Geometrical Image Compression with Bandelets

Erwan Le Pennec[a] and Stéphane Mallat[ab]

[a] Ecole Polytechnique, CMAP, 91128 Palaiseau Cedex, France
[b] Courant Institute, NYU, New-York, USA

ABSTRACT

This paper introduces a new class of bases, called bandelet bases, which decompose the image along multiscale vectors that are elongated in the direction of a geometric flow. This geometric flow indicates the direction in which the image grey levels have regular variations. The image decomposition in a bandelet basis is implemented with a fast subband filtering algorithm. Bandelet bases lead to optimal approximation rates for geometrically regular images. For image compression, the bandelet basis geometry is optimized with a fast best basis algorithm. Comparisons are made for image compression with wavelet bases.

1. INTRODUCTION

Developing efficient signal processing algorithms for signal compression, noise removal or inverse problems, often requires to build sparse representations where the signal is precisely approximated with few parameters. Such representations take advantage of some form of regularity, which indicate that some values can be predicted from their neighbors. Representations in orthonormal bases have been shown to be particularly efficient for images and in particular block cosine bases and wavelet bases have lead to the two image compression standard JPEG and JPEG-2000. They are constructed with separable products of one dimensional bases and are composed of vectors having a square support. Such bases do not take into account an important source of redundancy in images, which is their geometric regularity. Sharp image transitions such as edges are expensive to represent although one could reduce their cost by taking into account the fact that they often have a piecewise regular evolution across the image support. Integrating the geometric regularity in the image representation is therefore a key challenge to improve state of the art applications to image compression, denoising or inverse problems. Very different approaches have been studied, including edge detection techniques[1–3] or the construction of bases such as curvelets[4] or contourlets.[5] By reviewing previous approaches, Section 3 explains the difficulties to create stable and efficient geometric representations.

This paper constructs a new representation, that decomposes the image over a basis of *bandelets*, which are elongated multiscale vectors that are adapted to the image geometry. A major difficulty is to optimize the calculated geometry. In computer vision, it is well known that the geometry provides crucial information to analyze the information content of images, but hundreds of papers have already been published on edge detection, with no final consensus on how to do it. We argue that edges are ill-defined and often not appropriate to describe the image geometry. Moreover, for complex images, an optimal geometry can only be defined in the context of a precise application such as image compression or noise removal, where the performance of the geometry can be measured.

To construct bandelet bases, the image geometry is defined as a field of locally parallel vectors which indicates directions in which the image gray level have regular variations. We call *geometric flow* this vector field. One can associate a bandelet basis to any geometric flow. The geometric flow is not calculated a priori from the image, but a posteriori by optimizing the bandelet basis for a particular application. Finding the optimal geometry is therefore casted as a best basis search among a predefined dictionary of bases, in order to optimize a particular application. This paper studies applications to image compression.

Bandelet bases are obtained with a *bandeletization* of warped wavelet bases, which takes advantage of the image regularity along the geometrical flow. Section 4 explains how to construct such bases together with their

Further informations:
Corresponding Author : E. L.P. (lepennec@cmap.polytechnique.fr)
This work was partly supported by the NSF grant IIS-0114391.

Visual Communications and Image Processing 2003, Touradj Ebrahimi, Thomas Sikora,
Editors, Proceedings of SPIE Vol. 5150 (2003) © 2003 SPIE · 0277-786X/03/$15.00

geometric flow, and Section 5 shows how to discretize this basis and studies applications to image compression. It explains how to compute an optimal geometry with a fast algorithm, that requires $O(N^2(\log_2 N)^2)$ operations for an image of N^2 pixels. Numerical results show that optimized bandelet bases improve the image compression results that are obtained by wavelet bases.

2. NON-LINEAR APPROXIMATION OF IMAGES WITH WAVELETS

Sparse signal approximations can be obtained by decomposing signals in an orthonormal basis. Let us first consider the case of continuous spatial parameter images $f(x_1, x_2)$ before concentrating on discrete images. In an orthonormal basis $\mathcal{B} = \{g_m\}_{m\in\mathbb{N}}$, an image f can be approximated by the partial sum

$$f_M = \sum_{m\in I_M} \langle f, g_m\rangle\, g_m\,,$$

where I_M is an index set of M elements. To minimize the error

$$\|f - f_M\|^2 = \sum_{m\notin I_M} |\langle f, g_m\rangle|^2 \tag{1}$$

the set I_M should correspond to the M largest inner products. This is equivalent to threshold with a threshold T_M that is adjusted to keep M coefficients:

$$I_M = \{m \in \mathbb{N} \;:\; |\langle f, g_m\rangle| > T_M\}\,.$$

For discrete images of N^2 pixels, the same approximations can be implemented in an orthonormal basis $\mathcal{B}$ of $\mathbb{R}^{N^2}$. For compression applications, the inner products are not thresholded but quantized and coded. Yet, it has been shown in Falzon and Mallat[6] that for a uniform quantization of step T_M, at high compression rates the quadratic distortion D is proportional to $\|f - f_M\|^2$ and the total bit budget R is proportional to M. The distortion rate $D(R)$ thus has an asymptotic decay that is the same as the approximation error $\|f - f_M\|^2$ as a function of M. The efficiency of thresholding estimators that remove additive white noises by representing the signal in the basis $\mathcal{B}$ also depends upon this approximation error.[7] For both applications, given some prior information on the properties of f, we thus want to find a basis $\mathcal{B}$ where $\|f - f_M\|^2$ converges quickly to zero when M increases. This is the case if there exists a small constant C and a large exponent α with

$$\|f - f_M\|^2 \leq C\, M^{-\alpha}\,. \tag{2}$$

Wavelet bases have been shown to be particularly efficient to approximate images. A separable wavelet basis is constructed from a one-dimensional wavelet $\psi(t)$ and a scaling function $\phi(t)$ which are dilated and translated

$$\psi_{j,m}(t) = \frac{1}{\sqrt{2^j}}\,\psi\Big(\frac{t - 2^j m}{2^j}\Big)$$

and

$$\phi_{j,m}(t) = \frac{1}{\sqrt{2^j}}\,\phi\Big(\frac{t - 2^j m}{2^j}\Big)\quad.$$

The resulting family of separable wavelets

$$\Big\{\phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2)\;\;,\;\;\psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2)\;\;,\;\;\psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2)\Big\}_{j\in\mathbb{Z},(m_1,m_2)\in\mathbb{Z}^2} \tag{3}$$

is an orthonormal basis of $\mathbf{L^2}(\mathbb{R}^2)$. To construct a basis over a subset Ω of $\mathbb{R}^2$, one must keep the wavelets whose support are inside Ω and modify appropriately the ones whose support intersect the boundary of Ω.[8] We shall

still write $\phi_{j,m}$ and $\psi_{j,m}$ the modified scaling functions and wavelets at the boundary, and the resulting basis of $\mathbf{L^2}(\Omega)$ can be written

$$\left\{\phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2) \;\;,\;\; \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2) \;\;,\;\; \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2)\right\}_{(j,m_1,m_2)\in I_\Omega} \tag{4}$$

where I_Ω is an index set that depends upon the geometry of the boundary of Ω.

If the image $f(x_1, x_2)$ is uniformly regular, which is measured by the fact that it is $\mathbf{C}^\alpha$ (α times continuously differentiable) and if the wavelet ψ has $p > \alpha$ vanishing moments then one can prove[9] that there exists a constant C such that the approximation f_M from M wavelets satisfies

$$\|f - f_M\|^2 \leq C\, M^{-\alpha} \;. \tag{5}$$

This decay rate is optimal is the sense that one can not find a basis for all $\mathbf{C}^\alpha$ functions f satisfy $\|f - f_M\|^2 = O(M^{-\beta})$ with $\beta > \alpha$.[9] However, wavelet bases are not the only bases to achieve the optimal rate (5).

If f is $\mathbf{C}^\alpha$ ($\alpha > 1$) everywhere but along curves of finite length where is is discontinuous, then the discontinuities create many fine scale wavelet coefficients of large amplitude and the error decay (5) is no more valid. However, one can still prove that there exists a constant C such that

$$\|f - f_M\|^2 \leq C\, M^{-1} \;. \tag{6}$$

This result extends to all images having a bounded total variation, which means that their level sets have a finite average length. Moreover, wavelet bases are optimal for bounded variation functions in the sense that there exists no basis that leads to an approximation error (2) with a decay exponent $\alpha > 1$ over all bounded variation functions.[10] Keeping fine scale wavelet coefficients near singularities can be interpreted as building an adaptive grid approximation where the image is locally approximated at a scale adapted to the image structure. For discrete images decomposed in a discrete wavelet basis, this adaptation leads a more sparse representation than with a block cosine basis, which explains why wavelets are more efficient more image compression.

Despite the optimality of wavelets for bounded variation functions, one can often improve the approximation performance of wavelet bases for images, by observing that the level sets of many images not only have a finite average but are regular geometric curves. Exploiting this geometric regularity can allow us to improve the representation. This is well illustrated by a simple example. Let Ω a subset of $[0,1]^2$ whose boundary $\partial\Omega$ is a piecewise $\mathbf{C}^2$ curve, with a finite number of corners, as illustrated in Figure 1(a). Suppose that $f(x_1, x_2)$ is a $\mathbf{C}^2$ function inside and outside Ω, which is discontinuous along $\partial\Omega$. One can verify that there exists C_1 and C_2 such that the wavelet approximation f_M satisfies

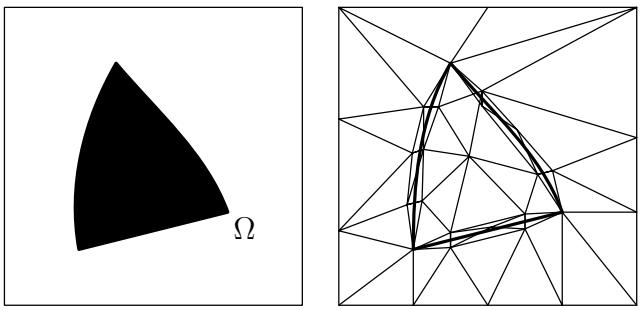
$$C_1\, M^{-1} \;\leq \|f - f_M\|^2 \leq C_2\, M^{-1} \;.$$

This can be improved with representations adapted to the image geometry. A simple example is obtained with a piece-wise linear approximation constructed over an adapted triangulation illustrated in Figure 1(b). The boundary $\partial\Omega$ is covered with narrow triangles whose widths are $O(M^{-2})$, and the inside and outside of Ω are covered by large triangles so that the total number of triangles is M. Over such a triangulation, one can construct a piecewise linear approximation f_M of f that satisfies

$$\|f - f_M\|^2 \leq C\, M^{-2} \;. \tag{7}$$

In this case the decay rate exponent $\alpha = 2$ is better than with wavelets and this approximation decay is identical to the one obtained when f is $\mathbf{C}^2$ over its whole support. The existence of discontinuities does not degrade the asymptotic decay of the approximation.

This example shows that exploiting the geometrical image regularity can lead to much smaller approximation errors for a fixed number of approximation elements M. One must also include the fact that these approximation elements (triangles) are defined by multiple parameters (orientations, width, length), which can be incorporated in the constant C in (7) and which does not affect the asymptotic decay rate. The major difficult is of course that images are much more complex, and adaptive triangulations are extremely hard to construct for such complex



(a) (b)

Figure 1. (a) Indicator function of a domain Ω, in black. (b) Adapted triangulation that covers the boundary with narrow triangles.

functions. Moreover, one would like to extend this result for regularity indexes $\alpha \geq 2$. If the boundary $\partial\Omega$ is a $\mathbf{C}^\alpha$ curve and if f is bounded and $\mathbf{C}^\alpha$ inside and outside Ω then one would like to find a geometrical approximation f_M from M elements such that

$$\|f - f_M\|^2 \leq C\, M^{-\alpha} \ . \tag{8}$$

We shall see that bandelet bases are able to achieve this optimal decay rate.

3. GEOMETRIC IMAGE REPRESENTATIONS

The construction of geometric image representation is a very active research area where many beautiful, and innovative ideas have been tested. Summarizing the different approaches will help understand the major difficulties.

In the computer vision community, Carlsson[1] proposed in 1988 an edge based image representation which measures the image jumps across curves in the images, called edges. An image approximation is then calculated by imposing the same jumps along the edge and by computing values between edges with a diffusion process. Many edge based image representations have then been elaborated along similar ideas,[11,12] with different edge detection procedures and image approximations using jump models along these edges. To refine these models, multiscale edge representations using wavelet maxima[2] or an edge adapted multiresolution[3] have also been studied. Edge based image representations with non-complete orthonormal families of foveal wavelets or finger prints have also been introduced and studied to reconstruct the main image edge structures. To stabilize the edge detection, global optimization procedures have also been elaborated by Donoho,[13] Shukla et al.[14] and Wakin et al.[15] The optimal configuration of edges is then calculated with an image segmentation over dyadic squares using fast dynamic programming algorithms over quad-trees.

A major difficulty that face all edge based approaches is that sharp image transitions often do not correspond to discontinuous jumps along edge curves. On one hand, the optical diffraction produces an averaging effect which blurs the grey level discontinuities along occlusion boundaries, and on the other hand many sharp transitions are produced by texture variations that are not aggregated along geometrical curves. Edge based algorithms are thus often not more efficient than a standard wavelet image representation over a wide range of approximation precision.

All the approaches previously described are adaptive in the sense that the image representation is adapted to a geometry estimated from the image. Surprisingly, a remarkable result of Candes and Donoho[4] shows that one can construct a non adaptive representation that takes advantage of the image geometrical regularity by decomposing it in a fixed basis or frames of curvelets. Curvelet families are composed of multiscale elongated and rotated functions that defines bases or frames of $\mathbf{L^2}(\mathbb{R}^2)$. They proved that that an approximation f_M with M curvelets of an image f having discontinuities (blurred or not) along $\mathbf{C}^2$ curves produces an error that satisfy

$$\|f - f_M\|^2 \leq C\, M^{-2}\, (\log_2 M)^3 \ . \tag{9}$$

By comparing this to (8) we see that this approximation result is nearly asymptotically optimal up the $(\log_2 M)^3$ factor. Do and Vetterli[5] used similar ideas to construct contourlets that can be computed with a perfect reconstruction filter bank procedure. However, the beautiful simplicity due to the non-adaptivity of curvelets has a cost: curvelet approximations loose their excellent properties when the image is composed of edges which are not $\mathbf{C}^2$. For example, if the discontinuities are along piecewise regular curves which has corners or junctions, then the approximation decay (9) is not valid anymore. If edges are along irregular curves of finite length (bounded variation functions) then curvelets approximations are not as precise as wavelet approximations. Finally, if the edges are along curves whose regularity is $\mathbf{C}^\alpha$ with $\alpha > 2$ then the approximation decay rate exponent remains 2 and does not reach the optimal value α.

In image processing applications, we generally do not know in advance the geometrical image regularity. It is therefore necessary to find approximation schemes that can adapt themselves to varying degrees of regularity. Our goal is thus to construct an adaptive image approximation f_M of f, with M parameters, which satisfies an optimal decay rate $\|f - f_M\| \leq C\, M^{-\alpha}$. The exponent α is a priori unknown and specifies the geometrical image regularity.

4. BANDELETS ALONG GEOMETRIC FLOWS

Instead of describing the image geometry through edges, which are most often ill-defined, the geometry of images is characterized by a field of vectors which locally give a direction where the image has regular variations. This vector field is called a *geometric flow*. Section 4.1 constructs bandelet bases over appropriate geometric flows. Section 4.2 explains relates the optimization of the geometric flow to the precision of image approximations with few parameters.

4.1. Block Bandelet Basis

This section describes the construction of bandelet bases from a wavelet basis that is warped along the geometric flow, to take advantage of the image regularity along this flow. Conditions are imposed on the geometric flow to obtain orthonormal bandelet bases.

In a region Ω, a geometric flow is a vector field $\vec{\tau}(x_1, x_2)$ defined at each $(x_1, x_2) \in \Omega$, which indicates locally a direction in which the image intensity f has regular variations. If the image intensity is uniformly regular in the neighborhood of a point then this direction is not uniquely defined. Some form of global regularity is therefore imposed on the flow to specify it uniquely. To construct orthogonal bases with the resulting flow, a first regularity condition imposes that the flow is either parallel vertically, which means that $\vec{\tau}(x_1, x_2) = \vec{\tau}(x_1)$, or parallel horizontally and hence $\vec{\tau}(x_1, x_2) = \vec{\tau}(x_2)$. To maintain enough flexibility, this parallel condition is imposed within subregions Ω_i of the image support. The image support $\mathcal{S}$ is partitioned into regions $\mathcal{S} = \cup_i \Omega_i$, and within each Ω_i the flow is either parallel horizontally or vertically. If the image intensity f is uniformly regular over a whole region Ω_i then a geometric flow is meaningless and is therefore not defined.

Figure 2 gives an example where the image is partitioned into square regions that are small enough so that each region Ω_i includes at most one contour. In each region including a contour piece, the flow is parallel to the tangents of the contour curve. Bandelets are constructed in these regions by warping separable wavelet bases so that they follow the lines of flow, and by applying a *bandeletization* procedure that takes advantage of the image regularity along the geometric flow. The next section explains how to optimize this image segmentation and compute the flow over each region.

If there is no geometric flow over a region Ω, which indicates that the image restriction to Ω has an isotropic regularity then this restriction is approximated in the separable wavelet basis (4) of $\mathbf{L^2}(\Omega)$. If a geometric flow is calculated in Ω, this wavelet basis is replaced by a bandelet basis. We first explain how to construct the bandelet basis when the flow is parallel in the vertical direction: $\vec{\tau}(x_1, x_2) = \vec{\tau}(x_1)$. We normalize the flow vectors so that it can be written $\vec{\tau}(x_1) = (1\,,\, c'(x_1))$. Let $x_{\min} = \inf_{x_1}\{(x_1, x_2) \in \Omega\}$. A flow *flow line* is defined as an integral curve of the flow, whose tangents are parallel to $\vec{\tau}(x_1)$. Since the flow is parallel vertically, a flow line is a set of point $(x_1, x_2) \in \Omega$ which satisfy $x_2 = c(x_1) + c_0$ where

$$c(x_1) = \int_{x_{\min}}^{x_1} c'(x_1)\, dx_1 \ ,$$

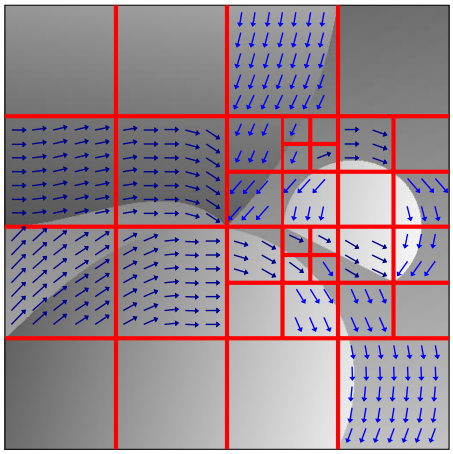


Figure 2. Example of an adapted dyadic squares segmentation of an image and and its corresponding flow.

and c_0 is a translation parameter. By construction of the flow, the image grey level has regular variations along these flow lines. To take advantage of this regularity with wavelets, the separable wavelets in (4) are warped with an operator W performing translations along x_2. The warped image

$$Wf(x_1, x_2) = f(x_1, x_2 + c(x_1))$$

is regular along the horizontal lines for x_2 fixed and x_1 varying. Over the warped region

$$\Omega' = W\Omega = \{(x_1, x_2) \ : \ (x_1, x_2 + c(x_1)) \in \Omega\}$$

we define a separable orthonormal wavelet basis of $\mathbf{L^2}(\Omega')$:

$$\left\{\phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2) \ \ , \ \ \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2)) \ \ , \ \ \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2))\right\}_{(j,m_1,m_2)\in \mathbf{I}_{\Omega'}} . \tag{10}$$

Since the warping operator W is an orthogonal operator, applying its inverse to each of these wavelets yields an orthonormal basis of $\mathbf{L^2}(\Omega)$, that is called a *warped wavelet basis*:

$$\left\{\phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2 - c(x_1)) \ \ , \ \ \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2 - c(x_1)) \ \ , \ \ \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2 - c(x_1))\right\}_{(j,m_1,m_2)\in \mathbf{I}_{\Omega'}} . \tag{11}$$

Warped wavelets are separable along the x_1 variable and along the $x_2' = x_2 - c(x_1)$ variable which follows the geometric flow lines within Ω.

The flow is calculated so that f is regular along the flow lines in Ω. Suppose that $f(x_1, x_2 + c(x_1))$ is $\mathbf{C}^\alpha$ function of x_1 for all x_2 fixed, within Ω. Since $\psi(t)$ has $p > \alpha$ vanishing moments, one can verify[16] that

$$|\langle f(x_1, x_2)\,,\, \phi_{j,m_1}(x_1)\psi_{j,m_2}(x_2 - c(x_1))\rangle| = O(2^{j(\alpha+1)})$$

and

$$|\langle f(x_1, x_2)\,,\, \psi_{j,m_1}(x_1)\psi_{j,m_2}(x_2 - c(x_1))\rangle| = O(2^{j(\alpha+1)}) \ .$$

However, the third type of wavelet coefficients have a slower decay when the scale 2^j decreases:

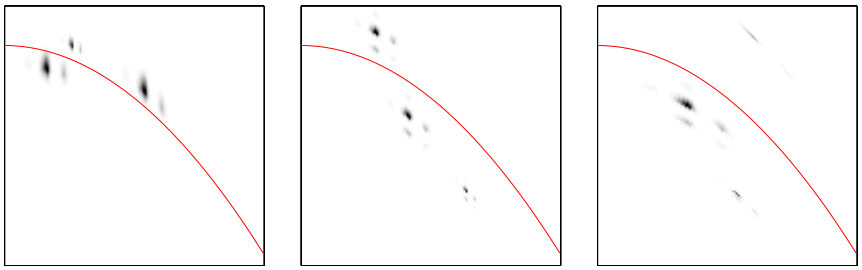
$$|\langle f(x_1, x_2)\,,\, \psi_{j,m_1}(x_1)\phi_{j,m_2}(x_2 - c(x_1))\rangle| = O(2^j) \ , \tag{12}$$

because ϕ has no vanishing moment and thus can not take advantage of the regularity of f along the flow lines.

To improve this result, it is necessary to replace the family of orthogonal scaling functions $\{\phi_{j,m_2}(x_2')\}_{m_2}$ with $x_2' = x_2 - c(x_1)$, by an equivalent family of orthogonal functions, that have vanishing moments and can thus take

advantage of the regularity of f along the flow lines. We know that $\{\phi_{j,m_2}(x_2')\}_{m_2}$ is an orthonormal basis of a multiresolution space which also admits an orthonormal basis of wavelets $\{\psi_{l,m_2}(x_2')\}_{l>j,m_2}$. This suggests replacing the orthogonal family $\{\psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2')\}_{j,m_1,m_2}$ by the family $\{\psi_{j,m_1}(x_1)\,\psi_{l,m_2}(x_2')\}_{j,l>j,m_1,m_2}$ which generates the same space. This is called a *bandeletization*, which can be implemented with a simple discrete wavelet transform.[17] The functions $\psi_{j,m_1}(x_1)\psi_{l,m_2}(x_2')$ are called bandelets because their support is parallel to the flow lines and is more elongated ($2^l > 2^j$) in the direction of the geometric flow. Figure 3 gives examples of such bandelets. Inserting these bandelets in the warped wavelet basis (11) yields a bandelet orthonormal basis of $\mathbf{L^2}(\Omega)$:

$$\left\{\phi_{j,m_1}(x_1)\psi_{j,m_2}(x_2-c(x_1))\;\;,\;\;\psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2-c(x_1))\;\;,\;\;\psi_{j,m_1}(x_1)\,\psi_{l,m_2}(x_2-c(x_1))\right\}_{j,l>j,m_1,m_2}. \quad (13)$$



(a) (b) (c)

Figure 3. The dark spots display the amplitude of bandelets in a region Ω having a vertically parallel flow. (a): Bandelets $\phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2-c(x_1))$. (b): Bandelets $\psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2-c(x_1))$. (c): Bandelets $\psi_{j,m_1}(x_1)\,\psi_{l,m_2}(x_2-c(x_1))$.

If $f(x_1, x_2+c(x_1))$ is a $\mathbf{C}^\alpha$ function of x_1 for all x_2 fixed in Ω then one can prove[16] that the bandelet coefficients are much smaller than the warped wavelet coefficients (13) at fine scales:

$$|\langle f(x_1,x_2)\,,\,\psi_{j,m_1}(x_1)\psi_{l,m_2}(x_2-c(x_1)\rangle| = O(\min(2^j, 2^{l(\alpha+1)}))\;.$$

This decay is sufficient to obtain approximation error from the largest bandelet coefficients which has the optimal decay rate (8).

If the geometrical flow in Ω is parallel in the horizontal direction, meaning that

$$\vec{\tau}(x_1,x_2) = \vec{\tau}(x_2) = (c'(x_2)\,,\,1)$$

then the same construction applies by inverting the roles of the variables x_1 and x_2. Let $x_{\min} = \inf_{x_2}\{(x_1,x_2)\in\Omega\}$ and

$$c(x_2) = \int_{x_{\min}}^{x_2} c'(x_2)\,dx_2\;.$$

A warped wavelet basis is constructed from a separable wavelet basis of $\Omega' = \{(x_1,x_2)\;:\;(x_1+c(x_2),x_2)\in\Omega\}$, and is defined by:

$$\left\{\phi_{j,m_1}(x_1-c(x_2))\,\psi_{j,m_2}(x_2)\;\;,\;\;\psi_{j,m_1}(x_1-c(x_2))\,\phi_{j,m_2}(x_2)\;\;,\;\;\psi_{j,m_1}(x_1-c(x_2))\,\psi_{j,m_2}(x_2)\right\}_{(j,m_1,m_2)\in\mathbf{I}_{\Omega'}} \quad (14)$$

The bandeletization replaces each family of scaling functions $\{\phi_{j,m_1}(x_1-c(x_2))\}_{m_1}$ by a family of orthonormal wavelets that generates the same space. The resulting bandelet orthonormal basis of $\mathbf{L^2}(\Omega)$ is:

$$\left\{\psi_{l,m_1}(x_1-c(x_2))\,\psi_{j,m_2}(x_2)\;\;,\;\;\psi_{j,m_1}(x_1-c(x_2))\,\phi_{j,m_2}(x_2)\;\;,\;\;\psi_{j,m_1}(x_1-c(x_2))\,\psi_{j,m_2}(x_2)\right\}_{j,l>j,m_1,m_2}. \quad (15)$$

Given a partition of the image support $\mathcal{S} = \cup_i \Omega_i$ with the corresponding geometric flow, this strategy defines a bandelet or wavelet (if there is no flow) orthonormal basis in each $\mathbf{L^2}(\Omega_i)$. The union of these bases is a block orthonormal basis of $\mathbf{L^2}(\mathcal{S})$. The orthogonality of the wavelets and bandelets can be relaxed. If the one-dimensional wavelet ψ and the scaling function ϕ yield a biorthogonal orthogonal wavelet basis[18] then the same construction defines a biorthogonal bandelet basis of each $\mathbf{L^2}(\Omega_i)$.[16]

4.2. Optimized Geometry for Approximations

A major difficulty is to compute an appropriate image geometry. For image approximation, the best geometry is the one that leads to the best approximation. This means finding an approximation f_M from M parameters that minimizes the approximation error $\|f - f_M\|$. For bandelet approximations, the M parameters include the bandelet coefficients that are used to compute f_M as well as the parameters that specify the image partition and the geometric flow in each region.

To represent the image partition with few parameters, and be able to compute an optimal partition with a fast algorithm, we restrict ourselves to partitions in squares of varying dyadic sizes. A dyadic squares image segmentation is obtained by successive subdivisions of square regions into four squares of twice smaller width. For a square image support of width L, a square region of width $L\,2^{-j}$ is represented by a node at the depth j of a quad-tree. A square subdivided into four smaller squares corresponds to a node having four children in the quad-tree. Figure 4 gives an example of a dyadic square image segmentation with the corresponding quad-tree.

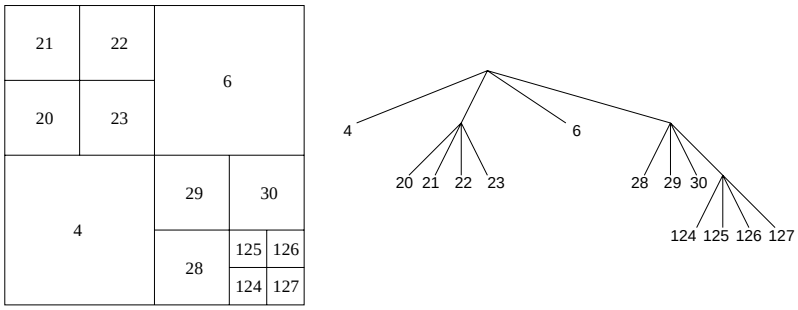


Figure 4. Example of dyadic square image segmentation. Each leave of the corresponding quad-tree corresponds to a square region having the same index number.

In each region Ω of the segmentation, one must decide if there should be a geometric flow, if this flow should be parallel in the horizontal or in the vertical direction, and what should be this flow. If there is a flow, it should be optimized to guarantee that the image has regular variations along the flow lines. This optimization is performed by minimizing the partial derivatives of a filtered image along the flow. Given a regularizing filter $\theta(x_1, x_2)$ that will be specified later, we minimize a flow energy:

$$\mathcal{E}(\vec{\tau}) = \int_{\Omega} \left| \frac{\partial (f \star \theta)(x_1, x_2)}{\partial \vec{\tau}(x_1, x_2)} \right|^2 dx_1\, dx_2 \ . \tag{16}$$

If the geometrical flow is chosen to be parallel in the vertical direction then $\vec{\tau}(x_1, x_2) = (1\,,\, c'(x_1))$ and the resulting flow energy (16) can be written:

$$\mathcal{E}(\vec{\tau}) = \int_{\Omega} \left| f \star \frac{\partial \theta}{\partial x_1}(x_1, x_2) + c'(x_1)\, f \star \frac{\partial \theta}{\partial x_2}(x_1, x_2) \right|^2 dx_1\, dx_2 \ . \tag{17}$$

To compute a flow that depends essentially upon the the sharpest image variations, we choose a separable filter $\theta(x_1, x_2) = \theta_0(x_1)\, \theta_1(x_2)$, where $\theta_0(x_1)$ is a high-pass filter and $\theta_1(x_2)$ is a low-pass filter. The choice of these filters depends upon the application, and in denoising applications they must must be optimized as part of the global optimization of the geometry.

A flow parallel in the horizontal direction can be written $\vec{\tau}(x_1, x_2) = (c'(x_2), 1)$ and the resulting flow energy is

$$\mathcal{E}(\vec{\tau}) = \int_{\Omega} \left| c'(x_2)\, f \star \frac{\partial \theta}{\partial x_2}(x_1, x_2) + f \star \frac{\partial \theta}{\partial x_1}(x_1, x_2) \right|^2 dx_1\, dx_2 \ , \tag{18}$$

with $\theta(x_1, x_2) = \theta_1(x_1)\, \theta_0(x_2)$.

In approximation or compression applications, the flow must be represented by a a limited number of parameters, and $c'(t)$ is calculated as an expansion over translated box splines functions $b(x)$ dilated by a scale

factor 2^l:

$$c'(t) = \sum_n \alpha_n \, b(2^{-l}t - n) \ .$$

A box spline $b(t)$ of degree m is obtained by convolving the indicator function $\mathbf{1}_{[-1/2,1/2]}$ with itself $m+1$ times. The parameters α_n are computed by minimizing the quadratic forms (17) or (18) depending upon the orientation of the flow, which is done by solving the corresponding linear systems. The scale parameter 2^l which defines the regularity of the flow is another parameter that must also be optimized.

The mathematical study[16] explains how to compute a segmentation and optimize the scale 2^l of the geometric flow to minimize the approximation error $\|f - f_M\|^2$ for a fixed number M of parameters, including the bandelet coefficients and all coefficients needed to specify the geometric flow. This is performed with a fast dynamic programming algorithm that is explained in Section 5.2 in the context of image compression. If the image f has contours that are $\mathbf{C}^\alpha$ curves which meet at corners but are not tangential, and if f is $\mathbf{C}^\alpha$ away from these curves then with an appropriate modification of the boundary wavelets, one can then prove[16] that the resulting optimal approximation satisfies the optimal asymptotic decay rate:

$$\|f - f_M\|^2 \leq C \, M^{-\alpha} \ ,$$

with no prior knowledge on the value of α. In this paper, we concentrate on discrete fast algorithms and applications to image compression.

5. FAST DISCRETE BANDELET TRANSFORM AND IMAGE COMPRESSION

5.1. Fast Discrete Bandelets Transform

Bandelets in a region Ω are computed by applying a bandeletization to warped wavelets in Ω, which are separable wavelets along a fixed direction (horizontal or vertical) and along the flow lines. A fast discrete bandelet transform can therefore be computed by using a fast separable wavelet transform along this fixed direction and along the image flow lines. The block bandelet basis of Section 4.1 is constructed with separate warped wavelet bases inside each region. In image processing applications, this creates border effects when modifying the corresponding bandelet coefficients. To avoid these border effects, we use a discrete warped wavelet transform which goes across the region boundaries while keeping perfect reconstruction properties.

The fast discrete bandelet transform associated to an image partition $\cup_i \Omega_i$ includes three steps:

- An image resampling, that computes the image sample values along the flow lines in each region Ω_i of the partition.
- A warped wavelet transform with a subband filtering along the flow lines, which goes across the region boundaries. The subband filtering procedure uses the lifting scheme on irregular grids introduced by Bernard.[19]
- A bandeletization that transforms the warped wavelet coefficient to compute bandelet coefficients along the flow lines.

The fast inverse bandelet transform includes the three inverse steps:

- An inverse bandeletization that recovers the warped wavelet coefficient along the flow lines.
- An inverse warped wavelet transform with an inverse subband filtering.
- An inverse resampling which computes the image samples along the original grid from the samples along the flow lines in each region Ω_i.

The fast algorithms that implement the three steps of this discrete bandelet transform and their inverse require a total of $O(N^2)$ operations for an image of N^2 pixels. A detailed description can be found in the reference.[17]

5.2. Application to Image Compression

A major difficulty of geometrical representations is to compute an appropriate geometry from the image. For a bandelet transform, the geometry is defined by the image partition in regions Ω_i and by the geometric flow within each region. The notion of optimal geometry can be well defined in the context of a precise application, and we shall concentrate here on image compression. The optimization of the geometry for noise removal is studied in.[17]

The goal is to compute a geometry that optimizes the distortion-rate of an image compression transform code in the corresponding bandelet basis. The best geometry is defined as the one that yields the most compressed image code. This optimization requires to establish the link between the image geometry and the distortion-rate of the image coder. The geometry is optimized by a fast algorithm that minimizes the distortion-rate with $O(N^2(\log_2 N)^2)$ operations for an image of N^2 pixels, because the geometry is structured by aggregating nearly independent building blocks. To concentrate on the properties of the bandelet transform itself, we use a relatively simple transform coder with a scalar quantization and an entropy coding of all coefficients. A comparison is made with the same coder applied to a wavelet transform.

Let $\mathcal{D} = \{\mathcal{B}^\gamma\}_{\gamma\in\Gamma}$ be the dictionary of all possible biorthogonal bandelet bases that we write $\mathcal{B}^\gamma = \{g_m^\gamma\}_{1\leq m\leq N^2}$. The index γ is associated to the geometry of the basis, which specifies the image partition in dyadic squares $[1,N]^2 = \cup_i \Omega_i$ and the geometric flow in each square Ω_i. We denote by $\widetilde{\mathcal{B}}^\gamma = \{\tilde{g}_m^\gamma\}_{1\leq m\leq N^2}$ the biorthogonal bandelet basis. For any $\gamma \in \Gamma$

$$f = \sum_{m=1}^{N^2} \langle f, g_m^\gamma\rangle\, \tilde{g}_m^\gamma \ .$$

Finding the best geometry for image compression can be interpreted as a search for a best bandelet basis in a specified dictionary.

The transform code is implemented with a uniform scalar quantizer $Q(x)$ with bins of size Δ, using a zero-bin twice larger than the others:

$$Q(x) = \begin{cases} 0 & \text{if } |x| \leq \Delta \\ (\lfloor x/\Delta \rfloor + 1/2)\times\Delta & \text{otherwise} \end{cases} .$$

The restored image from quantized coefficients is:

$$\tilde{f} = \sum_{m=1}^{N^2} Q(\langle f, g_m^\gamma\rangle)\, \tilde{g}_m^\gamma \ .$$

The resulting distortion is calculated with a Euclidean squared norm $D = \|f - \tilde{f}\|^2$. The total number of bits R to code $\tilde{f}$ is equal to the number of bits R_c to code the N^2 quantized coefficients $\{Q(\langle f, g_m^\gamma\rangle)\}_{1\leq m\leq N^2}$ plus the number of bits to code the geometry. The distortion D depends upon R through the value of Δ and through the choice of the geometry.

The geometric flow in a region Ω_i is a vector field $\vec{\tau}_i[n_1,n_2]$ defined over the image sampling grid. If the flow is parallel vertically then it can be written

$$\vec{\tau}_i[n_1,n_2] = \vec{\tau}_i[n_1] = (1\ ,\ c_i'[n_1])\ ,$$

where $c_i'[n_1]$ is the relative displacement of the image grey levels in Ω_i along the line n_1 with respect to the line $n_1 - 1$. If there is a flow, it should be optimized to guarantee that the image has regular variations along the flow lines. This optimization is performed by minimizing the partial derivatives of a filtered image along the flow. Given a regularizing filter $\theta(x_1,x_2)$, we minimize

$$\mathcal{E}(\vec{\tau}) = \sum_{(n_1,n_2)\in\Omega_i} \left| f \star \frac{\partial\theta}{\partial x_1}[n_1,n_2] + c_i'[n_1]\, f \star \frac{\partial\theta}{\partial x_2}[n_1,n_2] \right|^2 . \tag{19}$$

We construct θ from the one-dimensional wavelet and scaling function $\theta(x_1, x_2) = \psi(2^2 x_1)\,\phi(x_2)$. If the geometric flow is parallel horizontally in Ω_i then the flow can be written $\vec{\tau}_i[n_1, n_2] = (c_i'[n_2]\,, 1)$ and this flow vector is calculated by minimizing the quadratic image variations along the flow:

$$\mathcal{E}(\vec{\tau}) = \sum_{(n_1,n_2)\in\Omega_i} \left| f \star \frac{\partial\theta}{\partial x_2}[n_1, n_2] + c_i'[n_2]\, f \star \frac{\partial\theta}{\partial x_1}[n_1, n_2] \right|^2 , \tag{20}$$

with $\theta(x_1, x_2) = \phi(x_1)\,\psi(2^2 x_2)$.

Since the geometric flow is assumed to be regular, the displacement $c_i'[p]$ is specified by its decomposition coefficients α_n over a family of translated box splines, which are dilated by a scale factor 2^l:

$$c_i'[p] = \sum_n \alpha_n\, b(2^{-l}p - n)\ . \tag{21}$$

In a square region Ω_i of width 2^k, there are 2^{k-l} box spline coefficients α_n. The coefficients α_n that minimize (19) or (20) are computed by solving the linear systems associated to this quadratic minimization. These coefficients α_n are uniformly quantized. The quantization step adjusts the precision of the geometrical displacement $c_i'[p]$. It is set to be of the order of 1/8 of a pixel. The quantized coefficients are coded with a fixed length code.

To optimize the coder, we use the Lagrangian approach proposed by Ramchandran and Vetterli,[20] which finds the best basis that minimizes $D(R) + \lambda R$, where λ is a Lagrange multiplier. If $D(R)$ is convex, which is usually the case, by letting λ vary we are guaranteed to minimize $D(R)$. If $D(R)$ is not convex, then this strategy leads to a $D(R)$ that is at most a factor 2 larger than the minimum.

For a given image and parameter λ, we want to find the image segmentation and the geometric flow which defines a bandelet basis that minimizes $D(R) + \lambda R$. This optimization is implemented by a fast dynamic programming algorithm which uses the fact that

$$D + \lambda R \approx \sum_i \Big(D_i + \lambda R_i \Big)\ , \tag{22}$$

where $D_i = \|f - \tilde{f}\|^2_{\Omega_i}$ is the Euclidean norm restricted to a region Ω_i of the image partition, and R_i is the number of bits needed to code the bandelet coefficients and the geometry in Ω_i. Since the $\{\Omega_i\}_i$ define a partition of the image support $[1, N]^2$, we clearly have $D = \sum_i D_i$. The paper[17] describes a coder of the quantized bandelet coefficients and of the geometric information so that $R \approx \sum_i R_i$, from which we derive (22). In order to minimize $D + \lambda R$, we must adjust the size of the quantization bin Δ. One can verify[17] that the optimum is obtained when

$$\Delta = \sqrt{2\,\lambda\,\gamma_0}\ .$$

To find the bandelet basis that minimizes $D + \lambda R$, we begin by computing the geometric flow which minimizes $D_i + \lambda R_i$ in all possible dyadic regions Ω_i. This means finding if the geometric flow should exist or not, if it should be parallel horizontally or vertically, and what is the best scale parameter 2^l. All these possibilities are tested in each Ω_i and we retain the one that minimizes $D_i + \lambda R_i$. One can verify[17] that it requires $O((\log_2 N)^2\, N^2)$ operations.

At this point we know the optimal geometry within each possible dyadic region Ω_i of the image, and the corresponding minimum value $D_i + \lambda R_i$. Observe that for any image region Ω_i, a partition into subregions $\Omega_i = \cup_l \Omega_l$ gives a better distortion rate if

$$D_i + \lambda R_i \geq \sum_l D_l + \lambda R_l\ .$$

The optimal image partition into dyadic regions that minimizes the overall distortion rate $D + \lambda R$ can thus be computed with a bottom up algorithm along the branches of the segmentation quad-tree as in.[13–15]

The final image code is obtained by decomposing the image in the bandelet basis associated to the optimized partition and its geometric flow. To evaluate the performance of this bandelet compression algorithm, we compare

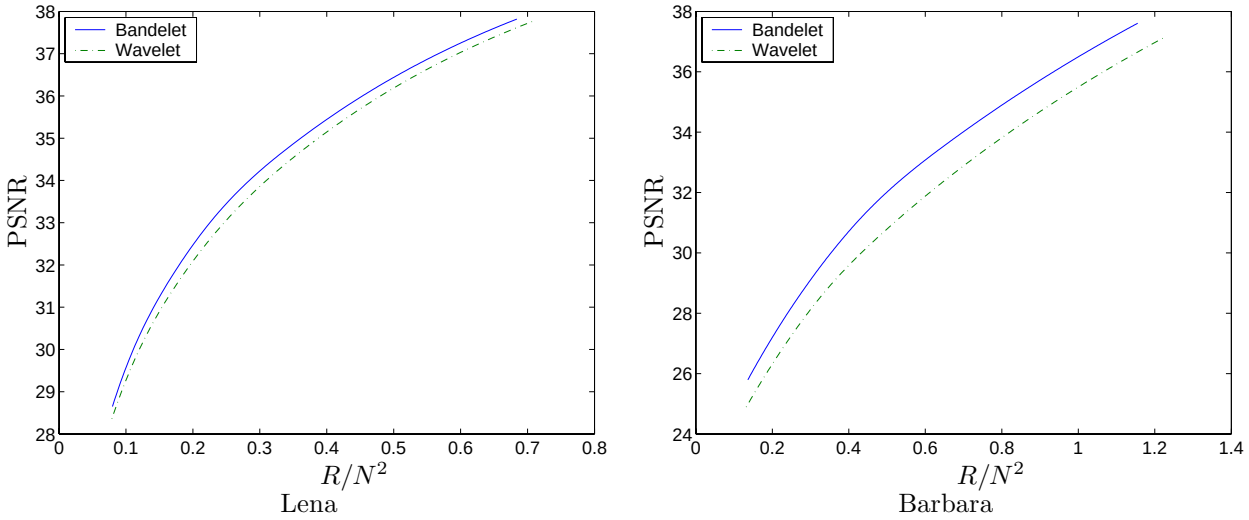


Figure 5. Distortion-rate the bandelet coder (full lines) and the wavelet coder (dashed lines) for the Lena and Barbara images. The distortion D is given by the PSNR in dB and the rate per pixel R/N^2 in bit per pixel. Over most bit rates, the bandelet coder reduces the distortion by approximatively .4 dB for Lena and by 1.1 dB for Barbara.

the distortion rate curve $D(R)$ with the distortion rate obtained in a 7/9 biorthogonal wavelet basis,[18] with the same quantization and entropy coding procedure. We do not incorporate the bit-plane strategy and the contextual coding procedure of JPEG-2000 to compare more easily the performance of the bandelet and wavelet bases themselves. Similar bit plane and contextual coding procedure can also be applied to bandelet coefficients.

Figure 5 shows these distortion rate curves for Lena and Barbara. The bandelet coder outperforms the wavelet coder by about .4 dB for Lena and 1.1 dB for Barbara. It is important to observe that this remains valid for a bit rate R/N^2 going from .15 bit per pixel to 1 bit per pixel, which covers the whole range of practical applications. From a visual quality point of view, the difference of performance between the two coders is more impressive as it can be seen from the images shown in Figure 6. Eventhough the bandelet coder introduces errors, the restored images have a regular geometry along the direction of the computed flow, and the resulting error is hardly visible. On the contrary, wavelets introduce visible ringing effects that are distributed the square grids of the wavelet sampling, which partly destroys the geometrical regularity. As a result, the bandelet compressed images have a better visual quality than their wavelet counterparts.

The main inefficiency of the current bandelet scheme comes from boundary effects between regions having different geometric flow. We use a bandelet transform that includes vectors that go across regions and thus produces no compression artefacts at the boundary of such regions. However, as it can be seen in Figure 2, the direction of the flow is typically discontinuous across the boundaries of the dyadic square regions. As a consequence of these orientation discontinuities, the transform does not capture fully the image regularity at these locations. Removing this geometric discontinuity with a multiscale flow is currently under investigation.

6. CONCLUSION

A central idea in the construction of bandelets is to define the geometry as a vector field, as opposed to a set of edge curves. This vector field plays the same role as motion vectors in video image sequences. It indicates the direction of displacement of grey level values, not in time but in space. Like in video image coding, this geometry is simplified by an image segmentation in squares, whose sizes are adapted to the local image structures.

The geometry of bandelet bases is not calculated a priori but by optimizing the resulting application, whether it is image compression or noise removal,[17] with a fast best basis search algorithm. As a result, bandelet bases improve the image compression and noise removal results obtained with wavelet bases. For video image sequences, a three-dimensional time-space geometric flow should be defined to construct bandelet bases that are adapted to

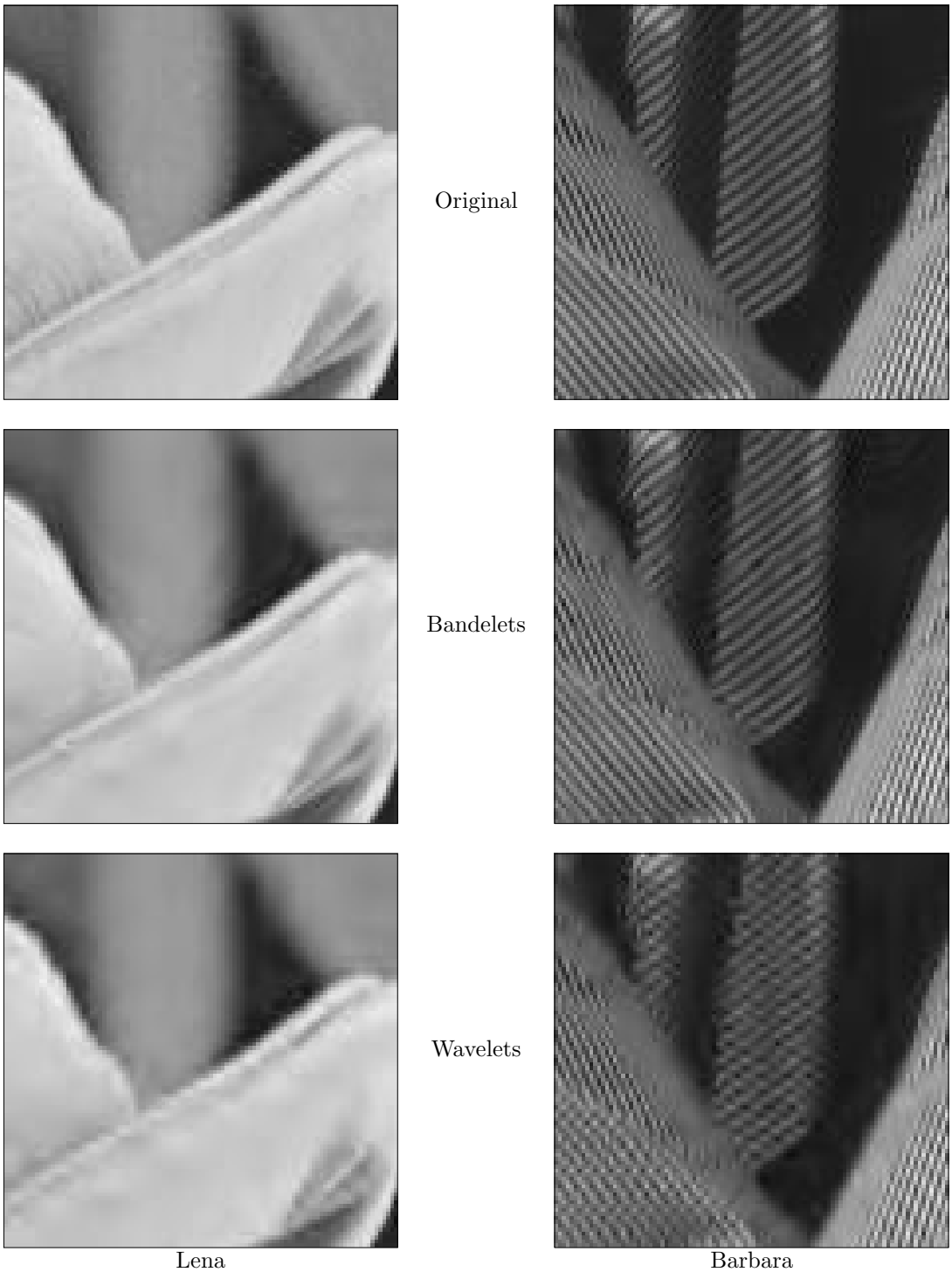


Figure 6. Zoom on Lena (left) and Barbara (right) images, reconstructed with a bandelet coder and a wavelet coder using the same number of bits. The bit budget is $R/N^2 = .22$ bit per pixel for Lena and $R/N^2 = .45$ bit per pixel for Barbara (on 512^2 pixels). The PSNR is respectively 30.0 dB and 29.6 dB with a bandelet and a wavelet compression of Lena. For Barbara, the PSNR is respectively 31.2 dB and 30.1 dB with a bandelet and a wavelet compression.

the space-time geometry of the sequence. This is a possible approach to improve the current video compression standard.

REFERENCES

1. S. Carlsson, "Sketch based coding of grey level images," *Signal Processing* **15**(1), pp. 57–83, 1988.
2. S. Mallat and S. Zhong, "Wavelet transform maxima and multiscale edges," in *Wavelets and their Applications*, B. R. et al., ed., Jones and Bartlett, Boston, 1992.
3. A. Cohen and B. Matei, "Nonlinear subdivisions schemes: Applications to image processing," in *Tutorial on multiresolution in geometric modelling*, A. Iske, E. Quack, and M. Floater, eds., Springer, 2002.
4. E. Candès and D. Donoho, "Curvelets: A surprisingly effective nonadaptive representation of objects with edges," in *Curves and Surfaces fitting*, L. L. Schumaker, A. Cohen, and C. Rabut, eds., Vanderbilt University Press, 1999.
5. M. N. Do and M. Vetterli, "Contourlets," in *Beyond Wavelets*, J. Stoeckler and G. V. Welland, eds., Academic Press, 2002. (to appear).
6. F. Falzon and S. Mallat, "Analysis of low bit rate image transform coding," *IEEE Transaction on Signal Processing* , January 1998.
7. D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika* **81**, pp. 425–455, Décembre 1994.
8. A. Cohen, W. Dahmen, and R. DeVore, "Multiscale decompositions on bounded domains," Tech. Rep. 113, IGPM, May 1995.
9. R. DeVore, "Nonlinear approximation," *Acta. Numer.* **7**, pp. 51–150, 1998.
10. A. Cohen, A. DeVore, and H. Petrushec, P.and Xi, "Non linear approximation and the space BV(R2)," *Amer. J. Math.* (121), pp. 587–628, 1999.
11. J. Elder, "Are edges incomplete?," *International Journal of Computer Vision* **34**(2/3), pp. 97–122, 1999.
12. X. Xue and X. Wu, "Image representation based on multi-scale edge compensation," in *IEEE Internat. Conf. on Image Processing*, 1999.
13. D. Donoho, "Wedgelets: Nearly-minimax estimation of edges," *Ann. Statist* **27**, pp. 353–382, 1999.
14. R. Shukla, P. L. Dragotti, M. N. Do, and M. Vetterli, "Rate-distortion optimized tree structured compression algorithms for piecewise smooth images," *IEEE Transactions on Image Processing* , January 2003. (submitted).
15. M. Wakin, J. Romberg, H. Choi, and R. Baraniuk, "Rate-distortion optimized image compression using wedgelets," in *IEEE International Conference on Image Processing*, September 2002.
16. E. Le Pennec and S. Mallat, "Non linear image approximation with bandelets," tech. rep., CMAP / École Polytechnique, 2003.
17. E. Le Pennec and S. Mallat, "Sparse geometrical image representation with bandelets," tech. rep., CMAP / École Poyltechnique, 2003.
18. A. Cohen, I. Daubechies, and J. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Comm. in Pure and Appl. Math* **45**, 1992.
19. C. Bernard and E. Le Pennec, "Adaptation of regular grid filterings to irregular grids," Tech. Rep. 500, CMAP / École Polytechnique, december 2002.
20. K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Transactions on Image Processing* **2**, pp. 160–175, April 1993.