



Journal of Applied and Computational Mechanics



Research Paper

Manifold Learning Algorithms Applied to Structural Damage Classification

Jersson X. Leon-Medina^{1,2}, Maribel Anaya³, Diego A. Tibaduiza⁴, Francesc Pozo¹

¹ Control, Modeling, Identification and Applications (CoDALab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya (UPC), Campus Diagonal-Besòs (CDB), Eduard Maristany, 16, Barcelona, 08019, Spain, Email: jersson.xavier.leon@upc.edu, francesc.pozo@upc.edu

² Departamento de Ingeniería Mecánica y Mecatrónica, Universidad Nacional de Colombia, Cra 45 No. 26-85, Bogotá, 111321, Colombia, Email: jxleonm@unal.edu.co

³ MEM (Modelling-Electronics and Monitoring Research Group), Faculty of Electronics Engineering, Universidad Santo Tomás, Bogotá 110231, Colombia, Email: maribelanaya@usantotomas.edu.co

⁴ Departamento de Ingeniería Eléctrica y Electrónica, Universidad Nacional de Colombia, Cra 45 No. 26-85, Bogotá, 111321, Colombia, Email: dtibaduizab@unal.edu.co

Received March 26 2020; Revised August 21 2020; Accepted for publication August 27 2020.

Corresponding author: Francesc Pozo (francesc.pozo@upc.edu)

© 2021 Published by Shahid Chamran University of Ahvaz

Abstract. A comparative study of four manifold learning algorithms was carried out to perform the dimensionality reduction process within a proposed methodology for damage classification in structural health monitoring (SHM). Isomap, locally linear embedding (LLE), stochastic proximity embedding (SPE), and laplacian eigenmaps were used as manifold learning algorithms. The methodology included several stages that comprised: data normalization, dimensionality reduction, classification through K-Nearest Neighbors (KNN) machine learning model and finally holdout cross-validation with 25% of data for training and the remaining 75% of data for testing. Results evaluated in an experimental setup showed that the best classification accuracy was 100% when the methodology uses isomap algorithm with a hyperparameter k of 170 and 8 dimensions as a feature vector at the input to the KNN classification machine.

Keywords: Structural health monitoring, Manifold learning, Feature extraction, Machine learning, Dimensionality reduction, Damage classification.

1. Introduction

The use of data-driven algorithms for structural health monitoring (SHM) allows continuous monitoring and online damage identification in structures subjected to changes in its operational and environmental conditions during its lifetime. The principal goal of SHM is oriented to the development of efficient methodologies to process the data obtained directly from the structures under inspection and provide results associated with the different levels of the damage identification process. Some advantages in the use of data from the structure are that data are obtained from a sensor network, which is permanently installed to the structure, allowing to know the state of the structure at any time; it is possible to identify different kinds of damages; and allow the possibility of enhancing the information about the detected damage such as size, position, among others. However, one of the main challenges in SHM with data-driven algorithms is the development of algorithms or methodologies with high accuracy that allows for avoiding false damage identification [1]. Damage identification process includes different levels; one of these is the damage classification where pattern recognition strategies are often applied.

To improve classification within the pattern recognition process, the feature extraction stage needs to be explored, as it compresses the original data and reduces the dimension thereof, extracting the main characteristics in a new feature space that can be used as input to the classification machine. In most cases, the nature of the acquired data during experimental tests contains non-linear features due to the kind of material to inspect, the inspection methods, among others. In this sense, linear techniques such as principal component analysis (PCA) [2], which have been used to perform the reduction of dimensionality, causes loss of relevant information and the possibility of detecting important patterns by data-driven algorithms [3]. The data obtained from a sensor network is governed by an underlying manifold and it reveals their dynamic nature [4]. Manifold learning algorithms are used in order to process the high dimensionality of data obtained by the sensor network. As this mapping is trained, a low-dimensional joint manifold of the data is implicitly learned [5]. Different manifold learning methods have been created to handle non-linearity dimensionality reduction, such as locally linear embedding (LLE), isomap, laplacian eigenmaps, kernel PCA among others. These methods have been used for image and audio classification tasks [6]. Within SHM, Principal Component Analysis-PCA has turned out to be a traditional technique to analyze data [7]. Regarding manifold learning applied in SHM, in 2015 Dervilis *et al.* [8] conducted a damage identification study that used LLE to reduce the dimensionality of data obtained in an SHM study of a bridge. In that work, the authors propose to isolate the effect of temperature changes from the



healthy structure and the structure with damage, using gaussian process regression. Yildiz et al. [9], compared PCA, isomap, landmark isomap (L-Isomap), laplacian eigenmaps, fast maximum variance unfolding (FastMVU), stochastic neighbor embedding (SNE) and t-distributed stochastic neighbor embedding (t-SNE) algorithms for dimensionality reduction of real-world datasets. The obtained results indicate that the most efficient algorithms among the dimensional reduction algorithms were laplacian eigenmaps, FastMVU and t-SNE algorithms.

In this work we follow the research line started some years ago in the field of pattern recognition, digital signal processing and applications of machine learning algorithms into SHM systems. Some of the algorithms that the authors have used for multivariate data analysis and signal processing are: PCA [6], self-organizing maps (SOM) [10], K- nearest neighbors (KNN)[2], artificial immune system (AIS) [11], fuzzy c-means clustering, support vector machines (SVM) [12], classification and regression trees (CART) [12], and hierarchical nonlinear principal component analysis (h-NLPCA) [12]. The subspace learning and manifold learning algorithms have been used satisfactory to solve different kind of classification problems in electronic tongues [13,14], electronic noses [15,16], damage classification in SHM [17].

Tests performed with piezoelectric sensors for SHM applications yield high dimensional data. In contrast, the manifold learning techniques find low-dimensional embedding structures within high-dimensional data. As a contribution, this work presents a damage detection and classification methodology which uses multivariate analysis and machine learning algorithms from a pattern recognition point of view. Isomap, laplacian eigenmaps, LLE and Stochastic Proximity Embedding (SPE) methods are used as dimensionality reduction tools to extract the non-linear features coming from the piezoelectric sensors in SHM. The four manifold learning techniques were compared because Isomap and SPE are global methods, and laplacian eigenmaps and LLE are local methods. This work also evaluates the influence of the number of dimensions used as the input of the classification machine. Validation of the classification is performed by the hold-out cross-validation method, splitting the data into two sets: a training set of 25% and a test set of 75%. Likewise, the model is evaluated with different data to avoid overfitting. Well levels of accuracy have been achieved in the classification carried out by the KNN method, and performance measures are calculated to indicate in a single number such classification goodness. The methodology is tested with data from an aluminum plate with a piezoelectric sensor network and results show that it is possible to demonstrate its usefulness for the damage detection and classification process in structures inspected with sensors.

This paper is organized as follows. Firstly, the theoretical background is described. Secondly, the developed methodology of damage classification is described. Afterward, the results and discussion section are defined from the validation using a plate with 3 different damages. Finally, the conclusions section highlights the main outcomes of this work.

2. Theoretical Background

A large amount of data coming from the structural monitoring system constitutes in some cases non-linear manifolds in the high-dimensional observation space. Some pattern recognition methods that have used linear techniques to process the data in SHM ignore the subtleties of the manifolds that describe raw data, such as concavities and protuberances [18], which generates inconveniences to achieve a highly accurate recognition.

In 2002, de Ridder and Duin [19] concluded that the local method of manifold learning LLE is only useful for few dimensions and that the classifiers have better performance when they use main components obtained by Linear PCA. The reason can be that in reality the data have noise, enough intrinsic characteristics and high curvature in the observation space and the embedded space, whereas current manifold learning methods depend on the selection of parameters [18]. Embedding methods of manifold learning are divided into global and local embedding algorithms. On the one hand, Isomap [20], as a global algorithm, assumes that isometric properties should be preserved in both the observation space and the intrinsic embedding space in the affine sense. On the other hand, laplacian eigenmaps [21] make emphasis on the preservation of local neighbor structure.

Dimensionality reduction techniques can be used for data compression and retain as much information as possible from the whole feature space [22]. In some cases, the resulting characteristics may represent the maximization of variation within the available data. This maximization can become a problem when observations representing the normal condition are inconsistent, since the corresponding data groups are scattered through the feature space. The dimensionality reduction should be used not only for visualization or as preprocessing in very high dimensional data, but also as a general pre-processing technique in numerical data to increase the classification performance [23].

2.1 Laplacian Eigenmaps

Laplacian eigenmaps is a manifold learning algorithm that preserves local properties of the manifold to construct a low-dimensional data [21]. For this purpose, the local properties are based on the pairwise distances between near neighbors [3]. Laplacian eigenmaps creates a neighborhood graph G in which every data point x_i and its k nearest neighbors are connected and these distances are minimized. The distances have weights that contribute differently to the cost function. The distance in the low-dimensional space between a data point and its first nearest neighbor contributes more to the cost function than the distance between the data point and its second nearest neighbor. The minimization of the cost function is defined as an eigenproblem. For all points x_i and x_j in graph G that are connected by an edge, the weight of the edge is computed using the next Gaussian kernel function eq. (1), leading to a sparse adjacency matrix W .

$$w_{ij} = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \tag{1}$$

In the computation of the low-dimensional representations y_i , the cost function minimized is defined as:

$$\phi(Y) = \sum_{ij} (y_i - y_j)^2 w_{ij} \tag{2}$$

In the cost function, large weights w_{ij} correspond to small distances between the data points x_i and x_j , hence the difference between their low-dimensional representations y_i and y_j highly contributes to the cost function. As a result, nearby points in the high-dimensional space are brought closer in the low-dimensional representation. The computation of the degree matrix M and the laplacian graph L of the graph W allows formulating the minimization problem as an eigenproblem [24]. The degree matrix M of W is a diagonal matrix, whose entries are the row sums of W ($m_{ii} = \sum_j w_{ij}$). The laplacian graph L is computed by $L = M - W$. Considering eq. (2), the cost function can also be computed as:



$$\phi(Y) = \sum_i y_i^2 m_{ii} + \sum_j y_j^2 m_{jj} - 2 \sum_{ij} y_i y_j w_{ij} = 2Y^T M Y - 2Y^T W Y = 2Y^T L Y \quad (3)$$

Afterward, minimizing $\phi(Y)$ is proportional to minimizing $Y^T L Y$. The low-dimensional data representation Y can thus be found by solving the generalized eigenvector problem [2]:

$$L v = \lambda M v \quad (4)$$

for the d smallest nonzero eigenvalues. The d eigenvectors v_i form the low-dimensional data representation Y .

2.2 Isomap

Isomap [20] preserves pairwise geodesic distances between datapoints x_i by constructing a neighborhood graph G , in which every data point x_i is connected with its k nearest neighbors in the dataset X . The shortest path between two points in the graph forms a good estimate of the geodesic distance between these two points, and can be computed using Dijkstra's or Floyd's shortest-path algorithm [3]. The geodesic distances between all data points in X are computed, thereby forming a pairwise geodesic distance matrix. The low-dimensional representations y_i of the data points x_i in the low-dimensional space Y are calculated by applying multidimensional scaling (MDS) [25] in the resulting matrix. Isomap follows the general multidimensional scaling philosophy by attempting to preserve the global geometric properties of the underlying nonlinear manifold [26]. MDS represents a collection of non-linear dimensionality reduction techniques that retain the pairwise distances between data points. The quality of the mapping is expressed in the stress function, a measure of the error between the pairwise distances in the low and high dimension representation of the data.

2.3 Locally Linear Embedding

LLE [27] is a local technique for dimensionality reduction that is similar to Isomap in that it also constructs a graphical representation of data points. Unlike Isomap, it tries to preserve only the local properties of the data, making LLE less sensitive to short circuit than Isomap [3]. LLE describes the local properties of the manifold around a data point x_i by writing the data point as a linear combination W_i of its closest neighbors x_{ij} . Therefore, LLE adjusts to a hyperplane through the x_i data point and its closest neighbors, assuming that the manifold is locally linear. The local linearity assumption implies that the reconstruction weights W_i of the data points x_i are invariant for translation, rotation and rescaling. Due to the invariance of these transformations, any linear mapping of the hyperplane to a space of smaller dimensionality preserves the reconstruction weights in the space of smaller dimensionality. In consequence, finding the representation of d -dimensional data Y is equivalent to minimizing the cost function

$$\varphi(Y) = \sum_i (y_i - \sum_{j=1}^k w_{ij} y_{ij})^2 \quad (5)$$

It can be observed that the coordinates of the low-dimensional representations y_i that minimize this cost function can be found by calculating the eigenvectors corresponding to the smallest non-null eigenvalues of the product of $(I - W)$. In this formula, I is the identity matrix $n \times n$.

2.4 Stochastic Proximity Embedding

SPE [28] is an iterative algorithm that minimizes the gross stress function of MDS. SPE differs from MDS in the efficient rule it uses to update the current estimate of the representation of low-dimensional data [2]. SPE minimizes the gross stress function MDS as

$$\varphi(Y) = \sum_{ij} (d_{ij} - r_{ij})^2 \quad (6)$$

where r_{ij} is the proximity between the high-dimensional data points x_i and x_j and d_{ij} is the euclidean distance between its low-dimension counterparts y_i and y_j in the current approximation of the embedded space. SPE can be easily performed to hold only the distances on a neighborhood graph G defined in the data, setting d_{ij} and r_{ij} to 0 if $(i, j) \notin G$.

3. Damage Classification Methodology

Data organization follows the guidelines provided in the work of Vitola et al. in 2018 [2], which describes the structural health monitoring system with all the sensor data fusion that includes information from all the actuation phases.

The signal processing methodology to treat the information captured by the sensors in the SHM system begins with the unfolding stage, according to the number of sensors. Then, the experimental samples and the measurements in time of the n receiving sensors would remain. Afterward, the data is scaled to decrease the existence of outliers, using Mean-Centered Group Scaling (MCGS) method [29], for more details, please refer to [2,11]. The next step is the feature extraction, in this work four nonlinear manifold learning algorithms are compared to reduce the dimensionality of the data and create the feature vector at the input of the classification machine. The machine learning algorithm used here is KNN with a single neighbor (Fine KNN) and Euclidean distance. A hold-out cross-validation method is used to validate the methodology splitting the data into two sets: a training set of 25% and a test set of 75% avoiding overfitting. The classification process results in a confusion matrix that is analyzed with the accuracy of the confusion matrix as performance metric expressed with a single number. This metric is described in the work of Ballabio et al. [30]. The steps of the developed damage classification methodology are shown in the Fig.1.



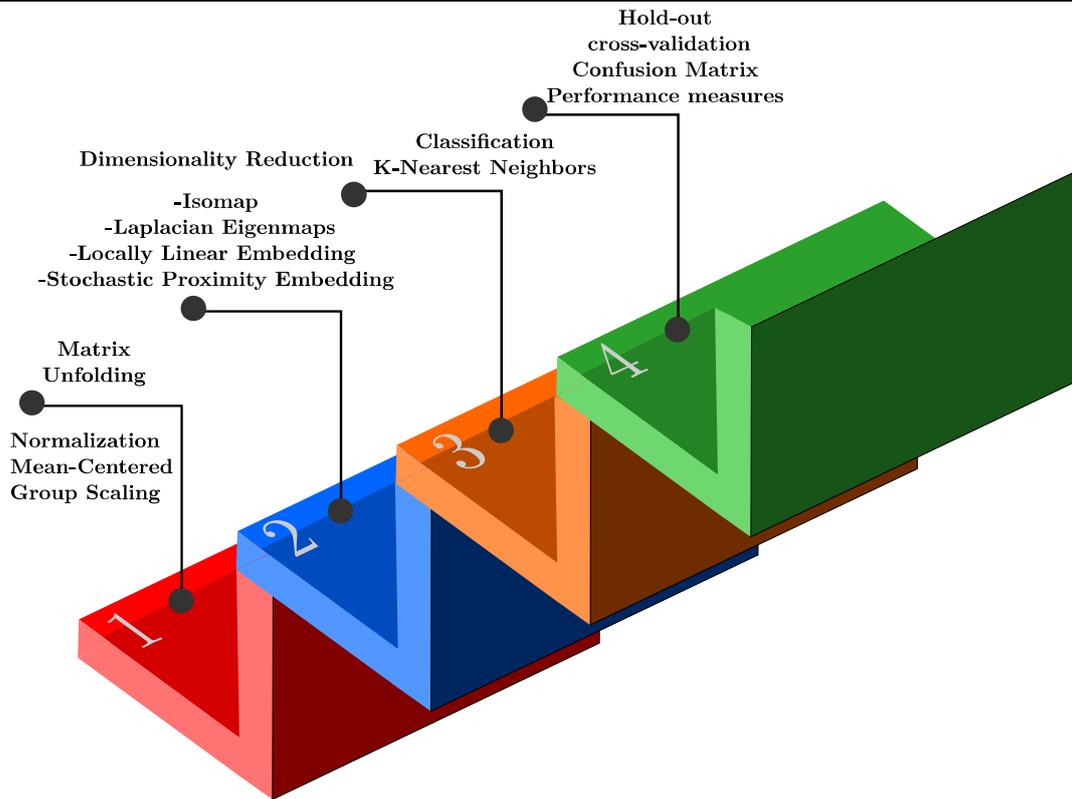


Fig. 1. Main steps of the damage classification methodology.

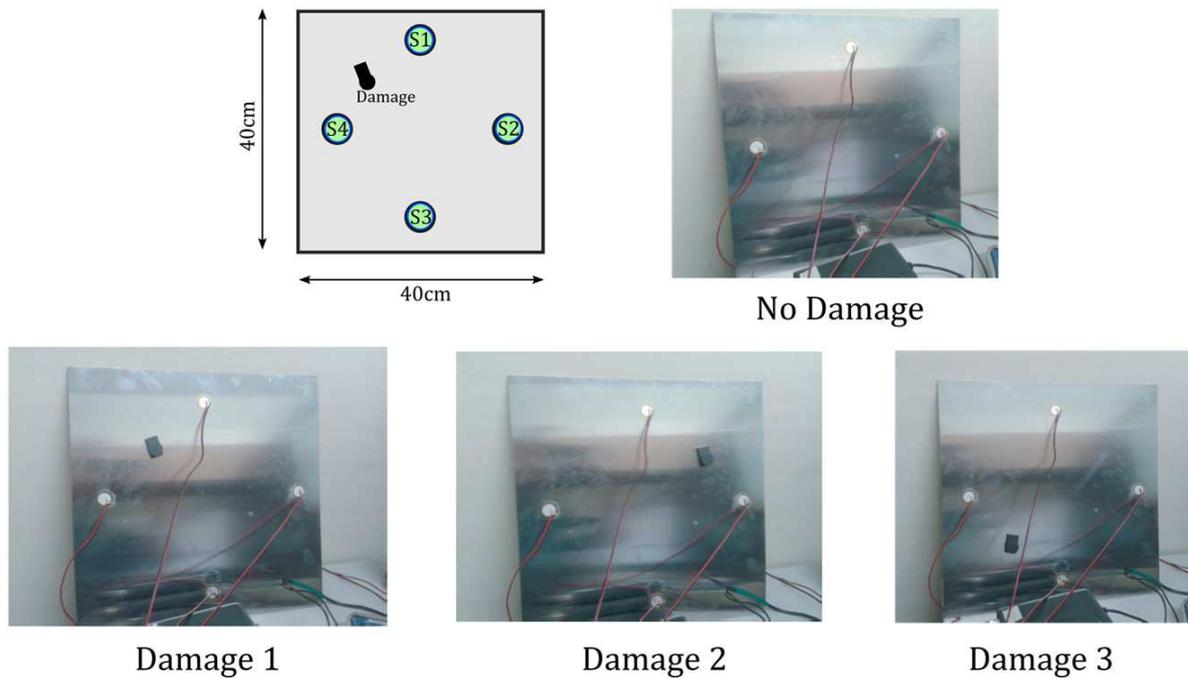


Fig. 2. Aluminium plate instrumented with the four piezoelectric sensors. Plate without damage and with damage 1, 2 and 3.

4. Experimental Setup

As experimental setup, an aluminum plate with dimensions 40cm x 40cm, and a thickness of 0.2cm was used for the validation of the developed damage classification methodology. This plate was instrumented with four piezoelectric sensors [2]. Three different kinds of damage were performed in the structure and the specimen was isolated from the noise and vibration that affect the laboratory. Each damage consists on a added mass located over the structure to modify the propagation of the waves. The inspection system considers the use of a sensor network that is distributed on the surface of the structure. In this work, a piezoelectric sensor network is used to works in several actuation phases. Each actuation phase is defined by the use of a piezoelectric transducer as an actuator, and the rest of the piezoelectric work as sensors. The schematic configuration of the considered structure is shown in Fig. 2.



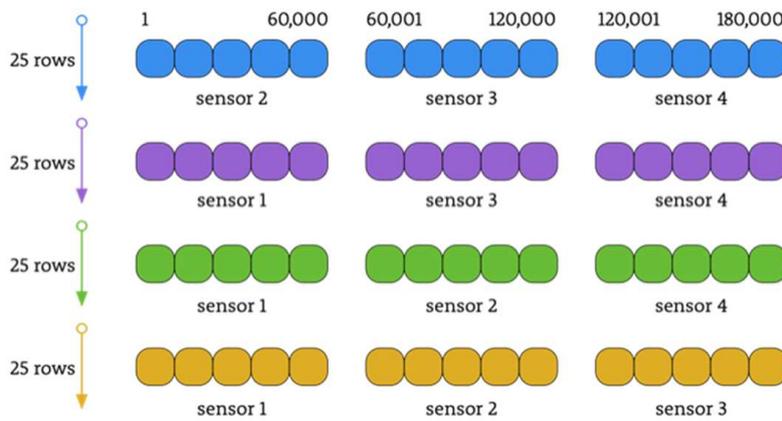


Fig. 3. Organization of raw data for each one of the structural states.

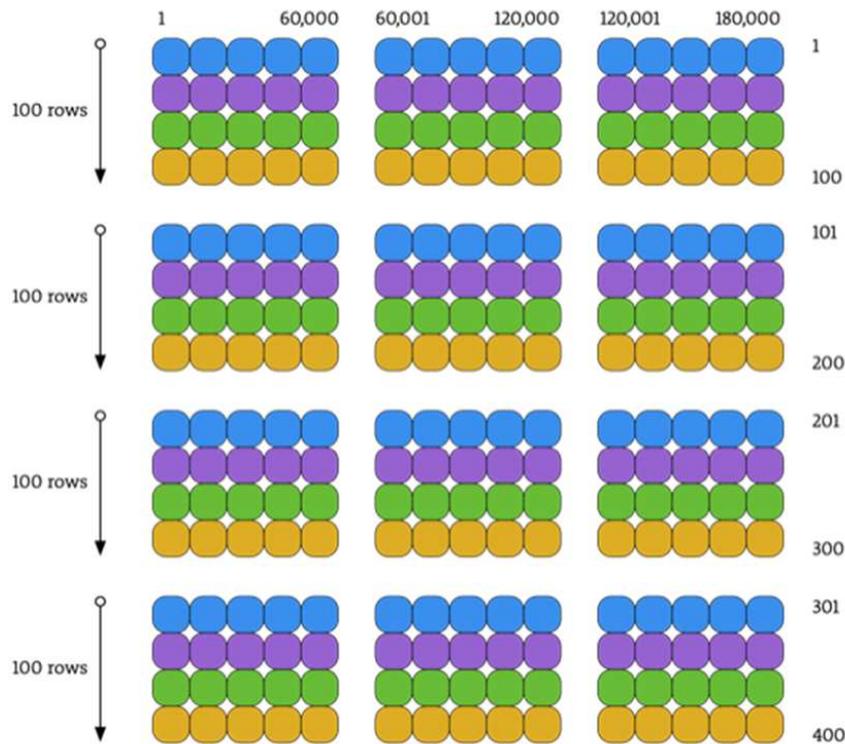


Fig. 4. Organization of raw data including the four different types of plates.

This information is collected and organized in a matrix per actuator. There are four different structural states to classify, first, the plate without damage and second, the plates with damage 1, damage 2 and damage 3. For each classification, 25 experimental samples were obtained by each one of the four sensors functioning as actuators. As a result, 100 experimental samples were recorded for each of the structural states as can be seen in Fig. 3.

Each signal captured by each sensor consists of 60.000 samples and 3 piezoelectrics are used as sensors in each actuation phase. This information is unfolded one after another to complete 180.000 measures, a total of 400 experimental samples for the four different types of plates. Finally, raw matrix information of size 400 x 180.000 is organized and illustrated in Fig. 4.

5. Results and Discussion

Performing feature extraction in the raw data helps to identify the information that will be used at the input of the machine-learning algorithm for classification. The manifold learning algorithms used for dimensionality reduction were SPE, LLE, isomap, and laplacian eigenmaps. Fig. 5 shows how the manifold learning algorithms groups the different types of plates to be classified and separates them, so that intra-class distance is minimized by agglomerating the data of the same class, and the inter-class distance is maximized, facilitating the work of the classification algorithm. Fig. 5a shows the three dimensional scatter plot of the features obtained using the SPE algorithm. It can be note that Fig.5a differs from the rest of the methods, since the classes are distributed over all the domain and they dont form clusters. In contrast, the results from the LLE, isomap and laplacian eigenmaps methods are show in Fig. 5b, Fig. 5c and Fig. 5d respectively. The latter are characterized by scattering the features in such a way a particular class is concentrated near to (0,0,0) point and the other classes are located in more outlying regions located on the axes radially. These behavior is desired as it helps the task performed by the classifier by separating the classes into different clusters.



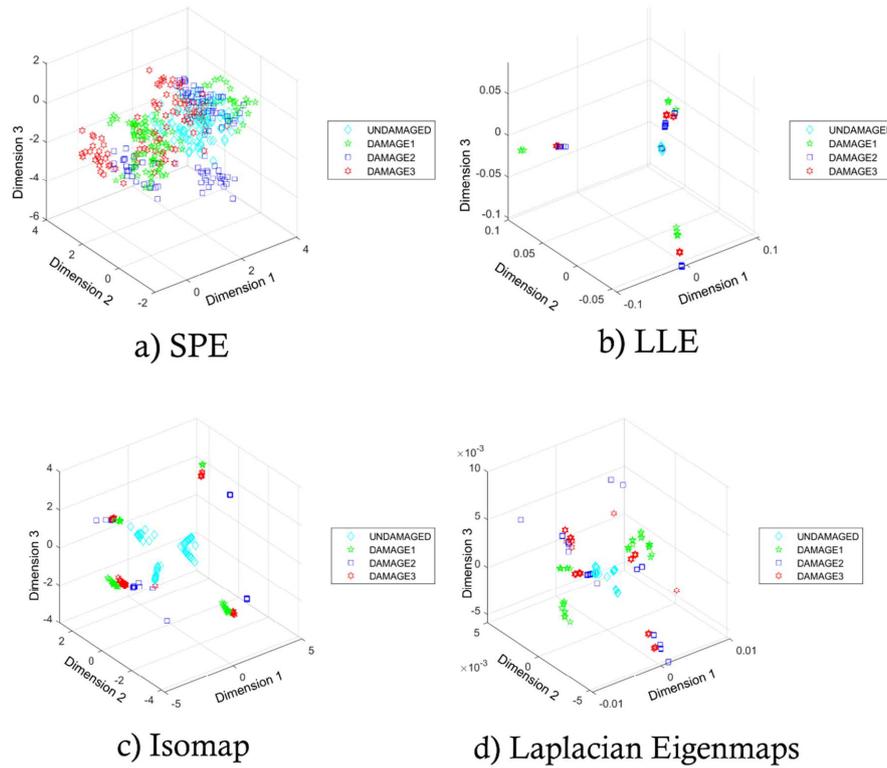


Fig. 5. First three dimensions scatter plot of each manifold learning algorithm used in the dimensionality reduction stage of the damage classification methodology.

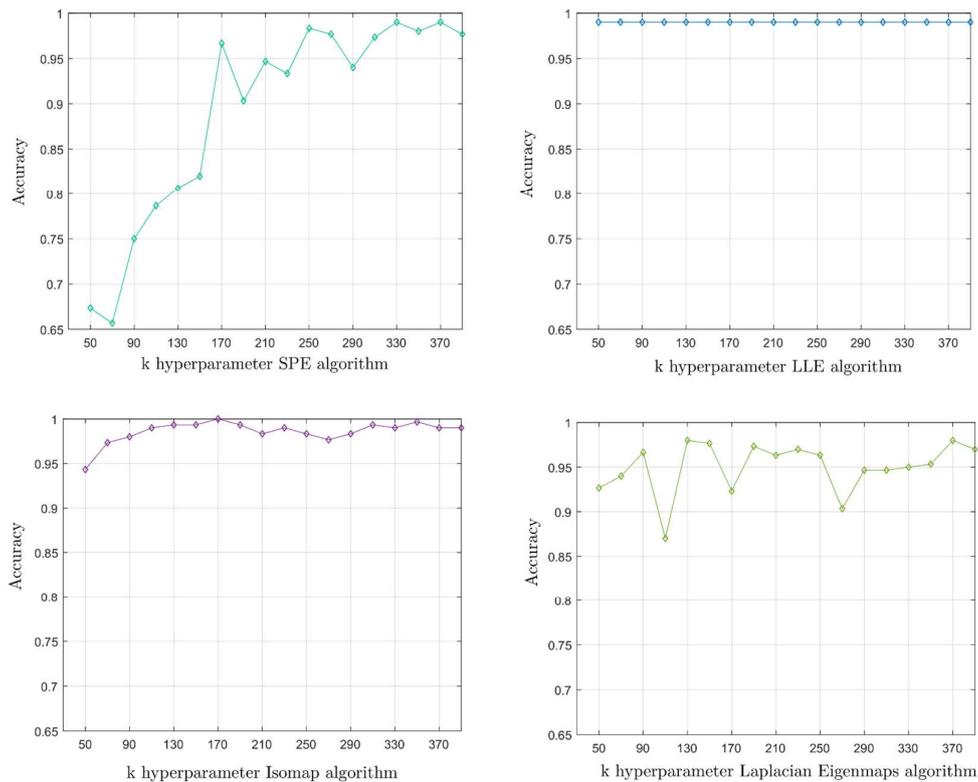


Fig. 6. Behavior of the classification accuracy due to variation of hyperparameter k in each manifold learning algorithm.

Each dimensionality reduction algorithm used has in common the hyperparameter k , which is used to construct the neighborhood graph G to compress the data to a target dimension l . Fig. 6 illustrates the variation of the hyperparameter k , fixing the target dimension of each manifold learning algorithm equal to eight $d = 8$. This tuning process was performed to find the best hyperparameter k within a range of 50 to 400 neighbors. Results in Fig. 6. show the classification accuracy behavior versus the number of k neighbors for each manifold learning algorithm.



Table 1. Accuracy according to the number of dimensions obtained by the manifold learning algorithms as input to the KNN classifier machine

# Dimensions	k =170	k =330	k =130	k =70
	Isomap	SPE	Laplacian eigenmaps	LLE
2	0.9800	0.5933	0.9400	0.9867
3	0.9867	0.8933	0.9433	0.9867
4	0.9933	0.9633	0.9300	0.9900
5	0.9933	0.9700	0.9200	0.9900
6	0.9933	0.9600	0.9833	0.9900
7	0.9967	0.9867	0.9867	0.9900
8	1.0000	0.9900	0.9800	0.9900
9	1.0000	0.9867	0.9800	0.9900
10	1.0000	0.9867	0.9833	0.9900
11	1.0000	0.9867	0.9667	0.9900
12	0.9967	0.9900	0.9633	0.9900
13	0.9967	0.9833	0.9600	0.9900
14	0.9967	0.9900	0.9433	0.9900
15	0.9967	0.9900	0.9300	0.9900
16	0.9967	0.9900	0.9500	0.9900
17	0.9967	0.9833	0.9233	0.9900

Table 2. Results of the confusion matrix of KNN classification considering laplacian eigenmaps k =130 with eight (8) dimensions

Actual class	Predicted class			
	NOT	DMG1	DMG2	DMG3
NOT	75	0	0	0
DMG1	0	75	0	0
DMG2	1	0	74	0
DMG3	1	1	3	70

Table 3. Results of the confusion matrix of KNN classification considering SPE k =330 and LLE k = 70 with eight (8) dimensions

Actual class	Predicted class			
	NOT	DMG1	DMG2	DMG3
NOT	75	0	0	0
DMG1	0	75	0	0
DMG2	2	0	73	0
DMG3	1	0	0	74

Table 4. Results of the confusion matrix of KNN classification considering Isomap k =170 with eight (8) dimensions

Actual class	Predicted class			
	NOT	DMG1	DMG2	DMG3
NOT	75	0	0	0
DMG1	0	75	0	0
DMG2	0	0	75	0
DMG3	0	0	0	75

Previous works have demonstrated that the number of principal components obtained by PCA used for creating a feature vector at the input of a classification machine learning algorithm affects the final classification accuracy [2,13]. In this work, an analysis of varying the number of feature vector dimensions at the input of the KNN method is done. Table 1 shows the results of classification accuracy for the four methods of manifold learning, comparing from 2 to 17 dimensions. It is observed that methodology with Isomap performs the best accuracy in the classification when eight dimensions are used, reaching 100% of accuracy, while laplacian eigenmaps, reached 98.67% of accuracy with seven dimensions, finally LLE and SPE reached 99% of classification accuracy with four and eight dimensions respectively.

After applying the classification procedure by Fine KNN and validating the results with the holdout cross-validation method enunciated in the introduction, the confusion matrix in Table 2 illustrates the results obtained by the laplacian eigenmaps algorithm obtaining a classification accuracy of 98.67%.

After applying the classification procedure by Fine KNN and validating the results with the holdout cross-validation method, the confusion matrix in Table 3 illustrates the results obtained by the SPE and LLE algorithms obtaining a classification accuracy of 99%.

After applying the classification procedure by Fine KNN and validating the results with the holdout cross-validation method, the confusion matrix obtained by the Isomap method is shown in Table 4. This method outperforms the results obtained, reaching a 100% of classification accuracy.

6. Conclusion

The damage classification methodology developed in this work obtained excellent percentages of accuracy. In the feature extraction stage, four different non-linear algorithms were compared from manifold learning: SPE, LLE, isomap, and laplacian eigenmaps. The best behavior of classification accuracy was achieved by the combination of isomap algorithm and the KNN classifier. The methodology allowed successfully to classify different damages in a flat aluminium plate with an accuracy of 100%. Laplacian eigenmaps obtained 98.67% accuracy when considering k =130 and seven dimensions at the input of the KNN classification algorithm. SPE k =330 and LLE k =70 reached 99% of classification accuracy with eight dimensions. Finally, isomap was the best manifold learning algorithm used to dimensionality reduction reaching the best classification accuracy of 100% with k =170 and eight dimensions at the input of KNN classifier. This good behavior evidences that the dimensionality reduction algorithms allow grouping the different types of damages and facilitate the process of distinction made by KNN. Finally, the results of the confusion matrices obtained were analyzed by the classification accuracy as a performance measure that expresses the goodness of classification in a single number. This is useful when it is necessary to optimize the process and identify, for



example, the best number of dimensions necessary to obtain the greatest accuracy in the classification.

Two important aspects of the damage classification methodology were the tuning process of the hyperparameter k , the characteristic of each algorithm of manifold learning when building a neighborhood graph G and the influence of the number of target dimensions bringing for the manifold learning algorithms at the input of the KNN classifier. The comparative study shows that the LLE algorithm is invariant to the change of k parameter reaching 99% of classification accuracy with 8 dimensions to form the feature vector at the input to the KNN classifier algorithm. On the other hand, the influence of the k hyperparameter inside the SPE algorithm evidences an increase of the classification accuracy when k increases and when the number of target dimensions increases. For more complex structures with multiple sensors and nonlinearities, the analysis would be tried dividing the whole data by the different actuation phases. Following a dimensionality reduction procedure by each actuation phase would be tried to finally join the reduce data by each actuation phase. Parallel computing techniques could be explored to simultaneously solve each actuation phase of the damage classification problem. As future work structures with irregular boundaries and internal stiffeners, which cause reflections of elastic waves and a much more complex response at the PZT sensors will be tried.

Author Contributions

J.X. Leon-Medina developed the current damage classification methodology with its stages, programmed the code, participated in the analysis of results and contributed to the writing of the paper. M. Anaya contributed to analyze the measurements, interpreted the results and designed the figures and art work. D.A. Tibaduiza developed the current damage classification methodology with its stages, took part in the pattern recognition tasks, contributed to the analysis of results and to the writing of the paper. F. Pozo developed the current damage classification methodology with its stages, took part in the pattern recognition tasks, designed the figures and art work, contributed to the analysis of results, and to the writing of the paper. All authors discussed the results, reviewed, and approved the final version of the manuscript.

Acknowledgments

We thank to Jaime Vitola for useful discussions with respect to the experimental set-up. The authors thank the Editor and the anonymous reviewers for their valuable comments and suggestions.

Conflict of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Funding

The current research is financed with resources from the FONDO DE CIENCIA TECNOLOGÍA E INNOVACION FCTeI DEL SISTEMA GENERAL DE REGALÍAS SGR. The authors express their gratitude to the Administrative Department of Science, Technology and Innovation – Colciencias with the grant 779 – “Convocatoria para la Formación de Capital Humano de Alto Nivel para el Departamento de Boyacá 2017” for sponsoring the research presented herein. Jersson X. Leon-Medina is grateful with Colciencias and Gobernación de Boyacá for the PhD fellowship. Jersson X. Leon-Medina also thank “Becas de estancias de investigación Secretaría General Iberoamericana (SEGIB) - Fundación Carolina 2019.” This work has been partially funded by the Spanish Agencia Estatal de Investigación (AEI) - Ministerio de Economía, Industria y Competitividad (MINECO), and the Fondo Europeo de Desarrollo Regional (FEDER) through the research project DPI2017-82930-C2-1-R; and by the Generalitat de Catalunya through the research project 2017 SGR 388.

Nomenclature

PCA	Principal Component Analysis	KNN	K-Nearest Neighbors
LLE	Locally Linear Embedding	G	Neighborhood graph
SPE	Stochastic Proximity Embedding		

References

- [1] Leon, J. X., Pineda Muñoz, W. A., Anaya, M., Vitola, J., Tibaduiza, D. A., Structural Damage classification using machine learning algorithms and performance measures, *12th International Workshop on Structural Health Monitoring-IWSHM 2019*, Stanford, California, USA, 2019.
- [2] Vitola, J., Pozo, F., Tibaduiza, D. A., Anaya, M., A sensor data fusion system based on k-nearest neighbor pattern classification for structural health monitoring applications. *Sensors*, 17(2), 2017, 417.
- [3] Van der Maaten, L. J. P., An introduction to dimensionality reduction using matlab. *Report*, 1201(07-07), 2007, 62.
- [4] Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., Munch, S., Detecting causality in complex ecosystems. *Science*, 338(6106), 2012, 496-500.
- [5] Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R., Rosen, M. S., Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697), 2018, 487-492.
- [6] Van Der Maaten, L. J. P., Van Den Hh, J., *Dimensionality reduction: A comparative review*. Tilburg, Netherlands: Tilburg Centre for Creative Computing, Tilburg University, Technical Report. 2009.
- [7] Tibaduiza, D. A., Mujica, L. E., Rodellar, J., Güemes, A., Structural damage detection using principal component analysis and damage indices. *Journal of Intelligent Material Systems and Structures*, 27(2), 2016, 233-248.
- [8] Dervilis, N., Antoniadou, I., Cross, E. J., Worden, K., A Non-linear Manifold Strategy for SHM Approaches. *Strain*, 51(4), 2015, 324-331.
- [9] Yildiz, K., Çamurcu, A. Y., Dogan, B., Comparison of dimension reduction techniques on high dimensional datasets. *The International Arab Journal of Information Technology*, 15(2), 2018, 256-262.
- [10] Tibaduiza, D. A., Torres-Arredondo, M. A., Mujica, L. E., Rodellar, J., Fritzen, C. P., A study of two unsupervised data driven statistical methodologies for detecting and classifying damages in structural health monitoring. *Mechanical Systems and Signal Processing*, 41(1-2), 2013, 467-484.
- [11] Anaya, M., Tibaduiza, D. A., Pozo, F., Detection and classification of structural changes using artificial immune systems and fuzzy clustering. *International Journal of Bio-Inspired Computation*, 9(1), 2017, 35-52.
- [12] Tibaduiza, D., Torres-Arredondo, M. A., Vitola, J., Anaya, M., Pozo, F., A damage classification approach for structural health monitoring using machine learning. *Complexity*, 2018, Article ID 5081283.
- [13] Leon-Medina, J. X., Cardenas-Flechas, L. J., Tibaduiza, D. A., A data-driven methodology for the classification of different liquids in artificial taste recognition applications with a pulse voltammetric electronic tongue, *International Journal of Distributed Sensor Networks*, 15(10), 2019, 1-18.



- [14] Zhang, L., Wang, X., Huang, G. B., Liu, T., Tan, X., Taste recognition in E-tongue using local discriminant preservation projection. *IEEE transactions on Cybernetics*, 49(3), 2018, 947-960.
- [15] Leon-Medina, J. X., Anaya, M., Pozo, F., Tibaduiza, D. A., Application of manifold learning algorithms to improve the classification performance of an electronic nose 2020 *IEEE International Instrumentation & Measurement Technology Conference (I2MTC)*, Dubrovnik, Croatia, 2020, 1-6.
- [16] Zhang, L., Tian, F. C., A new kernel discriminant analysis framework for electronic nose recognition. *Analytica Chimica Acta*, 816, 2014, 8-17.
- [17] Agis, D., Pozo, F., A frequency-based approach for the detection and classification of structural changes using t-SNE. *Sensors*, 19(23), 2019, 5097.
- [18] Zhang, J., Li, S. Z., Wang, J., Manifold learning and applications in recognition. In *Intelligent multimedia processing with soft computing* Springer, Berlin, Heidelberg, 2005, 281-300.
- [19] De Ridder, D., Duin, R. P., *Locally Linear Embedding for Classification*, Technical Reports, Delft University of Technology, The Netherlands, 2002.
- [20] Tenenbaum, J.B., Mapping a manifold of perceptual observations. In *Advances in Neural Information Processing Systems*, Vol. 10, Cambridge, MA, USA. The MIT Press, 1998, 682-688.
- [21] Belkin, M., Niyogi, P., Laplacian Eigenmaps and spectral techniques for embedding and clustering. In *Advances in Neural Information Processing Systems*, volume 14, Cambridge, MA, USA. The MIT Press, 2002, 585-591.
- [22] Bull, L. A., K. Worden, R. Fuentes, G. Manson, E. J. Cross, Dervilis, N., Outlier ensembles: A robust method for damage detection and unsupervised feature extraction from high-dimensional data. *Journal of Sound and Vibration*, 453, 2019, 126-150.
- [23] Plastria, F., De Bruyne, S., Carrizosa, E., Dimensionality Reduction for Classification, *4th International Conference on Advanced Data Mining and Applications*, ser. ADMA '08, Chengdu, China, 2008, 411-418.
- [24] Anderson, W.N., Morley, T.D., Eigenvalues of the Laplacian of a graph. *Linear and Multilinear Algebra*, 18(2), 1985, 141-145.
- [25] Cox, T. and M. Cox., *Multidimensional scaling*. Chapman and Hall, London, UK, 1994.
- [26] Ma, Y., and Y. Fu., *Manifold learning theory and applications*. CRC press, Boca Raton, FL, USA, 2011.
- [27] Roweis, S.T., Saul, L.K., Nonlinear dimensionality reduction by Locally Linear Embedding. *Science*, 290(5500), 2000, 2323-2326.
- [28] Agrafiotis, D. K., Stochastic proximity embedding. *Journal of Computational Chemistry*, 24(10), 2003, 1215-1221.
- [29] Pozo, F., Vidal, Y., Salgado, O. Wind turbine condition monitoring strategy through multiway PCA and multivariate inference. *Energies*, 11(4), 2018, 749.
- [30] Ballabio, D., F. Grisoni, Todeschini, R., Multivariate comparison of classification performance measures. *Chemometrics and Intelligent Laboratory Systems*, 174, 2018, 33-44.

ORCID iD

Jersson X. Leon-Medina  <https://orcid.org/0000-0002-9198-1996>

Maribel Anaya  <https://orcid.org/0000-0002-0241-4771>

Diego A. Tibaduiza  <https://orcid.org/0000-0002-4498-596X>

Francesco Pozo  <https://orcid.org/0000-0001-8958-6789>



© 2021 Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (<http://creativecommons.org/licenses/by-nc/4.0/>).

How to cite this article: Leon-Medina J.X., Anaya M., Tibaduiza D.A., Pozo F. Manifold Learning Algorithms Applied to Structural Damage Classification, *J. Appl. Comput. Mech.*, 7(SI), 2021, 1158-1166. <https://doi.org/10.22055/JACM.2020.33055.2139>

