ARTICLE IN PRESS YDSPR:2229

Digital Signal Processing ••• (••••) •••–•••

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Short-time Fourier transform with the window size fixed in the frequency domain

Carlos Mateo [a,*], Juan Antonio Talavera [b]

[a] Institute for Research in Technology, School of Engineering (ICAI), Universidad Pontificia Comillas, Madrid, 28015, Spain
[b] School of Engineering (ICAI), Universidad Pontificia Comillas, Madrid, 28015, Spain

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Short-time Fourier transform
Time–frequency domain
Window size
Multi-resolution
Electrocardiogram

ABSTRACT

The Short-Time Fourier Transform (STFT) is widely used to convert signals from the time domain into a time–frequency representation. This representation has well-known limitations regarding time–frequency resolution. In this paper we use the basic concept of the Short-Time Fourier Transform, but fix the window size in the frequency domain instead of in the time domain. This approach is simpler than similar existing methods, such as adaptive STFT and multi-resolution STFT, and in particular it requires neither the band-pass filter banks of multi-resolution techniques, nor the evaluation of local signal characteristics of adaptive techniques. Three case studies are analyzed and the results show that the proposed method allows better identification of signal components compared to standard STFT, multi-resolution STFT and Adaptive Optimal-Kernel Time Frequency Representation, although the method is not computationally efficient in its present form. Some synthetic and real world signals are used to demonstrate the effectiveness of the proposed technique.

1. Introduction

Fourier Transform (FT), as well as its discrete signal versions converts a signal from the time domain into the frequency domain [1,2]. For signals whose frequency content changes over time it is even more appropriate to analyze them in the time–frequency domain, although this may introduce some additional redundant information. A fundamental property of time–frequency representations (TFR) is related to the manner in which they depend upon the signal [3]. This dependency may be linear, quadratic or non-linear. All linear TFRs, such as the Short-Time Fourier Transform (STFT) and wavelets, satisfy the superposition or linearity principle. The STFT has been widely used for processing signals, for example in image processing [4], speech [5], engineering [6,7], biology and medicine [8]. The STFT adds a time dimension by segmenting a non stationary signal into many frames that should contain quasi-stationary parts, and uses a window function to reduce the side lobes in the spectra. However, this transform has the drawback of having a fixed window size. On one hand, long windows provide better frequency resolution but poor time resolution. On the other hand, short windows are appropriate for better time resolution but not for lower frequency resolution [9]. The properties of standard STFT have been extensively studied, including inversion of the transform to reconstruct the original signal, as well as shift invariance and rotation invariance properties [10–12]. The reassignment method can be used to improve the readability of time–frequency and time-scale representations [13,14]. This method creates a modified version of the representation, by moving the time–frequency points away from the location where they are computed to a more appropriate one. Synchrosqueezing is a special case of reassignment with the additional advantage of allowing for reconstruction [15,16]. It can also be used to analyze the dynamics of the periodicity of a given signal, such as time-varying frequency and amplitude [17]. However, synchrosqueezing does not change the qualitative behavior of the TFR and does not increase the time or frequency resolution of the transform, improving only the readability of the TFR [18].

Besides, many alternatives to the standard STFT have been proposed [19–24], including interference removal [25]. Also currently under study is the idea of improving the resolution of STFT by combining STFTs of different window lengths [22]. Multi-resolution is a method whereby the signal is divided into frequency bands and each band is processed with a different window size [26,27]. However, a band-pass filter has to be applied to separate the bands where different window sizes are used. In order to address the problems of the fixed window size of STFT, adaptive techniques (ASTFT) propose an adjustment of the window size depending on

* Corresponding author.
E-mail address: cmateo@comillas.edu (C. Mateo).

https://doi.org/10.1016/j.dsp.2017.11.003
1051-2004/

Please cite this article in press as: C. Mateo, J.A. Talavera, Short-time Fourier transform with the window size fixed in the frequency domain, Digit. Signal Process. (2017), https://doi.org/10.1016/j.dsp.2017.11.003

local signal characteristics [28,29]. However, they usually suffer heavy computational complexity [30]. Adaptive STFT techniques propose, for example, using a large window when the derivative of the instantaneous frequency varies smoothly and a narrow window when it varies sharply [29]. Unlike these techniques, which use different window sizes for different time instants, we propose the use of different window sizes for different frequencies. Our approach would be closer to multi-resolution techniques [27], insofar as it uses different window sizes for different frequencies. However, unlike multi-resolution, our approach does not require the use of band-pass filter banks and is not limited to a reduced number of window sizes.

In this paper, the basic concept of STFT is maintained; however, the window size is fixed in the frequency domain (STFT-FD) rather than in the time domain. Our objective is the same as that of multi-resolution; using small windows for high frequencies and large windows for low frequencies. Longer windows on low frequencies, allow getting better frequency resolution [27]. At the higher frequencies, lower frequency resolution is required, and smaller windows allow having more accurate time resolution to correctly detect transients. The key element for achieving this objective with the proposed approach is to define the window size as a fixed number of cycles of each frequency, rather than a fixed time length. With this approach, the window size is not adjusted depending on the local signal characteristics, as in ASTFT. However, as the window size is frequency dependent, it is possible to address the well-known problems of the STFT without increasing the required computation time excessively. With the selection of a window in the frequency domain, band-pass filter banks are not required.

The Continuous Wavelet Transform, which is based on wavelet analysis, can be considered another TFR which can be used as an alternative to STFT [31]. Wavelets rely on the use of a mother wavelet function which can be scaled and shifted in order to correlate with the anomalies or events of the signals. In this sense, the proposed approach has similarities with wavelets but its core methodology is different. The S transform is an extension of the ideas of the Continuous Wavelet Transform that overcomes the STFT by using a moving and scalable localizing Gaussian window [32]. Similarly to the S transform, our approach also provides frequency-dependent resolution and has a direct relationship with the Fourier spectrum. Other authors propose combining several techniques [33]. STFT, wavelets, time–frequency varying autoregressive process and kernel estimators are all part of the methodological approach in [34]. In order to better understand algorithmic and performance differences with wavelets, results are compared to the Morlet wavelet transform in the complementary parallel paper in SoftwareX [35].

On the other hand, in this paper, the results of our proposed approach are contrasted with standard STFT, with multi-resolution STFT [24] and with an Adaptive Optimal-Kernel Time Frequency Representation (AOK-TFR) [36], which can be classified as a signal-dependent time–frequency representation, whose kernel changes over time to be able to match the local signal characteristics. Section 2 describes the methodology and presents a formulation. Section 3 proposes two case studies, the results of which are shown in Section 4. Finally, Section 5 summarizes the main conclusions.

2. Methodology

We consider a time-constrained discrete signal $x(t)$, where t is a discrete index representing time or space and NS is the number of coefficients, which corresponds to the number of samples in the case of a sampled signal.

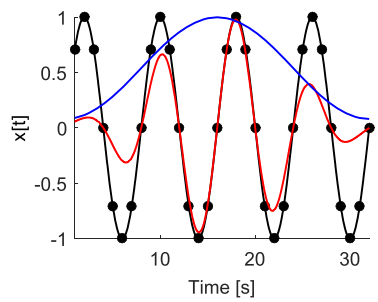

Fig. 1. Representation of a frequency component with $p = 8$ samples/cycle, $\mathrm{NC} = 4$ cycles, $\mathrm{NW}(f) = 32$ samples. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

We consider the following window.

$$w(n), \quad n \in \{1, \ldots, \mathrm{NW}(f)\} \tag{1}$$

We have defined the window with a time index n, which starts at 1; it also has a parameter $\mathrm{NW}(f)$ which is the number of samples of the window.

If the sampling interval is Ts, then the frequency corresponding to a given number of samples per cycle (p) is:

$$f = \frac{1}{p \cdot \mathrm{Ts}} \tag{2}$$

In this paper we set the window size as frequency dependent. In particular, we define NC as the number of periods or cycles within the window function. The following equation determines the window size (in number of samples) for each frequency.

$$\mathrm{NW}(f) = p \cdot \mathrm{NC} = \frac{\mathrm{NC}}{f \cdot \mathrm{Ts}} \tag{3}$$

Window size will depend on frequency. For example, if NC is set to 4, for a frequency of $f = 50$ Hz, and period $T = 20$ ms, the window size will be 4 cycles (80 ms). For a frequency of $f = 100$ Hz, $T = 10$ ms, the window size will be 4 cycles (40 ms). The parameter NC determines how locally we consider the STFT-FD. Low NC will mean that we compute just a few cycles of every frequency component. High NC values mean that we consider many cycles to compute the transform. In this way, with the proposed approach we can simultaneously adapt the representation to several frequencies. Khan et al. indicate that an ideal window would be one that gives maximum normalized energy [22]. The problem is then finding the optimal window size. However, this falls outside the scope of the present paper.

As a non-square window is also applied to the signal, it is not recommended to use a value of $\mathrm{NC} = 1$. In our case studies, we have selected a value of NC of 4 or 8 in order to have at least a couple of cycles before the window function softens the transform.

Let $X(t, f)$ be the STFT-FD transform of the discrete signal $x(t)$. For a given instant t and frequency f, the number of samples to consider will be dependent on frequency. The number of samples of the window will be NC times the number of samples per cycle (p) of each frequency component.

Several types of windows can be used, as in traditional STFT. In our case we have used the N-point symmetric Hamming window. Fig. 1 shows an example of a frequency component (in black) with $p = 8$ samples/cycle; considering $\mathrm{NC} = 4$ cycles. The window function and their product are also represented in blue and red respectively. The size is $\mathrm{NW}(f) = 32$ samples.

The product of the sinusoid and the Hamming function (represented in red) would be similar to the mother wavelet concept in wavelets, and scaled with the p parameter for different frequencies, similarly to wavelets which are also scaled. In the same way

Please cite this article in press as: C. Mateo, J.A. Talavera, Short-time Fourier transform with the window size fixed in the frequency domain, Digit. Signal Process. (2017), https://doi.org/10.1016/j.dsp.2017.11.003