AGGLOMERATIVE CLUSTERING OF FEATURE DATA FOR IMAGE SEGMENTATION

D.P. Mukherjee[1,2], Partha Pratim Mohanta[2] and Scott T. Acton[1]
[1] Virginia Image and Video Analysis, Electrical and Computer Engineering, University of Virginia
[2] Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta

ABSTRACT

We propose an image segmentation model using an agglomerative clustering technique. The clustering is performed within a feature matrix where intensity and boundary relations are defined between neighboring segments. The iterative segment agglomeration process satisfies a cluster aggregation property and does not utilize *a priori* knowledge of the number of clusters present in the image. The performance of the algorithm is demonstrated on a number of color images and compared with a similar algorithm.

1. INTRODUCTION

Segmentation of digital imagery is a vital first step for automated image analysis applications such as content based image retrieval. We are proposing an image segmentation model utilizing the concept of *agglomerative clustering*. In our clustering process, we assume no *a priori* knowledge of the final number of clusters present in the image. The existing image segmentation techniques exploit both perceptual and statistical models. A promising approach discussed in [3] applies anisotropic diffusion in the chromatic and lightness channels of a color image followed by k-means clustering. While anisotropic diffusion may compromise edge localization, the approach that we outline in this paper does not distort the original edge map. Also, the k-means clustering method requires specification of number of clusters in the image, which is found automatically in the paradigm presented here.

The segmentation process specified in [2] consists of two distinct steps – the first being color quantization and conversion of pixel labels into quantized class labels. This is followed by clustering based on spatial proximity of class maps. In contrast, we propose a unified approach of agglomerative clustering in which clustering is not performed on the image itself but on a feature matrix derived based on the homogeneity relations between neighboring segments.

Our agglomerative clustering model is inspired by a model used to cluster psychological observation data [5]. The model was based on the definition of similarity between features. Given the number of classes and assuming that each observed datum maintains strict membership in any particular class, the clustering model in [5] minimizes the error between theoretical and empirical estimates of the similarity between classes.

Instead of minimizing the error between theoretical and empirical estimates of similarity, we agglomerate image segments by maximizing an aggregation index. The index is derived from a feature matrix wherein relations between segments are specified. The segmentation model is described in Section 2, and the agglomeration technique is presented in Section 3. The construction of feature matrix is given in Section 3.1

The model in [5] requires knowledge of possible number of clusters. One could start with each individual pixel as a potential class and then initiate agglomeration from that stage, but such an approach is computationally expensive. We have used an area morphological technique [1] to initiate agglomeration process, as discussed in Section 3.2. The result obtained after agglomeration is shown and compared to another method in Section 4.

2. SEGMENTATION MODEL

The major premise on which this segmentation model is built is that the data set can be aggregated. This is logical given the fact that the purpose of the image segmentation is to find homogeneous image segments representing meaningful components. We now define the "aggregability" property.

Let I be a set of image segments. For n iso-intensity segments in I, we can define an $n \times n$ feature matrix p describing the relation between the segments. For every pair of spatially neighboring segments $i \in I$ and $j \in I$ a relation p_{ij} is calculated. This relation defines how close these segments are in terms of agglomeration into a cluster. Calculation of p_{ij} is explained in Section 3.1. We can consider p_{ij} a measure of affinity between segments i and j. Let us define $p(j/i)$ using $p(j/i) = p_{ij} / p_{i+}$. This ratio can be interpreted as affinity of segment j to i versus the total affinity of segment i to its entire neighborhood. Let us also define that $p(j) = p_{+j} / \Sigma p_{ij}$ is the share of j towards its neighbor with respect to total transactions between segments in I. Therefore the ratio $p(j/i)/p(j)$ specifies the affinity of j to i compared to its overall contribution in p. We define

$$\mathrm{M}_{ij} = p(j/i)/p(j) - 1, \qquad (1)$$

as the index of agglomeration. For $\mathrm{M}_{ij} > 0$, $p(j/i)$ is greater than $p(j)$ and hence agglomeration is favored as contribution of j towards i is greater than its overall average. In the case where $\mathrm{M}_{ij} = 0$, there is no difference of status so far as agglomeration of segments i and j is concerned while for $\mathrm{M}_{ij} < 0$, agglomeration is not desirable. The index of agglomeration is motivated by the data scatter measure in a relational matrix, such as the Pearson chi-square index [4]. In the next section, we implement the clustering following the index of agglomeration.

3. AGGLOMERATION

Let s_{it} denote an indicator function that is unity-valued when segment i belongs to cluster t. For m possible clusters at any point of agglomeration, the agglomerative clustering model is given by:

$$p_{ij} = \sum_{t=1}^{m} a_t s_{it} s_{jt} + e_{ij}. \qquad (2)$$

The average feature value in the cluster t to which segments i and j may possibly be merged is termed as a_t. The error incurred in the proposed model is e_{ij}. Assuming an orthogonal relationship between elements of equation (2), the above relationship can be extended for all the segments present in the image:

$$\sum_{i,j \in I} p_{ij}^2 = \sum_{t=1}^{m} a_t^2 \sum_{i,j \in I} s_{it} s_{jt} + \sum_{i,j \in I} e_{ij}^2. \qquad (3)$$

By definition of s_{it} and given that a segment can belong to only one cluster, we have $\Sigma_{i,j \in I} s_{it} s_{jt} = |S_t||S_t - 1|$. The cardinality of non-zero values for indicator function s_{it} is given by $|S_t|$. Note that the diagonal similarities are not included in feature matrix p. So, in order to minimize the error in (3), given that p_{ij} is constant for a specific topology of the segments in the image space, the following expression should be maximized:

$$\omega(I) = \sum_{t=1}^{m} a_t^2 |S_t||S_t - 1|. \qquad (4)$$

For a pair of clusters i and j, if $d = \omega(I(i,j)) - \omega(I)$ is positive, the pair is selected as a possible candidate for agglomeration. The terms $\omega(I(i,j))$ denotes the value of ω as defined in (4) after agglomerating segments i and j. All segment pairs are ordered in descending values of d. In that list, the first segment pair that satisfies $\mathrm{M}_{ij} > 0$ is merged. The merged cluster intensity is reassigned as the weighted mean intensities of the clusters those are merged. The feature matrix p is then recalculated according to the topology of merged clusters. The process is repeated till constraints as in the equations (4) and (1) is valid for cluster merging. The accuracy of the process depends on evaluation of p. In the next section we describe the evaluation of p.

3.1. Feature Matrix

As defined in Section 2, the feature matrix p estimates a measure of similarity between segments by which agglomeration is performed. We have considered two features to measure the similarity between segments. The first one is the similarity in intensity/color. The affinity between two segments increases if the two neighboring segments are similar in intensity/color. The second factor that influences agglomeration in the proposed scheme is the extent of boundary overlap between two neighboring segments. We assert that the segment that has larger common boundary with another segment should have an increased priority in agglomeration. The boundary feature introduces an asymmetric similarity value in the relational matrix. A large segment with relatively smaller common boundary to a neighboring small segment has less affinity toward the smaller segment than the smaller segment has to the larger. Where $p_{ij} \neq p_{ji}$ we have taken $\min(p_{ij}, p_{ji})$ as the affinity between segments i and j.

The p matrix is initialized using

$$p_{ij} = \exp(-\Delta I) + \kappa(\exp(\Delta B) - 1), \qquad (5)$$

where, ΔI is the absolute ($[0,1]$ normalized) intensity difference and ΔB is the ratio of the length of the common boundary with respect to the total boundary length of the segment. The boundary of the segment is defined taking 8-connectivity. We have used exponential function to evaluate p_{ij}. We are currently investigating other non-linear functions to improve the evaluation of p_{ij} so that agglomeration could be achieved in fewer numbers of iterations. The variable κ is the regularization parameter and its value could be assessed from the range of ΔI and ΔB of equation (5).

3.2 Initialization

Agglomerating within the feature matrix is computationally less expensive as compared to clustering within the actual image. Nevertheless the complexity depends on the number of segments present initially in the image space. So, a need exists for an initialization process by which comparatively fewer numbers of clusters are present in the image to start agglomeration and also the formation of these initial sets of clusters should abide by the principles of agglomeration. We utilize an area morphology based preprocessing technique that would satisfy these preconditions [1].

Consider the thresholded decomposition of the image $I \subset Z^2$, where $I(x,y)$ is the pixel intensity at spatial location (x,y). Then $L(x,y)=1$ if $I(x,y) \geq t$, 0 otherwise. The parameter t ranging from 1 to 255 generates 255 binary images for 8-bit intensity data. For each such binary image we consider 1-pixels as foreground connected components and 0-value pixels as background connected components. Stacking of these level sets reconstructs the original image [1].

During area open operation, foreground connected component of size less than a pre-specified minimum area is replaced by background pixels. Identical operation is carried out in the area close operation except that connected component of background pixels is replaced by foreground ones. Concatenation of both of these operations is defined as area open close operation and is defined as $I \,\hat{\circ}\, s \,\hat{\bullet}\, s$ where s is the minimum length of the connected component defined by the user. The area open close gives equal weight to contrast treating equally both increasing and decreasing intensity levels.

The area open close (AOC) operation removes segments of size less than a pre-specified limit. Next we show that this operation actually aggregates pixels towards homogeneous components.

Let us consider two connected components C_1 and C_2 at $I(x,y) \geq t_1$ and $I(x,y) \geq t_2$ respectively, such that $t_2 > t_1$. If C_1 contains C_2 at t_1 and given that $|C_2|$ is less than a pre-specified area s, the connected component C_2 is replaced by background pixels. So, the segment containing C_1 is aggregated to a more homogeneous region. However, if C_2 survives in spite of area limitation and t_1, t_2 are close in intensity scale, further agglomeration of segments is required. The result of the area morphology followed by agglomerative clustering (AC) is shown next.

4. RESULTS

We have experimented with a number of color images, three of which are reported in this paper. Fig. 1(a) is an HSV color image with V-component image shown in Fig. 1(b). The area open closed image is shown in Fig. 1(c). The area limit for area open close operation is given in Table 1. Fig. 1(d) is the result of segmentation after 9 iterations while Fig. 1(e) is the result obtained using algorithm in [2]. For Fig. 1(e) and for all the subsequent results generated using [2] we have used the default scale and merging parameters. Color quantization thresholds are also set as recommended in [2]. The result generated using agglomerative clustering appears more logical in a semantic sense, especially for two additional segments in Fig. 1(e) in the left-hand side of the segmented image

Fig. 2(a) shows an HSV image of clouds with the V-component image in Fig. 2(b), area open closed image in Fig. 2(c) and segmented image in Fig. 2(d). The result produced using [2] is shown in Fig. 2(e). Similarly for the original image of Fig. 3(a), the segmented image, executed on the processed V-component image in Fig. 3(b), is shown in Fig. 3(c). The result for comparison is shown in Fig. 3(d). The results obtained clearly show reduced number of image segments without compromise of the intricate edge details. In fact, for Fig. 3(c), the building contours are clearly indicated after the segmentation process.

Table 1 presents the effectiveness of the initialization showing the reduction of number of components first after initialization using area open close followed by agglomeration.

The data scatter of the feature matrix is evaluated over the iterations following the principles of agglomeration as presented in Section 2. The minimization of Pearson-chi square index [4] during agglomeration as shown in Fig. 4 justifies the agglomeration.

5. CONCLUSION

We have proposed agglomerative clustering based image segmentation tool and we have shown that useful results can be obtained for color images. We are currently investigating extension of the agglomeration using perceptual color spaces. Another future direction involves the use of shape information in deriving segments.

Acknowledgment

This work was supported in part by the National Science Foundation (0121596).

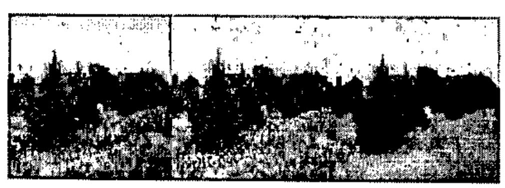

Fig. 1(a) Original Image | Fig. 1(b) Intensity image | Fig. 1(c) After AOC

Fig. 1(d) AC Result Fig. 1(e) Result from [2]

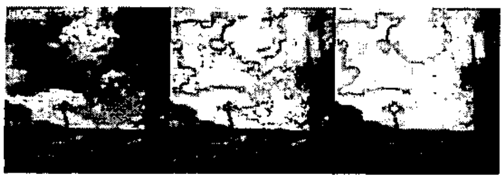

Fig. 2(a) Original Image Fig. 2(b) Intensity image Fig. 2(c) After AOC

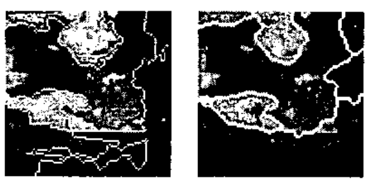

Fig. 2(d) AC Result Fig. 2(e) Result from [2]

Fig. 3(a) Original Image Fig 3(b) After AOC on intensity image

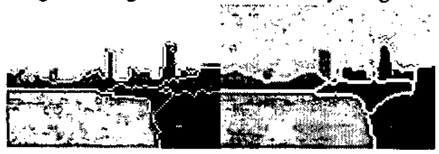

Fig. 3(c) Agglomerative Clustering Result Fig. 3(d) Result from [2]

Table 1: Number of segments in different stages of agglomeration (s is the area in the AOC operation).

Images	Original	After AOC	Final
Tree	5713	550 (s=200)	3
Cloud	6619	540 (s=100)	8
River	4713	420 (s=100)	5

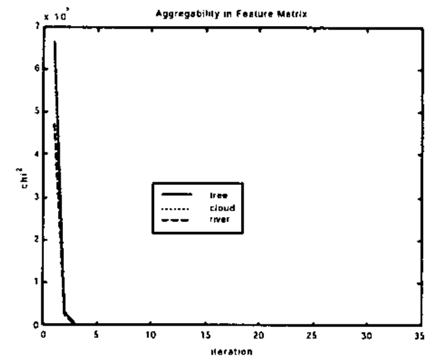

Fig. 4: Pearson-chi square index for each iteration of agglomeration.

References:

[1] S T Acton and D P Mukherjee, "Scale Space Classification using Area Morphology," *IEEE Trans. Image Proc.*, vol. 9, pp. 623-635, 2000.

[2] Y Deng, B S Manjunath and H Shin, "Color Image Segmentation," *Proc. CVPR '99*, vol. 2, Fort Collins, CO, June 1997, pp. 446-51.

[3] L. Lucchese and S K Mitra, "Color Image Segmentation through Independent Anisotropic Diffusion of Complex Chromaticity and Lightness," *Proc. ICIP*, Thessaloniki, Greece 2001.

[4] Boris Mirkin, *Mathematical Classification and Clustering*, Kluwer Academic Publishers, Boston, 1996.

[5] R N Shepard and P Arabie, "Additive Clustering: Representation of Similarities as Combinations of Discrete Overlapping Properties," *Psychological Review*, vol. 86, pp. 87-109, March 1979.