Image Retrieval Based on Hierarchical Locally Constrained Diffusion Process

Xianhua Zeng* , Meng Hu , and Suwen Zhu
Chongqing Key Laboratory of Computational Intelligence
Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China
**Email: zengxh@cqupt.edu.cn*

***Abstract*—To address the problem of high matrix computational cost in traditional diffusion process methods for large scale image retrieval, we propose a novel image retrieval algorithm based on hierarchical locally constrained diffusion process (HLCDP) by combining algebraic multigrid and diffusion ranking on manifold. In HLCDP, all retrieved images of the image database are built into a bottom-to-top hierarchical structure by selecting the representative images. Then the similarity among images on the top layer is diffused by using locally constrained diffusion process, and the affinities (i.e. ranking scores) between query images and top-layer representative images are interpolated to all images on the bottom layer to obtain the affinities between the query images and all of the images in database. Our method is evaluated on image retrieval tasks by comparing with locally constrained diffusion process (LCDP), self-smoothing operator (SSO) and self-diffusion (SD). The experimental results on *MPEG7* data set and *ImageCLEFmed2005* data set demonstrate that the proposed method improves the retrieval performance and reduces the retrieval time consumption.**

Keywords—Hierarchical structure; algebraic multigrid; Manifold learning; Diffusion process

I. Introduction

With the rapid development of communication technology and multimedia technology, the amount of images rapidly increases. How to retrieve images the users needed quickly from large scale images has a very important research value. Due to the development of machine learning technology, manifold learning [1], [2], [3] has been widely used in content based image retrieval. In traditional image retrieval methods based manifold learning [4], [5], [6], [7], image matching in retrieval tasks is only to calculate the affinity values between pairwise images, and then rank the most similar images according to these values. These methods ignore the intrinsic manifold structure of all images. To improve these issues, researchers begin to focus the influence of all images in the image database on the affinities between pairwise images, and then propose diffusion process methods.

The diffusion process methods make good use of the underlying manifold among all images in database to improve the retrieval performance. For example, label propagation [8] propagates the labeled information from labeled samples to unlabeled samples by using the diffusion process through the underlying manifold of the database. Yang [9] proposes a Locally Constrained Diffusion Process (LCDP). The LCDP algorithm replaces the connected graph of all images in database with K nearest neighbor graph to reduce the effect of noise. This method mainly considers the local structure of the manifold. There are also some diffusion methods focusing on the global structure. Jiang [10] proposes Self-Smoothing Operator (SSO), diffusing similarity of all images through the connected graph. Inspired by SSO, Wang [11] proposes a self-Diffusion (SD) method that improves the diffusion performance by adding a unit matrix at each diffusion step. In general, diffusion methods use the Gaussian kernel function to compute the similarity among images and structure the connected graph, then diffuse similarity through the graph. However, there are some deformed graph structures. Wang [12] proposes the shortest path Propagation (SPP), which is constructed by the graph based on the shortest path between images. This method improves the retrieval performance markedly. But this method result in an increase in runtime, since each of the images to be retrieved needs to be computed separately. Compared with traditional image retrieval methods based manifold learning, these above diffusion process methods have improved the retrieval efficiency. But these diffusion process methods require a large number of iterative calculations to converge. These methods are improper for large scale images.

In order to solve the problems of high matrix calculation and large amount of iterative computation of large-scale images, and to guarantee a better retrieval accuracy, we propose a novel image retrieval algorithm based on hierarchical locally constrained diffusion process which is derived from algebraic multigrid [13], [14], [15], [16] and diffusion ranking on manifold. Our algorithm builds all retrieved images in database into a bottom-to top hierarchical structure. Next, we use locally constrained diffusion process to obtain the ranking scores between query images and top-layer representative images. Finally, the ranking scores are interpolated to all retrieved images on the bottom layer. Compared with the locally constrained diffusion process (LCDP) algorithm, the hierarchical locally constrained diffusion process (HLCDP) algorithm proposed in this paper reduces the time consumption and improves the retrieval accuracy.

II. Related Work

Most of the diffusion-based methods have a general

978-1-5386-3120-1/17 $31.00 © 2017 IEEE
DOI 10.1109/ICBK.2017.45

framework [17]. First, we generally calculate the similarity matrices of pairwise images, and make a graph structure. Each image is a node of the graph, and the weight of the edge connected two nodes represents the similarity of the two nodes. Then a transition matrix is defined to represent the probability from one node to another. And next a diffusion process update scheme is defined for diffusing the similarity among images by adapting the update step. Finally, we can acquire the affinity matrix of all images.

Given a sample set of images $X=\{x_1^T, x_2^T, \ldots, x_n^T\} \in R^{n\times m}$, where n is the number of sample images and m is feature dimension of each image. A connected graph $G=(X,E)$ is constructed, where the nodes of the graph G represents the sample points and E represents the weights of the edges connecting the two nodes. The Gaussian kernel function is always used to compute the weights between two nodes $E(i,j)$ in the graph:

$$E(i\ ,\ j)=\exp\left\{-d^2(i\ ,\ j)/\left(2\sigma^2\right)\right\} \tag{1}$$

where $d^2(i\ ,\ j)$ represents the distance between image x_i and image x_j and σ is a parameter of Gaussian kernel function.

The transition matrix commonly used in the diffusion process is a random walk transition matrix [8], [10], [11], which is defined as

$$P=D^{-1}E \tag{2}$$

where D is a diagonal matrix with $D(i,i)=\Sigma^n_{j=1}E(i,j)$, and the transition matrix P is an non-symmetric matrix, and $\Sigma^n_{j=1}P(i,j)=1$.

In Global Page Rank (GPR) [18], the diffusion process is defined as

$$f_{t+1}=f_tP \tag{3}$$

where f_t is the ranking score vector after t steps of diffusion process. On the basis of Global Page Rank, the label propagation (LP) [8] combines semi-supervised learning to extend the diffusion process in (3) by fixing the query information $f(i)=1$ after each diffusion step. Therefore, the diffusion process of LP algorithm includes two steps: 1) $f_{t+1}=f_tP$; 2) $f_{t+1}(i)=1$ if x_i is a query image.

In the above diffusion process methods, calculating the transition probability from one image to another, we commonly consider all the paths between these two images. The influence of the noisy nodes is ignored completely, which will affect the diffusion performance. In order to solve this problem, LCDP algorithm [9] replaces the original global connected graph G with a K nearest neighbor graph G_K. According to the above graph G, we construct a K nearest neighbor graph $G_K(V,E_K)$, whose weights of the edges are defined as follows

$$E_K(i\ ,\ j)=\begin{cases}E(i\ ,\ j) & ,\ x_j\in KNN(x_i)\\ 0 & ,\ otherwise\end{cases} \tag{4}$$

and then the transition matrix is defined as

$$P_K(i\ ,\ j)=\frac{E_K(i\ ,\ j)}{\sum_j E_K(i\ ,\ j)} \tag{5}$$

Compared with the transition matrix P, the transition matrix P_K can reduce the effect of noisy nodes effectively. However there is a problem that the range of intersection between $KNN(x_i)$ and $KNN(x_j)$ is too small to lead two similar nodes that have no transition path. For this problem, the LCDP algorithm defined the diffusion process in a novel manner. Its main idea is to connect the point x_i and point x_j by finding two neighbor points k in $KNN(x_i)$ and l in $KNN(x_j)$, the definition is as follows

$$P_{KK}^{t+1}=\sum_{k\in KNN(x_i),l\in KNN(x_j)}P(x_i\ ,\ x_k)P_{KK}^t(x_k\ ,\ x_l)P(x_l\ ,\ x_j) \tag{6}$$

Equation (6) is expressed as a matrix:

$$P_{KK}^{t+1}=P_KP_{KK}^t(P_K)^T \tag{7}$$

The LCDP algorithm decreases the influence of noisy nodes by replacing the original connected graph with the K nearest neighbor graph. This algorithm solves the problem that there is no transition path between two similar images x_i and x_j when $KNN(x_i)$ of image x_i and $KNN(x_j)$ of image x_j have no intersection.

The existing diffusion process methods consume a large amount of time in the matrix operation, and require a large number of iterative calculations. With the increase of the size of the image database, the retrieval time will increase greatly. To address this problem, we propose a novel image retrieval method based on hierarchical locally constrained diffusion process (HLCDP). Our method combines the idea of algebraic multigrid and ranking scores on manifold learning. We construct the original retrieved images of database into a bottom-to-top hierarchical structure. Then we diffuse the similarity between query images and the top-layer representative images by exploiting locally constrained diffusion process method. Finally we interpolate to all images of bottom layer. Compared with the locally constrained diffusion process, the time consumption of our proposed method is reduced, and our method also improves the retrieval performance .

III. HIERARCHICAL LOCALLY CONSTRAINED DIFFUSION PROCESS

The diffusion process methods consider the influence of all the images in database on the similarity between the pairwise images. They exploit the intrinsic structure of the images database to improve the retrieval results. However, it takes a lot of time for matrix computation and computational iterative steps on the original images database directly. And a large number of iterations are required in the diffusion process. To improve the above problems, we propose a hierarchical locally constrained diffusion process (HLCDP) algorithm. This algorithm exploits the affinity values between pairwise images

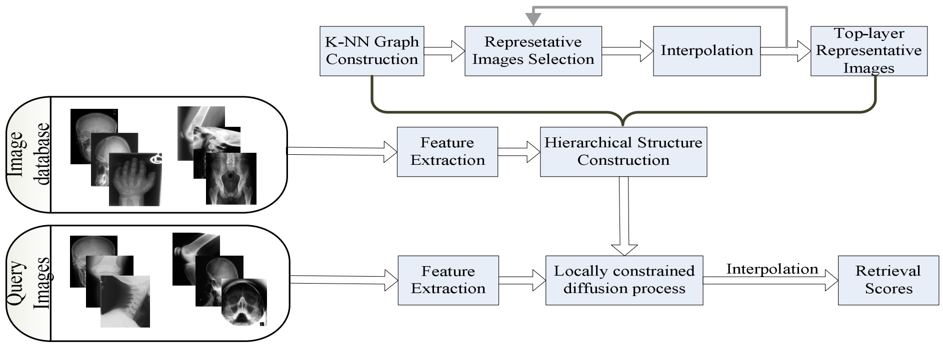

Figure 1. Framework of the proposed method

in the database to select the representative images and transforms the affinity values of the current layer to the next layer by interpolation matrix. Therefore, to transform the similarity among different layers, we construct the image database into a bottom-to-up hierarchical structure [14], [15], [16]. Then we use the LCDP [9] algorithm to obtain affinity matrix between query images and the representative images on the top layer. Last, the affinity matrix acquired on top layer is interpolated to the bottom layer and finally obtain the affinity matrix between query images and all the images. This method can reduce the time cost of high matrix computation and improve the retrieval performance. The framework of the proposed HLCDP is shown in Figure 1.

A. *Hierarchical Structure of Images*

To reduce the high computational cost, our method considers to construct the original one-layer structure of image database into hierarchical structure. First, we select the representative images from all of images on the current layer, which are regarded as candidate images on next layer. Next, the rest of the images can be selected in the same manner. This process forms a bottom-to-up pyramid, in which the most influential and representative images are at the top, and all images of database are at the bottom.

Let $X=\{x_1^T, x_2^T, \ldots, x_g^T, x_{g+1}^T, \ldots, x_n^T\} \in R^{n\times m}$ be a sample set of images, where n is the number of sample images and m is feature dimension of each image in the sample set. $X^L=\{x_1^T, x_2^T, \ldots, x_g^T\}$ represents the g retrieved images, and $X^U=\{x_{g+1}^T, x_{g+2}^T, \ldots, x_n^T\}$ represents the (n-g) query images. First, we construct the k nearest neighbor graph $G^{[0]}(V^{[0]}, W^{[0]})$, where $V^{[0]}$ represents the g images, and $W^{[0]}$ is the similarity matrix. The similarity matrix $W^{[0]}=[w_{ij}]_{g\times g}$ of all images in the database is determined as

$$W_{ij}^{[0]} = \begin{cases} \exp\left(-\alpha \left\| x_i - x_j \right\|_2\right) & , x_j \in KNN(x_i)\, or\, x_i \in KNN(x_j) \\ 0 \quad \cdots & , otherwise \end{cases} \tag{8}$$

where α is the parameter of the kernel function.

The representative images are selected to construct a bottom-to-up hierarchical pyramid structure, in which the non-representative are strongly connected with representative images. The definition of connection relationship between image x_i and image x_j is described as

$$W_{ij} \geq \theta\, Max_{k\neq i}\{W_{ik}\},\ 0 < \theta \leq 1 \tag{9}$$

where θ is the strength threshold, reflecting the strength of the relationship between neighboring images. Only the representative images are selected to participate in the subsequent operation and the non-representative images are ignored, which can reduce the amount of calculation.

The representative images selected of the next layer are represented as $V^{[s]} \in V^{[s-1]}$ (s=1,2,...) , which are strongly connected with non-representative images of current layer. In each layer, we establish an interpolation matrix to transform the similarity among representative images in current layer to next layer instead of computing the similarity matrix directly by (8). The interpolation matrix $Q^{[s-1]}$ from (s-1)-*th* layer to s-*th* layer is defined as

$$\begin{cases} Q_{ik}^{[s-1]} = \dfrac{W_{ik}^{[s-1]}}{\sum_k W_{ik}^{[s-1]}} & , i \notin V^{[s]}, k \in V^{[s]} \\ \begin{cases} Q_{ii}^{[s-1]} = 1 \\ Q_{ij}^{[s-1]} = 0 \end{cases} & , i \in V^{[s]}, j \neq i \end{cases} \tag{10}$$

As the representative images $V^{[s]}$ selected, the corresponding similarity matrix $W^{[s]}$ can be calculated by the interpolation matrix $Q^{[s-1]}$, the formula is as follows

$$W^{[s]} = Q^{[s-1]T} W^{[s-1]} Q^{[s-1]} \tag{11}$$

Thus, according to this strategy described as (11), the similarity matrix among images of each layer can be calculated by the following formula

$$W^{[s]} = \left(Q^{[s-1]}\right)^T \cdots \left(Q^{[2]}\right)^T \left(Q^{[1]}\right)^T \left(Q^{[0]}\right)^T W^{[0]} Q^{[0]} Q^{[1]} Q^{[2]} \cdots Q^{[s-1]} \tag{12}$$

B. *Diffusion Process on Top Layer*

As the selected representative images (counted as u) on top layer, if the distance among these images are calculated directly using the Gaussian kernel function, which may differ from the true distance, and is also double counting. Thus, we use the similarity matrix $W^{[s]}$ which has computed by (12) as the similarity among the top-layer representative images. Due to the large difference among the top-layer representative images, the performance of diffusion process will be affected by noisy image nodes. So we use the LCDP algorithm to diffuse the similarity among query images and the representative images on the top layer, which replaces the global connection graph with the K nearest neighbor graph to reduce the effect of noisy image nodes.

First, we directly use the similarity matrix $W^{[s]}$ among all top-layer images to construct a K nearest neighbor graph. The K nearest neighbor matrix is described as follows

$$W_K^{[s]}(x_i , x_j)=\begin{cases} W_{ij}^{[s]} & , x_j \in KNN(x_i) \\ 0 & , otherwise \end{cases} \tag{13}$$

Currently, the K nearest neighbor graph only contains the top-layer representative images, and reflects the relationship among the top-layer representative images. Considering the relationship between the query images and the top-layer representative images, we add the query images to the K nearest neighbor graph by exploiting the Euclidean distance to find the K nearest images of query images. For a query image $y_i(y_i \in X^U)$, the definition of its K nearest neighbor images $\{x_j, j=1,...,K\}$ that belong to the top-layer represented images $V^{[s]}$ is as follows

$$W_e(j)=\begin{cases} \exp\left[-d^2(x_i , x_j)/\sigma^2\right], & x_j \in KNN(y_i) \\ 0 & , otherwise \end{cases} \tag{14}$$

where σ is the parameter of the kernel function. To add the query image y_i to the K nearest neighbor graph of top-layer representative images, we construct a new weight matrix defined as

$$W_K = \begin{bmatrix} W_K^{[s]} & W_e \\ W_e^T & 1 \end{bmatrix} \tag{15}$$

where W_e is a u-dimension vector.

The transition matrix $P_K(i,j)$, which represents the probability of diffusion from one image x_i to another image x_j, is defined

$$P_K(i , j)=\frac{W_K(i , j)}{\sum_j W_K(i , j)} \tag{16}$$

For diffusing the similarity among K nearest neighbor graph, we utilize the update strategy in LCDP [11] algorithm, which is defined by update as follows

$$W_{t+1} = P_K W_t P_K^T \tag{17}$$

where the initialization of our diffusion process is W_K. Let W^* be the final result when the (17) convergences, where the first u values of the vector W^*_{u+1} which is the $(u+1)$-*th* row of matrix W^* corresponds to the affinity vector represented as $F_i^{[s]}$ between the query image y_i and all top-layer representative images after diffusing. The affinities between the rest of the query images and top-layer representative images can be obtained in the same manner. The affinity matrix of the $(n-g)$ query images and the u top-layer representative images is represented as $F^{[s]}$.

C. Similarity Transformation among Top-to-bottom Hierarchical Structure

After obtaining the affinity matrix $F^{[s]}$ between the query images and the top-layer representative images on the top layer, we need to interpolate the affinity matrix $F^{[s]}$ to all images in bottom layer along the top-to-bottom hierarchical structure by utilizing interpolation matrix Q. The interpolation formula is defined as follows

$$F^{[s-1]} = F^{[s]}\left(Q^{[s-1]}\right)^T \tag{18}$$

where $F^{[s]}$ represents the affinity matrix between the query images and the representative images in current layer, and $F^{[s-1]}$ represents the affinity matrix between the query images and the representative images in the next layer. The affinity matrix between the query images and the representative images of each layer can be process in this strategy. So the affinity matrix between the query images and all the images in the images database can be obtained by

$$F^{[0]} = F^{[s]}\left(P^{[s-1]}\right)^T\left(P^{[s-2]}\right)^T\cdots\left(P^{[1]}\right)^T\left(P^{[0]}\right)^T \tag{19}$$

D. Summary of HLCDP Algorithm

In this paper we propose a hierarchical local constrained diffusion process algorithm. It constructs a hierarchical structure for all images in database. Only the representative images of top layer are used with query images to diffuse similarity by LCDP algorithm. The affinity matrix obtained on top layer is interpolated to the bottom layer and finally we obtain the affinity matrix between query images and all the images. The algorithm of the proposed method is summarized as follows:

Algorithm: Hierarchical Locally constrained Diffusion Process (HLCDP)

Input: The image database $X^L=\{x_1{}^T,x_2{}^T,\ldots,x_g{}^T\}\in R^{g\times m}$
The query images $X^U=\{x_{g+1}{}^T,x_{g+2}{}^T,\ldots,x_n{}^T\}\in R^{(n-g)\times m}$
Output: Corresponding retrieved images of the query images

Step1. Construct a k nearest neighbor graph $G^{[0]}=(V^{[0]},W^{[0]})$ of images X^L, $V^{[0]}$ represents all images of the images database, $W^{[0]}$ using (8) represents a similarity matrix among all images in the images database ;
Step2. Build a bottom-to-top hierarchical structure of all images in the images database ;
for s=1:n
(1) For the representative images $V^{[s-1]}$ of upper layer, select the representative images $V^{[s]}$ of current layer using

(9), and the representative images $V^{[s]}$ are treated as candidate images of the next layer ;

(2) Calculate the interpolation matrix $Q^{[s-1]}$ from (s-1)-*th* layer to *s-th* layer using (10) ;

(3) Update the representative images $V^{[s]}$ corresponding similarity matrix $W^{[s]}$ using (11) ;

end for

Step3. Obtain the affinity matrix $F^{[s]}$ of query images and representative images on top layer ;

for i=1:*(n-g)*

(1) Obtain the K nearest neighbor similarity matrix among the top-layer representative images using (13) ;

(2) Add the query image $x_i(x_i\in X^U)$ to the graph of top-layer representative images, and construct a new weight matrix W_K using (14) ;

(3) On the top layer, the similarly among query image and representative images are diffused using (17) ;

end for

Step4. Obtain the final affinity matrix $F^{[0]}$ between the query images and all retrieved images in database by the interpolation matrix Q;

for s=n:1

$$F^{[s-1]} = F^{[s]}\left(Q^{[s]}\right)^T$$

end for

Step5. Output corresponding retrieved images of the query images by ranking $F^{[0]}$, where $F_i^{[0]}$ correspond the ranking scores of query image $x_i(x_i\in X^U)$.

E. Analysis of HLCDP Algorithm

As our proposed HLCDP algorithm, whose diffusion process is based on a hierarchical structure, the major cost is in the calculation of ranking scores between top-layer representative images and query images and interpolating the ranking scores to bottom layer. The time complexity of calculating ranking scores of images on top layer is $O(N^{[C]}mt_A)$, where m is the feature dimension of each image, t_A is the number of iterations, and $N^{[C]}$ is the number of the top-layer representative images. The time complexity of interpolating the ranking scores from *s-th* layer to (s-1)-*th* layer is $O(N^{[s]}N^{[s-1]})$, where $N^{[s]}(s=1,2,...,C)$ is the number of the representative images on the *s-th* layer, and C denotes the layer number. Thus, the time complexity of HLCDP algorithm is $O(N^{[C]}mt+N^{[C]}N^{[C-1]}+N^{[C-1]}N^{[C-2]}+\cdots+N^{[2]}N^{[1]})$. For LCDP algorithm, whose diffusion process is on the original all images of the image database, the complexity is $O((N^{[1]})^2mt_B)$, where t_B is the iteration number.

The images on the first layer are all retrieved images in database. There is $N^{[C]}<N^{[C-1]}<\cdots<N^{[1]}$. The time complexity of HLCDP satisfies

$$\begin{aligned}&O\left(N^{[C]}mt_A+N^{[C]}N^{[C-1]}+N^{[C-1]}N^{[C-2]}+\cdots+N^{[2]}N^{[1]}\right)\\&<<O\left(N^{[C]}mt_A\right)+(C-1)O\left(\left(N^{[1]}\right)^2\right)<O\left(\left(N^{[1]}\right)^2mt_B\right)\end{aligned} \tag{20}$$

From above (20), our HLCDP algorithm can reduce the time complexity effectively.

IV. EXPERIMENTS

To demonstrate the validity of HLCDP, we perform experiments on image retrieval in this section. In the experiments, we mainly use two common evaluation criteria which are precision rate and recall rate to evaluate the performance of proposed method, and carry out LCDP [9], SSO [10] and SD [11] as comparative methods. In this paper, the *MPEG7* data set [17] and *ImageCLEFmed2005* data set [19] are used in image retrieval.

A. Experiments on MPEG7 data set

The *MPEG7* data set contains 1,400 silhouette images divided into 70 shape classes. There are 20 shapes in each class. Every shape in the database is submitted as a query. Retrieval accuracy is measured by bull's eye score. Every shape in the database is submitted as a query, and the top 40 shapes are reported for each shape.

In the experiments, we use the shape matching algorithm in [21] to compute the pairwise distances between these shapes. The size of nearest neighbors k is set as 15. To make the *s-th* layer' representative images be approximately half of the (s-1)-*th* layer' images, we set strength threshold θ to 0.99. We set the layers of hierarchical structure to 4. As a result, the 1400 shape images are built into 4 hierarchical structure. The number of images on each layer are 1400, 889, 574 and 384 from bottom-to-top respectively. On the top layer, different size of K results different retrieval performance. To analyze the influence of different K values on retrieval results, we set K from 1 to 15 in the experiments. Bull's eye scores of different K on the *MPEG7* data set are shown in Figure 2.

As can be seen in Figure 2, HLCDP method's retrieval performance is sensitive to the size of K. From K=1, The bull's eye scores are improved with the increasing of K. From K=10, the bull's eye score is 100%. To obtain the best retrieval performance, we fix K to 10 in the next experiments on the *MPEG7* data set.

The proposed HLCDP algorithm is compared with the LCDP, SSO and SD to test validity of HLCDP. The results of four methods' bull' eye scores are shown in Table 1. From the Table 1, it can be seen HLCDP algorithm and LCDP algorithm both have the best retrieval performance with 100% bull' eye score while the bull' eye score of SSO and SD are 99.87% and 99.89% respectively on *MPEG7* data set. In contrast, the proposed HLCDP algorithm has better performance. This shows that our proposed is effective.

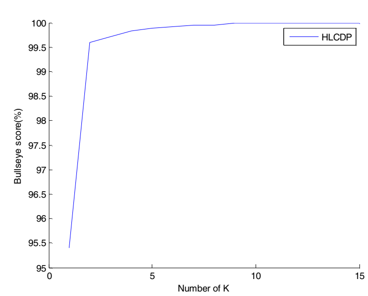

Figure 2. Bull's eye scores of different K on the MPEG7 data set

TABLE 1. COMPARISON OF RETRIEVAL PERFORMANCE ON MPEG7

Method	Bull's eye score(%)
AIR[23]	93.67
AIR+SSO	99.87
AIR+SD	99.89
AIR+LCDP	100
AIR+HLCDP	100

B. Experiments on ImageCLEFmed2005 data set

The above experiments are done on images that the query images are inside the database. In this section, we do experiments with the query images that are outside the database on *ImageCLEFmed2005* data set. In total, this data set contains 57 categories, while there are only categories that have over 110 images. We select randomly 110 images of each category from these categories. The first 10 images of each category are selected as the test set. The rest 1,500 images are used as the training set. Then the images of test set and training set are scaled to a size of 128 * 64, and these images are extracted the visual features by using histogram of Oriented Gradient (HOG) [20]. The hierarchical structure of the image is set to 4 layers. The strength threshold θ for each layer selecting representative images is set to 0.95. The value k of k nearest neighbor graph constructed of all images on bottom is set to 20.

In the experiments, the bottom layer of this structure is the 1500 training images. The 1500 training images are built into 4-layer hierarchical structure. The number of images on each layer are 1500, 890, 564 and 377 from bottom-to-top respectively. On the top layer, we diffuse the similarity between 150 query images and 377 representative images by using LCDP to obtain the affinities. The retrieval accuracy is different when the value K of K nearest neighbor graph changes. We set K to 10, 20, 30, and 40 to observe the influence of different value K on retrieval precision.

As can be seen from Figure3, the HLCDP shows the best retrieval precision when the value K is 20. In the next experiments, we fixed K to 20. In diffusion process methods existing, the number of iterations t also has an important effect on the retrieval results. The number of iterations is too large or small, which will lead to the decrease of the retrieval accuracy. Figure4 shows the relationship between the number of iterations and the precision of the proposed method and the other three methods for comparing respectively when the number of top retrieved samples is 50.

Figure4 shows the evolution of precision over different range of t in the retrieval task. Obviously, the HLCDP method' retrieval results is the best when the number of iterations t is 40. But LCDP algorithm, SD algorithm and SSO algorithm diffuse similarity over large t. The retrieval performances are increasing until the number of iterations of these three methods are around 4000, 3000 and 600 respectively. Different from LCDP algorithm, SD algorithm and SSO algorithm whose diffusion process are performed in the original 1500 images and 150 query images, our HLCDP algorithm builds the 1500 images into a four-layer hierarchical structure in which there are 337 representative images on the top layer. We only do diffusion process computation between the 337 representative images and 150 query images. With the images decreasing for diffusion process, the number of iterations also reduced.

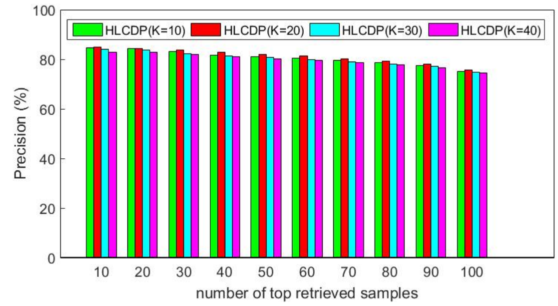

Figure 3. The relationship between the precision and the number of top retrieved samples when different K values are set (K=10,20,30,40)

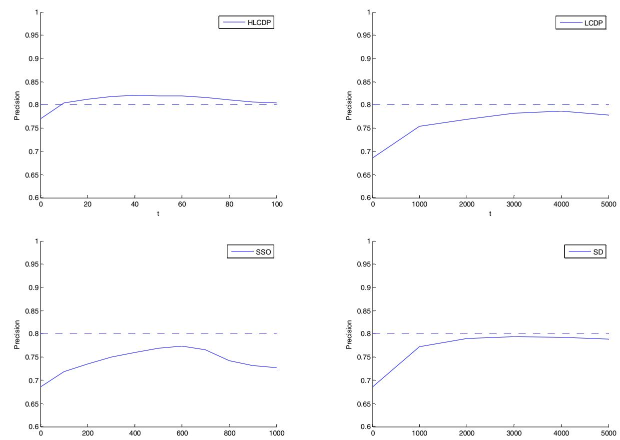

Figure 4. When the number of top retrieved samples is 50, precision of four methods over t

The proposed method is compared with the LCDP, SD and SSO algorithms, and our experiments evaluate these four methods from the precision and recall rate, which are shown in Figure 4 and Figure 5. It can be seen from Figure 4 the precision and recall rate of our method are lower than those of the other three methods when the number of top retrieved samples is 10. But with the increasing of the number of top retrieved samples, the precision and recall rate of our method (84.47% , 16.89%) are approximate to those of LCDP (84.47% , 16.93%) and still lower than that of SD (88.33% , 17.67%) and SSO (86.33% , 17.27%). When the number of top retrieved samples is 30, the precision and recall rate of our method (83.78% , 25.13%) are higher than that of LCDP (82.18% , 24.65%) and SSO (82.84% , 24.85%), and slightly lower than SD (84.6% , 25.38%). Since the number of top retrieved samples is 40, the precision and recall rate of our method (82.73% , 33.09%) are higher than the other three methods which are LCDP (80.32% , 32.13%), SSO (80.1% , 32.04%), SD (81.93% , 32.77%). With the number of top retrieved samples increasing from 40 to 100, the retrieval performance of our proposed is always better than other three methods, as shown in Figure5 and Figure 6. These experiments can verify the effectiveness and practicability of HLCDP in image retrieval.

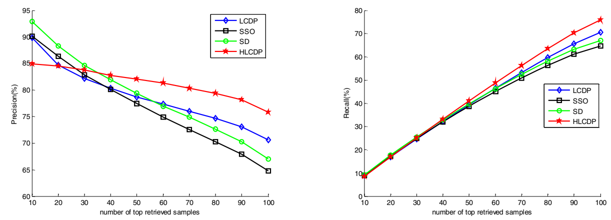

Figure 5. Precision (left) and recall (right) of our proposed method and other three methods on the ImageCLEFmed2005 data (*K*=20)

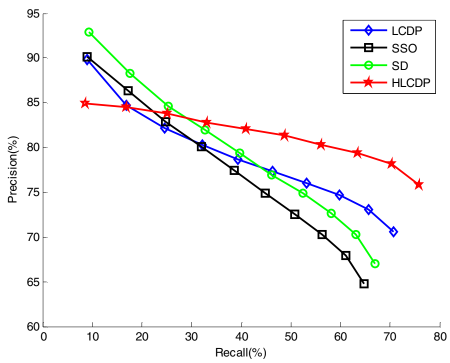

Figure 6. Precision recall curve on the ImageCLEFmed2005 data (*K*=20)

V. CONCLUSION

In this paper we propose an image retrieval method based on hierarchical locally constrained diffusion process, which is derived from diffusion ranking on manifold with the use of algebraic multigrid to solve the high matrix calculations and the large amount of iterations computation in traditional diffusion process methods. The proposed method builds all retrieved images into a bottom-to-top hierarchical structure. The affinities between query images and top-layer images are obtained by LCDP, and then these affinities are interpolated to the bottom layer. Finally we obtain the affinities between the query images and all retrieved images in database. Our HLCDP algorithm is compared with LCDP, SSO, SD on *MPEG7* data set and *ImageCLEFmed2005* data set. The experimental results demonstrate that our approach not only have advantage on reducing computation cost but also improve the retrieval accuracy. In reality, some images of the images data set are labeled. For making global use of the labeled information to improve retrieval performance further, we consider to combine our proposed method with semi-supervision learning in future.

ACKNOWLEDGMENT

This work was supported by the National Key Research and Development Program of China (Grant no. 2016YFB1000905), the National Natural Science Foundation of China (Grant no. 61672120), the Chongqing Natural Science Foundation Program (Grant no. cstc2015jcyjA40036) and Postgraduate Scientific Research and Innovation Project of Chongqing (Grant no. CYS15173).

REFERENCES

[1] S. T. Roweis, and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," Science,vol. 290, no.5500, pp. 2323-2326, 2000.

[2] J. B. Tenenbaum, V. De. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, vol. 290, no.5500, pp. 2319-2323, 2000.

[3] H. S. Seung, and D. D. Lee, "The manifold ways of perception," Science, vol. 290, no. 5500, pp. 2268-2269, 2000.

[4] J. R. He, M. J. Li, and H. J. Zhang, "Manifold-ranking based image retrieval," ACM International Conference on Multimedia. ACM, pp. 9-16, 2004.

[5] J. R. He, M. J. Li, and H. J. Zhang, "Generalized manifold-ranking-based image retrieval," IEEE Transactions on Image Processing, vol. 15, no. 10, pp. 3170-3177, 2016.

[6] F. Wang, G. Er, Q. Dai. "Inequivalent manifold ranking for content-based image retrieval," pp. 173-176, 2008.

[7] B. Xu, J. J. Bu, and C. Chen, "Efficient manifold ranking for image retrieval," 34th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, pp. 525-534, 2011.

[8] X. Zhu, Z. Ghahramani, and T. J. Mit. "Semi-Supervised Learning with Graphs," International Joint Conference on Natural Language Processing. pp. 2465 - 2472, 2005.

[9] X. Yang, S. Koknar-Tezel, and L. J. Latecki. "Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval," IEEE Conference on Computer Vision & Pattern Recognition, pp. 357-364, 2009.

[10] J. Jiang, B. Wang, and Z. Tu. "Unsupervised metric learning by Self-Smoothing Operator," IEEE International Conference on Computer Vision. IEEE, pp. 794-801, 2011.

[11] B. Wang, and Z. Tu. "Affinity learning via self-diffusion for image segmentation and clustering," IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, pp. 2312-2319, 2012:

[12] J. Wang, Y. Li, X. Bai, Y. Zhang, C. Wang, and N. Tang, "Learning context-sensitive similarity by shortest path propagation," Pattern Recognition, vol. 44, no. 10, pp. 2367-2374, 2011.

[13] A. Brandt, S. Mccormick, and J. Ruge, "Algebraic multigrid (AMG) for sparse matrix equations," Sparsity & Its Applications, vol.13, no. 3, pp. 257-284, 1984.

[14] M. Galun, E. Sharon, R. Basri., and A. Brandt, "Texture Segmentation by Multiscale Aggregation of Filter Responses and Shape Elements," IEEE International Conference on Computer Vision, vol. 3, pp. 716-723, 2003.

[15] E. Sharon, A. Brandt, and R. Basri, "Fast multiscale image segmentation," Computer Vision and Pattern Recognition, vol. 1, pp. 70-77, 2000.

[16] E. Sharon, M. Galun, D. Sharon, R. Basri, and A. Brandt, "Hierarchy and adaptivity in segmenting visual scenes," Nature, vol. 442, no. 7104, pp. 810-813, 2006.

[17] M. Donoser, and H. Bischof, "Diffusion Processes for Retrieval Revisited," Computer Vision & Pattern Recognition, pp. 1320-1327, 2013.

[18] L. Page, S. Brin, R. Motwani., and T. Winograd, "The PageRank citation ranking: Bringing order to the web," Stanford InfoLab, 1999.

[19] A. Rorissa, P. Clough, W. Hersh, A. Rorissa, and M Ruiz, "ImageCLEF and imageCLEFmed: Toward standard test collections for image storage and retrieval research," Proceedings of the American Society for Information Science & Technology, vol. 43, no. 1, pp. 1-6, 2007.

[20] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," Computer Vision and Pattern Recognition, vol. 1, pp. 886-893, 2005.

[21] R. Gopalan, P. Turaga, and R Chellappa, "Articulation-invariant representation of non-planar shapes," European Conference on Computer Vision Conference on Computer Vision, vol. 6313, pp. 286-299, 2010.