Introduction of EM Algorithm into Color Image Segmentation

Tatsuya Yamazaki
ATR Adaptive Communications Research Laboratories,
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288 Japan
yamazaki@acr.atr.co.jp

Abstract - A color image segmentation method is proposed on the basis of the maximum likelihood (ML) estimation. The observed color image is considered as a mixture of multi-variate normal densities and the number of densities is assumed to be known. The EM algorithm is introduced in order to estimate and improve the parameters of the mixture of densities recursively. The initial parameters for the EM algorithm are estimated by the multi-dimensional histogram and the minimum distance clustering methods. After the parameter estimation, the segmented image is obtained by the conventional ML method. Consequently, since no random selection is used for initial parameter estimation, the proposed method is stable and useful for unsupervised image segmentation applications. The performance of the algorithm is demonstrated by real color image segmentation experiments.

I. INTRODUCTION

Hitherto many image compression techniques have been developed and been used to reduce the resource requirements for transmitting data [1]-[3]. However mobile communications in wireless environments require much higher compression ratio to save network resources. One possible way to realize the higher compression ratio is to discriminate objects in an image and compress only the necessary objects for users. Scenes change time-dependently and contents of the scene also change. So it is difficult to prepare a priori information for the contents and image segmentation methods on the basis of the observed data information is required for discrimination of objects.

Image segmentation methods can be categorized into edge detection base and pixel classification base [4]. A disadvantage of the former is that unclosed regions may appear, whereas a disadvantage of the latter is that isolated classification may appear. The unclosed regions sometimes become crucial for object detection. On the other hand the isolated classification can be removed by some technique of filtering.

When a classification method uses sample image data which are representative of each class, it is called supervised, while a classification method without using sample data is called unsupervised. We shall investigate the unsupervised classification problem, that is also called clustering problem, from the viewpoint of pixel classification base. One of well-known clustering techniques is the k-means algorithm [6], whose problem is the random initial selection. For example, two segmentation results of 248x248 color image (Fig. 1) are shown in Figs. 2 and 3. Segmentation was carried out by the same k-means algorithm with different initial estimates assuming that the number of clusters is 6. The difference of cluster assignments between Figs. 2 and 3 is shown in Fig. 4, which shows that about 26.6% of total number of pixels were assigned to different clusters.

Figure 1. Original "Peppers" color image.

In this paper a color image segmentation method is proposed on the basis of the maximum likelihood (ML) estimation for the clustering problem. The difficulty lies in how to estimate the parameters of the likelihood functions and the number of segmentation. The expectation and maximization (EM) algorithm is introduced to improve the parameter estimation. The initial estimates for the EM algorithm are given by the histogram-based initial estimation, which can avoid the problem of random initial selection, and the minimum distance clustering method. The image segmentation is performed by the conventional ML method.

The remainder of the paper is organized as follows. Section II describes the EM algorithm briefly and the parameter estimation of multi-variate normal densities by the EM algorithm. In Section III an unsupervised segmentation algorithm for color images is proposed on the basis of the ML estimation. Section IV presents experimental results to show the performance of the proposed algorithm, and Section VI concludes by summarizing the paper.

II. THE EM ALGORITHM

The EM algorithm is an iterative algorithm for calculating the maximum-likelihood or maximum-a-posteriori estimates when the observations can be viewed as incomplete data. Each iteration of the algorithm consists of an expectation step followed by a maximization step. Details can be found in [5].

Figure 2. Segmentation result by the k-means algorithm.

Figure 3. Segmentation result by the k-means algorithm.

Figure 4. Difference between Fig.2 and Fig. 3.

The color image segmentation problem is formalized in the framework of the EM algorithm. $\boldsymbol{y}_j = (y_j^R, y_j^G, y_j^B)$ (j=1,..., N) is the j-th element in a color image observed in the RGB color space, where y_j^X is a scalar value observed on the X plane. $\boldsymbol{y}_j$ is supposed to be drawn from a mixture of multi-variate normal density functions whose parameters (mean vectors and covariance matrices) are unknown. Here, the number of densities is assumed to be M and the form of a mixture of multi-variate normal density functions is

$$p(\mathbf{y_j}) = \sum_{i=1}^{M} \alpha_i p_i(\mathbf{y}_j; \mu_i, \Sigma_i), \tag{1}$$

where $p_i(y_j; \mu_i, \Sigma_i)$ denotes the i-th multi-variate normal density with a mean vector μ_i and a covariance matrix Σ_i. α_i is the proportion of the $_i$-th normal density in the mixture and $\Sigma_{i=1}^{M} \alpha_i = 1$.

Let $\boldsymbol{y}$=($\boldsymbol{y}_1$,..., $\boldsymbol{y}_N$) denote the incomplete-data set consisting of N independent and identically distributed (i.i.d.) observations which is generated according to (1). We associate with $\boldsymbol{y}_j$ an unobservable M-dimensional indicator vector $\boldsymbol{x}_j=(x_{j1},...,x_{jM})$, whose entries are all zero except for the k-th entry, that is, $\boldsymbol{y}_j$ has actually been generated by the k-th normal density of the mixture. Thus, $\boldsymbol{z}=(\boldsymbol{x}_j, \boldsymbol{y}_j)$ (j=1,..., N) denote the complete-data set. Let $\phi=(\alpha_1,..., \alpha_{M-1}, \mu_1,..., \mu_M, \Sigma_1,..., \Sigma_M)$ be the parameter vector to be estimated. Assuming that $\phi^{(p)}$, the p-th estimated value of phai, has been obtained, $\phi^{(p+1)}$ is given by (2)-(5) under the EM algorithm framework.

$$\mu_j^{(p+1)} = \frac{\sum_{i=1}^{N} a_{ij}^{(p)} \mathbf{y}_i}{\sum_{i=1}^{N} a_{ij}^{(p)}} \tag{2}$$

$$\Sigma_j^{(p+1)} = \frac{\sum_{i=1}^{N} a_{ij}^{(p)} \left(\mathbf{y}_i - \mu_j^{(p)}\right)^T \left(\mathbf{y}_i - \mu_j^{(p)}\right)}{\sum_{i=1}^{N} a_{ij}^{(p)}} \tag{3}$$

for j=1,..., M. And,

$$\alpha_j^{(p+1)} = \frac{1}{N} \sum_{i=1}^{N} a_{ij}^{(p)} \tag{4}$$

for j=1,..., M-1, where

$$a_{ij}^{(p)} = \frac{\alpha_j p\left(\mathbf{y}_i; \mu_j^{(p)}, \Sigma_j^{(p)}\right)}{\sum_{j=1}^{M} \alpha_j p\left(\mathbf{y}_i; \mu_j^{(p)}, \Sigma_j^{(p)}\right)} \tag{5}$$

III. THE SEGMENTATION ALGORITHM

The proposed clustering algorithm is depicted in Fig. 5.

In STEP 1, initial estimates $\phi^{(0)}$ are calculated using a multi-dimensional histogram and the minimum distance clustering method. The multi-dimensional histogram is constructed by dividing the RGB color space into intervals and counting the number of elements of the observed color image in each RGB subspace, which is called a bin. Then the M highest-density bins are selected and the averages of xobserved elements belonging to the bins are calculated. The average is also called the centroid. If M centroids are not obtained because of narrow intervals, the histogram is rebuilt with wider intervals and the centroids are recalculated. The minimum distance clustering uses the criterion (6) to label all of the observed elements. $\phi^{(0)}$ can be calculated using the labels by the minimum distance clustering.

$$\text{Label of the } j\text{-th elements } = \min_i \left\| \mathbf{C}_i - \mathbf{y}_j \right\|, \qquad (6)$$

where $\boldsymbol{C}_i$ denotes the i-th centroid and means a distance measure. The Euclid distance is used in this paper.

In STEP 2, the EM algorithm is iteratively carried out with the initial estimates $\phi^{(0)}$ and the update equations (2)-(5). The EM algorithm converges when difference of old estimates and new estimates are less than some threshold thetaconv and the final estimates ϕ^{EM} is obtained. The EM algorithm contributes to the segmentation algorithm by way of improving the parameters of the mixture of densities on the basis of the ML criterion.

Finally, in STEP 3, the image segmentation is carried out by the conventional ML method using ϕ^{EM}. The j-th element is labeled L_j according to (7).

$$L_j = \max_i \frac{\exp\left\{-\frac{1}{2}\left(\mathbf{y}_j - \mu_i^{EM}\right)\cdot\left(\Sigma_i^{EM}\right)^{-1}\cdot\left(\mathbf{y}_j - \mu_i^{EM}\right)^T\right\}}{\left|\Sigma_i^{EM}\right|^{-\frac{1}{2}}} \qquad (7)$$

In (7) the constant factor is omitted and T denotes the transpose of a matrix.

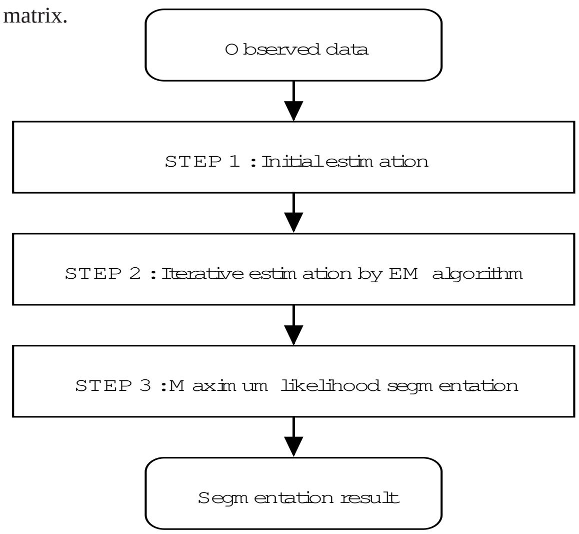

Figure 5. Segmentation algorithm.

The details of the second advantage is as follows. The number of densities, that is the number of segmentation, M must be given by the programmer at the beginning of the segmentation algorithm, but it is difficult to expect how many segmentations are included in the observed image because of the existence of noises or negligible small segmentations. One possible solution is to set M a little larger than the expected value and the negligible components of densities are weeded out when their parameter α_i's converge to 0 during the EM algorithm iterations. This selection of components depends on the initial estimates for the EM algorithm, however, and we just confirmed the selection experimentally.

IV. EXPERIMENTS

The proposed segmentation algorithm was applied for segmentation of a color image. In the first experiment the "Peppers" color image shown in Fig. 1 is used. The number of densities M was assumed to be 6. The EM algorithm converged after 276 iterations and Fig. 6 shows the final ML segmentation result followed by the 3x3 mode filtering operation in [7]. The final estimates of the means and the proportions are shown in Table 1.

In the second experiment the "Hair band" color image shown in Fig.7 is used. The number of densities M was assumed to be 4 in this case. The EM algorithm converged after 64 iterations and Fig. 8 shows the final ML segmentation result followed by the 3x3 mode filtering operation. The final estimates of the means and the proportions are shown in Table 2.

Figure 6. Segmentation result of the "Peppers" image.

Table 1. Final parameter estimates of the means and the proportions for the "Peppers" image.

		Cluster 1	Cluster 2	Cluster 3
R		195.4	159.0	150.4
G		52.1	205.5	34.7
B		40.1	147.7	34.6
Proportion		0.158	0.183	0.085

		Cluster 4	Cluster 5	Cluster 6
R		114.1	209.7	97.7
G		165.2	131.1	79.2
B		83.8	109.7	39.3
Proportion		0.317	0.044	0.213

Figure 7. Original “Hair band” color image.

Figure 8. Segmentation result of the “Hair band” image.

Table 2. Final parameter estimates of the means and the proportions for the “Hair band” image.

		Cluster 1	Cluster 2	Cluster 3	Cluster 4
R		29.4	90.9	46.1	81.7
G		36.5	71.8	39.5	90.5
B		38.9	62.1	39.7	109.4
Proportion		0.390	0.149	0.269	0.192

V. CONCLUSION

A segmentation algorithm for color images was proposed. In the algorithm the observed color image is considered as a mixture of multi-variate normal densities and segmentation is carried out by the ML estimation. The initial parameters for the mixture are obtained by the multi-dimensional histogram and the minimum distance clustering methods. Then the EM algorithm is introduced to improve the initial estimates and possibly reduce the unlikelysegmentation.

One advantage of the proposed algorithm is that it gives a stable solution for the color image segmentation problem, while the results of the segmentation methods that use a random initialization, e.g. the k-means algorithm, may differ according to the selection of initial parameters. This is more striking in the real images such as Fig. 1 than the artificial images, because of the existence of neutral tints and noises.

Although the number of densities, that is the number of segmentation, M is assumed to be known in this paper, it may be estimated during the iterations of the EM algorithm. The details are as follows. At the begging of the proposed algorithm, M is set a little larger than the expected value. Then the negligible components of densities are weeded out when their parameter α_i's converge to 0 during the EM algorithm iterations. When the EM algorithm converges, M also converges to a certain value. Actually when we set M=8 for the “Peppers” image segmentation experiment, reduction of M occurred and M converged to 6 finally. This is an experimental result and it is needed to analyze the mechanism of reduction as the further study.

VI. ACKNOWLEDGMENTS

The author would like to thank Dr. Bokuji Komiyama, the president of ATR adaptive communications research laboratories, and Dr. Jun Matsuda, head of Department 1, for their encouragement.

REFERENCES

[1] Draft ITU-T Recommendation H.263, “Video Coding for Low Bitrate Communication,” Dec. 1995.

[2] ISO/IEC JTC1/SC29/WG11 N1582, “MPEG-4 Video Verification Model Version 6.0,” Feb. 1997.

[3] ISO/IEC JTC1/SC2/WG10 CD 10918, “Digital Compression and Coding of Continuous-tone Still Images,” Jan. 1991.

[4] M. P. Ekstrom, editor, *Digital Image Processing Techniques*, New York: Academic Press, 1984.

[5] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum Likelihood from Incomplete Data via the EM Algorithm,” *J. Roy. Statistics Soc*., vol. 39, no. 1, pp. 1-38, 1977.

[6] G. A. Carpenterer and S. Grossberg, *Pattern Recognition by Self-Organizing Neural Networks*, Cambridge, MA: MIT Press, 1991.

[7] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, vol. 1, second edition, New York: Academic Press, 1982.