Image descriptor based on local color directional quaternionic pattern

Shan Zhao
Jing Liu

Image descriptor based on local color directional quaternionic pattern

Shan Zhao* and Jing Liu
Henan Polytechnic University, School of Computer Science and Technology, Jiaozuo, China

Abstract. Most of the existing local binary pattern methods discard local color differences by holding their binary information. And local spatial information is also neglected. To address these problems, a robust color image descriptor local color directional quaternionic pattern (LCDQP) is proposed. In the descriptor, the color distance map (CDM) is generated in the RGB color space to capture the color distribution among three channels. According to the distribution of CDM elements and their neighbors, four edge models are defined to describe the change trend of the original image. Then, based on this, the LCDQP strings are gained according to the distribution of four models in the four directions of 0 deg, 45 deg, 90 deg, and 135 deg. Finally, an effective quaternionic code method is adopted to construct the LCDQP descriptor. The proposed descriptor not only captures the local color features but also reflects the spatial structure information. Experiments on four representative databases demonstrate that the proposed descriptor is superior to other state-of-the-art approaches. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JEI.28.4.043003]

Keywords: image descriptor; local binary pattern; local color directional quaternionic pattern.

Paper 190141 received Feb. 14, 2019; accepted for publication Jun. 13, 2019; published online Jul. 3, 2019.

1 Introduction

Image descriptors are important in the fields of image processing, pattern recognition, and computer vision. A good descriptor should be distinctive and it should capture the most representative characteristics of the image appearance. A large number of image descriptors have been proposed and successfully applied to different research fields, such as image retrieval, face recognition, image classification, and many others. Texture, a prominent property of images, can be recognized in the form of small repeated patterns, and it is a key component for human visual perception. Many texture descriptors have been developed for texture analysis in recent years. Local descriptors are efficient texture descriptors, and have achieved excellent performance in a variety of computer vision applications. As a kind of local descriptor, local binary pattern (LBP), first proposed by Ojala et al,[1,2] has attained more attention in recent years. The LBP method has been used widely due to the high performance and effective computation. It has been successfully applied to many computer fields, such as description of salient regions,[3] face recognition,[4,5] texture analysis, classification,[6–8] and so on. LBP has proven to be a robust and computationally simple approach to describe local structures. However, it discards local intensity differences by holding their binary information with higher complexity. Furthermore, color information is also ignored because only gray information is used.

To improve the performance of LBP, lots of modified versions of LBP descriptors have been proposed. Heikkilä et al.[9] introduced the center-symmetric local binary pattern descriptor for matching and object category classification. The texture descriptor appears to be more robust to illumination and occlusion by combining the good properties of

the SIFT and LBP. Guo et al. developed a completed local binary pattern scheme,[10] which combines three features for rotation invariant texture classification. Significant performance improvement can be achieved. Tan and Triggs proposed local ternary pattern (LTP),[4] which quantizes the image pixel gray-level differences between center and neighbor pixel into a ternary value instead of a binary one. Murala et al.[11] proposed a new algorithm using directional local extrema patterns for content-based image retrieval applications, which extracts the directional edge information, based on local extrema in 0 deg, 45 deg, 90 deg, and 135 deg directions in an image, and shows a significant improvement. Huang et al.[12] proposed local circular pattern (LCP), which improves the LBP and its variants by replacing the binary quantization with a clustering method, resulting in higher discriminative ability as well as better robustness to noise. Ryu et al.[13] presented local directional ternary pattern (LDTP) for face recognition, which efficiently uses the directional information and ternary pattern in order to take advantage of the robustness of edge patterns in the edge region while overcoming weaknesses of edge-based methods in smooth regions. Fadaei et al.[14] proposed local derivative radial pattern for texture representation, which is based on gray-level difference of pixels along a line and their weighted combinations. Khadiri et al.[15] combined the three level descriptions from LTP and the directional features from LDP to form LDTP for texture classification.

The above discussed LBPs have proven to be highly discriminative features for texture classification and image feature description. However, because they threshold at exactly the value of the central pixel they tend to be sensitive to noise, particularly in near-uniform image regions, and smooth weak illumination gradients. Furthermore, these works deal with integrating texture in gray images, fewer works are dedicated to color LBP features. As we know, color is an important property that contains essential

---

*Address all correspondence to Shan Zhao, E-mail: zhaoshan@hpu.edu.cn

information of images. Existing works have demonstrated that color information can provide more discriminant information and a better invariance to illumination changes for image representation. To extract LBP from color images, one strategy is to convert the color image into the corresponding gray-scale image or to simply use one color component for further processing. However, some color information still has been ignored. Another strategy is to select an appropriate color space, then extract the feature of each spectrum with LBP, separately, and finally concatenate the feature vector of each spectrum together into one for color image representation. However, it ignores the complex relationship between different spectra of the color images. Furthermore, it can increase the dimensionality of the pattern and result in higher complexity. Much work has been attempted to extend the original LBP descriptor to the color image representation. CaLBP applied LBP operator independently on each of the different color spectra,[16] and LBP histogram of each spectrum was concatenated to form final image feature vector. Lee et al.[17,18] proposed a local color vector binary pattern (LCVBP) that is derived by applying LBP to the color image norm and angular features, respectively. Quaternionic local ranking binary pattern used quaternionic representation of the color images to obtain color texture features.[19] Completed local similarity pattern (CLSP) described the color image as the co-occurrence of its image pixel color quantization information and the local color image textural information.[20] Dubey et al.[21] proposed a method for image description with multichannel decoded LBPs by introducing two schemas for the combination of the LBPs from more than one channel. Singh et al.[22] used a plane to threshold color pixels in the neighborhood of a local window into two categories, and the uniform patterns of the two proposed descriptors are combined to yield another lower color descriptor. Yu et al.[23] mimicked the visual perception of human to describe images by constructing a set of multitrend descriptors that are encoded with binary codes.

Although the descriptors above offer a great improvement on the performance of LBP for the color images, it is obvious that they neglect the spatial relation of the color frequency of each channel. The edge information is not combined with the local descriptors. All of them can reduce the efficiency. In order to capture the cross-channel co-occurrence information and combine the color and spatial correlation information with the texture feature, we propose a simple and robust descriptor local color directional quaternionic pattern (LCDQP), by which four edge patterns for a color image are derived and the spatial distribution of patterns is considered. Our proposed descriptor can be summarized as follows. First, three color components are combined to create a color distance map (CDM). Four edge models are defined according to the distribution of the neighbor elements of CDM. Then, the statistical distribution of four edge models in four directions of 0 deg, 45 deg, 90 deg, and 135 deg is achieved. Finally, a quaternionic code method is presented to express the statistical strings and LCDQP is introduced. Experiments show that LCDQP is more efficient than the traditional descriptors. The major intention of this paper is to contribute a valuable new framework that combines color and texture with spatial information to progress the performance of the descriptor. The contributions of this paper are as follows:

- CDM is presented, which combines the difference of R, G, and B three channels of color images, to increase the effectiveness of color components to discriminate color image.
- According to the value difference between elements of CDM, four models are defined and their distribution of four directions is derived from CDM, which fully captures the edge and texture of local image region.
- Quaternionic code method is introduced for the statistical strings, which can make full use of the spatial distribution information.

This paper is organized as follows. Section 2 introduces the related works. The proposed method is introduced in Sec. 3. In Sec. 4, the performance of proposed method is evaluated on four databases. Section 5 concludes the paper.

2 Related Works

In this section, we will review original LBP and LTP. LBP quantizes the image pixel gray-level differences between center and neighbor pixel into a binary one. LTP quantizes the differences into a ternary value instead of the binary one. Inspired by the LDP and LTP, the LCDQP strings are achieved by quantizing the differences between center and neighbor element of CDM into a quaternionic value in four directions. In this sense, we introduce the LBP and LTP in Sec. 2.

Indeed, LBP and LTP are two descriptors for an image, which describe the image with histogram not an image. In this section, in order to verify the efficiency of LBP and LTP from another aspect, the LBP image and LTP image are given. The LBP image is an image whose pixel value is its LBP value. In the same way, the LTP image is an image whose pixel value is expressed by its LTP value. From LBP image and LTP image, we can see that the abundant texture information of original image is described well in them.

2.1 Local Binary Pattern

LBP is an effective method for feature extraction through encoding the pixelwise information in texture images that characterizes the local structure. For one pixel in a gray image, the values of its eight neighboring pixels are operated by the value of the central pixel as threshold. Precisely, for each neighboring pixel, the result will be set to one if its value is no less than the value of the central pixel, otherwise the result will be set to zero. The LBP code of the central pixel is then obtained by multiplying the results with weights given by powers of two and summing them up together. The LBP operator may be extended to a circular neighboring pixels.

One of the important advantages of LBP is its invariance against monotonic gray level changes caused by illumination variations. It is computationally simple and very discriminant, which makes it reach high performance in texture classification. The original LBP operator is defined as follows:[2]

$$\text{LBP}_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \qquad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad (1)$$
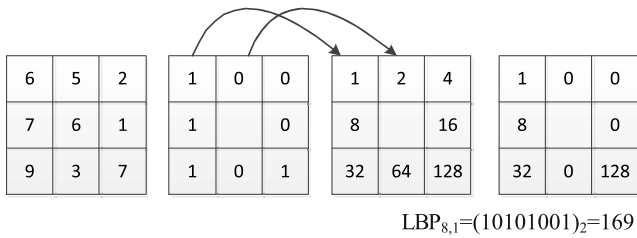
LBP$_{8,1}$=(10101001)$_2$=169

Fig. 1 Illustration of LBP$_{8,1}$.

Fig. 3 Same LBP value with large different gray blocks.

where g_c and g_p are the gray values of center pixel and its neighborhoods, respectively; P and R are the number of neighboring pixels and radius from neighboring pixels, respectively. The distribution is approximated by the local neighborhood differences in order to make the operator invariant to changes of the mean gray value. An example of LBP for a given 3×3 pattern is given in Fig. 1. The original images and their LBP images are given in Fig. 2.

LBP encodes the relationship between gray values of center pixel and neighboring pixels into 0 and 1. So it is a simple and efficient local descriptor for describing texture. But it lacks the local intensity difference and results in the same LBP value block with large different gray block. An example is given in Fig. 3.
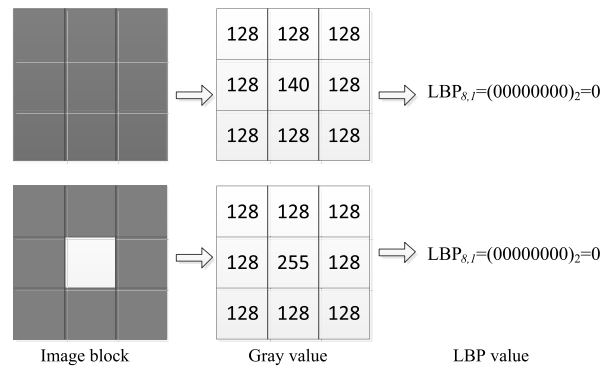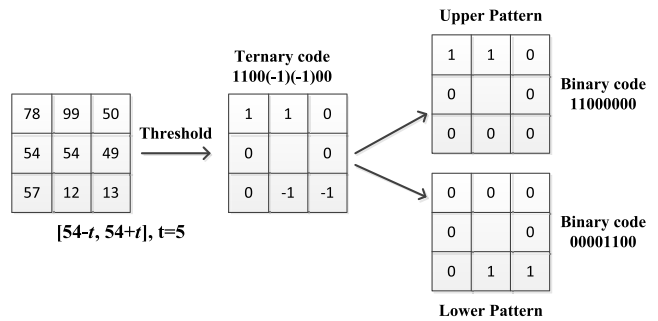
2.2 Local Ternary Pattern

LTP extends LBP to three-value codes,[4] in which gray-levels in a zone of width $\pm t$ around i_c are quantized to zero, ones above this are quantized to $+1$, and ones below it to -1. The indicator S(u) is replaced with a three-valued function and the binary LBP code is replaced by a ternary LTP code, as shown in Eq. (2). Here, t is a user-specified threshold, so LTP codes are more resistant to noise, but no longer strictly invariant to gray-level transformations.

Fig. 4 Illustration of the LTP.

The LTP encoding procedure and LTP images are shown in Figs. 4 and 5:

$$\mathrm{S}(u, i_c, t) = \begin{cases} 1, & u \geq i_c + t \\ 0, & |u - i_c| < t \\ -1, & u \leq i_c - t \end{cases} \tag{2}$$
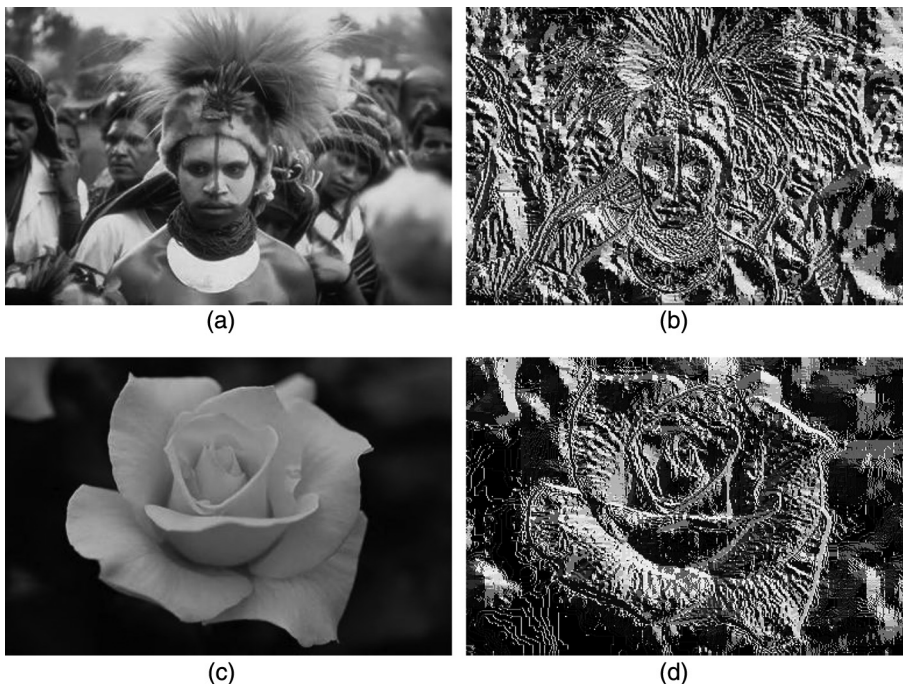

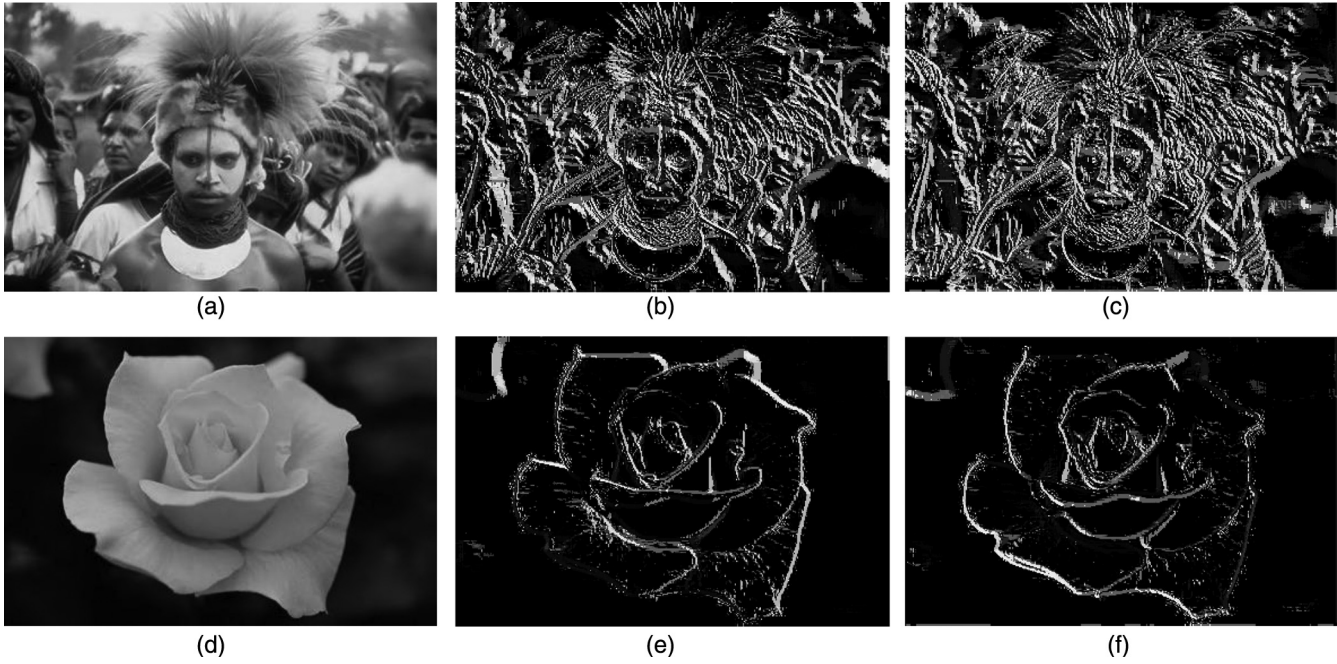**Fig. 2** Original images and LBP images: (a) and (c) original images, (b) and (d) LBP images.

Fig. 5 Original images and LTP images: (a) and (d) original images, (b) and (e) LTP positive images, (c) and (f) LTP negative images.

3 Local Color Directional Quaternionic Pattern

In our proposed descriptor, to effectively explore the inter-channel correlation of pixel values of three color channels at the same location, we create a CDM combining the joint information of each channel to improve the performance of descriptor according to the color distance of three color channels.

In order to construct CDM, it is necessary to select a color space. Here, several color spaces can be imagined, such as RGB, HIS, HSV, YUV, and so on. Because pixel value of R, G, B describes the quantized strength of the light at different wavelengths, so it can be directly compared. Even though some other color spaces achieve better performance than RGB color space, their pixel values are not quantitatively comparable.[5] Take YUV color space as an example, Y describes luminance information, whereas U and V describe chrominance information.[24] Thus Y, U, and V have totally different physical meanings and direct comparisons between pixel values of Y, U, and V may cause some problems. Therefore, the RGB color space is chosen in our method.

Supposed I is an RGB color image, I_p is one of the pixels whose color value is v_p, where $v_p = (v_{p,1}, v_{p,2}, v_{p,3})$ for three channels R, G, and B, p is the index of the pixels. The mean of channel are v_{mean}, similarly, where $v_{\text{mean}} = (v_{\text{mean},1}, v_{\text{mean},2}, v_{\text{mean},3})$ for three channels R, G, and B. The color distance of three channels is calculated between the pixel I_p and its mean gray of each channel as follows:[20]

$$d_p = D(v_p - v_{\text{mean}}) = \frac{|v_p - v_{\text{mean}}|}{|v_p| + |v_{\text{mean}}|}$$

$$= \frac{\sqrt{\sum_{k=1}^{N} (v_{p,k} - v_{\text{mean},k})^2}}{\sqrt{\sum_{k=1}^{N} (v_{p,k})^2} + \sqrt{\sum_{k=1}^{N} (v_{\text{mean},k})^2}}, \quad (3)$$

where N is the number of the color channels, here, for RGB color space, $N = 3$.

An example of the construction of CDM is given in Fig. 6. We take the first pixel as an example, its gray value in three channels is 178, 105, and 88. The mean of three channels R, G and B is 156, 150, and 82. Then, the value of the first element of CDM is as follows:

$$d_1 = \frac{\sqrt{(178 - 156)^2 + (105 - 150)^2 + (88 - 82)^2}}{\sqrt{178^2 + 105^2 + 88^2} + \sqrt{156^2 + 150^2 + 82^2}}$$

$$= 0.111. \quad (4)$$

The CDM can embody the relation among three channels and effectively represent the variance of pixels change, thus improving the effectiveness of the descriptor. Most of study
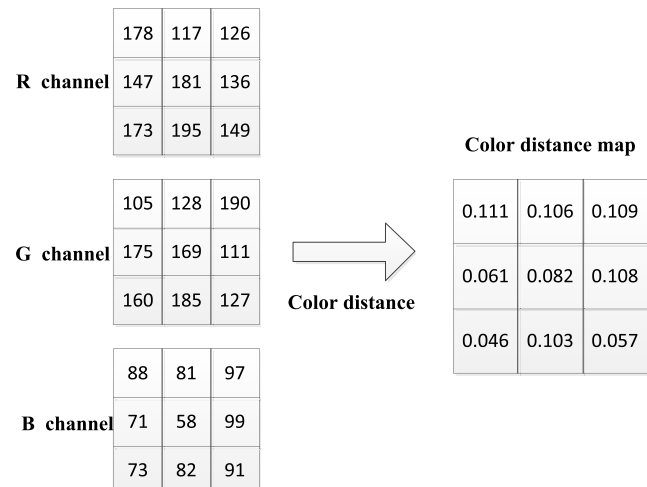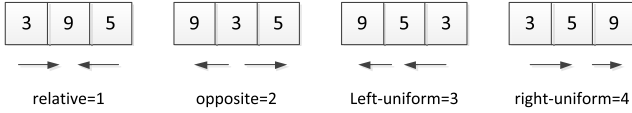

Fig. 6 Illustration of the construction of CDM.

Fig. 7 Definition of four local edge models.

proved that the directional edge information is very important in the image features. So, in order to introduce the edge information into the descriptor and exploit the intrinsic relation of image pixels as much as possible, after constructing the CDM, we define four local edge models according to the different trend between one element and its neighbors in CDM, which can embody the change trends of image pixels. The four local edge models are relative, opposite, left-uniform, and right-uniform. The relative model is denoted as 1 when the two neighbor elements are both smaller than the center element. The opposite model is denoted as 2 when the two neighbor elements are both greater than the center element. In the same way, if the values of the neighbor elements and center element are from large to small, the change trend is defined as left-uniform model and denoted as 3. If the values are from small to large, the change trend is defined as right-uniform model and denoted as 4. Figure 7 shows the four edge models of the local structure for specified region of an original image.

Four edge models effectively reveal the change trend of pixels and edge orientations. It can reflect the local information and make the variance of the different blocks. Based on the definition, LCDQP string in 0 deg, 45 deg, 90 deg, and 135 deg directions is given as

$$\text{LCDQP}[p(e_c)]|_\alpha = \{D_\alpha(e_c); D_\alpha(e_1); D_\alpha(e_2); \ldots, D_\alpha(e_8)\},$$
(5)

where α is the angle, $\alpha = 0$ deg, 45 deg, 90 deg, 135 deg, e_c is the center element of CDM, and $e_i(i = 1, 2, 3\ldots 8)$ is the neighbor elements of e_c. $p(e_c)$ is the value of the element of "e_c" in CDM. Similar to LTP, the function of $D_\alpha(e_c)$ is defined as follows:

$$D_\alpha(e_c) = F[p'(e_i), p'(e_{i+4})];$$

$$i = \left(1 + \frac{\alpha}{45}\right) \quad \forall \, \alpha = 0 \text{ deg}, 45 \text{ deg}, 90 \text{ deg}, 135 \text{ deg},$$
(6)

$$F[p'(e_i), p'(e_{i+4})] = \begin{cases} 1 \, p'(e_i) \geq 0 \text{ and } p'(e_{i+4}) \geq 0 \\ 2 \, p'(e_i) < 0 \text{ and } p'(e_{i+4}) < 0 \\ 3 \, p'(e_i) < 0 \text{ and } p'(e_{i+4}) \geq 0 \\ 4 \, p'(e_i) \geq 0 \text{ and } p'(e_{i+4}) < 0 \end{cases},$$
(7)

$$p'(e_i) = p(e_c) - p(e_i); \quad i = 1, 2, \ldots, 8.$$
(8)

We can get an LCDQP string for each direction including four local edge models. For these strings, in order to extract the spatial distribution of edge models, we adopt a new code method that chooses one of edge models to construct subfeatures separately. Then, four binary series are obtained. Based on this, we calculate the value of binary series by multiplying weights $2^i(i = 0, 1, \ldots 8)$. Then one direction feature is gained by four numbers. An example of the extraction process of LCDQP for a local block in 0 deg, 45 deg, 90 deg, and 135 deg directions is shown as Fig. 8. Figure 9 shows an example of original image (as shown in Fig. 1) and its LCDQP images. From Figs. 8 and 9, it is observed that the proposed descriptor can yield more edge direction and spatial information and thus can achieve better performance.
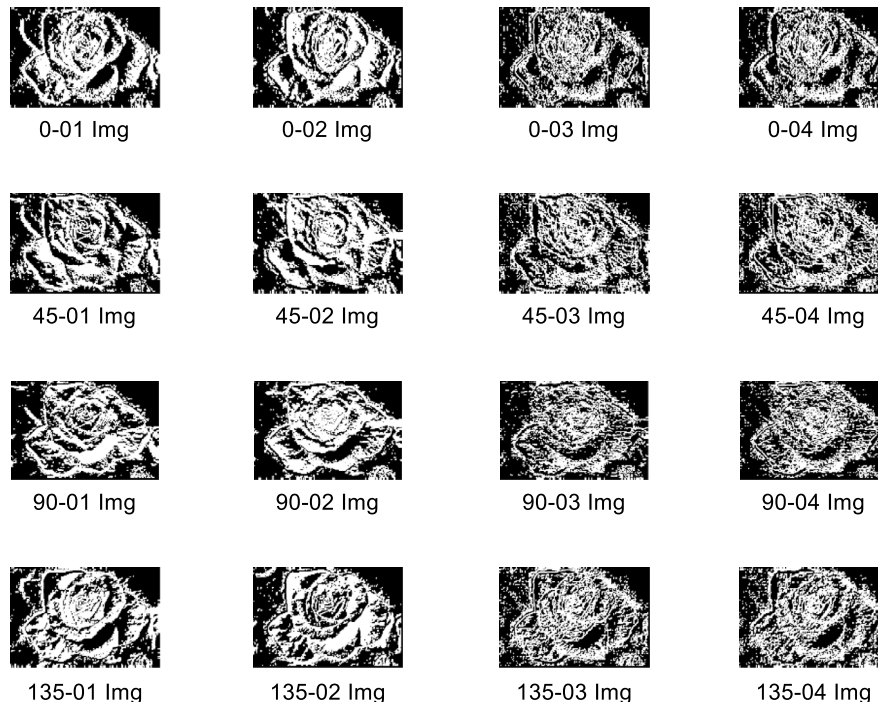
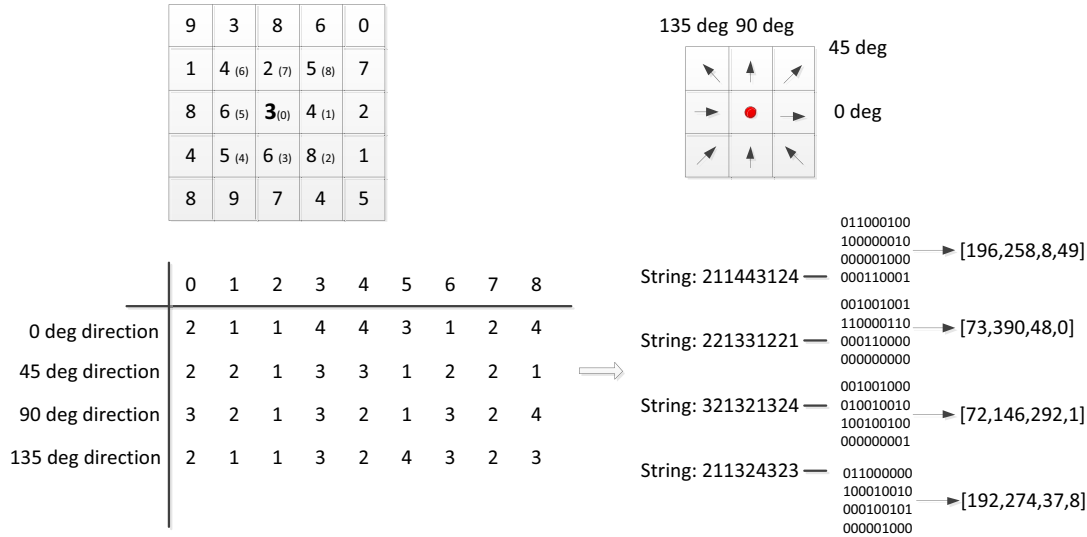**Fig. 8** Example of original image and its LCDQP images.

Fig. 9 Example of construction of LCDQP descriptor for a local block.

In this way, the given image is converted to LCDQP images with values ranging from 0 to 511. Finally, the whole image is represented by concatenating the four directions of feature vector into a histogram, which is supported by Eq. (9):

$$H_{\text{LCDQP}}(l)$$
$$= \{H_{\text{LCDQP}|\alpha}(l), \alpha\epsilon(0 \deg, 45 \deg, 90 \deg, 135 \deg)\}. \tag{9}$$

Suppose the size of input image is $M \times N$, and

$$H_{\text{LCDQP}|\alpha}(l) = \sum_{j=1}^{M}\sum_{k=1}^{N} f(\text{LCDQP}(j,k)|\alpha, l); l \in [0, 511], \tag{10}$$

$$f(x, y) = \begin{cases} 1 & x = y \\ 0 & \text{else} \end{cases}. \tag{11}$$

$H_{\text{LCDQP}|\alpha}$ is a histogram of the image represented by LCDQP. The maximum value of LCDQP is 511 and the dimensionality of H_{LCDQP} is $4 \times 4 \times 512$ bins.

The LCDQP descriptor is proposed on the basis of the CDM; therefore, the color information is combined effectively. What is more, the directional edge information is extracted by the proposed four edge models. The spatial structure can be taken into account because of the spatial code method for the LCDQP string. Therefore, LCDQP captures more spatial information as compared with other LBPs. The flowchart of the proposed method is shown in Algorithm 1.

4 Experiments

In this section, we describe various experiments to evaluate the performance of LCDQP against several state-of-the-art LBP-based methods in image retrieval and image classification. The abbreviations of these methods are given below.

- LCDQP: the proposed method.
- LBP:[1] local binary pattern.
- LTP:[4] local ternary pattern.

Algorithm 1 The construction flowchart of the LCDQP descriptor.

Input: color image

Output: LCDQP descriptor

Step 1. Upload an input image

Step 2. CDM generation. In RGB color space, utilize the difference between the value of pixels and its mean of each channel to generate the CDM.

Step 3. Four edge models definition. According to the distribution of the elements in CDM, four models are defined.

Step 4. LCDQP strings introduction. In four directions of 0 deg, 45 deg, 90 deg, and 135 deg, the distribution of the four models is taken into account to gain LCDQP strings.

Step 5. LCDQP descriptor construction. A spatial code method is adopted to generate LCDQP descriptor to describe the edge, texture, and spatial information. Finally, concatenate four directions of subdescriptors into histogram to describe the color image.

- CaLBP:[16] color local binary pattern
- LCVBP:[18] local color vector binary pattern
- CLSP:[20] completed local similarity pattern
- MDLBP:[21] multichannel decoded local binary patterns.
- LDP:[25] local directional pattern.
- LOOP:[26] local optimal oriented pattern.
- LTrP:[27] local tetra patterns.
- NRLBP:[28] noise-resistant local binary pattern.

4.1 Image Databases

For comparison, we adopt four public image databases: Corel-1K, Corel-5K, Outex_TC_00010, and GeorgiaTech face image database. Corel-1K and Corel-5K are the subsets of Corel-10K image database. All of the images of Corel-

10K have been preclassified into different categories each of size 100 by domain professionals. Some researchers think that Corel database meets all the requirements to evaluate an image retrieval system because of its large size and heterogeneous content. Outex_TC_00010 is one of the test suites of Outex texture image database. GeorgiaTech face image database is a well-known face image database. The details of the four databases are discussed below. A few samples of the databases are given in Fig. 10.

- **Corel-1K:** The database contains 1000 color images, which are divided into 10 classes of 100 JPG images each. Each class contains images with resolution either 256×384 pixels or 384×256 pixels. The 10 classes are African people, beach, building, bus, dinosaur, elephant, flower, horse glacier, and food.

- **Corel-5K:** Corel-5K consists of the first 5000 images of Corel-10K dataset and there are 5000 images from diverse contents such as sunset, beach, flower, building, car, horses, mountains, fish, food, door, etc. Each category contains 100 images of size 192×128 or 128×192 in the JPEG format.

- **Outex_TC_00010:** Outex_TC_00010 contains 4320 color images including 24 classes of 180 BMP images each, both in the form of surface textures and natural scenes. The collection of surface textures exhibits well-defined variations to a given reference in terms of illumination, rotation, and spatial resolution. Reliable manual segmentation of natural scenes is included in the database.

- **GeorgiaTech:** The face image database contains images of 50 persons, which are taken in two or three sessions at different times. For each person in this database, 15 color images were captured with clutter backgrounds. The resolution of these images was 640×480, pixels and the size of face within these images was around 150×150.

4.2 Performance Evaluation

In the experiment, to appraise the performance of the proposed method, two measurements, namely, average precision (AP) and average recall (AR) are adopted from usual retrieval system. Let I_n be the quantity of relevant images retrieved from top N position, M be the total number of images in the database that are similar to query image, and n be the whole amount of images retrieved from the database:

$$AP(q) = \frac{1}{N} \sum_{n=1}^{N} P(n), \tag{12}$$

$$AR(q) = \frac{1}{N} \sum_{n=1}^{N} R(n). \tag{13}$$

Here, $P(n)$ and $R(n)$ are the precision and recall for retrieval top n images, $n = 1, 2, \ldots, N$:

$$P(n) = \frac{I_n}{N}, \tag{14}$$

$$R(n) = \frac{I_n}{M}. \tag{15}$$

4.3 Similarity Measure

The performance of a descriptor depends not only on effective features but also on strong similarity measures or distance metrics. There are several similarity measures. Different aspects on choosing distance measure have also been analyzed. One factor of choosing the certain distance metric is the excessive computation time. When numerous feature vectors are generated, some distance metrics absorb more computing resources than others. Second, it is crucial to select a distance metric that is consistent with the human mind of resemblance.

In the experiment, we choose three distance measures and perform an experiment to investigate which distance measure is better suited for the proposed descriptor with the help of AP and AR in the image database. The three distance measures are Euclidean distance, $D1$ distance, and Manhattan distance. The definitions are as below.

Let $Q = (q_1, q_2, \ldots, q_n)$ and $P = (p_1, p_2, \ldots, p_n)$ represent the LCDQP feature vectors of the query image and database image, respectively. The dimension of the feature is n.

(1) Euclidean distance: Query matching in image retrieval system exploits this metric owing to its effectiveness. It measures the distance value between two feature descriptors and it can be calculated as

$$D_{\text{Euc}}(P, Q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}. \tag{16}$$

(2) $D1$ distance: Amount of normalized color histogram variation between every two images is computed by $D1$ distance. $D1$ distance is as follows:

$$D_{D1}(P, Q) = \sum_{i=1}^{n} \left| \frac{p_i - q_i}{1 + p_i + q_i} \right|. \tag{17}$$

(3) Manhattan distance: This metric is also called the city block distance metric. This distance metric is computed by the sum of differences between two feature vectors and can be calculated as

$$D_{\text{Man}}(P, Q) = \sum_{i=1}^{n} |p_i - q_i|. \tag{18}$$

Tables 1 and 2 show the comparison of the three various distance metrics with number of images retrieved in term of AP on Corel-1K and Corel-5K. From Tables 1 and 2, we can see that the LCDQP using $D1$ distance measures outperforms other distance measures due to the fact that $D1$ measures a normalized color histogram, and it is a standard way of comparing images that is very fast, reliable, and easy to implement. This measure is resilient to scaling, shifting, and rotating or moving, etc. Euclidean and Manhattan are pixel-by-pixel comparison methods, and if the image is shifted and rotated slightly, they can lead to large difference. From the results of this experiment, it is drawn that $D1$ distance is better suited for LCDQP. So, in the later experiments, $D1$ distance is chosen for the similarity measure for LCDQP.
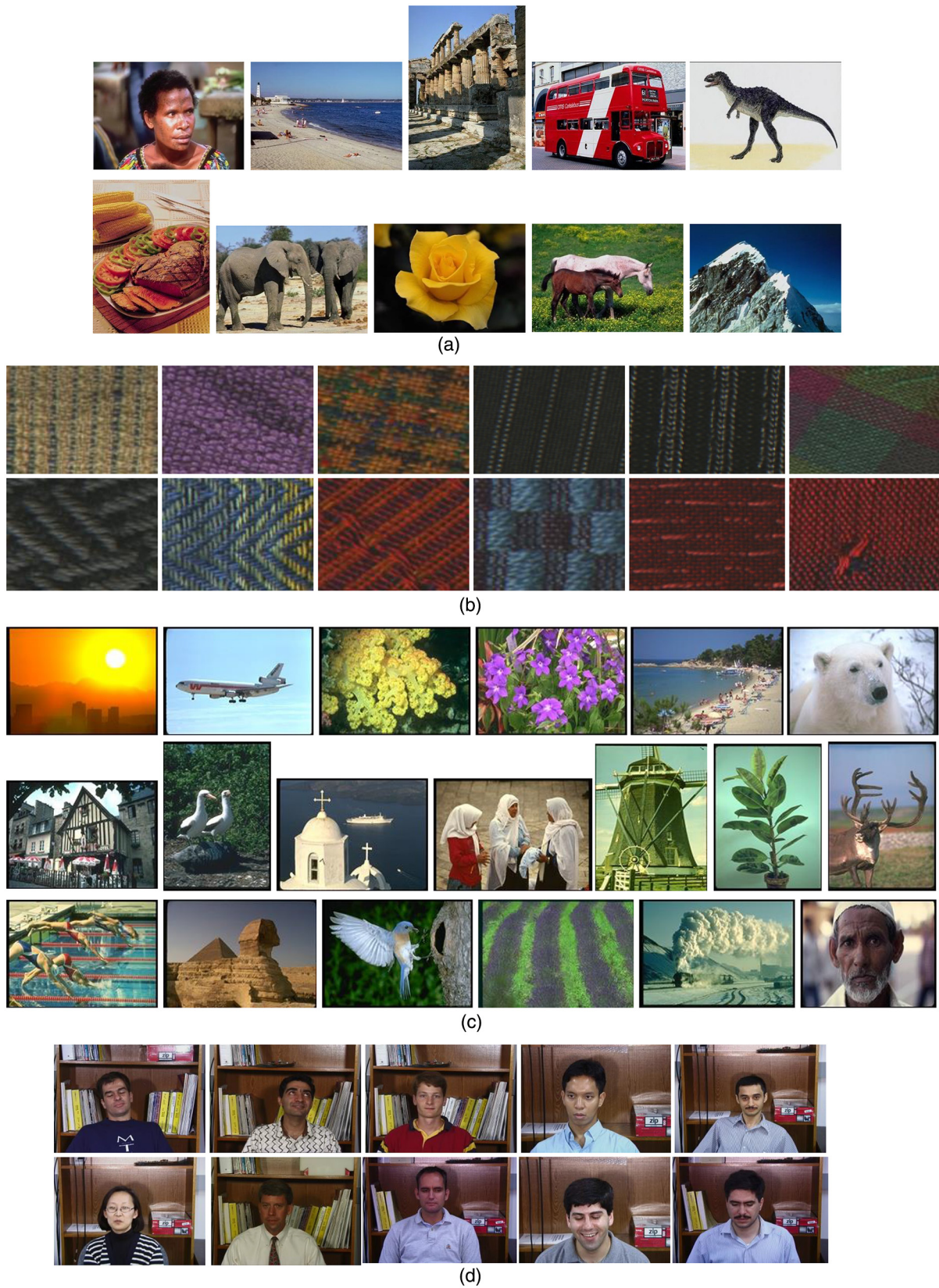
Fig. 10 Samples from four image databases: (a) Corel-1K, (b) Corel-5K, (c) Outex_TC_00010, and (d) GeorgiaTech.

Table 1 Results of AP (%) with different distance metrics on Corel-1K database for LCDQP.

No. of images retrieved	Euclidean	D1	Manhattan
10	84.12	97.65	94.68
20	88.96	95.37	93.96
30	91.40	95.32	94.49
40	91.72	94.80	94.16
50	91.13	94.01	93.71
60	90.07	92.90	91.99
70	89.22	90.48	90.18
80	88.29	89.29	89.02
90	86.11	87.97	87.53
100	85.41	86.37	86.21

Table 2 Results of AP (%) with different distance metrics on Corel-5K database for LCDQP.

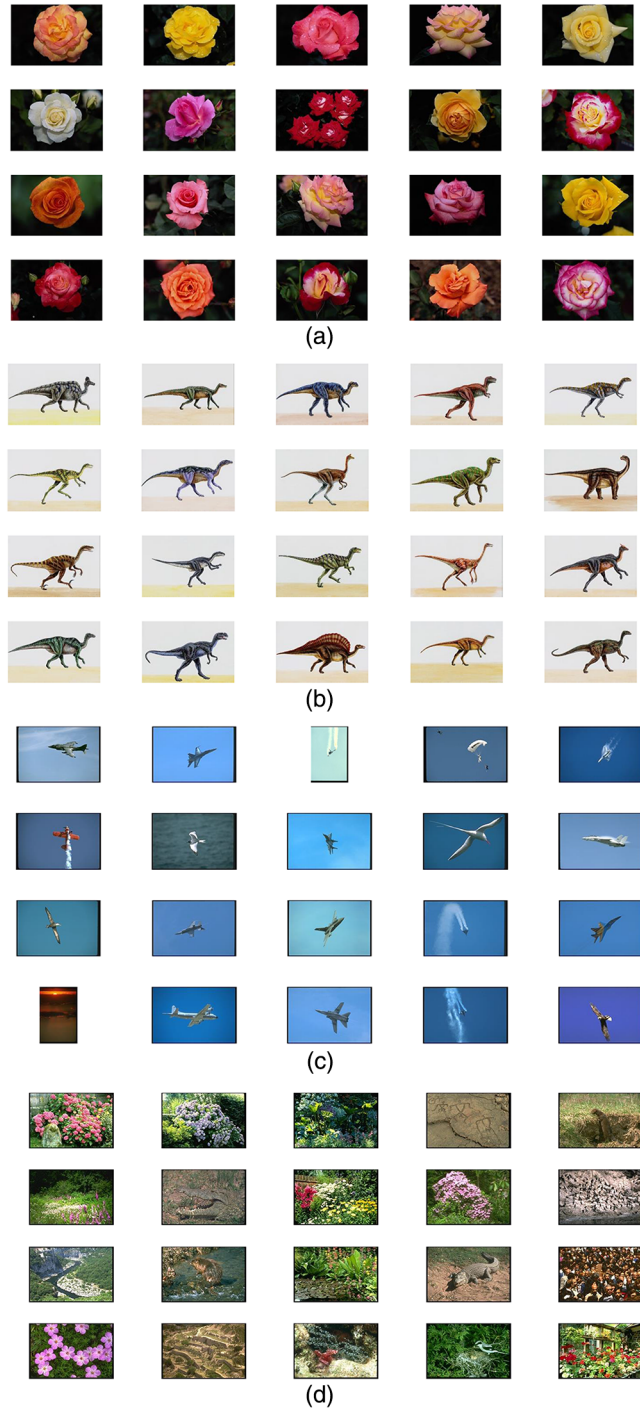| No. of images retrieved	Euclidean	$D1$	Manhattan
10	66.67	85.52	64.89
20	60.47	74.57	52.92
30	54.24	66.79	52.65
40	49.06	62.81	46.75
50	49.05	57.00	46.31
60	44.89	54.83	41.47
70	38.76	52.87	36.45
80	33.45	51.03	31.47
90	32.21	49.40	32.22
100	30.85	44.17	30.96

Fig. 11 First top 20 retrieval results images with top left as query image. (a), (b) Corel-1K; (c), (d) Corel-5K.

4.4 Experimental Results on Four Databases

4.4.1 Experiment 1

In this experiment, to evaluate the performance of LCDQP, we choose some images from Corel-1K and Corel-5K as query images, then retrieve the similar images. The first top 20 retrieval results are shown in Fig. 11. The first image is the query image and the rest of images are the retrieval results. From the retrieved images, it is observed that LCDQP can retrieve similar images effectively from the databases for the query images.

4.4.2 Experiment 2

In experiment 2, some images are chosen as query images from the Corel-1K. For each query image, the system collects n database images with the shortest image matching distance. If the retrieved image belongs to the same category as that of the query image, we can say that the system has appropriately identified the expected image, or else, the system has failed to find the expected image. LBP operator and LTP operator are extracted based on the gray image, which is transformed from the color image. LDP and LTrp are

Table 3 Results of different descriptors in terms of AP on Corel-1K database.

Category	AP ($n = 20$) (%)								
	LBP	LDP	LTP	LOOP	MDLBP	LTrP	NRLBP	OC-LBP	LCDQP
Africans	31.38	42.09	30.96	60.65	72.37	62.05	64.32	58.93	70.32
Beaches	63.79	60.49	62.63	59.52	56.35	51.62	57.67	63.19	63.71
Buildings	48.13	39.86	49.41	56.61	65.30	54.55	57.49	72.49	68.96
Buses	89.84	87.70	84.78	92.56	91.32	90.00	92.98	92.12	93.49
Dinosaurs	96.93	96.28	92.04	93.00	95.13	93.18	92.64	97.41	96.93
Elephants	49.95	58.35	50.65	53.31	61.06	53.44	53.07	64.22	62.24
Flowers	77.55	85.99	93.28	85.16	90.63	86.60	86.13	90.64	97.38
Horses	89.50	89.90	85.26	74.64	84.57	70.83	76.81	80.62	82.74
Mountains	43.63	36.35	40.95	43.97	47.33	44.45	41.70	48.23	51.19
Foods	29.31	30.61	34.75	56.85	69.49	61.96	59.80	62.46	60.81
Average	62.00	62.76	62.47	67.63	73.36	66.87	68.26	73.03	74.77

computed in second-order. Out of the two MDLBP operators, the decoder operator performs better than the adder operator. Hence, we consider the decoder operator for the performance comparison. For a fair comparison, the scale of (8,1) is used by all the compared methods. Therefore, the subscript "P, R" of relative methods is removed in all the expressions. The performance of all the descriptors is measured in terms of AP and AR in a particular class. Here, for AP, the first 20 retrieval results are considered for the performance evaluation so that the precision value of the retrieved results can vary from 0 to 20. For AR, the 100 retrieval results are considered. Tables 3 and 4 illustrate the performance of LCDQP and other existing methods on Corel-1K for 10 classes.

Table 4 Results of different descriptors in terms of AR on Corel-1K database.

Category	AR ($n = 100$) (%)								
	LBP	LDP	LTP	LOOP	MDLBP	LTrP	NRLBP	OC-LBP	LCDQP
Africans	13.87	27.13	17.24	31.67	36.51	34.23	35.89	28.4	36.3
Beaches	29.75	28.21	29.8	30.78	30.21	22.14	32.82	34.2	33.27
Buildings	20.69	16.41	18.53	23.06	33.70	23.59	25.65	27.6	28.99
Buses	62.59	55.54	57.48	61.19	55.58	58.21	70.16	70.4	63.18
Dinosaurs	91.21	88.79	78.5	81.61	75.89	77.89	74.03	87.7	78.53
Elephants	22.81	26.80	20.14	23.82	27.19	21.76	22.83	27.7	24.71
Flowers	43.42	60.80	65.04	59.77	51.51	59.11	49.13	58.9	80.93
Horses	51.19	52.37	44.66	36.33	39.98	31.81	36.48	37.4	45.69
Mountains	22.93	17.52	23.72	22.13	26.51	21.60	22.65	23.6	24.45
Foods	10.61	13.83	19.66	27.43	35.85	34.23	29.53	27.2	26.59
Average	36.91	38.74	37.48	39.78	41.29	38.46	39.92	42.31	44.26

From the results, it is clearly visible that the LCDQP takes averagely 74.77% in AP and 44.26% in AR on Corel-1K database, which is higher than the other descriptors. Furthermore, among 10 classes, the LCDQP has less pertinent retrieval results in mountains, foods, elephants, and buildings because most of the regions of these images are occupied by complex similar color. The flowers, dinosaurs, buses, and horses class images are significantly retrieved (∼97.38%, 96.93%, 93.49%, and 82.74% in AP) compared with other class images with the help of the color, texture, and edge information combined with the descriptor, thus achieving significant performance. The reason for the good results is that extrainformation is extracted by our descriptor. LCDQP uses the joint three channels to gain the color information and extracts four models according to the edge and direction information to distinguish more complex local structures of texture.

4.4.3 Experiment 3

In the meantime, for each image database, 60 images are chosen as query image to estimate the performance of the eight descriptors that belong to six categories and 10 images each categories. The precision-recall curve (P-R) of LCDQP and other existing methods on Core-1K, Corel-5K, and Outex_TC_00010 databases are given in Fig. 12, which provides the overall best results. We can see that the trend of retrieval performance of all methods is changed with respect to the sizes of the Corel-1K and Corel-5K databases. As the size of the database increases, there is a drop in the value of precision and recall for all methods. The relative drop is almost the same. It is evident from the graph that the proposed method achieves better significant results for all values of recall than other descriptors on three databases.

4.4.4 Experiment 4

To further estimate the recognition performance of LCDQP, we compared LCDQP with those state-of-the-art methods including CaLBP, LCVBP, and CLSP. The experimental data of these methods comes from Ref. 20. For a fair comparison, on the same GeorgiaTech image database, the recognition rate of LCDQP is achieved. Experimental results of these methods are illustrated in Table 5. From Table 5, we can see that recognition performance of LCDQP is higher than the three methods. Although the recognition rate of LCDQP is higher than other methods, the feature vector length is obviously longer than other descriptors, which may slow down the recognition speed, especially for very large databases. In our future work, we will consider how to reduce the dimensionality of feature vector and how to improve the computing efficiency.

5 Conclusion

Inspired by LBP and LTP, this paper proposes a descriptor-LCDQP for color image. The cross-channel co-occurrence information of three color channels is extracted by constructed CDM. The edge information is introduced by the definition of four edge models in CDM and its distribution in four directions. The effective spatial quaternionic code method makes the descriptor more compact. The patterns not only effectively describe color and textures but also take the spatial information into account. The experimental results on
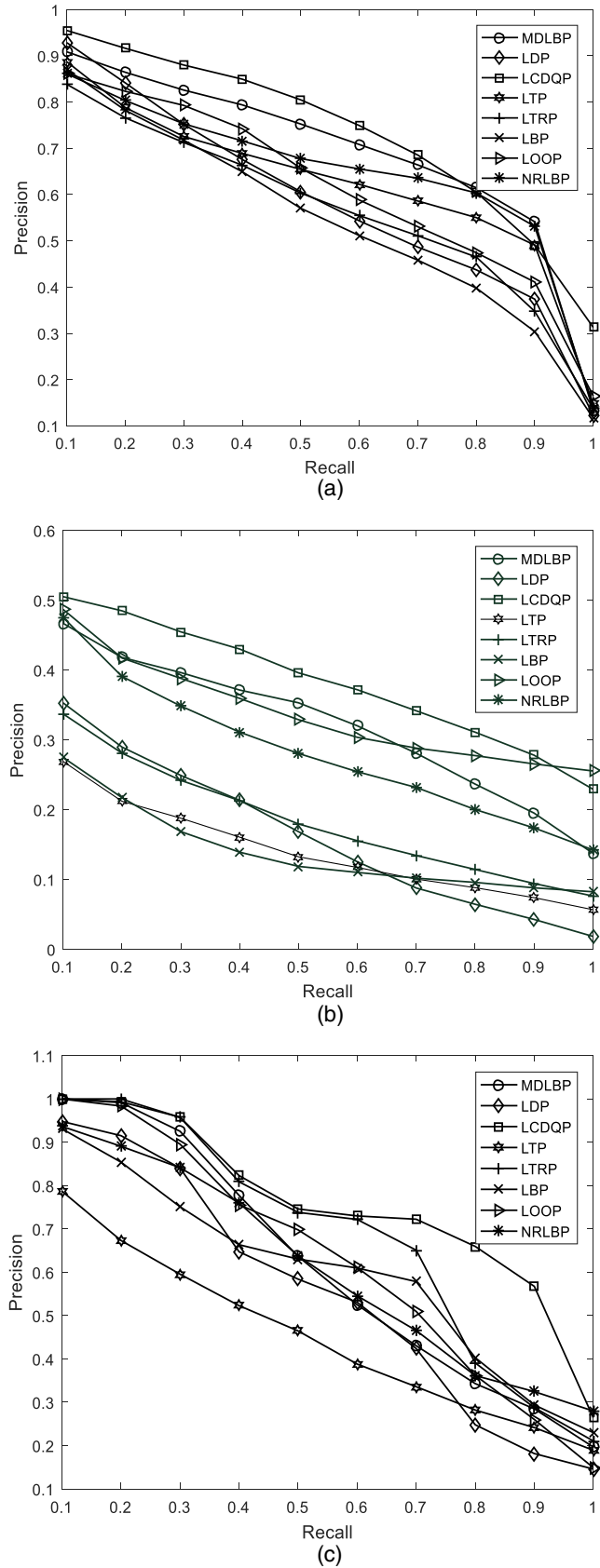

Fig. 12 P-R curves for the eight descriptors on three databases: (a) Corel-1K, (b) Corel-5K, and (c) Outex_TC_00010.

Table 5 Recognition rate (%) comparison with three methods.

Method	Recognition rate (%)
CaLBP	81.86
LCVBP	85.13
CLSP	89.68
LCDQP	94.32

the four databases validate that LCDQP can achieve higher performance than the state-of-the-art algorithms.

References

1. T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognit.* **29**(1), 51–59 (1996).
2. T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002).
3. M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognit.* **42**(3), 425–436 (2009).
4. T. Xiaoyang and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.* **19**(6), 1635–1650 (2010).
5. Z. Lu, X. Jiang, and A. Kot, "A novel LBP-based color descriptor for face recognition," in *Int. Conf. Acoust. Speech and Signal Process. (ICASSP2017)*, IEEE, pp. 1857–1861 (2017).
6. K. Wang et al., "Local binary circumferential and radial derivative pattern for texture classification," *Pattern Recognit.* **67**, 213–229 (2017).
7. M. H. Shakoor and R. Boostani, "Radial mean local binary pattern for noisy texture classification," *Multimedia Tools Appl.* **77**(16), 21481–21508 (2018).
8. Z. Dan et al., "An improved local binary pattern for texture classification," *Int. J. Light Electron Opt.* **125**(20), 6320–6324 (2014).
9. M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with center-symmetric local binary patterns," in *Indian Conf. Comput. Vision, Graph. Image Process. (ICVGIP2006)*, Springer, pp. 58–69 (2006).
10. Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Process.* **19**(6), 1657–1663 (2010).
11. S. Murala, R. Maheshwari, and R. Balasubramanian, "Directional local extrema patterns: a new descriptor for content based image retrieval," *Int. J. Multimedia Inf. Retrieval* **1**(3), 191–203 (2012).
12. D. Huang et al., "Local circular patterns for multi-modal facial gender and ethnicity classification," *Image Vision Comput.* **32**(12), 1181–1193 (2014).
13. B. Ryu et al., "Local directional ternary pattern for facial expression recognition," *IEEE Trans. Image Process.* **26**(12), 6006–6018 (2017).
14. S. Fadaei, R. Amirfattahi, and M. R. Ahmadzadeh, "Local derivative radial patterns: a new texture descriptor for content-based image retrieval," *Signal Process.* **137**, 274–286 (2017).
15. I. El Khadiri et al., "Local directional ternary pattern: a new texture descriptor for texture classification," *Comput. Vision Image Understanding* **169**, 14–27 (2018).
16. J. Y. Choi, K. N. Plataniotis, and Y. M. Ro, "Using color local binary pattern features for face recognition," in *Int. Conf. Image Process. (ICIP2010)*, IEEE, pp. 4541–4544 (2010).
17. S. H. Lee et al., "Local color vector binary pattern for face recognition," in *Int. Conf. Image Process. (ICIP2011)*, IEEE, pp. 2997–3000 (2011).
18. S. H. Lee et al., "Local color vector binary patterns from multichannel face images for face recognition," *IEEE Trans. Image Process.* **21**(4), 2347–2353 (2012).
19. R. Lan, Y. Zhou, and Y. Y. Tang, "Quaternionic local ranking binary pattern: a local descriptor of color images," *IEEE Trans. Image Process.* **25**(2), 566–579 (2015).
20. J. Li, N. Sang, and C. Gao, "Completed local similarity pattern for color image recognition," *Neurocomputing* **182**, 111–117 (2016).
21. S. R. Dubey, S. K. Singh, and R. K. Singh, "Multichannel decoded local binary patterns for content-based image retrieval," *IEEE Trans. Image Process.* **25**(9), 4018–4032 (2016).
22. C. Singh, E. Walia, and K. P. Kaur, "Color texture description with novel local binary patterns for effective image retrieval," *Pattern Recognit.* **76**, 50–68 (2018).
23. L. Yu et al., "Multi-trend binary code descriptor: a novel local texture feature descriptor for image retrieval," *Signal Image Video Process.* **12**(2), 247–254 (2018).
24. J. Yang, C. Liu, and L. Zhang, "Color space normalization: enhancing the discriminating power of color spaces for face recognition," *Pattern Recognit.* **43**(4), 1454–1466 (2010).
25. B. Zhang et al., "Local derivative pattern versus local binary pattern: face recognition with higher-order local pattern descriptor," *IEEE Trans. Image Process.* **19**(2), 533–544 (2010).
26. T. Chakraborti et al., "LOOP descriptor: local optimal oriented pattern," *IEEE Signal Process. Lett.* **25**(5), 635–639 (2018).
27. S. Murala, R. P. Maheshwari, and R. Balasubramanian, "Local tetra patterns: a new feature descriptor for content-based image retrieval," *IEEE Trans. Image Process.* **21**(5), 2874–2886 (2012).
28. J. Ren et al., "Noise-resistant local binary pattern with an embedded error-correction mechanism," *IEEE Trans. Image Process.* **22**(10), 4049–4060 (2013).

Shan Zhao is an assistant professor at Henan Polytechnic University. She received her BS and MS degrees in computer application from Jiaozuo Mining Institute and Henan Polytechnic University in 1996 and 2004, respectively, and her PhD in computer application from Xi'dian University in 2007. She is the author of more than 40 journal papers and two specialized books. Her current research interests include image processing and pattern recognition. She is a member of the China Computer Federation.

Jing Liu is an assistant professor at Henan Polytechnic University. She received her BS and MS degrees in computer application from Zhengzhou University of Technology and Henan Polytechnic University in 2000 and 2008, respectively. Her current research interest is image processing.