NFFT: Algorithm for irregular sampling

Akshay Gulati and Robert J. Ferguson

Summary

The nonuniform discrete Fourier transform (NDFT), used in many processing schemes, can be computed using a fast algorithm known as the non uniform fast Fourier transform (NFFT). The NFFT is not a new algorithm, but it is an approximation scheme that can be use to calculate an approximate spectrum. In one dimension, computational complexity of the NFFT is O(N log N) which is a dramatic improvement from the $O(N^2)$ complexity of the NDFT. This algorithm can be easily extended to higher dimensions. The approximate spectrum is calculated using a simple algorithm scheme which involves convolution of an irregularly sampled signal with a truncated Gaussian in the spatial domain. A new empirical expression based on numerical experiment for the analytical Gaussian width is proposed. Synthetic data examples, some with analytical solutions, demonstrate the utility and validity of this approach. The approximate spectrum obtained can be use further in a reconstruction algorithm. This algorithm removes the bottle neck from forward process by replacing NDFT with NFFT in many conventional processing algorithms.

Introduction

The problem of analyzing a signal P (t_j) having irregularly spaced measurements is common in geophysics (Ferguson, 2006). Faulty equipment, errors in positioning, obstacles, and noise sources can be reasons for irregularly spaced measurements. Several approaches can be found which are based on some kind of interpolation techniques, but most of these approaches do not handle the data optimally. Transformation methods which include Radon transformations have been used to handle the problem of irregular sampling (Clarebout,1992).In the parabolic Radon transform, two CMP gathers are combined to improve offset sampling, and thus differences between mid point positions are ignored. Similarly, hyperbolic and linear Radon transforms (Clarebout,1992) as well as the parabolic Radon transform are suitable for estimating frequencies at irregular nodes, but they suffer aliasing problems due to sparse sampling. Prediction error filtering is another method which is used to interpolate missing traces, but it works only on regularly sampled data, and it only corrects for aliasing. if sampling is irregular, the result will be erroneous. Ronen et al (1991) suggests a method for DMO stacks which can handle irregular sampling, but it is still not efficient for very large seismic data volumes. Approximate regularization/datuming (Ferguson, 2006) allows extrapolation of data recorded on an irregular grid onto a regular grid, but it requires a velocity model.

My aim in this paper is to solve the forward case with a faster algorithm, as done by Duijndam and Schonewille (1999) and Greengard and Lee (2004). I will begin with the theory behind the algorithm structure for NFFT. After reviewing the basic algorithms, improvements are suggested in the NFFT algorithm. Then, using synthetic examples to validate the algorithm and I show that the NFFT is approximately 100 times faster than the NDFT, and that the approximate spectrum obtained using NFFT gives a good approximation to the original spectrum.

Theory

The NFFT algorithm can be numerically expressed in following steps: convolution, FFT, and deconvolution. Two parameters are significant in our algorithm, one is numerical width *q* of the truncated filter, and other is analytical width *b* for Gaussian filter. Convolution with the short Gaussian filter *g(x)* is carried out to make the signal approximately band-limited according to

$$p_g(m) = g(x) * p(x) \qquad (1)$$

where $p_g(m)$ is the result is the result of spatial convolution. Equation 1 can be written as multiplication in the Fourier domain as

$$Pg(m) = G(m) * P(m) \qquad (2)$$

where $p_g(m)$ is the Fourier spectrum of $p_g(m)$ in Fourier domain. For efficiency Gaussian need to be truncated, thus generating n samples for $p_g(m)$ where

$$n = -\text{int}\,eger\left(\frac{q+1}{2}\right)+1, \ldots\ldots\ldots., N + \text{int}\,eger\left(\frac{q+1}{2}\right)-1 \qquad (3)$$

and where *integer(x)* truncates to the largest integer smaller than x for x≥0. The algorithm is initialized at $p_g(n)$, where subscript g indicates we apply a truncated Gaussian and keep updating by summation of the N shifted filters. This summation of N shifted filter can be given by

$$p_g(n) \leftarrow p_g(n) + \Delta x p_n g(n\Delta x - x_n) \qquad (4)$$

Equation (4) spreads the irregular samples onto a regular grid. The sampling is $p_g(n) = \Delta x p_n g(n\Delta x)$, similar to equation (5) in the Fourier domain which can be written as

$$P_g(m) = \sum_{l \in Z} P(m + lN) G(m + lN) \qquad (5)$$

When $P_g(m)$ is broadband, aliasing will occur when $G(m + lN) \neq 0$ for any $l \neq 0$. It is suggested that removal of the aliasing requires making the signal periodic.

$$p_g(n) = \sum_{l=-\infty}^{\infty} p_g(n + lN), n = 0,1,2,....,N-1 \qquad (6)$$

where $p_g(n + lN) = 0$ outside the interval given by equation (3). Convolution of the signal followed by the discrete transform can be represented by

$$P_g(m)_{FFT} = \sum_{n=0}^{N-1} p_g(n) e^{\frac{j2\Pi nm}{N}}, m = \frac{N}{2}, \ldots\ldots\ldots, \frac{N}{2} - 1 \qquad (7)$$

where $P_g(m)_{FFT}$ is the spectrum obtained using the FFT. Finally correction for convolution is carried out by deconvolution in the Fourier domain according to

$$P(m) = \frac{P_g(m)_{FFT}}{G(m)} \qquad (8)$$

where $P(m)$ is the approximate spectrum and $G(m)$ has been defined by Duijndam and Schonewille (1999) as

$$G(m) = \sqrt{\frac{\Pi}{b}} e^{\frac{m^2}{4b}} m = -\frac{M}{2} - 1 : \frac{M}{2} - 1 \qquad (9)$$

Computational cost for calculating the approximate spectrum can be written as

t=Convolution(qN) + FFT (NlogN) + Deconvolution(N) (10)

which is much faster than the conventional cost of $O(N^2)$. The numerical width q and the analytical width b affect the performance of the low pass filter. Increasing the numerical width of Gaussian increases the accuracy of the results. Rms error is calculated in Figure 1 and Figure 2 between slow NDFT and NFFT for optimal *q* and *b*. On the basis of numerical experiments, an improvement is suggested with respect to the width of the gauss pulse. Empirically, an analytical gaussian width can be written as

$$b=\frac{4.75}{q.dx} \qquad (11)$$

where dx is spacing of regular grid.

Examples

Algorithms cannot be applied, unless they are compared with some analytically available solution. For this purpose we are applying our algorithm to the analytically available solution of a Ricker wavelet and compared with its frequency domain representation. Figure (3), shows an application of the NFFT on the uniformly sampled Ricker wavelet which is giving the same result in Figure 3(b) as the analytical solution in Figure 3(c) of the wavelet. Figure 4(c) displays the effect of increasing the number of sample points, and makes it non-uniform. Increasing the number of sample points elevates the noise in the spectrum, although all frequencies are obtained. This presence of noise can be explained from the fact that in NFFT, columns of the Fourier matrix are not orthogonal to each other. Also, some noise will be observed when the signal is undersampled .Similarly, Figure 5 shows a seismic trace with with 5000 sample point and 25% decimated and its approximated frequency spectra, RMS error of 10^{-1} wrt FFT of the original spectrum. Accuracy and speed of NFFT can be given by Table 1 and 2. Table 1 compares the computational performance of three different algorithm FFT, NDFT(Dujhdham, 1999), and NFFT in seconds and table 2 gives rms error wrt original spectrum

Table 1: Computational Performance (Seconds)

N=M	FFT	NDFT	NFFT
100	0.0078	5.4	15
1000	.013	48.2	34.4
5000	0.299	1125	96.6
10000	0.645	4390	190.5
15000	0.971	10020	258

Table 2: Error Analysis

M=N	NFFT	NDFT
100	1.2E-04	1.63E-05
1000	1.7E-03	2.0E-05
2000	1.8E-03	2.2E-05
5000	2.2E-03	8.7E-04
7000	2.0E-03	7.9E-04

Conclusions

NFFT can be a major part of conventional algorithms. The tremendous improvement over NDFT allows us to handle large impractical and uneconomical data sets in an efficient way. 100 times speed up is achieved by NFFT over NDFT. A new generation of fast algorithms can be developed based on the NFFT approach, which can take advantage of irregular sampling to handle the problem of aliasing.

Acknowledgements

I would like to thank my supervisor Dr. R.J.Ferguson, for many discussions. I also thank Dr. K. Innanen and Dr. L. Lines for their comments on the work. Finally I thank the CREWES sponsors for the financial support.

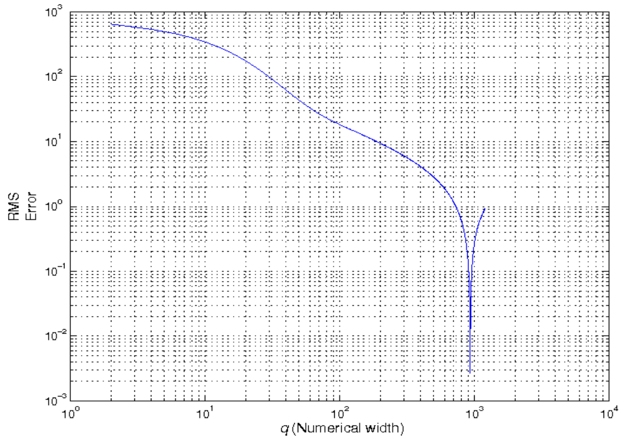

FIG 1. RMS Error for varying q

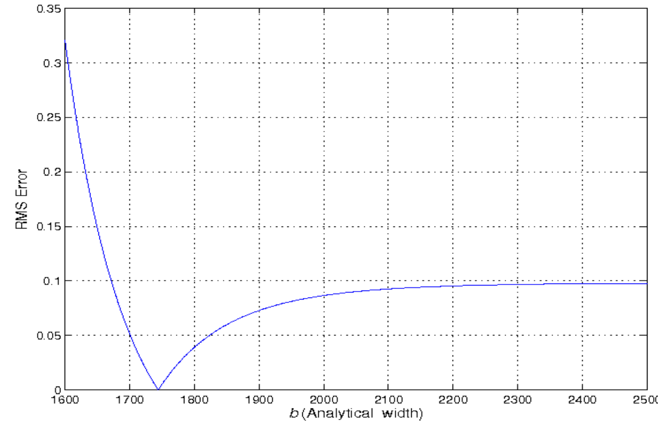

FIG 2. RMS Error for varying Analytical

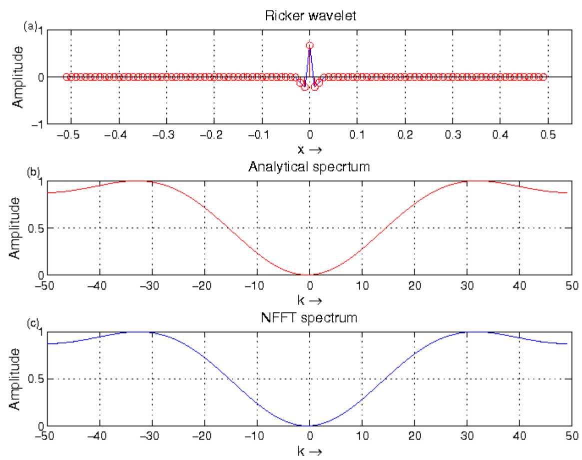

FIG 3. a) Ricker wavelet b) Analytical spectrum

c) NFFT spectrum

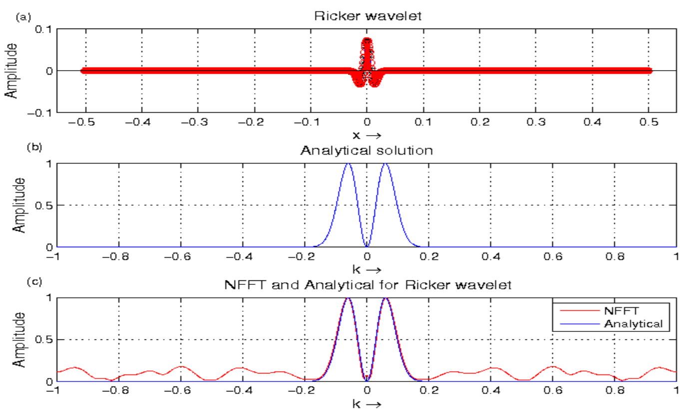

FIG 4. a) Ricker wavelet b) Analytical spectrum

c) NFFT spectrum

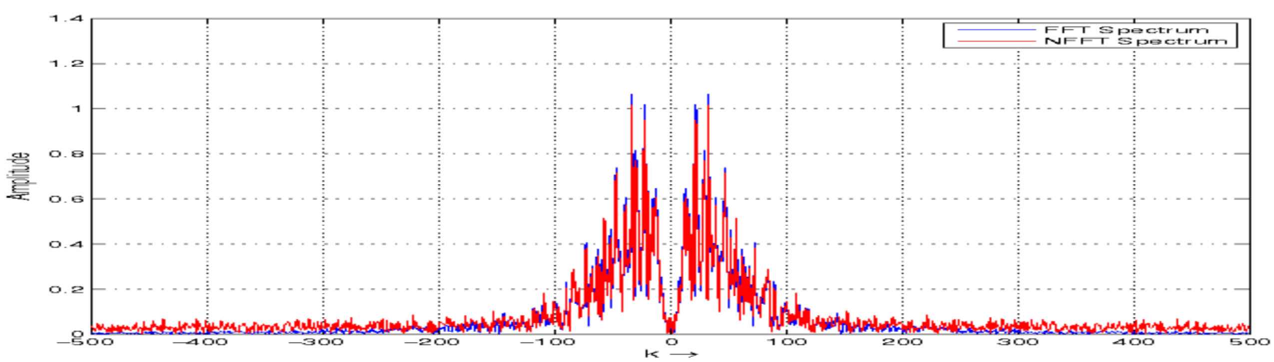

FIG 5. Comparison of Original spectrum and NFFT spectrum for Seismic trace

References

1.Clarebout, J., 1992, EARTH SOUNDING ANALYSIS : Processing vs inversion (PVI): Cambridge university press

2. Duijndam, A.J.W., and Schonewilliw, M. A,1999, Nonuniform fast Fourier transform Geophysics, 64, 551 - 573

3.Ferguson, R. J., 2006, Regularization and datuming of seismic data by weighted, damped least square: Geophysics, **71**, 67–76

4.Greengard, L., and Lee, J.Y., 2004, Accerlating the non uniform fast Fourier transform: SIAM Review.,46,443-454

5.Ronen, J., Sorin, V., and Bale, R., 1991, Spatial dealiasing of 3-d seismic reflection: Geophy. J. Int, **105**, 503–511.

6. Gulati, A and Ferguson, R. J. ,2009, NFFT: Algorithm for irregular sampling: CREWES Research Report, 21