AN IMPROVED SPECTROGRAM USING THE MULTIANGLE CENTERED DISCRETE FRACTIONAL FOURIER TRANSFORM

Juan G. Vargas-Rubio

Universidad Autónoma Metropolitana Azcapotzalco

México, DF 02200, MÉXICO

jgvr@correo.azc.uam.mx

Balu Santhanam

University of New Mexico

Albuquerque, NM 87131

bsanthan@ece.unm.edu

ABSTRACT

The spectrogram is a useful time-frequency analysis tool for non stationary signal analysis. This tool however, is based upon a multicomponent sinusoidal model over a signal analysis frame and is not suitable when for example the frequency content over the frame is chirping. Recently the centered version of the discrete Fractional Fourier Transform was shown to possess the capability to concentrate a linear chirp signal in a few transform coefficients. In this paper, we present a modified version of the spectrogram that incorporates the centered discrete Fractional Fourier transform and its multiangle version that instead decomposes the signal over the analysis frame into multiple chirp signals. Simulation results that study the efficiency of this improved spectrogram and its application to the analysis of harmonically related chirps and bat echolocation signals are presented.

1. INTRODUCTION

The spectrogram for discrete–time signals is based on the computation of the DFT over a windowed segment of the signal, and is a useful time-frequency analysis tool. This decomposition using the DFT however, inherently assumes a multicomponent sinusoidal signal model and that the frequency content of the signal is stationary over the analysis frame [1]. This however, may not be the case for example in signals such as bird chirps, bat echolocation signals, and human speech signals, where the fundamental frequency and its harmonics are chirping, i.e., non stationary.

Variations of the DFT, such as the *discrete Fractional Fourier transform* (DFRFT) and in particular the centered version of the DFRFT based on the Grunbaum tridiagonal commutor [4] have been shown to posses the capability to concentrate a chirp signal in a few transform coefficients [5]. The multiangle version of this CDFRFT (MA-CDFRFT) and the ensuing chirp rate versus frequency representation were shown to be useful time-frequency analysis tools for multicomponent chirp signals [6].

In this paper, we propose a modified version of the spectrogram that, instead of assuming a multicomponent sinusoidal model over the analysis frame, uses the CDFRFT and the MA-CDFRFT to decompose the signal frame into the superposition of harmonically related chirp signals. The notion of the slanted spectrum that is better suited towards the analysis of harmonically related chirps is introduced. Simulation results describing the application of this modified spectrogram to the analysis of synthetic chirps and to the analysis of natural bat echolocation signals is presented to illustrate the improvements over the conventional spectrogram.

2. THE CDFRFT AND THE MA-CDFRFT

The CDFRFT, as defined in [5], is the discrete fractional Fourier transform based on the Grünbaum commuting matrix [3, 4]. This transform has the ability to concentrate linear chirps in a few coefficients if we choose the angle α that corresponds to the chirp rate of the signal according to the empirical relation

$$\mathbf{c_r} = 2\frac{\tan(\alpha - \pi/2)}{N} + \sqrt{2}\frac{(\alpha - \pi/2)}{N}, \qquad (1)$$

where $\mathbf{c_r}$ is the chirp rate defined as the coefficient of the quadratic term of the phase, and α is the angle of the CDFRFT. Computing the CDFRFT for the set of angles $\alpha_r = \frac{2\pi r}{N}, r = 0, 1, \ldots, N - 1$, using the fast computation method developed in [5] we obtain the MA-CDFRFT. The
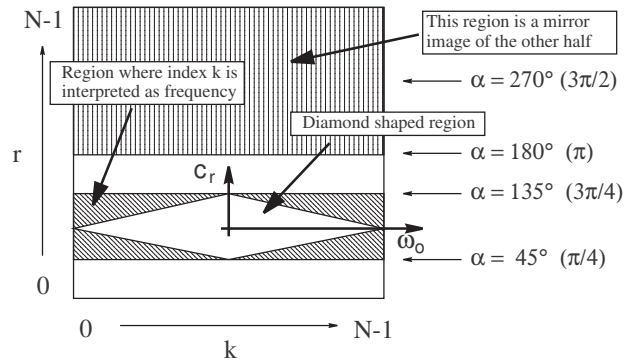

Figure 1: Map of $\mathbf{X}_k[r]$ that shows the region where the index k is interpreted as frequency.

MA-CDFRFT is a matrix, $\mathbf{X}_k[r]$, that contains a representation of the signal in which time and frequency share the same axis (index k), and depending on the value of α we interpret the index as time or frequency. The diamond shaped region in the $\mathbf{c_r}$–ω_0 space is the region where we have a good localization of the chirp rate and average frequency for a signal with constant chirp rate. Fig. 1 shows a

representation of the matrix that indicates where the region is located in matrix $\mathbf{X}_k[r]$.

When we compute the MA-CDFRFT of a complex sinusoidal signal we always obtain two peaks, one in the lower half and other in the upper half of $\mathbf{X}_k[r]$, and if the signal has a constant frequency, the peaks will be located exactly at the rows corresponding to angles $\alpha=90°$ and $\alpha=270°$ that correspond to the centered DFT (CDFT), and its inverse respectively. With a real signal we obtain four peaks due to the relevant symmetries. A signal with a chirp rate different than zero will produce peaks at angles different than $90°$ and $270°$, and they will be above or below those values depending on the sign of the chirp rate.

As an example, consider the signal

$$x[n] = e^{j(0.8n+0.004(n-127/2)^2)}, \quad n = 0, \ldots, 127.$$

For the quadratic term of the phase, we use $(n - 127/2)^2$ instead of n^2 in order to have an average frequency equal to the frequency of the linear term. The image of the absolute value of the MA-CDFRFT is given in Fig. 2(a), where we can observe the two peaks in the whole transform. Fig. 2(b) is a zoomed–in image to show the region where one of the peaks occurs. From this figure, we can see that the maximum in the lower part is located at $r=35$ and $k=80$. A signal with zero chirp rate would have the peak at $r=32$. This example depicts the capability of the MA-CDFRFT to concentrate a chirp in a few transform coefficients and the utility of the chirp rate/frequency representation.

3. APPLICATIONS TO HARMONICALLY RELATED CHIRPS

In nature, we observe many signals whose spectrum can be modelled as a fundamental frequency with harmonics, i.e., components whose frequency is a multiple of the fundamental frequency. If the fundamental frequency is constant, the DFT of the signal will display equally spaced peaks at the fundamental and harmonic frequencies. Many of the signals we encounter in nature display a fundamental frequency that is changing with time, i.e., they have some chirp rate. Common examples of these signals are bird chirps, bat echolocation signals, and speech. There are also artificial sounds, like machine vibrations and noises, that can be modelled in the same way.

A characteristic of these types of signals is that when the fundamental frequency has a given chirp rate, each harmonic has a chirp rate that is also a multiple of the chirp rate of the fundamental, that is, every harmonic has a different chirp rate. We can model a signal of this type with a sum of complex signals that is

$$x[n] = \sum_{k=1}^{M} e^{j(\omega_k n + c_{rk} n^2)}, \tag{2}$$

where the conditions for the components to be harmonically related are

$$\omega_k = k\,\omega_1, \quad c_{rk} = k\,c_{r1}, \quad k \geq 2. \tag{3}$$

It is easy to show that the conditions above produce harmonically related signals by obtaining the *instantaneous frequency* (IF) of each component.

When we apply the MA-CDFRFT to a signal with harmonically related components, we observe that the peaks corresponding to the different chirp components are aligned in a straight line. To illustrate this, we obtain the MA-CDFRFT of the signal

$$x[n] = \sum_{k=1}^{3} e^{j(0.5kn+0.001k(n-127/2)^2)}, \tag{4}$$

with $n = 0, \ldots, 127$. This signal consists of three components harmonically related according to our definition. The fundamental frequency component has an average frequency of 0.5 radians and a chirp rate of 0.001. The other two components have parameters twice and three times those values retaining a harmonic relation with the fundamental component for all values of n. Fig. 3(a) shows the image of the magnitude of the MA-CDFRFT for the signal. We can see that the three peaks are aligned along a straight line with a positive slope. If we observe the magnitude of the CDFT of the signal shown in Fig. 3(b) that corresponds to row 32 of the image (horizontal line), it is very difficult to determine that we have three harmonics, but if we take the values of the MA-CDFRFT along the line that goes through the peaks (slant line) to construct what we call a *slanted* spectrum, we obtain a much better representation of the three components that is shown in Fig. 3(c).

Although the empirical relation between the chirp rate and the angle α of Eq.(1) is not linear, the deviation from a straight line is relatively small, and for this reason we find the peaks of each component falling approximately on a straight line. The process of obtaining the *slanted* spectrum is analogous to the notion of focusing in optics, the slant line can be seen as the equivalent of the image plane where the best focus is obtained.

4. IMPROVED SPECTROGRAM

The conventional spectrogram used for analysis of human speech signals computes the *short–time Fourier transform* (STFT) over small windowed segments of the signal, where the frequency content is assumed stationary over the frame. The assumption of stationarity and the DFT based decomposition, further allows the signal segment to be modelled as a multicomponent sinusoidal signal [1]. To construct a spectrogram with the *slanted* spectrum we require the values of the MA-CDFRFT along the line where the peaks are aligned. In order to do this we need to compute the equation of the line. Assuming that the signal has a dominant component, that is, a component that has a larger amplitude than the others, the equation of the line can be determined by using the position of the largest peak in the MA-CDFRFT. The position of the largest peak, $[k_p, r_p]$, and the position of the point at zero frequency and zero chirp rate, $[(N-1)/2, N/4]$, define the equation of a line that provides the value of the index r of the slant line, r_s in terms of index k. The equation is

$$r_s = \left(\frac{r_p - \frac{N}{4}}{k_p - \frac{N-1}{2}} \right) k + \frac{N}{4} - \left(\frac{r_p - \frac{N}{4}}{k_p - \frac{N-1}{2}} \right) \left(\frac{N-1}{2} \right) \tag{5}$$

With the equation of the line we can select the values of the MA-CDFRFT to obtain the *slanted* spectrum. Repeating
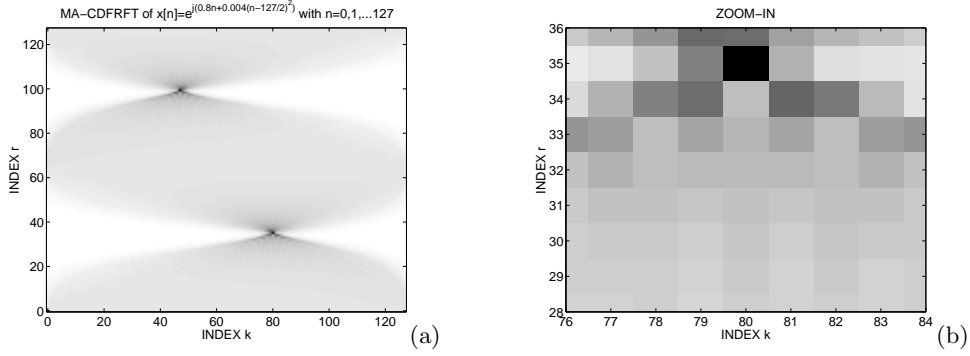
Figure 2: Images that shows the magnitude of the MA-CDFRFT applied to the signal $x[n] = e^{j(0.8n+0.004(n-127/2)^2)}$ with $n = 0, \ldots, 127$. (a) is the whole area of the transform, and (b) is a zoom–in to the region that contains one of the peaks and shows that the peak is located at the coordinate $r=35$ and $k=80$. Note that $r=32$ corresponds to the CDFT.
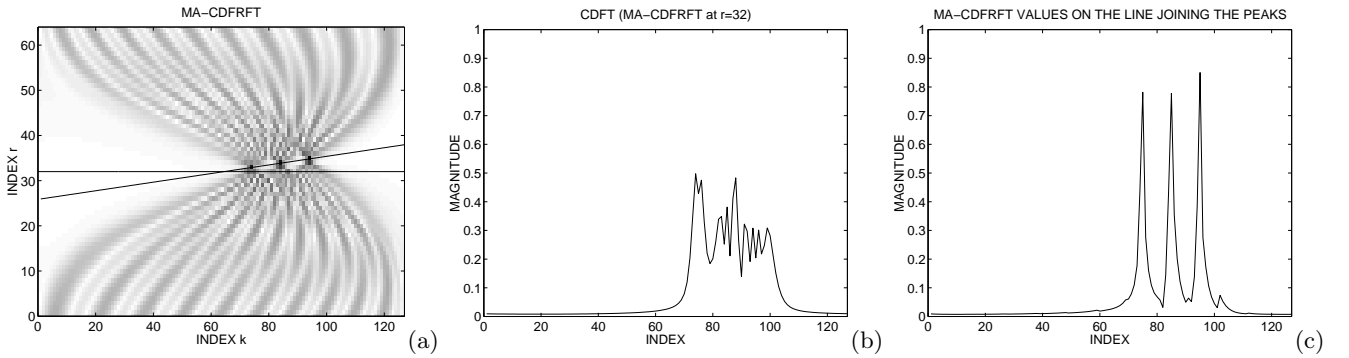

Figure 3: (a) Magnitude of the MA-CDFRFT applied to the signal $x[n] = \sum_{k=1}^{3} e^{j(0.5kn+0.001k(n-127/2)^2)}$, (b) is the CDFT of the signal, and (c) is the plot of the values obtained along the line defined by the peaks shown in (a).

this process of finding the largest peak and computing the equation of the line over each frame at different positions in time we can construct a spectrogram that is sharper than the one based on the DFT.

As a first example, we use the same signal parameters defined in Eq. (4), over the duration, $0 \leq n \leq 499$. For this signal, we compute the regular spectrogram using a CDFT of 128 points (time step of one sample, overlap of 127 samples). We also compute a spectrogram based on the *slanted* spectrum using the same parameters and utilizing the search for the largest peak to define the line. In both cases, we use a 128 point Hanning window. The resulting spectrograms are shown in Fig. 4. We observe that the spectrogram obtained using the *slanted* spectrum produces a sharper representation of the different components. We also observe that the third component loses its sharpness as it approaches the largest frequency. This effect is attributable to the corresponding peak falling well outside the diamond shaped region, the nonlinearity of the relation between the chirp rate and the angle begins to dominate when this happens, and the peak of that component does not fall close to the straight line. The method of detecting the peak and producing the equation of the line through

the peaks works well if the components of the signal fall within or close to the diamond shaped region. There are some problems associated with the discrete nature of the position of the peaks that causes slight jumps in the slope of the slant line that manifest as noise in the spectrogram, but from the figure we can observe that the representation is sharper than with the CDFT. As a final example, we show the analysis of a real world signal, the bat echolocation signal from Rice University[1] that is a signal frequently used to show the abilities of TF representations. The components of this signal are not linear chirps, yet the *slanted* spectrum gives sharper peaks than the DFT. Fig. 5 shows this signal and the magnitude of its CDFT.

We compute a regular spectrogram with the CDFT, and a spectrogram using the values of the MA-CDFRFT along the line defined by the largest peak (the slope of the line changes at every sample). In both cases, we use a 128 point transform size, a 127 sample overlap, and a 128 point Hanning window. The results are shown in Fig. 5. We observe that the spectrogram computed using the MA-CDFRFT is
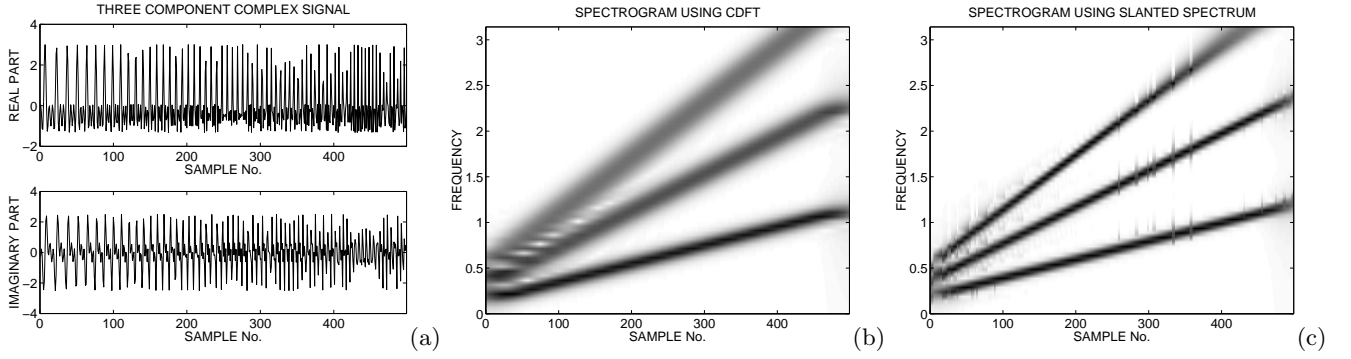
Figure 4: Spectrograms of the complex chirp signal shown in (a), The spectrogram in (b) is done using the CDFT. The spectrogram in (c) uses the *slanted* spectrum.
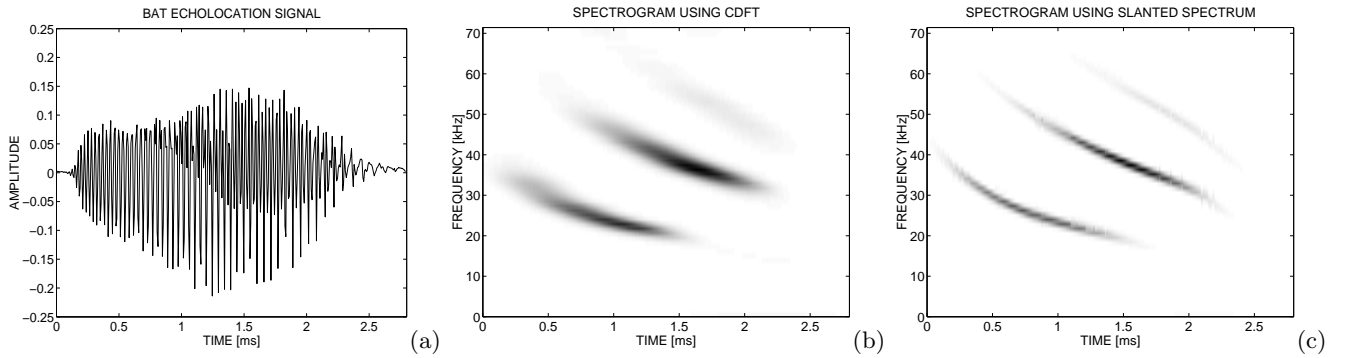

Figure 5: Spectrograms of the bat echolocation signal shown in (a). The spectrogram in (b) uses the CDFT, and the improved spectrogram in (c) uses the *slanted* spectrum from the MA-CDFRFT.

sharper as in the previous example.

5. CONCLUSION

In this paper, we have presented a modified version of the spectrogram based on the multiangle version of the centered DFRFT. We introduced the notion of the slanted spectrum, utilizing the CDFRFT, that is more appropriate for harmonically related chirp signals. This slanted spectrum in conjunction with a multicomponent chirp model over the analysis frame was used to define this modified spectrogram which produces sharper features in the time-frequency plane in comparison to the regular spectrogram that assumes a multicomponent sinusoidal model over the same analysis frame. This modified spectrogram was then applied to multicomponent chirps with harmonically related components and a real bat echolocation signal to quantify the improvement in resolution.

REFERENCES

[1] R. J. McAulay and T. F. Quatieri, "Speech Analysis and Synthesis based on a Sinusoidal Representation," *IEEE Trans. ASSP*, Vol. 34, No. 4, pp. 744-754, 1986.

[2] D. H. Mugler and S. Clary, "Discrete Hermite Functions and The Fractional Fourier Transform," *in Proc. Int. Conf. Sampl. Theo. and Appl. Orlando Fl*, pp. 303-308, 2001

[3] S. Clary and D. H. Mugler, "Shifted Fourier Matrices and Their Tridiagonal Commutors," *SIAM Jour. Matr. Anal. & Appl.*, Vol. 24, No. 3, pp. 809-821, 2003.

[4] B. Santhanam and J. G. Vargas-Rubio, "On the Grunbaum Commutor Based Discrete Fractional Fourier Transform," *Proc. of ICASSP–04*, Vol. II, pp. 641-644, May, 2004.

[5] J. G. Vargas-Rubio and B. Santhanam, "The Centered Discrete Fractional Fourier Transform and Linear Chirp Signals," *Proc. of 11th IEEE DSP Workshop*, pp. 163 - 167, 2004.

[6] J. G. Vargas-Rubio and B. Santhanam, "On the Multiangle Centered Discrete Fractional Fourier Transform," *to appear, IEEE Sig. Process. Letters.*, Nov 2004.

[7] J. G. Vargas-Rubio, "The Centered Discrete Fractional Fourier Transform, properties, computation, and application to linear chirp signals," Ph.D Dissertation, The University of New Mexico, 2004