Multiscale full waveform inversion using GPU
Jian Mao*, Ru-Shan Wu, University of California at Santa Cruz; Baoli Wang, Xi’an Jiaotong University

Summary

We develop an efficient multi-scale time domain full waveform inversion (FWI) method using Graphics Processing Unit (GPU). Benefited from the acceleration of GPU, the intensive computational cost of time domain FWI becomes affordable on our desktop. As the geological targets in subsurface area are usually multi-scale, we propose a multi-scale time domain FWI method using GPU to invert the model more efficiently and effectively. The strategy for multi-scale FWI is very similar to frequency domain FWI. Therefore we make a comparison between the results of our multi-scale method and the frequency domain FWI results. We use 2D Overthrust model and 2D Marmousi model to demonstrate the validity of our method.

Introduction

Seismic full waveform inversion (FWI) becomes more and more popular in nowadays, because it provides us promising results to solve inverse problems, such as inverting the model parameters (Vp, Vs, density and anisotropic parameters) from the observed data. Standard approach is to convert the original problem to an optimization problem, which minimizes some objective functional. The objective functional usually consists of errors between the observed data and the prediction from models and prior information about the model. The data error can be defined in a L2-norm least squares sense. The process of finding the solution is usually iterative. A starting model is needed and it is continuously updated until a minimum of the objective functional is reached.

Generally speaking, the full waveform inversion can be classified in two categories: one is frequency domain approaches (Pratt et al., 1998, Pratt, 1999, Pratt and Shipp, 1999, Sirgue and Pratt, 2004, Operto et al., 2007, Ben-Hadj-Ali et al., 2008, Virieux et al., 2011), the other is implemented in time domain (Boonyasiriwat et al., 2010, Bunks et al., 1995, Gauthier et al., 1986, Luo and Schuster, 1991, Schuster and Quintus-Boz), 1993). We can see a lot of promising results by methods from both categories. However, there still remains many issues in FWI theory and implementations, such as how to build good initial model, how to make the convergence better, how to improve the computation efficiency and so on. In recent years, a lot of progress has been made for FWI, such as Laplacian domain FWI(Bae et al., 2010, Shin and Ha, 2008, Shin and Min, 2006), preconditioning for FWI (Causse et al., 1999, Hu et al., 2011, Porsani and Pomponet Oliveira, 2008), FWI using plane source (Vigh and Starr, 2008) or multi-shot and phase encoding technique (Ben-Hadj-Ali et al., 2011, Choi and Alkhalifah, 2011, Guitton and Diaz, 2011, Wang and Gao, 2010, Mora, 1987), multiscale FWI (Boonyasiriwat et al., 2010, Boonyasiriwat et al., 2009, Bunks et al., 1995, Operto et al., 2004), FWI using compressed sensing techniques (Herrmann and Li, 2012, Mallat, 1989). There are also some attempts to invert the elastic constants and anisotropic parameters (Virieux et al., 2011, Bae et al., 2010, Brossier et al., 2010). In addition to the advances in algorithms, the acceleration of FWI can be achieved by hardware advances such as the development of graphics processing unit (GPU) for numerical calculation.

As we know, the geological models are usually composed of structures of multi scales. Thus, it is natural to conduct FWI in a multi-scale way. For the frequency domain FWI, the strategy from low frequency to high frequency is to invert the model from large scale to small scale, resulting in very good inversion result (Sirgue and Pratt, 2004, Ben-Hadj-Ali et al., 2008, Operto et al., 2007). In the time domain FWI, similar multi-scale procedure can be achieved by applying multiple filters to generate data with different frequency band (Bunks et al., 1995, Boonyasiriwat et al., 2009).

In this work, we present a GPU-based multi-scale FWI using conjugate gradient algorithm in time domain. Since the computational efficiency using GPU is much better than CPU-based FWI, we can do the multi-scale FWI with a very short turn-around period. Both 2D Overthrust model and Marmousi model are tested using the proposed method, and the results are compared with those obtained from frequency domain method.

Time domain Full Waveform Inversion

As we know, FWI is a method based on minimizing an objective function which measures the difference between the simulated and the acquired data. The objective functional of records can be defined as the L2-norm of residuals between them, which can be written as (Gauthier, Virieux et al. 1986)

\[
\min_{\nu} S(\nu) = \sum_{i=1}^{N_s} \sum_{j=1}^{N_r} \left[ p(x_i, t; x_j) - \tilde{p}(x_i, t; x_j) \right],
\]

(1)

where \( S(\nu) \) is the misfit function, \( \min S(\nu) \) means to minimize the objective function, \( \nu \) represents the geological model (velocity model), \( N_s \) and \( N_r \) stand for the number of sources and the number of receivers for the corresponding sources. \( T \) is the total time length of the data record. \( p(x_i, t; x_j) \) and \( \tilde{p}(x_i, t; x_j) \) are the simulated data and the observed data respectively.
Multiscale full waveform inversion using GPU

Obviously, FWI belongs to high dimensional nonlinear inversion problem. We can use Gradient methods, such as the steepest descent method, conjugate gradient method or Newton gradient method to solve these inversion problems. An initial velocity model needs to be provided at the beginning and the velocity is updated iteratively as follow

\[ v_{n+1} = v_n + \alpha_n \frac{\partial S}{\partial v} , \]  

where \( n \) is the index of iteration steps, \( \alpha_n \) is the iterative step length obtained by a linear search technique, and \( \frac{\partial S}{\partial v} \) is the gradient of the objective functional with respect to velocity model, which can be calculated by zero-lag correlation of the forward-propagated source wavefields and the backward-propagated residual wavefields as follows

\[ \frac{\partial S}{\partial v} = 2 \frac{\partial p}{\partial \mu} \int_0^t \lambda \frac{\partial p}{\partial t} dt , \]  

where \( p \) and \( \lambda \) are the forward-propagated source wavefields and the backward-propagated residual wavefields, respectively. Both of these wavefields satisfy the two-way wave equation as follow

\[ \frac{1}{v^2} \frac{\partial^2 p}{\partial t^2} = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} + f \]  
\[ p(x,z,t=0) = 0 \]  
\[ \frac{d}{dt} p(x,z,t=0) = 0 \]  

\[ \frac{1}{v^2} \frac{\partial^2 \lambda}{\partial t^2} = \frac{\partial^2 \lambda}{\partial x^2} + \frac{\partial^2 \lambda}{\partial z^2} + (p - \hat{p}) \]  
\[ \lambda(x,z,t=T) = 0 \]  
\[ \frac{d}{dt} \lambda(x,z,t=T) = 0 \]  

where \( f \) is the source function.

Based on the FWI theory, the flow chart of FWI is given in Figure 1. From the work flow, we can see that it needs at least two simulations in each iteration step. In fact, we need at least three, because we use the line search technique to calculate the step length and the saving boundary technique to avoid intensive data movement between CPU and GPU. From Figure 1, we note that most of the computational cost of FWI comes from the forward and backward simulations. Since these simulations are very time-consuming, FWI is computational demanding for research and for industry application. Therefore, we implemented FWI on GPU.

![Flow chart of FWI](image)

**Multi-scale strategy for time domain GPU FWI**

Since FWI on GPU has much better improvement in the computational time, we propose to use this GPU based method to do multi-scale FWI. The strategy for multi-scale FWI can be summarized as follow

1. Start FWI from the initial model
2. Use low frequency data to do the GPU FWI, obtaining the large scale components of the model
3. Use the output model from step 2 and higher frequency data to do GPU FWI, refining the smaller scale component
4. Loop step 3 over the frequency bands till the inverted result is satisfied

As the subsurface contains a number of components from large scale to small scale, this multi-scale strategy will give us the best result for FWI. This method is quite similar to the frequency domain FWI strategy, where the model is inverted from low frequency to high frequency. In the numerical test section, we will compare the results generated by these two methods.

**Numerical examples**

1. Frequency domain FWI for 2D Overthrust Model

We first run the frequency domain FWI method to invert the 2D Overthrust model. We use the FWT package from SEISCOPE consortium (http://seiscope.oca.eu/) to generate these results. The Overthrust model is shown in Figure 2. The grid size is \( 801 \times 187 \) with an interval of 25m for both horizontal and depth direction. In this test, we use 201 shots and a full aperture data from 201 receivers equally distributed on the surface. The initial model is a smoothed version of the true model, but within the top 100m depth, we use true velocity model as the starting model (shown in Figure 3). We test the sequential frequency domain method.
Multiscale full waveform inversion using GPU

to invert model with 7 frequency groups (3.5Hz, 4.7Hz, 7.2Hz, 9.6Hz, 13.3Hz, 16.9Hz, 20.6Hz), and we select four of them to plot in Figure 4. From the initial model, the velocity is inverted from low frequency to high frequency. We can see that large scale components are inverted by low frequency data and the small scale parts are inverted by high frequency data.

Figure 2 The 2D Overthrust model

Figure 3 Initial velocity model

Figure 4 Inverted velocity using frequency domain FWI with 7 frequency groups (from top to bottom: 3.5Hz, 4.7Hz, 9.6Hz, 20.6Hz)

2. Multi-scale time domain FWI for Overthrust Model

First, we test single scale GPU FWI. We use the same initial model, but the acquisition geometry is slightly different from that used in the frequency domain FWI test. We use 51 shots and a full aperture of 401 receivers evenly distributed on the surface. Figure 5 shows the single-scale FWI results, which are the inverted models with 30 iterations using the data centered at 3Hz, 5Hz and 10Hz respectively. From these results, we can clearly see the different performances of data in different frequency bands. Similarly, the large scale component in the model can be inverted by low frequency band data, while the data of high frequency band can be used to invert fine structures. Even with a large number of iterations, we still cannot reach to a good result by using the 10Hz data (shown in Figure 6). This is the motivation for us to test multi-scale FWI.

Figure 5 Inverted velocities by single-scale FWI (from top to bottom the frequency band used are 3Hz, 5Hz and 10Hz)

Figure 6 Inverted velocity by single-scale using 10Hz data with 300 iterations

Then we test the proposed multi-scale GPU FWI method. Figure 7 shows the results using the strategy we proposed in the last section. We can see the velocity model is much better reconstructed using the multi-scale inversion. Figure 8 compares the true velocity model, the frequency domain result and our multi-scale result. Our result is better in the deeper area. We also select two vertical lines (x=1.6km and x=13.75km) to compare the true, initial and inverted velocity profiles and the results are shown in Figure 9. From this figure, we see the inverted velocity profiles match the true velocity model very well.
Multiscale full waveform inversion using GPU

Figure 7 Velocities inverted by multi-scale FWI (from top to bottom are the results after 3Hz, 5Hz, 7Hz and 10Hz inversion)

Figure 8 Comparison of true velocity (top), frequency domain FWI result (middle) and time domain multi-scale FWI result (bottom)

3. Multi-scale FWI for Marmousi model

We further test the Marmousi model. This model has a grid size of 767 × 247 with a grid space of 12m. The acquisition system is composed of 51 shots and a full aperture with 383 receivers evenly distributed on the surface. Figure 10 is the true velocity model. Figure 11 shows the initial model and the inverted model by our method.

Figure 9 Comparison of true, initial and multi-scale FWI inverted velocity (left: x=1.6km, right: x=13.75km)

Figure 10 True velocity of 2D Marmousi model

Figure 11 Initial velocity (top) and the inverted velocity (bottom)

Conclusions

We develop an efficient multi-scale time domain full waveform inversion method with the implementation on GPU. We use 2D Overthrust model and 2D Marmousi model to demonstrate the validity of our method. From the results, we can see our method is comparable with frequency domain multi-scale method, and is even better in some cases. In the future, we will try to extend it in 3D case and parallelize the method using GPU clusters.

Acknowledgement

We thank Xiaobi Xie, Jinghua Gao, Yingcai Zheng, Yu Geng, Rui Yan, Fuchun Gao and Chuck Mosher for helpful discussions. The work is supported by WTOPI (Wavelet Transform On Propagation and Imaging for seismic exploration) Project at University of California, Santa Cruz.
EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2012 Annual International Meeting, SEG, Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES


Shin, C., and W. Ha, 2008, A comparison between the behavior of objective functions for waveform inversion in the frequency and Laplace domains: Geophysics, 73, no.5, VE119–VE133.


