Enhanced irregular seismic interpolation using approximate shrinkage operator and Fourier redundancy

Pengliang Yang ⁎, Jinghuai Gao ⁎

Xi’an Jiaotong University, National Engineering Laboratory for Offshore Oil Exploration, Xi’an 710049, China

Abstract

Article history:
Received 29 October 2014
Received in revised form 2 February 2015
Accepted 3 February 2015
Available online 13 February 2015

Keywords:
Sparsity
Seismic trace interpolation
Iterative shrinkage–thresholding (IST)
Taylor series
Redundant transform

In this paper, we propose an approximate shrinkage operator in the iterative shrinkage–thresholding (IST) algorithm to interpolate the nonuniformly sampled seismic traces. The key to designing the approximate shrinkage operator is the use of Taylor series. The redundant Fourier transform is employed to enhance the reconstruction performance, inspired by the theory of spectral compressive sensing. Numerical experiments using 3D real data and 5D synthetic data demonstrate the superiority of the proposed method.

1. Introduction

Fully recording the reflection seismic data in a multidimensional acquisition process is unrealistic for many reasons: a finite number of active recording channels, surface obstacles, as well as some other physical or financial constraints (Curry, 2008). Therefore, seismic data reconstruction plays an important role at the data processing stage.

To interpolate the regularly sampled seismic data, the linear prediction error filters (or referred to as PEF) have been utilized (Spitz, 1991; Porsani, 1999; Wang, 2002; Gullenay, 2003, just to name a few). These methods interpolate high frequencies by extracting the low-frequency nonaliased data. To address the nonuniformly sampled seismic data with missing traces, a popular way is using certain types of transforms such as Fourier transform (Sacchi et al., 1998), Radon transform (Trand and Ulrych, 2002), curvelet transform (Herrmann and Hennenfent, 2008; Yang et al., 2012a), as well as seislet transform (Liu and Fomel, 2010). There are some well known algorithms which fall under this umbrella, such as MWNI (minimum weighted norm interpolation) (Liu and Sacchi, 2004), ALFT (anti-leakage Fourier transform) (Xu et al., 2005), POCS (projection onto convex sets) (Ahma and Kabir, 2006) and MSAR (multistep autoregressive) (Naghizadeh, 2007).

The irregular sampling pattern has received more attention recently (Trickett et al., 2010). It leads to a highly reduced cost in the process of data acquisition. It is particularly promising to reconstruct the sparsely sampled seismic record without losing the essential information, combining a priori knowledge with certain transforms. This has been proved by the theory of compressive sensing (CS) (Donoho, 2006; Candès, 2006). Some convincing results have been obtained using curvelet transform and iterative shrinkage–thresholding (IST) algorithm (Herrmann and Hennenfent, 2008; Yang et al., 2012a).

This paper is intended to propose an approximate shrinkage operator-based IST interpolation method for the irregular sampled seismic data on regular grid. The principle of the design of the approximate shrinkage operator is using Taylor series. Besides the threshold, the method introduces another parameter p, which can be flexibly chosen according to the user’s experience and requirement, to obtain the competitive results as using many existing thresholding operators. The reconstruction performance of the proposed method will benefit a lot from the use of redundant Fourier transform. A 5D synthetic data and a 3D field data will be utilized to demonstrate the value of the proposed method.

2. Seismic data interpolation using IST algorithm

2.1. Problem formulation

The observed seismic record $d_{obs}$, including many missing samples, is connected with the complete seismic data $d$ to be recovered via the relation

$$d_{obs} = Md,$$

in which $M$ denotes a mask matrix with diagonal entries 1 for the observed samples and 0 otherwise. To recover the missing traces...
of the seismic data, one needs to project the seismic data into certain transform domain, that is,
\[ d = Ax, \]  
where \( A \) corresponds to a synthetic operator. It leads to
\[ d_{\text{obs}} = Md = MAx = Kx, \]
where \( K = MA \). The complete seismic data \( d \) can be recovered by minimizing the following inverse problem
\[ \min_x J(x) = \frac{1}{2} \| d_{\text{obs}} - Kx \|^2_2 + \lambda R_p(x), \]
where \( R_p(x) \) is always an \( \ell_p \)-norm regularization term. Currently, a popular choice is \( p = 1 \) which is associated with sparseness constrained data reconstruction, see more details in Herrmann and Hennenfent (2008) and Hennenfent et al. (2010). Eq. (4) is a formulation since the complete seismic data \( d \) can be synthesized from its representation coefficients, \( d = Ax \) where \( x \) is the minimizer of the optimization problem.

The above problem can be solved by IST algorithm (Figueiredo and Nowak, 2003; Daubechies et al., 2004):
\[ x^{(k+1)} = T_{\lambda} \left[ x^{(k)} + K^* \left( d_{\text{obs}} - Kx^{(k)} \right) \right], \]
where \( k \) is the iteration number, \( K^* \) stands for the adjoint of \( K \). \( T_{\lambda} \) is an element-wise shrinkage-thresholding operator. Particularly, for \( p = 1 \) it corresponds to the soft thresholding defined in Eq. (8). It is noteworthy that when the synthetic operator \( A \) is taken as a tight frame, i.e., \( A^* A = I \), the IST interpolation algorithm is closely equivalent to the projection onto convex sets (POCS) method (Yang et al., 2012a, 2013b),
\[ d^{(k+1)} = d_{\text{obs}} + (I-M) AT_{\lambda} [A^* d^{(k)}], \]
which solves
\[ \min_d J(d) = \frac{1}{2} \| d_{\text{obs}} - Md \|^2_2 + \lambda R_p (A^* d). \]

Eq. (7) is referred to as the analysis formulation since it directly analyzes the complete seismic record \( d \) as the unknown (Elad et al., 2007). Obviously, the POCs method updates in data domain, while the IST algorithm does updating in the model domain. These approaches are popular in many applications, such as image reconstruction and inpainting (Gunturk et al., 2002; Guleryuz, 2006a, 2006b; Cai and Chan, 2008), and seismic data interpolation (Abma and Kabir, 2006).

Minimizing an inverse problem with the objective function \( E(x) = \frac{1}{2} \| d_{\text{obs}} - Kx \|^2 \) can often be divided into 3 steps:

- apply the forward operator \( K \) to calculate the residual: \( r^k = d_{\text{obs}} - Kx^k \);
- apply the adjoint operator \( K^* \) to find the gradient: \( g^k = K^* r^k \);
- update the model according to the gradient \( g^k \) and the step length \( \mu \): \( x^{k+1} = x^k + \mu g^k \), where \( \mu \) can be found via minimization of \( J(x) \) with respect to \( \mu \).

Due to the penalty \( R_p(x) \), the IST interpolation algorithm requires an additional step, namely the shrinkage-thresholding, to enforce the sparsity of the transform domain coefficients. The step length is set to be 1 allowing for the use of tight frame \( A \) and the mask operator \( M \), which makes the spectral radius of \( K \) is 1 to guarantee the convergence of the algorithm (Daubechies et al., 2004). Shaping regularization (Fomel, 2007, 2008) provides us a general and flexible framework so that we do not need to care too much about the exact expression of the penalty function \( R_p(x) \) when a particular kind of shaping operator is used. In what follows, we introduce the approximate shrinkage operator as a nonlinear shaping operator (Fomel, 2008) in the framework of IST algorithm.

2.2. Approximate shrinkage operators

The a priori knowledge about sparseness in transform domain always holds well from a large number of informed studies (Lustig et al., 2008). Shrinkage–thresholding becomes a useful tool to filter out the unwanted small values to promote sparsity. The \( \lambda \)-constraint entails the soft thresholding in IST algorithm (Figueiredo and Nowak, 2003; Daubechies et al., 2004):
\[ \text{Soft}_\lambda(x) := \begin{cases} x - \lambda & |x| \geq \lambda \\ 0 & |x| < \lambda. \end{cases} \]

Even though soft thresholding has been widely used for sparsity-promoting regularization, some other alternatives exist when the regularization term is specified in various forms, such as hard thresholding operator, Stein thresholding operator, as well as the newly proposed \( p \)-norm thresholding operator. Hard thresholding is simply pointwise truncation of the coefficients:
\[ \text{Hard}_\lambda(x) := \begin{cases} x & |x| \geq \lambda \\ 0 & |x| < \lambda. \end{cases} \]

Stein thresholding is very similar to soft thresholding (Peyre, 2010; Mallat, 2009):
\[ \text{Stein}_\lambda(x) := x \max \left( 1 - \left( \frac{|x|}{\lambda} \right)^2, 0 \right). \]

Stein thresholding does not suffer from the bias of soft thresholding (Peyre, 2010), i.e.,
\[ |\text{Stein}_\lambda(x) - x| \rightarrow 0, |\text{Hard}_\lambda(x) - x| \rightarrow \lambda, \text{ if } x \rightarrow \infty. \]

The recent advance on nonconvex optimization (Chartrand, 2012; Chartrand and Wohlberg, 2013) shows that the shrinkage operator in IST algorithm (Eq. (5)) can be generalized as
\[ \text{pThresh}_{p}(x) = \frac{x}{|x|} \max \left( 0, |x| - \lambda^{2-p} |x|^{p-1} \right). \]

with a parameter \( p < 2 \). The pThresh operator can be considered as a generalized version of previous thresholding operators. When \( p = 1 \), it is exactly the soft thresholding. For \( p = 0 \), it is the so-called Stein thresholding. This thresholding operator does not suffer the bias as long as \( p \neq 1 \). A very interesting thing about pThresh operator is that IST algorithm still works well when \( p \) is negative, \( p < 0 \). As shown in Fig. 1, pThresh operator quickly approaches hard thresholding when \( p \rightarrow -\infty \).

As Peyre (2010) reported, “It is possible to devise many thresholding nonlinearities that interpolate between the hard and the soft thresholder. Depending on the statistical distribution of the coefficients, these thresholders might produce slightly better results.” For the convenience of observation, we express all the above shrinkage operators as
\[ T_\lambda(x) = x \cdot \alpha, \]
where $a$ is a non-negative factor less than or equal to 1:

$$a = \begin{cases} 1 : 0, & \text{Hard} \\ 1 - \frac{\lambda}{|x|} : 0, & \text{Soft} \\ 1 - \frac{\lambda^2}{|x|^2} : 0, & \text{Stein} \\ 1 - \frac{\lambda}{|x|^p} : 0, & \text{pThresh} \end{cases}$$

(14)

where $\{A:B\}$ is an if-else statement in C-code style: The expression equals $A$ if the statement $(\cdot)$ is true, and $B$ otherwise.

It is tempting for us to design an approximate shrinkage operator by multiplying a factor less than 1 on the original data, which may lead to better reconstruction performance in seismic data interpolation. The observations from these existing operators motivate us to propose the approximate shrinkage operator based on Taylor series. Several forms of Taylor series are listed as follows:

$$1 - \ln(1 + z) = 1 - z + \frac{1}{2} z^2 - + \frac{(-1)^n}{n} z^n - + |z| < 1$$

(15)

$$\frac{1}{1 + z} = 1 - z + z^2 - + (-1)^n z^n - + |z| < 1$$

(16)

$$\exp(-z) = 1 - z + \frac{1}{2!} z^2 - + + \frac{(-1)^n}{n!} z^n - + |z| < \infty.$$  

(17)

Obviously, the 1st order approximation of all the above functions is $1 - z$, while hard thresholding operator is merely a zero-order approximation. If we take $z = \left(\frac{x}{C_{16}/C_{17}}\right)^{2-p}$, a variety of operators whose 1st order approximations are the existing thresholding operators can be obtained. However, in this paper we use the exponential approximant allowing for the convergence radius applicable to a wide range of parameter choices of $p$, leading to

$$a = \frac{|x| \geq \lambda?}{\exp\left(-\left(\frac{\lambda}{|x|}\right)^{2-p}\right) : 0}.$$  

(18)
Note that for any $x \neq 0$, the exponential factor is less than 1, $\exp\left(-\left(\frac{\lambda}{|x|}\right)^{2-p}\right)<1$. We directly multiply $a = \exp\left(-\left(\frac{\lambda}{|x|}\right)^{2-p}\right)$ with $x$, resulting in the proposed shrinkage function

$$\text{expThresh}_p(x) = x \exp\left(-\left(\frac{\lambda}{|x|}\right)^{2-p}\right)$$

in which we define $\text{expThresh}_p(0) = 0$. Similar to the pThresh operator in Eq. (12), besides the threshold $\lambda$ we have another parameter $p$ which can be flexibly chosen according to our needs. The most significant property of the expThresh operator is that it is free of non-differentiable singularity at the threshold point $\lambda$, exhibiting smooth transition between small and large values, according to Fig. 2.

2.3. Fourier redundancy from zero-padding

The previous subsection shows the generality of IST algorithm using shrinkage operators, which can be understood as an iterative nonlinear shaping procedure. Meanwhile, the reconstruction performance of the non-uniformly sampled data is closely related to the transform used. The redundant transform is of special preference in sparsity-promoting regularizations (Cai and Chan, 2008; Starck et al., 2007; Elad et al., 2005): It preserves energy and ensures an isometric relation between the input signal and the output coefficients; the supports of these frames at different scales may locally overlap with the missing blocks. Thus, partial loss of data can be tolerated without adverse effects (Herrmann and Hennenfent, 2008; Cai and Chan, 2008; Yang et al., 2012a).

Although many other nice transforms are available, in this work we use Fourier transform as ideal dictionary for the sparse representation of the oscillating features of the seismic data, based on the recognition that seismic data can be deemed to be the convolved result of the underground layer structure and the seismic wavelet. Inspired by the newly proposed spectral compressive sensing (SCS) theory (Duarte and Baraniuk, 2013), we here employ zero-padding technique to enforce fast Fourier transform (FFT) to be redundant (Yang et al., 2012b, 2013a). In practice, the amount of zeros padded for FFT has to balance the computational efficiency and the performance improvement of reconstruction. Zero-padding a vector $x \in \mathbb{R}^m$ to be $y \in \mathbb{R}^m_0$ can be mathematically expressed as

$$y = \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} I_m \\ O_{(m'-m) \times m} \end{bmatrix} x$$

where $I_m \times m$ and $O_{(m'-m) \times m}$ are the identity and zero matrix. Thus, the zero-padding operator is

$$P = \begin{bmatrix} I_m \\ O \end{bmatrix}$$

and the adjoint is

$$P^\dag = \begin{bmatrix} I_m \\ O \end{bmatrix}.$$
Therefore, the composite operator using zero-padded FFT will be $A = FP \in \mathbb{R}^{m \times m}$ which can be used as a tight frame for IST data interpolation because

$$A^* A = P' F' FP = [I_{m \times m} 0] [I_{m \times m} 0] = I_{m \times m}.$$  \hspace{1cm} (23)

where $F$ is the Fourier operator in a space with more elements, $F' F = I_{m' \times m'}$.

Actually, FFT with zero-padding is a very common technique, which allows much finer frequency partition on unit circle to better exhibit the detailed information of the signal in frequency domain. However, that is not the only point we would like to illustrate. What we want to emphasize is: Zero-padding technique provides us a simple yet useful way to bring redundancy into an (any) orthogonal transform. It will greatly improve the iterative thresholding reconstruction, as we will see in the numerical demonstration. When the transform is not FFT any more, we cannot simply understand things with the corresponding concepts in frequency domain. The rigorous proof above shows that the orthogonal transform $F$ can be Fourier transform, wavelet transform, or any other orthogonal transform we plan to use. This paper is using zero-padded FFT to obtain Fourier redundancy. Other kinds of redundancy using a specific orthogonal transform (like wavelet transforms) are available from zero-padding.

### 3. Numerical results

In the numerical examples, we use a user-defined percentage to determine the thresholding quantile during the iterations, which can efficiently be computed via Hoare’s algorithm. To quantify the interpolated results, we define the signal-to-noise ratio as

$$\text{SNR} = 10 \log_{10} \frac{d^2}{||d - \hat{d}||^2} \text{ (dB)}$$  \hspace{1cm} (24)

where $d$ and $\hat{d}$ are the true and estimated complete data, respectively.

#### 3.1. 5D synthetic example

Using the method of Jin (2010), we generate a 5D data including two 5D super-planes, which satisfy $y = x_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4$ and stretch according to two sets of slope $(a_1, a_2, a_3, a_4)$. It has 100 points in temporal direction and is of size $10 \times 10 \times 10 \times 10$ in spatial coordinates. The 5D synthetic seismic data is then generated by convolving a Ricker wavelet with the 5D data volume. We display a 4D section and a 3D section of the 5D data in Figs. 3 and 4. Due to the difficulty of display, the next dimension is appended after the previous one. After randomly removing 90% traces, we obtain the observation to be interpolated, as show in Fig. 5.

![Fig. 8. Comparison of corresponding 3D section reconstruction with and without Fourier redundancy.](image-url)
Owing to the large percentage of missing traces, the SNR of the observation is very low, only $-9.54$ dB. Based on this synthetic observation, the IST algorithm using the proposed shrinkage operator is compared with the method using other existing thresholding operators. As shown in Table 1, the resulting SNR honors the superiority and necessity of the proposed shrinkage operator in high dimensional data reconstruction.

![Comparison of corresponding 3D section reconstruction with and without Fourier redundancy in the presence of random noise (variance = 0.01).](image)

**Fig. 9.** Comparison of corresponding 3D section reconstruction with and without Fourier redundancy in the presence of random noise (variance = 0.01).

A real 3D shot gather is decimated randomly using a percentage of 50%, leading to the observation with SNR = $-0.495$ dB.

![A real 3D shot gather is decimated randomly using a percentage of 50%, leading to the observation with SNR = $-0.495$ dB.](image)

**Fig. 10.** A real 3D shot gather is decimated randomly using a percentage of 50%, leading to the observation with SNR = $-0.495$ dB.
especially when a large percentage of the seismic traces are missing. Compared to many other operators, expThresh operator exhibits better reconstruction performance.

The Fourier redundancy can further improve the obtained result \((p = 0.78)\) above. The SNR based on redundant Fourier transform (spatially padded to be \(16 \times 16 \times 16 \times 16\)) is higher than that of non-redundant transform (original spatial size \(10 \times 10 \times 10 \times 10\)). It can be seen clearly that the large amount of noise using non-redundant FFT has been significantly attenuated in redundant Fourier reconstructed data volume, by comparing Fig. 6 with Fig. 7. The 3D sections as well as the error panels extracted from the corresponding volume are shown in Fig. 8.

It is very common that the field data is contaminated by noise. We repeat the synthetic reconstruction experiment by adding random noise with variance = 0.01 and extract the corresponding 3D section for comparison. As shown in Fig. 9, the Fourier redundancy is of particular importance to combat the strong noise and improve the interpolation outcome.

### 3.2. 3D real data interpolation

The second example is a 3D field data reconstruction using the proposed shrinkage operator based IST algorithm. As illustrated in Fig. 10, the real 3D shot gather is decimated randomly with a percentage of 50%. The reconstruction results using FFT with and without redundancy, as well as the corresponding error panels are shown in Figs. 11 and 12. As a trade-off between computational consumption and improved reconstruction performance, here we use the double zero-padding strategy proposed in Yang et al. (2012b, 2013a): The number of points \(n_i\) in ith spatial axis is padded to be \(2 \cdot 2^\log_2(n_i)\). From Fig. 11, the Fourier redundancy enhances the reconstruction a lot, making the interpolated seismic events more visible while introducing less noise. The computed SNR further demonstrates this point.

We tested 3D interpolation with much larger decimating percentage (90%). Unfortunately the reconstruction fails not only with the method using the proposed shrinkage operator, but also with the methods using other existing thresholding operators. It honors the necessity of 5D...
interpolation for field data. In fact, the inferred study shows that in 5D the decimating can be sparse enough (sometimes sparser than 90% missing rate) (Trickett et al., 2010).

4. Conclusions and discussions

This paper shows that the basic principle of designing an approximate shrinkage operators is the use of Taylor series, in the context of seismic interpolation using IST algorithm. Even if many different approximate functions exist, the exponential shrinkage operator is combined with redundant Fourier transform to better reconstruct 5D synthetic data and 3D field data. The numerical results show that Fourier redundancy plays an important role in the improvement of interpolation result. The best reconstruction performance can be obtained with a good choice of the additional parameter $p$, which should be determined by experimental test.

Although numerically working well, it is valuable to establish the exact expression of the penalty for the objective function in the future when the proposed exponential shrinkage operator is taken in the shaping regularization, which is out of the scope of this paper. Besides the Fourier transform, many other transforms can also be padded with zeros to enable redundancy and directly applied to the IST framework using the proposed shrinkage operator. Furthermore, any composite transform $A$ obtained by combining two tight frames $A_1$ and $A_2$, $A = A_1A_2$, is applicable because the fact that $A_1^*A_1 = I$ and $A_2^*A_2 = I$ leads to $A^*A = I$.

Acknowledgments

The first author would like to thank Sergey Fomel and Karl Schleicher for the valuable discussions on shrinkage operators and shaping-based interpolation algorithms during his visit to Bureau of Economic Geology, The University of Texas at Austin. We thank the Madagascar open software package (http://www.ahay.org) for visualization of the results.

References