A NEW TRADEMARK IMAGE RETRIEVAL ALGORITHM BASED ON MULTIPLE FEATURES FUSION

YUHUA LI, XU HAN, HUI LIANG *, PU LI, XIAYANG SHI, AND JIAN CHANG

ABSTRACT. The existing single feature based methods cannot effectively retrieve trademark images. Therefore, in this paper, we propose a novel trademark image retrieval algorithm making use of fused features to solve the problem, which combines shape features, improved SIFT features, and color features. Our idea is to first extract the local shape features of the input trademark image, and then use the improved SIFT algorithm to detect local key features. After that, the dominant color descriptor is utilized to describe the global color features of the input image. Based on the above extracted different sorts of features, a fused feature vector with different weights describing the overall information of the input trademark image is derived. Finally, the cosine distance is employed to compute the similarity between the searched trademark and the databases. To evaluate the performance of the proposed algorithm, we have conducted a set of query experiments on a data set containing more than 12K trademark images. Experimental results show that the proposed method has great performance in trademark image retrieval with good robustness, and is significantly better than the latest algorithms in query accuracy and efficiency. It also has practical value in the area of well-known trademark detection and trademark copyright protection.

1. INTRODUCTION

With the rapid development of information technology and the diversification of information expression methods, as a kind of rich and intuitive form of multimedia information, image has become more and more popular in daily life. Correspondingly, a large number of image databases have emerged and the number of images is sharply increasing. So it also become more difficult for people to search specific images which meeting their personal demands and preferences from the massive image databases. Traditional keyword based retrieval [1, 2] methods and manual classification of images are far from meeting the practical requirements for image search. Therefore, how to accurately and quickly find the desired images in large volume images has become one of the research hotspots in information retrieval field.

2010 *Mathematics Subject Classification.* 68U10.

Key words and phrases. Trademark image retrieval, shape feature, improved SIFT, dominant color descriptor, multiple feature fusion.

*Corresponding Author: Hui Liang (hliang@zzuli.edu.cn). The authors were jointly supported by the National Natural Science Foundation of China (62072414, U1504608, 61802352), and the Foundation and Cutting-Edge Technologies Research Program of Henan Province (212102210540, 192102210294, 202102210178, 202102210397).

To solve the problem, researchers have proposed various search schemes, one famous of them is content based image retrieval (CBIR) [3], as CBIR provides an intuitive and effective way. So CBIR has been extensively used in many fields, especially, trademark image retrieval [4, 5] has a significant practical application as a subset of CBIR [6, 7, 8]. In addition, in recent years, deep learning technology has also been used for the task of trademark images, and it is also very powerful in terms of retrieval performance. For example, Lan et al. [9], combined the Metric Convolutional Neural Network (CNN) and traditional manual features to retrieve trademark images. However, CNN require a large quantity of data for training in the process of realizing retrieval tasks. Generally speaking, the traditional manual features are more suitable for trademark images, since the trademark images are relatively scarce. Secondly, the features of trademark images are very clear, and improvements on the basis of traditional feature extraction algorithms are sufficient to achieve retrieval effects.

According to the characteristics of the trademark image itself and comparing it with other types of images, we know that the factors that determine the characteristics of a trademark image mainly include shape characteristics [10, 11] and color characteristics [12], and also scale variations [13, 14]. Accordingly the trademark image retrieval completed by single feature is not accurate enough, so we use fused features combining shape features, improved SIFT features with scale invariance, color features to handle this problem.

The proposed algorithm first use the PHOG algorithm [15, 16] to extract the shape features of the trademark image. Then, the improved SIFT algorithm is used to extract the local feature points of the trademark image. Finally, a dominant color descriptor is introduced to extract global color features. At the end, these features are fused with different weights, and normalized to obtain the last feature vector. Also, in order to accurately measure the performance of the algorithm, we introduced several metrics to measure the performance of the proposed method [17, 18, 19, 20].

The rest of this paper is organized as follows. In section 2, The existing literatures on trademark retrieval are presented. In Section 3 we first give the algorithm flowchart of this paper, and then introduce the feature extraction and retrieval algorithms separately. Followed that in Section 4 we give evaluation metrics and experimental results. Finally, this paper is summarized in Section 5.

2. Related Work

Before the process of trademark image retrieval, we simply review some related researches. Phan et al. [21] designed a retrieval scheme utilizing color edge gradient co-occurrence histograms. This scheme expands the color edge-related histogram object detection strategy, and can effectively search for trademarks with different colors. Because it only counts the color information, it does not consider other characteristics of the trademark, so it has a poor search effect for trademarks with similar colors or different shapes. Wei et al. [22] used synthetic features to describe the overall shape and internal structure for trademark retrieval. The method first extracts edges using a Canny edge detector and then normalizes the shape. Although the global shape features and local features are grasped, the color features are ignored, and trademark images with similar shapes but different color contents cannot be distinguished well. Biswas et al. [23] proposed a way to decompose image shape features into simple joint points, and establish distributed features for the shape to achieve the extraction effect. Although this method is easy to calculate the shape feature of the image, it still ignores the color, an important feature of the trademark. Some other cases about matching trademark image similarity through shape features, such as Ye et al. [24] proposed a hierarchical normalized descriptor (HV) for shape features. Anuar et al. [25] used a combination of global descriptors and local descriptors, which means to combine Zernike moments and edge gradient co-occurrence matrix, to improve retrieval performance. And the machine learning algorithm proposed by Trappey et al. [26] automatically detects the shape information of the text in the trademark image. Tripathi et al. [27] developed a trademark retrieval algorithm based on HSV correlation and SIFT operator. By combining shape features and introducing scale-invariant feature transform (SIFT) to represent image features, it is possible to effectively identify various changing images. However, the traditional SIFT method mainly relies on the gradient direction of the pixels in the local area in the process of calculating the main direction of the feature. It can be seen that some existing trademark image retrieval methods have certain defects in some aspects.

We use PHOG to extract the shape features of the trademark images, because compared with traditional methods such as wavelet modulus maxima algorithm [28], edge density histogram [29], and HOG, although PHOG does not have good rotation invariance, it is suitable for trademark retrieval. Some scholars used it for face detection, because facial images are generally upward, and there are few scale changes, and trademark images have similar situations. Moreover, PHOG can not only describe the global shape and local details of the image, but also reflect the spatial information of the images, which is very important for the retrieval of trademark images. In general, when extracting local features of the images, the effects of translation, rotation and zoom on the images should also be considered. Some traditional algorithms such as vector gradient calculation method, PCA-based gradient reduction algorithm, and color space model preprocessed color gradient algorithms. All these methods are affected by the rotation, translation, and scaling of the images, and the number of feature vectors is too large. To this end, inspired by Lowes SITF algorithm, we adopted it as improved SIFT algorithm through modified it from 128 to 64 dimensions, thereby reducing the time complexity. With respect to color feature extraction, we employed the dominant color descriptor, the primary color descriptor (DCD), which has a lot of application [30]. It has two main components, one is the representative color and the other is the percentage of each color. The algorithm is considered as the most suitable color feature extraction method.

Based on the previous results of image retrieval research, the fused feature vector of a trademark image is obtained, with adjusted weights of different features. Utilizing the fused feature vector representation of trademark images, we can further retrieve different types of trademark images. The advantages of the proposed method are proved through comparing the methods only using color feature, or using only shape feature, or using fused shape and color feature.

3. Algorithm and Methodology

Recent studies have shown that in the process of extracting trademark image features, the methods of feature extraction determine the final accuracy of trademark image retrieval. However, different algorithms to extract features have different limitations in terms of calculation volume, complexity, and stability, which leads to a low efficiency of image retrieval. In particular, SIFT features have some shortcomings especially in the calculation complexity. So we have made corresponding improvements to these methods, aiming to improve retrieval accuracy and efficiency.

In this section, we describe the algorithm and processing steps which were used in the proposed retrieval method and illustrate the algorithm with a flowchart, as is shown in Fig. 1.

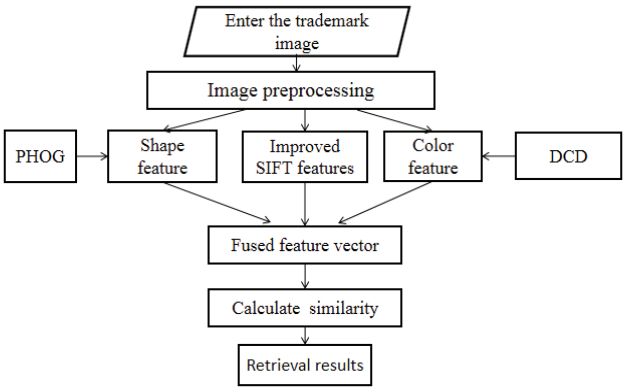

Fig. 1. Flowchart of trademark image retrieval method proposed in this article.

3.1. Image Preprocessing.

The purpose of image preprocessing is to remove interference factors in the images and retain useful information. The reliability of preprocessing can improve feature point detection, feature point positioning, and feature matching. Whether this stage is handled properly is related to the subsequent work smooth or not. Therefore, image preprocessing is the pioneer of the whole system and the prerequisite for ensuring subsequent work. The pre-processing process generally has steps such as image size scaling, noise reduction, etc. Generally, the length and width ratio of a trademark image is 1, and it follows a normal distribution within 100 to 300 pixels, so the length and width of the trademark image are set to 256×256 pixels.

In the process of image data collection, noise is often accompanied by the appearance of noise, which will affects the image itself and hinder the accurate extraction of image features. Therefore, we have to reduce the noise of the image. Some of the existing noise reduction methods are more representative. For example, scholars such as Simi et al. [31] and Kuppusamy et al. [32] used Beltrami filters and improved non-local mean (NLM) filters to reduce the noise of nuclear magnetic resonance images. Joseph et al. [33] used bilateral filters to denoise medical images.

These methods have produced good results in the process of denoising. For trademark images, however, since the quality of the image itself is already very good, in the process of noise reduction, simple preprocessing is enough. Traditional filtering algorithms like median filtering are sufficient to eliminate noise in trademark images. In this article, we used a two-dimensional median filter to improve the quality of trademark images.

3.2. **Extraction of PHOG Features.** The traditional HOG [34] can accurately represent the edge structure of the target image with the spatial distribution of the given image in terms of shape feature extraction, but it does not consider the impact of image division on different spatial scales on retrieval performance. The PHOG shape feature extraction algorithm is improved on the basis of the original extraction method of HOG feature. By using the PHOG feature, the detection time is reduced, and at the same time the ability which is used to describe shape features is to be strengthened, meantime the difference between the background and the target is also increased.

In the PHOG feature extraction process of the trademark image, we first calculate the gradient information of each pixel in the image. Then divide the image into multiple cell regions, and use a gradient histogram with 9 directions to count the gradient magnitude of the pixels in each cell, so as to generate a feature descriptor for each cell, and adjust the histogram of the direction gradient of each cell to get the PHOG features. Adjust according to Equation (3.1) and Equation (3.2) to increase the overall contrast of the gradient histogram of the cell to enhance PHOG features, as shown in Fig. 2. Finally, combine multiple cells into a block area, and then connect the descriptors of the cells in the block area in series to generate the descriptor of the block area,

$$g_{s,Avg} = \frac{\sum_{i=1}^{b} g_s(x,y)}{b} \tag{3.1}$$

$$g'_s = \begin{cases} a_1 g_s(x,y), & g_s \geq g_s(x,y) \\ a_2 g_s(x,y), & g_s < g_s(x,y) \end{cases} \tag{3.2}$$

where $g_s(x,y)$ is the gradient magnitude of the point (x,y) in the cell which belongs to the block s, and $g_{s,Avg}(x,y)$ denotes its average gradient magnitude. Here b is the number of pixels in the cell, a_1 and a_2 are adjustment coefficients respectively, requiring $a_1 \geq a_2$, and $a_1 > 1$, $a_2 > 0$. When $a_1 > a_2$, the gradient whose value is greater than the average value will be increased by a_1 proportion, and the pixel gradient whose gradient value is less than the average value will be reduced by a_2 proportion. After adjusting the gradient amplitude above, each variance of the cell histogram makes the differences in the cell more prominent. The feature enhancement process is shown in Fig. 2.

In terms of the principle of the above PHOG algorithm, we extract the shape features of a trademark image. First we divide the image into N local areas called "blocks", and then continue dividing each area into smaller spatial areas called "units", following that extract the PHOG features from the local area of 16×16 pixels and calculate a direction gradient histogram with 9 directions at a 4×4 local

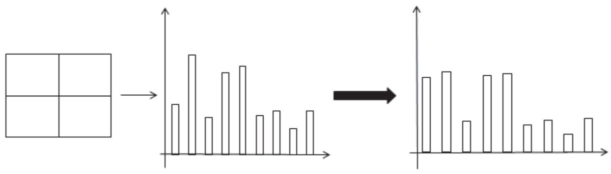

FIG. 2. Illustration of the feature enhancement process.

unit. Next we further divide the gradient direction range into k bins, express the value of k bins as $\phi_k(x, y)$ in Equation (3.3), where G(x,y) represents the gradient value of a point (x,y) on the image, and finally combine the gradient amplitude Equation (3.4) to get the characteristic value of PHOG feature .

$$\phi_k(x,y) = \begin{cases} G(x,y), & if\, \tan^{-1}[\frac{G_x(x,y)}{G_y(x,y)}] \in bin(k) \\ 0, & otherwise \end{cases} \tag{3.3}$$

$$(C,B,k) = \frac{\sum_{(x,y)\in C} \phi_k(x,y)}{\sum_{(x,y)\in B} G(x,y)} \tag{3.4}$$

3.3. **Improved SIFT Algorithms.** The SIFT algorithm firstly performs feature detection on the scale space, determines the position and scale of the key points. Then it uses the main direction of the key point neighbors as the directional feature of the point, to realize the independence of the operator from the scale and direction, and establish the description feature vector, the matching of the feature vector, etc. SIFT has an unparalleled advantage in the invariant feature extraction of images [35]. However, when SIFT extracts feature points, the dimension of feature descriptors make the number of feature vector too much, so we improve it through changing the original dimensions from 128 to 64, which can effectively reduce the time complexity.

For the feature point P, as shown in Fig. 3(a), in the scale space in which it is located, the 16×16 neighbors centered on P are selected, each small square represents a pixel, and then the 16×16 pixels are merged into 4×4 sub-regions, as shown in Fig. 3(b). For each sub-area, the cumulative histogram of its 16 pixels in the 8 gradient directions is calculated, as shown in Fig. 3(c). The final feature point produces a total of 128 sets of data, which is the 128-dimensional feature vector of the SIFT feature descriptor. We replace the square area of SIFT with a circular area. Since the circular area has rotation invariance, it can save the calculation of the main direction in SIFT, reduce the time complexity. We take the number of rings $r = 8$, where r sub-rings have $8 \times r$ elements, so there are 8×8 elements in total, as shown in Fig. 3(d).

We use the improved SIFT algorithm to extract the local feature points of the trademark image, and obtain a feature histogram describing the image through BOW model.

3.4. **Dominant Color Descriptor.** The dominant color descriptor (DCD) is a descriptor for image color features defined in MPEG-7 [36]. It gives a more efficient

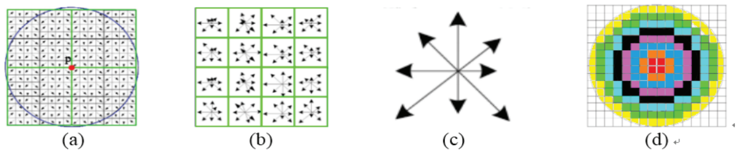

FIG. 3. Schematic diagram of descriptor structure based on improved SIFT feature.

and compact description of the color distribution for an image, and stores only the most important color information with a small amount of memory. In the process of color feature extraction for an image, generally the RGB color space is selected. However, for the true color image, there are many types of colors in RGB space. In order to improve the color quantization robustness of the trademark images, we use the LAB [37] color quantization model. The quantization effect is illustrated in Fig. 4. It can be seen that the color information presented after quantization is more robust and in line with human perception.

FIG. 4. Comparison of two color quantization effects. (a) represents the original example logo image, (b) represents quantization results using LAB method.

In DCD model, colors are quantized into eight dominant color descriptors M according to the human color perception. So the image can be redefined to satisfy the DCD description as follows.

$$M = \{(a_i, b_i), i = 1, \ldots, N, N < 8\} \tag{3.5}$$

where, a_i represents the number of dominant colors, and b_i represents the percentage of dominant colors. Because the percentage of dominant colors is not only affected by their number, but also by the number of other non-dominant colors. Therefore, in order to eliminate the interference of other colors on the dominant color, it is also necessary to use a normalization method to calculate the value of the dominant color's proportion b_i in set B.

$$B = \{\frac{b_1}{\sum_N^1 b_i}, \ldots, \frac{b_N}{\sum_N^1 b_i}\}, \{i = 1, \ldots, N\} \tag{3.6}$$

The following set E is the feature expression used when extracting dominant color features.

$$E = \{(a_i, b_i), i = 1, \ldots, N, N < 8\}, \quad \sum_N^1 b_i = 1 \tag{3.7}$$

We use the dominant color descriptor to represent features from different color regions of the image in Fig. 4(b), and the results are shown in Fig. 5. Where (a) represents the regional characteristics of the dominant color of red, (b) represents the regional characteristics of the dominant color of yellow, (c) represents the regional characteristics of the dominant color of blue, (d) represents the regional characteristics of the dominant color of white.

FIG. 5. Dominant color descriptor to extract features of different color regions.

3.5. **Feature Fusion.** Multi-feature fusion is the process of combining two or more feature vectors into a single feature vector. These feature attributes have a certain complementarity. Here, we adopt a more ideal feature level fusion method to integrate multiple features of the trademark image. By adjusting the proportion of each feature's weight, feature vectors capable of expressing image information are organically combined to comprehensively describe the features of the image. In terms of time complexity calculation, it is mainly used for feature extraction. The feature level fusion method will not increase the time complexity due to the dimensionality of the feature vector. The specific implementation steps are narrated as follows.

First, the combination of weight values is set as $\omega = [\omega_i, \omega_{ij}, \omega_{ijk}]$, where $\omega_i = 1/I$, represents the initial value of each feature in feature vector of an image, here I represents the total number of features. And based ω_i, ω_{ij} is proposed to represent the adjusted weight values according to the similarities of each feature in retrieval experiments, and the initial $\omega_{ij} = 1/J_i$, J_i indicates the dimension of feature f_i. $\omega_{ijk} = 1/K_{ij}$ is obtained through further subdividing the weight intervals and K_{ij} is acquired through the retrieval experiments. In these experiments first record M related images in the returned image set for each feature, and then calculate the K_{ij}, the standard deviation of the K column of the matrix $M \times K$ which consisted of eigenvectors of M.

Next we normalize the shape feature, improved SIFT feature and color feature of the trademark image. After that, under the distribution of different weight values, similarity detection is performed on these three features, and the weight corresponding to the highest point of similarity is found. Finally, according to the weight under the corresponding different features, the fusion is carried out by weighted average. The feature vector of the image is lastly represent as follows.

$$V_f = \frac{\omega_1 V_p + \omega_2 V_s + \omega_3 V_c}{\omega_1 + \omega_2 + \omega_3} \tag{3.8}$$

where V_f is the fused feature vector, $_i$ represents the weights, and V_i represents the feature vector of the three different features. According to our repeated experiments on various features of the trademark image, a more appropriate weight coefficient is obtained, in which shape feature occupies 0.46, improved SIFT feature occupies 0.32, and color feature occupy 0.22. According to the fused feature vector, the similarity between two different trademark images is obtained.

4. Experiments

In this section, we first describe the collected trademark dataset, and then give several methods for evaluating the performance of the algorithm in this paper. By setting up comparative experiments, the multi-feature fusion trademark image retrieval algorithm proposed is compared with the existing single feature based and dual features based retrieval algorithms. Experimental results proved the superiority and reliability of the proposed algorithm.

4.1. **Datasets.** For the trademark image datasets used in our experiments, a part of images came from Trademark Office of the State Intellectual Property Office, and a part of were acquired from some websites, like Fickr, mainly means FlickrLogos-32. Among them, the FlickrLogos-32 trademark image data set is divided into 32 types of trademark images, containing a total of 8,240 images, all of which have a relatively flat surface. This dataset is currently very popular in the task of searching trademark images. The collected trademark image dataset includes various types of trademarks such as school emblems, sporting goods, automobile brands, and clothing brands. In order to fully illustrate the reliability of the algorithm in this paper, we have also added some non-trademark images to the dataset, and performed a retrieval experiment on the trademark images on a total of more than 12K dataset. Fig. 6 shows some examples of the dataset.

Fig. 6. Some examples of trademark images in our datasets.

4.2. **Evaluation Metrics.** In order to verify the reliability and advantages of this algorithm for trademark image retrieval, we used several widely used metrics to evaluate the performance of trademark retrieval experiments. They are Precision Recall Plot (PR), mean Average Precision (mAP), and Discounted Cumulative Gain (DCG). First, precision recall plot (PR), therein the recall rate, refers to the ratio of the number of valid images in the search results to the total number of valid images in the image library. The higher the recall rate, the fewer images in the image library that are related to the image to be searched but have not been detected. Accuracy, also called retrieval accuracy, is the ratio of the total number of valid images in the query results to the total number of images returned. Precision (usually abbreviated to P), and Recall (usually abbreviated to R), are defined as follows.

$$P = \frac{T_p}{T_p + F_p} \tag{4.1}$$

$$P = \frac{T_p}{T_p + F_N} \tag{4.2}$$

The mean Average Precision (mAP) is a popular single number metric used to measure the performance in the task of information retrieval and object detection. It is the mean value on the basis of Average Precision (AP) representing the average of all precisions at all ranks where relevant searched objects found. The mAP can be computedas follow.

$$mAP = \frac{1}{|Q_R|} \sum_{q \in Q_R} AP(q) \tag{4.3}$$

where q means a query, Q_R denotes the whole image set and $|Q_R|$ represents the number of the whole image set.

DCG evaluates the retrieval performance based on the position of the image retrieval results in the returned list. Images with a high degree of relevance appear at the front of the returned image list than those at the back of the ranked list. The DCG value at a particular position p in the returned image list can be computer as follows.

$$DCG_p = rel_1 + \sum_{2}^{p} \frac{rel_1}{\log_2 i} \tag{4.4}$$

where rel_1 represents the query relevance score for the searched i-th image in returned result list.

4.3. **Experiment Results and Comparisons.** In order to comprehensively evaluate the performance of the retrieval scheme in this article, and to the retrieval results are more clear, we choose four representative retrieval algorithms, an alphabetic contour-based descriptor for shape-based image retrieval [17], here abbreviated as (DP), image retrieval using color and texture fused features [18], here abbreviated as (HSV). Retrieval of trademark images based on the combination of SIFT and color features [19], here abbreviated as (SCF), as well as the method using shape and key local color features [20], here abbreviated as (WMM+LCH), for comparative experiments. By comparing the search results, we can clearly see the accuracy of the algorithm in this paper and other search methods. As for retrieval complexity, we also give the corresponding time used respectively with statistical data. First, a comparative experiment is performed with the DP [17] that uses only shape features to retrieve images. We chose the BMW logo as the trademark for the query, and compared the first 4 images found. As shown in Fig. 7.

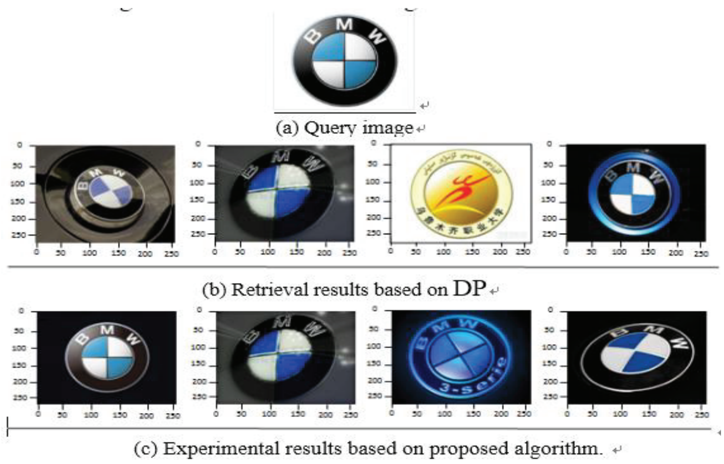

FIG. 7. The comparative experimental results obtained by the proposed method and the retrieval algorithm DP [17].

DP uses an alphabetic contour-based descriptor method to describe the shape features of the image, and by using dynamic programming to find the best similar segment of the shape sequence of the two images. It can be seen from the retrieval results that in the search results returned by DP, there is an unwanted trademark. Although this trademark and the given query are both circular in the outer shape, their colors are quite different. The DP method is a little hard to balance the weight of the shape feature and the color feature when extracting the features, resulting in retrieval errors due to different colors. According to the mAP metric, obtained by the DP algorithm is 0.687, and the mAP value obtained by the proposed algorithm is approximately 0.88. It can be seen that the trademark retrieval method proposed in this article can extract the shape and color features more accurately.

The following comparative experiment was conducted using HSV [18]. We select the NIKE trademark as the query image, the test results are shown in Fig. 8.

In HSV algorithm, the quantized HSV color space model is utilized to extract features in the form of color histogram, and then the global color histogram and the local color histogram are compared to obtain the feature vector describing the color information of the image. From the results of the retrieval, it can be seen that the third and fourth images retrieved in Fig. 8(b) are obviously not the trademark images we need. The reason is that the trademark is searched only by the proportion of color features in the image, which largely ignores the main feature of shape. As long as the proportion of a certain color proportion is large enough, it is considered the same trademark image, which is obviously incorrect. Similarly, according to the mAP performance evaluation standard, we can calculate that the mAP value with the method in the literature [18] is only 0.55, which is quite different from the retrieval result in the ideal state.

Next, we conducted a comparative experiment by employing SCF (SIFT and color features) in [18]. We select the KFC trademark with irregular shape characteristics as the image to be retrieved, and the experimental results are shown in Fig. 9.

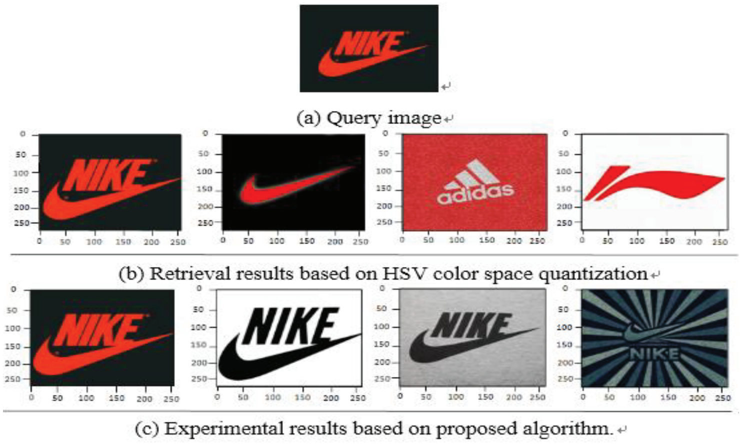

FIG. 8. The comparative experimental results obtained by the proposed method and the retrieval algorithm in [18].

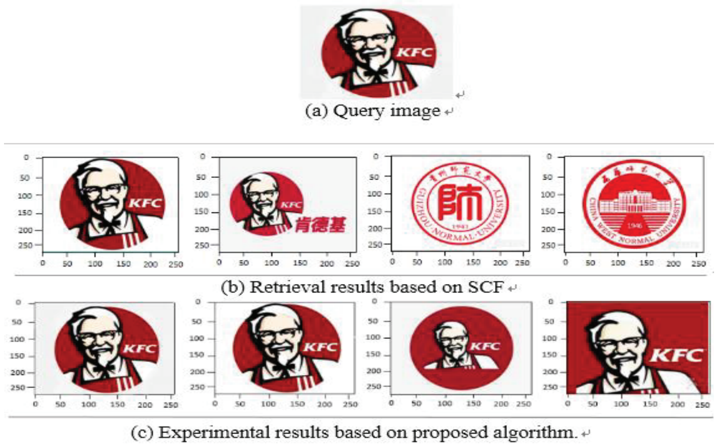

FIG. 9. The comparative experimental results obtained by the proposed method and the retrieval algorithm in [19].

For the KFC trademark image, we compared the retrieval algorithm in this article with the SCF algorithm. It can be seen that in Fig. 9(b), although the latter two images remain highly similar to the original image in terms of SCF (SIFT and color features), the results were not satisfied, due to the mismatch results. This is because the local shape features are not fully extracted, and the SIFT and color features occupy most of the proportion.

Finally, a comparative experiment was conducted using WMM+LCH in [20]. We chose the trademark with the zoom ratio as the query mark, and the test result is shown in Fig. 10.

In [20], Wavelet Modulus Maxima (WMM) is used to extract the shape features of the image, and the Local Cumulative Histogram (LCH) is used to describe the shape features of the image. Then the shape and color features are combined according to a

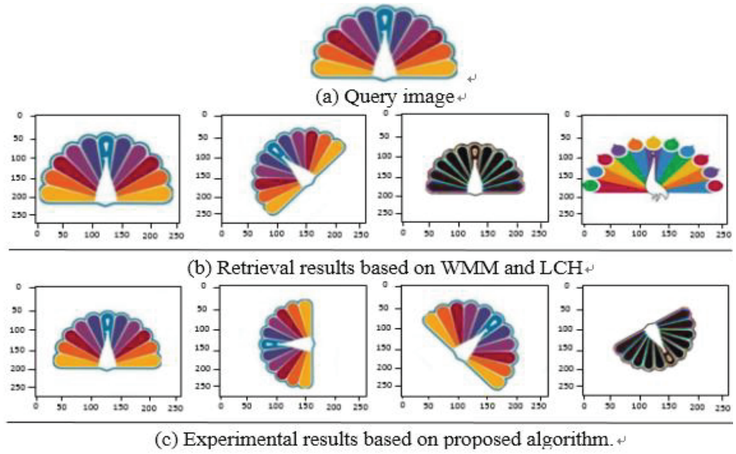

FIG. 10. The comparative experimental results obtained by the proposed method and the retrieval algorithm in [20].

certain weight to obtain a feature vector describing the overall characteristics of the image. By comparing the obtained search results, we can find that the fourth image retrieved in Fig. 10(b) is obviously not the trademark image we need. Although it is very similar in shape and color to the query image, when the trademark image undergoes changes in spatial position such as rotation, translation, etc., the retrieval error of the algorithm in [20] will occur. In this paper, SIFT features are further added on the basis of shape and color features, and SIFT algorithm is improved. We replace the square area of SIFT with a circular area. Because the circular area has rotation invariance, the global features are not affected by the local features. According to the DCG evaluation criteria, we can see that in the retrieved images, the images ranked at positions 1 and 2 have reached great similarity, and at position 3, the similarity obtained by the retrieval method in this paper is significantly higher than the similarity obtained the method combination of WMM and LCH, the corresponding DCG value is larger. And the mAP value obtained is also higher than the method WMM+LCH's 0.750.

The above are three representative comparative experiments that we have selected. It can be seen that the proposed algorithm has achieved ideal results according to retrieval accuracy and efficiency. In order to further test the performance of the algorithm in this paper, we also made the Precision-Recall plot of each algorithm as shown in Fig. 11.

According to Fig. 11, it can be seen that the proposed algorithm taking into account the shape, SIFT and color characteristics of the trademark at the same time, achieved great retrieval results. Also, after a series of retrieval experiments, we can obtained the weight coefficients of each feature, which can be used to fuse the these features. The results of the experiments clearly demonstrated the retrieval performance of the proposed method in retrieving trademark images. Moreover, it also showed the existing methods for trademark retrieval through single or dual features still have certain limitations.

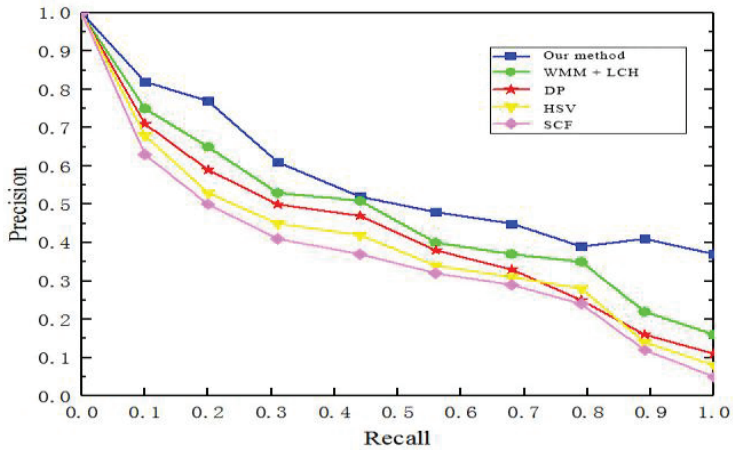

Fig. 11. P-R plot comparisons of the proposed algorithm with others in performance.

5. Conclusion

Trademark image retrieval has an important application as a subset of CBIR. So in this paper, a trademark image retrieval algorithm is presented, which based on the way of multiple features fusion. The algorithm first utilizes PHOG to represent the shape feature of target images, and then an improved SIFT method is utilized to describe the internal details of them. Lastly, the dominant color descriptor is used to describe the global color features of the input images. By extracting the shape, SITF, color features of the trademark images and setting different weights, the three features are totally fused together. In this research, we demonstrate the benefits of the proposed algorithm in respect of the latest methods through comparing the experimental results and performance. It is noted that our method has not only great performance, but also good efficiency in the real retrieval experiments for trademark retrieval.

References

[1] R. Zanibbi, G. Agam and C. Viard-Gaudin, *Keyword and image-based retrieval of mathematical expressions*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. **7874** (2011), 1–10.

[2] Y. Li, Z. He, S. Wang, Z. Wang and W. Huang, *Multideep Feature Fusion Algorithm for Clothing Style Recognition*, Wireless Communications and Mobile Computing. **2021** (2021), 14.

[3] V. A. Wankhede and P. S. Mohod, *Content-based image retrieval from videos using CBIR and ABIR algorithm*, 2015 Global Conference on Communication Technologies, 2015, 767–771.

[4] X. Li, J. Yang and J. Ma, *Recent developments of content-based image retrieval (CBIR)*, Neurocomputing, 2021.

[5] O. Tursun, S. Denman and S. Sivapalan, *Component-based Attention for Large-scale Trademark Retrieval*, IEEE Transactions on Information Forensics and Security, 2019.

[6] Y. J. Wang and C. F. Zheng, *Trademark Image Retrieval Based on Shape and Key Local Color Features*, 2009 Second International Conference on Information and Computing Science, **2** (2009), 325–328.

[7] Reddy G T, Reddy M. P. K., Lakshmanna K, et al, *Analysis of dimensionality reduction techniques on big data*, IEEE Access, **8** (2020), 54776–54788.

[8] Y. Li, S. Xu and X. Luo, *A new algorithm for product image search based on salient edge characterization*, Journal of the Association for Information Science and Technology, **65** (2014), 2534–2551.

[9] T. Lan, X. Feng and L. Li, *Similar Trademark Image Retrieval Based on Convolutional Neural Network and Constraint Theory*, 2018 Eighth International Conference on Image Processing Theory, 2018.

[10] F. A. Andalo, P. A. Miranda and R. D. Torres, *Shape feature extraction and description based on tensor scale*, Pattern Recognition, **43** (2010), 26–36.

[11] J. P. Eakins, K. J. Riley and J. D. Edwards, *Shape feature matching for trademark image retrieval*, Conference on image and video retrieval, 2003, 28–38.

[12] W. Chen, W. Liu and M. Chen, *Adaptive Color Feature Extraction Based on Image Color Distributions*, In IEEE Transactions on Image Processing, **19** (2010), 2005–2016.

[13] W. Zhong, Q. Ke, M. Isard and J. Sun, *Bundling features for large scale partial-duplicate web image search*, 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, 25–32.

[14] Z. F. Shao, L. I. Deren and X. Q. Zhu, *A multi-scale and multi-orientation image retrieval method based on rotation-invariant texture features*, Science in China Series F: Information Sciences, **54** (2011), 732–744.

[15] S. Saad, M. S. Yasein, M. H. Mousa and H. Nassar, Pedestrian detection using shape context and PHOG, 2014 9th International Conference on Computer Engineering & Systems, 2014, 277–281.

[16] K. Deepika, I. Ruth, S. Keerthana, B. Sathya Bama, S. Avvailakshmi and A. Vidhya, *Robust plant recognition using Graph cut based flower segmentation and PHOG based feature extraction*, 2012 International Conference on Machine Vision and Image Processing, 2012, 44–47.

[17] A. T. Anaraki, U. U. Sheikh, A. Rahman and Z. Omar, *An alphabetic contour-based descriptor for shape-based image retrieval*, 2017 IEEE International Conference on Signal and Image Processing Applications, 2017, 145–148.

[18] S Liu, D Liu, G Srivastava, D Połap, M Woźniak, *Overview and methods of correlation filter algorithms in object tracking*, Complex & Intelligent Systems, **3** (2020), 1–23.

[19] Li, Y., Zhan, Z. H., Lin, S., Zhang, J., and Luo, X, *Competitive and cooperative particle swarm optimization with information sharing mechanism for global optimization problems*, Information Sciences, **293** (2015), 370–382.

[20] W Fu, S Liu, G Srivastava, *Optimization of big data scheduling in social networks*, Entropy, **21** (2019), 902.

[21] R. Phan and D. Androutsos, *Content-based unconstrained logo and trademark retrieval in color image databases with Color Edge Gradient Co-occurrence Histograms*, 2009 16th International Conference on Digital Signal Processing, 2009, 1–6.

[22] C. Wei, Y. Li, W. Chau and C. Li, *Trademark image retrieval using synthetic features for describing global shape and interior structure*, Pattern Recognition, **42** (2009), 386–394.

[23] S. Biswas, G. Aggarwal and R. Chellappa, *An Efficient and Robust Algorithm for Shape Indexing and Retrieval*, In IEEE Transactions on Multimedia, **54** (2010), 372–385.

[24] B. Ye, Z. Cai, T. Lan and Y. Xiao, *A Novel Shape Representation Method for Complex Trademark Image*, In IEEE Access, **7** (2019), 53800–53811.

[25] F. M. Anuar, R. Setchi and Y. K. Lai, *Trademark image retrieval using an integrated shape descriptor*, Expert Systems with Applications, **40** (2013), 105–121.

[26] C. V. Trappey, A. J. C. Trappey and C. C. Lin, *Intelligent trademark similarity analysis of image, spelling, and phonetic features using machine learning methodologies*, Advanced Engineering Informatics, **45** (2020), 101–120.

[27] C. Tripathi and A. Nigam, *Trademark image retrieval using weighted combination of sift and HSV correlogram*, Journal of Computer Applications in Technology, **54** (2016), 61–67.

[28] O. Prakash and A. Khare, *CT and MR Images Fusion Based on Stationary Wavelet Transform by Modulus Maxima*, International Conference on Computational Vision & Robotics, 2014, 199–204.

[29] M. Mignotte, *An Energy-Based Model for the Image Edge-Histogram Specification Problem*, IEEE Transactions on Image Processing, **21** (2012), 379–386.

[30] J. Zhang, Y. Barhomi and T. Serre, *A new biologically inspired color image descriptor*, European conference on computer vision, 2012, 312–324.

[31] V. R. Simi, D. R. Edla, J. Joseph and V. Kuppili, *Analysis of controversies in the formulation and evaluation of restoration algorithms for MR Images*, Expert Systems with Applications, **135** (2019), 39–59.

[32] P. G. Kuppusamy, J. Joseph and S. Jayaraman, *A customized nonlocal restoration schemes with adaptive strength of smoothening for magnetic resonance images*, Biomedical Signal Processing and Control, **49** (2019), 160–172.

[33] J. Joseph and R. Periyasamy, *An image driven bilateral filter with adaptive range and spatial parameters for denoising Magnetic Resonance Images*, Computers & Electrical Engineering, **69** (2018), 782–795.

[34] V. Bisot, S. Essid and G. Richard, *HOG and subband power distribution image features for acoustic scene classification*, 2015 23rd European Signal Processing Conference, 2015, 719–723.

[35] H. Soyel and H. Demirel, *Facial expression recognition based on discriminative scale invariant feature transform*, Electronics Letters, **46** (2010), 343-3-45.

[36] P. Maheshwary and N. Srivastav, *Retrieving Similar Image Using Color Moment Feature Detector and K-Means Clustering of Remote Sensing Images*, International conference on computer and electrical engineering, 2008, 821–824.

[37] M. E. Celebi, *Improving the performance of k-means for color quantization*, Image and Vision Computing, **29** (2011), 260–271.

(Yuhua Li) Software Engineering College, Zhengzhou University of Light Industry, Zhengzhou 450001, China
E-mail address: `llyhua@126.com`

(Xu Han) School of Business Administration, Xinyang Agriculture and Forestry University, Xinyang 464000, China; Software Engineering College, Zhengzhou University of Light Industry, Zhengzhou 450001, China
E-mail address: `zzqyhanxu@163.com`

(Hui Liang) Software Engineering College, Zhengzhou University of Light Industry, Zhengzhou 450001, China
E-mail address: `hliang@zzuli.edu.cn`

(Pu Li) Software Engineering College, Zhengzhou University of Light Industry, Zhengzhou 450001, China
E-mail address: `superlipu@163.com`

(Xiayang Shi) Software Engineering College, Zhengzhou University of Light Industry, Zhengzhou 450001, China
E-mail address: `aryang123@163.com`

(Jian Chang) National Centre for Computer Animation, Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, United Kingdom
E-mail address: `jchangsuper@163.com`