CONSISTENT TIME-FREQUENCY REPRESENTATIONS

D. Scott Roberts and Douglas L. Jones

University of Illinois
Coordinated Science Laboratory
1308 W. Main Street
Urbana, IL 61801

ABSTRACT

This paper develops nonnegative, high-resolution time-frequency representations (TFRs) that correspond with intuitive notions of energy distribution. These so-called *consistent TFRs* require the desired representation to be consistent with a set of spectrogram-based energy measurements. By formulating the desired representation as the solution to a constrained optimization problem, it can be solved using a gradient-projection technique. The consistent TFR demonstrates superior performance compared to existing techniques on a variety of test signals and biological data. The result is a high-resolution, nonnegative, intuitively-satisfying TFR that should prove to be an excellent tool for exploratory data analysis.

1. INTRODUCTION

One of the primary goals of time-frequency analysis is to obtain a joint representation of the signal that shows how energy is distributed in the time-frequency plane. The Wigner-Ville distribution (WVD) is one of the most fundamental time frequency representations.

$$WVD(t, f) = \int s(t + \frac{\tau}{2})s^*(t - \frac{\tau}{2})e^{-j2\pi f\tau}d\tau \qquad (1)$$

The WVD provides an ideal TFR for monocomponent Gaussian signals of the form:

$$s(t) = e^{-\frac{1}{2}(at^2+2bt+c)} \qquad (2)$$

where $a, b \in \mathbb{C}$ with $Re[a] > 0$ and $c \in \mathbb{R}$. The resulting TFR has high resolution, is everywhere nonnegative, and has an energy distribution that corresponds well with intuitive notions for such a signal. Unfortunately, this is the only such signal for which the WVD is everywhere nonnegative [1]. Other monocomponent signals have a WVD that has negative values, and signals with multiple components exhibit cross-terms. Figure 1 shows a WVD of two closely-spaced parallel windowed chirp signals. The representation still has the desired behavior in the autocomponent regions, but now has a large cross-term at the mid-time, mid-frequency points of the autoterms. The cross-term is oscillatory in nature and takes on negative values.

Much of the early work in time-frequency dealt with the suppression of cross-terms. Many approaches result in the smoothing of the WVD to filter out the oscillatory cross-terms. This is seen by looking at Cohen's class for bilinear TFRs as a WVD convolved with a two-dimensional filter.

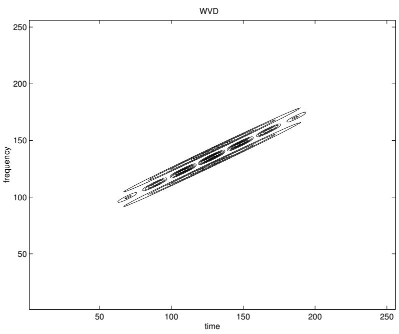


Fig. 1. WVD of windowed chirps

$$C(t, f) = \int\int\int \phi(t - t', f - f')WVD(t', f')dt'df' \qquad (3)$$

where the filter $\phi(\theta, \tau)$ is called the *kernel* of the representation. Different kernels result in different representations. Looking at Cohen's class in the correlative domain is useful for understanding the effect of the filter.

$$C(\tau, \theta) = \Phi(\tau, \theta)AF(\tau, \theta) \qquad (4)$$

where

$$AF(\tau, \theta) = \int s(t + \frac{\tau}{2})s^*(t - \frac{\tau}{2})e^{-j2\pi\theta t}dt \qquad (5)$$

is the ambiguity function (AF).

Ideally, an exploratory TFR would be nonnegative everywhere and retain the high-resolution autoterm behavior of the WVD autoterm regions without exhibiting cross-terms. Essentially, an exploratory TFR should have a *quasi-linear* behavior; a multicomponent signal should have a TFR that is similar to the superposition of the individual TFRs. Figure 2 shows the *consistent TFR*, which is developed in this paper, of the parallel chirp signals. The consistent TFR has excellent resolution, is nonnegative, and has energy placement exactly corresponding to intuitive notions. A variety of approaches to the problem of finding an "ideal" TFR exist, and are mentioned next.

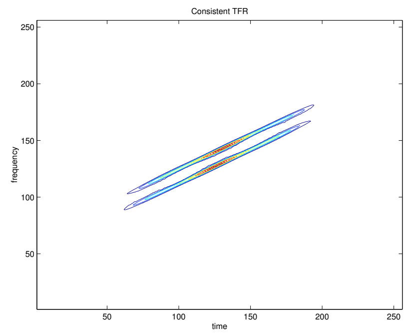


Fig. 2. Consistent TFR of windowed chirps

2. PREVIOUS APPROACHES

There have been many attempts to obtain an "ideal" TFR. While there is certainly some disagreement by what is meant by ideal, most of the approaches strive for nonnegativity, high-resolution of the autoterms, and suppression of cross-terms.

Cross-term energy lies away from the τ, θ axis [2]. A kernel with low-pass properties can be utilized to suppress the cross-terms. If the autoterms and cross-terms are not easily separable, the filter either passes some cross-term energy or filters out some of the autoterm energy, which results in a loss of autoterm resolution. More advanced techniques, such as the adaptive optimal-kernel (AOK) method of Jones and Baraniuk [3], adaptively design the kernel for each signal. This method works well on a wider variety of signals, but it is still a low-pass filtering operation and has limitations.

- **Filtering Approaches.** Filtered versions of Cohen's class were mentioned previously including more advanced methods such as AOK [3].
- **Cohen-Posch positive distributions.** Cohen and Posch [4] provided an early example that demonstrates the existence of a family of nonnegative, marginal-satisfying TFRs.
- **Minimum cross-entropy representations.** Loughlin, Pitton, and Atlas proposed an information-theoretic citerion for choosing a specific TFR from the class of Cohen-Posch positive distributions.
- **Positive RIDs.** Sang, Williams, and O'Neill [5] propose a method to develop positive, marginal-satisfying quadratic distributions with reduced cross-terms.

Many of these approaches are strongly based on satisfying marginals. Marginals are certainly an important mathematical property, but enforcing them along with nonnegativity can obscure the form of the autocomponents. If the time domain signal has very small amplitude at a certain instant, the marginal-satisfying, nonnegative representation must be near zero across all frequencies. This results in autoterms with a segmented structure that counteract the notion of quasi-linearity. Since the goal of quasi-linearity conflicts with that of marginals, the following TFR does not enforce marginal properties.

3. CONSISTENT TFRS

The WVD has high-resolution autoterm behavior that is excellent for an exploratory representation. The negative values of the WVD and excessive cross-terms limit the usefulness of the WVD for real data. The ideal TFR would mimic the autoterm behavior of the WVD and eliminate the cross-terms in some way other than a smoothing operation. The spectrogram is nonnegative and has an energy distribution that is very intuitive. Unfortunately, the spectrogram has poor concentration due to the inherent windowing operation [6]. The spectrogram and WVD are related via Moyal's formula:

$$\int\int C_{x_1}(t,f,\Pi)C^*_{x_2}(t,f,\Pi)dtdf = \left|\int x_1(t)x_2^*(t)dt\right|^2 \quad (6)$$

or expressed as an inner product,

$$< C_{x1}, C_{x_2} >= | < x_1, x_2 > |^2 \quad (7)$$

This relationship holds for all $C(t, f, \Pi)$ that have $|\Pi(\tau, \theta)| = 1$. The WVD is the only member of Cohen's class that satisfies this relationship. If x_1 is chosen as the signal and x_2 as the time-shifted and frequency-shifted window of the spectrogram,

$$< WVD_s, WVD_{w_{t,f}} >= | < s, w_{t,f} > |^2 \quad (8)$$

This relationship states that the inner-product of the WVDs of the signal and the window is equal to the value of the spectrogram at that t, f location. In other words, the energy overlap of the signal's WVD and the window's WVD is consistent with the energy measurement defined by the spectrogram. This is a desirable property, since it implies that intuitive notions of energy distribution are preserved. Unfortunately, the WVD satisfies this relationship for all windows because it is allowed to take on negative values. If the window is chosen so that it does not encompass both autoterms and cross-terms simultaneously in the time-frequency plane, then Moyal's formula will be valid in autoterm regions without utilizing negative values of the WVD. An ideal TFR that behaves like the WVD in autoterm regions can be developed. Additionally, if the TFR is constrained to be nonnegative, cross-term regions will be removed since they require the negative values of the WVD to satisfy Moyal.

The consistent TFR, X, is developed as the solution of a constrained optimization problem that attempts to minimize the deviation from Moyal's formula according to

$$Cost = \sum_k \sum_{t,f} \left|< X, W_{k,t,f} > - | < x, w_{k,t,f} > |^2\right|^2 \quad (9)$$

such that $X \geq 0 \, \forall t, f$. In this expression, X is the desired TFR of the signal, and $W_{k,t,f}$ is the TFR of window k that has been shifted by t, f. In the time domain, $w_{k,t,f}$ is the t, f shifted window k, and x is the signal. The cost function can be minimized more easily be expressing it in the τ, θ domain.

$$Cost = \sum_k \sum_{\tau,\theta} |\mathbb{F}[X]\mathbb{F}[W(-t,-f)] - AF_x(-\theta,\tau)AF_w(\theta,\tau)|^2 \quad (10)$$

$\mathbb{F}$ represents a Fourier transform. W is the WVD of the window. This is plausible since our windows are Gaussian functions and the WVD of a Gaussian exactly corresponds to intuition. After some manipulation, the cost function is

$$Cost = \sum_{\tau,\theta} \left(\sum_k |AF_{W_k}|^2\right) |\mathbb{F}[X] - AF_X(-\theta,\tau)|^2 \quad (11)$$

Now the cost function is in a form where a derivative with respect to the unknown distribution can easily be taken.

$$\frac{\partial Cost}{\partial \mathbb{F}[X]} = \sum_k \left(|AF_{W_k}|^2\right) 2(\mathbb{F}[X] - AF_X(-\theta,\tau)) \qquad (12)$$

The set of nonnegative TFRs is a convex set. It is also relatively straightforward to show that the cost function is a convex function. Weierstrass' theorem ensures the existence of a global minimum for a convex function over a convex set. In addition, the cost function is strictly convex over the set of nonnegative representations, the set of interest, which implies that the global minimum is unique. A method for obtaining the solution is implemented next.

This problem can be solved using a gradient-project algorithm as follows. The gradient-project technique is well suited to this problem, since the nonnegativity constraint can easily be imposed with a simple projection operation.

- Start with some initial representation for X
- Update X using the gradient as in Eq. 12
- Transform back into t, f domain
- Remove any negative values
- Transform back to τ, θ domain and iterate

4. CHOICE OF WINDOWS

The remaining issue to be dealt with is the choice of the window constraints with which to be consistent. In the required simplification of the cost function, the Fourier transform of the window representation is taken to be the AF. This implies that the ideal TFR of the window is the WVD. Since only Gaussian signals have a WVD that is nonnegative and corresponds to intuitive notions of energy distribution, the windows must be Gaussian. Gaussian windows also have a minimum time-bandwidth product, which is a useful attribute since the role of the window is to isolate a region in the time-frequency plane.

Selecting individual windows can be done by examining the cross-term behavior of the spectrogram and WVD. This issue is dealt with in great detail in [7]. The WVD of a multi-component signal will exhibit cross-terms at the mid-time and mid-frequency points of the autocomponent pairs. The amplitudes of the cross-terms can be twice as large as the product of the peak amplitude of each autocomponent. The spectrogram, however, exhibits cross-terms at the intersection of two autocomponents. The amplitude can be as large as twice the product of the two autocomponent amplitudes at the intersection. It is easy to see why cross-terms affect the WVD so severely, because they can have very large amplitudes even for components with minimal overlap. This is not the case with the spectrogram. If constraint windows are chosen such that the resulting spectrogram does not have common cross-term regions with the WVD of the signal, the autoterm behavior of the WVD will be retained. The window does not have to be perfectly matched to the signal; it just needs to not be so poorly matched as to smear the autoterms together. The windows can be chosen by hand with *a priori* signal information. They can also be selected automatically by using the radially-Gaussian kernel (RGK) design procedure that is part of the Jones and Baraniuk AOK [8] procedure as described in [9]. The RGK method works well in practice.

5. RESULTS: "CQ" SPEECH SEGMENT

This signal consists of a male speaker uttering the sound "CQ." Speech is one of the more challengin signals for time-frequency representations because it has time-varying, spectrally diverse components with wide dynamic ranges. Figure 3 shows the results for the consistent TFR and several other commonly used TFRs for comparison. The consistent TFR does an excellent job of showing both wideband and narrow-band information, providing excellent resolution of the formants and good time localization that clearly gives excellent pitch-period information. None of the other representations perform as well on the complex speech signal. More detailed results and other examples can be found in [9].

6. CONCLUSION

Consistent nonnegative TFRs demonstrate very high concentration and resolution of all signal components, superior cross-term suppression, and quasi-linear behavior. The performance as an exploratory visual analysis tool equals or exceeds that of any other TFR for all signals we have examined, including challenging real signals such as speech and multicomponent signals of biological origin.

7. REFERENCES

[1] F. Soto and P. Claverie, "When is the Wigner function of multi-dimensional systems nonnegative?," *Journal of Mathematical Physics*, vol. 24, no. 1, pp. 97–100, January 1983.

[2] P. Flandrin, "Some features of time-frequency representations of multicomponent signals," in *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, 1984, pp. 41B.4.1–41B.4.4.

[3] D. L. Jones and R. G. Baraniuk, "An adaptive optimal-kernel time-frequency representation," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2361–2371, October 1995.

[4] L. Cohen and T. Posch, "Positive time-frequency distribution functions," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-33, no. 1, pp. 31–38, 1985.

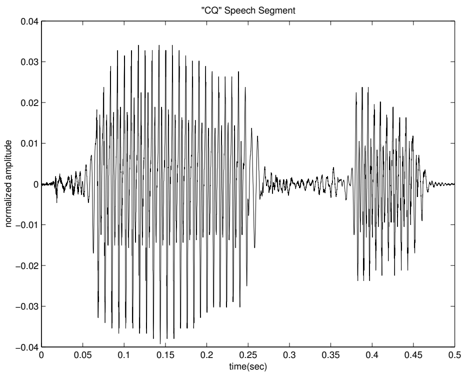
[5] H. Sang, W. J. Williams, and J. C. O'Neill, "An algorithm for positive time-frequency distibutions," in *IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, June 1996, pp. 165–168.

[6] D. L. Jones and T. W. Parks, "A resolution comparison of several time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 40, no. 2, pp. 413–420, February 1992.

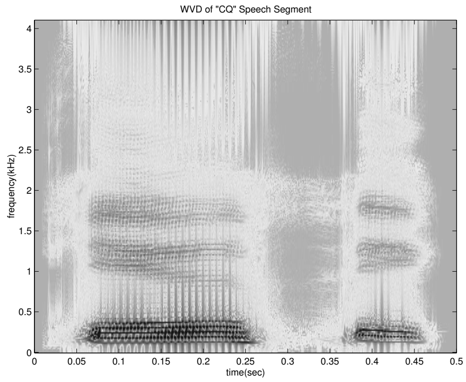
[7] S. Kadambe and G. F. Boudreaux-Bartels, "A comparison of the existence of 'cross-terms' in the Wigner distribution and the squared magnitude of the wavelet transform and the short-time Fourier transform," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2498–2517, October 1992.

[8] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 41, no. 4, pp. 1589–1602, April 1993.

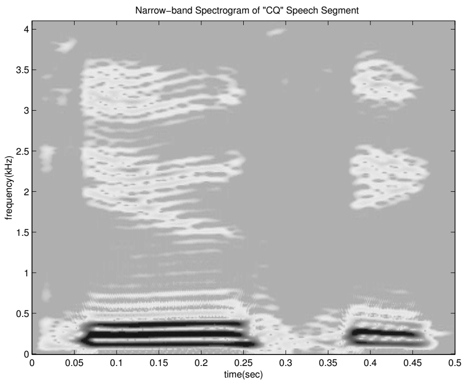
[9] D. S. Roberts, "Applications and extensions of quadratic signal representations," Ph.D. dissertation, University of Illinois at Urbana–Champaign, 2002.



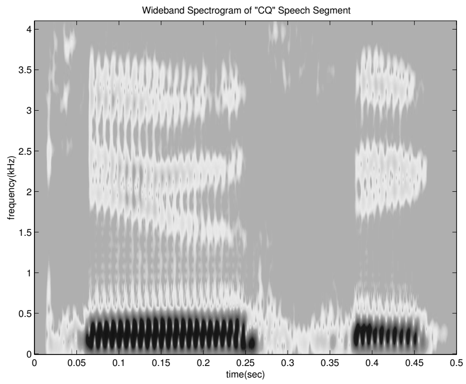
(a) Signal



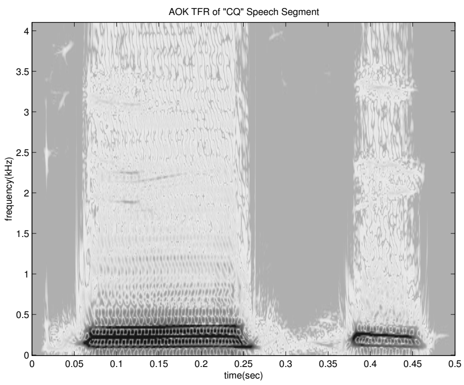
(b) WVD



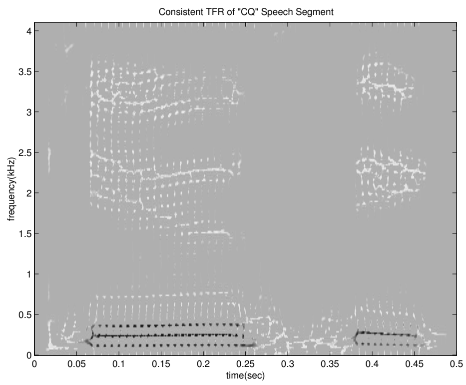
(c) Narrow-band spectrogram using length-255 Hamming window



(d) Wideband spectrogram using length-63 Hamming window



(e) AOK with kernel volume = 3.5



(f) Consistent TFR using a Gaussian window set

Fig. 3. Comparison of techniques for “CQ” speech segment