Elastic wave mode decomposition in anisotropic media with convolutional neural network

Introduction

Seismic anisotropy is widely distributed in the subsurface of Earth. In an anisotropic medium, P and S waves are intrinsically coupled due to the nature of elastic wave propagation. For the elastic reverse time migration (ERTM), it is essential to isolate the P and S wave modes before applying the imaging condition in order to avoid the image crosstalks. However, the polarization direction of P or S waves in anisotropic media, which is crucial for the mode separation, is no longer parallel or perpendicular to the propagation direction and will be spatially varying with the change of model parameters. Therefore, the wave mode separation or vector decomposition of elastic wavefields becomes infeasible because of the expensive computational cost to calculate the polarization.

To efficiently apply the P/S wave mode decomposition in heterogeneous anisotropic media, the general idea is to find an approximation to the spatially varying polarization of the P or S wave mode, then the P and S waves are decomposed by projecting the wavefield along the reconstructed polarization direction. Yan and Sava (2011) referred to the idea of phase shift offset plus interpolation (PSPI) and proposed an efficient algorithm. Cheng and Fomel (2014) introduced a low-rank approximation algorithm that can decompose wavefields with high accuracy and efficiency. Wang et al. (2018) extended the low-rank approximation method through the domain decomposition of the model to achieve better efficiency. However, the computational cost of wave-mode decomposition is still expensive, especially when the rank of the model is quite high. In recent years, the deep/machine learning method has achieved wide success in seismic data processing. Particularly, Kaur et al. (2019) proposed a deep learning method with a generative adversarial network to separate scalar qP- and qSV-waves. Xiong et al. (2020) employed the convolutional neural network (CNN) to obtain the separated P/S recording of the multi-component data. All these applications are proved effective and efficient for the P/S mode separation.

In this abstract, we proposed a CNN-based method to achieve P and S wave mode decomposition for anisotropic media. The network is carefully designed, including the dilated convolution and anti-aliased CNN, to suit this task. The labelled data for training are generated by the low-rank algorithm. We use a small number of labels with small size for training and then test the well-trained network on a large-size model. The synthetic examples demonstrate the accuracy and efficiency of the CNN-based wave mode decomposition.

Method

The elastic wave mode decomposition expects to isolate the vector qP- or qSV-wavefield (qPx-, qPz-, qSVx- and qSVz- for 2D case) from the original vector wavefield. From the perspective of computer vision, it can be regarded as a denoising task, in which the X- and Z-components of the displacement wavefield are the two-channel input of the network and the expected output are the qSVx- and qSVz-wavefield. Then, the qPx- and qPz-components are reconstructed by simply subtracting the qSV waves from the original wavefields. Note, it is also feasible to first predict the qP- wavefield with network and then reconstruct qSV-wavefield with subtraction.

We use U-Net (Ronneberger et al., 2015) as the backbone of our network to achieve wave mode decomposition, in which an encoder-decoder architecture is included to extract the feature maps with different scales. As shown in Figure 1, the basic convolution block consists of two 3*3 convolution layers, each followed by a ReLU activation function. In the stage of encoding, we use an anti-aliased CNN (Zhang, 2019) to avoid the aliasing effects introduced by the down-sampling. The feature maps are doubled after each down-sampling operation. At the end of encoder, we stack four convolution layers with kernel size 3*3 to ensure the high-level features are captured. Note that the dilated convolution is used to expand the receptive field without adding additional parameters. In the stage of decoding, we use the transposed convolution with kernel size 4*4 and stride 2 to apply the up-sampling. The skip-pathways concatenate the feature maps of the same spatial resolution in the encoder and decoder to recover fine-grained details of the target objects. Finally, the predicted vector wavefields are generated through a convolutional layer with kernel size 1*1.
Considering the amount of calculation and the data size, we randomly crop all the data into patches of uniform size for training. The validation and testing stages are also implemented on data patches. The sliding window technique is used to complete the inference of the whole data so that our decomposition method can process seismic data of any size. In this way, the small-size labels are sufficient for training, which is very important for low-rank algorithm to efficiently create the labelled data.

**Examples**

In this section, we implement the proposed network with BP 2007 TTI model, as shown in Figure 2a. The tilt angle is set to zero here, so it is a VTI medium. To collect the dataset for training, we cut the model into small blocks with a size of 3.2*3.2 km. In each block, the elastic wave equation is solved with the staggered grid finite difference method, and the low-rank algorithm (Cheng and Fomel, 2014) is used to obtain the vector qP- and qSV-waves. For each block, 23 snapshots are extracted with an even sample interval along the time axis as the labelled data. We randomly select 20 blocks for training and 8 for validation. During training, we randomly crop the X- and Z-components of the total wavefield and the qSVx- and qSVz-components into 256*256 patches as the inputs. We use L1-norm as the loss function and Adam algorithm as the optimizer. The initial learning rate is set to 0.0005 and decays exponentially with the number of epochs. The entire training process is conducted using Pytorch framework. We choose signal-to-noise ratio (SNR) as a quantitative indicator to evaluate the effectiveness of our method.

We trained our network for 100 epochs and chose the model that performed best on the validation set as the final result. In addition, we find that using vector qSV-wavefields as the prediction targets of the network can achieve better performance. In this case, the vector qP-wavefields can be obtained by subtracting qSVx- and qSVz-components from the total wavefield, respectively. We test the trained model on the rest samples located at different blocks to verify its performance and robustness. Figure 2b and 2c show the displacement wavefields of a certain time slice in the testing set. The corresponding vector qSV- and qP-wavefield predicted by the CNN-based method are shown in Figure 2d~2g, which are comparable to the results using the low-rank algorithm shown in Figure 2h~2k. According to the amplitude comparisons at x=1 km (Figure 3), we find that the predicted vector wavefields are almost the same as the references. The average SNR of the predicted qSVx- and qSVz-components in the testing set are 30.707 dB and 28.389 dB, respectively. The reconstructed vector qP-wavefields through subtraction have an average SNR of 40.179 dB for qPx- and 41.005 dB for qPz-component. We also counted the average SNR of each time slice for all testing data. As shown in Figure 4, the network has a stable performance for every time step. To further evaluate the trained network, we apply it to the data modelled with a large size of 11.2*11.2 km. Figure 5 shows a satisfactory result with only small residuals in the decomposed wavefield. It means that the proposed method is robust to train with small-size labels and implement on the wavefield with arbitrary size.

Moreover, on the final test with a large model size (11.2*11.2 km), the average inference time of the network for a time slice is only 1.2 s on a single GPU, outperforming the low-rank algorithm which takes almost 10s on an 8-core CPU speeding up by OpenMP. Besides, the computational cost of the low-rank algorithm highly depends on the rank of the decomposition operator that is related to the

---

**Figure 1** The network for vector wave-mode decomposition. The inputs are the X- and Z-components of the displacement wavefields and the outputs are the vector qP- or qSV-wavefields.
Figure 2 BP model and decomposed results of the multicomponent wavefield for a small block at $t=0.7$ s. (a) Vertical $P$-wave velocity of BP model. The red box roughly indicates the block’s location. (b)–(c) Vector elastic displacement wavefields. (d)–(g) the predicted vector $qSV$- and $qP$-wavefields using the network. (h)–(k) the ground truth of $qSV$- and $qP$-waves obtained by the low-rank algorithm.

Figure 3 Amplitude comparisons of (a) $qSVx$- and (b) $qSVz$-components at $x=2$ km from Figure 2. The black line represents the amplitude of the decomposed data using low-rank algorithm, while the red dot-dash line indicates the prediction by network.

Figure 4 The average SNR of the predicted wavefields for each time slice. The red and blue lines indicate the SNR of the predicted $qSVx$- and $qSVz$-components. The orange and grey lines represent the SNR of the $qPx$- and $qPz$-components obtained by subtraction.
complexity of the geological model. That means the rank of the decomposition operator may reach a quite high value when the anisotropic parameters smoothly vary in the model. But there is no such problem with the CNN-based method. These results illustrate that the proposed method can achieve wave mode decomposition accurately and efficiently.

Figure 5 The predicted results of the network on a large size model. (a) Geological structure of the large block. (b)–(c) The total displacement wavefields. (d)–(e) The decomposed vector wavefields using the trained network without fine-tuning.

Conclusions

We proposed a deep learning algorithm for wave mode decomposition of the anisotropic wavefields. The designed network first predicts the X- and Z-components of the qSV-wavefield and then reconstruct the qP-wavefields by subtraction. Numerical examples demonstrate the robustness and generalization of our method. The results also inspire us to cut small blocks for training and then apply the network to the data of large models to improve efficiency and flexibility. We expect the approach to provide an efficient way for the subsequent imaging or inversion in anisotropic media.

Acknowledgments

We acknowledge the supports from the National Key R&D Program of China (2018YFC0310104), the National Natural Science Foundation of China (42074157 & 41674117). We thank SEG for providing the synthetic model for this work.

References


