Loughborough University Institutional Repository

A fast image retrieval method designed for network big data

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: YANG, J. ...et al., 2017. A fast image retrieval method designed for network big data. IEEE Transactions on Industrial Informatics, In Press.

Additional Information:

Metadata Record: https://dspace.lboro.ac.uk/2134/24204

Version: Accepted for publication

Publisher: © IEEE

Please cite the published version.

A Fast Image Retrieval Method Designed for Network Big Data

Jiachen Yang, *Member, IEEE,* Bin Jiang, Baihua Li, Kun Tian and Zhihan Lv, *Member, IEEE,*

Abstract—In the field of big data applications, image information is widely used. The value density of information utilization in big data is very low, and how to extract useful information quickly is very important. So we should transform the unstructured image data source into a form that can be analyzed. In this paper, we proposed a fast image retrieval method which designed for big data. First of all, the feature extraction method is necessary and the feature vectors can be obtained for every image. Then, it is the most important step for us to encode the image feature vectors and make them into database, which can optimize the feature structure. Finally, the corresponding similarity matching is used to determined the retrieval results. There are three main contributions for image retrieval in this paper. New feature extraction method, reasonable elements ranking and appropriate distance metric can improve the algorithm performance. Experiments show that our method has a great improvement in the effective performance of feature extraction and can also get better search matching results.

Index Terms—Big data, image retrieval, feature ranking, distance learning.

I. INTRODUCTION

FA ced with the big media data, we need to organize and manage these semi-structured and even unstructured data effectively [1]–[4]. With the development of information retrieval technology [5]–[7] and commercial search technology [8] in recent years, database indexing and text retrieval techniques has become a general pattern. However, the present image retrieval performance still cannot meet the requirement of expectations. Compared with text retrieval, image retrieval is still in exploratory stage [9]–[11].

At present, there are two main types of image retrieval: text based image retrieval (TBIR) and content-based image retrieval (CBIR). Text based image retrieval makes use of artificial marks to avoid visual analysis. Through the searching and matching for texts, text based image retrieval establishes the correlation link between images [12], [13]. However, with the increase in size of the image database, the limitations of the method based on artificial annotation are also shown. Manual

annotation requires two much manpower and time consumption [14]. when the size of the image database increases to a certain extent, manual labeling for each image will become impossible to achieve. At the same time, due to the humans' understanding to image content often has a strong subjectivity, the description for the same image will be different in manual annotation [15]–[17].

In order to overcome the shortcomings of the methods based on manual annotation, researchers put forward the content based image retrieval [18], [19]. Different from the text based image retrieval, content based image retrieval directly gets visual vectors of the images to find out the similar characteristics. According to the difference in the feature scale, it is divided into the method based on global features and the method based on local features. Content based image retrieval can extract the relevant visual statistical characteristics such as colors and textures. According to the distance measurement between the feature vectors, images can be matched with each other [20]–[22].

However, global features are susceptible to the background and noise, especially when the object is partially occluded or there is geometric distortion. It is difficult to establish effective image matching. In addition, local features are irrelevant with its scale, rotation and translation, which can be used to establish local similarity matching methods, even with complex background or geometric distortion. In recent years, local features draw more attentions [23]. In order to fully express the visual content of images, an image may need thousands of local features, which will increase the image similarity matching process time. It is important to extract more effective features and design corresponding index mechanism to reduce the processing time [24]–[26].

In the research on large-scale image retrieval, the local feature vectors are mapped. In order to build image correlation matching, the similarity distance between images should be measured. Content based image retrieval achieve the image quantization based linear expression. In general, the inverted index mechanism significantly increases the efficiency of the large scale images matching [27]–[29]. However, it will reduce the visual discrimination ability of local features. Because of the ambiguity of quantization errors, unrelated images may also share the same visual words and thus causing the errors in image matching. This situation is becoming increasingly apparent and it greatly affects the accuracy of retrieval results. In order to ensure the efficiency of the existing quantitative method and further improve the accuracy of quantitative definition, more efficient image representation models need to be put forward [30].

In this paper, we proposed a fast image retrieval method designed for big data. Considering the local features and ranking the vectors, we can get a more accurate retrieval results. The main contributions of this paper are as followed.

1) First of all, this paper is innovative in dealing with the image features. When images are represented as feature vectors, the weight relationship between the different elements will affect the accuracy of retrieval. In this paper, the detailed work was done, and the feature vectors were sorted according to the contribution level, which could make the best use of the effective information.

2) It is well known that the image recognition based on deep learning will take a lot of processing time, and the different image database will affect the matching results. In this paper, based on the traditional distance matching algorithm, the matching levels of retrieved images were calculated. This method can not only guarantee the accuracy of the retrieval system, but also save a lot of retrieval time.

3) In this paper, we improve feature extraction methods for image retrieval. Gray co-occurrence matrix, color co-occurrence matrix and difference between pixels of scan pattern are used as the foundation. In addition, local binary pattern and histogram of oriented gradient are firstly to be used with other features. The increase in the number of features can help us to improve the performance of the retrieval algorithm. It also provides enough basic data for feature dimension reduction.

The rest of this paper is organized as followed. Section II provides the background and motivation for feature extraction. Section III formulates the fast image retrieval problem and presents the proposed approach. In Section IV and V, we show the experimental results and discuss the comparison. Finally, conclusions are drawn in Section VI.

II. BACKGROUND AND MOTIVATION

As a new research filed, the content based image retrieval has a great commercial potential. Visual features are very important for images, which are based on the prior knowledge of the image content. In addition, visual features are closely related with the specific application areas. Based on content-based image retrieval technology, visual features can be extracted automatically or semi automatically to improve the robustness of the system and ensure the quality of image retrieval.

The characteristic information of images can be classified into two categories. Low level feature attribute contains color feature, texture feature and shape feature. Semantic feature attribute contains interpretation and identification of objects. However, due to the current development of computer vision and image understanding, the image retrieval technique is still unable to realize semantic image attribute. Key of image retrieval is still extracting low-level feature attribute information.

Shape features are the most significant features for images. At the same time, it is also the most difficult to describe. For humans, it is easy to identify the objects in dominant position. For computers, they do not the ability. And shape features can help [31]. In the field of image retrieval, methods for shape feature extraction can be divided into two directions: extraction based on edge attributes and extraction based on region shape features [32]. The main idea of extraction for edge shape features is based on the principle of edge detection. Edge shape features are applied mainly in images whose edge is clear and separated. Perimeter, angle, eccentricity and points of interest are used as descriptors. Fourier descriptor, invariant moment, autoregressive model, the central chain code and the curvature scale space are often used in the analysis methods [33], [34].

Color attributes are the most simple image features, which can be used in the field of image retrieval [35]. The common retrieval methods can be divided into color histogram, color correlogram, color coherence vector, color matrix and so on. Through color channel decomposition for images, color histogram is used for calculating and comparing similarity between two images. In the content based image retrieval technology, color correlogram is more effective than the traditional color histogram. In other words, color histogram just simply reflects the proportion of different values in every color channel. However, the color correlogram not only reflects the global statistical relationship, but also expresses the spatial transformation between color and distance. In addition, color matrix mainly uses the mean, variance and skewness as the color feature information of images. Color matrix retrieval method has been proved to be representative as the distribution of color information.

Texture features are mainly extracted to describe the spatial relationship between the pixels in local area. Generally speaking, the methods for texture features extraction can be divided into the following four categories: description based on the statistical method, description based on structure method, description based on spectrum method and description based on model method [36]. Specifically speaking, feature description based on statistical method mainly considers distribution on pixel values. The most common statistical methods often describe the texture characteristics based on the gray histogram. In order to make full use of the spatial distribution on pixel values, co-occurrence matrix was proposed to describe the texture feature. The method considers the information on distance and direction between pixels. By constructing a co-occurrence matrix, correlation information can be used as a component of image feature vectors. In addition, texture features based on human visual perception was put forward including the following attributes: contrast, likeness, roughness and so on. Description based on structure method considers regular spatial structure arrangement.

Obviously, it will consume a large amount of time and storage space to find similarity for every image pairs in a database with too much images. In order to improve the retrieval efficiency, information dimensionality must be reduced as much as possible. With the explosive development of big data, image size has gradually become larger. Therefore, image dimension reduction processing is the premise and foundation of the content based image retrieval.

Different distance calculation methods can be used in image retrieval. For example, Manhattan distance is often used. This method is simple and easy to understand. The relationship
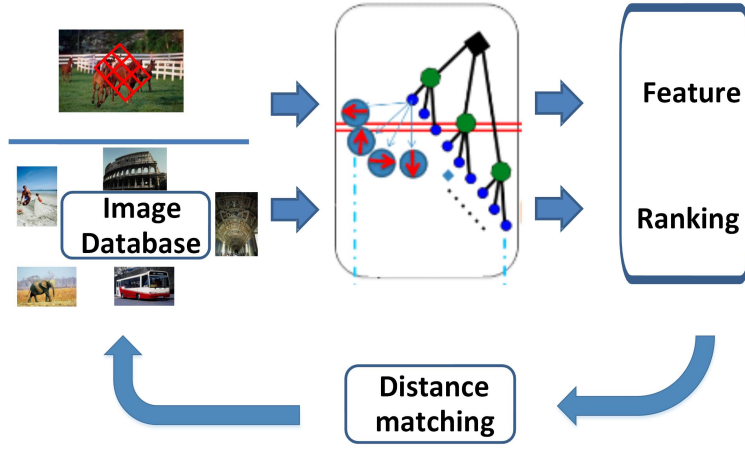
Fig. 1. Proposed content-based image retrieval method, this method is divided into three steps. Specific image features (GCM,CCM,DBPSP,LBP and HoG) are extracted and dimension reduction is necessary.

between the two feature vectors can be computed as Equ.1.

$$d(q, x_j) = \sum_{i=1}^{i=m} |x_{ij} - q_i| \qquad (1)$$

Euclidian distance is another method and the results can be obtained as Equ.2.

$$d(q, x_j) = \sqrt{\sum_{i=1}^{i=m} (x_{ij} - q_i)^2} \qquad (2)$$

In addition, cross correlation, minimum mean distance, statistical measure are also can be used in this problem, as shown in Equ.3, 4 and 5.

$$r(q, x_j) = \frac{\sum_{i=1}^{i=m} x_{ij} q_i}{\left| \sum_{i=1}^{i=m} (x_{ij})^2 \right|^{\frac{1}{2}} \left| \sum_{i=1}^{i=m} (q_i)^2 \right|^{\frac{1}{2}}} \qquad (3)$$

$$z_j = \frac{\sum_{i=1}^{i=m} x_{ij}}{m} \qquad (4)$$

$$s_j = \sum_{i=1}^{i=m} \left[\frac{Max(x_{ij}, q_i)}{Min(x_{ij}, q_i)} - 1 \right]^2 \qquad (5)$$

III. IMAGE RETRIEVAL WITH DIMENSION REDUCTION

Based on the above content, this paper proposed an image retrieval method. This method is divided into three steps, as shown in Fig.1. First, we extracted specific image features. The retrieval time was increased due to the large number of extracted features, and it was necessary to make feature dimension reduction as the second step. At last, new distance theory was put forward for image analysis and ultimately we got the corresponding retrieval images.

Gray co-occurrence matrix is actually obtained through the probability ($P(i, j)$) of gray co-occurrence pair($g_{k,l}, g_{m,n}$) in four directions (θ) at different distances (d) to describe the texture feature. Given an offset distance for gray image, there will be four gray co-occurrence matrixes. In addition, the size of the gray co-occurrence matrix is $Q * Q$, and Q is the maximum value of the image. So the element value of gray co-occurrence matrix is determined by number of pixels in certain offset distance and angle.

$$P(i, j, d, 45^o) = ((k, l), (m, n)) \in (M_x \times M_y) \times (M_x \times M_y) \qquad (6)$$

$$|k - m| = d, |l - n| = d, g_{k,l} = i, g_{m,n} = j \qquad (7)$$

In theory, the gray co-occurrence matrix is a two-dimensional matrix in 256×256. In this way, the quantization scale and computation cost of gray co-occurrence matrix is relatively large. In practice, we can quantify gray value by classification in different level $N(N = 8, 16, 32, 64, ...)$, so gray co-occurrence matrix can be used in the actual image processing. In this paper, N is set to be 8.

In addition, gray co-occurrence matrix considers the image variation information on direction, spacing and amplitude, but it does not directly provide different texture feature value. We must make a quantitative description on smooth texture, roughness, contrast and similar information. In this paper, four other statistical characteristic are used to solve this problem. We mainly use the angular second moment, contrast, correlation and inverse difference, which give a more accurate feature extraction on images.

To reflect the texture of the local image features, we can use a 3×3 window scaning in the image with four directions respectively, the statistics of the adjacent four pixel value changes, as shown in Fig.2 and 3. All features can be divided into seven types, in accordance with the frequency of values, so we can get a 7×7 two-dimensional matrix, known as CCM [37].

$$M_i(u, v) = M_i(u, v | \delta_x, \delta_y)$$
$$= M_i(P_i[x, y], P_i[x + \delta_x, y + \delta_y]) \qquad (8)$$

$$m_i(u, v) = \frac{M_i(u, v)}{N_i} \qquad (9)$$

Fig. 2. 3×3 window for color co-occurrence matrix (CCM)

Fig. 4. Rotation invariant local binary pattern

Fig. 3. 7 motifs color co-occurrence matrix (CCM)

$$N_i = \sum_{u=0}^{6} \sum_{v=0}^{6} M_i(u,v) \tag{10}$$

Color co-occurrence matrix (CCM) can only detect the direction of local texture in images, but can not carefully detect the complexity of textures. Difference between pixels of scan pattern (DBPSP) can be used to solve the problem.

$$d^m = |P_m - P_{m+1}| + |P_{m+1} - P_{m+3}| + |P_{m+3} - P_{m+2}| \tag{11}$$

$$AR_i = \frac{1}{N_i} \sum_{N_i}^{j} d_i^j(x,y) \tag{12}$$

Local binary pattern (LBP) is a kind of operator used to describe the local texture feature of images. It has remarkable advantages such as rotation invariance and gray scale invariance. At first, LBP is proposed used for texture feature extraction. Moreover, the extracted features are local textures in images [38].

Original LBP operator is defined with the 3×3 window, gray value center pixel is compared with adjacent pixels, if the surrounding pixel value is greater than the center pixel value, the pixel is labeled as 1, otherwise it is labeled as 0. So, 8-bit binary number (a total of 256 codes) is produced with 3×3 neighborhood with eight points , that is to be the LBP value of window center pixel, and the value can be used to reflect the texture information.In this paper, we make use of the rotation invariant LBP operator as shown in Fig.4.

Histogram of Oriented Gradient (HoG) is used for object detection in computer vision and image processing as a feature descriptor [39]. It is composed of the gradient direction histograms in local areas. In addition, hog feature can be combined with support vector machine as the classifier in image recognition [40], [41].

The retrieval time will be increased due to the large number of features extracted, it is necessary to make feature dimension reduction as the second step. In this step, maximal mutual information criterion is used as the dimensionality reduction algorithm [42].

$$I(x,y) = \sum_{i,j} p(x_i, y_j) log \frac{p(x_i, y_j)}{p(x_i)p(y_j)} \tag{13}$$

$$maxV = \frac{1}{|S|^2} \sum_{i \in S} I(c, x_i) \tag{14}$$

$$minW = \frac{1}{|S|^2} \sum_{i,j \in S} I(x_i, y_j) \tag{15}$$

$$max\phi = (V - W) \tag{16}$$

By this method, we can sort the image features according to the importance. There are many other calculation method for

the distance between two vectors, weighted Euclidian distance is calculated by Equ.17, 18 and 19.

$$d(q, x_j) = \sqrt{\sum_{i=1}^{i=m} \rho_i (x_{ij} - q_i)^2} \qquad (17)$$

$$\rho_i = \frac{n}{\sum_{j=1}^{n} (x_{ij} - \bar{z}_i)^2} \qquad (18)$$

$$\bar{z}_i = \frac{\sum_{n}^{j=1} x_{ij}}{n} \qquad (19)$$

Based on the distance between images, different calculation on image pairs can be used for the retrieval results.

IV. PERFORMANCE EVALUATION AND DATABASES

A. Performance evaluation

It is very important to evaluate the advantages and disadvantages of the image retrieval system. In recent years, a number of indicators to measure retrieval methods have been proposed. In this paper, we use the most common precision and recall to carry out test analysis [43]. Precision refers to the ratio between retrieval similar images number TP with the retrieval total images number TP + FP , which is used to measure the precision of image retrieval system. Recall refers to the ratio between retrieval similarity image number TP in the database with all the total number TP + FN, which is used to measure the recall of image retrieval system. These two indicators can be calculated by Equ.20 and 21.

$$precision = \frac{TP}{TP + FP} \qquad (20)$$

$$recall = \frac{TP}{TP + FN} \qquad (21)$$

In other words, the precision is the ratio between the retrieved images and all the retrieved images. The recall is the ratio between the retrieved relevant images and all the related images in the image database. Generally speaking, the more relevant, the better in feedback of retrieved images.

However, precision and recall are mutually restricted. For example, there are 1000 images in the image database, and only 100 images are relevant with the query image.If the first retrieved 10 images are correct, the precision is 1, but the recall is only 0.1. At the other extreme, if all the 1000 images in the database have been retrieved, recall is 1, but precision is only 0.1. In this paper, the precision-recall curve analysis is used to assess the various retrieval methods, on the other hand, calculation of comprehensive evaluation index is used to evaluate the performance of the analysis.

B. Databases

In this section, we will investigate the retrieval performance both with the proposed model and the other models for comparison. In addition, experiments were conducted on three sets of image database, which proved the validity of the proposed method.

Fig. 6.   Samples for Oxford Flowers database

1) WANG dataset: As the first experiment in this paper, WANG dataset has 1000 images [44], [45]. There are 10 kinds of these images, each of which is 100. The images in Fig.5 are the representation of the each categry. The database can be downloaded in http://wang.ist.psu.edu/docs/related/. In addition, these 1000 images can be divided into two parts, 900 images are used as training images, and the remaining 100 are used as test images. Of course, the proportion for each type of image is the same.

2) Oxford Flowers dataset: Oxford Flowers dataset is used for the second experiment [46]. In this database, the researchers chose a number of very common flowers in the UK, a total of 17 categories. For each category, 80 representatives of the picture are chosen. Some samples are shown in Fig.6. As similar with first experiment, pictures will also be divided into two categories: training and testing, in which the 60 images are used to train, and the other 20 images are used to test.

3) CIFAR-10 dataset: As the third experiment, CIFAR-10 dataset [47] contains 60,000 images, which are labeled subsets of the 80 million tiny images dataset. 50000 images are used to train and the other 10000 images are used to test.

V. EXPERIMENTAL RESULTS

A. Compared with the results of four models

To demonstrate the advantages of this method, we introduce some other methods: Jhanwar [48], Huang [49], Lin [50] and ElAlami [31], [51]. Jhanwar [48] present MCM (motif co-occurence matrix) as a feature for CBIR (contented based images retrieval), the MCM is used to measure the distance between images information in both color and texture. Huang [49] put forward a contented based images retrieval using the similarity between image textures. Sub-band gradient vectors are used as feature to combine the texture similarity. Lin [50] combine three features for contented based images retrieval. Color co-occurrence matrix is useful for measuring the similarity between images. Elalami [31], [51] presented a model depending the most relevant features.

The results using different retrieval models obtained on the Wang database are listed in the Tables II and III. And the

TABLE I
DEFINITION OF PRECISION AND RECALL

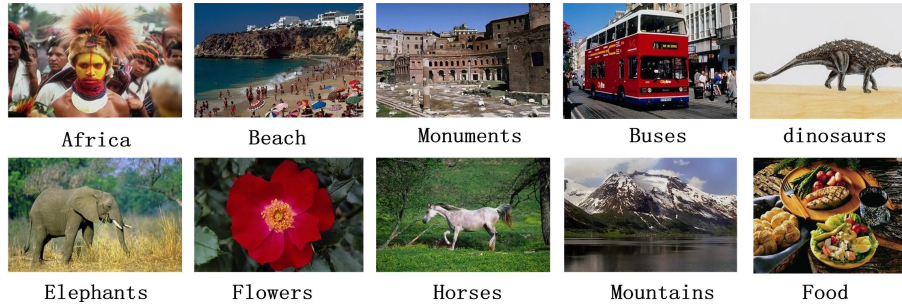|	Relevant	Irrelevant	Total
Retrieved	TP	FP	Predicted Positive
No retrieved	FN	TN	Predicted Negative
Total	Actual Positive	Actual Negative	TP+FP+FN+TN

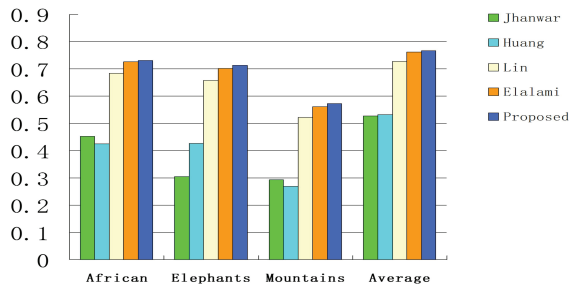Fig. 5. Samples for Wang database

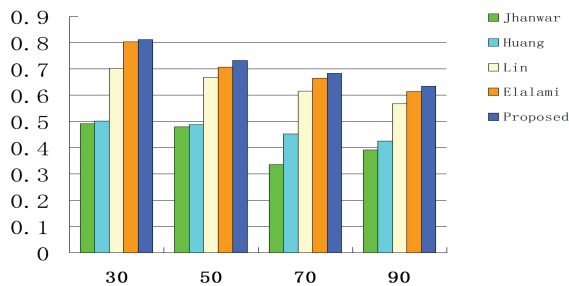
Fig. 7. Specific comparisons in Wang database.

Fig. 8. Specific comparisons in Oxford Flowers database.

results obtained on the Oxford Flowers database are listed in the Tables IV and V. In additions, Fig.7 and 8 show the specific comparisons. It can be seen from the experimental results that our method can get higher precision and recall in most of the retrieval process.

B. Comparison of the retrieval time of the different models

The time consumption of image retrieval process is an important index in retrieval system. Three databases are tested used in this paper respectively using different five methods. Finally the results obtained are shown in the Tables VI, VII

and VIII. It can be seen from the experimental results that our method can get result quickly in most of the retrieval process.

Usually, the retrieval time and the number of feature extraction are reverse. The more features, the higher the accuracy of retrieval, but it will be a waste of time cost. On the contrary, the time cost will be reduced. Experiments show that our method can find a good balance between the retrieval accuracy and the time cost.

In this part of the method, a large number of features are extracted, but the contribution of each feature to the retrieval result is different. As an innovation point, the importance of feature ordering is used to reduce the feature dimension, and to reduce the retrieval time as far as possible.

The proposed features based on co-occurrence information quantitative characteristic acceleration can get the best retrieval performance results , which proves the co-occurrence information between visual features not only to guarantee the effective matching of similarity between images, have also contributed to improve the efficiency of quantitative character, especially when asked to quantify the time overhead is smaller. The acceleration and optimization effect is more obvious.

C. The relationship between precision and number of feature

Compared with other methods, the proposed method extracts many kinds of features, which will have a negative impact on the retrieval time. So it is necessary to discuss the impact of the number of features on the whole algorithm. Through the arrangement of this method, the features will be sorted based on their weight. According to this order, we use the extracted features from less to more, and observe the change trend of the retrieval rate. The experimental results are shown in Fig.9.

Using different number of features, we obtained precision for the statistics under the conditions of different searching number. Experiments show that, no matter how the retrieval situation changes, if the number of features increase to a certain extent, it can not improve the results when continuing to increase feature number .

TABLE II
THE COMPARISON OF AVERAGE PRECISION FOR DIFFERENT MODELS APPLIED IN WANG DATABASE

	African	Beach	Monuments	Buses	Food	Dinosaurs	Elephants	Flowers	Horses	Mountains	Average
Jhanwar [48]	0.453	0.398	0.374	0.741	0.369	0.915	0.304	0.852	0.568	0.293	0.527
Huang [49]	0.424	0.446	0.411	0.852	0.427	0.587	0.426	0.898	0.589	0.268	0.533
Lin [50]	0.683	0.540	0.562	0.888	0.733	0.992	0.658	0.891	0.803	0.522	0.727
Elalami(2011) [31]	0.703	0.561	0.571	0.876	0.741	0.987	0.675	0.914	0.834	0.536	0.739
Elalami(2014) [51]	0.726	0.593	0.587	**0.891**	0.772	**0.993**	0.702	**0.928**	0.856	0.562	0.761
Proposed	**0.730**	**0.612**	**0.597**	0.890	**0.781**	0.990	**0.713**	0.916	**0.862**	**0.573**	**0.766**

TABLE III
THE COMPARISON OF AVERAGE RECALL FOR DIFFERENT MODELS APPLIED IN WANG DATABASE

	African	Beach	Monuments	Buses	Food	Dinosaurs	Elephants	Flowers	Horses	Mountains	Average
Jhanwar [48]	0.115	0.121	0.127	0.092	0.129	0.072	0.132	0.087	0.102	0.135	0.111
Huang [49]	0.126	0.113	0.132	0.099	0.122	0.104	0.119	0.093	0.103	0.152	0.116
Lin [50]	0.141	0.192	0.174	0.121	0.132	0.101	0.149	0.112	0.134	0.213	0.146
Elalami(2011) [31]	0.153	0.198	0.182	0.116	0.138	0.098	0.156	0.118	0.139	0.228	0.152
Elalami(2014) [51]	0.161	0.203	0.191	0.126	**0.148**	0.109	**0.163**	0.129	0.144	**0.236**	0.161
Proposed	**0.173**	**0.221**	**0.197**	**0.132**	0.135	**0.113**	0.159	**0.134**	**0.146**	0.229	**0.164**

TABLE IV
THE COMPARISON OF AVERAGE PRECISION FOR DIFFERENT MODELS APPLIED IN OXFORD FLOWERS DATABASE

	20	30	40	50	60	70	80	90	100
Jhanwar [48]	0.603	0.491	0.484	0.479	0.458	0.336	0.427	0.391	0.372
Huang [49]	0.621	0.502	0.491	0.488	0.462	0.453	0.442	0.426	0.391
Lin [50]	0.764	0.701	0.682	0.668	0.643	0.616	0.597	0.568	0.531
Elalami(2011) [31]	0.793	0.765	0.742	0.687	0.656	0.632	0.605	0.583	0.556
Elalami(2014) [51]	0.821	0.804	0.767	0.706	0.687	0.665	0.642	0.614	0.589
Proposed	**0.830**	**0.812**	**0.797**	**0.732**	**0.703**	**0.683**	**0.659**	**0.634**	**0.596**

TABLE V
THE COMPARISON OF AVERAGE RECALL FOR DIFFERENT MODELS APPLIED IN OXFORD FLOWERS DATABASE

	20	30	40	50	60	70	80	90	100
Jhanwar [48]	0.115	0.121	0.127	0.092	0.129	0.072	0.132	0.087	0.092
Huang [49]	0.126	0.113	0.132	0.099	0.122	0.104	0.119	0.093	0.103
Lin [50]	0.141	0.192	0.174	0.121	0.132	0.101	0.149	0.112	0.134
Elalami(2011) [31]	0.153	0.198	0.182	0.116	0.138	0.098	0.156	0.118	0.139
Elalami(2014) [51]	0.161	0.203	0.191	0.126	**0.148**	0.109	**0.163**	0.129	0.144
Proposed	**0.173**	**0.221**	**0.197**	**0.132**	0.135	**0.113**	0.159	**0.134**	**0.146**

TABLE VI
THE COMPARISON OF RETRIEVAL TIME FOR DIFFERENT MODELS APPLIED IN WANG DATABASE

	10	20	30	40	50	60	70	80	90	100
Jhanwar [48]	1.53	1.62	1.72	1.91	1.76	1.85	1.73	1.71	1.66	1.87
Huang [49]	1.50	1.47	1.51	1.51	1.54	1.52	1.51	1.49	1.53	1.55
Lin [50]	0.81	0.82	0.81	0.80	0.91	0.85	0.82	0.84	0.78	0.86
Elalami(2011) [31]	0.62	0.62	0.59	0.55	0.56	0.55	0.53	0.51	0.56	0.57
Elalami(2014) [51]	0.50	0.52	0.51	**0.51**	0.54	0.49	0.51	0.47	**0.48**	0.45
Proposed	**0.49**	**0.48**	**0.50**	0.53	**0.51**	**0.47**	**0.49**	**0.46**	0.53	**0.42**

TABLE VII
THE COMPARISON OF RETRIEVAL TIME FOR DIFFERENT MODELS APPLIED IN OXFORD FLOWERS DATABASE

	34	68	102	136	180	214	248	282	316	360
Jhanwar [48]	3.11	3.14	3.08	2.92	3.14	2.93	3.18	3.21	3.15	3.31
Huang [49]	2.52	2.61	2.65	2.33	2.53	2.62	2.32	2.34	2.43	2.11
Lin [50]	1.72	1.71	1.71	1.74	1.81	1.88	1.76	1.82	1.77	1.87
Elalami(2011) [31]	1.12	1.11	1.05	1.12	1.14	1.13	1.23	1.22	1.31	1.19
Elalami(2014) [51]	0.92	0.88	**0.95**	0.87	0.89	**0.83**	0.79	0.92	**0.93**	0.89
Proposed	**0.89**	**0.84**	1.02	**0.84**	**0.83**	0.86	**0.76**	**0.91**	1.02	**0.83**

TABLE VIII
THE COMPARISON OF RETRIEVAL TIME FOR DIFFERENT MODELS APPLIED IN CIFAR-10 DATABASE

	100	200	300	400	500	600	700	800	900	1000
Jhanwar [48]	99.1	96.2	92.8	91.6	94.8	91.9	92.2	102.5	103.4	102.7
Huang [49]	75.1	72.3	73.9	71.1	74.9	73.4	75.2	73.4	73.1	72.9
Lin [50]	40.2	42.1	43.5	41.3	44.1	41.3	42.5	42.1	43.2	42.8
Elalami(2011) [31]	29.9	27.4	28.1	31.2	27.8	29.2	26.6	25.8	25.2	28.8
Elalami(2014) [51]	16.4	13.4	12.6	**13.5**	14.1	17.2	13.6	12.6	14.6	**18.2**
Proposed	**16.1**	**12.2**	**11.4**	14.7	**13.3**	**16.4**	**12.2**	**11.1**	**13.1**	19.4
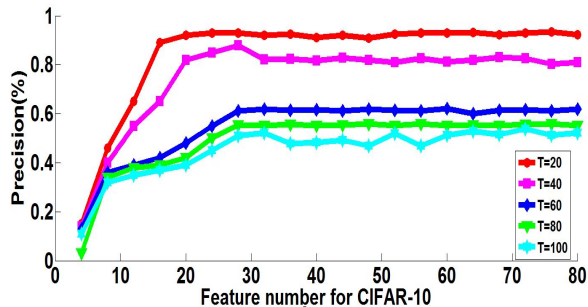
Fig. 9. The relationship between precision and number of feature in CIFAR-10 database.
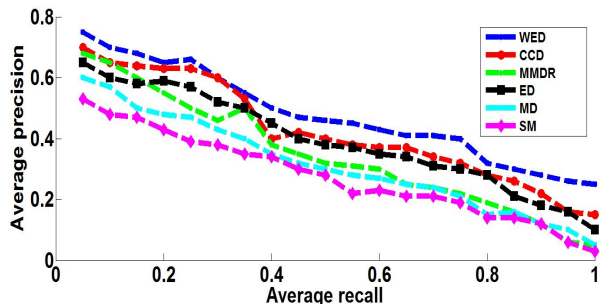
Fig. 10. Average precision and recall at different distance metrics

D. Average precision at different distance metrics

In pattern recognition, different distance detection methods have different effects on the detection results. In Sec.II, we list several commonly used distance calculation methods,such as Manhattan distance, Euclidian distance, Weighted Euclidian distance, Cross correlation, the minimum mean distance and the statistical measure. This part mainly discusses the influence of distance detection method on the detection accuracy. The experimental results are shown in Fig.10.

Experimental results show that, no matter what distance calculation method, the precision and the recall are contradictory, with the increase of the number of searches, both show the change in the opposite direction.

VI. CONCLUSION

On the basis of the existing methods for image retrieval, this paper focuses on the extraction of new features, which makes it more closely related to the image characters. First, we extract specific image features. Then the retrieval time will be increased due to the large number of features extracted, it is necessary to make feature dimension reduction. At last, new distance theory is put forward for image analysis, and ultimately get the best image related to retrieval samples. Features such as GCM and DBPSP are used. In addition, LBP, HoG and other features which have been successfully used in other fields are combined in this paper. Most importantly, a feature selection and ranking method plays an important role in this algorithm. At the same time, experiments show that our method has a great advantage in the retrieval time, which can meet the needs of real-time retrieval.

As a kind of concise and natural visual expression method, combination of features proposed has the following three advantages: it can be through the location information of local features and the relationship between each other including detection and extraction, which does not require supervised training alone and classification learning, and can effectively avoid the combinatorial explosion caused by random combination of features for it. Local area image provides a multi-scale visual expression, both contour description and preliminary clear details, compared to the single visual words, it has stronger visual expression and the ability to distinguish between local features. The relative position information implied in the combination of features can be naturally integrated into the large-scale image of the inverted index structure, geometric consistency check corresponding both to ensure the efficiency and ensure the accuracy of the matching. Of course, more features can still be extracted, in the future work, this part of the content remains to be studied.

REFERENCES

[1] Y. Sun, H. Song, A. J. Jara, and R. Bie, "Internet of things and big data analytics for smart and connected communities," *IEEE Access*, vol. 4, pp. 1–10, 2016.

[2] N. Cordeschi, M. Shojafar, D. Amendola, and E. Baccarelli, *Energy-saving QoS resource management of virtualized networked data centers for Big Data Stream Computing*, 2015.

[3] C. Lin, Z. Song, H. Song, Y. Zhou, Y. Wang, and G. Wu, "Differential privacy preserving in big data analytics for connected health," *Journal of Medical Systems*, vol. 40, no. 4, pp. 1–9, 2016.

[4] E. Baccarelli, N. Cordeschi, M. Panella, S. Mohammad, and J. Stefa, "Energy-efficient dynamic traffic offloading and reconfiguration of networked data centers for big data stream mobile computing: review, challenges, and a case study," *IEEE Network*, vol. 30, no. 2, pp. 54–61, 2016.

[5] Y. Liu, D. Zhang, G. Lu, and W. Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition*, vol. 40, no. 1, pp. 262–282, 2007.

[6] O. A. B. Penatti, F. B. Silva, E. Valle, V. Gouet-Brunet, and R. D. S. Torres, "Visual word spatial arrangement for image retrieval and classification," *Pattern Recognition*, vol. 47, no. 2, pp. 705–720, 2014.

[7] X. Zhang, W. Liu, M. Dundar, S. Badve, and S. Zhang, "Towards large-scale histopathological image analysis: hashing-based image retrieval." *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 496–506, 2015.

[8] X. Yang, Y. Zhang, T. Yao, C. W. Ngo, and T. Mei, "Click-boosting multi-modality graph-based reranking for image search," *Multimedia Systems*, vol. 21, no. 2, pp. 217–227, 2014.

[9] R. Brunelli and O. Mich, "Image retrieval by examples," *IEEE Transactions on Multimedia*, vol. 2, no. 3, pp. 164–171, 2000.

[10] J. Han and K. K. Ma, "Rotation-invariant and scale-invariant gabor features for texture image retrieval," *Image and Vision Computing*, vol. 25, no. 9, pp. 1474–1481, 2007.

[11] G. Cena, I. C. Bertolotti, A. Valenzano, and C. Zunino, "Evaluation of response times in industrial wlans." *IEEE Transactions on Industrial Informatics*, vol. 3, no. 3, pp. 191–201, 2007.

[12] C. Zhang, J. Y. Chai, and R. Jin, "User term feedback in interactive text-based image retrieval." in *SIGIR 2005: Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, Salvador, Brazil, August*, 2005, pp. 51–58.

[13] I. Ahamd and T. S. Jang, "Old fashion text-based image retrieval using fca," in *International Conference on Image Processing*, 2003, pp. III-33–6 vol.2.

[14] W. Li, L. Duan, D. Xu, and W. H. Tsang, "Text-based image retrieval using progressive multi-instance learning," *Proceedings*, vol. 58, no. 11, pp. 2049–2055, 2011.

[15] Z. Lv, J. Chirivella, and P. Gagliardo, "Bigdata oriented multimedia mobile health applications," *Journal of Medical Systems*, vol. 40, no. 5, pp. 1–10, 2016.

[16] A. Mehmood and H. Song, "Smart energy efficient hierarchical data gathering protocols for wireless sensor networks," *Smartcr*, vol. 5, 2015.

[17] L. D. Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.

[18] C. H. Lin, D. C. Huang, and Y. K. Chan, "Fast color-spatial feature based image retrieval methods," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11 412–11 420, 2011.

[19] M. J. Hsiao, Y. P. Huang, T. Tsai, T. W. Chiang, M. J. Hsiao, and T. W. Chiang, "An efficient and flexible matching strategy for content-based image retrieval," *LIFE SCIENCE JOURNAL-ACTA ZHENGZHOU UNIVERSITY OVERSEAS EDITION*, no. 1, pp. 99–106, 2008.

[20] D. L. Swets, "and juyang weng. using discriminant eigenfeatures for image retrieval," *IEEE Tpami*, vol. 18, no. 8, pp. 831–836, 1996.

[21] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: a power tool for interactive content-based image retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 5, pp. 644–655, 1998.

[22] R. Yong, T. S. Huang, and S. F. Chang, "Image retrieval: Current techniques, promising directions, and open issues ," *Journal of Visual Communication and Image Representation*, vol. 10, no. 1, pp. 39–62, 1999.

[23] N. Mitrovic, M. T. Asif, J. Dauwels, and P. Jaillet, "Low-dimensional models for compressed sensing and prediction of large-scale traffic data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 1–6, 2015.

[24] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349–1380, 2000.

[25] F. Li, Q. Dai, and W. Xu, "Improved similarity-based online feature selection in region-based image retrieval," in *IEEE International Conference on Multimedia and Expo, ICME 2006, July 9-12 2006, Toronto, Ontario, Canada*, 2006, pp. 349–352.

[26] W. Jiang, W. Er, Q. Dai, and J. Gu, "Similarity-based online feature selection in content-based image retrieval," *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 702–12, 2006.

[27] J. Yang, R. Xu, J. Cui, and Z. Ding, "Robust visual tracking using adaptive local appearance model for smart transportation," *Multimedia Tools and Applications*, pp. 1–14, 2016.

[28] J. Yang, H. Wang, Z. Lv, W. Wei, H. Song, M. Erol-Kantarci, B. Kantarci, and S. He, "Multimedia recommendation and transmission system based on cloud platform," *Future Generation Computer Systems*, 2016.

[29] J. Yang, S. He, Y. Lin, and Z. Lv, "Multimedia cloud transmission and storage system based on internet of things," *Multimedia Tools and Applications*, pp. 1–16, 2015.

[30] L. D. Xu, "Enterprise systems: State-of-the-art and future trends." *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, pp. 630–640, 2011.

[31] M. E. Elalami, "A novel image retrieval model based on the most relevant features," *Knowledge-Based Systems*, vol. 24, no. 1, pp. 23–32, 2011.

[32] J. Yue, Z. Li, L. Liu, and Z. Fu, "Content-based image retrieval using color and texture fused features," *Mathematical and Computer Modelling An International Journal*, vol. 54, no. 3-4, pp. 1121–1127, 2011.

[33] J. M. Guo and H. Prasetyo, "Content-based image retrieval using features extracted from halftoning-based block truncation coding." *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 1010–24, 2015.

[34] A. Sarmah, R. Mandal, and M. Bhattacharya, "Content based image retrieval using color, shape and texture," *International Journal of Advanced Research in Computer Science*, 2014.

[35] M. Yasmin, M. Sharif, I. Irum, and S. Mohsin, "An efficient content based image retrieval using ei classification and color features," *Journal of Applied Research and Technology*, vol. 12, no. 5, pp. 877–885, 2014.

[36] G. H. Liu, J. Y. Yang, and Z. Y. Li, "Content-based image retrieval using computational visual attention model," *Pattern Recognition*, vol. 48, no. 8, pp. 2554–2566, 2015.

[37] G. Qiu, "Color image indexing using btc," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 12, no. 1, pp. 93–101, 2003.

[38] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor." *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 19, no. 2, pp. 533–544, 2010.

[39] Y. P. Chen, S. Z. Li, and X. M. Lin, "Fast hog feature computation based on cuda," in *Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on*, 2011, pp. 748–751.

[40] Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using lbp variance (lbpv) with global matching," *Pattern Recognition*, vol. 43, no. 3, pp. 706–719, 2010.

[41] X. Wang, T. X. Han, and S. Yan, "An hog-lbp human detector with partial occlusion handling," *Proceedings*, vol. 30, no. 2, pp. 32–39, 2009.

[42] B. Wu, Z. G. Xiong, Y. Z. Chen, and Y. D. Zhao, "Classification of quickbird image with maximal mutual information feature selection and support vector machine," *Procedia Earth and Planetary Science*, vol. 1, no. 1, pp. 1165–1172, 2009.

[43] H. Mller, W. Mller, D. M. Squire, S. Marchand-Maillet, and T. Pun, "Performance evaluation in content-based image retrieval: overview and proposals," *Pattern Recognition Letters*, vol. 22, no. 5, pp. 593–601, 2001.

[44] J. Li and J. Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1075–1088, 2003.

[45] J. Z. Wang, J. Li, and G. Wiederholdy, "Simplicity: Semantics-sensitive integrated matching for picture libraries," *Pattern Analysis and Machine Intelligence IEEE Transactions on*, vol. 23, no. 9, pp. 947–963, 2001.

[46] M. E. Nilsback and A. Zisserman, "A visual vocabulary for flower classification," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, pp. 1447–1454.

[47] A. Krizhevsky, "Learning multiple layers of features from tiny images," 2012.

[48] N. Jhanwar, S. Chaudhuri, G. Seetharaman, and B. Zavidovique, "Content based image retrieval using motif cooccurrence matrix," *Image and Vision Computing*, vol. 22, no. 14, p. 12111220, 2002.

[49] P. W. Huang and S. K. Dai, "Image retrieval by texture similarity," *Pattern Recognition*, vol. 36, no. 3, pp. 665–679, 2003.

[50] C. H. Lin, R. T. Chen, and Y. K. Chan, "A smart content-based image retrieval system based on color and texture feature," *Image and Vision Computing*, vol. 27, no. 11, p. 658665, 2009.

[51] M. E. Elalami, "A new matching strategy for content based image retrieval system," *Applied Soft Computing*, vol. 14, no. 1, pp. 407–418, 2014.
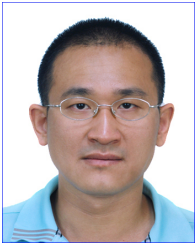
Jiachen Yang received the M.S. and Ph.D. degrees in communication and information engineering from Tianjin University, Tianjin, China, in 2005 and 2009, respectively. He is currently a professor at Tianjin University. He is also a visiting scholar with the Department of Computer Science, School of Science, Loughborough University, U.K. His research interests include stereo camera, stereo vision research, pattern recognition, stereo image displaying, and quality evaluation.

Bin Jiang received the B.S. and M.S. degree in communication and information engineering from Tianjin University, Tianjin, China, in 2013 and 2016. He is currently pursuing the Ph.D. degree at the School of Electronic Information Engineering, Tianjin University, Tianjin, China. His research interests include pattern recognition, stereo vision research, stereo image displaying and quality evaluation.

Baihua Li received her B.S. and M.S. degrees in Electronic Engineering from Tianjin University and her PhD degree in Computer Science from Aberystwyth University. She has worked at Tianjin University and Manchester Metropolitan University before she joined the Department of Computer Science at Loughborough University. Her research emphasizes innovations and novel applications of internet of things, computer vision and pattern recognition techniques in various fields. More than 50 papers have been published in high impact journals and conferences of international standard, such as Pattern Recognition, IEEE Trans Syst Man Cybern and IEEE Trans Biomed Eng.

Kun Tian received her B.S. at Xi'an University of Architecture and Technology in 1996 and her M.S. at Beijing Institute of Technology in 2005. She has worked at National Key Laboratory of Science and Technology on Aerospace Intelligence Control,Beijing,P.R. China. His research interests include pattern recognition, image classification and retrieval.

Zhihan Lv received his PhD degree from the Ocean University of China in 2012. He worked in CNRS (France) as Research Engineer, Umea University (Sweden) as Postdoc Fellow, Fundacion FIVAN (Spain) as Experience Researcher. He was a Marie Curie Fellow in European Union's Seventh Framework Programme LANPERCEPT. He has also been an assistant professor in Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China since 2012. His research mainly focuses on Virtual Reality, Augmented Reality, Multimedia, Computer Vision, 3D Visualization, Graphics, Serious Game, Human-Computer Interaction, Networks, Bigdata, Software Engineering.