ANALYSIS OF STRONGLY MODULATED MULTICOMPONENT SIGNALS WITH THE SHORT-TIME FOURIER TRANSFORM

T. Oberlin, S. Meignen

Laboratoire Jean Kuntzmann,
University of Grenoble and CNRS, France

S. McLaughlin, Fellow IEEE

School of Engineering and Physical Sciences,
Heriot-Watt University, Edinburgh, Scotland, UK

ABSTRACT

This paper addresses the issue of the retrieval of the components of a multicomponent signal from its short-time Fourier transform. It recalls two popular reconstruction methods, and extends each of them for the case of strong frequency modulation, by taking into account the second derivative of the phase. Numerical experiments illustrate the improvement and compare the methods.

***Index Terms*—** multicomponent signals, short-time Fourier transform, ridge analysis, chirps

1. INTRODUCTION

Many signals from the physical world, e.g. speech or physiological records, can be modelled as a sum of amplitude- and frequency-modulated (AM/FM) waves often termed modes. In the past few decades, there has been an increasing interest in designing new accurate representations and processing methods for these types of signals.

The short-time Fourier transform (STFT) is one of the simplest linear time-frequency (TF) transform, and is well adapted to process such signals. The starting point for the analysis of these signals is to compute their short-time Fourier transform (STFT), whose local maxima in frequency appear to draw so-called *ridges* in the TF plane. Based on a first order approximation of the phase of the components making up the signal, a mode retrieval algorithm is obtained by considering the STFT on each ridge (see for instance [1, 2]). However, this first-order approximation appears to be particularly sensitive to the ridge estimation, and may be no longer valid either when the modes are not well separated or when they are contaminated by noise. For this reason, it may be sensible to consider the transform not only on the ridge but in its vicinity. This key idea was exploited by the *synchrosqueezing transform* designed for both the STFT [3] and the continuous wavelet transform [4]. An implementation of these ideas was proposed in [5] and consisted of integrating the transform in frequency close to the ridge with applications to multicomponent signal denoising [6].

Yet, these methods were designed for weakly frequency modulated modes which is not relevant in many applications (e.g. in radar or speech processing). This paper thus aims to investigate how to extend these two mode retrieval methods to strongly frequency modulated modes. To do so, we first recall the approximation of the STFT of multicomponent signals under the weak modulation hypothesis, then we extend the reconstruction methods to strong frequency modulations. Finally, numerical simulations validate the proposed reconstruction extensions, and underline the differences between the reconstruction on the ridge and in its vicinity.

2. THE SHORT-TIME FOURIER TRANSFORM

Linear TF representations are particularly adapted to process modulated signals. We define here the STFT for tempered distributions and recall its first-order approximation that generates two different mode reconstruction methods.

2.1. Definitions

In the following, we denote by $L^1(\mathbb{R})$ the space of real integrable functions, and by $\mathcal{S}(\mathbb{R})$ and $\mathcal{S}'(\mathbb{R})$ the Schwartz class and the space of tempered distributions, respectively. χ_X stands for the indicator function of the set X, and $\bar{z}$ is the complex conjugate of z. Given a signal $s \in L^1(\mathbb{R})$, we define its Fourier transform by:

$$\hat{s}(\xi) = \int_{\mathbb{R}} s(t)\, e^{-2i\pi\xi t}\, dt. \tag{1}$$

Taking a window $g \in \mathcal{S}(\mathbb{R})$, the (modified) STFT of a signal $s \in \mathcal{S}'(\mathbb{R})$ is defined by

$$\begin{aligned} V_s(b,\eta) &= s * \overline{g_\eta}(b) \text{ with } g_\eta(t) = g(t)e^{2i\pi\eta t} \\ &= \int_{\mathbb{R}} s(t)g(t-b)e^{-2i\pi\eta(t-b)}\, dt \end{aligned} \tag{2}$$

where the convolution has to be understood in the distributional sense. If the window is the Gaussian function $g(t) = \sigma^{-\frac{1}{2}} e^{-\pi\frac{t^2}{\sigma^2}}$ then it is called the Gabor transform. The STFT admits the following synthesis formula:

$$s(t) = \frac{1}{g(0)} \int_{\mathbb{R}} V_s(t,\eta)\, d\eta \tag{3}$$

Remark: equation (3) is true only when $\eta \mapsto V_s(b,\eta)$ is integrable, which will always be the case in this paper.

Hereafter, we will assume that g (and thus $\hat{g}$) is real and even, and that $\hat{g}$ admits a global maximum at 0.

2.2. Monochromatic waves and Multicomponent Signals

A *quasi-monochromatic signal* is a modulated wave $h(t) = a(t)e^{2i\pi\phi(t)}$, where $a'(t)$ and $\phi''(t)$ are small compared to $\phi'(t)$. The STFT of such a wave forms a *ridge* in the TF plane, i.e. a curve with equation $(b, \phi'(b))$ (see [7] for instance). More precisely, we get the first-order approximation

$$V_h(b,\eta) = h(b)\hat{g}(\eta - \phi'(b)) + O(\frac{|a'(b)|}{|\phi'(b)|}, \frac{|\phi''(b)|}{|\phi'(b)|}), \quad (4)$$

where $O(X,Y)$ is the Landau notation for a quantity of the order of X and Y. A multicomponent signal can be modeled as a sum of quasi-monochromatic waves: $s(b) = \sum_{k=1}^{K} s_k(b) = \sum_{k=1}^{K} a_k(b)e^{2i\pi\phi_k(b)}$, where $|a_k'(b)|, |\phi_k''(b)| \ll \phi_k'(b)$. By linearity of the STFT and if the components are *separated*, i.e. $|\phi_k'(b) - \phi_l'(b)|$ is high enough for $k \neq l$, the STFT of s is not spoiled by mode interferences, and one is then able to detect the ridge associated to each mode and then proceed with mode reconstruction. Under this separation hypothesis and without any loss of generality, one can focus on the reconstruction of a single mode $h(t) = a(t)e^{2i\pi\phi(t)}$, for which two types of reconstruction procedure are available. The first one uses the value of the STFT on the ridge, i.e.

$$h(b) \approx \frac{1}{\hat{g}(0)} V_h(b, \phi'(b)). \quad (5)$$

Hereafter, we denote this reconstruction method by RR (for Ridge Reconstruction). Another way to reconstruct the modes is to locally integrate in frequency the STFT, namely:

$$h(b) \approx \frac{1}{g(0)} \int_{\phi'(b)-\Delta}^{\phi'(b)+\Delta} V_h(b,\eta)\, d\eta, \quad (6)$$

where Δ is the radius of the support of $\hat{g}$. When $\hat{g}$ is not compactly supported (typically when g is the Gaussian window), one can introduce a threshold ε and design Δ so as to select only the coefficients such that $|V_h(b,\eta)| > \varepsilon |V_h(b,\phi'(b))|$. As an illustration, for the Gaussian window we obtain

$$\Delta = \sqrt{\frac{-\log \varepsilon}{\pi\sigma^2}}. \quad (7)$$

The major benefit of this integration is that when h is contaminated by noise, one can adjust ε and thus Δ to the noise level, improving the reconstruction by making it less sensitive to the estimation of the instantaneous frequency $\phi'(b)$. A formula of type (7) has been successfully used in [5, 6] in a wavelet context, but Δ was fixed whatever the noise level. We denote this method by IR (for Integral Reconstruction).

However, many multicomponent signals contain strongly frequency modulated modes (e.g. chirps used in radar processing [8] or those associated to bat echolocations [9]), making this first-order approximation of the phase irrelevant for modes reconstruction. Figure 1 illustrates these difficulties, displaying the magnitude of the Gabor transform and its values for a fixed time, both for a weakly and a strongly modulated wave. For the former signal it is clear that equation (6) is relevant because we extract most of the information but irrelevant for the latter. How to take into account strong frequency modulations is the subject of the following sections.

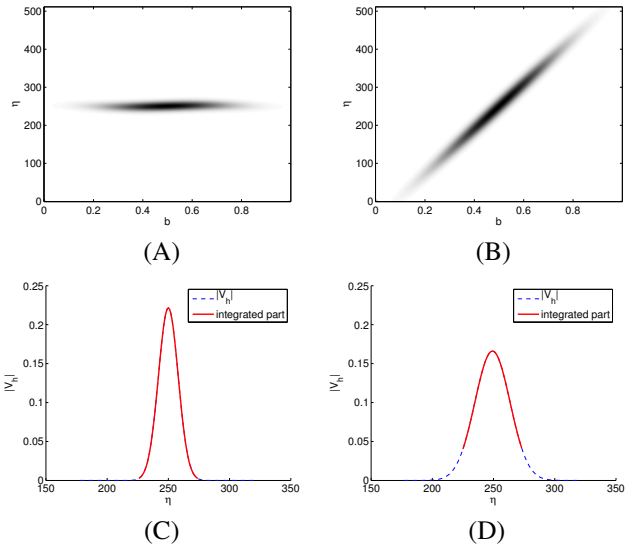

Fig. 1. Illustration of the IR method for two different signals. (A): the magnitude of the Gabor transform of a weakly modulated chirp, with $\sigma = 0.05$, (B): idem for a strongly modulated one, (C): slice of the graph displayed in A for $b = 0.5$, along with the selected coefficients (integrated part) when $\varepsilon = 0.01$, (D): idem as (C) but for the graph displayed in (B).

3. A NEW MODEL BASED ON SECOND ORDER APPROXIMATION

We consider here signals $h(t) = a(t)e^{2i\pi\phi(t)}$ which behave locally as linear chirps (i.e., waves with a linear instantaneous frequency $\phi'(t)$), so that the approximation of section 2 is no longer valid. To estimate $V_h(b,\eta)$ we now use the following approximation instead:

$$\tilde{h}_b(t) = h(b)e^{2i\pi[\phi'(t)(t-b)+\frac{1}{2}\phi''(b)(t-b)^2]}, \quad (8)$$

that is we consider a second order approximation of the phase. The following section studies the new approximation of the STFT induced by replacing h by $\tilde{h}_b$. Then, closed-form expression and reconstruction formulae are derived both for the Gabor transform.

3.1. Second order approximation of the phase and STFT

The linear chirp approximation leads to

$$\begin{aligned} V_{\tilde{h}_b}(b,\eta) &= \int_{\mathbb{R}} \tilde{h}_b(t) g(t-b) e^{-2i\pi\eta(t-b)}\, dt \\ &= h(b) \int_{\mathbb{R}} g(t) e^{i\pi\phi''(b)t^2} e^{-2i\pi(\eta-\phi'(b))t}\, dt \\ &= h(b)\widehat{gc_b}(\eta-\phi'(b)), \qquad (9) \\ &\quad \text{where } c_b(t) = e^{i\pi\phi''(b)t^2}. \end{aligned}$$

Remark: this is similar to equation (4), except that the window g is modulated by the pure linear chirp c_b.

3.2. A closed-form expression for the Gabor transform

To study the Gabor transform, we need to compute the Fourier transform of a Gaussian modulated linear chirp.

Proposition 3.1. *Consider the function $u(t) = e^{-\pi z t^2}$, where $z = re^{i\theta}$ with $\cos\theta > 0$, so that the function is integrable. Then its Fourier transform is*

$$\hat{u}(\xi) = r^{-\frac{1}{2}} e^{-i\frac{\theta}{2}} e^{-\frac{\pi}{re^{i\theta}}\xi^2}. \qquad (10)$$

Proof. One can proceed as in the case when z is real: it suffices to differentiate u and consider the Fourier transform of the obtained differential equation (see Appendix A of [10] for instance). □

Theorem 3.1. *The magnitude of the Gabor transform of $\tilde{h}_b$ admits the following closed-form expression:*

$$|V_{\tilde{h}_b}(b,\eta)| = h(b)\sigma^{\frac{1}{2}}(1+\sigma^4\phi''(b)^2)^{-\frac{1}{4}} e^{-\frac{\pi\sigma^2(\eta-\phi'(b))^2}{1+\sigma^4\phi''(b)^2}}. \qquad (11)$$

This shows that the magnitude of the Gabor transform of a linear chirp is a Gaussian function centered in $\eta = \phi'(b)$. The difference with equation (4) lies in the magnitude and the width of this Gaussian.

Proof. Equation (9) and Proposition 3.1 give

$$V_{\tilde{h}_b}(b,\eta) = h(b)\sigma^{-\frac{1}{2}} r^{-\frac{1}{2}} e^{-i\frac{\theta}{2}} e^{-\frac{\pi}{r}e^{-i\theta}(\eta-\phi'(b))^2}, \qquad (12)$$

with $r = (\frac{1}{\sigma^4}+\phi''(b)^2)^{\frac{1}{2}}$ and $\theta = \arctan(-\phi''(b)\sigma^2)$. Using the identity $\cos\arctan x = \frac{1}{\sqrt{1+x^2}}$, one finally gets (11). □

3.3. Reconstruction formulae based on second order approximation

Even if $\phi''(b)$ is not negligible, Theorem 3.1 shows that a modulated chirp creates a ridge centered in $\eta = \phi'(b)$. Thus, for mode reconstruction one just needs to adapt methods RR and IR taking into account equation (12). This leads to the following new ridge reconstruction method (RR2):

$$h(b) \approx V_h(b,\phi'(b))\sqrt{\sigma r}\, e^{i\frac{\theta}{2}}, \qquad (13)$$

where r and θ are those defined in equation (12).
Remark: this formula is also stated in [2], but based on a stationary phase approximation.

To design a reconstruction based on frequency integration in that new context, one needs to compute the value Δ such that $|V_{\tilde{h}_b}(b,\phi'(b)-\Delta)| = \varepsilon|V_{\tilde{h}_b}(b,\phi'(b))|$, which gives

$$\Delta = \sqrt{\frac{-log(\varepsilon)(1+\sigma^4\phi''(b)^2)}{\pi\sigma^2}}. \qquad (14)$$

Compared to the first-order approximation (7), this amounts to multiplying Δ by a factor $(1+\sigma^4\phi''(b)^2)^{\frac{1}{2}}$. This method will be denoted by IR2 hereafter.

4. NUMERICAL RESULTS

This section presents numerical experiments to validate the reconstruction algorithms based on the second-order approximation of the phase. For the sake of simplicity we only consider monocomponent signals, but the proposed techniques also apply to well separated multicomponent signals. We assume first the true values for ϕ' and ϕ'' to be known to compare the methods based on first and second order approximations of the phase, then we study the stability of RR2 and IR2 when ϕ'' is perturbed. For the sake of reproducible research, all codes used to create the figures of this paper can be downloaded from [11].

4.1. A comparative test

Let a family of signals depending on a parameter c be defined for $t \in [0,1]$ by

$$h_c(t) = e^{-10\pi(t-0.5)^2} e^{2i\pi(250t+\frac{500}{\pi^2}c\sin(\pi t))} \qquad (15)$$

We consider 1024 equi-spaced samples on $[0,1]$ and make c vary between 0 and 1, so that $\sup_t|\phi''(t)|$ varies between 0 and 500. Note that when $c \neq 0$, h_c is not a pure wave, but can well be approximated by a linear chirp.

We now test the four methods described before on the signal h_c, by displaying the SNR corresponding to the reconstruction of h_c as a function of c on Figure 2. We recall that the SNR of the reconstructed signal $\hat{h}$ from signal h writes

$$SNR(\hat{h}) = 20\log_{10}\frac{\|h\|_2}{\left\|\hat{h}-h\right\|_2}. \qquad (16)$$

One observes that methods IR and IR2 on one hand, and RR and RR2 on the other, give the same error when c is low. But when c increases, taking ϕ'' into account greatly improves the results.

4.2. Parameter ε and noise

We then aim to emphasize the specifics of method IR2, namely the influence of parameter ε. It is clear that in the

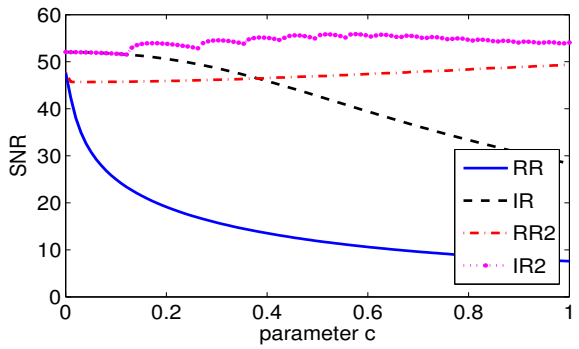

Fig. 2. SNR associated with the reconstruction of h_c as a function of c, for each method, the parameter ε being set to 0.01.

noise-free case, the results are all the better when ε is low. However, for real signals contaminated by noise this is questionable. As an illustration of this, we display on Figure 3 the SNR after reconstruction (see equation (16)) for signal $h_{0.7}$, as a function of ε for four different noise levels and for method IR2. We consider Gaussian white noise, and the noise level is characterized by its SNR, computed from equation (16) by replacing $\hat{h}$ by the noisy signal. For each noise level, the SNR after reconstruction increases slowly with ε until an optimal value is reached, and then decreases rapidly. This optimal value is strongly related to the noise level.

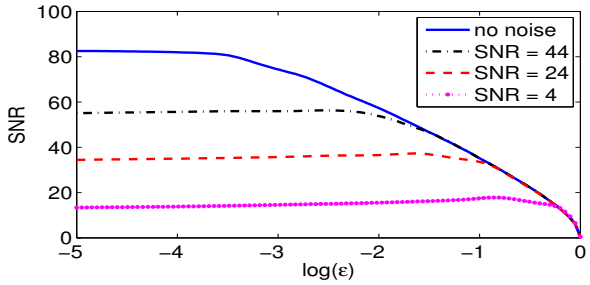

Fig. 3. SNR of $h_{0.7}$ as a function of ε, for the method IR2. Different noise levels are considered with SNRs 44, 24 and 4.

4.3. A comparison between IR2 and RR2

When dealing with real signals, neither $\phi'(b)$ nor $\phi''(b)$ are perfectly estimated, which impacts on each reconstruction method. However, as $\phi''(b)$ only appears in the estimation of the integration support, IR2 method appears to be less sensitive to these estimates. To compare both methods, Figure 4 (A) displays the reconstruction of $h_{0.7}$ as a function of the noise level (measured by its SNR) when $\phi''(t)$ is assumed to be known. Methods RR2 and IR2 are investigated with three different values of ε. The results show that method RR2 seems a bit more competitive when the noise level is high, whereas IR2 with a low ε performs better with low noise. To illustrate the sensitivity to ϕ'' issue, we make the same computations except that $\phi''(t)$ is biased by a constant term equal to 12 (see Figure 4 (B)). This simple example shows that when ϕ'' is not correctly estimated, RR2 reconstruction is strongly affected but not IR2.

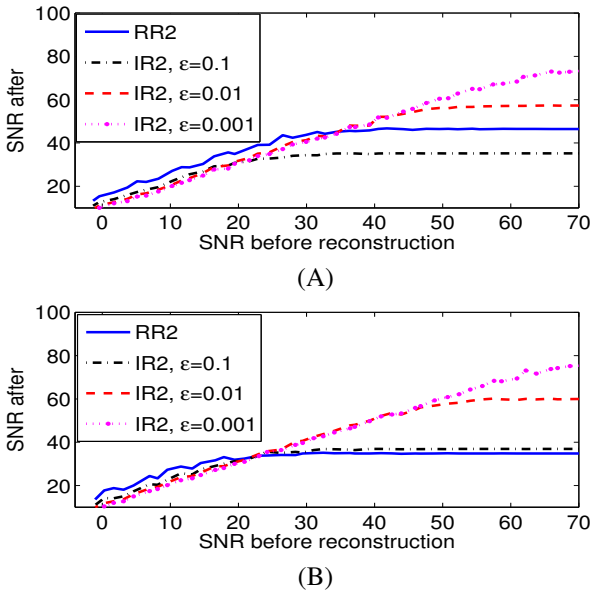

Fig. 4. SNR of $h_{0.7}$ as a function of the noise level (i.e., SNR before denoising) for methods RR2 and IR2 with different values of ε, (A): one uses the true value for ϕ'', (B): one uses a perturbed estimate.

5. CONCLUSION

This paper designed new methods for the analysis of strongly modulated multicomponent signals with the STFT. It showed that simple closed-form expressions can be derived for the Gabor transform of linear chirps, leading to improved reconstructions methods. Finally, it compared numerically ridge reconstructions with the integral reconstructions on a quite simple though informative signal, showing that the latter is more sensitive to the ridge detection.

6. RELATION TO PRIOR WORK

This paper compared classical ridge reconstruction (RR) with the integral (IR) approach. It showed that a simple 2^{nd}-order expansion of the phase leads to more efficients formulations when the frequency modulation is high. Whereas RR2 was already used in [2] (but derived through the stationary phase approximation), the method IR2 is new. It is also the first time that those two methods are compared when applied to exemplar signals.

7. REFERENCES

[1] R.A. Carmona, W.L. Hwang, and B. Torrésani, "Multiridge detection and time-frequency reconstruction," *Signal Processing, IEEE Transactions on*, vol. 47, no. 2, pp. 480–492, 1999.

[2] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrésani, "Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies," *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 644–664, 1992.

[3] G. Thakur and H.T. Wu, "Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples," *SIAM Journal on Mathematical Analysis*, vol. 43, no. 5, pp. 2078–2095, 2011.

[4] 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.

[5] S. Meignen, T. Oberlin, and S. McLaughlin, "A new algorithm for multicomponent signals analysis based on synchrosqueezing: With an application to signal sampling and denoising," *Signal Processing, IEEE Transactions on*, vol. 60, no. 11, pp. 5787 –5798, nov. 2012.

[6] S. Meignen, T. Oberlin, and S. McLaughlin, "Multicomponent signal denoising with synchrosqueezing," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, aug. 2012, pp. 660 –663.

[7] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1999.

[8] M.I. Skolnik, *Introduction to Radar System*, McGraw-Hill Education, 2003.

[9] Y. Kopsinis, E. Aboutanios, D.A. Waters, and S. McLaughlin, "Time-frequency and advanced frequency estimation techniques for the investigation of bat calls," *Journal of the Acoustical Society of America*, pp. 1124–1134, 2010.

[10] D.W. Kammler, *A first course in Fourier analysis*, Cambridge University Press, 2008.

[11] T. Oberlin, *Matlab Code*, 2012, `http://www-ljk.imag.fr/membres/Thomas.Oberlin/codeIC13.tar.gz`.