A SIMPLE SCHEME FOR ADAPTING TIME-FREQUENCY REPRESENTATIONS

Douglas L. Jones[◇] *and Richard G. Baraniuk*[□ *]

[◇] Coordinated Science Laboratory
University of Illinois
1308 West Main Street
Urbana, IL 61801
E-mail: d-jones@csl.uiuc.edu

[□] Department of Electrical and Computer Engineering
Rice University
P.O. Box 1892, Houston, TX 77251–1892
E-mail: richb@rice.edu

Appears in *IEEE Transactions on Signal Processing*, Vol. 42, No. 12, Dec. 1994

Abstract

Signal-dependent time-frequency representations, by adapting their functional form to fit the signal being analyzed, offer many performance advantages over conventional representations. In this paper, we propose a simple, efficient technique for continuously adapting time-frequency representations over time. The procedure computes a short-time quality measure of the representation for a range of values of a free parameter and estimates the optimal parameter value maximizing the quality measure via interpolation. Many representations, including the short-time Fourier transform, the cone-kernel distribution, and the continuous wavelet transform, support adaptation, at a computational cost of only a few times that of the corresponding static representations.

*Sponsored by the National Science Foundation, Grant No. MIP 90-12747, the Joint Services Electronics Program, Grant No. N00014-90-J-1270, and the Sound Group of the Computer-based Education Research Laboratory at the University of Illinois.

1 Introduction

Most time-frequency representations (TFRs) employ some kind of smoothing kernel, window, or filter to reduce noise and cross-components [1,2]. The choice of kernel dramatically affects the appearance and quality of the resulting TFR. It has been shown that, according to several different measures of performance, the optimal kernel depends on the signal being analyzed [3–5]. Therefore, utilization of any fixed kernel severely limits the class of signals for which the resulting time-frequency representation can perform well.

As an example, consider the signal shown in Fig. 1. It contains several narrow pulses, two sinusoids that overlap in time, and a Gaussian component. Figure 2 illustrates three short-time Fourier transforms (STFTs) of this signal computed using Gaussian windows of varying lengths. A short window (Fig. 2(a)) matches the pulse components well, but smears the sinusoidal and Gaussian components in the frequency direction. A medium-length window (Fig. 2(b)) matches the Gaussian component well, but smears the pulse and sinusoidal components in both the time and frequency directions. A long window (Fig. 2(c)) matches the sinusoidal components, but smears the pulse and Gaussian components in the time direction. These figures illustrate the fundamental drawback of the STFT: it is impossible to obtain simultaneously good time and good frequency resolution using a single fixed window. The continuous wavelet tranform [6] suffers from the same tradeoff, although in this case the tradeoff is a function of frequency. Moreover, a related tradeoff between time-frequency resolution and cross-term suppression applies to the kernel function in all bilinear (Cohen's class) time-frequency representations. In short, no TFR employing a fixed window, wavelet,or kernel performs well for all signals.

Due to this fundamental limitation of fixed windows or kernels, several researchers have developed signal-dependent or adaptive TFRs (see [3–5, 7–9] and the references in [4]). These methods often exhibit performance far surpassing that of fixed-kernel representations; however, they are either very computationally expensive or perform only off-line, block analysis of short signals. Thus, a need exists for simple, time-adaptive, computationally efficient TFRs suitable for real-time, on-line applications.

This paper presents a new scheme for adapting a TFR with a single free parameter. Section 2 describes the optimization formulation, which is based on maximizing the short-time concentration of the TFR. The proposed method supports on-line parameter adaptation over time, and applies to many classes of TFRs. For example, it can optimize the instantaneous window length in a STFT, the τ-extent of the cone-kernel distribution [10], or the

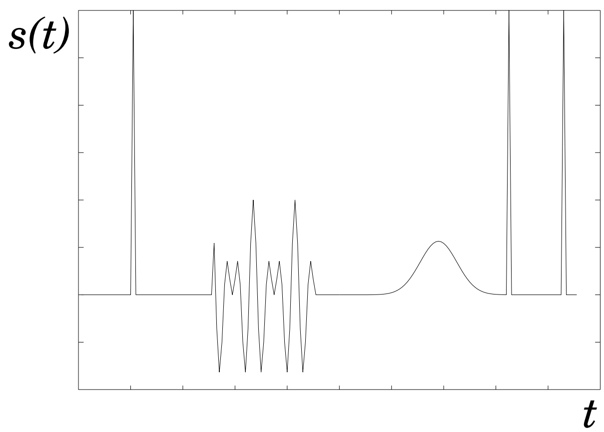


Figure 1: Test signal containing, from left to right, a unit pulse, the sum of two gated sinusoids, a Gaussian, and two unit pulses.

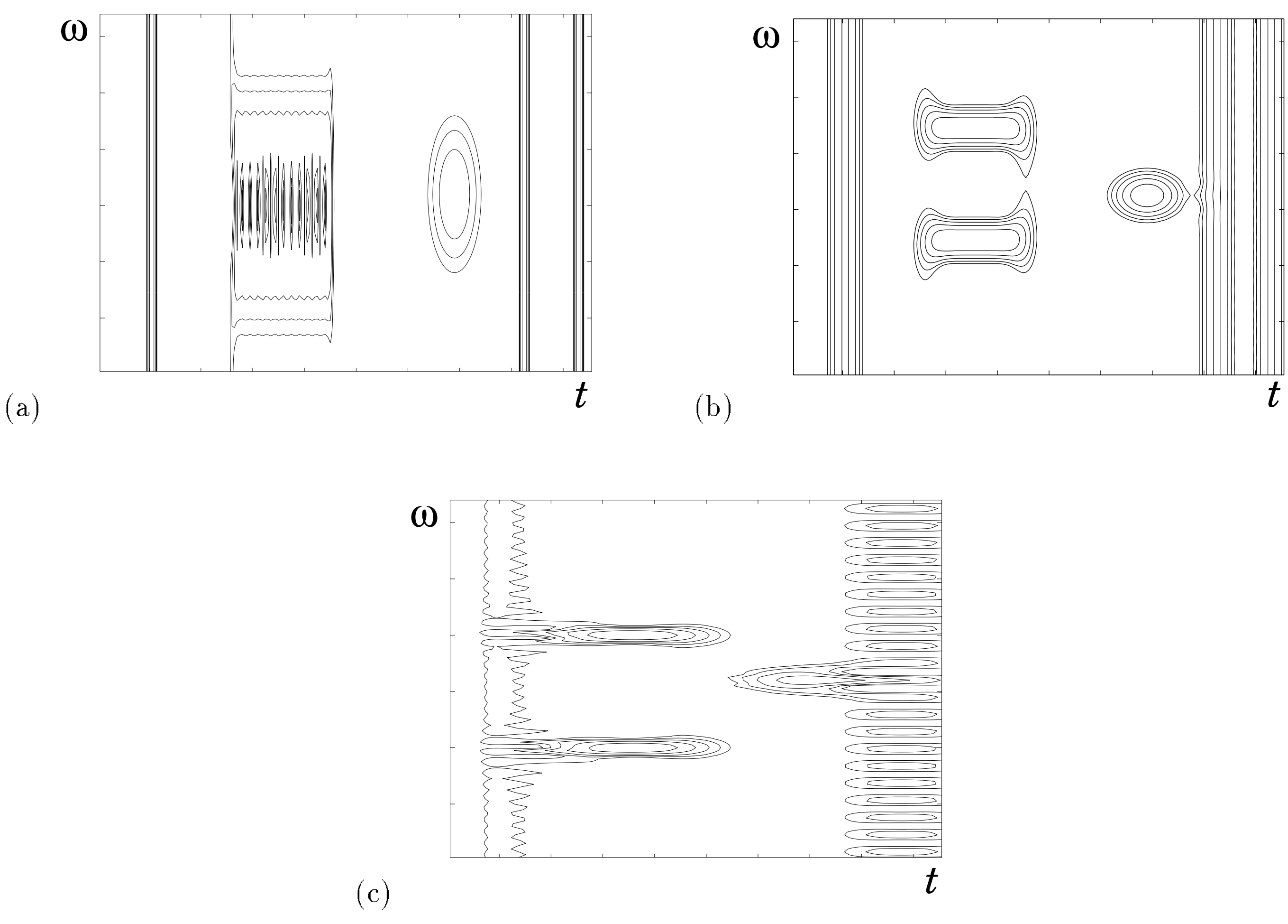


Figure 2: STFTs of the test signal computed using (a) short (3 sample), (b) medium (23 sample), and (c) long (47 sample) Gaussian windows.

"Q-factor" of the mother wavelet in a continuous wavelet transform. A fast algorithm, described in Section 3, reduces the computational cost of the adaptive scheme to only a few times that of an equivalent fixed-kernel TFR, making it viable for most TFR applications. As demonstrated in the examples of Section 4, the performance of the adaptive TFRs often greatly exceeds that of fixed-kernel representations.

2 The Adaptive Scheme

The development of a signal-dependent or adaptive TFR requires a means of determining an appropriate window or kernel function without extensive *a priori* knowledge of the signal characteristics. Procedures based on mathematical optimality criteria appear most promising [3–5]. In these approaches, optimization problems incorporating the signal to analyze yield an optimal adaptive window or kernel.

One of these successful approaches to adapting TFRs maximizes the concentration of a parameterized STFT. The algorithm described in [3] optimizes two STFT window parameters (window length and chirp rate) at each time-frequency location but, while effective, generally requires about three orders of magnitude more computation than a simple fixed-window STFT. We propose here a simpler concentration-based adaptive procedure that supports only time adaptivity of a single parameter and finds only an approximate maximizer of concentration. Due to this simplification, this approach requires only a few times the computation of a fixed-parameter transform. The adaptive technique developed here is not limited to adapting just STFT parameters; most TFRs parameterized by a single parameter p can be optimized.

To obtain a simple adaptive TFR from a TFR $D_p(\tau,\Omega)$ parameterized by a single parameter p, we adjust p over time to maximize a measure of short-time time-frequency concentration

$$C(t,p) \;=\; \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |D_p(\tau,\Omega)\; w(\tau-t)|^4 \, d\tau\, d\Omega}{\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |D_p(\tau,\Omega)\; w(\tau-t)|^2 \, d\tau\, d\Omega\right)^2}\,. \tag{1}$$

Here $w(\tau)$ is a one-dimensional window function centered at $\tau = 0$. The optimal time-varying parameter of the TFR is thus defined as

$$p^\star(t) \;=\; \operatorname{argmax}_p C(t,p). \tag{2}$$

The basis for this optimization approach rests on the intuition that a parameter providing high time-frequency concentration results in good time-frequency localization and high resolution. The ratio of the L_4 norm to the L_2 norm of the TFR favors "peaky" distributions that place as much signal energy into as small a region of time-frequency as possible, thus achieving a concentrated representation.

The local concentration measure $C(t,p)$ is a simplified version of the concentration measure utilized in [3] and is similar to kurtosis in statistics and also to a heuristic definition of entropy used for minimum entropy deconvolution in seismic signal processing [11]. Several other concentration, entropy, or peakiness measures have been studied in seismic applications [12]; for the purposes of adaptive time-frequency analysis, they yield essentially identical results and can be used interchangeably.

3 Efficient Implementation

Ideally, the optimal TFR parameter $p^\star(t)$ at each time t would be selected by computing the short-time concentration $C(t,p)$ as a function of the continuous-valued parameter and choosing the maximizing value. However, determination of the concentration for a large number of parameter values could be quite expensive, because each value of the parameter corresponds to a different TFR. We propose that the short-time concentration be computed only for several discrete values $p_i, i = 1, \ldots, P$, of the parameter, over a range of values that includes its maximum and minimum acceptable values. The resulting short-time concentrations $C(t,p_i)$ represent samples of $C(t,p)$, so an estimate of the optimal parameter value $p^\star(t)$ can be obtained by interpolating between the samples $C(t,p_i)$ to find the optimal value of p. Fortunately, as a rule the concentration measure is well-behaved and slowly varying with p; thus $C(t,p)$ can be sampled very coarsely with very little degradation in the final result. Experimentally, sampling as coarsely as once or twice per octave and using quadratic or cubic polynomial interpolation yields excellent results. Once the optimal parameter value $p^\star(t_0)$ is computed at time t_0, the time slice $D_{p^\star(t_0)}(t_0,\omega)$ of the optimal-parameter TFR can be computed based on that value.

Computation of the short-time concentrations represents the major expense of the adaptive algorithm. Fortunately, all of the information in $D_p(t,\omega)$ necessary to compute $C(t,p)$ can be summarized in two one-dimensional functions, obtained as the L^4 and L^2 norms of each time slice of the transform

$$c_4(\tau,p) \;=\; \int_{-\infty}^{\infty} |D_p(\tau,\Omega)|^4 \, d\Omega, \qquad c_2(\tau,p) \;=\; \int_{-\infty}^{\infty} |D_p(\tau,\Omega)|^2 \, d\Omega. \tag{3}$$

Using (3), the short-time concentration measure can be computed as

$$C(t,p) \;=\; \frac{\int_{-\infty}^{\infty} c_4(\tau,p)\,|w(\tau-t)|^4\,d\tau}{\left(\int_{-\infty}^{\infty} c_2(\tau,p)\,|w(\tau-t)|^2\,d\tau\right)^2} \;:=\; \frac{C_4(t,p)}{(C_2(t,p))^2}\,. \tag{4}$$

If the window function $w(\tau)$ has compact support (as it will in practice), then the integrals in (4) are over finite intervals, and only the portions of $c_4(\tau,p)$ and $c_2(\tau,p)$ centered around time t need be stored.

In practice, we use TFRs $D_p(n,k)$ discretized on a rectangular grid in time and frequency. The integrals in (3) are then computed as sums, for each time index n, over the frequency index k. With a rectangular weighting window of extent $\pm M$ samples, the numerator integral $C_4(t,p)$ in (4) reduces to the recursion

$$C_4(n,p) \;=\; \sum_{n-M}^{n+M} c_4(n,p) \;=\; C_4(n-1,p) + c_4(n+M,p) - c_4(n-M-1,p), \tag{5}$$

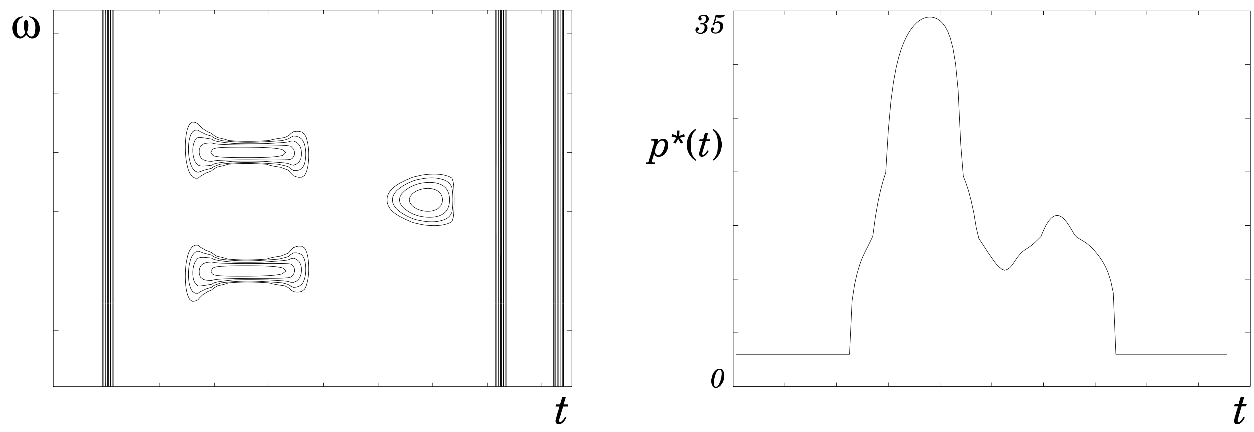
which can be updated at negligible cost. A similar recursion holds for the denominator integral $C_2(t,p)$. Finally, due to the inherent smoothness of $D_p(t,\omega)$, only modest oversampling in time and frequency is necessary to accurately estimate the concentration. Thus the test TFRs $D_{p_i}(n,k)$ used in practice can be coarsely sampled, greatly reducing the overall computational cost of the adaptive algorithm.

The total computational cost of this algorithm equals the cost of computing P coarsely sampled fixed-parameter TFRs with parameter values p_i, plus the minor cost of updating and interpolating the short-time concentration measures using (3) and (5), plus the cost of computing the TFR using the optimal parameter. The value of P necessary for excellent performance is usually less than ten, and the coarsely sampled, fixed-p TFRs cost substantially less to compute than the full-resolution optimal-parameter TFR. The total cost of the adaptive approach thus ranges from only two to ten times that of a fixed-parameter TFR.

4 EXAMPLES

4.1 Adaptive Short-Time Fourier Transform

As illustrated earlier in Fig. 2, the choice of STFT window length greatly affects the resulting TFR. No fixed window can work well for signal components of widely varying duration. Figure 3(a) presents an adaptive-window STFT of the 190-sample signal shown in Fig. 1.



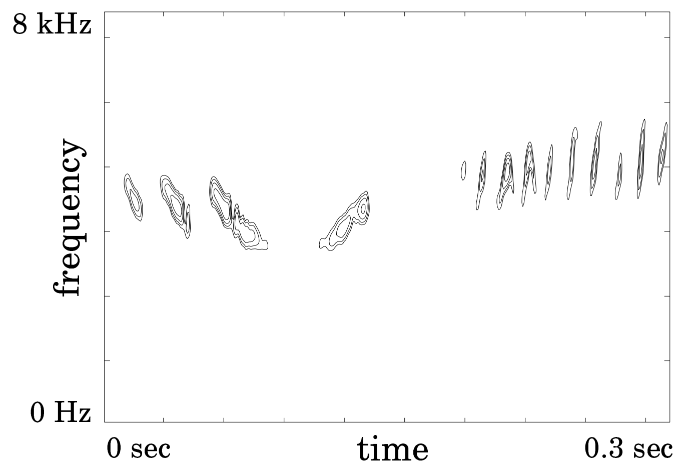
(a) (b)

Figure 3: (a) Optimal-window STFT of the test signal from Fig. 1. (b) Optimal time-varying window length $p^\star(t)$. The window length was varied to maximize the value of the short-time kurtosis concentration measure (4).

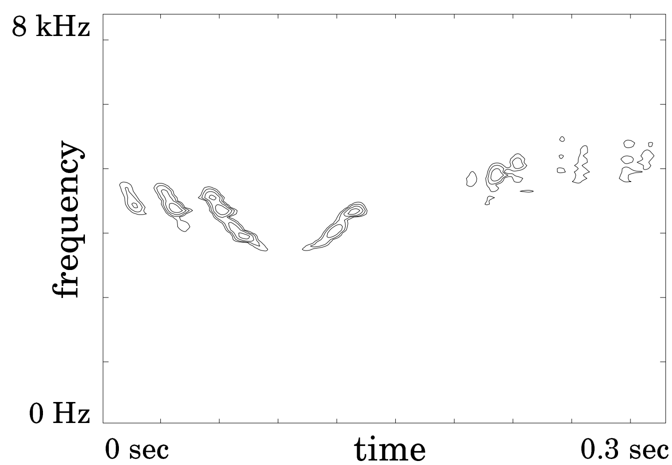
The length of a Gaussian window was selected as the adjustable parameter p for the algorithm of Section 3. The short-time concentration (4) for five different window lengths ($p = 3, 6, 12, 24, 48$ samples) was computed at each time location, and the optimal window length was interpolated using a cubic polynomial. The length of the weighting window w in (4) was set to 61 samples. (The weighting window length should always be at least slightly longer than the longest trial window length p_P, to provide unbiased concentration estimates.) In comparison to the STFTs of Fig. 2, Fig. 3(a) exhibits the benefits of adaptivity by closely matching the time-duration of the most significant signal component at each time. Figure 3(b) shows the optimal window length as a function of time.

The abrupt right edge of the Gaussian component in Fig. 3(a) results from its proximity to the second pulse function, whose large energy causes the adaptation to optimize for it rather than for the Gaussian component as it enters the short-time concentration window. This behavior is characteristic of this family of adaptive TFRs, in that the adaptation tends to optimize the performance for the largest-energy component in the short-time concentration window w. The performance may not be well optimized for nearby components of smaller energy if they have very different time-frequency characteristics.

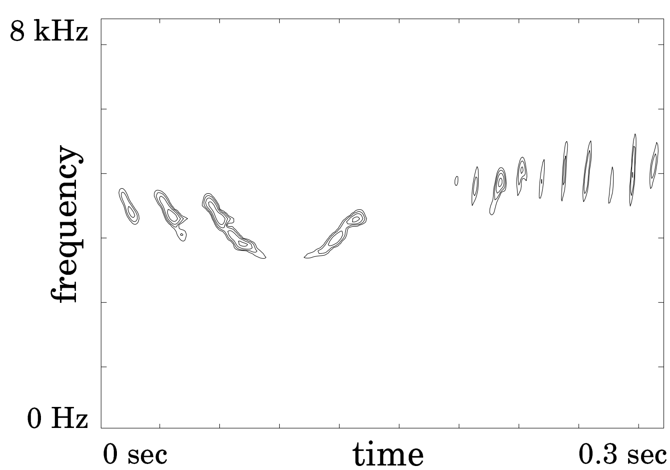
Figure 4 depicts several STFTs of the bird song of the Cerulean Warbler (*Dendroica cerulea*) [13]. Figure 4(a) shows the STFT computed using a short, 64-point Gaussian window, and Fig. 4(b) shows the STFT computed using a longer, 256-point Gaussian window. The short window is best for the ending "buzzy" portion of the signal, while the longer



(a)



(b)



(c)

Figure 4: Comparison of several TFRs of a portion of the bird song of a Cerulean Warbler (*Dendroica cerulea*). The sampling rate was 16 kHz. (a) STFT computed using a short (64-sample) Hamming window. (b) STFT computed using a long (256-sample) Hamming window. (c) STFT computed using the optimal time-varying window.

window is more concentrated in the chirping portion. The adaptive-window STFT, computed using window lengths between these two extremes, is given in Fig. 4(c). The length of the weighting window w was set to 200 samples. This figure illustrates the benefits of time adaptivity, since neither fixed-window STFT works well for all portions of the signal.

4.2 Adaptive Cone-Kernel TFR

Heretofore, we have discussed mainly adaptive STFTs. However, the adaptive scheme proposed above is quite general and applies to any TFR that can be parameterized in terms of a single parameter. An important example is the cone-kernel distribution [10], which exhibits the important property of preservation of outer time support. A "τ-extent" parameter characterizes the cone-kernel proposed in [10]; the quality of the resulting TFR often depends greatly on this parameter. Figures 5(a), (b), and (c) illustrate the cone-kernel distribution of the test signal of Fig. 1 for three choices of τ-extent (9, 65, and 128 samples, respectively).

Figure 5(d) illustrates an adaptive cone-kernel distribution in which the τ-extent parameter is adapted to minimize a measure of *short-time entropy*, which we define as

$$\begin{aligned} E(t,p) \;&=\; -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a_r(t)^{-1}\,|D_p(\tau,\Omega)\,w(\tau-t)|^r \\ &\quad\cdot \log\left(a_r(t)^{-1}\,|D_p(\tau,\Omega)\,w(\tau-t)|^r\right)\,d\tau\,d\Omega. \end{aligned} \tag{6}$$

Here, $a_r(t)$ is a time-varying normalization factor

$$a(t) \;=\; \int_{-\infty}^{\infty} |D(\tau,\Omega)\,w(\tau-t)|^r\,d\tau\,d\Omega, \tag{7}$$

and $r > 0$ is a constant. The magnitude of $D_p(\tau,\Omega)$ is used in the entropy calculation, because the cone-kernel distribution can take on negative values. Alternative approaches could be based on the positive part of $D_p(\tau,\Omega)$ or on short-time versions of the generalized Rényi entropy discussed in [14]. Regardless of the exact details, short-time entropies can be computed efficiently in a manner similar to (3), (4). For Fig. 5(d), six different τ-extents, from $p_1 = 5$ to $p_6 = 128$, were used, and the optimal value was interpolated using a quadratic fit. The kurtosis concentration measure (4) yields almost identical results. The optimal time-varying τ-extent parameter is graphed in Fig. 5(e).

While long cone lengths can yield good results for isolated components of short duration, artifacts due to interactions between different components in the cone interval can appear. For example, the medium cone-length distribution of Fig. 5(b) contains artifacts between the Gaussian component and the following impulse. The adaptive technique avoids these artifacts by using a short-length cone over that region.

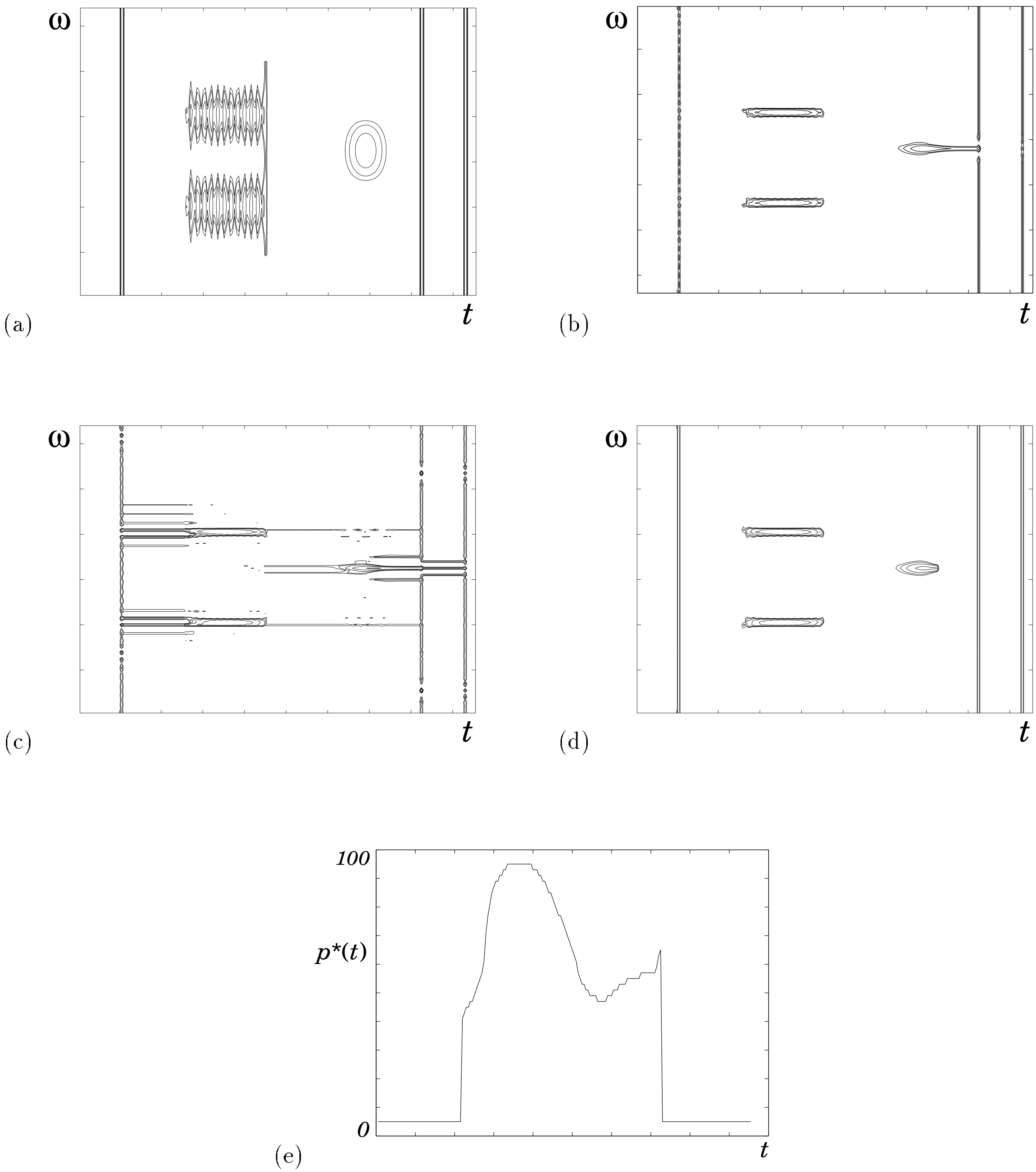


Figure 5: Cone-kernel distributions of the test signal computed using (a) small (9 sample), (b) medium (65 sample), (c) large (128 sample), and (d) the optimal time-varying τ-extent parameters. The optimal τ-extent parameter, graphed in (e), is varied to minimize the value of the short-time entropy concentration measure (6).

4.3 Adaptive Wavelet Transform

As a final example, consider the continuous wavelet transform (CWT) [6]. The CWT is an attractive alternative to the STFT in applications where the durations of the signal components scale inversely with their frequency. Given a bandpass wavelet function g whose Fourier transform is centered at frequency ω_0 and has bandwidth B, the CWT of a signal s is defined as

$$D(t,\omega) \;=\; \int_{-\infty}^{\infty} s(u)\, g^*\!\left(\frac{\omega}{\omega_0}(u-t)\right) du. \tag{8}$$

The CWT is a constant-Q TFR, with the Q-factor of the analysis being the center-frequency-bandwidth quotient ω_0/B. The Q of the wavelet has a large effect on the quality of the representation of the CWT in time-frequency. Setting Q too small results in excessive smearing of signal components in the frequency direction; setting Q too large results in excessive smearing of signal components in the time direction. One commonly used wavelet with adjustable center frequency and bandwidth is given by[1]

$$g_{\omega_0,\sigma}(t) = \exp(-t^2/2\sigma^2)\cos(\omega_0 t); \tag{9}$$

Either ω_0 or σ can be adapted to maximize concentration in the CWT. The required CWTs can be efficiently computed using the algorithm in [15].

5 Conclusions

This paper presents a simple procedure for the automatic optimization and time adaptation of many TFRs with a kernel/window/wavelet adjustable by a single parameter. Even this modest amount of adaptivity often yields greatly improved results, and in some cases achieves most of the benefits promised by more complicated signal-dependent time-frequency representations. This procedure generally requires only a few times the cost of a fixed-parameter distribution and supports on-line, real-time computation. Two or more paramenters can be adapted using a multi-dimensional version of this algorithm, but the cost increases geometrically with the number of parameters. By supporting adaptivity at a cost comparable with traditional methods, this technique makes adaptive TFRs viable for most applications.

[1]Note that this wavelet does not exactly satisfy the wavelet admissibility condition [6].

Acknowledgments

The authors wish to thank a reviewer for helpful comments and François Auger for finding an error in the manuscript. We are greatly indebted to Henry Kwok and Michael Kramer for implementing efficient Matlab code for the fast STFT algorithm, and to Quang Tran for testing this code. This work was supported by the Joint Services Electronics Program, Grant No. N00014-90-J-1270, the National Science Foundation, Grant No. MIP 90–12747, and the Sound Group of the Computer-based Education Research Laboratory at the University of Illinois.

References

[1] L. Cohen, "Time-frequency distributions — A review," *Proceedings of the IEEE*, vol. 77, pp. 941–981, July 1989.

[2] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, April 1992.

[3] 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, pp. 2127–2135, December 1990.

[4] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 41, pp. 1589–1602, April 1993.

[5] R. G. Baraniuk and D. L. Jones, "A radially Gaussian, signal-dependent time-frequency representation," *Signal Processing*, vol. 32, pp. 263–284, June 1993.

[6] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Magazine*, vol. 8, pp. 14–38, October 1991.

[7] D. L. Jones and R. G. Baraniuk, "An adaptive optimal kernel time-frequency representation," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing — ICASSP '93*, 1993.

[8] P. J. Loughlin, J. W. Pitton, and L. E. Atlas, "An information theoretic approach to positive time-frequency distributions," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing — ICASSP '92*, pp. V125–V128, 1992.

[9] G. Jones and B. Boashash, "Window matching in the time-frequency plane and the adaptive spectrogram," in *IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, (Victoria, BC, Canada), pp. 87–90, October 1992.

[10] Y. Zhao, L. E. Atlas, and R. J. Marks, "The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp. 1084–1091, July 1990.

[11] R. A. Wiggins, "Minimum entropy deconvolution," *Geoexploration*, vol. 16, pp. 21–35.

[12] D. Donoho, "On minimum entropy deconvolution," in *Applied Time-Series Analysis II* (D. F. Findley, ed.), pp. 556–608, New York: Academic Press, 1981.

[13] D. J. Boorer, *Common Bird Songs (Audio Cassette).* New York: Dover, 1984.

[14] P. Flandrin, R. G. Baraniuk, and O. Michel, "Time-Frequency complexity and information," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing — ICASSP '94*, 1994.

[15] D. L. Jones and R. G. Baraniuk, "Efficient approximation of continuous wavelet transforms," *Electronics Letters*, vol. 27, pp. 748–750, April 25, 1991.