Przegląd i analiza istniejących metod cyfrowego przetwarzania sygnałów niestacjonarnych

Autorzy: Urbanek J., Jabłoński A., Barszcz T

Sprawozdanie z wykonania zadania nr 1 projektu badawczego IP2012 061572 pt. „Opracowanie nowatorskich narzędzi analizy sygnałów wibroakustycznych generowanych przez maszyny wirnikowe pracujące w ekstremalnie zmiennych warunkach operacyjnych”

2013-9-19

1 Introduction

In recent years it had been widely noticed that analysis of nonstationary processes can be noteworthy source of information and can significantly overcome classical stationary approach [1]. Nonstationary signals are suitable for modelling processes in many fields of applications including communications, speech and audio, mechanics, geophysics, climatology, solar and space physics, optics, and biomedical engineering [2, 3]. It is due to the fact that nonstationary approach is more general than the classical Fourier analysis or the power spectrum-based approach. Several approaches that generalize the traditional concepts of Fourier analysis have been considered, including time-frequency, time-scale/wavelet analysis and cyclostationary signal analysis. In addition, techniques such as adaptive system and signal analysis, empirical mode decomposition, and other data-driven methods have been used with the purpose of modelling nonstationary phenomena.

Following chapter presents review of the most fundamental methods of signal processing dedicated for analysis of nonstationarities. For the purpose of this review presented methods has been divided into few general subgroups:

- Time-frequency methods
- Wavelets
- Instantaneous frequency tracking
- Empirical Mode Decomposition
- Cyclostationarity

Because the purpose of described research task is to select methods suitable for analysis of vibration signals generated by machinery operating under extremely varying operational conditions, introductory sub chapter will present exemplary signal used for testing of presented methods. Following task includes testing of selected, the most popular methods for witch computing tools are widely available and well documented. Therefore, not all of described methods will be tested on simulated signals.

Chapter 2 presents time-frequency based methods including spectrogram, instantaneous power spectrum and Wigner-Ville distribution. Chapter 3 contains brief presentation of time-scale/wavelet analysis methods for nonstationary signals. Chapter 4 includes review of methods dedicated for estimation of instantaneous frequency based on the signal itself as it is in authors opinion very important aspect of identification of nonstationary processes. Chapter 5 contains presentation of empirical mode decomposition, nowadays, one of the most powerful data-driven techniques for nonstationary signals decomposition. Finally, chapter six presents methods based on analysis of cyclostationarity. Additionally, it presents state-of-art attempts to generalization of cyclostationarity-based approach.

1.1 Test signals

Two generated non-stationary signals were used for test purposes. First signal consist linear chirp while the second the set of non-equally spaced pulses. Chirp signal was chosen as a popular example of non-stationary signal widely used in many works regarding analysis of non-stationarities. On the other hand it can be treated as a wage approximation of shaft-related vibration component generated during the run-up process. Frequency of the chirp signal changed linearly from 100Hz to 400Hz in one second. Sampling frequency was equal to 1kHz.

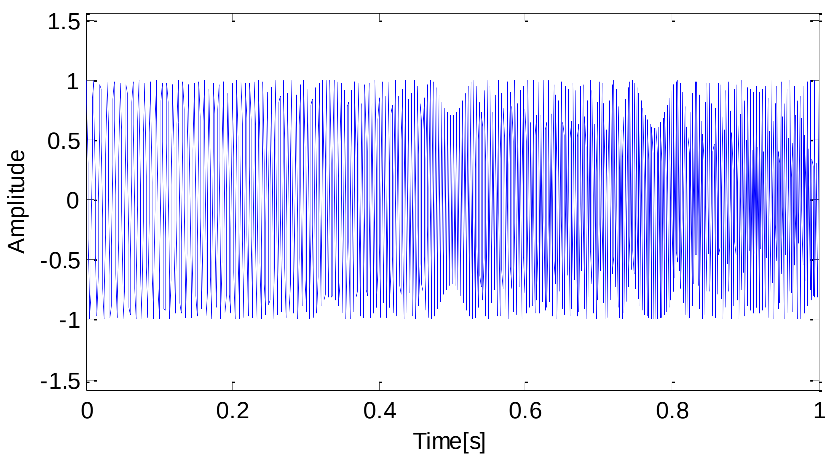


Fig.1. Time view of generated linear chirp signal.

The second signal used for further experimentation consists of the series of pulse waves of exponentially decreasing amplitude. Each pulse could be described as a short-time narrowband noise between 200Hz and 350Hz. Distance between the pulses was linearly decreasing meaning that characteristic frequency of pulses occurrence was linearly increasing. Characteristic frequency was from 10Hz to 40Hz and the time of the signal was equal to 1s. Sampling frequency was equal to 1kHz. Similarly to the first test signal, the second one can be understood as a wage approximation of rolling element bearing characteristic component generated during the run-up of machinery.

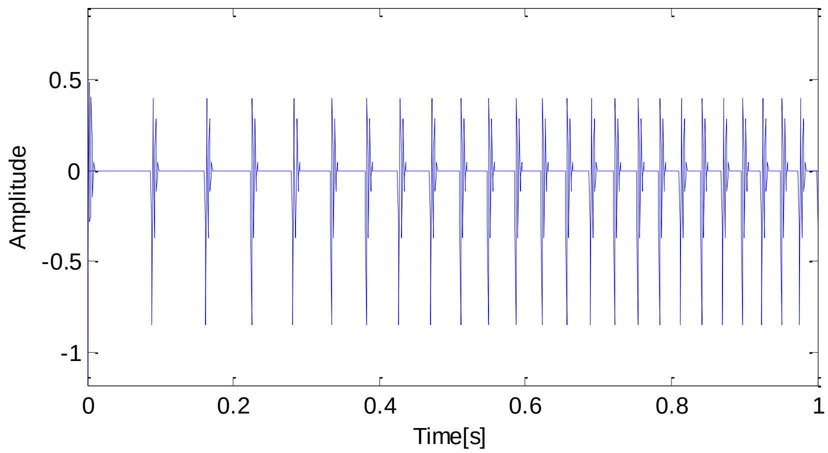


Fig.2. Time view of generated pulse signal.

2 Time-frequency representations

In general, time-frequency representation of the signal allows to observe how instantaneous energy of the signal is distributed in both; time and frequency domain. There is a large variety of time-frequency analysis methods and they all have proven its usefulness in the field of nonstationary signals. Fundamental introduction to this approach can be found in Ref. 4 and 5 where basics of relatively large selection of methods is explained.

Ref.6 for a first time introduces Instantaneous Power Spectrum (IPS) as a practical three dimensional representation of signal power. Although IPS was later used for practical nonstationary signals analysis, due to relative complexity of the method itself it was quickly replaced by other methods, less demanding in the terms of computational power.

One of the simplest, yet powerful methods of for time-frequency analysis of the signal is a spectrogram defined as a squared magnitude of short-time Fourier transform (STFT) of the signal x(t),

$$X(t, f) = \int_0^t x(t)\gamma(t-\tau)e^{-j2\pi ft}dt \quad , \qquad (1)$$

where $\gamma(t)$ is the window function [5].

In general one can define the short-time Fourier transform in terms of the output of an arbitrary bank of filters. However, we shall restrict ourselves to the much simpler case of identical, symmetric, bandpass filters uniformly spaced in frequency. The result of these simplifications is to allow the use of a single lowpass filter (window function) w(n) which determines all of the properties of the filter bank.

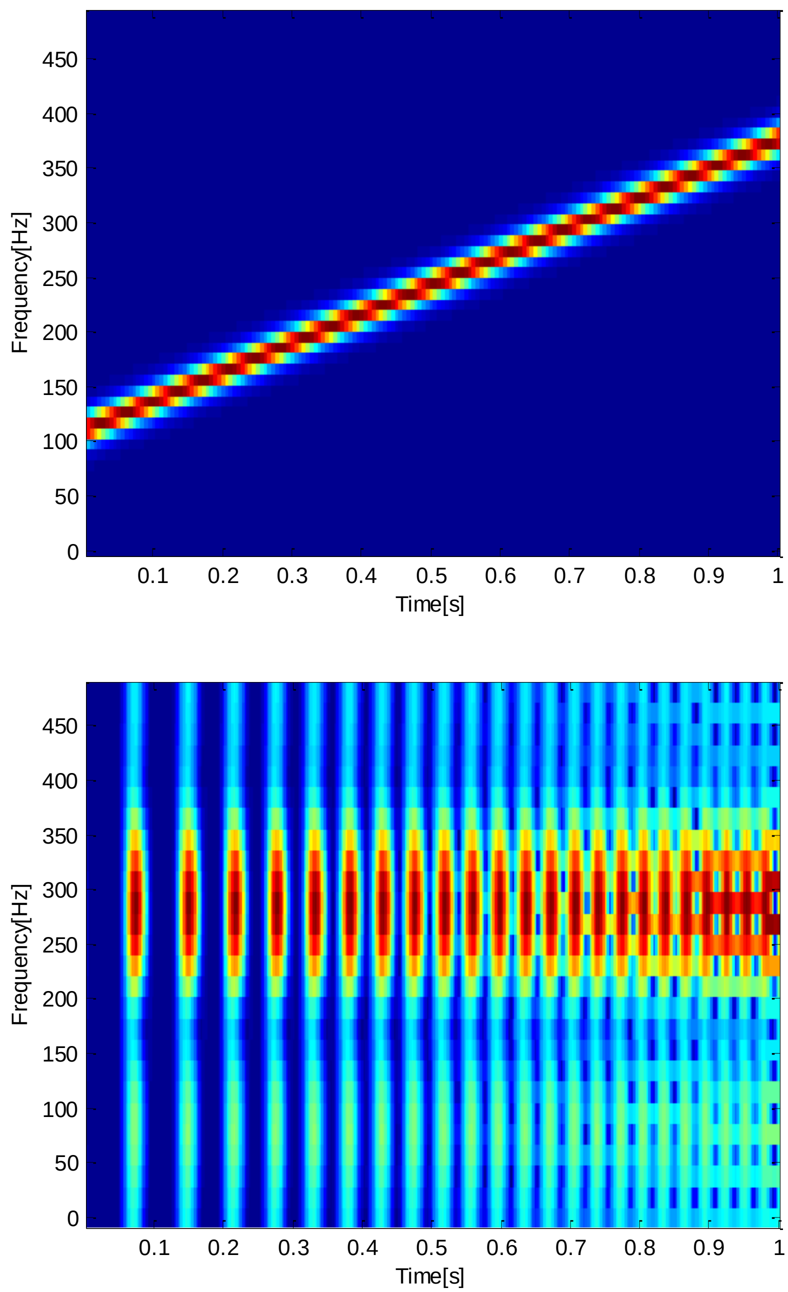


Fig.3. Exemplary spectrogram representation of test signals. Top – linear chirp, bottom – pulse signal.

Spectrograms calculated for generated test signals can be seen in fig.3. It can be noticed that for a chirp signal changes of a frequency can be clearly identified with the accuracy given by resulting frequency resolution. Frequency resolution of a spectrogram representation can be serious obstacle in identification of varying-frequency non-stationary signals. On the other hand, the one can notice that representation of pulse signal appears accurate in the terms of frequency resolution. However, due to the length of the time window used for calculation of STFT some more frequently occurring pulses might be unrecognizable.

More advanced time-frequency representations that base on the generalized concept of STFT can be found in Refs. 7 - 8. For improved accuracy of representation in frequency domain adaptive short-time Fourier transform [7] can be used. It has been proven useful especially for nonstationary signals components identification.

Another method for improvement of the time-frequency representation based on STFT was introduced in Ref.8 where the idea of reassignment vector field for time-frequency segmentation was presented. Presented approach was later on applied in the field of voice recognition, vibro-mechanics and geology. Reassignment is a non-linear method, which creates a new time-frequency representation by moving the spectrogram values away from their computation place. Reassignment focuses energy components by moving each time-frequency location to its group delay and instantaneous frequency, which represent more accurately the component energy. The obtained reassignment vector field associated to a given spectrogram describes how time-frequency locations are reassigned. On a pattern boundary, spectrogram energy will be moved inward the pattern. Consequently, all reassignment vector located on a pattern boundary will aim at the pattern. Considering that the boundaries variations are smooth enough, it can be assumed that reassignment vectors are locally parallel. The idea is to look for time-frequency locations where associated reassignment vectors and all of its neighbors have close angles. Due to the spectrogram discretization, reassignment vectors associated to contiguous time-frequency locations may aim at different points, leading to slightly different angles. In order to avoid this problem, reassignment vectors with close angles means that the absolute difference of their angles is below a threshold. This parameter also allows tracking the boundary's variation.

The concept of generalized demodulation approach to time-frequency projections can be found in Refs. 9-11. The one can notice that significant accuracy of representation of nonstationary signals on time-frequency plane can be obtain using presented approach. The demodulated analytic signal is projected onto the time-frequency plane so that, as closely as possible, each component contributes exclusively to a different 'tile' in a wavelet packet tiling of the time-frequency plane, and at each time instant the contribution to each tile definitely comes from no more than one component. A single reverse demodulation is then applied to all projected components. The resulting instantaneous frequency of each component in each tile is not constrained to a set polynomial form in time, and is readily calculated, as is the corresponding Hilbert energy spectrum [9].

Another approach to time-frequency representation of the signals is Vigner-Wille distribution [9-13]. In discrete time, the Wigner-Ville spectrum is the discrete Fourier transform of the autocovariance function. The Wigner-Ville Distribution (WVD) is defined as

$$W(t, f) = \int_{-\infty}^{\infty} x(t + \tfrac{\tau}{2}) x^*(t + \tfrac{\tau}{2}) e^{-2\pi f \tau} dt$$

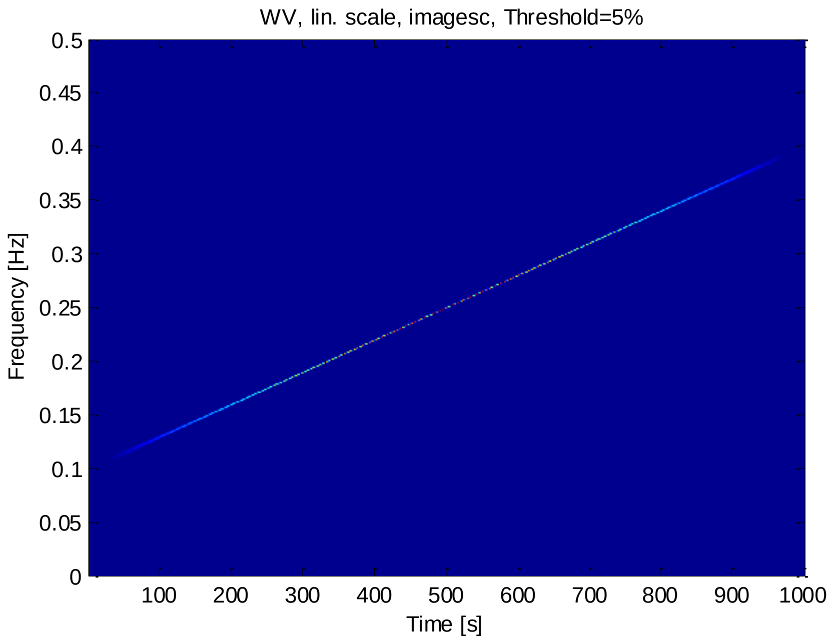
where x(t) is the signal, t the time, f the frequency and τ is the lag. Due to the quadratic nature of the distribution, the application of the WVD is limited by the presence of interference terms. These can be described considering elementary mono-components z(t) and g(t) and the WVD is then given by:

$$W_{z+g} = W_z(t, f) + W_g(t, f) + 2\,\mathrm{Re}[W_{z,g}(t, f)]$$

Wz(t; f), and Wg(t; f) are the auto-terms and Re[Wz,g(t; f)] represents the cross-terms observed between z(t) and g(t) which may lead to an erroneous visual interpretation of the time-frequency representation. Since interference terms are oscillatory they can be attenuated by means of a smoothing operation by a 2-D smoothing kernel in the Fourier domain. This is the method adopted for Smoothed Pseudo Wigner-Ville and Choi-Wlliams Distribution.

Wigner-Ville distributions calculated for exemplary signals are presented in the fig. 4. For linear chirp signal it can be noticed that representation of instantaneous frequency is extremely accurate. That is a well known property of Wigner-Ville distribution [9]. Unfortunately, linear frequency modulated signal is the only one for which cross-terms wont occur in Wigner-Ville representation.

Analysis of the exemplary pulse signal on the other hand is almost impossible using traditional Wigner-Ville distribution. It is mostly to the strong presence of cross-terms in the resulting representation. Wigner-Ville distribution is being widely applied in many fields of signal processing. Even though Vigner-Wille distribution allows for relatively high frequency resolution, unwanted cross-terms might make proper interpretation of the results impossible.



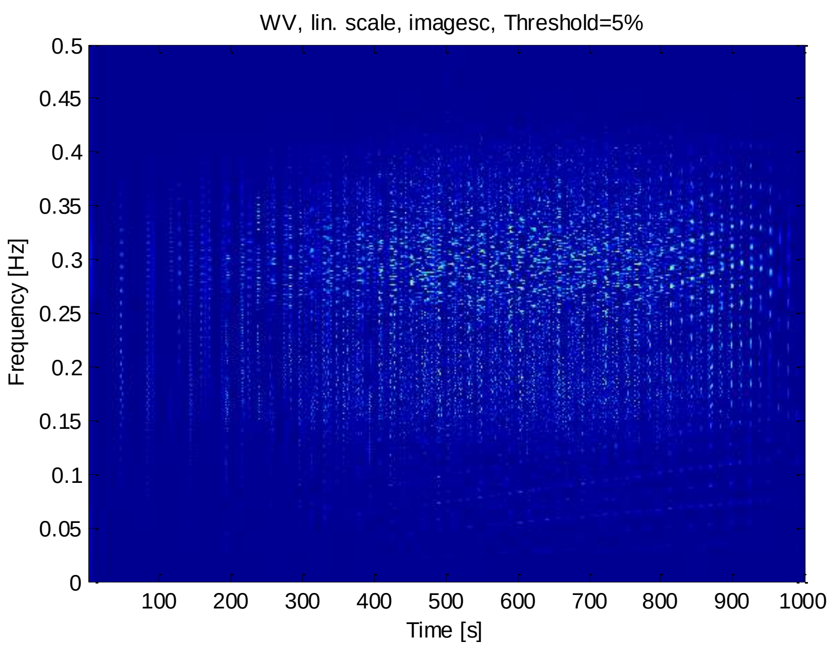


Fig.4. Exemplary Wigner-Ville distribution of the test signals. Top – linear chirp, bottom – pulse signal.

Another powerful method of accurate time-frequency representation of non-stationary signals is the method based on so-called time-frequency dictionaries [14]. An algorithm called matching pursuit decomposes any signal into a linear expansion of waveforms that belong to a redundant dictionary of functions. These waveforms are selected in order to best match the signal structures. Although a matching pursuit is nonlinear, like an orthogonal expansion, it maintains an energy conservation, which guaranties its convergence. Although, it application might be limited due to computational power caused by "trial and error" approach, the method itself appears to be both useful and accurate in the terms of identification. Methods based on called time-frequency dictionaries can lead directly to wavelet analysis described in the next chapter.

3 Wavelets

Wavelet analysis is becoming a common tool for analyzing localized variations of power within a time series. By decomposing a time series into time–scale space, one is able to determine both the dominant modes of variability and how those modes vary in time. Wavelet transform has been widely used for analysis of nonstationary, time-variant data [15, 16]. It is well known that wavelets analyze signals locally across various scales that correspond to frequencies. The continuous wavelet transform is defined as:

$$W_g(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t) g^* \left(\frac{t-b}{a} \right) \mathrm{d}t, \qquad (2)$$

where *x(t)* is the original signal, *g(t)* is the mother wavelet, *b* is a translation parameter indicating the locality, *a* is a dilation or scale parameter, and '*' indicates the complex conjugate [17]. Each value of the transform is normalized in Eq.(2) by the factor $1/\sqrt{|a|}$ to ensure that the integral energy given by each translated and dilated wavelet is independent of the dilation *a*. Various wavelet functions can be used in practice. The wavelet transform decomposes the signal into different scales with different levels of resolution by dilating a single prototype function, the mother wavelet. Furthermore, a mother wavelet has to satisfy that it has a zero net area, which suggest that the transformation kernel of the wavelet transform is a compactly support function (localized in time), thereby offering the potential to capture of low-energy pulses which normally occur in a short period of time. The continuous wavelet transform is one of the most suitable signal processing tools for detection of singularities, as discussed in Ref.18 -19. The Mexican hat wavelet is one of the best wavelet functions used for that task [20]. It was originally developed for multiscale edge detection in computer vision. The amplitude of the continuous wavelet transform exhibits large values at small scales for signal local features such as discontinuities, sharp frequency changes, and impulses.

As in previous sections, tests were preformed using generated signals. Mexican hat wavelet was used as a mother wavelet for the continuous wavelet transform. For the chirp signal (fig. 5. - top) component of linear frequency modulation can be recognized between 66 and 105 scale. However, beside that =, resulting scalogram is rich in other less intense components, witch makes its interpretation relatively