2015

Adaptive Graph Construction for Isomap Manifold Learning

Loc Tran
Old Dominion University

Zezhong Zheng
University of Electronic Science and Technology of China

Guoquing Zhou
Guilin University of Technology

Jiang Li
Old Dominion University, jli@odu.edu

Karen O. Egiazarian (Ed.)

See next page for additional authors

Original Publication Citation

Tran, L. Zheng, Z., Zhou, G., & Li, J. (2015) Adaptive graph construction for Isomap manifold learning. In K.O. Eglazarian, S.S. Agaian, & A.P. Gotchev (Eds.), *Image Processing: Algorithms and Systems XIII, Proceedings of SPIE-IS&T Electronic Imaging, SPIE Vol. 9399* (939904). SPIE. https://doi.org/10.1117/12.2082646

Adaptive Graph Construction for Isomap Manifold Learning

[1]Loc Tran, [2]Zezhong Zheng, [3]Guoqing Zhou and [1,3]Jiang Li

[1]Department of Electrical and Computer Engineering
Old Dominion University, Norfolk, VA 23529
[2]School of Resources and Environment
University of Electronic Science and Technology of China, Chengdu, China
[3]Guangxi Key Laboratory for Spatial Information and Geomatics
Guilin University of Technology, Guilin, China

ABSTRACT

Isomap is a classical manifold learning approach that preserves geodesic distance of nonlinear data sets. One of the main drawbacks of this method is that it is susceptible to leaking, where a shortcut appears between normally separated portions of a manifold. We propose an adaptive graph construction approach that is based upon the sparsity property of the ℓ_1 norm. The ℓ_1 enhanced graph construction method replaces k-nearest neighbors in the classical approach. The proposed algorithm is first tested on the data sets from the UCI data base repository which showed that the proposed approach performs better than the classical approach. Next, the proposed approach is applied to two image data sets and achieved improved performances over standard Isomap.

1. INTRODUCTION

In manifold learning, the goal is to learn a linear representation of nonlinear structures present in quantitative data. The methods used for this often incorporate a simple neighborhood selection algorithm such as k-nearest neighbors. Rigid approaches such as this may inadequately reflect the intrinsic structure of nonlinear manifolds since the optimal number of neighbors selected may vary at different points on the manifold. We introduce a novel adaptive neighborhood selection approach for manifold graph construction and apply it on classical Isomap.

The Isomap algorithm consists of three steps. First, a neighborhood is created to form a neighborhood graph. Next, a geodesic distance matrix is calculated from the neighborhood graph. This matrix will consist of pairwise distances between every sample of the data set where the distances are calculated by paths from the neighborhood graph. Finally, dimensionality reduction is performed on the distance matrix to form the lower dimensionality space while preserving the geodesic distances found in the second step. A successful manifold using Isomap depends largely on the formation of the neighborhood graph.[1]

One of the major drawbacks of the Isomap algorithm is that it is susceptible to leaking or short circuiting.[2] This can happen if neighbors are selected such that a short cut is created between two areas that do not follow the intrinsic structure of the manifold. As a result, large geodesic distances are misrepresented as short distances because of a poorly formed graph. As an example, refer to Figure 1. The intrinsic distance between the two points along the intrinsic structure is large. However if there were neighbors directly between the two points, the distance calculated will adhere more with the dotted line producing an erroneous geodesic distance calculation. The leaking problem is especially prevalent in noisy data sets. The graph construction is thus a very critical step for the Isomap method to create a successful manifold.

Because of the short circuiting problem, it is undesirable to have a large neighborhood. In the original graph construction method, the nearest neighbors are chosen from k-nearest neighbors using Euclidean distance. The selection of k is critical in creating a good graph. If k is too large the risk of short circuiting is higher. On the other hand, if k is too small, the graph may not be fully connected. Another problem that occurs is that a good k value for one location may be inappropriate for another location. In this paper, these problems are addressed with an adaptive neighbor selection method implemented with the ℓ_1-norm.

Further author information: (Send correspondence to Jiang Li: jli@odu.edu.

Image Processing: Algorithms and Systems XIII, edited by Karen O. Egiazarian, Sos S. Agaian,
Atanas P. Gotchev, Proc. of SPIE-IS&T Vol. 9399, 939904 · © 2015 SPIE-IS&T ·
CCC code: 0277-786X/15/$18 · doi: 10.1117/12.2082646

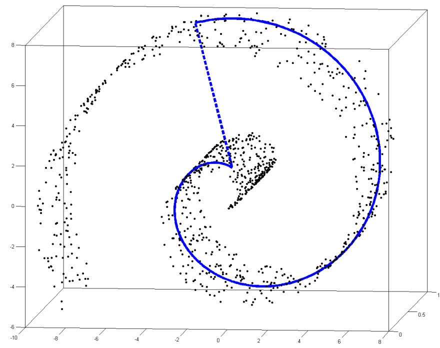

Figure 1. Swiss roll data set.

In the literature, it has been shown that sparse representation has applications in a wide range of computer vision and pattern recognition including face and object recognition,[3] compressive sensing,[4,5] subspace learning,[6] medical image analysis,[7] and dictionary learning.[8] This study's method to alleviating the problems associated with short circuited graphs revolves around finding a sparse set of neighbors for each point.

2. METHOD

Consider a high dimensional data set X that we want to connect via a graph. Now consider a target sample x_i which will need to be connected to a few neighbors. A reconstruction cost function is used such that:

$$\min_{\omega_i \geq 0} \frac{1}{2} \left\| X\omega_i - x_i \right\|_2^2 + \lambda \left\| \omega_i \right\|_1$$

The first term is the reconstruction error while the second term enforces the ℓ_1 sparsity. Here, λ coefficient controls the contribution of the sparsity term. The additional constraint is added so that there is at least one non-zero weight. Any sample, x_j from X that has a non-zero weight, ω_{ij}, will be considered a neighbor.

In this optimization above, every point in X is a potential neighbor for x_i. Since the number of samples will generally be high, the computational expense could also be a problem. It is safe to assert that the neighborhood should also be close with respect to the ℓ_1 norm. Therefore, the computational problem can be avoided by limiting the number of samples that could be potential neighbors. The neighborhood is first confined to the k nearest neighbors. This will guarantee that neighbors will not be selected far away from the target point and also speeds up the ℓ_1 optimization. Thus, the final optimization becomes:

$$\min_{\omega_i \geq 0} \frac{1}{2} \left\| X^{\{i\}}\omega_i - x_i \right\|_2^2 + \lambda \left\| \omega_i \right\|_1 \tag{1}$$

where X_i are the k nearest neighbors of x_i. The entire graph is found by finding neighbors for all samples in the data set X. The remaining steps of the Isomap algorithm follow normally. Pairwise distances are found using Dijkstra's algorithm. Eigendecomposition of the pairwise distance matrix will give the low dimensional embedding.

3. EXPERIMENTS

First, we test the proposed approach with quantitative benchmark data sets from the University of California Irvine (UCI) repository.[9] These experiments aim to show the proposed approach give an improvement over classical Isomap. Next, the proposed approach is applied to common image data sets.

3.1 UCI data set repository

The graph construction of the proposed ℓ_1 method is tested using five data sets from the UCI repository, namely the Ionosphere, Musk, Sonar, WDBC, and Wine data sets. Manifold learning was performed using the entire data set using the proposed ℓ_1 method and with the standard Isomap method.

Classification was performed on the learned manifold using k-nearest neighbor classification with $k = 5$. In each experiment, 2/3 of the samples are used for training. The remaining 1/3 of the samples are used for testing. The number of nearest neighbors for graph construction is varied from 5 to 15 with each data set. In the case of the proposed ℓ_1 neighbor selection approach, the number of nearest neighbors represents the size of the maximum size of the neighborhood selected. Due to the sparsity property of the proposed graph construction, the actual number of nearest neighbors may be smaller. The parameter controlling the sparsity, λ of the ℓ_1 term was set to $\lambda = .1$. Each experiment was conducted 10 times and the average classification accuracy was recorded.

3.2 UCI data set results

Figures 2 show the results from the experiments. Here, the vertical axis denotes the average classification accuracy. The horizontal axis denote the number of nearest neighbors used for graph construction. In the standard Isomap method, this is a fixed for each neighborhood while this is the maximum number of neighbors, k, in the ℓ_1 method.

The proposed graph construction method showed improvement compared over standard Isomap. In Figures 2, the results for the proposed method were consistently improved for wdbc, wine, and sonar. The ionosphere and musk data sets showed a comparable result between the two methods. Thus, the proposed method either performs better or comparable to the original.

Also from the figures, it can be seen that the proposed method is more robust to the number of nearest neighbors selected. For example in the wine data set of Figure 2, the classification accuracy for classical Isomap degrades after 8 nearest neighbors while the accuracy of the proposed remains relatively stable. The performance drop of the standard Isomap approach may be due to the manifold leakage. This suggests that the ℓ_1-based approach is effective in overcoming the leaking problem.

3.3 IMAGE DATA SETS

In the previous section, numerical data from the UCI machine learning repository are used to individually test the sparse ℓ_1-enhanced graph construction and embedding steps. Next, the proposed manifold learning system is applied to two image data sets; the Columbia University Image Library object database (COIL-20), and the Yale University face database (Yale). These image data sets were selected because of their non-linearity and prevalence in manifold learning literature.

3.3.1 COIL-20 database

The COIL-20[10] image database consists of objects placed on a motorized turntable against a black background. Images were taken while the object was rotated 360 degrees. Thus, the object's pose is altered. A total of 72 images were taken of each object with an rotation interval of 5 degrees. Figure 3 shows sample images of the 20 objects in the COIL-20 database. To normalize the images in this data set, each image is down-sampled to 40×40 pixels.

For this data set, a small number of samples are selected for each subject. Separate experiments are taken with 5, 10, and 15 training samples for each subject. Since there are 20 objects/classes in this data set, the number of training samples are 100, 200, and 300 respectively. Isomap is used to find the manifold from the training samples using the ℓ_1 graph construction method described in the proposed method. The remaining data samples are embedded into the manifold skeleton using the ℓ_1 embedding technique. Classification is performed using k-nn nearest neighbor classification with $k = 1$. Each experiment is repeated 10 times. The average classification accuracy and standard deviation is recorded.

Table 1 shows the results. Figure 4 shows the first two dimensions for the manifold of the proposed approach. Each color represents a different class. The separation of clusters suggest that manifold is unfolded correctly.

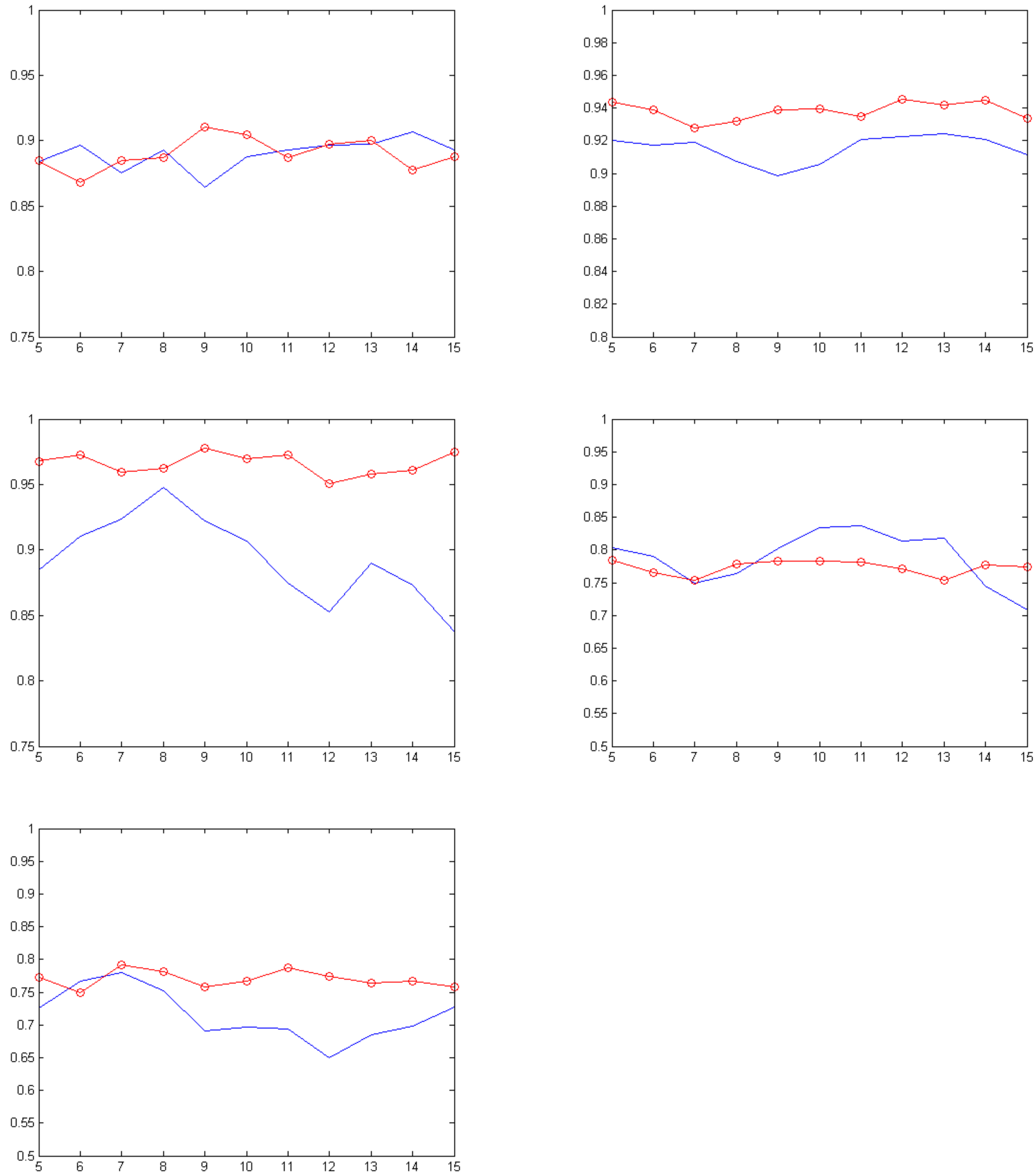

Figure 2. Classification accuracy of databases in UCI repository. Horizontal axis shows the number of nearest neighbors k. The red line with "o" line markers denotes proposed method. Blue line represents Isomap. From left to right, top to bottom. Ionosphere, wdbc, wine, musk, and sonar.

Figure 3. Sample images of 20 objects in COIL-20

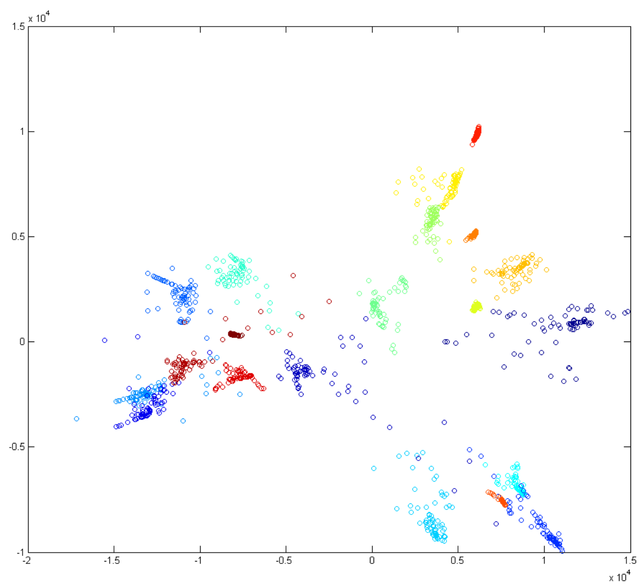

Figure 4. First 2 dimensions of manifold space for COIL-20

Table 1. COIL-20 classification accuracy. Parens show standard deviation.

Method	5 Training Samples	10 Training Samples	15 Training Samples
PCA	81.88 (2.39)	89.14 (1.00)	93.43 (0.73)
LTSA	53.99 (3.09)	55.86 (2.68)	55.45 (3.79)
Isomap	75.11 (3.19)	83.29 (2.76)	88.87 (1.35)
LLE	72.06 (4.07)	84.40 (7.04)	91.73 (0.98)
DONPP	**87.36**	94.11	**97.04**
Proposed ℓ_1	87.17 (1.30)	**94.37 (1.32)**	**97.04 (0.57)**

Figure 5. Sample images of one subject in the Yale face database

Table 2. Yale face database classification accuracy. Parens show standard deviation.

Method	3 Training Samples	5 Training Samples	7 Training Samples
PCA	74.74 (2.65)	79.01 (3.49)	82.46 (3.37)
LTSA	69.67 (5.40)	72.2 (5.60)	80.98 (2.81)
Isomap	68.43 (3.44)	73.63 (3.70)	76.07 (4.18)
LLE	69.32 (7.06)	71.98 (4.49)	76.90 (6.81)
DONPP	68.75	77.44	84.33
Proposed ℓ_1	**81.07 (2.11)**	**85.05 (2.50)**	**88.03 (2.45)**

From Table 1, the proposed approach showed a marked improvement over classical Isomap. The results show that the proposed approach is better than all methods tested except for DONPP,[11] which produced a comparable result. Compared to other methods in each test, the proposed approach's classification accuracy is at least one standard deviation higher. For 5 training samples per class, DONPP is marginally better while the proposed approach is marginally better for 10 training samples per class. The classification accuracy of these two methods are the same for 15 training samples.

3.3.2 Yale database

The Yale database[12] is comprised of 15 subjects. The samples within each subject vary with facial expression and lighting. There are 11 images for each subject for a total of 166 sample images. For each subject, there is one image per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink. A sample set of sample images for one subject of the Yale database is displayed in Figure 5. As with COIL-20, the Yale database is also down-sampled to 40×40 pixels.

For this data set, a small number of samples are selected for each subject. Separate experiments are taken with 3, 5, and 7 training samples for each subject. Since there are 15 objects/classes in this data set, the number of training samples are 45, 75, and 105 respectively. Isomap is used to find the manifold from the training samples using the ℓ_1 graph construction method described in the Proposed method. The remaining data samples are embedded into the manifold skeleton using the ℓ_1 embedding technique. Classification is performed using k-nn nearest neighbor classification with $k = 1$. Each experiment is repeated 10 times. The average classification accuracy and standard deviation is recorded.

Table 2 shows the results. Figure 6 shows the first two dimensions for the manifold of the proposed approach. Each color represents a different class. The separation of clusters suggest that manifold is unfolded correctly. For this image data set, the proposed approach produces a higher classification accuracy than the other methods in each of the tests.

4. CONCLUSIONS

In this study, we developed a new nonlinear graph construction algorithm for Isomap. The main contribution of this study was the introduction of a novel ℓ_1-based neighbor selection algorithm used for graph construction. From these experiments, the adaptive ℓ_1 neighborhood selection showed improvement over the standard methods. The adaptive neighbor selection may alleviate issues of leaking from standard Isomap. This proposed approach was then tested versus manifold learning algorithms from the literature using two image data sets; Yale face data set, and COIL-20 object recognition data set. The proposed approach showed improvement over the linear methods and sparse projection approaches. Compared with DONPP, the proposed approach achieved comparable results with the COIL-20 data set while achieving a better average classification accuracy with the Yale data set.

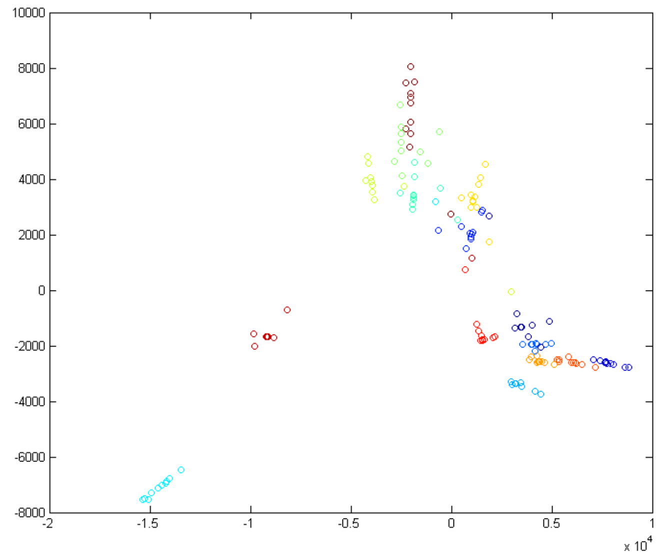

Figure 6. First 2 dimensions of manifold space for yale

REFERENCES

[1] Tenenbaum, J. B., de Silva, V., and Langford, J. C., "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science* **290**, 2319–2323 (Dec. 2000).

[2] van der Maaten, L., Postma, E. O., and van den Herik, H. J., "Dimensionality reduction: A comparative review," (2008).

[3] Wright, J., Yang, A., Ganesh, A., Sastry, S., and Ma, Y., "Robust face recognition via sparse representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **31**, 210–227 (Feb 2009).

[4] Candes, E., Romberg, J., and Tao, T., "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on* **52**, 489–509 (Feb 2006).

[5] Baraniuk, R. G., "More is less: Signal processing and the data deluge," *Science* **331**(6018), 717–719 (2011).

[6] Yang, B. and Chen, S., "Sample-dependent graph construction with application to dimensionality reduction," *Neurocomputing* **74**(1-3), 301 – 314 (2010).

[7] Yuan, L., Wang, Y., Thompson, P. M., Narayan, V. A., and Ye, J., "Multi-source learning for joint analysis of incomplete multi-modality neuroimaging data," in [*Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*], *KDD '12*, 1149–1157, ACM, New York, NY, USA (2012).

[8] Wright, J., Ma, Y., Mairal, J., Sapiro, G., Huang, T., and Yan, S., "Sparse representation for computer vision and pattern recognition," *Proceedings of the IEEE* **98**(6), 1031–1044 (2010).

[9] Bache, K. and Lichman, M., "UCI Machine Learning Repository," (2013).

[10] Nene, S. A., Nayar, S. K., and Murase, H., "Columbia object image library (coil-20)," Tech. Rep. CUCS-005-96, Columbia University (1996).

[11] Zhang, T., Huang, K., Li, X., Yang, J., and Tao, D., "Discriminative orthogonal neighborhood-preserving projections for classification," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* **40**(1), 253–263 (2010).

[12] Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J., "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," (1997).