ADAPTIVE STFT WITH CHIRP-MODULATED GAUSSIAN WINDOW

Soo-Chang Pei and Shih-Gu Huang

Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan

ABSTRACT

In this paper, we propose an adaptive STFT (ASTFT) with adaptive chirp-modulated Gaussian window. The window is obtained from rotating Gaussian function in time-frequency plane by fractional Fourier transform (FRFT). It is completely adaptive where the two parameters, FRFT rotation angle and Gaussian variance, are signal-dependent. The angle dependents on the chirp rate of the signal. The variance is determined by the chirp rate and its first derivative. Considering the input may be multicomponent, a chirp-modulated Gaussian window with time-frequency-varying angle and variance is developed. The proposed ASTFT has very high energy concentration and less interference between components in noiseless and noisy environments.

Index Terms— Adaptive time-frequency analysis, shorttime Fourier transform, concentration measure, fractional Fourier transform, instantaneous frequency.

1. INTRODUCTION

No single time-frequency representation (TFR) with fixed values of parameters can be the universal choice for all kinds of signals. In order to achieve high energy concentration for a variety of signals, adaptive (signal-dependent) TFRs are attracting more attentions in recent years. A comprehensive introduction of adaptive TFRs has been provided in [1, 2]. When designing an adaptive TFR, the choice of the type of TFR and the design of adaptive approach are the two core concerns. Concentration measure [3–6], adaptive parametric time-frequency analysis based on atomic decomposition [7–9] and adaptive kernel based on chirp rate [10–13] are some popular adaptive approaches.

In [13], an adaptive STFT (ASTFT) based on chirp rate was proposed. The ASTFT uses Gaussian window function with window width (i.e. variance) dependent on the input signal. In this paper, we further improve the energy concentration of the ASTFT by using chirp-modulated Gaussian window instead, which is obtained from rotating the Gaussian window (with variance ρ) in time-frequency plane by the fractional Fourier transform (FRFT) [14, 15] (with rotation angle θ). It is completely adaptive, where its two parameters ρ and θ are signal-dependent. The optimal FRFT angle depends on the slope of the instantaneous frequency, i.e. chirp rate, of the signal. The optimal variance is determined by the chirp rate and its first derivative. Considering the input may be a multicomponent signal, there would be multiple chirp rates at the same time instants. And it follows that some components may have different values of optimal FRFT angles and variances at the same time instants. Accordingly, a chirp-modulated Gaussian window with time-frequency-varying FRFT angle and variance is developed.

The idea of rotating window function in the time-frequency plane can be applied to other kinds of TFRs and other kinds of window functions. Here, we adopt the STFT with Gaussian window function because the STFT is a linear TFR with additivity and reversibility properties. Besides, there is a closed-form expression for the chirp-modulated Gaussian function and a closed-form solution to its optimal variance.

The chirp-modulated Gaussian window (chirplet) has been used in the TFR based on concentration measure [3] and TFRs based on atomic decomposition [7,9]. However, their computational complexity is considerable because there's no closed-form solutions to their optimization problems. Besides, their adaptive approaches are less robust to noise. Highly noisy environment may yield high energy concentration in the noise part instead of the signal part.

2. PRELIMINARY

An linear TFR can be expressed as

$$\text{TFR}_x(t,f) = \int_{-\infty}^{\infty} x(\tau) h^*(t-\tau) e^{-j2\pi f\tau} d\tau, \quad (1)$$

where $h(\tau)$ is the window function. To design an adaptive window function for the STFT, one popular approach is based on the chirp rate, i.e. first derivative of the instantaneous frequency. In [11], the adaptive window function is time-varying

$$h_{\sigma(t)}(\tau) = \frac{1}{\sqrt{2\pi\sigma^2(t)}} e^{-\frac{1}{2\sigma^2(t)}\tau^2}.$$
 (2)

When $\sigma^2(t) = \sigma^2$ is fixed, it reduces to the conventional STFT. In (2), $\sigma^2(t)$ is determined by the chirp rate f'_{inst} :

$$2\sqrt{2\ln 2}\sigma(t) = \max_{l} 2l$$
 s.t. $\int_{t-l}^{t+l} |f'_{\text{inst}}(\tau)| d\tau \le \xi.$ (3)

However, this ASTFT is not completely adaptive because the threshold ξ is user-defined, and the estimation of the instan-

taneous frequency $f_{inst}(t)$ is not signal-dependent. Plus, there is no proof that (3) is optimal in some sense.

In [13], a more powerful ASTFT is developed using the following window function:

$$h_{\sigma(t,f)}(\tau) = \frac{1}{\sqrt{2\pi\sigma^2(t,f)}} e^{-\frac{1}{2\sigma^2(t,f)}\tau^2},$$
 (4)

where $\sigma^2(t, f)$ is time-frequency-varying because there may be multiple different chirp rates at the same time instant for multicomponent signals. The optimal value of $\sigma^2(t, f)$ is

$$\sigma^{2}(t,f) = \frac{1}{2\pi |f'_{\text{inst}}(t,f)|},$$
(5)

where $f_{\text{inst}}(t, f)$ is the instantaneous frequencies. If the input is a monocomponent linear FM signal, then the relation reduces into $\sigma^2(t) = 1/(2\pi |f'_{\text{inst}}(t)|)$. This relation is more concise and more accurate than that in (3). Besides, an adaptive low-complexity algorithm was proposed in [13] for instantaneous frequency estimation.

3. ASTFT WITH CHIRP-MODULATED GAUSSIAN WINDOW

The time-frequency resolution of the ASTFT in [13] is still limited by the Heisenberg uncertainty principle. To solve this problem, we replace the Gaussian window by chirpmodulated Gaussian window.

3.1. Chirp-Modulated Gaussian Function

FRFT [14, 16] can produce a rotation in the time-frequency plane. The FRFT with rotation angle $\theta \neq 0$ is defined as

$$x_{\theta}(u) = \sqrt{\frac{1-j\cot\theta}{2\pi}} e^{j\frac{\cot\theta}{2}u^2} \int_{-\infty}^{\infty} e^{-j\csc\theta ut + j\frac{\cot\theta}{2}t^2} x(t) dt.$$

Consider a Gaussian function with variance ρ , denoted by $g_{\rho}(t)$. Performing the FRFT to rotate $g_{\rho}(t)$ clockwise in the time-frequency plane with angle θ , we have $g_{\rho,\theta}(u)$. Normalizing $g_{\rho,\theta}(u)$ leads to the chirp-modulated Gaussian function:

$$h_{\rho,\theta}(u) = c_0 e^{-\frac{1}{2}\frac{\rho(1+\cot^2\theta)}{1+\rho^2\cot^2\theta}u^2} e^{-j\frac{1}{2}\frac{\cot\theta(\rho^2-1)}{1+\rho^2\cot^2\theta}u^2}, \quad (6)$$

where $c_0 = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\rho(1+\cot^2\theta)}{1+\rho^2\cot^2\theta}}$, and note that $\cot(\arctan(a)) = 1/a$. Because $h_{\rho,\theta\pm\pi}(u) = h_{\rho,\theta}(u)$, let $\theta \in (-\pi/2, \pi/2]$.

3.2. Wy Using Chirp-Modulated Gaussian Window?

ASTFT using chirp-modulated Gaussian window with timevarying variance $\rho(t)$ and FRFT angle $\theta(t)$ is defined as

$$X_{\rho(t)}^{\theta(t)}(t,\omega) = \int_{-\infty}^{\infty} x\left(\tau\right) h_{\rho(t),\theta(t)}^{*}(t-\tau) e^{-j\omega\tau} d\tau.$$
(7)



Fig. 1. The straight line shows the instantaneous frequency of the input signal. The ellipses are the 3 dB contour plots of the WVDs of (a) Gaussian windows with different variances and (b) chirp-modulated Gaussian windows with same variance but different FRFT angles.

First, consider the simplest case that x(t) is a linear FM signal, say, $x(t) = \exp\left(j(\frac{a}{2}t^2 + bt)\right)$. Then, the variance and FRFT angle don't need to change over time, i.e. $\rho(t) = \rho$ and $\theta(t) = \theta$. The WVD of x(t) is a straight line in the time-frequency plane, as illustrated in Fig. 1. The WVD of the original Gaussian window (i.e. $\theta = 0$) is a 2D Gaussian function with shape depending on the variance ρ . In Fig. 1(a), the ellipses represent the 3 dB contour plots of the 2D Gaussian functions with different ρ . The squared magnitude of the ASTFT can be interpreted as the 2D convolution of the WVDs of x(t) and $h_{\rho,\theta}(t)$, i.e. the straight line blurred by the 2D Gaussian function. Accordingly, one needs to find the optimal variance (shape) such that the blurring effect is minimal. In [13], it has been derived that the optimal solution depends on the chirp rate and is $\rho = 1/|a|$ in this example.

Next, consider the chirp-modulated Gaussian window $h_{\rho,\theta}(t)$. The ellipses in Fig. 1(b) shows the 3 dB contour plots of the WVDs of $h_{\rho,\theta}(t)$ with the same variance ρ but different rotation angles θ . It is obvious that the optimal θ depends on the chirp rate and is $\theta = -\arctan(a)$ in this example. Comparing Fig. 1(b) with Fig. 1(a), the chirp-modulated Gaussian window yields less blurring effect than the original Gaussian window if the variance is large enough.

3.3. Optimal Time-Varying FRFT Angle

Consider a more general form that $x(t) = A(t) \exp(j\varphi(t))$. The instantaneous frequency and the chirp rate are $\omega_{inst}(t) = \varphi'(t)$ and $\omega'_{inst}(t) = \varphi''(t)$, respectively. The signal in Fig. 1 has $\omega_{inst}(t) = at + b$ and $\omega'_{inst}(t) = a$, and the optimal FRFT angle is $\theta = -\arctan(a)$. If the chirp rate changes over time, the optimal FRFT angle should also change over time:

$$\theta_{\rm opt}(t) = -\arctan\left(\omega_{\rm inst}'(t)\right) = -\arctan\left(\varphi''(t)\right).$$
(8)

In Fig. 2(a), an example is given to illustrate the idea. The FRFT angle depends on the slope of the instantaneous frequency, i.e. the chirp rate. This idea can be applied to different kinds of window functions, but there may be no closedform expression for the FRFT of the window function.



Fig. 2. The straight line shows the instantaneous frequency of the input signal. The ellipses are the 3 dB contour plots of the WVDs of the chirp-modulated Gaussian windows with (a) fixed variance but time-varying FRFT angle, and (b) time-varying FRFT angle and time-varying variance.

3.4. Optimal Time-Varying Variance

In Fig. 1(b), larger variance (flatter ellipse) can yield higher energy concentration (less blurring effect). However, it is not true in Fig. 2(a). At time-frequency point p_2 , too large variance makes the energy concentration lower rather than higher. Accordingly, the variance should depends on the chirp rate and its derivatives, as illustrated in Fig. 2(b).

Recall the ASTFT defined in (7). At time instant t, the spectrogram $\left|X_{\rho(t)}^{\theta(t)}(t,\omega)\right|^2$ is a function of frequency ω and is normalized into $\zeta_t(\omega)$. Then, $\zeta_t(\omega)$ can be deemed as a distribution function. The variance of the distribution, also known as *instantaneous bandwidth* [17], is given by

$$B_t^2 = \int_{-\infty}^{\infty} \left(\omega - \langle \omega \rangle_t\right)^2 \zeta_t(\omega) d\omega, \qquad (9)$$

where $\langle \omega \rangle_t$ is the mean of the distribution, also known as *local frequency*. Obviously, smaller instantaneous bandwidth leads to higher energy concentration. Therefore, the optimal value of $\rho(t)$ occurs when B_t^2 is minimal.

Consider the general case that $x(t) = A(t) \exp(j\varphi(t))$. To simplify the optimization problem, assume the amplitude A(t) varies slowly and the high-order derivatives of the phase are close to zero, i.e. $\varphi^{(n)}(t) \approx 0$ for $n \geq 4$. Then, the instantaneous bandwidth reduces into

$$B_t^2 = \frac{1}{2} \left[1 + (\varphi'')^2 \right] \rho^{-1} + \frac{1}{8} \frac{(\varphi''')^2 \left[(\varphi'')^2 \rho^{-1} + \rho \right]^2}{\left[1 + (\varphi'')^2 \right]^2}, \quad (10)$$

where the argument t is omitted from $\varphi(t)$ and $\rho(t)$ for expression simplification. Minimal B_t^2 occurs when $\frac{\partial}{\partial \rho}B_t^2 = 0$, which leads to the following quartic eqaution:

$$(\varphi''')^2 \rho^4 - 2 \left[1 + (\varphi'')^2 \right]^3 \rho - (\varphi''')^2 (\varphi'')^4 = 0.$$
 (11)

 $\rho_{\rm opt}$ is the positive real root of the above equation, i.e.

$$\rho_{\rm opt} = S + \sqrt{-S^2 - q/(2S)},$$
(12)



Fig. 3. The solid lines are the instantaneous frequencies. The solid ellipses are the 3 dB contour plots of the WVDs of chirp-modulated Gaussian windows with optimal FRFT angles and optimal variances, while the dashed ellipses are those with interpolated FRFT angles and interpolated variances.

where $q = -(1 + {\varphi''}^2)^3 / {\varphi'''}^2$, $r = -{\varphi''}^4$, $S = \sqrt{\frac{Q}{12} + \frac{r}{Q}}$, and $Q = \sqrt[3]{6(9q^2 + \sqrt{81q^4 - 48r^3})}$. If other kind of window function is used, there may be no closed-form solution to ρ_{opt} , and then one needs to numerically search the optimal value of ρ such that B_t^2 is minimized.

In practice, the environment is usually noisy. The inaccurate A(t) and high-order $\varphi^{(n)}(t)$ may decrease rather than increase the accuracy of the variance. That's why we let the optimal variance in (12) depend only on $\varphi''(t)$ and $\varphi'''(t)$. However, in low noise environments, the performance may be improved without the assumptions that A(t) varies slowly and $\varphi^{(n)}(t) \approx 0$ for $n \geq 4$. In this case, one needs to numerically search the optimal variance. Even so, (12) can still be used as the initial estimate to fasten the optimization process under certain conditions.

3.5. Time-Frequency-Varying Chirp-Modulated Gaussian Window

For multicomponent signals, there would be multiple different chirp rates at the same time instant. Each component has its own optimal FRFT angle and optimal variance; Accordingly, an ASTFT with time-frequency-varying chirpmodulated Gaussian window is introduced:

$$X_{\rho(t,\omega)}^{\theta(t,\omega)}(t,\omega) = \int_{-\infty}^{\infty} x(\tau) h_{\theta(t,\omega),\rho(t,\omega)}^{*}(t-\tau) e^{-j\omega\tau} d\tau.$$
(13)

It is very difficult to obtain the optimal solutions to $\rho(t, \omega)$ and $\theta(t, \omega)$. Instead, a method similar to that in [13] is used.

A signal with two components is considered as an example. The solid lines in Fig. 3 represent the instantaneous frequencies. It is obvious that these two components has different chirp rates, and thus they have different optimal FRFT angles ($\theta_{opt,1}(t)$ and $\theta_{opt,2}(t)$) and different optimal variances ($\rho_{opt,1}(t)$ and $\rho_{opt,2}(t)$). For time-frequency (TF) points on the first line, $\theta_{opt,1}(t)$ and $\rho_{opt,1}(t)$ are used. For TF points on the second line, $\theta_{opt,2}(t)$ and $\rho_{opt,2}(t)$ are adopted. For TF



Fig. 4. Contour plots of (a) optimized S-transform [23], (b) adaptive LPFT [24], (c) ASTFT using original Gaussian window [13], and (d) proposed ASTFT. All components have time-varying amplitude uniformly between 0 and 20.

points between the two lines, FRFT angle between $\theta_{opt,1}(t_i)$ and $\theta_{opt,2}(t_i)$ and variance between $\rho_{opt,1}(t_i)$ and $\rho_{opt,2}(t_i)$ are reasonable choices. See the dashed ellipses in Fig. 3. For low complexity, we simply interpolate the values of the FRFT angle and the variance for TF points between lines.

3.6. Instantaneous Frequency Estimation

Instantaneous frequency (IF) estimation is indispensable because the proposed window function relies on the chirp rate. An overview of some estimation methods is given in [1,18]. In this paper, we modify the estimation algorithm in [13]. Instead of ridge detection method, Viterbi algorithm [19–21] is utilized to trace the optimal route that accumulates most energy in the time-frequency plane. One can use other kind of TFR and other kind of more robust estimation algorithm to improve the performance. For example, the estimation algorithm in [22] performs well in complicated environments, but the cost is higher computational complexity.

Obviously, the performance (resolution and robustness) of the proposed ASTFT is affected by the accuracy of the IF estimation algorithm. Thus, the proposed ASTFT might not be able to yield a more accurate IF estimate. However, it is more useful in applications like component separation and extraction because it is linear, additive and has very high resolution.

4. SIMULATIONS

For a fair comparison, the proposed ASTFT is only compared with some linear adaptive TFRs, including optimized Stransform [23], adaptive local polynomial Fourier transform (LPFT) [24] and ASTFT with original Gaussian window [13].

Consider a noiseless signal where the two components have time-varying amplitude uniformly distributed between 0 and 20. The contour plots of the four adaptive TFRs are shown in Figs. 4(a)-(d), respectively. The optimized



Fig. 5. Contour plots of (a) optimized S-transform [23], (b) adaptive LPFT [24], (c) ASTFT using original Gaussian window [13], and (d) proposed ASTFT. The amplitude of each component is random uniformly distributed between 2 and 3 at each time instant, and the SNR is 7 dB.

S-transform in (a) cannot achieve high enough energy concentration even though the kernel function is optimized. The adaptive LPFT in (b) has good resolution in some timefrequency regions but poor resolution in the others. Comparing (c) and (d), we can find out that rotating the Gaussian window can significantly enhance the energy concentration. Besides, the proposed ASTFT in (d) has the very clear resolution with little interference between components. We use the concentration measure (CM) in [13] to quantitatively evaluate the energy concentration. The higher the CM, the more concentrated the TFR's energy. The values of CM are 0.0088, 0.0083, 0.0088 and 0.0221 in (a)-(d), respectively.

In real world, the environment is usually noisy. Consider a noisy signal with three components. The amplitude of each component is random uniformly distributed between 2 and 3 at each time instant, and the SNR is 7 dB. The simulation result is shown in Fig. 5. Due to the noise, all the adaptive TFRs are more or less adaptive to the noise part instead of the signal part in some time-frequency regions. However, the proposed ASTFT is more robust to noise than the other adaptive TFRs. In this example, the values of CM are 0.0085, 0.0084, 0.0085 and 0.0226 in (a)-(d), respectively.

5. CONCLUSION

We propose an ASTFT with chirp-modulated Gaussian window, which is obtained from rotating the Gaussian window in the time-frequency plane by the FRFT. We derive the closedform expressions for the chirp-modulated Gaussian window and the optimal solutions of its two parameters, i.e. FRFT angle and variance. Considering the input may have multiple components, time-frequency-varying FRFT angle and variance are developed. The proposed ASTFT greatly outperforms the ASTFT in [13] and some other adaptive linear TFRs in noiseless and noisy environments.

6. REFERENCES

- E. Sejdić, I. Djurović, and J. Jiang, "Time-frequency feature representation using energy concentration: an overview of recent advances," *Digital signal processing*, vol. 19, no. 1, pp. 153–183, 2009.
- [2] B. Boashash, *Time-frequency signal analysis and processing: a comprehensive reference*, Academic Press, 2015.
- [3] D. L. Jones and T. W. Parks, "A high resolution data-adaptive time-frequency representation," Acoustics, Speech and Signal Processing, IEEE Transactions on, vol. 38, no. 12, pp. 2127–2135, 1990.
- [4] L. Stanković, "A measure of some time-frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621–631, 2001.
- [5] S. Aviyente and W. J. Williams, "Minimum entropy time-frequency distributions," *Signal Processing Letters*, *IEEE*, vol. 12, no. 1, pp. 37–40, 2005.
- [6] I. Djurović, E. Sejdić, and J. Jiang, "Frequency-based window width optimization for S-transform," AEU-International Journal of Electronics and Communications, vol. 62, no. 4, pp. 245–250, 2008.
- [7] A. Bultan, "A four-parameter atomic decomposition of chirplets," *IEEE Transactions on Signal Processing*, vol. 47, no. 3, pp. 731–745, 1999.
- [8] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001.
- [9] O. A. Yeste-Ojeda, J. Grajal, and G. Lopez-Risueno, "Atomic decomposition for radar applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 1, pp. 187–200, 2008.
- [10] H. Kawahara, I. Masuda-Katsuse, and A. De Cheveigne, "Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds," *Speech communication*, vol. 27, no. 3, pp. 187–207, 1999.
- [11] J. Zhong and Y. Huang, "Time-frequency representation based on an adaptive short-time Fourier transform," *Signal Processing, IEEE Transactions on*, vol. 58, no. 10, pp. 5118–5128, 2010.
- [12] T. K. Hon and A. Georgakis, "Enhancing the resolution of the spectrogram based on a simple adaptation procedure," *Signal Processing, IEEE Transactions on*, vol. 60, no. 10, pp. 5566–5571, 2012.

- [13] S.-C. Pei and S.-G. Huang, "STFT with adaptive window width based on the chirp rate," *Signal Processing*, *IEEE Transactions on*, vol. 60, no. 8, pp. 4065–4080, 2012.
- [14] H. M. Ozaktas and D. Mendlovic, "Fourier transforms of fractional order and their optical interpretation," *Opt. Commun.*, vol. 101, no. 3, pp. 163–169, 1993.
- [15] S.-C. Pei and S.-G. Huang, "Fast discrete linear canonical transform based on CM-CC-CM decomposition and FFT," *IEEE Transactions on Signal Processing*, vol. 64, no. 4, pp. 855–866, 2016.
- [16] S. C. Pei and S.-G. Huang, "Reversible joint Hilbert and linear canonical transform without distortion," *IEEE transactions on signal processing*, vol. 61, no. 17-20, pp. 4768–4781, 2013.
- [17] Leon Cohen, *Time-frequency analysis*, vol. 299, Prentice hall, 1995.
- [18] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, and M. Daković, "Instantaneous frequency in time–frequency analysis: Enhanced concepts and performance of estimation algorithms," *Digital Signal Processing*, vol. 35, pp. 1–13, 2014.
- [19] L. Stankovic, I. Djurovic, A. Ohsumi, and H. Ijima, "Instantaneous frequency estimation by using Wigner distribution and Viterbi algorithm.," in *ICASSP* (6), 2003, pp. 121–124.
- [20] C. Conru, I. Djurovic, C. Ioana, A. Quinquis, and L. Stankovic, "Time-frequency detection using Gabor filter bank and Viterbi based grouping algorithm," in *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.* IEEE, 2005, vol. 4, pp. iv–497.
- [21] I. Djurović, "Viterbi algorithm for chirp-rate and instantaneous frequency estimation," *Signal Processing*, vol. 91, no. 5, pp. 1308–1314, 2011.
- [22] H. Zhang, G. Bi, W. Yang, S. G. Razul, and C. M. S. See, "IF estimation of FM signals based on time-frequency image," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 326–343, 2015.
- [23] A. Moukadem, Z. Bouguila, D. O. Abdeslam, and A. Dieterlen, "A new optimized Stockwell transform applied on synthetic and real non-stationary signals," *Digital Signal Processing*, vol. 46, pp. 226–238, 2015.
- [24] I. Djurović, T. Thayaparan, and L. Stanković, "Adaptive local polynomial Fourier transform in isar," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 129–129, 2006.