Sparse Time-Frequency-Frequency-Rate Representation for Multicomponent Nonstationary Signal Analysis

1st Wenpeng Zhang
College of Electronic Science
National University of Defense Technology
Changsha, China
zhangwenpeng@hotmail.com

2nd Yaowen Fu
College of Electronic Science
National University of Defense Technology
Changsha, China
fuyaowen@nudt.edu.cn

3rd Yuanyuan Li
College of Electronic Science
National University of Defense Technology
Changsha, China
liyuanyuan16@nudt.edu.cn

***Abstract*—Though high resolution time-frequency representations (TFRs) are developed and provide satisfactory results for multicomponent nonstationary signals, extracting multiple ridges from the time-frequency (TF) plot to approximate the instantaneous frequencies (IFs) for intersected components is quite difficult. In this work, the sparse time-frequency-frequency-rate representation (STFFRR) is proposed by using the short-time sparse representation (STSR) with the chirp dictionary. The instantaneous frequency rate (IFRs) and IFs of signal components can be jointly estimated via the STFFRR. As there are permutations between the IF and IFR estimates of signal components at different instants, the local k-means clustering algorithm is applied for component linking. By employing the STFFRR, the intersected components in TF plot can be well separated and robust IF estimation can be obtained. Numerical results validate the effectiveness of the proposed method.**

***Keywords*—multicomponent nonstationary signal, time-frequency-frequency-rate representation; short-time sparse representation; instantaneous frequency estimation; local k-means clustering algorithm**

I. Introduction

Nonstationary signal analysis has received extensive attention in the fields of speech processing, biomedical applications, radar, telecommunications, etc [1-2]. For example, the echo-location pulse emitted by the brown bat is a typical multicomponent nonstationary signal, which contains multiple chirps [3]. Generally, the instantaneous frequency (IF) or instantaneous frequency rate (IFR) of the signal is used for nonstationary signal characterization.

Time-frequency representations (TFRs) which can jointly provide time and frequency information are widely employed for IF estimation [1-2]. The short-time Fourier transform (STFT) is developed by performing Fourier transform on the local patches of the signal. The STFT is proved to be useful, but its resolution is limited by the window length. The Wigner-Ville distribution (WVD) possesses optimal TF resolution for a chirp signal. However, the application of WVD is restricted as it suffers from the auto-term inference for nonlinear frequency modulated signals and the cross-term inference for multicomponent signals. Later, the modifications of WVD, e.g., the polynomial Wigner-Ville distribution and the L-Wigner distribution, are developed to reduce the unwanted inferences and to get a high TF resolution [1-2]. The chirp-based TFR which represents the TFR by the sum of Gaussian chirplet components is free of inferences as well [3-5].

In addition to TFRs, time-frequency rate representations (TFRRs) are proposed for instantaneous frequency rate (IFR) estimation. In [6], the cubic phase function (CPF) is proposed and it is proved that the energy of the CPF is concentrated along the FR law of the signal for polynomial phase signals (PPSs) with terms less than third order. Similar to the WVD, as a bilinear transform, the CPF suffers from the cross term as well. To suppress the cross term, the smoothed high-resolution time-frequency rate representation (SHR-TFRR) is proposed by introducing an FR window [7].

Neither the TFR nor the TFRR can provide joint IF and IFR estimation simultaneously. Take the TFRR for example, the IF of a signal is re-calculated by the integral of the IFR estimate [8]. To the best of our knowledge, the cross quadratic spectrum (XQS) is the sole time-frequency-frequency rate representation (TFFRR) which provides accurate characterization for both IF and IFR [9]. The XQS is constructed by performing the quadratic Fourier transform (QFT) on the forward and backward segments followed by averaging over the products of the forward and backward QFTs.

After obtaining the representation of a signal in the TF or the TFR plot, IF or IFR estimation can be achieved by extracting the ridges. For the monocomponent signal, either the component with the maximum energy or the first conditional moment at each time instant is capable of providing IF estimates [1]. When it comes to the multicomponent signal, the Viterbi algorithm [8, 10] with defined path penalty function is proposed to obtain high accurate IF and IFR estimates in a high noise environment. However, this method fails when there are intersections between different signal components.

In this work, the sparse time-frequency-frequency rate representation (STFFRR) is proposed for multicomponent nonstationary analysis. The STFFRR is constructed through the short-time sparse representation (STSR) with the chirp dictionary, which is more accurate than the STFT and can simultaneously provide IF and IFR characterization. Furthermore, the employment of sparse representation (SR) allows direct joint IF and IFR estimation instead of

ISBN 978-90-827970-1-5

performing ridge extracting approaches, which is the main difference from the existing dense representations (e.g., XQS).

The signal components appear as separated and non-intersected curves in the TFFR plot and there are permutations between their IF and IFR estimates at different instants. The local k-means clustering algorithm is then applied for component linking and the permutations between components are resolved.

The rest of this paper is organized as follows. In section II, the signal model and the proposed method is proposed. A multicomponent nonstationary signal is employed to evaluate the performance of the proposed method in section III. Conclusions are provided in Section IV.

II. The Sparse Time-Frequency-Frequency-Rate Representation

A. The Multicomponent Nonstationary Signal

The following expression gives an analytic multicomponent nonstationary signal,

$$x(t)=\sum_{i=1}^{L} a_i(t)\exp\left(j\varphi_i(t)\right) \tag{1}$$

where $x_i(t)=a_i(t)\exp\left(j\varphi_i(t)\right)$ is the ith component, $a_i(t)$ and $\varphi_i(t)$ are the instantaneous amplitude (IA) and instantaneous phase (IP) of $x_i(t)$, respectively. The instantaneous frequency (IF) of $x_i(t)$ is defined by the first derivative of the IP,

$$f_i(t)=\frac{1}{2\pi}\frac{d\varphi_i(t)}{dt} \tag{2}$$

And the instantaneous frequency rate (IFR) is defined by the second derivative of the IP,

$$\Omega_i(t)=\frac{1}{2\pi}\frac{d^2\varphi_i(t)}{dt^2} \tag{3}$$

The purpose of multicomponent nonstationary signal analysis is to obtain a representation (TFR or TFRR) in which the energy is concentrated along the IFs or IFRs of each components. Then further processing can be carried out based on the representation or the IF and IFR estimates.

In this work, we only consider multicomponent nonstationary signals with continuous IFs and IFRs, i.e., $f_i(t)\in C^1[0,T]$, $\Omega_i(t)\in C[0,T]$, where T is the signal duration. For $i\neq j$, $\left(f_i(t),\ \Omega_i(t)\right)=\left(f_j(t),\ \Omega_j(t)\right)$ requires that these two components to be tangent in the TF plot (which is a more strict condition compared with the intersection in the TF plot). Thus the probability of intersection in TFFR plot decreases and components appear as non-intersected curves in the TFFR space for most of the time.

B. The TFFRR of Three Typical Signals

In this Subsection, the IFs and IFRs of the sinusoidal signal (SS), the chirp signal (CS), and the sinusoidal frequency modulated signal (SFMS) are derived, respectively. Further insight of the characteristics of the TFFRR can be obtained from these three signals.

The SS signal is given as

$$x_a(t)=A_a\exp\{j2\pi f_a t\} \tag{4}$$

The IF and IFR of $x_a(t)$ can be expressed as

$$f_a(t)=f_a \tag{5}$$

$$\Omega_a(t)=0 \tag{6}$$

Thus, the TFFRR of $x_a(t)$ is

$$c_a=(t,f_a,0) \tag{7}$$

c_a is a line perpendicular to the frequency axis in the TF plane.

The CS signal is given as

$$x_b(t)=A_b\exp\left\{j2\pi\left(f_b t+0.5\gamma_b t^2\right)\right\} \tag{8}$$

The IF, IFR, TFFRR of $x_b(t)$ can be expressed as

$$f_b(t)=f_b+\gamma_b t \tag{9}$$

$$\Omega_b(t)=\gamma_b \tag{10}$$

$$c_b=(t,f_b+\gamma_b t,\gamma_b) \tag{11}$$

c_b is a line as well.

Proposition 1: The sinusoidal signal and the chirp signal cannot be intersected in the TFFR plot.

Proof: Assume that c_a and c_b are intersected at point $(t,f_a,0)=(t,f_b+\gamma_b t,\gamma_b)$, this equation collapses into $(f_b,\gamma_b)=(f_a,0)$. Thus c_a and c_b would be the same line, which is contrary to the definition of c_a and c_b, and Proposition 1 is proved.

Proposition 2: Two chirp signals cannot be intersected in the TFFR plot.

Proof: Assume that c_b and c_b' are intersected at point $(t,f_b+\gamma_b t,\gamma_b)=(t,f_b'+\gamma_b' t,\gamma_b')$, this equation collapses into $(f_b,\gamma_b)=(f_b',\gamma_b')$. Thus c_b and c_b' would be the same line, which is contrary to the definition of c_b and c_b', and Proposition 2 is proved.

The SFMS signal is given as

$$x_c(t)=A_{c.1}\exp\left\{j2\pi A_{c.2}\sin\left(2\pi f_{c.1}t+\varphi_c\right)\right\}\exp\{j2\pi f_{c.2}t\} \tag{12}$$

$$f_c(t)=f_{c.2}+2\pi f_{c.1}A_{c.2}\cos\left(\left(2\pi f_{c.1}t+\varphi_c\right)\right) \tag{13}$$

$$\Omega_c(t)=-\left(2\pi f_{c.1}\right)^2 A_{c.2}\sin\left(\left(2\pi f_{c.1}t+\varphi_c\right)\right) \tag{14}$$

Thus, the IF and IFR of the SFMS signal are both a sinusoidal function with respect to time and their relation can be expressed by

$$\frac{\left(f_c(t)-f_{c.2}\right)^2}{\left(2\pi f_{c.1}A_{c.2}\right)^2}+\frac{\left(\Omega_c(t)\right)^2}{\left(\left(2\pi f_{c.1}\right)^2 A_{c.2}\right)^2}=1 \tag{15}$$

It can be known from (15) that $c_c=\left(t,f_c(t),\Omega_c(t)\right)$ is a helix curve.

ISBN 978-90-827970-1-5

C. The Short-Time Sparse Representation

Before giving the detail of the STSR, we briefly review the definition of the STFT. The STFT of $x(t)$ is

$$\rho_x(t,f)=\int_{-\infty}^{\infty} x(t)h(\tau-t)\exp(-j2\pi f\tau)d\tau \qquad (16)$$

where $h(t)$ is a symmetric window function centered at $t=0$. The performance of the STFT is limited if the frequency content of a signal varies fast in a short time.

To get a more accurate model for a nonstationary signal, it is preferred to perform quadratic Fourier transform on the signal patch [9, 11]. In this work, the STSR is employed and the local patch of the signal is approximated by multiple chirps. Thus, the STSR can be used for multicomponent signals and allows direct joint IF and IFR estimation.

Denote the local patch of $x(t)$ at instant t as

$$p_x(\tau;t)=x(t)h(\tau-t);\qquad t-\frac{l_w}{2}\le\tau\le t+\frac{l_w}{2} \qquad (17)$$

where l_w is the window length.

The chirp signal is defined as

$$s_k(\tau)=\exp\left\{j2\pi f_k\tau+j\pi\beta_k\tau^2\right\};\qquad -\frac{l_w}{2}\le\tau\le\frac{l_w}{2} \qquad (18)$$

where f_k, β_k are the central frequency and chirp rate, respectively.

$p_x(\tau;t)$ can be represented by chirps as follows

$$p_x(\tau;t)=\sum_{k=1}^{Q}a_k s_k(\tau)+e(\tau) \qquad (19)$$

where a_k is the complex amplitude of $s_k(\tau)$; $e(\tau)$ is the error term.

Equation (19) is the classical SR problem and can be solved by the orthogonal matching pursuit (OMP), the basis pursuit (BP), etc. The special property of chirps can be utilized to accelerate the computation and we refer to the approximate maximum likelihood estimation (MLE) method in [5] which is indeed a modification of the OMP. The parameters required to be calculated are $\{(a_k,f_k,\beta_k)\}_{k=1}^{Q}$ (These parameters has little difference with that of [5] and there are some modifications in the following introduced procedure.).

The modified OMP algorithm is detailed as follows.

1) Set $k=1$, denote the residual signal as $r_k(\tau)=p_x(\tau;t)$

2) Estimate $\hat{a}_k,\hat{f}_k,\hat{\beta}_k$ via the approximate MLE method

Estimate the chirp rate by

$$\hat{\beta}_k=\arg\max_{\beta}\int\left|S_\beta(f)\right|^4 df \qquad (20)$$

where $S_\beta(\omega)$ is the Fourier transform of the chirp rotated signal

$$S_\beta(f)=\int r_k(\tau)\exp\left\{-j\pi\beta\tau^2\right\}\exp\left\{-j2\pi f\tau\right\}d\tau \qquad (21)$$

Then frequency estimate of this chirp is given

$$\hat{f}_k=\arg\max_{f}\left|S_{\hat{\beta}_k}(f)\right| \qquad (22)$$

And the complex amplitude can be obtained by computing the normalized inner product

$$\hat{a}_k=\frac{\left\langle r_k(\tau),\hat{s}_k(\tau)\right\rangle}{\left\langle \hat{s}_k(\tau),\hat{s}_k(\tau)\right\rangle} \qquad (23)$$

where $\hat{s}_k(\tau)=\exp\left\{j2\pi\hat{f}_k\tau+j\pi\hat{\beta}_k\tau^2\right\}$.

3) Update the residual signal

$$r_{k+1}(\tau)=r_k(\tau)-\hat{a}_k\hat{s}_k(\tau) \qquad (24)$$

4) Set $k=k+1$, repeat step 2), 3) until k equals a predetermined number Q or the energy of residual signal is small then a threshold ε.

By performing the modified OMP algorithm on each signal patch $p_x(\tau;t)$, we get the STSR of $x(t)$ as

$$\Theta=\left\{\left\{\left(a_k^t,f_k^t,\beta_k^t\right)\right\}_{k=1}^{Q}\right\}_{t=0}^{T} \qquad (25)$$

where Q is the component number, T (N for the discrete version) is the time duration of $x(t)$. The component number can be obtained by a prior information or estimated by using short-term time-frequency renyi entropy [12].

D. Component Linking by Local K-means Clustering

From the aforementioned discussions, it is clear that the probability of intersection for the curves corresponding to different components are much lower in the TFFR plot compared with that in the TF plot. This property makes the linking of components much easier. One can simply assign a point at an instant in TFFR plot to the same component index of its nearest point at a previous instant. A more robust approach would be to use more points at several previous instants to make the decision. In this paper, the local k-means clustering algorithm, i.e., perform k-means clustering on the sliding window of the STSR results, is employed to achieve this.

The linking of components by local k-means clustering is detailed as follows.

Input: STSR results $\Theta=\left\{\left\{\left(a_k^n,f_k^n,\beta_k^n\right)\right\}_{k=1}^{Q}\right\}_{n=1}^{N}$, clustering window length l_s.

Output: $\Lambda=\left\{\left\{\left(a_k^n,f_k^n,\beta_k^n,y_k^n\right)\right\}_{k=1}^{Q}\right\}_{n=1}^{N}$, where y_k^n is the cluster label.

Initialize: $m=1$, initial cluster centroids are

$C_m=\left\{\left(f_1^1,\beta_1^1\right),\left(f_2^1,\beta_2^1\right),...,\left(f_Q^1,\beta_Q^1\right)\right\}$, $\Lambda=\varnothing$.

1) Gather the IF estimates and IFR estimates from instant m to instant $m+l_s-1$, denote the set as $\Theta_m=\left\{\left(a_k,f_k,\beta_k,n_k\right)\right\}_{k=mQ}^{m+Q\cdot l_s-1}$.

2) Perform k-means clustering on Θ_m with the initial cluster centroids C_m, the clustering result are

$Y_m = \{y_k\}_{k=m}^{m+Q\cdot l_s-1}$ where y_k is the cluster label of data point (f_k, β_k), and the new cluster centroids $\hat{C}_m$.

The distance metric adopted herein is the Euclidean distance of IF and IFR, which is defined as

$$d_{ij} = \sqrt{(f_i - f_j)^2 + (\beta_i - \beta_j)^2} \tag{26}$$

3) Store the clustering result in Λ with the form of $(a_k^n, f_k^n, \beta_k^n, y_k^n)$, update the initial cluster centroid $C_{m+l_s} = \hat{C}_m$.

4) Set $m = m + l_s$, repeat step 1) to step 3) until $m=N$.

III. EXAMPLES

A multicomponent signal is utilized to validate the proposed approach.

$$x(t) = \sum_{i=1}^{3} x_i(t) + e(t) \tag{27}$$

where $x_1(t)$ is a SS signal; $x_2(t)$ is a CS signal; $x_3(t)$ is a SFMS signal; $e(t)$ is zero mean, white Gaussian stationary noise with variance σ_e^2.

$$x_1(t) = \exp\{j2\pi\cdot(-120)t\} \tag{28}$$

$$x_2(t) = \cos(2\pi\cdot 0.05t)\exp\{j2\pi(200t + 0.5\cdot(-300)t^2)\} \tag{29}$$

$$x_3(t) = \sin\left(2\pi\cdot 0.1t + \frac{\pi}{2}\right) \cdot\exp\{j2\pi\cdot 10\sin(2\pi t - \pi)\}\exp\{j2\pi\cdot 80t\} \tag{30}$$

The duration of $x(t)$ is $T = 1\,\mathrm{s}$; the sampling frequency is $f_s = 500$ Hz. The discrete version of (27) is denoted as $x(n)$, where $n = 0,1,\cdots,N-1$, and N is the signal length. The signal to noise ratio (SNR) is defined by

$$\mathrm{SNR} = 10\log 10\frac{\sum_{n=0}^{N-1}|x(n)|^2}{N\sigma_e^2} \tag{31}$$

A 10 dB noise contaminated signal is first used for performance evaluation and the results are plotted in Fig.1-3. Fig.1 shows the STFT of the signal with an 85-length hamming window. The energy of the SS signal is more concentrated along the IF than the two other components and appears to be the strongest though these three components have similar IAs. This indicates that the STFT is not accurate enough for dealing with non-sinusoidal signals.

Fig. 2 shows the result of the STSR and its projections on the TF and the TFR plane; the amplitudes of the components are encoded in the color. In this example, the signal components appear as non-intersected curves, i.e., two lines and a helix curve, in the TFFR plot. IF and IFR can be directly obtained from the STFFRR. As it can be seen from the two upper subfigures, the disturbance of the IFR estimates are more obvious compared with the IF estimates. This means that the IFR estimates are less accurate. Furthermore, the IAs of the three component are similar. The experiment results show that the STSR is suitable for processing multicomponent nonstationary signals.

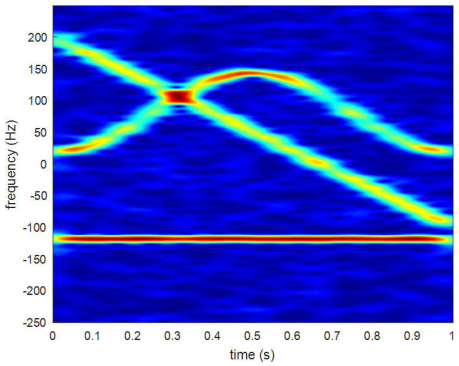


Fig. 1. STFT of the signal.

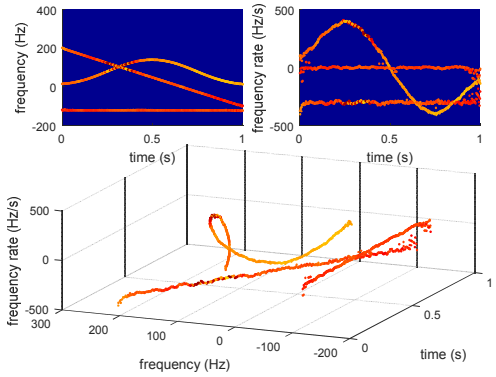
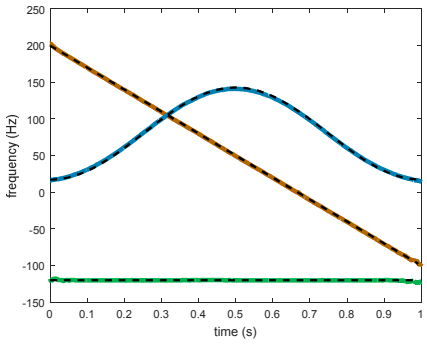
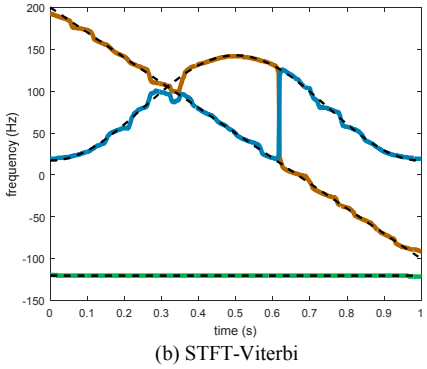


Fig. 2. STSR results. Upper left plot: TF representation; Upper right plot: TFR representation; lower plot: TFFR representation.



(a) STSR-k-means



(b) STFT-Viterbi

Fig. 3. IF estimates. The dashed lines represent the true IF and the solid lines represents the IF estimates.

ISBN 978-90-827970-1-5

Fig. 3 shows the IF estimates by performing local k-means clustering algorithm on the STSR and the Viterbi algorithm on the STFT [10]. It is evident that the proposed STSR-k-means algorithm can exactly extract each individual components with negligible IF derivations. While the STFT-Viterbi algorithm fails to recognize the intersected CS signal and SFMS. The IF estimates of these two components are swapped at 0.3s and 0.6s and with larger derivations. It should be pointed out that the Viterbi algorithm has the potential to track intersected IF components providing that the IFRs are utilized. However, the Viterbi algorithm is a global searching algorithm which is dependent on the defined parameterized path penalty function whose parameters are difficult to choose and optimize. Thus the STSR-Viterbi algorithm is not performed here.

This result shows that the STSR-k-means algorithm can process the multicomponent nonstationary signal with a high performance, while the STFT-Viterbi fails to achieve this. The permutations between intersected components in the TF plot are resolved by using IFR as the additional information and the local k-means algorithm for component linking.

To further demonstrate the performance of STSR, Monte Carlo simulation is carried out for SNR varying from -5 dB to 15 dB. 100 simulations is performed for each SNR. Fig. 4 shows the mean square errors (MSEs) of the IF estimates of the aforementioned three signals; a comparison with the STFT is made. The MSE of the IF estimates is given by

$$MSE = \frac{1}{NM}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left|\hat{f}^m(n) - f(n)\right|^2 \qquad (32)$$

where $f(n)$ are the true IFs of the signal and $\hat{f}^m(n)$ are the IF estimates, n is the discrete time, the superscript m represents the mth simulation.

Except for the SS signal, the MSEs of IF estimates of the STSR are much lower than that of the STFT for all noise levels. The average improvements for the CS signal and the SFMS signal are 3.3 and 1.3 in log 10 scale, respectively.

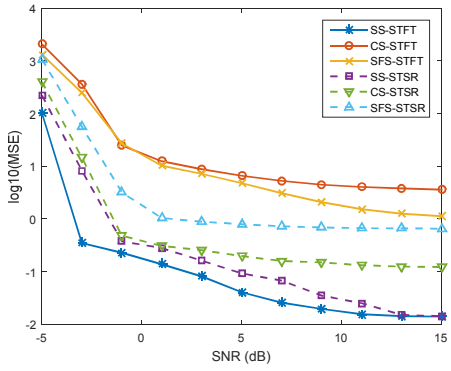


Fig. 4. MSE of IF estimates versus SNR

IV. CONCLUSIONS

In this work, the STFFRR is proposed by using the STSR with the chirp dictionary. By employing the chirp dictionary, the STFFRR can well modeled nonstationary signals with a high accuracy. Compared with the existing representations, the STFFRR has two distinct properties. Firstly, as a SR-based representation, the STFFRR can provide direct joint IF and IFR estimation. Secondly, the signal components can only be intersected in the TFFR plot when they are tangent in the TF plot. Thus, the probability of intersection decreases and the signal components appear as separated and non-intersected curves for most of the time. By performing local k-means clustering algorithm on the IF and IFR estimates, the permutations between intersected components in the TF plot are solved. Results show that robust component separation and IF estimation are obtained.

The TFFRR of three typical signals, i.e., the SS signal, the CS signal and the SFMS signal are derived as well. The SS signal and the CS signal appear as lines in the TFFR plot, and the SFMS signal appears as a helix line. It is shown that the SS signal and the CS signal, and two CS signals can never be intersected in TFFFR. The special properties of TFFRR, i.e., the curve shapes of different types of signals can be used to develop more sophisticated nonstationary signal processing methods. Further research on how to achieve this will be carried out.

REFERENCE

[1] L. Cohen, Time-Frequency Analysis: Theory and Applications. Upper Saddle River, NJ, USA: Prentice-Hall PTR, 1995.

[2] L. Stankovic, I. Djurovic, S. Stankovic, et al, "Instantaneous frequency in time–frequency analysis: Enhanced concepts and performance of estimation algorithms," Elsevier Digital Signal Process., vol.35, pp.1-13, Dec. 2014.

[3] S. Qian, D. Chen "Adaptive chirplet based signal approximation," in IEEE International Conference on Acoustics, Speech and Signal Processing, vol.3, pp.1781-1784, May. 1998.

[4] Q. Yin, S.Qian, A. Feng,A, " Fast Refinement for Adaptive Gaussian Chirplet Decomposition," IEEE Trans. Signal Process., vol.50, no.6, pp. 1298-306, Jun. 2002.

[5] J. C. O'Neill, P. Flandrin, W. C. Karl, "Sparse Representations with Chirplets via Maximum Likelihood Estimation," 1999. Available: http://tfd.sourceforge.net/Papers/chirp.pdf.

[6] P. O'Shea, "A new technique for instantaneous frequency rate estimation," IEEE Signal Process. Lett., vol. 9, no. 8, pp. 251–252, Aug. 2002.

[7] L. Zuo, M. Li, Z. Liu, et al, "A High-Resolution Time-Frequency Rate Representation and the Cross-Term Suppression," IEEE Trans. Signal Process., vol. 64, no. 10, pp. 2463-2474, May, 2016.

[8] I. Djurovic, "Viterbi algorithm for chirp-rate and instantaneous frequency estimation," Elsevier Signal Process., vol. 91, no.5, pp.1308-14, May. 2011.

[9] S.S. Abeysekera, "Time-frequency and time-frequency-rate representations using the cross quadratic spectrum," IEEE Tencon Spring Conference, pp.500-504, 2013.

[10] I. Djurovic, L. Stankovic, "An algorithm for the wigner distribution based instantaneous frequency estimation in a high noise environment," Elsevier Signal Process., vol.84, no.3, pp.631-43, Mar. 2010.

[11] I. Djurovic, T. Thayaparan, L. Stankovic, "Adaptive Local Polynomial Fourier Transform in ISAR," EURASIP J. Applied Signal Process., vol. 2006, pp.1298-313,2006.

[12] V. Sucic, N. Saulig, B. Boashash, "Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Rényi entropy," EURASIP J. Adv. Signal Process., vol. 2011, no. 1, 2011. Available: https://doi.org/10.1186/1687-6180-2011-125.

ISBN 978-90-827970-1-5 © EURASIP 2018