Evaluation of Nonlinear Dimensionality Reduction Techniques for Classification of Hyperspectral Images

Evgeny Myasnikov[1]

Samara University, Moskovskoe Shosse 34, Samara, Russia, 443086
mevg@geosamara.ru
http://ssau.ru

Abstract. Nonlinear dimensionality reduction techniques are becoming increasingly popular in the analysis of hyperspectral images. In this work, some such methods are evaluated as a preliminary stage to the classification of hyperspectral images. The list of methods to be studied includes Isomap, Locally Linear Embedding, Laplacian Eigenmaps, Nonlinear Mapping, and also the linear principal component analysis technique. We study the performance of nonlinear methods with both the Euclidean distance and the Spectral angle mapper (SAM) dissimilarity measures. Analyzed methods are evaluated in terms of the classification accuracy and runtime. The experiments are carried out using the well-known hyperspectral scenes.

Keywords: Hyperspectral image, Dimensionality reduction, Isomap, Locally Linear Embedding, Laplacian Eigenmaps, Nonlinear Mapping, Principal Component Analysis

CEUR-WS.org/Vol-2268/paper16.pdf

1 Introduction

Hyperspectral images are three-dimensional arrays with two spatial dimensions and one spectral dimension. A pixel in a hyperspectral image can be considered as a vector containing a number (typically, up to a few hundred) of components corresponding to different wavelengths.

Being compared to color or gray-scale images, hyperspectral images suggest extended opportunities, for example, to detect materials in a depicted scene or substantially improve the accuracy of classification. However, the use of hyperspectral images is accompanied by increased costs for storage, transmission, and processing of such images. For this reason, an important task is to eliminate the redundancy of such images, while maintaining the quality of the solutions to applied problems.

The most widely used solution to the above task consists in the use of dimensionality reduction techniques. As such a technique, in most cases, the principal component analysis is used. Nevertheless, nonlinear dimensionality reduction techniques are becoming increasingly popular in the last years.

In this work, some such methods are evaluated as a preliminary stage to the classification of hyperspectral images. The list of methods to be studied includes Isomap [5], Locally Linear Embedding (LLE) [7], Laplacian Eigenmaps (LE) [8], Nonlinear Mapping (NLM) [3], and also the linear principal component analysis technique [1].

We study the performance of nonlinear methods with both the Euclidean distance and the Spectral angle mapper (SAM) dissimilarity measure as they have been used in hyperspectral image analysis most often. The reason for choosing the above dimensionality reduction methods was the frequency of application of these methods in the analysis of hyperspectral images and the possibility of embedding the SAM measure.

Analyzed methods are evaluated in terms of the classification accuracy and runtime. The experiments are carried out on the well-known hyperspectral scenes.

The paper has the following structure. Section 2 is devoted to the brief description of methods used in the paper. Section 3 describes the results of experiments. The paper ends up with the conclusion.

2 Methods

In this study, we use the following dimensionality reduction techniques:

- Principal Component Analysis (PCA) technique [1] is the most well-known linear dimensionality reduction technique, which is used in the wide range of applications. This method searches for a linear projection into the subspace of a smaller dimension that maximizes the variance of data.
- Nonlinear Mapping (NLM) method is based on the principle of preserving the pairwise distances between datapoints. While the basics of this method were developed in 1960-s in works by J.B. Kruskal [2] and J.W. Sammon [3], here we use a different version of the method [10], which differs to the base method in PCA-based initialization and stochastic gradient descent.
- Isomap method was introduced by J.B. Tenenbaum at al. in papers [4, 5]. The main idea of this method consists in the use of geodesic distances instead of Euclidean distances in classical metric multidimensional scaling (MDS). Here we use the Landmark Isomap method [6], which is a faster version of this algorithm.
- Locally Linear Embedding (LLE) technique was introduced by S.T. Roweis and L.K. Saul in the paper [7]. This technique is based on the idea that each particular datapoint and its neighbors lie close to a locally linear patch of the nonlinear manifold, and can be reconstructed as a linear combination of its neighbors in both high-dimensional and embedding spaces.
- Laplacian Eigenmaps technique was introduced by M. Belkin and P. Niyogi in the paper [8]. This technique is based on the eigenvalue decomposition of the graph Laplacian matrix.

In all the above nonlinear techniques (except PCA), it is assumed that the Euclidean distance is used as a dissimilarity measure. As we said in the Introduction, in this paper, we also embed the SAM measure [9] in the above nonlinear techniques.

In particular, for the Nonlinear Mapping technique, we replace the calculation of Euclidean distances in hyperspectral space with the calculation of SAM measures. According to [10], it means approximation of spectral angles by the Euclidean distances in the embedding space (SAED technique). In the ISOMAP technique we use SAM measures to construct the neighborhood graph, that is to find neighbor points, and to initialize weights of edges. In the Locally linear embedding method, we use SAM measures only to find neighbor points. In the Laplacian Eigenmaps technique, we use SAM measures both to find neighbor points, and to define the heat kernel.

3 Experiments

Datasets For the reported study, we used several well-known hyperspectral image scenes [14], which supplied with groundtruth segmentation: Salinas, Indian pines, Botswana, and Kennedy space center. In this paper we describe the experimental results for two well-known scenes, namely, Salinas and Kennedy space center (Figure 1).

Both hyperspectral image scenes were acquired using the AVIRIS sensor. The first scene contains 512×217 pixels, and 224 spectral bands. In our experiments we used the image containing 204 spectral bands, in which some spectral bands were discarded due to a high noise and water absorption. As the Salinas scene contains more than 100 thousand pixels, and it was necessary to perform a lot of runs of nonlinear dimensionality reduction techniques, for our experiments we used regularly sampled test image, which was masked with the provided groundtruth image. The classified pixels of the groundtruth image are divided into 16 classes.

Kennedy space center scene contains 512×614 pixels. The version containing 176 spectral bands was used in the experiments. The groundtruth image contains information only on a small amount of pixels, so we applied groundtruth mask, and did not use any sampling. The classified pixels of the groundtruth image are divided into 13 classes.

Experimental setup. To perform the experiments we used PCA implementation provided with Matlab, C++ implementation of Nonlinear Mapping method, and for LLE, Laplacian Eigenmaps and Isomap, we used Matlab Toolbox for Dimensionality Reduction [11].

A laptop based on Intel Core i7-6500U CPU 2.5 GHz, 12 Gb RAM was used to perform experimental studies.

Evaluation. The k-Nearest Neighbor (k-NN) classifier and the Support Vector Machine (SVM) were used in this study. To measure the quality of classification we used the overall classification accuracy, defined as the proportion of the correctly classified pixels of the test set.

Fig. 1. False grayscale images for test hyperspectral image scenes (contrasted): Salinas (left) and Kennedy space center (right).

The whole set of ground truth samples was divided into a training (60 percents) and a test (40 percents) subsets in our experiments. The dimensionality of the reduced space ranged from 3 to 30.

Experimental results. The results of the experimental study for the NN classifier are shown in Figure 2. As it can be seen from the figure, the use of the SAM measure was preferable for the Salinas and Kennedy space center hyperspectral scenes, and for almost all the considered nonlinear techniques, as SAM provided a better quality of classification compared to Euclidean distance. This observation is also confirmed for two other scenes involved in the experiments.

In all the considered cases, for the NN classifier, the best results were obtained using the Nonlinear Mapping technique. The linear PCA technique provided similar or slightly worse results than the Nonlinear Mapping combined with Euclidean distances. PCA outperformed LLE, LE, and Landmark Isomap methods on Salinas and two other test hyperspectral scenes, except the Kennedy space center scene.

In KSC scene, the linear PCA technique performed much worse than nonlinear techniques for the dimensionality of the reduced space up to 25. In this scene, we can also observe that the Nonlinear Mapping technique in combination with Euclidean distances loses advantages over other nonlinear techniques based on SAM measures for the dimensionality of the reduced space up to 20. Thus, on the one hand, we see a significant advantage of using the SAM measure over the Euclidean distance for this scene. On the other hand, we can explain obtained results by substantially nonlinear properties of this dataset. It noted earlier that the discrimination of land cover for the KSC scene is difficult due to the similarity of spectral signatures for certain vegetation types [14].

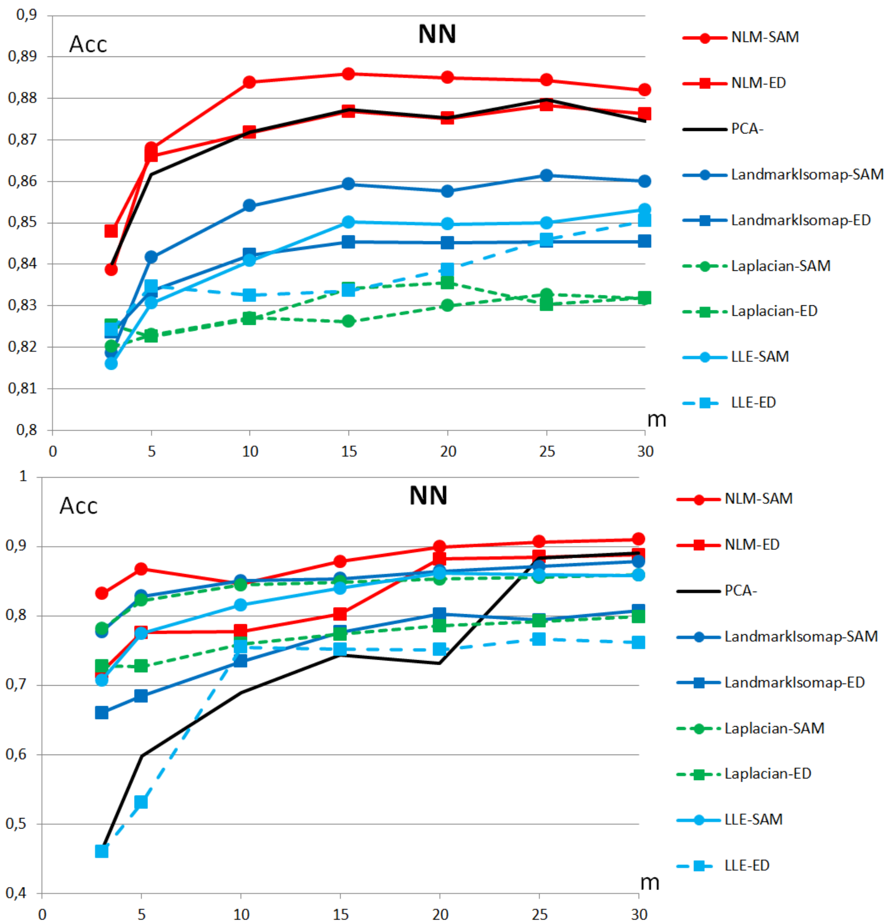

Fig. 2. Dependence of the classification accuracy *Acc* for the 1-NN classifier on the dimensionality m for the Salinas (top) and Kennedy space center (bottom) hyperspectral scenes.

The results of the experimental study for the SVM classifier are shown in Figure 3. The experimental results showed that in many considered cases (some results are not shown in the figures) the PCA was a preferable choice. The Nonlinear Mapping performed a bit worse for the SVM classifier. But again, the performance of the PCA technique was drastically reduced on the Kennedy space center scene for the dimensionality of the reduced space up to 25. This indicates the importance of the careful selection of the output dimensionality.

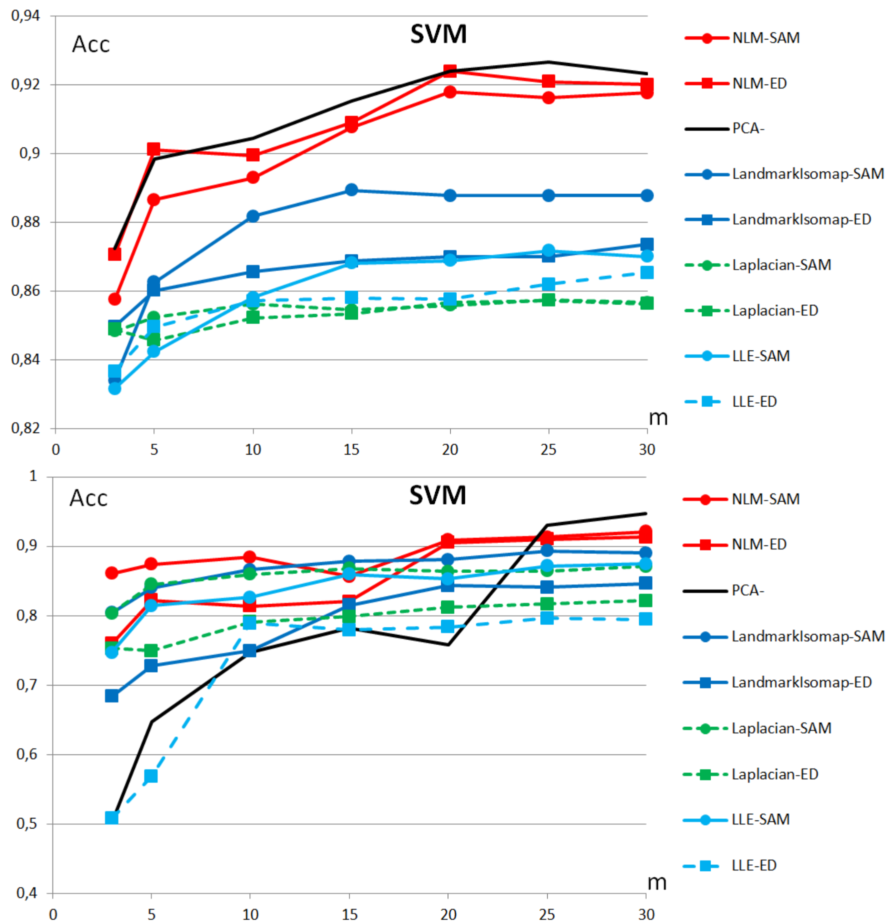

Fig. 3. Dependence of the classification accuracy *Acc* for the SVM classifier on the dimensionality m for the Salinas (top) and Kennedy space center (bottom) hyperspectral scenes.

It is worth noting that three graph-based dimensionality reduction techniques, namely Isomap, LLE, and Laplacian Eigenmaps have the mataparameter, which defines the number of neighbors. In the above experiments, we used the default value equal to 12. We found that these techniques are very sensitive to the choice of this parameter. This was especially evident for the LLE method. In some cases, the classification accuracy could be substantially (by some percent) improved over the reported above values by the good choice of the considered metaparameter. But for other cases, the same value could provide worse results.

In any case, we were not able to outperform the Nonlinear Mapping by varying this parameter in the reasonable range from 10 to 100 with the step equal to 10. The example dependency of the classification accuracy on the metaparameter k is shown in Figure 4.

The same Figure 4 shows the runtime of the considered techniques. The timing for the PCA technique is not shown as it was about 0.1 sec. that is negligible compared to the considered nonlinear methods. The run time of Landmark Isomap, LLE and Laplacian Eigenmaps is less than the run time of the Nonlinear Mapping technique for the minimum considered value of k, but it raises fast with the growth of k. So the Nonlinear Mapping becomes faster for $k > 30$. Moreover, there are approaches, which allow to speed-up this technique [12, 13]. It is worth noting, however, that timings depend hardly on the hardware and implementation.

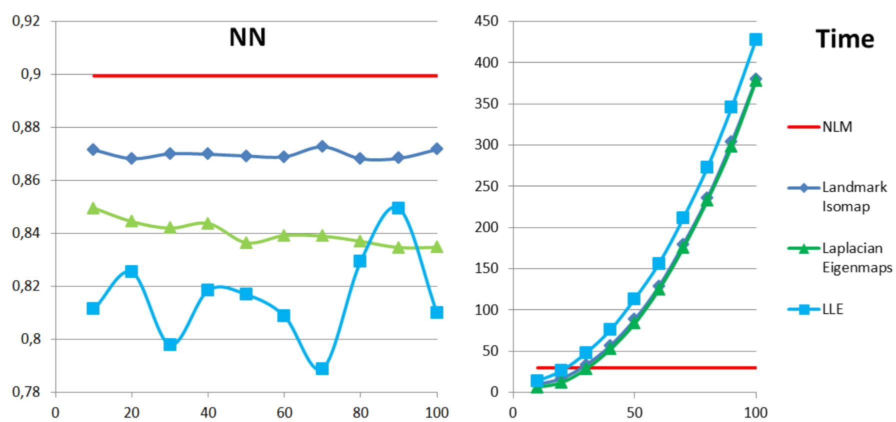

Fig. 4. Dependence of the classification accuracy *Acc* for the NN classifier on the metaparameter k (left), and the dependence of the dimensionality reduction run time on the metaparameter k (right) for the Kennedy space center hyperspectral scene.

4 Conclusion

In this paper, we studied several popular nonlinear dimensionality reduction techniques in the task of per-pixel hyperspectral image classification. We showed that the Nonlinear Mapping technique could be considered as a reasonable choice when the nearest neighbor classifier is used.

While in many cases the combination of the PCA technique with SVM classifier provides nice results, for complex hyperspectral scenes containing substantial nonlinear effects the traditional PCA technique could be a bad choice. In such cases, it is necessary to carefully choose the output dimensionality, and consider the possibility of using the nonlinear dimensionality reduction techniques.

The main drawback of the nonlinear methods is their high computational complexity, which is expressed by their long run time, which exceeds the runtime of the PCA technique by orders of magnitude.

Acknowledgments The reported study was funded by RFBR according to the research project no. 18-07-01312-a.

References

1. K. Fukunaga: Introduction to Statistical Pattern Recognition. London: Academic Press, 2nd edn. (2003)
2. Kruskal J. B.: Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. Psychometrika, 29, P. 1-27. (1964)
3. Sammon, J.W.: A nonlinear mapping for data structure analysis. Transactions on Computers. Vol. 18. P. 401-409. (1969)
4. J. Tenenbaum: in Advances in Neural Information Processing, 10. M. Jordan, M. Kearns, S. Solla (Eds.). MIT Press, Cambridge, MA. P. 682-688. (1998)
5. J. B. Tenenbaum, V. de Silva, J. C. Langford: A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science, P. 2319-2323. (2000)
6. J. B. Tenenbaum, V. de Silva: Global versus local methods in nonlinear dimensionality reduction. Advances in Neural Information Processing Systems, Vol. 15, P. 705-712. (2003)
7. S. T. Roweis and L. K. Saul: Nonlinear Dimensionality Reduction by Locally Linear Embedding. Science, Vol. 290. P. 2323-2326. (2000)
8. M.Belkin, P.Niyogi: Laplacian eigenmaps and spectral techniques for embedding and clustering. Advances in Neural Information Processing Systems, 14. P. 585-591.(2001)
9. Kruse, F.A., Boardman, J.W., Lefkoff, A.B., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., and Goetz, A.F.H.: The Spectral Image Processing System (SIPS) Interactive Visualization and Analysis of Imaging Spectrometer Data. Remote Sensing of Environment, 44. P. 145-163. (1993)
10. Myasnikov, E.: Nonlinear Mapping Based on Spectral Angle Preserving Principle for Hyperspectral Image Analysis. In: Felsberg M., Heyden A., Krger N. (eds) Computer Analysis of Images and Patterns. LNCS 10425, pp. 416-427. (2017)
11. L.J.P. van der Maaten, E.O. Postma, and H.J. van den Herik. Dimensionality Reduction: A Comparative Review. Tilburg University Technical Report, TiCC-TR 2009-005. (2009)
12. Myasnikov, E.: The use of Interpolation Methods for Nonlinear Mapping. In: Chmielewski L., Datta A., Kozera R., and Wojciechowski K. (eds) Computer Vision and Graphics. LNCS 9972, pp. 649-655. (2016)
13. Myasnikov, E.V.: Fast Techniques for Nonlinear Mapping of Hyperspectral Data. Proc. SPIE 10341, Ninth International Conference on Machine Vision (ICMV 2016), 103411D. (2017)
14. Hyperspectral Remote Sensing Scenes. Available from: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes