GEAR FAULT DIAGNOSIS USING TIME-FREQUENCY METHODS

DIAGNOSTIC DE DEFAUTS D'ENGRENAGE PAR ANALYSE TEMPS-FREQUENCE

Mir Saeed Safizadeh[1] *Aouni A. Lakis*[2] *Marc Thomas*[3]

[1]Institute for Aerospace Research, National Research Council, Canada, Ottawa, ON K1A 0R6

[2]Department of Mechanical Engineering, École polytechnique de Montréal, Montreal, Quebec, H3C 3A7

[3]Department of Mechanical Engineering, École de Technologie Supérieure, Montreal, Quebec, H3C 1K3

ABSTRACT

Time-frequency analysis has been found to be effective in monitoring the transient or time-varying characteristics of machinery vibration signals, and therefore its use in machine condition monitoring is increasing. This paper proposes the application of time-frequency methods, which can provide more information about a signal in time and in frequency and gives a better representation of the signal than the conventional methods in machinery diagnosis. In this paper, we review the machine diagnosis techniques based on the verification of classical vibration parameters. Then the necessity of using time-frequency analysis in machinery diagnostics is discussed. Finally, the theory of the Short-Time Fourier Transform, the Wigner-Ville distribution and the Wavelet transforms are briefly studied and their advantages are shown by some practical examples.

RESUME

L'analyse temps-fréquence s'est montrée efficace dans la surveillance de machines produisant des signaux transitoires ou variables / fluctuant en fonction du temps, et son utilisation se popularise dans le contexte de la surveillance des machines. Cet article propose l'utilisation des méthodes temps-fréquence pour leur capacité à mieux illustrer le contenu fréquentiel des variations du signal temporel que les méthodes conventionnelles normalement utilisées en diagnostic de machines. Les techniques de diagnostic fondées sur les paramètres classiques de vibration son discutées. La nécessité d'utilisation de l'analyse temps-fréquence est ensuite présentée. Finalement, la théorie de la transformée SFT (Short-Time Fourier Transform), de la distribution Wigner-Ville et de la transformée d'Ondelettes sont étudiées et leurs avantages respectifs sont démontrés par des exemples pratiques.

1. INTRODUCTION

The objective of predictive maintenance is to detect and identify defaults in the earlier steps in order to have the necessary time to schedule repairs with minimum disruption to operations and production [1]. The key factor of the predictive maintenance is diagnostic. The diagnostic is based on a systematic inspection in vibration signal to find all susceptible defects, which may affect the machine.

There are several conventional methods, which have been applied for a long time to fault detection and identification. Some of these methods provide a representation of signals in time domain and others provide a representation in frequency domain [2]. For example, overall level measurement is the most common vibration measurement in use in time domain. It is a simple and inexpensive type of measurement to undertake and there are charts available which indicate the levels deemed acceptable, for example VDI 2056. Great many indicators have been also developed for machine condition monitoring and fault detection, such as *crest factor* and *Kurtosis [2]*. The crest factor is the

ratio of the peak on the rms signal [3].
The Kurtosis is defined as the 4th order moment of the time signal distribution:

$$Y_{kurt} = \frac{\frac{1}{N}\sum_{k=1}^{N}\left(y_k - Y_m\right)^4}{\left(\frac{1}{N}\sum_{k=1}^{N}\left(y_k - Y_m\right)^2\right)^2} \qquad (1)$$

where y_k is the sampled signal for k = 1 to N, et Y_m is the mean signal.

If the decision criteria based on the time analysis allow for diagnosing a default, they don't allow for identifying its cause. In addition, we need to take into consideration not only the increase in the power of the signal, but also the development of its form and a spectral analysis is needed.

An alternative techniques have been also applied to verifying the variation in the form of a signal such as *Cepstrum* and the *envelop method* (Hilbert transform) of the narrow band of the signal. Cepstrum allows for detecting repetitive impacts in the time domain by identifying the impact period. It is the inverse Fourier transform of the logarithmic spectrum of the signal:

$$C(x(t)) = TF^{-1}(\log(X(\omega))) \qquad (2)$$

The analysis of the analytical signal, a complex signal in which the imaginary part is equal to the Hilbert transform of the real part, allows for identifying the amplitude and phase modulation. The Hilbert transform can be described as the following:

$$H(x(t)) = \frac{1}{\pi}\int_{-\infty}^{+\infty} x(\tau)\frac{1}{t-\tau}d\tau \qquad (3)$$

In all of these methods, it is assumed that signal is stationary but this assumption is not always true. In some cases, when defects begin, vibration signal becomes non-stationary and in this case, the conventional methods (FFT) are not applicable. On the other hand, there are presently several types of variable speed rotating machinery for which the stationary or pseudo-stationary vibration signals cannot be assumed. In recent years, a number of new analysis methods have been developed in the field of signal processing called joint time-frequency analysis methods. The time-frequency analysis not only enables us to represent the signal in three dimensions (time-frequency-amplitude) but also permit us to detect and follow the development of the defects, which generate weak vibration power. Although the time-frequency methods are regarded as advanced diagnostic techniques, which offer high sensitivity to faults and a good diagnostic capability, there is a little tendency to use these methods in the field of machinery diagnostics. The research into the development of the theory of joint time-frequency methods and other non stationary signal processing methods are fast progressing but it seems that some works are needed to motivate the industrial people and show them the capability of these new methods in condition monitoring of mechanical systems.

The objective of this work is on the one hand, to demonstrate the accuracy which can be obtained by using the joint time-frequency analysis methods in field of machinery diagnostics and on the other hand, to introduce an in-house user-friendly time-frequency software which has been developed to facilitate the use of time-frequency methods by engineers whether or not they are familiar with time-frequency analysis.

2. TIME-FREQUENCY ANALYSIS

The primary objective of all research into signal processing is to find an efficient method, which would generate results rapidly and clearly, and in a manner which could be relatively easily interpreted. Using the time-frequency representation of the signal energy is one of the attempts to show a signal in three dimensions and obtain clear interpretation.

2.1 Short-Time Fourier Transform

The short-time Fourier transform (STFT) was the first time-frequency method, which was applied by Gabor [4] in 1946 to speech communication. The STFT may be considered as a method that breaks down the non-stationary signal into many small segments, which can be assumed to be locally stationary, and applies the conventional FFT to these segments.

The STFT of a signal $s_t(\tau)$ is achieved by multiplying the signal by a window function, $h(\tau)$, centered at τ, to produce a modified signal. Since the modified signal emphasises the signal around time τ, Fourier Transforms will reflect the distribution of frequency around that time.

$$S_t(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-j\omega\tau} s(\tau)h(\tau - t)d\tau \tag{4}$$

The energy density spectrum at time τ may be written as follows:

$$P(t,\omega) = \left|S_t(\omega)\right|^2 = \left|\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-j\omega\tau} s(\tau)h(\tau - t)d\tau\right|^2 \tag{5}$$

For each different time, we get a different spectrum and the ensemble of these spectra provides the time-frequency distribution $P(t,\omega)$, which is called Spectrogram. The major disadvantage of the STFT is the resolution tradeoff between time and frequency. Resolutions in time and frequency will be determined by the width of window $h(\tau)$. A large window width provides good resolution in the frequency domain, but poor resolution in the time domain. Conversely, a small window width provides good resolution in the time domain and poor resolution in the frequency domain, following the Heisenberg principle. This limitation of the STFT is arising from using a single window for all frequencies and therefore, the resolution of analysis is the same at all locations in the time-frequency plane (figure 2-a).

2.2 Wavelet Transforms

The wavelet transform is another linear time-frequency representation, similar to the spectrogram but with more flexibility in time and frequency resolution. In the STFT, the length of window function will remain constant during the analysis of the signal. In the wavelet transform, by translation and dilation / contraction of a window function called the mother wavelet function, we build up a family of window functions of variable lengths:

$$\psi_{s\tau}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \tag{6}$$

where $\psi(t)$, s and τ are respectively a mother wavelet function, the scale of wavelet transform, and time shift. The wavelet transform is defined as

$$W_{\psi}x(s,\tau) = \int_{-\infty}^{\infty} x(t)\overline{\psi}_{s\tau}(t)dt \qquad (7)$$

where $W_{\psi}x(s,\tau)$ are called wavelet coefficients.

The variable window length property of the wavelet transform gives us the possibility of having the time and frequency resolutions dependent on the frequency under consideration. Figure 2 illustrates this point by showing the cells of resolution in the time-frequency plane for the STFT and the wavelet transform.

One important advantage of the wavelet transform is its ability to carry out local analysis. This property is of significant value in revealing any small change in the signal and distinguishes the wavelet transform from other signal analysis techniques. If we consider the result obtained by applying the wavelet transform on a Dirac pulse at time $t_0 = 0.1 sec$ (Figure 3), we see a triangular shape, which points at $t = t_0$ in the time-frequency plane. An impulse excites all the frequencies. Figure 3 shows that the signal is more localized in high frequencies than in low frequencies.

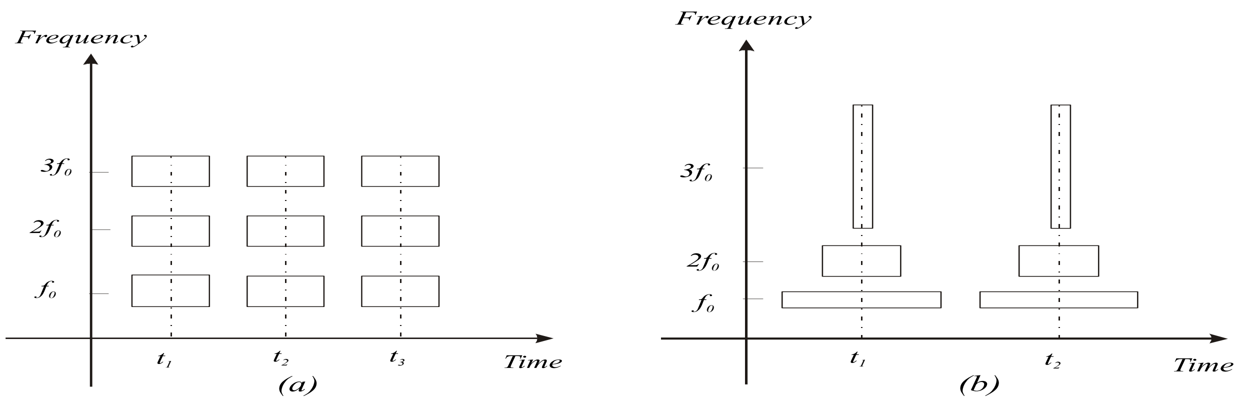

Figure 2: Time-frequency plane of (a) the STFT (b) the wavelet transform

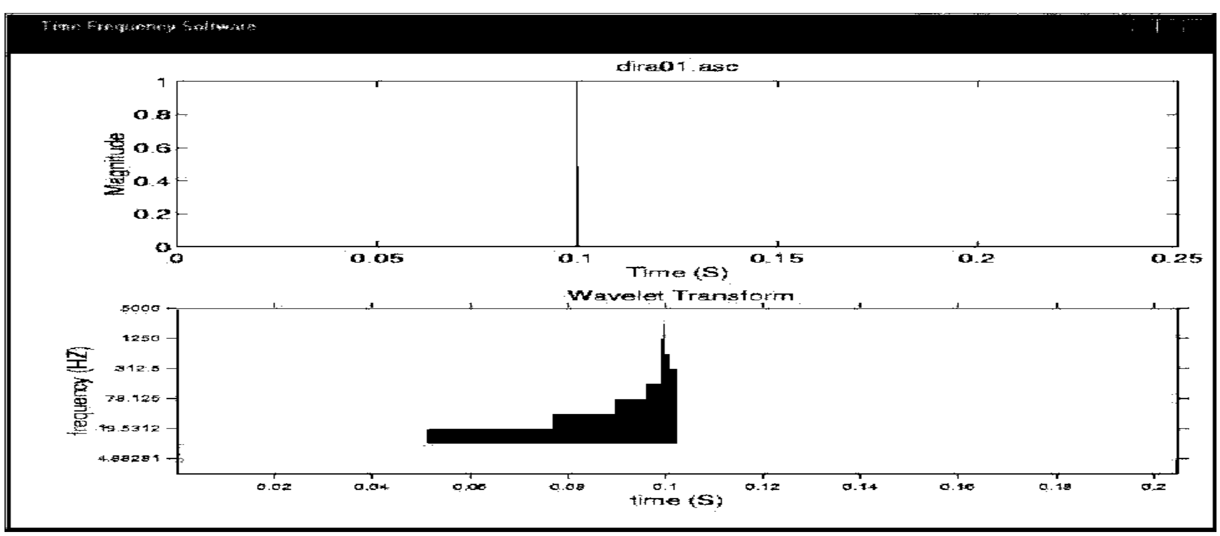

Figure 3: wavelet transform of a Dirac function

The variable time and frequency resolution of the wavelet transform is one of its advantages; however, in the discrete wavelet transform, the frequency axis has logarithmic scale (octave). The octave scale of the frequency axis does not permit either fine frequency resolution of the high frequencies. This characteristic of the frequency axis in the wavelet transform makes it a specialized method to be used for signals, which contain long-duration events at the low frequencies and short-duration events at the high frequencies. The octave scale of the frequency axis in the wavelet transform may at times be considered to be a disadvantage of this method. To resolve the inconvenience of the wavelet transform, another method based on the principle of the wavelet transform has been introduced. This method is called the wavelet packet transform, and gives a frequency axis with linear scale at the expense of losing the excellent time resolution of the high frequencies of the wavelet transform. Figure 4 shows the wavelet packet transform of a Dirac pulse at time $t_0 = 0.1 sec.$

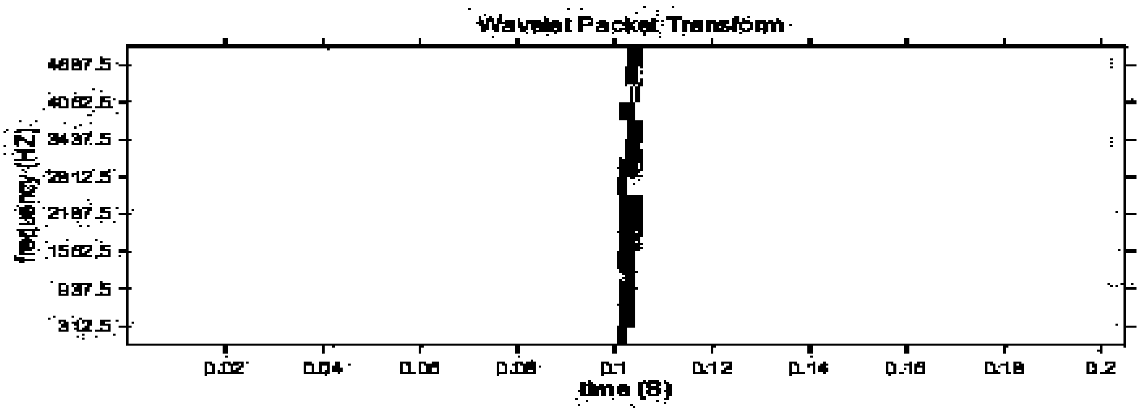

Figure 4: Wavelet packet transform of a Dirac function

2.3 Wigner-Ville Distribution and Cohen's Class Time-Frequency Distributions

One interesting time-frequency energy distribution is the Wigner-Ville distribution (WVD) [5], which has recently been applied to the field of mechanical signal analysis. This distribution is a bilinear function, in contrast to the transforms discussed above, which are linear transforms. In a linear transform, the similarity of the signal to a window function is measured using the correlation function; on the other hand, the Wigner-Ville distribution is the Fourier transform of the instantaneous auto-correlation of the signal. Thus, its time-frequency representation is independent of the window function.

If the instantaneous correlation, $R_x(\tau, t_0)$, at time t_0 with a time lag τ, is defined as

$$R_x(\tau, t_0) = \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+\tau} x(t - \tau/2)\, x(t + \tau/2)\, dt \tag{8}$$

Its Fourier transform may be written as follows:

$$S_x(\omega, t_0) = WVD_x(\omega, t_0) = \int_{-\infty}^{\infty} R_x(\tau, t_0)\, e^{-i\omega\tau} d\tau \tag{9}$$

The WVD satisfies a large number of desirable mathematical criteria and has excellent resolution in the time and frequency domains, but it has two major problems. First, it is not always non-negative, which, since energy is always positive, makes it difficult to interpret the Wigner-Ville representation as the energy distribution of the signal in the time-frequency plane. Secondly, because it is bilinear, it produces interference terms or cross terms for multicomponent signals [6]. The interference term is located between two components of a multicomponent signal in the time-frequency representation, and it oscillates with a frequency proportional to the distance between these two components, as shown in Figure 5 for two parallel chirps. In numerical method, we cannot use a signal from -∞ to +∞, and therefore we use a window function to cut the signal in the time domain. This time-window version of the WVD is called the pseudo-WVD [5]. Windowing in the time domain provides some smoothing in the frequency direction of the WVD and reduces the interference terms oscillating perpendicularly to the frequency axis, but at the expense of loosing many properties of the WVD.

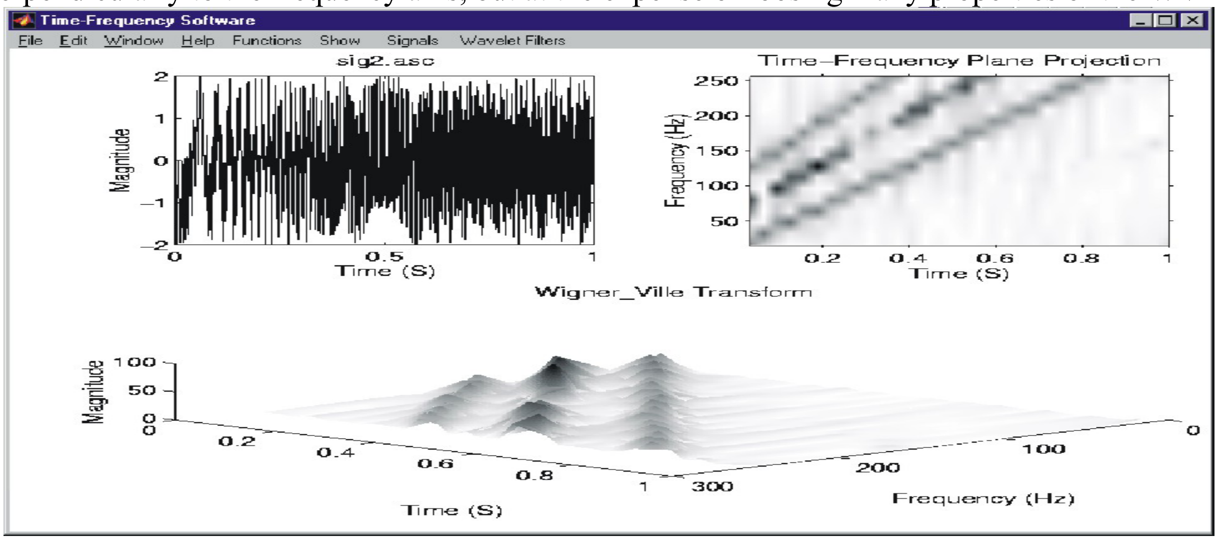

Figure 5: Wigner-Ville distribution of two parallel chirps

In addition to the interference terms, the alias problem may affect the discretization of WVD if the signal is real-valued and sampled at the Nyquist rate. To prevent this problem, Ville [7] suggested using the analytical signal, a complex signal in which the imaginary part is equal to the Hilbert transform of the real part. With the analytical signal, the spectral domain will be [0, ½] of the real signal and consequently the aliasing will not happen. On the other hand, since the spectral domain is divided by two, the number of components in the time-frequency plane is also reduced by half. In addition, application of the analytical signal eliminates the negative part of the frequency axis, so that the interference terms generated between negative and positive frequency components are eliminated, leading to a considerable decrease in the number of interference terms. Since the development of the WVD, there have been several attempts to find other formulas to express the energy of the signal in the time-frequency plane. Cohen classified these formulas by giving a general formula for all time-frequency energy distributions. This formula is defined as:

$$WD_x(t,\omega) = \iiint x(u+\tau/2)\, x(u-\tau/2)\, \varphi(\theta,\tau)\, e^{i(\theta u-\theta t-\tau\omega)} du\, d\tau\, d\theta \qquad (10)$$

where θ and τ are respectively a frequency lag and a time lag.

In addition, $\varphi(\theta,\tau)$ is a kernel function that, when changed, gives different time-frequency distributions with different properties. One desirable choice for the kernel function is a separable smoothing function in both the time and frequency domains which attenuates the interference terms of the WVD in both the frequency and time directions. The distribution attained in this way is called the smoothed-WVD, and is defined as:

$$SWVD_x(t,\omega) = \iint WVD_x(u,\xi)\, \Phi(t-u, \omega-\xi)\, du\, d\xi \qquad (11)$$

where $\Phi(t,\omega)$ is a two dimensional smoothing function.

The smoothed-WVD may be considered as an intermediate distribution between the STFT and the WVD. It has some of their advantages and none of their problems. The WVD provides the best resolution in time and in frequency, but produces some significant interference terms in the time and in frequency directions. The STFT is a linear transform and does not suffer from interference terms, but it is unable to give satisfactory resolution simultaneously in time and in frequency. The smoothed-WVD provides the best compromise between these two problems: interference terms and resolution in time and frequency. Figure 6 shows the STFT and the smoothed Wigner-Ville distribution of two parallel chirps. This figure shows that the smoothed-WVD provides better resolution and clearer representation of the signal than the STFT.

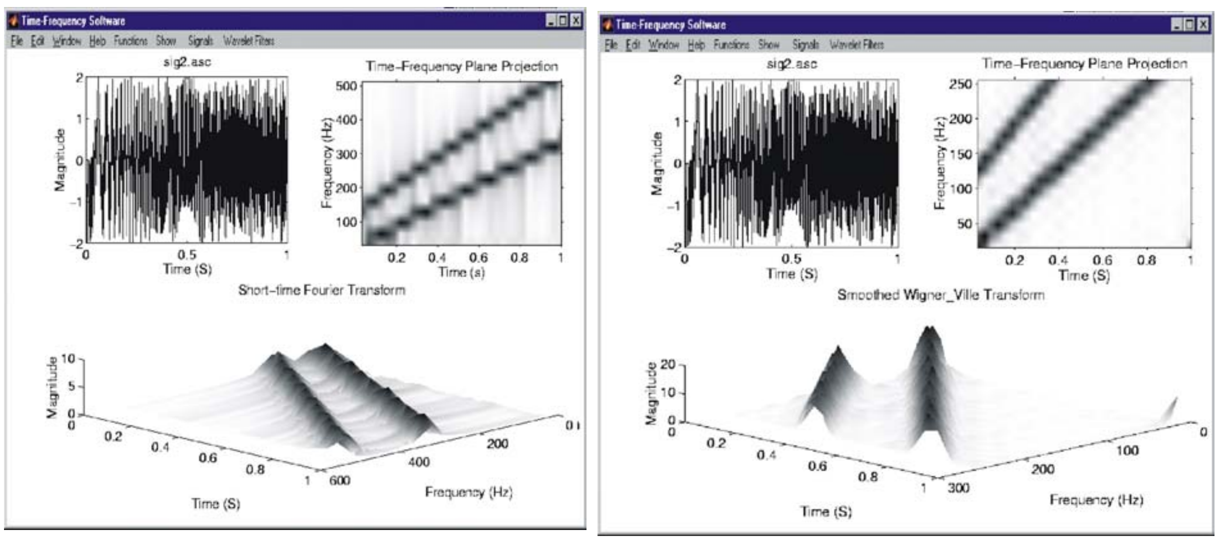

Figure 6: STFT and smoothed Wigner-Ville distribution of two parallel chirps

3. SOFTWARE FOR TIME-FREQUENCY ANALYSIS

Today, one of the most important factors limiting the progress of machine diagnostic techniques is the lack of familiarity of mechanical engineers with new signal processing methods. The complicated theory of time-frequency analysis and the absence of operational software for time-frequency analysis restrict engineers from using these methods in machine diagnosis. An in-house user-friendly software has been developed in collaboration with IMS [3] to facilitate the use of time-frequency methods by engineers whether or not they are familiar with time-frequency analysis [8]. This software permits the use of different methods of time-frequency analysis such as the Short-Time Fourier Transform, the Wigner-Ville Distributions, and the Wavelet Transforms. In addition, it provides different kinds of wavelet transforms, for example: the wavelet transform, the wavelet packet transform and the adaptive wavelet transform. In addition, a new technique of "zoom in wavelet transform" makes possible to obtain very satisfactory frequency resolution. This program has been developed especially for the diagnosis of defects in machinery and includes most of the commonly used methods of time-frequency analysis. The program has some interesting options, which are of considerable practical value in such cases. For example, denoising by wavelet transform, which is an important tool in the analysis of noisy signals, allows the user to obtain an improved time-frequency representation.

4. GEAR FAULT DETECTION USING TIME-FREQUENCY ALGORITHMS

In this section, the efficiency of the time-frequency methods in an industrial case is demonstrated. This case comes from the defective gear-train of a hoist drum in a large shovel operating at an open-pit iron mine. The data are measured by International Measurement Solutions company [3] in order to diagnose the problem in the machine. A minimum length of time is required to perform FFT analysis of each process. The time resolution required will depend on the period of each tooth mesh and the desired level of accuracy. Sometimes, it is not possible to measure the signal for long enough to provide the periodicity of shock in the FFT spectrum. In this particular case, the process did not even last one revolution of the driven gear. The case was investigated by time-frequency distribution precisely because it is known that time-frequency methods do not need as much time signal as the FFT spectrum. Figures 7 shows, respectively, the time signal of the damaged gearbox measured by IMS company and its spectrum, the STFT, the Wigner-Ville distribution and the smoothed Wigner-Ville distribution. The spectrum of the signal displays a large peak around 200 Hz with a lot of sidebands and some smaller peaks in the vicinity of 400 Hz, 800 Hz and 1200 Hz. The frequencies of 200 Hz and 400 Hz have been identified as gear mesh frequencies of the gear box.

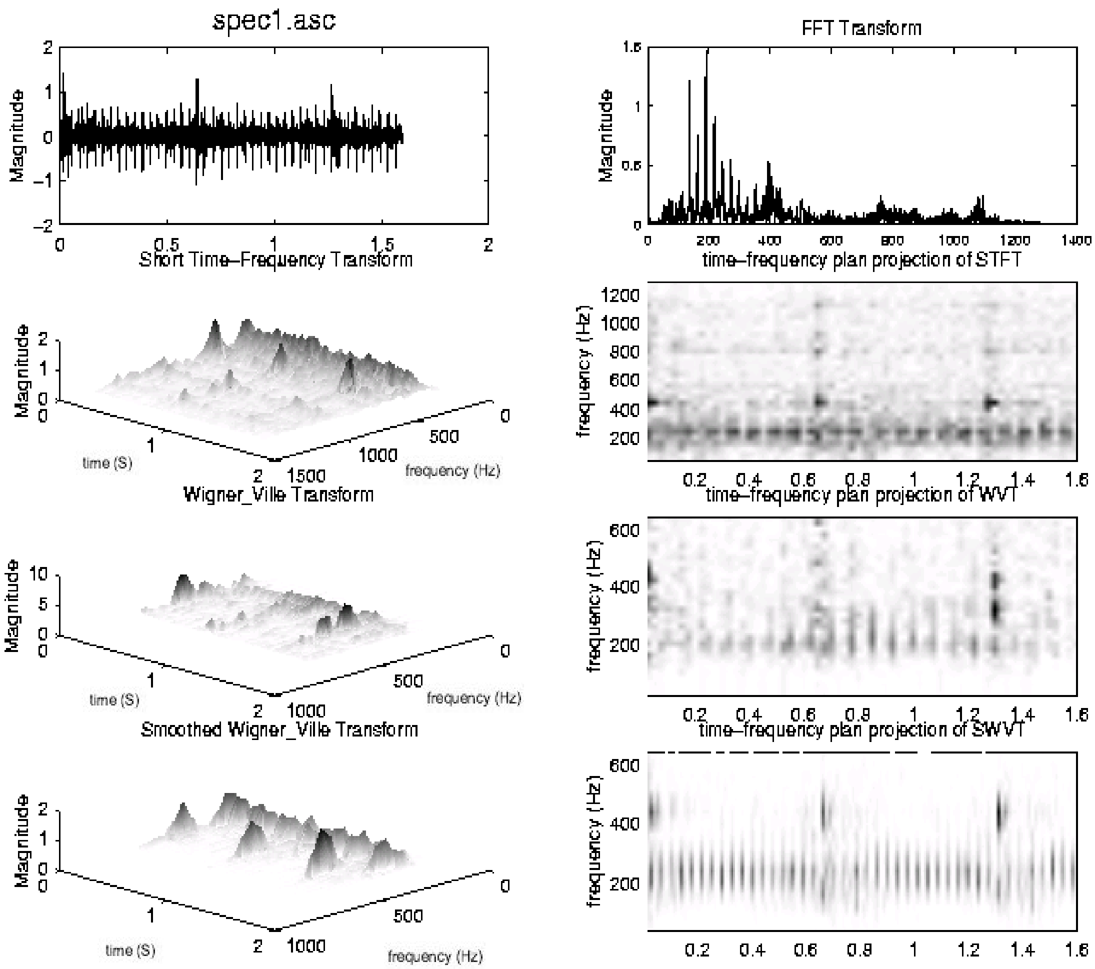

Figure 7 Time, spectrum, STFT, Wigner-Ville distribution and smoothed Wigner-Ville distribution, of defective gearbox signal.

Gears generate a mesh frequency equal to the number of teeth on the gear multiplied by the rotational speed of the shaft driving it. A high vibration level at the mesh frequency may be caused by tooth error, wear of the meshing surfaces or any other problem that would cause the profiles of meshing teeth to deviate from their ideal geometry. Usually, the amplitude at the gear mesh frequency is not used to detect a gear a gear damage because others operating parameters such as loads can affect this amplitude. Instead of that, analyzing the sidebands and the harmonics of the gear mesh frequency can identify the default. Sidebands at the mesh frequency are the result of a modulation that typically is due to a failure of mating teeth or to a bent or misaligned shaft [2]. For example, a cracked tooth will, due to its weakened mechanical condition, deflect more under load than the other (healthy) teeth when it goes into mesh. This results in a signal with amplitude modulation. Thus, an increasing level in the sidebands spaced with rotation speed in the frequency spectrum results from the cracked tooth. On the other hand, a bad backlash will generate a vibration at the second harmonic of the gear mesh frequency and wear of the meshing surfaces will produce a lot of harmonics of the gear mesh frequency accordingly with the severity of the wear [2]. Thus, we can suspect from the spectrum analysis, a broken teeth or a bent shaft (due to sidebands) and probably wear of the meshing surface (due to gear mesh harmonics). The Short-Time Fourier transform clearly displays time-frequency representation of the signal. There are a gear meshing frequency at approximately 200 Hz and some pulses at approximately 400 Hz that appears at each 0.66 sec (1.5 Hz). It is known that pulses can appear in the vibration signal of a gearbox if there is a broken tooth. Then we may suspect that there is a broken tooth in this gearbox at the gear mesh frequency of 400 Hz and it is also possible to find which gear has the default from its rotating speed (1.5 Hz). The Wigner-Ville distribution did not provide a good representation of the signal due to the cross terms, which are generated between the signal components. The smoothed Wigner-Ville shows similar characteristics of the signal even more clearly than the STFT and we can calculate the frequency of repetition of pulses with more precision by the smoothed Wigner-Ville than by the STFT. The wavelet transform of the signal (figure 8) shows the three repetitive pulses in the frequency band 320-640 Hz. The frequency resolution is too poor for clearly identifying the gear mesh frequency. The frequency of the periodicity of the signal may be calculated from the wavelet transform more precisely than from the STFT, because the time resolution in this band of the wavelet transform is finer than in the STFT. But in the three-dimensional representation of the signal, the STFT provides better representation than does the mean square wavelet.

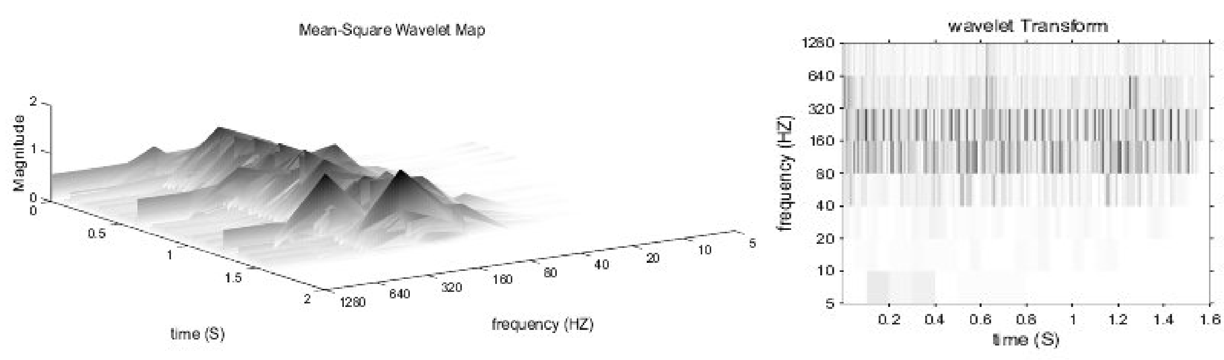

Figure 8 Wavelet transform of defective gearbox signal.

The wavelet packet transform (figure 9) provides not only better frequency resolution, but also better time-frequency representation (three-dimensional) than does the wavelet transform.

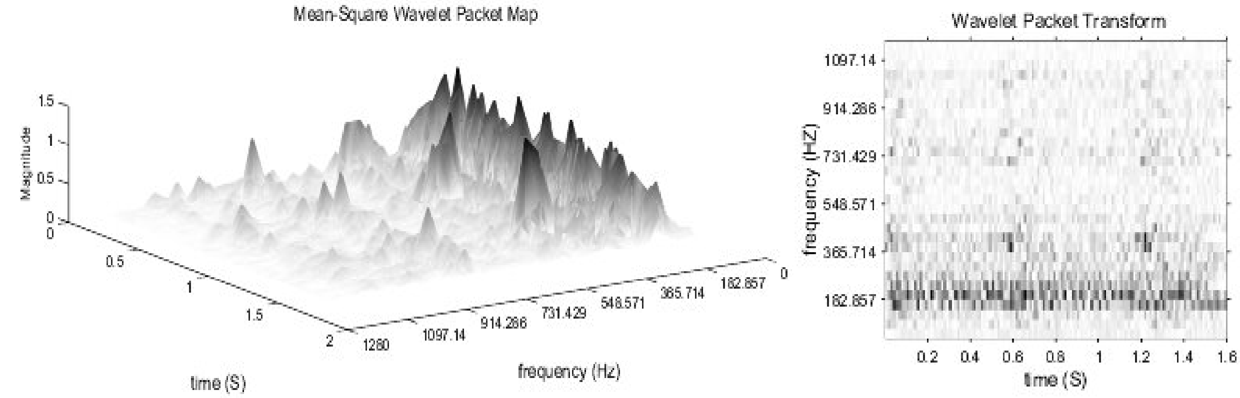

Figure 9 Wavelet packet transform of defective gearbox signal.

5. CONCLUSION

It has been shown that, although the majority of conventional methods may give good results when detecting a single fault in various simple elements of machines, no single technique can provide all the answers for all cases. It is difficult to decide which method gives the best result, in particular when the precise type of fault is not known. Time-frequency analysis provides a means to accurately identify the changing frequencies that occur with degradation; these spectral changes in turn reflect the state of the process. In this paper, a number of time-frequency methods that can be used to analyze non-stationary and time-varying signals have been described. The advantages and disadvantages of each method of time-frequency analysis have been discussed, and the benefits to be obtained from the application of these techniques in the monitoring and fault-detection of machinery have been highlighted. An in-house user-friendly time-frequency software has been introduced in this work. This software has been developed by authors in collaboration with IMS to analyze of non stationary signals which may come from machine. Finally, the advantages of the time-frequency methods have been demonstrated by using these methods on vibration signals from an industrial gearbox. The application on gear box has shown that the smoothed Wigner-Ville distribution, Short-Time Fourier transform and Wavelet packet transform were the best methods for diagnosing and locating a broken tooth in the analyzed case.

6. REFERENCES

1. Thomas M., Décembre 1986, Les vibrations comme indicateur du bon fonctionnement d'une machine, Québec Industriel
2. Thomas M., 2002, Fiabilité, maintenance prédictive et vibration de machines, Publications ETS, Montréal, 600 p.
3. Archambault R., A new method for reliable detection, diagnosis and prognosis of bearing faults, IMS, Pointe Claire, Québec, 4p.
4. D. Gabor, "*Theory of Communication*," J. IEEE (London), Vol. 93, pp. 429-457, 1946.
5. T.A.C.M. Classen and W.F.G. Mecklenbrauker, "*The Wigner Distribution - A Tool for Time-Frequency Signal Analysi Continuous Time Signals,*" Philips J. Res., Vol. 35, pp. 217-250, 1980.
6. P. Flandrin, "*Some features of time-frequency representation of multicomponent Signals,*" in Proc. IEEE 1984 Int. Conf. Acoust., Speech, Signal Processing (San Diego, CA), PP. 41.B.4.1-4, Mar. 1984.
7. J. Ville, "*Théorie et Applications de la Notion de Signal Analytique,*" Câbles et Transmission, Vol.

2A, pp. 61-74, 1948.
8. M.S. Safizadeh, A.A. Lakis and M. Thomas, Time-Frequency Algorithms and Their Application, International Journal of Computers and Their Applications, Vol. 7, No. 4, PP. 1-20, Dec. 2000.