AN EFFICIENT TIME-DOMAIN FULL WAVEFORM INVERSION USING THE EXCITATION AMPLITUDE METHOD

AHREUM KIM, DONGHYUN RYU and WANSOO HA

Department of Energy Resources Engineering, Pukyong National University, 45 Yongso-Ro, Nam-Gu, Busan 48513, South Korea. wansooha@pknu.ac.kr

(Received December 20, 2016; revised version accepted August 5, 2017)

ABSTRACT

Kim, A., Ryu, D. and Ha, W., 2017. An efficient time-domain full waveform inversion using the excitation amplitude method. *Journal of Seismic Exploration*, 26: 481-498.

Time-domain full waveform inversion algorithms store the source wavefield to calculate the cross-correlation between the source and receiver wavefields to obtain the gradient direction. Saving the full source wavefield imposes an enormous burden on computer memory resources. We apply the excitation amplitude method to a full waveform inversion to reduce the memory overload. This method removes the time dimension of the source wavefield by only exploiting the maximum amplitude signals. By adopting the excitation amplitude method to store the source wavefield and to calculate the cross-correlation, we can reduce the memory requirement for a full waveform inversion by three orders of magnitude. Since the excitation amplitude method cannot handle multipathing within the source wavefield, a gradient direction obtained using this method is an approximation to the original gradient. Nevertheless, synthetic inversion examples using the Marmousi and overthrust models demonstrate the efficiency and accuracy of the proposed scheme.

KEY WORDS: full waveform inversion, time domain, excitation amplitude, sparse cross-correlation.

INTRODUCTION

Full waveform inversion (FWI) is a method used to recover subsurface information by minimizing an objective function defined using the differences between the observed and numerically modeled wavefields (Lailly, 1983; Tarantola, 1984). Although its results are promising, FWI requires many computational resources in order to invert large quantities of seismic data. The back-propagation or adjoint-state techniques are efficient methods for calculating the gradient direction of the objective function (Tarantola, 1984; Plessix, 2006).

0963-0651/17/$5.00 © 2017 Geophysical Press Ltd.

In these methods, the gradient direction is calculated using the cross-correlation between the forward propagated source wavefield and the backward propagated receiver wavefield. The backward propagation is accomplished in a time-reversed manner. We require the last time sample of the source wavefield at the start of the backward propagation as well as the first time sample at the end of the backward propagation step. Therefore, we need to store the entire source wavefield to calculate the gradient direction. This is not a severe problem for small-scale 2D FWI. However, the requisite memory to save the full source wavefield is substantially large for 3D FWI as well as for reverse time migration (RTM), which utilizes the same cross-correlation algorithm as FWI.

Several researchers have tried to resolve this problem by enhancing the computational method to decrease the memory demand. Wavefield reconstruction methods store small parts of the forward wavefield and reconstruct the full source wavefield at the backward propagation stage. Symes (2007) applied the checkpointing method (Griewank, 1992; Griewank and Walther, 2000) to RTM. The checkpointing method uses buffers to store parts of the forward modeled wavefield. One buffer is the memory required to save the wavefield of a single time step. The buffers are continuously updated while reconstructing the forward modeled wavefield at the backward propagation step. The author demonstrated that recomputing the forward modeled wavefield could reduce the memory load. The load distribution between the memory and CPU depends on the buffer size. In a 2D RTM example, the author reduced the memory requirement to 32 buffers at the cost of the recomputation ratio, which was approximately 3, and suggested that a scheme with 36 buffers could reduce storage by more than one order of magnitude for 3D RTM. Anderson et al. (2012) further applied the checkpointing method to FWI. The checkpointing method can be classified as a type of forward wavefield reconstruction method that utilizes initial conditions (Nguyen and McMechan, 2015).

Boundary values can also be used to reconstruct the forward modeled wavefield (Dussaud et al., 2008). Yang et al. (2015) used a boundary value reconstruction method in their FWI applications. This method only stores the forward modeled wavefield at the boundary area and reconstructs the forward modeled wavefield using the saved values at the backward propagation stage. Both of these reconstruction methods that utilize initial and boundary conditions ultimately increase the computation to reduce the memory burden.

Non-reconstruction methods are also used. Sun and Fu (2013) suggested a subsampling method utilizing the Nyquist sampling theory in their RTM applications. One such sparse cross-correlation method that is applied to RTM is the excitation amplitude imaging condition using the maximum amplitude signal of the source wavefield (Nguyen and McMechan, 2013). Sparse cross-correlation methods reduce the memory burden at the cost of accuracy by approximating the original result through subsampling. The excitation amplitude

method is an extreme example of a sparse cross-correlation method since it uses only one time sample for each spatial grid. However, Nguyen and McMechan (2015) demonstrated that these algorithms utilized in order to avoid storing the full source wavefield eventually yield RTM results similar to those obtained from full cross-correlation methods. Kalita and Alkhalifah (2016) further demonstrated the possibility of using this scheme for FWI.

The excitation amplitude method can be compared with the maximum intensity projection used for medical imaging disciplines (Wallis et al., 1989). This method projects a 3D medical image onto a 2D plane using the maximum intensity voxel. The excitation amplitude method removes the time dimension from the source wavefield using samples containing the maximum amplitude. One inherent difference of this method is the loss of information of the reduced dimension. The maximum intensity projection method does not retain depth information, whereas the excitation amplitude method stores the time information of the extracted samples for use in the subsequent cross-correlation procedure. The excitation amplitude method can also be compared with the most energetic method of Kirchhoff migration (Nichols, 1996; Shin et al., 2003) because the proposed method employs the most energetic traveltime and amplitude from the source wavefield. Compared with the Kirchhoff migration method, the excitation amplitude method does not implement analysis time windows (Shin et al., 2003) and uses only one time sample containing the maximum amplitude.

In this study, we adopt the excitation amplitude approach of RTM for FWI. Following a brief introduction of the FWI algorithm, we examine the efficiency and accuracy of the proposed scheme using the synthetic Marmousi (Versteeg, 1994) and overthrust velocity models.

REVIEW OF THE FULL WAVEFORM INVERSION ALGORITHM

Before we discuss the subsampling schemes, we briefly review the time-domain waveform inversion algorithm suggested by Tarantola (1984). The acoustic wave equation can be expressed as

$$[1/v^2(\mathbf{x})][\partial^2 u(\mathbf{x},t)/\partial t^2] = \nabla^2 u(\mathbf{x},t) + f(\mathbf{x},t) \quad , \tag{1}$$

where $u(\mathbf{x},t)$ is the pressure wavefield, $f(\mathbf{x},t)$ is the source function, and $v(\mathbf{x})$ is the velocity of the P-wave. The objective function using the l_2 norm is

$$E(\mathbf{m}) = \tfrac{1}{2} \sum_{s,r} \int_0^{T_{max}} [u_s(\mathbf{x}_r,t) - d_s(\mathbf{x}_r,t)]^2 dt \quad , \tag{2}$$

where $u_s(\mathbf{x}_r,t)$ is the forward modeled wavefield obtained using eq. (1), $d_s(\mathbf{x}_r,t)$ is the observed wavefield, T_{max} is the maximum recording time, and $\mathbf{m}$ is the model parameter vector. The model parameter is the P-wave velocity in the acoustic case. The subscripts s and r indicate the source and receiver, respectively (Tarantola, 1984). The gradient direction of the objective function with respect to a model parameter can be expressed as

$$\nabla_{m(\mathbf{x})}E = -[2/v^3(\mathbf{x})] \sum_s \int_0^{T_{max}} [\partial^2 u_s(\mathbf{x},t)/\partial t^2]\lambda_s(\mathbf{x},t)dt$$

$$= \sum_s \int_0^{T_{max}} S_s(\mathbf{x},t)\lambda_s(\mathbf{x},t)dt \quad , \tag{3}$$

where the source wavefield or the virtual source $S(\mathbf{x},t)$ is the second derivative of the forward modeled wavefield multiplied by the velocity term and $\lambda(\mathbf{x},t)$ is the receiver wavefield (Tarantola, 1984; Plessix, 2006). We can obtain $\lambda(\mathbf{x},t)$ by propagating the differences between the modeled and observed wavefields at the receiver position in a time-reversed manner. The integral term indicates a zero-lag cross-correlation between the source and receiver wavefields. After calculating the gradient direction, we can update the model parameter as

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha_k \mathbf{H}_{pd}^{-1} \nabla_m E \quad , \tag{4}$$

where k is the iteration number, α is the step length, and $\mathbf{H}_{pd}$ is the diagonal element of the damped pseudo-Hessian matrix (Shin et al., 2001; Ha et al., 2012).

FWI USING THE EXCITATION AMPLITUDE METHOD

We adopt the excitation amplitude method in the cross-correlation process (Nguyen and McMechan, 2013) to calculate the gradient direction and the pseudo-Hessian [eqs. (3) and (4)]. The excitation amplitude method uses only one sample from the source wavefield that contains the maximum amplitude. We need to save only one sample (S_E) and the corresponding time (T_E) for each grid as follows:

$$S_E(\mathbf{x}) = S[\mathbf{x}, T_E(\mathbf{x})] \quad , \tag{5}$$

where S is the source wavefield and

$$T_E(\mathbf{x}) = \arg\max_t |S(\mathbf{x},t)| \quad . \tag{6}$$

The argmax operator returns the time at which the source wavefield attains the maximum amplitude (Nguyen and McMechan, 2015). Therefore, the gradient can be calculated through multiplication as

$$\nabla_{m(\mathbf{x})} E = \sum_{s} s_{E,s}(\mathbf{x}) \lambda_s[\mathbf{x}, T_{E,s}(\mathbf{x})] \quad . \tag{7}$$

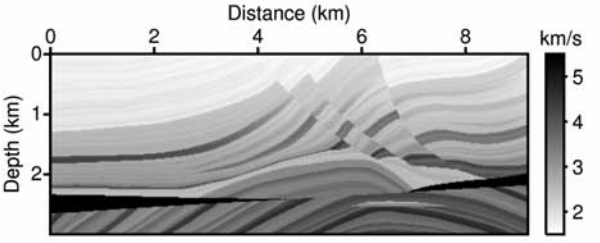
The memory demand can be reduced from $4N_gN_t$ bytes to $2 \times 4N_g$ bytes for single precision by adopting the proposed method, where N_g represents the total number of space grids and N_t is the number of time samples. The memory required is directly related to the time sampling rate. Smaller sampling rates increase the number of time samples; however, the sampling rate must be sufficiently small to satisfy the Courant-Friedrichs-Lewy condition for stable modeling (Courant et al., 1967). Generally, the memory requirement can be reduced by three orders of magnitude in consideration of numerical stability issues. Note that this method cannot account for the multipathing effect. Although only one time sample is used, the RTM results indicate that this method can increase the resolution of the RTM image by reducing the background low-wavenumber artifacts from the full cross-correlation (Nguyen and McMechan, 2013, 2015).

NUMERICAL EXAMPLES

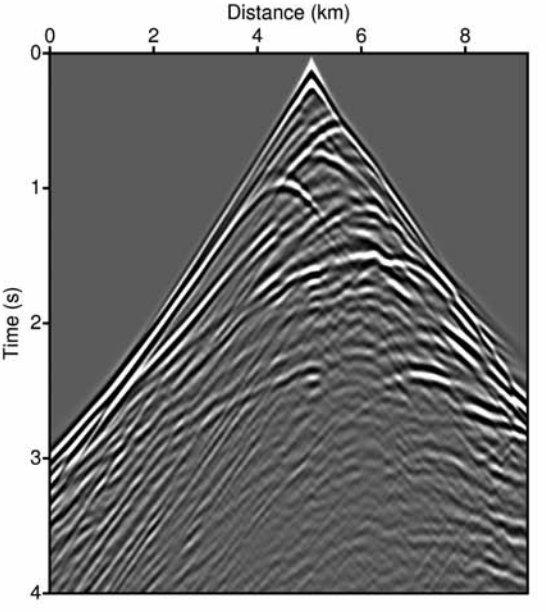
We tested the proposed methods using the Marmousi model (Versteeg, 1994) shown in Fig. 1a. The grid size was 16 m and the numbers of grids in the x- and z-directions were 576 and 188, respectively. We generated 180 shots using eighth-order and second-order finite-difference algorithms in space and time, respectively. Fig. 1b shows one of the shot gathers used as the observed data. Each shot gather contains 576 receivers at an interval of 16 m. The sampling rate was 1 ms, and the maximum recording time was 4 s. The maximum frequency of the Ricker source wavelet was 15 Hz.

We inverted the observed data using fourth-order and second-order finite-difference algorithms in space and time, respectively, both without and with the subsampling scheme. Fig. 1c illustrates a smooth velocity model used as the initial estimate. The memory required to store one single precision full source wavefield is 1,653 MB. Since the memory demand is not substantially high in this 2D example, we saved the full source wavefield within the random access memory (RAM).

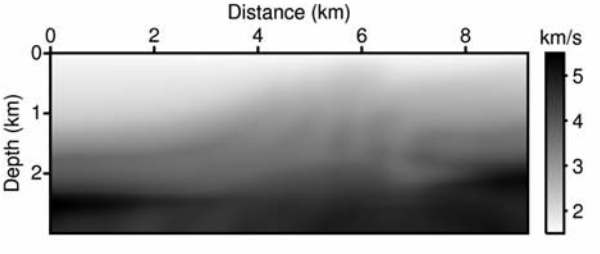
Fig. 2 shows a trace of the source and receiver wavefields from a shot that was extracted at the center of the model during the first iteration. While the full cross-correlation method uses all of the samples within the trace, the excitation amplitude method only uses one sample that contains the maximum



(a)



(b)



(c)

Fig. 1. (a) The Marmousi velocity model (Versteeg, 1994). (b) A shot gather obtained using the Marmousi velocity model. (c) A smooth initial velocity model.

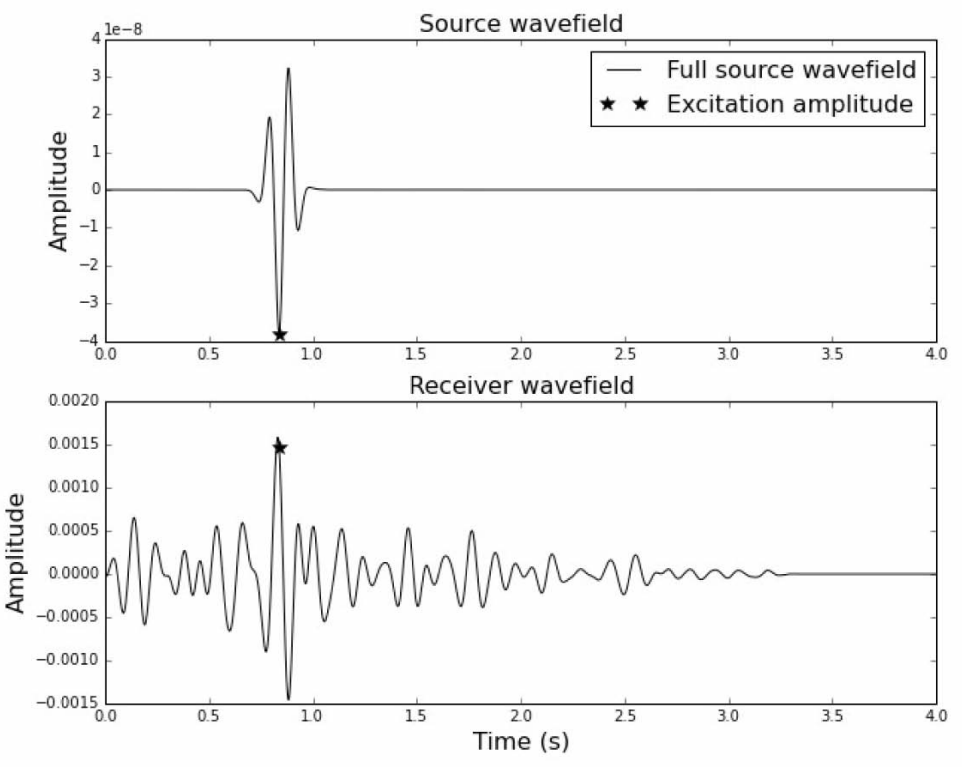
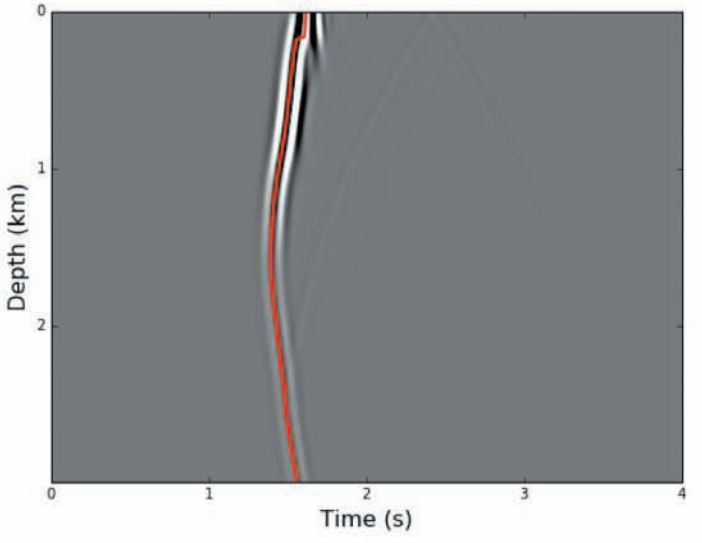


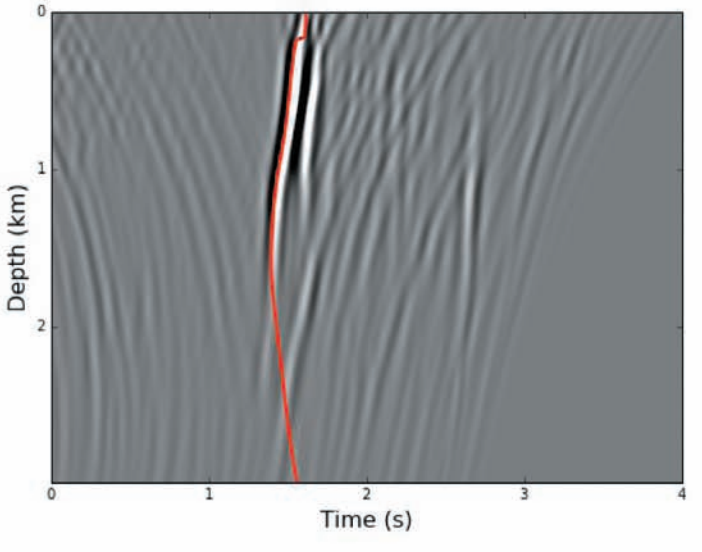
Fig. 2. The source (top) and receiver (bottom) wavefields used for the cross-correlation of the full cross-correlation and excitation amplitude methods at the center grid.

amplitude in the source wavefield. Fig. 3 displays the profiles of the source and receiver wavefields. The profiles were extracted at a distance of 2.3 km from the left margin of the model when a surface shot was detonated at the center. The red line delineates the occurrence time of the maximum amplitude signals from the source wavefield. The excitation amplitude method only uses signals along the red line to calculate the gradient direction. A discontinuity was observed along the line because the line follows the maximum amplitude signals regardless of their sign.

Fig. 4a shows the maximum amplitude values with the occurrence time contours. We multiplied the maximum amplitude source wavefield by the corresponding receiver wavefield (Fig. 4b) to obtain a one-shot gradient, as shown in Fig. 5a. The gradient direction resembles the gradient from the full cross-correlation (Fig. 5b). By stacking multiple one-shot gradients, we can obtain the velocity update direction from the first iteration. Although there are

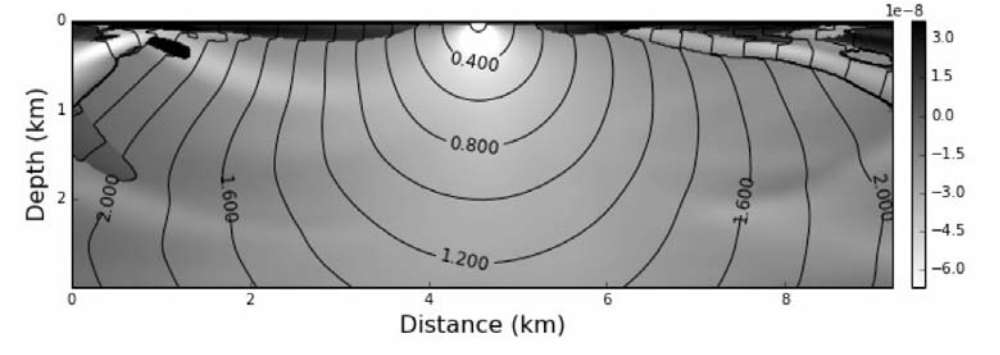


(a)

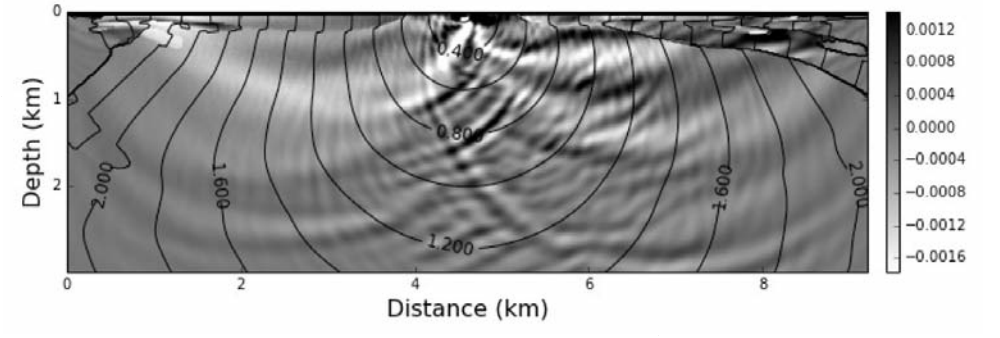


(b)

Fig. 3. Profiles of (a) the source wavefield and (b) the receiver wavefield from a shot detonated at a distance of 4.6 km from the left. The profiles are extracted at a distance of 2.3 km from the left.



(a)



(b)

Fig. 4. (a) The maximum amplitude wavefield of the source wavefield [$S_E(x)$] and the corresponding traveltime [$T_E(x)$]. (b) The receiver wavefield extracted using the excitation amplitude traveltime.

discontinuities near the surface within the one-shot gradients due to the discontinuous excitation amplitude, they disappear on the velocity update direction due to shot-stacking. When the quantity of shots is small, the discontinuities can affect the inversion results. The update directions are shown in Fig. 5c and 5d and reveal similar subsurface structures although the actual values are slightly different.

Fig. 6 shows the inversion results from the excitation amplitude method and the full cross-correlation method. The memory required to store the source wavefield is less than 1 MB for the excitation amplitude method. This memory requirement constitutes a storage reduction of approximately three orders of magnitude. The inversion result from the excitation amplitude method is similar

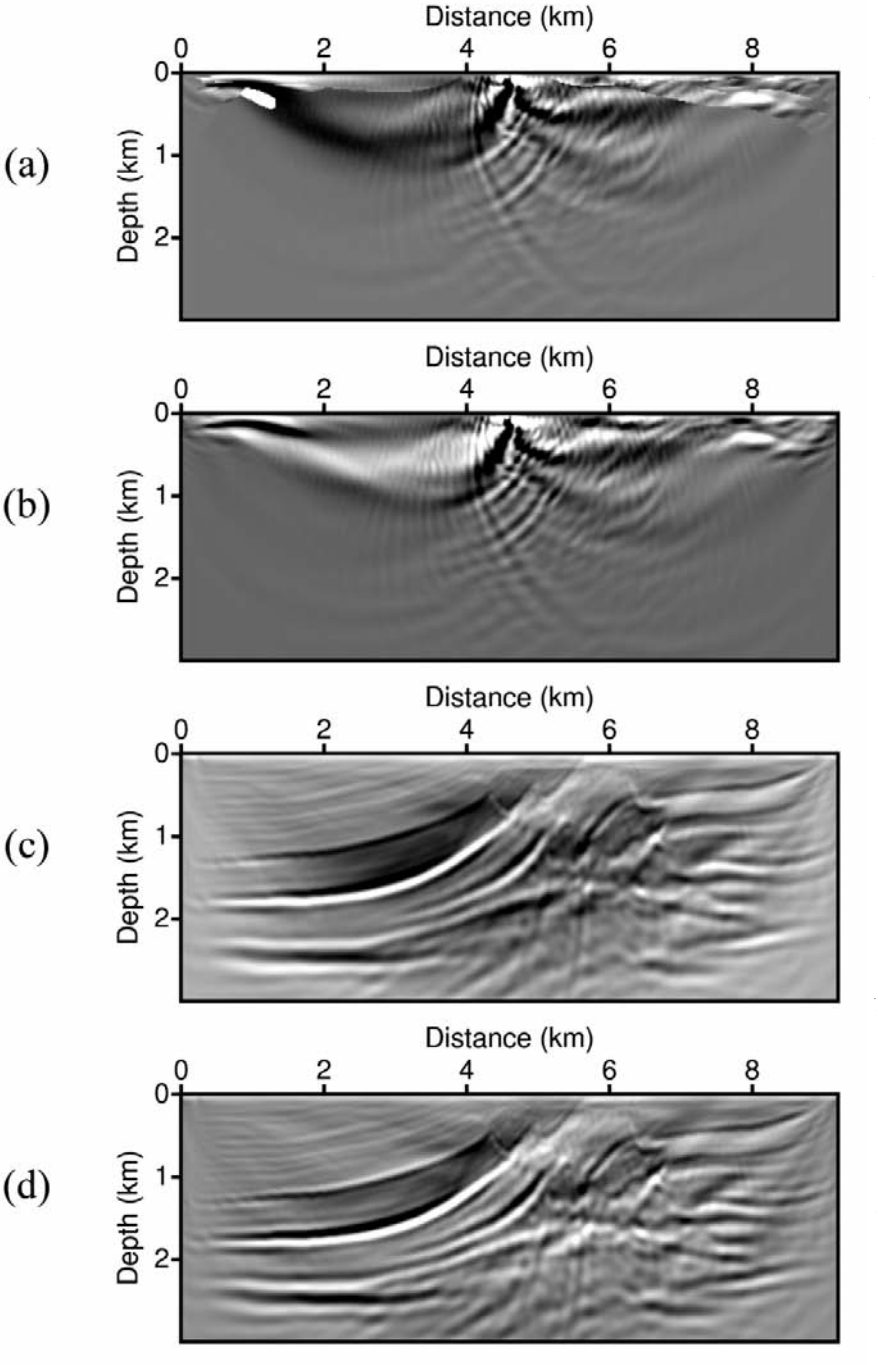


Fig. 5. The velocity update directions using one shot from the first iteration of (a) the excitation amplitude method and (b) the full cross-correlation method. The velocity update directions using all of the shots from the first iteration of (c) the excitation amplitude method and (d) the full cross-correlation method.

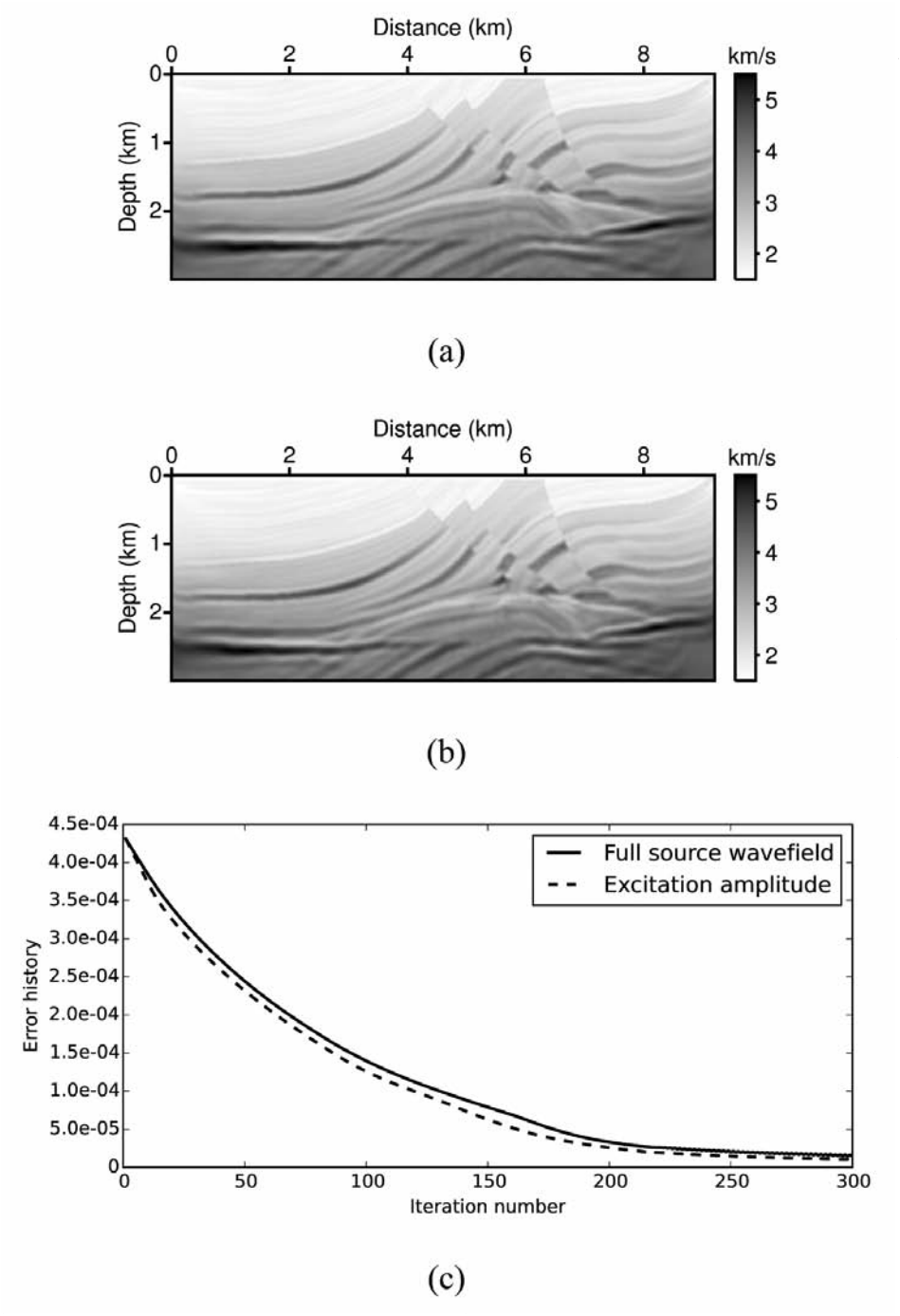


Fig. 6. The inversion results obtained using (a) the excitation amplitude method and (b) the full time samples. (c) The error histories from the full cross-correlation and excitation amplitude inversion methods.

to that from the full cross-correlation method. The error histories (Fig. 6c) suggest that the results obtained when using the excitation amplitude method are slightly better than the results obtained when using the full source wavefield. Note that this result may not be generalized for other datasets.

To examine the effects of discontinuities within the one-shot gradients, we reduced the number of observed shots to 10 and generated another dataset. The interval between the shots was 800 m. We inverted the dataset using the same inversion settings as before. Fig. 7 demonstrates that we can observe the discontinuities and acquisition footprints on the stacked gradient direction near

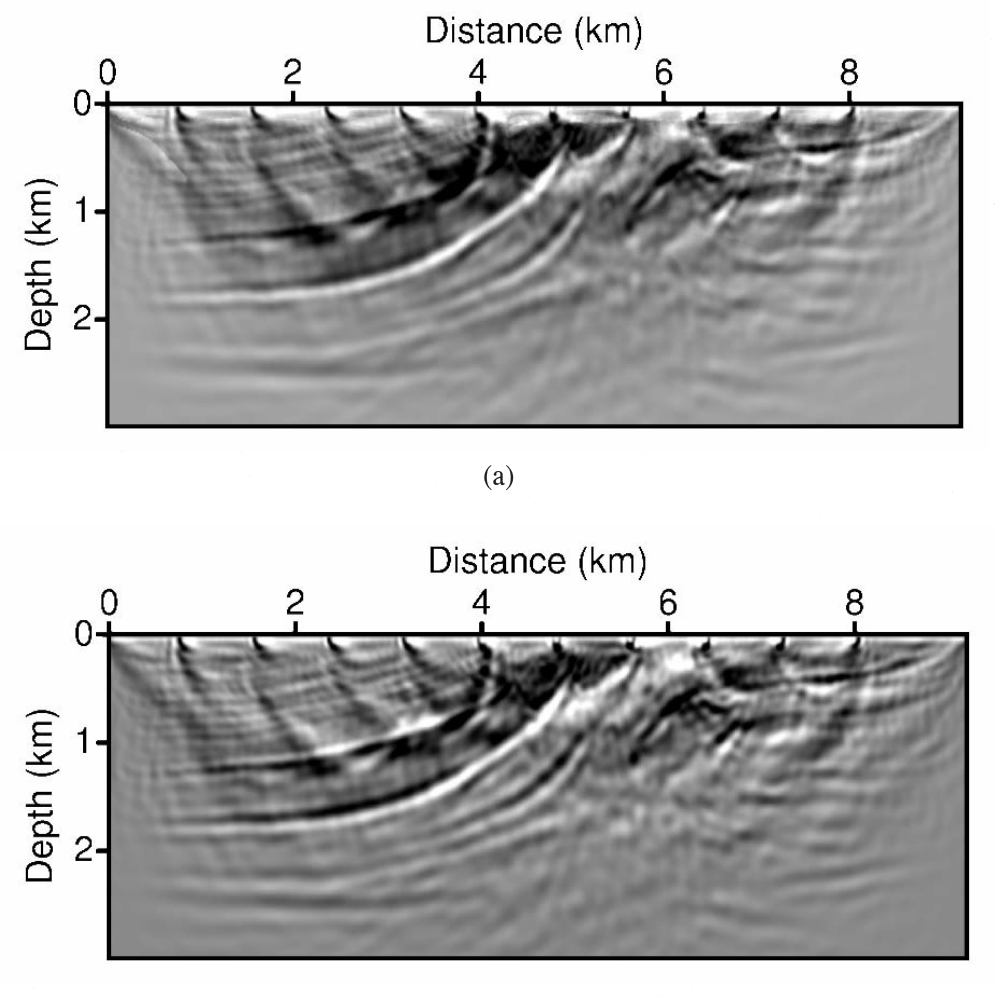


Fig. 7. The velocity update directions using 10 shots from the first iteration of (a) the excitation amplitude method and (b) the full cross-correlation method.

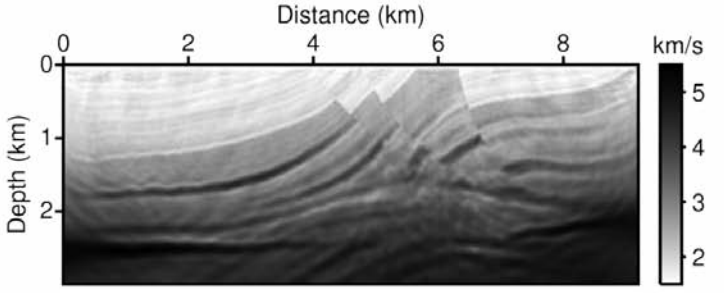
the surface. On the other hand, the discontinuities of the gradient direction were not obvious in the previous example that utilized 180 shots (Fig. 5c). Fig. 8 presents the inversion results. The result obtained using the excitation amplitude method is slightly inferior to that from the full cross-correlation method; however, the effects of the discontinuities are limited because the positions of the discontinuities change as the velocity is updated during the inversion process.

We also tested the excitation amplitude method using an overthrust velocity model (Fig. 9a). We detonated 399 shots and simulated the wave propagation for 8 s to generate the observed data. The time sampling rate was 1 ms. The maximum frequency of the source wavelet was 15 Hz and the grid size was 25 m. The number of grids was 801 by 187. We used a scheme incorporating an eighth-order and second-order finite-difference algorithm in space and time, respectively.

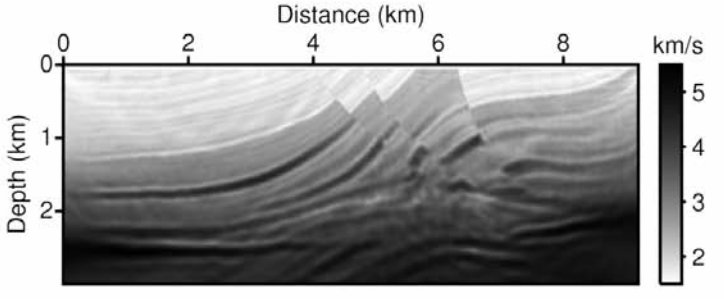
We inverted the observed data starting from a smooth initial velocity model (Fig. 9b). The memory required to save the full source wavefield is 4,572 MB, while that required for the excitation amplitude method is 1.2 MB. Figs. 10a and 10b show the inversion results obtained from using the excitation amplitude and the full cross-correlation methods. We can see that their results and error histories (Fig. 10) are similar to each other. Thus, we confirm that the excitation amplitude method can successfully approximate the full cross-correlation method in the Marmousi and overthrust examples.

DISCUSSION

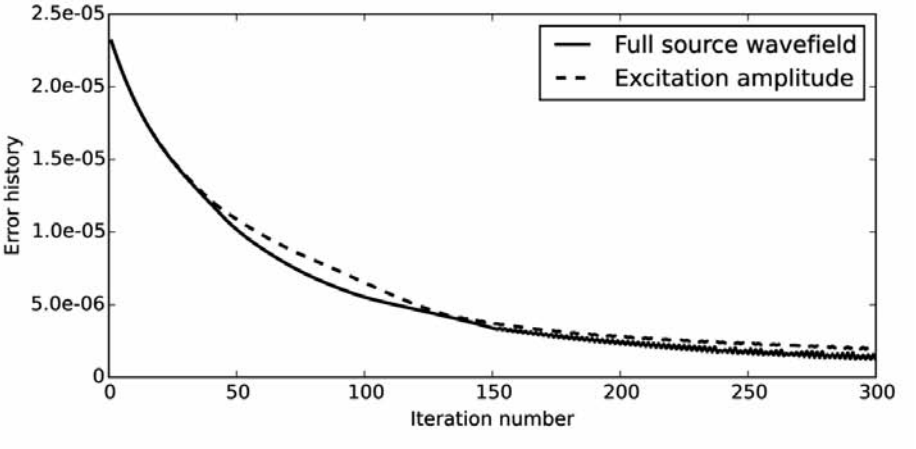
3D FWIs with tens of thousands of shots require immense computational resources. Although full 3D inversions are not affordable in our system, we can calculate the requisite memory burden to store the full source wavefield using the SEG/EAGE 3D salt model (Aminzadeh et al., 1994) with receivers at all of the surface grids. The grid size is 20 m and the number of grids is 676 $\times$ 676 $\times$ 201. When the maximum recording time is 10 s and the time sampling interval is 2 ms, the memory required to store the source wavefield is 1,711 GB. This wavefield must be saved to hard disk drives in many systems. Since the speed of the input/output (I/O) of a hard disk drive is much slower than that of RAM, a 3D inversion can present as an I/O-bound problem rather than as a CPU-bound problem, depending on the system. The memory required is less than 0.7 GB when using the excitation amplitude method, which eliminates the need for disk I/O in order to store the source wavefield. The inversion time can be reduced greatly when compared to that without subsampling since the excitation amplitude method does not require disk I/O. If multiple CPUs are used in a computing node, slow disk I/O can represent a severe bottleneck.



(a)

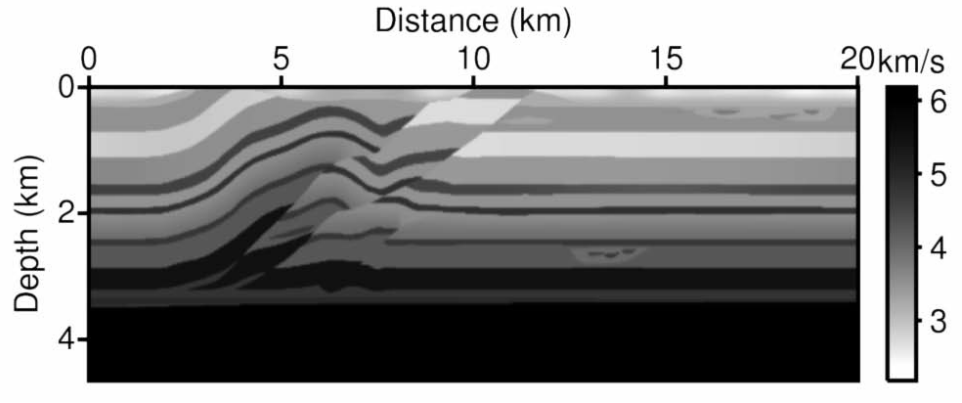


(b)

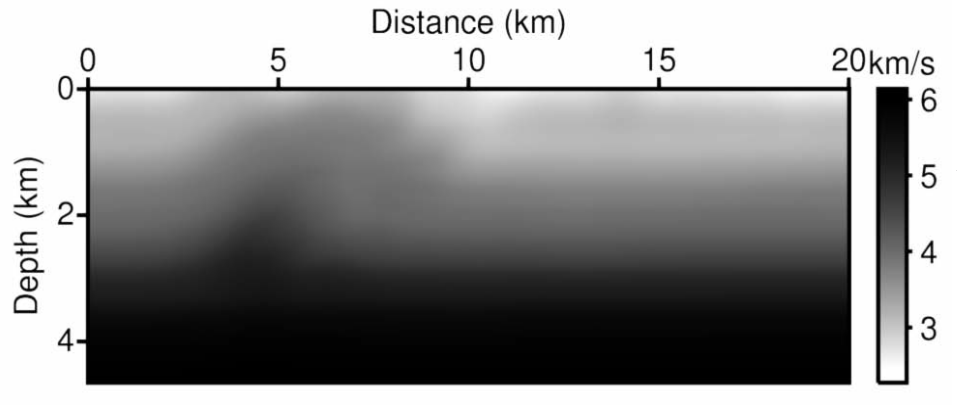


(c)

Fig. 8. The inversion results of the dataset containing 10 shots obtained using (a) the excitation amplitude method and (b) the full time samples. (c) The error histories from the full cross-correlation and excitation amplitude inversion methods.



(a)



(b)

Fig. 9. (a) The overthrust velocity model and (b) a smooth initial velocity model.

In an elastic FWI using multi-component data, the memory demand can be increased several times. When disk I/O is a limited factor, wavefield reconstruction methods (Anderson et al., 2012; Yang et al., 2015) can be used to further reduce the storage demand by sacrificing computational resources. Domain decomposition over several computing nodes can also be used in such a case to reduce or eliminate disk I/O; however, it will introduce communication burdens. On the other hand, the excitation amplitude method can reduce both memory demands and computation time by approximating the full cross-correlation algorithm to calculate the gradient direction.

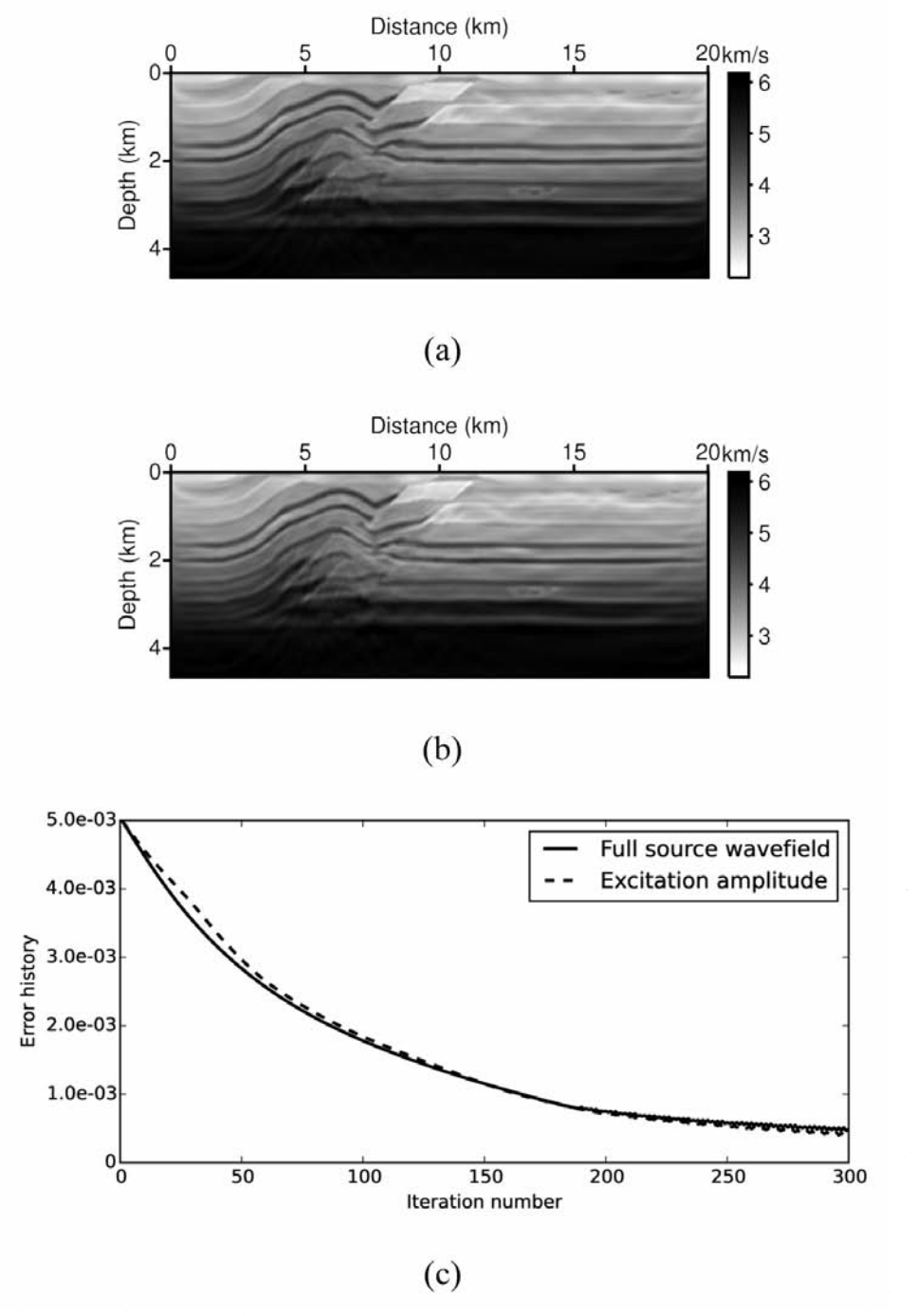


Fig. 10. The inversion results obtained using (a) the excitation amplitude method and (b) the full time samples. (c) The error histories from the full cross-correlation and excitation amplitude inversion methods.

One obvious limitation of the excitation amplitude method is the presence of multipathing within the source wavefield. In this regard, the excitation amplitude method can be compared to early arrival waveform tomography or high-frequency traveltime tomography methods (Sheng et al., 2006), which ignore multipathing in their objective functions. One significant difference is that the excitation amplitude method employs full wavefields in the objective function and solely approximates the cross-correlation procedure. As observed in the numerical examples, the excitation amplitude method yields higher-resolution results relative to those from early arrival tomography. A remedy to this multipathing problem is to adopt sparse cross-correlation using several peak values for each space grid (Nguyen and McMechan, 2015). In this sparse cross-correlation method, the memory burden is increased depending upon the number of peak values to account for multipathing.

CONCLUSIONS

We implemented an excitation amplitude method toward a time-domain FWI to calculate the gradient direction. We can reduce the memory demand by three orders of magnitude by adopting this method. As one can expect, the accuracy of the subsampling method depends on the subsampling interval. The excitation amplitude method uses only one sample that contains the maximum amplitude from the source wavefield for each grid point to calculate the cross-correlation between the source and receiver wavefields. Although the proposed method uses an approximated gradient direction, it can remove the memory burden in saving the source wavefield while maintaining the inversion results. Numerical examples suggest that the excitation amplitude method can be an efficient alternative to FWI using the full cross-correlation algorithm. Future research will involve applying the proposed method to elastic FWIs, 3D FWIs, and the inversions of field datasets.

ACKNOWLEDGMENTS

This research was supported by the Supercomputing Center/Korea Institute of Science and Technology Information (KSC-2014-C1-015), the Basic Research Project (16-3313) of the Korea Institute of Geoscience and Mineral Resources (KIGAM), which was funded by the Ministry of Science, ICT and Future Planning of Korea, and the Energy Resources Development Program of the Ministry of Trade, Industry and Energy.

REFERENCES

Aminzadeh, F., Burkhard, N., Nicoletis, L., Rocca, F. and Wyatt, K., 1994. SEG/EAEG 3-D modeling project: 2nd update. The Leading Edge, 13: 949-952.

Anderson, J.E., Tan, L. and Wang, D., 2012. Time-reversal checkpointing methods for RTM and FWI. Geophysics, 77(4): S93-S103.

Courant, R., Friedrichs, K. and Lewy, H., 1967. On the partial difference equations of mathematical physics. IBM J. Res. Developm., 11: 215-234.

Dussaud, E., Symes, W.W., Williamson, P., Lemaistre, L., Singer, P., Denel, B. and Cherrett, A., 2008. Computational strategies for reverse-time migration. Expanded Abstr., 78th Ann. Internat. SEG Mtg., Las Vegas: 2267-2271.

Griewank, A., 1992. Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation. Optimizat. Meth. Softw., 1: 35-54.

Griewank, A. and Walther, A., 2000. Algorithm 799: revolve: an implementation of checkpointing for the reverse or adjoint mode of computational differentiation. ACM Transact. Mathemat. Softw., 26: 19-45.

Ha, W., Chung, W. and Shin, C., 2012. Pseudo-Hessian matrix for the logarithmic objective function in full waveform inversion. J. Seismic Explor., 21: 201-214.

Kalita, M. and Alkhalifah, T., 2016. Full-waveform inversion using the excitation representation of the source wavefield. Expanded Abstr., 86th Ann. Internat. SEG Mtg.,....: 1084-1088.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations. Conf. Inverse Scatter.: Theory and Applications. Soc. Industr. Appl. Mathemat., Philadelphia: 206-220.

Nguyen, B.D. and McMechan, G.A., 2013. Excitation amplitude imaging condition for prestack reverse-time migration. Geophysics, 78(1): S37-S46.

Nguyen, B.D. and McMechan, G.A., 2015, Five ways to avoid storing source wavefield snapshots in 2D elastic prestack reverse time migration. Geophysics, 80(1), S1-S18.

Nichols, D.E., 1996. Maximum energy traveltimes calculated in the seismic frequency band. Geophysics, 61(1): 253-263.

Plessix, R.-E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. Geophys. J. Internat., 167, 495-503.

Sheng, J., Leeds, A., Buddensiek, M. and Schuster, G.T., 2006. Early arrival waveform tomography on near-surface refraction data. Geophysics, 71(4): U47-U57.

Shin, C., Jang, S. and Min, D., 2001. Improved amplitude preservation for prestack depth migration by inverse scattering theory. Geophys. Prosp., 49: 592-606.

Shin, C., Ko, S., Marfurt, K.J. and Yang, D., 2003. Wave equation calculation of most energetic traveltimes and amplitudes for Kirchhoff prestack migration. Geophysics, 68: 2040-2042.

Sun, W. and Fu, L.-Y., 2013. Two effective approaches to reduce data storage in reverse time migration. Comput. Geosci., 56: 69-75.

Symes, W.W., 2007. Reverse time migration with optimal checkpointing. Geophysics, 72(5): SM213-SM221.

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

Versteeg, R., 1994. The Marmousi experience: Velocity model determination on a synthetic complex data set. The Leading Edge, 13: 927-936.

Wallis, J.W., Miller, T.R., Lerner, C.A. and Kleerup, E.C., 1989. Three-dimensional display in nuclear medicine. IEEE Transact. Medic. Imag., 8: 297-230.

Yang, P., Gao, J. and Wang, B., 2015. A graphics processing unit implementation of time-domain full-waveform inversion. Geophysics, 80(3), F31-F39.