Research Article

Image Retrieval Based on the Combination of Region and Orientation Correlation Descriptors

Guangyi Xie, Zhe Huang, Baolong Guo , Yan Zheng, and Yunyi Yan

Research Institute of Intelligent Control & Image Engineering, Xidian University, Xi'an 710071, China

Correspondence should be addressed to Baolong Guo; blguo@xidian.edu.cn

Received 23 November 2019; Revised 2 February 2020; Accepted 13 May 2020; Published 10 June 2020

Academic Editor: Eduard Llobet

Copyright © 2020 Guangyi Xie et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A large number of growing digital images require retrieval effectively, but the trade-off between accuracy and speed is a tricky problem. This paperwork proposes a lightweight and efficient image retrieval approach by combining region and orientation correlation descriptors (CROCD). The region color correlation pattern and orientation color correlation pattern are extracted by the region descriptor and the orientation descriptor, respectively. The feature vector of the image is extracted from the two correlation patterns. The proposed algorithm has the advantages of statistic and texture description methods, and it can represent the spatial correlation of color and texture. The feature vector has only 80 dimensions for full color images specifically. Therefore, it is very efficient in image retrieving. The proposed algorithm is extensively tested on three datasets in terms of precision and recall. The experimental results demonstrate that the proposed algorithm outperforms other state-of-the-art algorithms.

1. Introduction

The rapid and massive growth of digital images requires effective retrieval methods, which motivates people to research and develop effective image storage, indexing, and retrieval technologies [1–4]. Image retrieval and indexing have been applied in many fields, such as the internet, media, advertising, art, architecture, education, medical, biological, and other industries. The text-based image retrieval process first manually labels the image with text and then uses keywords to retrieve the image. This method of retrieving an image based on the degree of character matching in the image description is time-consuming and subjective. The content-based image retrieval method overcomes the shortcomings of the text-based method, starting from the visual characteristics of the image (color, texture, shape, etc.) and finding similar images in the image library (search range). According to the working principle of general image retrieval, there are three keys to content-based image retrieval: selecting appropriate image features, adopting effective feature extraction methods, and accurate feature matching strategies.

Texture is an important and difficult-to-describe feature in images. Aerial, remote sensing pictures, fabric patterns, complex natural landscapes, and animals and plants all contain textures. Generally speaking, the local irregularity in the image and the macroscopic regularity are called textures, and the areas with repetitiveness, simple shapes, and consistent intensity are regarded as texture elements. After local binary pattern (LBP) [5], there are many similar methods proposed in recent years, i.e., local tridirectional patterns [6], local energy-oriented pattern [7], 3D local transform patterns [8], local structure cooccurrence pattern [9], local neighborhood difference pattern [10], etc.

Color histogram is the most commonly used and most basic method in color characteristics; however, it loses the correlation between pixel points. To solve this problem, many researchers have come up with their own visual models. Color correlogram [11] and color coherence vector (CCV) [12] characterize the color distributions of pixels and the spatial correlation between pair of colors. The gray cooccurrence matrix [13, 14] describes the cooccurrence relationship between the values of two pixels. Mehmood et al. present an image representation based on the weighted

average of triangular histograms (WATH) of visual words [15]. This approach adds the image spatial contents to the inverted index of the bag-of-visual words (BoVW) mode.

1.1. Related Works. Color, texture, and shape are prominent features of an image, but a single feature usually has some limitations. To overcome these problems, some researchers have proposed multifeature fusion methods, which utilize two or more features simultaneously. In [16], Pavithra et al. proposed an efficient framework for image retrieval using color, texture, and edge features. Fadaei et al. proposed a new content-based image retrieval (CBIR) scheme based on the optimised combination of the color and texture features to enhance the image retrieval precision [17]. Reta et al. put forward color uniformity descriptor (CUD) in the Lab color space [18]. Color difference histograms (CDH) count the perceptually uniform color difference between two points under different backgrounds with regard to colors and edge orientations in the Lab color space [19]. Taking advantage of multiregion-based diagonal texture structure descriptor for image retrieval is proposed in the HSV space [20]. In [21], Feng et al. proposed multifactor correlation (MFC) to describe the image, which includes structure element correlation (SEC), gradient value correlation (GVC), and gradient orientation correlation (GDC). Wang and Wang proposed SED [22], which integrates the advantages of both statistical and structural texture description methods, and it can represent the spatial correlation of color and texture. Singh et al. proposed BDIP+ BVLC+CH (BBC) [23], which is represented by a combination of texture feature block difference of inverse probabilities (BDIP) and block variation of local correlation coefficients (BVLC) and color histograms. In [24], the visual contents of the images have been extracted using block level discrete cosine transformation (DCT) and gray level cooccurrence matrix (GLCM) in RGB channel, respectively. It can be represented as DCT+ GLCM. In addition, local extrema cooccurrence pattern for color and texture image retrieval is proposed in [25].

According to the texton theory proposed by Julesz [26], many scholars have proposed texton-based algorithms. Texton cooccurrence matrix (TCM) [27], a combination of a trous wavelet transform (AWT) and Julesz's texton elements, is used to generate the texton image. Further, texton cooccurrence matrix is obtained from texton image which is used for feature extraction and retrieval of the images from natural image database. Multitexton histogram (MTH) integrates the advantages of cooccurrence matrix and histogram, and it has a good discrimination power of color, texture, and shape features [28]. Correlated primary visual texton histogram features (CPV-THF) is proposed for image retrieval [29]. Square Texton Histogram (STH) is derived based on the correlation between texture orientation and color information [30].

1.2. Main Contributions. Considering that color, texture, and uniformity features are of relevant importance in recognition of visual patterns [17–21], an algorithm proposed in this paper combines region and orientation correlation descrip-

tors (CROCD). This method entails two compact descriptors that characterize the image content by analyzing similar color regions and four orientation color edges in the image. It is based on the HSV color space since it is in better agreement with the visual assessments [20]. Contrasting with other approaches, CROCD features have the advantage of balancing operation speed and accuracy.

The rest of the paper is organized as follows. In Section 2, the overall introduction and workflow of the algorithm are presented. Section 3 explains the proposed algorithm in detail. Experimental results are obtained in Section (3). Finally, the whole work is concluded in Section 4.

2. Region Correlation and Orientation Correlation Descriptors

There are different objects in an image. The same object is usually a certain area made up of the same or approximate color, which constitutes the texture of the internal area of the object. The edges of an object have distinct color differences from the surrounding ones, and the edges of every object are the same or similar in color. Based on the above analysis, this paper presents a method of combining region color correlation descriptor and orientation color correlation descriptor. This method is also an effective method of combining color, texture, and edges to retrieve images. Firstly, the color image is quantified and coded, and then, the region color correlation pattern is calculated by the region descriptor; after that, the region correlation vector is calculated. Secondly, the orientation color correlation pattern is obtained by the orientation descriptor, and the color correlation histogram of the four orientations is obtained by statistics of the correlation pattern. The orientation color correlation vector of the image is calculated. The feature vector of image is obtained by concatenating the two-color correlation vectors of region and orientation. Finally, use similarity distance measure for comparing the query feature vector and feature vectors of database and sort the distance measure, then produce the corresponding images of the best match vectors as final results. The workflow of the proposed algorithm is shown in Figure 1.

3. The Algorithm Process

3.1. Image Color Quantization. Common color spaces for images are RGB, HSV, and Lab. Among them, the HSV space is a uniform quantized space, which could mimic human color perception well; thus, many researchers use it for image processing [17, 20–22, 25]. The HSV color space is defined in terms of three components: hue (H), saturation (S), and value (V). H component describes the color type which ranges from 0 to 360. S component refers to the relative purity or how much the color is polluted with white color which ranges from 0 to 1. V component is used for the amount of black that is mixed with a hue or represents the brightness of the color. It also ranges 0–1.

Image color quantization is a common method in image processing, especially in image retrieval. Assuming that the same objects are detected, the color will be slightly different

FIGURE 1: Workflow of the proposed algorithm.

due to the influence of light, environment, and background. These effects can be eliminated by quantization with appropriate bins. On the other hand, quantization in image processing can also make the operation simple and reduce the operation time.

Therefore, giving a color image $I(x, y)$, the quantization is presented as follows [22]:

- (1) Nonuniformly quantize the H , S , and V channels into 8, 3, and 3 bins, respectively, as equations (1), (2), and (3):

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

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

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

- (2) Calculate the value of every point according to formula (4).

$$L = Q_s Q_v H + Q_v S + V, \quad (4)$$

where Q_s, Q_v are the quantization bins of color S and V , respectively. As mentioned above, both S and V are quantified into 3 bins, respectively, so

both values are 3. Substitute them into equation (4) to get the following formula:

$$L = 9H + 3S + V \quad (5)$$

- (3) Obtain the quantized color image. The quantized image is denoted by I_Q , and $I_Q(x, y) \in L_i$ as follows:

$$L_i = \{(x, y) \mid (x, y) \in I_Q, 0 \leq i \leq 71\} \quad (6)$$

This set of points will be used for color statistics of the region and orientation descriptor, respectively, and the dimension of the quantized image I_Q is denoted by bins.

3.2. Region Correlation Descriptor. The concept of texton element is proposed by Julesz [26]. Texton is an important concept in texture analysis. In general, textons are defined as a set of blobs or emergent patterns sharing a common property all over the image.

The features of an image have close relation to the distribution of textons. Different textons form different images. If the textons in the image are small and the color tone difference between adjacent textons is large, the image may have a smooth texture. If the texton is large and composed of multiple points, the image may have a rough texture. At the same time, a smooth or rough texture is also determined by proportion of textons. If the textons in the image are large and have only a few types, distinct shapes may be formed. In fact, textons can be simply expressed by region correlation descriptors in a way [19]. Five region correlation templates are presented here, as shown in Figure 2. The shaded portion of the 2×2 grid indicates that these values are the same.

The process of extracting the region color correlation pattern I_R is shown in Figure 3. Figure 3(a) is a schematic diagram of a descriptor. The template moves from top to bottom, left to right, in two steps throughout the image I_Q . When the values in the grayscale frame where the image and template coincide are the same, these pixels are the color correlation region. The other templates are used successively to obtain the result pattern of that template. The corresponding shaded parts of the five templates in the quantization pattern I_Q are retained, and the rest are left blank to obtain the

FIGURE 2: Region correlation templates.

regional color correlation pattern I_R , as shown in Figure 3(c). Calculate its histogram, constitute a quantization vector, and get the region color correlation vector $H(I_R)$.

3.3. Orientation Correlation Descriptor. The orientation templates are shown in Figure 4, which can be used to detect the lines with the same color in the orientations of horizontal, vertical, diagonal, and antidiagonal, respectively. In other words, the edge information of an image can be detected. Figure 5 shows the operation diagram of horizontal, vertical, diagonal, and antidiagonal descriptors from top to bottom. These templates move through the whole image I_Q from top to bottom, left to right, in two steps. When the values in the grayscale frame where the image and template coincide are the same, the two pixels are the color correlation pixels of the orientation. The corresponding shadow part of the four orientation template in quantization pattern I_Q is retained, and the rest part is left blank to obtain quantization pattern I_O , as shown in Figure 5(d). Then, the quantization histogram of each orientation is counted, and the color correlation vector of the orientation is calculated. For the sake of illustration, only three quantization elements are taken as examples in Figure 5. In practice, it is the quantized value of image (0, bin-1). The specific steps are as follows:

- (1) Construct a statistical matrix of 4x bins. Each row of the matrix represents the orientation of horizontal, vertical, diagonal, and antidiagonal, respectively, and the number of columns is the bins of quantization
- (2) In the orientation color correlation pattern I_O , if it meets one of the orientation descriptor conditions, add 1 to the corresponding quantization value in the matrix.

$\text{Num}_{ij} = \text{Num}_{ij} + 1$ ($i \in (1, 4)$ represents one of the orientation descriptors, and $j \in (0, \text{bin}-1)$ represents the value of quantization)

- (3) Calculate the mean and standard deviation of each orientation descriptor according to equations (7) and (8), then get an 8-dimensional vector $T(I_o)$ to represent the orientation correlation vector of the image, $T(\mu_1, \sigma_1; \mu_2, \sigma_2; \dots; \mu_4, \sigma_4)$.

$$\mu_i = \sum_{j=1}^{\text{Bin}} \text{Num}(i, j) \times j, \quad (7)$$

$$\sigma_i = \left[\sum_{j=1}^{\text{Bin}} (\text{Num}(i, j) \times j - \mu_i)^2 \right]^{1/2} \quad (8)$$

According to the above steps, the orientation correlation vectors obtained in Figure 5 are (3, 2.65, 3.33, 2.52, 2, 1, 3.67, 2.52).

3.4. Composition of Feature Vector. The objects may have the same texture, but the edge characteristics of the objects may be different. The two factors can complement each other to improve the retrieval accuracy. The region correlation descriptor represents the texture features of an object and mainly represents the texture features of some areas inside the object, and the features are 72 dimensions. The orientation correlation descriptor represents the edge characteristics of the object. Different objects usually have different edge distributions. By taking the respective averages and standard deviations of the colors in the four directions of the horizontal, vertical, diagonal, and diagonal edges, the average color value and color offset in the four edge directions can be expressed and the object edge features are only represented by 8-dimensional feature vectors, which can improve the retrieval efficiency. Therefore, the region correlation descriptor in these two operators works better, and the later experimental part also proves that.

In Section 4.4, the experiments demonstrated that quantizing the HSV color space into 72 color bins nonuniformly is well suitable for our proposed algorithm. Therefore, $H(I_R)$ can represent the histogram of the region correlation image obtained by the region correlation descriptor, leading to a 72 dimensional vector. $T(I_O)$ can represent the orientation correlation image obtained by the orientation correlation descriptor, leading to an 8-dimensional vector. Finally, the two vectors are concatenated into a vector to obtain an 80-dimensional vector representing the image. Figure 6 shows two images and their own feature vectors of CROCD.

4. Experimental Results

4.1. Experimental Dataset. For the purpose of experimentation and verification, experiments are conducted over the benchmark Corel-1K, Corel-5K, and Corel-10K datasets. (1) 1K dataset (as shown in Figure 7(a)), with a size of 384×256 (or 256×384), contains 10 categories of original residents, beaches, buildings, public buses, dinosaurs, elephants, flowers, horses, valleys, and food, with 100 images for each category, and a total of 1000 images. (2) 5K dataset (shown in Figure 7(b)), with a size of 187×126 (or 126×187), contains 50 categories of images, including lion, bear, vegetable, female, castle, and fireworks, with 100 images for each category, a total of 5,000 images. (3) 10K dataset (as shown in Figure 7(c)), with a size of 187×126 (or 126×187), contains 100 category images of flags, stamps, ships, motorcycles, sailboats, airplanes, and furniture and 100 images of each category, a total of 10,000 images. In this section, we evaluate the performance of our method by these Corel datasets.

4.2. Performance Evaluation Metrics. The performance of an image retrieval system is normally measured using precision P_T and recall P_R for retrieving top T images defined by formula (9) and (10), respectively, where n is the number of relevant images retrieved from top T positions and R is the total

FIGURE 3: The process of extracting regional color correlation pattern I_R : (a) operation diagram of a descriptor; (b) five components of I_R ; (c) the final I_R pattern.

FIGURE 4: Orientation template: (a) horizontal; (b) vertical; (c) diagonal; (d) antidiagonal.

number of images in the dataset that are similar to the query image. Precision is used to describe the accuracy of algorithm query. Recall is used to describe the comprehensiveness of algorithm query. The higher the precision and recall are, the better the function of the algorithm is. Precision and recall are the most extensive evaluation criteria for evaluating query algorithms.

$$P_T = \frac{n}{T}, \quad (9)$$

$$P_R = \frac{n}{R}. \quad (10)$$

In these experiments, we randomly selected 10 images from each category. In other words, 100, 500, and 1,000 images are selected randomly from three datasets, respectively, as query images to compare various results.

4.3. Similarity Measure. In the content-based image retrieval system, the retrieval precision and recall are not only related

to the extracted features but also related to the similarity measurement. So, choosing an appropriate measure for our algorithm is a key step. In this experiment, we compared several common similarity criteria, such as Euclidean, L1, weighted L1, Canberra, and χ^2 .

There are two feature vectors $x = (x_1, x_1, \dots, x_n)^T$ and $y = (y_1, y_1, \dots, y_n)^T$ extracted from images; their similarity measures can be expressed as

$$\text{L1 distance : } D(x, y) = \sum_i |x_i - y_i|,$$

$$\text{Euclidean distance : } D(x, y) = \sum_i (x_i - y_i)^2,$$

$$\chi^2 \text{ statistics : } D(x, y) = \sum_i \frac{(x_i - y_i)^2}{x_i + y_i}, \quad (11)$$

$$\text{Canberra distance : } D(x, y) = \sum_i \frac{|x_i - y_i|}{x_i + y_i},$$

$$\text{Weighted L1 distance : } D(x, y) = \sum_i \frac{|x_i - y_i|}{1 + x_i + y_i}.$$

Calculate the value according to the above formulas and sort it from smallest to largest. The smaller the value is, the more similar the two images are. Table 1 shows the comparison results of different distance measurement methods. The test dataset is Corel-1K, and the statistical precision and recall are taken, respectively, when the total returned images from 10 to 30. It can be seen that the commonly used Euclidean distance is not good, while weighted L1 is the best.

4.4. Retrieval Performance. Different color spaces and quantization methods are both used to evaluate the performance of

FIGURE 5: The process of extracting orientation color correlation pattern I_O and calculating histogram of orientation descriptor: (a) orientation descriptor template; (b) quantized color image I_Q ; (c) color correlation pattern in each orientation; (d) the final orientation color correlation pattern I_O ; (e) Histogram statistics of every orientation descriptor.

the proposed algorithm. Experimental results reveal why the HSV space and nonuniform quantization are chosen.

The average precision and recall of HSV, RGB, and Lab are shown in Table 2. Images returned in the experiment range from 10 to 30. When color quantization is increased from 45 to 225 dimensions in the Lab color space, the precision and recall of the proposed method are both increased on the whole. There are the same in two other color spaces. On the other hand, the more quantization will increase the noise; thus, the precision and recall of the proposed method are both decreased when the quantization is 225 in the Lab color space. The highest precision of the top-10 image retrieval results is 79.2% and 71.5% in the RGB and Lab spaces, respectively. The best results are seen in the HSV space, which range from 78.7% to 83.2%. The precision of uniform quantization is not more than 81%; thus, we chose the HSV space of 72-dimensional quantization nonuniformly.

In order to test our proposed algorithm, we compared the algorithms proposed by CDH [19], SED [22], BBC [23], DCT + GLCM [24], TCM [27], and MTH [28] on Corel-1K and

compared the retrieval precision and recall of 10 categories when the top retrieval image is 15, as shown in Table 3. Five of the ten classes in the proposed method are the best, and its average precision and recall are obviously higher than other algorithms.

In addition, the average precision and recall curve of the algorithm and other algorithms on Corel-1K dataset is shown in Figure 8. According to the results, the average precision of the proposed algorithm has been significantly improved from DCT+ GLCM, CDH, TCM, BBC, SED, and MTH up to 11.6%, 9.74%, 7%, 5.54%, 5.27%, and 4.27%, respectively, when the top retrieval image is 15. Moreover, the area enclosed by the P-R curve of the proposed algorithm is the largest. Therefore, the precision and recall of the proposed algorithm are higher than the other six algorithms. Based on these analyses, this method has better robustness.

To illustrate the universality of the algorithm, the precision and recall of the algorithm and other algorithms on Corel-5K and Corel-10K dataset are shown in Tables 4 and 5, respectively. When tested on Corel-5K and Corel-10K

FIGURE 6: Two examples of CROCD feature vectors: (a) branches; (b) dinosaurs.

datasets, the $\bar{P}_{10}$ of the proposed method is 60.2% and 50.02%, respectively, which are superior to the other six algorithms. To give an intuitive view, Figure 9 shows the P-R curves of the seven algorithms. It can also be seen from the figure that the algorithm proposed in this paper has the best effect.

The region correlation descriptor (RCD) and orientation correlation descriptor (OCD) in the CROCD algorithm make different contributions to the retrieval results. Retrieval results of region correlation vector, orientation correlation vector, and their combination (CROCD) are shown in Table 6 on the datasets Corel-1K, Corel-5K, and Corel-10K when the returned image is 15. In the dataset Corel-1K, the precision of RCD and OCD is 71.42% and 38.54%, respectively. The combination of the two, that is, CROCD is 78.07%, and the precision is increased by 6.65%. In the datasets Corel-5K and Corel-10K, the precision of CROCD increased by 5.49% and 5.43%, respectively, compared with the bigger one between RCD and OCD. So, in both the region correlation vector and the orientation correlation vector, the region correlation vector makes a major contribution to the final retrieval result. The results of orientation correlation vector alone are not very good, but after combining with region correlation vector, the proposed algorithm is better than other state-of-the-art retrieval methods. For an intuitive display, the contents of Table 6 are shown in Figure 10.

Figure 11 shows four images retrieved by CROCD from dataset Corel-10K and lists the first 30 returned images according to their similarity to the query images. The first 30 images returned from the tree branch (Figure 11(a)) and

dinosaur (Figure 11(b)) images are related to the query images, respectively. And, of course, not all query images of these two categories have such effect, but it can be shown that the proposed algorithm has the superiority to those objects which have the obvious color and texture in the similar background. Of the 30 returned images in the snow mountain category (Figure 11(c)), 27 were returned correct. Those incorrect images (enclosed by the rectangular box), the three billow images, have similar colors and textures as snow mountains. Machinery category (Figure 11(d)) also has the 27 returned correct. In the three images returned by the error (enclosed by the rectangular box), they have similar textures and colors to the query image.

4.5. Computational Complexity. The complexity of the proposed algorithm consists of the amount of calculations required to complete a retrieval which is divided into three parts: query image and database image feature extraction, similarity measurement, and ranking retrieval.

As for feature extraction, the calculation amount of extracting the correlation features of the region is $K \times 17 M \times N$, and the calculation amount of extracting the correlation features of orientation correlation is $K \times (5 M \times N + 16 L + 8)$, and the total is $K \times (22 M \times N + 16 L + 8)$, which is $K \times [O(MN) + O(L)]$, where M and N are the length and width of the image. L is the dimensions of the image color quantization space. The variable K represents the total number of images in the dataset.

As for similarity measurement, the weighted 1 criterion is adopted, and the calculation amount is $K \times (4D - 1)$, that is,

Corel-1K dataset

(a)

Corel-5K dataset

(b)

Corel-10K dataset

(c)

FIGURE 7: Some images in the three Corel datasets.

TABLE 1: Retrieval results of CROCD with different distance measures for Corel-1K.

Similarity measure	Precision (%)					Recall (%)				
	10	15	20	25	30	10	15	20	25	30
WeightL1	83.20	78.07	74.60	71.96	69.80	8.32	11.71	14.92	17.99	20.94
L1	75.80	71.93	69	66.24	64.10	7.58	10.79	13.80	16.56	19.23
Euclidean	68.40	64.60	61.70	59.32	57.70	6.84	9.69	12.34	14.83	17.31
Canberra	80.90	76.07	72.15	69	66.54	8.09	11.41	14.43	17.25	19.96
χ^2	77.50	73.60	70.85	68.80	66.67	7.75	11.04	14.17	17.20	20

The best retrieval results are shown in bold, which means that CROCD has the best performance on this condition.

TABLE 2: Retrieval results of CROCD in different color spaces for Corel-1K dataset.

Color space		Precision (%)					Recall (%)				
		10	15	20	25	30	10	15	20	25	30
HSV	72 (nonuniform)	83.20	78.07	74.60	71.96	69.80	8.320	11.71	14.92	17.99	20.94
	72	80.60	77.60	73.75	70.56	68.77	8.060	11.64	14.75	17.64	20.63
	108	79.50	74.73	71.45	68.80	67.07	7.950	11.21	14.29	17.20	20.12
	128	81	75.80	72.90	70.40	68.03	8.100	11.37	14.58	17.60	20.41
	192	78.70	73.93	71.15	69.24	67.27	7.870	11.09	14.23	17.31	20.18
RGB	16	72.90	68.13	65.10	62.20	59.70	7.290	10.22	13.02	15.55	17.91
	32	79.20	74.74	71.55	68.48	66.43	7.920	11.21	14.31	17.12	19.93
	64	77.90	73.93	70.25	67.48	65.20	7.790	11.09	14.05	16.87	19.56
	128	78.50	74.47	70.60	67.92	65.57	7.850	11.17	14.12	16.98	19.67
Lab	45	64.50	60.80	57	54.60	51.67	6.450	9.120	11.40	13.65	15.50
	90	71.50	65.47	62.10	59.24	56.67	7.150	9.820	12.42	14.81	17
	180	69.70	65.33	61.90	59.64	57.13	6.970	9.800	12.38	14.91	17.14
	225	69.80	65.13	62.30	59.52	57.10	6.980	9.770	12.46	14.88	17.13

The best retrieval results are shown in bold, which means that CROCD has the best performance on this condition.

TABLE 3: Retrieval results with different methods for each category of Corel-1K.

Category	Precision (%)								Recall (%)							
	CDH	SED	BBC	DCT+	GLCM	MTH	TCM	CROCD	CDH	SED	BBC	DCT+	GLCM	MTH	TCM	CROCD
African	78.67	78.67	72.67	60	72.67	76	84.67	11.80	11.80	10.90	9	10.90	11.4	12.70		
Beach	46.67	33.33	42.67	46	40.67	58	44.67	7	5	6.4	6.900	6.100	8.7	6.700		
Building	68.67	78	63.33	48	76.67	57.33	81.33	10.30	11.70	9.5	7.200	11.50	8.6	12.20		
Bus	77.33	88.67	84	71.33	80.67	86.67	90.67	11.60	13.30	12.6	10.70	12.10	13	13.60		
Dinosaur	96	99.33	96	97.33	98	96	99.33	14.40	14.90	14.40	14.60	14.70	14.4	14.90		
Elephant	56	56.67	46.67	52	70	48	62	8.400	8.500	7	7.800	10.50	7.20	9.300		
Flower	90.67	97.33	98	94	84	88.67	94.67	13.60	14.60	14.70	14.10	12.60	13.30	14.20		
Horse	68.67	83.33	88	96	90	78	93.33	10.30	12.50	13.20	14.40	13.50	11.7	14		
Mountain	22.67	46	54.67	46	52	46.67	48	3.400	6.900	8.200	6.900	7.800	7.0	7.200		
Food	78	66.67	79.33	54	73.33	75.33	82	11.70	10	11.90	8.100	11	11.30	12.30		
Average	68.33	72.80	72.53	66.47	73.80	71.07	78.07	10.25	10.92	10.88	9.97	11.07	10.66	11.71		

The best retrieval results are shown in bold.

the order of $K \times O(D)$. The dimension of the feature vector is D .

As for sort and search, the quick sort method is used. The calculation amount for sorting and searching the relevant images from the dataset is $O(K \log_2 K) + O(\log_2 K)$ [24].

The total amount of calculation is

$$K \times [O(MN) + O(L) + O(D)] + O(K \log_2 K) + O(\log_2 K). \quad (12)$$

FIGURE 8: Performance comparison for Corel-1K.

TABLE 4: Retrieval results with different methods for Corel-5K.

Methods	Precision (%)					Recall (%)				
	10	15	20	25	30	10	15	20	25	30
MTH	54.26	48.85	45.33	42.66	40.37	5.43	7.33	9.07	10.66	12.11
TCM	54.71	49.34	46.04	43.35	40.98	5.47	7.40	9.21	10.84	12.29
CDH	52.52	47.76	44.48	42.16	40.13	5.25	7.16	8.9	10.54	12.04
SED	58.84	52.91	49.29	46.20	43.71	5.88	7.94	9.86	11.55	13.11
BCC	58.40	52.51	48.78	45.91	43.79	5.84	7.88	9.76	11.48	13.14
DCT+ GLCM	54.32	49.43	45.75	42.3	39.33	5.43	7.41	9.15	10.57	11.80
CROCD	60.20	54.60	50.75	47.83	45.23	6.02	8.19	10.15	11.96	13.57

The best retrieval results are shown in bold.

TABLE 5: Retrieval results with different methods for Corel-10K.

Methods	Precision (%)					Recall (%)				
	10	15	20	25	30	10	15	20	25	30
MTH	43.48	37.97	34.77	32.11	30.15	4.35	5.7	6.95	8.03	9.04
TCM	44.27	39.21	35.88	33.38	31.49	4.43	5.88	7.18	8.35	9.45
CDH	43.71	38.27	34.86	32.53	30.55	4.37	5.74	6.97	8.13	9.17
SED	49.08	42.89	39.09	36.2	33.93	4.91	6.43	7.82	9.05	10.18
BCC	47.05	41.35	37.80	35.09	33.13	4.71	6.20	7.56	8.77	9.94
DCT+ GLCM	45.20	39.99	36.39	33.12	30.68	4.52	6.00	7.28	8.28	9.20
CROCD	50.02	44.51	40.67	37.86	35.66	5.0	6.68	8.13	9.46	10.7

The best retrieval results are shown in bold.

The speed of extracting similar images to the query image depends on the feature vector length of the image. Lengthy feature vector takes more time in calculating the difference between query image and database images. The comparison

of feature vector of the proposed method with other methods has been given in Table 7 for speed evaluation. Also, feature extraction time for one image has been given in Table 7 for all methods including the proposed method. These

FIGURE 9: The average retrieval performance comparison for (a) Corel-5K; (b) Corel-10K.

TABLE 6: Comparison of the proposed RCD, OCD, and CROCD on different datasets.

Methods	Precision (%)			Recall (%)		
	Corel-1K	Corel-5K	Corel-10K	Corel-1K	Corel-5K	Corel-10K
OCD	38.54	19.04	14.89	5.78	2.86	2.23
RCD	71.42	49.11	39.08	10.71	7.37	5.86
CROCD	78.07	54.60	44.51	11.71	8.19	6.68

The best retrieval results are shown in bold.

FIGURE 10: Comparison of precision (%) and recall (%) of the proposed RCD, OCD, and CROCD on different datasets: (a) Corel-1K; (b) Corel-5K; (c) Corel-10K.

experiments are conducted on dataset Corel-10K with Matlab R2016b on a Windows 10 machine equipped with an Intel i7-9700 CPU 3.0 GHz and 16 Gb of RAM.

As demonstrated in the table, the proposed method is slightly slower than SED but faster than the other methods

for feature extraction. The feature vector length of the proposed method is slightly longer than the DCT+ GLCM but shorter than other methods. Moreover, the proposed method outperforms the other methods in terms of accuracy as mentioned in different datasets.

(a)

(b)

FIGURE 11: Continued.

FIGURE 11: Image retrieval results on different images: (a) 9327; (b) 6125; (c) 647; (d) 4828.

5. Conclusions

In this paper, the proposed method is an effective approach for color, texture, and edge image retrieval. Firstly, the color

image is quantized into 72 bins, and then, the color correlation pattern of the region is calculated using the region descriptor. The orientation color correlation pattern which reflects edges of objects in an image is obtained using the

TABLE 7: Feature vector length, feature extraction, and image retrieval time of different methods.

Method	Feature vector length	Feature extraction time (s)	Image retrieval time (s)	Total times (s)
MTH	$64 + 18 = 82$	0.1955	3.2475	3.443
TCM	$9 \times 10 = 90$	0.4728	3.2521	3.725
CDH	$90 + 18 = 108$	0.3330	3.2980	3.631
SED	$72 + 4 \times 72 = 360$	0.1419	3.3106	3.453
BCC	$16 + 16 + 64 = 96$	0.5949	3.2679	3.863
DCT+ GLCM	$(16 + 4) \times 3 = 60$	0.2635	3.2267	3.490
CROCD	$72 + 2 \times 4 = 80$	0.1627	3.2472	3.410

The best retrieval results are shown in bold.

orientation descriptor. Furthermore, the color correlation histogram of the four orientations is obtained by statistics of the correlation pattern, and then, the orientation color correlation vector is calculated. The characteristic vector of the image is obtained by combining two vectors of region and orientation. Finally, the similarity ranking of query images is obtained by similarity comparison. Experiments show that the proposed method has the advantages of balancing high speed and high precision compared with similar algorithms. It is often difficult to extract a single closed shape in a natural image, but there are many partial contours of objects in a natural image. If an efficient description method of partial contours can extract feature information and integrate it with our proposed method, it should be able to improve retrieval effect. The next step is to combine the features of color, texture, and shape for retrieval, so as to further improve the retrieval effect. Besides, voting-based scoring, ranking on manifold [31] or other ranking methods [32] for image retrieval will be used instead of distance based merely for measurement criteria.

Data Availability

Data are available on request. Please contact Guangyi Xie to request the data.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

All the authors contributed to this study. G.X. performed the conceptualization, writing of the original draft, and editing; Z.H. did the investigation and designed the network and experiments; Y.Z. analyzed the data and investigation; B.G. and Y.Y. contributed to funding acquisition, project administration, and instruction.

Acknowledgments

This research is supported financially by the National Natural Science Foundation of China (Grant Nos. 61571346 and 61671357).

References

- [1] W. Zhou, H. Li, J. Sun, and Q. Tian, "Collaborative index embedding for image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 5, pp. 1154–1166, 2018.
- [2] C. Iakovidou, N. Anagnostopoulos, M. Lux, K. Christodoulou, Y. Boutalis, and S. A. Chatzichristofis, "Composite description based on salient contours and color information for cbir tasks," *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 3115–3129, 2019.
- [3] Z. Shabbir, A. Irtaza, A. Javed, and M. T. Mahmood, "Tetragonal local octa-pattern (T-LOP) based image retrieval using genetically optimized support vector machines," *Multimedia Tools and Applications*, vol. 78, no. 16, pp. 23617–23638, 2019.
- [4] Y. Zheng, B. Guo, Y. Yan, and W. He, "O2O method for fast 2D shape retrieval," *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 1–5378, 2019.
- [5] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [6] M. Verma and B. Raman, "Local tri-directional patterns: a new texture feature descriptor for image retrieval," *Digital Signal Processing*, vol. 51, pp. 62–72, 2016.
- [7] G. M. Galshetwar, L. M. Waghmare, A. B. Gonde, and S. Murala, "Local energy oriented pattern for image indexing and retrieval," *Journal of Visual Communication and Image Representation*, vol. 64, article 102615, 2019.
- [8] A. B. Gonde, S. Murala, S. K. Vipparthi, R. Maheshwari, and R. Balasubramanian, "3D local transform patterns: a new feature descriptor for image retrieval," in *Proceedings of International Conference on Computer Vision and Image Processing, CVIP 2016*, pp. 495–507, Roorkee, India, February 2016.
- [9] K. Zhang, F. Zhang, J. Lu, Y. Lu, J. Kong, and M. Zhang, "Local structure co-occurrence pattern for image retrieval," *Journal of Electronic Imaging*, vol. 25, no. 2, article 023030, 2016.
- [10] M. Verma and B. Raman, "Local neighborhood difference pattern: a new feature descriptor for natural and texture image retrieval," *Multimedia Tools and Applications*, vol. 77, no. 10, pp. 11843–11866, 2018.
- [11] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms," in *Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 762–768, San Juan, PR, USA, February 1997.

- [12] G. Pass, R. Zabih, and J. Miller, "Comparing images using color coherence vectors," in *Proceedings of the 1996 4th ACM International Multimedia Conference*, pp. 65–73, Boston, MA, USA, November 1996.
- [13] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 6, pp. 610–621, 1973.
- [14] D. Srivastava, B. Rajitha, S. Agarwal, and S. Singh, "Pattern-based image retrieval using GLCM," *Neural Computing and Applications*, pp. 1–14, 2018.
- [15] Z. Mehmood, T. Mahmood, and M. A. Javid, "Content-based image retrieval and semantic automatic image annotation based on the weighted average of triangular histograms using support vector machine," *Applied Intelligence*, vol. 48, no. 1, pp. 166–181, 2018.
- [16] L. K. Pavithra and T. S. Sharmila, "An efficient framework for image retrieval using color, texture and edge features," *Computers & Electrical Engineering*, vol. 70, pp. 580–593, 2018.
- [17] S. Fadaei, R. Amirfattahi, and M. R. Ahmadzadeh, "New content-based image retrieval system based on optimised integration of DCD, wavelet and curvelet features," *IET Image Processing*, vol. 11, no. 2, pp. 89–98, 2017.
- [18] C. Reta, J. A. Cantoral-Ceballos, I. Solis-Moreno, J. A. Gonzalez, R. Alvarez-Vargas, and N. Delgadillo-Checa, "Color uniformity descriptor: an efficient contextual color representation for image indexing and retrieval," *Journal of Visual Communication and Image Representation*, vol. 54, pp. 39–50, 2018.
- [19] G. H. Liu and J. Y. Yang, "Content-based image retrieval using color difference histogram," *Pattern Recognition*, vol. 46, no. 1, pp. 188–198, 2013.
- [20] W. Song, Y. Zhang, F. Liu et al., "Taking advantage of multi-regions-based diagonal texture structure descriptor for image retrieval," *Expert Systems with Applications*, vol. 96, pp. 347–357, 2018.
- [21] L. Feng, J. Wu, S. Liu, and H. Zhang, "Global correlation descriptor: a novel image representation for image retrieval," *Journal of Visual Communication and Image Representation*, vol. 33, pp. 104–114, 2015.
- [22] X. Wang and Z. Wang, "A novel method for image retrieval based on structure elements' descriptor," *Journal of Visual Communication and Image Representation*, vol. 24, no. 1, pp. 63–74, 2013.
- [23] C. Singh and K. Preet Kaur, "A fast and efficient image retrieval system based on color and texture features," *Journal of Visual Communication and Image Representation*, vol. 41, pp. 225–238, 2016.
- [24] N. Varish and A. K. Pal, "A novel image retrieval scheme using gray level co-occurrence matrix descriptors of discrete cosine transform based residual image," *Applied Intelligence*, vol. 48, no. 9, pp. 2930–2953, 2018.
- [25] M. Verma, B. Raman, and S. Murala, "Local extrema co-occurrence pattern for color and texture image retrieval," *Neurocomputing*, vol. 165, pp. 255–269, 2015.
- [26] B. Julesz, "Textons, the elements of texture perception, and their interactions," *Nature*, vol. 290, no. 5802, pp. 91–97, 1981.
- [27] A. B. Gonde, R. P. Maheshwari, and R. Balasubramanian, "Texton co-occurrence matrix: a new feature for image retrieval," in *Proceedings of the 2010 Annual IEEE India Conference: Green Energy, Computing and Communication*, pp. 1–5, Kolkata, 2010.
- [28] G. H. Liu, L. Zhang, Y. K. Hou, Z. Y. Li, and J. Y. Yang, "Image retrieval based on multi-texton histogram," *Pattern Recognition*, vol. 43, no. 7, pp. 2380–2389, 2010.
- [29] A. Raza, H. Dawood, H. Dawood, S. Shabbir, R. Mehboob, and A. Banjar, "Correlated primary visual texton histogram features for content base image retrieval," *IEEE Access*, vol. 6, pp. 46595–46616, 2018.
- [30] A. Raza, T. Nawaz, H. Dawood, and H. Dawood, "Square texton histogram features for image retrieval," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 2719–2746, 2019.
- [31] S. Liu, J. Wu, L. Feng et al., "Perceptual uniform descriptor and ranking on manifold for image retrieval," *Information Sciences*, vol. 424, pp. 235–249, 2018.
- [32] Z. Liu, S. Wang, L. Zheng, and Q. Tian, "Robust ImageGraph: rank-level feature fusion for image search," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3128–3141, 2017.