Time-Frequency Feature Representation Using Energy Concentration: An Overview of Recent Advances

Ervin Sejdić, Igor Djurović and Jin Jiang

Abstract— **Signal processing can be found in many applications and its primary goal is to provide underlying information on specific problems for the purpose of decision making. Traditional signal processing approaches assume the stationarity of signals, which in practice is not often satisfied. Hence, time or frequency descriptions alone are insufficient to provide comprehensive information about such signals. On the contrary, time-frequency analysis is more suitable for nonstationary signals. Therefore, this paper provides a status report of feature based signal processing in the time-frequency domain through an overview of recent contributions. The feature considered here is energy concentration. The paper provides an analysis of several classes of feature extractors, i.e., time-frequency representations, and feature classifiers. The results of the literature review indicate that time-frequency domain signal processing using energy concentration as a feature is a very powerful tool and has been utilized in numerous applications. The expectation is that further research and applications of these algorithms will flourish in the near future.**

I. Introduction

Signal processing is often used for feature extraction and classification in medical disease diagnosis [1]-[3], industrial process control [4], fault detection [5] and many other fields. The primary goal of signal processing in these applications is to provide underlying information on specific problems for decision making [6]. These techniques can be classified either as time, frequency or time-frequency domain based algorithms. At the classification level, there also exist several different methodologies. Typical approaches along with sample features used in extraction and classification are shown in Fig. 1. Understanding of the problem at hand is crucial in deciding which framework to employ for feature analysis. Some features, such as amplitude levels in the time domain, are easily extracted and classified, but are susceptible to noise. Others, such as energy concentration in the time-frequency domain, even though require more involved operations, can lead to more robust feature extraction and more accurate classification. Furthermore, not every feature yields plausible conclusions. For example, in the analysis of heart sounds, which are nonstationary, the amplitude rarely provides conclusive information. The intensity of the recorded heart sounds is affected by many factors, which are not necessarily pathological. On the other hand, the amplitude in the time domain will provide sufficient information when considering control of the liquid level in a tank. Therefore, depending upon whether the phenomenon under analysis is stationary or nonstationary, and on the nature of the desired feature, different algorithms have to be used. The question is what signal processing algorithms should be used for feature analysis in a given situation? The answer simply depends on *a priori* knowledge about the phenomenon under consideration. Parametric signal processing algorithms can be used for feature extraction and classification if an accurate model of the signal exists in a selected representation space [7]. However, such modeling techniques have limitations as well. Modeling of nonstationary signals is more difficult and consistent parametric models often do not exist, except in very few special cases, e.g., mono or multi component chirp signals [8]. Most of the signals encountered in practice do not satisfy the sta-

Digital Signal Processing, Vol. 19, No. 1, Jan. 2009.

tionarity conditions, which explains the growing interest in nonstationary signal processing.

A. Time-Frequency Analysis

Time-frequency analysis (TFA) is of great interest when the signal models are unavailable. In those cases, the time or the frequency domain descriptions of a signal alone cannot provide comprehensive information for feature extraction and classification. The time domain lacks the frequency description of the signals. The Fourier transform of the signal cannot depict how the spectral content of the signal changes with time, which is critical in many nonstationary signals in practice. Hence, the time variable is introduced in the Fourier based analysis in order to provide a proper description of the spectral content changes as a function of time. Therefore, the basic goal of the TFA is to determine the energy concentration along the frequency axis at a given time instant, i.e., to search for joint time-frequency representation of the signal [9]. In an ideal case, the time-frequency transform would provide direct information about the frequency components occurring at any given time by combining the local information of an "instantaneous frequency spectrum"with the global information of the temporal behaviour of the signal [10], [11].

The time-frequency representations (TFRs) can be classified according to the analysis approaches. In the first category, the signal is represented by time-frequency (TF) functions derived from translating, modulating and scaling a basis function having a definite time and frequency localization. For a signal, $x(t)$, the TFR is given by

$$TF_x(t,\omega) = \int_{-\infty}^{+\infty} x(\tau)\phi^*_{t,\omega}(\tau)d\tau = \langle x, \phi_{t,\omega}\rangle \tag{1}$$

where $\phi_{t,\omega}$ represents the basis functions (also called the TF atoms) and * represents the complex conjugate. The basis functions are assumed to be square integrable, $\phi_{t,\omega} \in \mathbf{L^2}(\mathbb{R})$, i.e., they have finite energy [12]. Short-time Fourier transform (STFT) [10], wavelets [12], [13], and matching pursuit algorithms [12], [14] are typical examples in this category.

Cohen's idea of time-frequency distributions (TFD), originally proposed in [15], represents the second category of TFRs. This approach characterizes the TFR by a kernel function. The properties of the representation are reflected by simple constraints on the kernel that produces the TFR with prescribed, desirable properties [9]. A mathematical description of these TFRs can be given by

$$\begin{aligned} &TFD_x(t,\omega) \\ &= \frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\left(u+\frac{1}{2}\tau\right) \\ &\quad \times x^*\left(u-\frac{1}{2}\tau\right)\phi(\theta,\tau) \\ &\quad \times e^{-j\theta t - j\tau\omega + j\theta u} du d\tau d\theta \end{aligned} \tag{2}$$

where $\phi(\theta,\tau)$ is a two-dimensional kernel function, determining the specific representation in this category, and hence, the properties of the representation. Wigner distribution, Choi-Williams distribution, and spectrogram are some of the methods commonly used for obtaining the TFDs [9].

Extensive review of TFRs and their properties is beyond the scope of this paper; however, an interested reader is referred to the following excellent sources [10]-[13], [16]-[37] for details.

B. Feature Based Signal Processing and TFA

The main goal of the TFA of a signal is to determine the energy distribution along the frequency axis at each time instant [9]. Effects of TF transforms on energy distribution are considered by using three sample signals: $x_1(t)$ - a signal with four short transients; $x_2(t)$ - a linear chirp; and $x_3(t)$ - a signal with sinusoidally modulated frequency. The TF domain representations of the signals are obtained by four different TFRs: STFT, S-transform [38], S-method [39], and Wigner distribution (WD) as shown in Fig. 2.

Several observations can be made by comparing the respective TFRs. The STFT provides constant concentration at all frequencies. The S-transform provides good concentration at lower frequencies, but poor concentration at higher frequencies. The S-method provides overall good concentration at all frequencies,

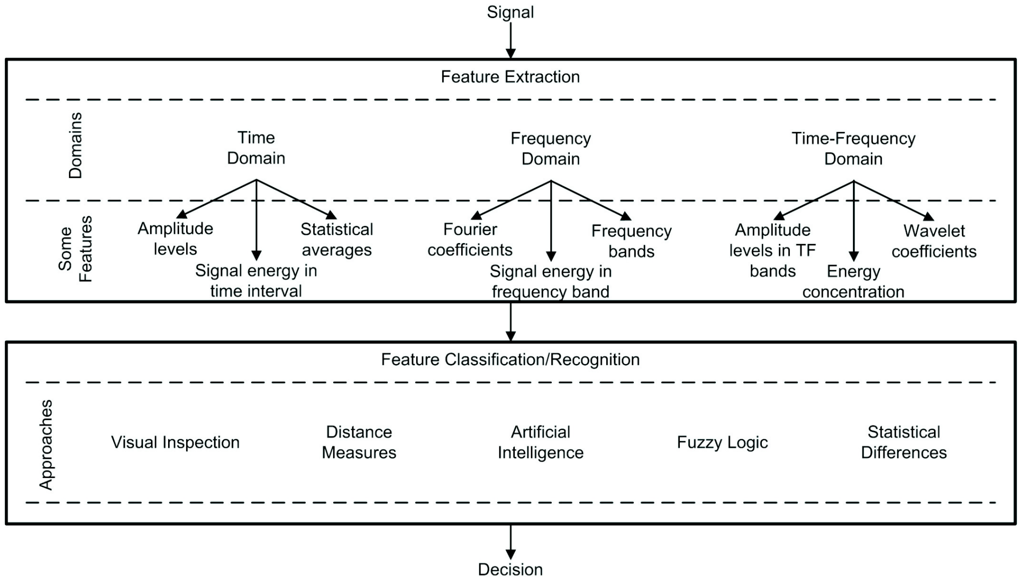

Fig. 1. Signal processing for pattern classification in a typical application.

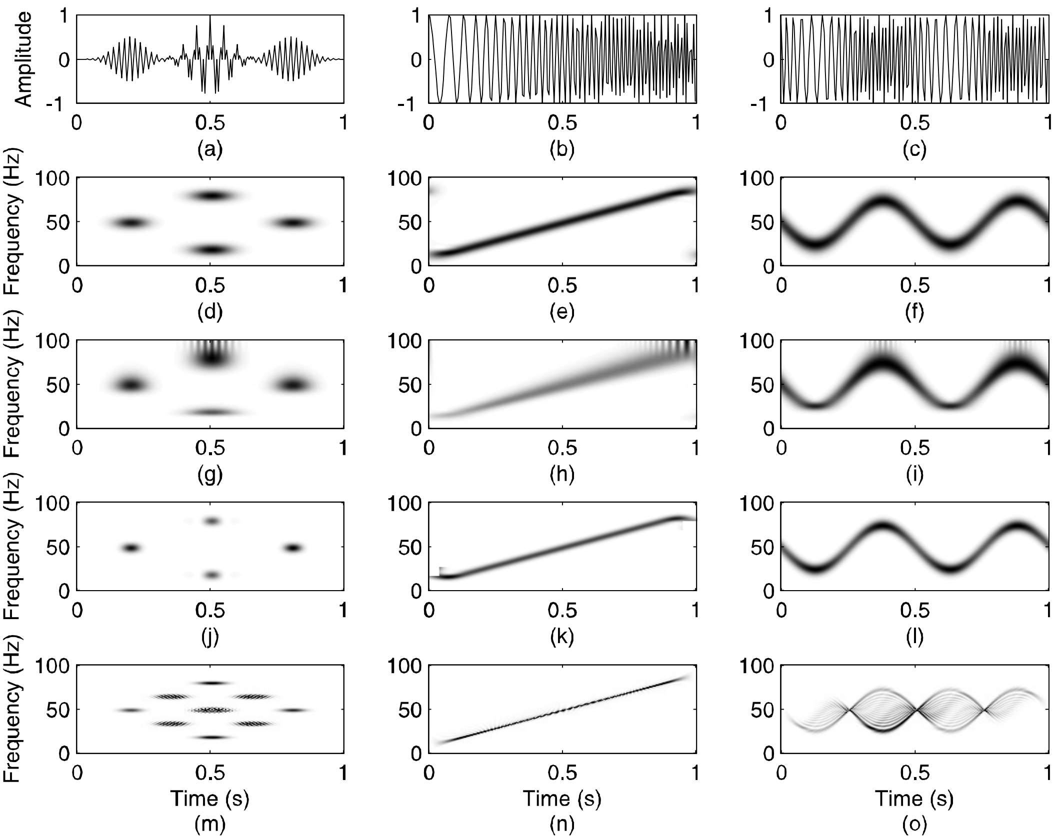

Fig. 2. Sample TFRs: (a) $x_1(t)$; (b) $x_2(t)$; (c) $x_3(t)$; (d) STFT of $x_1(t)$; (e) STFT of $x_2(t)$; (f) STFT of $x_3(t)$; (g) S-transform of $x_1(t)$; (h) S-transform of $x_2(t)$; (i) S-transform of $x_3(t)$; (j) S-method of $x_1(t)$; (k) S-method of $x_2(t)$; (l) S-method of $x_3(t)$; (m) WD of $x_1(t)$; (n) WD of $x_2(t)$; (o) WD of $x_3(t)$.

but it is nonivertible, which may pose a problem if a synthesis of the entire or a part of the signal is required. The Wigner distribution suffers from cross terms for multicomponent signals. Furthermore, this distribution may also suffer from inner interferences for monocomponent signals as shown in Fig. 2(o). These simple examples show that no single TFR can be ideal for all possible applications. The choice of a particular TFR depends on specific applications at hand. However, the TFA offers what other time or frequency techniques are unable to do. Simultaneous analysis of a signal in time and frequency domains has proved to be the key to successful extraction and classification of signals with different characteristics in numerous applications.

One of the simplest feature based signal processing procedures in TFA is via energy concentration. The idea is to analyze behaviour of the energy distribution, i.e., the concentration of energy at certain time instant or certain frequency band or more generally, in some particular time and frequency region. Such analysis is capable of revealing more information about a particular phenomenon for diagnostic purposes. However, if the energy concentration in the TF domain is used as a feature for extraction, classification and/or recognition, the following questions have to be answered. For example, can enhanced concentration of the STFT be achieved? More generally, is it possible to enhance the energy concentration in the TF domain for a variety of TFRs such that they resemble as closely as possible to an ideal TFR? Also, if the energy concentration in a certain TF band is used as a feature in a classification process then how does one carry out the classification procedure? Should existing classification techniques be used? Or should new classification schemes be developed which rely strictly on the TFR? The rest of this paper provides a literature overview on the development in the field of feature based signal processing in the TF domain, and also provides some answers to the above questions.

Two research streams prevail in the literature as shown in Fig. 3. The first stream relies on enhancement of the energy concentration in the TF domain. The idea is that the properly optimized energy concentration will simplify the decision-making process. From a pattern recognition point of view, this approach essentially means increasing the resolution of the feature extractor. The second stream deals with the development of new classification schemes relying on TFR of the signal. For example, it has been shown that the accuracy of a correlation based classifier can be enhanced if certain pre-processing of the signal is carried out.

C. Organization of the Paper

This paper has been divided into six sections. Section II provides an overview of the TF algorithms relevant to the scope of this paper. These algorithms have appeared in the literature dating back to 1990's. Earlier developments of the TF techniques have been reviewed in excellent papers by Cohen [19] and Hlawatsch [22]. Section III provides a review of the classification schemes based on TFRs. An application example is shown in Section IV, where the accuracy of instantaneous frequency (IF) estimation for different TFRs is examined. General remarks and future directions regarding the feature analysis based on the energy concentration in the TF domain are presented in Section V. Conclusions are drawn in Section VI followed by an extensive list of references.

A reader should keep in mind of the followings while reading this paper: First, the paper provides an overview of algorithms for only one-dimensional signals. The overview of the algorithms based on the artificial intelligence methods or multidimensional signals (i.e. images) is beyond the current scope. Second, some of the algorithms considered herein have previously been reviewed, mostly in the form of edited books. For the sake of completeness, they are still included.

II. TFR as a Feature Extractor

Signal processing using energy concentration as a feature in the TF domain essentially consists of evaluating a TFR of the given signal. If the energy concentration in the TFR is closer to that of the ideal TFR, more likely it will produce more accurate classification re-

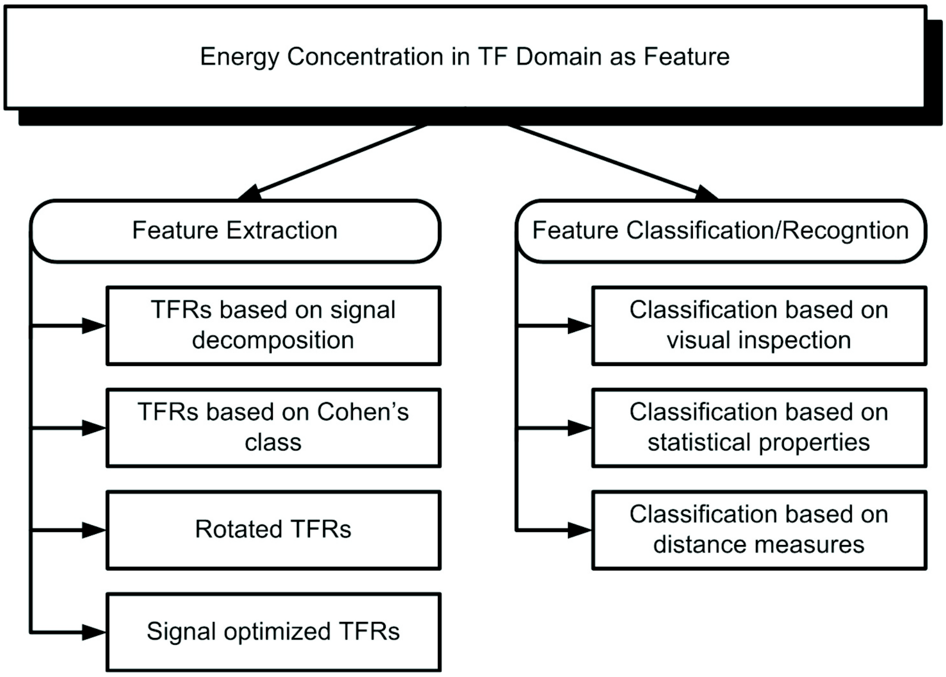

Fig. 3. Overview of feature extraction and classification procedures based on the energy concentration in the TF domain.

sults. Hence, a lot of research has focused on how to obtain more concentrated energy distribution.

Research activities reported in the literature can be summarized in the following four aspects: The first two deal with the development of new TFRs based on either signal decomposition or Cohen's idea. The third relies on so-called rotated TFRs, in which the TF plane is rotated to a certain angle in order to align the analysis axis with the signal components. The fourth relates to the signal optimized transform. A possible approach in obtaining the signal optimized transform is to employ a concentration measure in order to optimize the behaviour of a parameter. For example, the window length in the short-time Fourier transform can be optimized for every signal in order to achieve higher energy concentration [40]. Another approach to signal optimized transform is to design the TF transform optimized for classification. For example, the kernel of the transform is directly optimized in the TF domain to yield a classifier with a higher accuracy [41]. Even though the TFA represents a clear framework for the analysis of the energy concentration in time and frequency domains, there are still some problems as outlined by sample examples in the previous section. This section provides an overview of these approaches with emphasis on recent developments.

A. Signal Decomposition Based TFRs

The signal decomposition based TFRs are often used to describe energy concentration since they do not have cross term issues as those TFRs based on Cohen's idea. Cross terms can cause problems at the classification stages. The methods for decomposition range from classical such as STFT, wavelet transform to some newer methods such as:

• multiresolution Fourier transform (MFT) [42]:

$$\phi_{t,\omega}(\tau) = \sqrt{s}h(s(\tau - t)) \exp(-j\omega\tau) \tag{3}$$

where $h(\cdot)$ is a window function and s is the scale similar to one used in the wavelet analysis;

• S-transform [38]:

$$\phi_{t,\omega}(\tau) = h(\tau - t, \sigma(\omega)) \exp(-j\omega\tau) \quad (4)$$

where $h(\cdot)$ is a Gaussian window function and $\sigma(\omega)$ is the standard deviation of the Gaussian window;

• short-time harmonic transform (STHRT) [43], [44]:

$$\phi_{t,\omega}(\tau) = h(t - \tau)\varphi_u^{(1)}(\tau)\exp(-j\omega\varphi_u(\tau)) \quad (5)$$

where $\varphi_u(\tau)$, known as the unit phase function, is the phase function of the fundamental divided by its nominal IF and $\varphi_u^{(1)}(\tau)$ is the first-order derivative of $\varphi_u(\tau)$;

• short-time Hartley transform (STHT) [45]:

$$\phi_{t,\omega}(\tau) = h(t - \tau)\text{cas}(\omega\tau) \quad (6)$$

where $\text{cas}(\cdot) = \cos(\cdot) + \sin(\cdot)$.

It should be mentioned that the S-transform can be considered a special case of the MFT with the Gaussian window. In fact, the S-transform adds a constraint by restricting the window width of MFT. Because MFT is a function of three independent variables, it becomes difficult to be used as a tool for analysis [46].

Some properties of these techniques are summarized in Table I. The choice of a feature extractor, i.e., the TFR, depends on an application. Different techniques have unique properties.

A hyperbolic FM signal, $x(t) = \exp(j20\pi \ln(11|t| + 1))$, is used to examine the effects of a variable window width over a constant window. The signal is analyzed with STFT and the S-transform. The TFRs are shown in Fig.s 4(a) and 4(b). The S-transform provides a more concentrated representation than the STFT does due to the fact that the window for the S-transform is wider at lower frequencies and narrower at higher frequencies. However, the S-transform does not always yield satisfactory results as depicted in Fig.s 2(e) and 2(h), where higher energy concentration for the linear FM signal is achieved with the STFT.

The advantage of the TFA of the harmonic signal, $x(t) = \exp(j2\pi(10t + 5t^2)) + \exp(j2\pi(20t+5t^2))+\exp(j2\pi(30t+5t^2))$, with the STHRT over the STFT is depicted in Fig.s 4(c) and 4(d). These graphs represent TFRs of sample harmonic signal which consists of three linear FM signals. The STHRT yields a higher concentration in comparison to the STFT for the harmonic signals as expected. Furthermore, the STHRT provides a localized impulse-train spectrum for signals that are comprised of time-varying harmonics. However, a severe limitation for this transform is that $\phi_u(\tau)$ has to be known in advance. Otherwise, an exhaustive search procedure is required to determine the unit phase function [44].

Hardware implementation of most signal decomposition based techniques requires separate implementation for the forward and backward transforms. This may add to the cost of the implementation [45]. However, for STHT, any hardware built to compute the forward transform can be used for the inverse transform without any modification, because the Hartley transform kernel is the same for both the forward and the backward transforms.

Some shortcomings identified in Table I have been addressed in the literature. A generalized S-transform is introduced to allow greater control over the window function. This generalization also allows nonsymmetric windows to be used [47], [48]. Several window functions are considered including two forms of exponential functions, amplitude and phase modulations by cosine functions, a bi-Gaussian window [49], a complex phase function [50], and a subclass of complex windows [51]. The bi-Gaussian window is introduced to resolve time resolution associated with the Gaussian window. The long front tapers of the Gaussian window degrade the time resolution of event onsets [49]. By joining two non-symmetric half-Gaussian windows, this problem can be resolved. The phase and the amplitude modulation resolve the issue for complex windows which could produce a misleading amplitude spectrum in the TF domain. Unless corrected by proper modulation, the complex windows can produce an IF in the TF domain that is not equal to the true IF [50], [51]. The solution to the problem of the constant window width associated with the STHT is proposed in the

TABLE I

PROPERTIES OF THE SIGNAL DECOMPOSITION TECHNIQUES FOR REPRESENTING ENERGY CONCENTRATION IN THE TF DOMAIN.

Method	Advantages	Disadvantages
STFT	Very simple for implementation.	Constant window width limits time-frequency resolution.
Wavelet analysis	Variable resolution.	Does not maintain the absolute phase of the signal components. A scale to frequency conversion is dependent on a mother wavelet.
MFT	Variable resolution. Absolute phase of each component is maintained.	Complex requirements for the window function. Choice of scale might require oversampling.
S-transform	Variable resolution. Absolute phase of each component is maintained.	Single window function.
STHRT	Good energy concentration obtained for the harmonic signals.	$\varphi_u(\tau)$ has to be known or precisely estimated.
STHT	Easy for hardware implementation.	Same disadvantages as STFT.

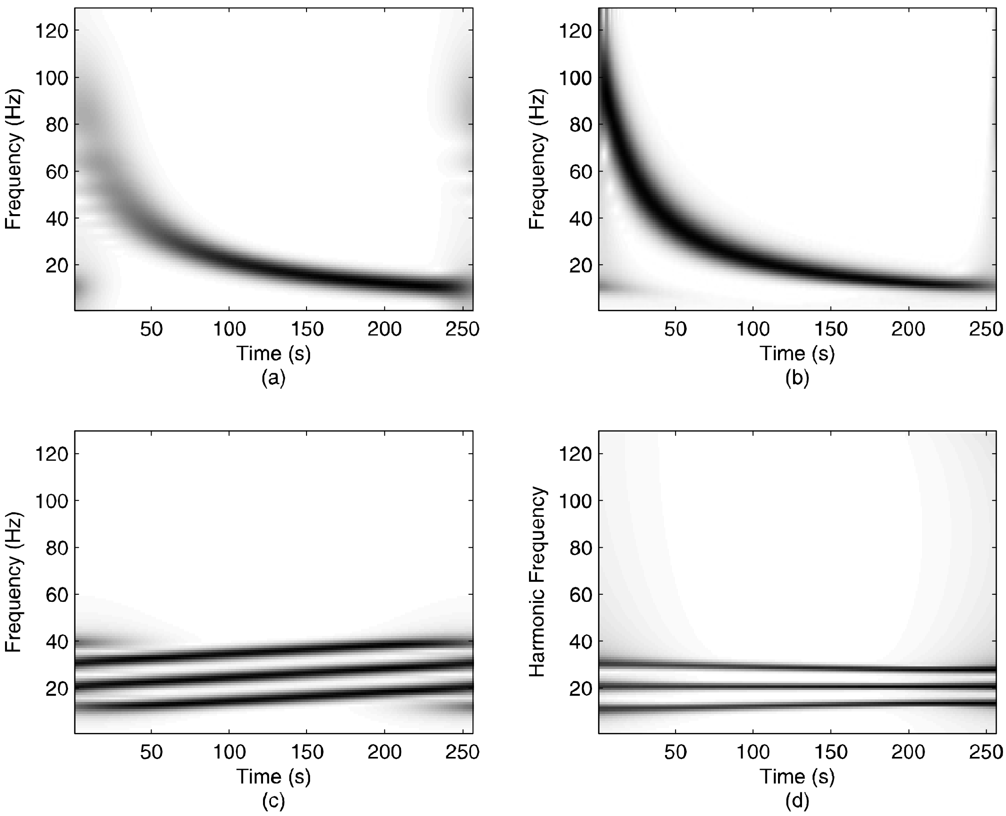

Fig. 4. A comparison of four signal decomposition techniques based TFRs: (a) STFT of a sample hyperbolic signal; (b) S-transform of a sample hyperbolic signal; (c) STFT of a sample harmonic signal; (d) STHRT of a sample harmonic signal.

form of a Hartley S-transform. The Hartley S-transform introduces a variable window width framework for the Hartley analysis [52]. However, only one window function is introduced as for the Fourier S-transform.

B. Feature Representation Based on Cohen's Class of TFR

A lot of research has been done for feature representation and extraction based on Cohen's TFR. Many significant contributions have been made and some are listed below. The attractiveness of these representations is based on the fact that, when cross terms and inner interferences are minimized, these transforms can produce very high resolution representations. A classical example is a TFA of a linear FM signal as shown in Fig. 2. The energy concentration obtained by Wigner distribution is significantly higher than the concentrations obtained by the STFT or the S-transform.

The problems with feature extractors based on Cohen's class are cross terms and inner interferences, which can lead to the ambiguous representation of a signal in the TF domain. Hence, most of the research conducted in this area attempts to reduce the effects of cross terms. The classification accuracy is significantly diminished by the cross terms, especially for multicomponent signals. The cross terms can be reduced or eliminated by introducing a kernel function $\phi(\theta, \tau)$. To show how different kernels can reduce the effects of the cross terms, let's rewrite Cohen's class of the TFRs in terms of the ambiguity function, $A(\theta, \tau)$. The ambiguity function is defined as [9]:

$$A(\theta, \tau) = \int_{-\infty}^{+\infty} x\left(u + \frac{1}{2}\tau\right) x^*\left(u - \frac{1}{2}\tau\right) e^{j\theta u} du \tag{7}$$

and the Cohen's class can then be rewritten as

$$TFD_x(t, \omega) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(\theta, \tau) \times \phi(\theta, \tau) e^{-j\theta t - j\tau\omega} d\tau d\theta. \tag{8}$$

This reformulation provides an easier understanding of the auto and cross terms location. The ambiguity function can be considered as a joint TF autocorrelation function. All auto terms are located along and around the ambiguity domain axis, and hence the maximum occurs around the origin. For the nonoverlapping components, the cross terms are dislocated further from the axis [22].

The framework of reduced interference distribution (RID), introduced in [53] and [54], summarizes the efforts of different kernels. Kernels are designed in the ambiguity domain as low-pass filters to suppress and eliminate the efforts of cross terms, and to obtain the desired properties of the TFRs. Some of the proposed distributions following the idea of the RID class are listed below:

- Born-Jordan distribution [9] with

$$\phi(\theta, \tau) = \frac{\sin(\theta\tau/2)}{\theta\tau/2}. \tag{9}$$

- Choi-Williams distribution [55] with

$$\phi(\theta, \tau) = \exp\left(-\frac{\theta^2\tau^2}{\sigma^2}\right) \tag{10}$$

where σ is a scaling factor.

- Zhang-Sato distribution [56] with

$$\phi(\theta, \tau) = \exp\left(-\frac{\theta^2\tau^2}{\sigma^2}\right) \cos(2\pi\beta\tau) \tag{11}$$

where σ and β are parameters. For $\beta = 0$ a Choi-Williams distribution is obtained, since σ is defined in the same manner as for the Choi-Williams distribution.

- Radial Butterworth Distribution [57] with

$$\phi(\theta, \tau) = \frac{1}{1 + \left(\frac{\theta^2+\tau^2}{r_o}\right)^M} \tag{12}$$

where r_o and M are adjustable parameters with constraints $r_o \neq 0$ and $M \in \mathbb{Z}^+$.

- Bessel distribution [58] with

$$\phi(\theta, \tau) = \frac{J_1(2\pi\alpha\theta\tau)}{\pi\alpha\theta\tau} \tag{13}$$

where J_1 is the first kind Bessel function of order one and $\alpha > 0$ is a scaling factor.

• Generalized exponential distribution [59], [60]

$$\phi(\theta,\tau)=\exp\left(-\left(\frac{\theta}{\theta_1}\right)^{2N}\left(\frac{\tau}{\tau_1}\right)^{2M}\right) \quad (14)$$

where N, M are positive integers, and θ_1, τ_1 are positive frequency and time scaling constants, respectively, chosen such that $\phi(\theta_1,\tau_1)=\exp(-1)$.

• Multiform tiltable exponential distribution [61] with

$$\phi(\theta,\tau)=\exp\left\{-\pi\left[\mu^2\left(\tau/\tau_o,\theta/\theta_0,\alpha,r,\beta,\gamma\right)\right]^{\lambda}\right\} \quad (15)$$

where

$$\begin{aligned}&\mu\left(\tau/\tau_o,\theta/\theta_0,\alpha,r,\beta,\gamma\right)\\&=\left(\tau/\tau_0\right)^2\left(\theta/\theta_0\right)^{2\alpha}+\left(\tau/\tau_o\right)^{2\alpha}\left(\theta/\theta_0\right)^2\\&\quad+2r\left\{\left[\left(\tau/\tau_o\right)\left(\theta/\theta_0\right)\right]^{\beta}\right\}^{\gamma}\end{aligned} \quad (16)$$

and the parameters have the following properties: α is a nonnegative power, λ is a positive power, τ_0 is a positive time lag scaling constant, θ_0 is a positive frequency lag scaling constant, r is a tilt or rotation given by $r\in[-1,1]$, and β and γ are coupled powers.

• S-method [39] with

$$\begin{aligned}\phi(\theta,\tau)=P\left(-\frac{\theta}{2}\right)*_\theta\int_{-\infty}^{\infty}w\left(u+\frac{\tau}{2}\right)\\ \times w^*\left(u-\frac{\tau}{2}\right)\exp(-j\theta u)du\end{aligned} \quad (17)$$

where $*_\theta$ represents a convolution in θ, $P(\theta)$ is a smoothing function and $w(t)$ is a window function used for the STFT.

• Distribution for multicomponent linear FM signals [62] with

$$\phi(\theta,\tau)=\Pi\left(\frac{\theta-\chi\tau}{b}\right) \quad (18)$$

where χ is a frequency modulation rate, b is the width in the direction of θ and $\Pi(\xi)=1$ for $|\xi|\le 1/2$.

• A time-lag kernel distribution [63]

$$\phi(\theta,\tau)=|\tau|^{\alpha}\frac{2^{2\alpha-1}}{\Gamma(2\alpha)}\Gamma(\alpha+j\pi\theta)\,\Gamma(\alpha-j\pi\theta) \quad (19)$$

where α is a bounded parameter such that $0<\alpha\le 1$, and $\Gamma(z)$ is the Gamma function of z.

• Hyperbolic distribution [64]:

$$\phi(\theta,\tau)=\frac{1}{\cosh(\beta\theta\tau)} \quad (20)$$

where β is a parameter to control the exponential terms of the hyperbolic function.

Furthermore, two subclasses of RID based TFDs are also proposed for discrete signals [65], [66]. The RID kernels which can be implemented recursively are proposed in [65]. These kernels allow simultaneously recursive implementations of the local autocorrelation. In [66], high resolution kernels based on the Prony's method are introduced.

It is important to mention that all the kernels presented above, except the kernel for the Born-Jordan distribution, contain one or more adjustable parameters. This implies that for a given kernel the parameter(s) can be chosen such that the resulting kernel produces a representation similar to a representation obtained by some other kernel with the same number of parameters. Having the opportunity to "fine tune" the kernel generally represents an advantage for feature extraction. In a given application, the kernel can be optimized to achieve maximal reduction of the cross term effects. As an example, variations of some of the parameters for the kernel proposed in [61] are depicted in Fig. 5. However, finding a proper value of the parameter(s), yielding the highest energy concentration in the TF domain, can also represent an additional computational burden.

It should be mentioned that not every kernel can produce satisfactory results in all applications. Some kernels are only proposed for certain specific classes of signals, such as the kernel defined by (18) [62]. In addition, it should be noted that the Cohen's class of representations can only achieve the ideal TFR of the signal if the IF of the signal is a linear function (e.g. a linear FM signal) [67] as depicted in Fig. 2. If the IF variations are of higher order, no signal independent distribution from Cohen's class can produce the ideal representation [67]. Therefore, it is worthwhile to mention the generalization of Cohen's class

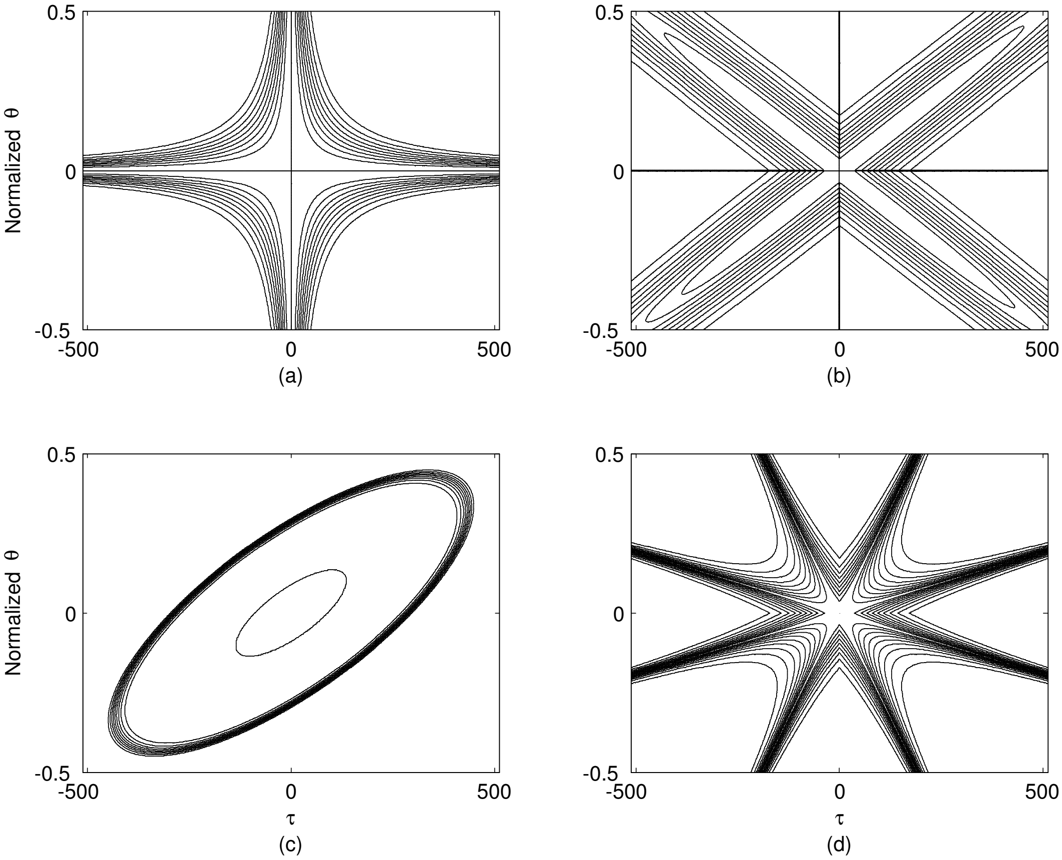

Fig. 5. The tiltable, generalized exponential kernel for various values of parameters: (a) $\lambda = 1/2$, $\alpha = 1$, $r = 0$, $\beta = 1$, $\gamma = 1$, $\tau_0 = 200$, $\theta_0 = 0.2$; (b) $\lambda = 1/2$, $\alpha = 0.002$, $r = -1$, $\beta = 2$, $\gamma = 1/2$, $\tau_0 = 200$, $\theta_0 = 0.2$ (c) $\lambda = 8$, $\alpha = 0$, $r = -0.75$, $\beta = 1$, $\gamma = 1$, $\tau_0 = 300$, $\theta_0 = 0.3$ (d) $\lambda = 1/2$, $\alpha = 0$, $r = -1.5$, $\beta = 2$, $\gamma = 1/2$, $\tau_0 = 200$, $\theta_0 = 0.2$.

representations proposed in the form of the L-class distributions [68] - [73] in the context of feature extraction for signals with a higher order IF variation. These distributions represent higher order representations, i.e., the order higher than second, with diminished inner interference effects and enhanced resolution in comparison to the Cohen's class. The problem of the cross terms becomes more profound. However, these cross terms can be diminished or completely eliminated by careful recursive implementation of a L-class distribution by using the STFT [68]. Some further improvements are proposed in the forms of a pseudo-quantum signal representation [74], and a "complex time" TFD [75], [76].

In addition to reducing the effects of cross terms, the kernels presented here have other properties on the resulting TFRs. These properties are usually selected in advance by the designers. Furthermore, there exist design methods for constructing new kernel functions with specific application oriented properties. A summary of some kernel design methods is given in Table II.

C. Rotation of the TF Plane

The feature extractors based on the rotation of the TF plane have been introduced to improve energy concentration for signals whose components are not aligned with either the time or the frequency axis [83]. As an example, let's compare the TFRs obtained by the rotation of the TF plane with some standard

TABLE II
Properties of the kernel design methods in the literature.

Kernel Design Method	Advantages	Disadvantages
POCS method [77]	Two or more design constraints can be satisfied simultaneously.	Constraints have to be chosen carefully, otherwise questionable results may be obtained.
Frequency transformation method [78]	Produced kernels can have efficient cascade implementation.	Not every kernel produced by the FTM is amenable to cascade filter implementation in the time-frequency plane.
Design via point and derivative constraints [79]	Kernels with various constraints could be easily constructed.	Only applicable to discrete kernels. The design procedure may yield a conflict between time and frequency marginal properties.
Kernels with desired auto-term properties [80]	Kernel is optimized for the signal auto-term.	It has to be recalculated for every class of signals.
Minimum variance kernels [81], [82]	Kernel satisfies the TF constraints and provides the minimum variance for the power spectrum estimate for the Gaussian white noise processes [81] or additive circular complex white noise processes [82].	Only minimizes the average variance. The method is optimal for noisy signals.

approaches presented earlier, e.g. the STFT and Wigner distribution. Let's assume a sample signal consisting of three linear FM components. The Wigner distribution (WD) is capable of achieving the ideal energy concentration of the linear FM signal as shown in Fig. 2. However, in this case, the TFR obtained by WD suffers from the effects of cross terms as shown in Fig. 6(b). The advantage of the STFT in this case is that it does not contain cross terms. However, the energy concentration of each component is severely degraded in comparison to the representation obtained by the Wigner distribution. The TFR obtained by the local polynomial Fourier transform (LPFT) enhances the concentration of the components in comparison to the STFT, and it does not contain any cross terms as shown in Fig. 6(d).

The TFA based on the rotation of the TF plane can be achieved in several ways:

- Fractional Fourier transform (FRFT) [84], [85], [86]:

$$F_\alpha(u) = \begin{cases} \sqrt{\frac{1-j\cot\alpha}{2\pi}}e^{j(u^2/2)\cot\alpha} \times \int_{-\infty}^{+\infty} x(t) \times e^{j(t^2/2)\cot\alpha - jut\csc\alpha}dt & \alpha \bmod \pi \neq 0 \\ x(t) & \alpha \bmod 2\pi = 0 \\ x(-t) & (\alpha+\pi) \bmod 2\pi = 0 \end{cases} \tag{21}$$

The standard Fourier transform is a special case of the FRFT with a rotation angle $\alpha = \pi/2$.

- Local polynomial Fourier transform (LPFT) [87] - [91]:

$$LPFT_x(t,\overline{\omega}) = \int_{-\infty}^{+\infty} x(t+\tau)w(\tau)\times \exp\left(-j\omega_1\tau - j\omega_2\tau^2/2 - ... - j\omega_M\tau^M/M\right)d\tau \tag{22}$$

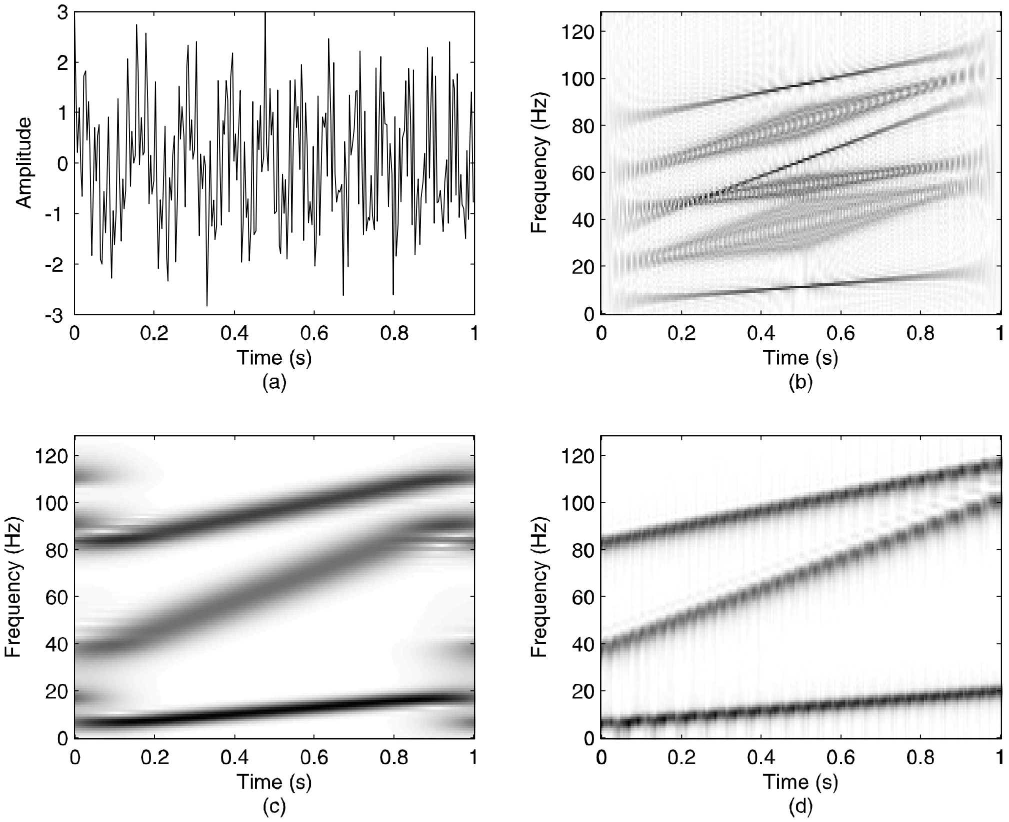

Fig. 6. Sample signal analysis with several TFRs: (a) time-domain representation of a sample signal consisting of 3 linear FM components; (b) Wigner-Ville distribution of the sample signal; (c) STFT representation of the sample signal; (d) LPFT representation of the sample signal.

where $\overline{\omega} = (\omega_1, \omega_2, ..., \omega_M)$. The LPFT enables one to estimate both the time-varying frequency and its derivatives. The technique is based on fitting a local polynomial approximation of the frequency which implements a high-order nonparametric regression.

• Radon-Wigner distribution (RWD) [92] - [95]:

$$RWD(r, \vartheta) = \mathcal{R}\left[WV_x(t, \omega)\right] = \int WV_x(t, \omega_o + mt)dt\bigg|_{m=-1/\tan(\vartheta);\omega_o=r/\sin(\vartheta)} \tag{23}$$

where $\mathcal{R}\left[f(x, y)\right] = \int f(r\cos\vartheta - s\sin\vartheta; r\sin\vartheta + s\cos\vartheta)ds$ and r and s represent x and y axes rotated counterclockwise by an angle ϑ.

Summary of some properties associated with these approaches is given in Table III. It should be mentioned that the FRFT corresponds to the rotation of a class of TFRs as along as $\Psi(t, f) = \mathcal{F}^{-1}_{\theta\to t, \tau\to f}\left\{\phi(\theta, \tau)\right\}$ is rotational symmetric [96]. The FRFT based TFRs also have marginals associated with them [97] in analogy to the TFRs based on the standard Fourier transform.

Even though the RWD is considered a tool for the rotation of the TF plane at a certain angle, the RWD was developed primarily for detection and classification of multicomponent linear FM signals in noise. This approach reduces the task of tracking straight lines in the TF plane to locating the maxima in a 2-D plane. It is also interesting to mention that the ambiguity function can be obtained as an inverse Fourier transform of the RWD.

The presented approaches for rotation of the TFRs are similar in principle. The relationship

TABLE III

Properties of the approaches for the rotation of the TF plane.

Approach	Advantages	Disadvantages
FRFT	Allows representation of a signal on the orthonormal basis formed by chirps.	$\cot(\alpha)$ can take enormous values and oversampling may be needed to satisfy the sampling theorem.
LPFT	Provides generalization of the FRFT to any order of the polynomial IF.	A drawback of the LPFT is the increase in dimensionality, i.e., an increase of the calculation complexity.
RWD	Excellent for establishing the direction of the linear FM modulated signal in the ambiguity plane.	Not suitable for long data records, and the segmentation of such records is needed. Analyzed in depth only for the WD.

between FRFT and RWD has been studied in [98], and it is shown that the Radon-Wigner distribution is the squared modulus of the fractional Fourier transform:

$$RW\left[x(t)\right] = \left|FRFT\left[x(t)\right]\right|^2. \qquad (24)$$

To establish the relationship between the FRFT and the LPFT, the FRFT can be written as:

$$F_\alpha(u) = \sqrt{\frac{1-j\cot\alpha}{2\pi}} e^{j(u^2/2)\cot\alpha}$$
$$\times \int_{-\infty}^{+\infty} x_w(\tau) e^{j(\tau^2/2)\cot\alpha - ju\tau\csc\alpha} d\tau \qquad (25)$$

where $x_w(\tau) = x(t+\tau)w(\tau)$. For $M = 2$, $\omega_1 = u\csc\alpha$, and $\omega_2 = \cot\alpha$ in (22), equation (25) can be expressed in terms of the LPFT as:

$$F_\alpha(u) = \sqrt{\frac{1-j\cot\alpha}{2\pi}}$$
$$\times e^{j(u^2/2)\cot\alpha} LPFT_x(t,\omega_1,\omega_2). \qquad (26)$$

From these equations it can be seen that the LPFT provides a broad generalization of the FRFT.

Several different feature extractors have been proposed using the rotated TF domain framework. A fractional-Fourier-domain realization of the weighted Wigner distribution (i.e. S-method) [99] and of Gabor expansion [100]-[103] are introduced in several publications. The LPFT is also implemented for a polynomial Wigner distribution [104], and the extension to the L-Wigner distribution is presented in [105]. Several other generalizations to and modifications of the rotated TFA are also proposed in the literature such as: unitary similarity transformations [106], a four-parameter atomic decomposition of chirplets [107], joint fractional representations [108], [109], generalization of the FRFT into the linear canonical transform [110], and the tomography TF transform defined as the inverse Radon transform of the FRFT [111]. Also, efficient algorithms to compute uniformly spaced samples of the Wigner distribution and the ambiguity function located on arbitrary line segments are proposed in [112][113].

D. Signal Dependant TFRs

The feature extractors described in the previous sections deal with several concepts regarding the improvement of energy concentration: reducing the effects of spectral leakage; diminishing the effects of cross terms; and aligning the axis of analysis with the principal axis of the signal components. However, can a single feature extractor be optimal for all signals? Unfortunately not, since a major drawback of all fixed mappings is that, for each mapping, the resulting TFR is satisfactory only for a limited class of signals. Thus, the enhanced concentration in the TF domain is desirable for a variety of classes of signals. Concentrated components generally overlap or interfere with other nearby components as little as possible, and yield a "sharp" representation. The maximal concentration also implies that components are confined as closely as possible to their proper support in the TF domain. Hence, this is why signal dependent TFRs are important. It has to be mentioned that these techniques are generally nonlinear and non-quadratic due to the nature of the computation process. In this subsection, an overview is provided only for signal dependent representations, which are based on the two classes of the TFRs mentioned in the Introduction.

The signal dependent TFRs are available in several forms in the literature. These representations differ in their adopted forms. They are based on:

- concentration measures [40] - [124]
- reassignment methods [125] - [128]
- signal optimized kernels/windows [129] - [142].

Some properties of each approach are summarized in Table IV.

The concentration measure approach examines the effects of certain parameter variations on the energy concentration of the signal in the TF domain. The parameter value yielding the highest energy concentration is chosen for the signal dependent TFR. The development of the concentration measure can be divided into two groups based either on the distribution norms or on the entropy of the distributions. The initial research in the development of the measures based on the distribution norms has been carried out by Jones and Parks [40], [114]. They proposed a measure based on the STFT for signal concentration that allows the fully automated determination of the optimal basis parameters. The concentration measure (CM) is given by:

$$CM = \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |STFT(t,\omega)|^4 \, dt d\omega}{\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |STFT(t,\omega)|^2 \, dt d\omega\right)^2}. \tag{27}$$

The concentration measure in (27) favours those components with higher concentration. However, for multicomponent signals, a local measure is required to determine the concentration of the dominant component at each location in the TF domain. Eqn. (27) can be turned into a local concentration measure by multiplying the squared magnitude of the short-time Fourier transform by a "localization weighting function" [115].

A solution to the problem in the Jones-Parks measure is proposed by Stanković. The concentration measure proposed in [116] does not discriminate low concentrated components with respect to the highly concentrated ones within the same distribution, and it is given by:

$$CM = \left(\sum_{k=1}^{N}\sum_{n=1}^{N} |TFR_x(n,k)|^{1/p}\right)^p \tag{28}$$

where $TFR_x(n,k)$ is a discrete version of any of the TFRs.

A different notion of the quantification of the TFR appeared in the literature around the same time as the Jones-Parks measure. Williams *et al* considered how the information measures, such as the Shannon or Rényi information measure, could be used to provide information on TFDs [117]. The Shannon information measure is appropriate only for positive TFRs. The Rényi measure conforms closely to the visually based notion of complexity when inspecting TFRs and can be used for other TFRs [118]. For Cohen's class of the TFRs, the Shannon information measure is given as

$$H(TF_x(t,\omega)) = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} TF_x(t,\omega)\log_2 TF_x(t,\omega) dt d\omega \tag{29}$$

and the Rényi measure as

$$R_\alpha(TF_x(t,\omega)) = -\frac{1}{1-\alpha}\log_2 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} TF_x(t,\omega) dt d\omega \tag{30}$$

where $\alpha > 0$, and the Shannon entropy is recovered as the limit of R_α, as $\alpha \to 1$. A detailed study of the properties and some potential applications of the Rényi TF information measures, with emphasis on the mathematical foundations for quadratic TFRs can be found in [121]. It should also be noted that the Rényi measure is sensitive to the amplitude and phase variations in the signal components

[119]. However, it has been shown that the expected value of the third-order Rényi entropy has well defined upper and lower bounds in the presence of white noise [122]. Effects of three concentration measures on the TFR of a signal consisting of a sinusoidally FM and linear FM components, $x(t) = \exp(j20\pi t + j30\pi t^2) + \exp(j5\pi\cos(4\pi t) + j150\pi t)$, are depicted in Fig. 7.

It is also necessary to mention a resolution performance measure [123], [124]. The resolution performance measure allows the design of high-resolution TFRs for multicomponent signals. However, this measure requires extensive knowledge of the signal and representation attributes such as: the average amplitudes of the mainlobes; sidelobes; cross terms; and the components relative frequency separation of any two consecutive components of multicomponent signals. Thus, it may be difficult to implement in practice.

The energy concentration of the signal components in the TF domain is tackled from another perspective by a so-called reassignment method. The reassignment method, initially proposed in [125] for a spectrogram, and later on, generalized for any TF method in [126], [127], [128], creates a modified version of a representation by moving its values away from where they are computed to produce a better localization of the signal components. In order to perform such an operation, for each point in the TF plane, one calculates the center of gravity for the signal energy such as:

$$\widehat{t}(t,\omega) = t - \frac{\int\int uTFR(t-u,\omega-\Omega)dud\Omega}{\int\int TFR(t-u,\omega-\Omega)dud\Omega} \quad (31)$$

$$\widehat{\omega}(t,\omega) = \omega - \frac{\int\int \Omega TFR(t-u,\omega-\Omega)dud\Omega}{\int\int TFR(t-u,\omega-\Omega)dud\Omega}. \quad (32)$$

Given these centers of gravities, the reassigned TFR is obtained by

$$RTFR(t,\omega) = \int\int TFR(\tau,\upsilon)\delta(t-\widehat{t}(\tau,\upsilon))\delta(\omega-\widehat{\omega}(\tau,\upsilon))d\tau d\upsilon \quad (33)$$

where $\delta(t)$ is a Dirac function. However, it is noticed that the technique is highly sensitive to noise, and some modifications to the original algorithm have been proposed [143] - [146]. The reassignment method is also computationally expensive. A fast algorithm that allows the recursive evaluation of TFDs modified by the reassignment method is introduced [147].

The first two approaches to signal dependent TFRs are based upon the fact that an optimized representation is found for each new signal. Another stream of research in this area is based on the development of the signal dependent kernels/windows for a class of signals through an optimization design procedure. The initial research has been conducted for so-called radially Gaussian distributions [129], [130]. The problem of finding the optimized kernel boils down to finding the optimal $\sigma(\psi)$ for radially Gaussian functions for the given signal. Therefore, the optimization problem can be posed as:

$$\max_{\Phi} \int_0^{2\pi} \int_0^{\infty} |A(r,\psi)\Phi(r,\psi)|^2 \, r dr d\psi \quad (34)$$

with a constraint that the energy of $\Phi(r,\psi)$ must be finite, where $r = \sqrt{\theta^2+\tau^2}$ and $A(r,\psi)$ is the ambiguity function of the signal in the polar coordinates. The technique performs well in the presence of additive noise, which suggests that it may prove useful for the automatic detection of unknown signals in noise. A generalization of the idea to any other type of kernels is shown in [131], and in [133], where the kernel is further optimized locally for each signal component. Computationally effective procedure for the optimal kernel design is given in [132], and a procedure that adapts the kernel over time is presented in [134]. Similar approach based on the idea that the kernel should be optimized for classification has been proposed in [41]-[141],[148],[149]. The idea is that once the kernel is optimized to extract discriminant features among different classes, the classification process will yield more accurate results. The optimal kernel for classification, Φ, is the solution as given below:

$$\widehat{\Phi}(\theta,\tau) = \arg\max_{\Phi} d(TFR_1, TFR_2) \quad (35)$$

where $d(TFR_1, TFR_2)$ is a distance between the two TFRs, TFR_1 and TFR_2 represent

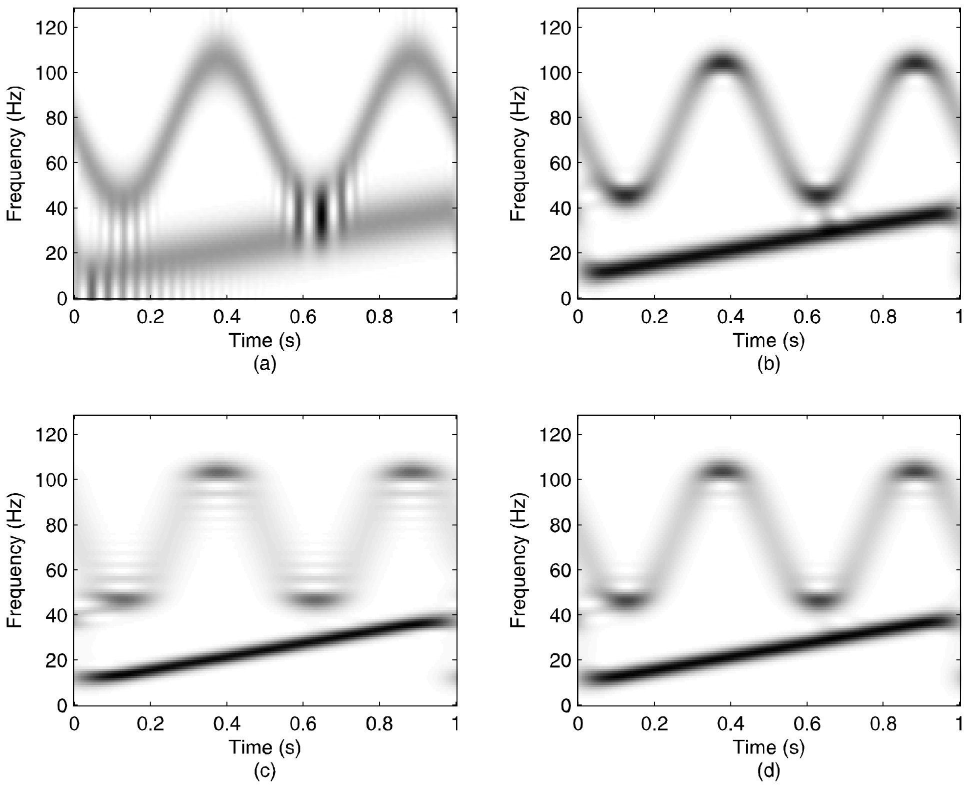

Fig. 7. Several TFRs of a sample signal consisting of a linear FM component and sinusoidally modulated component: (a) spectrogram (b) spectrogram according to the Stanković measure; (c) spectrogram according to the Jones-Parks measure; (d) spectrogram according to the Rényi entropy.

TABLE IV
Some properties of the approaches for obtaining signal dependent TFRs.

Approach	Advantages	Disadvantages
Concentration measure	Usually easy to implement. Good energy concentration can be obtained.	It has to be calculated for each signal.
Reassignment methods	Excellent energy concentration can be obtained.	Computationally expensive. Sensitive to noise.
Signal optimized kernels/windows	It does not need recalculation for every signal, but it is rather based on class of signals.	Needs careful implementation when working with signals in noisy environment.

TFRs of two signals belonging to different classes. Also a variety of distance measures can be implemented such as: Euclidian distance, correlation, a broad family of dissimilarity measures that is given by the f-divergences (e.g. Kolmogorov distance and Bhattacharyya distance) and the L_q distances based on the normalized TFRs.

While developing the signal optimized windows or kernels, it is important to mention that the bias and the variance of the estimated signal parameters in the presence of noise is dependent on the window/kernel used [142]. Hence, choosing appropriate parameters for the window/kernel is critical in order to achieve accurate estimation. In particular, the optimal choice of the window size based on asymptotic formulas for the bias and the variance can resolve the bias-variance tradeoff usual for nonparametric estimation. However, in practice, such an optimal estimator is difficult to implement because the optimal window size depends on the unknown smoothness of the IF. In [142] an algorithm is presented, which determines a time-varying data-driven window size for local polynomial periodrogram. The algorithm is then able to provide an accurate estimate that is close to what can be achieved if the smoothness of the IF is known in advance. The developed algorithm uses only the formulas for the variance of the estimate. This approach has also been applied to other TFRs as shown in Table V.

III. Signal Classification/Recognition Based on Energy Concentration in the TF Domain

In signal processing, linear or non-linear transformations are used to enhance features for improved classifications [178]. The previous section discussed how to extract the energy concentration of signals in the TF domain. Classification/recognition based on the extracted features will be discussed in this section. In situations where a statistical model (such as Gaussian distribution) is known, the optimal classification procedure can be developed. Often, however, no statistical model is available. In these cases, the application of the optimal classifier would require an estimation of the relevant probability density functions. Hence, a large set of signal realizations may be required for learning purpose [6]. If the set is small, suboptimal procedures may have to be used. As pointed out in the Introduction, for the nonstationary signals it is necessary to use a model-free representation space in which the differences between different features are emphasized and the similarities are de-emphasized [178].

A. TFA in Classification Process

TFR-based classification methods are preferred because TFRs have discriminant capabilities for signals belonging to different signal classes. This situation is often encountered in practical applications [149]. Also, the main advantage of the TF domain based classification is the flexibility to form the feature vector in 2D representations. The question is how to perform classification/recognition based on energy concentration in the TF domain.

Before analyzing possible approaches, let's consider sample energy concentration patterns depicted in Fig. 8. The patterns represent phenomena, which are manifested through short duration transients. These patterns can be nonoverlapping as shown in Fig.s 8(a)-(c) or overlapping as shown in Fig. 8(d). The nonoverlapping patterns can be easily classified through frequency or time domain filtering. However, what happens if the two sample patterns are overlapping in frequency and time domain, such as Fig. 8(d)? Classification of such patterns becomes more involved either in frequency or time domain alone. In such a situation, the energy concentration in the TF domain can effectively be used as the feature for classification purpose. The classification based on energy distribution in the TF domain can be performed in two ways:

- by visual inspection of the patterns in the TF domain;
- by development of classification schemes.

It has been shown in [179] - [229] that the differences amongst different patterns can be best revealed in the TF domain. However, in some cases the differences are not always obvious with all the feature extractors presented. For example, in the analy-

TABLE V
DIFFERENT TFRS DEVELOPED BASED ON THE SIGNAL DEPENDENT TFRS APPROACHES.

Approach	TFR
Concentration measures	Optimization of various TFDs [120], [122],[150]-[155]. Also signal dependent TFR analysis based on the FRFT [156]-[158], the LPFT [159]-[162], and the Radon-Wigner transform [163].
Reassignment methods	STFT, wavelet transform, pseudo Wigner distribution, smoothed pseudo Wigner distribution, RID [125]-[128], S-method [164].
Signal dependent kernels or windows	Signal dependent kernels/basis for various representations [165]-[170]. The optimal choice of the window length based on the asymptotic formulae for the variance and bias is used for: the pseudo Wigner distribution [171]-[174], L-Wigner distribution [175], robust M-periodogram [176], spectrogram [177].

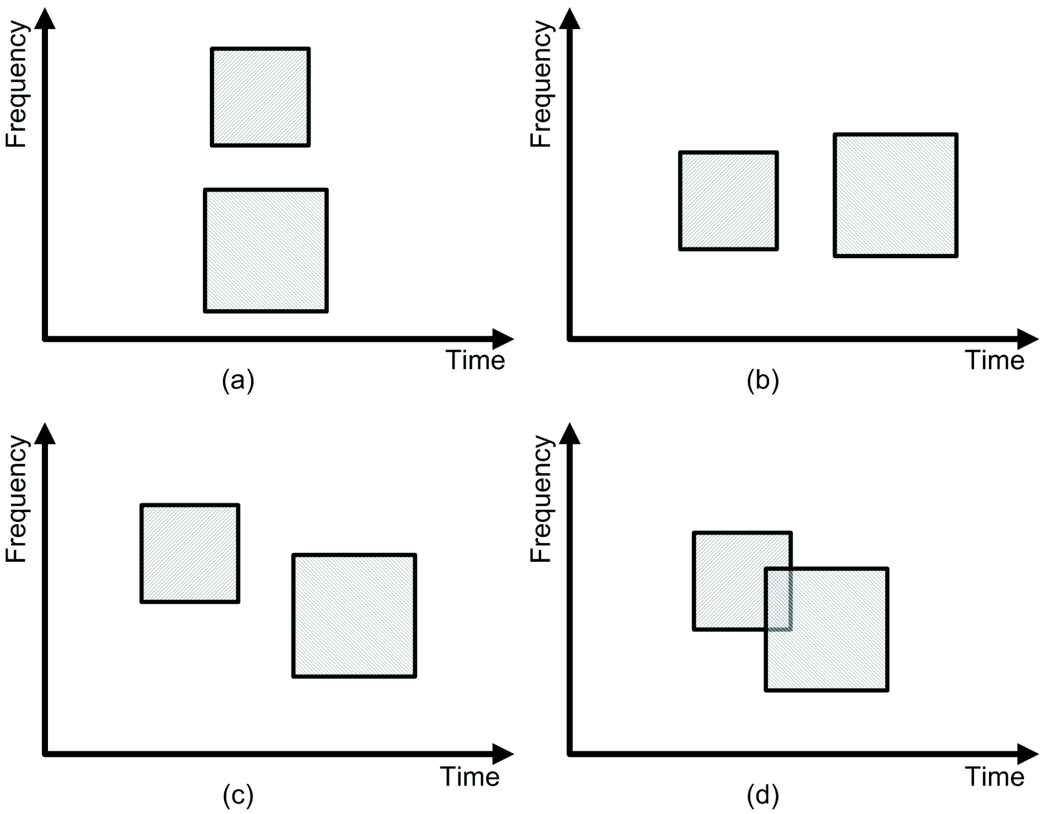

Fig. 8. Illustrative pattern scenarios in the TF domain: (a) patterns occupying the same time band; (b) patterns occupying the same frequency bands; (c) patterns partly occupying the same frequency band, but not intersecting; (d) patterns overlaping on some time and frequency bands.

sis of some heart sounds, it has been noticed that the S-transform provides visual representation emphasizing the morphological differences amongst the sounds with a sharper time-frequency concentration than the STFT or the continuous wavelet analysis [192], [203]. Does it mean that the S-transform is the optimal feature extractor for heart sounds? Not necessarily, since some feature extractors from Cohen's class can also provide a sharp TF concentration of the same sounds, if the effects of the cross terms are eliminated [229]. In addition to the choice of a suitable feature extractor, this approach has other limitations. First, it is not an automated decision process. It relies on human expertise, and also requires some initial training to recognize the differences among patterns. Furthermore, consecutive classifications require human intervention. Hence, they are difficult to be implemented as a stand alone software/hardware product. Needless to say, such decision process is prone to human errors.

The second approach to feature classification relies on an automated feature classifier, which makes independent decisions. Such a classifier makes the decision based on features represented in terms of energy concentration. The decision making process is usually based on statistical differences among patterns [230] - [256] or distance measures among patterns [257] - [262]. The statistical differences among patterns can be measured in several ways such as correlation [248], [256], linear discriminant analysis [241], mutual information [242], to name a few. It should noted that the choice of a feature extractor can have significant influence on the final results: some are better, and some are worse [248] [256]. The implementation of distance measures as feature classifiers can be viewed as a mathematical extension of the classification based on visual inspection. The extracted patterns are simply classified based on the "distances" from the given templates for different classes. The choice of feature extractor is spread across the spectrum of the extractors presented in Section II, such that the signal decomposition based TFRs [259], [261], Cohen's TFRs [258], [259], [260], [262] and the signal dependent TFRs [257] can all be used.

The feature classification based on statistical differences or distance measures can be seen as a favourable approach. A logical question is which of these classifiers can lead to the most accurate results. The answer is rather difficult. The accuracy depends on applications. Choosing an effective template often requires familiarity with the problem. Also, accuracy depends on the choice of feature extractor used.

References sorted according to different fields of applications, which use the energy concentration in the TF domain as features are summarized in Table VI. The columns represent the four types of feature extractors. It is interesting to note that some feature extractors such as the rotated TFRs have limited fields of applications.

An example with two simple templates is used to show the advantage of the TF based classifiers over their time domain counterparts. The time domain and the TF domain representations of the templates are depicted in Fig. 9. The templates have identical low frequency content. The transients present in the signal denote two different phenomena, which are desired to be classified. TF boundaries of the transient parts are given by $T_1 = \{(t,\omega) : t \in [0.54, 0.6], \omega \in [120\pi, 180\pi]\}$ and $T_2 = \{(t,\omega) : t \in [0.55, 0.65], \omega \in [100\pi, 160\pi]\}$, respectively. Furthermore, each phenomenon consists of three short duration sinusoids with frequencies within the frequency boundaries defined by the templates.

Unknown signals are generated with equal probability of belonging to either class. These signals have the same low frequency content as the templates. The frequencies of the three short duration sinusoids are generated with uniform probability for the given sets. The signals are classified with the time-domain based Euclidean distance and the TF domain based Euclidean distance [257]. The distance between the signals and templates are calculated. The classification is done based on the shortest distance between the signals and the respective templates. An error rate, defined as the incorrect classification of the unknown signal, is calculated for 10000 trials. The results show that the time domain classification produces an er-

TABLE VI
REFERENCES SORTED ACCORDING TO APPLICATIONS AND FEATURE EXTRACTORS USED.

Application	Signal Decomposition TFR	Cohen's TFR	Rotated TFR	Signal Dependent TFR
Biomedical Signal Analysis	[182] [183] [184] [192] [195] [197] [200] [202] [203] [210] [225] [227] [231] [235] [248] [249] [250] [251] [252]	[179] [180] [181] [185] [186] [190] [191] [193] [206] [229] [232] [258]		[196] [208] [239] [242] [257]
Mechanical Signal Analysis	[188] [189] [198] [204] [207] [209] [220] [221] [222] [226] [228] [236] [255] [261]	[205] [214] [215] [216] [220] [233]		[194] [212] [223]
Power Systems Analysis	[213] [240] [243] [244]	[218] [262]		
Speech and music processing	[254] [259]	[247] [259]		[241] [259]
Radar and sonar signal processing	[201] [217] [230]	[187] [199] [211] [238] [245] [253] [260]	[219] [246]	

ror rate of approximately 33 %. The TF classifier produces an error rate of approximately 11 %, which is three times smaller than the time domain classifier.

IV. Feature Extraction Error Analysis: An Application Example of IF Estimation

It is well known that the choice of a feature extractor affects the classification accuracy. The effects can be as simple as a limited resolution obtained by a representation, but can be as complicated as nonlinearities of IF of a signal. To diminish these effects different representations have been introduced as shown throughout Section II. However, the question still is how accurately a representation can extract energy concentration. The answer to this question lies in the error introduced by the extractor in the classification process. Therefore, it is desirable to understand the estimation error introduced by a TFR in order to approximate the minimum classifier resolution.

The rest of this section provides a description of an approach that examines the extraction accuracy of TFRs. The focus is on the IF estimation based on the maximum of energy concentration. However, for the sake of completeness, a quick overview of other TF based estimation methods is given as well. Interested readers should refer to [263] and [264] for details.

A. Estimation of IF Using TFA

In some applications, the accurate estimation of the maximum of energy concentration is important for two reasons. First, it is well known that the location of the maximum energy concentration in the TF domain corresponds to the IF of a signal [9]. Second, the IF can be used as a mean to classify different phenomena (e.g. [239]).

The problem of estimating the IF using the TF techniques has been studied extensively

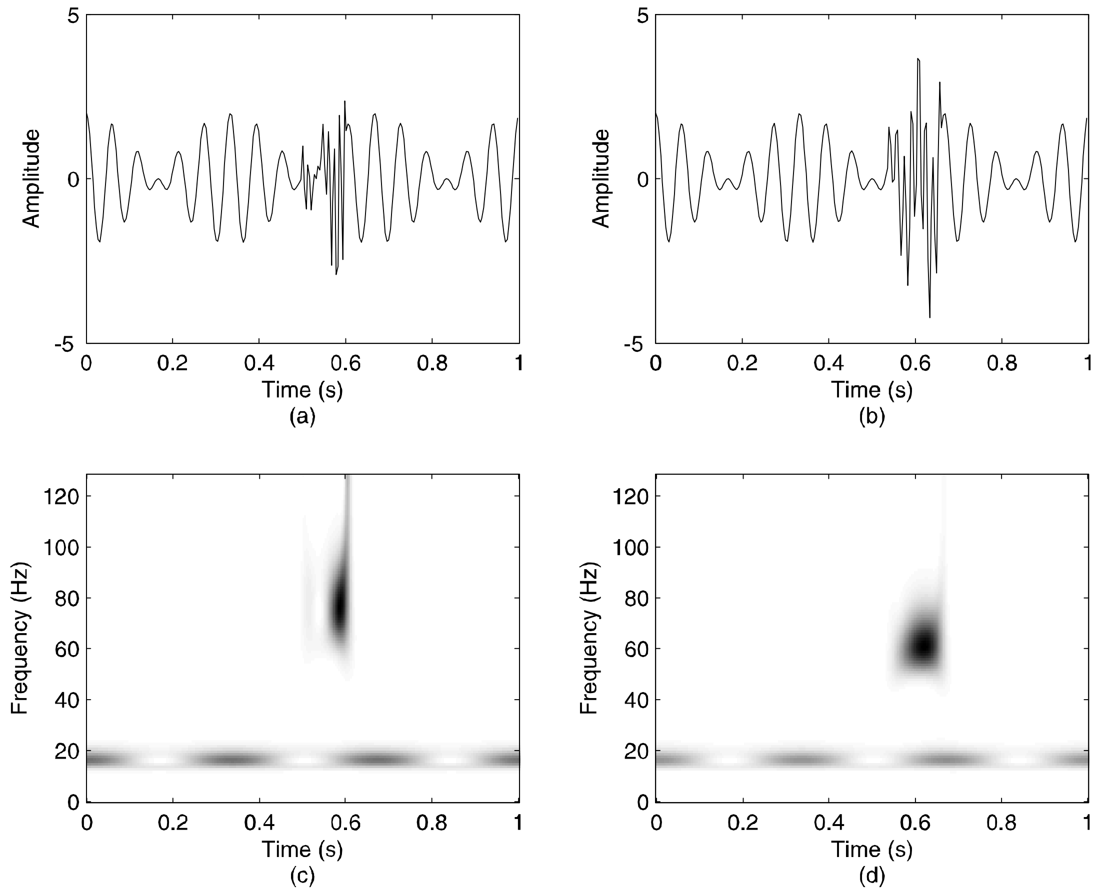

Fig. 9. Time domain and the TF domain representations of two templates: (a) time domain representation of the first template; (b) time domain representation of the second template; (c) TFR of the first template; (d) TFR of the second template.

in past years [70], [90], [142], [171], [172], [174],[175], [177], [265]-[307]. The IF can be estimated as a first moment of the TFR

$$\omega(t) = \frac{\int_{-\infty}^{+\infty} \omega TF_x(t,\omega)d\omega}{\int_{-\infty}^{+\infty} TF_x(t,\omega)d\omega} \tag{36}$$

or based on the position of the maximum value of the energy concentration in the TF domain as

$$\omega(t) = \arg\max_{\omega} \left[|TF_x(t,\omega)|\right]. \tag{37}$$

The first moment provides an unbiased estimate of the IF of a signal [268],[276]. The presence of additive noise leads to the serious degradation of the first moment estimate. It may have a high statistical variance even at high values of input SNR [269]. The first moment estimate is not affected by the multiplicative noise [300]. The maximum value estimate is greatly affected by the multiplicative noise when the power spectral density of the noise has a maximum at a frequency other than DC [300].

The maximum value estimate is hence used for the signals contaminated with the additive noise. It is based on the detection of a distribution maxima positions. This estimate is also prone to some estimation errors. The sources of estimation error are:

- bias;
- random deviation of the maxima within the auto-term caused by a small noise;
- large random deviations due to false maxima

detection outside the auto-term caused by a high noise.

In [171], authors have developed an approach to examine effects of the first two estimation errors for signals contaminated with the additive noise. They showed that the estimator bias and variance are highly signal dependent. Also, the bias generally caused by the IF non-linearity is proportional to a power of the lag window length. The variance caused by the noise is a decreasing function of the lag window length. Thus, the bias-to-variance trade-off exists, producing the minimal mean squared error. The effects of large random deviation due to false maxima detection outside the auto-term caused by the high noise are considered in [289],[304]. This error occurs when some points outside the signals' auto-term surpass values inside the auto-term, due to the influence of a relatively high noise. It has been shown that this kind of error, when it appears, dominates over other sources of error.

The approach based on the examination of the estimation error due to bias and random deviations within auto-term has also been used to examine the IF estimator based on the maximum of the energy concentration for various TFRs such as the L-Wigner distribution [175], spectrogram [294], reduced interference distributions, the L-class, and signal-dependent optimal TFRs [297], shift covariant class of quadratic TFDs [302], the S-method [301]. This approach is also extended to a combination of multiplicative and additive noise for pseudo Wigner-Ville distribution [291], and similar results are obtained. However, when the standard deviation of the multiplicative noise is larger than its mean, the noise can deteriorate the phase of the signal significantly making the use of TF techniques difficult [286].

V. Remarks and Future Perspectives

This paper provides an overview of methods dealing with energy concentration in the TF domain. The scope of the paper is restricted to only the methods that are based on analytical algorithms, that is, artificial intelligence based algorithms have not been considered for space reasons.

The theoretical developments behind the different extractors are comprehensive. Based on the reviewed literature it is difficult to foresee major contributions changing the field drastically in years to come. Our expectation is that most of the focus will be given to higher order representations briefly mentioned in Section II-2. These transformations provide high concentration representations of the signals with higher order IF modulations. Their significance will be especially pronounced in fields like spectroscopy, radar signal analysis, optics, and biomedical signal processing in years to come.

On the contrary to feature extraction, feature classification in the TF domain still lacks comprehensive development. The variety of practical problems requiring different classification approaches limits the development of a unifying classification framework. For example, a classifier performing well in one application may not necessarily provide good results in another. However, at least for similar problems stemming from different applications fields comprehensive studies should be carried out to compare different existing classification approaches. In such a way, some benchmark performances can be established against which future contributions can be compared.

An expansion of TF methods in different applications are expected to dominate the future contributions. Let's refer back to Table VI. The classical methods based on signal decomposition approaches and Cohen's class are widely used in different application fields. However, it is interesting to note the rare applications of rotated and signal dependent TFRs. It does not necessarily mean that such representations do not provide valid results. For example, it is expected to see increased application of rotated TFRs in speech and music processing, biomedical signal processing and mechanical vibrations analysis. Some problems stemming from such applications actually require the employment of such advanced TF transforms. Similar situations can be seen for signal dependent representations.

VI. Conclusions

The TFA provides a powerful framework for the extraction and classification of nonstationary phenomena in signals as shown in this paper. This paper summarized research results using energy concentration as a feature in the TF domain in a period from early 1990s until now.

Choice of feature extractors in the TF domain, and the feature classifier is highly application-dependent. There is no single TFR that can be claimed to be "the optimal" for all applications. It can be concluded that:

- The signal decomposition based TFRs are implemented in applications when it is not desirable to deal with the cross terms imposed by TFRs that are based on Cohen's idea. The STFT and the wavelet analysis, even though widely applied, do have limitations. Some newer techniques such as the S-transform, the MFT, the STHT or the STHRT provide a framework which enables an improved concentration of the signals in comparison to standard techniques.
- The feature extractors based on Cohen's idea are more suitable when high resolution representation of the feature is required. However, the implementation has to be carefully considered. The kernel function should be optimized for the given application in order to diminish the effects of cross terms. This kernel optimization process can represent an additional computational burden, which is an addition to that of signal decomposition techniques.
- The rotation of the TF plane is used to ensure that the principal axis of the analysis is aligned with the principal axis of the signal components. Several approaches have been introduced to implement such rotation: fractional Fourier transform, linear polynomial Fourier transform and Radon-Wigner distribution. It has been shown that the Radon-Wigner distribution corresponds to the magnitude square of the FRFT of the signal, while the LPFT is a broad generalization of the FRFT.
- The signal dependent TFRs overcome potential shortcomings of fixed mapping representations, which can yield optimized representations only for limited classes of signals. These signal dependent representations can yield higher energy concentration for wider variety of signals. Furthermore, these representations have higher computation cost associated with them. The signal dependent representations can be realized in several ways.

The signal classification using the energy concentration in the TF domain as features is a well researched area, and based on the work of this paper, the following can be concluded:

- The TF based classifiers are more accurate than time- or frequency-domain based classifiers.
- The TF based classification can be performed either by the visual inspection of energy concentration patterns, or by automated processes relying on the measures of distances between the signals and the corresponding template.

As an application example, the framework for the IF estimation error analysis based on the maximum energy concentration is examined as well. Such a framework is important for applications using the IF in the classification of different phenomena.

This paper provides a concise summary of the work in this field in recent years. The results indicate that the TF domain signal processing using energy concentration as a feature is a very powerful tool and has been applied to many fields of applications. It is expected that further research and applications of existing schemes will flourish in the near future.

Acknowledgements

Ervin Sejdić and Jin Jiang would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for financially supporting this work.

References

[1] R. E. Challis and R. I. Kitney, "Biomedical signal processing (in four parts) - part 1: Time-domain methods," *Medical and Biological Engineering and Computing*, vol. 28, no. 6, pp. 509–524, Nov. 1990.

[2] ——, "Biomedical signal processing (in four parts) - part 2: The frequency transforms and their inter-relationships," *Medical and Biological Engineering and Computing*, vol. 29, no. 1, pp. 1–17, Jan. 1991.

[3] ——, "Biomedical signal processing (in four parts) - part 3: The power spectrum and coher-

ence function," *Medical and Biological Engineering and Computing*, vol. 29, no. 3, pp. 225–241, Mar. 1991.

[4] G. Stephanopoulos and C. Han, "Intelligent systems in process engineering: a review," *Computers and Chemical Engineering*, vol. 20, no. 6, pp. 743–791, Jun. 1996.

[5] M. Mahmoud, J. Jiang, and Y. Zhang, *Active Fault Tolerant Control Systems: Stochastic Analysis and Synthesis*. Berlin: Springer, 2003.

[6] A. R. Webb, *Statistical Pattern Recognition*, 2nd ed. West Sussex, England: Wiley, 2002.

[7] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering, and Array Processing*. Boston: McGraw-Hill, 2000.

[8] S. Mukhopadhyay and P. Sircar, "Parametric modelling of non-stationary signals: a unified approach," *Signal Processing*, vol. 60, no. 2, pp. 135–152, Jul. 1997.

[9] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, N.J: Prentice Hall PTR, 1995.

[10] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Boston: Birkhäuser, 2001.

[11] LJ. Stanković, "An analysis of some time-frequency and time-scale distributions," *Annales des Telecommunications*, vol. 49, no. 9/10, pp. 505–517, Sep./Oct. 1994.

[12] S. G. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. San Diego: Academic Press, 1999.

[13] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics, 1992.

[14] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[15] L. Cohen, "Generalized phase-space distribution functions," *Journal of Mathematical Physics*, vol. 7, no. 5, pp. 781–786, May 1966.

[16] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner distribution - a tool for time frequency signal analysis: Part I Continuous time signals," *Philips Journal of Research*, vol. 35, no. 3, pp. 217–250, 1980.

[17] ——, "The Wigner distribution - a tool for time frequency signal analysis: Part II Discrete time signals," *Philips Journal of Research*, vol. 35, no. 4/5, pp. 276–300, 1980.

[18] ——, "The Wigner distribution - a tool for time frequency signal analysis: Part III Relations with other time-frequency signal transformations," *Philips Journal of Research*, vol. 35, no. 6, pp. 372–389, 1980.

[19] L. Cohen, "Time-frequency distribution - a review," *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941–981, Jul. 1989.

[20] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Magazine*, vol. 8, no. 4, pp. 14–38, Oct. 1991.

[21] M. Farge, "Wavelet transforms and their applications to turbulence," *Annual Review of Fluid Mechanics*, vol. 24, pp. 395–457, Jan. 1992.

[22] F. Hlawatsch and G. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, no. 2, pp. 21–67, Apr. 1992.

[23] P. M. Bentley and J. T. E. McDonnell, "Wavelet transforms: an introduction," *Electronics and Communication Engineering Journal*, vol. 6, no. 4, pp. 175–186, Aug. 1994.

[24] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[25] A. Cohen and J. Kovačević, "Wavelets: The mathematical background," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 514–522, Apr. 1996.

[26] N. Hess-Nielsen and M. V. Wickerhauser, "Wavelets and time-frequency analysis," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 523–540, Apr. 1996.

[27] P. Guillemain and R. Kronland-Martinet, "Characterization of acoustic signals through continuous linear time-frequency representations," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 561–585, Apr. 1996.

[28] M. Akay, Ed., *Time-Frequency and Wavelets in Biomedical Signal Processing*. Piscataway, NJ: IEEE Press, 1998.

[29] S. Qian and D. Chen, "Joint time-frequency analysis," *IEEE Signal Processing Magazine*, vol. 16, no. 2, pp. 52–67, Mar. 1999.

[30] P. Flandrin, *Time-frequency/time-scale analysis*. San Diego: Academic Press, 1999.

[31] L. Debnath, Ed., *Wavelet Transform and Time-Frequency Signal Analysis*. Boston: Birkhäuser, 2001.

[32] ——, *Wavelets and Signal Processing*. Boston: Birkhäuser, 2003.

[33] A. Papandreou-Suppappola, Ed., *Applications in Time-Frequency Signal Processing*. Boca Raton: CRC Press, 2003.

[34] P. Bello, "Time-frequency duality," *IEEE Transactions on Information Theory*, vol. 10, no. 1, pp. 18–33, Jan. 1964.

[35] F. Hlawatsch, "Duality and classification of bilinear time-frequency signal representations," *IEEE Transactions on Signal Processing*, vol. 39, no. 7, pp. 1564–1574, Jul. 1991.

[36] A. Papandreou-Suppappola, F. Hlawatsch, and G. F. Boudreaux-Bartels, "Quadratic time-frequency representations with scale covariance and generalized time-shift covariance: A unified framework for the affine, hyperbolic, and power classes," *Digital Signal Processing*, vol. 8, no. 1, pp. 3–48, Jan. 1998.

[37] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*. Amsterdam: Elsevier, 2003.

[38] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: The S-transform," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 998–1001, Apr. 1996.

[39] LJ. Stanković, "A method for time-frequency analysis," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 225–229, Jan. 1994.

[40] D. L. Jones and T. Parks, "A high resolution data-adaptive time-frequency representation," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Process-*

ing (ICASSP 1987), vol. 12, Dallas, Texas, USA, Apr. 6–9, 1987, pp. 681–684.

[41] L. E. Atlas, J. Droppo, and J. McLaughlin, "Optimizing time-frequency distributions for automatic classification," in *Proc. of SPIE Conference on dvanced Signal Processing: Algorithms, Architectures, and Implementations VII*, vol. 3162, San Diego, CA, USA, Jul. 28, 1997, pp. 161–171.

[42] R. Wilson, A. D. Calway, and E. R. S. Pearson, "A generalized wavelet transform for Fourier analysis: The multiresolution Fourier transform and its application to image and audio signal analysis," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 674–690, Mar. 1992.

[43] F. Zhang, G. Bi, Y. Q. Chen, and Y. Zeng, "Harmonic transform," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, vol. 6, Salt Lake City, UT, USA, May 7–11, 2001, pp. 3537–3540.

[44] F. Zhang, G. Bi, and Y. Chen, "Harmonic transform," *IEE Proceedings - Vision, Image, and Signal Processing*, vol. 151, no. 4, pp. 257–263, Aug. 2004.

[45] J.-C. Liu and T. Lin, "Short-time Hartley transform," *IEE Proceedings - Part F: Radar and Signal Processing*, vol. 140, no. 3, pp. 171–174, Jun. 1993.

[46] R. G. Stockwell, "S-transform analysis of gravity wave activity from a small scale network of airglow imagers," Ph.D. dissertation, The University of Western Ontario, London, Ontario, Canada, Sep. 1999.

[47] P. D. McFadden, J. G. Cook, and L. M. Forster, "Decomposition of gear vibration signals by the generalized S-transform," *Mechanical Systems and Signal Processing*, vol. 13, no. 5, pp. 691–707, Sep. 1999.

[48] C. R. Pinnegar and L. Mansinha, "The S-transform with windows of arbitrary and varying shape," *Geophysics*, vol. 68, no. 1, pp. 381–385, Jan./Feb. 2003.

[49] ——, "The bi-Gaussian S-transform," *SIAM Journal on Scientific Computing*, vol. 24, no. 5, pp. 1678–1692, 2003.

[50] ——, "Time-local Fourier analysis with a scalable, phase-modulated analyzing function: the S-transform with a complex window," *Signal Processing*, vol. 84, no. 7, pp. 1167–1176, Jul. 2004.

[51] C. R. Pinnegar, "A new subclass of complex-valued S-transform windows," *Signal Processing*, vol. 86, no. 8, pp. 2051–2055, Aug. 2006.

[52] C. R. Pinnegar and L. Mansinha, "Time-frequency localization with the Hartley S-transform," *Signal Processing*, vol. 84, no. 12, pp. 2437–2442, Dec. 2004.

[53] W. J. Williams and J. Jeong, "New time-frequency distributions: theory and applications," in *Proc. of 1989 IEEE International Symposium on Circuits and Systems*, vol. 2, Portland, OR, USA, May 8-11, 1989, pp. 1243–1247.

[54] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Transactions on Signal Processing*, vol. 40, no. 2, pp. 402–412, Feb. 1992.

[55] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 6, pp. 862–871, Jun. 1989.

[56] B. Zhang and S. Sato, "A time-frequency distribution of Cohen's class with a compound kernel and its application to speech signal processing," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 54–64, Jan. 1994.

[57] D. Wu and J. M. Morris, "Time-frequency representations using a radial Butterworth kernel," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, PA, USA, Oct. 25–28, 1994, pp. 60–63.

[58] Z. Guo, L.-G. Durand, and H. C. Lee, "The time-frequency distributions of nonstationary signals based on a Bessel kernel," *IEEE Transactions on Signal Processing*, vol. 42, no. 7, pp. 1700–1707, Jul. 1994.

[59] A. Papandreou and G. F. Boudreaux-Bartels, "Generalization of the Choi-Williams distribution and the Butterworth distribution for time-frequency analysis," *IEEE Transactions on Signal Processing*, vol. 41, no. 1, pp. 463–472, Jan. 1993.

[60] E. J. Diethorn, "The generalized exponential time-frequency distribution," *IEEE Transactions on Signal Processing*, vol. 42, no. 5, pp. 1028–1037, May 1994.

[61] A. H. Costa and G. F. Boudreaux-Bartels, "Design of time-frequency representations using a multiform, tiltable exponential kernel," *IEEE Transactions on Signal Processing*, vol. 43, no. 10, pp. 2283–2301, Oct. 1995.

[62] N. Ma, D. Vray, P. Delachartre, and G. Gimenez, "Time-frequency representation of multicomponent chirp signals," *Signal Processing*, vol. 56, no. 2, pp. 149–155, Jan. 1997.

[63] B. Barkat and B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2232–2239, Oct. 2001.

[64] K. N. Le, K. P. Dabke, and G. K. Egan, "Hyperbolic kernel for time-frequency power spectrum," *Optical Engineering*, vol. 42, no. 8, pp. 2400–2415, Aug. 2003.

[65] M. G. Amin, "Recursive kernels for time-frequency signal representations," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 16–18, Jan. 1996.

[66] M. G. Amin and W. J. Williams, "High spectral resolution time-frequency distribution kernels," *IEEE Transactions on Signal Processing*, vol. 46, no. 10, pp. 2796–2804, Oct. 1998.

[67] L. Cohen, "Distributions concentrated along the instantaneous frequency," in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations*, vol. 1348, San Diego, CA, USA, Jul. 10, 1990, pp. 149–157.

[68] LJ. Stanković, "A multitime definition of the Wigner higher order distribution: L-Wigner

distribution," *IEEE Signal Processing Letters*, vol. 1, no. 7, pp. 106–109, Jul. 1994.

[69] ——, "A method for improved distribution concentration in the time-frequency analysis of multicomponent signals using the L-Wigner distribution," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1262–1268, May 1995.

[70] LJ. Stanković and S. Stanković, "An analysis of instantaneous frequency representation using time-frequency distributions - generalized Wigner distribution," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 549–552, Feb. 1995.

[71] LJ. Stanković, "A time-frequency distribution concentrated along the instantaneous frequency," *IEEE Signal Processing Letters*, vol. 3, no. 3, pp. 89–91, Mar. 1996.

[72] ——, "L-class of time-frequency distributions," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 22–25, Jan. 1996.

[73] ——, "S-class of time-frequency distributions," *IEE Proceedings - Vision, Image, and Signal Processing*, vol. 144, no. 2, pp. 57–64, Apr. 1997.

[74] ——, "Highly concentrated time-frequency distributions: Pseudo quantum signal representation," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 543–551, Mar. 1997.

[75] S. Stanković and LJ. Stanković, "Introducing time-frequency distribution with a "complex-time" argument," *Electronics Letters*, vol. 32, no. 14, pp. 1265–1267, Jul. 1996.

[76] LJ. Stanković, "Time-frequency distributions with complex argument," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 475–486, Mar. 2002.

[77] S. Oh, R. J. Marks, and L. E. Atlas, "Kernel synthesis for generalized time-frequency distributions using the method of alternating projections onto convex sets," *IEEE Transactions on Signal Processing*, vol. 42, no. 7, pp. 1653–1661, Jul. 1994.

[78] E. J. Zalubas and M. G. Amin, "Time-frequency kernel design by the two-dimensional frequency transformation method," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2198–2202, Sep. 1995.

[79] M. Amin and J. Carroll, "Time-frequency kernel design via point and derivative constraints," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1994)*, vol. 4, Adelaide, SA, Australia, Apr. 19–22, 1994, pp. 309–312.

[80] LJ. Stanković, "Auto-terms representation by the reduced interference distributions: A procedure for kernel design," *IEEE Transactions on Signal Processing*, vol. 44, no. 6, pp. 1557–1563, Jun. 1996.

[81] S. B. Hearon and M. G. Amin, "Minimum-variance time-frequency distribution kernels," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1258–1262, May 1995.

[82] M. G. Amin, "Minimum variance time-frequency distribution kernels for signals in additive noise," *IEEE Transactions on Signal Processing*, vol. 44, no. 9, pp. 2352–2356, Sep. 1996.

[83] M. J. Bastiaans, T. Alieva, and LJ. Stanković, "On rotated time-frequency kernels," *IEEE Signal Processing Letters*, vol. 9, no. 11, pp. 378–381, Nov. 2002.

[84] D. Mendlovic and H. M. Ozaktas, "Fractional Fourier transforms and their optical implementation: part I," *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, vol. 10, no. 9, pp. 1875–1881, Sep. 1993.

[85] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3084–3091, Nov. 1994.

[86] C. Capus and K. Brown, "Fractional Fourier transform of the Gaussian and fractional domain signal support," *IEE Proceedings - Vision, Image, and Signal Processing*, vol. 150, no. 2, pp. 99–106, Apr. 2003.

[87] V. Katkovnik, "A new form of the Fourier transform for time-varying frequency estimation," in *Proc. of URSI International Symposium on Signals, Systems, and Electronics (ISSSE '95)*, San Francisco, USA, Oct. 25–27, 1995, pp. 179–182.

[88] ——, "A new form of the Fourier transform for time-varying frequency estimation," *Signal Processing*, vol. 47, no. 2, pp. 187–200, Nov. 1995.

[89] ——, "Local polynomial periodograms for signals with the time-varying frequency and amplitude," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1996)*, vol. 3, Atlanta, GA, USA, May 7–10, 1996, pp. 1399–1402.

[90] ——, "Nonparametric estimation of instantaneous frequency," *IEEE Transactions on Information Theory*, vol. 43, no. 1, pp. 183–189, Jan. 1997.

[91] ——, "Discrete-time local polynomial approximation of the instantaneous frequency," *IEEE Transactions on Signal Processing*, vol. 46, no. 10, pp. 2626–2637, Oct. 1998.

[92] J. Wood and D. T. Barry, "Radon transformation of time-frequency distributions for analysis of multicomponent signals," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1992)*, vol. 4, San Francisco, CA, USA, Mar. 23–26, 1992, pp. 257–260.

[93] ——, "Radon transformation of the Wigner spectrum," in *SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations III*, vol. 1770, San Diego, CA, USA, Jul. 20, 1992, pp. 358–375.

[94] J. Wood and D. Barry, "Tomographic time-frequency analysis and its application toward time-varying filtering and adaptive kernel design for multicomponent linear-fm signals," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2094–2104, Aug. 1994.

[95] ——, "Radon transformation of time-frequency distributions for analysis of multicomponent signals," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3166–3177, Nov. 1994.

[96] H. M. Ozaktas, N. Erkaya, and M. A. Kutay, "Effecct of fractional Fourier transformation on time-frequency distributions belonging to the Cohen class," *IEEE Signal Processing Letters*, vol. 3, no. 2, pp. 40–41, Feb. 1996.

[97] X.-G. Xia, Y. Owechko, B. H. Soffer, and R. M.

Matic, "On generalized-marginal time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2882–2886, Nov. 1996.

[98] A. W. Lohmann and B. H. Soffer, "Relationships between the Radon-Wigner and fractional Fourier transforms," *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, vol. 11, no. 6, pp. 1798–1801, Jun. 1994.

[99] LJ. Stanković, T. Alieva, and M. J. Bastiaans, "Fractional-Fourier-domain weighted Wigner distribution," in *Proc. of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, Singapore, Aug. 6–8, 2001, pp. 321–324.

[100] A. Akan and L. Chaparro, "Discrete rotational Gabor transform," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Paris, France, Jun. 18–21, 1996, pp. 169–172.

[101] A. Akan, V. Shakhmurov, and Y. Çekiç, "A fractional Gabor transform," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1994)*, vol. 6, Adelaide, SA, Australia, Apr. 19–22, 1994.

[102] A. Akan and Y. Çekiç, "A fractional Gabor expansion," *Journal of the Franklin Institute*, vol. 340, no. 5, pp. 391–397, Nov. 2003.

[103] Y. Zhang, B.-Y. Gu, B.-Z. Dong, G.-Z. Yang, H. Ren, X. Zhang, and S. Liu, "Fractional Gabor transform," *Optics Letters*, vol. 22, no. 21, pp. 1583–1585, Nov. 1997.

[104] LJ. Stanković, "Local polynomial Wigner distribution," *Signal Processing*, vol. 59, no. 1, pp. 123–128, May 1997.

[105] C. Hory, C. Mellet, J.-C. Valiere, and C. Depollier, "Local polynomial time-frequency transform formulation of the pseudo L-Wigner distribution," *Signal Processing*, vol. 81, no. 1, pp. 233–237, Jan. 2001.

[106] R. G. Baraniuk and D. L. Jones, "Unitary equivalence: A new twist on signal processing," *IEEE Transactions on Signal Processing*, vol. 43, no. 10, pp. 2269–2282, Oct. 1995.

[107] A. Bultan, "A four-parameter atomic decomposition of chirplets," *IEEE Transactions on Signal Processing*, vol. 47, no. 3, pp. 731–745, Mar. 1999.

[108] O. Akay and G. Boudreaux-Bartels, "Joint fractional representations," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Pittsburgh, PA, USA, Oct. 6–9, 1998, pp. 417–420.

[109] O. Akay and G. F. Boudreaux-Bartels, "Joint fractional signal representations," *Journal of the Franklin Institute*, vol. 337, no. 4, pp. 365–378, Jul. 2000.

[110] S.-C. Pei and J.-J. Ding, "Relations between fractional operations and time-frequency distributions, and their applications," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1638–1655, Aug. 2001.

[111] F. Zhang, G. Bi, and Y. Q. Chen, "Tomography time-frequency transform," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1289–1297, Jun. 2002.

[112] A. K. Özdemir and O. Arakan, "A high resolution time-frequency representation with significantly reduced cross-terms," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2000)*, vol. 2, Istanbul, Turkey, Jun. 5–9, 2000, pp. 693–696.

[113] A. K. Özdemir and O. Arikan, "Fast computation of the ambiguity function and the Wigner distribution on arbitrary line segments," *IEEE Transactions on Signal Processing*, vol. 49, no. 2, pp. 381–393, Feb. 2001.

[114] D. L. Jones and T. W. Parks, "A high resolution data-adaptive time-frequency representation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 12, pp. 2127–2135, Dec. 1990.

[115] D. L. Jones and R. Baraniuk, "A simple scheme for adapting time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3530–3535, Dec. 1994.

[116] LJ. Stanković, "A measure of some time-frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621–631, Mar. 2001.

[117] W. J. Williams, M. L. Brown, and A. O. Hero, "Uncertainty, information and time-frequency distributions," in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations II*, vol. 1566, San Diego, CA, USA, Jul. 24, 1991, pp. 144–156.

[118] P. Flandrin, R. G. Baraniuk, and O. Michel, "Time-frequency complexity and information," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1994)*, vol. 3, Adelaide, SA, Australia, Apr. 19–22, 1994, pp. 329–332.

[119] O. Michel, R. G. Baraniuk, and P. Flandrin, "Time-frequency based distance and divergence measures," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, PA, USA, Oct. 25–28, 1994, pp. 64–67.

[120] T. H. Sang and W. J. Williams, "Rényi information and signal dependent optimal kernel design," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1995)*, vol. 2, Detroit, MI, USA, May 8–12, 1995, pp. 997–1000.

[121] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1391–1409, May 2001.

[122] S. Aviyente and W. J. Williams, "Minimum entropy time-frequency distributions," *IEEE Signal Processing Letters*, vol. 12, no. 1, pp. 37–40, Jan. 2005.

[123] V. Sucic and B. Boashash, "Parameter selection for optimising time-frequency distributions and measurements of time-frequency characteristics of non-stationary signals," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, vol. 6, Salt Lake City, UT, USA, May 7–11, 2001, pp. 3557–3560.

[124] B. Boashash and V. Sucic, "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distribu-

tions," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1253–1263, May 2003.

[125] K. Kodera and R. G. C. De-Villedary, "Analysis of time-varying signals with small BT values," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-26, no. 11, pp. 64–76, Feb. 1978.

[126] F. Auger and P. Flandrin, "Generalization of the reassignment method to all bilinear time-frequency and time-scale representations," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1994)*, vol. 4, Adelaide, SA, Australia, Apr. 19–22, 1994, pp. 317–320.

[127] ——, "The why and how of time-frequency reassignment," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, PA, USA, Oct. 25–28, 1994, pp. 197–200.

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

[129] R. Baraniuk and D. Jones, "A radially-Gaussian, signal-dependent time-frequency representation," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1991)*, vol. 5, Toronto, Ontario, Canada, Apr. 14–17, 1991, pp. 3181–3184.

[130] R. G. Baraniuk and D. L. Jones, "Signal-dependent time-frequency analysis using a radially Gaussian kernel," *Signal Processing*, vol. 32, no. 3, pp. 263–284, Jun. 1993.

[131] ——, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 41, no. 4, pp. 1589–1602, Apr. 1993.

[132] ——, "A signal-dependent time-frequency representation: Fast algorithm for optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 134–146, Jan. 1994.

[133] M. J. Coates and W. J. Fitzgerald, "Regionally optimised time-frequency distributions using finite mixture models," *Signal Processing*, vol. 77, no. 3, pp. 247–260, Sep. 1999.

[134] D. L. Jones and R. Baraniuk, "An adaptive optimal-kernel time-frequency representation," *IEEE Transactions on Signal Processing*, vol. 43, no. 10, pp. 2361–2371, Oct. 1995.

[135] R. N. Czerwinski and D. L. Jones, "Adaptive cone-kernel time-frequency analysis," *IEEE Transactions on Signal Processing*, vol. 43, no. 7, pp. 1715–1719, Jul. 1995.

[136] J. McLaughlin, J. Droppo, and L. Atlas, "Class-dependent, discrete time-frequency distributions via operator theory," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1997)*, vol. 3, Munich, Germany, Apr. 21–24, 1997, pp. 2045–2048.

[137] J. Droppo and L. Atlas, "Application of classifier-optimal time-frequency distributions to speech analysis," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Pittsburgh, PA, USA, Oct. 6–9, 1998, pp. 585–588.

[138] B. Gillespie and L. Atlas, "Optimization of time and frequency resolution for radar transmitter identification," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1999)*, vol. 3, Phoenix, AZ, USA, Mar. 15–19, 1999, pp. 1341–1344.

[139] ——, "Data-driven time-frequency classification techniques applied to tool-wear monitoring," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2000)*, vol. 2, Istanbul, Turkey, Jun. 6–9, 2000, pp. 649–652.

[140] B. W. Gillespie and L. E. Atlas, "Optimizing time-frequency kernels for classification," *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 485–496, Mar. 2001.

[141] M. Wang and A. V. Mamishev, "Classification of power quality events using optimal time-frequency representations - part 1: Theory," *IEEE Transactions on Power Delivery*, vol. 19, no. 3, pp. 1488–1495, Jul. 2004.

[142] V. Katkovnik, "Adaptive local polynomial periodogram for time-varying frequency estimation," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Paris, France, Jun. 18–21, 1996, pp. 329–332.

[143] E. Chassande-Mottin, F. Auger, and P. Flandrin, "Supervised time-frequency reassignment," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Paris, France, Jun. 18–21, 1996, pp. 517–520.

[144] A. Loza, N. Cunugurujuh, and D. Bull, "A simple scheme for enhanced reassignment of the smoothed pseudo Wigner-Ville representation of noisy signals," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2003)*, vol. 6, Apr. 6–10, 2003, pp. 457–460.

[145] E. Chassande-Mottin, I. Daubechies, F. Auger, and P. Flandrin, "Differential reassignment," *IEEE Signal Processing Letters*, vol. 4, no. 10, pp. 293–294, Oct. 1997.

[146] I. Daubechies and F. Planchon, "Adaptive Gabor transforms," *Applied and Computational Harmonic Analysis*, vol. 13, no. 1, pp. 1–21, Jul. 2002.

[147] C. Richard and R. Lengellé, "Joint recursive implementation of time-frequency representations and their modified version by the reassignment method," *Signal Processing*, vol. 60, no. 2, pp. 163–179, Jul. 1997.

[148] M. Davy and C. Doncarli, "Optimal kernels of time-frequency representations for signal classification," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Pittsburgh, PA, USA, Oct. 6–9, 1998, pp. 581–584.

[149] M. Davy, C. Doncarli, and G. F. Boudreaux-Bartels, "Improved optimization of time-frequency-based signal classifiers," *IEEE Signal Processing Letters*, vol. 8, no. 2, pp. 52–57, Feb. 2001.

[150] P. J. Loughlin, J. W. Pitton, and L. E. Atlas, "Construction of positive time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2697–2705, Oct. 1994.

[151] P. Loughlin, J. Pitton, and B. Hannaford, "Approximating time-frequency density functions

via optimal combinations of spectrograms," *IEEE Signal Processing Letters*, vol. 1, no. 12, pp. 199–202, Dec. 1994.

[152] S. I. Shah, P. J. Loughlin, L. F. Chaparro, and A. El-Jaroudi, "Informative priors for minimum cross-entropy positive time-frequency distributions," *IEEE Signal Processing Letters*, vol. 4, no. 6, pp. 176–177, Jun. 1997.

[153] P. Argoul and T. P. Le, "Instantaneous indicators of structural behaviour based on the continuous Cauchy wavelet analysis," *Mechanical Systems and Signal Processing*, vol. 17, no. 1, pp. 243–250, Jan. 2003.

[154] W. J. Williams and T. Sang, "Adaptive RID kernels which minimize time-frequency uncertainty," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, PA, USA, Oct. 25–28, 1994, pp. 96–99.

[155] S. Aviyente and W. J. Williams, "Minimum entropy approach to denoising time-frequency distributions," in *Proc. of SPIE - Advanced Signal Processing Algorithms, Architectures, and Implementations XI*, vol. 4474, San Diego, CA, USA, Aug. 1, 2001, pp. 57–67.

[156] F. Zhang, Y. Q. Chen, and G. Bi, "Adaptive harmonic fractional Fourier transform," *IEEE Signal Processing Letters*, vol. 6, no. 11, pp. 281–283, Nov. 1999.

[157] C. Capus and K. Brown, "Short-time fractional Fourier methods for the time-frequency representation of chirp signals," *The Journal of the Acoustical Society of America*, vol. 113, no. 6, pp. 3253–3263, Jun. 2003.

[158] LJ. Stanković, T. Alieva, and M. J. Bastiaans, "Time-frequency signal analysis based on the windowed fractional Fourier transform," *Signal Processing*, vol. 83, no. 11, pp. 2459–2468, Nov. 2003.

[159] M. Daković, I. Djurović, and LJ. Stanković, "Adaptive local polynomial Fourier transform," in *Proc. of 11th European Signal Processing Conference (EUSIPCO 2002)*, vol. 2, Toulouse, France, Sep. 3–6, 2002, pp. 603–606.

[160] LJ. Stanković and S. Djukanović, "Order adaptive local polynomial FT based interference rejection in spread spectrum communication systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 6, pp. 2156–2162, Dec. 2005.

[161] I. Djurović, "Robust adaptive local polynomial Fourier transform," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 201–204, Feb. 2004.

[162] Y. Wei and G. Bi, "Efficient analysis of time-varying multicomponent signals with modified LPTFT," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 8, pp. 1261–1268, 2005.

[163] S. Krishnamachari and W. J. Williams, "Adaptive kernel design in the generalized marginals domain for time-frequency analysis," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1995)*, vol. 3, Detroit, MI, USA, Apr. 19–24, 1994, pp. 341–344.

[164] I. Djurović and LJ. Stanković, "Time-frequency representation based on the reassigned S-method," *Signal Processing*, vol. 77, no. 1, pp. 115–120, Aug. 1999.

[165] B. Ristić and B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform," *IEEE Transactions on Signal Processing*, vol. 41, no. 5, pp. 1996–2008, May 1993.

[166] N. S. Rao and P. S. Moharir, "A signal-dependent evolution kernel for Cohen class time-frequency distributions," *Digital Signal Processing*, vol. 8, no. 3, pp. 158–165, Jul. 1998.

[167] Y. Zhang, M. Amin, and G. Frazer, "High-resolution time-frequency distributions for manoeuvring target detection in over-the-horizon radars," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 299–304, Aug. 2003.

[168] Q. Jiang, S. S. Goh, and Z. Lin, "Local discriminant time-frequency atoms for signal classification," *Signal Processing*, vol. 72, no. 1, pp. 47–52, Jan. 1999.

[169] A. Papandreou-Suppappola and S. B. Suppappola, "Analysis and classification of time-varying signals with multiple time-frequency structures," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 92–95, Mar. 2002.

[170] L.-K. Shark and C. Yu, "Design of matched wavelets based on generalized Mexican-hat function," *Mechanical Systems and Signal Processing*, vol. 86, no. 7, pp. 1451–1469, Jul. 2006.

[171] V. Katkovnik and LJ. Stanković, "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2315–2326, Sep. 1998.

[172] LJ. Stanković and V. Katkovnik, "Algorithm for the instantaneous frequency estimation using time-frequency distributions with adaptive window width," *IEEE Signal Processing Letters*, vol. 5, no. 9, pp. 224–227, Sep. 1998.

[173] ——, "Wigner distribution of noisy signals with adaptive time-frequency varying window," *IEEE Transactions on Signal Processing*, vol. 47, no. 4, pp. 1099–1108, Apr. 1999.

[174] LJ. Stanković, I. Djurović, and R.-M. Laković, "Instantaneous frequency estimation by using the Wigner distribution and linear interpolation," *Signal Processing*, vol. 83, no. 3, pp. 483–491, Mar. 2003.

[175] LJ. Stanković and V. Katkovnik, "Instantaneous frequency estimation using higher order L-Wigner distribution with data-driven order and window length," *IEEE Transactions on Information Theory*, vol. 46, no. 1, pp. 302–311, Jan. 2000.

[176] I. Djurović, V. Katkovnik, and LJ. Stanković, "Instantaneous frequency estimation based on the robust spectrogram," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, vol. 6, Salt Lake City, UT, USA, May 7–11, 2001, pp. 3517–3520.

[177] S. Chandra-Sekhar and T. V. Sreenivas, "Adaptive spectrogram vs. adaptive pseudo-Wigner-Ville distribution for instantaneous frequency estimation," *Signal Processing*, vol. 83, no. 7, pp. 1529–1543, Jul. 2003.

[178] D. H. Kil and F. B. Shin, *Pattern Recognition and Prediction with Applications to Signal Characterization*. Woodbury, NY: AIP Press, 1996.

[179] H. P. Zaveri, W. J. Williams, L. D. Iasemidis, and J. C. Sackellares, "Time-frequency representation of electrocorticograms in temporal lobe epilepsy," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 5, pp. 502–509, May 1992.

[180] F. Peyrin, B. Karoubi, D. Morlet, F. Dupont, P. Rubel, P. Desseigne, and P. Touboul, "Application of the Wigner distribution to the detection of late potentials in ECG," in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations III*, vol. 1770, San Diego, CA, USA, Jul. 20, 1992, pp. 418–428.

[181] Z.-Y. Lin and J. D. Z. Chen, "Time-frequency representation of the electrogastrogram - application of the exponential distribution," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 3, pp. 267–275, Mar. 1994.

[182] O. Meste, H. Rix, P. Caminal, and N. V. Thakor, "Ventricular late potentials characterization in time-frequency domain by means of a wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 7, pp. 625–634, Jul. 1994.

[183] T. Mzaik and J. M. Jagadeesh, "Signal component separation using the wavelet transform," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, PA, USA, Oct. 25–28, 1994, pp. 560–563.

[184] P. Bentley and J. McDonnell, "Analysis of heart sounds using the wavelet transform," in *Proc. of IEEE 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Engineering Advances: New Opportunities for Biomedical Engineers)*, vol. 2, Baltimore, MD, USA, Nov. 3–6, 1994, pp. 1304–1305.

[185] G.-C. Jang, C.-K. Cheng, J.-S. Lai, and T.-S. Kuo, "Using time-frequency analysis technique in the classification of surface EMG signals," in *Proc. of IEEE 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Engineering Advances: New Opportunities for Biomedical Engineers)*, vol. 2, Baltimore, MD, USA, Nov. 3–6, 1994, pp. 1242–1243.

[186] J. Wood and D. Barry, "Time-frequency analysis of the first heart sound," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 2, pp. 144–151, Mar./Apr. 1995.

[187] S. Abrahamson, B. Brusmark, H. C. Strifors, and G. C. Gaunaurd, "Aspect dependence of time-frequency signatures of a complex target extracted by impulse radar," in *Record of the IEEE 1995 International Radar Conference*, Alexandria, VA, USA, May 8–11, 1995, pp. 444–449.

[188] W. J. Wang and P. D. McFadden, "Application of orthogonal wavelets to early gear damage detection," *Mechanical Systems and Signal Processing*, vol. 9, no. 5, pp. 497–507, Sep. 1995.

[189] R. Burnett, J. F. Watson, and S. Elder, "The application of modern signal processing techniques for use in rotor fault detection and location within three-phase induction motors," *Signal Processing*, vol. 49, no. 1, pp. 57–70, Feb. 1996.

[190] J. Wood and D. Barry, "Time-frequency analysis of skeletal muscle and cardiac vibrations," *Proceedings of the IEEE*, vol. 84, no. 9, pp. 1281–1294, Sep. 1996.

[191] P. R. White, W. B. Collis, and A. P. Salmon, "Time-frequency analysis of heart murmurs in children," in *Proc. of IEE Colloquium on Time-Frequency Analysis of Biomedical Signals*, London, UK, Jan. 27, 1997, pp. 3/1–3/4.

[192] M. Varanini, G. De Paolis, M. Emdin, A. Macerata, S. Pola, M. Cipriani, and C. Marchesi, "Spectral analysis of cardiovascular time series by the S-transform," in *Computers in Cardiology 1997*, Lund, Sweden, Sep. 7–10, 1997, pp. 383–386.

[193] A. Haghighi-Mood and J. N. Torry, "Time-frequency analysis of systolic murmurs," in *Proc. of Computers in Cardiology 1997*, Lund, Sweden, Sep. 7–10, 1997, pp. 113–116.

[194] H. Oehlmann, D. Brie, M. Tomczak, and A. Richard, "A method for analysing gearbox faults using time-frequency representations," *Mechanical Systems and Signal Processing*, vol. 11, no. 4, pp. 529–545, Nov. 1997.

[195] F. C. Jandre and M. N. Souza, "Wavelet analysis of phonocardiograms: differences between normal and abnormal heart sounds," in *Proc. of IEEE 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 4, Chicago, IL, USA, Oct./Nov. 30–2, 1997, pp. 1642–1644.

[196] Z. Wang, Z. He, and J. D. Z. Chen, "Optimized overcomplete signal representation and its applications to time-frequency analysis of electrogastrogram," *Annals of Biomedical Engineering*, vol. 26, no. 5, pp. 859–869, Sep. 1998.

[197] W. Qiao, H. H. Sun, W. Y. Chey, and K. Y. Lee, "Continuous wavelet analysis as an aid in the representation and interpretation of electrogastrographic signals," *Annals of Biomedical Engineering*, vol. 26, no. 6, pp. 1072–1081, Nov. 1998.

[198] D. Boulahbal, M. F. Golnaraghi, and F. Ismail, "Amplitude and phase wavelet maps for the detection of cracks in geared systems," *Mechanical Systems and Signal Processing*, vol. 13, no. 3, pp. 423–436, May 1999.

[199] G. C. Gaunaurd and H. C. Strifors, "Applications of time-frequency signature analysis to target identification," in *Proc. of SPIE Conference on Wavelet Applications VI*, vol. 3723, Orlando, FL, USA, Apr. 6, 1999, pp. 78–90.

[200] G. Olmo, F. Laterza, and L. L. Presti, "Matched wavelet approach in stretching analysis of electrically evoked surface EMG signal," *Signal Processing*, vol. 80, no. 4, pp. 671–684, Apr. 2000.

[201] V. C. Chen and R. D. Lipps, "Time frequency signatures of micro-Doppler phenomenon for feature extraction," in *Proc. of SPIE Conference on Wavelet Applications VII*, vol. 4056, Orlando, FL, USA, Apr. 26, 2000, pp. 220–226.

[202] B. Gramatikov, J. Brinker, S. Yi-chun, and N. V.

Thakor, “Wavelet analysis and time-frequency distributions of the body surface ECG before and after angioplasty,” *Computer Methods and Programs in Biomedicine*, vol. 62, no. 2, pp. 87–98, Apr. 2000.

[203] G. Livanos, N. Ranganathan, and J. Jiang, “Heart sound analysis using the S-transform,” in *Computers in Cardiology 2000*, Cambridge, MA, Sep. 24–27, 2000, pp. 587–590.

[204] S.-H. Yoon, B. Kim, and Y.-S. Kim, “Helicopter classification using time-frequency analysis,” *Electronics Letters*, vol. 36, no. 22, pp. 1871–1872, Oct. 2000.

[205] N. Baydar and A. Ball, “Detection of gear deterioration under varying load conditions by using the instantaneous power spectrum,” *Mechanical Systems and Signal Processing*, vol. 14, no. 6, pp. 907–921, Nov. 2000.

[206] B. Boashash, M. Mesbah, and P. Colditz, “Newborn EEG seizure pattern characterisation using time-frequency analysis,” in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, vol. 2, Salt Lake City, UT, USA, May 7–11, 2001, pp. 1041–1044.

[207] W. J. Wang, “Wavelets for detecting mechanical faults with high sensitivity,” *Mechanical Systems and Signal Processing*, vol. 15, no. 4, pp. 685–696, Jul. 2001.

[208] S. Yoo, J. R. Boston, T. E. Rudy, C. M. Greco, and J. K. Leader, “Time-frequency analysis of temporomandibular joint (TMJ) clicking sounds using radially Gaussian kernels,” *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 8, pp. 936–939, Aug. 2001.

[209] W. Q. Wang, F. Ismail, and M. F. Golnaraghi, “Assessment of gear damage monitoring techniques using vibration measurements,” *Mechanical Systems and Signal Processing*, vol. 15, no. 5, pp. 905–922, Sep. 2001.

[210] Z. Zhang, H. Kawabatab, and Z.-Q. Liu, “Electroencephalogram analysis using fast wavelet transform,” *Computers in Biology and Medicine*, vol. 31, no. 6, pp. 429–440, Nov. 2001.

[211] H. C. Strifors, G. C. Gaunaurd, and A. Sullivan, “Influence of soil properties on time-frequency signatures of conducting and dielectric targets buried underground,” in *Proc. of SPIE Conference on Automatic Target Recognition XII*, vol. 4726, Orlando, FL, USA, Apr. 2, 2002, pp. 15–25.

[212] Z. Peng, F. Ghu, and Y. He, “Vibration signal analysis and feature extraction based on reassigned wavelet scalogram,” *Journal of Sound and Vibration*, vol. 253, no. 5, pp. 1087–1100, Jun. 2002.

[213] P. K. Dash, B. K. Panigrahi, and G. Panda, “Power quality analysis using S-transform,” *IEEE Transactions on Power Delivery*, vol. 18, no. 2, pp. 406–411, Apr. 2003.

[214] G. Meltzer and Y. Y. Ivanov, “Fault detection in gear drives with non-stationary rotational speed - part II: the time-frequency approach,” *Mechanical Systems and Signal Processing*, vol. 17, no. 2, pp. 273–283, Mar. 2003.

[215] ——, “Fault detection in gear drives with non-stationary rotational speed - part I: the time-frequency approach,” *Mechanical Systems and Signal Processing*, vol. 17, no. 5, pp. 1033–1047, Sep. 2003.

[216] I. Yesilyurt, “Fault detection and location in gears by the smoothed instantaneous power spectrum distribution,” *NDT and E International*, vol. 36, no. 7, pp. 535–542, Oct. 2003.

[217] S. Panagopoulos and J. J. Soraghan, “Small-target detection in sea clutter,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 7, pp. 1355–1361, Jul. 2004.

[218] Y.-J. Shin, E. J. Powers, W. M. Grady, and A. Arapostathis, “Determination of transient disturbance energy flow in electric power systems via cross time-frequency distribution,” in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations XIV*, vol. 5559, Denver, CO, USA, Aug. 4, 2004, pp. 258–265.

[219] M. Bennett, S. McLauglin, T. Anderson, and N. McDicken, “Filtering of chirped ultrasound echo signals with the fractional Fourier transform,” in *Proc. of 2004 IEEE Ultrasonics Symposium*, vol. 3, Montréal, Canada, Aug. 23–27, 2004, pp. 2036–2040.

[220] J. Zou and J. Chen, “A comparative study on time-frequency feature of cracked rotor by Wigner-Ville distribution and wavelet transform,” *Journal of Sound and Vibration*, vol. 276, no. 1-2, pp. 1–11, Sep. 2004.

[221] S. Loutridis and A. Trochidis, “Classification of gear faults using Hoelder exponents,” *Mechanical Systems and Signal Processing*, vol. 18, no. 5, pp. 1009–1030, Sep. 2004.

[222] S. Loutridis, “A local energy density methodology for monitoring the evolution of gear faults,” *NDT and E International*, vol. 37, no. 6, pp. 447–453, Sep. 2004.

[223] H. Yang, J. Mathew, and L. Ma, “Fault diagnosis of rolling element bearings using basis pursuit,” *Mechanical Systems and Signal Processing*, vol. 19, no. 2, pp. 341–356, Mar. 2005.

[224] J. Azaña, “Time-frequency (Wigner) analysis of linear and nonlinear pulse propagation in optical fibers,” *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 10, pp. 1554–1565, 2005.

[225] Y. Goren, L. R. Davrath, I. Pinhas, E. Toledo, and S. Akselrod, “Individual time-dependent spectral boundaries for improved accuracy in time-frequency analysis of heart rate variability,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 1, pp. 35–42, Jan. 2006.

[226] J.-D. Wu and J.-C. Chen, “Continuous wavelet transform technique for fault signal diagnosis of internal combustion engines,” *NDT and E International*, vol. 39, no. 4, pp. 304–311, Jun. 2006.

[227] S. Assous, A. Humeau, M. Tartas, P. Abraham, and J.-P. L'Huillier, “S-transform applied to laser Doppler flowmetry reactive hyperemia signals,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1032–1037, Jun. 2006.

[228] C. Cristalli, N. Paone, and R. Rodriguez, “Mechanical fault detection of electric motors by laser vibrometer and accelerometer measure-

ments," *Mechanical Systems and Signal Processing*, vol. 20, no. 6, pp. 1350–1361, Aug. 2006.

[229] P. Raković, E. Sejdić, LJ. Stanković, and J. Jiang, "Time-frequency signal processing approaches with applications to heart sound analysis," in *Computers in Cardiology 2006*, Valencia, Spain, Sep. 17–20, 2006, pp. 197–200.

[230] C. H. Chen, "Application of wavelet transforms to ultrasonic NDE and remote-sensing signal analysis," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Philadelphia, PA, USA, Oct. 25–28, 1994, pp. 472–475.

[231] S. Blanco, R. Q. Quiroga, O. A. Rosso, and S. Kochen, "Time-frequency analysis of electroencephalogram series," *Physical Review E*, vol. 51, no. 3, pp. 2624–2631, Mar. 1995.

[232] P. Bentley, J. McDonnell, and P. Grant, "Classification of native heart valve sounds using the Choi-Williams time-frequency distribution," in *Proc. of IEEE 17th Annual Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, Montréal, Canada, Sep. 20–23, 1995, pp. 1083–1084.

[233] G. Meltzer and Y. Y. Ivanov, "Identification of blunt-process at wood-milling tools by time-frequency analysis," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Pittsburgh, PA, USA, Oct. 6–9, 1998, pp. 449–452.

[234] E. Grall-Maës and P. Beauseroy, "Features extraction for signal classification based on Wigner-Ville distribution and mutual information criterion," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Pittsburgh, PA, USA, Oct. 6–9, 1998, pp. 589–592.

[235] T. S. Leung, P. R. White, J. Cook, W. B. Collis, E. Brown, and A. P. Salmon, "Analysis of the second heart sound for diagnosis of paediatric heart disease," *IEE Proceedings - Science, Measurement and Technology*, vol. 145, no. 6, pp. 285–290, Nov. 1998.

[236] B. Yazici and G. B. Kliman, "An adaptive statistical time-frequency method for detection of broken bars and bearing faults in motors using stator current," *IEEE Transactions on Industry Applications*, vol. 35, no. 3, pp. 442–452, Mar./Apr. 1999.

[237] C. Baudet, O. Michel, and W. J. Williams, "Detection of coherent vorticity structures using time-scale resolved acoustic spectroscopy," *Physica D: Nonlinear Phenomena*, vol. 128, no. 1, pp. 1–17, Apr. 1999.

[238] P. Chevret, N. Gache, and V. Zimpfer, "Time-frequency filters for target classification," *The Journal of the Acoustical Society of America*, vol. 106, no. 4, pp. 1829–1837, Oct. 1999.

[239] S. Krishnan, R. M. Rangayyan, G. D. Bell, and C. B. Frank, "Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 6, pp. 773–783, Jun. 2000.

[240] O. Poisson, P. Rioual, and M. Meunier, "Detection and measurement of power quality disturbances using wavelet transform," *IEEE Transactions on Power Delivery*, vol. 15, no. 4, pp. 1039–1044, Jul. 2000.

[241] S. Hainsworth, M. Macleod, and P. Wolfe, "Analysis of reassigned spectrograms for musical transcription," in *Proc. of IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, New Platz, NY, USA, Oct. 21–24, 2001, pp. 23–26.

[242] E. Grall-Maës and P. Beauseroy, "Mutual information-based feature extraction on the time-frequency plane," *IEEE Transactions on Signal Processing*, vol. 50, no. 4, pp. 779–790, Apr. 2002.

[243] P. Purkait and S. Chakravorti, "Pattern classification of impulse faults in transformers by wavelet analysis," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 9, no. 4, pp. 555–561, Aug. 2002.

[244] P. K. Dash, B. K. Panigrahi, D. K. Sahoo, and G. Panda, "Power quality disturbance data compression, detection, and classification using integrated spline wavelet and S-transform," *IEEE Transactions on Power Delivery*, vol. 18, no. 2, pp. 595–600, Apr. 2003.

[245] G. Turhan-Sayan, "Natural resonance-based feature extraction with reduced aspect sensitivity for electromagnetic target classification," *Pattern Recognition*, vol. 36, no. 7, pp. 1449–1466, Jul. 2003.

[246] M. Levonen and S. McLaughlin, "Fractional Fourier transform techniques applied to active sonar," in *Proc. of OCEANS 2003*, vol. 4, San Diego, CA, USA, Sep. 22–24, 2003, pp. 1894–1899.

[247] A. Franzen and I. Y. Gu, "Classification of bird species by using key song searching: a comparative study," in *Proc. of IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, Washington, DC, USA, Oct. 5–8, 2003, pp. 880–887.

[248] E. Sejdić and J. Jiang, "Comparative study of three time-frequency representations with applications to a novel correlation method," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, vol. 2, Montréal, Canada, May 17–21, 2004, pp. 633–636.

[249] S. M. Debbal and F. Bereksi-Reguig, "Analysis of the second heart sound using continuous wavelet transform," *Journal of Medical Engineering and Technology*, vol. 28, no. 5, pp. 151–156, Jul./Aug. 2004.

[250] T. Wang, J. Deng, and B. He, "Classification of motor imagery EEG patterns and their topographic representation," in *Proc. of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, San Francisco, USA, Sep. 1–5, 2004, pp. 4359–4362.

[251] W. Wang, J. Pan, and H. Lian, "Decomposition and analysis of the second heart sound based on the matching pursuit method," in *Proc. of the 7th International Conference on Signal Processing (ICSP 2004)*, vol. 3, Beijing, China, Aug./Sep. 31–4, 2004, pp. 2229– 2232.

[252] A. Bernjak, A. Stefanovska, V. Urbančič-Rovan, and K. Ažman-Juvan, "Quantitative assessment

of oscillatory components in blood circulation: classification of the effect of aging, diabetes, and acute myocardial infarction," in *Proc. of SPIE Conference on Advanced Biomedical and Clinical Diagnostic Systems III*, vol. 5692, San Jose, CA, USA, Jan. 23, 2005, pp. 163–173.

[253] H. C. Strifors, T. Andersson, D. Axelsson, and G. C. Gaunaurd, "A method for classifying underground targets and simultaneously estimating their burial conditions," in *Proc. of SPIE Conference on Automatic Target Recognition XV*, vol. 5807, Orlando, FL, USA, Mar. 29, 2005, pp. 112–121.

[254] Y. Amit, A. Koloydenko, and P. Niyogi, "Robust acoustic object detection," *The Journal of the Acoustical Society of America*, vol. 118, no. 4, pp. 2634–2648, Oct. 2005.

[255] A. G. Rehorn, E. Sejdić, and J. Jiang, "Fault diagnosis in machine tools using selective regional correlation," *Mechanical Systems and Signal Processing*, vol. 20, no. 5, pp. 1221–1238, Jul. 2006.

[256] E. Sejdić and J. Jiang, "Selective regional correlation for pattern recognition," *IEEE Transactions on Systems, Man and Cybernetics - Part A*, vol. 37, no. 1, pp. 82–93, Jan. 2007.

[257] S. Aviyente, L. A. W. Brakel, R. K. Kushwaha, M. Snodgrass, H. Shevrin, and W. J. Williams, "Characterization of event related potentials using information theoretic distance measures," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 5, pp. 737–743, May 2004.

[258] C. Martinez, I. Vincent, C. Ioncarli, and P. Guiheunect, "Comparison of classification methods applied to CNAPs," in *Proc. of IEEE 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Engineering Advances: New Opportunities for Biomedical Engineers)*, vol. 2, Baltimore, MD, USA, Nov. 3–6, 1994, pp. 1238–1239.

[259] C. Delfs and F. Jondral, "Classification of piano sounds using time-frequency signal analysis," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1997)*, vol. 3, Munich, Germany, Apr. 21–24, 1997, pp. 2093–2096.

[260] A. Lauberts and T. Andersson, "Classification of buried land mines using combined matched filters on data sequences collected by a hand-held ground-penetrating radar," in *Proc. of SPIE Conference on Subsurface and Surface Sensing Technologies and Applications III*, vol. 4491, San Diego, CA, USA, Jul. 30, 2001, pp. 31–40.

[261] H. Zheng, Z. Li, and X. Chen, "Gear fault diagnosis based on continuous wavelet transform," *Mechanical Systems and Signal Processing*, vol. 16, no. 2-3, pp. 447–457, Mar./May 2002.

[262] Y.-J. Shin and P. Crapse, "Development of transient power quality indices based on time-frequency distribution," in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations XV*, vol. 5910, San Diego, CA, USA, Aug. 2, 2005, p. 59100F.

[263] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal - part 1: Fundamentals," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992.

[264] ——, "Estimating and interpreting the instantaneous frequency of a signal - part 2: Algorithms and applications," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 540–568, Apr. 1992.

[265] G. Jones and B. Boashash, "Generalized instantaneous parameters and window matching in the time-frequency plane," *IEEE Transactions on Signal Processing*, vol. 45, no. 5, pp. 1264–1275, May 1997.

[266] C. Lee and L. Cohen, "Instantaneous mean qualities in time-frequency analysis," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1988)*, vol. 4, New York, NY, USA, Apr. 11–14, 1988, pp. 2188–2191.

[267] B. Boashash, B. Lovell, and P. Kootsookos, "Time-frequency signal analysis and instantaneous frequency estimation: methodology, relationships and implementations," in *Proc. of IEEE International Symposium on Circuits and Systems (ISCAS '89)*, vol. 2, Portland, OR, USA, May 8–11, 1989, pp. 1237–1242.

[268] L. Cohen and C. Lee, "Instantaneous frequency and time-frequency distributions," in *Proc. of IEEE International Symposium on Circuits and Systems (ISCAS '89)*, vol. 2, Portland, OR, USA, May 8–11, 1989, pp. 1231–1234.

[269] P. Rao and F. J. Taylor, "Estimation of instantaneous frequency using the discrete Wigner distribution," *Electronics Letters*, vol. 26, no. 4, pp. 246–248, Feb. 1990.

[270] J. Jeong, G. S. Cunningham, and W. J. Williams, "Instantaneous frequency and kernel requirements for discrete time-frequency distributions," in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations*, vol. 1348, San Diego, CA, USA, Jul. 10, 1990, pp. 170–180.

[271] M. A. Poletti, "Instantaneous frequency and conditional moments in the time-frequency plane," *IEEE Transactions on Signal Processing*, vol. 39, no. 3, pp. 755–756, Mar. 1991.

[272] G. Jones and B. Boashash, "Instantaneous quantities and uncertainty concepts for signal-dependent time-frequency distributions," in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations II*, vol. 1566, San Diego, CA, USA, Jul. 24, 1991, pp. 167–178.

[273] P. J. Kootsookos and B. C. L. B. Boashash, "A unified approach to the STFT, TFDs, and instantaneous frequency," *IEEE Transactions on Signal Processing*, vol. 40, no. 8, pp. 1971–1982, Aug. 1992.

[274] B. C. Lovell, R. C. Williamson, and B. Boashash, "The relationship between instantaneous frequency and time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 41, no. 3, pp. 1458–1461, Mar. 1993.

[275] J. Jeong, G. S. Cunningham, and W. J. Williams, "The discrete-time phase derivative as a definition of discrete instantaneous frequency and its relation to discrete time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 341–344, Jan. 1995.

[276] B. Tacer and P. Loughlin, "Instantaneous frequency and time-frequency distributions," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1995)*, vol. 2, Detroit, MI, USA, May 09–12, 1995, pp. 1013–1016.

[277] B. Ristić and B. Boashash, "Instantaneous frequency estimation of quadratic and cubic FM signals using the cross polynomial Wigner-Ville distribution," *IEEE Transactions on Signal Processing*, vol. 44, no. 6, pp. 1549–1553, Jun. 1996.

[278] V. Valeau, J. C. Valiere, P. Herzog, L. Simon, and C. Depollier, "Instantaneous frequency tracking of a sine wave phase modulation signal," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Paris, France, Jun. 18–21, 1996, pp. 501–504.

[279] M. K. Emresoy and A. El-Jaroudi, "Iterative instantaneous frequency estimation and adaptive matched spectrogram," *Signal Processing*, vol. 64, no. 2, pp. 157–165, Jan. 1998.

[280] F. Çakrak and P. J. Loughlin, "Instantaneous frequency estimation of polynomial phase signals," in *Proc. of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Pittsburgh, PA, USA, Oct. 6–9, 1998, pp. 549–552.

[281] J. Gao, X. Dong, W.-B. Wang, Y. Li, and C. Pan, "Instantaneous parameters extraction via wavelet transform," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 867–870, Mar. 1999.

[282] P. Bonato, Z. Erim, S. H. Roy, and C. J. D. Luca, "Comparison of time-frequency-based techniques for estimating instantaneous frequency parameters of nonstationary processes," in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, vol. 3807, Denver, CO, USA, Jul. 19, 1999, pp. 625–636.

[283] C. Wang and M. G. Amin, "Time-frequency distribution spectral polynomials for instantaneous frequency estimation," *Signal Processing*, vol. 76, no. 2, pp. 211–217, Jul. 1999.

[284] B. Barkat and B. Boashash, "Instantaneous frequency estimation of polynomial FM signals using the peak of the PWVD: Statistical performance in the presence of additive Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2480–2490, Sep. 1999.

[285] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2000)*, vol. 2, Istanbul, Turkey, Jun. 6–9, 2000, pp. 657–660.

[286] B. Barkat and B. Boashash, "IF estimation of linear FM signals corrupted by multiplicative and additive noise: a time-frequency approach," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2000)*, vol. 2, Istanbul, Turkey, Jun. 6–9, 2000, pp. 661–664.

[287] L. Cohen, "Instantaneous frequency and group delayof a filtered signal," *Journal of the Franklin Institute*, vol. 337, no. 4, pp. 329–346, Jul. 2000.

[288] H. K. Kwok and D. L. Jones, "Improved instantaneous frequency estimation using an adaptive short-time fourier transform," *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2964–2972, Oct. 2000.

[289] I. Djurović and LJ. Stanković, "Influence of high noise on the instantaneous frequency estimation using quadratic time-frequency distributions," *IEEE Signal Processing Letters*, vol. 7, no. 11, pp. 317–319, Nov. 2000.

[290] R. G. Baraniuk, M. Coates, and P. Steeghs, "Hybrid linear/quadratic time-frequency attributes," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 760–766, Apr. 2001.

[291] B. Barkat, "Instantaneous frequency estimation of nonlinear frequency-modulated signals in the presence of multiplicative and additive noise," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2214–2222, Oct. 2001.

[292] V. Ivanović, M. Daković, I. Djurović, and LJ. Stanković, "Instantaneous frequency estimation by using time-frequency distributions," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, vol. 6, Salt Lake City, UT, USA, May 7–11, 2001, pp. 3521–3524.

[293] I. Djurović and LJ. Stanković, "Robust Wigner distribution with application to the instantaneous frequency estimation," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 2985–2993, Dec. 2001.

[294] LJ. Stanković, V. N. Ivanović, and M. Daković, "Performance of spectrogram as IF estimator," *Electronics Letters*, vol. 37, no. 12, pp. 797–799, Jun. 2001.

[295] Z. M. Hussain and B. Boashash, "Multicomponent IF estimation: a statistical comparison in the quadratic class of time-frequency distributions," in *Proc. of IEEE International Symposium on Circuits and Systems (ISCAS 2001)*, vol. 2, Sydney, NSW, Australia, May 6–9, 2001, pp. 109–112.

[296] G. Viswanath and T. V. Sreenivas, "IF estimation using higher order TFRs," *Signal Processing*, vol. 82, no. 2, pp. 127–132, Feb. 2002.

[297] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent fm signals using quadratic time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1866–1876, Aug. 2002.

[298] J. D. Harrop, S. N. Taraskin, and S. R. Elliott, "Instantaneous frequency and amplitude identification using wavelets: Application to glass structure," *Physical Review E*, vol. 66, no. 2, pp. 026 703–1–026 703–9, Aug. 2002.

[299] L. Angrisani and M. D'Arco, "A measurement method based on a modified version of the chirplet transform for instantaneous frequency estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 51, no. 4, pp. 704–771, Aug. 2002.

[300] G. Azemi, B. Senadji, and B. Boashash, "Instantaneous frequency estimation of frequency modulated signals in the presence of additive and multiplicative noise: Application to mobile communication systems," in *Proc. of the*

XI European Signal Processing Conference (EUSIPCO'02), vol. 3, Toulouse, France, Sep. 3–6, 2002, pp. 441–444.

[301] M. Daković, V. N. Ivanović, and LJ. Stanković, "On the S-method based instantaneous frequency estimation," in *Proc. of Seventh International Symposium on Signal Processing and Its Applications (ISSPA 2003)*, vol. 1, Paris, France, Jul. 1–4, 2003, pp. 605–608.

[302] V. N. Ivanović, M. Daković, and LJ. Stanković, "Performance of quadratic time-frequency distributions as instantaneous frequency estimators," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 77–89, Jan. 2003.

[303] S. Chandra-Sekhar and T. V. Sreenivas, "Effect of interpolation on PWVD computation and instantaneous frequency estimation," *Signal Processing*, vol. 84, no. 1, pp. 107–116, Jan. 2004.

[304] I. Djurović and LJ. Stanković, "An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment," *Signal Processing*, vol. 84, no. 3, pp. 631–643, Mar. 2004.

[305] V. Valeau, J.-C. Valiére, and C. Mellet, "Instantaneous frequency tracking of a sinusoidally frequency-modulated signal with low modulation index: application to laser measurements in acoustics," *Signal Processing*, vol. 84, no. 7, pp. 1147–1165, Jul. 2004.

[306] L. Angrisani, M. D'Arco, R. S. L. Moriello, and M. Vadursi, "On the use of the warblet transform for instantaneous frequency estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 4, pp. 1374–1380, Aug. 2005.

[307] S. Krishnan, "A new approach for estimation of instantaneous mean frequency of a time-varying signal," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 17, pp. 2848–2855, 2005.