Analysing Transient Vibration Signals with Parametric Time-Frequency Methods

Padovese, L. R
Department of Mechanical Engineering, Escola Politécnica, University of São Paulo
Av. Prof. Mello Moraes 2231, 05508-900 São Paulo, SP, Brasil
Email: lrpadove@usp.br

Abstract. This paper presents parametric time-frequency methods as a tool for transient vibration signal analysis. These methods are based on autoregressive models of the signal. A time-frequency method based on the classical Wigner-Ville Distribution is used for comparison purposes. The characteristics, advantages and disadvantages, of non-parametric and parametric methods are discussed and their performances are compared by analyzing actual data. The superiority of time-frequency resolution of the parametric methods is pointed out over that of the non-parametric ones. Two practical examples of transient vibration signals illustrate the discussion: the detection of lubricant film collapse in a journal bearing and the vibration signal issued from a damaged gearbox.

Keywords. time-frequency, parametric methods, transient vibrations

1. Introduction

Non-stationary signals, are presented in many fields of mechanical engineering area, particularly in the vibrational behavior of machinery. Machine health condition monitoring, monitoring of machining process, non destructive testing, system identification are examples of areas where the spectral analysis of transient signals are common, important and necessary.

Several time-frequency representation techniques (TFRs) have been developed, in the last decades, to allow access to the time-frequency energy behavior (amplitude and/or frequency modulations) of non-stationary signals. The most commonly used are those issued from the Cohen Class, particularly those based on a linear transform - the short term Fourier Transform (STFT) - and on a quadratic transform - the Wigner-Ville Distributions (WVD) and their relatives. Many authors have used this non-parametric TFR in order to analyse transient mechanical vibration signals. Of particular interest to this study are the works of Baydar & Ball (2001), Staszewski et all (1997), and Oehlmann et all (1997), where gear faults detection is discussed using non-parametric TFRs.

Despite the WVD potentiality to reach higher time-frequency resolution than that of STFT; WVD presents a major drawback which is the presence of cross-interference terms among the time-frequency patterns. This side effect compromises the visual interpretation of the WVD image. This problem, its consequences and the ways to overcome these difficulties will be discussed later in this paper.

Another different approach to access TFR is the use of parametric spectral estimations. This approach provides a way to overcome the major drawbacks of the linear and quadratic transform, mainly the poor time-frequency resolution and interference terms. Parametric methods are known by their high spectral resolution and it is possible to have a very good spectral estimation with very short time signal.

Initially, this work discusses the highlights of non-parametric TFRs. Next, time-frequency methods based on autoregressive models are presented and their most important characteristics are analyzed. Both classes of methods, parametric and non-parametric, are compared and their advantages and disadvantages are discussed. Two practical examples of transient vibration signals are used to illustrate the discussion: the vibration signal issued from lubricant film collapse in a journal bearing and the vibration signal issued from a damaged helical gearbox.

2. Non-Parametric Time-Frequency Representations

It is possible to classify the non-parametric TFR methods in two classes. The first one is based on the short-Fourier transform. The second employs quadratic transforms such as the Wigner-Ville Distribution (WVD) and its relatives.

2.1 The Fourier transform approach

A linear TFR based on Fourier transform (FT) can be reached by pre-windowing the signal around a chosen time, calculating its FT, and proceeding in the same way for each instant. This transform is known as Short Time-Frequency Transform (STFT). A quadratic form related with the STFT can be obtained by taking the square of this transform. It is known as spectrogram (SPEC), and measures the spectral energy density of the signal in the time-frequency plan.

$$\mathrm{SPEC(t,f)} = \left| \mathrm{STFT(t,f)} \right|^2 = \left| \int_{-\infty}^{\infty} x(\tau) h^*(\tau - t) e^{-j 2\pi f \tau} d\tau \right|^2 \tag{1}$$

The STFT time resolution is determined by the length of the selected sliding window, and the frequency resolution is determined by the bandwidth of the window. The best frequency resolution is achieved with the natural window and defined as $\Delta f = 1/D$, where D is the time duration of the window. Any other different window will degrade the resolution (Harris, 1978). The product $\Delta f \times D \geq 1$ measures the joint time-frequency resolution of the STFT method. This resolution limitation is the most significant drawback of this TFR. As others major STFT problems, it can be cited (a) the implicit windowing problem that causes the "leakage" phenomenon, and (b) the impossibility, when working with short time data, of evaluating periodogram averaging for good power spectra estimation (Marple, 1989).

2.2 The quadratic approach

The Cohen class is a general formulation for non-parametric time-frequency distribution, which includes the Wigner-Ville Distribution (WVD) and relatives (Boashash & Chen, 1992; Flandrin, 1999; Cohen, 1989, Qian & Chen, 1999). The STFT can be considered as a special case of the Cohen class. The Wigner-Ville Distribution can be defined as:

$$\mathrm{WVD(t,f)} = \int_{-\infty}^{\infty} x(t+\tau/2)x^*(t-\tau/2)e^{-j2\pi f \tau}d\tau \qquad (2)$$

Since the value of the WVD is determined by all the values of the signal (and therefore, not limited by a time window) the WVD overcomes the STFT tradeoff between time and frequency resolution (the hypothesis of short-term stationarity is not necessary). This improvement comes at a price of the appearance of spectral cross-terms, which come from WVD bilinear characteristic (Marple, 1998). This spectral interference is critical in multicomponent signals, since it makes difficult the distinction of weaker signal components and it masks spectral features.

To overcome this major drawback of the WVD, several modifications have been proposed and can be found in literature (Boashash & Chen, 1992; Flandrin, 1999). One of them, the Smooth Pseudo Wigner Ville Distribution (SPWVD), is of particular interest to this work since it will be used later to analyse experimental results. The SPWVD can be defined as:

$$\mathrm{SPWVD(t,f)} = \int_{-\infty}^{\infty} h(\tau) \int_{-\infty}^{\infty} g(t-\eta)\, x(\eta+\tau/2)\, x^*(\eta-\tau/2)\, d\eta\, e^{-j2\pi f \tau} d\tau \qquad (3)$$

where g(t) is the time smoothing window and h(t) the frequency smoothing window. With the introduction of these two windows it is possible to attenuate, to smooth the interference terms presented in the WVD, by independently choosing the type of window and its length (Conforto & D'Alessio, 1999b; Baydar & Ball, 2001, Staszewski et all, 1997).

This solution for minimizing the interference terms introduces losses in time-frequency resolution. The WVD gives the best time-frequency resolution, but has serious interference term problems. The STFT does not have interference terms but, on the order hand, presents worse time and frequency resolution. The SPWVD is a compromise between these two extremes. The results from the SPWVD depend on the choice of the type and size of the two smoothing windows that can imply in some trial and error calculations. The final performance depends on the signal composition and noise, and how well the tradeoff between cross-terms and resolution can be managed for revealing the signal information.

It is worth noticing that the signal sampling must be carefully done in order to use with the WVD or SPWVD. The Fourier transform of a signal, sampled at the Nyquist frequency, is periodic with the period of the sampling frequency. Because of the quadratic nature of the WVD the distribution is periodic, the period being half that of the Fourier Transform. This implies that it is necessary to increase the signal sample rate by a factor of 2 or else to use the analytical signal. This last option is more adequate since it brings another advantage, the elimination of part of interference terms, those that come from interference between positive and negative frequencies, since the analytical signal spectrum does not have negative frequencies (Moss et all, 1989).

3 Parametric Approach

3.1 The parametric model - Background

Many deterministic and random signals can be well approximated by a linear prediction model:

$$\mathrm{x(n)} = \tilde{\mathrm{x}}\mathrm{(n)} + \mathrm{w(n)} = \sum_{k=1}^{p} a_k x(n-k) + w(n) \qquad (4)$$

This equation represents an autoregressive (AR) model of the signal x(n). It can also be seen as a discrete form of linear differential equation of order *p* that models the signal, or else as an equation of a stationary linear all-poles filter driven by a white noise. The parameters a_k are the coefficients of prediction and w(n) is a white noise. The x(n) value is expressed as a predicted value $\tilde{x}(n)$, plus a white noise, where $\tilde{x}(n)$ is obtained as a weighted sum of *p* past values.

Once the coefficients a_k have been estimated, it can be shown that the Power Spectrum Density (PSD) of an AR model (P_{AR}) can be obtained by:

$$P_{AR}(f) = \frac{P_w}{\left| 1 + \sum_{k=1}^{p} a_k \, e^{-j2\pi f k} \right|^2} \tag{5}$$

where P_w is the power of the white noise or, in another way, the prediction error (since *w(n)* express the difference between *x(n)* and $\tilde{x}(n)$). Equation (5) is also known as Maximum Entropy Spectrum.

The value of P_w depends on the method chosen to estimate the prediction coefficients. Several methods can be found in the literature to estimated the *p* coefficients a_k. Among them, two are of special interest in this work, because of their characteristics: the covariance and Burg methods. The covariance method is understood, in this work, as the unconstrained least-square method, with minimization of the forward and backward prediction error power. The Burg method minimizes the forward and backward prediction errors in the least square sense, with the constraint that the AR coefficients need to satisfy the Levinson-Durbin recursion. More details can be found in Proakis & Manolakis, (1996). The Burg method has three major disadvantages: it presents spectral peak locations highly dependent on the initial phase, mainly for short signals; it suffers from spectral line splitting, mainly at high signal to noise ratio and it introduces spurious peaks for high order models. The covariance method does not suffer, or is only slightly affected, by these drawbacks. The only major advantage of the Burg method to the Covariance one is that the model obtained with Burg is always stable, and for the covariance, there is no guaranty of stability. However, for spectral analysis the stability of the model is not a problem (Proakis & Manolakis, 1996).

The parametric approach for calculating the signal spectrum has a major advantage, which is its high frequency resolution (Proakis & Manolakis, 1996; Marple, 1989). Since the AR model allows extrapolating the known values of the signal (over the analyzing time window), it avoids the periodicity hypothesis inherent of traditional methods. As a consequence, the P_{AR} does not present side lobe effects, and offers better frequency resolution.

An AR model implies in *a priori* knowledge (or assumption) about the process from which the signal is taken. This a priori information is expressed by the selection of the model order *p*. A good choice of this parameter is essential for a good spectral estimation. A too high order will introduce spurious frequencies and too low will smooth the spectra. Several ways can be envisage to overcome this problem. The first and intuitive approach is to calculate different spectra with increasing order, and to seek for a minimum in the prediction error power. However, the prediction error power for the least-square method (in which the Burg and covariance methods are based) decreases monotonically with increasing order *p*, which makes this approach unsuitable. The second approach is based on the use of two well known criteria, proposed by Akaike, for selecting the model order. They are the final prediction error criterion (FPE) and the Akaike information criterion (AIC) (Proakis & Manolakis, 1996). The third approach comes from the fact that the AR order is related with the number of system degrees of freedom that are presented in the signal. In this way, the order can be estimated by analyzing a traditional signal spectrum by observing the number of the main frequencies. This procedure will depend on the signal to noise ratio and on the nature of the non stationarity presented in the signal. In the present work, these last two approaches were used in a complementary way

3.2 The time moving window parametric time-frequency methods

The fact that it is possible to obtain high spectral resolution with short time windows can be used in order to construct a parametric TFR with high time-frequency resolution. This TFR can be reached by using a time moving window across the signal and then calculating the AR spectra for each time window, with Eq. 5, in a similar procedure of that for STFT TFR (Lesniak & Niitsuma,1996, Conforto & D'Alessio, 1999b, Fargetton et all, 1980; Nadine,1986). This approach assumes signal stationarity over the time window and therefore is appropriated to analyze weakly nonstationary signals.

This time-frequency representation is constructed by plotting the P_{AR} (calculated by eq. 5) at each time instant *tw*, corresponding to the center of the moving time window. Hence, P_{AR} [tw,f] is a matrix of size (*Ntw* x *p*), where *Ntw* is the number of time windows across the signal and *p* is the model order. If time calculation is important, the estimation of the covariance matrix can be implemented with a recursive algorithm (Fargetton et all, 1980; Nadine, 1986).

The tradeoff between time and frequency resolution, inherent to the STFT, and that between interference terms and time-frequency resolution, inherent to the WVD, does not exist in the case of the parametric TFR, and its final time-frequency resolution is higher than in the STFT and WVD cases.

Two different methods for AR coefficient calculations can be used, as previously discussed, to build the parametric time-frequency methods: the Burg and the Covariance. Each of these parametric TFRs has its own properties, inherited from the properties of the Burg and Covariance methods.

Despite the fact that a shorter time window does not imply in a degradation of frequency resolution, it remains to solve the problem of establishing how shorter the time window length can be. A too short window can not have enough information for a good spectral estimation. There are no specific rules to orient the taking of this decision and some tests need to be done. By experience, it was adopted as a start value the triple of the order as a minimum value for the window length (in number of points).

Two different methods for AR coefficient calculations are used in this work, the Burg and the Covariance, to build two parametric time-frequency methods: the ARBTF and the ARCTF, respectively. Each of these parametric TFRs has its own properties, inherited from the properties of the Burg and Covariance methods, previously discussed.

3.3 The time–varying parametric time-frequency methods

A more complex parametric TFR approach, known as time-varying parametric method, is obtained by adapting the model coefficients to the time-varying behavior of the signal (Lesniak & Niitsuma, 1996; Qian & Chen, 1999; Girault et all, 2000; Gustafsson, et all, 1994; Conforto & D'Alessio, 1999a, Conforto & D'Alessio, 1999b). These coefficiens are considered as a linear combination of time-varying deterministic basis functions $f_i(tn)$, and become function of time:

$$a_k(tn) = \sum_{i=0}^{d} a_{k,i} f_i(tn) \tag{6}$$

where d is the basis dimension, tn the time instant where a_k are calculate and $a_{k,i}$ are constant coefficients of the expansion. Several basis functions can be used such as Taylor series, Legendre polynomials, prolate spheroidal functions, etc..(Conforto & D'Alessio, 1999a, Conforto & D'Alessio, 199b; Gustafsson et all, 1994; Girault et all, 2000). This improvement makes possible to follow rapidly varying spectra, but it comes at a price of introducing two more *a priori* assumptions from the signal: the basis function set and the basis dimension d. The time-varying time-frequency PSD equation becomes:

$$\mathrm{P_{ARtv}}(\mathrm{tn}, \mathrm{f}) = \frac{P_w}{\left|1 + \sum_{k=1}^{p} a_k(\mathrm{tn})\, e^{-j2\pi f k}\right|^2} \tag{7}$$

4. Case Studies and Discussion

The aim of this section is to evaluate the performance of the described time-frequency methods, by analyzing actual transient signals. The performance aspects, like frequency and time resolution, and frequency bias, will be compared and discussed. The signals represent two failure problems in rotating machinery and contain time transient events.

The TFRs that will be compared are the SPWVD , ARCTF and ARBTF. These TFRs are presented in hot color, where the highest amplitudes are represented as red and the lowest as blue.

4.1 Detection of lubricant film collapse in journal bearing

This case study use vibration transient signal carrying information about the collapse of the lubricant film in a radial journal bearing, caused by a radial overload on the shaft. The measurement was done at a nominal constant shaft speed of 3400 rpm (57 Hz), and using a sample frequency of 1050 Hz. At the initial time the shaft was unload, and further the radial load was increased until twice the maximum nominal load allowed for the journal bearing be reached. It is expected that, during the loading process, the hydrodynamic flow in the bearing collapses into the boundary lubrication regime, changing the vibration behavior of the system bearing/shaft. Figure 1 (a) shows the signal in the time domain.

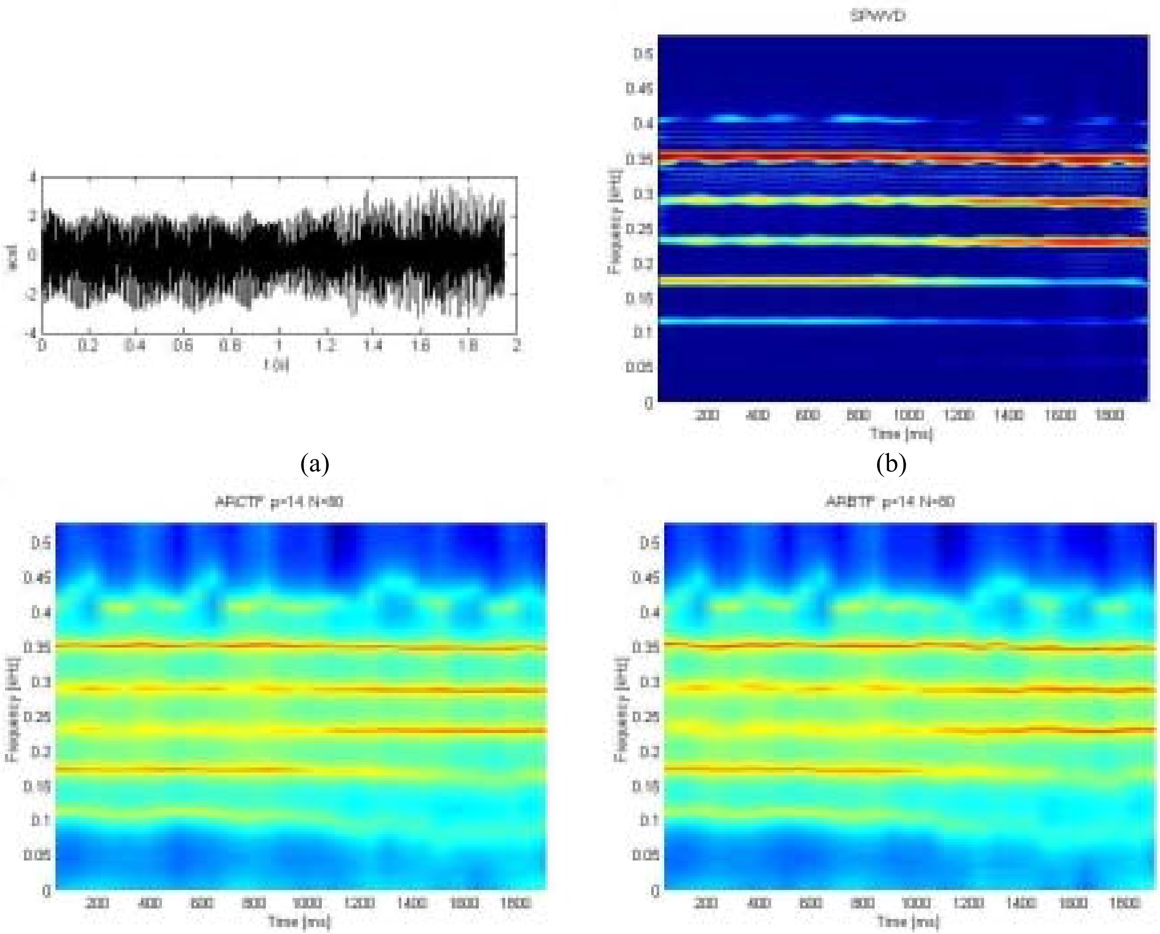
Figure 1 also presents three time-frequency representations (SPWVD, ARCTF and ARBTF) in which it is possible to observe the spectral modulation caused by the shaft speed. This modulation is represented by an equal frequency spacing (57 Hz) among the spectral lines. Each of these spectral lines represents shaft speeds harmonics (from the 2nd till the 6th).

All the TFR panels in Figure 1 allow recognizing the instant when the lubricant film collapse (≈ 1.5 s), and show the influence of the lubrication regime change (hydrodynamic to boundary) in the system dynamic. Some spectral lines are amplitude modulated after the transition. The energy of the 4th and 5th harmonic increases after the collapse, while that of the 2nd and 3rd decreases, what represents the beginning of the boundary lubrication regime.

Comparing the performance of the three time-frequency methods, all of them present enough information to allow studying the problem. It seems that the dynamic range between strongest and weakest signal components is greater for the parametric methods than for the quadratic TFR. At least, the detection of small signal levels is degraded by the interference terms, which can be seen in the SPWVD panel at high frequencies.

Regarding the computational time consumption, it is worth noticing that the parametric methods are far faster than the quadratic one.

The ARBTF representation shows some spectral fluctuation when compared with the ARCTF. This behavior is the effect of the phase influence on the spectral estimation (as discussed previously).



(a) (b)

(c) (d)

Figure 1 Time-frequency representations of the journal bearing signal. (a) time signal; (b) SPWVD (time and frequency smoothing: Hamming window = 120 samples), (c) ARCTF (*p* = 14, time window = 80), (d) ARBTF (*p* = 14, time window = 80);

4.2 Detection of tooth defect in a gearbox

In this second case, the transient signal represents the vibration behavior of a double reduction helical gearbox whose first pinion (31 teeth) has a missing tooth. The signal was sampled at 5.12 kHz, and lasts for 0.4 s. The shaft speed of the defected pinion is1400 rpm (23,3 Hz).

Figure 2 (a) and (b) show the time signal for the cases of toothless and normal gearbox conditions. Figure 2 (c) to (g) present their TFRs, where it is possible to observe the gearmesh frequency (722 Hz) and its 2nd and 3rd harmonics as straight frequency lines. These gearmesh harmonics are particularly noticeable in the normal gear case. The toothless failure introduces new spectral patterns to the normal time-frequency behavior, by affecting the meshing characteristics on mating gears. These patterns, a kind of dashed spectral lines, repeat itself at a certain time interval, which is equal to one shaft turn (almost 43 ms), and they are related with the missing contact of the faulty tooth.

It is worth observing that only the parametric TFRs (e) and (g) have enough time-frequency resolution to clearly evidence the one shaft turn time transient. In the SPWVD panel, Figure 2(c), it is possible to observe that some spectral patterns, associated with the failure, are presented, although not with the amount of details that the parametric TFRs allow. Both ARCTF and ARBTF present very close results.

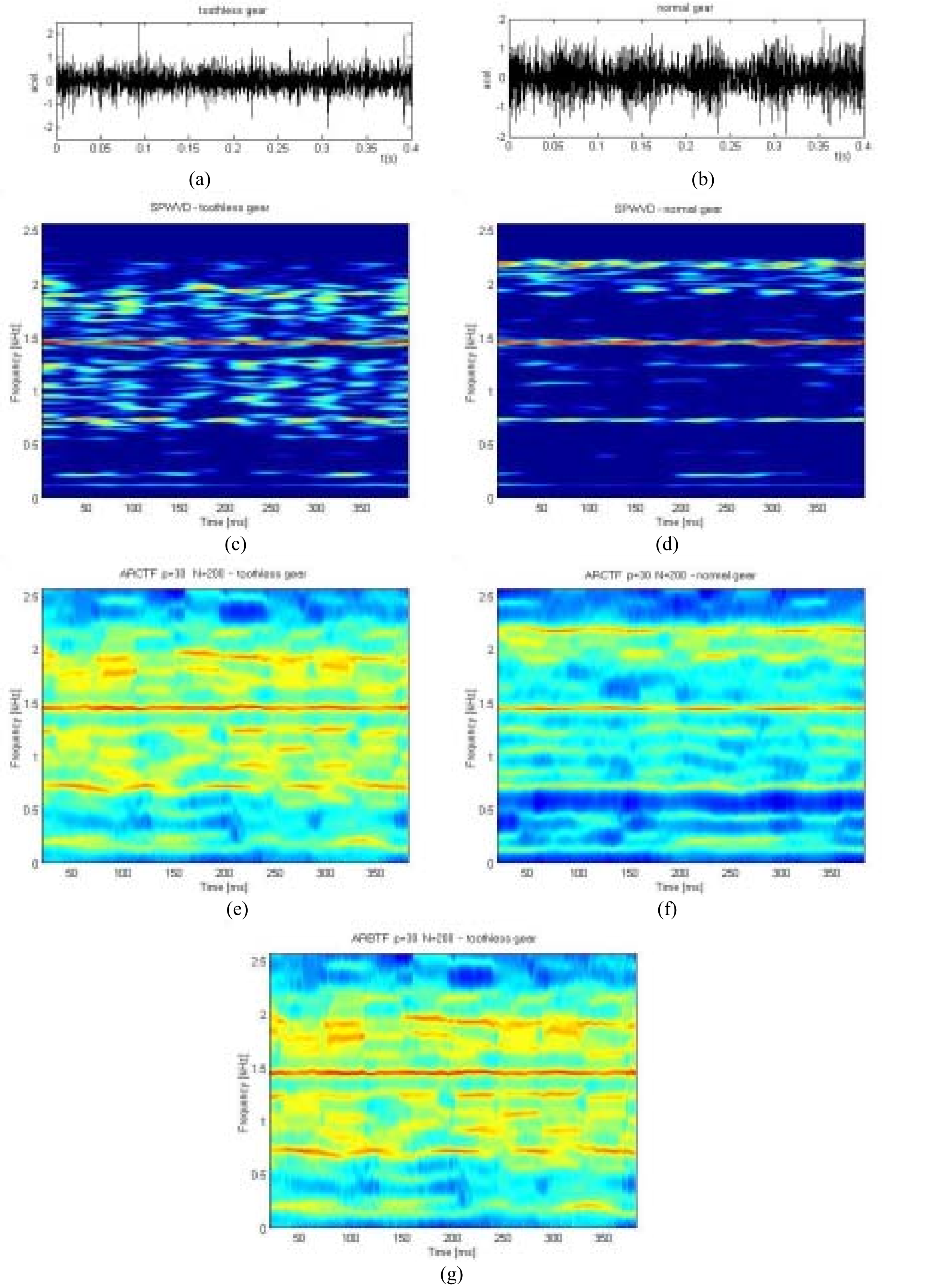


Figure 2 Time-frequency representations of the thoothless and normal condition gearbox. (a) toothless signal; (b) normal signal; (c) toothless gear SPWVD (time and frequency smoothing Hamming window = 250 samples), (d) normal gear SPWVD (time and frequency smoothing Hamming window = 250 samples); (e) toothless gear ARCTF (p = 30, time window = 200 samples), (f) normal gear ARCTF (p=30, time window = 200 samples); (g) toothless gear ARBTF (p = 30, time window = 200 samples)

5. Conclusions

In this paper parametric time-frequency techniques were proposed as a tool for analyzing transient vibration. These techniques are based on an autoregressive model approach to calculate the PSD of a moving time window signal. Two methods were used to estimate the AR coefficients: one based on the Burg method and the other, based on the Covariance one. For comparison purpose a well known quadratic time-frequency method, the Smooth Pseudo Wigner-Ville based on the Wigner-Ville Distribution, was employed

To analyze the performance of these methods, two examples of actual transient vibration signal were used.

By analyzing the actual data, it was possible to observe that the parametric methods have a performance equal or superior to that of the quadratic one. The time-frequency resolution and amplitude representation capability of the parametric techniques are superior to that of SPWVD, with a computational time consumption far lower. The parametric methods do not present the "time-frequency resolution – interference terms" tradeoff, typical of the Wigner-Ville Distributions. On the other hand, the parametric TFRs need a priori information (or assumption) about the signal, represented by the model order specification.

This paper shows that parametric time-frequency methods here presented offer a satisfactory and an alternative tool for time-frequency analysis of transient vibration signals.

Several improvements can be considered for the parametric time-frequency methods. One of them is the use of a variable time window length, adaptable to the stationarity changes of the signal. In this way, another improvement can be made by using time varying AR coefficients, what would allow the AR model to adapt to the time-varying characteristics of the signal. In this paper, a fixed AR order was used for all the moving time windows. Since the frequency pattern of the signal changes on time, the optimization of the order value for each time window would improve the methods.

6. Acknowledgements

The authors wish to thank CAPES for the partial financial support (Process PROCAD 0136-01-08).

7. References

Baydar, N.; Ball, A.; 2001; "A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville Distribution"; Mechanical Systems and Signal Processing, Vol. 15, No.6, pp. 1091-1107

Boashash, B., S.; Chen, D.; 1992; "Time-frequency signal analysis – methods and applications"; Halsted Press, NY.

Cohen L.; 1989; "Time-frequency distribution – a review"; Proc. Of the IEEE, Vol. 77, No. 7, July

Conforto, S.; D'Alessio, T.; 1999a; "Optimal estimation of power spectral density by means of a time-varying autoregressive approach"; Signal processing, Vol. 72, pp. 1-14

Conforto, S.; D'Alessio, T; 1999b; "Spectral Analysis for non-stationary signals from mechanical measurements: a parametric approach"; Mechanical Systems and Signal processing, Vol. 13 No. 3, pp. 395-411

Fargetton, H.; Gendrin, R.; Lacoume, J.L.; 1980; Adpative methods for spectral analysis of time varying signals; in Signal Processing: Theories and Applications; Edited by M.Kunt and F. de Coulon; North-holland Publishing Company.

Flandrin, P.; 1999; "Time-frequency/time scale"; Academic Press, S. Diego

Girault, J.M.; Kouamé, D.; Ouahabi, A.; Patat; F.; 2000; "Estimation of the blood Doppler frequency shift by a time-varying parametric approach"; Ultrasonics, Vol. 38, pp. 682-687

Gustafsson, F.; Gunnarsson, S.; Ljung, L.; 1994; "On time-frequency resolution of signal properties using parametric techniques"; Proc. 33rd Conf. On Decision and Control; pp. 2259-2264

Harris, F. J.; 1978; "On the use of windows for harmonic analysis with the discrete Fourier Transform"; Proc. Of the IEEE, Vol. 66 No. 1, pp. 51-83

Lezniak, A.; Niitsuma, H.; 1996; "Resolution of time-frequency analysis for seismic crosshole measurements"; J. of Applied Geophysics, Vol. 35, pp. 105-115

Marple, S.L.; 1998; "Time-Frequency signal analysis: issues and alternative methods"; Proc. IEEE-SP Intern. Symp. on Frequency and Time-Scale Analysis, pp. 329 -332

Marple, S.L.; 1989; "A tutorial overview of modern spectral estimation"; ICASSP-89., Vol. 4; pp. 2152 -2157

Moss, J.C.; Adamopoulos, P.G.; Hammond, J.K.; 1989; "Time-frequency distributions: a modification applied to the pseudo-Wigner-Ville distribution and the spectrogram"; IEEE Int. Symp. On Circuits and Systems, Vol. 2, pp. 1256-1259

Nadine, N.; 1986; An AR spectral analysis of non-stationary signals; Signal processing, Vol. 10, pp 61-74

Oehlmann, H.; Brie, D.; Tomczak, M.; Richard, A.; 1997; "A method for analysing gearbox faults using time-frequency representations"; Mechanical Systems and Signal Processing, Vol. 11 No. 4, pp. 529-545

Qian, S.; Chen, D.; 1999; "Understanding the nature of signals whose power spectra change with time – joint analysis"; IEEE Signal Processing Magazine, March, pp. 53-67

Proakis, J.G.; Manolakis, D.G.; 1996; "Digital signal processing: principles, algorithms and applications", 3rd. ed.; Upper Saddle River, Prentice Hall

Staszewski-2, W.; Worden, K.; Tomlinson, G.R.; 1997; "Time-Frequency analysis in gearbox fault detection using the Wigner-Ville Distribution and Pattern Recognition", Mechanical Systems and Signal Processing, Vol. 11, No. 5, pp. 673-692