Contents

0.1 Adaptive Directional Time-Frequency Distributions [0]

Quadratic time-frequency distributions are often assessed in terms of their ability to resolve closely spaced signal components while suppressing undesired cross-terms. Cross-terms are usually suppressed by employing 2D smoothing [1], but such smoothing deteriorates the resolution of auto-terms. Section 5.7 has shown that TFDs based on separable kernels offer higher resolution as compared to the spectrogram due to the extra degree of freedom to independently control smoothing along time and frequency axis. However, these methods fail to provide the optimum energy concentration, when the signal energy is concentrated in a certain direction in the (t, f) domain. In such scenarios, (t, f) kernels rotated along the direction of energy distribution can significantly suppress cross-terms while retaining the resolution of auto-terms [2]. This method of rotating a kernel along the major axis of the auto-terms cannot be applied to signals having multiple directions of energy distribution in the (t, f) domain [2]. In order to deal with such situation, the shape of the smoothing kernel could be adapted locally either at each time instant (e.g., the adaptive optimal kernel approach [3]) or at each (t, f) point (e.g., the adaptive fractional spectrogram [4]).

0.1.1 The Need for Adaptive Directional Kernels

Let us consider two sample cases where signal representations in the (t, f) domain can be improved if adaptive kernels are adopted. Let us first consider a signal composed of two parallel linear frequency modulated (LFM) chirps and two tones. Figure 0.1.1 illustrates the Wigner-Ville distribution (WVD) and the extended modified B-distribution (EMBD) of the given signal. The (t, f) representation obtained by the WVD is more complicated for interpretations due to the presence of cross-terms. The EMBD has improved the signal representation by reducing cross-terms generated due to the interaction between LFM components and tones. However, the EMBD still contains cross-terms generated due to the interaction between two close components i.e. cross-terms appearing in between the IFs of two close LFM signals and cross-terms appearing in between the IFs of two close tones.

To further suppress cross-terms in the EMBD, a single directional kernel cannot be used for directional smoothing as the LFM signals and two tones are not along the same direction as shown in Figure 0.1.2(a). For example, directional smoothing along the direction of the major axis of the two tones suppresses the cross-terms appearing due to the tone interaction, but fails to suppress cross-terms between two LFM signals as illustrated in Figure 0.1.2(b). In this scenario, successive application of two directional filters, one parallel to the tones and another one parallel to the LFM signals, can remove the cross-terms without degrading the resolution of auto-terms as illustrated in Figure 0.1.2(c).

[0]Author: **Nabeel Khan, Boualem Boashash**, Department of Electrical Engineering, Qatar University (nabeelalikhan@qu.edu.qa and boualem@qu.edu.qa), **Ervin Sejdić**, Department of Electrical and Computer Engineering, University of Pittsburgh (esejdic@ieee.org). Reviewers: S.

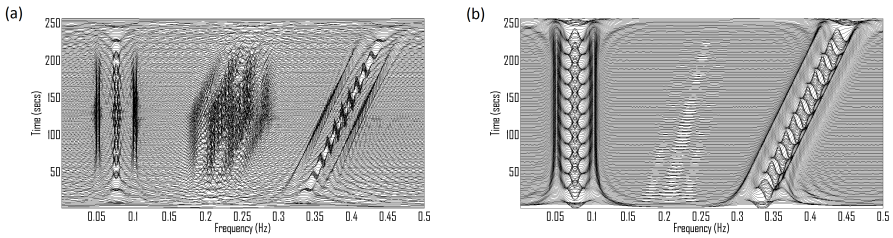

Fig. 0.1.1: TFDs of a multi-component signal composed of 2 tones and 2 LFM components: (a) the WVD of a four component signal; (b) the EMBD of the same four component signal.

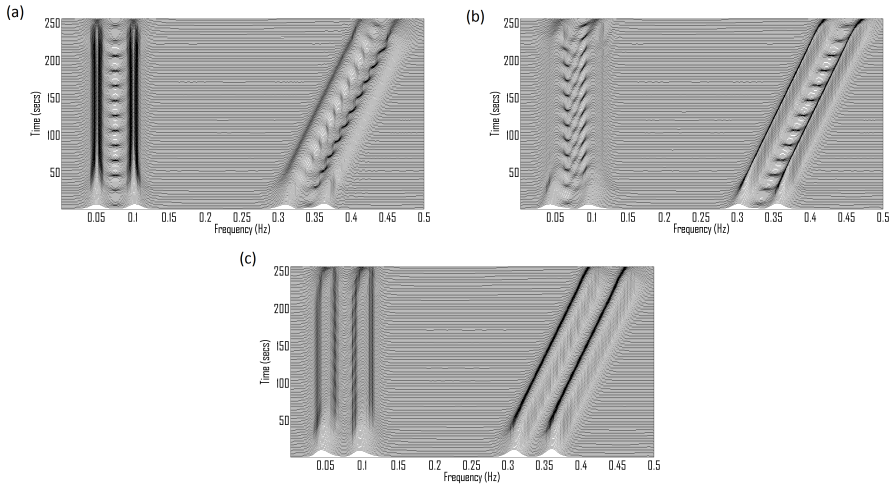

Fig. 0.1.2: Directionally smoothed TFDs of a multi-component signal composed of 2 tones and 2 LFM components: (a) the EMBD filtered by the directional Gaussian filter parallel to time axis; (b) the EMBD filtered by the directional Gaussian filter parallel to the major axis of LFM signals; (c) a TFD obtained by first applying the directional filter along the time axis and then along the major axis of the LFM signal components.

A more complicated case involves a signal composed of two quadratic frequency modulated component as shown in Figure 0.1.3. The WVD is difficult to interpret because of both inner and outer interferences. The EMBD, by suppressing inner interference terms, shows the two quadratic FM signal components and oscillatory cross-terms between the two components. It is observed that the direction of oscillation of cross-terms changes with time; the cross-terms cannot therefore be suppressed by the application of one or two directional filters as done in earlier examples. In this scenario, we need a smoothing kernel that adapts itself locally at each time instant or at each (t, f) point as discussed in the following sections.

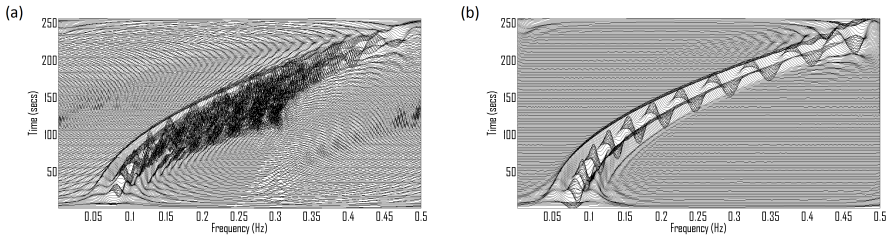

Fig. 0.1.3: TFDs of a multi-component signal composed of two quadratic chirps: (a) the WVD of a signal composed of two quadratic chirps; (b) the EMBD of the two component signal.

0.1.2 Adaptive Directional (t, f) Distributions

Adapting the orientation and the shape of a smoothing kernel can be performed using the following two approaches:

1. Time varying local adaptation: The shape of a smoothing kernel is adapted for each time instant, excluding the f variable.

2. (T, f) varying local adaptation: The shape of a smoothing kernel is adapted for each (t, f) point, so taking advantage of the extra degree of freedom provided by the full (t, f) plane.

Accounting for both cases, and others, a general formulation of an adaptive TFD may be given as:

$$\rho^{(adapt)}(t, f) = \rho(t, f) *_t *_f \gamma(t, f) \tag{0.1.1}$$

where $*_t$ represent convolution along the time axis, $*_f$ represents convolution along the frequency axis, $\gamma(t, f)$ is a smoothing kernel, $\rho(t, f)$ is a TFD, and $\rho^{(adapt)}(t, f)$ is the adaptive TFD.

0.1.2.1 Time-Varying Local Adaptation

The adaptive optimal kernel (AOK) TFD defined in Chapter 5.3. adapts the shape of its smoothing kernel at each time instant using the short time ambiguity function. In this case, the windowed WVD as computed from the short time ambiguity function, is the time-varying (t, f) smoothing kernel related to the Doppler lag kernel by the following expression [3] (or Chapter 5.3 in this book):

$$\gamma_t(t, f) = \mathfrak{F}^{-1}_{t \leftarrow \nu} \mathfrak{F}_{f \leftarrow \tau} e^{-\frac{\nu^2 + \tau^2}{2\sigma^2(\phi)}} \tag{0.1.2}$$

where $\sigma^2(\phi)$ controls the spread of the kernel along direction based on the constraints described in Chapter 5.3. $\mathfrak{F}$ and $\mathfrak{F}^{-1}$ represent the forward and backward Fourier operators, respectively. The optimization of the kernel shape is carried out by maximizing the correlation of the kernel with the short time ambiguity domain of a given multi-component signal at each time instant. The computational cost of

this algorithm is $O(M^3)$ [5], where M is the number of samples in the frequency domain.

0.1.2.2 (t, f) Varying Local Adaptation

Adaptive Fractional Spectrogram Another approach is to define an adaptive fractional spectrogram by analyzing a signal using an adaptive window such that the window parameters are optimized at each (t, f) point [4]:

$$AFS(t,f) = \left| \int s(\tau) h_{\alpha,\sigma}(\tau - t) e^{-j2\pi f\tau} d\tau \right|^2 \tag{0.1.3}$$

where

$$h_{\alpha,\sigma}(t) = \frac{e^{j\alpha/2}}{\sqrt{j \sin\alpha}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2\sigma^2}} e^{j\pi \frac{(u^2+t^2)\cos\alpha - 2ut}{\sin\alpha}} du \tag{0.1.4}$$

is the signal dependent analysis window, α is the chirp rate or the rotation order of the analysis window, and σ is the standard deviation of the Gaussian window. Both α and σ are adapted based on the observation that the ideal window for all those points that lie along the IF of signal components is the one that maximizes the correlation of the signal with the window. This criterion is expressed as:

$$h_{\alpha,\sigma}(t) = \underset{h_{\alpha,\sigma}(t)}{\arg\max} \left| s(\tau) h_{\alpha,\sigma}(\tau - t) e^{-j2\pi f\tau} d\tau \right|^2 \tag{0.1.5}$$

This criterion is not optimal for those points that do not exactly lie on the instantaneous frequency of the signal component. Hence, the adaptive fractional spectrogram leaks the signal energy at those (t, f) points where no signal is present. As it is difficult to find an analytical solution to eqn. (0.1.5), this equation is numerically solved using the procedure outlined in [4].

In general, the implementation of the adaptive fractional spectrogram requires the computation of K spectrograms. The computational cost of computing a single spectrogram is $O(NM \log M)$, where M is the number of samples in the frequency domain and N is the length of the signal. The total computational cost of the adaptive fractional spectrogram becomes then $O(KNM \log M)$.

Directional Gaussian Filter Based TFD Based on the sample signals shown in Figures 0.1.1, 0.1.2 and 0.1.2, there are two major observations:

- Cross-terms oscillate along their major axis.
  The main implication of this observation is that smoothing the cross-terms along their major axis will reduce them, as smoothing acts as a low-pass filter. The cut-off frequency of such a low-pass filter is dependent on the separation of the two closest signal components, as the rate of oscillation of cross-terms is directly proportional to the spacing between signal components. Note that a narrow-band filter with a longer duration would be better in resolving close frequency components, but such a filter spreads the signal energy for short-time components.

- Auto-terms vary slowly along their major axis.
  The major implication of this observation is that the smoothing (i.e., a low-pass filter) operation along the major auto-term axis does not significantly degrade the auto-terms appearance as the auto-terms appear in the pass region of the smoothing (filtering) operation. A general 2D smoothing has more severe effects on the appearance of auto-terms.

The above observations imply that smoothing along the major axis of auto- and cross-terms will reduce cross-terms without severely degrading the auto-terms. Hence, the problem of estimating the direction of the smoothing kernel is now reduced to the estimation of local directions of major axes (i.e. at each (t, f) point). Both auto-terms and cross-terms appear as ridges in the magnitude squared (t, f) domain as the squaring operation removes cross-term oscillations. The local directions can be estimated by maximizing the convolution value for a directional kernel convolved with a magnitude squared quadratic TFD as a directional kernel would have a maximum value when parallel to ridges (i.e., the major axis of cross-terms and auto-terms). Mathematically,

$$\theta(t,f) = \arg\max_{\theta} \left| |\rho(t,f)|^2 *_t *_f \gamma_\theta(t,f) \right|^2 \tag{0.1.6}$$

Based on the above expression, a smoothing kernel should therefore have the following properties:

- Good localization: The kernel should have the maximum response when aligned with ridges. This implies that the kernel should have low pass characteristics along its major axis. The kernel should also be symmetric along its major axis to match the shape of a ridge.
- Minimal number of false responses: The kernel output should be zero for non-ridge points as otherwise spectral leakages would be observed at (t, f) points where no signal is present.

The double derivative directional Gaussian filter (DGF), commonly used in image processing approximately fulfills the above properties. It is defined as [6]:

$$\gamma_\theta(t,f) = \frac{d^2}{df_\theta^2} e^{-a^2 t_\theta^2 - b^2 f_\theta^2} \tag{0.1.7}$$

where $t_\theta = t\cos(\theta) + f\sin(\theta)$ and $f_\theta = -t\sin(\theta) + f\cos(\theta)$. Note that DGF has low pass characteristics along its major axis (i.e. $e^{-a^2 t_\theta^2}$) while it performs second order differentiation along its minor axis (i.e. $\frac{d^2}{dt_\theta^2} e^{-b^2 f_\theta^2}$). This filter has a maximum response when parallel to ridges and its response reduces to zero as the filter is orthogonal to ridges. The parameters a and b control smoothing along the major and minor ridge axes. Small parameter values imply more smoothing along the respective axis while large values imply less smoothing. Optimal parameter values

depend on the nature of a signal. For example, we have experimentally observed that $a = 3$ and $b = 8$ provide good results for the analysis of real-life EEG seizure signals.

0.1.2.3 Algorithm for implementing a DGF-based TFD

To implement a DGF-based TFD the following steps should be adopted:

1. Analyze a given signal using a quadratic TFD.
2. Convolve the absolute squared quadratic TFD using K DGFs:

$$\rho_k(t,f) = |\rho(t,f)|^2 *_t *_f \gamma_{\theta_k}(t,f) \tag{0.1.8}$$

   where $\theta_k = 2\pi k/K$ for $0 \leq k \leq K-1$. K is the total number of quantization levels for the angle θ. In this study, 16 quantization levels were used.

3. Estimate $\theta(t,f)$ at each (t,f) point by choosing k that maximizes the squared magnitude of a directionally smoothed TFD.

$$\theta(t,f) = \frac{2\pi}{k} \arg\max_k |\rho_k(t,f)|^2 \tag{0.1.9}$$

4. Using $\theta(t,f)$, calculate the DGF based adaptive TFD by convolving the adaptive DGF, $\gamma_\theta(t,f)$, with the quadratic TFD as expressed in eqn. (0.1.1).

To understand the computational cost of implementing the DGF based TFD, let us consider the steps involved:

1. Computation of a quadratic TFD with the computational cost equal to $O(NM \log M)$.
2. Estimation of the direction of the smoothing kernel: This step involves K convolution operations, so the computational cost of this step is $O(KMN)$ as the computational cost of the single convolution operation is $O(MN)$ if there are MN points in the (t,f) domain.
3. Convolving the quadratic TFD with the adaptive DGF: The computational cost of this step is $O(MN)$ as it involves a single convolution operation.

Therefore, the total computation cost is equal to the sum of the computational costs of the all the above mentioned steps i.e. $O(NM \log M + (K+1)MN)$.

0.1.3 Illustrative Examples

In order to assess the performance of these adaptive algorithms, we consider a multi-component signal composed of two quadratic chirps and two crossing LFM components:

$$x(n) = x_1(n) + x_2(n) + x_3(n) + x_4(n) \tag{0.1.10}$$

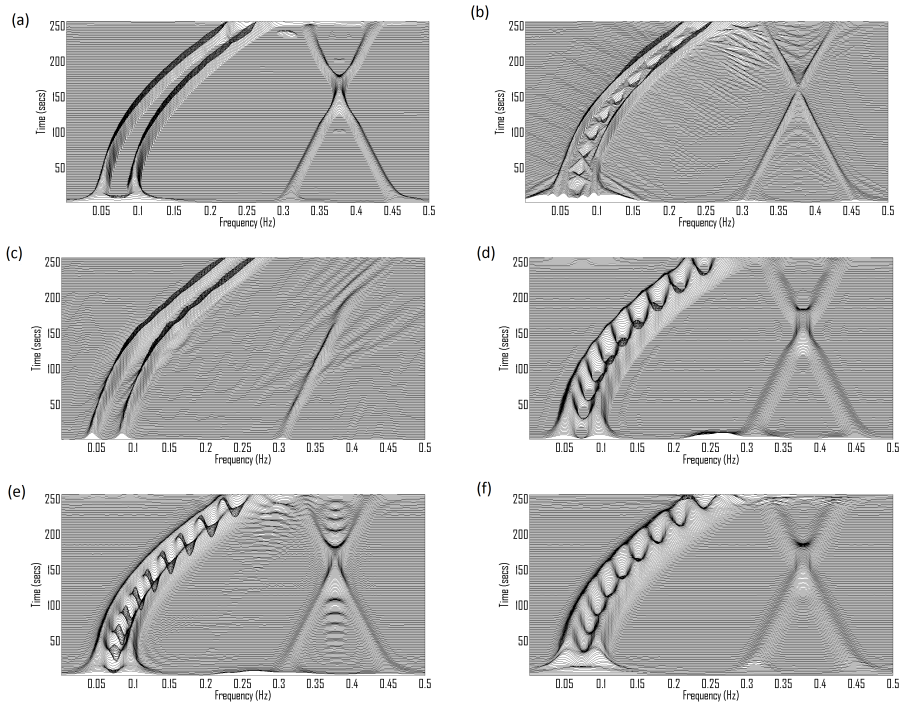

Fig. 0.1.4: TFDs of a multi-component signal composed of two quadratic chirps and two crossing LFM components: (a) DGF based TFD ($a = 3$, $b = 12$); (b) adaptive fractional spectrogram; (c) AOK-TFD; (d) compact kernel TFD ($C = 1$, $D = 0.06$, $E = 0.1$); (e) EMBD ($\alpha = 0.07$, $\beta = 0.15$); (f) spectrogram (the Hamming window with its length equal to 65 points).

where $x_1(n) = 2\cos\left(0.1\pi n + 0.000002\pi n^3\right)$, $x_2(n) = 2\cos\left(0.18\pi n + 0.000002\pi n^3\right)$ and $x_3(n) = \cos\left(0.6\pi n + 0.0005\pi n^2\right)$ and $x_4(n) = \cos\left(0.9\pi n - 0.0005\pi n^2\right)$ for $0 \leq n \leq 255$. The signal is challenging for most TFDs as the instantaneous frequencies of its components have non-linear variations. Furthermore, the signal energy is not uniformly distributed among all signal components i.e. the quadratic chirps carry more energy as compared to the crossing LFM components. To illustrate the performance of various TFDs, this signal is analyzed using several TFD approaches as shown in Figure 0.1.4.

Figure 0.1.4 indicates that the fixed kernel TFDs including the spectrogram, the EMBD, and the compact kernel TFD have failed to resolve the two closely spaced FM quadratic chirps as these methods cannot adapt the direction of the smoothing kernel. The adaptive fractional spectrogram has resolved closely spaced quadratic chirps, but it suffers from the energy leakage problem i.e. we can observe the signal energy between the instantaneous frequencies of two quadratic chirps. The AOK-TFD resolves two quadratic chirps, but fails to concentrate energy for crossing LFM components as the shape of adaptive optimal kernel is adapted at each time instant,

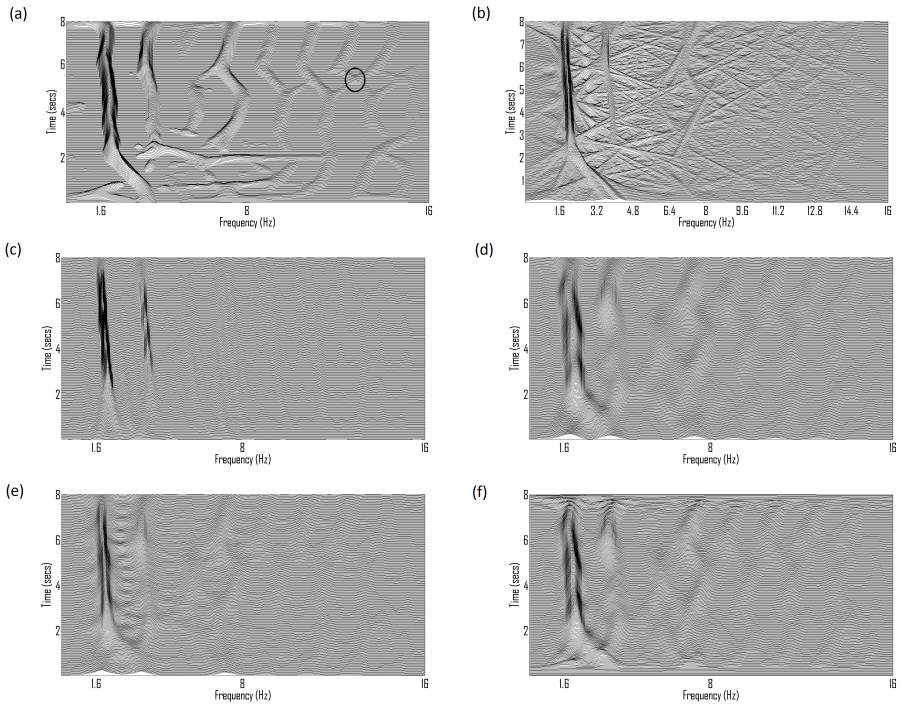

Fig. 0.1.5: TFDs of a sample EEG signal: (a) DGF based TFD ($a = 3$, $b = 8$) (The black circle indicates the region where two signal components intersect); (b) adaptive fractional spectrogram; (c) AOK-TFD; (d) compact support kernel based TFD ($C = 1$, $D = 0.06$, $E = 0.1$); (e) EMBD ($\alpha = 0.15$, $\beta = 0.05$); (f) spectrogram (the Hamming window with its length equal to 95 points).

but not at each (t, f) point. Hence, the AOK-TFD fails to concentrate energy for weak signal components. On the other hand, the DGF based TFD has resolved closely spaced signal components and it gives clear representation for crossing LFM components.

Let's now consider a real-life 8 second long seizure EEG signal. Figure 0.1.5 shows that the compact support kernel TFD and the AOK-TFD provide high energy concentration for the two strong signal components, but fail to concentrate signal energy for the weak components lying at higher frequencies. The adaptive fractional spectrogram is difficult to interpret because of the leakage of the signal energy at those (t, f) points where there is no signal. The DGF based TFD provides a clear representation of all the signal components including the weak signal components. The DGF based TFD also reveals that the components of the seizure signal intersect each other (see the black circle in figure 0.1.5 (a)). This information cannot be obtained from other TFDs due to their poor resolutions of weak signal components.

0.1.4 Summary and Conclusions

A class of adaptive directional TFDs is described in this article. These TFDs adapt the direction of a smoothing kernel based on the distribution of the signal energy in the (t, f) domain. The analysis of the location and orientation of cross-terms in quadratic TFDs define the following constraints for adapting the direction of a smoothing kernel:

- For parallel LFM signals, a global directional kernel aligned with the major axis of auto-terms can reduce cross-terms without deteriorating the resolution of auto-terms.
- For LFM signals with more than one direction of energy concentration, successive applications of multiple directional filters should be performed.
- For non-linearly frequency modulated signals, local adaptation of the smoothing kernel at each (t, f) point is needed.

To achieve a high-resolution (t, f) distribution with significantly reduced cross-terms, one can adapt the direction of the DGF along the major axis of ridges in the (t, f) domain to exploit the fact that such smoothing removes cross-terms without degrading the resolution of auto-terms. Experimental results confirmed that the adaptive TFD based on DGF outperforms other adaptive and fixed kernel TFDs in terms of its ability to resolve closely spaced signal components while removing cross-terms for both synthetic and real-life EEG seizure signals.

References

[1] B. Ristić and B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform," *IEEE Transactions on Signal Processing*, vol. 41, pp. 1996–2008, May 1993.

[2] M. J. Bastiaans, T. Alieva, and LJ. Stanković, "On rotated time-frequency kernels," *IEEE Signal Processing Letters*, vol. 9, pp. 378–381, Nov. 2002.

[3] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Fast algorithm for optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 42, pp. 134–146, Jan. 1994.

[4] N. A. Khan and B. Boashash, "Instantaneous frequency estimation of multicomponent nonstationary signals using multiview time-frequency distributions based on the adaptive fractional spectrogram," *IEEE Signal Processing Letters*, vol. 20, pp. 157–160, Feb. 2013.

[5] M. Abed, A. Belouchrani, M. Cheriet, and B. Boashash, "Time-frequency distributions based on compact support kernels: Properties and performance evaluation," *IEEE Transactions on Signal Processing*, vol. 60, pp. 2814–2827, June 2012.

[6] Y. Liu, L. Mejias, and Z. Li, "Fast power line detection and localization using steerable filter for active UAV guidance," in *Proc. of 12th International Society for Photogrammetry and Remote Sensing (ISPRS2012)*, (Melbourne, Australia), pp. 491–496, Aug. - Sept. 25-1, 2012.