Pattern Recognition Letters xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Color image segmentation using morphological clustering and fusion with automatic scale selection

O. Lézoray *, C. Charrier

Université de Caen Basse Normandie, GREYC CNRS UMR 6072, Ensicaen, Équipe Image, 6 Bd. Maréchal Juin, F-14050 Caen, France

ARTICLE INFO

Article history:
Received 14 March 2006
Received in revised form 27 August 2008
Available online xxxx

Communicated by Y.J. Zhang

PACS:
07.05.Pj
87.19.Dd

Keywords:
Color
Segmentation
Clustering
Mathematical morphology
Scale selection

ABSTRACT

In this paper, a color image segmentation method considering pairwise color projections is proposed. Each pairwise projection is analyzed according to an unsupervised morphological clustering which looks for the dominant colors of a 2D histogram. This leads to obtaining three segmentation maps combined by superposition after being simplified. The superposition process itself producing an over-segmentation of the image, a pairwise region merging is performed according to a similarity criterion up to a termination criterion. To fully automate the segmentation, an energy function is proposed to quantify the segmentation quality. The latter acts as a performance indicator and is used all over the segmentation to tune its parameters: the scale of the unsupervised morphological clustering and the termination criterion of region merging. Experimental results are conducted on a reference image database and comparisons with state-of-the-art algorithms.

1. Introduction

Image segmentation is an essential step in image content understanding since the quality of the image visual objects interpretation depends on the results of that very image segmentation. This is a difficult problem the solution of which is not universal. Color image segmentation methods can be roughly categorized into two main families, depending on wether the distribution of pixel colors is analyzed in the image or in the color domains (Cheng et al., 2001; Plataniotis and Venetsanopoulos, 2000).

It is generally assumed that homogeneous regions in the image domain give rise to clusters in the color domain. Clusters in color domain can be extracted by means of color histogram analysis (Shafarenko et al., 1998; Gillet et al., 2001; Vannoorenberghe et al., 1999; Busin et al., 2004; Celenk, 1990) or cluster identification (Scheunders, 1997; Uchiyama and Arbib, 1994; Agarwal et al., 2005). A segmentation map is obtained by labeling connected pixels in the image domain that belong to the same class in the color domain. Such an analysis often leading to over-segmentation, spatial refinement is needed to simplify the segmentation (Deng and Manjunath, 2001).

For the case of color histogram analysis, one can consider separately the histograms of each component, this is the marginal approach (Busin et al., 2004; Vannoorenberghe et al., 1999). One can also directly consider the whole color histogram (Géraud et al., 2001), few works consider that issue because of the complexity of the approach. To have reliable and competing histogram clustering strategies, several researchers (Kurugollu et al., 2001; Lucchese and Mitra, 2001; Xue et al., 2003) have proposed to use an intermediate representation called bi-marginal that considers pairwise combinations of color components to have a partial view of the correlation of color components. Moreover, this has the advantage of lower complexity as compared to clustering the whole color histogram. We consider such a pairwise projection representation for segmenting color images.

In our proposed approach, the three pairwise *RGB* color combinations are independently treated according to an unsupervised clustering with automatic scale selection. Three segmentation maps are then obtained, intersected and further simplified by region merging. The bi-marginal segmentation relies on an unsupervised morphological clustering which clusters the image into its representative colors. The proposed clustering being multi-scale, a scale selection algorithm is proposed. In our algorithm, we have

* Corresponding author. Tel.: +33 33775517; fax: +33 233771167.
E-mail addresses: olivier.lezoray@unicaen.fr (O. Lézoray), christophe.charrier@unicaen.fr (C. Charrier).
URL: http://www.info.unicaen.fr/~lezoray (O. Lézoray).

0167-8655/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2008.11.005

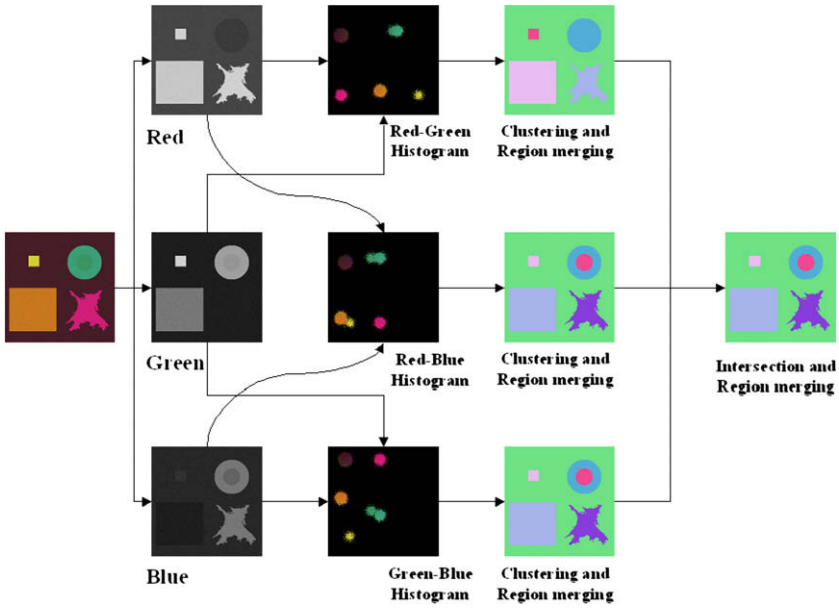

Fig. 1. General scheme of the proposed bi-marginal approach to color image segmentation.

been very careful to solve this problem and we have developed a robust criterion as the basis for scale selection. Moreover, this criterion leads all the segmentation steps and is used to perform both scale selection and region merging. The complete segmentation algorithm is parameter-free and is illustrated by Fig. 1. Clustering operates in pairwise color domains and region merging is a spatial-based post-processing to recover the spatial coherence of colors in the image domain.

The remainder of the paper is organized as follows. In Section 2, we discuss how to perform a clustering of a color image and we propose an unsupervised morphological clustering of bivariate histograms. Section 3 presents a robust criterion as the basis for scale selection of the proposed bi-marginal clustering. Section 4 presents how the bi-marginal segmentations are combined and simplified by automatic and adaptive region merging. The paper ends with Sections 5 and 6 providing experimental results, discussion and conclusion.

2. Clustering colorimetric information

To perform a clustering of a color image, several strategies can be used. They generally differ on the dimension of the data used to cluster the image but they share the same assumptions: homogeneous regions in the image plane give rise to explicit clusters in a colorimetric projection representation (1D or 3D). Usually, the spatial repartition of the colors in a color space is used to cluster an image assuming that the colors of the objects are grouped around dominant color clusters in a colorimetric projection. A cluster therefore corresponds to a class of pixels which have similar color properties. In the 1D case, the histograms of each band are considered separately (marginal approach) and the method is reduced to finding thresholds dissecting the band histograms (Celenk, 1990; Kurugollu et al., 2001; Lezoray and Cardot, 2002; Busin et al., 2004). Since one has one 1D histogram per color band, several clusterings are obtained and combined (Vannoorenberghe et al., 1999). In the 3D case, the vectorial aspect of color is taken into account (vectorial approach) and the clustering is considered as the classification of multi-spectral data (Soille, 1996; Géraud et al., 2001; Park et al., 1998; Postaire et al., 1993). Both 1D and 3D methods suffer from the determination of the number of classes (number of dominant colors), which is usually assumed to be known.

An interesting alternative to 1D and 3D clustering methods relies on the use of pairwise projections and especially bivariate histograms (2D histograms) which use two color bands together (pairwise associations) namely *RG*, *RB* and *GB* in the *RGB* color space. This can bring several advantages (Matas, 1995; Kurugollu et al., 2001; Xue et al., 2003; Lezoray and Cardot, 2003; Lezoray, 2004). The paucity of the data encountered in the 3D case is partially overcome and the search complexity is drastically reduced. Moreover, it partially uses the spatial repartition of colors and offers an intermediate method to the 1D and 3D ones. Another advantage to be considered is the fact that a 2D histogram is nothing more than a grey-level image. Therefore, classical and fast grey-level image processing algorithms can be used to cluster a 2D histogram. One can refer to (Kurugollu et al., 2001) for a comparison of 1D, 2D and 3D histogram-based methods for clustering color images which shows the advantages of 2D histograms. Moreover, in the context of skin detection, the advantage of using 2D histograms (e.g chromaticity histograms) is now established over other methods (Kakumanu et al., 2007). Fig. 2 presents, for a synthetic image, its 3D histogram and the three 2D histograms obtained by band-pair projections. Fig. 2 is composed of 6 regions with different shapes, sizes and colors (corrupted by gaussian noise). One can see that the projections obviously do not contain the same information: in some projections, clusters are not separable and they do not appear in the same areas for different projections.

One has to note that a clustering is not a segmentation in itself. A clustering produces a clustering map in which each pixel is associated to a class. Since a color cluster in a color space is not necessarily associated to connected colored pixels in the image plane, the clustering can assign the same class to pixels that are not connected. It is therefore necessary to perform a labeling of the clustering map to obtain a real segmentation.

2.1. Unsupervised morphological clustering

In this Section, we propose a new method for clustering 2D histograms obtained from color images. This method is unsupervised and relies on mathematical morphology operations (Serra, 1982) to

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005

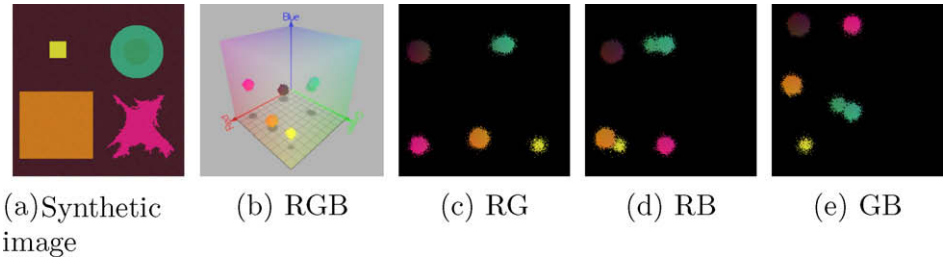

Fig. 2. A synthetic image, its 3D color space projection and the corresponding projections on band-pair planes (2D histograms).

partition an image in a given number of classes. Morphological classifiers consider a histogram as a probability density function which is represented by an image. The advantage of using morphological classifiers relies on the fact that no assumption is used on the distributions since assuming they are Gaussian ones is usually very far from reality. The use of a watershed for clustering has been proposed by SOILLE (Soille,1996, 2004). SOILLE's method has an *a priori* parameter which is the number of classes and it might be more suitable to have a way of determining this number. Therefore, we propose an unsupervised morphological clustering where the watershed operates on a distance function to the class centers (dominant colors), the number of the latter determining the number of classes.

Our proposed 2D histogram is close to SOILLE's method for 1D histogram. Fig. 3 presents its principles. Since a 2D histogram $\mathscr{H}$ is usually noisy, the latter is smoothed with a symmetrical exponential filter φ_β (with a given parameter $\beta \in [0,1]$ which levels the smoothing effect: the lower β, the smoother). The result

$$\mathscr{H}^{(1)} = \varphi_\beta(\mathscr{H}) \tag{1}$$

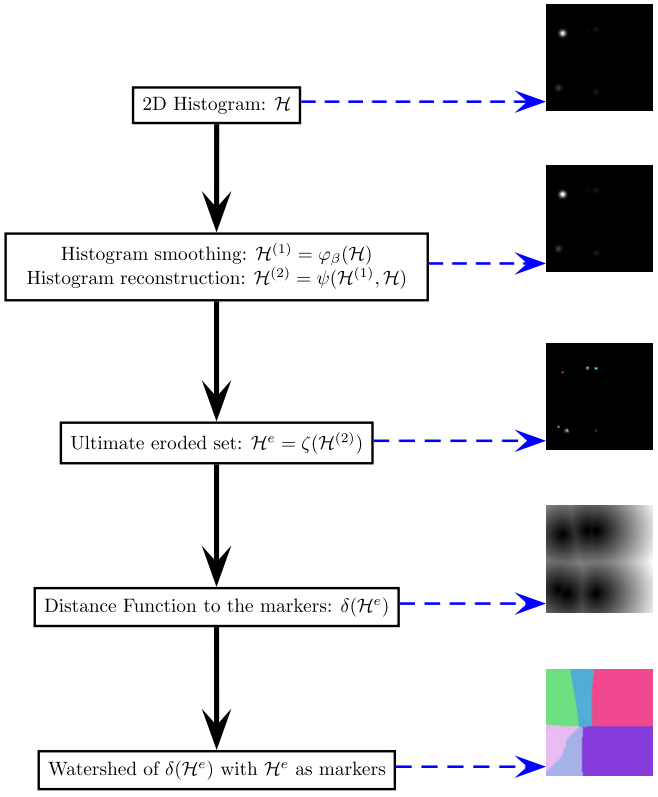

Fig. 3. The proposed unsupervised morphological clustering: a view of each processing result on the 2D histogram is provided (dotted arrows).

is reconstructed (by a morphological reconstruction process ψ) in the original histogram image $\mathscr{H}$ to obtain a noiseless and regularized version of the histogram

$$\mathscr{H}^{(2)} = \psi(\mathscr{H}^{(1)}, \mathscr{H}). \tag{2}$$

Dominant colors corresponding to maxima in a 2D histogram, an erosion operation applied on an histogram reduces each cluster to its main colors: the clusters can be progressively thinned until minimal connected components appear (Gillet et al., 2000). The iteration of this property is used to extract the dominant colors and the ultimate eroded set

$$\mathscr{H}^e = \zeta(\mathscr{H}^{(2)}) \tag{3}$$

enables to automatically extract from a histogram all its dominant colors (corresponding to coherent clusters in the histogram). The ultimate eroded set provides the centers of the classes of a 2D histogram without any assumption on their number. These centroids are labeled and used as markers for a watershed operating on a distance function δ to the markers

$$W_{\mathscr{H}} = \mathit{Watershed}(\delta(\mathscr{H}^e)). \tag{4}$$

From the clustered 2D histogram, a clustering map is obtained since each region in the clustered histogram image corresponds to a set of colors in the original image: to each (RGB) vector in the image plane is associated the corresponding label in the clustered 2D histogram. This is depicted by Fig. 4b–d where the three segmentations of the three band-pair projections of Fig. 2a are shown. One can notice that for the *RG* projection, the obtained clustering cannot distinguish the two close greens because they are too mixed in the corresponding projection. Each segmentation map is labeled to obtain a partition of the image in regions and not only a clustering. The number of regions in each segmentation map can therefore be higher than the number of classes since color clusters in a projection are not necessarily connected in the image.

3. Scale selection by energy minimization

The proposed unsupervised morphological clustering method enables to partition a bivariate histogram into a number of regions, the number of which is automatically determined by the extraction of its dominant colors. This number of regions (i.e the number of classes) is not fixed and depends on the content of the reconstructed version of the histogram. The histogram is simplified by reconstruction and parameter β gives the smoothing factor of the exponential filter. The stronger the smoothing, the lower the number of regions of the clustering and identically, the lower the smoothing, the higher the number of regions (high number of classes). The β parameter is therefore a multi-scale parameter which enables to level the complexity of the final segmentation map since it levels the number of regions that will be extracted by the clustering. This parameter can be fixed *a priori* but for general segmentation purposes, a more accurate tuning of this parameter is needed. For some images, high values of β (low smoothing, numerous

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005

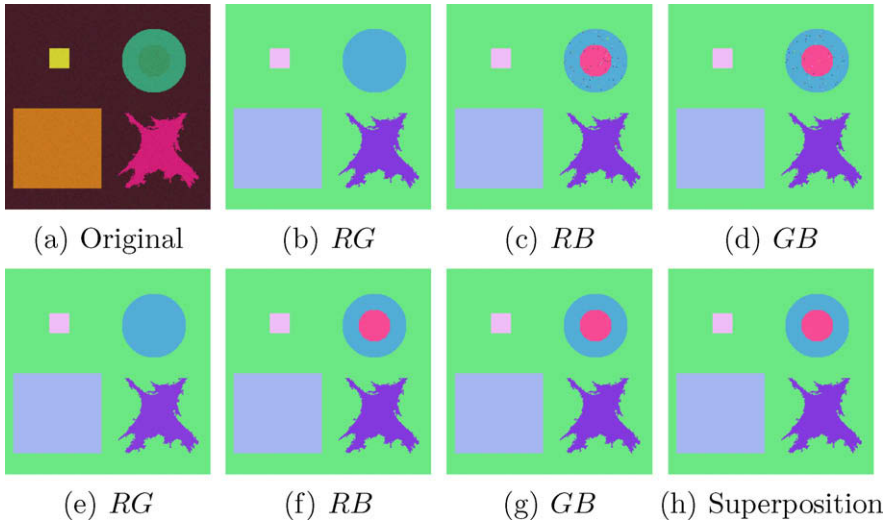

Fig. 4. Segmentation maps (b–d) of a color image (a), their simplification by area region merging (e–g) and their superposition (h).

regions) will be needed and for others, low values (high smoothing, few regions). We propose to tune automatically this parameter to find the best suited level of simplification of the histogram according to the obtained final segmentation. To tune β, we have to be able to compare the outputs of two different clustering segmentations for two different values of β. Therefore, we want to quantitatively compare segmentation maps. This problem is difficult and is the same as to define a similarity measure between two segmentations. Variational methods for image segmentation try to minimize a segmentation energy E which can be regarded as a measure of the segmentation quality (Koepfler et al., 1994). It is a general agreement that the smaller the energy, the better the segmentation. An energy function E can be defined as (Guigues and Men, 2003; Poggio et al., 1985) $E : \mathcal{M} \times \mathcal{I} \rightarrow R^+$ on the set of all the possible segmentations $\mathcal{M}$ and on the image to segment $\mathcal{I}$. The solution of the segmentation problem is the same as to find the segmentation $\mathcal{M}^*$ which minimizes the energy E:

$$\mathcal{M}^*(\mathcal{I}) = \arg\min_{\mathcal{M}_k \in \mathcal{M}} E(\mathcal{M}_k, \mathcal{I}). \tag{5}$$

Since there surely is no unique solution to the problem of image segmentation, a trade-off has to be found between the quality of the segmentation and its complexity. The quality of a segmentation describes its fidelity to the original data of the image and the complexity describes the simplicity of the model. Segmentation energies can often be expressed as follows:

$$E(\mathcal{M}, \mathcal{I}) = D(\mathcal{M}, \mathcal{I}) + C(\mathcal{M}, \mathcal{I}) \tag{6}$$

This expression of segmentation energies expresses a data fidelity term $D(\mathcal{M}, \mathcal{I})$ and a regularization term $C(\mathcal{M}, \mathcal{I})$. For instance, in the reference MUMFORD SHAH functional (Mumford and Shah, 1989), the term D is a mean squared error and the term C is the total length of the boundaries (C depends only on the model $\mathcal{M}$ in this case). A lot of segmentation energies can fit this way of expressing the energy: Markov random fields (MRF) (Li, 1995) but also segmentation quality evaluators. For instance (Boykov et al., 2001),

$$E(\mathcal{M}, \mathcal{I}) = E_{data}(\mathcal{M}, \mathcal{I}) + E_{smooth}(\mathcal{M}, \mathcal{I}) \tag{7}$$

where $E_{data}(\mathcal{M}, \mathcal{I})$ measures the disagreement between the model $\mathcal{M}$ and the data $\mathcal{I}$, $E_{smooth}(\mathcal{M}, \mathcal{I})$ measures the extent to which $\mathcal{M}$ is not piecewise smooth. Energy-based models have several advantages: an energy measure is a measure of segmentation quality and looking at a minimum of the energy provides a way to tune parameters. One important thing to point out is that we cannot assert that a global minimum can be found for energy-based models (Li, 1995). We propose an energy model inspired from MRF and segmentation quality measures. The energy we consider to quantify the quality of the segmentation $\mathcal{M}$ of an image $\mathcal{I}$ into regions $\mathcal{R}_i$, $i \in [1, |\mathcal{M}|]$ is the following one:

$$E(\mathcal{M}, \mathcal{I}) = \alpha_d \times \sum_{i=1}^{|\mathcal{M}|} \mathcal{D}_{\mathcal{R}_i}(\mathcal{M}, \mathcal{I}) + \alpha_c \times \sum_{i=1}^{|\mathcal{M}|} \left[\frac{1}{|\phi(\mathcal{R}_i)|} \sum_{\mathcal{R}_j \in \phi(\mathcal{R}_i)}^{|\phi(\mathcal{R}_i)|} d(\mathcal{R}_i, \mathcal{R}_j) \right] \tag{8}$$

with $\alpha_c = 100 \times \frac{\sqrt{|\mathcal{M}|}}{h \times w}$ and $\alpha_d = \frac{1}{|L| \times |\mathcal{M}|}$. $h \times w$ denotes the size of the image $\mathcal{I}$, $\phi(\mathcal{R}_i)$ the set of adjacent regions to the region $\mathcal{R}_i$, $|\phi(\mathcal{R}_i)|$ its cardinal, $|\mathcal{M}|$ the number of regions of the segmentation model $\mathcal{M}$, $|L|$ is the number of edges present in the region adjacency graph associated to the model $\mathcal{M}$. $\mathcal{D}_{\mathcal{R}_i}(\mathcal{M}, \mathcal{I})$ measures how the region $\mathcal{R}_i$ of the model $\mathcal{M}$ fits the initial data of the image $\mathcal{I}$. This measure can be performed with a piecewise or a Gaussian modeling of the regions of the model $\mathcal{M}$. $d(\mathcal{R}_i, \mathcal{R}_j)$ represents the distance measure between two regions and several distribution distances can be used. In this paper, Gaussian modeling of regions and Mahalanobis distance between regions are considered. A region $\mathcal{R}_i$ has $|\mathcal{R}_i|$ pixels denoted by the set $(X_1, X_2, \ldots, X_{|\mathcal{R}_i|})$ where each pixel is described by X_k^j, $j \in [1, 3]$ (three components). $\overline{\mathcal{R}_i}$ is the average color of the region $\mathcal{R}_i$. The covariance matrix V is given by

$$V(j,k) = \frac{1}{|\mathcal{R}_i|} \sum_{p=1}^{|\mathcal{R}_i|} (X_p^j - \overline{\mathcal{R}_i^j})(X_p^k - \overline{\mathcal{R}_i^k}) \tag{9}$$

and λ_j is its jth eigenvalue. For a gaussian modeling (Guigues, 2003),

$$\mathcal{D}_{\mathcal{R}_i}(\mathcal{M}, \mathcal{I}) = |\mathcal{R}_i| log(detV) = |\mathcal{R}_i| \sum_{j=1}^{3} log(\lambda_i). \tag{10}$$

The data fidelity term $D(\mathcal{M}, \mathcal{I})$ is the same as for MRF (Li, 1995; Boykov et al., 2001), weighted by the normalization factor α_c proposed by Liu and Yang (1994) to penalize segmentation models with too many regions. This term takes into account the homogeneity of the regions by means of a similarity measure from each pixel to its associated region.

$C(\mathcal{M}, \mathcal{I})$ describes the complexity of the segmentation model by describing the complexity of the region adjacency graph (RAG) associated to the segmentation model. This complexity is evaluated in terms of connected regions similarities and in terms of

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005

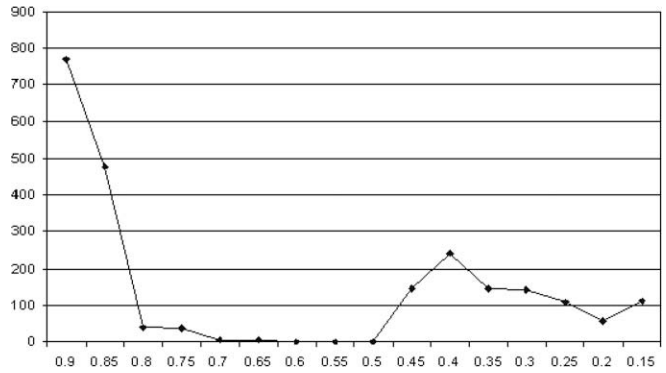

Fig. 5. The energy according to β for the *RB* 2D histogram of Fig. 2a.

the topological structure of the RAG. The data fidelity term is globally decreasing over a set of nested partitions whereas the term $C(\mathcal{M},\mathcal{I})$ is globally increasing. These two antagonist terms reflect the trade-off that one has to face with between data fidelity and model complexity. The minimization of the proposed energy $E(\mathcal{M},\mathcal{I})$ can be performed so as to find a local minimum which corresponds to a trade-off between fidelity and complexity, the segmentation being neither too fine, nor too coarse.

This energy formulation can be used to tune the parameter β of the above proposed unsupervised morphological clustering. If one takes decreasing values of this parameter, different clustering scales of the image $\mathcal{I}$ are obtained and the obtained segmentations have a decreasing number of regions. The segmentations are not necessarily nested from one level to another but the lower β, the coarser the segmentation. The energy evaluated on a set of segmentations obtained for different values of β will therefore progressively decrease (convex-like energy). We propose to evaluate the energy for values of β by steps of 0.05 starting from 1 (we recall here that the lower β the stronger the regularization effect). Successive clusterings are computed for successive values of β and the process stops when a local minimum of the energy is found. By this method one can automatically find the best suited level of clustering and the number of classes of the clustering is automatically determined. Fig. 5 presents several values of the energy for the clustering of the *RB* 2D histogram of Fig. 2a, the best value of β being 0.6. Our energy model does not evaluate an obtained clustering in the color space domain but in the image plane domain which enables us to take into account the spatial repartition of the colors in the image. The clustering will therefore be more accurate since we try to construct classes of pixels that correspond to connected regions in the image: this is a colorimetric clustering tuned on spatial connectedness properties. Fig. 4b–d present the clustering segmentation results for the synthetic image depicted by Fig. 2a. It is important to note that scale selection is done separately for each of the three bivariate histograms and clustering scales can be different. This ensures the fact that fine scales are retained only when necessary (typically when clusters are separable only in a given bivariate histogram, see Fig. 2 for green colors).

4. Fusion and simplification by region merging

The above clustering method is applied to the three 2D histogram images obtained from a color image and three different segmentation maps are obtained $(\mathcal{I}_1,\mathcal{I}_2,\mathcal{I}_3)$. The resulting segmentations have to be combined altogether to provide the final segmentation of the image. Fig. 6 presents the fusion and simplification strategy we have adopted. Since the 2D histograms do not hold the same information, different color clusters can be extracted from the histograms and the segmentations maps can be totally different. Therefore, before combining the three segmentations, each one is simplified by region merging according to their area. The area for which regions can merge is considered as being equal to 9 (the size of a 3×3 structuring element). Fig. 4e–g show the simplification effect of the clustering segmentation maps depicted by Fig. 4b–d. From the three simplified segmentation maps $(\mathcal{I}'_1,\mathcal{I}'_2,\mathcal{I}'_3)$, a fourth $\mathcal{I}_4$ is generated by superposition (Fig. 4h) (Xue et al., 2003):

$$\mathcal{I}_4 = (\mathcal{I}'_1 \wedge \mathcal{I}'_2) \wedge \mathcal{I}'_3, \tag{11}$$

where $\wedge$ is the operator that computes the superposition of two segmentation maps. The superposition of two segmentation maps (two partitions) is the largest partition that is a sub-partition of both partitions. A typical drawback of the superposition as a fusion operator is to produce over-segmentation. A lot of small pixel differences can occur even if the segmentation maps are very close, therefore a simplification is needed. Fig. 7b–d present the three segmentation maps after clustering and area region merging of Fig. 7a. One can see that although the segmentations seem very close, their superposition is over segmented (Fig. 7e). Indeed, the segmentation still contains too many regions to describe in a compact way the original color image. Clustering neglects the spatial context of the image plane since the only coherence of color clusters has been used. This has a strong effect on the produced segmentation since color clusters do not necessarily correspond to significant regions in the image one wants to segment. Since it is not possible, with a clustering approach used for pre-segmentation, to guarantee that color clusters

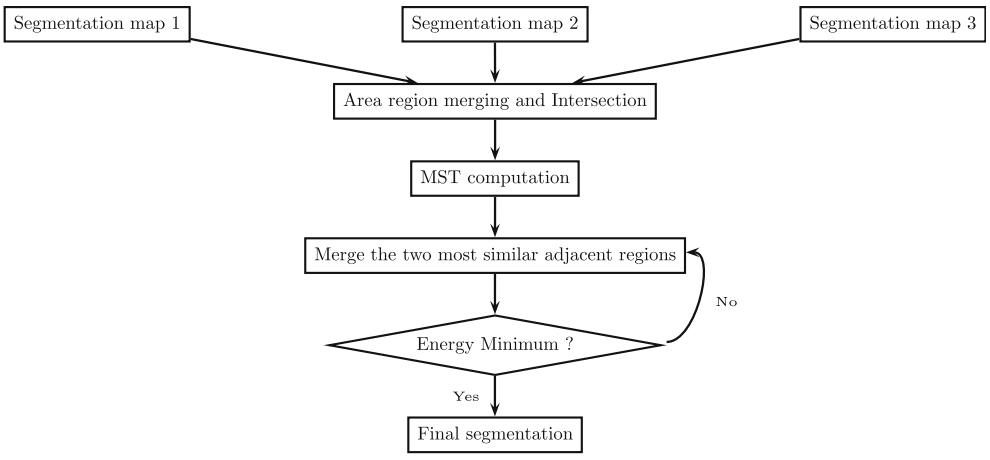

Fig. 6. Fusion and simplification of segmentation maps obtained from bivariate histograms.

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005

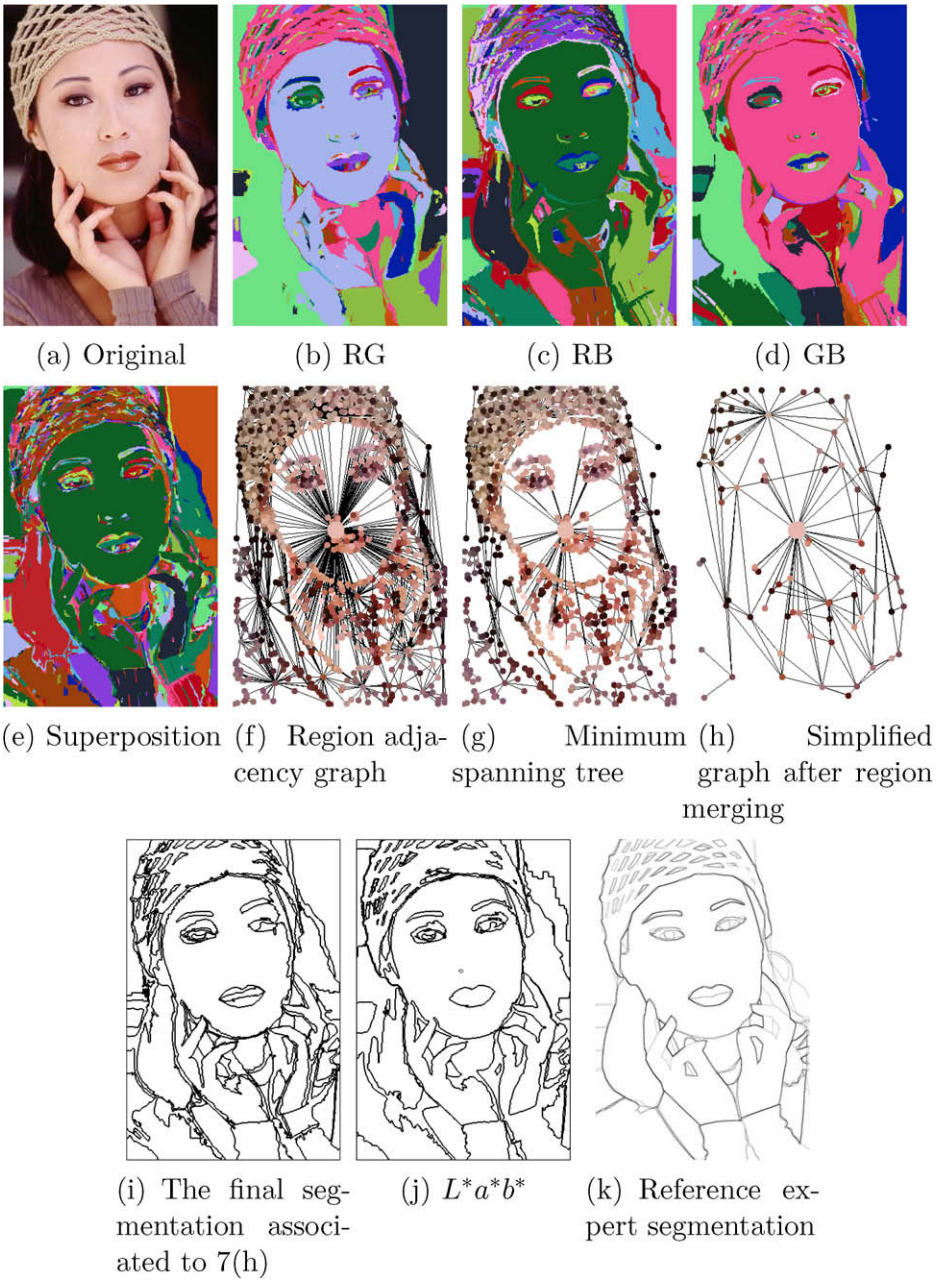

Fig. 7. Illustration of the successive steps of the proposed method on a sample image (see text).

do correspond to effective regions in the image, a further step of simplification is needed (Makrogiannis et al., 2005). To merge similar regions which have been extracted as different ones by the clustering, the spatio-colorimetric information in the image plane has to be taken into account: this corresponds to perform region merging. Therefore, a region adjacency graph is associated to the segmentation map (Fig. 7f). The edges of the graph are weighted by the distance between the regions (as defined by $d(\mathscr{R}_i, \mathscr{R}_j)$ in the previous Section). This graph is simplified by computing its Minimum Spanning Tree (MST, see Fig. 7g) to have a sparser representation (Haxhimusa et al., 2005). Then, the edges of the MST are ranked in a priority queue according to their weights and region merging is performed by merging pairwise the most similar regions (as defined in (Garrido et al., 1998)). The merging continues until a termination criterion is reached (Salembier and Garrido, 2000; Makrogiannis et al., 2001). Moreover, no merging criterion is used: regions successively merge while accessing the priority queue. We propose to use once again the energy formulation (Eq. 8) as a termination criterion: the presence of a local minima assessing the termination. With respect to the formulation of the energy, the research space covered by the region merging is different from the research space covered by the unsupervised morphological clustering. Therefore, the energy criterion can be used once again for scale selection. For clustering, the scale selection covers a part of color space domain, whereas for region merging a part of the image domain is covered. With this automatic region merging, one recovers the spatial coherence between regions and the region merging is stopped when a satisfying segmentation is obtained: neither too coarse, nor too fine (Fig. 7h–i).

5. Results and discussion

In this Section, quantitative and qualitative results of the proposed method are provided in contrast to two state-of-the-art

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005

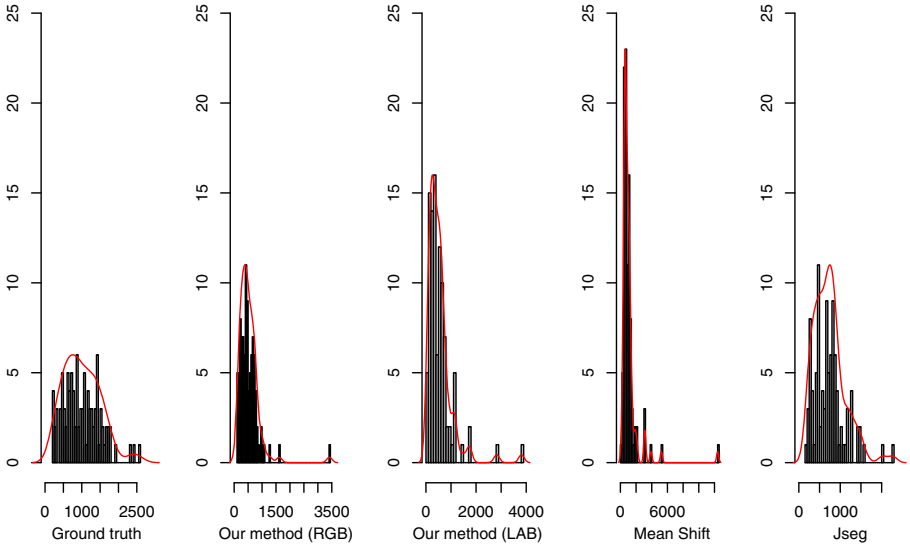

Fig. 8. MSE measure histograms.

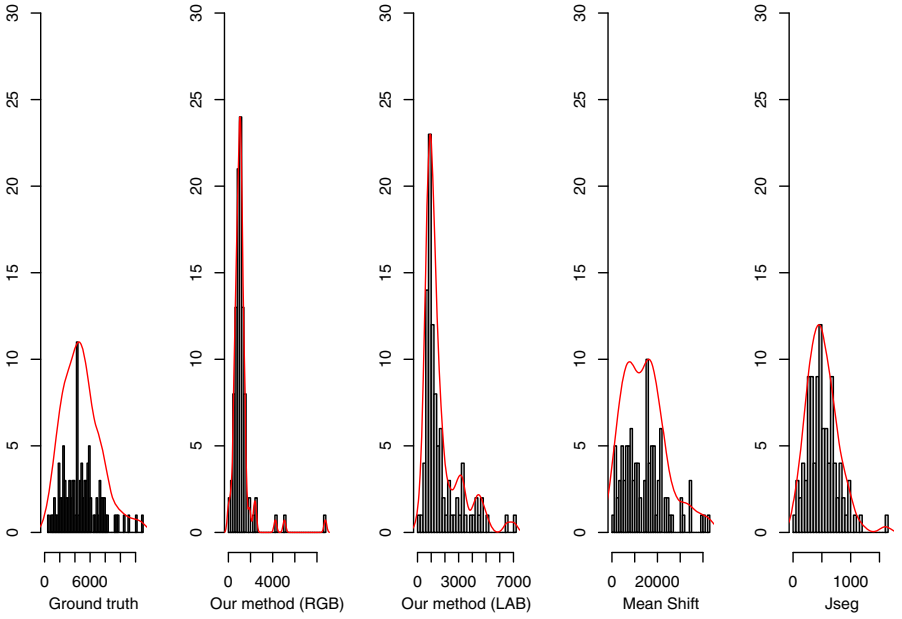

Fig. 9. Energy measure histograms.

methods (mean shift (Comaniciu and Meer, 2002) and Jseg (Deng and Manjunath, 2001)[1]). Mean shift detects modes of the probability density of color occurring jointly in image and color domains. Jseg proceeds to the multi-scale analysis in the image domain of a clustering map obtained from quantization in color domain. Training set (200 images) of the Berkeley Segmentation Data Set (Martin et al., 2001) is considered. To have a comparison between segmentation methods, a similar methodology to (Haxhimusa et al., 2005) is adopted. For each of the images in the test set, we have calculated two measures for evaluating the segmentation quality: the mean squared error (MSE) between the color median image of the segmentation and the original, and our energy evaluation measure. Finally, to compare segmentation methods, histograms of evaluation methods are computed on the whole test set (Figs. 8 and 9 with estimated histogram modes superimposed). To compare methods, we compare histograms with respect to the histogram of evaluation measures of human segmentations. For the case of our segmentation method, we also study the influence of color space and provide results in *RGB* and $L^*a^*b^*$ color spaces. Fig. 7j presents the segmentation of Fig. 7a in the $L^*a^*b^*$ color space, one can see that although being close to the segmentation in the *RGB* color space (Fig. 7h), some differences occur. $L^*a^*b^*$ color space is very popular because it provides better distance measures between colors. However, it is difficult to say if this is essential for segmentation purposes.

In terms of MSE (Fig. 8), the lower the value, the closer the segmentation is to the original. Jseg has the closest histogram to the human segmentations. Mean shift provides the worst results. Our method performs better in *RGB* than in $L^*a^*b^*$. Moreover, results of our method in *RGB* are close to the one obtained with Jseg. In terms of MSE, segmentation methods can be classified with this

[1] With the code obtained from pages: http://vision.ece.ucsb.edu/segmentation/jseg/software/ and http://www.caip.rutgers.edu/riul/research/code/EDISON/index.html.

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005

ordering: Jseg $\geqslant$ Our_{RGB} $\geqslant$ $\text{Our}_{L^*a^*b^*}$ $\geqslant$ meanshift where $\geqslant$ means"better than".

MSE measure favors over-segmentation and does not really reflect the complexity of a segmentation. On the opposite, our energy measure also takes into account the complexity of the segmentation in addition to a data attachment term. Fig. 9 shows histograms of our energy evaluation measure for the considered segmentation methods. A look at the histogram of human segmentations shows that these segmentations have a high complexity. Since images of the data set are natural images, this seems natural that a reasonable level of complexity has to be considered for having good segmentations (in terms of human evaluation). In fact, humans have also made a compromise between precision and complexity. If we analyze the histograms of the segmentation methods with respect to the histogram of human segmentations, we can see that Jseg tends to under segment images whereas Mean Shift produces over-segmentation. Our method seems to offer a best compromise between precision and complexity. Moreover, results in *RGB* color space are more consistent than in $L^*a^*b^*$. In terms of a trade-of between complexity and precision, segmentation methods can be classified with this ordering: Jseg $\leqslant$ Our_{RGB} $\geqslant$ $\text{Our}_{L^*a^*b^*}$ $\geqslant$ meanshift.

Fig. 10 shows segmentation results on sample images of the Berkeley segmentation data set. We can see the effects that were shown in terms of histograms: Jseg produces under-segmentation, Mean Shift produces over-segmentation and our method corresponds to a trade-off between a coarse and a fine segmentation. Moreover, the boundaries of the objects in segmented images are visually better delineated with our method. To compare the

Fig. 10. Sample segmentation results on images of the Berkeley segmentation benchmark data set. First row: original images, second row: ground truth, third row: *RGB* segmentations, fourth row: $L^*a^*b^*$ segmentations, fifth row: mean shift segmentation, sixth row: Jseg segmentation.

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005

computational complexity of our algorithm with Mean Shift and Jseg, we have measured the segmentation computation time on all images of the Berkeley segmentation data set (on a standard computer with 2.4 GHz Intel Xeon Processor and 1GB of RAM). The average segmentation time for one image is 10.58 ± 5.5 s for our algorithm, $14.52 \pm 3.3s$ for Jseg and 11.76 ± 3.7 s for mean shift.

Finally, robustness against Gaussian and impulse noise of our segmentation method is studied and compared with mean shift and Jseg. The reference image Parrots is considered. Segmentations and their evaluation are computed in terms of PSNR (Fig. 11) and energy (Fig. 12) for different amounts of noise ($\sigma \in [0,20]$ for Gaussian noise and the percentage of corrupted pixels belongs to $[0,20]$% for impulse noise). In terms of PSNR, the quality of segmentation is degraded with noise introduction. However, obtained segmentation tend to be of better quality as noise level increases. This can be surprising but can be explained while having a look at the energy evaluation (Fig. 12). One can see that the energy increases with noise assessing an increase of segmentation complexity. This is particularly the case for Gaussian noise. However, the method appears to be more robust to impulse noise.

Regarding the results obtained with Mean Shift and Jseg, our algorithm has a similar behavior for small portions of noise. For Gaussian noise, Means Shift and Jseg enable to obtain segmentation of lower complexities. For impulse noise, our algorithm obtains results similar to Mean Shift in terms of segmentation complexity while being better in terms of PSNR. Results obtained with Jseg always correspond to under segmentation. To conclude these experiments, our algorithm is competitive for the case of impulse noise but not Gaussian noise. However, if one modifies the energy formulation (Eq. 8) to decrease the importance of the data fidelity term $D(\mathscr{M},\mathscr{I})$ with parameter α_d, better results will be obtained for the case of Gaussian noise.

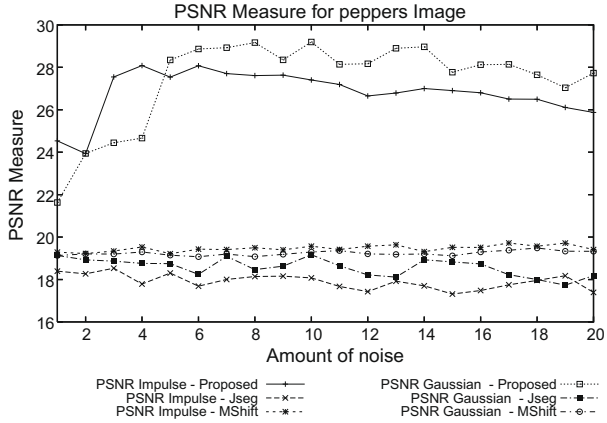

Fig. 11. PSNR measure on peppers image.

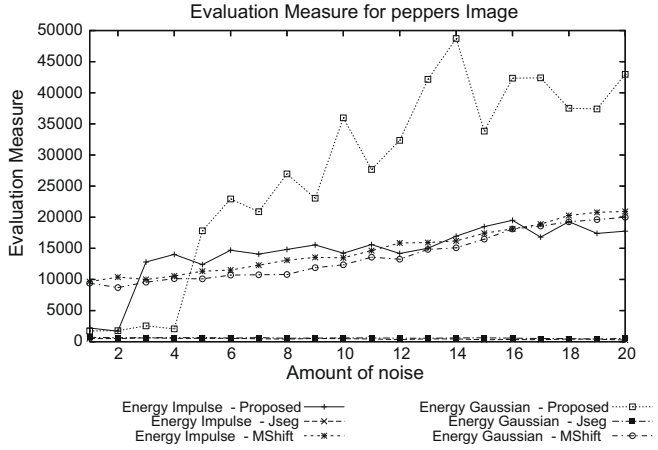

Fig. 12. Energy measure on peppers image.

6. Conclusion

In this paper, a new approach to color image segmentation has been proposed. Unsupervised morphological clustering is used to analyze bivariate histograms. This clustering being multi-scale, scale parameter is tuned by minimizing an energy model blending together segmentation fidelity and segmentation complexity. After intersecting the obtained region maps, a merging of regions is performed with as termination criterion the same energy model. Segmentation results are compliant with the energy properties: a trade-off between fidelity and complexity is obtained. Experimental results show that the proposed method is a good compromise in-between state-of-the-art methods.

References

Agarwal, S., Madasu, S., Hanmandlu, M., Vasikarla, S., 2005. A comparison of some clustering techniques via color segmentation. In: ITCC, vol. 2, pp. 147–153.

Boykov, Y., Veksler, O., Zabih, R., 2001. Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Machine Intell. 23 (11), 1222–1239.

Busin, L., Vandenbroucke, N., Macaire, L., Postaire, J.-G., 2004. Color space selection for unsupervised color image segmentation by histogram multithresholding. In: Proc. IEEE Internat. Conf. on Image Processing (ICIP'04), pp. 203–206.

Celenk, M., 1990. A color clustering technique for image segmentation. Comput. Vision Graphics Image Process. 52, 145–170.

Cheng, H.-D., Jiang, X., Sun, Y., Wang, J., 2001. Color image segmentation: Advances and prospects. Pattern Recognition 34 (12), 2259–2281.

Comaniciu, D., Meer, P., 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Machine Intell. 24 (5), 603–619.

Deng, Y., Manjunath, B.S., 2001. Unsupervised segmentation of color-texture regions in images and video. IEEE Trans. Pattern Anal. Machine Intell. 23 (8), 800–810.

Garrido, L., Salembier, P., Garcia, D., 1998. Extensive operators in partition lattices or image sequence analysis. Signal Process. 6 (2), 157–180.

Géraud, T., Strub, P., Darbon, J., 2001. Color image segmentation based on automatic morphological clustering. In: IEEE Internat. Conf. on Image Processing, vol. 3, pp. 70–73.

Gillet, A., Macaire, L., Botte-Lecocq, C., Postaire, J.-G., 2000. Fuzzy unsupervised color image segmentation. In: Proc. CGIP, pp. 141–146.

Gillet, A., Macaire, L., Botte-Lecocq, C., Postaire, J.-G., 2001. Color image segmentation by fuzzy morphological transformation of the 3d color histogramm. In: FUZZ-IEEE, p. 824.

Guigues, L., 2003. Modèles multi-échelles pour la segmentation d'images. Ph.D. Thesis, Université de Cergy-Pontoise.

Guigues, L., Men, H.L., 2003. Scale-sets image analysis. In: Proc. IEEE Internat. Conf. on Image Process. (ICIP'03), vol. 2, pp. 45–48.

Haxhimusa, Y., Ion, A., Kropatsch, W.G., Illetschko, T., 2005. Evaluating minimum spanning tree based segmentation algorithms. In: Proc. 11th Internat. Conf. on Computer Analysis of Images and Patterns (CAIP 2005), September, LNCS, vol. 3691, pp. 579–586.

Kakumanu, P., Makrogiannis, S., Bourbakis, N.G., 2007. A survey of skin-color modeling and detection methods. Pattern Recognition 40 (3), 1106–1122.

Koepfler, G., Lopez, C., Morel, J.-M., 1994. A multiscale algorithm for image segmentation by variational method. SIAM J. Numer. Anal. 31 (1), 282–299.

Kurugollu, F., Sankur, B., Harmanci, A., 2001. Color image segmentation using histogram multithresholding and fusion. Image Vision Comput. 19 (13), 915–928.

Lezoray, O., 2004. An unsupervised color image segmentation based on morphological 2D clustering and fusion. In: Proc. CGIV'2004, pp. 173–177.

Lezoray, O., Cardot, H., 2002. Histogram and watershed based segmentation of color images. In: Proc. CGIV'2002, pp. 358–362.

Lezoray, O., Cardot, H., 2003. Hybrid color image segmentation using 2D histogram clustering and region merging. In: Proc. ICISP'2003, vol. 1, pp. 22–29.

Li, S., 1995. Markov Random Field Modeling in Computer Vision. Springer-Verlag.

Liu, J., Yang, Y.-H., 1994. Multiresolution color image segmentation. IEEE Trans. Pattern Anal. Machine Intell. 16 (7), 689–700.

Lucchese, L., Mitra, S., 2001. Color image segmentation: A state-of-the-art survey. In: Indian National Science Academy (INSA-A). Image Process. Vision Pattern Recognition 67 (2), 207–221.

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005

Makrogiannis, S., Economou, G., Fotopoulos, S., 2001. A graph theory approach for automatic segmentation of color images. In: Internat. Workshop on Very Low Bit-rate Video, pp. 162–166.

Makrogiannis, S., Economou, G., Fotopoulos, S., Bourbakis, G., 2005. Segmentation of color images using multiscale clustering and graph theoretic region synthesis. IEEE Trans. Systems Man Cybernet. A 35 (2), 224–238.

Martin, D., Fowlkes, C., Tal, D., Malik, J., 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proc. 8th Internat. Conf. on Computer Vision, July, vol. 2, pp. 416–423.

Matas, J., 1995. Colour-based object recognition. Ph.D. Thesis, University of Surrey.

Mumford, D., Shah, J., 1989. Optimal approximations by piecewise smooth functions and associated variational problems. Comm. Pure Appl. Math. 17 (4), 577–685.

Park, S., Yun, I., Lee, S., 1998. Color image segmentation based on 3-D clustering: Morphological approach. Pattern Recognition 31 (8), 1061–1076.

Plataniotis, K., Venetsanopoulos, A., 2000. Color Image Processing and Applications. Springer-Verlag.

Poggio, T., Torre, V., Koch, C., 1985. Computational vision and regularization theory. Nature 317, 314–319.

Postaire, J., Zhang, R., Lecocq-Botte, C., 1993. Cluster analysis by binary morphology. IEEE Trans. Pattern Anal. Machine Intell. 15 (2), 170–180.

Salembier, P., Garrido, L., 2000. Binary partition tree as an efficient representation for image processing, segmentation and information retrieval. IEEE Trans. Image Process. 9 (4), 561–576.

Scheunders, P., 1997. A genetic c-means clustering algorithm applied to color image quantization. Pattern Recognition 30 (6), 859–866.

Serra, J., 1982. Image Analysis and Mathematical Morphology. Academic Press, London.

Shafarenko, L., Petrou, H., Kittler, J., 1998. Histogram-based segmentation in a perceptually uniform color space. IEEE Trans. Image Process. 7 (9), 1354–1358.

Soille, P., 1996. Morphological partitioning of multispectral images. J. Electron. Imaging 18 (4), 252–265.

Soille, P., 2004. Morphological Image Analysis: Principles and Applications. Springer-Verlag.

Uchiyama, T., Arbib, M.A., 1994. Color image segmentation using competitive learning. IEEE Trans. Pattern Anal. Machine Intell. 16 (12), 1197–1206.

Vannoorenberghe, P., Colot, O., Brucq, D.D., 1999. Color image segmentation using dempster-shafer's theory. In: Internat. Conf. on Image Processing, vol. 4, pp. 300–303.

Xue, H., Géraud, T., Duret-Lutz, A., 2003. Multi-band segmentation using morphological clustering and fusion application to color image segmentation. In: Proc. IEEE Internat. Conf. on Image Processing (ICIP'03), vol. 1, pp. 353–356.

Please cite this article in press as: Lézoray, O., Charrier, C., Color image segmentation using morphological clustering and fusion ..., Pattern Recognition Lett. (2008), doi:10.1016/j.patrec.2008.11.005