IMPROVED SPECTROGRAMS USING THE DISCRETE FRACTIONAL FOURIER TRANSFORM

Oktay Agcaoglu, Balu Santhanam, and Majeed Hayat

Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, New Mexico 87131
oktay, bsanthan@unm.edu, hayat@ece.unm.edu

ABSTRACT

The conventional spectrogram is a commonly employed, time-frequency tool for stationary and sinusoidal signal analysis. However, it is unsuitable for general non-stationary signal analysis [1]. In recent work [2], a slanted spectrogram that is based on the discrete Fractional Fourier transform was proposed for multicomponent chirp analysis, when the components are harmonically related. In this paper, we extend the slanted spectrogram framework to non-harmonic chirp components using both piece-wise linear and polynomial fitted methods. Simulation results on synthetic chirps, SAR-related chirps, and natural signals such as the bat echolocation and bird song signals, indicate that these generalized slanted spectrograms provide sharper features when the chirps are not harmonically related.

***Index Terms*—** Spectrogram, DFrFT, slanted spectrogram, non-harmonic chirps.

1. INTRODUCTION

The conventional spectrogram is based on the discrete Fourier transform, which has sinusoids as basis functions, therefore it is unsuitable for signals having non-stationary component. The discrete Fractional Fourier transform (DFrFT) shows promise for the analysis of signals frequency chirping (i.e. linearly changing) [3], such as radar, speech, bat echolation and bird songs.

The slanted spectrogram was proposed as a DFrFT based spectrogram in [2]. This spectrogram provides sharper features in comparison to the conventional spectrogram for non-stationary signals with harmonically related chirps. The conventional spectrogram assumes a multicomponent sinusoidal model over the analysis frame while the slanted spectrogram replaces that sinusoidal assumption with a multicomponent harmonically related chirp model. However, most real data contains non-harmonically related chirps.

In this paper, two DFrFT based spectrograms are proposed using polynomial fitted and piece-wise linear approaches, which provide sharper features than the slanted spectrogram. Simulation results for a non-harmonically related complex chirp signal, bat echolation, bird song and radar data are presented to quantify the improvement of the proposed spectrograms.

2. DISCRETE FRACTIONAL FOURIER TRANSFORM

The Fractional Fourier transform (FrFT) is a generalization of the conventional Fourier transform. If time and frequency are considered as orthogonal axes, then the Fourier transform is a 90° rotation in this plane, while the FrFT can generate signal representations at any angle of rotation in the plane [4]. The eigenfunctions of the FrFT are Hermite-Gauss functions, which result in a kernel composed of chirps. When The FrFT is applied on a chirp signal, an impulse in the chirp rate-frequency plane is produced [5]. The coordinates of this impulse gives us the center frequency and the chirp rate of the signal.

Discrete versions of the Fractional Fourier transform (DFrFT) have been developed by several researchers and the general form of the transform is given by [3]:

$$\mathbf{X}_\alpha = \mathbf{W}^{\frac{2\alpha}{\pi}}\mathbf{x} = \mathbf{V}\Lambda^{\frac{2\alpha}{\pi}}\mathbf{V}^H\mathbf{x} \tag{1}$$

where $\mathbf{W}$ is a DFT matrix, $\mathbf{V}$ is a matrix of DFT eigenvectors and $\boldsymbol{\Lambda}$ is a diagonal matrix of DFT eigenvalues. The DFT matrix $\mathbf{W}$ has repeated eigenvalues, so there is no unique set of eigenvectors for the solution of this equation. In this paper, we use quantum mechanics in finite dimensions (QMFD) basis because it provides the largest valid interval in the sense of connectivity and adjacency [6]. The DFrFT is computed using the fast algorithm described in [7], which requires the transform size to be multiples of 4.

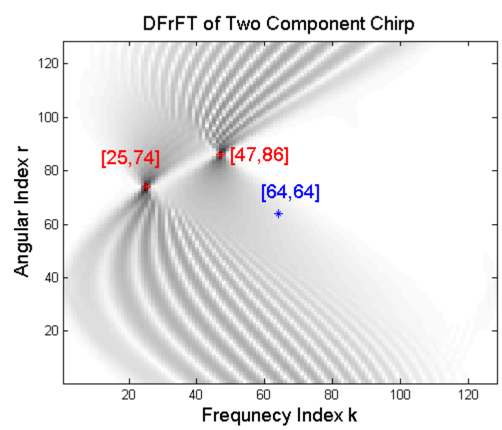

Fig. 1. DFrFT magnitude of a two component chirp signal whose constituent components are non-harmonically related, the peak locations indicate the chirp rates and the carrier frequencies. The peaks do not lie on the same line with the zero chirp-rate, zero frequency coordinates (shown by blue point) because of being non-harmonically related

The DFRFT has the ability to transform a linear chirp into a peak in the chirp-rate, frequency plane. Signal model for linear chirps shown at equation 2 where c_{rk} is chirp rates, w_k is the carrier frequencies and M is the number of chirps.

$$x[n] = \sum_{k=1}^{M} \exp\left(j\left[w_k + c_{rk}\left(n - \frac{N-1}{2}\right)^2\right]\right) \tag{2}$$

This work was supported by the U.S. Department of Energy (Award No. DE-FG52-08NA28782) and by the National Science Foundation (Award No. IIS-0813747).

The conditions to be harmonically related are:

$$w_k = k\,w_1, \quad c_{rk} = k\,c_1, \quad k \geq 2 \tag{3}$$

When these conditions are satisfied, it is possible to construct a line that passes through all the peaks and the zero chirp-rate, zero frequency coordinates of the DFrFT spectrums. However, this is not the case for most of the time. When a signal has non-harmonically related chirps, the peaks do not lie on a single line that passes through the zero chirp-rate, zero frequency coordinates. In order to better understand this concept, we look at a simple example with a chirp signal comprised of two non-harmonic components with chirp rates $c_{1,2} = [0.015, 0.006]$ and carrier frequencies $w_{1,2} = [1, 2]$ respectively.

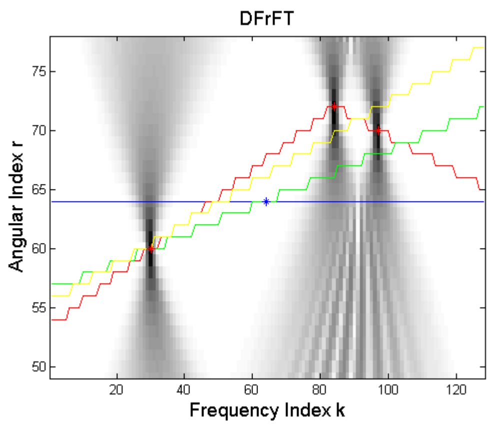

Fig. 2. DFrFT magnitude for the three component chirp signal whose constituent components are non-harmonically related. The piece-wise linear spectrogram (red) passes through all the peaks, the polynomial spectrogram (yellow) passes on one peak and very close to the other two, the slanted spectrogram (green) benefits only one peak and ignores the other two, the conventional spectrogram (blue) ignores all the peaks. (Zoomed in angular axis)

The motivation behind the proposed modified spectrograms is that for a general multicomponent chirp signal, the line connecting the zero chirp-rate, zero frequency coordinates to the largest peak may not pass through all the spectral peaks. Therefore, a DFT based spectrogram extracts spectrogram lines for each frame that do not pass through all the peaks, thereby producing more shallow and blurry results. The slanted spectrogram attains the benefit of passing all peaks only if the chirp are harmonically related so they lie on the same line. The proposed spectrograms, the polynomial fitted and the piece-wise linear methods take into accounts all the peaks and attempt to benefit all the peaks to get sharper spectrogram without requirement of harmonically related.

3. IMPROVED SPECTROGRAMS

3.1. Rewiew of Slanted Spectrogram

The slanted spectrogram uses the zero chirp-rate, zero frequency coordinates and the highest peak coordinates on the DFrFT to calculate a line equation. The points on this line populate the spectrogram for that frame. The coordinates of the largest peak, [kp, rp], and the zero chirp-rate, zero frequency coordinates, $[(N-1)/2, N/4]$, 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.

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

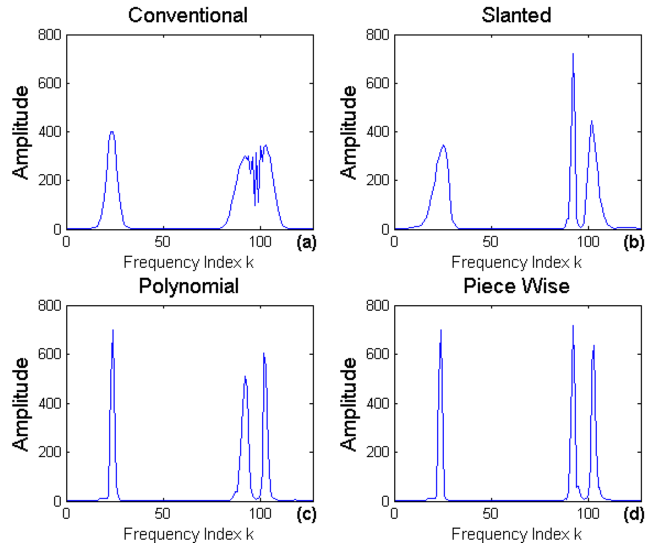

Fig. 3. An example of extracted spectrogram for a frame of the non-harmonically related complex chirp signal. (a) is the conventional, (b) is the slanted. The polynomial fitted (c) and the piece-wise linear (d) extract peaks higher and sharper in comparison to (a) and (b).

By applying this procedure over each frame and finding the largest peak and computing the equation of the line, a spectrogram with sharper features than the conventional spectrogram is constructed [2]. This method works well if the chirp peaks lie on a single line, i.e. if the chirps are harmonically related. However, this is not always the case. When chirps rates are not harmonically related, the slanted spectrogram fails to produce sharp features because it ignores certain peaks.

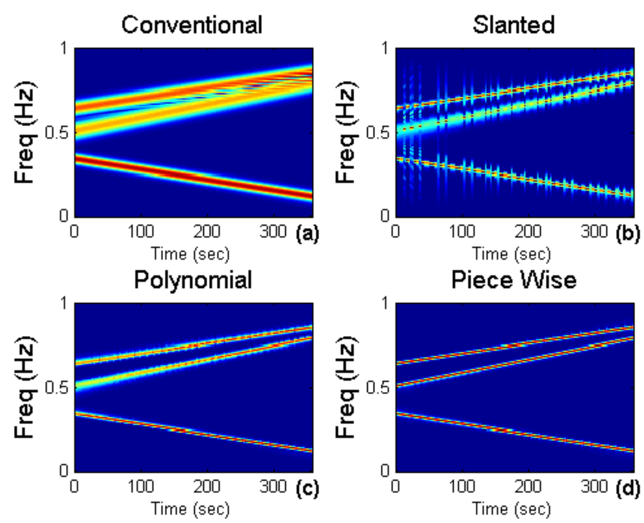

Fig. 4. Spectrograms of the non-harmonically related complex chirp signal. (a) is the conventional spectrograms and (b) is the slanted spectrogram. The polynomial fitted (c) and the piece-wise linear (d) provide sharper features in comparison to (a) and (b).

3.2. Piece-Wise Linear

As we mentioned before, chirps of the analysis signal manifest as peaks in the DFrFT. When we have prior knowledge of the number of chirps in the signal, we can use the location of all the peaks while calculating the spectrogram rather than using only the largest peak.

This enables us to construct a sharp spectrogram even if the peaks do not lie on the same line. The piece-wise linear spectrogram uses the coordinates of the chirp peaks to calculate line segments. Let $[k_p^i, r_p^i]$ be the coordinates of peaks, M is the number of chirps and N is size of DFrFT, piece-wise linear spectrogram for a frame, r_{pw}, calculated as:

$$
\begin{aligned}
r_{pw}^1(k) &= \left(\frac{r_p^2 - r_p^1}{k_p^2 - k_p^1}\right)(k - k_p^1) + r_p^1 \qquad 0 \le k < k_p^2 \\
r_{pw}^2(k) &= \left(\frac{r_p^3 - r_p^2}{k_p^3 - k_p^2}\right)(k - k_p^2) + r_p^2 \qquad k_p^2 \le k < k_p^3 \\
r_{pw}^i(k) &= \left(\frac{r_p^{i+1} - r_p^i}{k_p^{i+1} - k_p^i}\right)(k - k_p^i) + r_p^i \qquad k_p^i \le k < k_p^{i+1} \\
&\;\;. \\
&\;\;. \\
&\;\;. \\
r_{pw}^{M-1}(k) &= \left(\frac{r_p^M - r_p^{M-1}}{k_p^M - k_p^{M-1}}\right)(k - k_p^{M-1}) + r_p^{M-1} \quad k_p^M \le k < N
\end{aligned}
\tag{5}
$$

Applying this method to each frame, we construct the piece-wise linear spectrogram that is sharper than the slanted spectrogram. For real signals, the DFrFT spectrum is mirrored, therefore we can only have an even number of peaks, while for a complex signal we can have either an even or odd number of peaks. In the special case, a complex signal having only one chirp component, the line is calculated using the the zero chirp-rate, zero frequency coordinates, which is exactly the same as the slanted spectrogram.

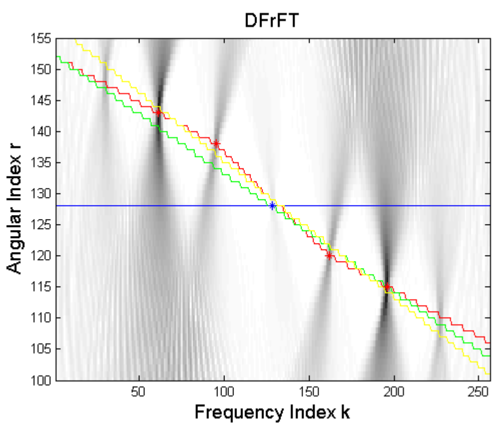

Fig. 5. DFrFT magnitude of a frame of the bat echolation signal, blue is the conventional, green is the slanted, yellow is the polynomial fitted and red is the piece-wise linear spectrogram. The spectrogram lines are close to each other because the chirps are almost harmonically related. (Zoomed in angular axis)

3.3. Polynomial Fitted

For the polynomial fitted spectrogram, we calculate nth order polynomial that fits best to the points of the chirp peaks in least square sense. The polynomial that is found gives us the polynomial fitted spectrogram segment for that frame. Applying the same procedure for each frame for different time instances, the polynomial fitted spectrogram is constructed. The order of polynomial is a free variable and the user can choose the order depending on the number of chirps and the characteristics of the chirps.

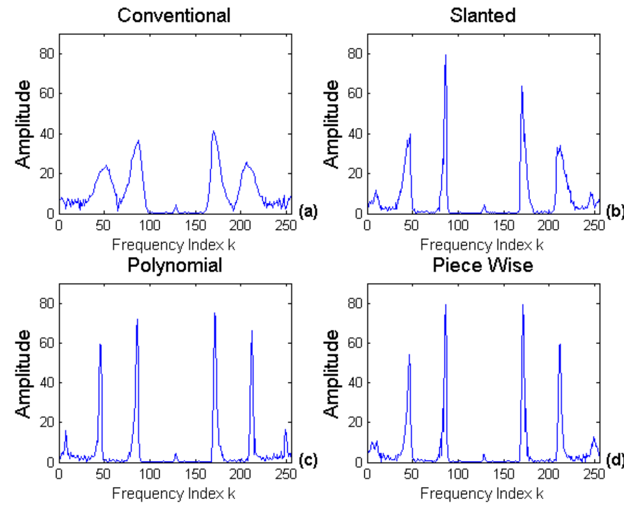

Fig. 6. An example of extracted spectrogram for a frame of the bat echolation signal. Even though chirps are almost harmonically related, still the peaks are higher and sharper for the polynomial fitted (c) and the piece-wise linear (d) in comparison to the slanted (b) and the conventional (a) spectrograms.

4. SIMULATION EXAMPLES

In nature, there are many signals whose frequencies are changing with time, e.g. chirping. Some common examples of these non-stationary signals are bird chirps, bat echolocation signals, dolphin sounds and speech. There are also artificial signals, like machine vibrations, radar waveforms, and continuous phase modulation waveforms.

4.1. Non-Harmonic Complex Chirp Signal

As a first simulation example, we generate a three components non-harmonically related complex chirp signal according to the signal model at Eq. 2. For carrier frequencies $w_1 = 0.05$, $w_2 = 1$ and $w_3 = 0.9$ and chirp rates $c_{r1} = 0.0039$, $c_{r2} = 0.003$ and $c_{r3} = 0.0029$, respectively and for $0 \le n \le 355$.

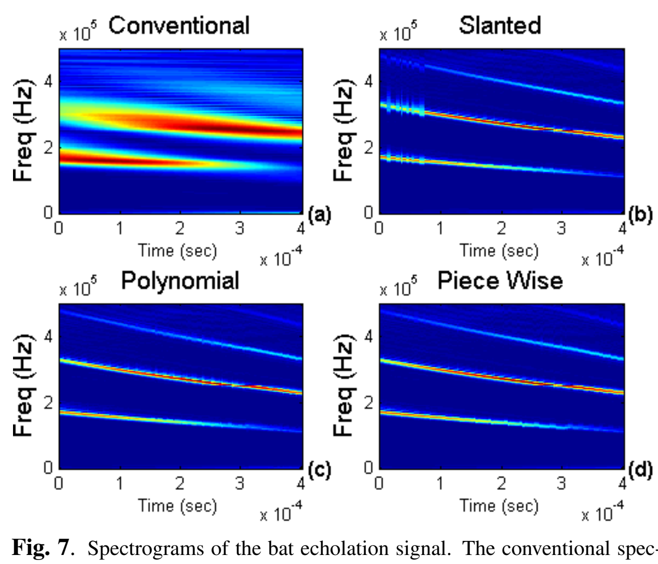

Fig. 7. Spectrograms of the bat echolation signal. The conventional spectrogram (a) is very blurry, the slanted spectrogram (b) is sharper than the conventional spectrogram but suffers from local blurring. The polynomial fitted (c) and the piece-wise linear (d) provide sharper features and clearer spectrograms.

For this signal, we calculated the conventional spectrogram, the slanted spectrogram and the proposed polynomial fitted and piece-wise linear spectrogram using 128 point DFrFT. For all cases, a 128 point Hanning window and an overlap of 127 is used. The order of the polynomial is 1 for the simulation.

Figure 2 shows the calculated DFrFT for an frame. Because the chirps are non-harmonic, the spectrum results in a series of peaks that do not lie on a single line. The chirp peaks are marked with red dots, the zero chirp-rate, zero frequency coordinates are shown with a blue dot, and lines show the extracted spectrograms for this frame.

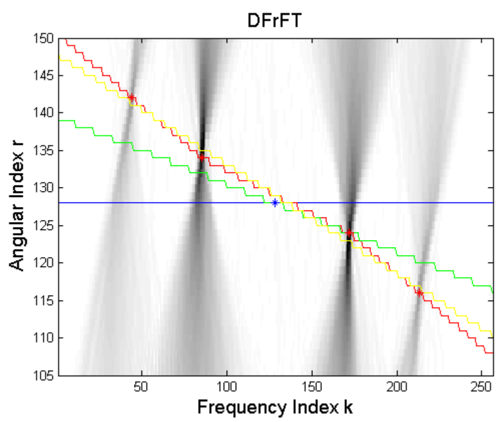

Fig. 8. DFrFT magnitude of a frame of the loon bird song data, blue is the conventional, green is the slanted, yellow is the polynomial fitted and red is the piece-wise linear spectrogram. (Zoomed in angular axis)

Figure 3 shows an example frame for the extracted features for the conventional, the slanted and the proposed spectrograms. The conventional spectrogram fails to separate the second and the third peaks. The slanted spectrogram produces a single sharp peak, while the polynomial fitted and the piece-wise linear were able to provide sharper and higher peaks for all three components. This is due to the fact that slanted spectrogram only accounts for the highest peak whereas the polynomial fitted and the piece-wise linear take into account all of the peaks.

Figure 4 shows the conventional and the modified spectrograms for a multicomponent complex chirp signal whose constituent components are non-harmonically related. The polynomial fitted and the piece-wise linear spectrograms have significantly sharper features in comparison to the conventional spectrogram. The slanted spectrogram has a lot of irregularities and glitches, which is to be expected because the chirps are not harmonically related, so the line of slanted method does not pass through all the chirp peaks of the DFrFT. Overall, we can say the piece-wise linear method works best for this signal.

4.2. Bat Echolation

The bat echolation signal from the Rice University database[1] [8] is a real world signal used frequently to show capabilities of time frequency transforms. Bats find their way by sending these signals and receiving them back. It was shown in [2] that the application of the slanted spectrogram to this bat echolocation signal produces sharper features than the conventional spectrogram.

[1]The authors wish to thank Curtis Condon, Ken White, and Al Feng of the Beckman Institute of the University of Illinois for the bat data and for permission to use it in this work

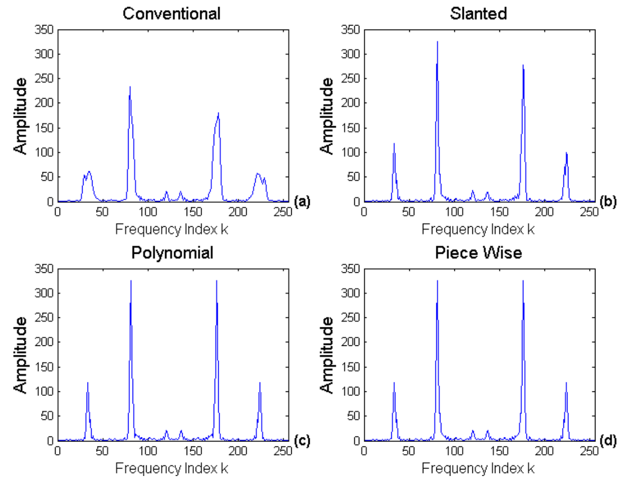

Fig. 9. An example of extracted spectrogram for a frame of the loon bird song data. The modified spectrograms (b,c,d) have higher and sharper peaks in comparison to the conventional spectrogram (a). The chirps are almost harmonically related, so the improvement from the slanted (b) to the polynomial (c) and the piece-wise (d) is not significant.

We calculated the conventional, the slanted, and the proposed polynomial fitted and piece-wise linear spectrograms of the bat echolation signal using 256 point DFrFT. For all cases, a 256 point Hanning window is used and number of overlap is 255. The order of the polynomial is 1 for the simulation.

Figure 6 depicts an example of extracted spectrogram slices. The associated chirp rates are almost harmonically related, therefore the slanted spectrogram provides sharper peaks in comparison to the conventional spectrogram. However, only one peak of the slanted spectrogram is as sharp as the polynomial fitted and piece-wise polynomial peaks. This is due to the fact that it utilizes only the highest peak and ignores the other peaks.

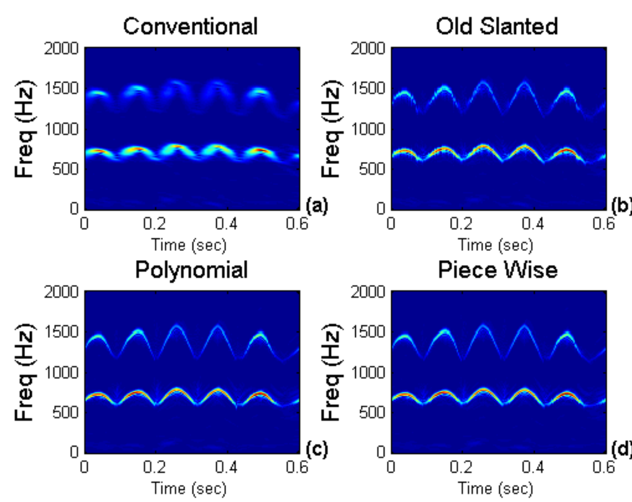

Fig. 10. Spectrograms of the loon bird song data. The conventional spectrogram (a) is very blurry, the slanted spectrogram (b) is sharper than the conventional spectrogram but has blurring over several time instances . The polynomial fitted (c) and the piece-wise linear (d) provide sharper features and clearer spectrograms.

The calculated spectrograms are shown at Fig. 7. The DFrFT based spectrograms are significantly sharper than the conventional spectrogram. The proposed methods produce clearer spectrograms in comparison to the slanted spectrogram. The local blurring on the slanted spectrogram is due to the fact that over several time instances, the bat echolation signal fails to satisfy harmonically related chirp assumption. The polynomial fitted and the piece-wise linear results are quite similar for this signal.

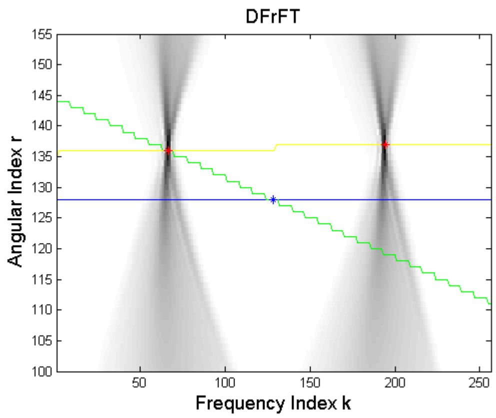

Fig. 11. DFrFT magnitude of a frame of the SAR signal, blue is the conventional, green is the slanted, yellow is the polynomial fitted and red is the piece-wise linear spectrogram. The line from the slanted spectrogram passes through the zero chirp-rate, zero frequency coordinates and one peak, but completely misses the second peak. (Zoomed in angular axis)

4.3. Bird Song Data

Bird songs are examples of natural signals with chirp content. We applied the spectrograms to a bird song of loon from the Stony Brook University database[2] [9]. We calculated the conventional, the slanted, the polynomial fitted and the piece-wise linear spectrograms of the bird song of loon using 256 point DFrFT. For all cases, a 256 point Hanning window is used and number of overlap is 255. The order of the polynomial is 1 for the simulation.

When the chirp components are exactly harmonically related, theoretically the slanted, the polynomial fitted with order 1 and the piece-wise linear would give the same output, due to the fact that all the peaks lie on the same line. The constituent chirps of the loon bird song data are almost harmonically related, therefore we expect the modified spectrograms to give similar results among themselves and a sharper spectrogram in comparison to the conventional spectrogram. Figure 9 shows an example of the extracted slices for all spectrograms. The slanted spectrogram produces more or less the same sharp features and spectral peaks as the proposed methods. This is due to the almost harmonically related structure of the chirps.

The calculated spectrograms are shown on Fig. 10. The DFrFT based spectrograms produces sharper features in comparison to the conventional spectrogram. The piece-wise, the polynomial fitted and the slanted spectrogram have similar performance. However, the slanted spectrogram has some local blurring because the chirps of the loon bird song data fail to satisfy to be harmonically related for some time instances and this manifests as glithches in the slanted spectrogram.

[2]The authors wish to thank Tony Phillips of Stony Brook Mathematics Department and the Institute for Mathematical Sciences for the loon bird song data and for permission to use it in this work

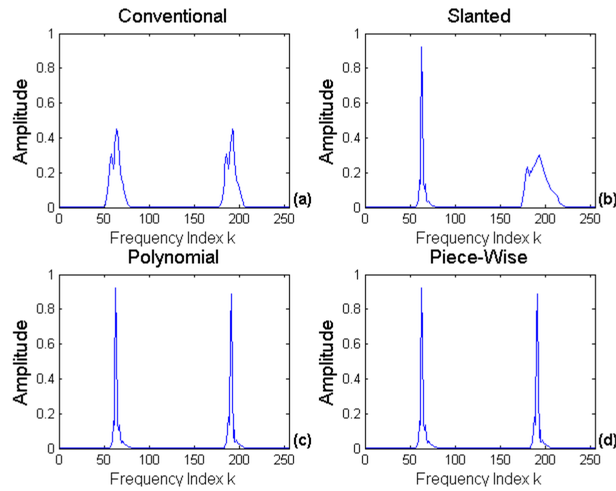

Fig. 12. An example of extracted spectrogram for a frame of the SAR signal. The slanted spectrograms (b) has only one higher and sharper peak in comparison to the conventional spectrogram (a). The polynomial fitted (c) and the piece-wise linear (d) provide higher and sharper peaks for the other two peaks.

4.4. Synthetic Aperture Radar

The last simulation example is a synthetic aperture radar signal. Recently, it has been shown at [10] that the DFrFT can be a useful tool to detect vibrations of an object via SAR. When ground targets have small vibrations, they introduces phase modulation in returned signals of the SAR. After applying standard pre-processing of the return signals, these small vibrations of the ground targets show themselves on the SAR signal as the quasi-instantaneous chirp rates [10].

We use a SAR signal containing two vibrating objects with amplitudes of 1 cm and 2 mm at vibrating frequencies 2 Hz and 4 Hz, respectively with system parameters described by [10] for the noise free case. Even though the initial sampling frequency is 3.216 MHz, a downsampling is applied on the data at preprocessing, the signal that modified spectrograms are applied to has a sampling rate of 40 KHz. Again, we refer to the original work [10] for specifics on the DFRFT technique and further inquiry about this application.

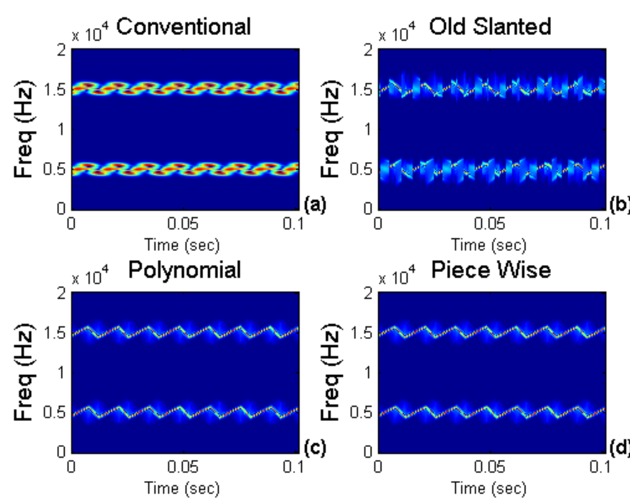

Fig. 13. Spectrograms of the SAR signal. The conventional spectrogram (a) and the slanted spectrogram (b) are very blurry. The polynomial fitted (c) and the piece-wise linear (d) provide sharper features and clearer spectrograms.

We calculated the spectrograms of the SAR signal using 256 point DFrFT. For all cases, a 256 point Hanning window is used and number of overlap is 255. The order of the polynomial is 1 for the simulation.

The effectiveness of the modified spectrograms are described by the simulation results. The radar signal is a complex signal with two chirp rates. This is a special condition where the polynomial with order 1 and the piece-wise spectrogram method produce the same outputs, due to the fact that there is only one line passing through two distinct points. Therefore, in this special case, the polynomial fitted and the piece-wise linear gives the same results.

5. CONCLUSION

In this paper, we have presented two generalized spectrograms based on polynomial and piece-wise linear fitting of the peaks of the discrete Fractional Fourier transform. The proposed spectrograms are able to accommodate all the peaks in the DFrFT spectrum, while the slanted spectrogram incorporated just the largest peak. This is appropriate when the components are harmonically related. However, when the chirps are non-harmonically related, the slanted spectrogram suppresses the other peaks and produces local blurring. The proposed modified spectrograms accommodate all the spectral peaks and do not suffer from the blurring problem, thereby producing sharper features. These generalized fractional spectrograms were applied to a multicomponent non-harmonically related complex chirp signal, a real bat echolation signal, a real bird song, and synthetic aperture radar data with vibrating targets to quantify the improvement.

6. 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, Aug. 1986.

[2] J. G. Vargas-Rubio and B. Santhanam, “An Improved Spectrogram Using The Multiangle Centered Discrete Fractional Fourier Transform,” *Proc. of ICASSP 2005*, pp. 505-508.

[3] B. Santhanam and J. H. McClellan, “The discrete rotational Fourier transform,” *Sig. Process.*, Vol. 44, No. 4, pp. 994-998, Apr., 1996.

[4] L. B. Almeida, “An introduction to the angular Fourier transform,” *Proc.of ICASSP–1993*, Vol. 3, pp. 257-260, Apr., 1993.

[5] B. Santhanam and T. S. Santhanam, ‘On discrete Gauss-Hermite functions and eigenvectors of the discrete Fourier transform,” *Sig. Process.*, Vol. 88, No. 11, pp. 2738-2746, 2008.

[6] D. J. Peacock and B. Santhanam, “Multicomponent subspace chirp parameter estimation using discrete fractional Fourier analysis,” *Proc. of IASTED–2011*, pp. 326-333, Dec., 2011.

[7] J. G. Vargas-Rubio and B. Santhanam, “On the multiangle centered discrete fractional Fourier transform,” *IEEE Sig. Process. Lett.*, Vol. 12, No. 4, pp. 273-276, 2005.

[8] C. Condon, K. White, and A. Feng, (2013, Apr), Bat Echolocation Chirp. Rice University. [Online]. Available: http://www.ece.rice.edu/dsp/software/bat.shtml.

[9] T. Phillips, (2013, Apr), Songs and calls of some New York State birds. Stony Brook University. [Online]. Available: http://www.math.sunysb.edu/ tony/birds/loons.htm.

[10] Q. Wang, M. Pepin, R. J. Beach, R. Dunkel, T. Atwood, B. Santhanam, W. Gerstle, A. W. Doerry and M. M. Hayat, “SAR-Based Vibration Estimation Using the Discrete Fractional Fourier Transform,” *IEEE Trans. on Geosc. and Rem. Sens.*, Vol. 50, No. 10, pp. 4145-4156, Oct., 2012.