Signal Processing 132 (2017) 170–176

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Reassigned time–frequency representations of discrete time signals and application to the Constant-Q Transform

Sébastien Fenet[a,*], Roland Badeau[b], Gaël Richard[b]

[a] Anthéa SAS, rue des Verchères, 01320 Chatillon-la-Palud, France
[b] LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, 46 rue Barrault, 75013 Paris, France

ARTICLE INFO

Keywords:
Reassigned time–frequency representations
Reassigned CQT
Discrete signals
Synchrosqueezing

ABSTRACT

In this paper we provide a formal justification of the use of time–frequency reassignment techniques on time–frequency transforms of discrete time signals. State of the art techniques indeed rely on formulae established in the continuous case which are applied, in a somehow inaccurate manner, to discrete time signals. Here, we formally derive a general framework for discrete time reassignment. To illustrate its applicability and generality, this framework is applied to a specific transform: the Constant-Q Transform.

1. Introduction

Time–frequency reassignment has received great attention over the last decades, especially for the task of sinusoidal parameter estimation in noisy data. Numerous methods have been developed based on Fourier analysis [6,1,7], on subspace decomposition [32,33] or on more general models such as AM/FM models [2,8]. Time–frequency reassignment methods aim at providing enhanced time–frequency representations with an improved resolution in both time and frequency. To this end, these methods propose to assign the energy computed at some time–frequency point in the signal to a different point in the time–frequency plane that depends on the window used for the spectral computation.

Time–frequency reassignment methods emerge from the idea first proposed by Kodera [22]. This original approach uses the phase information of the time–frequency representation and remains difficult to use in practice. Later on, Auger and Flandrin [6] proposed a new closed-form solution to this problem which applies to a wide variety of time–frequency representations and relies on much more straightforward computations of the reassigned indexes. This work has opened the door to the use of time–frequency reassignment in numerous domains such as physics [26], radar imaging [30] or audio [28]. It has also led to the development of numerous extensions and adaptations of the original method [21,5,27].

Another solution to the problem, named *synchrosqueezing*, has been proposed by Daubechies and Maes in [13]. Notably because it offers the ability to reconstruct the time signals, the technique has drawn a lot of interest and has become the root of multiple applications and enhancements [12,23,3]. Although synchrosqueezing was initially presented as a distinct technique from Auger and Flandrin's reassignment method, the strong connection that exists between the two has been clarified in [4].

Traditionally, reassignment calculations are carried out in the context of continuous time signals (see [18] for a review) while most applications involve discrete time signals. In practice, all results obtained in the continuous time case are applied, without detailed justification, to the discrete case.

In this work, we propose a formal framework for the computation of the reassigned transforms which fully takes into account the discrete time aspect. Our approach consists of expressing the magnitude of the time–frequency transform of a discrete signal as a mass function in the time–frequency plane and in assigning the energy to the centre of mass of this representation. Interestingly, the obtained mathematical expressions are very similar to the classical expressions proved in the continuous time case. The main advantage of our approach is that we obtain exact expressions of the solution when discrete signals are considered. This opens the door to the implementation of exact solutions and, should the implementation require an approximate solution, we are able to characterise the introduced error. To some extent, this work also gives a formal justification of the common approximation made when applying continuous time formulae to discrete time signals.

The paper is organised as follows. The mathematical model and the derivations of closed-form expressions for reassigned time and frequency indexes are provided in Section 2. We then discuss in the subsequent section the merit of the proposed solution. An application of the framework to a specific time–frequency representation, the Constant Q Transform (CQT), is finally proposed in Section 4.

* Corresponding author.
E-mail address: Sebastien.Fenet@Zoho.com (S. Fenet).

http://dx.doi.org/10.1016/j.sigpro.2016.10.008
Received 14 April 2016; Received in revised form 7 October 2016; Accepted 13 October 2016
Available online 14 October 2016
0165-1684/

2. Mathematical model

2.1. Traditional time–frequency representations

As stated in the introduction, our aim is to derive the mathematical formulation of the reassigned time–frequency representation of a discrete numerical signal.

We thus consider a discrete time signal x. Such a signal maps any discrete index $n \in \mathbb{Z}$ to a complex value $x_n \in \mathbb{C}$. Its frequency content $X(\xi)$ is defined for any normalised frequency $\xi \in \mathbb{R}/\mathbb{Z}$ by the Discrete Time Fourier Transform (we use the standard notation $\mathbb{R}/\mathbb{Z}$ to indicate that the normalised frequency is defined modulo 1).

The information that one reads in a time–frequency representation of x is the amount of energy in x at time $t \in \mathbb{R}$ and normalised frequency $\nu \in [0; 1]$. In order to evaluate this energy, the scalar product between x and a kernel is computed. The kernel consists of a windowing sequence $h^{t,\nu}$, centred on t, multiplied by the harmonic function of frequency ν. Let us note that the (t, ν) exponent makes it explicit that the windowing sequence depends on the time of interest and may also depend on the frequency of interest. More precisely, the transforms of x that fall within the scope of this paper can be written in the following form:

$$\mathcal{T}^{x,h^{t,\nu}}(t, \nu) = \sum_{n\in\mathbb{Z}} h_n^{t,\nu} x_n \mathrm{e}^{-j2\pi\nu n} \tag{1}$$

The time–frequency representation at time t and frequency ν is finally obtained by considering the squared magnitude of the transform:

$$s(t, \nu) = |\mathcal{T}^{x,h^{t,\nu}}(t, \nu)|^2. \tag{2}$$

It is interesting here to recall the Heisenberg–Gabor limit [17] that constrains the design of the windowing sequence $h^{t,\nu}$. More precisely, the Gabor limit states that there is a trade-off between the temporal and spectral resolutions when representing a signal in the time–frequency plane. In practice, adjusting the support of the windowing sequence is a direct way to tune this trade-off. A wide support will result in a precise frequency resolution with a poor temporal resolution. Conversely, a narrow support will provide a good temporal resolution at the cost of the frequency resolution. In order to ensure consistency, we consider that the windowing sequences are of finite support and that they are normalised by the size of their supports. For instance, with h being a continuous window function of finite temporal support, the windowing sequence is defined by $h_n^{t,\nu} = h(n - t)$ in the case of a Short-Term Fourier Transform (STFT) or by $h_n^{t,\nu} = \nu h(\nu(n - t))$ in the case of a Constant-Q Transform (CQT) for a set of frequencies ν within $[0; \frac{1}{2}]$ (see Section 4 for more details).

In addition to constraining the design of the window, the choice of a given time–frequency transform also determines the set of time–frequency points (t, ν) at which the representation is evaluated. Typically, a Short-Time Fourier Transform with a temporal hop size Δ_t and a spectral hop size Δ_ν is obtained with the following set of points: $\{(t_0 + k\Delta_t, \nu_0 + k'\Delta_\nu) \text{ for } (k, k') \in \mathbb{N}^2)\}$. In contrast, the Constant-Q Transform, whose spectral geometric progression is often denoted by $2^{\frac{1}{r}}$ (r being the number of bins per octave), is obtained with the set $\{(t_0 + k\Delta_t, \nu_0 2^{\frac{1}{r}k'}) \text{ for } (k, k') \in \mathbb{N}^2\}$. In these expressions, t_0 naturally denotes the lowest time index of the representation and ν_0 the lowest frequency bin.

Let us make explicit here that in the following derivations we will be using the notation $\overline{z}$ for the complex conjugate of z and the symbol * for the discrete convolution operator.

2.2. Time–frequency representations as mass functions

Reassignment techniques rely on the idea that time–frequency representations, at a given point (t, ν), can be written as the sum of a mass function defined on the time–frequency plane (n, ξ). Given our context, which involves a discrete time axis and a periodic frequency axis, we have $(n, \xi) \in \mathbb{Z} \times (\mathbb{R}/\mathbb{Z})$. We thus look for an expression of the form:

$$s(t, \nu) = \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \Phi^{t,\nu}(n, \xi) d\xi \tag{3}$$

where the function $\Phi^{t,\nu}$ is real-valued.

Let us define W, a discrete version of Rihaczek's ambiguity function [31], for any sequence $\varphi \in \ell^1(\mathbb{Z})$, time index $n \in \mathbb{Z}$ and frequency $\xi \in \mathbb{R}/\mathbb{Z}$ by:

$$W_\varphi(n, \xi) = \sum_{\tau\in\mathbb{Z}} \varphi_{n+\tau} \overline{\varphi_n} \mathrm{e}^{-j2\pi\xi\tau}. \tag{4}$$

Proposition 1. *The time–frequency representation $s(t, \nu)$ of a discrete time signal $x \in \ell^1(\mathbb{Z})$, as defined in Eq. (2), can be written as the sum of a mass function, as in Eq. (3), with:*

$$\Phi^{t,\nu}(n, \xi) = \Re\{W_{h^{t,\nu}}(n, \nu - \xi) W_x(n, \xi)\}.$$

Putting things together, this means that the time–frequency representation of x at point (t, ν) can be written in the following form:

$$s(t, \nu) = \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \Re\{W_{h^{t,\nu}}(n, \nu - \xi) W_x(n, \xi),\} d\xi \tag{5}$$

Proof. Let us evaluate the following expression:

$$E(t, \nu) = \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} W_{h^{t,\nu}}(n, \nu - \xi) W_x(n, \xi) d\xi.$$

We have:

$$E(t, \nu) = \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \left\{ \sum_{\tau_1\in\mathbb{Z}} h_{n+\tau_1}^{t,\nu} \overline{h_n^{t,\nu}} \mathrm{e}^{-j2\pi(\nu-\xi)\tau_1} \sum_{\tau_2\in\mathbb{Z}} x_{n+\tau_2} \overline{x_n} \mathrm{e}^{-j2\pi\xi\tau_2} \right\} d\xi.$$

Knowing that $h^{t,\nu}$ is of finite support and that x is in $\ell^1(\mathbb{Z})$, Fubini's theorem ensures that the summations can be permuted:

$$E(t, \nu) = \sum_{n,\tau_1,\tau_2\in\mathbb{Z}} \left\{ h_{n+\tau_1}^{t,\nu} \overline{h_n^{t,\nu}} x_{n+\tau_2} \overline{x_n} \mathrm{e}^{-j2\pi\nu\tau_1} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \mathrm{e}^{-j2\pi\xi(\tau_2-\tau_1)} d\xi \right\}.$$

Knowing that, for any integer k, we have:

$$\int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \mathrm{e}^{-j2\pi\xi k} d\xi = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

the above expression can be rewritten with respect to a single shift variable τ:

$$E(t, \nu) = \sum_{n,\tau\in\mathbb{Z}} h_{n+\tau}^{t,\nu} \overline{h_n^{t,\nu}} x_{n+\tau} \overline{x_n} \mathrm{e}^{-j2\pi\nu\tau}.$$

By the substitution $\tau \mapsto m - n$ we get:

$$E(t, \nu) = \sum_{n,m\in\mathbb{Z}} h_m^{t,\nu} \overline{h_n^{t,\nu}} x_m \overline{x_n} \mathrm{e}^{-j2\pi\nu(m-n)} = \left\{ \sum_{m\in\mathbb{Z}} h_m^{t,\nu} x_m \mathrm{e}^{-j2\pi\nu m} \right\} \left\{ \overline{\sum_{n\in\mathbb{Z}} h_n^{t,\nu} x_n \mathrm{e}^{-j2\pi\nu n}} \right\} = |\mathcal{T}^{x,h^{t,\nu}}(t, \nu)|^2.$$

Altogether, we have proved that: