ACOUSTIC FULL WAVEFORM INVERSION USING DISCRETE COSINE TRANSFORM (DCT)

SUMIN KIM[1], WOOKEEN CHUNG[1,2] and JONGHYUN LEE[3]

[1] *Department of Convergence Study on the Ocean Science and Technology, Ocean Science and Technology (OST) School, Korea Maritime and Ocean University, Busan, South Korea.*
[2] *Department of Energy and Resource Engineering, Korea Maritime and Ocean University, Busan, South Korea.*
[3] *Department of Civil and Environmental Engineering, and Water Resources Research Center, University of Hawaii at Manoa, Honolulu, HI 96822, U.S.A.*
jonghyun.lee@hawaii.edu

(Received August 10, 2020; accepted February 28, 2021)

ABSTRACT

Kim, S., Chung, W.K. and Lee, J.H., 2021. Acoustic full waveform inversion using Discrete Cosine Transform (DCT). *Journal of Seismic Exploration*, 30: 365-380.

Full waveform inversion (FWI) has been implemented widely to reconstruct high-quality velocity model in the subsurface. However, due to the advance in geophysics data acquisition with increasing demand for high dimension velocity model estimation, computational costs become prohibitive if not impossible. To alleviate computational burdens, we incorporate discrete cosine transform, one of the most widely used compression techniques in image processing, into FWI while estimation results are still kept comparable to those from the full-model based FWI. The unknown velocity fields are transformed in the DCT domain and only a small number of DCT coefficients are included in FWI to describe common velocity model features without losing much reconstruction accuracy. Generally, DCT coefficients can be chosen specific window (e.g., square window). However, there can be more dominant DCT coefficients out of this specific window. To take more dominant DCT coefficients, we sort the absolute value of DCT coefficients in descending order and determine DCT coefficients with compression ratio. Through the comparison reconstructed velocity models, our proposed method generate more accurate velocity model than case of using square window. We investigate the applicability of our DCT-based FWI method to two numerical examples. It is shown that the proposed method can reduce the computational cost significantly and produce satisfactory results. Through these results, we expect that our FWI method can contribute to enhance computational efficiency for FWI with enormous amount of unknown parameters.

KEY WORDS: acoustic, seismic, full waveform inversion (FWI), Discrete Cosine Transform (DCT).

09634-0651/21/$5.00

INTRODUCTION

Full waveform inversion (FWI) is one of the promising seismic imaging techniques that enable to obtain high-resolution velocity information in the subsurface. To reconstruct subsurface velocity model, FWI typically performs minimization of a misfit function that is defined as the difference between modeled and observed data (Laily, 1983; Tarantola, 1984; Pratt et al., 1998; Virieux and Operto, 2009). In spite of its ability to recover high-resolution velocity models, FWI has been regarded as challenging because of 1) the large number of forward model run proportional to the number of discretization grids or the number of observations, and 2) the computational costs to construct, store and solve the linear system whose size is a function of the number of discretization grids or the number of observations. To overcome these computational challenges, many techniques from linear algebra/optimization community such as back-propagation (e.g., Tarantola, 1984) and limited-memory BFGS (e.g., Nocedal and Wright, 2006) have been applied for FWI.

Although several efficient methods can handle intermediate-scale FWI problems successfully, it is still difficult to implement FWI to large-scale with a tremendous amount of seismic data, high dimensional problems with high-resolution survey with shorter spatial intervals. Thus, the computational burden grows dramatically, especially when processing the observation data, constructing Jacobian and Hessian matrices, and computing the linear system for model updates. To address these issues, a number of studies have been suggested for efficient FWI implementation by reducing the number of observations and/or the number of unknowns while the final inverse solution is kept close to one obtained from full-model based FWI. One of widely used data reduction approaches is simultaneous-source FWI method (Krebs et al., 2009; Gao et al., 2010; Ben-Hadj-Ali et al., 2011; Son et al., 2012), which uses simultaneous-source gather that is made from a lot of source gather with encoding scheme. Habashy et al. (2011) propose a source and receiver compression technique by checking the amount of redundancy in the data. These techniques can significantly reduce the computational costs and storage by compressing the number of receivers as well as sources. Recently, the compressive sensing technique have also been applied to FWI for enhancing efficiency by subsampling data in specific domain (Li et al., 2012, 2016; Zhu et al., 2017).

High-resolution velocity model reconstruction is another challenging factor for computational burden. With high-resolution survey with fine spatial grid intervals, the number of unknowns in velocity model can increase tremendously. It requires that huge computational environment is needed for obtaining high-resolution velocity model. In petroleum and groundwater engineering, dimension reduction of the unknown subsurface properties has been investigated recently (Jafarpour and McLaughlin, 2008; Lee et al., 2016). However, only few studies have investigated the applicability dimension reduction in FWI, for example, Lin et al. (2012) utilized a discrete wavelet transform to decrease the number of unknowns in velocity model during FWI. They choose a few dominant Daubechies

wavelet basis vectors to perform model compression and reconstruct model successfully. While they mentioned Daubechies 4 wavelet basis performed better than Harr basis in their own study, no other wavelet basis was not tested for performance comparison in a systematic way. In this study, we propose a new FWI method using discrete cosine transform (DCT) for efficient velocity model reconstruction. DCT is widely used to compress image files (e.g. JPEG) in image processing applications for its flexibility and asymptotically convergence to optimal Karhunen-Loeve transform (Jain, 1989). In geophysics, DCT has been applied to enhance efficiency for data processing (Zhou and Li, 2013; Dalmau et al., 2014; Zhu et al., 2015). Zhou and Li (2013) try to reduce the burden of large capacity of data by compressing data in FWI on ground-penetrating radar (GPR) data. For removing noise signal in seismic data, DCT can be used to make base dictionary (i.e., basis vectors) for sparsity-promoting dictionary learning (Zhu et al., 2015; Chen, 2017). Furthermore, DCT can be applied on top of CS techniques to exploit the sparsity structures in the subsurface (Baraniuk and Steeghs, 2017; Zhu et al., 2017). Since DCT has been previously used for data compression in FWI, we investigate its applicability on reduction of unknowns in the velocity model. In our FWI algorithm, the Gauss-Newton method is used to minimize the misfit function and computational bottleneck mostly comes from large Jacobian and Hessian matrix constructions since their dimension is proportional to the number of unknowns. Thus, reducing the number of unknowns is important to implement FWI efficiently.

In following sections, we begin brief explanation about DCT and its characteristics. Then we give a detailed account for the problem on size reduction. Also, essential parts of our FWI algorithm implementation in DCT domain is delineated. By comparing the reconstructed velocity models obtained from fixed square window and our approach, we confirm the accuracy of reconstructed velocity model and the applicability of our FWI algorithm. After that we implement our FWI algorithm to two synthetic models followed by discussion of results.

THEORY

Discrete Cosine Transform

Among several discrete cosine transform (DCT) formulations, generally, DCT-II and DCT-III are widely used as forward DCT and inverse DCT respectively. Two-dimensional forward and inverse DCT can be expressed as follows (Jain, 1989):

$$V(p,q) = c(p)c(q)\sum_{x=0}^{M-1}\sum_{z=0}^{N-1} v(x,z)\cos\frac{\pi(2x+1)p}{2M}\cos\frac{\pi(2z+1)q}{2N}, \quad (1)$$

$$v(x,z) = \sum_{p=0}^{M-1}\sum_{q=0}^{N-1} c(p)c(q)V(p,q)\cos\frac{\pi(2x+1)p}{2M}\cos\frac{\pi(2z+1)q}{2N}, \quad (2)$$

where $v(x,z)$ is the original velocity model, p,q is the coordinate indexes (integer) in transformed domain, $V(p,q)$ is the transformed velocity model of $v(x,z)$, M,N is the number of grid in x and z-axis, and $c(p)$, $c(q)$ are transform coefficient.

$$c(p) = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \le p \le M-1, \end{cases} \quad (3)$$

$$c(q) = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \le q \le N-1. \end{cases} \quad (4)$$

DCT has a useful characteristic for data compression. It is that almost of energies in DCT domain are concentrated near low frequency (Lam and Goodman, 2000). In other words, when data are transformed using DCT, most of dominant DCT coefficients are located near low frequency. Even if we discard DCT coefficients at high frequency, we can reconstruct data similar to original data. Because DCT coefficients at high frequency are less dominant than its near low frequency to reconstruct data. For this reason, DCT is mostly used for compressing image.

In DCT domain, each DCT coefficient corresponds to a specific frequency. Fig. 1 shows the true velocity model used in this study and its DCT coefficients. It is worth noting that most of energies in DCT domain are concentrated near low frequency (Fig. 1b) so that the true model can be reconstructed with only a small number of DCT coefficients. Fig. 2 shows several DCT bases nearby DC component. Each DCT basis consist of cosine function with its natural frequency (or period). The model is reconstructed by summing the basis with appropriate coefficients.

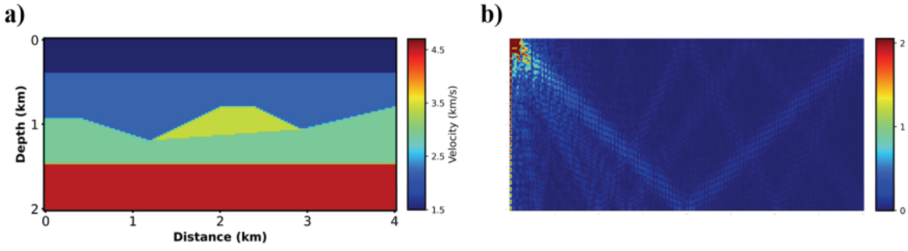

Fig. 1. True velocity model and transformed velocity model. a) Simple velocity model with syncline and big anomaly, b) transformed velocity model (this figure was depicted in absolute value and scaled by 200 times of maximum value).

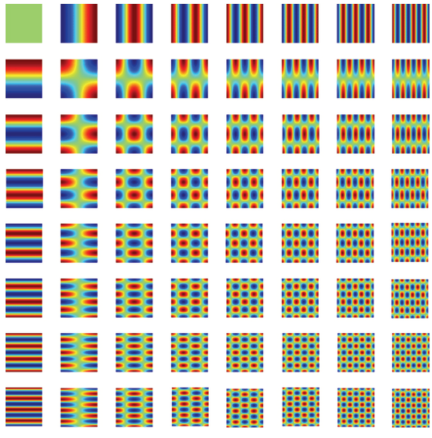

Fig. 2. 8 x 8 discrete cosine transform bases nearby DC component.

DCT-based dimension reduction in FWI

Compression properties in DCT allow dimension reduction of unknown velocity in FWI. A velocity model can be transformed in the DCT domain as below:

$$\mathbf{m}' = \mathcal{F}\mathbf{m}\,, \tag{5}$$

where $\mathbf{m}$ is $M \times N$ original unknown velocity model, $\mathbf{m}'$ is $M \times N$ unknown DCT coefficients vector transformed from the original velocity model and $\mathcal{F}$ is forward discrete cosine transform matrix. For reducing dimension in FWI, we only use dominant DCT coefficients consisting of

velocity model. In other words, we select the DCT coefficient based on the absolute value of the DCT coefficients. Generally, dominant DCT bases are located nearby low frequency. Dominant DCT coefficients can be chosen with specific window such as square box window at low frequency. However, there can be more dominant DCT coefficients out of this specific window. To use more dominant DCT coefficients with same number of DCT coefficients, we first sort the absolute value of current estimated DCT coefficients in descending order as:

$$\tilde{\mathbf{m}} = \mathcal{S}(\mathbf{m}'), \tag{6}$$

where $\tilde{\mathbf{m}}$ is $MN\times1$ sorted DCT coefficients vector in descending, $\mathcal{S}$ is sorting operator. Then we determine n_p that is the number of DCT coefficients used in the inversion as:

$$n_p = int\left[\left(1-\frac{r_{comp}}{100}\right)MN\right], \tag{7}$$

where r_{comp} is compression ratio and M, N is the number of grids in x and z-axis in finite difference scheme with regular grid interval, respectively. Eq. (7) indicates the number of DCT coefficients defined as compression ratio r_{comp}. n_p is generally chosen to a very small value compared to MN with a high compression ratio r_{comp} such as 5% or 10%. We select the number of DCT coefficients with Eq. 7 in DCT coefficients array. Fig. 3 shows our strategy for determining number of DCT coefficients. After determining the number of DCT coefficients for FWI, we use these DCT coefficients for updating the velocity model in final step at each iteration.

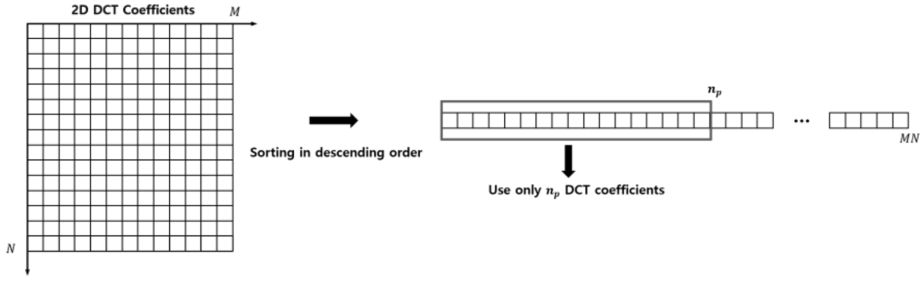

Fig. 3. A mimetic diagram of our strategy for reducing unknowns of velocity model.

Full waveform inversion using Gauss-Newton method on DCT coefficients

FWI is generally implemented by updating unknowns through gradient descent or Gauss-Newton methods to minimize a nonlinear misfit function iteratively. Misfit function based on the *l*2-norm is given by Eq. 8,

$$\mathrm{E} = \frac{1}{2}\sum_{nshot} \|\mathbf{u} - \mathbf{d}\|_2^2 , \tag{8}$$

where E is the misfit function, $\mathbf{u}$ and $\mathbf{d}$ are the modeled and recorded common shot gather, respectively. Because we reduced the number of unknowns of velocity model via DCT, we can apply the Gauss-Newton method in FWI. It can be implemented with few DCT coefficients. For performing FWI with Gauss-Newton scheme, we computed the Jacobian matrix directly. To construct Jacobian matrix, we calculate Fréchet derivative numerically,

$$\frac{\partial \boldsymbol{u}}{\partial \widetilde{\mathrm{m}}_k} \approx \frac{\boldsymbol{u}[\mathcal{F}^{-1}(\widetilde{\mathbf{m}} + \Delta\widetilde{\mathrm{m}}_k)] - \boldsymbol{u}[\mathcal{F}^{-1}\widetilde{\mathbf{m}}]}{\Delta\widetilde{\mathrm{m}}_k}, \tag{9}$$

where $\frac{\partial \boldsymbol{u}}{\partial \widetilde{\mathrm{m}}_k}$ is Fréchet derivative about DCT coefficient, $\mathcal{F}^{-1}$ is inverse discrete cosine transform operator and $\Delta\widetilde{\mathrm{m}}_k$ is perturbation of DCT coefficient at k-th point in $\widetilde{\mathbf{m}}$. With $\frac{\partial \boldsymbol{u}}{\partial \widetilde{\mathrm{m}}_k}$ obtained from Eq. (9), we can construct Jacobian matrix as follows:

$$\mathbf{J} = \left[\frac{\partial \boldsymbol{u}}{\partial \widetilde{m}_1} \quad \frac{\partial \boldsymbol{u}}{\partial \widetilde{m}_2} \quad \cdots \quad \frac{\partial \boldsymbol{u}}{\partial \widetilde{m}_{n_p}}\right], \tag{10}$$

where $\mathbf{J}$ is $n_{rcv}\times\ n_p \times nt$ Jacobian matrix that consist of Fréchet derivative about DCT coefficient, n_{rcv} is number of receivers, and nt is number of time series. After constructing Jacobian matrix, we can also determine gradient direction and construct approximate Hessian matrix via eq. (11) and eq. (12), respectively.

$$\frac{\partial \mathrm{E}}{\partial \widetilde{\mathbf{m}}} = \mathbf{J}^T \Delta\mathbf{d} . \tag{11}$$

$$\mathbf{H_a} = \mathbf{J}^T\mathbf{J} , \tag{12}$$

where $\frac{\partial \mathrm{E}}{\partial \widetilde{\mathbf{m}}}$ is gradient direction, $\Delta\mathbf{d}$ is residual vector defined by $\mathbf{u} - \mathbf{d}$, $\mathbf{H}_{\boldsymbol{a}}$ is $n_p \times n_p$ approximate Hessian matrix and superscript T is transpose operator. Now, we can obtain the update vector $\Delta\widetilde{\mathbf{m}}$:

$$\Delta\tilde{\mathbf{m}} = (\mathbf{H}_a + \lambda\mathbf{I})^{-1}\frac{\partial \mathrm{E}}{\partial \tilde{\mathbf{m}}}, \tag{13}$$

where λ is damping term, defined as 1% of the absolute of largest value of $\mathbf{H}_a$. With update vector $\Delta\tilde{\mathbf{m}}$, we can finally obtain updated velocity model by transforming $\tilde{\mathbf{m}}$ via inverse DCT.

$$\mathbf{m}^{k+1} = \mathcal{F}^{-1}(\tilde{\mathbf{m}}^k + \Delta\tilde{\mathbf{m}})\,. \tag{14}$$

At every iteration, the velocity model $\mathbf{m}$ is obtained by transforming inversely updated DCT coefficients $\tilde{\mathbf{m}}$. In summary, our FWI algorithm is implemented as a flowchart shown in Fig. 4. Firstly, velocity model is transformed by DCT, and the number of unknown parameters is determined with compression ratio from Eq. 7. After that, Jacobian and Hessian matrix are constructed by calculating the Fréchet derivative about DCT coefficient and updating the selected DCT coefficients. Lastly, updated velocity model is obtained by transforming inversely the updated DCT coefficients.

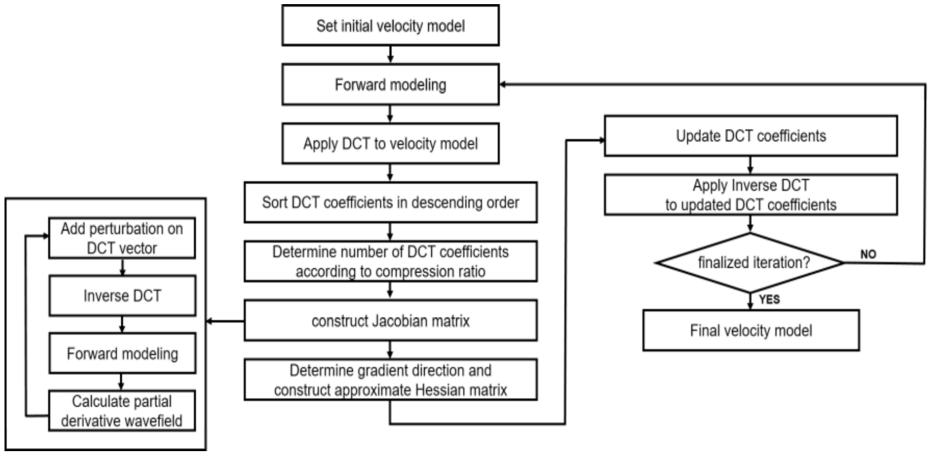

Fig. 4. A flowchart of our FWI approach in DCT domain.

NUMERICAL EXAMPLES

In this part, we investigate the applicability and efficiency of our FWI algorithm to synthetic velocity model reconstruction. Two synthetic models are tested: 1) simple velocity model with syncline and big anomaly, 2) modified Marmousi-2 model.

Model 1: simple velocity model with syncline and big anomaly

1) Comparison the reconstructed velocity model

Before testing our FWI algorithm, we first compare the reconstructed velocity model obtained from DCT coefficients selected by fixed square window and our approach. Fig. 1a is the true velocity model used in this section. The model has three layers with big trapezoidal velocity anomaly in the center. The dimensional scale of this velocity model is 4 km × 2 km. The number of grids of velocity model is 201 × 101. The thickness of water layer is 0.38 km. We assume that the velocity information of water layer is known as 1.5 km/s hence the part of water layer is not included in DCT and FWI process. Accordingly, the number of total unknowns is 16,281. With a compression ratio r_{comp} of approximately 90 percent, we use 1,600 unknowns in this test for comparing reconstructed velocity models between case of sorting in descending order and fixed square window. In case of fixed square window, we set square window 40 × 40 in DCT domain at near frequency and choose DCT bases in window. Fig. 5 shows chosen 1,600 selected DCT coefficients through our approach and fixed square window and reconstructed models, respectively. When comparing reconstructed velocity models, velocity model obtained from our approach is more similar to true model (Fig. 1a) and has less artifacts. We calculate pixel-wise RMS (Root-Mean Squares) error with true model in entire zone. The equation for calculating RMS error is given by

$$E_{RMS} = \sqrt{\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}(m_{true}(x,y) - m(x,y))^2}\,. \tag{16}$$

where E_{RMS} is pixel-wise RMS error between true and estimation model, m_{true} is true velocity model and m is input velocity model for calculating error. In case of our approach and fixed square window, pixel-wise RMS error of reconstructed model is 0.03628 and 0.08945, respectively. From this pixel-wise RMS error, we can tell that velocity model reconstructed by our approach is more approximated to the true model than case of using fixed square window. Fig. 6 shows reconstructed velocity models with increasing number of used DCT bases. The more DCT bases are included, the closer reconstructed velocity model is to true velocity model. Through comparing reconstructed velocity model with same number of bases, we can also find that proposed method is more accurate than case of using DCT bases chosen with fixed square window for reconstructing velocity model. It means that sorted basis case has more dominant bases than case using chosen fixed window with same number of DCT bases. Successfully reconstructed models indicate that we may be able to reconstruct the velocity model by estimating the first few DCT coefficients instead of all the velocity values at the grids. With this result, we expect that FWI with our approach can generate satisfactory inversion result by updating DCT coefficients. So, we implemented FWI on DCT coefficients selected by our approach.

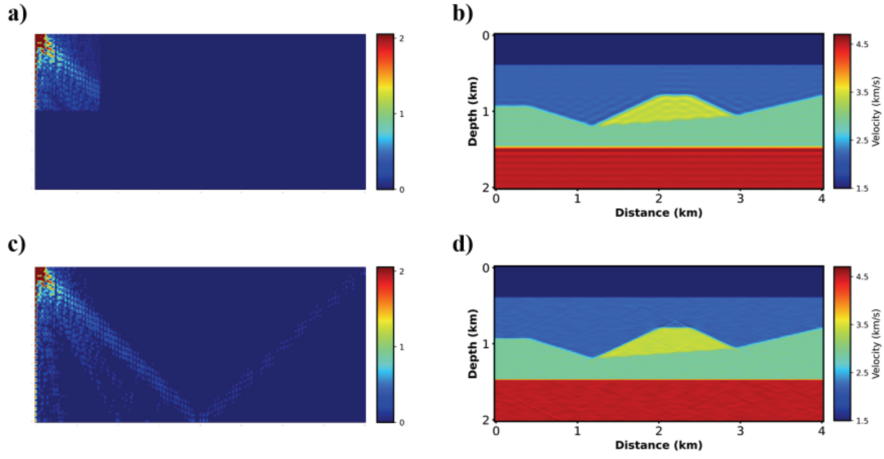

Fig. 5. Selected DCT coefficients and reconstructed velocity model. a) selected 1,600 DCT coefficients via fixed square window 40 × 40, b) reconstructed velocity model using a), c) selected 1,600 DCT coefficients via our approach, and d) reconstructed velocity model using c).

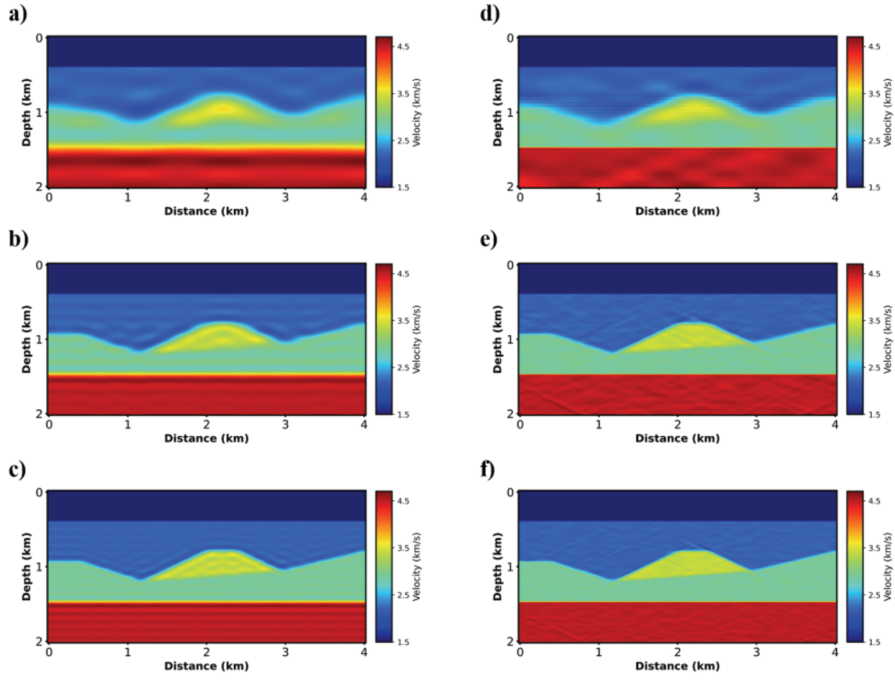

Fig. 6. Comparison of the reconstructed velocity model obtained from the case of chosen fixed square window and our approach. a-c) reconstructed model with DCT bases selected by square window (a: 10 × 10, b: 20 × 20, c: 30 × 30 bases, respectively), d-f) reconstructed velocity model with DCT bases selected by our approach (d: 100, e: 400, f: 900 bases, respectively).

2) Implementation of FWI with DCT

We apply our FWI algorithm to simple velocity model with syncline and anomaly. For performing our algorithm, we used 21 shots and 201 receivers. They were deployed with equidistant intervals, 0.2 km and 0.02 km, respectively. The first derivative Gaussian wavelet was used as source wavelet and its maximum frequency is 12.5 Hz. We used 1,600 DCT coefficients that were selected by our approach. Fig. 7a shows that initial velocity model. We used smoothed velocity model as initial model for implementing our algorithm. The number of iterations of inversion process is 10th. Fig. 7b is inverted velocity models from our FWI algorithm, respectively. When comparing with reconstructed velocity models (Fig. 5d), our inversion result is very satisfactory despite of a small number of unknowns. Big anomaly is inverted well, and especially boundary between anomaly and syncline layer can be identified. Fig. 8 is normalized error between modeled data and observed data. Error value reaches nearly 0 percent after finishing FWI process. From this result, we could confirm that FWI with a small number of DCT coefficients selected by our approach can produce the satisfactory inversion results.

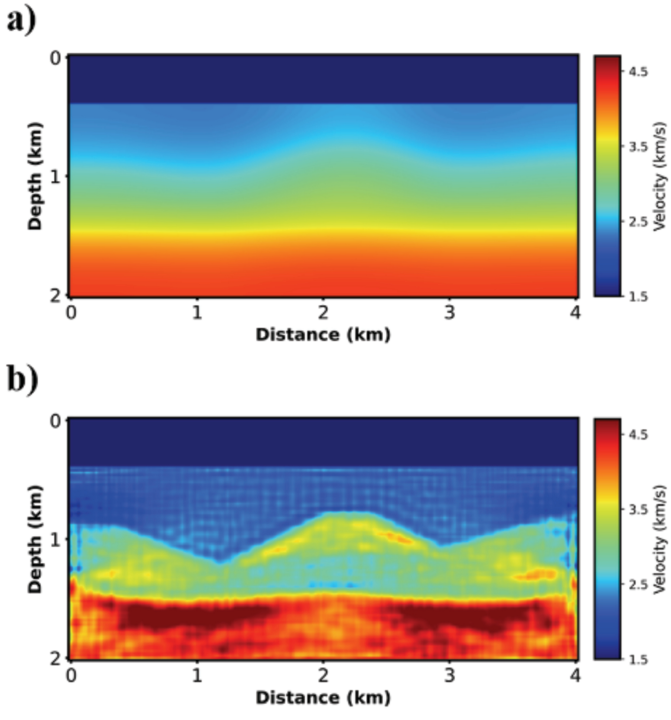

Fig. 7. a) initial velocity model, b) FWI results after 10 iterations obtained by our approach.

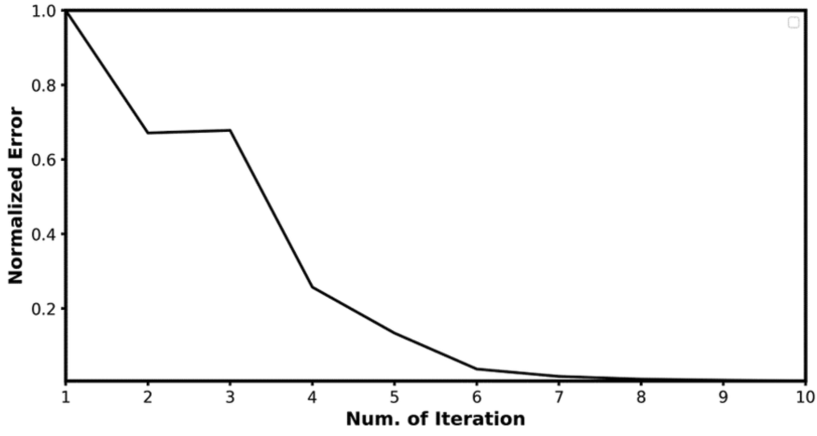

Fig. 8. Normalized RMS error curve on our FWI algorithm.

Model 2: modified Marmousi-2 model

For considering complex geological structure to test our FWI algorithm, we also applied our FWI algorithm to modified Marmousi-2 model (Martin et al., 2006). Fig. 9a is true modified Marmousi-2 velocity model. The dimensional scale of this modified Marmousi-2 velocity model is 9 km × 3.5 km and spatial interval is 0.02 km. The number of grids of this velocity model is 451 × 176. However, as mentioned earlier, we only performed our FWI algorithm without water layer. The water layer is a thickness about 0.44 km, its number of grid is 23. Thus, the net number of grid is 451 × 153 (total: 69,003). With set compression ratio r_{comp} as 92.5 percent in this FWI case, we only use 5,175 unknowns for FWI in DCT domain. Fig. 9b indicates reconstructed modified Marmousi-2 velocity model. It seems to similar with true velocity model (Fig. 9a). To implement our FWI algorithm on this modified Marmousi-2 model, we used 31 shots and 451 receivers. The shots were exploded with 0.3 km interval and receivers were placed with equidistant spatial intervals, 0.02 km. Fig. 10a shows initial velocity model for FWI. After implementing FWI to 14 iteration, the inverted velocity model is obtained as Fig. 10b. The inversion result delineates high-resolved not only shallow geologic structure but deep part. We note that inverted velocity model has high-resolution result despite of a small number of unknowns compared to true velocity model (Fig. 9a). Fig. 11 shows the normalized error curve in data-side as the result of our FWI algorithm. After 14 iteration, error reaches nearly 1 percent.

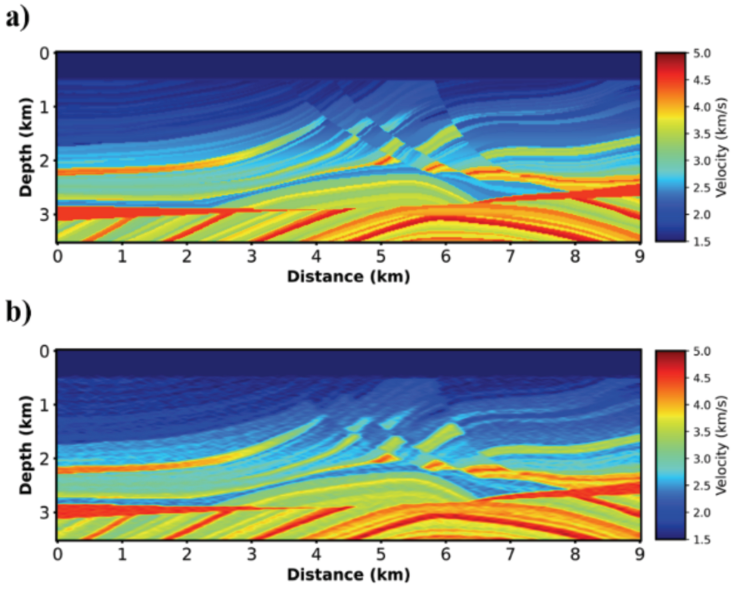

Fig. 9. a) true modified Marmousi-2 model, b) reconstructed modified Marmousi-2 velocity model according to compression ratio for 92.5 percent.

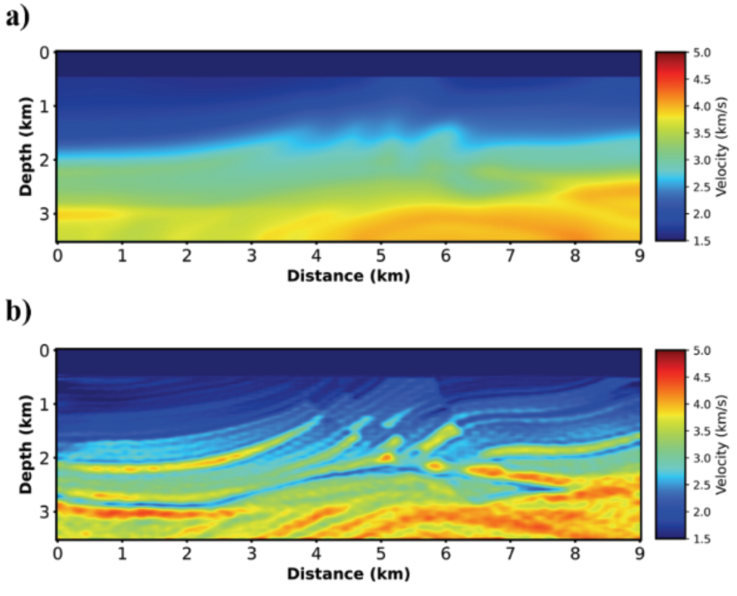

Fig. 10. a) initial velocity model for testing our FWI algorithm, b) inverted velocity model after 14 iteration.

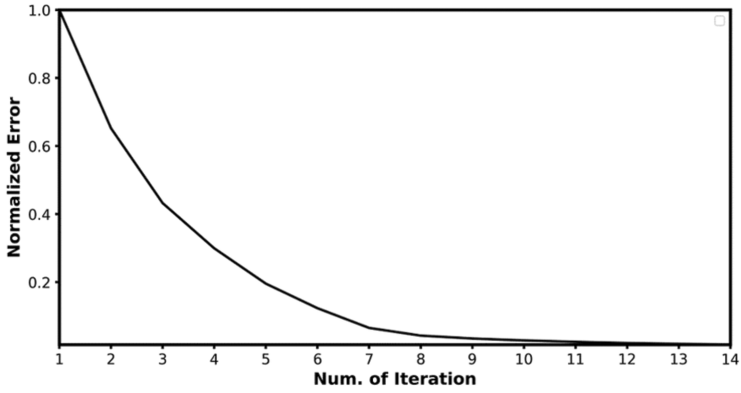

Fig. 11. Normalized error curve as the result of our FWI algorithm on modified Marmousi-2 model.

CONCLUSIONS

We present FWI algorithm with reduced number of unknowns of velocity model. The number of unknowns can be reduced by applying discrete cosine transform (DCT) on velocity model. After transforming velocity model, DCT coefficients are sorted in descending order and chosen based on compression ratio. By sorting DCT coefficients, we can use more dominant DCT bases than case of using DCT bases chosen with window (e.g., square window) in same number of DCT coefficients. This means that our proposed method is more accurate for reconstructing velocity model. With this approach, we only use a small number of unknowns for FWI. Through reconstructed model using a few bases, we could identify the applicability of our FWI algorithm. For testing our FWI algorithm, we implement our FWI algorithm on simple velocity model and modified Marmousi-2 model. These inversion results delineate that our FWI algorithm can be generate satisfactory results despite of a small number of unknowns. Through these results, we think that our FWI algorithm can be a cornerstone for enhancing computational efficiency in FWI. We expect that our FWI algorithm will have a great strength for 3D FWI. Thus, expanding our FWI algorithm to 3D FWI is needed as further research. Although we succeeded to reduce the number of unknowns considerably for FWI, our approach is needed to enhance efficiency. Because we calculated directly Fréchet derivative and constructed gradient direction, approximate Hessian matrix. For enhancing efficiency of our FWI algorithm, efficient optimization method (e.g., preconditioned conjugate-gradient method, limited-memory BFGS method) or techniques such as back-propagation technique are required to construct gradient direction or Hessian matrix efficiently in our FWI algorithm.

ACKNOWLEDGEMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No.20182510102470). Jonghyun Lee was supported by the Hawai'i Experimental Program to Stimulate Competitive Research (EPSCoR) provided by the National Science Foundation Research Infrastructure Improvement (RII) Track-1: 'Ike Wai: Securing Hawai'i's Water Future Award OIA 1557349 and Faculty Research Participation Program at the U.S. Engineer Research and Development Center, Coastal and Hydraulics Laboratory administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. department of Energy and ERDC.

REFERENCES

Baraniuk, R.G. and Steeghs, P., 2017. Compressive sensing: A new approach to seismic data acquisition. The Leading Edge, 36: 642-645.

Ben-Hadj-Ali, H., Operto, S. and Virieux, J., 2011. An efficient frequency-domain full waveform inversion method using simultaneous encoded sources. Geophysics, 76: R109-R124.

Chen, Y., 2017. Fast dictionary learning for noise attenuation of multidimensional seismic data. Geophys. J. Internat., 209: 21-31.

Dalmau, F.R., Hanzich, M., de la Puente, J. and Gutiérrez, N., 2014. Lossy data compression with DCT transforms. EAGE Workshop on High Performance Computing for Upstream, HPC30.

Gao, F., Atle, A. and Williamson, P., 2010. Full waveform inversion using deterministic source encoding. Expanded Abstr., 80th Ann. Internat. SEG Mtg., Denver: 1013-1017.

Habashy, T.M., Abubakar, A., Pan, G. and Belani, A., 2011. Source-receiver compression scheme for full-waveform seismic inversion. Geophysics, 76: R95-R108.

Jafarpour, B. and McLaughlin, D.B., 2008. History matching with an ensemble Kalman filter and discrete cosine parameterization. Computat. Geosci., 12: 227-244.

Jain, A.K., 1989. Fundamentals of Digital Image Processing. Prentice Hall, Englewood Cliffs, NJ.

Krebs, J.R., Anderson, J.E., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A. and Lacasse, M.D., 2009. Fast full-wavefield seismic inversion using encoded sources. Geophysics, 74: WCC177-WCC188.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application, Society for Industrial and Applied Mathematics, Expanded Abstr.: 206-220.

Lam, E.Y. and Goodman, J.W., 2000. A mathematical analysis of the DCT coefficient distributions for images. IEEE Transact. Image Process., 9: 1661-1666.

Lee, J., Yoon, H., Kitanidis, P.K., Werth, C.J. and Valocchi, A.J., 2016. Scalable subsurface inverse modeling of huge data sets with an application to tracer concentration breakthrough data from magnetic resonance imaging. Water Resour. Res., 52: 5213-5231.

Li, X., Aravkin, A.Y., van Leeuwen, T. and Herrmann, F.J., 2012. Fast randomized full-waveform inversion with compressive sensing. Geophysics, 77: A13-A17.

Li, X., Esser, E. and Herrmann, F.J., 2016. Modified Gauss-Newton full-waveform inversion explained - Why sparsity-promoting updates do matter. Geophysics, 81: R125-R138.

Lin, Y., Abubakar, A., and Habashy, T. M., 2012. Seismic full-waveform inversion using truncated wavelet representations. Expanded Abstr., 82nd Ann. Internat. SEG Mtg., Las Vegas: 1-6..

Martin, G. S., Wiley, R., and Marfurt, K. J., 2006. Marmousi2: An elastic upgrade for Marmousi. The Leading Edge, 25: 156-166.

Nocedal, J., and Wright, S., 2006. Numerical optimization. Springer Science & Business Media.

Pratt, R.G., Shin, C. and Hick, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. Geophys. J. Internat., 133: 341-362.

Son, W., Pyun, S., Jang, D., Park, Y. and Shin, C., 2012. A new algorithm adapting encoded simultaneous-source full waveform inversion to the marine-streamer data. Expanded Abstr., 82nd Ann. Internat. SEG Mtg., Las Vegas: 1-5.

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

Virieux, J. and Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics. Geophysics, 74: WCC1-WCC26.

Zhou, H. and Li, Q., 2013. Increasing waveform inversion efficiency of GPR data using compression during reconstruction. J. Appl. Geophys., 99: 109-113.

Zhu, L., Liu, E. and McClellan, J.H., 2015. Seismic data denoising through multiscale and sparsity-promoting dictionary learning. Geophysics, 80: WD45-WD57.

Zhu, L., Liu, E. and McClellan, J.H., 2017. Sparse-promoting full-waveform inversion based on online orthonormal dictionary learning. Geophysics, 82: R87-R107.