

Visualizing Data using t-SNE

Taran Lynn, Xiaoli Yang, Xiaoxin Chen

October 20, 2020

Isometric Mapping (Isomap)

An extension of Multi-dimensional Scaling or Kernel PCA

- ▶ Nearest neighbor search
- ▶ Shortest-path graph search
- ▶ Partial eigenvalue decomposition

Locally Linear Embedding (LLE)

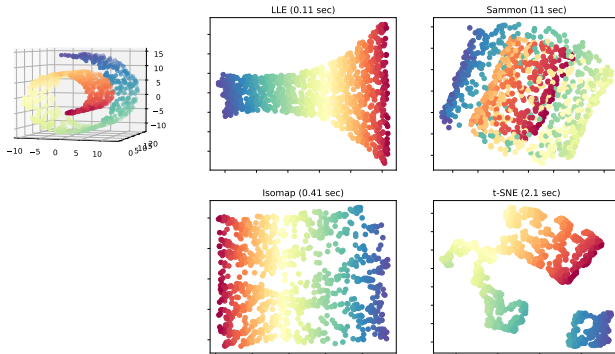
A series of local Principal Component Analyses

- ▶ Nearest neighbor search
- ▶ Weight Matrix Construction
- ▶ Partial Eigenvalue Decomposition

Sammon Mapping

Algorithm Comparison

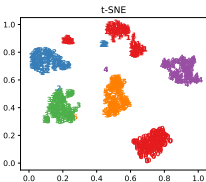
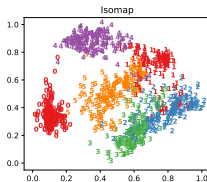
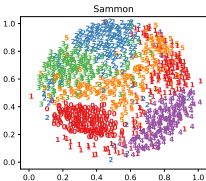
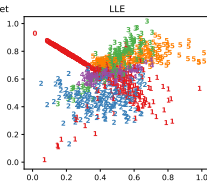
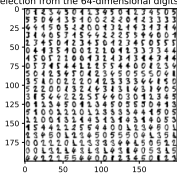
Manifold Learning with 1000 points, 10 neighbors



Algorithm Comparison

MNIST dataset, 30 neighbors

A selection from the 64-dimensional digits dataset



Weakness of t-SNE

- ▶ Dimensionality reduction for other purposes
- ▶ Curse of intrinsic dimensionality
- ▶ Non-convexity of the t-SNE cost function

Conclusion

Computational complexity: $O(n^2)$

Memory complexity: $O(n^2)$