This full text paper was peer-reviewed at the direction of IEEE Instrumentation and Measurement Society prior to the acceptance and publication.

Observer-based discrete Gabor transform

Bence Ország, László Sujbert
Department of Measurement and Information Systems,
Budapest University of Technology and Economics,
Magyar tudósok krt. 2., 1521 Budapest, Hungary
orszag.bence.laszlo@gmail.com, sujbert@mit.bme.hu

***Abstract*—Time-frequency representations are frequently used in signal processing applications. This paper presents a framework for recursive implementation of biorthogonal nonstationary discrete Gabor transforms. These transforms can achieve a non-uniform frequency resolution unlike the well known Fourier transform. Typically they are realized with finite impulse response filters. This paper shows an observer-based recursive implementation of these transforms based on Hostetter's approach. In addition it reviews the construction of generalized Gabor frames and the conditions of their invertibility in detail. The design of the observer's parameters are discussed and multiple examples are given to illustrate the properties and capabilities of both the design process and the observer.**

***Index Terms*—signal processing, Gabor transform, observer, frames**

I. Introduction

It is common to use time-frequency representations to determine important signal characteristics when strict time or frequency representations are not adequate. The most obvious way to construct such a representation is to separate the time domain signal into multiple (possibly overlapping) shorter segments then calculate their Fourier transforms. This is the essence of the short-time Fourier transform (STFT) also known as Gabor transform.

This idea can be generalized to achieve non-uniform frequency resolutions which might be mandated by the problem specification. For example in audio signal processing a transform with logarithmic resolution might be a better choice over a linear one because it reflects the resolution of the human auditory system better. The tools of frame theory gave rise to the construction of nonstationary Gabor transforms [1]. Near arbitrary frequency resolutions can be achieved with them, moreover they are invertible. Gabor transforms were successfully applied to solve a variety of problems in signal processing. A practical example is audio denoising [2]. Furthermore an invertible constant-Q transform can be constructed as a special case of Gabor-transforms [3] and it can be implemented in real-time applications if the signal is processed in a blockwise manner [4].

Gabor transforms can be interpreted as a filtering operation with finite impulse response (FIR) filters, which can be implemented naively with the help of convolutions. The novelty of our paper is that it presents a real-time, recursive implementation of Gabor transforms with the help of Hostetter's observer theory [5]. The observer recursively estimates the transformed signal as well as the signal itself. The estimation is refined based on the reconstruction error which leads to desirable numerical properties.

In section II the relevant parts of frame theory and the definition of discrete Gabor transforms and it's generalizations are presented. Section III presents the conceptual signal model, it's observer and their properties. Section IV contains the derivation of the main results, the recursive estimation of Gabor transforms with an observer. Section V illustrates the method with examples. Finally section VI concludes the paper.

II. Gabor transform

This section follows [3] and [4] in the presentation of the theory of Gabor transforms.

A. Frames

Frame theory [6] is mainly concerned with signal representations. It is a generalization of the theory of (orthogonal) bases. Frames might be constructed with greater flexibility to satisfy a specification by allowing "redundancy" in the signal representation. Given a set of φ_k atoms ($k = 0, \cdots, K-1$) the goal is to find w_k expansion coefficients for a signal $\mathbf{v}$ such that

$$\mathbf{v} = \sum_k w_k \varphi_k = \mathbf{\Phi w} \tag{1}$$

Frame theory provides conditions which ensure that $\mathbf{v}$ is reconstructible from w_k. In this article we assume that all signals are elements of $\mathbb{C}^N$, i.e. they are N periodic and discrete time, represented by a column vector. The n^{th} entry in the vector is $\mathbf{v}[n]$. In this setting all results can be proven with the tools of linear algebra. We denote the inner product of $\mathbf{u}$ and $\mathbf{v}$ as

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^H \mathbf{u} = \sum_n \mathbf{u}[n]\overline{\mathbf{v}[n]} \tag{2}$$

where the overline means complex conjugation. The atoms can be ordered into a matrix called synthesis operator with dimensions $N \times K$:

$$\mathbf{\Phi} = \begin{pmatrix} \varphi_0 & \varphi_1 & \cdots & \varphi_{K-1} \end{pmatrix} \tag{3}$$

The frame operator can be defined as a Hermitian matrix with dimensions $N \times N$:

$$\mathbf{S} = \mathbf{\Phi \Phi}^H \tag{4}$$

If it is invertible then the set of atoms is a frame. In this case the dual frame elements are defined by

$$\widetilde{\mathbf{\Phi}} = \mathbf{S}^{-1}\mathbf{\Phi} \tag{5}$$

A discrete transformation can be defined with the matrix called analysis operator which provides the expansion coefficients:

$$\mathbf{w} = \widetilde{\mathbf{\Phi}}^H \mathbf{v} \tag{6}$$

The reconstruction is perfect in the sense of (1) because

$$\mathbf{v} = \mathbf{S}\mathbf{S}^{-1}\mathbf{v} = \mathbf{\Phi}\mathbf{\Phi}^H\mathbf{S}^{-1}\mathbf{v} = \mathbf{\Phi}\left(\mathbf{S}^{-1}\mathbf{\Phi}\right)^H\mathbf{v} = \mathbf{\Phi}\widetilde{\mathbf{\Phi}}^H\mathbf{v} \tag{7}$$

When $K = N$, a frame implements a biorthogonal transform.

B. Gabor transform

All frames are constructed from a set of atoms. It is convenient to create these in a structured manner. The set of Gabor atoms can be generated with the help of an initial φ window function, an a translation, and a b modulation parameter.

$$\{\mathbf{M}_{mb}\mathbf{T}_{la}\varphi\}_{(l,m)\in[0,N/a)\times[0,N/b)} \tag{8}$$

where $\mathbf{T}_\tau$ and $\mathbf{M}_\omega$ denote a circular time-shift by τ and a modulation by ω, i.e.

$$\mathbf{T}_\tau\mathbf{v}[n] = \mathbf{v}[n-\tau] \qquad \mathbf{M}_\omega\mathbf{v}[n] = \mathbf{v}[n]\cdot e^{j\frac{2\pi}{N}\omega n} \tag{9}$$

where $n-\tau$ is calculated modulo N.

$\frac{N}{a}, \frac{N}{b} \in \mathbb{Z}$ must be satisfied, in that case $m = 0, \cdots, \frac{N}{b}-1$ and $l = 0, \cdots, \frac{N}{a}-1$ so the total number of atoms is $K = \frac{N^2}{ab}$. The two index variables foreshadow that atoms are localized in time and frequency.

For a particular choice of φ, a and b the atoms can be organised into the matrix $\mathbf{\Phi}$. If the induced frame operator is invertible then the Gabor transform of a $\mathbf{v}$ signal is (6). It is an important result in the theory of Gabor frames that the dual atoms can be generated with $\widetilde{\varphi}$, a and b.

The k^{th} element of the expansion coefficient vector becomes

$$\mathbf{w}[k] = \sum_{n=0}^{N-1}\mathbf{v}[n]\overline{\widetilde{\varphi}[n-la]}e^{-j\frac{2\pi}{N}mb} \tag{10}$$

which is the STFT of the input signal. This means that the a and b parameters adjust the time and frequency resolution of the transform. Note that the value of k depends on l, m and on the order of the atoms in $\mathbf{\Phi}$, for example

$$k = m + \frac{N}{b}l \tag{11}$$

C. Nonstationary Gabor transform

Ordinary Gabor transforms are limited to linear time-frequency resolutions by construction. In the case of nonstationary Gabor transforms (NSGT) this restriction can be lifted. The set of atoms is

$$\{\mathbf{M}_{mb_l}\varphi_l\}_{(l,m)\in[0,L)\times[0,M_l)} \tag{12}$$

where $\{\varphi_l\}$ is a set of windows and $\{b_l\}$ is a set of modulation parameters. The (12) atoms form a frame if their frame operator is invertible. These atoms are generated by the modulation of the φ_l windows by b_l. The number $M_l = \frac{N}{b_l}$ must be an integer. The number of atoms derived from φ_l is M_l which means the total number of atoms in (12) is $K = \sum_l M_l$. For a particular choice of windows and modulation parameters the atoms can be organized into a matrix $\mathbf{\Phi}$. One possibility can be seen below:

$$\mathbf{\Phi} = \begin{pmatrix}\mathbf{M}_{0b_0}\varphi_0 & \mathbf{M}_{1b_0}\varphi_0 & \cdots & \mathbf{M}_{(M_0-1)b_0}\varphi_0 & \mathbf{M}_{0b_1}\varphi_1 & \cdots\end{pmatrix} \tag{13}$$

If the induced frame operator is invertible, then the Gabor transform of a signal can be defined like (6).

$$\mathbf{w}[k] = \sum_{n=0}^{N-1}\mathbf{v}[n]\overline{\widetilde{\varphi}_l[n]}e^{-j\frac{2\pi}{N}mb_l} \tag{14}$$

In this case the frequency resolution remained linear but the time resolution can be nonlinear which is determined by the φ_l windows. This property is desirable when processing unevenly sampled signals. It should be noted that ordinary Gabor transforms are a special case of NSGTs.

D. Painless invertibility

Previously it was assumed that $\mathbf{S}$ is invertible but to successfully construct an NSGT it is necessary to prove it. Gabor transforms can be defined more generally on Hilbert spaces where the invertibility of $\mathbf{S}$ is not enough, in a practical implementation it must be computable in a straightforward way. This was achieved in [7] where the authors could give conditions which ensure that $\mathbf{S}$ is diagonal. In our case it is enough to show that every φ_l has a finite support $N_l \leq N$, and every $b_l \leq \frac{N}{N_l}$. From these it can be asserted that $\mathbf{S}$ is diagonal. We prove this statement in Theorem 1 based on [8] (see appendix).

As a corrolarry it is easy to show that the dual frame of (12) can be generated with $\{\widetilde{\varphi}_l\}$ and $\{b_l\}$, where

$$\widetilde{\varphi}_l[n] = \frac{\varphi_l[n]}{\mathbf{S}^{-1}[n,n]} = \frac{\varphi_l[n]}{\sum_{l=0}^{L-1} M_l\left|\varphi_l[n]\right|^2} \tag{15}$$

E. Discrete Fourier transform

To develop further the theory of generalized Gabor transforms, it is necessary to introduce the discrete Fourier transform and it's properties. The unitary discrete Fourier transform (DFT) can be represented with $\mathcal{F} \in \mathbb{C}^{N\times N}$, where

$$\mathcal{F}[n,m] = \frac{1}{\sqrt{N}}e^{-j\frac{2\pi}{N}nm} \tag{16}$$

The inverse transform is

$$\mathcal{F}^{-1} = \mathcal{F}^H \tag{17}$$

The DFT of a signal can be written briefly as $\dot{\mathbf{v}} = \mathcal{F}\mathbf{v}$. The translation and modulation theorems of the Fourier transform take the following form

$$\mathcal{F}\mathbf{T}_\tau = \mathbf{M}_{-\tau}\mathcal{F} \qquad \mathcal{F}\mathbf{M}_\omega = \mathbf{T}_\omega\mathcal{F} \tag{18}$$

F. Nonlinear frequency resolution

The nonstationary Gabor frames introduced in the previous sections are adaptive in time resolution, but frequency adaptive frames are straightforward to construct based on these. The set

$$\{\mathbf{T}_{ma_l}\varphi_l\}_{(l,m)\in[0,L)\times[0,M_l)} \tag{19}$$

where $M_l = \dfrac{N}{a_l} \in \mathbb{Z}$ is a frame if and only if $\{\mathbf{M}_{-ma_l}\dot{\varphi}_l\}$ is a frame. The frame operators of the set in (19) and $\{\mathbf{M}_{-ma_l}\dot{\varphi}_l\}$ are

$$\mathbf{R} = \mathbf{\Phi}\mathbf{\Phi}^H \qquad \mathbf{S} = \dot{\mathbf{\Phi}}\dot{\mathbf{\Phi}}^H \tag{20}$$

It is easy to prove that $\mathbf{R}$ is similar to $\mathbf{S}$ in the sense of linear algebra. The former is invertible only if the latter is invertible.

$$\mathbf{R}^{-1} = \mathcal{F}^{-1}\mathbf{S}^{-1}\mathcal{F} \tag{21}$$

In this case

$$\mathbf{w}[k] = \sum_{n=0}^{N-1} \mathbf{v}[n]\overline{\widetilde{\varphi}_l[n - ma_l]} \tag{22}$$

The time resolution is linear, while the frequency resolution is given by the set of windows, just as desired.

G. Biorthogonality condition

Frame theory gives a flexible generalization of bases but it can be used to construct biorthogonal transforms as a special case of painless nonstationary Gabor transforms. For them to be biorthogonal it is enough to ensure that $\sum_l M_l = N$ as it was noted in section II-A. Expanding the definition of M_l the equation can be rewritten as

$$\sum_l \frac{1}{a_l} = 1 \tag{23}$$

In practice this is a restricting condition because it is not enough to satisfy the above equation, but for all l a_l and $\dfrac{N}{a_l}$ must be an integer.

H. Filter bank interpretation

Every NSGT designed with nonlinear frequency resolution can be interpreted as a filter bank. It has two main parts. The first is the analysis bank which decomposes the input signal into several – possibly decimated – frequency bands. The other part is the analysis bank which interpolates and sums the components to restore the signal. Arbitrary signal processing can take place between the analysis and synthesis banks, but if it is omitted then the output signal is identical to the delayed input. This is called the perfect reconstruction property and it is ensured by the invertibility of the NSGT.

The analysis bank is realized by L FIR filters with coefficients equal to the complex conjugate of the time flipped dual atoms (see (22), where φ_l is from (19)). The output of the analysis filters can be decimated by a factor given by the time shift parameters. Finally the synthesis filters are also FIR, these are given by φ_l.

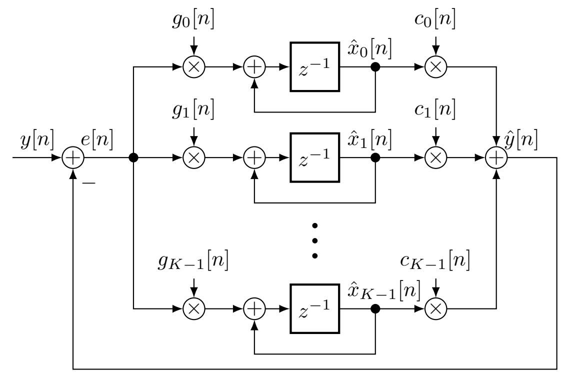

Fig. 1. An observer for recursive transformations.

III. A COMMON STRUCTURE FOR RECURSIVE DISCRETE TRANSFORMS

A. Conceptual signal model

The conceptual signal model [9] is a linear system whose $y[n]$ output is acquired by the summation of $c_k[n]$ atoms weighted by $x_k[n]$ state variables. The $n = 0, \cdots, N-1$ values serve as discrete time instants and $k = 0, \cdots, K-1$ is used to enumerate the atoms. In this setting it is possible to construct all discrete signals with period N with the right choice of atoms. The referenced article assumes that all signals are discrete and have period N and the conceptual signal model has constant state variables. Let's denote

$$\mathbf{x}[n] = \begin{pmatrix} x_0[n] & x_1[n] & \cdots & x_{K-1}[n] \end{pmatrix}^T \tag{24}$$

$$\mathbf{c}[n] = \begin{pmatrix} c_0[n] & c_1[n] & \cdots & c_{K-1}[n] \end{pmatrix}^T \tag{25}$$

This means that the output of the system is

$$y[n] = \mathbf{c}^T[n]\mathbf{x}[n] \tag{26}$$

B. Observer

The observer which was introduced and analyzed in [9] can be seen in Fig. 1. It tries to reconstruct the $y[n]$ input signal by refining the $\hat{\mathbf{x}}[n]$ estimated state variables based on the reconstruction error with the help of the $g_k[n]$ dual atoms. The latter is denoted by

$$\mathbf{g}[n] = \begin{pmatrix} g_0[n] & g_1[n] & \cdots & g_{K-1}[n] \end{pmatrix}^T \tag{27}$$

With the notation introduced so far, the time course of the estimated state variables can be given by the following equation:

$$\hat{\mathbf{x}}[n+1] = \hat{\mathbf{x}}[n] + \mathbf{g}[n]\mathbf{c}^T[n](\mathbf{x}[n] - \hat{\mathbf{x}}[n]) \tag{28}$$

It was proven in [9] that $\hat{\mathbf{x}}[n] = \mathbf{x}[0]$ after N time steps (or less) if $\mathbf{g}[n]$ and $\mathbf{c}[n]$ form a biorthogonal basis.

IV. Observer-based Gabor transform

A. Estimation with observer

Section III-A introduced the atoms of a conceptual signal model. The output samples of the system can be organized into a vector in $\mathbb{C}^N$ if the state variables are constant, $\mathbf{y} = \mathbf{C}\mathbf{x}[0]$.

$$\mathbf{C} = \begin{pmatrix} c_0[0] & c_1[0] & \cdots & c_{K-1}[0] \\ c_0[1] & c_1[1] & \cdots & c_{K-1}[1] \\ \vdots & \vdots & \ddots & \\ c_0[N-1] & c_1[N-1] & & c_{K-1}[N-1] \end{pmatrix} \quad (29)$$

Similarly the dual atoms can be arranged into a matrix

$$\mathbf{G} \;=\; \begin{pmatrix} g_0[0] & g_0[1] & \cdots & g_0[N-1] \\ g_1[0] & g_1[1] & \cdots & g_1[N-1] \\ \vdots & \vdots & \ddots & \\ g_{K-1}[0] & g_{K-1}[1] & & g_{K-1}[N-1] \end{pmatrix} \quad (30)$$

By comparing (29) and (3) one can see that both $\mathbf{C}$ and $\mathbf{\Phi}$ are constructed from the atoms as columns. Likewise there is a correspondence between $\mathbf{G}$ and $\widetilde{\mathbf{\Phi}}^H$. Succinctly

$$c_k[n] = \varphi_k[n] \qquad g_k[n] = \overline{\widetilde{\varphi}_k[n]} \quad (31)$$

The two equations above connect biorthogonal Gabor transforms and the conceptual signal models. An observer can be constructed for said transforms which can recursively estimate the state variables.

B. Designing an observer-based Gabor transform

If the NSGT has nonlinear frequency resolution then based on (22) the generator atoms can be interpreted as FIR filters with time reversed impulse responses. The DFT of these impulse responses are the transfer functions sampled in regular intervals. The condition for painless invertibility in this case states that the individual amplitude characteristics of the filters must have bounded support with length N_l, and their sum must be nonzero on all frequencies. It also requires that the a_l time shift parameters satisfy the Nyquist-Shannon sampling theorem. The number of filters is $l = 0, \cdots, L-1$. To achieve biorthogonality (23) needs to hold.

The first step in the design process is to construct the $\dot{\varphi}_l$ windows directly in the frequency domain with center frequencies positioned (possibly irregularly) according to a specification and with $N_l \leq N$ finite support. The second step is to check the invertibility of the frame operator and that the time shift parameters $a_l = \dfrac{N}{N_l}$ are integers. If this is the case then the biorthogonality condition automatically holds, and the windows partition the whole frequency domain into disjoint segments. Finally the

$$c_k[n] = \varphi_l[n - m a_l] \qquad g_k[n] = \overline{(\mathbf{R}^{-1}\varphi_l)[n - m a_l]} \quad (32)$$

atoms of the conceptual signal model can be calculated, where

$$k = m + \sum_{i=0}^{l-1} M_i \quad (33)$$

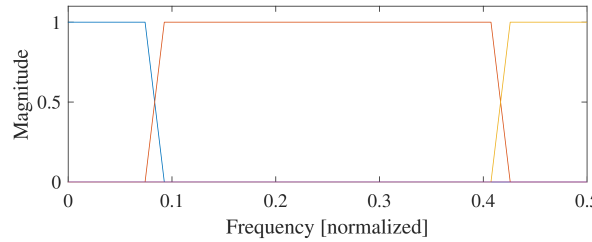

Fig. 2. Amplitude characteristics of the filters used in the three band transform example. The sampling frequency is normalized to 1.

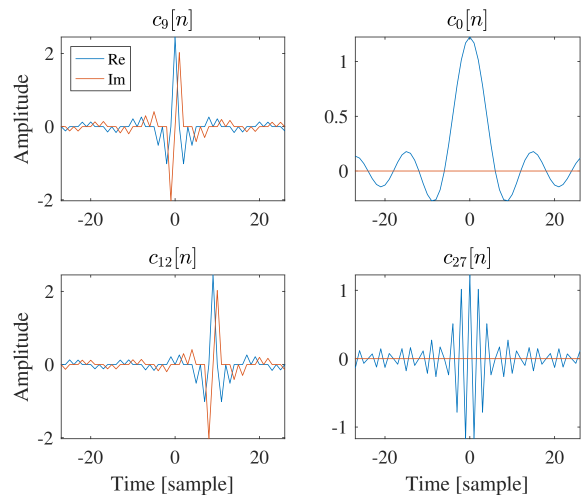

Fig. 3. Some of the $c_k[n]$ atoms corresponding to the three band transform example. The two on the left are the impulse responses of the bandpass filter, they differ only in a time shift. The upper right corresponds to the impulse response of the lowpass filter while the bottom right to the highpass filter.

V. Examples

A. Three band transform

This example shows the usage of the observer-based method on an intentionally simple specification. L, the number of generator atoms – filters – is chosen to be 4. The first is a lowpass filter, while the second and fourth form a complex conjugate pair, and the third one is a highpass filter. The $\dot{\varphi}_l$ atoms were chosen to be rectangular windows. The length of the supports were constructed with biorthogonality in mind, they are computed from the time shift parameters

$$a_0 = 6 \qquad a_1 = 3 \qquad a_2 = 6 \qquad a_3 = 3 \quad (34)$$

To ensure that all $\dfrac{N}{a_l}$ is integer, N was selected to be 54. The amplitude characteristics of the filters can be seen on Fig. 2. The calculation of the basis/reciprocal basis is straightforward based on (32) and (33). Fig. 3. illustrates some atoms of the conceptual signal model, which are the impulse responses and their time shifts. For $k = 0, \cdots, 8$ $c_k[n] = \varphi_0[n - 6m]$, for $k = 9, \cdots, 26$ $c_k[n] = \varphi_1[n - 3m]$, and so on. The real part of the observed signal and the output during a simulation are depicted on Fig. 4 and Fig. 5. The conceptual signal model's state variables are initialized with random real values. The

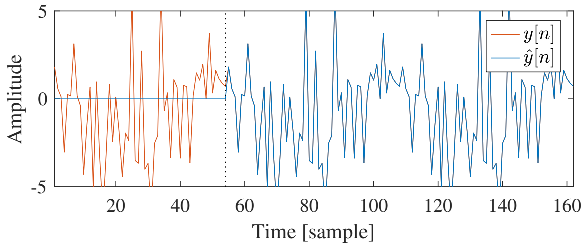

Fig. 4. The real part of the output of the conceptual signal model and the observer in the three band transform example. The state variables are random generated real values with normal distribution. The dotted vertical line marks the N^{th} timestep.

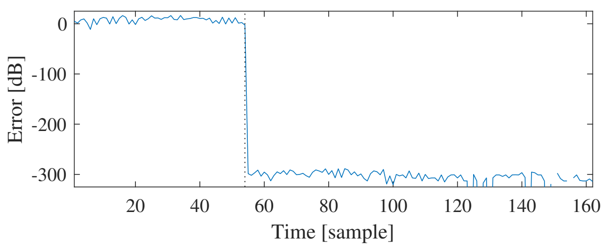

Fig. 5. The reconstruction error in the three band transform example. The dotted vertical line marks the N^{th} timestep.

observer has a delay of N samples, afterwards it reconstructs the signal perfectly.

B. Discrete Fourier transform

To obtain the DFT as a special case, the K number of $\varphi_k[n]$ windows should be equal to N then they can be specified as $\dot{\varphi}_k[n] = \mathbf{e}_k[n]$, where $\mathbf{e}_k$ are the standard basis vectors of $\mathbb{C}^N$. This results in equal support lengths for all l $M_l = 1$. Equation (23) is satisfied with $a_l = \dfrac{N}{M_l} = N$.

The atoms of the conceptual signal model are

$$c_k[n] = (\mathcal{F}^{-1}\dot{\varphi}_k)[n-ma_l] = (\mathcal{F}^{-1}\mathbf{e}_k)[n-mN] = \frac{1}{\sqrt{N}}e^{j\frac{2\pi}{N}kn} \tag{35}$$

where a_l is effectively a constant with value N. Before computing the dual atoms the inverse of the frame operator should be determined. $\mathbf{S} = \dot{\mathbf{\Psi}}\dot{\mathbf{\Psi}}^H = \mathbf{I}$ and $\mathbf{R} = \mathcal{F}^{-1}\mathbf{S}^{-1}\mathcal{F} = \mathbf{I}$ which leads to

$$g_k[n] = \overline{(\mathbf{R}^{-1}\varphi_k)[n-ma_l]} = \overline{\varphi_k[n-mN]} = \frac{1}{\sqrt{N}}e^{-j\frac{2\pi}{N}kn} \tag{36}$$

C. Third-octave filter bank

Filter banks are frequently used in audio signal processing algorithms. Third-octave filter banks are especially useful during the analysis of speech or music. [3] introduces the invertible constant-Q transform (CQT) as a possible implementation of these filter banks. The invertibility ensures the perfect reconstruction of the analyzed signal. This property can be utilized in the development of noise reduction methods.

The signal-to-noise ratio (SNR) of a noisy microphone can be improved with the help of a third-octave filter bank

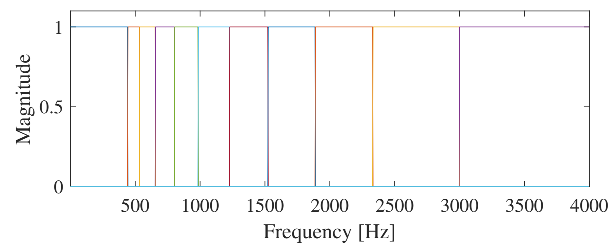

Fig. 6. Amplitude characteristics of the filters used in the constant-Q transform example. The sampling frequency is 8 kHz.

[2]. It works on the assumption that the useful signal is well localized in frequency – which holds reasonably well for speech signals – while the noise is wideband. Furthermore the signal to noise ratio is already positive. Estimating the power of the noisy signal components generated by the filter bank and comparing it to the estimated level of the noise makes it possible to detect and suppress the unwanted signal components with tresholding.

A third-octave CQT can be approximated by a biorthogonal Gabor transform, it's design is illustrated with the following example. The number of filters are $L = 20$. There is a lowpass and highpass filter and 9-9 third-octave bandpass filters with approximately geometrically spaced center frequencies. The sampling frequency is 8 kHz. The cutoff frequency of the lowpass filter is 450 Hz, while the cutoff frequency of the high pass filter is 3000 Hz. The bandpass filters have center frequencies $500 \cdot 2^{k/3}$ Hz, where $k = 0, \cdots, 8$. The bands above half of the sampling frequency have the mirrored amplitude characteristics from the matching ones from below. This is illustrated in Fig. 6. The a_l time shift parameters were determined by an iterative process by hand to minimize their least common multiple while satisfying (23). This leads to a value of 2376 for N.

Fig. 7. shows the response of the observer for a triangle wave input. It may seem that the input is amplitude modulated but it is just a consequence of the non-coherent sampling. It is important to note that several state variables are corresponding to one filter. If one of the harmonic components of the triangle wave is in the passband of the filter then the value of the state variables will be affected accordingly. Based on this knowledge the estimation of the power content of each signal component can be carried out.

D. Comparison of the naive and observer-based method

The implementation of a NSGT with the help of FIR filters is straightforward. But in certain cases it becomes numerically sensitive. Using the filter bank interpretation, when the signal is reconstructed during the synthesis, the sum of the upsampled components might have some uncertainty. This effect is amplified if the calculations are carried out with fixed point representations. In contrast, the observer-based method has significantly lower numerical sensitivity [10] which makes it a suitable choice where floating point representations cannot be used.

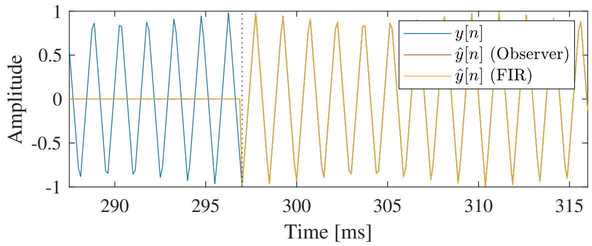

Fig. 7. The outputs of the FIR and observer-based third-octave filter banks. The input signal is a 674 Hz non-coherently sampled triangle wave. The dotted vertical line marks the N^{th} timestep.

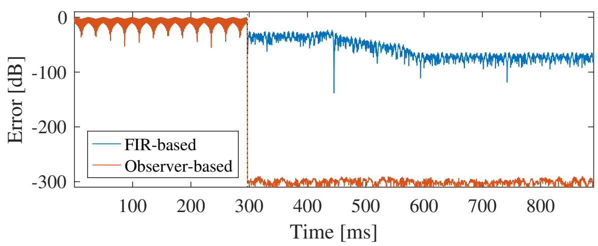

Fig. 8. The comparison of the reconstruction errors acquired as the difference of the delayed input and the output of the filter banks. The dotted vertical line marks the N^{th} timestep.

If the coherent sampling of the input signal cannot be ensured the observer-based method can estimate the signal with greater precision. This claim is supported by Fig. 8 which clearly shows that the naive FIR-based implementation is able to reconstruct the signal but the error settles just below -65 dB. On the other hand the error obtained with the use of the observer does not exceed -290 dB. Furthermore it is able to reconstruct arbitrary non-periodic signals which is due to the fact that the observer is refining it's estimation based on a feedback. [4] describes a method to calculate the samples of a filter bank implemented with an NSGT in real time which handles non-periodic signals as well. It is computationally efficient because it is utilizing the FFT, but it suffers from blocking artifacts. In contrast, the observer based method can estimate these samples in a recursive manner avoiding these artifacts.

VI. Conclusion

The paper reviewed the generalizations of Gabor transforms and their theoretical background. It showed a correspondence between a subset of NSGTs and filter banks. Then it described a conceptual signal model and an observer which can estimate the parameters of the aforementioned model. As a main result a method was given to the observer-based implementation of biorthogonal nonstationary Gabor transforms. This enables the recursive, real-time computation of the transform coefficients with low numerical sensitivity. Lastly three examples are given to illustrate the method. The first two examples demonstrate the design process with the implementation of a simple filter bank and the DFT. Finally a third-octave filter bank is presented and used to examine the desirable numerical properties of the observer-based method compared to the naive approach. It has been showed that it can reconstruct the signal with greater precision due to the fact that the observer is refining it's estimation based on a feedback.

Appendix

A. Diagonality of the frame operator

Theorem 1: Given a nonstationary Gabor frame, if every φ_l has an $N_l \leq N$ finite support and every $b_l \leq \dfrac{N}{N_l}$ then the induced $\mathbf{S}$ frame operator is diagonal.

Proof:

$$\begin{aligned}\mathbf{S}[p,q] &= (\mathbf{\Phi}\mathbf{\Phi}^H)[p,q] = \\ &= \sum_{l=0}^{L-1} \varphi_l[p]\overline{\varphi_l[q]} \sum_{m=0}^{M_l-1} e^{j\frac{2\pi}{M_l}m(p-q)} = \\ &= \sum_{l=0}^{L-1} \varphi_l[p]\overline{\varphi_l[q]} M_l \delta[\mathrm{mod}\,(p-q, M_l)] \end{aligned} \quad (37)$$

The first two equations expand definitions and uses the linearity of the inner sum. The last equation can be derived from the formula for the sum of geometric series. On the main diagonal $\delta[\mathrm{mod}\,(0, M_l)] = 1$, where $\delta[n]$ is the Dirac delta function, so the sum depends only on $\varphi_l[p]$. Off the main diagonal there are two separate cases. If $\delta[\mathrm{mod}\,(p-q, M_l)] = 0$ then the sum is trivially zero. But if $\delta[\mathrm{mod}\,(p-q, M_l)] = 1$ then it is also zero because $|p-q| \geq N_l \leq M_l = \dfrac{N}{b_l}$ so either $\varphi_l[p]$ or $\varphi_l[q]$ or both will lie outside of the support of φ_l. ■

References

[1] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. A. Velasco, "Theory, implementation and applications of nonstationary gabor frames," *Journal of Computational and Applied Mathematics*, vol. 236, no. 6, pp. 1481–1496, 2011. [Online]. Available: https://doi.org/10.1016/j.cam.2011.09.011

[2] K. Siedenburg and M. Dörfler, "Audio denoising by generalized time-frequency thresholding," in *Audio Engineering Society Conference: 45th International Conference: Applications of Time-Frequency Processing in Audio*. Audio Engineering Society, 2012.

[3] G. A. Velasco, N. Holighaus, M. Dörfler, and T. Grill, "Constructing an invertible constant-q transform with nonstationary gabor frames," in *Proceedings of the 14th International Conference on Digital Audio Effects, DAFx 2011*, 09 2011.

[4] 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, pp. 775–785, 2013. [Online]. Available: http://doi.org/10.1109/TASL.2012.2234114

[5] G. Hostetter, "Recursive discrete fourier transformation," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 28, pp. 0–190, 1980. [Online]. Available: http://doi.org/10.1109/tassp.1980.1163389

[6] J. Kovacevic and A. Chebira, "Life beyond bases: The advent of frames (part i)," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 86–104, 2007. [Online]. Available: https://doi.org/10.1109/MSP.2007.4286567

[7] I. Daubechies, A. Grossmann, and Y. Meyer, "Painless nonorthogonal expansions," *Journal of Mathematical Physics*, vol. 27, no. 5, pp. 1271–1283, 1986. [Online]. Available: https://doi.org/10.1063/1.527388

[8] M. Dörfler, "Time-frequency analysis for music signals: A mathematical approach," *Journal of New Music Research*, vol. 30, pp. 3–12, 03 2001. [Online]. Available: http://doi.org/10.1076/jnmr.30.1.3.7124

[9] G. Péceli, "A common structure for recursive discrete transforms," *IEEE Transactions on Circuits and Systems*, vol. 33, pp. 1035–1036, 1986. [Online]. Available: http://doi.org/10.1109/tcs.1986.1085844

[10] G. Péceli, "Resonator-based digital filters," *IEEE Transactions on Circuits and Systems*, vol. 36, no. 1, pp. 156–159, 1989. [Online]. Available: http://doi.org/10.1109/31.16588