GRAPH-CUT RATE DISTORTION ALGORITHM FOR CONTOURLET-BASED IMAGE COMPRESSION

M. Trocan, B. Pesquet-Popescu

GET - Télécom Paris
Signal and Image Processing Dept.
46, rue Barrault, 75634 Paris, France
{trocan, pesquet}@tsi.enst.fr

J.E. Fowler

Mississippi State University
Electrical & Computer Engineering Dept.
Box 9571, Mississippi State, MS 39762 USA
fowler@ece.msstate.edu

ABSTRACT

The geometric features of images, such as edges, are difficult to represent. When a redundant transform is used for their extraction, the compression challenge is even more difficult. In this paper we present a new rate-distortion optimization algorithm based on graph theory that can encode efficiently the coefficients of a critically sampled, non-orthogonal or even redundant transform, like the contourlet decomposition. The basic idea is to construct a specialized graph such that its minimum cut minimizes the energy functional. We propose to apply this technique for rate-distortion Lagrangian optimization in subband image coding. The method yields good compression results compared to the state-of-art JPEG2000 codec, as well as a general improvement in visual quality.

Index Terms— subband image coding, rate - distortion allocation

1. INTRODUCTION

The compression of natural images is still a challenge for the researchers and industry. The geometric features of images, such as edges, characterized by abrupt changes in pixel intensity, are difficult to represent. The wavelet transform has been succesfully used for image representation [1], due to its energy compaction capacities and compression efficiency [2]. Unfortunately, wavelets are good tools only for catching point discontinuities of the signal. They fail in the case of line and curve singularities, so often present in images. This is due to the fact that standard wavelet transforms applied to images are separable.

In order to overcome the problem of edge representation, Minh N. Do and Martin Vetterli have defined a new family of geometrical wavelets, called contourlets [3]. Thanks to them, one can represent the class of smooth images with discontinuities along smooth curves in a very efficient and sparse way. The theory of geometrical wavelets has progressed in many directions, giving the definitions of wedgelets [4], beamlets [5], curvelets [6], directionlets [7] and others [8, 9], as well as their corresponding fast transforms.

The main advantage of these new decompositions lies in the fact that they possess all the nice properties of classical wavelets, that is space localization and scalability (catching global as well as local characteristics of a signal in a single bitstream) and, additionally, the geometrical wavelet transforms have strong directional characteristics. They can be successfully used in image segmentation and noise removal, as well as in image compression: as shown in [10], the codec based on wedgelets gives better performance in image compression than the JPEG2000 standard.

In this paper we present a new rate-distortion optimization algorithm based on graph theory that can encode efficiently the contourlet coefficients. As described in [11], problems that arise in computer vision can be naturally expressed in terms of energy minimization. The basic idea is to construct a specialized graph for the energy function to be minimized such that the minimum cut on the graph also minimizes the energy (either globally or locally). The minimum cut, in turn, can be computed very efficiently by max flow algorithms [12]. These methods have been successfully used for a wide variety of vision problems [13, 14], including image restoration, stereo and motion, image synthesis, image segmentation, voxel occupancy, multicamera scene reconstruction and medical imaging. We propose to use the graph-cut mechanism for the minimization of the rate-distortion Lagrangian function. In order to do this, we have designed the graph starting from the subbands given by the contourlet decomposition, where the nodes are given by the spatial subbands and the edges between them are given by their spatial position. As it will be shown by the experimental results, the method gives good compression results compared to the state-of-art JPEG2000 codec, as well as a general improvement in visual quality.

This paper is organised as follows: Section 2 describes the interest of contourlets for image compression. The graph-cut rate-distortion algorithm used for contourlet coefficients coding is presented in Section 3. Some experimental results are presented in Section 4. Finally, conclusions and future work directions are given in Section 5.

2. CONTOURLETS IN IMAGE COMPRESSION

In [3] is presented a geometrical transform that preserves the interesting features of classical wavelets, namely the multiresolution and local characteristics of the signal and at the expense of a spatial redundancy, it better represents the directional features of the image. As shown in Fig.1, the transform is a multiscale and directional decomposition using a combination of a Laplacian pyramid (LP) and a directional filter bank (DFB). Bandpass images from the LP are passed to a DFB so that directional information can be retrieved. This results into a filter bank structure, named contourlet filter bank, which decomposes images into directional subbands at multiple scales. As its redundancy is given only by the LP transform, it has an upper limit of $4/3$, which makes the scheme more appropriate for compression than other geometrical transforms. Another reason for which we have considered this scheme is that contourlets can be approximated with less coefficients than the wavelets; that is, for a contourlet basis, the approximation error for keeping only the M most significant coefficients is:

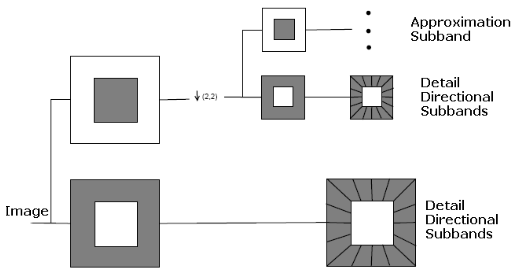

Fig. 1. Contourlet filter bank

$$\|f - f_{M_{contourlet}}\| = O((logM)^3 M^{-2}) \qquad (1)$$

which is smaller than the one obtained on the wavelet basis:

$$\|f - f_{M_{wavelet}}\| = O(M^{-1}). \qquad (2)$$

3. GRAPH-CUT RATE-DISTORTION ALGORITHM

As mentioned in the introduction, the max-flow/min-cut algorithm has been successfully used in computer vision for solving different energy minimization problems. In this paper we propose to apply this technique for rate-distortion Lagrangian optimization in subband image coding.

Generally, for a graph $G = (V, E)$, where V/E is the set of vertices/edges, which have two special vertices (terminals), $q_1, q_2 \in V$, a $q_1 - q_2$ cut is defined as a partition of the vertices in V into two disjont sets Q_1 and Q_2 such that $q_1 \in Q_1$ and $q_2 \in Q_2$. The cost of the cut is given by the sum of costs c of all edges linking Q_1 to Q_2 (the cut-edges in Fig.2), i.e:

$$C(Q_1, Q_2) = \sum_{u \in Q_1, v \in Q_2, (u,v) \in E} c(u, v) \qquad (3)$$

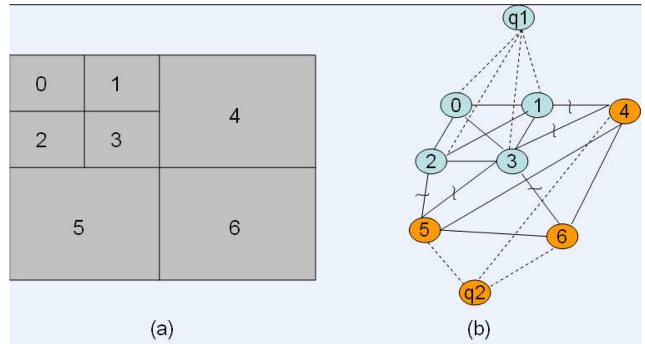

Fig. 2. Two-level spatial decomposition scheme (a) and the corresponding graph-cut repartition of two quantizers (b) (q_1 partition in blue, q_2 partition in orange, where the regular edges are with full lines and the terminal edges with dashed ones).

The multi-terminal min-cut problem can be formulated as follows: given an undirected graph G having $V = N \cup Q$ vertices (where N denote the regular nodes and Q the terminal ones), E edges (each having associated a cost) , find a partitioning of the regular nodes in the graph such that: (a) each partition is connected to one terminal node (i.e. all the regular vertices in a partition are connected to the same terminal node) and (b) the sum of the costs of the edges between any two disjoint partitions (denoted by C) is minimal. A simple approach for this graph can be obtained by seeing as the regular (planar) vertices the decomposition subbands (N), which are connected between them following their 2D geometrical position ($E - N \times Q$), and each terminal node being connected to all the vertices ($N \times Q$) (Fig.2). So one can distinguish two connection types: one between regular vertices and the other one between the terminal nodes and the vertices. The minimum cut corresponds to the cut with the smallest cost C; as proved in [12], C can be found as the solution of the maximal-flow for the given graph. As the graph has been geometrically designed, we define in the following the costs associated to the two types of edges, which coincide with the definition of the energy function to be minimized.

Consider the problem of coding an image at a maximal rate R_{max} with a minimal distortion D. Each image consists of a fixed number of coding units (e.g., in our case, the contourlet spatial subbands), each of them coded with a different quantizer q_i, $q_i \in Q$ where Q is the quantizers set. Let $D_i(q_i)$ be the distortion of subband i when quantized with q_i, and let $R_i(q_i)$ be the number of bits required for coding it. The problem can now be formulated as: find $\min \sum_i D_i(q_i)$, such that $\sum_i R_i(q_i) = R \leq R_{max}$.

In Lagrange-multiplier framework, this constrained optimization problem can be written as:

$$\min \sum_i \left(D_i(q_i) + \lambda R_i(q_i) \right), \quad R \leq R_{max} \qquad (4)$$

where the choice of λ measures the relative importance of distortion, respectively rate for the optimization and which can be determined using a binary search. The advantage of problem formulation in Eq. (4) is that the sum and the minimum operator can be exchanged to:

$$\sum_i \min\left(D_i(q_i) + \lambda R_i(q_i)\right), \quad R \leq R_{max} \tag{5}$$

This formulation obviously reveals that the global optimization can now be carried out independently for each spatial subband, making an efficient implementation feasible.

The distortion D between the original image x and the quantized one, $\widehat{x}$ can be written as the L^2 norm, i.e. $D = \|x - \widehat{x}\|^2$. If in the reconstructed image $\widehat{x}$ we highlight the contribution of each subband, $\widehat{x} = \sum_i \widehat{x}_i$, where $\widehat{x}_i$ is the contribution per subband, then we can also write the image in a similar way, $x = \sum_i x_i$. However, here x_i is completely arbitrary. In the case of a linear basis, it may become $x_i = \sum_k \langle x, \widetilde{e}_{k,i}\rangle e_{k,i}$, where $\widetilde{e}_{k,i}$, $e_{k,i}$ are the analysis, respectively synthesis elements of the biorthogonal basis. Then we have:

$$D = \left\| \sum_i (\widehat{x}_i - x_i) \right\|^2 = \sum_{i,i'} \langle \widehat{x}_i - x_i, \widehat{x}_{i'} - x_{i'} \rangle \tag{6}$$

In a first approximation, we can consider only the diagonal terms, i.e.:

$$D \cong \sum_i \|x_i - \widehat{x}_i\|^2 \tag{7}$$

which amounts at estimating the distortion between the contribution to the image and to the quantized image only of the i^{th} subband. This means we can reconstruct the image only from i^{th} subband coefficients (the others being set to zero).

In [11], O. Veksler et al. propose two graph-cut based algorithms able to reach a minimum for an energy function of the form:

$$\min E(f) = E_{data}(f) + E_{smooth}(f) \tag{8}$$

where E_{smooth} is a smoothness constraint, while E_{data} measures the distortion introduced by the partitioning f with respect to the original data. Without taking into consideration the rate constraint, one can easily translate the Eq.(7) in E_{data} (i.e. $E_{data} = D$). Because E_{data} can be arbitrary chosen, with only the positiveness constraint, we add to it the rate factor, that is:

$$E_{data} = \sum_i (D_i(q_i) + \lambda R_i(q_i)) \tag{9}$$

We can define:

$$E_{smooth} = \sum_{n_1, n_2 \in \mathcal{N}} \mathcal{V}_{n_1,n_2}(q_{n_1}, q_{n_2}) \tag{10}$$

where $\mathcal{N}$ represents the 2D neighborhood system of the nodes and $\mathcal{V}_{n_1,n_2}(q_{n_1}, q_{n_2})$ measures the cost of assigning the quantizers q_{n_1}, q_{n_2} to the adjacent nodes n_1, n_2. We define $\mathcal{V}$ as the Potts interaction penalty, i.e. :

$$\mathcal{V} = \beta T(q_{n_1} \neq q_{n_2}) \tag{11}$$

where T is a boolean operator (e.g. its value equals 1 if the argument is true and 0 otherwise) and β is a real constant which enforces or diminues the smoothing. As can be seen, the definition of E_{smooth} is consistent, as for two strongly correlated subbands the same quantizer choice is imposed. Moreover, it is a metric on the quantizers space, so the $q - expansion$ algorithm [11] can be used in order to minimize E. As one can remark, for a terminal node edge (i.e. link between a quantizer q and a given subband vertex i) the cost is given by the sum between the distortion induced by that quantizer to the image and the number of bits needed to transmit the quantized subband i, $D_i(q_i) + \lambda R_i(q_i)$. For a neighborhood edge (i.e. link between two neighbor vertices), the cost is 0 if the two nodes are quantized at the same scale or β otherwise. Moreover, this cost is dynamically computed for each possible partitioning of the graph.

Once the graph construction and the energy function to be minimized have been defined, the algorithm starts with an initial (random) partitioning f (where f is a set of quantizers, $f : Q \to Q$) of the graph. For each quantizer $q \in Q$ finds $\widehat{f}$ as the quantizers repartition which minimizes $E(f')$ (i.e. $\widehat{f} \leftarrow \min E(f')$) among f' within one $q - expansion$ of f (where f' denotes the possibilites of linking the terminal node q to the planar nodes that are not linked to it in the initial f partitioning). This operation is repeated this until $E(\widehat{f})$ no longer decreases. $\widehat{f}$ is efficiently found as being the best quantizer repartition, because its cost corresponds to a minimal-cut over the constructed graph. The computation complexity is polynomial; however, in practice the running time is nearly linear for graphs with many short paths between two terminal nodes, such as the one we have modelled for our problem.

4. EXPERIMENTAL RESULTS

For our simulations, we have considered two representative test images: “Zoneplate” (512x512 pixels) and “Mandrill” (512x512 pixels), which have been selected for their texture characteristics.

We have used dead-zone scalar quantization, with $q \in \{2^0, \ldots, 2^{10}\}$ and a 5-level contourlet decomposition, where the coarsest three decomposition levels consists in a 9/7 separable wavelet transform (i.e. 3 directions) and the finest two levels are represented with a 16 and respectively 32 bands directional filter. One can remark that the algorithm can also be used with vector quantizers and the coefficient space be further partitioned into macroblocks. As shown in Fig.[3, 4], both the numerical and visual quality are improved; for the

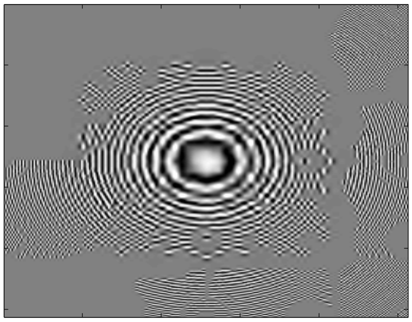

Fig. 3. Zoneplate image compressed at 0.2 bpp with JPEG2000 (PSNR = 11.12dB)

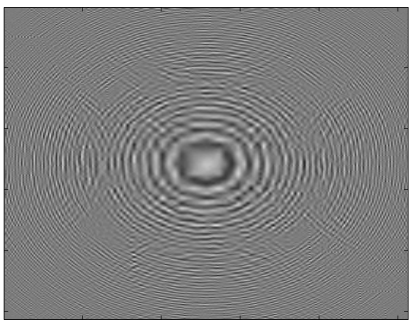

Fig. 4. Zoneplate image compressed at 0.2 bpp with the proposed method (PSNR = 12.66dB)

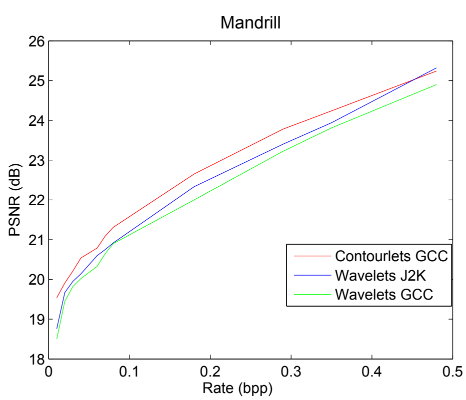

Fig. 5. Rate-distortion comparison for Mandrill image

same transmission rate (e.g. 0.2 bpp), one can remark more than 1.5 dB improvement, even though our method employs a redundant transform. Similar results are also depicted in Fig. 5. Note that for rate estimation in the allocation algorithm we have used a simple (non-contextual) arithmetic coder [15], while JPEG2000 codec uses a highly optimized contextual coder.

5. CONCLUSION

In this paper we have presented a graph-based method for rate-distortion optimization in subband image coding. The experiments show that it can encode efficiently the contourlet coefficients at low bitrates, improving both the visual and numerical quality. Moreover, the proposed method can be further used with vector quantizers and the graph design could be developed to represent the subbands at macroblock level.

6. REFERENCES

[1] "Information technology – JPEG 2000 image coding system," Tech. Rep., ISO/IEC 15444-1, 2000.

[2] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, 1989.

[3] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. on Image Proc.*, vol. 14, pp. 2091–2106, 2005.

[4] D. L. Donoho, "Wedgelets: nearly minimax estimation of edges," Tech. Rep., Statistics Department, Stanford University, 1997.

[5] X. Huo, J. Chen, and D.L. Donoho, "JBEAM: coding lines and curves via digital beamlets," *Data Compression Conference, 2004. Proceedings. DCC 2004*, pp. 449–458.

[6] M. N. Do and M. Vetterli, "Pyramidal directional filter banks and curvelets," *Proceedings of the IEEE International Conference on Image Processing*, Thessaloniki, Greece, Oct.2001.

[7] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, and P.L. Dragotti, "Directionlets: Anisotropic multi-directional representation with separable filtering," *IEEE Trans. on Image Proc.*, vol. 15, pp. 1916–1933, 2006.

[8] A. Cohen and B. Matei, "Nonlinear subdivison schemes : Applications to image processing," *Tutorials on Multiresolution in Geometric Modelling, A. Iske, E. Quak et M.S. Floater eds.,Springer Verlag*, pp. 93–97, 2002.

[9] G. Piella, B. Pesquet-Popescu, H. Heijmans, and G. Pau, "Combining seminorms in adaptive lifting schemes and applications to image analysis and compression," *Journal of Mathematical Imaging and Vision*, vol. 24, 2006.

[10] M. Wakin, J. Romberg, Hyeokho Choi, and R. Baraniuk, "Geometric tools for image compression," *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1725 – 1729, 2002.

[11] Y. Boykov, O. Veksler, and R Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1222 – 1239, 2001.

[12] L. Ford and D. Fulkerson, "Flows in networks," *Princeton University Press*, 1962.

[13] V. Kolmogorov and R Zabin, "What energy functions can be minimized via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 147 – 159, 2004.

[14] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1124 – 1137, 2004.

[15] M. Nelson and J.-L. Gailly, *The Data Compression Book*, M&T Books, New York, 1995.