Reassignment and synchrosqueezing for general time-frequency filter banks, subsampling and processing

Nicki Holighaus[a,*], Zdenek Prusa[a], Peter L. Søndergaard[b]

[a] *Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12–14, 1040 Vienna, Austria*
[b] *Oticon A/S, 2765 Smørum, Denmark*

Abstract

In this contribution, we extend the reassignment method (RM) and synchrosqueezing transform (SST) to arbitrary time-frequency localized filters and, in the first case, arbitrary downsampling factors. A sufficient condition for the invertibility of the SST is provided. RM and SST are techniques for deconvolution of short-time Fourier and wavelet representations. In both methods, the partial phase derivatives of a complex-valued time-frequency representation are used to determine the instantaneous frequency and group delay associated to the individual representation coefficients. Subsequently, the coefficient energy is *reassigned* to the determined position. Combining RM and frame theory, we propose a processing scheme that benefits both from the improved localization of the reassigned representation and the frame properties of the underlying complex-valued representation. This scheme is particularly interesting for applications that require low redundancy not achievable by an invertible SST.

1. Introduction

Time-frequency representations such as the short-time Fourier transform (STFT) [1, 2] are ubiquitous in signal analysis. In particular their squared magnitude, the *spectrogram* is frequently used to determine the local frequency content of an analyzed signal. The spectrogram however provides a biased, smoothed representation highly dependent on the chosen STFT window function. The smoothing effect in the spectrogram is subject to Heisenberg's uncertainty inequality and thus cannot be arbitrarily reduced.

Therefore, various alternative time-frequency representations have been proposed. The quadratic time-frequency representations in Cohen's class [3, 4] are given by the Wigner-Ville distribution (WVD) [5, 6] convolved with a smoothing kernel. While the spectrogram suffers from a large amount of smoothing, the WVD produces undesirable interference terms. Cohen's class representations often sit between these two extrema and are designed with a certain trade-off between smoothing and interference attenuation in mind, e.g. the smoothed pseudo WVD [7].

*Corresponding author
Email addresses: nicki.holighaus@oeaw.ac.at (Nicki Holighaus), zdenek.prusa@oeaw.ac.at (Zdenek Prusa), peter@sonderport.dk (Peter L. Søndergaard)

Instead of reducing the overall smoothing, representations are often designed to locally provide a good trade-off between time and frequency smoothing with respect to expected signal characteristics. Such representations include (non-uniform) filter banks (FB) [8, 9], in particular, wavelet filter banks [10, 11] (often referred to as time-scale representations), as well as nonstationary Gabor transforms [12, 13] or the representations presented in [14, 15].

Here, we recall the RM originally proposed by Kodera et al. [16] to obtain a sharper time-frequency picture. RM attempts the deconvolution of the spectrogram by means of the partial derivatives of the phase of the underlying STFT. The resulting *reassigned spectrogram* achieves a sharp representation similar to the WVD, but with little interference. Reassignment was made popular by Auger and Flandrin [17, 18], who generalized the method beyond the spectrogram to general Cohen's class time-frequency representations and an equivalent class of time-scale representations. Most importantly, they discovered an efficient means of computation and the application of reassignment. In contrast to the previously mentioned methods, the reassigned spectrogram is not bilinear and even approximate reconstruction is a nontrivial task, although it has been attempted, see [19] for a method based on additive sound modeling or [20]. Therefore, although a valuable analysis tool, RM is deemed unsuitable for signal processing so far.

The SST is a variant of RM dealing with the reconstruction problem, proposed by Daubechies et al. [21, 22, 23]. The main differences between RM and SST are that the latter acts on a complex valued time-frequency representation, hence preserving phase information, and only performs *frequency reassignment*. For recent work on the implementation and application of synchrosqueezing transforms, see e.g. [24, 25, 26, 27] and references therein, or [28] for an approach trying to fuse the invertibility of SST with the superior localization of RM.

In this contribution, we extend RM and SST to filter banks. Furthermore, we demonstrate that the reassigned FB representations, with respect to a *(over-)complete system*, can be used as an interface for processing. Hence, the benefits of a sharpened representation can be combined with the invertibility of the underlying filter bank representation.

2. Preliminaries

Although the results presented in this contribution are applicable for discrete signals, we will present the theoretical material for continuous time signals $f \in \mathbf{L}^2(\mathbb{R})$ to stay consistent with previous publications on RM and SST. We use the following unitary variant of the Fourier transform $\mathcal{F} : \mathbf{L}^2(\mathbb{R}) \mapsto \mathbf{L}^2(\mathbb{R})$, $\hat{f}(\xi) := \mathcal{F}f(\xi) = \int_{\mathbb{R}} f(t)e^{-2\pi i\xi t}\, dt$.

Localized functions: We will often require a function to be *well-localized* around a point $x \in \mathbb{R}$ in a rather informal manner. We consider a function *well-localized* around x if it decays sufficiently fast away from (a small area around) x. Although none of the methods presented below strictly require any localization, all

of them benefit from filters (or window functions) with unique global time and frequency maxima x and ω and better than linear decay around those maxima. Since practically every common time-frequency relevant filter has these properties, this is no severe restriction.

Short-time Fourier transforms, analysis filter banks and frames: For a signal $f \in \mathbf{L}^2(\mathbb{R})$ and window function $g \in \mathbf{L}^2(\mathbb{R})$, let $g_{x,\omega} = g(\cdot - x)e^{2\pi i\omega(\cdot - x)}$. The STFT of f with respect to g is given by

$$V_g f(x,\omega) = \langle f, g_{x,\omega}\rangle = \sqrt{\mathcal{S}_g f(x,\omega)}e^{2\pi i\phi(x,\omega)}, \tag{1}$$

where $\mathcal{S}_g f = |\mathcal{V}_g f|^2$ is the spectrogram and $\phi : \mathbb{R}^2 \to \mathbb{R}$ is the phase of the STFT. The collection $\mathcal{G}(g) := \{g_{x,\omega}\}_{x,\omega\in\mathbb{R}}$ is the (continuous) Gabor system with respect to g. As usual, we assume that $g, \hat{g}$ are well-localized around 0.

Let $\mathbf{g} = \{g_m \in \mathbf{L}^2(\mathbb{R})\}_{m\in\mathbb{Z}}$, $\mathbf{a} = \{a_m \in \mathbb{R}^+\}_{m\in\mathbb{Z}}$ a sequence of functions and time steps, respectively. The system $\{g_{n,m} \ : \ a_m \in \mathbb{R}^+\}_{n,m\in\mathbb{Z}}$ with $g_{n,m} = g_m(\cdot - na_m)$ is a *filter bank (FB)*. We call $\{g_{x,m}\}_{x\in\mathbb{R},m\in\mathbb{Z}}$ with $g_{x,m} = g_m(\cdot - x)$ a *time-continuous filter bank*.

For convenience, we assume that g_m is continuous, well-localized around time 0 and $\widehat{g_m}$ is well-localized around center frequency ω_m with $\omega_j < \omega_m$ if $j < m$ and $\lim_{m\to-\infty}\omega_m = -\infty$, $\lim_{m\to\infty}\omega_m = \infty$. The associated *filter bank analysis* is given by

$$c_{n,m} := c_f[n,m] := \langle f, g_{n,m}\rangle. \tag{2}$$

Wavelet or other filter banks with an accumulation of center frequencies at frequency 0 are completely valid FBs for all the results presented here, but require a slightly less straightforward indexing scheme.

A filter bank forms a frame, if there are constants $0 < A \le B < \infty$, such that $A\|f\|_2^2 \le \|c_f\|_2^2 \le B\|f\|_2^2$, for all $f \in \mathbf{L}^2(\mathbb{R})$. The frame property guarantees the stable invertibility of the coefficient mapping by means of a dual frame $\{\widetilde{g_{n,m}}\}_{n,m\in\mathbb{Z}}$, i.e.

$$f = \sum_{n,m} c_f[n,m]\widetilde{g_{n,m}}, \text{ for all } f \in \mathbf{L}^2(\mathbb{R}). \tag{3}$$

3. Proposed method and statement of results

In this section, we show how to perform reassignment and synchrosqueezing for filter banks, show how the inverse reassignment map can be used in conjunction with invertible filter banks and provide a sufficient condition enabling inversion from synchrosqueezed filter bank coefficients. The derivation of the formulas and conditions can be found in the subsequent sections.

Let $\{g_{n,m} \ : \ a_m \in \mathbb{R}^+\}_{n,m\in\mathbb{Z}}$ be a FB with center frequencies ω_m. We define $g_m^T(t) := tg_m(t)$ and $g_m^F(t) := e^{2\pi i\omega_m t}(e^{-2\pi i\omega_m t}g_m)'(t)$, for all $t \in \mathbb{R}$, $m \in \mathbb{Z}$. With

$$c_f^T[n,m] = \langle f, g_m^T(\cdot - na_m)\rangle \text{ and } c_f^F[n,m] = \langle f, g_m^F(\cdot - na_m)\rangle, \tag{4}$$

we obtain the reassignment operators

$$\mathbf{x}_0(na_m, \omega_m) = na_m + \mathbf{Re}\left(c_f^T[n,m]/c_f[n,m]\right) \quad \text{and} \quad \boldsymbol{\omega_0}(na_m, \omega_m) = \omega_m - \mathbf{Im}\left(c_f^F[n,m]/c_f[n,m]\right), \tag{5}$$

whenever $c_f[n,m] \neq 0$. For applications like mode detection, these estimates of *instantaneous time and frequency* might be sufficient. However, a reassigned representation on the same time-frequency grid $(na_m, \omega_m)_{n,m\in\mathbb{Z}}$ is often desired.

We obtain the *reassigned filter bank* (RFB)

$$c_f^{\mathcal{R}}[n,m] := \sum_{(l,k)\in L_{n,m}} |c_f[l,k]|^2, \quad \text{with } L_{n,m} := \{(l,k)\in\mathbb{Z}^2 \ : \ (\mathbf{n_0}[l,k], \mathbf{m_0}[l,k]) = (n,m)\}. \tag{6}$$

Here, with $\mathbf{x}_0$ and $\boldsymbol{\omega_0}$ as in Eq. (5),

$$\mathbf{m_0}[l,k] := \arg\min_{m\in\mathbb{Z}} |\omega_m - \boldsymbol{\omega_0}(la_k, \omega_k)| \quad \text{and} \quad \mathbf{n_0}[l,k] := \left\lfloor \frac{\mathbf{x}_0(la_k, x_k)}{a_{\mathbf{m_0}[l,k]}} \right\rceil. \tag{7}$$

Synchrosqueezing does not admit the modification of the coefficients time positioning. Therefore, additional care would be required to define a synchrosqueezed filter bank associated to $\{g_{n,m}\}_{n,m\in\mathbb{Z}}$. However, the invertibility property is lost in the presence of intra-channel downsampling and hence, the RFB should be preferred in any case.

Clearly, the time-continuous FB $\{g_{x,m}\}_{x,m\in\mathbb{Z}}$ admits the computation of $\mathbf{x}_0, \boldsymbol{\omega_0}$ at all points (x, ω_m) and we define the associated RFB by

$$c_f^{\mathcal{R}}(x,m) := \sum_{k\in\mathbb{Z}} \int_{\tilde{L}_{x,m,k}} |c_f(y,k)|^2 \, dy, \quad \text{with } \tilde{L}_{x,m,k} := \{y\in\mathbb{R} \ : \ (\mathbf{x}_0(y,\omega_k), \mathbf{m_0}(y,k)) = (x,m)\}. \tag{8}$$

In this setting, we can also define the *synchrosqueezed filter bank* (SSFB)

$$c_f^{\mathcal{S}}(x,m) := \sum_{k\in K_{x,m}} c_f(x,k), \quad \text{with} \quad K_{x,m} := \{k\in\mathbb{Z} \ : \ \mathbf{m_0}(x,k) = m\}. \tag{9}$$

Here, $\mathbf{m_0}(x,k) := \arg\min_{m\in\mathbb{Z}} |\omega_m - \boldsymbol{\omega_0}(x,\omega_k)|$.

The framework above includes Gabor and wavelet filter banks, the latter after a suitable change of the center frequency indexing scheme. Thus, it can be considered a generalization of the RM and the SST to a larger class of representations. In fact, filter bank structure is completely optional for the RM. If a function g is well-localized around time x and frequency ω, we can define $g^T(t) := (t-x)g(t)$ and $g^F(t) := e^{2\pi i\omega t}(e^{-2\pi i\omega t}g)'(t)$. Then the reassignment operators can be defined as before and, given a family of such functions, a variant of $c_f^{\mathcal{R}}$ is easily obtained.

Frames and the inverse reassignment map: The RM does not provide a means of perfect (or even good quality approximate) reconstruction. Therefore, processing a signal directly by modifying the RFB is not feasible. However, the additional information provided by the instantaneous time and frequency estimates provided by Eq. (5) may prove useful in many a signal processing task.

Thus, we propose to use the RFB as a processing interface for invertible FB representations. Such FBs are characterized by the frame property. By storing the original FB coefficients c_f and the *inverse reassignment map*

$$\mathbf{iR}[n,m] = \{(l,k) \in \mathbb{Z}^2 \ : \ (\mathbf{n_0}[l,k], \mathbf{m_0}[l,k]) = (n,m)\}, \tag{10}$$

a processing tool can allow the user to interact with the RFB $c_f^{\mathcal{R}}$ only, while the software employs $\mathbf{iR}$ to apply the requested processing onto c_f. Since the FB used forms a frame, the processed signal is obtained simply by frame synthesis, see Eq. (3).

In this framework, the frame property guarantees that the synthesis operation (and therefore the processing) is stable, while the RFB provides an intuitive representation incorporating the instantaneous time and frequency estimates provided by Eq. (5). Examples of such processing can be found in Section 5.

It should be noted that the idea of using the inverse reassignment map for processing is not completely new, see [20] for a crude variant of the method proposed here. However, this usage of Eq. (5) is surprisingly rare and neither subchannel downsampling, nor invertibility of the underlying representation are systematically considered.

Invertible synchrosqueezed filter banks: The SST has been introduced for continuous wavelet and short-time Fourier representations to circumvent the RM's lack of invertibility. A reconstruction formula is obtained by exploiting the fact that under suitable conditions on the analysis wavelet, the Dirac δ distribution can be used as synthesis wavelet. In the wavelet context, the corresponding reconstruction formula is known as Morlet's formula [29]. For a time-continuous filter bank $\{g_{x,m}\}_{x\in\mathbb{R},m\in\mathbb{Z}}$, a simple condition enables reconstruction of all $f \in \mathbf{L}^2(\mathbb{R})$ from from their SSFB $c_f^{\mathcal{S}}$.

Proposition 1. *Let $\{g_{x,m}\}_{x\in\mathbb{R},m\in\mathbb{Z}}$ be a time-continuous filter bank with $\sum_{m\in\mathbb{Z}} |\widehat{g_m}| \in \mathbf{L}^\infty(\mathbb{R})$. The following equation holds*

$$\mathcal{F}^{-1}\left(\hat{f} \sum_{m\in\mathbb{Z}} \widehat{g_m}\right) = \sum_{m\in\mathbb{Z}} c_f^{\mathcal{S}}(\cdot,m) \quad a.e. \tag{11}$$

In particular, if $0 < A \leq \sum_{m\in\mathbb{Z}} \widehat{g_m}$ almost everywhere,

$$f = \mathcal{F}^{-1}\left(\left(\sum_{m\in\mathbb{Z}} \widehat{g_m}\right)^{-1} \widehat{\left(\sum_{m\in\mathbb{Z}} c_f^{\mathcal{S}}(\cdot,m)\right)}\right) \quad a.e. \tag{12}$$

Proof. Equation (12) is an immediate consequence of Eq. (11). To show Eq. (11), we first assume $f \in \mathbf{L}^2(\mathbb{R}) \cap \mathcal{F}\mathbf{L}^1(\mathbb{R})$. We have

$$\begin{aligned}\sum_{m\in\mathbb{Z}} \sum_{k\in K_{x,m}} |c_f(x,k)| &= \sum_{m\in\mathbb{Z}} |c_f(x,m)| \leq \sum_{m\in\mathbb{Z}} \int_{\mathbb{R}} |\hat{f}(\xi)||\widehat{g_m}(\xi)| \, d\xi \\ &= \int_{\mathbb{R}} |\hat{f}(\xi)| \sum_{m\in\mathbb{Z}} |\widehat{g_m}(\xi)| \, d\xi \leq \left\| \sum_{m\in\mathbb{Z}} |\widehat{g_m}| \right\|_\infty \|\hat{f}\|_1 < \infty.\end{aligned}$$

Therefore, we can change summation order and interchange sum and integral in the following chain of equations.

$$\sum_{m\in\mathbb{Z}}\sum_{k\in K_{x,m}} c_f(x,k) = \sum_{m\in\mathbb{Z}} c_f(x,m) = \sum_{m\in\mathbb{Z}}\int_{\mathbb{R}} \hat{f}(\xi)\widehat{g_m}(\xi)e^{2\pi i x\xi}\, d\xi$$
$$= \int_{\mathbb{R}} \hat{f}(\xi)e^{2\pi i x\xi}\sum_{m\in\mathbb{Z}}\widehat{g_m}(\xi)\, d\xi = \mathcal{F}^{-1}\left(\hat{f}\sum_{m\in\mathbb{Z}}\widehat{g_m}\right),$$

where we used Parseval's identity and the definition of $g_{n,m}$ in the second step. For general $f \in \mathbf{L}^2(\mathbb{R})$, Eq. (11) follows by density. □

Since (intra-channel) downsampling necessarily introduces aliasing, there is no straightforward way to obtain an invertible SSFB for a filter bank $\{g_{n,m}\}_{n,m\in\mathbb{Z}}$, even if it is uniform, i.e. $a_m = a \in \mathbb{R}^+$ for all $m \in \mathbb{Z}$.

A remark on implementation. Proposition 1 has a straightforward analogue in the setting of discrete signals. In that setting, aliasing occurs during filtering whenever the downsampling factor is larger than 1. Thus, the analogue to Proposition 1 covers the case where all downsampling factors equal 1. Clearly, the frame method described for the RFB is also applicable for the SSFB, but in that setting, it seems more reasonable to utilize possibly higher sparsity of the RFB. Furthermore, storing only the non-negative RFB coefficients noticeably reduces space complexity.

4. Derivation of the reassignment operators

The derivation of the reassignment operators, given in Eq. (5), below should serve as indication that Eq. (5) makes sense and provides the desired result, as verified by the Examples in Section 5.

Consider a function $g \in \mathbf{L}^2(\mathbb{R})$, well-localized around x with its Fourier transform $\hat{g}$ well-localized around ω. Then

$$\tilde{g}(t) := e^{-2\pi i\omega t}g(t+x) \tag{13}$$

is well-localized around 0 in time and frequency and we have

$$V_{\tilde{g}}f(x,\omega) = \langle f, \tilde{g}_{x,\omega}\rangle = \langle f, g\rangle. \tag{14}$$

For the STFT $V_{\tilde{g}}f$, the reassignment operators are defined as $\mathbf{x}_0(x,\omega) = x - \frac{\partial}{\partial\omega}\phi(x,\omega)$ and $\boldsymbol{\omega_0}(x,\omega) = \frac{\partial}{\partial x}\phi(x,\omega)$, where ϕ is the phase of $V_{\tilde{g}}f$ as in Eq. (1).

Auger and Flandrin [17] have shown that the reassignment operators can be expressed as the pointwise product of the STFTs with respect to 3 window functions, depending on the window $\tilde{g}$:

$$\mathbf{x}_0(x,\omega) = x + \mathbf{Re}\left(\mathcal{V}_{\tilde{g}_T}f(x,\omega)/\mathcal{V}_{\tilde{g}}f(x,\omega)\right) = x + \mathbf{Re}\left(\frac{\mathcal{V}_{\tilde{g}_T}f(x,\omega)\overline{\mathcal{V}_{\tilde{g}}f(x,\omega)}}{\mathcal{S}_{\tilde{g}}f(x,\omega)}\right) \quad \text{and} \tag{15}$$

$$\boldsymbol{\omega_0}(x,\omega) = \omega - \mathbf{Im}\left(\mathcal{V}_{\tilde{g}'}f(x,\omega)/\mathcal{V}_{\tilde{g}}f(x,\omega)\right) = \omega - \mathbf{Im}\left(\frac{\mathcal{V}_{\tilde{g}'}f(x,\omega)\overline{\mathcal{V}_{\tilde{g}}f(x,\omega)}}{\mathcal{S}_{\tilde{g}}f(x,\omega)}\right), \tag{16}$$

whenever $\mathcal{S}_{\tilde{g}} f(x,\omega) \neq 0$ and 0 else. Here, $\tilde{g}'(t) = \frac{\partial}{\partial t}\tilde{g}(t)$ is the derivative and $\tilde{g}_T(t) := t\tilde{g}(t)$ is a time weighted version of $\tilde{g}$.

Note that differentiation is a translation-invariant operator and time weighting is invariant under modulation, i.e. multiplication with $e^{2\pi i\xi t}$ for $\xi \in \mathbb{R}$. Therefore, with $g^T(t) := (t-x)g(t)$ and $g^F(t) := e^{2\pi i\omega t}(e^{-2\pi i\omega t}g)'(t)$ as in Section 3, we obtain

$$\mathcal{V}_{\tilde{g}_T} f(x,\omega) = \langle f, g^T \rangle, \quad \text{and} \quad \mathcal{V}_{\tilde{g}'} f(x,\omega) = \langle f, g^F \rangle. \tag{17}$$

Given a filter bank or continuous filter bank, simply set $g = g_m$ to derive Eq. (5).

5. Discussion

In the previous sections, we have shown how to apply reassignment and synchrosqueezing techniques to general filter banks with time-frequency localized filters. Figure 1 shows that reassignment achieves the expected result. For demonstration purposes, we chose here a perceptually motivated non-uniform FB adapted to the ERB scale [30], the so called ERBlet FB [31, 32]. The crucial point here is that the method works with non-standard FBs not covered by earlier results, not the specific FB choice.

We prefer to show the RFB over the SSFB in most cases, since the latter performs similar to (in the absence of high frequency modulation or pulse-like signal components) or worse than the RFB in terms of localizing and concentrating the signal representation. Additionally, the usage of the RFB in the examples serves to show that intra-channel subsampling is not detrimental to the reassignment performance.

Some applications do not require intra-channel subsampling and are thus suitable for the computation of the SSFB. Still, there might be reasons to consider the usage of the RFB over the SSFB. Although not strictly necessary, subsampling results in a considerable reduction of time and space complexity by significantly reducing the redundancy of the underlying representation, cf. Example 1. In the presence of high frequency modulation and/or pulse-like signal components, the RFB will usually provide a better localized representation. Finally, the synthesis formula for the SSFB essentially amounts to using a synthesis FB comprised of Dirac δ distributions. It is well known that such a synthesis is not very robust to subchannel processing. This has also been observed in practice [33]. Synthesis with a dual FB frame on the other hand is known to provide robust synthesis in the presence of processed coefficients [34].

5.1. Notes on implementation

We provide implementations for filter bank reassignment and synchrosqueezing in the LTFAT Toolbox [35], available at `ltfat.github.io` since version 2.1.1. The phase derivatives are provided by `filterbankphasegrad`. The RFB and SSFB are computed by `filterbankreassign` and `filterbanksynchrosqueeze`, respectively. For examples, please visit our website `http://ltfat.github.io/notes/041`, where scripts reproducing all Figures and Experiments can be found. The scripts can be run in MATLAB as well as in GNU Octave, a free MATLAB alternative.

5.2. Examples

We apply the RFB and/or SSFB in three different processing setups. All the experiments based on the RFB use the inverse reassignment map $\mathbf{iR}[n,m] = \{(l,k) \in \mathbb{Z}^2 \ : \ (\mathbf{n_0}[l,k], \mathbf{m_0}[l,k]) = (n,m)\}$ and the invertibility of the underlying FB representation, as described in Section 3.

Example 1 attempts denoising by hard thresholding on the reassigned (or synchrosqueezed) representation. An excerpt of the second test signal in Figure 1, sampled at 44.1 kHz, is contaminated with uniformly distributed noise with expected value 0 to a SNR of 0 dB, where $\mathrm{SNR}(f_{\mathrm{noisy}}) = 20\log_{10}(\frac{\|f_{\mathrm{clean}}\|}{\|f_{\mathrm{clean}} - f_{\mathrm{noisy}}\|})$ dB.

For each representation the optimal threshold in terms of SNR has been determined manually in 1 dB steps. Note that this setup retains a relatively large number of isolated, noisy coefficients and a slightly larger threshold leads to perceptually better results. For this example, an invertible constant-Q FB [36, 37, 38] with 48 bins per octave is used, resulting in a total of 300 channels spaced between 0 Hz and 22.05 kHz. Considering only nonzero coefficients, the SSFB has a redundancy factor of ≈ 429, while the complete processing setup for the RFB, i.e. the nonzero RFB coefficients, phase partial derivatives and the original FB representation has a redundancy of ≈ 9.3. Therefore the usage of the RFB and the frame method provides a massive reduction in space complexity, also leading to a noticeable drop in computation time, although the latter may vary depending on the hardware. Most importantly, the resulting denoised signals have SNR of ≈ 9.5 dB and ≈ 10.6 dB for the synchrosqueezing- and reassignment-based results, respectively. Although the increase in SNR is not very large, the perceptual quality of the reassignment-based result is significantly better (see website). Although no significant improvement over thresholding the constant-Q coefficients directly has been observed, this demonstrates that the proposed processing frameworks, in particular the one based on RFB, are robust and reliable. Moreover, it has been found that structured thresholding on the SST coefficients can improve denoising performance over block-thresholding [23]. Similar performance can be expected for RFB and SSFB. Note that reassignment-based denoising on the undecimated filter bank did not yield significant improvements over the decimated variant. The intermediate representations can be seen in Figure 2 and the experimental setup and results can be reproduced using the MATLAB script `exp1_denoise.m` provided on the webpage.

Example 2 attempts mode decomposition on a synthetic signal similar to those treated in previous synchrosqueezing contributions, cf. [28, 26, 22]. The test signal is given by the sum $s = C_1 + C_2 + C_3$ with

$$C_1(t) = \cos(4800\pi t + 350\cos(10\pi t)), \quad C_2(t) = \cos(20000\pi t - 5400\cos(2\pi(t-1))),$$

$$C_3(t) = \cos(750\pi t + 80\cos(4\pi(t-0.03))).$$

with $t \in [0,1[$ and sampled at 44.1 kHz. The ERBlet filter bank employed for signal analysis has 6 bins per ERB resulting in a total of 256 channels spaced between 0 Hz and 22.05 kHz. A set of binary masks was designed to separate the 3 signal components in the RFB and SSFB. All 3 signal representations are shown

in Figure 3, alongside local reconstruction errors of components s_2 and s_3 at critical points. Reconstruction quality is good for all 3 signal components and both approaches, but Table 1 shows that overall, the RFB provides better results. The experimental setup can be reproduced using the MATLAB script `exp2_modes.m` provided on the webpage.

Example 3 attempts the isolation of a single note in a musical signal. This is similar to the previous experiment but in a more realistic setting. A binary mask for the isolation of the second note in a glockenspiel signal was designed and applied to the RFB. The component was synthesized using the frame method described in Section 3. The improved separation of signal components in the RFB somewhat simplifies the mask design process. Instead of constructing the mask manually, techniques such as ridge detection in the RFB could possibly assist the mask design process. Since the isolation quality cannot be evaluated objectively, we provide the results in the form of a re-analysis of the separated components, see Figure 4. However, the resulting audio signals can be obtained by executing the MATLAB script `exp3_separate.m` provided on the webpage.

6. Conclusion

We have motivated and demonstrated the application of the reassignment method and synchrosqueezing to general filter banks. For the reassigned filter bank, we also allow arbitrary downsampling factors, providing significant and in real application often crucial, redundancy reduction. We provide a sufficient condition for the invertibility of the synchrosqueezed filter bank representation and propose a processing framework that is able to harness the improved time-frequency localization of the reassigned filter bank, while preserving the perfect reconstruction properties of the underlying filter bank representation. Thus, we overcome a major perceived weakness of the reassignment method.

Efficient implementations of the presented methods have been included in the open-source MATLAB/Octave toolbox LTFAT, and freely available to the community. The figures and examples in this contribution can be easily reproduced using the script files we provide, see `http://ltfat.github.io/notes/041`.

A number of methods exist, trying to further refine the RM and/or SST for various purposes, e.g. multi-tapered reassignment/synchrosqueezing [39, 40]. With the proposed methods, such refinements can also be applied in the more general filter bank setting.

7. Acknowledgement

This work was supported by the Austrian Science Fund (FWF) START-project FLAME ("Frames and Linear Operators for Acoustical Modeling and Parameter Estimation"; Y 551-N13). The authors would like to thank the reviewers for their valuable comments.

References

[1] D. Gabor, Theory of communication, J. IEE 93 (26) (1946) 429–457.

[2] K. Gröchenig, Foundations of Time-Frequency Analysis, Appl. Numer. Harmon. Anal., Birkhäuser, Boston, MA, 2001.

[3] L. Cohen, Time-frequency distributions - a review, Proc. IEEE 77 (7) (1989) 941–981.

[4] L. Cohen, Time-Frequency Analysis: Theory and Applications., Prentice Hall Signal Processing Series, Prentice Hall, 1995.

[5] E. P. Wigner, On the quantum correction for thermodynamic equilibrium., Phys. Rev., II. Ser. 40 (1932) 749–759.

[6] J. d. Ville, et al., Théorie et applications de la notion de signal analytique, Cables et transmission 2 (1) (1948) 61–74.

[7] W. Martin, P. Flandrin, Wigner-ville spectral analysis of nonstationary processes, Acoustics, Speech and Signal Processing, IEEE Transactions on 33 (6) (1985) 1461–1470.

[8] M. Vetterli, Filter banks allowing perfect reconstruction, Signal Process. 10 (1986) 219–244.

[9] S. Akkarakaran, P. Vaidyanathan, Nonuniform filter banks: new results and open problems, in: Beyond wavelets, Vol. 10 of Studies in Computational Mathematics, Elsevier, 2003, pp. 259–301.

[10] M. Vetterli, J. Kovačević, Wavelets and subband coding, Vol. 87, Prentice Hall PTR Englewood Cliffs, New Jersey, 1995.

[11] S. Mallat, A wavelet tour of signal processing: The sparse way, Third Edition, Academic Press, 2009.

[12] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, G. A. Velasco, Theory, implementation and applications of nonstationary Gabor frames, J. Comput. Appl. Math. 236 (6) (2011) 1481–1496.

[13] M. Liuni, P. Balazs, A. Röbel, Sound Analysis and Synthesis Adaptive in Time and Two Frequency Bands, in: Proc. of the 14th Int. Conference on Digital Audio Effects (DAFx-11), Paris, France, September 19-23, 2011.

[14] C. Herley, J. Kovacevic, K. Ramchandran, M. Vetterli, Tilings of the time-frequency plane: Construction of arbitrary orthogonal bases and fast tiling algorithms, Signal Processing, IEEE Transactions on 41 (12) (1993) 3341–3359.

[15] M. Dörfler, Quilted Gabor frames - A new concept for adaptive time-frequency representation, Adv. in Appl. Math. 47 (4) (2011) 668–687.

[16] K. Kodera, C. De Villedary, R. Gendrin, A new method for the numerical analysis of non-stationary signals, Physics of the Earth and Planetary Interiors 12 (2) (1976) 142–150.

[17] F. Auger, P. Flandrin, Improving the readability of time-frequency and time-scale representations by the reassignment method, Signal Processing, IEEE Transactions on 43 (5) (1995) 1068 –1089.

[18] E. Chassande Mottin, I. Daubechies, F. Auger, P. Flandrin, Differential reassignment, IEEE Signal Processing Letters 4 (10) (1997) 293 –294.

[19] K. Fitz, L. Haken, On the use of time: Frequency reassignment in additive sound modeling, Journal of the Audio Engineering Society 50 (11) (2002) 879–893.

[20] L. Han, M. D. Sacchi, L. Han, Spectral decomposition and de-noising via time-frequency and space-wavenumber reassignment, Geophysical Prospecting 62 (2) (2014) 244–257.

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

[22] I. Daubechies, J. Lu, H.-T. Wu, Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool, Applied and computational harmonic analysis 30 (2) (2011) 243–261.

[23] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H.-T. Wu, Time-frequency reassignment and synchrosqueezing: An overview, Signal Processing Magazine, IEEE 30 (6) (2013) 32–41. doi:10.1109/MSP.2013.2265316.

[24] G. Thakur, E. Brevdo, N. S. Fučkar, H.-T. Wu, The synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications, Signal Processing 93 (5) (2013) 1079 – 1094.

[25] A. Ahrabian, D. Looney, L. Stanković, D. P. Mandic, Synchrosqueezing-based time-frequency analysis of multivariate data, Signal Processing 106 (0) (2015) 331 – 341.

[26] C. Li, M. Liang, A generalized synchrosqueezing transform for enhancing signal time–frequency representation, Signal Processing 92 (9) (2012) 2264 – 2274.

[27] S. Wang, X. Chen, Y. Wang, G. Cai, B. Ding, X. Zhang, Nonlinear squeezing time–frequency transform for weak signal detection, Signal Processing 113 (0) (2015) 195 – 210.

[28] T. Oberlin, S. Meignen, V. Perrier, Second-order synchrosqueezing transform or invertible reassignment? towards ideal time-frequency representations, Signal Processing, IEEE Transactions on 63 (5) (2015) 1335–1344.

[29] M. Farge, Wavelet transforms and their applications to turbulence, Annual review of fluid mechanics 24 (1) (1992) 395–458.

[30] B. R. Glasberg, B. C. J. Moore, Derivation of auditory filter shapes from notched-noise data, Hear. Res. 47 (1990) 103–138.

[31] T. Necciari, P. Balazs, N. Holighaus, P. Sondergaard, The ERBlet transform: An auditory-based time-frequency representation with perfect reconstruction, in: Proceedings of the 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), 2013, pp. 498–502.

[32] T. Necciari, N. Holighaus, P. Balazs, Z. Prusa, A perfect reconstruction auditory filter bank for audio signal processing, submitted.

[33] T. Oberlin, S. Meignen, V. Perrier, On the mode synthesis in the synchrosqueezing method, in: Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European, IEEE, 2012, pp. 1865–1869.

[34] O. Christensen, Frames and Bases. An Introductory Course., Applied and Numerical Harmonic Analysis. Basel Birkhäuser, 2008.

[35] Z. Průša, P. L. Søndergaard, N. Holighaus, C. Wiesmeyr, P. Balazs, The Large Time-Frequency Analysis Toolbox 2.0, in: M. Aramaki, O. Derrien, R. Kronland-Martinet, S. Ystad (Eds.), Sound, Music, and Motion, Lecture Notes in Computer Science, Springer International Publishing, 2014, pp. 419–442.

[36] J. Brown, Calculation of a constant Q spectral transform, J. Acoust. Soc. Amer. 89 (1) (1991) 425–434.

[37] N. Holighaus, M. Dörfler, G. A. Velasco, T. Grill, A framework for invertible, real-time constant-Q transforms, IEEE Trans. Audio Speech Lang. Process. 21 (4) (2013) 775 –785.

[38] C. Schörkhuber, A. Klapuri, N. Holighaus, M. Dörfler, A Matlab toolbox for efficient perfect reconstruction time-frequency transforms with log-frequency resolution, in: Audio Engineering Society Conference: 53rd International Conference: Semantic Audio, AES, 2014.

[39] J. Xiao, P. Flandrin, Multitaper time-frequency reassignment for nonstationary spectrum estimation and chirp enhancement, Signal Processing, IEEE Transactions on 55 (6) (2007) 2851–2860.

[40] I. Daubechies, Y. Wang, H.-T. Wu, ConceFT: Concentration of Frequency and Time via a multitapered synchrosqueezed transform, arXiv preprint arXiv:1507.05366.

	s_1	s_2	s_3
RFB	52.4 dB	56.6 dB	54.5 dB
SSFB	44.6 dB	45.2 dB	53.2 dB

Table 1: (Example 2) Reconstruction errors for the individual components of the test signal s.

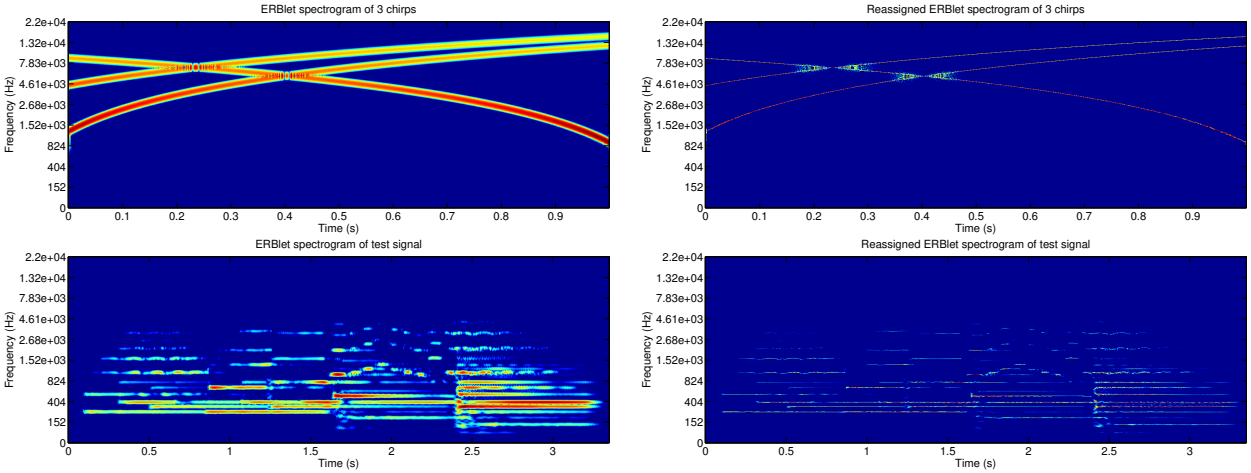

Figure 1: Reassignment of ERBlet filter bank representations: (top) Reassignment perfectly localizes linear chirps that are well separated in the time-frequency plane. Separation quality decreases in the vicinity of crossover points. (bottom) Reassignment of a violin and piano test signal.

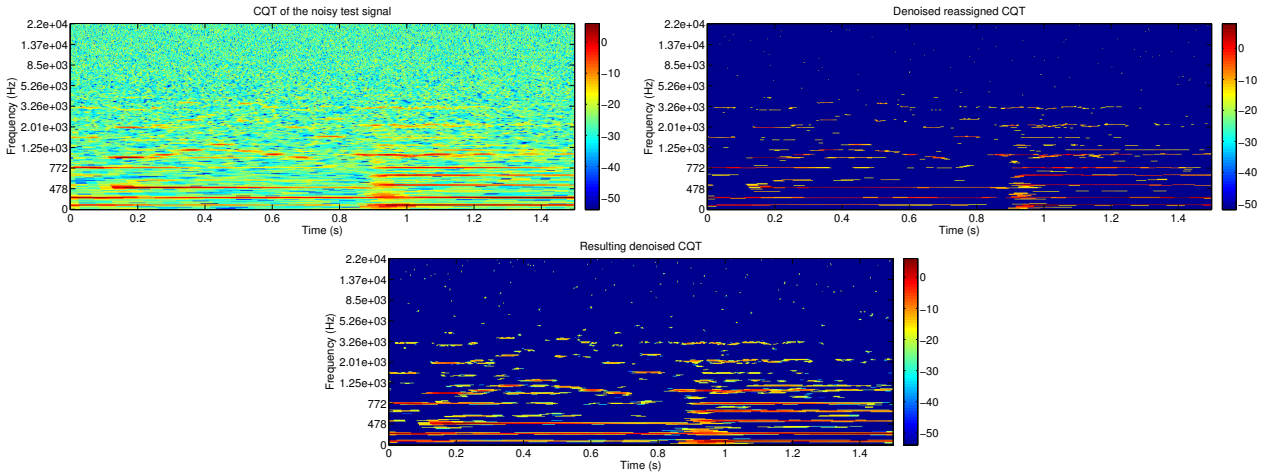

Figure 2: (Example 1) Spectrogram of the test signal contaminated with noise (0 dB SNR) alongside denoised reassigned coefficients and resulting denoised constant-Q representation (10.8 dB SNR).

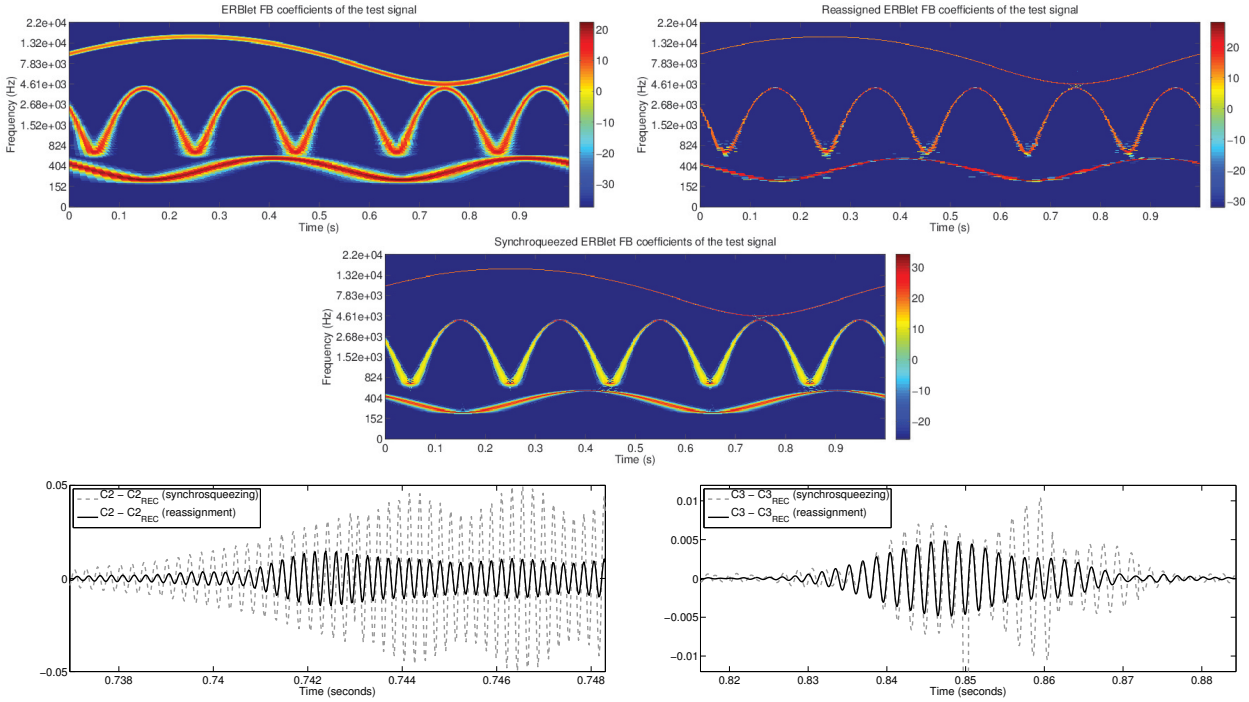

Figure 3: (Example 2) Filter bank, RFB and SSFB representations of the synthetic signal s (top/center). Local reconstruction error of C_2 and C_3 around the positions of least separation to C_1.

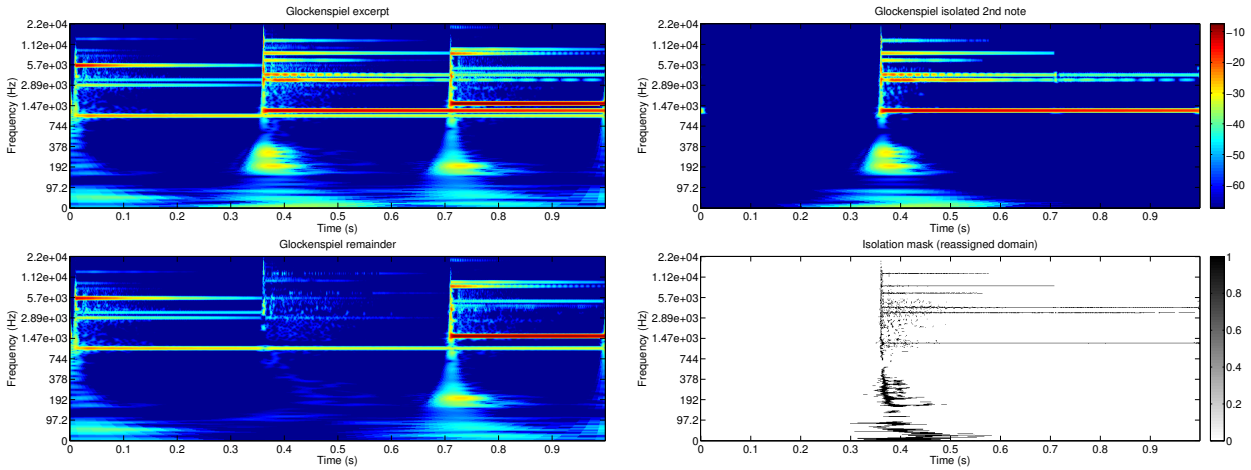

Figure 4: (Example 3) Separation of the 2nd glockenspiel note.