Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition Letters 25 (2004) 1133–1141
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Image segmentation based on maximum-likelihood estimation and optimum entropy-distribution (MLE–OED)

J. Xie *, H.T. Tsui

Department of Electronic Engineering, Computer Vision and Image Processing Laboratory, The Chinese University of Hong Kong, Shatin, Hong Kong

Received 28 April 2003; received in revised form 17 February 2004
Available online 16 April 2004

Abstract

A novel method based on MLE–OED is proposed for unsupervised image segmentation of multiple objects with fuzzy edges. It adjusts the parameters of a mixture of Gaussian distributions via minimizing a new loss function. The loss function consists of two terms: a local content fitting term, which optimizes the entropy distribution, and a global statistical fitting term, which maximizes the likelihood of the parameters for the given data. The proposed segmentation method is validated by experiments on both synthetic and real images. The experimental results show that the proposed method outperformed two popular methods.

Keywords: Segmentation; EM; Gradient; Entropy

1. Introduction

Image segmentation is a key basis of many higher level image processing activities such as visualization, compression, and image guided medical diagnoses. Numerous algorithms using different approaches have been proposed for image segmentation. These approaches include local edge detection (e.g. Ma and Manjunath, 1997), deformable curves (e.g. Xu and Jerry, 1998), morphological region-based approaches (e.g. Deng and Manjunath, 2001; Zhu and Alan, 1996; Daniel and Chuang, 2001), global optimization approaches on energy functions and stochastic model-based methods (e.g. Punam and Jayaram, 2001; Serge and Chad, 1998; Wang, 1998; Zhang et al., 2001).

Some intensity-based methods such as thresholding and histogram-based finite mixture models are easy to be formulated and fast. However they often fail to segment objects with low contrast or noisy images with varying background. It is noted that these methods do not make use of the spatial morphological information of images. Though there are some adaptive (local) thresholding algorithms, they also suffer from this trait.

On the other hand, some other methods such as morphological segmentation, region growing (e.g.

* Corresponding author.
E-mail addresses: jxie@ee.cuhk.edu.hk (J. Xie), httsui@ee.cuhk.edu.hk (H.T. Tsui).

0167-8655/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2004.03.013

Deng and Manjunath, 2001) and deformable curves (e.g. Xu and Jerry, 1998), mainly focus on spatial information such as local structures or regions. They have proven to be useful for many image segmentation applications. Zhu and Alan (1996) presented an algorithm called Region Competition to segment multi-band images. This algorithm combined the aspects of snakes/balloons and region growing. Deng and Manjunath (2001) proposed a criterion for "good" segmentation of color-texture images and video using region growing. Xu and Jerry (1998) used an improved snake with a new external force called GVF to segment the endocardial in the magnetic resonance images of human heart. But most of them need a manual initial input such as seeds or a flexible boundary model near the vicinity of the structure to be segmented. Meanwhile, these methods are often sensitive to these initial conditions. Moreover, all these methods cannot segment the entire image due to the lack of a global criterion.

There are also some segmentation methods that take into account both the spatial information and global intensity distribution properties. Zhang et al. (2001) proposed a hidden Markov random field model (HMRF) derived from the way in which the spatial information is encoded through the mutual influences of neighboring sites. They used an expectation maximization (EM) algorithm to fit the HMRF model. Serge and Chad (1998) modelled a joint distribution of color, texture and position features with a mixture of Gaussians and estimated the parameter of the model using the EM algorithm. Then the resulting pixel-cluster memberships provide a segmentation of the image. They appended the position of the pixel to the feature vector so that their segmentation method can consider both the spatial distribution of intensities and the spatial morphological information in the image. More recently, Punam and Jayaram (2001) proposed a thresholding method called Minimization of homogeneity- and uncertainty-based energy (MHUE). It is based on a postulate that in any scene with fuzzy boundaries, for optimum partitioning of object classes, points with high class uncertainty will have low region homogeneity. The main idea in the method of MHUE is the observation that the pixels with high class uncertainty accumulate mostly around object boundaries. A threshold is selected accordingly.

In this paper, we model the entire image as a mixture of Gaussians and construct a new loss function based on spatial information and global statistical criterion. In our proposed method (MLE–OED), an optimal segmentation is obtained by maximizing the likelihood of the model as well as optimizing the distribution of entropy over the image. This method can guarantee a good match of the model with the image and avoid to assign high class uncertainty to only these pixels on the "strongest" edges.

This paper is organized as follows: In Section 2.1, we describe the feature of the spatial information used in our method. In Section 2.2, we present and explain the formulation and concept of generalized intensity-based class uncertainty. The new loss function used in MLE–OED is proposed in Section 2.3. Experimental results and the comparison between our method with methods of MHUE and intensity-based EM are presented in Section 3. Finally, our conclusions are stated in Section 4.

2. Theory and algorithm

2.1. Measure of spatial information

There are many ways to quantify the spatial information of images such as pixel's coordinates, gradient, shape and texture. In the method of MHUE, the authors consider the region homogeneity as a property of every point in the given scene. This homogeneity depends on point adjacency and fuzzy affinity that were introduced in connection with the development of fuzzy connectedness principles. To segment images containing objects with fuzzy boundaries, we consider the edge information as an important factor for a good segmentation. Instead of taking gradient directly, we use the gradient vector flow (GVF), proposed by Xu and Jerry (1998). This will produce a better effect because the value of GVF will nearly equal to the gradient of the image when the gradient value is large, but not be equal to zero in a relative homogeneous region such as the un-

sharpen boundary areas. In the GVF framework, a two dimensional vector field $v(x,y) = [u(x,y), v(x,y)]$, called gradient vector field, is used to minimize the following energy:

$$E(v) = \int\int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |v - \nabla f|^2 \,\mathrm{d}x\,\mathrm{d}y, \quad (1)$$

where μ a blending parameter and ∇f the gradient of edge map. According to this objective function, areas with constant intensity are dominated by the partial derivatives of the vector field, resulting in a smooth flow map. On the other hand, when there are variations on the boundary space, the term that dominates the energy is the second one, leading to the vector $v = \nabla f$. A more detailed interpretation of this energy can be found in (Xu and Jerry, 1998). An example of gradient, GVF and region homogeneity employed in MHUE is shown in Fig. 1. We can see that there are many local structures missed in the region homogeneity map (Fig. 1(d)) while the gradient map (Fig. 1(b)) is more abrupt and coarse than the GVF map in Fig. 1(c).

2.2. Intensity-based class uncertainty

Intensity-based class uncertainty principles have been used by several researchers. Here, we give its generic formulation as we have used in our method. The idea behind intensity-based class uncertainty is to determine the uncertainty of classifying a point into a certain object class based on its scene intensity and the a priori knowledge of the intensity probability distributions of different classes. In general case, we consider a scene that consists of k objects S_i, $i = 1, 2, \ldots, k$. A gray scale image with N pixels has been captured to present the scene. Let n_i denote the number of the pixels from object S_i and $\sum_{i=1}^{k} n_i = N$. Assume that if we select a pixel x from the image, randomly, the probability that x belongs to S_i is n_i/N, i.e.,

$$p(x \in S_i) = \frac{n_i}{N}. \quad (2)$$

Let $p_i(g)$ denote the a priori probability that a pixel x of class S_i has the intensity value g:

$$p_i(g) = P(I(x) = g | x \in S_i), \quad (3)$$

where $I(x)$ represents the gray value of pixel x; $g = 0, 1, \ldots, (L-1)$, i.e., the gray level values of the gray scale image are discretized into L levels ranging from 0 to $L-1$ in unit step.

To segment the image into k objects, we model the scene as a mixture of k Gaussians in the intensity space. The PDF of pixel x in the input image is

$$P(x|\Theta) = \sum_{i=1}^{k} \alpha_i f_i(I(x)|\theta_i), \quad (4)$$

where

$$f_i(g|\theta_i) = \frac{1}{\sqrt{2\pi}\delta_i} \exp\left(-\frac{(g-u_i)^2}{2\delta_i^2} \right) \quad (5)$$

α_i's are mixing weights and $\sum_{i=1}^{k} \alpha_i = 1$. Θ refers to the collection of parameters $\Theta = (\alpha_1, \ldots, \alpha_k; u_1, \ldots, u_k; \delta_1, \ldots, \delta_k)$.

Then the a posteriori probability that pixel x belongs to the object S_i is

$$P_i(x|\Theta, g) = p(x \in S_i | \Theta, I(x) = g) = \frac{n_i p_i(g)}{N P(x|\Theta)} \quad (6)$$

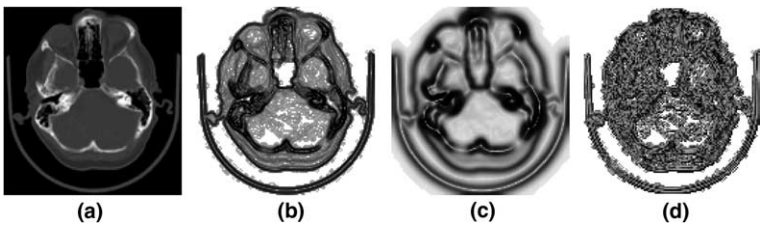

Fig. 1. Three spatial information representations of an axial non-contrast CT image: (a) original brain image, (b) the gradient map, (c) the map of GVF, (d) the map of region homogeneity in MHUE.

According to Shannon's theory, the uncertainty of classifying pixel x into a certain class is the entropy of the posterior probability in Eq. (6). So the class uncertainty of a given classification can be measured by the entropy expression

$$H(x|I(x)=g) = -\sum_{i=1}^{k} P_i(x|\Theta,g)\log P_i(x|\Theta,g). \tag{7}$$

Finally, for a given Θ, the class uncertainty of pixel x is given by Eqs. (2)–(7).

2.3. The loss function

In the maximum likelihood problem, we have a density function $P(X|\Theta)$ that is governed by Θ which is the set of parameters, and a sample data $X = \{x_1, x_2, \ldots, x_N\}$. Assume the data is independent and identically distributed with distribution in Eq. (4). The density for the sample is

$$P(X|\Theta) = \prod_{n=1}^{N} P(x_n|\Theta). \tag{8}$$

To find the Θ^* that maximizes $P(X|\Theta)$, the EM algorithm is originally described in its general form by Empster (1997). It is a general method of finding the maximum-likelihood estimation of the parameters of an underlying distribution from a given data set when the data is incomplete or has missing values. Several EM based segmentation methods have been proposed (e.g. Serge and Chad, 1998; Wells, 1994). However, these methods do not take advantage of the spatial information. On the other hand, if we pursue the segmentation only via minimizing the weighted sum of class uncertainty such as the entropy function in Eq. (12), it will lead that only the points on the strongest boundaries can get high class uncertainty. Moreover, this strategy cannot guarantee the global match between the model and the data.

In our method, we calculate the Θ^* which maximizes the likelihood of the parameters to the image. Meanwhile we pursue the purpose to make pixels with high class uncertainty accumulate mostly around the fuzzy boundaries. The loss function is defined as

$$L = L_{\text{entropy}} - \lambda L_{\text{likelihood}}, \tag{9}$$

where λ is a scale factor. L_{entropy} and $L_{\text{likelihood}}$ are the entropy term and likelihood term, respectively. They are defined as following:

$$\begin{aligned} L_{\text{entropy}} &= \sum_{i=1}^{N} (1-G(x_i))H(x_i) \\ &\quad + (1-H(x_i))G(x_i) \\ &= \sum_{i=1}^{N} H(x_i) - G(x_i)H(x_i) + G(x_i) \\ &\quad - H(x_i)G(x_i) \\ &= \sum_{i=1}^{N} (1-2G(x_i))H(x_i) + G(x_i), \end{aligned} \tag{10}$$

$$L_{\text{likelihood}} = \sum_{i=1}^{N} \log P(x_i), \tag{11}$$

where G is the normalized GVF value of pixels introduced in Section 2.1, $G \in [0,1]$; H is the normalized entropy (class uncertainty) of pixels described in Section 2.2, $H \in [0,1]$. Then as shown in Fig. 2, only when both G and H are high or low, L_{entropy} is low. Thus the low value of the entropy term indicates high class uncertainty is assigned to

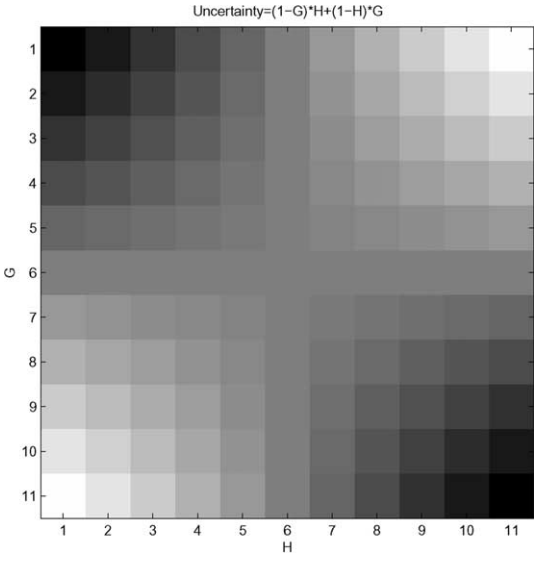

Fig. 2. The property of the entropy term.

pixels with high GVF values or low uncertainty to pixels with low GVF values.

For function L is governed by parameter set Θ, term G is a constant. We can rewrite the entropy term in Eq. (10) as

$$L_{\text{entropy}} = \sum_{i=1}^{N} (1 - 2G(x_i))H(x_i). \tag{12}$$

Finally, we get the loss function for an image with N pixels as

$$L(\Theta) = \sum_{i=1}^{N} (1 - 2G(x_i))H(x_i|\Theta) - \lambda \log P(x_i|\Theta). \tag{13}$$

3. Results and discussion

This section shows several examples of the MLE–OED segmentation method. Results have been compared with those of other two methods: the method using an intensity-based mixture of Gaussians model with EM optimization, and the MHUE method. In our experiments, the scale parameter λ was chosen as $1/N$ such that we could control the entropy and likelihood terms simultaneously.

All methods were implemented using C code. For the MLE–OED method, the GVF map was first computed for each image with the blending

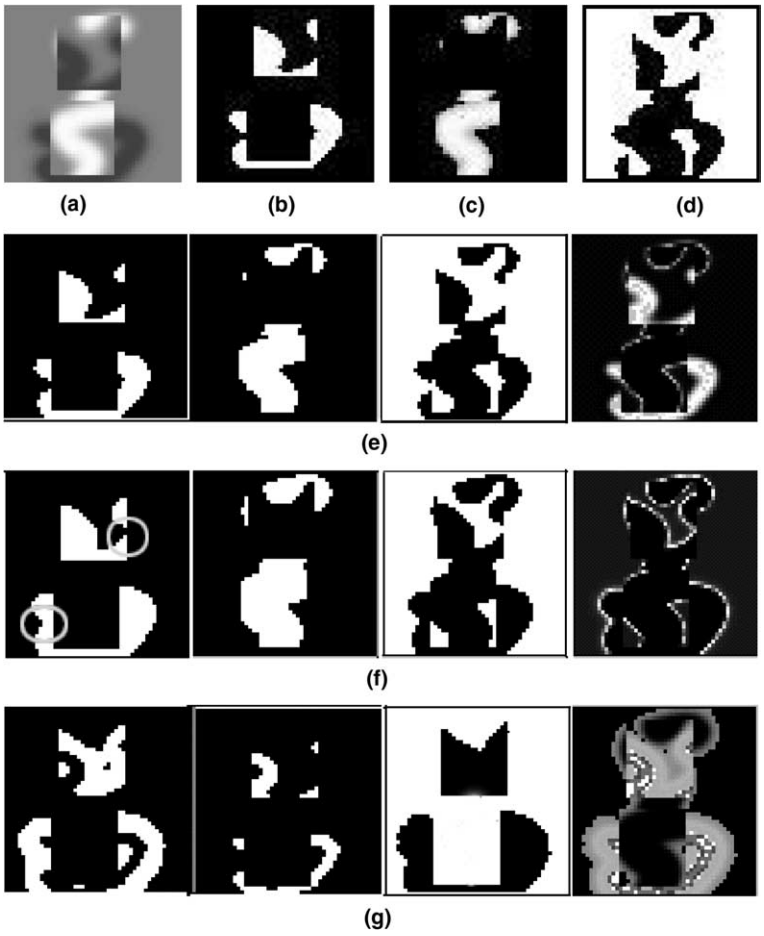

Fig. 3. The experiment on a simulated image: (a) original image, (b) manually segmented black object, (c) manually segmented white object, (d) manually segmented gray background, (e) results of method MLE–OED, (f) results of method EM, (g) results of method MHUE.

parameter $\mu = 0.2$. Then a generalizing step was applied to range G and H from 0 to 1. The means of each Gaussian density were constrained to range from 0 to 255 and the minimization of the loss function L is through the Quasi-Newton algorithm. To improve the efficiency and increase the accuracy, we applied the results Θ_{EM} of the intensity-based EM method as the initial values of the parameters in our method. With this start guess, we calculated set $S_i = \{x | f_i(I(x)|\theta_i) \geqslant f_j(I(x)|\theta_j)\}$, $j = 1, 2, \ldots, k, j \neq i$. Then we set the size of set S_i as the initial value of n_i.

The simulated test image in Fig. 3(a) contains a black object and a bright object on a gray background. Figs. 3(b)–(d) are the manual segmentation results where white pixels in each image represent the black object, the white object and the gray background, respectively. The segmentation results of method MLE–OED, EM and MHUE are shown in the first three images of Fig. 3(e)–(g). The fourth image is the entropy map of each method in which the brighter pixels have higher entropy value. We can see that, in the three entropy maps, high entropy values appeared around the boundaries as assumed. It is noted that the simple intensity-based EM method failed to get some details as marked on Fig. 3(f). On the other hand, in the results of method MHUE, high class uncertainty was only assigned to the few pixels with lowest homogeneity value (i.e., these pixels in the "strongest" boundaries). Thus,

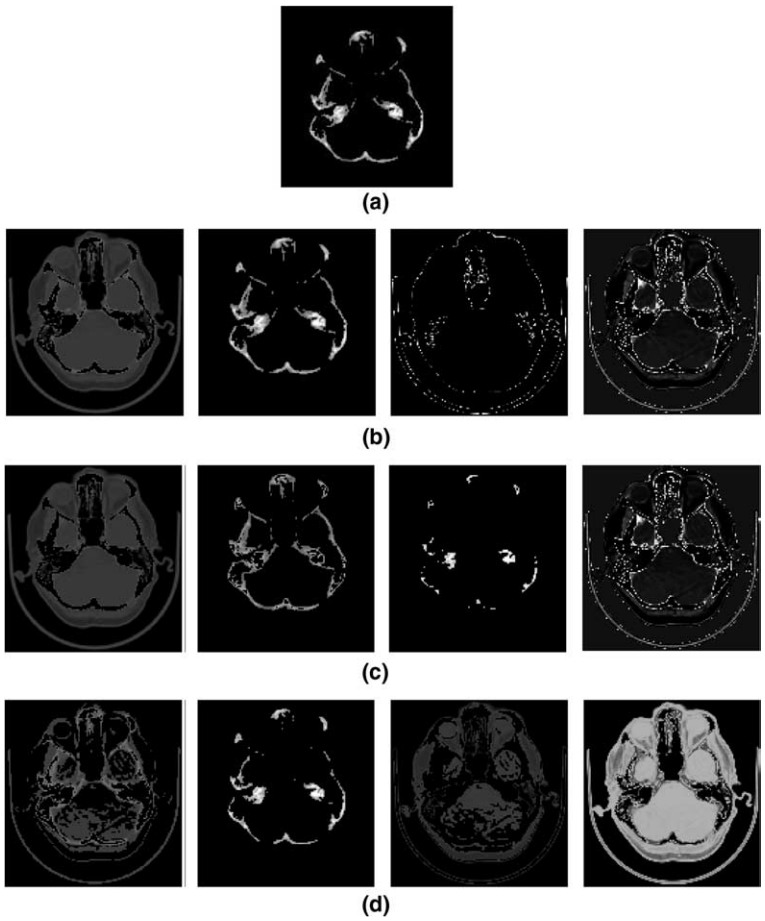

Fig. 4. The experiment on the CT image in Fig. 1(a): (a) manually segmented white object including some sulcus and lesions, (b) results of method MLE–OED, (c) results of method EM, (d) results of method MHUE.

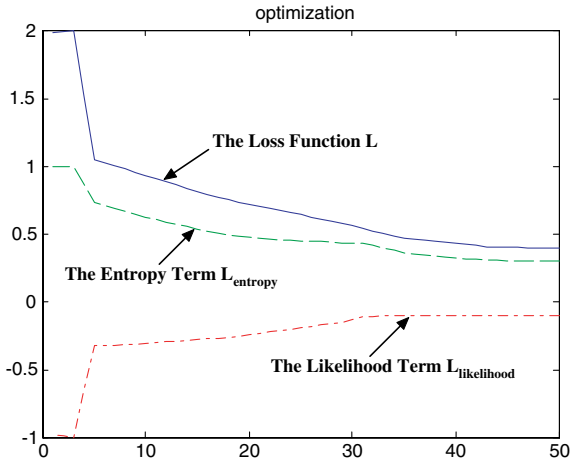

Fig. 5. The optimization process of the experiment on the CT image.

this method missed some relative weak edges and local structures were. This is why the bright object was not segmented correctly by the MHUE method.

Results of an experiment on a real image are shown in Fig. 4. The input image is in Fig. 1(a). It is an axial non-contrast CT image demonstrating lytic lesion in the left occiput region. Figs. 4(b)–(d) are the three segmented objects and the entropy map of each method. Referring to the manual segmented white object, including some sulcus and lesions, in Fig. 4(a), it is obvious that the result of method MLE–OED has more accuracy than those of the other two methods. The optimization process of this experiment is shown in Fig. 5 which illustrates that the likelihood of the model was increased round by round while the entropy term was reduced. This leaded a continuous decrease of the loss function L.

More segmentation results of our method are shown in Figs. 6 and 7. Fig. 6(a) is an endoscopy image about paraesophageal hernia involving distal body and antrum of stomach with organoaxial volvulus without evidence of obstruction. Fig. 6(b) is the manual segmentation of the distal body. Figs. 6(c)–(e) are the results of method EM, MHUE and MLE–OED, respectively. We can observe that the EM method missed the fuzzy area of the body while the MHUE method extracted more fake objects than the MLE–OED method.

Fig. 7(a) shows another test image. It is an angiographic image of the aortic arch demonstrating rapid extravasation of contrast from the apex of the aortic arch distal to the origin of the left subclavian artery. The manually segmented arch is shown in Fig. 7(b) and the results of the

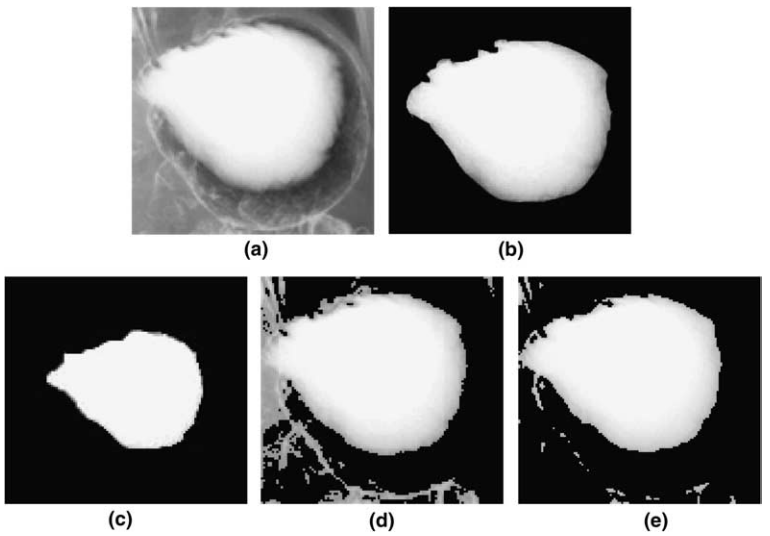

Fig. 6. The experiment on an endoscopy image: (a) the original image, (b) the manual segmentation of the distal body, (c) the result of method EM, (d) the result of method MHUE, (e) the result of method MLE–OED.

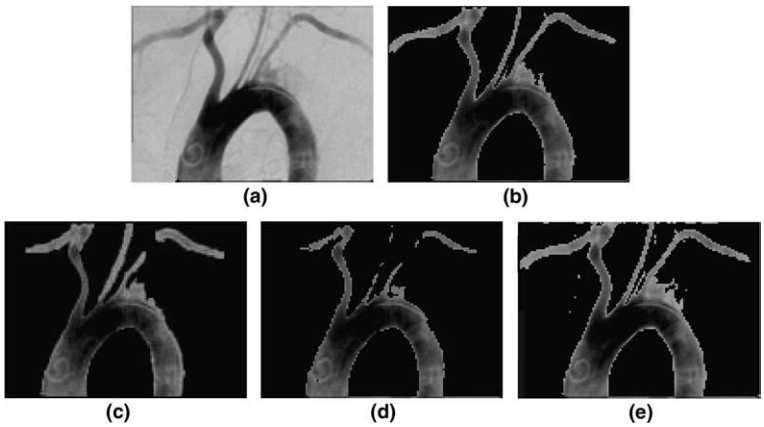

Fig. 7. The experiment on an angiographic image: (a) the original image, (b) the manual segmentation of the arch, (c) the result of method EM, (d) the result of method MHUE, (e) the result of method MLE–OED.

Table 1
Comparison of the three methods in the four experiments

Test ID	Image size (pixels)	Area error (%)			Computational time (s)		
		EM	MHUE	MLE–OED	EM	MHUE	MLE–OED
1	50×50	3.2	10.1	1.2	0.1	0.2	0.4
2	128×135	9.2	7.3	2.1	0.4	1.3	3.0
3	220×206	13.2	12.3	4.8	1.6	2.8	9.3
4	259×189	5.8	8.2	2.6	1.9	3.5	10.5
Mean	–	7.85	9.46	2.68	1.00	1.95	5.8
Std.	–	4.33	2.22	1.53	0.88	1.48	4.88

three methods are shown in Fig. 7(c)–(e), respectively. We can see both method EM and MHUE missed some weak structures while our method obtained a more complete segmentation.

To evaluate our method quantitatively, we compared all the results of the three methods with the manual segmentations in terms of area errors and computational time. From the visual values in Table 1, we observe that the area error of our method in the four experiments vary from 1.2% to 4.8% with mean of 2.68% which is smaller than the errors of the other two methods. On the other hand, it is noted that our method needs a higher computational time which contains the time for parameter initialization via EM method and the time for minimizing the loss function. The low speed occurred because our method needs to optimize the parameters of the model as well as the distribution of the classification uncertainty. Our experiments were performed on a 1.7-GHz Pentium IV PC.

4. Conclusion

This paper presents a novel method (MLE–OED) for gray image segmentation. Based on the assumption that pixels with high class uncertainty accumulate mostly around object boundaries, this method utilizes both the spatial morphological information and a global statistical criterion to implement an unsupervised multiple objects segmentation. One major contribution of this paper is the improvement of the MHUE method via modelling the image as a mixture of Gaussians. This makes our method be able to overcome the weakness of missing weak edges. Another important novelty of this paper is the proposed loss

function. The goal for this loss function is to optimize the parameters of the global model as well as the contribution of entropy over the whole image.

We applied the proposed segmentation method on both simulated and real images. The results of our method are compared with those of manual segmentation, the simple intensity-based EM method and the MHUE method. The results of simulated image segmentations show that the MLE–OED method performs better than the other two methods in terms of capturing morphological details and global matching. Moreover, the experimental results on real images show that the proposed method is valid for realistic applications. The current limitation of our method lies in the high computational time. Further research is needed to increase the speed of the optimization process.

Acknowledgements

This project is partially supported by the Hong Kong RGC Grant CUHK 4378/02E.

References

Daniel, G.-P., Chuang, G., 2001. Extensive partition operators, gray-level connected operators, and region merging/classification segmentation algorithms: Theoretical links. IEEE Trans. Image Process. 10 (9), 1332–1345.

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

Empster, A.P.D., 1997. Maximum likelihood from incomplete data via the Em algorithm. J. Roy. Statist. Soc. 29, 1–38.

Ma, W.Y., Manjunath, B.S., 1997. Edge flow: A framework of boundary detection and image segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, pp. 744–749.

Punam, K.S., Jayaram, K.U., 2001. Optimum image thresholding via class uncertainty and region homogeneity. IEEE Trans. Pattern Recog. Mach. Intell. 23 (7), 689–706.

Serge, B., Chad, C., 1998. Color- and texture-based image segmentation using EM and its application to content-based image retrieval. In: International Conference on Computer Vision, January, Bombay, India, pp. 675–682.

Wang, J.-P., 1998. Stochastic relaxation on partitions with connected components and its application to image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 20 (6), 619–636.

Wells, W., Kikinis, R., Grimson, W., Jolesz, F., 1994. Statistical intensity correction and segmentation of magnetic resonance image data. In: SPIE the Third Conference on Visualization in Biomedical Computing, October.

Xu, C.Y., Jerry, L.P., 1998. Snakes, shapes, and gradient vector flow. IEEE Trans. Image Process. 7 (3), 359–369.

Zhu, S.C., Alan, Y., 1996. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 18 (9), 884–900.

Zhang, Y.Y., Michael, B., Stephen, S., 2001. Segmentation of brain images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imag. 20 (1), 45–57.