

# A New Method for Color Feature Representation of Color Image in Content-Based Image Retrieval – 2D Projection Maps

Seok-Wun Ha, *Member, KIMICS*

**Abstract**—The most popular technique for image retrieval in a heterogeneous collection of color images is the comparison of images based on their color histogram. The color histogram describes the distribution of colors in the color space of a color image. In the most image retrieval systems, the color histogram is used to compute similarities between the query image and all the images in a database. But, small changes in the resolution, scaling, and illumination may cause important modifications of the color histogram, and so two color images may be considered to be very different from each other even though they have completely related semantics. A new method of color feature representation based on the 3-dimensional RGB color map is proposed to improve the defects of the color histogram. The proposed method is based on the three 2-dimensional projection map evaluated by projecting the RGB color space on the RG, GB, and BR surfaces. The experimental results reveal that the proposed is less sensitive to small changes in the scene and that achieve higher retrieval performances than the traditional color histogram.

**Index Terms**—Color, Image, Retrieval, Feature, Representation, Content-based, 2D, Projection, and Maps.

## I. INTRODUCTION

Color is a very important feature in extracting information from a color image, and color histogram comparison has recently become a popular technique for image or video retrieval [1]. The color histogram comparison is used as a preliminary step for database indexing in order to reduce the number of candidate images for next steps which could use other features (e.g. texture, shape) to compare database images with a given query image. In general, the RGB color space is inappropriate for image retrieval due to the fact that small changes in the resolution, scaling, and illumination may cause important modifications of the color histogram, and two color images may be considered to be very different from each other even though they have completely related semantics [2], and so other color spaces like HSI, YIQ, or LUV are often used for image retrieval [3, 4, 5].

Actually, the RGB color space is related to the color signal from the image capture devices like camera, video

camera, or scanner, and so it is need to use directly the RGB color space without a color space translation to other color spaces [1].

In this paper, 2 dimensional projection maps that is projected on the RG, GB, and BR surfaces from the RGB color space as a new color feature creation for image retrieval. When these proposed is used to compute the similarity distance between two images, these three 2-dimensional projection maps are connected and are transformed to a 1-dimensional projection vector.

This paper is organized as follows. The traditional color histogram method is described in Section 2, and the proposed 2-dimensional projection map method is described in Section 3. Some experimental results and comparisons are shown in Section 4, and then some concluding remarks are then presented in Section 5.

## II. TRADITIONAL COLOR HISTOGRAMS

Usually, color histogram calls histogram from RGB color space. A color histogram denotes the joint probabilities of the intensities of the three color channels. The color histogram is defined as follows:

$$h_{r,g,b}[r, g, b] = N \cdot \text{Pr}\{R=r, G=g, B=b\} \quad (1)$$

Where R, G and B are the three color channels and N is he number of pixels in the image [6].

The color histogram is computed by counting the number of pixels of each color within the image. Since the number of colors is finite, it is usually more convenient to transform the three channel histogram into a single variable histogram. Given an RGB image, one transform is given by  $m = r + N_r g + N_r N_g b$ , where  $N_r$ ,  $N_g$  and  $N_b$  are the number of bins for colors red, blue and green, respectively. This gives the single variable histogram by. For example, fig. 1 shows color histogram of a sunset image with quantization rate of 5 bits.

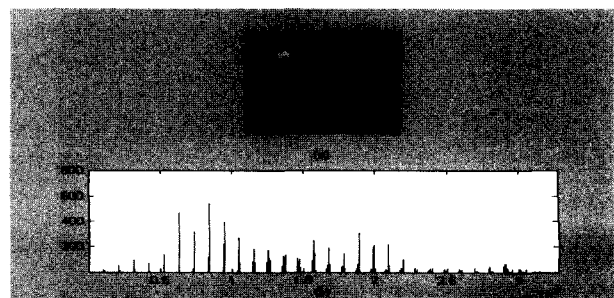


Fig. 1 A sunset image and its color histogram

Manuscript received April 7, 2004.

Seok-Wun Ha is with the Department of Computer Science, Gyeongsang National University, Jinju, Gyeongnam, Korea 660-701. (email:swaha@nongae.gsnu.ac.kr). He is a researcher in Research Institute of Computer Information & Communication.

This color histogram has a defect that a small change of color in the image may cause a great modification of the color histogram, and also color histogram may vary greatly by changes in lighting and large view angle. An effective way to retain light-independent color properties is to use only hue and saturation in HSI color space, and some kinds of improved methods to extract color features from the histogram of Hue component have studied [7, 8]. But because an additional mathematical conversion processing from RGB space to HSI space is requested to use the Hue histogram, it is need to use directly the RGB color space without a color space translation to other color spaces for fast retrieval from the image database. The histogram of the hue component is defined as follows:

$$h_H[h] = N \cdot \text{Pr ob}\{H = h\} \quad (2)$$

Where H is the hue component and N is the number of pixels in the image and Hue component has one value between 0 and 1. Fig. 2 shows the histogram of the hue component of the sunset image in fig. 1.

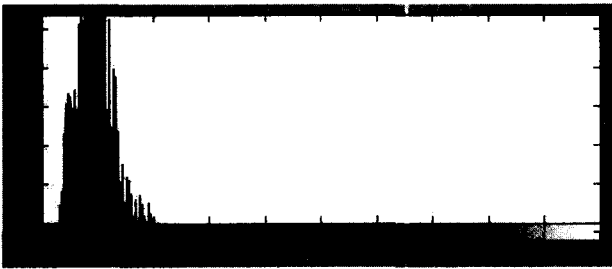


Fig. 2 Histogram of hue component of the sunset image

Because these traditional color histograms relate in frequency of pixels that have equal color, color histogram changes greatly to slight change of image contents. But, to slight change of image contents, if there is identical color continuously, feature of color may not change. In this study, we propose a new color feature representation method emphasizing in existence of such color.

### III. 2D PROJECTION MAPS

A new method that is proposed to represent color features of a color image efficiently is changing 3-dimensional RGB color space map to 2-dimensional projection maps. Fig. 3 shows 3-dimensional RGB color space map representing the three brightness values in R, G, and B channels of the sunset image.

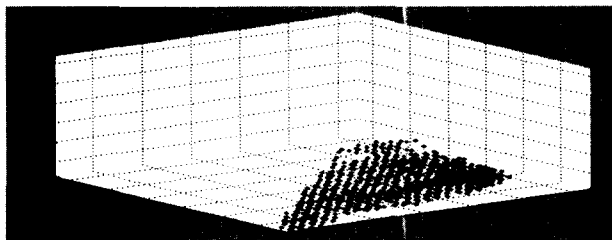


Fig. 3 3-dimensional RGB color space map of the sunset image in Fig. 1

Projecting the 3-dimensional color space map to 2-dimensional RG, GB, and BR surfaces make three 2-dimensional projection maps. First, to generate the projection map to RG surface, correspond the integer value of the R channel brightness to the x coordinate bin and the integer value of the G channel brightness to the y coordinate bin by one to one for one pixel in the image, then the position (x, y) in the 2-dimensional map has value of 1, and then the same procedure for the all pixels can make a 2-dimensional projection map. Through equal process, other projection maps to GB surface and BR surface,  $P_{RG}$ ,  $P_{GB}$ , and  $P_{BR}$  are formed as follows:

$$\begin{aligned} P_{RG}(x, y) &= \begin{cases} 1, & \text{at } x = v(r), y = v(g) \\ 0, & \text{at other bins} \end{cases} \\ P_{GB}(x, y) &= \begin{cases} 1, & \text{at } x = v(g), y = v(b) \\ 0, & \text{at other bins} \end{cases} \\ P_{BR}(x, y) &= \begin{cases} 1, & \text{at } x = v(b), y = v(r) \\ 0, & \text{at other bins} \end{cases} \end{aligned} \quad (3)$$

Where,  $v(r)$ ,  $v(g)$ , and  $v(b)$  are brightness values of one pixel in the image. Three 2-dimensional projection map of the sunset image are presented in Fig. 4.

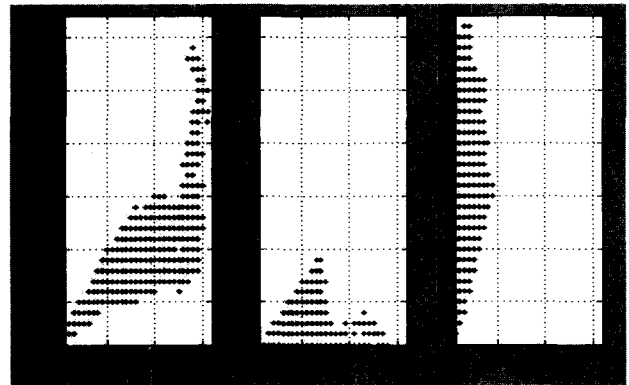


Fig. 4 Three 2-dimensional projection maps.

In the content-based image retrieval, in order to measure similarities between the query image and the object images in the database, these three 2-dimensional projection maps are connected and transformed to a 1-dimensional projection vector as follows:

$$I = [P_{RG}(1,1), \dots, P_{RG}(2^q, 2^q), P_{GB}(1,1), \dots, P_{GB}(2^q, 2^q), P_{BR}(1,1), \dots, P_{BR}(2^q, 2^q)] \quad (4)$$

Where, I is 1-dimensional projection vector, q is the number of quantization bits per pixel and q is 5 here.

Therefore, length of the 1-dimensional projection vector is  $2^q \times 2^q \times 3$ . Because length of the color histogram is  $2^q \times 2^q \times 2^q$ , length rate between color histogram and 2-dimension projection map is  $\frac{2^q}{3}$ , and in case of q=5, the

length rate is about 10.7. Therefore, we can find that the processing speed reduces greatly and will appear for the performance through follow experiments.

### IV. EXPERIMENTS AND RESULTS

The proposed method of color feature representation was compared with the traditional color histograms. Their performance to retrieve similar images from a heterogeneous collection of images has been used as comparison criterions. For the color histograms, because values that compose the color histogram are analogous, the Euclidean distance method is used to measure similarities between a query image and candidate images.

For the RGB color histogram and the histogram of Hue component, when the feature vector of the query image is  $C_p$  and one of the candidate image is  $C_c$ , the distance between these two images,  $ED_{histo}$ , is computed as follows.

$$\begin{aligned}
 C_p &= \{C_p(1), \dots, C_p(n)\} \\
 C_c &= \{C_c(1), \dots, C_c(n)\} \\
 ED_{histo} &= \sqrt{\sum_{i=1}^{2^q \times 2^q \times 2^q} (C_p(i) - C_c(i))^2}
 \end{aligned}
 \tag{5}$$

Where,  $n$  is length of histogram and is  $2^q \times 2^q \times 2^q$ .

For the 2-dimensional projection map, because its 1-dimensional projection vector is a bit pattern of binary value, the Hamming distance was used to measure similarities between a query image and candidate images., when the feature vector of the query image is  $I_p$  and one of the candidate image is  $I_c$ , the distance between these two images,  $HD_{2d}$ , is as follows.

$$\begin{aligned}
 I_p &= \{I_p(1), \dots, I_p(n)\} \\
 I_c &= \{I_c(1), \dots, I_c(n)\} \\
 HD_{2d} &= \sum_{i=1}^{2^q \times 2^q \times 3} abs(I_p(i) - I_c(i))
 \end{aligned}
 \tag{6}$$

Where,  $n$  is length of 1-dimensional projection vector and is  $2^q \times 2^q \times 3$ .

To estimate their performance, 5 images that the color contents are similar is selected from database that stores 100 sunset images and 5 images selected for querying are represented in fig. 5.



Fig. 5 Query images with similar colors

When these similar images was used as query images in turn, the experimental results were compared each other. First, when the first image of the selected sunset images was presented as a query image, fig. 6 shows the results of, (a) the color histogram, (b) the Hue histogram, and (c) the proposed 2-dimensional projection maps.

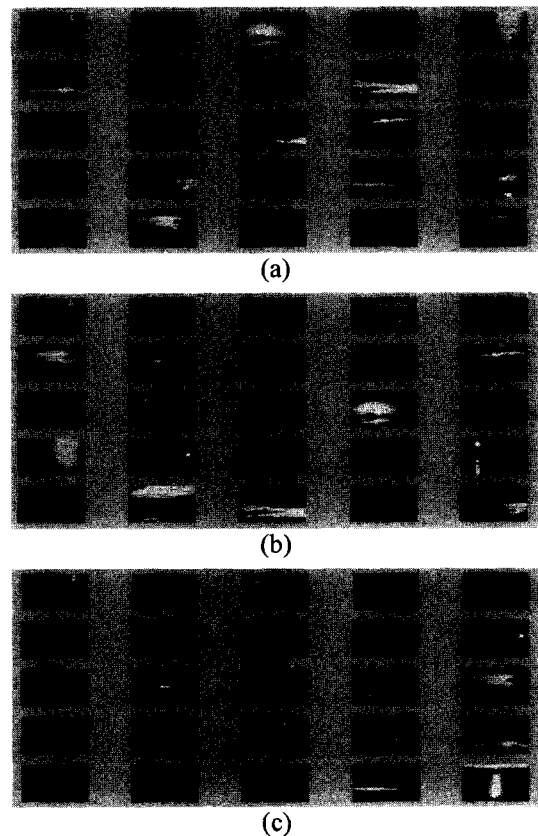


Fig. 6 Results of retrieval for first sunset image as a query, (a) the color histogram, (b) the Hue histogram, and (c) the proposed.

To compare their retrieval performance, for each color representation method from result of fig. 1, the number of query images that appear in priority order 1 to 5 was counted. Then, equal process was done for other query images and number of other query images that appear in priority order 5 was counted. The result was compared in table 1.

From table 1, average ratios of retrieval performance about each method are as following: 36% for color histogram, 56% for Hue histogram, and 76% for 2d projection map.

Table 1. Comparison of retrieval results for each color representation methods (in priority order 5).

Query	Methods	Represent ation Number (N)	Efficiencies $\frac{N}{5} \times 100(\%)$
q1	Color histogram	2	40
	Hue histogram	3	60
	The proposed	4	80
q2	Color histogram	2	40
	Hue histogram	2	40
	The proposed	4	80
q3	Color histogram	2	40
	Hue histogram	3	60
	The proposed	3	60
q4	Color histogram	1	20
	Hue histogram	2	40
	The proposed	4	80
q5	Color histogram	2	40
	Hue histogram	4	80
	The proposed	4	80

Then, equal process was done for other query images and number of other query images that appear in priority order 10 was counted. The result was compared in table 2.

For the number of other query images that appear in priority order 10, average ratios are as following: 44% for color histogram, 64% for Hue histogram, and 96% for 2d projection map.

Table 2. Comparison of retrieval results for each color representation methods (in priority order 10).

Query	Methods	Representat ion Number (N)	Efficiencies $\frac{N}{5} \times 100(\%)$
q1	Color histogram	3	60
	Hue histogram	3	60
	The proposed	5	100
q2	Color histogram	2	40
	Hue histogram	3	60
	The proposed	4	80
q3	Color histogram	2	40
	Hue histogram	3	60
	The proposed	5	100
q4	Color histogram	2	40
	Hue histogram	2	40
	The proposed	5	100
q5	Color histogram	2	40
	Hue histogram	5	100
	The proposed	5	100

Therefore, from the results of above experiments, when the 2d projection map that was proposed for representing color features of a color image was used in image retrieval, we found that the proposed has better performances than other two methods. We also compared their retrieval performances for flower, animal, nature, and plane images and same result was obtained. In fig. 7, fig. 8, fig. 9, and fig.10, we could know that the proposed 2d projection map could express better color characteristic of image.



Fig. 7 Results of retrieval by (a) the color histogram, (b) the Hue histogram, and (c) the 2-dimensional projection maps for flower images.



Fig. 8 Results of retrieval by (a) the color histogram, (b) the Hue histogram, and (c) the 2-dimensional projection maps for animal images.

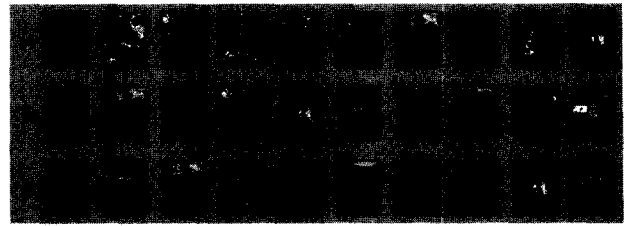


Fig. 9 Results of retrieval by (a) the color histogram, (b) the Hue histogram, and (c) the 2-dimensional projection maps for nature images.

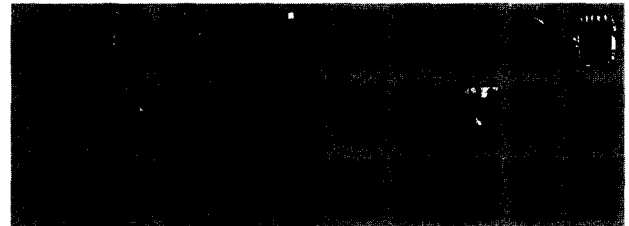


Fig. 10 Results of retrieval by (a) the color histogram, (b) the Hue histogram, and (c) the 2-dimensional projection maps for plane images.

#### IV. CONCLUSIONS AND FUTURE WORK

To improve color image retrieval performance using color feature, the 2-dimensional projection maps method has been proposed. For the given heterogeneous image collections, some experimental results show that the proposed method has better performance than the color histogram and the Hue histogram in image retrieval.

The proposed 2-dimensional projection maps are projections of RGB color space to three RG, GB, and BR surfaces, and they have more precise color information than the color histogram or Hue histogram.

Hereafter, a study for comparison between the proposed method and the traditional color feature creation methods would be done in image classification fields.

#### REFERENCES

- [1] A. B. Bimbo, Visual information retrieval, Morgan Kaufmann Publishers, Inc. pp. 81-116, 1999.
- [2] S. Sablak and T. E. Boulton, "Multilevel Color Histogram Representation of Color Images by Peaks for Omni Camera," Proceedings of the IASTED International Conference Signal and Image Processing, pp. 18-21, Oct., 1999.
- [3] A. Vellaikal and C. C. J. Kuo, "Content based image retrieval using multiresolution histogram representation," SPIE Digital Image Storage and Archiving Systems, vol. 2606, pp. 312- 323, 1995.
- [4] H. J. Zhang, Y. Gong, C. Y. Low, and S. W. Smoliar, "Image retrieval based on color feature: an evaluation study," SPIE Digital Image Storage and Archiving Systems, vol. 2606, pp. 212- 220, 1995.
- [5] X. Wan and C. C. J. Kuo, "Pruned octree feature for interactive retrieval," SPIE Multimedia Storage and Archiving Systems II, vol. 3229, pp. 182-193, 1997.

- [6] J. R. Smith and Shih-Fu Chang, "Tools and techniques for color image retrieval," <http://www.ctr.columbia.edu/~jrsmith/html/pubs/tatfeir/color.html>, Mar. 1996.
- [7] S. Sural, Gang Qian, and S. Pramanik, "Segmentation and histogram generation using the HSV color space for image retrieval," Proceedings of 2002 International Conference on Image Processing, Vol. 2, pp. 589-592, Sep. 2002.
- [8] Y. Tao and w. I. Grosky, "spatial color indexing: a novel approach for content-based image retrieval," *iee international conference on multimedia computing and systems*, vol. 1, pp. 530-535, jun. 1999.

**Seok-Wun Ha**

Received B.S., M.S., and Ph.D. degrees in Electronics Engineering from Pusan National University, Korea, in 1979, 1985, and 1995 respectively. He is with Gyeongsang National University and is currently an associate professor in the Dept. of Computer

Science. He has been a visiting researcher in University of California, Riverside, in 2002. He works as a researcher at Research Institute of Computer Information & Communication and Engineering Research Institute in Gyeongsang National University. His Major is computer vision and his interests are image processing, image retrieval, neural network, and data visualization.