initial signal depicted in Fig. 1. The errors of the obtained solutions are incorporated into Table I, from which it is seen that the quality of locally regularized solutions compares favorably with that of not only direct solutions but also of the quality of globally regularized solutions.

As an example of image reconstruction, the 8-bit initial image depicted in Fig. 4(a), which had square support region with dimensions 16 $\times$ 16, was reconstructed from its noisy 2-D Fourier transform phase. It turned out that even a very small noise level ($D = 10^{-12}$ rad^2) led to dramatic distortion of the image reconstructed by conventional direct algorithm [Fig. 4(b)] according to (9). The same effect took place for larger noise variances ($D = 10^{-10}$, 10^{-8} rad^2). On the contrary, global RA allows reconstruction of the image that has all the details of the initial one [Fig. 4(c)–(e)] that points to significant smoothing of all undesirable fluctuations in the reconstructed image.

V. Conclusion

In this correspondence, it was pointed out that the problem of the reconstruction of finite discrete signals and images from noisy phase samples of their Fourier transforms is ill-posed with respect to the phase distortion. The regularizing algorithms developed for reconstruction of discrete finite signals and images from their noisy Fourier transform phases have led to significant quality growth of the reconstruction compared with the existing conventional direct approach developed in [6] and [7]. These algorithms significantly (in some cases more than 100 times) improve the accuracy of the reconstruction that allows them to be used as a practical tool for signal reconstruction in conditions of the magnitude uncertainty in the frequency domain.

References

[1] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE*, vol. 69, pp. 529–541, May 1981.
[2] Y. T. Li and A. L. Kurkjian, "Arrival time determination using iterative signal reconstruction from the phase of the cross spectrum," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 502–504, Apr. 1983.
[3] A. Momose, T. Takeda, Y. Itai, and K. Hirano, "Tomographic image reconstruction using x-ray phase information," *Proc. SPIE*, vol. 2708, pp. 674–684, 1996.
[4] A. M. Akhmetshin and I. V. Lyuboshenko, "Introscopy of the layered structures using the phase of reflection coefficient," *Defectoscopia*, no. 12, pp. 64–69, Dec. 1995 (in Russian).
[5] I. V. Lyuboshenko and A. M. Akhmetshin, "Regularizing algorithm for wide-band introscopy using the noisy phase of the reflection coefficient of the probed structure," *Defectoscopia*, no. 1, pp. 71–75, Jan. 1997 (in Russian).
[6] M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Signal reconstruction from phase or magnitude," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 672–680, Dec. 1980.
[7] M. H. Hayes, "The reconstruction of a multi-dimensional sequence from the phase or magnitude of its Fourier Transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 140–154, Apr. 1982.
[8] A. M. Akhmetshin and I. V. Lyuboshenko, "Method for signal reconstruction from noisy phase spectrum," *Radiotech. Electron.*, no. 10, pp. 1229–1234, 1996 (in Russian).
[9] S. Shitz and Y. Y. Zeevi, "On the duality of time and frequency domain signal reconstruction from partial information," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1486–1498, Dec. 1985.
[10] C. Ma, "Novel criteria of uniqueness for signal reconstruction from phase," *IEEE Trans. Signal Processing*, vol. 39, pp. 989–992, Apr. 1991.
[11] A. V. Bacushinsky and A. B. Goncharsky, *Ill-Posed Problems, Numerical Methods with Applications*, Moscow, Russia: MGU, 1989 (in Russian).
[12] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*. New York: Halsted, 1977.
[13] M. T. Silvia and A. B. Weglein, "Method for obtaining a nearfield inverse scattering solution to the acoustic wave equation," *J. Acoust. Soc. Amer.*, vol. 69, pp. 478–482, 1981.
[14] Z. Zhou, R. M. Leahy, and J. Qi, "Approximate maximum likelihood hyperparameter estimation for Gibbs priors," *IEEE Trans. Image Processing*, vol. 6, pp. 844–861, June 1997.
[15] A. Mohammad-Djafari, "On the estimation of hyperparameters in Bayesian approach of solving inverse problems," in *Proc. ICASSP*, 1993, pp. 495–498.
[16] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed. New York: Wiley, vol. 1, 1970.
[17] K. Konstantinides and K. Yao, "Statistical analysis of effective singular values in matrix rank determination," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 757–763, May 1988.
[18] C. Y. Espy and J. S. Lim, "Effects of additive noise on signal reconstruction from Fourier transform phase," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 894–898, Aug. 1983.
[19] J. Park and A. D. Chave, "On the estimation of magnetotelluric response functions using singular value decomposition," *Geophys. J. R. Astr. Soc.*, vol. 77, pp. 683–709, 1984.

Cone–Kernel Representation versus Instantaneous Power Spectrum

G. Viswanath and T. V. Sreenivas

***Abstract*— The Cone–Kernel representation (CKR) and the instantaneous power spectrum (IPS) are two time–frequency representations (TFR's) where the cross-terms are localized in the region of auto-terms. Exploring their relationship for discrete sequences, we show that the IPS kernel is only two extreme terms of the CKR kernel. Further comparing other properties of these two TFR's, viz., invertibility, aliasing, and noise robustness, we show that there is aliasing in some transform domains for signals sampled at Nyquist rate and that the CKR has better noise robustness than the IPS.**

I. Introduction

Time–frequency representations (TFR's) are a more natural and effective method of representing nonstationary signals, such as speech, than quasi-stationary representations. Here, we consider a class of TFR's *whose cross-terms are localized in the region of auto-terms*. The Cone–Kernel representation (CKR), instantaneous power spectrum (IPS), and Rihachzek distribution share this property. This is a unique property among many TFR's because the interference terms generated by the quadratic TFR's for multicomponent signals get localized at the signal terms themselves and do not cause additional ambiguity about the number of signal components in the TFR. Although there are some limitations of these TFR's, such as not being a proper density function, these TF representations could be useful for accurately detecting time-varying signal components; in fact, it

Manuscript received July 25, 1996; revised June 9, 1998. The associate editor coordinating the review of this paper and approving it for publication was Prof. Moeness Amin.

The authors are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India (e-mail: tvsree@ece.iisc.ernet.in).

Publisher Item Identifier S 1053-587X(99)00158-0.

has been shown [1] that the CKR gives an accurate estimate of time discontinuities. In addition, for speech and music signals, frequency plays a more important role than amplitude in their perception. Hence, this class of TFR's hold better promise for speech and music spectral estimation. Interestingly, we find that CKR is related to IPS, which is a TFR that has been ignored for a long time because of the side-lobe leakage problem. The connection shown here would provide a means for trading complexity with performance within this class of TFR's.

The CKR of a signal $x(t)$ is defined as

$$\begin{aligned}\mathrm{CKR}(t,\omega) &= \int_{\tau=-\infty}^{\infty}\int_{s=t-(|\tau|/2)}^{t+(|\tau|/2)} \rho(\tau)x\left(s+\frac{\tau}{2}\right)x^*\left(s-\frac{\tau}{2}\right) \times \exp(-j\omega\tau)\,ds\,d\tau \\ &= \int_{\tau=-\infty}^{\infty}\int_{t}^{t+|\tau|} \rho(\tau)x(s)x^*(s-\tau)\exp(-j\omega\tau)\,ds\,d\tau \end{aligned} \quad (1)$$

where $\rho(t)$ is a finite length window. The IPS is also another TFR whose cross-terms are localized in the region of auto-terms. The IPS has been proposed as a representation of the power spectrum of a signal at each instant. It has also been shown to be a Cohen's class TFR, which is defined as [2]

$$\mathrm{IPS}_x(t,\omega) = \frac{1}{2}\int_{-\infty}^{\infty}[x(t)x^*(t-\tau)+x^*(t)x(t+\tau)] \times \exp(-j\omega\tau)\,d\tau. \quad (2)$$

The IPS has not been very popular because of the side lobe leakage, which is referred to as ringing. A smoothing window $\rho(t)$ is introduced in IPS to reduce ringing [3] (smoothing window introduces a filter with a broader bandwidth), resulting in a modified IPS given by

$$\mathrm{IPS}_x(t,\omega) = \frac{1}{2}\int_{-\infty}^{\infty}[x(t)x^*(t-\tau)+x^*(t)x(t+\tau)]\rho(\tau) \times \exp(-j\omega\tau)\,d\tau. \quad (3)$$

II. Discrete Domain CKR versus IPS

For practical applications, whichever TFR is chosen, the signals are discretized, and a finite window computation is often used. Considering such a situation, for a discrete sequence $x(n)$ and a symmetric window $\rho(n) \neq 0$, $-L \leq n \leq L$, $N = 2L+1$, the discrete CKR and IPS can be expressed as

$$\mathrm{CKR}(n,\omega) = \frac{1}{2L+1}\sum_{k=-L}^{L}\sum_{m=n-(|k|/2)}^{n+(|k|/2)} \rho(k)x\left(m+\frac{k}{2}\right) \times x^*\left(m-\frac{k}{2}\right)e^{-j\omega k} \quad (4)$$

$$= \frac{1}{2L+1}\left[\rho(0)x(n)x^*(n) + 2\,\mathrm{Re}\sum_{k=1}^{L}\sum_{m=n}^{n+k}\rho(k) \times x(m)x^*(m-k)e^{-j\omega k}\right] \quad (5)$$

$$\mathrm{IPS}(n,\omega) = \frac{1}{2}\left[\sum_{k=-L}^{L}(x(n)x^*(n-k) + x^*(n)x(n+k))\rho(k)e^{-j\omega k}\right] \quad (6)$$

$$= \frac{1}{2}\left[\rho(0)x^*(n)x(n) + 2\,\mathrm{Re}\sum_{k=1}^{L}\{x(n)x^*(n-k) + x^*(n)x(n+k)\}\rho(k)e^{-j\omega k}\right]. \quad (7)$$

Consider the CKR over a window of three samples, i.e., a cone length corresponding to $L = 1$ and IPS for a window of $N = 3$

$$\mathrm{CKR}(n,\omega) = \frac{1}{3}[\rho(0)x(n)x^*(n) + 2\rho(1)\,\mathrm{Re}(x(n)x^*(n-1) + x(n)x^*(n+1)e^{-j\omega})]$$

$$\mathrm{IPS}(n,\omega) = \frac{1}{2}[\rho(0)x(n)x^*(n) + 2\rho(1)\,\mathrm{Re}(x^*(n)x(n-1) + x^*(n)x(n+1)e^{-j\omega})].$$

It is clear that both are identical except for a scale factor. For a larger cone length, it can be seen that IPS is only a few terms of CKR. Let us consider the kernels for IPS and CKR as

$$\mathrm{IPS}(n,\omega) = \frac{1}{2}\sum_{k=-L}^{L}R_{\mathrm{IPS}}(n,k)\rho(k)\exp(-j\omega k)$$

where

$$R_{\mathrm{IPS}}(n,k) = x(n)x^*(n-k) + x^*(n)x(n+k) \quad (8)$$

$$\mathrm{CKR}(n,\omega) = \frac{1}{2L+1}\sum_{k=-L}^{L}R_{\mathrm{CKR}}(n,k)\rho(k)\exp(-j\omega k)$$

where

$$R_{\mathrm{CKR}}(n,k) = \sum_{m=n}^{n+k}x(m)x^*(m-k). \quad (9)$$

Thus, it turns out that IPS kernel is obtained by considering only the two terms corresponding to $m = n$ and $m = n+k$ in the expression for the CKR kernel. For CKR, the autocorrelation averaging is up to a lag equal to the cone length, whereas IPS has a fixed averaging using only the two extreme terms of the CKR kernel.

In the case of random signals (for the case in which the time-varying signal may be corrupted with stationary noise), considering the expectation of the TFR, we get

$$E[\mathrm{IPS}(n,\omega)] = \frac{1}{2}\sum_{k}\rho(k)E[x(n)x^*(n-k) + x^*(n)x(n+k)]\exp(-j\omega k) \quad (10)$$

$$= \frac{1}{2}\sum_{k}\rho(k)R_{\mathrm{IPS}}(n,k)\exp(-j\omega k) \quad (11)$$

where $R_{\mathrm{IPS}}(n,k)$ is the nonstationary autocorrelation function. Similarly

$$\begin{aligned}E[\mathrm{CKR}(n,\omega)] &= \frac{1}{2L+1}E\left(\sum_{k=-L}^{L}\rho(k)\sum_{u=n}^{n+k}x(u)x^*(u-k)\exp(-j\omega k)\right) \\ &= \frac{1}{2L+1}\left(\sum_{k=-L}^{L}(|k|+1)\rho(k)R(n,k)\exp(-j\omega k)\right).\end{aligned} \quad (12)$$

For a stationary signal, $R(n,k) = R(k)$ for both IPS and CKR. Such a case would be a windowed periodogram estimate, except that the window function is modified by $(|k|+1)$ for CKR, i.e., a circularly shifted Bartlett window. Assuming $\rho(k)$ to be a rectangular window, at best, the frequency resolution of CKR(n,ω) is half when compared with that of IPS(n,ω).

III. Properties of CKR and IPS

In the light of above similarities, it is interesting to compare the various properties of discrete IPS and discrete CKR. Some of the important properties are discussed below, which help in establishing the tradeoff between CKR and IPS.

A. Inversion

Proposition: For finite-duration signals, if the first nonzero value of the signal is known, then the signal can be recovered from the IPS. When the first nonzero value is not known, then we can recover the signal correct to a scale factor and constant phase shift. (The CKR has been shown [4] to be invertible only if all the signal samples are nonzero.)

Proof: We develop here an inversion algorithm for the discrete IPS that is similar to the inversion algorithm for the short-time Fourier transform magnitude [5]. Consider the IPS for a finite duration signal $x(n) \neq 0, |n| \leq L$, and zero otherwise. From the discrete samples of IPS(n, ω_k), $-L \leq n \leq L$, we can obtain the inverse discrete Fourier transform for each slice of n using a N-point DFT. Thus

$$\begin{aligned} R_{\text{IPS}}(n,k) &= [x(n)x^*(n-k) + x^*(n)x(n+k)]\rho(k) \\ &= \text{IDFT}[\text{IPS}(n,\omega_k)] \end{aligned} \tag{13}$$

where $-L \leq n \leq L$, $-L \leq k \leq L$. Let $x(-L)$ be the first nonzero sample. Considering $R_{\text{IPS}}(-L, 1)$, we can rearrange to get

$$x(-L+1) = \frac{\frac{R_{\text{IPS}}(-L,1)}{\rho(1)} - x^*(-L-1)x(-L)}{x^*(-L)} = \frac{R_{\text{IPS}}(-L,1)}{\rho(1)x^*(-L)} \tag{14}$$

because $x(-L-1) = 0$. Continuing this in a sequential manner

$$x(-L+2) = \frac{\frac{R_{\text{IPS}}(-L+1,1)}{\rho(1)} - x^*(-L)x(-L+1)}{x^*(-L+1)} \tag{15}$$

in which all the previous samples are known. However, the algorithm began only with a nonzero value of $x(-L)$. Thus, for a general complex signal, the reconstruction from IPS is accurate to within a constant factor and a fixed phase delay. The above sequential procedure assumes that $x(n) \neq 0$. However, if for any n $x(n)$ is zero, then $x(n+1)$ can be obtained from $R_{\text{IPS}}(n-1, 2)$ instead of $R_{\text{IPS}}(n, 1)$. This can be shown as

$$x(n+1) = \frac{\frac{R_{\text{IPS}}(n-1,2)}{\rho(2)} - x^*(n-3)x(n-1)}{x^*(n-1)} \tag{16}$$

where all the right-hand side terms are known. Extending this idea, if there are m consecutive zeros in the signal, i.e., $x(n)$ is computed to be zero for $n_1 \leq n < n_1 + m$ and $m < L - 1$; then, $x(n_1 + m)$ can be computed using $R(n_1 - 1, m + 1)$ instead of using previous equation. This is given by

$$x(n_1+m) = \frac{\frac{R_{\text{IPS}}(n_1-1,m+1)}{\rho(m+1)} - x^*(n_1-m-2)x(n_1-1)}{x^*(n_1-1)}. \tag{17}$$

Once a nonzero value is computed, the sequential inversion can continue using (14). Note that the above expression is valid up to $m = L - 2$, i.e., $L - 1$ consecutive zeros in the signal because $R_{\text{IPS}}(n, k)$ is defined only within the range of $|n|, |k| \leq L$.

For very long signals, IPS is computed by consecutive windowing of the signal. In such a case, there should be overlap between windows such that there is at least one nonzero sample in the region of overlap. The above sequential algorithm can be then applied by knowing the amount of overlap between adjacent windows. □

B. Aliasing

Proposition: For real signals sampled at Nyquist rate $\mathbf{f_s}$, cross-terms can cause aliasing in the transform domains of TFR if the bandwidth of the signal is greater than $\frac{\mathbf{f_s}}{\mathbf{4}}$. The aliasing is due to the cross-terms localized at the auto-terms in the TF domain. (It may be noted that there is no aliasing along the frequency axis in the TF domain.)

Proof: Consider a real signal $x(n) = A\cos(\omega_1 n + \phi_1))$ and its CKR. We find an oscillating term along time in the TF domain, whose frequency of oscillation is equal to twice the frequency of the signal, i.e.,

$$R_{\text{CKR}}(n,k) = 2(k+1)\cos(\omega_1 k) + \frac{2\cos(2\omega_1 n)}{\sin(\omega_1)}\sin(\omega_1(k+1)). \tag{18}$$

In the case of IPS, it can also be seen that there is a similar oscillating term whose frequency of oscillation is twice the signal frequency.

$$R_{\text{IPS}}(n,k) = 4\cos(\omega_1 k)(1 + \cos(2\omega_1 n)). \tag{19}$$

It may be noted that the frequency doubling does not cause any aliasing along the frequency dimension of the TF plane. However, there will be aliasing due to this frequency doubling in both the ambiguity domain and spectral domain along doppler and frequency directions, respectively (i.e., if the signal frequency is $\frac{\pi}{4} \leq \omega_1 \leq \pi$). Therefore, an analytic signal or signal sampled at twice the Nyquist rate $(F_s = 4F_{\max})$ is necessary to avoid aliasing in all the transform domains of both CKR and IPS. For signals in noise at low SNR, signal classification based on ambiguity domain is reported to be robust [6]. In such cases, aliasing could introduce errors and affect the performance. □

C. Noise Robustness

Proposition: For a signal in additive white noise, the CKR estimate tends to be unbiased for large cone lengths [4], whereas the IPS is a biased estimator. For sinusoids in circularly complex white Gaussian noise, CKR provides a better SNR than IPS.

Proof: Consider a signal $x(n) = s(n) + \sigma\eta(n)$, where $\eta(t)$ is stationary, zero mean unit variance and circularly complex white noise.[1] In addition, let $s(n) = A\exp(-j\omega_0 n)$; hence

$$E[\text{IPS}_{xx}(n,\omega)] = \text{IPS}_{ss}(n,\omega) + \sigma^2 = A^2 + \sigma^2. \tag{20}$$

Thus, IPS is a biased estimate, whereas CKR is an asymptotically unbiased estimator as shown in [4]. Let us define the SNR as the ratio of sinusoid variance to that of other components. The second moment of IPS is obtained as

$$\begin{aligned} &E[\text{IPS}^2(n,\omega)] \\ &= \sum_{k_1=-L}^{L}\sum_{k_2=-L}^{L}\Bigg[\overbrace{E[x(n)x^*(n-k_1)x^*(n)x(n-k_2)]}^{a} \\ &\quad + \overbrace{E[x^*(n)x(n-k_1)x^*(n)x(n+k_2)]}^{b} \\ &\quad + \overbrace{E[x^*(n)x(n+k_1)x^*(n)x(n-k_2)]}^{c} \\ &\quad + \overbrace{E[x^*(n)x(n+k_1)x^*(n)x(n+k_2)]}^{d}\Bigg]\exp(-j\omega(k_1-k_2)). \end{aligned} \tag{21}$$

Let us use the following new symbols: $s_1 = s(n)$, $s_2 = s(n-k_1)$, $s_3 = s(n)$, $s_4 = s(n-k_2)$ and $n_1 = \eta(n)$, $n_2 = \eta(n-k_1)$, $n_3 = \eta(n)$, $n_4 = \eta(n-k_2)$. Thus

$$\begin{aligned} a = \sum_{k_2=-L}^{L}\sum_{k_1=-L}^{L} &E[(s_1+n_1)(s_2^*+n_2^*)(s_3^*+n_3^*)(s_4+n_4)] \\ &\times\exp(-j\omega(k_1-k_2)) \end{aligned} \tag{22}$$

[1] It may be noted that we take the approach given in [7] in assuming a complex white Gaussian noise to simplify the expressions for the variance of noise terms.

which can be further simplified as

$$\begin{aligned} s_1 s_2^* s_1^* s_4 &= A^4 \exp(j\omega_1(k_1 - k_2)) \\ s_1 s_2^* E[n_1^* n_4] &= A^2\sigma^2 \exp(-j\omega_1 k_1)\delta(k_2) \\ s_1 s_2^* E[n_1^* n_4] &= A^2\sigma^2 \delta(k_1 - k_2) \\ s_2^* s_4 E[n_1 n_1^*] &= A^2\sigma^2 \exp(-j\omega_1(k_1 - k_2)) \\ s_1^* s_4 E[n_1 n_2^*] &= A^2\sigma^2 \exp(-j\omega_1 k_2)\delta(k_1) \\ E[n_1 n_1^*]E[n_2^* n_4] &= \sigma^4 \delta(k_1 - k_2) \\ E[n_1 n_2^*]E[n_1 n_4^*] &= \sigma^4 \delta(k_1)\delta(k_2). \end{aligned} \tag{23}$$

Substituting these into (22) and simplifying, we get

$$E[\text{IPS}^2(n,\omega)] = A^4 + (2L+4)A^2\sigma^2 + \sigma^4(2L+2). \tag{24}$$

Thus,

$$\text{SNR}_{\text{IPS}} = \frac{A^4}{(L+2)A^2\sigma^2 + \sigma^4(L+1)}. \tag{25}$$

Similarly, the CKR variance is given by

$$\begin{aligned} &E[\text{CKR}^2(n,\omega)] \\ &= \sum_{k_1=-L}^{L} \sum_{k_2=-L}^{L} \sum_{m_1=n}^{n+k_1} \sum_{m_2=n}^{n+k_2} (s_1 s_2^* s_3^* s_4 + s_1 s_2^* E[n_3^* n_4] \\ &\quad + s_1 s_3^* E[n_2^* n_4] + s_2^* s_4 E[n_1 n_3^*] + s_3^* s_4 E[n_1 n_2^*] \\ &\quad + E[n_1 n_2^*]E[n_3^* n_4] + E[n_1 n_3^*]E[n_2^* n_4]) \\ &\quad \times \exp(-j\omega_1(k_1 - k_2)). \end{aligned} \tag{26}$$

The arguments in the above summation can be simplified by direct substitution for s and n

$$\begin{aligned} s_1 s_2^* s_3^* s_4 &= A^4 \exp(+j\omega_1(k_1 - k_2)) \\ s_1 s_2^* E[n_3^* n_4] &= A^2\sigma^2 \exp(-j\omega_1 k_1)\delta(k_2) \\ s_1 s_3^* E[n_2^* n_4] &= A^2\sigma^2 \exp(-j\omega_1(m_1 + k_1)) \\ &\quad \times \exp(j\omega_2(m_2 + k_2)) \\ &\quad \times \delta(m_1 - m_2 + k_1 - k_2) \\ s_2^* s_4 E[n_1 n_3^*] &= A^2\sigma^2 \exp(j\omega_1(m_1 - k_1/2)) \\ &\quad \times \exp(-j\omega_1(m_2 - k_2/2)) \\ &\quad \times \delta(m_1 - m_2 + k_1 - k_2) \\ s_3^* s_4 E[n_1 n_2^*] &= A^2\sigma^2 \exp(-j\omega_1 k_2)\delta(k_1) \\ E[n_1 n_2^*]E[n_3 n_4^*] &= \sigma^4 \delta(k_1)\delta(k_2) \\ E[n_1 n_3^*]E[n_2^* n_4] &= \sigma^4 \delta(m_1 - m_2 + k_1 - k_2). \end{aligned} \tag{27}$$

Substituting these terms in (26) and simplifying, we get the variance and SNR as

$$\text{Var}_{\text{CKR}} = \frac{\sigma^4}{4} + A^2\sigma^2, \quad \text{SNR}_{\text{CKR}} = \frac{A^4(L^2 + 3L + 2)}{4A^2\sigma^2 + \sigma^4}. \tag{28}$$

We can see that the CKR has a significant SNR advantage over the IPS even at minimum cone length of $L = 2$.

Simulation: To verify the noise robustness property of CKR and IPS discussed above, a simulation experiment is conducted using a stationary signal in stationary noise situation. The CKR cone width is chosen as 128 and IPS window width also as 128. A sinusoid of frequency $\omega_1 = 900$ Hz, $F_s = 8$ KHz is chosen for simulations. The signal is zero padded, and a DFT of order $N = 8192$ is computed. The maximum value of spectrum at $n = 800$ is obtained for both CKR and IPS. The associated frequency is used as the estimate of instantaneous frequency. This is repeated over ten different noise realizations. The mean and variance of the estimate is shown in Table I. (The mean and variance are computed by finding the peaks at each instant of time.) The spectral discretization leads to an error of ≈1 Hz. We can see that, as expected, CKR has no bias upto 0 dB SNR, whereas IPS has a large bias, which increases with the amount of noise. In addition to the bias, we can also compute the variance of the frequency estimation error (different from the variance of the noise terms derived earlier), as indicative of the effect of noise terms on the signal component. From Table II, it can be seen that CKR provides an order of magnitude smaller variance than IPS, as expected. In addition, the variance diverges at about 10 dB lower for CKR than IPS.

TABLE I
MEAN AND VARIANCE OF IF ESTIMATE OF A SINUSOID AT FREQUENCY 900 Hz BASED ON PEAKS IN CKR AND IPS DOMAINS

	IPS		CKR	
SNR	mean	var	mean	var
-5	1211.9	819.97	903.51	96.89
0	1060.4	605.76	900.51	53.55
5	989.1	464.75	899.43	48.15
10	950.6	353.92	899.67	29
15	925.6	248.61	899.9	24.67
20	923.5	238.23	900.27	21.68
25	922.6	232.48	900.12	19.21
30	923.4	237.59	900.11	17.8

TABLE II
COMPARISON OF THE PROPERTIES OF THE CKR AND THE IPS

Properties	IPS	CKR
1. Invertibility	Invertible even if some signal terms are zero	Only if the signal is non-zero at all instants of time
2. Aliasing in transform domains for signals sampled at Nyquist rate	There is aliasing if signal frequency exceeds $\frac{f_s}{4}$	Same as in the IPS
3. Localization of cross-terms	Localized on auto-terms	Localized on auto-terms
4. Noise Robustness	SNR is significantly less than that of CKR. Biased estimator.	Depends on the cone-length. For large cone-length high SNR. Unbiased estimator.
5. Resolution	Better	Half that of IPS

IV. SUMMARY

We have explored the relation between two unique TFR's (the CKR and the IPS) whose cross-terms are localized in the region of auto-terms. Interestingly, they both have similarities in the structure of the kernel. The CKR kernel has the number of terms depend on the cone-length, whereas the IPS has a fixed number of two terms corresponding to the extreme terms of CKR. An inversion algorithm is developed for the IPS, thus making both the transforms invertible. In the presence of noise, the CKR is an unbiased estimator unlike the IPS and has a large advantage in terms of SNR in the transform domain. However, the frequency resolution of CKR is half that of IPS. For real signals, the oscillating component along the auto-terms of both CKR and IPS has twice the signal frequency, which can cause aliasing in the ambiguity domain.

REFERENCES

[1] L. Atlas, P. J. Loughlin, and J. W. Pitton, "New nonstationary techniques for analysis and display of speech transients," in *Proc. IEEE ICASSP*, 1990, pp. 385–388.

[2] M. J. Levin, "Instantaneous spectra and ambiguity function," *IEEE Trans. Inform. Theory*, vol. IT-10, Jan. 1964.

[3] R. D. Hippensteil and P. M. De Oliveira, "Time varying spectral estimation using the instantaneous power spectrum," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1752–1759, 1990.

[4] S. Oh and R. J. Marks, "Some properties of generalized time-frequency representation with the cone-kernel," *IEEE Trans. Signal Processing*, vol. 40, pp. 1735–1746, July 1992.

[5] S. H. Nawab and T. F. Quatieri, "Signal reconstruction from STFT magnitude," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 986–998, Aug. 1983.

[6] S. S. Abeysekhara and B. Boashash, "Methods of signal classification using the images produced by Wigner–Ville distribution," *Pattern Recogn. Lett.*, vol. 12, pp. 717–729, Nov. 1991.

[7] M. G. Amin, "Minimum variance time-frequency distribution kernels for signals in additive noise," *IEEE Trans. Signal Processing*, vol. 44, pp. 2352–2356, Sept. 1996.

Hybrid Linear/Bilinear Time-Scale Analysis

Martin Pasquier, Paulo Gonçalvès, and Richard Baraniuk

***Abstract*— We introduce a new method for the time-scale analysis of nonstationary signals. Our work leverages the success of the "time–frequency distribution series/cross-term deleted representations" into the time-scale domain to match wideband signals that are better modeled in terms of time shifts and scale changes than in terms of time and frequency shifts. Using a wavelet decomposition and the Bertrand time-scale distribution, we locally balance linearity and bilinearity in order to provide good resolution while suppressing troublesome interference components. The theory of frames provides a unifying perspective for cross-term deleted representations in general.**

I. Introduction

By displaying the time-varying frequency content of a nonstationary signal in terms of time and frequency variables, joint time–frequency and time-scale representations can reveal subtle features that remain hidden from other methods of analysis. Each type of representation matches a different class of signals. Time–frequency representations are covariant to time and frequency shifts and match signals with constant-bandwidth structure, such as narrowband radar signals [1]. Time-scale representations are covariant to time shifts and scale changes and match signals with proportional-bandwidth structure, such as wideband sonar and acoustic signals [2], [3]. Many different representations exist, both linear and nonlinear.

Linear representations, such as the short-time Fourier and Gabor time–frequency representations and the wavelet time-scale representation, offer the benefit of simple interpretation at the expense of poor resolution. Bilinear representations such as the Wigner time–frequency distribution [1] and the Bertrand time-scale distribution [3] were developed as high-resolution alternatives. While the nonlinearity of the bilinear distributions sharpens the representation of local signal structure, it also generates interference between widely separated components that degrade the representation of global structure. Traditionally, nonlinear interference due to these cross-components has been suppressed via smoothing over the time-frequency or time-scale planes [1], [2].

Manuscript received August 6, 1996; revised June 23, 1998. This work was supported by the National Science Foundation under Grant MIP-9457438 and the Office of Naval Research under Grant N00014-95-1-0849. The associate editor coordinating the review of this paper and approving it for publication was Dr. Nugun Erdol.

M. Pasquier is with the Société SEGIME—Groupe ALTRAN, Paris, France.

P. Gonçalvès is with INRIA Rocquencourt—Projet Fractales, Le Chesnay, France (e-mail: paulo.goncalves@inria.fr)

R. Baraniuk is with the Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005 USA (e-mail: richb@rice.edu).

Publisher Item Identifier S 1053-587X(99)00159-2.

In [4] and [5], Qian *et al.* introduced an alternative approach to time–frequency analysis that features an explicit and controllable linear versus bilinear tradeoff. First, the signal is represented in terms of a discrete sum of time–frequency concentrated atoms via a linear Gabor transform. Then, the Wigner distribution, evaluated on this linear signal decomposition rather than on the signal itself, separates into two distinct components: the Wigner auto-components of the atoms (the "quasilinear" part of the representation) and the Wigner cross-components of the atoms (the bilinear part of the representation). By limiting the number of cross-components entering into the sum, such a *cross-term deleted Wigner distribution* can locally control the degree of nonlinearity of the time–frequency representation and, furthermore, tune it for maximum concentration with minimum cross components.

This time–frequency decomposition performs very well, but it is matched only to signals possessing a constant-bandwidth structure. In this correspondence, we extend the concept of hybrid linear/bilinear analysis to the time-scale plane. Our approach is based on the linear wavelet transform and the bilinear Bertrand distribution. In the process of our development, we gain new insights into the procedure of Qian *et al*. In Section II, we briefly review their approach to quasilinearizing the Wigner distribution. In Section III, we transpose the problem of hybrid linear/bilinear analysis to time-scale and propose a frame-based solution. After discussing an implementation of this new method in Section IV, we close with conclusions in Section V.

II. Hybrid Time–Frequency Analysis

A *hybrid linear/bilinear* system for time–frequency analysis consists of three components:

1) a bilinear time-frequency mapping;
2) a discrete linear signal decomposition based on time-frequency concentrated "atoms";
3) a rule for determining which cross-components to include in the overall signal representation.

In [4] and [5], Qian *et al.* utilize the bilinear Wigner distribution, a linear Gabor transform with a Gaussian window, and a Manhattan distance criterion.

A. Wigner Distribution

The Wigner distribution is, in many senses, the central bilinear time–frequency distribution [1]. The cross-Wigner distribution of two signals r and s is defined as[1]

$$W_{r,s}(t,f) = \int r^*\Big(t - \frac{\tau}{2}\Big) s\Big(t + \frac{\tau}{2}\Big) e^{-j2\pi f \tau}\, d\tau. \tag{1}$$

When $s = r$, we have the Wigner distribution $W_s(t,f)$.

The excellent time–frequency localization properties of the *auto-components* of the Wigner distribution result from its bilinear,

[1] Throughout this paper, integration bounds run from $-\infty$ to $+\infty$.