Interpolation of nonstationary seismic records using a fast non-redundant S-transform
Mostafa Naghizadeh and Kris Inman

SUMMARY

We propose a fast and efficient method for interpolation of nonstationary seismic data. The proposed method utilizes a fast and non-redundant S transform to identify the space-wavenumber evolution of nonstationary spatial signals at each temporal frequency. The non-redundant nature of used transform renders a big computational advantage to the proposed interpolation method. Next, a least-squares fitting scheme is used to retrieve the optimal set of coefficients representative of the ideal interpolated data. For randomly sampled data on a regular grid we seek a sparse representation of the coefficients in order to retrieve the missing samples. Also, to interpolate the regularly sampled seismic data at a given frequency, we use a mask function derived from the coefficients of the low frequencies. Synthetic and real data examples are provided to examine the performance of the proposed method.

INTRODUCTION

The problems of seismic data reconstruction and interpolation have attained a special stature in the seismic data processing community in recent years. Reconstruction methods use available seismic traces, measured on irregular and/or coarsely sampled grids in space, to estimate data on a regularly and sufficiently sampled grid. An effective solution can open the door to application of multidimensional wave equation imaging and de-multiple algorithms to a data set, without having had to acquire it with the completeness these methods demand. In a useful method of interpolation/reconstruction we look for speed, stability in the presence of noise and aliasing, and ability to manage complex events. From the key issues in judging the performance of transform-based reconstruction methods and their management of complex events lies in their response to data nonstationarity.

In seismic data processing, nonstationarity means the frequency/wavenumber content of the signal varies in time/space. For instance, seismic sections which contain hyperbolic and parabolic (or any nonlinear) events produce nonstationary spatial signals in the f-x domain at a given frequency. Interpolation/reconstruction methods typically cope with nonstationary signals through spatial windowing. Inside sufficiently small spatial windows nonlinear seismic events appear linear or stationary. Hence, methods which assume stationarity such as those referenced above may be applied. Naghizadeh and Sacchi (2009) have proposed an alternative method providing a beyond-alias interpolation of nonstationary seismic data. It is essentially a modification of the f-x interpolation of Spitz (1991). In their method the windowing of the spatial axis is avoided through use of adaptive prediction filters. Although spatial windowing of data and the adaptive f-x interpolation method are capable of handling nonstationarity, they lead to computationally demanding algorithms.

On another note, there exists a range of transformations specifically designed to deal with identification and analysis of nonstationary behavior in signals that might be used as a basis for interpolation. The Gabor transform, for instance, is a class of short-time Fourier transforms that has been used for a range of processing tasks such as nonstationary deconvolution (Mardrage et al., 2003). Wavelet (Beylkin et al., 1991), curvelet (Candes et al., 2005; Candes and Donoho, 2004), and S-transform, are further examples. The S-transform (Stockwell et al., 1996) is a type of short-time Fourier transform in which the window size is frequency dependent. Larger windows are used for lower frequencies and smaller windows are used for higher frequencies. Consequently, the spectrum at a given frequency is estimated with a number of samples appropriate to that frequency. Two properties of the aforementioned transforms work against their potential as a base for an interpolation method. They are redundant (the size of the transformed data set is larger than that of the original data set) and computationally demanding.

However, recent work on the S-transform has led to a fast, non-redundant algorithm (Brown et al., 2010), renewing the possibility of developing an efficient and effective interpolation/reconstruction approach based on S-transform theory. They coined the name fast generalized Fourier transform (FGFT) for the proposed transform. In this paper, we develop such an approach and examine its behavior when applied to synthetic and field data sets. First, because the transform algorithm is new, we will describe a straightforward and intuitive frequency domain computation of the FGFT. We will then combine FGFT with a least-squares fitting principle to formulate our FGFT Interpolation method. Synthetic and field data examples are examined.

BACKGROUND THEORY

Generalized Fourier transform

The S-transform of a time signal \( g(t) \) is defined as follows (Stockwell et al., 1996):

\[
S(\tau, f) = \int_{-\infty}^{\infty} g(t) \frac{|f|}{\sqrt{2\pi}} e^{-(\frac{\tau}{2})(\frac{1}{f^2})} e^{-2\pi i f \tau} dt, \tag{1}
\]

where \( \tau \) and \( f \) are the time and frequency coordinates, respectively. The term \( \frac{|f|}{\sqrt{2\pi}} e^{-(\frac{\tau}{2})(\frac{1}{f^2})} \) is a Gaussian window which depends on time lag and frequency. The width of the Gaussian window decreases with increasing frequency. This results in finer frequency resolution for low frequencies and finer time resolution for high frequencies. If the window function is set to unity, equation 1 reverts to the ordinary Fourier transform. The S-transform can be generalized by replacing the Gaussian window with other functions, such as Gabor and B-Spline windows. Brown et al. (2010) have used the term “general Fourier-family transform” to refer to this entire group.

Equation 1 is a formula for calculation of the S-transform of an input signal \( g(t) \) in the time domain. The same calculation can be carried out in the frequency domain. In fact, the S-transform in the Fourier domain turns out to be simpler to derive and easier to implement than its time domain counterpart. Also, the frequency domain implementation of S-transform provides a framework for the FGFT method. The equivalence of the time and frequency domain S-transforms, as well as the derivation of the method and algorithm details, are presented by Brown et al. (2010). Since this work is outside of typical geophysical literature, we begin with a brief overview of the frequency-domain S-transform algorithm, exemplifying it with the chirp function illustrated in Figure 1a.

The procedure of S-transform in the frequency domain is as follows:

1. Transform \( g(t) \) to the Fourier domain to obtain \( G(f) \). Figures 1a and 1b show the original chirp function and its Fourier domain representation, respectively.

2. Create a data matrix, \( G(f, f') \), using repeated instances of \( G(f) \), each time shifting its elements (Figure 1c). This matrix is referred to as the data in the \( \alpha \)-domain, where \( f' \) represents the frequency shift. To be specific, the column of data at
frequency shift zero is the original Fourier domain representation of the data in Figure 1b. The other columns are shifted copies of Figure 1b.

3. Create a window matrix, $W(f, f')$, the same size as the $\alpha$-domain representation of the data (Figure 1d). The window function can be any symmetric smooth function such as Gaussian, Hanning, B-spline, etc. The size of windows grows linearly from lower to higher frequencies. The windows are chosen to be wider for high frequencies and narrower for low frequencies to gain proper resolution in time-frequency analysis (Brown et al., 2010).

4. Multiply $G(f, f')$ and $W(f, f')$ element by element to window the data in the $\alpha$-domain, obtaining $G' (f, f')$ (Figure 1e).

5. Apply a 1D inverse Fourier transform along each row (i.e., the $f'$ axis) of $G'(f, f')$ to obtain the $S$-Transform of the original data (Figure 1f).

The chirp function in Figure 1a has more low frequency content at the start of the time signal and more high frequency content at the end. The $S$-transform of the chirp function in Figure 1f reflects this evolution with local spectra trending toward high frequency as time increases.

Figure 1: Frequency domain implementing of S-Transform. a) Original chirp function. b) Fourier transform of (a). c) $\alpha$-domain matrix built by shifting spectrum in (b). d) Window function with the same size as the $\alpha$-domain representation of the data in (c). e) Windowed $\alpha$-domain obtained by element by element multiplication of (c) and (d). f) S-transform of data obtained by applying inverse Fourier transform on each row of (e).

Fast generalized Fourier transform (FGFT)

The frequency-dependence of the resolution of the $S$-transform suggests that computational efficiency can be increased by adopting sampling criteria varying from high to low frequency. Brown et al. (2010) used a dyadic segmentation approach, through which high frequencies (which are inherently low resolution) are coarsely sampled and low frequencies (which are inherently high resolution) are finely sampled in the $\alpha$-domain.

Figure 2 illustrates schematically an application of the FGFT algorithm. Figure 2a represents a time signal containing 16 samples. A fast Fourier transform is applied to the time signal to obtain the frequency domain representation illustrated in Figure 2b. In the frequency domain the signal is dyadically segmented (dashed boxes), and within each segment inverse Fourier transforms are applied to the data. In this article, we perform the segmentation with the window sizes which are growing by power of 2. However, depending on the application, one can deploy other forms of segmentation. It is also important to mention that each segment needs to be properly tapered on the edges using a tapering window, for instance Gaussian window, to avoid the ringing effects of the Fourier transform. Notice that smaller windows are used for low frequencies and larger windows for high frequencies.

The output is illustrated in Figure 2c. Each individual inverse Fourier transform is represented by a particular symbol in Figure 2c (square, diamond, triangle, circle). Assuming a real time signal, we expect a symmetric outcome. So we may then focus our attention on the right hand side of the output, indicated with an underbrace.

Next, to properly represent the time-frequency behavior of the data, the underbraced FGFT coefficients must be arrayed in a 2D plot. Figure 2d illustrates this arrangement of FGFT coefficients. Each individual element of a given inverse Fourier transform is distinguished from the others in that group via size. Hence each FGFT coefficient is uniquely represented by a symbol type and size. Figure 2d illustrates how these outputs are distributed. We note that there is better time resolution in the high frequencies and better frequency resolution in the low frequencies. The inverse or adjoint FGFT is performed by reversing the order of operations from Figures 2a-c, by replacing inverse Fourier transforms with Fourier transforms and vice versa. The algorithms for both forward and adjoint FGFT are described by Brown et al. (2010).

Figure 2: Graphical representation of implementing FGFT. a) Original signal with 16 time samples. b) Fourier transform of original data in (a). c) FGFT representation of data after applying inverse Fourier transform on each dashed box of data in (b). d) The time-frequency interpretation of FGFT coefficients in (c) only for positive frequencies.

Figure 3a shows the same chirp function illustrated in Figure 1a. Figure 3b shows the FGFT of Figure 3a. Figure 3c illustrates the 2D array of FGFT coefficients after proper up-scaling, i.e., the time-frequency decomposition of the chirp. Because the original chirp input is real, we only include the positive frequencies for this example. Figure 3d illustrates the adjoint FGFT acting on the FGFT coefficients of the chirp function. The adjoint FGFT recovers the original data within a small error level produced by the windowing step.
FGFT Interpolation

The conjugate gradients method does not require the explicit knowledge of $\mathbf{G}$ in matrix form. It requires the action of operators $\mathbf{G}^T$ and $\mathbf{G}$ on a vector in the coefficient and data spaces, respectively (Claerbout, 1992). The goal of the proposed algorithm is to find the coefficients $\mathbf{g}$ that minimize $J$, and use them to reconstruct the data via the adjoint FGFT operator $\mathbf{d} = \mathbf{G}^T \mathbf{g}$.

Derivation of the weight function $W$ for irregularly sampled data

For a randomly sampled signal one can impose a sparsity constraint to the FGFT coefficients. This corresponds to a weight function which diminishes small amplitude coefficients and preserves large ones (Sacchi et al., 1998). The large FGFT coefficients correspond with coherent features in a data set. By suppressing the small FGFT coefficients, the coherent data features will be reintroduced across regions where data was initially missing.

We adapt an iteratively Re-weighted Least Squares (Scales and Gersztenkorn, 1988) approach to achieve this purpose. The algorithm is summarized as follow

$$W^0 = I$$

For $k = 1, 2, 3, \ldots$

$$\hat{g}^k = \arg \min_{g} \left( \| \mathbf{d}_{\text{obs}} - \mathbf{T} \mathbf{G}^T \mathbf{W}^{k-1} \mathbf{g} \|^2 + \mu^2 \| \mathbf{g} \|^2 \right)$$

$$W^k = \text{diag}(|\hat{g}^k|)$$

End

where diag$(\cdot)$ builds a diagonal matrix from the absolute values of a vector. The weight function was initiated with an identity matrix and was updated by the absolute value of FGFT coefficients after each least-squares fitting. The updating step of IRLS method serves to amplify the high amplitude FGFT coefficients and eliminate the low amplitude ones. Our tests show that using a small number of internal iteration for conjugate gradients (3 or 4) and more iterations of external re-weighting, produces optimal results. Using more internal iterations for conjugate gradients could lead to spurious high frequency features.

EXAMPLES

Randomly missing samples

Figure 4a shows an original synthetic seismic section with three hyperbolic events composed of 202 traces. Figure 4b shows the section of missing data after randomly eliminating 50% of the traces. Figure 4c shows the FGFT reconstructed data after applying Equation 6 for each single frequency. Notice that the non-redundant nature of FGFT interpolation create small amplitude artifacts in the interpolated data. This means there is a trade-off between the resolution and speed of the interpolation method. Figures 4d, 4e, and 4f show the f-k spectra of Figures 4a, 4b, and 4c, respectively.

Regularly missing samples

We next demonstrate how the half frequency information in multidimensional seismic data, 2D in this case, can provide weights necessary to interpolate data with regularly missing samples. Merging the basic arguments of Spitz (1991) and Naghizadeh and Sacchi (2009) with the FGFT methodology, a procedure for interpolating regularly sampled nonstationary seismic data emerges:

1. Transform the original data from t-x to f-x domain.
2. Compute the FGFT of the data at a given frequency, $d(f)$, along all spatial axes, to obtain $g(f)$.
3. Create the weight function $W(f)$ to have a value one for coefficients larger than a threshold value and zero elsewhere.
4. For the frequency $f' = 2f$, interleave zero values between each pair of available spatial samples to obtain $d_{\text{obs}}(f')$. 

FGFT Interpolation

The interpolation problem is under-determined and hence to solve it we require some prior information. To provide this, let us represent the FGFT coefficients $\mathbf{g}$ of the desired, fully sampled signal $\mathbf{d}$ by

$$\mathbf{g} = \mathbf{G} \mathbf{d}.$$  (2)

where $\mathbf{G}$ represents the forward FGFT operator. The adjoint FGFT operator $\mathbf{G}^T$ can furthermore be used to express the desired interpolated data $\mathbf{d}$ in terms of $\mathbf{g}$ as follows

$$\mathbf{d} \approx \mathbf{G}^T \mathbf{W} \mathbf{g}.$$  (3)

where we have introduced a diagonal weight function $\mathbf{W}$ that preserves a subset of FGFT coefficients. The observed data $\mathbf{d}_{\text{obs}}$ is related to the fully sampled data $\mathbf{d}$ by a sampling matrix $\mathbf{T}$, therefore we have

$$\mathbf{d}_{\text{obs}} \approx \mathbf{T} \mathbf{G}^T \mathbf{W} \mathbf{g}.$$  (4)

Let us assume for the moment that the operator $\mathbf{W}$ is known. The system of equations in equation 4 is under-determined (Menke, 1989) and therefore, it admits an infinite number of solutions. A stable and unique solution can be found by minimizing the following cost function (Tikhonov and Goncharsky, 1987)

$$J = \| \mathbf{d}_{\text{obs}} - \mathbf{T} \mathbf{G}^T \mathbf{W} \mathbf{g} \|^2 + \mu^2 \| \mathbf{g} \|^2.$$  (5)

where $\mu$ is the trade-off parameter. We minimize the cost function $J$ using the method of conjugate gradients (Hestenes and Stiefel, 1952).

Figure 3: a) Original chirp signal. b) The FGFT coefficients of (a). c) 2D plot of FGFT coefficients clearly showing the time-frequency distribution of chirp signal. d) The recovered chirp signal by applying inverse FGFT on (b).
FGFT Interpolation

In order to investigate the performance of FGFT interpolation method on real data, we chose a near offset section from a Gulf of Mexico data set. The original data which contains 145 equally spaced traces is shown in Figure 5a. Figure 5b shows the interpolated data using the FGFT interpolation method. The interpolated section contains 290 traces. To provide a comparison, we also interpolated the original data using the f-x adaptive interpolation method of Naghizadeh and Sacchi (2009). The result is shown in Figure 5c. The run-time was 9 seconds for FGFT interpolation and 180 seconds for the f-x adaptive interpolation. Therefore, speed was significantly increased with the FGFT interpolation method. The interpolated traces using the FGFT interpolation method contain more noisy features than when using the f-x adaptive interpolation method. However, the performance of both methods was comparable. These operations were implemented in Matlab and run on a single quad-core notebook computer.

DISCUSSION AND CONCLUSIONS

The FGFT method is an effort to recover missing samples of a nonstationary signal without the necessity of windowing in spatial directions. For a randomly sampled stationary signal on a regular grid, seeking a sparse representation in the Fourier domain leads to successful reconstruction of missing samples. This is because a sparse Fourier representation is a property of stationary signals. This is not true of nonstationary signals. A mixture of regular sampling and a nonstationary signal further complicates the task of reconstruction by introducing artifacts in the Fourier domain. The commonly used approach to tackle this issue is to divide the signal into small windows, within which data have stationary properties. One needs to choose an optimal length for window sizes as well as overlap between windows for smooth reconstruction of missing samples. The FGFT interpolation method combines all these tasks into one operator to obtain a sparse representation of FGFT coefficients. The FGFT interpolation handles random distribution of samples on a regular grid, windowing and nonstationarity problems simultaneously. The sparsity constraint on the FGFT domain is not effective in eliminating the aliased energy of regularly sampled data, especially for high frequencies.

The computational cost of FGFT interpolation is very low compared to the alternative methods such as f-x adaptive interpolation (Naghizadeh and Sacchi, 2009). The FGFT method is memory-efficient because of its non-redundant nature. The speed advantage of FGFT comes with loss of resolution both in time/space and frequency/wavenumber axes. However, for the spatial interpolation of seismic data the resolution of FGFT is expected to be sufficient in most cases.

ACKNOWLEDGEMENTS

We thank the financial support of the sponsors of the Consortium for Research in Elastic Wave Exploration Seismology (CREWES) at the University of Calgary.
EDITED REFERENCES
Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2011 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES


Tikhonov, A. N. and A. V. Goncharsky, 1987, Ill-posed problems in the natural sciences: MIR.