Reducing noise in the time-frequency representation using sparsity promoting kernel design

Branka Jokanović, Moeness G. Amin, and Yimin D. Zhang

Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA

ABSTRACT

Missing samples in the time domain introduce noise-like artifacts in the ambiguity domain due to their de facto zero values assumed by the bilinear transform. These artifacts clutter the dual domain of the time-frequency signal representation and obscures the time-frequency signature of single and multicomponent signals. In order to suppress the artifacts influence, we formulate a problem based on the sparsity aware kernel. The proposed kernel design is more robust to the artifacts caused by the missing samples.

Keywords: Missing samples, sparsity, time-frequency kernel design, time-frequency signal representation

1. INTRODUCTION

Nonstationary signals are frequently encountered in practice, including radar, sonar and biomedical applications [1]- [4]. These signals are traditionally analysed in the time-frequency domain in order to adequately address their time-varying spectral behaviour.

Reduced interference distributions (RIDs) are well-known methods used for the time-frequency representation (TFR) of nonstationary signals [5]- [8]. These distributions can be conveniently defined through the use of the ambiguity domain. In this domain, the signal auto-terms are clustered around the origin, while the undesirable cross-terms are dislocated form the origin. Since the positions of auto-terms and cross-terms are different for each signal, it is not simple to decouple these terms and obtain a proper TFR. Various kernels of low-pass filter characteristic have been designed with the goal of suppressing the cross-terms, while preserving the signal auto-terms [3], [4].

Traditional time-frequency analysis assumes that the analyzed signals are uniformly sampled with the rate equal to or higher than the Nyquist rate. As a result, RIDs do not consider the randomly undersampled or incomplete data. Recently, compressive sensing (CS) is introduced as a framework for the analysis of data that is sampled at a rate significantly lower than Nyquist [9], [10]. Even though CS has been extensively studied in various fields, possible benefits in other areas such as the time-frequency analysis are of growing interest.

Time-frequency signature and ambiguity function of frequency modulated signals are sparse in their respective domains and when viewed through a window. This property has allowed compressive sensing and sparse reconstruction to play an important role in enhancing the data time-frequency distributions (TFDs) [11]- [16]. Joint-variable nonstationary signal representations have greatly benefited from the local sparsity of polynomial phase signals which has put the high resolution TFD in a new context different from the one traditionally considered in this area [17], [18].

This paper introduces a novel kernel design that incorporates the sparsity property into the kernel definition. By imposing the sparsity in the time-frequency domain, the kernel avoids the clutter regions in the ambiguity domain. This in turn mitigates the noise associated with the missing samples in the time-frequency domain and produces a better signal representation than the one obtained with sparsity-unaware kernel.

This paper is organized as follows. Section 2 reviews the RIDs and the role of the traditional kernels. The proposed design is presented in Section 3. Simulation results are provided in Section 4, while the conclusion is given in Section 5.

Further author information: (E-mail: branka.jokanovic@villanova.edu)

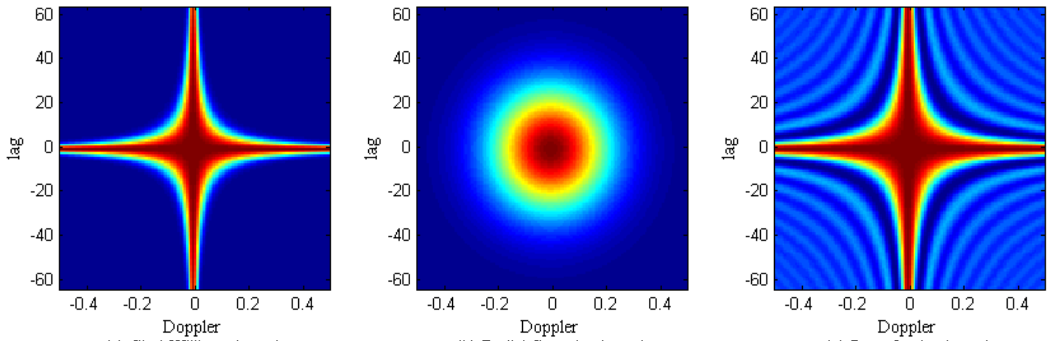

Figure 1. Commonly used kernels in time-frequency analysis.

2. REDUCED INTERFERENCE DISTRIBUTIONS

The Wigner-Ville distribution (WVD) is one of the simplest TFRs. The WVD of $x(n)$ is defined as,

$$WVD(n,k) = \sum_{m=-N/2}^{N/2-1} x(n+m)x^*(n-m)e^{-\jmath 2\pi mk/N}, \tag{1}$$

where m denotes time lag, N is the signal length, while n and k denote discrete time and frequency, respectively. The WVD provides high energy concentration in the time-frequency plane and is an ideal TFR for linearly frequency modulated signals, i.e., chirps. However, for multicomponent signals, the WVD exhibits the presence of cross-terms which can be mistaken as signal components. These cross-terms are a consequence of the bilinear product in (1). In order to suppress the cross-terms and achieve high energy concentration in the time-frequency domain, RIDs are introduced. RID can be defined as,

$$RID(n,k) = \sum_{p=-N/2}^{N/2-1} \sum_{m=-N/2}^{N/2-1} A(p,m)C(p,m)e^{-\jmath 2\pi np/N}e^{-\jmath 2\pi mk/N}, \tag{2}$$

where p denotes Doppler frequency, $C(p,m)$ is the kernel, whereas $A(p,m)$ is the ambiguity function of signal $x(n)$. The ambiguity function can be formulated as,

$$A(p,m) = \sum_{n=-N/2}^{N/2-1} x(n+m)x^*(n-m)e^{-\jmath 2\pi np/N}. \tag{3}$$

Various kernels have been defined in order to address the issue of reducing the effect of cross-terms. Figure 1 illustrates several kernels well-known in the time-frequency analysis. Their formulation is given in Table 1. We can notice that all kernels assume some form of a low-pass filter in the ambiguity domain. Figure 2 shows the WVD and RID based on the Choi-Williams kernel of a signal consisting of a sinusoid and a chirp. We can notice that strong cross-terms, which exist in the WVD, are removed in the RID.

3. SPARSITY PROMOTING KERNEL DESIGN

In this section, we present the kernel design which promotes sparsity in the time-frequency distribution and the sparsity awareness in the kernel design is a novel step which is introduced in an attempt to reduce the artifacts in the TFR caused by the missing samples in the time domain. The kernel design is formulated as an optimization problem.

Table 1. Definition of some commonly used time-frequency kernels.

Time-frequency kernels	$C(p,m)$
Choi-Williams	$e^{\frac{-p^2m^2}{\sigma}}$
Radial Gaussian	$e^{\frac{-(p^2+m^2)}{\sigma}}$
Born-Jordan	$\frac{\sin(pm/2)}{pm/2}$

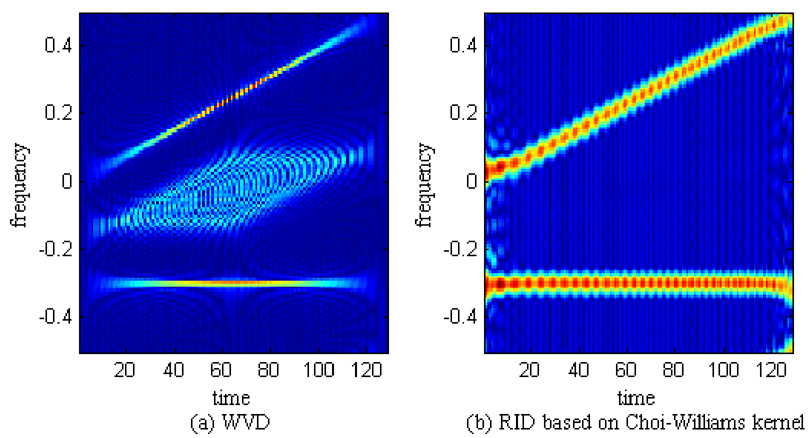

Figure 2. TFRs of a signal consisting of a sinusoid and a chirp.

Let us first briefly examine the effect of missing samples on the ambiguity domain. Signal with missing samples can be represented as the sum of the original signal $s(n)$ and the set of missing samples at random positions n_i,

$$x(n) = s(n) - \sum_{n_i} \delta(n - n_i)s(n). \tag{4}$$

If we denote $\sum_{n_i} \delta(n - n_i)s(n)$ as $v(n)$, we can write the following expression for the ambiguity function of $x(n)$,

$$A_x(p,m) = \sum_{n=-N/2}^{N/2-1} [s(n+m)s^*(n-m) - v(n+m)s^*(n-m) - v^*(n-m)s(n+m) + v(n-m)v^*(n-m)]e^{-\jmath 2\pi np/N}, \tag{5}$$

i.e.,

$$A_x(p,m) = A_s(p,m) + \sum_{n=-N/2}^{N/2-1} [-v(n+m)s^*(n-m) - v^*(n-m)s(n+m) + v(n-m)v^*(n-m)]e^{-\jmath 2\pi np/N}. \tag{6}$$

We can notice that only the first term in (6) corresponds to the desired ambiguity function, while the other terms are artifacts caused by the missing samples. Further details about the noise effect caused by the missing samples can be found in [14]. As can be seen from (6), the noise pattern in the ambiguity domain depends on the values of the missing samples values and their positions. Therefore, some parts of the ambiguity function will be more distorted than others. This motivates the search for points which are reliable for TFR estimation. Model in (4) provides another information which facilitates the search for the reliable points. Namely, missing samples can be seen as impulses in the time domain. It is known that a portion of energy of impulses in the time domain is located along the $\tau = 0$, i.e., the lag axis in the ambiguity domain. Thus, by avoiding the points along the lag axis, a part of noise is removed. This is shown in the Figure 3 where RIDs are based on the Choi-Williams kernel. It should be noted that other signal components, which are not impulses, can also occupy some part of lag axis. However, since they are overlapping with impulses, we can neglect them under the assumption that the power of impulses is higher than that of auto-terms.

In the case of multicomponent signals the influence of cross-terms should be reduced, hence the search for more reliable points is performed within the specified kernel region. That is, we assume that the kernel shape is fixed in advance and that the cross-terms are negligible or eliminated. For a given kernel shape, the search for reliable points is formulated as follows,

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \| \mathbf{a}_C \odot \mathbf{x} - \mathbf{a}_C \|_2 \\ \text{subject to} \quad & \| \mathbf{W}(\mathbf{x} \odot \mathbf{a}_C) \|_1 \leq K \\ & \mathbf{x}_L \leq \mathbf{x} \leq \mathbf{x}_U \end{aligned} \tag{7}$$

where symbol $\circ$ denotes Hadamard product, $\mathbf{a}_C$ is the vectorized version of the ambiguity function $\mathbf{A}_C$ which is obtained after applying a kernel ($\mathbf{A}_C = \mathbf{A} \odot \mathbf{C}$). $\mathbf{W}$ is the discrete 2D Fourier transform matrix, K is a parameter for adjusting the sparsity, whereas $\mathbf{x}_L$ and $\mathbf{x}_U$ denote the lower and upper limits of $\mathbf{x}$.

The task of the optimization problem is to find the mask $\mathbf{x}$, which when applied along with some of the traditional kernels, will minimize the clutter in TFR caused by the missing samples. In order to perform that task, we formulate the objective function and the constraints in the following manner. The objective function is the error between the noisy ambiguity function and the masked ambiguity function. The role of the objective function is to preserve the energy of auto-terms since they contain the desired signal information. On the other hand, these auto-terms are distorted so the sparsity constraint is added in order to reduce the clutter in TFR. Namely, by specifying the sparsity of the TFR, the kernel mask puts the highest weight on the points which alleviate that property, while the influence of other points is suppressed. In this way, the artifacts are reduced since they compromise the sparsity of TFR. Additionally, we include the lower and upper limits for the mask. Also, in the implementation we specify that kernel assume zero value along the lag axis.

4. SIMULATION RESULTS

In this section, we demonstrate the effectiveness of the proposed kernel using multicomponent signal. The results show that artifacts in the TFR are significantly reduced by using the proposed kernel design. We include RID using Choi-Williams kernel as one of the traditional kernels.

We observe a multicomponent signal which consists of two sinusoids. 30% of samples are randomly missing in time. The RID based on the use of Choi-Williams kernel is shown in Figure 4(d). Some of the noise is removed by imposing zeros along the kernel lag axis. TFR obtained after applying the proposed approach on the Choi-Williams kernel is shown in Figure 4(b). We can notice that the kernel design obtained in the optimization process described in Section 3 (Figure 4(a)) provides less cluttered TFR when compared to the RID based on the non-optimized kernel (Figure 4(c)). Even though the kernel does not perfectly recover the auto-terms, noise is significantly removed when compared to the results in Figure 4(d).

5. CONCLUSION

This paper introduces kernel design with the aim of reducing the artifacts in the TFR caused by missing samples or random undersampling. Namely, missing samples in the time domain distort the ambiguity function and the TFR. In order to reduce the artifacts in the time-frequency domain, the proposed kernel searches for the points in the ambiguity domain responsible for a sparse TFR. In so doing, TFR becomes less cluttered when compared to the traditional RIDs.

REFERENCES

[1] P. Flandrin, M. Amin, S. McLaughlin, and B. Torresani, "Time-frequency analysis and applications [from the guest editors]," *IEEE Signal Process. Magazine*, vol. 30, no. 6, pp. 19–150, 2013.

[2] M. I. Skolnik, *Radar Systems.* McGraw Hill, 2001.

[3] B. Boashash, *Time Frequency Analysis.* Gulf Professional Publishing, 2003.

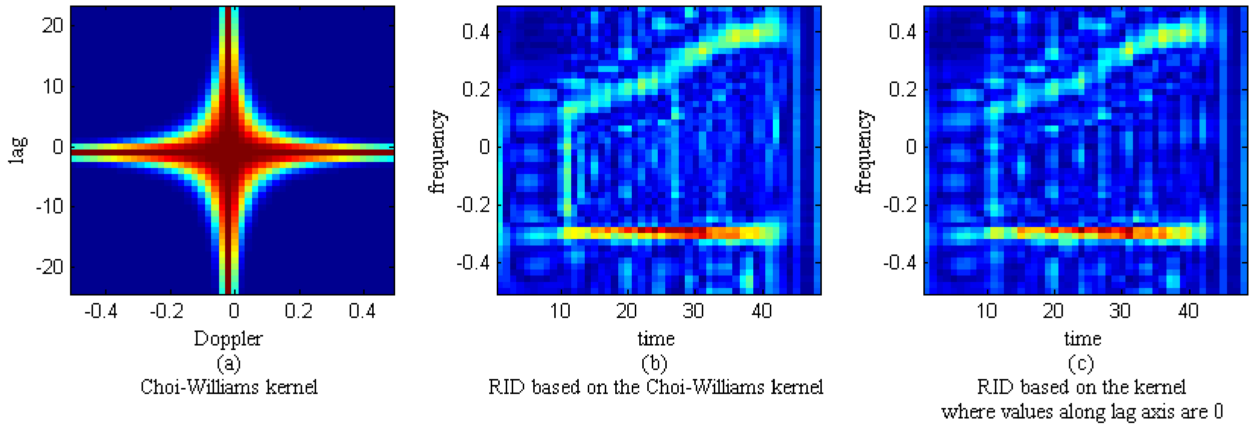

Figure 3. Reducing the effect of impulses by imposing the zero values along the lag axis in the ambiguity domain.

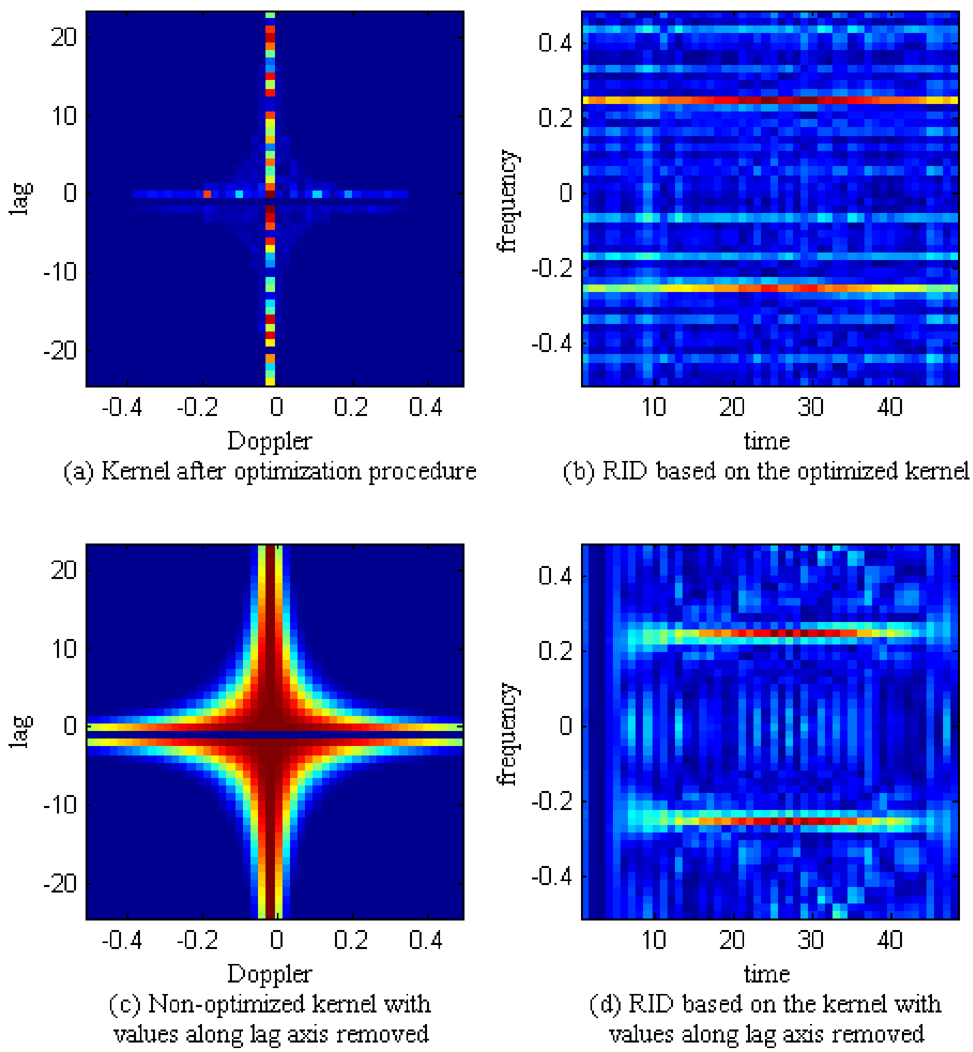

Figure 4. TFR using the proposed approach and RID based on Choi-Williams kernel.

[4] L. Cohen, *Time-Frequency Analysis: Theory and Applications.* Prentice-Hall, Inc., 1995.

[5] H. I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, 1989.

[6] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 402–412, 1992.

[7] M. G. Amin, "Spectral decomposition of time-frequency distribution kernels," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1156–1165, 1994.

[8] J. W. Pitton and L. E. Atlas, "Discrete-time implementation of the cone-kernel time-frequency representation," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1996–1998, 1995.

[9] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory.*, vol. 52, no. 4, pp. 1289–1306, 2006.

[10] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[11] P. Flandrin and P. Borgnat, "Time-frequency energy distributions meet compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 2974–2982, 2010.

[12] B. Jokanović, M. Amin, and S. Stanković, "Instantaneous frequency and time-frequency signature estimation using compressive sensing," in *Proc. SPIE*, vol. 871418, 2013.

[13] L. Stankovic, I. Orovic, S. Stankovic, and M. Amin, "Compressive sensing based separation of nonstationary and stationary signals overlapping in time-frequency," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4562–4572, Sept 2013.

[14] Y. D. Zhang, M. G. Amin, and B. Himed, "Reduced interference time-frequency representations and sparse reconstruction of undersampled data," in *Proc. European Signal Proc. Conf.*, 2013.

[15] M. G. Amin (ed.), *Compressive Sensing for Urban Radar.* CRC Press, 2014.

[16] L. Stankovic, S. Stankovic, and M. Amin, "Missing samples analysis in signals for applications to L-estimation and compressive sensing," *Signal Processing*, vol. 94, pp. 401–408, 2014.

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

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