GeoConvention 2014 FOCUS adapt refine sustain

Full waveform inversion with phase encoded deconvolution imaging condition

Wenyong Pan, Kris Innanen, Gary Margrave*
CREWES, University of Calgary
wpan@ucalgary.ca

Summary

Full waveform inversion (FWI) is a very important method for estimating the subsurface parameters. While it suffers from extensively computational cost, slow convergence rate, etc, which impede its practical application. In our implementation, to reduce the computational cost, a linear source encoding strategy is used for the gradient calculation. The gradient is a poorly scaled reverse time migration image based on cross-correlation imaging condition. The Hessian matrix can enhance the poorly scaled gradient considerably while it is extremely expensive to calculate the full Hessian directly. The Hessian matrix is thought to carry out source illumination compensation, which is actually equivalent to diagonal pseudo-Hessian. We can also construct the source illumination using phase encoding strategy. Thus, preconditioning the phase encoded gradient using phase encoded diagonal pseudo-Hessian forms phase encoded deconvolution imaging condition. These strategies can reduce the computational cost and improve the convergence rate of FWI. We carried out a numerical experiment with a portion of Marmousi model and analyzed the effectiveness of the proposed strategies.

Introduction

Full waveform inversion (FWI), formulated as a least-squares inverse problem, seeks to minimize the difference between the observed data and synthetic data (Virieux and Operto, 2009). It has drawn much recent attention for its powerfulness in estimating the subsurface parameters with high resolution. To fulfill its potential, FWI algorithms must cope with significant issues: computational burden, slow convergence rate and cycle skipping being important amongst these.

The efficient gradient calculation strategy proposed by Lailly (1983) and Tarantola (1984), the adjoint state method, allows us to construct the gradient by finding the zero-lag cross-correlation between the forward modelled wavefields and the back-propagated wavefields. In this approach, we only need $2N_S$ forward simulations to calculate the gradient, where N_S indicates the number of sources. However, for a 3D or a large 2D problem, the computational cost remains very high. The simultaneous source technique has been introduced to address this problem further, at the cost of the introduction of cross-talk artifacts. These artifacts can in turn be suppressed with a sufficient number of ray parameters, which are controlled by the take-off angle at the surface location. By combining these strategies, the number of forward simulations in one FWI iteration becomes $2N_p$, where N_p is the number of ray parameters.

FWI suffers from a slow convergence rate. The gradient absent the inverse Hessian does not account for geometric spreading effects (Shin et al., 2001). Gradient based method, which replaces the Hessian matrix with an identity matrix, suffers from crude scaling, which slows convergence. The un-scaled image can be enhanced considerably with the inverse Hessian (Pratt et al., 1998), but direct inclusion of the inverse Hessian matrix is computationally unfeasible. The full Hessian matrix is thought to carry out illumination compensation, which is actually equivalent to the diagonal pseudo-Hessian. We also employed the phase encoding technique to construct the source illumination. The phase encoded gradient with phase encoded source illumination becomes phase encoded deconvolution imaging condition. Furthermore, to avoid the local minimum and improve the chance that the global minimum can be achieved, the multi-scale approach was also applied by increasing the frequency band from low to high iteratively.

In this paper, the basic theory for least squares inverse problem is reviewed firstly. And then we introduce how to construct the deconvolution based gradient using phase encoding technology. Finally, these strategies are practiced on a modified Marmousi model to prove the effectiveness of the proposed strategies.

Theory and Method

The least-squares inverse problem

As a least-squares local optimization, FWI seeks to minimize the difference between the synthetic data and observed data (Lailly, 1983; Tarantola, 1984) and update the model iteratively. The misfit function ϕ is given in a least-squares norm:

$$\varphi\left(s_0^{(n)}(r)\right)=\frac{1}{2}\int d\omega\left\{\sum_{r_s,r_g}\left\|\delta P\right\|_2\right\}, \tag{1}$$

where $s_0^{(n)}(r)$ are the model parameters, the square of the slowness in the n th iteration. δP mean the data residuals, the difference between the observed data and synthetic data. $\left\|\cdot\right\|_2$ indicates the $l-2$ norm. Applying a second order Taylor-Lagrange development of the misfit function and then taking partial derivative with respect to the model parameters can obtain the model perturbation, which consists of inverse Hessian and the gradient. Then the model can be updated iteratively.

Phase encoded gradient

The gradient is the first order partial derivative of the misfit function with respect to the model parameters and it can be obtained by applying a zero-lag cross-correlation between the forward modeling wavefields and back-propagated data residuals, which avoids the direct computation of the partial derivative wavefields. However, it is still very expensive to calculate the gradient using traditional shot-by-shot method. The gradient construction is equivalent to a migration process. We can employ the phase encoding technology to construct the gradient by exciting the sources simultaneously. So, the linear phase encoded gradient can be expressed as (Pan et al., 2013a, b and c):

$$\tilde{g}\left(r,p^g,\omega\right)=\sum_{\omega}\sum_{r_s}\sum_{r_s^{'}}\sum_{j=1}^{N_p^g}\Re\left\{\omega^2\left|f(\omega)\right|^2\tilde{G}\left(r,r_s,\omega\right)\bar{G}^*\left(r,r^{'}_s,\omega\right)e^{i\omega p_j^g\left(x_s-x_s^{'}\right)}\right\}, \tag{2}$$

The phase encoding technology can reduce the computational cost effectively but unfortunately introduce cross-talk artifacts. Slant stacking over sufficient ray parameters can disperse these cross-terms reasonably.

Phase encoded deconvolution imaging condition

The Hessian matrix can enhance the poorly scaled gradient considerably. However, direct calculation of the Hessian matrix is extremely expensive. The Hessian matrix is thought to carry out illumination compensation, which is equivalent to the diagonal part of the pseudo-Hessian, proposed by Shin et al., (2001). In this research, we used a chirp phase encoding technique to construct the diagonal pseudo-Hessian (or source illumination):

$$\tilde{H}\left(p^{H},\omega\right)=\sum_{\omega}\sum_{r_s}\sum_{r_s^{'}}\sum_{j=1}^{N_p^H}\Re\left\{\omega^2 G\left(r^{'},r_s,\omega\right)G^{*}\left(r^{''},r^{'}_{s},\omega\right)e^{i\omega\left(p_j^H+\varepsilon\Delta p\right)\left(x_s-x_s^{'}\right)}\right\}, \qquad (3)$$

Similarly, the phase encoded pseudo-Hessian also suffers from cross-talk noise. Slant stacking over sufficient ray parameters can suppress these crosstalk artifacts. The diagonal pseudo-Hessian can be obtained when $r^{'}=r^{''}$ in equation (3). We know that the gradient is equivalent a reverse time migration image with cross-correlation imaging condition. Hence, the phase encoded gradient preconditioned by the phase encoded diagonal pseudo-Hessian becomes phase encoded deconvolution imaging condition:

$$\tilde{I}\left(r,p^{g},p^{H},\omega\right)=\frac{\sum_{\omega}\sum_{r_s}\sum_{r_s^{'}}\sum_{j=1}^{N_p^g}\Re\left\{\omega^2\left|f\left(\omega\right)\right|^2\tilde{G}\left(r,r_s,\omega\right)\bar{G}^{*}\left(r,r^{'}_{s},\omega\right)e^{i\omega p_j^g\left(x_s-x_s^{'}\right)}\right\}}{\sum_{\omega}\sum_{r_s}\sum_{r_s^{'}}\sum_{j=1}^{N_p^H}\Re\left\{\omega^2 G\left(r,r_s,\omega\right)G^{*}\left(r,r^{'}_{s},\omega\right)e^{i\omega\left(p_j^H+\varepsilon\Delta p\right)\left(x_s-x_s^{'}\right)}\right\}+\lambda A_{\max}}, \qquad (4)$$

where $\lambda A_{\max}$ is the damping term to make the deconvolution based gradient stable.

Examples

The strategies proposed in this research are practiced and applied on a modified Marmousi model. The model has 180x767 grid cells with the same grid interval of 5m in horizontal and vertical. 380 point sources are distributed on the surface with a source interval of 10m from 0 to 3800m and 767 receivers are deployed on the surface with a receiver interval of 5m from 5m to 5750m. The source function is a Ricker wavelet with a dominant frequency of 30Hz. The ray parameter range used for linear phase encoding method is [-0.3s/km, 0.3s/km] with a step of 0.1s/km are tested for comparison. The lowest frequency band used is [0Hz, 5Hz] and the frequency band increases by 5Hz for every 10 iterations.

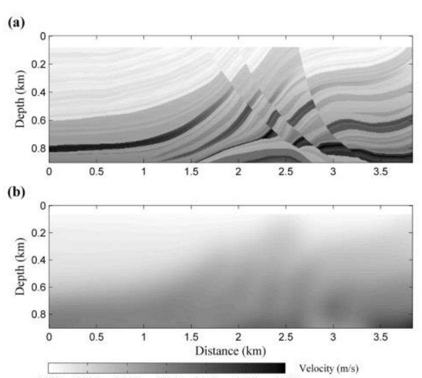

Figure 1: (a) True velocity model; (b) Initial Velocity Model.

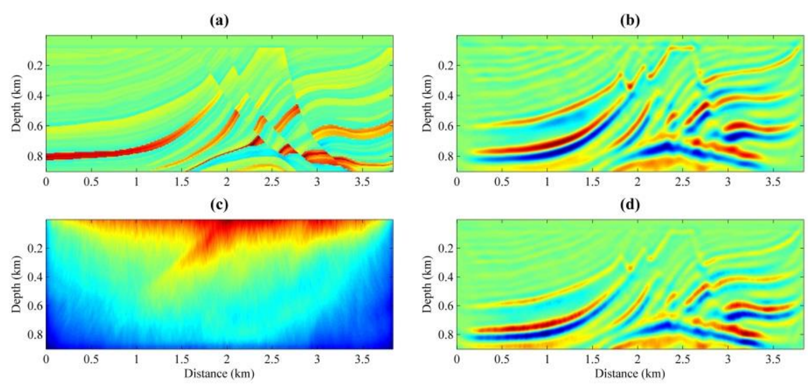

Figure 2: (a) True model perturbation; (b) Gradient based method; (c) Source illumination; (d) Deconvolution based gradient.

Figure 1a shows the exact P-wave velocity model and Figure 1b shows the smoothed P-wave velocity model used as the reference model in this research. Figure 2a shows the normalized true model perturbation. Figure 2c show the phase encoded source illumination using the initial velocity model with 60 simulations. Figure 2b and d show the model perturbation estimation without precondition and with precondition respectively. We can see that the model perturbation estimation with precondition is more close to the true model perturbation. And we also calculate their errors as shown in Figure 3a and b. Two vertical lines at 1km and 3km are extracted for comparison, as shown in Figure 4. It can be seen that the normalized error with precondition, indicated by the blue lines are more close to zero, comparing with the red lines. Hence, we can conclude that the phase encoded deconvolution imaging condition method can estimate the model perturbation better.

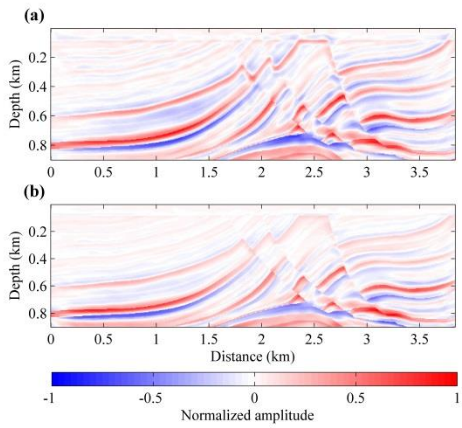

Figure 3: Model perturbation estimations errors by gradient based method (a) and deconvolution based gradient method.

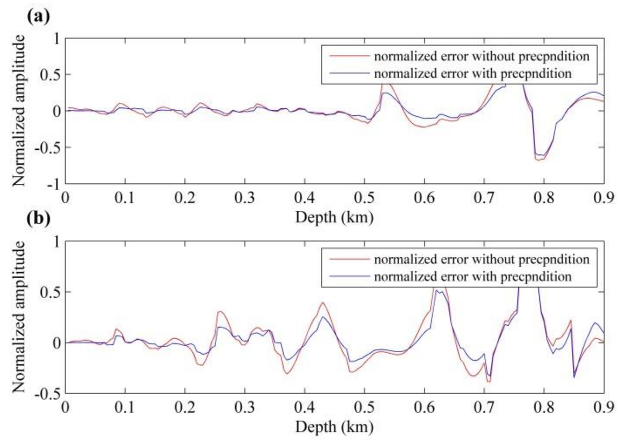

Figure 4: Model perturbation estimations errors at 1km (a) and 3km (b).

To reduce the computational cost further, we propose to use one ray parameter to construct the gradient in each iteration but vary the ray parameter regularly when the iteration proceeds. In this case, the model update can be balanced iteratively. Figure 5a and b show the FWI inversion results after 200 iterations using traditional gradient based method and deconvolution based gradient method. We can notice that the deep parts of the inversion result in Figure 5b is rebuilt better than that of Figure 5a.

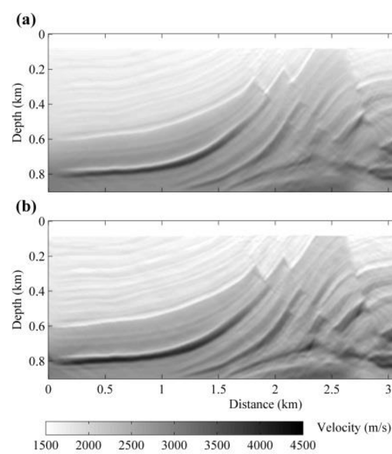

Figure 5: Inversion results by gradient based method (a) and deconvolution based gradient method.

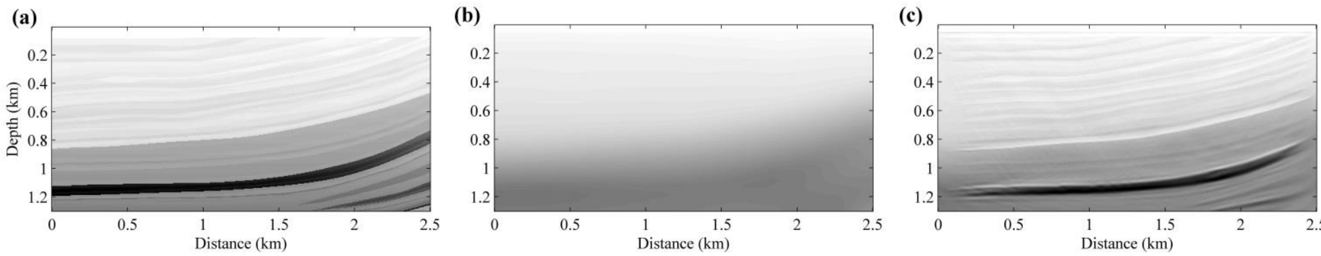

Figure 6: (a) True velocity model; (b) Initial velocity model; (c) Inversion result after 200 iterations with fixed ray parameter of -0.1s/km.

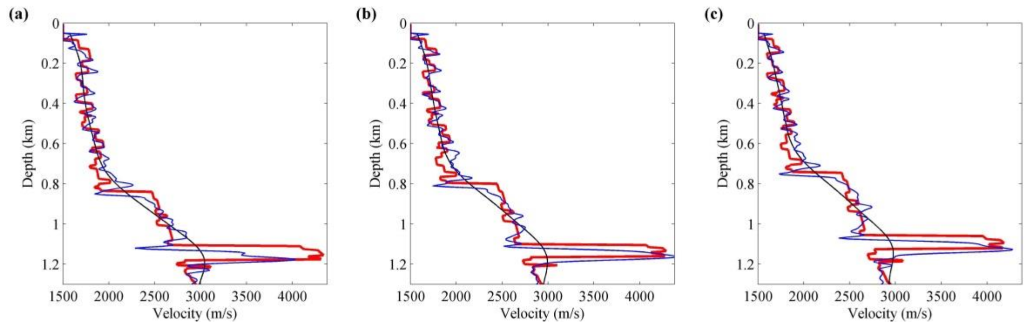

Figure 7: Inverted results at 0.5km (a), 1km (b) and 1.5km (c) respectively. The red and black lines indicate the true velocity model and initial velocity model respectively. The blue lines denote the inverted velocity model.

Generally the ray parameter range can be determined by the dip angles of the subsurface layers. So, if the subsurface layers are almost horizontal, it is possible for us to use just one single ray parameter to build the subsurface velocity model. We test this idea on another portion of Marmousi model, as shown in Figure 6. Figure 6a shows the true velocity model with almost horizontal subsurface layers. Figure 6b and c show the initial velocity model and inverted velocity after 200 iterations with ray parameter fixed at -0.1s/km. Figure 8a, b and c show the inversion result at 0.5km, 1km and 1.5km respectively. We can see that the velocity model is re-built very well.

Conclusions

From what we have discussed above, the conclusions can be achieved that the phase encoding strategy can reduce the computational burden of FWI effectively. Compared to gradient based method, the deconvolution gradient method with phase encoding can get a better inversion result with the same computational cost. For multisource FWI, if the subsurface layers are almost horizontal, we can obtain a very good inversion result for fixed ray parameter during iterations. The proposed strategies in this research can reconstruct the subsurface velocity model efficiently.

Acknowledgements

This research was supported by the Consortium for Research in Consortium of Elastic Wave Exploration Seismology (CREWES).

References

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration, Conference on Inverse Scattering, Theory and Applications, SIAM, Expanded Abstracts, 206–220.

Pan, W., Innanen, A. K., and Margrave, F. G., 2013a, Efficient pseudo Gauss-Newton full waveform inversion in the time-ray parameter domain: CREWES Research Report.

Pan, W., Innanen, A. K., and Margrave, F. G., 2013b, Full waveform inversion with phase encoded pseudo-Hessian: CREWES Research Report.

Pan, W., Innanen, A. K., and Margrave, F. G., 2013c, On the role of the deconvolution imaging condition in full waveform inversion: CREWES Research Report.

Pratt, R. G., Shin, C., and Hicks, G. J., 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Shin, C., Jang, S., and Min, M., 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical Prospecting, **49**, 592–606.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Virieux, A., and Operto, S., 2009, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **74**, WCC1–WCC26.