A RADIALLY-GAUSSIAN, SIGNAL-DEPENDENT TIME-FREQUENCY REPRESENTATION

Richard G. Baraniuk and Douglas L. Jones*

Department of Electrical and Computer Engineering
University of Illinois
Urbana, IL 61801

ABSTRACT

Each bilinear time-frequency distribution corresponds to a kernel function that controls its cross-component suppression properties. Selection of a fixed kernel limits the class of signals for which the distribution performs well; for analysis of a broad class of signals, a signal-dependent kernel is recommended. In this paper, a new procedure that is based on optimization criteria and radially-Gaussian functions is introduced for signal-dependent kernel design. Examples demonstrate excellent performance for a variety of multi-component signals.

1 INTRODUCTION

Time-Frequency Distributions (TFD's) are employed to analyze signals with time-varying frequency content in a wide variety of applications. A number of different representations have been developed. The bilinear distributions of Cohen's class can be interpreted as the two-dimensional Fourier transforms of weighted versions of the ambiguity function (AF) of the signal [1]; that is,

$$P(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\theta, \tau) \Phi(\theta, \tau) e^{-j\theta t - j\tau \omega} d\theta d\tau, \quad (1)$$

where A is the AF of the signal, given by

$$A(\theta, \tau) = \int_{-\infty}^{\infty} s(t + \frac{\tau}{2}) s^*(t - \frac{\tau}{2}) e^{j\theta t} dt. \quad (2)$$

The weighting function, Φ , is called the *kernel*; the product $C = A\Phi$ is known as the characteristic function of the distribution.

Since the AF is a bilinear function of the signal, it exhibits cross-components, which, if allowed to pass into the TFD, can obscure the true signal features, make interpretation of the distribution difficult and reduce auto-component resolution [2]. Therefore, the kernel is often selected to weight the AF such that the auto-components, which are centered at the origin of the (θ, τ) -plane, are passed, while the cross-components, which are located away from the origin, are suppressed. This is the specification of a two-dimensional lowpass filter.

Traditionally, TFD's have used fixed kernels, with fixed passband and stopband regions. However, specification of a fixed kernel limits the class of signals for which the time-frequency representation reflects the true nature of the signal.

That is, given any fixed kernel, it is always possible to find signals with either significant cross-component energy in the passband of the kernel, or significant auto-component energy in the stopband of the kernel.

To illustrate the limitations of fixed-kernel TFD's, consider a signal consisting of the sum of two chirp signals of large effective time envelope. The Wigner distribution [1] has a kernel

$$\Phi_{WD}(\theta, \tau) \equiv 1. \quad (3)$$

The Wigner distribution of the signal is shown in Fig. 1; although the auto-components are highly concentrated, there is a large cross-component. The Choi-Williams distribution [3], which has an exponential kernel

$$\Phi_{CW}(\theta, \tau) = e^{-\frac{\theta^2 \tau^2}{\rho}}, \quad (4)$$

works well for signals whose AF auto-components lie nearly parallel with the θ or τ axis, but poorly for signals with substantial frequency modulation (see Fig. 2), since the kernel severely truncates the auto-components of such signals.

Since the locations of the auto- and cross-components in the (θ, τ) -plane depend on the signal, we expect to obtain good performance for a broad class of signals only by using a *signal-dependent* kernel. Signal-dependent kernels have been considered by several authors; a number of references are given in [4]. In this paper, we introduce a new procedure for selecting a signal-dependent kernel that, given a signal, automatically designs a kernel that is optimal with respect to a set of performance criteria.

2 OPTIMAL KERNEL DESIGN

We formulate the signal-dependent kernel design procedure as an optimization problem. In general, the problem formulation involves the signal, and includes a set of constraints that define a class of kernels from which the optimal kernel is chosen, and a performance index that measures the quality of the time-frequency representation for each kernel. The kernel that maximizes the value of the performance measure is selected as the optimal kernel for the signal. Possible constraints could include, for example, constraints that force the kernel to suppress cross-components, satisfy the marginal distributions [3] or satisfy Moyal's formula [1]. Clearly, the choice of constraints and performance measure is crucial to the success of the method. However, once a satisfactory set of constraints and measure are found, kernel design for a wide range of signals reduces to solving an optimization problem.

A method for optimal, signal-dependent kernel design is discussed in [4]. In that paper, the constraints force the ker-

*This work was supported by the Music Group of the Computer-based Education Research Laboratory and the Joint Services Electronics Program, Grant No. N00014-90-J-1270.

nel to be bounded, radially nonincreasing and of finite volume, that is, to be a two-dimensional lowpass filter, while the performance index measures the energy passed from the AF into the TFD. The optimal kernel yields excellent results for a variety of signals; the optimal-kernel TFD for the chirp signal discussed in the Introduction is given in Fig. 3. A fundamental property of this technique is that the optimal kernel is a 1/0 mask, and, hence, can introduce ringing into the TFD. In applications where ringing is objectionable, suboptimal tapering of the kernel is necessary. In this paper, we formulate a new kernel design procedure using radially-Gaussian functions. This class of kernels is inherently tapered, and may potentially yield more desirable results for many signals.

2.1 A Radially-Gaussian Kernel

The signal-dependent kernel we propose is a two-dimensional, radially-Gaussian function whose spread parameter, σ , varies with radial angle, ψ :

$$\Phi(\theta, \tau) = e^{-\frac{\theta^2 + \tau^2}{2\sigma^2(\psi)}}. \quad (5)$$

The shape of the kernel is completely parameterized by the positive, one-dimensional function σ , so finding the optimal, radially-Gaussian kernel for a signal is equivalent to finding the optimal function, σ , for the signal. Since $\sigma(\psi)$ controls the spread along a radial slice of the kernel at angle ψ , we call it the *spread function*. Clearly, if σ is smooth, then Φ is also smooth. The kernel is easily expressed in polar coordinates by using $r = \sqrt{\theta^2 + \tau^2}$ as the radius variable:

$$\Phi(r, \psi) = e^{-\frac{r^2}{2\sigma^2(\psi)}}. \quad (6)$$

2.2 Continuous Optimization Formulation

Given a signal, the optimal kernel is defined as the radially-Gaussian function whose spread function, σ , solves the following optimization problem:

$$\max_{\Phi} \int_0^{2\pi} \int_0^\infty |A(r, \psi)|^2 \Phi(r, \psi)^2 r dr d\psi \quad (7)$$

subject to

$$\Phi(r, \psi) = e^{-\frac{r^2}{2\sigma^2(\psi)}} \quad (8)$$

$$\frac{1}{2\pi} \int_0^{2\pi} \int_0^\infty |\Phi(r, \psi)|^2 r dr d\psi \leq \alpha, \quad \alpha \geq 0, \quad (9)$$

where $A(r, \psi)$ is the AF of the signal in polar coordinates. The special structure of the kernel permits a simplification of the second constraint to

$$\frac{1}{2\pi} \int_0^{2\pi} \sigma^2(\psi) d\psi \leq \alpha. \quad (10)$$

Also, since the AF is symmetric about the origin,

$$|A(r, \psi)|^2 = |A(r, \psi + \pi)|^2, \quad (11)$$

σ must be determined only on the interval $[0, \pi)$.

The constraints and performance index are motivated by a desire to suppress cross-components, and pass auto-components with as little distortion as possible. The first constraint (8) limits the signal-dependent kernel to the class of radially-Gaussian kernels. Thus, the kernel is constrained to be a lowpass filter. Since the AF auto-components are centered at the origin, this encourages the kernel to preferentially pass auto-components. The second constraint (9), (10) limits the volume of the kernel to α , so that cross-components are suppressed. The performance measure (7) expresses a

desire to minimize auto-component distortion by passing as much auto-component energy as possible into the TFD for a kernel of fixed volume. An advantage of this formulation is that the constraints are insensitive to both the time-scale and orientation of the signal in time-frequency.

Conceptually, the computation of the optimal-kernel TFD involves first finding the AF of the signal. The optimization problem (7)–(10) is then solved for the optimal spread function, σ^* , which determines a radially-Gaussian kernel, Φ . The optimal-kernel TFD is then computed via (1).

2.3 Discrete Optimization Formulation

In practice, TFD's are computed at discrete time and frequency locations, so we reformulate the optimization problem by discretizing (7), (8) and (10). The polar-sampled, radially-Gaussian kernel is given by

$$\Phi_p(p, q) = e^{-\frac{(p\Delta_r)^2}{2\sigma_q^2(q\Delta_\psi)}}, \quad p = 0, \dots, P-1, \quad q = 0, \dots, Q-1, \quad (12)$$

where p and q are the radius and angle indices, and Δ_r and Δ_ψ are the radius and angle step sizes. The discrete kernel is parameterized by a positive spread vector, $\underline{\sigma} \in \mathbb{R}^Q$, whose elements correspond to samples of the spread function

$$\sigma_q = \sigma(q\Delta_\psi). \quad (13)$$

We define the optimal discrete kernel as the radially-Gaussian function whose spread vector solves the following optimization problem:

$$\max_{\Phi_p} \Delta_r^2 \Delta_\psi \sum_{q=0}^{Q-1} \sum_{p=1}^{P-1} p |A_p(p, q)|^2 \Phi_p(p, q)^2 \quad (14)$$

subject to

$$\Phi_p(p, q) = e^{-\frac{(p\Delta_r)^2}{2\sigma_q^2}} \quad (15)$$

$$\frac{\Delta_\psi}{2\pi} \sum_{q=0}^{Q-1} \sigma_q^2 \leq \alpha, \quad \alpha \geq 0. \quad (16)$$

The computation of the discrete, optimal-kernel TFD requires samples of the AF on a polar grid, $A_p(p, q)$. These are computed either directly, or via interpolation from a rectangularly-sampled AF. Solving the optimization problem (14)–(16) for the optimal spread vector, $\underline{\sigma}^*$, determines polar samples of a radially-Gaussian kernel, Φ_p . Interpolation of the kernel to a rectangular grid and multiplication by the rectangularly-sampled AF yields the characteristic function, whose two-dimensional discrete Fourier transform is the optimal-kernel TFD.

3 KERNEL SOLUTION

The constrained optimization problem (14)–(16) that defines the optimal kernel is nonlinear in the variables σ_q . In this section, we propose a simple iterative solution algorithm that is based on the principle of gradient ascent, with an additional projection step that ensures that the volume constraint (16) remains satisfied. Although the presence of local minima on the error surface is, at present, an open question, in experiments this method has converged to good solutions.

3.1 Gradient Ascent Algorithm

Gradient ascent methods solve optimization problems by iteratively updating an estimate, $\underline{g}(k)$, of the optimal solution such that the value of the performance measure increases each iteration. The elements of the gradient vector, $\nabla(k)$, at iteration k of the algorithm are obtained by differentiating the performance measure (14) with respect to σ_q :

$$\nabla_q(k) = \frac{\Delta_\psi^4 \Delta_\psi}{\sigma_q^3(k)} \sum_{p=1}^{P-1} p^3 |A_p(p, q)|^2 e^{-\frac{(p\Delta_\psi)^2}{\sigma_q^2(k)}}, \quad q = 0, \dots, Q-1. \quad (17)$$

The steepest ascent algorithm is defined by the following iteration:

$$\underline{g}(k+1) = \underline{g}(k) + \mu \nabla(k), \quad (18)$$

where the step size, μ , is a positive constant small enough to ensure stability of the iterations. The spread vector is initialized such that the kernel is circularly symmetric with volume α , that is,

$$\underline{g}(0) = \sqrt{\frac{2\pi\alpha}{Q\Delta_\psi}} [1, \dots, 1]^T. \quad (19)$$

Since $\nabla(k) \geq 0$ [see (17) and recall that $\underline{g} \geq 0$], the volume of the kernel will increase at each iteration of this algorithm, in violation of the kernel volume constraint. To ensure that the volume remains bounded by α , the spread vector is scaled every iteration using

$$\underline{g}(k+1) \leftarrow \underline{g}(k+1) \sqrt{\frac{2\pi\alpha}{\Delta_\psi \sum_q \sigma_q^2(k+1)}}. \quad (20)$$

Since this normalization leaves the ratios of the elements of $\underline{g}$ intact, it returns the volume of the kernel to α , without affecting its shape.

Each iteration of the gradient ascent algorithm involves $O[PQ]$ computations, so the total cost of finding the optimal kernel is $O[PQI]$, where I is the number of iterations required for convergence.

3.2 Kernel Volume Selection

Since α controls the volume of the optimal kernel, it also controls the tradeoff between cross-component suppression and auto-component smoothing. A set of reasonable upper and lower bounds for α is given by

$$1 \leq \alpha \leq 5. \quad (21)$$

There appears to be no benefit to smoothing more than a spectrogram, so the lower bound represents the volume of a spectrogram kernel. The upper bound can be heuristically derived by finding the value of α that corresponds to a 10% increase in the radius of support of a circularly-symmetric, Gaussian auto-component. Increasing α further has little effect on the auto-component, but, potentially, passes more cross-component energy into the TFD.

4 EXAMPLES

In order to compare the optimal-kernel TFD with the fixed-kernel TFD's discussed in the Introduction, the optimal radially-Gaussian kernel was computed for the same chirp signal. The rectangularly-sampled AF and kernel were of size 64×64 , the polar-sampled AF was of size 64×32 , the volume parameter was set to $\alpha = 2$, and a step size $\mu = 25$ was

used in the gradient ascent algorithm. The spread vector converged in 28 iterations. The optimal-kernel TFD is shown in Fig. 4; equal-energy contour plots of the rectangular-sampled AF and optimal kernel are given in Figs. 5 and 6. The cross-component visible in the fixed-kernel TFD's is virtually eliminated, yet the distribution is still quite concentrated. The main difference between the present result and the 1/0 optimal-kernel TFD of Fig. 3 is the absence of ringing.

The performance of the design procedure under noisy conditions was tested by embedding a signal composed of the sum of two chirp signals that cross in time-frequency in -3 dB, additive, white, Gaussian noise. The Wigner distribution of the signal is shown in Fig. 7; the optimal-kernel TFD is shown in Fig. 8. The noise and cross-component are suppressed, while the auto-components remain quite concentrated.

5 CONCLUSION

In this paper, we have presented an optimization formulation for designing signal-dependent kernels that are based on radially-Gaussian functions. The method possesses several attractive properties. First, the procedure is based on optimality criteria and is not *ad hoc*. Second, the procedure is automatic. Third, the optimization criteria are formulated so that the resulting TFD is insensitive to the time-scale and orientation of the signal in time-frequency. Finally, the examples demonstrate that the optimal-kernel TFD offers excellent performance for a larger class of signals than any current fixed-kernel representation. Worthy of special mention is the excellent performance in the presence of substantial additive noise, which suggests that the technique may prove useful for automatic detection of unknown signals in noise.

The computational cost of finding the optimal-kernel TFD is $O[PQI]$. Since, in experiments, the value of I has typically been approximately 30, the cost of this technique is only a few times greater than both fixed-kernel methods and the 1/0 optimal kernel method.

The optimization formulation of this paper is reminiscent of that in [4], with the additional constraint that the kernel be a radially-Gaussian function. This constraint results in an optimal-kernel TFD with less ringing. However, the inherent smoothness of the radially-Gaussian kernel could result in less cross-component suppression for signal components that are closely spaced in time-frequency.

REFERENCES

1. L. Cohen, "Time-Frequency Distributions — A Review," *Proceedings of the IEEE* 77(7), pp. 941-981, July 1989.
2. P. Flandrin, "Some Features of Time-Frequency Representations of Multicomponent Signals," *IEEE ICASSP-1984*, pp. 41.B.4.1-41.B.4.4, 1984.
3. 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 ASSP-37*(6), pp. 862-871, June 1989.
4. R.G. Baraniuk and D.L. Jones, "Optimal Kernels for Time-Frequency Analysis," *SPIE 1990 Symposium on Optical and Optoelectronic Applied Science and Engineering*, San Diego, California, July 8-13, 1990.

Figure 1. Wigner distribution

Figure 2. Choi-Williams distribution

Figure 3. Optimal 1/0 kernel distribution

Figure 4. Optimal radially-Gaussian kernel TFD

Figure 5. Ambiguity function

Figure 6. Optimal radially-Gaussian kernel

Figure 7. Wigner distribution of signal in noise

Figure 8. Optimal radially-Gaussian kernel TFD