Region-Based Image Retrieval Using a Joint Scalable Bayesian Segmentation and Feature Extraction

Sarra Sakji-Nsibi, Amel Benazza-Benyahia
University of Carthage, Higher School of Communication of Tunis (SUP´COM), COSIM Research Lab., Tunisia
Email: sakji.sarra@gmail.com, benazza.amel@supcom.rnu.tn

***Abstract*—In this paper, a region based system is designed for textured image retrieval. A scalable joint Bayesian segmentation and feature extraction in the wavelet transform domain is performed. The segmentation map and the extracted region features are refined by exploiting more decomposition levels. In order to account for spatial dependencies, Markov Random Field (MRF) is employed to model the prior distribution of the segmentation map at each scale. Moreover, a coarse to fine resolution retrieval procedure is proposed. Experimental results carried out on remote sensing images corroborate the gain achieved by the proposed indexing method. Moreover, the resort to an adaptive smoothing parameter reflecting the image homogeneity improves the gain provided by the proposed approach.**

I. Introduction

Conventional Content-Based Image Retrieval (CBIR) systems represent the visual content of an image by extracting salient features from the whole image. However, most images contain several objects with various characteristics possibly scaled differently. Hence, a single global signature computed for the entire image could not sufficiently capture the relevant properties of individual regions which results in a significant semantic gap. Region-Based Image Retrieval (RBIR) systems aim to overcome such limitation by considering an image as a set of regions represented by their local features [1], [2], [3], [4], [5], [6], [7]. This is intuitively closer to the perception of human visual system. Furthermore, several RBIR systems operate in the Wavelet Transform (WT) domain since the latter provides a multiscale image representation consistent with the human visual system [8]. This is the reason why we focus in this work on WT-based RBIR. The design of such systems presents several challenging issues. The first one is the choice of a segmentation method. The second issue concerns the definition of salient features that faithfully reflect the characteristics of each region. Finally, the retrieval step should find images in the database that have similar categories of regions to those of the query image. For instance, in [1], the K-means algorithm based on Mahalanobis distance is applied on the approximation coefficients of color images. Each resulting region is characterized by its size, its centroid and the cross-color covariance matrices defined for the WT subbands at the coarsest level. The matching of query image regions and those of a database image is performed thanks to a similarity measure based on Bhattacharyya distance. In [4], the K-means allows to find the areas of interest in the WT domain. The energy of the WT coefficients and, the barycentric coordinates are chosen as features. In [5], the image is segmented into homogeneous regions which are transformed with shape adaptive discrete wavelet transform (SA-DWT). Then, local color and texture features are extracted. In [6], the Fuzzy C-means (FCM) algorithm is used to segment the image into regions and, robust features are defined in each region. In [7], the image is decomposed by the Haar WT, the K-means algorithm is applied to the approximation subband. The energies of the resulting regions at the coarsest scale are chosen as descriptors. One common limitation of these approaches relies in using only one level of WT decomposition for segmentation and/or extraction of signatures. To alleviate this drawback, in [3], the K-means also clusters the WT coefficients at the coarsest scale but the region features vectors consist of the energy region in all the WT subbands at all the scales. In our recent work [9], we have defined a hierarchical segmentation and descriptor extraction in the WT domain. The segmentation is refined and, the signatures become more representative of the visual context as we exploit finer scales. The resulting tree-structured organization of regions and signatures allows to design a coarse-to-fine fast region matching at the retrieval step. Note that all aforementioned works do not account for spatial dependencies during the feature extraction. In different image processing applications, Markov Random Fields (MRF) [10] were found to be efficient tools to account for spatial dependencies [11], [12], [13]. They can operate in the WT domain thanks to multiresolution MRF models [14], [15], [12]. In this paper, we aim at designing a RBIR in the WT domain. The novelty of our approach consists in designing a joint multiscale segmentation and feature extraction within a Bayesian framework to capture the spatial dependencies. Another contribution concerns the design of a coarse-to-fine scale retrieval procedure. The remainder of this paper is organized as follows. In Sec. II, the prior statistical models of the segmentation map and of the WT coefficients are presented. In Sec. III, the scalable joint segmentation and feature extraction method is explained. In Sec. IV, the hierarchical retrieval procedure is presented. Finally, experimental results are given in Sec. V and some conclusions are drawn in Sec. VI.

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

II. Proposed method

A. A general overview

First, a J-stage WT is applied to an image of size $N \times N$. Then, a joint segmentation and feature extraction are performed from coarse to fine scales. They are firstly performed on the low-pass subband a_J, resulting in K regions characterized by their related features. However, detail coefficients $w_{j,o}$ at orientations $o = 1, 2, 3$ at each scale $j = J, \ldots, 1$ contain image details absent from the approximation subband and are used to characterize the texture at the segmentation step as well as the feature extraction one. This motivates to exploit at each scale j, the information carried by $w_{j,o}$ to derive a segmentation map x_j common to the 3 wavelet subbands and, the related feature vector $\mathbf{F}_j$. In our work, a Bayesian framework is adopted by computing a maximum a posteriori (MAP) estimate of x_j from the WT coefficients. Therefore, we need to model the prior distribution and the likelihood function.

B. Prior distribution of the segmentation map

A segmentation map x_j, such as $x_j(m,n) \in \{1, \ldots, K\}$ for each spatial position $(m,n) \in \{1, \ldots, N_j\}$, $N_j = \frac{1}{N/2^j}$, is viewed as a realization of a random process X_j representing the field of classification labels for each (m,n) at scale j. We assume that the classification label at a given pixel only depends on the labels of its neighbors in a second order neighborhood system. Hence, X_j is considered as a Markov random process and the prior distribution of X_j is a Gibbs distribution given by:

$$\mathsf{p}(x_j) = \frac{1}{Z} \exp[- \sum_{\{(m,n),(m',n')\} \in \mathcal{C}} V(x_j(m,n), x_j(m',n'))] \tag{1}$$

where Z is a normalization constant, $\mathcal{C}$ is the set of pairwise cliques $\{(m,n),(m',n')\}$ and $V(x_j(m,n), x_j(m',n'))$ is the potential function of the clique $\{(m,n),(m',n')\}$ given by $V(x_j(m,n), x_j(m',n')) = -\beta\, \delta(x_j(m,n) - x_j(m',n'))$ where $\delta(\cdot)$ is the impulse function and, β is a positive parameter which controls the degree of image homogeneity.

C. Likelihood function

The likelihood function $\mathsf{p}(w_{j,1}, w_{j,2}, w_{j,3}/x_j)$ is the probability density function of the magnitudes of wavelet coefficients$(w_{j,1}, w_{j,2}, w_{j,3})$ conditionally to the segmentation map x_j. At a given scale j, the observations in different orientations given the labels are assumed to be independent:

$$\begin{aligned} &\mathsf{p}(w_{j,1}, w_{j,2}, w_{j,3}/x_j) = \\ &\prod_{(m,n)=(1,1)}^{(N_j,N_j)} \mathsf{p}((w_{j,o}(m,n))_{o=1}^{3}/x_j(m,n)) \\ &= \prod_{o=1}^{3} \prod_{k=1}^{K} \prod_{(m,n)\in\mathcal{P}_k} \mathsf{p}(w_{j,o}(m,n)/x_j(m,n) = k) \end{aligned} \tag{2}$$

where $\mathcal{P}_k$ is the set of positions (m,n) inside the class k. Due to the sparsity of the WT coefficients inside a region k, the density $\mathsf{p}(w_{j,o}(m,n)/x_j(m,n) = k)$ could be reflected for each position (m,n) by a Gamma distribution. Denoted by $\mathsf{f}_{j,o,k}$, it is defined for every v in $\mathbb{R}^+$ by:

$$\mathsf{f}_{j,o,k}(v) = v^{\alpha_{j,o,k}-1} \frac{e^{-\gamma_{j,o,k} v}}{{\gamma_{j,o,k}}^{\alpha_{j,o,k}} \Gamma(\alpha_{j,o,k})} \tag{3}$$

where $\alpha_{j,o,k}$ and, $\gamma_{j,o,k}$ are respectively the scale and shape parameters and $\Gamma(\cdot)$ is the Gamma function. Hence, the feature vector is $\mathbf{F}_j = (\mathbf{f}_j(1)^\intercal, \ldots, \mathbf{f}_j(K)^\intercal)^\intercal$ where $\mathbf{f}_j(k) = (\alpha_{j,1,k}, \gamma_{j,1,k}, \alpha_{j,2,k}, \gamma_{j,2,k}, \alpha_{j,3,k}, \gamma_{j,3,k})^\intercal$ is the feature vector of region k.

III. Scalable joint Bayesian segmentation and feature extraction

As shown in Figure 1, a MAP estimation of the segmentation map is performed in a scalable way from the coarse approximation subband to the finer WT subbands and, the related feature vectors are consequently computed.

A. Proceeding at the approximation subband

The initial segmentation map of the low-pass subband results from the region growing algorithm [16]. The considered attributes are the mean values computed over a sliding window around each approximation coefficient. The vector describing the approximation subband is $\mathbf{F}_{\text{app}} = (\mu(1), \ldots, \mu(K))^\intercal$ where $\mu(k)$ is the mean of the approximation coefficients in the region $k = 1, \ldots, K$.

B. Bayesian segmentation of the detail subbands

First, at each scale $j = J, \ldots, 1$ an initial segmentation map $x_j^{(0)}$ is inherited from the immediately coarser scale $j+1$ if $j < J$ or from the approximation subband if $j = J$. Then, the final segmentation map x_j and the feature vector $\mathbf{F}_j$ are computed by maximizing the posterior density function of x_j conditionally to $(w_{j,1}, w_{j,2}, w_{j,3})$ which can be expressed thanks to the Bayes rule by:

$$x_j = \arg\max_x \mathsf{p}(w_{j,1}, w_{j,2}, w_{j,3}/x)\mathsf{p}(x). \tag{4}$$

The segmentation map x_j and, consequently the feature vector $\mathbf{F}_j$ are estimated in an iterative way thanks to the ICM method [17]. More precisely, given the regions defined by $x_j^{(0)}$, the parameters of their distributions are estimated according to the Maximum Likelihood (ML) criterion to form the initial feature vector $\mathbf{F}_j^{(0)}$. At iteration it, position (m,n), the MAP estimate of $x_j^{(it)}(m,n) = k$ is computed conditionally to $w_{j,1}(m,n), w_{j,2}(m,n), w_{j,3}(m,n)$ and, to $x_j^{(it-1)}(\mathcal{N}_{m,n})$ where $\mathcal{N}_{m,n}$ is the second order neighborhood of (m,n). This is equivalent to minimize the local energy $E_{m,n}^{(it)}(k)$ given by:

$$\begin{aligned} E_{m,n}^{(it)}(k) = &- \sum_{o=1}^{3} \ln \mathsf{p}(w_{j,o}(m,n)/x_j(m,n) = k) \\ &+ \sum_{(m',n')\in\mathcal{N}_{m,n}} V(k, x_j^{it-1}(m',n')). \end{aligned} \tag{5}$$

Once the map $x_j^{(it)}$ is computed at iteration it, $\mathbf{F}_j^{(it)}$ is estimated in the sense of the ML from $x_j^{(it)}$ and $(w_{j,1}, w_{j,2}, w_{j,3})$.

C. Adaptive smoothing parameter

Generally, a predefined value is chosen for the smoothing parameter β. However, β plays a key role as it reflects the contribution of the neighborhood in the segmentation by controlling the region homogeneity degree. Hence, it is important to adjust its value according to the image content, and possibly at each scale. To this end, in [18], [19], it is suggested to find the value β_j that maximizes the likelihood $\mathsf{p}(x_j/w_{j,1}, w_{j,2}, w_{j,3})$:

$$\beta_j = \arg\max_{\beta} \mathsf{p}(w_{j,1}, w_{j,2}, w_{j,3}/x_j)\mathsf{p}(x_j). \tag{6}$$

However, $\mathsf{p}(x_j)$ includes the unknown constant Z which depends on β_j and whose evaluation is unfeasible in practical situations. In [18], [19], the proposed alternative is to maximize a Pseudo Maximum Likelihood (PML) criterion $\mathcal{J}_{\text{PML}}$ corresponding to the product of local conditional distributions for all the positions (m,n):

$$\mathcal{J}_{\text{PML}}(\beta_j) = \prod_{(m,n)=(1,1)}^{(N_j,N_j)} \mathsf{p}((w_{j,o}(m,n))_{o=1}^{3}/x_j(m,n)) \times \mathsf{p}(x_j(m,n)/x_j(\mathcal{N}_{m,n})) \tag{7}$$

After some algebraical steps, it can be proved that β_j is the solution of the following equation:

$$\sum_{(m,n)=(1,1)}^{(N_j,N_j)} U_{m,n}(x_j(m,n)) = \sum_{(m,n)=(1,1)}^{(N_j,N_j)} \frac{\sum_{k=1}^{K} U_{m,n}(k)\mathsf{p}_{m,n}(k)e^{\beta_j U_{m,n}(k)}}{\sum_{k=1}^{K} \mathsf{p}_{m,n}(k)e^{\beta_j U_{m,n}(k)}} \tag{8}$$

where $U_{m,n}(k)$ is the number of neighbors in $\mathcal{N}_{m,n}$ that have the label $k \in \{1,...,K\}$ and, $\mathsf{p}_{m,n}(k) = \mathsf{p}((w_{j,o}(m,n))_{o=1}^{3}/x_j(m,n) = k)$. The previous step involves an observed map x_j and also the knowledge of $\mathsf{p}(w_{j,1}, w_{j,2}, w_{j,3}/x_j)$ associated to $\mathbf{F}_j$. However, in our case, x_j and $\mathbf{F}_j$ are unknown. Hence, we propose to couple the estimation of β_j with that of x_j and $\mathbf{F}_j$ in an iterative way thanks to Algorithm 1. More precisely, based on the regions defined by $x_j^{(0)}$ and the initial feature vector $\mathbf{F}_j^{(0)}$ estimated according to the ML criterion, an initial parameter $\beta_j^{(0)}$ is estimated according to the PML criterion. At iteration it and position (m,n), $x_j^{(it)}(m,n)$ is updated by minimizing the local energy $E^{(it)}$ whose expression involves $x_j^{(it-1)}$, $\mathbf{F}_j^{(it-1)}$ and, $\beta_j^{(it-1)}$. Once the new $x_j^{(it)}$ and consequently $\mathbf{F}_j^{(it)}$ are obtained, a new value of the PML estimator $\beta_j^{(it)}$ is derived from $x_j^{(it)}$ and, $\mathbf{F}_j^{(it)}$. The algorithm is stopped when the variation $|\beta_j^{(it)} - \beta_j^{(it-1)}|$ becomes below a given threshold ε.

IV. Scalable retrieval procedure

A. Matching procedure

The goal consists in finding the images I^{db} in the database whose regions are the most similar to those of a query image I^{q}. To this end, we proceed in two steps: we match the approximation subbands then the detail ones. Each class $k = 1,\ldots,K^{\text{q}}$ in the approximation subband of I^{q} is firstly assigned to a class $Q[k] \in \{1,\ldots,K^{\text{db}}\}$ in the approximation subband of I^{db} whose prototype is the closest to that of k: $Q[k] = \arg\min_{k' \in \{1,\ldots,K^{\text{db}}\}} \mathcal{D}(\mu(k), \mu(k'))$ where $\mathcal{D}$ denotes the Normalized Euclidean Distance (NED). Hence, we define the similarity $\mathcal{S}_{\text{app}}(I^{\text{q}}, I^{\text{db}}) = \mathcal{D}(\mathbf{F}^{\text{q}}_{\text{app}}, \mathbf{F}^{\text{db}}_{\text{app}}[I^{\text{q}}])$ between the approximation subband of I^{q} and that of I^{db} where $\mathbf{F}^{\text{db}}_{\text{app}}[I^{\text{q}}] = (\mu(Q[1]),\ldots,\mu(Q[K^{\text{q}}]))^{\intercal}$ is formed by the mean values $\mu(Q[k])$ of the matched regions $Q[k]$ of I^{db} to the regions $k = 1,\ldots,K^{\text{q}}$ of I^{q}. The matching procedure continues by exploiting the information contained in the wavelet coefficients by assigning to each class k of I^{q} at scale j a class $Q_j[k]$ of I^{db} at the same scale j if it satisfies the condition: $Q_j[k] = \arg\min_{k' \in \{1,\ldots,K^{\text{db}}\}} \mathcal{D}(\mathbf{f}^{\text{q}}_j(k), \mathbf{f}^{\text{db}}_j(k'))$ where $\mathbf{f}^{\text{q}}_j(k)$ and $\mathbf{f}^{\text{db}}_j(k')$ are respectively the descriptors of regions k and k' at scale j. Thus, it is easy to define the similarity $\mathcal{S}_j(I^{\text{q}}, I^{\text{db}}) = \mathcal{D}(\mathbf{F}^{\text{q}}_j, \mathbf{F}^{\text{db}}_j[I^{\text{q}}])$ between the detail subbands of I^{q} and those of I^{db} at scale j, where $\mathbf{F}^{\text{db}}_j[I^{\text{q}}] = (\mathbf{f}^{\text{db}}_j(Q_j[1])^{\intercal},\ldots,\mathbf{f}^{\text{db}}_j(Q_j[K^{\text{q}}])^{\intercal})^{\intercal}$.

B. Coarse-to-fine scale retrieval

The search is performed in a coarse-to-fine resolution way until the user is satisfied. It starts by only considering the information at the approximation level: the system returns the R images I^{db} of the database that minimize $\mathcal{S}_{\text{app}}(I^{\text{q}}, I^{\text{db}})$. If the user is not satisfied, the additional information carried by the WT subbands from scale J to a chosen scale J_{u} at which he stops the search procedure is exploited. In this case, the R images I^{db} of the database that minimize $\mathcal{S}_{\text{app}}(I^{\text{q}}, I^{\text{db}}) + \sum_{J}^{J_{\text{u}}} \mathcal{S}_j(I^{\text{q}}, I^{\text{db}})$ are returned.

V. Experimental results

A. Validation of the segmentation approach

Several synthetic images were tested to assess the reliability of our segmentation approach. For reasons of space, the illustration are just given for 2 images with respectively 2 and 5 different textures. A two-stage ($J = 2$) 5/3 lifting transform is retained. Then, the scalable segmentation is performed, the parameter β being firstly set manually so as $\beta \in \{0, 2\}$. Note that if $\beta = 0$, no spatial dependency is considered in the segmentation. In a second phase, β_j at scales $j \in \{1, 2\}$ are adjusted by PML estimation. Figure 2 shows the result of the segmentation of the approximation subband and that of the WT subbands at scales $j \in \{1, 2\}$ for the different values of β (or estimated β_j). For both images, we note that the choice of the smoothing parameter has a significant influence on the relevance of the segmentation. Besides, the best segmentation results are obtained when the β_j are estimated according to the PML criterion. Moreover, the segmentation results are improved by accounting for finer detail subbands.

B. Retrieval results

To test retrieval performances we use a training database consisting of the first channels of 451, 128 $\times$ 128 SPOT3

satellite images representing regions belonging to one of the following categories (urban, field, water, mountain, aeroport). $J = 2$ scales of a 5/3 lifting scheme are applied to each database image, then for different preset values of β, scalable joint segmentation and feature extraction is performed. All the images in the database are considered as queries. Figure 3 (**a** ,**b**) provides the plot of average precision PR *versus* average recall RC computed over all the queries for $\beta = 0$ and $\beta = 0.2$ and, in three cases where the user stops the search at the approximation level, and at the detail levels $J_{\mathrm{u}} = J$ and $J_{\mathrm{u}} = J - 1$. It indicates that the performances are improved when we account for the detail coefficients and especially for finer scales. Figure 3 (**c**,**d**) provides the plot of average precision and recall obtained for different numbers R of returned images for $\beta = 0.2$. It shows that the superiority of considering all scales ($J_{\mathrm{u}} = 1$) becomes more significant for bigger values of R. Besides, Table I proves that accounting for further scales requires more computational time to analyze an image. Figure 4 (**a**,**b**) shows the evolution of average precision PR computed for $R = 10$ returned images and for different values of fixed β. The performances are the best for a particular value β^*. However, Figure 4 (**c**,**d**) indicates that the best performances are provided when β_j are estimated by PML.

C. Comparison with state-of-art approaches

The first part of Figure 5 shows improvement of performances given by RBIR in comparison with standard CBIR approach based on the same nature of features. The second part of Figure 5 shows that our approach also outperforms the three WT-based RBIR systems: Hirbir [3], Windsurf [1] and Amoda [7], except for the case of $\beta = 0$ when spatial dependencies are ignored.

VI. Conclusion

In this paper, a new RBIR system accounting for spatial dependencies in the WT coefficients is presented. The novelty relies on the joint scalable Bayesian segmentation and feature extraction in the WT domain. Moreover, the scalable organization of features enables a coarse to fine scale retrieval procedure. Simulations show improvement of retrieval performances when a multiscale strategy is considered. Adapting the smoothing parameter achieves further improvement of the performances. Comparison with standard CBIR and WT-based RBIR systems proves that a region based approach accounting for spatial dependencies has a significant impact on retrieval performances which is more significant in the case of adaptive smoothing parameter.

References

[1] S. Ardizzoni, I. Bartolini, and M. Patella, "Windsurf: Region-based image retrieval using wavelets," in *DEXA*, Florence, Italy, 1999.

[2] A. Natsev, R. Rastogi, and K. Shim, "WALRUS: a similarity retrieval algorithm for image databases," *IEEE Trans. on T-KDE*, vol. 16, no. 3, pp. 301–316, 2004.

[3] Y. Sun and S. Ozawa, "HIRBIR: A hierarchical approach to region-based image retrieval," *Mult. Syst. J.*, vol. 10, no. 6, pp. 559–569, 2005.

[4] Xugang, F. Yin, and C. Wei, "Wavelet transformation for content-based image retrieval combine G-regions of interest," *Springer, Ad. in Intell. and Soft Comp.*, vol. 129, pp. 285–291, 2012.

[5] L. Belhallouche, K. Belloutata, and K. Kpalma, "A new approach to region based image retrieval using shape adaptive discrete wavelet transform," *IJIGSP*, vol. 1, no. 8, pp. 1–14, 2016.

[6] K. Ventakasalam and P. Ranjendran, "Effective rbir fuzzy c-means segmentation haar wavelet with user interactive multi threshold robust features vector," *AJIT*, vol. 2, no. 15, pp. 223–231, 2016.

[7] N. Amoda and R. K. Kulkarni, "Efficient image retrieval using region based image retrieval," *SIPIJ*, vol. 4, no. 3, pp. 17–9, 2013.

[8] E. Salari and Z. Ling, "Texture segmentation using hierarchical wavelet-decomposition," *Patt. Rec.*, vol. 28, no. 12, pp. 1819–1824, 1995.

[9] S. Sakji-Nsibi and A. Benazza-Benyahia, "Wavelet-based image retrieval using hierarchical segmentation and region matching," in *ISIVC*, Marrakech, Morocco, 2014.

[10] S. Z. Li, *Markov random field modeling in image analysis*, 3rd ed. Springer, 2009.

[11] S. Pelizzari and J. M. Bioucas-Dias, "Bayesian segmentation of oceanic SAR images, application to oil spill detection," *IEEE TGRS*, 2010.

[12] A. R. Backes, L. C. Gerhardinger, E. B. N. Jdo, and O. Bruno, "Medical image retrieval and analysis by Markov random fields and multiscale," *Phys. Med. Biol.*, vol. 60, no. 3, pp. 1125–1139, 2015.

[13] R. Reulkea and A. Lippokb, "MRF-based texture segmentation for road detection," in *ISPRS*, Beijing, China, 2008.

[14] V. Gupta, L. Kumar, and U. Kumari, "Color satellite image segmentation using Markov random field and multiresolutional wavelet transform," *IJCAES*, vol. 1, no. 1, pp. 75–, 2011.

[15] H. Noda, M. N. Shirazib, and E. Kawaguchia, "MRF-based texture segmentation using wavelet decomposed images," *Patt. Rec.*, vol. 35, no. 4, pp. 771–782, 2002.

[16] A. B. M. Faruquzzaman, N. R. Paiker, J. Arafat, M. A. Ali, and G. Sorwar, "Literature on image segmentation based on split and merge techniques," in *ICITA*, Cairns, Australia, 2008.

[17] J. Besag, "On spatial interaction and the statistical analysis of lattice systems," *JRSS*, vol. 139, no. 2, pp. 192–236, 1974.

[18] J. Gimenez, A. Frery, and A. Flesia, "Inference strategies for the smoothness parameter in the potts model," in *IGARSS*, Melbourne, VIC, 2013.

[19] ——, "When data do not bring information: A case study in markov random fields estimation," *IEEE J. Select. Topics In Appl. Earth Obs. Rem. Sens.*, vol. 8, no. 1, pp. 195–203, 2015.

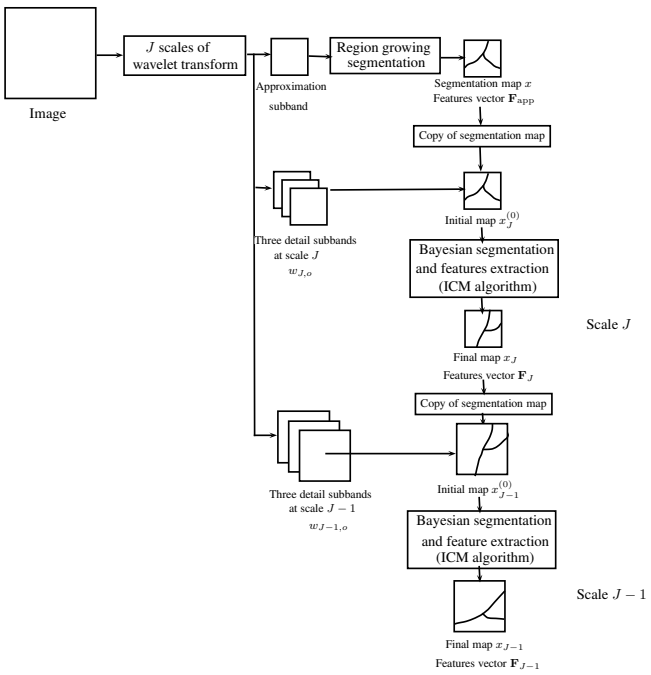


Fig. 1. Joint scalable segmentation and feature extraction.

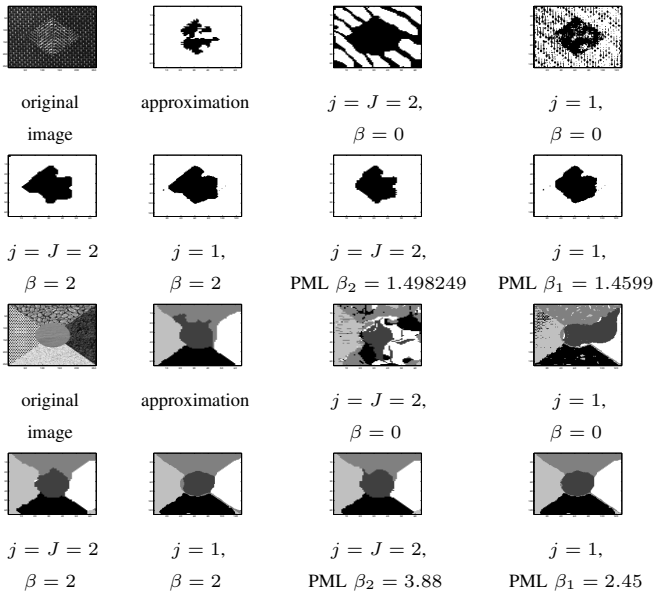
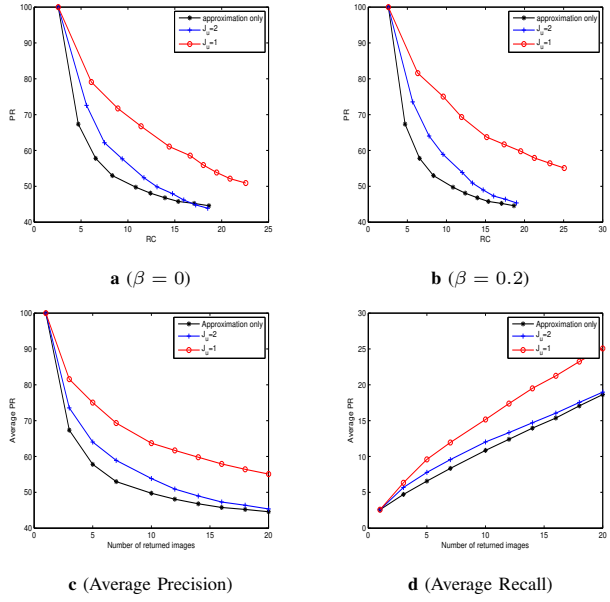


Fig. 2. Segmentation results on a synthetic images at different scales and for different β

.



a ($\beta = 0$) **b** ($\beta = 0.2$)

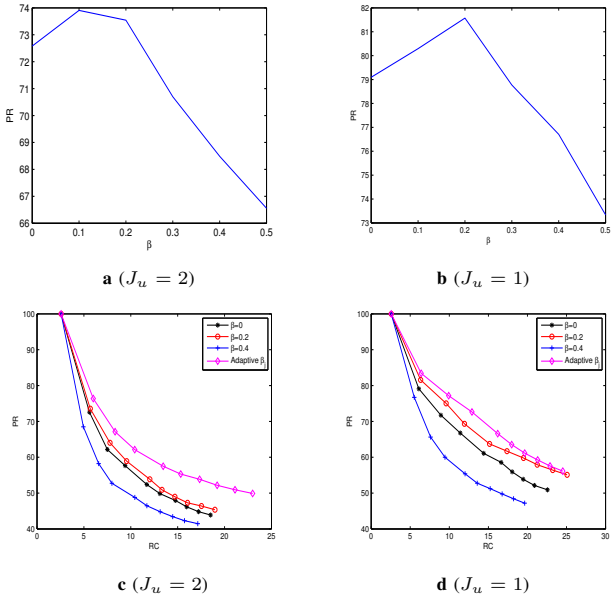
c (Average Precision) **d** (Average Recall)

Fig. 3. Influence on performances of the number of considered scales.

TABLE I

CPU TIME REQUIRED TO ANALYZE AN IMAGE ON A COMPUTER WITH 4 GB OF RAM AND A PROCESSOR INTEL (R) CORE (TM) 5200U CPU 2.20 GHZ.

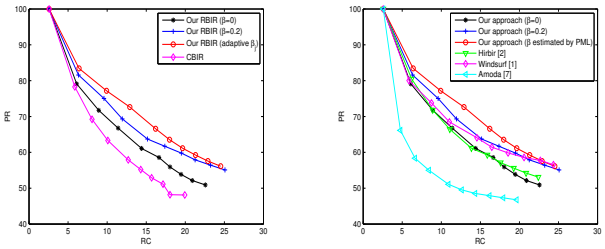
	Approximation only	$J_u = 2$	$J_u = 1$
CPU time	0.1656	1.0453	2.7516



a ($J_u = 2$) **b** ($J_u = 1$)

c ($J_u = 2$) **d** ($J_u = 1$)

Fig. 4. Influence of the parameter β on retrieval performances.



CBIR approach Approaches in the literature

Fig. 5. Comparison between our approach ($J_u = 1$) and state-of-the art approaches.

Algorithm 1 Modified ICM algorithm.

Require: $x_j^{(0)}$, ε
  $\mathbf{F}_j^{(0)}$ = ML estimator from$(x_j^{(0)}, w_{j,1}, w_{j,2}, w_{j,3})$
  $\beta_j^{(0)}$ = PML estimator from$(x_j^{(0)}, \mathbf{F}_j^{(0)})$
  Δ=1;
  $it = 1$
  while $\Delta > \varepsilon$ **do**
    for $m = 1, \ldots, N_j$ **do**
      for $n = 1, \ldots, N_j$ **do**
        for $k = 1, \ldots, K$ **do**
          compute $E_{m,n}^{(it)}(k)$
        end for
        $x_j^{(it)}(m,n) = \arg \min_{k=1,\ldots,K} E_{m,n}^{(it)}(k)$
      end for
    end for
    $\mathbf{F}_j^{(it)}$ = ML estimator from $(x_j^{(it)}, w_{j,1}, w_{j,2}, w_{j,3})$
    $\beta_j^{(it)}$ = PML estimator from $(x_j^{(it)}, \mathbf{F}_j^{(it)})$
    $\Delta = \mid \beta_j^{(it)} - \beta_j^{(it-1)} \mid$
    $it \longleftarrow it + 1$
  end while