Best Basis Compressed Sensing

Gabriel Peyré

Best Basis Compressed Sensing

Gabriel Peyré

Ceremade, Université Paris Dauphine,
Place du Marchal De Lattre De Tassigny,
75775 Paris Cedex 16 France
gabriel.peyre@ceremade.dauphine.fr,
http://www.ceremade.dauphine.fr/∼peyre/

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 random 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 sounds and geometrical images show that adaptivity is indeed crucial to capture the structures of complex natural signals.

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, most natural signals f are characterized by very different prior assumptions, such as a decomposition with few elements on some fixed orthogonal basis $\mathcal{B}$. This is the case for a sparse expansion of a sound in a 1D local Fourier basis or the compression of a natural 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 $\Phi f = \{\langle f, \varphi_i \rangle\}_{i=1}^n$ on a set of fixed vectors $\varphi_i \in \mathbb{R}^N$. 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 [1] and Donoho [2].

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 [3] 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 sounds or natural images. For instance the orthogonal wavelet transform [4] lacks of translation and rotation invariance and is not efficient to compress geometric images [5,6]. 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 [4].

However, to capture the complex structure of sounds or the geometry 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.

Sound processing. Atoms with a broad range of localizations in space and frequency are needed to represent the transient features that exist in sounds such as the one depicted in figure 3, (a). Local cosine bases [4] divide the time axis in segments that are adapted to the local frequency content of the sound. Other kinds of dictionaries of 1D bases have been proposed, such as the wavelet packets dictionary [7].

Images and geometry. The set of cartoon images is a simple model that captures the sketch content of natural images. Figure 4, (a), shows such a geometrically regular image, which contains smooth areas surrounded by regular curves. The curvelet frame of Candès and Donoho [8] can deal with such a regularity and enjoys a better approximation rate than traditional isotropic wavelets. This result can be enhanced using a dictionary of locally elongated functions that follow the image geometry. Bandelets bases of Le Pennec and Mallat [6] 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 [9]. 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 [11] 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. This paper only deals with orthogonal best basis search but 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

Notations. The ℓ^p norms are defined by

$$\|f\|_{\ell^p}^p = \sum_i |f[i]|^p \qquad \text{and} \qquad \|f\|_{\ell^0} = \#\left\{i \setminus f[i] \neq 0\right\}.$$

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 $\mathbb{R}^N$. A cost $\mathrm{pen}(\lambda)$ is defined as a prior complexity measure associated to each basis $\mathcal{B}^\lambda$ and 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 [7], a best basis $\mathcal{B}^{\lambda^*}$ adapted to a signal minimizes a Lagrangian $\mathcal{L}$

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

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 noise level due to acquisition errors or approximate sparsity, whereas C_0 is a scaling constant that can be tuned for a specific dictionary $\mathcal{D}_\Lambda$. Note that t is estimated iteratively during the recovery algorithms presented in subsections 2.2 and 3.2.

Practical best basis computation. For a dictionary $\mathcal{D}$ that enjoys a multiscale structure, the optimization of $\mathcal{L}$ is carried out with a fast procedure. This algorithm uses the fact that if a basis splits as $\mathcal{B}^\lambda = \mathcal{B}^{\lambda_0} \cup \mathcal{B}^{\lambda_1}$, the Lagrangian satisfies $\mathcal{L}(f,\lambda,t) = \mathcal{L}(f,\lambda_0,t) + \mathcal{L}(f,\lambda_1,t)$. It is thus only needed to compute the value of $\mathcal{L}$ on elementary bases from which each basis $\mathcal{B}^\lambda$ can be decomposed. The number of such elementary bases is usually much smaller than the total number of bases. A bottom-up regression algorithm is then used to search for λ^*, see [7,6] for practical examples of this process.

2 Compressed Sensing Reconstruction

In this paper, the sampling matrix $\Phi = [\varphi_0, \dots, \varphi_{n-1}]^{\mathrm{T}}$ is defined by random points $\{\varphi_i\}_i$ of unit length although other random sensing schemes can be used, see [1].

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

$$f^* = \operatorname*{argmin}_{g \in \mathbb{R}^N} \|\Psi\, g\|_{\ell^0} \quad \text{subject to} \quad \Phi g = y. \tag{2}$$

where $\Psi = [\psi_0, \dots, \psi_{N-1}]^{\mathrm{T}}$ is the transform matrix defined by $\mathcal{B}$. The combinatorial optimization (2) is NP-hard to solve and convexification of the objective function is introduced in the following basis pursuit [12] formulation

$$f^* = \operatorname*{argmin}_{g \in \mathbb{R}^N} \|\Psi\, g\|_{\ell^1} \quad \text{subject to} \quad \Phi g = y. \tag{3}$$

Candès and Tao in [1] and Donoho in [2] 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 $\|\Psi f\|_{\ell^0} \leqslant k$ then if $n \geqslant C\, \log(N)\, k$ one recovers $f^* = f$.

To deal with noisy measurements $y = \Phi f + w$, where w is a Gaussian noise of variance t^2, one can turn the constrained formulations (3) into a penalized variational problem

$$f^* = \operatorname*{argmin}_{g \in \mathbb{R}^N} \big(\frac{1}{2}\|\Phi g - y\|_{\ell^2}^2 + t\|\Psi g\|_{\ell^1}\big). \tag{4}$$

The Lagrange multiplier t accounts both for stabilisation against noise and approximate sparsity, which is common in practical applications.

This approximate sparsity is characterized by the decays of the best M-term approximation f_M of f in basis $\mathcal{B}$ defined by

$$f_M = \sum_{m \in I_t} \langle f,\, \psi_m \rangle \psi_m \qquad \text{where} \qquad I_t = \{m \setminus |\langle f,\, \psi_m \rangle| > t\}$$

and $M = \mathrm{Card}(I_t)$. This can also be understood as an approximation using a thresholding of the coefficients of f in $\mathcal{B}$ that are below t. The approximation power of $\mathcal{B}$ is measured by the decay of the approximation error $\|f - f_M\| \leqslant C\, M^{-\alpha}$. The parameter α is a measure of the smoothness of f with respect to the basis $\mathcal{B}$.

The function f with an approximation exponent α can be seen as a k-sparse function f corrupted by a deterministic noise of amplitude $k^{-\alpha}$. One thus needs to choose a threshold $t \sim p^{-\alpha}$ for a compressed sensing scenario where $p \geqslant$

$C \log(N)\, k$. Candès et al. [1] shows that the ℓ^1 minimization (4) leads to a recovery error of order $n^{-\alpha}$ up to some logarithmic factors.

The relationship $t \sim p^{-\alpha}$ (up to logarithmic factors) is not precise enough to be useful in real applications, but reveals the insight behind the regularization formulation (4). When the number n of acquired samples decreases, the noise in the function reconstructed with basis pursuit (3) increases and the regularization imposed by the sparsity $\|\Psi f\|_{\ell^1}$ needs to be increased.

2.2 Iterative Thresholding For Sparsity Minimization

The recovery procedure suggested by equation (4) corresponds to the inversion of the operator Φ under sparsity constraints on the observed signal f. In this paper we follow Daubechies et al. that propose in [13] an algorithm based on iterated thresholding to perform such regularized inversion. This algorithm had been previously used for image restauration by Figueiredo and Nowak [14] and is derived in [15] as the iteration of two projections on convex sets.

Other greedy algorithms such as orthogonal matching pursuit (OMP) [16] and stagewise orthogonal matching pursuit [17] (StOMP) have been used to solve the compressed sensing reconstruction in a fast way. These algorithms do not fit very well into the best-basis extension exposed in section 3, mainly because OMP is too computationally intensive for imaging applications and because StOMP performed a very agressive iterations (which speed up computation but bias the choice of best-basis toward the initially chosen basis).

We make repeated use of the soft thresholding operator defined in an orthogonal basis $\mathcal{B} = \{\psi_m\}_m$

$$S_t(g) = \sum_m s_t(\langle g,\, \psi_m \rangle)\, \psi_m \qquad \text{where} \qquad s_t(x) = \text{sign}(x)(|x| - t)_+.$$

The algorithm. The steps of the algorithm are

- **Initialization.** Set $s = 0$, $f_0 = 0$.
- **Step 1: Updating the estimate.** Set $f_{s+1} = f_s + \Phi^{\mathrm{T}}(y - \Phi\, f_s)$. This update is an orthogonal projection on the convex set $\{g \setminus \Phi g = y\}$.
- **Step 2: Denoising the estimate.** Estimate the noise level σ_s using a median estimator in the transformed domain $\sigma_s = \text{median}(|\Psi f_{s+1}|)/0.6745$ and set the threshold to $t = 3\sigma_s$. Compute $f_{s+1} = S_t(f_{s+1})$ the thresholding of f in basis $\mathcal{B}$. This step corresponds to a projection on a ℓ^1 ball $\{g \setminus \|\Psi g\|_{\ell^1} \leqslant c\}$ for some c that depends on t.
- **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 strategy of updating t through iterations has been adopted by Donoho et al. in their extension of orthogonal matching pursuit [17]. The median estimator of the noise is based on the assumption, accurate in practice, that the current estimate f_{s+1} is corrupted by a Gaussian noise which diminishes during iterations.

3 Best Basis Compressed Sensing

3.1 Variational Formulation

To enhance the quality of the recovery in real life compressed sensing applications, the use of redundant frames has been proposed by several authors [18,15]. Redundancy can impact compressed sensing efficiency, because random sensing vectors can be correlated once inverted through the frame operator Ψ, since it is not orthogonal. This paper uses a different approach relying on orthogonal transforms which does not lower the recovery property of the sampling matrix Φ. It also allows a more efficient approximation through the use of a best basis in a large dictionary.

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 $\mathcal{B}^\lambda$. This is coherent with the best-basis approximation scheme of [5,6], although penalization is not strictly required in classical statistical estimation. The original recovery procedure (4) is replaced by the following minimization

$$f^* = \operatorname*{argmin}_{g\in\mathbb{R}^N} \min_{\lambda\in\Lambda} \big(\frac{1}{2}\|\Phi g - y\|_{\ell^2}^2 + t\,\|\Psi^\lambda\, g\|_{\ell^1} + C_0\, t^2\ \mathrm{pen}(\lambda)\big), \tag{5}$$

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

3.2 Best Basis Recovery Algorithm

Searching in the whole dictionary $\mathcal{D}$ for the best basis that minimizes formulation (5) 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. This leads to 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.** Set $f_{s+1} = f_s + \Phi^{\mathrm{T}}(y - \Phi\, f_s)$.
- **Step 2: Denoising the estimate.** Compute the noise level

$$\sigma_s = \mathrm{median}(|\Psi f_{s+1}|)/0.6745$$

  and set the threshold to $t = 3\sigma_s$. Compute $f_{s+1} = S_t(f_{s+1})$ where S_t is the threshold operator at t in the basis $\mathcal{B}^{\lambda_s}$.
- **Step 3: Update best basis.** Compute $\lambda_{s+1} = \mathrm{argmin}_\lambda\ \mathcal{L}(f_{s+1}, \lambda, t)$. 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.

4 Best Local Cosine Basis Compressed Sensing

4.1 Adapted Local Cosine Transform

For each scale $j > 0$, the set of locations $\{0, \dots, N-1\}$ is subdivided using $N/2^j$ intervals $[x_p^j, x_{p+1}^j]$, where the endpoints are given by $x_p^j = 2^{-j}Np - 1/2$. A local cosine basis $\mathcal{B}_j^p = \{\psi_k^{jp}\}_k$ is defined for each of these intervals using

$$\forall\, k = 0, \dots, N/2^j - 1, \quad \forall\, \ell,$$
$$\psi_k^{jp}[\ell] = b\big(2^j(\ell - x_p^j)\big)\sqrt{\frac{2}{2^{-j}N}}\cos\left[\pi\big(k+\frac{1}{2}\big)\frac{\ell - x_p^j}{2^{-j}N}\right],$$

where b is a smooth windowing function that satisfies some compatibility conditions [4]. A local cosine basis $\mathcal{B}^\lambda$ of $\mathbb{R}^N$ is parameterized by a binary tree segmentation λ of $\{0, \dots, N-1\}$, see [4]. The set of leaves $L = \{L_j^p\}$ of λ are indexed by their depth $j > 0$ and position p in the binary tree, see figure 1, left. The basis $\mathcal{B}^\lambda$ is the union of the various elementary bases $\mathcal{B}^\lambda = \bigcup \{\mathcal{B}_j^p \setminus L_j^p \in L\}$, see figure 1, right. For this local cosine basis dictionary, the penalization $\mathrm{pen}(\lambda)$ is defined as the number of leaves in the binary tree λ.

The decomposition of a signal f in each of the basis $\mathcal{B}_j^p$ is computed in $O(N\log(N)^2)$ time using FFT. From these atomic decompositions, a best basis $\mathcal{B}^{\lambda^*}$ that minimizes (1) can be extracted by a tree pruning procedure in $O(N)$ time, see [7,4].

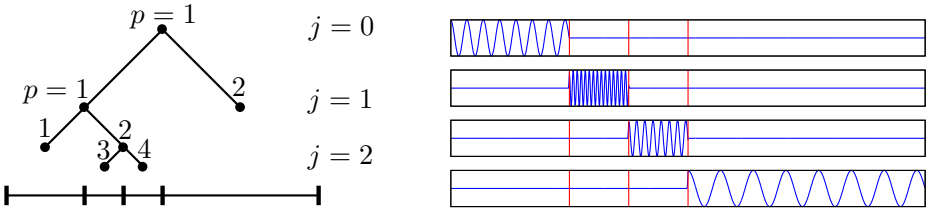

Fig. 1. *A dyadic tree λ defining a spatial segmentation (left) ; some local cosine basis functions ψ_k^{jp} of the basis $\mathcal{B}^\lambda$ (right).*

4.2 Numerical Results

A synthetic sparse signal $f = (\Psi^\lambda)^{-1}h$ is generated using a random local cosine basis $\mathcal{B}^\lambda$ and a random signal of spikes h with $\|h\|_{\ell^0} = 30$, see figure 2, (a). The signal recovered by the non-adaptative algorithm of subsection 2.2 in an uniform cosine basis $\mathcal{B}^{\lambda_0}$ is significantly different from the original, figure 2, (b). This is due to the fact that f is less sparse in $\mathcal{B}^{\lambda_0}$, since $\|\Psi^{\lambda_0} f\|_{\ell^0} = 512$

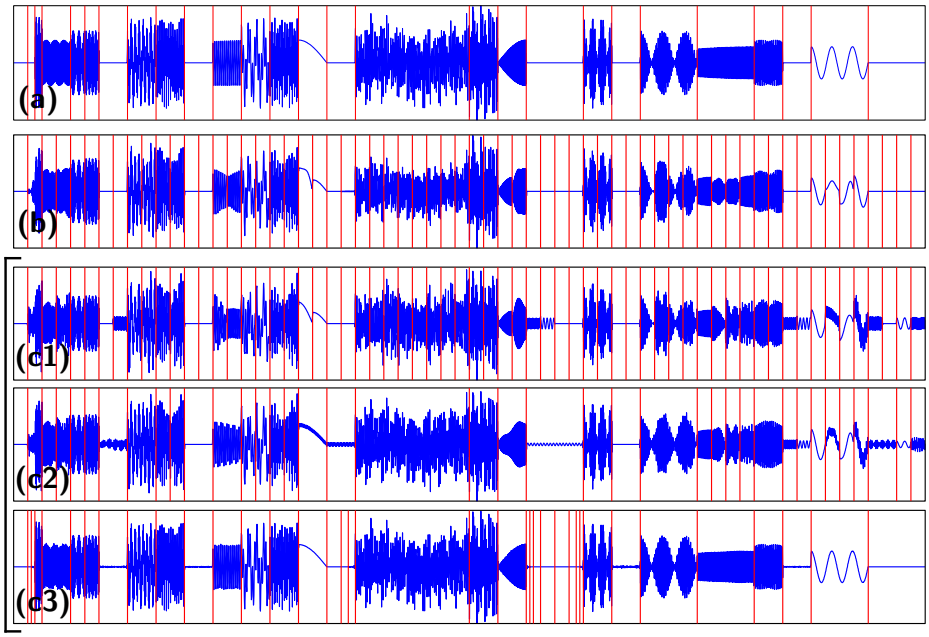

Fig. 2. *(a) synthetic sound signal with 30 random cosine atoms $N = 4096$; (b) recovery using a fixed cosine basis ; (c1) first iteration of the best basis recovery algorithm, $n = N/3$; (c2) iteration $s = 5$; (c3) iteration $s = 20$.*

and $\|\Psi^{\lambda_0} f\|_{\ell^1} \approx 2.8\|\Psi^{\lambda} f\|_{\ell^1}$. During the iterations of the algorithm presented in subsection 3.2, the estimated best basis $\mathcal{B}^{\lambda_s}$ evolves in order to match the best basis $\mathcal{B}^{\lambda}$, see figure 2, (c1–c3). The recovered signal (c3) is nearly identical to f.

On figure 3 one can see a real sound signal of a tiger howling, together with the signals recovered from fixed basis and adapted basis iterations. Although the final adapted basis is not the same as the one of the original signal, it still provides an improvement of 2dB with respect to a fixed spatial subdivision.

5 Best Bandelet Basis Compressed Sensing

5.1 Adapted Bandelet Transform

The bandelet bases dictionary was introduced by Le Pennec and Mallat [6,19] to perform adaptive approximation of images with geometric singularities, such as the cartoon image in figure 4, (a). We present a simple implementation of the bandelet transform inspired from [20].

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 4, (a). The bandelet transform corresponding to this basis applies independently over each square S of the image either

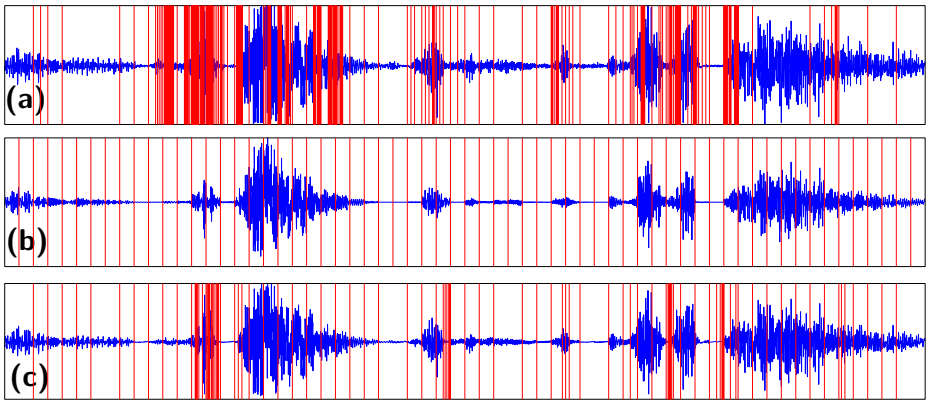

Fig. 3. *(a) sound signal of a tiger howling, together with the best spatial segmentation, $N = 32768$; (b) recovery using fixed local cosine basis, $n = N/3$ (PSNR=19.24dB) ; (c) recovery using best cosine basis, $n = N/3$ (PSNR=21.32dB)*

• 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 $p_{x^{(0)}} \leqslant p_{x^{(1)}} \leqslant \ldots \leqslant p_{x^{(m-1)}}$. This ordering allows to turn the image $\{f[x]\}_{x \in S}$ defined over S into a 1D signal $f_{1\mathrm{D}}[i] = f[x^{(i)}]$, see figure 4, (c). The bandelet transform of the image f inside S is defined as the 1D Haar transform of the signal $f_{1\mathrm{D}}$, see figure 4, (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 4 (f).

In order to restrict the number of tested geometries θ for a square $S \in Q$ containing $\#S$ pixels, we follow [20] 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 [20], 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.

5.2 Numerical Results

The geometric image depicted in figure 5, (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

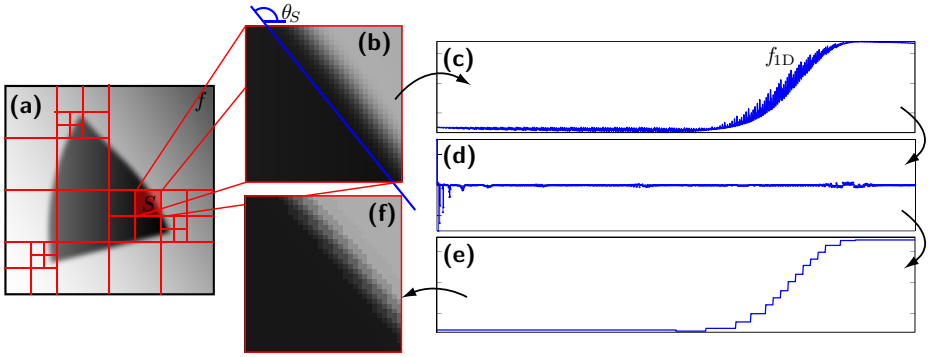

Fig. 4. *(a) a geometric image together with some adapted dyadic segmentation Q ; (b) a square S together with some adapted direction θ_S ; (c) the 1D signal f_{1D} obtained by mapping the pixels values $f(x^{(i)})$ on a 1D axis ; (d) the 1D Haar coefficients of f_{1D} ; (e) the 1D approximation obtained by reconstruction from the 20 largest Haar coefficients ; (f) the corresponding square approximated in bandelet.*

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 5, (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.

6 Conclusion

The best basis framework presented in this paper allows to recover signals with complex structures from random measurements. This approach is successful for natural sounds and geometric images that contain a broad range of sharp transitions. 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 lowers the computational burden and the numerical instabilities. 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.

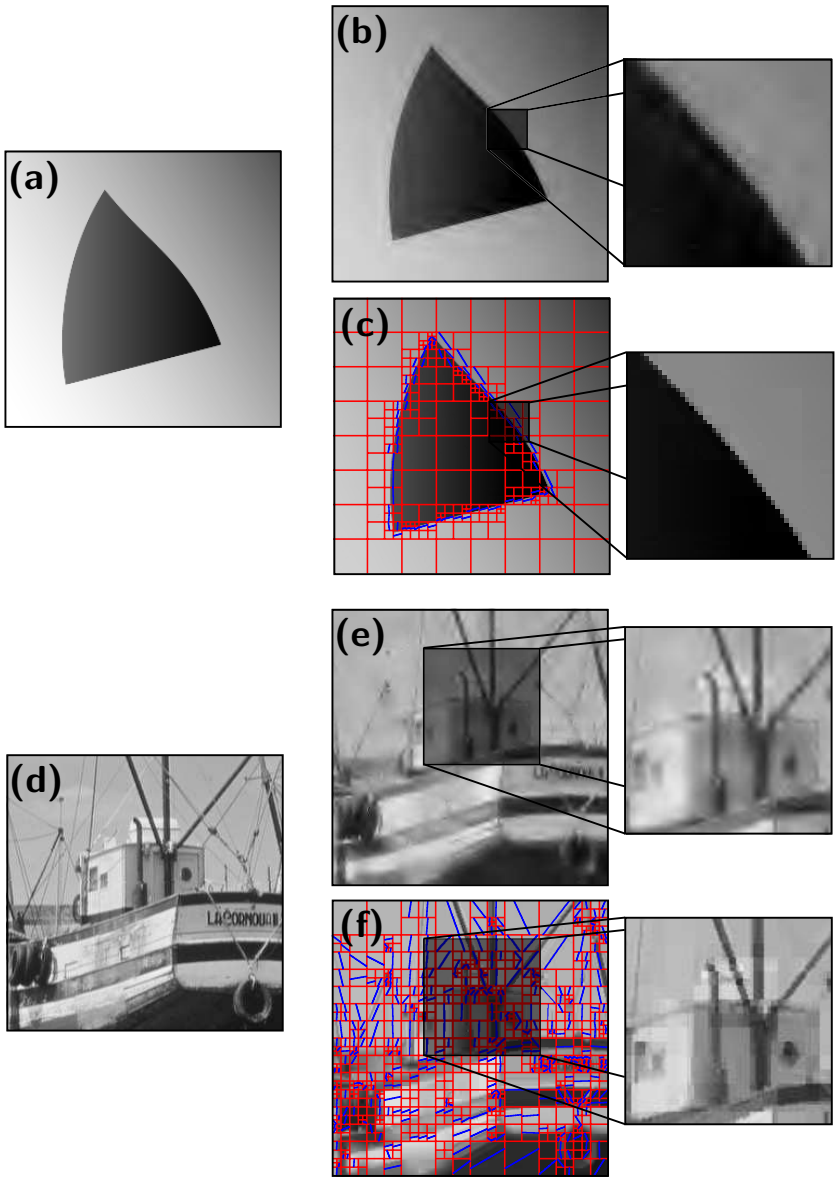

Fig. 5. *(a/d) original image ; (b/e) compressed sensing reconstruction using the wavelet basis, $n = N/6$ (b:PSNR=22.1dB, e:PSNR=23.2dB) ; (c/f) reconstruction using iteration in a best bandelet basis (c: PSNR=24.3dB, f:PSNR=25.1dB).*

References

1. Candès, E., Romberg, J., Tao, T.: Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inform. Theory (2004) Submitted.
2. Donoho, D.: Compressed sensing. IEEE Transactions on Information Theory **52**(4) (2006) 1289–1306
3. Skarda, C., Freeman, W.: Does the brain make chaos in order to make sense of the world? Behavioral and Brain Sciences **10** (1987) 161–165
4. Mallat, S.: A Wavelet Tour of Signal Processing. Academic Press, San Diego (1998)
5. Donoho, D.: Wedgelets: Nearly minimax estimation of edges. Annals of Statistics **27**(3) (1999) 859–897
6. Le Pennec, E., Mallat, S.: Bandelet Image Approximation and Compression. SIAM Multiscale Modeling and Simulation **4**(3) (2005) 992–1039
7. Coifman, R., Wickerhauser, V.: Entropy-based algorithms for best basis selection. IEEE Trans. Inform. Theory **IT–38**(2) (1992) 713–718
8. Candès, E., Donoho, D.: New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. Comm. Pure Appl. Math. **57**(2) (2004) 219–266
9. Hubel, D., Wiesel, T.: Receptive fields and functional architecture of monkey striate cortex. Journal of Physiology (London) **195** (1968) 215–243
10. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive-field properties by learning a sparse code for natural images. Nature **381**(6583) (1996) 607–609
11. Lee, T.S.: Computations in the early visual cortex. J Physiol Paris **97**(2-3) (2003) 121–139
12. Chen, S., Donoho, D., Saunders, M.: Atomic decomposition by basis pursuit. SIAM J. Sci. Comp. **20**(1) (1998) 33–61
13. Daubechies, I., Defrise, M., Mol, C.D.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Comm. Pure Appl. Math **57** (2004) 1413–1541
14. Figueiredo, M., Nowak, R.: An em algorithm for wavelet-based image restoration. IEEE Transactions on Image Processing **12**(8) (2003) 906–916
15. Candès, E., Romberg, J.: Practical signal recovery from random projections. IEEE Trans. Signal Processing (2005) Submitted.
16. Tropp, J., Gilbert, A.C.: Signal recovery from partial information via orthogonal matching pursuit. Preprint (2005)
17. Donoho, D.L., Tsaig, Y., Drori, I., Starck, J.L.: Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit. Preprint (2006)
18. Donoho, D., Tsaig, Y.: Extensions of compressed sensing. Preprint (2004)
19. Le Pennec, E., Mallat, S.: Sparse geometric image representations with bandelets. IEEE Transactions on Image Processing **14**(4) (2005) 423–438
20. Peyré, G., Mallat, S.: Surface compression with geometric bandelets. ACM Transactions on Graphics, (SIGGRAPH'05) **24**(3) (2005)