Control Eng. Practice, Vol. 1, No. 5, pp. 771-778, 1993
Printed in Great Britain.

0967-0661/93 $6.00 + 0.00

PATTERN CLASSIFIER FOR FAULT DIAGNOSIS OF HELICOPTER GEARBOXES

H. Chin*, K. Danai* and D.G. Lewicki**

**Department of Mechanical Engineering, University of Massachusetts, Amherst, MA 01003, USA*
***U.S. Army Research Laboratory, Vehicle Propulsion Directorate, NASA Lewis Research Center, Cleveland, OH 44135, USA*

Abstract. Application of a diagnostic system to a helicopter gearbox is presented. The diagnostic system is a nonparametric pattern classifier that uses a *multi-valued influence matrix (MVIM)* as its diagnostic model and benefits from a fast learning algorithm that enables it to estimate its diagnostic model from a small number of measurement-fault data. To test this diagnostic system, vibration measurements were collected from a helicopter gearbox test stand during accelerated fatigue tests and at various fault instances. The diagnostic results indicate that the MVIM system can accurately detect and diagnose various gearbox faults so long as they are included in training.

Key Words. Failure detection; monitoring; pattern recognition.

1. INTRODUCTION

Helicopter drive trains are significant contributors to both maintenance cost and flight safety incidents. Drive trains comprise almost 30% of maintenance costs and 16% of mechanically related malfunctions that often result in the loss of aircraft (Chin and Danai, 1991a). As such, it is crucial that faults be diagnosed in-flight so as to prevent loss of lives.

Fault diagnosis of helicopter power trains is based primarily on vibration monitoring. As such, considerable effort has been directed toward the identification of features of vibration that are affected by specific faults (e.g., Pratt, 1986; Mertaugh, 1986), and the development of signal processing techniques that can quantify such features through the parameters they estimate. For example, the crest factor of vibration, which represents the peak-to-rms ratio of vibration, has been shown to increase with localized faults such as gear tooth cracks (Braun, 1986). The main problem with this approach, however, is that due to the complexity of helicopter gearboxes and the interaction between their various components, the individual parameters estimated from vibration measurements do not provide a reliable basis for diagnosis.

As an alternative to single-parameter based diagnosis, fault signatures consisting of many parameters can be established using pattern classification techniques (Pau, 1981; Gallant, 1987). Among the various pattern classifiers used for diagnosis, artificial neural nets are the most notable due to their nonparametric nature (independence of the probabilistic structure of the system), and their ability to generate complex decision regions. However, neural nets generally require extensive training to develop the decision regions (diagnostic model). In cases such as helicopter power trains, where adequate data may not be available for training, artificial neural nets may misdiagnose the fault.

This paper demonstrates the application of a diagnostic method that can establish the fault signatures based on a small number of measurement-fault data. This method uses nonparametric pattern classification to estimate its diagnostic model so, like artificial neural nets, is independent of the probabilistic structure of the system. This method utilizes a *multi-valued influence matrix (MVIM)* as its diagnostic model which provides indices for diagnosability of the system and variability of the fault signatures (Danai and Chin, 1991). These indices are used as feedback to improve fault signatures through adaptation (Chin and Danai, 1991b).

To train and test the MVIM method, vibration data reflecting the effect of various helicopter gearbox faults was obtained at NASA. This vibration data was then processed through a micro-computer customized for vibration signal processing. The parameters obtained from this signal processor were then utilized to train the MVIM method and test its performance. Diagnostic results indicate that the MVIM method can correctly diagnose various faults when they are included in the training set.

2. THE MVIM METHOD

Measurements are processed in the MVIM method as illustrated in Fig. 1. They are usually filtered first to obtain a vector of processed measurements P, then they are converted to binary numbers through a flagging operation (i.e., abnormal measurements characterized by '1' and normal ones by '0'), and finally they are analyzed through the diagnostic model. In the MVIM method, flagging is performed by a Flagging Unit that is tuned according to measures of diagnosability and fault signature variability obtained from MVIM so as to improve the fault signatures. The MVIM method is explained in detail in (Danai and Chin, 1991) and (Chin and Danai, 1992). The overall concept will be discussed here for completeness.

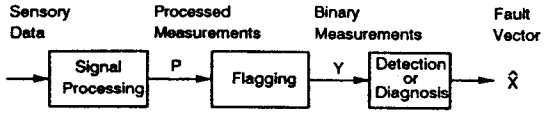

Fig. 1: Processing of measurements in the MVIM method

2.1. Fault Signature Representation

Fault signatures in the MVIM method are represented by the n unit-length columns $\bar{V}_j \in \mathcal{R}^m$ of a multi-valued influence matrix (MVIM) $\bar{A}$:

$$\bar{A} = [\bar{V}_1, \ \ldots, \ \bar{V}_j, \ \ldots, \ \bar{V}_n] \qquad (1)$$

where m denotes the number of measurements, and n represents the number of faults. Based upon this influence matrix, the faults can be ranked according to their possibility of occurrence by the closeness of their influence vector $\bar{V}_j$ to the vector of flagged measurements Y (see Fig. 2). Note that in MVIM, m is usually larger than n so as to enhance the uniqueness of fault signatures. However, $m > n$ is not a necessary condition for this method, due to the underlying assumption that only one fault is present at a time. In the MVIM method, the vector of *diagnostic certainty measures*, $\hat{X}$, which ranks the faults according to their possibility of occurrence is defined as

$$\hat{X} = \{\hat{x}_1, .., \hat{x}_j, .., \hat{x}_n\}^T = \cos\{\alpha_1, .., \alpha_j, .., \alpha_n\}^T \quad (2)$$

where the $\hat{x}_j$ represent the individual *diagnostic certainty measures*, and the α_j denote the individual angles between the influence vectors $\bar{V}_j$ and the flagged measurement vector Y (see Fig. 2).

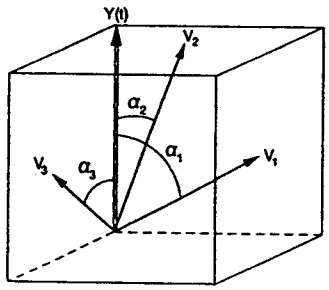

Fig. 2: Schematic of diagnostic reasoning in the MVIM method, illustrated in three dimensional space

2.2. Estimation of $\bar{A}$

The influence vectors in Eq. (1) are not known a-priori and need to be estimated. In the MVIM method, the error in diagnosis is used as the basis to estimate/update the influence vectors. For this purpose, the fault signatures are updated recursively after the occurrence of each fault to minimize the sum of the squared diagnostic error associated with that fault (Danai and Chin, 1991).

2.3. Fault Signature Evaluation

One of the unique features of the MVIM method is its ability to evaluate quantitatively the uniqueness and variability of fault signatures, so that these quantitative measures can be used to improve the flagging operation. In the MVIM method, the uniqueness of fault signatures is characterized by the closeness of pairs of influence vectors. For this purpose, a diagnosability matrix $\mathcal{D}$ is defined to represent the closeness of the orientation of individual influence vectors as

$$\mathcal{D} = \sin\left[\cos^{-1}\left(\hat{\bar{A}}^T \hat{\bar{A}}\right)\right] \qquad (3)$$

where $\hat{\bar{A}}$ is the estimated influence matrix. Matrix $\mathcal{D}$ in Eq. (3) is a 0 diagonal n dimensional symmetric matrix representing the sine of the angles between pairs of influence vectors. The index of diagnosability d is defined as the smallest off-diagonal component of matrix $\mathcal{D}$ to denote the closest pair of fault signatures.

In the MVIM method, the variability of fault signatures is defined by their variance. For this purpose, the variance matrix $\hat{\Sigma}$ associated with $\hat{\bar{A}}$ is estimated to provide a measure of the variations of individual components of $\hat{\bar{A}}$. Since in the MVIM method the components of $\hat{\bar{A}}$ are adjusted recursively, $\hat{\Sigma}$ can be readily estimated during training (Chin and Danai, 1991b):

$$\hat{\Sigma} = [\hat{\sigma}_{ij}] = \left[\frac{1}{N_j}\sum_{k_j=1}^{N_j}(\bar{y}_i(k_j) - \hat{\bar{a}}_{ij}(k_j))^2\right], \qquad (4)$$

where $k_j = 1, 2, \ldots, N_j$ represents the number of occurrences of the jth fault, the $\bar{y}_i$ represent the components of the normalized measurement vector $\bar{Y}$, and the $\hat{\bar{a}}_{ij}$ denote the components of $\hat{\bar{A}}$. The index of fault signature variability v in the MVIM method is defined as the largest component of the variance matrix, which represents the largest variability in the components of matrix $\hat{\bar{A}}$.

2.4. Flagging

The influence matrix $\bar{A}$ is estimated based on the values of the flagged measurement vector Y. Thus, before the influence matrix is used for diagnostic reasoning, the integrity of the flagging operation needs to be ensured. Flagging in the MVIM method is performed by a Flagging Unit that is tuned to improve the diagnosability of the system and reduce the variability of the fault signatures (Chin and Danai, 1991b).

The Flagging Unit uses a sample set of measurement-fault vectors to tune its parameters iteratively. After each pass through the training batch, the Flagging Unit counts the total number of false alarms and undetected faults, and estimates the uniqueness and variability of the estimated fault signatures from the current values of the influence matrix, so that it can use these measures as feedback in the next adaptation round. Adaptation stops when the total number of false alarms and undetected faults are minimized, and the uniqueness and consistency of fault signatures are enhanced.

The Flagging Unit processes the residuals as follows (see Fig. 3). The residual vector $P \in \mathcal{R}^m$ is first passed through Hard-Limiter I (consisting of a vector of m thresholds, h_{1i}, $i = 1, \ldots, m$) to produce a binary vector $Z \in B^m$. This vector is then multiplied sequentially by the normalized columns of the Quantization Matrix

$$Q = [Q_1, \ \ldots, \ Q_i, \ \ldots, \ Q_m] \qquad (5)$$

and then thresholded by Hard-Limiter II as

$$y_i = \begin{cases} 1 & \text{when } \mathbf{Z}^T \bar{\mathbf{Q}}_i \geq h_{2i} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

to produce the individual components of the flagged measurement vector $\mathbf{Y} \in \mathcal{B}^m$. The vectors $\bar{\mathbf{Q}}_i$ in Eq. (5) represent the normalized columns of the Quantization Matrix $\mathbf{Q}$ associated with individual measurements, and the h_{2i} denote the thresholds of Hard-Limiter II associated with individual measurements. Note that although the processing of residuals by the Flagging Unit is similar to residual processing in model-based methods, there are fundamental differences as well. The two methods are similar from the point of view that both perform hard-limiting (by the Hard-Limiters) to produce 'structured residuals', and residual transformation (by projecting Z on the individual columns of the Quantization Matrix) to produce 'direction-fixed residuals' (Gertler, 1991). The differences are that (i) the Flagging Unit does not utilize pre-specified thresholds or quantization vectors like model-based methods, and (ii) it employs hard-limiting and residual transformation together, whereas in model-based methods these operations are commonly used independently. Training of the Flagging Unit comprises adjusting the thresholds of Hard-Limiters I and II and the normalized quantization vectors $\bar{\mathbf{Q}}_i$, as explained in detail in (Chin and Danai, 1992). It should also be noted that conversion of measurements to binary numbers is the simplest form of flagging, which may result in loss of information. To provide more resolution, the Flagging Unit can be modified such that the flagged measurements will have values between 0 and 1.

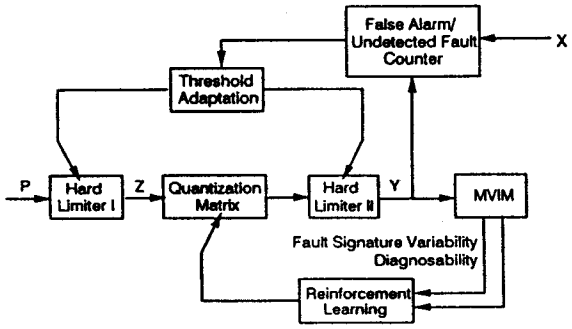

Fig. 3: Schematic of the Flagging Unit

3. EXPERIMENTAL

Vibration data were collected at NASA Lewis Research Center. Various component failures in an OH-58A main rotor transmission were produced during the experiments (Lewicki et al., 1992). The configuration of the transmission which was tested in the NASA 500-hp Helicopter Transmission Test Stand is shown in Fig. 4. The vibration signals were recorded from eight piezoelectric accelerometers (frequency range of up to 10 KHz) using an FM tape recorder. The signals were recorded once every hour for about one to two minutes per recording (at the tape speed of 30 in/sec, providing a bandwidth of 20 KHz). Two magnetic chip detectors were also used to detect the debris caused by component failures. The location and orientation of the accelerometers are shown in Fig. 5.

In these experiments, accelerated fatigue tests were performed. The transmission was run under a constant load

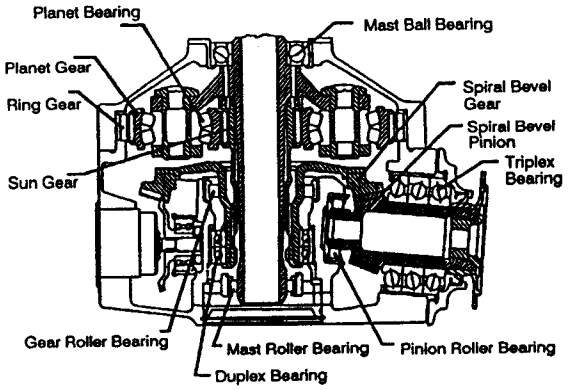

Fig. 4: Configuration of the OH-58A main rotor transmission

and was disassembled/checked periodically or when one of the chip detectors indicated a failure. A total of five tests were performed, where each test was run between nine to fifteen days for approximately four to eight hours a day. New components were used at the start of each test. When a component fault was detected during a test, it was replaced with a new one for the remainder of the test. Among the eleven failures occurring during these tests (see Table 1), there were three cases of planet bearing failure, three cases of sun gear failure, two cases of top housing cover cracking, and one case each of spiral bevel pinion, mast bearing, and planet gear failure. Insofar as fault detection during these tests, the chip detectors were reliable in detecting failures in which a significant amount of debris was generated, such as the planet bearing failures and one sun gear failure. The remaining failures were detected during routine disassembly and inspection.

Table 1: Faults occurring during the experiments

Test #	Number of Days	Failures
1	9	Sun gear tooth pit
		Spiral bevel pinion scoring
2	9	None
3	13	Planet bearing spall
		Top housing cover crack
		Planet bearing spall
		Mast bearing micropitting
4	15	Planet bearing spall
		Sun gear tooth pit
5	11	Sun gear teeth spalls
		Planet gear tooth spall
		Top housing cover crack

4. SIGNAL PROCESSING

In order to identify the effect of faults on the vibration data, the vibration signals obtained from the five tests were digitized and processed by a commercially available signal analyzer (Stewart Hughes, 1987) with four processing modules: (1) *Statistical Analysis (STAT)*, (2) *Baseband Power Spectrum Analysis (BBPS)*, (3) *Bearing Analysis (BRGA)*, and (4) *Signal Averaging Analysis (SGAV)*.

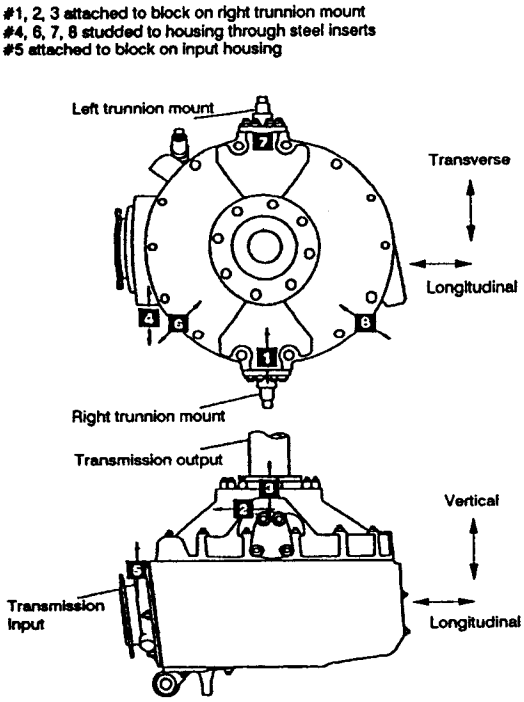

Fig. 5: Location of the accelerometers on the test stand

For analysis purposes, only one data record per day was used for each test. The data records were taken at the beginning of the day unless a fault was reported. When a fault was detected, the record was taken right before the fault incident to ensure that the data record reflected the fault. Also, in order to reduce estimation errors, each data record was partitioned into sixteen segments, and parameters were estimated for each segment and averaged over these segments. The data records as well as the parameters obtained from the above processing modules were then transferred to a personal computer for further analysis (Chin, 1992). Note that the objective of this paper is to demonstrate the MVIM pattern classification scheme, not to develop or verify individual diagnostic algorithms. The algorithms described next were used to produce inputs to the MVIM system, but may not have been optimized for transmission health monitoring.

- *Statistical Analysis.* It is generally believed that the probability density function of the vibration amplitude is near Gaussian when machinery is healthy, and that its shape changes when a defect appears. The *Statistical Analysis* module of the signal analyzer estimates parameters that would characterize such change. The parameters obtained from this module were the *skewness*, *kurtosis*, *crest factor*, and *peak-to-peak value* of vibration data.

- *Baseband Power Spectrum Analysis.* Spectrum analysis (or frequency domain analysis) is perhaps the most widely used technique in vibration signal processing, as failures such as unbalance, misalignment, wear, and bearing spalling produce a clear change in the spectrum (e.g., see (Dewell and Mitchell, 1984; Randall, 1982; Taylor, 1980; Lees and Pandey, 1980)). However in complex machinery where the background noise masks the basic distress signal, changes in the spectra cannot be easily distinguished (Pratt, 1986). The *Baseband Power Spectrum Analysis* module provides several parameters that can be associated with the frequencies generated by individual components of the transmission. The parameters obtained from this module were the *root-mean-square*, *white spectrum*, *rice frequency*, *comparison analysis*, *metacepstral analysis*, and *tone analysis* (Stewart Hughes, 1987).

- *Bearing Analysis.* The vibration energy of bearing elements is usually lower than those produced by gears, shafts, and sometimes noise. As such, bearing faults cannot be readily detected through abnormalities in the bearing tone. However, since bearing faults such as spalling produce time domain impulses which modulate the bearing shaft frequency over a wide range of frequencies, there are features of high-frequency vibration that would reflect bearing faults (Mathew and Alfredson, 1984; Braun and Datner, 1979). The *Bearing Analysis* module is designed to extract such features. This module uses a *heterodyner* to demodulate the vibration signals and obtain an amplitude envelope (e.g., see Courrech and Gaudet, 1985). The module then calculates the power spectrum of this envelope (i.e., spectral envelope) so that its various features (parameters) can be estimated for bearing fault detection (Dyer and Stewart, 1978; Yhland and Johansson, 1970). The parameters obtained from this module were: *envelope band energy*, *envelope kurtosis value*, *envelope base energy*, and *envelope tone energy* (Stewart Hughes, 1987).

- *Signal Averaging Analysis.* The vibration data usually contains signals generated by different components of the gearbox, as well as background noise. It is therefore necessary to filter the signal such that it can be related to a particular component of interest. This can be achieved through *signal averaging* (or synchronous time averaging) which averages the digitized time signals over a large number of revolutions synchronous with the speed of a particular component (McFadden and Smith, 1985; McFadden, 1987). Signal averaging is particularly useful for the identification of localized faults like a gear tooth spall (McFadden, 1986). Once the averaged signals are obtained, they can be further analyzed for features that reflect various gear faults. The parameters obtained from this module were: *FM1*, *FM1A-B*, *FM4A*, *FM4B*, and *comparison analysis parameters* (Stewart Hughes, 1987).

5. IMPLEMENTATION AND RESULTS

As mentioned in Section 2, the MVIM method requires a set of measurements during normal operation and at various fault incidents to estimate the diagnostic model. As such, the exact times for the fault incidents of the five tests needed to be established before the measurements could be used for training and testing the MVIM method. Unsupervised learning *(Kohonen's Feature Mapping)* (e.g., see (Kohonen, 1989)) was first used to classify individual parameters into normal and abnormal. The exact times of

fault incidents were then established through correlating these parameters with the faults which had been detected at the end of each test.

The status of various faults during the five tests are shown in Table 2 where no-fault cases are denoted by x_0. In Test #1, spiral bevel pinion failure (x_4) is estimated to have been present on days 5-9 with sun gear failure (x_1) also occurring on day 9. No failures occurred in Test #2, so all the nine days for this test are marked as normal. In Test #3, planet bearing failure (x_2) was established to have been present on days 3-4 and 11-12, with housing crack (x_3) and mast bearing failure (x_6) occurring on day 9 and day 13, respectively. Test #4 is esimated to have contained planet bearing failure on days 11-12 and sun gear failure on days 14-15. In Test #5, housing crack is assumed to have been present on days 7-11, with sun gear failure (x_1) and planet gear failure (x_5) occurring on days 10-11 and day 11, respectively.

Table 2: Association of data from each day of the five tests with no-fault and various fault cases. The no-fault cases are denoted as x_0 and the six faults are represented as x_1: sun gear failure, x_2: planet bearing failure, x_3: housing crack, x_4: spiral bevel pinion failure, x_5: planet gear failure, x_6: mast bearing failure

	Fault Status				
Day	Test #1	Test #2	Test #3	Test #4	Test #5
1	x_0	x_0	x_0	x_0	x_0
2	x_0	x_0	x_0	x_0	x_0
3	x_0	x_0	x_2	x_0	x_0
4	x_0	x_0	x_2	x_0	x_0
5	x_4	x_0	x_0	x_0	x_0
6	x_4	x_0	x_0	x_0	x_0
7	x_4	x_0	x_0	x_0	x_3
8	x_4	x_0	x_0	x_0	x_3
9	x_4, x_1	x_0	x_3	x_0	x_3
10			x_0	x_0	x_3, x_1
11			x_2	x_2	x_3, x_1, x_5
12			x_2	x_2	
13			x_6	x_0	
14				x_1	
15				x_1	

The configuration of the MVIM system as applied to fault diagnosis of the OH-58A main rotor transmission is illustrated in Fig. 6. As shown in this figure, two MVIMs are trained for each accelerometer, one MVIM to perform detection (i.e., to determine whether a fault has occurred or not), and a diagnostic MVIM to isolate the fault. The 54 parameters obtained from the signal analyzer were used to train and test the MVIM system. The detection MVIM contained only two columns to characterize the no-fault and fault signatures, whereas the diagnostic MVIM contained seven columns, one characterizing the no-fault signature and the other six representing the signatures of individual faults. Note that the two MVIMs can be perceived as filters with different resolutions. In order to integrate the results from the MVIMs associated with the eight accelerometers, a voting scheme was utilized.

Tests #3 and #4 contained most of the failure modes (i.e., 4 out of 6). Therefore, the parameters from these two

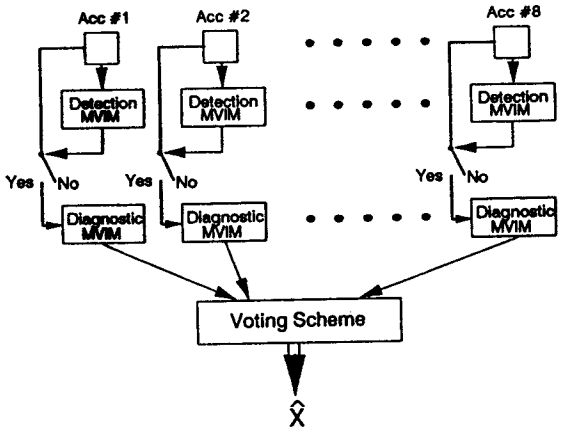

Fig. 6: Configuration of the MVIM system as applied to the OH-58A main rotor transmission

tests were used to train the MVIMs. Note that not all of the failure modes were included in training, so the test results were not expected to be perfect. For training the detection MVIMs, only the 19 parameters from the *STAT*, *BBPS*, and *BRGA* modules were used. Previous studies on this data show that these 19 parameters are adequate for detection (Chin, 1993). For training the diagnostic MVIMs, all of the 54 parameters were utilized.

The initial values of the detection MVIMs (19 × 2) and diagnostic MVIMs (54 × 7) were each set to 0 (i.e., matrices with all zero entries), and the initial values of the Quantization Matrices (19 × 19 for detection) and (54 × 54 for diagnosis) were set to identity matrices. The initial threshold levels for Hard-Limiter I were set at the mean plus one standard deviation of the corresponding parameter, and for Hard-Limiter II, they were set at 0.5. The maximum number of epochs for training the detection and diagnostic MVIMs was set to 50. After each epoch, the detection/diagnostic performance of MVIMs within the training set was tested. Training was stopped once perfect detection/diagnosis was achieved, to avoid overtraining (Hertz et al., 1991). The number of epochs used for individual detection MVIMs were: 8, 5, 50, 37, 50, 15, 50, and 50 for accelerometers #1 to #8, respectively, whereas for diagnostic MVIMs they were: 50, 1, 2, 2, 26, 50, 50, and 50. According to the number of epochs used for individual MVIMs, it is clear that the detection MVIMs associated with accelerometers #3, #5, #7, and #8 did not achieve perfect detection within the training set. Similarly, the diagnostic MVIMs associated with accelerometers #1, #6, #7, and #8 did not achieve perfect diagnosis within the training set.

The MVIMs trained on Tests #3 and #4 were evaluated for all of the tests. For this purpose, the 19 parameters from each of the eight accelerometers were first passed through the corresponding detection MVIM for all of the five tests to reflect the occurrence of faults. Once a fault was posted by a detection MVIM, the set of 54 parameters from that accelerometer was passed through the corresponding diagnostic MVIM. At the final stage, the diagnostic certainty measures obtained from the eight diagnostic MVIMs were consolidated by a voting scheme.

The voting scheme utilized weights which reflected the

speed of convergence of individual fault signatures during training. That is,

$$weight = \frac{1}{number\ of\ epochs}. \quad (7)$$

As such those influence vectors which converged faster were assigned larger weights and vice versa. Zero weights were assigned to the influence vectors which did not converge during training, and unity weights to those which converged within one epoch. The weights associated with individual fault signatures for the eight diagnostic MVIMs, based on training with Tests #3 and #4, are shown in Table 3.

Table 3: Weights of individual fault signatures for the eight diagnostic MVIMs. The x_i are the same as indicated in Table 2. The '-' denotes that the fault signature was not present in the training set

		Weights						
MVIM	Fault	x_0	x_1	x_2	x_3	x_4	x_5	x_6
1		1	1	0	0	-	-	0.02
2		1	1	1	1	-	-	1
3		1	0.5	0.5	1	-	-	1
4		0.5	0.5	0.5	0.5	-	-	1
5		1	0.02	0	0	-	-	0
6		0.5	0	0	1	-	-	0.5
7		1	0	0	0.5	-	-	0.33
8		1	0.02	0	1	-	-	0.33

To illustrate the implementation of the voting scheme, consider fault diagnosis in Test #5. The eight detection MVIMs trained based on Tests #3 and #4 were first utilized to detect the presence of faults in Test #5. The detection results obtained from these MVIMs are shown in Table 4. Although the results vary among the MVIMs for days 2-11, they all indicate normality on day 1. After detection of faults by each detection MVIM, its diagnostic MVIM, which was already trained on Test #3 and #4, was used to identify the faults. The diagnostic results produced by individual MVIMs are shown in Table 5. As expected, the results are not consistent. In order to consolidate the results, the voting scheme was used to assign more weight to those MVIMs which had more quickly established the fault signature associated with the diagnosed fault. For this purpose, the results in Table 5 were multiplied by the weight of the associated influence vector (see Table 3) and were added together. The diagnostic results for Test #5 are shown in Table 6, where the faults with the highest score are shown by an asterisk.

The diagnostic results obtained from the above diagnostic system are shown in Table 7. The results indicate that the MVIM system was able to produce perfect diagnostics for Tests #3 and #4, on which it was trained, and that it diagnosed 88% of the faults in all of the tests. Specifically, the results in Table 7 indicate that the MVIM system produced two false alarms (on day 4 of Test #1 and day 6 of Test #5), and five misdiagnoses (on days 5-8 of Test #1 and day 11 of Test #5). In addition, this system produced equal diagnostic certainty measures for the no-fault case and sun gear failure on day 10 of Test #5 and could only diagnose one of the faults on day 9 of Test #1, and on days 10 and 11 of Test #5. However, it should be noted that faults x_4 and x_5 were not included in training, so no

Table 4: Detection results obtained by the eight detection MVIMs for Test #5. '0's denote no-fault and '1's denote fault cases

Test #5		Estimated Fault Status							
Day	MVIM	1	2	3	4	5	6	7	8
1		0	0	0	0	0	0	0	0
2		0	0	0	1	0	0	0	0
3		0	0	0	1	0	0	0	0
4		1	0	0	1	1	0	0	0
5		0	0	0	1	0	0	0	0
6		1	0	1	1	0	0	1	0
7		1	0	1	0	0	0	0	1
8		0	0	0	1	1	0	1	0
9		1	1	1	1	0	0	1	1
10		1	0	1	1	1	0	0	0
11		1	1	1	1	0	1	1	1

Table 5: Individual diagnostic results obtained by the eight diagnostic MVIMs for Test #5 before the application of the voting scheme. The '-' denotes that diagnosis was not performed, based on the detection results. The x_i are the same as indicated in Table 2

Test #5		Estimated Faults							
Day	MVIM	1	2	3	4	5	6	7	8
1		-	-	-	-	-	-	-	-
2		-	-	-	x_0	-	-	-	-
3		-	-	-	x_0	-	-	-	-
4		x_2	-	-	x_0	x_6	-	-	-
5		-	-	-	x_0	-	-	-	-
6		x_2	-	x_6	x_0	-	-	x_3	-
7		x_2	-	x_3	-	-	-	-	x_1
8		-	-	-	x_3	x_6	-	x_3	-
9		x_2	x_2	x_6	x_0	x_3	-	x_3	x_3
10		x_2	-	x_1	x_0	x_3	-	-	-
11		x_2	x_2	x_6	x_3	x_3	x_2	x_1	x_2

fault signatures could be estimated for them. The correct diagnostic rate of MVIM, with these two faults excluded would be over 95%, which is quite noteworthy considering that the MVIM system was trained on a small number of measurement-fault data with very few repetitions of each fault.

To compare the diagnostic performance of MVIM with a conventional pattern classifier, the parameters obtained from the signal analyzer were also used to train and test a Bayes system similar to the one in Fig. 6. In this system, Bayes classifiers were used in place of individual MVIMs. The diagnostic model in a Bayes classifier is the probability density matrix B whose components b_{ij} are estimated according to the maximum likelihood formula (Duda and Hart, 1973)

$$\hat{b}_{ij} = \frac{1}{N_j} \sum_{k_j=1}^{N_j} y_i(k_j). \quad (8)$$

In the above equation, $k_j = 1, 2, \ldots, N_j$ represents the number of occurrences of the jth fault, x_j, and the y_i denote the individual flagged measurements. According to Bayes classification theory, the discriminant functions which minimize the average probability of error for statis-

Table 6: Sums of the weighted diagnostic results obtained from the eight diagnostic MVIMs for Test #5. The '-' denotes that diagnosis was not performed, based on the detection results. The '' denotes the "winner"of each day. The x_i are the same as indicated in Table 2*

Test #5	Sums of Weighted Diagnostic Results							
Day	Fault	x_0	x_1	x_2	x_3	x_4	x_5	x_6
1		-	-	-	-	-	-	-
2		0.5*	0	0	0	0	0	0
3		0.5*	0	0	0	0	0	0
4		0.5*	0	0	0	0	0	0
5		0.5*	0	0	0	0	0	0
6		0.5	0	0	0.5	0	0	1*
7		0	0.02	0	1*	0	0	0
8		0	0	0	1*	0	0	0
9		0.5	0	1	1.5*	0	0	1
10		0.5*	0.5*	0	0	0	0	0
11		0	0	1*	0.5	0	0	1*

tically independent y_i are defined as

$$g_j(\mathrm{X}) = Prob(x_j|\mathrm{Y}) = \prod_{i=1}^{n} b_{ij}^{y_i}(1-b_{ij})^{1-y_i} \qquad (9)$$

where $g_j(\mathrm{X})$ represents the discriminant function associated with x_j and Y is the vector of flagged measurements. According to Bayes decision theory, the largest discriminant function in the above set is attributed to the occurred fault (Duda and Hart, 1973).

To implement the Bayes method in diagnosis, a Bayes system with two Bayes classifiers were used for each accelerometer, one to perform detection and the other to isolate the fault. The detection Bayes classifier, similar to MVIM, contained only two columns to characterize the no-fault and fault signatures, whereas the diagnostic Bayes classifier contained seven columns, one characterizing the no-fault signature and the other six representing the signatures of individual faults. For consistency, the parameters from Tests #3 and #4 were used to train the Bayes classifiers. The parameters from the two tests were first flagged to produce y_i, with threshold values set at the mean plus one standard deviation of the corresponding parameters. Flagging was found to improve the performance of the Bayes classifiers. The flagged measurements were then used to obtain the probability density matrices $\hat{\mathrm{B}}$ according to Eq. (8).

The performance of the Bayes system was evaluated for all of the five tests. For this purpose, the fifty-four parameters from each of the eight accelerometers were first flagged according to the threshold values obtained from Tests #3 and #4. Like MVIM, the nineteen parameters from each of the eight accelerometers were then passed through the corresponding detection Bayes classifier, and once the presence of a fault was indicated by the detection classifier, all of the fifty-four parameters from that accelerometer were passed through the corresponding diagnostic Bayes classifier for fault isolation. The diagnostic results obtained from the eight diagnostic Bayes classifiers were then averaged to produce the estimated faults. Note that since training for Bayes takes only one epoch, a voting scheme such as the one devised for MVIM could not be utilized.

Table 7: Estimated fault status from each day of the five tests by the MVIM system. Same notations are adopted as in Table 2

	Estimated Fault Status (MVIM System)				
Day	Test #1	Test #2	Test #3	Test #4	Test #5
1	x_0	x_0	x_0	x_0	x_0
2	x_0	x_0	x_0	x_0	x_0
3	x_0	x_0	x_2	x_0	x_0
4	x_3	x_0	x_2	x_0	x_0
5	x_3	x_0	x_0	x_0	x_0
6	x_3	x_0	x_0	x_0	x_6
7	x_3	x_0	x_0	x_0	x_3
8	x_3	x_0	x_0	x_0	x_3
9	x_1	x_0	x_3	x_0	x_3
10			x_0	x_0	x_0,x_1
11			x_2	x_2	x_2, x_6
12			x_2	x_2	
13			x_6	x_0	
14				x_1	
15				x_1	

The diagnostic results obtained from the Bayes system for all of the five tests are shown in Table 8, with the actual faults indicated inside parentheses. The results indicate that although the Bayes system was able to produce near perfect diagnostics for Tests #3 and #4 (except for the false alarm on day 8 of Test #3), it did not detect any of the faults in Tests #1 and #5. This indicates that the Bayes classifier was particularly dependent on the training set.

Table 8: Estimated fault status from each day of the five tests by the Bayes system. Same notations are adopted as in Table 2

	Estimated Fault Status (Bayes System)				
Day	Test #1	Test #2	Test #3	Test #4	Test #5
1	x_0	x_0	x_0	x_0	x_0
2	x_0	x_0	x_0	x_0	x_0
3	x_0	x_0	x_2	x_0	x_0
4	x_0	x_0	x_2	x_0	x_0
5	x_0	x_0	x_0	x_0	x_0
6	x_0	x_0	x_0	x_0	x_0
7	x_0	x_0	x_0	x_0	x_0
8	x_0	x_0	x_2	x_0	x_0
9	x_0	x_0	x_3	x_0	x_0
10			x_0	x_0	x_0
11			x_2	x_2	x_0
12			x_2	x_2	
13			x_6	x_0	
14				x_1	
15				x_1	

6. CONCLUSION

An efficient fault diagnostic system based on the MVIM method is introduced and applied to a helicopter gearbox from which only a small number of measurement-fault data was available. This diagnostic system utilizes two levels of isolation and integrates the results obtained from various accelerometers through a voting scheme. The diagnostic

results indicate that the MVIM system correctly detected all of the fault incidents, with only two false alarms posted. The results further indicate that this system correctly diagnosed all of the faults it was trained for.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to Sikorsky Aircraft Company for its continued support of this project and NASA for providing the experimental data. This work was supported in part by the National Science Foundation (Grants No. DDM-9015644 and MSS-9102149).

REFERENCES

Braun, S. (1986). *Mechanical Signature Analysis - Theory and Applications*, Academic Press, New York, NY.

Braun, S., and B. Datner (1979). "Analysis of Roller/Ball Bearing Vibrations," *ASME J. of Mechanical Design*, Vol. 101, Jan., pp. 118–125.

Chin, H. (1992). *Vibration Analysis of an OH-58A Main Rotor Transmission*, Technical Report, Department of Mechanical Engineering, University of Massachusetts, Amherst, MA 01003.

Chin, H. (1993). *A Non-parametric Pattern Classifying Diagnostic Method and Its Application*, Ph.D. Dissertation, Department of Mechanical Engineering, University of Massachusetts, Amherst, MA 01003.

Chin, H., and K. Danai (1991a). "Fault Diagnosis of Helicopter Power Train," *Proc. of the 17th Annual NSF Grantees Conference in Design and Manufacturing Systems Research*, Austin, TX, January, pp. 787–790.

Chin, H., and K. Danai (1991b). "A Method of Fault Signature Extraction for Improved Diagnostics," *ASME J. of Dynamic Systems, Measurement, and Control*, Vol. 113, No. 4, pp. 634–638.

Chin, H., and K. Danai (1992). "Improved Flagging for Pattern Classifying Diagnostic Systems," *IEEE Trans. on Systems, Man, and Cybernetics*, in press.

Courrech, J., and M. Gaudet (1985). "Envelope Analysis - the Key to Rolling-Element Bearing Diagnosis," Bruel & Kjaer Application Notes.

Danai, K., and H. Chin (1991). "Fault Diagnosis with Process Uncertainty," *ASME J. of Dynamic Systems, Measurement, and Control*, Vol. 113, No. 3, pp. 339 –343.

Dewell, D. L., and L. D. Mitchell (1984). "Detection of a Misaligned Disk Coupling Using Spectrum Analysis," *J. of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 106, Jan., pp. 9–16.

Duda, R. O., and P. E. Hart (1973). *Pattern Classification and Scene Analysis*, Wiley, New York, NY.

Dyer, D., and R. M. Stewart (1978). "Detection of Rolling Element Bearing Damage by Statistical Vibration Analysis," *ASME J. of Mechanical Design*, Vol. 100, April, pp. 229–235.

Gallant, S. I. (1987). "Automated Generation of Connectionist Expert Systems For Problems Involving Noise and Redundancy," *Proc. of AAAI Workshop on Uncertainty.*

Gertler, J. (1991). "Analytical Redundancy Methods in Fault Detection and Isolation," *IFAC Safeprocess Symposium*, Baden-Baden, Germany, Sept., pp. 10–13.

Hertz, J., A. Krogh, and R. G. Palmer (1991). *Introduction to the Theory of Neural Computation*, Addison-Wesley, Redwood City, CA.

Kohonen, T. (1989). *Self-Organization and Associative Memory*, Springer-Verlag, Berlin, Germany.

Lees, A. W., and P. C. Pandey (1980). "Vibration Spectra from Gear Drives," *Proc. of 2nd Int. Conf. on Vibration in Rotating Machinery*, Sept., pp. 103–108.

Lewicki, D. G., H. J. Decker, and J. T. Shimski (1992). *Full-Scale Transmission Testing to Evaluate Advanced Lubricants*, Technical Report, NASA TM-105668, AVSCOM TR-91-C-035, NASA Lewis Research Center, Cleveland, OH 44135.

Mathew, J., and R. J. Alfredson (1984). "The Condition Monitoring of Rolling Element Bearings Using Vibration Analysis," *J. of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 106, July, pp. 447–453.

McFadden, P. D. (1986). "Detecting Fatigue Cracks in Gears by Amplitude and Phase Demodulation of the Meshing Vibration," *J. of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 108, April, pp. 165–170.

McFadden, P. D. (1987). "Examination of A Technique for the Early Detection of Failure in Gears by Signal Processing of the Time Domain Average of the Meshing Vibration," *Mechanical Systems and Signal Processing*, Vol. 1, No. 2, pp. 173–183.

McFadden, P. D., and J. D. Smith (1985). "A Signal Processing Technique for Detecting Local Defects in a Gear From the Signal Average of the Vibration," *Proc. of Institution of Mech. Engineers*, Vol. 199, No. C4, pp. 287–292.

Mertaugh, L. J. (1986). "Evaluation of Vibration Analysis Techniques for the Detection of Gear and Bearing Faults in Helicopter Gearboxes," *Mechanical Failure Prevention Group 41th Meeting*, Oct., pp. 28–30.

Pau, L. F. (1981). *Failure Diagnosis and Performance Monitoring*, Marcel Dekker, New York, NY.

Pratt, J. L. (1986). "Engine and Transmission Monitoring - A Summary of Promising Approaches," *Mechanical Failure Prevention Group 41th Meeting*, Oct., pp. 229–236.

Randall, R. B. (1982). "A New Method of Modeling Gear Faults," *ASME J. of Mechanical Design*, Vol. 104, April, pp. 259–267.

Stewart Hughes (1987). *Transmission Systems Analysis for the MSDA, MM35: 2nd edition*, Stewert Hughes Limited, Southhampton, U. K.

Taylor, J. I. (1980). "Identification of Bearing Defects by Spectral Analysis," *ASME J. of Mechanical Design*, Vol. 102, April, pp. 199–204.

Yhland, E., and L. Johansson (1970). "Analysis of Bearing Vibration," *Aircraft Engineering*, Dec., pp. 18–20.