The Application of Extended Geodesic Distance in Head Poses Estimation

Bingpeng Ma[1,3], Fei Yang[1,3], Wen Gao[1,2,3], and Baochang Zhang[2]

[1] Institute of Computing Technology,
Chinese Academy of Sciences, Beijing 100080, China
[2] Department of Computer Science and Engineering,
Harbin Institute of Technology, Harbin, China, 150001
[3] Graduate School of the Chinese Academy of Sciences, Beijing 100039, China

Abstract. This paper we proposes an extended geodesic distance for head pose estimation. In ISOMAP, two approaches are applied for neighborhood construction, called k-neighbor and ϵ-neighbor. For the k-neighbor, the number of the neighbors is a const k. For the other one, all the distances between the neighbors is less than ϵ. Either the k-neighbor or the ϵ-neighbor neglects the difference of each point. This paper proposes an new method called the kc-neighbor, in which the neighbors are defined based on c time distance of the k nearest neighbor, which can avoid the neighborhood graph unconnected and improve the accuracy in computing neighbors. In this paper, SVM rather than MDS is applied to classify head poses after the geodesic distances are computed. The experiments show the effectiveness of the proposed method.

1 Introduction

Dimension reduction techniques are widely used for the analysis of complex sets of data, such as face images. For face images, classical dimensionality reduction methods include Eigenface[1], Linear Discriminant Analysis(LDA)[2], Independent Component Analysis(ICA)[3], etc, all of which are linear methods. The linear methods have their limitations. On one hand, they cannot reveal the intrinsic distribution of a given data set. On the other hand, if there are changes in poses, facial expression and illumination, the projections may not be appropriate and the corresponding reconstruction error may be much higher. For a pair of points on the manifold, their Euclidean distance may not accurately reflect their intrinsic similarity and, consequently, is not suitable for determining intrinsic embedding or pattern classification. For example, Fig.1 is the data points sampled of Swissroll[4]. The Euclidean distance between point **x** and point **y** is deceptively small in the three-dimensional input space though their geodesic distance on a intrinsic two-dimensional manifold is large.

The recently proposed ISOMAP[5], LLE[6] and Laplacian Eigenmaps[7] algorithms are popular non-linear dimensionality reduction methods. The ISOMAP method computes pair-wise distances in the geodesic space of the manifold, and then performs classical Multidimensional Scaling(MDS)[8] to map data points
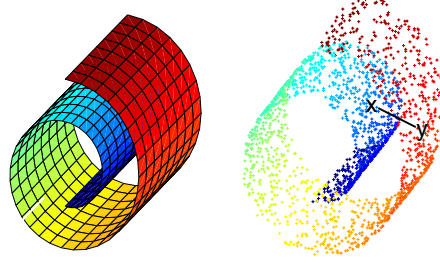
Fig. 1. The data points of Swissroll

from their high-dimensional input space to low-dimension coordinates of a non-linear manifold. In ISOMAP, the geodesic distances can reflect the intrinsic low-dimensional geometry of the manifold, but it can't reduce dimension when the number of samples is very large. And, MDS is applied for visualization in low-dimension, which can't deal with non-linear data.

In this paper, kc-neighbor is applied to compute the geodesic distances for the head-pose estimation, which is necessary in a variety of applications such as face recognition. The problem is difficult because it is an estimation for an inherently three dimensional quantity from two dimensional image data.

In this paper, each face image with a certain pose is considered as a point in the high-dimension manifold. First the neighborhood is constructed using the kc-neighbor method. Then the geodesic distance are computed for all the pairwise points. Finally, SVM is applied to classify each point into pose classes using the geodesic distances from other points. Compared with k-neighbor and ϵ-neighbor of ISOMAP, kc-neighbor can correctly reflect the relation between each point and its neighbors, and SVM classifiers can improve the accuracy of the pose estimation. Experimental results on data sets show that kc-ISOMAP improves estimation accuracy.

The remaining part of this paper is organized as follows. In section 2, we describe kc-neighbor. In section 3, we introduce the SVM classifiers. Then, two databases are used to evaluate the performance of kc-neighbor in Section 4. Finally we conclude this work in section 5.

2   The Extended Geodesic Distance

ISOMAP's global coordinates provide a simple way to analyze and manipulate the high-dimensional observation in terms of their intrinsic nonlinear degrees of freedom. In ISOMAP, nonlinear features are extracted based on estimating geodesic distances and embedded by MDS. The basic idea is that for the neighbor points on a manifold, the Euclidean distances provide a fair approximation of geodesic distances, whereas for faraway points the geodesic distances are estimated by the shortest pathes through neighboring points.

The construction of neighborhood is a critical step in the ISOMAP. Neighbors should be local in the sense that the Euclidean distances are fair approximation of the geodesic distances. Tenenbaun et. al.[5] proposed two methods for neighborhood construction, called k-ISOMAP and ϵ-ISOMAP. The k-ISOMAP means defines the graph G over all data points by connecting points $\mathbf{x}_i$ and $\mathbf{x}_j$ if $\mathbf{x}_i$ is one of the k nearest neighbors of $\mathbf{x}_j$. In ϵ-ISOMAP method, the graph G is defined by connecting each point to all the points within the fixed radius ϵ. The neighborhood relation is symmetric by definition and the numbers of neighbors is different for each point. The choice of an appropriate ϵ is a difficult task. If ϵ is too small the resulting graph becomes sparse and the unconnected subgraphs often exist, while if ϵ is too large the idea of connecting local patches gets lost. In both cases the approximation error increases.

Due to the inhomogeneous density of the samples it seems more data-sensitive to define the k nearest points of x_i as its neighbors. The k-neighbor method will not generate any isolated point.But if more than k points cluster, they will form an unconnected subgraph. Furthermore, the rule is not symmetric in the sense that x_j is a neighbor point of x_i does not necessarily imply that x_i is also a neighbor point of x_j, so that G has to be symmetrized afterwards.
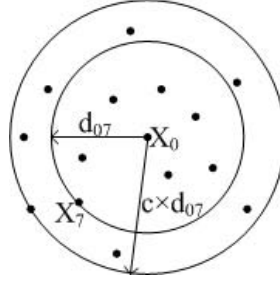

Fig. 2. kc-neighbor. $\mathbf{x}_7$ is the 7th nearest neighbor and the radius is d_{07}. In kc-neighbor method, all the points whose radius less than c times the d_{07} are $\mathbf{x}_0$'s neighbors.

To consider changes in the sample density, kc-neighbor method is presented in this paper. In this method the neighbors of a point $\mathbf{x}_i$ include all the points that lie inside the ϵ-ball with the radius ϵ equal to c times distance between $\mathbf{x}_i$ and its k-th nearest neighbor. If a point $\mathbf{x}_i$ have k neighbors in k-ISOMAP, and d_{07} is the distance of the k-th neighbor, we define all the points which are closer than c times d_{07} as the neighbors of $\mathbf{x}_i$. Three reasons lead us to present this idea. First, the sample density varies, so a fixed rule will not apply effectively to all points. When using k-neighbor, we think all the points have the same number of neighbors. As to ϵ-neighbor, all the neighbor points are within the same distance. Second, compared with k-neighbor, kc-neighbor can avoid unconnected subgraph because we specify the different numbers of neighbors for different points. Compared with ϵ-neighbor, kc-neighbor uses a dynamic ϵ for different points and makes all the points have roughly the same number

of neighbors. At last, kc-neighbor does not increase the computing complexity because we use the sort result when finding neighbors.

Based on the kc-neighbor, we present the kc-ISOMAP method. Compared with the k-ISOMAP, the main difference of the kc-ISOMAP is using kc-neighbor to replace the k-neighbor. Given a training set $\{\mathbf{x}_i, i = 1, \ldots, m\}$, the first step of kc-ISOMAP determines the nearest neighbors for point $\mathbf{x}_i$ based on the Euclidean distances $d_X(\mathbf{x}_i, \mathbf{x}_j)$ in the input space $\mathbb{X}$. These neighborhood relations are represented as a weighted graph G in which

$$d_G(\mathbf{x}_i, \mathbf{x}_j) = \begin{cases} d_X(\mathbf{x}_i, \mathbf{x}_j), & \text{if } \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ are neighbors} \\ \infty, & \text{otherwise} \end{cases} \tag{1}$$

In the second step, kc-ISOMAP estimates the geodesic distances $d_M(\mathbf{x}_i, \mathbf{x}_j)$ between all pairs of points on the manifold $\mathbb{M}$ by computing their shortest path distance $d_G(\mathbf{x}_i, \mathbf{x}_j)$ in the graph G. In generally, Floyd-Warshall algorithm is used to compute the geodesic distances $d_M(\mathbf{x}_i, \mathbf{x}_j)$:

$$d_M(\mathbf{x}_i, \mathbf{x}_j) = min\{d_G(\mathbf{x}_i, \mathbf{x}_j), d_G(\mathbf{x}_i, \mathbf{x}_p)) + d_G(\mathbf{x}_p, \mathbf{x}_j)\} \tag{2}$$

3   SVM Classification

In ISOMAP, after computing geodesic distances, MDS is applied at the aim of visualization in low-dimension. From a non-technical point of view, the purpose of MDS is to provide a visual representation of the pattern of proximities (i.e., similarities or distances) among a set of objects, which does not contribute to the improvement of the classification accuracy. In this paper, SVM classifers are used to replace MDS after computing the geodesic distances.

SVM is a quadratic optimization problem in order to maximize the margin between examples of two classes either in the original input space or in an implicitly mapped higher dimensional space by using kernel functions. Though new kernels are being proposed by researchers, we still use the basic RBF(radial basis function) kernel.

Generally, SVM are used for 2-class problems. In this paper, we use "one against one" approach to slove the k-class problem. In this approach each classifier is constructed to seperate a two classes, and totally $k(k-1)/2$ classifiers are constructed and a voting strategy is used to get the "winner" class.

4   Experiment

We test kc-ISOMAP method using the public FERET[11] database and the CAS-PEAL[12] database. Fig.3 and Fig.4 show some subjects in the FERET and CAS-PEAL database. The FERET database contains 1400 images of 200 persons, varying in facial expression(open/closed eyes, smiling/non smiling), and each person has seven horizontally poses $\{-40°, -25°, -15°, 0°, +15°, +25°, +40°\}$. The persons of FERET database come from Asia, Europe, and Africa.

The CAS-PEAL database contains seven poses $\{-45°, -30°, -15°, 0°, +15°, +30°, +45°\}$ of 1400 persons. In order to compare the results with FERET database, we use a subset of the CAS-PEAL database including 1400 images of 200 persons with the subject ID ranging from 401 to 600. Unlike FERET, the persons of CAS-PEAL are all from Asian.

Fig. 3. Face images in the FERET database   **Fig. 4.** Face images in the CAS-PEAL database

We first label the positions of eyes by hand, and then crop the images to 32×32 pixels. We use histogram equalization to reduce the influence of lighting, represent each image by a raster scan vector of the intensity values, and finally normalize each vector to be zero-mean unit-variance vectors.

In experiment, we use cross-validation in order to avoid over-training. We first sort the images in the database by file name, and divide them into 3 parts. One part is taken as testing set and the other two as training set. Repeat three times so that each part has been taken as testing set. All the testing results are the mean results of all testing sets.
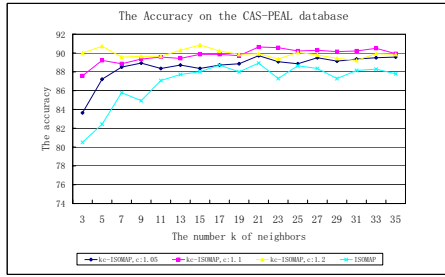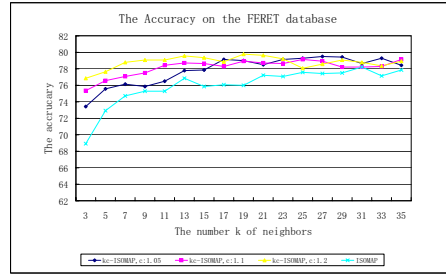
In computing geodesic distance, it is difficult to compute geodesic distance of new samples. In our experiment, we use the general ways which are used in ISOMAP and LLE. We first compute the geodesic distances of all samples, which donot consider the difference between the training samples and the testing samples. And then the geodesic distances can be divided two parts: the training set and the testing set. In actual application, the geodesic distances of the testing samples can be computed by executing the Floyd-Warshall algorithm.

We compare the following three methods: the P-k-ISOMAP, the k-ISOMAP and kc-ISOMAP. In the P-k-ISOMAP, we first use PCA to reduce the dimension of the images from 1024 to 245, which reserves 99.9 percent of the total energy of eigenvalues, and then use the new samples to compute the geodesic distances. In the the k-ISOMAP, we directly use the images to compute the geodesic distances. In all the three methods, SVM classifiers are performed for pose estimation and different ks are used to find the influence caused by k. And three different values of $c(1.05, 1.1, 1.2)$ are selected to discover the influence of c in kc-ISOMAP. The experiment results are shown in Tab.1, Fig.5 and Fig.6.

Form the table and the figures, we can know kc-ISOMAP have an improvement in pose estimation on both the FERET database and the CAS-PEAL database. The accuracy rate of both the ISOMAP and kc-ISOMAP improve remarkably with the increase of the number of neighbors when the number of neighbors is ranging from 3 to 9, but it tends to stabilize with the increase of the number of the neighbors. It means that the selection of k is very important

Table 1. Error Rate Comparison of Different Pose Estimation

method	CAS-PEAL	FERET
P-k-ISOMAP $k = 7$	13.79	24.21
P-k-ISOMAP $k = 14$	12.64	23.52
P-k-ISOMAP $k = 21$	11.21	22.78
k-ISOMAP $k = 7$	14.29	25.35
k-ISOMAP $k = 14$	11.86	23.35
k-ISOMAP $k = 21$	11.07	22.78
kc-ISOMAP(c=1.1) k=7	11.14	22.93
kc-ISOMAP(c=1.1) k=14	10.86	21.36
kc-ISOMAP(c=1.1) k=21	9.29	21.21

Fig. 5. The results of pose estimation on the CAS-PEAL database

Fig. 6. The results of pose estimation on the FERET database

for the preservability of the pose manifold. If the number of neighbors is much smaller, the structure of the manifold can not be maintained, at this instance, the improvement of kc-ISOMAP is more apparent. With the increase of k, the predominance of kc-ISOMAP is decreased, but kc-ISOMAP always can obtain better accuracy than k-ISOMAP, which means kc-ISOMAP can maintain the structure of manifold better because kc-ISOMAP more care about the neighborhood relation of the each samples. And from Fig.5 and Fig.6, we can know the results of the difference c are nearly equal, which means that the value of c is not important, but it can cause the dynamic number of the neighbors and then improve the accuracy.

5 Conclusion and Future Work

This paper proposes a novel method to extend the geodesic distance in ISOMAP. Compared with the traditional geodesic distance, this method considers the dynamic number of neighbors for each point, which makes the relation of neighbors more correctly. And after computing the geodesic distances, it apply SVM classifiers replace MDS because the MDS is the method to preserve the feature of

samples, which can not improve the accuracy rate of classify and SVM is the best classifier when the number of the training samples is enough to find the correct support vectors. The experiment shows kc-ISOMAP can improve the accuracy of the poses estimation.

Acknowledgements

References

1. M.Turk, and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, (3) 71-86,1991.
2. P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specfic linear projection", IEEE Trans. on PAMI, Vol. 19, No. 7, 711-720, 1997.
3. Marian Stewart Bartlett, Terrence J. Sejnowski, "Independent components of face images: A representation for face recognition", Proceedings of the 4th Annual Jount Symposium on Neural Computation, Pasadena, CA, May 17, 1997.
4. Ming-HsuanYang, "Extended Isomap for Classification", ICPR (3) 2002: 615-618.
5. J.B.Tenenbaum, V.de Silva, and J.C.Langford, "A global geometric framework for nonlinear dimensionality reduction", Science 290: 2319-2323, 2000.
6. Roweis.S. and Saul.L., "Nonlinear dimensionality reduction by locally linear embedding". Science 290:2323-2326, 2000
7. M. Belkin and P. Niyogi, "Laplacian eigenmaps and spetral techniques for embedding and clustering", Advances in Neural Information Processing Systems, vol. 15, 2001.
8. Trevor F. Cox and Michael A. A. Cox, "Multidimensional Scaling", CRC Press, 2000.
9. Cortes,C. and Vapnik, V, "Support vector network", Machine Learning, 20:273:297, 1995.
10. Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm
11. Phillips P.J., Moon H., etc. The FERET evaluation methodology for face - recognition algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(10):1090-1104.
12. Wen Gao, Bo Cao, Shiguang Shan, Xiaohua Zhang, Delong Zhou, The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations, technical report of JDL, 2004, http://www.jdl.ac.cn/ peal/peal_tr.pdf.