Pattern Recognition 46 (2013) 188–198

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Content-based image retrieval using color difference histogram

Guang-Hai Liu [a,b,*], Jing-Yu Yang [b]

[a] College of Computer Science and Information Technology, Guangxi Normal University, Guilin 541004, China
[b] School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China

ARTICLE INFO

Article history:
Received 24 November 2011
Received in revised form
18 May 2012
Accepted 7 June 2012
Available online 15 June 2012

Keywords:
Image retrieval
$L^*a^*b^*$ color space
Edge orientation detection
Color difference histogram

ABSTRACT

This paper presents a novel image feature representation method, namely color difference histograms (CDH), for image retrieval. This method is entirely different from the existing histograms; most of the existing histogram techniques merely count the number or frequency of pixels. However, the unique characteristic of CDHs is that they count the perceptually uniform color difference between two points under different backgrounds with regard to colors and edge orientations in $L^*a^*b^*$ color space. This method pays more attention to color, edge orientation and perceptually uniform color differences, and encodes color, orientation and perceptually uniform color difference via feature representation in a similar manner to the human visual system. The method can be considered as a novel visual attribute descriptor combining edge orientation, color and perceptually uniform color difference, as well as taking the spatial layout into account without any image segmentation, learning processes or clustering implementation. Experimental results demonstrate that it is much more efficient than the existing image feature descriptors that were originally developed for content-based image retrieval, such as MPEG-7 edge histogram descriptors, color autocorrelograms and multi-texton histograms. It has a strong discriminative power using the color, texture and shape features while accounting for spatial layout.

1. Introduction

With the development of digital image processing technology, it has become imperative to find a method to efficiently search and browse images from large image collections. Generally, three categories of methods for image retrieval are used: text-based, content-based and semantic-based. In daily life, people search for images mainly via search engines such as Google, Yahoo, etc., which are based mainly on text keyword searches. Prompted by market demand for search services, image retrieval has become an extremely active research area in the field of pattern recognition and artificial intelligence. Current image retrieval techniques are usually based on low-level features (e.g., color, texture, shape, spatial layout), but low-level features often fail to describe high-level semantic concepts; that is, 'semantic gap' exists between high-level concepts and low-level features. To reduce this 'semantic gap', researchers have adopted machine-learning techniques to derive high-level semantics [1–3]. In addition, some researchers extract low-level features by simulating the mechanisms of the primary visual cortex [4–6]. Based on current advances in artificial intelligence and cognitive science, semantic-based image retrieval techniques remain limited. In this paper, we mainly focus on content-based image retrieval; based on low-level features, the image representation techniques used in this method are an effective way of integrating low-level features into a whole.

It is well known that visual perceptual differences between two colors in $L^*a^*b^*$ color space are related to a measure of Euclidean distance, but the representation of this attribute for image representation and its use for content-based image retrieval need to be further studied. To address this problem, this paper presents a new method of feature representation for content-based image retrieval, namely the color difference histogram (CDH). The CDH can be viewed as a general visual attribute descriptor and has the discrimination power of low-level features and spatial layout. The CDH was designed for color image analysis and its use can achieve higher retrieval precision than those of the MPEG-7 edge histogram descriptor (EHD) [9], color autocorrelograms (CAC) [8] and the Multi-texton histogram (MTH) [18], and can represent the perceptually uniform color difference between colors and edge orientations and take spatial information in $L^*a^*b^*$ color space into consideration.

The remainder of this paper is organized as follows. In Section 2, related work is introduced. The proposed algorithm is presented in Section 3; in Section 4, the performances of the MPEG-7 edge

* Corresponding author at: College of Computer Science and Information Technology, Guangxi Normal University, Guilin 541004, China.
Tel.: +86 0773 5811621.
E-mail addresses: liuguanghai009@163.com (G.-H. Liu), yangjy@mail.njust.edu.cn (J.-Y. Yang).

0031-3203/$ - see front matter Crown Copyright © 2012 Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2012.06.001

histogram descriptor, color autocorrelograms, the Multi-texton histogram and the proposed algorithm are compared based on two Corel datasets containing 15,000 images. Section 5 concludes the paper.

2. Related work

The visual system extracts information from the environment and transforms it into a neural code that results in perception [7]; color, texture and shape information are the most common types of visual information, and it is widely used in CBIR systems. Classical CBIR techniques are based on two types of visual features: global and local features. Global feature-based algorithms aim at the whole image as visual content, e.g., color, texture and shape whereas local feature-based algorithms focus mainly on key points or salient patches. Various algorithms have been designed for the extraction of global and local features.

Color is a wavelength-dependent perception [7] and has become a very important visual cue for image retrieval and object recognition. Color histograms are invariant to orientation and scale, and this feature makes it more powerful in image classification. Color histogram-based image retrieval is easy to implement and has been well studied and widely used in CBIR systems. However, color histograms characterize the spatial structures of images with difficulty. Therefore, several color descriptors have been proposed to exploit spatial information, including compact color moments, the color coherence vector, and color correlograms [8]. In the MPEG-7 standard, color descriptors consist of a number of histogram descriptors, such as the dominant color descriptor, the color layout descriptor, and a scalable color descriptor [9,10].

Texture is one of the most important characteristics of an image and still not an authority definition. Texture features are also widely used in CBIR systems. Various algorithms have been designed for texture analysis, such as gray level co-occurrence matrices [11], the Tamura texture feature [12], the Markov random field model [13], Gabor filtering [14], and local binary patterns [15]. The MPEG-7 standard adopts three texture descriptors: the homogeneous texture descriptor, the texture browsing descriptor and the edge histogram descriptor [9,10]. In practice, texture features can be combined with color features to improve discrimination power, thus yielding improved retrieval performance. One of the most commonly used methods is to join gray-level texture features and color features. Some algorithms can ultimately combine color and texture together; these include integrative co-occurrence matrices [16], the texton co-occurrences matrix [17], the multi-texton histogram [18], the color edge co-occurrence histogram [19], and the micro-structure descriptor [20]. The use of classical texture descriptors can be extended to color images by combining the results in color channels. In general, computing Gabor features separately for each channel can be used as a color texture descriptor.

In addition to color and texture features, shape features are also used in CBIR because humans can recognize objects solely based on their shapes. Classical methods of describing shape features include the use of moment invariants, Fourier transforms coefficients, edge curvature and arc length [21,22]. In MPEG-7, three shape descriptors are used for object-based image retrieval; these are the 3-D shape descriptor, region-based shapes derived from Zernike moments, and the curvature scale space (CSS) descriptor [10]. In many cases, shape feature extraction requires the use of image segmentation; this remains difficult, thereby limiting its application in practice.

Local image feature extraction and description have been attracting increasing attention in recent years. Several types of local descriptors have been reported in the literature [23–27], where the technique is termed 'scale-invariant feature transform' (SIFT) [23] is the most popular form of local feature representation and can tolerate certain levels of illumination changes, perspective distortions, and image transformations, and are very robust to occlusion. Recently, Bag-of-visual words-based methods and its variants, which are derived from local features such as keypoints and salient patches, have been proposed for object recognition and scene categorization [28–32]; in essence, these methods borrow techniques from text retrieval. It has been demonstrated that bag-of-visual words representation can result in improved object recognition and scene categorization performance. Unfortunately, because visual words are usually obtained by implementing clustering, which imposes heavy computational burdens, the bag-of-visual words technique has limitations, e.g., a lack of semantic information, the ambiguity of visual words, and a very high vector dimension. In practical applications, the discrimination power of visual words cannot compare with that of text words.

In this paper, we adopt the perceptually uniform color difference between two points under different color and edge-orientation backgrounds for image representation, but without any model training, clustering implementation or image segmentation. The proposed algorithm pays extra attention to color, edge orientation and perceptually uniform color difference, and combines the use of these features.

3. The color difference histogram (CDH)

Psychophysical and neurobiological studies indicate that the human visual system is very sensitive to color and edge orientation [44,45]. The perceptually uniform color difference between colors and edge orientations cover on rich variety of visual information, it is very useful information and plays an important role in image content analysis and understanding. However, to our knowledge, few articles have been published on how to apply the perceptually uniform color difference between colors and edge orientations to image representation and image retrieval. To this end, we propose a new descriptor for image retrieval in this paper. This descriptor combines the use of orientation, color and color difference and considers the spatial layout without the use of any image segmentation or learning processes.

In describing the proposed color difference histogram (CDH)-based image retrieval scheme, we first briefly describe the $L^*a^*b^*$ color space. Second, we describe the detection of edge orientation. Third, we describe the color quantization to be implemented in $L^*a^*b^*$ color space, and finally, we describe image features use perceptually uniform color difference.

3.1. $L^*a^*b^*$ Color space

Color is a very important visual attribute. In digital processing, the *RGB* color space is most commonly used; *R*, *G* and *B* components are highly correlated, and therefore, chromatic information is not directly fit for use. *CLE* $L^*a^*b^*$ was designed to be perceptually uniform [21,22]. Due to its high uniformity with respect to human color perception, the CLE $L^*a^*b^*$ color space is a particularly good choice for determining the difference between colors, and the difference between two color points can be measured as a Euclidean distance [21]. In the *CLE* $L^*a^*b^*$ color space, the L^*, a^* and b^* components are computed obtained through a non-linear mapping of *XYZ* coordinates. This conversion was performed using standard *RGB* to *CLE* $L^*a^*b^*$ transformations

as follows [21,22]:

$$\begin{cases} L^* = 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16 & \text{for } \frac{Y}{Y_n} > 0.008856 \\ L^* = 903.3\left(\frac{Y}{Y_n}\right)^{1/3} & \text{for } \frac{Y}{Y_n} \le 0.008856 \end{cases} \tag{1}$$

$$a^* = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \tag{2}$$

$$b^* = 200\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \tag{3}$$

With

$$\begin{cases} f(u) = u^{1/3} & \text{for } u > 0.008856 \\ f(u) = 7.787u + \frac{16}{116} & \text{for } u \le 0.008856 \end{cases} \tag{4}$$

where

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{5}$$

where X_n, Y_n and Z_n are the values of X, Y and Z for the illuminant (reference white point) and $[X_n, Y_n, Z_n] = [0.950450,\ 1.000000,\ 1.088754]$ in accordance with the illuminant D65. D65 is a commonly used standard illuminant defined by the International Commission on Illumination (CIE). For details, please refer to [21].

3.2. Edge orientation detection in $L^*a^*b^*$ color space

Edge orientation has a strong influence on human image perception and can represent object boundaries and texture structures [21,22,43], thereby providing most of the semantic information in the image. In this paper, we adopted a computationally efficient algorithm for edge orientation detection in $L^*a^*b^*$ color space.

Full color images include red, green and blue channels. If a full color image is converted to a gray-scale image and the gradient magnitude and orientation are detected based on the gray-scale image, much chromatic information will be lost. In [33], Zenzo has proposed a method for gradient computation using a full-color image. The core idea involves extending the concept of a gradient to the vector maximum rate of a scalar function $f(x, y)$ at coordinates (x, y) [22]. To efficiently detect the edges caused by chromatic changes in a perceptually uniform color space, we adopt the following method for edge orientation detection.

Let l, a and b be unit vectors along the L^*, a^* and b^* axes in $L^*a^*b^*$ color space; then, we adopt the following vectors for a full color image $f(x, y)$ [22,33]:

$$u = \frac{\partial L^*}{\partial x} l + \frac{\partial a^*}{\partial x} a + \frac{\partial b^*}{\partial x} b \tag{6}$$

$$v = \frac{\partial L^*}{\partial y} l + \frac{\partial a^*}{\partial y} a + \frac{\partial b^*}{\partial y} b \tag{7}$$

g_{xx}, g_{yy} and g_{xy} are defined as the dot products of these vectors [22,33]:

$$g_{xx} = u^T u = \left|\frac{\partial L^*}{\partial x}\right|^2 + \left|\frac{\partial a^*}{\partial x}\right|^2 + \left|\frac{\partial b^*}{\partial x}\right|^2 \tag{8}$$

$$g_{yy} = v^T v = \left|\frac{\partial L^*}{\partial y}\right|^2 + \left|\frac{\partial a^*}{\partial y}\right|^2 + \left|\frac{\partial b^*}{\partial y}\right|^2 \tag{9}$$

$$g_{xy} = u^T v = \frac{\partial L^*}{\partial x}\frac{\partial L^*}{\partial y} + \frac{\partial a^*}{\partial x}\frac{\partial a^*}{\partial y} + \frac{\partial b^*}{\partial x}\frac{\partial b^*}{\partial y} \tag{10}$$

The partial derivatives required for implementing the above vectors can be computed using Sobel operators. We use the Sobel operator because it is less sensitive to noise and has a small computational burden [22]. Let $I(x, y)$ be an arbitrary vector in the $L^*a^*b^*$ color space. Using the above notations, it can be seen that the direction of the maximum rate of the change of $I(x, y)$ is

$$\varphi(x,y) = \frac{1}{2}\arctan\left(\frac{2g_{xy}}{g_{xx} - g_{yy}}\right) \tag{11}$$

The value of the rate of change at (x, y) in the direction of $\varphi(x,y)$ given by

$$G(x,y) = \left\{\frac{1}{2}[(g_{xx} + g_{yy}) + (g_{xx} - g_{yy})\cos 2\varphi + 2g_{xy}\sin 2\varphi]\right\}^{1/2} \tag{12}$$

Because $\tan(a) = \tan(a \pm \pi)$, if φ_0 is a solution to Eq. (11), then $\varphi_0 \pm \pi/2$ will also be a solution. Furthermore, $G_\varphi = G_{\varphi+\pi}$; therefore, $G(x, y)$ has to be computed only for values of φ in the half-open interval $[0,\pi)$. Because Eq. (11) provides two values 90° apart, this equation associates a pair of orthogonal directions with each point (x, y) [22]. This process can be represented as follows:

$$G_1(x,y) = \left\{\frac{1}{2}[(g_{xx} + g_{yy}) + (g_{xx} - g_{yy})\cos 2\varphi_0 + 2g_{xy}\sin 2\varphi_0]\right\}^{1/2} \tag{13}$$

$$G_2(x,y) = \left\{\frac{1}{2}[(g_{xx} + g_{yy}) + (g_{xx} - g_{yy})\cos 2(\varphi_0 + \pi/2) + 2g_{xy}\sin 2(\varphi_0 + \pi/2)]\right\}^{1/2} \tag{14}$$

F_{max} is the maximum along one of these directions, and F_{min} is the minimum along the other. Thus, we can obtain the maximum gradient and minimum gradients shown in Fig. 1.

$$F_{max} = \max(G_1, G_2) \tag{15}$$

$$F_{min} = \min(G_1, G_2) \tag{16}$$

Because φ_0 and $\varphi_0 \pm \pi/2$ affect the gradient values, the maximum gradient is considered as the final gradient in practical applications; therefore the direction of $\varphi(x,y)$ is defined as

$$\varphi(x,y) = \begin{cases} \varphi_0 & \text{if } F_{max} = G_1(x,y) \\ \varphi_0 + \pi/2 & \text{if } F_{max} = G_2(x,y) \end{cases} \tag{17}$$

To facilitate implementation, we project it into the interval $[0, 2\pi]$. After the edge orientation $\varphi(x,y)$ of each pixel has been computed,

Fig. 1. Edge detection: (a) original full color image, (b) maximum gradient image and (c) minimum gradient image.

the orientations are uniformly quantized into m bins, where $m \in 6,12,18,24,30,36$. Denotes $\theta(x,y)$ as edge orientation map, where $\theta(x,y)=\phi, \phi \in 0,1,...,m-1$. For instance, if we set $m=6$, it is equal to that all edge orientations are uniformly quantized to the range of 0°, 60°, 120°, 180°, 240°, 300°.

In Section 4.5, the experimental results demonstrated that the 18 bins used in the $L^*a^*b^*$ color space are more suitable for our framework. Indeed, the orientations are quantized into 18 bins, each corresponding to angle intervals of 20°.

3.3. Color quantization in the $L^*a^*b^*$ color space

Color perception is a central component of primate vision, is fundamental to our perception of the world and can facilitate object perception and recognition. Humans can discern thousands of color shades and intensities, but only two dozen shades of gray [22]. To extract color information and simplify manipulation, color quantization needs to be implemented. The task of color quantization is to select and assign a limited set of colors for representing a given color image with maximum fidelity [21].

Color quantization has a close relationship with color spaces. Many color spaces have been proposed and used for image retrieval and object recognition. However, given the variety of color spaces available, it is difficult to choose the most appropriate color space for image retrieval. The choice of color space is also an important step in many image retrieval and object recognition algorithms. In terms of digital processing, however, *RGB* color space is the space most commonly used in practice and is straightforward [21]; however, color differences cannot be measured in *RGB* color space in a manner that is close to human color perception. In this work, we use the $L^*a^*b^*$ color space, which is quantized into 90 colors.

In Section 4.5, we describe experiments demonstrating that the $L^*a^*b^*$ color space is more suitable for our framework. Given a color image of size $M \times N$, we uniformly quantize the L^* channel into 10 bins and the a^* and b^* channels into 3 bins; therefore, $10 \times 3 \times 3 = 90$ color combinations are obtained. Denote by $C(x,y)$ the quantized image, where $0 < x < M$, $0 < y < N$.

3.4. Feature representation

Color, edge orientation and uniform color difference are closely related to human perception. It is an important challenge that feature representation use color difference and take into account both the spatial information of color and edge orientation cues because fewer algorithms use uniform color differences between two colors and edge orientations for feature representation. Indeed, differences inevitably exist between the results of edge orientation quantization and color quantization. Different edge orientations or colors may be assigned to the same values even if they are not similar to each other. In addition, two similar edge orientations or colors may be assigned to different values. The measurement of this difference is an important problem. It is well known that perceptually uniform color differences can be measured in a way that is close to human color perception [22]. Based on this idea, we propose a novel image feature representation method termed the color difference histogram (CDH) for image retrieval. The proposed algorithm can be expressed as follows:

The values of a quantized image $C(x,y)$ are denoted as $w \in 0,1,\ldots,W-1$. Denote neighboring pixels locations by (x, y) and (x', y') and their color index values as $C(x,y)=w_1$ and $C(x',y')=w_2$. The values of an edge orientation image $\theta(x,y)$ are denoted as $v \in 0,1,\ldots,V-1$. The angles at (x, y) and (x',y') are denoted by $\theta(x,y)=v_1$ and $\theta(x',y')=v_2$. For neighboring pixels, whose distance is D and whose respective quantization numbers for the color and edge orientations are W and V, we define the color difference histogram (CDH) as follows:

$$H_{color}(C(x,y)) = \begin{cases} \sum\sum \sqrt{(\Delta L)^2+(\Delta a)^2+(\Delta b)^2} \\ \text{where } \theta(x,y)=\theta(x',y'); \ \max(|x-x'|,|y-y'|)=D \end{cases} \tag{18}$$

$$H_{ori}(\theta(x,y)) = \begin{cases} \sum\sum \sqrt{(\Delta L)^2+(\Delta a)^2+(\Delta b)^2} \\ \text{where } C(x,y)=C(x',y'); \ \max(|x-x'|,|y-y'|)=D \end{cases} \tag{19}$$

Experimentally, we combine $H_{color}(C(x,y))$ and $H_{ori}(\theta(x,y))$ as the final feature vector H.

$$H=[H_{color}(0),H_{color}(1),\ldots,H_{color}(W-1),H_{ori}(0),H_{ori}(1),\ldots,H_{ori}(V-1)] \tag{20}$$

where ΔL, Δa and Δb are the respective color differences between two pixels in the L^*, a^* and b^* channels. In the color difference histogram, $H_{color}(C(x,y))$ can represent the perceptually uniform color difference between neighboring edge orientations using color index information as a constraint, leading to a 90-dimensional vector. $H_{ori}(\theta(x,y))$ can represent the perceptually uniform color difference between neighboring color indexes with edge orientation information as a constraint, leading to an 18-dimensional vector; in total, a $90+18=108$-dimensional vector is obtained for the final image features during image retrieval.

In this manner, orientation and perceptual color information are combined into a unified framework, and both spatial layouts are considered. Experimental results shown in Section 4.5 demonstrated that a 108-dimensional vector and the distance parameter $D=1$ is most suitable for our proposed framework.

In [34], a color difference histogram is used for video segmentation with different meanings based on two color difference channels, namely, the *R–G* and *G–B* channels. However, it should be stressed that the proposed color difference histogram (CDH) is entirely different from that described in [34] and the existing histogram. The proposed color difference histogram uses perceptually uniform color differences as the values in the histogram. However, the existing histogram techniques merely focus on the frequency or number of pixels and use these as the histogram values.

The proposed algorithm consists of two special histograms types, which are computed in a parallel manner under the background of colors and orientations. Fig. 2 shows two examples of the proposed CDH. In the proposed algorithms, only edge orientations and color index values that are the same are selected to calculate the color difference histogram, rather than all of them. There are two reasons for this: (1) this method is inspired by views that there are mechanisms of selective attention in the human visual cortex in the context of a biased competition account of attention [35,44]. (2) This method may result in a very high-dimensional vector if different edge orientations and colors are considered in the proposed algorithm.

4. Experiments and results

In this section, we demonstrate the performance of the proposed algorithm using two Corel datasets. In these experiments, we randomly selected 20 images from each category as query images. The performance was evaluated based on the average results of each query. For fair comparison, we selected algorithms that were originally developed for image retrieval, such as the edge histogram descriptor (EHD) [9], color autocorrelograms (CAC) [8], and the multi-texton histogram (MTH) [18]

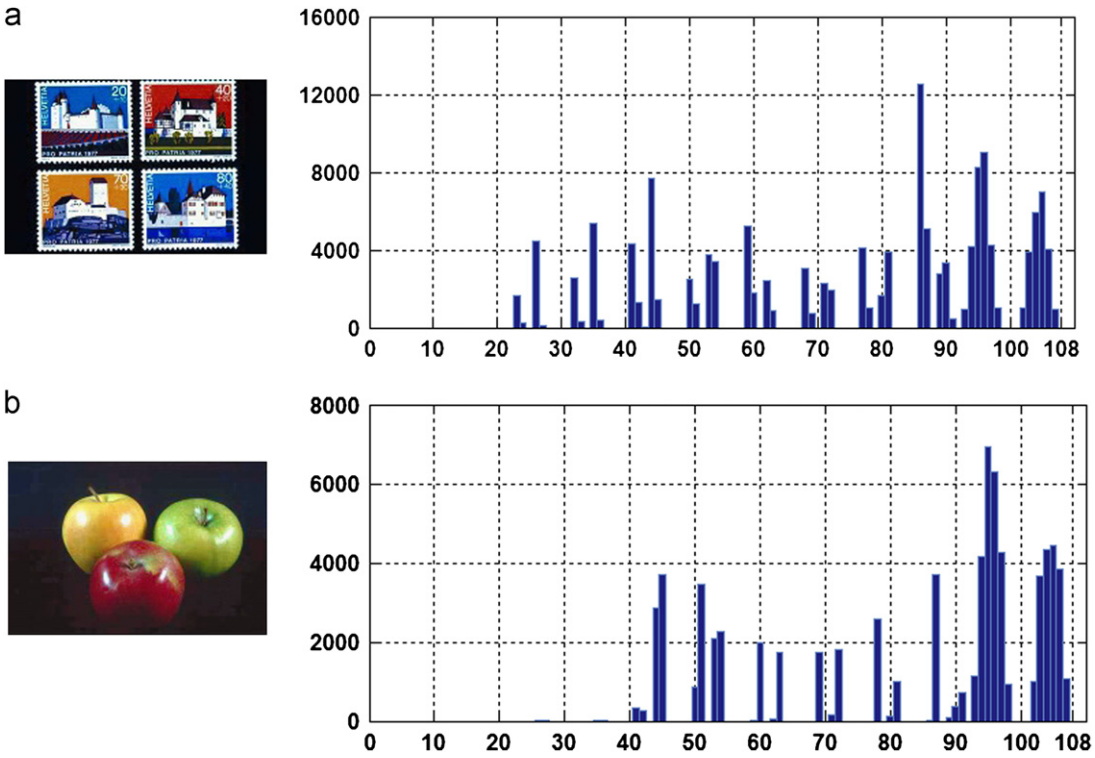

Fig. 2. Two examples of CDH: (a) stamps and (b) fruit. The horizontal axis corresponds to the index values for edge orientation and color (where values in the range 1–90 denote color index values and values in the range 91–108 denote edge orientation index values). The vertical axis corresponds to perceptually uniform color difference values.

because these adopted edge orientations or color information for image representation without the use of image segmentation and model training. An online image retrieval system using the proposed algorithm is available at: http://www.ci.gxnu.cn/cbir/.

4.1. Datasets

Various datasets are popularly used for various purposes in the field of image research; these include the Corel dataset, the Brodatz texture dataset, the OUTex texture dataset, the Coil-100 dataset, the ETH-80 dataset, the Caltech 101 dataset and the PASCAL VOC dataset. The Corel image dataset is the most commonly used dataset to test image retrieval performance, and the Brodatz texture dataset and the OUTex texture dataset are also widely used in texture-based image retrieval or texture analysis. However, other databases are mainly used for object recognition or scene categorization. Self-collected datasets can also be used for image retrieval.

The Corel image database contains a large amount of images containing various contents, ranging from animals and outdoor sports to natural scenes. Two Corel subsets are used in our image retrieval systems. All Corel images were obtained from Corel Gallery Magic 20, 0000 (8 CDs). The first subset is the Corel-5K dataset, which contains 50 categories covering 5000 images including diverse content such as fireworks, bark, microscopy images, tiles, food textures, trees, waves, pills and stained glass. Every category contains 100 images of size 192×128 or 128×192 in JPEG format. The second dataset is the Corel-10K dataset, which contains 100 categories covering 10,000 images including diverse content such as sunsets, beaches, flowers, buildings, cars, horses, mountains, fish, food, and doors. Each category contains 100 images of size 192×128 or 128×192 in JPEG format.

4.2. Distance metric

Retrieval accuracy not only depends on strong feature representation, but also on good similarity measures or distance metrics. The measurement of image content similarity remains problematic. In this paper, we extend the Canberra distance as a distance metric [36]. For each template image in the dataset, an M-dimensional feature vector $T=[T_1, T_2,\ldots,T_M]$ is extracted and stored in the database. Let $Q=[Q_1, Q_2,\ldots,Q_M]$ be the feature vector of a query image; then, the distance metric between them is simply calculated as follows:

$$D(T,Q)=\sum_{i=1}^{M}\frac{|T_i-Q_i|}{|T_i+u_T|+|Q_i+u_Q|} \tag{21}$$

where $u_T=\sum_{i=1}^{M}T_i/M$ and $u_Q=\sum_{i=1}^{M}Q_i/M$. For the proposed CDH, $M=108$ for color images. The class label of the template image that yields the smallest distance will be assigned to the query image.

In our experiments, the performances of the proposed distance metric will be compared to those of the commonly used distance or similarity metrics, such as the Canberra distance, x^2 statistics, the L_1 distance, the L_2 distance, the histogram intersection, the Cos Correlation and the Jeffrey divergence [36–38].

4.3. Performance metrics

In our experiments, we use the *Precision* and *Recall* curves, a performance metric commonly used in information retrieval [39,40]. *Precision* and *Recall* is defined as follows:

$$P(N)=I_N/N \tag{22}$$

$$R(N)=I_N/M \tag{23}$$

where I_N is the number of images retrieved in the top N positions that are similar to the query image, M is the total number of images in the database that are similar to the query, and N is the total number of images retrieved. In our image retrieval system, $N=12$ and $M=100$. Higher average precision and recall indicates better retrieval performance.

In addition to the *Precision* and *Recall* curves, the feature vector dimension is also very important for evaluating image retrieval performances. It is clear that every user wants their image search to be fast and accurate. Vector dimension affects the speed of image retrieval. Given the same or similar precision, the algorithms with the lowest vector dimension may be considered to have the highest performance, especially for large image datasets.

4.4. Implementation details

In the *RGB* color space, the color difference between two colors is defined as follows:

$$D_{RGB}(P_1,P_2)=\sqrt{(\Delta R)^2+(\Delta G)^2+(\Delta B)^2} \tag{24}$$

where ΔR denotes the color difference between two points P_1 and P_2 in channel R, and so on.

Because the HSV color space is based on the cylinder coordinate system, the values of its three channels, $H\in(0,360)$, $S\in[0,1]$ and $V\in[0,1]$, in this interval are unsuitable for edge orientation detection in the proposed framework; therefore, the HSV color space should first be transformed to the Cartesian coordinate system. Suppose that (H,S,V) is a random dot in the cylinder coordinate system and (H',S',V') is the transformation of (H,S,V) in the Cartesian coordinate system, where$H'=S\cos(H)$, $S'=S\sin(H)$, and $V'=V$. Then, the color difference between two colors is defined as follows:

$$D_{HSV}(P_1,P_2)=\sqrt{(\Delta H')^2+(\Delta S')^2+(\Delta V')^2} \tag{25}$$

where $\Delta H'$ denotes the color difference between two points P_1 and P_2 in channel H', and so on.

In this paper, the proposed algorithm is designed for color image retrieval. However, the MPEG-7 edge histogram descriptor (EHD) was originally designed for the retrieval of gray texture images. For fair comparison, we apply them to each of the R, G and B channels for feature extraction. Each channel leads to an 80-dimensional feature vector for the EHD. The final feature vector dimension of the EHD is $80\times3=240$. In the MPEG-7 edge histogram descriptor, image blocks whose edge strengths exceed 11 are used in computing the histogram, and L1 measure is used to compute the distance between two edge histograms.

4.5. Retrieval performance

In the experiments, we first demonstrate the reason why we adopted the $L^*a^*b^*$ color space for the proposed framework and confirm the final quantization number for color and edge orientation. Second, we demonstrate that the distance metric proposed in Eq. (21) is more suitable as the color difference histogram. Third, we will test the effect of the distance parameter on the color difference histogram. Finally, the retrieval performances are compared.

Different quantization numbers for color and edge orientation are used to test the performance of the CDH in the $L^*a^*b^*$, *RGB* and *HSV* color spaces. The values of precision and recall are listed in Tables 1–3. It can be seen from those data that the CDH algorithm performs the best in the $L^*a^*b^*$ color space, and the precision of the CDH ranges from 49% to 57%. When the quantization numbers for color and edge orientation are 90 and 18, respectively, the precision of the CDH is 57.23% for the Corel-5K dataset. When the quantization number for color is increased to $10\times4\times4=160$, the performance of the CDH is reduced, because as the color quantization number is increased, too many noisy features may be obtained, which will not enhance the description power.

In addition to the $L^*a^*b^*$ color space, the *RGB* and *HSV* color spaces also adopt uniform quantization. In the *RGB* color space, the total color quantization number is at least $4\times2\times2=16$ bins, and this number gradually increases to $8\times4\times4=128$ bins. The *HSV* color space could mimic human color perception well; therefore, many researchers use this space for color quantization. In the *HSV* color space, the total number of bins is at least $8\times3\times3=72$ bins, and this number gradually increases to $12\times4\times4=192$ bins.

Table 1
The average retrieval precision and recall of the CDH with different quantization numbers for color and edge orientation using the Corel-5K dataset in the $L^*a^*b^*$ color space.

The quantization number for color	The quantization number for edge orientation											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
180	54.12	55.36	56.60	56.96	57.13	57.13	6.49	6.64	6.79	6.84	6.85	6.85
160	49.41	51.73	51.80	51.64	51.16	50.45	5.93	6.21	6.22	6.20	6.14	6.05
90	54.81	56.58	57.23	56.87	57.10	56.70	6.58	6.79	6.87	6.83	6.85	6.80
45	52.46	53.50	53.07	52.85	52.26	52.01	6.30	6.42	6.37	6.34	6.27	6.24

Table 2
The average retrieval precision and recall of the CDH with different quantization numbers for color and edge orientation using the Corel-5K dataset in the *RGB* color space.

The quantization number for color	The quantization number for edge orientation											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
128	51.59	53.32	53.57	53.94	53.81	53.75	6.19	6.40	6.43	6.47	6.46	6.45
64	46.89	52.94	52.87	52.81	53.16	52.65	5.63	6.35	6.35	6.34	6.38	6.32
32	50.05	52.20	51.66	51.09	50.87	50.49	6.01	6.26	6.20	6.13	6.11	6.06
16	46.60	48.40	48.18	48.21	48.04	47.86	5.59	5.81	5.78	5.79	5.77	5.74

Table 3
The average retrieval precision and recall of the CDH with different quantization numbers for color and edge orientation using the Corel-5K dataset in the HSV color space.

The quantization number for color	The quantization number for edge orientation											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
192	51.21	53.38	54.03	54.48	55.11	55.34	6.15	6.41	6.48	6.54	6.61	6.64
128	52.80	54.35	55.19	55.80	56.01	56.18	6.34	6.52	6.62	6.70	6.72	6.74
108	52.20	53.03	53.77	54.34	54.52	54.81	6.26	6.36	6.45	6.52	6.54	6.58
72	52.58	53.09	53.91	54.23	54.42	54.82	6.31	6.37	6.47	6.51	6.53	6.58

Table 4
The average retrieval precision and recall of CDH with different distance or similarity metrics.

Dataset	Performance	Distance or similarity metrics							
		Our distance metric	Canberra	x^2 statistics	L_1	L_2	Histogram intersection	Cos Correlation	Jeffrey divergence
Corel-5K	Precision (%)	57.23	56.40	56.21	53.06	53.06	24.28	45.01	45.96
	Recall (%)	6.87	6.76	6.74	6.37	6.37	2.93	5.40	5.52
Corel-10K	Precision (%)	45.24	44.30	44.47	42.05	42.05	16.48	36.34	35.08
	Recall (%)	5.43	5.32	5.34	5.05	5.05	2.01	4.36	4.22

Given the same or similar retrieval precision and considering that the number of color quantization determines the feature vector dimension, we select the $L^*a^*b^*$ color space for color quantization in the proposed CDH scheme. However, it should be stressed that this does not mean that the $L^*a^*b^*$ color space will perform better than the other color spaces for other image retrieval methods. It only validates that the $L^*a^*b^*$ color space is better suited to the proposed algorithm. Indeed, the *HSV* color space is widely used in image retrieval and object recognition and achieves good performance [9,19,41]. In the proposed framework, the *HSV* color space can also provide much better results than the *RGB* color space.

In practice, color histograms based on a given color space may perform very well. However, the proposed algorithm is entirely different from the existing color histograms, in that it does not merely focus on color information; instead, color, edge orientation and color difference are all considered. Color information is merely one of the important factors used in deciding which color space is better suited to the proposed algorithm.

Indeed, precision in the *HSV* color space is similar to that in the $L^*a^*b^*$ color space. Given the similar precision, the vector dimension of the HSV color space is greater than that of the $L^*a^*b^*$ color space. As seen from Tables 1–3, increasing the equalization number of color and edge orientation does not always enhance the description power. Based on the results shown in Table 1 and to balance the retrieval precision and vector dimensions, we set the final quantization numbers for color and edge orientation in the proposed algorithm to 90 and 18, respectively. In other words, the final vector dimension is $90+18=108$ for the proposed algorithm during image retrieval.

We then validate the performance of the proposed distance metric and other popular distances or similarity metrics in the proposed algorithm. As seen from Table 4, the proposed distance metric performed better than distance or similarity metrics such as the Canberra distance, x^2 statistics, the L_1 distance, the L_2 distance, histogram intersection, the Cos Correlation and the Jeffrey divergence [36–38]. x^2 Statistics and the Canberra distance provided much better precision in retrieval experiments. We can also see that the L_1 and L_2 distances yield the same result as the proposed CDH method; however, the L_1 distance is much more computationally efficient, but this comes at the price of losing the rotation invariant property [19]. Euclidian distance is not always the best metric, because the distances in each dimension are squared before summation, placing great emphasis on features that are greatly dissimilar. It is clear that histogram intersection performs worst and is not suitable for the proposed algorithm. The major reason is that the color difference histogram is entirely different from the existing histograms and takes as its values the perceptually uniform color differences. However, the values of most of the existing histograms are the number or frequency of pixels.

The proposed distance metric can be considered as the improved Canberra distance. The Canberra distance is mostly used for data that are scattered around the origin [36], and the same values of $|T_i-Q_i|$ can come from different pairs of T_i and Q_i; using the weight parameter $(1/(T_i-Q_i)$ as the weight can reduce the opposite forces. When the weight parameter is taken into account and u_T and u_Q are used as smoothing factors, as shown in Eq. (21), the performance is increased.

To test the effect of the distance between two points on the perceptually uniform color difference, we experimented with various distance parameter values D ($D=1, 2,\ldots, 9$). The resulting average retrieval precision values are shown in Fig. 3. The average retrieval precision of the CDH is approximately 57% for the Corel-5K dataset and approximately 45% for the Corel-10K dataset. As seen in Fig. 3, the proposed algorithm achieves the best performance when the distance parameter $D=2$ for the Corel-5K dataset and $D=1$ for the Corel-10K dataset. Because the difference in the resulting precision between $D=1$ and 2 is small in the Corel-5K dataset, we adopt $D=1$ for image representation. When $D=1$, the CDH only counts the perceptually uniform color difference between two neighboring pixels under the color and edge orientation backgrounds. It is clear that the perceptually uniform color difference between two neighboring pixels has the best discrimination power.

The average precision and recall curves are plotted in Fig. 4. The average precision and recall of every image category is the varying parameter for each curve. There are 50 and 100 image categories in the Corel-5K and Corel-10K datasets, respectively. The horizontal axis corresponds to the recall, whereas the vertical

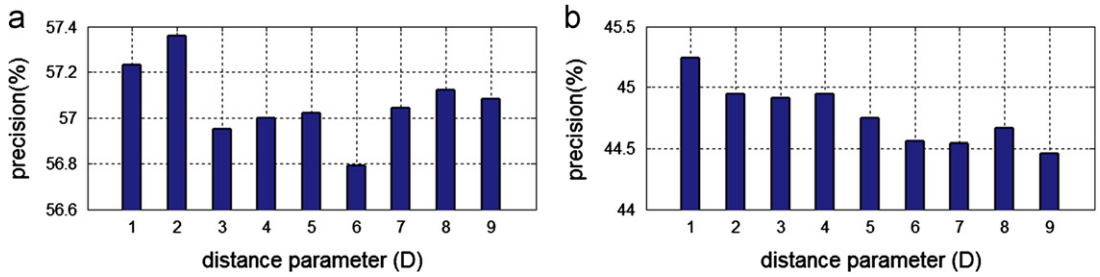

Fig. 3. The average retrieval precision of the CDH algorithm for various values of the distance parameter *D*: (a) the Corel-5K dataset and (b) the Corel-10K dataset.

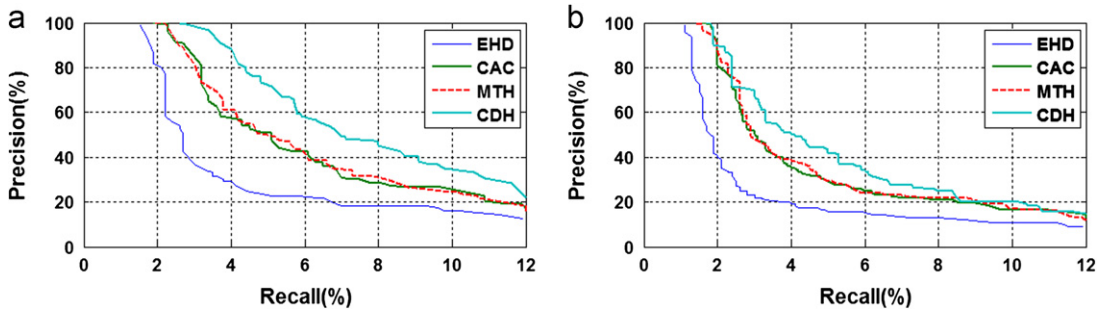

Fig. 4. The precision and recall curves of the EHD, CAC, MTH and CDH algorithms: (a) the Corel-5K dataset and (b) the Corel-10K dataset.

Table 5
The average retrieval precision and recall results using the two Corel datasets.

Datasets	Performance	Method			
		EHD	CAC	MTH	CDH
Corel-5K	Precision (%)	39.46	49.05	49.84	57.23
	Recall (%)	4.74	5.89	5.98	6.87
Corel-10K	Precision (%)	32.31	40.94	41.44	45.24
	Recall (%)	3.88	4.92	4.97	5.43

axis corresponds to precision. If the average retrieval precision and recall are higher, the curves will go gar from the original of coordinate. It can be seen from Table 5 and Fig. 4 that the proposed algorithm outperforms the EHD, CAC and MTH algorithms. It should be stressed that the proposed method performs this well using only 108 dimensional vectors, whereas the vector dimensions of the EHD and CAC algorithms are 240 and 256, respectively, higher than that of the proposed algorithm. Only the vector dimension of the MTH algorithm is lower than that of the CDH algorithm.

Figs. 5 and 6 show two retrieval examples using the Corel 10K dataset. In Fig. 5, the query image is the image of a stamp, and the top 11 retrieved images show a good match in texture and color to the query image. In Fig. 6, the query image is the image of fruit, which contains obvious shape features. All of the top 12 retrieved images show good matches between their shape and color to the query image. However, it should be emphasized that the two retrieval examples are only used to validate that the proposed algorithm has discrimination power regarding color, texture and edge features and do not suggest that all queries in the datasets can provide such high retrieval accuracy.

The edge histogram descriptor (EHD) captures the spatial distribution of local edges and is an efficient descriptor for images with a heavy textural presence even when the underlying texture is not homogeneous. The computation of this descriptor is straightforward and generates a histogram of the main edge directions within blocks of fixed size [42]. It can also work as a shape descriptor as long as the edge field contains the true object boundaries and is not saturated by the background texture [9]. In practice, the EHD can be used not only to represent the spatial distribution of local edges in natural images, but also to describe shape features; however, it is very sensitive to object or scene distortions. Thus, the retrieval performance may be unsatisfactory [42].

In [8], Huang et al. proposed the use of color correlograms for image indexing and retrieval using a way of the joint probability. Indeed, a subset of color correlograms, namely, color autocorrelograms, is used in image retrieval due to the high-dimensional vector of color correlograms. Huang's color autocorrelograms are based on color quantization in the RGB color space and lead to a 256-dimensional vector with 4 distance sets (1, 3, 5 and 7). The use of color autocorrelograms captures the spatial correlation of colors in an image, thus overcoming the major disadvantages of the classical histogram method, and this algorithm outperforms both the traditional histogram method and the color coherent vector method [8]. However, it should be stressed that color autocorrelograms achieve very good performance in uniform color space, such as the HSV, $L^*a^*b^*$ or Luv color space, but because color autocorrelograms come from 4 distance sets, they may result in very high dimensional vectors. Color provides powerful information and is only one image attribute, and edge orientation is also a very important factor in image representation that is ignored when using color autocorrelograms. Besides, different color images may obtain the same color autocorrelograms, and this can result in opposite forces.

MTH combines first-order and second-order statistics into an entity for texton analysis that is based on Julesz's texton conception [43] and encodes color and edge orientation as feature representations; thus, the texture discrimination power is greatly increased. MTH can represent the spatial correlation of edge orientation and color based on texton analysis [18]. Thus, its performance is better than that of EHD, which is well known for its textural feature representation in MPEG-7. The MTH algorithm is based on four special texton types; however, the four special

Fig. 5. An example of image retrieval using the CDH algorithm on the Corel 10K dataset. The query is an image of stamps, and 11 images are correctly retrieved and ranked within the top 12 images. (The top-left image is the query image, and the similar images returned include the query image itself.)

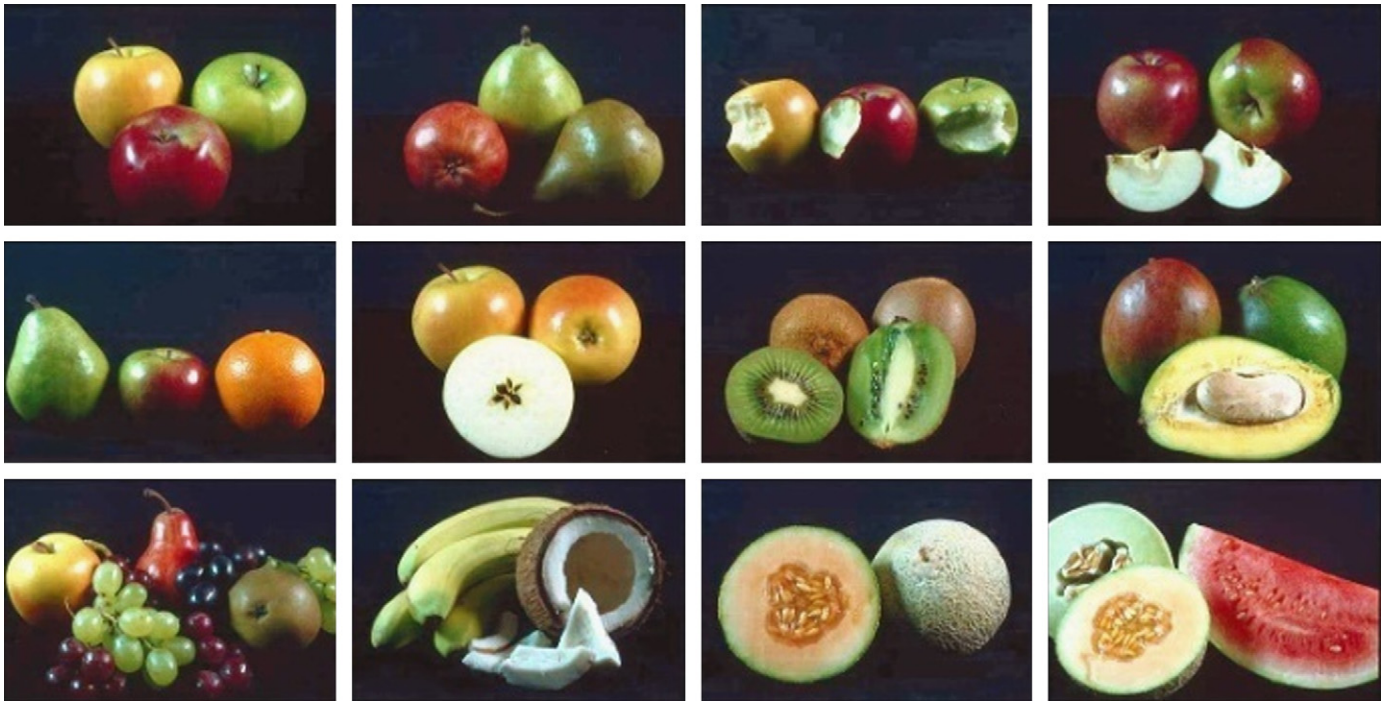

Fig. 6. An example of image retrieval using the CDH algorithm on the Corel 10K dataset. The query is an image of fruit, and all the returned images are correctly retrieved and ranked within the top 12 images. (The top-left image is the query image, and the similar images returned include the query image itself.)

texton types are only some of the many texton types in natural images. It cannot fully represent the content of texton images. The most important factor ignored by MTH is the perceptually uniform color difference, which can be measured in a similar way to human color perception.

The proposed algorithm analyzes the perceptually uniform color difference between neighboring colors and edge orientations based on two special histogram types in $L^*a^*b^*$ color space and overcomes the disadvantage of MTH, which discards perceptual color information. This algorithm can represent the perceptually uniform color

difference between colors and edge orientations, and contains the spatial information between them. Therefore, this algorithm can provide better performance than that of MTH and EHD. The proposed algorithm can be considered as an improved MTH because it considers the same neighboring colors and edge orientations as texton types and is not just limited to four special texton types; therefore, the image discrimination power is increased. The most significant improvement of the CDH over the multi-texton histogram is that it considers perceptually uniform color differences in feature representation, and this is similar to human color perception.

Based on the synthetic analysis of retrieval precision and vector dimension, the CDH algorithm performs better than the EHD, CAC and MTH algorithms.

5. Conclusion

In this paper, we have proposed a novel image feature representation method, namely, the color difference histogram (CDH), which is used to describe image features for image retrieval. This histogram is entirely different from existing histograms, and the proposed color difference histogram uses the perceptually uniform color difference as the histogram values. However, most of the existing histogram techniques merely focus on the frequency or number of pixels, which are used as the histogram values. In the proposed algorithm, orientation and perceptual color information have been combined in the unified framework, and both of their spatial layouts have been considered. This algorithm can be considered as a generalized low-level feature representation without the need for any image segmentation or model training. The vector dimension of the proposed algorithm is only 108 and the algorithm is therefore very efficient for image retrieval. The proposed algorithm can be considered as an improved multi-texton histogram. The most significant improvement of the proposed algorithm over multi-texton histograms is that it considers perceptually uniform color differences during feature representation, and this is similar to human color perception. Our experimental results have demonstrated that it is much more efficient than representative image feature descriptors, such as MPEG-7 edge histogram descriptors, color autocorrelograms or the multi-texton histogram. It has good discrimination power of color, texture, shape features and spatial layout.

Acknowledgments

This research was supported by the Guangxi Natural Science Foundation of China (no. 2011GXNSFB018070) and the National Natural Science Fund of China (no. 90820306, no. 61202272). The authors would like to thank the anonymous reviewers for their constructive comments.

References

[1] A. Vailaya, M.A.T. Figueiredo, A.K. Jain, H.J. Zhang, Image classification for content-based indexing, IEEE Transactions on Image Processing 10 (1) (2001) 117–130.
[2] V.N. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.
[3] S. Tong, E. Chang, Support vector machine active learning for image retrieval, in: Proceedings of the ACM International Conference on Multimedia, Ottawa, Canada, 2001, pp. 107–118.
[4] C. Grigorescu, N. Petkov, M.A. Westenbeg, Contour detection based on No classical receptive filed inhibition, IEEE Transactions on Image Processing 12 (7) (2003) 729–739.
[5] M. Ursino, G. Emiliano, L. Cara, A model of contextual interaction and contour detection in primary visual cortex, Neural Networks 17 (5–6) (2004) 719–735.
[6] G. Papari, N. Petkov, An improved model for surround suppression by steerable filters and multilevel inhibition with application to contour detection, Pattern Recognition 44 (9) (2011) 1999–2007.
[7] S. Schwartz, Visual Perception: A Clinical Orientation, fourth edition, McGraw-Hill Medical, 2009.
[8] J. Huang, S.R. Kumar, M. Mitra, et al., Image indexing using color correlograms, in: IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 762–768.
[9] B.S. Manjunath, J.-R. Ohm, V.V. Vasudevan, A. Yamada, Color and texture descriptors, IEEE Transactions on Circuit and Systems for Video Technology 11 (6) (2001) 703–715.
[10] B.S. Manjunath, P. Salembier, T. Sikora, Introduction to MPEG-7: Multimedia Content Description Interface, John Wiley & Sons Ltd, 2002.
[11] R.M. Haralick, K. Shangmugam, I. Dinstein, Textural feature for image classification, IEEE Transactions on Systems, Man and Cybernetics SMC-3 (6) (1973) 610–621.
[12] H. Tamura, S. Mori, T. Yamawaki, Texture features corresponding to visual perception, IEEE Transactions on Systems, Man, and Cybernetics 8 (6) (1978) 460–473.
[13] G. Cross, A. Jain, Markov random field texture models, IEEE Transactions on Pattern Analysis and Machine Intelligence 5 (1) (1983) 25–39.
[14] B.S. Manjunathi, W.Y. Ma, Texture features for browsing and retrieval of image data, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (8) (1996) 837–842.
[15] T. Ojala, M. Pietikanen, T. Maenpaa, Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (7) (2002) 971–987.
[16] C. Palm, Color texture classification by integrative co-occurrence matrices, Pattern Recognition 37 (5) (2004) 965–976.
[17] G.-H. Liu, J.-Y. Yang, Image retrieval based on the texton co-occurrence matrix, Pattern Recognition 41 (12) (2008) 3521–3527.
[18] G.-H. Liu, L. Zhang, et al., Image retrieval based on multi-texton histogram, Pattern Recognition 43 (7) (2010) 2380–2389.
[19] J. Luo, D. Crandall, Color object detection using spatial-color joint probability functions, IEEE Transactions on Image Processing 15 (6) (2006) 1443–1453.
[20] G.-H. Liu, Z.-Y. Li, L. Zhang, Y. Xu, Image retrieval based on micro-structure descriptor, Pattern Recognition 44 (9) (2011) 2123–2133.
[21] W. Burger, M.J. Burge, Principles of Digital image processing: Core Algorithms, Springer, 2009.
[22] R.C. Gonzalez, R.E. Woods, Digital Image Processing, 3rd edition, Prentice-Hall, 2007.
[23] D.G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60 (2) (2004) 91–110.
[24] Y. Ke, R. Sukthankar, PCA-SIFT: a more distinctive representation for local image descriptors, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, 2004, pp. 506–513.
[25] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (10) (2005) 1615–1630.
[26] H. Bay, T. Tuytelaars, L.V. Gool, SURF: speeded up robust features, in: European Conference on Computer Vision, vol. 1, 2006, pp. 404–417.
[27] S. Belongie, J. Malik, J. Puzicha. Shape matching, Object recognition using shape contexts, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (4) (2002) 509–522.
[28] J. Sivic, A. Zisserman, Video Google: a text retrieval approach to object matching in videos, in: IEEE International Conference on Computer Vision, vol. 2, 2003, pp. 1470–1477.
[29] S. Lazebnik, C. Schmid, J. Ponce, Beyond bags of features: spatial pyramid matching for recognizing natural scene categories, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, 2006, pp. 2169–2178.
[30] E. Nowak, F. Jurie, B. Triggs, Sampling strategies for bag-of features image classification, in: European Conference on Computer Vision, 2006, pp. 490–503.
[31] J. Philbin, O. Chum, M. Isard, J. Sivic, A. Zisserman, Lost in quantization: improving particular object retrieval in large scale image databases, in: IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8.
[32] J.C.V. Gemert, C.J. Veenman, A.W.M. Smeulders, J.M. Geusebroek, Visual word ambiguity, IEEE Transactions on Pattern Analysis and Machine Intelligence 32 (7) (2010) 1271–1283.
[33] S.D. Zenzo, A note on the gradient of a multi-image, Computer Vision, Graphics, and Image Processing 33 (1) (1986) 116–125.
[34] C.F. Lam, M.C. Lee, Video segmentation using color difference histogram, in: Proceeding of the International Workshop on Multimedia Information Analysis and Retrieval, London, UK, 1998.
[35] R. Desimone, Visual attention mediated by biased competition in extrastriate visual cortex, Philosophical Transactions of the Royal Society B 353 (1998) 1245–1255.
[36] G.N. Lance, W.T. Williams, Mixed-data classificatory programs I—agglomerative systems, Australian Computer Journal 1 (1) (1967) 15–20.
[37] Y. Rubner, J. Puzicha, C. Tomasi, J.M. Buhmann, Empirical evaluation of dissimilarity measures for color and texture, Computer Vision and Image Understanding 84 (1) (2001) 25–43.
[38] S. Antani, R. Kasturi, R. Jian, A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video, Pattern recognition 35 (4) (2002) 945–965.

[39] H. Müller, W. Müller, D.G. Squire, S.M. Maillet, T. Pun, Performance evaluation in content-based image retrieval: overview and proposals, Pattern Recognition Letters 22 (5) (2001) 593–601.
[40] Y. Yang, An evaluation of statistical approaches to text categorization, Information Retrieval 1 (1–2) (1999) 69–90.
[41] L.M Brown. Example-based color vehicle retrieval for surveillance, in: Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance, 2010, pp. 91–96.
[42] C.S. Won, D.K. Park, S.-J. Park, Efficient use of MPEG-7 edge histogram descriptor, ETRI Journal 24 (1) (2002) 23–30.
[43] B. Julesz, Textons, the elements of texture perception and their interactions, Nature 290 (5802) (1981) 91–97.
[44] S. Kastner, L.G. Ungerleider, The neural basis of biased competition in human visual cortex, Neuropsychologia 39 (12) (2001) 1263–1276.
[45] M.S. Livingstone, D.H. Hubel, Anatomy and physiology of a color system in the primate visual cortex, The Journal of Neuroscience 4 (1) (1984) 309–356.

Guang-Hai Liu is currently an associate professor with the College of Computer Science and Information Technology, Guangxi Normal University in China. He received Ph.D degree from the School of Computer Science and Technology, Nanjing University of Science and Technology (NUST). In 2011, He was engaged as an evaluation expert of science and technology project of Guangxi, china. His current research interests are in the areas of image processing, pattern recognition and artificial intelligence.

Jing-Yu Yang received the B.S. Degree in Computer Science from Nanjing University of Science and Technology (NUST), China. From 1982 to 1984 he was a visiting scientist at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign. From 1993 to 1994 he was a visiting professor at the Department of Computer Science, Missourian University in 1998; he worked as a visiting professor at Concordia University in Canada. He is currently a professor and Chairman in the department of Computer Science at NUST. He is the author of over 100 scientific papers in computer vision, pattern recognition and artificial intelligence. He has won more than 20 provincial awards and national awards. His current research interests are in the areas of image processing, robot vision, pattern recognition and artificial intelligence.