IMAGE COMPRESSION WITH BEST LOCAL COSINE BASES

Y. Huang and I. Pollak

Purdue University
School of ECE
West Lafayette, IN 47907

ABSTRACT

We apply the framework of multitree dictionaries introduced in [7, 8], to design a novel image coder based on lapped orthogonal bases of local cosines. We show that our image coder outperforms both the quadtree-based approach of [13] and SPIHT [16] for textured images.

1. INTRODUCTION

A number of research efforts have recently concentrated on developing adaptive algorithms for representing and approximating signals in overcomplete dictionaries. This paper addresses the *best basis problem*—or, more generally, the *best representation problem*: given a signal, a dictionary of representations, and an additive cost function, the aim is to select the representation from the dictionary which minimizes the cost for the given signal. This paradigm has been successfully used for problems in compression [7, 10, 13, 15], estimation [3, 4, 9, 14], and time-frequency (or space-frequency) analysis [5, 6, 8].

The original papers on best basis search [2] considered the wavelet packet bases and bases of local cosines on dyadic intervals. In each of these two cases, all the bases in the dictionary can be organized using a single tree: a dyadic tree in 1-D and a quadtree in 2-D. This organization was exploited in [2] to devise a fast recursive tree pruning algorithm to find the best basis for any additive cost function.

Since then, a number of efforts [1,4,5,10] have sought to lift the restrictions that a fixed dyadic/quadtree structure imposes on the underlying dictionary. In the present paper, we build on one such effort reported in [7, 8] where we developed a new framework of multitree dictionaries. We use this framework to design a novel image coder based on lapped orthogonal bases of local cosines. We apply our image coder to the compression of textured images and show that our new coder outperforms both the quadtree-based coder of [13] and SPIHT [16].

This work was supported in part by a National Science Foundation (NSF) CAREER award CCR-0093105 and an NSF grant IIS-0329156.

2. BACKGROUND

2.1. Best Basis Search Problem.

The general best basis search problem is formulated, for example, in [2, 11]. We consider a dictionary $\mathcal{D}$ that is the union of orthonormal bases for $\mathbb{C}^N$, $\mathcal{D} = \bigcup_{\lambda\in\Lambda} B^\lambda$, where each basis B^λ is a family of N vectors, $B^\lambda = \{g_m^\lambda\}_{1\le m\le N}$. The cost of representing a signal f in a basis B^λ is defined as

$$C(f, B^\lambda) = \sum_{m=1}^{N} \Phi\left(\frac{|\langle f, g_m^\lambda\rangle|^2}{||f||^2}\right), \qquad (1)$$

where Φ is application dependent. Any basis which achieves the minimum of the cost $C(f, B^\lambda)$ over all the bases in the dictionary, is called the best basis.

2.2. Local Cosines.

A local cosine family [12] is defined using cosine functions multiplied by overlapping smooth windows. For each discrete interval $[u, v-1] \subset \mathbb{Z}$ of length $a = v-u$, we define a window function $\beta_{u,v}$ which gradually ramps up from zero to one around u and goes down from one to zero around $v-1$:

$$\beta_{u,v}(t) = \begin{cases} r\left(\frac{t-(u-1/2)}{\eta}\right) & \text{if } u-\frac{1}{2}-\eta \le t < u-\frac{1}{2}+\eta \\ 1 & \text{if } u-\frac{1}{2}+\eta \le 1 < v-\frac{1}{2}-\eta \\ r\left(\frac{(v-1/2)-t}{\eta}\right) & \text{if } v-\frac{1}{2}-\eta \le t \le v-\frac{1}{2}+\eta \\ 0 & \text{otherwise,} \end{cases}$$

where the parameter $\eta \in \mathbb{R}$ controls how fast the window tapers off, and r is a monotonically increasing profile function. We always set $u - v \ge 2\eta$. Following [12], we define the discrete local cosine family $\mathcal{B}_{u,v}$ as follows:

$$\mathcal{B}_{u,v} = \left\{ \frac{\beta_{u,v}(n)\sqrt{2}}{\sqrt{v-u}} \cos\frac{\pi(k+\frac{1}{2})(n-(u-\frac{1}{2}))}{v-u} \right\}_{k=0}^{v-u-1},$$

where $n \in \mathbb{Z}$ is a discrete parameter. It can be shown [12] that this set of signals is orthonormal. A 2-D orthonormal local cosine family $\mathcal{B}_P$ can similarly be defined for any rectangular domain P.

3. MULTITREE DICTIONARIES AND IMAGE COMPRESSION.

3.1. A Multitree Local Cosine Dictionary.

We consider all images supported on a discrete rectangular domain $Q \subset \mathbb{Z}^2$. Any partition $\lambda = \{P_1, P_2, \ldots, P_d\}$ of this domain into rectangular tiles $P_1, P_2, \ldots, P_d$ where each tile is larger than $2\eta \times 2\eta$, induces the 2-D local cosine basis $B^\lambda = \mathcal{B}_{P_1} \cup \mathcal{B}_{P_2} \cup \ldots \cup \mathcal{B}_{P_d}$. We restrict our choice of tilings to those which can be obtained by recursively splitting rectangles into pairs of subrectangles. Such a splitting process can be represented as a binary tree whose root corresponds to the entire image and whose every node corresponds to a unique rectangular region of the image. The resulting local cosine dictionary which consists of all the bases corresponding to all such tilings, is easily shown to be a multitree dictionary [7]. The problem of finding the global minimum of the cost $C(f, B^\lambda)$ (Eq. (1)) can therefore be efficiently solved using our search algorithm presented in [7].

3.2. A New Image Coder.

To obtain our new image coder, we fuse our best local cosine basis search with several aspects of the compression strategy in [15]. The input image is partitioned into square blocks; for each block, the optimal tiling λ^* is found, and every tile is encoded. When looking for the best tiling, we optimize with respect to the rate-distortion cost [15] $D^\lambda + \gamma R^\lambda$, where R^λ is the number of bits it takes to encode the image if we use the tiling λ, D^λ is the total distortion, and γ is a parameter. In order to use the optimal search algorithm of [7], we assume that the cost $D^\lambda + \gamma R^\lambda$ has the form (1): $D^\lambda + \gamma R^\lambda = \sum_{P \in \lambda} (D(P) + \lambda R(P))$, where $D(P)$ and $R(P)$ are the distortion and rate, respectively, for the tile P, and the summation is performed over all the tiles in the tiling. For each tile, we follow a procedure outlined in [7] which finds the local cosine coefficients, quantizes, and entropy-codes the coefficients, and also codes the structure of the optimal tree. If several different quantization strategies are available for each tile, we can use the algorithm of [7] to choose the best quantization strategy concurrently with choosing the best tiling. Note also that the iterative procedure described in [15] can be used to adjust λ so as to minimize D subject to a fixed bit budget, and a similar procedure can be used to minimize the rate subject to a fixed distortion.

3.3. Examples

Feasibility of local cosine best basis compression for textured images was shown in [13] where a quadtree-based algorithm was developed for extracting the best local cosine basis. It was demonstrated in [13] that the resulting image coder outperforms SPIHT [16]—which is a state of the art embedded wavelet coder—on highly textured images. Our algorithm provides more flexibility in the choice of tilings, and therefore outperforms the algorithm of [13], as shown in Fig. 1. We use three 512×512 images from [13]: "fingerprint," "Barbara," and "clown," shown in the top row of Fig. 1. Rate-distortion curves are plotted in the middle row of the figure; in addition, the bottom row shows these curves with rates as percentages of the rate achieved by our algorithm. For the "fingerprint" and "Barbara" images, our algorithm typically achieves about 5% more compression than the quadtree-based algorithm of [13] and up to 35% more compression than SPIHT. For the "clown" image which, as remarked in [13], is more wavelet-friendly, our algorithm performs similarly to SPIHT (except for very low bit rates where SPIHT is significantly better) and outperforms the quadtree-based algorithm of [13] by 5-20%.

4. REFERENCES

[1] N.N. Bennett. Fast algorithm for best anisotropic Walsh bases and relatives. *J. of Appl. and Comput. Harmonic Analysis*, 8:86-103, 2000.

[2] R.R. Coifman and M.V. Wickerhauser. Entropy based algorithms for best basis selection. *IEEE Trans. Inf. Th.*, 38(2):713-718, March 1992.

[3] D.L. Donoho and I.M. Johnstone. Ideal denoising in an orthonormal basis chosen from a library of bases. *Comptes Rendus Acad. Sci.*, Ser. I 319:1317-1322, 1994.

[4] D. Donoho. CART and best-ortho-basis: A connection. *Ann. Stat.*, 25:1870-1911, 1997.

[5] C. Herley, J. Kovačević, K. Ramchandran, and M. Vetterli. Tilings of the time-frequency plane: construction of arbitrary orthogonal bases and fast tiling algorithms. *IEEE Trans. Sig. Proc.*, 41(12):3341-3359, Dec. 1993.

[6] Y. Huang, I. Pollak, C.A. Bouman, and M.N. Do. New algorithms for best local cosine basis search. In *Proceedings* of ICASSP, May 17-21, 2004, Montreal, Quebec. `www.ece.purdue.edu/~ipollak/icassp04.pdf`

[7] Y. Huang, I. Pollak, and C.A. Bouman. Image compression with multitree tilings. In *Proceedings* of ICASSP, March 18-23, 2005, Philadelphia, PA. `www.ece.purdue.edu/~ipollak/icassp05.pdf`

[8] Y. Huang, I. Pollak, M.N. Do, and C.A. Bouman. Optimal tilings and best basis search in large dictionaries. In *Proc. 37th Asilomar Conference on Signals, Systems, and Computers*, Nov. 9-12, 2003, Pacific Grove, CA.

[9] H. Krim and J.-C. Pesquet. On the statistics of best bases criteria. In *Wavelets and Statistics*, Lecture Notes in Statistics, A. Antoniadis, Ed., pp. 193-207. Springer-Verlag, 1995.

[10] M. Lindberg and L.F. Villemoes. Image compression with adaptive Haar-Walsh tilings. In *Wavelet Applications in Signal and Image Processing VIII*, Proc. SPIE 4119, 2000.

[11] S.G. Mallat. *A Wavelet Tour of Signal Processing*, Second Edition. Academic Press, 1999.

[12] H. Malvar. *Signal Processing with Lapped Transforms.* Artech House, 1992.

[13] F. Meyer. Image compression with adaptive local cosines: A comparative study. *IEEE Trans. Im. Proc.*, 11(6):616–629, June 2002.

[14] P. Moulin. Signal estimation using adapted tree-structured bases and the MDL principle. In *Proc. IEEE-SP Int. Symp. TFTS*, pp. 141-143, Paris, June 1996.

[15] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Trans. Im. Proc.*, 2(2):160-175, Apr. 1993.

[16] A. Said and W. A. Pearlman. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. Circ. Syst. Vid. Tech.*, 6(3):243–250, 1996.

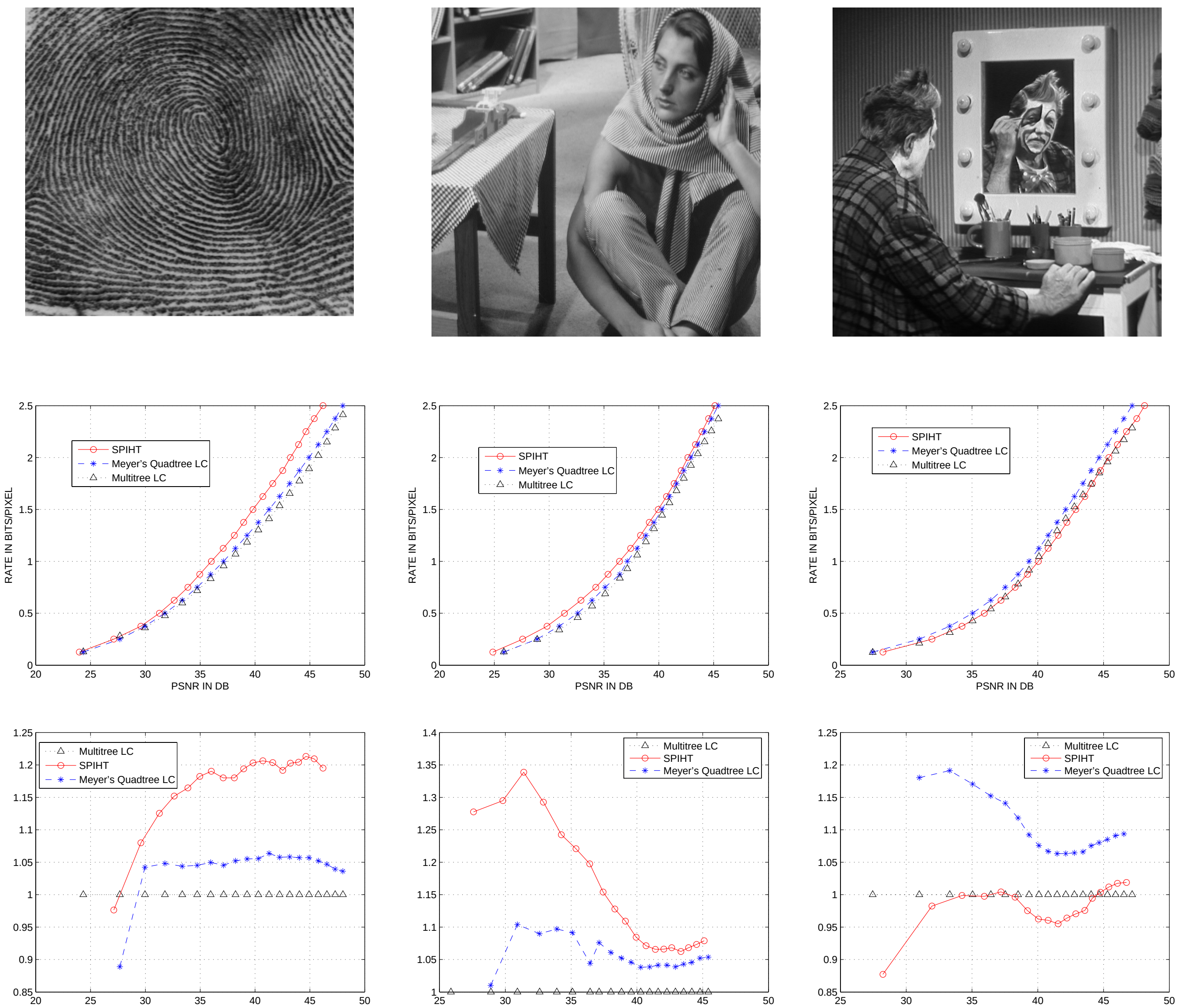

Fig. 1: Top row: three images; middle row: rate-distortion curves for these images, for our algorithm (dotted line with triangles), the algorithm in [13] (dashed line with stars), and SPIHT (solid line with circles); bottom row: rate-distortion curves with bit rates as percentages of the bit rate for our algorithm.