Geometric Estimation with Orthogonal Bandlet Bases

Gabriel Peyré, Erwan Le Pennec, Charles H Dossal, Stéphane Mallat

Geometrical Image Estimation with Orthogonal Bandlet Bases

Gabriel Peyré[a], Erwan Le Pennec[b], Charles Dossal[c], Stéphane Mallat[d]

[a]CEREMADE, Université Paris Dauphine;
Place du Maréchal De Lattre De Tassigny
75775 Paris Cedex 16 (FRANCE)

[b]LPMA, Université Paris Diderot - Paris 7,
2, place Jussieu,
75251 PARIS Cedex 05 (FRANCE)

[c]LaBAG, Université Bordeaux 1,
351, cours de la Libération,
F-33405 Talence cedex (FRANCE)

[d]CMAP, École Polytechnique
Route de Saclay,
91128 Palaiseau Cedex (FRANCE)

ABSTRACT

This article presents the first adaptive quasi minimax estimator for geometrically regular images in the white noise model. This estimator is computed using a thresholding in an adapted orthogonal bandlet basis optimized for the noisy observed image. In order to analyze the quadratic risk of this best basis denoising, the thresholding in an orthogonal bandlets basis is recasted as a model selection process. The resulting estimator is computed with a fast algorithm whose theoretical performance can be derived. This efficiency is confirmed through numerical experiments on natural images.

Keywords: Image denoising, bandlets, minimax estimation, model selection

Introduction

Image estimation within the framework of an additive white noise is a classical problem in image processing and statistics. The goal is to recover an image f of N pixels from a noisy observation

$$Y = f + \varepsilon W$$

where W is a gaussian white noise with unit variance and ε is a known parameter that controls the variance of the noise. Denoising the image f from the observed data Y implies the computation of an estimator F that depends only on the observed data Y. The performance of the estimation is measured here using the average quadratic risk $E(\|f - F\|^2)$ where the expectation is computed over the random noise W.

Geometric images can be modeled as C^α images outside a set of edge curves that are themselves C^α. For such geometric images, Korostelev and Tsybakov[1] proved that for any estimator F, the quadratic risk $E(\|f - F\|^2)$ cannot decay faster with the noise level ε than $\varepsilon^{2\alpha/(\alpha+1)}$. This article proposes an estimator that reaches this asymptotic decay for geometric images up to a logarithmic factor. This estimator is adaptive in the sense that it does not require an a priori knowledge of the smoothness α.

This estimator is a thresholding operator in an orthogonal basis adapted to the noisy data. Namely, bandlet dictionaries of orthogonal bases offer an asymptotically optimal representation for geometric images. Results from model selection theory allows this optimality to carry over to the setting of estimation in gaussian noise. This estimator is computed with a fast analysis and synthesis algorithm and its performance are estimated for the denoising of natural images. Theoretical details of this construction are exposed in.[2]

Contact: gabriel.peyre@ceremade.dauphine.fr

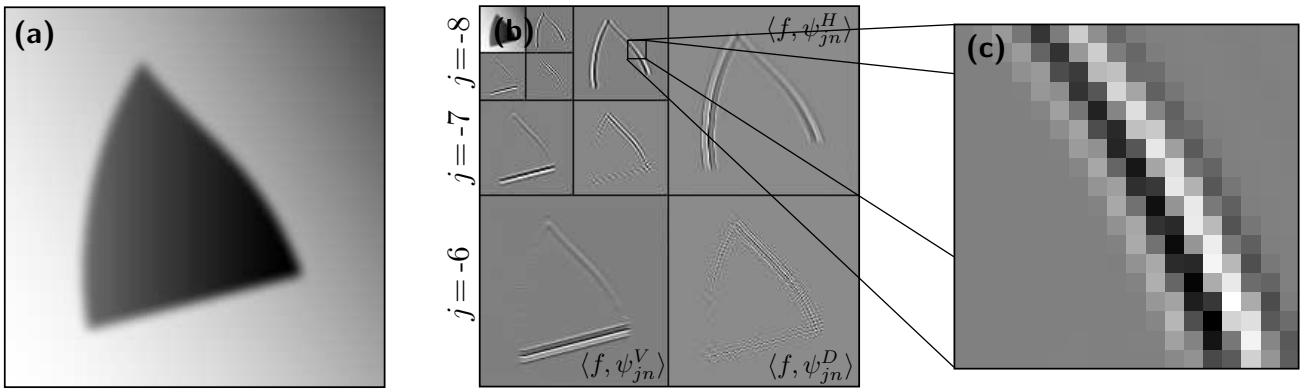


Figure 1. (a) A geometrically regular function. (b,c) Its wavelet coefficients, where the white color corresponds to large positive coefficients and the black corresponds to large negative coefficients. These coefficients are arranged by scale 2^j and orientation k.

1. ESTIMATION AND APPROXIMATION OF GEOMETRIC IMAGES

1.1 White Noise Model

A simple non-linear estimator can be constructed with a thresholding operator in an orthogonal basis $\mathcal{B} = \{b_\nu\}_\nu$ of $\mathbb{R}^N$, where N is the number of pixels in the image. This thresholding is defined as $F = S_T(Y, \mathcal{B})$ where the operator S_T is

$$S_T(g, \mathcal{B}) = \sum_{|\langle g,\, b_\nu \rangle| \geqslant T} \langle g,\, b_\nu \rangle b_\nu. \tag{1}$$

Donoho and Johnstone[3] show that the threshold level $T = \gamma\sqrt{\log_2(N)}\varepsilon$ is optimal when ε goes to zero as soon as γ is large enough: the corresponding quadratic risk is smaller up to a logarithmic factor than the one obtained by the projection on the best possible subspace.

The quality of the approximation of an image f in $\mathcal{B}$ is classically assessed using the M-term approximation error $\|f - f_M\|$ where

$$f_M = S_T(f) \qquad \text{and} \qquad M = \#\left\{\nu \setminus |\langle f,\, b_\nu \rangle| \geqslant T\right\}.$$

Donoho and Johnstone[3] prove the following link between the estimation risk and the M-term approximation

$$\|f - f_M\|^2 \leqslant C_1 M^{-\alpha'} \quad \Longrightarrow \quad E(\|f - F\|^2) \leqslant C_2 \log_2(N) \varepsilon^{\frac{2\alpha'}{\alpha'+1}}, \tag{2}$$

where C_1 and C_2 are constants that depend only on f. This relationship shows that the construction of a basis $\mathcal{B}$ that can approximate well the set of functions f one wishes to denoise induces an efficient thresholding estimator.

This analysis is however restricted to the use of a single orthogonal basis $\mathcal{B}$. In order to approximate and denoise complex geometrical patterns, one would like to use more advanced constructions. This article shows how to use a dictionary of orthogonal bases in order to estimate optimally geometrical images.

1.2 Geometrically Regular Images

Functions of $L^2([0,1]^2)$ that are C^α outside a set of curves that are themselves C^α constitute a simple model for geometric images. This model has been introduced by Korostelev and Tsybakov in order to study the estimation of geometric edges in images.[1] To account for the blurring inherent to the diffraction of the optical device, we consider the set of images $f = \tilde{f} * h$ where $\tilde{f}$ is C^α outside of C^α edges and h is an unknown smooth and compactly supported kernel. Figure 1 shows an example of such a $\mathrm{C}^\alpha - \mathrm{C}^\alpha$ geometrically regular function.

Korostelev and Tsybakov[1] show that the quadratic risk $E(\|F - f\|^2)$ cannot decay faster than

$$O(\varepsilon^{2\alpha/(\alpha+1)}) \quad \text{for all } \mathrm{C}^\alpha - \mathrm{C}^\alpha \text{ geometrically regular functions,} \tag{3}$$

when ε goes to zero. Using this upper bound and equation (2), they show that the optimal asymptotic approximation speed for geometric images is $\|f - f_M\|^2 \leqslant O(M^{-\alpha})$.

The result (3) is based on an estimation of the Kolmogorov complexity of the set of geometric functions. They propose an estimator that achieves this risk for non blurred geometric images. It is based on a detection of the edges and thus is not adapted to blurred geometric images or natural images. This article derives an alternate construction based on a decomposition in an adapted orthogonal basis. This construction comes with a fast algorithm suitable for image denoising and with a provable estimation bound that matches the one of Korostelev and Tsybakov.

1.3 Geometric Estimation and Geometric Approximation

Wavelet approximation. The thresholding estimator of Donoho and Johnstone[3] performs well on natural images and its translation invariant extension is often preferred in practice.[4] This isotropic regularization over wavelets the wavelet domain is close to the total variation minimization of Rudin, Osher and Fatemi,[5] see for instance[6] for a variational formulation of the wavelet denoising.

Approximation in an orthogonal wavelet basis reaches the bound (3) only on the restricted class of uniformly C^α functions, see.[7] The asymptotic decay of the wavelet approximation for functions having discontinuities along curves is $\|f - f_M\|^2 = O(M^{-1})$ which is also the decay for the larger class of bounded variations functions. Wavelets are thus not optimal to approximate geometrically regular images because their isotropic support fails to capture the regularity of the edge curves.

Geometric Estimation. The wavelet decomposition of geometric patterns exhibits dependencies among neighboring coefficients over both the spacial and the scale domain. Figure 1 (c) indeed shows how the wavelet coefficients are smoothly varying along an edge. Simoncelli and co-workers[8] have studied the non-gaussian and non-independent behavior of groups of wavelet coefficients.

In order to enhance the results of denoising by a thresholding over the wavelet domain, one has to take into account local dependencies among the coefficients. In this article, we compare our results to the BLS-GSM algorithm of Portilla et al.[9] but other algorithms exploit the inter and intra-scale redundancies of wavelet coefficients, see for instance.[10–12]

This line of research makes use of a clever statistical modeling to enhance the estimation process and has proved effective for natural images denoising. The performance of these models is however difficult to analyze and there is no underlying functional model for the image to recover. This paper proposes to capture the redundancy of the wavelet representation using a retransformation of these multiscale coefficients. This bandletization process avoids the need of a complex statistical model and an independent thresholding of the resulting coefficients leads to an optimal estimation for geometric images.

Other approaches take into account the geometry of the underlying image. This includes anisotropic diffusion,[13] non-local filtering[14] optimized local transforms learned from data[15] and more complex filtering processes based on local transforms.[16]

Geometric approximation. For C^α geometrically regular functions, one wants to find an approximation scheme with M parameters which yields an error that decays like $M^{-\alpha}$, as in a wavelet approximation of uniformly C^α functions. Indeed, although these functions may be discontinuous, one can take advantage of the regularity of their edge curves.

To exploit the geometric image regularity along edge curves, the image has to be decomposed over functions having both an elongated support and vanishing moments. The curvelets of Candès and Donoho[17] are such elongated functions and define a frame. Candès and Donoho prove that for $\alpha = 2$ the best M-term approximation f_M of a C^α-geometrically regular function f in curvelets satisfies

$$\|f - f_M\|^2 = O(\log^3(M)\, M^{-2}). \tag{4}$$

Curvelet approximations are nearly optimal for $\alpha = 2$, but one does not reach the $M^{-\alpha}$ optimal bound for $\alpha > 2$. A thresholding estimator in a curvelet frame has been used successfully to denoise natural images.[18]

The warped bandlet transform, introduced by Le Pennec and Mallat[19] uses an adaptive scheme to approximate an image. A dictionnary of bandlet bases parametrized by their geometry is defined and a fast algorithm allows to find a best

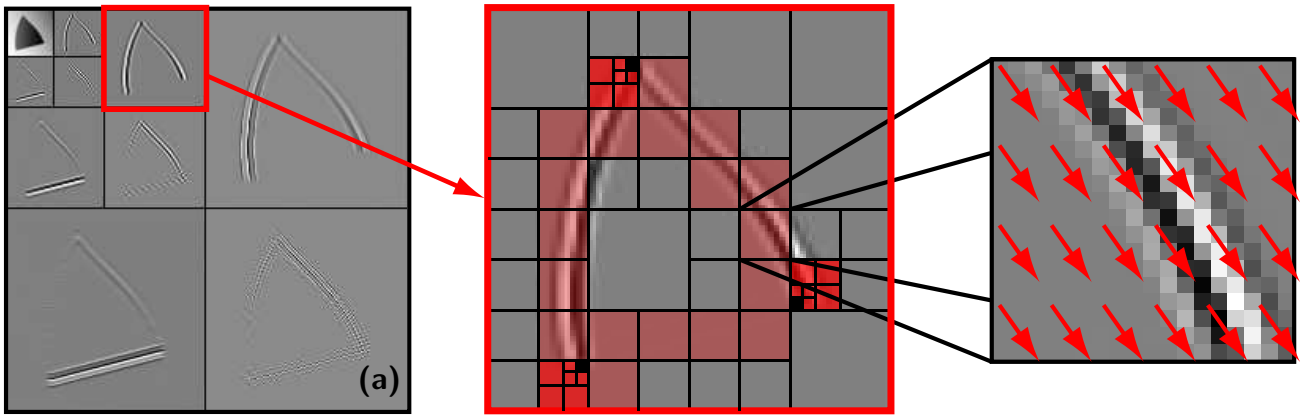


Figure 2. (a) Wavelet coefficients of an image. (b) Dyadic segmentation of the wavelet coefficients of a geometrically regular image. (c) An adapted flow is computed on each square of the segmentation.

basis adapted to the function to approximate. The resulting approximation scheme results in an error $\|f - f_M\|^2$ that decays like $O(M^{-\alpha})$ where M is the total number of parameters to describe both the approximation and the best bandlet basis, see.[19]

This warped bandlet construction however does not correspond to an orthogonal decomposition, which leads to both theoretical and algorithmic difficulties when approximating and estimating a geometric image. The energy conservation of an orthogonal transformation is indeed a key ingredient to analyze the thresholding estimation. This orthogonality is important to extend the bound (2) to a dictionary of bases using model selection theory. This paper uses the construction of Mallat and Peyré[20] that builds orthogonal bandlets by retransforming the coefficients of a wavelet decomposition.

2. ORTHOGONAL BANDLET BASES

Mallat and Peyré construct a dictionary of orthogonal bases[20] that offers the same $O(M^{-\alpha})$ error decay as the original warped bandlets[19] for the class of C^α geometric images. These orthogonal bases are defined on top of an orthogonal wavelet transform. The wavelet coefficients are retransformed using an adaptive bandletization that enhances the original wavelet representation. This bandletization step is defined by an Alpert multiwavelet transform along an optimized flow that follows closely the geometry of the underlying image.[21]

2.1 Orthogonal Bandlets Parametrization

Each orthogonal bandlet basis $\mathcal{B}(\lambda) = \{b_\nu\}_\nu$ is parametrized using a geometry $\lambda \in \Lambda$ that specifies, for each scale 2^j and each orientation k of the wavelet transform,

- a dyadic segmentation of the corresponding wavelet coefficients,
- a flow, that indicates the approximate geometric direction over each square of the segmentation that contains an edge.

The bandlets are obtained through an orthogonal retransformation of the wavelet coefficients inside each square that contains an edge. This retransformation is the decomposition of each set of wavelet coefficients on an orthogonal basis of Alpert multi-wavelets.[21]

This Alpert retransformation of the wavelet coefficients corresponds to the decomposition of the original image on a set of elongated basis functions b_ν. These bandlet functions capture the directional regularity of contours because the Alpert retransformation follows the estimated geometric flow. Figure 2 shows an example of a typical geometric parameter $\lambda \in \Lambda$ for a bandlet basis adapted to a synthetic geometric image. On can see on (a) that the wavelet coefficients are close to zero in regular areas, but large coefficients emerge near edges. On (b) the dyadic segmentation isolates squares where a single geometric direction can be robustly represented using a vector field, as shown on (c).

Bandlet bases are gathered in a dictionary of orthogonal bandlet bases indexed by a geometry. The efficiency of these bases is linked to the use of two fast algorithms. The first one performs the analysis (decomposition) and synthesis (reconstruction) of a function f in some given basis $\mathcal{B}(\lambda)$. The second algorithm searches in the whole dictionary of $\mathcal{B}(\lambda)$ for a best basis $\mathcal{B}(\lambda^\star)$ adapted to some function f one wishes to approximate.

2.2 Bandlet Approximation Algorithm

Given a geometry $\lambda \in \Lambda$, the fast bandlet transform algorithm decomposes an image f in the bandlet basis $\mathcal{B}(\lambda) = \{b_\nu\}_\nu$. This algorithm, detailed in,[20] computes first an orthogonal wavelet transform followed by a directional Alpert transform in each square of the segmentation described by λ.

The best M-terms approximation f_M of f in this basis is obtained by keeping the M largest coefficients. This is equivalent to defining f_M with the thresholding operator S_T using an adapted threshold T

$$f_M = S_T(f, \mathcal{B}(\lambda)) = \sum_{|\langle f,\, b_\nu\rangle| \geqslant T} \langle f,\, b_\nu\rangle b_\nu \tag{5}$$

where $M = \#\{\nu \setminus |\langle f,\, b_\nu\rangle| \geqslant T\}$ and where S_T is defined in (1). The threshold T is set to the value of the M^{th} coefficient ordered by magnitudes. The image f_M is obtained from the set of thresholded coefficients with a fast inverse bandlet transform. This algorithm is similar to the forward transform and applies the inverse Alpert directional transform followed by the inverse orthogonal wavelet transform.

2.3 Best Basis Search Algorithm

Given a fixed number of parameters M, one can search for the best bandlet basis that minimizes the approximation error $\|f - f_M\|$. This constrained problem is difficult to solve, but one can relax it and minimize the corresponding Lagrangian that penalizes the approximation error with the number of terms M as followed

$$\mathcal{L}(f, T, \mathcal{B}(\lambda)) = \|f - f_M\|^2 + T^2 M, \qquad \text{where} \qquad f_M = S_T(f, \mathcal{B}(\lambda)) \tag{6}$$

where S_T is defined in (1).

The best basis algorithm allows to optimize the Lagrangian $\mathcal{L}$ over the set of bases $\mathcal{B}(\lambda)$ for $\lambda \in \Lambda$. Thanks to the dyadic refinement of the segmentation composing the geometries, this algorithm avoids to test every possible geometry and has an overall complexity of $O(NT^{-2(p-1)^2})$ where $p \geqslant \alpha$ is the number of vanishing moments of the transform.

For a fixed threshold T, this algorithm computes the adapted geometry

$$\lambda^\star = \operatorname*{argmin}_{\lambda \in \Lambda} \mathcal{L}(f, T, \mathcal{B}(\lambda)).$$

The approximation $f_M = S_T(f, \mathcal{B}(\lambda^\star))$ in the best basis $\mathcal{B}(\lambda^\star)$ is optimal for the constrained approximation problem of f with M coefficients. If f is a $\mathrm{C}^\alpha - \mathrm{C}^\alpha$ geometrically regular function, Peyré and Mallat[20] show that

$$\|f - f_M\|^2 = O(M^{-\alpha}). \tag{7}$$

This result is adaptive since one does not need to know a priori the regularity parameter α in order to reach the optimal decay of the error.

3. GEOMETRIC ESTIMATION WITH BANDLETS

Equation (2) shows that the performance of a thresholding estimator F of f from a noisy observation $Y = f + \varepsilon W$ is linked to the approximation properties of the chosen basis. A thresholding of Y in the best bandlet basis associated to f with the threshold $T = \gamma\sqrt{\log_2(N)}\varepsilon$ should thus give an optimal estimation for $\mathrm{C}^\alpha - \mathrm{C}^\alpha$ geometrically regular functions. Such a basis is however not available in practice since it depends on the unknown noiseless image f. One thus has to estimate an adapted basis from the noisy observation Y. This section uses tools from model selection to show that this strategy indeed leads to a nearly optimal estimator.

3.1 Bandlet Estimator

The bandlet estimator F is defined in two steps. One first searches for the best basis $\mathcal{B}(\lambda^\star)$ associated to the noisy observation Y and to the threshold T,

$$\mathcal{B}(\lambda^\star) = \operatorname*{argmin}_{\mathcal{B}(\lambda)} \mathcal{L}(Y, T, \mathcal{B}(\lambda)),$$

where the Lagrangian is defined in equation (6). This estimator is then computed as

$$F = S_T(Y, \mathcal{B}(\lambda^\star)) \qquad \text{where} \qquad T = \gamma\sqrt{\log_2(N)}\varepsilon \tag{8}$$

where γ is a large enough constant, and the thresholding operator is defined in equation (1). The determination of $\mathcal{B}(\lambda^\star)$ is performed using the fast best basis search algorithm outlined in section 2. The decomposition of Y in $\mathcal{B}(\lambda^\star)$ and the reconstruction after the thresholding also use the fast bandlet processing algorithms described in section 2.2.

3.2 Model Selection and Bandlets

In order to analyze the statistical properties of this estimator, one cannot use the classical theory of thresholding in an orthogonal basis. Our estimator can however be re-casted within the framework of the model selection.[22]

This theory proposes a non-linear statistical estimator that is the projection of the observations Y on a vector space $m^\star$, solution of the following penalized optimization problem

$$m^\star = \operatorname*{argmin}_{m\in\mathcal{M}} \|Y - P_m Y\|^2 + T^2 \dim(m),$$

where P_m is the orthogonal projector on m. In this problem, $\mathcal{M}$ is the set of available models that are low dimensional vector spaces that should approximate well the original data f. The estimator F of f using the optimal model is defined as

$$F \stackrel{\text{def.}}{=} P_{m^\star} Y.$$

The model selection theory allows to control the quadratic risk $E(\|f - F\|^2)$ as long as the number of vectors used to generate all models m in $\mathcal{M}$ grows polynomially with the dimension N. In such a case, one has the following result[22]

$$E(\|f - F\|^2) \leqslant C\Big(\min_{m\in\mathcal{M}} \|f - P_m f\|^2 + \gamma^2 \log_2(N)\varepsilon^2 \dim(m) + N^{-1}\Big). \tag{9}$$

When the set $\mathcal{M}$ of all the models is the set of sub-spaces spanned by the vectors of some orthogonal basis $\mathcal{B} = \{b_\nu\}$, the best model corresponds to the thresholding

$$m^\star = \operatorname{span}\left\{ b_\nu \setminus |\langle Y,\, b_\nu\rangle| \geqslant T \right\}.$$

In the general case, the computation of $m^\star$ is however a difficult task. This article focusses on a highly structured set of models that are built from the bandlet dictionary. This structure allows to apply a fast best basis exploration algorithm to select the optimal model.

For the bandlet estimator, $\mathcal{M}$ is the union of all the subspaces spanned by bandlets in the different bases

$$\mathcal{M} \stackrel{\text{def.}}{=} \bigcup_{\lambda\in\Lambda} \bigcup_{k>0} \bigcup_{(\nu_i)_i \subset \mathcal{B}(\lambda)} \operatorname{span}\left(b_{\nu_1}, \ldots, b_{\nu_k}\right).$$

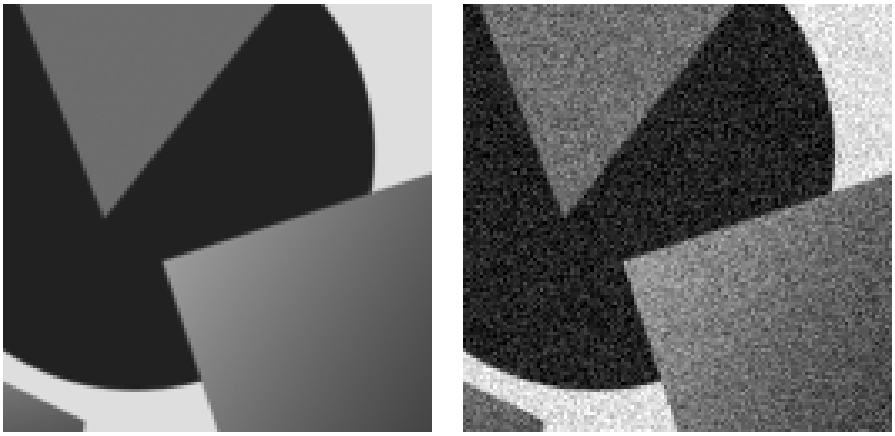
One thus has

$$P_{m^\star} Y = S_T(Y, \mathcal{B}(\lambda^\star)) = F.$$

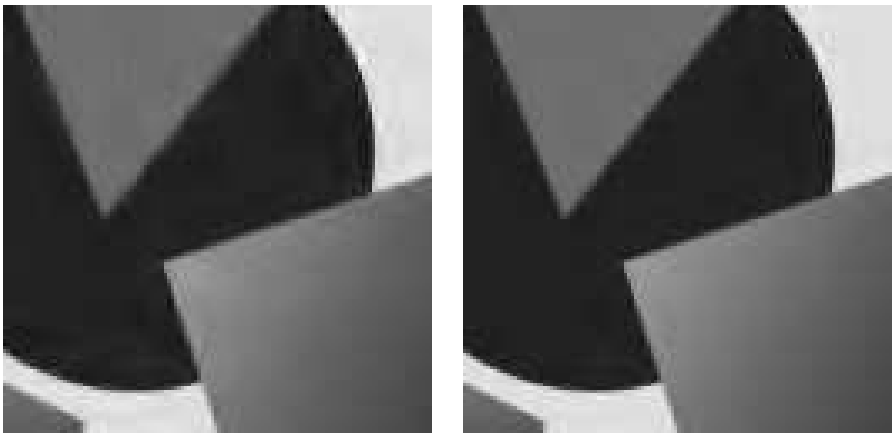
where the best basis $\mathcal{B}(\lambda^\star)$ associated to Y is defined in equation (3.1). In the case of the bandlet model selection problem, the optimal model $m^\star$ can be computed using the fast best basis search algorithm exposed in section 2.3.

The number of bandlet functions that generate $\mathcal{M}$ is polynomial in the number of pixels N of the image f one wishes to recover. One can thus apply the model selection result of equation (9), that can be reformulated using the Lagrangian of equation (6)

$$E(\|f - F\|^2) \leqslant C\Big(\min_{\mathcal{B}(\lambda)} \mathcal{L}(f, \gamma\sqrt{\log_2(N)}\varepsilon, \mathcal{B}(\lambda)) + N^{-1}\Big). \tag{10}$$



Original image | Noisy image (22.06dB)



Wavelets (36.03dB) | Bandlets (38.27dB)

Figure 3. Visual comparison of the wavelets and the bandlet estimators on a geometric image.

3.3 Bandlet Estimation Theorem

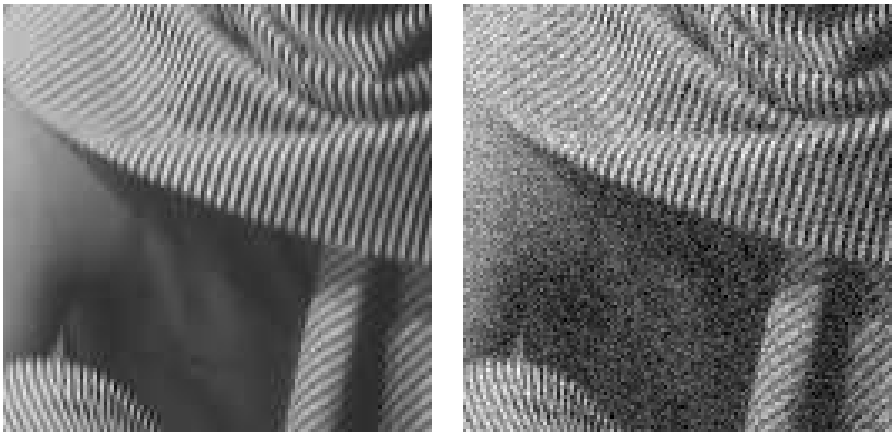
In order to assess the quality of the bandlet estimator, one needs to combine the model selection result (10) with the approximation result (7). This leads to the following theorem of quasi-optimality of the bandlet estimator, see[2] for a detailed proof of this result.

THEOREM 1. *For any $C^\alpha - C^\alpha$ geometrically regular function f, it exists a constant C such that, for any noise level ε, the bandlet estimator F satisfies*

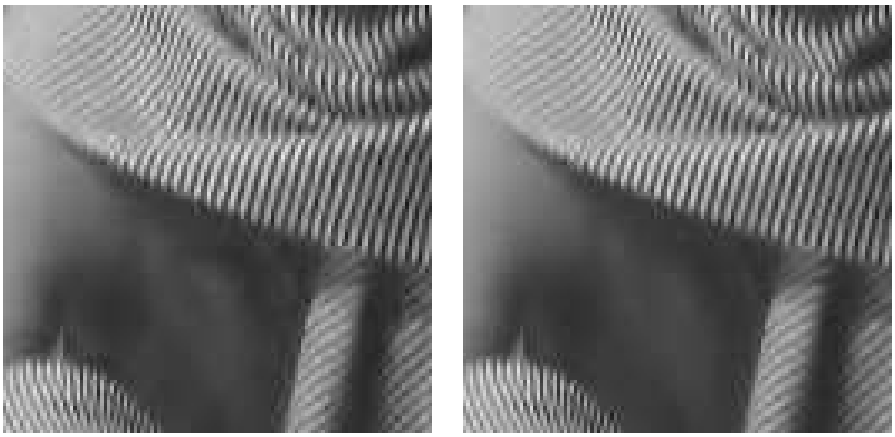
$$E(\|f - F\|^2) \leqslant C\Big(\log_2(N)\varepsilon^{\frac{2\alpha}{\alpha+1}} + N^{-1}\Big).$$

This estimator exploits the adaptivity of the bandlets since it does not require the knowledge of the regularity parameter α. This result can be enhanced for ε^2 that is large with respect to N^{-1} by replacing the $\log_2(N)$ term by $\log_2(\varepsilon)$.

In order to reach this optimal risk decay, one has to exploit the anisotropic regularity of the geometry. The use of an orthogonal decomposition is however not mandatory and other constructions are possible. The non adaptive estimator proposed by Korostelev and Tsybakov[1] is for instance based on a direct estimation of the edges. The use of orthogonal bases eases the mathematical analysis of the scheme and underlines the link between estimation and approximation. Using the framework of model selection, this link can be used in a seamless manner as long as one can control both the approximation rate of the set of bases and the overall cardinality of the set of vectors to avoid over-fitting the noise.



Original image | Noisy image (22.06dB)



BLS-GSM (29.41dB) | Bandlets (29.32dB)

Figure 4. Visual comparison of the wavelet and bandlet estimators on a natural image.

4. NUMERICAL RESULTS

Theorem 1 gives the asymptotic optimality of the bandlet estimator when the noise level ε vanishes. In real applications, this noise level can be far from zero and the asymptotic theorem can be inefficient to describe the quality of an estimator. Furthermore, natural images can deviate a lot from the ideal model of $\mathrm{C}^\alpha - \mathrm{C}^\alpha$ geometrically regular images because of the presence of textured areas that are not pure edges.

A translation invariant estimator. In order to enhance the results of the bandlet estimator and make it approximately translation invariant, we apply the denoising process to 4 translated versions of the noisy observation and average the results. This process is similar to the cycle-spinning algorithm used in translation invariant wavelet denoising.[4]

In order to further increase the translation invariance of the estimator, we have enhanced the Alpert transform[21] that is at the core of the bandletization process. This Alpert decomposition can be turned into a translation invariant redundant frame using cycle-spinning or using a fast decomposition similar to the translation invariant Haar transform. This redundant bandletization is implemented in a toolbox freely available online.[23]

Denoising results. Figure 5 quantifies the numerical results of the denoising process using the PSNR, defined as

$$\mathrm{PSNR}(f, g) = -20 \log_{10}(\|f - g\|_2 / \|f\|_\infty),$$

for $g = Y$ (noisy PSNR) and $g = F$ (denoised PSNR).

Three different estimators have been used:

– a thresholding operator in a translation invariant wavelet frame,[4]
– the BLS-GSM estimator,[9] that uses an advanced statistical modeling for the wavelet coefficients,
– the translation invariant bandlet estimator.

For geometric images (figure 3 and figure 5, right) the bandlet estimator gives better results than the other estimators. In particular, the recovered edges are sharper and show less ringing artifacts. This is due to the anisotropic support of the adapted bandlet functions that follow closely the geometry of the underlying image.

For a complex natural image (figures 4 and figure 5, left) the bandlet estimator gives results similar to the state of the art (BLS-GSM method[9]). The bandlet reconstruction recovers well sharp transitions and directional textures, but is less efficient for complex texture patterns and regular areas, where a statistical modeling is more adapted.

REFERENCES

1. A. Korostelev and A. Tsybakov, *Minimax Theory of Image Reconstruction*, vol. 82, Spinger, 1993.
2. C. Dossal, E. L. Pennec, and S. Mallat, "Denoising with bandelets," *Preprint* , 2007.
3. D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika* **81**, pp. 425–455, Dec 1994.
4. R. R. Coifman and D. Donoho, "Time-invariant wavelet de-noising," in *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, eds., *Lecture Notes in Statistics* **103**, pp. 125–150, Springer-Verlag, (New York), 1995.
5. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**, pp. 259–268, 1992.
6. A. Chambolle, R. D. Vore, N. Lee, and B. Lucier, "Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Transcations on Image Processing* **7**, pp. 319–335, March 1998.
7. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, 1998.
8. E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *2001* **24**, pp. 1193–1215, Annual Review of Neuroscience.
9. J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, "Image denoising using a scale mixture of Gaussians in the wavelet domain," *IEEE Trans Image Processing* **12**, pp. 1338–1351, November 2003.
10. A. Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single and multiband image denoising," *IEEE Transactions on Image Processing* **15**, pp. 654–665, March 2006.

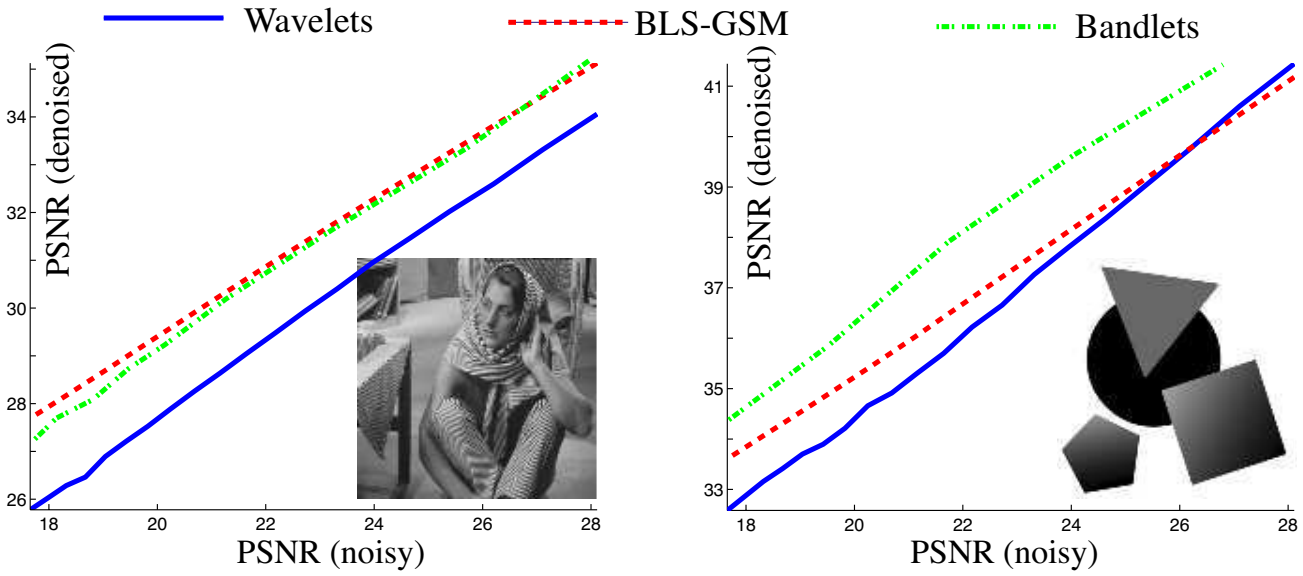


Figure 5. Comparison of the three methods for various noise levels.

11. L. Sendur and I. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Transactions on Signal Processing* **50**, pp. 2744–2756, November 2002.
12. F. Luisier, T. Blu, and M. Unser, "A new sure approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Transactions on Image Processing* **16**, pp. 593–606, March 2007.
13. D. Tschumperlé and R. Deriche, "Vector-valued image regularization with PDEs: A common framework for different applications," *IEEE Trans. Pattern Anal. Mach. Intell* **27**(4), pp. 506–517, 2005.
14. A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation* **4**(2), pp. 490–530, 2005.
15. M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Processing* **15**, pp. 3736–3745, Dec. 2006.
16. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3d transform-domain collaborative filtering," *IEEE Transactions on Image Processing* **16**, August 2007.
17. E. Candes and D. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise c_2 singularities," *Comm. Pure Appl. Math.* **57**(2), pp. 219–266, 2004.
18. J. L. Starck, E. J. Candes, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Processing* **11**, pp. 670–684, June 2002.
19. E. Le Pennec and S. Mallat, "Bandelet Image Approximation and Compression," *SIAM Multiscale Modeling and Simulation* **4**(3), pp. 992–1039, 2005.
20. S. Mallat and G. Peyré, "Orthogonal bandelet bases for geometric image approximation," in *Preprint CEREMADE*, 2006.
21. B. Alpert, *Wavelets and Other Bases for Fast Numerical Linear Algebra*, pp. 181–216. C. K. Chui, editor, Academic Press, San Diego, CA, USA, 1992.
22. A. Barron, L. Birge, and P. Massart, "Risk bounds for model selection via penalization," *Probab. Th. Rel. Fields* **113**, pp. 301–413, 1999.
23. G. Peyré and S. Mallat, "Bandlets toolbox, available on Matlab Central." `http://www.mathworks.com/matlabcentral/`, 2005.