ADAPTIVE DIRECTIONAL IMAGE COMPRESSION WITH ORIENTED WAVELETS

François G. Meyer, and Ronald R. Coifman

Department of Mathematics, Yale University, New Haven CT, 06520, USA.

ABSTRACT

We construct a new adaptive basis that provide precise frequency localization and good spatial localization. We develop a compression algorithm that exploits this basis to obtain the most economical representation of an image in terms of textured patterns with different orientations, frequencies, sizes, and positions. The technique directly works in the Fourier domain and has potential applications for compression of richly textured images.

1. INTRODUCTION

Edges and textures in an image can exist at all possible locations, orientations, and scales. The ability to efficiently analyze and describe textured patterns is thus of fundamental importance for image analysis and image compression. Wavelets provide an octave based decomposition of the Fourier plane with a poor angular resolution. Wavelet packets make it possible to adaptively construct an optimal tiling of the Fourier plane, and they have been used for image compression [1]. However the tensor product of two real valued wavelet packets is always associated with four symmetric peaks in the frequency plane. It is therefore not possible to selectively localize a unique frequency. Directionally oriented filter banks [2] have been used for image compression and image analysis. They do not allow however an arbitrary partitioning of the Fourier plane. In order to obtain a better angular resolution than the standard wavelet packets we expand the Fourier plane into windowed Fourier bases [3]. The method results in an expansion of the image into a set of *brushlets*. A brushlet is a function reasonably well localized with only one peak in frequency. Furthermore, the brushlet is a complex valued function with a phase. The phase of the bi-dimensional brushlet provides valuable information about the orientation of the brushlet. We can adaptively select the size and locations of the brushlets in order to obtain the most concise and precise representation of an image in terms of oriented textures with all possible directions, frequencies, and locations. We demonstrate that this new basis can be used for directional image analysis and to efficiently compress richly textured images.

2. BRUSHLET BASIS

We are interested in a local "time frequency" analysis of the Fourier transform of a signal. We explain here how to perform a time-frequency analysis with windowed Fourier bases. In order to analyze the local frequency content of a signal, we use a smooth window function to localize the segment of interest. Then a local Fourier analysis is performed inside each interval. We want to construct orthonormal bases with good "time-frequency" localization. We know from the Ballian-Low theorem [4] that we cannot use windowed exponentials of the form

$$g_{n,m}(x) = e^{im\omega_0 x} g(x - nt_0) \tag{1}$$

In order to circumvent the obstacle raised by the Ballian-Low theorem various Wilson bases have been constructed that use sines and cosines rather than exponential [5, 6]. We are interested in using exponentials, because the phase of the exponential will provide information about the direction of the pattern when describing images in two dimensions. Therefore we will use the smooth localized orthonormal exponential bases defined in [3]. These functions are exponentials with good localization in both position and Fourier space. We consider a cover $\mathbb{R} = \bigcup_{n=-\infty}^{n=+\infty} [a_n, a_{n+1}[$. We write $l_n = a_{n+1} - a_n$, and $c_n = (a_n + a_{n+1})/2$. Around each a_n we define a neighborhood of radius ε. Let r be a ramp function such that

$$r(t) = \begin{array}{ll} 0 & \text{if} \quad t \leq -1 \\ 1 & \text{if} \quad t \geq 1 \end{array} \tag{2}$$

and $r^2(t) + r^2(-t) = 1, \qquad \forall t \in \mathbb{R}$ We introduce the steepness factor $s = \frac{\varepsilon}{l_n}$. Let v_s be the bump function supported on $[-s, s]$

$$v_s(t) = r(\frac{t}{s}) r(-\frac{t}{s}) \tag{3}$$

Let b_s be the window function supported on $[-\frac{1}{2} - s, \frac{1}{2} + s]$

$$\begin{array}{rcll} b_s(t) & = & r^2(\frac{1}{s}\{t + \frac{1}{2}\}) & \text{if} \quad t \in [-\frac{1}{2} - s, -\frac{1}{2} + s] \\ & = & 1 & \text{if} \quad t \in [-\frac{1}{2} + s, \frac{1}{2} - s] \\ & = & r^2(\frac{1}{s}\{\frac{1}{2} - t\}) & \text{if} \quad t \in [\frac{1}{2} - s, \frac{1}{2} + s] \end{array} \tag{4}$$

We consider the collection of exponential functions

$$e_{j,n}(x) = \frac{1}{\sqrt{l_n}} e^{-2i\pi j(\frac{x - a_n}{l_n})}.$$

We can construct a basis of smooth localized orthonormal exponential functions $u_{j,n}$, where each $u_{j,n}$ is supported on $[a_n - \varepsilon, a_{n+1} + \varepsilon]$ and is given by [3]

$$\begin{array}{rcl} u_{j,n}(x) & = & b_s(\frac{x - c_n}{l_n}) e_{j,n}(x) \\ & + & v_s(\frac{x - a_n}{l_n}) e_{j,n}(2a_n - x) \\ & - & v_s(\frac{x - a_{n+1}}{l_n}) e_{j,n}(2a_{n+1} - x) \end{array} \tag{5}$$

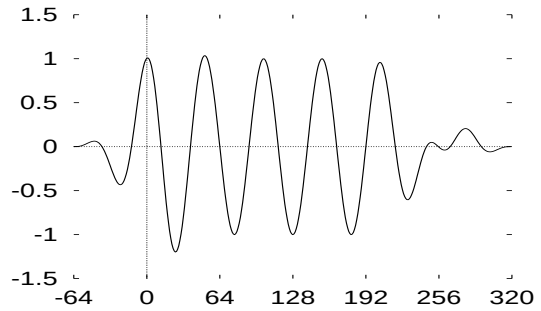

Figure 1: Real part of the windowed exponential function $u_{j,n}$, with $a_n = 0$, $a_{n+1} = 256$, and $\varepsilon = 64$.

Figure 1 shows the real part of the function $u_{j,n}$ with $a_n = 0$, $a_{n+1} = 256$, $l_n = 256$, and $j = 5$.

Theorem 1 [3] *The collection* $\{u_{j,n}\ j, n \in \mathbb{Z}\}$ *is an orthonormal basis for* $L^2(\mathbb{R})$.

3. ORTHONORMAL BRUSHLET BASES

The orthonormal windowed Fourier bases can be used to perform a time frequency analysis of an image. For a number of applications, it is more relevant to perform a time frequency analysis of the Fourier transform of the signal. This analysis corresponds to finding all the patterns in the image with a given orientation, and frequency. In order to decompose the image into different oriented patterns we expand the Fourier transform into windowed Fourier bases. Our construction permits to build a new set of reasonably well localized functions with only one peak in frequency.

3.1. One dimensional case

Let $w_{j,n}$ the inverse Fourier transform of $u_{j,n}$. Since the Fourier transform is a unitary operator, we have

Lemma 1 $\{w_{j,n}, j, n \in \mathbb{Z}\}$ *is an orthonormal basis for* $L^2(\mathbb{R})$.

We call $\{w_{j,n}\}$ the orthonormal brushlet basis. From (5) we have

$$\begin{aligned} w_{j,n}(x) &= \sqrt{l_n} e^{2i\pi a_n x} e^{i\pi l_n x} \left\{ (-1)^j \widehat{b}_s(l_n x - j) \right. \\ &- \left. 2i\, sin(\pi l_n\, x) \widehat{v_s}(l_n x + j) \right\} \qquad (6) \end{aligned}$$

We note in (6) that l_n appears as a scaling factor of the analysis. $w_{j,n}$ has an expression similar to a wavelet, however as opposed to a real valued wavelet, $w_{j,n}$ is a complex valued function with a phase. The phase encodes the frequency and the orientation of the brushlet pattern in the two-dimensional case. b_s and v_s are even real valued functions, thus $\hat{b}_s$ and $\hat{v}_s$ are also even real valued functions. The function $w_{j,n}$ is composed of two terms. Since $|\hat{v}_s(x)| \leq s$, the second term can be made as small as possible. However, when s tends to zero the first term is not localized anymore. There is a tradeoff between the localization of $\widehat{b}_s$ and the magnitude of the second term. We choose s such that the brushlet function is mainly localized around j/l_n. Figure 2 shows the graph of the real part of $w_{j,n}$ for a particular choice of r.

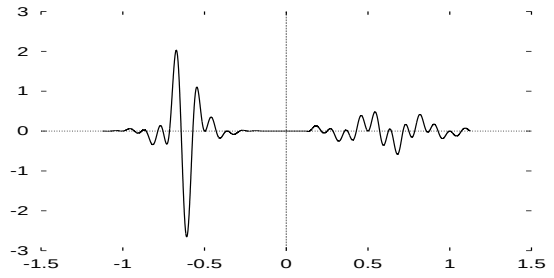

Figure 2: Orthonormal brushlet $w_{j,m}$ with $\varepsilon = 4$, $j/l_n = 5/8$, and $c_n = 7$. The window of the orthonormal brushlet has many oscillations.

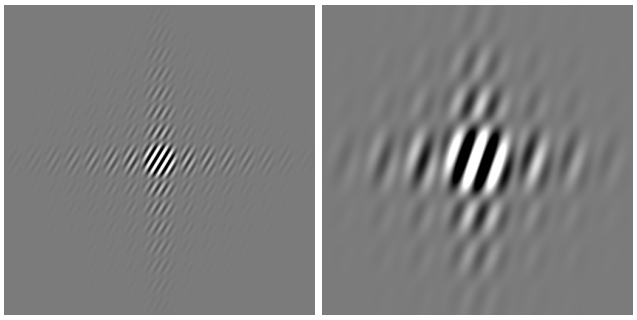

Figure 3: Two dimensional brushlet basis functions $\{w_{j,m} \otimes w_{k,n}\}$ A good spatial resolution corresponds to a $\hat{b}$ with a small support, and is thus associated with a poor frequency resolution as shown on the left. A good frequency resolution corresponds to a b with a small support, and is thus associated with a poor spatial resolution as shown on the right.

3.2. Two-dimensional case

In the two-dimensional case we define two partitions of $\mathbb{R}$, $\bigcup_{n=-\infty}^{n=+\infty} [a_m, a_{m+1}[$, and $\bigcup_{m=-\infty}^{m=+\infty} [b_n, b_{n+1}[$. We write $h_m = a_{m+1} - a_m$, and $l_n = b_{n+1} - b_n$. We then consider the tiling obtained by the lattice cubes $[a_m, a_{m+1}[\otimes [b_n, b_{n+1}[$. We consider the separable tensor products of bases $w_{j,m}$, and $w_{k,n}$. We have

Lemma 2 *The sequence* $w_{j,m} \otimes w_{k,n}$ *is an orthonormal basis for* $L^2(\mathbb{R}^2)$.

3.3. Adaptive tiling of the Fourier plane

The Fourier transform $\hat{f}$ of the image f is computed using an FFT. $\hat{f}$ is hermitian-symmetric, therefore we only retain the upper half of the Fourier plane $\{(\nu, \xi), \xi \geq 0\}$ for coding. As explained in [3] we can adaptively select the size and location of the windows $[a_m, a_{m+1}[\otimes [b_n, b_{n+1}[$ with the best basis algorithm. We divide the Fourier plane into four sub-squares, and we consider the brushlet basis associated

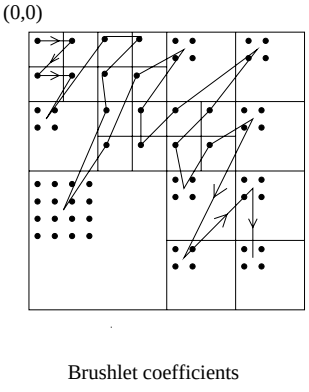

Figure 4: We order all the brushlet coefficients associated with the same region in the spatial domain using a zigzag pattern in the Fourier plane.

with this tiling. Instead of calculating the inner product of $\hat{f}$ with $u_{j,m} \otimes u_{k,n}$ we fold the image around the horizontal and vertical lines associated with the tiling as explained in [3]. We then calculate inside each block the 2-D FFT of the folded block, and obtain the brushlet coefficients. We then further decompose each square into four sub-squares, and consider the brushlet basis associated with this finer tiling. By applying this decomposition recursively we obtain a homogeneous quadtree-structured decomposition. For each subblock, or node of the quadtree, we calculate the set of coefficients associated with the brushlets living on the subblock. We associate a cost for each node of the tree, based on the set of coefficients. We then find an optimal segmentation of the Fourier space, using a divide and conquer algorithm.

3.4. Zig-zag scanning and entropy coding

The brushlet coefficients are quantized with uniform quantizers. In order to exploit the correlation between brushlet coefficients in different subbands, we order all the brushlet coefficients associated with the same region in the spatial domain. The coefficients are ordered by increasing frequency order by scanning the quadrant with a zig-zag pattern as shown in Fig. 4. Since the magnitude of the terms in a zig-zag sequence decreases with an exponential decay, we encode a terminating symbol after the last non-zero coefficient to indicate that the remaining coefficients are zeros. This represents a zero-tree like extension of the algorithm proposed in [7]. After zig-zag ordering, the coefficients are then coded using variable length coding. The alphabet that describes the variable length encoding is entropy coded with an adaptive arithmetic coder. The first term of a zig-zag scan corresponds to a DC coefficient. The DC coefficients of adjacent spatial locations are still correlated, and are therefore differentially encoded. We have implemented the coder and decoder, and an actual bit stream was created for each experiment.

	Barbara
Compression	PSNR (dB)
8:1	35.30
16:1	30.86
32:1	25.15
67:1	23.47
82:1	23.08
135:1	22.06
334:1	20.31

Table 1: Coding results for 8bpp. 512x512 Barbara

4. EXPERIMENTS

We present the results of the algorithm using two test images that are difficult to compress: 512x512 "Barbara", and 512x512 "Mandrill". The performance of the algorithm are summarized in Tables 1 and 2. Figure 5 shows the Mandrill image coded with a compression ratios of 100:1, with a PSNR = 21.02dB, and the optimal tiling of the upper half of the Fourier plane. We note that the segmentation is not symmetric, reflecting some significant oriented textures in the image. We also note that even at a compression ratio of 100:1 the Mandrill still keeps its high frequency features such as the whiskers. In order to emphasize the performance of the algorithm, we have used the EZW algorithm of Shapiro [7] to compress the image a fingerprint image. Figure 7 shows the result of the compression with EZW at a compression ration of 120:1, with a PSNR = 17.42dB. Figure 6 shows the result of the compression with our algorithm at a compression ration of 120:1, with a PSNR = 19.90dB. The associated tiling of the Fourier plane is also shown. We note that most of the details have been smeared by EZW, while our algorithm keeps the structure of the fingerprint with a much better PSNR.

5. REFERENCES

[1] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Trans. on Image Processing*, pages pp 160–175, April 1993.

[2] R.H. Bamberger and M.J.T. Smith. A filter bank for the directional decomposition of images: theory and design. *IEEE Trans. on Signal Processing*, pages 882–893, April 1992.

[3] M.V. Wickerhauser. *Adapted Wavelet Analysis from Theory to Software*. A.K. Peters, 1995.

[4] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.

[5] H. Malvar. Lapped transforms for efficient transform/subband coding. *IEEE Trans. Acoust. Sign. Speech Process.*, Vol 38:969–978, 1990.

[6] R.R. Coifman and Y. Meyer. Remarques sur l'analyse de fourier à fenêtre. *C.R. Acad. Sci. Paris I*, pages pp. 259–261, 1991.

[7] J.M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. on Signal Processing*, pages 3445–3462, Dec. 1993.

	Mandrill
Compression	PSNR (dB)
8:1	28.28
15:1	25.34
30:1	23.14
58:1	21.76
81:1	21.26
121:1	20.71
206:1	20.19

Table 2: Coding results for 8bpp. 512x512 Mandrill

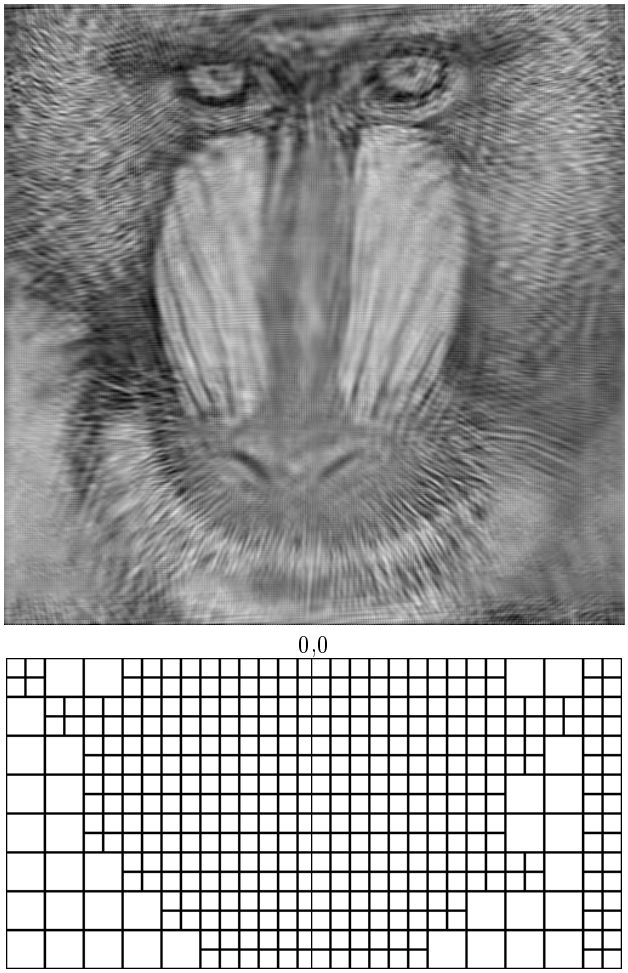

Figure 5: Top: Mandrill, compression 100:1, PSNR = 21.02dB. Bottom: optimal tiling of the upper half of the Fourier plane; the horizontal axis point toward the right, and the vertical axis points upwards.

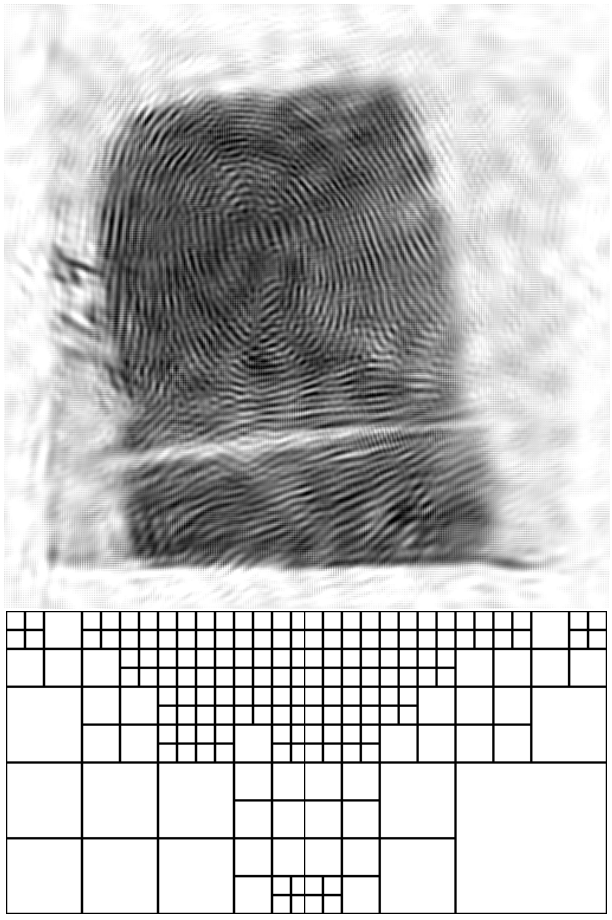

Figure 6: Compression 120:1 with brushlets, PSNR = 19.90dB ; tiling of the upper half of the Fourier plane.

Figure 7: Fingerprint, compression 120:1 with EZW, PSNR = 17.42dB.