IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 21, 2020, accepted July 31, 2020, date of publication August 10, 2020, date of current version August 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3015285

Combination of Dominant Color Descriptor and Hu Moments in Consistent Zone for Content Based Image Retrieval

GUANGYI XIE, BAOLONG GUO, (Senior Member, IEEE), ZHE HUANG, YAN ZHENG, AND YUNYI YAN, (Member, IEEE)
Institute of Intelligent Control & Image Engineering, Xidian University, Xi'an 710071, China

Corresponding author: Baolong Guo (blguo@xidian.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61571346 and Grant 61671357.

ABSTRACT The rapidly increasing number of digital images requires effective retrieval. Meanwhile, the dominant color descriptor has been widely used in image processing. Due to the influence of lighting and other factors, the same color in nature may have some different changes. The human eye is usually more sensitive to zones of consistent color, often identifying objects by zones of consistency. Therefore, the proposed method in this paper first applies the texton template to detect and extract the consistent zone of an image, and calculates the dominant color descriptor feature on the pixels in this consistent zone. Besides, the translation and rotation invariance of the Hu moments feature is applied to extract the shape information in the same consistent zone of the image. Finally, the combination of the dominant dolor descriptor and the Hu moments is used for content-based image retrieval. The algorithm proposed in this paper is tested on three data sets: Corel-1k, Corel-5k and Corel-10k, and the experimental results show that it is superior to the current content-based image retrieval methods.

INDEX TERMS Image retrieval, dominant color descriptor, Hu moments, consistent zone.

I. INTRODUCTION

As mobile phones, video cameras, and other devices are capable of shooting and storing a large number of digital images, images on the Internet have shown a geometric progression. Managing, finding and retrieving images have become a difficult and urgent problem. Earlier years, manually managing and annotating images in large databases was time-consuming, labor-intensive, and error-prone. Thus, it is very necessary to build an efficient content-based image retrieval system [1]–[9].

People use high-level semantics to recognize images, while machines use low-level visual features. The gap between high-level semantics and low-level features is called the semantic gap. In order to close this gap, many models have recently been proposed to solve this problem. For example, the visual vocabulary based, SIFT feature-based, visual phrase-based, and deep feature-based image retrieval systems. Image retrieval based on visual words, using Sift, SURF, or HOG to extract features from the image of the data set, clustering to obtain many visual words. The method in paper [10] represents a modified curvelet transform (MCT) and its combination with vocabulary tree (VT) for feature collection and retrieval of the images from the database. The MCT uses the ridgelet transform as a component step and implements curvelet sub-bands using a filter bank of Gabor wavelet filters. Descriptor vectors (energy histogram vectors) of each image are indexed using vocabulary tree. Visual phrase model is proposed to overcome the high-dimensional and quantization errors of the bag of visual words (BovW) model. Based on the well-known BovW model, Ouni *et al.* propose three different methodologies [11] inspired by the visual phrase model effectiveness and a compression technique which ensures the same effectiveness for retrieval than the BoVW model. Due to the large-scale use of neural networks in the field of image classification and recognition, different types of deep neural networks are proposed for image retrieval. This paper [12] proposes a new bilinear CNN-based architecture using two parallel CNNs as feature extractors. The activations of convolutional layers are directly used to

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaogang Jin.

extract the image features at various image locations and scales. The network architecture is initialized by deep CNNs sufficiently pre-trained on a large generic image dataset then fine-tuned for the CBIR task. Kang *et al.* propose a hashing scheme [13] for fast SIFT feature-based image matching and retrieval. Firstly, a training process of the hashing function is done involving geometric and topological information. second, it is applied that a geometry enhanced similarity evaluation considers both global and details of an image in evaluating. However, the computational complexities of vocabulary-based and deep-learning-based methods are high due to clustering implantation and multi-layer processing.

The content-based image retrieval method handles the inherent characteristics of the image (color, shape, texture, etc.). Color features have been widely used in CBIR systems because they have the advantage of being insensitive to translation and rotation. For example, the main color descriptor (DCD) [14], Color correlogram [15] and color coherence vector (CCV) [16] are widely used in CBIR systems. The dominant color descriptor (DCD) is widely applied in the image retrieval taken as one of MPEG-7 color descriptors. But objects belonging to the same kind in nature may have different colors and similar shapes. For example, white horse and black horse, red peony and yellow peony, etc. The DCD only considers the color features of the image, and does not use the shape features of the object. Hu moments have translation, rotation, and scale invariance [17], and have been widely used in object shape recognition.

A new method is proposed in this paper which combines the advantages of color and shape for image retrieval using the features of both DCD and Hu moments (DCD-HM). However, the same color in nature may have some different changes under the sunlight. And the human eye is usually more sensitive to zones of consistent color, often identifying objects by zones of consistency. Therefore, an image is first quantified and then applied the texton template to detect and extract the consistent zone in this paper. Finally, the features of DCD and Hu moments (DCD-HM) are extracted in this consistent zone to obtain better retrieval results.

II. RELATED WORK

As an important feature extraction method, Hu moments are used in many aspects such as graphics, medical images and human activity recognition. Shu *et al.* presented details of a comparative analysis on several modifications of the original Hu moment invariants which are used to describe and retrieve two-dimensional (2D) shapes with a single closed contour [18]. Zhang *et al.* used wavelet entropy (WE) and Hu moment invariants (HMI) for feature extraction to detect pathological brains from normal brains obtained by magnetic resonance imaging (MRI) scanning [19]. An algorithm of image recognition techniques, including Hu invariant moment, texture features, lateral Fourier transform and Daubechies (DBn) wavelet transform, was used to describe the features of defects of sewer pipe [20]. In paper [21], an explicit formula was derived which uses the first two Hu moment invariants to compute a shape ellipticity measure, i.e. to evaluate how much a planar shape differs from an ellipse. The ellipticity measure computed by this formula is invariant concerning for translation, rotation and scaling transformations. Also, the highest possible value is obtained if and only if the shape considered is an ellipse. Zhang proposed Hu moment invariants (HMI) were extracted from a specific magnetic resonance (MR) brain image [22]. To improve the recognition rate, a human behavior recognition method based on weighted modified Hu moments will be put forward in the paper [23].

Min *et al.* proposed a dominant color descriptor (DCD) to improve the performance of image retrieval [24]. In paper [25] proposed a color-based descriptor which describes well image contents, integrating both global features provided by dominant color and local features provided by color correlogram for image retrieval. In paper [26], a new semantic feature extracted from dominant colors (weight for each DC) is proposed. A top-down descriptor was proposed called the Spatial Dominant Color Descriptor (SDCD) in paper [27]. In the extraction of dominant colors, a dynamic quantization by Gaussian Mixture Models (GMMs) was adopted. The number of dominant colors is determined automatically using the Bayesian Information Criterion (BIC). A novel image retrieval scheme in the DCT domain was presented in paper [28]. Firstly, dominant color blocks are selected from the image's DCT coefficients, then the quantized color histogram with a spatial relationship is constructed.

According to the texton theory proposed by Julesz [29], many scholars have proposed texton-based algorithms. Multi-texton histogram (MTH) integrates the advantages of co-occurrence matrix and histogram, and it has good discrimination power of color and shape features [30]. Correlated primary visual texton histogram features (CPV-THF) are proposed for image retrieval [31], CPV-THF integrates the visual content and semantic information of the image by finding consistents among the colour, texture orientation, intensity, and local spatial structure information of an image. Based on the texton theory, box-shaped structural elements are designed for image texture analysis in the paper. Square Texton Histogram (STH) [32] is derived based on the consistent between texture orientation and color information. According to Julesz's texton theory, 'Square Texton' templates are proposed for Image texture analysis. It was presented in color difference histograms (CDH) [33] that counts the perceptually uniform color difference between two points under different backgrounds concerning colors and edge orientations in Lab color space. The novel proposed approach [34] uses texture, color and shape features to make retrieval. Firstly, shape detection is done based on top-hat transform to detect and crop main object parts of the image. Secondly, use Color Local Binary Patterns (CLBP) and local variance as discriminant operators for texture features. Finally, the log likelihood ratio is used to retrieve mostly closing matching images for the query. The proposed

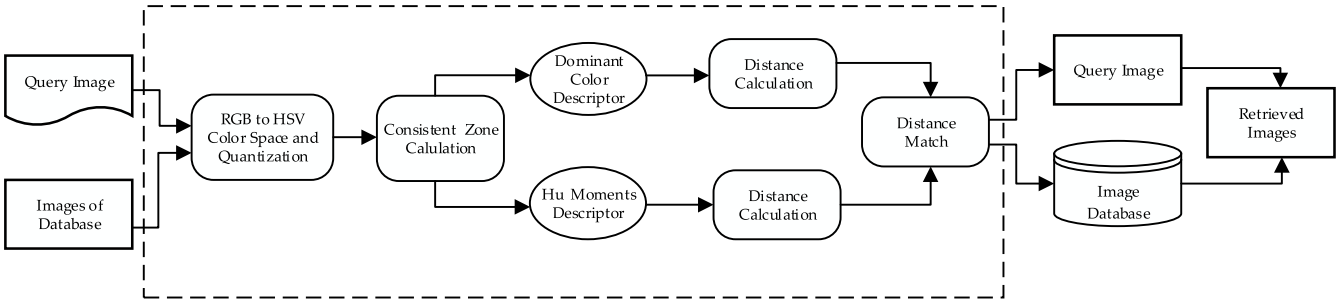


FIGURE 1. The workflow of the proposed algorithm.

CBIR techniques [35] are described and constructed based on RGB color with neutrosophic clustering algorithm and Canny edge method to extract shape features, YCbCr color with discrete wavelet transform and Canny edge histogram to extract color features, and gray-level co-occurrence matrix to extract texture features.

This paper presents a new method (DCD-HM) for image retrieval based on color and shape features. The main contributions are:

(1) It has been proposed that A new method combines dominant color descriptor and Hu moments for image retrieval.

(2) Extract the consistent zone of an image according to Julesz 's texton theory.

(3) The cross-interval method is used to compare image similarities and improve the accuracy of image retrieval.

The remainder of this paper is organized as follows. In section 3, the overall introduction and workflow of the algorithm are illustrated, and details are also described. Experimental results are shown in Section 4. Finally, the whole work is concluded in Section 5.

III. THE PROPOSED METHOD

The dominant color descriptor is widely used in the field of computer vision such as image retrieval. It quantizes the color space of an image into a limited number of intervals, and then calculates the average value and percentage of pixels that fall in each interval. Use these two features to represent the color features of the image. But the dominant color descriptor considers all pixels on the whole image, it is not targeted. According to the texton theory proposed by Julesz [29], an image is constructed according to a regular texton. Therefore, the consistent zone of the image is obtained through texton detection, and the feature extraction of the dominant color descriptor is performed on these pixels in the consistent zone. In addition, Hu moments not only have translation, rotation, and scale invariance, but also have the advantages of describing the shape of an object. The image blocking method is often used in image processing to improve the effectiveness of a method. The seven values of the Hu moments of each sub-block of an image, therefore, are obtained in the proposed method. Since the values of the first several moments play a large role in the image retrieval process, they are in granting greater weights in the distance comparison, which can also further improve the accuracy of retrieval.

The workflow of the proposed algorithm is shown in Figure 1. It can be seen from the figure that the image is changed from RGB space to HSV space firstly, and then the color space is quantized and encoded. Secondly, the consistent characteristics of the image are obtained using the texton template, and noise interference can be eliminated. Thirdly, calculate the vectors of the dominant color descriptor and Hu moments in consistent zone of the image, in order to obtain color and shape feature information respectively. Finally, obtain the feature distance between the query image and the library image through the DCD color information and the Hu moments shape information, and combine them to find an image result close to the query image.

A. IMAGE COLOR QUANTIZATION

The HSV space is a quantized space with uniform colors and can mimic human color perception, so it is used in many image processing methods [30]–[33]. Image color quantization is also a commonly used method in image retrieval. The color of the same object will be slightly different due to the influence of light and environment. These influences can be eliminated by dividing the appropriate quantization interval. In addition, quantization can also reduce the complexity of the operation, thereby reducing the operation time. If a color image I (x, y) is given, it is quantized according to the following rules [36]:

(1) The H, S and V channels are measured non-uniformly according to equations (1), (2) and (3) respectively:

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

FIGURE 2. Texton consistent templates.

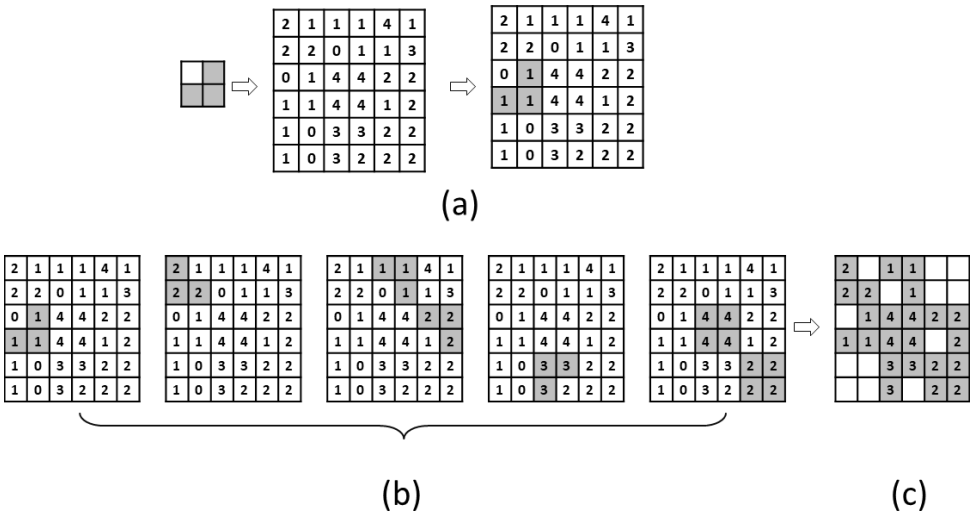


FIGURE 3. The process of extracting color consistent zone pattern I_Z: (a) Operation diagram of a texton template; (b) Five components of I_Z; (c) The final I_Z pattern.

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

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

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

$$\mathrm{L} = \mathrm{Q}_s\mathrm{Q}_v H + \mathrm{Q}_v S + V \tag{4}$$

where $\mathrm{Q}_s, \mathrm{Q}_v$ are the quantization bins of color S and V respectively. Both S and V are quantified into 3 bins respectively as mention above. Substitute them into equation (4) to get the following equation:

$$\mathrm{L} = 9H + 3S + V \tag{5}$$

Get the quantized image of the input image. Recorded as I_Q, and $\{(x, y)|I_Q \in [0, 71]\}$.

B. CONSISTENT ZONE DESCRIPTOR

According to the texton theory proposed by Julesz [29], due to the visual characteristics of the human eye, related structures in the image can be extracted with small regular patterns, which are texton. The "continuous theory" [31] pointed out that people can recognize patterns from small uniform areas. Therefore, using a small-size texton template can increase the probability of finding more regular patterns in the image. As shown in Figure 2, we show five 2×2 texton templates [30]. The gray part in each template indicates that these values are the same.

Figure 3 shows the process of extracting the color consistent zone of an image. Figure 3(a) is a schematic diagram of the first template to extract the color consistent zone map in the image. The template moves from top to bottom and left to right in two steps in the entire quantized image I_Q. In this process, when the values in the quantized image I_Q that fall in the template gray box are the same, the values of these pixels are retained. The corresponding color consistent zone of the template can be obtained. Then use other templates, in turn, to obtain the color consistent of their template respectively, as shown in Figure 3(b). Finally, the color consistent zone detected by each template is retained in the quantized color image I_Q to obtain the final color consistent zone map. As shown in Figure 3(c).

C. FEATURE EXTRACTION OF DOMINANT COLOR DESCRIPTOR

In dominant color descriptor (DCD), the color space, HSV, is divided into partitions known as course partitions. Each partition has two main components, including the partition center and percentage which are computed as [27]:

$$\mathrm{C}_i = (\mathrm{X}_i^H, \mathrm{X}_i^S, \mathrm{X}_i^V)\overline{X_i} = \frac{\sum_{X \in C_i} X}{\sum_{X \in C_i} 1} \tag{6}$$

$$p_i = \frac{\sum_{X \in C_i} 1}{\sum_{X \in C} 1} \tag{7}$$

According to equations (6) and (7), the feature vector representation of the DCD in consistent zone I_Z is obtained.

As shown in Figure 4, the square dots indicate the specific positions of the query image in different quantization intervals, and the dots indicate the position of an image to be detected in the database in each quantization interval. For ease of explanation, only four intervals are shown in the figure. It can be seen from the figure that if the conventional distance method of the same interval is used, the distance between the two images is the sum of the solid blue line segments shown in the figure. If the cross-interval calculation method is adopted, the sum of the red dotted line segments shown in the figure can be obtained, and the distance is significantly smaller than the sum of the solid blue line segments. The following distance measurement methods are proposed:

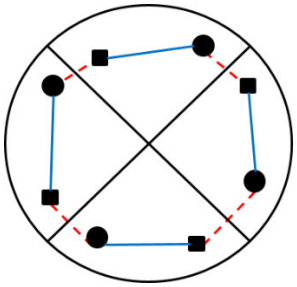


FIGURE 4. Schematic diagram of cross-interval distance measurement.

(1) Assuming that there are N quantized intervals, set $\mathrm{Sum}_{Color} = 0$, calculate the distance between the two intervals according to the following equation;

$$\mathrm{D}_{Color}(Q_k, \mathrm{I}_j) = \sqrt{\sum_{i=1}^{3} \left(\frac{p_{\mathrm{Q}_k}^2 + p_{\mathrm{I}_j}^2}{2} \right) \cdot \left(Z_{\mathrm{Q}_k}(i) - Z_{\mathrm{I}_j}(i) \right)^2} \quad (8)$$

where $Z_{\mathrm{Q}_k}(i)$ is i^{th} center of k^{th} interval in the DCD feature of query image Q, $i = [1, 2.3]$. p_{Q_k} is the ratio of the pixels in the interval k and the whole pixels in all intervals of the query image Q. $Z_{\mathrm{I}_j}(i)$ is i^{th} center of j^{th} interval in the DCD feature of the image in the database, $i = [1, 2.3]$. p_{I_j} is the ratio of the pixels in the interval j and the whole pixels in all intervals of the image to be inspected in the database. p_{Q_k}, p_{I_j} are added to the equation to allow the more pixels in the interval to have a greater effect on the calculation results.

(2) For the obtained N×N matrix, find the minimum value in the matrix, let $\mathrm{Sum}_{Color} = \mathrm{Sum}_{Color} + \min(\mathrm{D}_{Color}(Q_k, \mathrm{I}_j))$, and then eliminate the row and column where the minimum value lies.

(3) Repeat step (2) until the calculation stops when the matrix is empty.

Finally, the distance between query image Q and the image in the database I is $\mathrm{D}_{Color}(Q, \mathrm{I}) = \mathrm{Sum}_{Color}$.

D. FEATURE EXTRACTION OF HU MOMENTS

Color features can describe the distance of colors, but objects of the same color may not be the same type of objects. Besides color, applying the shape features of the consistent zone can improve the retrieval effect. Moments are invariant to rotation, translation, and scaling in characterizing image shapes. Moments are therefore widely used to describe the shape characteristics of objects. Hu constructs 7 moments with translation, rotation and scale invariance [17], but in discrete state only has invariance to rotation and translation, and does not have scale-invariance, and Hu moment has stronger sensitivity to contrast.

For digital images $I_{M\times N}$, the image size is $M \times N$. The gray level at the pixel (x, y) is $f(x, y)$, and the defined $p + q$ moment is:

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y) \quad (9)$$

The $p + q$ central moment of the image is defined as:

$$\delta_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (10)$$

where $\bar{x} = m_{10}/m_{00}, \bar{y} = m_{01}/m_{00}$ is the center of mass coordinates.

The normalized central moment of the image can be expressed as:

$$\eta_{pq} = \frac{\delta_{pq}}{\delta_{00}^{1+(p+q)/2}}, \quad p + q = 2, 3, \cdots \quad (11)$$

Hu applied the algebraic invariant theory to the above normalized central moment, and constructed the following 7 translation and rotation invariant moments:

$$\begin{cases} \phi_1 = \eta_{20} + \eta_{02} \\ \phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\ \qquad -3(\eta_{21} + \eta_{03})^2] \\ +(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \qquad +4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\ \qquad -3(\eta_{21} + \eta_{03})^2] \\ +(3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{cases} \quad (12)$$

Seven moments are obtained to describe the shape information.

The shape feature of the query image is denoted as $\mathrm{S}_{\mathrm{Q}_i} = (\mathrm{S}_{\mathrm{Q}_{i1}}, \mathrm{S}_{\mathrm{Q}_{i2}}, \cdots, \mathrm{S}_{\mathrm{Q}_{im}})$, where $i = 1, 2, \cdots, 7$. The shape feature of the image in the database is denoted as $\mathrm{S}_{\mathrm{I}_i} = (\mathrm{S}_{\mathrm{I}_{i1}}, \mathrm{S}_{\mathrm{I}_{i2}}, \cdots, \mathrm{S}_{\mathrm{I}_{im}})$, where $i = 1, 2, \cdots, 7$. The weighted distance between the two vectors is given by

$$\mathrm{D}_{\mathrm{shape}}(\mathrm{Q}, \mathrm{I}) = \sum_{i=1}^{7} \sum_{j=1}^{\mathrm{m}} \lambda_i \frac{\left| \mathrm{S}_{\mathrm{Q}_{ij}} - \mathrm{S}_{\mathrm{I}_{ij}} \right|}{\max(\mathrm{S}_{\mathrm{Q}_{ij}}, \mathrm{S}_{\mathrm{I}_{ij}})} \quad (13)$$

where $S_{Q_{ij}}$ is j^{th} feature of i^{th} Hu moment in the query image Q and m is feature vector length. There are 7 moments in the Hu moments, which satisfy $\sum_{i=1}^{7} \lambda_i = 1$.

E. DISTANCE MEASUREMENT

When the query image is Q and the image to be inspected in the image database is I, the feature weighted distance is used here for the two images. The distance between DCD color information features of the query image Q and the library image I is D_{Color}. Their Hu moments shape information feature distance is D_{shape}, and the final distance of the two images is:

$$D(Q, I) = \omega_C D_{Color}(Q, I) + \omega_S D_{shape}(Q, I) \quad (14)$$

Among them, ω_C and ω_S are adjustable weights and satisfy $\omega_C + \omega_S = 1$; The choice of ω_C and ω_S is determined according to the role of the main color information and shape information in the retrieval. For the query image and each image in the database, calculate the value $D(Q, I)$ according to the above equation, and these values sort from small to large. For distance measure the smaller the value, the more similar the two images.

F. ALGORITHM FLOW

(1) Input the image and change it from RGB space to HSV space. Quantization is performed then.

(2) Find the consistent zone in the image.

(3) By equations (6) and (7), it is obtained that the DCD vector represents the color information in consistent zone of the image. Then the distance between the query image and the image in the database is calculated according to the cross-interval distance measurement method mentioned in this paper.

(4) According to equations (9)-(12), it is calculated that the Hu moments vector describes the shape information in consistent zone of the image. The Then the distance between the query image and the image in the database is calculated according to the equation (13).

(5) The distances obtained in steps (3) and (4) are combined to obtain the final returned image relevant to the query image by sorting these distance values from small to large.

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL DATASET

In order to conduct experiments and verification, the proposed algorithm was tested on the Corel-1k, Corel-5k and Corel-10k datasets. (1)The Corel-1k dataset contains 10 types of images. There are African, beach, building, bus, dinosaur, flower, elephant, horse, mountain, and food. There are 100 images in each category, for a total of 1,000 images. The image size is 384 × 256 (or 256 × 384). (2) Corel-5K data set, its size is 187 × 126 (or 126 × 187), containing images of 50 categories, including lions, bears, sika deer, drinks, fighting, bonsai, etc., each category 100 images, a total of 5,000 images. (3) Corel-10k data set, the size is 187 × 126 (or 126 × 187). It contains 100 categories, with road signs, poker, butterflies, pistols, musical instruments, aircraft and other category images, each category 100 Images, a total of 10,000 images.

B. PERFORMANCE EVALUATION METRICS

The image retrieval method is based on the precision, recall and F-measure three parameters, which can evaluate the algorithm's ability to retrieve images. In the retrieval process, for a given query image, the retrieval results will return many images, the total number is recorded as N. The images related to the query image are called related images, and the number is n. R is the total number of images related to the query image in the dataset. The calculation equations of precision rate P_r^N and recall rate R_r^N are as follows:

$$P_r^N = n/N \quad (15)$$

$$R_r^N = n/R \quad (16)$$

Here r is the ID number of the retrieved image in the database. The calculation equations for the precision rate and recall rate of each category are

$$P_{avg}^N(c) = \frac{1}{n_1} \sum_{r=1}^{n_1} p_r^N \quad (17)$$

$$R_{avg}^N(c) = \frac{1}{n_1} \sum_{r=1}^{n_1} R_r^N \quad (18)$$

Here c represents each category number, which n_1 is the total number of images in the number category used to query. In addition, the total precision and recall rate of the database are evaluated as follows:

$$p^N = \frac{1}{n_2} \sum_{c=1}^{n_2} p_{avg}^N(c) \quad (19)$$

$$R^N = \frac{1}{n_2} \sum_{c=1}^{n_2} R_{avg}^N(c) \quad (20)$$

Here n_2 is the total number of categories in the database. The total recall rate is also called the average retrieval rate (ARR).

F-measure based on precision and recall is calculated as follows:

$$F_{measure} = \frac{2 \times P^N \times R^N}{P^N + R^N} \quad (21)$$

In order to measure the capability of the proposed method, 10 images from each category are randomly selected. In other words, randomly select 100, 500, and 1000 images from the three datasets corel-1k, corel-5k, and corel-5k respectively as query images. Average precision, recall and F-measure have been calculated for each database, and graphs have been plotted.

TABLE 1. Retrieval results of DCD in different quantization levels for Corel-1k dataset.

quantization levels	Precision%					Recall%				
	10	20	30	40	50	10	20	30	40	50
HSV(6_3_3)	71.53	65.07	61.88	58.71	55.86	7.150	13.01	18.56	23.49	27.93
HSV(8_3_3)	72.60	66.93	63.37	60.74	57.92	7.260	13.39	19.01	24.30	28.96
HSV(12_3_3)	73.79	66.07	62.24	59.14	56.44	7.380	13.21	18.67	23.66	28.22
HSV(8_4_4)	71.87	67.12	63.61	60.63	57.43	7.190	13.42	19.08	24.25	28.72
HSV(12_4_4)	72.79	67.35	63.92	60.49	57.01	7.280	13.47	19.18	24.20	28.51
HSV(16_4_4)	72.60	67.67	63.86	60.69	57.75	7.260	13.53	19.16	24.28	28.88
Ununiform HSV(8_3_3)	**74.94**	**68.92**	**65.52**	**61.14**	**58**	**7.490**	**13.78**	**19.66**	**24.46**	**29**

The best retrieval results are shown in bold

TABLE 2. Retrieval results of HM with different weight values for Corel-1k.

weight values							Precision（%）	Recall%（%）
λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7		
1	0	0	0	0	0	0	43.12	5.17
0	1	0	0	0	0	0	38.60	4.63
0	0	1	0	0	0	0	41.64	5
0	0	0	1	0	0	0	39.18	4.7
0	0	0	0	1	0	0	37.12	4.45
0	0	0	0	0	1	0	36.88	4.43
0	0	0	0	0	0	1	17.58	2.11
1/7	1/7	1/7	1/7	1/7	1/7	1/7	47.47	5.7
0.25	0.25	0.2	0.1	0.1	0.05	0.05	51.17	6.14

The best retrieval results are shown in bold.

C. PERFORMANCE EVALUATION AND PARAMETER ESTIMATION OF THE PROPOSED DESCRIPTOR

In this section, the estimation of the parameters that can enhance the retrieval accuracy of DCD-HM is discussed. Various experiments are conducted to identify the effect for the color quantization of DCD, each moment in HM, distance measure, and the combination of DCD and HM.

1) PERFORMANCE EVALUATION ON THE DIFFERENT COLOR QUANTIZATION OF DCD

Different color quantization methods are used to evaluate the performance of the proposed algorithm. The average precision and recall are shown in Table 1. The images returned in this experiment ranged from 10 to 50. When the color quantization increases and the image return value is 10, the retrieval precision rate is between 71.53% and 74.49%, and the ununiform HSV(8_3_3) can obtain the best result.

2) PERFORMANCE EVALUATION ON THE COMBINATION OF DIFFERENT MOMENTS IN HM

There are 7 moments in the Hu moments. Different moments play different roles on the retrieval results, and we assign different values to the seven moments, which satisfy $\sum_{i=1}^{7} \lambda_i = 1$. The experimental results are shown in the Table 2. We can see that when the returned image is 12, and $\lambda_1 = \lambda_2 = 0.25, \lambda_3 = 0.2, \lambda_4 = \lambda_5 = 0.1, \lambda_6 = \lambda_7 = 0.05$, HM obtains the maximum precision of 51.17%.

3) PERFORMANCE EVALUATION ON THE DIFFERENT DISTANCE METRIC

The two feature vectors $q = (q_1, q_2, \cdots q_m)^T$ and z $= (z_1, z_2, \cdots z_m)^T$ extracted from images, their distance measures can be expressed as:

TABLE 3. Retrieval results of DCD-HM with different distance measures for Corel-1k.

Distance measure	Precision (%)					Recall (%)				
	10	20	30	40	50	10	20	30	40	50
WeightL1	72.89	65.90	61.50	57.83	55.41	7.290	13.18	18.45	23.13	27.71
L1	59.43	54.13	50.12	47.20	44.57	5.940	10.83	15.04	18.88	22.28
Euclidean	63.92	56.07	51.65	48.14	45.28	6.390	11.21	15.49	19.26	22.64
Canberra	71.36	63.41	60.28	56.88	54.09	7.140	12.68	18.08	22.75	27.04
χ^2	60.04	53.37	49.14	46.76	43.94	6	10.67	14.74	18.71	21.97
Proposed Weights	**80.40**	**74.05**	**70.07**	**66.40**	**64.02**	**8.040**	**14.81**	**21.02**	**26.56**	**32.01**

The best retrieval results are shown in bold.

a. L1 distance:

$$D(q, z) = \sum_{s=1}^{m} |q_s - z_s|$$

b. Euclidean distance:

$$D(q, z) = \sqrt{\sum_{s=1}^{m} (q_s - z_s)^2}$$

c. χ^2 statistics:

$$D(q, z) = \sum_{s=1}^{m} \frac{(q_s - z_s)^2}{q_s + z_s}$$

d. Canberra distance:

$$D(q, z) = \sum_{s=1}^{m} \frac{|q_s - z_s|}{|q_s| + |z_s|}$$

e. Weighted L1 distance

$$D(q, z) = \sum_{s=1}^{m} \frac{|q_s - z_s|}{1 + q_s + z_s}$$

In the content-based image retrieval system, the retrieval precision and recall rate are not only determined by the extracted features, but also have a great relationship with the distance measure. Therefore, choosing the right metric is also an extremely critical step.

In the experiment, the retrieval performance of the proposed algorithm is compared under several commonly used distance standards. Euclidean, L1, weighted L1, Canberra metrics are selected. The retrieval results are shown in Table 3, from which it can be seen that the proposed weights distance metric has achieved the best results.

4) PERFORMANCE EVALUATION ON THE COMBINATION OF DCD AND HM

ω_C and ω_S are adjustable weights and satisfy $\omega_C + \omega_S = 1$; The choice of ω_C and ω_S is determined according to the role of the color information and shape information in the retrieval. If value of ω_C or ω_S increases, the corresponding features play a greater role in the final retrieval results. When the returned image is 12 on the dataset corel_1K, the retrieval results of different ω_C and ω_S values are shown in the **Table 4**. From the table, we can see that when $\omega_S = 0.4$, $\omega_C = 0.6$, the maximum precision is 79.36%.

TABLE 4. Retrieval results of DCD-HM with different ω_C and ω_S for Corel-1k.

ω_S	ω_C	Precision (%)	Recall (%)
0	1	73.58	8.83
0.1	0.9	75.33	9.04
0.2	0.8	77.17	9.26
0.3	0.7	78.83	9.46
0.4	**0.6**	**79.36**	**9.52**
0.5	0.5	78.75	9.45
0.6	0.4	77.31	9.28
0.7	0.3	74.58	8.95
0.8	0.2	71.5	8.58
0.9	0.1	64.33	7.72
1	0	51.17	6.14

D. RETRIEVAL PERFORMANCE

When calculating the DCD, the quantized image is 72-dimensional, 72 intervals in an image. Obtain the average value of HSV in each interval and the ratio of the pixels in the interval to the total number of pixels on the image, and the

TABLE 5. Retrieval results with different methods for each category of Corel-1k.

Category	Precision (%)					Recall (%)				
	MTH	CDH	CPV-THF	STH	DCD-HM	MTH	CDH	CPV-THF	STH	DCD-HM
African	65.83	69.17	64.17	75.83	**77.5**	7.9	8.3	7.7	9.1	**9.3**
beach	46.67	47.5	44.17	43.33	**69.17**	5.6	5.7	5.3	5.2	**8.3**
building	68.33	70.83	**80**	66.67	67.5	8.2	8.5	**9.6**	8	8.1
bus	72.5	78.33	**85.83**	60	77.5	8.7	9.4	**10.30**	7.2	9.3
dinosaur	99.17	**100**	**100**	**100**	**100**	11.90	**12**	**12**	**12**	**12**
elephant	70	**77.5**	64.17	61.67	69.17	8.4	**9.3**	7.7	7.4	8.3
flower	85.83	**100**	90	85.83	95.83	10.30	**12**	10.80	10.3	11.5
horse	95	85.83	95.83	92.5	**98.33**	11.40	10.30	11.5	11.1	**11.8**
mountain	50	26.67	50	40.83	**55.83**	6	3.2	6	4.9	**6.7**
food	**78.33**	76.67	72.50	66.67	76.67	**9.4**	9.2	8.7	8	9.2
average	73.17	73.25	74.67	69.33	**78.75**	8.78	8.79	8.96	8.32	**9.45**

The best retrieval results are shown in bold

TABLE 6. Retrieval results with different methods for Corel-1k.

methods	Precision%					Recall%				
	10	20	30	40	50	10	20	30	40	50
MTH	74.80	66	60.33	56.68	52.94	7.480	13.20	18.10	22.67	26.47
CDH	75.1	66.9	62	58.47	55.46	7.51	13.38	18.6	23.39	27.73
CPV-THF	76.60	70.60	66.10	63.05	60.20	7.660	14.12	19.83	25.22	30.10
STH	71.50	64.35	59.83	56.35	53.98	7.150	12.87	17.95	22.54	26.99
DCD-HM	**80.40**	**74.05**	**70.07**	**66.40**	**64.02**	**8.040**	**14.81**	**21.02**	**26.56**	**32.01**

The best retrieval results are shown in bold

DCD color vector is expressed as $72 \times 4 = 288$ dimensions. In the process of calculating Hu moments, in order to obtain better results, the image was divided into blocks. For example, in the Corel-1k data set, the image size is 384×256 (or 256×384). For the convenience of calculation, the image is all changed to 256×384, and the block is divided into 17×17 pixels, ignoring the extra pixel area around the image. The image is divided into 15×22 sub-blocks. Find the moments of each sub-block, that is, get a $15 \times 22 \times 7 = 2310$-dimensional vector representing the shape feature of the image. Therefore, the vector finally obtained by the proposed method is expressed as a feature vector of 2310(HM) + 288(DCD) = 2598 dimensions.

When comparing the distance between the query image of the Hu moment vector and the image to be inspected in the database, the weights of 7 invariant moments satisfy $\sum_{i=1}^{7} \lambda_i = 1$, and $\lambda_1 = \lambda_2 = 0.25, \lambda_3 = 0.2, \lambda_4 = \lambda_5 = 0.1, \lambda_6 = \lambda_7 = 0.05$. When synthesizing the distance feature between Hu moments and DCD, $\omega_C = \omega_S = 0.5$ is obtained. These parameters are set empirically.

1) COREL-1K DATA SET

The results of the proposed algorithm and other algorithms are shown in the Table 5 when the returned image is 12 for each category of Corel-1K. From the table, it can be seen that the proposed algorithm has achieved the best result in the five categories of African, beach, dinosaur, horse and mountain, and the average precision and recall rate are also the highest.

Table 6 highlights the precision rate and recall rate of the proposed algorithm (CDC-HM), MTH, CPV-THF, STH, and CDH when the returned images from 10 to 50 on the data set corel-1k. It can be seen that when the returned images are 10, the precision of the proposed method has been increased from STH, MTH, CDH and CPV-THF up to 8.9%, 5.6%, 5.3%, and 3.8% respectively.

Figure.5 shows the performance of these different methods according to the number of images retrieved. Precision–recall

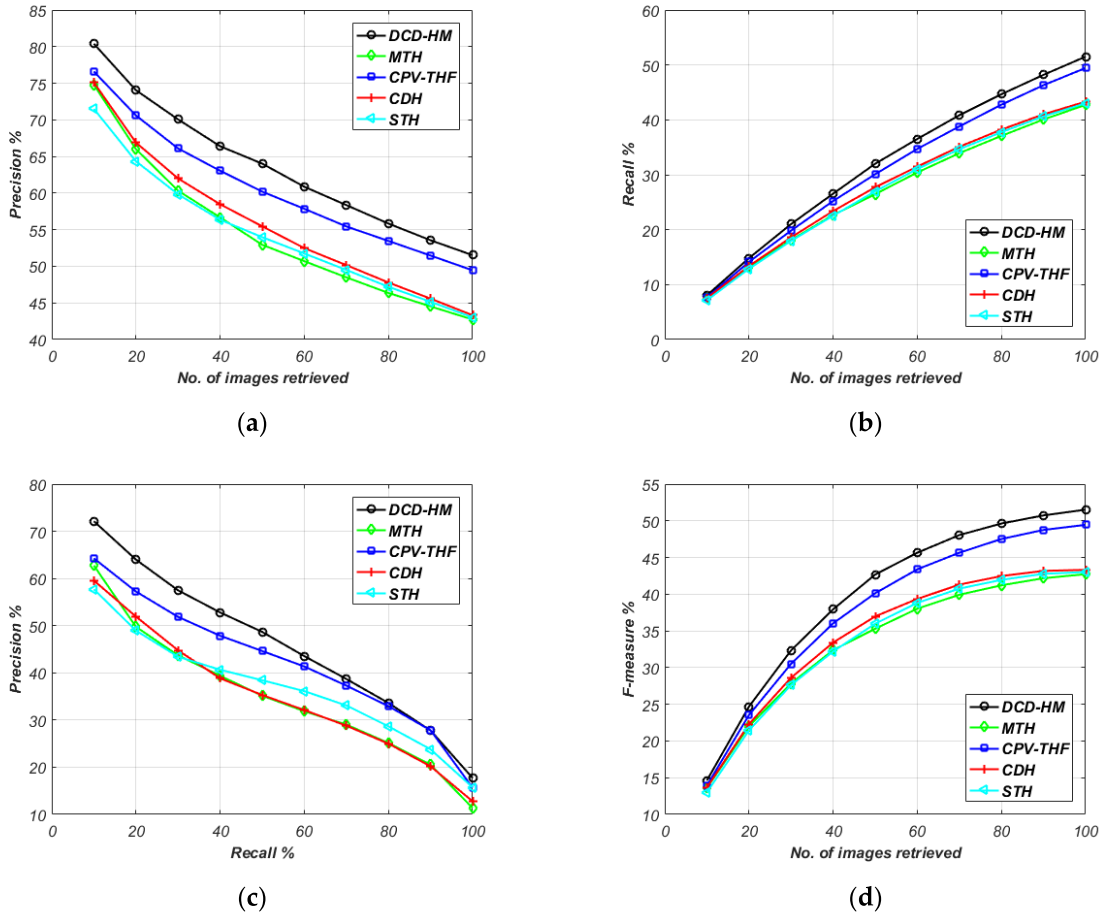


FIGURE 5. The retrieval performance comparison for Corel-1k: (a) Average precision; (b) Average recall; (c) Precision–recall curve; (d) F-measure curve.

TABLE 7. Retrieval results with different methods for Corel-5k.

methods	Precision%					Recall%				
	10	20	30	40	50	10	20	30	40	50
MTH	54.43	46.15	41.32	37.87	35.01	5.440	9.230	12.40	15.15	17.51
CDH	51.80	44.33	40.04	37.18	34.84	5.180	8.870	12.01	14.87	17.42
CPV-THF	57.84	49.40	44.66	41.11	38.33	5.780	9.880	13.40	16.44	19.17
STH	56.02	47.17	42.54	39.47	37.26	5.600	9.430	12.76	15.79	18.63
DCD-HM	**63.47**	**53.92**	**48.89**	**44.84**	**41.77**	**6.350**	**10.78**	**14.67**	**17.94**	**20.88**

The best retrieval results are shown in bold

curve and F-measure curve have also been shown in this Figure∘ It can be seen from the figure that the area enclosed by the PR curve of the proposed algorithm is the largest. It indicates that the proposed method is better than other methods.

2) COREL-5K DATA SET

Use equations (15)-(20) to calculate the average precision and recall rate. And use equation (21) to calculate the F-measure. The retrieval results of different algorithms on the Corel-5k database such as precision, recall, PR curve and F-measure curve are shown in Figure 6. It can be clearly seen from the figure that the proposed algorithm is superior to other algorithms. Table 7 illustrates the precision and recall results in the Corel-5k database when the returned images from 10 to 50. It can be seen from the table that the precision of the proposed method has been significantly improved. The proposed method precision has been significantly improved

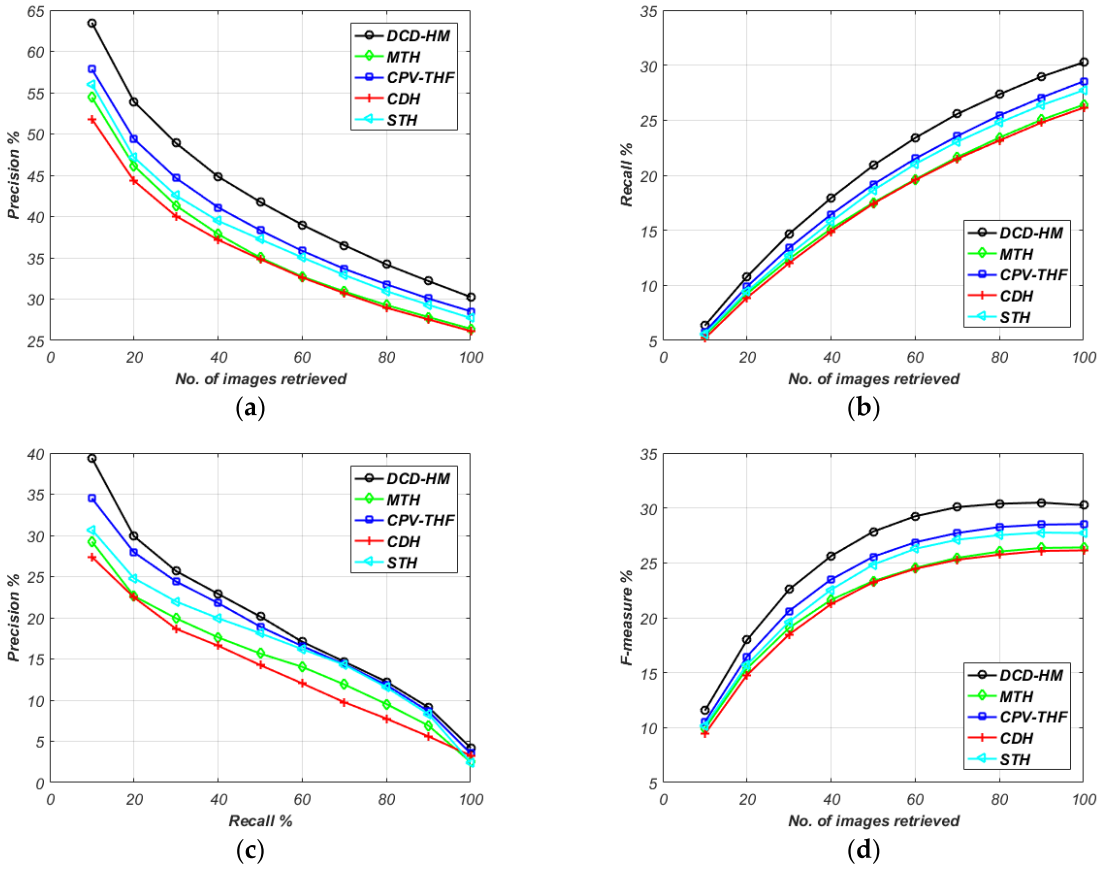


FIGURE 6. The retrieval performance comparison for Corel-5k: (a) Average precision; (b) Average recall; (c) Precision–recall curve; (d) F-measure curve.

TABLE 8. Retrieval results with different methods for Corel-10k.

methods	Precision%					Recall%				
	10	20	30	40	50	10	20	30	40	50
MTH	47.27	38.85	34.66	31.73	29.46	4.730	7.770	10.40	12.69	14.73
CDH	46.93	39.95	35.84	32.69	30.51	4.690	7.990	10.75	13.08	15.25
CPV-THF	51.86	41.98	36.86	33.31	30.69	5.190	8.400	11.06	13.32	15.35
STH	49.50	40.76	36.85	33.78	31.62	4.950	8.150	11.06	13.51	15.81
DCD-HM	**55.90**	**46.17**	**41.18**	**37.23**	**34.37**	**5.590**	**9.230**	**12.35**	**14.89**	**17.19**

The best retrieval results are shown in bold

from STH, MTH, CDH and CPV-THF up to 7.45%, 9.04%, 11.67 and 5.63% respectively when the returned images are 10.

3) COREL-10k DATA SET

Similar to the previous database, use equations (15) – (21) to calculate the precision, recall, and F-measure. The retrieval results of different algorithms on the data set Corel-10k are shown in Figure 7. Table 8 illustrates the precision and recall results in the Corel-10k database when the returned images from 10 to 50. It can be seen from the above figure and table that the precision of the proposed method has been significantly improved. The proposed method precision has been considerably raised from STH, MTH, CDH and CPV-THF up to 6.4%, 8.65%, 8.97% and 4.04% respectively when the returned images are 10.

Figure 8 shows some results of the proposed algorithm in the corel-10k data set, which are cold drinks, fighting and

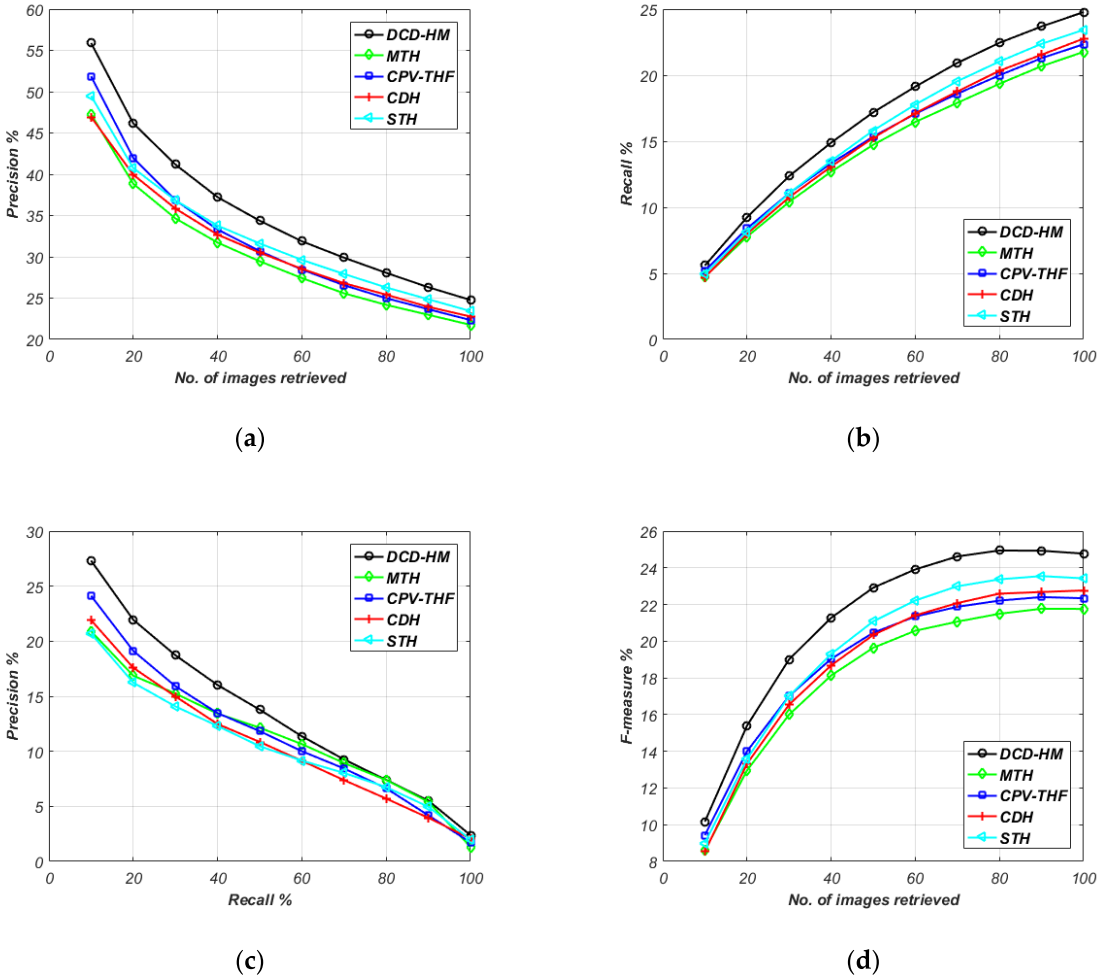


FIGURE 7. The retrieval performance comparison for Corel-10k: (a) Average precision; (b) Average recall; (c) Precision–recall curve; (d) F-measure curve.

TABLE 9. Comparison of proposed DCD-HM, DCD and HM on different datasets.

Methods	Corel-1k (%)		Corel-5k (%)		Corel-10k (%)	
	Precision	Recall	Precision	Recall	Precision	Recall
HM	51.17	6.14	36.83	4.42	30.33	3.64
DCD	73.58	8.83	55.33	6.64	48.92	5.87
DCD-HM	**78.75**	**9.45**	**60.05**	**7.26**	**53.17**	**6.38**

The best retrieval results are shown in bold

fireworks. The first image in the upper left corner is the query image and also the most matching image returned by the database image. The search matching images are arranged from left to right and from top to bottom according to the similarity. From these figures, it can be seen that the database images that are closer to the query image in color and shape are ranked in front of the returned image, indicating that the proposed algorithm has the retrieval characteristics of combining color and shape.

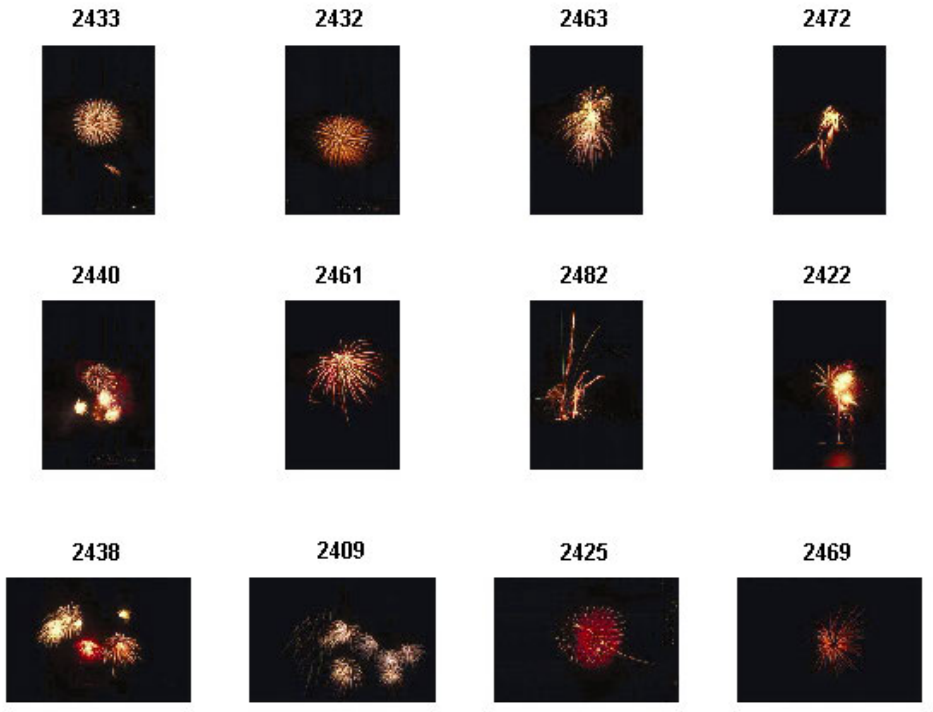
E. RETRIEVAL PERFORMANCE OF DCD AND HU MOMENTS

The proposed method (DCD-HM) in this paper consists of both the dominant color descriptor (DCD) color vector and the Hu moments (HM) shape vector. DCD and Hu moments play different roles in the algorithm for image retrieval. When the returned image is 12, the retrieval results of DCD, HM and DCD-HM applied to corel-1k, corel-5k and corel-10k three data sets are shown in Table 9. In the data set corel-1k, we can

(a)

(b)

FIGURE 8. Image retrieval results on different images: (a) 1702; (b) 1812.



(c)

FIGURE 8. *(Continued.)* **Image retrieval results on different images: (c) 2433.**

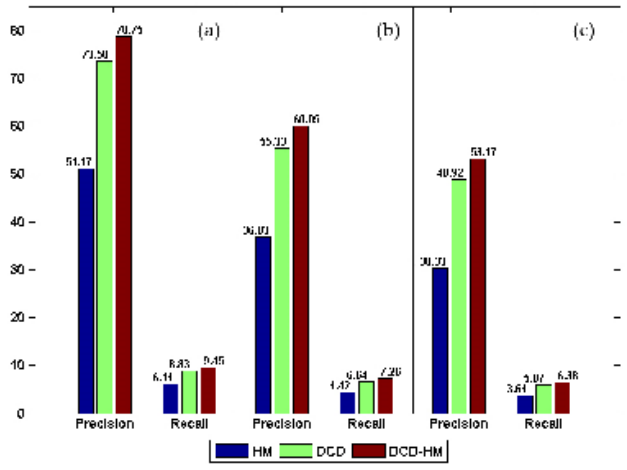


FIGURE 9. **Comparison of precision (%) and recall (%) of proposed HM, DCD and DCD-HM on different datasets: (a) Corel-1K; (b) Corel-5K; (c) Corel-10K.**

see that the precision of HM is 51.17%, the precision of DCD is 73.58%, and the precision of DCD-HM is 78.75%, which is 5.17% higher than the DCD. In Corel-5k and Corel-10k, the precision of DCD-HM is 4.72% and 4.25% higher than the DCD, respectively. It can be seen from the table that the precision of DCD is higher than HM. After combining DCD and HM, the DCD-HM algorithm is superior to the current algorithms. In order to give an intuitive effect, Figure 9 is a comparison chart of the precision of the DCD, HM and DCD-HM methods on the three data sets when the returned image is 12.

TABLE 10. **Feature vector length of different methods.**

Method	Feature vector length
CDH	90+18=108
MTH	64+18=82
CPV-THF	127+75+40=242
STH	45+127=172
DCD-HM	288+2310=2598

The vector length of DCD-HM, CDH, STH, CPV-THF and MTH are shown in Table 10. The feature vector length of the proposed method is the longest than the other methods. Moreover, the proposed method outperforms the other methods in terms of precision as mentioned in different datasets.

V. CONCLUSION

The proposed method (DCD-HM) in this paper combines the characteristics of color and shape. It extracts the average color value of the image quantization interval through the dominant color descriptor, and can improve the retrieval accuracy of

the image by comparing the distance across the interval. The Hu moments feature represents the shape characteristics in an image. The proposed method of combining DCD and HM has the advantages of color and shape detection. Tested on corel-1k, corel-5k and corel-10k three data sets, the experimental results show that its precision and recall rate are better than the state-of-the-art content-based image retrieval methods.

Image retrieval methods can be used in many other computer vision applications, for example skin detection [37], face detection [38], etc. The next step is to apply the proposed algorithm in these areas to improve the versatility of the algorithm.

ACKNOWLEDGMENT

The authors would like to thank the editor and anonymous reviewers for their valuable comments on this article.

AUTHOR CONTRIBUTIONS

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

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

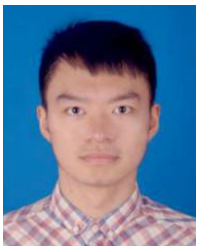
[1] W. Zhou, H. Li, J. Sun, and Q. Tian, "Collaborative index embedding for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, pp. 1154–1166, May 2018.
[2] C. Iakovidou, N. Anagnostopoulos, M. Lux, K. Christodoulou, Y. Boutalis, and S. A. Chatzichristofis, "Composite description based on salient contours and color information for CBIR tasks," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 3115–3129, Jun. 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 Appl.*, vol. 78, no. 16, pp. 23617–23638, Aug. 2019.
[4] Y. Zheng, B. Guo, Y. Yan, and W. He, "O2O method for fast 2D shape retrieval," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5366–5378, Nov. 2019.
[5] L. K. Pavithra and T. S. Sharmila, "An efficient framework for image retrieval using color, texture and edge features," *Comput. Electr. Eng.*, vol. 70, pp. 580–593, Aug. 2018.
[6] 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," *J. Vis. Commun. Image Represent.*, vol. 54, pp. 39–50, Jul. 2018.
[7] W. Song, Y. Zhang, F. Liu, Z. Chai, F. Ding, X. Qian, and S. C. Park, "Taking advantage of multi-regions-based diagonal texture structure descriptor for image retrieval," *Expert Syst. Appl.*, vol. 96, pp. 347–357, Apr. 2018.
[8] C. Singh and K. Preet Kaur, "A fast and efficient image retrieval system based on color and texture features," *J. Vis. Commun. Image Represent.*, vol. 41, pp. 225–238, Nov. 2016.
[9] 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," *Int. J. Speech Technol.*, vol. 48, no. 9, pp. 2930–2953, Sep. 2018.
[10] A. B. Gonde, R. P. Maheshwari, and R. Balasubramanian, "Modified curvelet transform with vocabulary tree for content based image retrieval," *Digit. Signal Process.*, vol. 23, no. 1, pp. 142–150, 2013.
[11] A. Ouni, T. Urruty, and M. Visani, "A robust CBIR framework in between bags of visual words and phrases models for specific image datasets," *Multimedia Tools Appl.*, vol. 77, no. 20, pp. 26173–26189, Oct. 2018.
[12] A. Alzu'bi, A. Amira, and N. Ramzan, "Content-based image retrieval with compact deep convolutional features," *Neurocomputing*, vol. 249, pp. 95–105, Aug. 2017.
[13] C. Kang, L. Zhu, X. Qian, J. Han, M. Wang, and Y. Y. Tang, "Geometry and topology preserving hashing for SIFT feature," *IEEE Trans. Multimedia*, vol. 21, no. 6, pp. 1563–1576, Jun. 2019.
[14] N.-C. Yang, W.-H. Chang, C.-M. Kuo, and T.-H. Li, "A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval," *J. Vis. Commun. Image Represent.*, vol. 19, no. 2, pp. 92–105, Feb. 2008.
[15] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Image indexing using color correlograms," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, San Juan, PR, USA, 1997, pp. 762–768.
[16] G. Pass, R. Zabih, and J. Miller, "Comparing images using color coherence vectors," in *Proc. 4th ACM Int. Conf. Multimedia*, Boston, MA, USA, 1996, pp. 65–73.
[17] J. unić, K. Hirotab, and P. L. Rosin, "A Hu moment invariant as a shape circularity measure," *Pattern Recogn.*, vol. 43, no. 1, pp. 47–57, 2010.
[18] X. Shu, Q. Zhang, J. Shi, and Y. Qi, "A comparative study on weighted central moment and its application in 2D shape retrieval," *Information*, vol. 7, no. 1, p. 10, Mar. 2016.
[19] Y. Zhang, S. Wang, P. Sun, and P. Phillips, "Pathological brain detection based on wavelet entropy and hu moment invariants," *Bio-Med. Mater. Eng.*, vol. 26, no. s1, pp. 1283–1290, Aug. 2015.
[20] X. Ye, J. Zuo, R. Li, Y. Wang, L. Gan, Z. Yu, and X. Hu, "Diagnosis of sewer pipe defects on image recognition of multi-features and support vector machine in a southern chinese city," *Frontiers Environ. Sci. Eng.*, vol. 13, no. 2, pp. 27–39, Apr. 2019.
[21] D. uniá and J. uniá, "Shape ellipticity from HU moment invariants," *Appl. Math. Comput.*, vol. 226, pp. 406–414, Jan. 2014.
[22] Y. Zhang, J. Yang, S. Wang, and Z. Dong, "Pathological brain detection in MRI scanning via Hu moment invariants and machine learning," *J. Exp. Theor. Artif. Intell.*, vol. 29, no. 2, pp. 299–312, 2017.
[23] C. H. Liang and Q. Chang, "Weighted modified hu moment in human behavior recognition," *Adv. Mater. Res.*, vols. 765–767, pp. 2603–2607, Sep. 2013.
[24] R. Min and H. D. Cheng, "Effective image retrieval using dominant color descriptor and fuzzy support vector machine," *Pattern Recognit.*, vol. 42, no. 1, pp. 147–157, Jan. 2009.
[25] A. Fierro-Radilla, K. Perez-Daniel, M. Nakano-Miyatake, and J. Benois, "Dominant color correlogram descriptor for content-based image retrieval," *Proc. SPIE*, vol. 9443, Mar. 2015, Art. no. 94431I.
[26] A. Talib, M. Mahmuddin, H. Husni, and L. E. George, "A weighted dominant color descriptor for content-based image retrieval," *J. Vis. Commun. Image Represent.*, vol. 24, no. 3, pp. 345–360, Apr. 2013.
[27] I. B. Rejeb, S. Ouni, and E. Zagrouba, "Image retrieval using spatial dominant color descriptor," in *Proc. IEEE/ACS 14th Int. Conf. Comput. Syst. Appl. (AICCSA)*, Hammamet, Tunisia, Oct. 2017. pp: 788-795.
[28] L. Xia, "Image retrieval based on dominant color and spatial relationship in DCT domain," *Int. J. Appl. Math. Statist.*, vol. 50, no. 20, pp. 516–524, vol. 2013.
[29] B. Julesz, "Textons, the elements of texture perception, and their interactions," *Nature*, vol. 290, no. 5802, pp. 91–97, Mar. 1981.
[30] G.-H. Liu, L. Zhang, Y.-K. Hou, Z.-Y. Li, and J.-Y. Yang, "Image retrieval based on multi-texton histogram," *Pattern Recognit.*, vol. 43, no. 7, pp. 2380–2389, Jul. 2010.
[31] 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.
[32] A. Raza, T. Nawaz, H. Dawood, and H. Dawood, "Square texton histogram features for image retrieval," *Multimedia Tools Appl.*, vol. 78, no. 3, pp. 2719–2746, Feb. 2019.
[33] G. H. Liu and J. Y. Yang, "Content-based image retrieval using color difference histogram," *Pattern Recognit.*, vol. 46, no. 1, pp. 188–198, 2013.
[34] F. Tajeripour, M. Saberi, and S. F. Ershad, "Developing a novel approach for content based image retrieval using modified local binary patterns and morphological transform," *Int. Arab J. Inf. Technol.*, vol. 12, pp. 574–581, Nov. 2015.
[35] M. K. Alsmadi, "Content-based image retrieval using color, shape and texture descriptors and features," *Arabian J. Sci. Eng.*, vol. 45, no. 4, pp. 3317–3330, Apr. 2020, doi: 10.1007/s13369-020-04384-y.
[36] X. Wang and Z. Wang, "A novel method for image retrieval based on structure elements' descriptor," *J. Vis. Commun. Image Represent.*, vol. 24, pp. 63–74, Jan. 2013.

[37] S. Fekri Ershad, "An innovative skin detection approach using color based image retrieval technique," *Int. J. Multimedia Appl.*, vol. 4, no. 3, pp. 57–65, Jun. 2012.
[38] M. Verma and B. Raman, "Local Tri-directional patterns: A new texture feature descriptor for image retrieval," *Digit. Signal Process.*, vol. 51, pp. 62–72, Apr. 2016.

GUANGYI XIE received the B.S. degree in electronic surveillance from the Armed Police College of Engineering, Xi'an, China, in 1995, and the M.S. degree in signal and information processing from Xidian University, Xi'an, in 2005. He is currently pursuing the Ph.D. degree with the School of Aerospace Science and Technology, Intelligent Control and Image Engineering Institute. His research interests include image processing, machine learning, computer vision, and pattern recognition.

BAOLONG GUO (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in communication and electronic systems from Xidian University, Xi'an, China, in 1984, 1988, and 1995, respectively.

From 1998 to 1999, he was a Visiting Scholar with Doshisha University. He is currently a Full Professor with the School of Aerospace Science and Technology, Xidian University. He has published over 100 academic articles and organized several international conferences as the Co-Chair. His current research interests include pattern recognition, intelligent information processing, image processing, and video communication.



ZHE HUANG received the B.S. degree in automation engineering from the School of Electro-Mechanical Engineering, Xidian University, in 2016, where he is currently pursuing the M.S. degree with the School of Aerospace Science and Technology, Intelligent Control and Image Engineering Institute. His main research interests include deep learning and image processing.

YAN ZHENG received the B.S. degree from Northeast Petroleum University, Daqing, China, in 2012. He is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, Xidian University, Xi'an, China. His research interests include computer vision and machine learning.

YUNYI YAN (Member, IEEE) received the Ph.D. degree in engineering from the Department of Electronic Science and Technology, Xidian University, in 2008.

From 2010 to 2011, he was appointed as a Visiting Scholar with McGill University, Canada. He is currently working with the School of Aerospace Science and Technology, Xidian University, where he is also an Assistant Professor. He mainly involved in the research of control science and control engineering and electronic science and technology. In recent years, he has published more than 30 articles in the IEEE TRANSACTIONS ON DEVICE AND MATERIALS RELIABILITY, the *Journal of Internet Technology*, and other important academic journals, including more than 20 SCI / EI searches, 17 authorized invention patents, and undertaken the National Natural Science Foundation of China and the 863 Program more than ten items. His current interests include space load radiation resistance reliability design, new network intelligent system design, and embedded precision industrial signal detection technology.