Computers and Electrical Engineering xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Computers and Electrical Engineering

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

# A new content-based image retrieval technique using color and texture information ☆

Xiang-Yang Wang *, Hong-Ying Yang [1], Dong-Ming Li

*School of Computer and Information Technology, Liaoning Normal University, Dalian 116029, China*

ARTICLE INFO

*Article history:*
Available online xxxx

ABSTRACT

Feature extraction and representation is one of the most important issues in the content-based image retrieval. In this paper, we propose a new content-based image retrieval technique using color and texture information, which achieves higher retrieval efficiency. Firstly, the image is transformed from RGB space to opponent chromaticity space, and the characteristics of the color contents of an image is captured by using Zernike chromaticity distribution moments from the chromaticity space. Secondly, the texture features are extracted using a rotation-invariant and scale-invariant image descriptor in Contourlet domain, which offers an efficient and flexible approximation of early processing in the human visual system. Finally, the combination of the color and texture information provides a robust feature set for color image retrieval. Experimental results show that the proposed color image retrieval is more accurate and efficient in retrieving the user-interested images.



## 1. Introduction

Nowadays, with increased digital images available on Internet, efficient indexing and searching becomes essential for large image archives. Traditional annotation heavily relies on manual labor to label images with keywords, which unfortunately can hardly describe the diversity and ambiguity of image contents. Hence, content-based image retrieval (CBIR) [1] has drawn substantial research attention in the last decade. CBIR usually indexes images by low-level visual features such as color and texture. The visual features cannot completely characterize semantic content, but they are easier to integrate into mathematical formulations [2]. Extraction of good visual features which compactly represent a query image is one of the important tasks in CBIR.

Color is widely regarded as one of the most expressive visual features, and as such it has been extensively studied in the context of CBIR, thus leading to a rich variety of descriptors. As conventional color features used in CBIR, there are color histogram, color correlogram, and dominant color descriptor (DCD) [1,3,4]. A simple color similarity between two images can be measured by comparing their color histograms. The color histogram, which is a common color descriptor, indicates the occurrence frequencies of colors in the image. The color correlogram describes the probability of finding color pairs at a fixed pixel distance and provides spatial information. Therefore color correlogram yields better retrieval accuracy in comparison to color histogram [3]. DCD is MPEG-7 color descriptors. DCD describes the salient color distributions in an image or a region of interest, and provides an effective, compact, and intuitive representation of colors presented in an image. However, DCD similarity matching does not fit human perception very well, and it will cause incorrect ranks for images with similar color

☆ Reviews processed and recommended for publication to Editor-in-Chief by Deputy Editor Dr. Ferat Sahin.

* Corresponding author. Tel.: +86 0411 85992415; fax: +86 0411 85992323.
*E-mail addresses:* wxy37@126.com (X.-Y. Wang), yhy_65@126.com (H.-Y. Yang).
[1] Tel.: +86 0411 85992415.

0045-7906/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compeleceng.2013.01.005

distribution [5]. In Ref. [6], Yang et al. presented a color quantization method for dominant color extraction, called the linear block algorithm (LBA), and it has been shown that LBA is efficient in color quantization and computation. For the purpose of effectively retrieving more similar images from the digital image databases (DBs), Lu et al. [7] uses the color distributions, the mean value and the standard deviation, to represent the global characteristics of the image, and the image bitmap is used to represent the local characteristics of the image for increasing the accuracy of the retrieval system. Aptoula et al. [8] presented three morphological color descriptors, one making use of granulometries independently computed for each subquantized color and two employing the principle of multiresolution histograms for describing color, using respectively morphological levelings and watersheds.

Textures are psycho-physically perceived by the human visual system (HVS), particularly, on the aspects of orientation and scale of texture patterns. Texture is also an important visual feature that refers to innate surface properties of an object and their relationship to the surrounding environment. Many objects in an image can be distinguished solely by their textures without any other information. In conventional texture features used for CBIR, there are statistic texture features using gray-level co-occurrence matrix (GLCM), Markov random field (MRF) model, simultaneous auto-regressive (SAR) model, Wold decomposition model, edge histogram descriptor (EHD), etc. Recently, BDIP (block difference of inverse probabilities) and BVLC (block variation of local correlation coefficients) features have been proposed which effectively measure local brightness variations and local texture smoothness, respectively [9]. These features are shown to yield better retrieval accuracy over the compared conventional features. Kokare et al. [10] designed a new set of 2D rotated wavelet by using Daubechies eight tap coefficients to improve the image retrieval accuracy. The 2D rotated wavelet filters that are non-separable and oriented, improves characterization of diagonally oriented textures. In Ref. [11], He et al. presented a novel method, which uses non-separable wavelet filter banks, to extract the features of texture images for texture image retrieval. Compared to traditional tensor product wavelets (such as db wavelets), the new method can capture more direction and edge information of texture images. Tzagkarakis et al. [12] described the design of a rotation-invariant texture retrieval system that exploits the non-Gaussian heavytailed behavior of the distributions of the subband coefficients, representing the texture information via a steerable pyramid. Han et al. [13] proposed a scale-invariant Gabor representations, where each representation only requires few summations on the conventional Gabor filter impulse responses, and the texture features are then extracted from these new representations for conducting scale-invariant texture image retrieval.

Most of the early studies on CBIR have used only a single feature among various visual features. However, it is hard to attain satisfactory retrieval results by using a single feature because, in general, an image contains various visual characteristics. Recently, active researches in image retrieval using a combination of color and texture features have been performed [1,2,14]. In Ref. [15], two-dimensional or one-dimensional histograms of the CIELab chromaticity coordinates are chosen as color features, and variances extracted by discrete wavelet frames analysis are chosen as texture features. In scheme [16], an efficient approach for querying and retrieval by multiple visual features is proposed. The approach employs three specialized histograms (i.e. distance, angle, and color histograms) to store feature-based information. Choraś et al. [17] developed original CBIR methodology that uses Gabor filtration, in which the texture features based on thresholded Gabor features, and color features based on histograms are calculated. Lin et al. [18] propose three image features for use in image retrieval. The first image feature is based on color distribution and is called an adaptive color histogram (ACH). The second and third image features, called adaptive motifs co-occurrence matrix (AMCOM) and gradient histogram for adaptive motifs (GHAM), are based on color and texture features, respectively. Chun et al. [19] proposed a CBIR method based on combination of multiresolution color and texture features. As its color features, color autocorrelograms of the hue and saturation component images in HSV color space are used. As its texture features, BDIP and BVLC moments of the value component image are adopted. In Ref. [20], an automatic content-based video shot indexing framework is proposed employing five types of MPEG-7 low-level visual features (color, texture, etc.). Kebapci et al. [21] presented a content-based image retrieval system for plant image retrieval, intended especially for the house plant identification problem. The suitability of various well-known color and texture features was studied, and some new texture matching techniques are introduced. Hiremath et al. [22,23] presented novel retrieval frameworks for combining multiple image information, in which the local color and texture descriptors are captured in a coarse segmentation framework of grids.

In this paper, we propose a new content-based image retrieval technique using Zernike chromaticity distribution moments and rotation-scale invariant Contourlet texture feature, which achieves higher retrieval efficiency. The rest of this paper is organized as follows. Section 2 presents Zernike chromaticity distribution color moments extraction. Section 3 describes the Contourlet transform and rotation-scale invariant texture representation. Section 4 contains the description of similarity measure for image retrieval. Simulation results in Section 5 will show the performance of our scheme. Finally, Section 6 concludes this presentation.

## 2. The Zernike chromaticity distribution moment

In general, color is one of the most dominant and distinguishable low-level visual features in describing image. Many CBIR systems employ color to retrieve images, such as QBIC system and VisualSEEK. In the proposed image retrieval method, we capture the characteristics of the color contents of an image by using Zernike chromaticity distribution moments directly from the chromaticity space. It is shown that the set of Zernike chromaticity distribution moments can provide a compact, fixed-length and computation effective representation of the color contents of an image, and only a small fixed number of compact Zernike chromaticity distribution moments need to be stored to effectively characterize the color content of an image.

### 2.1. The Zernike moments

Zernike moments consist of a set of complex polynomials [24] that form a complete orthogonal set over the interior of the unit circle, $x^2 + y^2 \leqslant 1$. If the set of these polynomials is denoted by $\{V_{nm}(x, y)\}$, then the form of these polynomials is as follows:

$$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho)\exp(jm\theta) \tag{1}$$

where $\rho = \sqrt{x^2+y^2}, \theta = \tan^{-1}(y/x)$. Here $n$ is a non-negative integer, $m$ is restricted to be $|m| \leqslant n$ and the radial Zernike polynomial $R_{nm}(\rho)$ is defined as the following

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} \frac{(-1)^s (n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} \tag{2}$$

Like any other orthogonal and complete basis, the Zernike polynomial can be used to decompose an analog 2D signal $f(x, y)$

$$f(x,y) = \sum_{n=0}^{\infty} \sum_{\{m:|m|\leq n\}} A_{nm} V_{nm}(x,y) \tag{3}$$

where $A_{nm}$ is the Zernike moments of order $n$ with repetition $m$, whose definition is

$$A_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2\leq 1} f(x,y) V^*_{nm}(x,y)dxdy \tag{4}$$

It should be pointed out that in case of 2D signal, (4) cannot be applied directly, but rather, its approximate version has to be employed. For instance, given a 2D signal of size $M \times N$, its Zernike moments are computed as

$$\widehat{A}_{nm} = \frac{n+1}{\pi} \sum_{i=1}^{M}\sum_{j=1}^{N} h_{nm}(x_i,y_j) f(x_i,y_j) \tag{5}$$

where the value of $i$ and $j$ are taken such that $x_2^i + y_2^j \leqslant 1$ and

$$h_{nm}(x_i,y_j) = \int_{x_i-\frac{\Delta x}{2}}^{x_i+\frac{\Delta x}{2}} \int_{y_i-\frac{\Delta y}{2}}^{y_i+\frac{\Delta y}{2}} V^*_{nm}(x,y)dxdy \tag{6}$$

where $\Delta x = \frac{2}{M}, \Delta y = \frac{2}{N} \cdot h_{nm}(x_i,y_j)$ can be computed to address the nontrivial issue of accuracy. In this research, we adopt the following formulas which are most commonly used in literature to compute Zernike moments of discrete 2D signals

$$A_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)V^*_{nm}(x,y) = \frac{n+1}{\pi}\sum_r\sum_\theta f(r\cos\theta, r\sin\theta)R_{nm}(r)\exp(-jm\theta) \tag{7}$$

The orthogonality and completeness of the Zernike yield the following formula for reconstructing the 2D signal

$$\hat{f}(x,y) = \sum_{n=0}^{n\max}\sum_{m=-n}^{n} A_{nm}V_{nm}(x,y) \tag{8}$$

### 2.2. The advantages of Zernike moments

The reason we use Zernike moments for image retrieval is that they have some very important properties, i.e., their magnitudes are invariant under rotation and flipping. We now elaborate on these invariance properties.

(1) The invariance properties. If 2D signal $f(\rho, \theta)$ is rotated $\alpha$ degrees counterclockwise, the 2D signal under rotation is $f'(\rho, \theta) = f(\rho, \theta - \alpha)$, and the Zernike moments of can be expressed by $A'_{nm} = A_{nm}\exp(-jm\alpha)$, which leads to $|A'_{nm}| = |A_{nm}|$. Therefore, the magnitudes of Zernike moments are robust to rotation. The magnitudes of Zernike moments also have perfect robustness to flipping.
(2) Good robustness against noise. Zernike moments have been proven to be superior to other moment functions such as Legendre moments in terms of their feature representation capabilities. Zernike moments can offer more feature vectors than Legendre moments, and they are less sensitive to noise than the conventional moments.

(3) Orthogonal property. Zernike moments are orthogonal. Orthogonal moments have been proven to be more robust in the presence of noise, and they are able to achieve a near zero value of redundancy measure in a set of moment functions.
(4) Reconstruction property. Zernike moments can be used to reconstruct the 2D signal.
(5) Multilayer expression. For a 2D signal, the low-order moments of the Zernike moments can express the outline of the 2D signal; and the high-order moments of the Zernike moments can express the detail of the 2D signal. Fig. 1 shows the binary image (2D signal) R and its reconstructed image. From left to right, the orders are 5, 10, 15, 20, and 25, respectively. It is not difficult to see that the low-order Zernike moments of digital image contain contour, but the high-order ones give the detail.

On the whole, the Zernike moments is an ideal region-based shape descriptor.

### 2.3. Opponent chromaticity space

A color space, also known as color signal, defines a set of attributes that uniquely identify a color. Color spaces are then important, as they set the distribution that colors present for different objects, which is fundamental in color classification and color retrieval. In addition, each color space provides with features that may be better suited for specific problems, such as varying lighting conditions or noisy environments. Color spaces can be classified into linear color space (including RGB, CMY, XYZ, YIQ, YUV) and non-linear color space (including Nrgb, Nxyz, L*a*b*, L*u*v*, HSV). Linear color spaces are based on a set of primary colors, and describe colors in terms of a weighed sum of the intensities of these primary colors. In contrast, non-linear color spaces can include more properties of color spaces that may help humans or different image processing techniques to better describe colors [25,26].

The most common representation for digital images is the RGB linear color space. The RGB color space used in computer monitors describes colors in terms of three primary colors: red, green and blue. A linear combination of these components will produce all the colors that can be shown in such screens. For RGB color space, the direct utilization of the $(R,G,B)$ triplet for image indexing is unreliable due to its susceptibility to change of brightness. Hence, the $(R,G,B)$ triplet are mapped to brightness independent chromaticities prior to indexing

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \tag{9}$$

Since $b = 1 - r - g$, the two-dimensional chromaticities (r, g) is sufficient to describe the color content of the image. To avoid the case of $R = G = B = 0$, which will cause an undefined division, we utilize the transformations $R \to R + \delta$, $G \to G + \delta$, $B \to B + \delta$, and modify the equation to

$$r = \frac{R+\delta}{R+G+B+3\delta}, \quad g = \frac{G+\delta}{R+G+B+3\delta}, \quad b = \frac{B+\delta}{R+G+B+3\delta} \tag{10}$$

where $\delta$ is an arbitrary small number. The brightness independence property is not affected since $R, G, B \geqslant \delta$. In our case, we choose $\delta = 0.01$.

In color research it is well known that a primal encoding of color in RGB co-ordinates is less uniform than an opponent encoding [25,26]. The opponent color space has three axes: white–black, yellow–blue and red–green axes. Each axis is mutually uncorrelated and so conveys independent information. The opponent chromaticity space

$$(rg, yb) = \left(r - g, \frac{r}{2} + \frac{g}{2} - b\right) \tag{11}$$

is more uniform and so can be more efficiently characterized by the Zernike moments (see following subsection). Note that the white–black axis is not used since it encodes brightness information and will not be included in the chromaticity histogram. It is straightforward to show that $-1 \leqslant rg \leqslant 1$, $-1 \leqslant yb \leqslant 1$.

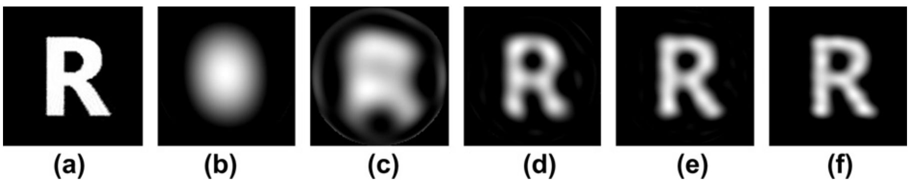



**Fig. 1.** The binary image (2D signal) R and its reconstructed image: (a) Origin image, (b) the reconstructed image (order is 5), (c) the reconstructed image (order is 10), (d) the reconstructed image (order is 15), (e) the reconstructed image (order is 20), and (f) the reconstructed image (order is 25).

### 2.4. The Zernike chromaticity distribution moment

Let $I(i, j) = [R(i, j), G(i, j), B(i, j)]$ represent an $M \times N$ color image, where $i = 0, 1, \ldots, M-1$, $j = 0, 1, \ldots, N-1$. We first map the RGB triplets to the opponent chromaticity space, and have

$$I(i,j) = [rg(i,j), yb(i,j)] \quad (i = 0,1,\ldots,M-1;\ j = 0,1,\ldots,N-1)$$

Then, we compute the Zernike chromaticity distribution moments of color image in opponent chromaticity space

$$\begin{aligned} A_{nm} &= \frac{n+1}{\pi}\sum_r\sum_\theta I(r\cos\theta, r\sin\theta)R_{nm}(r)\exp(-jm\theta) \\ &= \frac{n+1}{\pi}\sum_r\sum_\theta [rg(r\cos\theta, r\sin\theta), yb(r\cos\theta, \sin\theta)]R_{nm}(r)\exp(-jm\theta) \end{aligned} \tag{12}$$

Generally, the lower order Zernike chromaticity distribution moment represents slow-changing components in the opponent chromaticity space while the higher order one represents the fast changing components, and the zero order Zernike chromaticity distribution moment is a constant regardless of the image. So, some low order Zernike chromaticity distribution moments are selected as color feature set for image retrieval in this paper, and the color feature vector based on Zernike chromaticity distribution moments is given by

$$F_C = (A_{10}, A_{01}, A_{11}, A_{02}, A_{20}) \tag{13}$$

## 3. The rotation-scale invariant Contourlet texture feature

Most natural surfaces exhibit texture, which is an important low-level visual feature. Texture recognition will therefore be a natural part of many computer vision systems. In this paper, we propose a new rotation-invariant and scale-invariant texture representation for image retrieval based on Contourlet transform.

### 3.1. The Contourlet transform

The Contourlet transform is a true 2D transform defined in the discrete form to capture the contour information in all directions. It not only possesses the features of multiscale and time–frequency localization, but also offers anisotropy. It is able to capture significant information about an object of interest using a smaller description, when compared to wavelet transform. Contourlet transform decomposes image into bandpass directional subbands by pyramid directional filter banks (PDFBs). There are two stages in this processing, namely subband decomposition and directional decomposition. In this double filter bank, the Laplacian pyramid is first used to capture the point discontinuities, then followed by a directional filter bank to link point discontinuities into linear structures. The overall result is an image expansion using basic elements like contour segments. We would like to point out that the Contourlet transform is a multiscale and multidirection image representation, which allows for different and flexible number of directions at each scale [27].

#### 3.1.1. Pyramid frames

One way to obtain a multiscale decomposition is to use the Laplacian pyramid (LP) introduced by Burt and Adelson. The LP decomposition at each level generates a downsampled lowpass version of the original and the difference between the original and the prediction, resulting in a bandpass image.

#### 3.1.2. Iterated directional filter banks

The DFB is efficiently implemented via an $l$-level binary tree decomposition that leads to $2^l$ subbands with wedge-shaped frequency partitioning. In 1992, Bamberger and Smith constructed a 2-D directional filter bank (DFB) that can be maximally decimated while achieving perfect reconstruction [27].

Do and Vetterli proposed a new construction for the DFB that avoids modulating the input image and has a simpler rule for expanding the decomposition tree [27].

#### 3.1.3. Pyramid directional filter banks

The Laplacian pyramid only provides multiscale analysis without directionality. Since the directional filter bank (DFB) was designed to capture the high frequency (representing directionality) of the input image, the low frequency content is poorly handled. Combining the Laplacian pyramid and the directional filter bank, we can obtain an efficient image representation.

Fig. 2 illustrates the 3-level Contourlet decomposition for Zoneplate image. We notice that only Contourlets that match with both location and direction of image contours produce significant coefficients.

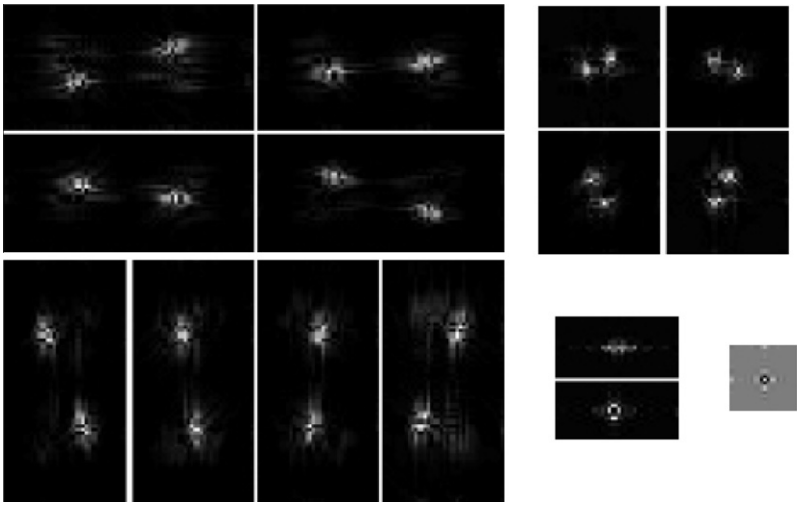


**Fig. 2.** The directional subbands and lowpass subband of three levels Contourlet decomposition for Zoneplate image.

### 3.2. The rotation-scale invariant Contourlet texture feature

In order to describe the texture characteristic of image, we use the rotation-scale invariant standard deviation and entropy of the Contourlet transformed image to capture the relevant image content into feature vectors.

We first extract the luminance component from the color image, and then perform a $S$ level Contourlet transform on the luminance component, so we can obtain a low-pass subband $A(x,y)$ and a series of high-pass subbands $D_{mn}(x,y)$ at scale $m$ = , 0, 1, ..., $S-1$ and orientation $n = 0, 1, \ldots, K-1$. The energy distribution ($E(m,n)$) and mean ($\mu_{mn}$) of the high-pass subbands $D_{mn}(x,y)$ at scale $m$, and at orientation $n$ are defined as

$$E(m,n) = \sum_x \sum_y |D_{mn}(x,y)|$$
$$\mu_{mn} = \frac{1}{MN} E_{mn}(x,y) \tag{14}$$

Additionally, the standard deviation ($\sigma_{mn}$) and entropy ($H_{mn}$) of the energy distributions are found as follows

$$\sigma_{mn} = \sqrt{\frac{1}{MN} \sum_x \sum_y (|D_{mn}(x,y)| - \mu_{mn})^2}$$
$$H_{mn} = -\sum_x \sum_y D_{mn}^2(x,y) \log(D_{mn}^2(x,y)) \tag{15}$$

where $M$ and $N$ denote the height and width of the input image respectively, $D_{mn}(x,y)$ is the Contourlet coefficient within the given high-pass subband.

So, the corresponding Contourlet texture feature vector ($F_T$) is defined by using the standard deviation and entropy as texture feature elements. It is denoted as

$$F_T = (\sigma_{00}, H_{00}, \sigma_{01}, H_{01}, \cdots, \sigma_{S-1K-1}, H_{S-1K-1})$$

Obviously, the dimensionality of the texture feature vectors depends on the number of scales ($S$) and on the number of orientations ($K$) considered during the Contourlet transform. Next, we discuss the rotation-scale invariant Contourlet texture feature.

Rotation-invariant texture representation is achieved by computing the dominant orientation of the images followed by feature alignment. The dominant orientation (DO) is defined as the orientation with the highest total energy across the different scales considered during image decomposition [28]. It is computed by finding the highest accumulated energy for the $K$ different orientations considered during image decomposition

$$DO_i = \max\{E_0^{(R)}, E_1^{(R)}, \ldots, E_{K-1}^{(R)}\}$$

where $i$ is the index where the dominant orientation appeared, and

$$E_n^{(R)} = \sum_{m=0}^{S-1} E(m,n), \quad n = 0, 1, \ldots, K-1$$

Here, each $E_n^{(R)}$ covers a set of high-pass subbands $D_{mn}(x, y)$ at different scale $m$ but at same orientation $n$.

Finally, rotation-invariance is obtained by shifting circularly feature elements within the same scales, so that first elements at each scale correspond to dominant orientations. As an example, let $F_T$ be a feature vector obtained by using a Contourlet transform with $m = 2$ scales, and $n = 2$ orientations

$$F_T = (\sigma_{00}, H_{00}, \sigma_{01}, H_{01}; \sigma_{10}, H_{10}, \sigma_{11}, H_{11}, \sigma_{12}, H_{12})$$

Suppose that the dominant orientation appears at index $i = 1$ ($DO_{i=1}$), thus the rotation-invariant texture feature vector, after feature alignment, is represented as follows

$$F'_T = (\sigma_{01}, H_{01}, \sigma_{00}, H_{00}; \sigma_{11}, H_{11}, \sigma_{12}, H_{12}, \sigma_{10}, H_{10})$$

Similarly, scale-invariant texture representation is achieved by finding the scale with the highest total energy across the different orientations (dominant scale). For this purpose, the dominant scale (DS) at index $i$ is computed as follows

$$DS_i = \max\{E_0^{(S)}, E_1^{(S)}, \ldots, E_{S-1}^{(S)}\}$$

where $E_m^{(S)}$ denotes the accumulated energies across the $S$ different scales

$$E_m^{(S)} = \sum_{n=0}^{K-1} E(m,n), \quad m = 0, 1, \ldots, S-1$$

Here, each $E_m^{(S)}$ covers a set of high-pass subbands $D_{mn}(x, y)$ at different orientations for each scale. As an example, let $F_T$ be a feature vector obtained by using a Contourlet transform with $m = 2$ scales, and $n = 2$ orientations

$$F_T = (\sigma_{00}, H_{00}, \sigma_{01}, H_{01}; \sigma_{10}, H_{10}, \sigma_{11}, H_{11}, \sigma_{12}, H_{12})$$

By supposing that the dominant scale was found at index $i = 1$ (second scale in the Contourlet decomposition), its scale-invariant texture feature vector, after feature alignment, is represented as

$$F''_T = (\sigma_{10}, H_{10}, \sigma_{11}, H_{11}; \sigma_{12}, H_{12}, \sigma_{00}, H_{00}, \sigma_{01}, H_{01})$$

According to the dominant orientation and dominant scale simultaneously, we can obtain the rotation-scale invariant Contourlet texture feature by feature alignment (In this paper, we set $m = 2$ and $n = 2$)

$$F_T = (\sigma_{00}, H_{00}, \sigma_{01}, H_{01}; \sigma_{10}, H_{10}, \sigma_{11}, H_{11}, \sigma_{12}, H_{12}) \tag{16}$$

## 4. Similarity measure

After the color and texture feature vectors are extracted, the retrieval system combines these feature vectors, calculates the similarity between the combined feature vector of the query image and that of each target image in an image DB, and retrieves a given number of the most similar target images.

### *4.1. Color feature similarity measure*

The color feature similarity is given by

$$S_{Color}(Q,I) = ((A_{10}^Q - A_{10}^I)^2 + (A_{01}^Q - A_{01}^I)^2 + (A_{11}^Q - A_{11}^I)^2 + (A_{02}^Q - A_{02}^I)^2 + (A_{20}^Q - A_{20}^I)^2)^{1/2} \tag{17}$$

where $A_{ij}^Q$ denotes the color feature of query image $Q$, and $A_{ij}^I$ denotes the color feature of target image $I$.

### *4.2. Texture feature similarity measure*

We give the texture feature similarity as follows

$$\begin{aligned} S_{Texture}(Q,I) = ((&\sigma_{00}^Q - \sigma_{00}^I)^2 + (H_{00}^Q - H_{00}^I)^2 + (\sigma_{01}^Q - \sigma_{01}^I)^2 + (H_{01}^Q - H_{01}^I)^2 + (\sigma_{10}^Q - \sigma_{10}^I)^2 + (H_{10}^Q - H_{10}^I)^2 + (\sigma_{11}^Q \\ &- \sigma_{11}^I)^2 + (H_{11}^Q - H_{11}^I)^2 + (\sigma_{12}^Q - \sigma_{12}^I)^2 + (H_{12}^Q - H_{12}^I)^2)^{1/2} \end{aligned} \tag{18}$$

where $\sigma_{ij}^Q$ and $H_{ij}^Q$ denote the rotation-scale invariant Contourlet texture feature of query image $Q$, $\sigma_{ij}^I$ and $H_{ij}^I$ denote the rotation-scale invariant Contourlet texture feature of target image $I$.

So the distance used for computing the similarity between the query feature vector and the target feature vector is given as

$$S(I,Q) = w_C S_{Color}(Q,I) + w_T S_{Texture}(Q,I) \qquad (19)$$
$$w_C + w_T = 1$$

where $w_C$ and $w_T$ are the weights of the color and texture features respectively.

When retrieving images, we firstly calculate the similarity between the query image and each target image in the image DB, and then sort the retrieval results according to the similarity value.

## 5. Experimental results

In this paper, we propose a new and effective color image retrieval scheme for combining color and texture information, which achieve higher retrieval efficiency. To evaluate the performance of the proposed algorithm, we conduct an extensive set of CBIR experiments by comparing the proposed algorithm to several state-of-the-art image retrieval approaches [16,23].

### 5.1. Evaluation setup and dataset

The color image retrieval systems have been implemented in MATLAB 7.0 environment on a Pentium 4 (2 GHz) PC. To check the retrieval efficiency of proposed method, we perform experiments over 6000 images from 150 categories of the COREL photo gallery, in which each category contains 100 images. Every database image is of size $256 \times 384$ or $384 \times 256$. Corel images have been widely used by the image processing and CBIR research communities. They cover a variety of topics, such as “Flowers”, “Buses”, “Beach”, “Elephants”, “Sunset”, “Buildings”, and “Horses”. Fig. 3 shows some image examples in this dataset.

In order to evaluate the retrieval performance, at the same time considering that the number of similar images returned from retrieval system is not suitable for using general Precision and Recall standard, we adopt Precision and Normal Recall proposed by Tan et al. as evaluation standard [29]. Normal Precision and Normal Recall are defined as below

$$P_{Normal} = 1 - \frac{\sum_{i=1}^{L}(\log r_i - \log i)}{\log \frac{N!}{(N-L)!L!}},$$
$$R_{Normal} = 1 - \frac{\sum_{i=1}^{L}(r_i - i)}{(N-L)L}$$

where $N$ is the number of image retrieved ($N$ is set as 20 in the simulation experiment), $L$ is the number of relevant images retrieved, and $r_i$ is the serial number of each relevant images retrieved.

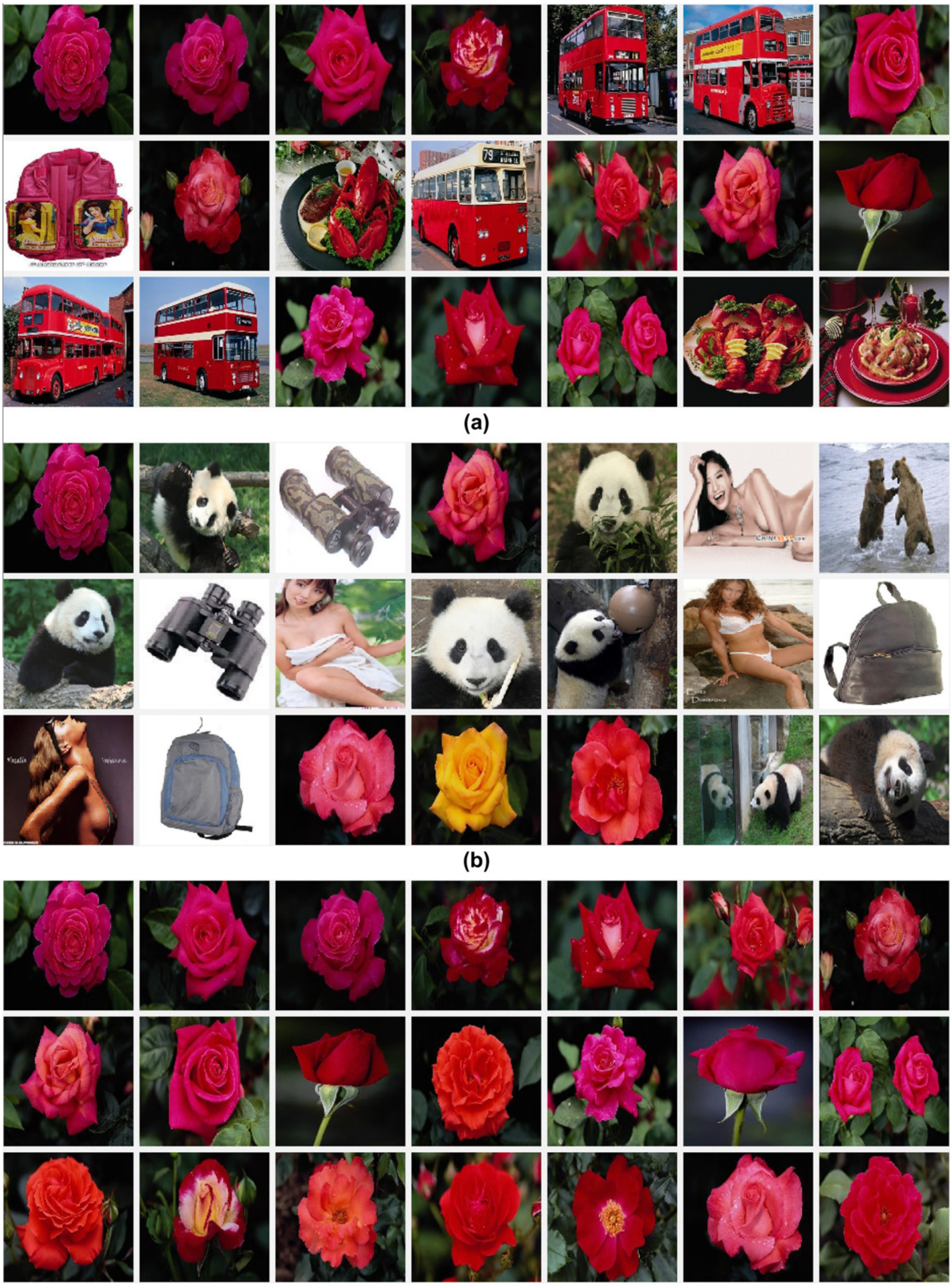
### 5.2. The performance of feature selection

Generally in a CBIR system, images are represented by multiple visual features. Color information is the most informative feature because of its robustness with respect to scaling, rotation, perspective, and occlusion. Texture information can be another important feature and previous studies have shown that texture structure and orientation fit well the model of human perception.

In our image retrieval, the Zernike chromaticity distribution moments are use to capture the color content, and the rotation-scale invariant image descriptor in Contourlet domain are extracted to capture the texture information. To evaluate the overall performance of the proposed image feature in retrieval, a number of experiments were performed on our image retrieval. The parameter settings are: $w_C = w_T = 1/2$.

In Figs. 4 and 5, we demonstrate our retrieval results with the Zernike chromaticity distribution moments only, the rotation-scale invariant Contourlet texture feature only, both Zernike chromaticity distribution moments and rotation-scale

**Fig. 3.** Some images from the COREL dataset used in our experiments.



(a)

(b)

(c)

**Fig. 4.** Our image retrieval results (Flower): (a) By taking the Zernike chromaticity distribution moments only, (b) by taking the rotation-scale invariant Contourlet texture feature only, and (c) by taking both Zernike chromaticity distribution moments and rotation-scale invariant Contourlet texture feature.


**Fig. 5.** Our image retrieval results (Elephant): (a) By taking the Zernike chromaticity distribution moments only, (b) by taking the rotation-scale invariant Contourlet texture feature only, and (c) by taking both Zernike chromaticity distribution moments and rotation-scale invariant Contourlet texture feature.

Please cite this article in press as: Wang X-Y et al. A new content-based image retrieval technique using color and texture information. Comput Electr Eng (2013), http://dx.doi.org/10.1016/j.compeleceng.2013.01.005


**Fig. 6.** The image retrieval results (Horse) using different schemes: (a) The retrieval scheme [16], (b) the retrieval scheme [23], and (c) the proposed retrieval method.

Please cite this article in press as: Wang X-Y et al. A new content-based image retrieval technique using color and texture information. Comput Electr Eng (2013), http://dx.doi.org/10.1016/j.compeleceng.2013.01.005


**Fig. 7.** The image retrieval results (Bus) using different schemes: (a) The retrieval scheme [16], (b) the retrieval scheme [23], and (c) the proposed retrieval method.

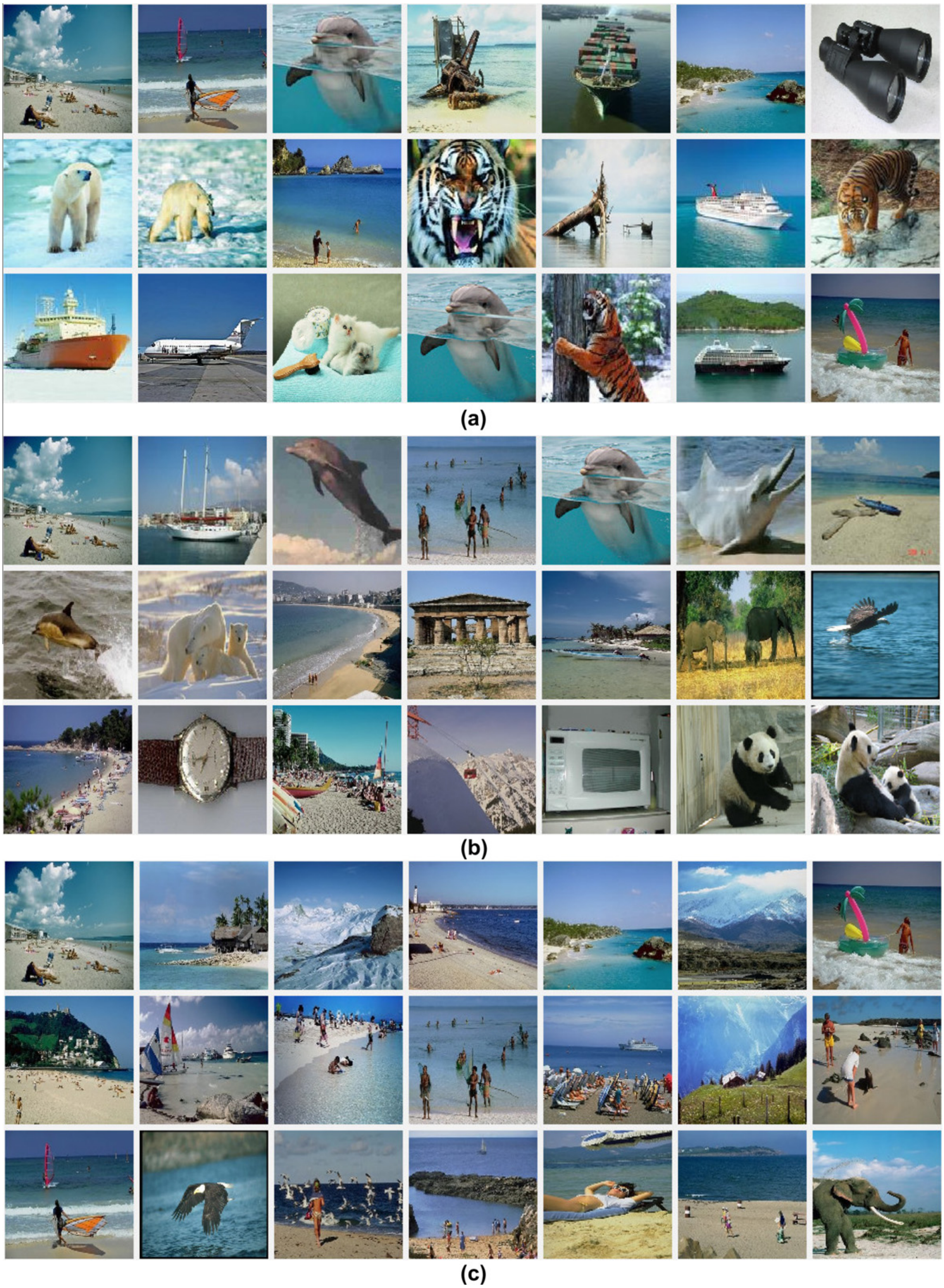


**Fig. 8.** The image retrieval results (Beach) using different schemes: (a) The retrieval scheme [16], (b) the retrieval scheme [23], and (c) the proposed retrieval method.

invariant Contourlet texture feature, respectively. The image at the top of left-hand corner is the query image; other 20 images are the retrieval results.

### 5.3. Comparative performance evaluation

We report experimental results that show the feasibility and utility of the proposed algorithm and compare its performance with two state-of-the-art image retrieval approaches [16,23]. To simulate the practical situation of online users, the sequence of query images used in all the experiments is generated at random.

Figs. 6–8 show the image retrieval results using the scheme [16], scheme [23], and the proposed method. The image at the top of left-hand corner is the query image; other 20 images are the retrieval results. Table 1 shows the number of correct retrieval images for different methods.

**Table 1**
The number of correct retrieval images for different methods.

| The query image | Scheme [16] | Scheme [23] | The proposed method |
|---|---|---|---|
| Horse | 8 | 9 | 18 |
| Bus | 9 | 9 | 20 |
| Beach | 6 | 8 | 16 |

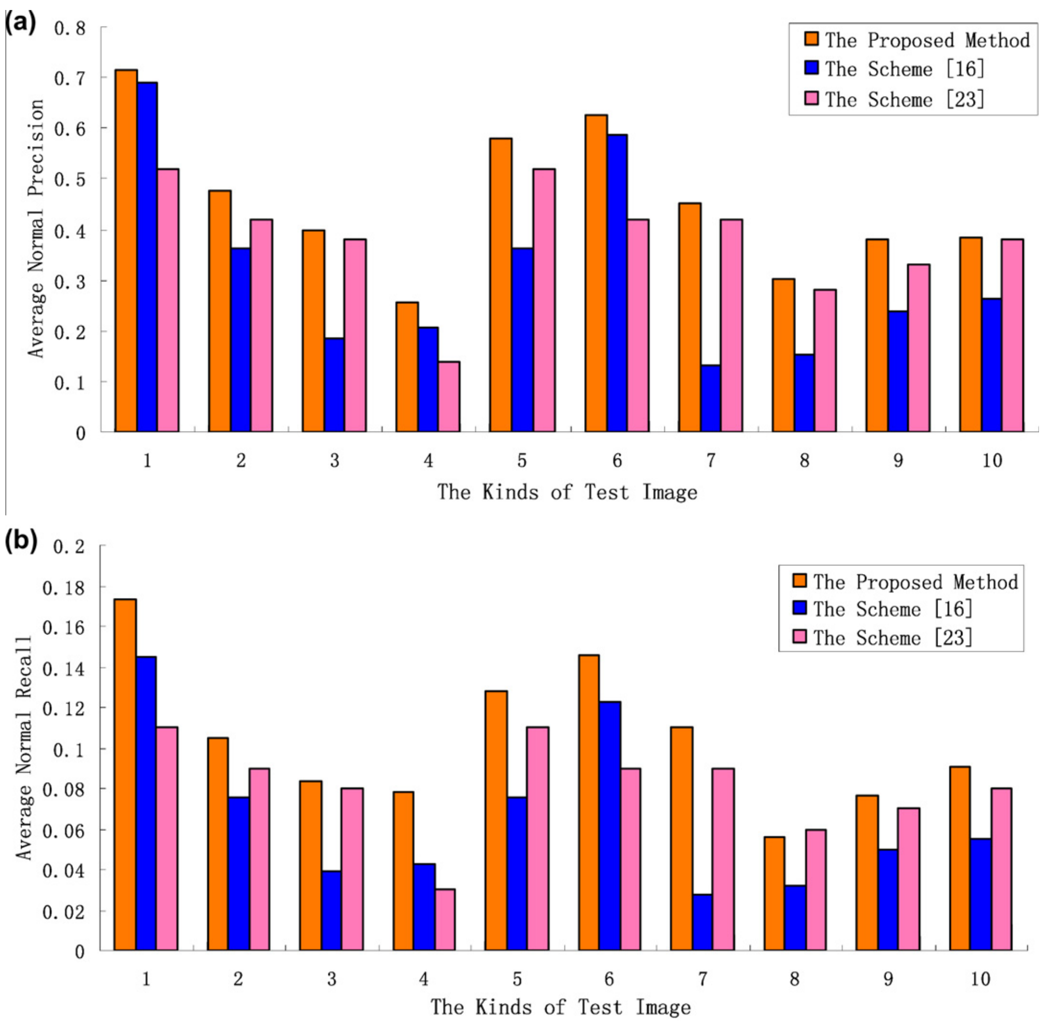



**Fig. 9.** The average retrieval performance of three schemes: (a) The average normal precision and (b) the average normal recall.

In order to further confirm the validity of the proposed algorithm, we randomly selected 50 images as query images from the above image database (The tested 10 semantic class includes bus, horse, flower, dinosaur, building, elephant, people, beach, scenery, and dish). Each kind is extracted five images, and each time returns the first 20 most similar images as retrieval results. To each kind of image, the average normal precision, and the average normal recall of five times query results are calculated. These values are taken as the retrieval performance standard of the algorithm, as shown in Fig. 9.

According to the Figs. 4–9 we can see that our color image retrieval is more accurate and efficient in retrieving the user-interested images, and is superior to the state-of-the-art image retrievals recently proposed in the literature.

## 6. Conclusion

CBIR is an active research topic in image processing, pattern recognition, and computer vision. In this paper, a CBIR method has been proposed which uses the combination of Zernike chromaticity distribution moments and rotation-scale invariant Contourlet texture descriptor. Experimental results showed that the proposed method yielded higher retrieval accuracy than the other conventional methods with no greater feature vector dimension. In addition, the proposed method almost always showed performance gain in of average normal precision, average normal recall, and average retrieval time over the other methods. As further studies, the proposed retrieval method is to be evaluated for more various DBs and to be applied to video retrieval.

## Acknowledgement

This work was supported by the National Natural Science Foundation of China under Grant Nos. 61272416, 60873222, 60773031, the Open Foundation of State Key Laboratory of Information Security of China under Grant No. 04-06-1, and Liaoning Research Project for Institutions of Higher Education of China under Grant Nos. 2008351 & L2010230.

## References

[1] Datta Ritendra, Joshi Dhiraj, Li Jia, Wang James Z. Image retrieval: ideas, influences, and trends of the new age. ACM Computing Surveys 2008;40(2):1–60.
[2] Vogel J, Schiele B. Performance evaluation and optimization for content-based image retrieval. Pattern Recognition 2006;39(5):897–909.
[3] W. Rasheed, Y. An, S.B. Pan, I. Jeong, J. Park, J. Kang, Image retrieval using maximum frequency of local histogram based color correlogram, in: Second Asia International Conference on Modeling & Simulation(AICMS 08), Kuala Lumpur, May 2008, pp. 322–326.
[4] A.B. Penatti, R. da Silva Torres, Color descriptors for web image retrieval: a comparative study, in: The XXI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI '08), Campo Grande, Brazil, 12–15 October 2008, pp. 163–170.
[5] Min Rui, Cheng HD. Effective image retrieval using dominant color descriptor and fuzzy support vector machine. Pattern Recognition 2009;42(1):147–57.
[6] Yang Nai-Chung, Chang Wei-Han, Kuo Chung-Ming, Li Tsia-Hsing. A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval. Journal of Visual Communication and Image Representation 2008;19(2):92–105.
[7] Zu-Chuen Lu, Chang Chin-Chen. Color image retrieval technique based on color features and image bitmap. Information Processing and Management 2007;43(2):461–72.
[8] Aptoula E, Lefèvre S. Morphological description of color images for content-based image retrieval. IEEE Transactions on Image Processing 2009;18(11):2505–17.
[9] Chun YD, Seo SY, Kim NC. Image retrieval using BDIP and BVLC moments. IEEE Transactions on Circuits and Systems for Video Technology 2003;13(9):951–7.
[10] Kokare Manesh, Biswas PK, Chatterji BN. Texture image retrieval using rotated wavelet filters. Pattern Recognition Letters 2007;28(10):1240–9.
[11] He Zhenyu, You Xinge, Yuan Yuan. Texture image retrieval based on non-tensor product wavelet filter banks. Signal Processing 2009;89(8):1501–10.
[12] Tzagkarakis G, Beferull-Lozano B, Tsakalides P. Rotation-invariant texture retrieval via signature alignment based on Steerable sub-Gaussian modeling. IEEE Transactions on Image Processing 2008;17(7):1212–25.
[13] Han Ju, Ma Kai-Kuang. Rotation-invariant and scale-invariant Gabor features for texture image retrieval. Image and Vision Computing 2007;25(9):1474–81.
[14] Deselaers T, Keysers D, Ney H. Features for image retrieval: an experimental comparison. Information Retrieval 2008;11(2):77–107.
[15] Liapis S, Tziritas G. Color and texture image retrieval using chromaticity histograms and wavelet frames. IEEE Transactions on Multimedia 2004;6(5):676–86.
[16] Saykol E, Gudukbay U, Ulusoy O. A histogram-based approach for object-based query-by-shape-and-color in image and video databases. Image and Vision Computing 2005;23(13):1170–80.
[17] Choraś RS, Andrysiak T, Choraś M. Integrated color, texture and shape information for content-based image retrieval. Pattern Analysis & Applications 2007;10(4):333–43.
[18] Lin Chuen-Horng, Lin Wei-Chih. Image retrieval system based on adaptive color histogram and texture features. The Computer Journal 2011;54(7):1136–47.
[19] Chun Young Deok, Kim Nam Chul, Jang Ick Hoon. Content-based image retrieval using multiresolution color and texture features. IEEE Transactions on Multimedia 2008;10(6):1073–84.
[20] B. Rachid, H. Benoit, Perplexity-based evidential neural network classifier fusion using MPEG-7 low-level visual features, in: ACM International Conference on Multimedia Information Retrieval 2008 (MIR 2008), October 27–November 01, 2008, Vancouver, BC, Canada, pp. 336–341.
[21] Kebapci H, Yanikoglu B, Unal G. Plant image retrieval using color, shape and texture features. The Computer Journal 2011;54(9):1475–90.
[22] P.S. Hiremath, J. Pujari, Content based image retrieval using color, texture and shape features, in: Proceedings of the 15th International Conference on Advanced Computing and Communications, Guwahati, India, December 2007, pp. 780–784.
[23] Wang Xiang-Yang, Yong-Jian Yu, Yang Hong-Ying. An effective image retrieval scheme using color, texture and shape features. Computer Standards & Interfaces 2011;33(1):59–68.
[24] Wood J. Invariant pattern recognition: a review. Pattern Recognition 1996;29(1):1–17.
[25] Yap PT, Paramesran R. Content-based image retrieval using Legendre chromaticity distribution moments. IEE Proceedings Vision, Image and Signal Processing 2006;153(1):17–24.
[26] Raziel A, Erik M, Alejandro AL, Ricardo SO. Accurate Color Classification and Segmentation for Mobile Robot. Tech-Open Access Publisher; 2007.

Please cite this article in press as: Wang X-Y et al. A new content-based image retrieval technique using color and texture information. Comput Electr Eng (2013), http://dx.doi.org/10.1016/j.compeleceng.2013.01.005

[27] Do MN, Vetterli M. The Contourlet transform: an efficient directional multiresolution image representation. IEEE Transactions on Image Processing 2005;14(12):2091–106.
[28] Javier AMZ, Joao PP, Neucimar JL. Learning how to extract rotation-invariant and scale-invariant features from texture images. EURASIP Journal on Advances in Signal Processing 2008;S1687–6172:1–15.
[29] Tan KL, Ooi BC, Yee CY. An evaluation of color-spatial retrieval technique for large image database. Multimedia Tools and Applications 2002;14(1):55–78.

**Xiang-Yang Wang** is currently a professor with the Multimedia and Information Security Laboratory, School of Computer and Information Technology, Liaoning Normal University, China. His research interests lie in the areas of information security, image processing, pattern recognition, and computer vision.

**Hong-Ying Yang** is currently a professor with the School of Computer and Information Technology at the Liaoning Normal University, China. Her research interests include signal processing and communications, digital multimedia data hiding.

**Dong-Ming Li** received the B.E. degree from the School of Computer and Information Technology, Liaoning Normal University, China, in 2009, where he is currently pursuing the M.S.E. degree. His research interests include image retrieval and signal processing.

Please cite this article in press as: Wang X-Y et al. A new content-based image retrieval technique using color and texture information. Comput Electr Eng (2013), http://dx.doi.org/10.1016/j.compeleceng.2013.01.005