HAL
open science

Reassigned Time-Frequency Representations of Discrete Time Signals and Application to the Constant Q Transform

Sébastien Fenet, Roland Badeau, Gael Richard

To cite this version:

Sébastien Fenet, Roland Badeau, Gael Richard. Reassigned Time-Frequency Representations of Discrete Time Signals and Application to the Constant Q Transform. Signal Processing, 2017, 132, pp.170-176. 10.1016/j.sigpro.2016.10.008 . hal-01370542

HAL Id: hal-01370542

https://hal.science/hal-01370542v1

Submitted on 8 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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.*

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.

Keywords: Reassigned time-frequency representations, reassigned CQT, discrete signals, synchrosqueezing

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 [1, 2, 3], on subspace decomposition [4, 5] or on more general models such as AM/FM models [6, 7]. 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

*Corresponding author, e-mail address: Sebastien.Fenet@Zoho.com

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 [8]. This original approach uses the phase information of the time-frequency representation and remains difficult to use in practice. Later on, Auger and Flandrin [1] 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 [9], radar imaging [10] or audio [11]. It has also led to the development of numerous extensions and adaptations of the original method [12, 13, 14].

Another solution to the problem, named *synchrosqueezing*, has been proposed by Daubechies and Maes in [15]. 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 [16, 17, 18] . 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 [19].

Traditionally, reassignment calculations are carried out in the context of continuous time signals (see [20] 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.

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 Tranform (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 depends 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:

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

It is interesting here to recall the Heisenberg-Gabor limit [21] 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, 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 $\left[0;\frac{1}{2}\right]$ (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: $\{(t_0 + k\Delta_t,\, \nu_0 + k'\Delta_\nu)$ 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 $\{(t_0 + k\Delta_t,\ \nu_0 2^{\frac{1}{r}k'})$ 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 [22], 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 equation (2), can be written as the sum of a mass function, as in equation (3), with:*

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

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

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

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:

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

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:

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

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:

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

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

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

Altogether, we have proved that:

$$|\,\mathcal{T}^{x,h^{t,\nu}}(t,\nu)|^2 = \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} W_{h^{t,\nu}}(n,\nu-\xi)\, W_x(n,\xi)\, d\xi \ . \tag{6}$$

By applying the real part operator on both sides of the above equality, using its linearity and knowing that $|\,\mathcal{T}^{x,h^{t,\nu}}(t,\nu)|$ is real, we get:

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

□

We have thus proved that the time-frequency representation of x at a given time t and frequency ν, which we have defined as the squared magnitude of the transform, can be expressed as the sum of a real-valued function in the time-frequency plane. In accordance with the context of discrete time signals, we have a discrete time axis (leading to a discrete summation over the time index) and a continuous and periodic frequency axis (leading to a continuous and finite summation over the frequency variable). It is also interesting to note that the contributions of the window $h^{t,\nu}$ and the signal x are quite separated and symmetric. They are induced through the same transform W that applies separately on the window and on the signal. We may now consider the function that is summed as a mass function and, as such, look for its centre of mass in the time-frequency plane.

2.3. Reassigned time index

Given the formulation (3) of the time-frequency representation at point (t,ν) as the two-dimensional sum of a mass function, it is possible to evaluate the centre of mass of this distribution in the time-frequency plane. By definition, the first coordinate of the centre of mass is given by the following formula:

$$\hat{t}^{t,\nu} = \frac{\sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} n\Phi^{t,\nu}(n,\xi)d\xi}{\sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \Phi^{t,\nu}(n,\xi)d\xi} \quad . \tag{7}$$

Proposition 2. *The temporal coordinate of the centre of mass in the time-frequency plane of the time-frequency representation evaluated at point (t,ν), defined in equation (7), can also be expressed as:*

$$\hat{t}^{t,\nu} = t + \Re\left\{ \frac{\mathcal{T}^{x,\,r^t h^{t,\nu}}(t,\nu)}{\mathcal{T}^{x,h^{t,\nu}}(t,\nu)} \right\} \tag{8}$$

where r^t is the unit ramp centred on t: $r^t_n = n - t$.

PROOF. Let us evaluate the following quantity:

$$E'(t,\nu) = \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} n\Phi^{t,\nu}(n,\xi)d\xi \ . \tag{9}$$

We have:

$$E'(t,\nu) = \Re\left\{ \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \sum_{\tau_1\in\mathbb{Z}} n h_{n+\tau_1}^{t,\nu} \overline{h_n^{t,\nu}} \, \mathrm{e}^{-j2\pi(\nu-\xi)\tau_1} \sum_{\tau_2\in\mathbb{Z}} x_{n+\tau_2} \, \overline{x_n} \, \mathrm{e}^{-j2\pi\xi\tau_2} \, d\xi \right\}.$$

By using the same grouping and the same substitution as in the proof of Proposition 1, the expression becomes:

$$E'(t,\nu) = \Re\left\{ \sum_{n,m\in\mathbb{Z}} n \, h_m^{t,\nu} \, \overline{h_n^{t,\nu}} \, x_m \, \overline{x_n} \, \mathrm{e}^{-j2\pi\nu(m-n)} \right\} \ .$$

Expanding n as $t+n-t$, we have:

$$\begin{aligned} E'(t,\nu) &= t\,\Re\left\{ \overline{\sum_{n\in\mathbb{Z}} h_n^{t,\nu} \, x_n \, \mathrm{e}^{-j2\pi\nu n}} \sum_{m\in\mathbb{Z}} h_m^{t,\nu} \, x_m \, \mathrm{e}^{-j2\pi\nu m} \right\} \\ &\quad + \Re\left\{ \overline{\sum_{n\in\mathbb{Z}} (n-t) \, h_n^{t,\nu} \, x_n \, \mathrm{e}^{-j2\pi\nu n}} \sum_{m\in\mathbb{Z}} h_m^{t,\nu} \, x_m \, \mathrm{e}^{-j2\pi\nu m} \right\} \end{aligned}$$

This can be more concisely written as:

$$E'(t,\nu) = t \, |\, \mathcal{T}^{x,h^{t,\nu}}(t,\nu)|^2 + \Re\left\{ \overline{\mathcal{T}^{x,r^t h^{t,\nu}}(t,\nu)} \ \mathcal{T}^{x,h^{t,\nu}}(t,\nu) \right\} \tag{10}$$

Given the definitions (7) and (9), we know that dividing $E'(t,\nu)$ by $\sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \Phi^{t,\nu}(n,\xi)d\xi$ yields $\hat{t}^{t,\nu}$. Besides, we know from equations (2) and (3) that:

$$\begin{aligned} \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \Phi^{t,\nu}(n,\xi)d\xi &= |\, \mathcal{T}^{x,h^{t,\nu}}(t,\nu)|^2 \\ &= \mathcal{T}^{x,h^{t,\nu}}(t,\nu) \, \overline{\mathcal{T}^{x,h^{t,\nu}}(t,\nu)} \ . \end{aligned} \tag{11}$$

Thus, dividing equation (10) by $\sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \Phi^{t,\nu}(n,\xi)d\xi$, we finally get:

$$\hat{t}^{t,\nu} = t + \Re\left\{ \frac{\mathcal{T}^{x,\,r^t h^{t,\nu}}(t,\nu)}{\mathcal{T}^{x,h^{t,\nu}}(t,\nu)} \right\}$$

□

Let us recall that we are currently looking for the centre of mass of the mass function $\Phi^{t,\nu}$ from which the time-frequency representation $s(t,\nu)$ stems. What we have done here is establishing a simple expression of the time coordinate $\hat{t}^{t,\nu}$ of the centre of mass. The original definition of the centre of mass, shown in equation (7), indeed leads to quite complex computations while our calculations have showed that the time coordinate of the centre of mass can be expressed mostly as the division of two transforms. The benefits of this expression will be further discussed in section 3 but it is already remarkable that we could express this coordinate in a nice and simple way.

2.4. Reassigned frequency

Similarly to what has been done for the time coordinate, the expression of the time-frequency representation at point (t,ν) as the sum of a mass function $\Phi^{t,\nu}$, proved in Proposition 1, allows to determine the frequency coordinate of the centre of mass of $\Phi^{t,\nu}$. The latter is defined by:

$$\hat{\nu}^{t,\nu} = \frac{\sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \xi \Phi^{t,\nu}(n,\xi) d\xi}{\sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \Phi^{t,\nu}(n,\xi) d\xi} \ . \tag{12}$$

Proposition 3. *The frequency coordinate of the centre of mass in the time-frequency plane of the time-frequency representation evaluated at point (t,ν), defined in equation (12), can also be expressed as:*

$$\hat{\nu}^{t,\nu} = \nu - \frac{1}{2\pi} \Im \left\{ \frac{\mathcal{T}^{x,h^{t,\nu}*g}(t,\nu)}{\mathcal{T}^{x,h^{t,\nu}}(t,\nu)} \right\} \tag{13}$$

where g is the differentiator filter, i.e. $g_n = \frac{(-1)^n}{n}$ for $n \neq 0$ and $g_0 = 0$.

Proof. Let us evaluate the following expression:

$$E''(t,\nu) = \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \xi \Phi^{t,\nu}(n,\xi) d\xi \ . \tag{14}$$

We have:

$$E''(t,\nu) = \Re\left\{ \sum_{n\in\mathbb{Z}} \int_{\xi=\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \xi \sum_{\tau_1\in\mathbb{Z}} h_{n+\tau_1}^{t,\nu}\, \overline{h_n^{t,\nu}}\, \mathrm{e}^{-j2\pi(\nu-\xi)\tau_1} \sum_{\tau_2\in\mathbb{Z}} x_{n+\tau_2}\, \overline{x_n}\, \mathrm{e}^{-j2\pi\xi\tau_2}\, d\xi \right\}.$$

By the substitution $\xi \mapsto \xi' + \nu$ we get:

$$\begin{aligned} E''(t,\nu) &= \nu\, |\, \mathcal{T}^{x,h^{t,\nu}}(t,\nu)|^2 \\ &\quad + \underbrace{\Re\left\{ \sum_{n\in\mathbb{Z}} \int_{\xi'=-\frac{1}{2}}^{\frac{1}{2}} \xi' \sum_{\tau_1\in\mathbb{Z}} h_{n+\tau_1}^{t,\nu}\, \overline{h_n^{t,\nu}}\, \mathrm{e}^{j2\pi\xi'\tau_1} \sum_{\tau_2\in\mathbb{Z}} x_{n+\tau_2}\, \overline{x_n}\, \mathrm{e}^{-j2\pi(\xi'+\nu)\tau_2}\, d\xi' \right\}}_{E_2''(t,\nu)}. \end{aligned}$$

In the above expression, let us call the second operand $E_2''(t,\nu)$. We now have:

$$E''(t,\nu) = \nu\, |\, \mathcal{T}^{x,h^{t,\nu}}(t,\nu)|^2 + E_2''(t,\nu)\ . \tag{15}$$

Adequately reordering the summations in $E_2''(t,\nu)$, the latter can be written as:

$$E_2''(t,\nu) = \Re\left\{ \sum_{n,\tau_1,\tau_2\in\mathbb{Z}} h_{n+\tau_1}^{t,\nu}\, \overline{h_n^{t,\nu}}\, x_{n+\tau_2}\, \overline{x_n}\, \mathrm{e}^{-j2\pi\nu\tau_2} \int_{\xi'=-\frac{1}{2}}^{\frac{1}{2}} \xi'\, \mathrm{e}^{j2\pi\xi'(\tau_1-\tau_2)}\, d\xi' \right\}.$$

Let us define the sequence g in the following way:

$$g_n = \int_{\xi'=-\frac{1}{2}}^{\frac{1}{2}} 2j\pi\xi'\, \mathrm{e}^{j2\pi\xi' n}\, d\xi' = \begin{cases} \frac{(-1)^n}{n} & \text{if } n \neq 0 \\ 0 & \text{if } n = 0 \end{cases}\ . \tag{16}$$

The expression of $E_2''(t,\nu)$ consequently becomes:

$$E_2''(t,\nu) = \Re\left\{ \frac{j}{2\pi} \sum_{n,\tau_1,\tau_2\in\mathbb{Z}} h_{n+\tau_1}^{t,\nu}\, \overline{h_n^{t,\nu}}\, x_{n+\tau_2}\, \overline{x_n}\, \mathrm{e}^{-j2\pi\nu\tau_2}\, g_{\tau_2-\tau_1} \right\}.$$

We can then use the substitution:

$$\begin{cases} \tau_2 & \mapsto & l-n \\ \tau_1 & \mapsto & l-n-m \end{cases}$$

to get the following equation.

$$E_2''(t,\nu) = \Re\left\{ \frac{j}{2\pi} \sum_{n,m,l\in\mathbb{Z}} h_{l-m}^{t,\nu}\, \overline{h_n^{t,\nu}}\, x_l\, \overline{x_n}\, \mathrm{e}^{-j2\pi\nu(l-n)}\, g_m \right\}.$$

As far as the terms in m are concerned, we observe that:

$$\sum_{m\in\mathbb{Z}} h^{t,\nu}_{l-m} g_m = (h^{t,\nu} * g)_l \ .$$

The whole expression can thus be rewritten:

$$E''_2(t,\nu) = \Re\left\{ \frac{j}{2\pi} \sum_{n\in\mathbb{Z}} \overline{h^{t,\nu}_n \, x_n \, \mathrm{e}^{-j2\pi\nu n}} \sum_{l\in\mathbb{Z}} (h^{t,\nu} * g)_l \, x_l \, \mathrm{e}^{-j2\pi\nu l} \right\} .$$

By identifying the transforms in the above expression of $E''_2(t,\nu)$ and reinjecting in equation (15), we get:

$$E''(t,\nu) = \nu |\, \mathcal{T}^{x,h^{t,\nu}}(t,\nu)|^2 - \frac{1}{2\pi} \Im\left\{ \overline{\mathcal{T}^{x,h^{t,\nu}}(t,\nu)} \, \mathcal{T}^{x,h^{t,\nu}*g}(t,\nu) \right\} .$$

Taking into account equation (11) and given the definition (14) of $E''(t,\nu)$, it is only needed to divide the latter by $|\,\mathcal{T}^{x,h^{t,\nu}}(t,\nu)|^2$ to obtain $\hat{\nu}^{t,\nu}$, as defined in equation(12). This finally leads to:

$$\hat{\nu}^{t,\nu} = \nu - \frac{1}{2\pi} \Im \left\{ \frac{\mathcal{T}^{x,h^{t,\nu}*g}(t,\nu)}{\mathcal{T}^{x,h^{t,\nu}}(t,\nu)} \right\}$$

□

Similarly to what we have done in Section 2.3 for the time coordinate, we have obtained here a simple expression of the frequency coordinate of the centre of mass of $\Phi^{t,\nu}(n,\xi)$. Although the mathematical steps that led to this expression are rather different from the ones performed in the case of the first coordinate, the final result in equation (13) is quite symmetric with the one in equation (8). In a similar fashion, the frequency coordinate is expressed as a simple combination of two transforms. This expression of $\hat{\nu}^{t,\nu}$ will be further discussed in the next section.

3. Discussion

In the previous section, we have obtained exact expressions of the coordinates $(\hat{t}^{t,\nu}, \hat{\nu}^{t,\nu})$ of the centre of mass of $\Phi^{t,\nu}$. The latter is a mass function

whose summation in the time frequency plane gives $s(t,\nu)$ *i.e.* the value of the time-frequency representation of x at time t and frequency ν. The obtained expressions are valid for any time-frequency representation of a discrete signal with any kind of windowing. Time-frequency reassignment consists of building an enhanced time-frequency representation of x simply by reassigning the value $s(t,\nu)$ to the time-frequency point $(\hat{t}^{t,\nu}, \hat{\nu}^{t,\nu})$.

3.1. Practical interest of the expressions

The major advantage of the presented formulae for $\hat{t}^{t,\nu}$ and $\hat{\nu}^{t,\nu}$ is that they provide a way of computing the reassigned index with the computation of only three transforms, one of which is the 'original' one ($\mathcal{T}^{x,h^{t,\nu}}(t,\nu)$). The presented reassignment tool is thus very powerful since it only involves an increase of the computation time by a factor 3 compared to the original representation, which does not seem a very high price to pay considering the tremendous gain of resolution.

In detail, in order to compute the reassigned time index as in equation (8), one has to compute an additional transform of x with a window function which consists of the original window function multiplied by the unit ramp function. The real part of the division of this ramp-windowed transform by the original transform provides the time shift. In order to compute the reassigned frequency index as in equation (13), one has to compute the transform of x with a window sequence which is the original window convolved with some discrete filter g. Let us note that this modified window can be pre-computed once so that calculating the transform of x with this window has the same complexity as with the original window. The frequency shift is proportional to the imaginary part of the transform with the convolved window divided by the original transform.

3.2. Comparison with previously established expressions

The transition from the mathematical proof of the time-frequency reassignment formulae in the context of continuous time signals [1] to the one for discrete time signals is not straightforward. The continuous proof indeed relies on the

rewriting of the spectrogram as a summed product of Wigner-Ville transforms [23, 24, 25]. However, since the Wigner-Ville transform involves non-integer shifts, it cannot be directly applied to discrete-time signals. We have overcome the issue by introducing in equation (4) the transform W, which can be seen as a discrete version of Rihaczek's ambiguity function [22]. This transform is fully applicable to discrete time signals but it lacks the symmetric design of the Wigner-Ville transform. As a result, the transform is not bounded to real values. Fortunately, we could still step back to the expression of a real-valued mass function by applying the real operator on equation (6). Another pitfall when switching to discrete time signals is the choice of ξ's integration interval in equation (3). Given the spectral periodicity of discrete signals, it is clear enough that the interval should be of unit length. However, the localisation of the interval is not obvious. It does not intervene in the calculation of the reassigned time index but it has a meaningful impact when deriving the reassigned frequency index. Our calculations have shown that, in a rather expectable and logical way, the interval of integration should be centred around ν when studying the time-frequency representation at point (t, ν).

In spite of the meaningfully dissimilar calculations steps that led to them, the formulae that we finally obtain for $\hat{t}^{t,\nu}$ and $\hat{\nu}^{t,\nu}$ are very similar to the ones obtained in the continuous case by Auger and Flandrin [1]. The expression of the reassigned time index is indeed the same, with the same appearance in the calculation of the window multiplied by the ramp function. One meaningful contribution of our work, though, is that our expressions have been derived in the context of discrete time signals. There is thus no approximation when applying it to real signals and we have formally proved that the expression is valid. As far as the reassigned frequency index is concerned, the continuous proof, which relies on an integration by parts, leads to an expression that brings the derivative of the original window into play. On the other hand, our calculations lead to a new window which is the original one convolved with a filter g. For that matter, it is very interesting to note that the filter g that we obtain is the ideal differentiator filter. It is indeed clear in the definition of g, given in equation

(16), that it is the inverse Fourier Transform of $2j\pi\xi'$ which is the frequency response of the perfect differentiator. The similarity between our expression and the one from Auger and Flandrin then becomes clear. To some extent, we may indeed consider that the "derivative" of h approximates the convolution of h by g. We however take full advantage here of our discrete time approach. We indeed have an exact formulation for the reassigned frequency index whereas Auger and Flandrin propose to compute the "derivative" of h. The derivative is quite an ill-defined concept for discrete time signals and applying their formula raises the question of the best way to estimate it. Our work brings an answer to this, until now, open question. The technically correct way of using Auger and Flandrin formula is to estimate the derivative with an ideal differentiator filtering.

3.3. From reassignment to synchrosqueezing

As stated in the introduction the pioneering approach of Kodera [8] suggests to compute the reassigned time and frequency indexes by using the partial derivatives of the phase of the time-frequency representation. In a similar spirit, the synchrosqueezing method [15] makes use of the partial derivatives of the phase to compute the reassigned frequency index. A notable difference of synchrosqueezing, however, is that only the frequency index is reassigned.

Our method, on the other hand, looks for the centre of mass of the distribution $\Phi^{t,\nu}$ related to the time-frequency representation $s(t,\nu)$. The equivalence between the computation of the reassigned indexes thanks to the partial derivatives of the phase and the computation of the same indexes through the centre of mass has however been proved in [1].

As a result, as explained in [19], the synchrosqueezing method can be computed within our framework in a straightforward way. First, the reassigned frequency $\hat{\nu}^{t,\nu}$ is determined according to equation (13). Finally, the *complex value* $\mathcal{T}^{x,h^{t,\nu}}(t,\nu)$ is additively reassigned to the point $(t,\hat{\nu}^{t,\nu})$. The differences with the reassignment method that we detailed in the first place are thus clear

and limited: only the frequency index is reassigned and the reassignment is operated on the complex value of the transform instead of its squared magnitude.

4. Application: A reassigned Constant-Q Transform

As explained, the formal framework derived above is valid for a wide variety of time-frequency transforms. To illustrate its applicability, we propose to demonstrate the merit of the reassignment strategy on the Constant-Q Transform. The interested reader will find a Python implementation of this reassigned Constant-Q Transform on the authors' homepages[1].

4.1. Presentation of the transform

The Constant-Q Transform is a common tool in the field of audio processing. Its numerous fields of application include main melody extraction [26], audio-fingerprinting [27], chords detection [28, 29]. It was first proposed in [30] by Brown, who aimed at designing a time-frequency representation that mimics our perception of sound. To this end, the Constant-Q Transform has frequency bins with geometrically spaced centre frequencies. Moreover, their frequency resolution is inversely proportional to the centre frequency. Its direct calculation, however, is computationally heavy compared to the Short Time Fourier Transform which can rely on the Fast Fourier Transform. Hence the proposition by Brown and Puckette of a faster implementation in [31] and a follow-up in [32] by Klapuri. Another pitfall of the initial transform is its non-invertibility. As a consequence, several authors have focused on designing an invertible version of the Constant-Q Transform [33, 34, 35, 36]. Very recently, the concept of reassigned Constant-Q Transform has appeared in [37] and [12]. The point of view is however slightly different from the one adopted here since both works present the Constant-Q Transform as an application example of the theory derived in [12]. In the latter, reassignment formulae are specifically derived for filter banks.

[1] googledrive.com/host/0B1DUW3X2T63neXhKWmxrTW5nUlE/ressources.html

Since they rely on the work by Auger and Flandrin [1], the derivations are only carried out for continuous-time signals.

More precisely, in the context of the Constant-Q spectrogram, one evaluates the amount of energy in x at any time t for a set of frequencies $\{\nu_k\}$. The ν_k's are distributed according to the following law:

$$\nu_k = 2^{\frac{k}{r}} \nu_0$$

The parameter r specifies the resolution of the transform (the bigger r is, the more numerous the frequency bins are) and ν_0 is the centre of the lowest frequency bin of the transform. A meaningful interest of the Constant-Q Transform is that the ν_k's can be aligned on the note frequencies of the Western scale. The parameter r then corresponds to the number of bins per octave, or equivalently, $r/12$ to the number of bins per note.

As far as the windowing functions are concerned, the Constant-Q Transform scales them with respect to frequency:

$$h_n^{t,\nu} = \nu h(\nu(n-t))$$

where h is a window function. The above relationship implies that the temporal support of the window decreases as ν_k increases. This equivalently means that the temporal resolution of the transform increases with frequency, and, according to the Heisenberg-Gabor principle [21], that its frequency resolution decreases.

The results presented thereafter have been obtained with the following set of parameters:

$$\begin{cases} r = 36 \\ \nu_0 = 2.45 \times 10^{-3} \end{cases}$$

and h^{t,ν_0} is a Hann window whose length is about 22500 samples. The analysed sound has a sampling rate of 44100 Hz, which means that the lowest frequency bin of the transform corresponds to the continuous frequency 107.9 Hz and that the support of h^{t,ν_0} is half a second long. With this setup the frequency bins of the transform are aligned on the note frequencies of the Western scale,

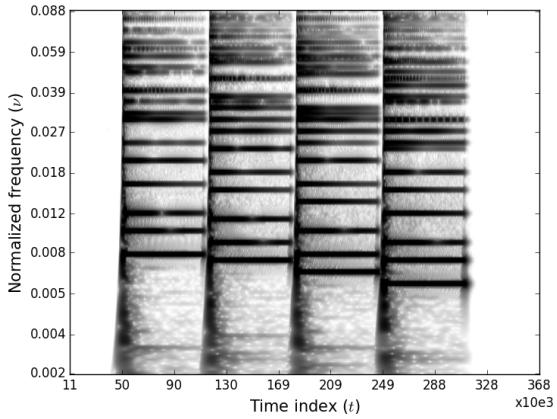

Figure 1: Spectrogram of a piano sound, sampled at 44100 Hz, obtained by means of the traditional Constant-Q Transform. The X axis is the time, it is labelled with the integer indexes of the sampled signal. The Y axis represents the frequency. It is noticeable that the latter follows a logarithmic scale.

with a resolution of 3 bins per note. More specifically, the three lowest bins of the transform are centre on 110 Hz, which corresponds to the frequency of the musical note A_2.

4.2. Considerations on the resolution

Figure 1 shows the spectrogram of a piano sound obtained by means of the traditional Constant-Q Transform. We can see that the frequency resolution of the transform is well adapted to the musical sound since every frequency of the sound is well resolved. As mentioned, this transform has a better frequency resolution in low frequencies, where the musical frequencies are closer to each other. Of course, the counterpart of this better frequency resolution is a loss in temporal resolution. As a result, in spite of the fact that all frequencies of the piano sound have been played simultaneously, one can observe in the spectrogram that the energy in low frequencies starts earlier than in the higher frequencies. Similarly, the energy in low frequencies vanishes later. This is a typical artifact of the Constant-Q Transform due to these resolution considerations.

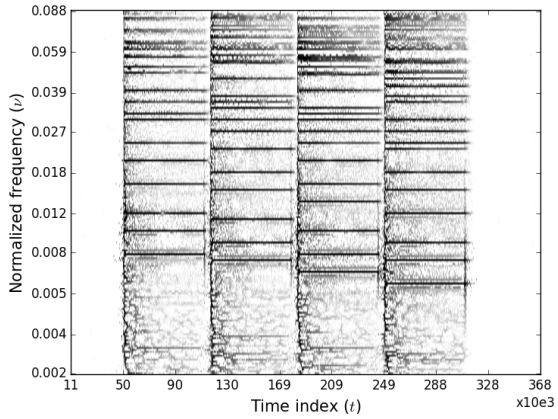

Figure 2: Spectrogram of the same piano sound as in Figure 1 obtained by means of the reassigned Constant-Q Transform.

4.3. Application of the reassignment tool

It is quite straightforward to apply the reassignment formulae to the Constant-Q Transform. When evaluating the energy in x at time t and frequency ν_k, the process is the following. We first compute the magnitude of the Constant-Q Transform at (t, ν_k). We then compute the Constant-Q Transform with the ramp-multiplied windows. We finally compute the Constant-Q Transform with the g-convolved windows. Thanks to equations (8) and (13), we obtain the indexes $\hat{t}^{t,\nu_k}$ and $\hat{\nu}^{t,\nu_k}$. In the end, the energy computed at point (t, ν_k) is additively reassigned to the point $(\hat{t}^{t,\nu_k}, \hat{\nu}^{t,\nu_k})$ in the time-frequency plane.

Figure 2 shows the reassigned Constant-Q Transform of the same piano sound as in Figure 1. One can see that the temporal resolution has much improved in the second figure. The typical artifacts of the widening energy bursts in low frequencies seem to be accurately handled in the reassigned spectrogram. Besides, one can see that the frequency resolution has also improved. Even if the original tool's design ensures a suitable frequency resolution for sound applications, we can observe that the reassigned transform has much more localised (in terms of frequency) energy bursts.

5. Conclusion

In this work we have derived a formal framework for the time-frequency reassignment of discrete time signals. The principle of time-frequency reassignment is to express the time-frequency transform of interest as the summation of a real-valued mass function in the time-frequency plane. The energy is then reassigned to the centre of mass of this function. The first derivations carried out in this paper result in an explicit expression of the mass function which is valid for a wide variety of time-frequency transforms. Based upon the latter, an expression of the coordinates of the centre of mass could be obtained. Interestingly, it consists of very simple combinations of three time-frequency transforms, each of them using a specific window. This result makes time-frequency reassignment a rather generic, simple and easily computable technique.

To illustrate the applicability of the derived framework as well as the benefits of time-frequency reassignment, the obtained results have then been applied to the Constant-Q Transform. This transform, commonly used in audio processing, has geometrically-spaced frequency bins with a variable resolution. This makes it a well adapted tool for the analysis of sound. This frequency setup, however, comes at the cost of time resolution, which has the serious drawback of being variable with frequency. Once reassigned though, the representation provides a much better resolution, in time and frequency. A Python implementation of this enhanced tool is available on the authors' homepages[2]. It is our belief that the proposed reassigned Constant-Q Transform has a strong potential for numerous audio processing applications, including pitch/multipich estimation, onset detection and chords transcription.

References

[1] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representation by the reassignment method," *IEEE Transactions*

[2] googledrive.com/host/0B1DUW3X2T63neXhKWmxrTW5nUlE/ressources.html

on Signal Processing, vol. 43, no. 5, pp. 1068–1088, May 1995.

[2] M. Abe and J. O. Smith III, "Design criteria for simple sinusoidal parameter estimation based on quadratic interpolation of FFT magnitude peaks," in *Audio Engineering Society 117th Convention*, San Francisco, CA, Oct. 2004.

[3] M. Betser, P. Collen, B. David, and G. Richard, "Review and discussion on STFT-based frequency estimation methods," in *Audio Engineering Society 120th Convention*, Paris, France, May 2006.

[4] R. Roy, A. Paulraj, and T. Kailath, "Esprit - a subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 5, pp. 1340–1342, Oct 1986.

[5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar 1986.

[6] M. Abe and J. O. Smith III, "AM/FM rate estimation for time-varying sinusoidal modeling," in *International Conference on Acoustics, Speech and Signal Processing*, vol. 3, Philadelphia, PA, Mar. 2005, pp. 201–204.

[7] M. Betser, P. Collen, G. Richard, and B. David, "Estimation of frequency for AM/FM models using the phase vocoder framework," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 505 – 517, Feb. 2008.

[8] K. Kodera, R. Gendrin, and C. de Villedary, "Analysis of time-varying signals with small BT values," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 1, pp. 64–76, Feb 1978.

[9] M. Niethammer, L. J. Jacobs, J. Qu, and J. Jarzynski, "Time-frequency representations of lamb waves," *Journal of the Acoustical Society of America*, vol. 109, no. 5, pp. 1841–1847, 2001.

[10] S. S. Ram, Y. Li, A. Lin, and H. Ling, "Doppler-based detection and tracking of humans in indoor environments," *Journal of the Franklin Institute*, vol. 345, no. 6, pp. 679 – 699, 2008.

[11] G. Peeters, "Template-based estimation of time-varying tempo," *EURASIP Journal of Applied Signal Processing*, no. 1, p. 158, Jan. 2007.

[12] N. Holighaus, Z. Pruša, and P. L. Søndergaard, "Reassignment and synchrosqueezing for general time–frequency filter banks, subsampling and processing," *Signal Processing*, vol. 125, pp. 1–8, 2016.

[13] F. Auger, E. Chassande-Mottin, and P. Flandrin, "Making Reassignment Adjustable: the Levenberg-Marquardt Approach," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, Mar. 2012, pp. 3889–3892.

[14] T. Oberlin, S. Meignen, and V. Perrier, "Second-order synchrosqueezing transform or invertible reassignment? towards ideal time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1335–1344, 2015.

[15] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," *Wavelets in medicine and biology*, pp. 527–546, 1996.

[16] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243 – 261, 2011.

[17] C. Li and M. Liang, "A generalized synchrosqueezing transform for enhancing signal timefrequency representation," *Signal Processing*, vol. 92, no. 9, pp. 2264 – 2274, 2012.

[18] A. Ahrabian, D. Looney, L. Stanković, and D. P. Mandic, "Synchrosqueezing-based time-frequency analysis of multivariate data," *Signal Processing*, vol. 106, pp. 331 – 341, 2015.

[19] F. Auger, P. Flandrin, Y. T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H. T. Wu, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 32–41, Nov. 2013.

[20] S. Hainsworth and M. Macleod, "Time frequency reassignment: A review and analysis," Cambridge University Engineering Department, Tech. Rep. 459, 2003.

[21] D. Gabor, "Theory of communication," *Journal of Institute of Electrical Engineers*, vol. 93, pp. 429–457, 1946.

[22] A. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 369–374, May 1968.

[23] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, pp. 749–759, Jun 1932.

[24] J. Ville, "Théorie et Applications de la Notion de Signal Analytique," *Câbles et Transmission*, vol. 2A, pp. 66–74, 1948.

[25] T. A. C. M. Classen and W. F. G. Mecklenbrauker, "The Wigner Distribution: A Tool for Time-Frequency Signal Analysis - Part III: Relations with other time-frequency transformations." *Philips Journal of Research*, vol. 35, no. 6, pp. 372–389, 1980.

[26] B. Fuentes, A. Liutkus, R. Badeau, and G. Richard, "Probabilistic model for main melody extraction using constant-Q transform," in *IEEE Conference on Acoustics, Speech, and Signal Processing*, Kyoto, Japan, Mar. 2012, pp. 5357–5360.

[27] S. Fenet, G. Richard, and Y. Grenier, "A scalable audio fingerprint method with robustness to pitch-shifting," in *International Society for Music Information Retrieval*, Miami, USA, Oct. 2011, pp. 121–126.

[28] S. H. Nawab, S. A. Ayyash, and R. Wotiz, "Identification of musical chords using constant-Q spectra," in *IEEE Conference on Acoustics, Speech, and Signal Processing*, vol. 5, Salt Lake City, UT, May 2001, pp. 3373–3376.

[29] C. Harte, M. Sandler, and M. Gasser, "Detecting Harmonic Change in Musical Audio," in *1st ACM Workshop on Audio and Music Computing Multimedia*, Santa Barbara, California, USA, Oct. 2006, pp. 21–26.

[30] J. C. Brown, "Calculation of a constant Q spectral transform," *Journal of the Acoustical Society of America*, vol. 89, no. 1, pp. 425–434, Jan. 1991.

[31] J. C. Brown and M. S. Puckette, "An efficient algorithm for the calculation of a constant Q transform," *Journal of the Acoustical Society of America*, vol. 92, no. 5, pp. 2698–2701, Nov. 1992.

[32] C. Schörkhuber and A. Klapuri, "Constant-Q transform toolbox for music processing," in *7th Sound and Music Computing Conference*, Barcelona, Spain, Jul. 2010.

[33] G. A. Velasco, N. Holighaus, M. Dörfler, and T. Grill, "Constructing an invertible constant-Q transform with non-stationary Gabor frames," in *International Conference on Digital Audio Effects*, Paris, France, 2011, pp. 93–99.

[34] D. Fitzgerald, M. Cranitch, and M. T. Cychowski, "Towards an inverse constant Q transform," in *120th Convention of the Audio Engineering Society*, Paris, France, May 2006.

[35] N. Holighaus, M. Dörfler, G. A. Velasco, and T. Grill, "A Framework for Invertible, Real-Time Constant-Q Transforms," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 4, pp. 775–785, April 2013.

[36] A. Nagathil and R. Martin, "Optimal Signal Reconstruction from a Constant-Q Spectrum," in *IEEE Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, Mar. 2012, pp. 349–352.

[37] Z. Pruša and N. Holighaus, "Real-Time Audio Visualization With Reassigned Non-uniform Filter Bank," in *International Conference on Digital Audio Effects*, Brno, Czech Republic, Sep. 2016.