TUNING OF EXPERT SYSTEMS FOR STRUCTURAL DAMAGE DETECTION THROUGH DIFERENTIAL EVOLUTIONARY ALGORITHMS *

Rodolfo Villamizar[1], Jhonatan Camacho[1], Luis Eduardo Mujica[2], Magda L Ruiz[2]

Abstract— **this paper explores the use of Self Organizing Maps (SOM) combined with Principal Component Analysis in order to obtain an expert system able to detect damages in structures. In addition, a differential evolutive algorithm was used to tune properly the neural network parameters, during the training stage. Structural dynamical time responses are reduced, by using principal component analysis, in order to construct a database case for the SOM network. The expert system was tested using experimental data set supplied by the CODALAB group. These data correspond to time signals from several piezoelectric (PZT) sensors attached on the surface of an aircraft turbine blade. Damages are simulated by adding a mass on the surface of the structure at different positions and by exciting the blade by means of high-frequency vibration signals. Numerical results for each scenario are presented and discussed, where feasibility and potential use of this formulation in structural health monitoring is demonstrated, where low identification error indexes were obtained.**

I. INTRODUCTION

In recent years, several methodologies have been used to solve the Structural Health Monitoring (SHM) paradigm. The scope of the algorithms are categorized according to the SHM levels, which allow to define the degree of damage identification **[1]**: i.) Detection ii.)Localization iii.) Quantification iv) Prognosis. Some approaches include the application of digital signal processing techniques as Fourier transform (FT) or discrete wavelet transform (DWT) in order to accomplish SHM tasks. By applying digital signal processing it is possible to identify deviations of vibrational modes in the time/frequency domain **[2]**. Vibrational-based SHM by using DSP techniques have been found in applications for damage detection of bolted flange joints in pipelines for example **[3]**. However, if the SHM scope includes classification of different damage types, it is recommended to use artificial intelligent techniques, such as neural networks, support vector machines or Bayesian methods **[4]**. Neural networks have been trained to detect, locate and quantify mass changes in an Aluminum Beam Structure and young modulus reduction in a Cantilever Truss structure **[5]**. A critical step for the success of SHM algorithms by using artificial intelligence techniques is to tune several parameters involved in the training stage. In this sense, optimization based on genetic algorithms can be used in order to avoid poor performance of the SHM algorithm **[6]**.

Because SHM methodologies involve management of a huge data volume corresponding to signals from the sensors attached to the structure, statistical reduction techniques can be suitable. In this sense, Principal Component Analysis (PCA) has demonstrated to be successful for damage identification applications **[7]**. PCA based methods correlate data recorded from sensors in order to obtain the most significant features in a reduced space. In addition, if the mathematical projection is used as a model for sensing data, statistical indexes can be used to differentiate diverse damage types **[8]**.

Thus, in this paper an approach for automatically tuning parameters of a structural assessment algorithm experimented by CoDalab research group is discussed. The algorithm includes PCA statistics and SOM network, and the automatic tuning is obtained by using differential evolutive algorithms. Experimental data for validation purposes correspond to time vibrational data of a turbine blade structure recorded by using an active piezo-electric system.

II. DAMAGE DETECTION AND CLASSIFICATION ALGORITHM: TECHNICAL FORMULATION

Fig. 1 summarizes the structural damage assessment algorithm used in this paper.

*Research supported by Universidad Industrial de Santander, COLCIENCIAS and CoDaLab group.

[1]Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones (E3T), Universidad Industrial de Santander (UIS). Grupo de Control Electrónica Modelado y Simulación (CEMOS). Santander, Colombia. Corresponding autor: rovillam@uis.edu.co phone: +57(7) 6344000. Second email: Camacho.navarro.jhonatan@gmail.com

[2] Department of Applied Mathematics III, Escola Universitària d'Enginyeria Tècnica Industrial deBarcelona (EUETIB), Universitat Politécnica de Catalunya (UPC) BARCELONATECH, Comte d'Urgell 187, E-08036, Barcelona, Spain. Emails: luis.eduardo.mujica@upc.edu and magda.ruiz@upc.edu.

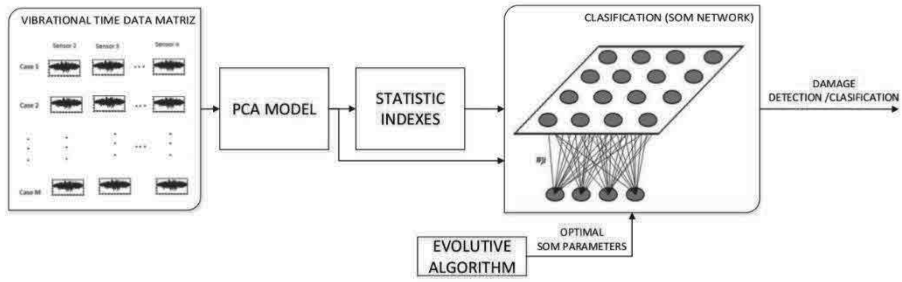

Figure 1. PCA based methodology for damage assessment

An active Piezo-Electric System with N nodes is used to obtain the structural time vibrational response. The structure is excited with a burst high frequency signal on one of the N nodes selected as an actuation point, such as is shown in Fig. 2

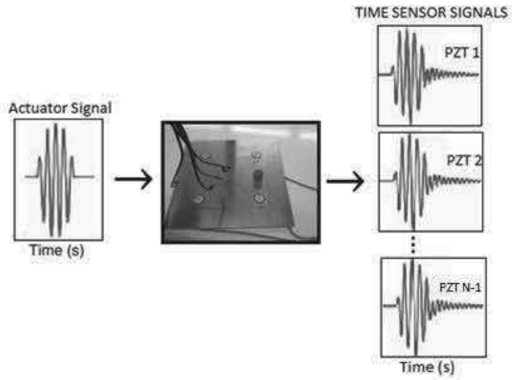

Figure 2. Active Piezo-Electric System

The vibrational time data matrix is obtained by organizing the experimental data from $N-1$ PZT nodes, which operate as sensors for different damage cases. In this sense, r repetitions for similar configuration and condition of the experiment are stored by using p time samples for all M damage cases produced in the structure. Then, PCA is applied to the undamaged data in order to build a PCA model for the healthy structure. The undamaged principal component space is used to project Damage cases onto the reduced representation. The steps followed to obtain a representation for undamaged state are:

i. Normalize the undamaged DataCase Matriz $\boldsymbol{X_{und}}$ by applying GroupScaling (GS) method **[9]**.
ii. Determine the first $\boldsymbol{l}$ principal components by using the NIPALS iterative algorithm **[10]**.

After previous steps are accomplished, the PCA model for the undamaged state is defined by the following elements:

i. Normalization components: means ($\hat{\mu}_i$) and deviations ($\hat{\sigma}_i$). According to GS procedure are obtained $N * t_s$ mean values and N deviations.
ii. The $\boldsymbol{l}$ loading vectors (principal components): $\varphi_i]^{Und} = [\phi_1 \quad \dots \quad \phi_l]$
iii. The variance of principal components (Eigenvalues):$\lambda_i = [\lambda_1 \quad \dots \quad \lambda_l]$

By using a PCA model a projection (scores) of remaining Normalize Damage Cases onto the Undamaged PCA reduced space can be obtained:

$$Z_{damage} = \varphi_i]^{Und} * \left(\frac{X_{Damage}-\hat{\mu}_i}{\hat{\sigma}_i}\right) \quad (1).$$

In addition, two relevant statistics indexes can be calculated: The $\boldsymbol{T^2}$ statistic index measures variations inside the own PCA model and the $\boldsymbol{Q}$ statistic index is a squared 2-norm that measures deviations of the observation respect to the lower-dimensional PCA representation. They are used by their capability to detect deviations of the current vibrational response respect to the undamaged one. Thus, the scores and PCA indexes are used as features to train a SOM Network in order to facilitate visualization of different damage types and to assist classification tasks. The result interpretations is carried out by using graphical tools

III. SOM AUTOMATIC TUNING BY USING DIFFERENTIAL EVOLUTIVE ALGORITHM

The success of the damage classification algorithm depends of an appropriated SOM training. The Table 1 presents the SOM algorithm parameters that are required to be tuned in order to obtain a high quality SOM according to the indexes exposed in that table **[11]**.

TABLE I. SOM PARAMETERS AND INDEXES

SOM parameters	Description	SOM Quality Indexes	Description
Normalization method	Data normalization avoids false dominant clusters. Options: variance/linear range/logarithm/logistic.	**Topographical error**	It is a measurement of topology preservation**.** It should be near to zero
Output neurons number	It is the *clusters* number	**Distortion**	Shows how well each neuron represents the input data
Grid structure	Local topology map. Options: Rectangular/Hexagonal		
Map shape	Local topology map. Options: Laminate/Cylindrical/Toroid	**Histogram uniformity**	It is measurement of the cases distribution in the clusters. Ideally, each cluster should be containing cases of the same type and there is not be empty clusters
Neighborhood function	Interactions between reference vectors. Affects the precision and generalization of the SOM network. Options: *Gaussian*/ cut *Gaussian* / Bubble.		

Thus, a methodology for automatic tuning of the SOM parameters, it is proposed by minimizing a fitness function, which is a weighted sum of SOM quality indexes and damage identification error:

$$f\left(\vec{x}_{i,j}\right) = \sum_{i=1}^{n} w_i * q_i \ + \sum_{j=1}^{M} w_j * e_j \qquad (2)$$

Where $\vec{x}_{i,j}$ Is a vector containing a combination of SOM parameters; w_i, w_j are weighting factors, e_j the classification error for each damage type and q_i are the SOM quality indexes. The fitness function is minimized by applying Differential Evolutive Algorithm (DE).

DE is a stochastic algorithm based on population and focused on a scheme for generating trial parameter vectors. Thus, the main operator consist of adding the weighted difference between two population vectors to a third vector. DE algorithm has been tested on several optimization tasks with successful rates **[12]**. Fig. 3 summarizes the operation mode for DE algorithm

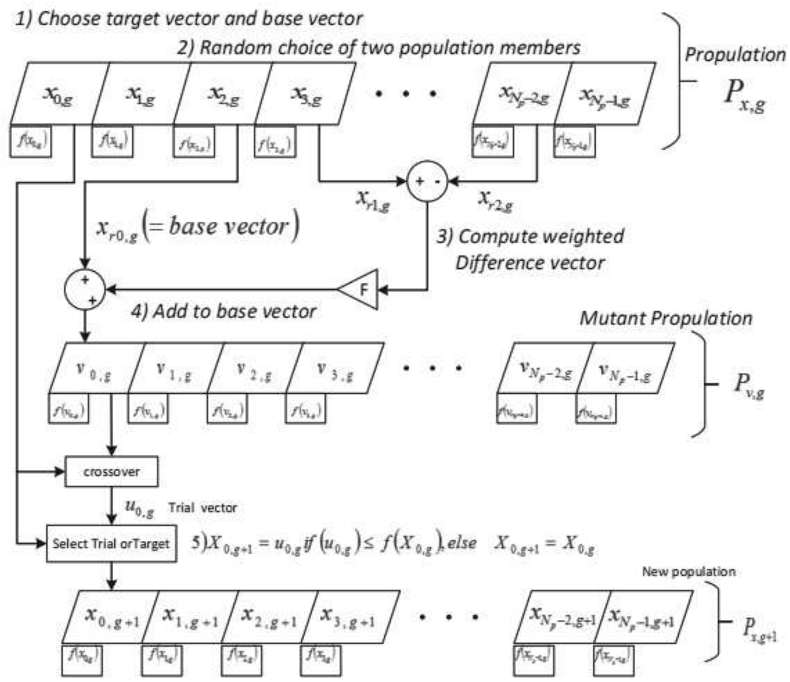

Figure 3. Differential Evolutive Algorithm Operation **[13]**

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The test structure is a turbine blade of a commercial aircraft manufactured by a homogenous material with a similar density than titanium (3.57 g/mL), whose experimental data where supplied by the research group CoDAlab of the Universitat Politécnica de Catalunya (UPC). Time vibrational response was recorded by using an active piezoelectric system with seven PZT nodes distributed over the surface structure (see Fig. 4). PZT 1 was excited by means of a burst signal of 3 peaks and 350 kHz and the other PZTs were used as sensors.

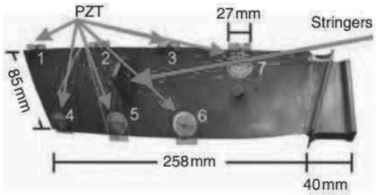

Figure 4. Aircraft turbine blade and active PZT system

Damages were induced by adding masses at several locations shown in Fig. 5. 19 undamaged cases were used to build the PCA model, where 18 principal components were maintained. 100 experiments (10 per each 0f 9-damage types D1-D9 and 10 undamaged cases) were conducted in order to evaluate the performance of the fault detection algorithm.

Figure 5. Locations of simulated damages

Fig. 6 depicts the damage detection indexes T^2 and Q in a scatter plot. Shapes and colors represent different damage types. Original data corresponds to the undamaged cases used to build the PCA model and labeled with tag 'orig'.

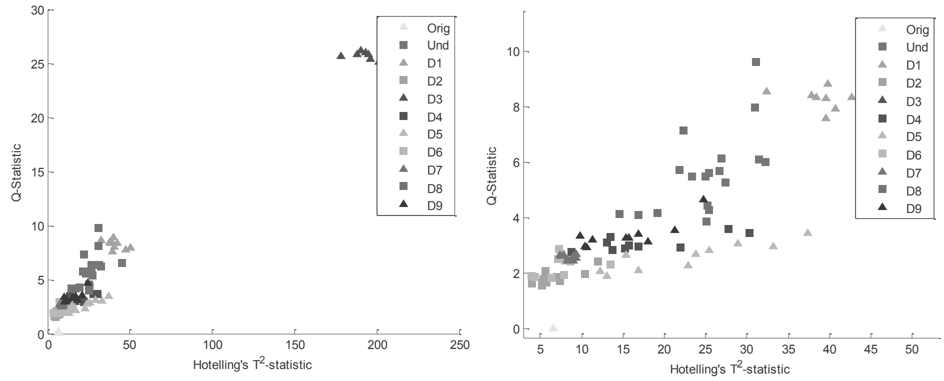

Figure 6. Q and T^2 indexes for damaged cases

In Fig. 6, it is observed that undamaged cases (Orig, Und) are clearly separated from damaged cases (D1-D9). Then, presence or absence of damages can be easily detected by using a PCA model. On other hand, discrimination of damages could be complex for some groups, because they appear quite overlapped. Because only PZT 1 is being excited, damages 5, 6 and 7 are the most difficult to be identify. Also, damages 8 and 9, which take into account quantification performance, appear to be overlapped with their similar located damages 1 and 4. In order to take into account possible nonlinear relations between features, a SOM network was built to map them onto a 2D cluster representation; Figure 7 depicts the resulting Map

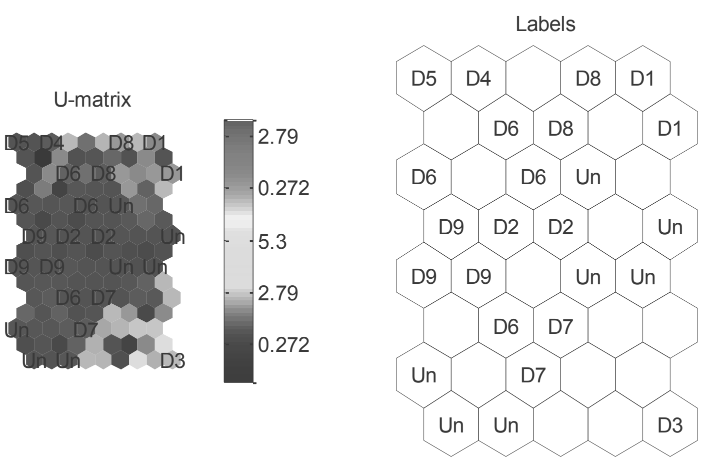

Figure 7. SOM network using default-training values

The SOM was trained by using default values: map size: [8 5]; lattice: 'hexa'; shape: 'sheet'; norm method: 'var'; neigh: 'gaussian'. The Final quantization error: 1.865 and Final topographic error: 0.000. For SOM training purposes 14/29 undamaged cases and the half of damaged cases were used. According to Figure 7, the SOM network has 15/40 empty clusters, which influence empty labeling for validation cases. The damage cases grouped in each cluster are specified in Table 2.

TABLE II. DAMAGE CASES CLUSTER

#→ SOM Clúster							
#	**Cases**	**#**	**Cases**	**#**	**Casos**	**#**	**Cases**
1	'Un',D8','D8'	**11**	'D7'	**21**	'Un', 'D2'	**31**	Empty
2	'D8'	**12**	'D6'	**22**	'Un', 'D2'	**32**	Empty
3	'D8','D9'	**13**	'D6'	**23**	'D9'	**33**	Empty
4	'Un','D9'	**14**	Empty	**24**	'D9'	**34**	'Un','Un','Un','Un','D7'
5	'D4','D5','D9'	**15**	'D6','D6'	**25**	Empty	**35**	'Un', 'Un'
6	Empty	**16**	'D4','D4'	**26**	'D1','D1','D7'	**36**	'Un','Un','Un'
7	'D5', 'D5'	**17**	'D5'	**27**	Empty	**37**	'Un','Un','Un','Un'
8	'D4','D4','D5'	**18**	'D1','D1','D1'	**28**	'D2',D7'	**38**	'Un', 'Un', 'D2'
9	'D8'	**19**	Empty	**29**	Empty	**39**	Empty
10	Empty	**20**	'D6'	**30**	'D2','D7'	**40**	'D3','D3','D3','D3','D3'

In a more detailed view of Table 2, it is clearly identifiable that damage types D3, D6, D9 and undamaged cases appear in separate groups. Fig. 8 shows the labels assigned by SOM network to the training/validation data

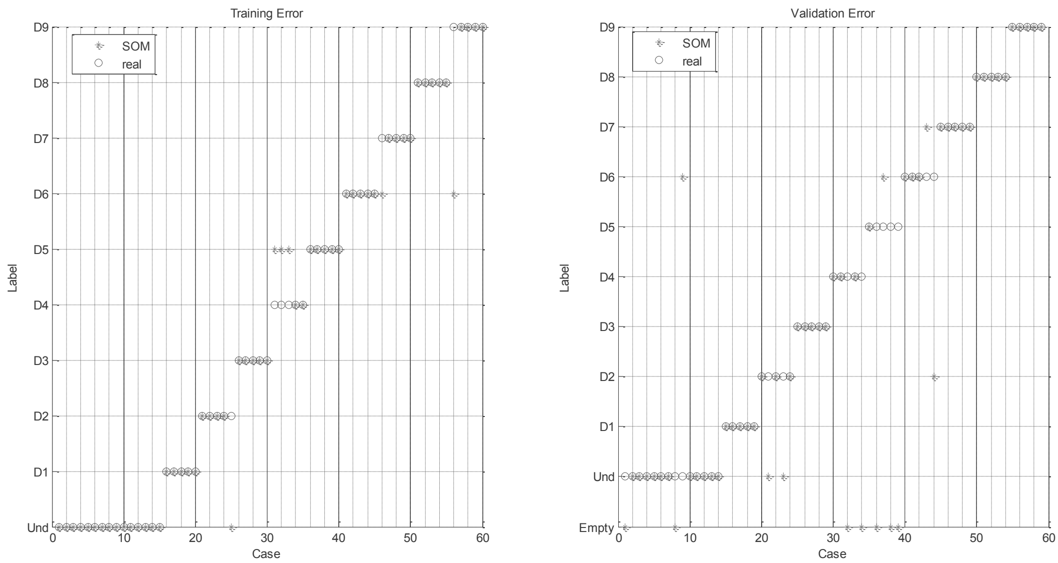

Figure 8. Class assignment by using SOM default-training values

In Figure 8, the training and validation errors correspond to 10% and 22.0339% respectively. Because the SOM empty clusters, 7/59 test data appear without labels and any damage type was assigned to seven validation cases. To improve the SOM quality, 10.000 iterations of a DE algorithm were executed in order to minimize the sum of training and validation errors. The main parameters to solve the optimization problem were set to: CR: 0.2000, F: 0.5629, VTR: 0 and NP: 200. To solve DE algorithm takes about two hours. Figs. 9 and 10 illustrate the evolution of the fitness function for all SOM parameters options.

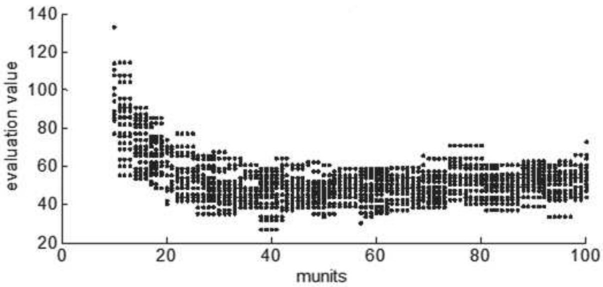

Figure 9. All evaluation values over parameter munits

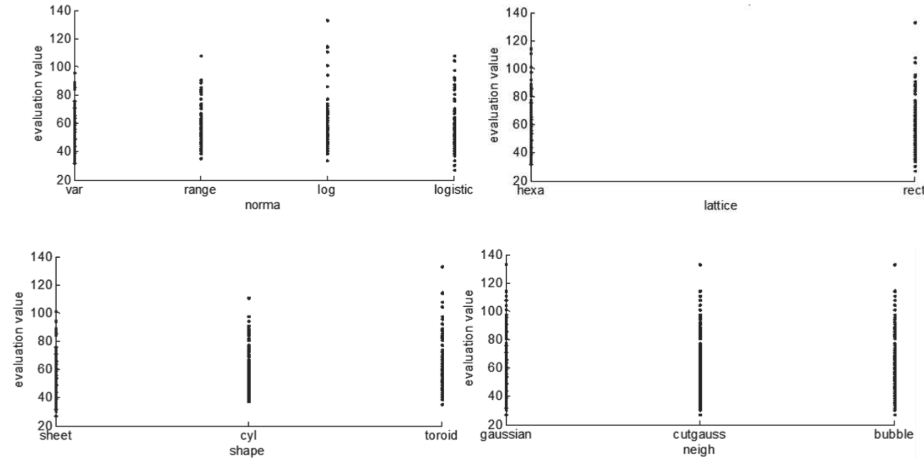

Figure 10. All evaluation values over parameters Norma, lattice, shape and neighborhood

The results obtained in the optimization process show that the munits parameter affects the identification error in major proportion. The SOM cluster numbers must be at least 20 and approximately 40 in order to obtain low identification errors. Otherwise, all other SOM parameters (Norma, lattice, shape and neighborhood) allows low identification errors for their different options. The best configuration for the SOM network parameters, which are suggested by DE algorithm, correspond to norm method: 'logistic', neigh: 'gaussian', msize: [8 5], lattice: 'rect', and shape: 'sheet'. Fig. 11 depicts the final resulting map.

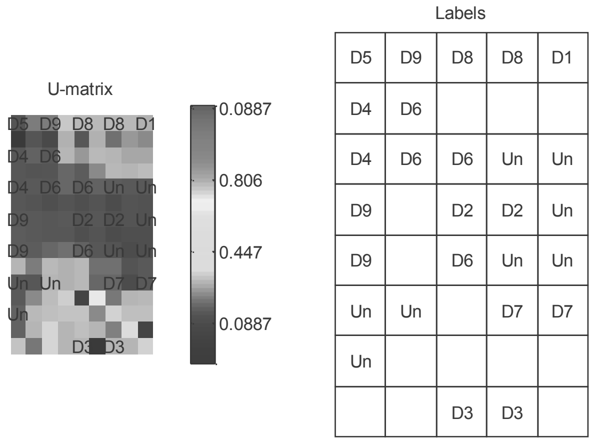

Figure 11. Optimal SOM network

The resulting quantization error, topological error and Distortion measure are 0.3896, 0.0333 and 1.6440 respectively. The distance values illustrated in the U-matrix indicate that damage types are separate by better-defined boundaries. In addition, it is observed that empty clusters were reduced to 13/40. Fig. 12 shows the labels assigned by SOM network to the training/validation data and using the optimal tuning parameters.

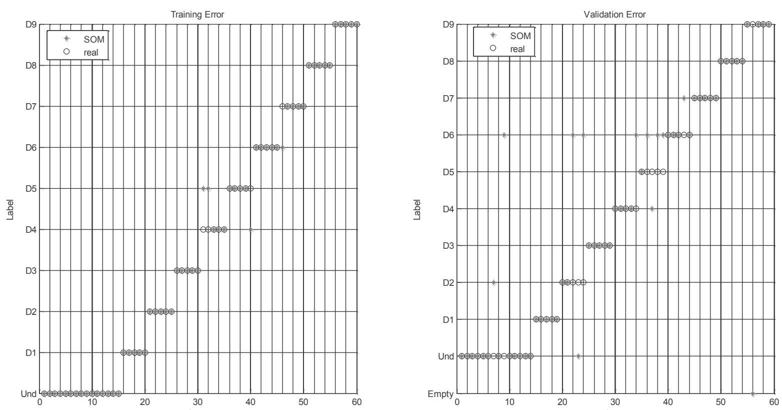

Figure 12. Class assignment by using SOM optimal-training values

The associated training and validation errors were 6.6667% and 20.3390%. In this case, only one validation case was bad classified as empty label. Table 3 details the distribution of damage types in each SOM cluster.

TABLE III. OPTIMAL DAMAGE CASES CLUSTER

#→ SOM Clúster							
#	**Cases**	#	**Cases**	#	**Cases**	#	**Cases**
1	'D5' 'D4' 'D4' 'D5' 'D5' 'D5' 'D5'	**11**	'D6' 'D6'	**21**	'D6' 'D6'	**31**	Empty
2	'D4' 'D4' 'D4' 'D5'	**12**	Empty	**22**	Empty	**32**	'D3' 'D3'
3	'D4' 'D4'	**13**	Empty	**23**	Empty	**33**	'D1' 'D1' 'D1' 'D1' 'D1' 'D1'
4	'D9' 'D9' 'D9''D9'	**14**	'Un' 'Un'	**24**	'D3' 'D3' 'D3' 'D3' 'D3'	**34**	Empty
5	'D9' 'D9'	**15**	Empty	**25**	'D8' 'D8' 'D8'	**35**	'Un' 'Un' 'Un'
6	'Un' 'Un'	**16**	Empty	**26**	Empty	**36**	'Un' 'Un' 'Un' 'Un' 'Un'
7	'Un' 'Un' 'Un' 'Un'	**17**	'D8' 'D8' 'D8' 'D8'	**27**	'Un' 'Un'	**37**	'Un' 'Un' 'Un'
8	Empty	**18**	Empty	**28**	'D2' 'D2'	**38**	'D7' 'D7' 'D7' 'D7'
9	'D9' 'D9'	**19**	'D6' 'D6' 'D6' 'D7'	**29**	'Un' 'Un'	**39**	Empty
10	'D6' 'D6'	**20**	'D2' 'D2' 'D2' 'D2' 'D2'	**30**	'D7' 'D7'	**40**	Empty

In Table 3, it is observed that overlapping between damage types was reduced and separation groups are defined most clearly. However, for damage type 5 only one cluster contains majority (not absolute) of them and real cases are misclassified, where most are classified as damage type 6. Damage 6 has several clusters in the final SOM and most of misclassified cases are labeled as D6 because his high probability of classification.

V. CONCLUSION

Because no previous information is available about the best parameters for tuning a SOM network, it is recommended to apply evolutive strategies in order to obtain good classification tools. The results detailed in this paper show that DE algorithm works adequately for SHM applications. By using default-training parameters, the identification error is around 22%, which is low, however the DE algorithm reduced it to 20% by exploring all options in an optimal strategy. Although the error classification is only reduced 2%, it is observed that SOM characteristics are improved. Thus, empty labels appear less often and generalization error is enhanced because validation test decrease. Also, in the SOM network only a few of empty clusters are present and damage types are grouped in more separable classes. Finally, it is important to emphasize that if not proper tuning parameters are defined for each application, high identification errors can be found as is shown in the obtained results.

ACKNOWLEDGMENT

This work has been developed as part of the *Monitorización y Detección de Defectos en Estructuras usando Algoritmos Expertos Embebidos* research project, financed by the Departamento Administrativo de Ciencia y Tecnología Francisco José de Caldas – COLCIENCIAS and Banca Mundial in the Universidad Industrial de Santander, Colombia. *Agosto* 2012-*Febrero* 2015.

REFERENCES

[1] FARRAR, Charles R.; WORDEN, Keith. An Introduction to Structural Health Monitoring. En New Trends in Vibration Based Structural Health Monitoring. Springer Vienna, 2011. p. 1-17.

[2] FAN, Wei; QIAO, Pizhong. Vibration-based damage identification methods: a review and comparative study. Structural Health Monitoring, 2011, vol. 10, no 1, p. 83-111.

[3] RAZI, Pejman. DEVELOPMENT OF A VIBRATION-BASED HEALTH MONITORING STRATEGY FOR ONSHORE AND OFFSHORE PIPELINES. 2013.

[4] LU, Pengzhen; CHEN, Shengyong; ZHENG, Yujun. Artificial intelligence in civil engineering. Mathematical Problems in Engineering, 2012, vol. 2012.

[5] MUJICA, L. E., et al. A hybrid approach of knowledge-based reasoning for structural assessment. Smart Materials and Structures, 2005, vol. 14, no 6, p. 1554.

[6] ALMEIDA, Leandro Maciel; LUDERMIR, Teresa Bernarda. An Improved Method for Automatically Searching Near-optimal Artificial Neural Networks. In: International Joint Conference on Neural Networks (IJCNN2008), 2008, Hong Kong. IEEE World Congress on Computational Intelligence (WCCI2008), 2008.

[7] MAJID, Nazatul Aini Abd, et al. Aluminium process fault detection by multiway principal component analysis. Control Engineering Practice, 2011, vol. 19, no 4, p. 367-379.

[8] MUJICA, L. E., et al. Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. Structural Health Monitoring, 2011, vol. 10, no 5, p. 539-553.

[9] Ruiz, M. L. Multivariate Statical Process Control and Case-Based Reasoning for Situation. Assessment of Sequencing Batch Reactors, 2008. Universitat de Girona, 168.

[10] WOLD, Svante, et al. Multi-way principal components-and PLS-analysis.Journal of chemometrics, 1987, vol. 1, no 1, p. 41-56.

[11] KOHONEN, Teuvo. Self-organizing maps. Springer, 2001.

[12] Storn, R., "On the Usage of Differential Evolution for Function Optimization" NAFIPS 1996, Berkeley, pp. 519 - 523.
[13] http://www1.icsi.berkeley.edu/~storn/code.html. Accessed on 1/03/2014.