Orthogonal Bandlet Bases for Geometric Images Approximation

Stéphane Mallat, Gabriel Peyré

Orthogonal Bandelet Bases for Geometric Images Approximation

STÉPHANE MALLAT
CMAP, École Polytechnique

AND

GABRIEL PEYRÉ
CEREMADE, Université de Paris - Dauphine

Abstract

This paper introduces orthogonal bandelet bases to approximate images having some geometrical regularity. These bandelet bases are computed by applying parametrized Alpert transform operators over an orthogonal wavelet basis. These bandeletization operators depend upon a multiscale geometric flow that is adapted to the image at each wavelet scale. This bandelet construction has a hierarchical structure over wavelet coefficients taking advantage of existing regularity among these coefficients. It is proved that C^α-images having singularities along C^α-curves are approximated in a best orthogonal bandelet basis with an optimal asymptotic error decay. Fast algorithms and compression applications are described.

1 Introduction

Wavelet bases are suboptimal to approximate natural images because they cannot take advantage of the geometrical regularity of image structures. Indeed, wavelets have a square support and are thus not adapted to the anisotropic regularity of geometrical elements including edges. Several frames such as the curvelets of Candès and Donoho [4] and the warped bandelets of Le Pennec and Mallat [14] have been introduced to improve the approximation performances of wavelets. The image is decomposed over vectors that are elongated and have vanishing moments in various directions to take advantage of existing image regularity along specific directions. Asymptotic theorems give better approximation error decays in these frames compared to wavelet bases, but curvelets and warped bandelets do not seem to clearly improve the numerical approximation capabilities of wavelets for most natural images.

The human visual system suggests a different hierarchical approach to geometric image representation. Hubel and Wiesel [13] showed in the 1960s that the V1 visual cortex region includes simple cells that have a quasi-linear response relative to the input visual stimuli received on the retina. The response of a simple

Communications on Pure and Applied Mathematics, Vol. LXI, 0001–0040 (PREPRINT)

cell can thus be interpreted as an inner product between the retina image and an "impulse response." The support of these impulse responses, also called *receptive fields*, are well localized in the retina image. Many experiments [7] have measured these impulse responses, which are oscillating functions similar to wavelets. More recent physiological experiments have shown that a geometrical integration appears through horizontal connections between these simple cells [2, 12] within the columnar structure discovered by Hubel and Wiesel. These horizontal connections vary depending upon the image properties. Horizontal connections are also involved in the perception of geometric illusions such as the Kaniza triangle [15].

From a mathematical point of view, the question raised by these recent physiological models is to understand if one can construct hierachical image representations from wavelet coefficients to take advantage of geometric image regularity. Defining a geometry on wavelet coefficients also offers the flexibility to let this geometry depend upon the image scale. This can be important for textures having multiscale structures following different geometries at each scale.

We prove that defining a hierarchical geometric representation from wavelet coefficients has a number of mathematical and algorithmic advantages over direct decompositions such as curvelet and warped bandelet frames. As opposed to these previous constructions, the resulting bandelet bases are orthogonal and inherit the regularity of the wavelets they are constructed from. The geometry can also be adapted at each scale. These bases are derived from a wavelet basis with a cascade of orthogonal operators that define a discrete bandeletization, which leads to a fast algorithm.

The key property that enables us to construct a hierachical representation from wavelet coefficients is given in Section 4, which proves that geometric regularity is preserved over sets of orthogonal wavelet coefficients. It produces a form of redundancy among wavelet coefficients. A bandeletization operator transforms these orthogonal wavelet coefficients to take advantage of their regularity with vanishing moments along appropriate directions. This bandeletization is implemented with an Alpert transform along parametrized multiscale geometric flows. The succession of the orthogonal wavelet transform and adapted bandeletization is equivalent to a decomposition in an adapted orthogonal bandelet basis.

To best approximate an image f from M coefficients, a "best" bandelet basis is constructed by optimizing the geometrical parameters of the bandeletization. For images f that are C^α besides a set of C^α-curves, the main theorem proves that an approximation f_M from M parameters in an optimized bandelet basis satisfies

$$\|f - f_M\|^2_{\mathrm{L}^2} = O(M^{-\alpha}).$$

Similar to the original result proved for warped bandelets [14], this approximation scheme is adaptive with respect to the regularity exponent α, which is a priori unknown. However, it introduces no boundary artifacts, and proofs are simpler because it takes advantage of the approximation properties of the underlying wavelet basis.

For discrete images with N pixels, this bandelet construction also defines orthogonal bases of $\mathbb{R}^N$. Section 7 describes a fast algorithm that decomposes an image in a best bandelet basis with $O(NM^\kappa)$ operations, where M is the number of parameters used for approximation and $\kappa = (\alpha+1)(p-1)^2$. For a compression application where the number M of parameters scales like N^γ, $\gamma < 1$, the overall complexity is $O(N^{1+\gamma\kappa})$.

For compression applications, images are decomposed in a best bandelet basis, and the resulting coefficients are quantized and entropy coded. For images that are discretizations of C^α-functions outside C^α-edges, Section 8 proves that the error introduced by this compression scheme decays like $\log^\alpha(R)R^{-\alpha}$, where R is the number of bits of the compressed code. The resulting distortion-rate curve approaches the Kolmogorov asymptotic lower bound up to a logarithmic factor.

2 Wavelet Approximation of Geometrically Regular Images

Donoho [9] introduced a cartoon image model where $f(x)$ for $x \in [0,1]^2$ is C^α over regions whose boundaries are piecewise C^α-curves. To incorporate the diffraction blur produced by the optics of a camera, this model is refined in [14] with a convolution by an unknown regular kernel. An example is shown in Figure 2.1(a). The following definition formalizes this model:

DEFINITION 2.1 A function $f \in \mathrm{L}^2([0,1]^2)$ is said to be C^α-*geometrically-regular with a scale* $s > 0$ if $f = \tilde{f} * h$ where $\tilde{f} \in \mathrm{C}^\alpha(\Lambda)$ for $\Lambda = [0,1]^2 - \{\gamma_i\}_{1\le i\le G}$. The blurring kernel h is C^α, supported in $[-s,s]^2$ with $\|h\|_{\mathrm{C}^\alpha} \le s^{-(2+\alpha)}$. The edge curves γ_i are C^α and do not intersect tangentially. For $s = 0$, the same definition is valid for $f = \tilde{f}$.

The image $\tilde{f}$ is typically discontinuous along the edge curves γ_i that may correspond to the boundaries of objects in the observed scene. The convolution with the blurring kernel h accounts for the diffraction phenomenon. The scale parameter s of h may be arbitrarily small. For an open set Ω in $\mathbb{R}^2$, $\mathrm{C}^\alpha(\Omega)$ is the space of α-Holderian functions, and $\|f\|_{\mathrm{C}^\alpha(\Omega)}$ is the usual norm on this space.

An isotropic wavelet orthogonal basis of $\mathrm{L}^2([0,1]^2)$ is obtained by translating and dilating three mother wavelets $\{\psi^H, \psi^V, \psi^D\}$ (for the horizontal, vertical, and diagonal directions) [16]. We consider compactly supported wavelets having a support in $[-K,K]^2$. Let p be the number of vanishing moments of these wavelets. Inner products in this basis of $\mathrm{L}^2([0,1]^2)$ are written

(2.1)
$$f_j^k[n] \stackrel{\text{def}}{=} \langle f, \psi_{jn}^k \rangle \quad \text{with} \quad \begin{cases} k \in \{H,V,D\}, \quad j < 0, \\ n = (n_1,n_2) \in \{0,\dots,2^{-j}-1\}^2, \\ \psi_{jn}^k(x) = 2^{-j}\psi^k(2^{-j}x_1 - n_1, 2^{-j}x_2 - n_2), \end{cases}$$

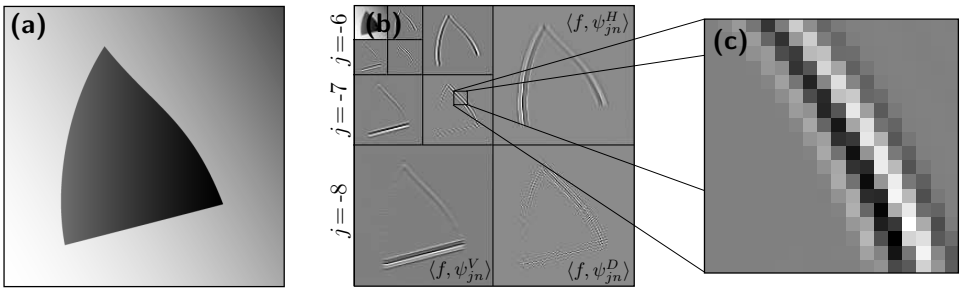

FIGURE 2.1. (a) Example of a C^α-geometrically-regular function f. (b) Orthogonal wavelet coefficients at different scales 2^j. (c) Zoom over wavelet coefficients located in a square including a singularity curve.

with appropriate modifications of ψ^k_{jn} to maintain their support in $[0,1]^2$, as explained by Cohen, Daubechies, and Vial in [6]. In the following, we shall drop the orientation index $k \in \{H, V, D\}$ to simplify notation.

Let M be the number of wavelet coefficients above a threshold T. The M-term wavelet approximation of f is

$$f_M = \sum_{|f_j[n]| \geq T} f_j[n]\, \psi_{jn}.$$

If $f \in C^\alpha([0,1]^2)$ with $\alpha < \beta$, then wavelet coefficients are small at fine scales, and one can prove [18] that the approximation error satisfies

$$\|f - f_M\|^2_{L^2} = O(M^{-\alpha}). \tag{2.2}$$

An image that is C^α-geometrically-regular is usually not uniformly C^α over $[0,1]^2$ because of the discontinuities across edges. For such an image f, the approximation error $\|f - f_M\|_{L^2}$ is dominated by wavelet coefficients $f_j[n]$ corresponding to wavelets ψ_{jn} whose support intersects the singularity curves of f. As a consequence the error $\|f - f_M\|^2_{L^2}$ decays like M^{-1} as opposed to $M^{-\alpha}$. These edge wavelet coefficients thus need to be retransformed in order to reduce this approximation error.

3 Review of Geometric Image Approximations

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 the geometry of their edge curves.

3.1 Curvelet Frame

In order to exploit the geometric image regularity along edge curves, the image is decomposed over functions having vanishing moments along several directions and a support that is elongated. The curvelets of Candès and Donoho [4] are such elongated functions that define a frame. Candès and Donoho proved that for $\alpha = 2$ the best M-term approximation f_M of a C^α-geometrically-regular function f satisfies

$$\|f - f_M\|^2_{L^2} = O(\log^3(M)\, M^{-2}). \tag{3.1}$$

Curvelet approximations are nearly optimal for $\alpha = 2$, but one does not reach the $M^{-\alpha}$ optimal bound for $\alpha > 2$. For bounded variation functions, curvelet approximations do not reach the optimal error decay $\alpha = 1$ obtained by wavelet bases. This is due to the nonadaptivity of curvelet frame geometry.

To define an orthonormal basis of $\mathbb{R}^N$, Do and Vetterli [8] have introduced a modified construction with contourlets implemented with a multiscale and directional filter bank. However, contourlets do not satisfy the asymptotic decay error property (3.1) of curvelets.

3.2 Adaptive Schemes

Instead of decomposing the image in a fixed basis or frame, adaptive schemes adapt the approximation procedure to an estimated geometry calculated from the image.

The wedgelet scheme of Donoho [9] divides the image support into adapted dyadic squares as in Figure 5.2. Over each square the image is approximated with a "wedge" that is constant on each side of a straight line that approximates the image edge in the square. This approach is generalized by Shukla et al. [21] with polynomials separated by polynomial curves. A classification and regression tree (CART) [3] algorithm is used to optimize the dyadic image segmentation.

These approximation schemes can reach the same error decay (3.1) as curvelets only if the edges are discontinuities with no blurring, and the image segmentation in dyadic squares introduces blocking artifacts in the approximation that are discontinuities at the boundaries of the squares.

3.3 Capturing Wavelet Regularity along Edges

Wavelet approximations of geometrically regular functions are inefficient because edges create many large-amplitude wavelet coefficients. To improve wavelet representations, several approaches have been proposed to further transform wavelet coefficients along edges. Wakin et al. [22] and Dragotti and Vetterli [11] perform a vector quantization of wavelet coefficients. Following the work of Matei and Cohen on adaptive essentially non-oscillatory (ENO) lifting [17], new lifting

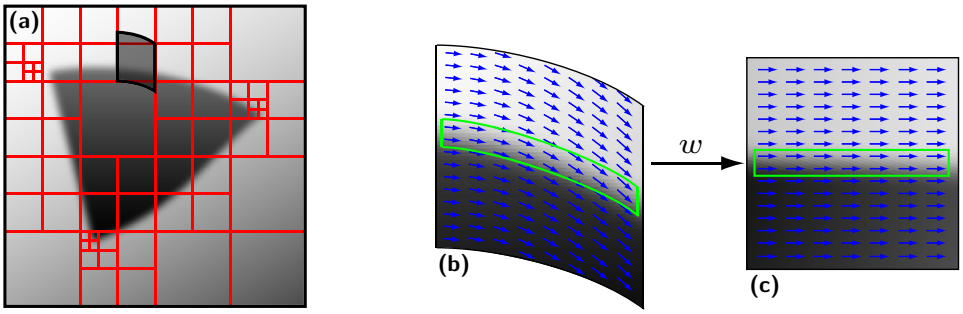

FIGURE 3.1. (a) Image segmentation in dyadic squares, inside each of which there is a single edge. (b) Zoom on a band shown in (a) with its geometric flow nearly parallel to the edge. (c) This band is warped into a square where the flow becomes horizontal. A separable anisotropic wavelet with a rectangular support in this square corresponds to a bandelet parallel to the flow in the original band.

schemes have also been introduced to predict wavelet coefficients from their neighbors [5]. These approaches are mostly algorithmic and do not lead to approximation theory. They require detecting edge curves, which makes it difficult to obtain optimal approximation results when the image is blurred.

3.4 Warped Bandelet Approximation

The warped bandelet transform, introduced by Le Pennec and Mallat [14], also uses dyadic square segmentation to approximate functions. In each square the geometry is defined by finding not an edge location but an orientation along which the image has regular variations. This orientation is defined by a vector field called geometric flow, which is nearly parallel to the edge as shown in Figure 3.1(a). To take advantage of the image regularity along the flow, a larger band parallel to the flow is warped into a rectangle as in Figure 3.1(b). The image in the band is warped into a rectangular image whose flow is either horizontal or vertical. Decomposing this warped image over an anisotropic separable orthogonal wavelet basis is equivalent to decomposing the original image in orthogonal bandelets obtained by warping these wavelets [14].

Figure 3.1 illustrates this process. The union of these bandelet bases over all the bands defines a bandelet frame of $\mathrm{L}^2([0,1]^2)$. Le Pennec and Mallat proved that decomposing a C^α-geometrically-regular image over a best bandelet frame yields a bandelet approximation that satisfies $\|f - f_M\|^2_{\mathrm{L}^2} = O(M^{-\alpha})$ where M is the total number of bandelet coefficients and parameters that specifies the geometric flow and segmentation. Difficulties and boundary issues appear for warped bandelets because adjacent image squares are typically warped with different geometric flows. The resulting bandelets are discontinuous at these boundaries and the orthogonality is lost.

This paper introduces new orthogonal bandelet bases obtained with orthogonal operators applied directly to wavelet coefficients. A multiscale geometric flow is defined over wavelet coefficients, which avoids all boundary issues and maintains the orthogonality of this transform. The next section analyzes the regularity of wavelet coefficients, which is at the core of this construction.

4 Polynomial Approximations of Wavelet Coefficients

To improve wavelet approximations, this section studies the regularity of wavelet coefficients located along edges. This regularity suggests approximating these wavelet coefficients with piecewise polynomials in the wavelet coefficient domain. For C^α-geometrically-regular functions, the resulting approximation scheme yields a nearly optimal approximation error and gives the basic principles of bandelet approximations constructed over wavelet bases.

4.1 Anisotropic Regularity of Wavelets Coefficients

Orthogonal wavelet coefficients of f are calculated through convolutions with scaled wavelets

$$\langle f, \psi_{jn} \rangle = f_j(2^j n)$$

$$\text{where} \quad f_j(x) \stackrel{\text{def}}{=} f * \psi_j(x) \quad \text{and} \quad \psi_j(x) = \frac{1}{2^j}\psi(-2^{-j}x).$$

The convolution guarantees that f_j is at least as regular as ψ_j. The function f_j also inherits the regularity of f. In the following, we suppose that the wavelet ψ is C^α.

If f is C^α-geometrically-regular with a scale s in the sense of Definition 2.1, then it can be written $f = \widetilde{f} * h$, where $\tilde{f}$ is piecewise C^α with singularities along piecewise C^α-curves and where h is a regularization kernel that is supported in $[-s, s]$. As a consequence,

$$f_j = \tilde{f} * h * \psi_j = 2^j \tilde{f} * h_j \tag{4.1}$$

where $h_j = 2^{-j} h * \psi_j$ is a new regularization kernel. Since ψ_j has a support in $[-K2^j, K2^j]^2$, the new kernel h_j has a support in $[-s_j, s_j]^2$ with $s_j = s + K2^j$.

4.2 Local Warping

In the neighborhood of an edge, the values of the wavelet coefficients $f_j(x)$ are regular when moving nearly parallel to the edge. Such displacement directions are specified by a geometric flow that is a vector field. Following the original bandelet approach [14], the geometric regularity of f_j is characterized after a warping that transforms the geometric flow in a horizontal or vertical vector field.

Locally an edge can be parametrized horizontally or vertically. The geometric flow is a vector field that is also parametrized horizontally or vertically and which is constant in the other direction. Figure 4.1(c) shows an example in a square $S \subset [0, 1]^2$ of length $\lambda > 0$, where the edge curve is parametrized horizontally by $x_2 =$

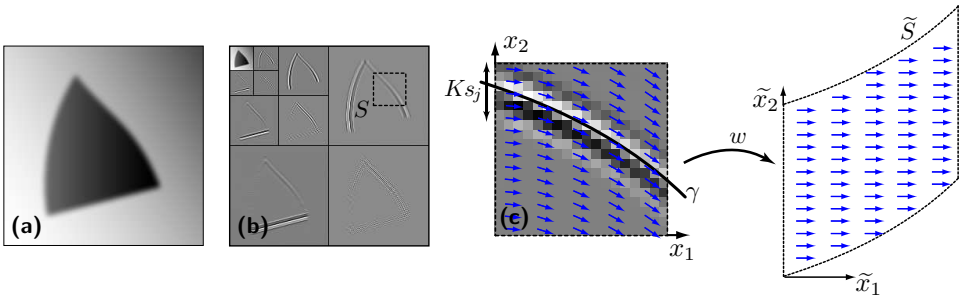

FIGURE 4.1. (a) Original image f. (b) Wavelet coefficients of f. (c) Zoom on wavelet coefficients in a square S including an edge. A geometric flow is a vector field nearly parallel to the edge curve. The warping w aligns the flow horizontally or vertically.

$\gamma(x_1)$. For $(x_1, x_2) \in S$ the coordinates of a flow vector can be written $(1, \tilde{\gamma}'(x_1))$, where $\tilde{\gamma}'$ specifies the flow direction. If the edge is parametrized vertically, then the geometric flow is parallel vertically and hence can be locally written $(\tilde{\gamma}'(x_2), 1)$.

Let us consider the warping operator

$$(\tilde{x}_1, \tilde{x}_2) = w(x_1, x_2) \stackrel{\text{def}}{=} (x_1, x_2 - \tilde{\gamma}(x_1)), \tag{4.2}$$

where $\tilde{\gamma}'$ is the derivative of $\tilde{\gamma}$. The curve $x_2 = \tilde{\gamma}(x_1)$ is an integral curve of the flow. This warping transforms the geometric flow in a horizontal flow as illustrated in Figure 4.1(c). If the flow direction $\tilde{\gamma}'$ is sufficiently close to the edge direction γ', then Proposition 4.1 gives upper bounds on the partial derivatives of the warped wavelet coefficients $f_{j\mathrm{W}}(\tilde{x}) \stackrel{\text{def}}{=} f_j(w^{-1}(\tilde{x}))$. A similar result is proved in [14].

PROPOSITION 4.1 *Let f be a C^α-geometrically-regular function with a scale s. Suppose that f has only one singularity curve parametrized horizontally by $x_2 = \gamma(x_1)$. Let $s_j = s + K2^j$. There exists a constant $C > 0$ such that for any 2^j if, over a square S of length $\lambda \leq s_j^{1/\alpha}$ a polynomial flow direction $\gamma'(x)$ satisfies*

$$\forall (x_1, x_2) \in S = S_1 \times S_2, \quad |\gamma'(x_1) - \tilde{\gamma}'(x_1)| \leq (1 + \|\gamma\|_{\mathrm{C}^\alpha})\lambda^{\alpha-1}, \tag{4.3}$$

then the resulting warped wavelet coefficients $f_{j\mathrm{W}}(\tilde{x}) \stackrel{\text{def}}{=} f_j(w^{-1}(\tilde{x}))$ satisfy

$$\forall i_1 \leq \alpha,\ \forall i_2 \leq \beta,\ \forall \tilde{x} \in w(S), \quad \left| \frac{\partial^{i_1+i_2} f_{j\mathrm{W}}}{\partial x_1^{i_1}\, \partial x_2^{i_2}}(\tilde{x}) \right| \leq C 2^j (1 + \|\gamma\|_{\mathrm{C}^\alpha}^{\alpha}) s_j^{-i_1/\alpha - i_2}. \tag{4.4}$$

PROOF: Condition (4.3) implies that there exists an integral curve $\tilde{\gamma}$ of the flow $\tilde{\gamma}'$ that satisfies $\|\gamma - \tilde{\gamma}\|_{\mathrm{L}^\infty} \leq (1 + \|\gamma\|_{\mathrm{C}^\alpha})\lambda^\alpha$.

The derivatives of $\gamma - \tilde{\gamma}$ are first bounded using condition (4.3). Let γ_1 be a Taylor expansion of degree $\alpha - 1$ of γ inside S_1. It satisfies $\|\gamma^{(i)} - \gamma_1^{(i)}\|_{\mathrm{L}^\infty} \leq \|\gamma\|_{\mathrm{C}^\alpha} \lambda^{\alpha - i}$. The derivatives of $\gamma - \tilde{\gamma}$ can be bounded using

$$\|\gamma^{(i)} - \tilde{\gamma}^{(i)}\|_{\mathrm{L}^\infty} \leq \|\gamma^{(i)} - \gamma_1^{(i)}\|_{\mathrm{L}^\infty} + \|\gamma_1^{(i)} - \tilde{\gamma}^{(i)}\|_{\mathrm{L}^\infty}.$$

The second term is bounded using an expansion in $\{\theta_m\}_{m=0}^{\alpha-1}$, the orthogonal family of Lagrange polynomials on S_1,

$$\gamma_1^{(i)} - \tilde{\gamma}^{(i)} = \sum_{m=0}^{\alpha-1} \langle \gamma_1 - \tilde{\gamma}, \theta_m \rangle \theta_m^{(i)}$$

and thus

$$\|\gamma_1^{(i)} - \tilde{\gamma}^{(i)}\|_{\mathrm{L}^\infty} \leq \alpha \|\gamma_1 - \tilde{\gamma}_1\|_{\mathrm{L}^\infty} \max_m \|\theta_m\|_{\mathrm{L}^1} \|\theta_m^{(i)}\|_{\mathrm{L}^\infty}.$$

There exists a constant C_θ independent of λ such that $\|\theta_m\|_{\mathrm{L}^1} \leq C_\theta \lambda^{1/2}$ and $\|\theta_m^{(i)}\|_{\mathrm{L}^\infty} \leq C_\theta \lambda^{-i-1/2}$. Using the fact that $\|\gamma_1^{(i)} - \tilde{\gamma}^{(i)}\|_{\mathrm{L}^\infty} \leq \|\gamma_1^{(i)} - \gamma^{(i)}\|_{\mathrm{L}^\infty} + \|\gamma^{(i)} - \tilde{\gamma}^{(i)}\|_{\mathrm{L}^\infty} \leq (1 + 2\|\gamma\|_{\mathrm{C}^\alpha}) \lambda^{\alpha - i}$, one has

$$\forall i \leq \alpha, \quad \|\gamma^{(i)} - \tilde{\gamma}^{(i)}\|_{\mathrm{L}^\infty} \leq C_1 \lambda^{\alpha - i} \leq C_1 s_j^{1 - i/\alpha} \tag{4.5}$$

for a constant C_1.

Equation (4.1) shows that $f_j = 2^j \tilde{f} * h_j$ where $h_j = 2^{-j} h * \psi_j$ has a support in $[-s_j, s_j]^2$. The proof of inequality (4.4) is performed by expanding the derivatives of the convolution

$$\begin{aligned} f_{j\mathrm{W}}(x) &= 2^j \int \tilde{f}(x_1 - u_1, x_2 + \tilde{\gamma}(x_1) - u_2) h_j(u) \mathrm{d}u \\ &= 2^j \int \tilde{f}(A(x,u)) h_j(B(x,u)) \mathrm{d}u \end{aligned}$$

where

$$\begin{aligned} A(x,u) &\stackrel{\text{def}}{=} (x - u_1, x_2 + \gamma(x_1 - u_1) - u_2), \\ B(x,u) &\stackrel{\text{def}}{=} (u_1, u_2 + \tilde{\gamma}(x_1) - \gamma(x_1 - u_1)). \end{aligned}$$

Taking derivatives leads to

$$\begin{aligned} &\frac{\partial^{i_1 + i_2} f_{j\mathrm{W}}}{\partial x_1^{i_1} \partial x_2^{i_2}}(x) = \\ &\quad 2^j \int \sum_{d=0}^{i_1} \binom{d}{i_1} \frac{\partial^{i_1 - d}}{\partial {x_1}^{i_1 - d}} [\tilde{f}(A(x,u))] \frac{\partial^d}{\partial {x_1}^d} \left[\left(\frac{\partial^{i_2} h_j}{\partial {x_2}^{i_2}} \right) (B(x,u)) \right] \mathrm{d}u. \end{aligned} \tag{4.6}$$

By hypothesis the function $x_1 \mapsto \tilde{f}(A(x,u))$ is regular, and there exists a constant C such that

$$\left| \frac{\partial^{i_1-d}}{\partial {x_1}^{i_1-d}} [\tilde{f}(A(x,u))] \right| \leq C \|f\|_{C^\alpha(\Lambda)} \max(\|\gamma\|_{C^\alpha}^\alpha, 1).$$

The second term of equation (4.6) is bounded with the Faà di Bruno formula for the derivatives of a composition,

$$\frac{\partial^d}{\partial {x_1}^d} \left[\left(\frac{\partial^{i_2} h_j}{\partial {x_2}^{i_2}} \right) (B(x,u)) \right] = \sum_{(k_s)_s} \frac{d!}{k_1! \cdots k_d!} \frac{\partial^{k+i_2} h_j}{\partial^k x_1 \partial^{i_2} x_2} (B(x,u)) \prod_{s=1}^{d} \left(\frac{\tilde{\gamma}^{(s)}(x_1) - \gamma^{(s)}(x_1 - u_1)}{s!} \right)^{k_s}, \tag{4.7}$$

where the sum is on all d-tuple $(k_s)_s$ such that $\sum s k_s = d$ and where $k \stackrel{\text{def}}{=} \sum k_s$. The deviation on the geometry is bounded using

$$\begin{aligned} &|\tilde{\gamma}^{(s)}(x_1) - \gamma^{(s)}(x_1 - u_1)| \\ &\leq \underbrace{|\tilde{\gamma}^{(s)}(x_1) - \gamma^{(s)}(x_1)|}_{\leq C_1 s_j^{1-s/\alpha}} + \underbrace{|\gamma^{(s)}(x_1) - \gamma^{(s)}(x_1 - u_1)|}_{\leq \|\gamma\|_{C^\alpha} s_j} \qquad (4.8) \\ &\leq C_1 \max(1, \|\gamma\|_{C^\alpha}^\alpha) s_j^{1-s/\alpha}. \end{aligned}$$

Using the fact that

$$\left| \frac{\partial^{k+i_2} h_j}{\partial^k x_1 \partial^{i_2} x_2} (B(x,u)) \right| \leq C_\psi s_j^{-2-i_2-k}, \tag{4.9}$$

where C_ψ is a constant that depends only on ψ, expression (4.7), together with bounds (4.8) and (4.9), leads to

$$\begin{aligned} \left| \frac{\partial^d}{\partial {x_1}^d} \left[\frac{\partial^{i_2} h_j}{\partial {x_2}^{i_2}} (B(x,u)) \right] \right| &\leq C \sum_{(k_s)_s} s_j^{-2-i_2-k} \prod_{s=1}^{d} s_j^{k_s(1-\frac{i}{\alpha})} \\ &\leq C \max(1, \|\gamma\|_{C^\alpha}^\alpha) s_j^{-2-i_2-\frac{\alpha}{d}}. \end{aligned}$$

The term corresponding to $d = i_1$ thus dominates in equation (4.6), and one obtains the bound (4.4) by using the fact that the size of the support of h_j is s_j^2. □

4.3 Polynomial Regression over Band-Shaped Domains

The regularity of warped wavelet coefficients along edges implies that these coefficients can be approximated with piecewise polynomials. This section explains how to construct these polynomials to approximate edge wavelet coefficients at

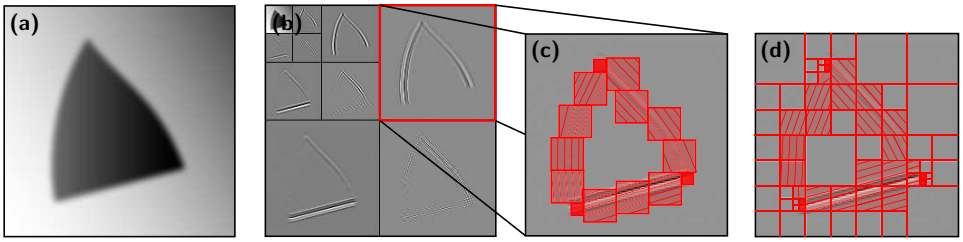

FIGURE 4.2. (a) A geometrically regular image. (b) Wavelet coefficients. (c) A nondyadic segmentation into squares together with bands of size $\lambda \times \mu$ over each square crossing the singularities. (d) A dyadic subdivision of the coefficients together with a dyadic subdivision into bands.

each scale 2^j. We suppose that f is C^α-geometrically-regular, with a regularization scale s. A regularization by a scale $s \neq 0$ is essentially equivalent to translating the wavelet scale 2^j by s and replacing 2^j by $s_j = s + K2^j$. We suppose that $s = 0$ to simplify explanations.

This section gives some insights to understand why an appropriate piecewise polynomial approximation of the wavelet coefficients of f leads to an approximation f_M of f that satisfies

$$\|f - f_M\|_{L^2}^2 \leq C \log(M)^\alpha M^{-\alpha}, \tag{4.10}$$

where M is the total number of parameters needed to specify f_M. Technical details are omitted to carry out the main ideas. A more precise theorem is proved in Section 5.

The set $\mathcal{E}$ of edge wavelet coefficients $f_j[n] = f_j(2^j n) = \langle f, \psi_{jn} \rangle$ corresponds to wavelets ψ_{jn} whose support intersects a single edge of f. Since the support of ψ_{jn} is included in $[2^j n - K2^j, 2^j n + K2^j]$, it corresponds to the set of points $2^j n$ that are at a distance smaller than $K2^j$ from an edge. This set $\mathcal{E}$ of edge wavelet coefficients is first segmented in squares of length λ such that in any such square S, the corresponding edge is parametrized either horizontally or vertically. Figure 4.2(c) shows an example of such a segmentation. In the following, a horizontal parametrization $x_2 = \gamma(x_1)$ is assumed.

The width λ of the squares is chosen in order to match the precision of the approximation. Over each square S, an approximated flow $\tilde{\gamma}'$ is defined as the Taylor expansion of degree $\alpha - 2$ of γ inside S. It satisfies inside S

$$\|\gamma' - \tilde{\gamma}'\|_{L^\infty} \leq \|\gamma\|_{C^\alpha} \lambda^{\alpha - 1}. \tag{4.11}$$

The warping $w(x_1, x_2) = (x_1, x_2 - \tilde{\gamma}(x_1))$ maps the flow $\tilde{\gamma}'$ onto a horizontal flow. As shown on Figure 4.3(d), the warped domain $w(S)$ is subdivided into horizontal bands of length λ and width μ that will be adjusted. This set of warped bands defines a segmentation of the square S into bands that follow the approximated flow $\tilde{\gamma}'$; see Figure 4.3(e).

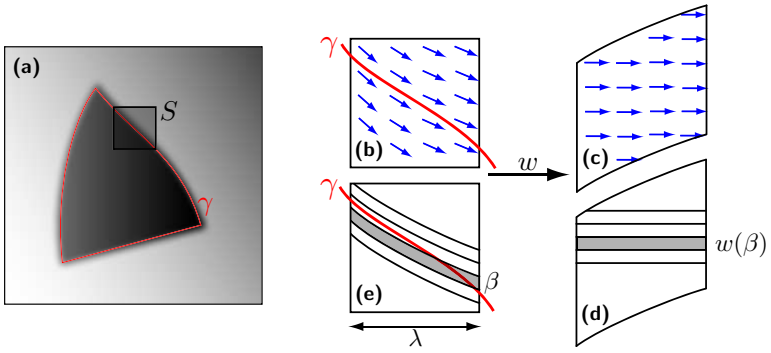

FIGURE 4.3. (a) A geometrically regular image with a square S of width λ on which the geometry γ is parametrized horizontally. (b) Approximated geometric flow $\tilde{\gamma}'$. (c) Horizontal flow over the warped domain $w(S)$. (d) A horizontal band $w(\beta)$ of size $\lambda \times \mu$ in the warped domain. (e) The corresponding band β.

4.4 Polynomial Approximation

The size $\lambda \times \mu$ of the bands is set by analyzing the polynomial approximation error over each band $\beta \subset S$. Wavelet coefficients $f_j[n]$ for $2^j n \in \beta$ are approximated by a polynomial P_β of degree $\beta - 1$ defined over the warped domain $w(\beta)$. Choosing P_β as a Taylor expansion of the warped function $f_{j\mathrm{W}}$ inside $w(\beta)$ leads to an error

$$\begin{aligned} \forall x \in \beta, \quad |f_j(x) - P_\beta(w(x))| &= |f_{j\mathrm{W}}(w(x)) - P_\beta(w(x))| \\ &\leq \sum_{i_1+i_2=\alpha} \left\| \frac{\partial^\alpha f_{j\mathrm{W}}}{\partial x_1^{i_1} \partial x_2^{i_2}} \right\|_\infty \lambda^{i_1} \mu^{i_2}. \end{aligned} \tag{4.12}$$

Condition (4.11) allows us to use Lemma 4.1 to bound the derivatives of $f_{j\mathrm{W}}$ and leads to

$$\begin{aligned} \forall x \in \beta, \quad |f_j(x) - P_\beta(w(x))| &\leq C(1 + \|\gamma\|_{\mathrm{C}^\alpha}^\alpha)\, 2^j \sum_{i_1+i_2=\alpha} s_j^{-i_1/\alpha - i_2} \lambda^{i_1} \mu^{i_2} \\ &\leq C'(2^j s_j^{-1} \lambda^\alpha + 2^j s_j^{-\alpha} \mu^\alpha). \end{aligned}$$

To minimize this approximation error bound, the ratio width : length $= \mu/\lambda$ of the bands is chosen so that $\forall x \in \beta$

$$s_j^{-1} \lambda^\alpha = s_j^{-\alpha} \mu^\alpha \quad \Longrightarrow \quad |f_j(x) - P_\beta(w(x))| \leq 2C' 2^j s_j^{-1} \lambda^\alpha. \tag{4.13}$$

This defines the width μ of the bands as a function of their length λ.

Since the number of bands is

$$M_j = \frac{\text{area of all bands}}{\text{area of a band}} = \frac{LK2^j}{\mu\lambda},$$

where L is the total length of the edge curves of f, the chosen length λ can be expressed as a function of the number of bands M_j,

$$\lambda = \mu s_j^{1/\alpha-1} \quad \Longrightarrow \quad \lambda = (CK)^{-1} s_j^{-1} M_j^{-1}. \tag{4.14}$$

4.5 Approximation Error

Let $\tilde{f}_j[n] = P_\beta(w(2^j n))$ be the approximated wavelet coefficients. This construction is carried over each band β inside $\mathcal{E}$, which defines a discrete piecewise polynomial approximation $\tilde{f}_j$ over the set of coefficients located near edges. Let $M_0 > 0$ be some fixed integer. For locations $(2^j n)$ outside the set $\mathcal{E}$ of edge coefficients, the approximated coefficients are set to $\tilde{f}_j[n] = f_j[n]$ if $f_j[n]$ is one of the M_0 highest coefficients and $\tilde{f}_j[n] = 0$ otherwise.

In the following, $\mathcal{E}$ denotes the set of coefficients located near edge curves and $\mathcal{R} \cup \mathcal{C}$ are the remaining coefficients, where $\mathcal{R}$ stands for the coefficients over regular areas and $\mathcal{C}$ for corner coefficients where two curves are crossing. For each scale $2^j \leq 2^{j_0} \stackrel{\text{def}}{=} M_0^{-\alpha}$, the number of bands is fixed to $M_j = M_0$. By using equations (4.13) and (4.14), the approximation error can be computed for the set of edge coefficients

$$\begin{aligned}
&\sum_{j \geq j_0} \|f_j - \tilde{f}_j\|^2_{\ell^2(\mathcal{E})} \\
&\quad \leq \sum_{j \geq j_0} (\#\text{ coefficients in bands}) \max_{\beta, (2^j n) \in \beta} |f_j[n] - P_\beta(w(2^j n))|^2 \\
&\quad \leq \sum_{j \geq j_0} LK 2^{-j} (2C' 2^j s_j^{-1} \lambda^\alpha)^2 \leq \sum_{j \geq j_0} C_1 s_j^{-1} \lambda^{2\alpha} \\
&\quad \leq \sum_{j \geq j_0} C_1 M_j^{-\alpha} \leq C M_0^{-\alpha} \log(M_0).
\end{aligned} \tag{4.15}$$

The scale $2^{j_0} = M_0^{-\alpha}$ ensures that the error for the remaining scale $2^j < 2^{j_0}$ is bounded by $O(M_0^{-\alpha})$. The error of the wavelet approximation of a regular image with M_0 coefficients decays like $O(M_0^{-\alpha})$ as stated in equation (2.2). One can thus prove that the error for the remaining coefficients in region $\mathcal{R} \cup \mathcal{C}$ satisfies

$$\sum_{j \geq j_0} \|f_j - \tilde{f}_j\|^2_{\ell^2(\mathcal{R} \cup \mathcal{C})} = O(M_0^{-\alpha}). \tag{4.16}$$

Let

$$f_M \stackrel{\text{def}}{=} \sum_{j,n} \tilde{f}_j[n] \psi_{jn}.$$

From equations (4.15) and (4.16) we get that

$$\|f - f_M\|^2_{\mathrm{L}^2} = \sum_j \|f_j - \tilde{f}_j\|_2^2 = O(M_0^{-\alpha} \log(M_0)). \tag{4.17}$$

4.6 Number of Parameters

Let M be the total number of parameters that specifies f_M. One has $M = M_B + M_G$ where M_B is the number of polynomial coefficients needed to specify the approximated coefficients $\tilde{f}_j[n]$, and M_G is the number of geometric coefficients needed to specify the polynomial bands.

For each relevant scale $2^j \geq 2^{j_0}$, the number of polynomial coefficients is

$$\frac{p(p+1)}{2} M_j = \frac{p(p+1)}{2} M_0$$

so that

$$M_B = M_0 + \sum_{j \geq j_0} \frac{p(p+1)}{2} M_j = C \log(M_0) M_0.$$

To specify each band β, one needs to record the coordinates of the segmentation square S containing β and the coefficients of the adapted polynomial flow $\tilde{\gamma}'$. Since $\tilde{\gamma}'$ is a one-dimensional polynomial of degree $p-1$, the number of geometric coefficients at each scale $2^j \geq 2^{j_0}$ is proportional to M_j, and thus M_G is proportional to $M_0 \log(M_0)$.

Summing the number of polynomial and geometric parameters leads to

$$M = M_B + M_G = C \log(M_0) M_0.$$

Combining this result with the error bound (4.17) leads to the global approximation bound (4.10),

$$\| f - f_M \|_{\mathrm{L}^2}^2 \leq C \log(M)^{\alpha+1} M^{-\alpha}.$$

The technical details are skipped since the bandelet basis construction presented in Section 5 gives a constructive proof of these results without the suboptimal $\log(M)^\alpha$ factor.

5 Orthogonal Bandelet Bases

The scheme presented in the previous section does not provide an effective algorithm for computing the adapted geometric flows $\tilde{\gamma}'$ since the location of the edges γ is unknown. It also does not describe a way to estimate the optimal size $\lambda \times \mu$ of the bands since the Holder exponent α is unknown.

We introduce a computational scheme where wavelet coefficients are transformed by an orthogonal bandeletization operator that decomposes these coefficients over a discrete Alpert basis. This Alpert basis depends upon the segmentation and geometrical flows computed at the corresponding scale. This geometry is calculated with a best basis search strategy that optimizes the resulting image approximation. The resulting adaptive bandeletization establishes connections between orthogonal wavelet coefficients with a geometry that is adapted to the image. In that sense, this approximation scheme has similarities with the horizontal connections observed in the V1 visual cortex between simple cells computing wavelet coefficients.

5.1 Alpert Bases

For fast computations, wavelet image coefficients are segmented with dyadic squares. Figure 4.2(d) shows an example of such dyadic subdivisions. Wavelet coefficients located in a square $S \subset [0,1]^2$ of length λ at a given scale 2^j are approximated by decomposing these coefficients in an Alpert basis and thresholding the resulting coefficients. This Alpert basis [1] depends upon the geometric flow calculated in this square. It is defined from a multiresolution constructed over the space $\ell^2(S)$ of wavelet coefficients in S, with piecewise polynomials over bands of dyadic widths that are parallel to the geometric flow. We shall see that thresholding the transformed wavelet coefficients in this Alpert basis is equivalent to automatically adjusting the optimal width $\mu = s_j^{1/\alpha-1}\lambda$ without any knowledge of α.

A geometric flow direction $\tilde{\gamma}'$ is assumed to be known over S. The warping operator w in (4.2) warps S into $\tilde{S}$ as shown in Figure 5.1(a). Points $x_n \overset{\text{def}}{=} 2^j n \in S$ are warped onto $\tilde{x}_n \overset{\text{def}}{=} w(2^j n)$. The space $\ell^2(S)$ is the set of sampled functions $\{g(x_n)\}_{2^j n \in S}$. Similarly, $\ell^2(\tilde{S})$ denotes functions sampled in the warped domain $\{\tilde{g}(\tilde{x}_n)\}_{2^j n \in S}$. A set of wavelet coefficients $\{f_j[n]\}_{2^j n \in S}$ are sample values of $f_j(x)$ at points x_n or of $\tilde{f}_j(\tilde{x})$ at points $\tilde{x}_n$.

To define a multiresolution, for each scale 2^ℓ, the warped square $\tilde{S}$ is recursively subdivided into $2^{-\ell}$ horizontal bands $\tilde{S} = \bigcup_{i=0}^{2^{-\ell}-1} \tilde{\beta}_{\ell,i}$. This process is illustrated in Figure 5.1. The scheme divides $\tilde{\beta}_{\ell,i} = \tilde{\beta}_{\ell-1,2i} \cup \tilde{\beta}_{\ell-1,2i+1}$ by looking for a horizontal cut ensuring that $\tilde{\beta}_{\ell-1,2i}$ and $\tilde{\beta}_{\ell-1,2i+1}$ contain the same number of points. The recursive subdivision is stopped at the scale $\ell = L$ such that $2^L(\lambda 2^{-j})^2 \leq p(p+1)/2$.

Each band $\beta_{\ell,i} \overset{\text{def}}{=} w^{-1}(\tilde{\beta}_{\ell,i})$ in the original square S has a width roughly equal to $\lambda 2^\ell$ and contains $2^\ell(\lambda 2^{-j})^2$ sampling points. Note that some bands near the boundary of S might be disconnected.

Alpert multiresolution spaces $\tilde{V}_\ell \subset \ell^2(\tilde{S})$ are defined for each scale 2^ℓ, $L \leq \ell \leq 0$, by

$$\tilde{V}_\ell \overset{\text{def}}{=} \left\{ \tilde{g} \in \ell^2(\tilde{S}) \;\middle|\; \begin{array}{l} \forall (2^j n) \in \beta_{\ell,i},\ \tilde{g}(\tilde{x}_n) = P_i(\tilde{x}_n) \text{ with } P_i \\ \text{a polynomial and } \deg(P_i) < p. \end{array} \right\}.$$

The space $\tilde{V}_\ell$ is composed of discrete vectors sampled from piecewise polynomials. There is no continuity requirement on these underlying continuous functions so that vectors of $\tilde{V}_\ell$ can exhibit jump discontinuities across bands.

These spaces are embedded since $\tilde{V}_\ell \subset \tilde{V}_{\ell-1}$. An orthogonal basis $\{\tilde{h}_{\ell,i,k}\}_{i,k}$ of each space $\tilde{V}_\ell$ is defined using discrete Legendre polynomials where $k = (k_1, k_2)$ with $k_1 + k_2 < p$ indexes the polynomial degree and $0 \leq i < 2^{-\ell}$ indexes the

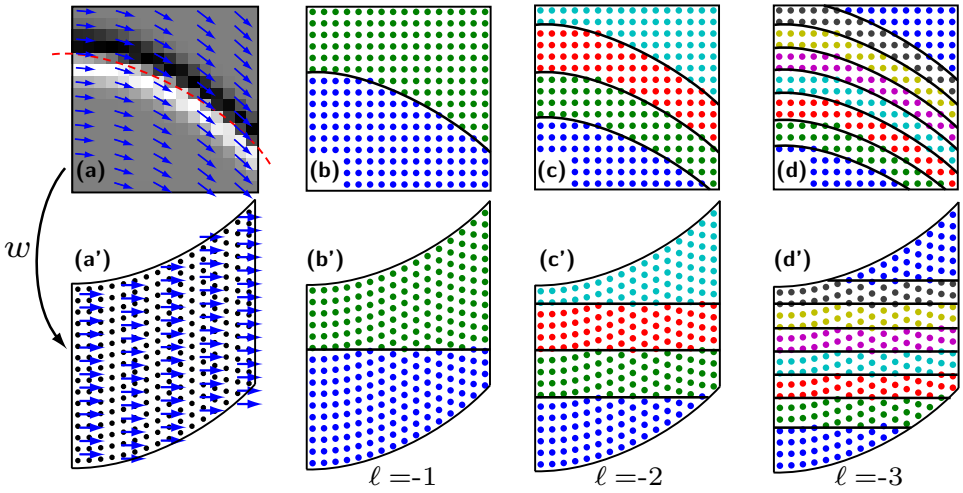

FIGURE 5.1. (a) Wavelet coefficients over a square S near an edge curve. The adapted flow $\tilde{\gamma}'$ is depicted as arrows. (a′) The warping maps the flow onto a horizontal flow. (b)–(d) Refinements of the segmentation in bands. (b′)–(d′) Corresponding horizontal segmentation over the warped domain.

position. Basis vectors $\{\tilde{h}_{\ell,i,k}\}_k$ are obtained by Gram-Schmidt orthogonalization of the set of monomial $\{\tilde{P}_k\}_k$ vectors defined by

$$\forall \tilde{x}_n \in \tilde{\beta}_{\ell,i}, \quad \tilde{P}_k(\tilde{x}_n) = (\tilde{x}^1)^{k_1}(\tilde{x}^2)^{k_2} \quad \text{where } \tilde{x}_n = (\tilde{x}^1, \tilde{x}^2).$$

Alpert wavelets $\{\tilde{\Psi}_{\ell,i,k}\}_{i,k}$ are an orthogonal basis of the orthogonal complement $\tilde{W}_\ell$ of $\tilde{V}_\ell$ in $\tilde{V}_{\ell-1}$ that satisfies $\tilde{V}_{\ell-1} = \tilde{V}_\ell \oplus^\perp \tilde{W}_\ell$. The Alpert wavelet vectors $\{\tilde{\Psi}_{\ell,i,k}\}_k$ are computed by Gram-Schmidt orthogonalization of the family

$$\{\tilde{h}_{\ell-1,2i,k} - \tilde{h}_{\ell-1,2i+1,k}\}_{k_1+k_2<p} \subset \tilde{V}_{\ell-1}$$

against the family $\{\tilde{h}_{\ell,i,k}\}_{k_1+k_2<p} \subset \tilde{V}_\ell$. The numerical computation of the decomposition of a vector on this Alpert basis is carried over by a fast algorithm described in Appendix B. This algorithm involves the orthogonalization over low-dimensional spaces and thus avoids the numerical burden of orthogonalizing directly the vectors $\tilde{h}_{\ell,i,k}$.

The resulting multiwavelet vectors $\tilde{\Psi}_{\ell,i,k}$ are sampled from piecewise polynomial functions that are discontinuous at the middle of the band $\tilde{\beta}_{\ell,i}$. Each vector $\tilde{\Psi}_{\ell,i,k}$ has vanishing moments over the warped domain since it is orthogonal to $\tilde{V}_\ell$,

$$\forall k_1 + k_2 < p, \quad \sum_n \tilde{\Psi}_{\ell,i,k}(\tilde{x}_n)(\tilde{x}_n)^k = 0$$

where $(\tilde{x}_n)^k \stackrel{\text{def}}{=} (\tilde{x}^1)^{k_1}(\tilde{x}^2)^{k_1}$ for each point $\tilde{x}_n = (\tilde{x}^1, \tilde{x}^2)$ in the warped domain. The orthogonal basis $\{\tilde{\Psi}_{\ell,i,k}\}_{\ell,i,k}$ of $\ell^2(\tilde{S})$ defines an orthogonal Alpert basis of $\ell^2(S)$ by

$$\Psi_{\ell,i,k}(x_n) \stackrel{\text{def}}{=} \tilde{\Psi}_{\ell,i,k}(\tilde{x}_n).$$

Note that this definition over the original domain S does not involve any interpolation.

In the following, $m = (i,k)$ indexes the $p(p+1)2^{\ell-1}$ Alpert wavelets $\{\Psi_{\ell,m}\}$ at a scale 2^ℓ, and we consider m as an integer. The orthogonal Alpert basis $\mathcal{B}(S,\tilde{\gamma}')$ of $\ell^2(S)$ is defined by

$$\mathcal{B}(S,\tilde{\gamma}') \stackrel{\text{def}}{=} \{\Psi_{\ell,m} \mid L \le \ell \le 0 \text{ and } 0 \le m < p(p+1)2^{\ell-1}\}.$$

In the following we write $\Psi_{\ell,m}[n] = \Psi_{\ell,m}(x_n)$, the coordinates of the discrete Alpert vector.

The following proposition gives the normalization of the Alpert basis vectors, which is used to find upper bounds for Alpert coefficients.

PROPOSITION 5.1 *There exists a constant C_Ψ such that for any flow $\tilde{\gamma}'$ defined over S, the Alpert basis $\mathcal{B}(S,\tilde{\gamma}') = \{\Psi_{\ell,m}\}_{\ell,m}$ satisfies*

$$\begin{aligned} \|\Psi_{\ell,m}\|_2 = 1, \quad \|\Psi_{\ell,m}\|_\infty \le C_\Psi \lambda^{-1} 2^{-\frac{\ell}{2}} 2^j, \\ \|\Psi_{\ell,m}\|_1 \le C_\Psi \lambda 2^{\frac{\ell}{2}} 2^{-j}. \end{aligned} \tag{5.1}$$

PROOF: The support of an Alpert vector $\Psi_{\ell,m}$ is $\beta_{\ell,m}$, and its cardinal is

$$\operatorname{Card}(\beta_{\ell,m}) = \frac{\{\#\text{ points in } S\}}{\{\#\text{ bands}\}} = \lambda^2 2^\ell 2^{-2j}.$$

Let

$$I_{\ell,m} = \left\{ n \;\middle|\; |\Psi_{\ell,m}[n]| \ge \frac{1}{2}\|\Psi_{\ell,m}\|_\infty \right\}. \tag{5.2}$$

One has the following bound:

$$1 = \|\Psi_{\ell,m}\|_2 \ge \sum_{n \in I_{\ell,m}} |\Psi_{\ell,m}[n]|^2 \ge \frac{1}{2} \operatorname{Card}(I_{\ell,m}) \|\Psi_{\ell,m}\|_\infty^2. \tag{5.3}$$

Asymptotically, the coefficients $\Psi_{\ell,m}[n]$ are samples from one of $p(p+1)/2$ piecewise polynomials scaled by a factor 2^ℓ. One thus has

$$\forall \ell, m, \quad \operatorname{Card}(I_{\ell,m}) \ge C_\Psi \operatorname{Card}(\beta_{\ell m}),$$

where C_b is a constant that does not depend on ℓ or m. As a result,

$$\|\Psi_{\ell,m}\|_\infty \le \sqrt{2} C_\Psi \operatorname{Card}(\beta_{\ell,m})^{-\frac{1}{2}}, \tag{5.4}$$

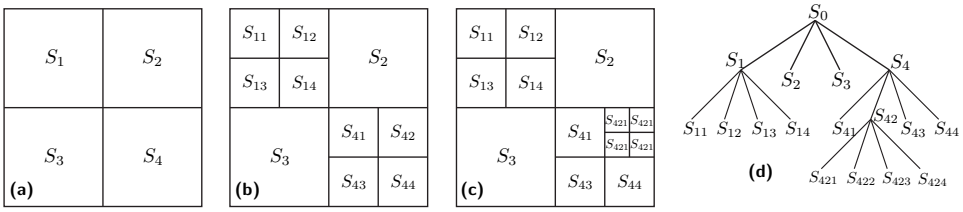

FIGURE 5.2. (a)–(c) Construction of a dyadic segmentation by successive subdivisions. (d) Quadtree representation of the segmentation. Each leaf of the tree indicates a square in the segmentation and can be represented with a binary string whose length is proportional to the depth of the tree.

which implies (5.1) by inserting (5.4) and (5.2) in (5.3). The bound on the ℓ^1-norm is obtained using

$$\|\Psi_{\ell,m}\|_1 \leq \mathrm{Card}(\beta_{\ell,m})\|\Psi_{\ell,m}\|_\infty.$$

□

5.2 Bandeletization of Wavelet Coefficients

Segmentation of Wavelet Coefficients

For each scale 2^j, a segmentation $\mathcal{S}_j = \{S\}_{S\in\mathcal{S}_j}$ subdivides $[0,1]^2$ into non-overlapping squares S of width bigger than 2^j. Following the ideas of Section 4.3, the length λ of the squares must be adapted to the approximation precision.

In order to do so, the segmentation $\mathcal{S}_j$ is enforced to contain only squares of dyadic lengths 2^k for $k \leq 0$. A segmentation of $[0,1]^2$ using dyadic squares is obtained by a recursive subdivision of the original square into four squares of equal sizes. On Figure 5.2 one can see the subdivision steps leading to the construction of a dyadic subdivision $\mathcal{S}_j$, together with the tree representing the subdivision process. Each square $S \in \mathcal{S}_j$ of size $2^k \times 2^k$ corresponds to a leaf of the tree at depth k.

For a geometrically regular image f, an adapted segmentation $\mathcal{S}_j$ should encapsulate the singularity curves in a set of squares whose size λ matches the approximation precision. Junctions between singularity curves should be covered by small squares as in Section 4.3, and the remaining domain should be covered by the largest possible squares.

The squares of the dyadic segmentation $\mathcal{S}_j$ are partitioned into several groups:

- The set $\mathcal{E}(\mathcal{S}_j) = \mathcal{E}^H(\mathcal{S}_j) \cup \mathcal{E}^V(\mathcal{S}_j)$ of edge squares. By definition, a horizontal (respectively, vertical) edge square $S \in \mathcal{E}^H$ (respectively, $S \in \mathcal{E}^V$) is a square smaller than $\lambda(T)$ that is at a distance less than $s_j \stackrel{\text{def}}{=} s + K2^j$ from one and only one edge curve. This curve is supposed to be parametrized horizontally (respectively, vertically) by $x_2 = \gamma(x_1)$

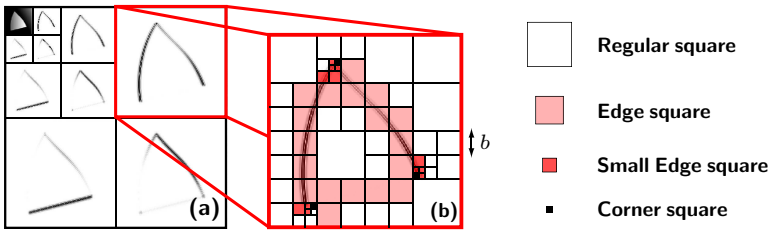

FIGURE 5.3. (a) Wavelet transform of a geometrically regular function. (b) Example of a segmentation $\mathcal{S}_j$ adapted to the geometry of the function.

(respectively, $x_1 = \gamma(x_2)$) with

$$|\gamma'| \leq 2. \tag{5.5}$$

Over such a square, the adapted flow $\tilde{\gamma}'$ should be designed to approximate γ'.

- The set $\mathcal{C}(\mathcal{S}_j)$ of corner squares. By definition, a corner square is a square that contains the junction of two curves.
- The set $\mathcal{R}(\mathcal{S}_j)$ of regular squares. These are the squares that do not contain any edge curves.

Figure 5.3 shows an example of such an adapted dyadic segmentation.

Bandelet Basis

Inside each square $S \in \mathcal{S}_j$, the choice of discrete Alpert basis $\mathcal{B}(S, \tilde{\gamma}'_S)$ constructed in Section 5.1 depends on the choice of a geometric flow $\tilde{\gamma}'_S$.

Over regular square $S \in \mathcal{R}(\mathcal{S}_j)$, a basis adapted to the geometrically image f should not transform the wavelet coefficients. For such a square S, the flow $\tilde{\gamma}'_S$ is undefined and the projection onto $\mathcal{B}(S, \tilde{\gamma}'_S)$ leaves the wavelet coefficients in S unchanged.

A dyadic segmentation together with the adapted flows $\Gamma_j = (\mathcal{S}_j, \{\tilde{\gamma}'_S\}_{S \in \mathcal{S}_j})$ specifies a bandeletization basis $\mathcal{B}(\Gamma_j)$ of the whole space of wavelet coefficients at a scale 2^j,

$$\mathcal{B}(\Gamma_j) \stackrel{\text{def}}{=} \bigcup_{S \in \mathcal{S}_j} \mathcal{B}(S, \tilde{\gamma}'_S).$$

A discrete Alpert vector $\Psi_\nu \in \mathcal{B}(\Gamma_j)$ is thus specified by $\nu = (j, S, \tilde{\gamma}'_S, \ell, m)$ where

- 2^j is a scale of the two-dimensional wavelet transform,
- $S \in \mathcal{S}_j$ is a square of width $\lambda = 2^{-L/2}\, 2^j$,

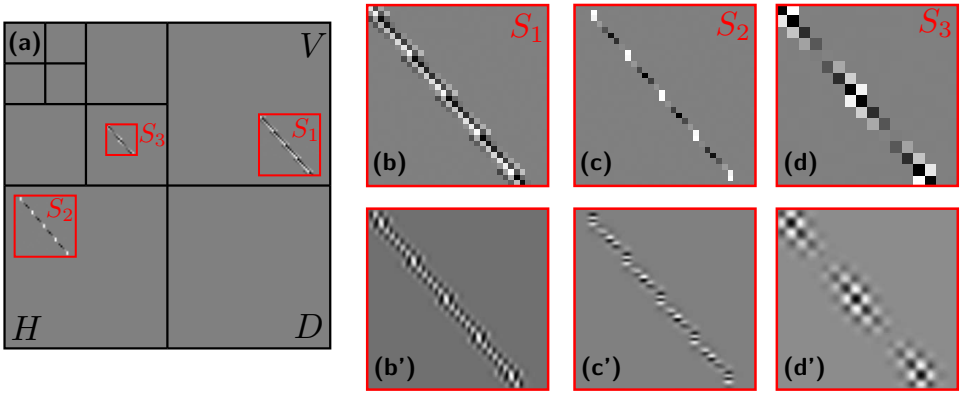

FIGURE 5.4. (a) Localization on the wavelet domain of the squares S_i on which each Alpert wavelet vector is defined. (b)–(d) Discrete Alpert vectors $\Psi_{\ell i}$ for various scales 2^ℓ. (b′)–(d′) Corresponding bandelet functions $b_{\ell i}$.

- $\tilde{\gamma}'_S$ is a geometric flow,
- $\ell \in \{L, \dots, 0\}$ and $m \in \{0, \dots, p(p+1)2^{-\ell-1} - 1\}$ are the scale and index of a discrete Alpert vector $\Psi_{\ell,m} \in \mathcal{B}(S, \tilde{\gamma}'_S)$ and

$$\forall (2^j n) \in S \in \mathcal{S}_j, \quad \Psi_v[n] = \Psi_{\ell,m}[n] \quad \text{where } \mathcal{B}(S, \gamma'_S) = \{\Psi_{\ell,m}\}_{\ell,m}.$$

The coefficients $\Psi_v[n]$ are the coordinates of a bandelet function $b_v \in \mathrm{L}^2([0,1]^2)$ in the wavelet basis. This function is defined by

$$b_v(x) = \sum_n \Psi_v[n] \psi_{jn}(x). \tag{5.6}$$

This function is called a bandelet. Indeed, it is a combination of wavelets along a band and its support is thus also along a band as illustrated in Figure 5.4. Bandelets are obtained from an orthogonal wavelet basis with an orthogonal transformation that we call a "bandeletization." If we apply this transformation to each scale 2^j,

$$\mathcal{B}(\Gamma) \stackrel{\text{def}}{=} \bigcup_{j \le 0} \{b_v \mid \Psi_v \in \mathcal{B}(\Gamma_j)\} \quad \text{where } \Gamma \stackrel{\text{def}}{=} \bigcup_{j \le 0} \Gamma_j$$

is an orthogonal basis of $\mathrm{L}^2([0,1]^2)$.

Bandelets are as regular as the underlying wavelets because they are obtained in (5.6) as a finite linear combination of wavelets of the same scale 2^j. If the wavelets have a compact support, then the resulting bandelets have a compact support. The support of bandelets overlap in the same way that the support of wavelets overlap. Setting to zero bandelet coefficients does not create any blocking artifact because bandelets are regular. This is particularly important for reconstructing image approximations with no artifacts.

6 Best Bandelet Basis Approximation

The set of bandelet bases defines a dictionary of orthogonal bases

$$\mathcal{D} = \{\mathcal{B}(\Gamma)\}_{\Gamma \in G}.$$

In order to approximate a function f with M parameters, one would like to find a bandelet basis $\mathcal{B}(\Gamma^\star) = \{b_\nu\}_\nu$ adapted to f. This adapted basis should be chosen in order to minimize the approximation error $\|f - f_M\|$ of f in $\mathcal{B}(\Gamma^\star)$ defined by

$$f_M \stackrel{\text{def}}{=} \sum_{|\langle f, b_\nu\rangle| > T} \langle f, b_\nu\rangle b_\nu \tag{6.1}$$

where the number of coefficients is $M = M_B + M_{\mathcal{D}}$, where

$$M_B \stackrel{\text{def}}{=} \text{Card}\{\nu \mid |\langle b_\nu, f\rangle| \geq T\} \tag{6.2}$$

is the number of bandelet coefficients above the threshold and $M_{\mathcal{D}}$ is the number of coefficients needed to specify $\Gamma^\star$ in G. Our goal is to choose $\mathcal{B}(\Gamma^\star)$ such that

$$\|f - f_M\|_{\mathrm{L}^2} = O(M^{-\alpha}).$$

6.1 Bandelet Basis Dictionary

The number of bandelet bases in the dictionary is reduced to a finite size by parametrizing the flows $\tilde{\gamma}'_S$ up to a precision T^2 and by limiting the bandeletization to the scales $2^j > T^2$.

Parameterizing the Geometry

Inside a square S of length λ, a geometric flow direction $\tilde{\gamma}'$ is parametrized by a polynomial of degree $p-2$ whose coefficients are quantized at a precision τ:

$$\tilde{\gamma}'(x) = \sum_{i=0}^{p-2} \frac{a_i \tau}{\lambda^{i+1}} x^i \quad \text{with } a_i \in \mathbb{Z}. \tag{6.3}$$

In the following, the precision is set to

$$\tau = \tau(\lambda) \stackrel{\text{def}}{=} \frac{\lambda^p}{p-1}, \tag{6.4}$$

which ensures that a flow γ' can be approximated by a flow $\tilde{\gamma}'$ quantized as in equation (6.3) with $\|\gamma' - \tilde{\gamma}'\|_{\mathrm{L}^\infty} = O(\lambda^{\alpha-1})$.

The adapted geometric flow $\tilde{\gamma}'$ is thus chosen in the finite set

$$\mathcal{G}(S) \stackrel{\text{def}}{=} \left\{ \tilde{\gamma}'(x) = \sum_{i=0}^{p-2} \frac{a_i \tau(\lambda)}{\lambda^{i+1}} x^i \;\middle|\; a_i \in \mathbb{Z} \text{ and } |a_i| \leq \frac{2\overline{C}\lambda}{\tau(\lambda)} \right\}, \tag{6.5}$$

where the constant $\overline{C}$ is set so that the following bound on the polynomial expansion holds:

$$\forall x \in [0,1],\ |a_0 + a_1 x + \cdots + a_{\alpha-1} x^{\alpha-1}| \leq 1 \quad \Longrightarrow \quad \forall i,\ |a_i| \leq \overline{C}. \tag{6.6}$$

Construction of the Dictionary

The dictionary $\mathcal{D}_j$ of bandeletization bases at a given scale 2^j is composed of the bases $\mathcal{B}(\Gamma_j)$ for all possible dyadic segmentations $\mathcal{S}_j$ and geometric flow directions $\tilde{\gamma}'_S \in \mathcal{G}(S)$ inside the squares S of the segmentation:

$$\mathcal{D}_j \stackrel{\text{def}}{=} \{\mathcal{B}(\Gamma_j) \mid \Gamma_j = (\mathcal{S}_j, \{\tilde{\gamma}'_S\}) \text{ and } \forall S \in \mathcal{S}_j,\ \tilde{\gamma}'_S \in \mathcal{G}(S)\}.$$

Let $T > 0$ be the nonlinear approximation threshold. A finite bandelet dictionary $\mathcal{D}_{T^2}$ of bases of $\mathrm{L}^2([0,1]^2)$ is constructed by using bandeletization bases for the first wavelet scales $2^j \geq T^2$ and using the wavelet basis functions for the remaining scales:

$$\mathcal{D}_{T^2} \stackrel{\text{def}}{=} \{\mathcal{B}(\Gamma) \mid \forall 2^j \geq T^2,\ \mathcal{B}(\Gamma_j) \in \mathcal{D}_j\}.$$

6.2 Best Basis Approximation

A Lagrangian minimization computes a bandelet basis $\mathcal{B}(\Gamma^\star)$ whose segmentation $\{\mathcal{S}_j\}_j$ and geometric flows $\{\tilde{\gamma}'_S\}_{S \in \mathcal{S}_j}$ are adapted to f.

Number of Coefficients

Let $T > 0$ be some approximation threshold and

$$\mathcal{B}(\Gamma) = \{b_\nu\}_\nu \in \mathcal{D}_{T^2} \quad \text{where } \Gamma = \bigcup_j \Gamma_j \tag{6.7}$$

be a bandelet basis, where $\Gamma_j = (\mathcal{S}_j, \{\tilde{\gamma}'_S\}_S)$ are the parameters that describe the basis at each scale. The thresholded approximation at T in this basis is defined by equation (6.1). The number of parameters M needed to describe f_M is decomposed as

$$M = M_B + M_{\mathcal{D}} \quad \text{where } M_{\mathcal{D}} = M_G + M_S = \sum_{j,k} M_{Gj} + M_{Sj}. \tag{6.8}$$

The number of bandelet coefficients M_B is determined by equation (6.2) and for each scale 2^j:

- M_{Sj} is the number of parameters needed to specify the segmentation $\mathcal{S}_j$. The dyadic segmentation is described using a quadtree structure, as shown in Figure 5.2. The quadtree is coded using one coefficient per node to specify whether it is an interior node or if it corresponds to a square S that is either a horizontal edge $S \in \mathcal{E}^H(\mathcal{S}_j)$, a vertical edge $S \in \mathcal{E}^V(\mathcal{S}_j)$, or regular square $S \in \mathcal{R}(\mathcal{S}_j)$. Corner squares $S \in \mathcal{C}(\mathcal{S}_j)$ are treated as regular squares since no bandeletization is performed on the wavelet coefficients of these squares. One has

$$M_{Sj} \leq \mathrm{Card}(\mathcal{S}_j). \tag{6.9}$$

- M_{Gj} is the number of parameters needed to specify the geometric flows $\tilde{\gamma}'_S$ over all the dyadic squares $S \in \mathcal{S}_j$. The adapted flow $\tilde{\gamma}'_S$ is parametrized with $p-1$ polynomial coefficients. The number of geometric coefficients is thus

$$M_{Gj} \leq (p-1)\,\mathrm{Card}(\mathcal{S}_j). \tag{6.10}$$

Lagrangian Minimization

A best basis is computed by minimizing the Lagrangian

$$\mathcal{L}(f, \mathcal{B}(\Gamma), T) \stackrel{\text{def}}{=} \sum_{|\langle b_\nu, f\rangle| < T} |\langle b_\nu, f\rangle|^2 + T^2 M, \tag{6.11}$$

as introduced by Donoho in [9]. A similar Lagrangian is optimized by Le Pennec and Mallat in [14]. The best bandelet basis $\mathcal{B}(\Gamma^\star)$ adapted to f is defined by

$$\mathcal{B}(\Gamma^\star) \stackrel{\text{def}}{=} \operatorname*{argmin}_{\mathcal{B}(\Gamma)\in\mathcal{D}_{T^2}} \mathcal{L}(f, \mathcal{B}(\Gamma), T). \tag{6.12}$$

This best bandelet basis can be computed with a fast algorithm described in Section 7. The following theorem gives the approximation rate of a geometrically regular function f in this adapted bandelet basis $\mathcal{B}(\Gamma^\star)$.

THEOREM 6.1 *Let f be a C^α-geometrically-regular function. There exists C such that for any $T > 0$ the M-parameter approximation f_M in the best bandelet basis $\mathcal{B}(\Gamma^\star)$ satisfies*

$$\|f - f_M\|^2_{\mathrm{L}^2} \leq C M^{-\alpha}. \tag{6.13}$$

This theorem states that the approximation of geometrically regular functions in a best bandelet basis recovers the asymptotic $M^{-\alpha}$ decay of the approximation of uniformly regular functions in a wavelet basis. This asymptotic decay rate is thus optimal.

PROOF: The proof relies on Lemma 6.2, which constructs an optimized bandelet basis within the dictionary that yields an approximation error from M parameters that decays in $O(M^{-\alpha})$.

LEMMA 6.2 *Let f be a C^α-geometrically-regular function of Definition* 2.1. *There exists C such that for all $T > 0$ there exists a bandelet basis $\mathcal{B}(\Gamma) \in \mathcal{D}_{T^2}$ in which the thresholded approximation f_M of f at T in this basis satisfies*

$$\|f - f_M\|^2_{\mathrm{L}^2} \leq C\, T^{\frac{2\alpha}{\alpha+1}} \quad \text{and} \quad M \leq C\, T^{-\frac{2}{\alpha+1}}. \tag{6.14}$$

The proof of Lemma 6.2 is in Appendix A. Theorem 6.1 is derived by showing that the best basis that minimizes the Lagrangian is nearly as good as the optimized bandelet basis provided by Lemma 6.2.

Lemma 6.2 provides an adapted bandelet basis $\mathcal{B}(\Gamma) \in \mathcal{D}_{T^2}$ such that

$$\mathcal{L}(f, \mathcal{B}(\Gamma^\star), T) \leq \mathcal{L}(f, \mathcal{B}(\Gamma), T) \leq C\, T^{\frac{2\alpha}{\alpha+1}},$$

so

$$\|f - f_M\|_{\mathrm{L}^2}^2 \leq \mathcal{L}(f, \mathcal{B}(\Gamma^\star), T) \leq C\, T^{\frac{2\alpha}{\alpha+1}} \tag{6.15}$$

and

$$M T^2 \leq \mathcal{L}(f, \mathcal{B}(\Gamma^\star), T) \leq C\, T^{\frac{2\alpha}{\alpha+1}}. \tag{6.16}$$

Combining these equations proves equation (6.13) in Theorem 6.1. □

7 Fast Bandelet Approximation

This section describes the fast transform of a discretized image in a best bandelet basis. A Matlab implementation of this transform is available [19].

A discretized image $\overline{f}$ of $N \times N$ pixels is obtained by projecting a function $f \in \mathrm{L}^2([0,1]^2)$ onto a set of orthogonal scaling functions $\{\phi_{Jn}\}_n$ at a resolution $2^J = N^{-1}$

$$\forall n \in \{0, \ldots, N-1\}^2, \quad \overline{f}[n] \stackrel{\text{def}}{=} \langle f, \phi_{Jn} \rangle$$
$$\text{where } \phi_{Jn}(x) \stackrel{\text{def}}{=} 2^{-J} \phi(2^{-J} x - n).$$

In the following, ϕ is assumed to be the scaling function associated with the two-dimensional wavelet functions used for the construction of the bandelet bases. This allows us to consider that the discrete coefficients $f_j[n]$ computed with the fast orthogonal wavelet transform are inner products $\langle f, \psi_{jn} \rangle$ of the underlying continuous function with the wavelet basis defined in (2.1). This hypothesis simplifies the explanations since the continuous approximation results of Section 6.2 carry over without modification in the discrete setting. These theoretical results are still valid for an arbitrary scaling function ϕ as long as it is C^α regular and has a compact support, using arguments similar to those of [14].

The forward fast discrete bandelet transform decomposes the discrete image $\overline{f}$ in a best bandelet basis of $\mathbb{R}^{N \times N}$, which is equivalent to the decomposition of the underlying function f on a best bandelet basis of $\mathrm{L}^2([0,1]^2)$.

(1) *Two-dimensional wavelet transform.* The wavelet coefficients $\{f_j[n]\}_{j,n}$ are computed using a discrete wavelet transform of the image $\overline{f}$. The complexity of the fast wavelet transform is $O(N^2)$ [16].

(2) *Fast Alpert transform.* For each scale $2^j \geq T^2$, for each dyadic square S of length λ larger than 2^j, and for each geometric flow $\tilde{\gamma}' \in \mathcal{G}(S)$, let $\mathcal{B}(S, \tilde{\gamma}') = \{\Psi_{\ell,m}\}_{\ell,m}$. The Alpert coefficients $\langle f_j, \Psi_{\ell,m} \rangle$ of the wavelet coefficients f_j of f inside S are computed using the fast Alpert transform described in Appendix B. The complexity of the Alpert transform for each square is $O(m^2)$ where $m^2 = (2^{-j}\lambda)^2$ is the number of coefficients in the square S.

(3) *Lagrangian Minimization* For each scale 2^j, the dyadic segmentation $\mathcal{S}_j$ is computed using a fast bottom-up algorithm similar to the CART regression procedure [3]. This involves computing the value of the Lagrangian $\mathcal{L}(f_j, \mathcal{B}(S, \tilde{\gamma}'), T)$ for all the squares S and geometries $\tilde{\gamma}' \in \mathcal{G}(S)$. This set of Lagrangians is stored in a full quadtree that is pruned by the regression tree algorithm as detailed in [14]. The resulting dyadic segmentation $\mathcal{S}_j$ corresponding to the tree together with the set of optimized geometries $\tilde{\gamma}'_S$ for each $S \in \mathcal{S}_j^\star$ are the best basis parameter for the scale 2^j.

Numerical Complexity

The Lagrangian minimization requires the computation of the Alpert transform over each square S of dyadic length for each geometry of $\mathcal{G}(S)$. The bottom-up procedure that builds the quadtree has a negligible complexity. For each scale 2^j, the complexity of computing the Alpert transform over the set of all squares of length λ and all geometries is

$$\underbrace{(1/\lambda)^2}_{\text{\# squares}} \times \underbrace{C_{\mathcal{A}}(2^{-j}\lambda)^2}_{\substack{\text{complexity of} \\ \text{Alpert transform}}} \times \underbrace{C_{\mathcal{G}}\lambda^{-(p-1)^2}}_{\text{\# geometries}} = C\,2^{-2j}\,T^{-2(p-1)^2},$$

since equation (6.5) shows that the cardinal of $\mathcal{G}(S)$ is proportional to $\lambda^{-(p-1)^2}$ and that $\lambda \geq 2^j \geq T^2$. As the number of such widths λ and scales 2^j is proportional to $|\log_2(T)|$, the overall complexity of the bandelet transform is $O(N^2\,T^{-2(p-1)^2})$. Since the number of parameters M used for the approximation in the best bandelet basis scales like $T^{-(\alpha+1)/2}$, the complexity of the algorithm is $O(NM^\kappa)$ with $\kappa = (\alpha+1)(p-1)^2$.

The complexity of the algorithm scales linearly in the number of pixels as for the classical wavelet transform. It requires testing an exhaustive set of local geometries with a precision related to M, which makes the algorithm slower than an orthogonal wavelet transform. In [20] the authors describe an application to surface compression and propose the use of bandelets with one vanishing moment, thus replacing the Alpert transform by an orthogonal Haar transform. The resulting compression algorithm competes favorably with the state of the art, and the use of one vanishing moment is enough for this kind of geometrical data. The overall complexity of the resulting scheme is roughly equal to a dozen orthogonal wavelet transforms for a typical compression scenario.

8 Best Bandelet Basis Compression

An image f is compressed in a bandelet basis $\mathcal{B}(\Gamma) = \{b_\nu\}_\nu \in \mathcal{D}_{T^2}$ by quantizing and coding its transformed coefficients and by coding the geometric parameters $\Gamma = \{\Gamma_j\}_j$ that describe the basis, where $\Gamma_j = (\mathcal{S}_j, \{\tilde{\gamma}'_S\}_{S\in\mathcal{S}_j})$. The restored

image from the compressed code is

$$f_R \stackrel{\text{def}}{=} \sum_{\nu} Q_T(\langle f, b_\nu \rangle)\, b_\nu, \tag{8.1}$$

where Q_T is a uniform quantizer defined by

$$Q_T(x) = qT \quad \text{if } (q-\tfrac{1}{2})T \leq x \leq (q+\tfrac{1}{2})T. \tag{8.2}$$

The coding distortion is

$$D(R) \stackrel{\text{def}}{=} \|f - f_R\|^2_{\mathrm{L}^2}.$$

Since $|x - Q_T(x)| \leq T/2$ and $Q_T(x) = 0$ if $|x| \leq T/2$, one has

$$D(R) = \|f - f_R\|^2_{\mathrm{L}^2} = \sum_{\nu} |\langle f, b_\nu \rangle - Q_T(\langle f, b_\nu \rangle)|^2 \tag{8.3}$$

$$\leq \sum_{|\langle f, b_\nu \rangle| < T/2} |\langle f, b_\nu \rangle|^2 + \frac{1}{4} M_B T^2 \tag{8.4}$$

$$\leq \|f - f_M\|^2_{\mathrm{L}^2} + \frac{1}{4}\, M_B T^2, \tag{8.5}$$

which links the distortion $D(R)$ with the nonlinear approximation f_M obtained with a thresholding at $T/2$ as defined in (6.1). The number of coefficients $M = M_B + M_S + M_G$ is computed following Section 6.2.

The bit budget of this transformed code is

$$R \stackrel{\text{def}}{=} R_B + R_S + R_G = \sum_j (R_{Bj} + R_{Sj} + R_{Gj})$$

where:

- R_{Bj} is the number of bits that is needed to code the bandelet coefficients $\langle f, b_\nu \rangle = \langle f_j, \Psi_\nu \rangle$ for a single scale 2^j. Since there are 2^{-2j} bandelet coefficients at a scale 2^j, the index of each of the M_{Bj} nonzero quantized coefficients is coded using $\log_2(2^{-2j})$ bits per coefficient. For a bounded image f, one has

  $$|\langle f, b_\nu \rangle| \leq \|f_j\|_\infty \|\Psi_\nu\|_1 \leq 2^j \|f\|_{\mathrm{L}^\infty} \|\psi\|_{\mathrm{L}^1} C_b \leq \|f\|_{\mathrm{L}^\infty} \|\psi\|_{\mathrm{L}^1} C_b.$$

  Consequently, there exists a constant C_1 such that the quantized amplitudes $Q_T(\langle f, b_\nu \rangle)$ are coded using less than $\log_2(C_1/T)$ bits per coefficient. The number of bits to code the bandelet coefficients is thus bounded by

  $$R_{Bj} \leq M_{Bj}(\log_2(2^{-2j}) + \log_2(C_1/T)).$$

  The usual scale restriction $2^j > T^2$ implies that $R_{Bj} \leq C M_{Bj} |\log_2(T)|$.
- R_{Sj} is the number of bits needed to code the quadtree segmentation $\mathcal{S}_j$ for a single scale 2^j. One needs to differentiate between interior, vertical edge, and horizontal edge nodes, so 2 bits per segmentation coefficient is

needed, and thus $R_{Sj} = 2M_{Sj}$, where M_{Sj} is the number of coefficients needed to specify the dyadic segmentation $\mathcal{S}_j$, as described in Section 6.2.

- R_{Gj} is the number of bits needed to code the adapted geometric flow $\tilde{\gamma}'_S \in \mathcal{G}(S)$ inside each square S of each quadtree $\mathcal{S}_j$. Each geometric coefficient M_{Gj} is quantized and equation (6.5) shows that there are $\mathrm{Card}(\mathcal{G}(S)) = C_{\mathcal{G}}\lambda^{-(1-p)^2}$ possible quantized geometries where $C_{\mathcal{G}}$ is a constant. As $\lambda \geq 2^j$, the condition $2^j \geq T^2$ implies

$$R_{Gj} \leq M_{Gj} \log_2(C_{\mathcal{G}}\lambda^{-(1-p)^2}) \leq C M_{Gj} |\log_2(T)|.$$

Using this coding scheme, the total bit budget is thus

$$R = \sum_{2^j \geq T^2} (R_{Bj} + R_{Sj} + R_{Gj}) \leq C M |\log(T)|,$$

where $M = \sum_j (M_{Bj} + M_{Sj} + M_{Gj})$ is the total number of coefficients needed to specify f_M as described in Section 6.2.

In order to minimize $D(R)$, equation (8.5) shows that one should use the Lagrangian minimization of (6.12) with a Lagrange multiplier equal to $T/2$. An adaptive compression of the image is thus performed by using the best bandelet basis defined by

$$\mathcal{B}(\Gamma^*) \stackrel{\text{def}}{=} \operatorname*{argmin}_{\mathcal{B}(\Gamma) \in \mathcal{D}_{T^2}} \mathcal{L}(f, \mathcal{B}(\Gamma), T/2).$$

One needs to compute the distortion $D(R)$ in this basis and link this distortion with the number of bits R.

The bounds of equation (6.15) shows that the thresholding approximation f_M at $T/2$ in the bandelet basis $\mathcal{B}(\Gamma^*)$ satisfies

$$\|f - f_M\|^2_{\mathrm{L}^2} \leq C(T/2)^{\frac{2\alpha}{\alpha+1}} \quad \text{with } M \leq C(T/2)^{-\frac{2}{\alpha+1}}. \tag{8.6}$$

The scheme used to code the bandelet coefficients and the geometric parameters ensures that $R \leq CM|\log(T)|$. Combining this result with the bounds of equations (8.5) and (8.6) proves the following theorem:

THEOREM 8.1 *Let f be a C^α-geometrically-regular function. There exists $C > 0$ such that for any $T > 0$, the compressed image f_R in the best bandelet basis $\mathcal{B}(\Gamma^*)$ satisfies*

$$\|f - f_R\|^2_{\mathrm{L}^2} \leq C \log_2(R)^\alpha R^{-\alpha}.$$

The class of geometrically regular functions contains the class of uniformly C^α-functions, for which the Kolmogorov bound decreases like $R^{-\alpha}$; see [10]. This theorem thus proves that the asymptotic coding error decay in an adapted bandelet basis reaches the Kolmogorov lower bound for geometrically regular functions up to a $|\log(R)|^\alpha$ factor.

Appendix A: Proof of Lemma 6.2

PROOF: A bandelet basis $\mathcal{B}(\Gamma) \in \mathcal{D}_{T^2}$ is built by choosing, for each scale 2^j, a discrete Alpert basis $\mathcal{B}(\Gamma_j)$. At each scale 2^j the thresholded approximation f_{jM_j} of the wavelet coefficients f_j at T in $\mathcal{B}(\Gamma_j) = \{\Psi_\nu\}_\nu$ is defined by

$$f_{jM_j} = \sum_{|\langle f_j, \Psi_\nu\rangle| \geq T} \langle f_j, \Psi_\nu\rangle \, \Psi_\nu .$$

The number of parameters is defined following equation (6.8) by $M_j = M_{Gj} + M_{Sj} + M_{Bj}$. Using the fact that $\langle f, \Psi_\nu\rangle = \langle f_j, b_{\ell,m}\rangle$ for some index (ℓ, m), one can decompose the approximation error as

$$\| f - f_M \|_{\mathrm{L}^2}^2 = \sum_j \| f_j - f_{jM_j} \|_2^2 \quad \text{where } M \stackrel{\text{def}}{=} \sum_j M_j .$$

In order to build a bandelet basis adapted to the function f, one has to consider three approximation modes depending on the scale.

For fine scales. $2^j < 2^{j_0} \stackrel{\text{def}}{=} T^{\frac{2\alpha}{\alpha+1}}$, a standard result, states that $\| f_j \|_2^2 \leq C 2^j$ for a constant C that depends only on f. As $|f_j[n]| \leq \| f \|_{\mathrm{L}^\infty} \| \psi \|_{\mathrm{L}^1} 2^j$, one has, for T small enough, $f_j[n] = 0$, so $M_j = 0$ and

$$\sum_{j<j_0} \| f_j - f_{jM_j} \|_2^2 \leq \sum_{j<j_0} C 2^j \leq 2C 2^{j_0} = 2C T^{\frac{2\alpha}{\alpha+1}} \quad \text{with} \sum_{j<j_0} M_j = 0.$$

For coarse scales. $2^j > 2^{j_1} \stackrel{\text{def}}{=} T^{1/(\alpha+1)}$. There are fewer than 2^{-2j_1} coefficients so

$$\sum_{j>j_1} \| f_j - f_{jM_j} \|_2^2 \leq 2^{-2j_1} T^2 \leq T^{\frac{2\alpha}{\alpha+1}} .$$

Since there is no segmentation and no flow, $M_{Sj} = M_{Gj} = 0$ and hence

$$\sum_{j>j_1} M_j \leq 2^{-2j_1} \leq T^{\frac{-2}{\alpha+1}} .$$

For intermediate scales. $2^{j_0} \leq 2^j \leq 2^{j_1}$. A bandeletization basis $\mathcal{B}(\Gamma_j)$ adapted to f is used to approximate the coefficients. Such a basis is provided by Lemma A.2. This lemma proves that a thresholding in this basis gives the following error:

$$\begin{aligned}\sum_{j=j_0}^{j_1} \| f_j - f_{jM_j} \|_2^2 &\leq \sum_{j=j_0}^{j_1} C\left(s_j^{\frac{\alpha-\beta}{\delta}} T^{\frac{4\alpha\beta}{\delta}} + 2^{2j\alpha}\right) \quad \text{where } \delta \stackrel{\text{def}}{=} 2\alpha\beta + \alpha + \beta \\ &\leq 2C\left(2^{j_0\frac{\alpha-\beta}{\delta}} T^{\frac{4\alpha\beta}{\delta}} + 2^{2j_1\alpha}\right) \leq 4C T^{\frac{2\alpha}{\alpha+1}} .\end{aligned}$$

Combining all these bounds proves (6.14) in Lemma 6.2. □

A.1 Adapted Segmentation of Wavelet Coefficients

In order to match the approximation precision given by T, the optimal width of the square in an adapted segmentation $\mathcal{S}_j$ is set to

$$\lambda(T) \stackrel{\text{def}}{=} C_0 T^{\frac{2\beta}{\delta}} s_j^{\frac{\beta+1}{\delta}} \quad \text{and} \quad \delta \stackrel{\text{def}}{=} 2\alpha\beta + \alpha + \beta, \tag{A.1}$$

where C_0 is a constant that depends only on f and whose value is tuned during the proof of Lemma A.3. This definition of the optimal width is similar to the one given in equation (4.14), but it is parametrized by the threshold value T.

Section 5.2 has introduced the partition of squares $S \in \mathcal{S}_j$ into edge squares $S \in \mathcal{E}(\mathcal{S}_j)$, corner squares $S \in \mathcal{C}(\mathcal{S}_j)$, and regular squares $S \in \mathcal{R}(\mathcal{S}_j)$. Only edge squares $S \in \mathcal{E}$ contain an adapted flow $\tilde{\gamma}$, since in the other squares the original coefficients are not modified. Ideally, one would like to have only regular squares and squares of width λ. But topological constraints (such as edge crossings and corners) and curvature variations (horizontal edges that become vertical) force us to subdivide some squares of width λ into smaller ones. The set of edge squares is thus subdivided as $\mathcal{E} = \mathcal{E}_\lambda \cup \tilde{\mathcal{E}}_\lambda$ where $\mathcal{E}_\lambda$ is the set of squares length $\lambda(T)$ and $\tilde{\mathcal{E}}_\lambda$ contains squares of smaller lengths. The following lemma shows that one can build a dyadic segmentation with a small number of these subdivided squares $S \in \tilde{\mathcal{E}}_\lambda$:

LEMMA A.1 *Let f be a geometrically regular function. There exists a constant C such that for all $\lambda > 0$, there exists a dyadic segmentation $\mathcal{S}_j$ of $[0,1]^2$ into squares of width larger than 2^j that has the following properties*:

$$\begin{aligned} &\text{Card}(\mathcal{S}_j) \leq C(\lambda^{-1} + |\log_2(\lambda/2^j)|), && \text{Card}(\mathcal{E}_\lambda(\mathcal{S}_j)) \leq C\lambda^{-1}, \\ &\text{Card}(\tilde{\mathcal{E}}_\lambda(\mathcal{S}_j)) \leq C|\log_2(\lambda/2^j)|, && \text{Card}(\mathcal{C}(\mathcal{S}_j)) \leq C. \end{aligned}$$

PROOF: Following [14], a dyadic image segmentation is performed by iteratively labeling edge squares $S \in \mathcal{E}$ and corner squares $S \in \mathcal{C}$ while removing temporary unlabeled squares $S \in \mathcal{T}$. The algorithm proceeds as follows:

- Initialization: label the square $S = [0,1]^2$ as temporary $S \in \mathcal{T}$.
- Step 1: Split into four every temporary square $S \in \mathcal{T}$ and remove S from $\mathcal{T}$.
- Step 2: Label, in the following order, each new subdivided square S as follows:
  - A *regular square* $S \in \mathcal{R}(\mathcal{S}_j)$ if it is at a distance larger than s_j from all edges.
  - A *corner square* $S \in \mathcal{C}(\mathcal{S}_j)$ if its size is smaller than 2^j.
  - A *horizontal edge square* $S \in \mathcal{E}^H(\mathcal{S}_j)$ if a single horizontal edge component γ is closer than s_j to S. Following the edge square definition of Section 5.2, the curve is supposed to be parametrized horizontally with $|\gamma'| \leq 2$.

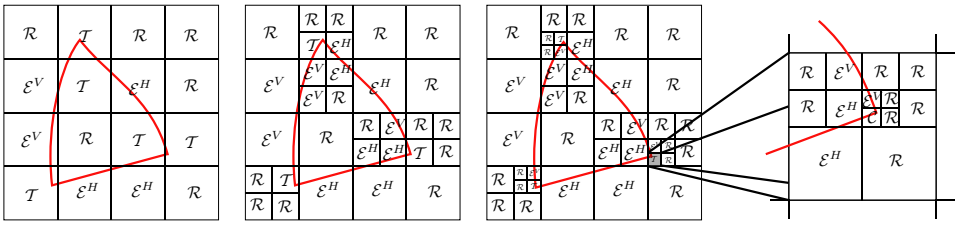

FIGURE A.1. Recursive subdivision of dyadic squares with the corresponding labeling $S \in \{\mathcal{R}, \mathcal{E}^V, \mathcal{E}^H, \mathcal{C}\}$.

  - A *vertical edge square* $S \in \mathcal{E}^V(\mathcal{S}_j)$ if a single vertical edge component γ is closer than s_j to S.
  - A *temporary square* $S \in \mathcal{T}$ otherwise.
- Step 3: Go to step 1 if any temporary squares remain.

Figure A.1 illustrates this process.

The first set of subdivision steps subdivides squares that contain edges until the square length reaches $\lambda(T)$. Due to the regularity of the curves, the number of resulting edge squares is of order $O(\lambda^{-1})$. The nontangency condition of the curves also ensures that after a constant number of steps, only a constant number of squares near junctions are labeled as temporary.

In the neighborhood of the junctions, the recursive splitting continues during $\log(\lambda/2^j)$ steps until the length of the squares is 2^j. Since there is only a finite number of such junctions, the number of the small edge squares is of order $O(\log(\lambda/2^j))$. □

The following lemma constructs an adapted bandeletization basis for a given scale 2^j. It uses the construction of an adapted dyadic segmentation together with an adapted quantized geometric flow that closely matches the real geometry.

LEMMA A.2 *Let f be a C^α-geometrically-regular function. There exists C such that for any $T > 0$, if 2^j satisfies*

$$T^{\frac{2\alpha}{\alpha+1}} \overset{\text{def}}{=} 2^{j_0} \leq 2^j \leq 2^{j_1} \overset{\text{def}}{=} T^{\frac{1}{\alpha+1}},$$

then there exists a bandeletization basis $\mathcal{B}(\Gamma_j) \in \mathcal{D}_j$ such that the thresholding approximation f_{jM_j} at T of f_j in this basis satisfies

$$\text{(A.2)} \qquad \begin{cases} \|f_j - f_{jM_j}\|_2^2 \leq C \max\left(s_j^{(\alpha-\beta)/\delta} T^{4\alpha\beta/\delta}, 2^{2j\alpha}\right), \\ M_j \leq C s_j^{(\alpha-\beta)/\delta} T^{-2(\alpha+\beta)/\delta} \quad \textit{with } \delta \overset{\text{def}}{=} 2\alpha\beta + \alpha + \beta \end{cases}$$

PROOF: The optimal width $\lambda = \lambda(T)$ of the edge squares is defined in equation (A.1). Lemma A.1 provides a dyadic segmentation of $[0,1]^2$ with edge squares conforming as much as possible to this optimal length λ. A bandeletization basis

$\mathcal{B}(\Gamma_j)$ is defined by choosing the bandelet basis $\mathcal{B}(S, \tilde{\gamma}'_S)$ provided by Lemma A.3 over edge squares $S \in \mathcal{E}_\lambda(\mathcal{S}_j) \cup \tilde{\mathcal{E}}_\lambda(\mathcal{S}_j)$ and by keeping the original wavelet coefficients over the remaining squares. Wavelet coefficients inside a square S are denoted $f_j[n]$, and the corresponding thresholded approximation with $M(S)$ coefficients is denoted $f_{jM(S)}$. One has the following error:

$$\|f_j - f_{jM_j}\|_2^2 = \sum_{S\in\mathcal{R}(\mathcal{S}_j)} \|f_j - f_{jM(S)}\|^2_{\ell^2(S)} + \sum_{S\in\mathcal{C}(\mathcal{S}_j)} \|f_j - f_{jM(S)}\|^2_{\ell^2(S)} + \sum_{S\in\mathcal{E}_\lambda(\mathcal{S}_j)} \|f_j - f_{jM(S)}\|^2_{\ell^2(S)} + \sum_{S\in\tilde{\mathcal{E}}_\lambda(\mathcal{S}_j)} \|f_j - f_{jM(S)}\|^2_{\ell^2(S)}$$

with

$$M_{Bj} = \sum_{S\in\mathcal{R}(\mathcal{S}_j)} M(S) + \sum_{S\in\mathcal{C}(\mathcal{S}_j)} M(S) + \sum_{S\in\mathcal{E}_\lambda(\mathcal{S}_j)} M(S) + \sum_{S\in\tilde{\mathcal{E}}_\lambda(\mathcal{S}_j)} M(S).$$

In *regular squares* $S \in \mathcal{R}(\mathcal{S}_j)$, we saw in equation (2.2) that

$$\forall (2^j n) \in S, \quad |\langle f, \psi_{jn}\rangle| \le C_f 2^{j(1+\alpha)},$$

where C_f is proportional to $\|f\|_{C^\alpha(\Lambda)}$. Let $2^{j_\star}$ be the cutoff scale defined by

$$2^{j_\star} \stackrel{\text{def}}{=} (C_f)^{-\frac{1}{\alpha+1}} T^{\frac{1}{\alpha+1}}.$$

If $j < j_\star$, one has $|f_j[n]| < T$; thus $\sum_{S\in\mathcal{R}(\mathcal{S}_j)} M(S) = 0$. If $j \ge j_\star$, one has

$$\sum_{S\in\mathcal{R}(\mathcal{S}_j)} M(S) \le 2^{-2j_\star} \le (C_f)^{\frac{2}{\alpha+1}} T^{-\frac{2}{\alpha+1}} \le (C_f)^{\frac{2}{\alpha+1}} s_j^{\frac{\alpha-\beta}{\delta}} T^{-2\frac{\alpha+\beta}{\delta}}.$$

In both cases one has

$$\sum_{S\in\mathcal{R}(\mathcal{S}_j)} \|f_j - f_{jM(S)}\|^2_{\ell^2(S)} \le 2^{-2j} \big(\max_{S\subset\mathcal{R}(\mathcal{S}_j)} |f_j[n]|^2\big) \le (C_f)^2 2^{2j\alpha}.$$

In *corner squares* $S \in \mathcal{C}(\mathcal{S}_j)$, there is a constant number C of coefficients and thus

$$\sum_{S\in\mathcal{C}(\mathcal{S}_j)} \|f_{jM(S)} - f_j\|^2_{\ell^2(S)} \le C\, T^2 \le C s_j^{\frac{\alpha-\beta}{\delta}} T^{\frac{4\alpha\beta}{\delta}} \quad \text{and} \quad \sum_{S\in\mathcal{C}(\mathcal{S}_j)} M(S) \le C,$$

which satisfies bounds (A.2).

In *edge squares* $S \in \mathcal{E}_\lambda(\mathcal{S}_j)$ of size λ, Lemma A.3 bounds the bandelet approximation error and the number of coefficients needed. Since there are fewer than $C\lambda^{-1}$ such squares, one has

$$\sum_{S\in\mathcal{E}_\lambda(\mathcal{S}_j)} \|f_j - f_{jM(S)}\|^2_{\ell^2(S)} \le (C\lambda^{-1})\lambda s_j^{\frac{\alpha-\beta}{\delta}} T^{\frac{4\alpha\beta}{\delta}} \tag{A.3}$$

together with

$$\sum_{S \in \mathcal{E}_\lambda(\mathcal{S}_j)} M(S) \leq (C\lambda^{-1})\lambda s_j^{\frac{\alpha-\beta}{\delta}} T^{-2\frac{\alpha+\beta}{\delta}}, \tag{A.4}$$

and since

$$\lambda(T) \stackrel{\text{def}}{=} C_0\, T^{\frac{2\beta}{\delta}} s_j^{\frac{\beta+1}{\delta}} \quad \text{and} \quad \delta \stackrel{\text{def}}{=} 2\alpha\beta + \alpha + \beta, \tag{A.5}$$

one get (A.2) by inserting (A.5) into (A.3) and (A.4).

In *small edge squares* $S \in \tilde{\mathcal{E}}(\mathcal{S}_j)$ of size less than λ, Lemma A.3 still applies. The number of bandelet coefficients $M(S)$ is bounded by the number of coefficients needed for squares of optimal size λ. Since there are fewer than $|\log_2(\lambda/2^j)| \leq C_1|\log_2(T)|$ such squares, bounds (A.2) still hold.

Geometric coefficients. By combining these bounds together, one gets the error bound of equation (A.2). Equations (6.9) and (6.10) show that

$$\begin{aligned} M_{Sj} + M_{Gj} &\leq p\,\mathrm{Card}(\mathcal{S}_j) \leq C\lambda^{-1} \leq CC_0^{-1}T^{-2\frac{\beta}{\delta}} s_j^{-\frac{\beta+1}{\delta}} \\ &\leq CC_0^{-1} s_j^{\frac{\alpha-\beta}{\delta}} T^{-2\frac{\alpha+\beta}{\delta}} \underbrace{\left(T^{-2\frac{\alpha}{\delta}} s_j^{-\frac{1+\alpha}{\delta}}\right)}_{\leq 1 \text{ since } s_j \geq T^{2\alpha/(\alpha+1)}} \leq (CC_0^{-1}) s_j^{\frac{\alpha-\beta}{\delta}} T^{-2\frac{\alpha+\beta}{\delta}}, \end{aligned}$$

which gives the bound of equation (A.2) for the number of coefficients. □

The technical analysis of the bandelet approximation in an edge square S is detailed in Lemma A.3. In this lemma, f_j is the set of wavelet coefficients inside S, $f_{jM(S)}$ is the thresholded approximation of f_j at T in basis $\mathcal{B}(S, \tilde{\gamma}')$, and $M(S)$ is the number of needed bandelet coefficients.

LEMMA A.3 *Let f be a C^α-geometrically-regular function. There exists $C > 0$ such that for any $T > 0$ and 2^j satisfying*

$$T^{\frac{2\alpha}{\alpha+1}} \stackrel{\text{def}}{=} 2^{j_0} \leq 2^j \leq 2^{j_1} \stackrel{\text{def}}{=} T^{\frac{1}{\alpha+1}}, \tag{A.6}$$

for any edge square S of width $\lambda \stackrel{\text{def}}{=} \lambda(T)$ there exists an adapted geometric flow $\tilde{\gamma}' \in \mathcal{G}(S)$ such that

$$\begin{cases} \|f_j - f_{jM(S)}\|^2_{\ell^2(S)} \leq C\lambda s_j^{(\alpha-\beta)/\delta} T^{4\alpha\beta/\delta} \\ M(S) \leq C\lambda s_j^{(\alpha-\beta)/\delta} T^{-2(\alpha+\beta)/\delta} \quad \textit{with } \delta \stackrel{\text{def}}{=} 2\alpha\beta + \alpha + \beta. \end{cases} \tag{A.7}$$

PROOF: The edge curve is parametrized in S by $x_2 = \gamma(x_1)$. Let $\tilde{\gamma}'_0$ be a Taylor polynomial expansion of degree $\alpha-2$ of γ' inside S. An adapted polynomial flow $\tilde{\gamma}' \in \mathcal{G}(S)$ quantized with a precision $\tau = \lambda^\alpha/(p-1)$ is defined by

$$\tilde{\gamma}'_0(x) = \sum_{i=0}^{p-2} \frac{\alpha_i}{\lambda^{i+1}} x^i \quad \text{and} \quad \tilde{\gamma}'(x) = \sum_{i=0}^{p-2} \frac{a_i \tau}{\lambda^{i+1}} x^i \quad \text{with } a_i = Q_\tau(\alpha_i),$$

where the uniform quantizer Q_τ is defined by

$$Q_\tau(x) = q\tau \quad \text{if } (q-\tfrac{1}{2})\tau \le x \le (q+\tfrac{1}{2})\tau.$$

Since the edge square definition of equation (5.5) enforces that $|\gamma'| \le 2$, one has

$$\forall x \in [0,1], \quad \left|\sum_{i=0}^{p-2} \left(\frac{\alpha_i}{\lambda}\right) x^i\right| \le 2 \quad \Longrightarrow \quad |a_i| \le 2\overline{C}\,\frac{\lambda}{\tau},$$

where the constant $\overline{C}$ is defined by equation (6.6) and thus $\tilde{\gamma}' \in \mathcal{G}(S)$. Furthermore, this adapted flow satisfies

$$\begin{aligned} \|\tilde{\gamma}' - \gamma'\|_{\mathrm{L}^\infty} &\le \|\tilde{\gamma}' - \tilde{\gamma}'_0\|_{\mathrm{L}^\infty} + \|\tilde{\gamma}'_0 - \gamma'\|_{\mathrm{L}^\infty} \\ &\le (p-1)\frac{\tau}{\lambda} + \|\gamma\|_{\mathrm{C}^\alpha}\lambda^{\alpha-1} \le (1+\|\gamma\|_{\mathrm{C}^\alpha})\lambda^{\alpha-1}. \end{aligned} \tag{A.8}$$

The set of points x such that $|\gamma(x_1) - x_2| \le K s_j$ is denoted $\mathcal{T}$. The support of an Alpert vector is $\beta_{\ell,m}$.

A.2 Bounding Bandelet Coefficients

There are two kinds of Alpert vectors $\Psi_{\ell,m} \in \mathcal{B}(S, \tilde{\gamma}')$:

- If $\beta_{\ell,m}$ does not intersect $\mathcal{T}$, the fact that f is C^α-regular away from edges implies

$$\forall (2^j n) \notin \mathcal{T}, \quad |f_j[n]| = |\langle f, \psi_{jn}\rangle| \le C_f 2^{j(1+\alpha)}, \tag{A.9}$$

  where C_f is proportional to $\|f\|_{\mathrm{C}^\alpha(\Lambda)}$, which leads to

$$|\langle f_j, \Psi_{\ell,m}\rangle| \le C_f \|f_j\|_\infty \|\Psi_{\ell,m}\|_1 \le (C_f C_\Psi) 2^{j\alpha} 2^{\frac{\ell}{2}} \lambda$$

  where $\|\Psi_{\ell,m}\|_1$ is given by equation (5.1). A bandelet cutoff scale 2^{L_0} is defined by

$$(C_f\, C_\Psi) 2^{j\alpha} 2^{L_0/2} \lambda = T \quad \Longrightarrow \quad 2^{L_0} \stackrel{\text{def}}{=} (C_f C_\Psi)^{-2} 2^{-2j\alpha} T^2 \lambda^{-2},$$

  hence

$$2^\ell \le 2^{L_0} \quad \Longrightarrow \quad |\langle f, \Psi_{\ell,m}\rangle| \le T. \tag{A.10}$$

- If $\beta_{\ell,m}$ intersects $\mathcal{T}$, one needs to use the regularity of the warped function $f_{j\mathrm{W}}$. Let $P(x)$ be a Taylor polynomial expansion of degree p of the function $f_{j\mathrm{W}}$ inside $w(\beta_{\ell,m})$, where w is the warping operator defined by equation (4.2). The definition (A.1) of $\lambda(\tau)$ together with the bound $s_j \ge 2^j \ge T^{2\alpha/(\alpha+1)}$ implies that $\lambda(\tau) \le C_0 s_j^{1/\alpha}$. Equation (A.8) then allows us to apply Proposition 4.1, which provides bounds on the derivatives of $f_{j\mathrm{W}}$. The construction of the band detailed in Section 5.1 enforces

that $w(\beta_{\ell,m})$ is of size $\lambda \times \mu$ where $\mu \stackrel{\text{def}}{=} \lambda 2^{\ell}$. Similarly to equation (4.12), the error is bounded by

$$\begin{aligned}\forall \tilde{x} \in w(\beta_{\ell,m}), \quad |f_{j\mathrm{W}}(\tilde{x}) - P(\tilde{x})| &\leq \sum_{(i_1,i_2)\in I_\alpha^p} \left\| \frac{\partial^\alpha f_{j\mathrm{W}}}{\partial x_1^{i_1} \partial x_2^{i_2}} \right\|_{\mathrm{L}^\infty} \lambda^{i_1} \mu^{i_2} \\ &\leq C(1+\|\gamma\|_{\mathrm{C}^\alpha}^\alpha) \sum_{(i_1,i_2)\in I_\alpha^p} 2^j s_j^{-\frac{i_1}{\alpha-i_2}} \lambda^{i_1} \mu^{i_2} \\ &\leq C_{\mathrm{W}} 2^j \left(s_j^{-1}\lambda^\alpha + s_j^{-\beta}(2^\ell\lambda)^\beta\right),\end{aligned}$$

where the constant C_{W} depends on $\|\gamma\|_{\mathrm{C}^\alpha}$ and $\|f\|_{\mathrm{C}^\alpha(\Lambda)}$ and where the set of indices is

$$I_\alpha^p \stackrel{\text{def}}{=} \left\{ (i_1,i_2) \;\middle|\; \begin{array}{l} i_1+i_2 = p \text{ and } i_1 < \alpha, \\ \alpha + i_2 \leq p \text{ and } i_1 = \alpha \end{array} \right\}.$$

The result is that for all $\tilde{x} = w(x) \in w(\beta_{\ell,m})$,

$$f_{j\mathrm{W}}(\tilde{x}) = P(\tilde{x}) + \epsilon(\tilde{x}) \quad \text{with } \|\epsilon\|_{\mathrm{L}^\infty} \leq C_{\mathrm{W}} 2^j \left(s_j^{-1}\lambda^\alpha + s_j^{-\beta}(2^\ell\lambda)^\beta\right). \tag{A.11}$$

The bandelets inner products are bounded using the fact that $\Psi_{\ell,m}$ is orthogonal to the space of discrete warped polynomials of degree $p-1$ on $w(\beta_{\ell,m})$,

$$\begin{aligned}\langle f, \Psi_{\ell,m}\rangle = \sum_{2^j n \in \beta_{\ell,m}} f_j[n]\Psi_{\ell,m}[n] &= \sum_{2^j n \in \beta_{\ell,m}} f_{j\mathrm{W}}(w(2^j n))\Psi_{\ell,m}[n] \\ &= \sum_{2^j n \in \beta_{\ell,m}} \epsilon(w(2^j n))\Psi_{\ell,m}[n],\end{aligned}$$

which leads to the bound

$$\begin{aligned}|\langle f, \Psi_{\ell,m}\rangle| &\leq \|\epsilon\|_{\mathrm{L}^\infty} \|\Psi_{\ell,m}\|_1 \\ &\leq (C_{\mathrm{W}} C_\Psi) \lambda^{\beta+1} s_j^{-\beta} 2^{\frac{\ell}{2}} \max\left(s_j^{\beta-1}\lambda^{\alpha-\beta} 2^{\ell\beta}\right),\end{aligned} \tag{A.12}$$

where $\|\Psi_{\ell,m}\|_1$ is bounded by equation (5.1). A bandelet cutoff scale 2^{L_1} is defined by

$$s_j^{\beta-1}\lambda^{\alpha-\beta} = 2^{L_1\beta} \quad \Longrightarrow \quad 2^{L_1} \stackrel{\text{def}}{=} s_j^{1-1/\beta} \lambda^{\frac{\alpha}{\beta}-1}.$$

The constant C_0 that defines the optimal length $\lambda(T)$ in equation (A.1) is set to

$$C_0 \stackrel{\text{def}}{=} (C_{\mathrm{W}} C_\Psi)^{-\frac{\delta}{2p}},$$

which implies

$$2^\ell \leq 2^{L_1} \quad \Longrightarrow \quad |\langle f, \Psi_{\ell,m}\rangle| \leq (C_{\mathrm{W}} C_\Psi) \lambda^{\frac{\delta}{2\beta}} s_j^{-\frac{1+\beta}{2}} \leq T. \tag{A.13}$$

A.3 Bounding $M(S)$

The number of coefficients above the threshold

$$M(S) \stackrel{\text{def}}{=} \text{Card}(J_T) \quad \text{where } J_T \stackrel{\text{def}}{=} \{(\ell,m) \mid |\langle f, \Psi_{\ell,m}\rangle| \geq T\}$$

can be bounded using the sets

$$J_0 \stackrel{\text{def}}{=} \{(\ell,m) \mid \text{Supp}(\Psi_{\ell,m}) \cap \mathcal{T} = \varnothing \text{ and } 2^\ell \geq 2^{L_0}\}$$

and

$$J_1 \stackrel{\text{def}}{=} \{(\ell,m) \mid \text{Supp}(\Psi_{\ell,m}) \cap \mathcal{T} \neq \varnothing \text{ and } 2^\ell \geq 2^{L_1}\}$$

since equations (A.10) and (A.13) imply that $J_T \subset (J_0 \cup J_1)$.

One has

$$\begin{aligned}\text{Card}(J_0) \leq \sum_{\ell \geq L_0} 2^{-\ell} &\leq 2(2^{-L_0}) \\ &\leq 2(C_f C_\Psi)^2 \lambda \underbrace{2^{2j\alpha}}_{\leq T^{2\alpha/(\alpha+1)}} \underbrace{\lambda}_{\leq T^{2\beta/\delta}} T^{-2} \\ &\leq 2(C_f C_\Psi)^2 \lambda T^{-2\frac{\alpha}{\alpha+1}\frac{\beta+1}{\delta}} \leq 2(C_f C_\Psi)^2 \lambda s_j^{\frac{\alpha-\beta}{\delta}} T^{-2\frac{\alpha+\beta}{\delta}}.\end{aligned}$$

For each scale 2^ℓ, there are fewer than $K2^j/(\lambda 2^\ell)$ bandelets in $\{\Psi_{\ell,m}\}_m$ that intersect $\mathcal{T}$, so

$$\begin{aligned}\text{Card}(J_1) \leq \sum_{\ell \geq L_1} \frac{K2^j}{\lambda 2^\ell} &\leq 2K2^j \lambda^{-1} 2^{-L_1} \\ &\leq 2K2^j s_j^{\frac{1}{\beta}-1} \lambda^{-\frac{\alpha}{\beta}} \leq 2Ks_j^{\frac{1}{\beta}} \lambda^{-\frac{\alpha}{\beta}} \leq 2K\lambda s_j^{\frac{\alpha-\beta}{\delta}} T^{-2\frac{\alpha+\beta}{\delta}}.\end{aligned}$$

The number of coefficients is thus bounded by

$$M(S) \leq \text{Card}(J_0) + \text{Card}(J_1) \leq C_2 \lambda s_j^{\frac{\alpha-\beta}{\delta}} T^{-2\frac{\alpha+\beta}{\delta}}.$$

A.4 Bounding $\|f_j - f_{jM(S)}\|_2^2$

One has

$$\begin{aligned}\|f_j - f_{jM(S)}\|_2^2 &= \sum_{(\ell,m)\notin J_T} |\langle f_j, \Psi_{\ell,m}\rangle|^2 \\ &= \sum_{(\ell,m)\in (J_0\cup J_1)\setminus J_T} |\langle f_j, \Psi_{\ell,m}\rangle|^2 \\ &\quad + \sum_{(\ell,m)\in \widetilde{J}_o \cup \widetilde{J}_1} |\langle f_j, \Psi_{\ell,m}\rangle|^2,\end{aligned} \tag{A.14}$$

where the set of coefficients are split according to

$$\widetilde{J}_0 \stackrel{\text{def}}{=} \{(\ell, m) \mid \text{Supp}(\Psi_{\ell,m}) \cap \mathcal{T} = \varnothing \text{ and } 2^\ell < 2^{L_0}\}$$

and

$$\widetilde{J}_1 \stackrel{\text{def}}{=} \{(\ell, m) \mid \text{Supp}(\Psi_{\ell,m}) \cap \mathcal{T} \neq \varnothing \text{ and } 2^\ell < 2^{L_1}\}.$$

The first part of the error in equation (A.14) is bounded using

$$\begin{aligned}\sum_{(\ell,m)\in(J_0\cup J_1)\setminus J_T} |\langle f_j, \Psi_{\ell,m}\rangle|^2 &\leq (\text{Card}((J_0 \cup J_1) \setminus J_T)T^2 \\ &\leq (\text{Card}(J_0 \cup J_1)T^2 \leq C_2\lambda s_j^{\frac{\alpha-\beta}{\delta}} T^{\frac{4\alpha\beta}{\delta}}.\end{aligned}$$

Bounding the contribution of $\widetilde{J}_0 \cup \widetilde{J}_1$ to the error using directly equation (A.12) leads to a suboptimal result. The linear space

$$V \stackrel{\text{def}}{=} \text{Span}\{\Psi_{\ell,m} \mid (\ell, m) \in J_0 \cup J_1\}$$

allows us to write the second part of the error in equation (A.14) as a projection

$$\sum_{(\ell,m)\in\widetilde{J}_o\cup\widetilde{J}_1} |\langle f_j, \Psi_{\ell,m}\rangle|^2 = \|f_j - \text{P}_V(f_j)\|_2^2.$$

The following decomposition

$$f_j = f_j^0 + f_j^1 \quad \text{where} \begin{cases} \forall (2^j n) \in \mathcal{T} \;\; f_j^0[n] = 0 \text{ and } f_j^1[n] = f_j[n], \\ \forall (2^j n) \notin \mathcal{T} \;\; f_j^0[n] = f_j[n] \text{ and } f_j^1[n] = 0, \end{cases}$$

leads to

$$\begin{aligned}\|f_j - \text{P}_V(f_j)\|_2 &\leq \|f_j^1 - \text{P}_V(f_j^1)\|_2 + \|f_j^0\|_2 + \|\text{P}_V(f_j^0)\|_2 \\ &\leq \|f_j^1 - \text{P}_V(f_j^1)\|_2 + 2\|f_j^0\|_2.\end{aligned}$$

The norm of f_j^0 is bounded using the fact that the wavelet coefficients are bounded by equation (A.9) outside $\mathcal{T}$,

$$\begin{aligned}\|f_j^0\|_2^2 &\leq \underbrace{(\lambda^2 2^{-2j})}_{\text{neighboring points in } S} (C_f)^2 2^{2j(\alpha+1)} \\ &\leq (C_f)^2\lambda \underbrace{2^{2j\alpha}}_{\leq T^{2\alpha/(\alpha+1)}} \underbrace{\lambda}_{\leq T^{2\beta/\delta}} \\ &\leq (C_f)^2\lambda T^{\frac{2\alpha}{\alpha+1}+\frac{2\beta}{\delta}} s_j^{\frac{\beta+1}{\delta}} \leq (C_f)^2\lambda T^{\frac{4\alpha\beta}{\delta}} s_j^{\frac{\alpha-\beta}{\delta}}.\end{aligned}$$

Equation (A.11) allows us to decompose $f_{j\text{W}} = P + \epsilon$ inside each warped domain $w(\beta_{L_0,m})$. The fact that warped discrete polynomials P belong to V implies that the pointwise error is bounded by

$$\forall (2^j n) \in \beta_{L_0,m}, \quad |f_j^1[n] - \text{P}_V(f_j^1)[n]| \leq 2\|\epsilon\|_{\text{L}^\infty} \leq 4C_{\text{W}} 2^j s_j^{-1}\lambda^\alpha.$$

Since the number of points $(2^j n) \subset \mathcal{T}$ where $f_j^1[n] \neq 0$ is $(Ks_j/\lambda)(b2^{-j})^2$, one gets the estimate

$$\begin{aligned}\|f_j - \mathrm{P}_V(f_j)\|_2^2 &\leq \left(\frac{Ks_j}{\lambda}\right)(\lambda 2^{-j})^2(2\|\epsilon\|_{\mathrm{L}^\infty})^2 \\ &\leq C s_j^{-1}\lambda^{2\alpha+1} \leq C\lambda T^{\frac{4\alpha\beta}{\delta}} s_j^{\frac{\alpha-\beta}{\delta}}.\end{aligned}$$

□

Appendix B: Fast Alpert Transform

Let $\mathcal{B}(S, \tilde{\gamma}') = \{\Psi_{\ell,m}\}_{\ell,m}$ be an Alpert bandeletization basis over a square S containing wavelet coefficients $f_j[n]$ for $(2^j n) \in S$. An adapted Alpert transform [1] computes with a fast algorithm the coefficients $\{\langle f_j, \Psi_{\ell,m}\rangle\}_{\ell,m}$ of f_j with a complexity of $O(N)$ where N is the number of input wavelet coefficients $f_j[n]$.

For $x \in \mathbb{R}^2$, let $x^t = x_1^{t_1} x_2^{t_2}$ and $|t| = t_1 + t_2$. The two-dimensional monomials are indexed using

$$\forall t \text{ such that } |t| < p, \quad \tilde{t} \overset{\text{def}}{=} \frac{1}{2}|t|(|t|+1) + t_2 \in \left\{0, \ldots, \frac{p(p+1)}{2} - 1\right\}.$$

In the following, the sampling locations of wavelet coefficients are written as $x_n \overset{\text{def}}{=} 2^j n \in [0,1]^2$ and the warped points $\tilde{x}_n \overset{\text{def}}{=} w(x_n)$ where the warping is defined by (4.2). If A and B are two matrices, $[A; B]$ denotes the concatenation along columns, $[A, B]$ the concatenation along rows, and $\mathrm{diag}(A, B)$ the concatenation along the diagonal.

B.1 Polynomial Inner Product

The dot product of two polynomials P and Q defined over a band $\tilde{\beta}_{\ell,m}$ is

$$\langle P, Q\rangle_{\ell,m} \overset{\text{def}}{=} \sum_{x_n \in \beta_{\ell,m}} P(\tilde{x}_n) Q(\tilde{x}_n) = \overline{P}^{\mathsf{T}} A_{\ell,m} \overline{Q},$$

where $\overline{P}$ is the column vector of coefficients of P in the basis of monomials

$$P(x) = \sum_{|t|<p} \overline{P}_{\tilde{t}}\, x^p.$$

For all m, the symmetric matrices $A_{\ell,m}$ of size $p(p+1)/2 \times p(p+1)/2$ are the matrices of the dot products over $\tilde{\beta}_{\ell,m}$ for discrete polynomial vectors expressed in the basis of monomials, defined by

$$\forall m = 0, \ldots, 2^\ell - 1, \quad (A_{\ell,m})_{\widetilde{s,t}} \overset{\text{def}}{=} \sum_{x_n \in \beta_{\ell,m}} (\tilde{x}_n)^{s+t}. \tag{B.1}$$

These matrices are computed iteratively during the Alpert transform.

B.2 Initialization

The finest partition $S = \bigcup_{m=0}^{2^{-L}-1} \beta_{Lm}$ is computed by recursive splits as explained in Section 5.1. For all $m \in \{0, \dots, 2^{-\ell} - 1\}$, one needs to compute the polynomial P_{Lm} of degree p that interpolates f_j on β_{Lm},

$$\forall x_n \in \beta_{Lm}, \quad P_{Lm}(\tilde{x}_n) = f_j[n].$$

These polynomials correspond to the finest scale representation of f_j in the basis of local monomials. The matrices A_{Lm} must also be calculated with (B.1) for $\ell = L$.

B.3 Computation of Alpert Coefficients

For each scale $\ell = L, \dots, -1$ and each m, the pair of polynomials $(P_0, P_1) \stackrel{\text{def}}{=} (P_{\ell,2m}, P_{\ell,2m+1})$ is decomposed into a sum of a polynomial $P \stackrel{\text{def}}{=} P_{\ell+1,m}$ defined on the whole band $\beta_{\ell+1,m}$ and a residual (Q_0, Q_1) that is orthogonal to polynomials of degree less than $p-1$. The polynomials P_0 and Q_0 are restricted to $\tilde{\beta}_{\ell,2m}$, and the polynomials P_1 and Q_1 are restricted to $\tilde{\beta}_{\ell,2m+1}$, whereas the polynomial $P = P_0 - Q_0 = P_1 - Q_1$ is restricted to $\tilde{\beta}_{\ell+1,m}$. In the following we denote $A_0 = A_{\ell,2m}$, $A_1 = A_{\ell,2m+1}$, and $A \stackrel{\text{def}}{=} \text{diag}(A_0, A_1)$.

The residual polynomials (Q_0, Q_1) are computed as the projection of (P_0, P_1) onto a set of orthogonal couples of polynomials $\{(h_t^0, h_t^1)\}_{0 \le |t| < p}$. Each couple $h_t \stackrel{\text{def}}{=} (h_t^0, h_t^1)$ is the piecewise polynomial that interpolates the bandelet vector $\Psi_{\ell \tilde{i}}$ for $\tilde{i} \stackrel{\text{def}}{=} p(p+1)i/2 + \tilde{t}$ over the set of locations $\tilde{\beta}_{\ell+1,m} = \tilde{\beta}_{\ell,2m} \cup \tilde{\beta}_{\ell,2m+1}$,

$$\forall \tilde{x}_n \in \tilde{\beta}_{\ell,2m+\epsilon}, \quad b_{\ell,\tilde{i}}[n] \stackrel{\text{def}}{=} h_t^\epsilon(\tilde{x}_n) \quad \text{for } \epsilon \in \{0,1\}.$$

The bandelet piecewise polynomials h_t are computed by satisfying two criteria:

- *Orthogonality*. For each $0 \le |t|, |t'| < p$, one should have

$$\begin{aligned} \delta_{t,t'} &= \sum_{x_n \in \beta_{\ell+1,i}} h_t(\tilde{x}_n) h_{t'}(\tilde{x}_n) \\ &= \sum_{x_n \in \beta_{\ell,2i}} h_t^0(\tilde{x}_n) h_{t'}^0(\tilde{x}_n) + \sum_{x_n \in \beta_{\ell,2i+1}} h_{t'}^1(\tilde{x}_n) h_t^1(\tilde{x}_n) \end{aligned} \tag{B.2}$$

  where δ is the Kronecker symbol.

- *Vanishing moments*. For each $0 \le |s|, |t| < p$, h_t is orthogonal to the monomial x^s defined on $\tilde{\beta}_{\ell,m}$,

$$\begin{aligned} 0 &= \sum_{x_n \in \beta_{\ell+1,i}} h_t(\tilde{x}_n)(\tilde{x}_n)^s \\ &= \sum_{x_n \in \beta_{\ell,2i}} h_t^0(\tilde{x}_n)(\tilde{x}_n)^s + \sum_{x_n \in \beta_{\ell,2i+1}} h_{t'}^1(\tilde{x}_n)(\tilde{x}_n)^s. \end{aligned} \tag{B.3}$$

One can express conditions (B.2) and (B.3) in matrix form in the basis of the monomials as

$$\text{(B.4)}\quad \big(\text{(B.2)} \iff H^{\mathsf{T}} A H = \mathrm{Id}\big) \quad \text{and} \quad \big(\text{(B.3)} \iff [A_0, A_1] H = 0\big),$$

where the column of index $\tilde{t}$ of the matrix H of size $p(p+1) \times p(p+1)/2$ is $[\overline{h}_t^0; \overline{h}_t^1]$. Conditions (B.4) mean that

$$A^{1/2} H \text{ is an orthogonal basis of the kernel of } [A_0, A_1] A^{-1/2}.$$

Matrix H can thus be computed with a constant number of operations proportional to $(p(p+1)/2)^3$. The residual is computed by $[\overline{Q}_0; \overline{Q}_1] = H[\overline{P}_0; \overline{P}_1]$, and the low-scale polynomial is $P = P_0 - Q_0$. The bandelet coefficients are the projections of (Q_0, Q_1) onto the computed basis using the dot product defined by A,

$$\langle f_j, \Psi_{\ell \widetilde{m}} \rangle = \langle h_t^0, Q_0 \rangle_{\ell, 2m} + \langle h_t^1, Q_1 \rangle_{\ell, 2m+1} = (\overline{h}_t)^{\mathsf{T}} A [\overline{Q}_0; \overline{Q}_1]$$

for $\widetilde{m} \stackrel{\text{def}}{=} p(p+1)i/2 + \tilde{t}$. The dot product matrix for the next scale is computed using $A_{\ell+1,m} = A_{\ell,2m} + A_{\ell,2m+1}$.

A constant number of operations is needed to compute each bandelet matrix H and to update matrices $A_{\ell,m}$ for each scale 2^ℓ. Since this process is repeated $N/2^\ell$ times for each scale, the overall complexity of the algorithm is $O(N)$ to transform N wavelet coefficients $f_j[n]$.

Bibliography

[1] Alpert, B. K. A class of bases in L^2 for the sparse representation of integral operators. *SIAM J. Math. Anal.* **24** (1993), no. 1, 246–262.

[2] Bosking, W. H.; Zhang, Y.; Schofield, B.; Fitzpatrick, D. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* **17** (1997), no. 6, 2112–2127.

[3] Breiman, L.; Friedman, J.; Olshen, R. A.; Stone, C. J. *Classification and regression trees*. Wadsworth Statistics/Probability Series. Wadsworth, Belmont, Calif., 1984.

[4] Candès, E. J.; Donoho, D. L. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.* **57** (2004), no. 2, 219–266.

[5] Claypoole, R. L.; Davis, G. M.; Sweldens, W.; Baraniuk, R. G. Nonlinear wavelet transforms for image coding via lifting. *IEEE Trans. Image Process.* **12** (2003), no. 12, 1449–1459.

[6] Cohen, A.; Daubechies, I.; Vial, P. Wavelets on the interval and fast wavelet transforms. *Appl. Comput. Harmon. Anal.* **1** (1993), no. 1, 54–81.

[7] DeAngelis G. C.; Ohzawa, I.; Freeman, R. D. Receptive-field dynamics in the central visual pathways. *Trends Neurosci.* **18** (1995), no. 10, 451–458.

[8] Do, M. N.; Vetterli, M. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Trans. Image Process.* **14** (2005), no. 12, 2091–2106.

[9] Donoho, D. L. Wedgelets: nearly-minimax estimation of edges. *Ann. Statist.* **27** (1999), no. 3, 859–897.

[10] Donoho, D. L. Counting bits with Kolmogorov and Shannon. Technical Report No. 2000-38, Department of Statistics, Stanford University, December 2000. Available online at: `www-stat.stanford.edu/reports/abstracts/00-38.ps`

[11] Dragotti, P. L.; Vetterli, M. Wavelet footprints: theory, algorithms, and applications. *IEEE Trans. Signal Process*. **51** (2003), no. 5, 1306–1323.

[12] Field, D. J.; Hayes, A.; Hess, R. F. Contour integration by the human visual system: evidence for a local "association field." *Vision Res*. **33** (1993), no. 2, 173–193.

[13] Hubel, D. H.; Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol*. **195** (1968), no. 1, 215–243.

[14] Le Pennec, E.; Mallat, S. Bandelet image approximation and compression. *SIAM Multiscale Modeling and Simulation* **4** (2005), no. 3, 992–1039.

[15] Lee, T. S. Computations in the early visual cortex. *J. Physiol. Paris* **97** (2003), no. 2-3, 121–139.

[16] Mallat, S. *A wavelet tour of signal processing*. Academic Press, San Diego, 1998.

[17] Matei, B.; Cohen, A. Nonlinear subdivison schemes: applications to image processing. *Tutorials on multiresolution in geometric modelling*, 93–97. Springer New York, Secaucus, N.J., 2002.

[18] Meyer, Y. *Wavelets and operators*. Cambridge Studies in Advanced Mathematics, 37. Cambridge University Press, Cambridge, 1992.

[19] Peyré, G. Bandelets toolbox. Available online at: `http://www.mathworks.com/matlabcentral/fileexchange/loadAuthor.do?objectType=author&objectId=1094083`

[20] Peyré, G.; Mallat, S. Surface compression with geometric bandelets. *ACM Transactions on Graphics (TOG)* **24** (2005), no. 3, 601–608.

[21] Shukla, R.; Dragotti, P. L.; Do, M. N.; Vetterli, M. Rate-distortion optimized tree-structured compression algorithms for piecewise polynomial images. *IEEE Trans. Image Process*. **14** (2005), no. 3, 343–359.

[22] Wakin, M. B.; Romberg, J. K.; Choi, H.; Baraniuk, R. G. Wavelet-domain approximation and compression of piecewise smooth images. *IEEE Trans. Image Process*. **15** (2006), no. 5, 1071–1087.

GABRIEL PEYRÉ
CEREMADE
Université de Paris - Dauphine
Place du Maréchal De Lattre
De Tassigny
75775 Paris Cedex 16
FRANCE
E-mail: `gabriel.peyre@ceremade.dauphine.fr`

STÉPHANE MALLAT
CMAP
École Polytechnique
91128 Palaiseau Cedex
FRANCE
E-mail: `Stephane.Mallat@polytechnique.fr`

Received December 2006.
Revised May 2007.