J. Vis. Commun. Image R. 24 (2013) 63–74

Contents lists available at SciVerse ScienceDirect

J. Vis. Commun. Image R.

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

A novel method for image retrieval based on structure elements' descriptor

Wang Xingyuan *, Wang Zongyu

Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China

ARTICLE INFO

Article history:
Received 26 June 2012
Accepted 24 October 2012
Available online 31 October 2012

Keywords:
Texture descriptor
Color descriptor
Quantization color space
Content based image retrieval
Structure elements' descriptor
Structure elements' histogram
Texton detection
HSV color space

ABSTRACT

In this paper, structure elements' descriptor (SED) – a novel texture descriptor, is proposed. SED can effectively describe images and represent image local features. Moreover, SED can extract and describe color and texture features. The image structure elements' histogram (SEH) is computed by SED, and HSV color space is used (it has been quantized to 72 bins). SEH integrates the advantages of both statistical and structural texture description methods, and it can represent the spatial correlation of color and texture. The results demonstrate that the method has a better performance than other image retrieval methods in the experiments.

1. Introduction

In recent years, image database is increasing fast. It is difficult to capture useful images in the huge image database, and it becomes worse that we frequently obtain useful images through the Internet which is widely spreading. Existing methods are usually difficult to change the awkward situation, so it is urgent to find a new way to retrieve images accurately. Thus image retrieval becomes an important topic in image processing and pattern recognition domain. Generally speaking, images can be retrieved in three ways: text based, content based and semantic based [1–6]. Text based retrieval approach is used widely, such as Baidu and Google, and we can retrieve images using keywords that are annotated on images. However, with this method we often obtain images that are unrelated to our expected results, and the results of the image retrieval rely on our understanding of the query images. There are two drawbacks of this approach. Firstly, images in the database are annotated manually by annotators, it is time-consuming for a huge image database and requires much human labor. Secondly, the results of retrieval are inaccurate, because they are related to the understanding of the query images. The second approach, e.g., content based image retrieval (CBIR) has been proposed in the early 1990's [7–14]. This approach is to retrieve images using low-level features like color, texture and shape that can represent an image. With this method, we firstly extract the low-level features on the example image, then compute the similarity between the query image and the images in image dataset, finally sort images by similarity and the top images will be displayed. CBIR has been shown more effectively and subjectively than text based approach [15–17]. The final approach is semantic based method. CBIR also fails to describe the semantic concepts, so researchers proposed some methods for image retrieval by using relevance feedback algorithms. The relevance feedback algorithms capture user's preferences and bridge the semantic gap [18,19], and the results are closer to human perception.

In CBIR, color, texture and shape are the most important features. HSV color space is widely used in extracting color features. In this space, hue is used to distinguish color, saturation is the percentage of white light added to a pure color, and value refers to the perceived light intensity [2,20]. The advantage of HSV color space is that it is closer to human conceptual understanding of colors. In order to cut down the computing complexity and extract the color features in efficient way, the HSV color space are quantized to 72 bins in general [20].

Generally, in CBIR, the descriptors represent the image based on color, shape or texture and are used to describe images. Various algorithms have been designed to extract features for image retrieval. The multi-texton histogram (MTH) is proposed for image retrieval, MTH integrates the advantages of co-occurrence matrix and histogram by representing the attribute of co-occurrence matrix using histogram [1]. A novel image feature detecting and describing method called micro-structure descriptor (MSD) is

* Corresponding author.
E-mail addresses: wangxy@dlut.edu.cn (X. Wang), chenxizongyu@163.com (Z. Wang).

1047-3203/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jvcir.2012.10.003

proposed. MSD is built based on the underlying colors in microstructures with similar edge orientation [2]. In [7], three types of image features are proposed to describe the color and spatial distributions of an image. In these features, the K-means algorithm is adopted to classify all of the pixels in an image into several clusters according to their colors. The MPEG-7 edge histogram descriptor (EHD) is extracted on spatial distribution of edges, and it is an efficient texture descriptor for images [21]. The edge orientation autocorrelogram (EOAC) is proposed for shape based descriptor [22]. A very effective method has been proposed to detect and describe local features in images, called scale-invariant feature transform (SIFT) [23]. The texton co-occurrence matrices (TCM) can describe the spatial correlation of textons for image retrieval which is proposed in [24]. In [25], an adaptive color feature extraction scheme is proposed by considering the distribution of an image, the binary quaternion-moment-preserving (BQMP) threshold technique is used.

In this paper, a novel texture descriptor SED is proposed, the image is described by SED. The structure elements are defined by five structure elements denoting five directions, respectively. SED can effectively represent image local features, and can extract and describe color and texture features. SEH is computed by SED, and HSV color space that has been quantized to 72 bins is used. SEH integrates the advantages of both statistical and structural texture description methods, and it can represent the spatial correlation of color and texture. Experiments showed that the SED method has higher retrieval precision and recall ratio than the MTH, MSD and CSD3 method.

The rest of the paper is organized as follows. Section 2 details the method of quantizing the HSV color space to 72 bins. Section 3 describes the descriptor of SED and the method of extracting SEH of the image. The similarity measure approach is defined in Section 4. The experimental results and comparisons are presented in Section 5. Section 6 is the conclusion of the paper.

2. Color quantization in HSV color space

HSV color space is widely used in color feature extracting. In this space, hue is used to distinguish color, saturation is the percentage of white light added to a pure color and value refers to the perceived light intensity. The advantage of HSV color space is that it is closer to human conceptual understanding of colors. It is well known that color provides powerful information for image retrieval method, but the human eye cannot perceive a large number of colors at the same time, but it is able to distinguish similar colors well [2,20].

As is known to all that texture is an important low-feature and becomes a useful feature in CBIR. In order to extract color, texture features and more useful information, we quantize the HSV color space to several bins to represent all colors and extract SED on every bin. So the texture is extracted with combined color feature. In this way, the feature vector of the image can extract color and texture features simultaneously, therefore it can represent the image more accurately and has a better performance in CBIR.

In order to cut down the computing complexity and distinguish similar colors well, the HSV color space is quantized to 72 bins in this paper, as the method used in [20].

We extract the point in the image, obtain its RGB values, and convert the RGB color space to the HSV color space which is closer to human conceptual understanding of color.

According to the characteristics of HSV color space, $H \in [0, 360]$, $S \in [0, 1]$ and $V \in [0, 1]$, we quantize HSV color space to 72 bins as follows:

(1) Divide hue into 8 zones, saturation into 3 zones and values into 3 zones by follow equations:

$$H = \begin{cases} 0, & H \in [0,24] \cup [345,360] \\ 1, & H \in [25,49] \\ 2, & H \in [50,79] \\ 3, & H \in [80,159] \\ 4, & H \in [160,194] \\ 5, & H \in [195,264] \\ 6, & H \in [265,284] \\ 7, & H \in [285,344]. \end{cases} \tag{1}$$

$$S = \begin{cases} 0, & S \in [0,0.15] \\ 1, & S \in (0.15,0.8] \\ 2, & S \in (0.8,1]. \end{cases} \tag{2}$$

$$V = \begin{cases} 0, & V \in [0,0.15] \\ 1, & V \in (0.15,0.8] \\ 2, & V \in (0.8,1]. \end{cases} \tag{3}$$

(2) Construct the color feature of the image into one-dimension based on the following classification:

$$P = Q_s Q_v H + Q_v S + V. \tag{4}$$

In Eq. (4), Q_s and Q_v are the number of the quantization of color space component S and V, respectively. In this paper, S and V are quantized into three levels, therefore $Q_s = 3,\ Q_v = 3$. The following equation can be inferred from Eq. (4).

$$P = 9H + 3S + V. \tag{5}$$

(3) Compute the point set L_i.

Generally, let M be an $m \times n$ image and convert RGB color space to HSV color space, we suppose I is the result that has been quantized by HSV color space before. $i\,(0 \leqslant i \leqslant 71)$ is used to denote the value that has been quantized to 72 bins, respectively. And L_i denotes the point set where the value of I is i, which can be computed by the following equation:

$$L_i = \{(x,y)|(x,y) \in I, I(x,y) = i, 0 \leqslant i \leqslant 71\}, \tag{6}$$

where $I(x, y)$ is the quantized value which is located on (x, y).

In this way, convert three components of the HSV to a one-dimensional vector, which has 72 main colors. SED is used to represent the image on 72 kinds of main colors and compute SEH.

3. Feature extraction

3.1. The structure elements' descriptor (SED)

The contents in digital images are very significant so that comparing them directly is feasible for applications in image retrieval. Color, texture and shape play an important role in content based image retrieval. The image is represented by low level features such as color, texture and shape descriptor. However, the local structures of images from the same class often show a certain amount of similarities. In some sense, we may think that the meaningful content of images is composed of many structure elements, if we extract these structure elements and describe them effectively, they can serve as common bases for the comparison and analysis for different images [2]. In this way, the image can be represented effectively by these structure elements. On the other hand, the texture can be decomposed into elementary units, and analysis texture orientation plays an important role in computer vision and image processing [1,24]. But with these methods, only the texture is obtained.

In [24], original images are quantized into 256 colors in RGB color space and used to extract textons. TCM describes an image by its gradient information and color information, and TCM does not combine the texture feature with the color feature while

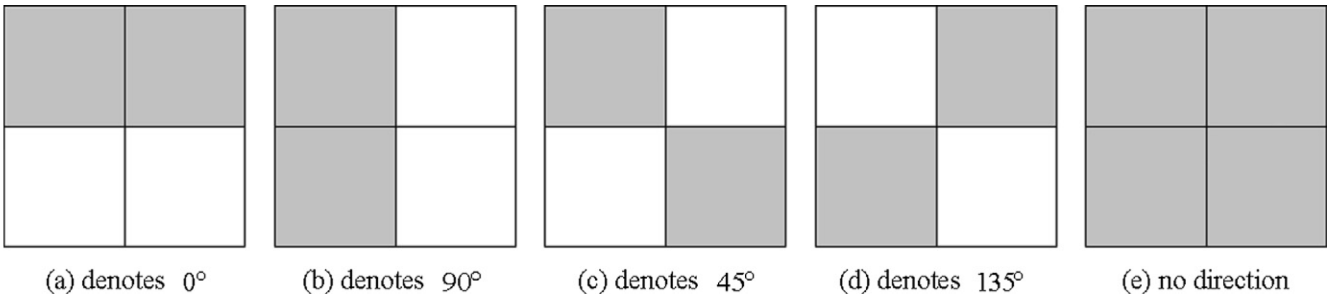

Fig. 1. Five structure elements in SED.

extracting the feature vector of images, thus the discrimination power of TCM is not high enough for image retrieval in large scale image dataset. In this paper, quantized HSV color space is used to extract the color feature and texture feature simultaneously. In this way, the feature would be extract more effectively, so this method has enough discrimination power for image retrieval in large scale image dataset.

In this paper, the image is represent by SED. One main problem of the SED is how to define structure elements. It is well known that human visual system is sensitive to color and orientation, and orientation plays an important role in image description. Therefore, the structure elements are defined by five structure elements denoting five directions, respectively. SED can effectively represent image features, and can extract and describe color and texture features simultaneously. SED has five 2×2 matrixes and is displayed in Fig. 1.

The original color image is quantized into 72 colors in HSV color space, five structure element templates are used to detect the textons. To obtain a final described image of structure elements, we use a simple three-step strategy described as follows (in order to explain the principle in a simple way, we suppose that the image has only two values).

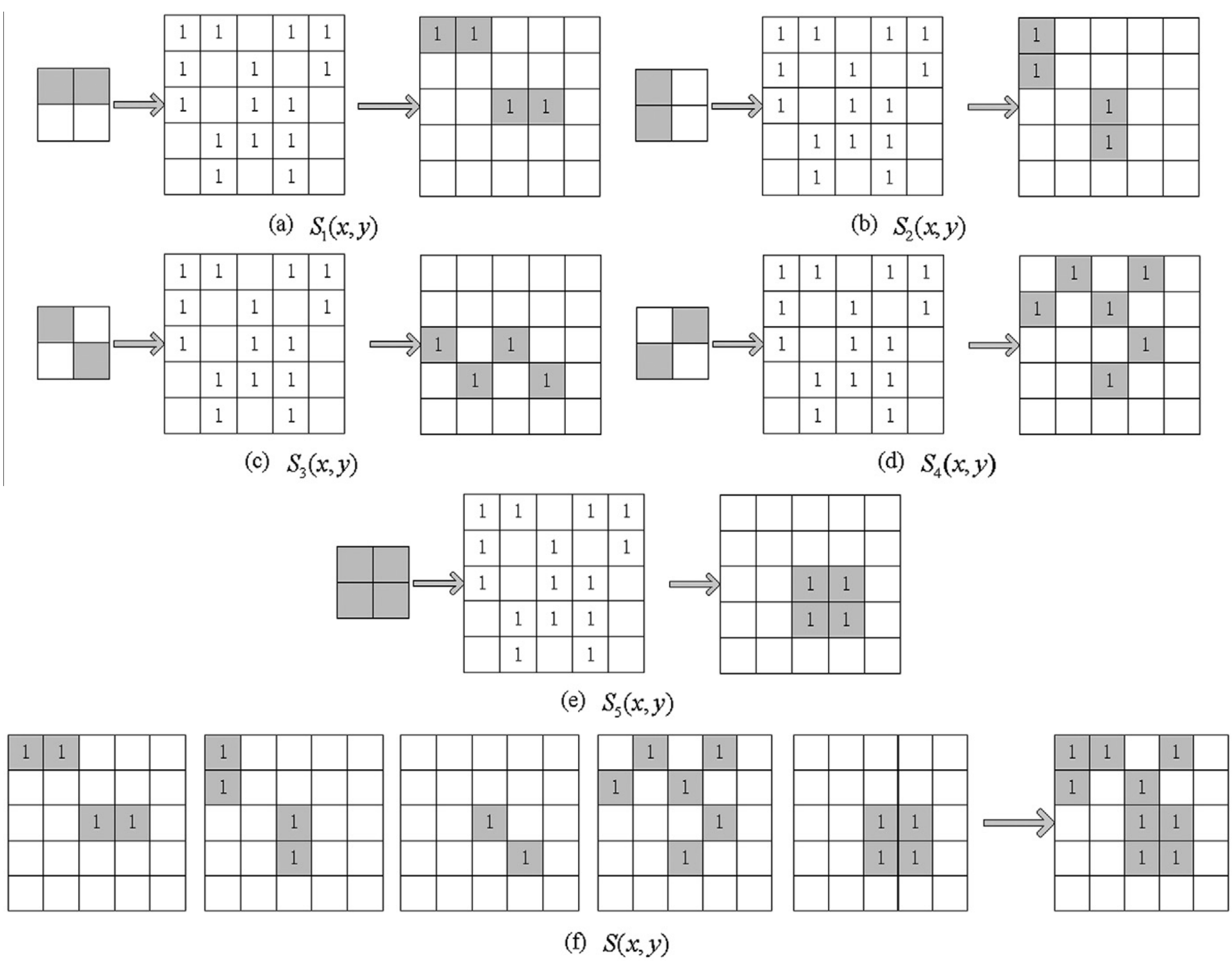

Fig. 2. The process of extracting SED map: (a), (b), (c), (d) and (e) are processes of extracting SED map using five structure elements, respectively. (f) is the final SED map.

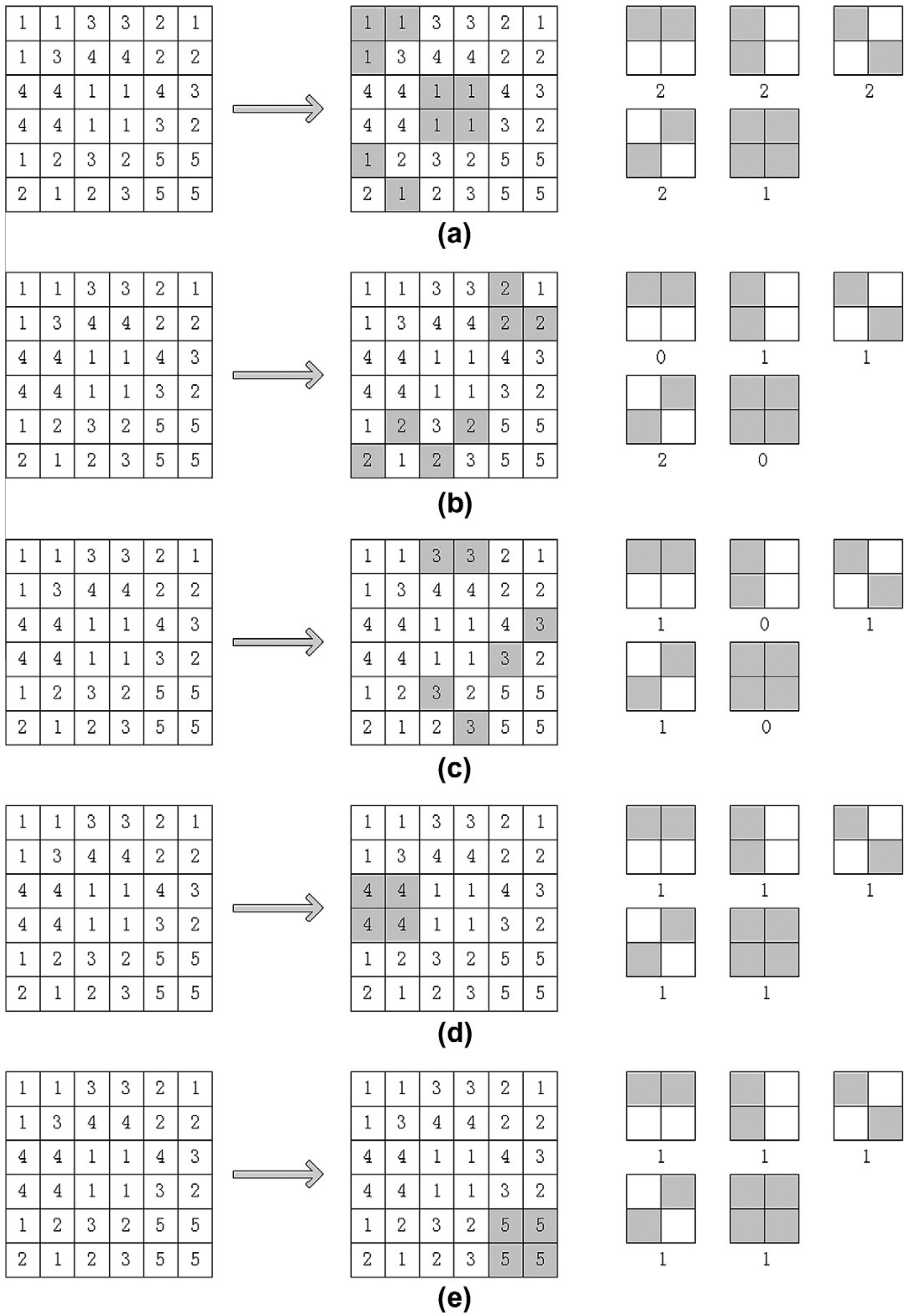

Fig. 3. SEH extraction method. (a), (b), (c), (d) and (e) are the processes of extracting SEH of five bins, respectively. The number of which under the structure element is the number of the structure element of every bin.

(1) Starting from the origin (0, 0), move the 2×2 SED from left to right and top to bottom with 2-step length.

(2) If the structure element matches the value of the image (match means that the values of the image in the corresponding structure element are equal), the value will be reserve, or not we will give up the value. Then we will obtain a SED map, denoted by $S_i(x, y)(1 \leqslant i \leqslant 5)$, because there are five structure elements. Specially, the other four structure elements should be detected when the structure element that denotes no direction has been detected, because no direction means that every direction is possible. In other words, a point can be considered as every direction, the reason is like that the other four structure elements should be detected when the structure element that denotes no direction has been detected.

(3) The final SED map, denoted by $S(x, y)$, is obtain by fusing five SED maps based on the following rule:

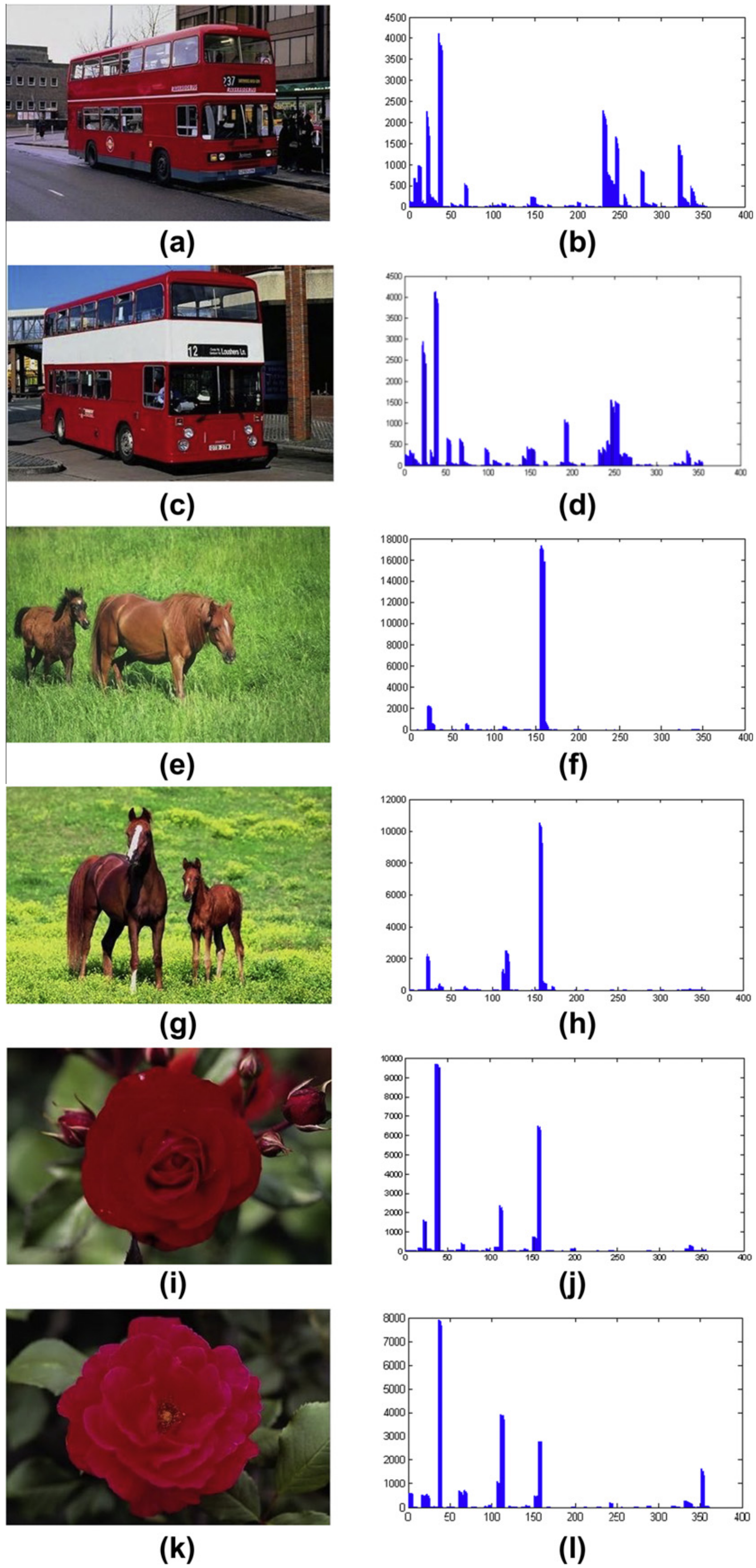

Fig. 4. The SEH of the images. (a), (c), (e), (g), (i) and (k) are the images and (b), (d), (f), (h), (j) and (l) are the corresponding SEH.

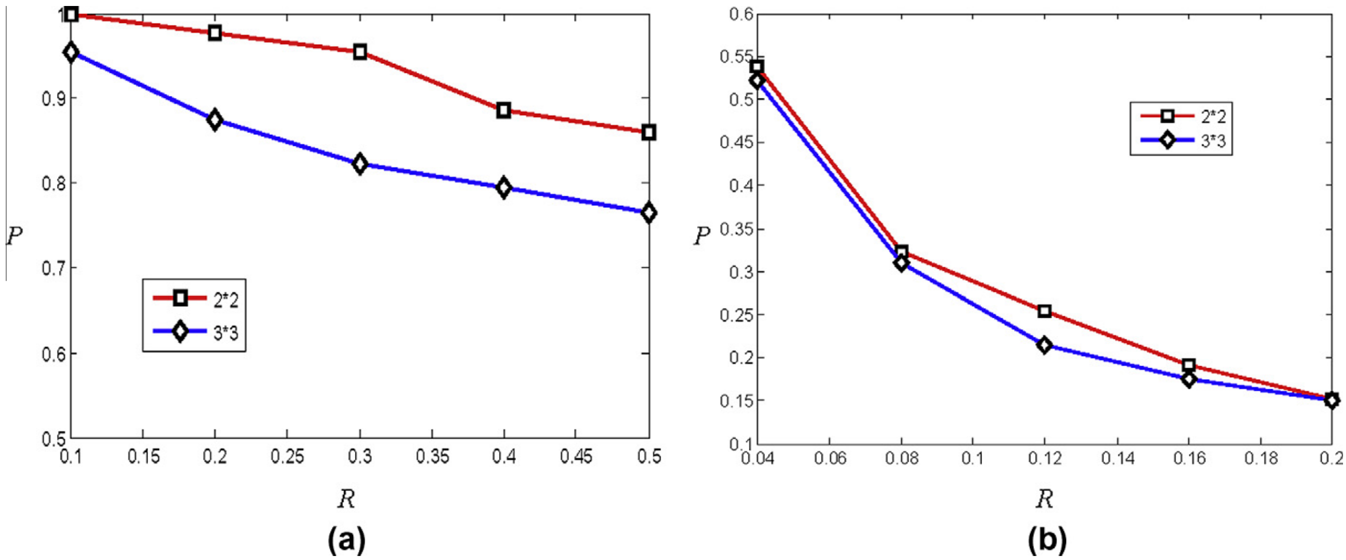

Fig. 5. The average retrieval precision and recall results of using 2×2 structure elements and 3×3 structure elements. (a) Corel-1000 dataset is used. (b) Corel-10000 dataset is used.

$$S(x,y) = \{(x,y)|S_1(x,y) \cup S_2(x,y) \cup S_3(x,y) \cup S_4(x,y) \cup S_5(x,y)\}. \tag{7}$$

Fig. 2 uses an example to illustrate the above SED map extraction process. Fig. 2(a) displays the process of extracting SED $S_1(x,y)$, Fig. 2(b), (c), (d) and (e) are the extracting processes of SED maps $S_2(x,y)$, $S_3(x,y)$, $S_4(x,y)$ and $S_5(x,y)$, respectively. Fig. 2(f) shows the fusion of the five maps to final SED map $S(x,y)$.

In the proposed algorithm, we use SED to describe the original image. SED has five structure elements, and the structure elements are used to detect the texton in the image, respectively. The SED map carefully explains the method of how to represent an image, effectively.

3.2. Structure elements' histogram (SEH)

The method of using SED has been displayed, and SED is used to detect the texton in the image. The SED map can be treated as a texture map of the image, so only the texture map is obtained and it cannot be feature vectors. The next step is to describe and extract the feature vectors of the image. In this part, SEH will be proposed and used to describe and compute feature vectors of the image in HSV color space.

The RGB color space has been quantized to 72 kinds of colors in HSV color space, we use $L_i(0 \leqslant i \leqslant 71)$ to denote the bin of HSV color space. SED is used to analyses the texture feature on image, and L_i is used to represent the color features on image. In order to extract texture and color features at the same time, SEH will be extracted on every bin L_i in HSV color space. To obtain the SEH of the image, we use a simple two-step strategy described as follows (in order to explain the principle in a simple way, we suppose that the image has been quantized to five bins and the size of image is 6×6).

(1) Count the number of SED on five SED maps when the value is L_i (here the number of bin is five and $0 \leqslant i \leqslant 4$, in real experiment the number of bin is 72 and $0 \leqslant i \leqslant 71$). Specially, the other four structure elements should be counted when the structure element that denotes no direction has been counted, because no direction means that every direction is possible. In the other word, a point can be considered as every direction, the reason is like that the other four structure elements should be count when the structure element that denotes no direction has been counted.

(2) Compute SEH based on the number of SED.

Fig. 3 is a simple example to explain the principle of extracting SEH. Fig. 3(a) shows the process of computing SEH of the simple image where the value is 1, and the result of SEH is $\{2,2,2,2,1\}$. Fig. 3(b)–(e) are the processes of computing SEH where the values are 2, 3, 4, 5. The final SEH of simple image is $\{2,2,2,2,1,0,1,1,2,0,1,0,1,1,0,2,1,1,1,1,1,1,1,1,1\}$.

SEH is extract from 72 bins in HSV color space, the five structure elements are extracted in every bin, so there are 360 elements extracted in HSV color space. Fig. 4 shows the SEH of the six examples image, and Fig. 4(a) and (c) are similar images, Fig. 4(e) and (g) are similar images, Fig. 4(i) and (k) are similar images, Fig. 4(b), (d), (f), (h), (j) and (l) are SEH of the corresponding image, respectively.

From Fig. 4, we can conclude that the SEH are similar when the images are similar to each other, such as Fig. 4(a) and (c), and that the SEH are very different when the images have much difference from each other, such as Fig. 4(a) and (e). Therefore, we can conclude that the SEH can represent image effectively.

SEH has been computed, at this time, the pixel number of the image will change when the image is scaled, therefore, the SEH should also change. In other words, images would be dissimilar when the image is scaled. In order to solve this problem better, we find that the proportion of pixels is same when the image is scaled. Therefore, the normalization should be taken to solve this problem. $L_i(0 \leqslant i \leqslant 71)$ denotes the bin of HSV color space, $|E_{1i}|$, $|E_{2i}|$, $|E_{3i}|$, $|E_{4i}|$ and $|E_{5i}|$ denote the number of SED on i bin, respectively. The normalization is defined as follows:

$$e_{ki} = \frac{|E_{ki}|}{\sum_{j=1}^{5}|E_{ji}|}, \tag{8}$$

where e_{ki} is the number of SED that has been normalized, k $(1 \leqslant k \leqslant 5)$ is the serial number of five structure elements in SED.

SEH has been normalized, at this time, the images are similar when the image is scaled, because the proportion of the SEH is same when the image is scaled.

4. Similarity measure

We use feature vector to represent an image, the two images will be similar when two images have similar feature vector. There are five structure elements in SED. In SEH, the number of structure

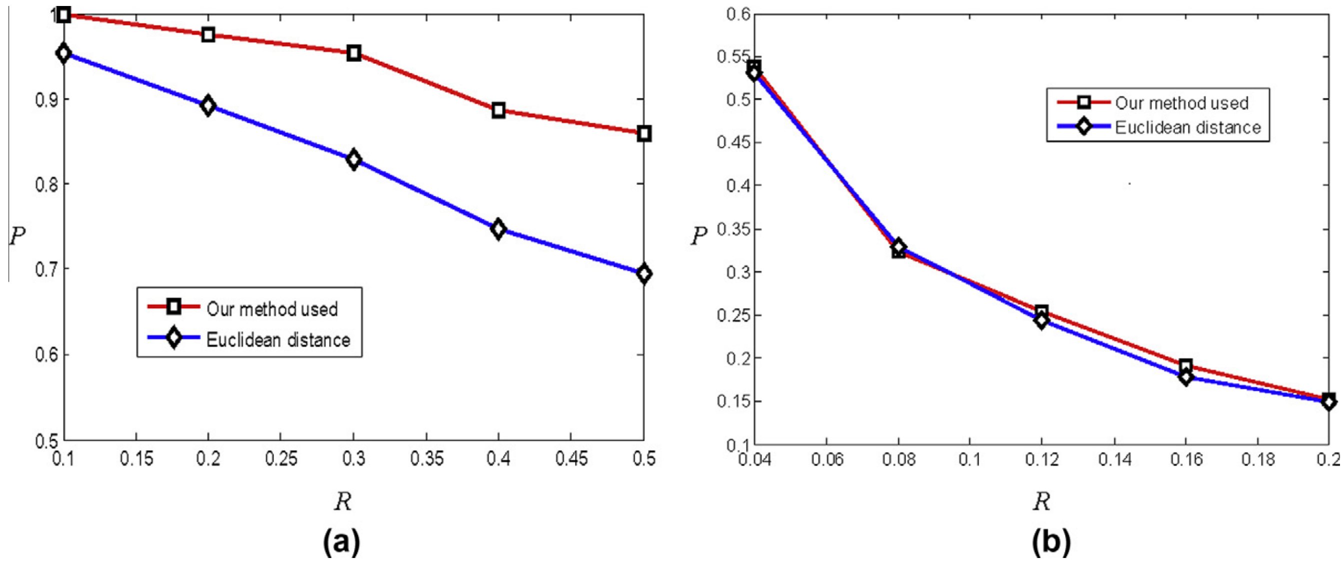

Fig. 6. The average retrieval precision and recall results of our method and Euclidean distance. (a) Corel-1000 dataset is used. (b) Corel-10000 dataset is used.

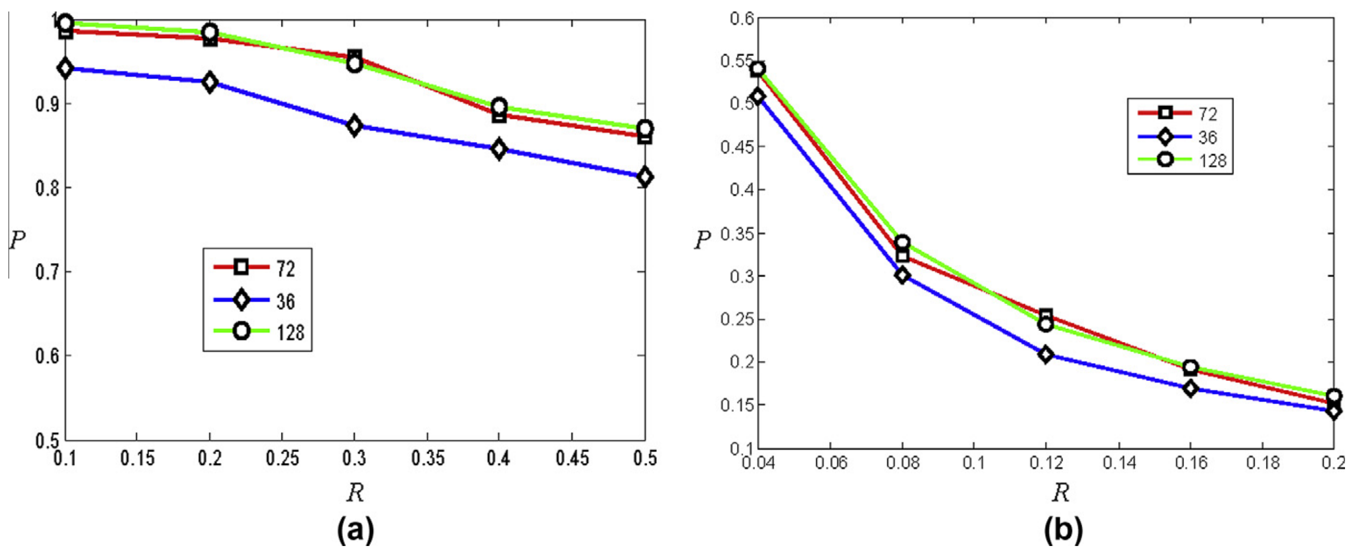

Fig. 7. The average retrieval precision and recall results on 72, 36 and 128 bins is used, while quantize HSV Color space. (a) Corel-1000 dataset is used. (b) Corel-10000 dataset is used.

elements is respectively counted in every bins and SEH has been normalized for similarity measure.

For each template image in the database, a 360-dimensional feature vector $F = \{f_1, f_2, \ldots, f_i, \ldots, f_{360}\}$ will be extracted and stored in the file. $T = \{t_1, t_2, \ldots, t_i, \ldots, t_{360}\}$ is the feature vector of the query image.

In this paper, the distance between the query image and the dataset's image is defined as follows:

$$D(F,T) = \sum_{i=1}^{360} \frac{|f_i - t_i|}{1 + f_i + t_i}. \tag{9}$$

The above formula is simple to compute the similarity, which needs no square and square root operations. Our method can use this formula to effectively measure similarity. The Euclidean distance is widely used in similarity measuring method for image retrieval, the formula is displayed as follows.

$$D'(F,T) = \sqrt{\sum_{i=1}^{360} (f_i - t_i)^2}. \tag{10}$$

The Euclidean distance measuring method is more complex than the method we use. Moreover, the method we use is more effective than Euclidean distance. And our similarity measuring method has a better performance than Euclidean distance. The compared results will be displayed in our experiments.

5. Experimental results

There are many image datasets CBIR systems, such as UCID [26], NISTER [27], Corel image dataset, etc., the widely used is Corel image dataset. In our algorithm, experiments are carried out by using two Corel image datasets. The first one is Corel-1000 dataset, which is divided into 10 categories including landscapes, horses, elephants, human beings, bus, flowers, buildings, mountains, food and dragons, and each category contains 100 images. Another one is Corel-10000 dataset, which contains 100 categories including beach, car, fish, door sunset, etc., and each category contains 100 images. Experimental images cover a variety of contents.

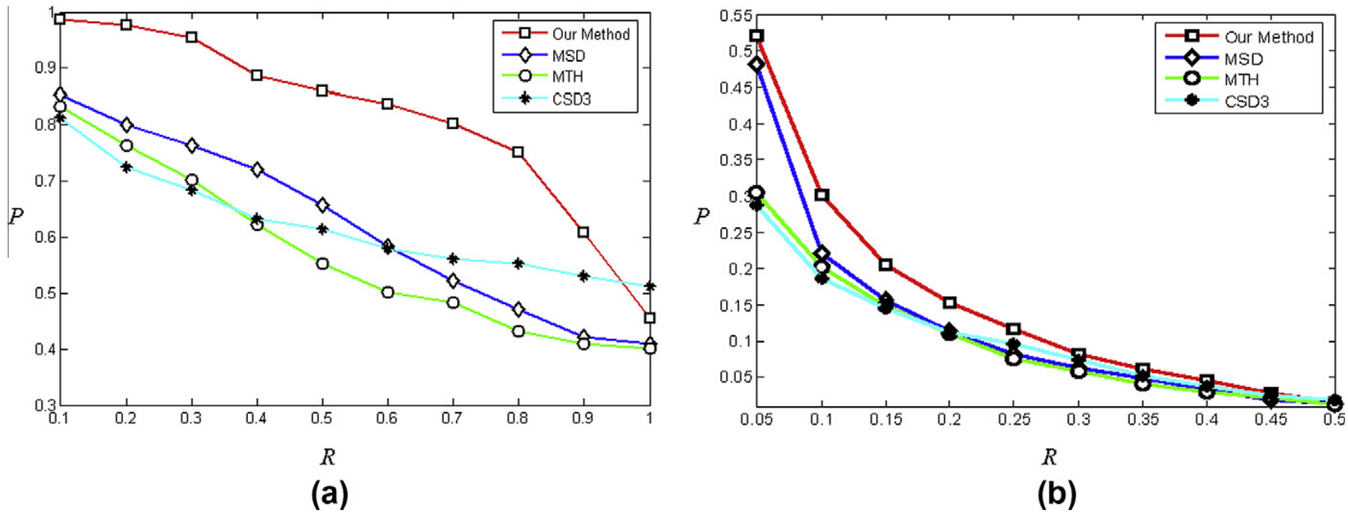

Fig. 8. The average retrieval performance comparison of four methods. (a) Corel-1000 dataset is used. (b) Corel-10000 dataset is used.

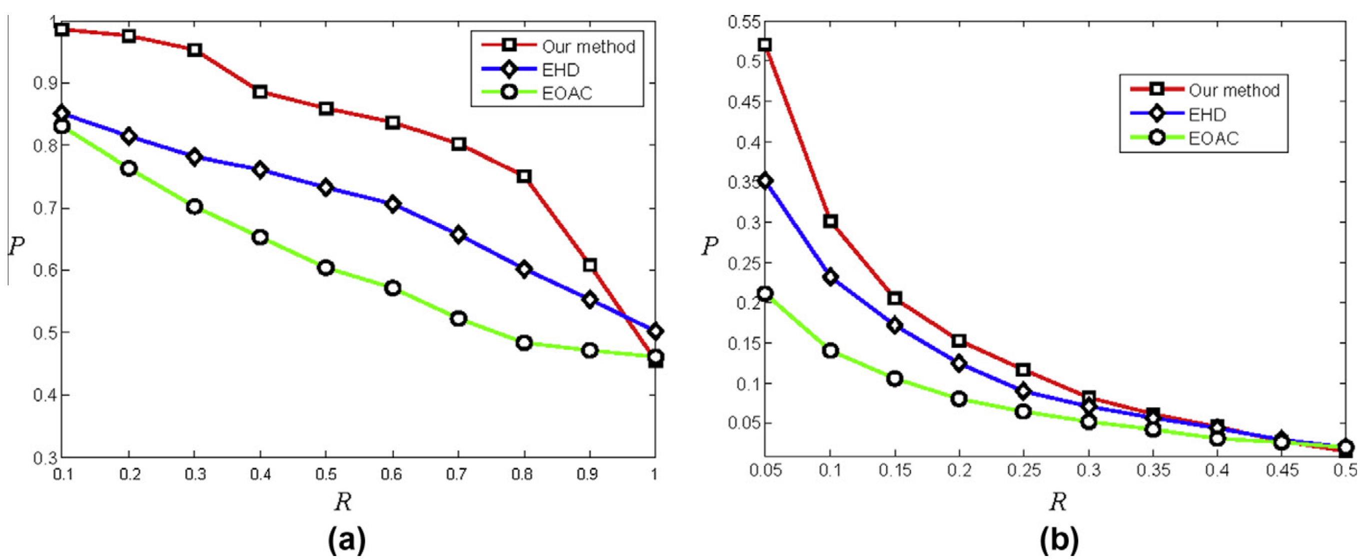

Fig. 9. The average retrieval performance comparison of the three methods. (a) Corel-1000 dataset is used. (b) Corel-10000 dataset is used.

In our experiment, we randomly choose 10 images from each category in the Corel-1000 dataset and use them as query images. For each category, we compute the precision and recall percentage of images. Then, the mean precision-recall pair percentage is computed by the precision-recall pair that has been obtained by the retrieval results of 10 random images. In Corel-10000 dataset, 20 kinds of categories are chosen randomly from 100 categories. We choose 10 images from each category that has been chosen and use them as query images. Then, the mean precision-recall pair percentage is computed.

The commonly used performance measurement, precision-recall pair, is used for evaluation of retrieval performance. The precision P is defined as ratio between the number of the retrieved relevant images M and the total number of the retrieved images N; it measures the accuracy of the retrieval. Recall R is defined as ratio between the number of the retrieved relevant images M and the total number of the relevant images S of the whole database; it measures the robustness of the retrieval. Precision P and recall R are computed by the following equations:

$$P = \frac{M}{N}, \tag{11}$$

$$R = \frac{M}{S}. \tag{12}$$

Fig. 5 displays the results of comparing the experimental results of using 2×2 structure elements and 3×3 structure elements. In experiment, the performance of using 2×2 structure elements is higher than using 3×3 structure elements, because more useful information can be extracted when 2×2 structure elements are used.

Fig. 6 displays the experimental results of comparing our similarity measuring method with the Euclidean distance measuring method. From the results, we can conclude that the performance of our method is higher than Euclidean distance similarity measuring method. Although the Euclidean distance is widely used in similarity measuring, our similarity measuring method is more appropriate than Euclidean distance and takes a better performance in experiment.

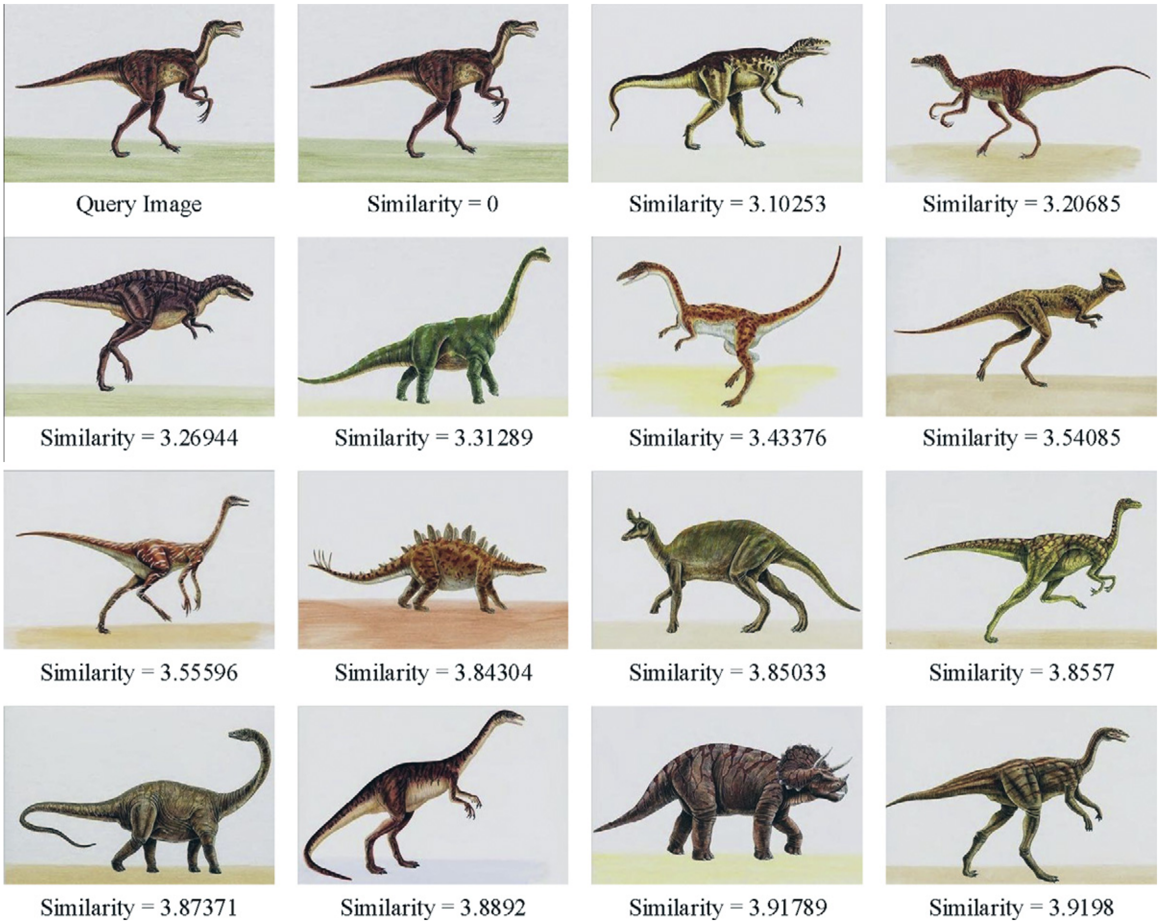

Fig. 10. Image retrieval for dragons.

Fig. 7 displays the experimental results of using different number of bins to which HSV Color space is quantized. In the experiment, the HSV Color space is quantized to 36, 72 and 128 bins, respectively. Fig. 7 shows that the performance of using 36 bins is lower than using 72 and 128 bins, the performance of using 72 bins is as good as using 128 bins, and that the complexity of using 128 bins is higher than using 72 bins.

As we all know, local feature descriptors, such as SIFT and SURF descriptors, were originally developed for texture classification and image matching. In the SIFT and SURF descriptors, there are 128 and 64 dimensional feature vector in each keypoint [2]. Therefore, a very high dimensional descriptor will be created when the SIFT and SURF descriptors are used in large scale image retrieval, because an image will have many keypoints. It is not suitable to use SIFT and SURF for image retrieval directly.

Fig. 8 displays the experimental results of comparison between our method and methods MTH in [1], MSD [2] and CSD3 in [7]. The retrieval results show that the proposed method has a better performance in image retrieval than other two methods. In Fig. 8(a), the top 10 images are relevant to the query image in our proposed method. In MSD, MTH and CSD3, the precisions of top 10 images are 85%, 83% and 81%, respectively. The precisions reach 39% and 37% in MSD and MTH, and the recalls reach 100%. In CSD3, the precision reaches 51%, which is better than our method, and the recall reaches 100%. But our method has a better performance in global. Fig. 8(b) shows that the performance of the proposed method is higher than other methods when Corel-10000 dataset is used. Therefore, we can conclude that the performance of our proposed method is higher than other three methods.

From Fig. 8, the proposed method outperforms MSD [1], MTH [2] and CSD3 [7]. In the proposed method, the feature vector is extracted by SEH and combines the color information with SED. SED is composed of five structure elements that denote the five different directions, and the texture feature can be extracted more effectively. Seventy-two kinds of main colors are quantized in HSV color space, and SEH is extracted on every bin. Therefore, SEH can represent images in color and texture simultaneously.

In MSD [1], the micro-structure descriptor is proposed, the MSD is defined by the underlying colors in micro-structures with similar edge orientation, and the feature vector is extracted on edge orientation, so the feature vector is the part feature of the image. In proposed method, the feature vector is the global feature of the image. In MTH [2], the multi-texton histogram is used for representation images, it works directly on natural images as a shape descriptor and hard to fully represent the content of images. In CSD3 [7], the method describes the image by color and spatial distributions. In MTH and CSD3, the feature vector is extracted on single feature, shape or color. In proposed method, the feature vector is extracted on color and texture features simultaneously, and it can represent more effectively. Therefore, the proposed method outperforms MSD [1], MTH [2] and CSD3 [7].

Fig. 9 displays the experimental results of comparison among the proposed method, MPEG-7 descriptors EHD (edge histogram descriptor) [21] and EOAC (edge orientation autocorrelogram) method [22]. From Fig. 9, we can conclude that the performance of proposed method is higher than EHD and EOAC.

EHD captures the spatial distribution of edges and it is useful for image matching [21]. EHD can represent shape information better,

Fig. 11. Image retrieval for buses.

but it cannot well represent the color and texture information. EHD is useful for image matching, but the matching process is very slow, so it is not suitable for image retrieval. Therefore, EHD is not a good idea for image retrieval.

Correlogram is defined in [28], and it is used for image indexing and comparison. EOAC is an improved correlogram. EOAC represents edge features with their orientations and correlation between neighboring edges, and it can well represent the shape information of the image, but the color and texture features will be lost [22]. EOAC is not appropriate for image retrieval with texture or color features. The spatial correlation of edge can represent image features partially. So the proposed method outperforms EHD and EOAC.

Finally, three examples of image retrieval are displayed and Corel-1000 dataset is used. The similarity is computed between the query image and the image database, and we then ranks them. The top 15 images are displayed in our experiment. Three different kinds of images are tested with our proposed retrieval method. The results are displayed in Figs. 10–12.

From experiment, the precision of retrieving by dragon and bus is higher than the precision of retrieving by landscape. The top 15 retrieval images of dragons and buses are relevant to the query image, entirely. And there are 12 images relevant to the query image of landscape when the top 15 images of landscapes are retrieved. The reason is that images of dragon and bus have significant regions, and these images have significant texture, but the texture of landscape is not significant. Therefore, we can conclude that the precision is higher when the query image has more significant regions and stronger texture than those images whose regions are not significant.

In summary, the proposed method is more suitable for these images that have significant regions, significant objects and strong texture, such as, dragons, buses and flowers. For these images, our method will have a better performance and higher precision.

6. Conclusion

In this paper, a novel texture descriptor SED is proposed. SED is an effective descriptor for describing the image. SED can effectively represent image local features, and it can extract and describe color and texture features. SED is composed of five structure elements which denote five directions. The image's SEH is computed by SED, and HSV color space is used which has been quantized to 72 bins. There are five structure elements in SED, and the SED is extracted in every bins, therefore, the feature vector is a 360-dimensional feature vector. SEH integrates the advantages of both statistical and structural texture description methods, and represents the spatial correlation of color and texture. In the similarity measuring method, a simple method is used to calculate the similarity between the query image and images in the database. The simple method has a better performance than other methods, such as Euclidean distance measuring method. The retrieval performance of our approach based on SED is tested and compared with several other retrieval methods. The experimental results show that the proposed algorithm has a better performance than other retrieval methods.

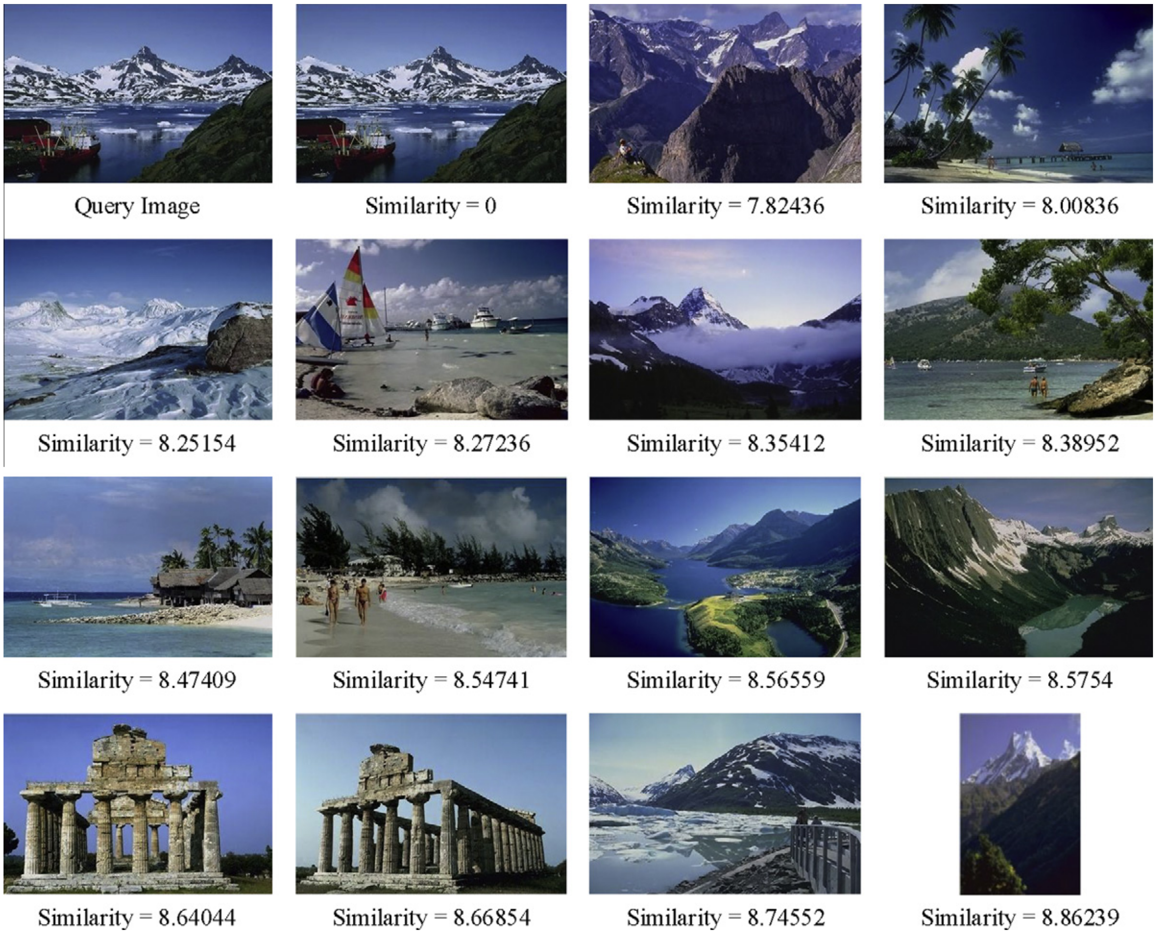

Fig. 12. Image retrieval for landscapes.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (Nos. 61173183, 60973152, and 60573172), the Superior University Doctor Subject Special Scientific Research Foundation of China (No. 20070141014), Program for Liaoning Excellent Talents in University (No. LR2012003), the National Natural Science Foundation of Liaoning Province (No. 20082165) and the Fundamental Research Funds for the Central Universities (No. DUT12JB06).

References

[1] G.H. Liu, L. Zhang, Y.K. Hou, Z.Y. Li, J.Y. Yang, Image retrieval based on multi-texton histogram, Pattern Recognition 43 (7) (2010) 2380–2389.
[2] G.H. Liu, Z.Y. Li, L. Zhang, Y. Xu, Image retrieval based on micro-structure descriptor, Pattern Recognition 44 (9) (2011) 2123–2133.
[3] R. Kwitt, P. Meerwald, A. Uhl, Efficient texture image retrieval using copulas in a Bayesian framework, IEEE Transactions on Image Processing 20 (7) (2011) 2063–2077.
[4] E. Aptoula, S. Lefèvre, Morphological description of color images for content-based image retrieval, IEEE Transactions on Image Processing 18 (11) (2009) 2505–2517.
[5] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, C. Roux, Fast wavelet-based image characterization for highly adaptive image retrieval, IEEE Transactions on Image Processing 21 (4) (2012) 1613–1623.
[6] M. Heikkilä, M. Pietikäinen, C. Schmid, Description of interest regions with local binary patterns, Pattern Recognition 42 (3) (2009) 425–436.
[7] C.H. Lin, D.C. Huang, Y.K. Chan, K.H. Chen, Y.J. Chang, Fast color-spatial feature based image retrieval methods, Expert Systems with Applications 38 (9) (2011) 11412–11420.
[8] G. Michèle, Z. Bertrand, Body color sets: a compact and reliable representation of images, Journal of Visual Communication and Image Representation 22 (1) (2011) 48–60.
[9] Z. Konstantinos, E. Kavallieratou, P. Nikos, Image retrieval systems based on compact shape descriptor and relevance feedback information, Journal of Visual Communication and Image Representation 22 (5) (2011) 378–390.
[10] A. El-ghazal, O. Basir, S. Belkasim, Invariant curvature-based Fourier shape descriptors, Journal of Visual Communication and Image Representation 23 (4) (2012) 622–633.
[11] T. Celik, T. Tjahjadi, Bayesian texture classification and retrieval based on multiscale feature vector, Pattern Recognition Letters 32 (2) (2011) 159–167.
[12] N.D. Thang, T. Rasheed, Y.K. Lee, S.Y. Lee, T.S. Kim, Content-based facial image retrieval using constrained independent component analysis, Information Sciences 181 (15) (2011) 3162–3174.
[13] K. Zhan, H.J. Zhang, Y.D. Ma, New spiking cortical model for invariant texture retrieval and image processing, IEEE Transactions on Neural Networks 20 (12) (2009) 1980–1986.
[14] R. Kwitt, A. Uhl, Lightweight probabilistic texture retrieval, IEEE Transactions on Image Processing 19 (1) (2010) 241–253.
[15] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, C. Roux, Adaptive nonseparable wavelet transform via lifting and its application to content-based image retrieval, IEEE Transactions on Image Processing 19 (1) (2010) 25–35.
[16] S. Murala, R.P. Maheshwari, R. Balasubramanian, Local tetra patterns: a new feature descriptor for content-based image retrieval, IEEE Transactions on Image Processing 21 (5) (2012) 2874–2886.
[17] X.F. He, Laplacian regularized D-optimal design for active learning and its application to image retrieval, IEEE Transactions on Image Processing 19 (1) (2010) 254–263.
[18] W. Bian, D.C. Tao, Biased discriminant Euclidean embedding for content-based image retrieval, IEEE Transactions on Image Processing 19 (2) (2010) 545–554.
[19] D.S. Zhang, M.M. Islam, G.J. Lu, A review on automatic image annotation techniques, Pattern Recognition 45 (1) (2012) 346–362.

[20] J.L. Liu, D.G. Kong, Image retrieval based on weighted blocks and color feature, in: International Conference on Mechatronic Science, Electric Engineering and Computer, Jilin, 2011, pp. 921–924.
[21] 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.
[22] F. Mahmoudi, J. Shanbehzadeh, Image retrieval based on shape similarity by edge orientation autocorrelogram, Pattern Recognition 36 (8) (2003) 1725–1736.
[23] D.G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60 (2) (2004) 91–110.
[24] G.H. Liu, J.Y. Yang, Image retrieval based on the texton co-occurrence matrix, Pattern Recognition 41 (12) (2008) 3521–3527.
[25] W.T. Chen, W.C. Liu, M.S. Chen, Adaptive color feature extraction base on image color distributions, IEEE Transactions on Image Processing 19 (8) (2010) 2005–2016.
[26] G. Schaefer, M. Stich, UCID – an uncompressed colour image database, in: Conference on Storage and Retrieval Methods and Applications for Multimedia, San Jose, 2004, pp. 472–480.
[27] D. Nistér, H. Stewénius, Scalable recognition with a vocabulary tree, in: IEEE Conference on Computer Vision and Pattern Recognition, Washington, 2006, pp. 2161–2168.
[28] J. Huang, S.R. Kumar, M. Mitra, W.J. Zhu, R. Zabih, Image indexing using color correlograms, in: IEEE Conference on Computer Vision and Pattern Recognition, San Juan, 1997, pp. 762–768.