Color Image Segmentation Using Anisotropic Diffusion and Agglomerative Hierarchical Clustering

Daehee Kim[1], Yo-Sung Ho[1], and B.S. Manjunath[2]

[1] Kwangju Institute of Science and Technology
1 Oryong-dong Puk-gu, Kwangju, 500-712, Korea
{kimdh, hoyo}@kjist.ac.kr
[2] University of California
Santa Barbara, CA 93106-9560, USA
manj@ece.ucsb.edu

Abstract. A new color image segmentation scheme is presented in this paper. The proposed algorithm consists of image simplification, region labeling and color clustering. The vector-valued diffusion process is performed in the perceptually uniform LUV color space. We present a discrete 3-D diffusion model for easy implementation. The statistical characteristics of each labeled region are employed to estimate the number of total clusters and agglomerative hierarchical clustering is performed with the estimated number of clusters. Since the proposed clustering algorithm counts each region as a unit, it does not generate oversegmentation along region boundaries.

1 Introduction

Image segmentation refers to the operation of partitioning an image into separate regions, each of which is homogeneous with respect to some image features. Most image segmentation algorithms have been proposed for gray-scale images, while segmentation of color images has received much less attention from the scientific community until recent years.

Fast and accurate segmentation is necessary for authoring multimedia content and for high-level image analysis, such as object recognition, image understating, and scene interpretation. If we can achieve good object segmentation in generic images, it will enrich the current MPEG-4 and MPEG-7 standards. While MPEG-4 enables object-based functionalities, MPEG-7 is related to content-based indexing and retrieval of multimedia databases.

In this paper, we propose a new segmentation algorithm for color images that is applicable to various types of images. The proposed algorithm consists of three operations: image simplification, region labeling, and color clustering. According to Skarbek's classification [1], our proposed algorithm can be classified as a mixture of pixel-based segmentation and region-based segmentation. In this paper, we propose a discrete 3-D diffusion model with a shock filter and a new clustering method that does not require any priori known information. These subtasks are explained in detail in the following sections.

Y.-C. Chen, L.-W. Chang, and C.-T. Hsu (Eds.): PCM 2002, LNCS 2532, pp. 759-766, 2002.

2 Image Segmentation

2.1 Morphological Simplification

Since images have several objects of homogeneous regions and too specific small structures are not interesting in semantic analysis, we can initially perform structure simplification and region texture simplification. In order to simplify image structures, the morphological open-close and close-open operations by reconstruction filters are used in the RGB color space independently [2].

In this paper, we employ the morphological opening and closing operations by reconstruction filters to preserve object boundaries during image simplification. These filter pairs can remove fractional regions that are smaller than a certain size and contours of the remaining objects in the image are well-preserved [2]. The opening and closing operations by reconstruction filters can be described by partial reconstruction. Once unwanted details in the texture of the regions are removed, the image is homogenized in terms of region textures.

2.2 Anisotropic Diffusion

In order to simplify region textures, we can apply an image smoothing operation in homogeneous regions, but not across region boundaries. This requirement can be satisfied with an anisotropic diffusion filter or a scale-space filter [3]. The diffusion is carried out in the perceptually uniform LUV color space [4].

If each component in the LUV color space is diffused independently, it evolves with different smoothing directions and intensities. Therefore, the vector-valued diffusion is performed in the 3-D LUV color space by

$$\frac{\partial \mathbf{I}(x,y,t)}{\partial t} = div(\rho(x,y,t)\nabla \mathbf{I}(x,y,t)) \tag{1}$$

where $\mathbf{I}(x, y, t)$ is a vector in the LUV color space and ρ is the diffusion or conduction coefficient.

For the scalar case, the gradient of the gray level is used as an edge estimator and the conduction coefficient is defined as a decreasing function of the magnitude of gradient to prevent diffusion from affecting region boundaries. Therefore, a vector image diffusion process needs to define the local variation of $\|d\mathbf{I}\|^2$ [5].

$$\|d\mathbf{I}\|^2 = \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix}^T \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \text{ where } g_{ij} = \frac{\partial \mathbf{I}}{\partial x_i} \cdot \frac{\partial \mathbf{I}}{\partial x_j} \tag{2}$$

The two eigenvalues of $[g_{ij}]$ are the extrema of $\|d\mathbf{I}\|^2$ and the orthogonal eigenvectors η and ξ are the corresponding variation directions.

Sapiro-Ringach proposed an anisotropic vector diffusion PDE [6]:

$$\frac{\partial \mathbf{I}}{\partial t} = g(\lambda_+ - \lambda_-)\mathbf{I}_{\xi\xi} \tag{3}$$

Eq. (3) includes the second derivative of a vector in the tangential direction to the maximal variation and a decreasing function of the difference between two eigenvalues. However, since the tangential direction is continuous and vector-valued images are on the digitized grid, a bilinear or geometrical interpolation is needed; however, such an interpolation operation may introduce additional image burring or distortion. In addition, calculation of eigenvalues increases computational complexity.

The scalar-valued diffusion also requires gradient and tangential directions [7]. In order to alleviate these problems, Perona-Malik's resistor model [3] can be employed; however, this model was originally developed for scalar-valued images. In this model, each pixel is connected to four neighboring pixels with resistors whose resistances increase as the pixel differences of each connection increase.

In this paper, we extend Perona-Malik's resistor model to vector-valued images. Since the diffusion equation relates the amount of the temporal variation to the amount of the spatial variation, it is more reasonable to employ the eight-connected resistor model rather than the four-connected model. In addition, we describe resistances between each connection with the local variation $\|d\mathbf{I}\|$ in Eq. (2) for vector-valued images. The local variation can be expressed as the sum of two eigenvalues of $[g_{ij}]$, which can be calculated by

$$\|d\mathbf{I}\|^2 = \sum_{n=1}^{3} \|\nabla I_n\|^2 \tag{4}$$

where n represents each space of the LUV color space.

Therefore, the discretized 3-D diffusion equation of Eq. (1) is described by

$$\begin{aligned} \mathbf{I}_{y,x}^{t+1} = \mathbf{I}_{y,x}^{t} &+ \frac{1}{8}(r_N d\mathbf{I}_N + r_S d\mathbf{I}_S + r_E d\mathbf{I}_E + r_W d\mathbf{I}_W) \\ &+ \frac{1}{8\sqrt{2}}(r_{NW} d\mathbf{I}_{NW} + r_{SW} d\mathbf{I}_{SW} + r_{NE} d\mathbf{I}_{NE} + r_{SE} d\mathbf{I}_{SE}) \end{aligned} \tag{5}$$

where N, S, E, W are the subscripts for the north, south, east, and west directions, respectively. r_x is a admittance function of each direction, and $d\mathbf{I}_x$ represents eight neighbor differences, defined by

$$\begin{aligned} d\mathbf{I}_N &= \mathbf{I}_{y-1,x} - \mathbf{I}_{y,x} & d\mathbf{I}_S &= \mathbf{I}_{y+1,x} - \mathbf{I}_{y,x} \\ d\mathbf{I}_E &= \mathbf{I}_{y,x+1} - \mathbf{I}_{y,x} & d\mathbf{I}_W &= \mathbf{I}_{y,x-1} - \mathbf{I}_{y,x} \\ d\mathbf{I}_{NW} &= \mathbf{I}_{y-1,x-1} - \mathbf{I}_{y,x} & d\mathbf{I}_{NE} &= \mathbf{I}_{y-1,x+1} - \mathbf{I}_{y,x} \\ d\mathbf{I}_{SW} &= \mathbf{I}_{y+1,x-1} - \mathbf{I}_{y,x} & d\mathbf{I}_{SE} &= \mathbf{I}_{y+1,x+1} - \mathbf{I}_{y,x} \end{aligned} \tag{6}$$

In order to compensate for differing distances to pixel's neighbors, we scale the differences of the diagonal directions by $1/\sqrt{2}$.

Although we employ an anisotropic diffusion based on the resistor model, there are some leakages at the region boundaries. In order to address this problem, a shock filter that enhances blurred edges can be included in the entire diffusion equation. A shock filter for scalar-valued images [7] is given by

$$\frac{\partial I}{\partial t} = -\text{sign}(\nabla^2 I)|\nabla I| \tag{7}$$

For vector-valued images, a vector-valued shock filter should be also controlled by the local variation in the LUV color space because of the same reason in Eq. (1).

$$\frac{\partial \mathbf{I}_{Shock}}{\partial t} = -(1 - r(\|\nabla \mathbf{I}\|)) \begin{pmatrix} \text{sign}(\nabla^2 I_L)\|\nabla I_L\| \\ \text{sign}(\nabla^2 I_U)\|\nabla I_U\| \\ \text{sign}(\nabla^2 I_V)\|\nabla I_V\| \end{pmatrix} \tag{8}$$

where r is an admittance function.

The discretized version of Eq. (8) is implemented using the difference of each connection and the weighted sum of the differences in the same way as in Eq. (5). The final diffusion equation is given by

$$\frac{\partial \mathbf{I}(x,y,t)}{\partial t} = div(r\nabla \mathbf{I}) + \alpha \frac{\partial \mathbf{I}_{Shock}}{\partial t} \tag{9}$$

where α is inversely proportional to the number of iterations to guarantee system stability against the enhancement of a shock filter. The operation of Eq. (9) makes texture more homogeneous.

3 Region Labeling and Clustering

Some color segmentation algorithms have performed clustering in a color space directly; however, those algorithms show coarse segmentation results near object boundaries, because those algorithms use only the color histogram and the similarity measure; it does not consider the region information of objects.

In order to alleviate these problems, we employ a gradient-based watershed algorithm and obtain partitioned regions of the simplified image, resulted from Section 2. The watershed algorithm is region-growing and region-labeling algorithm that assigns a unique label to each region. Various watershed algorithms have been proposed in the literatures. In this paper, we use an immersion-based watershed algorithm because it is simple and computationally efficient [8]. The color gradient image, the input image to the watershed step, is generated by $\|d\mathbf{I}\|$.

Our goal here is to find the best representative clusters in the LUV color space and to preserve the boundaries of objects. Most clustering algorithms need a priori information, such as the number of clusters and the initial mean vectors, or a certain threshold value given by input parameters [9, 10, 11]. Furthermore, segmentation results of clustering algorithms are often dependent on the initial conditions.

In the proposed method, we estimate the number of color clusters based on the region information. The proposed clustering algorithm consists of two steps: estimation of the number of clusters, and agglomerative hierarchical clustering.

In order to estimate the number of clusters required, we assume that each vector component in each region resulted from the watershed algorithm obeys a Gaussian distribution $N(\mu_i, \sigma_i)$, where i indicates one component among the LUV color space. This assumption is valid because a vector in the each diffused region has a small Euclidian distance to the mean vector in the corresponding region. Under the hypothesis H_0 that two points are in the same region, the difference d_i between two points in the same region obeys a zero mean Gaussian distribution $N(0, \sqrt{2}\sigma_i)$.

$$p(d_i \mid H_0) = \frac{1}{\sqrt{4\pi\sigma_i^2}} \exp\left(-\frac{d_i^2}{4\sigma_i^2}\right) \tag{10}$$

Since the Euclidian distance is a good representation of the color distance in the LUV color space that is perceptually uniformly distributed, we consider the normalized following test statistic.

$$\theta = \sum_{i=1}^{3} \frac{d_i^2}{2\sigma_i^2} \tag{11}$$

where i is the index for each space of the LUV color space. In Eq. (11), the test statistic θ has a χ^2 probability density function with three degrees of freedom.

$$p(\theta \mid H_0) = \frac{1}{\sqrt{2\pi}} \theta^{\frac{1}{2}} \exp\left(-\frac{\theta}{2}\right) \tag{12}$$

With the known distribution $p(\theta|H_0)$, the decision whether or not two points are in the same region can be made by a significance test [12]. For this purpose, we specify a false alarm rate α.

$$\alpha = \Pr(\theta > \theta_{th} \mid H_0) = 0.1 \tag{13}$$

In Eq. (13), if the interval $[0, \theta_{th}]$ is a $100(1-\alpha)$ confidence interval for the parameter θ, the hypothesis H_0 is rejected if and only if θ is not in the interval $[0, \theta_{th}]$. In practice, the relationship between α and θ_{th} is in look-up tables [12]. In this work, we set θ_{th} to be 6.25 because we select the false alarm rate α to be 0.1.

If we obtain n partitioned regions from the watershed algorithm, we assign a central vector to each region. A central vector is a representative vector of each region, which is the average vector of a corresponding region. The test statistic θ is now evaluated by the differences between the central vector of a certain region to be tested and the central vectors of the reference regions where the test statistic θ is normalized by variances of the reference regions. Whenever it exceeds θ_{th}, this region is declared as this region is not homogeneous to the previous reference regions. Thus, the estimated number of clusters increases by one.

Fig. 1 shows the estimation procedure to find the number of clusters. The reference regions are updated and the final number of the reference regions is the estimated number of clusters. All regions obtained by the watershed algorithm are sorted with the region size in the descending order. The largest region is the initial reference region $\mathbf{T}_1$ and each region $\mathcal{R}_i$ is represented with the mean vector $\mathbf{m}_i$. In Fig. 1, each region $\mathbf{T}_j$ is the reference region, c is the estimated number of clusters and n is the number of regions obtained from the watershed algorithm.

After the number of clusters is estimated, we merge the regions generated from the watershed scheme by an agglomerative hierarchical clustering algorithm [9]. As described in Fig. 2, this procedure is terminated when the estimated number of clusters is obtained and it returns the final color clusters.

Therefore, the final color clusters can segment images based on colors and regions and the proposed clustering algorithm does not require any a priori information. Since

this clustering algorithm reflects the shape of the region, it does not generate oversegmentation along region boundaries.

Algorithm 1: Estimation of number of clusters
begin initialize n, $\mathbf{T}_1 \leftarrow R_1$, $R_i \leftarrow \{\mathbf{m}_i\}$, $c \leftarrow 1$, $i \leftarrow 1$
do $i \leftarrow i+1$
find nearest $\mathbf{T}_j$ to R_i among c reference regions based on Euclidian Distance
compute θ
if($\theta > \theta_{th}$)
$c \leftarrow c+1$, $\mathbf{T}_c \leftarrow R_i$
else
merge $\mathbf{T}_j \leftarrow R_i$ and $\mathbf{T}_j$
recompute $\mathbf{m}_j$ on $\mathbf{T}_j$
recompute variances of each space on $\mathbf{T}_j$
until $i=n$
return c
end

Fig. 1. Estimation of Number of Clusters

Algorithm 2: Clustering based on Region
begin initialize c, $c^* \leftarrow n$, $R_i \leftarrow \{\mathbf{m}_i\}$
do $c^* \leftarrow c^*-1$
find nearest clusters, say, R_i and R_j based on Euclidian Distance
merge R_i and R_j
until $c = c^*$
return c clusters
end

Fig. 2. Agglomerative Hierarchical Clustering

4 Experimental Results

In order to evaluate our proposed algorithm, we have performed computer simulations on many different kinds of images. Fig. 3 shows the simulation results from the MOTHER AND DAUGHTER image. After we have applied morphological filtering, anisotropic diffusion, watershed, estimation of number of color clusters, and agglomerative hierarchical clustering operations on the sequence, Fig. 3©, Fig. 3(e), Fig. 3(g) and Fig. 3(h) are obtained by the proposed algorithm.

While Fig. 3(b) is the diffused image by Eq. (9) without the morphological filtering, Fig. 3(c) is the diffused image after morphological filtering. In Fig. 3(c), we can ignore specific detail image structures, which are not interesting for semantic

image analysis because of small regions. Fig. 3(d) and Fig. 3(e) are output images by the watershed algorithm on the diffused image without morphological filtering and with morphological filtering, respectively. We note that Fig. 3(e) shows a much simpler output than Fig. 3(d).

Fig. 3(f) shows the segmentation result based on the clustering using a color histogram [10]. Fig. 3(f) demonstrates oversegmentation at region boundaries because the clustering algorithm has directly been applied on the diffused or low-pass filtered images in a certain color space and the diffused or low-pass filtered images can generate the mixed colors along the region boundaries. On the other hand, the final result of our proposed scheme does not generate oversegmentation at region boundaries and it can ignore the unwanted details, as shown in Fig. 3(g).

Algorithm 1 estimates that the MOTHER AND DAUGHTER image has twelve color clusters and Algorithm 2 makes twelve color clusters. Fig. 3(h) displays the segmented image that is represented by the final twelve color clusters.

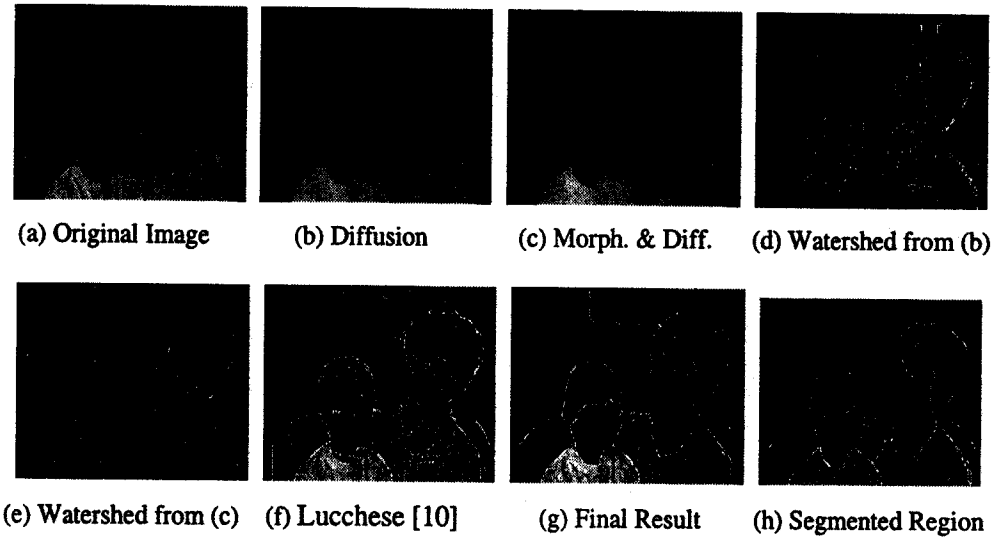

Fig. 3. Result for MOTHER AND DAUGHTER

The proposed algorithm works well on other types of images, as shown in Fig. 4. In addition, unlike the previous works [10, 11], this algorithm operates without any color quantization and is independent on the positions of the initial mean vectors.

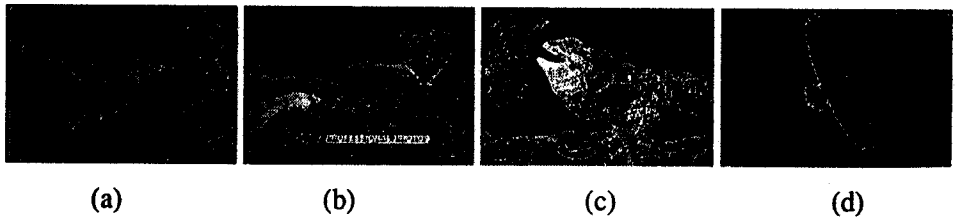

Fig. 4. Segmentation Results for Various Images

5 Conclusions

In this paper, a new region-based color segmentation algorithm has been presented. The proposed algorithm consists of image simplification, region labeling and color clustering. In order to take advantage of the perceptually uniformly distributed LUV color space, we perform the vector-valued diffusion process in the LUV color space. In addition, statistical characteristics of each labeled region are employed to estimate the number of total clusters and agglomerative hierarchical clustering is performed with the estimated number of clusters. Therefore, the proposed clustering algorithm does not require any a priori information. The proposed algorithm is also stable because it does not need any initial conditions for the clustering operation. Since the proposed clustering algorithm reflects the shape of the region, it does not generate oversegmentation along region boundaries.

Acknowledgement. This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fiber-Optic Networks (UFON) Research Center at Kwangju Institute of Science and Technology (K-JIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

References

1. Skarbek, W., Koschan, A.: Colour Image Segmentation – A Survey. Technical Report 94-32, Technical University of Berlin (1994)
2. Salembier, P., Pardas, M.: Hierarchical Morphological Segmentation for Image Sequence Coding. IEEE Trans. Image Processing, Vol. 3, No. 5 (1994) 639-651
3. Perona, P., Malik, J.: Scale-Space and Edge Detection using Anisotropic Diffusion. IEEE Trans. PAMI, Vol. 12, No. 7 (1990) 629-638
4. Plataniotis, K.N., Venetsanopoulos, A.N.: Color Image Processing and Applications. Springer, New York (2000)
5. Zenzo, S.D.: A note in the Gradient of a Multi-Image. CVGIP, Vol. 33 (1986) 116-125
6. Sapiro, G., Ringach, D.L.: Anisotropic Diffusion of Multivalued Images with Applications to Color Filtering. IEEE Trans. Image Processing, Vol. 5, No 11, (1996) 1582-1585
7. Guichard, F., Moisan, L., Morel, J.-M.: A Review of P.D.E. Models of Image Processing and Image Analysis. Journal de Physique IV, Vol. 12 (2002) 137-154
8. Vincent, L., Soille, P.: Watersheds in Digital Spaces: an Efficient Algorithm based on Immersion Simulations. IEEE Trans. PAMI, Vol. 13, No.5 (1991) 583-598.
9. Duta, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. John Wiley & Sons Inc., Singapore (2001)
10. Lucchese, L., Mitra, S.K.: Unsupervised color image segmentation. Proc. IEEE Workshop on Multimedia Signal Processing (1998) 33-38
11. Deng, Y., Manjunath B.S., Shin, H.: Color Image Segmentation. Proc. of IEEE Conf. on CVPR, Vol. 2 (1999) 446-451
12. Montgomery, D.C.: Design and analysis of Experiments 5th Edition. John Willy & Sons Inc., New York (2001)