Block-based graph-cut rate allocation for subband image compression and transmission over wireless networks

Maria Trocan
Institut Supérieur
d'Electronique de Paris
21 rue d'Assas, 75006 Paris
maria.trocan@isep.fr

Beatrice Pesquet-Popescu
Telecom ParisTech
37-39 rue Dareau, 75013 Paris
beatrice.pesquet@telecom-paristech.fr

James E. Fowler
Mississippi State University
U.S.A.
fowler@ece.msstate.edu

Charles Yaacoub
USEK
Liban
charlesyaacoub@usek.edu.lb

ABSTRACT

The compression of natural images and their transmission over multi-hop wireless networks still presents many challenges for the researchers and industry. In this paper we present a new block-based rate-distortion optimization algorithm that can encode efficiently the coefficients of a critically sampled, non-orthogonal or even redundant transform. 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 block-based 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.

1. INTRODUCTION

Nowadays, the majority of image-compression algorithms use wavelet transforms, attempting to exploit all the signal redundancy that can appear within and across the different subbands of a spatial decomposition. The wavelet transform has been succesfully used for image representation [1], due to its energy compaction capacities and compression efficiency [2]. However, efficiency of a coding scheme highly depends also on bit allocation. In order to maximize the compression efficiency, high-complexity subband-based image-compression schemes, as the state-of-the-art compression standard, JPEG2000 [1], may be used in wireless networks.

In this paper we present a rate-distortion optimization based on graph cuts, which can compress efficiently the coefficients of a critically sampled or even redundant, non-orthogonal transform. As described in [3, 4, 5], problems that arise in computer vision can be naturally expressed in terms of energy minimization. Each of these methods consists in modelling a graph for an energy type, such that the minimum cut minimizes globally or locally that functional. Usually, these graph constructions are dense and complex, designing the energy function at pixel level. For example, in [6] the graph cut provides a clean, flexible formulation for image segmentation. With a grid design, the graph provides a convenient manner to encode simple local segmentation decisions and presents a set of powerful computational mechanisms to extract global segmentation from these simple local (pairwise) pixel similarities. Good energy-optimization results based on graph cuts were obtained in image restoration [7], as well as in motion segmentation [8], texture synthesis in image and video [9], etc. As it will be shown by the experimental results, the method gives good compression results compared to the state-of-the-art JPEG2000 codec. The paper is organized as follows: Section 2 describes the solution for rate-distortion optimization using graph-cuts, by modeling distortion energy interactions at block level. Some experimental results obtained with the proposed methods for both wavelet-based and edge-oriented contourlet-based image coding are presented in Section 3. Finally, conclusions and future work directions are drawn in Section 4.

2. IMAGE COMPRESSION USING GRAPH CUTS

As mentioned in the introduction, we propose to use the graph-cut mechanism for the minimization of the rate- distortion Lagrangian function and thus find the optimal set of quantizers satisfying the imposed constraints. To this aim, we have designed a specialized graph able to represent a subband decomposition taking into consideration the correlations between subbands in a multiresolution approach.

In the following, we express the Lagrangian functional as a discrete sum accumulating the contribution of each coding unit (subband or block) in terms of rate and distortion induced by the quantization. Moreover, the graph model is planar, and the energy function we intend to optimize is convex, so the minimum graph cut can be found in polynomial time.

2.1 Graph design

Consider the weighted graph $G = (V, E, W)$, with V-vertices, E-edges and positive edge weights W, which have not only two, but a set of terminal nodes, $Q \in V$. Recall that a subset of edges $\mathcal{E}_C \in E$ is called a *multiway cut* if the terminal nodes are completely separated in the induced graph $G(\mathcal{E}_C) = (V, E - \mathcal{E}_C, W)$ and no proper subset of $\mathcal{E}_C$ separates the terminals in $\mathcal{E}_C$. If C is the cost of the multiway cut, then the multi-terminal min-cut problem is equivalent to finding the minimum-cost multiway cut. For our optimization problem, the terminals are given by a set of quantizers Q, and the coding units give the rest of the vertices $V - Q$. The edges and their weights/capacities will be defined in the following depending on the coding strategy (subband or block coding) and the distortion functional.

In [7], Y. Boykov *et al.* find the minimal multiway cut by succesively finding the min-cut between a given terminal and the other terminals. This approximation guarantees a local minimization of the energy function that is close to the optimal solution for both concave and convex energy functionals. As the rate-distortion Lagrangian lies on a convex curve (i.e. $D(R)$), we propose to use the method in [7] for its optimization.

2.2 Lagrangian rate-distortion functional

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 (spatial subbands or blocks of coefficients), each of them coded with a different quantizer q_i, $q_i \in Q$ (Q being the quantizers set). Let $D_i(q_i)$ be the distortion of the coding unit i when quantized with q_i, and let $R_i(q_i)$ be the number of bits required for its encoding. The problem can now be formulated as: find $\min \sum_i D_i(q_i)$, such that $\sum_i R_i(q_i) = R \le R_{max}$.

In the Lagrange-multiplier framework, this constrained optimization is written as the equivalent problem:

$$\min \sum_i \left(D_i(q_i) + \lambda R_i(q_i)\right), \quad R \le R_{max} \tag{1}$$

where the choice of the Lagrangian parameter $\lambda > 0$ measures the relative importance between distortion and rate for the optimization and which can be determined using a binary search. The advantage of problem formulation in Eq. (1) is that the sum and the minimum operator can be exchanged to:

$$\sum_i \min \left(D_i(q_i) + \lambda R_i(q_i)\right), \quad R \le R_{max} \tag{2}$$

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

2.2.1 Rate estimation

For the rate estimation of the quantized coding units we consider a non-contextual arithmetic coder [10], which uses a zero-order entropy model, where the M quantized coefficients of a given coding unit are random i.i.d. variables following a Gaussian distribution. Thus, the zero-order entropy (H) estimation in bits/variable (i.e., coefficient) is obtained as:

$$H = -\sum_{i=1}^{M} p_i \log_2 p_i, \tag{3}$$

where p_i is the probability of the i^{th} coefficient. The resulting entropy estimate per coding unit is weighted by the size of the coding unit in order to obtain the total entropy of the quantized image.

2.2.2 Distortion estimation

The distortion D between the original image x and the quantized one, $\hat{\mathbf{x}}$, is estimated in the following as the L^2 norm, i.e. :

$$D = \|\boldsymbol{x} - \widehat{\boldsymbol{x}}\|^2 . \tag{4}$$

This model will be futher developed, in order to obtain a good distortion estimate in the spatial domain, rather than in the transform domain, as is usually done for orthonormal transforms.

2.3 Graph design with cross-correlation distortion at the block level

Recall that we have written the distortion D between the original image, x, and the quantized one, $\widehat{x}$, as the L^2 norm, i.e. $D = \|\boldsymbol{x} - \widehat{\boldsymbol{x}}\|^2$. In a first approximation [11], we have considered only the diagonal terms, i.e.:

$$D_I \cong \sum_i \|x_i - \widehat{x}_i\|^2 \tag{5}$$

which amounts to estimating the distortion between the contribution to the image and to the quantized image of only the i^{th} subband.

In a second approximation, we have also considered the *cross-correlation* terms, i.e.:

$$D \cong D_I + \sum_i \sum_{i' \in \mathcal{N}(i)} \langle \widehat{x}_i - x_i, \widehat{x}_{i'} - x_{i'} \rangle \tag{6}$$

where $\mathcal{N}(i)$ is a neighborhood of i, containing closely correlated subbands. Indeed, given the limited support of the wavelets, the closer in scale and frequency are the subbands, the higher the correlation among them. In practice, this neighborhood could be described by the geometrical position of the subbands in a multiresolution decomposition (where only the vertical and horizontal directions are considered), or by simply linking the subbands in a chain-manner, one after another (for example, in Fig. 1, the neighborhood relations are indicated by the black edges in the graph). Thus, Eq. (6) can be written as:

$$D = \sum_i \underbrace{\|x_i - \widehat{x}_i\|^2}_{D_i} + \sum_i \sum_{i' \in \mathcal{N}(i)} \underbrace{\langle \widehat{x}_i - x_i, \widehat{x}_{i'} - x_{i'} \rangle}_{D_{i,i'}} \tag{7}$$

We have shown in [12] that in this case, the function to be

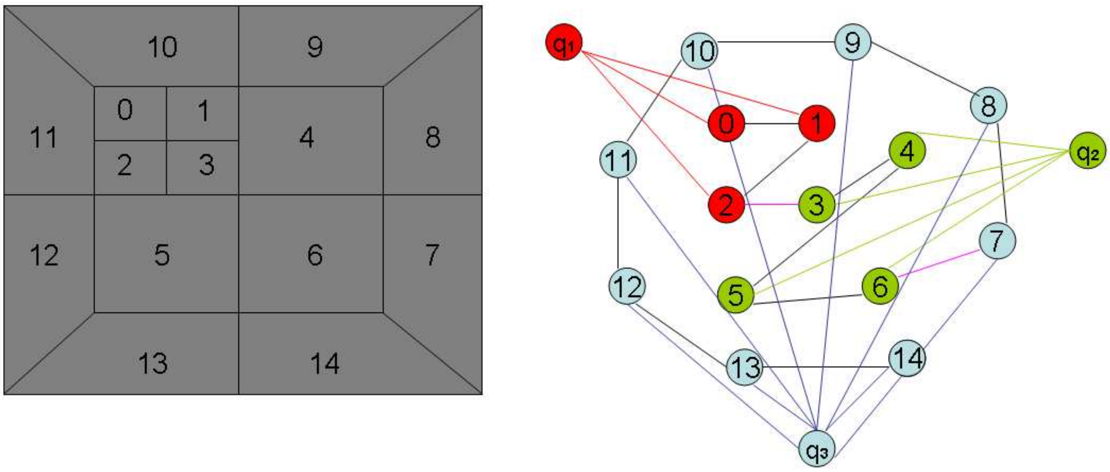

Figure 1: Contourlet decomposition with three levels (left) and three-way graph-cut repartition (right) (q_1 partition in red, q_2 partition in green, q_3 partition in blue, where the regular edges are with full black lines, terminal links in colors and the cut-edges in gray lines).

minimized is:

$$\min \sum_i \left(\underbrace{\|x_i - \widehat{x}_i\|^2 + \lambda R(i)}_{E_{data}(i)} + \underbrace{\sum_{i' \in \mathcal{N}(i)} \langle \widehat{x}_i - x_i, \widehat{x}_{i'} - x_{i'} \rangle}_{E_{smooth(i)}} \right) \quad (8)$$

In the following, we propose to extend the subband level distortion estimation presented in [12] to the block level (Fig. 2). This extension comes naturally, as the smaller the coding unit, the more correlated in amplitude are the coefficients within it. At block level, Eq. (8) becomes:

$$\min \sum_{i=1}^{X} \sum_{j=1}^{N_b} \underbrace{\|x_{i,j} - \widehat{x}_{i,j}\|^2 + \lambda R(i,j)}_{E_{data}(i,j)} + \underbrace{\sum_{(i',j') \in \mathcal{N}(i,j)} |\langle \widehat{x}_{i,j} - x_{i,j}, \widehat{x}_{i',j'} - x_{i',j'} \rangle|}_{E_smooth(i,j)} \quad (9)$$

where X, respectively N_b represent the number of subbands, respectively blocks in each subband, $x_{i,j}$ denotes the image reconstucted only from the j^{th} block of the i^{th} subband and $\langle \widehat{x}_{i,j} - x_{i,j}, \widehat{x}_{i',j} - x_{i',j'} \rangle$ measures the correlation between the neighbour blocks.

The minimization of the energy function defined above is equivalent to the best partition of quantizers per subbands blocks. Note that for E_{smooth} we have used the sum of absolute values of cross-correlation terms, in order to ensure that our regular vertices will have associated positive weights. Our graph will have therefore $B = X \times N_b - 1$ regular vertices. The neighbourhood system, $\mathcal{N}$, contains

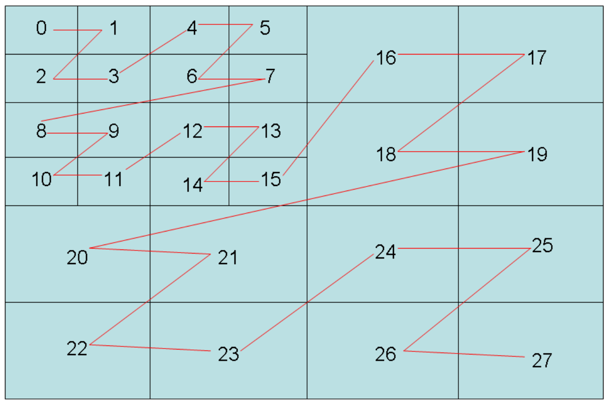

Figure 2: Block graph design: two-level wavelet decomposition with four-blocks subband division and chain network design for the regular vertices.

now only position correlation links E_N(i.e., edges between neighbour blocks, as described in Fig.2). The geometrical model can be described as: $G = (V, E)$ where $V = B \cup Q$, $E = E_N \cup E_Q$ and Q/E_Q represent the quantizers set/the links between block nodes and quantizers. For the terminal links, E_Q, the weights are given by the direct costs in terms of distortion and rate induced by the quantization (i.e., the edge between block b and quantizer q, (b, q), has the associated weight $w_{b,q} = D_b(q) + R_b(q)$). The capacity between two regular neighbour blocks ($(b_i, b_{i'}) \in E_N$) is defined as the absolute value of the cross-correlation distortion induced by the current quantization of these blocks.

3. APPLICATION TO SUBBAND IMAGE COMPRESSION

In the following, we propose to apply the proposed graph-cut minimization model to subband image compression. Some results are drawn in the framework of classical separable wavelet image coding, as well as for a geometrical transform, namely the contourlet decomposition [13]. Note that the method can be applied to almost any existing decomposition (wavelets, Xlets, subbands, blocks, may them be critically sampled / redundant etc.).

3.1 Wavelet image compression with graph-cuts

Due to their energy compaction efficiency, the biorthogonal filter banks are the most used in image compression [1]. This is the reason for which we consider in our simulation framework both the 5/3 and 9/7 filter banks for the spatial decomposition.

3.1.1 *Experimental results*

For our simulations, we have considered two representative test images: Barbara (512x512 pixels) and Mandrill (512x512 pixels), which have been selected for the difficulty to encode their texture characteristics.

We have used dead-zone scalar quantization, with $q \in \{2^0, \ldots, 2^{10}\}$. The dead-zone has twice the width of the other quantization intervals. All the images have been decomposed over five spatial levels with the floating-point 5/3 and 9/7 filter banks. Note that for rate estimation in the allocation algorithm we have used a simple (non-contextual) arithmetic coder [10], while JPEG2000 codec [1] uses a highly optimized contextual coder. The JPEG2000 results have been obtained with the Kakadu framework.

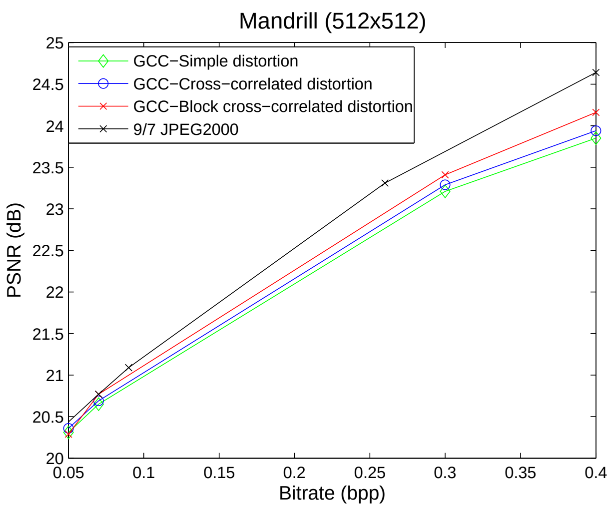

Figure 3: Rate-distortion comparison for Mandrill image with 9/7 wavelet subband decomposition

As it can be remarked from Fig. 3 and Fig. 4, the results obtained with the 9/7 wavelet subband decompostion of JPEG2000 are between 0.5 and 1.5 dB higher than those obtained with the proposed graph-cut rate-distortion algorithm. This situation can be explained by the fact that the 9/7 filter bank is very close, from an energy partition point of view, to an orthonormal decomposition. As illustrated in Fig. 5 and Fig. 6, our method seems to better cope with non-orthogonal decompositions at very low bitrates (≤ 0.1 bpp).

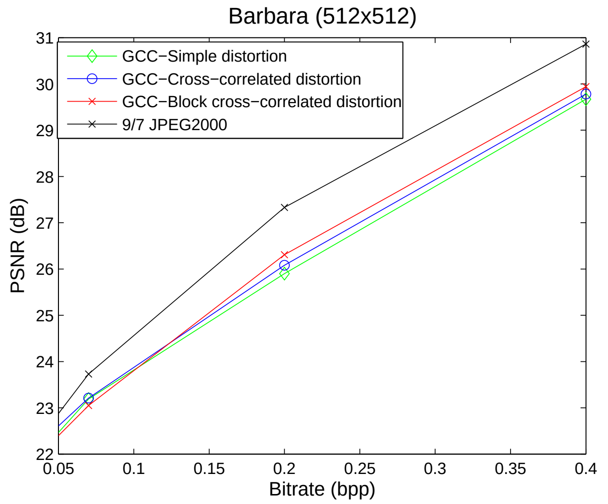

Figure 4: Rate-distortion comparison for Barbara image with 9/7 wavelet subband decomposition

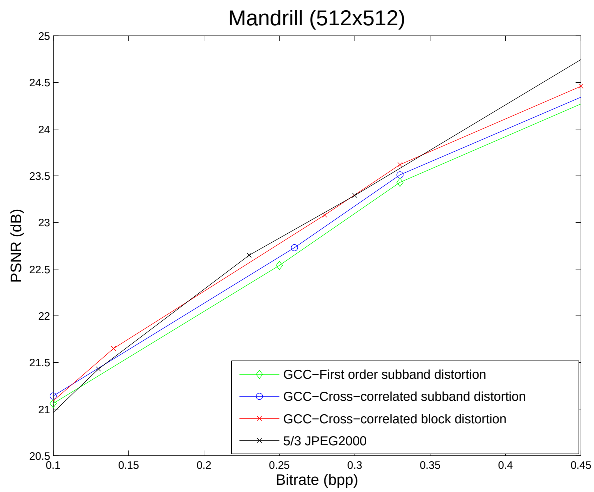

Figure 5: Rate-distortion comparison for Mandrill image with 5/3 wavelet subband decomposition

One can remark that distortion approximation at subband level taking into account the cross-correlation among subbands always leads to better results than the simple model without cross-correlation terms, by using a more realistic correlation model. Moreover, the finer level of representation for the coding units, the higher the correlation among these units, as it can be remarked from the presented results, having an average gain of 0.25 dB over the preceeding rate-distortion curve obtained with a subband-level cross-correlated distortion model.

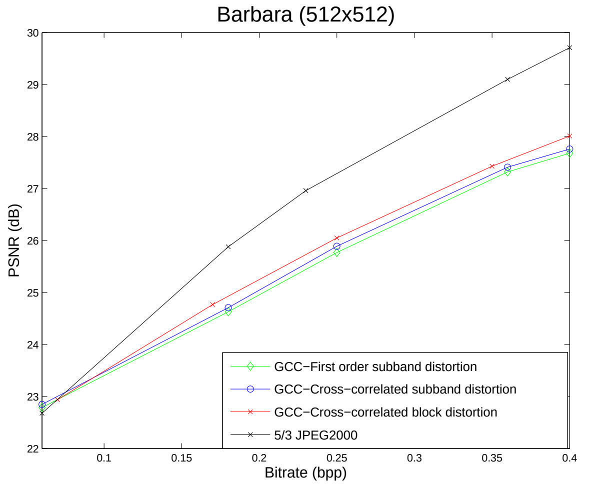

Figure 6: Rate-distortion comparison for Barbara image with 5/3 wavelet subband decomposition

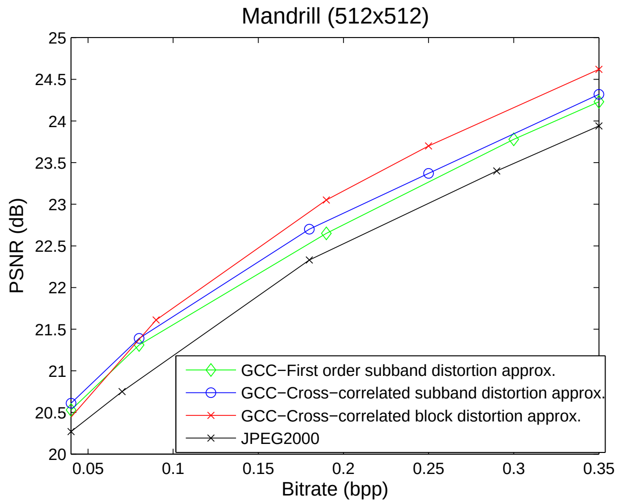

Figure 7: Rate-distortion comparison for Mandrill image with contourlet subband decomposition

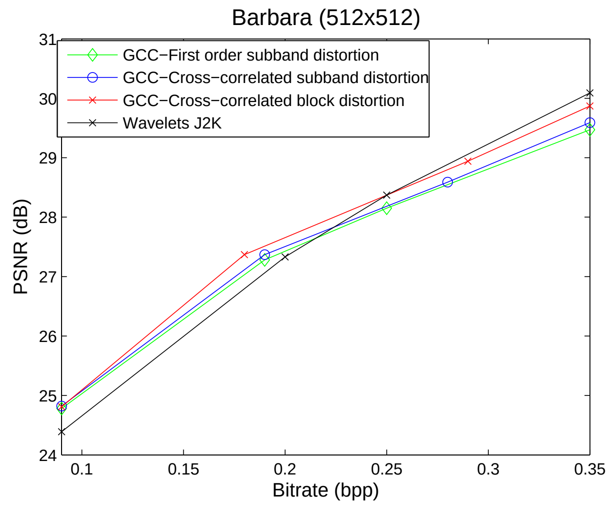

Figure 8: Rate-distortion comparison for Barbara image with contourlet subband decomposition

3.2 Contourlet image compression with graph-cuts

The drawback of separable wavelets is the limited orientation selectivity, as they fail to capture the geometry of the image edges. 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 [13]. With contourlets, one can represent the class of smooth images with discontinuities along smooth curves in a very efficient and sparse way. These decompositions have been successfully applied in image segmentation and noise removal, as well as in image compression: as shown in [14], the codec based on wedgelets gives better performance in image compression than the JPEG2000 standard at very low rate.

3.2.1 Experimental results

For a better comparison, we have considered the same test images: Barbara (512x512 pixels) and Mandrill (512x512 pixels). We have used dead-zone scalar quantization, with $q \in \{2^1, \ldots, 2^{10}\}$ and a 5-level contourlet decomposition, where the coarsest three decomposition levels consist of a 9/7 separable wavelet transform (i.e., 3 directions), and the finest two levels are represented with a 16- and 32-band biorthogonal directional filter. The efficiency of this hybrid scheme has been proved in [15].

As shown in Figs. 7 and 8, our method surpasses JPEG2000 at low bitrates, even though it employs a redundant transform. Note that for the rate estimation in the allocation algorithm we have used a simple (non-contextual) arithmetic coder [10], while JPEG2000 codec uses highly optimized contextual coder.

4. CONCLUSION

In this paper we have presented a block-based graph-cut method for rate-distortion optimization in image coding. Its great advantage is that it can be applied to decompositions which are not necessarily orthonormal. As shown by experimental results, it can efficiently encode both wavelet and contourlet coefficients compared to standard RD coding tools, enhancing thus the wireless transmission efficiency. Moreover, the proposed method could be further used with vector quantizers.

5. 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] 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.

[4] 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.

[5] Y. Boykov and O. Veksler, "Graph cuts in vision and graphics: Theories and applications," *Handbook of Mathematical Models in Computer Vision, Springer-Verlag*, pp. 79–76, 2006.

[6] Y. Boykov and G.Funka-Lea, "Graph cuts and efficient N-D image segmentation," *International Journal of Computer Vision*, vol. 70, pp. 109–131, 2006.

[7] 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.

[8] T. Schoenemann and D. Cremers, "Near real-time motion segmentation using graph cuts," *K. Franke et al.(Eds.): DAGM 2006, Springer-Verlag*, pp. 455–464, 2006.

[9] Y. Zhang, X. Ji, D. Zhao, and W. Gao, "Video coding by texture analysis and synthesis using graph cut," *In Proceedings of Pacific-Rim Conference on Multimedia, China*, 2006.

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

[11] M. Trocan, B. Pesquet-Popescu, and J.E. Fowler, "Graph-cut rate distortion algorithm for contourlet based image compression," *Proc. of the IEEE Int. Conf. on Image Processing, San Antonio, Texas U.S.A.*, Sept. 2007.

[12] M. Trocan and B. Pesquet-Popescu, "Graph-cut rate distortion optimization for subband image compression," *Proc. of the IEEE European Signal Processing Conference, EUSIPCO'07*, Poznan, Poland, Sept. 2007.

[13] 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.

[14] 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.

[15] V. Chappelier, C. Guillemot, and S. Marinkovic, "Image coding with iterated contourlet and wavelet transforms," *Proc. of IEEE International Conference on Image Processing*, vol. 5, pp. 3157– 3160, 2004.