The following manuscript was published in

Color Image Segmentation Using Self-Organizing Map Algorithm

Hui-Yu Huang* Yung-Sheng Chen †
Wen-Hsing Hsu *‡

* Department of Electrical Engineering National Tsing Hua University
101 sec. 2, Kuang Fu Road, Hsinchu, Taiwan 300, R.O.C.
† Department of Electrical Engineering Yuan Ze University
135 Yuan-Tung Road, Chung-Li, Taiwan 320, R.O.C.
‡ Institute of Information Science, Academia Sinica
Taipei, Taiwan 115, R.O.C.

Abstract

A color image segmentation methodology based on Self-Organizing Map (SOM) is proposed. This developed method takes into account the color similarity and spatial relationship of objects within an image. According to the features of color similarity, an image is first segmented into coarse cluster regions. The resulting regions are further treated by computing the spatial distance between any two cluster regions, and SOM with a labeling process is applied. In this paper, the selection of parameters for the SOM algorithm was also investigated experimentally. The experimental results show that the proposed system is feasible, and that the segmented object regions are similar to those perceived by human vision.

Keywords: color image segmentation, Self-Organizing Map (SOM), spatial relationship, labeling process.

*corresponding author address: Department of Electrical Engineering National Tsing Hua University, 101 Sec. 2, Kuang Fu Road, Hsinchu, Taiwan 300, R.O.C. TEL: +886-3-5710300, FAX: +886-3-5715971, Email address: whhsu@snoopy.ee.nthu.edu.tw
1 Introduction

Image segmentation is the process by which an original image description is partitioned into some meaningful regions. It is usually assumed that the objects are represented by regions which are homogeneous and well-defined in some sense. Segmentation is also a significant task in image understanding. Many researchers have focused their attention on the quality of segmentation results and the development of techniques which are generally based on the engineering viewpoint. From the viewpoint of vision, human beings can focus intuitively on some objects and automatically cluster objects in the given image. In this paper, we present an approach to color image segmentation, which is related to color information processing. Color is a perceptual phenomenon related to human response to different wavelengths in the visible electromagnetic spectrum [1]. Human beings intuitively feel that color is an important part of their visual experience, and color is useful for a powerful processing in computer vision. Color image segmentation techniques can roughly be categorised as techniques which chromatically divide an image space and those which cluster a color feature space derived from an image. The color segmentation scheme adopted in this paper belongs to the latter category. We treat color image segmentation as the problem of classifying color pixels and spatial relationships.

Many approaches to color segmentation have been developed, including fuzzy c-means (FCM) [2, 3, 4], region growing and merging [5, 6, 7], region growing and edge information [8]–[11], the Markov random field (MRF) [12]–[15], the neural network [16, 17, 18], histogram-based methods [19]–[24], and the physically-based method [25]. For a more detailed review of gray and color image segmentation methods, the reader can refer to a general review [26]. Lim and Lee [2] proposed a color segmentation method which combines the thresholding and FCM techniques. This segmentation strategy is mainly a kind of coarse to fine technique. In the coarse segmentation stage, coarse segmentation is performed using the thresholding technique. Those pixels which are not segmented by in the coarse segmentation process are further segmented using the FCM in the fine segmentation stage. Histogram analysis for the setting of thresholds has also been used successfully in image segmentation. However, with complex natural images,
thresholds are difficult to decide, and the computational cost of processing an image is high. In addition, Noordam et al. [3] used a semi-supervised FCM technique which simultaneously adds geometrical information during clustering in order to obtain the fine segmentation. Adding spatial information can improve the separation of two overlapping clusters in spectral domain corresponding to two different objects in the spatial domain. Thus, the spatial information is an important factor in image segmentation. Huang et al. [12], and Liu & Yang [13] used the scale space filter (SSF) for coarse segmentation and the MRF for refinement. Andrey and Tarroux [14] proposed selectionist relaxation to solve the problem of segmenting MRF modeled textures in unsupervised mode. Owing to the MRF model-based image segmentation, it is hampered by computational complexity. Kim et al. [15] proposed a new genetic algorithm (GA)-based segmentation that can improve the computational efficiency. Meanwhile, edge information is used in another approach to image segmentation. An edge is a place where there is more or less a sharp change in one or more characteristic features. One challenge with this method is the continuity of detected edges. Segments must be enclosed by continuous edges, a difficult task to perform by detecting edges. Therefore, edge detection alone is not a segmentation process. It is not appropriate for outdoor images because the object boundary of such an image usually presents a rough contour. Some approaches [7]–[11] combining edge detection and region growing in color images have been proposed. Region growing and merging methods [5, 6] process the image recursively or iteratively, which may be expensive in terms of computational time and memory.

Neural networks have been adopted in techniques for color image segmentation. The Hopfield networks are one of the widely using techniques [17, 18]. Campadelli et al. [17] proposed two different algorithms: a Hopfield network and a single network according to the the number of clusters obtained using histogram analysis. The histogram method adopts a scale-space histogram filtering technique to obtain optimal decision boundaries for the peaks in the histogram, and then the number of clusters can be obtained by histogram analysis. Cheng and Sun [24] proposed a histogram of the homogeneity domain and color feature domain to process color image segmentation. This method uses the homogeneity feature which takes into account local
and global information to obtain the homogeneous regions, and the final segmentation result was obtained using region growing method. However, in a natural outdoor image, the number of clusters may be difficult to determine using the histogram technique. In spite of the fact that the histogram-based approach to color image segmentation is a very simple and low-level method, it has been used to obtain good results [20, 21], especially in the image indexing and retrieval tasks. However, spatial knowledge is usually lost, meaning that the spatial relations between the parts of an image cannot be used. In order to solve this problem, Chen [27, 28] proposed an unsupervised approach which combines color similarity and spatial information to effectively analyze the clustering problem. Their model is applicable mainly to the problem of colour blindness plate (CBP) image segmentation. The principal idea of this method is to recognize meaningful dotted patterns from a complex CBP image, based on the viewpoint of human visual perception. Hence, we will extend this scheme and further apply it to outdoor image segmentation. In area of visual processing [29], some results of psychophysical experiments show the existence of a hierarchy of visual features based on the relations between image contours. The human visual system appears to be preattentively, selectively sensitive to image contours. Further, any complete vision system will make use of two types of processing: One is low-level processing (or passive processing), which generally is data-driven; the other processing at the highest levels (or active processing), which generally involves considerable top-down, knowledge-based computation. Therefore, in the field of computer vision, researchers have attempted to utilize this human perception viewpoint. Thus, applications which incorporate color image segmentation are becoming increasingly prevalent nowadays.

In this paper, a system based on the human vision viewpoint is proposed for application to color image segmentation. The method used in this system mainly utilizes the self-organizing map (SOM), in which some characteristics are extremely close to those of the human vision mechanism [30]. Our approach simultaneously takes into account color similarities in the color space and spatial relationships. The purpose of this system is to imitate what human vision perceives when human beings observe an image. In accordance with the color features, the image is first segmented into limited clusters, and the segmented clusters are further found by
computing the spatial relationships.

The selection of parameters for the SOM algorithm, i.e., the number of output nodes and iterations, will be discussed and analyzed experimentally. Using the Mean-Square-Error (MSE) measurement, these parameters can be selected. The MSE measurement is obtained by computing the color difference between the resultant image and the original one. The rest of this paper is organized as follows. Section 2 describes the proposed method, including the modeling assumption, the HVC (hue, value, chroma) color space, the SOM algorithm and the relationships of spatial features. Section 3 presents some experimental results and includes a brief discussion. Finally, a conclusion is given in Section 4.

2 The proposed method

Our method further extends the technique in [28], and applies it to outdoor image segmentation. The main goal of color image segmentation is to simulate what regions human eyes attend when they look at an image. The way human beings perceive an image is simulated by the SOM algorithm, which is based on color features and spatial relationships. Using color features, the SOM algorithm can first segment the image into a set of limited color clusters, defined as color planes. That is, a color plane is a piece-wise constant approximation of the color distribution in a color cube. The spatial relationships are obtained by computing the spatial distance between any two color clusters. The SOM algorithm and labeling process are then utilized again to classify the color clusters into the segmented cluster \( f_{\text{cluster}(c)} \). The system flowchart of our approach is shown in Fig. 1. This approach principally consists of the designed model, the SOM algorithm with the labeling process, and computation of the spatial distance. The system is further described in the following.

2.1 Approach to modeling design

The image is represented by a two-dimensional (2-D) color function \( f(i, j) \), where \( i \) and \( j \) denote spatial coordinates. The pixels \( f(i, j) \) of a color image can be seen as points in the three-dimensional (3-D) color space. The vector \( f(i, j) \) is related to the relative magnitude of the
tristimulus values hue($h$), value($v$), and chroma($c$) color components. It can be denoted by a three-element color vector of $f(i, j) = (h_{i,j}, v_{i,j}, c_{i,j})^T$. Segmentation is a process based on pixel classification. The image $[f]$ is segmented into subsets, called color planes $[f]_k$, $k = 1, 2, \ldots, K$, by assigning the individual color vectors to classes. Each color plane $[f]_k$ possessing the same image size contains one set of pixels with similar color vectors $(h^{(k)}, v^{(k)}, c^{(k)})^T$ and another set of pixels with zero vectors $(0, 0, 0)^T$. Therefore, if the $K$ color vectors of $[f]$ have been found, then the original $[f]$ can be approximately obtained as the sum of $[f]_k$, $k = 1, 2, \ldots, K$:

$$[f] \approx [f]_1 + [f]_2 + \ldots + [f]_K = \sum_{k=1}^{K} [f]_k,$$

where the color plane $[f]_k$ can be approximately defined as

$$f_k(i, j) = \begin{cases} (h^{(k)}, v^{(k)}, c^{(k)})^T, & \text{if } f(i, j) \approx (h^{(k)}, v^{(k)}, c^{(k)})^T, \\ (0, 0, 0)^T, & \text{otherwise.} \end{cases}$$

That is, a color image usually consists of a set of clusters, and each cluster is composed of some color planes.

The clustering process can be performed by measuring the spatial distance between any two color planes, $[f]_h$ and $[f]_k$, ($h \neq k$), based on their spatial relationships. The smaller the distance between two color planes, the greater the possibility that both belong to the same cluster. The greater the distance, the more likely they belong to different clusters. Based on the spatial distance information and relationships, the task of segmenting a color image is reduced to that of classifying color planes into several clusters. Assume that there are $C$ clusters; the segmented color image can be further denoted as follows:

$$[f] \approx \sum_{c=1}^{C} [f]_{\text{cluster}(c)},$$

where

$$[f]_{\text{cluster}(c)} = \sum_{k \in \text{cluster}(c)} [f]_k.$$

### 2.2 The HVC color space

The detection of segmented regions matching a given color feature is a frequently conducted task in image processing. Diverse color identification schemes have been proposed and used.
A color space which describes colors close to human perception is crucial for calculating color differences corresponding to color perceptual differences. The RGB color space has been widely adopted because of its simplicity in implementation [31]. However, the RGB color space, which is normally used in the frame grabber of a color-processing system, does not obtain direct information about a color. It has been shown to be unable to separate luminance and chromatic components. In addition, these values corresponding to RGB components are perceptually non-uniform. For example, changes of color perception are non-linear with respect to numerical changes.

The HVC (hue, value, chroma) color space completely separates the luminance and chromatic components which represent with hue the color type, with value the luminance, and the chroma the colorfulness. There are several ways to transform the RGB into HVC color space. In this paper, transformation is conducted based on the CIEXYZ and CIEL*a*b* color spaces. Assuming a set of RGB color values is given, the transformation of a color value from the RGB color space into the HVC color space is conducted as follows [1].

1. RGB color values are first converted to the CIEXYZ tristimulus values using the following formula

\[
X = 0.49 \times R + 0.31 \times G + 0.20 \times B, \tag{5}
\]

\[
Y = 0.17697 \times R + 0.8124 \times G + 0.01063 \times B, \tag{6}
\]

\[
Z = 0.01 \times G + 0.99 \times B. \tag{7}
\]

In this color space, \( Y \) represents the luminance of the color.

2. CIEXYZ color values are then transformed into uniform color coordinates with CIEL*a*b* and expressed as

\[
L^* = 116 \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16, \quad \text{for } \frac{Y}{Y_n} > 0.008856,
\]

\[
= 903.3 \left( \frac{Y}{Y_n} \right), \quad \text{for } \frac{Y}{Y_n} \leq 0.008856,
\]

\[
a^* = 500 \left[ \left( \frac{X}{X_n} \right)^{\frac{1}{3}} - \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} \right],
\]

\[
b^* = 200 \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} / \left[ \left( \frac{X}{X_n} \right)^{\frac{1}{3}} + \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} + \left( \frac{Z}{Z_n} \right)^{\frac{1}{3}} \right],
\]

where \( X_n, Y_n, Z_n \) are standard color values.
\[ b^* = 200 \left[ \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - \left( \frac{Z}{Z_n} \right)^{\frac{1}{3}} \right], \]  

where \( X_n, Y_n, Z_n \) are the values of \( X, Y, Z \) for the appropriately chosen reference white. Here, these values are selected as 95.046, 100.0, and 108.8826, based on the CIE standard illuminant \( D65 \).

3. The HVC color values are obtained by

\[
\begin{align*}
H_{ab} & = \arctan \left( \frac{b^*}{a^*} \right), \\
V & = 116 \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16, \\
C_{ab}^* & = \left( (a^*)^2 + (b^*)^2 \right)^{\frac{1}{2}}.
\end{align*}
\]  

2.3 The \( SOM_1 \) algorithm

\( SOM \) [30] is mainly used to visualize and interpret high-dimensional signal data sets by mapping them to low-dimensional space based on a competitive learning scheme. \( SOM \) consists of an input layer and an output layer. The number of nodes in the input vector is the same as the dimension of the input vector. Furthermore, the structure of the output layer can consist of 1-D or 2-D connected nodes, depending on the application. The output nodes are connected to each input node with some weights. Based on nearest neighbor competition and the weight adaptation procedure, the index of the winning node is taken as the output of \( SOM \). The Hebbian learning rule is used to adjust the weights of the winning node and its neighborhood nodes. The Hebbian learning rule originated from Hebb’s postulate of learning is described in [32]. This rule consists of two-part rule expanded as follows. 1. \textit{If two neurons on either side of a synapse (connection) are activated simultaneously, then the strength of that synapse is selectively increased.} 2. \textit{If two neurons on either side of a synapse are activated asynchronously, then that synapse is selectively weakened or eliminated.} And Haykin [32] defined a Hebbian synapse as a synapse that has four key mechanisms, time-dependent, local, interactive, and conjunctional or correlational mechanisms. Time-dependent mechanism is that the modifications in a Hebbian synapse depend on the exact time of occurrence of the presynaptic and postsynaptic activities.
The local mechanism enables a neural network which makes of Hebbian synapses to perform unsupervised learning. In interactive mechanism, the occurrence of a change in a Hebbian learning depends on both presynaptic and postsynaptic activity levels. A Hebbian learning is that the condition for a change in synaptic efficiency is the conjunction of presynaptic and postsynaptic activities. According to these properties, the Hebbian learning rule increases its strength with positively correlated presynaptic and postsynaptic activities, and decreases its strength when activities are uncorrelated. Hence, we use this learning rule to adjust the weights of nodes in SOM neural network. SOM is highly related to vector quantization (VQ) and K-mean clustering. One good characteristic of SOM is partial data density preservation when it is properly trained.

In this paper, only a 1-D node of the output layer is adopted. The SOM is first used to classify the pixels of color image $[f]$ into several clusters, and these represented clusters are implemented by means of a labeling process. Here, the labeling process is to assign a number which can serve as an index to each output node of SOM. Thus, the output nodes generate the respective color planes. See [28, 33] for detailed definitions and computations. We will briefly describe these definitions and computations in the following. Consider $N$ input nodes and $M$ output nodes; then, we have a matrix whose element $w_{nm}^{t}$, $1 \leq n \leq N$, $1 \leq m \leq M$, denotes the weight from input node $n$ to output node $m$ at time $t$. Let $W_{m}^{t}$ be the topological neighborhood which is the set of nodes in the neighborhood of node $m$ at time $t$. $W$ decreases slowly in size over time. Then, at time $t$, the distance $l_{m}$ between the input $x_{n}^{t} = (h_{n}, v_{n}, c_{n})^t$ and each output node $m$ is used as the Euclidean distance, given by

$$l_{m} = \sum_{n=1}^{N} \|x_{n}^{t} - w_{nm}^{t}\|, \quad 1 \leq m \leq M,$$

and the output node $c$ with minimum $l_{c}$ is obtained as

$$l_{c} = \min_{m=1}^{M} l_{m}.$$

Then, the weights are used to update all the neurons in the neighborhood around the winning
neuron node \( c \) ranged by \( W_c^t \). The updated formula is expressed as

\[
   u_{nm}^{t+1} = \begin{cases} 
   w_{nm}^t + h_{cm}^t(x_n^t - w_{nm}^t), & 1 \leq n \leq N, \ m \in W_c^t, \\
   w_{nm}^t, & \text{otherwise}. 
   \end{cases} 
\]

(12)

Here, \( t \) is an integer, the discrete-time coordinate, and \( h_{cm}^t \) is a neighborhood kernel function defined over the 1-D array points. Generally, \( h_{cm}^t \) can be written in terms of a step function. Here, we adopt the useful Hanning window [27, 34], which is re-expressed as follows

\[
   h_{cm}^t = \begin{cases} 
   \alpha^t \left[ 1 - sin\left( \frac{\pi (c-m)}{|W_c^t|} \right) \right], & |c - m| \leq \frac{|W_c^t|}{2}, \\
   0, & \text{otherwise}, 
   \end{cases} 
\]

(13)

where \( \alpha^t \) is a gain term \((0 < \alpha^t < 1)\) that decreases a positive variation \( \Delta \alpha \) in time, i.e., \( \alpha^{t+1} = \alpha^t - \Delta \alpha \) and the parameter \( |W_c^t| \) defines the width of the kernel. With increasing \( |c - m| \), \( h_{cm} \to 0 \). If \( |c - m| \to 0 \), \( h_{cm} \to \alpha^t \). In other words, the kernel function \( h_{cm}^t \) attains its maximum value at the winning neuron \( c \) for which the distance \( |c - m| \) is 0. And the amplitude of the kernel function \( h_{cm}^t \) decreases with increasing distance \( |c - m| \), decaying to zero for \( |c - m| \geq |W_c^t|/2 \). Hence, the Hanning window function defined in Equation (13) is used to update the synaptic weight vector \( u_{nm}^t \) of neuron \( n \) at distance \( |c - m| \) from the winning neuron \( c \). Both \( \alpha^t \) and \( |W_c^t| \) are some monotonically decreasing functions of time, and their exact forms are not critical; they could thus be selected as linear. Effective choices for these functions and their parameters can only be determined experimentally [30].

In the SOM algorithm, the selection of parameters, namely, the number of iterations and output nodes, is an important task. Moreover, the effectiveness of the mapping function used with the learning procedure of the SOM algorithm depends mainly on how the parameters are selected. Unfortunately, there is no theoretical basis for the selection of these parameters. They are usually determined by means of a trial and error process or based on experience [30]. In our system, in order to obtain these parameters, the Mean-Square-Error (MSE) measurement, which is easy to implement, is adopted. Let \( f(i, j) \) be an original image denoted by a three-element color vector \((h_{i,j}, v_{i,j}, c_{i,j})^T\), and let \( f'(i, j) = (h'_{i,j}, v'_{i,j}, c'_{i,j})^T \) obtained by the SOM algorithm represent the resultant image. The MSE is defined as

\[
   MSE = \frac{\sum_{i=1}^{h} \sum_{j=1}^{w} ||f(i, j) - f'(i, j)||}{h \times w}, 
\]

(14)
where

$$\| f(i, j) - f'(i, j) \| = ((h_{i,j} - h'_{i,j})^2 + (v_{i,j} - v'_{i,j})^2 + (c_{i,j} - c'_{i,j})^2)^{\frac{1}{2}}.$$  \hspace{1cm} (15)

Here $h$ and $w$ denote the image’s height and width, respectively. The MSE measurement is obtained by computing the difference between two points representing colors in the HVC color space of the original image and the resultant image. Thus, we can obtain the appropriate range of these parameters by means of the MSE measurement. In accordance with the MSE value estimated by our experiments, it represents such a situation that if the number of output nodes increases, then the MSE value becomes smaller and smaller and varies slightly. That is, it is not easy to distinguish between the resultant image and the original one. Furthermore, the smaller the number of output nodes, the greater the variation of the MSE. That is, the MSE shows the specific difference between the resultant image and the original one.

The MSE corresponding to the number of iterations encounters the same situation. In order to obtain a lower MSE, a higher number of iterations may be selected; however, a long period of time will be needed to achieve convergence in the learning stage, and the results and performance of the system will be affected. Hence, we will discuss the selection of these parameters in this paper.

2.4 The relationships among spatial feature

According to the principle of segmentation of color images mentioned previously, it is important to find which cluster the color plane belongs to. This can be done by calculating the spatial distance $d_{hk}$ between any two color planes, $[f]_h$ and $[f]_k$, $(h \neq k)$. The smaller the distance between two color planes, the higher the probability that both belong to the same cluster and vice versa. The detailed procedure is given in [27, 28, 33]. In the following, we will briefly review this process.

Let the distance $d_{hk}$ between two non-zero pixels $f_h(i, j)$ and $f_k(u, v)$ be Euclidean as follows:

$$d(i, j; u, v) = ((i - u)^2 + (j - v)^2)^{\frac{1}{2}},$$  \hspace{1cm} (16)
and let the distance between the non-zero pixel \( f_h(i, j) \) and the color plane \([\mathbf{f}]_k\) be expressed as
\[
d(i, j; [\mathbf{f}]_k) = \min_{\forall \mathbf{f}_h(u, v) \neq (0, 0, 0)^T} d(i, j; u, v). \tag{17}
\]
Then, the distance between color planes \([\mathbf{f}]_h\) and \([\mathbf{f}]_k\) is defined as
\[
d_{hk} = \frac{1}{n([\mathbf{f}]_h)} \sum_{\forall \mathbf{f}_h(i, j) \neq (0, 0, 0)^T} d(i, j; [\mathbf{f}]_k). \tag{18}
\]
Here, \(n([\mathbf{f}]_h)\) presents the number of non-zero pixels belonging to the color plane \([\mathbf{f}]_h\). It serves as a normalizing factor. After computing \(d_{hk}\), we can measure the distance from \([\mathbf{f}]_h\) to \([\mathbf{f}]_k\) and from \([\mathbf{f}]_k\) to \([\mathbf{f}]_h\). Since \([\mathbf{f}]_h\) and \([\mathbf{f}]_k\) are different, \(n([\mathbf{f}]_h)\) and \(n([\mathbf{f}]_k)\) are also different. This implies that \(d_{hk}\) will not usually be equal to \(d_{kh}\). Thus, a more appropriate measure is calculated based on the average \(\bar{d}_{hk} = (d_{hk} + d_{kh})/2\) and \(\bar{d}_{hk} = \bar{d}_{kh}\).

### 2.5 The SOM₂ Algorithm

In order to implement Equations (3)-(4) and the final segmented results, the SOM algorithm will be used again. The second SOM₂ algorithm is similar to SOM₁ with a labeling process except for the difference in input data. That is, the input data becomes the vector form \((x_1, x_2, \ldots, x_K)^T\). According to the previous results given in Subsection 2.4, the average distance \(\bar{d}^k = (\bar{d}_{hk}, h, k = 1, 2, \ldots, M)\) can be obtained. The \(\bar{d}^k\) is input nodes with the SOM₂ algorithm. After performing Equations (10)-(12), we can obtain the \(k\)-element distance vector, denoted as \(d^{(m)} = (d_1^{(m)}, d_2^{(m)}, \ldots, d_k^{(m)}), m = 1, 2, \ldots, M_2\). Hence, \(K\) input nodes and \(M_2\) output nodes are utilized in the SOM algorithm (SOM₂). This distance vector can serve as an index matrix; if a distance column vector of the input data \(\bar{d}^k\) closely matches the \(m\)th node of the index matrix, then the input vector is labeled \(m\).

In other words, the \(C\) best-matching nodes selected from the \(M_2\) output nodes are obtained in the labeling process, where \(C \leq M_2\). This implies that, in general, the \(K\) color planes are classified into \(C\) clusters. Each cluster is represented by the straightforward sum of the color planes belonging to the same cluster as Equation (4) indicates.

The proposed system is briefly summarized as follows. The purpose of our system is to perform and represent color image segmentation. In this approach, the principal color planes
are first obtained by means of an unsupervised segmentation algorithm (SOM). In order to obtain the segmented clusters, the spatial distance between any two color planes is further computed. In addition, the SOM algorithm is applied again to the former distance vectors. The final segmented clusters can thus be obtained.

3 Experiments and discussion

3.1 Setting parameters

Figs. 2(a) and (b) show the MSE value evaluated based on the HVC color space corresponding to the number of output nodes and iterations for the SOM algorithm, respectively. The MSE used by Equations (14)-(15) is obtained by computing the color difference between the resultant image and the original image. Following [30], the selection of parameters is not critical, and there is no theoretical basis; rather, it depends on the image data. The appropriate selection of these parameters can only be determined experimentally. In our experiments on outdoor images, these parameters were decided by measuring the MSE. Fig. 2 clearly shows that the MSE tends to be stable and to vary smoothly when the number of output nodes exceeds about 70. If the number of output nodes is higher than this value, the resultant image will very approximate to the original image. That is to say, human beings are difficult to distinguish the resultant and original images. Thus, the resultant image and the original one are very close to those observed by human beings. However, if this parameter is too large, the MSE will vary slightly, of course, because human vision will not easily to distinguish the difference between the resultant image and original one. This is because some output nodes representing colors are very similar; that is, according to the SOM characteristics, many the output nodes are clustered to the same cluster by competitive learning of the SOM. In spite of the fact that the MSE will become smaller and smaller when a large number is selected, some of these output nodes will be very close to each other. Too many output nodes will cause the computation load to be too heavy and affect system efficiency. Hence, the number of output nodes must not be large.
The number of iterations of the SOM algorithm also affects the result and convergence time of the neural network. In order to obtain a lower MSE and better clustering result, a higher number of iterations should be selected; however, this will result in a long computing time and can affect the performance of the system. Therefore, based on our experimental measurement results for outdoor images, the number of output nodes may be selected in the range from 50 to 100 output nodes; the other parameter, the number of iterations, can be assigned in the range from 10 to 30. In our experiments, the former parameter was set to 80, and the latter one was set to 20. We will describe the procedure used in the SOM algorithm and a series procedures in the following subsection.

3.2 Training and testing

An original color image \([\mathbf{f}]\) whose size was 256 \(\times\) 256 pixels is shown in Fig. 3(a). Each pixel which was transformed into the HVC color space by means of Equations (5)-(9) was a three-element color vector of the form \(f(i, j) = (h_{i,j}, v_{i,j}, c_{i,j})^T\). There were three input nodes, i.e., \(N = 3\), we set the number of output nodes (representing color planes) to \(M = 80\), the initial width of window to \(|W_{\text{initial}}| = 7\), the gain to \(\alpha = 0.1\) and the initial value of the weights to a random number in the range within 0 and 1 in our experiments. With Equations (10)-(12) for the SOM$_1$ algorithm, three input nodes and 80 output nodes were used in the SOM$_1$ as shown in Fig. 1. Next, the labeling process was performed on the original color image as shown in Fig. 3(a). If a pixel of the color image matched the \(m\)th node closely, then the pixel was labeled \(m\). Therefore, the original image could be approximately indicated to sum of some color planes by means of Equation (1). Fig. 3(b) displays the resultant image of merging 80 output nodes corresponding to Fig. 3(a). Obviously, Fig. 3(b) is very close to the original image.

In order to obtain the clustering results, the spatial relationships were then utilized. After performing spatial distance computation between any two color planes using Equations (16)-(18), we obtained the distance matrix shown in Table 1(a), \([\mathbf{f}]_h\) and \([\mathbf{f}]_k\), \((h \neq k)\), \(h, k = 1, 2, \ldots, M\). By averaging as described in the previous section, the symmetrical distance matrix was obtained and is shown in Table 1(b). Table 1 only presents the top 10 elements in the
distance matrix and the symmetrical distance matrix, respectively. Therefore, the 80-element
distance vectors of the form \( \tilde{d}^{(k)} = (\tilde{d}_{1k}, \tilde{d}_{2k}, \ldots, \tilde{d}_{80k})^T, k = 1, 2, \ldots, 80 \) were obtained.

According to the previous results, the second SOM\(_2\) had 80 input nodes and \( M_2 = 15 \) output nodes (here the gain was \( \alpha = 0.1 \), the width of the window was \( |W_{\text{initial}}| = 2 \), and the initial weights were also assigned as a random number in the range between 0 to 1 ), and the labeling process was performed. Based on a SOM mechanism, the SOM\(_2\) possesses the same property of SOM\(_1\). The selection of the number of output nodes in SOM\(_2\) is determined by experience. Thus, the number of output nodes must be smaller than the output nodes of SOM\(_1\). This is because too many output nodes will cause the long computing time and some of these output nodes are of uselessness. Hence, in order to increase effect of our experiments, it was set to 15. Each output node in this stage was an 80-element distance vector of the form \( \tilde{d}^{(m)} = (\tilde{d}_1^{(m)}, \tilde{d}_2^{(m)}, \ldots, \tilde{d}_{80}^{(m)})^T, m = 1, 2, \ldots, M_2 \). If a distance column vector of the input data closely matched the \( m\th \) node displayed in Table 2, then the input vector was labeled \( m \). The results show that the 80 color planes resorted by label value were classified into 4 clusters. Thus, \( C = 4 \),

\[
[f]_{\text{cluster}(1)} = \sum_{k=1}^{5} [f]_{k},
\]

\[
[f]_{\text{cluster}(2)} = \sum_{k=6}^{23} [f]_{k},
\]

\[
[f]_{\text{cluster}(3)} = \sum_{k=24}^{50} [f]_{k},
\]

\[
[f]_{\text{cluster}(4)} = \sum_{k=51}^{80} [f]_{k}.
\] (19)

Then, we had the 4 segmented clusters illustrated in Figs. 3(c)-(f), respectively. (The other experimental results are shown in Figs. 8-10.)

We will now compare our results with those obtained using the histogram-based technique [19] and region growing by merging method [35]. Figs. 4-7 show the results. Figs. 4-7(a) are testing images, and the segmented results obtained using the histogram-based technique, region growing and merging method, and our proposed method are shown in Figs. 4-7(b), 4-7(c) and
4-7(d), respectively. In addition, each color in the results was regarded as a cluster region. As for the histogram-based method, its advantage is that it does not require information about the image in advance. However, this approach suffers from a lack of spatial relations. Owing to this lack of spatial relations, it is obvious that some clusters should belong to the same cluster but were segmented out individually, such as those on the region boundaries shown in Figs. 4-7(b). The images shown in Figs. 4-7(b) were segmented into 8, 9, 7, and 8 clusters, respectively. Therefore, if only the color information to perform color image segmentation, poor results may be obtained. As for the region growing and merging method, the main technique is to identify various region that have similar features. The image is first divided into atomic regions using the color difference between pixels, and then similar adjacent regions are merged sequentially until the adjacent regions become sufficiently different. This approach can obtain the good results in homogeneous regions. However, many regions in an image usually represent the non-uniform or varied property, and these regions simultaneously possess the property of global spatial relation not only the neighborhood. Furthermore, the merging condition depending on a different boundary is not easily decided. However, this is an important factor for obtaining merging regions. If only the adjacent regions to perform color image segmentation, many small regions belonging to the same region cannot be merged together as shown in Figs. 4-7(c). The images shown in Figs. 4-7(c) were segmented into 90, 92, 39, and 56 clusters, respectively. From the segmented results, it is seen that the results obtained using our proposed method were the principal regions in the image. Based on the color and spatial features, the original image can be segmented into primary clusters. That is, each segmented cluster contains color and spatial information. The color information is mainly derived from the results of the SOM$_1$ algorithm, and the spatial information is then calculated based on the spatial distance between any two color planes and using the SOM$_2$ algorithm. The features, color and spatial information are preserved in each segmented cluster. Furthermore, each segmented cluster is perceived first, intuitively, to represent the global meaning of the image, and the global meaning may be represented by a linguistic variable and/or inference rules. In other words, the segmented clusters still contain the fundamental information of the original image which is very helpful
for high-level image understanding (i.e., linguistic descriptions).

To conduct a performance evaluation of the histogram-based method, region growing and merging method, and our proposed method, we also adopted the MSE for measuring the color difference in the HVC color space between the original image and the segmented image, as shown in Table 3. Table 3 shows the numbers of clusters and the MSE values obtained using the different methods. From the MSE values, it is seen that our purposed method for outdoor images can obtain a lower color difference. This is because each segmented cluster consists of some color planes among 80 color planes when our proposed system is employed (see the illustration of Equation (19) for reference). Although, our results consist of a few clusters, they contain many color planes that belong to the same cluster based on the color feature and spatial information. Hence, the MSE for the color difference is lower than the results obtained using the histogram-based method, region growing and merging method.

When human beings look at an image, some regions are perceived in accordance with the primary-view perception [36], such as the sky, building, and grass in the image shown in Fig. 3(a). The experimental results obtained using our system primarily consist of these segmented regions, which are acceptable for the primary-view perception. Hence, the segmented results might be further interpreted as having linguistic meaning based on human knowledge. For example, Figs. 3(c), (e), and the other two, (d) and (f), might serve as "sky," "grass," and "building" regions, respectively. These regions can be clustered into meaningful groups according to the compactness structure of these objects.

Fig. 8(a) mainly consists of some object regions, including sky, a building, and grass. The segmented results almost represents these regions. Figs. 8(b), (d), and (e) might be interpreted as "sky," "grass," and "building," respectively. Fig. 8(c) can be ignored because it cannot make up a grouping, and also difficult for human beings to perceive. The other experimental results shown in Figs. 9 and 10 might also be interpreted the global meaning of objects. Although some objects, such as the mountain, are not clearly perceived, this will not extremely affect interpretation of the image.
3.3 Discussion

The proposed system is briefly discussed as follows. The main goal of our system is to represent the regions that human eyes attend when they look an image. That is, the aim is to simulate the primary-view visual perception described in [36], which approach was designed to deal with gray scale images. These displayed regions and their relationships are produced to represent the global meaning of an image, and this global meaning can be further explored by means of the linguistic variables and inference rules. The principal color planes are first obtained using an unsupervised segmentation algorithm \(SOM_1\). In order to obtain segmented clusters, the spatial distance between any two color planes is computed and the unsupervised algorithm \(SOM_2\) is applied again to the previous distance vectors. The main problem with the proposed system is that the segmented clusters may not be precisely produced since the current system is performed using color and spatial information and may be lacking the other useful information. If some useful information (e.g., texture, knowledge base with attention mechanism, and region size or contrast of regions) is added, the performance should be further improved.

As a result, in our proposed system, the color image, including the principle regions, can be coarsely segmented. Basically, the represented regions are very similar to those perceived by human beings. These regions provide helpful information which can be used to further interpret the image meaning using human semantic descriptions.

4 Conclusions

In this paper, we have described our color image segmentation system based on the viewpoint of human vision and performed using the \(SOM\) algorithm. The principal characteristic of this method is that it takes into account both color similarity and spatial relationships. The goal of this approach is to present the regions that human vision perceives when human beings observe an image. Furthermore, the represented region is a rough segmentation region different from the precise segmentation region of the conventional literature indicated. In the first step of our segmentation method, color information based on human vision perception is first extracted in
$SOM_1$. Hence, in this step, it can be regarded as a pure color segmentation. In the second process of using $SOM_2$, spatial information based on computing the spatial distance between any two color planes obtained from the $SOM_1$ is taken into consideration. It is an important factor for further segmenting color regions in our approach.

To sum up, in our system, the $SOM_1$ algorithm with a labeling process is used to classify pixels based on the color similarity of an image into some color planes. The spatial-distance relationships and the $SOM_2$ with a labeling process are exploited to further classify the color planes into segmented clusters. The segmented result shows coarse segmentation, not crisp segmentation. This result may still misclassify some clusters by using color information and spatial information to process the color image segmentation. However, this result is acceptable from the viewpoint of human observation. It provides primary information which can then be used in object interpretation and analysis. The experimental results show that the proposed system is valid and feasible.

Consequently, our system provides some primary information that can be used to interpret object features and semantic descriptions. In the future, we will further investigate and develop a complete system for content-based image processing with semantic descriptions and regions association.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Council of Republic of China under Grant No. NSC 89-2218-E-007-025.

References


Authors

Hui-Yu Huang was born in Taichung, Taiwan on 1968. She received B.S. degree from Feng Chia University in 1992, and M.S. degree in Electrical Engineering Yuan-Ze Institute of Technology, Chung-Li, Taiwan, R.O.C. in 1994. She is currently active as Ph.D. student of Department of Electrical Engineering, National Tsing Hua University, Taiwan. Her research interests include fuzzy logic, neural networks, pattern recognition and color image processing.

Yung-Sheng Chen was born in Taiwan, on 30 June 1961. He received BS degree from Chung Yuan Christian University in 1983, and the MS and Ph.D. degrees from National Tsing Hua University, Taiwan, in 1985 and 1989, respectively, all in electrical engineering. He received a Best Paper Award from the Chinese Institute of Engineers in 1989. He is a member of the IEEE, ACM, and IPPR of ROC. In 1991, he joined the Electrical Engineering Department of the Yuan-Ze Institute of Technology, Taoyuan, Taiwan, ROC, where he is now a Professor. Since 1998, his name is listed in the Who’s Who of the World. His interests include human visual perception, neural model, fuzzy computing, computer vision, and circuit design.
Wen-Hsing Hsu was born in Taipei, on 17 May 1950. He received B.S. degree in electrical engineering from National Cheng Kung University in 1972, M.E. and Doctor of Engineering degrees in Electrical Engineering from Keio University, Tokyo, Japan, in 1978 and 1982, respectively. In 1982, he joined the Department of Electrical Engineering at National Tsing Hua University, and is currently a professor there. He also joined the Institute of Information Science as an associate research fellow in 1984 and is currently a research fellow there. Since August 1990, he was been the Deputy General Director of Opto-Electronics and Systems Lab., Industrial Technology Research Technology Research Institute. He received a Technology Discovery Award from the Sun Yat-Sen Cultural Foundation in 1986, and was honored as an Outstanding Person on Information Science in the Republic of China in 1986, and one of the ten outstanding engineers in the Republic of China by the Chinese Institute of Engineers in 1988. He also received an Outstanding Research Award from the National Science Council, Republic of China in 1988, 1989 and 1990 and a Best Paper Award from the Chinese Institute of Engineers in 1989. His research interests include image processing, pattern recognition, parallel algorithm, neural computer, fuzzy theory, and computer security. Dr. Hsu is a member of the Institute of Electronics, Information and Communication Engineers (IEICE), Information Processing Society of Japan, and IEEE Computer Society of USA.
Color Image

Transformation of Color Space

$SOM_1$ Algorithm with Labeling Process

Color planes

Computation of Spatial Distance

$SOM_2$ Algorithm with Labeling Process

Segmented Clusters

Fig. 1: Flowchart of the proposed method.
Fig. 2: (a) MSE (estimated based on the HVC color space) versus the number of output nodes. (b) MSE (estimated based on the HVC color space) versus the number of iterations, respectively.
Fig. 3: (a) Original image. (b) Result of merging 80 output nodes (color planes). (c)-(f) The segmented results.
Fig. 4: (a) Original image. (b) Result of segmentation using the histogram-based method [19]. (c) Result of segmentation using region growing and merging method [35]. (d) Result of segmentation using the proposed method.
Fig. 5: (a) Original image. (b) Result of segmentation using the histogram-based method [19]. (c) Result of segmentation using region growing and merging method [35]. (d) Result of segmentation using the proposed method.
Fig. 6: (a) Original image. (b) Result of segmentation using the histogram-based method [19]. (c) Result of segmentation using region growing and merging method [35]. (d) Result of segmentation using the proposed method.
Fig. 7: (a) Original image. (b) Result of segmentation using the histogram-based method [19]. (c) Result of segmentation using region growing and merging method [35]. (d) Result of segmentation using the proposed method.
Fig. 8: (a) Original image. (b)-(e) The segmented results.
Fig. 9: (a) Original image. (b)-(d) The segmented results.
Fig. 10: (a) Original image. (b)-(c) The segmented results.
Table 1: (a) The top 10 elements of the distance matrix $\vec{d}^{(k)} = (d_{1k}, d_{2k}, \ldots, d_{80k})^T$, $k = 1, 2, \ldots, 80$, computed using Equations (16)-(18). (b) By averaging, the symmetrical distance matrix can be obtained, and the partial results are shown here.

(a)

<table>
<thead>
<tr>
<th>[f]_h</th>
<th>k=1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>h = 1</td>
<td>0.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.3</td>
<td>1.0</td>
<td>1.0</td>
<td>1.4</td>
<td>1.7</td>
<td>1.3</td>
<td>3.9</td>
</tr>
<tr>
<td>2</td>
<td>1.4</td>
<td>0.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.6</td>
<td>1.0</td>
<td>2.6</td>
</tr>
<tr>
<td>3</td>
<td>2.5</td>
<td>1.2</td>
<td>0.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>1.7</td>
<td>1.0</td>
<td>4.5</td>
</tr>
<tr>
<td>4</td>
<td>9.7</td>
<td>1.3</td>
<td>1.0</td>
<td>0.0</td>
<td>1.2</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>15.4</td>
</tr>
<tr>
<td>5</td>
<td>7.9</td>
<td>1.3</td>
<td>1.0</td>
<td>0.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>19.6</td>
</tr>
<tr>
<td>6</td>
<td>2.5</td>
<td>1.1</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>0.0</td>
<td>1.0</td>
<td>1.8</td>
<td>1.1</td>
<td>7.4</td>
</tr>
<tr>
<td>7</td>
<td>13.7</td>
<td>1.3</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>0.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>18.1</td>
</tr>
<tr>
<td>8</td>
<td>17.2</td>
<td>3.1</td>
<td>1.7</td>
<td>1.0</td>
<td>1.1</td>
<td>1.9</td>
<td>1.0</td>
<td>0.0</td>
<td>1.0</td>
<td>25.0</td>
</tr>
<tr>
<td>9</td>
<td>6.8</td>
<td>1.4</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.1</td>
<td>1.0</td>
<td>0.0</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.7</td>
<td>1.6</td>
<td>1.6</td>
<td>2.0</td>
<td>3.8</td>
<td>2.2</td>
<td>1.1</td>
<td>1.1</td>
<td>1.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

(b)

<table>
<thead>
<tr>
<th>[f]_h</th>
<th>k=1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>h = 1</td>
<td>0.0</td>
<td>1.2</td>
<td>1.8</td>
<td>5.5</td>
<td>4.4</td>
<td>1.8</td>
<td>7.6</td>
<td>9.5</td>
<td>4.1</td>
<td>2.8</td>
</tr>
<tr>
<td>2</td>
<td>1.2</td>
<td>0.0</td>
<td>1.1</td>
<td>1.1</td>
<td>1.2</td>
<td>1.1</td>
<td>1.2</td>
<td>2.3</td>
<td>1.2</td>
<td>2.1</td>
</tr>
<tr>
<td>3</td>
<td>1.8</td>
<td>1.1</td>
<td>0.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.7</td>
<td>1.0</td>
<td>3.1</td>
</tr>
<tr>
<td>4</td>
<td>5.5</td>
<td>1.1</td>
<td>1.0</td>
<td>0.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>8.7</td>
</tr>
<tr>
<td>5</td>
<td>4.4</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0</td>
<td>0.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>11.7</td>
</tr>
<tr>
<td>6</td>
<td>1.8</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.0</td>
<td>0.0</td>
<td>1.0</td>
<td>1.9</td>
<td>1.1</td>
<td>4.8</td>
</tr>
<tr>
<td>7</td>
<td>7.6</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>9.6</td>
</tr>
<tr>
<td>8</td>
<td>9.5</td>
<td>2.3</td>
<td>1.7</td>
<td>1.0</td>
<td>1.1</td>
<td>1.9</td>
<td>1.0</td>
<td>0.0</td>
<td>1.0</td>
<td>13.0</td>
</tr>
<tr>
<td>9</td>
<td>4.1</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.0</td>
<td>4.3</td>
</tr>
<tr>
<td>10</td>
<td>2.8</td>
<td>2.1</td>
<td>3.1</td>
<td>8.7</td>
<td>11.7</td>
<td>4.8</td>
<td>9.6</td>
<td>13.0</td>
<td>4.3</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Table 2: Result of output nodes obtained using the SOM₂ algorithm. Each node is an 80-element distance vector of the form \( \mathbf{d}^{(m)} = (d_1^{(m)}, d_2^{(m)}, \ldots, d_{80}^{(m)})^T \), \( m = 1, 2, \ldots, 15 \). Here the top 10 ranking results are listed.

<table>
<thead>
<tr>
<th>Vector element</th>
<th>Output Node</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m = 1 )</td>
<td>2</td>
</tr>
<tr>
<td>( d_1^{(m)} )</td>
<td>38</td>
</tr>
<tr>
<td>( d_2^{(m)} )</td>
<td>20</td>
</tr>
<tr>
<td>( d_3^{(m)} )</td>
<td>16</td>
</tr>
<tr>
<td>( d_4^{(m)} )</td>
<td>17</td>
</tr>
<tr>
<td>( d_5^{(m)} )</td>
<td>16</td>
</tr>
<tr>
<td>( d_6^{(m)} )</td>
<td>18</td>
</tr>
<tr>
<td>( d_7^{(m)} )</td>
<td>16</td>
</tr>
<tr>
<td>( d_8^{(m)} )</td>
<td>18</td>
</tr>
<tr>
<td>( d_9^{(m)} )</td>
<td>16</td>
</tr>
<tr>
<td>( d_{10}^{(m)} )</td>
<td>43</td>
</tr>
</tbody>
</table>

Table 3: Comparison obtained using the results of the histogram-based method, region growing and merging method, and the proposed method. The MSE is evaluated based on the color difference in HVC color space.

<table>
<thead>
<tr>
<th>Image</th>
<th>Number of clusters</th>
<th>MSE</th>
<th>Image</th>
<th>Number of clusters</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Histogram -based</td>
<td>Region</td>
<td>Our</td>
<td>Histogram -based</td>
<td>Region</td>
</tr>
<tr>
<td></td>
<td>method</td>
<td>growing</td>
<td>proposed</td>
<td>method</td>
<td>growing</td>
</tr>
<tr>
<td>Fig. 4(a)</td>
<td>8</td>
<td>90</td>
<td>4</td>
<td>20.034</td>
<td>29.928</td>
</tr>
<tr>
<td>Fig. 5(a)</td>
<td>9</td>
<td>92</td>
<td>4</td>
<td>19.946</td>
<td>32.408</td>
</tr>
<tr>
<td>Fig. 6(a)</td>
<td>7</td>
<td>39</td>
<td>3</td>
<td>25.732</td>
<td>31.464</td>
</tr>
<tr>
<td>Fig. 7(a)</td>
<td>8</td>
<td>56</td>
<td>2</td>
<td>20.954</td>
<td>27.660</td>
</tr>
</tbody>
</table>