Influence of JPEG2000 Standard Compression on Color Image Segmentation Quality

Christophe CHARRIER and Olivier LEZORAY

Université de Caen - Basse Normandie
LUSAC EA-2607, Groupe Vision et Analyse d'Images
120, route de l'exode, 50000 Saint-Lô, France
Phone : +33 (0)2 33 77 11 66
Fax : +33 (0)2 33 77 11 67
E-mail: c.charrier@chbg.unicaen.fr, Olivier.Lezoray@info.unicaen.fr

Abstract

To preserve real time application of retrieval process for wide color images (e.g., biomedical, radar or satellite images), the influence of the JPEG2000 compression scheme on the quality of segmentation process is investigated. To perform this investigation, four color segmentation techniques have been applied on five different color images. These images are composed of the original image and four compressed images at different rates. The four retained segmentation schemes are 1) color watershed, 2) color watershed followed by a RAG merging, 3) color quadtree and 4) color quadtree followed by a RAG merging. To assess the effects of both the compression and the compression rate on the results of the segmentation, the accuracy of a given segmentation at a fixed compression rate is quantized using three error metrics: the Mean Average Error, the PSNR and the MSE (Mean Square Error).

The obtained results demonstrate that a compression level equal to 0.75 bpp can be reached whitout affect the segmentation process.

1 Introduction

Information research and/or indexing process from database of digital images are actually highly challenging problems. The amount of manipulated data exponentially increases, especially in the case of biomedical (e.g., cell analysis, tumor cell detection, etc...) radar and satellite images. For those kinds of images, the size can easily reach 40 Mb and more. Thus for e-learning applications and/or request on accessible database *via* Internet, the size of large images may imply a growth of the response time. Actually, typical image retrieval schemes can be decomposed in three steps: 1) the computation of the feature vector from the request image, 2) the search of similar vectors according to the minimization of specific criteria and 3) the display of associated images to these similar vectors. In order to preserve the real time characteristic of these applications, the first step has to be as quick as possible, even in the case of large request images. One can easily get the picture of the consequences of high response time in the case of teleoperation for example.

One way to preserve the real time operation of such applications is to use a compression technique in order to reduce the size of the request image. This scheme presents, *a priori*, two major advantages: 1) a gain of the space needed to store wide images, and 2) a reduction of the transmission requirements of result images. Yet, data compression is an encoding process used to reduce the amount of information while preserving an "acceptable quality". The major problem addressed is the characterization of the term "acceptable quality". Indeed, in order to obtain interesting compression rate, *i.e.*, superior to 10:1 bpp, a lossy compression is unescapable and thus degradations appear within reconstructed images. Under such assumptions, one has to determine the impact of the reduction of the original amount of information on the quality of the feature vector construction. The challenge is to determine the tradeoff between the compression level and the quality of the feature vector, *i.e.*, the quality of the result of the segmentation.

In this paper, the compression scheme used is the baseline JPEG2000 standard [1, 2]. Actually, this algorithm allows to reach high compression levels while maintaining high image quality [3]. Four segmentation methods (among the most widely used) have been investigated to determine the influence of the JPEG2000 compression algorithm on the results of trial segmentation schemes. These schemes are: color watershed, color quadtree and both associated

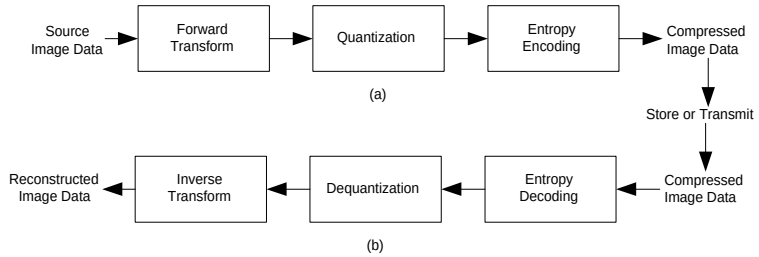


Figure 1: Block diagram of the baseline JPEG-2000 codec (source [3]).

with Region Adjacency Graph. To measure the quality of the results, three different segmentation error metrics have been used.

2 The baseline JPEG2000 Compression Scheme

In March 1997 a new call for contributions was launched for the development of a new standard for the compression of still images: the JPEG2000. This project, JTC2 1.29.14 (15444), was intended to create a new image coding system for different types of still images (bi-level, gray-level, color, multi-component), with different characteristics (natural images, scientific, medical, remote sensing, text, rendered graphics, etc) allowing different imaging models (client/server, real-time transmission, image library archival, limited buffer and bandwidth resources, etc) preferably within a unified system.

The JPEG2000 standard provides a set of features that are of importance to many high-end and emerging applications by taking advantage of new technologies. It addresses areas where current standards fail to produce the best quality or performance and provides capabilities to markets that currently do not use compression. The markets and applications better served by the JPEG2000 standard are the Internet, color facsimile, printing, scanning (consumer and prepress), digital photography, remote sensing, mobile, medical imagery, digital libraries / archives and Ecommerce. Each application area imposes some requirements that the standard should fulfill.

For the most part, the JPEG-2000 standard is written from the point of view of the decoder. That is, the decoder is defined quite precisely with many details being normative in nature (i.e., required for compliance), while many parts of the encoder are less rigidly specified [1]. So, the baseline architecture can be described by the block diagram presented in Figure 1. The encoder proceeds as follows (Fig. 1(a)). The discrete wavelet transform is first applied on the source image data. The transformed coefficients are then quantized and entropy coded, before forming the output codestream (bitstream). The decoder is the reverse of the encoder (Fig. 1(b)). The codestream is first entropy decoded, dequantized and inverse discrete transformed, thus resulting in the reconstructed image data.

At first, it should be mentioned that the standard works on image tiles. The term “tiling” refers to the partition of the original (source) image into rectangular non-overlapping blocks (tiles), which are compressed independently, as though they were entirely distinct images (Fig. 2). Prior to computation of the forward discrete wavelet transform (DWT) on each image tile, all samples of the image tile component are DC level shifted by subtracting the same quantity (i.e. the component depth). DC level shifting is performed on samples of components that are unsigned only. If color transformation is used, it is performed prior to computation of the forward component transform.

Otherwise it is performed prior to the wavelet transform. Arithmetic coding is used in the last part of the encoding process. The MQ coder is adopted in JPEG2000. This coder is basically similar to the QM-coder adopted in the original JPEG standard [4]. The MQ-coder is also used in the JBIG-2 standard [5].

It should be noted here that the basic encoding engine of JPEG2000 is based on EBCOT (Embedded Block Coding with Optimized Truncation of the embedded bitstreams) algorithm, which is described in detail in [6, 7].

3 Color Image Segmentation Schemes

To segment color images, different segmentation schemes well-known in the literature are selected: watershed, quadtree and both associated with Region Adjacency Graph. The principles of each one of these schemes will now be reviewed.

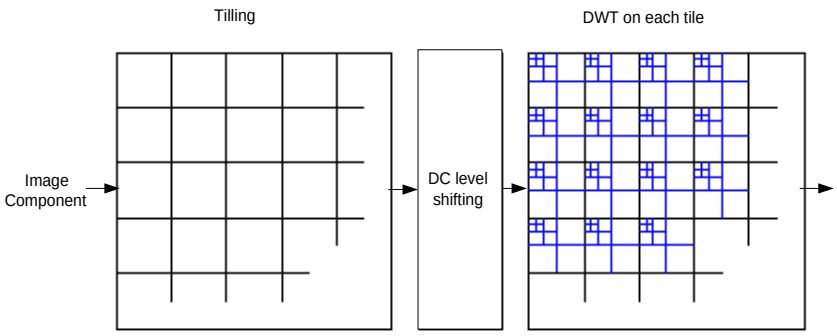


Figure 2: Tiling, DC level shifting and DWT of each image tile component(source [3]).

3.1 Color watershed

The first scheme proposed to segment color images is similar to the one proposed in [8] except that the simplification step has been removed (in order to see the effects of the compression). The segmentation acts in three stages. First the color gradient is computed and the minima of the gradient are extracted. From these minima, a color watershed is performed to obtain the final image partition. To compute the gradient of a color image I, Di Zenzo [9] considers this image as a surface and looks for its local variations of $\|dI^2\|$:

$$\|dI^2\| = \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix}^T \begin{bmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \qquad \text{with} \qquad g_{ij} = \frac{\partial I}{\partial x_i}\frac{\partial I}{\partial x_j}$$

The two eigenvalues of g_{ij} are the extrema of $\|dI^2\|$ and the orthogonal eigenvectors η, ξ are the corresponding variation directions:

$$\begin{cases} \lambda_{+/-} = \left(g_{11} + g_{22} \pm \sqrt{(g_{11} - g_{22})^2 + 4g_{12}^2}\right)/2 \\ \eta = \frac{1}{2}\arctan(2g_{12}/(g_{11} - g_{12})) \\ \xi = \eta + \pi/2 \end{cases}$$

λ_+ and λ_- respectively correspond to the maximum and minimum local variations in the η and ξ directions. Then, several vector gradient norms $\|\nabla I(p)\|$ can be defined:

$$\|\nabla I(p)\| = \begin{cases} \sqrt{\lambda_+ - \lambda_-} & [10] \\ \sqrt{\lambda_+ + \lambda_-} & [11] \\ \sqrt{\lambda_+} & [12] \end{cases}$$

The first norm was proposed in [10] and fails to detect corners where $\lambda_+ = \lambda_-$. The second norm was proposed in [11] and gives more importance to some corners (but not all). The third norm is a direct extension of the gradient norm definition (the value of the maximum direction) and doesn't favor any corner [12]. The latter gradient norm has to be preferred to the others. From the gradient image, the h-minima are extracted: they correspond to the minima of height h in the image. This operation is a classic mathematical morphology one and is performed by a reconstruction of the $\|\nabla I(p)\|$ image by the $\|\nabla I(p)\| - h$ image. In this paper, h is fixed to 3 to extract all the reliable minima.

From these minima, a region growing method is used to achieve the flooding of the initial markers (the minima) into all the image. To this aim a color watershed [13] using local and global criterion is used. The first criterion is expressed as the gradient norm at the location of a point p and the second one is expressed by the difference between the mean color of a region and the color of the point p. These two criteria are blended together and provide a trade-off between preserving the mean global color of a region and letting the region grow towards its natural boundary (depending on the gradient). This is expressed by the following aggregation function which defines the aggregation probability of a pixel to a neighbor region. Let $\overline{I(R_i)}$ denote the mean color vector of the region R_i for the image I, the $I(p)$ vector giving the color of a pixel p and $\nabla I(p)$ the color gradient. The aggregation function is expressed as [14, 15]:

$$f(p, R_i) = (1 - \alpha)\|\overline{I(R_i)} - I(p)\| + \alpha\|\nabla I(p)\|$$

At the beginning of the growth, only the initial minima are labeled. Progressively unlabeled pixels adjacent to a region are processed and merged until the total flooding of the image. The initial value of α is 0 and at each iteration k, $V_k = \sum f(p, R)$ is computed for all the unlabeled pixels p. V_0 gives the initial value for all the unlabeled pixels of the image. At a given iteration k and after the processing of all the considered unlabeled pixels, V_k is computed and a new value for α is given by $\alpha = V_k/V_0$. To avoid high variations of α between the iterations, α is considered to be the mean of all the previous values of α, including the new computed one [15].

3.2 Color quadtree

Quadtree segmentation is known to produce simple image partitions [16], subject to rigid geometrical constraints. It segments the image into different regions according to a uniformity criterion. The obtained regions are all based on square non-overlapping blocks. The principle is the following: if a square block is not homogeneous (i.e its uniformity is lower than a given threshold), the block is split into four equal ones and the algorithm iterates on the obtained blocks. In this paper the uniformity criterion used is the following:

$$UC(R_i) = \sum_{j=1}^{3} \left[1 - \left(\frac{\sigma(I_{C_j}(R))}{\|\overline{I_{C_j}(R)}\|} \right)^2 \right]$$

where I_{C_j} denotes the jth channel of the image I. σ and $\|\cdot\|$ respectively denote the standard deviation and the mean of the color channel. Square regions having a uniformity criterion lower than 2.9 are split.

(a) (b) (c) (d)

Figure 3: Examples of the COLUMBIA image database

3.3 Region Adjacency Graph merging

The two above segmentation schemes have a serious drawback since they tend to over-segment images. A merging strategy is therefore needed to simplify the segmentation into fewer regions (with preserved color homogeneity). We have used the method proposed by Mumford and Shah [17] based on the following principle. It operates directly on the RAG of the initial segmentation [18, 19, 20]. For each pair of adjacent regions the energy provided by their fusion is computed. The two adjacent regions providing the minimum of energy reduction are merged since they are the most similar pair of regions of the whole image. This merging operation is iteratively performed. Adjacent regions having an energy higher than a threshold (10^4 in this paper) are not merged. The energy variation is defined as follows:

$$\Delta E = \frac{\Theta(R_i) \cdot \Theta(R_j)}{\Theta(R_i) + \Theta(R_j)} \cdot \|\overline{I(R_i)} - \overline{I(R_j)}\|^2 - 2 \cdot \alpha \cdot \delta(R_i, R_j)$$

where $\Theta(R_i)$ denotes the number of pixels of the region R_i and $\delta(R_i, R_j)$ denotes the number of boundary pixels shared by the two regions R_i and R_j. α is fixed to 0.2.

3.4 Segmentation schemes

Four segmentation schemes have been retained in this paper. They operate on color images and each scheme uses different assumptions on the color image: block homogeneity, region homogeneity and edge preservation, reduction of the over-segmentation with the Mumford-Shah energy. The four retained segmentation schemes are 1) color watershed, 2) color watershed followed by a RAG merging, 3) color quadtree and 4) color quadtree followed by a RAG merging. Using different segmentation schemes will enable us to assess the effect of compression on each one of the above-mentioned schemes. All the more, since they use different assumptions on the image to segment it, these properties

Figure 4: first row: compressed images, row 2: LPE, row 3: LPE+RAG, row 4: Quadtree, row 5:Quadtree+RAG

might be modified according to the image compression rate and we propose to study the effect of compression on the image segmentation result.

4 Segmentation Quality Evaluation

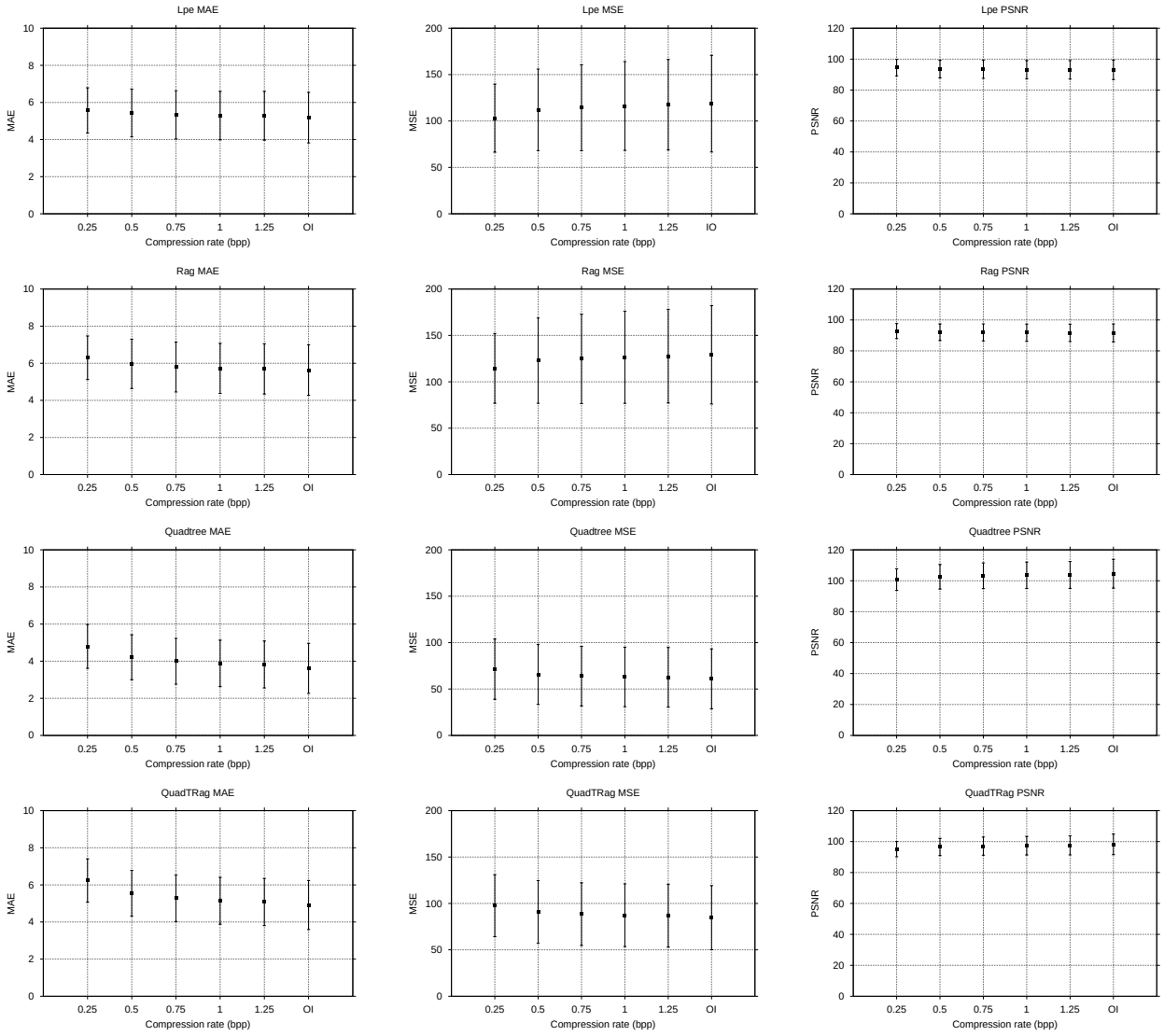
4.1 Experimental Setup

The used color images come from the COLUMBIA image database [21]. Figure 3 presents a sample of this database. Five different compression rates have been used : 0.25, 0.5, 0.75, 1 and 1.25 bpp. These levels range from a poor to an acceptable image quality with respect to the quality scale defined within the recommendation UIT-R BT.500-7 [22].

Figure 4 presents an example of obtained results when applying the four segmentation schemes on the original image and the compressed images according to the compression levels used.

4.2 Error metrics

To assess the effects of both the compression and the compression rate on the results of the segmentation, we have to quantify the accuracy of a given segmentation at a fixed compression rate. The image is sequentially compressed, uncompressed and segmented. The evaluation of a segmentation result is a difficult task without any ground truth



(a) (b) (c)

Figure 5: Results obtained for (a) MAE, (b) MSE and (c) PSNR

reference and a single numerical measure cannot capture the whole description of the image provided by the segmentation. We propose to use three different segmentation error metrics: the MAE (Mean Average Error), the PSNR (Peak Signal to Noise Ratio) and the MSE (Mean Square Error) [23]. They are defined by the following expressions, where I and J denote two color images of size $h.w$:

$$MAE(I,J) = \frac{1}{3.h.w}\sum_{k=1}^{3}|I_{C_k} - J_{C_k}| \tag{1}$$

$$MSE(I,J) = \frac{1}{3.h.w}\sum_{k=1}^{3}\|I_{C_k} - J_{C_k}\| \tag{2}$$

$$PSNR(I,J) = 10log_2\left(\frac{255^2}{MSE(I,J)}\right) \tag{3}$$

where I_{C_k} and J_{C_k} respectively represent the kth color channel of the image I and J, and $|.|$ and $\|.\|$ respectively denote the L_1 and L_2 norms.

If an image is compressed at a given rate, a different segmentation can be obtained for each one of the four segmentation schemes above-mentioned. A segmentation is a region map and from this latter a quantized image can be created where the color associated to each pixel is the mean color value of its surrounding region. If the segmentation is considered as stable, the quantized image for a given segmentation rate might be as close as possible to the quantized one obtained by the same segmentation on the original image. This is assessed by measuring the error between the reconstructed quantized image and the original one (not compressed). If the compression has no or small effects on the segmentation then the error remains stable: the obtained error is close to the error obtained by segmenting the original image. On the other hand, a variation of the error metric states that the compression has altered the segmentation.

4.3 Results

The obtained results are shown in Figure 5. These results represent the mean and the standard deviation of the three error metrics used to quantify the quality of the segmentation in respect with the compression rate. The first remark is obvious: the compression directly influences the quality of the segmentation, and this whatever the considered method. Nevertheless, several remarks can be formulated depending on the used segmentation scheme:

- a stability for each one of the three error metrics can be observed between the original image and the compression one with a level equal to 0.75 bpp,
- below this rate, a significant change from curves is visible,
- generally, measures increase with the compression rate. Nevertheless, for some results, a decrease of values can be observed, especially for the color watershed MSE and the color watershed followed by a RAG merging MSE (two first figures of the column (b)). This can be explained by the following remark: when compressing images, a loss of data is implied. Thus for high level compression, a severe loss of information is performed. From the compression side, this implies that small areas have disappeared in bigger regions. When applying any one of the tested segmentation schemes, the error concerning the classification of pixels that belong to these regions decreases, and consequently the error metric too (Cf. second row of Figure 4, where one observes less regions for a compression level equal to 0.25 bpp than for 0.75 bpp). Yet, the compressed image has noticeable and annoying artefacts (Cf. first row of Figure 4) that badly influence the segmented image in respect with the original segmented image.

  Similar remarks can be formulated for the quadtree PSNR and the quadtree followed by a RAG merging PSNR (two last figures of the column (c)).

From these remarks, a compression rate equal to 0.75 bpp can be reached without degrading the quality of the segmentation in respect with the obtained results from the original image.

5 Conclusion

The influence of the JPEG2000 compression scheme on the quality of segmentation process has been studied in order to preserve real time application of retrieval process for wide color images. To perform this investigation, four color segmentation techniques have been applied on five different color images. These images are composed of the original image and four compressed images at different rates. The four retained segmentation schemes are 1) color watershed, 2) color watershed followed by a RAG merging, 3) color quadtree and 4) color quadtree followed by a RAG merging.

The obtained results demonstrate that a compression level equal to 0.75 bpp can be reached without degrading the quality of the segmented image according to those obtained from the original image.

References

[1] ISO/EIC JTC 1/SC 29/WG1, “JPEG2000 PartI final draft international standard,” tech. rep., N1855, August 18 2000.

[2] ISO/IEC 15444-2, “Information technology JPEG 2000 image coding system: Extensions,” tech. rep., ISO/EIS JTC1/SC29/WG1, Aug. 2002.

[3] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: an overview," *IEEE Transactions on Consumer Electronics*, vol. 46, pp. 1103–1127, Nov. 2000.

[4] G. K. Wallace, "The JPEG still image compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, Feb. 1992.

[5] I. J. N1359, "Information technology - coded representation of picture and audio information - lossy/lossless coding of bi-level images," tech. rep., 14492 Final Committee Draft, July 1999.

[6] D. Taubman, "High performance scalable image compression with EBCOT," in *Proc. IEEE Int. Conference Image Processing*, vol. III, (Kobe, Japan), pp. 344–348, Oct. 1999.

[7] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, pp. 1158–1170, July 2000.

[8] O. Lezoray, A. Elmoataz, H. Cardot, and M. Revenu, "A color morphological segmentation," in *Proceedings of International Conference on Color in Graphics and Image Processing*, pp. 170–175, 2000.

[9] S. DiZenzo, "A note on the gradient of a multi-image," *Computer Vision, Graphics and Image Processing*, vol. 33, pp. 116–126, 1986.

[10] G. Sapiro and D. Ringach, "Anisotropic diffusion of multivalued images with applications to color filtering," *IEEE transactions on Image Processing*, vol. 5, no. 11, pp. 1582–1585, 1996.

[11] P. Blomgren and T. Chan, "Color tv: Total variation methods for restoration of vector-values images," *IEEE Transactions on Image Processing, Special Issue on Partial Differential Equations and Geometry-Driven diffusion in Image Processing and Analysis*, vol. 7, no. 3, pp. 304–309, 1998.

[12] D. Tschumperle and R. Deriche, "Restauration d'images vectorielles par edp," in *Proceeding of RFIA*, pp. 247–256, 2000.

[13] L. Vincent and P. Soille, "Watersheds in digital spaces : An efficient algorithm based on immersions simulations," *IEEE transactions on PAMI*, vol. 13, no. 16, pp. 583–598, 1991.

[14] P. Belhomme, A. Elmoataz, P. Herlin, and D. Bloyet, "Generalized region growing operator with optimal scanning : application to segmentation of breast cancer images," *Journal of Microscopy*, vol. 186, pp. 41–50, 1997.

[15] O. Lezoray and H. Cardot, "Histogram and watershed based segmentation of color images," in *Proceedings of CGIV'2002*, pp. 358–362, 2002.

[16] H. Samet and R. Weber, "On encoding boundaries with quadtrees," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 3, pp. 365–369, 1984.

[17] D. Mumford and J. Shah, "Boundary detection by minimizing functionals," in *Proceedings of CVPR*, vol. 1, pp. 22–26, 1985.

[18] K. Saarinen, "Color image segmentation by a watershed algorithm and region adjacency graph processing," in *Proceedings of ICIP*, vol. 3, pp. 1021–1025, 1994.

[19] L. Garrido, P. Salembier, and D. Garcia, "Extensive operators in partition latties or image sequence analysis," *Signal Processing*, vol. 6, no. 2, pp. 157–180, 1998.

[20] K. Haris, S. Efstratiadis, and N. Maglaveras, "Hybrid image segmentation using watersheds and fast region merging," *IEEE transactions on Image Processing*, vol. 7, no. 2, pp. 1684–1699, 1998.

[21] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (COIL-100)," tech. rep., CUCS-006-96, Feb. 1996.

[22] R. U.-R. BT.500-10, "Méthodologie d'évaluation subjective de la qualité des images de télévision," tech. rep., UIT, 2000.

[23] K. N. Platoniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*. Springer, 2000.