INEXHAUSTIVE REGION SEGMENTATION BY ROBUST CLUSTERING

Naoyuki Ichimura

Electrotechnical Laboratory
1-1-4 Umezono, Tsukuba, Ibaraki, 305 JAPAN
e-mail:ichimura@etl.go.jp

ABSTRACT

An *inexhaustive* region segmentation method using a novel *robust clustering* algorithm is proposed in the present paper. The term 'inexhaustive' means that this method segments *only* homogeneous and major regions in the image. Therefore, the pure features of the major regions that are the important clues in a recognition process can be obtained.

The finite mixture model is used to represent the distribution of the features. The region segmentation is formulated as parameter estimation of the model. The robust clustering algorithm is used in the estimation. The number of major regions is estimated from changes of the number of outliers as a function of the number of components. Experimental results for the real images are shown.

1. INTRODUCTION

The problem of region segmentation includes issues varying from a low-level to a high-level process in image analysis[1]. An ideal region segmentation for the high-level process is the one that can give the pure features of the regions as low-level cues[4]. In particular, the major regions in the image should be extracted because they are important clues in the high-level process. However, the contamination of their features sometimes occurs. This is caused by merging of the heterogeneous pixels along the boundaries and the inhomogeneous portions of details. Hence, a method that does not segment such pixels and portions is needed. The concept of such segmentation, i.e. *inexhaustive region segmentation*, was already discussed in [4]. However, the method that realizes the segmentation has not been explored.

In the present paper, the inexhaustive region segmentation method using a novel robust clustering algorithm is proposed. The clustering approach to region segmentation [2] [3] is suitable rather than the region growing to segment the major structures. Only major regions can be segmented by extracting the clusters in the feature space while rejecting the outliers corresponding to the inhomogeneous portions. The experimental results for the real images show that only major regions can be segmented by the proposed method without *a priori* knowledge about the scene.

2. THE FINITE MIXTURE MODEL

Let $V = \{\boldsymbol{x}_i\}_{i=1}^{N} \subseteq \boldsymbol{R}^n$ be a set of N feature vectors obtained from the image. The following *finite mixture model* is used to represent the distribution of V.

$$f(\boldsymbol{x}\,|\boldsymbol{\Theta}) = \sum_{i=1}^{r} \lambda_i \alpha_i(\boldsymbol{x}\,|\boldsymbol{\theta}_i) \tag{1}$$

where r is the number of component density $\alpha_i\,()$, $\boldsymbol{\Theta} = \{\{\lambda_i\}_{i=1}^{r}\ \{\boldsymbol{\theta}_i\}_{i=1}^{r}\}$ is a set of parameters, $\{\lambda_i\}_{i=1}^{r}$ are the mixing proportions such that $\sum_{i=1}^{r}\lambda_i = 1, 0 \leq \lambda_i \leq 1$, and $\boldsymbol{\theta}_i$ is a set of parameters of $\alpha_i\,()$.

The following multivariate normal density function is used as the density of the i-th component C_i.

$$\alpha_i(\boldsymbol{x}\,|\boldsymbol{\theta}_i) = (2\pi)^{-n/2}\,|\boldsymbol{V}_i|^{-1/2} \cdot \exp\left\{-1/2 \cdot (\boldsymbol{x}-\boldsymbol{m}_i)^t\,\boldsymbol{V}_i^{-1}\,(\boldsymbol{x}-\boldsymbol{m}_i)\right\} \tag{2}$$

where $\boldsymbol{\theta}_i = \{\boldsymbol{m}_i, \boldsymbol{V}_i\}$ is a set of parameters, and $\boldsymbol{m}_i$ $(n \times 1)$,$\boldsymbol{V}_i$ $(n \times n)$ are the mean and the covariance matrix of C_i, respectively.

3. NOVEL ROBUST CLUSTERING ALGORITHM FOR INEXHAUSTIVE REGION SEGMENTATION

3.1. Robust Clustering Algorithm

Each component of (1) corresponds to a region, and the component distribution represents the distribution of the features in the region. Hence, region segmentation is formulated as parameter estimation of (1).

The objective function of the proposed clustering algorithm for parameter estimation is defined as follows:

$$L = \sum_{i=1}^{N}\sum_{j=1}^{r} z_{ij}\log\lambda_j\alpha_j(\boldsymbol{x}_i\,|\theta_j) - \xi\left(\sum_{j=1}^{r}\lambda_j - 1\right) \tag{3}$$

where the first term is the log-likelihood[†] of (1), the second term is the constraint of the mixing proportions, the parameter ξ is a Lagrange multiplier, and $\boldsymbol{z}_i$ is a vector such that

$$\begin{aligned} \boldsymbol{z}_i &= (z_{i1}, \ldots, z_{ir}) \quad (i = 1, \ldots, N) & (4) \\ z_{ij} &= \begin{cases} 1 & \boldsymbol{x}_i \in C_j \\ 0 & \boldsymbol{x}_i \notin C_j \end{cases} & (5) \end{aligned}$$

The EM algorithm [5] [6] [7] has been used to estimate the parameters maximizing such objective function. The proposed algorithm is also based on the EM algorithm; however, an outlier rejection process is incorporated. Only major regions can be segmented by rejecting the outliers in V. The proposed algorithm is shown as follows:

[Step 1] Determine $\boldsymbol{z}_i$ under the current approximation of the parameters $\boldsymbol{\Theta}^c = \left\{ \{\lambda_j^c\}_{j=1}^r, \{\boldsymbol{\theta}_j^c\}_{j=1}^r \right\}$ as follows:

$$z_{ij} = \begin{cases} 1, & \lambda_j^c \alpha_j \left(\boldsymbol{x}_i \middle| \boldsymbol{\theta}_j^c \right) > \lambda_k^c \alpha_k \left(\boldsymbol{x}_i \middle| \boldsymbol{\theta}_k^c \right) \\ & j, k = 1, \ldots, r, j \neq k \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

This step divides V into the components $\{C_j\}_{j=1}^r$. The component number assigned to $\boldsymbol{x}_i$ is denoted by $C(\boldsymbol{z}_i)$, and the number of feature vectors in C_j is denoted by N_j.

[Step 2] Calculate the squared Mahalanobis distances.

$$\begin{aligned} d_i^2 &= \left(\boldsymbol{x}_i - \boldsymbol{m}_{C(\boldsymbol{z}_i)}^c \right)^t \boldsymbol{V}_{C(\boldsymbol{z}_i)}^c{}^{-1} \left(\boldsymbol{x}_i - \boldsymbol{m}_{C(\boldsymbol{z}_i)}^c \right) \\ i &= 1, \ldots, N & (7) \end{aligned}$$

[Step 3] In each component C_j, extract the $100\alpha N_j$ $(0 < \alpha \leq 1.0)$ feature vectors in accordance with ascending order of the d_i^2[‡]. The subset of the extracted feature vectors in C_j is denoted by $J_{j,\alpha}$.

[Step 4] Calculate the following parameters.

$$\begin{aligned} \boldsymbol{m}_{j,\alpha} &= \boldsymbol{m}_{J_{j,\alpha}} & (8) \\ \boldsymbol{V}_{j,\alpha} &= \left(d_{j,\alpha}^2 / \chi_{n,\alpha}^2 \right) \boldsymbol{V}_{J_{j,\alpha}}, j = 1, \ldots, r & (9) \end{aligned}$$

where $\chi_{n,\alpha}^2$ is a 100α percent point of χ^2 distribution with n degrees of freedom, $d_{j,\alpha}^2$ is the maximum squared Mahalanobis distance in $J_{j,\alpha}$, and $\boldsymbol{m}_{J_{j,\alpha}}$, $\boldsymbol{V}_{J_{j,\alpha}}$ are the mean and the covariance matrix of $J_{j,\alpha}$, respectively.

[Step 5] Recalculate d_i^2 using $\boldsymbol{m}_{C(\boldsymbol{z}_i),\alpha}$ and $\boldsymbol{V}_{C(\boldsymbol{z}_i),\alpha}$.

[Step 6] Determine the weights $\{w_i\}_{i=1}^N$ as follows:

$$w_i = \begin{cases} 1, & d_i^2 < \chi_{n,0.975}^2 \\ 0, & d_i^2 \geq \chi_{n,0.975}^2 \end{cases} \quad (10)$$

[Step 7] Calculate a set of parameters $\boldsymbol{\Theta}^+$.

$$\begin{aligned} \lambda_j^+ &= W_j / N & (11) \\ \boldsymbol{m}_j^+ &= \sum_{i=1}^N w_i z_{ij} \boldsymbol{x}_i / W_j & (12) \\ \boldsymbol{V}_j^+ &= \sum_{i=1}^N w_i z_{ij} \left(\boldsymbol{x}_i - \boldsymbol{m}_j^+ \right) \left(\boldsymbol{x}_i - \boldsymbol{m}_j^+ \right)^t / W_j & (13) \end{aligned}$$

where $W_j = \sum_{i=1}^N w_i z_{ij}, j = 1, \ldots, r$

[Step 8] If the parameters are unchanged or the number of iterations exceeds the prescribed number, then terminate this algorithm. Otherwise,update Θ^c by Θ^+ and go to step1.

This algorithm is constructed by two main parts. One includes step1 and step7 that are based on the EM algorithm, the other includes the remaining steps that reject the outliers. The squared Mahalanobis distance recalculated in step5 is given as follows:

$$\begin{aligned} d_i^2 = \left(\chi_{n,\alpha}^2 / d_{C(\boldsymbol{z}_i),\alpha}^2 \right) \cdot & \quad (14) \\ \left(\boldsymbol{x}_i - \boldsymbol{m}_{J_{C(\boldsymbol{z}_i),\alpha}} \right)^t \boldsymbol{V}_{J_{C(\boldsymbol{z}_i),\alpha}}{}^{-1} \left(\boldsymbol{x}_i - \boldsymbol{m}_{J_{C(\boldsymbol{z}_i),\alpha}} \right) \end{aligned}$$

This equation shows that the algorithm rejects the outliers by following two stages:(i) calculate the squared Mahalanobis distances using the parameters obtained from $J_{j,\alpha}$,(ii) modify these distances based on the difference between the 100α percent point of χ^2 distribution obtained from the model and that obtained from $J_{j,\alpha}$.

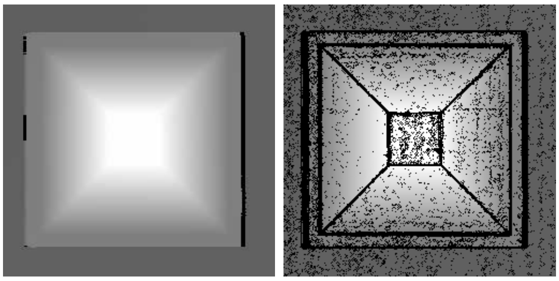
3.2. Comments on Some Parameters

First, we discuss the parameter α. The explanation of step3 and equation (14) show that the parameter α controls the effect of the feature vectors in calculating the distances used in (10). That is, if α is less than 0.5, the effect of the assumption of the component distribution is dominant in calculating (14), while if α is greater than 0.5, the effect of the feature vectors is dominant.

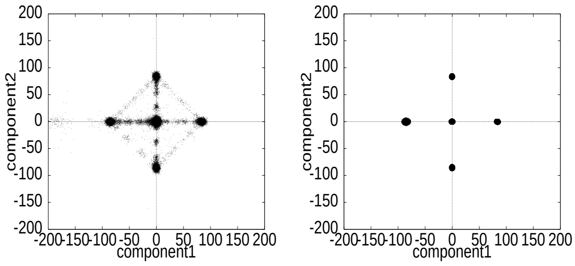
Next, we discuss the number of components r. In the robust clustering, if the number of components is greater than the number of major regions, the surplus components tend to divide the major regions rather than to represent the outliers, since the algorithm rejects them. As a result, if we observe changes of the number of outliers as a function of the number of components, saturation may occur. The number of components r is determined by the saturated number.

[†] The method using this log-likelihood is called the classification likelihood approach[7].

[‡] This extraction method is called the multivariate trimming [8] [9].

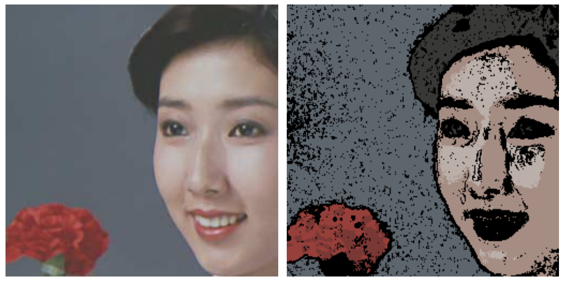


(a) (b)

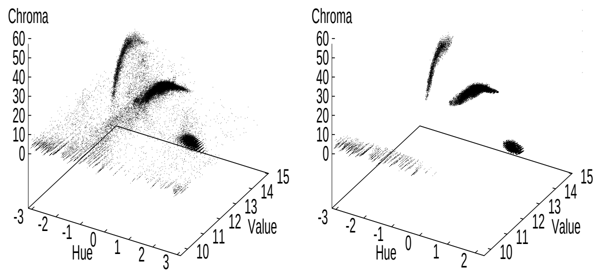


(c) (d)

Figure 1: Experimental results for the range image. (a)original image (256×256). (b)inexhaustive region segmentation result. (c)feature space of (a). (d)feature space of (b). In figure (b), the feature vectors of the black regions were treated as outliers.



(a) (b)



(c) (d)

Figure 2: Experimental results for the face image. (a)original color image (256×256). (b)inexhaustive region segmentation result. (c)feature space of (a). (d)feature space of (b).

4. EXPERIMENTS

The experimental results for the three real images are shown. In all the experiments, the parameter α was 0.5.

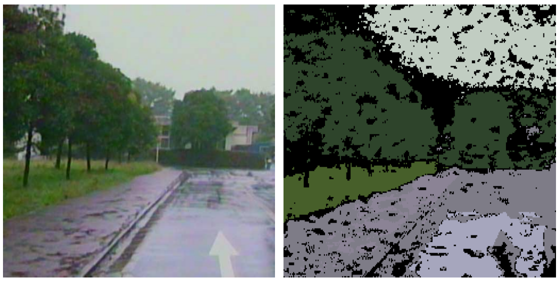
In the range image shown in Fig.1(a), the surface normal vector at each pixel was used as a feature. The number of components was estimated at five (Fig.4 (a)). This is equal to the number of planes that have different directions. Figure 1(c) and (d) show that only the clusters corresponding to the major regions were extracted. Thus the pixels that might be contaminated by the noises and the heterogeneous pixels along the boundaries were rejected (Fig.1(b)).

In the color image of the woman's face shown in Fig.2(a), the Munsell color space (H,V,C)[11] was used as a feature. The number of major regions was estimated at six (Fig.4 (b)). Figure 2 (b) shows that each of the segmented regions can be regarded as the major region. The pure features of the regions were obtained (Fig.2(c) and (d)). Owing to their pureness, they are useful in matching with the stored prototype to label the regions. The inhomogeneous portions, e.g. the eyes and the mouth, were rejected(Fig.2(b)).

In the color image of the outdoor rainy scene shown in Fig. 3 (a), the Munsell color space and the coordinates of each pixel were used as the features. The number of major regions was estimated at six (Fig. 4 (c)). Only major regions were segmented, and the following portions were rejected: the buildings, the small puddles on the road, the traffic sign, and the boundaries between the sky and the trees(Fig.3 (b)).

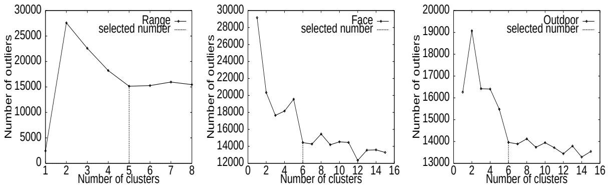
The estimated component densities for the major regions can be used with the Markov random field(MRF) model [12] to compose the maximum a posteriori(MAP) estimation. Figure 5 shows the MAP segmentation results for Fig.2 (a) and Fig.3 (a). Using the MRF model as a prior distribution, we can take account of the spatial relation. The smooth regions, therefore, could be obtained. Nevertheless, the inhomogeneous portions corresponding to the extreme outliers were preserved.

To segment the inhomogeneous portions detected by this method, the local spatial analysis should be used. In the present work, the region growing that is suitable for local analysis is used. Figure 6 shows that the images were exhaustively segmented in accordance with the complexity of each portion. The choice of the threshold used in the growing had little influence for the global segmentation result, because the growing was carried out only in the restricted area, i.e. detected inhomogeneous portions.



(a) (b)

Figure 3: Experimental results for the outdoor rainy scene image. (a)original color image (512×400). (b)inexhaustive region segmentation result.



(a) (b) (c)

Figure 4: Changes of the number of outliers as a function of the number of components. (a)Range. (b)Face. (c)Outdoor.

5. CONCLUSIONS

The inexhaustive region segmentation method using the novel robust clustering algorithm was proposed. The experimental results showed that only major regions were segmented without *a priori* knowledge about the scene. The number of major regions could be estimated from changes of the number of outliers. The estimated component densities were used with the MRF model in the MAP segmentation. The images were exhaustively segmented in accordance with the complexity of each portion using the proposed method and the restricted region growing.

(a) (b)

Figure 5: MAP segmentation results. (a)Face. (b)Outdoor.

(a) (b)

Figure 6: Segmentation result by combining the restricted region growing. (a)Face. (b)Outdoor.

6. REFERENCES

[1] T. Kanade: "Region Segmentation: Signal vs Semantics", CGIP, Vol.13, pp.279-297 (1980)

[2] K.S. Fu and T.K. Mui: "A Survey on Image Segmentation", Pattern Recognition, Vol.13, pp.3-16 (1981)

[3] R.M. Haralick and L.G. Shapiro: "Image Segmentation Techniques", CVGIP, Vol.29, pp.100-132 (1985)

[4] T. Fujimori and T. Kanade: "An Approach to Knowledge-Based Interpretation of Outdoor Natural Color Road Scenes", in C.E. Thorpe ed. "Vision and Navigation The Carnegie Mellon Navlab", Kluwer Academic Publishers, pp.39-81 (1988)

[5] A.P. Dempster, N.M. Laird and D.B. Rubin: "Maximum Likelihood from Incomplete Data via the EM Algorithm", J.R.Statist.Soc. B, Vol.39, No.1, pp.1-38 (1977)

[6] R.A. Redner and H.F. Walker: "Mixture Densities,Maximum Likelihood and the EM Algorithm", SIAM Review, Vol.26, No.2, pp.195-239 (1984)

[7] G.J. McLachlan and K.E. Basford: "Mixture models", MARCEL DEKKER,INC. (1988)

[8] J.D. Jobson: "Applied Multivariate Data Analysis Volume II : Categorical and Multivariate Methods", Springer-Verlag (1992)

[9] P.J. Rousseeuw and B.C. Zomeren: "Unmasking Multivariate Outliers and Leverage Points", Journal of the American Statistical Association , Vol.85 , No.411 , pp.633-639 (1990)

[10] J.M. Jolion, P. Merr and S. Bataouche: "Robust Clustering with Applications in Computer Vision", IEEE Trans. PAMI , Vol.13 , No.8 , pp.791-802 (1991)

[11] M. Miyahara and Y. Yoshida:"Mathematical Transform of (R,G,B) Color Data to Munsell(H,V,C) Color Data", The journal of the ITE of Japan, Vol.43, No.10 , pp.1129-1136 (1989) (in Japanese)

[12] R. Kindermann and J.L. Snell: "Markov Random Fields and their Applications" , American Mathematical Society (1980)

[13] J. Besag: "On the Statistical Analysis of Dirty Pictures" , J.R.Statis.Soc. B, Vol.48, No.3, pp.259-302 (1986)