TENSOR SUPER-RESOLUTION FOR SEISMIC DATA

Songjie Liao[†], *Xiao-Yang Liu*[*], *Feng Qian*[†], *Miao Yin*[†], *Guang-Min Hu*[†]

[†]University of Electronic Science and Technology of China, China
[*]Columbia University, USA

ABSTRACT

In this paper, we propose a novel method for generating high-granularity three-dimensional (3D) seismic data from low-granularity data based on tensor sparse coding, which jointly trains a high-granularity dictionary and a low-granularity dictionary. First, considering the high-dimensional properties of seismic data, we introduce tensor sparse coding to seismic data interpolation. Second, we propose that the dictionary pairs trained by low-granularity seismic data and high-granularity seismic data have the same sparse representation, which are used to recover high-granularity data with the high-granularity dictionary. Finally, experiments on the seismic data of an actual field show that the proposed method effectively perform seismic trace interpolation and can improve the resolution of seismic data imaging.

***Index Terms*—** Tensor super-resolution, seismic data interpolation, tensor sparse coding.

1. INTRODUCTION

Seismic data are generally sparsely sampled in spatial directions, resulting in poor lateral resolution of the seismic record [1]. For high resolution seismic data imaging, seismic trace interpolation is an important task.

There are a considerable number of conventional methods for seismic trace interpolation. Spitz [1] implemented an interpolation method based on predictive filtering in the f-x domain. Claerbout [2] studied the prediction error filtering seismic interpolation method in the t-x domain, and achieved data reconstruction with false frequency. Sacchi [3] implemented five-dimensional interpolation under the idea of inversion. Recently, Jia [4] proposed a machine learning interpolation method for seismic data. However, these methods ignore structural characteristics of three-dimensional (3D) seismic data.

Super-resolution (SR) is a classic issue in image signal processing, and there are promising methods based on dictionary learning to generate high-resolution data [5, 6, 7]. Recently, Christian [8] proposed to use generative adversarial networks (GAN) for image SR. Considering that formation parameters have lateral continuity and seismic data located in different traces have similarities, a sparse dictionary can be extracted from seismic data [9], and namely the seismic data have sparse representation. Compressed sensing theory has been used for seismic data [10]. However, these proposed SR methods convert data into vectors which ignore spatial correlation of multi-dimensional data [11, 12], which carry rich geophysics information [13] that can help us obtain higher resolution.

In order to capture spatial correlation of the multi-dimensional data, we propose a seismic trace interpolation method based on tensor sparse coding [14, 15] for 3D seismic data, which trains a high-granularity and low-granularity dictionary pair. First, we introduce tensor sparse coding for seismic data. Second, since structural features are approximately revealed in the collected low-granularity seismic data and high-granularity seismic data, we propose that the data pairs have the same sparse representation, and the sparse coefficients of the low-granularity data are combined with the high-granularity dictionary to recover the high-granularity data. Finally, experiments on the seismic data of an actual field show that the proposed method promotes the resolution of seismic data imaging.

The remainder of this paper is organized as follows: Section 2 introduces the notations and preliminary used in our paper. The system model and problem formulation are given in Section 3. Section 4 presents the solution algorithm and generates the high-granularity seismic data of an actual field in Section 5.

2. NOTATION AND PRELIMINARY

A third-order tensor is denoted as $\mathcal{X} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$. The expansion of $\mathcal{X}$ along the third dimension is represented as $\underline{\mathcal{X}} = \left[\mathcal{X}^{(1)}; \mathcal{X}^{(2)}; \cdots; \mathcal{X}^{(N_3)}\right]$. The transpose of tensor $\mathcal{X}$ is denoted as $\mathcal{X}^\dagger$, where ${\mathcal{X}^\dagger}^{(1)} = {\mathcal{X}^{(1)}}^T$, and ${\mathcal{X}^\dagger}^{(k)} = {\mathcal{X}^{(N_3+2-k)}}^T$, $2 \le k \le N_3$, and the superscript "^{T}" represents the transpose of matrices. The discrete Fourier transform (DFT) along the third dimension of $\mathcal{X}$ is denoted as $\widetilde{\mathcal{X}}$. Let $[N]$ denote the set $\{1, 2, \cdots, N\}$.

The Frobenius and ℓ_1-norm norms of tensors are denoted as $\|\mathcal{X}\|_F = \left(\sum_{i,j,k} |\mathcal{X}(i,j,k)|^2\right)^{(1/2)}$, and $\|\mathcal{X}\|_1 =$

This work was supported by the Natural Science Foundation of China (41874155 and U1562218) and the Fundamental Research Funds for the Central Universities (ZYGX2016J219).

$\sum_{i,j,k} |\mathcal{X}(i,j,k)|$, respectively.

Furthermore, we introduce the following definitions.

Definition 1. *(t-product) [16] The tensor-product between $\mathcal{D} \in \mathbb{R}^{N_1 \times r \times N_3}$ and $\mathcal{Z} \in \mathbb{R}^{r \times N_2 \times N_3}$ is a tensor $\mathcal{X} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$, where $\mathcal{X}(i,j,:) = \sum_{q=1}^{r} \mathcal{D}(i,q,:) * \mathcal{Z}(q,j,:)$, and $*$ denotes the circular convolution operation.*

Remark 1. *The t-product $\mathcal{X} = \mathcal{D} * \mathcal{Z}$ can be efficiently computed in the frequency domain as:*

$$\widetilde{\mathcal{X}}^{(k)} = \widetilde{\mathcal{D}}^{(k)} \widetilde{\mathcal{Z}}^{(k)},\ k \in [N_3]\,. \tag{1}$$

Lemma 1. *[16] The t-product $\mathcal{X} = \mathcal{D} * \mathcal{Z}$ has an equivalent matrix-multiplication form as follows:*

$$\underline{\mathcal{X}} = \underline{\mathcal{D}}^c \underline{\mathcal{Z}}, \tag{2}$$

where $\underline{\mathcal{D}}^c$ is the circular matrix of $\mathcal{D}$ defined as follows:

$$\underline{\mathcal{D}}^c = \begin{pmatrix} \mathcal{D}^{(1)} & \mathcal{D}^{(N_3)} & \cdots & \mathcal{D}^{(2)} \\ \mathcal{D}^{(2)} & \mathcal{D}^{(1)} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \mathcal{D}^{(N_3)} \\ \mathcal{D}^{(N_3)} & \mathcal{D}^{(N_3-1)} & \cdots & \mathcal{D}^{(1)} \end{pmatrix}. \tag{3}$$

3. PROBLEM STATEMENT

Instead of transforming data into vectors, we represent 3D seismic data with a tensor $\mathcal{X} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ to train a tensor dictionary. The tensor dictionary learning model is as follows:

$$\begin{aligned} \min_{\mathcal{D},\mathcal{Z}} \ & \frac{1}{2} \|\mathcal{X} - \mathcal{D} * \mathcal{Z}\|_F^2 + \lambda' \|\mathcal{Z}\|_1\,, \\ s.t.\ & \|\mathcal{D}(:,j,:)\|_F^2 \le 1,\ j \in [r]\,, \end{aligned} \tag{4}$$

where $\mathcal{D} \in \mathbb{R}^{N_1 \times r \times N_3}$ is the tensor dictionary, and each lateral slice $\mathcal{D}(:,j,:)$ is a basis, while $\mathcal{Z} \in \mathbb{R}^{r \times N_2 \times N_3}$ represent the tensor coefficients. The parameter λ' balances the approximation error and the sparsity of tensor coefficients, and r is the number of atoms.

The deposition of the underground layers is often relatively stable, so the seismic responses of the same field usually have a certain degree of consistency and lateral continuity [9]. The observed low-granularity seismic data $\mathcal{Y}_l$ and high-granularity seismic data $\mathcal{X}_h$ from the same field are sparse and approximate. Based on this prior, we jointly train two dictionaries $\mathcal{D}_h$ and $\mathcal{D}_l$ by $\mathcal{Y}_l$ and $\mathcal{X}_h$, named high-granularity dictionary and low-granularity dictionary, respectively. $\mathcal{D}_h$ and $\mathcal{D}_l$ enforce $\mathcal{X}_h$ and $\mathcal{Y}_l$ to share the same sparse representation $\mathcal{Z}$. The joint dictionaries model is as follows:

$$\mathcal{D}_h = \underset{\mathcal{D}_h,\mathcal{Z}}{\operatorname{argmin}} \|\mathcal{X}_h - \mathcal{D}_h * \mathcal{Z}\|_F^2 + \lambda' \|\mathcal{Z}\|_1\,, \tag{5}$$

$$\mathcal{D}_l = \underset{\mathcal{D}_l,\mathcal{Z}}{\operatorname{argmin}} \|\mathcal{Y}_l - \mathcal{D}_l * \mathcal{Z}\|_F^2 + \lambda' \|\mathcal{Z}\|_1\,, \tag{6}$$

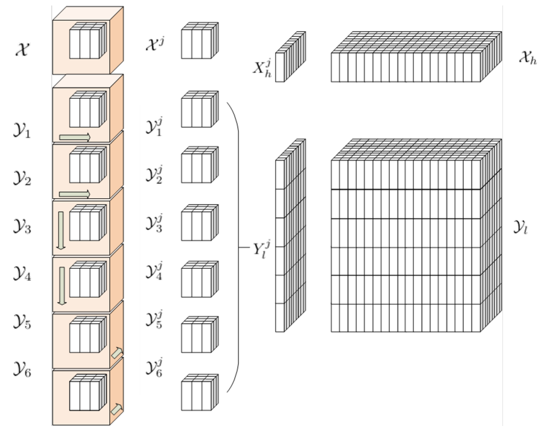


Fig. 1. Constructing low- and high-granularity seismic data training pairs.

where $\mathcal{D}_h$ is the high-granularity dictionary and $\mathcal{D}_l$ is the low-granularity dictionary. $\mathcal{Y}_l$ and $\mathcal{X}_h$ are the low-granularity and high-granularity seismic data, which are constructed via the following method.

Combining (5) and (6) to force the training pairs sharing the same sparse representation , we get:

$$\begin{aligned} \min_{\mathcal{D}_h,\mathcal{D}_l,\mathcal{Z}} \frac{1}{P} \|\mathcal{X}_h - \mathcal{D}_h * \mathcal{Z}\|_F^2 + \frac{1}{Q} \|\mathcal{Y}_l - \mathcal{D}_l * \mathcal{Z}\|_F^2 \\ + \lambda'(\frac{1}{P} + \frac{1}{Q}) \|\mathcal{Z}\|_1\,, \end{aligned} \tag{7}$$

where P and Q represent the numbers of the first dimension of $\mathcal{X}_h$ and $\mathcal{Y}_l$. $1/P$ and $1/Q$ balance the two cost terms of (5) and (6).

Equation (7) can be converted to the following form:

$$\min_{\mathcal{D},\mathcal{Z}} \|\mathcal{X} - \mathcal{D} * \mathcal{Z}\|_F^2 + \lambda \|\mathcal{Z}\|_1\,, \tag{8}$$

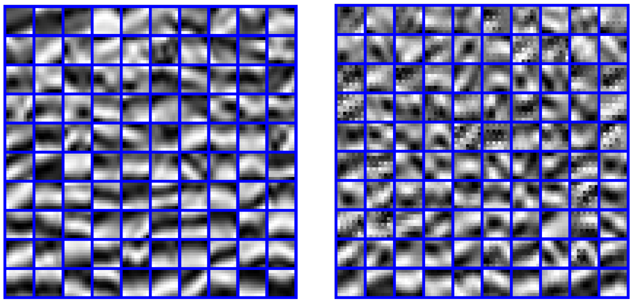
where

$$\mathcal{X} = \begin{bmatrix} \frac{1}{\sqrt{P}} \mathcal{X}_h \\ \frac{1}{\sqrt{Q}} \mathcal{Y}_l \end{bmatrix},\ \mathcal{D} = \begin{bmatrix} \frac{1}{\sqrt{P}} \mathcal{D}_h \\ \frac{1}{\sqrt{Q}} \mathcal{D}_l \end{bmatrix},\ \lambda = \lambda'(\frac{1}{P} + \frac{1}{Q}). \tag{9}$$

The problem is solved efficiently by alternately optimizing $\mathcal{D}$ and $\mathcal{Z}$ directly in tensor space, then (8) can be decomposed into the following two sub-problems: (i) solving the tensor sparse coefficients, which is solved by the iterative shrinkage threshold algorithm based on the tensor product; (ii) updating the tensor dictionary, which is solved by Lagrange dual method. The solution process is detailed in Section 4.

4. OUR SOLUTION

In order to solve problem (8), we alternately optimized the tensor dictionary $\mathcal{D}$ and the coefficients $\mathcal{Z}$ while fixing the other one.



(a) (b)

Fig. 2. Dictionaries learned from high-granularity and low-granularity seismic data. (a) is part of the high-granularity tensor dictionary with each atom of size 9×9, and (b) is the low-granularity tensor dictionary.

4.1. Tensor coefficients learning

In order to solve the tensor sparse coefficients, we fix $\mathcal{D} \in \mathbb{R}^{N_1 \times r \times N_3}$ and use the tensor-based iterative shrinkage threshold algorithm [14] to solve the tensor sparse representations $\mathcal{Z} \in \mathbb{R}^{r \times N_2 \times N_3}$ of seismic data $\mathcal{X} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$. Problem (8) is converted to the following problem:

$$\min_{\mathcal{Z} \in \mathbb{R}^{r \times N_2 \times N_3}} \frac{1}{2} \|\mathcal{X} - \mathcal{D} * \mathcal{Z}\|_F^2 + \lambda \|\mathcal{Z}\|_1 . \tag{10}$$

For convenience of the computation, (10) can be transformed to matrix-product by Lemma 1:

$$\min_{\underline{\mathcal{Z}} \in \mathbb{R}^{rN_2 \times N_3}} \frac{1}{2} \|\underline{\mathcal{X}} - \underline{\mathcal{D}}^c \underline{\mathcal{Z}}\|_F^2 + \lambda \|\underline{\mathcal{Z}}\|_1 . \tag{11}$$

We apply the iterative shrinkage thresholding algorithm based on tensor-product (ISTA-T) [14] to solve (10) directly. We first rewrite (10) as:

$$\min_{\mathcal{Z}} f(\mathcal{Z}) + \lambda g(\mathcal{Z}), \tag{12}$$

where $f(\mathcal{Z})$ stands for $\frac{1}{2} \|\mathcal{X} - \mathcal{D} * \mathcal{Z}\|_F^2$ which is continuous with the Lipschitz constant $L = \left\|\underline{(\mathcal{D}^\dagger * \mathcal{D})}^c\right\|_F^2$ [14], namely $L = \sum_{k=1}^{N_3} \left\|\widetilde{\mathcal{D}}^{(k)^H} \widetilde{\mathcal{D}}^{(k)}\right\|_F^2$, and $g(\mathcal{Z})$ stands for the sparsity constraint term $\|\mathcal{Z}\|_1$. $\mathcal{Z}$ is solved by iteration, and (12) can be rewritten as a linearized function around the previous estimation $\mathcal{Z}_p$ with the proximal regularization and the nonsmooth regularization. Thus, at the $(p+1)$-th iteration, $\mathcal{Z}_{p+1}$ is updated by

$$\begin{aligned} \mathcal{Z}_{p+1} = \underset{\mathcal{Z}}{\operatorname{argmin}}\, & f(\mathcal{Z}_p) + \langle \nabla f(\mathcal{Z}_p), \mathcal{Z} - \mathcal{Z}_p \rangle \\ & + \frac{L}{2} \|\mathcal{Z} - \mathcal{Z}_p\|_F^2 + \lambda g(\mathcal{Z}), \end{aligned} \tag{13}$$

where $\nabla f(\mathcal{Z})$ is the gradient defined in tensor space. Then (13) is equivalent to

$$\begin{aligned} \mathcal{Z}_{p+1} = \underset{\mathcal{Z}}{\operatorname{argmin}}\, & \frac{1}{2} \left\| \mathcal{Z} - (\mathcal{Z}_p - \frac{1}{L} \nabla f(\mathcal{Z}_p)) \right\|_F^2 \\ & + \frac{\lambda}{L} \|\mathcal{Z}\|_1 , \end{aligned} \tag{14}$$

Lastly, (14) can be solved by the proximal operator $\mathbf{Prox}_\beta(\mathcal{Z}_p - \frac{1}{L}\nabla f(\mathcal{Z}_p))$, where $\mathbf{Prox}_\beta$ is the soft-thresholding operator $\mathbf{Prox}_\tau(\cdot) \to \text{sign}(\cdot)\text{max}(|\cdot| - \tau, 0)$ [17]. To speed up the ISTA-T, an extrapolation operator was adopted [18].

4.2. Tensor dictionary learning

For learning the dictionary $\mathcal{D}$ while $\mathcal{Z}$ is fixed, the optimization problem is:

$$\begin{aligned} & \min_{\mathcal{D} \in \mathbb{R}^{N_1 \times r \times N_3}} \frac{1}{2} \|\mathcal{X} - \mathcal{D} * \mathcal{Z}\|_F^2 , \\ & s.t.\ \|\mathcal{D}(:, j, :)\|_F^2 \leq 1,\ [j] \in r, \end{aligned} \tag{15}$$

where atoms are coupled together due to the $*$ operation. Therefore, we transform (15) into the frequency domain by DTF which is decomposed into k nearly-independent problems (that are coupled only through the norm constraint) as follows:

$$\begin{aligned} & \min_{\widetilde{\mathcal{D}}^{(k)}, k \in [N_3]} \sum_{k=1}^{N_3} \left\| \widetilde{\mathcal{X}}^{(k)} - \widetilde{\mathcal{D}}^{(k)} \widetilde{\mathcal{Z}}^{(k)} \right\|_F^2 , \\ & s.t.\ \sum_{k=1}^{N_3} \left\| \widetilde{\mathcal{D}}^{(k)}(:, j) \right\|_F^2 \leq N_3,\ [j] \in r. \end{aligned} \tag{16}$$

Then, we adopted the Lagrange dual [19] method for solving (16) in frequency domain. First, we transform (16) into the following Lagrangian:

$$\begin{aligned} \mathcal{L}(\widetilde{\mathcal{D}}, \Lambda) = & \sum_{k=1}^{N_3} \left\| \widetilde{\mathcal{X}}^{(k)} - \widetilde{\mathcal{D}}^{(k)} \widetilde{\mathcal{Z}}^{(k)} \right\|_F^2 + \\ & \sum_{j=1}^{r} \lambda_j \left(\sum_{k=1}^{N_3} \left\| \widetilde{\mathcal{D}}^{(k)}(:, j) \right\|_F^2 - N_3 \right), \end{aligned} \tag{17}$$

where $\lambda_j \geq 0,\ j \in [r]$ is a dual variable, and $\Lambda = \text{diag}(\lambda)$.

Second, minimizing over $\widetilde{\mathcal{D}}$ analytically, we obtain the optimal formulation of $\widetilde{\mathcal{D}}$:

$$\widetilde{\mathcal{D}}^{(k)} = \left(\widetilde{\mathcal{X}}^{(k)} \widetilde{\mathcal{Z}}^{(k)^H} \right) \left(\widetilde{\mathcal{Z}}^{(k)} \widetilde{\mathcal{Z}}^{(k)^H} + \Lambda \right)^{-1},\ k \in [N_3]. \tag{18}$$

Substituting (18) into the Lagrangian $\mathcal{L}(\widehat{\mathcal{D}}, \Lambda)$, we obtain the Lagrange dual function $\mathcal{D}(\Lambda)$:

$$\mathcal{D}(\Lambda) = -\sum_{k=1}^{N_3} \mathbf{Tr} \left(\widetilde{\mathcal{D}}^{(k)^H} \widetilde{\mathcal{X}}^{(k)} \widetilde{\mathcal{Z}}^{(k)^H} \right) - N_3 \sum_{j=1}^{r} \lambda_j , \tag{19}$$

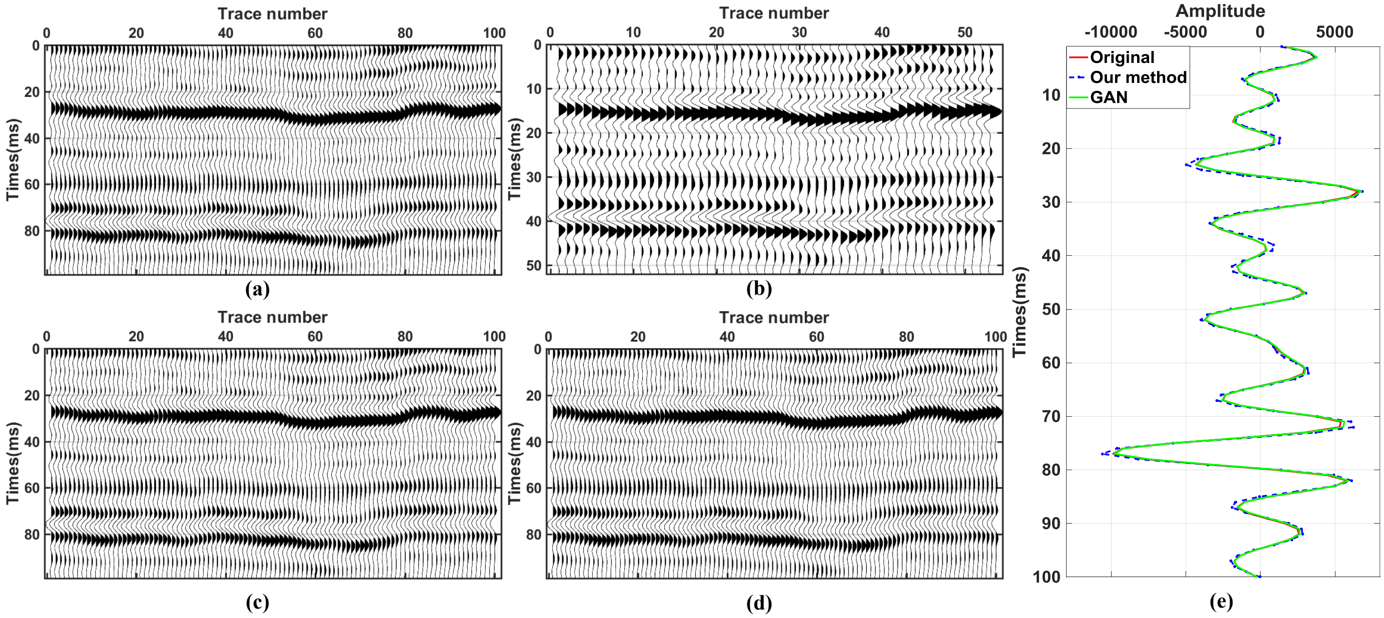


Fig. 3. (a) Original high-granularity seismic data. (b) Low-granularity seismic data. (c) Generated high-granularity seismic data by our method, and (d) generated by GAN. (e) Single trace seismic data reconstructed by our method and GAN.

which was solved by Newton's method. Once the dual variables was obtained, the dictionary can be recovered by Equation (18).

5. EVALUATION

In this section, we apply the proposed method to the seismic data of an actual field.

For the low-granularity seismic data to be reconstructed, collect a high-granularity 3D seismic data set $\mathcal{X} \in \mathbb{R}^{100\times20\times100}$ in the same field. Then, one-half sampling is applied to get the low-granularity seismic data $\mathcal{Y} \in \mathbb{R}^{50\times20\times50}$. For the low-granularity seismic data set, a feature extraction operator f_r is used to extract the gradient and curvature features in three directions of $\mathcal{Y}$, where $r = 1, 2, \cdots, 6$. With the convolution of f_r and $\mathcal{Y}$ along 6 directions, we obtain the low-granularity training data $\mathcal{Y}_l$. As the low- and high-granularity data sets are divided into 10000 cubes size of $\mathbb{R}^{3\times3\times3}$, respectively. And those cubes can be reshaped into $\mathcal{X}_h \in \mathbb{R}^{9\times10000\times81}$. Then, $\mathcal{Y}_l \in \mathbb{R}^{54\times10000\times81}$ can be constructed in the same way. The specific sample construction method is shown in Fig. 1.

Then, we trained a high-granularity dictionary $\mathcal{D}_h \in \mathbb{R}^{9\times512\times81}$ with the high-granularity seismic data $\mathcal{X}_h$ and a low-granularity dictionary $\mathcal{D}_l \in \mathbb{R}^{54\times512\times81}$ with the low-granularity seismic data $\mathcal{Y}_l$ which formed by extracting 6 features, showing in Fig. 2. The structures of two dictionary are very similar, except that the high-granularity dictionary has more concise features.

Seismic data were from the actual field to verify the proposed seismic trace interpolation algorithm. We selected a section to display and showed the interpolation results. Figure. 3(a) is the low-granularity seismic data, which are sampled at interval in space and time directions from the original high-granularity seismic data, which are shown in Fig. 3(b). Figure. 3(c) shows the result of the generated data by our method. In comparison, the high-resolution seismic data are generated using a GAN [8] generator. The generated data can reveal a weak reflector, which is important for seismic interpretation. The single trace waveform reconstruction by our method and the GAN [8] are shown in Figure. 3(e), where both results are shown reconstruction errors for high frequency parts.

6. CONCLUSION

In this paper, we proposed a tensor joint sparse coding method and applied it to 3D seismic trace interpolation. This method extends two-dimensional dictionary learning to three dimensions via tensor product, which effectively utilizes the spatial information of seismic data. Then, tensor sparse coefficients and tensor dictionary are alternately optimized by ISTA-T and Lagrange dual method. The ISTA-T is used to solve the tensor sparse coefficients, and the Lagrangian dual method is used to solve the tensor dictionary. Finally, experiments on the seismic data of the actual field showed that the proposed method could effectively perform seismic trace reconstruction and improve the resolution of seismic data.

References

[1] Simon Spitz, "Seismic trace interpolation in the fx domain," *Geophysics*, vol. 56, no. 6, pp. 785–794, 1991.

[2] JF Claerbout and DAVE Nichols, "Interpolation beyond aliasing by (t, x)-domain pefs," in *53rd EAEG Meeting*, 1991.

[3] Mauricio D Sacchi*, Jianjun Gao, Aaron Stanton, and Jinkun Cheng, "Tensor factorization and its application to multidimensional seismic data recovery," in *SEG Technical Program Expanded Abstracts 2015*, pp. 4827–4831. Society of Exploration Geophysicists, 2015.

[4] Yongna Jia and Jianwei Ma, "What can machine learning do for seismic data processing? an interpolation application," *Geophysics*, vol. 82, no. 3, pp. V163–V177, 2017.

[5] Loukianos Spyrou and Saeid Sanei, "Coupled dictionary learning for multimodal data: An application to concurrent intracranial and scalp eeg," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2349–2353.

[6] Jun-Jie Huang and Pier Luigi Dragotti, "A deep dictionary model for image super-resolution," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 6777–6781.

[7] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma, "Image super-resolution via sparse representation," *IEEE transactions on image processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[8] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew P Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al., "Photo-realistic single image super-resolution using a generative adversarial network.," in *CVPR*, 2017, vol. 2, p. 4.

[9] Bin She, Yaojun Wang, Jiandong Liang, Zhining Liu, Chengyun Song, and Guangmin Hu, "A data-driven avo inversion method via learned dictionaries and sparse representation," *Geophysics*, vol. 83, no. 6, pp. 1–91, 2018.

[10] Felix J Herrmann and Gilles Hennenfent, "Non-parametric seismic data recovery with curvelet frames," *Geophysical Journal of the Royal Astronomical Society*, vol. 173, no. 1, pp. 233–248, 2010.

[11] Xiao-Yang Liu and Xiaodong Wang, "Ls-decomposition for robust recovery of sensory big data," *IEEE Transactions on Big Data*, 2017.

[12] Zhen Wang, Zhiling Long, and Ghassan AlRegib, "Tensor-based subspace learning for tracking salt-dome boundaries constrained by seismic attributes," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1125–1129.

[13] Feng Qian, Miao Yin, Xiao-Yang Liu, Yao-Jun Wang, Cai Lu, and Guang-Min Hu, "Unsupervised seismic facies analysis via deep convolutional autoencoders," *Geophysics*, vol. 83, no. 3, pp. A39–A43, 2018.

[14] Fei Jiang, Xiao-Yang Liu, Hongtao Lu, and Ruimin Shen, "Efficient two-dimensional sparse coding using tensor-linear combination," *AAAI*, 2018.

[15] Xiao-Yang Liu, Shuchin Aeron, Vaneet Aggarwal, and Xiaodong Wang, "Low-tubal-rank tensor completion using alternating minimization," *(Major revision) IEEE Transactions on Information Theory*, 2018.

[16] Misha E Kilmer, Karen Braman, Ning Hao, and Randy C Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 1, pp. 148–172, 2013.

[17] Neal Parikh, Stephen Boyd, et al., "Proximal algorithms," *Foundations and Trends® in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[18] Yangyang Xu and Wotao Yin, "A fast patch-dictionary method for whole image recovery," *arXiv preprint arXiv:1408.3740*, 2014.

[19] Honglak Lee, Alexis Battle, Rajat Raina, and Andrew Y. Ng, "Efficient sparse coding algorithms," in *International Conference on Neural Information Processing Systems (NIPS)*, 2006, pp. 801–808.