FULLY ADAPTIVE MODE DECOMPOSITION FROM TIME-FREQUENCY RIDGES

*S. Meignen *, T. Oberlin †, and S. McLaughlin †,*

* Laboratoire Jean Kuntzmann, Grenoble, France
† INP–ENSEEIHT and IRIT, University of Toulouse, Toulouse, France
‡ Heriot Watt University, Edinburgh, UK.

ABSTRACT

In this paper, we consider ridge detection for multicomponent signal analysis. We introduce a new ridge detector based on a projection of the reassignment vector in a specific direction which is related to the geometry of the spectrogram magnitude. The ridge definition we introduce enables that of the basin of attraction associated with a ridge and then mode reconstruction. Simulations show better concentration of the information on the ridges obtained by our method compared to other existing ridge detectors that also make use of the reassignment vector.

Index Terms— multicomponent signals; short-time Fourier transform; reassignment; time-frequency; AM/FM; ridges

1. INTRODUCTION

The analysis of multicomponent signals has been at the heart of signal processing research for over 60 years. A challenge that has faced the signal processing community, and for which many approaches have been developed, is that of dealing with signals with multiple AM-FM components [1, 2]. Time-frequency (TF) analysis is central in the analysis of such signals, and many techniques have been developed within that framework: e.g. [3], the *synchrosqueezing transform* (SST) [4, 5], which enhances the TF representation associated either with the continuous wavelet transform (CWT) or the short-time Fourier transform (STFT), while enabling mode reconstruction. One key issue associated with the SST method is related to ridge estimation, since an estimate of the ridge associated with each mode prior to mode reconstruction is required. Many approaches have been proposed that take this approach, e.g. [6, 7] and, once the ridges are known, alternative techniques to SST have been developed for mode reconstruction by integrating the STFT coefficients in the vicinity of the detected ridges [8, 9].

In this paper, we are interested in ridge detection using the properties of the *reassignment vector* (RV), and then mode reconstruction based on the *basins of attraction* associated with the detected ridges. This approach to mode reconstruction was first proposed in [10] and then improved in [11, 12], but some difficulties remain when trying to assess a ridge associated wih a Dirac distribution, which can arise when the signal under consideration is discontinuous in time.

The goal of the paper is to propose a new ridge detector, based on the property of the reassignment vector, but which allows for AM-FM mode reconstruction in a fully adaptive way, whatever the types of the modes, including impulses or discontinuities. The paper is structured as follows, first we introduce basic definitions, followed by a short review of the ideas underlying the reassignment vector (RV). Then, we recall the principle of ridge detection based on RV and develop our new technique, then we assess the improvement brought by our new ridge detector on different type of signals finally drawing some conclusions.

2. DEFINITIONS

In this section, we provide some basic definitions which will be useful in the sequel. For a given signal $f \in L^1(\mathbb{R})$, its (modified) *STFT* is defined by

$$V_f^g(t,\omega) = \int_{\mathbb{R}} f(u)g(u-t)e^{-i2\pi\omega(u-t)}\,du, \quad (1)$$

where the window g is assumed to be real-valued. The spectrogram is then defined as $|V_f^g(t,\omega)|^2$. In the following, we will study more in details multicomponent signals f defined by:

$$f(t) = \sum_{k=1}^{K} f_k(t), \text{ with } f_k(t) = a_k(t)e^{i2\pi\phi_k(t)}. \quad (2)$$

3. REASSIGNMENT OF THE SPECTROGRAM

The principle of the *reassignment method* (RM) [13] is to compensate for the TF shifts induced by the 2D smoothing involved in defining the spectrogram. To do so, a meaningful TF location is first determined to which to assign the local energy given by the spectrogram. This corresponds to

The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-13-BS03-0002-01 (ASTRES).

the *centroid* of the distribution, whose coordinates are defined by $\hat{\tau}_f(t,\omega) := -\partial_\omega \arg V_f^g(t,\omega)$, and $\hat{\omega}_f(t,\omega) := \omega + \frac{1}{2\pi}\partial_t \arg V_f^g(t,\omega)$ [13]. Both quantities *locally* define an instantaneous frequency (IF) and a group delay (GD) and enable perfect localization of linear chirps [13]. An efficient procedure to compute them is:

$$\hat{\tau}_f(t,\omega) = t + \Re\left\{\frac{V_f^{tg}(t,\omega)}{V_f^g(t,\omega)}\right\},$$
$$\hat{\omega}_f(t,\omega) = \omega - \frac{1}{2\pi}\Im\left\{\frac{V_f^{g'}(t,\omega)}{V_f^g(t,\omega)}\right\}, \quad (3)$$

where tg stands for the function $tg(t)$ and $\Re\{Z\}$ (resp. $\Im\{Z\}$) is the real (resp. imaginary) part of the complex number Z. In that context, the reassignment vector (RV) is defined by:

$$RV(t,\omega) = \begin{pmatrix} \hat{\tau}_f(t,\omega) - t \\ \hat{\omega}_f(t,\omega) - \omega \end{pmatrix}. \quad (4)$$

When $g(t) = e^{-\pi t^2}$, $g'(t) = -2\pi t g(t)$ and the reassignment vector reads $RV(t,\omega) = \left(\Re\left\{\frac{V_f^{tg}(t,\omega)}{V_f^g(t,\omega)}\right\}, \Im\left\{\frac{V_f^{tg}(t,\omega)}{V_f^g(t,\omega)}\right\}\right)$. With such a window, $\nabla \log|V_f^g(t,\omega)|$ can be written in terms of RV using the following identities:

$$\partial_t V_f^g(t,\omega) = -V_f^{g'}(t,\omega) + 2i\pi\omega V_f^g(t,\omega) \quad (5)$$
$$\partial_\omega V_f^g(t,\omega) = -2i\pi V_f^{tg}. \quad (6)$$

Thus, we immediately obtain:

$$\nabla \log|V_f^g(t,\omega)| = \left(\frac{\partial_t|V_f^g(t,\omega)|}{|V_f^g(t,\omega)|}, \frac{\partial_\omega|V_f^g(t,\omega)|}{|V_f^g(t,\omega)|}\right)$$
$$= \left(-\Re\left\{\frac{V_f^{g'}(t,\omega)}{V_f^g(t,\omega)}\right\}, 2\pi\Im\left\{\frac{V_f^{tg}(t,\omega)}{V_f^g(t,\omega)}\right\}\right) = 2\pi RV(t,\omega)$$

In what follows, we use the notation $(L_t(t,\omega), L_\omega(t,\omega)) := RV(t,\omega)$. When the window $g(t) = e^{-a\pi t^2}$ is used instead, we may write: $\nabla \log|V_f^g(t,\omega)| = 2\pi\left(aL_t(t,\omega), \frac{1}{a}L_\omega(t,\omega)\right)$, meaning $\nabla \log|V_f^g(t,\omega)|$ can still be accessed via the appropriate renormalizations of the reassignment vector field.

4. DEFINITIONS OF CONTOUR POINTS

In this section, we investigate two different definitions of contour points. Detecting ridge points and linking them to build smooth contours is a challenging problem that has previously been considered in both STFT and wavelet settings [6, 11]. As we will see later, the proposed approach consists of projecting the RV, possibly renormalized, in a specific direction. Our aim is indeed to detect zero-crosssings of a signed function, which we can do in practice with the contourc Matlab function, the difficulty being to find a appropriate direction of projection.

4.1. Ridge definition based on reassignment vector projection

In the technique introduced in [10, 14] to define contours in the TF plane, points on a contour correspond to locations where the RV changes orientation rapidly, which happens when the contours, corresponding to the IF of a mode, is crossed. However, determining where those crossings are in a discrete setting is problematic, therefore it is preferable to project the RV in a specific direction, given by an angle θ, and then determine the location of the change of sign of the projection. These points, called *contour points* (CPs), are defined as the zeros of the following quantity

$$\langle \nabla \log|V_f^g(t,\omega)|, v_{\frac{\pi}{2}+\theta}\rangle, \quad (7)$$

where v_λ is the unit vector in the direction λ. This approach requires the direction θ to be known a priori and is therefore not well suited to determining CPs with varying orientations. Instead of imposing an orientation θ an alternative is to compute CPs as in [15]:

$$\alpha(t,\omega) := \langle \nabla \log|V_f^g(t,\omega)|, v_{\theta(t,\omega) \mod \pi}\rangle = 0 \quad (8)$$

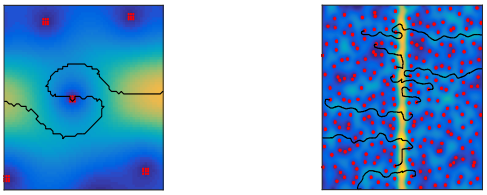
with $\theta(t,\omega)$ the argument of $\nabla \log|V_f^g(t,\omega)|$ and $(\theta(t,\omega) \mod \pi) \in [0,\pi[$. This approach was further studied in [11] and compared to a technique based on a study of the zeros of the spectrogram proposed in [16]. The rationale behind this formula is that $\alpha(t,\omega)$ corresponds to the signed magnitude of the (renormalized) RV. More precisely, it can be checked that $\alpha(t,\omega)$ is negative above a ridge (with finite slope) and positive below. This way, a new type of CPs is defined that is no longer dependent on a fixed angle θ.

However, this technique, called M_1 in the sequel, has several drawbacks, which are related. First, a zero of the spectrogram is a repulsive point for the vector field $\nabla \log|V_f^g(t,\omega)|$ [11], and the mod π computation induces $\alpha(t,\omega)$ to be zero on horizontal TF lines crossing that zero. This creates special structures in the vicinity of zeros as shown on Figure 1 A. The second limitation is that the technique fails to detect vertical ridges, because of the mod π factor. To illustrate this, we consider the STFT of a noisy Dirac and see how the contour detector works. The results depicted in Figure 1 B show that, the mod π creates artificial sign changes in $\alpha(t,\omega)$ preventing the detection of the actual contour.

4.2. Determination of ridge points based on differential geometry

Alternatively, the estimation of ridges and valleys of an image is an old and well-known problem in computer vision, for which a nice answer has been proposed [17] using the Hessian of the smoothed image, within the scale-space theory. Indeed, the quantity $\log(|V_f^g(t,\omega)|)$ can be viewed as an image, whose gradient is equal to:

$$G(t,\omega) = 2\pi(aL_t(t,\omega), \frac{1}{a}L_\omega(t,\omega)), \quad (9)$$



A B

Fig. 1. A: an illustration of the behaviour of a contour, computed with method M_1, in the vicinity of a zero of the spectrogram;B: STFT of a noisy Dirac distribution (SNR = 0 dB) on which the first 10 contours, according to their energy content and computed with method M_1, are superimposed.

and its Hessian to:

$$H(t,\omega) = \begin{pmatrix} a\partial_t L_t(t,\omega) & a\partial_\omega L_t(t,\omega) \\ \frac{1}{a}\partial_t L_\omega(t,\omega) & \frac{1}{a}\partial_\omega L_\omega(t,\omega) \end{pmatrix} = \begin{pmatrix} H_{tt} & H_{t\omega} \\ H_{\omega t} & H_{\omega\omega} \end{pmatrix}, \quad (10)$$

where, omitting the variable (t,ω):

$$\partial_t L_t = \Re\left\{-1+\frac{V^{tg}V_f^{g'}-V_f^{tg'}V_f^g}{(V_f^g)^2}\right\}, \partial_t L_\omega = 2\pi\Im\left\{\frac{V_f^{t^2g}V_f^g-(V_f^{tg})^2}{(V_f^g)^2}\right\}$$

$$\partial_\omega L_t = \frac{1}{2\pi}\Im\left\{\frac{V_f^{g''}V_f^g-(V_f^{g'})^2}{(V_f^g)^2}\right\}, \partial_\omega L_\omega = \Re\left\{\frac{V_f^{tg'}V_f^g-V_f^{tg}V_f^{g'}}{(V_f^g)^2}\right\}.$$

Taking into account the relation between g and its derivative, the computation of the Hessian matrix can be carried out using only one more STFT as those already used to compute L_t and L_ω. Since H is diagonal in an orthonormal real basis of eigenvectors, we denote by λ_p and λ_q, the eigenvalues of H and p and q the corresponding unitary eigenvectors, i.e. $p=(\cos(\beta),\sin(\beta))$ with

$$\cos(\beta) = \sqrt{\frac{1}{2}\left(1+\frac{H_{tt}-H_{\omega\omega}}{\sqrt{(H_{tt}-H_{\omega\omega})^2+4H_{t\omega}H_{\omega t}}}\right)}$$

$$\sin(\beta) = sign(H_{t\omega})\sqrt{\frac{1}{2}\left(1-\frac{H_{tt}-H_{\omega\omega}}{\sqrt{(H_{tt}-L_{\omega\omega})^2+4H_{t\omega}H_{\omega t}}}\right)}, \quad (11)$$

and $q=(\sin(\beta),-\cos(\beta))^T$. In that framework, $\lambda_p = p^THp$ and $\lambda_q = q^THq$. Then, we define $L_p(t,\omega) = \langle G(t,\omega),p\rangle$ (L_q being defined the same way from q), and compute the ridge and valley points (corresponding to two different types of CPs) as follows:

Ridges:$L_p(t,\omega)=0$ and $\lambda_p(t,\omega)<0$, $|\lambda_p(t,\omega)|\geq|\lambda_q(t,\omega)|$
Valleys:$L_p(t,\omega)=0$ and $\lambda_p(t,\omega)>0$, $|\lambda_p(t,\omega)|\leq|\lambda_q(t,\omega)|$.

This new formulation appears to get rid of the problem arising from the use of the mod π factor. However, there is still one thing that can give rise to instability: when $H_{t\omega}$ is too small, which is typically the case when a purely harmonic mode or a Dirac is considered, $sign(H_{t\omega})$ may change spuriously. Note that this change does not affect these two types of signals in the same manner: if $H_{t\omega}$ changes signs in the vicinity of an horizontal ridge (constant frequency), we get $\beta=0$ (since $\sin(\beta)=0$), and L_p corresponds to the projection in the direction $\frac{\pi}{2}$, while in case of a vertical ridge (i.e. a Dirac impulse) $\beta=\pm\frac{\pi}{2}$ (since $\sin(\beta)=\pm1$), and L_p corresponds to a projection either in the direction 0 or π.

So when considering this model the problem of instability with a Dirac pulses remains, but we note that the ridge detection is stable if β is far from $\frac{\pi}{2}$ (i.e $\cos(\beta)$ not too small). We therefore change the direction of projection when $\cos(\beta)$ becomes too low by using the following new estimate for the eigenvectors:

Algorithm 1

p defined with (11), if $\cos(\beta)<\gamma$ else by

$$\begin{cases} \cos(\beta) = \sqrt{\frac{1}{2}\left(1-\frac{H_{tt}-H_{\omega\omega}}{\sqrt{(H_{tt}-L_{\omega\omega})^2+4H_{t\omega}H_{\omega t}}}\right)}, \\ \sin(\beta) = sign(H_{t\omega})\sqrt{\frac{1}{2}\left(1+\frac{H_{tt}-H_{\omega\omega}}{\sqrt{(H_{tt}-H_{\omega\omega})^2+4H_{t\omega}H_{\omega t}}}\right)} \\ \beta = \text{atan}(\sin(\beta)/\cos(\beta)) \end{cases}$$

This method is denoted M_2 in the sequel. Looking at the above description, we see that p is replaced by a vector orthogonal to it only if $\beta=\frac{\pi}{2}$ (the absolute value between the original vector and the new one equals $2|\cos(\beta)||\sin(\beta)|$). Figure 2 A displays the STFT of the same noisy Dirac as the one in Figure 1 B, again with the first 10 contours, with respect to their energy content and computed with method M_2, superimposed. This illustrates the benefits of using method M_2 rather than M_1 for that particular type of signals.

4.3. Determination of basins of attraction using RV and mode reconstruction

Having determined the ridges associated with the modes making up the signal, we define the *basin of attraction* (BA) associated with a ridge, i.e. the set of coefficients associated with a given contour, as in [11]. Since the RV points to a ridge in its vicinity, we determine the BA of a given ridge as the set of points such that the RV points to that ridge. However, because the localization property of RV is only valid for linear chirps, and also because of the presence of noise, RV does not point exactly to a ridge. Therefore, it is proposed in [11] to associate with a given coefficient (t,ω) the closest ridge of point $(\hat{\tau}_f(t,\omega),\hat{\omega}_f(t,\omega))$. The BA corresponding to the most energetic contour of the STFT of a noisy Dirac is given in Figure 2 B. Once the BAs are computed, the corresponding modes can be retrieved as follows. Let $\mathcal{B}_i\subset\mathbb{R}^2$ be the BA associated with ridge i, then a local reconstruction technique

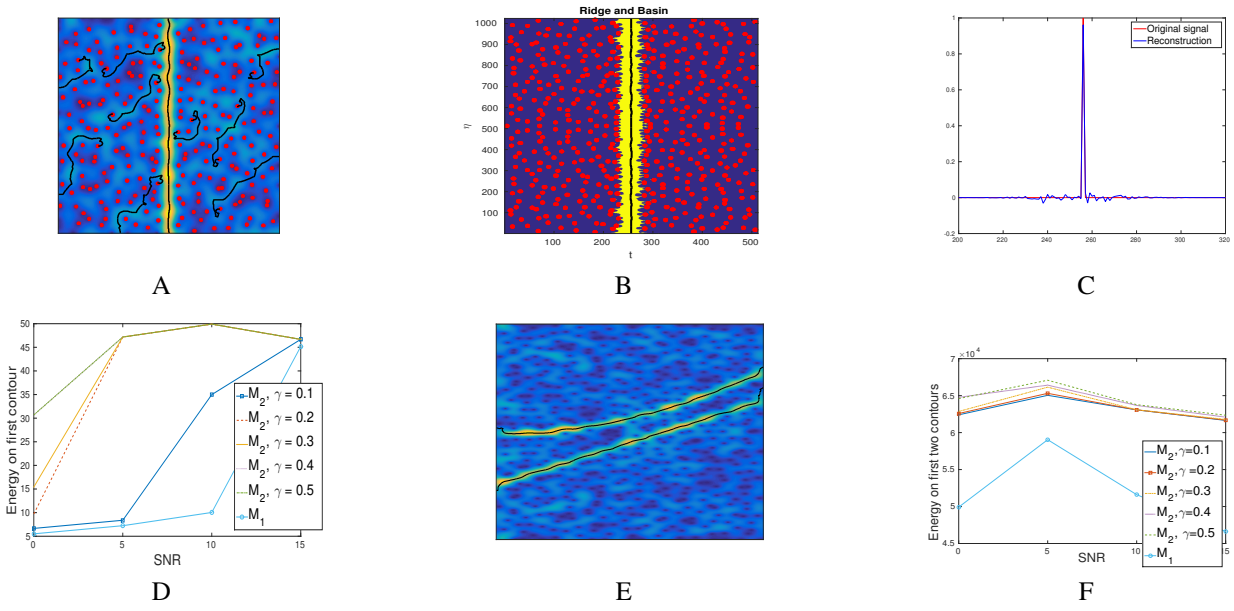


Fig. 2. A: Contours (the 10 most energetic ones) associated with a noisy Dirac using method M_2(SNR = 0 dB); B: Basin of attraction (in yellow) associated with the most energetic contour comuted on a noisy Dirac; C: Reconstructed signal based on the coefficients contained in the basin of attraction depicted in B along with the original Dirac distribution; D: Energy contained on the first contour for noisy Diracs (SNR in abscissa) and for the different methods (M_1 and M_2 with different values for γ); E: noisy STFT of linear and a polynomial chirp with the first two contours superimposed (SNR = 0 dB, $\gamma = 0.1$); F: Energy contained on the first two contours computed on E, for methods M_1 and M_2, for different SNRs

of mode f_i corresponding to ridge i can be achieved by:

$$f_i(t) = \frac{1}{g(0)} \int_{(t,\omega)\in\mathcal{B}_i} V_f^g(t,\omega)d\omega. \tag{12}$$

An illustration of the reconstruction procedure is given in Figure 2 C, where only the BA computed in Figure 2 B is used for mode reconstruction (in blue), along with the original signal (in red).

5. NUMERICAL RESULTS

In this section, we purposefully focus only on assessing the quality of the new ridge detector, which is the main novelty of the paper, depending on parameter γ (defined in Algorithm 1), the only one to be fixed in method M_2. We investigate the quality of the ridge estimation for different types of mode and when the noise level varies. To measure this, we consider the energy contained in the first K (the number of modes) contour as γ varies:

$$P(\gamma) = \sum_{i=1}^{K} \sum_{(t,\omega)\in\mathcal{C}_i^\gamma} |V_f^g(t,\omega)|^2, \tag{13}$$

where $\mathcal{C}_i^\gamma$ is the ith contour, according to its energy content, computed with method M_2 with parameter γ. We remark, that when M_2 is applied to a noisy Dirac (see Figure 2 D) the energy is better concentrated on the first contour, when the noise increases, when γ is chosen larger: the noise results in some instabilities in the direction of p, which varies around $\pi/2$, and, to choose a large enough γ enables to compensate for that. As far as the signal whose STFT is depicted in Figure 2 E, as expected, the sensitivity to γ is very low (see Figure 2 F), since the angle p with the horizontal axis is lower than $\frac{\pi}{3}$ (corresponding to $\gamma = 0.5$), and so, in Algorithm 1, one stays in the first case. Finally, it is worth noting that the proposed method M_2 always behaves better than M_1.

6. CONCLUSION

In this paper, we have presented a new algorithm to estimate structures called contours, associated with the STFT of multicomponent signals. This algorithm is based on the study of the geometry of STFT magnitude. We then defined basins of attraction associated with the contours ; these can be used to reconstruct the modes making up the signal. Numerical experiments show a better concentration of the STFT on the contours determined using this new technique rather than other methods based on the projection of reassignment vector.

REFERENCES

[1] K. Kodera, R. Gendrin, and C. Villedary, "Analysis of time-varying signals with small BT values," *IEEE*

Trans. Acoust., Speech and Sig. Proc., vol. 26, no. 1, pp. 64–76, 1978.

[2] P. Flandrin, *Time-frequency / Time-scale Analysis*, vol. 10, Academic press, 1998.

[3] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 520–538, 1992.

[4] I. Daubechies, J. Lu, and H.T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis*, 2010.

[5] T. Oberlin, S. Meignen, and 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.

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

[7] D. E. Newland, "Ridge and phase identification in the frequency analysis of transient signals by harmonic wavelets," *Journal of Vibration and Acoustics*, vol. 121, no. 2, pp. 149–155, 1999.

[8] 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, 2012.

[9] T. Oberlin, S. Meignen, and S. Mclaughlin, "A novel time-frequency technique for multicomponent signal denoising," in *Proceedings of the 21st European Signal Processing Conference (EUSIPCO-13)*, 2013, pp. 1–5.

[10] Y. Lim, B.G. Shinn-Cunningham, and T.J. Gardner, "Sparse contour representations of sound," *IEEE Sig. Proc. Let.*, vol. 19, no. 10, pp. 684–687, 2012.

[11] S. Meignen, T. Oberlin, P. Depalle, P. Flandrin, and S. McLaughlin, "Adaptive multimode signal reconstruction from time–frequency representations," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, 2016.

[12] S. Meignen, T. Gardner, and T. Oberlin, "Time-frequency ridge analysis based on reassignment vector," in *Proceedings of the 23st European Signal Processing Conference (EUSIPCO-15)*. EURASIP, 2015.

[13] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *Signal Processing, IEEE Transactions on*, vol. 43, no. 5, pp. 1068–1089, 1995.

[14] T. J. Gardner and M. O. Magnasco, "Sparse time-frequency representations," *Proc. Nat. Acad. Sci.*, vol. 103, no. 16, pp. 6094–6099, 2006.

[15] S. Meignen, T. Oberlin, and T. Gardner, "Time-frequency ridge analysis based on reassignment vector," in *Signal Processing Conference (EUSIPCO), 2015 Proceedings of the 21st European*. IEEE, 2015, pp. 1–5.

[16] P. Flandrin, "Time-frequency filtering based on spectrogram zeros," *IEEE Sig. Proc. Lett.*, vol. 22, no. 11, pp. 2137–2141, 2015.

[17] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 117–156, 1998.