

COLOR IMAGE FEATURE EXTRACTION USING COLOR INDEX LOCAL AUTO-CORRELATIONS

Takumi Kobayashi and Nobuyuki Otsu

National Institute of Advanced Industrial Science and Technology
Tsukuba, 1-1-1 Umezono, Japan

ABSTRACT

In this paper, we propose a method for extracting color image features, called color index local auto-correlations. Pixel color is quantized and described sparsely in a manner similar to the color indexing of color histograms. In addition, by utilizing spatial auto-correlations of the color indexes, the characteristics of color texture can be extracted more effectively than ordinary histogram-based methods. The proposed method has variants in terms of the color space and basic colors used for indexing colors. These various settings are comprehensively compared in the experiments of image retrieval and image classification, and are shown to exhibit favorable results compared to the other conventional methods.

Index Terms— Feature extraction, Color auto-correlation, Color indexing, Image recognition

1. INTRODUCTION

Color is an essential and typical representation for images and it is also an important key for distinguishing objects. Therefore, extracting features from color images is necessary in many tasks of computer vision, e.g., object recognition [1], image retrieval [2], and image matching [3]. In addition, humans perceive various impressions from color images, such as paintings and photographs. The relationships between color features and human perception are useful for human-computer interactions [4]. However, it is difficult to effectively extract features from color images. Several approaches have been so far made for color feature extraction.

In the color histogram [5] which has been a popular approach, pixel colors are quantized into several basic colors; this process is called *color indexing*. Since geometric characteristics are hardly extracted by the color histogram, color coherence vector (CCV) [2] extends the color histogram by considering spatial color distributions. In CCV, for each color index, i.e., the color associated with each histogram bin, the connected regions in the color are detected in the image and then the value of the histogram bin is divided into two components corresponding to the pixel counts in connected and disjoint regions, respectively. Thereby, the spatial distribution for each color index can be roughly characterized although

local patterns, such as texture, are not captured. Color higher-order local auto-correlation (Color HLAC) [4] extracts local spatial information by using the auto-correlations of RGB values of local pixels. However, various color patterns tend to result in the same values of auto-correlations (actually multiplications) of RGB values. Such ambiguity in the representation yields the degeneration in Color HLAC features. Thus, the representations by raw RGB values are considered to be inappropriate for use of auto-correlations. In addition, the other color representation, such as cyclic Hue value in HSV color space, is difficult to be dealt with in the Color HLAC.

In this paper, we propose a method of color feature extraction by combining the representations in color histograms and auto-correlations. In the method, colors are quantized and described sparsely by using various basic colors in a manner similar to the color indexing of color histograms, and then the auto-correlations of the color indexes is calculated; we call this method color index local auto-correlation (CILAC). Thus, the method can extract *co-occurrence* of colors which is related to color texture and contains richer information than *occurrence* in ordinary method based on color histogram. In addition, the ambiguities regarding Color HLAC described above can be reduced by using the sparse representations in the indexed colors. Although the dimensionality of CILAC increases, the computational cost is lower than that of Color HLAC in virtue of the sparse representation of colors.

various settings both of color space and color indexing are comparatively studied through the experiments of image retrieval and image classification.

2. COLOR INDEX LOCAL AUTO-CORRELATION

An overview of the proposed method CILAC is shown in Fig. 1. It consists of the following three steps. First, the pixel values, usually RGB values, are mapped into a color space, such as $L^*a^*b^*$ and HSV (Hue, Saturation and Value). Second, the pixel values in the color space are indexed based on several basic colors as is similar to the color histogram methods, and described as a sparse vector c . Finally, the auto-correlations of the vector pairs in local neighbors are calculated. We describe the details of each step in the following sections.



Fig. 1. Overview of the proposed method (CILAC)

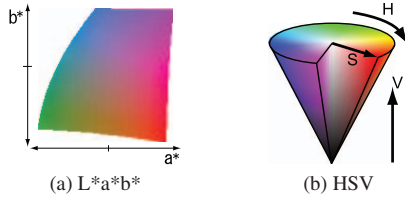


Fig. 2. Color spaces. (a) Slice image at $L^*=54$ in $L^*a^*b^*$, (b) HSV (best viewed in color)

2.1. Color Representing Spaces

We employ three types of color space: **RGB**, **$L^*a^*b^*$** and **HSV**. RGB is a typical representation for pixel color values. $L^*a^*b^*$ approximates human perception for colors and is constructed so as to be perceptually uniform (Fig. 2(a)). The metric in $L^*a^*b^*$ space reflects human perception. In HSV, colors are represented rather categorically, which is similar to the way of human perceiving colors. A cone-type HSV color space is used in this paper, as shown in Fig. 2(b). Hue is cyclically described as the circumference around the central cone axis. The color spaces of $L^*a^*b^*$ and HSV have been applied to several tasks [6, 7].

2.2. Color Indexing in CILAC

In order to consider co-occurrence by using auto-correlation, which is described in the next section, each pixel color represented in the color space is indexed into basic colors (D in number) by voting weights to several nearest basic colors. Then, the pixel color is described as D -dimensional vector c consisting of the weight values. The vector c has a sparse representation, i.e., only a few elements have non-zero values, which is similar to the categorical representation, e.g., “red” or “green.” Color indexing also establishes robustness to noise. The basic colors in the color space are determined as follows.

Predefined Colors (P7/P8/P15) We define several basic colors *a priori* based on knowledge about the color space. This indexing is abbreviated to **P7/P8/P15** of which digits indicates the number (D) of basic colors.

In RGB, basic colors are defined based on the cubic shape of color space. P8 consists of eight vertex points of the color cube, which are denoted as points of a square in Fig. 3(a). In addition to these eight basic colors, for P15, seven more basic colors are selected as well. These additional basic colors represent ambiguous colors: six centers on the faces and one at the center of the color cube in RGB, denoted as star points in Fig. 3(a).

In $L^*a^*b^*$ color space, we simply derive basic colors from those of RGB. The eight or fifteen basic colors (P8/P15) in RGB space are mapped into the $L^*a^*b^*$ space. Since the $L^*a^*b^*$ space is distorted due to its perceptual uniformity, the basic colors are not so equally located as those in the RGB space.

In HSV, basic colors are defined based on the cone shape of the color space. On the upper circle, six equally divided points (red, yellow, green, cyan, blue, magenta) and one white color are chosen, as depicted in Fig. 3(b). Only these seven basic colors disregarding V-values can compose the basic colors, denoted as P7. In the P7 indexing, on the assumption that darker pixels have less information, darker pixels are ignored as is the case with the other color features based on raw pixel values, such as Color HLAC [4], which is described later. On the other hand, P8 is simply constructed by adding black color, i.e., the point of the cone, to these seven basic colors, shown as square points in Fig. 3(c). For P15, seven more basic colors are added to P8, representing ambiguous colors: six equally divided points and one at the center in the middle circle in HSV, denoted as star points in Fig. 3(c).

Cluster Center Colors (C8/C15) In this indexing, the basic colors are determined *a posteriori* using a given dataset. Since the pixel colors may not actually be distributed so evenly in the color space, some of the predefined basic colors (Fig. 3) may be useless for color indexing. Therefore, by taking into account such bias of color distributions, pixel colors in all images of the given dataset are clustered, e.g., by k-means, into eight or fifteen clusters for basic colors **C8/C15**, respectively. The basic colors are determined as the cluster centers, all of which contain pixel samples, and are useful for indexing.

In color indexing, a pixel color can be sparsely described on the basis of the basic colors to which voting weights are calculated in the following two ways.

In the case of predefined basic colors in HSV color space, we employ the (tri-)linear weights to the nearest basic colors along HSV values. For instance, when using the basic colors of P8 (square points in Fig. 3(c)), pixel color $p = (h, s, v)$, of which $s, v < 1$, is described as

$$p = (1 - v)\mathbf{b} + v(1 - s)\mathbf{w} + vs \frac{h_R - h}{h_R - h_L} \mathbf{e}_L + vs \frac{h - h_L}{h_R - h_L} \mathbf{e}_R$$

where \mathbf{b}, \mathbf{w} are black and white colors, and $\mathbf{e}_L, \mathbf{e}_R$ are the nearest adjacent basic colors on the circumference (Fig. 3(b)) and their H-values are h_L, h_R ($h_L < h < h_R$), respectively. The coefficients for the nearest basic colors ($\mathbf{b}, \mathbf{w}, \mathbf{e}_L, \mathbf{e}_R$) are the weights which are definitely summed to 1. Only for P7 in HSV, the coefficient of black color (\mathbf{b}) is disregarded

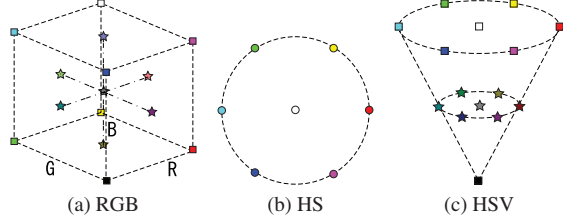


Fig. 3. Configuration of predefined basic colors

and thus the sum of the weights is not necessarily equal to 1, especially for the darker colors.

In the case of other indexing, k -nearest neighbor method is applied for calculating the voting weights. Consider k nearest basic colors with distances d_1, \dots, d_k (say, $k = 4$). By determining the largest weight for the nearest basic color as $w_1 = 1$, the $i (> 1)$ -th weight is defined as $w_i = \frac{d_1}{d_i}$. These weights are finally normalized by the sum: $\hat{w}_i = \frac{w_i}{1 + \sum_{m=2}^k w_m} = \frac{\prod_{j \neq i} d_j}{\sum_{m=1}^k \prod_{j \neq m} d_j}$, which are also summed to 1.

As a result, pixel color is described by means of D -dimensional vector c of which element values associated to the nearest basic colors are set to the voting weights and the others are 0. Thus, c is a sparse vector.

2.3. Auto-Correlations of CILAC

By using the color descriptor c , the zeroth and first order auto-correlations of colors are defined as follows:

$$\text{0th order } R_0(i) = \sum_{\mathbf{r}} c_i(\mathbf{r}) \quad (1)$$

$$\text{1st order } R_1(i, j, \mathbf{a}) = \sum_{\mathbf{r}} c_i(\mathbf{r}) c_j(\mathbf{r} + \mathbf{a}), \quad (2)$$

where \mathbf{a} is a displacement vector from the reference point $\mathbf{r} = (x, y)$, c_i is the i -th element value of the D -dimensional vector c , and $i, j \in \{1, \dots, D\}$. We define $\mathbf{a} \in \{(\Delta r, 0), (\Delta r, \Delta r), (0, \Delta r), (-\Delta r, \Delta r)\}$ where Δr indicates local spatial interval [8]. The configurations of $(\mathbf{r}, \mathbf{r} + \mathbf{a})$ are shown in Fig. 4.

The first order CILAC can extract co-occurrence of colors which is related to local texture, while the zeroth order corresponds to an ordinary color histogram. The first order (Eq.(2)) contains two kinds of auto-correlation: spatial auto-correlations derived from displacement vector \mathbf{a} , and color index auto-correlations derived from the products of the element values c_i, c_j . This is a formulation similar to that of Color HLAC, but we employ the auto-correlations of the quantized data (c), not the raw RGB values used in Color HLAC. This is based on the empirical fact that, in [8], the auto-correlations of binary values, i.e., quantized data, are better for establishing recognition than those of the pixel values themselves. Although the dimensionality of CILAC is high ($D + 4D^2$), the computational cost is rather low in virtue of the sparseness of c : the operations of Eq.(1) and

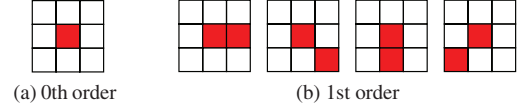


Fig. 4. Spatial auto-correlation patterns



(a) N-S dataset [9]



(b) Photograph and paintings

Fig. 5. Example images used in the experiments

(2) are applied to only a few non-zero elements of c . This is practically performed by storing only non-zero weights and their color indexes of c (sparse representation).

3. EXPERIMENTAL RESULTS

We apply the proposed method CILAC to two kinds of task: image retrieval [9] and image classification. The performances are compared to those of the other methods: CCV [2] and Color HLAC [4]. CCV can be applied to any color spaces and the number of basic colors is 64, as in [2]. The other parameters in CCV are determined so as to yield the best results as well as the spatial interval Δr in CILAC and Color HLAC. Note that Color HLAC allows only for RGB color space. Although the combination of color and shape image features, like HueSIFT [10], might work better in these experiments, any shape image features are not applied for accurate comparison of the color image features themselves.

For image retrieval, N-S dataset [9] is used. This is composed of 2550 objects and scenes, each of which is taken from four different viewpoints. Hence, the dataset contains 10200 images of 640×480 pixels (Fig. 5(a)). The images are retrieved by using extracted color feature vectors according to the distance from the query. The measure for evaluating performance is the average number of correct images among the four first images returned for the given query, as described in [9]. The results of the performances are shown in Fig. 6(a). By comparing the results of P7 and P8 in HSV, it is shown that the addition of the “black” basic color improves the performance. Thus, the black color has the discriminative information and should not be neglected. Clustering seems to be effective for color indexing in RGB and, especially, in $L^*a^*b^*$. In $L^*a^*b^*$ space, the predefined basis colors are not uniformly located, which would be ineffective for indexing. By applying clustering method, the cluster center colors can correctly capture the characteristics of color distribution in that space. In

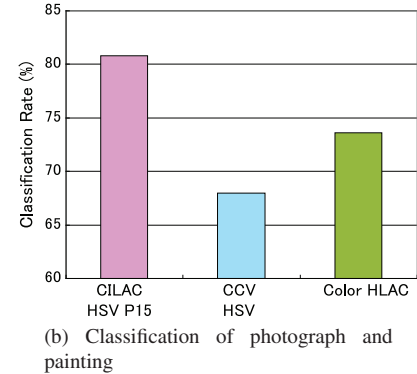
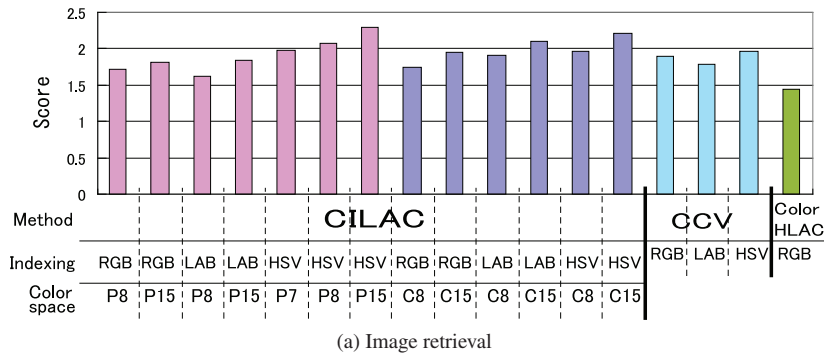


Fig. 6. Experimental results

addition, the metric of $L^*a^*b^*$ is derived from human perception and is suitable to the metric-based clustering method that we use. Among the various color spaces used in CILAC, HSV outperforms the others, showing that the (categorical) color representation of HSV is suitable for the CILAC method. The color indexing using predefined basic colors in HSV color space is simple, but CILAC with P15 in HSV is the most effective method producing the best result among all the methods. Although Color HLAC has the higher dimensionality (739 dimensions), its performance is the worst. These experimental results illustrate that co-occurrence of colors realized by auto-correlations of color indexes is effective feature.

Next, we try to classify photographs and paintings. This is a challenging task since both kinds of image have the same motif: natural outdoor scenes. We collected 60 paintings (including 30 Japanese paintings) and 60 photographs from world wide web (Fig. 5(b)¹). The classification is performed by linear SVM and 5-fold cross validation is employed for evaluation. As shown in Fig. 6(b), the result of CILAC is superior to the others. The characteristics of local textures extracted by CILAC would capture the subtle difference between these types of image.

4. CONCLUSIONS

We have proposed a method, color index local auto-correlation (CILAC), for extracting color image features using auto-correlations of color indexes. CILAC, which corresponds to co-occurrence of colors, can extract features of local color texture more discriminatively than the color histogram-based methods. CILAC is based on so general formulation that it has several variants in terms of both the selection of color space and the determination of basic colors in the color space for color indexing, as is the case with the color histogram. In the experiments of image retrieval and image classification, all these variants were comprehensively compared and

CILAC using fifteen predefined basic colors in HSV color space yielded a result superior to the other methods.

5. REFERENCES

- [1] Y-Y. Lin, T-L. Liu, and C-S. Fuh, "Local ensemble kernel learning for object category recognition," in *CVPR*, 2007, pp. 1–8.
- [2] G. Pass, R. Zabih, and J. Miller, "Comparing images using color coherence vectors," in *ACM international conference on Multimedia*, 1997, pp. 65–73.
- [3] A. E. Abdel-Hakim and A. A. Farag, "Csift: A sift descriptor with color invariant characteristics," in *CVPR*, 2006, pp. 1978–1983.
- [4] T. Kurita, S. Kato, I. Fukuda, and A. Itakura, "Sense retrieval on a image database of full color paintings (in japanese)," *Information Processing Society of Japan*, vol. 33, no. 11, pp. 1373–1383, 1992.
- [5] M. Swain and D. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.
- [6] S. Sural, G. Qian, and S. Pramanik, "Segmentation and histogram generation using the hsv color space for image retrieval," in *ICIP*, 2002, pp. II:589–592.
- [7] O. Hasegawa T. Toyoda, "Random field model for integration of local information and global information," *Pattern Analysis and Machine Intelligence*, vol. 30, no. 8, pp. 1483–1489, 2008.
- [8] N. Otsu and T. Kurita, "A new scheme for practical flexible and intelligent vision systems," in *IAPR Workshop on Computer Vision*, 1988.
- [9] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," in *CVPR*, 2006, pp. 2161–2168.
- [10] J. Weijer and C. Schmid, "Coloring local feature extraction," in *ECCV*, 2006, pp. 334–348.

¹<http://nihongo.wunderground.com/data/wximagenew/f/Feather3/4198.jpg>,
<http://www.alpha-net.ne.jp/users2/tenni811/bri.JPG>,
http://upload.wikimedia.org/wikipedia/commons/0/0a/The_Great_Wave_off_Kanagawa.jpg