Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Sound and Vibration 282 (2005) 475–508

JOURNAL OF SOUND AND VIBRATION

www.elsevier.com/locate/jsvi

A review of vibration-based techniques for helicopter transmission diagnostics

Paul D. Samuel*, Darryll J. Pines

Alfred Gessow Rotorcraft Center, Department of Aerospace Engineering, University of Maryland, College Park, MD 20742-3015, USA

Received 4 September 2003; accepted 28 February 2004
Available online 14 October 2004

Abstract

Over the past 25 years, much research has been devoted to the development of Health and Usage Monitoring (HUM) systems for rotorcraft gearbox and drivetrain components. The promise of HUM systems is the ability to provide accurate information regarding the condition of various flight critical components. This paper reviews the state of the art in vibration-based helicopter transmission diagnostics. The development of various damage detection techniques is discussed from a historical perspective, and the ability of these techniques to detect damage in helicopter transmissions is reviewed. Emerging research trends suggest that improvements in signal processing, sensor development and individual-tooth mesh waveform modelling could improve the performance of current and future helicopter transmission diagnostics.

1. Introduction

Over the past 25 years, much research has been focused on vibration-based transmission damage detection techniques. These techniques are applied to vibration signals collected from transducers, typically accelerometers, mounted to the transmission housing. The goal of these techniques is to enhance changes in the signal caused by damaged components within the

*Corresponding author. Fax: +1-301-314-9001.
E-mail address: pdsamuel@umd.edu (P.D. Samuel).

doi:10.1016/j.jsv.2004.02.058

transmission while remaining invariant in the presence of changes caused by normal variations in the operating condition of the transmission.

1.1. Background

Vibration-based transmission damage detection techniques are critical elements of a system designed to monitor the health of a helicopter, typically referred to as Health and Usage Monitoring (HUM) systems or, more recently, as Prognostic and Health Management (PHM) systems. A more general description of a HUM or PHM system must be provided in order to understand the role of a damage detection algorithm. Often, the terms HUM and PHM are used interchangeably. However, there is a subtle difference in their meaning. Specifically, the term PHM is more general, and was adopted to reflect an expansion of the desired capabilities of the system. To better understand when each term should be used, the desired capabilities and supporting technologies associated with each system should be delineated.

A HUM system is a combination of two complimentary subsystems: *health monitoring* and *usage monitoring*. Health monitoring is a procedure for determining the mechanical or functional condition of a component or system, specifically detecting and possibly diagnosing incipient damage or degradation that could ultimately lead to a system failure. Usage monitoring is a process by which the life consumption of various critical components and systems is determined by assessing operation hours and load history. Thus, a HUM system is expected to perform two specific tasks. First, it is expected to provide an accurate measure of the current usage of critical components so that statistical life models can be used to provide an accurate estimate of the point in time when the part should be replaced. Second, it is expected to indicate the presence of incipient damage that could lead to failure of a component prior to the end of its expected safe-life.

A PHM system incorporates a broader range of technologies. *Prognostics* is a general term that describes a process to predict the remaining safe-life of a component or system. Note that usage monitoring is one approach to prognostics. However, another potentially more accurate approach is to use information about the current health of the system along with damage propagation models to predict the remaining safe-life of a component once incipient damage has been detected. The term *Health Management* describes a procedure to handle the information gathered through both health monitoring and prognostics. This information can be used to provide an accurate picture of the current condition of the system and possibly suggest methods to extend the life of the system through, for example, limitations on flight maneuvers.

The ultimate goal of both PHM and HUM systems is to enable condition-based maintenance (CBM) for rotorcraft. A rotorcraft that uses CBM is said to have its maintenance *on-condition* [1]. CBM is a maintenance philosophy that states that system maintenance and component replacement should only occur when there is objective evidence of an impending failure. To implement CBM, a reliable, robust, accurate HUM or PHM system is required. To date, a sufficient HUM or PHM system has not been developed. Thus CBM has yet to be completely realized.

Currently, to ensure rotorcraft safety and reliability, manufacturers tend toward conservative safety factors on design and component safe-life estimates, and recommend frequent scheduled maintenance. More specifically, three techniques are used: damage tolerant or fail-safe designs, redundancy, and safe-life determination techniques [2]. Often a HUM system is incorporated essentially as a backup for these primary techniques.

However, two primary factors are motivating the development of enabling technologies for rotorcraft CBM. These are increased safety and reliability, and decreased operating cost.

1.2. Motivation

A rotorcraft is an extremely dynamic system. As such, all the flight critical components within the vehicle are exposed to extreme dynamic loads; sustained vibratory and impulsive loads. Thus to ensure rotorcraft safety and reliability, maintenance inspections, overhauls, and parts replacement must be frequently performed [3]. This is an expensive and time consuming task. Even given these measures, premature failure due to fatigue leads to many helicopter accidents, resulting in a fatal accident rate per mile flown for rotorcraft an order of magnitude higher than for conventional fixed-wing aircraft [4]. Also, in the United States, the annual number of rotorcraft accidents is approximately 2.5 times higher than that of fixed wing aircraft [5]. Thus both safety and cost are motivating factors for the development of HUM and PHM systems.

1.2.1. Cost

Helicopter operating costs are extremely high. Maintenance accounts for 24% of this cost, and insurance accounts for an additional 29% [4].

Placing rotorcraft maintenance on-condition would reduce the frequency of maintenance inspections. In addition, healthy components would not need to be discarded at the end of their estimated safe-life. Instead, they could remain in service until signs of failure are detected. Thus CBM allows a reduction in part replacement and a decrease in the number of hours required for maintenance, leading to a potentially significant reduction in maintenance costs.

Insurance costs are generally high due to large liability awards for accidents and high levels of danger, both real and perceived, associated with helicopter flight. However, the continuous monitoring of the condition of the rotorcraft means that a greater number of incipient failures will be detected, thus reducing the accident rate and increasing both the real and perceived level of safety. Hence, the implementation of CBM could yield a reduction in the insurance costs associated with rotorcraft operation.

1.2.2. Safety

Over the years, the National Transportation Safety Board (NTSB) has collected an extensive database of information on civil aviation accidents. From mid-1963 through 1997, they recorded a total of 8436 rotorcraft accidents and presented a report on each. This database provides invaluable information regarding the primary causes of rotorcraft accidents.

An exhaustive study of the NTSB helicopter accident reports from 1963 through 1997 was presented by Harris et al. [6]. A summary of the accident count and distribution of commercially manufactured helicopters over the period of the study is presented in Table 1.

The airframe/component/system failure/malfunction category covers the helicopter failures which could most likely be avoided through the use of CBM. It can be seen that for each type of helicopter, this category provides a moderately significant contribution to the total number of accidents. It should also be noted that within this category, drive train failures made the greatest contribution to the overall number of accidents as seen in Table 2.

It can be seen from these studies that the use of CBM could have a significant impact on the safety and reliability of rotorcraft.

1.3. Rotorcraft transmission health monitoring system

Because of the severe loading environment to which helicopters are subjected, helicopter transmissions suffer from a variety of potential damage modes. These include characteristic wearing patterns, scoring, fatigue cracks, pitting, spalling, housing cracks, bearing failures, loosening mounting bolts, as well as a host of other failure modes. Some possible transmission damage modes are displayed in Fig. 1. It should be noted that the damage shown in this figure is

Table 1
Summary accident count and distribution, 1963–1997 [6]

NTSB first event accident category	Single piston	Single turbine	Twin turbine
	Count (%)	Count (%)	Count (%)
Loss of engine power	1.554 (28.9)	704 (31.3)	39 (12.9)
In flight collision with object	953 (17.7)	298 (13.2)	43 (14.2)
Loss of control	625 (11.6)	284 (12.6)	40 (13.2)
Airframe/component/system failure/malfunction	639 (11.9)	282 (12.5)	89 (29.5)
Hard landing	483 (8.99)	140 (6.23)	8 (2.65)
In flight collision with terrain/water	443 (8.25)	143 (6.36)	8 (2.65)
Rollover/nose over	290 (5.40)	119 (5.29)	4 (1.32)
Weather	57 (1.06)	85 (3.78)	12 (3.97)
Other	327 (6.09)	192 (8.54)	51 (16.9)

Table 2
Airframe major system accident count and distribution, 1963–1997 [6]

Airframe major system	Single piston	Single turbine	Twin turbine
	Count (%)	Count (%)	Count (%)
Drive train—main	127 (19.9)	49 (17.3)	13 (14.6)
Drive train—tail	119 (18.6)	54 (19.1)	19 (21.3)
Main rotor	57 (8.9)	36 (12.6)	19 (21.3)
Tail rotor	124 (19.4)	52 (18.4)	10 (11.2)
Control system—main	63 (9.9)	29 (10.3)	11 (12.3)
Control system—tail	38 (5.9)	11 (3.9)	7 (7.9)
Airframe (fuselage, other subsystems)	64 (10.0)	41 (14.5)	8 (9.0)
Landing gear	24 (3.8)	2 (0.7)	2 (2.2)
Engine	7 (1.1)	3 (1.1)	0 (0)
Undetermined/other	16 (2.5)	5 (1.8)	0 (0)

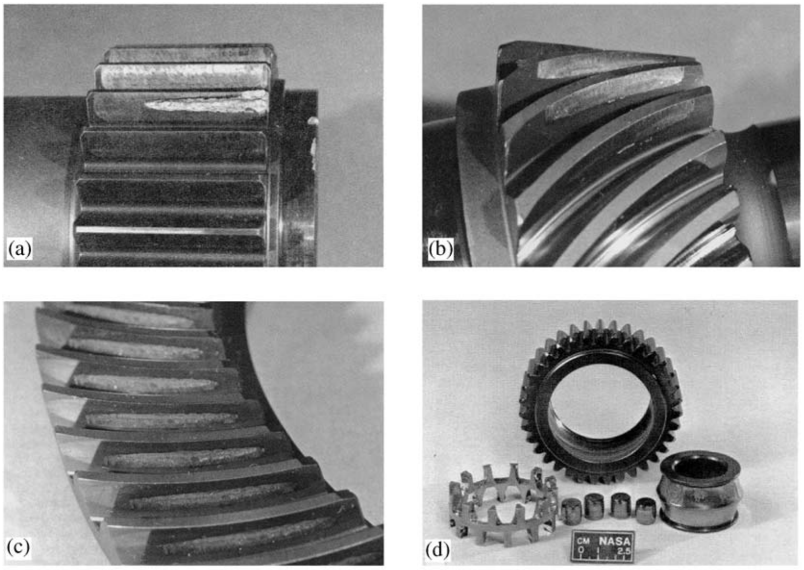

Fig. 1. OH-58A transmission faults [7]; (a) spalled sun gear; (b) scored spiral bevel pinion; (c) scored spiral bevel face gear; (d) spalled planetary bearing race.

the result of an advanced lubricant study where an OH-58 transmission mounted in a test rig was run at 117% of the maximum rated load for an extended period of time [7]. If left unchecked, many of these damage modes can lead to catastrophic failure of the transmission and thus the entire helicopter.

Transmission monitoring is a significant part of any PHM system. As stated earlier, a PHM system is composed of two components: prognostics and health management. For a transmission health monitoring system, each of these components can be subdivided into a number of sub-components, each with an associated supporting system or algorithm. A schematic of a transmission health monitoring system is given in Fig. 2. During normal operation, the PHM system acquires data from a number of sensors including accelerometers mounted to the transmission housing and a one-per-rev reference signal. The diagnostic function provides comprehensive integrated component-by-component mechanical diagnostics. The diagnostics focus on individual gears, bearings and shafts. The prognostic function then provides a prediction of the remaining life of each component. The condition of each component is monitored constantly during flight to ensure vehicle safety.

The initial research in the area of transmission damage detection was focused on vibration signal analysis, i.e. the analysis of transmission vibration signals using the various tools available in the signal processing community. At first, the statistical characteristics of the signal in the time domain were the primary focus of study. However, the field quickly expanded to include spectral analysis, time–frequency analysis and wavelet analysis. This field is continuing to grow. As new signal processing techniques emerge, they are applied to the transmission damage detection problem and their merit as potential damage detection techniques is assessed.

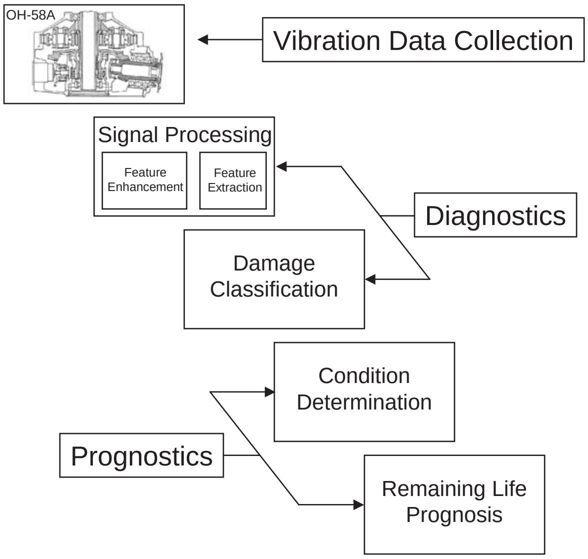

Fig. 2. Transmission PHM system schematic.

In the late 1980s, model-based transmission damage detection techniques began to appear. This approach is fundamentally different from the previous in that the technique is trained to recognize signals from a healthy transmission and then indicate when the vibration signal deviates from this nominal condition. Additionally, some techniques can be trained using data from both healthy and damaged transmission. Different types of damage are used so that the system cannot only detect the onset of damage, but also attempt to classify the type of damage that has occurred. Training, also referred to as learning or adaptation is typically autonomous.

Research into model-based damage detection focused primarily on the application of pattern classification (neural network) and mathematical modelling (autoregressive moving average) techniques. Typically, neural networks and expert systems do not use the vibration signal directly; instead, they are trained using the output from many vibration signal analysis techniques. In addition, they are trained using a data set that includes signals from the transmission operating under both healthy and damaged conditions. Mathematical modelling, on the other hand, models the vibration signal directly, and typically only learns the characteristics of a healthy signal. Wavelet-based adaptive signal processing, a technique that just recently has begun to receive the attention of the transmission diagnostics community, can be placed under the heading of mathematical modelling.

1.4. The promise of HUM systems

The promise of HUM systems is the ability to provide accurate information regarding the condition of various flight critical components. In March 1993, the first success of a HUM system was reported; a Stewart Hughes Ltd. HUM system detected its first serious problem in a Norwegian Boeing 234 Commercial Chinook, preventing it from take-off. After a maintenance inspection, it was found that a 10 mm bolt on a cross-shaft coupling between the engine and a gearbox had broken. However, whether or not the broken bolt would have caused an accident was not determined. Other successes have subsequently been reported in literature [1,4,8,9].

The following sections summarize many of the fundamental transmission damage detection techniques presented in the literature over the past 25 years.

2. Statistical damage detection metrics

2.1. Theory

The traditional techniques for vibration-based transmission damage detection are typically based on some statistical measurement of the energy of the vibration signal. For example, it was observed that the rms energy of the vibration signal may change in the presence of damage. However, in the mid 1970s, Stewart [10] began to more rigorously investigate the changes in a transmission vibration signal due to gear damage. He made the observation that under no fault conditions, a transmission vibration signal tends to be dominated by the mesh frequency and its harmonics, referred to as the regular meshing components of the signal, and that a noise floor is present which can generally be assumed to be Gaussian. He also noted the presence of sidebands about the regular meshing components, spaced at the rotational frequency of the gear of interest.

Based on these observations, McFadden [11] proposed a basic mathematical model of the transmission vibration signal. He stated that a signal from a perfect transmission can be approximated by

$$x_{\text{perfect}}(t) = \sum_{n=0}^{\infty} P_n \cos(n\omega t + \phi_n) + w(t), \tag{1}$$

where P_n is the amplitude of the nth harmonic, ω is the mesh frequency given by $\omega = 2\pi f_r T$, f_r is the gear rotation frequency, T is the number of teeth, ϕ_n is the phase angle of the nth harmonic, and $w(t)$ is the Gaussian white noise. However, imperfections in the manufacturing and assembly of the transmission lead to amplitude and phase modulation of the signal. These modulations have the effect of creating the sidebands about the regular meshing components. Thus, McFadden noted that the amplitude modulation function $a_n(t)$ and the phase modulation function $b_n(t)$ are periodic with the gear rotational frequency f_r and can be written as a sum of its harmonics as

$$a_n(t) = \sum_{m=0}^{\infty} A_{nm} \cos(m\omega_r t + \alpha_{nm}), \quad b_n(t) = \sum_{m=0}^{\infty} B_{nm} \cos(m\omega_r t + \beta_{nm}), \tag{2,3}$$

where $\omega_r = 2\pi f_r$. Substituting these modulation functions into $x_{\text{perfect}}(t)$ yields a more accurate model of a transmission vibration signal, given as

$$x(t) = \sum_{n=0}^{\infty} P_n[1 + a_n(t)] \cos[n\omega t + \phi_n + b_n(t)] + w(t). \tag{4}$$

Stewart [10] observed that in a healthy transmission, the amplitude of the sidebands, and thus the amplitude of the modulation functions, was small. However, he noted that the presence of a fault superimposes additional dynamics upon this signal and that different types of faults have different effects on the vibration signal. In major tooth faults, for example, the peak-to-peak value of $x(t)$ tends to increase. However, for heavily distributed damage, the peak-to-peak value can remain rather constant while the amplitude of the mesh frequency and its harmonics, P_n, decrease. He also made the important observation that the additional dynamics caused by the appearance of a fault tend to significantly increase the amplitude of the sidebands, in particular, in the presence of a local defect such as a single tooth failure. This corresponds to an increase in the amplitude of the modulation functions in McFadden's model.

To better observe the changes in the amplitude of the modulation functions, Stewart proposed the formation of a difference signal d, the statistical properties of which could then be analyzed. The difference signal is given as

$$d = x(t) - y_d(t), \tag{5}$$

where $x(t)$ is the original time synchronous signal and $y_d(t)$ is the signal containing the mesh frequencies, their harmonics and their first-order sidebands. Thus, according to McFadden's model, d is composed of the higher-order sidebands and Gaussian noise. Later, it was proposed that the first-order sidebands need not be removed [12]. This led to the formation of the residual signal, r, given as

$$r = x(t) - y_r(t), \tag{6}$$

where, in this case, $y_r(t)$ is the signal containing only the mesh frequencies and their harmonics.

These observations of the nature of the changes in the transmission vibration signal in the presence of damage form the foundation of the statistical transmission damage detection metrics which follow.

2.2. Diagnostic techniques

2.2.1. Root mean squared

The root mean squared (rms) [13] is defined as the square root of the average of the sum of the squares of the signal samples and is given by

$$\mathrm{RMS}_x = \sqrt{\frac{1}{N}\left[\sum_{i=1}^{N}(x_i)^2\right]}, \tag{7}$$

where x is the original sampled time signal, N is the number of samples and i is the sample index. The rms of a sine wave is defined as 0.707 times the amplitude of the signal.

2.2.2. Crest factor

The crest factor (CF) [14] is defined as the maximum positive peak value of the signal x divided by RMS_x and is given by

$$\mathrm{CF} = \frac{x_{0-pk}}{\mathrm{RMS}_x}, \tag{8}$$

where pk is the sample index for the maximum positive peak of the signal and x_{0-pk} is the value of x at pk. CF is a normalized measurement of the amplitude of the signal and is designed to increase in the presence of a small number of high-amplitude peaks, such as those caused by some types of local tooth damage.

2.2.3. Energy ratio

The energy ratio (ER) [14] is defined as the ratio of the RMS of the difference signal d and the RMS of the signal containing only the regular meshing components, y_d, and is given by

$$\mathrm{ER} = \frac{\mathrm{RMS}_d}{\mathrm{RMS}_{y_d}}. \tag{9}$$

ER is designed to increase in the presence of heavy uniform wear since it would be expected that in this case RMS_d would increase while RMS_{y_d} would decrease.

2.2.4. FM0

The parameter FM0 was developed by Stewart in 1977 as a robust indicator of major faults in a gear mesh [10]. Major changes in the meshing pattern are detected by comparing the maximum peak-to-peak amplitude of the signal to the sum of the amplitudes of the mesh frequencies and their harmonics. FM0 is given as

$$\mathrm{FM0} = \frac{\mathrm{PP}_x}{\sum_{n=0}^{H} P_n}, \tag{10}$$

where PP_x is the maximum peak-to-peak amplitude of the signal x, P_n is the amplitude of the nth harmonic, and H is the total number of harmonics in the frequency range. Notice that in cases where PP_x increases while P_n remains relatively constant, FM0 increases. Also, if P_n decreases while PP_x remains constant, FM0 also increases.

2.2.5. Kurtosis

The kurtosis is the fourth normalized moment of a given signal x and provides a measure of the peakedness of the signal, i.e. the number and amplitude of peaks present in the signal. It is given by

$$\mathrm{Kurtosis} = \frac{N \sum_{i=1}^{N} (x_i - \bar{x})^4}{\left[\sum_{i=1}^{N} (x_i - \bar{x})^2\right]^2}. \tag{11}$$

The fourth moment is normalized by the square of the variance. A signal consisting exclusively of Gaussian noise will have a kurtosis of approximately 3.

2.2.6. Energy operator

The energy operator [15] is found by first calculating the value $x_i^2 - (x_{i-1} * x_{i+1})$ for every point x_i of the signal x. At the end points, the signal is assumed to be a continuous loop. The energy operator is then computed by taking the normalized kurtosis of the resulting signal.

2.2.7. FM4

Developed by Stewart in 1977, the parameter FM4 was designed to complement FM0 by detecting faults isolated to only a limited number of teeth [10]. This is accomplished by first constructing the difference signal, d, given in Eq. (5). The normalized kurtosis of d is then computed. FM4 is given as

$$\mathrm{FM4} = \frac{N\sum_{i=1}^{N}(d_i - \bar{d})^4}{\left[\sum_{i=1}^{N}(d_i - \bar{d})^2\right]^2}, \tag{12}$$

where $\bar{d}$ is the mean of the difference signal, and N is the total number of data points in the time signal. FM4 is nondimensional and designed to have a nominal value of 3 if d is purely Gaussian. When higher-order sidebands appear in the vibration signal, FM4 will deviate from this value.

2.2.8. NA4

The parameter NA4 was developed in 1993 by Zakrajsek, Townsend, and Decker at the NASA Lewis Research Center as a general fault indicator which reacts not only to the onset of damage as FM4 does, but also to the continuing growth of the fault [12]. The residual signal r, given in Eq. (6), is first constructed. The quasi-normalized kurtosis of the residual signal is then computed by dividing the fourth moment of the residual signal by the square of its run time averaged variance. The run time averaged variance is the average of the residual signal over each time signal in the run ensemble up to the point at which NA4 is currently being calculated. NA4 is given as

$$\mathrm{NA4}(M) = \frac{N\sum_{i=1}^{N}(r_{iM} - \bar{r}_M)^4}{\left\{\frac{1}{M}\sum_{j=1}^{M}\left[\sum_{i=1}^{N}(r_{ij} - \bar{r}_j)^2\right]\right\}^2}, \tag{13}$$

where $\bar{r}$ is the mean of the residual signal, N is the total number of data points in the time signal, M is the number of the current time signal, and j is the index of the time signal in the run ensemble. Like FM4, NA4 is nondimensional and designed to have a nominal value of 3 if r is purely Gaussian.

2.2.9. M6A

The parameter M6A was proposed by Martin in 1989 as an indicator of surface damage on machinery components [16]. The underlying theory is the same as that of FM4. However, it is expected that M6A will be more sensitive to peaks in the difference signal due to the use of the sixth moment. M6A is given as

$$\mathrm{M6A} = \frac{N^2\sum_{i=1}^{N}(d_i - \bar{d})^6}{\left[\sum_{i=1}^{N}(d_i - \bar{d})^2\right]^3}. \tag{14}$$

Note that in this case, the moment is normalized by the cube of the variance.

2.2.10. M8A

The parameter M8A, also proposed by Martin in 1989, is designed to be yet more sensitive than M6A to peaks in the difference signal [16]. M8A uses the eighth moment normalized by the variance to the fourth power and is given as

$$\mathrm{M8A} = \frac{N^3\sum_{i=1}^{N}(d_i - \bar{d})^8}{\left[\sum_{i=1}^{N}(d_i - \bar{d})^2\right]^4}. \tag{15}$$

It should be mentioned that increased sensitivity to peaks is not necessarily a desired property in that an excessively sensitive parameter will yield many false alarms. Hence, the choice of damage indicator is not related exclusively to sensitivity.

2.2.11. NB4

The parameter NB4 was developed by Zakrajsek, Handschuh and Decker in 1994 as an indicator of localized gear tooth damage [17]. The theory behind NB4 is that damage on just a few teeth will cause transient load fluctuations different from those load fluctuations caused by healthy teeth, and that this can be seen in the envelope of the signal. As with NA4, NB4 uses the quasi-normalized kurtosis. However, instead of the difference signal, NB4 uses the envelope of the signal band-pass filtered about the mesh frequency. The envelope, s is computed using the Hilbert transform and is given by

$$s(t) = |[b(t) + i[H(b(t))]]|, \tag{16}$$

where $b(t)$ is the signal band-pass filtered about the mesh frequency, $H(b(t))$ is the Hilbert transform of $b(t)$, and i is the sample index.

2.2.12. NA4*

The parameter NA4* was developed in 1994 by Decker, Handschuh and Zakrajsek as an enhancement to NA4 [18]. In this case, the denominator of NA4 is statistically modified, i.e. when the variance of the residual signal exceeds a certain statistically determined value, the averaging stops and the denominator is locked. This modification was made based on the observation that as damage progresses from localized to distributed, the variance of the signal increases significantly, causing the kurtosis to settle back to nominal values after the initial indication of the onset of damage. By normalizing the fourth moment by the variance of a baseline signal from the transmission operating under nominal conditions, NA4* is provided with enhanced trending capabilities.

Since it was observed that the variance of a damaged transmission signal is greater than that of a healthy transmission signal, the decision to lock the denominator is made based on an upper limit, L, given by

$$L = \bar{v} + \frac{Z}{\sqrt{N}}\,\sigma, \tag{17}$$

where v is the mean value of previous variances, Z is the probability coefficient usually chosen for a normal distribution, σ is the standard deviation of the previous variances, and N is the number of samples. Z for a normal distribution can be found in any introductory statistics text. However, the actual choice of Z should be made based on experimentation as too small a value could lead to an over abundance of false alarms.

2.2.13. NB4*, M6A* and M8A*

As with the parameter NA4*, the parameters NB4* [19], M6A* [13] and M8A* use the statistically modified average variance to normalize their respective moments.

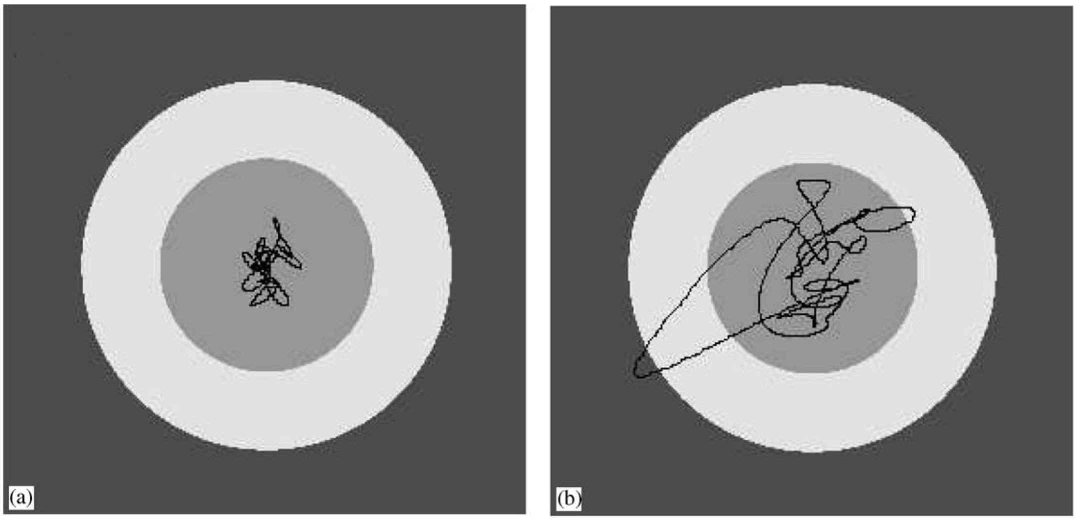

Fig. 3. Bulls-eye plots of Wessex main rotor gear box (WAK143) input pinions [22]; (a) undamaged. (b) 103 h prior to failure.

2.2.14. Demodulation

The original observation made by Stewart [10] that gear tooth damage causes an increase in the amplitude of the sidebands about the regular meshing components led to further investigations into the nature of the amplitude and phase modulation functions. It was proposed that the vibration signal could be demodulated to obtain separate approximations of the amplitude and phase modulation functions and that these approximations could subsequently be inspected to find early indications of gear damage [20,21]. This work was further refined by Blunt and Forrester [22] to produce a useful damage indicator referred to as a bulls-eye plot which indicates both amplitude and phase demodulations simultaneously. Fig. 3(a) shows a bulls-eye plot for a Wessex WAK143 main rotor gear box operating under healthy conditions, and Fig. 3(b) shows that for the same gear box with a cracked input pinion 103 h prior to complete failure. Note that the inner, middle, and outer regions of the plot indicate safe, caution and danger regions of operation, respectively. These figures clearly illustrate the changes in modulation due to damage.

3. Time–frequency analysis

3.1. Theory

In the late 1980s, investigations into the joint time–frequency characteristics of transmission vibration signals were begun [23]. It was observed that local gear faults such as spalling on a limited number of gear teeth produce sharp transients in the vibration signal of a transmission.

The statistical damage detection metrics tend to have an underlying assumption of signal stationarity and thus provide information about a signal which is averaged over time. As a result, the effect of local transient phenomena can be lost. On the other hand, nonstationary techniques provide information about the local time-domain properties of a signal. In general, a transmission vibration signature consists of three significant components: a sinusoidal component due to time varying loading, a broad-band impulsive component due to impact, and random noise. For an undamaged transmission, the sinusoidal components dominate. However, as damage propagates through the system, the sinusoidal components exhibit both modulation and reduction in amplitude [20,24]. In addition, both the broad-band impulsive components and the random noise become more prevalent [25]. It can be seen that the trends exhibited by the sinusoidal components are most visible in the frequency domain; the trends exhibited by the broad-band impulsive components are most visible in the time domain. Thus, in order to capture these trends, nonstationary, time–frequency analysis techniques are deemed appropriate [26].

Initial research into the use of joint time–frequency signal processing techniques for transmission damage detection was focused primarily on two of the standard time–frequency analysis techniques available in the signal processing community, the spectrogram and the Wigner–Ville distribution (WVD).

The spectrogram is the squared magnitude of the short-time Fourier transform (STFT) and provides the energy density spectrum of the signal as a function of time [27]. To investigate the frequency domain properties of a signal about time t, the signal $x(t)$ is first multiplied by a window function $h(t)$, centered at t. This leads to the formation of a windowed signal given as

$$x_\tau(t) = x(\tau)h(\tau - t). \tag{18}$$

The window function is chosen such that

$$x_\tau(t) \sim \begin{cases} x(\tau) & \text{for } \tau \text{ near } t, \\ 0 & \text{for } \tau \text{ far from } t. \end{cases} \tag{19}$$

The STFT about time t is the Fourier transform of $x_\tau(t)$, given as $X_t(\omega)$, and the energy density spectrum about time t is the squared magnitude of $X_t(\omega)$. Thus, the energy density spectrum about time t is given as

$$P_{\mathrm{SP}}(t,\omega) = |X_t(\omega)|^2 = \left|\frac{1}{\sqrt{2\pi}}\int \mathrm{e}^{-\mathrm{j}\omega\tau}\,x(\tau)h(\tau - t)\,\mathrm{d}\tau\right|^2. \tag{20}$$

For each time t, a different spectrum is obtained. These spectra can be combined to form the time–frequency distribution P_{SP} referred to as the spectrogram.

The WVD is a qualitatively different distribution than the spectrogram, even though it falls into the same general class of representations [27]. To obtain the WVD at time t, take the sum of the product of the signal before time $t, x(t - 1/2\tau)$ and the reversed signal after time t, $x(t + 1/2\tau)$. The reversal of the signal after t ensures that at time t, the points being multiplied are equidistant in time from t, specifically $\pm 1/2\tau$. Thus the WVD is given as

$$W(t,\tau) = \frac{1}{2\pi}\int x^*(t - 1/2\tau)x(t + 1/2\tau)\mathrm{e}^{\mathrm{j}\tau\omega}\,\mathrm{d}\tau. \tag{21}$$

Effectively, the computation of the WVD at time t can be thought of as folding the signal onto itself at t and determining if there is any overlap. It should be noted that times far from t are weighted equally with times near t. Thus the WVD is highly nonlocal.

An important property of the WVD that must be mentioned is the presence of cross-terms. Consider the WVD of the sum of two signals

$$x(t) = x_1(t) + x_2(t). \tag{22}$$

Substituting Eq. (22) into Eq. (21) it can be seen that

$$W(t,\omega) = W_{11}(t,\omega) + W_{22}(t,\omega) + W_{12}(t,\omega) + W_{21}(t,\omega) \tag{23}$$

and given that $W_{12}(t,\omega) = W^*_{21}(t,\omega)$, it is evident that

$$W_{12}(t,\omega) + W_{21}(t,\omega) = 2\,\mathrm{Re}\{W_{12}(t,\omega)\}, \tag{24}$$

where $2\,\mathrm{Re}\{W_{12}(t,\omega)\}$ is referred to as the cross-term. Thus

$$W(t,\omega) = W_{11}(t,\omega) + W_{22}(t,\omega) + 2\,\mathrm{Re}\{W_{12}(t,\omega)\}. \tag{25}$$

From Eq. (25) it can be observed that any signal which can be composed of the sum of two signals may have cross-terms present in the distribution. While these terms may be useful in some cases, e.g. for the identification of multicomponent signals [27], they tend to interfere with the interpretation of the distribution of transmission vibration signals. Despite this limitation, the WVD has been demonstrated to have utility in the transmission diagnostics.

The spectrogram and the WVD have both been investigated as potential time–frequency analysis tools for transmission diagnostics. These investigations are summarized below.

3.2. Diagnostic techniques

A significant advancement in the field of transmission diagnostics was made when Forrester introduced the use of time–frequency analysis of transmission vibration signals [23,28]. In this pioneering work, he used the WVD to process vibration data from a damaged Royal Australian Navy Wessex helicopter main rotor transmission. It was demonstrated that the WVD of the signal average, with no further enhancements or filtering, was able to reveal the fault. Fig. 4 shows the WVD for a Wessex WAK143 main rotor gear box with a cracked input pinion 42 h prior to failure. The disturbance around a frequency of 44 shaft-orders near a shaft angle of 100° is a result of the damage.

This work was followed by McFadden in the early 1990s. He investigated the properties of the WVD for both continuous and discrete signals [29]. Specifically, he investigated the effect of the cross-terms on the distribution computed from a transmission vibration signal. He subsequently applied the WVD to experimentally obtained time-domain averages of the vibration of damaged gears in a helicopter transmission, in order to detect early signs of damage. It was shown that the damage to the gears can be detected by visual inspection of the changes occurring in the pattern of the WVD.

Further research by McFadden demonstrated that the application of sliding weighting functions to the WVD could reduce the effect of the cross-terms [30,31]. Subsequently, he showed that gear damage in a helicopter transmission can be more easily detected by visual inspection of the changes occurring in the pattern of the weighted WVD more easily than they could with the original WVD.

Additional assessment of the WVD as a transmission damage detection technique was performed at the NASA Glenn Research Center [32,33]. They concluded that the WVD provided

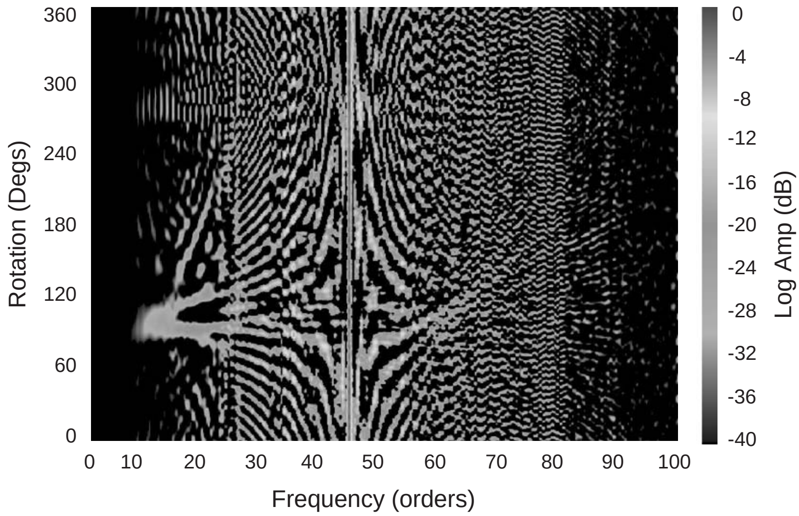

Fig. 4. WVD plot of Wessex transmission 42 h prior to failure [22].

vital information regarding the location and severity of gear tooth damage. Since this early work, a number of variances of the WVD have been investigated [34]. Research into the WVD ultimately led to the development of the gear diagnostic parameter NP4 [35,36]. The creation of a diagnostic parameter is desirable as it eliminates the requirement for visual inspection and interpretation of the distribution, enabling the use of a threshold for damage detection.

The parameter NP4 is the normalized kurtosis of the instantaneous signal power, $P_x(t)$ as computed using the WVD. The instantaneous signal power is given as

$$P_x(t) = \int |W(t,\omega)| \mathrm{d}\omega. \tag{26}$$

Thus, the parameter NP4 can be defined as

$$\mathrm{NP4} = \frac{N \sum_{i=1}^{N} (P_i - \bar{P})^4}{\left[\sum_{i=1}^{N} (P_i - \bar{P})^2\right]^2}, \tag{27}$$

where i is the sample index, P_i is the instantaneous signal power at sample index i, $\bar{P}$ is the mean of P, and N is the total number of data points in the time signal. NP4 is nondimensional and depends only on the shape of the power distribution.

Soon after the potential of time–frequency analysis was demonstrated through the use of the WVD, research began into the application of the spectrogram. McFadden applied the spectrogram to the calculation of the time–frequency distribution of a transmission vibration signal [37]. The selection of the window function was investigated, and the Gaussian function was found to be well suited to the calculation of the distribution. The spectrogram was then used to analyze experimentally obtained vibration signals from an industrial gearbox. McFadden concluded that the spectrogram was a potentially useful tool for the early detection of local gear damage. Further work demonstrated that the spectrogram could be treated as an image and that

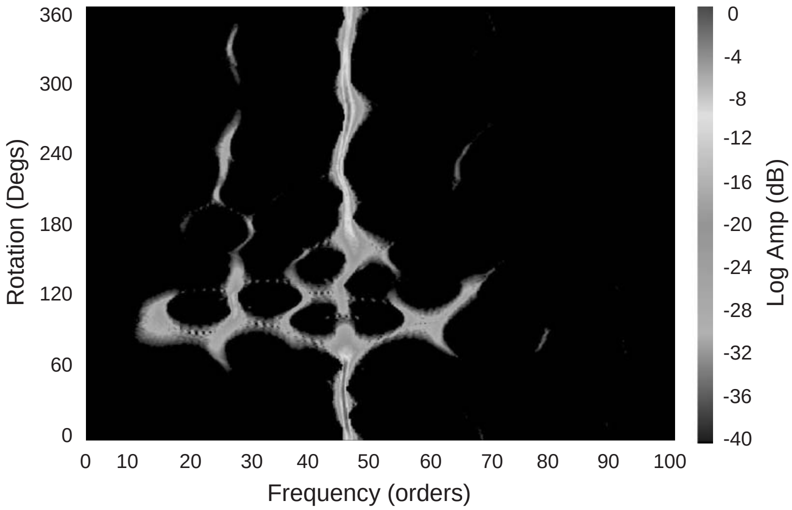

Fig. 5. Modified spectrogram plot of Wessex transmission 42 h prior to failure [39].

fault detection could then be accomplished using image processing techniques for feature extraction [38].

Subsequent research into the spectrogram was performed by Forrester in the mid 1990s [39]. He demonstrated that a modified spectrogram could be as sensitive to transmission faults as the WVD while eliminating the effect of the cross terms. Fig. 5 shows the modified spectrogram for a Wessex WAK143 main rotor gear box with a cracked input pinion 42 h prior to failure. Note that the modified spectrogram indicates damage more clearly than the WVD, shown in Fig. 4.

4. Wavelet analysis

4.1. Theory

The advent of wavelet analysis in the mid 1980s provided a new and powerful tool for the analysis of the joint time–frequency characteristics of transmission vibration signals. The wavelet transform is a time–frequency analysis method similar to the WVD, though it is independent of time and thus can truly describe local behavior. Although wavelet analysis falls under the general heading of time–frequency analysis, it merits an independent section due to the large volume of research into its use as a diagnostic technique.

Unlike the Fourier transform, where stationary complex basis functions are used to map the temporal signal into the frequency domain, the wavelet transform uses a new class of real and complex nonstationary basis functions, termed wavelets, which can be independently dilated and shifted as a function of time, to create a unique time–frequency map. The advantage of this method is that the frequency content of the signal can be analyzed without the loss of vital time-domain information. In addition, time and frequency are independent. Thus, the time–frequency representation is not affected by cross-terms.

The roots of the wavelet transform reach back to the development of quadrature mirror filters (QMF) in the signal processing community [40–44]. In the mathematics community, the classical definition of wavelets as translations and dilations of a single basis function was inaugurated in several key references [45–48]. The introduction of the fast wavelet transform and multiresolution analysis (MRA) provided a link between wavelets and QMFs [47,49,50]. This work laid the foundation for further development of wavelet theory [46,51,52].

The wavelet transform is a signal processing tool which allows both the time- and frequency-domain properties of a signal to be viewed simultaneously. At first glance this would seem to violate the uncertainty principle which states that, for a signal $x(t)$, the time density $|x(t)|^2$ and frequency density $|X(\omega)|^2$ cannot be made arbitrarily narrow simultaneously [27]. However, wavelet analysis offers a tradeoff: some loss of density in both time and frequency is tolerated so that time and frequency domain properties of the signal can be analyzed concurrently.

Performing a wavelet transform consists of convolving a signal with time shifted and dilated versions of a nonstationary wavelet basis function. Let the function $\psi_\lambda(t)$ be defined as a wavelet, where λ belongs to some set of indices Λ and describes the translation and dilation of the wavelet. Then the basis Ψ is given as $\Psi = \{\psi_\lambda | \lambda \in \Lambda\}$. After a basis has been chosen, the signal x can then be represented as a combination of wavelets, referred to as a wavelet series and given as

$$x = \sum_{\lambda} c_\lambda \psi_\lambda \tag{28}$$

with $c = \{c_\lambda | \lambda \in \Lambda\}$. The result of this process is a set of coefficients c_λ which are a function of time and frequency, or time and scale. Note that the change in scale of a wavelet, referred to as dilation, corresponds to a change in frequency; as a wavelet is dilated, i.e. stretched, its frequency decreases. The coefficients obtained from the wavelet transform can be used to form a mean square wavelet map which is a time–frequency representation of the energy of a signal. In addition, the coefficients computed as a wavelet of a fixed dilation is shifted along the time signal can be used to construct a new time signal. This new time signal will contain only those frequency components included in the wavelet. The continuous wavelet transform (CWT) requires that these coefficients be computed over all time and frequency. However, it has been found that if the dilation and position of the wavelets are chosen to be powers of two, then the transform can be computed much more efficiently with no loss of accuracy. This form of the wavelet transform is known as the discrete wavelet transform (DWT). It should be noted that the CWT is a non-orthogonal transformation whereas the DWT is orthogonal.

In general, wavelet basis functions, including the standard set developed by Daubechies [46], cannot be represented analytically, and thus even the DWT is computationally inefficient. However, in 1989, Mallat [49] presented a relatively fast, recursive algorithm for computing the DWT. This algorithm uses the wavelet basis function to form sets of filters, each consisting of a high-pass filter and a lowpass filter. The signal under consideration is then passed through the first set of these filters and downsampled. The result is two signals, each with about half the number of coefficients as the original signal. The first of these signals, which was formed using the lowpass filter and which consequently contains the low-frequency information, is referred to as the approximation (A_1). The second signal, which was formed using the highpass filter and which consequently contains the high-frequency information, is referred to as the *detail* (D_1). For the second recursion, the approximation is passed through the next set of highpass and lowpass filters, again based on the wavelet basis function, yielding a second approximation (A_2) and detail (D_2). This decomposition can be repeated until an approximation and detail, each consisting of only one coefficient, are formed.

Research into the use of the wavelet transform for transmission diagnostics has focused primarily on the assessment of the performance of various basis functions and various forms of the transform. Typically, a set of vibration signals is analyzed by the technique under consideration and the resulting distribution is visually inspected to determine whether changes due to damage are evident. However, research is beginning to focus on more intelligent applications of wavelets in an attempt to exploit the flexibility offered by the transform. Over the past 10 years, many papers have focused on the use of the wavelet transform for transmission diagnostics. A selection of significant research is summarized below.

4.2. Diagnostic techniques

Initial investigations into the use of the wavelet transform for transmission diagnostics began in the early 1990s [53–55], soon after research into the spectrogram and WVD demonstrated the potential of joint time–frequency vibration signal analysis techniques. This early research showed that the wavelet transform had the potential to out-perform the earlier time–frequency analysis techniques when applied to the problem of transmission diagnostics.

Ensuing investigations focused on the choice of basis function and the merits of the CWT versus the DWT. A comparison of compact and smooth wavelets for transmission fault detection was presented in 1995 by Wang and McFadden [56]. Generally, compact wavelets are not smooth and as a result tend to be better for lower accuracy approximations. On the other hand, smooth wavelets are better for achieving high numerical accuracy in signal reconstruction [57]. Both the highly compact Daubechies 4 wavelet and a smooth wavelet known as the harmonic wavelet, developed by Newland [58–60], were used in conjunction with the DWT to process vibration signals from a Westland helicopter transmission. It was demonstrated that the harmonic wavelet was slightly better suited to fault detection than the Daubechies 4 wavelet. However, Wang and McFadden determined that, for transmission damage detection, the nonorthogonal CWT with a Gaussian enveloped oscillation as a basis was more effective than the orthogonal DWT used with either the Daubechies or harmonic basis. The resulting work with the CWT was reported in 1996 [61].

Further investigation into the harmonic wavelet was performed by Samuel and Pines [62,63]. It was shown that in conjunction with the wavelet packet transform and some a priori knowledge of the fundamental frequencies of the transmission under consideration, the harmonic wavelet could be used to monitor the energy of the regular meshing components and the sidebands as a function of time. This research led to the development of a wavelet-based damage detection metric referred to as the normalized energy metric (NE) [26,64,65]. NE was based on the observation that with the onset of damage, energy tends to transfer from the regular meshing components to the sidebands. It was shown that NE could be used to detect and classify certain types of gear damage in helicopter transmissions.

The evaluation of the wavelet transform as a transmission diagnostic technique is an ongoing field of research. As new wavelet techniques emerge from the signal processing, image processing and mathematics communities, their merit as a potential part of a transmission diagnostic system is assessed. Many variations on the wavelet transform have been investigated and the results presented in the literature. A selection of these more recent investigations is included in the bibliography [24,66–71].

5. Neural networks

5.1. Theory

The majority of diagnostic techniques presented in the literature have demonstrated some ability to indicate the presence of damage in a transmission. However, each technique has associated strengths and limitations and is usually suited to the detection of a limited number of fault types. Thus, any one technique is generally insufficient for detecting all of the various types of damage that could occur in a transmission.

Model-based techniques offer a fundamentally different approach to the problem of transmission damage detection. The fundamental idea is that the technique is trained to recognize signals from a healthy transmission and then indicate when the vibration signal deviates from this nominal condition. In addition, some measure of the nature of the deviation can often be indicated, thus potentially enabling the type of damage to be determined.

Neural networks are the most commonly used technique for model-based diagnostics. A neural network is defined as a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use [72]. It possesses two fundamental properties: (1) the network obtains knowledge through a learning process, and (2) the interneuron connection strengths or weights are used to retain the knowledge. The fundamental theory underlying neural networks is extensive and has been well documented in numerous books and papers in the literature [72,73].

In general, a neural network damage detection system does not process the vibration signal directly. Instead, it takes as input the results of various processing techniques, learning how the techniques behave in the presence of healthy and faulty data. For example, vibration data could be processed using the statistical metrics presented above, and the output of these techniques could then be used as input for the network. This would allow the network to take advantage of the strength of each individual damage detection technique. The resulting trained network acts as a model of the transmission vibration data, and can be used to detect abnormalities in the signal associated with damage. In addition, neural networks offer the added benefit of fault classification as long as the processed data behaves differently in the presence of different types of damage.

Two standard neural network architectures are used in transmission diagnostics. These are the feed-forward backpropagation network and the Kohonen feature map, also referred to as the self organizing map (SOM).

The feed-forward backpropagation neural network, more simply referred to as the back-propagation network, typically consists of multiple layers of interconnected nodes: the input layer, one or more hidden layers, and the output layer. A diagram of the architecture of a typical backpropagation network is shown in Fig. 6. The backpropagation network can be trained using healthy data and data from different types of faults. It can then be used to determine the condition of future data, and classify any detected faults. If designed and trained correctly, a backpropagation network has the ability to generalize [72], which could allow the network to indicate the presence of faults slightly different from those encountered in the training set. This property could be particularly useful when multiple faults are encountered simultaneously. Training a backpropagation network must be supervised, i.e. when a training data set is input, a corresponding target output must also be provided.

The Kohonen feature map, on the other hand, uses unsupervised learning. It consists of two layers of nodes: an input layer and an output layer. The output layer takes the form of a two-dimensional grid and each grid node is connected to every node of the input layer. A diagram of the architecture of a typical Kohonen feature map is shown in Fig. 7. The organization of the

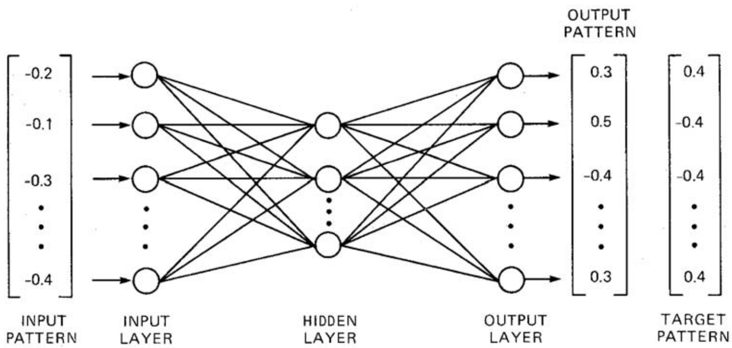

Fig. 6. Backpropagation network [73].

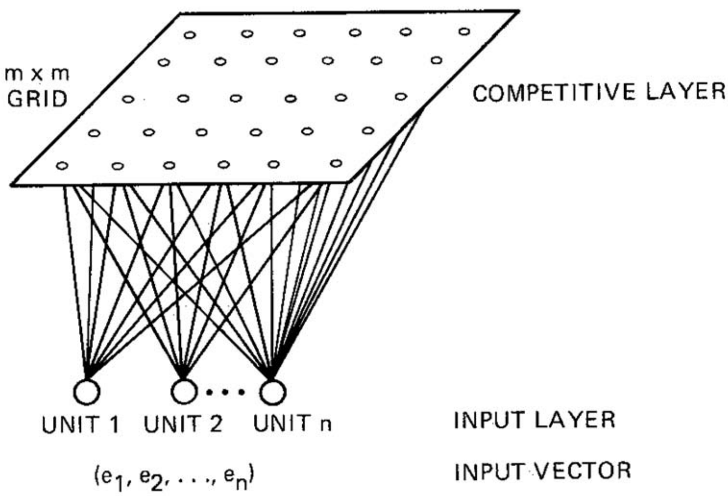

Fig. 7. Kohonen feature map [73].

training set is determined by the network. Data sets with different properties are mapped to different locations on the output grid. Unsupervised learning is a desirable property as it simplifies the training process. However, the backpropagation network is more powerful for classification.

Both architectures are commonly used in neural-network-based diagnostic systems. A selection of significant research into the use of neural networks for damage detection is summarized below.

5.2. Diagnostic techniques

In 1993, Harris proposed the use of the SOM for the detection and classification of damage in rotating machinery [74]. Input to the network consisted of frequency spectrum values at salient operating frequencies of the machine under consideration. Only data from the machine operating under nominal conditions was used for training. It was shown that the network could distinguish between vibration signals obtained from the machine under healthy and damaged conditions. In addition, it was proposed that limited damage classification could be performed if the network was trained using both healthy and faulty data. However, it was noted that faulty data are often difficult to obtain.

Subsequent work by MacIntyre et al. demonstrated that backpropagation neural networks could be used to effectively detect and classify damage in rotating machinery [75]. Just as in the work of Harris, salient frequency spectrum values were used as input and excellent classification performance was reported. However, the requirement for supervised learning was noted as a limitation of the system. Hence, the SOM was again investigated. This technique proved useful for the detection of faults, however, the inability to classify faults was noted as a limitation. Finally, the use of a hybrid neural network architecture was proposed. For this system, the output of the SOM would be used as input for the backpropagation network. It was suggested that this hybrid network could address the limitations associated with each individual neural network.

In 1996, Lopez et al. [76] demonstrated that helicopter transmission damage detection and classification could be performed by a backpropagation neural network-based diagnostic system. Input for the system was vibration data processed using the continuous wavelet transform. It was shown that for multiple sensors, a network could be trained for each sensor, and then a separate back-propagation neural network, termed a hierarchical neural network, could be used to perform sensor fusion and make a final decision regarding the health of the transmission.

The hybrid neural network architecture proposed by MacIntyre et al. [75] has become the most commonly used architecture for transmission damage detection and classification. Recent work by Essawy et al. demonstrated that a hybrid neural network architecture could be used to detect damage in a helicopter transmission [77]. Multiple sensors were located at various positions on the transmission housing. A neural network was trained for each sensor using vibration data that was preprocessed using the wavelet transform. The output from each network was then fed into a fuzzy logic-based sensor fusion algorithm which would make the final decision regarding the condition of the gearbox. Excellent classification results were reported. Also, in 2001, a hybrid neural network architecture was used by Samuel and Pines to demonstrate that the Normalized Energy (NE) damage detection metric, a metric based on the wavelet transform, could be used as input for a neural-network-based damage detection technique [65].

Investigations into the use of other neural network architectures for damage detection and classification have also been presented. In 1991, Chin and Danai [78,79] proposed a fault detection and classification system based on a nonparametric classifier that used a multi-valued influence matrix (MVIM) as its diagnostic model. This technique constitutes a simple neural network architecture that incorporates a fast learning algorithm based on error feedback so that only a limited training set is necessary. The MVIM provides indices describing both the level of diagnosis possible for the process and variability of the fault signatures. These indices are used as feedback. The MVIM is trained using an extensive set of statistical fault indicators and supervised learning is required. Further research into this technique demonstrated its potential as a damage detection and classification system [80–82]. Subsequent work by Jammu et al. [83–85] resulted in the development of an unsupervised diagnostic system referred to as a structure-based connectionist network (SBCN). The connectionist network used in this research consists of a single layer neural network. Weights are influenced by both the features associated with damage, termed featural influences, and the proximity of the sensor under consideration to the potentially faulty component, termed structural influences. Once again, the input to the network consisted of an extensive set of statistical indicators, computed from vibration data collected from multiple sensors on a healthy gearbox. The SBCN was shown to be more versatile than the MVIM method at detecting and classifying transmission faults since it does not depend on supervised learning.

6. Mathematical modelling

6.1. Theory

The idea behind mathematical modelling techniques is similar to that of neural network techniques. If vibration data collected from a healthy transmission can be sufficiently modelled, then the model can be used to filter vibration data collected from the operational transmission to indicate changes in the vibration signal due to damage. The primary ideological difference between mathematical modelling diagnostic techniques and neural network techniques is the form in which the data is input. As show above, neural network techniques are typically trained on data that have been somehow preprocessed, e.g. on the output of many different diagnostic metrics and signal processing techniques. On the other hand, mathematical modelling diagnostic techniques process the vibration data directly.

There are two primary mathematical modelling techniques that have been applied to transmission diagnostics. These are autoregressive (AR) modelling and autoregressive moving average (ARMA) modelling. ARMA modelling is a generalization of AR modelling. Both techniques are referred to as parametric spectral estimation procedures, and were developed to compute high-resolution spectral estimations of data for which a sufficiently long time record is difficult or impossible to acquire [86].

An AR model seeks to represent a time series x by a linear regression of x on itself plus an error series assumed to be noise with a Gaussian distribution. This model is given as

$$x_i = -\sum_{k=1}^{p} a_k x_{i-k} + e_i, \tag{29}$$

where p is the number of past inputs required to sufficiently model the signal, referred to as the model order, a_k are the AR coefficients, i is the sample index, and e_i is the Gaussian error series. The frequency response function $H(z)$ of the linear system is then given as

$$H(z) = \frac{1}{1 + \sum_{k=1}^{p} a_k z^{-k}}, \tag{30}$$

where $z = \exp(-\mathrm{j}2\pi f \Delta t)$ is the z transform. This is referred to as an AR model of order p. The model order p must be determined experimentally. Once the model order is selected, one of a number of algorithms may be used to determine the AR coefficients. Many algorithms are available in the literature; a summary of the fundamental techniques is provided by Rofe [87], and a list of more advanced techniques is presented by Zhuge et al. [88]. The choice of algorithm is a topic of research. It is also interesting to note that an AR model can be interpreted as a single layer neural network [89].

An ARMA model is a generalization of the AR model. The frequency response function given in Eq. (30) can more generally be given as

$$H(z) = \frac{1 + \sum_{k=1}^{q} b_k z^{-k}}{1 + \sum_{k=1}^{p} a_k z^{-k}}. \tag{31}$$

The resulting data sequence model is given by

$$x_i = -\sum_{k=1}^{p} a_k x_{i-k} + \sum_{k=0}^{q} b_k e_{i-k}. \tag{32}$$

This is referred to as an ARMA model of order p and q. Note that if all b_k are zero except b_0, then the ARMA model reduces to the AR model. If all a_k are zero, then the model reduces to the moving average (MA) model. The coefficients b_k are referred to as the MA coefficients. It has been established that determination of the AR coefficients is a linear problem whereas determination of the MA coefficients is a nonlinear problem [90,91] since the model error $b_k e_{i-k}$ and hence the MA coefficients b_k cannot be computed until the model has been fitted and the AR coefficients have been determined. Thus, the AR technique has an computational advantage over the ARMA technique. Once the model order and model coefficients have been determined, the model must be validated.

Mathematical modelling has received minimal attention by the transmission diagnostic community. However, interest in the technique is beginning to grow. The significant research is summarized in the following section.

6.2. Diagnostic techniques

Initial consideration of AR modelling for rotating machinery diagnostics was presented by Gersch et al. in 1983 [92]. In this study, a nearest-neighbor time series classification rule methodology was developed. AR modelling was used to compute the dissimilarity measure between a vibration signal collected under a known operating condition and one collected under an unknown operating condition. Once the dissimilarity measures were determined, the nearest-neighbor rule methodology could be used to classify the damage. Satisfactory classification results were reported. In 1985, Kitagawa and Gersch [93] presented a more in-depth consideration of AR modelling. This work investigated a time varying coefficient AR model approach for the modelling of nonstationary covariance time series. It was shown that complex time varying earthquake data could be characterized locally in time. It was then proposed that the technique could be used for machine diagnostics as it would enable the instantaneous spectra of a vibration signal to be estimated and inspected.

This research was extended by Zhuge et al. in 1990 [88] through a comparison between a time varying and time invariant techniques for AR coefficient estimation. The time varying technique investigated was based on Kalman filtering theory. This technique was found to be very expensive computationally. In comparison, the time invariant technique based on the assumption of local stationarity of the signal was found to be computationally inexpensive and thus fast to realize.

More recently, Ballie and Mathew presented a comparison of AR modelling, backpropagation neural networks and radial basis function networks for detecting faults in rolling element bearings [89]. For this study, the AR model coefficients were determined using a forward least squares algorithm. It was reported that the neural network technique generally outperformed the other two.

Rofe investigated ARMA modelling as a technique for fault detection that would compensate for changes in the transmission operating load [87]. It was shown through simulation that the ARMA modelling technique could effectively detect faults while compensating for load variations. This work was extended to AR modelling by Wang and Wong [94]. It was demonstrated through an analysis of vibration data collected from a transmission test rig that the AR modelling technique could be used to compensate for load variation while detecting gear faults. The computational advantages of AR modelling over ARMA modelling were discussed.

In 1997, Garga et al. proposed that for gear fault detection, only a small set of AR model coefficients was necessary [95]. It was demonstrated that the application of dimension reduction to the AR coefficient prior to diagnosis could significantly reduce the complexity of damage detection systems without a reduction in performance.

7. Adaptive signal processing

7.1. Theory

The advent of adaptive signal processing techniques enables the development of diagnostic techniques that combine both model-based techniques and joint time–frequency signal processing-based techniques. Specifically, the adaptive wavelet transform allows a given vibration signal to be analyzed using a set of wavelet basis functions chosen from a set of functions, termed a dictionary, consisting of multiple bases. The wavelet that best represents the signal at each point is chosen. This adaptation property enables the creation of a general diagnostic algorithm that is able to tailor itself to a specific transmission, thus potentially enhancing sensitivity to damage. Hence, a change in perspective is suggested for wavelet-based diagnostic algorithms: instead of using the wavelet transform to exclusively investigate changes in the characteristics of a vibration signal, a set of wavelets can be chosen that best represent a signal obtained from a known operating condition. This set of bases can then be compared to a signal obtained from the transmission during use and can thus be considered to be an empirical model of the transmission vibration signal. This approach is similar to that taken by AR- and ARMA-based diagnostic techniques. It should be noted that these adaptive signal processing techniques fall into the category of mathematical modelling along with AR and ARMA modelling. However, the flexibility of the wavelet transform allows more control over the adaptation process.

The goal of an adaptive representation algorithm is to find an approximation of a signal, in terms of a given over-complete dictionary of waveforms, that optimizes a desired characteristic. In an over-complete dictionary, some elements may have representations in terms of other elements and, as a consequence, any given signal representation is nonunique. Thus, the representation is referred to as adaptive, since there is the option of choosing the representation that best suits the application. Consider an over-complete dictionary $\mathbf{D}$ of waveforms ϕ_k where $k \in \mathbf{Z}$. Then, let Φ be the over-complete dictionary $\mathbf{D}$ made up of N bases $\Psi_b,\ b = 1, \dots, N$. Then the waveform ϕ_k is a wavelet $\psi_{\lambda,b}$ where $\lambda \in \Lambda$ and Λ is a two dimensional set of integers in $\mathbf{Z}^2$ with $\lambda = (j, l)$. An approximate, nonunique decomposition of a signal x formed using m wavelets, ϕ_k with $k = 1, \dots, m$ is then given as

$$x = \sum_{k=1}^{m} c_k \phi_k + R^{(m)}, \tag{33}$$

where c_k is the coefficient describing the waveform ϕ_k and $R^{(m)}$ is the residual, the difference between the signal approximation and the original signal. In general, no restrictions are placed on $\mathbf{D}$. For example, $\mathbf{D}$ could be composed of a combination of many smaller dictionaries. The result is a method for obtaining signal representations that is significantly more flexible than the classical wavelet transform. In addition, the resulting representations should be more sparse and thus more physically meaningful than typical wavelet transform representations. Consider a dictionary of wavelets constructed from transmission vibration data as mentioned in the previous section. Such an algorithm could decompose the vibration signal into a series of waveforms associated with either undamaged or damaged transmission components.

A number of methods for obtaining signal representations in over-complete dictionaries have been developed in recent years. These include the Method of Frames [96], Matching Pursuit [97], High Resolution Pursuit [98], and Basis Pursuit [99]. However, the majority of these techniques have yet to be applied to transmission diagnostics.

7.2. Diagnostic techniques

To date, minimal research has been focused on the use of adaptive signal processing techniques for transmission diagnostics. Liu and Ling [100] have shown that an extension of the Matching Pursuit algorithm in conjunction with a combination of classical wavelet dictionaries has the potential to be more effective than a single dictionary for rotating machinery diagnostics.

More recently, Samuel and Pines demonstrated that an adaptive version of the lifting scheme could be used to analyze individual tooth mesh waveforms in order to detect the presence of damage in a transmission [101]. The lifting scheme, a time-domain prediction-error realization of the wavelet transform, was developed by Wim Sweldens in the early 1990s as a method for creating new biorthogonal wavelets in settings where the Fourier transform could not be used

[102–104], such as on bounded domains and on curves and surfaces. Connections are cited between lifting and other independent research, particularly the work of Herley and Vetterli [105]. In later work, Sweldens and Daubechies demonstrated that all perfect reconstruction filterbanks could be formed with a sequence of lifting steps [106]. The lifting scheme provides a convenient framework for the implementation of new wavelet-based signal analysis techniques, and is easily modified for use in specific application areas with specific requirements. These characteristics make lifting an attractive option for the development of a wavelet-based signal modelling diagnostic technique.

A diagnostic technique based on the adaptive lifting work of Claypoole [107] was developed by Samuel and Pines [108]. In some cases, it was observed that the model, developed using healthy-state transmission vibration data, was able to effectively represent other sets of healthy-state data while failing to effectively represent the damaged-state vibration signal. The result was a small prediction error in the healthy-state case and a large prediction error in the damaged-state case. Thus it was shown that the prediction error could be potentially used to indicate the presence of damage.

It was subsequently found that, in many cases, adaptive lifting by itself yields a model that is too flexible to allow the effective distinction of healthy-state and damaged-state vibration signals. Hence, it was determined that adaptive lifting should be modified to improve the sensitivity to damage. Specifically, constraints on certain basis characteristics were deemed necessary to enhance the detection of local wave-form changes caused by certain types of gear damage. This type of adaptation, where in some cases the technique is constrained to the choice of some subset of bases within the dictionary, is referred to as *constrained adaptation*. Thus, an adaptive signal processing technique based on lifting and incorporating some form of constrains can be referred to as *constrained adaptive lifting* (CAL).

The goal of these constraints is to divide the tooth mesh waveform into simple, effectively monotonic functions, so that they can be approximated by low-order spline functions. The resulting model should be rigid to the appearance of higher-order features associated with damage, thus yielding large prediction errors:

- The individual tooth mesh wave-form is divided into analysis domains based on the slope inflection points of a healthy-state transmission vibration signal. The domains are fixed and used for subsequent analysis of future-state vibration signals. Lifting occurs only within each domain.
- The predictor P is constrained to be the lowest order least-squares spline approximation that provides a sufficiently small prediction error for a healthy-state transmission vibration signal. Both the order and the knots are fixed for the analysis of future-state vibration signals.

The adaptation process leading to the determination of the constrained properties and thus the development of the model occurs only offline. Once the model has been constructed, adaptation is no longer required. The model is then used to analyze vibration data collected from an operating transmission, and the prediction error vector is inspected for anomalies indicating damage.

In order to better quantify the prediction error vector generated by the CAL diagnostic algorithm, a metric, referred to as CAL4 was developed. Let the CAL prediction error vector for an individual tooth be denoted as ρ. Thus, for a gear g with N_g teeth, the prediction error vector ρ

for tooth P_n, $n = 1, \ldots, N_g$ is given as ρ_n. The prediction error vector for the entire gear is the then concatenation of the individual ρ_n, and is denoted $\mathbf{R}$. The metric CAL4 can be formed by taking the normalized kurtosis of $\mathbf{R}$. CAL4 is given as

$$\mathrm{CAL4} = \frac{N \sum_{i=1}^{N} (\mathbf{R}_i - \bar{\mathbf{R}})^4}{[\sum_{i=1}^{N} (\mathbf{R}_i - \bar{\mathbf{R}})^2]^2}, \tag{34}$$

where $\bar{\mathbf{R}}$ is the mean of $\mathbf{R}$ and N is the total number of data points in $\mathbf{R}$.

In CAL, the frequencies that are preserved in the prediction error vector are time varying, specifically, they vary from domain to domain. In addition, the frequencies which are preserved in $\mathbf{R}$ are chosen based on the characteristics of the individual tooth mesh waveform associated with the operation of the transmission, specifically, the conjugate action of the gear teeth. The set of bases model the vibration signal associated with healthy conjugate action, and thus, for a healthy transmission, the CAL prediction error is expected to contain only noise, and yield a value of 3.

It was demonstrated that the CAL diagnostic technique was capable of detecting damage on individual teeth of a planetary gearbox [101]. In addition, CAL4 was shown to be more effective than FM4 for the detection of tooth face damage.

8. Performance assessment of diagnostic techniques

Many of the transmission diagnostic techniques presented in this section have been validated using one of two primary transmission vibration data sets. These are referred to as the Westland data set and the OH-58 data set. The Westland data set was obtained as part of the Air Vehicle Diagnostics System (AVDS) project. Under this project, the Office of Naval Research funded a program of seeded fault testing on the aft-main transmission of the H-46 helicopter at Westland Helicopters in 1993 [109]. The OH-58 data set was collected from a test rig mounted OH-58 main transmission at the NASA Lewis Research Center in 1992 as part of a joint Navy/NASA/Army advanced lubrication evaluation program [7].

Unfortunately, a number of the techniques have been evaluated using internal transmission test rigs or proprietary data sets. Thus, the performance of these techniques cannot be compared using the results presented in the literature.

Research is continuing into comparing the performance of currently available transmission damage detection techniques [13,26,110]. However, there is currently no comprehensive, standard mechanism for the evaluation of new transmission diagnostic techniques. The need for a general means of comparison and evaluation has been recognized and is the focus of current research. A modular, real-time testbed for the evaluation of new damage detection techniques is under development by Conroy et al. [111]. This system is being developed using LabVIEW [112] and is designed to be general and easily modified so that any data format may be analyzed and damage detection techniques may be incorporated as a LabVIEW virtual instruments. A standard data set for the evaluation of transmission diagnostic systems is being created by Maynard et al. [113]. One H-46 aft transmission has been instrumented for initial tests and flight checks with successful results, and full squadron testing is planned. The Rotorcraft Industry Technology Association

(RITA), Inc. is creating a data base of vibration data for the evaluation of transmission diagnostic techniques. A standard method of the evaluation will be invaluable to the research community.

9. Emerging technologies

Research in the area of transmission health monitoring is constantly advancing the state of the art, and new technologies continue to emerge. In the area of vibration analysis, adaptive signal processing techniques have only begun to be explored, and new techniques must be investigated. For example, the Empirical Mode Decomposition (EMD) [114], a new signal processing technique that has proven useful in other signal processing applications such as financial time series analysis [115], beam and plate anomaly detection [116] and electrogastrogram (EGG) analysis [117], has yet to be evaluated. For a given signal, EMD yields a finite and often small number of Intrinsic Mode Functions (IMFs), where an IMF is defined as a function having a zero mean and an equal number of minima and maxima. Given the properties of an IMF, a Hilbert transform can be applied to obtain instantaneous frequencies as functions of time. The combination of EMD and the Hilbert transform is referred to as the Hilbert–Huang transform. The Hilbert–Huang transform is able to provide sharp indications of structures embedded within the signal such as those associated with gear tooth damage.

Another approach is a system that combines the output of vibration-based diagnostic techniques with diagnostics techniques based on data from various types of sensors such as oil debris monitors, load sensors and acoustic sensors, as well as other data types such as usage history. These techniques offer potentially significant improvements in robustness and accuracy. Thus, research is continuing into new sensor technologies such as acoustic arrays [118], fiber optic sensors, micro electric mechanical system (MEMS) wireless strain sensors and accelerometers, and embedded eddy current sensors that could provide new methods for transmission diagnostics which complement vibration signal analysis. Data fusion is drawing an increasing amount of attention as a method to process the output of multiple diagnostic algorithms. For example, the technique presented by Essawy et al. [77] should help prevent the occurrence of false alarms in on-board HUM systems. In addition, multidimensional health monitoring using techniques such as singular value decomposition (SVD) and principle component analysis (PCA) is being considered. For example, Cempel [119] has presented an introductory approach to multidimensional health monitoring of mechanical systems using SVD. Multidimensional health monitoring may offer more accurate diagnostics and lower false alarm rates than any single algorithm acting alone and thus is a promising area for future research.

10. Summary

In summary, over the past 25 years, much research has been devoted to the development of vibration-based transmission damage detection techniques. Many of these techniques have been shown to successfully detect damage in helicopter transmission systems, and some are incorporated in production HUM systems. However, there is still room for improvement, and thus research into new damage detection techniques is continuing.

References

[1] Health and monitoring system enhances helicopter safety, *Aviation Week and Space Technology* 139 (12) (1993) 94.

[2] J.D. Roe, D.G. Astridge, The health and usage monitoring of helicopter systems—the next generation, in: *Proceedings of the American Helicopter Society 41st Annual Forum*, Fort Worth, TX, 1985, pp. 175–187.

[3] D.C. Lombardo, Helicopter structures—a review of loads, fatigue design techniques and usage monitoring, Technical Report ARL-TR-15, Australian Department of Defense: Aeronautical Research Laboratory, May 1993.

[4] J. Land, C. Weitzman, How HUMS systems have the potential of significantly reducing the direct operating cost for modern helicopters through monitoring, in: *Proceedings of the American Helicopter Society 51st Annual Forum*, Fort Worth, TX, 1995, pp. 744–757.

[5] J.S. Kiddy, A Modal-base Damage Detection Technique for Helicopter Rotor Blades, PhD Thesis, University of Maryland, College Park, MD, USA, February 1999.

[6] F.D. Harris, E.F. Kasper, L.E. Iseler, US civil rotorcraft accidents, 1963 through 1997, Technical Report NASA/TM-2000-209597, USAAMCOM-TR-00-A-006, NASA and the US Army Aviation and Missle Command, December 2000.

[7] D.G. Lewicki, H.J. Decker, J.T. Shimski, Development of a full-scale transmission testing procedure to evaluate advanced lubricants, Technical Report NASA TP-3265, AVSCOM TR-91-C-026, NASA and the US Army Aviation Systems Command, August 1992.

[8] D.G. Astridge, Helicopter transmissions—design for safety and reliability, *Proceedings of the Institution of Mechanical Engineers. Part G, Journal of Aerospace Engineering* 203 (G2) (1989) 123–138.

[9] G. Marsh, The future of HUMS, *Avionics Magazine* (1996) 22–27.

[10] R.M. Stewart, Some useful analysis techniques for gearbox diagnostics, Technical Report MHM/R/10/77, Machine Health Monitoring Group, Institute of Sound and Vibration Research, University of Southampton, July 1977.

[11] P.D. McFadden, 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* 1 (2) (1987) 173–183.

[12] J.J. Zakrajsek, D.P. Townsend, H.J. Decker, An analysis of gear fault detection methods as applied to pitting fatigue failure data, Technical Report NASA TM-105950, AVSCOM TR-92-C-035, NASA and the US Army Aviation Systems Command, January 1993.

[13] H.J. Decker, D.G. Lewicki, Spiral bevel pinion crack detection in a helicopter gearbox, in: *Proceedings of the American Helicopter Society 59th Annual Forum*, Phoenix, AZ, 2003, pp. 1222–1232.

[14] N.S. Swansson, Application of vibration signal analysis techniques to signal monitoring, in: *Conference on Friction and Wear in Engineering*, Institution of Engineers, Australia, Barton, Australia, 1980, pp. 262–267.

[15] J. Ma, Energy operator and other demodulation approaches to gear defect detection, in: *Proceedings of the 49th Meeting of the Society for Machinery Failure Prevention Technology (MFPT)*, Virginia Beach, VA, 1995, pp. 127–140.

[16] H.R. Martin, Statistical moment analysis as a means of surface damage detection, in: *Proceedings of the Seventh International Modal Analysis Conference*, Society for Experimental Mechanics, Schenectady, NY, 1989, pp. 1016–1021.

[17] J.J. Zakrajsek, A review of transmission diagnostics research at NASA Lewis Research Center, Technical Report NASA TM-106746, ARL-TR-599, NASA and the US Army Research Laboratory, December 1994.

[18] H.J. Decker, R.F. Handschuh, J.J. Zakrajsek, An enhancement to the NA4 gear vibration diagnostic parameter, Technical Report NASA TM-106553, ARL-TR-389, NASA and the US Army Research Laboratory, July 1994.

[19] J.J. Zakrajsek, Transmission diagnostics research at NASA Lewis Research Center, Technical Report NASA TM-106901, ARL-TR-748, NASA and the US Army Research Laboratory, April 1994.

[20] P.D. McFadden, Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration, *Journal of Vibration, Acoustics, Stress, and Reliability in Design* 108 (1986) 165–170.

[21] C. Cempel, W.J. Staszewski, Signal demodulation techniques in vibroacoustical diagnostics of machinery, *Machine Dynamics Problems* 5 (1992) 161–173.

[22] D.M. Blunt, B.D. Forrester, Health monitoring of blackhawk and seahawk main transmissions using vibration analysis, in: *Proceedings of the Sixth Australian Aeronautical Conference/Second Pacific International Conference on Aerospace Science and Technology*, Melbourne, Australia, 1995.

[23] B.D. Forrester, Use of the Wigner–Ville distribution in helicopter transmission fault detection, in: *Proceedings of the Australian Symposium on Signal Processing and Applications-ASSPA 89*, Adelaide, Australia, 1989, pp. 78–82.

[24] M. Brennan, M.H. Chen, A.G. Reynolds, Use of vibration measurements to detect local tooth defects in gears, *Sound and Vibration* 31 (11) (1997) 12–17.

[25] B. Larder, An analysis of HUMS vibration diagnostic capabilities, in: *Proceedings of the American Helicopter Society 53rd Annual Forum*, Virginia Beach, VA, 1997, pp. 1308–1315.

[26] P.D. Samuel, D.J. Pines, D.G. Lewicki, A comparison of stationary and non-stationary transforms for detecting faults in helicopter gearboxes, *Journal of the American Helicopter Society* 45 (2) (2000) 125–136.

[27] L. Cohen, *Time–Frequency Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[28] B.D. Forrester, Analysis of gear vibration in the time–frequency domain, in: *Proceedings of the 44th Meeting of the Society for Machinery Failure Prevention Technology (MFPT)*, 1990, pp. 225–234.

[29] P.D. McFadden, W.J. Wang, Time–frequency domain analysis of vibration signatures for machinery diagnostics (I) introduction to the Wigner–Ville distribution, Technical Report OUEL 1859, Department of Engineering Science, Oxford University, 1990.

[30] P.D. McFadden, W.J. Wang, Time–frequency domain analysis of vibration signal for machinery diagnostics (II) the weighted Wigner–Ville distribution, Technical Report OUEL 1891, Department of Engineering Science, Oxford University, 1991.

[31] P.D. McFadden, W.J. Wang, Analysis of gear vibration signatures by the weighted Wigner–Ville distribution, in: *Proceedings of the Fifth International Conference on Vibrations in Rotating Machinery*, Institution of Mechanical Engineers, University of Bath, Bath, UK, 1992, pp. 387–393.

[32] F.K. Choy, S. Huang, J.J. Zakrajsek, R.F. Handschuh, D. Townsend, Vibration signature analysis of a faulted gear transmission system, Technical Report NASA TM-106623, NASA, 1994.

[33] F.K. Choy, V. Polyshchuk, J.J. Zakrajsek, R.F. Handschuh, D.P. Townsend, Analysis of the effects of surface pitting and wear on the vibrations of a gear transmission system, Technical Report NASA TM-106678, ARL-TR-520, NASA and the US Army Research Laboratory, July 1994.

[34] W.J. Williams, E.J. Zalubas, Helicopter transmission fault detection via time–frequency, scale and spectral methods, *Mechanical Systems and Signal Processing* 14 (4) (2000) 545–559.

[35] V.V. Polyshchuk, F.K. Choy, M.J. Braun, Gear fault detection with time–frequency based parameter NP4, in: *Proceedings of the Eighth International Symposium on Transport Phenomena and Dynamics of Rotating Machinery (ISROMAC-8)*, Honolulu, HI, 2000.

[36] V.V. Polyshchuk, F.K. Choy, M.J. Braun, New gear-fault-detection parameter by use of joint time–frequency distribution, *Journal of Propulsion and Power* 16 (2) (2000) 340–346.

[37] W.J. Wang, P.D. McFadden, Early detection of gear failure by vibration analysis—I. Calculation of the time–frequency distribution, *Mechanical Systems and Signal Processing* 7 (3) (1993) 193–203.

[38] W.J. Wang, P.D. McFadden, Early detection of gear failure by vibration analysis—II. Interpretation of the time–frequency distribution using image processing techniques, *Mechanical Systems and Signal Processing* 7 (3) (1993) 205–215.

[39] B.D. Forrester, Advanced Vibration Analysis Techniques for Fault Detection and Diagnosis in Geared Transmission Systems, PhD Thesis, Swinburne University of Technology, Melbourne, Australia, 1996.

[40] M.J.T. Smith, T.P. Barnwell, Exact reconstruction techniques for tree-structured subband coders, *IEEE Transactions on Acoustics, Speech and Signal Processing* 34 (3) (1986) 434–441.

[41] P.P. Vaidyanathan, Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having perfect reconstruction property, *IEEE Transactions on Acoustics, Speech and Signal Processing* 35 (2) (1987) 476–492.

[42] T.Q. Nguyen, P.P. Vaidyanathan, Two-channel perfect reconstruction FIR QMF structures which yield linear-phase analysis and synthesis filters, *IEEE Transactions on Acoustics, Speech and Signal Processing* 37 (1989) 676–690.

[43] M. Vetterli, Filter banks allowing perfect reconstruction, *Signal Processing* 10 (1986) 219–244.

[44] M. Vetterli, J. Kovačević, *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995.
[45] A. Grossmann, J. Morlet, Decomposition of hardy functions into square integrable wavelets of constant shape, *SIAM Journal on Mathematical Analysis* 15 (4) (1984) 723–736.
[46] I. Daubechies, The wavelet transform, time–frequency localization and signal analysis, *IEEE Transactions on Information Theory* 36 (5) (1990) 961–1005.
[47] Y. Meyer, *Wavelets and Operators*, Cambridge University Press, Cambridge, UK, 1992.
[48] I. Daubechies, A. Grossmann, Y. Meyer, Painless nonorthogonal expansions, *Journal of Mathematical Physics* 27 (5) (1986) 1271–1283.
[49] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11 (7) (1989) 674–693.
[50] S.G. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$, *Transactions of the American Mathematical Society* 315 (1) (1989) 69–87.
[51] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics* 41 (1988) 909–996.
[52] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61 Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.
[53] W.J. Wang, P.D. McFadden, Application of the wavelet transform to gearbox vibration analysis, in: *Proceedings of the 16th Annual Energy Sources Technology Conference and Exhibition, Structural Dynamics and Vibration Symposium*, ASME, Houston, TX, 1993, pp. 13–20.
[54] W.J. Staszewski, G.R. Tomlinson, Application of the wavelet transform to fault detection in a spur gear, *Mechanical Systems and Signal Processing* 8 (3) (1994) 289–307.
[55] W.J. Staszewski, G.R. Tomlinson, Fault detection procedures employing wavelet transform, in: *Proceedings of ISMA19—Tools for Noise and Vibration Analysis*, Leuven, Belgium, 1994, pp. 371–384.
[56] W.J. Wang, P.D. McFadden, Application of orthogonal wavelets to early gear damage detection, *Mechanical Systems and Signal Processing* 9 (5) (1995) 497–507.
[57] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, Second ed., Cambridge University Press, Cambridge, UK, 1992.
[58] D.E. Newland, *An Introduction to Random Vibrations, Spectral, and Wavelet Analysis*, Third ed., Longman, Essex, 1993.
[59] D.E. Newland, Wavelet analysis of vibration—part I: theory, *Journal of Vibration and Acoustics* 116 (1994) 409–416.
[60] D.E. Newland, Wavelet analysis of vibration—part II: wavelet maps, *Journal of Vibration and Acoustics* 116 (1994) 417–425.
[61] W.J. Wang, P.D. McFadden, Application of wavelets to gearbox vibration signals for fault detection, *Journal of Sound and Vibration* 192 (5) (1996) 927–939.
[62] P.D. Samuel, D.J. Pines, Health monitoring/damage detection of a rotorcraft planetary geartrain using piezoelectric sensors, in: *Proceedings of SPIE's Fourth Annual Symposium on Smart Structures and Materials*, San Diego, CA, 1997, pp. 44–53, paper No. 3041-05.
[63] P. Samuel, D.J. Pines, D.G. Lewicki, Fault detection in the OH-58A main transmission using the wavelet transform, in: *Proceedings of the 52nd Meeting of the Society for Machinery Failure Prevention Technology (MFPT)*, Virginia Beach, VA, 1998, pp. 323–335.
[64] D.J. Pines, P.D. Samuel, Fault classification in a helicopter gearbox using a normalized energy metric, in: *40th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, St. Louis, MO, 1999, paper Number AIAA-99-1547.
[65] P.D. Samuel, D.J. Pines, Classifying helicopter gearbox faults using a normalized energy metric, *Journal of Smart Materials and Structures* 10 (1) (2001) 145–153.
[66] S.T. Lin, P.D. McFadden, Gear vibration analysis by B-spline wavelet-based linear wavelet transform, *Mechanical Systems and Signal Processing* 11 (4) (1997) 603–609.
[67] D. Boulahbal, M.F. Golnaraghi, F. Ismail, Detection of fatigue cracks in gears with the continuous wavelet transform, Technical Report AGMA 97FTM3, American Gear Manufacturers Association, 1997.

[68] J.-C. Lo, R.E.V. Dyck, A.K. Garga, D.L. Hall, Fault prediction in transmissions using wavelet analysis, in: *Proceedings of the 52nd Meeting of the Society for Machinery Failure Prevention Technology (MFPT)*, Virginia Beach, VA, 1998, pp. 441–450.

[69] K. McClintic, K. Reichard, B. Nickerson, Feature prediction and tracking for monitoring the condition of complex mechanical systems, in: *Proceedings of the International Ship Symposium III*, Philadelphia, PA, 1999, pp. 227–237.

[70] S.M. Berri, J.M. Klosner, A new strategy for detecting gear faults using denoising with the orthogonal discrete wavelet transform (ODWT), in: *Proceedings of the 1999 ASME Design Engineering Technical Conference*, Las Vegas, NV, 1999, paper Number DETC99/VIB-8326.

[71] Y. Ohue, A. Yoshida, New evaluation method on gear dynamics using continuous and discrete wavelet transforms, in: *Proceedings of the 2000 ASME Deisgn Engineering Technical Conference*, Baltimore, MD, 2000, paper Number DETC2000/PTG-14433.

[72] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing, New York, 1994.

[73] J. Dayhoff, *Neural Network Architectures: An Introduction*, Van Nostrand Reinhold, New York, 1990.

[74] T. Harris, A. Kohonen, S.O.M. based machine health monitoring system which enables diagnosis of faults not seen in the training set, in: *Proceedings of the 1993 International Joint Conference on Neural Networks*, Nagoya, Japan, 1993, pp. 947–950.

[75] J. MacIntyre, P. Smith, T. Harris, A. Brason, Neural networks for intelligent machinery diagnostics, *Engineering Systems Design and Analysis* 100 (1994) 507–512.

[76] J.E. Lopez, I.A. Farber-Yeldman, K.A. Oliver, M.W. Protz, Hierarchical neural networks for improved fault detection using multiple sensors, in: *Proceedings of the American Helicopter Society 52nd Annual Forum*, Washington, DC, 1996, pp. 1752–1758.

[77] M.A. Essawy, S. Diwakar, S. Zein-Sabatto, A.K. Garga, Fault diagnosis of helicopter gearboxes using neuro-fuzzy techniques, in: *Proceedings of the 52nd Meeting of the Society for Machinery Failure Prevention Technology (MFPT)*, Virginia Beach, VA, 1998, pp. 293–302.

[78] K. Danai, H. Chin, Fault diagnosis with process uncertainty, *Journal of Dynamic Systems, Measurement and Control* 113 (3) (1991) 339–343.

[79] H. Chin, K. Danai, A method of fault signature extraction for improved diagnosis, *Journal of Dynamic Systems, Measurement and Control* 113 (4) (1991) 634–638.

[80] H. Chin, K. Danai, D. Lewicki, Fault detection of helicopter gear boxes using the multi-valued influence matrix method, Technical Report NASA TM-106100, AVSCOM TR-92-C-015, NASA and the US Army Aviation Systems Command, 1992.

[81] H. Chin, K. Danai, D. Lewicki, Pattern classifier for health monitoring of helicopter gearboxes, Technical Report NASA TM-106099, AVSCOM TR-92-C-033, NASA and the US Army Aviation Systems Command, 1992.

[82] H. Chin, K. Danai, D.G. Lewicki, Fault detection of helicopter gearboxes using the multi-valued influence matrix method, *Journal of Mechanical Design* 117 (2) (1995) 248–253.

[83] V.B. Jammu, K. Danai, D.G. Lewicki, Unsupervised connectionist network for fault diagnosis of helicopter gearboxes, in: *Proceedings of the American Helicopter Society 53rd Annual Forum*, Virginia Beach, VA, 1997, pp. 1297–1307.

[84] V.B. Jammu, K. Danai, D.G. Lewicki, Structure-based connectionist network for fault diagnosis of helicopter gearboxes, *Journal of Mechanical Design* 120 (1) (1998) 100–105.

[85] V.B. Jammu, K. Danai, D.G. Lewicki, Experimental evaluation of a structure-based connectionist network for fault diagnosis of helicopter gearboxes, *Journal of Mechanical Design* 120 (1) (1998) 100–105.

[86] J.S. Bendat, A.G. Piersol, *Random Data; Analysis and Measurement Procedures*, Second ed., Wiley, New York, 1986.

[87] S. Rofe, Signal processing methods for gearbox fault detection, Technical Report DSTO-TR-0476, Australian Department of Defense: Defence Science and Technology Organisation, February 1997.

[88] Q. Zhuge, Y. Lu, S. Yang, Non-stationary modeling of vibration signals for monitoring the condition of machinery, *Mechanical Systems and Signal Processing* 4 (5) (1990) 355–365.

[89] D.C. Baillie, J. Mathew, A comparison of autoregressive modeling techniques for fault diagnosis of rolling element bearings, *Mechanical Systems and Signal Processing* 10 (1) (1996) 1–17.

[90] S.L. Marple, *Digital Spectral Analysis: with Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

[91] S.M. Kay, *Modern Spectral Estimation: Theory and Application*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[92] W. Gersch, T. Brotherton, S. Braun, Nearest neighbor-time series analysis classification of faults in rotating machinery, *Journal of Vibration, Acoustics, Stress, and Reliability in Design* 105 (1983) 178–184.

[93] G. Kitagawa, W. Gersch, A smoothness prior time-varying AR coefficient modeling of nonstationary covariance time series, *IEEE Transactions on Automatic Control* 30 (1985) 48–56.

[94] W. Wang, A.K. Wong, A model-based gear diagnostic technique, Technical Report DSTO-TR-1079, Australian Department of Defense: Defence Science and Technology Organisation, December 2000.

[95] A.K. Garga, B.T. Elverson, D.C. Lang, Fault classification in helicopter vibration signals, in: *Proceedings of the American Helicopter Society 53rd Annual Forum*, Virginia Beach, VA, 1997, pp. 1316–1323.

[96] I. Daubechies, Time frequency localization operators: a geometric phase space approach, *IEEE Transactions on Information Theory* 34 (1988) 605–612.

[97] S. Mallat, Z. Zhang, Matching pursuit with time–frequency dictionaries, *IEEE Transactions on Signal Processing* 41 (12) (1993) 3397–3415.

[98] S. Jaggi, W.C. Karl, S. Mallat, A.S. Willsky, High resolution pursuit for feature extraction, Technical Report, Massachusetts Institute of Technology, November 1995.

[99] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, *SIAM Journal on Scientific Computing* 20 (1) (1998) 33–61.

[100] B. Liu, S. Ling, On the selection of informative wavelets for machinery diagnosis, *Mechanical Systems and Signal Processing* 13 (1) (1999) 145–162.

[101] P.D. Samuel, D.J. Pines, Helicopter transmission diagnostics using constrained adaptive lifting, in: *Proceedings of the American Helicopter Society 59th Annual Forum*, Phoenix, AZ, 2003, pp. 351–361.

[102] W. Sweldens, The lifting scheme: a new philosophy in biorthogonal wavelet constructions, in: A.F. Laine, M. Unser (Eds.), *Wavelet Applications in Signal and Image Processing III*, *Proceedings of SPIE*, vol. 2569, San Diego, CA, 1995, pp. 68–79.

[103] W. Sweldens, P. Schröder, Building your own wavelets at home, in: *Wavelets in Computer Graphics*, ACM SIGGRAPH Course Notes, 1996, pp. 15–87.

[104] W. Sweldens, The lifting scheme: a custom-design construction of biorthogonal wavelets, *Applied Computational Harmonic Analysis* 3 (2) (1996) 186–200.

[105] C. Herley, M. Vetterli, Wavelets and recursive filterbanks, *IEEE Transactions on Signal Processing* 41 (8) (1993) 2536–2556.

[106] I. Daubechies, W. Sweldens, Factoring wavelet transforms into lifting steps, *Journal of Fourier Analysis and Applications* 4 (3) (1998) 245–276.

[107] R.L. Claypoole, R.G. Baraniuk, R.D. Nowak, Adaptive wavelet transforms via lifting, in: *Proceedings of the IEEE Conference on Acoustics*, *Speech*, *and Signal Processing*, Seattle, WA, 1998, pp. 1513–1516.

[108] P.D. Samuel, D.J. Pines, Adaptive signal representations for helicopter transmission diagnostics, in: *Proceedings of the Third International Workshop on Structural Health Monitoring*, Stanford, CA, 2001, pp. 221–230.

[109] K.G. Church, R.R. Kolesar, M.E. Phillips, R.C. Garrido, Air vehicle diagnostic system CH-46 aft main transmission fault diagnostic—final report, Technical Report NRaD Technical Document 2966, Office of Naval Research, June 1997.

[110] P.J. Dempsey, M. Mosher, E.M. Huff, Threshold assessment of gear diagnostic tools on flight and test rig data, in: *Proceedings of the American Helicopter Society 59th Annual Forum*, Phoenix, AZ, 2003, pp. 1244–1262.

[111] J.K. Conroy, P.D. Samuel, D.J. Pines, Development of a real-time LabVIEW-based testbed for implementation of planetary gear diagnostic algorithms, in: *Proceedings of the American Helicopter Society 59th Annual Forum*, Phoenix, AZ, 2003, pp. 1347–1354.

[112] National Instruments LabVIEW software, Website www.labview.com.

[113] K.P. Maynard, C.S. Byington, G.W. Nickerson, M.V. Dyke, Validation of helicopter nominal and faulted conditions using fleet data sets, in: *Proceedings of the International Conference on Condition Monitoring*, University of Wales, Swansea, UK, 1999, pp. 129–141.

[114] N. Huang, Z. Shen, S.R. Long, M.C. Wu, E.H. Shih, Q. Zheng, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the Royal Society of London Serie A* (1998) 903–995.

[115] N.E. Huang, M.-L. Wu, W. Qu, S.R. Long, S.S.P. Shen, J.E. Zhang, Applications of Hilbert–Huang transform to non-stationary financial time series analysis, *Applied Stochastic Models in Business and Industry* 19 (2003) 245–268.

[116] S.T. Quek, P.S. Tua, Q. Wang, Detecting anomalies in beams and plate based on the Hilbert–Huang transform of real signals, *Smart Materials and Structures* 12 (2003) 447–460.

[117] H. Liang, Z. Lin, R.W. McCallum, Artifact reduction in electrogastrogram based on empirical mode decomposition method, *Medical and Biological Engineering and Computing* 38 (1) (2000) 35–41.

[118] A.A. Hood, D.J. Pines, Feasibility of a phased acoustic array for monitoring acoustic signatures from meshing gear teeth, *Journal of the Acoustical Society of America* 112 (6) (2002) 2849–2857.

[119] C. Cempel, Multidimensional condition monitoring of mechanical systems in operation, *Mechanical Systems and Signal Processing* 17 (6) (2003) 1291–1303.