On Self-organising Diagnostics in Impact Sensing Networks

Mikhail Prokopenko[1], Peter Wang[1], Andrew Scott[2], Vadim Gerasimov[1], Nigel Hoschke[2] and Don Price[2]

[1]CSIRO Information and Communication Technology Centre
[2]CSIRO Industrial Physics
Locked bag 17, North Ryde 1670, Australia
{firstname.lastname}@csiro.au

Abstract. Structural health management (SHM) of safety-critical structures requires multiple capabilities: sensing, assessment, diagnostics, prognostics, repair, etc. This paper presents a capability for *self-organising diagnosis* by a group of autonomous sensing agents in a distributed sensing and processing SHM network. The diagnostics involves acoustic emission waves emitted as a result of a sudden release of energy during impacts and detected by the multi-agent network. Several diagnostic techniques identifying the nature and severity of damage at multiple sites are investigated, and the self-organising maps (Kohonen neural networks) are shown to outperform the standard k-means algorithm in both time- and frequency domains.

1 Introduction

Structural health management (SHM) of complex, safety-critical structures such as aerospace vehicles will ultimately require the development of intelligent systems – systems that can process the data from large numbers of sensors; evaluate and diagnose detected damage; form a prognosis for the damaged structure; make decisions regarding response to or repair of the damage; initiate the required actions and monitor their effectiveness. CSIRO, with support from NASA, is investigating approaches to such sensor networks based on complex multi-agent systems principles that are expected to provide the desirable characteristics of robustness, reliability and scalability [1, 12]. The transition from conventional "hot spot" monitoring, which uses relatively few sensors and treats damage detection as a separate task from data analysis and prognosis, to comprehensive SHM employing very large numbers of diverse sensors integrated into the material microstructure, will necessitate the handling of massive amounts of diverse data, leading to potential failures due to information and communications loss, and information overload. In order to solve this problem, we propose an innovative multi-cellular sensor and communication network, with self-monitoring and self-diagnosing capabilities, leading to distributed self-assessment. Data will be processed locally, and only information relevant to other regions of the structure will be communicated.

Thus, the focus of our current research is on developing a capability for *self-organising diagnosis* by a group of autonomous sensing cells in a distributed sensing and processing network, and on demonstrating the outcomes on a large-scale hardware test-bed – the "Ageless" Aerospace Vehicles (AAV) Concept Demonstrator (CD). The AAV-CD consists of "cells", that not only form a physical shell for an aerospace vehicle, but also have sensors, logic, and communications. Currently, each cell contains 4 passive piezoelectric

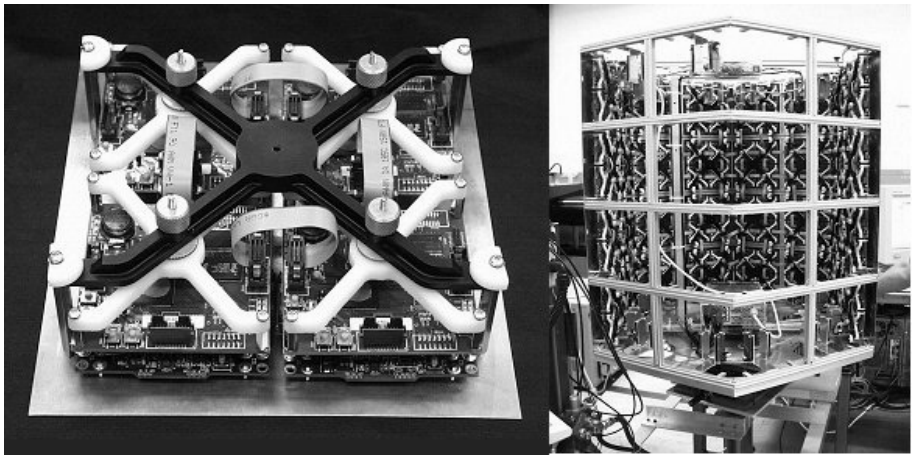

Fig. 1. A single aluminium panel with 4 cells (left) and a general view of the Concept Demonstrator (right), with four of the six sides populated.

polymer sensors, consisting of a 110 μm-thick film of PVDF (polyvinylidene fluoride) coated on both sides with a conductive gold layer, bonded to an aluminium skin panel in order to detect the elastic waves generated in the structure by impacts. The present structure of the AAV-CD is a hexagonal prism. A modular aluminium frame is covered by 220 mm x 200 mm, $1-$mm thick aluminium panels that form the outer skin of the structure. Each such panel contains four cells, and each of the six sides of the prism contains eight of these panels. The entire AAV Concept Demonstrator contains 48 panels, 192 cells, and a total of 768 sensors in the initial system (Figure 1).

Some of the distributed processing involves only interactions between local agents, while other processing requires the emergence of dynamic hierarchical structures. An example of the former is emergent spatial organisation, such as our previously reported work on the formation of impact boundaries [7, 6], evolvable recovery membranes [15] and impact networks around damaged areas [14], which can be used to characterize the extent of damage. The focus of this paper is to investigate diagnostic techniques identifying the nature and severity of damage at multiple sites, but at this stage, not identifying either the causes of damage or the required remedial actions. In particular, we consider the self-organising maps (SOMs or Kohonen neural networks) [9, 10], the Principal Component Analysis (PCA) technique [8], and their combinations. Each considered technique involves collecting input-vectors for several impact signal categories, a training phase using a half of the input-vectors and producing a clustering, and a testing phase when the other half is diagnosed by mapping to produced clusters. Appropriate metrics, e.g. recall and precision, are used to characterise the techniques performance.

At this stage, the diagnostic models are centralised: all the detected inputs are collected prior to the training and testing phases, so that processing is carried out outside the AAV-CD. Eventually, these techniques will be decentralised and embedded in the AAV-CD multi-cellular array, enabling an on-line assessment of damage type and severity. Our main objective is to evaluate the concept of self-organising diagnostics in the context of a self-monitoring impact sensing network. The following Section briefly reviews some background research. Section 3 describes the employed techniques, followed by experimental set-up and diagnostics results (Section 4) and conclusions (Section 5).

2 Background

Acoustic emission (AE) is an ultrasonic wave emitted as a result of a sudden release of energy during a deformation and failure process [3]. The released energy can be detected by sensors sensitive to displacement or velocity. A subsequent analysis of the arrival times of the signals at different sensor locations along with knowledge of the velocity of sound propagation can be used to triangulate the location of the damage mechanism [2, 13]. In some cases, detailed analysis of the acoustic emission signals can also provide information about the nature and severity of the damage. For example, AE methods are being investigated for on-board impact detection for the Space Shuttle — following the Columbia Shuttle accident which resulted from damage to the Shuttle wing's leading edge, caused by impact of foam insulation that broke off of the external tank during ascent [13].

Recently, neural networks have been successfully used in clustering and identifying damage-related AE signals in a highly noisy environment, achieving very good classification results for crack-related signals in the presence of strong time-varying noise and other interference [5] — in particular, the Kohonen network (SOM) was applied for this purpose. Artificial Neural Networks (ANN) and Support Vector Machines (SVM) were also recently considered as approaches to build classifiers in order to assess the structural integrity of shafts and pins based on ultrasonic signatures [11]. It is generally observed that a combination of classifiers may produce a more informative classification. Thus, a sequence of PCA and SOM is used in [5], while a hybrid ANN-SVM classifier is found more informative in [11].

3 Self-organising Maps

The very popular SOM algorithm was originally devised by Teuvo Kohonen [9, 10] as a model of the self-organisation of neural connections, with the ability to produce organisation starting from possibly total disorder. The Kohonen algorithm for self organising feature maps stores prototypes $\overrightarrow{w_{i,j}}$ (also known as codebook vectors) of the input-vectors of an euclidian input space X, $\overrightarrow{x}$, in the neurons (cells) of a neural layer R (a $2D$ grid). The prototype vectors $\overrightarrow{w_{i,j}}$ are usually initialized with random values. At each iteration $t \geq 0$, there are M updates: one for each input-vector $\overrightarrow{x(m)}$, $1 \leq m \leq M$ (we shall denote an update with an index $\tau = tM + m$). Every stimulus $\overrightarrow{x(m)}$ is mapped to the "winner" neuron with the position (i, j) in the neural layer R, which has the prototype vector $\overrightarrow{w_{i,j}}$ with the minimal euclidian distance in input space:

$$\|\overrightarrow{x(m)} - \overrightarrow{w_{i,j}(\tau)}\| = \min_{(k,l)\in R}(\|\overrightarrow{x(m)} - \overrightarrow{w_{k,l}(\tau)}\|) \ ,$$

where $\|.\|$ denotes the euclidian distance in input space. The winner neuron corresponds to the highest neural activity, being the "center of excitation". The prototype vector of the winner neuron is updated, becoming more sensitive to that type of input. This allows different cells to be trained for different types of data. The neighbors of the winner neuron adjust their prototype vector towards the input vector as well, but in a lesser degree, dependent on their distance from the winner. Usually a neighbourhood function $\eta_{(i,j),(k,l)}(t)$ that decreases both with time and distance between positions (i, j) and (k, l) on the planar neural layer (e.g., a radial symmetric Gaussian neighbourhood function), is used for this purpose:

$$\overrightarrow{w_{k,l}(\tau+1)} = \overrightarrow{w_{k,l}(\tau)} + \alpha(t)\ \eta_{(i,j),(k,l)}(t)\ [\ \overrightarrow{x(m)} - \overrightarrow{w_{k,l}(\tau)}\ ] \ .$$

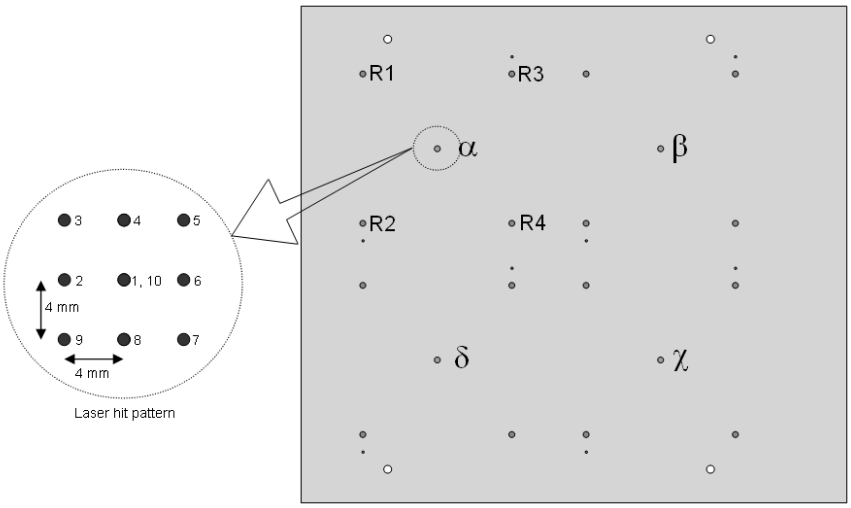

Fig. 2. A pattern of laser impacts.

The function $\alpha(t)$, such that $0 < \alpha(t) < 1$ determines the speed of learning and can be reduced during the learning process. The radius of the neighbourhood $\rho(t)$ decreases with each iteration as follows:

$$\rho(t) = \rho_0 \, e^{-\frac{t \ln \rho_0}{N}}$$

with N being the maximal number of iterations. In our experiments, a linear neighbourhood function outperformed the Gaussian function.

Since prototype vectors, whose neurons with positions (k, l) are close to the winner with the position (i, j), are always updated together, a topology preservation is achieved — neurons that are closer in the neural layer tend to respond to inputs that are closer in the input space. In other words, SOM learns a smooth mapping of the input space to the neural layer. However, the solutions obtained by SOM might be different due to different trajectories through the search space, and therefore the input-vectors can be reused many iterations ("epochs") during the training phase, in different orders.

4 Experimental Results

In order to carry out and evaluate self-organising diagnostics in the AAV impact sensing network, we collected the data containing the transients (voltage vs time) from PVDF transducers attached to a separate aluminium panel that is subject to laser or mechanical impacts, with high or low energy. Laser impacts have been produced by pulses with either the energy of 300 mJ or 30 mJ using a Q-switched Nd:YAG laser focused onto the surface to simulate particle impacts by ablating material from the aluminum. Impacts did not pierce the skin, but left a significant crater in the aluminum surface. The outputs were detected by four receivers while the laser impact site is changed over the panel in a specific pattern (Figure 2). The sensors were in quadrant α only, at positions $R1$, $R2$, $R3$ and $R4$. Ten laser impacts, for each energy level, were aimed near the centre of each quadrant equidistant from each, with each subsequent impact moving 4-mm in the sequence, indexed from 1 to 10, so that the first and the tenth strike at the centre.

The mechanical impacts resulted from a stainless steel-tipped pendulum bob. Again, the outputs were detected by four receivers, while striking the panel ten times at each

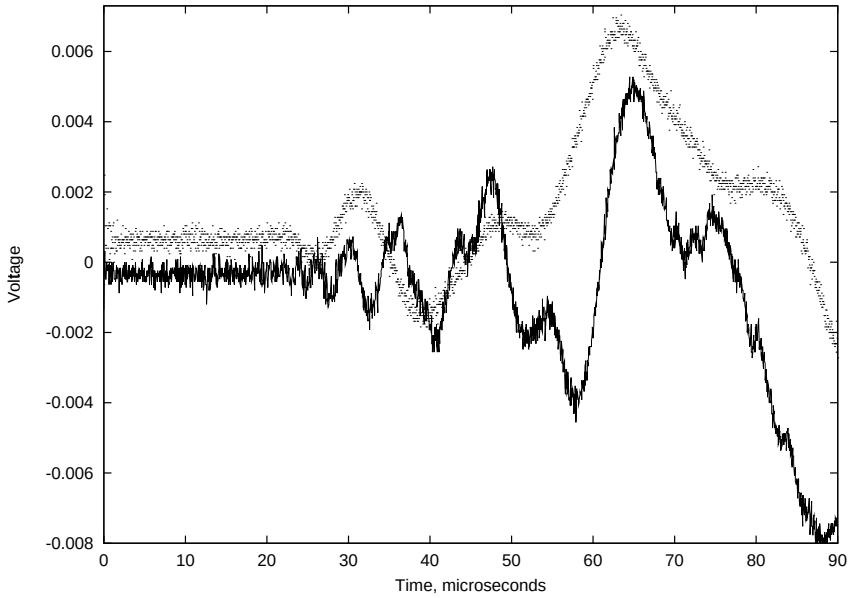

Fig. 3. A laser signal (high energy) in quadrant β, location 5, detected by sensor $R2$ is shown with solid lines. A pendulum signal (low energy) in quadrant α detected by sensor $R1$ is shown with dots.

quadrant. Ten pendulum impacts, for each energy level, were aimed at the centre of each quadrant — unlike laser shots, they were not moved around on a 4-mm grid.

The goal of clustering and diagnostics is to identify and distinguish between different impact types (laser vs pendulum) and different impact energies (high vs low) — the distances to impacts are not intended to be determined. In other words, the process of varying impact locations in the laser experiments does not pursue the goal of representing different distances in the data-set, but has a rather practical purpose of avoiding the panel fatigue from repeated laser impacts.

The collected data present a number of challenges for clustering and diagnostics. First of all, there are different impact types and different energy levels, creating 4 objective categories. For example, a high-energy impact from the quadrant β may be very similar to a low-energy impact from the quadrant α — especially if the signals of the same impact, detected by four sensors, are treated separately from each other. Laser impacts, in addition, vary due to the shifting impact locations: for instance, impacts at locations 1 and 10 differ from those at 3,5,7,9, and from those at 2,4,6,8. Moreover, there are signal reflections from the panel's edges. More precisely, in all of these signals, the first arrival corresponds to the S_0 Lamb wave (the lowest order extensional wave), whose faster low frequency components propagate at 5.3 mm$/\mu s$ in this material. This is followed by the A_0 (flexural) wave, for which the higher frequency modes propagate faster than the lower frequency components, at the Rayleigh wave velocity (3 mm$/\mu s$ in aluminum). The presence of both waves in the signal makes the laser impacts quite hard to cluster and properly diagnose. Figure 3 shows two signals: a laser with high energy, and a pendulum with low energy. Finally, the training set is relatively small containing 80 impacts detected by 4 sensors.

Each signal contains 2048 time-domain points. Moreover, the signals from the same impact detected by four sensors were concatenated into one input-vector — to better account for the varying distances. In practice such merging can be easily achieved after an AAV cell determines that its four sensors detected the same signal. The training (clustering) phase used signals corresponding to the impact locations from 6 to 10, leaving locations from 1 to 5 to the testing (diagnostics) phase.

In order to comparatively evaluate SOM performance, we used k-means clustering algorithm [4], augmented either with the Principal Component Analysis (PCA) [8], or Discrete Fourier Transform. The PCA transforms a data-set into a new coordinate system is such a way that the new first coordinate axis has the greatest possible variance of the data-set projection, the second axis has the second greatest possible variance of the projections, etc. The ordered new coordinates are called the Principal Components (PCs). In many applications the corresponding variance rapidly decreases from the first to the last PC. A reduction of data-set dimensionality with minimal information loss can be achieved by using only a few first PCs to represent the data. This is a common technique in image compression and pattern recognition. The PCA helps to improve performance and reduce memory demand of data processing algorithms. However, this technique is not optimized for clustering, which, in combination with the information loss, degrades the quality of the clustering results. Although the PCA is frequently used as a part of a processing sequence in data classification, the true reason to include it is not to improve the classification results, but to reduce the computational demands of the algorithms that follow the PCA. In our experiments we used Singular Value Decomposition (SVD) based algorithm to calculate the PCA.

The Discrete Fourier Transform (DFT) also changes the way the data-set is distributed. It does not guarantee a better clustering either. But if the data-set represents acoustic or electromagnetic signals, the spectral representation of the signals may lead to a better discrimination between different classes of the signals. This was the case with the data-set in our experiments. The K-means algorithm applied to the DFT-processed data-set led to slightly better results than the same algorithm applied to the raw data. To calculate the DFT we used the Fast Fourier Transform (FFT) algorithm embedded in Matlab.

In summary, the clustering was done with a) k-means algorithm applied to the input space; b) k-means algorithm applied to the PCA space; c) k-means algorithm applied to the frequency domain (using Fourier Transform of the input-vectors, FFT) d) SOM of the input space; e) SOM of the FFT space. A comparison between a) and b) is expected to show the benefits of using PCA as a preliminary step, while a comparison between a) and c) would highlight the benefits of FFT. On the other hand, a comparison between a) and d), as well as between c) and e) may support the usage of SOMs in diagnostics.

Input-vectors are labelled with 4 categories (“laser-high”, “laser-low”, “pendulum-high”, “pendulum-low”), and these labels are propagated to resultant clusters (using majority-voting), enabling diagnostics and its evaluation. We used several metrics for each scenario: recall, precision and effectiveness. The recall measures how many test input-vectors x were successfully matched to the correct cluster category out of y existing test vectors from this category: $\delta_i = x_i/y_i$. The precision measures how many test vectors were attracted to each clustered category in relation to the total size of the cluster: $\pi_i = x_i/z_i$, where z_i is the number of all test vectors placed in the cluster. For example, is the cluster 3 attracted $z_3 = 25$ vectors, out of which $x_3 = 19$ vectors belonged to

the correct category (which contains $y = 20$ vectors in total), then $\delta_3 = 19/20 = 0.95$ and $\pi_i = 19/25 = 0.76$. In other words, 95% of the test vectors from this category are correctly recalled (diagnosed) with the precision of 76%. The effectiveness is defined as a harmonic mean of recall and precision: $q_i = 2\delta_i\pi_i/(\delta_i + \pi_i)$, and we use the average effectiveness $q = \sum_{i=1}^{4} q_i/4$ in the comparative analysis.

Technique	Laser				Pendulum				
	High		Low		High		Low		
	δ_1	π_1	δ_2	π_2	δ_3	π_3	δ_4	π_4	q
a) k-means: k=6, no PCA	0.15	1.0	1.0	0.28	0.25	1.0	0.0	0.0	0.2739
k-means: k=60, no PCA	0.4	1.0	0.95	0.76	1.0	1.0	0.90	0.67	0.7955
b) k-means: k=32, 35 PCs	0.35	1.0	0.65	0.65	1.0	0.91	0.90	0.58	0.7067
c) k-means: k=50, FFT	0.45	1.0	0.85	0.77	1.0	1.0	1.0	0.69	0.8116
d) SOM: linear η, mean	0.42	0.95	0.83	0.61	0.99	0.98	0.93	0.80	0.7817
SOM: linear η, maximum	0.55	1.0	0.95	0.68	1.0	1.0	0.95	0.90	0.8570
e) SOM-FFT: mean	0.44	0.95	0.88	0.60	0.97	0.98	0.88	0.81	0.7814
SOM-FFT: maximum	0.55	1.0	0.95	0.70	1.0	1.0	1.0	0.91	0.8676

Table 1. Diagnostics: results of testing.

The results of diagnostics are summarised in table 1. The k-means algorithm applied to the PCA-processed data-set had inferior results comparing to the same algorithm applied to the original data. The number of clusters maximising the effectiveness was quite large with or without the FFT step ($k = 50$ and $k = 60$ respectively). FFT helped only marginally — however, there is a clear benefit in using it in the presence of noise. We trained 20 SOMs in the time-domain, each having a square layer of 400 neurons, with a linear neighbourhood function, and calculated the mean recall and precision over these 20 runs, followed by the effectiveness of these mean values (shown in the table). The average performance of these SOMs approached the best result of k-means, while the best SOM out of the trained set exceeded this result. The SOMs applied in the frequency domain have also shown only a marginal improvement, and their best SOM has also outperformed k-means with the FFT step. Thus, we can conclude that SOM technique is quite appropriate when the training set is small and in the presence of noise.

5 Conclusions

In this paper we evaluated several techniques aimed at self-organising diagnostics in impact sensing networks. The self-organising maps (Kohonen neural networks) have shown to be more promising, outperforming the standard k-means algorithm in both time- and frequency domains, and suggesting a way to embed self-organising diagnostics in distributed and decentralised SHM systems. The common feature of all experiments was a difficulty in recalling the "Laser-High" category. In fact, a better account of this category was a distinguishing contributor to the overall success of the SOM. It can be observed that the signals from this category are mostly placed in the "Laser-Low" category, reducing its precision. In other words, the energy levels are harder to recognise in the higher-velocity signals detected by remote sensors. This observation calls for a comprehensive analysis of optimal sensor density and layout, which is another subject of future work.

References

1. Abbott, D., B. Doyle, J. Dunlop, T. Farmer, M. Hedley, J. Herrmann, G. James, M. Johnson, B. Joshi, G. Poulton, D. Price, M. Prokopenko, T. Reda, D. Rees, A. Scott, P. Valencia, D. Ward, and J. Winter. Development and Evaluation of Sensor Concepts for Ageless Aerospace Vehicles. Development of Concepts for an Intelligent Sensing System. Technical Report NASA/CR-2002-211773, Langley Research Center, Hampton, VA, 2002.
2. Abbott, D., J. Dunlop, G. Edwards, T. Farmer, B. Gaffney, M. Hedley, N. Hoschke, P. Isaacs, M. Johnson, C. Lewis, A. Murdoch, G. Poulton, D. Price, M. Prokopenko, I. Sharp, A. Scott, P. Valencia, P. Wang, and D. Whitnall. Development and Evaluation of Sensor Concepts for Ageless Aerospace Vehicles. Report 5. Phase 2 — Implementation of the Concept Demonstrator. CSIRO Telecommunications and Industrial Physics. Confidential Report No. TIPP 2056, April 2004.
3. Drouillard, T. F. A history of acoustic emission. *Journal of Acoustic Emission*, vol. 14, no. 1, 1–34, 1996.
4. Elkan, C. Using the Triangle Inequality to Accelerate k-Means. In Proceedings of the 20th International Conference on Machine Learning, 147–153, Washington, DC, 2003.
5. Emamian, V., M. Kaveh, and A. Tewfik. Robust Clustering of Acoustic Emission Signals using the Kohonen Network. In Proceedings of the 25th IEEE International Conference on Acoustics, Speech and Signal Processing, Istanbul, Turkey, 2000.
6. Foreman, M., M. Prokopenko, and P. Wang. Phase Transitions in Self-organising Sensor Networks. In Banzhaf, W., Christaller, T., Dittrich, P., Kim, J.T. and Ziegler, J. (Eds.) *Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life*, 781–791, LNAI 2801, Springer Verlag, 2003.
7. Lovatt, H., G. Poulton, D. Price, M. Prokopenko, P. Valencia, and P. Wang. Self-organising Impact Boundaries in Ageless Aerospace Vehicles. In Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems, 249–256, Melbourne, 2003.
8. Jolliffe, I.T. *Principal Component Analysis*. New York, Springer, 1986.
9. Kohonen, T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59–69, 1982.
10. Kohonen, T. Analysis of a simple self-organizing process, *Biological Cybernetics*, 44, 135–140, 1982.
11. Lee, K., V. Estivill-Castro. A Hybrid Classification Approach to Ultrasonic Shaft Signals. In G.I. Webb, X. Yu (Eds.) *AI 2004: Advances in Artificial Intelligence, 17th Australian Joint Conference on Artificial Intelligence, Cairns, Australia, December 4-6, 2004, Proceedings*, LNCS 3339, 284–295, Springer 2004.
12. Price, D., A. Scott, G. Edwards, A. Batten, T. Farmer, M. Hedley, M. Johnson, C. Lewis, G. Poulton, M. Prokopenko, P. Valencia, and P. Wang. An Integrated Health Monitoring System for an Ageless Aerospace Vehicle. *Structural Health Monitoring 2003: From Diagnostics & Prognostics to Structural Health Management*, F.-K. Chang (Ed.), DEStech Publications, 310–318, 2003.
13. Prosser, W.H., S.G. Allison, S.E. Woodard, R.A. Wincheski, E.G. Cooper, D.C. Price, M. Hedley, M. Prokopenko, A. Scott, A. Tessler and J.L. Spangler. Structural health management for future aerospace vehicles. In Proceedings of the 2nd Australasian Workshop on Structural Health Monitoring, Melbourne, Australia, December 2004.
14. Wang, P., P. Valencia, M. Prokopenko, D. Price, and G. Poulton. Self-reconfigurable sensor networks in ageless aerospace vehicles. In Proceedings of the 11th International Conference on Advanced Robotics, 1098–1103, Coimbra, Portugal, 2003.
15. Wang, P., and M. Prokopenko. Evolvable Recovery Membranes in Self-monitoring Aerospace Vehicles. In Proceedings of the 8th International Conference on Simulation of Adaptive Behaviour, 509–518, Los Angeles, USA, July 2004.