Journal of Sound and Vibration (2001) **242**(5), 813–827
doi:10.1006/jsvi.2000.3390, available online at http://www.idealibrary.com on IDEAL®

STRUCTURAL DAMAGE DETECTION USING ARTIFICIAL NEURAL NETWORKS AND MEASURED FRF DATA REDUCED VIA PRINCIPAL COMPONENT PROJECTION

C. ZANG AND M. IMREGUN

Department of Mechanical Engineering, Imperial College of Science, Technology & Medicine, Exhibition Road, London SW7 2BX, England. E-mail: c.zang@ic.ac.uk, m.imregun@ic.ac.uk

(*Received* 11 *February* 2000, *and in final form* 31 *August* 2000)

This paper deals with structural damage detection using measured frequency response functions (FRFs) as input data to artificial neural networks (ANNs). A major obstacle, the impracticality of using full-size FRF data with ANNs, was circumvented by applying a principal component analysis (PCA)-based data reduction technique to the measured FRFs. The compressed FRFs, represented by their projection onto the most significant principal components, were then used as the ANN input variables instead of the raw FRF data. The output is a prediction for the actual state of the specimen, i.e., healthy or damaged. A further advantage of this particular approach was found to be the ability to deal with relatively high measurement noise, which is of common occurrence when dealing with industrial structures. The methodology was applied to the measured FRFs of a railway wheel, each response function having 4096 spectral lines. The available FRF data were grouped into x, y and z direction FRFs and a compression ratio of about 400 was achieved for each direction. Three different networks, each corresponding to a co-ordinate direction, were trained and verified using 80 PCA-compressed FRFs. Twenty compressed FRFs, obtained from further measurements, were used for the actual damage detection tests. Half of the test FRFs were polluted further by adding 5% random noise in order to assess the robustness of the method in the presence of significant experimental noise. The results showed that, in all cases considered, it was possible to distinguish between the healthy and damaged states with very good accuracy and repeatability.

1. INTRODUCTION

Since the introduction of parallel distributed processing by Rumelhart *et al.* [1], numerous artificial neural network (ANN) techniques have been applied to structural health monitoring and damage detection [2, 3]. Recent research indicates that neural networks, such as the radial basis function (RBF) and the multi-layer back propagation (BP), can be trained on measured frequency responses of healthy and damaged specimens to recognize the actual condition of the structure. For instance, Wu *et al.* [4] identified the damage in a three-storey building model by selecting the first 200 points of the frequency response functions (FRFs) as input to a BP neural network. Chaudhry and Ganino [5] used measured FRF data over some specified frequency range as input to a BP neural network to identify the presence and severity of delamination in debonded beams.

Although such work demonstrates the feasibility of training neural networks on FRF data for damage detection, a very significant hurdle remains: the size of the FRF data, which

0022-460X/01/200813 + 15 $35.00/0

is determined by the number of spatial response locations and the number of spectral lines, is too large for neural network applications to representative engineering problems. The direct use of such data will lead to neural networks with a very large number of input nodes, which in turn will require a very large number of connections. Such networks are known to be impractical, both in terms of training effort and convergence stability. To avoid large neural networks, one can use a subset of the available FRF data by considering fewer measurement points and by defining frequency windows. Atalla and Inman [6] updated a 15-DOF lumped-parameter system and a flexible frame structure using an RBF neural network, the input to which consisted of pre-selected frequency points from a number of frequency windows. An alternative approach, based on modal parameter input, is described by Levin and Lieven [7]. They reduced the size of the FRF data by performing a modal analysis first. An RBF network was successful in detecting errors in a cantilevered beam. Marwala and Hunt [8] combined the previous two techniques and studied a clamped beam using a *committee* of multi-layer perceptron (MLP) network. They concluded that better fault identification results were obtained by the combined approach than the individual ones. However, the computation time was increased because of the need to train two networks, one for modal parameter input and the other for reduced FRF input. Because of possible modal analysis errors, it was recognized that different weightings could be allocated to the two networks.

As mentioned earlier, the examples above highlight the need to find a more compact representation of the measured FRFs. The selection of data points from frequency windows and the use of fewer spatial locations may lead to a loss of important information. The alternative approach, the modal analysis route, though compact, is unlikely to succeed in the cases where an accurate determination of modal parameters is fraught with difficulties: high modal density, high damping, non-linear effects, etc. Accordingly, the main objective of this paper is to present a method for an efficient and accurate reduction of the FRF data so that neural network techniques can be applied routinely to structural damage detection. A further objective is to demonstrate the feasibility of applying the methodology to a representative industrial structure.

2. PRINCIPAL COMPONENT ANALYSIS

2.1. BASIC THEORY

Principal component analysis (PCA), developed by Jolliffe [9] and Bishop [10], is a linear data compression technique which is widely used in the fields of image processing, flow visualization, pattern recognition and time-series prediction. An application to structural dynamics is due to Hasselman and Anderson [11, 12] who presented a theoretical basis dealing with the derivation of modal metrics for use in non-linear model correlation, updating and uncertainty evaluation of response time histories. They also developed a frequency-domain tool for vibro-acoustic response predictions over a wide frequency range.

Principal component analysis can be viewed as a statistical technique for achieving a dimensionality reduction. Using an orthogonal projection, the original set of variables in an N-dimensional space is transformed into a new set of uncorrelated variables, the so-called principal components (PCs), in a P-dimensional space such that $P < N$. Here, the basic theory for FRF data will be developed, though generalization to any other quantity is straightforward.

Using all available FRF data, let matrix $[H(\omega)]_{M\times N}$ be formed which has M rows of FRFs, each with N frequency points. A typical element is denoted by $h_{ij}(\omega)$, indices i and j indicating the position in the matrix.

The mean response of the jth column can be defined as follows:

$$\bar{H}_j = \frac{1}{M}\sum_{i=1}^{M} h_{ij}(\omega). \tag{1}$$

The corresponding standard deviation S_j can be defined as

$$S_j^2 = \frac{1}{M}\sum_{i=1}^{M}(h_{ij}(\omega) - \bar{H}_j)^2. \tag{2}$$

A typical element $h_{ij}(\omega)$ of the FRF matrix $[H(\omega)]_{M\times N}$ can now be replaced by

$$\tilde{h}_{ij}(\omega) = \frac{h_{ij}(\omega) - \bar{H}_j}{S_j\sqrt{M}}. \tag{3}$$

The application of equation (3) to each element in turn yields a response variation matrix $[\tilde{H}(\omega)]_{M\times N}$.

The correlation matrix can now be defined as

$$[C]_{N\times N} = [\tilde{H}]^{\mathrm{T}}_{N\times M}[\tilde{H}]_{M\times N}. \tag{4}$$

By definition, the principal components are the eigenvalues and associated eigenvectors of the correlation matrix:

$$[C]\{\Psi_i\} = \lambda_i\{\Psi_i\}, \tag{5}$$

where i is the principal component index.

The first principal component, i.e., the highest eigenvalue and its associated eigenvector, represents the direction and the amount of maximum variability in the original data. The next principal component, which is orthogonal to the first component, represents the next most significant contribution from the original data, and so on.

2.2. FRF RECONSTRUCTION

Since $[C]$ is a square matrix, the number of principal components obtained from equation (5) will be N. The projection of the response variation matrix $[\tilde{H}(\omega)]_{M\times N}$ on the N principal components is given by

$$[A]_{M\times N} = [\tilde{H}(\omega)]_{M\times N}[\Psi]_{N\times N}. \tag{6}$$

The projection matrix $[A]$ and the eigenvector matrix $[\Psi]$ can be partitioned into two sub-matrices with P principal components and $(N-P)$ principal components:

$$\begin{aligned}[A]_{M\times N} &= [[A_1]_{M\times P} \vdots [A_2]_{M\times(N-P)}],\\ [\Psi]_{N\times N} &= [[\Psi_1]_{N\times P} \vdots [\Psi_2]_{N\times(N-P)}].\end{aligned} \tag{7}$$

The response variation matrix can now be reconstructed for P components only:

$$[\tilde{H}_R] = [A][\Psi]^{\mathrm{T}} = [[A_1]_{M\times P} \vdots [A_2]_{M\times(N-P)}][[\Psi_1]_{N\times P} \vdots [\Psi_2]_{N\times(N-P)}]^{\mathrm{T}}$$

$$\approx [A_1]_{M\times P}[\Psi_1]^{\mathrm{T}}_{P\times N}. \tag{8}$$

Finally, element $\tilde{h}_{ijR}$ of the reconstructed response variation matrix is used to obtain element h_{ijR} of the reconstructed FRF matrix:

$$h_{ijR}(\omega) = S_j\sqrt{M}\tilde{h}_{ijR}(\omega) + H_j. \tag{9}$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_N$ be the ordered eigenvalues of the correlation matrix $[C]$. A measure of the maximum variation from the mean value can be computed as

$$J_v = \sum_{i=1}^{N} \lambda_i. \tag{10}$$

If the first $P < N$ principal components are used to reconstruct the FRF data, a measure of the reconstruction error is given by

$$J_e = \sum_{i=1}^{N} \lambda_i - \sum_{i=1}^{P} \lambda_i = \sum_{i=P+1}^{\mathrm{N}} \lambda_i. \tag{11}$$

The relative reconstruction error becomes

$$E = \left(\sum_{i=P+1}^{N} \lambda_i\right) \Big/ \left(\sum_{i=1}^{N} \lambda_i\right). \tag{12}$$

3. DETAILS OF THE NEURAL NETWORK

Artificial neural networks provide a general, non-linear parameterized mapping between a set of inputs and a set of outputs. A network with three layers of weights and sigmoidal activation functions can approximate any smooth mapping and such a type will also be used here. A typical supervised feed-forward multi-layer neural network, known as a BP neural network, is schematically illustrated in Figure 1.

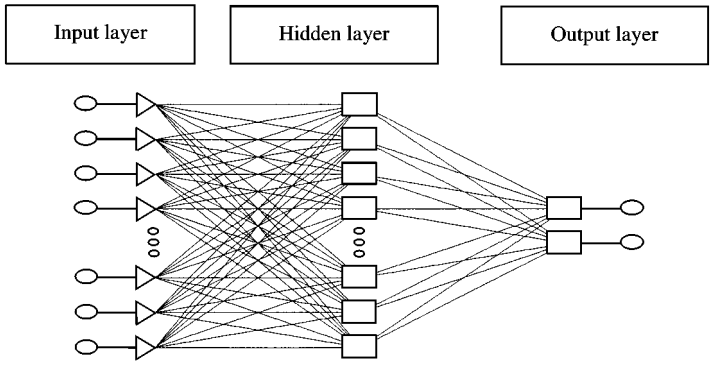

Figure 1. Typical BP neural network for damage detection.

The network consists of three types of layers:

- the input layer that receives the measured FRF data,
- the hidden layer which processes the data and
- the output layer that provides the result of the analysis, i.e., healthy or damaged.

The structural damage detection model is established through network training on known samples by the generalized delta learning algorithm. Based on s training samples, the learning algorithm is designed to minimize recursively an error function E_s of the form

$$E_s = \frac{1}{2}\sum_k (y_{sk} - o_{sk})^2, \tag{13}$$

where y_{sk} and o_{sk} are the desired output and the actual output vectors of the kth unit respectively.

Let the input and output vectors be denoted by $x = \{x_1, x_2, \ldots, x_n\}$, $o = \{o_1, o_2, \ldots, o_m\}$ respectively. The corresponding target output vector is $y = \{y_1, y_2, \ldots, y_m\}$ and $\{w_{ij}\}$ is the weight function between node i and j. A non-linear sigmoidal function f can be defined as

$$f(x) = \frac{1}{1 + \mathrm{e}^{-x}}. \tag{14}$$

The output of the kth node in the hidden and the output layers can be described by

$$o_k = f(net_k) = f\left(\sum_j w_{kj} o_j\right), \tag{15}$$

where the net_k is the input of the kth node.

The interconnection weights, adjusted in such a way that the prediction errors on the training set can be minimized, are given by

$$\begin{aligned} &\Delta_s w_{ji} = \eta \delta_{sj} o_{si}, \\ &\delta_{sj} = (y_{sj} - o_{sj})(1 - o_{sj}) o_{sj} \quad \text{(when node } j \text{ is in the output layer)}, \\ &\delta_{sj} = o_{sj}(1 - o_{sj}) \sum_k \delta_{sk} w_{kj} \quad \text{(when node } j \text{ is in the hidden layer)}, \end{aligned} \tag{16}$$

where $0 < \eta < 1$ is the learning rate coefficient, Δ is the actual change in the weight and δ is the node error.

In order to control the network oscillations during the training process, a momentum coefficient $0 < \alpha < 1$ is introduced to the definition of the weight change:

$$\Delta_s w_{ji}(k+1) = \eta \delta_{sj} o_{sj} + \alpha \Delta_s w_{ji}(k). \tag{17}$$

Once the change is computed, the new weight is given by

$$w_{ji}(k+1) = w_{ji}(k) + \Delta w_{ji}(k+1). \tag{18}$$

The training of a BP neural network is a two-step procedure. In the first stage, the network propagates input through each layer until an output is generated. The error between the

actual output o_{sk} and the target output y_{sk} is then computed using equation (13). In the second stage, the calculated error is transmitted backwards from the output layer and the weights are adjusted according to equations (17) and (18) in order to minimize the error. The training process is terminated when the error E_s is sufficiently small for each training sample.

4. CASE STUDY: RAILWAY WHEEL

4.1. PRELIMINARIES

The methodology above was validated using simple numerical test cases based on finite element models of beam and plate structures. Very good results were obtained in all the cases studied. It was then decided to investigate the more taxing case of a railway wheel. Railway wheels are subjected to random dynamic loads during their operation and hence the initiation of fatigue cracks is relatively common. For the particular type under study, circumferentially propagating cracks are known to occur at about $\frac{1}{3}$ radius and these can reach several centimetres in length. For both healthy and damaged specimens, several FRFs, each containing 4096 spectral lines, were measured in the x, y and z directions (Figures 2–4).

Although a visual inspection reveals that there are differences between the FRFs of the healthy and damaged specimens, an objective assessment of the actual damage state is not straightforward. Although there is a general trend for the healthy specimen FRFs to have fewer, better defined modes with higher response levels, there are significant variations from one FRF to the next. The approach proposed here is to design and train neural networks for damage detection. However, the size of the FRF data, both the number of measurements and 4096 spectral lines in each measurement, is prohibitive for a direct network use and a data reduction process is necessary. Due to high modal density and significant

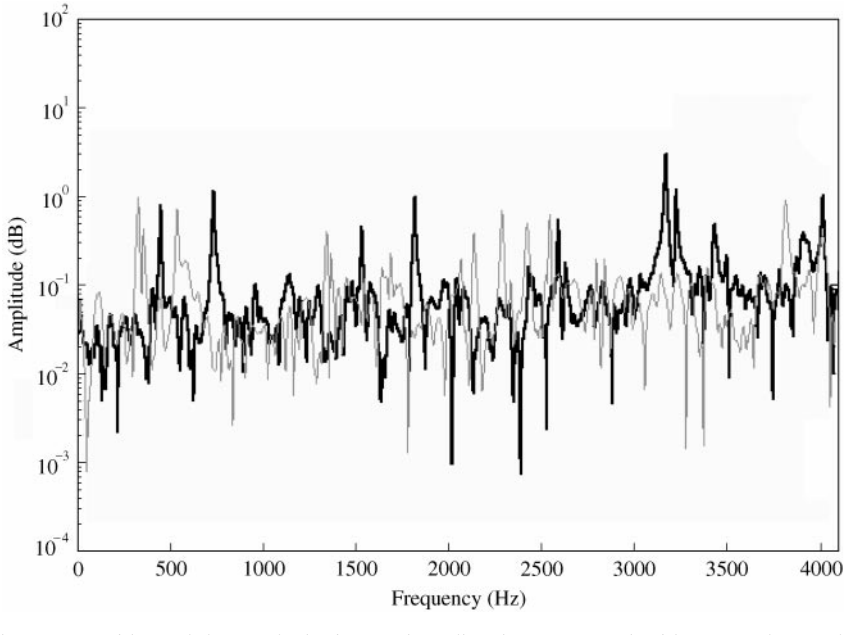

Figure 2. Healthy and damaged wheel FRFs in x direction: ———, healthy; ——, damaged.

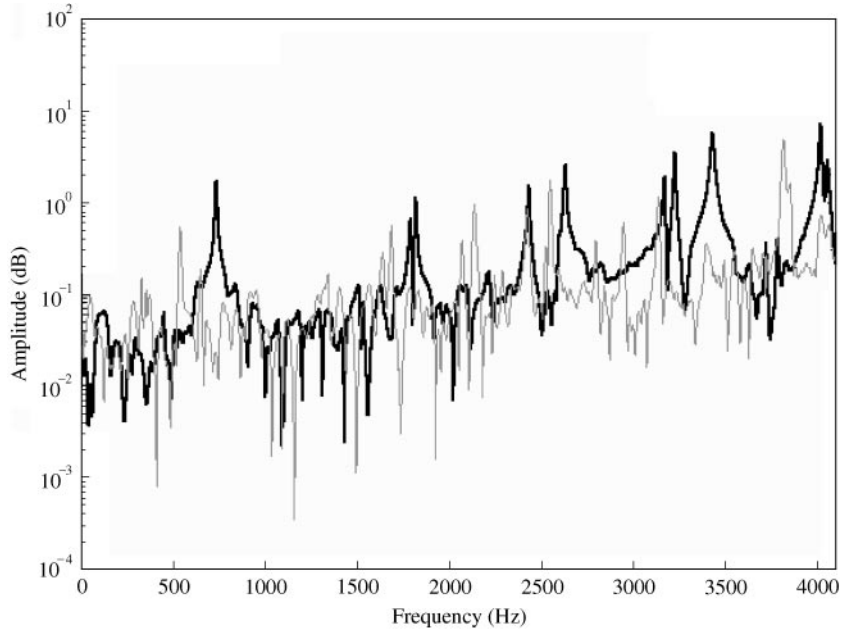

Figure 3. Healthy and damaged wheel FRFs in y direction: ———, healthy; ——, damaged.

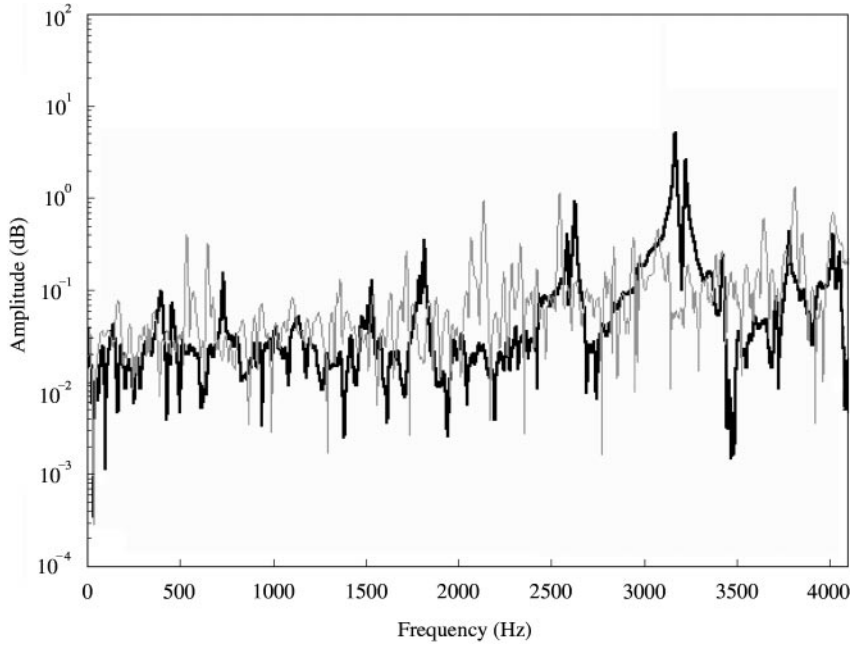

Figure 4. Healthy and damaged wheel FRFs in z direction: ———, healthy; ——, damaged.

measurement noise, a modal analysis is unlikely to yield accurate enough system parameters. Similarly, the selection of key spectral points to represent frequency windows is not straightforward because of the large variations in the dynamic behaviour. Therefore, the principal component reduction of the FRF data appears to be the only way forward.

TABLE 1

Principal component eigenvalues together with relative and cumulative percentages

No.	Principal component eigenvalue	Percentage in sum of all eigenvalues	Cumulative percentage
1	1753·0	42·80	42·80
2	886·8	21·65	64·45
3	468·4	11·44	75·89
4	320·1	7·81	83·70
5	209·2	5·11	88·81
6	141·6	3·46	92·27
7	100·3	2·45	94·72
8	76·4	1·87	96·59
9	41·3	1·01	97·60
10	17·6	0·43	98·03
20	2·5	0·06	99·30
30	1·1	0·03	99·69
40	0·5	0·01	99·85
50	0·2	0·01	99·95

4.2. PRINCIPAL COMPONENT ANALYSIS

Using equations (1)–(4), three correlation matrices were formed for each of the x, y and z directions. Each matrix had 80 rows consisting of 40 FRFs for the healthy wheel, and another 40 for the damaged one. The number of columns is equal to the number of spectral lines, here 4096.

The principal component eigenvalues are listed in Table 1, together with their relative and cumulative percentages. It is observed that the first 10 eigenvalues account for 98·03% of the variance of the response about the mean FRF values. Therefore, the reconstruction of the response using 10 principal components would represent a reduction of $4096:10 = 4096:1$ and the error would be 1·97%.

The FRFs of the healthy wheel were reconstructed using 1, 5 and 11 principal components and the results are plotted in Figure 5 together with the original ones. As expected, the match between the original and reconstructed FRFs improves with increasing number of principal components. Here, the use of 11 principal components (cumulative percentage 98·26%) was considered to be adequate. A very similar trend was observed for damaged wheel FRFs, also plotted in Figure 5.

The most significant principal components contain those features which are dominant in most of the frequency responses. It follows that random noise, which is not correlated with such global features, will be represented by the less significant components[†]. Thus, reconstructing the response by using the highest principal components only, should achieve not only data compression but also remove some of the noise. For the case of 50 principal components, the original, the reconstructed FRFs and their absolute difference is plotted in Figure 6.

4.3. IMPLEMENTATION OF THE NEURAL NETWORK

As mentioned earlier, available FRF data were grouped into three correlation matrices of order 80×4096. The selection of all 4096 spectral lines to define the input vector to the

[†] The argument probably does not apply to systematic errors caused by, for instance, a shaker misalignment.

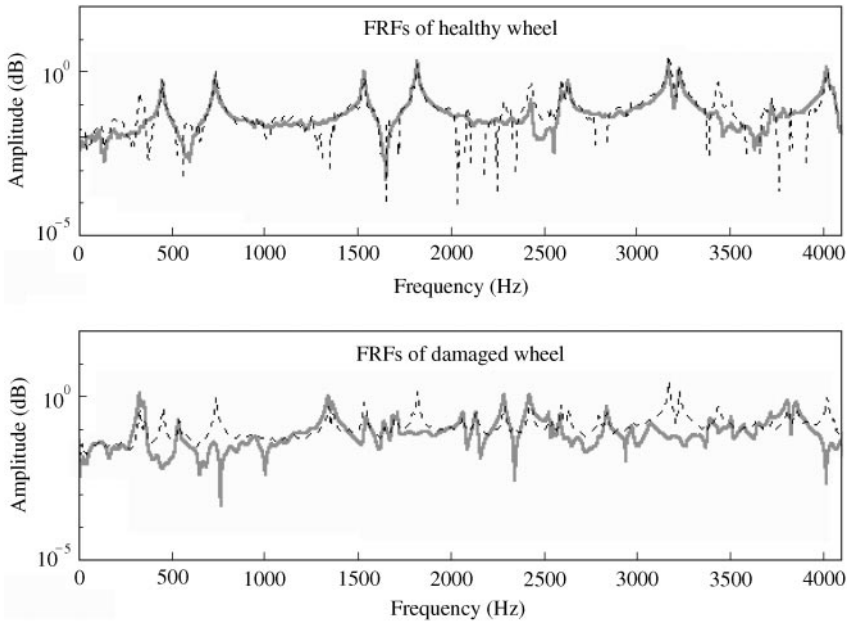

Figure 5a. Original (solid line) and the reconstructed FRFs (one principal component).

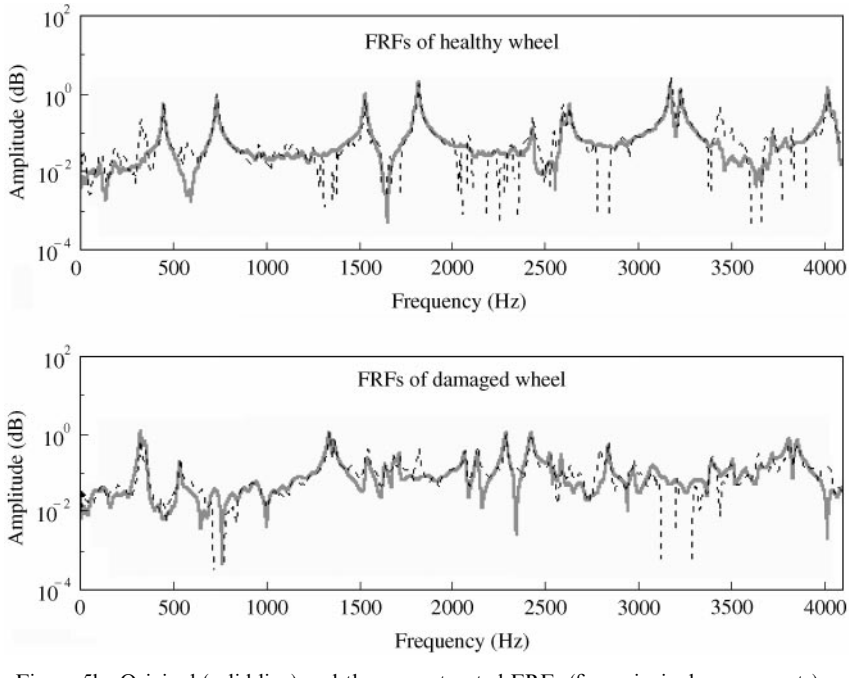

Figure 5b. Original (solid line) and the reconstructed FRFs (five principal components).

neural network is clearly impracticable. Even if the training time and convergence-related problems could be overcome by parallel processing on fast CPU arrays, network over-fitting would still remain a major obstacle. When the number of variables is much greater than the number of training samples, neural networks can focus on local details of

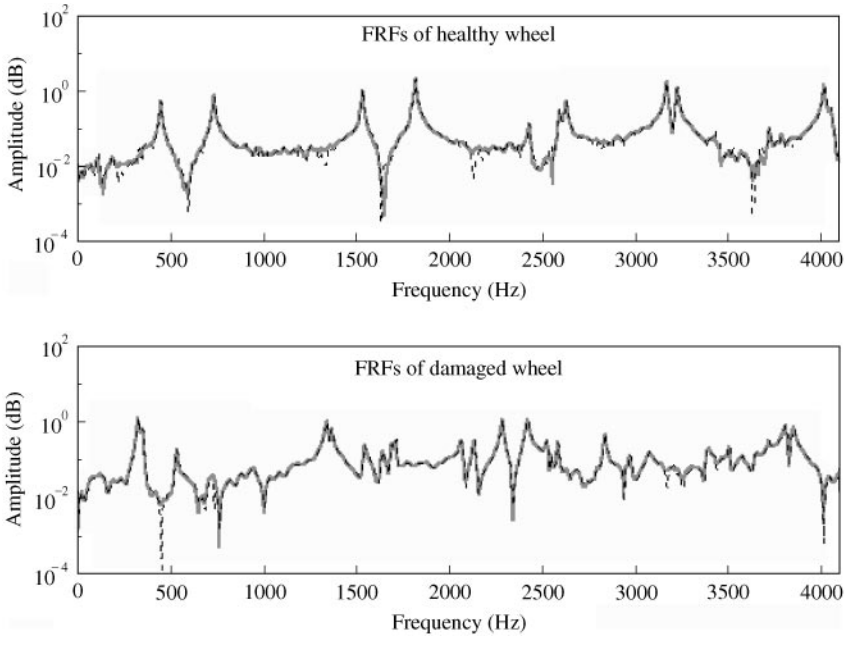

Figure 5c. Original (solid line) and the reconstructed FRFs (11 principal components).

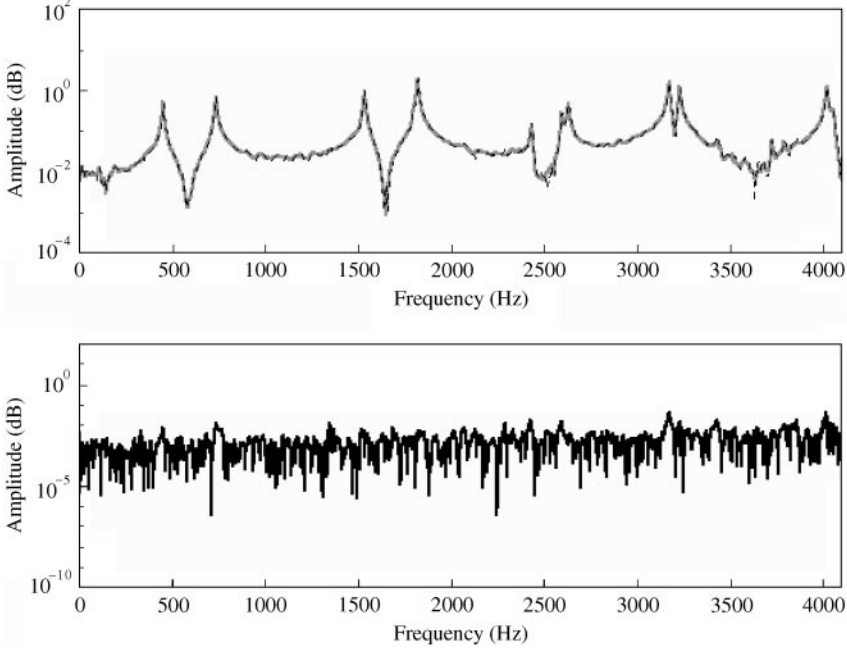

Figure 6. Original (solid line) and reconstructed (broken line) FRFs (upper plot) together with their difference (lower plot), representing measurement noise (50 principal components).

individual training samples which may well be meaningless in a global context. In such instances, a very large number of training samples may be needed, a feature that will be illustrated below.

The relationship between the number of training samples s, the number of input variables n, the number of output variables m and the number of hidden layer nodes h in hidden layer can be expressed as [12, 13]

$$s = 1 + h(n + m + 1)/m. \tag{19}$$

Here, there are $m = 2$ output variables since the specimen is either healthy or damaged, and $n = 4096$ input variables, each representing a spectral line of the corresponding FRF. Meanwhile, $h = 2049$ hidden nodes are needed to create the BP network. In this case, the required number of training samples becomes

$$s = 1 + 2049(4096 + 2 + 1)/2 \approx 4\,199\,427. \tag{19a}$$

In other words, about 4·2 million samples are needed to make the network numerically determined. Such a requirement cannot be met when dealing with real applications. The principal component analysis was applied to each of the $80 \times 4\,096$ FRF matrices to achieve an acceptable data reduction level. It was decided to retain 95% of the variance about the mean response values. Such an approach led to using seven principal components in the x direction, nine in y direction, and 13 in the z direction. Using equations (8) and (9), the original FRF data array of 80×4096 was reduced to 80×7 for the x direction, 80×9 for the y direction and 80×13 for the z direction. The reduced data sets were used as input vectors to three different neural networks, one in each of the x, y and z co-ordinate directions. The procedure is summarized in Figure 7.

For each neural network, the input consists of the PCA-compressed FRFs while the output is either [1, 0] (healthy) or [0, 1] (damaged). Although it is possible to use 1 output node only, 2 nodes were used to achieve better non-linear mapping and to improve the

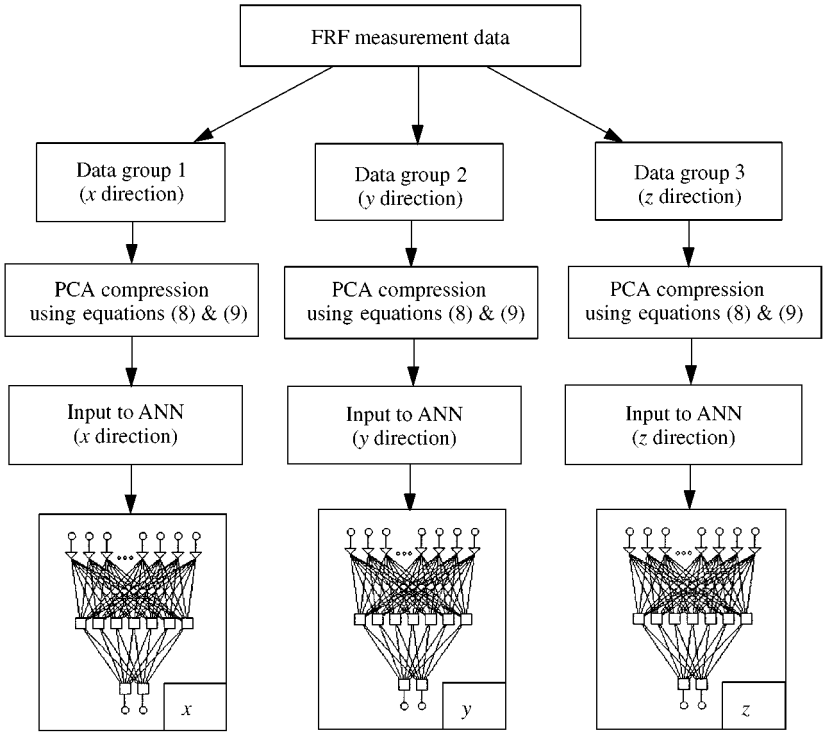

Figure 7. Preparation of FRF data for neural network input.

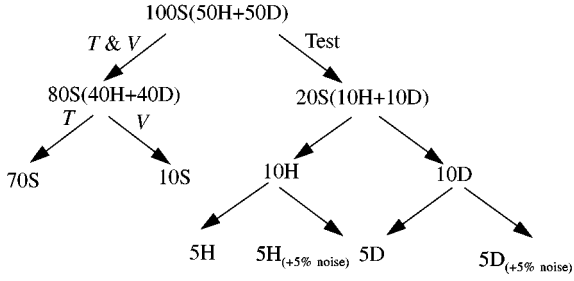

Figure 8. Data samples for training, verification, and testing in each group: S—samples, H—samples for healthy state, D—samples for damaged state, T—training, V—verification.

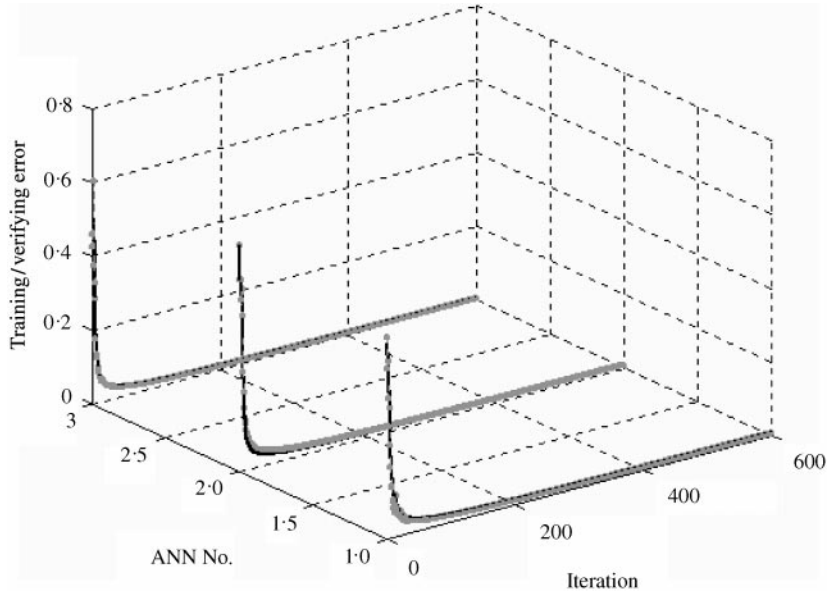

Figure 9. Training (lighter line) and verification (darker line) errors for the three neural networks.

damage detection resolution. In the general case, even more output nodes can be used for damage location but this will increase the number of training samples.

The training, verification and sample testing paths of the three neural networks are shown in Figure 8. Out of the 100 available FRFs, 80 were used for network training and verification. The learning and momentum rates were fixed at 0·6 and 0·3 respectively. As will be discussed later, the remaining 20 were used for the damage detection tests.

The error variation during the training and verification iterations is plotted in Figure 9. As can be seen from Figure 9, both the training and verification errors drop sharply after 50 iterations for all three networks. The RMS values of training and verification errors after 600 iterations are listed in Table 2. The results indicate that all three networks are stable and well trained. The output from the x direction network is plotted in Figure 10. It is easily seen that all samples represented by their x direction FRFs are classified correctly, both for training and verification.

TABLE 2

Network architecture and training/verification errors (I: input, H:hidden, O:output)

Network	NN structure	Training error (%)	Verification error (%)
x direction	7I-4H-2O	0·70	0·63
y direction	9I-5H-2O	0·56	0·86
z direction	13I-7H-2O	0·58	0·55

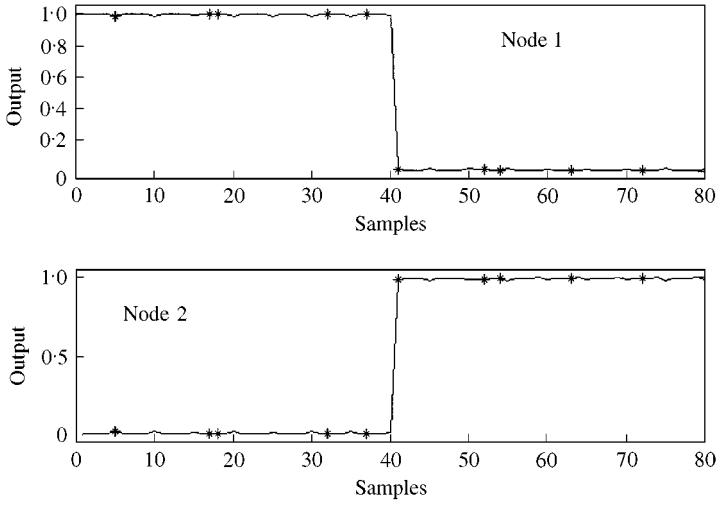

Figure 10. *x* direction network output for Nodes 1 and 2: Training sample (solid line), verification sample (asterisk).

4.4. DAMAGE DETECTION

After successful training and verification of all three networks, three sets of 20 FRFs, 10 FRFs from a healthy specimen and 10 FRFs from a damaged specimen, were selected in each of the *x*, *y* and *z* directions. The FRFs in each set were fed sequentially into the appropriate network, thus creating 20 tests in each co-ordinate direction. The output from all three networks is listed in Table 3. Half of the FRFs were polluted by adding 5% random noise and these correspond to Tests 6–10 and 16–20 in Table 3. In spite of added noise, all tests are seen to be successful, the output from Nodes 1 and 2 yielding a clear distinction between healthy and damaged specimens. It should be noted that robustness to noise is, in the main, provided by the PCA compression which acts as a filter (Figure 6). When using the noise-polluted FRFs, the truncation levels for PCA compression were kept the same as in section 4.3, namely 7, 9 and 13 principal components in the *x*, *y* and *z* directions.

5. CONCLUDING REMARKS

(1) The results indicate that the combination of FRF data reduction via the principal component analysis and the use of artificial neural networks provides a suitable methodology for damage detection. Once the size problem is circumvented, the training of the neural networks with FRF data from several damage configurations/specimens is

TABLE 3

(a) Output from the x direction network

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
Node 1	0·998	0·998	0·999	1·000	0·994	0·998	0·998	0·999	1·000	0·992
Node 2	0·000	0·001	0·002	0·000	0·007	0·000	0·002	0·004	0·000	0·011
	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
Node 1	0·002	0·002	0·001	0·000	0·006	0·002	0·002	0·001	0·000	0·008
Node 2	1·000	0·999	0·998	1·000	0·992	1·000	0·998	0·996	1·000	0·989

(b) Output from the y direction network

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
Node 1	0·995	0·996	0·995	0·995	0·995	0·980	0·979	0·976	0·995	0·986
Node 2	0·006	0·004	0·004	0·004	0·005	0·023	0·018	0·019	0·005	0·011
	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
Node 1	0·002	0·002	0·003	0·004	0·004	0·002	0·002	0·003	0·062	0·003
Node 2	0·998	0·998	0·997	0·996	0·996	0·998	0·998	0·996	0·923	0·996

(c) Output from the z direction network

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
Node 1	0·996	0·997	0·997	0·995	0·996	0·995	0·997	0·997	0·994	0·997
Node 2	0·003	0·003	0·004	0·005	0·005	0·004	0·003	0·004	0·006	0·004
	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
Node 1	0·008	0·006	0·008	0·011	0·007	0·018	0·011	0·012	0·017	0·009
Node 2	0·991	0·993	0·992	0·989	0·994	0·979	0·989	0·989	0·984	0·993

relatively straightforward. The routine availability of measured frequency response data and relatively modest computational requirements make the method well suited to on-line industrial applications.

(2) The extension of the methodology to damage location is somewhat more difficult. A fine spatial resolution is needed for damage location, a requirement that increases the number of input nodes. Similarly, significantly more output nodes will also be needed to monitor the state of various locations on the structure simultaneously. The increase in the input and output nodes will inevitably require more training samples, making data compression even more critical. In any case, preliminary studies with combined ANN methods and using an alternative approach based on self-organizing Kohonen maps led to some limited success in damage location and further work is in progress.

(3) The principal component analysis is a powerful tool for reducing the size of measured frequency response data. Such a reduction has clear benefits over data reduction via point selection or modal analysis. Even when a relatively small number of principal components is used, the PCA-compressed FRFs appear to retain most of the original information. This feature is likely to be suited to other signal processing applications.

(4) The principal component analysis has the potential for filtering unwanted measurement noise. Consequently, it is, in principle, possible to pre-process measured FRF data for better modal analysis. Such a route will be explored in a forthcoming paper.

ACKNOWLEDGMENTS

The financial support of the European Commission under the Brite/EuRam Programme (Contract No: BRPR-CT98-0688 AMADEUS) is gratefully acknowledged. The authors also wish to thank Centro Ricerche Fiat for supplying the measured FRF data.

REFERENCES

1. D. E. Rumelhart, G. E. Hinton and R. J. Williams 1986 *Parallel Distributed Processing*: *Explorations in the Microstructure of Cognition*; *Vol.* 1: *Foundations*, (D. E. Rumelhart and J. L. McClelland, editors). Cambridge, MA: The MIT Press. Learning internal representations by error propagation.
2. S. W. Doebling, C. R. Farrar, M. B. Prime and D. W. Shevitz 1996 *Los Alamos National Laboratory Report No. LA*-13070-*MS*. Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review.
3. P. Zeng, 1998 *Applied Mechanics Reviews* **51**, 173–197. Neural computing in mechanics.
4. X. Wu, J. Ghaboussi and J. H. Garrett 1992 *Computers and Structures* **42**, 649–659. Use of neural network in detection of structural damage.
5. Z. Chaudhry and A. J. Ganino 1994 *Journal of Intelligent Material Systems and Structures* **5**, 585–589. Damage detection using neural networks — an initial experimental study on debonded beams.
6. M. J. Atalla and D. J. Inman 1998 *Mechanical Systems and Signal Processing* **12**, 135–161. On model updating using neural networks.
7. R. I. Levin and N. A. J. Lieven 1998 *Journal of Sound and Vibration* **210**, 593–607. Dynamic finite element model updating using neural networks.
8. T. Marwala and H. E. M. Hunt 1999 *Mechanical Systems and Signal Processing* **13**, 475–490. Fault identification using finite element models and neural networks.
9. I. T. Jolliffe 1986 *Principal Component Analysis*. New York: Springer-Verlag.
10. C. M. Bishop 1995 *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.
11. T. K. Hasselman and M. C. Anderson 1998 *Proceedings of the 16th International Modal Analysis Conference*, *Santa Barbara*, *CA*, 644–651. Principal components analysis for non-linear model correlation, updating and uncertainty evaluation.
12. T. K. Hasselman and M. C. Anderson 1998 *Proceedings of the 16th International Modal Analysis Conference*, *Santa Barbara*, *CA*, 1285–1291. Linking FEA and SEA by principal components analysis.
13. C. Zang 1994 *Ph.D. Thesis*, *Southeast University*, *MED*, *China*. Studies on theory and method of monitoring & diagnosis for rotating machinery.