Novelty detection in airport baggage conveyor gear-motors using Synchro-squeezing transform and Self-organizing maps

Budhaditya Hazra[1], Shilpa Pantula[2], and Sriram Narasimhan[3]

[1,2]*Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON N2L 3G1, CANADA*

b2hazra@uwaterloo.ca
srpantul@uwaterloo.ca
sriram.narasimhan@uwaterloo.ca

ABSTRACT

A powerful continuous wavelet transform based signal processing tool named Synchro-squeezing transform (SST) has recently emerged in the context of non-stationary signal processing. Founded upon the premise of time-frequency (TF) reassignment, its basic objective is to provide a sharper representation of signals in the TF plane. Additionally, it can also extract the individual components of a non-stationary multi-component signal, which makes it attractive for rotating machinery signals. This work utilizes the decomposing power of SST transform to extract useful components from gear-motor signals in relevant sub-bands, followed by the application of standard rotating machinery condition indicators. For timely detection of faults in airport baggage conveyor gear-motors, a novelty detection technique based on the recently developed concepts of self-organizing maps (SOM) is applied on the condition indicators. This approach promises improved anomaly detection power than that can be achieved by applying condition indicators and SOM directly to the inherently complex raw-data. Data collected from the airport baggage conveyor gear-motors provides the test bed to demonstrate the efficacy of the proposed approach.

1. INTRODUCTION

Faults in gearmotors can lead to catastrophic failures in airport baggage handling system (BHS) infrastructure leading to substantial downtime, significant monetary losses and expensive replacement scheduling. Ensuring their smooth operation requires maintenance, so that any change in the condition such as deterioration or damage can be detected in a timely manner. This can be accomplished through a combination of signal processing of gear motor vibration signatures to detect faults and novelty detection to classify a healthy from a faulty state without the historical knowledge of faults. In this study, we propose a novelty detection algorithm for condition assessment of gearmotors. The algorithm utilizes the recently developed concepts of synchro-squeezing transform (Daubechies, Lu & Wu, 2011) and integrates it with the traditional condition monitoring indicators (Vecer, Kreidl & Smid, 2005) and self-organized map (SOM) (Kohonen, 1990) based novelty detector (Lee & Cho, 2005) to monitor the healthy and faulty states of the BHS gearmotors.

In order to accomplish accurate fault diagnosis, it is important that the acquired rotating machinery vibration signals have good signal to noise ratio and less complexity. The complexity of rotating machinery signals are attributed to individual contributions from different sources like gears, bearings, rotors, and motors etc. to the overall vibration response. The problem is compounded further in presence of noise and transients from faulty machine components. This can be addressed by considering the extraction of meaningful components from the mixed signals. The other alternative is to consider the signals directly as they evolve and avoid decomposing them into different components and instead make use of a spectral techniques (Jardine, Lin & Banjevic, 2006) to diagnose faults.

Traditional signal processing methods towards gear fault diagnosis comprises of non-parametric spectral analysis methods like the Fourier transform, Cepstrum analysis (Jardine et. al, 2006) and Envelope spectrum analysis (Antoni & Randall, 2011). Recent trends in fault diagnosis have witnessed a shift towards the application of time-frequency representation (TFR) methods like short-time Fourier transform (Cohen, 1996), wavelet transforms (Wang & Mcfadden, 1995) and Wigner-Ville distribution (Cohen, 1996) to accommodate the innate non-stationarity in the gear vibration signals. Time series based techniques like auto-regressive moving average models and their variants, applied using time-invariant coefficients (Zhan & Jardine, 2005); have provided attractive options in the family of

parametric spectral approaches. Since gearbox signals comprise mainly of time-varying frequency components and amplitudes, the use of Kalman filtering based techniques for modeling time varying ARMA models is worth mentioning (Zhan & Jardine, 2005). But all of the parametric models suffer from the problem of model order selection which is again an impeding factor as far as dealing with complex and noisy rotating machinery signals is concerned.

In the alternative approach of component extraction methods, blind source separation (BSS), requiring multiple channels of data, has been applied on numerous cases of machinery fault detection (Ypma, Leshem, & Duin, 2002). The families of signal decomposition methods like empirical mode decomposition (EMD) (Lei, Lin, He, & Zuo, 2013) and synchro-squeezing transform (SST) (Daubechies et. al, 2011), applied to single channel measurements holds significant promise in this regard. EMD is a powerful and robust signal processing technique that requires only one signal measurement. However, its robustness is sometimes encumbered by its poor performance in noise and requirement of intermittency criterion. Moreover, it is essentially an empirical method lacking in rigorous mathematical construct. Synchro-squeezing transform (Daubechies et. al, 2011) is a relatively new and promising signal processing tool based on the concepts of CWT. It can decompose noisy and non-stationary signals into its components without the restrictive requirement of intermittency criterion and provides a more robust alternative to EMD for gearbox signals (Liang & Li, 2012).

Signal processing alone is not adequate to address the complete problem of condition monitoring of gear boxes. It merely generates some diagnostic patterns. These patterns need to be processed through inference tools like pattern recognition, pattern classification, novelty detection etc. (Timusk, Lipsett & Mechefske, 2008). Novelty detection (also called anomaly/outlier detection) is the process of finding an unusual behavior in machinery that has not been observed before. It is essentially a two-stage process when applied in the context of condition monitoring of machines. The first stage entails learning or training, in which the novelty detector learns by utilizing the data from a machine in normal condition. After the training stage, the detector is fed with data from the machine in a running condition to get a novelty score. If data is similar to the training data in some sense, the novelty detector shows a similair trend and the novelty score is low. Novelty score increases when there is a deviation in the operating performance of the machine or an anomaly. Higher the novelty score, higher is the level of fault in the machinery (Worden, Manson & Fieller, 2000).

A self-organizing map (SOM) is a type of artificial neural network (ANN) trained using unsupervised learning to produce a low-dimensional, discretized representation of the input space called a map (Kohonen, 1990). SOM is based on nonlinear projection of the input space to some (usually lower dimensional) output space like Principal Component Analysis. Two properties of SOM widely applicable to condition monitoring are vector projection (VP) and vector quantization (VQ). Vector projection essentially involves projecting the multidimensional data to a lower dimensional space. VQ reduces the number of samples or substitutes them with representative centroids. The accuracy of the representation of the input data in a two-dimensional map can be used as the novelty score.

This paper is based on the application of SST to identify useful signal components and then apply condition indicators (Vecer et. al, 2005) and a SOM based novelty detector to detect faulty states in a BHS gearmotor. SST (Daubechies et. al, 2011) belongs to the family of time-frequency reassignment methods that not only provide a sharp TFR but also allows extraction of the individual components (intrinsic mode functions or IMFs) of a general multicomponent non-stationary signal like EMD, yet in a much more mathematically structured manner free of restrictive requirements of intermittency. The extracted IMFs are then utilized to assess the machine health condition by subsequent application of standard condition indicators and novelty detection, in keeping with the recent trends where better diagnosis results are reported when signal processing algorithms are used in conjunction with condition indicators (Hazra & Narasimhan, 2013) and novelty detectors (Timusk, Lipsett & Mechefske, 2008).

2. SYNCHRO-SQUEEZING TRANSFORM

Since its introduction in the context of speech signals SST has evolved into an EMD-like tool (Daubechies et. al, 2011) capable of decomposing a multi-component non-stationary signal into AM-FM components that resemble intrinsic mode functions (IMFs). It is in fact a type of time-frequency reassignment algorithm that works by reallocating the continuous wavelet transform (CWT) coefficients based on the frequency information, to obtain a sharper representation in the time-frequency plane.

To understand the basic idea of SST, it is instructive to review some of the concepts of CWT. CWT of a signal $s(t)$ can be mathematically defined as an inner product:

$$W_s(a, b) = \frac{1}{a} \int_R s(t) \overline{\psi\left(\frac{t - b}{a}\right)} dt \qquad (1)$$

where, $\psi(t)$ is the mother wavelet, a and b are scale and shift parameters, respectively. In the context of synchrosqueezing, an essential requirement for the mother wavelet is that it must have a unique peak frequency (Oberlin, Meignen, & Perrier, 2012). For such a wavelet, let us denote its central frequency as ω_ψ and let ϵ_ψ be the extremal value so that ψ is supported compactly in the interval $[\omega_\psi - \epsilon_\psi, \omega_\psi + \epsilon_\psi]$. As an example (Oberlin,

Meignen, & Perrier, 2012), for bump wavelet where $\psi \propto \exp\left(-\frac{1}{1-((2\pi\omega-\mu)/\sigma)^2}\right)$, $\omega_\psi = \mu$ and $\epsilon_\psi = \sigma$.

Let us consider an example (Daubechies et. al, 2011) of a purely harmonic sinusoidal signal to illustrate the working principle of SST. The wavelet transform of the signal $W_s(a, b)$ should in principle be concentrated around the frequency of the signal as a line of constant magnitude. However, in practice, it is observed that the wavelet transform is always smeared out around the horizontal line corresponding to the sinusoid frequency in the T-F plane. This problem can be addressed by estimating instantaneous frequency $\omega(a,b)$ for all values of (a, b), which is given by the following formula (Daubechies et. al, 2011):

$$\omega(a,b) = \frac{-i}{W_s(a,b)} \frac{\partial}{\partial b} W_s(a,b) \qquad (2)$$

The primary objective for calculating the instantaneous frequency $\omega(a, b)$ is that if the signal s(t) possesses an IMF like characteristic, or is of the form $s(t) = a(t)\cos\phi(t)$ (i.e. Hilbert transform is $Hs(t) \sim a(t)\sin\phi(t)$), then $\omega(a, b)$ calculated using Eq. (2) is approximately equal to $\phi'(t)$. This suggests that, for asymptotic signals, synchrosqueezing will give a single line on the time-frequency plane, at the value corresponding to the instantaneous frequency of an IMF. Implementation-wise the wavelet coefficients in $W_s(a, b)$ are computed only at discrete scales a_k and its synchrosqueezed counterpart $T_s(\omega, b)$ is determined at the centers ω_c of the successive bins $[\omega_c - \frac{1}{2}\Delta\omega, \omega_c - \frac{1}{2}\Delta\omega]$, by the following formula (Daubechies et. al, 2011):

$$T_s(\omega_c, b) = \frac{1}{\Delta\omega} \sum_{a_k : |\omega(a_k, b) - \omega_c| \leq \frac{\Delta\omega}{2}} W_s(a_k, b) \, a_k^{-3/2} \, \Delta a_k \qquad (3)$$

where, $\Delta\omega = \omega_c - \omega_{c-1}$ and $\Delta a_k = a_k - a_{k-1}$

The next step in the SST entails extraction of the IMFs. This involves extraction of one "ridge" from T_s by finding the curve c(t) with the largest energy subject to some optimization criterion (Oberlin, Meignen, & Perrier, 2012). Once the curve c(t) is known, the associated mode h can be estimated by summing the SST coefficients near that curve: at time t, according to the following equation:

$$h(t) = \int_{\omega \approx c(t)} T_s(\omega, t) d\omega \qquad (4)$$

The main problem in this approach is stability of curve extraction. Deviations of IMFs from their asymptotic behavior or contamination by noise make the extraction unstable (Oberlin, Meignen, & Perrier, 2012). In the present work, the authors follow the approach proposed by Oberlin et. al (2012). The method is based on utilizing a ridge near the frequency peak instead of using ω as suggested by some authors (Daubechies et. al, 2011). The ridge is defined by a set of coefficients as per the following equation (Oberlin et. al, 2012):

$$Y_t = \left\{ a \; : \; \frac{\omega_\psi - \epsilon_\psi}{c(t)} \leq a \leq \frac{\omega_\psi + \epsilon_\psi}{c(t)} \right\} \qquad (5)$$

The details of the procedure are beyond the purview of this work and the readers are referred elsewhere (Oberlin et. al, 2012).

2.1. Numerical example

Let us consider a mixture of one pure sinusoid and 2 AM-FM type signals. A gearbox in its pristine state can be represented by a pure sinusoid (Hazra & Narasimhan, 2013) whose frequency matches with the meshing frequency (shaft rotation frequency times the number of gear teeth of the gear). Meshing defects in gear are manifested by the appearance of the sidebands around the meshing harmonic which can be typically represented by amplitude modulating and frequency modulating (AM-FM) signal. Thus the signal can be written as:

s (t) = sin $(2\pi 110t)$ + [1 + 5.5cos $(2\pi 2t)$] sin ($2\pi\mathbf{160}t$ + 2sin $(2\pi 10t)$) + [1 + 5.5cos $(2\pi 2t)$] sin ($2\pi\mathbf{230}t$ + 2sin $(2\pi 10t)$) + 0.7randn (1, length (t)). (6)

Fig. 1 shows the plots of the signal s (t) and its Fourier spectrum and the also the spectra of its pure sinusoidal and AM-FM components. Fig 2 shows plots of recovered IMFs. It can be observed that the synchro-squeezing transform is able to extract all the 3 components with good accuracy.
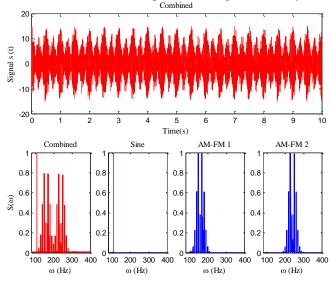

Figure 1: Signal and its components

We further consider the case of real data obtained from one of the accelerometers placed on the airport BHS gear motor (Fig. 6). The description of the BHS measurement program is mentioned in details in section. 5. Fig. 3 shows the acceleration and its synchro-squeezed time frequency representation. It is clear from Fig. 4 that SST is able to

extract the key significant energy IMFs from the gearbox signal. Furthermore, it can be noted that the extracted IMFs are consistent with the time frequency representation of the signal.
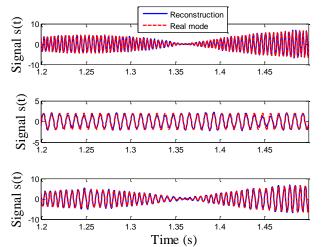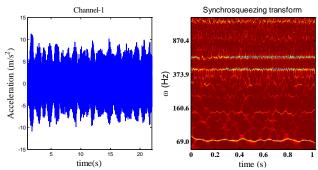

Figure 2: Recovered IMFs

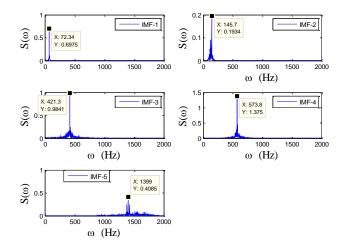**Figure 3:** Acceleration data of a BHS gearmotor & its Synchro-squeezing transform

Figure 4: Recovered IMFs from the BHS data

To compare the performance of CWT and SST, we consider the same signal with added noise of SNR=20. It can be

clearly observed from Fig. 5, that SST significantly reduces noise and is clearly able to delineate the AM-FM and the sinusoidal components. This property is particularly useful in dealing with noisy data. Thus, SST serves two important purposes; reduces noise and is able to decompose a non-stationary signal into its components.
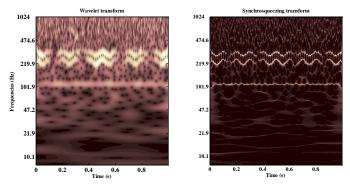

Figure 5: Comparative performance of SST and CWT in noisy data (SNR=20)

3. NOVELTY DETECTION

Self-Organizing Map (Kohonen, 1990), is used as the novelty detector in the present work. Majority of its applications are in the visualization of nonlinear relations of multidimensional data. It has also been applied in rotating machinery diagnostics (Timusk, Lipsett & Mechefske, 2008). SOM is a two-dimensional array containing neurons. A prototype vector (also called model or codebook vector), having same dimension as the input data set is associated with each neuron. This prototype vector approximates a subset of the sample vectors. During the training phase, sample vectors are assigned to the most similar prototype vector, also called best-matching unit (BMU). The algorithm trains itself in such a manner that similar input samples are mapped to the relatively close BMUs. The prototype vectors are updated iteratively during the training steps by selecting the sample randomly. The neighborhood kernel, whose radius decreases with training steps, determines the influence on the neighboring codebook vectors. Learning starts with rough learning phase having a big influence area and fast-changing codebook vectors, shifting gradually to a fine-tuning phase with small influence area and slowly adapting codebook vectors. This algorithm is referred to as sequential training or basic SOM.

SOM has also been applied to novelty detection (Lee & Cho, 2005). Given training set X, containing N normal patterns, SOM is trained to generate a set of codebook vectors $w = \{w_k | k = 1, 2, \ldots, K\}, K \ll N$. The codebook vector $m(x)$ of an input vector x and the Voronoi region S_k of each codebook vector w_k are defined as follows,

$$m(x) = w_k \Longleftrightarrow x \in S_k,$$

4

if, $\qquad \|w_k - x\|^2 < \|w_l - x\|^2, \forall l \neq k$ $\qquad$ (7)

Given a test pattern z, the Euclidean distance (quantization error) $e(z)$ between z and $m(z)$ is calculated as:

$$e(z) = \|z - m(z)\|^2 \qquad (8)$$

If this is greater than a threshold value, then it is considered to be novel. To identify the threshold value, the quantization errors corresponding to the training patterns are computed.

In the present problem, SOM is trained using condition indicators (CI) estimated using the first IMF and sum of first 3 IMFs obtained from the application of SST to the acceleration data of a gearmotor in relatively new health state. The codebook vectors and quantization errors are computed to set a threshold value. Data from gearmotor in non-normal state are fed to the algorithm to compute quantization errors. The quantization errors are calculated pointwise and the average of the quantization errors over a fixed size data window is considered in this study. If the test pattern has a mean quantization error more than the threshold, then the test pattern is identified to be novel. Mean quantization error, computed using the test set from machines in non-normal conditions is used as novelty score. Higher the novelty score, higher is the level of fault in the machine.

3.1. Condition indicators (CI)

To detect the condition of the gears and bearings, 4 condition indicators (Vecer et. al, 2005) which have been widely used in the literature, are chosen, namely: variance, kurtosis, crest factor and the energy operator. Only a brief description of the performance indicators is provided here.

- Variance: The variance of a signal is defined as:

$$\text{var(s)} = \frac{1}{N} \sum_{i=1}^{N} (s(t_i) - \mu)^2 \qquad (9.1)$$

where, μ is the mean of the signal $s(t)$ and N is the number of samples.

- Kurtosis: The kurtosis of a signal is the normalized fourth moment and is defined as:

$$\text{kurt(s)} = \frac{1}{N} \sum_{i=1}^{N} \frac{(s(t_i) - \mu)^4}{\sigma^4} \qquad (9.2)$$

where, σ is the standard deviation of the signal

- Energy operator: The energy operator for a signal s(t) is defined as [15]:

$$\text{EOP(s)} = \Delta x(i) = s(i+1)^2 - s(i)^2 \qquad (9.3)$$

where, and $\overline{\Delta x}$ is the mean value of Δx vector.

- EOP variance and EOP kurtosis are calculated in the same manner using the formulas (7.1) and (7.2)

- Crest factor: The crest factor (CF) for a signal s(t) is defined as:

$$CF = \frac{max(x(i)) - min(x(i))}{\sigma} \qquad (9.4)$$

4. PROPOSED ALGORITHM

The main steps of the proposed algorithm are as follows:
- Calculate the CWT W_s(a,b) of $s(t)$
- Calculate the instantaneous frequency $\omega(a, b)$
- Calculate the SST $T_s(\omega, b)$ over the TF plane
- Extract dominant curves from c(t) from $T_s(\omega, b)$
- Reconstruct the signal as a sum of components, one for each extracted dominant curve
- Apply the CI on the most dominant IMFs
- Apply SOM to the CI, treating the data from the new healthy state as the training set
- Calculate the mean quantization errors between the codebook vector and the subsequent windows of data

5. RESULTS FROM FIELD EXPERIMENTATION PROGRAM

Recently the authors have engaged in condition based maintenance program aimed at detecting faults in the Toronto Pearson airport baggage handling system (BHS) gearboxes. The main idea is to gather acceleration data from the BHS system gearboxes and develop sophisticated algorithms towards automated diagnostics and prognostics of the gearboxes. As a part of the field instrumentation and data acquisition programme, vibration data was collected for a few minutes on a particular day of every week starting from May-2013. The schematic of a typical gearmotor and the data-acquisition set-up is shown in Fig. 5. The sampling frequency was kept at 4000 Hz. The fundamental meshing frequency is approximately close to 80Hz depending upon the conveyer belt rpm which varies between 170 to 200 rpm. Data were collected from a selected gearmotor at its new and old health stages. Fourier spectra (Fig. 6) of the gearmotor data shows the presence of sidebands with significant energy in the data at the old health state compared to the new one.
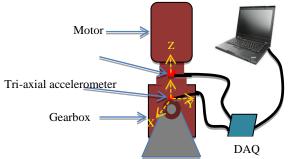

Figure 6: Schematic of the gearmotor and the instrumentation set-up
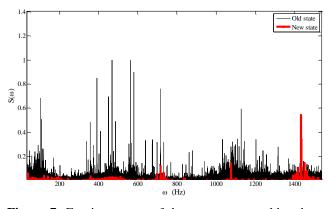
Figure 7: Fourier spectra of the gearmotor at old and new health states

The raw data from new healthy state gearmotor is concatenated to the data from the old state. The condition indicators (CI) are applied on the concatenated raw data, IMF-1 and the sum of first 3-IMFs obtained using SST. SST is applied on sub-windows of 4096 samples of data and concatenated to form the IMF vectors. The main idea of extracting first 3 IMFs and summing them up is that the first 3-IMFs contribute to most of the signal's energy content. The higher order IMFs contains contributions mostly from the noise. The condition indicators are estimated recursively considering every sample of the data. Fig. 8 shows the condition indicator for the raw data, IMF-1 and the sum of first 3 IMFs. It can be observed that there is a considerable jump in the CI values after approximately first 70000 samples of data. These samples represent the data obtained from the gearbox at the new state of health. The remaining part of CIs represents data at old health state.

The trend in the values of recursive variance, kurtosis and energy kurtosis for raw data is not as clear as it is for the corresponding CI values for IMF-1 and sum of the first 3 IMFs, indicating thereby that the combination of SST and CIs provides a better indication of comparative health states than the combination of raw data and CIs. The uncertainties associated in the estimates of the condition indicators by using raw data and the sum of IMFs 1, 2 & 3 is shown in Fig. 9. From Fig. 9, it can be observed that the variance-ratio (ratio of variances between the CIs corresponding to the old stage data and that corresponding to the new stage of data) of the condition indicators using raw data is much more compared to the same when estimated using the sum of IMFs. This clearly points towards more uncertainties associated with the use of condition indicators on the raw
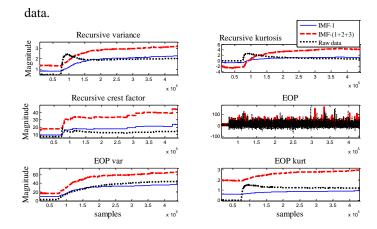
data.

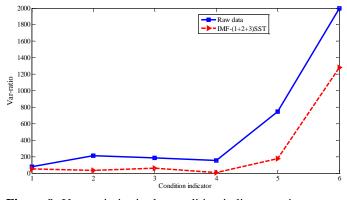**Figure 8:** Recursive condition indicators applied on the raw data and IMFs

Figure 9: Uncertainties in the condition indicator estimates

Novelty detection is applied as mentioned in the previous sections. First 70,000 samples of the raw data and also the IMF-1 is treated as the training set and the mean of the quantization error as given by Eq. 8 is estimated for the successive windows of test data. Fig. 10 shows the plot of the average quantization error. It can be clearly observed that the novelty score increases for the successive windows of data. This implies that the data from the successive windows represent a more novel or an anomalous state of data compared to the first window (training window). A health state is typically indicated by approximately constant values of Q_{avg} over successive windows. Novelty detection applied on the sum of the IMFs (1, 2 & 3) obtained using SST shows best and most consistent performance. It is closely followed by the performance of the sum of 3-IMFs extracted using SST with 10% added noise and the case with IMF-1 only. The performance of novelty detection applied on raw data degrades significantly. This is clear from the trend of the Q_{avg} values for the raw data in Fig. 10, which fails to establish health states or values that are nearly constant over successive windows or 3 consecutive windows. The performance of novelty detection is worst for raw data with 10% noise. Thus, it can be concluded that the

6

SOM novelty indicator is clearly able to distinguish between the relatively older and newer health states. Superior performance of SST even with added levels of noise is the key result.
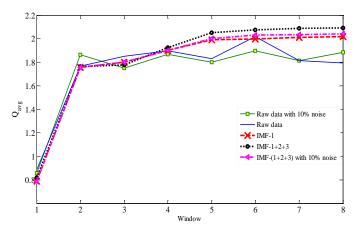

Figure 10: Average quantization error for the SOM based novelty detector

6. CONCLUSIONS

A new novelty detection algorithm towards fault diagnosis of airport baggage handling system gearmotors using a combination of synchro-squeezing transform, traditional rotating machine condition features and self-organized maps is presented. The reassignment property of synchro-squeezing transform allows for a better resolution of the signal features in the presence of noise. Subsequent application of curve extraction techniques along the ridges allowed EMD like decomposition of the signal into IMFs. Application of condition indicators recursively to the IMFs clearly shows the trend indicative of health degradation in the experimental data from the gearmotor. Application of SOM based novelty detector further delineated the new and the degraded health states of the airport baggage handling system gearmotor.

REFERENCES

Antoni, J. and Randall, R, B. (2011). Rolling element bearing diagnostics—A tutorial. *Mechanical Systems and Signal Processing*, 25(2), pp. 485-520.

Cohen, L. (1995). *Time-Frequency Analysis.* Prentice-Hall, Englewood Cliffs. NJ, USA.

Daubechies, I., Lu, J., and Wu, H. (2011). Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Applied and Computational Harmonic Analysis*, 30(2), pp. 243–261.

Hazra, B., and Narasimhan, S. (2013), Rotating machinery diagnosis using synchro-squeezing transform based feature analysis. *Proceedings of MFPT Conference.* May 13-17, Cleveland, Ohio.

Jardine, A. K. S., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition based maintenance. *Mechanical Systems and Signal Processing,* 20(7), pp. 1483-1510.

Kohonen. T. (1990), The Self-Organizing Map. *Proceedings of the IEEE,* 8(9), pp. 1464–1480.

Lee, H. J., & Cho, S. (2005). SOM-based novelty detection using novel data. *Intelligent Data Engineering and Automated Learning-IDEAL.* Springer, Berlin Heidelberg, pp. 359-366.

Lei, Y., Lin, J., He, Z., and Zuo, M, J., (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing,* 35(1-2), pp. 108-126.

Liang, M., and Li, C. (2012). A generalized synchro-squeezing transform for enhancing signal time–frequency representation. *Signal Processing,* 92(9), pp. 2264-2274.

Oberlin, T., Meignen, S., and Perrier, V. (2012), On the Mode Synthesis in the Synchrosqueezing method. *Proceedings of EUSIPCO*, Bucharest, Aug. 27-31.

Staszewski, W., Worden, K., and Tomlinson, G. (1997). Time frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition. *Mechanical Systems and Signal Processing*, 11(5), pp. 673–692.

Timusk, M., Lipsett, M., and Mechefske, C, K., (2008). Fault detection using transient machine signals. *Mechanical Systems and Signal Processing*. 22(7), pp. 1724-1749.

Vecer, P., kreidl, M., and Smid, R. (2005). Condition indicators for gearbox condition monitoring systems. *Acta Polytechnica*, 45(6), pp. 35-43.

Wang, W. and McFadden, P. (1995). Application of orthogonal wavelets to early gear damage detection. *Mechanical Systems and Signal Processing*, 9(5), 497–507.

Worden, K., Manson, G., & Fieller, N. R. J. (2000). Damage detection using outlier analysis. *Journal of Sound and Vibration*, 229(3), 647-667.

Ypma, A., Leshem, A., and Duin, R. (2002). Blind separation of rotating machine sources: bilinear forms and convolutive mixtures. *Neurocomputing*, 49(1-4), 349-68.

Zhan, Y. and Jardine, A. (2005). Adaptive autoregressive modeling of non-stationary vibration signals under distinct gear states Part 1: modeling. *Journal of Sound and Vibration*, 286(6), 429-450.

Budhaditya Hazra received his undergraduate degree from the National Institute of Technology, Durgapur, India, Master of Technology from the Indian Institute of Technology, Kanpur, India, and his PhD from the University of Waterloo, Canada, in 2010. He is currently a post-doctoral research associate at the University of Waterloo. His research interests are in the areas of system identification and machine diagnostics.

Shilpa Pantula received her undergraduate degree from the Indian Institute of Technology, Chennai, India in 2012. She is currently pursuing her Master of Applied Science at the University of Waterloo, Canada. Her research interests are in structural engineering and machine diagnostics.

Sriram Narasimhan received his PhD from Rice University, Houston, Texas, USA, in 2005. He is an Associate Professor in the departments of Civil and Environmental Engineering & Mechanical and Mechatronics Engineering at the University of Waterloo, Canada. His research interests are in the areas of system identification, control, and machine diagnostics. He is a member of ASCE, and serves on the editorial boards of ASCE Journal of Structural Engineering, Bridge Engineering and Structural Control and Health Monitoring.