell_2}^2 \tag{2}$$

holds for all K-sparse vectors x. Therefore, we can say that a matrix A obeys the RIP of order K if δ_K differs enough from one. When this condition is verified, A approximately preserves the Euclidean length of K-sparse signals, which in turn implies that K-sparse vectors cannot be in the null space of A. An alternative description of this property is to say that all subsets of K columns taken from A are in fact nearly orthogonal (they cannot be exactly orthogonal since we have more columns than rows).

Let's imagine we want to acquire K-sparse signals making use of matrix A. Suppose that δ_{2K} is sufficiently smaller than one. This indicates that all pair-wise distances between K-sparse signals must be well preserved in the measurement space, which means that

$$(1-\delta_{2K})\|x_1 - x_2\|_{\ell_2}^2 \le \|Ax_1 - Ax_2\|_{\ell_2}^2 \le (1+\delta_{2K})\|x_1 - x_2\|_{\ell_2}^2 \tag{3}$$

is true for all K-sparse vectors x_1, x_2 [1, 11].

2.3 Incoherence

Let's now consider $M < N$ linear measurements of x and a collection of test functions $\{\varphi_m\}_{m=1}^{M}$ such that $y[m] = \langle x, \varphi_m \rangle$. By stacking the measurements $y[m]$ into the $M \times 1$ vector y and the test functions φ_m^T as rows into an $M \times N$ sensing matrix Φ we can write

$$y = \Phi x = \Phi \Psi s = \Theta s. \tag{4}$$

A condition related with RIP is *incoherence*, which requires that the rows of Φ (the measurement or sensing matrix) cannot represent the columns of Ψ in a sparse way (and vice-versa).

Incoherence extends the duality between time and frequency and expresses the idea that an object having a sparse representation in Ψ must be spread out in the domain in which it was acquired. This incoherence property is verified for many pairs of bases, including, for instance, delta spikes and sine waves of the Fourier basis, or the Fourier basis and wavelets.

The coherence between the sensing basis Φ and the representation basis Ψ can be given by the following equation:

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1 \le k, j \le n} \left| \langle \varphi_k, \psi_j \rangle \right|, \tag{5}$$

which, in simple words, is measuring the largest correlation between any two elements of Φ and Ψ. CS is essentially interested in low coherence pairs. For instance, for the previously referred delta spikes and sine waves (time-frequency) pair, $\mu(\Phi, \Psi) = 1$, therefore, indicating maximal incoherence [1, 14].

A particular aspect of interest is that random matrices are largely incoherent with any fixed basis Ψ. This empowers the use of known fast transforms such as a Walsh, Hadamard, or Noiselet transform [15].

Furthermore, what is most remarkable about this concept is that it allows capturing information contained in a sparse signal in a very efficient way without the need to understand that signal.

3. HOW COMPRESSIVE SENSING WORKS

Compressive sensing addresses the inefficiencies presented by the *sample-then-compress* framework by directly acquiring a compressed signal representation, avoiding the intermediate stage of acquiring N samples [5]. CS bypasses the sampling process and directly acquires a condensed representation y consisting of M linear measurements. Furthermore, the measurement process is nonadaptive in that Φ does not depend on the signal x. The transformation from x to y is a dimensionality reduction and so loses information in general. In particular, since $M < N$, for a given y, there is an infinite number of x' such that $\Phi x' = y$, if there is no restriction on x'. The overwhelming capacity of CS is that Φ can be designed such that sparse/compressible x can be recovered exactly/approximately from measurements of y. To recover the signal x from the random measurements y, the traditional favorite method of least squares has been shown to fail with high probability. Instead, it has been demonstrated that using the ℓ_1 optimization [12]

$$\hat{s} = \arg\min \|s'\|_{\ell_1} \text{ such that } \Theta s' = y \tag{6}$$

it is possible to exactly reconstruct K*-sparse* vectors and closely approximate compressible vectors stably with high probability using just $M \ge O(K \cdot \log(N/K))$ random measurements [5, 6]. Minimizing the ℓ_1-norm subject to linear equality constraints can easily be recast as a linear program, also known as *basis pursuit*, which can find several alternative reconstruction techniques based on greedy, stochastic and variational algorithms [5, 9, 16, 17].

3.1 The geometry of ℓ_1-norm

The geometry of CS problems in $\Re^{\mathrm{N}}$ helps to visualize why ℓ_2 reconstruction fails to find the sparse solution that can be identified by ℓ_1 reconstruction. Figure 2 presents significant information to this subject. Part (a) illustrates the ℓ_2 ball in $\Re^3$ with a certain radius. It must be emphasized that this ball is isotropic. Part (b) represents the ℓ_1 ball in $\Re^3$, which is anisotropic ("pointy" along the axes).

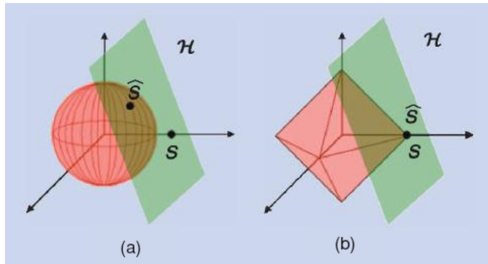

Figure 2 - Geometry of ℓ_1 recovery. (a) Visualization of the ℓ_2 minimization that does not find the sparse point of contact $\hat{s}$ between the ℓ_2 ball (hypersphere, in red) and the translated measurement matrix null space (in green). (b) Visualization of the ℓ_1 minimization solution that finds the sparse point of contact $\hat{s}$ with high probability thanks to the pointiness of the ℓ_1 ball [14].

The ℓ_2 minimizer $\hat{s}$ is the point from H closest to the origin. This point can be found by blowing the ℓ_2 ball until it bumps into H. Due to the random orientation of H (imposed by the randomness in matrix Θ), the closest point $\hat{s}$ will be away from the coordinate axes with high probability and, therefore, will not be sparse and will be far from the sparsest answer s (only one of its components is non-zero). In higher dimensions, this difference becomes even more significant. Paying attention to the part (b) of Figure 2, it can be seen that the point of intersection $\hat{s}$ is now defined by the vector that solves equation (6).

4. TRANSMISSIVE SINGLE-PIXEL CAMERA

Following the work we had already developed [18], it was our intent to extend its scope to the development of alternative configurations that would disregard the need of an active illumination source and operate in a transmissive mode rather than reflective. For that purpose, we idealized the use of a liquid crystal display (LCD) as the spatial modulation device for the light coming from the scene being imaged.

The scheme of the developed camera is depicted in Figure 3.

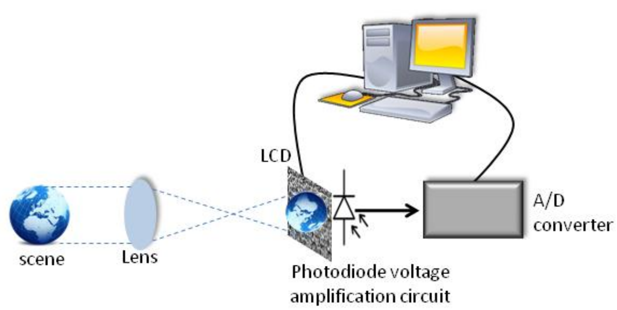

Figure 3 – Scheme of the transmissive single-pixel camera.

As it can be seen in Figure 3, by the means of a lens system it is formed an image of the scene being acquired on the surface of the LCD, behind which lies right the photodetector (our single-pixel) active region. This photodetector is part of a voltage amplification circuit whose output is connected to an analog-to-digital (A/D) converter that enables data acquisition by a computer for subsequent processing. The LCD is also controlled by the computer. Basically, what our system does is to apply random binary codes to the LCD and measure the corresponding output voltage from the amplification circuit. As we have already seen, based on CS theory, this process can be repeated K times, with K much smaller than the image full dimensionality. Afterwards, using combinatorial optimization algorithms, it is reconstructed the image that in conjunction with the used codes gave rise to the corresponding voltage measurements.

This transmissive single-pixel camera was then exploited on an optical microscope (see Figure 4), since its optical system is well known and allows an easy configuration of both the optical magnification and working distance. At the same time, with this microscope we could mount a conventional camera on the other optical path to acquire conventional images of the scenes being imaged. This was of great help while focusing and aligning the system, and also because we could easily acquire images for comparison with the results of the implemented single-pixel camera.

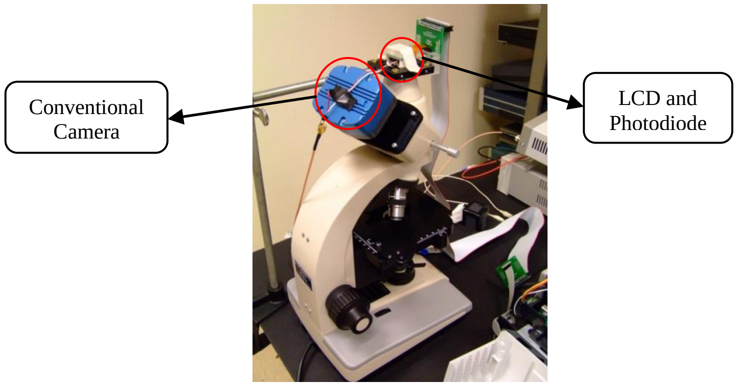

Figure 4 – Transmissive single-pixel camera mounted on an optical microscope, placed on the vertical optical path. Note the conventional camera (in blue) mounted on the other optical path.

The LCD, having a maximum resolution of 800 × 600 pixels (active area: 11.2 mm × 8.4 mm = 94.08 mm^2), was taken from an Epson® PowerLite S5 projector. The central part of the LCD (512 × 512 pixels) was used to apply the random codes. The photodiode (Si) was a Thorlabs FDS1010 (active area: 9.7 mm × 9.7 mm = 94.09 mm^2). The active areas values confirm the suitability of the LCD and the chosen photodiode to be used together to spatially encode and measure the light, respectively.

5. RESULTS AND DISCUSSION

One of the scenes acquired with our imaging system is represented in Figure 5. Since the conventional camera sensor size (1/2-inch CCD) was smaller than the LCD size, we had to stitch 4 separate pictures taken with the conventional camera in order to get an image containing the single-pixel camera field of view. For the case of the single-pixel camera, ideally, its field of view would correspond to the size of the LCD. In particular, the actual field of view was defined by a 512 × 512 pixels region centered on the LCD active region.

Figure 5 – Image of one of the acquired scenes. This image was obtained from the stitching of 4 separate pictures taken with the conventional camera due to its reduced sensor size (1/2-inch CCD) when compared to the size of the LCD. The red inset indicates the region acquired with our single-pixel camera.

Following, some results obtained with the setup of Figure 4 are presented.

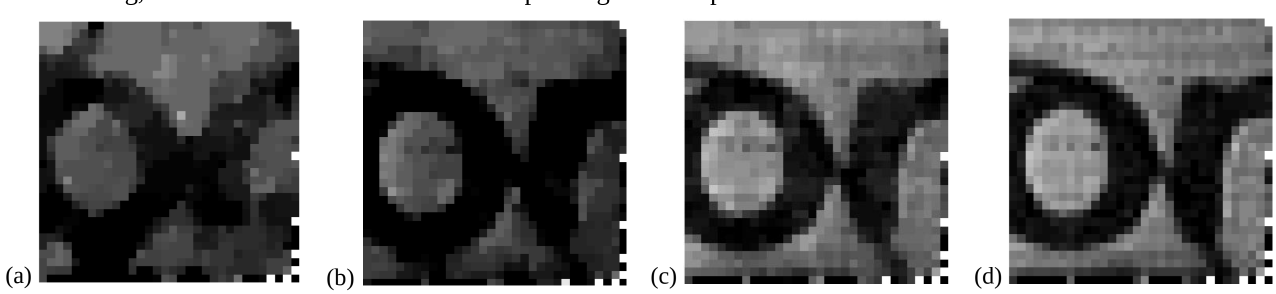

Figure 6 – Reconstruction of an image with 32 × 32 pixels (N = 1024) from: (a) 25% (K = 256); (b) 50% (K = 512); (c) 75% (K = 768); (d) 100% (K = 1024) measurements. For each reconstructed image, the PSNR has been calculated relatively to the image reconstructed using 100% (K = 1024) of the measurements: (a) PSNR = 12.85 dB; (b) PSNR = 14.34 dB; (c) PSNR = 25.89 dB.

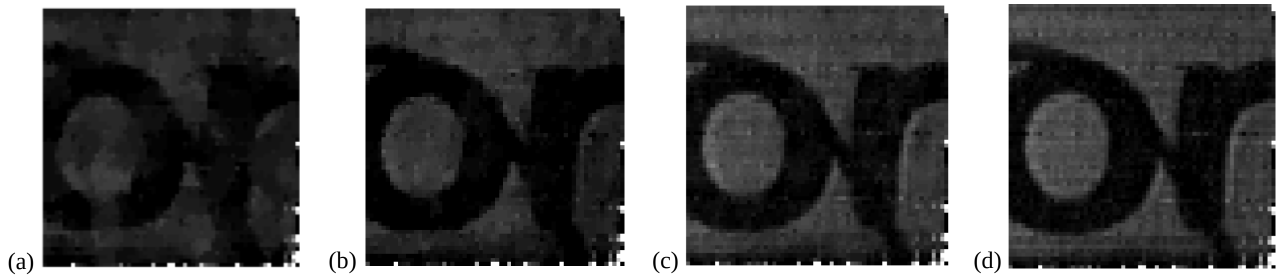

Figure 7 – Reconstruction of an image with 64 × 64 pixels (N = 4096) from: (a) 25% (K = 1024); (b) 50% (K = 2048); (c) 75% (K = 3072); (d) 100% (K = 4096) measurements. For each reconstructed image, the PSNR has been calculated relatively to the image reconstructed using 100% (K = 4096) of the measurements: (a) PSNR = 20.27 dB; (b) PSNR = 22.63 dB; (c) PSNR = 32.44 dB.

The results presented in Figure 6 and Figure 7, correspond to reconstructions of the scene contained within the red inset of Figure 5. From the analysis of these results, it is clear that the quality of the reconstructed images increases with the growth of the number of measurements used for reconstruction, as well as with the dimensionality of the reconstructed image. These facts are supported by the PSNR (Peak Signal-to-Noise Ratio) values presented for each of the reconstructed images relatively to the image of the same size reconstructed using a number of measurements equal to its full dimensionality, i.e., $K = N$. Also, as expected, the images with higher resolution exhibit smoother contours and are more perceptible.

As stated previously, random matrices are largely incoherent with any fixed basis Ψ and, therefore, for the measurement or sensing matrices, we have used Hadamard-based random binary codes. This choice revealed to be of great importance because the Hadamard-based random binary codes could be built *on-the-run* during the measurement and reconstruction phases, avoiding the memory waste that otherwise would be required to save all the matrices.

Using the NESTA software package[†] it took, in average, 0.8 seconds to reconstruct a 32 × 32 pixels image from 410 (~40%) measurements, and 1.8 seconds to reconstruct a 64 × 64 pixels image from 1640 (~40%) measurements. These results were obtained with MATLAB® on Windows Vista™ with an Intel® Core™2 Duo CPU @ 2.50GHz and 3GB of RAM.

6. CONCLUSIONS

In this work the revolutionary theory of Compressive Sensing has been described along with its principal concepts, providing a background to the understanding of the main subject here presented – a compressive sensing based transmissive single-pixel camera.

One of the advantages of the proposed single-pixel camera is the ability to operate under very low light intensities, much lower than those required by conventional cameras. This is due to the fact that usually a single photodetector exhibits much higher sensitivity than the pixels of a conventional image sensor. This fact will be further exploited in applications such as fluorescence microscopy, where light intensities in the wavelengths of interest are usually minute. Furthermore, for instance, in fluorescence microscopy and high dynamic range imaging, pixel bleeding is often a problem. With a camera like the one presented here we have the capability of pixel-wise controlling the gain, thus extending its application to these scenarios.

Another advantage relies on the fact that the information is acquired in an already compressed form. This avoids the waste of information often verified along the use of conventional imaging systems, which gather huge amounts of information that is then wasted through the application of compression standards, such as JPEG.

Since data are compressed since the beginning of the process, an efficient encryption method is brought into light, since the apparently random measurements will resemble noise and have no meaning for an observer that has no knowledge about their seed.

[†] This package is a collection of MATLAB® routines for solving the convex optimization programs central to compressive sensing and is available at http://www.acm.caltech.edu/~nesta/

In these approaches, most of the burden is placed on the reconstruction process, reducing the complexity of the hardware and acquisition phase. This constitutes another advantage since the required computational resources are widely available and with increasing tendency to perform better without significant increase of their costs.

The results obtained with our camera have corroborated the initial assumptions about its principle of operation and proved to be promising and motivating towards new developments.

Furthermore, in the future, our efforts will also be directed towards the refinement of the light source since the microscope's halogen lamp exhibits an emission spectrum that extends up to the near-infrared region which cannot be filtered by the liquid crystals. This reduces the dynamic range we can obtain with the random binary codes applied to the LCD, therefore reducing the contrast of the reconstructed images.

ACKNOWLEDGMENTS

This work was supported by a PhD grant (SFRH/BD/45380/2008) from Fundação para Ciência e Tecnologia. The author would also like to acknowledge Fundação Calouste Gulbenkian for the financial support provided.

REFERENCES

1. Candès, E.J. and M.B. Wakin, *An Introduction To Compressive Sampling.* Signal Processing Magazine, IEEE, 2008. **25**(2): p. 21-30.
2. Kirk, L.K., *Rethinking signal processing.* Commun. ACM, 2009. **52**(5): p. 13-15.
3. Claerbout, J.F. and F. Muir, *Robust modeling with erratic data.* Geophysics, 1973. **38**(5): p. 826-844.
4. Fadil, S. and W.S. William, *Linear Inversion of Band-Limited Reflection Seismograms*. 1986, SIAM. p. 1307-1330.
5. Candès, E.J., J. Romberg, and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information.* Information Theory, IEEE Transactions on, 2006. **52**(2): p. 489-509.
6. Candès, E., J. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements.* Communications on Pure and Applied Mathematics, 2006. **59**(8): p. 1207-1223.
7. D. Baron, et al., *Distributed Compressive Sensing [preprint]*. 2005.
8. Golbabaee, M. and P. Vandergheynst, *Distributed Compressed Sensing for Sensor Networks, Using p-thresholding*. 2009, HAL - CCSD.
9. Haupt, J. and R. Nowak, *Signal Reconstruction From Noisy Random Projections.* Information Theory, IEEE Transactions on, 2006. **52**(9): p. 4036-4048.
10. M. Lustig, D. L. Donoho, and J.M. Pauly. *Rapid MR imaging with compressed sensing and randomly under-sampled 3DFT trajectories*. in *14th Ann. Meeting ISMRM*. 2006. Seattle.
11. Candès, E.J., *Compressive sampling*, in *Int. Cong. Mathematicians*. 2006: Madrid, Spain. p. 1433-1452.
12. Candès, E.J. and T. Tao, *Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?* Information Theory, IEEE Transactions on, 2006. **52**(12): p. 5406-5425.
13. Candès, E.J. and T. Tao, *Decoding by Linear Programming.* IEEE Transactions on Information Theory, 2005. **51**(12): p. 4203-4215.
14. Baraniuk, R.G., *Compressive Sensing [Lecture Notes]*. Signal Processing Magazine, IEEE, 2007. **24**(4): p. 118-121.
15. Candès, E.J. and J. Romberg, *Sparsity and incoherence in compressive sampling.* Inverse Problems, 2007. **23**(3): p. 969-985.
16. Chen, S., D. Donoho, and M. Saunders, *Atomic Decomposition by Basis Pursuit.* SIAM Review, 2001. **43**(1): p. 129-159.
17. Tropp, J.A. and A.C. Gilbert, *Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit.* Information Theory, IEEE Transactions on, 2007. **53**(12): p. 4655-4666.
18. Magalhães, F., et al., *Active illumination single-pixel camera based on compressive sensing.* Appl. Opt., 2011. **50**(4): p. 405-414.