DESIGN OF COMPLICATED DUPLICATE IMAGE REPRESENTATION APPROACH BASED ON DESCRIPTOR LEARNING

Yongjiao Wang [1, 2], Xiaojie Du [1] and Lei Liang[2]

[1] School of Computer Science, University of Urban Construction, Pingdingshan, 467036, China

[2] National Engineering Laboratory for Fiber Optical Sensing Technology, Wuhan University of Technology, Wuhan, 430070, China

Submitted: July 15, 2014 ***Accepted: Mar. 30, 2015*** ***Published: June 1, 2015***

Abstract- In order to solve the low discrimination of image representations in complicated duplicate image detection, this paper presents a complicated duplicate image representation approach based on descriptor learning. This approach firstly formulates objective function as minimizing empirical error on the labeled data. Then the tag matrix and the classification matrix of training dataset are brought into the objective function to ensure semantic similarity. Finally, by relaxing the constraints, we can get the learning hashes. The learning hashes are used to quantify local descriptors of images into binary codes and the frequency histograms of binary codes are as image representations. Experimental results demonstrate that compared with the state-of-the-art algorithms, this approach can effectively improve the discrimination of image presentations by introducing semantic information.

Index terms: hashing; image; semi supervised learning;

I. INTRODUCTION

Duplicate image problem is becoming more and more common with the explosive growth of the Internet multimedia data. This application has brought new challenges to the diversity of images. For example image retrieval systems can returns more related image according to the query content. If the returned image contains a large number of duplicate content messages. It will browse a great inconvenience to the user. Because experienced people tend to only focus on the first few pages when users browse the contents. In addition multimedia data acquisition easy to been tampered because it efficiency and convenience. This gives the image copyright protection has brought a lot of problems. So people urgently need a kind of technology. It can fast, accurate and reliable to find all copies of a given image on the Internet. The traditional digital watermark technology is embedded in the additional information to the image. This can also be found duplicate images effectively. However, it must be in the image to embed the watermark in the image before release. For the published images are powerless. This greatly limits the application of digital watermarking technology. By extracting image invariant features to represent the image Based on the content of the duplicate image detection technology, Not only has good identification results and won't change the content of the image itself, thus has better applicability.

Although global features to simple duplicate image has better retrieval accuracy and scalability. But its universality is not strong. It can not be achieved cropping, rotation, translation and other geometric transformations. To solve this problem requires the use of more robust local feature. Compared with global features, image local features is stronger robustness for optical transformation and geometric transformation. It is widely used in complicated duplicate image detection. However the local feature is represented as a collection of images in high-dimensional feature vectors. This has brought great pressure to the data organization and management.

In recent years, the characteristics of bag of words model BOF is considered key technology solving large-scale complicated duplicate image retrieval. The core idea of this technique is the local descriptor quantization "visual words by K-means clustering". An image is represented as a set of "visual words". Since the quantization technique can avoid storage and relatively high dimension vector. Therefore, this technique can effectively reduce the memory usage and improve the retrieval efficiency. However, the quantization process of BOF has greatly reduced the image representation of the differentiated [1, 11-13, 17-18]. This is because the construction of the visual dictionary of the BOF model is mainly completed by the K-means algorithm. But when the dictionary size is over 105, the training is difficult to achieve [2, 14-15]. The suppression of a visual dictionary size reduces the discriminative visual words. Although the researchers have presented a layered K-means [3, 16], approximate K-means [4] optimal algorithm to solve this problem, the memory limited is still a fundamental reason that restricts the effectiveness of retrieval.

In order to solve the above problems, literature [5] use unsupervised learning of local descriptor quantization image into binary code to improve the image distinguishes ability. Visual dictionary as a result of unsupervised learning can be great. So visual word based on the unsupervised learning is more discriminative than based on Kmeans clustering definition. The related experiment results also prove this conclusion [5]. However, the unsupervised learning process in the generation of a visual dictionary does not consider the semantic information. While a large number of research results show that the local image descriptor, similarity metrics cannot ensure similarity in semantic [6]. Therefore, in order to get more precise image representation, can distinguish. This article through carries on the semi supervised learning to introduce semantic information of the training data set. It makes binary coding of image local descriptors and the corresponding visual words can describe the image content more accurate.

This paper adopts the semi supervised learning to generate image representation is mainly based on the consideration of the following three aspects. First semi supervised learning training is simple. It uses a small part of image annotation generated learning hash. It calculates Binary coded of local descriptor by learning hash. This not only reduces the image representation of the generation cost, but also in line with only a small part of the Internet image with annotation information reality. Second the semi supervised training process of learning can effectively by introducing semantic information. According to the semantic information generating learning hash can not only ensure the image descriptor similarity metrics but also it can ensure semantic similarity. In the dictionary generated process, it can map matching descriptors to the same binary coding and map not match local descriptor to the different binary code. It reduces the effect of quantization noise and descriptor noise. It guarantees that the distinction of local descriptor. Third Binary coded generated by local descriptor is less storage space .It is conducive to the expansion of the system. In real time query it can through the hash table and bit operation accelerated binary code comparison process.

Duplicate image with a variety different definitions in the existing literature. But there is no a repetitive based on image content discovery technology can effective against all definitions. If people do not know the definition content possible in advance, they can not target to improve the recognition rate of duplicate image. So the duplicate image definition is given before. According to the manual FOO et al found [1], duplicate image on the Internet is the most common of the source image scaling and cropping. Based on the multi resolution theory and image compression theory, image scaling transform will make the image content change. But its overall structure will remain the same. For this type of image transform using global features can achieve a higher retrieval precision. Geometric and optical transformation led by clipping transform due to the overall information of image change greatly. so we need more stable local features to identify. Therefore, The duplicate image is divided into the following two categories based on repetitive images found method.

(1) simple duplicate image

Definition 1: the simple duplicate image is a collection of duplicate images of various scaling transform, image storage format transform, watermark transform, color transformation generated by the same source.

Scaling: including image scaling and scaling.

Storage format transform: image storage format is JPEG, BMP, PNG, TIFF etc. Images in different formats conversion between images appear to be almost no change. But the image coding has changed.

Watermark: watermark information transformation is different with the same image. Such as Sina, Tencent micro logging service provider will automatically add the watermark information of each user to upload images when upload a picture.

Gray level transformation. It include color image into gray image or gray image is transformed into a color image. The application of many occasions often involves the use of different color space to represent the image. This needs to be able to unify measurement of image color.

Watermark transform, gray-scale transformation and storage format conversion are often associated with scaling appearance. At the same time also does not affect the overall image structure, so in our account.

(2) Complicated duplicate image

Definition 1: complicated duplicate image are different replica sets that source image be a variety of image transformed. Mathematics expressed as $I' = f(I)$ .I represents the source image. A collection of F represents one or more tolerable image transform. The new image [1] transform after the generation of representative. Can tolerate transform mainly refers to the optical image transform or geometric transformation.

Optical transform: It including image illumination, brightness and color transformation.

Geometric transformation: It also called space transformation. It is the coordinate mapping an image to a new location in the other image. It does not change the image pixel values. It is just in the image plane of pixel rearrangement. It includes cutting transform, image rotation transformation, translation transform, twist transformation etc.

II. SEMI SUPERVISED LEARNING OF LOCAL DESCRIPTOR

Although the Internet exist a large number of images, but only a small part of them include mark information (including image annotation). So the research on the of this section focus on researching how to realize the semi supervised learning according to the local descriptor and the mark information of small part image.

Traditional semi supervised learning [6] is extracted global features from whole the image. The global feature that mapping into binary code to increase the system real-time retrieval. It combines image represented as some bits, so it must be accompanied by a large amount of information loss. It is difficult to accurately describe the image. It eventually resulted in retrieval accuracy is hard to meet the needs of the actual retrieval. In order to solve this problem, in this section, through the construction of the training data set label matrix and classification matrix to semi supervised learning is introduced into the traditional learning process in the local descriptor. Because an image contain a large number of local descriptors, learning process for local descriptor is equivalent to repeat confirm duplicate image. It can effectively improve the accuracy of image retrieval.

The goal of semi supervised learning is to generate K hash function. It is used to convert the local descriptor matrix $X \in \mathbb{R}^{D \times N}$ of all images into binary code matrix $Y \in B^{K \times N}$. N represents the total number of local image descriptors. D represents the dimension of image descriptors, each column represents a local descriptor. This section uses the linear projection as Hashi function:

$$h_k(x_i) = \operatorname{sgn}(w_k{}^T x_i + b_k) = y_{ki}, k = 1, ..., K \cdot \tag{1}$$

$h_k()$ represents the k-th hash function. x_i represents the i-th local descriptor. $\operatorname{sgn}()$ represents a symbol function. w_k represents the projection vector. b_k is the data projection expectations. In this section $b_k = 0$. Y_{ki} represents the i-th local descriptor corresponding to the k-th digit code. It makes Hashi function cluster $H = [h_1(x_i), h_2(x_i), ..., h_K(x_i)]$. $W = [w_1, w_2, ..., w_K]$ is the K projection vector of this festival ask.

In order to ensure the semantic similarity, the algorithm using hashes function mapping the match local descriptor to the same binary coding. That is all for 1 or -1.It mapping mismatch Local descriptors to different binary encoding, which is respectively 1 and -1. Therefore the following objective optimization function is set up:

$$J(H) = \sum_k \{ \sum_{(x_i, x_j) \in M} h_k(x_i) h_k(x_j) - \sum_{(x_i, x_j) \in C} h_k(x_i) h_k(x_j) \} \tag{2}$$

M represents the local image descriptor x_i and x_j matching. C represents the local image descriptors x_i and x_j mismatch. In order to apply the formula (3-1) into a more compact matrix, it needs to calculate the training data of local descriptor tag matrix X_L and classification matrix S. This process is different with the traditional semi supervised learning [6]. In the traditional semi supervised the learning, each image marking information are known. It can easily building tag matrix and classification matrix according to the label information. To get the final learning vector W. But in this section, local descriptor tag information and classification information is unknown. Therefore it needs to rebuild the local descriptor tag matrix X_L and classification matrix S.

2.1 Tectonic marker matrix and classification matrix

The core idea of Semi supervised learning is to ensure the semantic similarity. It maps image with same tag information to the same binary coding construct hash function.

It maps Image with different marker information to different binary encoding. Use of this idea, in order to ensure the semantic similarity of local descriptors learning process .It needs to map match local descriptors to the same binary coding. Mismatch Local descriptors mapped to different binary encoding. This section constructs marker matrix X_L and classification matrix S as follows:

Definition 1: marked matrix X_L representation matrix composed of local descriptors by all markers image. Each column represents a local descriptor with marker information. L represents the number of local descriptors with marker information. The local descriptor tag information should be same with their label information of consistent image. That is to say if images tag information as "blue sky". Then all of its local descriptor tag information must be "blue sky".

Definition 2: classification matrix S representation matrix composed of all pair wise local descriptors with marker information according to the classification of the genus. This is a $L \times L$ dimensional matrix.

$$s_{ij} = \begin{cases} 1 & (x_i, x_j) \in M \\ -1 & (x_i, x_j) \in C \end{cases} \circ$$

Matching condition is x_i and x_j simultaneously meet the following three conditions:

1) x_i and x_j have the same label information.

2) image in the nearest neighbor matching , x_i and x_j in line with the nearest neighbor matching principle proposed by Lowe [7].

3) (x_i, x_j) is the interior point random consistency check.

When one of x_i and x_j satisfy the conditions as the mismatch:

1) x_i and x_j has a different marker information.

2) x_i and x_j have the same label information. But in nearest neighbor matching is not possible matching points.

3) x_i and x_j have the same label information. In the nearest neighbor matching point is also possible matching. But not random consistency checks of the interior point.

It can make matrix X_L and classification matrix S of local descriptors according to the definitions 1 and 2 . Bring them into the formula (2), a formula (3) established:

$$J(H)=\frac{1}{2}tr\{H(X_L)SH(X_L)^T\} \tag{3}$$

Hashi function $h_k(x_i)=\mathrm{sgn}(w_k{}^T x_i)$ into formula (3), formula (4) established:

$$J(W)=\frac{1}{2}tr\{\mathrm{sgn}(W^T X_L)S\,\mathrm{sgn}(W^T X_L)^T\} \tag{4}$$

The above process only considers the labeled data in image set. It is prone to over fitting phenomenon when it extended to the entire image set. In order to get more generalization form, it needs to use unlabeled data to increase adjustment of extra. To adjust the objective function using the entire training data set. A formula (5) established:

$$\begin{aligned} & H^* = \arg\max_H J(W) \\ \textit{subject to} \quad & \sum_{i=1}^{N} h_k(x_i)=0, k=1,\ldots,K \\ & \frac{1}{N}H(X_N)H(X_N)^T = I \end{aligned} \tag{5}$$

X_N represents matrix construct by all local descriptors of images in training set . Each column represents a local descriptor. N represents the number of all the local descriptor. I represent the unit matrix. A constraint representation the probability is equal and independent that each bit is 1 or -1 .Bit is refers to binary code generate by the semi supervised learning.

2.2 Solving the objective function

The balance and the independent segmentation of hash code is a problem of NP. So it need to relax the objective function and constraint conditions to obtain the approximate optimal solution for the formula (5) calculations. The detailed derivation as follows [6]:

1) Relax the objective function

The formula (5) to relax the sign restrictions on the objective function to get the target functions of the new formula (6):

$$J(W) = \frac{1}{2} tr(W^T X_L S X_L{}^T W) \tag{6}$$

2) Relax the balancing inhibitory conditions

By maximizing the variance to replace the balance of hash code training data base on segmentation problem, a formula (7) established:

$$J(W) = \frac{1}{2} tr[W^T X_L S X_L{}^T W] + \frac{\eta}{2} E[\|h_k(x) - u_k\|^2] \tag{7}$$

η is a positive scalar. It represents the adjustment factor of regular. u represent search hash code expectations. u is 0 when maximum variance. Therefore a formula (8) established:

$$J(W) = \frac{1}{2} tr[W^T X_L S X_L{}^T W] + \frac{\eta}{2} E[\|w_k{}^T x\|^2] \tag{8}$$

Formula $\sum_k E(\|w_k{}^T x\|^2) = \frac{1}{n} tr[W^T X_N X_N{}^T W]$ and η is an adjustment parameter, to bring them into (8), a formula (9) was established in:

$$J(W) = \frac{1}{2} tr[W^T X_L S X_L{}^T W] + \frac{\eta}{2} tr[W^T X_N X_N{}^T W] \tag{9}$$

Set $M = X_L S X_L{}^T + \eta X_N X_N{}^T$, formula (10) was established in:

$$J(W) = \frac{1}{2} tr\{W^T M W\} \tag{10}$$

3) Relax the orthogonal constraint conditions

The hash function is a limitation $WW^T = I$ .This restriction requires the solution vector between orthogonal. It can relax this condition and gets the hash function more general representation. a formula (11) was established in:

$$\begin{aligned} J(W) &= \frac{1}{2} tr\{W^T M W\} - \frac{\rho}{2} \|W^T W - I\|_F^2 \\ &= \frac{1}{2} tr\{W^T M W\} - \frac{\rho}{2} tr[(W^T W - I)^T (W^T W - I)] \end{aligned} \tag{11}$$

The objective function is non convex. So it is very difficult to find the global optimal solution. It can get the local optimal solution. A formula (12) was established:

$$WW^T W = (I + \frac{1}{\rho} M)W \tag{12}$$

ρ represent plus adjustment. It Can prove that Q is a positive definite matrix when $Q = I + \frac{1}{\rho} M$. Learning vector matrix at this point it is easy to solve. A formula (13) was established:

$$W = LU_k \tag{13}$$

L is positive definite matrix Q by using Choleshy decomposition, $Q == LL^T$. U_k is the former K feature vector of M.

2.3 Image representation

Put $W = [w_1, ..., w_K]$ into the hash function $h_k(x_i) = \mathrm{sgn}(w_k{}^T x_i), k = 1, ..., K$. It can obtain the K binary coding y_i of each locale descriptor. This will set the image representation into a group of binary code. Each binary code can be regarded as a visual word. All image binary code can constitute a visual dictionary. Image representation is the frequency histogram obtains by visual words according to the TF-IDF weighting strategy.

The semi supervised learning process can be effectively introducing semantic information. According to the semantic information generating learning hash can not only ensure the image descriptor similarity metrics and ensure semantic similarity. It can effectively reduce the effect of quantization noise and descriptor noise. Ensure the image can be differentiated.

III. RESULTS AND ANALYSIS

In order to validate the algorithm, it can carry out the following experiment. 1) The selection of parameters. Objective is to select the appropriate length of binary code to determine the size of the visual dictionary. 2) Retrieval accuracy comparison. The purpose is use the retrieval results validate the algorithm of image that can be differentiated.

3.1The experimental set up

(1) Experimental data set

The image data used in the experiments set comes mainly from four parts.

Data set 1: UKbench[8] image sets data set. This data set consists of 2550 groups of images. Each group contains 4 images of the same object or scene at different angle view. Each image has the obvious affine transformation.

Data set 2: Holidays [1] image sets data set. This data set consists of 1491 images. It is divided into 500 groups. Each image can be used to test the robustness of attack on the rotation, light and fuzzy.

Data set 3: Randomly download 100000 images from the flick .100 images were randomly selected as the test image.

Data set 4: in order to obtain a copy of the image, each test image by generating 33 image copies of various transformation software Stirmark. Specific transformation types are as follows: adding noise, affine transform, Gauss filtering, sharpening, cropping (50%, 60%, 70%, 80%, 90%), JPEG compression, median cut, the peak signal to noise ratio and linear remove, rotation, scaling (50450900) (25%, 50%, 75%).

In this experiment, the remaining 99900 images of the data set 3 and data set 1, 2 and 4 set together constitute the training database. The data set 1and the data set2 is used to provide mark information. Although UKbench image sets and Holidays image sets are not a strict sense duplicate image. But they are description in different conditions on the same object or scene. So it has a large number of correct matching descriptors. And because the image contains different conditions of transformation, so it has stronger robustness for various transformation attacks.

(2)Experimental process

The experiments were conducted in the same soft and hardware environment. Software environment is CentOS release 5.5, 2 Intel XeonCPU 2.4GHz, 12G RAM, g++ 4.1.2, matlab7.1. The local image descriptor extracted by Hessian-Affine detector and SIFTS descriptor [9]. Local descriptor matching between images by literature [10] presented the nearest neighbor matching principle. The nearest neighbor distance and the next nearest neighbor distance ratio ratio is set to 0.6.

In practical application, it is difficult to directly determine the merits of image representation method. Therefore, we use an indirect comparison method by constructing the same index structure to detect retrieval accuracy under different image. According to the actual search results to distinguish the quality of image representation, in the index structure of this paper, two tables T1 and T2 were constructed. T1 adopts the inverted index structure. Each keyword corresponds to a series of entity behind. The entity content is image ID contains the code and address information in the T2 table. The T2 table is obtained according to the image frequency histogram TF-IDF weight strategy. Frequency histogram is a sparse vector. In order to save storage space and reduce the comparison time, the algorithm only store binary coded images actual contain and the correspond weight. For the query image and the candidate image similarity computation, the two image binary code for the union expansion. Then calculate the two image similarity by the cosine of the angle between.

As shown in Figure 1, the algorithm first extracts the query local descriptors of image when the query image comes. Then use the second section generation learning put local descriptor quantization into binary code. Then it can find the candidate images contain the same code and the frequency histogram by querying the inverted index. Finally, it can determine the image similarity.

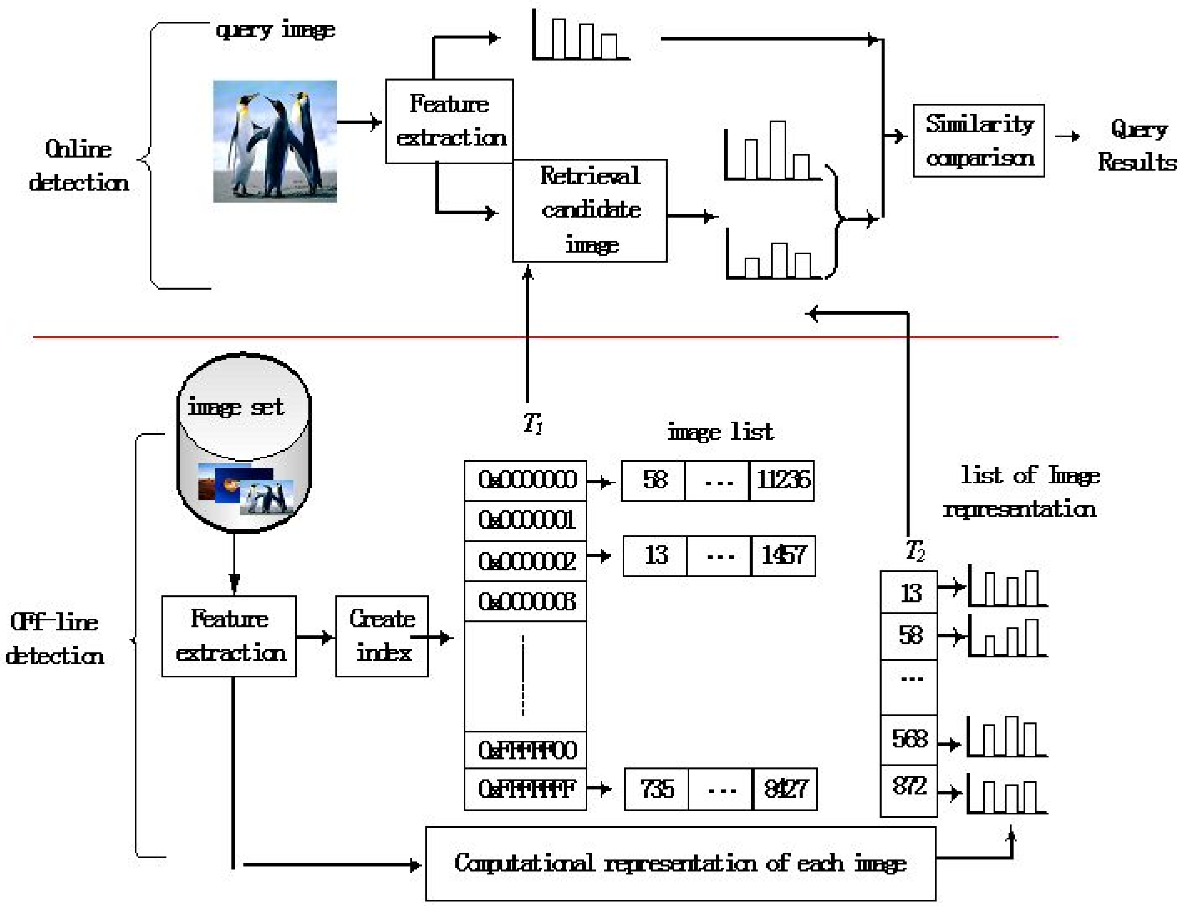

Fig. 1 Experimental Architecture

3.2 Selection of parameters

Experiments select the suitable parameters K by comparison retrieval accuracy of duplicate image with different code length. For each test image are generated by 33 different images copy. Therefore, it can directly use the former 33 images each query returns images to calculate the experimental average retrieval precision. A formula (14) established:

$$\text{Average retrieval accuracy} = \frac{\text{Number of repeat image contained in the former 33 images by each query returns}}{3300} \tag{14}$$

As shown in Figure 2, experiment has the best retrieval results when K=28. At this time the average retrieval precision is 93.6%. At beginning with binary code length increases, the average accuracy of the image retrieval will rise. But when the code length reaches a certain value, the precision of retrieval at this time decreased with the increase of the binary encoding length. This is because when the binary code length is small, the visual dictionary is small. There will be a large number of local descriptors are mapped to the same visual words. It reduce the discriminative local descriptor. The correct matched local descriptor is quantified in different visual words when the binary code length is too large. It reduces the retrieval accuracy of duplicate images.

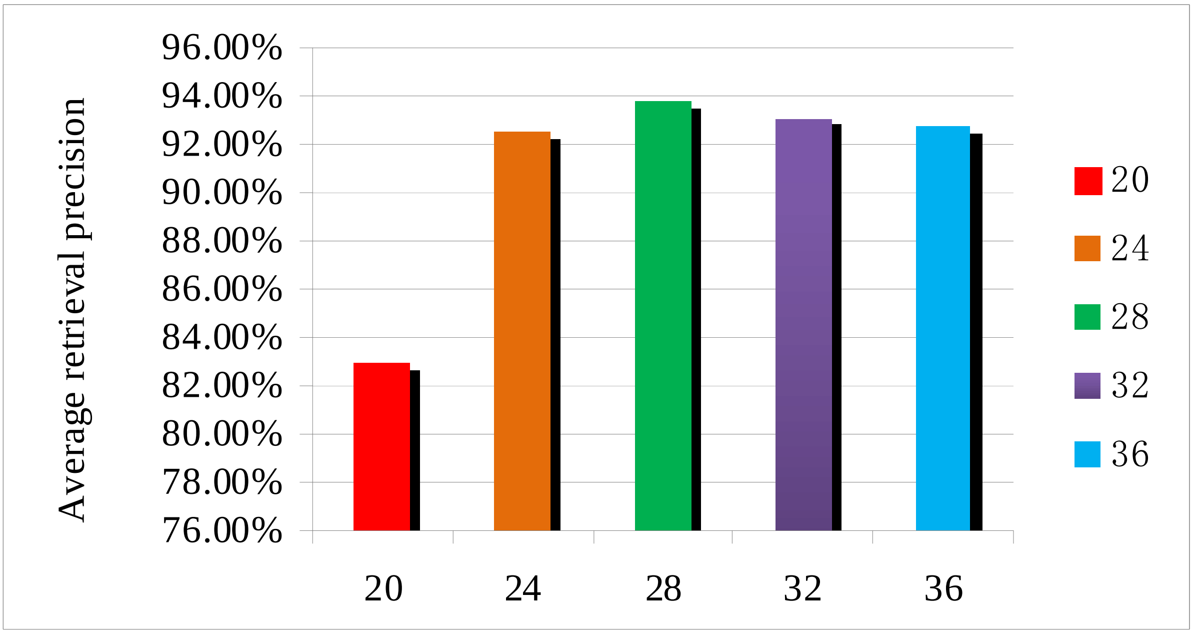

Fig 2. Binary code length selection

As shown in figure 3.The burden of duplicate image retrieval is aggravating gradually with the increase of K. Therefore, the value of K according to the average retrieval precision and retrieval time are two factors to consider.

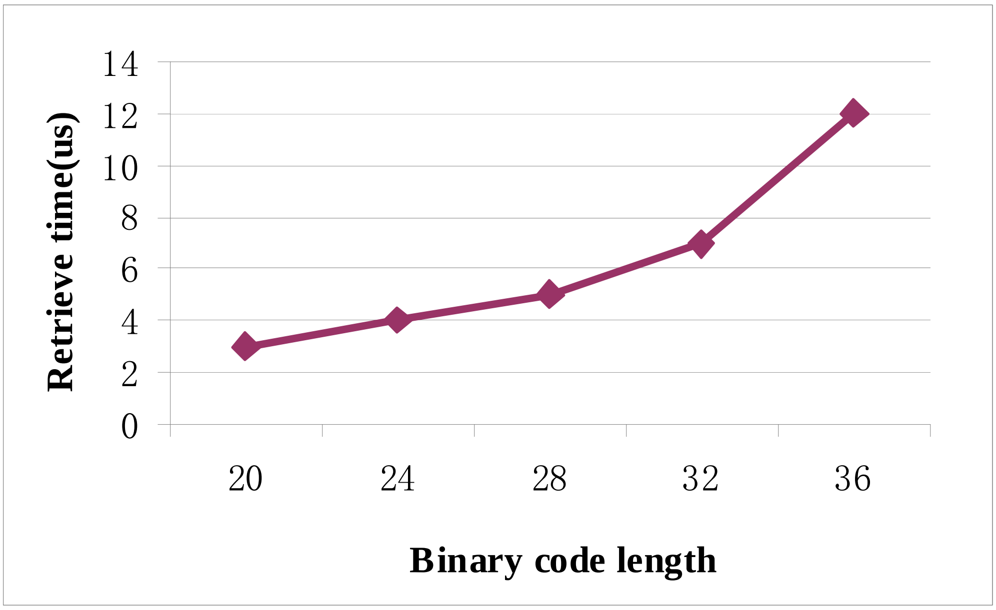

Fig 3. Retrieval time

Select the appropriate length of binary coding is very important in practical application. Retrieval precision different encoding length change with time, as shown in Fig 4. This article choose K=28.

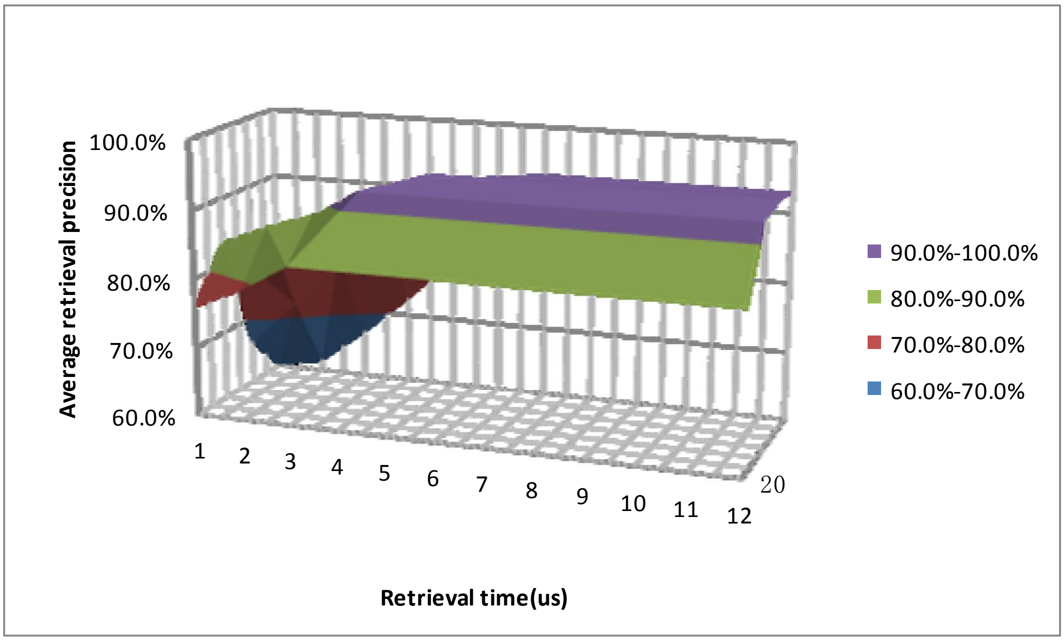

Fig 4. Surface plot

3.3 Retrieval accuracy comparison

In this section, compares this algorithm with the algorithm in literature [6] and [5]. We can see that this algorithm has the best retrieval results from Figure 5. This is because the literature [6] is use semi supervised learning transforms the global GIST features of images into binary code. In the inevitable process of image content quantification of high dimension into binary code in low dimensions with large amount of information loss, therefore the actual retrieval effect is not ideal. The average retrieval precision is only 68.5%. While this algorithm and literature [5] is extracted image local descriptors. It make the binary encoding set of local descriptor on contains more information. It can describe the image more accurately. The literature [5] is using unsupervised learning by entropy maximization principle to quantize the local descriptor into binary code. It does not consider the semantic information of images. It can not guarantee the actual semantic retrieval process similarity. The average retrieval precision is 87.2%. This algorithm introduces semantic information by the semi supervised learning. It can effectively reduce the image content and different semantics between the underlying characteristics. So it has better retrieval effect. The average retrieval precision is 93.6%.

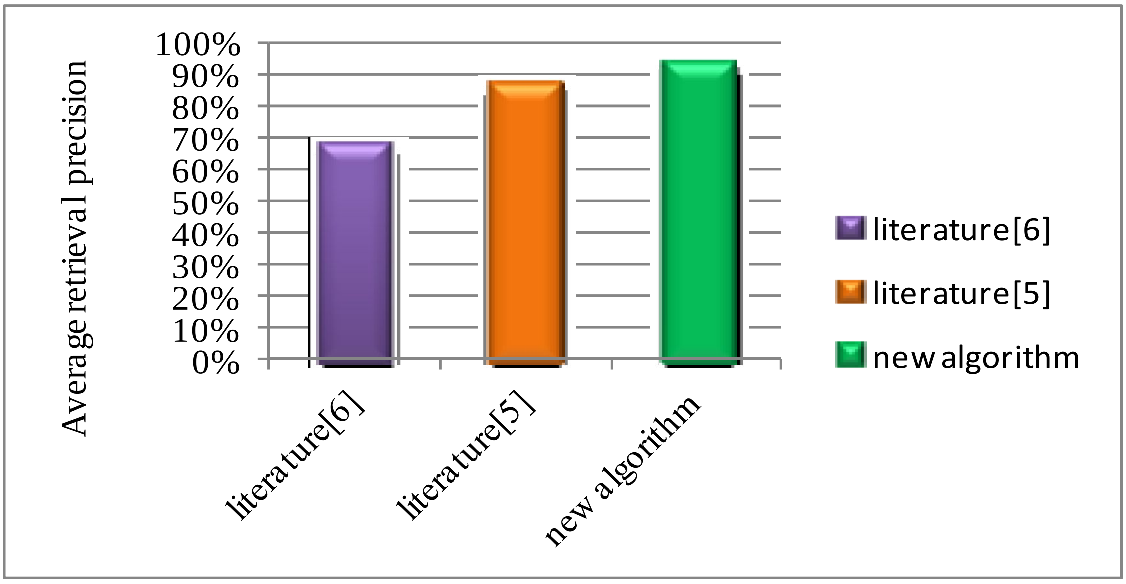

Fig. 5 comparison algorithm

In this paper average search time three algorithms were tested at different image number. As shown in Figure 6. The new algorithm with the number of images increases. The average search time is close to linear increase, while the other two algorithms are not. When the number of images is 12 million, the average retrieval time is only 0.7s. Therefore, this system has better scalability. Moreover, the average search time can reduce with cluster scale further expansion.

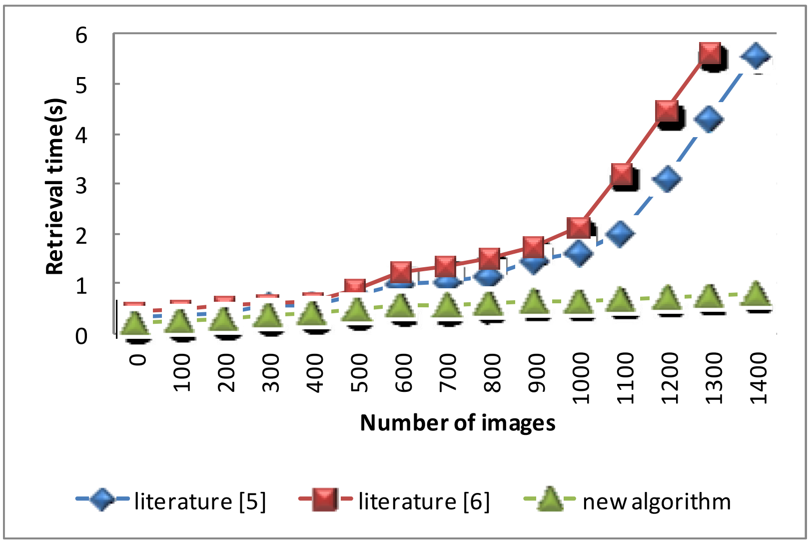

Fig. 6 Retrieval time comparison

IV. CONCLUSIONS

In order to solve the low discrimination of image representations in complicated duplicate image detection. This paper presents a complicated duplicate image representation approach based on descriptor learning. The method has two salient features. 1) Only a small part of the image has the characteristics of the annotation information in the internet. This paper presents a semi supervised learning to introduce semantic information. Through the generation of hash computation local descriptor visual words, not only ensures low generation overhead visual dictionary, but also ensure the semantic similarity. Reduce the effect of quantization noise and descriptor noise. Improve the distinguish ability of local descriptor. 2). the traditional semi supervised learning algorithm is applied to the learning process in the local descriptor by marking matrix constructing local descriptors and classification matrix. This improves the applicability of algorithm. The experimental results show that algorithm presented in this paper has better distinction.

Although the algorithm has better retrieval accuracy in application， there are still some deficiencies. For example: geometric information between local image descriptor is not considered. Therefore, in the follow-up research it can be improved and be optimized to improve the retrieval precision.

V. REFERENCES

[1] .Jegou H, Douze M, Schmid C. Hamming Embedding and Weak Geometric Consistency for Large Scale Image Search, In Proc. of the 10th European Conference on Computer Vision(ECCV), 2008: 304-317.

[2] San Pedro J, Siersdorfer S, Sanderson M. Content Redundancy in YouTube and its Application to Video Tagging, ACM Transactions on Information Systems(TOIS). 2009, 29(3):13-43.

[3] G. Sengupta, T.A. Win, C. Messom, S. Demidenko and S.C. Mukhopadhyay, “Defect analysis of grit-blasted or spray printed surface using vision sensing technique”, Proceedings of Image and Vision Computing NZ, Nov. 26-28, 2003, Palmerston North, pp. 18-23.

[4] Kumar A, Sabharwal Y, Sen Sandeep, A simple linear time (1 +)-approximation algorithm for k-means clustering in any dimensions, In Proc. of the 45th Annual IEEE Symposium on Foundations of Computer Science. 2004: 454-462.

[5] Ling HF, Yan LY, Zou FH, et al. Fast Image Copy Detection Approach Based on Local Fingerprint Defined Visual Words, Signal Processing. 2013, 93(8): 2328-2338.

[6] Wang J, Kumar S Chang SF. Semi-Supervised Hashing for Scalable Image Retrieval, In Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2010: 3424-3431.

[7] T. Ohji, S.C. Mukhopadhyay, M. Iwahara and S. Yamada, “Performance of Repulsive Type Magnetic Bearing System under Nonuniform Magnetization of Permanent Magnet”, IEEE Transactions on Magnetics, Vol. 36, No. 5, pp. 3696-3698, September 2000.

[8] Lowe D. Distinctive Image Features from Scale-invariant Key points, International Journal of Computer Vision. 2004, 60(2):91-110.

[9] Chandrasekhar V, Takacs G, Chen D, et al. CHoG: Compressed histogram of gradients a low bit-rate feature descriptor, IEEE Conference on Computer Vision and Pattern Recognition. 2009:789-801

[10] Dong W, Wang Z, Charikar M, et al. High-Confidence Near-Duplicate Image Detection, In Proc. of the 2th ACM International Conference on Multimedia Retrieval. 2012: 313-321.

[11] S.C. Mukhopadhyay, S. Deb Choudhury, T. Allsop, V. Kasturi and G. E. Norris, "Assessment of pelt quality in leather making using a novel non-invasive sensing approach", Journal of Biochemical and Biophysical methods, Elsevier, JBBM Vol. 70, pp. 809-815, 2008.

[12] Xie HT, Gao K, Zhang YD, et al. Pairwise Weak Geometric consistency for large scale image search[C].In Proc. of the 1th ACM International Conference on Multimedia Retrieval, Trento, Italy, 2011:186-195

[13] Liu, Z. ; Li, H. ; Zhou, W. ; Hong, R.; Tian, Q,Uniting Keypoints: Local Visual Information Fusion for Large Scale Image Search, IEEE Transactions on Multimedia, vol.pp,no.99, pp.1-6, 2015.

[14] S.C. Mukhopadhyay, F.P. Dawson, M. Iwahara and S. Yamada, "A Novel Compact Magnetic Current Limiter for Three Phase Applications", IEEE Transactions on Magnetics, Vol. 36, No. 5, pp. 3568-3570, September 2000.

[15] Kristen Grauman, Rob Fergus, Learning Binary Hash Codes for Large-Scale Image Search, Studies in Computational Intelligence, Vol.411, 2013, pp 49-87.

[16] Zhen Liu, Houqiang Li, Wengang Zhou, Ruizhen Zhao, Qi Tian, Contextual Hashing for Large-Scale Image Search, IEEE Transactions on Image Processing, vol. 23 , no.4, pp.1606 - 1614, 2014.

[17] Swati Agrawal, Ashish Sureka,Copyright Infringement Detection of Music Videos on YouTube by Mining Video and Uploader Meta-data, Lecture Notes in Computer Science, vol. 8302, 2013, pp 48-67.

[18] Ulges, A. ; Koch, M. ; Borth, D. ; Breuel, T.M,TubeTagger - YouTube-based Concept Detection, IEEE International Conference on Data Mining Workshops, pp.190 - 195, 2009.

[19] Mohammed El Agha, et al., Efficient and Fast Initialization Algorithm for K-means Clustering, I.J. Intelligent Systems and Applications, 2012, 1, 21-31.

[20] S. Yamada, K. Chomsuwan, S.C. Mukhopadhyay, M. Iwahara, M. Kakikawa and I. Nagano, "Detection of Magnetic Fluid Volume Density with a GMR Sensor", Journal of Magnetics Society of Japan, Vol. 31, No. 2, pp. 44-47, 2007.

[21] Leongwai Yie, Joel Than Chia Ming, Features of Sleep Apnea Recognition and Analysis, International Journal on Smart Sensing and Intelligent Systems, vol. 7, no. 2, pp, 481 – 497, 2014.

[22] Wenqing Che and Wang Tao, A Scene Recognition Algorithm Based on Multi-instance Learning, International Journal on Smart Sensing and Intelligent Systems, vol. 7, no. 2, pp,1470 – 1492, 2014.