IMAGE COMPRESSION WITH ADAPTIVE LOCAL COSINES : A COMPARATIVE STUDY

François G. Meyer

Electrical Engineering,University of Colorado at Boulder, francois.meyer@colorado.edu

ABSTRACT

The goal of this work is twofold: First, we demonstrate that an advantage can be gained by using local cosine bases to encode images that contain periodic textures. We designed a coder that outperforms a wavelet coder [1] on a large number of images. This new coder constitutes the first contribution of the paper. Second, we used our coder to compare the performance of several optimized bells [2, 3, 4] in terms of rate-distortion for a large collection of images.

1. INTRODUCTION

Many classes of images have very diffuse representations in a standard wavelet basis. Images with oscillatory patterns are examples of non wavelet-friendly signals. In principle, the description of oriented periodic patterns can be best performed with local windowed Fourier or cosine bases. A lot of work has been expended in the design of orthonormal and biorthogonal bases that exploit a combination of overlapping bells and a Fourier-like transform [2, 3, 4]. While several variations exist, one focuses in this paper on the following family of functions :

$$\psi_{n,k}(x) = b_n(x) \cos(j + \frac{1}{2})\pi x$$

If one accepts to use Riesz biorthogonal bases instead of orthonormal bases, then one has much freedom to choose the bell b_n. Several "optimal bells" were designed in [2, 4, 5]. The optimality of these bells has been claimed from a strictly theoretical point of view. The goal of this work is twofold. First, we demonstrate that an advantage can be gained by using local cosine bases over wavelets to encode images that contain periodic textures. We designed a coder that outperforms a wavelet coder [1] on a large number of images. This new coder constitutes the first contribution of the paper. Second, we used our coder to compare the performance of several bells [2, 3, 4] in terms of rate-distortion for a large collection of images.

2. LOCAL COSINE TRANSFORMS (LCT)

We review here the construction of smooth Localized Cosine Transforms (LCT) [2, 6, 7] in one dimension. Let $\bigcup_{n=-\infty}^{n=+\infty}[a_n, a_{n+1}[$ be a cover of $\mathbb{R}$. We define a neighborhood $[a_n - \alpha_n, a_n + \beta_n]$ around each point a_n ; where $\alpha_n > 0$, $\beta_n > 0$, and $a_n + \beta_n < a_{n+1} - \alpha_{n+1}$ (see Fig. 1). Let b_n be a sequence of bell functions such that:

$$\forall x \in [a_n - \alpha_n, a_n + \beta_n],$$
$$b_n(x)\, b_{n-1}(2a_n - x) + b_n(2a_n - x)\, b_{n-1}(x) \neq 0$$
$$\forall x \in [a_n + \beta_n, a_{n+1} - \alpha_{n+1}] b_n(x) \neq 0$$

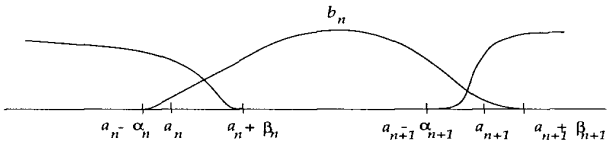

Fig. 1. The bell b_n lives over the interval $[a_n - \alpha_n, a_{n+1} + \beta_{n+1}]$.

Let

$$\theta_n(x) = \frac{1}{b_n(x)\, b_{n-1}(2a_n - x) + b_n(2a_n - x)\, b_{n-1}(x)}$$

the dual bell $\tilde{b}_n$ is defined as follows:

$$\tilde{b}_n(x) = \begin{cases} \theta_n(x) b_{n-1}(2a_n - x) & \text{if } a_n - \alpha_n \leq x \leq a_n + \beta_n \\ \dfrac{1}{b_n(x)} & \text{if } a_n + \beta_n \leq x \leq a_{n+1} - \alpha_{n+1} \\ \theta_{n+1}(x) b_{n+1}(2a_{n+1} - x) & \text{if } a_{n+1} - \alpha_{n+1} \leq x \leq a_{n+1} + \beta_{n+1} \\ 0 & \text{otherwise} \end{cases}$$

Let $c_{n,k}(x)$ be the family of basis functions of the DCT-IV:

$$c_{n,k}(x) = \sqrt{\frac{2}{a_{n+1} - a_n}} \cos\left((k + 1/2)\frac{x - a_n}{a_{n+1} - a_n}\right)$$

The local cosine basis functions $\psi_{n,k}$ and their dual $\tilde{\psi}_{n,k}$ are defined by

$$\psi_{n,k} = b_n(x)\, c_{n,k}(x) \qquad \tilde{\psi}_{n,k} = \tilde{b}_n(x)\, c_{n,k}(x).$$

$\psi_{n,k}$ and $\tilde{\psi}_{n,k}$ are Riesz biorthogonal bases, and one can use either one of the two bases to compute the coefficients of a function, and use the dual basis to reconstruct the function.

0-7803-6725-1/01/$10.00 ©2001 IEEE

3. CHOICE OF THE BELL FUNCTION

We first present a standard orthonormal bell, and then we describe several "optimized" bells that were discovered recently. These bells require that $a_n^- = a_{n-1}^+$. In the case where the bell is symmetric this condition implies that $a_n^+ = (a_n + a_{n+1})/2$, and that all intervals need to have the same size. For the rest of the paper, one will assume that

$$\forall n, \quad \begin{array}{c} a_{n+1} - a_n = l \\ a_{n+1}^- = a_n^+ = (a_n + a_{n+1})/2 \end{array} \tag{1}$$

All bells b_n are obtained from a prototype bell b by dilation by l, and translation by a_n. The prototype bell b is defined on $[-1/2, 3/2]$.

3.1. Orthonormal bell

Wickerhauser [3] defines for $x \in [-1/2, 1/2]$

$$b^M(x) = \sin \frac{\pi}{2}(x_M + 1/2) \tag{2}$$

with

$$x_0 = x \qquad x_j = \frac{1}{2}\sin(\pi x_{j-1}). \tag{3}$$

The bell is symmetric, and if $x \in [-1/2, 1/2]$ we define $b^M(x) = b^M(1-x)$. This bell gives rise to an orthonormal basis ($b_n = \tilde{b}_n$).

3.2. Optimized bell of Matviyenko

Gregory Matviyenko designed a family of bells that allow to minimize the number of coefficients $x_{0,j}$ needed to reconstruct constant functions for a given error ε. Matviyenko solve the following optimization problem

$$\min \sum_{n=M}^{\infty} | < 1, \psi_{n,k} > |^2 \tag{4}$$

under the constrain :

$$b(x) + b(-x) = 1 \quad \text{for all } 0 \le x < 1/2. \tag{5}$$

The optimized bell $b^M(x)$ is given by :

$$\frac{1}{2}(1 + \sum_{n=0}^{M-1} g_n \sin(n+1/2)\pi x) \text{ if } -0.5 \le x \le 0.5$$

$$\frac{1}{2}(1 + \sum_{n=0}^{M-1} (-1)^n g_n \cos(n+1/2)\pi x) \text{ if } 0.5 \le x \le 1.5$$

$$0 \quad \text{otherwise}$$

M controls the steepness of the bell. The g_n are calculated numerically in [2]. The Riesz bounds are 1 and 2 for all M.

3.3. Modulated Lapped Biorthogonal Transform (MLBT)

Henrique Malvar proposed [4] the following bell to improve the frequency selectivity of the basis functions :

$$b(x) = \frac{1 - \cos[(\frac{x+1}{4})^\alpha \pi] + \beta}{2 + \beta}.$$

The Riesz bounds are also excellent : 1 and 1.458. Fig.2 shows the frequency responses of the basis function $\psi_{n,16}$ for the three afore-mentioned bells. Malvar's bell has the most narrow main lobe, but Matviyenko's bell ($M = 3$) has the smallest side lobes, and the best stopband attenuation.

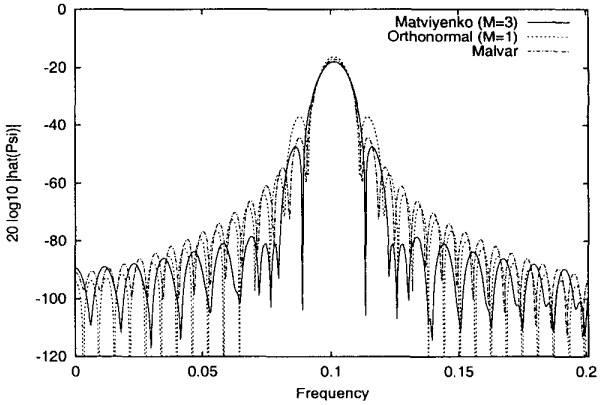

Fig. 2. Frequency response of $\psi_{n,16}$ for 3 different bells (512 samples)

4. DETAILS OF THE CODER

We use the best basis algorithm [8]to adaptively select the size and location of the 2-dimensional windows (blocks) according to a cost function $\mathcal{M}$. Our cost function $\mathcal{M}$ returns an estimate of the actual rate achieved by each block in the quadtree segmentation. It is composed of two complementary terms: (i) the cost of coding the sign and the magnitude of the non zero output levels of the scalar quantizer, and (ii) the cost of coding the locations of the non zero output levels (significance map). The distribution of the cosine coefficients is approximated with a Laplacian distribution, and we use a particularly efficient near optimal scalar quantizer [9]. In order to exploit large scale correlations that may exist between adjacent blocks, we gather together coefficients from all the blocks that have a similar two-dimensional frequencies. We start with the smallest frequency, and continue until the highest frequency. We implemented the Fast Local Cosine Transform (FLCT) coder and decoder, and an actual bit stream is generated by the coder. The FLCT code that was used for the experiments is available from our web site.

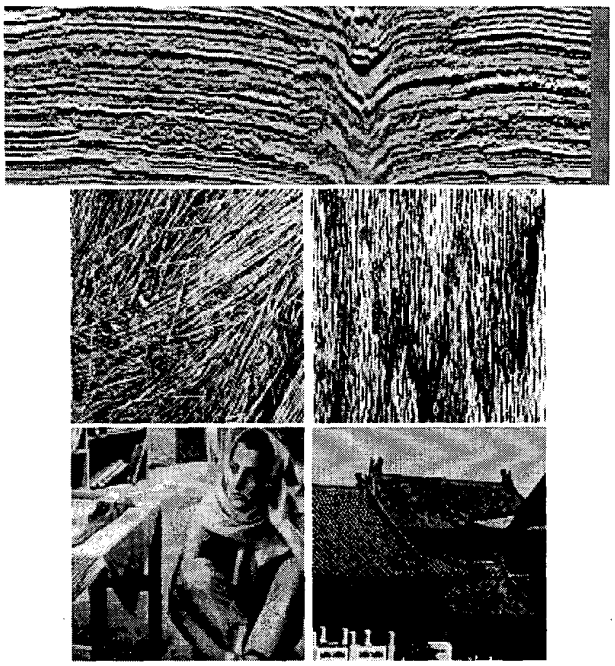

Fig. 3. Top (1) seismic image, 2nd row from left : (2) straw and (3) wood grain; last row, from left (4) Barbara, and (5) Roofs.

5. EXPERIMENTS

The following images have been used for the experiments (see Fig 3) :

1. Seismic image, 32 bpp 192 x 640. This image contains rapid oscillations, as well as discontinuities.

2. straw, 8 bpp 512 x 512. This image is part of Brodatz's book. It is extremely difficult to code : the intensity, as well as the orientation of the patterns are varying very rapidly.

3. wood grain, 8 bpp 512 x 512. This image is also part of Brodatz's book.

4. roofs, 8 bpp 512 x 512. This image is part of the MIT VisTex database. It contains a mixture of periodic texture (roofs), as well as smooth regions (façades and sky).

5. Barbara, 8 bpp 512 x 512.

We first report the results of experiments conducted with a fixed window size : 32×64 for the seismic image, and 32×32 for the straw, woodgrain, and roofs images. Because different bells can result in different best bases, one needs to separate the effect of the bell from the effect of the basis. In addition to comparing the three bells discussed in the previous section, we also included the results obtained with the abrupt cutoff bell : the characteristic function of the interval $[a_n, a_{n+1}]$, the other parts of the algorithm being equal. Finally, we also compared our coder to one of the best wavelet coder that was available to us: the SPIHT wavelet coder of Amir Said and William A. Pearlman [1].

Rate (bpp)	0.40	0.50	0.67	1.00	2.00
Orthonormal	**19.09**	**19.90**	**21.02**	**22.94**	**27.72**
Matviyenko	18.93	19.76	20.90	22.79	27.54
MLBT	19.06	19.87	21.00	22.91	27.68
No bell	18.05	18.88	20.05	22.05	26.97

Table 1. Seismic data : comparison of the bells

Seismic All bells performed equally well (see table 1), and resulted in a 1dB improvement over an abrupt cutoff bell (no bell). Interestingly, the optimized bell of Matviyenko did not result in the gain that could be expected with an image that contains many oscillatory patterns.

Rate (bpp)	0.125	0.25	0.50	0.75	1.00
Orthonormal	18.52	21.07	**23.94**	**25.78**	**27.24**
Matviyenko	**18.55**	21.05	23.87	25.69	27.12
MLBT	18.54	**21.08**	23.93	25.73	27.18
No bell	17.07	19.30	21.94	23.72	25.17
SPIHT	16.49	18.57	21.49	23.71	25.78

Table 2. Wood grain : comparison of the bells.

Wood grain All bells performed equally well (see table 2), and resulted in a significant improvement of 1.5 to 2dB over an abrupt cutoff bell. The LCT significantly outperformed SPIHT by 1.5 to 2.5 dB. In fact, our coder performs as well as SPIHT even without a smooth bell.

Rate (bpp)	0.125	0.25	0.50	0.75	1.00
Orthonormal	24.84	27.86	31.88	34.79	37.08
Matviyenko	24.90	27.90	31.91	34.76	36.97
MLBT	**24.90**	**27.91**	**31.96**	**34.82**	**37.18**
No bell	21.24	23.66	27.19	29.59	31.64
SPIHT	23.77	27.02	31.18	34.18	36.68

Table 3. Roofs : comparison of the bells.

Roof All the bells performed equally well (see table 3). The consequence of using an abrupt cutoff bell was a dramatic drop in the PSNR of 3.6 to 5.5 dB. Again, our coder outperforms SPIHT.

Rate (bpp)	0.125	0.25	0.50	0.75	1.00
Orthonormal	**14.38**	**16.28**	**18.85**	**20.75**	**22.34**
Matviyenko	14.30	16.18	18.75	20.64	22.21
MLBT	14.34	16.23	18.80	20.70	22.30
No bell	13.96	15.73	18.08	19.83	21.29
SPIHT	14.16	16.01	18.34	20.33	21.87

Table 4. Straw : comparison of the bells.

Straw As expected, the PSNR degrades very rapidly as the rate decreases. All bells performed equally well (see table 4). Even for this image, using an abrupt cutoff bell has the effect of a drop of 1dB in the PSNR. Our coder again outperformed SPIHT.

The second set of experiments was designed to include the effect of the best basis selection. We report here the results obtained with the image Barbara (see table 5). All the bells performed equally well, and resulted in a gain of 1 to 3 dB over an abrupt cutoff bell. Our coder outperformed SPIHT by 0.7 to 1dB. To further illustrate the gain obtained from using local cosine bases instead of the DCT, we compare in table 5 several DCT-based coders [10, 11, 12] that use very elaborate coding strategies. Despite its simple design our coder outperforms all these more complicated coders.

Rate (bpp)	0.125	0.25	0.50	0.75	1.00
Orthonormal	25.75	28.71	32.43	35.00	37.04
Matviyenko	25.76	28.82	**32.61**	35.04	37.03
MLBT	**25.87**	**28.87**	32.58	**35.10**	**37.11**
No bell	24.78	27.01	29.71	32.19	33.89
SPIHT	24.86	27.58	31.39	34.25	36.41
JPEG [10]		26.7	30.6	33.6	35.9
STQ [12]		26.5	30.4	32.4	34.9
EZDCT [11]	24.51	27.26	31.10	33.93	36.21

Table 5. Barbara : comparison of the bells.

Fig. 4. Barbara @ 0.25bpp (compression = 32), PSNR = 28.87dB

6. REFERENCES

[1] Amir Said and William A. Pearlman, "A new fast and efficient image codec based on set partioning in hierarchical trees," *IEEE Trans.on Circ.& Sys. for Video Tech.*, vol. 6, pp. 243–250, 1996.

[2] G. Matviyenko, "Optimized local trigonometric bases," *Applied and Comput. Harmo. Anal.*, vol. 3, pp. 301–323, 1996.

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

[4] H.S. Malvar, "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts," *IEEE Trans. on Signal Process.*, 46(4), pp. 1043–1053, 1998.

[5] K. Bittner, "Error estimates and reproduction of polynomials for biorthogonal local trigonometric bases," *Applied and Comput. Harmo. Anal.*, 6, pp. 75–102, 1999.

[6] P. Auscher, G. Weiss, and M.V. Wickerhauser, "Local sine and cosine bases of Coifman and Meyer," in *Wavelets-A Tutorial*. 1992, Academic Press.

[7] C.K. Chui and X. Shi, "Characterization of biorthogonal cosine wavelets," *J. Fourier Anal. Appl.*, 3(5), pp. 560–575, 1997.

[8] R.R. Coifman and M.V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. on Info. Theory*, 38(2), pp. 713–718, 1992.

[9] G.J. Sullivan, "Efficient scalar quantization of exponential and Laplacian random variables," *IEEE Trans. on Info. Theory*, 42(5), pp. 1365–1374, 1996.

[10] M. Crouse and K. Ramchandran, "Joint thresholding and quantizer selection for transform image coding", *IEEE Trans. on Image Process.*, 6, pp. 285–97, 1997.

[11] Z. Xiong, K. Ramchandran, M.T. Orchard, and Y. Zhang, "A comparative study of DCT- and wavelet-based image coding," *IEEE Trans. Circ. & Sys. for Video Tech.*, 9(5), pp. 692–695, 1999.

[12] G. Davis and S. Chawla, "Image coding using optimized significance tree," in *IEEE Data Compression Conference -DCC'97*, 1997, pp. 387–396.