Applying manifold learning for automatic sorting of facial datasets

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Mingchen Li[1], Shihong Zhang[1], Zinan Lin[1], Lige Zhao[1] {mingchen.li, shihong.zhang, zlinby, lige.zhao}@connect.ust.hk

[1] DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING, HKUST

Introduction

Machine learning techniques, focuses on revealing the underlying low-dimensional structure within high-dimensional data.

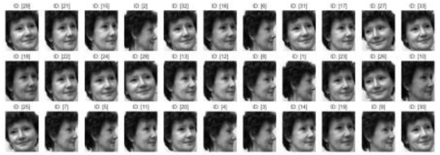
The investigation of face pose estimation is captivating as it allows us to infer an individual's attention and discern relevant behaviors.

The purpose of this project is to categorize the 33 images of a single person according to their facial orientation, employing nine different manifold learning techniques. Our approach entailed obtaining the embedding outcomes from each method and arranging them based on the magnitude of their first eigenvector. Ultimately, we compared the outcomes of the manifold learning methods employed.

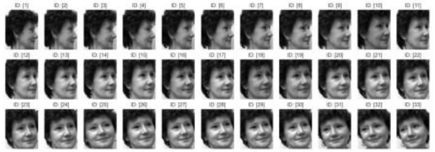
Dataset

The dataset contains 33 faces of the same person ($Y \in \mathbb{R}^{112\times92\times33}$) in different angles. The original data and the data adjusted to true values are shown below. All the snapshots labeled by the figure ID.

Original Data



Ground truth



Methodology

In this project, various manifold learning methods have been implemented to discover the essential features within a set of face images belonging to a single woman. The following section provides detailed information on each approach utilized.

Diffusion maps [1] aims to capture the underlying structure of high-dimensional data by modeling the diffusion process on a graph. It constructs a graph representation of the data, where each data point is connected to its nearest neighbors.

Multidimensional Scaling (MDS) [2] is a technique used to visualize the similarity or dissimilarity between objects based on their pairwise distances. It aims to find a low-dimensional representation of the data points in a way that preserves the pairwise distances as much as possible

ISOMAP (Isometric Feature Mapping) [3] focuses on preserving the geodesic distances between data points. It constructs a neighborhood graph based on the nearest neighbors and then approximates the geodesic distances on this graph. By performing classical multidimensional scaling on these geodesic distances, ISOMAP obtains a low-dimensional embedding that captures the underlying manifold structure.

Locally Linear Embedding (LLE) [4] aims to reconstruct the local linear relationships between neighboring data points. It starts by identifying the nearest neighbors for each data point and then seeks a low-dimensional representation where each data point can be reconstructed as a linear combination of its neighbors

Modified Locally Linear Embedding (MLLE) [5] is an extension of LLE that addresses some of its limitations. MLLE overcomes the sensitivity of LLE to noise and outliers by incorporating a two-step process. It first constructs a neighborhood graph similar to LLE, and then applies a global linear model to capture the relationships between the low-dimensional embeddings.

Local Tangent Space Alignment (LTSA) [6] focuses on preserving the local geometry of the data. It starts by computing the local tangent spaces for each data point based on its neighbors, and then aligns these tangent spaces to obtain a global low-dimensional representation.

t-Distributed Stochastic Neighbor Embedding (t-SNE) [7] focuses on preserving the local and global relationships between data points. It constructs a probability distribution over pairs of data points in the high-dimensional space and a similar distribution in the low-dimensional space.

Principal Component Analysis (PCA) [8] is a widely used linear dimensionality reduction technique that aims to find orthogonal axes, called principal components, along which the data exhibits the maximum variance. It projects the high-dimensional data onto a lower-dimensional subspace while preserving the most important information.

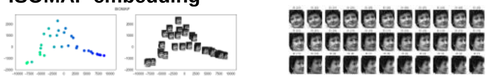
Uniform Manifold Approximation and Projection (UMAP) [9] is a nonlinear dimensionality reduction method that combines aspects of both graph-based and manifold-based approaches. It constructs a fuzzy topological representation of the data using a graph structure and optimizes a low-dimensional embedding that preserves both local and global structures.

Results

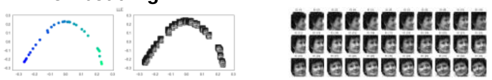
Diffusion map embedding

MDS embedding

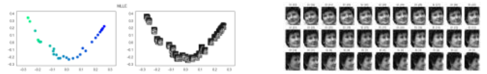
ISOMAP embedding



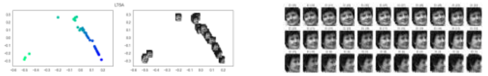
LLE embedding



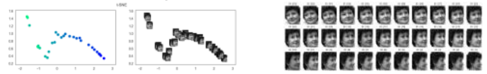
MLLE embedding



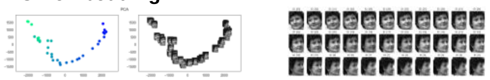
LTSA embedding



t-SNE embedding



PCA embedding



UMAP embedding

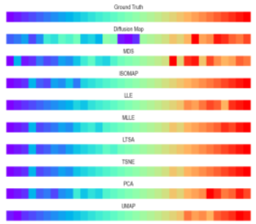
Evaluation

Evaluation criteria

The Kendall tau correlation coefficient is utilized as a measure of association or correlation between two ranked variables. It is used to determine the similarity of the orderings of the data points between the two variables.

Evaluation result (absolute value)

Diffusion: 0.674
MDS: 0.818
ISOMAP: 0.962
LLE: 0.970
MLLE: 0.962
LTSA: 0.973
TSNE: 0.973
PCA: 0.920
UMAP: 0.973



Conclusion

This project involved investigating the performance of various manifold learning methods in revealing the low-dimensional structure within high-dimensional data. The dataset used was small and featured variations in face orientation. Our experiments demonstrated that TSNE, LTSA, and UMAP exhibited nearly flawless performance. However, the diffusion map embedding method proved to be less robust and yielded significantly inferior results compared to the other methods.

Contribution

Mingchen Li: Experiment design, methodology, coding, data
Shihong Zhang: Experiment design, methodology, post, edit
Zinan Lin: Experiment design, methodology, presentation, video
Lige Zhao: Experiment design, methodology, slides, review

References

[1] Coifman, R. R., & Lafon, S. (2006). Diffusion maps. Applied and computational harmonic analysis, 21(1), 5-30.
[2] Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. Psychometrika, 17(4), 401-419.
[3] Tenenbaum, J. B., de Silva, V., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. science, 290(5500), 2319-2323.
[4] Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. science, 290(5500), 2323-2326.
[5] Zhang, Z., & Zha, H. (2004). Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. SIAM Journal on scientific computing, 26(1), 313-338.
[6] Zhang, Z., & Zha, H. (2002). Principal manifolds and nonlinear dimensionality reduction via local tangent space alignment. In Proceedings of the 2002 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Vol. 2, pp. II-II). IEEE.
[7] Maaten, L. V. D., & Hinton, G. (2008). Visualizing data using t-SNE. Journal of machine learning research, 9(Nov), 2579-2605.
[8] Jolliffe, I. T. (2002). Principal component analysis. Wiley Online Library.
[9] McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.