色知覚の３属性によるカラー画像の分割

COLOR IMAGE SEGMENTATION USING THREE PERCEPTUAL ATTRIBUTES

富永 昌治　大阪電気通信大学
Shoji TOMINAGA　Osaka Electro-Communication University

Abstract

This paper describes a method for segmenting a color image into some meaningful regions by means of the three color perceptual attributes of hue, lightness, and saturation. The segmentation technique is based on a recursive thresholding method using the three histograms. The Munsell color system is used as the color space for specifying human color perception. A color specification method is first presented to predict color perception of a measured image. Next, a practical segmentation procedure is described. A sequence of uniform color regions is extracted by a recursive thresholding of the three histograms. Finally, an experiment is made using a flower picture.

1. Introduction

Image segmentation of partitioning an image into a set of meaningful regions is a key step in image analysis. There have been developed several segmentation techniques for color images [1]-[7]. The most basic technique is a histogram thresholding method using multi-dimensional color features. However, the performance of segmentation depends not only on its segmentation algorithm, but greatly on color features used in its segmentation process. In fact, in the earliest work by Ohlander et al., a large set of nine color features was used to segment complex natural images. Ohta et al. conducted a systematic experiment of segmentation, and derived three effective color features in terms of a linear combination of R, G, and B intensities.

The human color perception can normally be described in terms of the three attributes of hue, lightness, and saturation. These attributes are psychological quantities on human impressions of color, which are different from the physically defined primary components of R, G, and B. We can believe that, by analyzing these quantities, we will actually be processing information which is perceptually important for a human observer. For example, hue may be useful for discriminating different materials in an image [8]. So far there have been some proposals to obtain the three perceptual attributes. The methods are based on nonlinear transformations of the observed signals of R, G, and B. Therefore two problems have often been pointed out in using these formulations for image segmentation. One is an insufficient accuracy in representing the perceptually meaningful quantities. The other is caused by computational instability in the nonlinear transformation.

The present paper describes a segmentation method by means of the Munsell three perceptual attributes. The Munsell color system provides a perceptually uniform color space defined in the three attributes called Munsell Hue, Value, and Chroma, which are determined on human perceptual experience of object colors. Although it is not easy to measure the color specifications, the Munsell system has been used widely, and other uniform color spaces for colorimetry have been designed to agree with this system [9]. The author has proposed a mapping method for transforming the observed color signals into the Munsell color system, so that the perceptual attributes of a color sample can be predicted quantitatively [10]. The mapping algorithm is computationally stable for any observations, and generates good color specifications.

2. Three Color Attributes

General meanings of hue, lightness, and saturation are given by CIE [9]. The Munsell color system gives a perceptually uniform color space in which an object color is identified in a cylindrical coordinate system of Munsell Hue, Value, and Chroma.

Hue (H) is the attribute represented by combinations of different wavelengths such as red, yellow, green, and so on. The major hues consist of R (red), Y (yellow), G (green), B (blue), P (purple), and the five half-way hues of YR, GY, BG, PB, and RP. Each major hue is divided into 10 points, and then the total 100-point hue scale is arranged at perceptually equal spacing as a hue circle. In this study, we represent Hue numerically by a real number H which starts with H=0 at 5R, and increases clockwise in the range of [0, 100].

Value (V) corresponds to lightness. The scale gives 10 perceptually equal steps ranging between ideal black (V=0) and ideal white (V=10).

Chroma (C) is similar to saturation or purity, and represents the amount of gray in the color with the same lightness. We have C=0 for an achromatic color, and have an increasing number in equal steps up to the maximum producible with pigments which is usually C=15 at most.

3. Color Specification Process

Fig. 1 shows a pictorial diagram of the color specification process to measure a color image, and predict the color perception in terms of H, V, and C. The objective material is any surface color samples like colored papers or color prints which are flexible. The imaging device is assumed to be a drum scanner which measures the optical densities of B/W (black-and-white), R, G, and B components of the reflected light from the surface. The present color specification process can also be applicable in the same manner to any other imaging devices with R, G, and B signals.

We define the effective reflectances as the observations corresponding to the tristimulus values of an imaging device. This definition is given by the ratios of reflected light intensity to incident light intensity for the respective image sensors as

$$\rho_i = \int E(\lambda)\rho(\lambda)\phi_i(\lambda)d\lambda \Big/ \int E(\lambda)\phi_i(\lambda)d\lambda \quad (1)$$

where

i: B/W, R, G, and B;

$E(\lambda)$: spectral distribution of the lamp;

$\rho(\lambda)$: spectral reflectance of the object;

$\phi_i(\lambda)$: spectral sensitivity of the sensors.

Note that these values are different from the CIE tristimulus values in their normalization. The spectral response curves $\phi_i(\lambda)$ are designed to describe the CIE color-matching functions by their linear transformations. In the real measurement, the effective reflectances can easily be obtained from an exponential conversion of the density signals.

A relationship is found between these reflectances and the perceptual attributes such that a cube-root coordinate system by using $\rho_i^{1/3}$ approximates a perceptually uniform space of the Munsell system. So we can define an observation space of the color signals by a five-dimensional vector

$$\underline{s} = [1, \rho_{B/W}^{1/3}, \rho_R^{1/3}, \rho_G^{1/3}, \rho_B^{1/3}]^T, \quad (2)$$

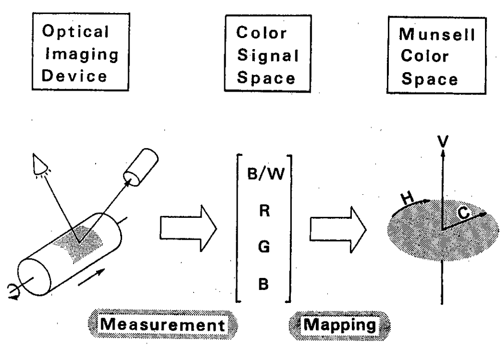

Fig. 1 Color specification process.

where the first element is a constant bias, and T denotes matrix transposition. While we describe the Munsell space (H, V, C) by a vector $\underline{p}$ in the three-dimensional rectangular coordinates.

$$\underline{p} = [C\ \sin(\frac{2\pi}{100}H),\ C\ \cos(\frac{2\pi}{100}H),\ V]^T. \quad (3)$$

Then a mapping from the observation space into the Munsell space is described by a linear equation

$$\underline{p} = \underline{F}\underline{s}. \quad (4)$$

The matrix $\underline{F}$ is a 3x5 transformation matrix defining the mapping, which is determined based on the measurements of many standard-color chips. The practical estimates of the attributes H, V, and C are obtained from the cylindrical representation of the estimated perceptual vector $\underline{p}$.

4. Segmentation Procedure

A color image expansion in terms of the three perceptual attributes can be performed by the mapping of measured signals into the Munsell color space pixel-by-pixel, so that each pixel is described by the three estimates of Hue, Value, and Chroma. The attribute images of Hue, Value, and Chroma can be displayed in accordance with the scale of the Munsell system.

The statistics of each attribute image are represented by a histogram. These three histograms are one-dimensional and mutually independent. However each has an individual feature as shown in Fig. 2. The Hue histogram h(H) is cyclic in a modulus of the domain [0, 100]. The Value histogram h(V) takes zero at both ends of the domain, because the ideal black V=0 and the ideal white V=10 are unrealized. The Chroma histogram h(C) descends steeply as C nears zero. The upper bound of the domain is not fixed.

The present image segmentation is based on a histogram thresholding method which examines independently the three one-dimensional histograms associated with H, V, and C, and then divides recursively a region into subregions of uniform colors by thresholding. The performance of segmentation depends on the method of splitting peaks appearing on the histograms. This histogram splitting process is considered as a cluster detection in the color space, since large regions with similar colors in an image create a cluster in the three-dimensional color coordinate system. Note that the present coordinate system is cylindrical, so that a cylindrical color block is extracted as a cluster. An illustration appears in

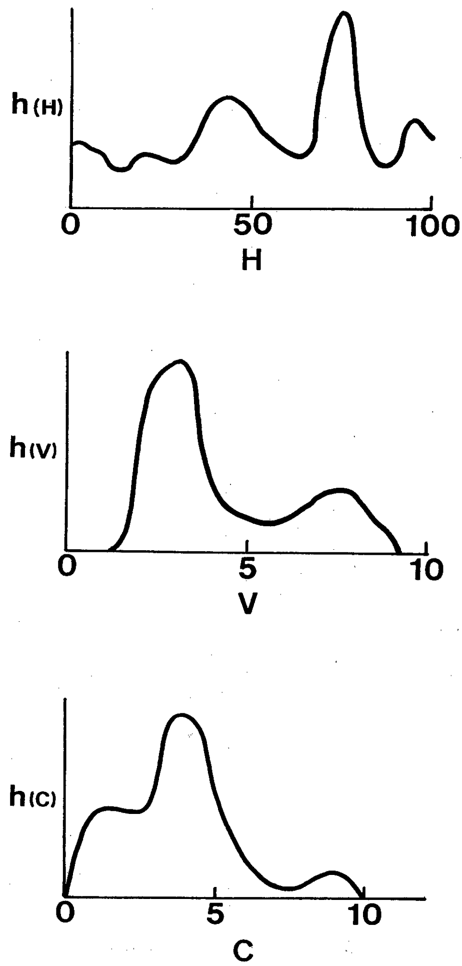

Fig. 2 Typical histograms.

Fig. 3. A color block is fun-shaped for chromatic color. This block is bounded by six coordinate surfaces which are determined by two pairs of vertical plane of constant H, horizontal plane of constant V, and cylindrical surface of constant C. While, the color block for achromatic color is a column which is determined by V and C only.

A practical procedure for image segmentation consists of the following steps.

1. First of all, the entire image is regarded as one region, and histograms are computed for each attribute of H, V, and C. In other cases, this computation is done for the specified regions to be segmented. The histograms are smoothed out by a moving average to eliminate small peaks. The smoothing operation is necessary because of the digitization effects and the noisy observations by a measuring system.

2. The most significant peak is found in a set of three histograms. The peak selection is based on the shape analysis of each peak on the histograms. First, some clear peaks are selected as the candidates from all peaks by means of a simple thresholding of peak parameters such as area and height. Next, a criterion function is

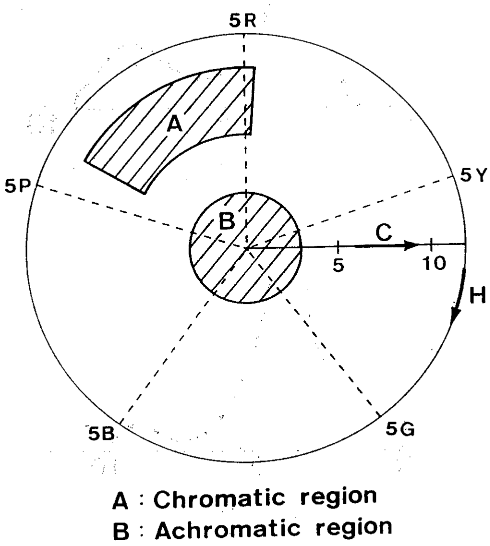

Fig. 3 Color blocks in the color space.

calculated for each peak in the candidate list. The definition is given as

$$f = \frac{S_p}{T_a} \times \frac{100}{fwhm} \tag{5}$$

Fig. 4 illustrates a typical peak to explain the criterion function. The notation S_p represents a peak area between two valleys V_1 and V_2, and FWHM is the full-width at half-maximum of the peak. Moreover T_a denotes the overall area of the histogram, that is, the total number of pixels in the specified image region. In real processing, we normalize all the domains of H, V, and C to [0,100]. Therefore, the criterion function of eq.(5) represents a combination of the largeness (S_p/T_a) of a peak and the sharpness (100/fwhm).

3. Thresholding of a color image is executed using two threshold values derived from the lower bound V_1 and the upper one V_2 for the most significant peak in the set of three histograms. This thresholding operation partitions an image region into two sets of subregions; one consists of subregions corresponding to the color attributes within the threshold limits, and the other is a set of subregions with the remaining attribute values. We extract the former set only, in which each pixel satisfies the peak condition.

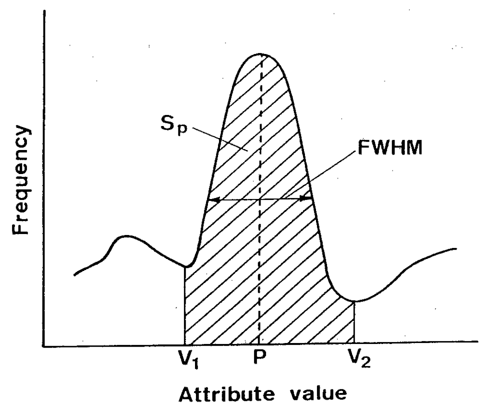

Fig. 4 Criterion for peak finding.

4. The thresholding process is repeated for the extracted subregions. The area of subregions decreases by further thresholding. This process leads to the detection of the most significant cluster in the color space as shown in Fig. 3. If all the histograms become monomodal, the cluster detection is to be finished. So one step segmentation is completed. Then a suitable label is assigned to the latest extracted subregions.

5. The image labeled by the above segmentation is smoothed on the basis of connectedness of pixels. This refinement is intended to smooth out noisy boundaries, and also eliminate small regions and short lines. The 8-connection property is used in our smoothing algorithm. This operation is not binary, but multilevel smoothing. We adopt no region mask to extract only a single connected region. A set of connected regions, which are located separately in an image, may remain for further segmentations. This is because a material dose not necessarily appear as one connected region in a scene, but sometimes is separated by an obstacle, so that the same color appears in different regions.

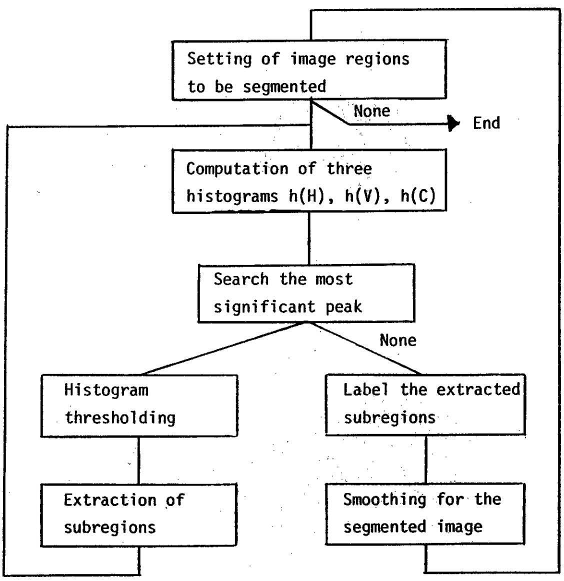

Fig. 5 Image segmentation process.

6. Steps 1 through 5 is repeated for the remaining regions. The segmentation process is terminated when an area of the regions is sufficiently small in comparison to the original image size, or no histogram has significant peaks. The remaining pixels without labeling are regarded as noisy fluctuations, and merged into the neighboring labeled regions of similar colors. The mean values of the color specifications are computed, and a color difference formula is used for choosing the nearest color region.

The above procedure is summarized as a flow chart in Fig. 5.

5. Experimental Results

A picture of a morning glory shown in Fig. 6 has been used. A 250x256 digital image was measured from the photographic paper by using a drum scanner. The pixel size is 0.3x0.3 (mm^2). Next, the measured images for B/W, R, G, and B components have been transformed into the Munsell color space. Fig. 7 shows the resulting color attribute images of Hue, Value, and Chroma. Each image is displayed in the same scale as the Munsell system; the Hue image is painted with 10 colors for the 10 major hue domains,

Fig. 6 Original picture.

and the Value and Chroma images are displayed respectively in the 10 and 13 brightness levels.

Histograms of these attribute images are shown in Fig. 8, where the horizontal axis denotes the attribute values, and the vertical axis gives the frequency of occurrence of each attribute in the corresponding image. The frequency is smoothed by the moving average of a window size 5, and the histogram shape is depicted after taking the logarithmic transformation of such a smoothed frequency.

The features of the three attribute histograms are as follows: The Hue histogram has some clear peaks H_1-H_4. The highest and most sharp peak H_3 corresponds to the background in the original picture. The peak H_1 is to the stems and strings, and H_4 is to the flower. Furthermore, H_2 and the succeeding relatively flat continuum correspond to the

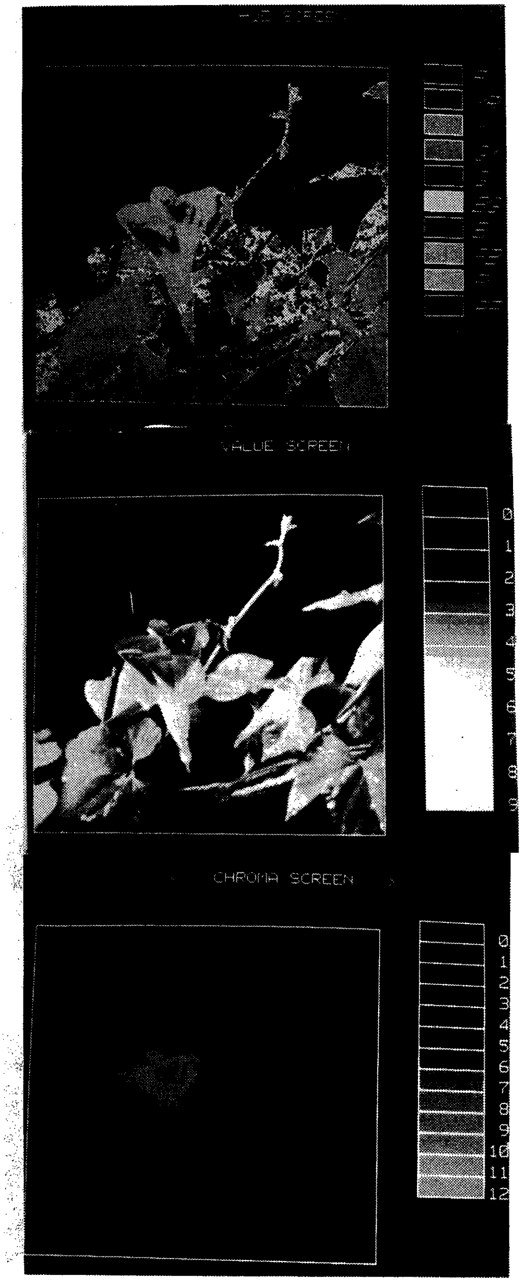

Fig. 7 Three attribute images

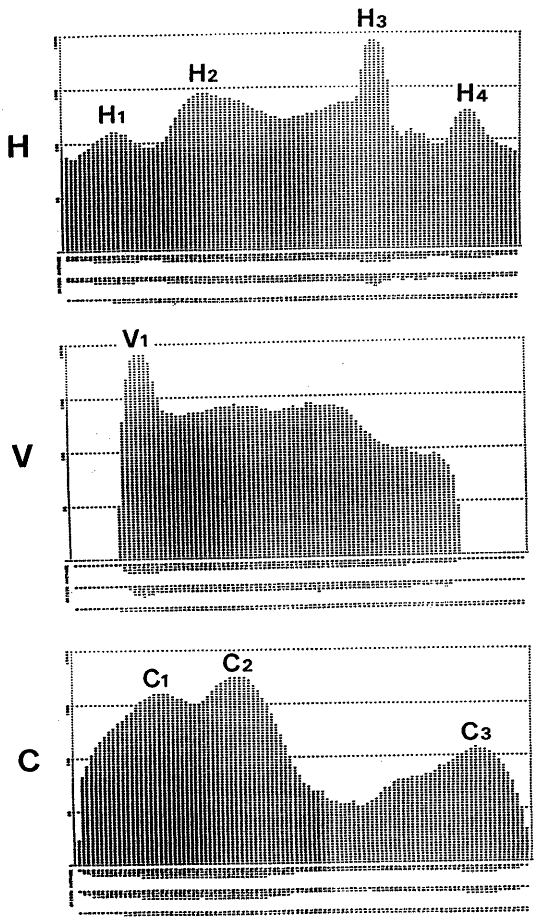

Fig. 8 Histograms of three attribute images.

leaves. The Value histogram has a single major peak V_1 which corresponds to the background. The Chroma image extracts clearly vividness of the flower. Its statistics are given by the peak area of C_3.

It is essential to recognize a great difference between the features by the two sets of (R, G, B) and (H, V, C). Fig. 9 shows the three histograms by the effective reflectances of R, G, and B for the same image. Note that all the histograms have nearly the same shape. A comparison of Figs. 8 and 9 shows a clear difference between the multimodal representation of the H, V, and C histograms, and the monomodal representation of the R, G, and B histograms. The former is much easier for thresholding. Thus the three color attributes H, V, and C are effective for the cluster detection and the image segmentation.

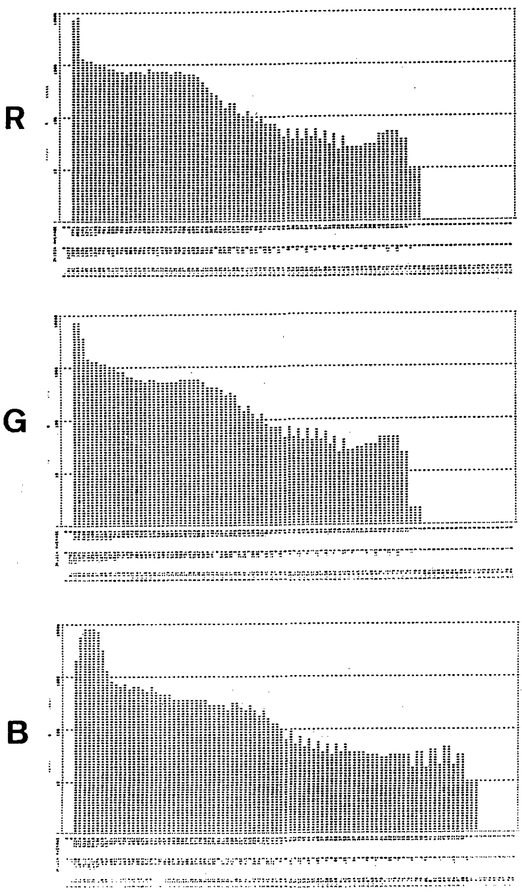

Fig. 9 Histograms of R, G, and B images.

Fig. 10 shows a segmentation result by applying the histogram thresholding procedure to the above histograms. The original image is segmented into 10 uniform color regions. A detailed inspection of the figure shows that the main elements of flower, stems, bud of flower, and background are well partitioned, and moreover most of the details in the original image are maintained. One problem has occurred at the lower part of flower, where white of the flower is merged with the right leaf. We have computed the average of the Munsell color specifications for each of the extracted color regions. These values are useful for the representative color specifications for the segmented image.

6. Conclusion

A method has been developed for segmenting a color image into subregions with perceptually uniform colors by means of the three color perceptual attributes. The color specification process was first presented to measure a color image, and predict the color perception in terms of Munsell Hue, Value, and Chroma. An image segmentation was executed based on a recursive thresholding method using the three attribute histograms. This operation corresponds to the recursive detection of a compact cluster in a color specification space. Experimental results using a flower picture have demonstrated the usefulness of the three color attributes and moreover the feasibility of the proposed procedure.

References

[1] R. Ohlander, K. Price, and D.R. Reddy, Picture segmentation using a recursive

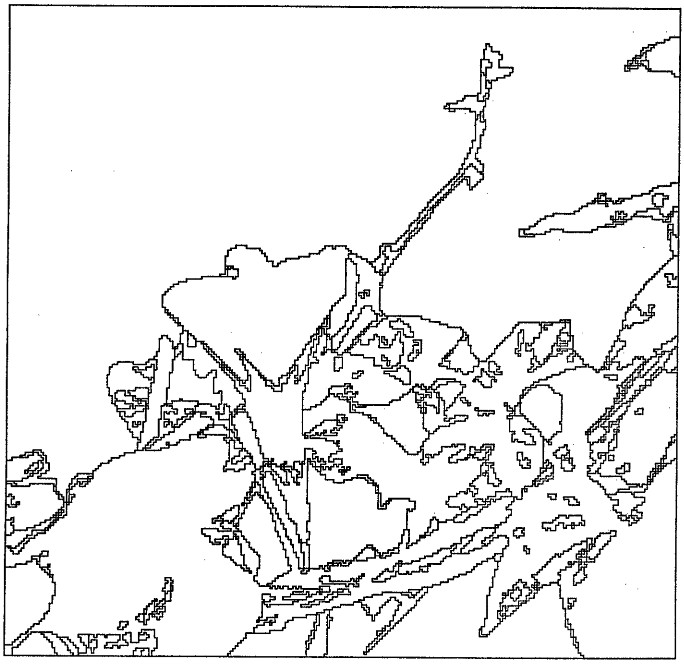

Fig. 10 Segmentation result.

region splitting method, Computer Graphics and Image Processing, vol. 8, pp. 313-333, 1978.

[2] A.R. Hanson and E.M. Riseman, Segmentation of natural scenes, Computer Vision System, New York: Academic Press, pp. 129-163, 1978.

[3] A. Sarabi and J.K. Aggarwal, Segmentation of chromatic images, Pattern Recognition, vol. 13, pp. 417-427, 1981.

[4] Y. Ohta, T. Kanade, and T. Sakai, Color information for region segmentation, Computer Graphics and Image Processing, vol. 13, pp. 224-241, 1980.

[5] M.D. Levine and S.I. Shaheen, A modular computer vision system for picture segmentation and interpretation, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. PAMI-3, pp. 540-556, 1981.

[6] N. Funakubo, Region segmentation of biomedical tissue image using color texture feature, Proc. 7th Int. Cof. on Pattern Recognition, pp. 30-32, 1984.

[7] J.M. Chassery and C. Garbay, An iterative segmentation method based on a contextual color and shape criterion, ibid, pp. 642-644, 1984

[8] J.M. Rubin and W.A. Richards, Color vision: Representing material categories, MIT A.I. MEMO, No.764, 1984.

[9] G.W. Wyszecki and W.S. Stiles, Color Sciences, 2nd. ed., New York: John Wiley&Sons, 1982.

[10] S. Tominaga, A mapping method for computer color vision based on the theory of colorimetry, Trans. of IPS of Japan, vol. 26, pp. 318-328, 1985.