COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 52, 145–170 (1990)

# A Color Clustering Technique for Image Segmentation

MEHMET CELENK

*Department of Electrical and Computer Engineering, Ohio University, Stocker Center, Athens, Ohio 45701*

Received July 1, 1988; accepted September 5, 1989

This paper describes a clustering algorithm for segmenting the color images of natural scenes. The proposed method operates in the 1976 CIE ($L^*$, $a^*$, $b^*$)-uniform color coordinate system. It detects image clusters in some circular-cylindrical decision elements of the color space. This estimates the clusters' color distributions without imposing any constraints on their forms. Surfaces of the decision elements are formed with constant lightness and constant chromaticity loci. Each surface is obtained using only 1D histograms of the $L^*$, $H°$, $C^*$ cylindrical coordinates of the image data or the extracted feature vector. The Fisher linear discriminant method is then used to project simultaneously the detected color clusters onto a line for 1D thresholding. This permits utilization of all the color properties for segmentation and inherently recognizes their respective cross correlation. In this respect, the proposed algorithm also differs from the multiple histogram-based thresholding schemes.


## 1. INTRODUCTION

The segmentation of images is the critical first step in image information extraction for computer vision systems [1]. This process determines the simple structural characteristics of objects' surfaces using detailed image domain properties. Although the identification of object and surface boundaries comes naturally to a human observer, accurate image segmentation has proved difficult and complex. Achieving an adequate segmentation result depends mainly on devising techniques to detect uniformity among the feature values of the picture points, and then isolating the areas of the picture exhibiting these uniformities. Several techniques have been proposed to accomplish this, such as edge detection, region growing, histogram thresholding, and clustering [2, 3]. Among these, the last two methods have been extensively used for segmenting the color pictures. For example, a multidimensional histogram thresholding scheme was suggested by Ohlander, Price, and Reddy [4], using threshold values obtained from three different color coordinate systems (RGB, YIQ, and hue-saturation-intensity). This method takes a region of the image, and using histograms of the nine redundant attribute values in this segment, determines a threshold in one color component to split the region into smaller parts. Some work on the segmentation of scenes using the 2D projection of color space onto the chromaticity triangle was reported by Tenenbaum *et al.* [5]. A 2D clustering technique described by Underwood and Aggarwal [6] and Ali and Aggarwal [7] is based on the projections of the ($X, Y, I$) normalized color space onto the *X-Y*, *X-I*, and *Y-I* planes. In this interactive method, rectangular type broad and refined color bandpass filters are used to detect insect infestations in citrus groves from infrared aerial photographs. Ali, Martin, and Aggarwal [8] further explored this approach for the segmentation of color aerial photographs using quadratic decision surfaces in an interactive operation. In a 3D clustering technique used by Schacter, Davis, and Rosenfeld [9], the 3D image histogram is



0734-189X/90 $3.00

stored in a binary tree, with the key being RGB values, and the information being the number of points with this key value. Clusters are detected if the number of points exceeds some threshold and if they lie within some distance from each other. A similar approach for 3D clustering in the $(X, Y, I)$ normalized color space was employed by Sarabi and Aggarwal [10] to develop a general interactive system for segmenting the color scene. Recently, Klinker, Shafer, and Kanade [33] described an approach to segment surfaces with color variations due to highlights and shading using 3D clustering.

From the image segmentation point of view, color clustering is a computationally expensive process [11], since the feature space is three dimensional. One solution to this problem is to project the feature space onto its lower dimensional subspaces (i.e., 2D planes) as described in [6, 7, 8]. Another alternative is to develop efficient methods of storing and processing the 3D color image data as reported in [9, 10]. A more effective method as described here is to project the color space onto the selected coordinate axes repeatedly until the image clusters are enclosed in some specified volume elements. From each projection of the space, we determine one particular surface of the decision elements. This estimates the clusters' distributions in the 3D color space without imposing any constraints on their forms. The goal here is to reduce the computational cost involved in maintaining many derived images and their respective histograms and in forming 3D decision surfaces.

The estimated color clusters are then projected onto a line for 1D thresholding using the information obtained from the 3D feature space. This permits utilization of all the property values of clusters for segmentation and inherently recognizes their respective cross correlation. This way, the region acceptance is not limited to the information available from one color component. Furthermore, the error rate associated with region acceptance is minimized in the Fisher sense [23]. In this respect, the proposed algorithm also differs from the multidimensional histogram thresholding technique, which will further be addressed in the paper.

Success of this method relies on the existence of 1D clusters in the initial projections of the space. An additional feature set is needed to give rise to such modes in the respective histograms in the absence of image clusters. For this purpose, a feature extraction algorithm is devised to obtain some domain characteristics by integrating the spatial and spectral information of image points in the local image areas. This integration provides a better clustering in the measurement space than the sensory data alone as demonstrated in the feature extraction phase of the method.

The remainder of this paper is organized as follows. We first give the background information for selecting the color space for the method. Operation of the overall approach is then described by an algorithm. The cluster detection process is explained in detail using a sample image. In the next two sections, a cluster isolation method based on the linear discriminant analysis is given and performance of the feature extraction algorithm is demonstrated. Results of segmenting three low contrast and low resolution pictures are then presented.

## 2. DESCRIPTION OF THE METHOD

### 1. *Background*

A color image is usually given by three values at every pixel, which correspond to the $R$ (red), $G$ (green), and $B$ (blue) tristimuli. In computer processing of color

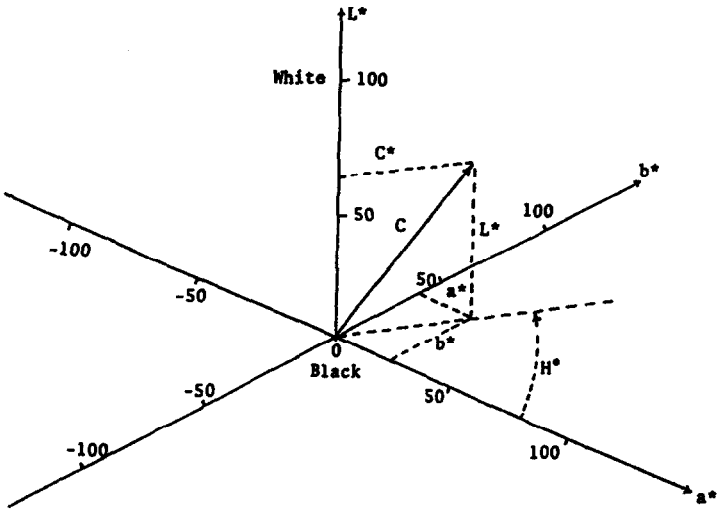


FIG. 1. Organization of the ($L^*$, $a^*$, $b^*$) uniform color coordinate system.

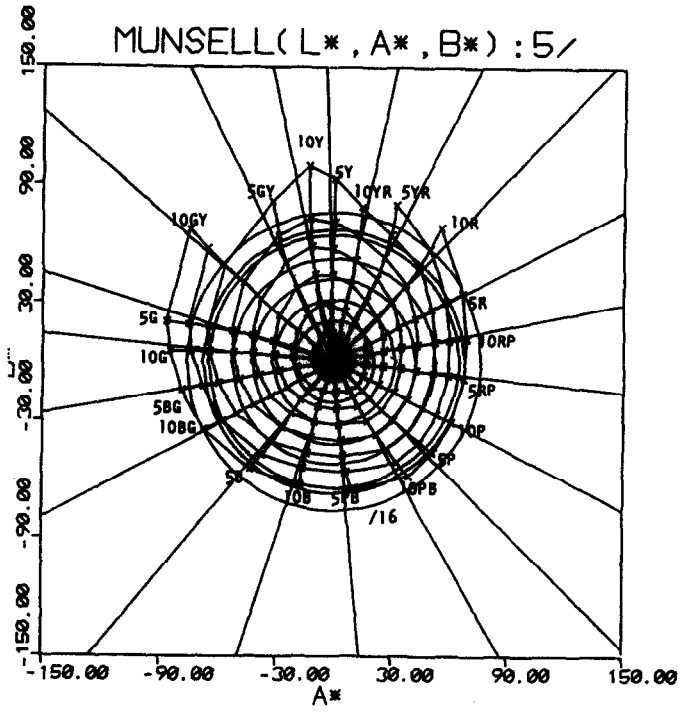


FIG. 2. The loci of constant hue and constant chroma of Munsell system and their approximated networks (radial lines and concentric circles) in the ($a^*$, $b^*$) plane at Munsell value of 5/.

pictures, several features can be calculated from the tristimuli for different purposes [12]. For example, Nevatia [13] concluded by extending the Hueckel operator for color images that the intensity and chromaticity coordinates gave better results than the $R, G, B$ values in detecting the color edges. Ohlander [14] used nine redundant color attributes in multihistogram region segmentation. He reported that the hue component was more useful than the $R, G, B, Y, I, Q$, intensity, and saturation components. Schacter, Davis, and Rosenfeld [9] suggested that the use of uniform color space such as the $(L^*, a^*, b^*)$ could improve the performance of color clustering. Ohta, Kanade, and Sakai [15] derived a set of effective color features, $\{(L^*, a^*, b^*)\}$ and $\{(R + G + B)/3, (R - B), (2G - R - B)/2\}$, by systematic experiments and comparative study in region segmentation. Recently, Tominaga [32] used three color perceptual attributes of the Munsell system for segmenting the color images of photographic prints.

For color clustering, it is desirable that the selected color features define a space possessing uniform characteristics [16, 17, 32]. The $(L^*, a^*, b^*)$ and $(U^*, V^*, W^*)$ color coordinate systems developed by the CIE (Commission Illumination d'Eclairage) in 1976 [18] approximately satisfy this property. It has been shown in [15] that the former space gives better results than the latter in segmenting the color pictures. This system is obtained from the $(R, G, B)$-primary system by converting the $(R, G, B)$ values into the $(X, Y, Z)$-nonphysical primary system [21] and applying a cube-root transformation to the $(X, Y, Z)$ values:

$$X = 2.7690R + 1.7518G + 1.1300B \tag{1.1}$$

$$Y = 1.0000R + 4.5907G + 0.0601B \tag{1.2}$$

$$Z = 0.0000R + 0.0565G + 5.5943B \tag{1.3}$$

$$L^* = 116[Y/Y_0]^{1/3} - 16, \qquad Y/Y_0 > 0.01 \tag{2.1}$$

$$a^* = 500\left[(X/X_0)^{1/3} - (Y/Y_0)^{1/3}\right], \qquad X/X_0 > 0.01 \tag{2.2}$$

$$b^* = 200\left[(Y/Y_0)^{1/3} - (Z/Z_0)^{1/3}\right], \qquad Z/Z_0 > 0.01, \tag{2.3}$$

where $X_0$, $Y_0$, and $Z_0$ are the $(X, Y, Z)$ values of the reference white. Here, they are selected 255 for 8-bit image data representation. Organization of the $(L^*, a^*, b^*)$ cube-root system is shown in Fig. 1. The cylindrical coordinates $(L^*, H^\circ, C^*)$ [22] of this space resemble the empirical Munsell color order system [19] and concur almost exactly with the accepted physiological model of color vision [20]. These coordinates, known as psychometric lightness, hue, and chroma, are given by (see also Fig. 1)

$$L^* = L^* \tag{3.1}$$

$$H^\circ = \tan^{-1}(b^*/a^*) \tag{3.2}$$

$$C^* = (a^{*2} + b^{*2})^{1/2}. \tag{3.3}$$

The loci of constant lightness, hue, and chroma describe the appearance of object colors under certain view and lighting conditions. Any horizontal section through the space would define a plane of constant lightness while any vertical plane originating at the achromatic $L^*$ axis would be a plane of constant hue. A

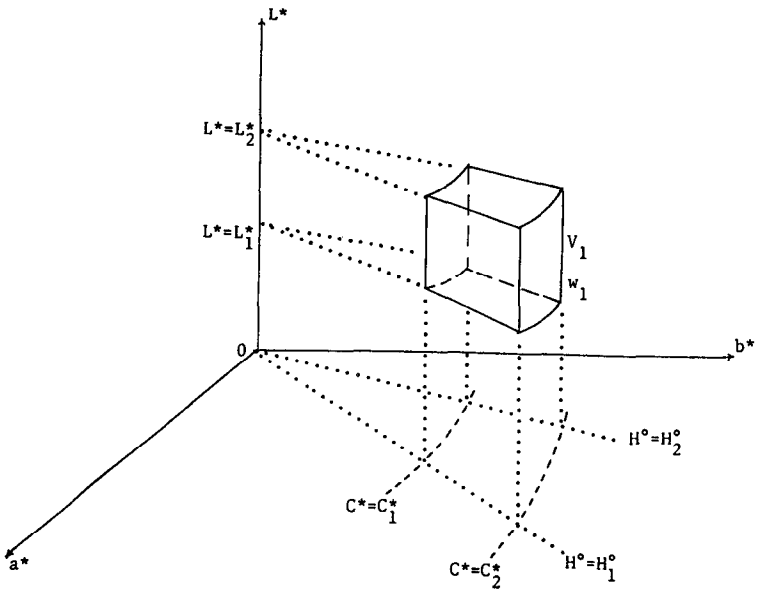


FIG. 3. A circular-cylindrical decision element $V_1$ enclosing an image cluster $w_1$ in the $(L^*, a^*, b^*)$ color space.

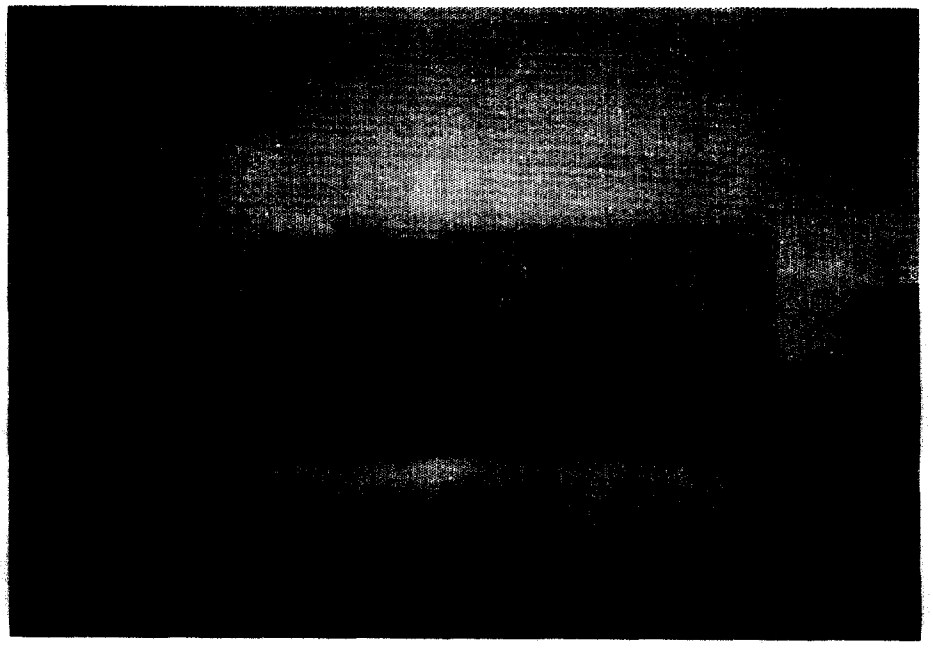

FIG. 4. Photograph of the computer printout (using 32 gray levels) of the intensity value of the test image (picture "11").

cylindrical section concentric with the $L^*$ axis, however, would constitute a surface of constant saturation in which all the object colors have the same purity. In Fig. 2, organization of the constant chromaticity network of the Munsell system are shown for twenty hues (varying from $5R$ to $10RP$) at eight chroma values (increasing from /2 to /16) in the $(a^*, b^*)$-plane for the Munsell value of 5/.

Here this network is explored for estimating the 3D color distributions of image clusters. For this purpose, the radial hue and concentric chroma loci are approximated by radial lines and concentric circles in the least square sense as shown in Fig. 2. They are used to estimate the clusters' distributions in some circular-cylindrical volume elements in the $(L^*, a^*, b^*)$ space. Figure 3 shows such a decision element which is bounded by two lightness planes ($L_1^*$ and $L_2^*$), two hue planes ($H_1°$ and $H_2°$), and two chroma cylinders ($C_1^*$ and $C_2^*$). It is clear from Fig. 3 that

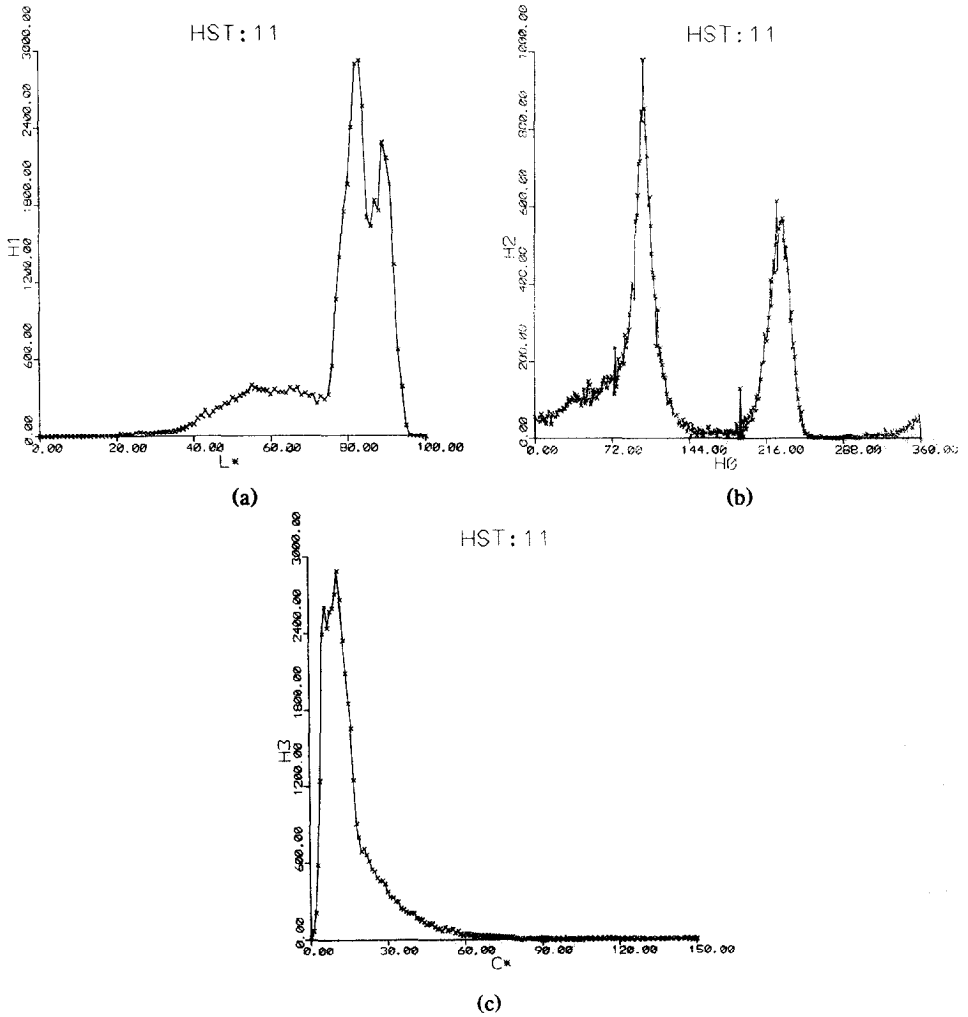


FIG. 5. 1D histograms of the $L^*$, $H°$, and $C^*$ cylindrical color components of the test image given in Fig. 4.

we need to determine the values of six loci to specify such a decision volume. This gives the nonparametric estimates of the clusters' distributions in the 3D color space. The goal here is to reduce the computational cost involved in forming 3D decision surfaces. In addition, use of the 1D histograms for detecting the 3D clusters makes the detection process computationally efficient. This is similar to the process described in [8], which relies on the use of elliptical and paraboloic decision surfaces to enclose image clusters in the *X-Y*, *X-I*, and *Y-I* plane projections of the $(X, Y, I)$ normalized color space. However, use of the chromaticity network results in simpler decision surfaces than those obtained in [8]. Furthermore, here the cluster detection is performed in the 3D color space instead of its lower dimensional subspaces as proposed in [8]. This permits utilization of all the property values of clusters for segmentation, and inherently recognizes their respective cross correlation. This way, the region acceptance is not limited to the information available from one color component.

2. *Description of the Overall Approach*

Operation of the method may be described by the following algorithm:

*Step* 1. Starting from the 1D histograms of the line projections of the space, detect the most prominent cluster (M1) and its neighbor (M2) in some ranges of a particular color component.

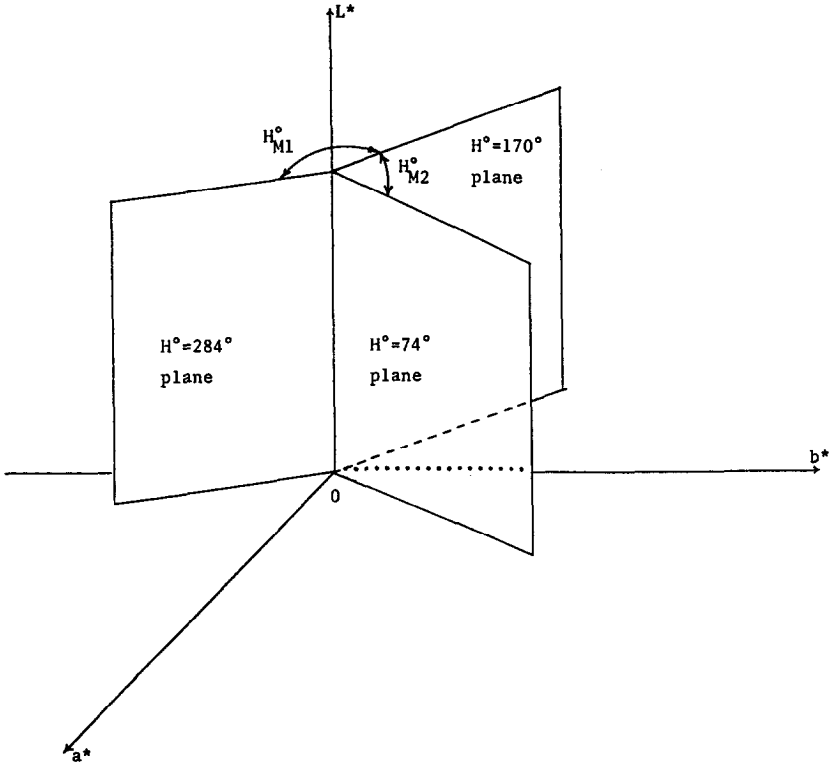


FIG. 6. Angular hue planes describing the hue ranges of the best cluster ($w_1$) and its spectral neighbor ($w_2$) detected for the test image.

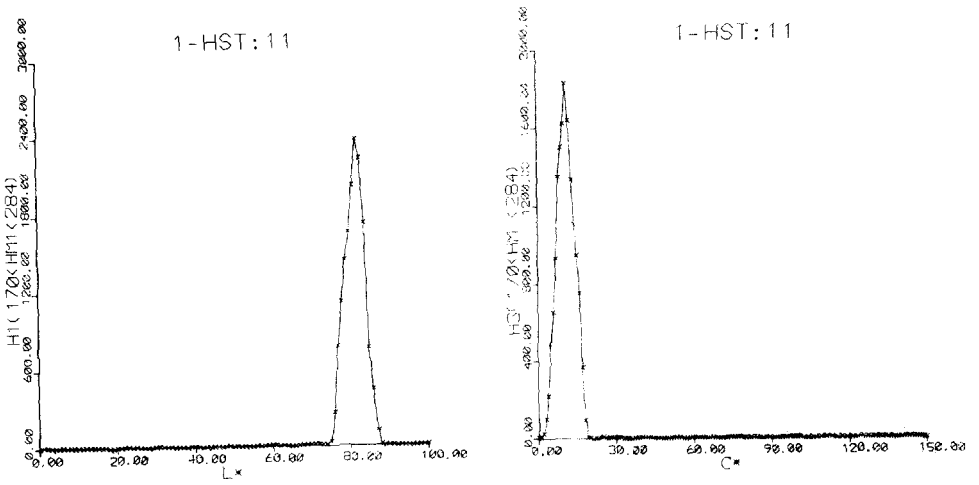


FIG. 7. 1D parametric histograms of the $L^*$ and $C^*$ color components of the test picture constructed for mode 1.

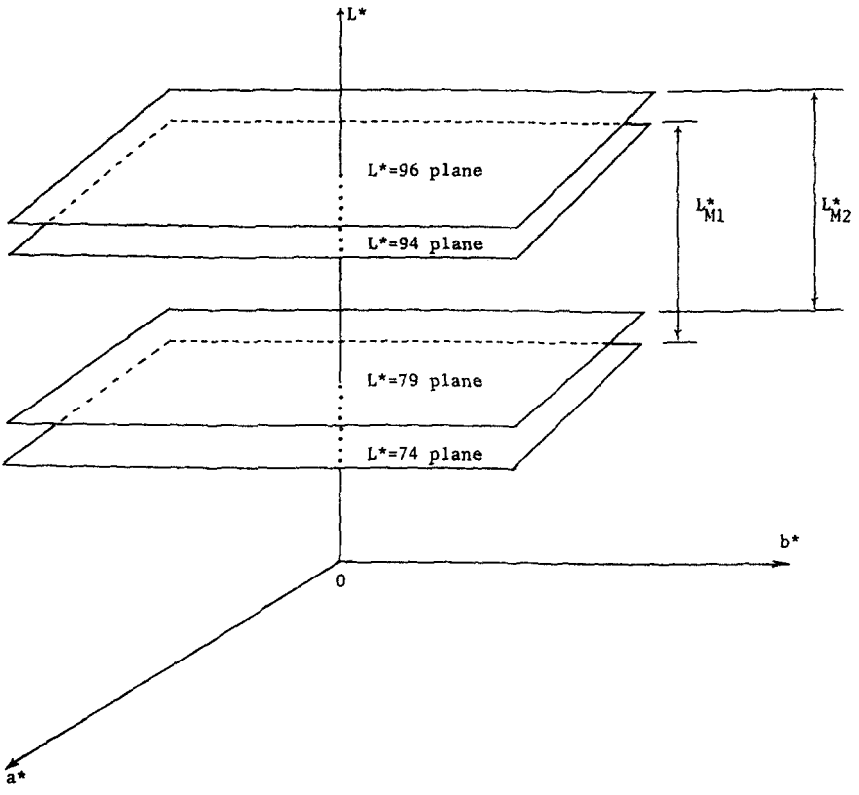


FIG. 8. Horizontal lightness planes describing the lightness ranges of $w_1$ and $w_2$.

*Step* 2. Based on this line information, project only a portion of the space onto the other two color coordinates and using their respective 1D histograms, determine the plane properties of the detected clusters.

*Step* 3. Based on the extracted plane information, project only a part of the space onto the last color coordinate and using the respective 1D histogram, find the space distributions (V1 and V2) of the detected modes.

*Step* 4. Project the estimated color distributions of the best mode and its neighbor onto the line of Fisher discriminant for 1D thresholding.

*Step* 5. Test the modality of 1D histograms of the remaining image points for a cluster. If a decisive 1D peak exists, proceed with Step 1; otherwise, extract an additional feature set and proceed with Step 1.

In the following sections, we first demonstrate the operation of the first three steps of the algorithm for the low resolution picture of Fig. 4. We then present the operation of step 4 for isolating the regions of the detected clusters of the test image.

### 3. *Cluster Detection*

Consider the 1D histograms of the $L^*$, $H°$, and $C^*$ color coordinates of the test image given in Fig. 5. From these distributions, the best decisive peak (M1) and its largest neighbor (M2) are located in the hue histogram. (For peak acceptance

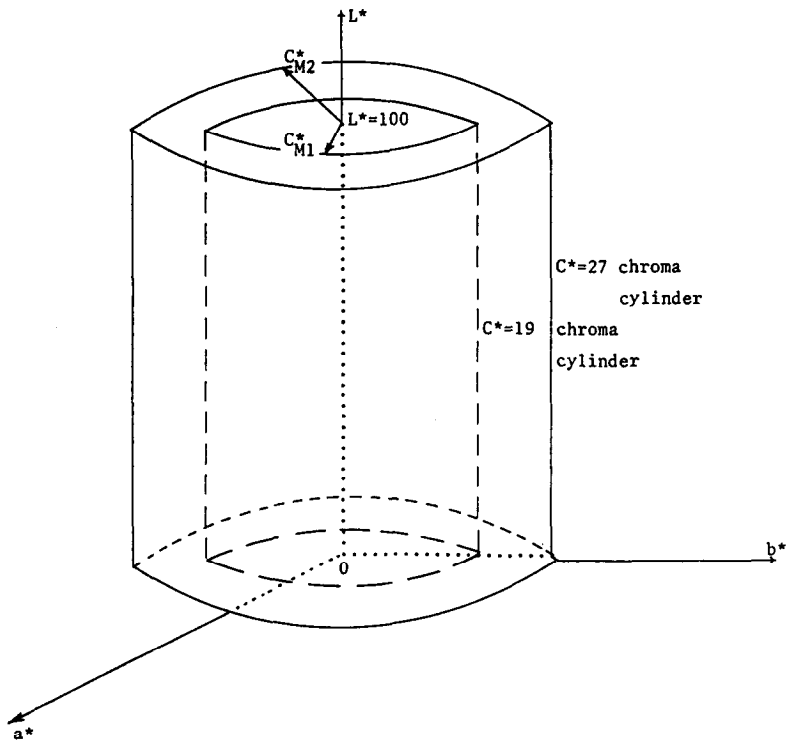


FIG. 9. Circular chroma cylinders corresponding to the chroma ranges of $w_1$ and $w_2$.

procedure see, for example, [30, 31, 34].) In the hue scale, the range of M1 ($H^{\circ}_{M1}$) varies from 170 to 284° and the range of M2 ($H^{\circ}_{M2}$) lies from 74 to 170°. These planes describe the hue variation of the most prominent cluster ($w_1$) and its largest spectral neighbor ($w_2$) as illustrated in Fig. 6.

The space distribution of $w_1$ is obtained using the 1D histograms of $L^*$ and $C^*$ color components of the image pixels whose hue values are in the hue range of the mode M1. These distributions for the test image are given in Fig. 7. Each possesses only one decisive peak. The lightness and chroma ranges ($L^*_{M1}$ and $C^*_{M1}$) of these peaks are given by $74 < L^* < 94$ and $0 < C^* < 19$. The lightness values 74 and 94 define two constant horizontal lightness planes in the space and the chroma value 19 specifies a constant circular cylinder around the lightness axis. These loci describe the lightness and chroma variation of $w_1$ as depicted in Figs. 8 and 9. It is clear that the respective loci of the constants (170°, 284°), (74, 94), and (0, 19) represent a circular-cylindrical decision volume $V_1$ in the ($L^*, a^*, b^*$) space (see Fig. 10) as the estimated color distribution of $w_1$.

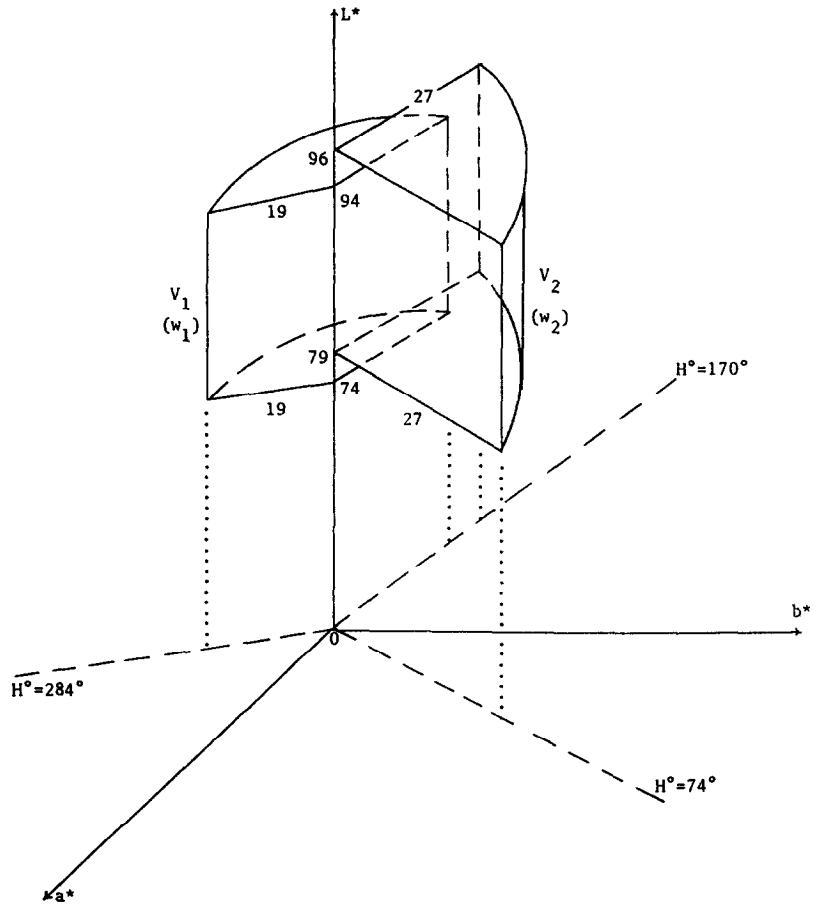


FIG. 10. The estimated color distributions ($V_1$ and $V_2$) of $w_1$ and $w_2$ in the ($L^*, a^*, b^*$) space.

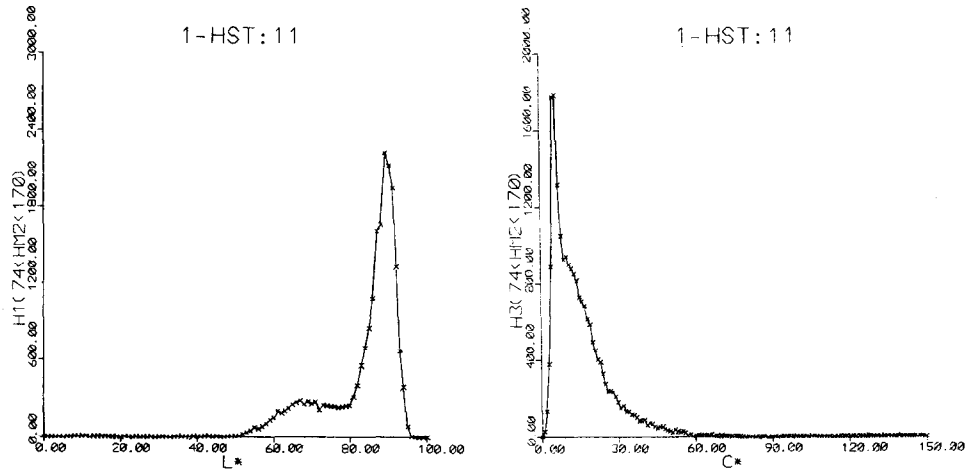


FIG. 11. 1D parametric histograms of the $L^*$ and $C^*$ color coordinates of the test image derived for mode 2.

The process described above is then repeated for estimating the color distribution of $w_2$. Figure 11 shows the 1D histograms of $L^*$ and $C^*$ color components of the image points whose hue values lie within the hue range of $w_2$. From these two distributions, the lightness and chroma ranges ($L^*_{M2}$ and $C^*_{M2}$) of $w_2$ are given by $79 < L^* < 96$ and $0 < C^* < 27$. The lightness values 79 and 96 define two horizontal planes in the space and the chroma value 27 describes an eccentric cylinder around the achromatic axis. These surfaces are also plotted in Figs. 8 and 9. The respective loci of the constants (74°, 170°), (79, 96), and (0, 27) also characterize a volume element $V_2$ in the space as the estimated distribution of $w_2$. This volume is also illustrated in Fig. 10 by showing its position with respect to that of $w_1$.

It is clear from Fig. 10 that the estimated distributions ($V_1$ and $V_2$) of $w_1$ and $w_2$ have significant overlap along the $L^*$ axis due to the lightness ranges. In order to minimize the effect of this overlap in region isolation, we describe a pixel classification technique in the next section based on the linear discriminant analysis.

## 4. *Cluster Isolation and Region Extraction*

In this step of the algorithm, the most prominent cluster $w_1$ whose color distribution was estimated by the volume element $V_1$ is isolated from its spectral neighbor $w_2$ using all the information gathered from the 3D color space. This isolation process is carried out using the Fisher linear discriminant function [23]. A similar study was also made by Ahuja *et al.* [24] and by Chen [25]. In [24], the Fisher criterion was employed for pixel classification using the gray levels in the neighborhood of a point as features. In [25], it was used to extract a target from the background in infrared and reconnaissance images. Here, this technique is explored for isolating the detected color clusters of natural scenes. This isolation is performed as a two-at-a-time class separation process as follows.

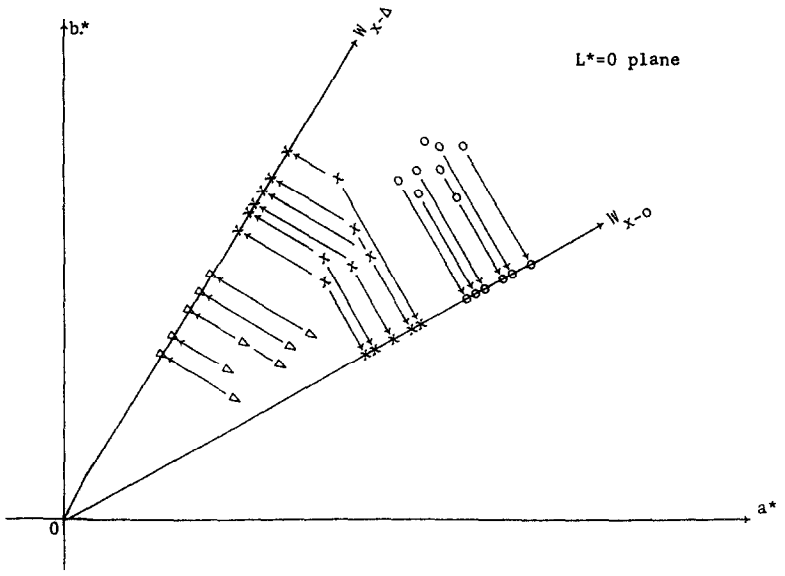


FIG. 12. Two-at-a-time class projection process for separating the class "×" from its neighbors "○" and "△" in the $(a^*, b^*)$ plane.

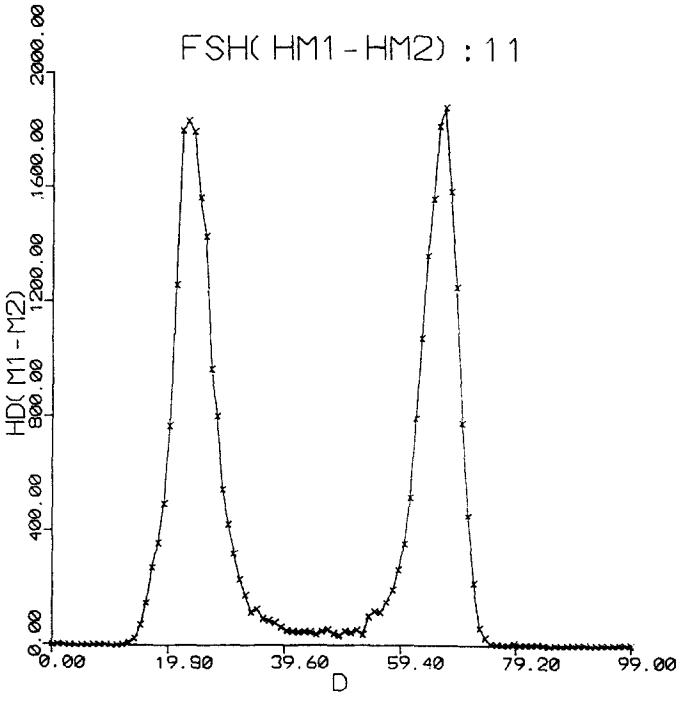


FIG. 13. 1D histogram of the projected values of color vectors belonging to $w_1$ and $w_2$, detected for extracting the image region "11:SKY."

The color vectors of $w_1$ and $w_2$ are projected onto a line so that the projected color points are well clustered and separated for 1D thresholding. Figure 12 illustrates this process in the $(a^*, b^*)$-chromaticity plane of the uniform color space. The line used for each projection is determined according to the Fisher criterion. For the two class problem, the line of projection is given by

$$W = [S_1 + S_2]^{-1} \cdot (M_1 - M_2), \tag{4}$$

where $(S_1, S_2)$ and $(M_1, M_2)$ are the covariance matrices and the mean vectors of classes 1 and 2, respectively. $W$ defines a direction in the feature space, which is independent of the class distribution functions. This means that, once the image

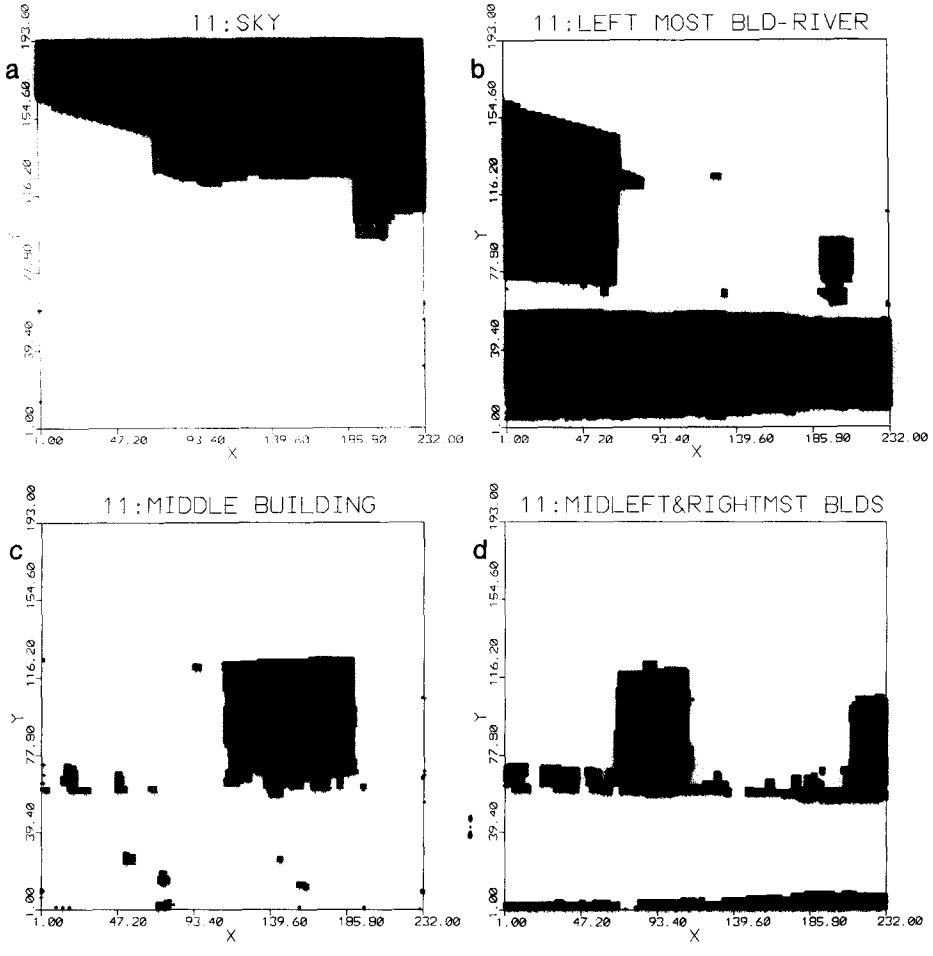


FIG. 14. Image regions resulting from the segmentation of the test image (picture "11"): (a) Image region "11:SKY"; (b) Image region "11:BLRI" (leftmost building and river) extracted by discarding the points of the detected region "11:SKY"; (c) Image region "11:APM" (midright building) extracted by excluding the pixels of regions in (a) and (b); (d) Image region "11:APS" (midleft and rightmost buildings) detected by discarding the points of regions in (a), (b), and (c).

clusters are detected in the feature space, there is no need for detailed statistical knowledge of their pattern vectors in order to determine $W$.

For the detected clusters of the test image, the line of projection is computed as $W = [-0.07407, 0.32801, -0.20976]^T$, where $T$ denotes the transpose of the row matrix. The color vectors of the image points, which are the elements of $w_1$ and $w_2$ (see Fig. 10), are then projected onto this line using

$$d(C) = W^T \cdot C. \tag{5}$$

Here, $d$ is the linear discriminant function and $C$ is the color vector of the image pixels belonging to the decision volumes $V_1$ or $V_2$.

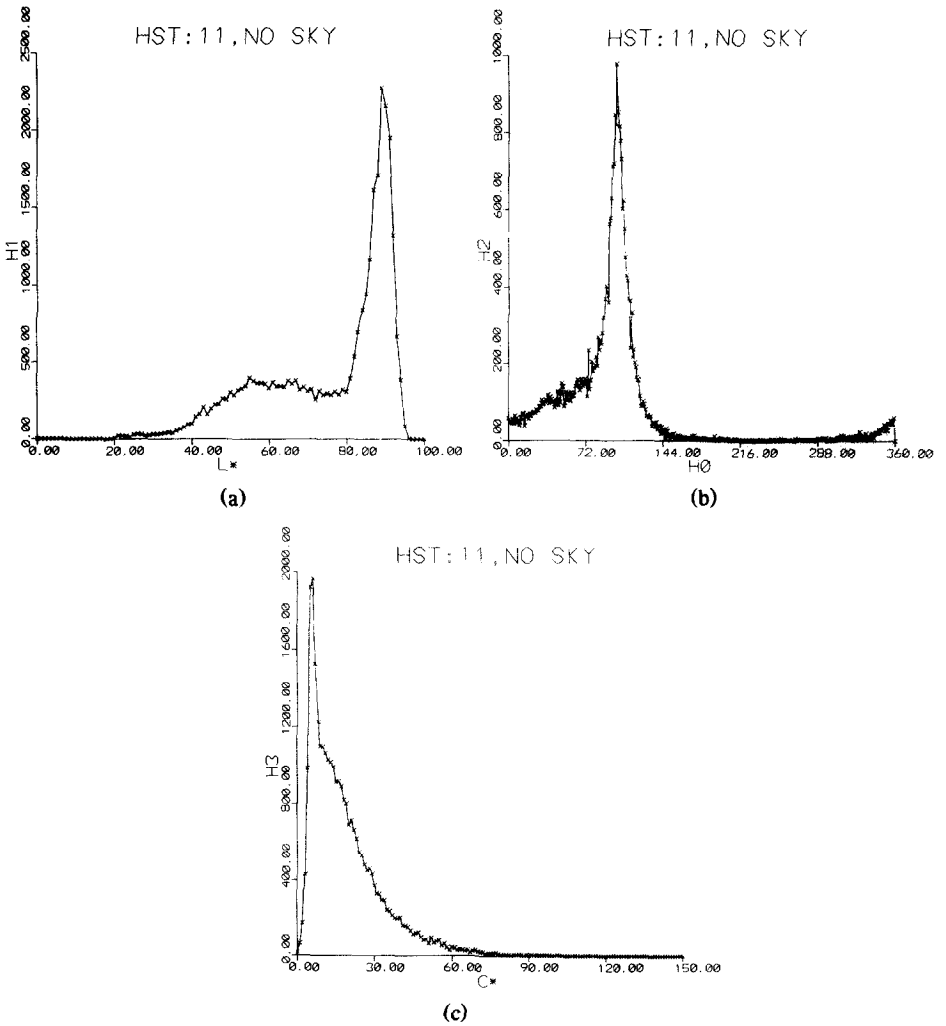


FIG. 15. 1D histograms of the $L^*$, $H°$, and $C^*$ color coordinates of the test image, constructed by excluding the points belonging to the sky.

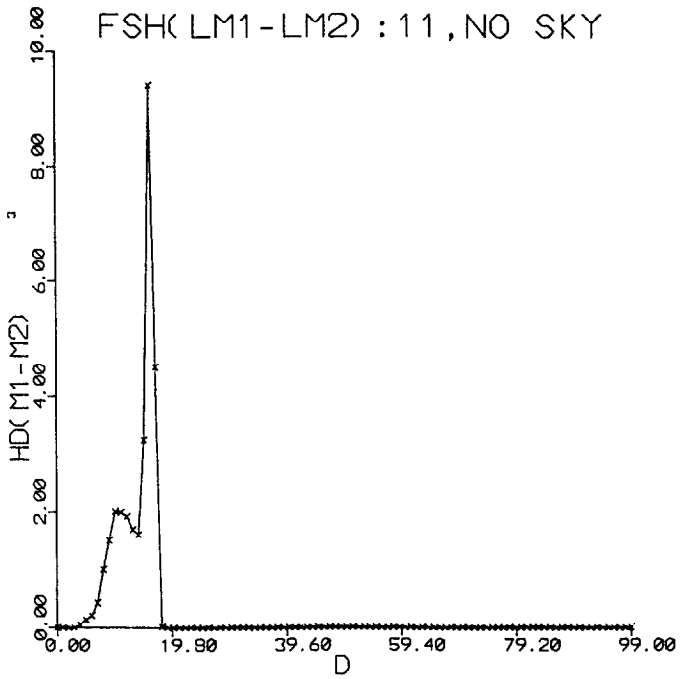


FIG. 16. 1D histogram of the projected values of color vectors belonging to the best mode and its neighbor, detected for extracting the image region "11:BLRI."

The 1D histogram of the projected data points is given in Fig. 13. The projected values of the class mean vectors are also computed as $d_{M1} = 65.775$ and $d_{M2} = 25.295$. Comparing the value of $d_{M1}$ with the valley between the histogram peaks, a threshold is set in this histogram as $d > 42$ to create a binary picture for the image region of $w_1$. This segment is further refined in the spatial domain using a set of binary smoothing templates [26]. The final result is shown in Fig. 14a, which corresponds to the sky.

The above detection and isolation algorithms are then applied to the remaining image points which yield the 1D histograms of Fig. 15. Two peaks in the lightness distribution are selected as the best mode ($M_1$) and its neighbor ($M_2$) which lie in the lightness range of $80 < L^* < 100$ and $0 < L^* < 80$. Projecting the image clusters due to these two peaks onto $W = [0.18654, -0.00102, 0.01216]^T$ results in a 1D histogram (Fig. 16) possessing two peaks. By setting a threshold at 14.5 of the histogram variable, a binary picture is created and further refined in the spatial domain using the smoothing templates. The resultant regions, called "11:BLRI," are shown in Fig. 14b, which correspond to the leftmost building and river.

Since the 1D histograms (see Fig. 17) of the remaining part of the test image do not exhibit any decisive peak, an additional feature set is needed to give rise to 1D clusters in the image histograms. For this, a statistical-structural feature extraction algorithm is devised to obtain some domain characteristics by combining the spatial and spectral information of image points in the local image areas. This

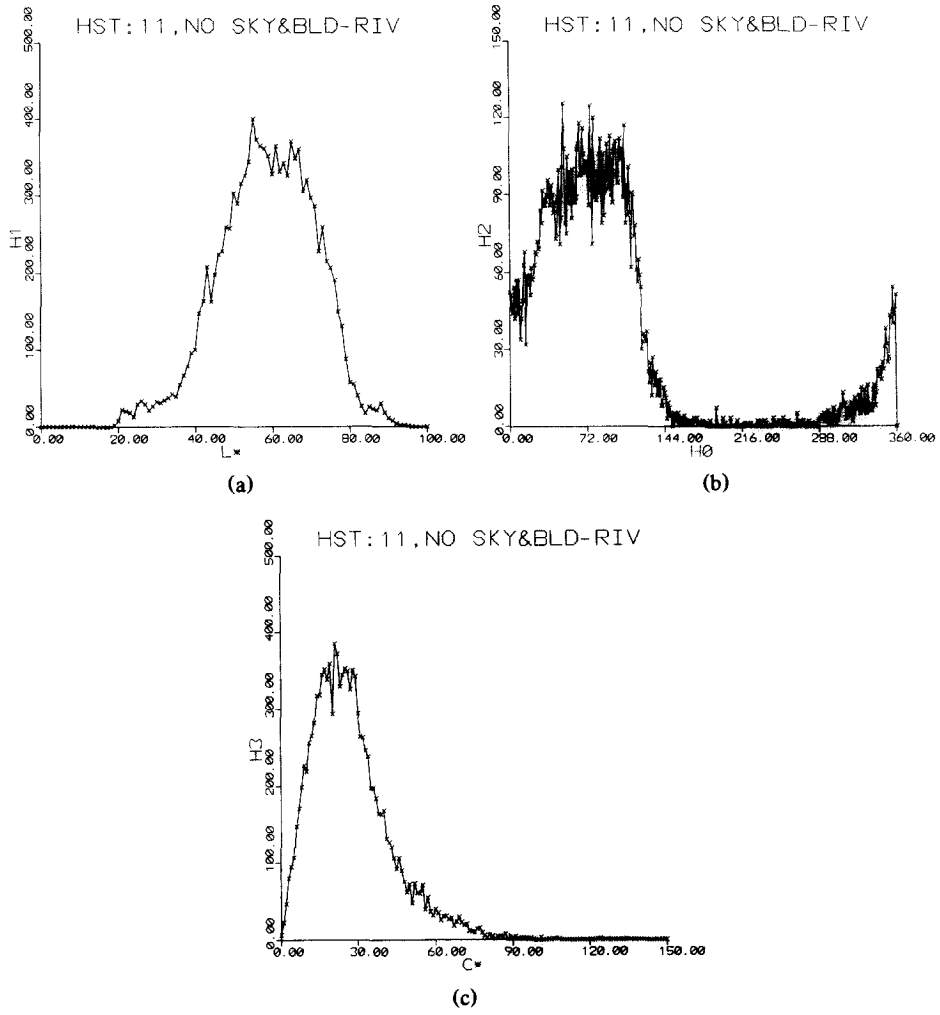


FIG. 17. 1D histograms of the $L^*$, $H°$, and $C^*$ color components of the test image, derived by discarding the points of the detected regions "11:SKY" and "11:BLRI."

integration provides better clustering in the measurement space than the sensory data alone as demonstrated in the following section.

## 5. *Feature Extraction Algorithm*

In this phase of the method, the remaining part of the test image is first partitioned into its atomic regions of maximally 4-connected pixels of the similar color. For this partitioning, every unprocessed pixel is initially assumed to be an atomic region by assigning it a region number. Starting from the first row and first column, the color vector of each unclassified point is compared with that of its left and top neighbors. If a similarity between the pixel in process and any of its specified neighbors in a $3 \times 3$ local window is detected, then the region number of

that element is modified to that of its similar neighbor. Here the similarity between an image pixel $(i, j)$ and its left or top neighbor $(m, n)$ is defined by

$$|L^*(i,j) - L^*(m,n)| < 10 \tag{6.1}$$

$$|H^\circ(i,j) - H^\circ(m,n)| < 10^\circ \tag{6.2}$$

$$|H^\circ(i,j) - H^\circ(m,n)| > 350^\circ \tag{6.3}$$

$$|C^*(i,j) - C^*(m,n)| < 10. \tag{6.4}$$

The outlined region growing technique is repeated at every unprocessed point of the test image. The total number of atomic regions created for the remaining points of the test picture is 2171.

After this initial partitioning, the algorithm computes the center of gravity of every atomic region $a_k$ using

$$i_k = \left\{ \sum_i \sum_j |C(i,j)| \cdot i \right\} \Big/ \left\{ \sum_i \sum_j |C(i,j)| \right\} \tag{7.1}$$

$$j_k = \left\{ \sum_i \sum_j |C(i,j)| \cdot j \right\} \Big/ \left\{ \sum_i \sum_j |C(i,j)| \right\}, \tag{7.2}$$

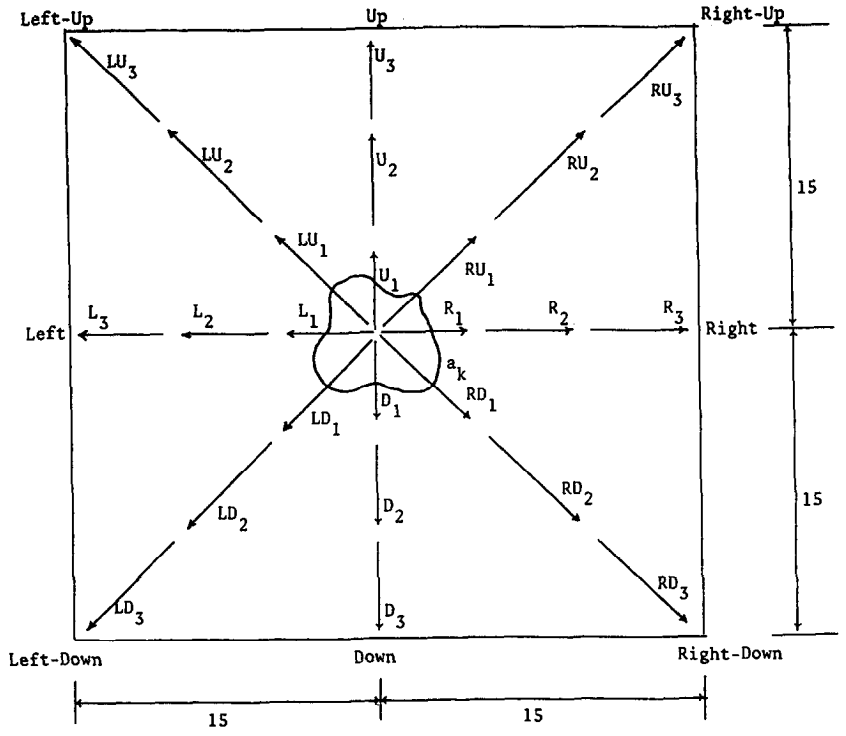


FIG. 18. Eight fixed directions defined in the 30 × 30 local window located at the center of gravity of an atomic region $a_k$.

where $|C(i, j)|$ is the magnitude of the color vector of a point $(i, j)$ in $a_k$. A $30 \times 30$ local window is then located at the center of the atomic region in process. Within this window, eight fixed directions are defined from the center of that atomic region (see Fig. 18). On each of these directions, similarity between the atomic region in question and its neighbors is tested according to the similarity criteria given by the Eqs. (6.1)–(6.4). If the continuity property is observed along these directions, the atomic region of interest is classified as being a part of a uniform region and its feature vector is computed in the corresponding direction(s). If the continuity search fails, then a possible periodic texture structure is sought to classify the atomic region as being a spectral primitive of a texture field and to compute its feature vector in the respective directions. In the absence of continuity

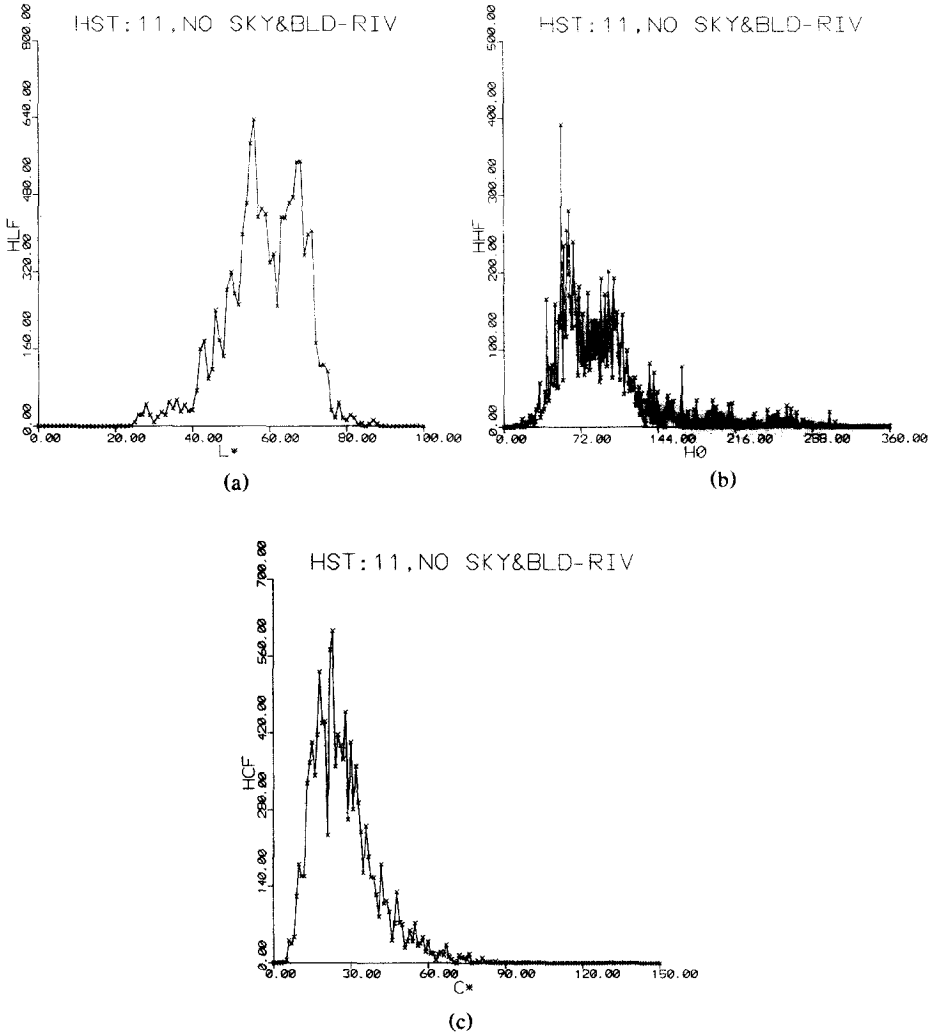


FIG. 19. 1D histograms of the $L^*$, $H°$, and $C^*$ components of the feature vector $f$, extracted for the remaining points of the test image.

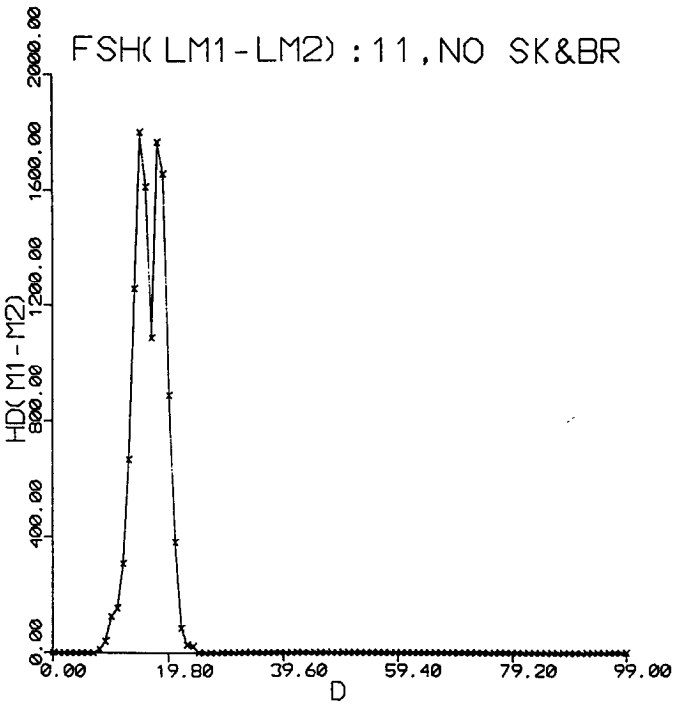


FIG. 20. 1D histogram of the projected values of feature vectors belonging to the modes detected for extracting the region "11:APM."

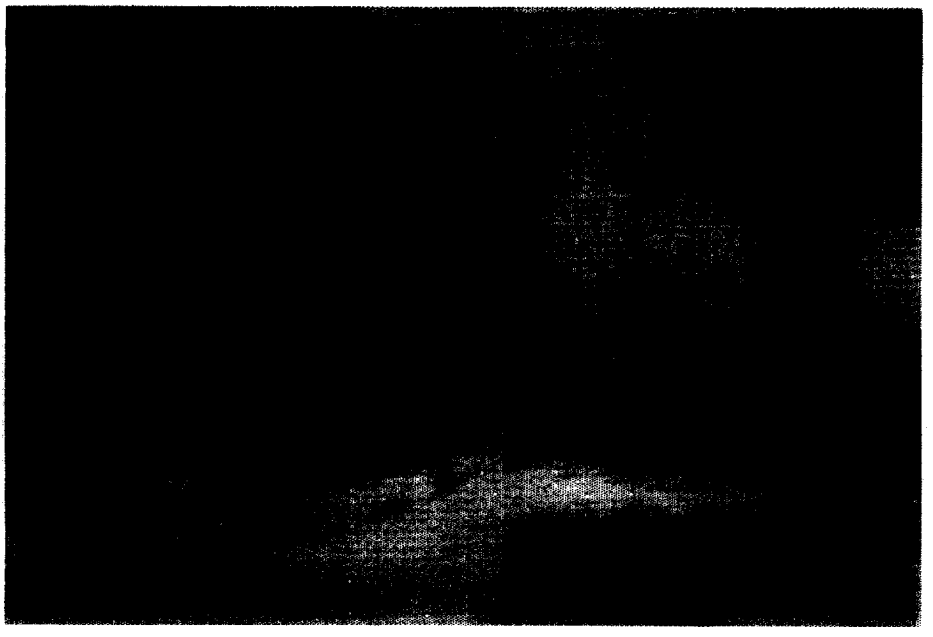

FIG. 21. Photograph of the computer printout (using 32 gray levels on a logarithmic scale) of the intensity value of picture "9."

and texture patterns, the atomic region is accepted as an isolated one and its feature assignment is made considering all of its touching neighbors.

Once an atomic region is classified as being a part of a visual pattern through the specified directions, then the feature vector $f$ is assigned to that atomic region by averaging the color vectors of the atomic regions in that visual pattern as described below:

$$f = E\{C_k \cup C_l\}. \tag{8}$$

Here, $C_k$ and $C_l$ are the color vectors of the atomic region $a_k$ and its neighbors in the determined directions, $E$ is the expected value operator, and $\cup$ represents the union of color vectors.

The above feature extraction technique yields a feature set $f$ for the remaining points of the test image that provides the 1D histograms of Fig. 19. Comparing the

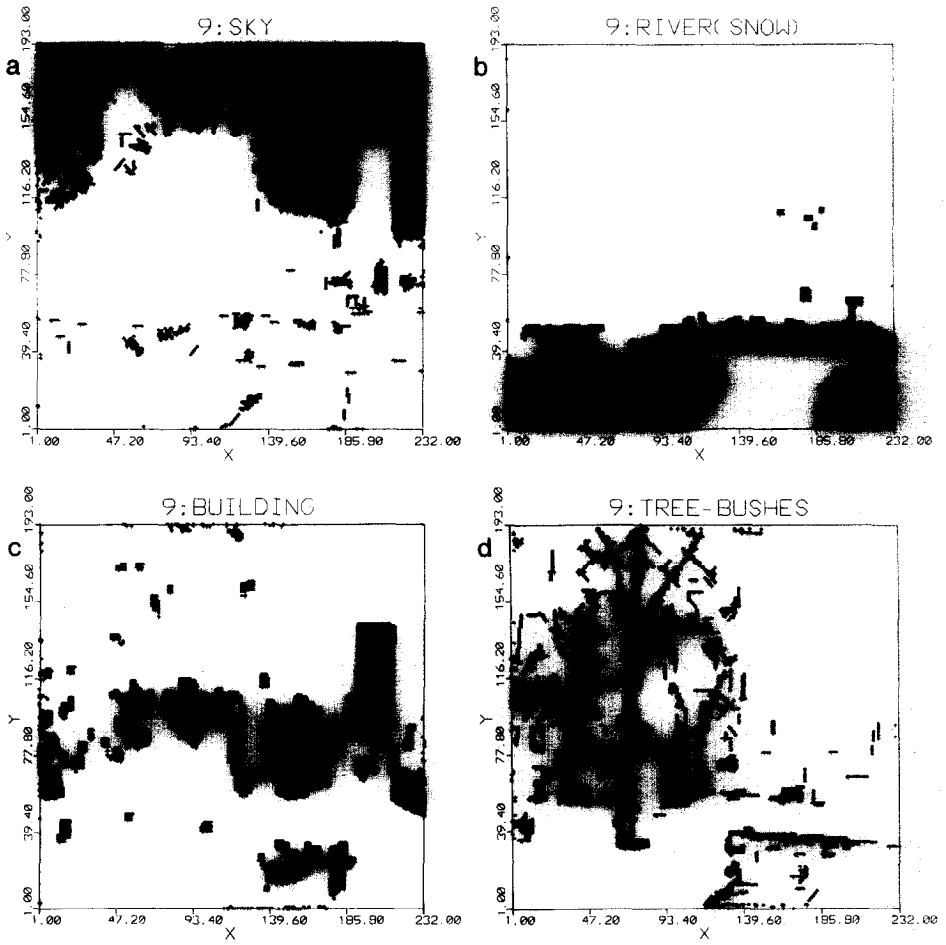


FIG. 22. Segmentation results of picture "9": (a) Image region "9:SKY" extracted for the sky; (b) Image region "9:RIV" corresponding to the river or snow; (c) Image region "9:BL" representing the building block; (d) Image region "9:TRE" obtained for the tree and bushes.

$L^*$ histogram of Fig. 19 with that of Fig. 17, it can be noticed that the $L^*$ component of $f$ gives a better clustering property than that of the sensory data. The former histogram possesses two 1D clusters in contrast to the latter. These decisive 1D peaks are defined in the lightness scale as $62 < L^* < 100$ and $0 < L^* < 62$. Projecting the 3D clusters due to these two peaks onto the line $W = [0.25864, 0.01177, 0.02304]^T$ results in the histogram of Fig. 20. By setting a threshold at 17.5, the resultant region, called "11:APM," is obtained as shown in Fig. 14c. This represents the building at the center of the test image.

The remaining points of the test image are accepted as a uniform region with respect to the extracted feature vector $f$. After the spatial refinement, the resultant regions, called "11:APS," are obtained as shown in Fig. 14d. They are the midleft and rightmost buildings in the scene.

## 3. COMPUTER IMPLEMENTATION AND SEGMENTATION RESULTS

The multiple histogram-based clustering technique presented in this paper was applied to several low contrast color images of the skyline of the city of Pittsburgh taken in winter. The spectral contents of the images are limited mainly to the bright colors. This property is observed mostly on the lightness and chroma distributions, which are concentrated in the lower values of chroma and the higher values of lightness (usually $C^* < 30$ and $L^* > 60$), respectively. The original images are given with the eight bits/pixel $R, G, B$ specification in $580 \times 700$ grid. In order to fit the image data to the computer system (DEC PDP-10) used in the implementation, the pictures were reduced to $193 \times 232$ grid by a factor of 3 in each dimension. This reduction significantly affected the spatial resolution of the

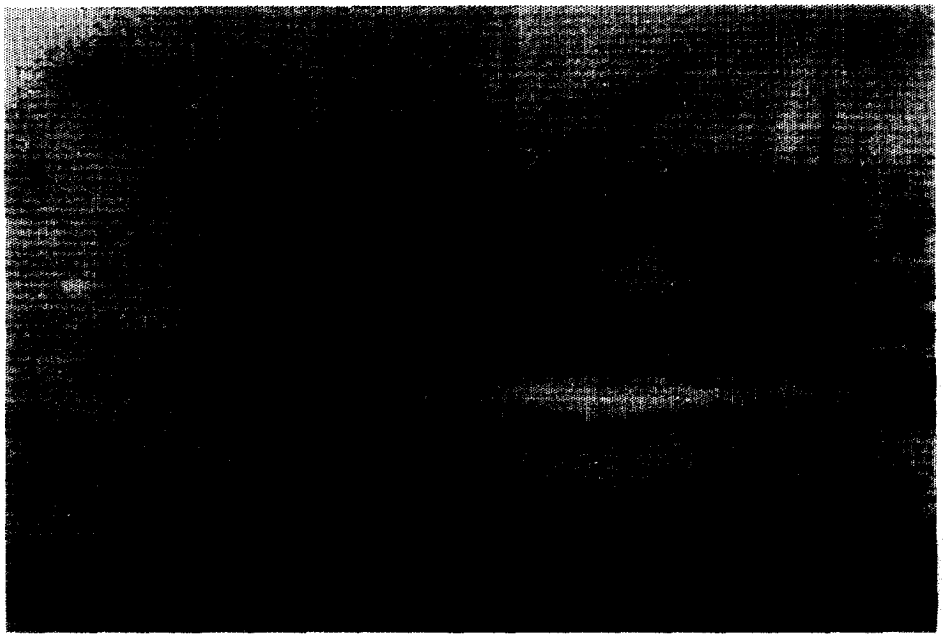

FIG. 23. Photograph of the computer printout (using 32 gray levels) of the intensity file of pic ture "7."

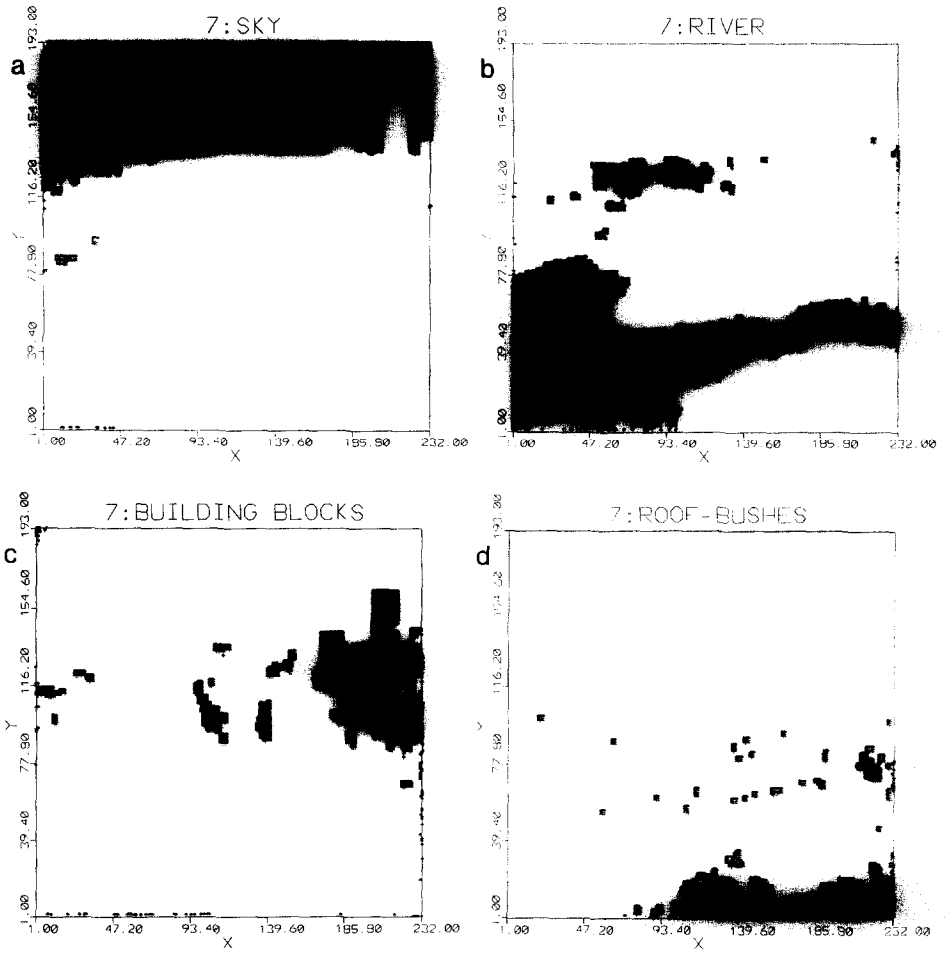


FIG. 24. Segmentation results of picture "7": (a) Image region "7:SKY" corresponding to the sky; (b) Image region "7:RIV" obtained for the river; (c) Image region "7:BLK" representing the building blocks; (d) Image region "7:RFBS" extracted for the roof and bushes.

resultant images. For example, the textured buildings located in the middle of the test image shown in Fig. 4, almost became a wide building block. Although the separation of these two buildings along a straight line is almost impossible by an observer, the algorithm gives a reasonably good segmentation result by locating this straight line between the buildings.

Four images were segmented in the computer implementation of the method with varying degrees of success [27–29]. These pictures are shown in Figs. 4, 21, 23, and 25. All the visually distinct areas of the test image given in Fig. 4 were successfully extracted as described earlier. Some of the processing results of the other poor resolution images, called picture "9," picture "7," and picture "10," are presented in Figs. 22, 24, and 26.

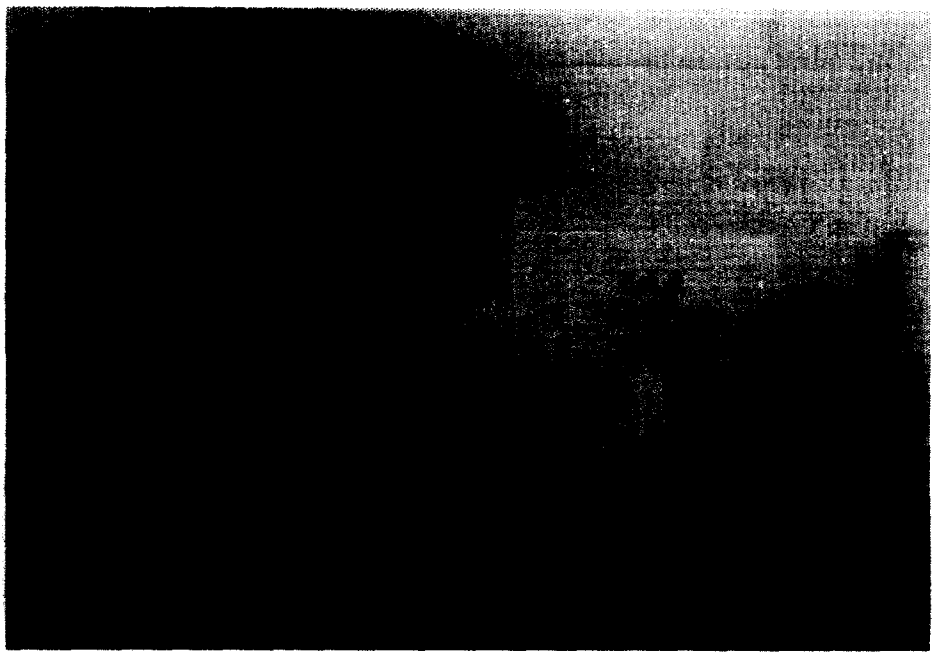

FIG. 25. Photograph of the computer printout (using 32 gray levels) of the intensity file of picture "10."

## 4. DISCUSSION AND CONCLUSION

In this paper, a clustering technique has been presented and its use in color picture segmentation has been described. The proposed method operates in the $(L^*, a^*, b^*)$ cube-root color coordinate system using the cylindrical coordinates. It detects the image clusters by means of estimating their distributions in some well-defined decision volumes of the constant lightness and constant chromaticity network. Use of the 1D image histograms in this estimation process for detecting 3D color clusters reduces the computational cost involved in maintaining many derived images and obtaining their 3D distributions.

For region isolation, the estimated color clusters are projected onto a line for 1D thresholding using the information obtained from the 3D feature space. This permits utilization of all the property values of clusters for segmentation and inherently recognizes their respective cross correlation. The line used for each projection is determined in accordance with the Fisher criterion to minimize the error rate associated with region acceptance. In these respects, the proposed algorithm also differs from the multidimensional histogram thresholding techniques. For example, this method first thresholds an image cluster in all the dimensions of a color space. It then extracts the corresponding image region based on the information obtained from the 3D color space. This way, the decision is not limited to the information available from one color component. To see the effect of this process, a 1D histogram thresholding technique is applied to the peak determined within 170 to 284° range of the hue histogram given in Fig. 5. By setting the threshold values at the specified hue limits of this peak, we have

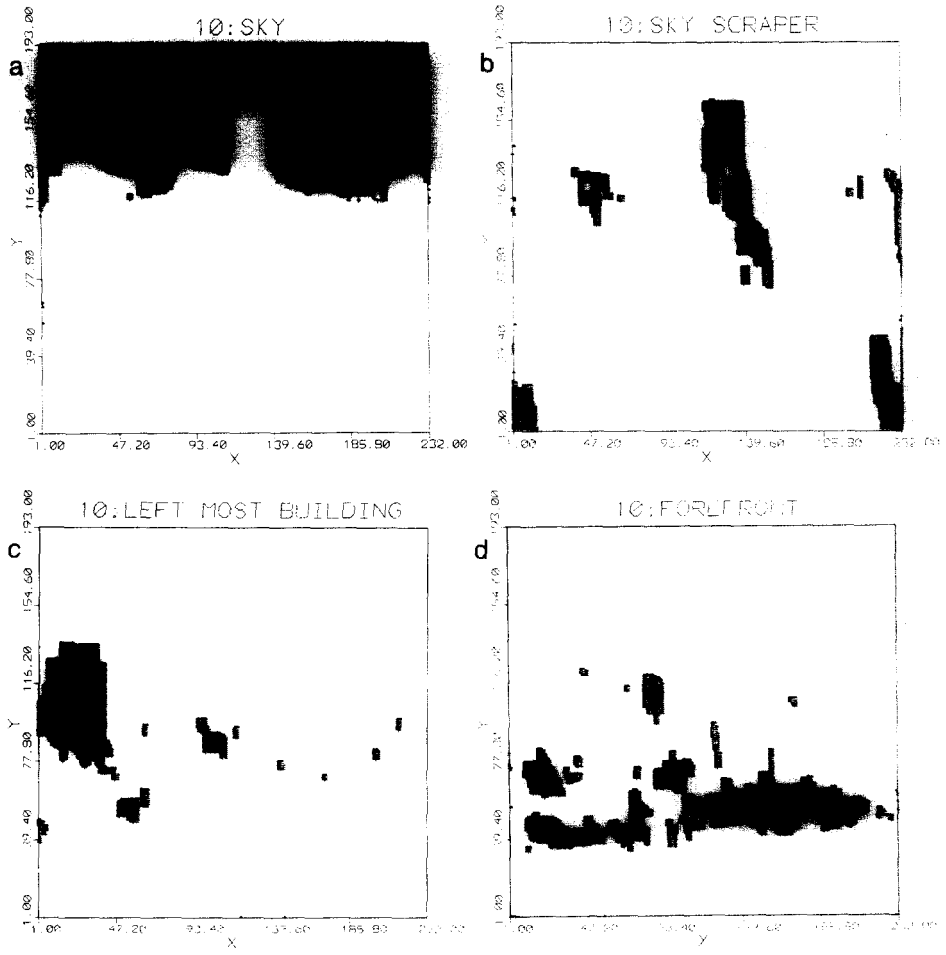


FIG. 26. Segmentation results of picture "10": (a) Image region "10:SKY" corresponding to the sky; (b) Image region "10:SKY SCRAPER" extracted for the sky scraper; (c) Image region "10:LEFTMOST BUILDING" representing the leftmost building; (d) Image region "10:FOREFRONT" obtained for the forefront.

obtained the image region in the spatial domain. Boundaries of the resultant region (sky) have severely deformed and many fine grain image segments have resulted along the boundary. The proposed technique, however, gave reasonably good results by locating the straight line to separate the sky from the buildings in the image.

The procedure described here does not use any a priori information about the scene domain or impose any constraints on the clusters' distributions. It also extracts the features most useful for the picture being processed. With these properties, the technique is not affected by substantial variations of input scenes. The only limitation the procedure appears to have, which is partially due to the limitations imposed by the computer used for the implementation presented here, is the rather lengthy computational time requirement (approximately 40 min) for

extracting the features. On the other hand, the time for extracting image clusters is relatively short, varying from 2 to 9 min depending on the number of regions detected previously. Much of the computational time may be attributed to the language used (FORTRAN).

Further research is needed on the following issues: The link between the presented method and the sensor characteristics and their physical models requires further research. The described method is overly dependent on histogram analysis. Its operation needs to be improved for segmenting scenes involving textures, shadows, inconsistent color appearancy, etc. Operation of the method should be tested on oddly shaped clusters (e.g., linear, long, elongated, etc.). To reduce the overall computation time, the array processor implementation of the method needs further investigation.

## APPENDIX: SYMBOLS

| | |
|---|---|
| $R, G, B$ | red, green, blue tristimuli |
| $X, Y, Z$ | CIE nonphysical primary system |
| $L^*, a^*, b^*$ | cube-root uniform color coordinate system |
| $L^*, H^\circ, C^*$ | metric lightness, hue angle and chroma |
| $L_i^*$ | a constant lightness plane |
| $H_i^\circ$ | a constant angular hue plane |
| $C_i^*$ | a constant circular chroma cylinder |
| M1 and M2 | histogram mode (peak) 1 and 2 |
| $L_{Mi}^*$ | lightness range of a peak in $L^*$ histogram |
| $H_{Mi}^\circ$ | hue range of a peak in $H^\circ$ histogram |
| $C_{Mi}^*$ | chroma range of a peak in $C^*$ histogram |
| $w_i$ | a color cluster |
| $V_i$ | a circular-cylindrical decision volume |
| $W$ | line of projection |
| $S_i$ | class covariance matrix |
| $M_i$ | class mean vector |
| $d(C)$ | Fisher linear discriminant function |
| $C$ | color vector |
| $a_k$ | an atomic region |
| $i_k$ and $j_k$ | center of gravity of the atomic region $a_k$ |
| $\|C\|$ | magnitude of the color vector $C$ |
| $C_k$ | average color vector of the atomic region $a_k$ |
| $f$ | extracted feature vector of an atomic region. |

## REFERENCES

1. E. M. Riseman and M. A. Arbib, Computational techniques in the visual segmentation of static scenes, *Comput. Graphics Image Process.* **6**, 1977, 221–276.
2. K. S. Fu and J. K. Mui, A survey on image segmentation, *Pattern Recognit.* **13**, 1981, 3–16.
3. R. M. Haralick and L. G. Shapiro, Image segmentation techniques, *Comput. Vision Graphics Image Process.* **29**, 1985, 100–132.
4. R. Ohlander, K. Price, and D. R. Reddy, Picture segmentation using a recursive region splitting method, *Comput. Graphics Image Process.* **8**, 1978, 313–333.
5. J. M. Tenenbaum, T. D. Garvey, S. Weyl, and H. C. Wolf, *An Interactive Facility for Scene Analysis Research*, Technical Note 87, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, CA, 1974.

6. S. A. Underwood and J. K. Aggarwal, Interactive computer analysis of aerial color infrared photographs, *Comput. Graphics Image Process.* **6**, 1977, 1–24.
7. M. Ali and J. K. Aggarwal, Automatic interpretation of infrared aerial color photographs of citrus orchards having infestations of insect pests and diseases, *IEEE Trans. Geosci. Electron.* **GE-15**, No. 3, 1977, 170–179.
8. M. Ali, W. N. Martin, and J. K. Aggarwal, Color-based computer analysis of aerial photographs, *Comput. Grapahics Image Process.* **9**, 1979, 282–293.
9. B. J. Schacter, L. S. Davis, and A. Rosenfeld, *Scene Segmentation by Cluster Detection in Color Space*, Technical Report 424, Computer Science Center, University of Maryland, 1975.
10. A. Sarabi and J. K. Aggarwal, Segmentation of chromatic images, *Pattern Recognit.* **13**, 1981, 417–427.
11. R. Gershon, Aspects of perception and computation in color vision, *Comput. Vision Graphics Image Process.* **32**, 1985, 244–277.
12. T. Kanade, Region segmentation: Signal vs semantics, *Comput. Graphics Image Process.* **13**, 1980, 279–297.
13. R. Nevatia, A color edge detector and its use in scene segmentation, *IEEE Trans. Systems Man Cybernet.* **SMC-7** No. 11, 1977, 820–826.
14. R. B. Ohlander, *Analysis of Natural Scenes*, Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA, 1975.
15. Y. Ohta, T. Kanade, and T. Sakai, Color information for region segmentation, *Comput. Graphics Image Process.* **13**, 1980, 222–241.
16. F. Bumbaca and K. C. Smith, Design and implementation of colour vision model for computer vision applications, *Comput. Vision Graphics Image Process.* **39**, 1987, 226–245.
17. G. B. Coleman, *Image Segmentation by Clustering*, Ph.D. thesis, USCIPI Report 750, Image Processing Institute, University of Southern California, CA, 1977.
18. A. L. Robertson, The CIE 1976 color difference formulae, *Color Res. Appl.* **2**, No. 1, 1977, 7–11.
19. G. Wyszecki and W. S. Stiles, *Color Science*, Wiley, New York, 1967.
20. R. H. Wallis, *Film Recording of Digital Color Images*, Ph.D. thesis, USCIPI Report 570, Image Processing Institute, University of Southern California, CA, 1975.
21. W. T. Wintringham, Color television and colorimetry, *Proc. IRE*, **39**, 1951, 1135–1172.
22. K. McLaren, The development of the CIE ($L^*, a^*, b^*$)-uniform color space, *J. Soc. Dyers Colour*, 1976, 338–341.
23. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.
24. N. Ahuja, A. Rosenfeld, and R. M. Haralick, Neighbor gray levels as features in pixel classification, *Pattern Recognit.* **12**, 1980, 251–260.
25. C. H. Chen, On the statistical image segmentation techniques, in *Proceedings, IEEE Conf. Pattern Recognition and Image Processing, 1981*, pp. 262–266.
26. M. Celenk and S. H. Smith, Gross segmentation of color images of natural scenes for computer vision systems, in *Proceedings, SPIE Annual Conf. Applications of Artificial Intelligence III, Orlando, 1986*, Vol. 635, pp. 333–344.
27. M. Celenk, *Parametric-Histogramming Technique for Gross Segmentation of Color Images of Natural Scenes*, Ph.D. thesis, Stevens Institute of Technology, Hoboken, NJ, 1983.
28. M. Celenk and S. H. Smith, Color image segmentation by clustering and parametric histogramming technique, in *Proceedings, IAPR Eight Int. Conf. Pattern Recognition, Paris, 1986*, pp. 883–886.
29. M. Celenk, Modular design of the segmentation unit of hierarchical computer vision systems, in *Proceedings, IEEE Int. Conf. Robotics and Automation, Raleigh, 1987*, Vol. 1, pp. 372–379.
30. M. Celenk, An adaptive machine learning algorithm for color image analysis and processing, in *Proceedings, MSTF Int. Conf. Manufacturing Science and Technology of the Future, Cambridge, 1987.*
31. M. Celenk, A parametric training algorithm for image understanding systems, in *Proceedings, SPIE Annual Conf. Applications of Artificial Intelligence V, Orlando, 1987*, Vol. 786, pp. 169–175.
32. S. Tominaga, Color image segmentation using three perceptual attributes, in *Proceedings, Conf. Computer Vision and Pattern Recognition, 1986*, pp. 628–630.
33. G. J. Klinker, S. A. Shafer, and T. Kanade, Image segmentation and reflection analysis through color, in *Proceedings, SPIE Annual Conf. Applications of Artificial Intelligence VI, Orlando. 1988*, Vol. 937, pp. 229–244.
34. M. Celenk, An adaptive machine learning algorithm for color image analysis and processing, *Int. J. Rob. Comput. Integr. Manuf.* **4**, Nos. 3/4, 1988, 403–412.