Order the faces by Diffusion Map, ISOMAP and LLE

Kang Lei

Department of Chemical and Biological Engineering (Bioengineering)

HKUST

lkangaa@connect.ust.hk

Abstract

In this paper, Diffusion map, MDS, ISOMAP and LLE are used to find the low-dimensional embedding of the high-dimensional face images. By comparing the scatter plots of the low-dimensional embedding and the face order graphs, we find that Diffusion map, ISOMAP and LLE can get correct low-dimensional embedding results while MDS performs poorly in this problem. In addition, we compare MDS with ISOMAP and ISOMAP with LLE respectively.

1. Introduction

An assumption that data concentrates around a low dimensional manifold in high dimensional spaces, leads to manifold learning or nonlinear dimensionality reduction, e.g. Diffusion Maps, ISOMAP and LLE. In this paper, the dataset contains 33 faces of the same person in different angles, so it can meet the basic assumption for manifold learning. We try to order the 33 faces using different manifold learning methods, e.g. Diffusion Maps, ISOMAP and LLE. In addition, a comparison between MDS method and manifold learning methods is made in this paper. We also compare the results of different manifold learning methods.

2. Datasets

33 faces of the same person ($Y \in R112\times92\times33$) in different angles.

https://github.com/yao-lab/yao-lab.github.io/blob/master/data/face.mat

3. Methodology

3.1 Diffusion Map

Diffusion maps is a dimension reduction technique that can be used to discover low dimensional structure in high dimensional data. It assumes that the data points, which are given as points in a high dimensional metric space, actually live on a lower dimensional structure. To uncover this structure, diffusion maps builds a neighborhood graph on the data based on the distances between nearby points. Then a graph Laplacian **L** is constructed on the neighborhood graph.

Standard diffusion maps approximates the differential operator:

$$\mathcal{L}f = \Delta f - 2(1-\alpha)\nabla f \cdot \frac{\nabla q}{q}$$

where Δ is the Laplace Beltrami operator, ∇ is the gradient operator and q is the sampling density. The normalization parameter α, which is typically between 0.0 and 1.0, determines how much q is allowed to bias the operator L.

3.2 ISOMAP

ISOMAP is an extension of MDS, where pairwise euclidean distances between data points are replaced by geodesic distances, computed by graph shortest path distances.

3.3 LLE

The algorithm assume that any date point in a high dimensional ambient space can be a linear combination of data points in its neighborhood. In other words, a data point has its neighborhood deciding its sufficient statistics. Alignment of such local linear structures can lead to a global unfolding of data manifolds, often described as fit locally and think globally.

4 Results and analysis

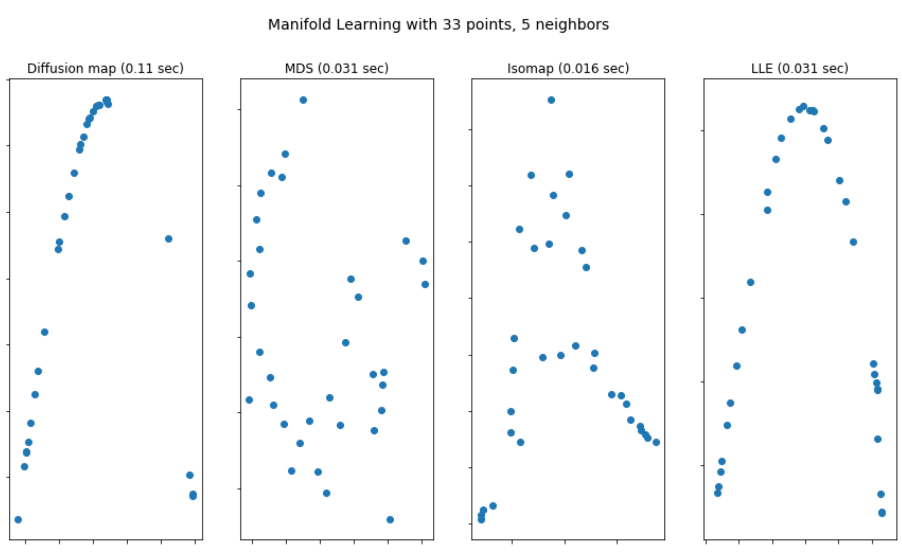

Figure.1 Scatter plots on two eigenvectors of four methods

4.1 Diffusion map explore

The Diffusion mapping is implemented in python.

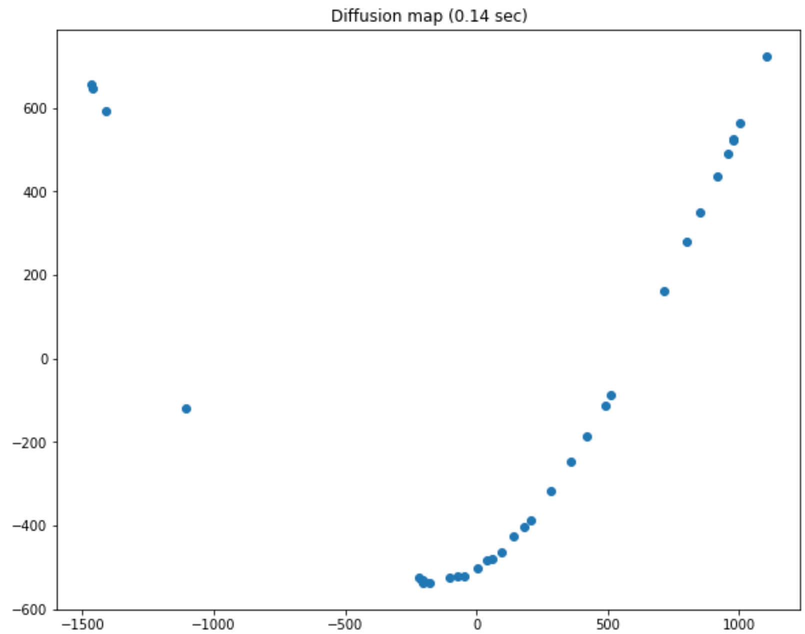

Figure.2 Diffusion map on two eigenvectors

After ordering the face on the first eigenvector, we get the following faces order.

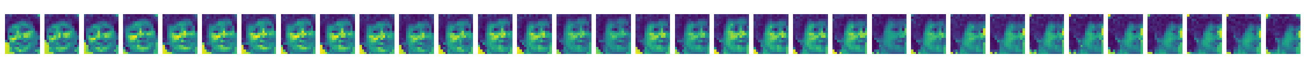

Figure.3 Diffusion map face order

From the figure below, it can been seen that the Diffusion map face order is right and it can successfully show the gradually changed angles of captures faces.

4.2 MDS explore

The MDS is implemented in python.

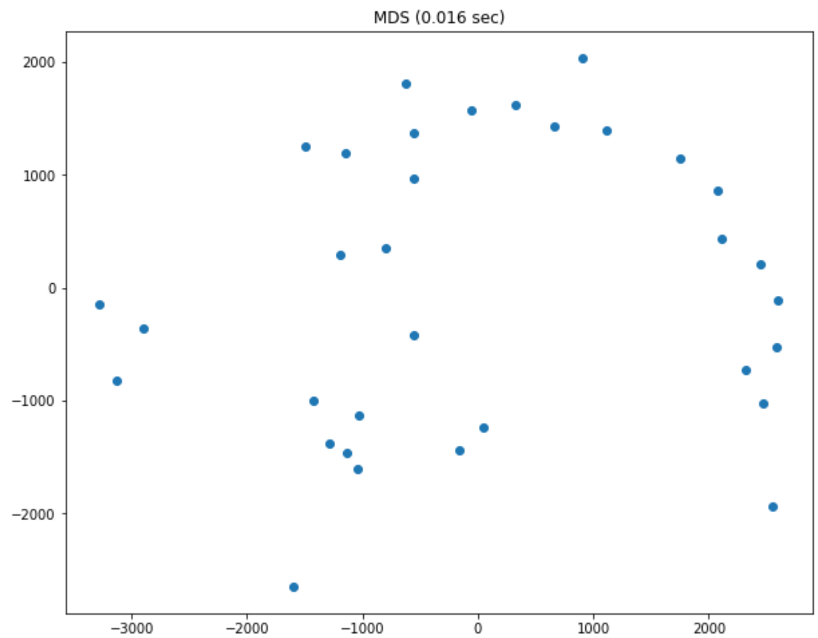

Figure.4 MDS on two eigenvectors

After ordering the face on the first eigenvector, we get the following faces order.

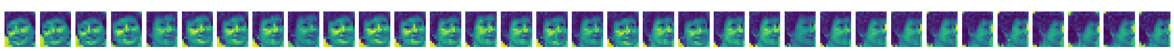

Figure.5 MDS face order

Because MDS is quite unstable on the data cloud, we take another two maps and show the faces sequence respectively.

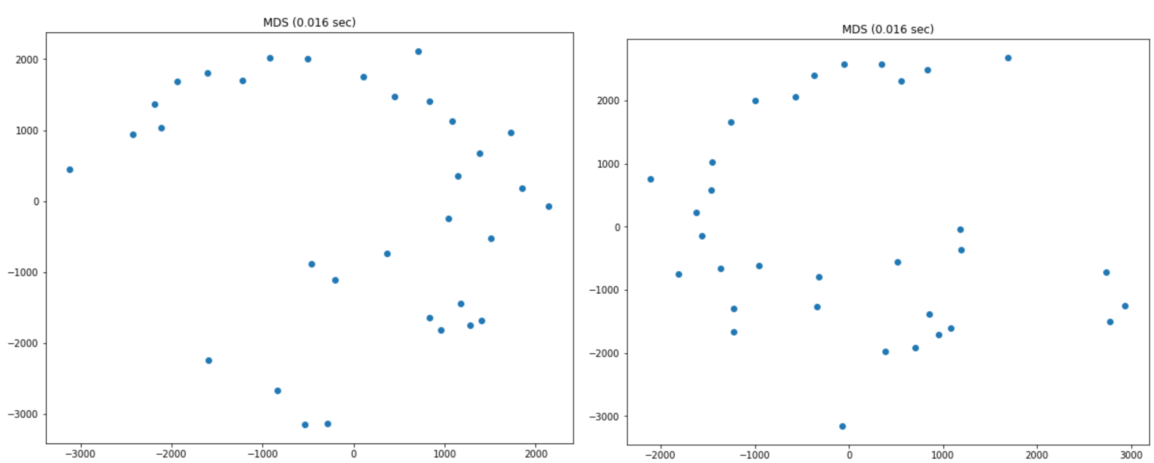

Figure.6 Another two MDS scatter plots

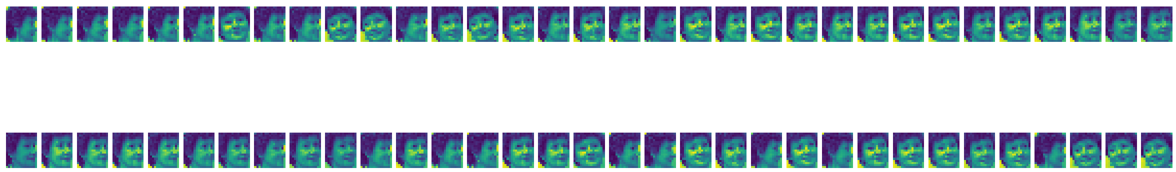

Figure.7 Another two MDS face orders

From the fig.6 and fig.7, we can see that MDS scatter plot is different for each execution of the algorithm and there are several face images are in the wrong position in the face order. This means that MDS performs poorly in nonlinear dimensionality reduction because it uses euclidean distance to calculate the points distance or similarity.

4.3 ISOMAP-embedding

The ISOMAP is implemented in python on the k = 5 nearest neighbor graph.

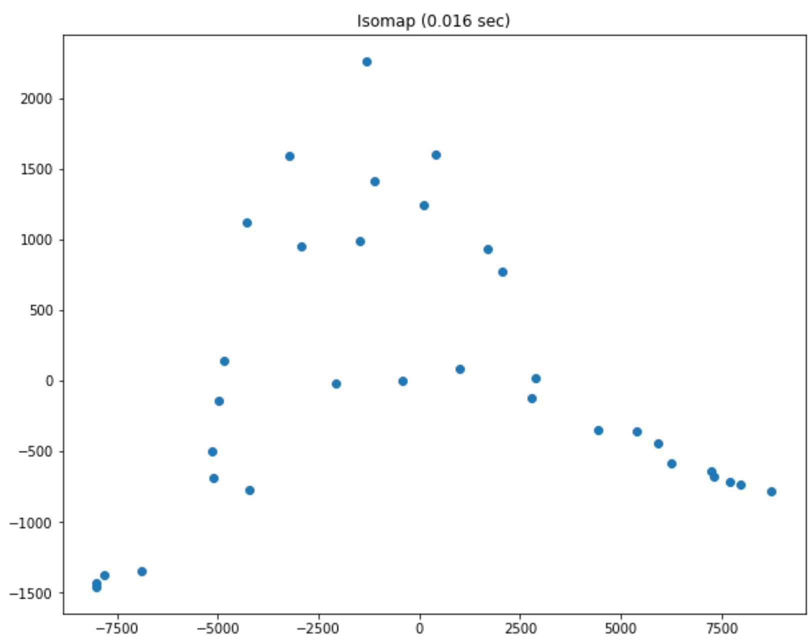

Figure.8 ISOMAP on two eigenvectors

After ordering the face on the first eigenvector, we get the following faces order.

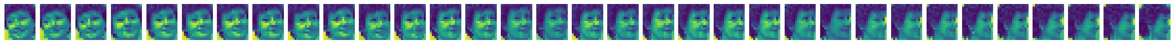

Figure.9 ISOMAP face order.

From the figure below, it can been seen that the ISOMAP face order is right and it can successfully represent the gradually changed angles of captures faces.

Comparison with MDS:

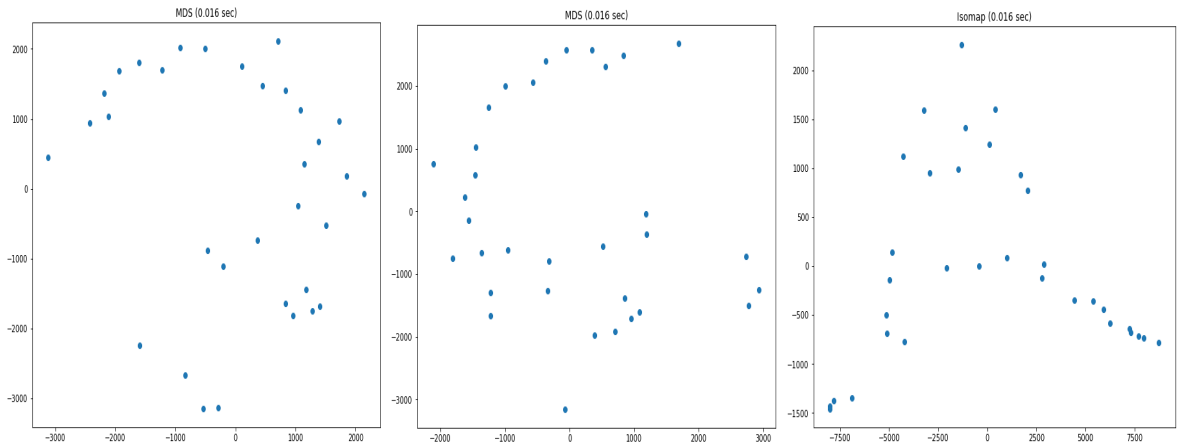

Figure.10 Left: MDS scatter plots, Middle: MDS scatter plots, Right: ISOMAP.

Figure.11 The first and second face order are from MDS, the third one is from ISOMAP.

From the fig.10 and fig.11, we can see ISOMAP is quite stable than MDS and the face order from ISOMAP can successfully show the gradually changed angles of captures images while the MDS face order makes several mistakes. ISOMAP extends MDS by incorporating the geodesic distances imposed by a weighted graph. That's the reason that ISOMAP performs better than MDS in manifold learning.

4.4 LLE-embedding

The ISOMAP is implemented in python on the $k = 5$ nearest neighbor graph.

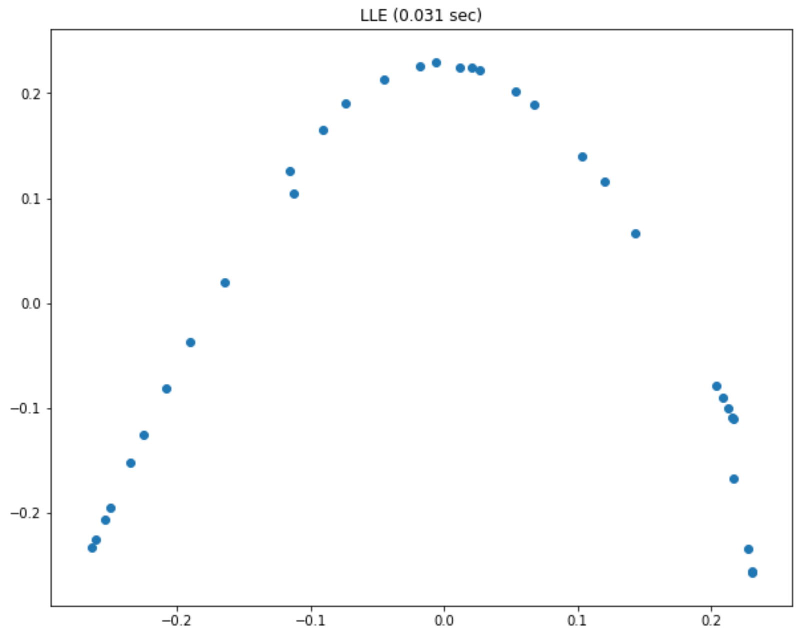

Figure.12 LLE on two eigenvectors

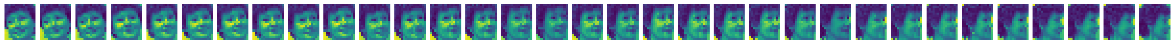

Figure.13 LLE face order

From the fig.13, it can been seen that the LLE face order is right and it can successfully reflect the gradually changed angles of captures faces.

Comparison with ISOMAP:

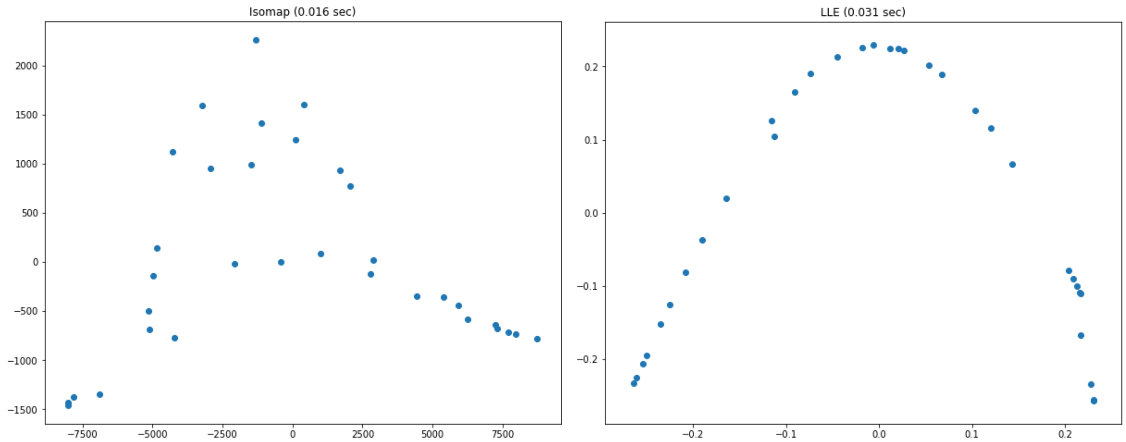

Figure.14 Left: ISOMAP scatter plots, Right: LLE scatter plots

Figure.15 The first is from ISOMAP, the second one is from LLE.

From fig.14, we can see that LLE has smoother curve than ISOMAP. This occurs because that LLE stiches the graph by finding a set of weights that perform local linear interpolations that closely approximate the data while ISOMAP does this by doing a graph traversal.

From fig.15, we can see both LLE and ISOMAP can get correct face order. But LLE spends nearly twice the time of ISOMAP which means it needs more computation.

5 Conclusion

In this paper, we correctly order the face by Diffusion map, ISOMAP and LLE. The results of MDS shows that it performs poor in manifold learning, but its extension-ISOMAP can get correct face order. Compared with ISOMAP, LLE can get smoother scatter plots but with more computation time. The two-dimensional scatter plots of Diffusion map and LLE looks to be similar.

6 Reference

[1] Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.