Application of Sparse Dictionary Learning to Seismic Data Reconstruction[1]

Hamid Reza Khatami[1], Mohammad Ali Riahi[2*], and Mohammad Mahdi Abedi[3]

[1]Petroleum, Mining and Materials Engineering, Islamic Azad University, Central Tehran Branch, Tehran, Iran.
[2] Institute of Geophysics, University of Tehran, Tehran, Iran.
[3] BCAM - Basque Center for Applied Mathematics, Spain.

*Corresponding author(s). E-mail(s): mariahi@ut.ac.ir, h-khatami@kiau.ac.ir, mm.abedy@gmail.com.

ABSTRACT

According to the principle of compressed sensing (CS), under-sampled seismic data can be interpolated when the data becomes sparse in a transform domain. To sparsify the data, dictionary learning presents a data-driven approach trained to be optimized for each target dataset. This study presents an interpolation method for seismic data in which dictionary learning is employed to improve the sparsity of data representation using improved K^{th} Singular Value Decomposition (K-SVD). In this way, the transformation will be highly compatible with the input data, and the data in the converted domain will be sparser. In addition, the sampling matrix is produced with the restricted isometry property (RIP). To reduce the sensitivity of the minimizer term to the outliers, we use the smooth L_1 minimizer as a regularization term in the regularized orthogonal matching pursuit (ROMP). We apply the proposed method to both synthetic and real seismic data. The results show that it can successfully reconstruct the missing seismic traces.

KEYWORDS: Compressed Sensing, Dictionary Learning, Optimization, Reconstruction, Sparsity.

[1] This is a preprint. This paper is published in Journal of Seismic Exploration, Volume 32, Issue 2, Pages 185 – 204. April 2023.

INTRODUCTION

Conventional digital data acquisition systems are based on the Shannon-Nyquist sampling theorem, which states that the characteristics of a band-limited analog signal can be determined exactly from a sequence of samples taken at a rate twice the signal's frequency bandwidth (Herrmann et al. 2011). This fundamental result enables lossless data processing of signals. This fundamental result enables lossless data processing of signals. However, the Shannon-Nyquist theorem dictates an excessively high sampling (Baraniuk and Steeghs 2017). When the signal of interest has some structure, such as when the signal allows a representation in the transform domain with few significant and many zero or insignificant coefficients, compressed sensing (CS) demonstrates that periodic sampling at Nyquist rates is far from optimal (Herrmann et al. 2011).

CS is the use of a few linear measurements to extract information from sparse data. While other researchers, to compress the data, modified the conventional approach of signal processing, in which a whole signal is generated and subsequently compressed (Donoho et al. 2005). However, It requires significant effort to reliably capture very large signals to compress them by eliminating the information.

Discrete cosine transform (DCT), Fourier transform, wavelet transform, and Curvelet transform are some examples of conventional methods. These transformation methods face some issues to represent data in a sparse form. For instance, DCT and Fourier transform is not well presenters of local characteristics of seismic data and Curve let transform cannot be selected based on the characteristics of seismic data. In contrast, the CS method mitigates these pitfalls appropriately. Recent works on CS combine the two processes in such a way that we can immediately perceive the signal or its fundamental components with a few linear measurements is an efficient method. There are numerous techniques for reconstructing sparse signals using CS, primarily based on two approaches: The first method, Basis Pursuit (BP) is based on optimization and is linearly solvable. The second strategy consists of greedy algorithms. A greedy algorithm is a simple, intuitive algorithm that is used in optimization problems. The algorithm makes the optimal choice at each step as it attempts to find the overall optimal way to solve the entire problem. Greedy. The main greedy methods are orthogonal matching pursuit (OMP) (Mallat and Zhang, 1993, Tropp and Gilbert, 2007), stage-wise orthogonal matching pursuit (StOMP) (Donoho et al. 2005), regularized orthogonal matching pursuit (ROMP) (Needell and Vershynin, 2010), and compact sampling matching pursuit (CoSaMP) (Needell and Tropp, 2009). According to the theoretical results and actual examples presented in the preceding sources, the computational speed of the greedy technique is its primary advantage. Several hybrid algorithms attempt to utilize the favorable aspects of both the basic pursuit and greedy approaches (Gilbert et al. 2005; Iwen, 2007). These hybrid approaches improve execution speed but put strict constraints on the measurement matrix.

The majority of real-world data are not sparse and must be mathematically changed to new domains before becoming sparse. Using dictionary learning, rather than conventional transformations such as Fourier, wavelet, radon, and Gabor, one might design a specialized domain to improve sparsity. Dictionary learning identifies data patterns and has applications in image processing, image classification, denoising, and audio processing, face recognition, signal classification, data categorization, and image analysis (Moura et al. 2017; Meng et al. 2017; Oguz et al. 2016). Numerous studies have also been carried out recently in the seismic community at DL (Yu et al. 2015; Chen et al. 2016; Zhou et al. 2016; Chen 2017; Nazari Siahsar et al. 2017a; Nazari Siahsar et al. 2017b; Zu et al. 2018; Wang et al. 2021). These methods focus mainly on solving the problems encountered in deblending, interpolation, and denoising of 2D/3D seismic data. The algorithm DL was first proposed by (Elad and Aharon, 2006) and has reached the state-of-art in denoising gray images. A dictionary typically represents the fundamental structures of the images and is learned in small areas (She et al. 2019).

Seismic data are mostly under-sampled in the spatial dimension. This affects seismic processing results so that they are demanded to be interpolated. Some researchers have interpolated seismic data depending on its sparseness in several transformation domains (Herrmann and Hennenfent, 2008). By assuming a transformed sparse wavefield, their solution contains a convergent function with smooth L1 constraints. This study uses CS and dictionary learning to interpolate missing seismic data. We investigate creating suitable sampling operators, the aliasing problem, the seismic data reconstruction, and improving the results using fewer data points. First, the concept of sparsity is explained and interpolation of seismic data in the sparse domain is discussed. In the end, constraints and methods for dictionary learning and transforming the data to the sparse domain are studied. Governing the idea of (Sun et al. 2019) we use ROMP to accelerate the calculations of dictionary learning while enhancing it. As known, smooth L1 regularization is used for doing box regression on some object detection systems and is less sensitive to outliers while conventional L1 regularization is sensitive to outliers. The smooth L1 tries to mimic the L1 optimizer while being smooth. The smoothness property allows for treatment as smooth continuous regularization, which is in general easier than a non-smooth optimizer. Smooth L1 regularization can be interpreted as a combination of L1 regularization and L2 regularization. It behaves like L1 regularization when the absolute value of the argument is high, and it behaves like L2 regularization when the absolute value of the argument is close to zero. The equation is:

L1; smooth={|x| if |x|>α;

$\frac{1}{|\alpha|}x^2$ if $|x| \leq \alpha$.

Therefore, as a regularization term of ROMP, the smooth L1 regularization instead of the L1 regularization is used. By using smooth L1 regularization our approach differs from Sun et al. (2019). The ROMP method using smooth L1 regularization was successfully applied to common seismic reflection data and the results and some suggestions for future works are presented.

Seismic data regularization is a popular ongoing research field. In addition to the mentioned methods, there are also other numerous methods such as neural networks or machine learning (Kaur et al. 2019) which each uses an underlying mathematical property of the data. In this paper, we focus on the sparsity of the data in one domain. The fact that the proposed algorithm learns the sparse domain from its input data makes it more promising than other transformations such as the Fourier transform.

THEORY

Consider x as complete noiseless seismic data and y as incomplete and noisy data. Interpolation could be defined as applying a mathematical operator as follow:

$$y = \phi x \tag{1}$$

where ϕ is a linear operator which maps a higher dimensional vector of $\mathbf{x} \in \mathbb{R}^n$ to a lower dimension vector of $\mathbf{y} \in \mathbb{R}^m$. In our case, y is the incomplete seismic data, while x represents the complete seismic data. Obtaining a higher-dimensional signal from a lower-dimensional signal has infinite solutions, in other words, equation (1) is an under-determined ill-conditioned problem.

Sparsity and inverse problem

Equation (1) is under-determined and has infinite solutions. By adding some constrain, we can decrease the number of solutions. A popular approach is that we suppose that the signal is sparse. Although the vast majority of signals are not sparse, most of their components are small enough to be neglected. Few linear measurements are used to rebuild sparse signals using sparse reconstruction techniques (Lan et al. 2022). A measured sample of signal x is the inner product of $x \in \mathbb{R}^n$ and a vector-like $\phi_i \in \mathbb{R}^n$. The combination of m signal measurements is equivalent to multiplying the sampling matrix ϕ by dimensions of m×n whose columns are ϕ_i to the signal. The sparse reconstruction is to recover a k-sparse signal x from a sampled vector $y = \phi x$. Even though the number of existing samples is less than the number of samples in the model, a unique solution can be extracted using an under-determined inverse problem and sparsity constraint.

$$\min_{\mathbf{x} \in \mathbb{R}^n} (\| \mathbf{x} \|_0) \quad \text{subjects to } \| \mathbf{y} - \phi x \|_2 < \varepsilon \tag{2}$$

If x is k-sparse (it is a linear combination of only K basis vectors), accordingly y is also k-sparse. Equation (2) is a restricted inverse problem, which contains two terms: the first term is to adopt the data and the second term is to adopt the result with the L0 regularization. The L0 regularization is not convergent and has many local minimums (Boyd, 2004) and its minimization is difficult. Therefore, in practice a weaker approximation of this regularization (i.e. smooth L1 regularization) is used:

$$\min_{\mathbf{x} \in \mathbb{R}^n} (\| \mathbf{x} \|_1) \quad \text{subject to } \| \mathbf{y} - \phi \mathbf{x} \|_2 < \varepsilon \tag{3}$$

The smooth L1 regularization is convergent and it could be minimized by using linear programming and inner point methods.

Condition of sparse reconstruction

If the columns of the sampling matrices ϕ create an orthogonal system, the inverse problem has a unique solution. Geometrically, ϕ must contain a batch of unit vectors in Hilbert space in which their mutual angles are equal. Therefore, these vectors have the lowest correlation. This property is known as an equiangular tight frame (ETF) (Strohmer and Heath, 2003). While using over-complete dictionaries in CS, ETF is valuable (Candes et al. 2008). They showed that the measurement matrix ∅ must satisfy the RIP (Restricted Isometry Property) (Candes et al., 2006);

$$(1-\delta_{2k}) \parallel \mathbf{x} \parallel_2^2 \leq \parallel \phi\mathbf{x} \parallel_2^2 \leq (1-\delta_{2k}) \parallel \mathbf{x} \parallel_2^2 \\ \delta_{2k} \in (0,1) \qquad (4)$$

Where δ2k is the smallest δ in the equation for which the above equation (4) is satisfied for 2k-sparse signals. If δ is much smaller than one, the above equation indicates that each k subset (or less) of the columns of ϕ is orthogonal and unitary. This property is called the restricted isometry property (RIP). Although equation (4) is an important theoretical basis for the analysis of the thin constraint of the matrix ϕ, due to the complexity of RIP calculations, its study is not easy even for medium-sized matrices. Thus, apart from certain types of matrices, such as sub-Gaussian, quasi-Fourier (Candes et al. 2006), and some structured matrices (Duarte and Eldar, 2011), it is difficult to check if the matrix follows RIP. Instead of RIP, (Candes et al. 2008) propose mutual coherence between the α_i and α_j coefficients (Candes et al. 2006), which is defined as follows:

$$\mu(\phi) = \max_{i\neq j} \frac{|< \alpha_i \cdot \alpha_j >|}{\|\alpha_i\|_2 \|\alpha_j\|_2}, \qquad i,j = 1,2,3,\ldots,N \qquad (5)$$

Where α_i and α_j represent the i^{th} and j^{th} column vectors of matrix ϕ. Mutual coherence μ(ϕ) measures the maximum similarity between columns of ϕ, and it has a major impact on the output of the compressed sensing algorithm. If this similarity is high, it affects the result of equation (3) and $\mu(\phi)$ must be small and near zero (Donoho, 2006).

CS AND SEISMIC DATA INTERPOLATION

In data reconstruction using intense measurement theory, two considerations must be made. First, according to equation 3, must be presented seismic data in a sparse form, and second, according to equation 5, the measurement matrix must meet the minimum mutual coherence condition. In this case, by applying a suitable algorithm, the data can be accurately interpolated.

Sparse Representation of the Seismic Data

Consider a simple two-dimensional geological model in which some layers have different densities and velocities of seismic waves. Synthetic seismic data of this model is created by using a Ricker wavelet with the dominant frequency of 30 Hz, while forty percent of the traces are randomly deleted figure 1a. The spatial distances of the traces are 20 meters and time intervals are 4 milliseconds. Figure 1b shows the Fourier transform of this section in the wave number-frequency domain, where due to random sampling, the frequency spectrum is disturbed so that a true reconstruction of the data with this spectrum is not possible. Using ϕ on each row of the complete section yields the samples of the incomplete section as shown in equation 1. In the example of figure 1, the number of the original points (N) is equal to 50 while the number of the sampled points (M) is equal to 30.

To calculate the original data x from the incomplete data y, the number of unknowns is larger than the number of information; therefore, this equation is under-determined. Conventional solutions to such systems do not provide a unique answer. The objective here is to use CS theory for reconstruction, which requires a sparse condition to be satisfied. In the spatial-temporal domain, seismic data are not subject to this requirement. Due to the discrete and irregular sampling intervals (figure 1a), the Fourier spectrum of the mentioned data and its incomplete data is disturbed, and the data is not properly sparse. On the other hand, due to the folding of the primary events, aliasing has generated artifact events. Consider a sparse and discrete signal x, with N samples. This signal can be considered as a linear combination of the orthogonal basis of Ψ.

$$\Psi^T = [\Psi_1, \Psi_2, \Psi_3, \cdots, \Psi_N] \tag{6}$$

where Ψ^T is the transpose of Ψ. Since the basis is orthogonal, we rewrite x as follows:

$$\mathbf{x} = \Psi\alpha = \sum_{i=1}^{N} \alpha_i \Psi_i \tag{7}$$

Where $\alpha_i =< \mathbf{x}, \Psi_i > \Psi$ is called sparse coefficients collection of x, if only k number of coefficients are nonzero (k << N). In this case, x is k-sparse. If the transformation coefficients of seismic data under an orthogonal transformation like Fourier or Curvelet are zero or close to zero, seismic data is sparse in the transformation domain.

Based on CS theory, using a transformation operator (Ψ) we transform the incomplete data x to a domain where the transformation coefficients are sparse. Incomplete data x to a domain where the transformation coefficients are sparse.

$$\alpha = \Psi_{\mathbf{x}}$$
$$\mathbf{x} = \Psi^H \alpha \tag{8}$$

The superscript H indicates the conjugate transpose. Combining equations (1) and (8) results in equation (9):

$$\mathbf{y} = \phi\mathbf{x} = \phi\Psi^H \alpha = \Theta\alpha \tag{9}$$

Where Θ is sensing matrices which must be singular so that we could find coefficient α from y. Equation (9) is under-determined and it has no unique solution. Since the coefficient α must be sparse, we rewrite (9) to an optimizing problem (Donoho, 2006; Candes et al. 2006):

$$\tilde{\alpha} = \text{argmin}_{\alpha} \parallel \alpha \parallel_1 \;\; s \cdot t \cdot \; \mathbf{y} = \Theta\alpha \tag{10}$$

Using basic pursuit or greedy approaches and calculating transformation coefficients $\tilde{\alpha}$ while keeping in mind the sparsity condition, one could estimate the seismic data:

$$\tilde{\mathbf{x}} = \Psi^H \tilde{\alpha} \tag{11}$$

Dictionary Learning

When a matrix or vector is transformed into a sparse domain, the signal can be reconstructed using a variety of reconstruction procedures (Needell and Tropp, 2009; Lotfi and Vidyasagar, 2018). Unlike common dictionaries such as Fourier or Wavelet, building and learning a dictionary that fits the input data enhances sparsity dramatically. Dictionary learning tries to use input data $\mathbf{Y} \in \mathbb{R}^{m\times l}$ and find a dictionary $\mathbf{D} \in \mathbb{R}^{m\times n}$ and sparse representation (Zhou et al. 2016) of the input data $\mathbf{X} \in \mathbb{R}^{n\times l}$ in a way that the difference between **Y** and **DX** is minimized. This is an optimizing problem as in Equation 12:

$$\min_{\mathbf{D},\mathbf{x}_i} \sum_{i=1}^{n} \|\mathbf{y}_i - \mathbf{D}\mathbf{x}_i\|_F^2 + \lambda\|\mathbf{x}_i\|_0 \tag{12}$$

Where $Y = [y_1, y_2, \dots, y_m] \in R^{m\times l}$ and $X = [x_1, x_2, \dots, x_l] \in R^{n\times l}$ are training data and the sparse representation coefficient vector of training data, respectively. $D = [d_1, d_2, \dots, d_n] \in R^{m\times n}$ represents the dictionary. To prevent large numbers in the dictionary and small sparse vectors, the size of atoms is restricted to one ($\|d_i\|_2 \leq 1$). The parameter λ regularizes the ratio between sparsity and minimization error and must beset (λ > 0). Solving L_0 is an NP-hard problem and equation 12 is not convergent (Sun et al. 2019*)*. If we use L_1 instead of L_0 sparsity is guaranteed (Donoho, 2006), and assuming a known **D** or a known **X**, this equation converts to a convergent optimization problem. However; if both **D** and **X** are unknown, this equation is not convergent. In Equation 12 if $m < n$ we say the dictionary is under-complete, and if $m > n$ dictionary is over-complete. In dictionary learning, it is assumed that the dictionary is over-complete. An over-complete dictionary that shows the sparse signal can be a usual transformation such that Fourier and wavelet or be defined and learned in a way that ideally transforms the existing signal to sparse form. Compared to predefined dictionaries, learned dictionaries create sparser representations.

The original data for this study is unavailable, thus we will reconstruct it by interpolating the missing traces. We solve the incompleteness of the data by filling the missing traces with their neighbor traces and calculating $\boldsymbol{\alpha}$ coefficients. Rewriting equation 12 for seismic data $\widehat{\boldsymbol{X}}$, we have:

$$\arg\min_{\mathbf{D},\mathbf{x}_i} \sum_{i=1}^{n} \|\hat{\mathbf{x}}_i - \mathbf{D}\alpha_i\|_F^2 + \lambda\|\alpha_i\|_0 \tag{13}$$

$\widehat{\boldsymbol{X}}$ is the transposition of the filled seismic data. L is the number of samples in the time direction or the length of a seismic trace. Most of the algorithms which solve equation (13) include two parts, dictionary learning and sparse estimation (sparse coding) which execute iteratively. Here, a method such as improved K-SVD can create the dictionary **D** (Shi et al. 2018). We utilized the ROMP, which is a method that has the advantages of BP and greedy methods. Since it finds several coefficients at each iteration, ROMP is also faster than OMP.

Interpolation of seismic data

Suppose that sampling matrix ϕ and sparse dictionary D are correct. In this case, it is necessary that at the sampling points ‖y− ϕ Dα‖2 < ε. Accordingly, we create sparse coefficients as follows:

$$\forall i, \alpha_i = \text{argmin}_y \parallel \text{y} - \phi D\alpha \parallel_2 \quad \text{s.t.} \parallel \alpha \parallel_2 < T, \quad i = 1,2,3 \cdots, L \tag{14}$$

Where T is the maximum number of non-zero elements in Dα. We use ROMP to solve Equation 14. While having sparse coefficients, sparse dictionaries, and sampling matrices, the original data can be reconstructed (**x** = ϕ**D**α). Algorithm 1 shows the interpolation of seismic data using ROMP.

Algorithm 1 Interpolation of seismic data with ROMP

Require: Φ measurement matrix, $\mathbf{Y}$ incomplete seismic section, $\mathbf{y}$ one column of $\mathbf{Y}$, $\mathbf{I}$ Set the initial index set as null, ϵ error threshold.

Residual $\mathbf{r} = \mathbf{y}$

for all rows of $\mathbf{Y}$ Repeat the followings **do**

Calculate observation vector: $\mathbf{u} = \Psi\mathbf{r}$

Select a set $\mathbf{J}$ of the k larger coordinates respect to the magnitude of the $\mathbf{u}$ vector or all of its nonzero coordinates, whichever set is smaller.

Among all subsets $\mathbf{J_0} \subset \mathbf{J}$ with comparable coordinates:

$$\forall i, j : i, j \in \mathbf{J_0} \rightarrow \mid \mathbf{u}(i) \mid \leq 2 \mid \mathbf{u}(j) \mid$$

Choose $\mathbf{J_0}$ with minimal energy

Add the set $\mathbf{J_0}$ to the index set: $\mathbf{I} \cup \mathbf{J_0} \rightarrow \mathbf{I}$,

Solve the least square problem:

$$\text{argmin}_{\alpha_t} \|\mathbf{y} - \Theta\alpha_t\|_2$$

Update the residual: $\mathbf{r} = \mathbf{y} - \Theta\alpha_t$

If size $\mathbf{I} > 2k$ or $\mathbf{r} = 0$ or $iterationnumbers > k$, stop the iterations and $\mathbf{x} = \mathbf{D}\alpha_t$

end for

After calculating all the columns, transpose the data to get the interpolated seismic data.

Calculation of the Sampling Matrix

One of the two operators of Θ in equation (9) is already calculated (Ψ^H = **D**). If we had the sampling matrix Θ, the mutual coherence of the columns of Θ must be minimum (Θ = ϕ**D**). Here we use this property to find a proper sampling matrix. There are several methods to optimize the mutual coherence of a matrix. To generate the sampling matrices ϕ, we use algorithm 2 (Sun et al. 2019):

APPLICATION OF THE PROPOSED METHOD

We apply our proposed method to a synthetic seismogram. Figure 1 shows the seismic section after randomly deleting 40 percent of the traces with its 2D Fourier transom section. Using the proposed algorithm, the missing traces of figure 1 are interpolated. The result is shown in Figure 2a. The existing data is reconstructed with high accuracy. The reconstructed missing traces are highly consistent with their neighboring traces, but there are some differences in the location of large gaps (Figure 2b).

25, 50, and 75 percent of the seismic traces of the studied section are deleted and the incomplete sections are interpolated. To fit with the CS theory, the sparsity is selected to be less than the existing seismic trace. The missing traces are reconstructed accurately because the input data is noiseless. For example, the difference between the original section and the reconstructed section in which 75 percent of seismic traces were deleted (Figure 3a), is shown in figure 3b. Most of the interpolation errors have occurred at the location of the missing traces and large gaps. So far, the seismic section was noiseless which usually is not the case for seismic data. Therefore, a basic criterion for the efficiency of seismic processing methods is their ability to handle different noises. In different stages of the seismic processing workflow, several methods are used for the attenuation of random and coherent noise.

Table 1 SNR of the sections in figure 4. The ratios are in dB.

Signal to Noise Ratio SNR		
Original Sections	Filled Sections	Interpolated Sections
10	3.41	8.2
1	-0.58	0.8
-5	-5.41	-4.7

Since the range of frequency content of the manmade seismic waves is known (Usually between 10 to 100 Hz and even less), filtering the noises using the Fourier transform is popular.

Different ratios of noise are added to the studied seismic section and the proposed algorithm is applied to them. Figure 4a shows this seismic section after adding random noise, in which its signal-to-noise ratio (SNR) is reduced to 10dB, and randomly 40 percent of the seismic traces are deleted. Following our algorithm, to start the dictionary learning, the missing traces are filled with their neighboring traces. Since for the filled traces the location of seismic events could differ from the real locations, replacing seismic traces would add coherent noise and this is the reason the SNR of the filled section is much lower than the actual data (3.41dB as presented in Table 1). Algorithm 3 is applied to the incomplete section and the interpolation is performed (figure 4b). SNR of the reconstructed section is 8.2dB which is much higher than the filled section (Table 1). Although this ratio is less than the ratio of the real data, practically the real seismic section is not available. Therefore, it could be said that interpolation increased the SNR of the existing data. The difference between the real section and the interpolated section shows that the errors are spread over the whole section. In the case of having coherent noise, these events could not be differentiated from seismic traces unless we apply some assumptions and conditions to the results.

According to the Shannon-Nyquist theory, a complete reconstruction requires a signal that has at least twice the maximum sampling frequency in the data. This theory assumes that the signal has a complete range of frequencies. Based on CS, the existing frequencies of the signal are sparse and the signal contains few frequencies. With this assumption, the theory of Shannon-Nyquist does not preclude signal reconstruction and aliasing will not occur. Although the f-k spectrums of the sections in figure 4 show the aliased event in the f-k spectrum, aliasing did not occur in the interpolated results.

The above procedure is repeated for incomplete sections with different SNR ratios while 40 percent of the traces are randomly deleted. Figures 4e and h show the interpolation results of the incomplete sections with SNR of 1 dB and -5 dB respectively. Same as in figure 4b, the missing traces are reconstructed and compared to the filled sections and SNRs are increased (table 1). Although the proposed method reconstructed the missing trace of low SNR sections, its ability reduces when noises are larger. Therefore, in this situation, it is necessary to de-noise the incomplete sections before the interpolation.

We test our method on real seismic data with more complicated structures. A shot-gather with 650 traces is selected. Each trace is sampled in 5 seconds with a time interval of 4 ms. Noisy traces and zero traces do not exist in this record. Some random noise, coherent noise, and ground roll exist in this shot gather. A 10 to 80 Hz band-pass filter was already applied to the data. 50 percent of the trace are randomly removed (figure 6. a). To have a clear plot, every 10 consecutive traces are omitted in the plotting. Empty trace locations are spread randomly in the entire offset and there are large gaps in the incomplete data. Figure 6. b shows the interpolation result of the incomplete shot-gather. Since the original data was filtered and has a relatively high SNR, the original traces are replaced in the interpolated section in the location of available traces in the incomplete data. In the location of missing traces and large gaps, the reconstructed seismic events are highly compatible with their neighbor traces. The frequency spectrum of the interpolated section in figure 6. c shows that some high-frequency noises are added to the data. The missing traces are reconstructed acceptably. As in the circled example areas in figure 6. b, most of the errors and seismic artifacts are related to steep seismic events and large gaps.

Figure 7. a, shows an Inline of the 3D F3 block seismic data. 50 percent of this data is removed (Figure 7. b). The proposed algorithm is used to interpolate the missing data. As in the result in figure 7. c, and the error section in figure 7.d the interpolation introduces some low-amplitude random noises in the data. However, the seismic events are reconstructed. To compare the results to another sparsity-based interpolation algorithm, the missing data is also interpolated using a Fourier-based interpolation method which is discussed in Jahanjooy et al. 2016 (Figure 7.e, and f). Although the result is promising, the Fourier interpolation creates more errors in regions with steep events.

CONCLUSIONS

To use CS, the data must have a sparse representation, which original seismic data do not have. The improved K-SVD is used to solve the problem as well as to generate the over-complete transformation matrix based on the efficacy of dictionary learning techniques. This dictionary generates a sparse data depiction of the seismic data in line with CS. It is optimized so that sampling matrices have little mutual coherence with the dictionary. A sparse data depiction is generated in the sparse domain using a smooth L1 minimizer term as a minimizer term of the matching pursuit method such as ROMP. The entire data is rebuilt using the sparse data representation and the over-complete dictionary as transform matrices. After applying it to different types of seismic data, we find that the suggested approach can interpolate seismic data with a high proportion of missing traces and low SNR. This accuracy declines in the case of complex and dense seismic events. In the transformation domain, seismic data is sparse with only a few non-zero coefficients and the aliasing does not occur in the reconstruction. The accuracy and speed of the improved K-SVD of seismic data could be improved by researching the impact of dictionary redundancy. In this research, the proposed method is applied to 2D seismic data. Studying the efficiency of this method on 3D data is considered for future research.

CONFLICT OF INTEREST: On behalf of all authors, the corresponding author states that there is no conflict of interest.

REFERENCES

Baraniuk, R.G. and Steeghs, P., 2017. Compressive sensing: A new approach to seismic data acquisition. The Leading Edge, 36: 642–645.

Boyd, S. and Vandenberghe, l., 2004. Convex Optimization. Cambridge University Press, New York.

Candes, E.J., Romberg, J. and Tao, T., 2006. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on information theory, 52:489–509.

Candes, E.J., Wakin, M.B. and Boyd, S.P., 2008. Enhancing sparsity by reweighted l 1 minimization. Journal of Fourier analysis and applications, 14:877–905.

Chen, Y., Ma, J. and Fomel, S., 2016. Double-sparsity dictionary for seismic noise attenuation. Geophysics, 81: V103–V116.

Chen, Y., 2017. Fast dictionary learning for noise attenuation of multidimensional seismic data. Geophysical Journal International, 209:21–31.

Donoho, D.L., 2006. Compressed sensing. IEEE Transactions on information theory, 52:1289–1306.

Donoho, D.L., Elad, M. and Temlyakov, V.N., 2005. Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Transactions on information theory, 52:6–18.

Duarte, M.F., and Eldar, Y.C., 2011. Structured compressed sensing: From theory to applications. IEEE Transactions on signal processing, 59:4053–4085.

Elad, M., and Aharon, M., 2006. Image denoising via sparse and redundant representations over learned dictionaries. IEEE Transactions on Image processing, 15:3736–3745.

Gilbert, A.C., Muthukrishnan, S. and Strauss, M., 2005. Improved time bounds for near-optimal sparse Fourier representations. In Wavelets XI, 5914: 398–412. SPIE. doi: 10.1117/12.615931

Herrmann, F.J. and Hennenfent, G., 2008. Non-parametric seismic data recovery with curvelet frames. Geophysical Journal International, 173:233–248.

Herrmann, F.J., Wason, H. and Lin, T., 2011. Compressive sensing in seismic exploration: an outlook on a new paradigm. CSEG Recorder, 36:19–33.

Iwen, M.A., 2007. A deterministic sub-linear time sparse Fourier algorithm via non-adaptive compressed sensing methods. arXiv: 0708.1211v1. https://doi.org/10.48550/arXiv.0708.1211

Jahanjooy, S., Nikrouz, R. and Mohammed, N., 2016. A faster method to reconstruct seismic data using the anti-leakage Fourier transform. Journal of Geophysics and Engineering, 13: 86-95.

Kaur, H., Pham, N. and Fomel, S., 2019. Seismic data interpolation using CycleGAN. In SEG technical program expanded abstracts. 2202-2206. Society of Exploration Geophysicists.

Lan, N.Y., Zhang, F. C. and Yin, X.Y., 2022. Seismic data reconstruction based on low dimensional manifold model, Petroleum Science. 19: 518-533.

Lotfi, M. and Vidyasagar, M., 2018. A fast noniterative algorithm for compressive sensing using binary measurement matrices. IEEE Transactions on Signal Processing, 66: 4079–4089.

Mallat, S.G. and Zhang, Z., 1993. Matching pursuits with time-frequency dictionaries. IEEE Transactions on signal processing, 41:3397–3415.

Meng, F., Yang, X., Zhou, C. and Li, Z., 2017. A sparse dictionary learning-based adaptive patch inpainting method for thick cloud removal from high-spatial resolution remote sensing imagery. Sensors, 17: 2130.

Nazari Siahsar, M. A., Gholtashi, S., Abolghasemi, V., and Chen, Y., 2017a. Simultaneous denoising and interpolation of 2d seismic data using data-driven non-negative dictionary learning. Signal Processing, 141: 309–321.

Nazari Siahsar, M. A., Gholtashi, S., Roshandel Kahoo A., Chen, W.,\ and Chen, Y., 2017b. Data-driven multitasks sparse dictionary learning for noise attenuation of 3d seismic data. Geophysics, 82: V385–V396. https://doi.org/10.1190/geo2017-0084.1

Needell, D. and Tropp, J.A., 2009. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. Applied and computational harmonic analysis, 26: 301–321.

Needell, D. and Vershynin, R., 2010. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. IEEE Journal of selected topics in signal processing, 4: 310–316.

Oguz, I., Zhang, L., Abramoff, M.D. and Sonka, M., 2016. Optimal retinal cyst segmentation from oct images. In Medical Imaging 2016: Image Processing. 9784: 375–381. Computer Science SPIE Medical Imaging.

She, B., Wang, Y., Liu, Z., Cai, H., Liu, W. and Hu. G.M., 2019, Seismic impedance inversion using dictionary learning-based sparse representation and nonlocal similarity. Interpretation, 7: SE51– SE67.

Shi, R., Bing Li, B. and Zhang, J., 2018, Modulated Signal Denoising Algorithm Based on improved K-SVD. IOP Conf. Series: Materials Science and Engineering 452: 1-6. 032064. doi:10.1088/1757-899X/452/3/032064.

Strohmer, T. and Heath, R.W., 2003. Grassmannian frame with applications to coding and communication. Applied and computational harmonic analysis., 14:257–275.

Sun, H.M., Jia, R.S., Zhang, X.L., Peng, Y. J. and Lu. X.M., 2019. Reconstruction of missing seismic traces based on sparse dictionary learning and the optimization of measurement matrices. Journal of Petroleum Science and Engineering, 175:719–727.

Tropp, J.A., and Gilbert, A.C., 2007. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Transactions on information theory, 53: 4655–4666.

Wang, H., Chen, W., Zhang, Q., Liu, X., Zu, S. and Chen, Y., 2020. Fast dictionary learning for high-dimensional seismic reconstruction. IEEE Transactions on Geoscience and Remote Sensing, 59: 7098-7108.

Yu, S., Ma, J., Zhang, X. and Sacchi, M.D., 2015. Interpolation and denoising of high-dimensional seismic data by learning a tight frame. Geophysics, 80: V119–V132.

Zhou, Y., Gao, J., Chen, W. and Frossard, P., 2016. Simultaneous source separation via patchwise sparse representation. IEEE Transactions on Geoscience and Remote Sensing, 54: 5271–5284.

Zu, S., Zhou, H., Wu, R., Mao, W. and Chen, Y., 2018. Hybrid-sparsity constrained dictionary learning for iterative deblending of extremely noisy simultaneous-source data. IEEE transactions on geoscience and remote sensing, 57: 2249–2262.

a) b)

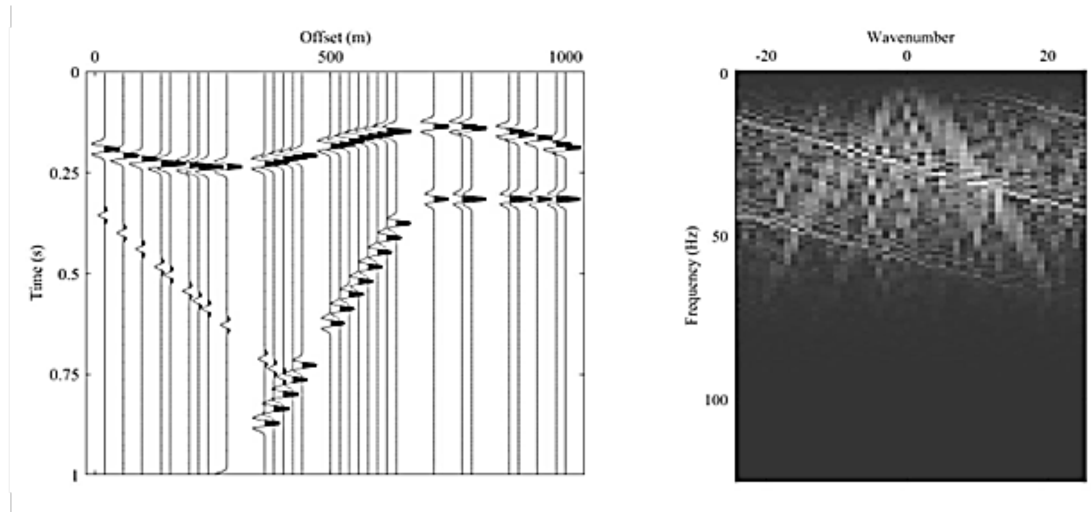

Fig. 1 (a) Seismic section after randomly deleting 40 percent of the traces. (b) the f-k spectrum of (a).

a) b)

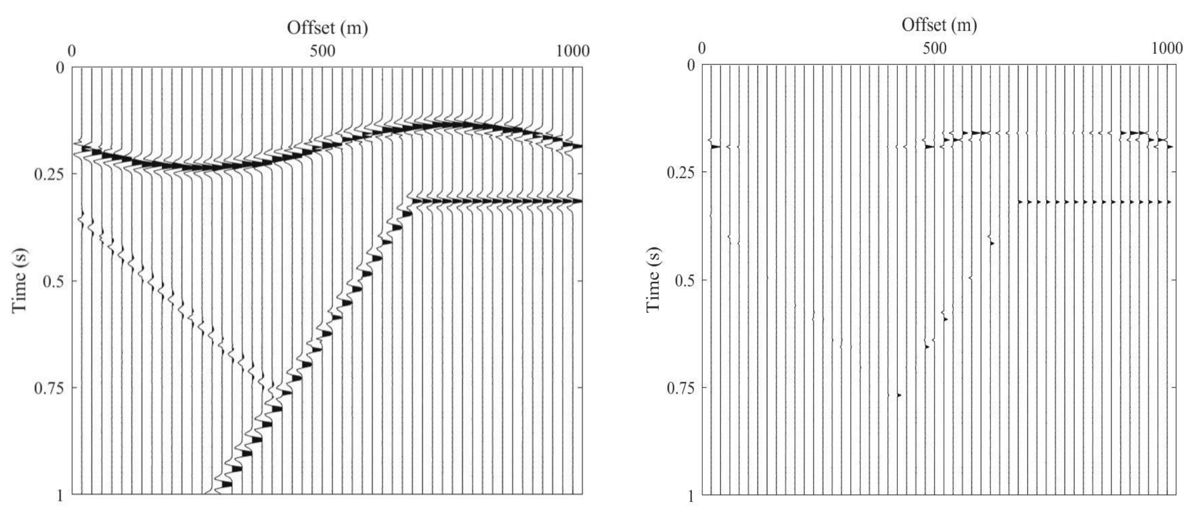

Fig. 2 Interpolation of figure 1a. (a) Interpolation using the proposed algorithm (b) Difference between the original section and the reconstructed one.

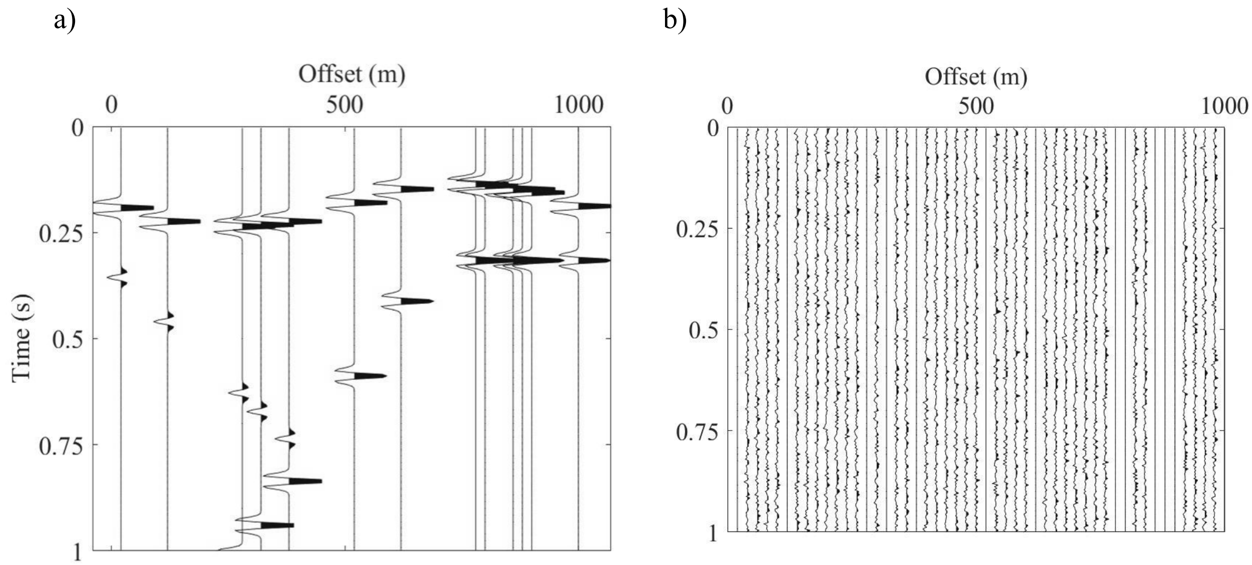

Fig. 3 (a) Seismic section after randomly deleting 75 percent of the traces. (b) Difference between the original section and the interpolated section.

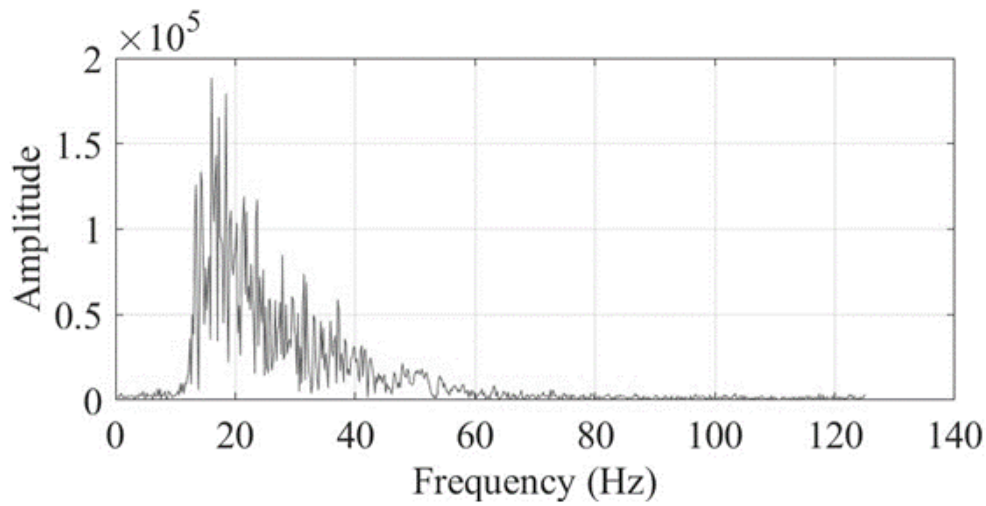

Fig. 4 Reconstruction of a synthetic seismic section. 40 percent of the traces are deleted. a, d, and g: the incomplete sections with SNR 10dB, 1dB, and -10dB respectively. b, e, and h: the interpolated sections related to a, d, and g. c, f, and i: F-K spectrum of the b, e, and h respectively.

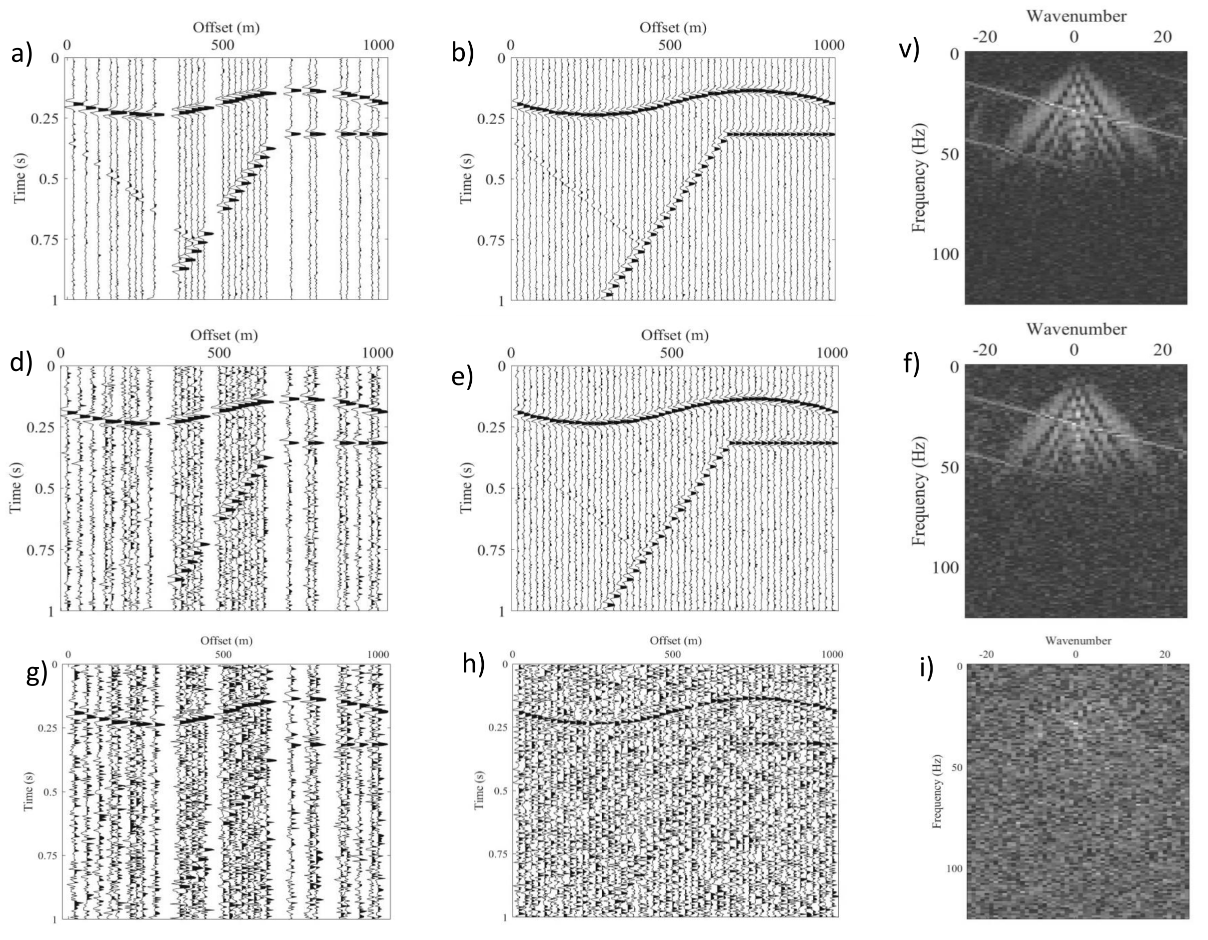

Fig. 5 The frequency spectrum of the original shot-gather.

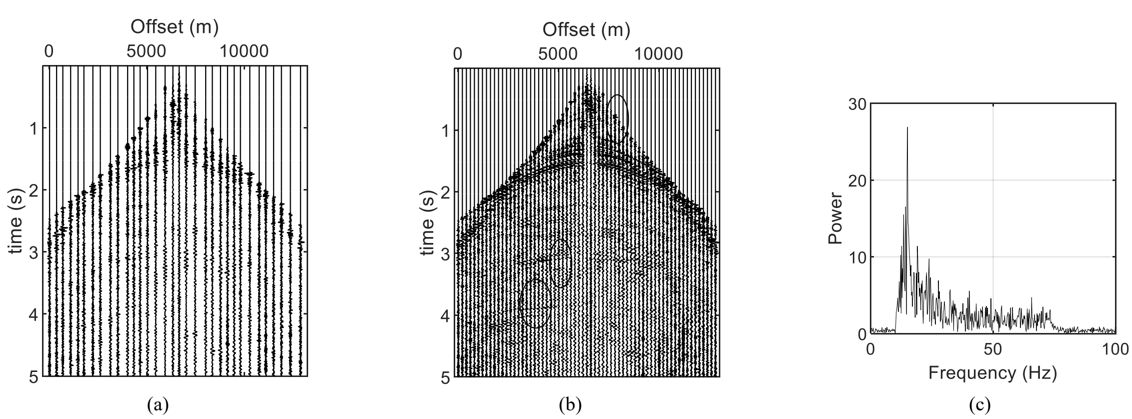

Fig. 6 (a) The shot-gather after randomly deleting 50 percent of the traces. (b) Interpolated shot-gather. (c) The frequency spectrum of b.

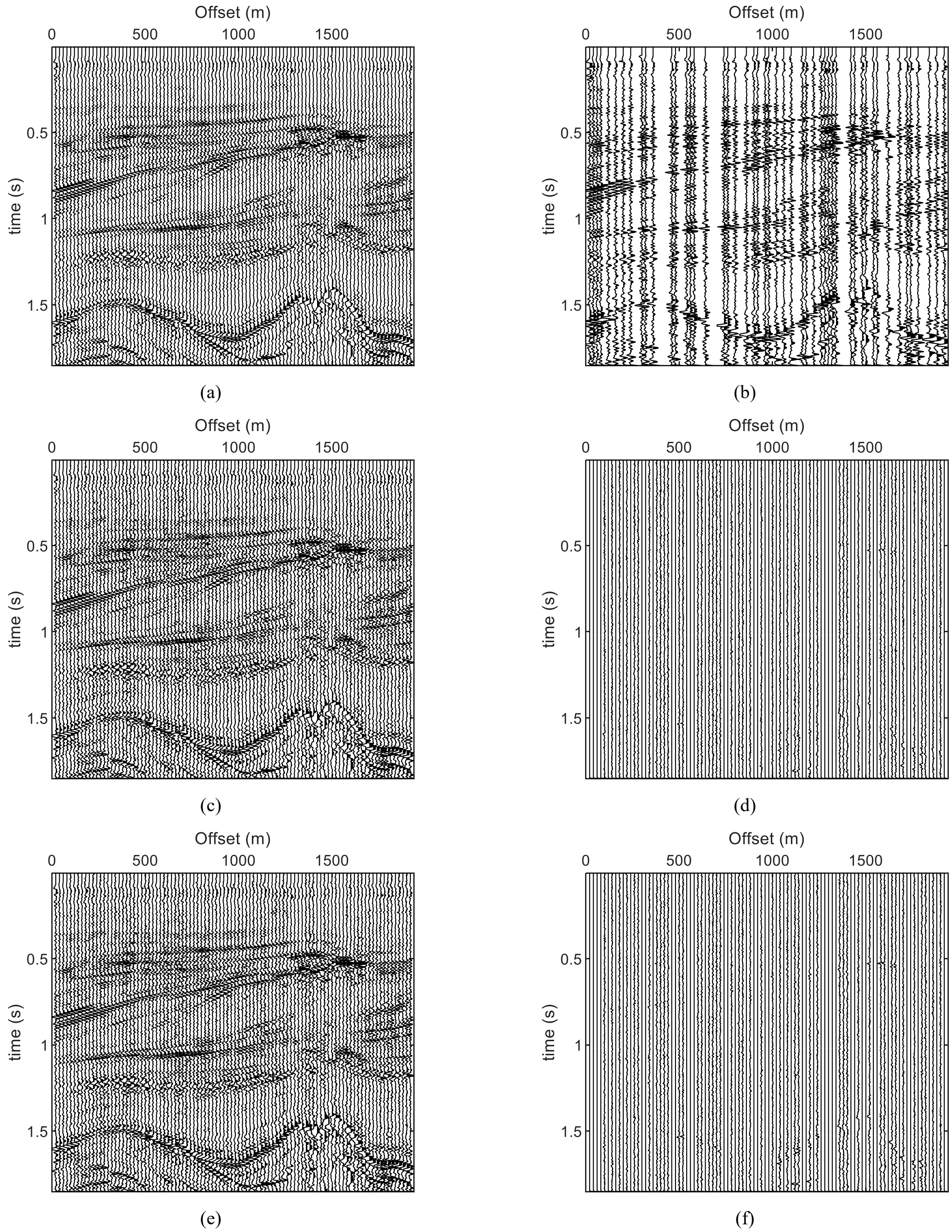

Figure 7.a, shows an Inline of the 3D F3 block seismic data. 50 percent of this data is removed (Figure 7. b). The proposed algorithm is used to interpolate the missing data. As in the result in figure 7. c, and the error section in figure 7.d the interpolation introduces some low-amplitude random noises in the data. However, the seismic events are reconstructed. To compare the results to another sparsity-based interpolation algorithm, the missing data is also interpolated using a Fourier-based interpolation method which is discussed in Jahanjooy et al. 2016 (Figure 7.e, and f).