

Isomap Based on the Image Euclidean Distance

Jie Chen¹ Ruiping Wang² Shiguang Shan² Xilin Chen² Wen Gao^{1,2}

¹School of Computer Science and Technology, Harbin Institute of Technology, Harbin, 150001, China

²ICT-ISVISION Joint R&D Laboratory for Face Recognition, Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100080, China
{chenjie, rpwang, sgshan, xlchen, wgao, gqcui}@jdl.ac.cn

Abstract

Scientists find that the human perception is based on the similarity on the manifold of data set. Isometric feature mapping (Isomap) is one of the representative techniques of manifold. It is intuitive, well understood and produces reasonable mapping results. However, if the input data for manifold learning are corrupted with noises, the Isomap algorithm is topologically unstable. In this paper, we present an improved manifold learning method when the input data are images—the Image Euclidean distance based Isomap (ImIsomap), in which we use a new distance for images called Image Euclidean Distance (IMED). Experimental results demonstrate a consistent performance improvement of the algorithm ImIsomap over the traditional Isomap based on Euclidean distance.

1. Introduction

Manifold facilitates scientists to deal with large volumes of high-dimensional data, such as global climate patterns, stellar spectra, or human gene distributions. It is especially useful when one confronts the problem of dimensionality reduction. Recently, some representative techniques have been proposed, such as Isomap [21], local linear embedding (LLE) [16], and Laplacian Eigenmap [2]. Meanwhile, Isomap is intuitive, well understood and produces reasonable mapping results [9, 12, 18, 19, 25]. Also, it is supported theoretically, such as its convergence proof [3] and it can recover the co-ordinates [6]. Besides, other improvements for Isomap are also presented, such as [4, 5, 6, 8, 10, 11, 13, 14, 15, 17, 20, 22, 23, 26].

However, Isomap suffers from the topological stability when the input data are noised. As discussed in [1], Balasubramanian and Schwartz present that the “Swiss roll” data, to which they have added a small amount of noise, Isomap becomes topologically unstable. Though it can be solved by finding suitable neighborhood size to yield topology-preserving embeddings, Tenenbaum indicates that it is an important area to improve the robustness of these dimensionality reduction algorithms confronted with high noise levels [21].

When the input data are images, we argue that for the manifold learning of images, an alternative distance measure is preferred to improve the robustness of Isomap in the presence of high noise level. It is because that a central problem in Isomap is to determine the distance between images. Among all the image metrics, Euclidean distance (ED) is commonly used due to its simplicity. However, this distance measure suffers from a high sensitivity even to small deformation. This phenomenon is caused by the fact that the Euclidean distance does not take into account that the objects \mathbf{x} , \mathbf{y} are images and the spatial relationships among pixels. The traditional Euclidean distance is only a summation of the pixel-wise

intensity differences and, consequently, small deformation may result in a large Euclidean distance. Therefore, in the improved Isomap algorithm, we adopt the IMED instead of ED for the computation of distances between pairs of images for Isomap.

The rest of this paper is organized as follows: The proposed method ImIsomap is described in section 2. Experimental results are presented in section 3, followed by conclusion in section 4.

2. IMED based Isomap

In this section, after a review of the image Euclidean distance, we discuss the improved Isomap.

2.1 The image Euclidean distance

Different from the traditional Euclidean distance, the IMED considers the spatial relationships of pixels. It characterizes by robust to small perturbation [24].

Given an $M \times N$ image, it is actually a point in an MN -dimensional image space with the base as e_1, e_2, \dots, e_{MN} . Let $g_{ij} = \langle e_i, e_j \rangle = \sqrt{\langle e_i, e_i \rangle} \sqrt{\langle e_j, e_j \rangle} \cos \theta_{ij}$, $i, j = 1, 2, \dots, MN$, where $\langle \cdot, \cdot \rangle$ is the scalar product and θ_{ij} is the angle between e_i and e_j . The Euclidean distance of two images \mathbf{x} , \mathbf{y} is written by $d_E^2(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{MN} g_{ij} (x_i - y_i)(x_j - y_j) = (\mathbf{x} - \mathbf{y})^T G (\mathbf{x} - \mathbf{y})$, where $G = (g_{ij})_{MN \times MN}$ is a MN^{th} order symmetric and positive definite matrix.

$$\text{A concrete form for } g_{ij} \text{ is } g_{ij} = \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{|P_i - P_j|^2}{2\sigma^2} \right\},$$

where σ is the width parameter; P_i, P_j , $i, j = 1, 2, \dots, MN$ is pixels and $|P_i - P_j|$ is the distance between P_i and P_j on the image lattice.

However, it is time consuming to compute the IMED $d_{IMED} = (\mathbf{x} - \mathbf{y})^T G (\mathbf{x} - \mathbf{y})$ for all pairs of images, especially for the large database. It can be greatly simplified by introducing a linear transformation. For a matrix G , it is decomposed as $G = G^{\frac{1}{2}} \Lambda G^{\frac{1}{2}}$, $G^{\frac{1}{2}} = \Gamma \Lambda^{\frac{1}{2}} \Gamma^T$. Here, Λ is a diagonal matrix consisting of the eigenvalues of G ; Γ is an orthogonal matrix whose column vectors are eigenvectors of G . All images \mathbf{x} , \mathbf{y} are then transformed. Let $\mathbf{u} = G^{\frac{1}{2}} \mathbf{x}$, $\mathbf{v} = G^{\frac{1}{2}} \mathbf{y}$, so $d_{IMED} = (\mathbf{x} - \mathbf{y})^T G^{\frac{1}{2}} \Lambda G^{\frac{1}{2}} (\mathbf{x} - \mathbf{y}) = (\mathbf{u} - \mathbf{v})^T (\mathbf{u} - \mathbf{v})$. Note that most elements of $G^{\frac{1}{2}}$ are nearly zero, so the linear

transformation in spatial domain is simplified by a 5×5 mask in a practical use.

2.2 The improved Isomap

Manifold enables us to visualize data, perform classification and cluster more efficiently. Recently, Isomap—one of the emergent representative techniques—is intuitive, well understood and can produce reasonable mapping results.

As discussed in [21], in order to learn the manifold of handwriting, Tenenbaum *et al.* indicate the input-space distances $d_x(i, j)$ are measured by tangent distance, a metric designed to capture the invariance relevant in handwriting recognition. We argue that for the manifold learning of images, an alternative distance measure is preferred. In order to improve the robustness of these dimensionality reduction algorithms in the presence of high noise level, we adopt the IMED instead of ED for the computation of distances between pairs of images.

The improved Isomap algorithm based on IMED (ImIsomap) is as follows:

The improved Isomap algorithm

Input: Given n data points in the high-dimensional input space $X = \{x_1, x_2, \dots, x_n\}$, the parameter ε or K (inherit from Isomap to compute the neighborhood).

Output: Coordinate vectors y_i in a d -dimensional Euclidean space Y that best represent the intrinsic geometry of the data.

Step 1: Performance the linear transformation for each input image: $u_i = G^{1/2} x_i$, $i = 1, 2, \dots, n$, where the width parameter σ is equal to 1 for the computation of g_{ij} ;

Step 2: Calculate the distances between pairs of images:

$$d_x = (x_i - x_j)^T G (x_i - x_j) = (u_i - u_j)^T (u_i - u_j)$$

$$i, j = 1, 2, \dots, n$$

Step 3: Construct the Neighborhood graph based on the IMED distances between pairs of images;

Step 4: Finds the shortest path through neighborhood graph, for each pair of non-neighboring data points, subject to the constraint that the path must hop from neighbor to neighbor;

Step 5: Construct low-dimensional embedding by the classical multidimensional scaling and output the coordinate vectors y_i .

3. Experiments

In this section, some experiments are carried out to prove that the topological stability of ImIsomap is much improved compared to Isomap. The first experiment is based on the face image set given by Tenenbaum; the second is based on a collected real image set.

3.1 Experiments on the virtual set

The first experiment is based on the face image set given by Tenenbaum in [21]. It consists of a sequence of 4096-dimensional vectors, representing the brightness values of 64 by 64 pixel images of a face rendered with different poses and lighting directions.

Applied to $N = 698$ raw images, Isomap can learn the

three-dimensional embedding of the data's intrinsic geometric structure. In general, the intrinsic dimensionality of the data can be estimated by looking for the "elbow" at which this curve ceases to decrease significantly with added dimensions. Note that in our case, we use the K nearest neighborhood and this parameter is determined experientially ($K=6$ for our case). And we denote *Elsomap* the Isomap based on ED, and *ImIsomap* the Isomap based on IMED.

In order to testify the robustness of the proposed method, we add several kinds of noises: 1) Gaussian noise (or zero-mean normally distributed noise), 2) blur, 3) affine deformation—rotation and shift.

First of all, we add some Gaussian noises into these rendered face images. As discussed in the first row of table 1, we change the variance of the Gaussian noises. When the variance is less than 0.03, both EIsomap and ImIsomap can work well. They can both estimate the right dimension of the database while the residual variance of the EIsomap is a little larger than that of the ImIsomap. However, when the variance of the Gaussian noises is larger than 0.03, EIsomap is out of work and ImIsomap perform correctly. When the noise variance belong to the interval $[0.04, 0.1]$, one may notice that the residual variance corresponding to the knee point of ImIsomap fluctuates in smaller extent compared to EIsomap.

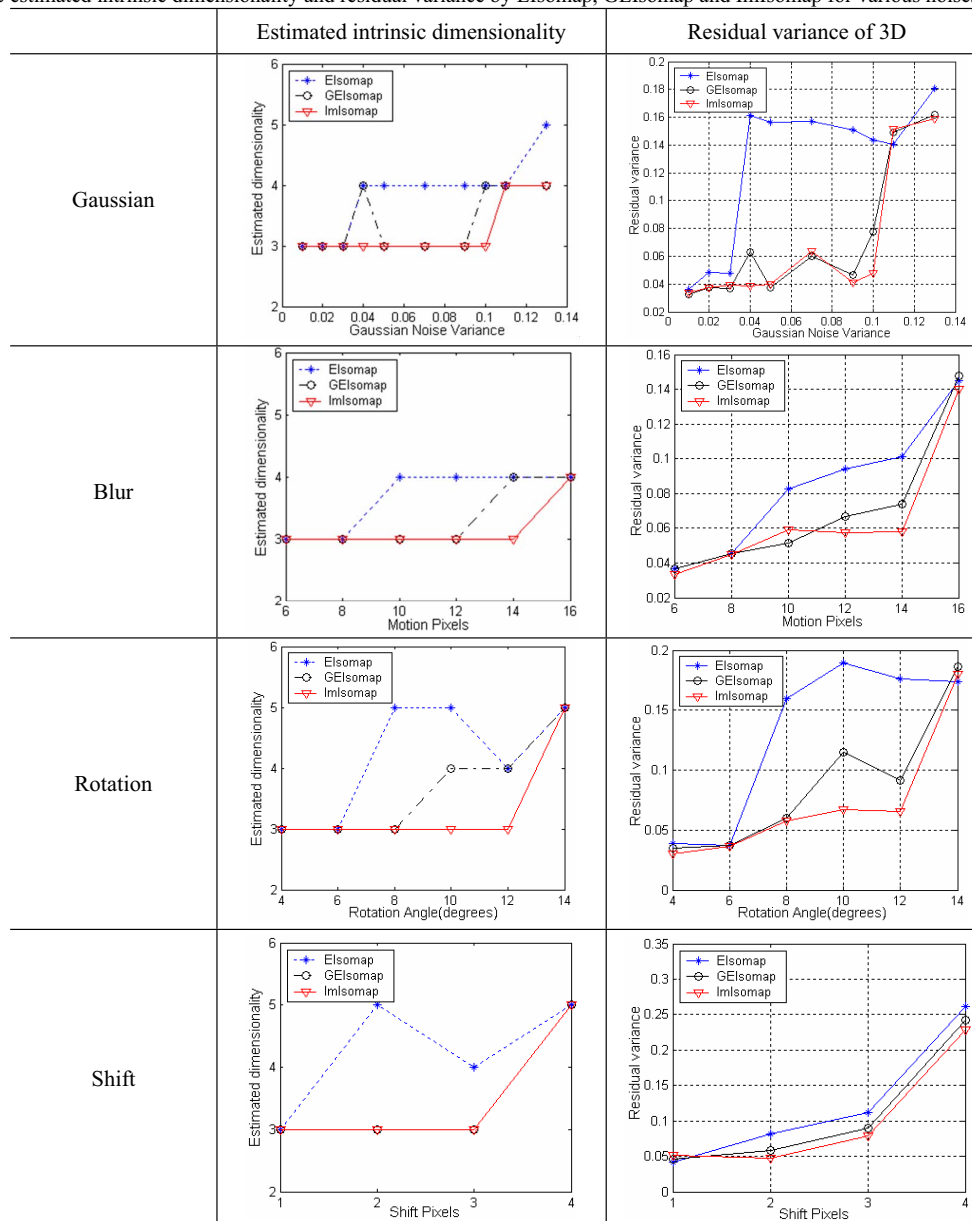
Considering the embedded Gaussian noises, one might filter those noisy images by the Gaussian filters. Subsequently, these filtered images can be learned by EIsomap (called GEIsomap for simplicity) to detect its intrinsic dimension. Herein, to compare with the results of ImIsomap, the window size of Gaussian filter is 5×5 .

The comparison of the detected dimension of the noisy images and the residual variance of the knee point among the EIsomap, GEIsomap and ImIsomap are also illustrated in table 1. One can conclude that the GEIsomap demonstrates its robustness to some extent when the type of the noise is known in advance. However, GEIsomap make a mistake — when the noise variance is 0.04, its estimated dimension is 4 instead of the true intrinsic dimension, 3. Furthermore, when the noise variance arrives at 0.09, GEIsomap can not work as well as EIsomap. It demonstrates that the ImIsomap has stronger robustness to the embedded Gaussian noises.

Besides the Gaussian noises, we also add other kinds of noise, such as blur and affine deformation—rotation and shift. The detected dimension and residual variance are also illustrated in table 1.

For the algorithm ImIsomap works better than EIsomap, we conclude that the distance metric IMED is much better than ED because that the former considers the spatial relation between the neighbor pixels compared to the latter. As to GEIsomap can improve the performance compared to EIsomap, we conclude that the Gaussian filter can smooth an image, which in fact modifies the pixel intensity according its neighborhoods by a given template. That is to say, it can perform partially the function of IMED. However, ImIsomap can improve the performance further than GEIsomap. It is because the IMED can decrease the influence of noise better than the Gaussian filter.

Table 1. The estimated intrinsic dimensionality and residual variance by Elsomap, GElsomap and ImIsomap for various noises



3.2 Experiments on the real set

In this section, we use a real set to testify the robustness of ImIsomap. This set of face images has only one freedom degree— rotation in plane for good visualization. This set consists of 180 face images under different rotation degree of a person. Two neighboring images have the space degree equal one. The rotation degree locate within the interval $[-90^\circ, 90^\circ]$. The resolution of each sample is of 64 by 64 pixels. Some examples are shown in Fig.1 (a). The detected intrinsic dimension and the embedding are illustrated in Fig 1 (b) and (c).

Likewise, some Gaussian noises with different variances are added into these images. Some 1D projections of the learned embedding of this set of images under different noise densities are shown in Fig. 2 (a) and (b). The estimated intrinsic dimensionality corresponding to different noise variances is illustrated in Fig. 2 (c). On account of the noises added into the face images, the 1D

projections of the face embeddings become curves, comparing to the line shown in Fig 1(c). Compared with the noiseless case, for various noise densities, projections by Elsomap, GElsomap and ImIsomap are all offset. In Fig. 2(d), we statistically compute the offset variances.

From Fig 2, one can conclude that the ImIsomap is more robust than GElsomap and Elsomap when confronted with the real and noisy images.

4. Conclusion

We have addressed the issue of manifold learning when the input data are noisy images. We present an improved manifold learning method — ImIsomap. The proposed method is tested on different corrupted images: Gaussian noise, blur and affine deformation with various densities. The experimental results demonstrate that ImIsomap can significantly improve the topological stability comparison to the traditional Isomap.

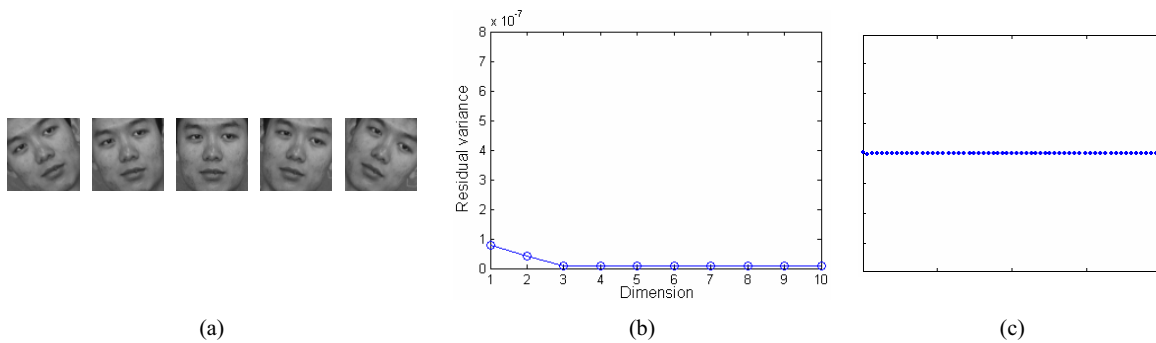


Fig. 1. The real set and their residual variance and embedding; (a) Some examples of the real set; (b) The residual variance; (c) the 1D projection of the real face images embedding.

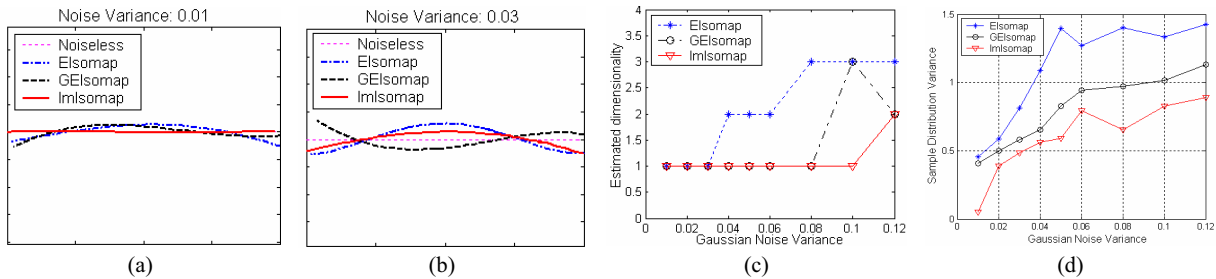


Fig.3. 1D projections and statistical results; (a) and (b) The projections of the real face images embedding after corrupted by the Gaussian noises with different variances; (c) The estimated intrinsic dimensionality corresponding to different noise variances; (d) Statistical variances of the projections offset corresponding to different noise variances compared with the noiseless case.

Acknowledgements

This research is partially sponsored by Natural Science Foundation of China under contract No.60332010, “100 Talents Program” of CAS.

Reference

- [1] M. Balasubramanian, E. L. Schwartz, J. B. Tenenbaum, V. de Silva and J. C. Langford. The Isomap Algorithm and Topological Stability. *Science*, 2002.
- [2] M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In *NIPS*, 2002.
- [3] M. Bernstein, V. de Silva, J. Langford, and J. Tenenbaum. Graph approximations to geodesics on embedded manifolds. *Technical report, Department of Psychology, Stanford University*, 2000.
- [4] M. Brand. Charting a manifold. In *NIPS*, 2003.
- [5] J. Costa and A. O. Hero, Geodesic entropic graphs for dimension and entropy estimation in manifold learning. *IEEE Trans. on Signal Processing*, pp. 2210-2221, 2004.
- [6] D. L. Donoho and C. Grimes. When does ISOMAP recover natural parameterization of families of articulated images? *Technical Report 2002-27, Stanford University*, 2002.
- [7] A. Efros, V. Isler, J. Shi, M. Visontai. Seeing through water. *NIPS* 2004.
- [8] D. R. Hundley and M. J. Kirby. Estimation of topological dimension. In *SIAM*, 2003.
- [9] O. Jenkins and M. Mataric. Automated derivation of behavior vocabularies for autonomous humanoid motion. In *Autonomous Agents and Multivalent Systems*, 2003.
- [10] O. Jenkins and M. Mataric. A Spatio-temporal Extension to Isomap Nonlinear Dimension Reduction. *ICML*, 2004.
- [11] N. A. Laskaris, A.A. Ioannides. Semantic geodesic maps: a unifying geometrical approach for studying the structure and dynamics of single trial evoked responses. *Clinical Neurophysiology*, 2002, 1209–1226.
- [12] M. H. Law, N. Zhang, A. K. Jain. Nonlinear Manifold Learning for Data Stream. In *SIAM*, pp. 33-44, 2004.
- [13] J. A. Lee, A. Lendasse and M. Verleysen. Curvilinear Distance Analysis versus Isomap, *European Symposium on Artificial Neural Networks*, pp.185–192, 2002.
- [14] F. M'emoli and G. Sapiro. Distance Functions and Geodesics on Points Clouds. *Institute for Mathematics and its Applications*, 2002.
- [15] K. Pettis, T. Bailey, A. K. Jain, and R. Dubes. An intrinsic dimensionality estimator from near-neighbor information. *PAMI*, pp.25-36, 1979.
- [16] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000.
- [17] S. T. Roweis, L. K. Saul, and G. E. Hinton. Global coordination of local linear models. In *NIPS*, 2002.
- [18] V. de Silva and J. B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. *NIPS*, pp. 705-712, 2003.
- [19] V. de Silva, J. B. Tenenbaum. Unsupervised learning of curved manifolds. *Nonlinear Estimation and Classification*, 2002.
- [20] Y. W. Teh and S. T. Roweis. Automatic alignment of local representations. In *NIPS*, pp. 841-848. 2003.
- [21] B. J. Tenenbaum, V. Silva, and J. Langford. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science*, pp.2319-2323, 2000
- [22] J. J. Verbeek, N. Vlassis, and B. Krose. Coordinating principal component analyzers. In *International Conf. on Artificial Neural Networks*, pp. 914-919, 2002.
- [23] J. J. Verbeek, N. Vlassis, and B. Krose. Fast nonlinear dimensionality reduction with topology preserving networks. In *10th European Symposium on Artificial Neural Networks*, pp.193-198, 2002.
- [24] L. Wang, Y. Zhang, and J. Feng. On the Euclidean Distance of Images. In *PAMI*, pp.1334-1339, 2005
- [25] M.-H. Yang. Face recognition using extended ISOMAP. In *ICIP*, pp.117-120, 2002.
- [26] H. Zha and Z. Zhang. Isometric embedding and continuum ISOMAP. In *ICML*, 2003.