Anisotropic Multi-scale Sparse Learned Bases for Image Compression

Angélique Drémeau, Cédric Herzet, Christine Guillemot and Jean-Jacques Fuchs

INRIA-Rennes Research Center, Campus de Beaulieu, 35042 Rennes Cedex, France

ABSTRACT

This paper proposes a new compression algorithm based on multi-scale learned bases. We first explain the construction of a set of image bases using a bintree segmentation and the optimization procedure used to select the image basis from this set. We then present the sparse orthonormal transforms introduced by Sezer *et al.*[1] and propose some extensions tending to improve the convergence of the learning algorithm on the one hand and to adapt the transforms to the coding scheme used on the other hand. Comparisons in terms of rate-distortion performance are finally made with the current compression standards JPEG and JPEG2000.

Keywords: Image compression, sparsity-distortion optimization, anisotropic learned basis.

1. INTRODUCTION

Recent results in image compression tend to show that adaptation of the transform to the local characteristics of the image allows a gain in performance. In practice, the optimization of the transform to the image characteristics can mainly be made at two levels: *i)* in the spatial domain by adapting the support of the transform; *ii)* in the transformed domain by adapting the atoms of the projection basis to the signal characteristics we want to describe.

Several contributions considering these general approaches can be found in the literature. Chen[2] suggests to use the DCT on blocks of variable size. Meyer[3] compares different lapped DCT transforms. At the same time, a number of new transform bases have emerged. The DCT in the JPEG scheme is so replaced by other transforms, better adapted to the local statistics of each block.[4] A similar idea is followed by Sezer *et al.*[1] where a library of bases is optimized on a training set to maximize the sparsity of the transform vectors. We can likewise refer to the well-known wavelets,[5] curvelets,[6] contourlets[7] and bandelets,[8] which are very effective at describing the edges of an image. Finally, a hybrid approach is also studied,[9] optimizing both the size of the blocks and the direction of a bandelet basis.

In this paper, we study an image compression method based on a different hybrid approach: the transform basis is selected (in a rate-distortion sense) in a set of bases made up of the concatenation of *local* multi-scale anisotropic (rectangular) bases. The libraries of local bases are optimized in a *sparsity-distortion* sense following a slightly different procedure than the one presented in Sezer *et al.*[1] In particular, we propose another initialization of the algorithm and another norm for the sparse optimization problem. We investigate also the integration of the knowledge of the quantizer used in the compression codec in the basis optimization procedure.

2. MULTI-SCALE BASES FOR IMAGE COMPRESSION

In this section, we first briefly recall the theoretical arguments which motivate the use of a set of bases matched to the local properties of an image. Then, we consider the construction of a set of image bases as the concatenation of local anisotropic bases. We finally present the optimization procedure we use to select the image basis.

E-mail: first-name.last-name@irisa.fr

2.1 Low bit rate compression and sparsity

We consider a compression scheme using an orthonormal basis $\mathbf{B} = \{\mathbf{b}_k\}_{k=1}^N$ and a quantizer $\mathcal{Q}$. An image block $\mathbf{f}$ of N pixels is then approximated by $\hat{\mathbf{f}}$ as

$$\hat{\mathbf{f}} = \sum_{k=1}^N \mathcal{Q}(\langle \mathbf{f}, \mathbf{b}_k \rangle) \mathbf{b}_k \tag{1}$$

where $\langle \mathbf{f}, \mathbf{b}_k \rangle$ denotes the scalar product between vector $\mathbf{f}$ and $\mathbf{b}_k$ and $\mathcal{Q}(\langle \mathbf{f}, \mathbf{b}_k \rangle)$ denotes the nearest value of $\langle \mathbf{f}, \mathbf{b}_k \rangle$ among the set of quantized points.

In that context, Mallat and Falzon[10] show that the number of nonzero quantized transform coefficients (*i.e.,* scalar products), say M, plays a crucial role in the characterization of the rate-distortion performance at low bit rates. On the one hand, they emphasize that the distortion D can be linked to the number of nonzero coefficients M via a function $\varphi(M)$ which modelizes the ability of the basis to approximate a signal with M projection coefficients, *i.e.,*

$$D = \varphi(M). \tag{2}$$

On the other hand, they show that the rate required to represent the quantized coefficients is proportional to M, *i.e.,*

$$R = \gamma M. \tag{3}$$

As a consequence, Mallat and Falzon's work leads to the following observation: at low bit rates, the rate-distortion performance depends on the ability of the basis to provide a good approximation of the signal with few coefficients.

In practice, different bases are often well-suited to different kinds of images. This suggests that enhanced rate-distortion performance could be achieved by adapting the projection basis to the image to compress. As mentioned in the introduction, we consider here a compression algorithm which selects the best basis (in a rate-distortion sense) among a set of bases built as the concatenation of anisotropic multiscale local "sparse" bases. We explain the construction of this set of bases and the optimization procedure used to select the image basis in the remainder of this section.

2.2 Bintree concatenation of local bases

A common way to form an image basis from local bases is to use a quadtree. This method leads to a set of bases which are the concatenation of square local bases whose support sizes can possibly be different (see Le Pennec et Mallat[9] for example). In this paper, we want to exploit anisotropic (rectangular) local basis supports. This is achieved by the use of a bintree.

With this additional degree of freedom, a bintree segmentation increases the number of possible image bases. On the one hand, we can intuitively assume that the more bases we have at our disposal, the more likely we are to select a basis which properly catches the local properties of the image. On the other hand, this leads to an increase of the bit rate since the choice of the selected basis has to be transmitted.

Therefore, in addition to the bit rate required to code the transform coefficients (expressed in (3) and noted R_c here), we have to take into account the encoding cost of the image basis, *i.e.,* the cost associated to the encoding of the tree specifying the local basis supports (noted R_s), and the cost associated to the encoding of the local basis indices (noted R_m). The total bit rate can then be expressed as:

$$R = R_c + R_s + R_m. \tag{4}$$

In the next section we will discuss how to select the basis which leads to the best compromise in terms of rate-distortion. The use of a bintree instead of a quadtree increases R_s but we can observe experimentally that this is compensated by a better local adaptivity, resulting in a gain in performance.

2.3 Basis selection in a rate-distortion sense

In the transform coding paradigm, two parameters can impact the rate-distortion performance: the transform basis and the quantization of the transform coefficients. Therefore, the rate-distortion optimization problem can be formalized as follows:

$$(\mathbf{B}^\star, \mathcal{Q}^\star) = \arg\min_{\mathbf{B}, \mathcal{Q}} \ D(\mathbf{B}, \mathcal{Q}) \quad subject\ to \quad R(\mathbf{B}, \mathcal{Q}) \leq R_t, \tag{5}$$

where $\mathbf{B}$ (resp. $\mathcal{Q}$) is a trial image basis (resp. quantizer) and R_t is a target rate we specify as a constraint.

Solving (5) is usually quite cumbersome. Instead, Shoham and Gersho[11] introduced the following simplified unconstrained problem:

$$(\mathbf{B}_\lambda^\star, \mathcal{Q}_\lambda^\star) = \arg\min_{\mathbf{B}, \mathcal{Q}} \ D(\mathbf{B}, \mathcal{Q}) + \lambda R(\mathbf{B}, \mathcal{Q}), \tag{6}$$

where λ is a Lagrangian multiplier. They showed that if $R(\mathbf{B}_\lambda^\star, \mathcal{Q}_\lambda^\star) = R_t$, then $(\mathbf{B}_\lambda^\star, \mathcal{Q}_\lambda^\star)$ is also a solution of the initial optimization problem (5).

This condition is unfortunately not always satisfied in practice. However, the approach proposed by Shoham and Gersho has been shown to lead to good performance in many contributions and will therefore be considered hereafter.

In our compression scheme, we consider a uniform scalar quantizer with a quantization step Δ and a deadzone equal to 2Δ. Optimizing $\mathcal{Q}$ is then equivalent to optimizing Δ. In this context, Le Pennec and Mallat[9] proved that, as long as assumption (3) is valid, the optimal quantization step, say $\Delta_\lambda^\star$, is related to λ as follows:

$$\Delta_\lambda^\star = \sqrt{\frac{4\,\gamma\,\lambda}{3}}. \tag{7}$$

With this result, problem (6) reduces to optimizing the rate-distortion function with respect to basis $\mathcal{B}$, *i.e.*,

$$\mathbf{B}_\lambda^\star = \arg\min_{\mathbf{B}} \ D(\mathbf{B}, \Delta_\lambda^\star) + \lambda\,R(\mathbf{B}, \Delta_\lambda^\star). \tag{8}$$

The solution of this problem can be found efficiently using dynamic programming[12] provided that: *i)* the set of image bases is the "tree" concatenation of local bases; *ii)* the distortion and the rate can be decomposed into local terms associated to each local basis. The first condition is obviously satisfied by our set of image bases, see section 2.2. The second one requires some additional coding assumptions.

First, we adopt a simple implementation of the bintree by assigning "1" to internal nodes and "0" to leaf nodes. Hence, R_s can be expressed as the sum of the bit rates required to code each tree leaf, say R_s^l, *i.e.*,

$$R_s = \sum_{l \in \mathcal{L}} R_s^l, \tag{9}$$

where $\mathcal{L}$ is the set of tree leaves.

Second, we assume that the choice of the local basis indices is encoded by a fixed-length code (FLC). Hence, denoting by L the length of the FLC code, we have

$$R_m = \sum_{l \in \mathcal{L}} R_m^l,$$
$$= |\mathcal{L}|\,L, \tag{10}$$

where $|\mathcal{L}|$ is the cardinal of $\mathcal{L}$. Note that the use of a FLC is in general quite suboptimal. This choice is made here only for complexity purpose since it allows to decompose R_m as a sum of local terms. In section 4, we will describe a more efficient way to encode the local basis indices. Here, (10) can therefore be seen as an upper bound on the rate which can actually be achieved by our practical scheme.

Finally, the rate required to encode the transform coefficients, R_c, can readily be expressed as a sum of local terms by using assumption (3). Mallat and Falzon evaluate experimentally γ at 5.5 for the DCT basis and 6.5 for the wavelet basis. For the sparse learned bases, we estimate* $R_c = 6.5\ M$ as a good approximation of the rate required to encode the transform coefficients.

Summarizing these observations, we finally obtain that the overall rate can be decomposed as a sum of "local" rates, *i.e.*, $R = \sum_{l \in \mathcal{L}} R^l$ with $R^l = R_c^l + R_s^l + R_m^l$. Standard dynamic programming methods[12] can therefore be applied to solve (8).

3. SPARSE LEARNED BASES

The procedure described above leads to the selection of an image basis made up of the bintree concatenation of local block bases. At each node of the bintree, the choice of the local basis is made over a set $\mathcal{B}_{local}$. The construction of this set is crucial for the performance achievable by the compression scheme. We follow here an approach similar to the one proposed by Sezer *et al.*[1] for the optimization of a set of orthogonal bases in a sparsity-distortion sense.

3.1 Sezer's algorithm

Let $\{\mathbf{f}^j\}_{j=1}^L$ be a set of image blocks of same dimension, say N, and let $\mathcal{B}_{local} = \{\mathbf{B}_1, \ldots, \mathbf{B}_K\}$ be a set of K local orthogonal bases to optimize. Each basis is optimized from a subset of the training blocks $\{\mathbf{f}^j\}_{j=1}^L$. More particularly, let $\mathcal{S}_i$ be the set of indices of the blocks involved in the optimization of $\mathbf{B}_i$. Similarly to the rate-distortion optimization exposed in the previous section, the optimization of the K bases in a sparsity-distortion sense can be written as an unconstrained problem depending on a Lagrangian multiplier μ:

$$\forall i \in \{1, \ldots, K\} \qquad \mathbf{B}_i^\star = \arg\min_{\mathbf{B}_i \in \mathbb{R}^{N \times N}} \left\{ \sum_{j \in \mathcal{S}_i} \min_{\mathbf{c}^j \in \mathcal{C}^j} \left\{ \|\mathbf{f}^j - \mathbf{B}_i \mathbf{c}^j\|_2^2 + \mu \|\mathbf{c}^j\|_p \right\} \right\} \qquad s.t. \qquad \mathbf{B}_i^T \mathbf{B}_i = \mathbf{I}, \quad (11)$$

where $\|\cdot\|_p$ denotes the l_p-norm of a vector ($p \in [0,1]$), $\mathbf{I}$ is the unity matrix and $\mathcal{C}^j$ is the set of possible values that $\mathbf{c}^j$ can take on. The solution of each minimization problem can be found by means of iterative conditional minimization over the $\mathbf{B}_i$'s and $\mathbf{c}^j$'s.

More particularly, Sezer *et al.* address the case where $\mathcal{C}^j = \mathbb{R}^N\ \forall j$ and $p = 0$. They show that the conditional minimization with respect to $\mathbf{c}_i^j$ with fixed $\mathbf{B}_i$ is then given by the following threshold operator:

$$\mathbf{c}_i^j(l) = \left\{ \begin{array}{ll} \mathbf{v}_i^j(l) & if \quad |\mathbf{v}_i^j(l)| > \sqrt{\mu} \\ 0 & otherwise \end{array} \right. \tag{12}$$

where $\mathbf{v}_i^j = \mathbf{B}_i^T \mathbf{f}^j$.

After convergence of the iterative conditional minimization, the subsets of blocks used to optimize each basis is updated as follows

$$\forall i \in \{1, \ldots, K\} \qquad \mathcal{S}_i = \left\{ j \in \{1, \ldots, L\} \,\middle|\, \mathbf{B}_i^\star = \arg\min_{\mathbf{B} \in \mathcal{B}_{local}^\star} \min_{\mathbf{c}^j \in \mathcal{C}^j} \{ \|\mathbf{f}^j - \mathbf{B}\mathbf{c}^j\|_2^2 + \mu \|\mathbf{c}^j\|_p \} \right\}, \tag{13}$$

where $\mathcal{B}_{local}^\star \triangleq \{\mathbf{B}_1^\star, \ldots, \mathbf{B}_K^\star\}$ is the set of bases optimized in (11).

3.2 Extensions and improvements

In this section, we introduce some modifications in Sezer's algorithm in order to improve its convergence and its performance in a rate-distortion sense. To study the relevance of these modifications, we apply the learned bases on 8×8 blocks and adopt a simple coding scheme, using Huffman codes to encode the quantized coefficients and a run-length encoder to code the indices of the nonzero coefficients. The rate-distortion performance achieved is then compared with Sezer's one.

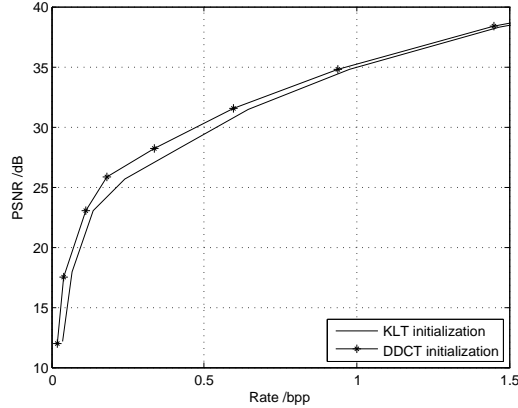*The value of γ actually depends on the practical coding implementation, see section 4.

Figure 1. Rate-distortion performance achieved by two different initializations: KLT and DDCT.

At the initialization step, Sezer *et al.* perform a classification of the training blocks into K subsets by means of image gradients. On each subset, the basis is then initialized to the Karhunen-Loeve transform (KLT) using the corresponding blocks. We proceed differently, by first initializing K bases independently of the training set and then linking the training blocks to one of the bases with (13). To this purpose, we propose to use the directional DCT (DDCT) introduced by Zeng *et al.*,[13] which provide better coding performance for image blocks that contain directional edges. Thus, using DDCT encourages an initial classification of the image blocks according to some directional similarities. In this paper, we want to exploit anisotropic supports. We thus extend the principle of DDCT to rectangular bases.

Fig. 1 presents the rate-distortion performance achieved by Sezer's algorithm initialized by KLT and by DDCT on image Barbara. The latter improves the PSNR by more than 1dB at low bit rates.
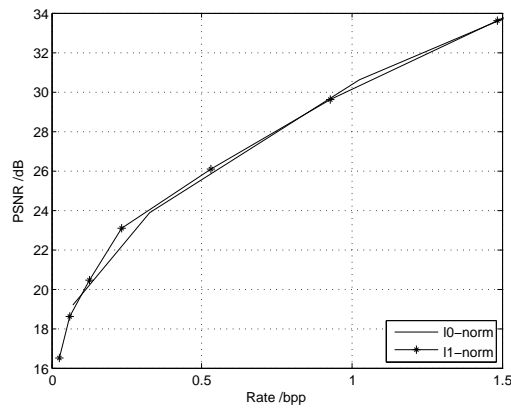

Figure 2. Rate-distortion performance achieved by two different optimization norms: l_0-norm and l_1-norm.

Another important issue is the choice of the l_p-norm used in (11) which can lead to different performance. The case $p = 0$ implements the actual "sparsity" criterion. However, since problem (11) is not convex (the equality constraints are not linear) there can exist several local minima. Hence, other choices of p may be advantageous in terms of convergence of the iterative optimization algorithm. In particular, we investigate the case $p = 1$, which de facto is often used instead of the l_0-norm in sparse standard problems. The solution of (11) with l_1-norm can also be obtained by a simple thresholding operation as shown by Lesage *et al.*:[14]

$$\mathbf{c}_i^j(l) = \begin{cases} \mathbf{v}_i^j(l) - \mu/2 & if \quad \mathbf{v}_i^j(l) > \mu/2 \\ 0 & if \quad |\mathbf{v}_i^j(l)| \leq \mu/2 \\ \mathbf{v}_i^j(l) - \mu/2 & if \quad \mathbf{v}_i^j(l) < -\mu/2 \end{cases} \tag{14}$$

where $\mathbf{v}_i^j = \mathbf{B}_i^T \mathbf{f}^j$.

The comparative rate-distortion curves are given in Fig. 2 on image Roof. Both algorithms are initialized with DDCT. We can see that the choice of the l_1-norm leads to a slight improvement of the PSNR.
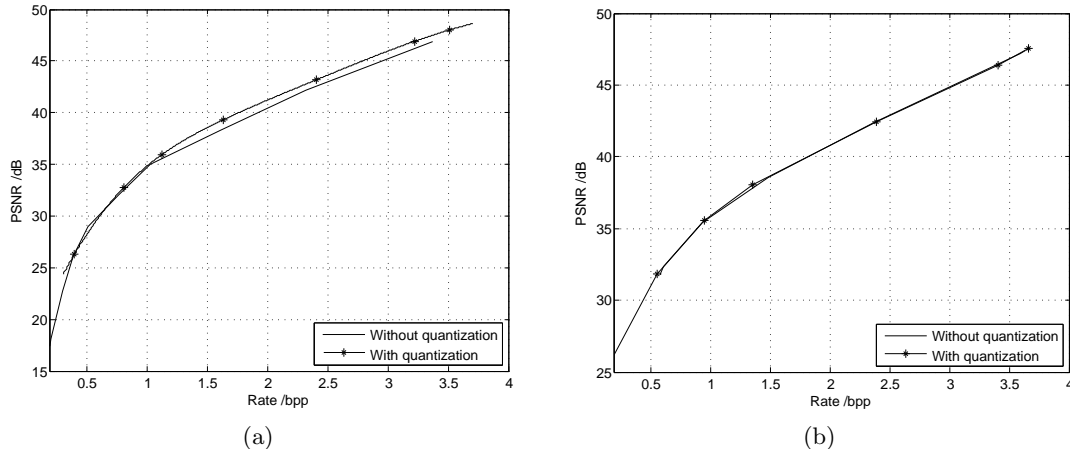

Figure 3. Rate-distortion performance achieved with and without integration of the quantization in the learning algorithm: (a) on 4×4 blocks, (b) on 8×8 blocks.

Finally, since the learned bases are used in an image compression context, it seems interesting to integrate the knowledge of the coding scheme used after the transform step. In practice, this comes to take the quantization into account and thus, to constraint the $\mathcal{C}^j$'s to be discrete sets of points. If we consider a uniform quantizer with a quantization step Δ and a deadzone equal to 2Δ, we have $\mathcal{C}^j = \{0\} \cup \{\pm(\frac{3}{2}\Delta + k\Delta)\}_{k\in\mathbb{N}}$. Using such a definition for the $\mathcal{C}^j$'s implicitly relates the set of optimized bases to the quantizer used in the considered coding scheme.

Fig. 3 analyses the relevance of the integration of quantization in the learning algorithm on image Peppers. The algorithms are initialized with DDCT and use the l_1-norm. For 4×4 blocks, it improves the rate-distortion performance at moderate-to-high bit rates (see Fig. 3(a)). However it is not the case at low bit rates, or for blocks of higher size as we can see in Fig. 3(b) for 8×8 blocks. Several reasons can explain these disappointing results. The most likely one involves the classification step used in the learning algorithm. Indeed, taking the quantization into account generates new instabilities: the gap between the real and the quantized values is higher at low bit rates and has an even more important impact on the rate-distortion performance because of the small number of nonzero coefficients. As the block size increases, this behaviour is reinforced by the sparsity constraint. As a consequence, we will not take the quantization into account in the final version of the learning algorithm used in the codec presented in section 4.

4. IMAGE CODEC: IMPLEMENTATION AND RESULTS

In this section, we detail the implementation of our image compression codec and illustrate its performance.

We consider a dictionary of image bases constructed as explained in section 3. The supports of the local bases range from 32×32 to 4×4 pixels. Depending of the size of the support, we used between 50000 and 400000 training blocks to optimize a set of 7 bases by the procedure and the improvements described in section 3.2. The bases are first initialized with DDCT whose directional modes correspond to the prediction modes of H.264.[15] Mode "1" stands for the conventional DCT. Finally, the transform basis is selected with the optimization procedure described in section 2.3.

The quantized transform coefficients are encoded with Huffman codes. The Huffman tables are optimized according to the size of the support of the local transforms. The indices of the nonzero coefficients $\mathcal{I}$ are encoded with a run-length encoder.

The encoding of the local basis indices is performed by means of a quadtree, as illustrated in Fig. 4. Fig. 4(a) represents the supports of the local bases. The number inside each support corresponds to the selected local basis. Fig. 4(b) represents the quadtree encoding of the corresponding indices. We proceed as follows. The

image is segmented into 4 square blocks of equal dimension. If all the local bases in a block have the same index, this block corresponds to a leaf of the quadtree and is labelled by the common index of the local bases in the block; otherwise the block is subdivided into 4 square blocks and so one.
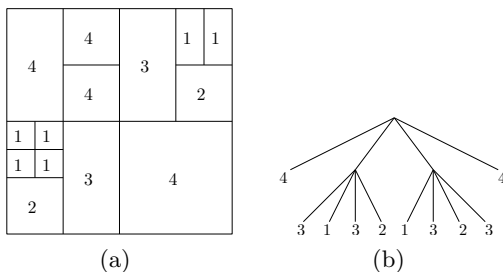

Figure 4. (a) Example of a bintree segmentation and (b) the corresponding quadtree encoding of the local basis indices

In Fig. 5, we illustrate the performance achieved by our codec for the compression of "Cameraman". Fig. 5(a) represents the supports of the local bases making up the image basis as well as their indexes for $\mu = 50$. Fig. 5(b) compares the rate-distortion performance obtained by our compression scheme to the one obtained by the "Sparse Learned Bases" restricted to 8×8 pixels blocks and the compression standards JPEG and JPEG2000. We can notice that the proposed codec outperforms JPEG by more than 1 dB at low bit rates. As shown in Table 1, this observation holds for all the images we have tested. As far as "Cameraman" is concerned, we can also observe that the proposed codec slightly outperforms JPEG2000 at low bit rates. Additional results in Table 1 show that the proposed codec exhibits a good behavior with respect to JPEG2000.
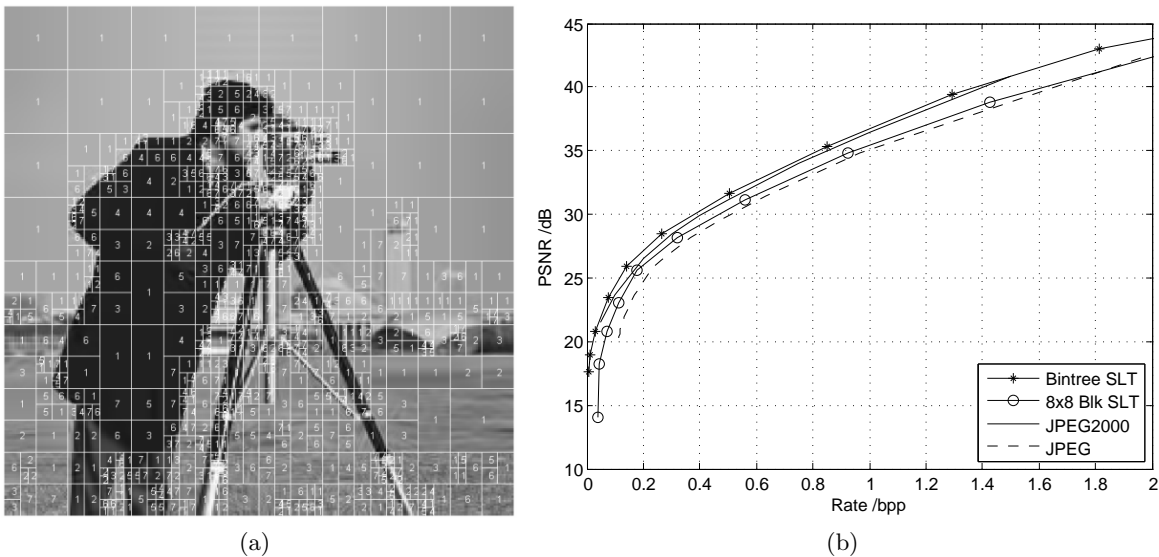

Figure 5. (a) Local basis supports obtained for "Cameraman" at $R = 0.51$ bit-per-pixel and PSNR= 31.57 dB. The number denotes the basis selected on each support. (b) Rate-distortion curves for the compression of "Cameraman" with the proposed codec ("Bintree SLT"), the "Sparse Learned Transform" restricted to 8×8 pixels blocks ("8×8 Blk SLT"), JPEG2000 and JPEG standards.

5. CONCLUSION

In this paper, we studied the performance of a "basis-adaptive" image compression codec. The set of bases used in the codec is built by the concatenation of local anisotropic sparse learned bases. The selection of the optimal basis is made by exploiting the bintree structure of the basis dictionary and using dynamic programming. The local anisotropic bases are learned according to the algorithm presented by Sezer *et al.*[1] We proposed and

Rate	Lena (512×512 pixels)			Barbara (512×512 pixels)			Roof (512×512 pixels)		
(bpp)	Codec	JPEG2000	JPEG	Codec	JPEG2000	JPEG	Codec	JPEG2000	JPEG
	(dB)	(dB)	(dB)	(dB)	(dB)	(dB)	(dB)	(dB)	(dB)
0.1	29.50	29.90	-	25.27	24.80	-	24.47	23.50	18.75
0.2	32.42	33.00	29.88	27.99	27.30	24.03	27.36	26.50	22.79
0.5	36.74	37.30	35.44	32.44	32.20	29.75	33.07	31.70	29.59
0.7	38.21	38.66	36.82	34.50	34.28	32.30	35.84	34.18	32.57
1.0	39.67	40.40	38.54	37.24	37.10	35.18	39.00	37.60	35.94
1.5	41.86	42.80	-	40.30	40.40	38.24	42.83	42.30	39.93

Table 1. Summary table of the rate-distortion performance for the compression of 3 different images with the proposed codec and the standards JPEG and JPEG2000.

introduced several modifications in this algorithm, involving the initialization of the bases and the l_p-norm used in the sparse optimization.

As far as the images tested are concerned, the proposed codec outperforms JPEG in terms of rate-distortion and is slightly superior to JPEG2000 in most cases.

REFERENCES

[1] Sezer, O. G., Harmanci, O., and Guleryuz, O. G., "Sparse orthonormal transforms for image compression," in [*Proc. IEEE Int'l Conference on Image Processing (ICIP)*], (October 2008). San Diego, CA.

[2] Chen, C.-T., "Adaptive transform coding via quadtree-based variable blocksize dct," in [*Proc. IEEE Int'l Conference on Acoustics, Speech and Signal Processing (ICASSP)*], (23-26 May 1989).

[3] Meyer, F. G., "Image compression with adaptive local cosines: A comparative study," *IEEE Trans. On Image Processing* **11**, 616–629 (June 2002).

[4] Helsingius, M., Kuosmanen, P., and Astola, J., "Image compression using multiple transforms," *Signal Processing* **15**, 513–529 (2000).

[5] Mallat, S., "A theory for multiresolution signal decomposition : the wavelet representation," *IEEE Trans. On Pattern Analysis Machine Intelligence* **11**, 674–693 (July 1989).

[6] Candes, E. J. and Donoho, D. L., "Curvelets: A surprisingly effective nonadaptive representation for objects with edges," tech. rep., Stanford University CA - Dept of Statistics (2000).

[7] Do, M. N. and Vetterli, M., "Contourlets: a directional multiresolution image representation," in [*Proc. IEEE Int'l Conference on Image Processing (ICIP)*], **1**, 357–360 (2002).

[8] LePennec, E. and Mallat, S., "Bandelet image approximation and compression," *SIAM MMS* **4**, 992–1039 (April 2005).

[9] LePennec, E. and Mallat, S., "Sparse geometric image representations with bandelets," *IEEE Trans. On Image Processing* **14**, 423–438 (April 2005).

[10] Mallat, S. and Falzon, F., "Analysis of low bit rate image transform coding," *IEEE Trans. On Signal Processing* **46**, 1027–1042 (April 1998).

[11] Shoham, Y. and Gersho, A., "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. On Acoustics, Speech and Signal Processing* **36**, 1445–1453 (September 1988).

[12] Ramchandran, K. and Vetterli, M., "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. On Image Processing* **2**, 160–175 (April 1993).

[13] Zeng, B. and Fu, J., "Directional discrete cosine transforms - a new framework for image coding," *IEEE Trans. On Circuits and Systems for Video Technology* **18**, 305–313 (March 2008).

[14] Lesage, S., Gribonval, R., Bimbot, F., and Benaroya, L., "Learning unions of orthonormal bases with thresholded singular value decomposition," in [*Proc. IEEE Int'l Conference on Acoustics, Speech and Signal Processing (ICASSP)*], **5**, v293–v296 (18-23 March 2005).

[15] ITU-T Rec. H.264 ISO/IEC 14496-10 (AVC), *Advanced Video Coding for Generic Audiovisual Services* (March 2005).