Random Sensing of Geometric Images

Gabriel Peyré

RANDOM SENSING OF GEOMETRIC IMAGES

Gabriel Peyré
CMAP, École Polytechnique
92128 Palaiseau CEDEX, France
gabriel.peyre@polytechnique.fr

ABSTRACT
This paper proposes an extension of compressed sensing that allows to express the sparsity prior in a dictionary of bases. This enables the use of the universal sampling strategy of compressed sensing together with an adaptive recovery process that adapts the basis to the structure of the sensed signal. A fast greedy scheme is used during reconstruction to estimate the best basis using an iterative refinement. Numerical experiments on geometrical images show that adaptivity is indeed crucial to capture the structures of complex natural signals.

KEY WORDS
Compressed sensing, best basis, geometry of images.

1 Introduction

1.1 Classical Sampling vs. Compressed Sensing

The classical sampling theory of Shannon is based on uniform smoothness assumptions (low frequency spectral content). Under band limited condition, finely enough sampled functions can be recovered from a set of n pointwise measurements.

However, being bandlimited is not a good model for natural images and one usually assume that such a function f has a decomposition with few elements on some fixed orthogonal basis $\mathcal{B}$. This kind of assumption is at the heart of the compression of an image using a wavelet expansion. Under such sparseness assumption, one can hope to use a much smaller number $n < N$ of measurements, which are linear projections $\Phi f = \{\langle f, \phi_i \rangle\}_{i=1}^{n}$ on a set of fixed vectors $\phi_i \in \mathbb{R}^N$. The price to pay for this compressed sampling strategy is a non-linear reconstruction procedure to recover f from the compressed representation Φf. This theory of compressed acquisition of data has been pushed forward during last few years conjointly by Candès and Tao [2] and Donoho [5].

In order for this recovery to be effective, one needs sensing vectors ϕ_i that are incoherent with the vectors of $\mathcal{B}$. A convenient way to achieve this property is to use random vectors ϕ_i, which cannot be sparsely represented in basis $\mathcal{B}$.

Application in imaging. Compressed sensing acquisition of data could have an important impact for the design of imaging devices where data acquisition is expensive. For instance in seismic or magnetic imaging one could hope to use few random projections of the object to acquire together with a high precision reconstruction.

Analogies in physiology. This compressed sampling strategy could potentially lead to interesting models for various sensing operations performed biologically. Skarda and Freeman [12] have proposed a non-linear chaotic dynamic to explain the analysis of sensory inputs. This chaotic state of the brain ensures robustness toward unknown events and unreliable measurements, without using too many computing resources. While the theory of compressed sensing is presented here as a random acquisition process, its extension to deterministic or dynamic settings is a fascinating area for future research in signal processing.

1.2 The Best Basis Approach

Frames vs. dictionary of bases. Fixed orthogonal bases are not flexible enough to capture the complex redundancy of natural images. For instance the orthogonal wavelet transform [9] lacks of translation and rotation invariance and is not efficient to compress geometric images [7]. It is thus useful to consider families of vectors that are redundant but offer a stable decomposition. For instance, frames of translation invariant wavelets have been used for image denoising and frames of rotation-invariant Gabor functions are useful to characterize textures [9].

However, to capture the complex structure of natural images, one needs a very large number of such elementary atoms. Frame theory suffers from both theoretical difficulties (lack of stability) and technical problems (computational complexity) when the number of basis vectors increases too much. To cope with these problems, one can consider a dictionary $\mathcal{D} = \{\mathcal{B}^\lambda\}_{\lambda\in\Lambda}$ of orthogonal bases $\mathcal{B}^\lambda$. Choosing an optimal basis in such a dictionary allows to adapt the approximation to the complex content of a specific signal.

Images and geometry. Cartoon images is a simple model that captures the sketch content of natural images. Figure 1, (a), shows such a geometrically regular image, which contains smooth areas surrounded by regular curves. The curvelet frame of Candès and Donoho [1] can deal with such a regularity and enjoys a better approximation rate than traditional isotropic wavelets. This results can be enhanced using a dictionary of locally elongated functions that follow the image geometry. Bandelets bases of Le Pennec and Mallat [7] provide such a geometric dictionary together with a fast optimization procedure to compute a basis adapted to a given image.

Adaptive biological computation. Hubel and Wiesel have shown that low level computation done in area V1 of the visual cortex are well approximated by multiscale oriented linear projections [6]. Olshausen and Field proposed in [10] that redundancy is important to account for sparse representation of natural inputs. However further non-linear processings are done by the cortex to remove high order geometrical correlations present in natural images. Such computations are thought to perform long range groupings over the first layer of linear responses [8] and thus correspond to an adaptive modification of the overall neuronal response. The best-basis coding strategy could thus offer a signal-processing counterpart to this neuronal adaptivity. In this paper we restrict ourselves to orthogonal best basis search. Future extensions of this approach to dictionaries of redundant transforms are likely to improve numerical results and better cope with biological computations.

1.3 Best Basis Computation

In the following we use the norms $\|f\|_{\ell^p}^p = \sum_i |f[i]|^p$ and $\|f\|_{\ell^0} = \#\{i \setminus f[i] \neq 0\}$. A dictionary $\mathcal{D}_\Lambda = \{\mathcal{B}^\lambda\}_{\lambda\in\Lambda}$ is a set of orthogonal bases $\mathcal{B}^\lambda = \{\psi_m^\lambda\}_m$ of $\mathbb{R}^N$. We associate a cost $\mathrm{pen}(\lambda)$ which is a prior complexity measure on each basis $\mathcal{B}^\lambda$ that satisfies $\sum_{\lambda\in\Lambda} 2^{-\mathrm{pen}(\lambda)} = 1$. A fixed weight $\mathrm{pen}(\lambda) = \log_2(M)$ can be used if the size M of $\mathcal{D}$ is finite. The parameter $\mathrm{pen}(\lambda)$ can be interpreted as the number of bits needed to specify a basis $\mathcal{B}^\lambda$. Following the construction of Coifman and Wickerhauser [3], a best basis $\mathcal{B}^{\lambda^*}$ adapted to a signal minimizes a Lagrangian $\mathcal{L}$

$$\lambda^* = \underset{\lambda\in\Lambda}{\mathrm{argmin}}\ \mathcal{L}(f,\lambda,t) \qquad (1)$$

$$\text{where} \qquad \mathcal{L}(f,\lambda,t) = \|\Psi^\lambda f\|_{\ell^1} + C_0\, t\ \mathrm{pen}(\lambda),$$

where $\Psi^\lambda = [\psi_0^\lambda, \ldots, \psi_{N-1}^\lambda]^{\mathrm{T}}$ is the transform matrix defined by $\mathcal{B}^\lambda$. This best basis $\mathcal{B}^{\lambda^*}$ is thus the one that gives the sparsest description of f as measured by the ℓ^1 norm.

The Lagrange multiplier $C_0\, t$ weights the penalization $\mathrm{pen}(\lambda)$ associated to the complexity of a basis $\mathcal{B}^\lambda$. The parameter t is the level of deterministic noise caused by not having a perfectly k-sparse function but rather a function whose coefficients in basis $\mathcal{B}^{\lambda^*}$ decay fast. This parameter is not available since it depends on the decay speed and on k (which in turn depends on the number of sensed vectors). However the recovery algorithm presented in section 2.3 does not need the exact value of t since an upper bound t_s is estimated during iterations.

For a dictionary $\mathcal{D}$ that enjoys a multiscale structure, the optimization of $\mathcal{L}$ is carried out with a fast procedure, see [3, 7] for practical examples of this process.

2 Compressed Sensing Reconstruction

In this paper, the sampling matrix $\Phi = [\phi_0, \ldots, \phi_{n-1}]^{\mathrm{T}}$ is defined by random points $\{\phi_i\}_i$ of unit length although other random sensing schemes can be used, see [2].

2.1 Basis Pursuit Formulation

Searching for the sparsest signal f^* in some basis $\mathcal{B} = \{\psi_m\}_m$ that matches the sensed values $y = \Phi f$ leads to consider a penalized variational problem

$$f^* = \underset{g\in\mathbb{R}^N}{\mathrm{argmin}} \left(\frac{1}{2}\|\Phi g - y\|_{\ell^2}^2 + t\|\Psi g\|_{\ell^1} \right). \qquad (2)$$

where $\Psi = [\psi_0, \ldots, \psi_{N-1}]^{\mathrm{T}}$ is the transform matrix defined by $\mathcal{B}$. The Lagrange multiplier t accounts both for stabilisation against noise and for approximate sparsity, which is common in practical applications.

Candès and Tao in [2] and Donoho in [5] have shown that, if f is sparse enough in some basis $\mathcal{B}$, f can be recovered from the sensed data $y = \Phi f$. More precisely, they show that it exists a constant C such that if $\|\Psi f\|_{\ell^0} \leq k$ then if $n \geq C\ \log(N)\, k$ one recovers $f^* = f$.

2.2 Best Basis Pursuit

The compressed sensing machinery is extended to a dictionary of bases $\mathcal{D}_\Lambda$ by imposing that the recovered signal is sparse in at least one basis of $\mathcal{D}_\Lambda$. To avoid using too complex basis the recovery process from noisy measurements takes into account a complexity $\mathrm{pen}(\lambda)$ of the optimal basis $\mathcal{B}^\lambda$. The original recovery procedure (2) is replaced by the following minimization

$$\begin{aligned} f^* = \underset{g\in\mathbb{R}^N}{\mathrm{argmin}}\ \min_{\lambda\in\Lambda} \Big(&\tfrac{1}{2}\|\Phi g - y\|_{\ell^2}^2 + t\,\|\Psi^\lambda g\|_{\ell^1} \\ &+ C_0\, t^2\ \mathrm{pen}(\lambda) \Big), \end{aligned} \qquad (3)$$

where the penalization $C_0\, t^2\ \mathrm{pen}(\lambda)$ is the same as in equation (1).

2.3 Iterative Thresholding For Sparsity Minimization

The recovery procedure suggested by equation (3) corresponds to the inversion of the operator Φ under sparsity constraints on the observed signal f. In this paper we extends the algorithm of Daubechies et al. [4] to the setting of a dictionnary of bases. Searching in the whole dictionary $\mathcal{D}$ for the best basis that minimizes formulation (3) is not feasible for large dictionaries, which typically contain of the order of 2^N bases. Instead we propose a greedy search for the best basis during the recovery process. We make repeated use of the soft thresholding operator defined in an orthogonal basis $\mathcal{B} = \{\psi_m\}_m$ by $S_t(g) = \sum_m s_t(\langle g,\, \psi_m\rangle)\, \psi_m$ where $s_t(x) = \mathrm{sign}(x)(|x| - t)_+$ in the following algorithm

- **Initialization.** Set $s = 0$, $f_0 = 0$ and choose $\lambda_0 \in \Lambda$ at random or using some default choice (such as a DCT basis in 1D or a wavelet basis in 2D).
- **Step 1: Updating the estimate.** Enforce condition $y = \Phi f_{s+1}$ using $f_{s+1} = f_s + \Phi^{\mathrm{T}}(y - \Phi\, f_s)$.
- **Step 2: Denoising the estimate.** Compute the value of the current threshold $t_s = 3\sigma_s$ using the estimator of the noise level $\sigma_s = \mathrm{median}(|\Psi f_{s+1}|)/0.6745$. Perform the denoising $f_{s+1} = S_{t_s}(f_{s+1})$ where S_{t_s} is the threshold operator at t_s in the basis $\mathcal{B}^{\lambda_s}$.
- **Step 3: Update best basis.** Update the best basis using $\lambda_{s+1} = \mathrm{argmin}_\lambda\ \mathcal{L}(f_{s+1}, \lambda, t_s)$. For typical dictionaries such as the ones considered in this paper, this minimization is carried out with a fast procedure, as seen in subsection 1.3.
- **Stopping criterion.** If $s < s_{\max}$, go to step 1, otherwise stop iterations. In all our experiments, the number of iterations is set to $s_{\max} = 20$.

The stochastic noise level t_s is due to random acquisition corrupting the current estimate f_s. It is computed though the median estimator and its value decays during iterations toward the deterministic noise value t that is due to approximate sparsity.

3 Best Bandelet Basis Compressed Sensing

3.1 Adapted Bandelet Transform

The bandelet bases dictionary was introduced by Le Pennec and Mallat [7] to perform adaptive approximation of images with geometric singularities, such as the cartoon image in figure 1, (a). We present a simple implementation of the bandelet transform inspired from [11].

A bandelet basis $\mathcal{B}^\lambda$ is parameterized by $\lambda = (Q, \{\theta_S\}_{S\in Q})$, where Q is a quadtree segmentation of the pixels locations and $\theta_S \in [0, \pi[\cup \Xi$ is an orientation (or the special token Ξ) defined over each square S of the segmentation, see figure 1, (a). The bandelet transform corresponding to this basis applies independently over each square S of the image either

- if $\theta_S = \Xi$: a 2D isotropic wavelet transform,
- if $\theta_S \neq \Xi$: a 1D wavelet transform along the direction defined by the angle θ_S.

We now detail the latter transform. The position of a pixel $x = (x_1, x_2) \in S$ with respect to the direction θ_S is $p_x = \sin(\theta_S)x_1 - \cos(\theta_S)x_2$. The m pixels $\{x^{(i)}\}$ in S are ranked according to the 1D ordering $p_{x^{(0)}} \leq p_{x^{(1)}} \leq \ldots \leq p_{x^{(m-1)}}$. This ordering allows to turn the image $\{f[x]\}_{x\in S}$ defined over S into a 1D signal $f_{\mathrm{1D}}[i] = f[x^{(i)}]$, see figure 1, (c). The bandelet transform of the image f inside S is defined as the 1D Haar transform of the signal f_{1D}, see figure 1, (d). This process is both orthogonal and easily invertible, since one only needs to compute the inverse Haar transform and pack the retrieved coefficients at the original pixels locations. Keeping only a few bandelet coefficients and setting the others to zero performs an approximation of the original image that follows the local direction θ_S, see figure 1 (f).

In order to restrict the number of tested geometries θ for a square $S \in Q$ containing $\#S$ pixels, we follow [11] and use the set of directions that pass through two pixels of S. The number of such directions is of the order of $(\#S)^2$. For this bandelet dictionary, the penalization of a basis $\mathcal{B}^\lambda$ where $\lambda = (Q, \{\theta_S\}_S)$ is defined as $\mathrm{pen}(\lambda) = \#Q + \sum_{S\in Q} 2\log_2(\#S)$, where $\#Q$ is the number of leaves in Q. A fast best basis search, described in [11], allows to define a segmentation Q and a set of directions $\{\theta_S\}_S$ adapted to a given image f by minimizing (1). This process segments the image into squares S on which f is smooth, thus setting $\theta_S = \Xi$ and squares containing an edge, where θ_S closely matches the direction of this singularity.

3.2 Numerical Results

The geometric image depicted in figure 2, (a) is used to compare the performance of the original compressed sensing algorithm in a wavelet basis to the adaptative algorithm in a best bandelet basis. Since the wavelet basis is not adapted to the geometric singularities of such an image, reconstruction (b) has strong ringing artifacts. The adapted reconstruction (c) exhibits fewer such artifacts since the bandelet basis functions are elongated and follow the geometry. The segmentation is depicted after the last iteration, together with the chosen direction θ_S which closely matches the real geometry. On figure 2, (d/e/f), one can see a comparison for a natural image containing complex geometric structures such as edges, junctions and sharp line features. The best bandelet process is able to resolve these features efficiently.

4 Conclusion

Using a dictionary of bases decouples the approximation process from the redundancy needed for adaptivity and requires the design of a penalization cost on the set of bases. This framework is not restricted to orthogonal bases, although it is a convenient mathematical way to ensure the compressed sensing recovery condition. This best basis approach to sensing and recovery is also a promising avenue for interactions between biological processing, where a deterministic or chaotic process is highly probable and signal processing, where randomization has proven useful to provide universal coding strategies.

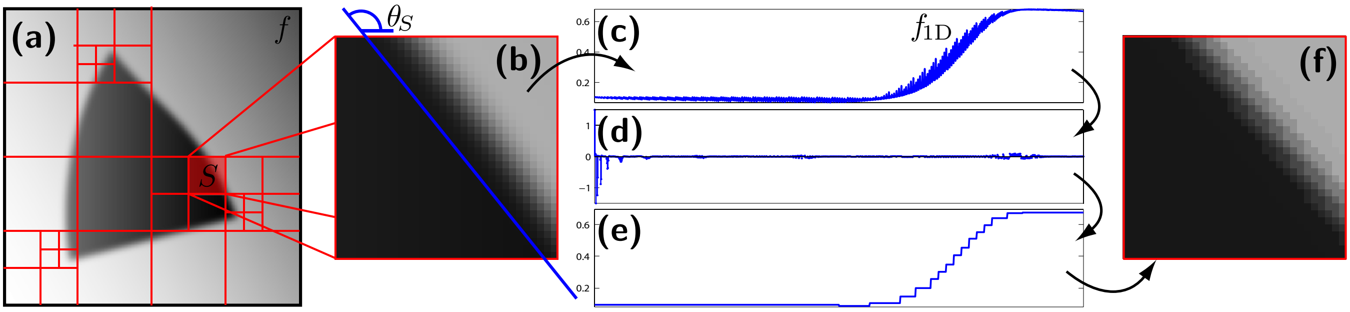

Figure 1. *(a) a geometric image together with some adapted dyadic segmentation Q ; (b) a square S together with some adapted direction θ_S ; (c) the 1D signal f_{1D} obtained by mapping the pixels values $f(x^{(i)})$ on a 1D axis ; (d) the 1D Haar coefficients of f_{1D} ; (e) the 1D approximation obtained by reconstruction from the 20 largest Haar coefficients ; (f) the corresponding square approximated in bandelet.*

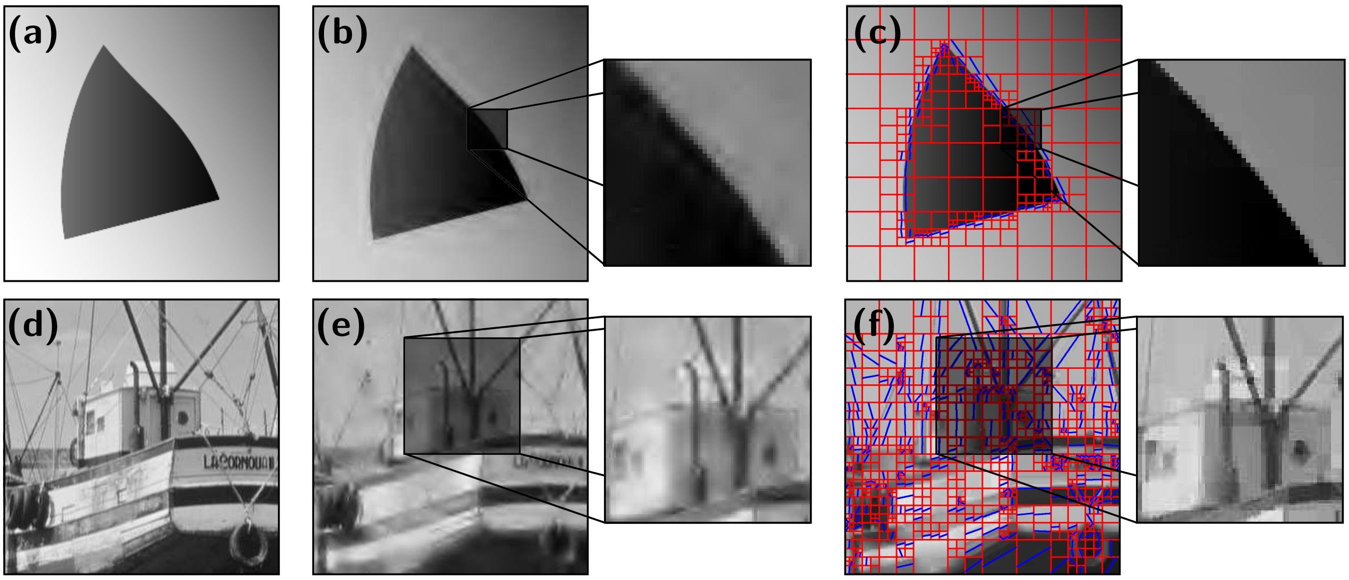

Figure 2. *(a/d) original image ; (b/e) compressed sensing reconstruction using the wavelet basis, $n = N/6$ (b: PSNR=22.1dB, b: PSNR=23.2dB) ; (c/f) reconstruction using iteration in a best bandelet basis (c: PSNR=24.3dB, f: PSNR=25.1dB).*

References

[1] E. Candès and D. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.*, 57(2):219–266, 2004.

[2] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 2004. Submitted.

[3] R. Coifman and V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Trans. Inform. Theory*, IT–38(2):713–718, Mar. 1992.

[4] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math*, 57:1413–1541, 2004.

[5] D. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[6] D. Hubel and T. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology (London)*, 195:215–243, 1968.

[7] E. Le Pennec and S. Mallat. Bandelet Image Approximation and Compression. *SIAM Multiscale Modeling and Simulation*, 4(3):992–1039, 2005.

[8] T. S. Lee. Computations in the early visual cortex. *J Physiol Paris*, 97(2-3):121–139, Mar-May 2003.

[9] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1998.

[10] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive-field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, June 1996.

[11] G. Peyré and S. Mallat. Surface compression with geometric bandelets. *ACM Transactions on Graphics, (SIGGRAPH'05)*, 24(3), Aug. 2005.

[12] C.A. Skarda and W.J. Freeman. Does the brain make chaos in order to make sense of the world? *Behavioral and Brain Sciences*, 10, 1987.