

# Multiple Manifold Learning by Nonlinear Dimensionality Reduction

Juliana Valencia-Aguirre<sup>1</sup>, Andrés Álvarez-Meza<sup>1</sup>, Genaro Daza-Santacoloma<sup>2</sup>,  
Carlos Acosta-Medina<sup>1,3</sup>, and César Germán Castellanos-Domínguez<sup>1</sup>

<sup>1</sup> Signal Processing and Recognition Group, Universidad Nacional de Colombia,  
Manizales, Colombia

<sup>2</sup> Faculty of Electronic Engineering, Universidad Antonio Nariño, Bogotá, Colombia

<sup>3</sup> Scientific Computing and Mathematical Modeling Group, Universidad Nacional de  
Colombia, Manizales, Colombia

{jvalenciaag, amalvarezme, gdazas, cdacostam, cgcastellanosd}@unal.edu.co

**Abstract.** Methods for nonlinear dimensionality reduction have been widely used for different purposes, but they are constrained to single manifold datasets. Considering that in real world applications, like video and image analysis, datasets with multiple manifolds are common, we propose a framework to find a low-dimensional embedding for data lying on multiple manifolds. Our approach is inspired on the manifold learning algorithm Laplacian Eigenmaps - LEM, computing the relationships among samples of different datasets based on an intra manifold comparison to unfold properly the data underlying structure. According to the results, our approach shows meaningful embeddings that outperform the results obtained by the conventional LEM algorithm and a previous close related work that analyzes multiple manifolds.

**Keywords:** Manifold learning, multiple manifolds, laplacian eigenmaps, video analysis.

## 1 Introduction

Often, in machine learning and pattern recognition literature, the nonlinear dimensionality reduction (NLDR) techniques are reviewed as learning methods for discovering an underlying low-dimensional structure from a set of high-dimensional input samples, that is, NLDR techniques unfold a non-linear manifold embedded within a higher-dimensional space. Nevertheless, most of the NLDR algorithms are constrained to deal with a single manifold, attaining unappropriate low-dimensional representations when input data lie on multiple manifolds, because the inter-manifold distance is usually much larger than the intra-manifold distance [1], moving apart each manifold from the others, regardless of whether the behavior among them is similar.

To our best knowledge, some few works [2,3,1] have proposed the application of the NLDR techniques to the analysis of multiple manifold datasets. Particularly, in [1] a framework to learn an embedded manifold representation from multiple

data sets called *Learning a Joint Manifold* (LJM) is presented, which finds a common manifold among the different data sets, without assuming some kind of correspondence between the different manifolds. However, the main drawback of this approach is that for obtaining suitable low dimensional representations, the input samples must be similar in appearance. When the multiple manifolds do not have a close resemblance among them, the LJM method fails to embed the data. On the other hand, the approach presented in [3] actually requires the use of correspondence labels among the samples in order align the data sets, in such case the complexity of the challenge is lower than when no one correspondence is assumed. A similar solution is proposed in [4].

Unlike these mentioned works for dealing with multiple manifolds, our work makes possible to analyze dissimilar objects/subjects in appearance but with a common behavior (similar motion), moreover our methodology allows to employ objects/subjects with different input dimensions and number of samples among manifolds. These features of our work are the major contribution to the state of the art. Our approach is inspired on the manifold learning algorithm Laplacian Eigenmaps - LEM [5], because its optimization problem has an analytic solution avoiding local minima, and few free parameters need to be fixed by user. Our approach can be employed to visually identify in a low-dimensional space the dynamics of a given activity, learning it from a variety of datasets. We test the method on two real-world databases, changing the number of samples and input dimensions per manifold. Our proposal is compared against both the conventional Laplacian Eigenmaps (LEM) [2] and the closest work found in the state of the art for multiple-manifold learning (LJM) [1]. Overall, our methodology achieves meaningful low dimensional representations, visually outperforming the results obtained by the other methods. This work is organized as follows. In Section 2, a brief description about LEM algorithm is presented. Section 3 introduces the proposed methodology for multiple manifold dimensionality reduction. In Section 4 the experimental results are described and discussed. Finally, in Section 5, we conclude about the obtained results.

## 2 Laplacian Eigenmaps – LEM

Laplacian Eigenmaps (LEM) is a NLDR technique based on preserving the intrinsic geometric structure of a manifold. Let  $\mathbf{X} \in \mathbb{R}^{n \times p}$  the input data matrix with row vectors  $\mathbf{x}_i$  ( $i = 1, \dots, n$ ). The LEM transformation finds a low-dimensional Euclidean space  $\mathbf{Y} \in \mathbb{R}^{n \times m}$ , with row vectors  $\mathbf{y}_i$  ( $m \ll p$ ). This algorithm has three main steps. First, an undirected weighted graph  $G$  with  $n$  nodes (one for each  $\mathbf{x}_i$ ) is built. Nodes  $i$  and  $j$  are connected by an edge  $E_{ij} = 1$ , if  $i$  is one of the  $k$  nearest neighbors of  $j$  (or viceversa) according to the Euclidean distance [2]. In the second step, a weight matrix  $\mathbf{W} \in \mathbb{R}^{n \times n}$  is calculated. For this purpose two alternatives variants can be considered: heat kernel or simple minded. In the heat kernel variant, if nodes  $i$  and  $j$  are connected, then  $W_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$ , being  $\kappa(\cdot, \cdot)$  a kernel function, otherwise,  $W_{ij} = 0$ . For the simple minded option,  $W_{ij} = 1$  if vertices  $i$  and  $j$  are connected by an edge, otherwise,  $W_{ij} = 0$ . Then, the  $\mathbf{L} \in \mathbb{R}^{n \times n}$  Laplacian graph is given by  $\mathbf{L} = \mathbf{D} - \mathbf{W}$ ,

where  $\mathbf{D} \in \mathbb{R}^{n \times n}$  is a diagonal matrix with elements  $D_{ii} = \sum_j W_{ji}$ . In the third step, the following objective function is minimized

$$\sum_{ij} (\mathbf{y}_i - \mathbf{y}_j)^2 W_{ij}, \tag{1}$$

which implies a penalty if neighboring points  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are mapped far apart. Finally, the LEM problem can be accomplished solving the generalized eigenvalue problem  $\mathbf{L}\mathbf{Y}_{:,l} = \lambda_l \mathbf{D}\mathbf{Y}_{:,l}$ ; where  $\lambda_l$  is the eigenvalue corresponding to the  $\mathbf{Y}_{:,l}$  eigenvector, with  $l = 1, \dots, n$ . First eigenvector is the unit vector with all equal components, while the remaining  $m$  eigenvectors form the embedded space.

### 3 Multiple Manifold Learning – MML

The NLDR techniques based on manifold learning fail when they look for a common low-dimensional representation for data lying on multiple manifolds. In this sense, we propose relate each input sample  $\mathbf{x}_i$  with  $C$  different manifolds that share a similar underlying structure. Let  $\Psi = \{\mathbf{X}^c\}_{c=1}^C$  an input manifold set, where  $\mathbf{X}^c \in \mathbb{R}^{n_c \times p_c}$ . Our goal is to find a mapping from  $\Psi$  to a low-dimensional space  $\mathbf{Y} \in \mathbb{R}^{n \times m}$  (with  $m \ll p_c$ , and  $n = \sum_{c=1}^C n_c$ ), which reveals both the intra manifold structure (relationships within manifold), and the inter manifold structure (relationships among manifolds). Consequently, a weight matrix  $\mathbf{A}$ , that takes into account both structures, can be computed as

$$\mathbf{A} = \begin{bmatrix} \mathbf{W}^1 & \mathbf{M}^{12} & \dots & \mathbf{M}^{1c} & \dots & \mathbf{M}^{1C} \\ \mathbf{M}^{21} & \mathbf{W}^2 & \dots & \mathbf{M}^{2c} & \dots & \mathbf{M}^{2C} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{M}^{c1} & \mathbf{M}^{c2} & \dots & \mathbf{W}^c & \dots & \mathbf{M}^{cC} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{M}^{C1} & \mathbf{M}^{C2} & \dots & \mathbf{M}^{Cc} & \dots & \mathbf{W}^C \end{bmatrix}, \tag{2}$$

where each  $\mathbf{W}^c \in \mathbb{R}^{n_c \times n_c}$  is the traditional LEM intra manifold weight matrix for each  $\mathbf{X}^c$  [1]. Furthermore, each  $\mathbf{M}^{cb} \in \mathbb{R}^{n_c \times n_b}$  ( $b = 1, \dots, C$ ) block is a soft correspondence matrix between  $\mathbf{X}^c$  and  $\mathbf{X}^b$ .

In [1] a methodology called Learning a Joint Manifold Representation (LJM) is proposed to unfold the data underlying structure from multiple manifolds, which calculates the matrix  $\mathbf{A}$  (equation (2)), computing the intra manifold structure matrices  $\mathbf{W}^c$  as in traditional LEM, and the inter manifold structure matrices  $\mathbf{M}^{cb}$  by solving a permutation matrix  $\mathbf{P}$ , which allows to find a maximum weight matching by permuting the rows of  $\mathbf{U}^{cb} \in \mathbb{R}^{n_c \times n_b}$ ,  $U_{qr}^{cb} = \kappa(\mathbf{x}_q, \mathbf{x}_r)$ ,  $\mathbf{x}_q \in \mathbf{X}^c$ , and  $\mathbf{x}_r \in \mathbf{X}^b$  ( $q = 1, \dots, n_c$ ;  $r = 1, \dots, n_b$ ). Nonetheless, LJM is quite sensitive to feature variability between samples of different manifolds, due to  $\mathbf{U}^{cb}$  is inferred in the high-dimensional space. Moreover, LJM is limited to analyze input matrices  $\mathbf{X}^c$  which belong to the same input dimension ( $p_1 = p_2 = \dots = p_c = \dots = p_C$ ), as can be seen in the calculation of each  $\mathbf{U}^{cb}$ .

In this work, we propose to identify the correspondence among data points from different manifolds without making a high-dimensional sample comparison. In other words, the similarities among observations of different manifolds are not directly calculated in each pair  $\mathbf{X}^c$  and  $\mathbf{X}^b$ . Therefore, we compute each soft correspondence matrix  $\mathbf{M}^{cb}$  in (2) as

$$M_{qr}^{cb} = \frac{\langle \mathbf{w}_q^c, \mathbf{w}_r^b \rangle}{\|\mathbf{w}_q^c\| \|\mathbf{w}_r^b\|}, \quad (3)$$

where  $\mathbf{w}_q^c \in \mathbb{R}^{1 \times n_c}$  and  $\mathbf{w}_r^b \in \mathbb{R}^{1 \times n_b}$  are row vectors of  $\mathbf{W}^c$  and  $\mathbf{W}^b$ , respectively. It is important to note that equation (3) is not well defined when  $n_c \neq n_b$ , thereby, we use a conventional interpolation method based on cubic splines for oversampling the lowest size vector to properly compute the inner product between  $\mathbf{w}_q^c$  and  $\mathbf{w}_r^b$ . Our approach for Multiple Manifold Learning (MML) aims to calculate the relationships among samples of different manifolds, comparing the intra manifold similarities contained in each  $\mathbf{W}^c$  (equation (3)). Finally, given the weight matrix  $\mathbf{A}$ , we minimize the following objective function

$$\sum_{ij} (\mathbf{y}_i - \mathbf{y}_j)^2 A_{ij}. \quad (4)$$

Solving equation (4) as in traditional LEM algorithm allows us to find a low-dimensional space  $\mathbf{Y}$  for data lying on multiple manifolds.

## 4 Experimental Results

We tested the conventional LEM algorithm [2], the LJM technique [1], and our proposed methodology MML on two real-world databases, in order to find a 2D low-dimensional representation ( $m = 2$ ) for data lying on multiple manifolds. The first database, the Columbia Object Image Library (COIL-100) [6], contains 72 RGB-color images, for each one of the 100 objects, in PNG format, which were taken at pose intervals of 5 degrees while the object is rotated 360 degrees. In this work, the following objects are used: Car, Frog and Duck. The image size is  $128 \times 128$ , which are transformed to gray scale. The second database is the CMU motion of body (Mobo) [7], which holds 25 individuals walking a treadmill. All subjects are captured using six high resolution color cameras distributed evenly around the treadmill. For concrete testing, we used the silhouette sequences of one gait cycle for slow walk of three persons, which are captured from a side view. The images are resized to  $80 \times 61$ . The Figure 1 shows some images samples of COIL-100 and Mobo databases.

Three types of experiments are performed. Firstly, we use the selected objects of COIL-100 with a same amount of observations per set ( $n_1 = n_2 = n_3 = 72$ ), and equal input dimensions ( $p_1 = p_2 = p_3 = 16384$ ). In this case, the number of nearest neighbors is fixed as  $k = 3$ . For the second experiment, we use the Mobo database, which leads input samples per manifold of different sizes:  $n_1 = 36$ ,  $n_2 = 40$ ,  $n_3 = 38$  and  $p_1 = p_2 = p_3 = 4880$ . The number of neighbors is set to  $k = 2$ . In order to test the algorithms on a dataset that contains

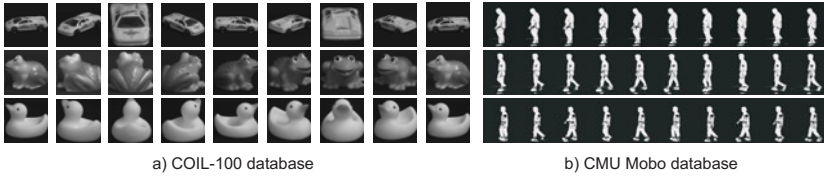


Fig. 1. Databases examples

multiple manifolds with high-variability in the input sample sizes, we use the COIL-100 but performing an uniform sampling of the observations, obtaining input spaces with  $n_1 = 72$ ,  $n_2 = 36$ ,  $n_3 = 18$  and  $p_1 = p_2 = p_3 = 16384$ . Here, the number of nearest neighbors are fixed as  $k_1 = 4$ ,  $k_2 = 2$ ,  $k_3 = 1$ . Finally, the third experiment aims to validate the proposed methodology for analyzing datasets with different amount of observations and input dimensions (image resolution). For this purpose, we employ the COIL-100 performing an uniform sampling of the observations, and resizing the images. Thence, the obtained input spaces have the following characteristics:  $n_1 = 72$ ,  $n_2 = 36$ ,  $n_3 = 18$  and  $p_1 = 16384$ ,  $p_2 = 8100$ ,  $p_3 = 2784$ .

According to the results presented in Figures 2(a), 3(a) and 4(a), traditional LEM is not able to find the correspondence among different datasets which are related to a common underlying data structure. For all the provided

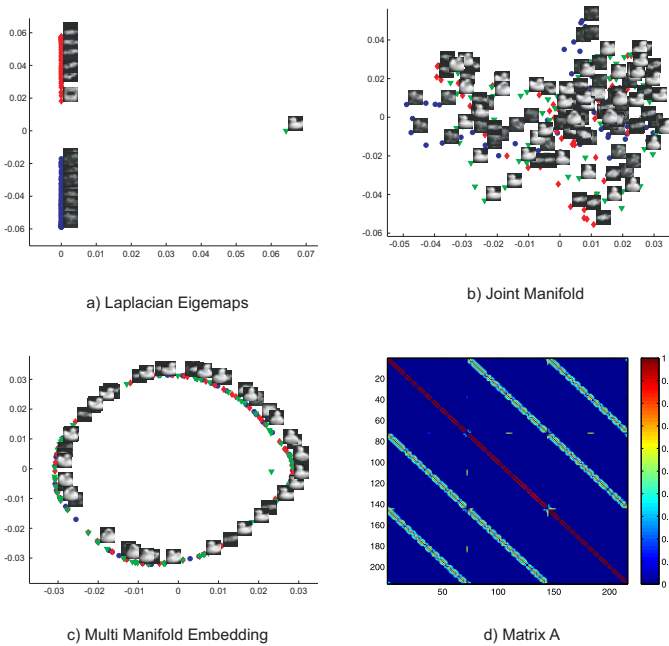


Fig. 2. Three objects, equal amount of observations

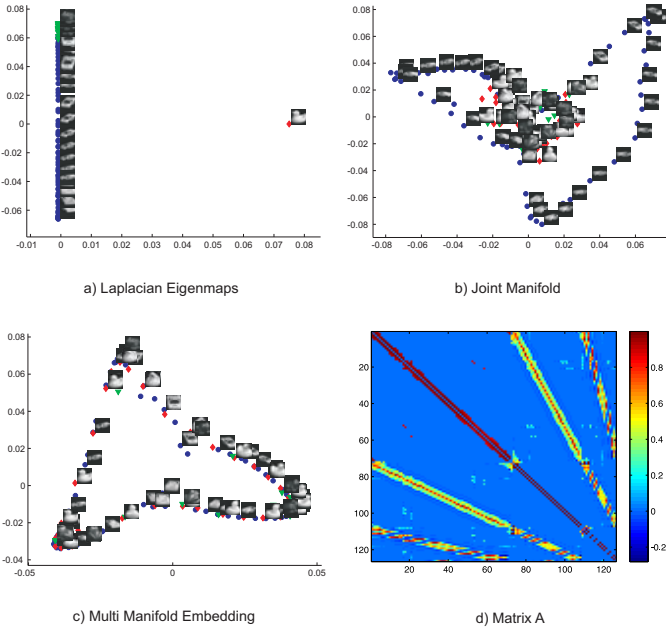


Fig. 3. Three objects, different amount of observations

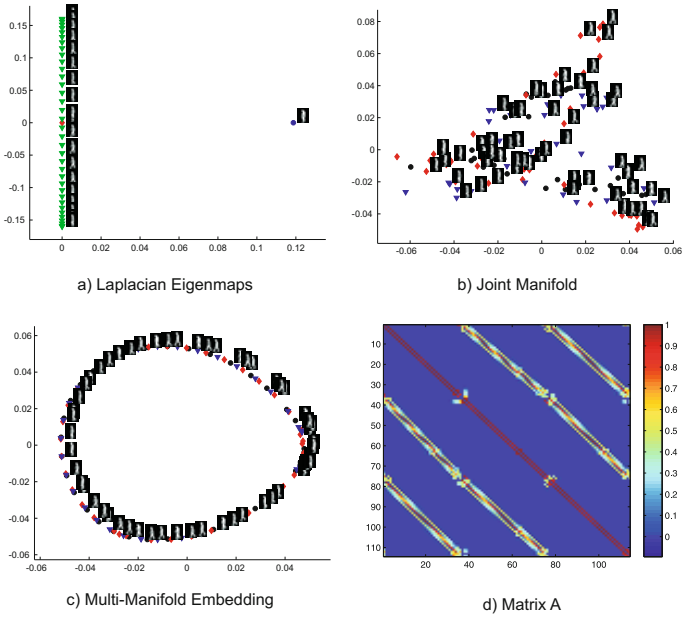
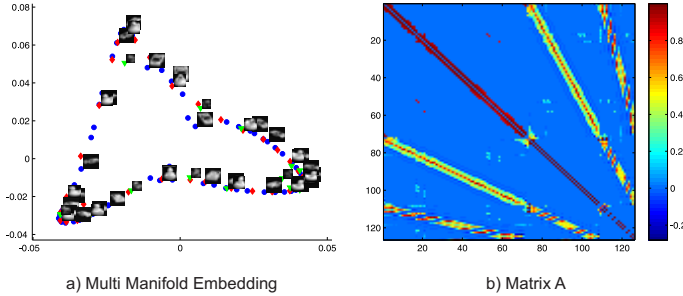


Fig. 4. Three Gait subjects, different amount of observations



**Fig. 5.** Different amount of observations and input dimensions (image resolution)

experiments, LEM performs a clustering of points for each manifold. That algorithm can not find a low-dimensional representation that unfolds the underlying data structure from multiple manifolds, due to the weight matrix  $\mathbf{W}$  in LEM is computed only considering pixel intensity similarities among frames. Again, taking into the account the attained results with the LJM technique (Figures 2(b), 3(b), and 4(b)), it can be seen how it attempts to find a correspondence among datasets but losing the intrinsic geometry data structure of the phenomenon (object motion). More precisely, for the COIL-100 database, the dynamic of the rotation is not reflected in the embedded space. Similar results are obtained for gait analysis in the Mobo database, although LJM tries to reveal the elliptical motion shape, it is not able to conserve a soft correspondence among samples. Note that the application of LJM technique is limited to analyze frames of video sharing a similar geometry, due to  $\mathbf{U}^{cb}$  is inferred in the high-dimensional space (pixels frame comparison). Overall, the LJM method can not properly learns the relationships among objects performing the same activity, it just develops well when the analyzed manifolds are similar in appearance.

Finally, the results obtained with the proposed methodology MML, demonstrate that the computed low-dimensional space exhibits the appropriated dynamic of a given activity, learning it from multiple datasets. Figures 2(c), 3(c), 4(c), and 5(a) show how this method learns the relationships among frames of videos related to a similar activity, unfolding the underlying data structure. The low-dimensional representations found by MML reflects the activity dynamics and the soft correspondence among points of different datasets. Furthermore, our approach identifies a soft correspondence among videos even when they do not share a common similarity appearance, number of observations, and/or resolution. This can be explained because the relationships among samples of different datasets are computed based on an intra manifold comparison (equation (3)) the samples are not directly compared on the high dimensional input space, instead of that, the samples are compared by means of their own similarity representations, which is the similarity between a sample and each one of the other samples on the same manifold. The Figures 2(d), 3(d), 4(d), and 5(b) confirm it.

## 5 Conclusion

In this paper a new NLDR methodology for finding a common low-dimensional representation from multiple datasets is presented. We proposed to calculate an unique embedding space in order to visually identify the dynamics of a given activity performed by a variety of objects/subjects. In other words, different manifolds that share a similar underlying structure are mapped to the same low-dimensional space. Our methodology is inspired on the manifold learning algorithm LEM, computing the relationships among samples of different datasets based on an intra manifold comparison to properly unfold the data underlying structure. According to the obtained results, our approach outperformed the original LEM method, and a previous similar work called LJM [1] that analyzes multiple manifolds. The main advantage of this proposed methodology is the possibility for analyzing dissimilar objects/subjects in appearance but with a common behavior (similar motion). Moreover our methodology allows to employ objects/subjects with different input dimensions and number of samples among manifolds. As future work, we are interested in apply our methodology to support human motion classification and identification of impairments, as well as for computer animation.

**Acknowledgments.** Research funded by ARTICA, the project 20201006570 and a M.sc. scholarship by UNAL, and the project 20201006594 funded by U. de Caldas and UNAL. GDS was supported by the project 20110108 - PI/UAN-2011-505gb UAN.

## References

1. Torki, M., Elgammal, A., Lee, C.S.: Learning a joint manifold representation from multiple data sets. In: ICPR, pp. 1068–1071 (2010)
2. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation* 15(6) (2003)
3. Liu, X., Lu, H., Li, W.: Multi-manifold modeling for head pose estimation. In: 17th IEEE International Conference on Image Processing, ICIP, pp. 3277–3280 (2010)
4. Ham, J.H., Lee, D.D., Saul, L.K.: Learning high dimensional correspondences from low dimensional manifolds. In: Workshop on The Continuum from Labeled to Unlabeled Data in Machine Learning and Data Mining, pp. 34–41 (2003)
5. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. In: NIPS, pp. 585–591 (2001)
6. Nene, S.A., Nayar, S.K., Murase, H.: Columbia object image library: Coil-100. Technical report, Department of Computer Science, Columbia University (1996)
7. Gross, R., Shi, J.: The cmu motion of body (mobo) database, tech. report cmu-ri-tr-01-18. Technical report, Robotics Institute, Carnegie Mellon University (2001)