Automatic Vibration Analysis
A Neural Network Approach

Mathew J. Boek[1], Andre S. Szczepanik[2] and Jacob L. Cybulski[3]

[1,2] Vimac Pty. Ltd.
10 Lynden Gve
Mt Waverley, Vic 3149, Australia
Ph: (03) 888 2207, Fax: (03) 807 4368

[3] Dept of Comp. Sci. and Eng.
La Trobe University
Bundoora, Vic 3083, Australia
Ph: (03) 479 1270, Email: jacob@latcs1.lat.oz.au

SUMMARY

The benefits of assessing machine operating condition with vibration signature analysis techniques are well established. Current methods require vibration analysis to be performed by a skilled engineer, and are generally performed off line. By viewing the vibration analysis task as a pattern classification problem, automatic techniques for interpreting the vibration signal can be employed. An advantage of the automatic approach is that machine condition can be continuously monitored. An automatic classification system for detecting and diagnosing certain rotor defects in rotating machinery based on *neural networks* is described. The system consists of data preprocessor, a feature extractor, and a multi-layer perceptron neural network classifier. . From the findings of experiments performed to test the system on classifying impeller faults on a desktop fan, it can be concluded that the neural network classifier methodology can act as a successful classifier of vibration spectra, if practical considerations are taken into account in the performance evaluation, and appropriate features are selected to represent the vibration signal.

1. INTRODUCTION

Modern machine condition monitoring is based on the analysis of machine vibration. It involves the collection and analysis of vibration signals to identify patterns of signal features which indicate deterioration in machine condition. Traditionally this procedure is conducted with the use of either a tape recorder or a portable data collector to record data to be later analysed off-line; we have developed software supporting this traditional approach and have gained experience in its application to specific industrial problems. More sophisticated systems allow on-line, real-time data collection and analysis; currently we are developing such a system.

The technique involves the following basic steps (areas that could potentially benefit from the application of "intelligent" information processing are outlined):

- **Specification** of the fundamental characteristics of the machine and its operating environment. (This step could use a machine type database, supporting type classification and machine instantiation with inheritance. Such machine organisation could greatly reduce the initial effort of machine specification).
- **Baseline Signature Recording** of a collection of measurements which can be used as a reference in all future data analysis.
- **Recording** of vibration signals
- **Feature Analysis** of signal features which are calculated from the collected vibration data. (This could be aided with the use of an intelligent data processor or expert system, which could generate the features automatically, independently of the operating environment of the machine).
- **Condition Analysis** of the collected signal features. Condition is usually assessed in terms of component wear, risk assessment, prediction of time to failure, or impact of current machine condition on product quality. (This step is the one which requires the most intelligence, and is the area targeted by the system described in this paper).
- **Recommendations** are produced to assist in the scheduling of machine maintenance. (An expert system, combined with a database of previous recommendations and machine condition, could aid in producing a new set of recommendations for the current situation).

In practice the association between complex patterns of signal features and fault conditions developing in machines is usually performed by highly skilled mechanical engineers. Their experience cannot be easily explained in terms of fundamental engineering principles or be transliterated into a set of written procedures which could become the basis for writing a computer program to be used in monitoring tasks. Nevertheless, modern expert system techniques provide tools, methods and techniques for capturing human expertise (McGraw and Harbinson-Briggs 1989), which could be used for building fully automatic on-line monitoring systems (Milne 1990) or off-line fault detection, identification and rectification systems (Bannister, R.H. and Moore, M.P. 1986). Neural networks (Rumelhart and McClelland 1986) offer an alternative approach to expert system knowledge acquisition techniques; they offer the additional advantages of an ability to learn implicit facts from examples and the simplicity of their implementation.

2. NEURAL NETWORKS FOR MACHINE CONDITION MONITORING

The condition analysis phase of machine monitoring can be viewed as a problem where a collection of highly correlated (and in some cases redundant) features of vibration signals could be associated with a number of machine conditions. The goal is to *recognise* incoming *patterns* (corresponding in this case to patterns of features derived from vibration signals) as belonging to one of a number of known classes (in this case machine states of operation).

Pattern recognition is often associated with the concept of a *classifier*, whose task is to identify the general class that a pattern belongs to. This corresponds to a *classification model* of pattern recognition (Pao, 1989) which consists of the following steps:

- Finding a set of features which describe a particular object or class
- Learning a mapping which maps the pattern of feature values of a particular instantiation of a class onto the class itself, using a set of labelled examples
- Classifying patterns of feature values using the learned mapping.

Many issues must be considered in the choice of a classifier: development and run-time costs, in terms of both time and money, generality, flexibility, complexity and the mode of processing (on-line or off-line). As is often the case, there are trade-offs that must be evaluated. Perhaps the most general and flexible automatic classifier is a well-designed knowledge-based (or expert) system, but development costs in terms of knowledge acquisition are generally high.

Another type of classifier learns classification rules automatically from a set of labelled training data. The generality and flexibility of this type of classifier depend on the learning algorithm employed, as well as the quality and size of the labelled data set. The principal advantage of such a classifier is that the cost of its development depends on how expensive it is to collect the data set; once done, the learning takes place with little or no intervention. Many different techniques can be employed to perform this type of classification, including various types of neural networks. The use of conventional statistical classification techniques for vibration signature analysis has been demonstrated (Hoffman and Fukunanga, 1971; Gersch, 1986).

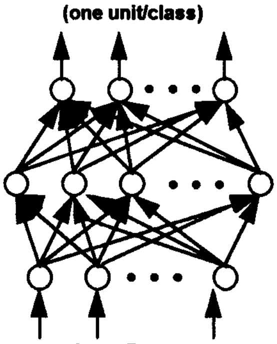

Figure 1 - Multi Layer Perceptron Architecture

Neural networks featured prominently in the literature of pattern recognition in the 1960s, until it was shown that they could not solve certain fundamental problems (Minsky and Papert, 1969). Much interest has recently been generated in using various neural network models for pattern classification. This is to a large extent due to the discovery of new and powerful learning algorithms. Recent applications of neural networks include sonar signal classification (Gorman and Sejnowski, 1988), and speech recognition (e.g. Waibel, Hanazawa, Hinton, Shikano and Lang, 1987).

Many neural network models can perform pattern classification, but by far the most often applied model is the *Multi-Layer Perceptron (MLP),* using the *Back-Propagation* learning algorithm (Rumelhart, Hinton and Williams, 1986). The basic architecture of this network consists of an *input layer* of processing *units*, connected to another, *hidden* layer of units, which can either connect to the output layer, or another hidden layer. There are no intra-layer connections, and no direct feedback from higher to lower layers; this type of network is sometimes called a *feedforward* network for this reason (Cf. figure 1).

Each processing unit receives a number of numeric inputs, and produces an output (or *activation*) which is a non-linear function (Cf. equation 2) of the weighted sum of the inputs (Cf. equation 1). The weights attached to each input can vary - it is through the variation in these weights that the network is able to learn.

The formula for calculating the net input to unit *i* is:

$$net_i = \sum_j act_j w_{ij} + bias_i \qquad (1)$$

where j ranges over all units connected to unit i, w_{ij} is the weight on the connection to unit i from unit j, act_j is the activation of unit j, and $bias_i$ is a *bias* for the unit, which can be considered as a weight from a unit that is always fully active. The bias term allows a unit to have a non-zero activation even if all weights and/or inputs from other units are zero.

The formula for calculating the output of unit *i*, given that its net input is net_i is:

$$act_i = \frac{1}{1+e^{-net_i}} \qquad (2)$$

Patterns are presented in the form of a vector of numbers to the input layer units. The input layer units compute their activation, and the hidden layer units compute their activation from the input layer activation. This continues until the output layer units compute their activation.

The task of network *training* is to set the values of the weights for each unit so that the network as a whole produces the desired output. In pattern recognition applications, the available data is usually divided into a training set and a test set. The training set data is used to train the network using a learning algorithm, and then the recognition performance of the network is tested using the data in the test set. For back-propagation, the training procedure generally is as follows:

1. Set all weights to small random values
2. Present a pattern to the input layer, and allow the network to compute its output
3. Compare the desired output with the actual output, and modify the weights of the output layer units proportionally to the magnitude of this error.
4. Modify the weights to the hidden layer units. Back-propagation works by distributing blame to a hidden unit based on the weight on the link between that hidden unit and an output layer unit, and the error made by that output unit. A measure of error for a hidden unit is found by multiplying the errors of all the output units that the hidden unit connects to, by the weights on the respective links, and summing the results.
5. Repeat steps 2 to 4 until all patterns in the training set have been presented.
6. Test the network using the patterns in test set. If the error of the network is still too high, repeat steps 2 to 5.

Several important issues must be considered in the application of neural networks. An extensive discussion of these is beyond the scope of this paper, however some are outlined below:

- *Determining Hidden Layer Configuration*
  It is clear how many units should be in the input layer and output layer, but it is not obvious how many units to use in the hidden layer(s).
- *Choosing Learning Parameters for Back-propagation*
  The formulae for back-propagation include a number of parameters governing essentially the amount of change to apply to weights on each presentation of the training data set. Too large a value, and the network may oscillate wildly about the solution space; too small a value and the network may take an unacceptably long time to find a solution.
- *Determining when to stop Training*
  This is an important issue, as too much training could result in the network becoming biased to the examples in the training set so that performance is excellent for the training data but poor on the test data. This phenomenon is known as *overfitting*.

3. FAN EXPERIMENTS

An experimental automatic classification system for detection and diagnosis of certain defects in rotating machinery was proposed and implemented with the use of neural networks (Boek, 1992). The system was tested using data generated by simulating faults (load imbalance, shaft imbalance, and a cracked blade) commonly found in heavy industrial fans by attaching small weights to the fan blades or cutting a blade on a 3-speed, 3-blade oscillating desktop fan.

Two accelerometers were attached to the fan, one in the radial and one in the axial plane of the shaft (cf. figure 3).

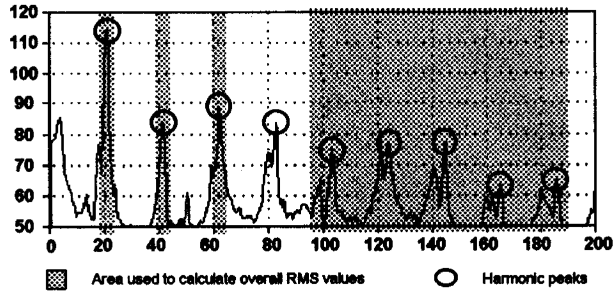

Figure 2 - Spectral feature extraction

The fan was run at a constant speed, and the signals from the accelerometers were processed by a charge amplifier, which fed the modified signals to a signal analyser. The analyser digitised the signal, transformed it from a time signal into a frequency spectrum, averaged a number of such spectra, and then passed the averaged spectrum to condition monitoring software (the M system, developed by Vimac Pty Ltd) for feature extraction, storage and analysis. The set of extracted features was then presented as input to the neural network classifier (based on the PDP package - McClelland and Rumelhart, 1988) for training and assessment of machine condition (cf. figure 3).
In our experiments the spectrum consisted of 400 values, which ranged between 0 and 130 dB. As every state of the fan was represented by two spectra, one from the vertical plane of vibration and one from the axial plane, a total of 800 values were available to describe each snapshot of fan condition. It is known, however, from reference to engineering principles that for rotating machines (such as a fan) these values consist mostly of noise. It is also known that the vibrational information important for detecting rotor defects is carried in the amplitudes of the rotational frequency and its associated harmonics. Thus, some of our experiments used only the peak values of the harmonics of the rotational frequency as a representation of the fan condition. This reduced the dimensionality of the input space from 800 to just 18. All spectral peaks extracted were scaled from the range 0..130 to a value between 0 and 1.

In an effort to further reduce the dimensionality of the input space, in our experiments we used the first 3 harmonics, with the higher harmonics summarised by a single *overall* value (the average RMS value of the high frequency end of the spectrum); see figure 2. Although dimensionality

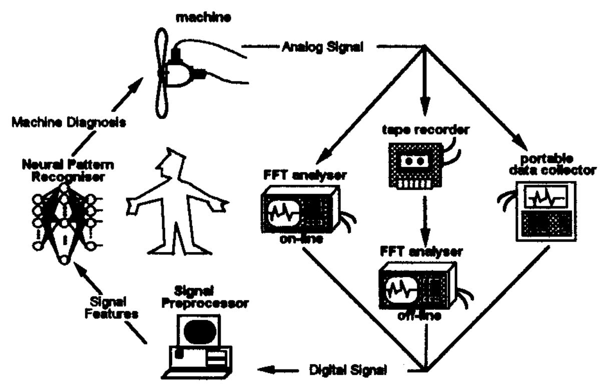

Figure 3 - Neural network vibration signal classification

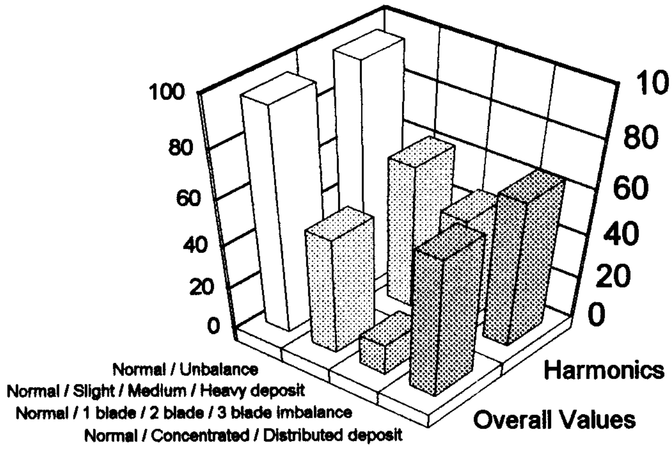

Figure 4 - Average test set performance

reduction was achieved, it was possible that some potentially important information was lost. As it is virtually impossible to determine in advance the precise feature values that characterise a specific machine condition the features selected for our experiments were determined by trying different combinations of harmonics, and observing which appeared to vary the most for different fan conditions.

Training was conducted using the back-propagation algorithm (see above). The neural network used for this project had one hidden layer containing between 5 and 20 processing units.

The performance of the network was tested on five classification tasks: detection of imbalance, classifying imbalance into subtypes (deposit on 1/2/3 blades, slight/medium/heavy or concentrated/distributed deposit) and distinguishing between imbalance and a cracked impeller blade. The average test set result across all classes for the first four tasks are presented on figure 4. As may be expected, the best results were obtained in the detection of imbalance, as this simple task could also be done by a simpler system than a neural network. In general, the results of the experiments to classify type of imbalance were quite poor, although in most cases the network detected that imbalance was present, but could not distinguish one type of imbalance from another. In order to put the results for these experiments in perspective, it should be noted that some of the classifications being attempted were of questionable practical importance; it is far more important to know that there is imbalance of some sort present in a machine than to know for example that the deposit is on one or two blades. This poor performance may also be explained by the limited number of examples available for training and testing the network. Good results, however, were obtained when the network had to distinguish between physically different faults, i.e. imbalance vs a crack in the blade. The good results for these experiments are the most important, as they show that a multi-layer perceptron neural network trained with back propagation is capable of developing rules to classify the complex signals provided by vibration spectra (with suitable preprocessing and feature extraction) to certain level of accuracy. The fact that the network coped successfully with data generated from a real fan, together with all the noise that is unavoidable in such an environment, also confirms to a certain degree the much heralded noise tolerance of neural pattern classification systems.

These experimental results should translate well to fault detection in and diagnosis of industrial fans, since such fans are more rigid in structure, and thus tend to produce less noise in the vibration signal. The transducers used in collecting the data for these experiments were attached to the plastic housing of the motor, and the impeller itself was made of flexible plastic. This flexibility introduces more noise into the signal, and thus in some sense the noise could be seen as "worst case" noise in these experiments.

A study was performed to compare the performance of an experienced mechanical engineer on the same data sets that were used for training and testing a network (figure 5 shows the results on the test set). This was done for the task in which the network had to classify the normal, imbalance, cracked impeller, and both-faults-together states of the fan. The human expert clearly outperformed the network for the cracked blade and both-faults-together classes. However, a network using overall value spectral representation obtained almost equivalent performance to that of the expert for the normal and imbalance classes. It should be noted that this result was obtained by the expert having access to the entire spectra of a faulty and a good machine condition as opposed to the restricted set of features used by the neural network. Further work in extracting better features could lead to improved performance by the network.

As the automation of machine condition analysis significantly reduces the dependence of monitoring tasks on vibration analysts, the proposed method is being investigated by the authors for the development of a system for continuous, on-line monitoring of machine vibrations. The software, currently under development, will be capable of rapid classification of vibration signals into a set of high-level messages and alarms associated with the current machine condition, process parameters and product quality. Due to the fact that vibration is often the very first sign of malfunction, the system will give the process operators an opportunity to detect, accurately assess and quickly correct

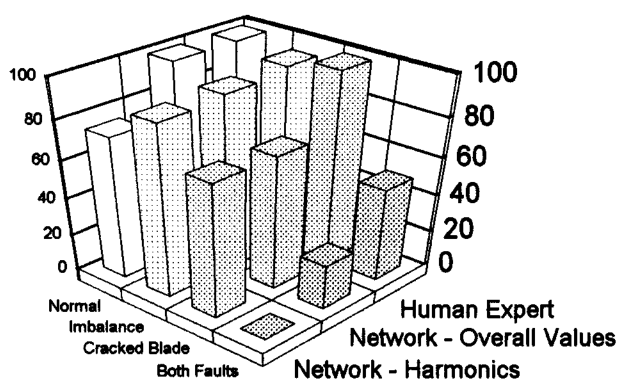

Figure 5 - Neural network vs human expert

deficient manufacturing processes well before conventional monitors could register a fault developing in the machine, process or product.

4. CONCLUSIONS

A neural network can act as a successful classifier of machine vibration spectra, if practical considerations are taken into account in the performance evaluation, and appropriate features are selected to represent the vibration spectra. This conclusion is supported by the findings of several experiments performed to detect and classify impeller faults on a desktop fan. The network's performance was comparable with that of a human expert for the same problem.

The principal benefit of using neural networks to perform the automatic detection and diagnosis of faults is that the system can learn automatically the criteria required to correctly perform this task. This is an important consideration, as the reason why expensive man-hours are required to perform machine condition monitoring is that vibration is a complex phenomenon. The vibrations for the same fault on different types of machines, or even the same type of machine but in a slightly different operating environment, can be quite dissimilar. This means that it is difficult to construct an automatic classifier that works reliably for a large number of machines and faults. Automatic learning means that once features are selected from the vibration signal, it is feasible to build separate detection and diagnosis systems for every circumstance. The philosophy of the approach is to build many simple but specialised devices to perform automatic diagnosis, rather than few, complex systems that can work for a broader range of problems. Multi-layer perceptron neural networks are compatible with this approach, as the neural mechanism itself need not change from one instance to another; it is only the strengths of the interconnections between the various computational units in the network that change. This is an important consideration for implementation of a classifier in hardware.

6. REFERENCES:

Bannister, R.H. and Moore, M.P. (1986): General Rotational Machinery Expert System. In *Research and Development in Expert Systems III: Proceedings of Expert Systems '86*, Bramer, M.A. (ed), British Computer Society Specialist Group on Expert Systems, Brighton, England, pp. 140-151.

Boek, M.J. (1992): An Application of Neural Networks to Machine Condition Monitoring, Master of Applied Science Thesis, Faculty of Applied Science, Royal Melbourne Institute of Technology, Melbourne, Australia.

Braun, S. (Ed) (1986): Mechanical Signature Analysis: Theory and Applications, Academic Press, London.

Gersch, W. (1986): Two applications of parametric time series modelling methods. In *Mechanical Signature Analysis: Theory and Applications.* Academic Press, Braun, S., Ch. 10.

Gorman, R.P. and Sejnowski, T.J. (1988): Analysis of Hidden Units in a Layered Network Trained to Classify Sonar Targets. *Neural Networks 1*, pp. 75-89.

Hoffman, R.L. and Fukunaga, K. (September 1971): Pattern Recognition Signal Processing for Mechanical Diagnostics Signature Analysis. *IEEE Transactions on Computers C-20*, 9 , pp. 1095-1100.

McClelland, J.L. and Rumelhart, D.E. (Eds) (1988): Explorations in Parallel Distributed Processing: A Handbook of Models, Programs and Exercises, MIT Press, Cambridge, Mass., Vol. 3.

McGraw, K.L. and Harbinson-Briggs, K. (1989): Knowledge Acquisition: Principles and Guidelines, Prentice-Hall Int. Ed., London.

Milne, R. (1990): Amethyst: Vibration Based Condition Monitoring. In *The Handbook of Expert System Applications in Manufacturing*. McGraw Hill, Keyes, J. and Maus, R (eds).

Minsky, M. and Papert, S. (1969): *Perceptrons: An Introduction to Computational Geometry,* MIT Press, Cambridge, Mass..

Pao, Y.H. (1989): *Adaptive Pattern Recognition and Neural Networks,* Addison Wesley.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986): Learning Internal Representations by Error Propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition Vol. 1 Foundations.* MIT Press, Rumelhart, D.E. and McClelland, J.L., pp. 318-362.

Rumelhart, D.E. and McClelland, J.L. (1986): PDP Models and General Issues in Cognitive Science. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition Vol. 1 Foundations.* MIT Press, Rumelhart, D.E. and McClelland, J.L., pp. 110-146.