Image retrieval based on colour and improved NMI texture features

Anyu Du, Liejun Wang & Jiwei Qin

Published online: 05 Sep 2019.

Submit your article to this journal ⬈

Article views: 288

View related articles ⬈

View Crossmark data ⬈

Taylor & Francis
Taylor & Francis Group

REGULAR PAPER

🔓 OPEN ACCESS ⊙ Check for updates

Image retrieval based on colour and improved NMI texture features

Anyu Du[a], Liejun Wang[b] and Jiwei Qin[c]

[a]School of Information Science and Engineering, Xinjiang University, Urumchi, People's Republic of China; [b]School of Software Engineering, Xinjiang University, Urumchi, People's Republic of China; [c]School of Education, Shaanxi Normal University, Xi'an, People's Republic of China

ABSTRACT

This paper proposes an improved method for extracting NMI features. This method uses Particle Swarm Optimization in advance to optimize the two-dimensional maximum class-to-class variance (2OTSU) in advance. Afterwards, the optimized 2OUSU is introduced into the Pulse Coupled Neural Network (PCNN) to automatically obtain the number of iterations of the loop. We use an improved PCNN method to extract the NMI features of the image. For the problem of low accuracy of single feature, this paper proposes a new method of multi-feature fusion based on image retrieval. It uses HSV colour features and texture features, where, the texture feature extraction methods include: Grey Level Co-occurrence Matrix (GLCM), Local Binary Pattern (LBP) and Improved PCNN. The experimental results show that: on the Corel-1k dataset, compared with similar algorithms, the retrieval accuracy of this method is improved by 13.6%; On the AT&T dataset, the retrieval accuracy is improved by 13.4% compared with the similar algorithm; on the FD-XJ dataset, the retrieval accuracy is improved by 17.7% compared with the similar algorithm. Therefore, the proposed algorithm has better retrieval performance and robustness compared with the existing image retrieval algorithms based on multi-feature fusion.

1. Introduction

Image retrieval has received more and more attention in recent years. It refers to the process of searching for a query image in a database, and then searching for and representing images related to user needs [1]. Image retrieval methods are mainly divided into two categories, text-based and content-based [2]. The early text-based image retrieval algorithm is a method for manually annotating image content and performing keyword search. However, this method had many shortcomings. For example, text-based image retrieval mainly relies on the standard information of images for retrieval, but manual indexing is hard, and this process is costly and time consuming [1], after which the researchers began to focus on content-based image retrieval technology. In the context of content-based image retrieval, the study of colour, texture, and shape features is relatively mature. In the beginning, the main research was on image retrieval based on single features. Such as: Su CH et al. [3] proposed an efficient image retrieval based on HSV colour space; Sulochana et al. [4] proposed texture-based image retrieval using framelet transform – Grey Level Co-occurrence Matrix (GLCM); Mishra et al. proposed an improved content-based image retrieval approach using Local Binary Pattern (LBP) [5]. However, searching based on a single feature can only be used to retrieve images with

strong corresponding features. For images with complex features, the retrieval efficiency is low. Therefore, the research of multi-features fusion retrieval increasingly attracts people's attention.

Later, Rahmaniansyah Dwi Putri et al. [6] proposed feature extraction of HSV colour and GLCM feature fusion; Swati Jain et al. proposed weighted fusion of LBP and CHH features [7]; Liu P et al. proposed fusion of colour histogram and LBP-based features for texture image retrieval and classification [8]. The accuracy of the above method is higher than that of a single feature, but it is bad for image processing such as smoothing, filtering, and compression. Then, Q Liu et al. [9] proposed an image NMI feature extraction based on pulse-coupled neural network. Although this method avoids poor image retrieval results caused by image smoothing and compression, the number of loop iterations needs to be determined artificially when extracting NMI features, resulting in the accuracy has declined. In recent years, researchers began to study how to obtain the number of iterations of the PCNN automatically, and then proposed to use the 2OTSU algorithm to automatically obtain the number of iterations of the loop. The time complexity of 2OTSU is relatively large. In order to solve this problem, the particle swarm optimization (PSO) algorithm is used to optimize the 2OTSU in this paper, and the optimized algorithm is introduced into

the PCNN to automatically obtain the NMI features of the image. This paper fused the features extracted by HSV, GLCM and LBP and the NMI feature extracted by the improved PCNN to improve the retrieval accuracy.

The rest of the paper is structured as follows: Section 2 reviews the related work of image retrieval; Section 3 describes proposed method; Section 4 explicates experimental results and analysis; Section 5 summarizes this paper.

2. Related work

The concept of "content-based image retrieval" was proposed by T. Kato in 1992. Where, he used colour and shape to describe the automatic image retrieval experiments of the databases [3]. The concept of CBIR has been widely used and has achieved great development. This paper proposed a new image retrieval method based on this concept. Its system block diagram is shown in Figure 1. The system begins to input the query image and enters the HSV feature extraction, GLCM feature extraction, LBP feature extraction and NMI feature extraction stages respectively, and weighted fusion of four features; afterwards, the similarity measurement stage is entered. At this stage, the query image matches the image in the database; when the matching process is completed, an image with similar matching values is selected in the database and the image query.

2.1. Colour feature extraction

Colour features are one of the basic visual features that humans perceive and distinguish between different objects. Every object in the world has its own colour characteristic. The colour features are very stable, and they are insensitive to rotation, translation, scale changes, and even various deformations. They showed strong robustness and is the most commonly used features of content-based image retrieval [10]. Colour histogram is the most commonly used colour features in image retrieval systems, while colour correlation diagram depicts the proportion of a certain number of colour pixels in an image, and also reflect the spatial correlation between different colours [11].

The choice of colour spaces and quantization levels are the two main aspects of colour feature extraction. The colour spaces are the multidimensional space in which different dimensions have different colour components. This paper used the HSV colour model because it is closer to the human eye's perception of colour. In the HSV model, H (hue) represents the chroma component in the model, S (saturation) represents the advantage of a particular hue in the hue, and V (value) represents the hue intensity [12]. Since human eyes have different perceptions of hue, saturation, and numerical values, the three components of H, S and V are quantified non-uniformly in this paper. Hue is divided into saturation and numerical values
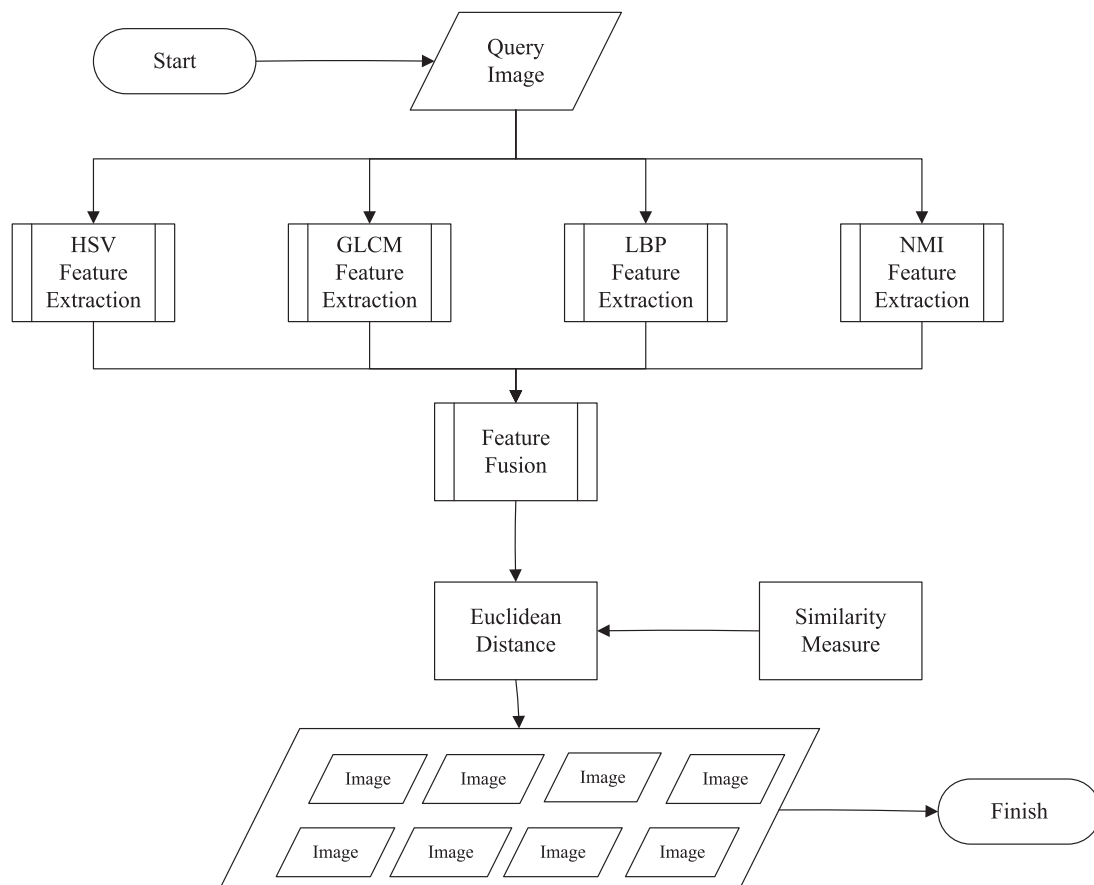

Figure 1. CBIR system flow diagram.

of 8 parts and quantized according to different colour ranges. After quantization, the three colour components are combined into the one-dimensional feature vector Z shown in equation (1):

$$Z = Q_S Q_V H + Q_V S + V \qquad (1)$$

where Q_S and Q_V, respectively, represent the quantization levels of S and V [1]. Let $Q_S = 2$, $Q_V = 2$, we can get formula (2):

$$Z = 4H + 2S + V \qquad (2)$$

Therefore, after quantization, the colour histogram vector can be expressed as a 1×32 one-dimensional vector.

2.2. Texture feature extraction

2.2.1. GLCM feature extraction

In 1973, Haralick et al. proposed a grey-level co-occurrence matrix for image classification [13]. And using GLCM to extract features require to steps: the first step was to calculate GLCM, and the second step was to calculate texture features using the calculated GLCM [11].

The grey level co-occurrence matrix of the image can reflect the integrated information of the direction of the image grey, the adjacent interval, and the amplitude change. Let $f(x, y)$ be a 2-bit digital image whose size is M×N, where Ng represents a grey level [12], then GLCM satisfies a certain spatial relationship, as shown in Equation (3):

$$P(x, y) = \#\{(x1, y1), (x2, y2) \in M \times N\} \qquad (3)$$

where the x in $\#(x)$ represents the number of elements in the set. Obviously, P is a Ng × Ng matrix. if the distance between $(x1, y1)$ and $(x2, y2)$ is d, and the angle between them and the horizontal axis coordinates is θ, then $P(i, j, d, \theta)$ and various inclinations and angles can be obtained. In this paper, we select four parameters of the co-occurrence matrix in four directions: contrast, correlation, energy, and uniformity to form a 1×4 vector.

2.2.2. LBP feature extraction

The local binary model was proposed by Ojala et al. It is an operator used to describe the local texture features of the image, and has significant advantages such as rotation invariance and grey scale invariance. It uses the local intensity of each pixel for feature vector extraction [13]. The initial LBP operator is defined as a 3×3 window. Using the centre pixel of the window as the threshold, the grey values of the adjacent eight pixels are compared with it. If the surrounding pixel value is greater than the value of the centre point, the pixel position is marked as 1, otherwise 0. In this way, 8 points in the 3×3 field can generate an 8-bit unsigned number. Then assign different weights based on their

positions and add them together to get an integer, which is the LBP value at that point. In addition, uniform and uneven images are distinguished based on the appearance. This paper uses a uniform local binary pattern to extract texture features.

2.2.3. NMI feature extraction

PCNN is widely used in image segmentation, but its parameters are numerous, and the number of iterations needs to be manually determined. For these situations, YD MA [14] proposed an automatic image segmentation method based on pulse coupled neural network and image entropy. However, this method is not ideal for low-contrast image segmentation. To solve this problem, this paper presents an automatic image segmentation method based on PCNN and 2OTSU, which is used to get the NMI features of the image and optimize the 2OTSU [15,16].

2.2.3.1. 2OTSU. The two-dimensional Otsu method is an efficient image segmentation method. It considers the grey information and spatial neighbouring information between pixels in the image at the same time. However, the computation of the optimal threshold vector for the two-dimensional Otsu method is very large. This paper uses Particle Swarm Optimization (PSO) to find the best two-dimensional threshold vector, where each particle represents a possible two-dimensional threshold vector and obtains the best two-dimensional threshold through the cooperation between the particles [17]. Experimental results show that the method not only can get the ideal segmentation result, but also can reasonably reduce the computational cost, so it is suitable for the real-time application.

2.2.3.2. PCNN. Pulse coupled neural network is a simplified neural network model constructed by Eckhorn in the 1990s based on the visual principle of cats. Compared with BP neural network and Kohonen neural network, PCNN does not require learning or training, and can extract effective information from a complex background. It has characteristics such as synchronous pulse distribution and global coupling [18]. And its signal form and processing mechanism are more in line with the physiological basis of the human visual nervous system [19]. Since the setting of PCNN parameters will greatly affect the segmentation effect, in order to reduce the complexity of the PCNN and increase the operating speed at the same time, a simplified PCNN model is generally used, and its formula is as follows:

$$F_{ij}(n) = I_{ij} \qquad (4)$$

$$L_{ij}(n) = \sum_{k,l} W_{kl} Y_{kl}(n-1) \qquad (5)$$
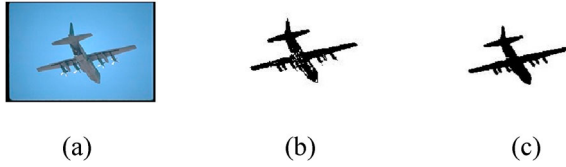
Figure 2. (a) Original (b) 2OTSU segment image (c) Optimized PCNN segment image.

$$U_{ij}(n) = F_{ij}(n)(1 + bL_{ij}(n)) \tag{6}$$

$$T_{ij}(n) = e^{-a}T_{ij}(n-1) + VY_{ij}(n) \tag{7}$$

$$Y_{ij}(n) = \begin{cases} 1, U_{ij}(n) > T_{ij}(n-1) \\ 0, \text{others} \end{cases} \tag{8}$$

where the parameters $W = \begin{pmatrix} \sqrt{1/2} & 1 & \sqrt{1/2} \\ 1 & 1 & 1 \\ \sqrt{1/2} & 1 & \sqrt{1/2} \end{pmatrix}$, $a = 0.1; b = 0.2; V = 0.2$.

In equation (4), L_{ij} and F_{ij} represent the original grey scale values of the image and the feedback domain input channels, respectively. In Equation (5), L_{ij} and Y_{kl} are the link domain channels and the pulse output, respectively; U_{ij} is expressed as the interior status signal of the bit in Equation (6). b is the strength coefficient of the link; T_{ij} in Equation (7) is expressed as a threshold threshold. In the PCNN model, b is variable. The larger value of b, the clearer the image is decomposed, and the more obvious the outline of the image. Conversely, the more detailed the image. The result as shown in Figure 2.

2.2.3.3. NMI feature. The NMI of the image has an invariant property, which is the normalized rotational inertia of the image has a good invariance to the geometric distortion (such as translation, rotation, scaling) [20]. The representation of the normalized moment of inertia is as follows:

$$i_c = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} iI_{ij}}{\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}}, j_c = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} jI_{ij}}{\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}} \tag{9}$$

$$\lambda = \frac{\sqrt{J_{i_c j_c}}}{m} = \frac{\sqrt{\sum_{i,j\in\Omega}\sum[(i-i_c)^2 + (j-j_c)^2]}}{\sum_{i,j\in\Omega}\sum Y_{ij}} \tag{10}$$

In Equation (9), (i_c, j_c) represents the point where the total mass of the image plane image is concentrated; in Equation (10), λ is the NMI eigenvalue, Y_{ij} is the binary image, and Ω is the region of $Y_{ij} = 1$ in the binary image. It can be seen that the eigenvalue λ of the NMI is the ratio of the moment of inertia of the binary image quality around its centre of gravity to its mass. For different binary images, different NMI features can be extracted, and NMI has the advantages of convenient extraction and low computational complexity compared to traditional image invariant features.

3. Proposed method

In this section, the study explains how to use the proposed method to establish a CBIR system. The proposed method has a third step: feature extraction, feature fusion and similarity measure [3].

3.1. Features extraction

The image is first processed using feature extraction. This paper uses HSV colour extraction, GLCM, LBP texture extraction and improved PCNN for NMI texture extraction.

In the HSV colour feature extraction stage, the image is converted to an HSV value, and the HSV image is converted into a histogram. This paper expresses the colour histogram vector as a 1×32 one-dimensional vector.

In the GLCM texture feature extraction phase, it is calculated respectively in four directions of 0°, 45°, 90° and 135° [3]. Subsequently, the four properties of the GLCM were used to calculate, including contrast, correlation, energy, and uniformity. The extraction process produces a 1×4 matrix.

In the LBP texture feature extraction phase, a 1×59 one-dimensional vector is generated during the extraction process.

In the NMI feature extraction phase, features are extracted using PCNN. However, the number of loop iterations of PCNN is N determined manually, so we introduce the optimized 2OTSU after particle swarm optimization into PCNN, and automatically get the number of iterations of loop. The length of the NMI feature vector is N.

3.2. Features fusion

After the four features have been extracted, they are weighted and fused. Because the physical meanings of colours and textures are different and cannot be directly compared, these feature vectors need to be internally and externally normalized prior to fusion. The Gaussian normalization method is used in this paper. Internal normalization means normalizing the components of the eigenvectors. External normalization is to normalize the feature distances to the same range of values to achieve the purpose of feature fusion [12]. In fusion, we assign weights based on the accuracy of the retrieved images, and then dynamically adjust the weights based on different types of images to obtain better fusion results. This paper assumes that HSV colour features, GLCM, LBP, and NMI weights respectively are ω_1, ω_2, ω_3 and ω_4. The weights of the features are shown as:

$$\omega_1 + \omega_2 + \omega_3 + \omega_4 = 1 \tag{11}$$

3.3. Similarity measure

The similarity measure is to measure the similarity between images. In this paper, when calculating similarity, the method used is Euclidean distance. The result of ED calculation is used as the parameter of image similarity. The smaller the ED value, the greater the similarity of the image. According to the result of sorting the ED value of the image, the ranking of the image is determined. X represents the feature vector of the image to be retrieved, y represents the feature vector of any image in the image database, and $i = 1, 2, 3, 4$, respectively, represent the four features of the image [10]. The Euclidean distance formula is as follows:

$$d_i = \sqrt{\left(\sum (x_i - y_i)^2 \right)} \qquad (12)$$

Similarity can be expressed as a formula:

$$s_i = \frac{1}{1 + d_i} \qquad (13)$$

The similarities of its four features are respectively s_1, s_2, s_3, s_4. The weight similarity formula is as follows:

$$s = \omega_1 s_1 + \omega_2 s_2 + \omega_3 s_3 + \omega_4 s_4 \qquad (14)$$

When similar images are retrieved, the weight is proportional to the accuracy; otherwise, it is inversely proportional.

4. Experimental results and analysis

The proposed method is verified on Corel-1k dataset, AT&T dataset and FD-XJ dataset, respectively.

4.1. Datasets

4.1.1. Corel-1k dataset

Corel-1000 Image Database: This database is used for various scientific papers on content-based image retrieval systems. The database contains 10 image classes, each containing 100 images. Courses include: Africa, beaches, buildings, racing, sunsets, flowers, elephants, marine life, aircraft and mountains. Each image has a size of 192×128 and the image format is JPEG [12].

4.1.2. AT&T dataset

Face database is used for face image retrieval. AT&T's face database contains 400 images of 40 topics. Each subject has 10 different facial expressions (open/closed eyes, smiling/not smiling) and facial details (eyeglasses/no glasses) images. These images were taken at different times and under different lighting conditions [13]. The size of each image in this database is 92×112.

4.1.3. FD-XJ dataset

The FD-XJ face image database is a Xinjiang ethnic minority face database created by the College of Information Science and Engineering of Xinjiang University. The database contains 3289 images of shoulders of Uygur and Kazak people in 253 Xinjiang in China. There are 13 images in each category. All images are collected in specific environments and include expressions, poses, accessories and lighting four variables conditions.

4.2. Results and analysis

This paper firstly obtains the single-retrieval precision of each feature for a class of images in each dataset, and compares the different capabilities of each feature of image retrieval. Then, when the fusion is performed, the weights of each feature are calculated using equations (11) to (14) according to the accuracy and fusion ratio of each single feature. And use the results of weights to fuse features. The weights in the three datasets of HSV, GLCM, LBP, PCNN, Improved PCNN, HSV + GLCM, HSV + GLCM + LBP, HSV + GLCM + PCNN, HSV + GLCM+ Improved NMI and our proposed method are shown in Tables 1–3.

This paper randomly selects an image from a class of image in three datasets as a query image. Subsequently, we select the first 5 images retrieved and evaluate their average accuracy and average recall rate [14]. The formulas for accuracy and recall are as

Table 1. Weights of feature fusion in Corel-1k dataset.

| Method | Corel-1k | | | |
	w1	w2	w3	w4
HSV	1	0	0	0
GLCM	0	1	0	0
LBP	0	0	1	0
PCNN	0	0	0	1
Improved PCNN	0	0	0	1
HSV + GLCM	0.47	0.53	0	0
HSV + GLCM + LBP	0.37	0.37	0.29	0
HSV + GLCM + PCNN	0.34	0.37	0	0.29
HSV + GLCM+ Improved PCNN	0.33	0.37	0	0.30
ours	0.26	0.29	0.22	0.23

Table 2. Weights of feature fusion in AT&T dataset.

| Method | AT&T | | | |
	w1	w2	w3	w4
HSV	1	0	0	0
GLCM	0	1	0	0
LBP	0	0	1	0
PCNN	0	0	0	1
Improved PCNN	0	0	0	1
HSV + GLCM	0.46	0.54	0	0
HSV + GLCM + LBP	0.28	0.33	0.39	0
HSV + GLCM + PCNN	0.30	0.35	0	0.35
HSV + GLCM + Improved PCNN	0.29	0.34	0	0.37
Ours	0.21	0.24	0.28	0.27

Table 3. Weights of feature fusion in FD-XJ dataset.

| Method	FD-XJ			
	w1	w2	w3	w4
HSV	1	0	0	0
GLCM	0	1	0	0
LBP	0	0	1	0
PCNN	0	0	0	1
Improved PCNN	0	0	0	1
HSV + GLCM	0.50	0.50	0	0
HSV + GLCM + LBP	0.31	0.32	0.37	0
HSV + GLCM + PCNN	0.31	0.32	0	0.37
HSV + GLCM + Improved PCNN	0.31	0.31	0	0.38
ours	0.23	0.23	0.26	0.28

follows:

$$P = \frac{(\text{Relenant images}) \cap (\text{Retrieved images})}{\text{Total number of images retrievevd}} \quad (15)$$

$$R = \frac{(\text{Relenant images}) \cap (\text{Retrieved images})}{\text{Total images of relevant in the database}} \quad (16)$$
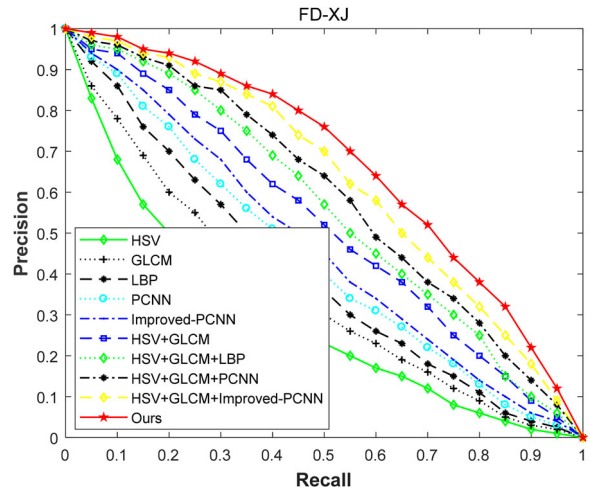
The features of the fusion are compared in three datasets. Their performance is displayed in the
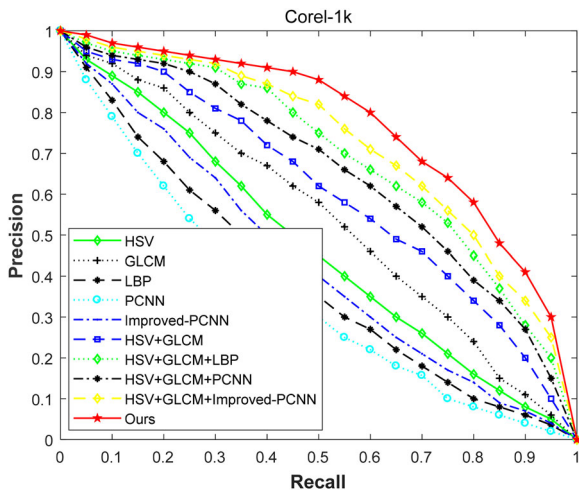
Figure 3. Precision-recall curve on Corel-1k.
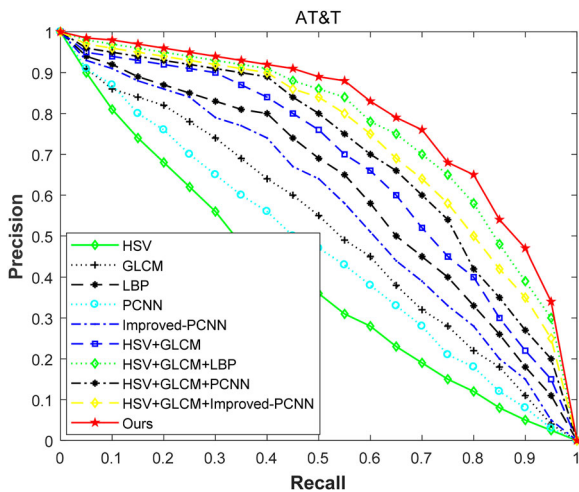
Figure 4. Precision-recall curve on AT&T.

Figure 5. Precision-recall curve on FD-XJ.

precision-recall curve, and the curve can reflect the overall index. The results are as shown in Figures 3–5.

The above three figures show the performance of each algorithm in the three datasets. From these, we can see that if the performance of a system is good, its curve should protrude as much as possible. In the experimental results, the image retrieval performance of a single feature is obviously worse than that of a multi-feature fusion image. And the performance of our proposed method is obviously better than that of commonly used multi-feature fusion methods.

Subsequently, in order to more clearly see the pros and cons of their performance, we give objective evaluation indicators for each system on three data sets.

Table 4. Objective evaluation indicators of system in Corel-1k dataset.

Method	Corel-1k			
	Precision	Pre-recall	F-measure	MAP
HSV	0.525	0.434	0.475	0.485
GLCM	0.561	0.515	0.537	0.537
LBP	0.445	0.411	0.427	0.423
PCNN	0.436	0.353	0.390	0.414
Improved PCNN	0.451	0.449	0.450	0.438
HSV + GLCM	0.623	0.564	0.592	0.602
HSV + GLCM + LBP	0.634	0.633	0.633	0.614
HSV + GLCM + PCNN	0.662	0.547	0.599	0.651
HSV + GLCM + Improved PCNN	0.674	0.595	0.632	0.651
Ours	0.697	0.691	0.694	0.684

Table 5. Objective evaluation indicators of system in AT&T dataset.

| Method	AT&T			
	Precision	Pre-recall	F-measure	MAP
HSV	0.463	0.395	0.426	0.442
GLCM	0.531	0.516	0.523	0.511
LBP	0.626	0.567	0.595	0.598
PCNN	0.523	0.431	0.473	0.503
Improved PCNN	0.594	0.538	0.565	0.564
HSV + GLCM	0.644	0.613	0.628	0.621
HSV + GLCM + LBP	0.724	0.676	0.699	0.706
HSV + GLCM + PCNN	0.709	0.591	0.645	0.677
HSV + GLCM + Improved PCNN	0.715	0.629	0.669	0.682
Ours	0.746	0.709	0.727	0.723

Table 6. Objective evaluation indicators of system in FD-XJ dataset.

Method	FD-XJ			
	Precision	Pre-recall	*F*-measure	MAP
HSV	0.406	0.278	0.330	0.379
GLCM	0.414	0.376	0.394	0.386
LBP	0.458	0.392	0.422	0.441
PCNN	0.479	0.426	0.451	0.453
Improved PCNN	0.481	0.469	0.475	0.461
HSV + GLCM	0.532	0.490	0.510	0.512
HSV + GLCM + LBP	0.547	0.6516	0.531	0.525
HSV + GLCM + PCNN	0.564	0.580	0.572	0.547
HSV + GLCM + Improved PCNN	0.596	0.518	0.554	0.579
Ours	0.629	0.618	0.618	0.603

Objective evaluation indicators include: accuracy (*P*), recall rate (*R*), *F*-measure, and MAP. The meaning of accuracy rate and recall rate has been given before. *F*-measure refers to the comprehensive evaluation index, which is a weighted harmonic average of *P* and *R*. The higher the value of *F*, the better the performance is proved; Map is to solve the limitations of *P*, *R*, and *F*-measure single point values. The higher the value of Map, the better of the performance. The results are shown in Tables 4–6.

Where, as can be seen from these tables, the value of each objective index of our proposed method is greater than that of other methods, which indicates that our proposed method performance is better than that of other systems.

Following, we compare the subjective evaluation indicators of several algorithms in three datasets. This paper randomly selects an image (a) as a query image from a class of image in each dataset, then the GLCM feature, HSV + GLCM features, LBP feature and the feature proposed by us are respectively used to perform image retrieval in the datasets, selecting the first 5 images retrieved the images are compared and the results are as follows.

In Figure 6, the other three methods each have an image does not belong to the same kind of image with the query image. The images retrieved by the method proposed in this paper belong to the same type of image with the query image; in Figure 7, the images retrieved by the GLCM feature and the HSV + GLCM features do not belong to the same type of image as the query image. The image retrieved by the method proposed in this paper belongs to the same category as the query image;
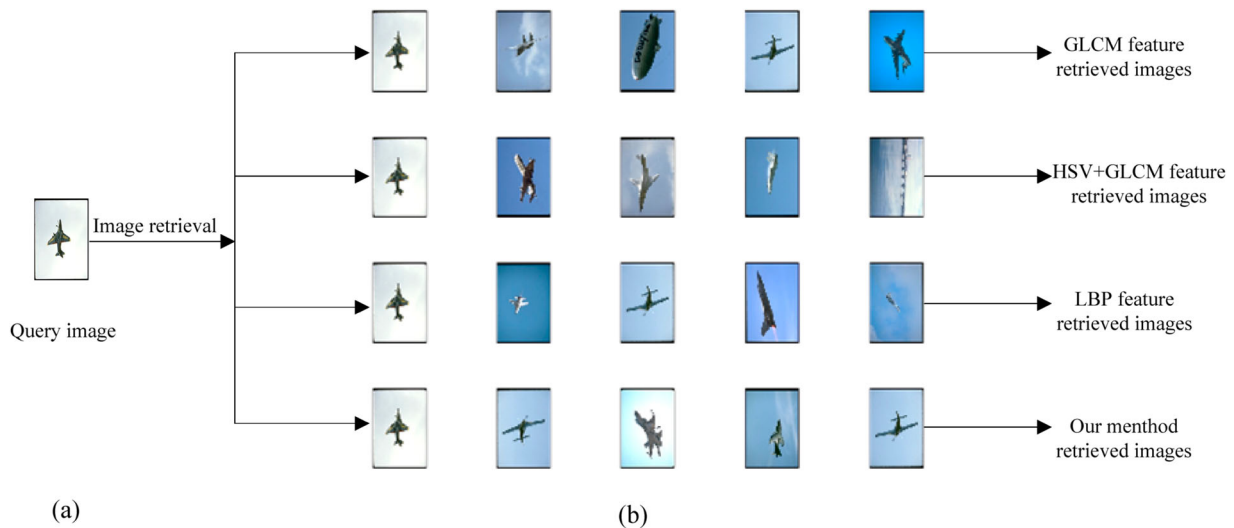
Figure 6. (a) Query image (b) images retrieved in Corel-1k.

Figure 7. (a) Query image (b) images retrieved in AT&T.

GLCM feature
retrieved images

HSV+GLCM feature
retrieved images

LBP feature
retrieved images

Our method
retrieved images

Query image

Image retrieval

(a)   (b)

Figure 8. (a) Query image (b) images retrieved in FD-XJ.

in Figure 8, in the first method, having two retrieved images and query images do not belong to one type. In the second and third methods, have one retrieval image separately and query image do not belong to the same type of image, however, the images retrieved by the method proposed in this paper and query image belong to the same type of image.

5. Conclusion

This paper presents a novel multi-feature fusion image retrieval method. The purpose of this method is to complete the image retrieval through the fusion of four features. The method is based on three steps, namely multi-feature extraction, dynamic retrieval of weights and similarity fusion [21]. In the feature extraction process, the number of loop iterations of the traditional NMI feature is difficult to determine. Therefore, this paper proposed an improved PCNN to extract NMI feature. That is, 2OTSU uses particle swarm optimization algorithm to optimize, and the optimized algorithm is introduced into PCNN to extract NMI feature. It is used to automatically obtain the number of loop iterations to extract the NMI features of the image. The advantages of NMI features are simple calculation, high accuracy, resistance to geometric distortion, and robustness to image operations such as filtering, smoothing, and compression. The proposed method was compared with HSV, GLCM, LBP, PCNN, Improved PCNN, HSV + GLCM, HSV + GLCM + LBP, HSV + GLCM + PCNN and HSV + GLCM + Improved PCNN. All methods were tested on Corel-1k dataset, AT&T dataset and FD-XJ face image dataset. From the theoretical and experimental results it can be concluded that image retrieval based on this method provides excellent retrieval accuracy. In future work, we would like to investigate more complex schemes to determine the weight of similarity fusion parts. It will also be our goal to reduce

the computational complexity through indexing techniques and to improve the framework that is more suitable for large datasets.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Mohamadzadeh S, Farsi H. Content-based image retrieval system via sparse representation. IET Comput Vision. 2016;10(1):95–102.
[2] Qazanfari H, Hassanpour H, Qazanfari K. A short-term learning framework based on relevance feedback for content-based image retrieval. 2017 3rd Iranian Conference on Intelligent systems and signal processing (ICSPIS), Shahrood; 2017. p. 136–140.
[3] Su C, Chiu H, Hsieh T. An efficient image retrieval based on HSV color space. 2011 International Conference on Electrical and Control Engineering, Yichang; 2011. p. 5746–5749.
[4] Sulochana S, Vidhya R. Texture based image retrieval using framelet transform-gray level co-occurrence matrix (GLCM). Int J Adv Res Artifi Intel. 2013;2(2): 68–73.
[5] Choobari BM, Mozaffari S. A robust content based image retrieval using local full-directional pattern (LFDP). 2017 5th Iranian Joint Congress on Fuzzy and Intelligent systems (CFIS), Qazvin; 2017. p. 178–183.
[6] Putri RD, Prabawa HW, Wihardi Y. Color and texture features extraction on content-based image retrieval. 2017 3rd International Conference on Science in information technology (ICSITech), Bandung; 2017. p. 711–715.
[7] Jain S, Zaveri T, Patel S. Weighted fusion of LBP and CCH features for effective content based image

retrieval. 2016 International Conference on signal processing and Communications (SPCOM), Bangalore; 2016. p. 1–5.

[8] Liu P, Guo JM, Chamnongthai K, et al. Fusion of color histogram and LBP-based features for texture image retrieval and classification. Inf Sci. 2017;390:95–111.

[9] Liu Q, Xu L-P, Y-D MA, et al. Image NMI feature extraction and retrieval method based on pulse coupled neural networks. Acta Autom Sin. 2010;36(7): 931–938.

[10] Srivastava D, Goel S, Agarwal S. Pipelined technique for image retrieval using texture and color. 2017 4th International Conference on Power, Control & Embedded systems (ICPCES), Allahabad; 2017. p. 1–6.

[11] Wang Y, Sui A, Fu W. Research on image retrieval technology based on image fingerprint and color features. 2016 International Conference on Audio, Language and image processing (ICALIP), Shanghai; 2016. p. 447–451.

[12] Wenfei D, Shuchun Y, Songyu L, et al. Image retrieval based on multi-feature fusion. 2014 Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control, Harbin; 2014. p. 240–243.

[13] Verma M, Raman B. Local tri-directional patterns: a new texture feature descriptor for image retrieval. Digit Signal Process. 2016;51:62–72.

[14] Ma Y-D, Liu Q, Qian Z. Automated image segmentation using improved PCNN model based on cross-entropy. Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong; 2004. p. 743–746.

[15] Cheng S, Wang L, Qin J, et al. Image segmentation algorithm based on fusion of group intelligent algorithm optimized OTSU-entropy and pulse coupled neural network. J Comput Appl. 2017;37(12): 3528–3535.

[16] Jieyu Z, Jindan H. An image segmentation algorithm research based on optimized PCNN. 2018 11th International Conference on Intelligent computation technology and Automation (ICICTA), Changsha; 2018. p. 86–90.

[17] Tang Y-G, Dong L, Guan X-P. Fast image segmentation based on particle swarm optimization and two-dimension Otsu method. Control Decis. 2007;22(2): 202–205.

[18] Guo X, Zhang M, Dai Y. Image of Plant Disease segmentation model based on pulse coupled neural network with Shuffle Frog Leap algorithm. 2018 14th International Conference on Computational Intelligence and Security (CIS), Hangzhou; 2018. p. 169–173.

[19] Zhou D, Zhou H, Gao C, et al. Simplified parameters model of PCNN and its application to image segmentation. Pattern Anal Appl. 2016;19(4):939–951.

[20] Liu C, Cui M, Yang Q, et al. Study on method for image retrieval based on feature fusion. 2011 4th IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, Beijing; 2011. p. 637–640.

[21] Chen Q, Ding Y, Li H, et al. A novel multi-feature fusion and sparse coding-based framework for image retrieval. 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), San Diego, CA; 2014. p. 2391–2396.