COMPUTER GRAPHICS AND IMAGE PROCESSING **13**, 222 – 241 (1980)

Color Information for Region Segmentation

YU-ICHI OHTA, TAKEO KANADE, AND TOSHIYUKI SAKAI

Department of Information Science, Kyoto University, Kyoto, Japan

Received October 11, 1978; accepted June 8, 1979

In color image processing various kinds of color features can be calculated from the tristimuli R, G, and B. We attempt to derive a set of effective color features by systematic experiments of region segmentation. An Ohlander-type segmentation algorithm by recursive thresholding is employed as a tool for the experiment. At each step of segmenting a region, new color features are calculated for the pixels in that region by the Karhunen Loeve transformation of R, G, and B data. By analyzing more than 100 color features which are thus obtained during segmenting eight kinds of color pictures, we have found that a set of color features, $(R + G + B)/3$, $R - B$, and $(2G - R - B)/2$, are effective. These three features are significant in this order and in many cases a good segmentation can be achieved by using only the first two. The effectiveness of our color feature set is discussed by a comparative study with various other sets of color features which are commonly used in image analysis. The comparison is performed in terms of both the quality of segmentation results and the calculation involved in transforming data of R, G, and B to other forms.

LIST OF SYMBOLS

Color Features

R, G, B	original tristimuli Red, Green, and Blue
Y, I, Q	color system for TV signal
D, S, H	D (intensity), S (saturation), H (hue)
r, g, b	normalized colors
X, Y, Z	*C.I.E.* X–Y–Z color system
L, a, b	L–a–b cube-root color system
U^*, V^*, W^*	U^*–V^*–W^* color system
X_1, X_2, X_3	color features obtained by the K.L. transformation
$I1, I2, I3$	$I1 = (R + G + B)/3$, $I2 = (R - B)/2$, $I3 = (2G - R - B)/4$
$I2', I3'$	$I2' = R - B$, $I3' = (2G - R - B)/2$
R', G', B'	reproduced color features

Segmentation Algorithm

S	a region represented by a mask

Color Feature Selection

Σ	covariance matrix
λ_i	eigenvalue of Σ
$\mathbf{W}_i$	eigenvector of Σ
w_{Ri}, w_{Gi}, w_{Bi}	elements of W_i
$\lVert\mathbf{W}_i\rVert$	norm of W_i
X_i	color feature obtained by the K.L. transformation
$\lvert w_R\rvert$	absolute value of w_R

Others

min (r, g, b) minimum value of r, g, and b

σ_{I1}^2 variance of the image of the color feature $I1$

0146-664X/80/070222-20 \$02.00/0

1. INTRODUCTION

In color image processing, color of a pixel is usually given as three values corresponding to the tristimuli R (red), G (green), and B (blue). Various kinds of color features such as intensity, saturation, and hue can be calculated from the tristimuli by using either linear or nonlinear transformations. Each color feature has its own characteristics. For instance, the set of $\{D$ (intensity), S (saturation), H (hue)$\}$ is convenient for representing the human color perception; the set of $\{Y, I, Q\}$ is used to efficiently encode color information in the TV signal; the normalized color $\{r, g, b\}$ is convenient to represent the color plane; etc.

It seems that in computer processing of color images, color features which are developed for other purposes have been used in different combinations for different purposes. Nevatia [1] extended the Hueckel operator for color edge extraction. He stated that the result obtained by using intensity $(D = R + G + B)$ and normalized colors ($r = R/D$ and $g = G/D$) was better than that obtained by using R, G, and B. Ohlander [2] employed nine redundant color features R, G, B, Y, I, Q, D, S, and H for color image segmentation. He reported that H was most useful and that Y, I, and Q were rarely used.

Kender [3] presented a very careful discussion of the behavior of linear and nonlinear color transformations to obtain color features such as hue, saturation, and normalized color from R, G, and B. He pointed out two problems: (1) nonlinear transformations like hue, saturation, and normalized color have nonremovable singularities, near which a small perturbation of input R, G, and B can cause a large jump in the transformed values; (2) the distribution of the nonlinearly transformed values can show spurious modes and gaps. By these reasons and from the computational point of view, he said that linear transformations such as Y, I, and Q would be more preferable than nonlinear ones.

It is an interesting and important problem to find color features which are suited for the segmentation process of color images by computer. One way to get such color features is to execute the segmentation by using various sets of color features and to compare the results. However, this way allows us to examine only predefined sets of color features. In this paper we attempt to derive a set of effective color features by systematic experiments in region segmentation. An Ohlander-type segmentation algorithm by recursive thresholding is employed as a tool for the experiments. At each step of segmenting a region, new color features are calculated for the pixels in that region by the Karhunen Loeve transformation of R, G, and B data. By analyzing the color features obtained in segmenting eight kinds of color pictures, we have found a set of effective color features. The effectiveness of our color feature set is discussed by a comparative study with various other sets of color features which are commonly used in image analysis.

2. SELECTION OF EFFECTIVE COLOR FEATURES

2.1. Region Splitting Using Plural Histograms

First of all, we briefly describe the segmentation algorithm which is employed in the experiments. Figure 1 shows a schematic diagram of the segmentation algorithm. The basic idea of the process is as follows: the whole image is first partitioned into subimages each of which is a connected region; then each subimage is further partitioned if possible; and this process iterates. Because of the

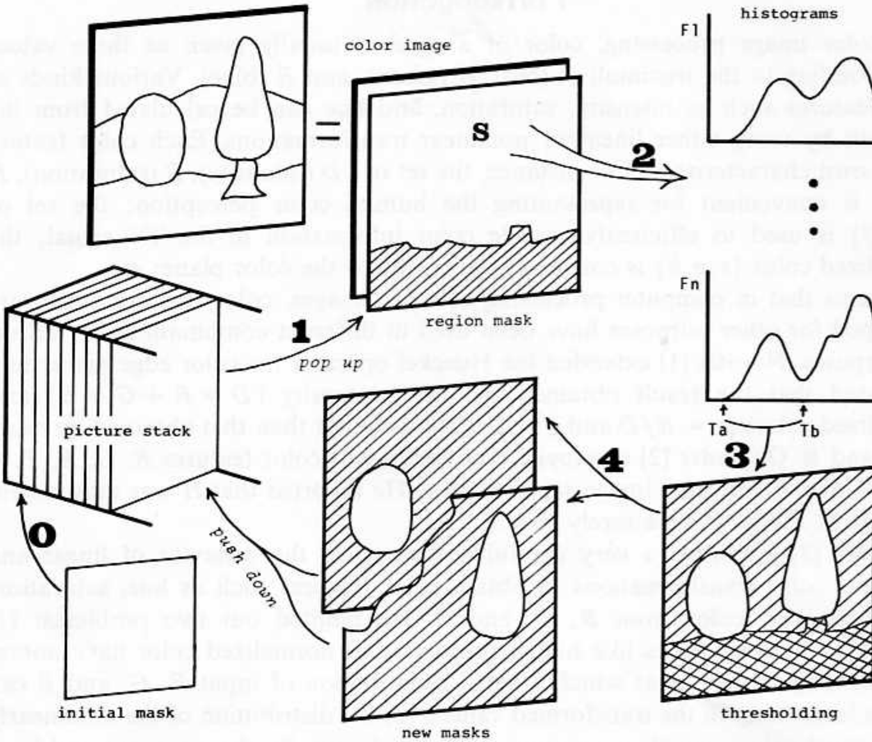


FIG. 1. Region segmentation by recursive thresholding.

recursive nature of the algorithm, a picture stack is used to store region masks. A region mask represents a connected region (the area without hatching in Fig. 1) which is to be examined for segmentation. The arrows with numbers shown in Fig. 1 represent the following operations.

0. A mask corresponding to the whole image is placed at the bottom of the stack.

1. One mask is taken from the top of the stack. Let S denote the region represented by the mask (the area without hatching).

2. Histograms of color features in the region S are computed.

3. If any of the histograms show conspicuous peaks, a pair of cutoff values which separate the peak in the histograms are determined at the position of valleys, and the image of the color feature corresponding to that histogram is thresholded using the cutoff values; thus the region S is partitioned. Otherwise, this region S is not partitioned further.

4. Connected regions are extracted. For each connected region, a region mask is generated and it is pushed down on the stack.

The operations from 1 to 4 are iterated until the picture stack becomes empty. In the operation of arrow 3, the cutoff values are selected by the following two criteria: Candidates of cutoff values are selected by evaluating the shape of peaks on the histograms; Bad cutoff values are rejected by verifying on the image the compactness of spatial distribution of the pixels belonging to the peak determined by the pair of cutoff values.

2.2. Computation of Color Features Using the K.L. Transformation

At the step of arrow 2 in Fig.1, color features for which the histograms are computed are often fixed to some features such as R, G, and B throughout the segmentation process. However, the color feature which has deep valleys on its histogram and has the largest discriminant power to separate the clusters in a given region need not be the same. In feature selection in pattern recognition theory, a feature is said to have large discriminant power if its variance is large. Thus we tried to derive color features with large discriminant power by the Karhunen Loeve transformation (K.L. transformation).

More specifically, let S be the region to be segmented, and Σ be the covariance matrix of the distributions of R, G, and B in S. Let λ_1, λ_2, and λ_3 be the eigenvalues of Σ and $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Let $\mathbf{W}_i = (w_{Ri} w_{Gi} w_{Bi})'$ for $i = 1$, 2, and 3 be the eigenvectors of Σ corresponding to λ_i, respectively. The color features X_1, X_2, and X_3 are defined as:

$$X_i = w_{Ri} \cdot R + w_{Gi} \cdot G + w_{Bi} \cdot B \qquad (\|\mathbf{W}_i\| = 1, i = 1, 2, \text{ and } 3). \tag{1}$$

It is well known that X_1, X_2, and X_3 are uncorrelated, and X_1 is the "best" feature in the sense that it has the largest variance (the value is λ_1). X_2 is the "best" one among the orthogonal ones to X_1. At each step of segmenting a region, three new color features X_1, X_2, and X_3 are calculated for the pixels in that region and used to compute the histograms. We call this scheme "segmentation by dynamic K.L. transformation."

Eight scenes shown in Fig. 2 are used for the experiments. Scene (a) shows a cylinder with color stripes illuminated from the front. The names of eight scenes are (a) cylinder, (b) building, (c) seaside, (d) girl, (e) room, (f) home, (g) auto, and (h) face. They were digitized with 256 × 256 spatial resolution and 6 bits density resolution for each of R, G, and B. The scenes of (a), (b), and (c) in Fig. 2 are digitized at Kyoto University. Scene (d) is from Southern California University and (e) through (h) are from Carnegie–Mellon University. Scenes (e), (g), and (h) are the images which Ohlander used in his experiment [2], except that the size and density resolution are reduced for our system. Scene (f) is almost the same as Ohlander's home scene except that there are some clouds in the sky.

Figure 3 shows results of segmentation by the dynamic K.L. transformation. Use of the "best" color features calculated adaptively at each step of segmenting a region produces satisfying results. In the cylinder scene, for example, the horizontal color stripes are separated almost completely and the vertical boundaries which separate the color stripes vertically because of the difference of intensity are relatively few. The scheme of using the K.L. transformation on the fly requires costly computation and is not very practical. Rather, our goal is to discover a set of color features with which we can achieve segmentation as good as those by the dynamic K.L. transformation.

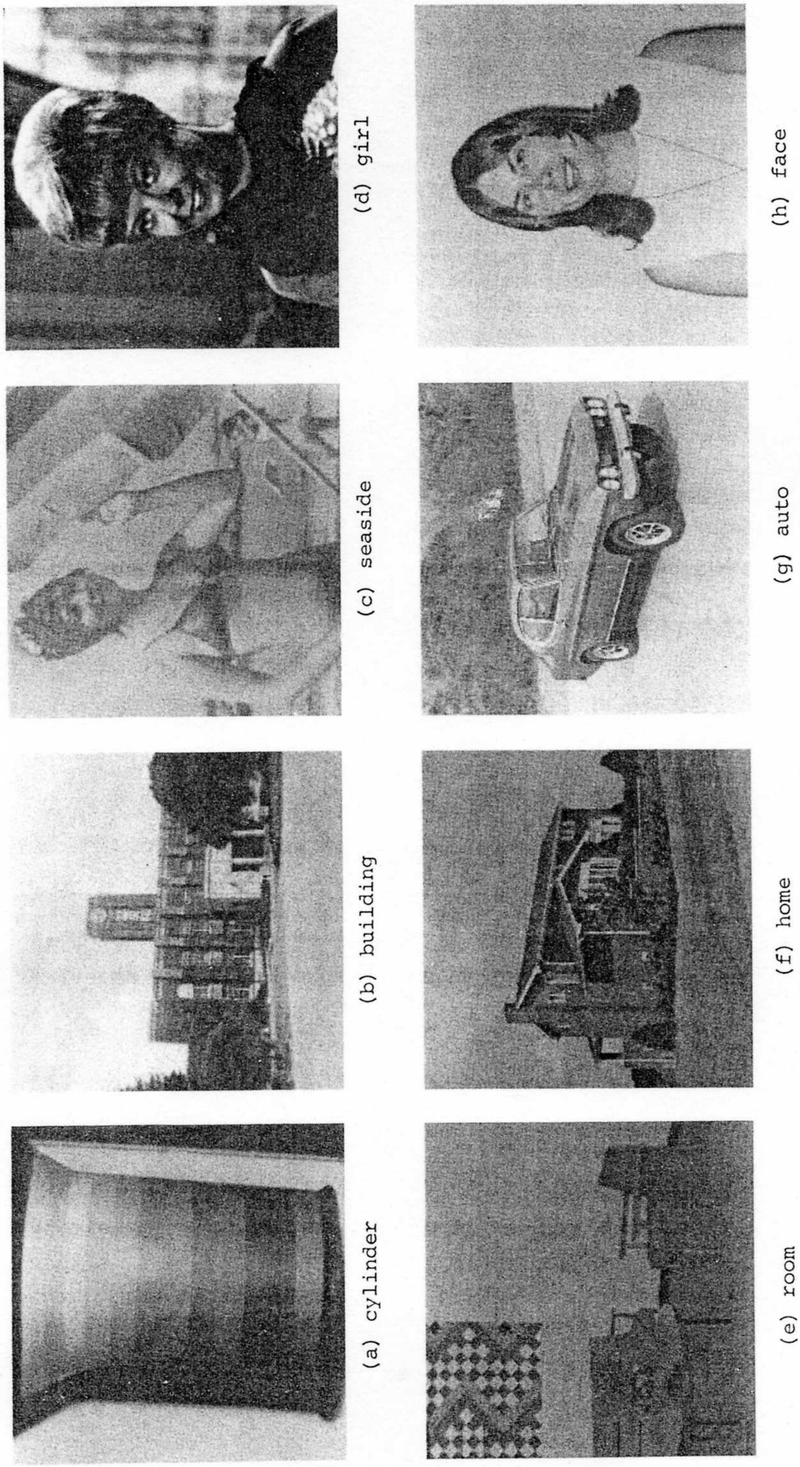


FIG. 2. Color scenes (reproduced in black and white) used for experiments.

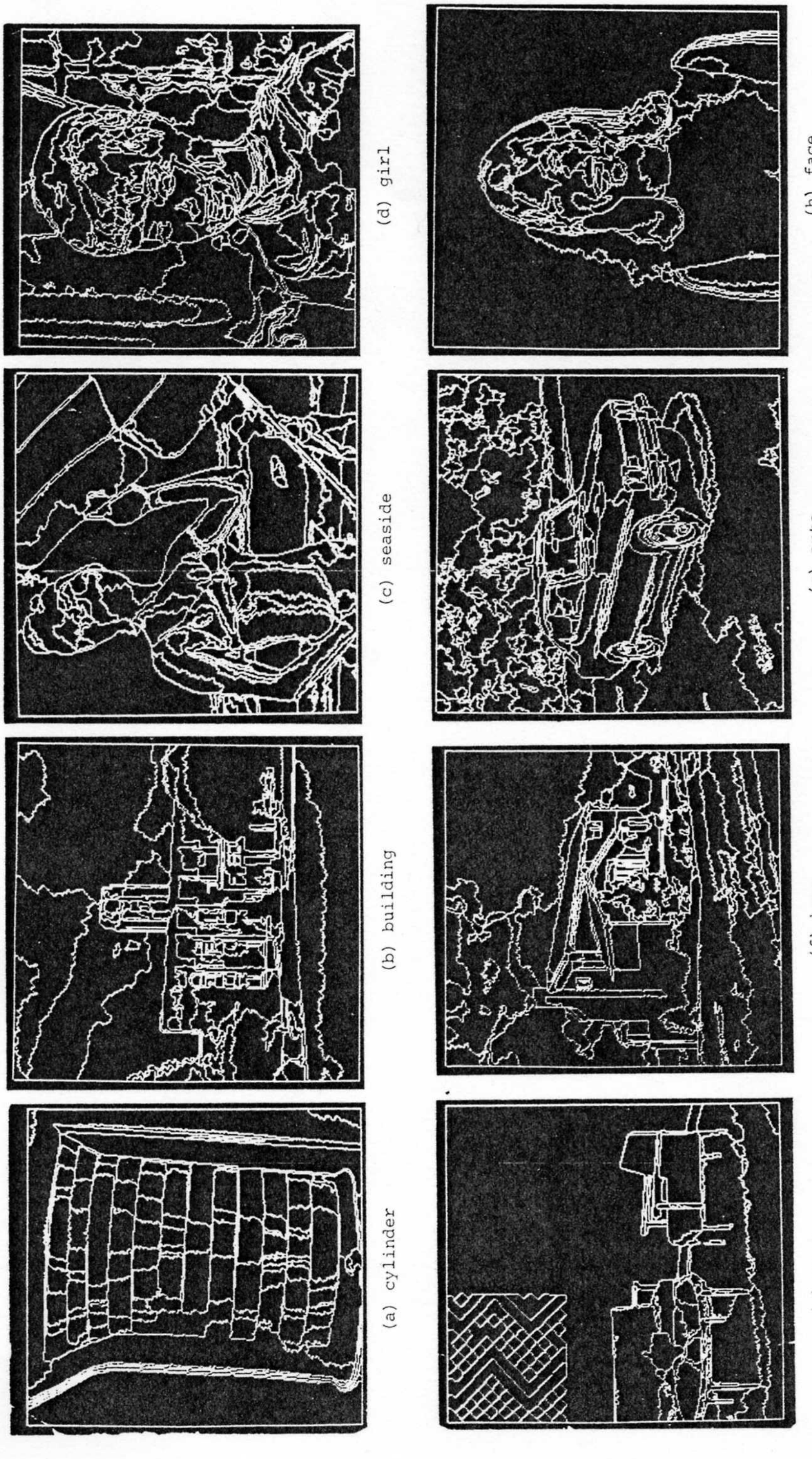


FIG. 3. Segmentation results by dynamic K.L. transformation.

2.3. *A Set of Effective Color Features*

Table 1 shows the eigenvectors of Σ for the whole image of each of the eight color scenes (reproduced in black and white) in Fig. 2. It is interesting that $\mathbf{W}_1$ is about $(\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3})^t$ for every scene. $\mathbf{W}_2$ is dominated by $(\frac{1}{2}\ 0\ -\frac{1}{2})^t$ or $(-\frac{1}{2}\ 0\ \frac{1}{2})^t$ and $\mathbf{W}_3$ by $(-\frac{1}{4}\ \frac{1}{2}\ -\frac{1}{4})^t$. Then it is possible to say that the three orthogonal color features, $I1 = (R+G+B)/3$, $I2 = (R-B)/2$ or $(B-R)/2$, and $I3 = (2G-R-B)/4$, are important components for representing color information.

To prove this experimentally, we analyzed the linear combinations of R, G, and B, which are used to find the cutoff values for thresholding in the segmentation process by the dynamic K.L. transformation for the eight scenes. We examined only the color features which are used in segmenting regions with an area larger than 1000 in such a manner that both split regions have areas larger than 200. The number of color features thus gathered is 109. The weight vectors of these color features are plotted on a $w_R - w_B$ plane as shown in Fig. 4. The weight vectors have been normalized so that $w_G \geq 0, |w_R| + |w_G| + |w_B| = 1$. For simplicity, each position of the vectors is plotted with the head letter of the name of the scene for which the color feature was used. The weight vectors corresponding to $\{I1, I2, I3\}$, $\{Y, I, Q\}$, $\{X, Y, Z\}$, $\{U, V, W\}$, and $\{R, G, B\}$ are indicated for reference. Contour lines are drawn to show equidistance from the reference points $I1$, $I2$, and $I3$. The curves are defined by:

$$\begin{aligned}
&\left(w_R - \tfrac{1}{3}\right)^2 + \left(w_G - \tfrac{1}{3}\right)^2 + \left(w_B - \tfrac{1}{3}\right)^2 = \epsilon && \text{(around } I1\text{)},\\
&\left(w_R - \tfrac{1}{2}\right)^2 + w_G^2 + \left(w_B + \tfrac{1}{2}\right)^2 = \epsilon && \text{(around } I2\text{)},\\
&\left(w_R + \tfrac{1}{2}\right)^2 + w_G^2 + \left(w_B - \tfrac{1}{2}\right)^2 = \epsilon && \text{(around } I2\text{)},\\
&\left(w_R + \tfrac{1}{4}\right)^2 + \left(w_G - \tfrac{1}{2}\right)^2 + \left(w_B + \tfrac{1}{4}\right)^2 = \epsilon && \text{(around } I3\text{)}\\
&(\epsilon = 1/81, 1/27, 1/9, 1/6).
\end{aligned} \tag{2}$$

Color features in the first quadrant have weight vectors such that $w_R, w_G, w_B > 0$. They correspond mainly to the intensity component and $I1 = (R+G+B)/3$ is the most typical feature of this quadrant. In the second and fourth quadrants, w_R and w_B have opposite signs and the color features in these quadrants represent the

TABLE 1

Eigenvectors of Σ for a Whole Image.[a]

	w_{R1}	w_{G1}	w_{B1}	w_{R2}	w_{G2}	w_{B2}	w_{R3}	w_{G3}	w_{B3}
Cylinder	0.269	0.363	0.367	0.469	0.095	−0.437	−0.308	0.461	−0.231
Building	0.269	0.340	0.391	0.479	0.103	−0.418	−0.296	0.485	−0.219
Seaside	0.258	0.380	0.362	−0.585	0.056	0.358	−0.176	0.464	−0.360
Girl	0.336	0.354	0.309	−0.493	0.193	0.314	−0.094	0.474	−0.436
Room	0.193	0.341	0.467	0.612	0.079	−0.310	−0.209	0.507	−0.284
Home	0.197	0.328	0.476	0.492	0.180	−0.328	−0.313	0.484	−0.204
Auto	0.304	0.317	0.378	0.239	0.309	−0.452	−0.514	0.450	0.036
Face	0.175	0.411	0.414	0.523	0.128	−0.349	−0.295	0.416	−0.289

[a] Normalized as $w_G \geq 0, |w_R| + |w_G| + |w_B| = 1$.

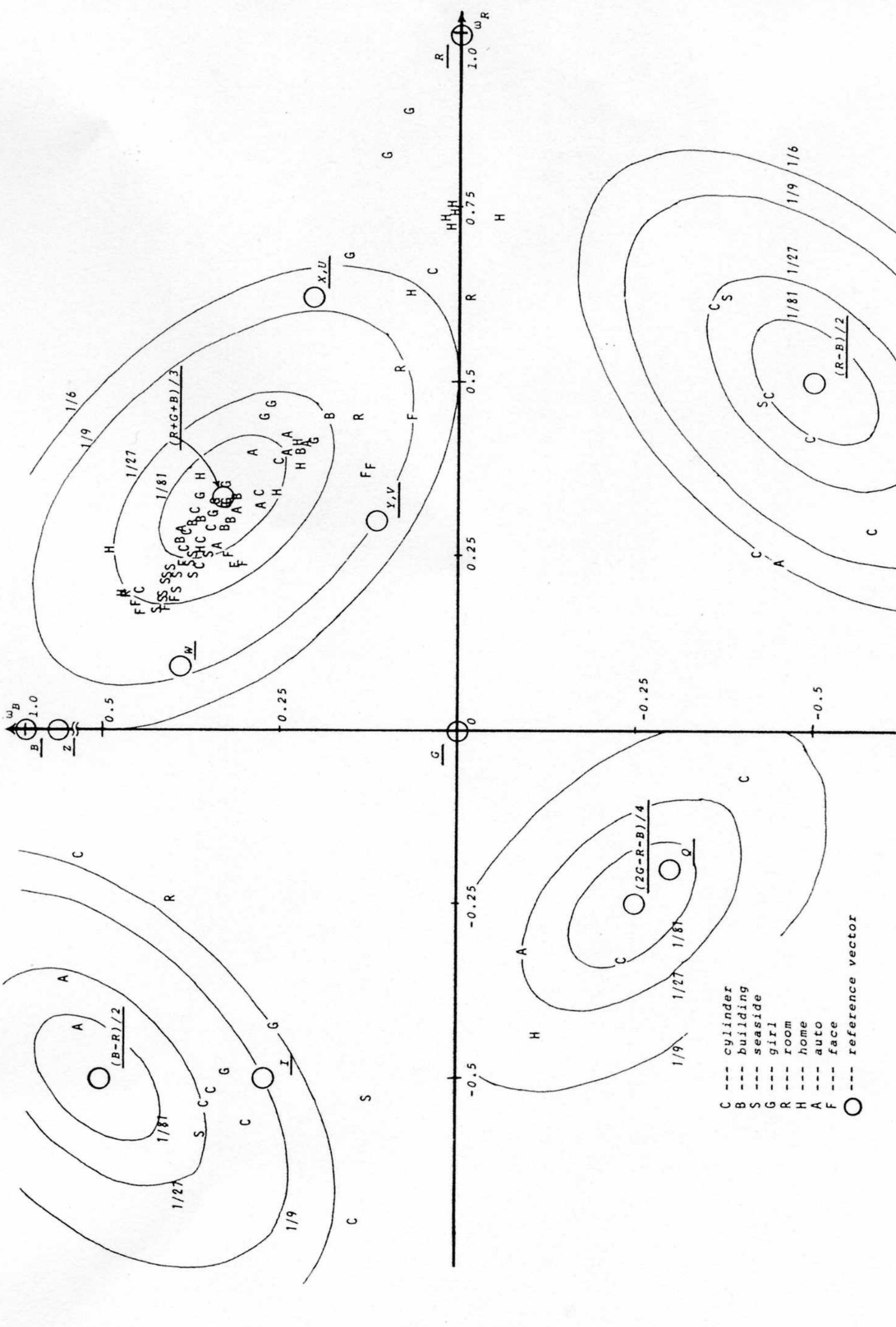


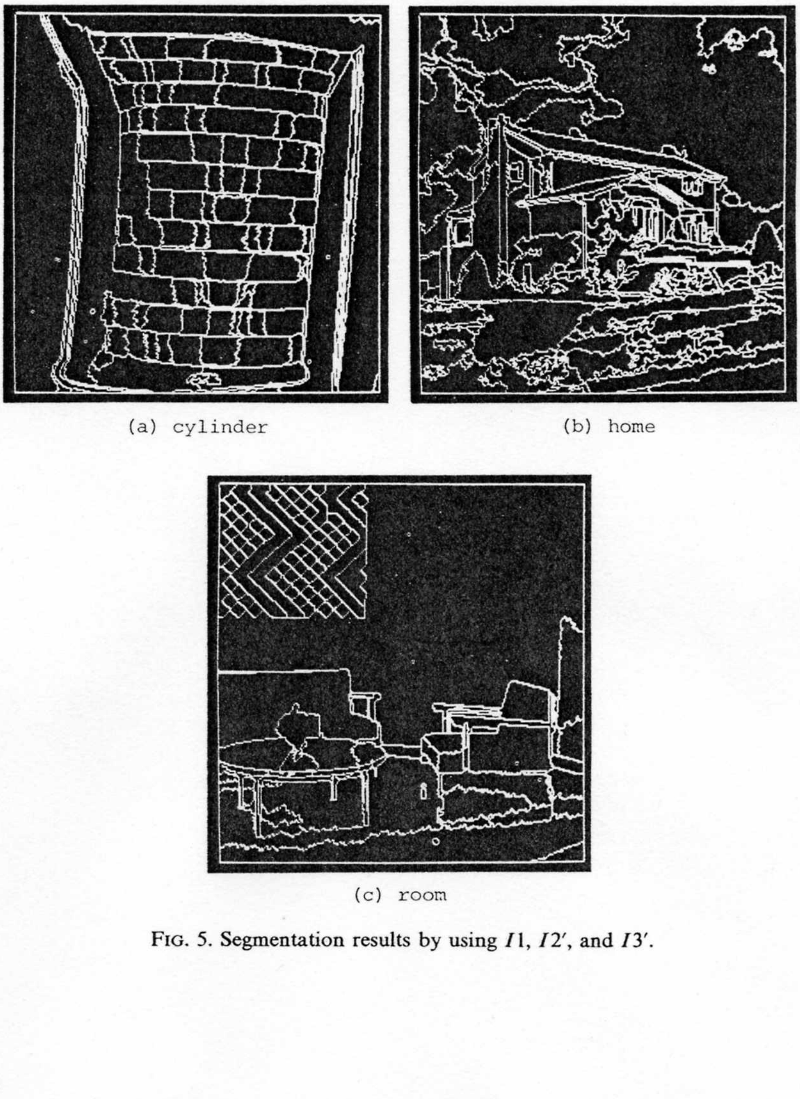
FIG. 4. Weight vectors of 109 color features used in the segmentation by dynamic K.L. transformation.

difference of R and B components. Most color features are in the first quadrant. This means that the intensity is the most important feature even in color image processing.

The weight vector of $I1$ is nearly at the center of the weight vectors in the first quadrant as shown in Fig. 4. $I2$ can be regarded as being at the center of the weight vectors in the second and the fourth quadrants. $I3$ will be a typical color feature in the third quadrant. Thus, it is possible to assume that every weight vector in the four quadrants can be approximated by the weight vectors of the three color features, $I1$, $I2$, and $I3$. The number of weight vectors in the first, second/fourth, and third quadrant in Fig. 4 are 83, 22, and 4, respectively. So, $I1$, $I2$, and $I3$ are assumed to be significant in this order.

We have performed the following experiments to verify the assumption described above.

(1) Segmentation by using the set of three color features. Figure 5 shows the results obtained by using the set of three fixed color features $I1$, $I2' = (R - B)$, and $I3' = (2G - R - B)/2$. They seem to be nondegraded compared to those obtained by dynamic K.L. transformation. This verifies the assumption that the set



(a) cylinder (b) home

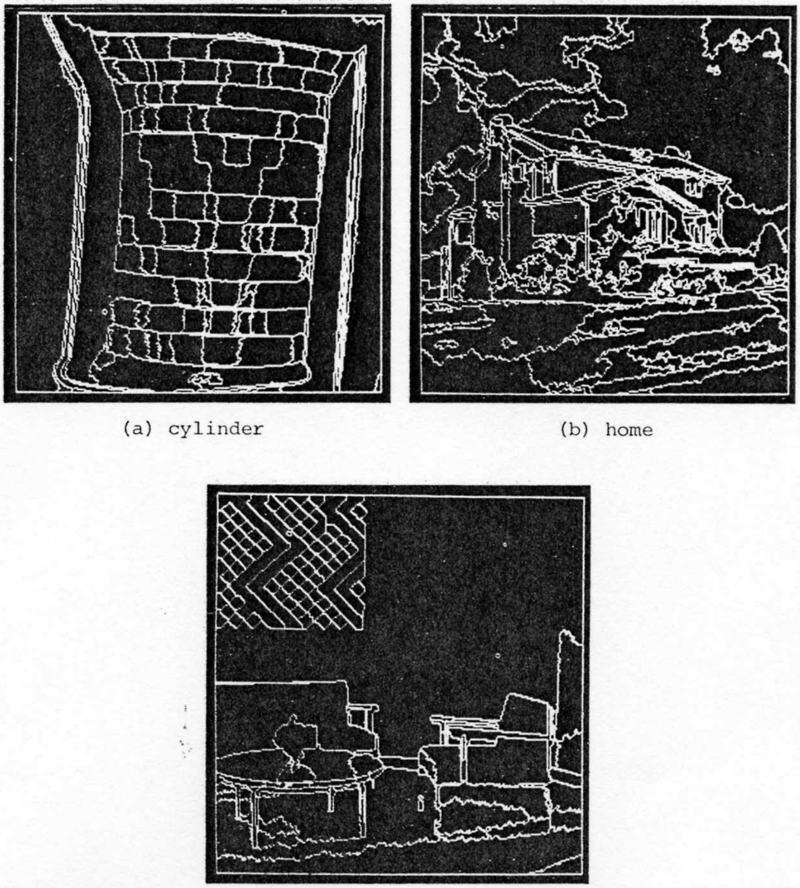
(c) room

FIG. 5. Segmentation results by using $I1$, $I2'$, and $I3'$.

of three color features $I1$, $I2'$, and $I3'$ can approximate all color features which are calculated as the "best" ones at each important step in segmenting the eight color scenes.

(2) Segmentation by using two color features. In Fig. 4 there are only four color features in the third quadrant, far fewer than in other quadrants. Thus the omission of $I3'$ will not significantly affect the quality of segmentation results. This is verified by the results shown in Fig. 6, which are obtained by using only two color features $I1$ and $I2'$. A picture indicating the boundaries in Fig. 5 missing in Fig. 6 is shown in Fig. 7 for cylinder scene in order to help visual comparison. Even in the cylinder scene which has two weight vectors in the third quadrant, the results of Figs. 5a and 6a are almost the same except that the fifth and sixth color stripes, golden yellow and orange, are not separated in Fig. 6a.

(3) Segmentation by using only one color feature I1. What quality of segmentation can be achieved by using only one color feature $I1$, i.e., black and white



(a) cylinder (b) home

(c) room

FIG. 6. Segmentation results by using $I1$ and $I2'$.

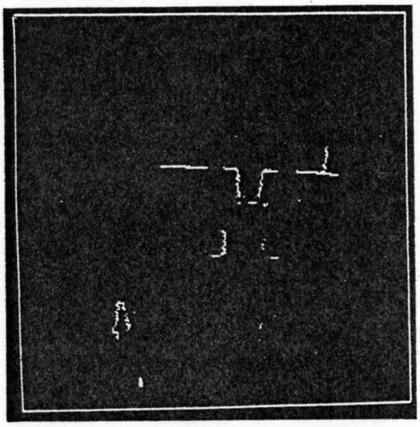
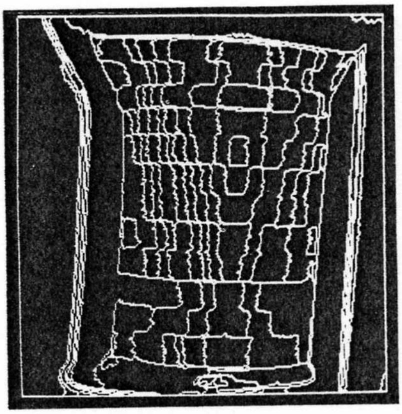
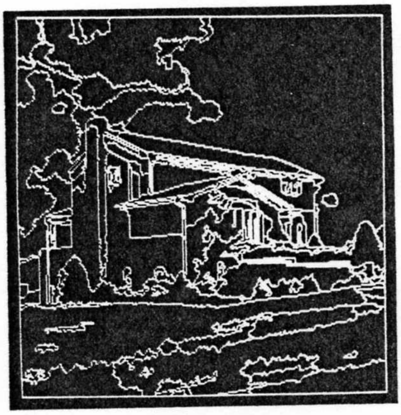


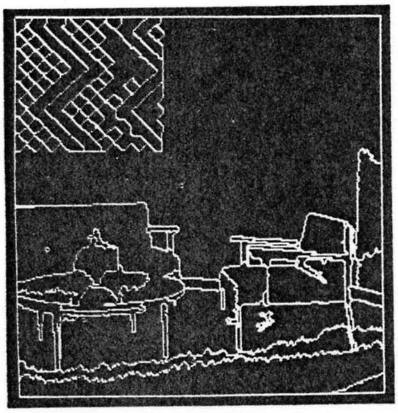
FIG. 7. Boundaries in Fig. 5a missing in Fig. 6a.



(a) cylinder



(b) home



(c) room

FIG. 8. Segmentation results by using only $I1$.

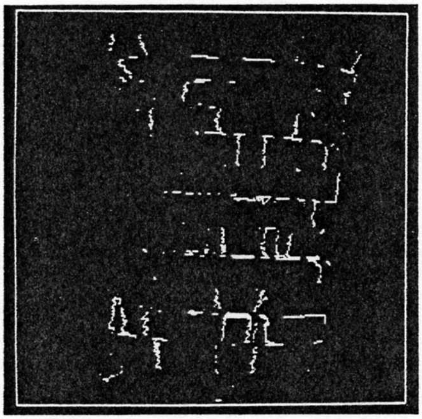


FIG. 9. Boundaries in Fig. 5a missing in Fig. 8a.

information? In the case of the cylinder scene, the fraction of the number of weight vectors in the first quadrant is 10 out of 22 weight vectors, and there are 10 weight vectors in the second and fourth quadrants. Therefore the quality of the segmentation for the cylinder scene will be significantly degraded by omitting the color feature $I2'$. In the other scenes, however, only a few weight vectors are in the second and fourth quadrants and the degradation will not be so significant. The results obtained using only $I1$ are shown in Fig. 8. Figure 9 shows the boundaries in Fig. 5a missing in Fig. 8a. Separation of color stripes in the cylinder scene is not done as well as expected. It is a natural outcome of missing distinguishability for color difference by omitting $I2'$ and $I3'$. In the case of the room scene, however, the quality of segmentation is not degraded as much, though there appears a little missegmentation. As for the home scene in Fig. 8b, there is no missegmentation, in the sense that the regions which should be segmented are all segmented.

3. COMPARISON OF COLOR FEATURES

3.1. Segmentation by Various Sets of Color Features

In order to discuss the effectiveness of the color feature set obtained in the previous section, the eight scenes used in the previous experiment are also segmented by using seven sets of color features which are commonly used in image analysis. The sets are $\{R, G, B\}$, $\{X, Y, Z\}$, $\{Y, I, Q\}$, $\{L, a, b\}$, $\{U^*, V^*, W^*\}$, $\{I1, r, g\}$, and $\{I1, S, H\}$. R, G, and B are the original tristimuli. X, Y, and Z correspond to the C.I.E. X–Y–Z primary color coordinate system. Y–I–Q is the color coordinate system for television signal. L–a–b color coordinate system is designed to agree with the Munsell color system. U^*–V^*–W^* is designed to obtain a color solid for which unit shifts in luminance and chrominance are uniformly perceptible [4]. $I1$, S, and H denote the intensity, saturation, and hue, respectively, and r and g denote the normalized colors. Other sets of color features such as $\{U, V, W\}$, $\{S, \theta, W^*\}$, and $\{u, v, V\}$ are not examined because they are similar to $\{X, Y, Z\}$, $\{I1, S, H\}$, and $\{I1, r, g\}$, respectively.

$\{Y, I, Q\}$ and $\{X, Y, Z\}$ are calculated from $\{R, G, B\}$ in our experiments by

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.500 & -0.230 & -0.270 \\ 0.202 & -0.500 & 0.298 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \tag{3}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.618 & 0.177 & 0.205 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.056 & 0.944 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{4}$$

The transformation matrices are not the standard ones: the weights for I, Q, X, and Z are rescaled to normalize the range of the transformed values to be the same as the original. $\{L, a, b\}$ and $\{U^*, V^*, W^*\}$ are defined as

$$\begin{aligned} L = W^* &= 25(100Y/Y_0)^{1/3} - 16, \\ a &= 500\left[(X/X_0)^{1/3} - (Y/Y_0)^{1/3}\right], \\ b &= 200\left[(Y/Y_0)^{1/3} - (Z/Z_0)^{1/3}\right], \\ U^* &= 13W^*(u - u_0), \\ V^* &= 13W^*(v - v_0), \end{aligned} \tag{5}$$

where $u_0 = 0.199$, $v_0 = 0.308$, $u = 4X/(X + 15Y + 3Z)$, $v = 6Y/(X + 15Y + 3Z)$, and X_0, Y_0, Z_0 are the X–Y–Z values for the reference white. Normalized colors, intensity, saturation, and hue are obtained as follows:

$$\begin{aligned} r &= R/(R + G + B), \qquad g = G/(R + G + B), \qquad b = B/(R + G + B), \\ I1 &= (R + G + B)/3, \\ S &= 1 - 3 \cdot \min(r, g, b), \\ H &= \arctan 2\left(\sqrt{3}\,(G - B), (2R - G - B)\right). \end{aligned} \tag{6}$$

There are two problems in using those color features for region segmentation. One is instability of nonlinear transformations. Normalized color, U^*, V^*, and saturation become unstable and meaningless when $R + G + B$ is small. Therefore, they are not used to compute histograms if $R + G + B$ is less than 30 in segmenting a region. Hue is unstable where saturation is near zero and is not used if $S \times (R + G + B)$ is less than 9. The other problem is caused by the fact that input data of R, G, and B is digitized. The histogram of the transformed values from digital input may have comblike structure. In order to prevent this, the input R, G, and B values are "undigitized" by adding a random number uniformly selected from the unit interval to each of them [3].

Figures 10 through 16 show the results of segmentation obtained by using the seven sets of color features.

3.2. Discussions

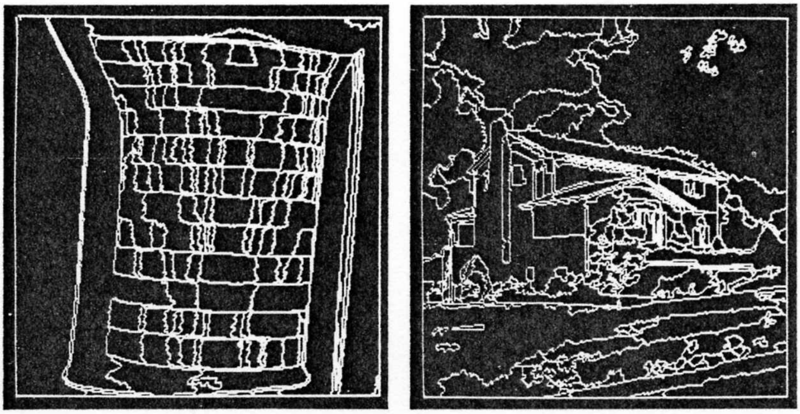
The effectiveness of a set of color features used in the segmentation process can be evaluated in terms of the quality of segmentation results and the behavior of the transformation from input tristimuli R, G, and B.

The method for evaluating the quality of segmentation results is very difficult. No quantative evaluation procedure has been established for segmentation of natural scenes. We adopted eyeballs as the most reliable tool at present. Pictures which indicate the difference between a pair of segmentation results are sometimes generated to help visual comparison as shown in Figs. 7 and 9. We think the segmentation errors which fail to split the regions that must be separated (undersegmentation) affect later processings more seriously than the errors which split the regions that need not be separated (oversegmentation). So the evaluation criteria are set more severely against undersegmentation than oversegmentation.

Use of the color feature set $\{R, G, B\}$ for segmentation requires no transformation. But, R, G, and B each have a strong factor of intensity and they are heavily correlated. Thus, spurious segmentations tend to happen because of the difference of intensity. This tendency clearly appears in Fig. 10a, which is segmented by using R, G, and B. Note that the vertical splitting of color stripes occurs more frequently in Fig. 10a than in Fig. 5a, which is segmented by the set of $I1$, $I2'$, and $I3'$.

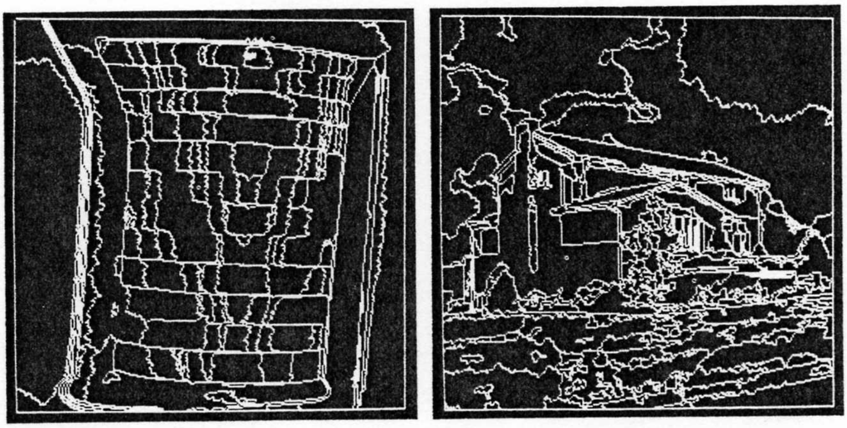
The weight vectors of X, Y, and Z are located in the first quadrant as shown in Fig. 4; i.e., they all have strong factors of intensity. This results in the similar segmentation to that obtained by using R, G, and B (see Figs. 10a and 11a). The separation of color stripes in Fig. 11a is worse than that in Fig. 10a because the weight vectors of X and Y are closer to the white point ($I1$) in Fig. 4 than those of R and G.

Y, I, and Q are in the first, fourth, and third quadrant, respectively (see Fig. 4). The segmentation results obtained by $\{Y, I, Q\}$ are similar to those obtained by $I1$, $I2'$, and $I3'$. In Fig. 12a, the uppermost two color stripes are not separated, while they are separated in Fig. 5a. The reason is that the weight vector of color feature I is located at a biased position in the fourth quadrant and it is not a good approximation of the color features in the second quadrant. A more important difference between $\{Y, I, Q\}$ and $\{I1, I2', I3'\}$ is in their calculation from $\{R, G, B\}$. The transformation of $\{Y, I, Q\}$ from $\{R, G, B\}$ needs floating multiplications. Furthermore, there is the possibility that spurious combs appear in the histograms of Y, I, and Q. In the case of $\{I1, I2', I3'\}$, all coefficients of the transformation



(a) cylinder (b) home

FIG. 10. Segmentation results by using $\{R, G, B\}$.

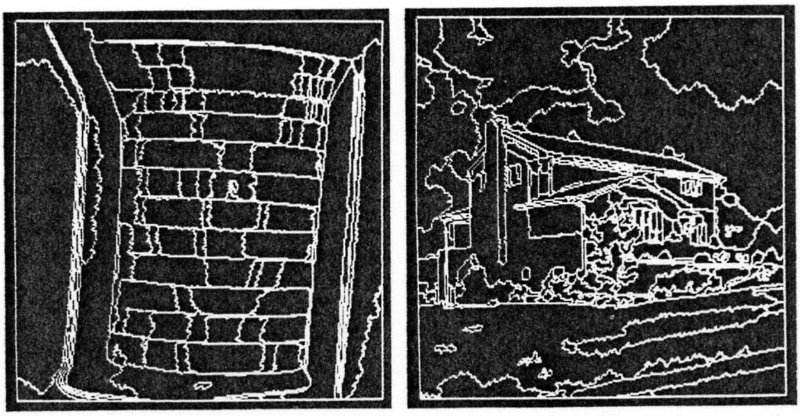


(a) cylinder (b) home

FIG. 11. Segmentation results by using $\{X, Y, Z\}$.

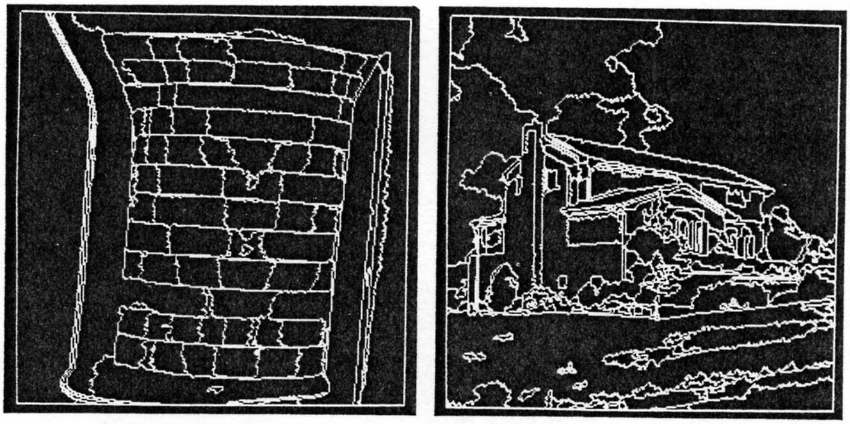
from R, G, and B have the form of $1/N$. This means that comblike structure never appears in the histograms of $I1$, $I2'$, and $I3'$. The calculation of $\{I1, I2', I3'\}$ from $\{R, G, B\}$ is far simpler than that of $\{Y, I, Q\}$. It can be performed by using addition and subtraction of integer numbers and shifting or simple table-value operations for scaling.

Figure 13 shows the result obtained by using the set of color features $\{L, a, b\}$ and Fig. 14 is the result by $\{U^*, V^*, W^*\}$. Both color coordinate systems L–a–b and U^*–V^*–W^* use cube-root features for luminance as in Eq. (5). This results in good performance of separation of color stripes at the left side of the cylinder, where intensity is dark and gradually changes. L, a, and b are based on Y, $(X - Y)$, and $(Y - Z)$ color features, which are located in the first, third, and fourth quadrant, respectively. On the other hand, U^*, V^*, and W^* are based on the $u - v - V$ normalized color coordinate system derived from U–V–W, and U, V,



(a) cylinder (b) home

FIG. 12. Segmentation results by using $\{Y, I, Q\}$.

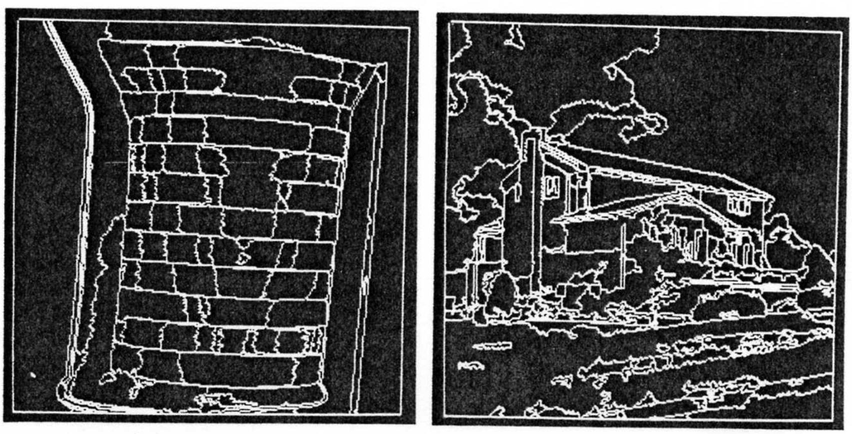


(a) cylinder (b) home

FIG. 13. Segmentation results by using $\{L, a, b\}$.

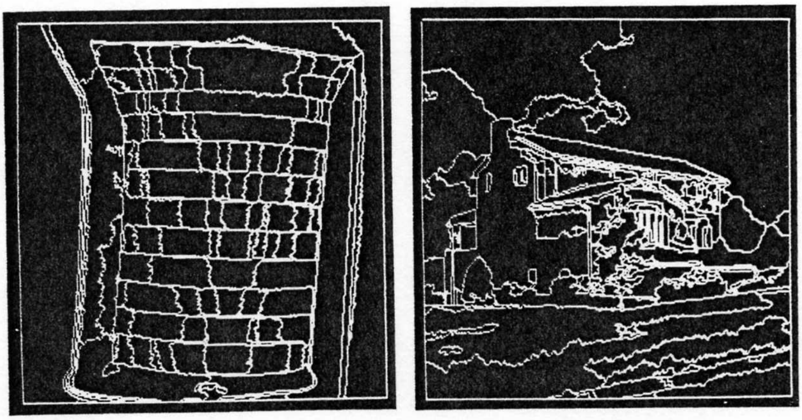
and W are all located in the first quadrant in Fig. 4. This causes the missegmentation at the borders of pale yellow and yellow stripes and the missegmentation at the border of golden yellow and orange stripes in the strongly illuminated part of cylinder surface in Fig. 14a while in Fig. 13a all color stripes are separated clearly. In the case of the home scene (and in other scenes), we cannot observe any significant difference among the segmentation results obtained by these two sets of features and other sets of features such as $\{I1, I2', I3'\}$, $\{R, G, B\}$, etc.

The set of color features $\{I1, r, g\}$ produces the result shown in Fig. 15. Use of color features normalized with intensity results in good segmentation at the dark part of cylinder scene as well as the results obtained by $\{L, a, b\}$ or $\{U^*, V^*, W^*\}$. Highly illuminated part of the border between the pale yellow and yellow stripes is not separated as in the result by $\{U^*, V^*, W^*\}$. Figure 16 shows the results by $\{I1, S, H\}$. They seem to be more degraded than those obtained by using $\{I1, r, g\}$ shown in Fig. 15. One reason is that the hue can be meaningful only in the limited



(a) cylinder (b) home

FIG. 14. Segmentation results by using $\{U^*, V^*, W^*\}$.

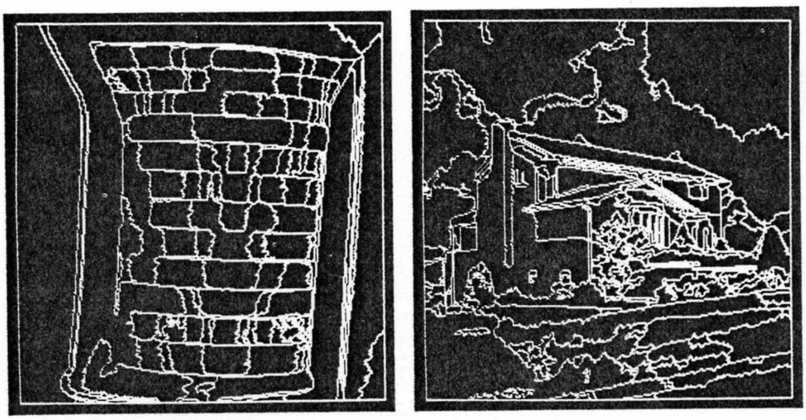


(a) cylinder (b) home

FIG. 15. Segmentation results by using $\{I1, r, g\}$.

cases. From the computational point of view, these nonlinear transformations incur far more cost than linear transformations.

Table 2 shows the variances, σ_{I1}^2, σ_{I2}^2, and σ_{I3}^2, of three components of a color image represented in the $I1$–$I2$–$I3$ color space. The variances are rescaled such that $\sigma_{I1}^2 + \sigma_{I2}^2 + \sigma_{I3}^2 = 100$ for each color image. The relation of $\sigma_{I1}^2 > \sigma_{I2}^2 > \sigma_{I3}^2$ exists for every color image. This relation corresponds to the relation of (the number of color features in the first quadrant) > (the number of color features in the second/fourth quadrant) > (the number of color features in the third quadrant) in Fig. 4. Thus, it can be said that color features with larger variance are more useful in region segmentation of a color image. However, there is no clear relation between the variance and the usefulness of a color feature in segmenting different kinds of color images. That is, the chromatic information is not always important for the segmentation process even in case of colorful scenes which have large variance for chromatic components. We think that the usefulness of a color feature



(a) cylinder (b) home

FIG. 16. Segmentation results by using $\{I1, S, H\}$.

TABLE 2

Variances of $I1$, $I2$, and $I3$ Images[a]

	σ_{I1}^2	σ_{I2}^2	σ_{I3}^2
Cylinder	92.4	6.5	1.1
Building	97.0	2.8	0.1
Seaside	80.6	17.0	2.4
Girl	85.8	10.4	3.8
Room	75.2	22.7	2.2
Home	76.7	19.5	3.8
Auto	89.9	6.4	3.7
Face	87.7	9.6	2.8

[a]Rescaled as $\sigma_{I1}^2 + \sigma_{I2}^2 + \sigma_{I3}^2 = 100$.

is greatly influenced by the structure of color scenes to be segmented. For instance, the variance of $I2$ is only 6.5 in the cylinder scene, where $I2$ plays an important role for segmentation, while it is 19.5 in the home scene, which was segmented well by using $I1$ alone. This phenomenon can be explained by the difference in the structure of the cylinder scene and home scene. The cylinder scene consists of curved surfaces, while the home scene mainly includes planar objects. On curved surfaces the intensity often gradually changes and can not be used as a useful feature for segmentation. This causes the chromatic information to be used frequently in the segmentation of the cylinder scene.

Table 3 shows the eigenvalues of Σ, λ_1, λ_2, and λ_3, obtained for the R, G, and B data in the whole image of each of the eight color scenes. The values have been rescaled such that $\lambda_1 + \lambda_2 + \lambda_3 = 100$ for each scene. λ_3 is very small for every scene and its maximum value is 3.6. This means that each color image can be approximated by two features X_1 and X_2 with mean square error of 3.6 at maximum. We tried to compose color images using only two features X_1 and X_2 by the inverse mapping of the K.L. transformation. Let R', G', and B' denote the color features obtained by the inverse transformation. We compared the color images composed by the R', G', and B' with the original color images. The following two facts were observed: (1) Although $R' - G' - B'$ color images are composed by using only two spectral features, they are good reproductions of the

TABLE 3

Eigenvalues of Σ for a Whole Image[a]

	λ_1	λ_2	λ_3
Cylinder	94.0	5.3	0.7
Building	99.2	0.7	0.1
Seaside	82.3	16.7	1.0
Girl	86.0	11.7	2.3
Room	81.8	16.1	2.1
Home	84.2	12.1	3.6
Auto	90.7	7.7	1.7
Face	97.4	2.2	0.4

[a]Rescaled as $\lambda_1 + \lambda_2 + \lambda_3 = 100$.

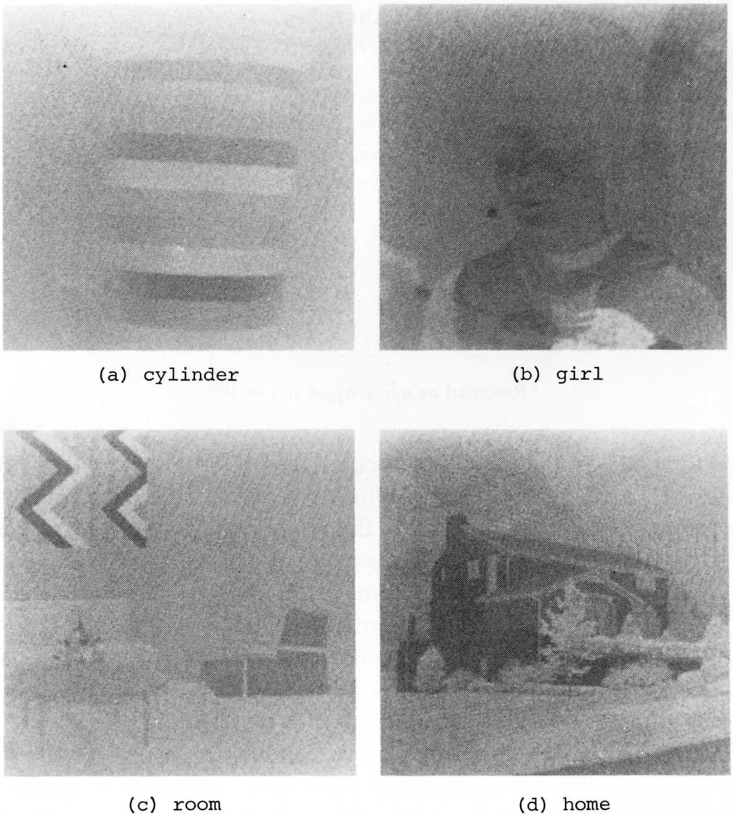


FIG. 17. Images corresponding to X_3.

original color images: (2) There is a tendency for the clarity of color in a small area in the color image to be heavily degraded. The first fact means that the color information in such scenes we have used is almost two dimensional. The second fact can be explained as follows. Figure 17 shows the images corresponding to the feature X_3; (a) cylinder, (b) girl, (c) room, and (d) home. The remarkably low contrast of the images is caused by the small values of λ_3. There exist small areas in them which have a gray value different from the average gray level. These areas have vivid colors which can not be represented by the two principal features X_1 and X_2. Thus, their colors are spoiled by neglecting X_3. To sum up, color images can be represented by using only two spectral features at the cost of spoiling the clarity of the colors within small areas.

4. CONCLUSION

We have made experiments and discussions about color information in region segmentation process. By means of systematic experiments of region segmentation, we have obtained a set of color features $I1 = (R + G + B)/3$, $I2' = (R - B)$, $I3' = (2G - R - B)/2$ which is effective for color image segmentation. The three features are significant in this order and in many cases good segmentations can be achieved by using only the first two. The transformation needed to derive them from R, G, and B is simple and it does not behave poorly when digitized input is used.

Comparisons are made between various sets of color features which are commonly used in image analysis. The characteristics of each set can be observed through the comparative experiments. The difference among segmentation results clearly appears in the case of the cylinder scene in which the separation of the color stripes is difficult. The color feature sets of $\{L, a, b\}$ and $\{I1, I2', I3'\}$ give good results in our experiments. But, in many cases for other scenes, no large difference is observed among the results obtained by using each of the eight sets of color features. We think this is because the color information in natural scenes is almost two dimensional (intensity and one chromatic feature) as shown in the experiment of color image reproduction from only two color features. That is, every set of color features can represent the color information with a fairly large margin for many scenes and can provide enough information for region segmentation. When every set of color features makes little difference in segmentation of a scene, the calculation involved in the coordinate transformation from R–G–B becomes an important factor to consider in determining the effectiveness of color coordinate systems for region segmentation. The set of color features derived in this paper is good in this regard as well as in the segmentation result. We think it is a useful color feature set for color image segmentation.

In this paper, the effectiveness of a color feature set was discussed in the framework of region splitting. The results will be valid in other domains of color image processing such as edge extraction.

REFERENCES

1. R. Nevatia, A color edge detector, in *Proceedings, Third International Joint Conference on Pattern Recognition*, pp. 829–832, 1976.
2. R. Ohlander, *Analysis of Natural Scenes*, Ph.D. thesis, Department of Computer Science, Carnegie–Mellon University, 1975.
3. J. Kender, *Saturation, Hue and Normalized Color: Calculation, Digitization Effects, and Use*, Technical Report, Department of Computer Science, Carnegie–Mellon University, 1976.
4. W. K. Pratt, *Digital Image Processing*, Wiley–Interscience, New York, 1978.