Available online at www.sciencedirect.com
SCIENCE DIRECT
Biomedical Signal Processing and Control 1 (2006) 33–43
Biomedical Signal Processing and Control
www.elsevier.com/locate/bspc

Time-frequency analysis of normal and abnormal biological signals

Seedahmed S. Mahmoud, Zahir M. Hussain, Irena Cosic, Qiang Fang *

School of Electrical and Computer Engineering, RMIT, 124 Latrobe Street, Melbourne, Vict. 3000, Australia

Received 24 November 2005; received in revised form 9 January 2006; accepted 7 February 2006
Available online 24 March 2006

Abstract

Due to the non-stationary, multicomponent nature of biomedical signals, the use of time-frequency analysis can be inevitable for these signals. The choice of the proper time-frequency distribution (TFD) that can reveal the exact multicomponent structure of biological signals is vital in many applications, including the diagnosis of medical abnormalities. In this paper, the instantaneous frequency (IF) estimation using four well-known TFDs is applied for analyzing biological signals. These TFDs are: the Wigner–Ville distribution (WVD), the Choi–Williams distribution (CWD), the Exponential T-distribution (ETD) and the Hyperbolic T-distribution (HTD). Their performance over normal and abnormal biological signals as well as over multicomponent frequency modulation (FM) signals in additive Gaussian noise was compared. Moreover, the feasibility of utilizing the wavelet transform (WT) in IF estimation is also studied. The biological signals considered in this work are the surface electromyogram (SEMG) with the presence of ECG noise and abnormal cardiac signals. The abnormal cardiac signals were taken from a patient with malignant ventricular arrhythmia, and a patient with supraventricular arrhythmia. Simulation results showed that the HTD has a superior performance, in terms of resolution and cross-terms reduction, as compared to other time-frequency distributions.

Keywords: Biological signal; ECG; EMG; Time-frequency analysis

1. Introduction

Traditionally, biological signals such as electrocardiogram (ECG), electroencephalogram (EEG) and electromyogram (EMG) are analyzed in the time-domain by skilled physicians. However, pathological conditions may not always be obvious in the time-domain signal. For example, in a subject with arrhythmia or a subject undergoing an epileptic seizure, certain rhythms become more prominent where they experience change in amplitude and frequency. Investigation of such biological signals can assist the surgeon in deciding on surgical intervention [1]. Sometimes, biological signals accompany by noise that may consist of artefact or environmental interference. These facts have motivated the use of frequency domain techniques, such as Fourier transform (FT), for analysis [2]. However, as the ECG and all other biological signals belong to the family of multicomponent nonstationary signals [3], accurate time-varying spectral estimates can be extremely difficult to obtain. However, a proper time-frequency distribution (TFD) can tackle this problem and reveal the multicomponent nature of such signals.

Time-frequency analysis plays a significant role in signal processing and biomedical engineering [4,5]. The instantaneous frequency (IF) is an important concept in time-frequency analysis, especially when analyzing multicomponent signals. The concept of the instantaneous frequency can be found in [6,4,5,7]. Every TFD has a ridge or concentration of energy in the time-frequency plane around the instantaneous frequency (IF) of each component. As such, it can be used to estimate the frequency variation of the ECG (or any biomedical) signal over time. This presentation provides information where the time-domain and frequency-domain may fail to produce. In particular, it uses to detect the QRS complex and arrhythmia [8]. Methods of IF estimation can be classified into two major categories: parametric and non-parametric. Parametric IF estimation methods are complicated and time-consuming, hence not suitable for real-time applications. Non-parametric IF estimation for multicomponent nonstationary signals is an important (and unresolved) issue in signal processing [6,9]. Although Fourier analysis can reveal the multicomponent

* Corresponding author. Tel.: +61 3 9925 2432; fax: +61 3 9925 2007.
E-mail addresses: mahmoud.seedahmed@rmit.edu.au (S.S. Mahmoud), zahir.hussain@rmit.edu.au (Z.M. Hussain), irena.cosic@rmit.edu.au (I. Cosic), john.fang@rmit.edu.au (Q. Fang).

1746-8094/$ – see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bspc.2006.02.001

nature of signals in some special cases (e.g., sum of sinusoids), it is only time-frequency analysis that can be used for general IF estimation for multicomponent signals through concentrating the signal energy in the time-frequency plane around the component IF laws [4]. There are many TFDs in active use, the most significant class of TFDs is known as the Quadratic Class or Cohen's Class [4]. However, quadratic time-frequency distributions suffer from the presence of cross-terms when used to analyze multicomponent signals [4,5,7]. Cross-terms are fictitious concentrations of energy, resulting from the quadratic nature of Cohen's Class, which can obscure the real features of interest in the signal. On the other hand, joint time-frequency resolution is another significant character that is different for different TFDs. In some applications (e.g., biomedical signal analysis), we may be confronted by multiple components with narrow separation in time, frequency or both; in such a case, many TFDs fail to reveal the true structure, as many components will overlap due to bad resolution. Considerable efforts have been made to define TFDs which reduce the effect of cross-terms while improving the time-frequency resolution [4,7]. However, there is always a compromise between these two requirements. TFD's have different performances in this respect and the choice of the proper TFD is application-dependent. This paper is organized as follows. In Section 2, we will explain the process of IF estimation, both in principle and numerically. In Section 3, the acquisition of normal and abnormal biological signals is explained. In Section 4, we will discuss four time-frequency distributions (TFDs) that will be used for the comparison purposes in this paper. Two of these TFDs are classified as time-only kernels distributions. Section 5 discusses the process of frequency estimation using this class, the T-distribution. Extensive performance comparison of the four TFDs over biological signals and noisy mono- and multicomponent FM signals will be presented in Section 6.

2. Instantaneous frequency estimation

Consider a real signal $s(t)$. To avoid aliasing in the digital implementation of the TFD for this signal, we always consider its analytic associate $z(t) = s(t) + j\hat{s}(t)$, where $\hat{s}(t)$ is the Hilbert transform of $s(t)$ [5].

Biological signals such as EEG can be described as a nonstationary random signal composed of an amplitude modulation-frequency modulation (AM-FM) part in additive stationary random noise, with low signal-to-noise ratio (SNR) [10,11]. As the AM variation is normally slow without sudden or abrupt changes, noisy multicomponent IF estimation techniques are applicable to biomedical signals. To verify the concept of IF, we consider an analytic FM signal of the form

$$z(t) = \alpha e^{j\phi(t)} + \in(t) \tag{1}$$

where the amplitude α is constant, $\varphi(t)$ is the phase of the analytic signal, and $\in(t)$ is a complex-valued white Gaussian noise with independent identically distributed (i.i.d.) real and imaginary parts with total variance $\sigma^2_{\in}$. The instantaneous frequency of $z(t)$ is given by the derivative of the phase as follows:

$$f_i(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt}. \tag{2}$$

We assume that $f_i(t)$ is an arbitrary, smooth and differentiable function of time with bounded derivatives of all orders.

The continuous time-frequency distribution of the analytic signal $z(t)$ associated with the original real signal $s(t)$ can be expressed as follows [5,6]:

$$\rho(t, f) = \underset{\tau \to f}{F}[G(t,\tau) \underset{(t)}{*} K_z(t,\tau)] \tag{3}$$

where $K_z(t, \tau) = z(t + \tau/2)z^*(t - \tau/2)$ is the instantaneous auto-correlation product, F is the Fourier transform, $G(t, \tau)$ is the time-lag kernel, and $*_{(t)}$ denotes time convolution. It is well-known that the kernel can completely characterize the TFD and its properties (e.g., resolution) [12]. The kernel can also be expressed in the Doppler-lag domain as follows:

$$G(t,\tau) = \underset{\nu \to t}{F^{-1}}\{g(\nu,\tau)\}. \tag{4}$$

For practical implementation, we need the discrete version of the TFD. First, in the discrete lag-domain, the TFD $\rho(t, f)$ can be expressed as follows:

$$\rho(t, f) = \int_{-\infty}^{\infty}\sum_{m=-\infty}^{\infty} K_z(u, 2mT)G(t-u, 2mT)e^{-j4\pi fmT}\,du \tag{5}$$

where m is an integer and T is the sampling interval. Second, if $\rho(t, f)$ is discretized over time and frequency we get:

$$\rho(n,k) = \sum_{l=-N}^{N-1}\sum_{m=-N}^{N-1} K_z(lT, 2mT)G(nT - lT, 2mT)e^{-j2\pi(km)/(2N)} \tag{6}$$

where $2N$ is the total number of signal samples. The implementation discrete frequency is given by $f_k = k/4NT$. Since all TFDs has a peak or a ridge around the IF, then the IF estimate will be a solution of the following optimization problem:

$$\hat{f}_i(t) = \arg[\max_f \rho(t, f)]; \qquad 0 \le f \le \frac{f_s}{2} \tag{7}$$

here $f_s = 1/T$ is the sampling frequency. The frequency estimation error is the difference between the actual value in Eq. (2) and the estimate in Eq. (7) as follows:

$$\Delta\hat{f}_i(t) = f_i(t) - \hat{f}_i(t) = \frac{\phi'(t)}{2\pi - \hat{f}_i(t)}. \tag{8}$$

The most important factors that decide the quality of estimation are the bias and the variance of the estimate. In the above IF estimate, the bias and variance can be described as follows:

$$B(\hat{f}_i(t)) = \varepsilon[\Delta\hat{f}_i(t)] = f_i(t) - \varepsilon[\hat{fi}(t)]$$
$$V(\hat{f}_i(t)) = \varepsilon[\Delta\hat{f}_i(t)]^2 = \varepsilon[\{f_i(t) - \hat{f}_i(t)\}^2]. \tag{9}$$