

Manifold-ranking Based Image Retrieval Using Natural Neighbor

Qingsheng Zhu, Zhi Chen and Cheng Zhang

*Chongqing Key Lab. of Software Theory and Technology
College of Computer Science, Chongqing University, Chongqing 400044, China
qszhu@cqu.edu.cn, 20121402014t@cqu.edu.cn, bootan@cqu.edu.cn*

Abstract

The manifold-ranking based method is widely used in semi-supervised learning, and its performance is closely related to the structure of the constructed graph. In this paper, we propose a novel graph structure named natural neighbor graph and an algorithm to construct it. We apply the new graph structure into the framework of manifold-ranking based image retrieval. The greatest superiority over k -NN based method is that the free parameter k need not to be explicitly specified any more. We have shown that the manifold ranking algorithm based on our proposed graph structure performs better than k -NN graph. Experiments demonstrate that the proposed algorithm outperforms the state-of-the-art algorithms.

Keywords: manifold-ranking; natural neighbor; image retrieval; k -NN

1. Introduction

Learning-based algorithms have become the mainstream in image, 3D model and other multimedia retrieval technologies [1], and they aim at solving the well-known "semantic gap" between low-level features and high-level semantic concepts. Since Tenenbaum [2] and Roweis [3] proposed the manifold learning framework, the semi-supervised learning (SSL) methods have been widely studied. By leveraging unlabeled data with certain assumptions, SSL is promising to build more accurate models than the methods which are achieved by purely supervised learning methods, because only a few labeled data can be obtained in practice [4]. As a major family of SSL, graph based SSL has become one of the most active research areas in semi-supervised learning community [5].

In 2004, Zhou [6-7] proposed a novel semi-supervised learning algorithm named manifold-ranking. In this algorithm, a connected graph is created first, in which vertex represents the data and edge represents the similarity between vertices, and then the score diffuses from the vertices to their neighbors. After several rounds, all the vertices get stable scores. Finally, all the data is ranked by the scores of the corresponding vertices in the graph. The assumption of the algorithm is that all the data are distributed in a low dimension manifold, which is embedded in the high dimension features space. Compared to the pairwise method, the main difference of the manifold-ranking based method is, it ranks all the data according to the manifold structures represented by the labeled and unlabeled data, which means that the algorithm ranks the data by considering local and global consistency simultaneously.

The performance of manifold-ranking depends on the quality of the constructed graph[8]. In this paper, we investigate various aspects of graph structures, which may affect the algorithm performance, and propose a new graph structure named natural neighbor (2N) graph as well as a corresponding constructing algorithm. Compared to k -NN or like k -NN based algorithms, the most difference of 2N based algorithms is that it need not to specify the parameter k , and meanwhile the number of neighbors of each vertex is variable which was first proposed in reference[9]. Experiments show that 2N is more suitable for manifold-ranking based image retrieval, 2D data clustering than the

original connected graph presented in paper[7] and k-nearest neighbor (k-NN) graph used in paper [8].

The main contributions of this paper can be summarized as follows:

- Propose the 2N graph and its corresponding constructing algorithm, so as to present the data manifold structure more effectively.
- Apply the new graph structure successfully into semi-supervised learning framework into image retrieval

The remainder of this paper is organized as follows. Related work is briefly reviewed in Section 2. The 2N graph and its corresponding constructing algorithm are described in Section 3 in detail. The experiment results on image retrieval are shown and discussed in Section 4. Finally, we conclude our method and discuss future work in Section 5.

2. Related Works

Based on Zhou's work [6-7], a lot of manifold-ranking based algorithms were proposed, and have been applied to image retrieval, document retrieval, 3D model retrieval as well as video semantic concept detection. For example, He [8,10] proposed a manifold-ranking approach to propagate the label information of image samples from positively labeled samples to unlabeled ones. Cai [11] created a model to enhance manifold-ranking based relevance propagation via mutual reinforcement between sentences and theme clusters of document. Furuya [12] introduced cross-domain manifold ranking for sketch-based image and 3D model retrieval. In paper [13], A manifold ranking method on key frames of video of semantic brain imaging space was proposed for video annotation.

Some modified versions of manifold-ranking based algorithms were proposed to improve the performance from different perspectives and were applied in different applications. Gui [14] added the label information, and adopted the geodesic distance instead of Euclidean distance as the measurement of differences between data points. To reduce the computational complexity, Zhang [15] proposed a multilevel scheme for speeding up the traditional manifold-ranking based method. In video semantic analysis, Tang [16] proposed the structure-sensitive manifold ranking (SSMR), which takes the local distribution differences into account to measure the pairwise similarity more accurately. Xia [17] presented the smooth consistency and constraint consistency to which an affinity graph must hold, and then defined the affinity graph respecting these consistencies in a regularization framework of manifold-ranking. This method is applicable to both unsupervised and semi-supervised spectral clustering.

Although the research of manifold-ranking is active, there is still limited work on the graph structure and its corresponding constructing method. If the graph cannot represent the manifold effectively, the algorithm performance is likely to fall down. Zhou's method employed the connected graph to represent the manifold structure. It first sorts the pairwise distances among points in ascending order, and then repeats connecting two points according to the ascending order until a connected graph is obtained. He [8] adopted k-NN graph instead of connected graph. It calculates the k-nearest neighbors for each point, and then connects two points if they are neighbors. The purpose of this modification is to ensure enough connection for each point while preserving the sparse property of the weighted graph as well. Another graph structure is called ε -NN graph. It connects two points if the distance between them is less than a user-defined threshold ε .

Several new graph structures were proposed in other applications. For instance, to solve the nonlinear data projection problem based on geodesic distance, Yang proposed the k-edge connected neighborhood graph, and presented two constructing method named k-MST [18] and Min-k-ST [19]. k-MST works by repeatedly extracting minimum spanning trees from the complete Euclidean graph of all data points. To make the set of extracted edges have a minimum total length and the graph is k-edge connected, Min-k-ST adopts a greedy algorithm which adds edges in a non-decreasing order of edge

length to the neighborhood graph if the end vertices of each edge are not yet k -edge-connected. Both of these two constructing methods are proven to be better than k -NN graph and ε -NN in geodesic distance computation.

3. Natural Neighbor Graph

3.1. The Definition of Natural Neighbor Graph

Besides complying with the constraints of graph theory, the graph structures in manifold-ranking must have explicit geometric meanings. In such constraints, the types of graph structures suitable for manifold learning are limited. k -edge connected neighbor graph and connected graph works well in geodesic distance computation. However, while focusing on guaranteeing the connectivity of graph, it does not pay enough attention to the whole graph manifold. k -NN graph and ε -NN graph are easy to construct, but the difficulty of finding a fitting parameter makes it impractical in manifold-ranking based methods. Take the toy data in 2D feature space for example. Figure 1 shows two moon data. The ideal manifold structure should be: the upper moon data is classified in one cluster and the lower moon data in another.

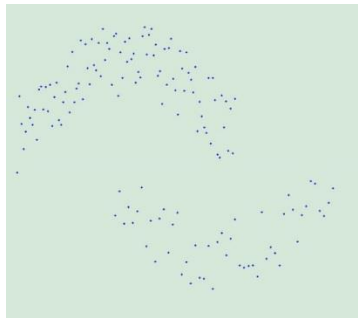


Figure 1. The Toy Data in 2D Feature Space

Figure2(a) shows that the graph structure which is constructed by k -NN based method where k equals to 5 is the ideal graph structure. However, when k is assigned to 15, there will be another situation as shown in Figure2(c). Figure2(b) shows the scores of each vertex after manifold ranking method run on the graph shown in Figure2(a). We use a hollow circle to represent a vertex in the figure. The circle within a red cross is the query point, and the diameter of each circle is proportional to the ranking score of its corresponding vertex if the ranking score is not zero, otherwise we use a point instead of circle to represent vertex. We can see that all vertices of the upper moon data get different non-zero scores and all vertices of the lower moon data get zero scores in Figure2(b) where k equals to 5 and it meets our common knowledge and experience. However, when k equals to 15, we can see that not only the upper moon data but also the lower moon data get non-zero scores and it is against our common sense. In all, k -NN graph have some limitations, and cannot represent the manifold structure well in some extent because of the free parameter k .

To reserve the advantages of connected graph and k -NN graph, and avoid their disadvantages, we propose a novel graph structure named Natural Neighbor (2N) graph. The idea of 2N is inspired from the real world observations that the neighbors should be acceptable by each other, similar to the friendship relations between individuals in social science psychology, naturally, some persons have more friends whereas some persons have few friends, the number of one's "true friends" should be determined by the number of how many other people are taken him or her as a friend, rather than the number of one's subjectively taken as. We denote the number of one's "true friend" as nb . 2N graph

is an undirected graph and a subgraph of the complete graph, in which each vertex represents a data point in feature space χ , and we firstly look for these data points who are the one to nb_i nearest neighbor of data point x_i , then connect them with data point x_i and the edge between two vertices x_i and x_j if they are connected represents their corresponding distance. The distance $d(x_i, x_j)$ can be any measure in the feature space, such as Manhattan distance (L1), Euclidean distance (L2), Mahalanobis and Chebychev (L_∞) distances.

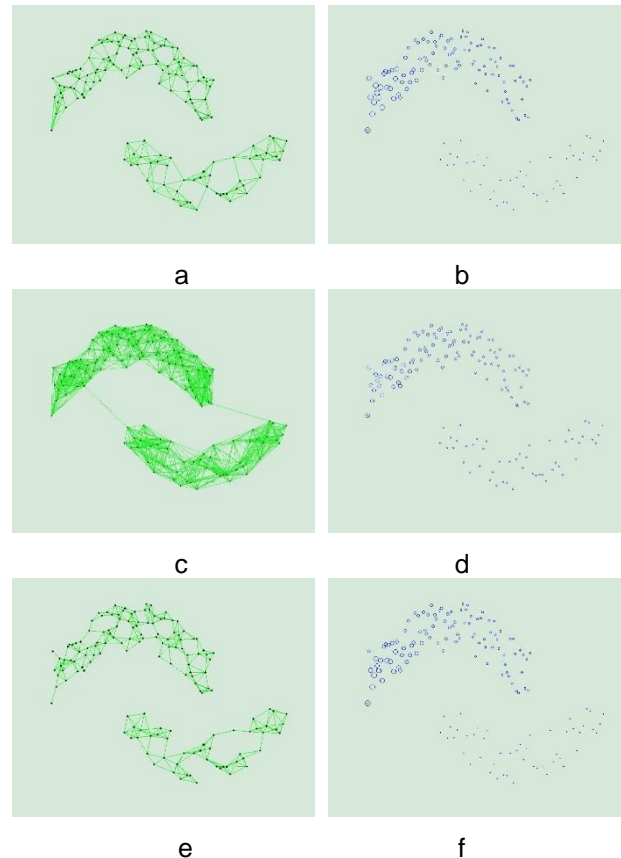


Figure 2. Comparison to Different k-NN Graph Based and 2N Graph Based

2N graph is more flexible than k-NN graph, for it does not constraint the degree of each vertex strictly to k , and in this case, the graph structure is more likely to capture the inherent manifold structure. Fig 2(a) and (c) have explained that the different k have quite different result. However Fig 2(e) has showed that 2N graph is an ideal graph the same as 5-NN graph, but it doesn't need to specify the free parameter k . In other words, the process of constructing 2N graph is automatic. The ranking result shown in Figure 2(f) gives more evidence: vertices closer to query vertex get higher scores within the upper moon data, and the vertices in lower moon data get zero scores, which is in accordance with our intuition. Although this is only a simple 2D dimension toy demo, and the feature spaces for image retrieval are usually in much higher dimension, we found that 2N graph outperforms k-NN graph distinctly in revealing manifold structure, and achieves better performance in manifold-ranking based retrieval, as demonstrated in Section 5.

3.2 The Constructing Method of 2N Graph

In 2N graph, the degree of each vertex is relative to its position comparing to other vertices. For example, if a vertex is too far away from all other vertices, it may be isolated. If a vertex is in a dense area, then it may have higher degree, because it is close to other vertices, and in this case they should be in the same cluster in high confidence. In this

paper, we design a greedy algorithm to approximately construct the proposed 2N graph, as illustrated in **Algorithm 1**.

Algorithm 1: Constructing algorithm of 2N graph

Input: data point set X
Output: the constructed 2N graph G

- 1 Initialize the 2N graph G without edges
- 2 Define integer value $r = 1$
- 3 Define integer vector $nb = 0$
- 4 Define boolean value $finding = true$
- 5 **while** $finding$ **do**
- 6 **for each** data point i **in** X **do**
- 7 Look for the r -th nearest neighbor of data point i and denote it as j .
- 8 Let $nb(j) = nb(j) + 1$
- 9 **end**
- 10 **if** $\forall i \in X$ $nb(i) \neq 0$ **then**
- 11 Let $finding = false$
- 12 **else**
- 13 Let $r = r + 1$
- 14 **end**
- 15 **end**
- 16 **for each** data point i **in** X **do**
- 17 **for each** $j=1$ **to** $nb(i)$ **do**
- 18 Look for the j -th nearest neighbor of data point i and denote it as k .
- 19 Add edge e_{ik} to G
- 20 **end**
- 21 **End**

We don't need to specify the free parameter k in this algorithm and this is main difference compared to k-NN based constructing algorithm. Another difference between our algorithm and k-NN constructing method is that we try to make the degree of each vertex to be proportional to its density, rather than constraining each vertex to be strictly k edge. Let denote size of data set X as s . The algorithm terminates when all nb doesn't equal to zero. This condition is certain to be satisfied because each point of G has the same the number of neighbors *i.e.* the number of neighbor of each point is $s-1$ when $r=s-1$ and this is the most terrible situation where not only it reach the worst time complexity, but also it has the most worse graph structure *i.e.* the constructed graph is a complete graph. Generally, this situation happens when there is an outlier point exists in the given data set X .

To cope with the most worse terrible situation mentioned above, we propose Algorithm 2 to improve Algorithm 1.

Algorithm 2: Constructing algorithm of 2N graph without outlier

Input: data point set X
Output: the constructed 2N graph G .

- 1 Initialize the 2N graph G without edges
- 2 Define integer value $r = 1$
- 3 Define integer vector $nb = 0$
- 4 Define boolean value $finding = true$
- 5 Define integer value $pre = 0$
- 6 Define integer value $cur = 0$
- 7 **while** $finding$ **do**

```

8   for each data point  $i$  in  $X$  do
9       Look for the  $r$ -th nearest neighbor of data point  $i$  and denote it as  $j$ .
10      Let  $nb(j) = nb(j) + 1$ 
11  end
12  Let  $pre = cur$  and  $cur$  equals to the number of the non-zero values in  $nb$ 
13  if  $\forall i \in X$   $nb(i) \neq 0$  or  $pre == cur$  then
14      Let  $finding = false$ 
15      Let  $pre = cur$ 
16  else
17      Let  $r = r + 1$ 
18  end
19  end
20  if  $\exists i \in X$   $nb(i) == 0$  then
21      Let  $nb = nb + 1$ ;
22  end
23  for each data point  $i$  in  $X$  do
24      for each  $j=1$  to  $nb(i)$  do
25          Look for the  $j$ -th nearest neighbor of data point  $i$  and denote it as  $k$ .
26          Add edge  $e_{ik}$  to  $G$ 
27      end
28  end
    
```

Algorithm 2 terminates if either of the two condition is satisfied: the first one is that there is no zero value exists in nb . This means that each vertex of X is taken as the neighbor of at least one other vertex. The second condition is that pre equals to cur . The variable cur and pre present respectively the number of the non-zero values in nb of current and previous loop. When $cur = pre$, we consider that the nb has become stable in high confidence. To guarantee the outlier if it exists is connected by at least one vertex, we should make sure that there is no zero exists in nb . Therefore we let $nb=nb+1$ if there is zero exists in nb .

3.3. The Manifold-Ranking Method Based on 2N

Zhou's original manifold-ranking algorithm [6-7], which takes the global information of the whole labeled and unlabeled dataset into consideration, is widely studied. However, it does not perform well in some cases due to the rough graph structure. In this part, we first apply our 2N graph structure in Zhou's method, and then analyze its validity.

Given a set of data points $= \{x_1, \dots, x_q, x_{q+1}, \dots, x_n\} \subset \mathbb{R}^m$, the first q points are the queries and the rest are the points that we want to rank according to their relevance to the queries. Let $\chi \times \chi \rightarrow \mathbb{R}$ denote a metric on χ , which assigns each points pair x_i and x_j a distance $d(x_i, x_j)$. Let $f: \chi \rightarrow \mathbb{R}$ denote a ranking function that assigns to each point x_i a ranking value f_i . We also define a vector $y = [y_1, \dots, y_n]^T$, in which $y_i=1$ if x_i is a query, and $y_i=0$ otherwise. The modified manifold-ranking algorithm based on 2N is illustrated in Algorithm 3.

Algorithm 3: Manifold-ranking algorithm based on 2N

Input: data point set X

Output: ranking score f_i^*

- 1 Calculate the pairwise distance $d(x_i, x_j)$ of all points in X
 - 2 Construct the 2N graph using **Algorithm 2**
 - 3 Form the affinity matrix W defined by $w_{ij} = \exp(-d^2(x_i, x_j)/(2\sigma^2))$
 - 4 Symmetrically normalize W by $S = D^{-1/2}WD^{1/2}$, in which D is the diagonal matrix with (i,i) -element equal to the sum of the i -th row of W
 - 5 Iterate $f(t+1) = \alpha Sf(t) + (1-\alpha)y$ until convergence, where α is a parameter in $[0,1)$
 - 6 Let f_i^* denote the limit of the sequence $f_i(t)$. Rank each point x_i according to its descend ranking score f_i^*
-

As stated in papers [20], defining a suitable affinity matrix W is of key importance. A commonly used distance function $d(x_i, x_j)$ is the L2 distance, which results in a Gaussian kernel for defining edge weights in W . However, based on the experimental results in [21], we can draw a conclusion that L1 distance can better approximate the perceptual difference between two images than other popular Minkowski distances when using color representation.

4. Experimental Result

To test the efficiency of our proposed graph structure in manifold-ranking based image retrieval, we employed the Corel dataset. Corel dataset is publicly available and widely used in evaluating image retrieval [22] and annotation methods recently [23]. Our version contains about 5000 small Corel photos. For the ground truth of the Corel dataset, there are total 50 categories, and in each category, there are about 100 images. We use each image in the whole database as a query, and average the results over the 5,000 queries. The precision vs. recall curve is used to evaluate the performance of various methods.

Feature selection is a large open problem and might have a great impact on the results. In our current version, the feature vector is simply made up of color histogram [19] since we focus on the relative performance comparison. Color histogram is obtained by quantizing the HSV color space into 128 bins. He [8] compared the retrieval results of manifold ranking methods based on connected graph and k-NN graph, and found the latter one is better, so we only compared our algorithm with He's. The parameters of the manifold-ranking based algorithms for k-NN graph and 2N graph were set as $\sigma=1$, $\alpha=0.99$, $k=200$.

Figure 3 which uses the first image as query point shows the comparison results for image retrieval using L1 distance and L2 distance in Color Histogram feature space. The manifold-ranking algorithm based on 2N graph without specifying the free parameter k also improves the retrieval performance, and the improvement for Color Histogram is up to 3.98% comparing to k-NN graph based algorithm. One screenshot of the retrieval results in our prototype system are depicted in Figure 4.

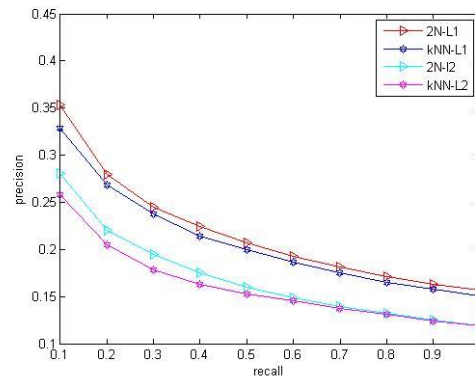


Figure 3. The Experiment Results in Image Retrieval

5. Conclusion and Future Work

In this paper, we presented a transductive learning algorithm based on manifold ranking with 2N and it was applied successfully in image retrieval. The variable number of neighbors of each vertex makes our graph structure more reliable and can reveal the inherent manifold structure more easily in high dimension feature spaces. Our experiments show that 2N is more suitable for the candidate graph structures of manifold-ranking based image retrieval.

Although 2N gained some success, we still need to improve method with respect to coping with outlier point. Based on the existing work, our future work may include: (1) Investigate how to efficiently reduce influence of an outlier point in the data set. An outlier is the main obstacle for the improvement of constructing 2N graph performance. (2) The method of constructing 2N graph is very simple and general. In our future work, we try to apply this method in different fields. (3) The experimental result on 2D toy data which is shown in Figure 2 has already exhibited the potentially automatic clustering ability of 2N graph structure. We will do further research on high dimension data clustering using 2N graph in our future work.



Figure 4. One Screenshot of the Retrieval Results

Acknowledgements

We would like to thank Ji Feng, Lijun Yang, Jinlong Huang and Langjun Duan for their valuable discussions and enlightening comments. This work was supported by National Natural Science Foundation of China (61272194) and Natural Science Foundation Project of Chongqing (CSTC2013JCYJA40049).

References

- [1] H. Kaur and K. Jyoti, "Survey of Techniques of High Level Semantic Based Image Retrieval", *IJRECT*, vol. 1, no. 2, (2013).
- [2] J. B. Tenenbaum, V. D. Silva and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction", *Science*, vol. 5500, no. 290, (2000).
- [3] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding", *Science*, vol. 5500, no. 290, (2000).
- [4] K. Duh and K. Kirchhoff, "Semi-supervised ranking for document retrieval", *Computer Speech & Language*, vol. 2, no. 25, (2011).
- [5] M. F. A. Hady and F. Schwenker, "Semi-supervised learning in Handbook on Neural Information Processing", Springer, (2013), pp. 215-239.
- [6] D. Zhou, O. Bousquet, T. N. Lal, J. Weston and B. Schölkopf, "Learning with local and global consistency", *Advances in neural information processing systems*, vol. 16, no. 16, (2004).
- [7] D. Zhou, J. Weston, A. Gretton, O. Bousquet and B. Schölkopf, "Ranking on data manifolds", *Advances in neural information processing systems*, vol. 16, (2004).
- [8] J. He, M. Li, H.-J. Zhang, H. Tong and C. Zhang, "Manifold-ranking based image retrieval", *Proceedings of the 12th annual ACM international conference on Multimedia*, (2004).
- [9] Z.-Y. Zhang and H.-Y. Zha, "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment", *Journal of Shanghai University (English Edition)*, vol. 4, no. 8, (2004).
- [10] J. He, M. Li, H.-J. Zhang, H. Tong, and C. Zhang, "Generalized manifold-ranking-based image retrieval, *Image Processing*", *IEEE Transactions on*, vol. 10, no. 15, (2006)
- [11] X. Cai and W. Li, "Mutually reinforced manifold-ranking based relevance propagation model for query-focused multi-document summarization, *Audio, Speech, and Language Processing*", *IEEE Transactions on*, vol. 5, no. 20, (2012).
- [12] T. Furuya and R. Ohbuchi, "Visual Saliency Weighting and Cross-Domain Manifold Ranking for Sketch-Based Image Retrieval", *MultiMedia Modeling*, (2014).
- [13] X. Ji, J. Han, X. Hu, K. Li, F. Deng, J. Fang, L. Guo, and T. Liu, "Retrieving video shots in semantic brain imaging space using manifold-ranking", *Image Processing (ICIP), 2011 18th IEEE International Conference on*, (2011).
- [14] J. Gui, D.-S. Huang, and Z. You, "An improvement on learning with local and global consistency", *Pattern Recognition*, 2008. *ICPR 2008. 19th International Conference on*, (2008).
- [15] C. Zhang and F. Wang, "A multilevel approach for learning from labeled and unlabeled data on graphs", *Pattern Recognition*, vol. 6, no. 43, (2010).
- [16] J. Tang, G.-J. Qi, M. Wang, and X.-S. Hua, "Video semantic analysis based on structure-sensitive anisotropic manifold ranking", *Signal Processing*, vol. 12, no. 89, (2009).
- [17] T. Xia, J. Cao, Y.-D. Zhang, and J.-T. Li, "On defining affinity graph for spectral clustering through ranking on manifolds", *Neurocomputing*, vol. 13, no. 72, (2009).
- [18] L. Yang, "K-edge connected neighborhood graph for geodesic distance estimation and nonlinear data projection", *Pattern Recognition*, 2004. *ICPR 2004. Proceedings of the 17th International Conference on*, (2004).
- [19] L. Yang, "Building k edge-disjoint spanning trees of minimum total length for isometric data embedding, *Pattern Analysis and Machine Intelligence*", *IEEE Transactions on*, vol. 10, no. 27, (2005).
- [20] Y. Bengio, P. Vincent, J.-F. Paiement, O. Delalleau, M. Ouimet, and N. LeRoux, "Learning eigenfunctions of similarity: linking spectral clustering and kernel PCA", *Dpartement d'informatique et recherche operationnelle, Universit de Montral, Tech. Rep*, (2003)
- [21] M. Kokare, B. Chatterji, and P. Biswas, "Comparison of similarity metrics for texture image retrieval. *TENCON 2003*", *Conference on Convergent Technologies for the Asia-Pacific Region*, (2003).
- [22] C. Wang, J. Zhao, X. He, C. Chen, and J. Bu, "Image retrieval using nonlinear manifold embedding", *Neurocomputing*, vol. 16, no. 72, (2009).
- [23] B.-K. Bao, B. Ni, Y. Mu, and S. Yan, "Efficient region-aware large graph construction towards scalable multi-label propagation", *Pattern Recognition*, vol. 3, no. 44, (2011).

Authors



Qingsheng Zhu, is a professor of Chongqing University. He received his PhD degree in Computer Science from Chongqing University in 2000. His research interests are service-oriented software engineering, data mining & outlier detection.



Zhi Chen, received his BS degree in Computer Sciences and Technology from Chongqing University in 2012. Now he is MS student of Computer Software and Theory, Chongqing University. His current research interests include different aspects of Data Mining.



Cheng Zhang, received his PhD degree in Computer Science from Chongqing University of China in 2010. He is an Associated Professor of College of Computer Science, Chongqing University, China. His research interests are the fields of Web Intelligence, Mobile Computation and New Web development technology.