Sparse optimization with directional DCT bases for image compression

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

INRIA Centre Rennes - Bretagne Atlantique, Campus universitaire de Beaulieu, 35000 Rennes, France

Context

State of the art: image compression

Adaptation of the transform to the image characteristics

- In the spatial domain by adapting the support of the transform,
  - DCT on blocks of variable size
  - Lapped DCT
- In the transformed domain by adapting the atoms of the projection basis,
  - Learned bases
  - Wavelets, curvelets, contourlets, bandelets, directional transforms
- Hybrid approaches.

Theoretical background: low bit rate compression and sparsity [1]

Dependency of distortion and rate on the number of nonzero quantized transform coefficients, say L

$$D = \varphi(L),$$
$$R = \gamma\, L,$$

where $\varphi(L)$ and γ depend on the basis.
⇝ 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.

Contributions of this paper

- Introduction of a new "Basis-adaptive" image compression codec.
- Adaptation of the basis made at both levels - spatial and transformed.
- Use of a set of bases built by the concatenation of local anisotropic directional DCT bases.
- Selection of the optimal basis made by exploiting a bintree structure, using dynamic programming.
- Comparisons in terms of rate-distortion performance are made between the proposed coder and the standard JPEG and JPEG2000.

Anisotropic directional DCT

Directional DCT [2]

- Separability property of the standard 2D-DCT transform: a 2D-DCT can be realized by performing successively a 1D-DCT on the columns and a 1D-DCT on the rows of the considered image block.
- Directional DCT (DDCT) are obtained by a modification of the scan order of the first 1D-DCT. Thus, to any directional mode corresponds a different scan order.

⇝ Example on a $8{\times}4$ px block:

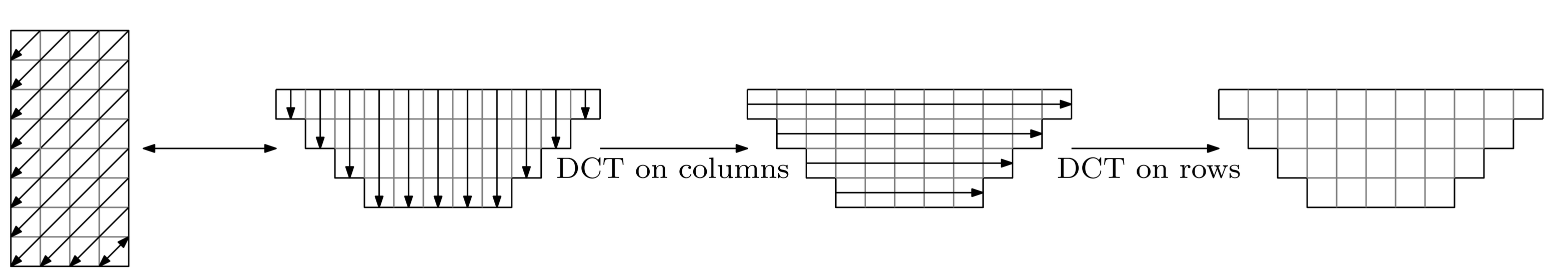

FIGURE 1: "Diagonal down-left" DCT transform on a rectangular block

Bintree concatenation of local DDCT

Trees are common ways to form a set of image bases from local bases. We prefer the use of a bintree segmentation, which:

⊕ Allows to exploit anisotropic (*rectangular*) basis supports,
⊕ Increases the number of possible image bases,
⊖ Increases the bit rate since the choice of the selected basis has to be transmitted, BUT, the additional cost is compensated by a better adaptivity, resulting in a gain in performance (observed experimentally).

Basis selection

Optimization problem

Given a target rate R_t, find the transform basis $\mathcal{B}^\star$ and the quantizer $\mathcal{Q}^\star$ which lead to the best rate-distortion compromise, *i.e.,* :

$$(\mathcal{B}^\star, \mathcal{Q}^\star) = \arg\min_{\mathcal{B},\mathcal{Q}} \; D(\mathcal{B},\mathcal{Q}) \quad \textit{subject to} \quad R(\mathcal{B},\mathcal{Q}) \leq R_t.$$

This is equivalent ([3]) to a simplified unconstrained problem depending on a Lagrangian multiplier λ:

$$(\mathcal{B}^\star_\lambda, \mathcal{Q}^\star_\lambda) = \arg\min_{\mathcal{B},\mathcal{Q}} \; D(\mathcal{B},\mathcal{Q}) + \lambda R(\mathcal{B},\mathcal{Q}).$$

In case of $\mathcal{Q}$, uniform scalar quantizer with quantization step Δ and deadzone equal to 2Δ, the optimal quantization step, $\Delta^\star_\lambda$, is related to λ as follows ([4]):

$$\Delta^\star_\lambda = \sqrt{\frac{4\,\gamma\,\lambda}{3}}.$$

Hence, we finally have to solve the following optimization problem:

$$\mathcal{B}^\star_\lambda = \arg\min_{\mathcal{B}} \; D(\mathcal{B}, \Delta^\star_\lambda) + \lambda\, R(\mathcal{B}, \Delta^\star_\lambda).$$

Dynamic programming

Optimization problem can be solved using dynamic programming provided that:
1. the set of image bases is the "tree" concatenation of local bases,
⇝ satisfied by our set of image bases,
2. distortion and rate can be decomposed into local terms associated to each local basis,
⇝ satisfied by the distortion but has to be verified for the rate.

Total bit rate:

$$R = R_c + R_s + R_m.$$

where R_c is the bit rate required to code the transform coefficients, R_s the encoding cost of the image basis support and R_m the cost associated to the encoding of the local directional modes.

- R_c proportional to the number of nonzero transform coefficients and thus satisfies condition 2.
- Simple implementaion of the bintree, "1" to internal nodes and "0" to leaf nodes, fulfills condition 2 for R_s.
- Assumption of FLC encoding of the directional modes allows to express R_m as a sum of local terms. In our practical scheme, we prefered another more efficient encoding so that the R_m used for the basis selection is an upper bound on the rate actually achieved.

Performance analysis

Implementation

- Set of bases: bintree concatenation of local DDCT bases. Supports of the local bases range from $32{\times}32$ to $4{\times}4$ pixels. For each support size, 7 directional modes are defined, relative to the prediction modes of H.264.
- Transform coefficients encoding: quantized values are encoded with Huffman codes, optimized according size of the support of the local bases. Indices of the nonzero coefficients are encoded with a run-length encoder.
- Directional modes encoding: performed by means of a quadtree.

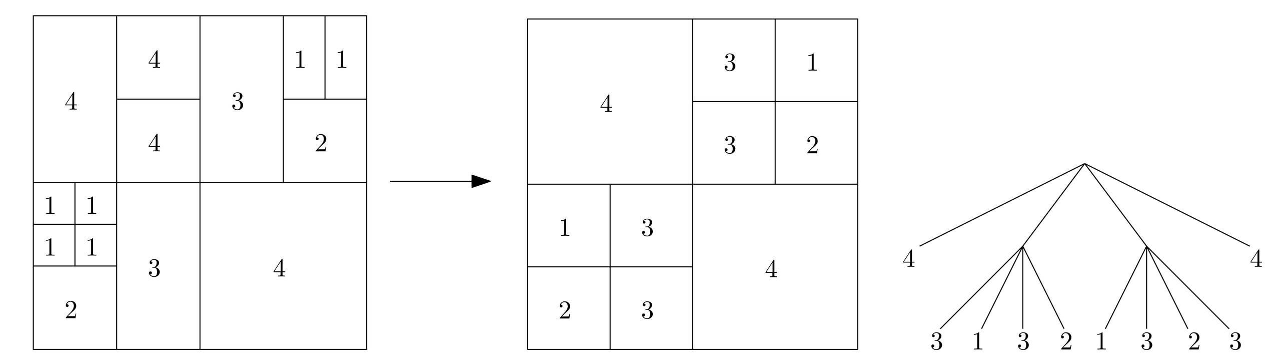

FIGURE 2: Example of a bintree segmentation and the corresponding quadtree encoding of the local basis indices

Results

Performance of the proposed codec is evaluated via the PSNR versus the rate.

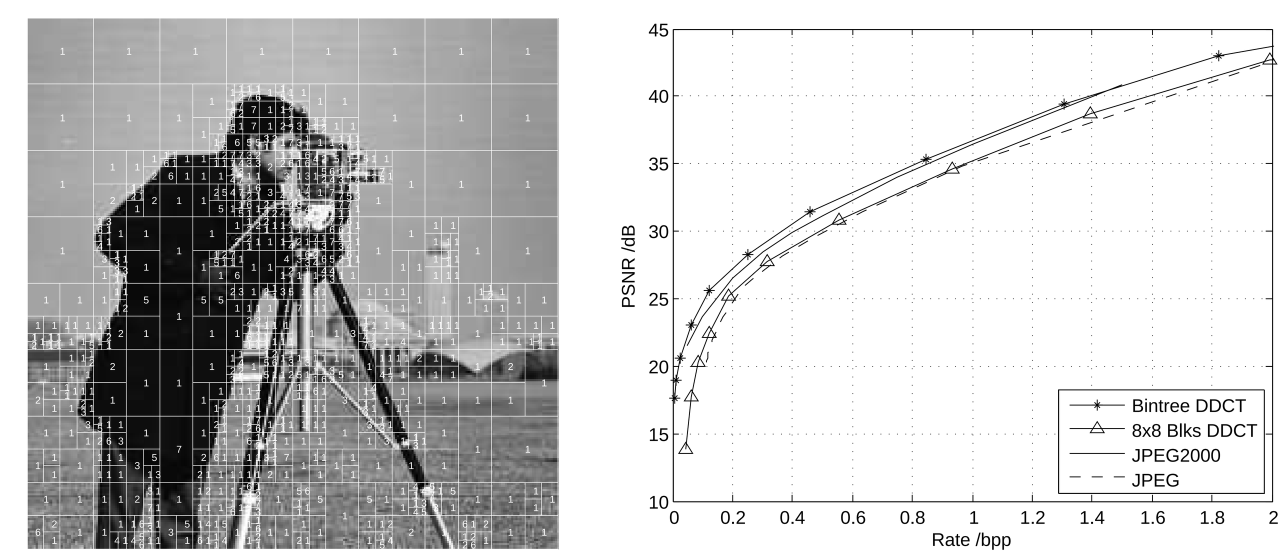

FIGURE 3: Left: bintree DDCT segmentation obtained for "Cameraman" at $R = 0.46$ bpp and PSNR$= 31.4$ dB. Right: rate-distortion curves obtained for the compression of "Cameraman" with the proposed coder, the DDCT applied on $8{\times}8$ blocks, the JPEG2000 standard and the JPEG standard (right)

Rate	Lena ($512{\times}512$ pixels)			Barbara ($512{\times}512$ pixels)			Roof ($512{\times}512$ pixels)		
(bpp)	Coder (dB)	JPEG2 (dB)	JPEG (dB)	Coder (dB)	JPEG2 (dB)	JPEG (dB)	Coder (dB)	JPEG2 (dB)	JPEG (dB)
0.1	29.48	29.90	-	25.31	24.80	-	24.63	23.50	18.75
0.2	32.65	33.00	29.88	28.14	27.30	24.03	27.53	26.50	22.79
0.5	36.76	37.30	35.44	32.69	32.20	29.75	33.25	31.70	29.59
0.7	38.36	38.66	36.82	34.91	34.28	32.30	35.98	34.18	32.57
1.0	39.70	40.40	38.54	37.29	37.10	35.18	39.03	37.60	35.94
1.5	41.83	42.80	-	40.38	40.40	38.24	42.95	42.30	39.93

FIGURE 4: Summary table of the rate-distortion performance for the compression of 3 different images with the proposed coder and the standards JPEG and JPEG2000.

References

[1] S. Mallat and F. Falzon, "Analysis of low bit rate image transform coding," *IEEE Trans. On Signal Processing*, vol. 46, no. 4, pp. 1027–1042, April 1998.

[2] B. Zeng and J. Fu, "Directional discrete cosine transforms - a new framework for image coding," *IEEE Trans. On Circuits and Systems for Video Technology*, vol. 18, no. 3, pp. 305–313, March 2008.

[3] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. On Acoustics, Speech and Signal Processing*, vol. 36, no. 9, pp. 1445–1453, September 1988.

[4] E. LePennec and S. Mallat, "Sparse geometric image representations with bandelets," *IEEE Trans. On Image Processing*, vol. 14, no. 4, pp. 423–438, April 2005.