Seismic Data Denoising and Interpolation Using Deep Learning

Introduction

Noise removal and interpolation play important roles in processing seismic data, and have been extensively explored. Typical works include Cadzow filter, fx prediction filtering, and sparsity based method. In recent years, various deep learning methods have been widely applied for improving seismic data quality. Zhang et al. (2020) integrated the CNN denoisers trained by natural images into the project-onto-convex-set (POCS) framework. Pham et al. (2020) generated complex inputs for denoising with U-Net architecture. Zhu et al. (2019) developed DeepDenoiser for denoising and decomposition. Fu et al. (2018) introduced a normalized weighted Gaussian filter to reconstruct seismic data with missing traces. Siahsar et al. (2017) used data-driven dictionary learning (DDL) method for sparse representation. Based on double-sparsity dictionary learning, Zhu et al. (2017) performed interpolation and denoising simultaneously. Si and Yuan (2018) based on residual learning of denoising convolutional neural network (DnCNN) for random noise attenuation.

In this work, we propose a so-called CBD-RDN network that simultaneously removes the noise and improves the data resolution. Specifically, we generate the training data by carefully choosing the acquisition geometry and data augmentation, which shows good performance in experiments.

Method

Let $d_{\text{obs}}$ be the observed data, $P$ be the downsampling operator and $n$ be noise, the goal of this work is to recover high resolution and noise-free data $d_c$ from $d_{\text{obs}} = Pd_c + n$ by using deep learning methods. By training the deep neural network in a carefully designed dataset, our method can output high quality $d_c$. In the following context, we give the details of network architecture and training data preparation.

Network architecture. The proposed CBD-RDN network combines residual dense network (RDN) (Zhang et al., 2018) and convolutional blind denoising network (CBDNet) (Guo et al., 2019) with an additional interpolation module and a modified noise estimation module. The detailed architecture is present in Figure 1 that contains a noise estimation block, residual dense block (bottom left) and interpolation block (bottom right). There are 64 filters of size $3 \times 3$ in each convolutional block.

![Figure 1 The proposed CBD-RDN network.](image1)

The construction of training data. Nine segments from the Marmousi model are selected for constructing the training pairs, and two different segments are chosen generating validation data. For each segment, we generate the seismograms using Devito (Louboutin et al., 2019) with 9 different shot locations evenly spaced on top and 100 receivers with a 40m interval. The source is set as a Ricker wavelet with 25 Hz central frequency, and we collect 1000 temporal sample points with

![Figure 2 The training (black), validation (red) and testing (yellow) velocity models.](image2)
sampling interval 1 ms. Moreover, we divide each simulated seismogram into 40 patches with size 40 × 40, thus there are $M = 9 \times 9 \times 40 = 3240$ training pairs in total. The chosen segments are shown in Figure 2.

**Loss function.** Let $F$ be the proposed CBD-RDN network with parameters $\theta$ and $\{(d^i_{obs}, d^i_c)\}_{i=1}^N$ be the training pairs, we determine $\theta$ by minimizing a loss function that consists of reconstruction loss $\mathcal{L}_{rec}$, asymmetric loss $\mathcal{L}_{asymm}$ and the total variation (TV) regularizer $\mathcal{L}_{TV}$. More specifically, each term in our loss is defined as

$$
\mathcal{L}_{rec} = \frac{1}{N} \sum_{i=1}^{N} \| F(d^i_{obs}, \theta), d^i_c \|_2^2,
$$

$$
\mathcal{L}_{TV} = \| \nabla_h \hat{\sigma}(d_{obs}) \|_2^2 + \| \nabla_v \hat{\sigma}(d_{obs}) \|_2^2,
$$

$$
\mathcal{L}_{asymm} = \sum_{i} \left| \alpha - I(\hat{\sigma}(d^i_{obs})-\sigma(d^i_{obs}) < 0) \right| \cdot (\hat{\sigma}(d^i_{obs}) - \sigma(d^i_{obs}))^2,
$$

where $\hat{\sigma}(d^i_{obs})$ and $\sigma(d^i_{obs})$ is the $i$-th estimated and true noise levels respectively, $I$ is a characteristic function, $\alpha$ is set in $(0, 0.5)$, $\nabla_h$ and $\nabla_v$ denotes the gradient operator along the horizontal and vertical direction respectively. The total loss function

$$
\mathcal{L} = \mathcal{L}_{rec} + \lambda_{asymm} \mathcal{L}_{asymm} + \lambda_{TV} \mathcal{L}_{TV},
$$

where $\lambda_{asymm}$ and $\lambda_{TV}$ are trade-off parameters among three loss functions.

**Two observations in training data preparation.** The performance of deep learning methods heavily depends on the quality and the diversity of training samples. In our experiments, we observe two components that are important for the reconstruction task. One is the consistency of frequency bands and acquisition geometry: the spectra that the synthetic and field data have are primarily overlapped; the frequency bands and acquisition geometry of the training set are close to that of the field data. The other observation is related to seismic data augmentation, which is a crucial pre-processing step in deep learning. Instead of using classical techniques in natural image processing, we add an opposite copy of other observation is related to seismic data augmentation, which is a crucial pre-processing step in deep learning.

**Synthetic experiments**

We test the performance of the proposed CBD-RDN networks for coherent noise removal and 4× interpolation problems using one segment chosen from the Marmousi model (see the yellow part in Figure 2) We test our network in the four tasks. **Task 1: coherent noise removal.** The noise $n$ is generated from some unexpected sources where we only consider direct waves here. **Task 2: random missing trace interpolation.** 30% traces are randomly removed in the test seismogram. **Task 3: 4× regular missing trace interpolation.** We regularly select 1 trace in every 4 traces. **Task 4: 4× regular missing trace interpolation + random missing trace interpolation + random noise removal.** We select the missing traces as Task 2 and Task 3, and add 3% Gaussian noise to the seismogram. The results are shown in Figure 3, the SNR are given in Table 1, and the data spectra in Task 1 are given in Figure 4. It is shown that our CBD-RDN method significantly outperforms the traditional curvelet based methods in terms of SNR and visual reconstruction quality. As the spectra of Task 1 show, the amplitude of curvelets result has changed, while it would behave much worse when the amplitude remains. Moreover, different from the curvelet based method which is labor-intensive and of high computational cost, the deep learning based method can automatically estimate the noise level and predict the results.

<table>
<thead>
<tr>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>noisy data</td>
<td>5.67/6.76</td>
<td>6.02/4.85</td>
<td>1.24/1.26</td>
</tr>
<tr>
<td>curvelets</td>
<td>12.81/12.72</td>
<td>16.68/7.62</td>
<td>20.14/16.91</td>
</tr>
<tr>
<td>CBD-RDN</td>
<td>20.62/14.52</td>
<td>20.68/11.21</td>
<td>21.51/27.18</td>
</tr>
</tbody>
</table>

**Table 1 SNR results (whole data/enlarged part).**

**Field data experiments**

We further apply the field data from the North Sea (Petroleum-Geo-Systems, 2017). The shot record we
choose has 1000 time samples and 256 receivers, with a 4 ms temporal sampling interval and 30m spatial sampling interval; the spectra are shown in Figure 5. The reflection wave is selected for the denoising problem, and the diving wave is selected for interpolation, as shown in Figure 6. The spatial sampling distance of our training data here is set to 30m to match the field data. Compared to the curvelet method, our CBD-RDN method obtains better results. As the spectra show, nearly all the information of the low frequency has been retained, which is crucial information of the full waveform inversion.

Conclusions

We propose a new convolutional neural network, named the CBD-RDN, for seismic data reconstruction, including tasks of denoising and interpolation. In the part of data preparation, we introduce strategies of consistence of frequency bands and data augmentation, and our method has better performance in both synthetic and field data.
Figure 6 Results of field data in random denoising and super-resolution. From top to bottom: denoising test in the reflection wave, and super-resolution test in the diving wave. From left to right: initial data, denoising results of curvelets, denoising results of our CBD-RDN method.

References


