Local Neighbor Propagation Embedding

Wenduo Ma[1], Hengzhi Yu[1], Shenglan Liu[2(✉)], Yang Yu[3], and Yunheng Li[1]

[1] Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, Liaoning, China

[2] School of Innovation and Entrepreneurship, Dalian University of Technology, Dalian 116024, Liaoning, China

liusl@dlut.edu.cn

[3] The Alibaba Group, Hangzhou 310051, Zhejiang, China

Abstract. Manifold learning occupies a vital role in the field of nonlinear dimensionality reduction and serves many relevant methods. Typical linear embedding methods have achieved remarkable results. However, these methods perform poorly on sparse data. To address this issue, we introduce neighbor propagation into Locally Linear Embedding (LLE) and propose a new method named Local Neighbor Propagation Embedding (LNPE). LNPE enhances the local connections between neighborhoods by extending 1-hop neighbors into n-hop neighbors. The experimental results show that LNPE can obtain faithful embeddings with better topological and geometrical properties.

Keywords: Local Neighbor Propagation · Manifold Learning · Geometrical Properties · Topological

1 Introduction

Over the past few decades, manifold learning has already attracted extensive attention and applied in video prediction [2], video content identification [14], spectral reconstruction [10], image super-resolution [3], etc. Manifold learning research can be divided into local methods such as Locally Linear Embedding [18], Local Tangent Space Alignment (LTSA) [25] and some global methods such as Isometric Mapping (ISOMAP) [20] and Maximum Variance Unfolding (MVU) [22]. Besides, compared with nonlinear manifold learning methods, some linear embedding methods are more effective for practical applications such as Locality Preserving Projection (LPP) [7], Neighborhood Preserving Embedding (NPE) [6], Neighborhood Preserving Projection (NPP) [15].

Locally linear methods in manifold learning such as LLE have been widely studied and applied. Based on LLE, many improved methods are proposed, such as Hessian Locally Linear Embedding (HLLE) [4], Modified Locally Linear Embedding (MLLE) [24] and Improved Locally Linear Embedding (ILLE) [23]. Among the LLE-based methods, HLLE has high computational complexity and MLLE does not work on weak-connected data. Actually, the computational complexity and robustness of algorithms are issues to be considered for

Q. Liu et al. (Eds.): PRCV 2023, LNCS 14427, pp. 239–249, 2024.
https://doi.org/10.1007/978-981-99-8435-0_19

many LLE-based improved methods. Inspired by Graph Convolutional Networks (GCNs), we propose a simple and unambiguous LLE-based improved method, named Local Neighbor Propagation Embedding (LNPE). In contrast to previous approaches, LNPE extends the neighborhood size through neighbor propagation layer by layer. This architecture enhances the topological connections within each neighborhood. From the view of global mapping, neighborhood interactions increase with neighbor propagation to improve the global geometrical properties. Experimental results verify the effectiveness and robustness of LNPE.

The remainder of this paper is organized as follows. We first introduce related work in Sect. 2. Section 3 presents the main body of this paper, which includes the motivation of LNPE, the mathematical background, the framework of LNPE and the analysis of computational complexity. The experimental results are presented in Sect. 4 to verify the effectiveness and robustness of LNPE and Sect. 5 summarizes our work.

2 Related Work

The essence of manifold learning is how to maintain the relationship corresponding to the intrinsic structure between samples in two different spaces. Researchers have done lots of work to measure the relationship from different aspects.

PCA maximizes the global variance to reduce dimensionality, while Multidimensional Scaling (MDS) [9] considers the low-dimensional distance between samples which is consistent with high-dimensional data. Based on MDS, ISOMAP utilize the shortest path algorithm to realize global mapping. Besides, MVU realizes an unfolding manifold through positive semi-definite and kernel technique and RML [11] obtains the intrinsic structure of the manifold with the Riemannian methods instead of Euclidean distance. Compared with ISOMAP, LLE represents local manifold learning, which obtains neighbor weights with locally linear reconstruction. Furthermore, Locally Linear Coordination (LLC) [17] constructs a local model and makes a global alignment and contributes to both LLE and LTSA. LTSA describes local curvature through the tangent space of samples and takes the curvature as the weight of tangent space to realize global alignment.

Another type of manifold structure preserving is graph-based embedding. Classical methods are still representative such as LE [1] and its linear version LPP. More related methods include NPE, Orthogonal Neighborhood Preserving Projections (ONPP) [8], etc. The graph-based method produces a far-reaching influence on machine learning and related fields. For instance, the graph is introduced into semi-supervised learning in LE, and LLE is also a kind of neighbor graph. Moreover, L-1 graph-based methods such as Sparsity preserving Projections (SPP) [16] and its supervised extension [5] are proposed with the wide application of sparse methods.

3 Local Neighbor Propagation Embedding

In this section, the idea of neighbor propagation is introduced gradually. First, we present the motivation for LNPE and describe the origin of LNPE. Then the basic algorithm prototype is briefly introduced. Finally, the LNPE framework is introduced.

3.1 Motivation

LLE, which is a classical method, should be representative of methods based on local information in manifold learning. In the case of simple data distribution, LLE tends to get satisfactory results. But once the data distribution becomes sparse, it is difficult for LLE to maintain topological and geometrical properties. The following items show the reasons.

1. The neighborhood size is hard to determine for sparse data. Inappropriate neighbors will be selected with a larger neighborhood size.
2. LLE focuses more on each single neighborhood, but is weak in the interaction between different neighborhoods. Thus, it is difficult to obtain the ideal effect in geometrical structure preservation.

There is a natural contradiction between Item 1 and Item 2. More specifically, the interaction of neighborhood information will be weakened by small neighborhoods inevitably. To improve the capability and robustness of LLE, we introduce neighbor propagation into LLE and propose Local Neighbor Propagation Embedding (LNPE), inspired by GCNs. The 1-hop neighborhoods are extended to n-hop neighborhoods through neighbor propagation, which will enhance the connections of points in different neighborhoods. In this way, LNPE avoids short circuits (Item 1) by setting a small neighborhood size k. Meanwhile, i-hop neighborhoods expand neighborhood size to depict local parts adequately by enhancing the topological connections. Furthermore, correlations between different neighborhoods (Item 2) are generated in neighbor propagation to produce more overlapping information, which is conducive to preserving the geometrical structure.

3.2 Mathematical Background

Based on LLE, LNPE introduces neighbor propagation to improve its applicability. Before starting the LNPE framework, we first review the original LLE.

Suppose $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \mathrm{R}^D$ indicates a high-dimensional dataset that lies on a smooth D-dimensional manifold, LLE tends to embed the intrinsic manifold from high-dimensional space into lower-dimensional subspace with preserving geometrical and topological structures. Based on the assumption of local linearity, LLE firstly reconstructs each high-dimensional data point $\mathbf{x}_i$ through linear combination within each neighborhood $\mathbf{N}_i$, where $\mathbf{N}_i$ indicates the k-nearest neighbors of $\mathbf{x}_i$ and $\mathbf{N}_i = \{\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_k}\}$. Then the reconstruction weights matrix $\mathbf{W}$ in high-dimensional space can be determined by minimizing the total reconstruction error ε_1 of all data points.

$$\varepsilon_1(\mathbf{W}) = \|\mathbf{X}\mathbf{W} - \mathbf{X}\|_F^2 \tag{1}$$

where $\mathbf{W} = [\vec{\mathbf{w}}_1, \vec{\mathbf{w}}_2, \cdots, \vec{\mathbf{w}}_n] \in \mathrm{R}^{n\times n}$, the i-th column vector $\vec{\mathbf{w}}_i$ indicates the reconstruction weights of data point $\mathbf{x}_i$ and $\|\cdot\|_F$ is the Frobenius norm. To remove the influence of transformations including translation, scaling, and rotation, a sum-to-one constraint $\vec{\mathbf{w}}_i^T\vec{1} = 1$ is enforced for each neighborhood.

Let $\mathbf{Y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_n\} \subset \mathrm{R}^d$ be the corresponding dataset in low dimensions. The purpose of LLE is to preserve the same local structures reconstructed in high-dimensional space. Then in the low-dimensional space, LLE chooses to utilize the same weights $\mathbf{W}$ to reproduce the local properties. The objective is to minimize the total cost function

$$\varepsilon_2(\mathbf{Y}) = \|\mathbf{Y}\mathbf{W} - \mathbf{Y}\|_F^2 \tag{2}$$

under the constraint $\mathbf{Y}\mathbf{Y}^T = \mathbf{I}$. Thus, the high-dimensional coordinates are finally mapped into lower-dimensional observation space.

3.3 Local Neighbor Propagation Framework

A faithful embedding is root in more sufficient within-neighborhood and between-neighborhood information of local and global structure. For LLE, the interactive relationship in single neighborhood is not enough to reproduce detailed data distribution, especially when the neighborhood size k is not large enough. Neighborhood propagation is introduced into LLE to intensify the topological connections within neighborhoods and interactions between neighborhoods.

High-Dimensional Reconstruction. Based on the single reconstruction in each neighborhood, LNPE propagates neighborhoods and determines propagating weight matrix. Similarly, we define the high-dimensional and low-dimensional data as $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset \mathrm{R}^D$ and $\mathbf{Y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_n\} \subset \mathrm{R}^d$, respectively. Suppose that we have finished the single reconstruction with LLE (Eq. 1) and got the weight matrix $\mathbf{W}_1$. In the meantime, the first reconstructed data $\mathbf{X}^{(1)} = \mathbf{X}\mathbf{W}_1$ are obtained from the first reconstruction with LLE. LNPE is to reuse the reconstructed data such as $\mathbf{X}_1$ to reconstruct the original data points in $\mathbf{X}$ again with the reconstructed data previously. Then the first neighborhood propagation is to utilize $\mathbf{X}_1$ to reconstruct original data $\mathbf{X}$ through

$$\varepsilon(\mathbf{W}_2) = \|\mathbf{X}\mathbf{W}_1\mathbf{W}_2 - \mathbf{X}\|_F^2 \tag{3}$$

where $\mathbf{W}_2$ corresponds to the weight matrix in the first neighbor propagation. Thus, we can simply combine the weight matrices $\mathbf{W}_1$ and $\mathbf{W}_2$ as $\mathbf{W} = \mathbf{W}_1\mathbf{W}_2$. Compared with $\mathbf{W}_1$ in LLE, $\mathbf{W}$ extends the 1-hop neighbors to 2-hop neighbors, which expands the neighborhood size through neighbor propagation. One of the most important advantages is that neighbor propagation can enhance topological connections while avoiding short circuits. Besides, a truth worth noting is that the multi-hop neighbors hold lower weights in the process of propagation. Then after $i-1$ propagations, each data point to be reconstructed establishes relations with its i-hop neighbors. The $(i-1)$-th neighbor propagation can be formulated as

$$\varepsilon(\mathbf{W}_i) = \|\mathbf{X}\mathbf{W}_1 \cdots \mathbf{W}_{i-1}\mathbf{W}_i - \mathbf{X}\|_F^2 \tag{4}$$

where $\mathbf{W}_i$ indicates the weight matrix in the $(i-1)$-th neighbor propagation. From the perspective of weight solution and optimization, the i-th neighbor propagation depends on all the first $i-1$ weight matrices and each weight matrix in the propagation must be determined in turn.

Global Low-Dimensional Mapping. Similar to LLE, the low-dimensional embedding in LNPE is to reproduce the high-dimensional properties determined in the reconstruction. LNPE aims to preserve all the topological connections from 1-hop neighborhoods to n-hop neighborhoods with n weight matrices. We define that the matrix product $\mathbf{P}_i = \mathbf{W}_1\mathbf{W}_2 \cdots \mathbf{W}_i$ is the product of weight matrices in the first $i-1$ neighbor propagation. And then the low-dimensional total optimization function can be expressed as

$$\varepsilon(\mathbf{Y}) = \sum_{i=1}^{t+1} \|\mathbf{Y}\mathbf{P}_i - \mathbf{Y}\|_F^2 \tag{5}$$

where the parameter $t+1$ denotes the total hop in the high-dimensional reconstruction with t neighbor propagations. According to the properties of F-norm, the low-dimensional objective Eq. 6 can be written as

$$\varepsilon(\mathbf{Y}) = \mathbf{Y}\left(\sum_{i=1}^{t+1}(\mathbf{P}_i - \mathbf{I})(\mathbf{P}_i - \mathbf{I})^T\right)\mathbf{Y}^T \tag{6}$$

where $\mathbf{I} \in \mathrm{R}^{n\times n}$ indicates the identity matrix. Under the constraint $\mathbf{Y}\mathbf{Y} = I$, the low-dimensional coordinates can be easily obtained by decomposing the target matrix. The detailed algorithm is shown as Algorithm 1.

Equation 5 and Eq. 6 show that LNPE aims to preserve all the learned properties in the high-dimensional reconstruction. Specifically speaking, for the sequence $i = 1, 2, \cdots, t$, a smaller i aims at maintaining the topological connections within each neighborhood, while a larger i pays more attention to expanding the neighborhood interactions between different neighborhoods.

3.4 Computational Complexity

The computational complexity of LNPE follows LLE. Calculating the k nearest neighbors scales as $O(Dn^2)$. In some special data distributions, the computational complexity can be reduced to $O(n \log n)$ with K-D trees [19]. Computing the weight matrix in $t+1$ reconstructions scales as $O((t+1)nk^3)$. Besides,

Algorithm 1. *LNPE Algorithm*

Require:
 high-dimensional data $\mathbf{X} \subset \mathrm{R}^D$;
 neighborhood size k;
 target dimensionality d;
 the neighbor propagation times t.
Ensure:
 low-dimensional coordinates $\mathbf{Y} \subset \mathrm{R}^d$;
 Find the k-nearest neighbors $\mathbf{N}_i = \{\mathbf{x}_{i_1}, \cdots, \mathbf{x}_{i_k}\}$ for each data point.
 Compute the weight matrix $\mathbf{W}_1$ with LLE.
 Initialize a zero matrix $\mathbf{M}$.
 for $e = 1 : t$ **do**
  Compute the matrix product $\mathbf{P}_e = \mathbf{W}_1\mathbf{W}_2 \cdots \mathbf{W}_e$
  Compute $\mathbf{M} = \mathbf{M} + (\mathbf{P}_i - \mathbf{I})(\mathbf{P}_i - \mathbf{I})^T$
  Compute the reconstruction data $\mathbf{X}^{(e)} = \mathbf{X}\mathbf{P}_e$
  for each sample $\mathbf{x}_i, i = 1, \cdots, n$ **do**
   Compute $\vec{\mathbf{w}}_i^{(e+1)}$ in $\mathbf{W}_{e+1}$ with $\mathbf{X}$ and $\mathbf{X}^{(e)}$ through minimizing $\|\mathbf{x}_i - \mathbf{X}^{(e)}\vec{\mathbf{w}}_i^{(e+1)}\|_2^2$
  end for
 end for
 Compute $\mathbf{M} = \mathbf{M} + (\mathbf{P}_{t+1} - \mathbf{I})(\mathbf{P}_{t+1} - \mathbf{I})^T$
 Solve $\mathbf{Y} = \arg\min_{\mathbf{Y}} \mathrm{Tr}(\mathbf{Y}\mathbf{M}\mathbf{Y}^T)$

computing matrices $\mathbf{P}$ and $\mathbf{M}$ scales as $O(p_1 n^3)$ with sparse matrix and the final calculation of eigenvectors of a sparse matrix has computational complexity $O(p_2 n^2)$, where p_1 and p_2 is parameters related to the ratio of nonzero elements in sparse matrix [21].

4 Experimental Results

In this section, we verify LNPE on both the synthetic and real-world datasets. To conduct a fair comparison, we set the same configuration for the comparison methods and LNPE.

4.1 Synthetic Datasets

On sparse data, curvature-based dimensionality reduction methods, such as HLLE [4] and LTSA [25], have performance degradation. This is because the instability of local relationship among sparse data increases, resulting in less smoothness of the local curvature. As shown in Fig. 1, HLLE [4] is unable to accomplish dimensionality reduction on the S-curve and the Swiss, while the dimensionality reduction performance of LTSA [25] is unstable. When the neighborhood size k is increased, the dimensionality reduction performance of LTSA improves. However, it is not a good tendency to obtain better dimensionality reduction performance by increasing the neighborhood size k, because too large

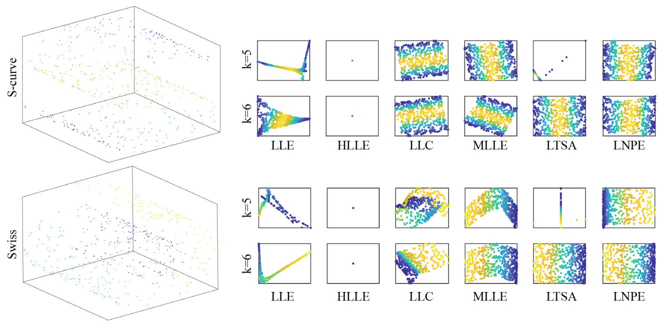

Fig. 1. The S-Curve with 500 samples and its embedding results and the Swiss with 500 samples and its embedding results.

k will lead to short circuits. For LLC [17], each point is represented as a linear combination of other points, where the weights used are controlled by a local constraint matrix. When the data become sparse, the constraint matrix is easily disturbed by noise and outliers, which leads to unstable and inaccurate weights. Unlike these methods, LNPE can better capture the relationship between data through local neighborhood propagation. This is because LNPE can not only better model the linear relationship in the neighborhood, but also strengthen the relationship between the center point and the points outside the neighborhood. LNPE can learn the topological and geometrical structure of the data and embed it in the low-dimensional space faithfully.

To further verify that LNPE has outstanding dimensionality reduction performance on data with complex topology, we select the Swiss-Hole and the Changing-Swiss for our experiments, as shown in Fig. 2. For complex data, the relationship between local points and global neighbors is not easy to maintain. With LLE, the connection of a single neighborhood is not sufficient to fully describe the original data distribution. Especially, LLE performs poorly when the neighborhood size k is not large enough, and too large k can cause short circuits. HLLE [4] has a tendency to depend on the neighborhood size k for the Swiss Hole and cannot accomplish the dimensionality reduction on the Changing-Swiss. Although LTSA [25] performs well on the Swiss Hole, it does not achieve satisfactory results on the Changing-Swiss. MLLE and LLC tend to project the high-dimensional data directly into the low-dimensional space, which leads to the inability to restore topological and geometrical structures well. LNPE is able to completely unfold the topological structure of the original data on the low-dimensional space.

4.2 Real-World Datasets

To evaluate LNPE on real-world applications, we conduct experiments on pose estimation tasks including the Statue face dataset and Teapot dataset. The Statue face dataset, which consists of 698 images with $64 \times 64 = 4096$ pix-

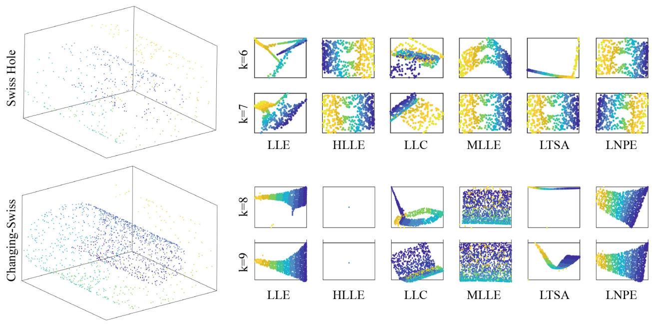

Fig. 2. The Swiss Hole with 500 samples and its embedding results and the Changing-Swiss with 1500 samples and its embedding results.

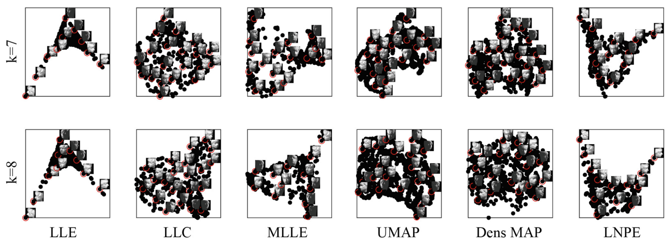

Fig. 3. The Statue Face dataset and its embedding results.

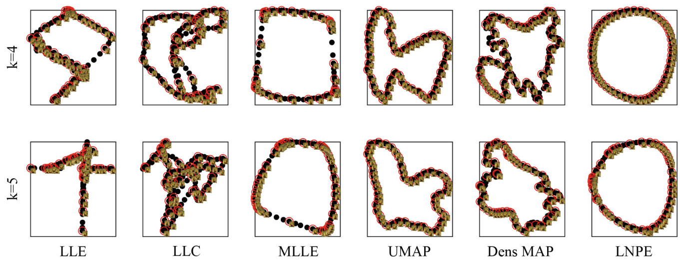

Fig. 4. The teapot dataset and its embedding results.

els, is first used in ISOMAP [20]. The Teapot dataset is a set of 400 teapot images in total [22]. Each image in the Teapot dataset can be seen as a high-dimensional vector consisting of 76 × 101 × 3 RGB pixels. Samples in these two datasets lie on smooth underlying manifolds. All images in the two datasets are tensed into their respective data spaces by bit-pose variations, and all data points obey a streamlined data distribution in this high-dimensional data space. To test the performance of LNPE on sparse data, we downsampled the number of samples on both datasets to half of the original which is feasible due to uniform distribution.

Figure 3 shows the low-dimensional embedding of LNPE and the comparison algorithms on the Statue face dataset, where the experimental results are compared under different choices of nearest neighbor ranges. Specifically, the LLC [17], UMAP [12] and DensMAP [13] perform poorly, and the topology and geometric structure of their embedding results are obviously destroyed. The LLE and MLLE can maintain good topology, but their geometric structure remains weak. In contrast, since MLLE employs a multi-weight structure, it has an advantage in geometric structure retention. The LNPE proposed in this paper is able to recover well the geometric structure at k=7, while it is better in geometric structure maintenance at k=8. Meanwhile, the dimensionality reduction results of LNPE do not exhibit significant topological damage, which indicates that the LNPE is more robust in the neighborhood range. To validate the efficacy of LNPE, we also compare our LNPE with other methods on the teapot dataset are shown in Fig. 4. The previous dimensionality reduction methods perform poorly when $k = 4$ and 5. In the same setup of neighbor, LNPE achieves better dimensionality reduction and obtains good geometric and topological properties.

Overall, compared with previous dimensionality reduction methods, LNPE can achieve competitive dimensionality reduction results with less dependence on the neighborhood size k. This is because it captures the global and local relationship of the data well and has good robustness to sparse data.

5 Conclusion

LLE-based methods usually fail in addressing sparse data where the local connections and neighborhood interactions are difficult to obtain. In this paper, we introduce neighbor propagation in LLE and propose Local Neighbor Propagation Embedding. LNPE expands 1-hop neighborhoods to n-hop neighborhoods, and can be solved iteratively based on the solution of LLE. Although computational complexity increases, LNPE enhances topological connections within neighborhoods and interactions between neighborhoods. Experimental results on both synthetic and real-world datasets show that LNPE improves embedding performance and is more robust than previous methods.

Author contributions. Wenduo Ma and Hengzhi Yu–These authors contributed equally to this work.

References

1. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. **15**(6), 1373–1396 (2003)
2. Cai, Y., Mohan, S., Niranjan, A., Jain, N., Cloninger, A., Das, S.: A manifold learning based video prediction approach for deep motion transfer. In: 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW) pp. 4214–4221 (2021). https://doi.org/10.1109/ICCVW54120.2021.00470

3. Dang, C., Aghagolzadeh, M., Radha, H.: Image super-resolution via local self-learning manifold approximation. IEEE Signal Process. Lett. **21**(10), 1245–1249 (2014). https://doi.org/10.1109/LSP.2014.2332118
4. Donoho, D.L., Grimes, C.: Hessian eigenmaps: locally linear embedding techniques for high-dimensional data. Proc. Natl. Acad. Sci. **100**(10), 5591–5596 (2003)
5. Gui, J., Sun, Z., Jia, W., Hu, R., Lei, Y., Ji, S.: Discriminant sparse neighborhood preserving embedding for face recognition. Pattern Recogn. **45**(8), 2884–2893 (2012)
6. He, X., Cai, D., Yan, S., Zhang, H.J.: Neighborhood preserving embedding. In: Tenth IEEE International Conference on Computer Vision (ICCV 2005) Volume 1, vol. 2, pp. 1208–1213. IEEE (2005)
7. He, X., Niyogi, P.: Locality preserving projections. In: Advances in Neural Information Processing Systems 16 (2003)
8. Kokiopoulou, E., Saad, Y.: Orthogonal neighborhood preserving projections. In: Fifth IEEE International Conference on Data Mining (ICDM 2005), pp. 8–pp. IEEE (2005)
9. Kruskal, J.B., Wish, M.: Multidimensional scaling, vol. 11. Sage (1978)
10. Li, Y., Wang, C., Zhao, J.: Locally linear embedded sparse coding for spectral reconstruction from rgb images. IEEE Signal Process. Lett. **25**(3), 363–367 (2018). https://doi.org/10.1109/LSP.2017.2776167
11. Lin, T., Zha, H.: Riemannian manifold learning. IEEE Trans. Pattern Anal. Mach. Intell. **30**(5), 796–809 (2008)
12. McInnes, L., Healy, J., Melville, J.: Umap: uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018)
13. Narayan, A., Berger, B., Cho, H.: Assessing single-cell transcriptomic variability through density-preserving data visualization. Nat. Biotechnol. (2021). https://doi.org/10.1038/s41587-020-00801-7
14. Nie, X., Liu, J., Sun, J., Liu, W.: Robust video hashing based on double-layer embedding. IEEE Signal Process. Lett. **18**(5), 307–310 (2011). https://doi.org/10.1109/LSP.2011.2126020
15. Pang, Y., Zhang, L., Liu, Z., Yu, N., Li, H.: Neighborhood preserving projections (NPP): a novel linear dimension reduction method. In: Huang, D.-S., Zhang, X.-P., Huang, G.-B. (eds.) ICIC 2005. LNCS, vol. 3644, pp. 117–125. Springer, Heidelberg (2005). https://doi.org/10.1007/11538059_13
16. Qiao, L., Chen, S., Tan, X.: Sparsity preserving projections with applications to face recognition. Pattern Recogn. **43**(1), 331–341 (2010)
17. Roweis, S., Saul, L., Hinton, G.E.: Global coordination of local linear models. In: Advances in Neural Information Processing Systems 14 (2001)
18. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290**(5500), 2323–2326 (2000)
19. Saul, L.K., Roweis, S.T.: An introduction to locally linear embedding. unpublished. http://www.cs.toronto.edu/~roweis/lle/publications.html (2000)
20. Tenenbaum, J.B., Silva, V.d., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000)
21. Van Der Maaten, L., Postma, E., Van den Herik, J., et al.: Dimensionality reduction: a comparative. J. Mach. Learn. Res. **10**(66–71), 13 (2009)
22. Weinberger, K.Q., Saul, L.K.: An introduction to nonlinear dimensionality reduction by maximum variance unfolding. In: AAAI, vol. 6, pp. 1683–1686 (2006)
23. Xiang, S., Nie, F., Pan, C., Zhang, C.: Regression reformulations of lle and ltsa with locally linear transformation. IEEE Trans. Syst. Man Cybern. Part B (Cybernetics) **41**(5), 1250–1262 (2011)

24. Zhang, Z., Wang, J.: Mlle: Modified locally linear embedding using multiple weights. In: Advances in Neural Information Processing Systems 19 (2006)
25. Zhang, Z., Zha, H.: Nonlinear dimension reduction via local tangent space alignment. In: Liu, J., Cheung, Y., Yin, H. (eds.) IDEAL 2003. LNCS, vol. 2690, pp. 477–481. Springer, Heidelberg (2003). https://doi.org/10.1007/978-3-540-45080-1_66