공학박사 학위논문

Multi-parametric Elastic Full Waveform Inversion Strategy for Seismic Data in the Frequency Domain

다변수 추출을 위한 탄성파 탐사 자료의 주파수 영역 탄성파 완전 파형 역산 전략

2014년 2월

서울대학교 대학원

에너지시스템공학부

오 주 원

Multi-parametric Elastic Full Waveform Inversion Strategy for Seismic Data in the Frequency Domain

다변수 추출을 위한 탄성파 탐사 자료의 주파수 영역 탄성파 완전 파형 역산 전략

지도 교수 민 동 주

이 논문을 공학박사 학위논문으로 제출함
2013 년 12 월

서울대학교 대학원
에너지시스템공학부
오 주 원

오주원의 박사 학위논문을 인준함
2014 년 1 월

위 원 장	신 창 수	(인)
부위원장	민 동 주	(인)
위 원	민 기 복	(인)
위 원	구 남 형	(인)
위 원	편 석 준	(인)

Abstract

Multi-parametric Elastic Full Waveform Inversion Strategy for Seismic Data in the Frequency Domain

Ju-Won Oh

Department of Energy Systems Engineering

The Graduate School

Seoul National University

As an alternative of the conventional hierarchical approach for the elastic full waveform inversion (FWI), I suggest new frequency-domain inversion strategy for multi-parametric elastic FWI of seismic data. The new inversion strategy consists of 4 inversion algorithms.

At first, the new parameterization using Poisson's ratio for a multi-parameter extraction is introduced. By analyzing the virtual source with mathematical and physical behaviors of its basis components, I verify that new parameterization using Poisson's ratio compensates some limitations of the conventional parameterizations, which are caused by absence of mode-converted and S-S waves in the partial derivative wavefields for the P-wave velocity.

Secondly, the spectral weighting scheme using the source-deconvolved backpropagated wavefields is developed to automatically control a spatial resolution of gradient direction depending on the thickness of subsurface layers.

Thirdly, for a noise reduction, I develop a spectral filtering scheme for the gradient direction using the denoise function, which is automatically calculated from seismic data and filters out relatively noise-contaminated frequency components during the inversion.

Finally, for a stable and accurate frequency-domain FWI, I suggest to apply the depth scaling scheme using the Levenberg-Marquardt method, in which the parameter-search moves from shallow to deep structures by reducing the damping factor.

By applying new inversion strategy to various numerical examples, I demonstrate that the new inversion strategy improves the stability and accuracy of the multi-parametric elastic FWI for random-noise included-data.

Keywords: Elastic full waveform inversion, Parameterization, Weighting method, Denoise function, Levenberg-Marquardt method
Student Number: 2010-23339

Contents

List of Tables

List of Figures

Chapter 1. Introduction

Seismic full waveform inversion (FWI) can be a useful method for inferring material properties such as velocities, which can be applied to background velocities during seismic migration. Although the rapid development of computer technology has made seismic waveform inversion more practical and a number of studies have been devoted to improving the accuracy and efficiency of waveform inversion (Tarantola 1984; Mora 1987; Pratt et al. 1998; Shin and Min 2006; Brossier et al. 2010; Herrmann 2010), some problems associated with seismic waveform inversion have yet to be resolved.

One of the main problems in multi-parametric full waveform inversion is that it is hard to recover subsurface parameters simultaneously because, as a number of parameters increases, the ill-posedness of the waveform inversion also increases. Several studies have devoted to increase the accuracy of multi-parametric FWI by finding optimal parameterization. For isotropic elastic case, Köhn et al. (2012) compared several isotropic elastic parameter groups and Jeong et al. (2012) suggested two-stage FWI strategy with a chain rule to recover accurate density. Recently, Prieux et al. (2013a) and Prieux et al. (2013b) tried the visco-acoustic and visco-elastic FWI using OBC (Ocean Bottom Cable) data by analyzing the radiation patterns of model parameters. For an elastic VTI case, which is more complex because 5 parameters should be recovered, Lee et al. (2010) tried recovering 4 parameters except the density by applying two-stage FWI with the coupling method to improve gradient for C_{11}. Koo et al. (2010) also suggested two-stage FWI using both isotropic and VTI elastic FWI. Kamath and Tsvankin (2013) tried to extract the VTI parameters from multi-components data including P-P waves and converted P-SV waves. However, there are still some problems to be resolved in VTI elastic FWI. For this reason, some groups have tried to recover VTI parameters from the acoustic VTI FWI (Gholami, et al., 2013a; Gholami et al., 2013b). However, most works have relied on phenomenological observations or some kinds of multi-stage FWI.

Another limitation of waveform inversion is that obtaining the global minimum solutions is not easy when the initial guesses deviate significantly from the true models. Several studies have proposed methods to overcome this problem. The frequency marching method (Sirgue and Pratt 2004; Kim et al. 2011) is one such method. In the frequency marching method, waveform inversion is performed by starting with low frequencies and moving to higher frequencies and is based on the fact that the objective function is not likely to have local minima at low frequencies (Bunks et al. 1995). The frequency marching method has yielded reasonable solutions for most cases to which it has been applied. However, in our experience, for salt models, which are one of the most challenging models in waveform inversion, the frequency marching method has not yielded good inversion results without good initial guesses. To provide good initial guesses for FWI, the Laplace-domain waveform inversion has been proposed (Shin and Cha 2008), which allows one to obtain reasonable solutions for salt models (Jeong et al. 2012). As an alternative to the frequency marching method, Lazaratos et al. (2011)

suggested the spectral shaping method, which makes the gradient have a similar spectrum to that of earth's subsurface using information recorded in local wells. However, most works also have relied on the multi-stage FWI or are expensive.

Some other problems must be resolved to make FWI practical. One of which is the effect of noise. Real field data include various coherent and incoherent noises. Because modeled data do not include noise, gradients that minimize residuals between noisy field data and noise-free modelled data can be distorted. In addition, noise can hide weak reflected signals with low signal-to-noise ratios. Several objective functions based on the l_1, l_2, Huber, and hybrid l_1/l_2 norms have been proposed as robust FWI algorithms for use with noise-contaminated data (Tarantola 1987; Huber 1973; Bube and Langan 1997; Ji, 2012). Pyun et al. (2009) suggested the l_1-like norm acoustic FWI algorithm, which shares features with the original l_1-norm inversion. Aravkin et al. (2011) proposed using Student's *t* distribution for a robust acoustic FWI. For elastic FWI, Brossier et al. (2010) showed that the l_1-norm objective function can be more convenient than the Huber and hybrid l_1/l_2 norms because additional effort is required to find suitable criteria for the Huber and hybrid l_1/l_2 norms. They reported reasonable results for data with uniform random noise using the l_1-norm FWI with a multistep strategy (Brossier et al. 2009a) and a dynamic time damping factor (Brossier et al. 2009b), under the assumption that the source wavelet spectrum is negligible when sequentially inverting single frequencies or groups of frequencies within a narrow bandwidth. However, the conventional l_1-norm elastic FWI does not guarantee stable solutions when the observed data are severely contaminated with incoherent random noise and when the inversion is performed simultaneously for the entire frequency band.

Xu et al. (2006) compared several source-independent waveform inversion algorithms for elastic FWI and asserted that IES (Iterative Estimation of Source signature) is better than the ATN (Average Trace Normalization) and STN (standard trace normalization) approaches for data containing random noise. Choi and Min (2012) compared the source-estimation logarithmic waveform inversion algorithm with the source-independent logarithmic waveform inversion. They found that the source-estimation logarithmic inversion can reduce the influence of random noise on the inversion result, whereas the source-independent logarithmic method yields slightly better results than the source-estimation logarithmic method for coherent noise. However, the inversion results from the elastic FWI for data with random noise are still unsatisfactory.

For another issues, that is to make frequency-domain FWI more stable and accurate, Wang and Rao (2009) suggested the layer-stripping method, in that the model parameters are sequentially inverted from the shallow to deep parts of given models. However, in the layer-stripping method, the parameter-searching region should be manually controlled with additional processes. On the other hand, Brossier et al. (2009b) suggested the dynamic time-damping strategy, in which the observed data are damped depending on arrival time and the parameter-searching is also performed from shallow to deep parts. However, in their method, the time-damping factors and the number of iteration for each time-damping factor are manually determined.

Because, as mentioned above, most previous works have relied on some hierarchical FWIs with too many stages to compensate various limitations of the frequency-domain multi-parametric FWI, the uncertainties of FWI increase due to human interventions at a finish of each stage. To overcome this limitation of conventional hierarchical approaches, throughout the paper, I introduce the new inversion algorithms, which can partially or totally replace the conventional multi-stage FWI.

This paper is organized as follows: In Chapter 2, the inverse problem for the single-frequency data and banded data will be discussed. In Chapter 3, the scattering patterns induced by model parameters are analyzed using a new concept (basis virtual source) and, based on the analysis, the new parameterization will be introduced, which provides robust and highly resolved inversion results for the isotropic elastic FWI. In Chapter 4, I analyze the limitations of conventional scaling methods and introduce the weighting method, in which the weighting factor is automatically calculated and reflects the information about the thickness of subsurface layer from the seismic data. In Chapter 5, I analyze the noise behavior in the frequency-domain FWI and introduce the denoise function, which acts like a conventional manual frequency filter but is determined automatically from the seismic data. In Chapter 6, for more stable and accurate frequency-domain FWI, the depth scaling method using the Levenberg-Marquardt method (Levenberg, 1944; Marquardt 1963) is introduced, which can be an alternative of the layer-stripping method (Wang and Rao, 2009) and time-damping strategy (Brossier et al. 2009b). Finally, in Chapter 7, the numerical examples for the elastic Marmousi-2 model using new inversion algorithm will be showed to check if the new inversion algorithm improves the accuracy and robustness of multi-parametric FWI without relying the multi-stage approaches.

Chapter 2. Inverse Problem

2.1 Inverse Problem for Single-Frequency Data

In frequency-domain waveform inversion using the Levenberg-Marquardt method (Levenberg, 1944; Marquardt, 1963; Lines and Treitel, 1984), the objective function, which measures the misfit between the single-frequency modeled and observed data based on the assumption that the error is a linear function with respect to the model parameter change, can be expressed by

$$M(\delta\mathbf{p},\beta)=\sum_{s}\left[\left(\mathbf{J}_s\delta\mathbf{p}-\mathbf{r}_s\right)^T\left(\mathbf{J}_s\delta\mathbf{p}-\mathbf{r}_s\right)^*+\beta\left(\delta\mathbf{p}^T\delta\mathbf{p}^*-p_0^2\right)\right], \quad (2\text{-}1)$$

where $\delta\mathbf{p}$ and β are the model parameter change vector and the damping factor (i.e., Lagrange multiplier), respectively and the term s denote the source. The superscripts T and * indicate the transpose and complex conjugate, respectively. According the Lines and Treitel (1984), when using the Levenberg-Marquardt method, we can apply a constraint that makes the energy of the parameter change vector, $\delta\mathbf{p}$, converge to a certain finite quantity, p_0. The term β damps out the changes in the model parameter vector by limiting the energy in the gradient vector. This approach is very useful for the frequency-domain FWI and I will discuss about several advantages of the Levenberg-Marquardt method in Chapter 6.

The terms $\mathbf{J}_s$ and $\mathbf{r}_s$ are the Jacobian matrix and the residual vector, respectively. The residual vector is defined as the set of differences between the observed data ($\mathbf{d}_s$) and the modeled data ($\mathbf{u}_s(\mathbf{p})$) obtained for the initial or assumed model, $\mathbf{p}$, as expressed by

$$\mathbf{r}_s = \mathbf{d}_s - \mathbf{u}_s(\mathbf{p}), \tag{2-2}$$

where

$$\mathbf{u}_s(\mathbf{p}) = \mathbf{S}^{-1}(\mathbf{p})\mathbf{f}_s \tag{2-3}$$

and $\mathbf{S}$ is the modeling operator. The Jacobian matrix, which is called partial derivative wavefields, is expressed by

$$\mathbf{J}_s = \begin{pmatrix} \frac{\partial \mathbf{u}_s(\mathbf{p})}{\partial p_1} & \cdots & \frac{\partial \mathbf{u}_s(\mathbf{p})}{\partial p_k} & \cdots & \frac{\partial \mathbf{u}_s(\mathbf{p})}{\partial p_{nm}} \end{pmatrix}, \tag{2-4}$$

where

$$\frac{\partial \mathbf{u}_s(\mathbf{p})}{\partial p_k} = \mathbf{S}^{-1}\left(-\frac{\partial \mathbf{S}}{\partial p_k}\mathbf{u}_s(\mathbf{p}) \right) = \mathbf{S}^{-1}\mathbf{f}_{s,k}^{v}(\mathbf{p}). \tag{2-5}$$

The terms $\mathbf{f}_{s,k}^{v}(\mathbf{p})$ and *nm* denote the virtual source vector for the k$^{\text{th}}$ model parameter and the number of model parameters, respectively.

The goal of the inverse problem is to find an optimal model parameter change vector ($\delta\mathbf{p}$) that minimizes the objective function expressed in eq. (2-1). To achieve this goal, we can differentiate the objective function with respect to the model parameter change vector. Consequently, the optimal parameter change vector, which is called the damped linear least-squares solution, can be obtained by solving the following normal equation:

$$\left[\sum_s \left(\mathbf{J}_s^T \mathbf{J}_s^* + \beta \mathbf{I} \right) \right] \delta\mathbf{p} = \sum_s \left(\mathbf{J}_s^T \mathbf{r}_s^* \right) \tag{2-6}$$

The left term without the damping factor ($\sum_s \left(\mathbf{J}_s^T \mathbf{J}_s^* \right)$) represents the approximate Hessian matrix obtained based on the assumption that the errors are linear; the right term denotes the gradient direction.

2.2 Inverse Problem for Banded Data

For the banded seismic data, the objective function described for the single-frequency data in eq. (2-1) can be rewritten as follows:

$$M(\delta\mathbf{p},\beta) = \sum_{\omega}\sum_{s}\left[\left(\mathbf{J}_{s,\omega}\delta\mathbf{p}-\mathbf{r}_s(\omega)\right)^T\left(\mathbf{J}_{s,\omega}\delta\mathbf{p}-\mathbf{r}_s(\omega)\right)^* + \beta\left(\delta\mathbf{p}^T\delta\mathbf{p}^* - p_0^2\right)\right], \quad (2\text{-}7)$$

where ω denote the angular frequency. As I did for the single-frequency FWI, the optimal solution, which minimizes the above objection function, can be obtained by solving the following normal equation:

$$\left[\sum_{\omega}\sum_{s}\left(\mathbf{J}_{s,\omega}^T\mathbf{J}_{s,\omega}^* + \beta\mathbf{I}\right)\right]\delta\mathbf{p} = \sum_{\omega}\sum_{s}\left(\mathbf{J}_{s,\omega}^T\mathbf{r}_s(\omega)^*\right) \quad (2\text{-}8)$$

This approach is accurate and satisfies the inverse theory for the banded data. I call this approach 'Conventional Scaling method-I (CS-I)' throughout the paper and will discuss the characteristics of the CS-I method in Chapter 4. However, the FWI using CS-I method generally provides poorly resolved inversion results, because, in seismic exploration using an active source, the frequency-domain inversion problem is generally affected by the source spectrum. In other words, the source spectrum acts as a weighting function during the inversion. This source-dependent property of frequency-domain seismic waveform inversion may degrade inversion results as Jang et al. (2009) showed. To avoid this limitation, Brossier et al. (2010) used a hierarchical approach, in which the inverse problem is solved for several frequencies moving from low-frequency to high-frequency bands. They assumed that, if the inversion is performed with mono-frequency or very narrow-band frequency data, the weighting effect of the source spectrum can be negligible. However, this approach may require more computational time than the simultaneous-frequency inversion because the inverse problem for each frequency group should be solved sequentially without the parallel computation over frequencies. On the other hand, Jang et al. (2009) suggested a scaling method for the simultaneous-frequency inverse problem of banded data by applying the Hessian matrix inside the frequency loop. They showed that this scaling method improves inversion results by minimizing the effect of source spectrum. This scaling method requires less computational time and human intervention than the hierarchical approach. However, for the banded data, this scaling method does not conform to the aforementioned inverse theory because. According to the inverse theory, the Hessian matrix should be applied outside the frequency loop as shown in eq. (2-8). For this reason, I will call this approach the 'combination of mono-frequency inverse problems' throughout the paper. The combination of mono-frequency inverse problems does not directly minimize the objective function in eq. (2-7). Instead, we first minimize the mono-frequency objective function, in eq. (2-1), to obtain mono-frequency gradient direction at each frequency and then combine those mono-frequency gradients to construct a total gradient direction, which is called 'banded gradient direction'. The combination of mono-frequency inverse problems can begin with the objective function expresses by

$$M = \sum_{\omega} \sum_{s} \left(\mathbf{r}_{s,\omega} - \mathbf{J}_{s,\omega} k_{\omega} \delta \mathbf{p}_{\omega} \right)^{T} \left(\mathbf{r}_{s,\omega} - \mathbf{J}_{s,\omega} k_{\omega} \delta \mathbf{p}_{\omega} \right)^{*}, \qquad (2\text{-}9)$$

where k_ω is a scaling factor, which determines the weight of each mono-frequency descent direction. As we did for the aforementioned inverse problem, the model parameter update at each frequency can be expressed using the mono-frequency descent direction as

$$\delta \mathbf{p}_\omega = \frac{1}{k_\omega} \left[\sum_s \mathbf{J}_{s,\omega}^T \mathbf{J}_{s,\omega}^* \right]^{-1} \sum_s \mathbf{J}_{s,\omega}^T (\mathbf{d}_{s,\omega} - \mathbf{u}_{s,\omega})^* . \quad (2\text{-}10)$$

We can combine the mono-frequency descent directions to obtain the banded descent direction in several ways. As performed by Jang et al. (2009), we can simply sum all the mono-frequency descent directions without normalization. In this case, the scaling factor, k_ω, is 1. In another way, as Ha et al. (2009) did, we can sum all the mono-frequency descent directions after the normalization, in which each mono-frequency descent vector is divided by its maximum absolute value. In this approach, the scaling factor is the maximum absolute value of each mono-frequency descent vector. However, the former provides inconsistent inversion results because of the spectral weighting effect of the Hessian matrix (for more details, refer Appendix A in Oh and Min [2013a]). Although the latter, which will be called 'Conventional Scaling method II (CS-II)' throughout the paper', provides more stable solutions, neither method guarantees the global minimum solution for the objective function in eq. (2-7) because their solutions are just the combination of solutions for the mono-frequency inverse problems. The influences and disadvantages of the conventional scaling methods will be discussed in Chapter 4.

One limitation of the FWI is the computational overburden related to the computation of the huge Jacobian matrix and the inverse of the Hessian matrix. Therefore, for a computational convenience, the back-propagation of the residuals is adopted (Pratt et al. 1998) to avoid computing Jacobian directly. With the same reason as that for the gradient, the diagonal of the new pseudo-Hessian matrix is used (Choi et al. 2008) instead of the approximate Hessian matrix although there are several ways to calculate the Hessian matrix approximately, such as the l-BFGS (Byrd et al. 1995), quasi-Newton (Nocedal, 1980) and the truncated Newton strategy (Métivier et al. 2012). Therefore, the approximated gradient direction (we call it 'descent direction' throughout the paper) for the CS-I and CS-II methods can be expressed by

$$\delta \mathbf{p}_{\text{CS-I}}^{(l)} = \left[\left\{ \sum_{\omega} \text{diag}\{\mathbf{H}_n(\omega)\} \right\} + \beta \mathbf{I} \right]^{-1} \left[\sum_{\omega} \nabla_{\mathbf{p}} E(\omega) \right] \qquad (2\text{-}11)$$

and

$$\delta \mathbf{p}_{\text{CS-II}}^{(l)} = \text{NRM}^2 \sum_{\omega} \text{NRM}^1 \left[\left(\text{diag}\{\mathbf{H}_n(\omega)\} + \beta(\omega)\mathbf{I} \right)^{-1} \nabla_{\mathbf{p}} E(\omega) \right], \qquad (2\text{-}12)$$

where

$$\mathbf{H}_n(\omega) = \sum_{s} \left[\mathbf{F}_s^{v}(\omega) \right]^T \mathbf{A} \left[\mathbf{F}_s^{v}(\omega) \right]^*, \qquad (2\text{-}13)$$

$$\text{diag}(\mathbf{A}) = \text{Re}\left\{ \sum_{i=1}^{ns} |g_{i,1}| \quad \sum_{i=1}^{ns} |g_{i,2}| \quad \cdots \quad \sum_{i=1}^{ns} |g_{i,np}| \right\}, \qquad (2\text{-}14)$$

and

$$\nabla_{\mathbf{p}} E(\omega) = \sum_{s} \text{Re}\left\{ \left[\mathbf{F}_s^{v}(\omega, \mathbf{p}) \right]^T \left[\mathbf{S}^{-1}(\omega, \mathbf{p}) \right]^T \left[\mathbf{d}_s(\omega) - \mathbf{u}_s(\omega, \mathbf{p}) \right]^* \right\}. \qquad (2\text{-}15)$$

The terms, $\mathbf{H}_n(\omega)$, g_i and $\nabla_{\mathbf{p}} E(\omega)$, are the new pseudo-Hessian matrix, the impulse response for the ith source and the mono-frequency gradient, respectively. The terms *ns* and *np* denote the number of sources and the number of nodal points, respectively. The term NRM^1 represents the normalizing operator, which divides the each single-frequency descent direction vector by its maximum absolute value. This normalizing operator makes the single-frequency descent directions contribute equally to the parameter update, $\delta \mathbf{p}^{(l)}$. The term NRM^2 denotes the normalizing operator for the total descent direction, which constrains the total descent direction to the range from -1 to 1. With the solution, $\delta \mathbf{p}^{(l)}$, the model parameter can be updated with step length (α) as follows:

$$\mathbf{p}^{(l+1)} = \mathbf{p}^{(l)} + \alpha \times \delta \mathbf{p}^{(l)} \quad . \qquad (2\text{-}16)$$

The modified version of the conjugate gradient method is applied to accelerate the convergence rate (Fletcher and Reeves 1964; Ha et al. 2009) and the finite-element method is used for forward modeling (Zienkiewicz and Taylor 2000). For a fair comparison, I assumed that the exact source wavelet is known because, in my experience, the results of the source wavelet estimation (Song et al. 1995) are different depending on the parameterization.

Chapter 3. Parameterization for Elastic FWI

3.1 Analysis on Radiation Patterns of Virtual Source

In Chapter 3, to find the best parameterization for multi-parametric FWI, the characteristics of virtual source will be discussed in its physical aspects. Because the virtual source is only term changed by the parameterization, the analysis on the behavior of the virtual source is required. Recently, several studies have tried to find the best parameterization in some cases analyzing the radiation patterns depending on the incidence and scattered angles, such as acoustic (Prieux et al., 2013a), acoustic VTI (Gholami et al., 2013a, Gholami et al., 2013b and Plessix and Cao, 2011), isotropic elastic (Köhn et al. 2012), acoustic-elastic coupled media (Prieux et al., 2013a and Prieux et al., 2013b). However, most of the previous works are too phenomenological to explain the details about behaviors of the virtual source. For these reasons, in this chapter, I discuss some generalized interpretation tools to understand the behavior of the virtual source depending on properties of medium and parameterizations by introducing new concept, which is called 'basis virtual source'. Before introducing basis virtual sources, I first discuss how the radiation patterns of the virtual source are determined in the FWI analyzing the scattering pattern of subsurface heterogeneities.

3.1.1 Previous works for the scattering of seismic wave

Before explaining the behaviour of virtual source as a point scatterer, previous studies about the scattering of seismic waves will be reviewed because the radiation pattern of the virtual source resembles the seismic scattering in a media with heterogeneous inclusion. The seismic waves are generally scattered by numerous subsurface heterogeneities depending on the wavelength and the length of the heterogeneity, which make the recorded seismic signals more complicated.

The mechanisms of scattering have been well resolved in various scientific fields like meteorology and optics (Cox, 2002; Bohren and Huffman, 1983) as well as in the seismology. According to the Reynolds (1997), the scattering patterns of seismic waves can be divided by four cases depending on the angular wavenumber (k) and the length of the heterogeneity (a). When the wavelength is too long compared to the length of a heterogeneity ($ka < 0.01$), the scattering effects are negligibly small and the medium can be regarded as quasi-homogeneous. When the wavelength becomes shorter or the length of a heterogeneity increases ($0.01 \le ka < 0.1$), the scattering is called 'Rayleigh scattering', in which the scattered wavefields propagate all the direction. For shorter wavelength or longer heterogeneity ($0.1 \le ka < 10$), the resonance scattering (so-called Mie scattering) can be occurred in which the scattered waves are dominant at the forward or backward directions. When the wavelength is too short compared to the length of a heterogeneity ($10 \le ka$), the scattered waves follow the geometric ray theory. The variations of scattered wavefields for the various values of *ka* are well described in Wu (1984).

The study about the scattering of seismic waves has been actively studied from 1980s in the pure seismology because recorded earthquake waves, which have very low frequency signals, tend to be easily scattered by numerous subsurface heterogeneities. If the seismic waves are affected by various kinds of scattering, they are damped due to the energy loss and contain lots of scattered artefacts. For these reasons, many geophysicists have devoted understanding the scattering patterns of seismic waves during few decades (Aki and Richards, 1980; Wu and Aki, 1985).

The scattered energy, which are generated by subsurface inclusions and recorded on the surface, are described by the Born approximation in the frequency-domain as follows (Snieder, 2002):

$$\begin{aligned} u_i^B(r) = \omega^2 \int G_{ij}^{(0)}(r,r')\Delta\rho(r')u_j(r')dV' \\ + \int G_{ij}^{(0)}(r,r')\partial_k^{'}\left(\Delta c_{nklj}(r')\partial_l^{'}u_j(r')\right)dV' . \\ - \int G_{ij}^{(0)}(r,r')n_j\Delta c_{nklj}(r')\partial_k^{'}u_l(r')dS' \end{aligned} \quad (3\text{-}1)$$

The terms, $u_i^B(r)$ and $G_{ij}^{(0)}(r,r')$, indicate the scattered wavefields and the Green's function from an arbitrary location, r', to a receiver, r, on a reference medium, respectively. The terms, $u_j(r')$ and $\partial_k^{'}$, denote the incidence wavefields at an arbitrary location (r') and a spatial differentiation, respectively. The first and second volume integral represents the recorded scattered displacement fields at a receiver (r) induced by a perturbation of densities and elastic coefficients at an arbitrary location, r', respectively. The surface integral at the third line in eq. (3-1) denotes the surface traction induced by a perturbation of elastic coefficients on the surface (S). Using the Born approximation, the scattered wavefields at a receiver induced by subsurface heterogeneities can be analytically calculated. For more details about how the Born approximation is derived, refer Snieder (2002).

Many previous works have devoted describing the scattering patterns of seismic waves using the moment-tensor description of the virtual source (Wu and Aki, 1985; Burridge et al., 1998). However, as a generalized interpretation tool for the virtual source, it is not enough because previous works describe the moment tensor of the virtual source only depending on subsurface parameter. In other words, conventional studies for the scattering patterns are specified for not the FWI but natural scattering phenomena. Because, in the FWI, the scattering patterns are mainly governed by the parameterization, the conventional analysis for the scattering should be modified to describe the numerical scattering phenomena induced by various kinds of the virtual source.

3.1.2 Virtual source in FWI as a point scatterer

As mentioned in the previous section, the virtual source is only term changed depending on the parameterization. Therefore, to understand the influence of different parameterization, we have to know the behavior of the virtual source during FWI. Because the virtual source can be regarded as a point scatterer, many geophysicists have tried to interpret its behavior based on the Born approximations and have provided reasonable interpretations, by plotting the radiation patterns of the virtual source for various incidence and scattered angles, for several cases (Gholami et al., 2013a; Gholami et al., 2013a; Prieux et al., 2013a; Prieux et al., 2013b). However, the scattering patterns of virtual source are too complicated, particularly in the elastic media. In addition, as the Born approximation (eq. (3-1)) shows, the scattering pattern are frequency-dependent and wavelength-dependent.

Other minor influences of the parameterization also disturb interpreting exact behavior of the virtual source. For examples, the degree of approximation for the Hessian matrix influences on the FWI because the Hessian matrix also consists of the virtual sources. Additionally, the different contribution of a step-length for different parameterizations also causes some differences. For these reasons, somebody can reach different results for the same parameterization depending on which inversion strategy they used. To reduce above minor influences, in this paper, the inversion approach is based on the steepest descent method and only the gradient direction obtained at the first iteration will be analyzed.

For more general interpretation tool for scattering patterns of the virtual source, we start to analyze the physical meaning of the virtual sources by comparing the elastic wave equation to the Born approximation. The displacement-based elastic wave equation in the time domain, particularly for the elastic VTI (Vertical Transverse Isotropic) media, can be expressed by

$$\begin{aligned}\rho\frac{\partial^2 u_x}{\partial t^2}&=\frac{\partial}{\partial x}\left[c_{11}\frac{\partial u_x}{\partial x}+c_{13}\frac{\partial u_z}{\partial z}\right]+\frac{\partial}{\partial z}\left[c_{44}\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right]\\ \rho\frac{\partial^2 u_z}{\partial t^2}&=\frac{\partial}{\partial z}\left[c_{13}\frac{\partial u_x}{\partial x}+c_{33}\frac{\partial u_z}{\partial z}\right]+\frac{\partial}{\partial x}\left[c_{44}\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right]\end{aligned}. \quad (3\text{-}2)$$

As discussed in the inverse problem, the virtual source (for comparison to basis virtual source that introduced next section, it will be called 'full virtual source') can be obtained by taking partial derivative of modelling operator with respect to the model parameter as expressed by eq. (2-5).

To make the full virtual source has a form like Born approximation, we try to express the elastic virtual source in the frequency domain, which follows simple ray path shown in Figure 3-1, for an arbitrary parameter (p) using wave propagators as shown in

$$f^{v}_{\text{elastic}}\approx\omega^2\frac{\partial m(i,j)}{\partial p}u_d(i,j)+w^{\mp}_{\tau-f}\left(\frac{\partial k_{e-\tau}(i,j)}{\partial p}\right)w^{+}_{d-e}u_d(i,j) \quad , \qquad (3\text{-}3)$$

where

$$w^{+}=\exp(-ik_x\ \ x-ik_z\Delta z) \qquad (3\text{-}4)$$

and

$$w^{\mp}=\exp(\pm ik_x\ \ x\pm ik_z\Delta z)\,. \qquad (3\text{-}5)$$

The operator, w^{+}, indicates the down- and right-going wave propagator and $w^{\mp}$ indicates the bidirectional wave propagator (Berkhout, 1985). The terms k_x and k_z are angular wavenumbers and the terms, x and z, are the propagation distance for *x* and *z* directions, which equivalent to the length of a heterogeneity, respectively. The terms, *m* and *k*, denote the mass and stiffness coefficients. The term, $-\omega^2$, represents the second order time derivative. The subscripts, *d*, *e*, *τ* and *f* denote the displacement, strain, stress and body force, respectively, and '-' in the subscript means that the physical quantities are converted by the operator from left to right.

Because the virtual source is assumed as a point scatterer located in the subsurface medium as shown in Figure 3-1, the surface integral in eq. (3-1) can be vanished and the volume integrals for whole domain can be ignored. In addition, because the scattered wavefields ($u_i^B(r)$) and Green's function ($G_{ij}^{(0)}$) in eq. (3-1) correspond to the Jacobian and inverse of modelling operators for the FWI in eq. (2-5), respectively, we notice that the full virtual source has the same form with that in the Born approximation. In addition, because wave propagators originate from the elastic wave equation in eq. (3-2), they represent the spatial derivatives with respect to *x* and *z*. In other words, the terms, w^{+} and $w^{\mp}$, can be regarded as the inner spatial derivatives and the outer derivatives in eq. (3-2), respectively. Using this expression, we can also regard the spatial derivatives of the displacement (so-called 'strain') determines the pattern of the incidence waves, on the other hand, the spatial derivatives of the stress determines the scattering patterns acting as body forces (For more details, refer section named 'Basis virtual source'). For these reasons, understanding the behavior of spatial derivatives in the virtual source is important to interpret the radiation patterns of the virtual source.

As well as the spatial derivative, the time derivative that corresponds to $-\omega^2$ in the frequency domain also need to be treated importantly. This is because, in the high frequency, the contribution of the time derivative becomes dominant in the total scattered wavefields and makes the scattering pattern more complicated. For the virtual source in the FWI, the length of heterogeneity corresponds to the size of grid because the virtual source acts as a point scatterer. When we perform the frequency-domain FWI, the grid size (Δh) is generally determined by the reasonably smallest grid numbers for one wavelength (G) in the range of avoiding numerical dispersions retaining the computational efficiency ($5 < G < 10$) as expressed by

$$\Delta h \leq \frac{v_{\min}}{G \times f_{\max}}, \tag{3-6}$$

where $f_{\max}$ and $v_{\min}$ indicate the maximum frequency and minimum velocity in the subsurface media. In this case, the value, ka, can be expressed as follows:

$$ka \leq \frac{2\pi f}{v} \times \frac{v_{\min}}{G \times f_{\max}}. \tag{3-7}$$

From above equation, we notice that the FWI is generally in the Mie scattering regime in which the scattering pattern is very complicated depending on the value, ka, except when the inversion-frequency (f) is enough low or the medium-velocity (v) is enough high.

This study is designed based on the fact that all the wave equations share similar time and spatial derivatives. For this reason, if the characteristics for the radiation pattern of the virtual source are classified by not the parameter but the time and spatial derivatives, it would provide more generalized insights for the behavior of the virtual source in most cases. For this reason, I developed the basis virtual sources, which are the mathematical and physical bases of the virtual source and are characterized depending on time and spatial derivatives.

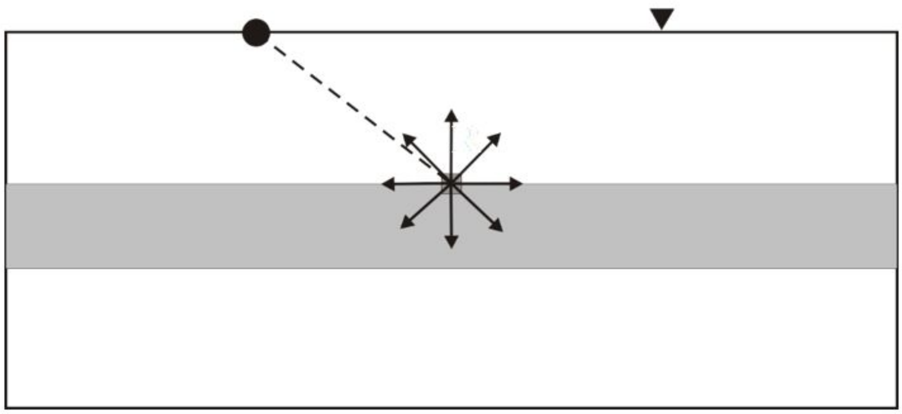


Figure 3-1 Schematic diagram illustrating the wave propagation scattered at a boundary of a subsurface layer. Black circle, inverted triangle and grey square indicate the source, receiver and subsurface reflector, respectively. The dashed and solid lines denote the incidence and scattered ray paths.

3.2 Basis Virtual Source

3.2.1 Mathematical Description of Basis Virtual Source

For more generalized interpretation tool for the behavior of the virtual source, the new concept, '*basis virtual source*', is introduced, which are classified depending on the time and spatial derivatives in the elastic wave equation. Because the staggered-grid FDM is specified to express the both stress and displacement components physically well, the staggered-grid (as shown in Figure 3-2a) is employed, in which all the physical parameters are located on the same grid as that of the normal stresses, to define the mathematical expression of the basis virtual sources (Graves, 1996). To figure out the mathematical shape of the basis virtual source at a glance, it would be helpful to analyze the displacement-based elastic wave equation as shown in eq. (3-2).

Based on the displacement-based elastic wave equation, we can divide the modeling operator by 8 kinds of spatial derivatives and 2 kinds of time derivatives. This classification is also valid in the Born approximation (eq. (3-1)) because it also consists of 8 kinds of scattering source in the second volume integral, depending on the directions (2 kinds) and components (2 kinds) of the spatial derivatives and recorded wavefields (2 kinds), and 2 kinds of scattering source in the first volume integral depending on the recorded wavefields. To distinguish 10 basis virtual sources, we can divide them into two groups depending on their direction: one group is named '*Horizontal Basis virtual source* (HB)' that acts along the horizontal direction and the other group is called '*Vertical Basis virtual source* (VB)', which acts along the vertical direction. Finally, the 10 basis virtual sources, which act on grids around the model parameter, can be briefly expressed by

$$\begin{aligned}
&\mathrm{HB}_{xxh} = \frac{\partial^2 u_x}{\partial x \partial x} && \mathrm{VB}_{zxh} = \frac{\partial^2 u_x}{\partial z \partial x} \\
&\mathrm{HB}_{xzv} = \frac{\partial^2 u_z}{\partial x \partial z} && \mathrm{VB}_{zzv} = \frac{\partial^2 u_z}{\partial z \partial z} \\
&\mathrm{HB}_{zzh} = \frac{\partial^2 u_x}{\partial z \partial z} && \mathrm{VB}_{xzh} = \frac{\partial^2 u_x}{\partial x \partial z} \\
&\mathrm{HB}_{zxv} = \frac{\partial^2 u_z}{\partial z \partial x} && \mathrm{VB}_{xxv} = \frac{\partial^2 u_z}{\partial x \partial x} \\
&\mathrm{HB}_{tth} = \frac{\partial^2 u_x}{\partial^2 t} && \mathrm{VB}_{ttv} = \frac{\partial^2 u_z}{\partial^2 t}
\end{aligned} \quad , \tag{3-8}$$

where first two subscripts, x, z and t, denote the two spatial and time derivatives, respectively, and the last subscripts, h and v, indicate that they are derived from horizontal and vertical displacements of incidence waves. Because, the virtual source is obtained by taking partial derivative with respect to model parameter, the parameter between the each primary derivative can be treated as a scalar. At an arbitrary point, (i, j), the vector for basis virtual sources can be expressed as follows:

$$\mathrm{HB}_{xxh}(i,j) = \begin{pmatrix} 0 & 0 & 0 & -M^4_{\mathrm{HB}_{xxh}} & M^5_{\mathrm{HB}_{xxh}} & 0 & 0 & 0 & 0 \end{pmatrix}^T \tag{3-9}$$

$$\mathrm{HB}_{xzv}(i,j) = \begin{pmatrix} 0 & 0 & 0 & -M^4_{\mathrm{HB}_{xzv}} & M^5_{\mathrm{HB}_{xzv}} & 0 & 0 & 0 & 0 \end{pmatrix}^T \tag{3-10}$$

$$\begin{aligned}
\mathrm{HB}_{zzh}(i,j) = &\begin{pmatrix} -M^1_{\mathrm{HB}_{zzh}} & -M^2_{\mathrm{HB}_{zzh}} & 0 & 0 & 0 & 0 & M^7_{\mathrm{HB}_{zzh}} & M^8_{\mathrm{HB}_{zzh}} & 0 \end{pmatrix}^T \\
&+ \begin{pmatrix} 0 & 0 & 0 & -M^4_{\mathrm{HB}_{zzh}} & M^5_{\mathrm{HB}_{zzh}} & 0 & 0 & 0 & 0 \end{pmatrix}^T
\end{aligned} \tag{3-11}$$

$$\begin{aligned}\mathrm{HB}_{zxv}(i,j) &= \begin{pmatrix} -M^{1}_{\mathrm{HB}_{zxv}} & -M^{2}_{\mathrm{HB}_{zxv}} & 0 & 0 & 0 & 0 & M^{7}_{\mathrm{HB}_{zxv}} & M^{8}_{\mathrm{HB}_{zxv}} & 0 \end{pmatrix}^{T} \\ &+ \begin{pmatrix} 0 & 0 & 0 & -M^{4}_{\mathrm{HB}_{zxv}} & M^{5}_{\mathrm{HB}_{zxv}} & 0 & 0 & 0 & 0 \end{pmatrix}^{T}\end{aligned} \tag{3-12}$$

$$\mathrm{HB}_{tth}(i,j) = \begin{pmatrix} 0 & 0 & 0 & -M^{4}_{\mathrm{HB}_{tth}} & -M^{5}_{\mathrm{HB}_{tth}} & 0 & 0 & 0 & 0 \end{pmatrix}^{T} \tag{3-13}$$

$$\mathrm{VB}_{zxh}(i,j) = \begin{pmatrix} 0 & -M^{2}_{\mathrm{VB}_{zxh}} & 0 & 0 & M^{5}_{\mathrm{VB}_{zxh}} & 0 & 0 & 0 & 0 \end{pmatrix}^{T} \tag{3-14}$$

$$\mathrm{VB}_{zzv}(i,j) = \begin{pmatrix} 0 & -M^{2}_{\mathrm{VB}_{zzv}} & 0 & 0 & M^{5}_{\mathrm{VB}_{zzv}} & 0 & 0 & 0 & 0 \end{pmatrix}^{T} \tag{3-15}$$

$$\begin{aligned}\mathrm{VB}_{xzh}(i,j) &= \begin{pmatrix} -M^{1}_{\mathrm{VB}_{xzh}} & 0 & M^{3}_{\mathrm{VB}_{xzh}} & 0 & 0 & 0 & -M^{7}_{\mathrm{VB}_{xzh}} & 0 & M^{9}_{\mathrm{VB}_{xzh}} \end{pmatrix}^{T} \\ &+ \begin{pmatrix} 0 & -M^{2}_{\mathrm{VB}_{xzh}} & 0 & 0 & M^{5}_{\mathrm{VB}_{xzh}} & 0 & 0 & 0 & 0 \end{pmatrix}^{T}\end{aligned} \tag{3-16}$$

$$\begin{aligned}\mathrm{VB}_{xxv}(i,j) &= \begin{pmatrix} -M^{1}_{\mathrm{VB}_{xxv}} & 0 & M^{3}_{\mathrm{VB}_{xxv}} & 0 & 0 & 0 & -M^{7}_{\mathrm{VB}_{xxv}} & 0 & M^{9}_{\mathrm{VB}_{xxv}} \end{pmatrix}^{T} \\ &+ \begin{pmatrix} 0 & -M^{2}_{\mathrm{VB}_{xxv}} & 0 & 0 & M^{5}_{\mathrm{VB}_{xxv}} & 0 & 0 & 0 & 0 \end{pmatrix}^{T}\end{aligned} \tag{3-17}$$

$$\mathrm{VB}_{ttv}(i,j) = \begin{pmatrix} 0 & -M^{2}_{\mathrm{VB}_{ttv}} & 0 & 0 & -M^{5}_{\mathrm{VB}_{ttv}} & 0 & 0 & 0 & 0 \end{pmatrix}^{T} \tag{3-18}$$

The terms, $M^{4}_{\mathrm{HB}_{zzh}}$, $M^{5}_{\mathrm{HB}_{zzh}}$, $M^{4}_{\mathrm{HB}_{zxv}}$, $M^{5}_{\mathrm{HB}_{zxv}}$, $M^{2}_{\mathrm{VB}_{xzh}}$, $M^{5}_{\mathrm{VB}_{xzh}}$, $M^{2}_{\mathrm{VB}_{xxv}}$ and $M^{5}_{\mathrm{VB}_{xxv}}$, are ignored because they are just the products of the second derivatives in staggered-grid FDM and are appeared only in the staggered-grid FDM. In each equation, the terms, M (the superscripts mean the nodal point as shown in Figure 3-2b), can be treated as momenta of basis virtual sources (physically they mean components of normal and shear strains) and can be expressed as follows:

$$M^{4}_{\mathrm{HB}_{xxh}} = -\frac{\partial u_x(i,j,t)}{\partial x} \qquad M^{5}_{\mathrm{HB}_{xxh}} = -\frac{\partial u_x(i,j,t)}{\partial x} \tag{3-19}$$

$$M_{\mathrm{HB}_{xzv}}^{4} = -\frac{\partial u_z(i,j,t)}{\partial z} \qquad M_{\mathrm{HB}_{xzv}}^{5} = -\frac{\partial u_z(i,j,t)}{\partial z} \tag{3-20}$$

$$\begin{aligned} M_{\mathrm{HB}_{zzh}}^{1} &= -\frac{\partial u_x(i-1,j,t)}{\partial z} & M_{\mathrm{HB}_{zzh}}^{2} &= -\frac{\partial u_x(i,j,t)}{\partial z} \\ M_{\mathrm{HB}_{zzh}}^{7} &= -\frac{\partial u_x(i-1,j,t)}{\partial z} & M_{\mathrm{HB}_{zzh}}^{8} &= -\frac{\partial u_x(i,j,t)}{\partial z} \end{aligned} \tag{3-21}$$

$$\begin{aligned} M_{\mathrm{HB}_{zxv}}^{1} &= -\frac{\partial u_z(i,j-1,t)}{\partial x} & M_{\mathrm{HB}_{zxv}}^{2} &= -\frac{\partial u_z(i,j-1,t)}{\partial x} \\ M_{\mathrm{HB}_{zxv}}^{7} &= -\frac{\partial u_z(i,j,t)}{\partial x} & M_{\mathrm{HB}_{zxv}}^{8} &= -\frac{\partial u_z(i,j,t)}{\partial x} \end{aligned} \tag{3-22}$$

$$M_{\mathrm{HB}_{tth}}^{4} = \frac{\partial u_x(i-1,j,t)}{\partial^2 t} \qquad M_{\mathrm{HB}_{tth}}^{5} = \frac{\partial u_x(i,j,t)}{\partial^2 t} \tag{3-23}$$

$$M_{\mathrm{VB}_{zxh}}^{2} = -\frac{\partial u_x(i,j,t)}{\partial x} \qquad M_{\mathrm{VB}_{zxh}}^{5} = -\frac{\partial u_x(i,j,t)}{\partial x} \tag{3-24}$$

$$M_{\mathrm{VB}_{zzv}}^{2} = -\frac{\partial u_z(i,j,t)}{\partial z} \qquad M_{\mathrm{VB}_{zzv}}^{5} = -\frac{\partial u_z(i,j,t)}{\partial z} \tag{3-25}$$

$$\begin{aligned} M_{\mathrm{VB}_{xzh}}^{1} &= -\frac{\partial u_x(i-1,j,t)}{\partial z} & M_{\mathrm{VB}_{xzh}}^{3} &= -\frac{\partial u_x(i,j,t)}{\partial z} \\ M_{\mathrm{VB}_{xzh}}^{4} &= -\frac{\partial u_x(i-1,j,t)}{\partial z} & M_{\mathrm{VB}_{xzh}}^{6} &= -\frac{\partial u_x(i,j,t)}{\partial z} \end{aligned} \tag{3-26}$$

$$\begin{aligned} M_{\mathrm{VB}_{xxv}}^{1} &= -\frac{\partial u_z(i,j-1,t)}{\partial x} & M_{\mathrm{VB}_{xxv}}^{3} &= -\frac{\partial u_z(i,j-1,t)}{\partial x} \\ M_{\mathrm{VB}_{xxv}}^{4} &= -\frac{\partial u_z(i,j,t)}{\partial x} & M_{\mathrm{VB}_{xxv}}^{6} &= -\frac{\partial u_z(i,j,t)}{\partial x} \end{aligned} \tag{3-27}$$

$$M_{\mathrm{VB}_{ttv}}^{2} = \frac{\partial u_z(i,j-1,t)}{\partial^2 t} \qquad M_{\mathrm{VB}_{ttv}}^{5} = \frac{\partial u_z(i,j,t)}{\partial^2 t} \tag{3-28}$$

Because, as mentioned before, the FWI is generally in the Mie scattering regime, the damping effects occurred when seismic waves propagate from one grid to an adjacent grid cannot be negligible contrary to the Rayleigh scattering. For this reason, all basis virtual sources can be regarded as a basis of the full virtual source. However, strictly saying, the two groups ([$\mathrm{HB}_{xxh}, \mathrm{HB}_{xzv}$] and [$\mathrm{VB}_{zxh}, \mathrm{VB}_{zzv}$]) are linear dependent. Therefore both members of these two groups cannot be called as a basis of the full virtual source separately. However, in my experience, it is just valid in the staggered-grid FDM and, in the other numerical schemes such as cell-based FDM (Min et al., 2004) and finite element method, all the basis virtual sources are mathematical bases of the full virtual source.

Using these basis virtual sources, the elastic virtual source for the arbitrary parameter, p_i, can be generalized as a linear combination of the basis virtual sources as follows:

$$\begin{aligned}\mathbf{f}_{p_i}^{\mathrm{full}} &= \frac{\partial c_{11}}{\partial p_i}\mathrm{HB}_{xxh} + \frac{\partial c_{33}}{\partial p_i}\mathrm{VB}_{zzv} + \frac{\partial c_{13}}{\partial p_i}\left(\mathrm{HB}_{xzv} + \mathrm{VB}_{zxh}\right) \\ &\quad + \frac{\partial c_{13}}{\partial p_i}\left(\mathrm{HB}_{zzh} + \mathrm{HB}_{zxv} + \mathrm{VB}_{xzh} + \mathrm{VB}_{xxv}\right) + \frac{\partial \rho}{\partial p_i}\left(\mathrm{HB}_{tth} + \mathrm{VB}_{ttv}\right)\end{aligned} \quad (3\text{-}29)$$

In eq. (3-29), we regard that the partial derivatives of the stiffness and mass coefficients with respect to the model parameter determine which basis virtual sources are activated or not. This mechanism depends on the parameterization. In addition, the eq. (3-29) means that the characteristics of full virtual source are determined by the combination of the behaviour of the basis virtual sources. In other words, if we know how the basis virtual sources behave during the FWI, we can forecast the behaviour of the full virtual source for each parameter even though we use different parameterization.

From eq. (3-29), we can guess that only difference caused by different assumption of medium is the coefficients of the linear combination. For this reason, the analysis for the isotropic basis virtual source can be also applied to elastic VTI, visco-elastic and visco-elastic VTI cases. In addition, because the group of basis virtual sources for the acoustic media is a subset of that for the elastic media, the radiation pattern of the basis virtual source through the acoustic or acoustic VTI media can be interpreted using the analysis for the elastic basis virtual sources by taking some characteristics related to only P-P reflections.

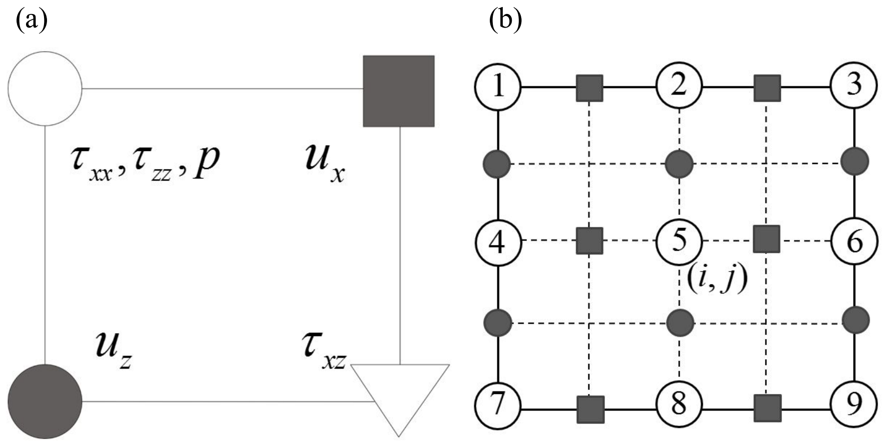


Figure 3-2 (a) The staggered-grid used for the analysis on the virtual source. All the physical parameters are located on the same grid with that of normal stresses (Graves, 1996). (b) The numbers for indicating the location of the entries of basis virtual sources through eqs. from (3-9) to (3-18)

3.2.2 Physical Description of Basis Virtual Source

To compute the gradient direction in the elastic FWI, the residual wavefields for horizontal and vertical directions are cross-correlated with horizontal and vertical displacements of partial derivative wavefields, respectively. Therefore, it would be helpful to analyze particle motions generated by the virtual source separately depending on the direction. To do so, in this part, the basis virtual source is physically described using moment tensor as previous studies did for the parameter (Wu and Aki, 1985; Burridge et al., 1998). Figures 3-3 and 3-4 show the moment tensor description of the horizontal and vertical basis virtual sources, respectively. The direction of moment tensor is determined when the first motion of incidence wave is positive and when the waves are attenuated propagating through the media. Because the incidence waves also oscillate, the direction of virtual sources is continuously changed. For this reason, I show only the first motion of P-P reflection (black solid arrows) and S-P reflection (black dashed arrows).

From the figures, we can notice that the basis virtual sources, which derived by horizontal normal stress like HB_{xxh} and HB_{xzv}, also act as horizontally tensional forces as shown in Figures 3-3a and 3-3b. This is because the scattering pattern is determined by the horizontal derivatives as shown in eq. (3-11). On the other hand, the basis virtual sources, VB_{zxh} and VB_{zzv} that are derived by the vertical normal stress, act as vertically tensional forces due to the vertical derivative. The shear stress-derived basis virtual sources (HB_{zzh}, HB_{zxv}, VB_{xzh} and VB_{xxv}) also acts like shear forces as Figures 3-3 and 3-4 show. If we focus the direction of the shear force along both the horizontal and vertical directions, we notice that they act like a double-coupled force together, which is the mechanism of earthquake.

On the other hand, as shown in Figures 3-3e and 3-4e, the HB_{tth} and VB_{ttv} act as a unidirectional force. This phenomenon is because these basis virtual sources do not have wave propagators and they only have the term with related to mass of particle. This might means that, if the incidence wave acts as an external force to the system, the particles which forced by HB_{tth} and VB_{ttv} interact with the external force following the law of universal gravitation between incidence waves and particles as shown in Figures 3-5 and 3-6. For this reason, the mass-induced basis virtual sources act like a universal gravitation between incidence waves and particles.

From these characteristics of basis virtual sources, we can classify all the basis virtual sources by four groups: horizontal normal stress-induced group (HB_{xxh} and HB_{xzv}), vertical normal stress-induced group (VB_{zxh} and VB_{zzv}), double-coupled force-like group (HB_{zzh}, HB_{zxv}, VB_{xzh} and VB_{xxv}) and universal gravitation-like group (HB_{tth} and VB_{ttv}). These grouping has some practical meanings, because, in many cases, the members of each group act together. However, in my opinion, each normal stress group should be divided because each member has different characteristics due to the S-wave. In Figures 3-3 and 3-4, the first S-P motions (black dashed arrows) are also displayed, which adds some complexities to the elastic FWI. For example, we can notice that the S-P motions of HB_{xxh}, which is induced by horizontal components of the incidence wave, and HB_{xxv}, which is induced by vertical components of incidence wave, are reversed. This reversed S-P motions, which are caused by the first motion of incidence SV waves as shown in Figure 3-7, cause some trade-off relationships during the FWI and make the FWI easily go to local minima according to circumstances as discussed in Appendix A.

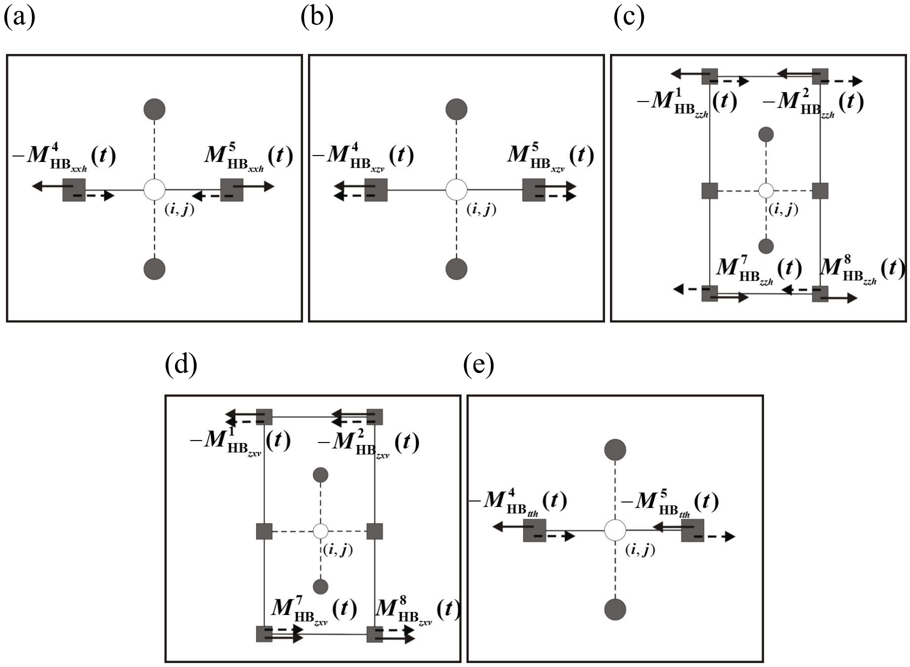


Figure 3-3 Moment tensor description of *horizontal* basis virtual sources: (a) HB_{xxh}, (b) HB_{xzv}, (c) HB_{zzh}, (d) HB_{zxv} and (e) HB_{tth}. Black solid and dashed lines indicate the first motion induced by incidence P- and S-waves, respectively.

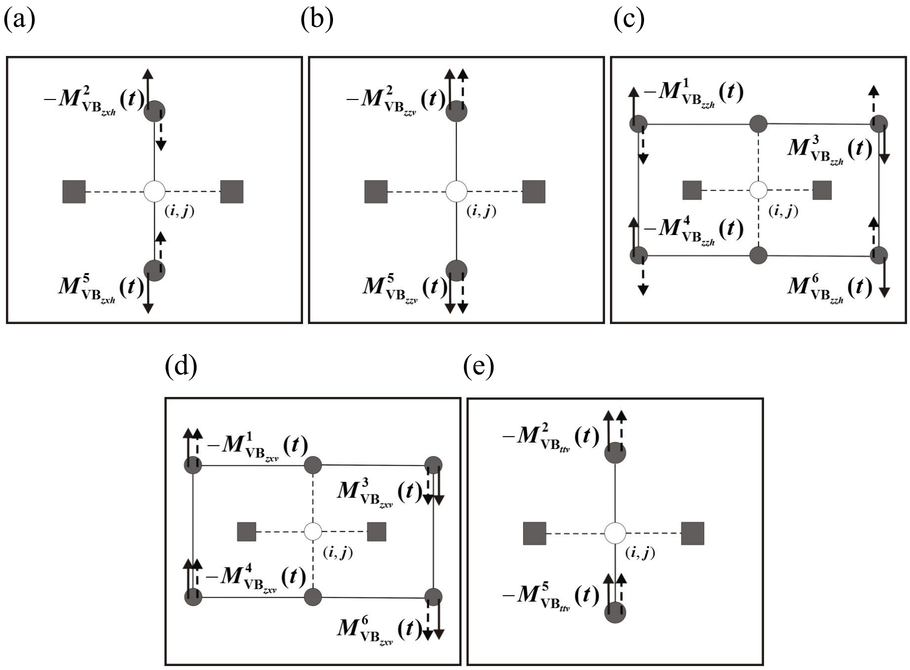


Figure 3-4 Moment tensor description of *vertical* basis virtual sources: (a) VB_{zxh}, (b) VB_{zzv}, (c) VB_{xzh}, (d) VB_{xxv} and (e) VB_{ttv}. Black solid and dashed lines indicate the first motion induced by incidence P- and S-waves, respectively.

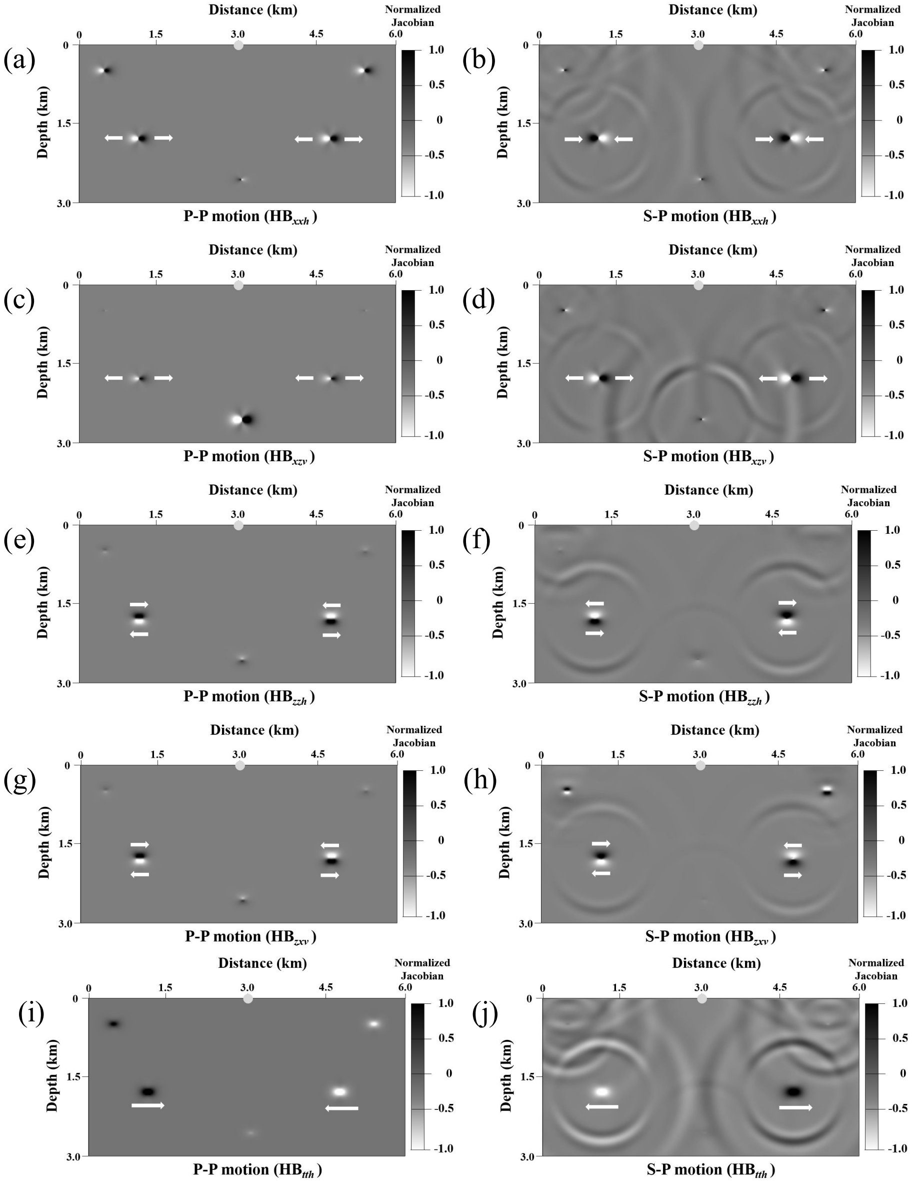


Figure 3-5 *Horizontal* displacements of Jacobian induced by HB_{xxh} (a and b), HB_{xzv} (c and d), HB_{zzh} (e and f), HB_{zxv} (g and h) and HB_{tth} (i and j) describing the first motion of P-P waves (a, c, e, g and i) and S-P waves (b, d, f, h and j), respectively. The circle and arrows indicate the source and the first motion.

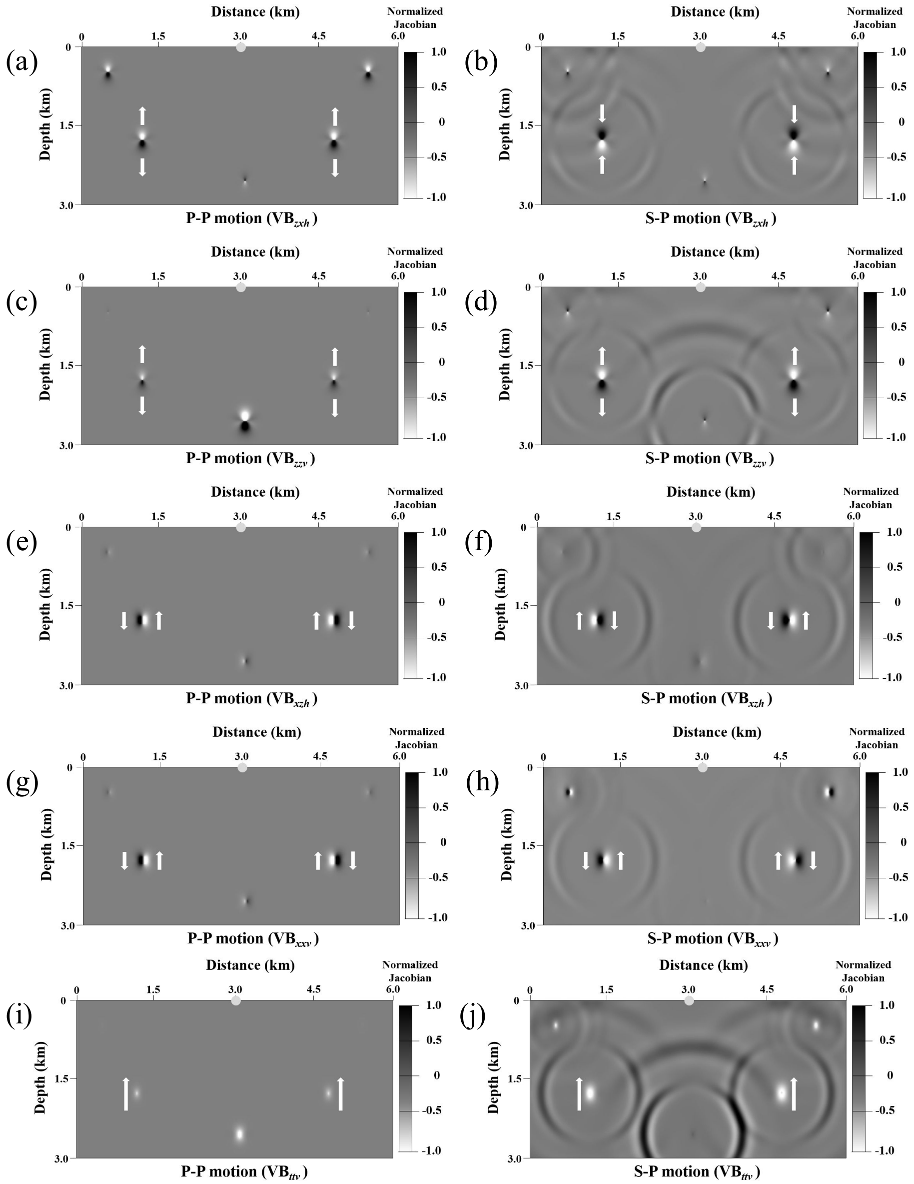


Figure 3-6 *Vertical* displacements of Jacobian induced by VB_{zxh} (a and b), VB_{zzv} (c and d), VB_{xzh} (e and f), VB_{xxv} (g and h) and VB_{ttv} (i and j) describing the first motion of P-P waves (a, c, e, g and i) and S-P waves (b, d, f, h and j), respectively. The circle and arrows indicate the source and the first motion.

Considering that the acoustic wave equation does not have shear stress, we can guess, in the acoustic media, the full virtual source only contains hydrostatic pressure or unidirectional pressure. Therefore, basis virtual sources in the acoustic media are HB_{xxh}, HB_{xzv}, HB_{tth}, VB_{zxh}, VB_{zzv} and VB_{ttv}. On the other hand, in elastic media, the behavior of full virtual source is more complex due to the shear stress-induced basis virtual sources.

To verify if basis virtual sources act like aforementioned forces, Jacobian induced by only each basis virtual source are calculated as shown in Figures 3-5 and 3-6. The five subsurface inclusions are assumed with the nearly same distances from the seismic source. By checking the first motion of Jacobian when the P-P and S-P waves are generated and the radiation pattern of Jacobian, we can notice that the behavior of each virtual source well reflects the scattering pattern of corresponding moment tensor. In addition to this phenomenon caused by directional property, we can also notice that the strengths of each basis virtual source are different over the domain related to the incidence angle.

To figure out this phenomenon, I discuss how each basis virtual source controls the pattern of incidence wave, which is related to the momentum of the basis virtual source as shown from eq. (3-19) to eq. (3-28). Because the momentum of basis virtual source is determined depending on the strength of strain induced by both incidence P- and S-waves, it is helpful to derive the strain using Helmholtz decomposition as follows:

$$u_x = \frac{\partial \Phi}{\partial x} - \frac{\partial \Psi}{\partial z} \qquad (3\text{-}30)$$

$$u_z = \frac{\partial \Phi}{\partial z} - \frac{\partial \Psi}{\partial x} \qquad (3\text{-}31)$$

The terms, Φ and Ψ, denote the scalar and vector potentials, respectively. From above two equations, we can derive the normal strains and components of shear strain as expressed by

$$\frac{\partial u_x}{\partial x} = \frac{\partial^2 \Phi}{\partial x^2} - \frac{\partial^2 \Psi}{\partial x \partial z}\ , \qquad (3\text{-}32)$$

$$\frac{\partial u_z}{\partial z} = \frac{\partial^2 \Phi}{\partial z^2} + \frac{\partial^2 \Psi}{\partial x \partial z}, \qquad (3\text{-}33)$$

$$\frac{\partial u_x}{\partial z} = \frac{\partial^2 \Phi}{\partial z \partial x} - \frac{\partial^2 \Psi}{\partial^2 z} \qquad (3\text{-}34)$$

and

$$\frac{\partial u_z}{\partial x} = \frac{\partial^2 \Phi}{\partial x \partial z} + \frac{\partial^2 \Psi}{\partial^2 x}. \qquad (3\text{-}35)$$

Figure 3-7 shows the incidence displacement fields and strains, which are calculated for all the elements when the source is located at middle of the surface. For the normal strains (Figure 3-7c and 3-7d), we can notice that the horizontal and vertical normal strains induced by P-waves have large values along the horizontal and vertical directions, respectively. On the other hand, the normal strains induced by S-waves have positive and negative values that corresponds to eqs. (3-32) and (3-33). This means that, if the basis virtual sources induced by horizontal and vertical normal strains act together, the coverage of incidence P-waves can be improved but the incidence S-waves are cancelled each other. In other words, the virtual source that consists of both horizontal and vertical normal strains cannot generate S-P and S-S motions as Tarantola (1984) showed. In addition, when the normal strain acts as tensional or compressional forces along both horizontal and vertical directions (isotropic explosive source), the P-S motions are also compensated each other as shown in Appendix A. As shown in Figures 3-7e and 3-7f, the components of shear strains also satisfy the eqs. (3-34) and (3-35).

From above results, I conclude that the pattern of incidence wave as a momentum of virtual source is determined by the inner spatial derivatives, which indicate the strain in elastic media.

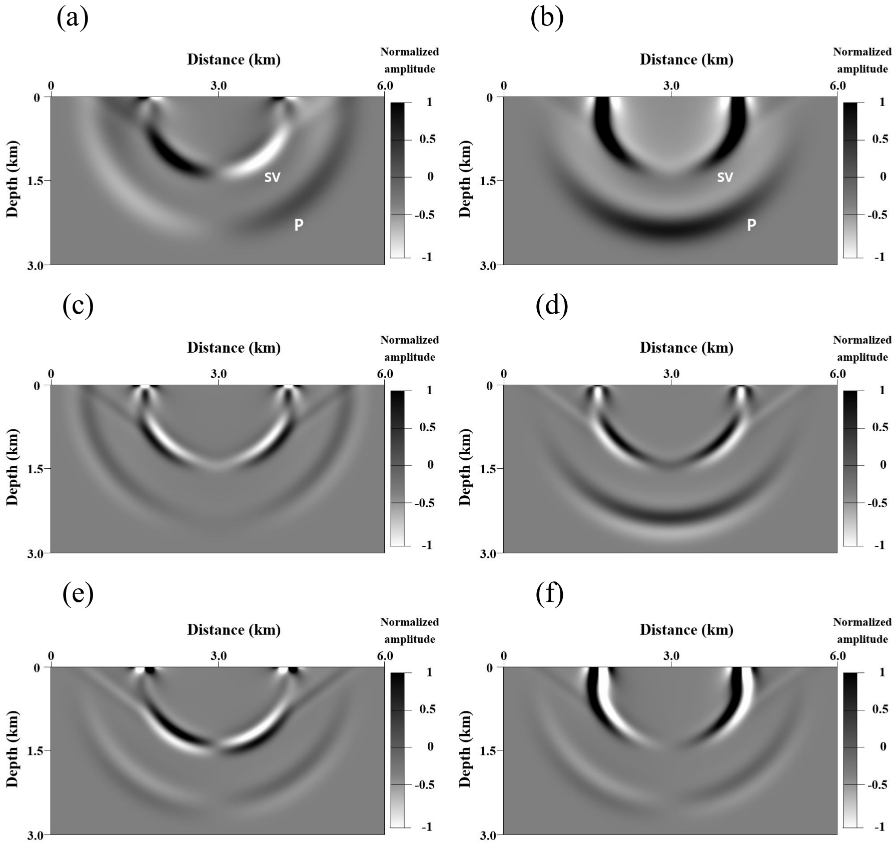


Figure 3-7 The (a) horizontal and (b) vertical displacements induced by a vertical body force on the surface and the distributions of resulting strains at subsurface medium: (c) $\frac{\partial u_x}{\partial x}$, (d) $\frac{\partial u_z}{\partial z}$, (e) $\frac{\partial u_x}{\partial z}$ and (f) $\frac{\partial u_z}{\partial x}$

3.2.3 Numerical Example for Simple Layer Model

To examine how the different kind of basis virtual source influence on the descent direction, I calculate the descent direction for the simple one layer model (Figure 3-8) at first iteration with steepest descent approach, in which the Hessian matrix is not applied. The frequency-domain FWI is applied using the finite element method for the forward modeling. The parameters used for the FWI are listed in Table 3-1. The initial guesses are the same as the background parameters so that only discrepancy is induced by the anomalous body.

In the frequency-domain FWI with backpropagation as shown in eq. (2-15), the gradient direction can be obtained by cross-correlating the virtual source and backpropagated wavefields, In this case, basis virtual source changes the sensitivity of gradient direction depending on its moment tensor, which originates by outer spatial derivatives. For example, if the moment tensor (horizontal derivative) of basis virtual source has horizontally tensional or compressional form, it provides gradient direction by measuring the horizontal variations of the backpropagated wavefields. On the other hand, the basis virtual source, whose moment tensor is a vertically tensional or compressional form, measures the vertical variations of the backpropagated wavefields. However, considering HB_{tth} and VB_{ttv} which have the unidirectional moment tensor, they measure the backpropagated wavefields themselves and provides poorly sensitive gradient direction. In other words, the sensitivity of gradient direction (or Jacobian) is determined by the moment tensor (outer spatial derivative) of the full virtual source.

To verify these phenomena, I calculate the single-shot-descent direction for the thick one layer model (Figure 3-8). Figures 3-10 and 3-11 show the descent direction obtained by 5 horizontal basis virtual sources and 5 vertical basis virtual sources, respectively. From figures, we can notice four main characteristics for behaviors of basis virtual sources. At first, we notice that each basis virtual source covers different area depending on the inner spatial derivatives that are strains, reflecting its main incidence angles as shown in Figure 3-7. This make us convince that, as mentioned in the previous subsection, the distribution of strains depending on the incidence angle determines the magnitude of basis virtual source. Figures 3-11 shows the descent direction of some groups for shear stress-induced basis virtual sources. We summed all the shear stress-induced basis virtual sources because, in most cases, all the shear stress-induced basis virtual sources (they share the same parameter) act together like double-coupled forces. From the descent direction obtained by the double-coupled forces (Figure 3-11c), we notice that the anomalous layer located at intermediate incidence angle is well described.

Secondly, each basis virtual source provides differently resolved descent direction because the sensitivity of the descent direction is also determined by whether corresponding basis virtual source measures horizontal (HB_{xxh} and HB_{xzv}, VB_{xzh} and VB_{xxv}) and vertical variations (VB_{zxh}, VB_{zzv}, HB_{zzh} and HB_{zxv}) of the backpropagated wavefields or backpropagated wavefields themselves (HB_{tth} and VB_{ttv}). For example, comparing the Figures 3-10b

(VB_{zzv}) and 3-10e (VB_{ttv}), we can notice that the VB_{ttv} provides less sensitive descent direction near the layer because it measures backpropagated wavefields themselves to describe the anomalous body. If we focus on the top boundary of layer in Figure 3-10e, the descent directions for the anomalous layer (black-colored) are nearly same magnitude (strictly, weaker due to damping effects) with cycle skipping artifacts (white-colored) above them. On the other hand, in Figure 3-10b, the magnitude of the descent direction for the anomalous layer obtained by the VB_{zzv} becomes stronger by measuring vertical variations of the backpropagated wavefields. This phenomenon is also well observed in Figures 3-9a for the HB_{xxh} and 3-9e for the HB_{tth}. From these results, I convince that the basis virtual sources obtained from time derivatives have the poorest sensitivity at the boundary of the anomalous layer among 10 basis virtual sources.

Thirdly, as Figures 3-10a (HB_{xxh}), 3-10b (HB_{xzv}), 3-10c (VB_{xzh}) and 3-10d (VB_{xxv}) show, the descent directions obtained from horizontal variation of the backpropagated wavefields have some problems to detect exact depth of the anomalous layer, which might be related to the Fresnel zone.

Fourthly, as Figures 3-9b and 3-10a show, the descent directions obtained from HB_{xzv} and VB_{zxh} tend to go to negative direction even though all the true parameters, except Poisson's ratio, are larger than initial guesses which means that the gradient should be positive at the anomalous layer. This is because these two basis virtual sources have strong trade-off depending on amplitudes of P-waves and S-waves in residual wavefields.

By synthesizing aforementioned four observations, we can convince that full virtual source, which consists of the normal strain-induced basis virtual sources (HB_{xxh}, VB_{zxh}, HB_{xzv} and VB_{zzv}) and double-coupled basis virtual sources (HB_{zzh}, HB_{zxv}, VB_{xzh} and VB_{xxv}) will provide wide-coverage and highly resolved gradient directions. From this idea, I find out some limitations of conventional parameterizations for isotropic media and suggest the new parameterization using the Poisson's ratio in the next subchapter. These analyses are also applicable to the elastic VTI case as shown in Appendix C.

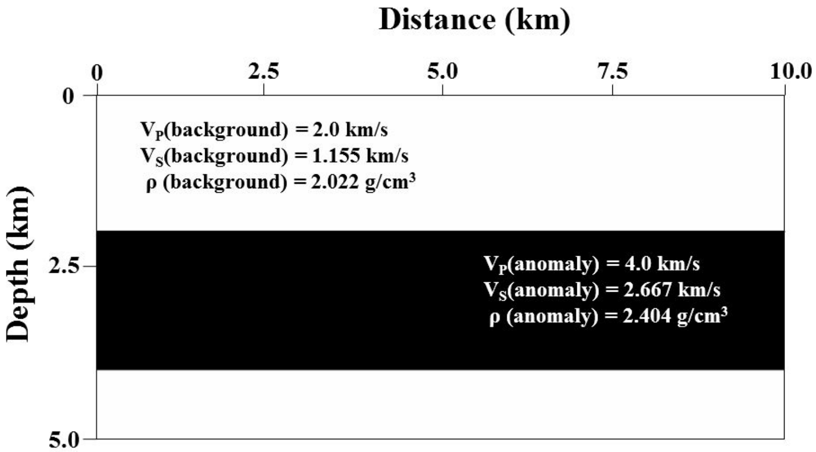


Figure 3-8 The simple one layer model.

Table 3-1 Inversion parameters used for one layer model.

Dimension	No. of shots	No. of receivers	Interval of receivers	Recording time	Maximum Frequency	Minimum Frequency
5.0 km × 2.5 km	1	500	0.01 km	3 sec	10 Hz	1/3 Hz

(a)
(b)
(c)

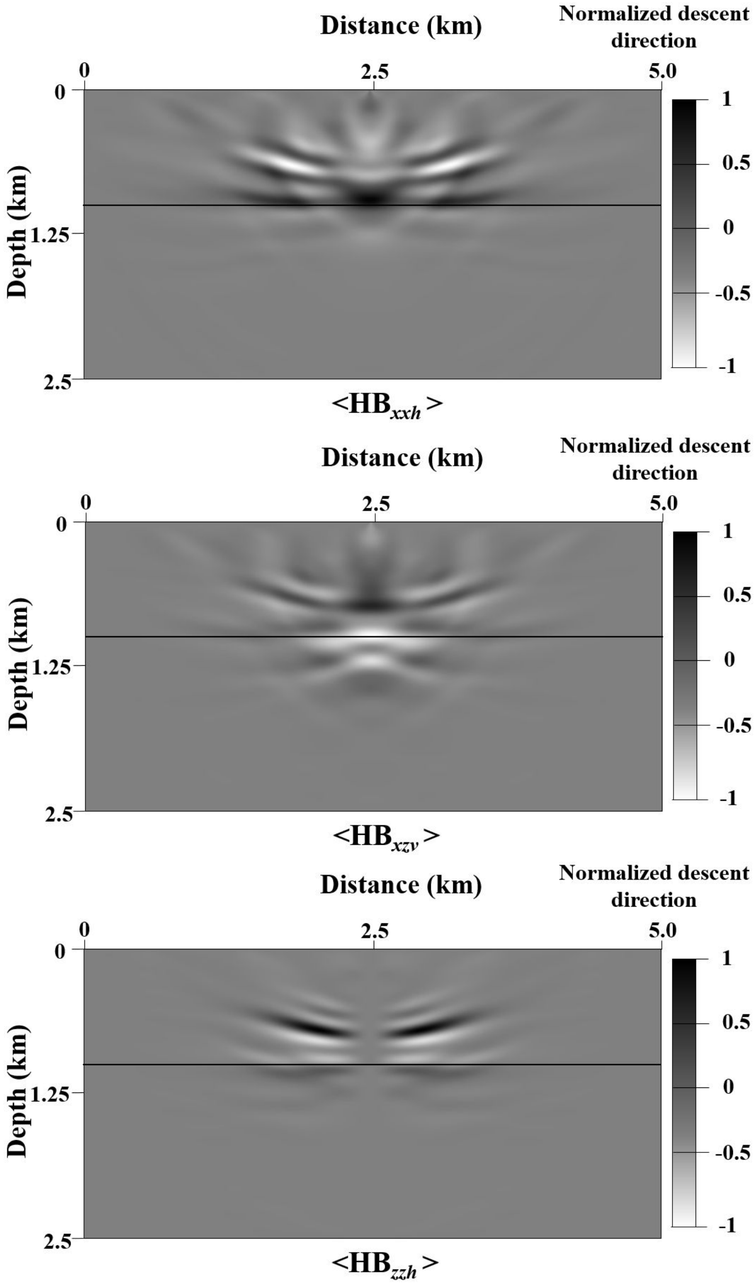
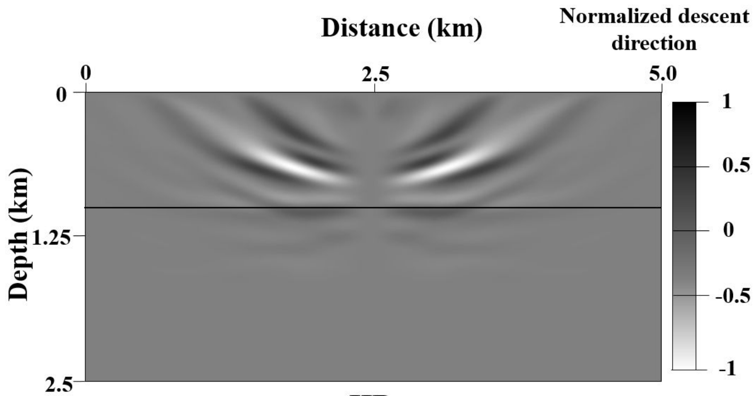


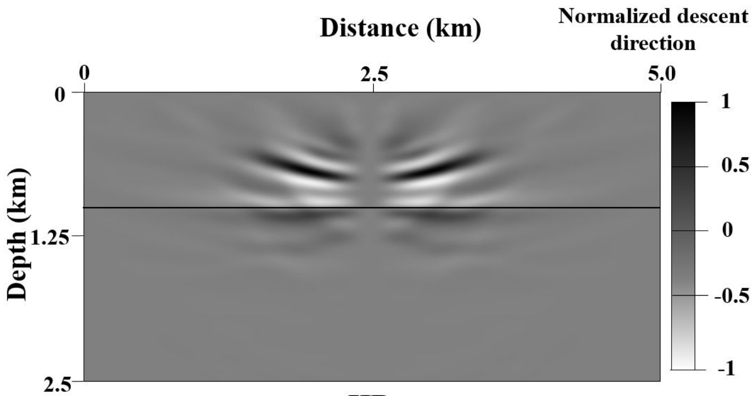
Figure 3-9 Single-shot descent directions obtained by using only (a) HB$_{xxh}$, (b) HB$_{xzv}$, (c) HB$_{zzh}$, (d) HB$_{zxv}$ and (e) HB$_{tth}$.

(d)



$<\mathbf{HB}_{zxv}>$

(e)



$<\mathbf{HB}_{tth}>$

Figure 3-9 (Continued)

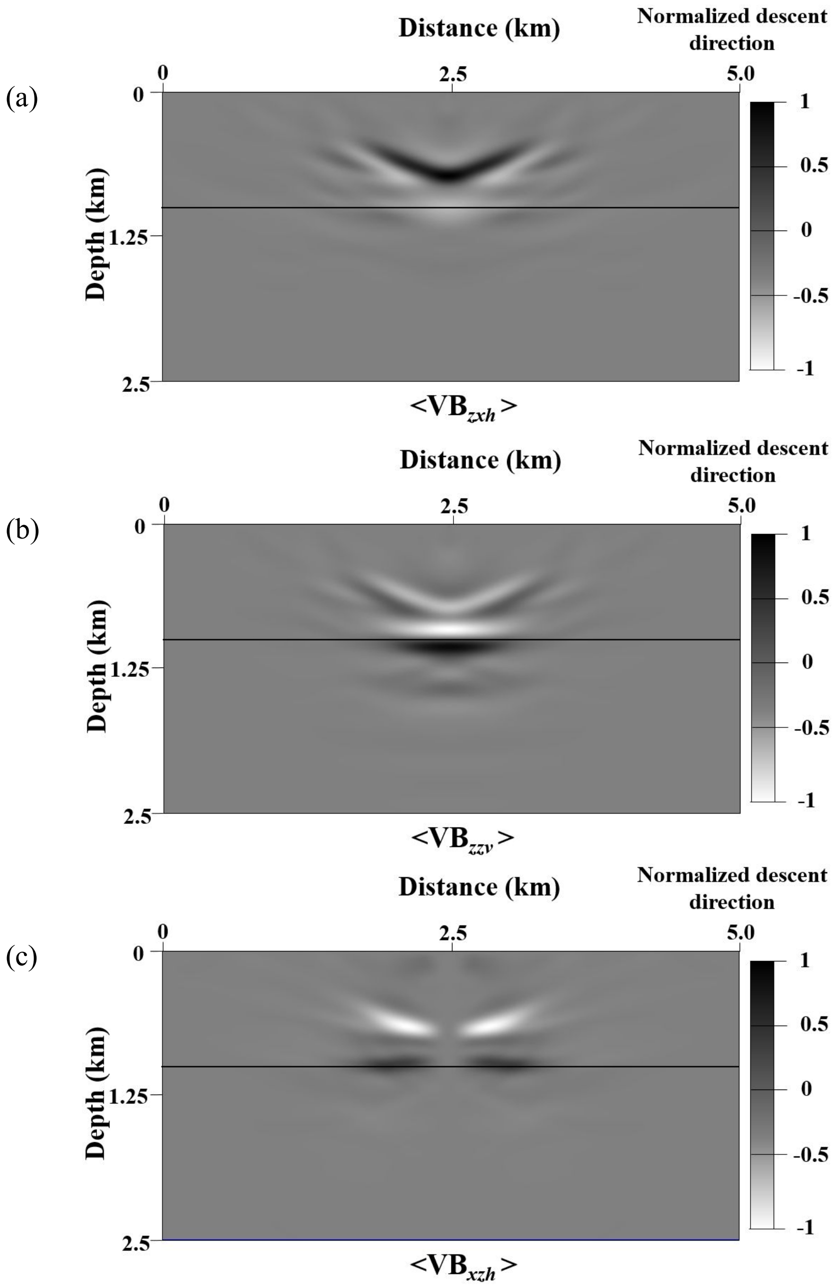


Figure 3-10 Single-shot descent directions obtained by using only (a) VB$_{zxh}$, (b) VB$_{zzv}$, (c) VB$_{xzh}$, (d) VB$_{xxv}$ and (e) VB$_{ttv}$.

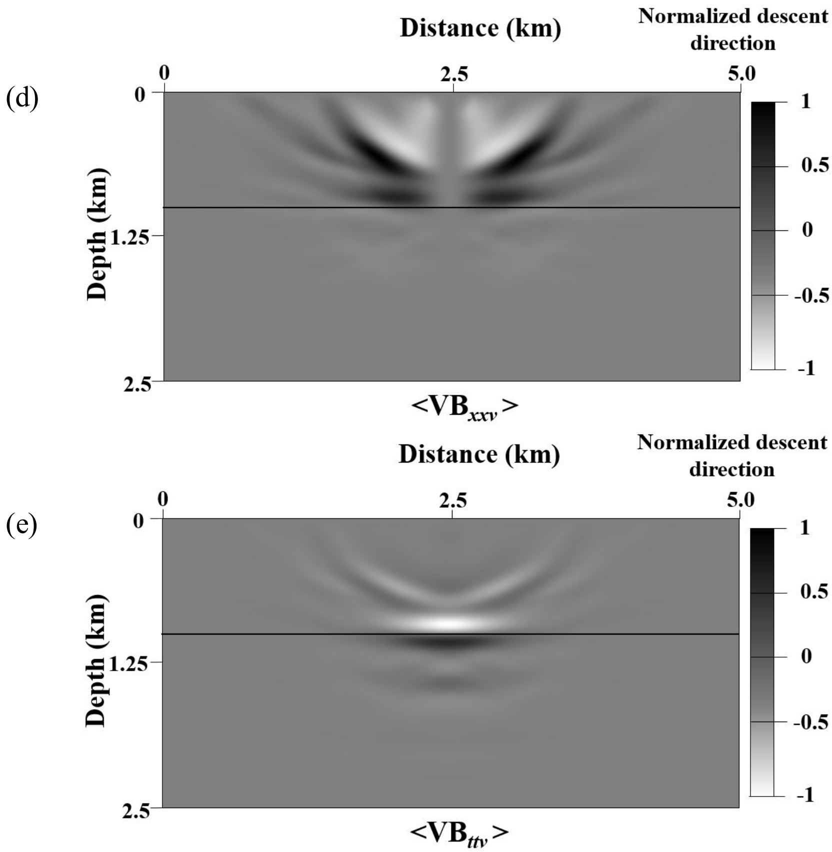


Figure 3-10 (Continued)

(a)

(b)

(c)

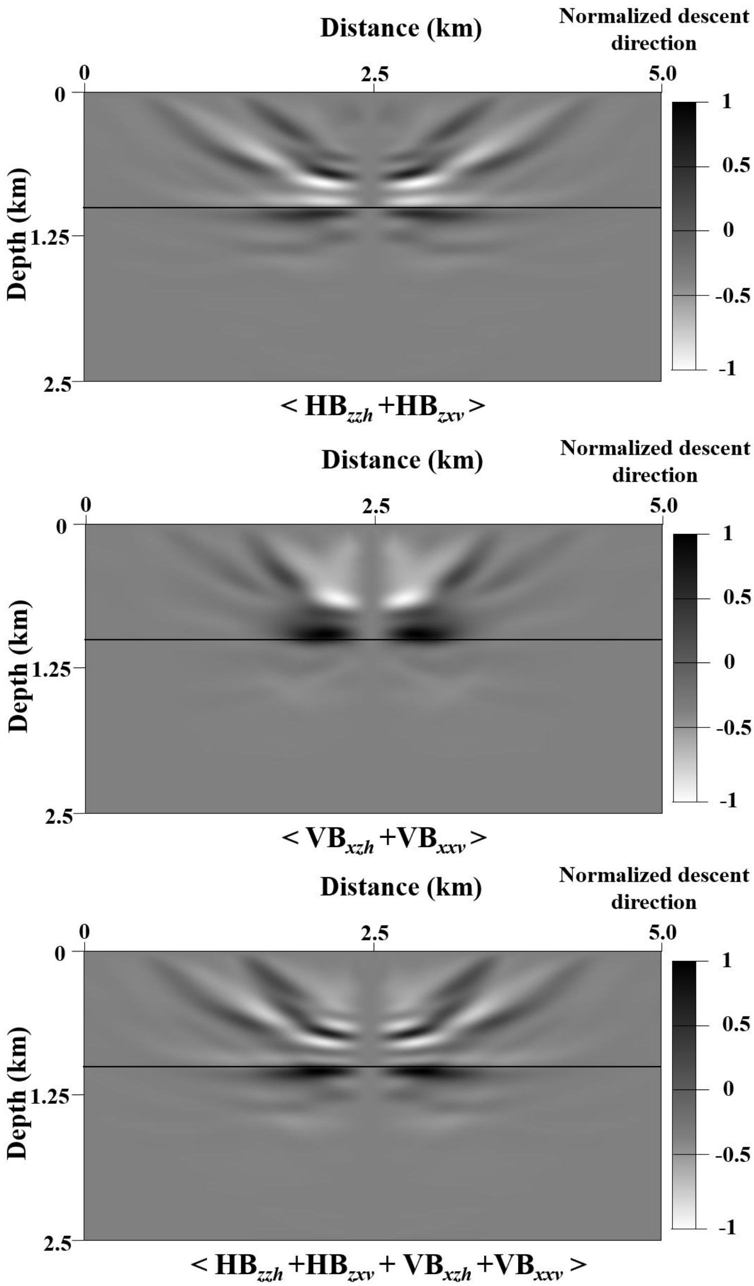


Figure 3-11 Single-shot descent directions obtained by using (a) HB$_{zzh}$+HB$_{zxv}$, (b) VB$_{xzh}$+VB$_{xxv}$ and (c) HB$_{zzh}$+HB$_{zxv}$+VB$_{xzh}$+VB$_{xxv}$ (double-coupled forces).

3.3 New Parameterization for Isotropic Elastic FWI

3.3.1 Conventional parameterization for isotropic elastic media

For isotropic elastic FWI, most previous studies have used two kinds of parameterizations. One uses the elastic wave equation parameterized by Lamé constants (λ and μ) and density (ρ), which I call 'Isotropic Parameter Group-I (IPG-1)' throughout the paper. The other parameterization method represents the elastic wave equation with P-wave (v_p) and S-wave (v_s) velocities and density, which will be called 'Isotropic Parameter Group-II (IPG-II)'.

The isotropic elastic wave equation using the IPG-I can be expressed by

$$\begin{aligned} \rho\frac{\partial^2 u_x}{\partial t^2} &= \frac{\partial}{\partial x}\left[(\lambda+2\mu)\frac{\partial u_x}{\partial x}+\lambda\frac{\partial u_z}{\partial z}\right]+\frac{\partial}{\partial z}\left[\mu\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right] \\ \rho\frac{\partial^2 u_z}{\partial t^2} &= \frac{\partial}{\partial z}\left[(\lambda+2\mu)\frac{\partial u_z}{\partial z}+\lambda\frac{\partial u_x}{\partial x}\right]+\frac{\partial}{\partial x}\left[\mu\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right] \end{aligned}, \quad (3\text{-}36)$$

in which other parameters, such as P-wave velocity (v_p), S-wave velocity (v_s), Poisson's ratio (v) that is defined as the ratio of transverse to longitudinal strains, can be calculated as follows:

$$v_p = \sqrt{\frac{\lambda+2\mu}{\rho}} \quad (3\text{-}37)$$

$$v_s = \sqrt{\frac{\mu}{\rho}} \quad (3\text{-}38)$$

$$\frac{v_s}{v_p} = \sqrt{\frac{1-2v}{2-2v}} \quad (3\text{-}39)$$

The isotropic elastic wave equation using the IPG-II can be expressed by

$$\begin{aligned}
\rho\frac{\partial^2 u_x}{\partial t^2} &= \frac{\partial}{\partial x}\left[\rho v_p^2\frac{\partial u_x}{\partial x}+\left(\rho v_p^2-2\rho v_s^2\right)\frac{\partial u_z}{\partial z}\right]+\frac{\partial}{\partial z}\left[\rho v_s^2\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right]\\
\rho\frac{\partial^2 u_z}{\partial t^2} &= \frac{\partial}{\partial z}\left[\rho v_p^2\frac{\partial u_z}{\partial z}+\left(\rho v_p^2-2\rho v_s^2\right)\frac{\partial u_x}{\partial x}\right]+\frac{\partial}{\partial x}\left[\rho v_s^2\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right]
\end{aligned} \quad . \text{(3-40)}$$

As previous studies have reported (Jeong et al. 2012; Köhn et al. 2012), the IPG-I provides better estimation for the Lamé constants or wave velocities but is weak at recovering density structures. On the other hand, the IPG-II provides much better parameter estimation when all the three parameters are simultaneously inverted. Based on this fact, Jeong et al. (2012) suggest to use the two-stage FWI strategy, in which the FWI is sequentially performed by the IPG-II with a chain rule following the IPG-I. Köhn et al. (2012) compared three parameter groups, which are (λ, μ, ρ), (v_p, v_s, ρ) and (I_p, I_s, ρ), respectively, and conclude that choice of the seismic velocities in favor of Lamé parameters can improve the image quality and that the model parameterization has mainly an influence on the density inversion result. However, in my opinion, the latter conclusion is only valid in the FWI for the low-frequency data because the behavior of high-frequency virtual sources is more complicated. The impedance-density parameter group is not considered in this paper, because the results might be nearly similar with those obtained using the seismic velocities as Köhn et al. (2012) insisted.

By contrast with previous study of Köhn et al. (2012), in this paper, three parameter groups will be discussed, those are aforementioned IPG-I and IPG-II and the new parameter group using the Poisson's ratio called 'Isotropic Parameter Group-III' throughout the paper. In the following section, the new parameterization method using Poisson's ratio will be introduced and the advantages of the new parameterization will be discussed.

3.3.2 New Parameterization using Poisson's ratio

When we use the parameter group using the Poisson's ratio instead of the S-wave velocity (IPG-III), the isotropic elastic wave equation can be expressed by

$$
\begin{aligned}
\rho \frac{\partial^2 u_x}{\partial t^2} &= \frac{\partial}{\partial x}\left[\rho v_p^2 \frac{\partial u_x}{\partial x} + \left(\rho v_p^2 - 2\rho v_p^2 \left(\frac{1-2\nu}{2-2\nu} \right) \right) \frac{\partial u_z}{\partial z} \right] \\
&+ \frac{\partial}{\partial z}\left[\rho v_p^2 \left(\frac{1-2\nu}{2-2\nu} \right) \left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x} \right) \right] \\
\rho \frac{\partial^2 u_z}{\partial t^2} &= \frac{\partial}{\partial z}\left[\rho v_p^2 \frac{\partial u_z}{\partial z} + \left(\rho v_p^2 - 2\rho v_p^2 \left(\frac{1-2\nu}{2-2\nu} \right) \right) \frac{\partial u_x}{\partial x} \right] \\
&+ \frac{\partial}{\partial x}\left[\rho v_p^2 \left(\frac{1-2\nu}{2-2\nu} \right) \left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x} \right) \right]
\end{aligned}
. \qquad (3\text{-}41)
$$

The S-wave velocity can be estimated using eq. (3-39). The IPG-III is designed so that the resolution and robustness of the gradient direction for the P-wave velocity, one of the important parameter for seismic data processing, can be enhanced maintaining the advantage of the IPG-II, that is good estimation of the density. This idea is derived based on the scattering theory of the basis virtual source as explained previous subchapters.

Using the linear combination of the basis virtual source, three full virtual sources for the IPG-I and IPG-II can be expressed as follows:

$$
\begin{aligned}
&\mathbf{f}_{\lambda_i}^{\text{full}} = \text{HB}_{xxh} + \text{HB}_{xzv} + \text{VB}_{zxh} + \text{VB}_{zzv} \\
&\mathbf{f}_{\mu_i}^{\text{full}} = \text{HB}_{xxh} + \text{HB}_{zzh} + \text{HB}_{zxv} + \text{VB}_{zzv} + \text{VB}_{xzh} + \text{VB}_{xxv}\,, \\
&\mathbf{f}_{\rho_i}^{\text{full}} = \text{HB}_{tth} + \text{VB}_{ttv}
\end{aligned}
\tag{3-42}
$$

$$
\begin{aligned}
&\mathbf{f}_{v_{p,i}}^{\text{full}} = 2\rho v_p\left(\text{HB}_{xxh} + \text{HB}_{xzv} + \text{VB}_{zxh} + \text{VB}_{zzv}\right) \\
&\mathbf{f}_{v_{s,i}}^{\text{full}} = -4\rho v_s\left(\text{HB}_{xzv} + \text{VB}_{zxh}\right) + 2\rho v_s\left(\text{HB}_{zzh} + \text{HB}_{zxv} + \text{VB}_{xzh} + \text{VB}_{xxv}\right) \\
&\mathbf{f}_{\rho_i}^{\text{full}} = v_p^2\left(\text{HB}_{xxh} + \text{VB}_{zzv}\right) + \left(v_p^2 - 2v_s^2\right)\left(\text{HB}_{xzv} + \text{VB}_{zxh}\right) \\
&\qquad + v_s^2\left(\text{HB}_{zzh} + \text{HB}_{zxv} + \text{VB}_{xzh} + \text{VB}_{xxv}\right) + \omega^2\left(\text{HB}_{tth} + \text{VB}_{ttv}\right)
\end{aligned}
\tag{3-43}
$$

The full virtual sources for the P-wave velocity, Poisson's ratio and the density in the IPG-III can be expressed by

$$
\begin{aligned}
&\mathbf{f}_{v_{p,i}}^{\text{full}} = 2\rho v_p\left(\text{HB}_{xxh} + \text{VB}_{zzv}\right) + 2\rho v_p\left(\frac{2\nu}{2-2\nu}\right)\left(\text{HB}_{xzv} + \text{VB}_{zxh}\right) \\
&\qquad + 2\rho v_p\left(\frac{1-2\nu}{2-2\nu}\right)\left(\text{HB}_{zzh} + \text{HB}_{zxv} + \text{VB}_{xzh} + \text{VB}_{xxv}\right) \\
&\mathbf{f}_{\sigma_i}^{\text{full}} = 4\rho v_p^2\left(\frac{1}{2-2\nu}\right)^2\left(\text{HB}_{xzv} + \text{VB}_{zxh}\right) \\
&\qquad - 2\rho v_p^2\left(\frac{1}{2-2\nu}\right)^2\left(\text{HB}_{zzh} + \text{HB}_{zxv} + \text{VB}_{xzh} + \text{VB}_{xxv}\right) \\
&\mathbf{f}_{\rho_i}^{\text{full}} = v_p^2\left(\text{HB}_{xxh} + \text{VB}_{zzv}\right) + v_p^2\left(\frac{2\nu}{2-2\nu}\right)\left(\text{HB}_{xzv} + \text{VB}_{zxh}\right) \\
&\qquad + v_p^2\left(\frac{1-2\nu}{2-2\nu}\right)\left(\text{HB}_{zzh} + \text{HB}_{zxv} + \text{VB}_{xzh} + \text{VB}_{xxv}\right) + \omega^2\left(\text{HB}_{tth} + \text{VB}_{ttv}\right)
\end{aligned}
\,, \tag{3-44}
$$

At first, as eqs. (3-43) and (3-44) show, the members of full virtual sources for the density in both IPG-II and IPG-III are the same, which means that the full virtual sources for the density show similar behaviors. On the other hand, in the IPG-I, the full virtual source for the density only consists of the HB_{tth} and VB_{ttv}, which induce unidirectional particle motions. These unidirectional basis virtual sources make the partial derivative wavefields less sensitive to variations of the residuals (Figure 3-10) and all kinds of reflected waves similarly contribute to the gradient direction (refer Appendix A). For these reasons, the inversion of the density in the IPG-I tends not to collaborate well with that of Lamé constants, which are induced by bi-direction particle motions. For the P-wave related parameters, λ and P-wave velocity, we notice that the virtual source only consists of normal stress-induced basis virtual sources. For this reason, the partial derivative wavefields only contain P-P scattered waves as we discussed in Chapter 3.2.2 and Appendix A. The absence of P-S, S-P and S-S scattered waves in partial derivative wavefields for the P-wave velocity causes three limitations. One is the poor resolution of the P-wave velocity in the IPG-II (or λ in the IPG-I). Second limitation is that the P-S, S-P and S-S reflected waves in residual wavefields act like coherent noise. Third one is high sensitivity of the gradient direction for the P-wave velocity to noises. These three limitations will be discussed in next subsection for the FWI with CTS model. To compensate above limitations, the IPG-III is designed so that the P-wave velocity provides highly resolved and robust gradient direction. This can be achieved by adopting Poisson's ratio as a direct parameter because the additional terms of full virtual source, those are the double-coupled basis virtual sources (HB_{zzh}, HB_{zxv}, VB_{xzh} and VB_{xxv}), generate P-S, S-P and S-S scattering.

For the S-wave related parameters, rigidity in the IPG-I, S-wave velocity in the IPG-II and the Poisson's ratio in the IPG-III, we can expect that the gradient direction of the rigidity is the best among the three parameter group because the other two parameters (S-wave velocity and Poisson's ratio) only have shear stress-induced basis virtual sources with HB_{xzv} and VB_{zxh}, which have strong trade-off effects (refer Appendix A).

3.3.3 Advantages of the New Parameterization

To verify one of the benefits of using the IPG-III, that is highly resolved P-wave image, I perform the elastic FWI for the synthetic CTS (Cross-Triangular-Square) model (Figure 3-12). This test is not the same inversion that done by Köhn et al. (2012). I just use the CTS model for the resolution analysis of the gradient (exactly, descent direction) image because various shaped-structures of the CTS model are good to determine which parameterization provides well resolved descent direction. The inversion settings are listed on Table 3-2. The initial guesses are the same as the background parameters so that only discrepancy is induced by the anomalous body. For more accurate analysis, in this numerical example, I apply single frequency (4 Hz) FWI with the approximate Hessian. Therefore, this approach is very close to the Gauss-Newton method except that I ignore the off-diagonal terms of the approximate Hessian matrix.

Figures 3-13a and 3-14a show the mono-frequency descent directions for the λ and the P-wave velocity obtained using the IPG-I and IPG-II, respectively, in the first iteration. I notice that the descent directions for the λ and P-wave velocity are the same because these two parameters are composed by the same basis virtual sources, those are HB_{xxh}, HB_{xzv}, VB_{zxh}

and VB_{zzv}, with same percentage (25 %) of each basis virtual source. This phenomenon indicates that, if some parameters have same basis virtual sources whether they are for the P-wave velocity or elastic constants, their behaviors are also similar during the FWI. However, the combination of these four basis virtual sources make the P-S, S-P and S-S motions be cancelled each other as shown in Appendix A. On the other hand, as shown in Figure 3-15a, the mono-frequency descent direction for the P-wave velocity obtained using the IPG-III is much improved, particularly around the cross structure, in the aspect of the spatial resolution. Other structures, triangular and rectangular-shaped structures, are also well resolved even though we only use 4 Hz component of data. This highly resolved descent direction for the P-wave velocity in the IPG-III can be achieved due to the additional terms of the full virtual source, those are the shear motion-induced basis virtual sources (HB_{zzh}, HB_{zxv}, VB_{xzh} and VB_{xxv}), because these double-coupled basis virtual sources generate not only P-P wave but also P-S, S-P and S-S waves that have shorter wavelength than P-P wave in mono-frequency (refer $\lambda_p = \frac{v_p}{\omega}$, $\lambda_s = \frac{v_s}{\omega}$ and Figure 3 in Virieux and Operto [2009]).

The other limitations of the IPG-I and IPG-II is high sensitivity to variations of Poisson's ratio and noises in observed data. Because, in the IPG-I and IPG-II, virtual sources for λ and P-wave velocity generate only P-P scattered waves, the maximum correlation is occurred when arrival times of P-P scattered wave in partial derivative wavefields and P-P (or S-P) reflected wave in residual wavefields are coincident. However, in the IPG-III, the maximum correlation is occurred when arrival times of all the P-P, P-S, S-P and S-S scattered waves are coincident with those in the residual wavefields because the virtual source for the P-wave velocity generates all kinds of scattered waves. This means that the maximum correlation for a perturbation of the P-wave velocity in the IPG-III is much larger than that in the IPG-II. For this reason, during the FWI, the gradient direction for the P-wave velocity in the IPG-II generates strong artifacts induced by zero-lag cross-correlation of P-P scattered waves with P-S, S-P and S-S reflected waves as shown in Figures 3-13a and 3-14a. For the same reason, the P-wave velocity in the IPG-II is also sensitive to the noise.

Figures 3-13b, 3-14b and 3-15b show the descent directions of the rigidity, S-wave velocity and Poisson's ratio obtained using the IPG-I, IPG-II and IPG-III, respectively. As similar with the previous example for the λ and P-wave velocity in the IPG-I and IPG-II, the descent directions of the S-wave velocity and Poisson's ratio obtained using the IPG-II and IPG-III, respectively, are also the same because they are composed by same basis virtual sources (HB_{xzv}, HB_{zzh}, HB_{zxv}, VB_{zxh}, VB_{xzh} and VB_{xxv}) with the same proportion of each basis virtual source (see eqs. (3-43) and (3-44)). However, the sign of the descent directions for the S-wave velocity and the Poisson's ratio is reversed because there exists some trade-off relations between the S-wave velocity and the Poisson's ratio. In other words, the negative direction of the descent direction for the Poisson's ratio provides same effects to the positive direction of the descent direction for the S-wave velocity, because a decrease in the Poisson's ratio causes an increase in the S-wave velocity as shown in eq. (3-39). On the other hand, as shown in Figure 3-14b, the descent direction for the rigidity also provides highly resolved image as that for the P-wave velocity in the IPG-III, because the full virtual source for the rigidity also has the double coupled basis virtual sources in addition to the normal stress-induced basis virtual sources (HB_{xxh} and VB_{zzv}) as shown in eq. (3-42).

Finally, Figures 3-13c, 3-14c and 3-15c show the descent directions of the density obtained using the IPG-I, IPG-II and IPG-III, respectively. We can notice that the descent directions for the density obtained using the IPG-II and IPG-III are also the same because they are also obtained by the same basis virtual sources as shown in eqs. (3-43) and (3-44) and both descent directions provide highly-resolved density image due to the contribution of the all kinds of basis virtual sources but, as Appendix B shows, the virtual source for density in the IPG-I and IPG-II acts like Mie scattering due to the time derivatives. On the other hand, the descent direction for the density obtained using the IPG-I is not sensitive at the boundary of the anomalous body because its basis virtual sources (HB_{tth} and VB_{ttv}) generate less sensitive Jacobian as mentioned before (Figure 3-10).

From these results for the CTS model, we can notice that the characteristics of the gradient direction for the various parameters are determined by included basis virtual sources, at least, in the first iteration. In addition, the highly resolved image for the gradient direction for the P-wave velocity can be achieved by adopting Poisson's ratio because shear stress-induced basis virtual sources additionally generates P-S, S-P and S-S scattered waves. As a result, in the IPG-III, the gradient directions for all the parameters can be improved than those for the conventional parameter groups because improvements of the P-wave velocity provide positive influences on FWIs for other parameters although there are some trade-off relations between S-wave velocity and Poisson's ratio.

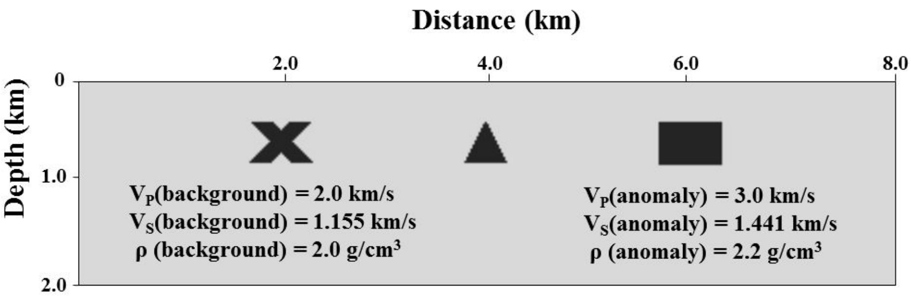


Figure 3-12 The CTS model for resolution analysis

Table 3-2 Inversion parameters used for the CTS model

Dimension	No. of shots	No. of receivers	Interval of shots	Interval of receivers	Recording time	Frequency
8.0 km × 2.0 km	200	400	0.04 km	0.02 km	3 sec	4 Hz

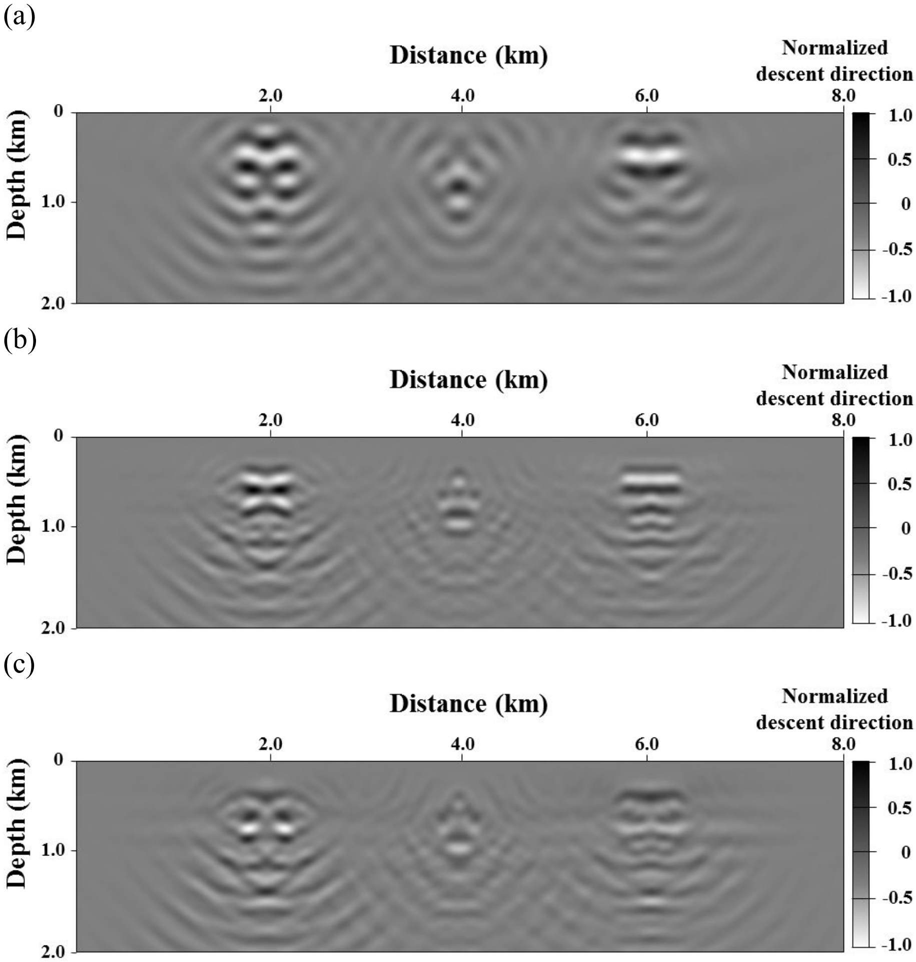


Figure 3-13 The normalized descent directions for the CTS model obtained at 1st iteration using the *IPG-I* : (a) λ , (b) μ and (c) ρ

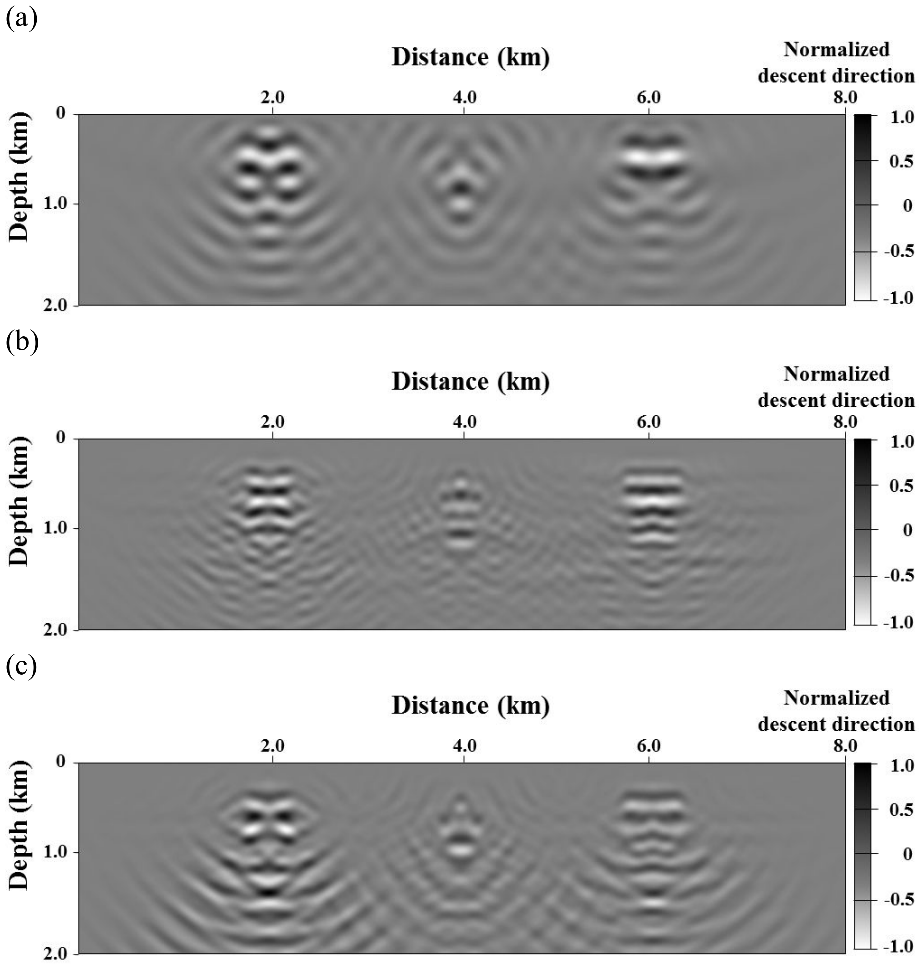


Figure 3-14 The normalized descent directions for the CTS model obtained at 1st iteration using the *IPG-II*: (a) v_p, (b) v_s and (c) ρ

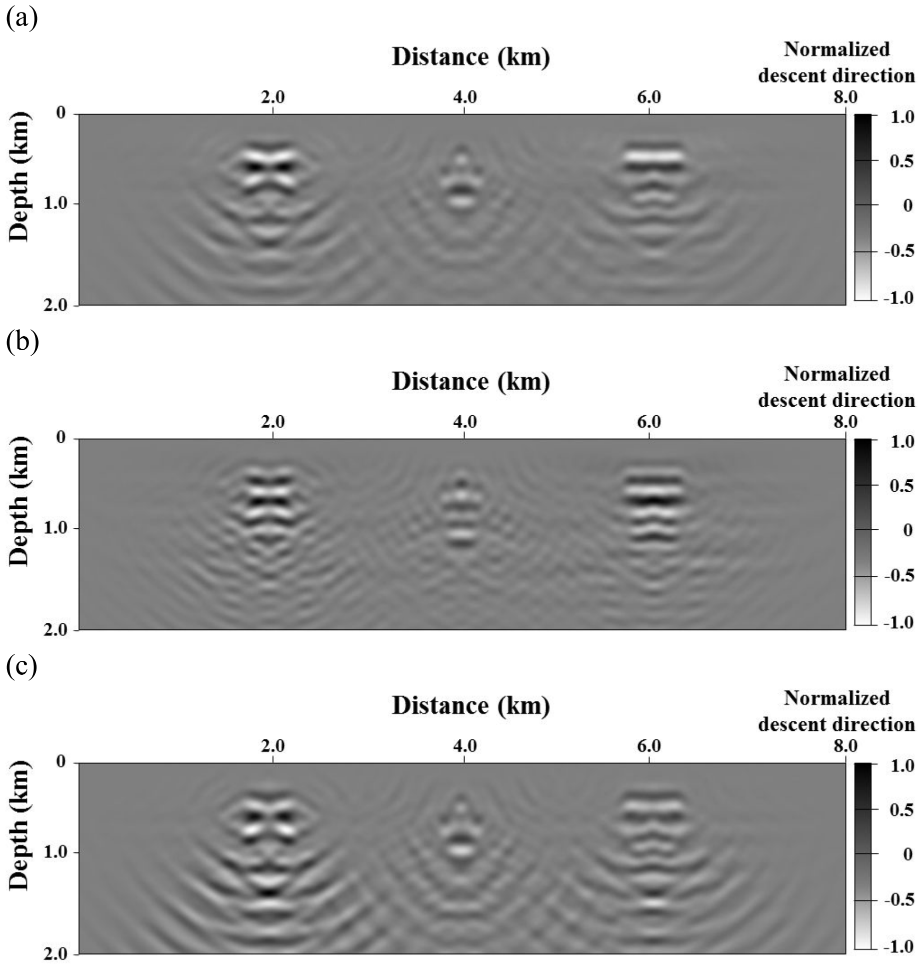


Figure 3-15 The normalized descent directions for the CTS model obtained at 1st iteration using the *IPG-III* : (a) v_p , (b) ν and (c) ρ

3.3.4 Numerical Examples for elastic Marmousi-2 Model

In this chapter, I demonstrate the superiority of the parameterization using the Poisson's ratio for more complex geologic structures by applying three parameterizations to the elastic Marmousi-2 model. Figure 3-16 shows the four parameter structures of the elastic Marmousi-2 model. Because the shallow structures of the original elastic Marmousi-2 model have very large Poisson's ratio, we need to modify only the Poisson's ratio of shallower part by multiplying some depth-variable factors (from 0.6 to 1.0) to satisfy the minimum grid numbers per wavelength for the finite-element method. Because, for this reason, the modified Poisson's ratio is not dramatically changed in this numerical example, the sensitivity of the parameter groups to the Poisson's ratio will not cause severe problems at each parameter group. The inversion parameters are listed on Table 3-3 and the new pseudo-Hessian matrix (Choi et al. 2008) is applied to pre-condition the gradient direction. The initial guesses for the P-wave velocity linearly increase from 1.5 at the top to 4.56 km/s at the bottom of the model. The initial guesses for the S-wave velocity and the density are estimated using the fixed Poisson's ratio (0.25) and Gardner's equation (Gardner et al., 1974), respectively as shown in Figure 3-17. Because of these poorly estimated initial guesses, I assume that very low frequency components are available. Although this assumption is unrealistic, the scaling method we use in this paper, in which each mono-frequency descent direction is normalized by its maximum value after the Hessian matrix is applied inside the frequency loop, provides high-frequency dominant descent direction (refer Chapter 4.1.2 or Oh and Min [2013a]). Therefore, the influence of these very low frequency data is not dominant and the high-frequency characteristics of each parameterization are well distinguished.

Figures 3-18, 3-19 and 3-20 show the recovered structures obtained using the IPG-I, IPG-II and IPG-III, respectively. We notice that, in this case, all the conventional parameterization methods provide reasonable inversion results, except that both density structures are little overestimated in the IPG-I and IPG-II. The Poisson's ratio obtained using the IPG-II is distorted because fixed step-length strategy is applied for both P-wave and S-wave velocity in the IPG-II. On the other hand, we can notice that, in the IPG-III (Figure 3-20), there are great improvements at the deeper structures in the aspects of spatial resolution. In addition, in the density structures, the salt structures that have high P-wave velocity and low density are well resolved in the IPG-III due to the improvements of the gradient direction for the P-wave velocity.

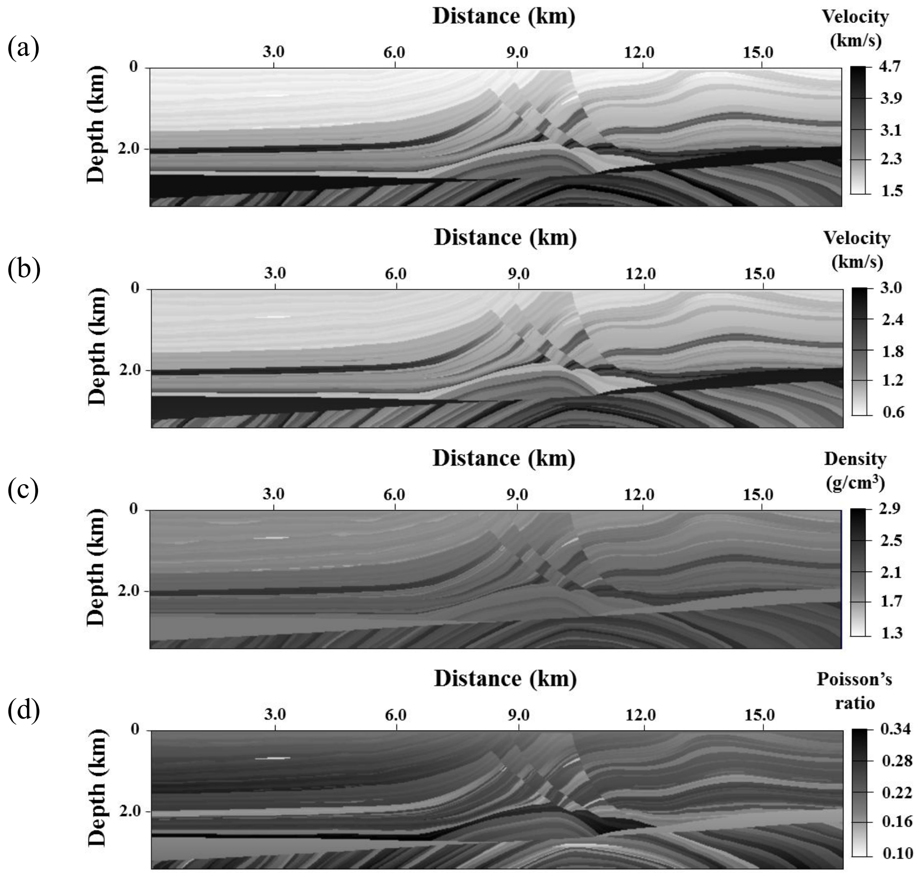


Figure 3-16 Modified elastic Marmousi-2 model: (a) P-wave velocity, (b) S-wave velocity, (c) density and (d) Poisson's ratio

Table 3-3 Inversion parameters used for elastic Marmousi-2 model

Dimension	No. of shots	No. of receivers	Interval of shots	Interval of receivers	Recording time	Maximum Frequency	Minimum Frequency
17.0 km × 3.04 km	212	850	0.08 km	0.02 km	6 sec	10 Hz	1/6 Hz

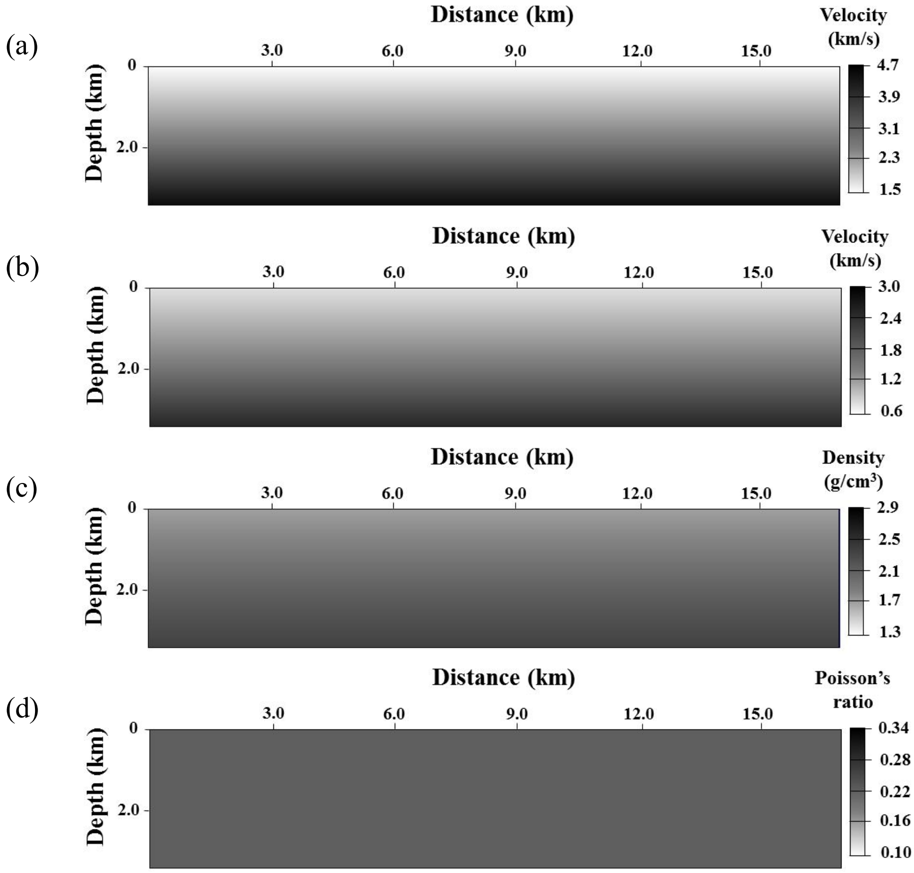


Figure 3-17 Initial models: (a) P-wave velocity, (b) S-wave velocity, (c) density and (d) Poisson's ratio

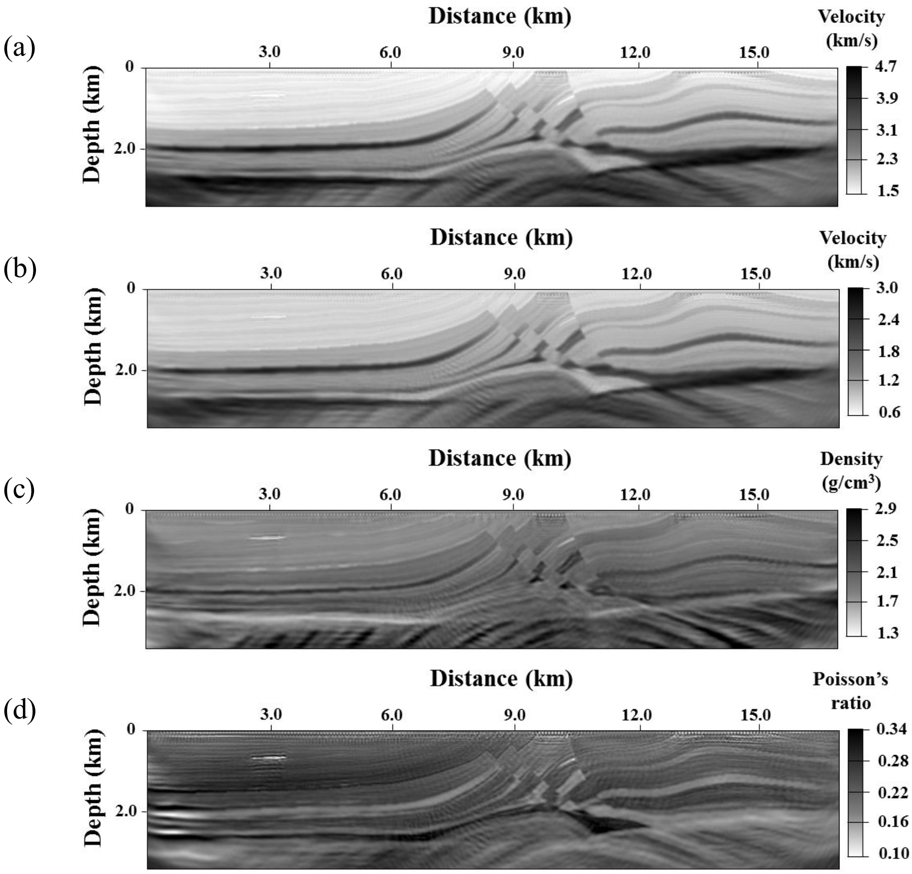


Figure 3-18 Final inverted models obtained by the *IPG-I*: (a) P-wave velocity, (b) S-wave velocity, (c) density and (d) Poisson's ratio

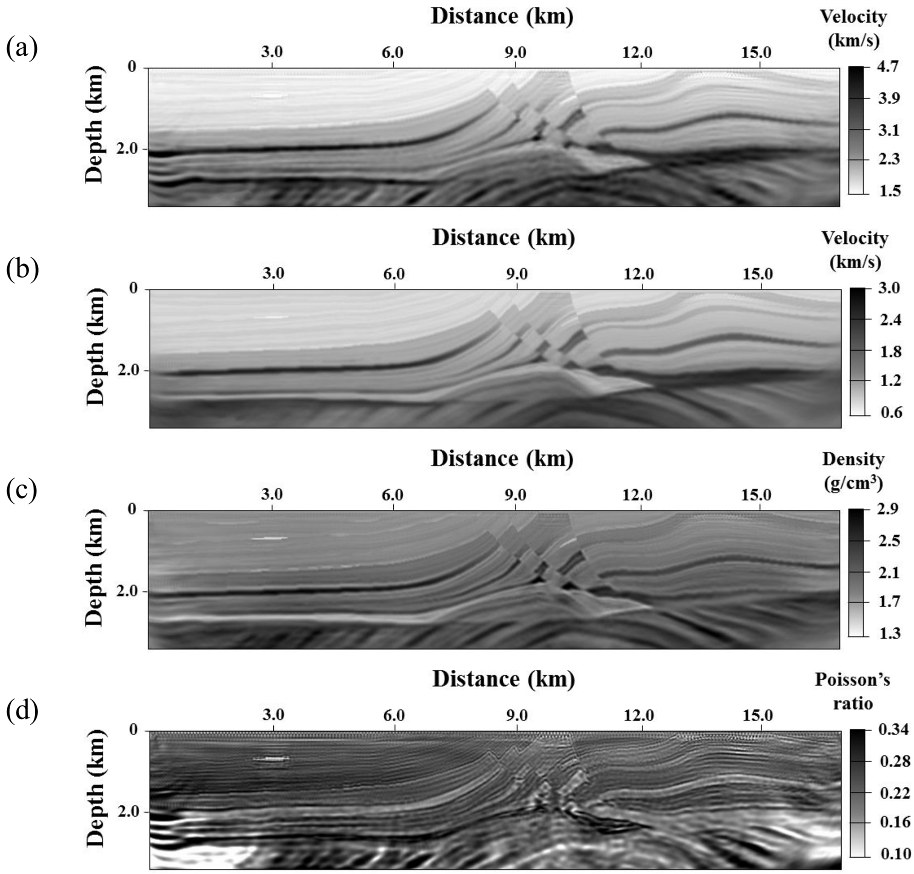


Figure 3-19 Final inverted models obtained by the *IPG-II*: (a) P-wave velocity, (b) S-wave velocity, (c) density and (d) Poisson's ratio

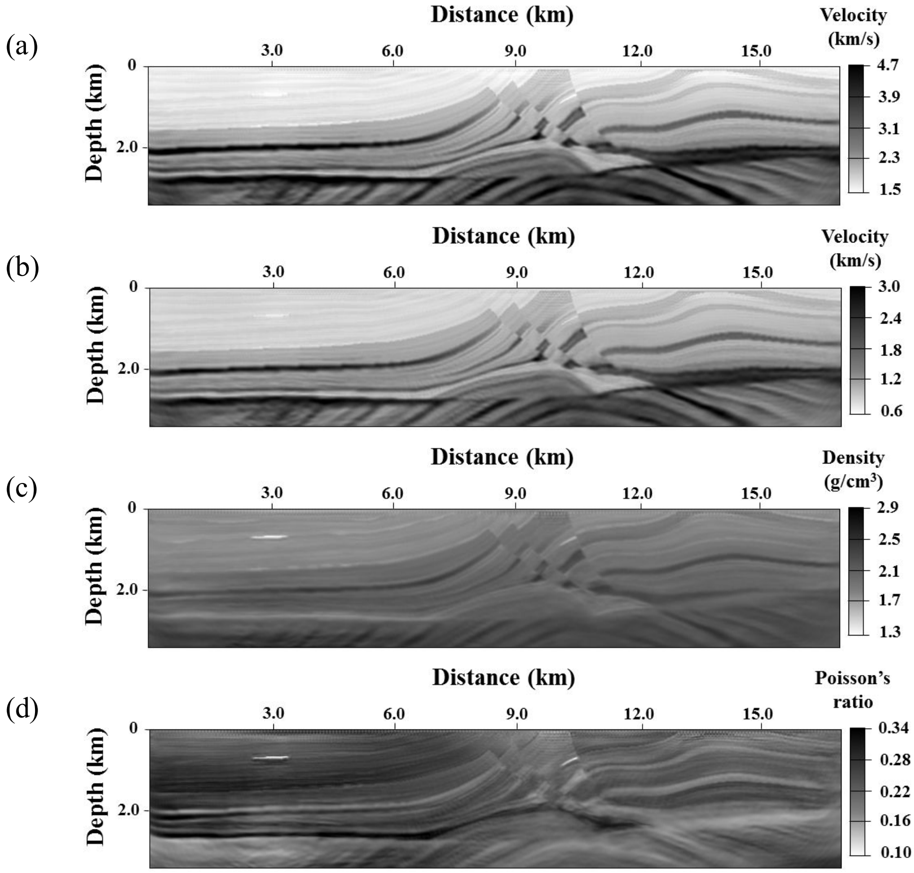


Figure 3-20 Final inverted models obtained by the *IPG-III*: (a) P-wave velocity, (b) S-wave velocity, (c) density and (d) Poisson's ratio

Chapter 4. Spectral Weighting Scheme

One of the main problems in waveform inversion is that obtaining the global minimum solutions is not easy when the initial guesses deviate significantly from the true models. Several studies have proposed methods to overcome this problem. The frequency marching method (Sirgue and Pratt 2004; Kim et al. 2011) is one such method. In the frequency marching method, waveform inversion is performed by starting with low frequencies and moving to higher frequencies and is based on the fact that the objective function is not likely to have local minima at low frequencies (Bunks et al. 1995). However, in my experience, the frequency marching method requires lots of human interventions and accurate criterions to finish each step. In addition, when we use frequency marching method, it takes more time than simultaneous approach under the assumption that the parallel computing is available.

As an alternative to the frequency marching method, Lazaratos et al. (2011) suggested the spectral shaping method, which makes the gradient have a similar spectrum to that of earth's subsurface using information of the impedance recorded in local wells. Although the spectral shaping method could be more accurate, it is too expensive because the spectral shaping function demands for several local wells.

In this Chapter, as an alternative to the frequency marching method and the spectral shaping method, I introduce the weighting method, in which the spectral weighting function can be automatically determined using the deconvolved backpropagated wavefields depending on the true and initial models without any prior information or human intervention (Oh and Min 2013a).

To design a waveform inversion algorithm that gives reliable inversion results regardless of the models, I first analyze, in the aspects of the scaling of gradient direction with Hessian matrix, why conventional waveform inversion algorithms are likely to converge to local minima. In general, waveform inversion is performed by constructing the objective function based on the residuals between the modeled and the field data and by computing the gradient to determine the model parameter update direction. When we apply frequency-domain waveform inversion to banded data, each single-frequency gradient is computed using the Jacobian and the residual and is then summed over the frequencies. Because the residual spectrum originate from the differences between the true and the assumed models (i.e., an initial or inverted model at each iteration step), the final gradient direction should appropriately describe the differences between the true and the assumed models. In general each gradient computed at each frequency contributes to recovering structures with different wavelength structures. In other words, for long-wavelength structures, the gradients obtained at low frequencies will be more important than the gradients at high frequencies and vice versa. Therefore, gradients computed over frequencies should most likely be properly weighted depending on the thickness-differences between the true and the assumed models to obtain reasonable inversion results close to the global minimum.

In this Chapter, by analyzing the characteristic of gradients over a range of frequencies, I address the limitations of some conventional elastic FWI algorithms and then propose a weighting method to overcome the limitations of these approaches. One problem associated with conventional waveform inversion is that the particular frequency components emphasized in the banded gradient direction are not dependent on the given models. In one case, high-frequency components are excessively weighted, and in another case, dominant frequencies of the source wavelets are always emphasized. These limitations of the conventional inversion algorithms will be discussed in following subchapter. This Chapter is edited version of Oh and Min (2013a).

4.1 Limitations of Two Conventional Scaling methods

Before introducing the weighting method, I analyze limitations of two conventional scaling methods in the aspect of the spatial resolution of the gradient direction. The first conventional scaling method is based on the general inverse theory for banded data as expressed by eq. (2-11), which is called 'CS-I method (Conventional Scaling method-I)' through the paper. The other conventional scaling method, which is called 'CS-II method (Conventional Scaling method-II; Ha et al. 2009), is defined by eq. (2-12). For easy analysis, I use two simplified geologic models shown in Figure 4-1. The thick rectangular-shaped model in Figure 4-1a is modeled on a salt structure, which has a low-velocity zone beneath the high-velocity salt body. The second model is modeled on interbedded structures consisting of three high-velocity layers (Figure 4-1b). P- and S-wave velocities and densities are shown in Figure 4-1, and the parameters used for the inversion are listed in Table 4-1. The S-wave velocities are generated so that the Poisson's ratio is constant at 0.25. The density is fixed at 2.0 g/cm^3 and is not updated during the FWI. For the initial guesses, we assume homogeneous models whose parameters are the same as those of the background media so that the only discrepancies between the true and initial models can be induced by anomalous bodies.

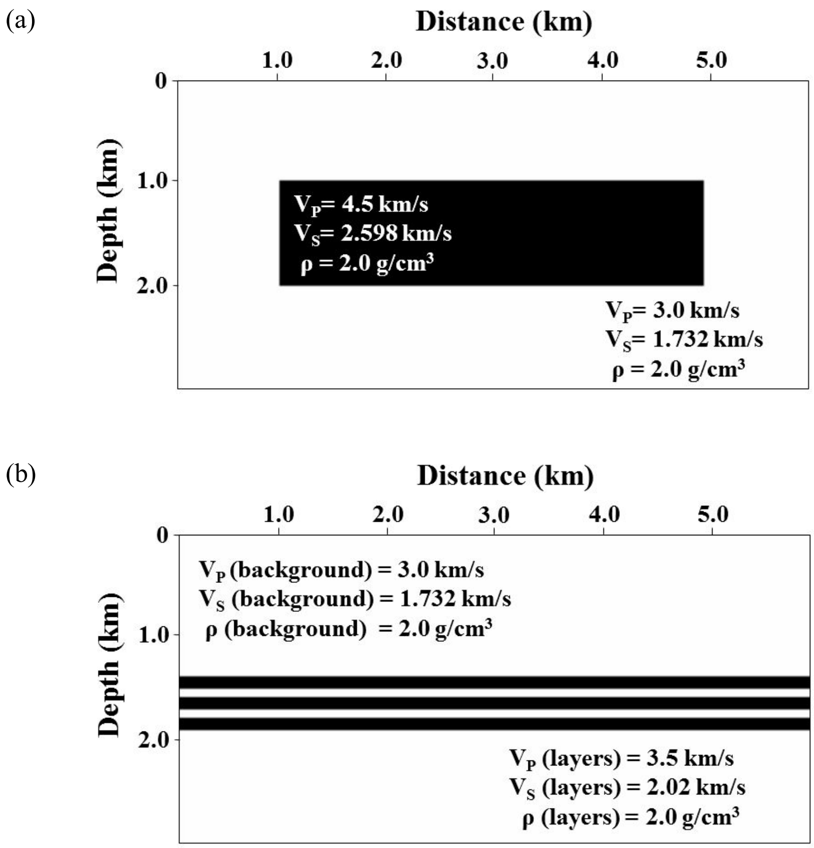


Figure 4-1 Simple models: (a) thick rectangular-shaped model and (b) thin-layers model.

Table 4-1 Inversion parameters used for the simple models for resolution analysis

Dimension	No. of shots	No. of receivers	Interval of shots	Interval of receivers	Recording time	Maximum Frequency	Minimum Frequency
6.0 km × 3.0 km	150	301	0.04 km	0.02 km	4 sec	10 Hz	0.25 Hz

4.1.1 Limitation of Conventional Scaling Method I

The limitation of CS-I method has been well reported by some previous works. As Jang et al. (2009) insisted, in the CS-I method, each mono-frequency gradient should be weighed by the source spectrum. For example, if the amplitude spectrum of the source wavelet is dominant at low- or high-frequency band, we obtain thick or thin gradient directions, respectively. On the other hand, the CS-II method, in which each mono-frequency gradient is pre-conditioned by the mono-frequency Hessian matrix inside the frequency loop, does not suffer from the source spectrum due to the cancellation of the source wavelet. To confirm this limitation of the CS-I method, we perform the elastic FWI for the thick rectangular-shaped model (Figure 4-1a). As Figure 4-2a shows, we suppose two types of source wavelets. The first source wavelet is the first derivative of Gaussian function, in which most energy is dominant at a quarter of maximum frequency as Figure 4-2b shows. The other source wavelet is the Ricker function, which maximum frequency is around one half of the maximum frequency (Figure 4-2b). As we mentioned before, because the source spectrum acts as a spectral weighting function in the CS-I method, the descent direction of the P-wave velocity obtained using the Ricker function is thinner than the descent direction obtained using the first derivative of Gaussian function (Figures 4-3a and 4-3b). On the other hand, in the CS-II method, the source spectrum does not influence on the descent directions as Figures 4-3c and 4-3d show.

(a)

(b)

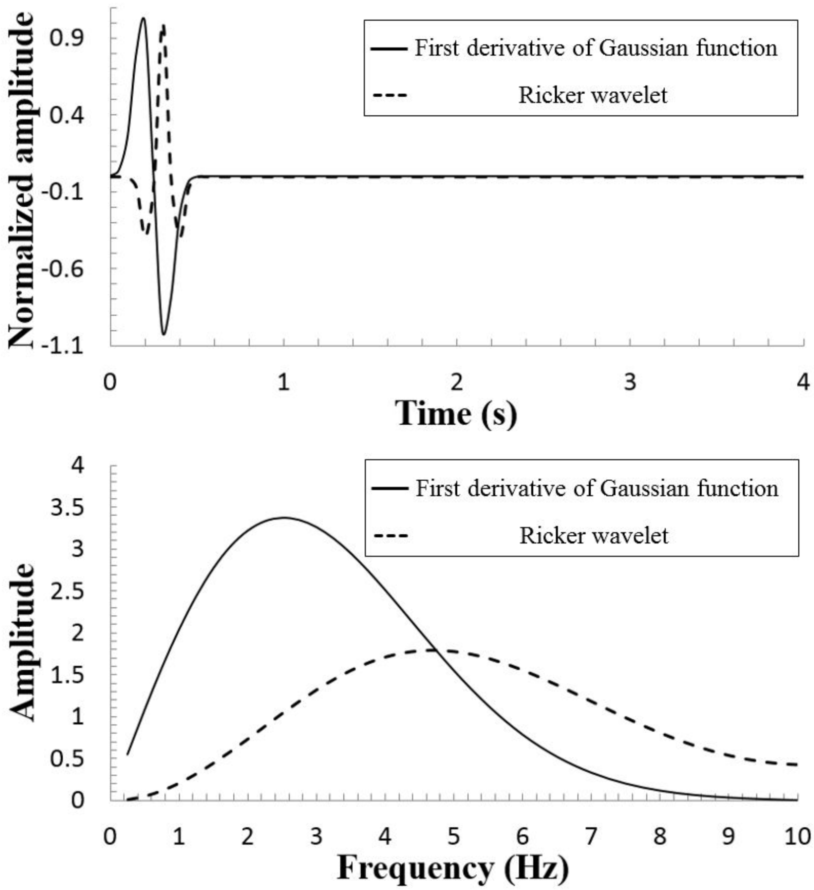


Figure 4-2 (a) Source wavelets of the first derivative of Gaussian function and Ricker function and (b) their amplitude spectra

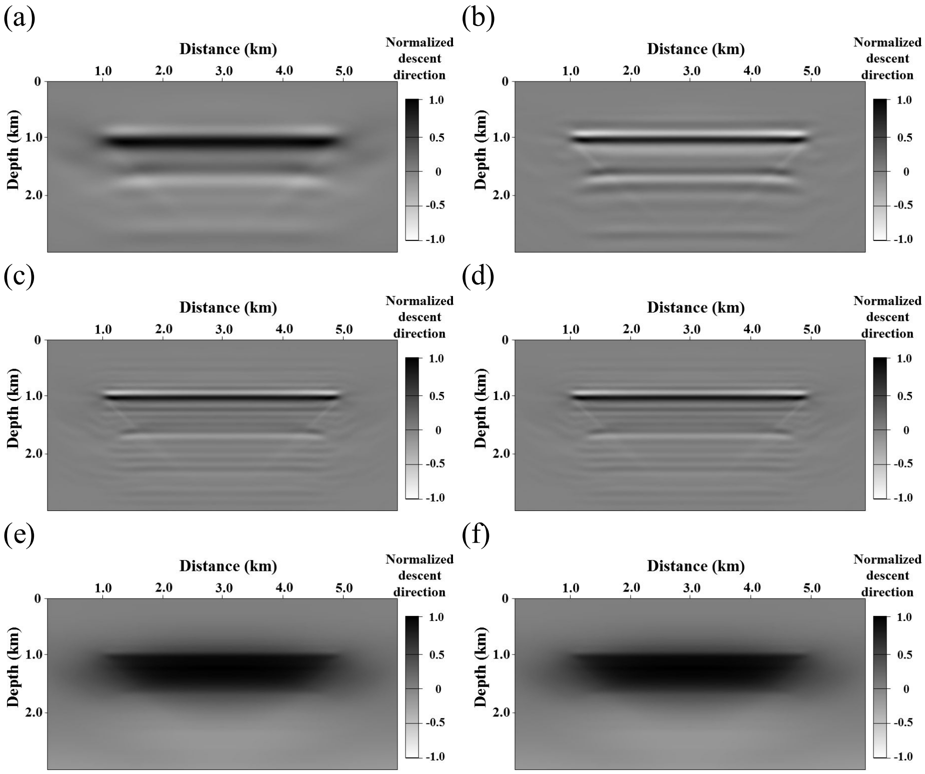


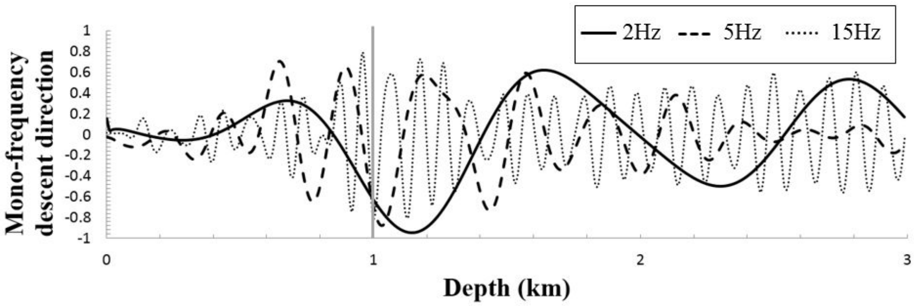
Figure 4-3 Descent directions of P-wave velocity for the thick rectangular-shaped model at the first iteration obtained using the CS-I (a and b), CS-II (c and d) and weighting methods (e and f) when the first derivative of Gaussian function (a, c and e) and the Ricker function (b, d and f) are used as a source wavelet

4.1.2 Limitation of Conventional Scaling Method II

Although the CS-II method provides source independent FWI under the assumption that the exact source wavelet can be estimated, the CS-II method also has a limitation related to controlling the spatial resolution of the gradient direction. To investigate the limitation of the CS-II method, we also perform the elastic FWI for the thick rectangular-shaped model.

Figure 4-4a shows the depth profiles of the single-frequency descent directions obtained by the CS-II method. As we can see, because each single-frequency descent directions are aligned at the top of the anomalous body, the banded descent direction, which is obtained by summing all the single-frequency descent directions after the normalization, has high-frequency dominant tendency at the boundary of the anomalous body as Figure 4-4b shows. Figures 4-5a and 4-5b show the banded descent direction for the P-wave velocity obtained by the CS-II method when the Nyquist frequency is 10 and 20 Hz, respectively. When the time sampling interval is smaller, we obtain a high-frequency descent direction, which is not appropriate to recover the long-wavelength structures even though we use the extremely low-frequency components (0.2 Hz) during the FWI.

(a)



(b)

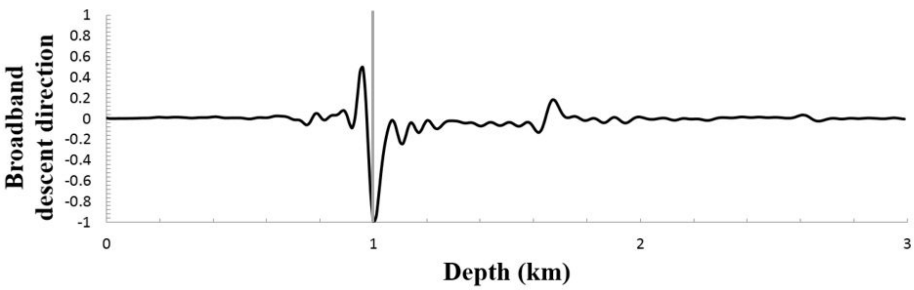


Figure 4-4 (a) Mono-frequency descent directions and (b) total descent direction summed from 0.25 to 20 Hz: Grey vertical line indicates the top of the anomalous body

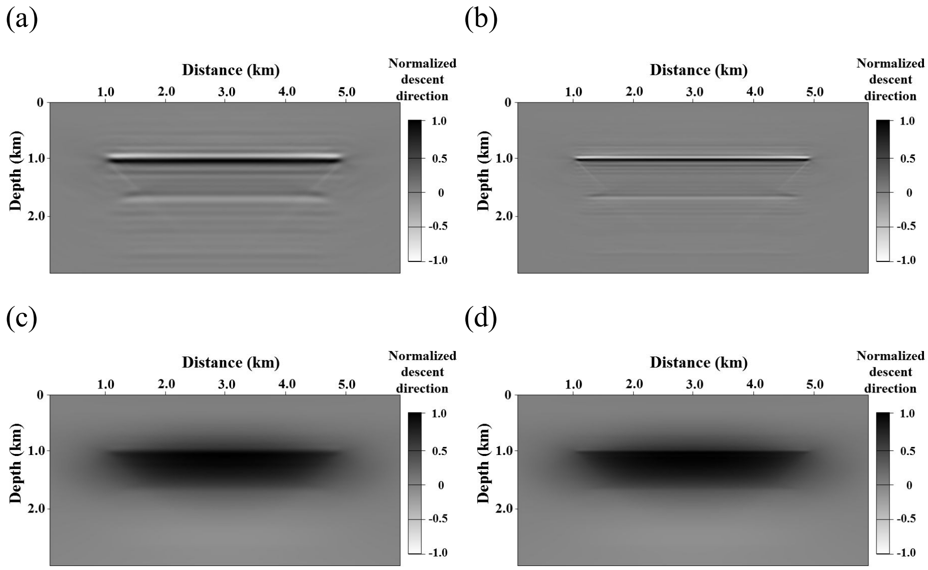


Figure 4-5 Descent directions of P-wave velocity for the thick rectangular-shaped model at the first iteration obtained using the CS-II (a and b) and weighting method (c and d) when the Nyquist frequency is 10 (a and c) and 20 Hz (b and d)

4.2 Analysis of the Jacobian and the residuals

In advance of analyzing the limitations of the conventional scaling methods, investigating the general characteristics of the gradient direction expressed by the Jacobian matrix and the residual helps us to describe the problems of the two conventional scaling methods and to examine the meaning and the importance of the weighting method in frequency-domain elastic FWI.

Figure 4-6 shows the real part of the vertical components of the transpose of Jacobian matrices obtained using the reciprocity theorem (Shin et al. 2001a) for the P-wave velocity at 1, 5 and 9 Hz when the source is located in the middle of the surface. The transpose of the Jacobian matrices in the frequency domain resembles a snapshot of the monochromatic waves. The Jacobian matrices at each frequency have aspects of fluctuation, and the widths of these fluctuations are related to dominant wavelengths or frequencies (see Figure 3 in Virieux and Operto [2009] for more details). Considering that each single-frequency Jacobian matrix has a different resolution associated with the width of the fluctuation, we can guess that, if the initial model is homogeneous, the low- and high-frequency Jacobian matrix would be good for recovering the thick rectangular-shaped and thin-layers models, respectively.

We can see in eq. (2-6) that the gradient is obtained by the cross-correlation between the single-frequency Jacobian matrix and the residual vector, which may mean that the contribution of each single-frequency Jacobian matrix to the banded gradient direction is controlled by the residual vector (strictly, the deconvolved residual vector).

Figure 4-7 schematically shows the propagation of waves reflected at subsurface layers that have different thicknesses. When the seismic waves encounter a layer in the subsurface, the low-frequency signals, whose wavelength is approximately larger than eight times the thickness of the layer in my examples (for more details, please see Widess [1973] and Kallweit and Wood [1982]), are transmitted through the layer without reflection, as shown in Figure 4-7b. Figure 4-8 shows the phases of monochromatic waves, which propagate through the homogeneous, thick rectangular-shaped and thin-layers models. Figure 4-8 shows that both the low- and high-frequency signals are reflected at the top of the thick layer (Figures 4-8c and 4-8d), while the low-frequency signals are transmitted through the thin layers and only the high-frequency signals are reflected (Figures 4-8e and 4-8f). For this reason, the low-frequency components of the residual wavefields are primarily associated with long wavelengths when the initial guesses are poorly estimated. Considering the feature of seismic attenuation, the low-frequency components of the residual wavefields may also include information about the deep structures. Conversely, the high-frequency components of the residual wavefields are related to the short-wavelength and the shallow structures. From these phenomena, we may say that the deconvolved residual vectors, to some extent, reflect the differences between the true and the inverted models during the inversion process.

Figure 4-9 shows the RMS errors calculated by the deconvolved residuals with a frequency at the first iteration of inversion for the two models. For the thick rectangular-shaped model, the RMS errors are dominant at the low-frequency band below 2 Hz, whereas for the thin-layers model, the RMS errors at high frequencies are relatively greater than the errors at low frequencies because the thin layers cannot be detected by low-frequency signals. However, an overall feature of the RMS errors for the thin-layers model is that they are relatively unchanged compared with those for the thick rectangular-shaped model. From these results, we can guess that if we apply the CS-II method to the thick rectangular-shaped model, the normalization will degrade the inversion results by relatively attenuating the effect of the important low-frequency components and by relatively emphasizing the effect of the high-frequency components. For the thin-layers model, the CS-II method will provide good results because the residuals are relatively uniform over frequencies, and the effect of normalization is not as great as for the thick rectangular-shaped model. On the other hand, if we use the CS-I method with the source-wavelet-convolved residuals, the CS-I method may not resolve the thin layers properly depending on the major frequency of the source wavelet as discussed in Chapter 4.1.

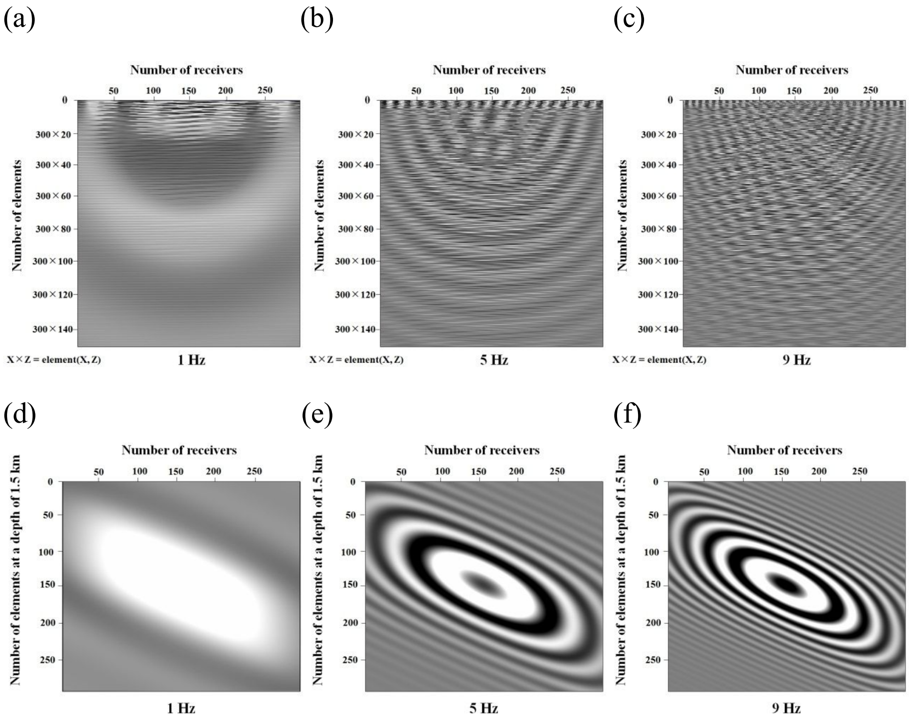
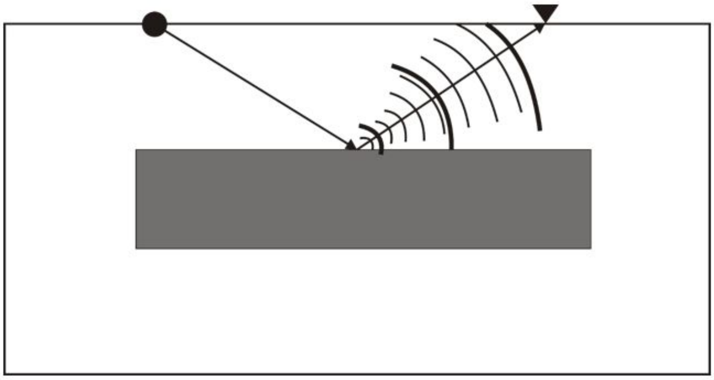


Figure 4-6 Vertical components of the transpose of the Jacobian matrices (a, b and c) and the partial derivative wavefields (d, e and f) extracted at a depth of 1.5 km for the P-wave velocity at 1 Hz (a and d), 5 Hz (b and e) and 9 Hz (c and f).

(a)



(b)

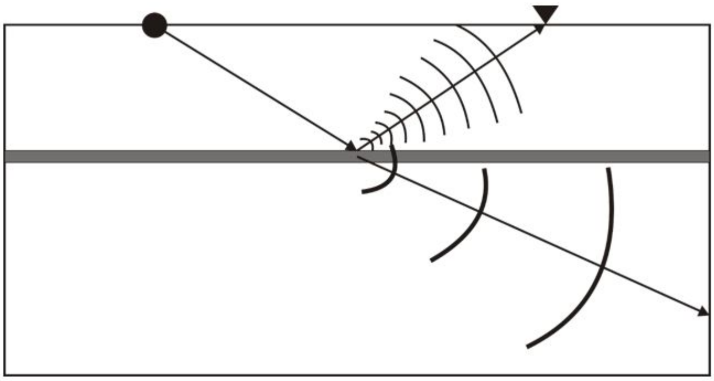


Figure 4-7 Schematic diagrams illustrating the propagation of waves reflected at the top of (a) thick and (b) thin layers. The black circle and inverted triangle indicate the seismic source and the receiver, respectively. The thick and thin arcs denote wavefronts of the low- and high-frequency wavefields, respectively.

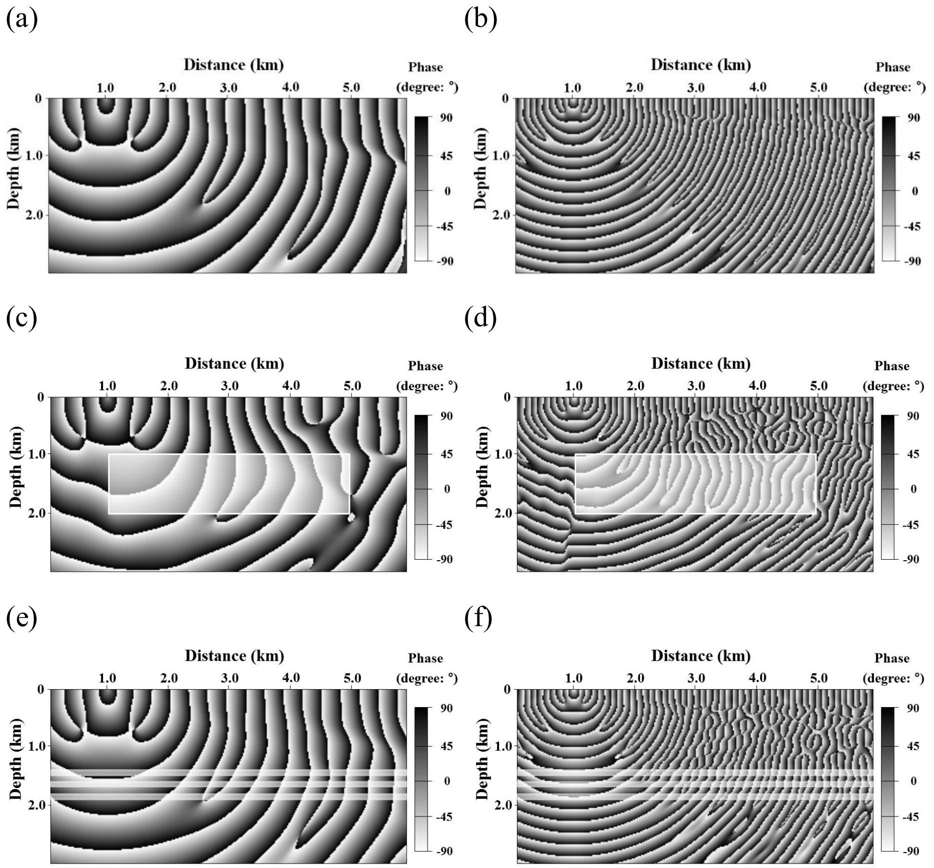


Figure 4-8 The phase of the monochromatic wavefields for the vertical components of the displacements obtained for the homogeneous initial (a and b), thick rectangular-shaped (c and d) and thin-layers models (e and f) at 2 (a, c and e) and 8 Hz (b, d and f). The white rectangles indicate the anomalous body.

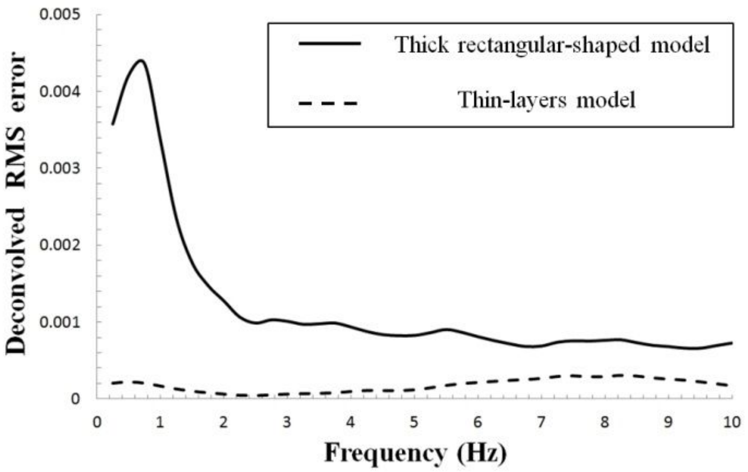


Figure 4-9 Deconvolved RMS errors with a frequency at the first iteration of inversion for the thick rectangular-shaped (solid line) and the thin-layers models (dashed line).

4.3 Weighting Factors using Deconvolved Wavefields

To overcome the limitations of the conventional scaling methods, I propose the weighting method, which can provide a geological model-induced spectral weighting effect for each single-frequency descent direction. The purpose of my weighting scheme is similar to previous works focused on gradient-shaping methods (Lancaster and Whitcombe 2000; Lazaratos and David 2009; Lazaratos et al. 2011). Lazaratos et al. (2011) addressed that the model generated by inversion should have a frequency spectrum of the earth's subsurface to assure good convergence and that this target spectrum can be derived by averaging the spectra of log curves recorded in local wells. Conversely, the weights in my weighting method are determined by spectral differences between the observed and the modeled data recorded during a seismic survey based on the analysis of the spectral wave propagation (Chapter 4.2) without using information recorded in local wells. In this subchapter, I discuss about the physical meaning of my weighting factors and how the weighting factors can determine the spatial resolution of gradient direction.

4.3.1 Frequency-domain FWI with Weighting Factors

The weighting method is designed based on the CS-II method so that the spectral weighting effects of the Hessian matrix (refer Appendix A in Oh and Min 2013a) and the source spectrum (Chapter 4.1.1) can be removed. If we apply the weighting factor to each single-frequency descent direction, the total descent direction can be written as

$$\delta \mathbf{p}_{\text{weighted}} = -\sum_{\omega} \gamma_{\omega} \left\{ \text{NRM}\left[\delta \mathbf{p}_{\omega} \right] \right\}, \tag{4-1}$$

where γ_{ω} is a weighting factor. Comparing eq. (4-1) with eqs. (2-9) and (2-10), we can notice that the objective function of the weighting method can be obtained by defining the scaling factor as the maximum absolute of the descent vector over the weighting factor. To give the model-induced spectral weighting effect, we can use two types of weighting factors as follows

$$\gamma_{\omega}^{\text{Type-I}} = \frac{\sum_{i=1}^{nr} \left| \tilde{r}_i \right|}{nr}, \tag{4-2}$$

and

$$\gamma_{\omega}^{\text{Type-II}} = \frac{\sum_{j=1}^{np} \left| \tilde{v}_j \right|}{np}, \tag{4-3}$$

where r_{i} and v_{j} indicate the residuals measures at the i^{th} receiver and the backpropagated wavefield recorded at the j^{th} nodal point, respectively. The tilde indicates that these measures are deconvolved, and the terms *nr* and *np* denote the number of receivers and nodal points, respectively. Each weighting factor is designed as an average of complex absolute values so that the weighting factors can quantify the spectral differences of the recorded signals between the observed and modeled data. The Type-I weighting factor has an effect of directly measuring the residuals of the reflected wavefields between observed and modeled data considering only forward wave propagation from sources to receivers. In contrast, the Type-II weighting factor approximately measures misfits in the square of the amplitudes of the reflected wavefields between the observed and the modeled data considering both forward and backward wave propagation. In other words, the Type-II weighting factor is proportional to the elastic energy at each frequency. The Type-II weighting factor calculated by the backpropagated wavefields plays a role in amplifying the weighting effect of the Type-I weighting factor and making the weighting effect more robust to the spectrum of noise. In this study, we use the Type-2 weighting factor.

In general, the backpropagated wavefields are generated by simultaneously backpropagating the residuals obtained for each shot gather. To avoid the complexity caused by combining numerous data for all the shot gathers, I only use the average of wavefields recorded at every nodal point by backpropagating the residuals measured when the source is located in the middle of the surface. Because we just use the residual and the backpropagred wavefields that were already calculated for the gradient, the weighting method does not increase computational costs a lot.

4.3.2 Physical Meanings of Weighting Factors

I discuss the physical meanings of two weighting factors. The Type-I weighting factor, as shown in eq. (4-2), is directly calculated using the deconvolved residuals; The Type-II weighting factor, as shown in eq. (4-3), is the backpropagated version of the Type-I weighting factor.

The residual vector can be written as

$$\mathbf{r} = \left[\mathbf{S}_{\text{true}}^{-1}\right]\mathbf{f}_{\text{true}} - \left[\mathbf{S}_{\text{estimated}}^{-1}\right]\mathbf{f}_{\text{estimated}} . \qquad (4\text{-}4)$$

Because each component of $\mathbf{S}^{-1}$ corresponds to Green's function, i.e., impulse response, we can say that the Type-1 weighting factor (eq. (4-2)) measures the spectral difference of the residuals between the observed and the modeled data. However, the physical meaning of Green's function is different depending on the data acquisition system of the observed data. Because, in eq. (4-4), the residual spectrum can be affected by the spectrum of Green's function, which physical meaning is determined by the physical properties of the source and recorded data. According to Berkhout (1985), the physical properties of Green's function can be divided by two kinds depending on the data acquisition system; reflectivity and admittance. In eq. (4-5), if the input (**I**) and output (**O**) have the same physical quantities, we call the operator (**M**) reflectivity.

$$\mathbf{MI} = \mathbf{O} \qquad (4\text{-}5)$$

Because this is the case of the common data acquisition system in marine survey (pressure to pressure), we call inverse of the modeling operator (i.e., Green's function) for the acoustic wave equation as a reflectivity matrix. However, the reflectivity tends to proportional to the angular frequency (for more details, refer to Chapter 6 in Lines and Newrick [2004]). These facts mean that, if we apply the Type-I weighting factor to marine data, the weighting effect might be dominant at high-frequency bands due to the spectrum of Green's function (i.e., reflectivity). For this reason, if someone want to apply the weighting method to marine data, some additional works are required to convert Green's function.

On the other hand, in eq. (4-5), if the input (**I**) is velocity fields and output (**O**) is pressure fields, we call the operator (**M**) impedance. In physics, the impedance is defined by the pressure field over the velocity field, which is the same physical meaning of the multiplication between the density and velocity. In seismology, the impedance is a measurement of how the wave can be easily impeded by subsurface layers. Because the high-frequency components of wave are easily impeded by subsurface layer (see Figures 4-7 and 4-8) and the pressure field is proportional to the time derivative of velocity (It means the multiplication of $i\omega$ in the frequency domain), we can guess that the high-frequency components of wave have larger impedance than the low-frequency components. On the other hand, the admittance, which is defined by the inverse of the impedance (velocity over pressure), is a measurement of how the wave can easily propagate through the subsurface layers and is inversely proportional to the angular-frequency for the same reason of the impedance.

As Berkhout (1985) explained, when we record the particle velocity fields induced by a pressure or body force source (i.e., the common data acquisition for land survey), the physical meaning of Green's function is the admittance, which is inversely proportional to the angular-frequency. For this reason, velocity components of Green's function induced by Pressure source (or body force source) plays a role in assigning large weights to the low-frequency bands, which gives similar effects to some smoothing techniques (Fichtner et al. 2009) and the multi-scale approach (Bunks et al. 1995, Kim et al. 2011). When we record the particle displacement as we did in this paper, this smoothing effect of Green's function can be amplified in the frequency domain due to the relationship of the time derivative of displacements to velocity. Because I use the CS-II method, in which the banded gradient direction has high-frequency characteristics as I discussed in Chapter 4.1, as an initial condition for the weighting, this smoothing effects will compensate the high-frequency dominance of the CS-II method. However, in my experience, the weighting effect of the Type-I weighting factor is not enough to compensate the high-frequency characteristics of the CS-II method. In addition, the Type-I weighting factor is not robust to noise spectrum because it is direct measure of the observed data. For these reasons, I also suggest the Type-II weighting factor, which is the amplified version of the Type-I weighting factor.

The backpropagation of the observed and the modeled data can be expressed by

$$\begin{aligned} v^{\text{observed}} &= \left[\mathbf{S}_{\text{true}}^{-1}\right]^{T} \mathbf{d}^{*} \\ &= \left[\mathbf{S}_{\text{true}}^{-1}\right]^{T} \left[\mathbf{S}_{\text{true}}^{-1}\right]^{*} \mathbf{f}_{\text{true}} \end{aligned} \tag{4-6}$$

and

$$
\begin{aligned}
v^{\text{modeled}} &= \left[\mathbf{S}_{\text{estimated}}^{-1}\right]^{T} \mathbf{u}^{*} \\
&= \left[\mathbf{S}_{\text{estimated}}^{-1}\right]^{T} \left[\mathbf{S}_{\text{estimated}}^{-1}\right]^{*} \mathbf{f}_{\text{estimated}}
\end{aligned}
\qquad (4\text{-}7)
$$

respectively. Because $\left[\mathbf{S}^{-1}\right]^{T}\left[\mathbf{S}^{-1}\right]^{*}$ consists of the products of Green's functions, we can say that the Type-2 weighting factor (eq. (4-3)) roughly measures the energy misfit (i.e., the square of amplitude difference) between the observed and the modeled data. Figure 4-10a describes the properties of the forward and the backward propagation of the wavefields. The forward-propagated wavefields recorded at receivers are reflected at the top of the anomalous body just once (e.g., solid trace), whereas the wavefields, which are propagated along the solid trace and are recorded at receivers, are, in turn, backpropagated along the dashed trace. If it is possible to use the same modelling operator for both the forward and the backward propagation, the forward and the backward propagation of the wavefields will be along the same paths. From these phenomena, we can guess that if we use backpropagated wavefields recorded at the original source position (black-filled circle in Figure 4-10), the original weighting effect of Type-I weighting factor will be amplified.

In eqs. (4-6) and (4-7), we assumed that we can compose the complex impedance matrix for true geological structures. However, in FWI, we backpropagate the observed data using the complex impedance matrix composed for the initial (or the estimated) models in which case the backpropagation of the residual vector can be expressed by

$$
\begin{aligned}
v^{\text{residual}} &= \left[\mathbf{S}_{\text{estimated}}^{-1}\right]^{T} \left[\mathbf{d}-\mathbf{u}\right]^{*} \\
&= \left[\mathbf{S}_{\text{estimated}}^{-1}\right]^{T} \mathbf{d}^{*} - \left[\mathbf{S}_{\text{estimated}}^{-1}\right]^{T} \mathbf{u}^{*}
\end{aligned}
\tag{4-8}
$$

The first term in eq. (4-8), $\left[\mathbf{S}_{\text{estimated}}^{-1}\right]^{T} \mathbf{d}^{*}$, has uncertainty, which means that the backpropagated wavefields of the observed data propagate along different ray paths from those used for the forward propagation. Let me assume that the true geological and the initial models have a thick and a thin rectangular anomalous body, as shown in Figures 4-10a and 4-10b, respectively. For convenience, I only consider the reflected waves. While the observed data experience reflection due to the thick anomalous body, the recorded data are backpropagated based on the thin-layer model (initial guess) and their low-frequency components (dotted lines in Figure 4-10b) may not be reflected because the layer of the initial guess is too thin compared to the wavelengths. As a result, the low-frequency components of the backpropagated wavefields cannot be recorded at the receiver for the virtual source (i.e., the original source position for forward propagation; black filled circle in Figure 4-10). In contrast, the high-frequency components of the backpropagated wavefields of the observed data (dashed lines in Figure 4-10b) can be reflected and recorded. This behavior of the backpropagated residual can have a negative influence on the spectral weighting effects if we only use the backpropagated residual recorded at the original source position used for the forward propagation. To avoid this problem, we use the average value of the backpropagated wavefields recorded at all the nodal points of the given model, as shown in eq. (4-3), to consider the energy of all the backpropagated seismic waves that are transmitted or reflected. By doing so, we can roughly measure the energy misfits between the true and the modeled seismic signals.

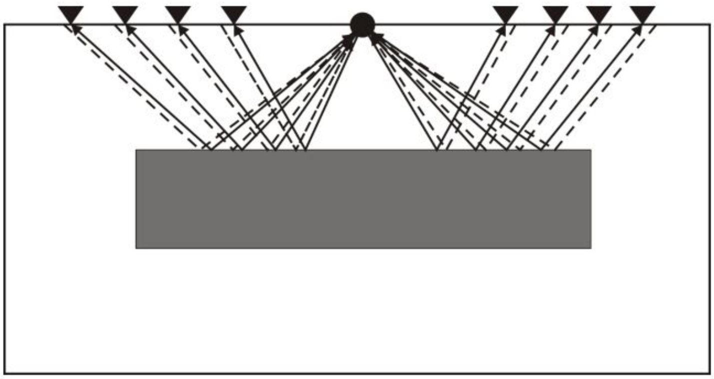
The other advantage of the Type-II weighting factor is the robustness of the weighting factor to the noise spectrum. When the observed data include some noise, eq. (4-6) can be rewritten as

$$v^{\text{observed}} = \left[\mathbf{S}_{\text{true}}^{-1}\right]^T \mathbf{d}_s^* + \left[\mathbf{S}_{\text{true}}^{-1}\right]^T \mathbf{d}_n^* \quad , \qquad (4\text{-}9)$$

Where the subscripts s and n denote the seismic signals and the noise, respectively. In contrast to the signals, because noise is only added during forward propagation to obtain the observed data (not during back propagation), the noise is relatively not amplified compared to the signals. Consequently, the back-propagation increases the spectral influence of the seismic signals on the weighting factor, which is the strength of the Type-2 weighting factor.

To investigate the sensitivity of the two weighting factors to the noise spectrum, I perform the elastic FWI for synthetic data with random noise for the thick rectangular-shaped model, whose signal-to-noise ratio gradually changes from 10 at the lowest frequency to 2 at the highest frequency. Figure 4-11 show the two weighting factors over frequencies. From the Figure 4-12, I notice that the Type-1 weighting factor, which is the direct measure of the deconvolved residual, is sensitive to the noise spectrum and does not work properly for the noise-included data. In contrast, the Type-2 weighting factor is less sensitive to noise and acts as a weighting function relatively well, as shown in Figure 4-11. For above two advantages of the Type-2 weighting factor, I only consider the Type-2 weighting factor through the paper.

(a)



(b)

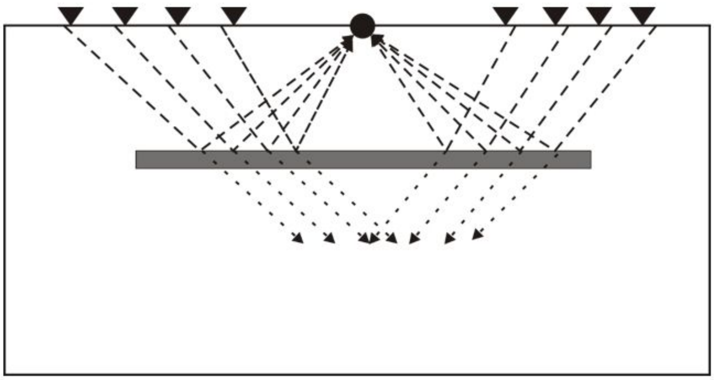


Figure 4-10 Schematic diagrams illustrating (a) the forward and backward propagations of the reflected waves by the thick rectangular-shaped anomalous body and (b) the backward propagation on the thin layer model. The black solid and dashed lines indicate the forward and backward wave paths, respectively. The dotted lines denote the backward wave paths for the low-frequency component. The black circle and inverted triangles indicate the seismic source and receivers, respectively.

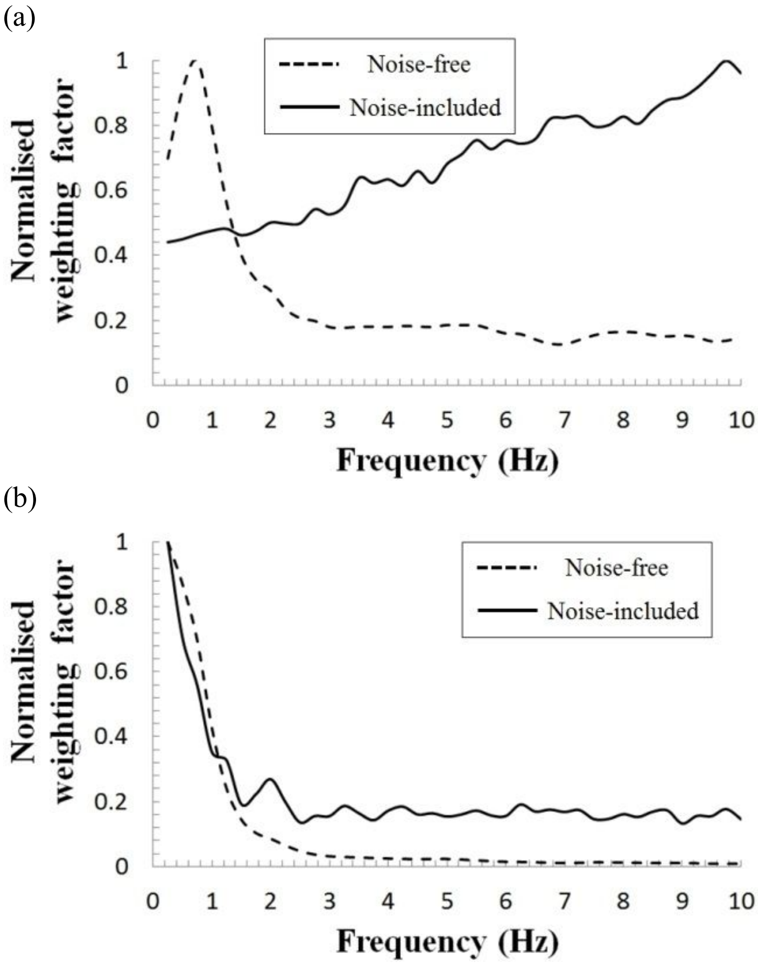


Figure 4-11 Distribution of the (a) Type-I weighting factor and (b) the Type-II weighting factor over frequencies without (dashed line) and with (solid line) random noise

4.3.3 Resolution Determination using Weighting Factor

To investigate whether the weighting method improves inversion results and compensate the limitations of two conventional scaling methods, I perform elastic FWI using the weighting method and compare the results with those of the CS-I and the CS-II methods. Figure 4-12 shows the banded descent directions of the P-wave velocity for two simplified models obtained at the first iteration using the CS-I, CS-II and weighting methods. In Figures 4-12c and 4-12d, I observe that the three thin layers are well detected in the banded descent directions obtained using the CS-II method, whereas the thick anomalous body is not detected due to the dominance of high-frequency components as we discussed in Chapter 4.2.2. In contrast, in the descent directions obtained using the CS-I method (Figures 4-12a and 4-12b), the three thin layers are not resolved well, although anomalous bodies in the rectangular shaped-model are resolved as thicker than those obtained using the CS-II method. From Figure 4-12, I note that the dominant frequency in the descent directions does not change even though the true model is different and the spectral distribution of the deconvolved residuals is different in each case (Figure 4-9). Therefore, neither method reflects the relative spectral differences between the true and the initial models properly. In other words, the single-frequency descent directions computed at each frequency are not appropriately weighted depending on the given model. Figures 4-12e and 4-12f illustrate the descent directions obtained at the first iteration using the weighting method. Unlike the CS-I and CS-II methods, both the targeted thick body and the thin layers are well reflected in the descent directions.

In summary, the two conventional scaling methods are quasi-monochromatic, which means that a certain frequency band is always emphasized in the descent direction regardless of the patterns of deconvolved residuals with frequency. This emphasis occurs because the normalization of the CS-II method excessively emphasizes high-frequency components, and in the CS-I method, the dominant frequency of the source wavelet is emphasized. Compared to the two conventional scaling method, the weighting method is much more flexible, and its banded descent directions appropriately describe the spectral variation of the deconvolved residuals, which are caused by the thickness difference between the true geological and the assumed models. In addition, the weighting method is not affected by the source spectrum (Figures 4-3e and 4-3f) and the Nyquist frequency (Figures 4-5c and 4-5d).

Figure 4-13 shows P-wave velocity models inverted by the CS-I, the CS-II and the weighting methods for the thick rectangular-shaped and thin-layers models. Because I fixed Poisson's ratio over the entire model, the inversion results for the S waves are very similar to the results for the P waves. For this reason, I do not display the S-wave velocity models. The velocity structures obtained using the CS-II method are too thin (Figure 4-13c) because the parameter update focuses on the upper part of the anomalous body due to the high-frequency patterns in the banded descent directions. For the same reason, I also observe cycle-skipping artifacts below the anomalous body (see Figure 1 in Bunks et al. (1995) and Figure 7 in Virieux and Operto (2009)). For the thin-layers model, whose deconvolved residuals are relatively unchanged, the velocities recovered using the CS-II method are in good agreement with the true velocities. In Figures 4-13a and 4-13b, we can see that the velocity structures inverted by the CS-I method are not satisfactory.

On the other hand, the velocity structures reconstructed by the weighting method are much improved (Figures 4-13e and 4-13f) for both cases because the frequency components required to effectively recover the true velocities are properly weighted in the early stage of inversion, which yields reliable banded descent directions, as shown in Figures 4-12e and 4-12f.

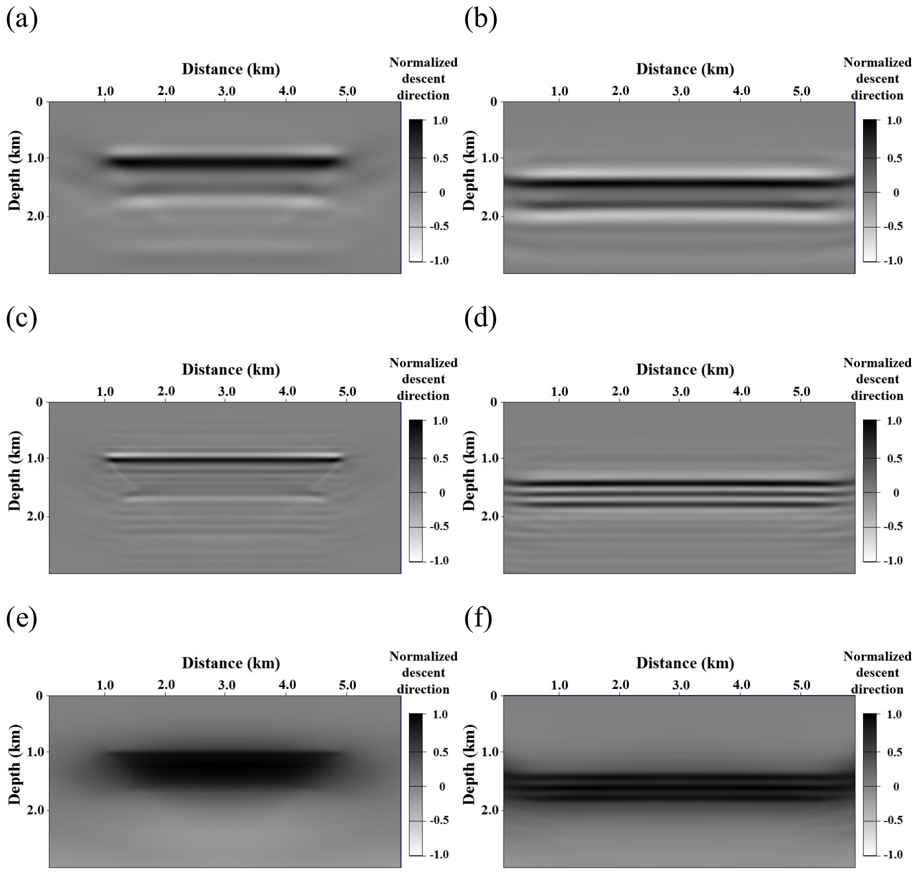
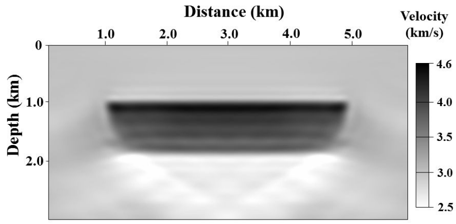
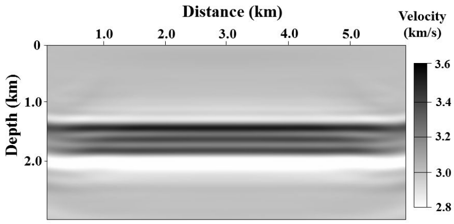


Figure 4-12 Descent directions of P-wave velocity for the thick rectangular-shaped (a, c and e) and thin layers models (b, d and f) at the first iteration obtained using the CS-I (a and b), CS-II (c and d) and weighting method (e and f)

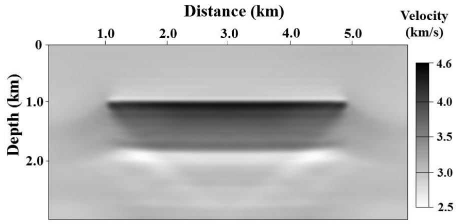
(a)



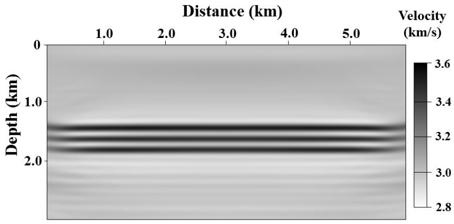
(b)



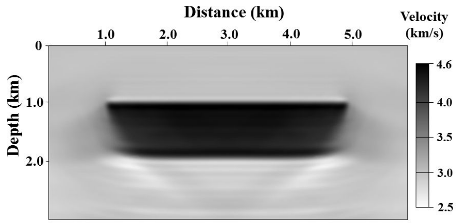
(c)



(d)



(e)



(f)

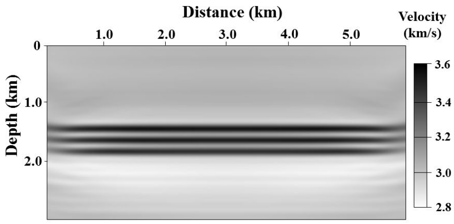


Figure 4-13 P-wave velocity structures for the thick rectangular-shaped (a, c and e) and thin-layers (b, d and f) models obtained at the 20th iteration using the CS-I (a and b), the CS-II (c and d), and the weighting (e and f) methods.

4.4 Numerical Example for SEG/EAGE Salt Model

To examine whether the weighting method can properly recover more complicated models, which have both long- and short-wavelength structures, we perform elastic FWI using the conventional and weighting methods for the SEG/EAGE salt model. Figure 4-14 shows a 2-D section of the SEG/EAGE salt model (AA' line). Parameters for the inversion are listed in Table 4-2. We use all of the frequencies in the range of 0.167 to 10 Hz with a frequency interval of 0.167 Hz although we know that reliable low-frequency signals are difficult to record in field explorations due to the noise (refer Chapter 5.1). It is well known that low-frequency components below 1 Hz are essential to properly invert salt models when initial velocities are poorly estimated (Abubakar et al. 2011). Plessix (2009) noted that low-frequency data can be obtained if we include the data recorded by OBS, which has been widely used in detecting earthquake waves. Nevertheless, using such low frequencies (0.167 Hz) may make the example of the salt model appear unrealistic. However, our purpose is to demonstrate if the weighting method can be effectively applied to the salt model, and we assumed the ideal case where very low frequencies are available. Figure 4-15 shows the velocity models inverted by the CS-I, the CS-II methods when we use the gradually increasing velocity and density models for the initial guesses. The initial P-wave velocities vary from 1.5 to 3.06 km/s, and the initial S-wave velocities are determined from the initial P-wave velocities and Poisson's ratios fixed at 0.25. In Figure 4-15, we observe that the salt bodies recovered by both the

CS-I and CS-II methods are thinner than the salt body of the true model and that the velocities below the salt body are higher than the velocities of the original versions. These phenomena are commonly encountered in conventional elastic waveform inversion for the salt model even though we also use very-low-frequency data. The results for the CS-I method (Figure 4-15a) have similar problems to the results from the CS-II method. On the other hand, the P-wave, S-wave velocity models and the density obtained using the weighting method (Figure 4-19) agree well with the true velocity models even though the initial models are poorly estimated, and the frequency selection strategy is not applied. In Figure 4-19, I confirm that, using the weighting method, we can recover the thickness of the salt body properly and that the background velocities below the salt body are also well reconstructed by controlling the spatial resolution automatically.

Figures 4-16, 4-17 and 4-18 show the banded descent directions for the P-wave velocity obtained by the CS-I, CS-II and the weighting methods at the 100th and 200th iterations. In the early stage of inversion, the low-frequency components of the descent directions are heavily weighted without any human intervention in the weighting method to reduce errors caused by the thick salt body. As iteration proceeds, the descent directions constructed by the weighting method become much more compatible with the true velocity model, while the images produced using the CS-I and the CS-II methods are not notably improved.

Figure 4-20b shows the normalized weighting factors extracted at the 100th and 200th iterations to confirm the relationship between the weighting factors and the descent directions shown in Figures 4-18 and 4-20. In the early stage of inversion, the weighting factors for low frequencies are greater than the factors for the high frequencies, which contributes to recovering the long-wavelength structures related to the salt body. Because large errors in the low-frequency band are reduced as the iterations proceed and the thick salt body is recovered to some extent due to good descent directions, the weights for the high-frequency band become relatively larger to restore the short-wavelength structures. This behavior of my weighting factor, in particular, the weights for the low-frequencies in early stage of in version, provides weighting effects similar to those of Lazaratos et al. (2011) and Routh et al. (2011). However, compared with their fixed shaping, the weights of my weighting method are automatically and flexibly determined depending on the spectral differences of the deconvolved residual during the FWI.

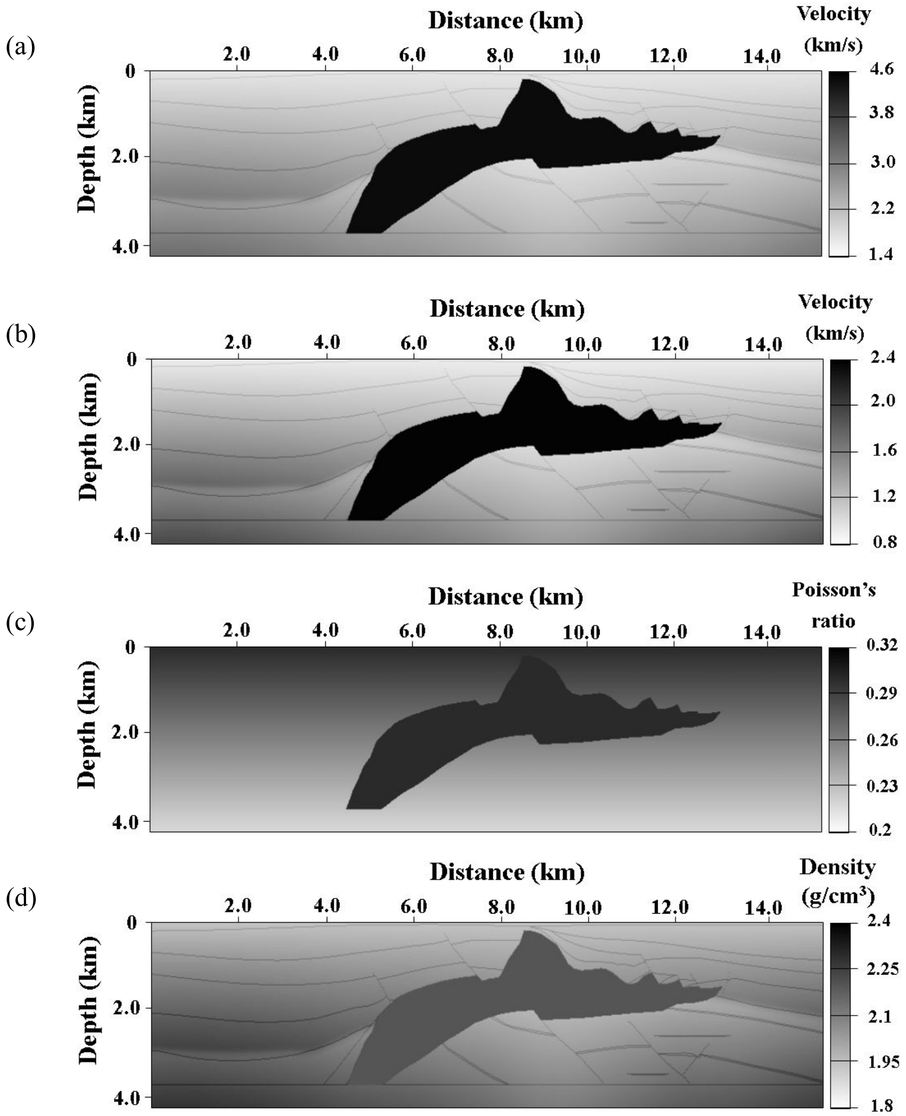
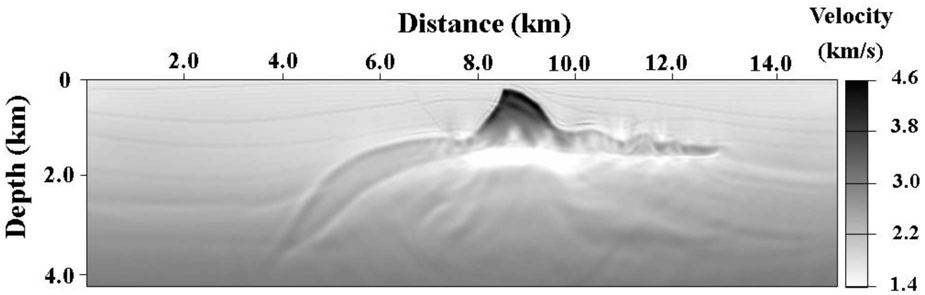


Figure 4-14 The SEG/EAGE salt model: (a) the P-, (b) the S-wave velocities, (c) the Poisson's ratios and (d) the densities.

(a)



(b)

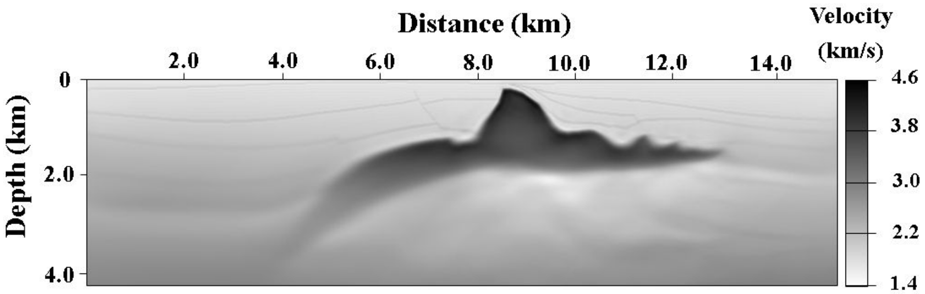
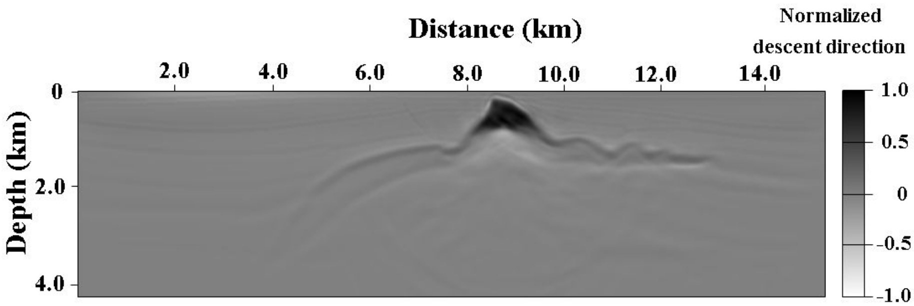


Figure 4-15 P-wave velocity structures obtained at the 300th iteration using (a) the CS-I and (b) the CS-II methods for the SEG/EAGE salt model when the initial P-wave velocity increases linearly from 1.5 to 3.06 km/s.

Table 4-2 Inversion parameters used for the 2D SEG/EAGE salt model

Dimension	No. of shots	No. of receivers	Interval of shots	Interval of receivers	Recording time	Maximum Frequency	Minimum Frequency
15.6 km × 4.2 km	390	781	0.04 km	0.02 km	6 sec	10 Hz	0.167 Hz

(a)



(b)

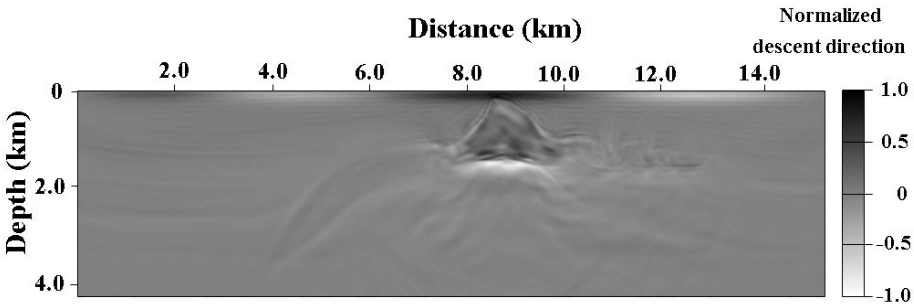


Figure 4-16 Descent directions for the P-wave velocity obtained at the (a) 100th and (b) 200th iterations using the *CS-I method*.

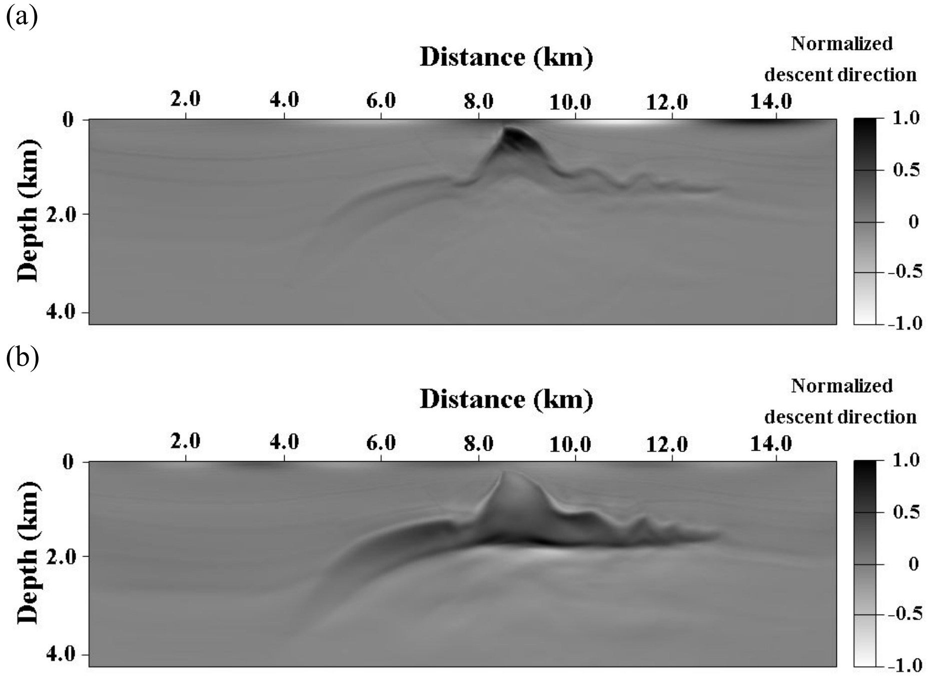
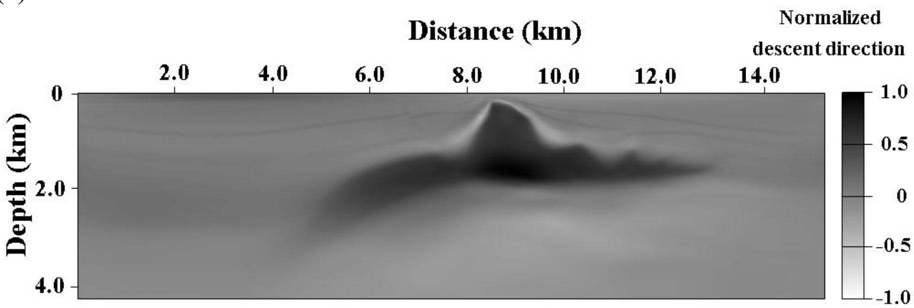


Figure 4-17 Descent directions for the P-wave velocity obtained at the (a) 100th and (b) 200th iterations using the *CS-II method*.

(a)



(b)

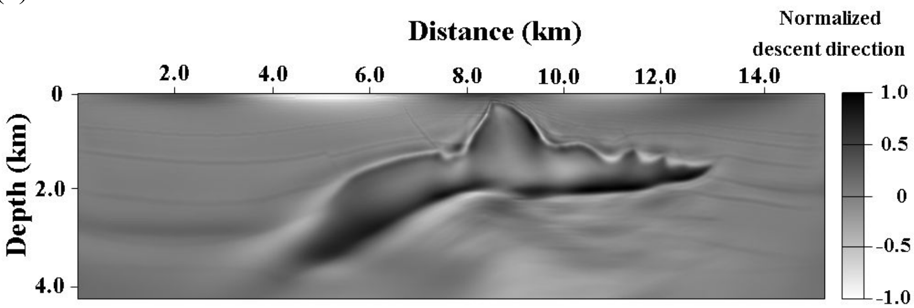


Figure 4-18 Descent directions for the P-wave velocity obtained at the (a) 100^{th} and (b) 200^{th} iterations using the *weighting method*.

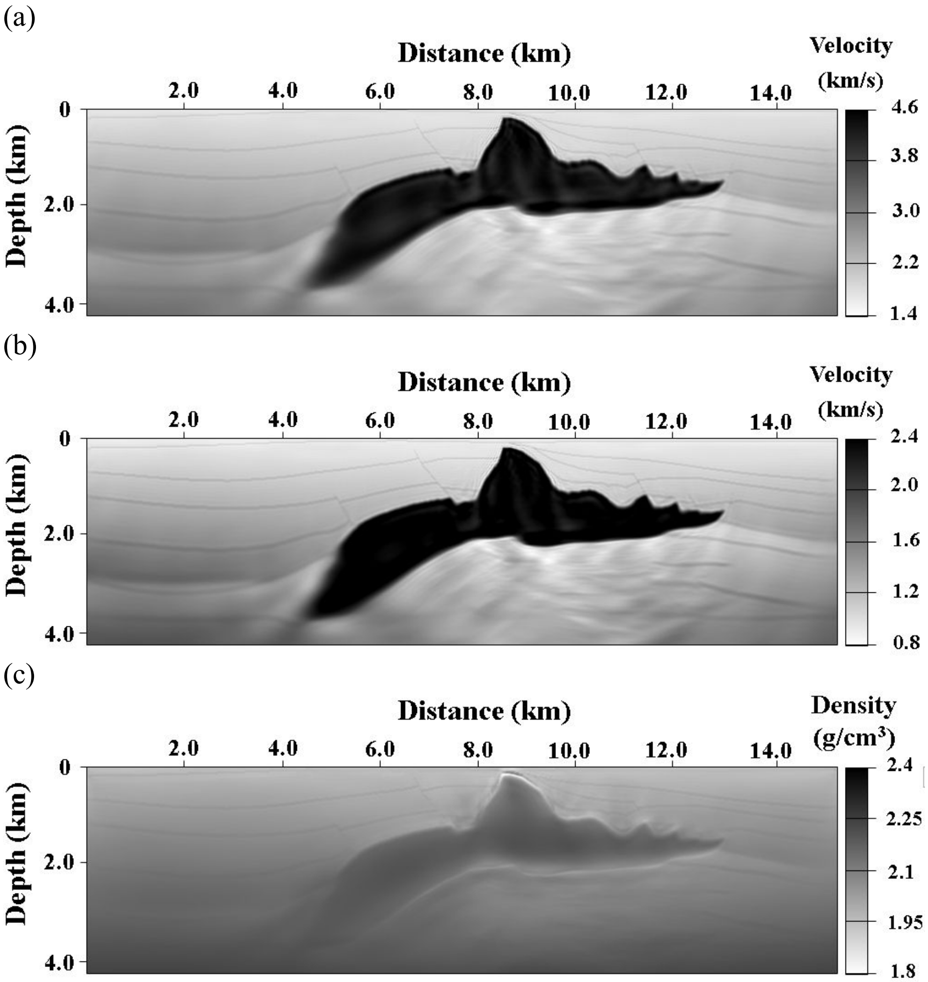
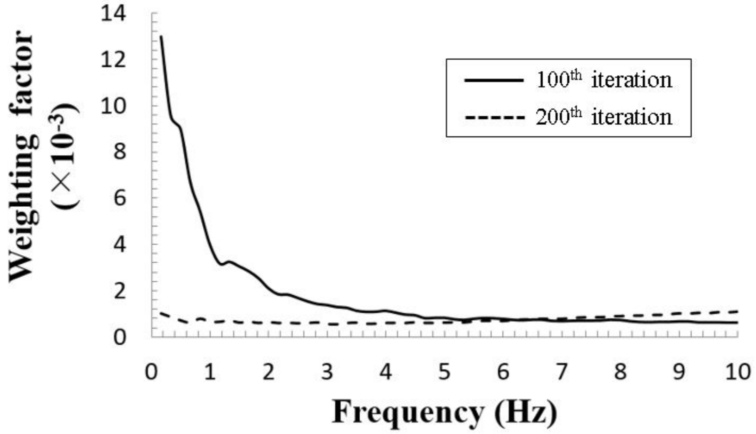


Figure 4-19 (a) P-, (b) S-wave velocity and (c) density structures obtained at the 300th iteration using the weighting method for the SEG/EAGE salt model when the initial P-wave velocity increases linearly from 1.5 to 3.06 km/s.

(a)



(b)

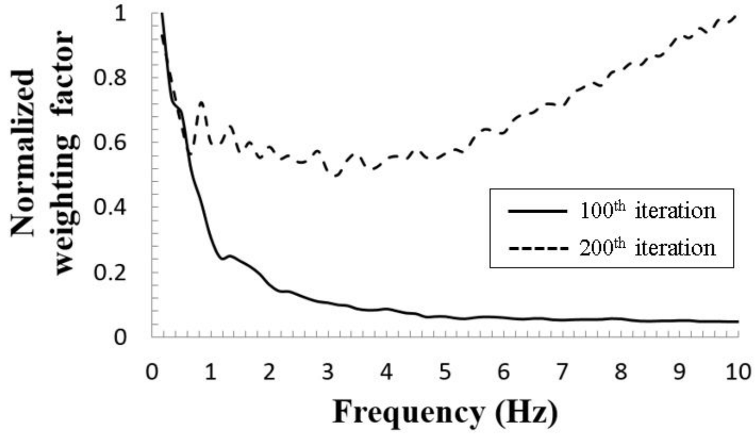


Figure 4-20 (a) The variation of the weighting factors through iterations and (b) that of the normalized weighting factor.

Chapter 5. Spectral Filtering Scheme

5.1 Noise-behavior in the frequency-domain FWI

Before introducing my spectral filtering scheme, I analyze the noise-behavior during frequency-domain FWI in the spectral and spatial aspects. In the inverse problem that we discussed in Chapter 2, the gradient direction is determined to minimize the misfit between the field data and modeled data. If the field data are noise-free, the model parameters obtained around the global minimum can properly describe subsurface structures because the modeled data are also noise-free. However, because field data are usually contaminated by noises, the model parameters determined around the global minimum can deviate from those of subsurface structures. In the former case, the gradient direction is only affected by seismic events, which will be called the 'pure seismic-gradient direction' throughout the paper. In the latter case, the gradient direction is influenced by both seismic events and noises, which will be referred to as the 'total gradient direction'.

In the noise-included inverse problem, because the observed data are expressed by

$$\mathbf{d} = \mathbf{d}^{event} + \mathbf{d}^{noise}, \qquad (5\text{-}1)$$

The total gradient direction for the single-frequency data can be written as follows:

$$\begin{aligned}\nabla_{\mathbf{p}}^{total} E(\omega) &= \sum_{s}\left[\mathrm{Re}\left\{\mathbf{J}^{T}\left(\mathbf{u}_{s}-\mathbf{d}_{s}^{event}\right)^{*}\right\}-\mathrm{Re}\left\{\mathbf{J}^{T}\left(\mathbf{d}_{s}^{noise}\right)^{*}\right\}\right] \\ &= \nabla_{\mathbf{p}}^{event} E(\omega)-\nabla_{\mathbf{p}}^{noise} E(\omega)\end{aligned}, \qquad (5\text{-}2)$$

where the superscripts *event* and *noise* indicate the seismic event and noise, respectively. In eq. (5-2), the total gradient direction in the noise-included inverse problem can be divided into two directions. The first direction, $\nabla_{\mathbf{p}}^{event} E(\omega)$, is the pure seismic-gradient direction, which corresponds to the aforementioned noise-free inverse problem for describing subsurface structures. The second direction, $\nabla_{\mathbf{p}}^{noise} E(\omega)$, is caused by the noise in the observed data, which will be referred to as the 'noise direction' throughout the paper. This direction corresponds to the reverse time migration (RTM) image obtained by only back-propagating the noise components. As demonstrated by eq. (5-2), the parameter change vector in the noise-included inverse problem can be distorted due to the artifacts induced by the RTM image of the noise components. The lower the S/N ratio is, the more distorted are the inversion results. In Appendix C, we discuss the noise direction caused by various types of noise.

In Figure 5-1a, the schematic diagram shows the relationship between the pure seismic and total gradient directions. In the noise-free inverse problem, the inversion process concentrates on finding the seismic global minimum solution, which is the goal of seismic inversion. However, when noise is included in the inverse problem, the inversion process converges to the total global minimum while updating the combination of the pure seismic-gradient direction and the noise direction (e.g., Figure 5-1a). In addition, the parameter update along the noise direction can yield numerous local minima during the inversion process. Consequently, the optimized solutions for noisy data can deviate from the seismic global minimum solutions (which are the desired solutions) due to the parameter updating along the noise direction. For this reason, to obtain reasonable solutions from noisy observed data, we must constrain the data-fitting process along the noise direction to make the model parameter converge to the seismic global minimum solution, as shown in Figure 5-1b.

To do so, the analysis of the noise behavior according to the noise distribution is required because real field data include various types of noise and because the noise distribution is not uniform. The noise distribution can be analyzed with respect to three factors: the receiver, the frequency and the arrival time (or depth). In Figure 5-2a, we display the noise distribution in the residual wavefields with a 3-D cube whose axes represent the receiver, the frequency and the arrival time. The noise distribution along the receiver axis describes the magnitude of the noise recorded at each receiver. The noise distribution along the frequency axis is dependent on the noise type. In other words, the waveform of noise determines the distribution of noise along the frequency axis. The noise distribution along the arrival-time axis indicates when the noise is dominantly recorded and which of the reflected waves (e.g., the waves reflected from shallow or deep structures) are highly contaminated by noise. When considering only primary reflected waves, the arrival time axis can be converted into the depth axis, as is done in migration.

The effect of the noise distribution for the receivers can be minimized by choosing an optimal objective function. Numerical studies have demonstrated that several objective functions, such as the l_1-norm, Huber norm, l_1/l_2 hybrid norm and Student's t-distribution, can improve the robustness of FWI for data including spike-shaped noise. For example, in FWI based on the l_1-norm objective function, normalizing the residuals measured at each receiver using their amplitudes plays a role in making the noise levels over all receivers commensurate with each other, consequently enhancing the robustness of the inversion for spike-shaped noise. However, although these objective functions minimize the effect of the noise distribution for the receivers, they do not consider the noise distributions for frequency and arrival time. This limitation may explain why conventional FWI fails to provide good inversion results for data containing random noise. Frequency-domain FWI is independently conducted for each single-frequency data set. Therefore, if noise dominates the data in certain frequency bands, the gradients at those frequencies can be distorted due to the severely noise-contaminated residual vectors. However, the conventional FWI methods do not prevent these distorted gradients from contributing to the final gradient direction.

In general, because the gradients for deeper structures have smaller values than those for shallower structures due to geometrical spreading effects, the model parameter updates are focused on shallow parts in the early iterations and then move to deeper parts as the iteration proceeds in the frequency-domain FWI. For noisy data, if shallow structures are recovered to some extent as the iteration proceeds, the percentage of noise increases in the residual wavefields for the early-arrival reflected waves, distorting the gradients for the shallower parts and degrading the inversion results. To obtain reasonable results from noisy data, the influence of the noise distributions for frequency and depth, as well as the receivers, should be suppressed.

In this chapter, I will introduce the denoise function, which helps to reduce the data-fitting procedure along the noise direction for frequency axis. With related to the technique to suppress the parameter updating along the noise direction for depth axis, in Chapter 6-2, I will discuss the depth scaling strategy using Levenberg-Marquardt method. For the easy analysis, I do not consider the spectral weighting scheme, which is introduced in Chapter 4. This Chapter is edited version of Oh and Min (2013b).

(a)

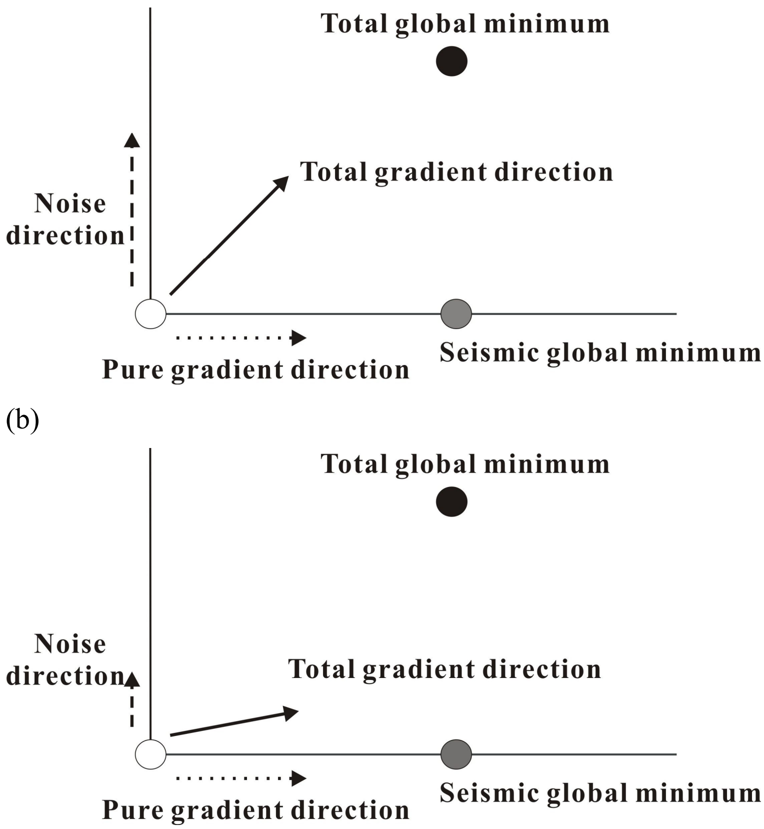


Figure 5-1 Schematic diagrams illustrating the parameter search directions in the noise-included FWI algorithm (a) without and (b) with constraints refraining from the data-fitting process along the noise direction. The dotted arrows indicate the pure seismic-gradient direction that leads to the seismic global minimum solution (gray filled circle) and the dashed arrows denote the noise direction. The solid arrows represent the total gradient direction leading to the total global minimum solution (black filled circle).

(a)

(b)

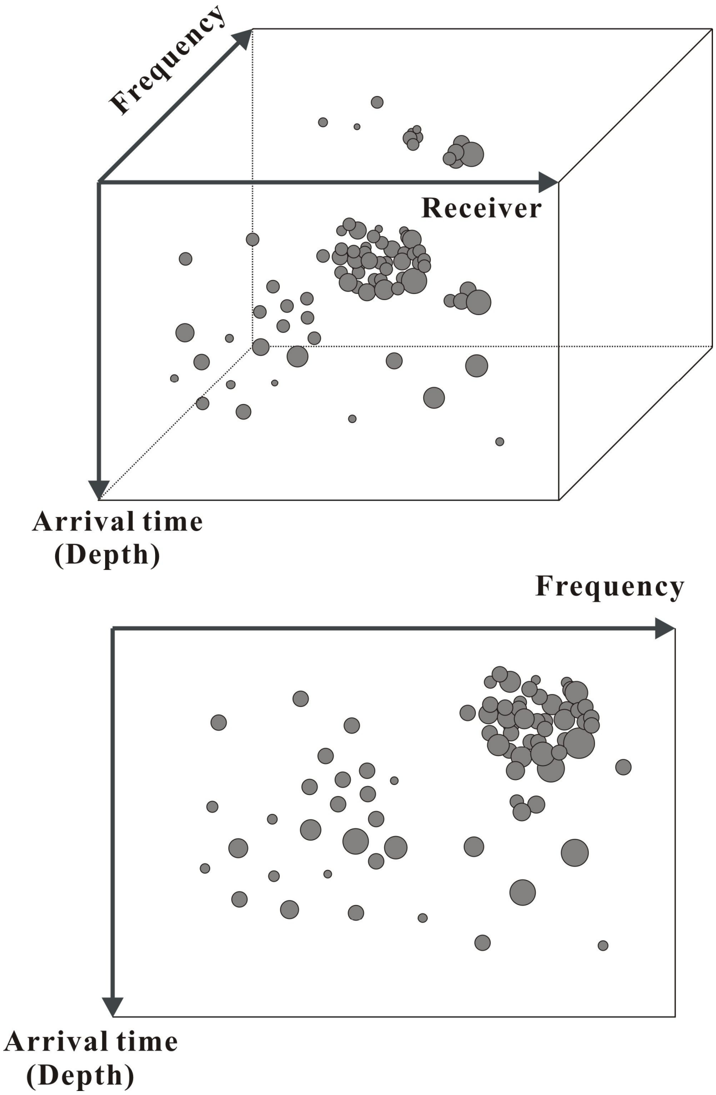


Figure 5-2 Schematic diagrams illustrating (a) a 3D noise-distribution model with frequency, receiver and arrival time (or depth) axes and (b) a 2D noise-distribution model with frequency and arrival time axes. The gray points indicate noise, and their size represents the intensity of the noise.

5.2 Review of Previous Studies on Broadband Seismic Noise

Before introducing the denoise function, we briefly review the previous studies of broadband seismic noise. During a seismic survey, various types of noise can be included in field data (Figure 5-3). Particularly, for a land seismic survey, recorded data can be severely contaminated by dispersive ground roll and ambient ground motion. When we apply the 2-D elastic FWI to field data, the 2-D approximation of 3-D field data can also be a major source of noise because of the amplitude loss caused by spherically expanding wavefronts and some coherent noises reflected from interfaces located out of the vertical plane including the survey line. To alleviate the effects of ground roll and coherent noise, several techniques, such as the f-k filtering method, the borehole seismic survey and the coherent noise removal technique (Abma 1995; Guitton 2003), have been actively studied.

The ambient ground motion has broad frequency range depending on the source of noise. According to Peck (2008), the ambient ground motion can come from two types of sources: natural and cultural. Natural sources include the wind and ocean. Although there are some regional variations, low-frequency ground motions (referred to as microseisms), which are caused by large-scale meteorological events or the wave motion of large bodies of water, typically have a dominant frequency below 0.5 Hz (Bard et al. 2003). Small scale wind also causes ground motions with various frequencies depending on the speed of the wind (Kanasewich 1990). Cultural sources of ground motion are mainly the result of human activities, such as the movement of vehicles and the operation of machinery, and these tend to produce high-frequency vibrations (Butler 1975).

To alleviate the effect of ambient ground motion, we can apply frequency filters to the Fourier-transformed recorded data to cut off undesired frequency components. To design an appropriate frequency-filter, we need to know the noise spectrum for the survey area. However, even though we have information about the noise spectrum, the frequency filter designed for the noise spectrum might not properly work during the FWI. When we design frequency filters, we define the passband as a trapezoid rather than a boxcar to avoid the Gibbs phenomenon (Yilmaz 2001). In this case, the sloping areas of the frequency filters may work against our original intention, depending on the scaling methods used in the inversion algorithms (refer Chapter 4 spectral weighting scheme or Oh and Min 2013a). For this reason, one option for robust FWI is to simply discard undesired frequency components during the inversion, as Shin and Min (2006) did. However, if the observed data have a small amount of noise at a certain frequency, this approach can degrade the inversion results or cause a loss of resolution. Moreover, when the noise is scattered over many frequencies, it may not be easy to filter out the noisy frequency components.

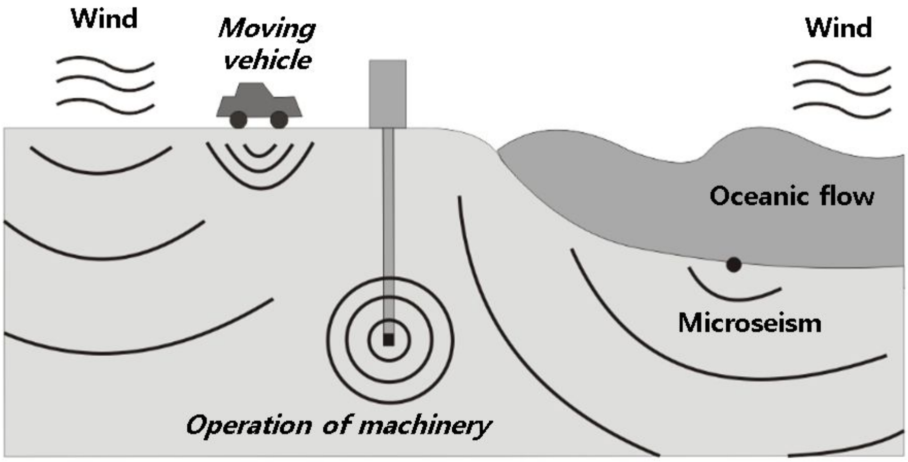


Figure 5-3 Schematic diagram illustrating the generation and propagation of the random noise. Black arcs represent the wavefronts of random noise. The cultural random noises are written in italics.

5.3 Spectral Filtering using Denoise function

5.3.1 Mathematical expression of the Denoise Function

To complement the limitations of frequency filters, I suggest using the denoise function (Oh and Min 2013b). Several studies have shown that some objective functions can improve the robustness of the full waveform inversion for outliers (Pyun et al. 2009) and coherent random noise (Brossier et al. 2010). However, in my experience, it is doubted that these objective functions are also robust to the incoherent random noise, such as the ambient ground motion. Because the spectrum of the ambient seismic noise is independent of the seismic source spectrum, the S/N ratios of the observed data at each frequency are determined by the combination of the source and noise spectra. Therefore, we need to introduce a denoise function during seismic waveform inversion to provide reasonable weights to each single-frequency descent (or gradient) direction, depending on its S/N ratio. Considering that field data contain noise and that synthetic data are noise-free, I construct the denoise function as

$$g(\omega) = \left[\frac{\sum_r \left| \sum_s u_{s,r}(\omega) \right|}{\sum_r \left| \sum_s d_{s,r}(\omega) \right|} \right]^e , \tag{5-3}$$

where s and r indicate the shot and receiver numbers, respectively, and e is a control factor to adjust the degree of noise suppression.

The denoise function is designed based on the characteristics of the seismic signal and the ambient ground motion. Considering the randomness of the noise sources (wind, oceans and human activities), I assume that the ambient ground motions are randomly recorded at each receiver of each shot gather. In contrast, the monochromatic seismic signals resemble sine or cosine curves, although they are damped depending on the propagation distance. We know that summing sine or cosine curves with different phases causes their amplitudes to cancel each other out in some places. Based on this fact, we can guess that summing the seismic signals with those recorded at adjacent shot gathers will have the effect of suppressing signals if the number of shots is enough. If the distance between shots is far enough, the degree of amplitude suppression is not big due to the damping effect of signals and thus amplitude level of summed signal can be maintained in the original level.

The denoise function is based on these properties of signal and random noise. For the denoise function, observed and modeled data are first summed over the entire shot (Step 1), which likely cancels out seismic signals due to their monochromatic property but also amplifies certain types of noise, particularly random noise. By summing the absolute values of summed data over the entire receiver (Step 2), we can roughly measure the ratio of signal to noise over different frequencies.

If we assume that field data are noise-free at certain frequencies, the denoise function can be approximated as 1 if the assumed velocity structure and the estimated source wavelet are close to the true velocity structure and the true source wavelet, respectively, as the iteration proceeds:

$$g(\omega)=\left[\frac{\sum_r\left|\sum_s u_{s,r}^{\mathrm{model}}(\omega)\right|}{\sum_r\left|\sum_s d_{s,r}^{\mathrm{model}}(\omega)\right|}\right]^e=1 \qquad (5\text{-}4)$$

Here, the superscript 'model' indicates that the respective variables are purely derived from geological models and do not include noise. For noisy data, the denoise function can be approximately proportional to the S/N ratio because monochromatic signals tend to be partially canceled out by those at adjacent shot gathers when they are added together (see Figures 5-7, 5-8 and 5-9). As a result, the denominator of eq. (5-4) can dominated by noise, which can be expressed as

$$g(\omega)=\left[\frac{\sum_r\left|\sum_s u_{s,r}^{\mathrm{model}}(\omega)\right|}{\sum_r\left|\sum_s d_{s,r}^{\mathrm{model}}(\omega)+d_{s,r}^{\mathrm{noise}}(\omega)\right|}\right]^e\approx\frac{s(\omega)}{n(\omega)}, \qquad (5\text{-}5)$$

where $s(\omega)$ and $n(\omega)$ are the spectra of the seismic signals and noises, respectively.

Introducing the denoise function into eq. (2-11) gives

$$\delta\mathbf{p}_{\mathrm{descent}}^{(l)}=-\mathrm{NRM}\sum_{\omega}\left\{g(\omega)\times\mathrm{NRM}\left[\left[\mathrm{diag}\{\mathbf{H}(\omega)\}\right]^{-1}\nabla_{\mathbf{p}}E(\omega)\right]\right\}. \qquad (5\text{-}6)$$

In eq. (5-6), we expect that the denoise function plays a role in weakening the influence of severely noise-contaminated frequency components on the total descent direction and acts similar to a frequency filter for descent (or gradient) direction.

The degree of noise suppression is also affected by the control factor, e. If the control factor is 0, the denoise function is 1, and the gradient vector is the same as that of the conventional waveform inversion. The larger the control factor is, the more strictly noise is suppressed, which means that the slope of the denoise function becomes steeper. An appropriate control factor can be chosen depending on which we prefer between noise suppression and spatial resolution because there is a trade-off between noise suppression and resolution loss in inversion results depending on the control factor.

5.3.2 Practical Aspects of the Denoise Function

We investigate the effect of the denoise function by applying the l_2-norm elastic FWI to synthetic data with several monochromatic random noises for the layered model shown in Figure 5-4. For simplicity, we assume that the Poisson's ratio and density are constant at 0.25 and 2.0 g/cm^3, respectively, for the entire model. The inversion parameters are listed in Table 5-1. The first derivative of the Gaussian function is used as a seismic source wavelet. The initial P- and S-wave velocities gradually increase from 1.5 to 4.5 km/s and from 0.866 to 2.581 km/s, respectively.

To investigate why the conventional FWI is weak for random noise produced by the ambient ground motion, we assume that the observed data include several types of monochromatic noise. We add monochromatic random noise only at integer frequencies, with the maximum amplitude set such that the spectral S/N ratio (i.e., maximum amplitude of signal over maximum amplitude of random noise) is 2 at all frequencies. Although this example is unrealistic, the approximation is useful for investigating problems of the conventional method and for assessing the sensitivity of the denoise function for the banded data. Figure 5-5 shows the real part of the Fourier-transformed true data, contaminated by monochromatic random noise at only integer frequencies. Because random noise is added at integer frequencies, the seismic signals at 5 and 10 Hz are contaminated by random noise, whereas those at 2.67 and 7.67 Hz are noise-free. Figure 5-6 shows single-frequency descent directions of P-wave velocity obtained when monochromatic random noise is added to observed data in the inversion. By comparing the descent directions obtained for the random noise-added data (Figure 5-6b) with those for the noise-free data (Figure 5-6a), I confirm that random noise distorts the single-frequency descent directions.

Figure 5-7a shows the amplitude spectra of the monochromatic random noise and the noise-free observed data. To consider all the data at each frequency, amplitudes of data and random noise are summed over whole receivers and shots. Figure 5-7b shows the variation of the denoise function during the inversion. In the early stage of the inversion, the denoise function for the noise-free frequency components deviates from 1 because the modelled data deviate from field data. However, as the iteration proceeds, the denoise function approaches 1, supporting eq. (5-4). In the noise-contaminated frequencies, the denoise function has relatively small values which indicates that the denoise function is proportional to the S/N ratio of the random noise-included data, as shown in eq. (5-5). In Figure 5-7b, the denoise function for monochromatic random noise resembles a notch filter. Based on these results, we expect that the denoise function will effectively filter out severely noise-contaminated gradients during the inversion.

However, we can also observe that the values at the low-frequency bands are relatively larger than those at the high-frequency bands, although the S/N ratio is the same over the entire frequency spectrum. This phenomenon can be explained by the spectral sensitivity of the denoise function. Figures 5-8 and 5-9 show the principle of the denoise function (i.e., Step 1 in eq. 14) at 10 and 5 Hz, respectively.

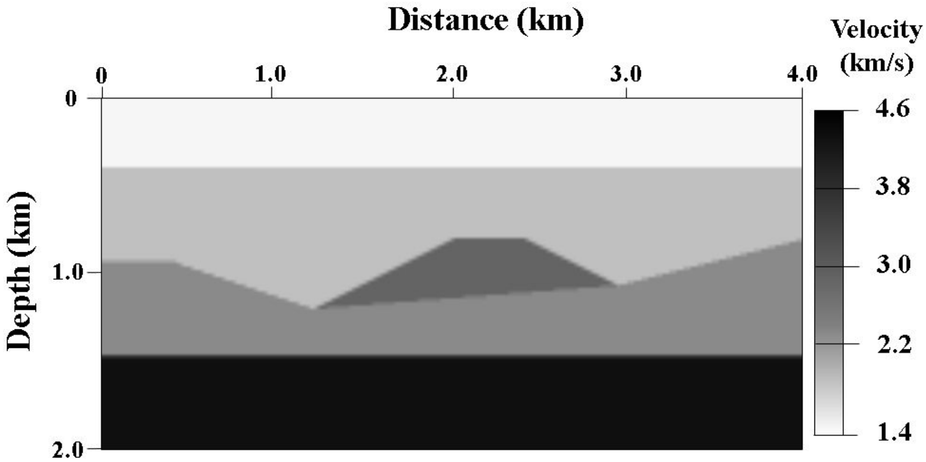
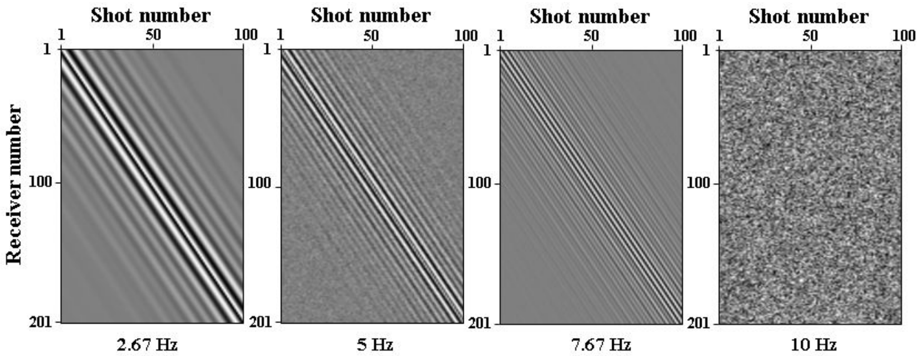


Figure 5-4 (a) P-wave and (b) S-wave velocity structures for the layered model.

Table 5-1 Parameters used in the inversion for the layered model

Dimension	No. of shots	No. of receivers	Interval of shots	Interval of receivers	Recording time	Maximum Frequency	Minimum Frequency
4.0 km × 2.0 km	100	201	0.04 km	0.02 km	3 sec	10 Hz	0.33 Hz

(a)



(b)

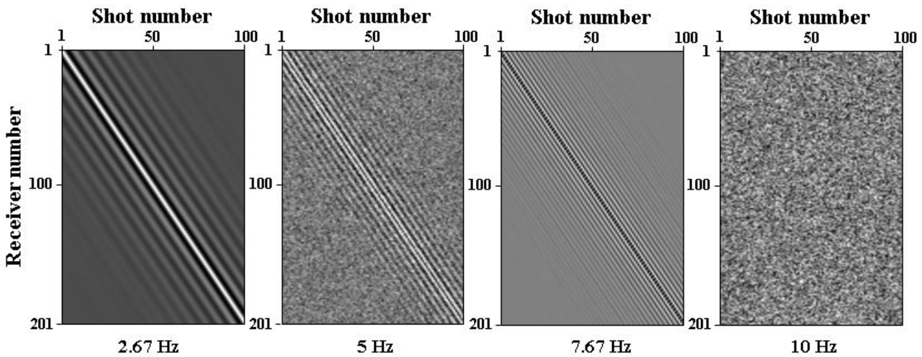
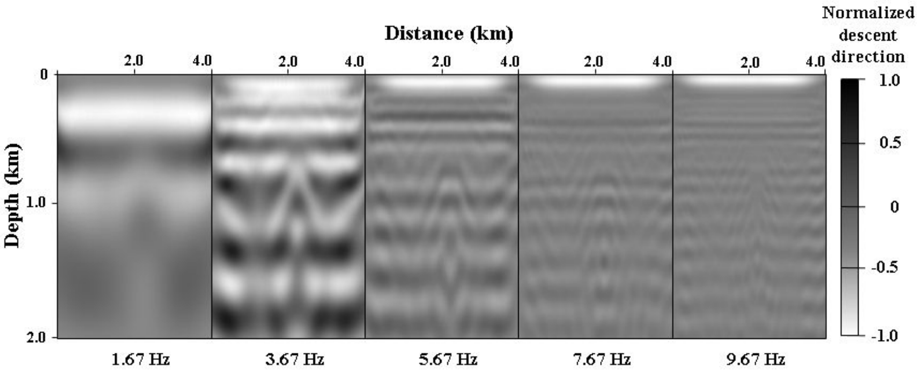


Figure 5-5 Real parts of the (a) horizontal and (b) vertical displacements of frequency-domain synthetic data with *monochromatic random noise* obtained for the layered model at 2.67 Hz, 5 Hz, 7.67 Hz and 10 Hz.

(a)



(b)

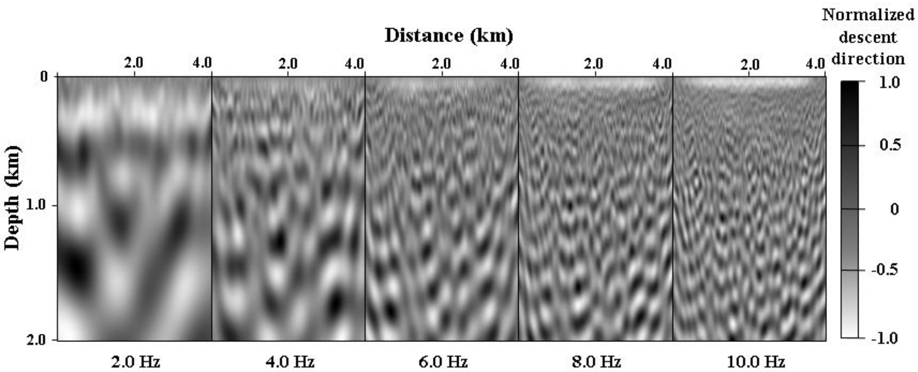
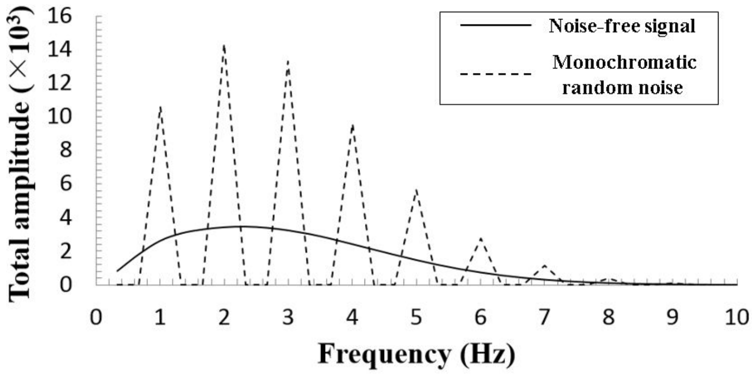


Figure 5-6 Mono-frequency descent directions of the P-wave velocity obtained at the 5th iteration for (a) noise-free data and (b) data containing *monochromatic random noise* obtained for the layered model.

(a)



(b)

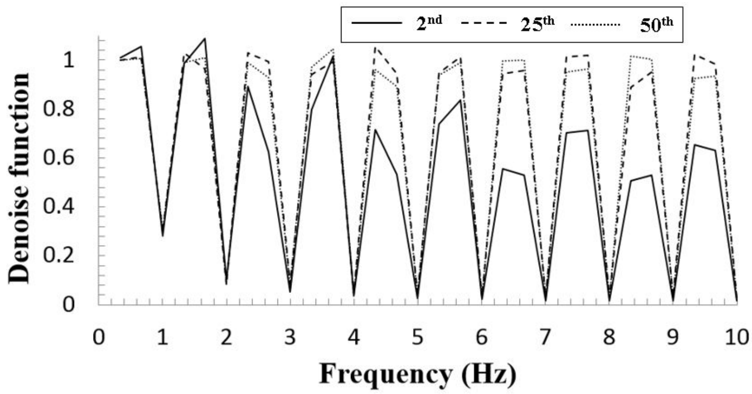


Figure 5-7 (a) Amplitude spectra of the *monochromatic random noise* (red line) and noise-free seismic signal generated using the *first derivative of the Gaussian function* (black line), summed over shots and receivers, and (b) spectra of the denoise functions (e=2) at the 2nd (solid black line), 25th (dashed black line) and 50th (solid red line) iterations.

(a)

(b)

(c)

(d)

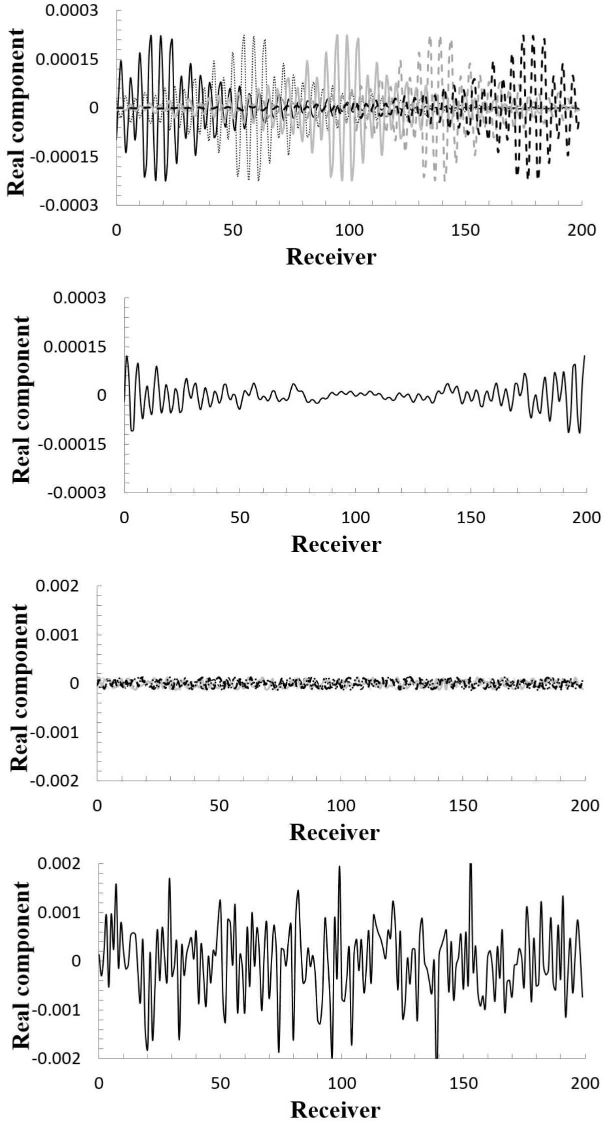


Figure 5-8 Real components of the horizontal displacements for the noise-free monochromatic signals (a and b) and *monochromatic random noises* (c and d) at *10 Hz* before (a and c) and after summation over 100 adjacent shot gathers (b and d). For visualisation, we only display 5 shot gathers for (a) and (c).

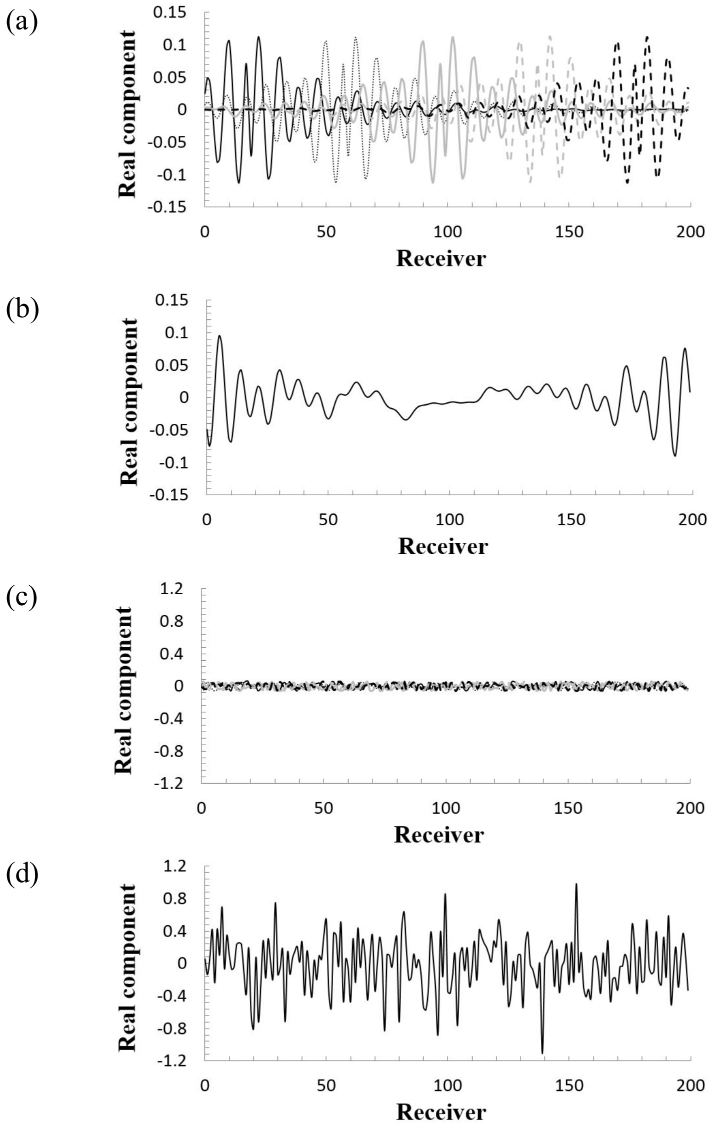


Figure 5-9 Real components of the horizontal displacements for the noise-free monochromatic signals (a and b) and *monochromatic random noises* (c and d) at *5 Hz* before (a and c) and after summation over 100 adjacent shot gathers (b and d). For visualisation, we only display 5 shot gathers for (a) and (c).

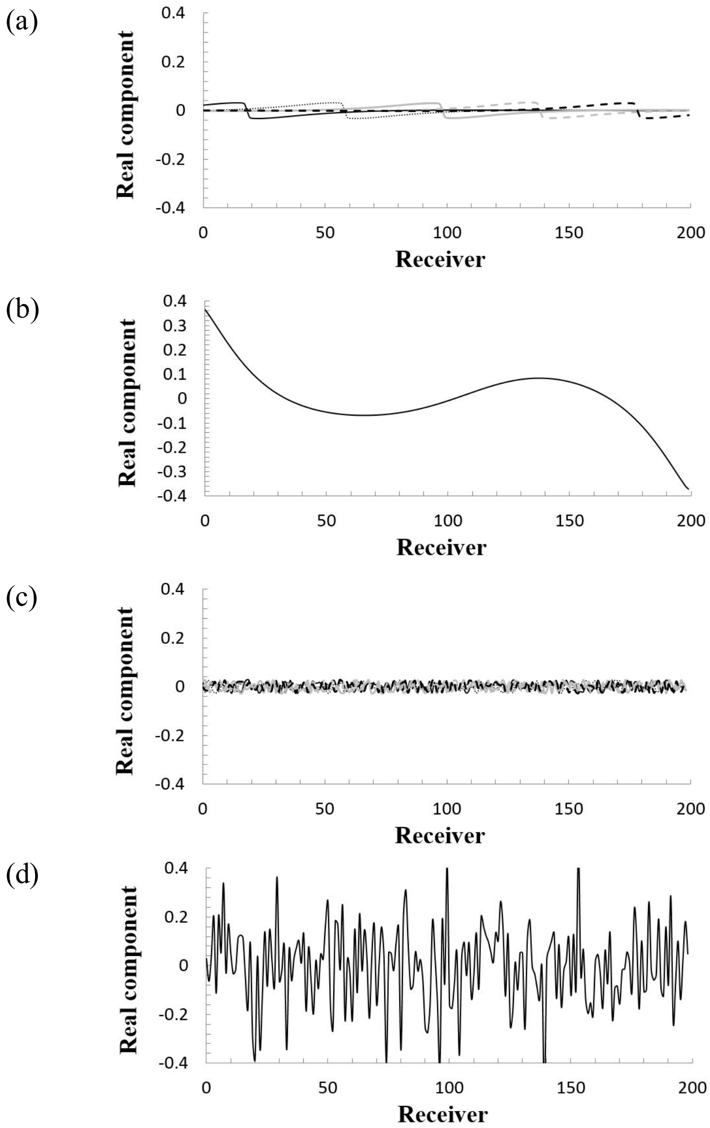
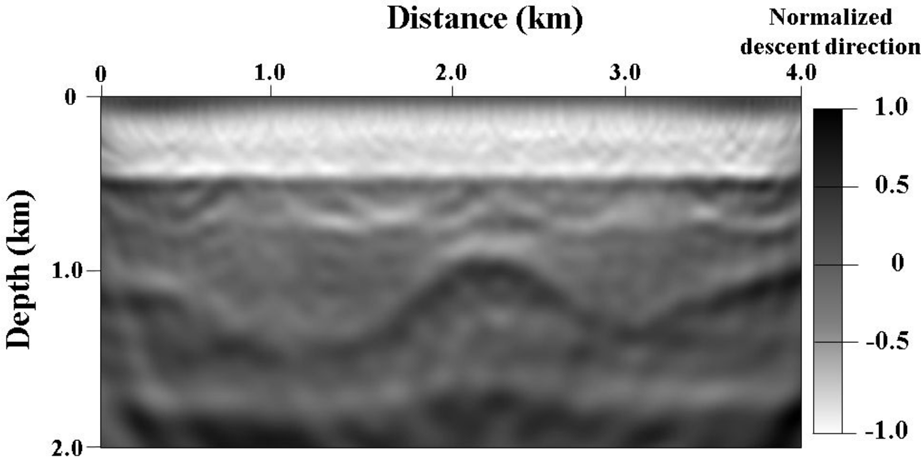


Figure 5-10 Real components of the horizontal displacements for the noise-free monochromatic signals (a and b) and *monochromatic random noises* (c and d) at *0.2 Hz* before (a and c) and after summation over 100 adjacent shot gathers (b and d). For visualisation, we only display 5 shot gathers for (a) and (c).

In Figures 5-8 and 5-9, we display raw monochromatic seismic signals, random noises and their respective summed data over 100 adjacent shot gathers. For raw data, we only show 5 shot gathers for visualization. As we mentioned in the previous section, the monochromatic seismic signals are cancelled out and their amplitudes are lowered, particular in the central part of the survey line, where superposition occurs most often (Figures 5-8b and 5-9b). In contrast, the random noise is amplified, although some reductions also occur (Figures 5-8d and 5-9d). Figure 5-10 shows raw and summed seismic signals and random noise at 0.2 Hz. Compared to the high-frequency examples, the extremely low-frequency seismic signals are amplified because of their long wavelengths. Based on Figures 5-8b, 5-9b and 5-10b, I can confirm that the sensitivity of the denoise function is inversely proportional to frequency. That is why the value of the denoise function increases at low frequency in Figure 5-7b even though I suppose the same S/N ratio of the observed data through all the frequencies. However, if we take the sum of only 5 shot gathers to calculate the denoise function, the amplification of the low-frequency signals will not be so large, and the sensitivity of the denoise function for the very low-frequency data will be improved. In other words, the interval of sources and the number of shot gathers summed for the denoise function will be important for obtaining a more reliable denoise function. Further study is needed on this issue.

Figure 5-11 compares the descent directions of the P-wave velocity obtained at the 5th iteration using the denoise function with those obtained without the denoise function (i.e., the conventional method). The descent direction obtained by the conventional method are distorted even though only 10 frequency components are contaminated by random noise. This occurs because all frequency components contribute to the total descent direction, irrespective of their noise contamination. However, in the FWI using the denoise function, the descent directions are much improved because the noise-dominated components are filtered out by the notch filter-like denoise function. Figure 5-12 shows the inversion results obtained with and without the denoise function. The inverted velocity structures obtained using the denoise function are also more compatible with true velocities than those obtained using the conventional waveform inversion. There results demonstrate that the denoise function is nearly proportional to the S/N ratio of the observed data contaminated by random noise, and they show that it prevents noise-contaminated components from affecting the total descent direction.

(a)



(b)

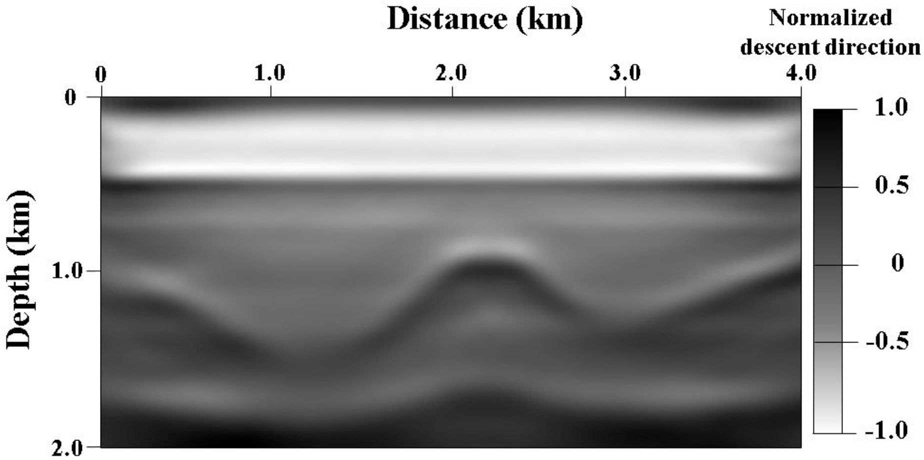
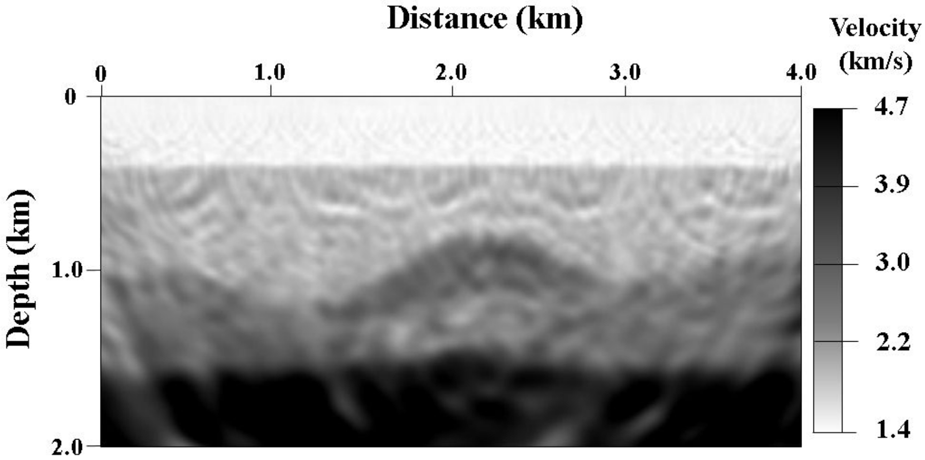


Figure 5-11 Gradient directions of the P-wave velocity obtained at the 5th iteration for data with *monochromatic random noise,* (a) without and (b) with the denoise function.

(a)



(b)

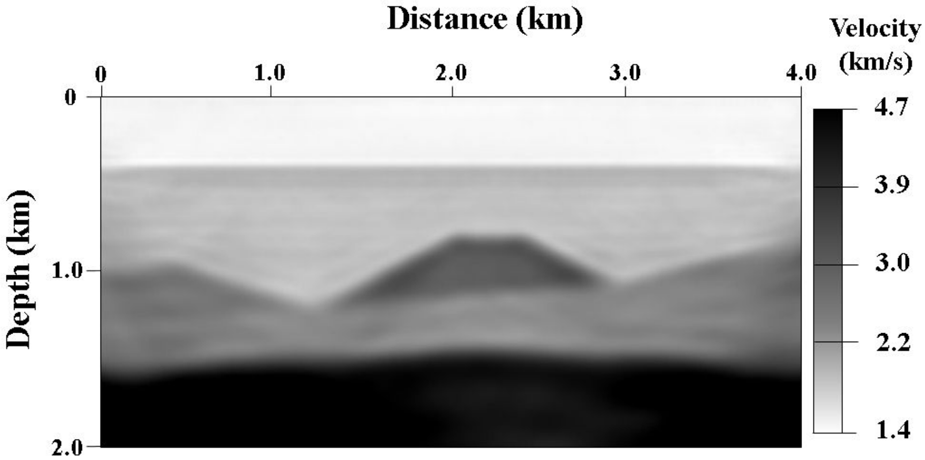


Figure 5-12 Recovered P-wave velocity structures at the 50th iteration (a) without and (b) with the denoise function for data with *monochromatic random noise* added at only integer frequencies.

5.4 Numerical Examples: Elastic Marmousi-2 Model

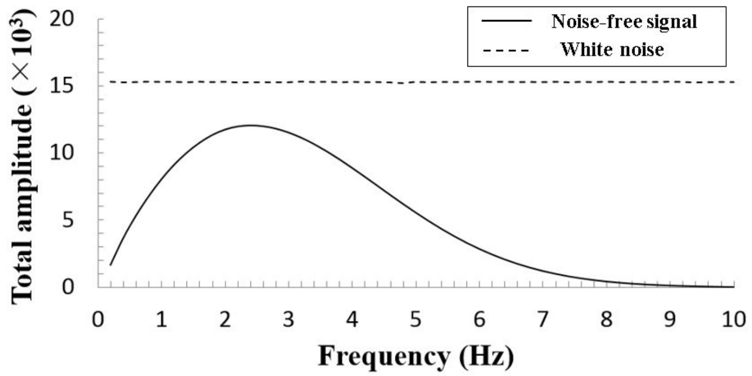
I need to demonstrate that the denoise function can improve inversion results for data contaminated by various types of random noise. To do this, we conduct the l_2-norm elastic FWI for the modified version of the elastic Marmousi-2 model (Figure 3-16) with linearly increasing initial models (Figure 3-17). We assume three types of random noise: white, low-frequency and high-frequency.

5.4.1 Effect of Source Spectrum

In the frequency-domain FWI, the influence of noise is determined not only by the noise spectrum but also by the source spectrum. To investigate the effect of the source spectrum during FWI of noisy data, we perform the elastic FWI for synthetic data with white random noise, which is randomly generated in the frequency-domain. We suppose two types of source spectra. One is the first derivative of the Gaussian function, whose energy is mainly concentrated around a quarter of the maximum frequency (Figure 4-3). The other source function is the Ricker wavelet, whose central frequency is approximately half of the maximum frequency as shown in Figure 4-3. Figure 5-13a shows total amplitude spectra of the random noise and noise-free signals. In Figure 5-13a, we observe that the high-frequency components are severely contaminated by the random noise due to the weak energy of the first derivative of the Gaussian function at high frequencies. Considering that high-frequency components mainly contribute to recovering short-wavelength structures in the inversion, we expect that short-wavelength structures are severely distorted by the random noise.

In this numerical example for white random noise, we can guess that the spectral S/N ratio is proportional to the source spectrum because the density of noise is uniform for the entire frequency. Figure 5-13b shows the denoise function over varying frequency for white random-noisy data during the inversion. The denoise function resembles the source spectrum, except at low-frequency bands. This is due to the insensitivity of the denoise function at low-frequency bands, as was mentioned before (Figure 5-10). In this case, the denoise function resembles a low-pass filter, which helps to suppress the effects of noisy high-frequency components of the descent direction on model parameter updates. In other words, the denoise function acts as a frequency-filter for the descent direction (or gradient direction) that is designed semi-automatically (except the control factor, e) from the spectra of the random noise in field data without any prior information. Figure 5-14 shows the descent direction of the P-wave velocity obtained with and without the denoise function at the 50^{th} iteration. The descent direction obtained without the denoise function (Figure 5-14a) is contaminated by the random noise because of the large contributions of high-frequency descent direction, which is induced by the normalizing operator in eq. (2-12). On the other hand, the effects of the random noise are not dominant in the descent direction obtained using the denoise function, which indicates that the noisy gradients for high frequencies are effectively filtered out, although some low-frequency artifacts are updated due to the inaccurate value of the denoise function at low-frequency bands, as shown in Figure 5-13b. Figure 5-15 shows the recovered velocity structures obtained using the FWI with and without the denoise function at the 130th iteration. In the conventional method (Figure 5-15a), the recovered P-wave velocities are also distorted by the random noise. Detailed structures are poorly inverted, particularly at greater depths. The P-wave velocity structure inverted using the denoise function (Figure 5-15b) are fairly compatible with the true P-wave velocity structure.

(a)



(b)

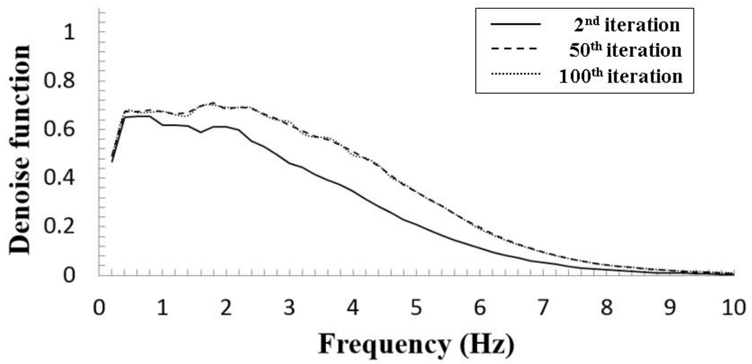
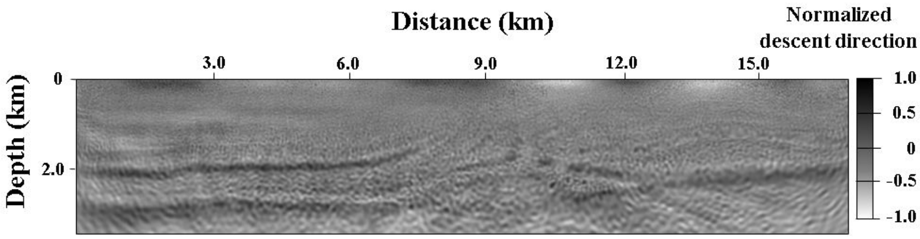


Figure 5-13 (a) Amplitude spectra of the *white random noise* (dashed line) and noise-free seismic signal generated using the *first derivative of the Gaussian function* (black line), summed over shots and receivers, and (b) spectra of the denoise functions (e=1) at the 2^{nd} (solid black line), 50^{th} (dashed black line), and 100^{th} (dotted black line) iterations.

(a)



(b)

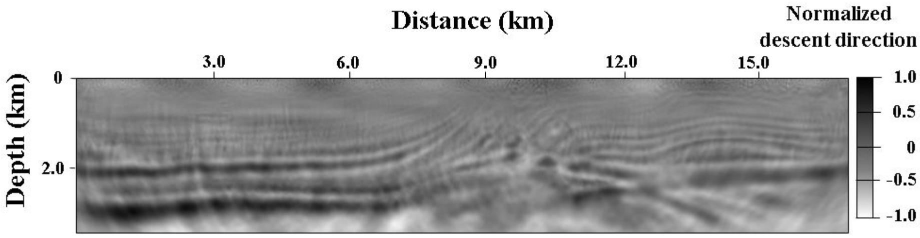
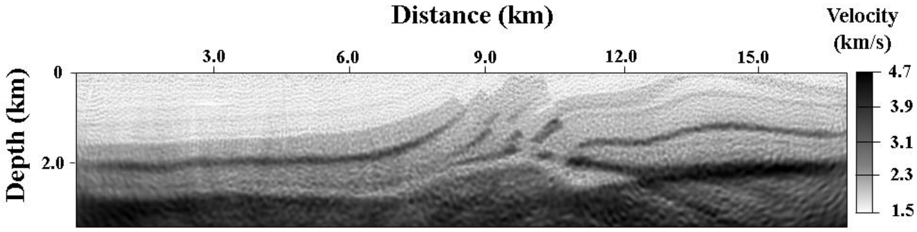


Figure 5-14 Descent directions of the P-wave velocity obtained at the 50^{th} iteration for the *white random noise*–included data generated using the *first derivative of the Gaussian function* (a) without and (b) with the denoise function.

(a)



(b)

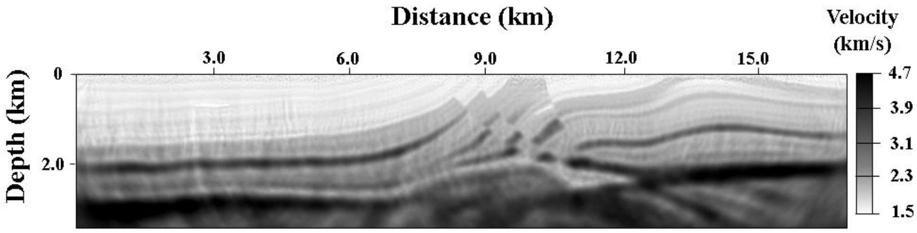


Figure 5-15 Recovered P-wave velocity structures at the 130th iteration (a) without and (b) with the denoise function for data with *white random noise* obtained by the *first derivative of the Gaussian function.*

To verify the behavior of the denoise function for different source wavelets, we also perform the frequency-domain elastic FWI for the observed data obtained by the Ricker wavelet with the same inversion setting as the previous example. The same white random noise is also added to the observed data, as shown in Figure 5-16a. In this case, the low-frequency components of the observed data are severely contaminated by the random noise due to the weak energy of the Ricker wavelet at the low-frequency bands (Figure 4-3). As shown in Figure 5-16b, the denoise function has a similar spectrum to that of the Ricker wavelet. Comparing Figure 5-16 with Figure 5-13, we note that the denoise function for the Ricker wavelet is different to that of the first derivative of the Gaussian function in two aspects. First, the denoise function for the Ricker wavelet does not have large errors at low frequencies as a result of the Ricker wavelet as the low-frequency band. As shown in eq. (5-5), if the signals ($u_{s,r}^{\text{model}}(\omega)$) are sufficiently weak compared to the noise ($d_{s,r}^{\text{noise}}(\omega)$), we can guess that the small variation of the signals does not cause any large variation in calculating the denoise function. The other dissimilarity is the unexpectedly large contribution at high-frequency band in spite of the low S/N ratio. The reason can be explained by a resonance frequency of denoise function. Because the denoise function is calculated by summing all the signals over entire shots, at each single frequency, the results are sensitive to the wavelength of seismic signals. If the velocities of surface are homogeneous in elastic media, the denoise function has two kinds of peaks that are caused by resonances of P-waves and S-waves, respectively, and can be expressed by

$$f_{resonance} = \frac{c}{\Delta s}. \tag{5-7}$$

where Δs indicates the interval of sources. For examples, if the P-wave and S-wave velocities are 1.5 and 1.0 km/s and shot interval is 0.1 km, the denoise function has peak values at every 10 Hz due to S-waves and 15 Hz due to P-waves. When we apply the FWI for marine data, the resonance frequency of the denoise function can be exactly calculated and removed during the FWI. However, this resonance frequency cannot be calculated and might not be severe in elastic media where seismic velocities at surface are heterogeneous. However, when we use homogeneous initial velocities at surface, the resonance frequency causes unexpected large values of the denoise function around resonance frequencies like this numerical examples. In contrast, in the example for the first derivative of the Gaussian function, such resonances did not arise because the first derivative of the Gaussian function has weak energy at high-frequency bands. Figures 5-17 and 5-18 show the descent directions of the P-wave velocity at the 100th iteration and the recovered P-wave velocity structures at the 200th iteration, respectively. The descent direction obtained without the denoise function (Figure 5-17a) is mainly distorted by low-frequency components of random noise because of the small S/N ratio at low frequencies, which is caused by the weak energy of the Ricker wavelet. In contrast, the effects of low-frequency random noise are not dominant in the descent direction obtained using the denoise function (Figure 5-17b) because the denoise function acts similar to a high-pass filter in this case.

Consequently, the P-wave velocities recovered using the denoise function (Figure 5-18) are quite compatible with true velocities. Some structures, such as the layers, the salt body, the unconformity, the anticlines above and below the unconformity and several faults in the central part of the model, are well recovered compared to the results of the conventional FWI. However, giving less weight to the noise-dominant low-frequency data makes the convergence rate slow, and the deeper structures are not properly recovered, although there are great improvements.

(a)

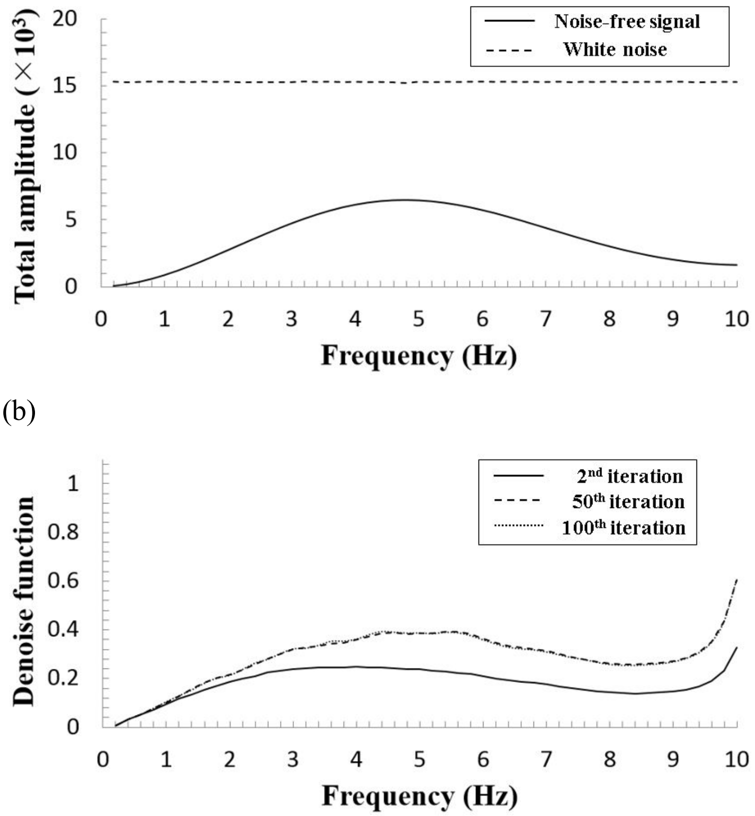
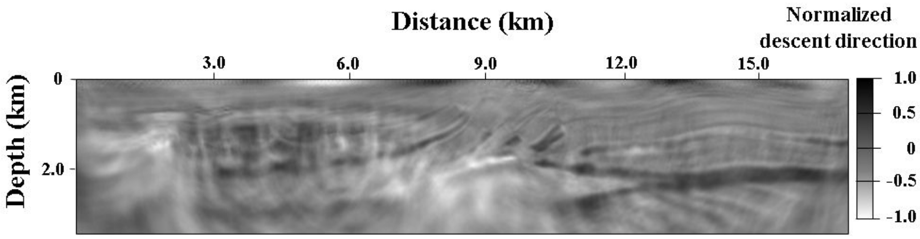


Figure 5-16 (a) Amplitude spectra of the *white random noise* (dashed line) and noise-free seismic signal generated using the *Ricker wavelets* (black line), summed over shots and receivers, and (b) spectra of the denoise functions (e=1) at the 2nd (solid black line), 50th (dashed black line), and 100th (dotted black line) iterations.

(a)



(b)

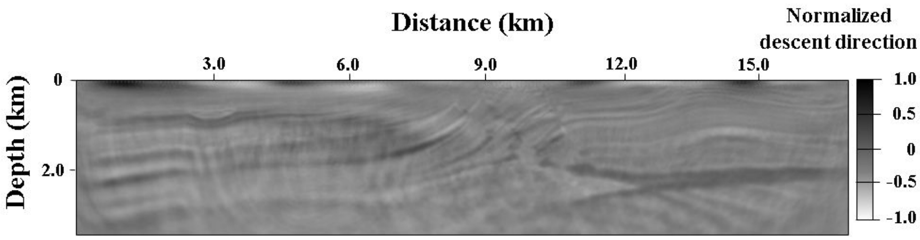
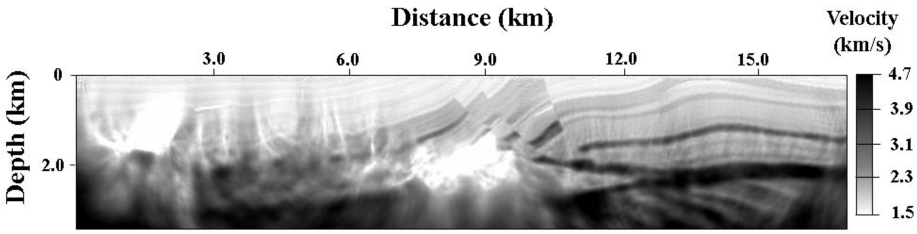


Figure 5-17 Descent directions of the P-wave velocity obtained at the 100th iteration for data with *white random noise* generated using the *Ricker wavelet* (a) without and (b) with the denoise function.

(a)



(b)

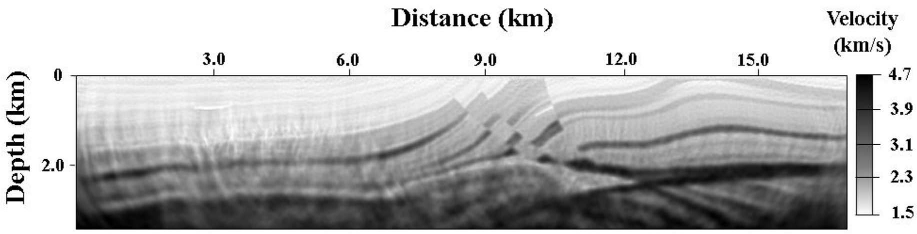


Figure 5-18 Recovered P-wave velocity structures at the 200th iteration (a) without and (b) with the denoise function for data obtained by the *Ricker wavelet* and included by *white random noise*.

5.4.2 Effect of Noise Spectrum

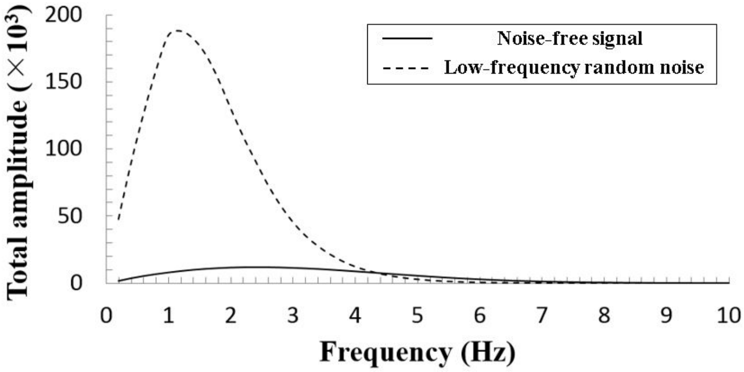
Generally, the random noise, which is produced by the ambient ground motion from natural sources such as wind and ocean microseisms, is dominant at the low-frequency band, as discussed above. Although the dominant frequency of noise changes depending on the site characteristics, we assume that the low-frequency random noise is dominant below 2 Hz. Figure 5-19 shows the spectra of noise-free seismic signal and the low-frequency random noise, which are used for the observed data.

Compared to the conventional method, the denoise function works quite well during the inversion and provides a reasonable solution, as shown by the descent directions (Figure 5-20) and the recovered P-wave velocities (Figure 5-21).

The high-frequency dominant random noise is generated after the ambient ground motion from cultural sources, such as moving vehicles or heavy machinery near the road and production well. Although this type of cultural noise can be transient, and thus dependent on time, we add the high-frequency random noise to the whole shot gathers, but the noise amplitudes are randomly determined. Figure 5-22 shows the amplitude spectra of the noise-free signal and random noise and the variations of the denoise function during the FWI. We observe that the denoise function well reflects the S/N ratio of the observed data. Using this low-pass filter-like denoise function, the calculated descent direction (Figure 5-23) is nearly noise-free, and the recovered P-wave velocities are in good agreement with the true P-wave velocities.

Based on these results for my four numerical examples, I am convinced that the conventional frequency-domain FWI does not properly cope with the spectral variations of the S/N ratio, which are determined by the noise and source spectra. In contrast, the denoise function reshapes the gradient (or descent) spectrum by giving less weight to the noise-dominant single frequency gradient directions, depending on the S/N ratio of the observed data, and provides a reasonable solution in the presence of the random noise with little human intervention.

(a)



(b)

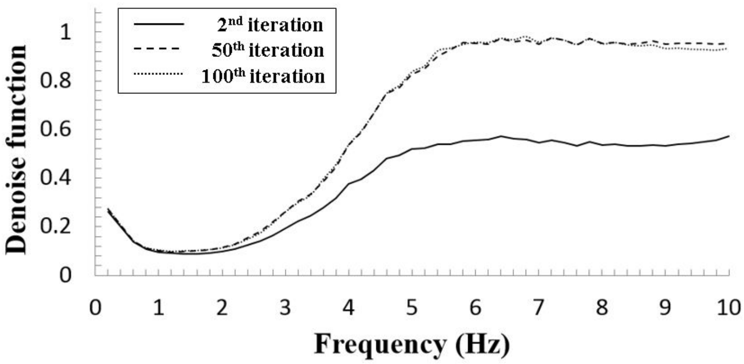
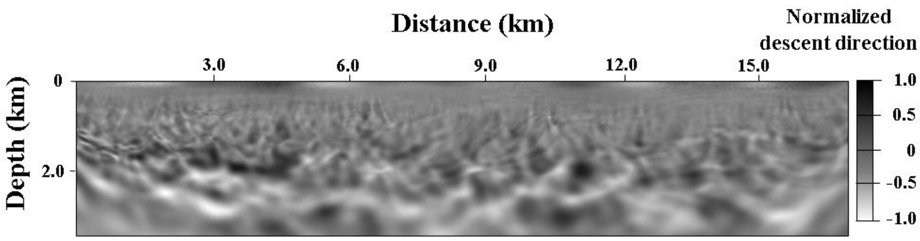


Figure 5-19 (a) Amplitude spectra of the *low-frequency random noise* (dashed line) and noise-free seismic signal generated using the *first derivative of the Gaussian function* (black line), summed over shots and receivers, and (b) spectra of the denoise functions (e=1) at the 2^{nd} (solid black line), 50^{th} (dashed black line), and 100^{th} (dotted black line) iterations.

(a)



(b)

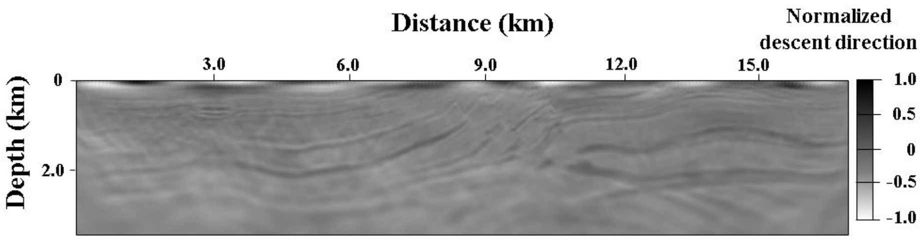
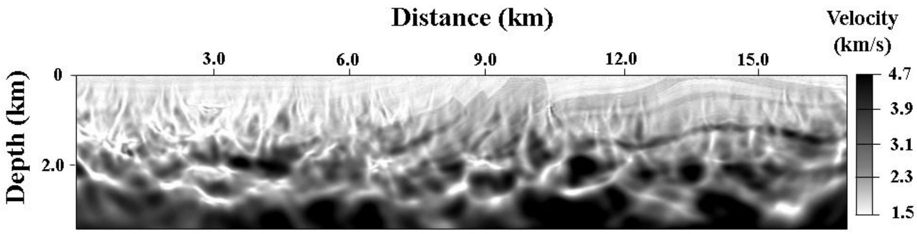


Figure 5-20 Descent directions of the P-wave velocity obtained at the 100th iteration (a) without and (b) with the denoise function for data with *low-frequency random noise* generated using the *first derivative of the Gaussian function*.

(a)



(b)

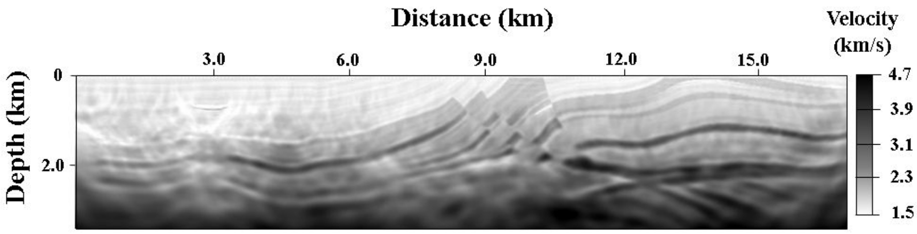
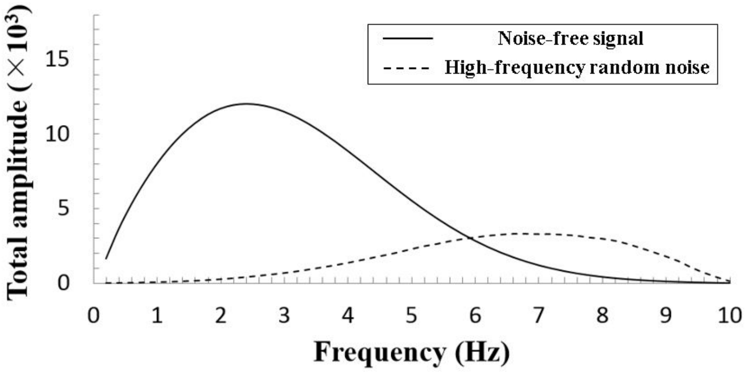


Figure 5-21 Recovered P-wave velocity structures at the 200^{th} iteration (a) without and (b) with the denoise function for data with *low-frequency random noise* obtained using the *first derivative of the Gaussian function*.

(a)



(b)

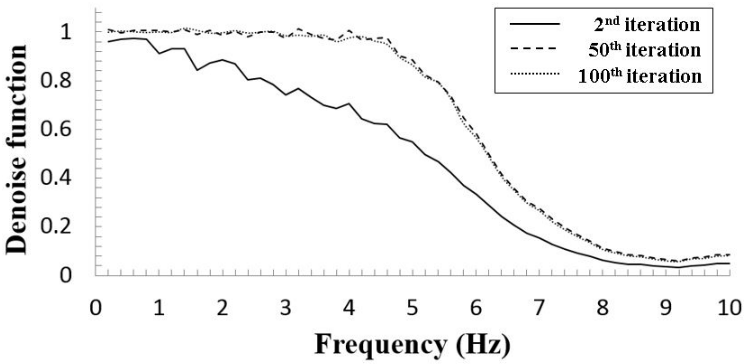
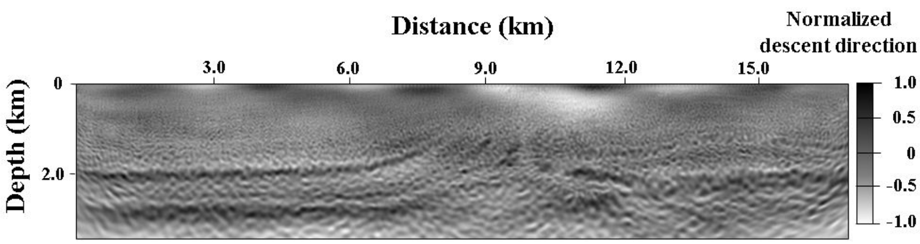


Figure 5-22 (a) Amplitude spectra of the *high-frequency random noise* (dashed line) and noise-free seismic signal generated using the *first derivative of the Gaussian function* (black line), summed over shots and receivers, and (b) spectra of the denoise functions (e=1) at the 2^{nd} (solid black line), 50^{th} (dashed black line), and 100^{th} (dotted black line) iterations.

(a)



(b)

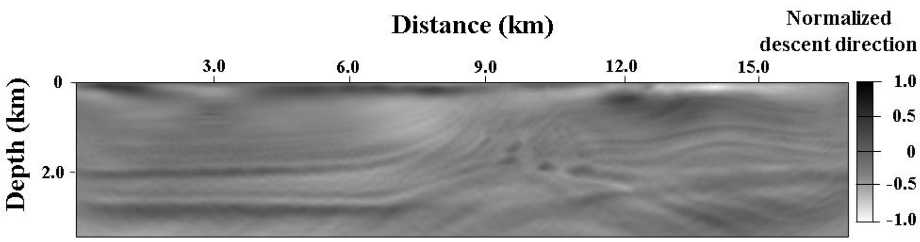
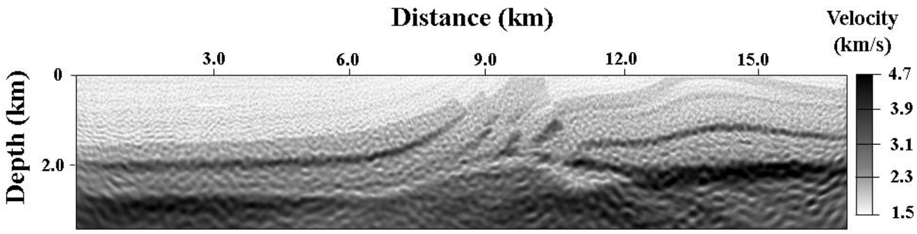


Figure 5-23 Descent directions of the P-wave velocity obtained at the 50th iteration (a) without and (b) with the denoise function for data with *high-frequency random noise* generated using the *first derivative of the Gaussian function*.

(a)



(b)

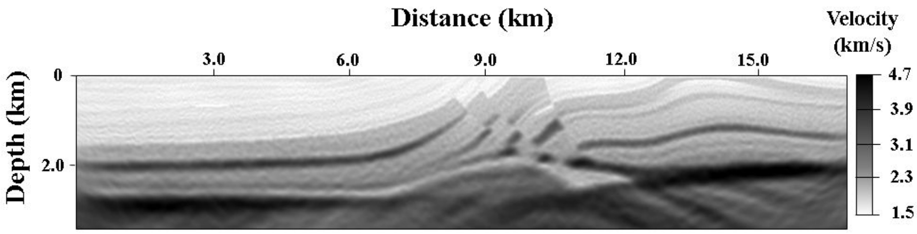


Figure 5-24 Recovered P-wave velocity structures at the 130th iteration (a) without and (b) with the denoise function for data with *high-frequency random noise* obtained using the *first derivative of the Gaussian function*.

Chapter 6. Depth Scaling Scheme

As I discussed in the beginning of Chapter 5, we also consider the noise distribution along the arrival time or depth axis to reduce the influence of the noise. Therefore, in addition to the spectral filtering scheme using the denoise function, I also suggest to use the Levenberg-Marquardt method with flexible damping factor. Because noise can hide weak reflections in certain time windows, as mentioned above, FWI might not provide reasonable solutions for certain depths. Wang and Rao (2009) used the layer-stripping scheme, in which they divided the model into two parts and then sequentially inverted the shallow and deep parts. Once the shallow part was recovered, it was not updated while the deeper part was updated. When they recovered the deeper layer, they replaced the early-time data with synthetic data generated based on the inverted model for the shallow part. Through these processes, Wang and Rao (2009) successfully applied the layer-stripping method to the Marmousi model and a field data set. However, one may hesitate to use the layer-stripping method because it requires additional effort due to the aforementioned additional processes. Moreover, one might obtain different final values near the boundary between the two recovered parts because they are inverted separately. As an alternative to the layer-stripping method, I suggest a depth scaling strategy in which I apply the constraint (i.e., the damping factor or Lagrange multiplier) of the damped least-squares method (i.e., the Levenberg-Marquardt method).

6.1 Reconsideration of Levenberg-Marquardt Method

At first, I discuss the effects of the damping factor in the Levenberg-Marquardt method (eq. (2-6)). Due to geometrical spreading effects, in the steepest descent method, the gradient for the deeper structures is small, which makes it hard to recover deeper structures. To compensate for this issue, we generally precondition the gradient direction by using the Hessian matrix, which plays a role in compensating for geometrical spreading effects (Choi et al. 2008). In this paper, I refer to this ability as the 'scaling power of the Hessian matrix'. The success of the depth scaling strategy depends on the scaling power of the Hessian matrix, i.e., the depth scaling strategy would not be working properly if we were to employ a poorly calculated Hessian matrix. Oh and Min (2012) compared various approximated versions of the Hessian matrix in terms of the matrix's scaling power and showed that the pseudo-Hessian (Shin et al., 2001b) matrix has a weaker scaling power than the new pseudo-Hessian matrix.

Following Shin et al. (2001b), we can define the damping factor as follows:

$$\beta = \frac{k}{100} \times \max|\mathbf{H}|, \qquad (6\text{-}1)$$

where $\max|\mathbf{H}|$ and k indicate the maximum value of the Hessian matrix and its percentage, respectively. Shin et al. (2001b) noted that the diagonal of the approximate Hessian matrix acts like an automatic gain control (AGC) for the subsurface image and that small damping factors over-amplify deeper structures. My depth scaling strategy uses this feature of the Hessian matrix.

However, because the computation cost of calculating the approximate Hessian matrix is too great, we use the new pseudo-Hessian matrix.

Figure 6-1 shows the synthetic graben structure, which is one of the oil-promising geologic structures and Table 6-1 shows the inversion parameters used during the FWI. Figure 6-2 shows the descent directions preconditioned by the diagonals of the approximate Hessian matrix, which are constrained by various damping factors. As Lines and Treitel (1984) explained, the main advantage of the Levenberg-Marquardt method is that, it provides good convergence in the early stage and efficiency in the late stage of the inversion by decreasing the damping factor as iteration proceeds. Corresponding to their statements, when k is 0 %, the descent direction provides highly resolved image in the deeper structures (but not accurate) and provides the same as that of the linear least-squares inversion. When k is large, the deep structures in the gradient images are poorly defined because the geometrical spreading effect is not properly compensated due to the constraint of the large damping factor. Using this behavior of the damping factor, we can control the parameter-recovering depth in the inversion. The parameters for shallow parts can be restored using large damping factors in the early stage and, in the later stage, deeper parts can be reconstructed primarily using small damping factors. I notice that the variable damping factor acts like a depth filter for gradients. In other words, we can hold the gradients for some parts for which we do not want to update the parameters, as in the layer-stripping scheme.

To determine the damping factor in a flexible manner, we can choose empirical values, as Oh and Min (2012) did. However, in this study, the damping factor is determined based on the convergence of the inversion process. The damping factor is set to 1 % of the maximum value of the new pseudo-Hessian matrix at the beginning of the inversion. Whenever the inversion process diverges (i.e., the trend for the RMS error increases), the damping factor decreases to 50 % of its former value (e.g., Lines and Treitel, 1984).

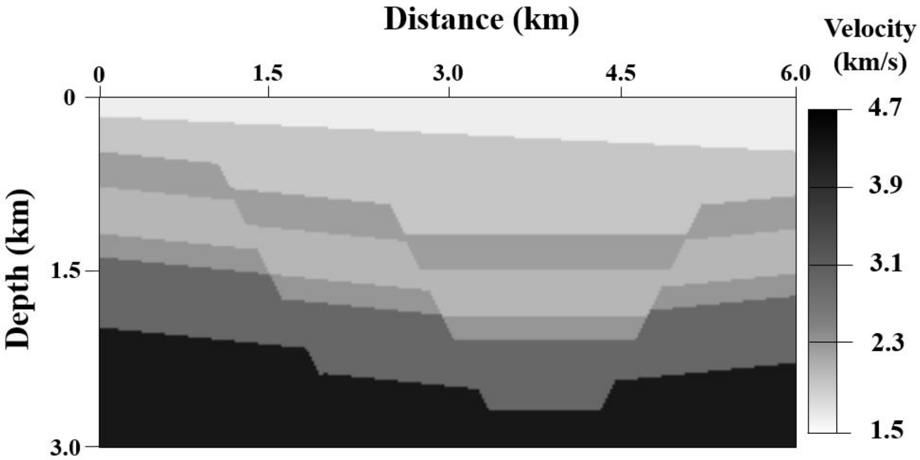


Figure 6-1 P-wave velocity of the synthetic graben structure. The true S-wave velocity and density are determined by the fixed Poisson's ratio (0.25) and equation of Gardner et al. (1974), respectively.

Table 6-1 Parameters used for the synthetic graben model

Dimension	No. of shots	No. of receivers	Interval of shots	Interval of receivers	Recording time	Maximum Frequency	Minimum Frequency
6.0 km × 3.0 km	150	301	0.04 km	0.02 km	4 sec	10 Hz	0.25 Hz

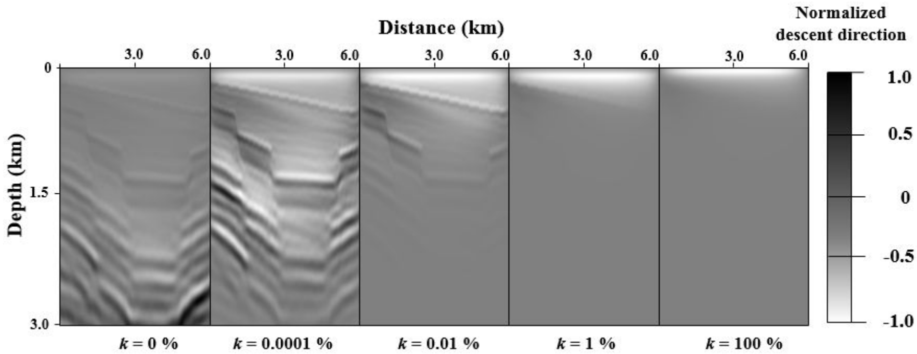


Figure 6-2 Descent directions obtained at the 1st iteration using various damping factors for the synthetic graben model with a known source wavelet when the diagonals of the approximate Hessian matrices are used as a pre-conditioner.

6.2 Advantages of the Depth Scaling Scheme

In this subchapter, I will discuss other two advantages of the depth scaling scheme. One advantage is that the Levenberg-Marquardt method is good for the mono-frequency FWI and is also good when the particular frequency bands are severely weighted during the inversion. The other advantage is that the depth scaling method can enhance the robustness of the FWI to the random noise. I verify these advantages comparing the depth scaling scheme with the fixed damping scheme as Sheen et al. (2006) did.

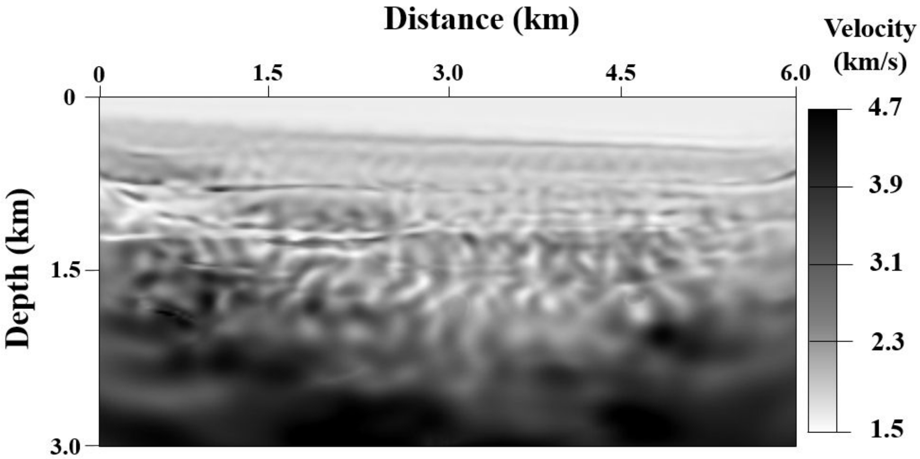
6.2.1 Advantage for mono-frequency FWI

In the mono-frequency FWI or the weighted-frequency FWI like the FWI with the weighting method described in Chapter 4, the inversion results can contain lot of artifacts because the Jacobian, that is a tool for resolving subsurface structures, behaves like monochromatic waves as shown in Figure 4-6. For example, as shown in Figure 3-16a, the mono-frequency descent directions fluctuates around boundaries of the subsurface layer and yields distorted results. However, if we apply the depth scaling strategy using the Levenberg-Marquardt method described previous chapter, these mono-frequency-induced artifacts can be reduced because the damping factor damps out the gradient direction below a certain depth as shown in Figure 6-2. Figure 6-3 shows the recovered P-wave velocity obtained by the mono-frequency FWI (3 Hz) with the fixed damping scheme and the depth scaling scheme. As the results show, the fixed damping scheme suffers from the wavelet-induced artifacts, on the other hand, the depth scaling scheme provides well described background velocity with less artifacts. Because the observed data generally are band-limited due to the noise or limitations of the measurement, when we perform the frequency-domain FWI, the depth scaling scheme should be employed for a stable FWI.

6.2.2 Advantage for random noise-included data

In conventional FWI with a fixed damping scheme, as performed by Sheen et al. (2006), even when the shallow parts have already been recovered and primarily the deeper parts are being inverted, the gradients for the shallow parts are still updated with large values. For noisy data, this procedure may play a role in degrading the inversion results for the shallow parts. This problem also can be compensated for by introducing the depth scaling scheme with a variable damping factor. To verify the noise suppression of the depth scaling scheme, I perform the FWI with random noisy data, in which the spectral S/N ratio (maximum amplitude of signal over maximum amplitude of noise) is 5. Figure 6-4 shows the P-wave velocity structures obtained by the FWI for the full bandwidth of the observed data with a fixed damping and depth scaling schemes. In the fixed damping scheme, the shallower structures suffer from the noise-induced artifacts, on the other hand, the damping scaling scheme provides relatively robust inversion results. From this numerical test, I notice that the depth scaling scheme should be applied with the denoise function (Chapter 5) for more robust FWI to the random noise.

(a)



(b)

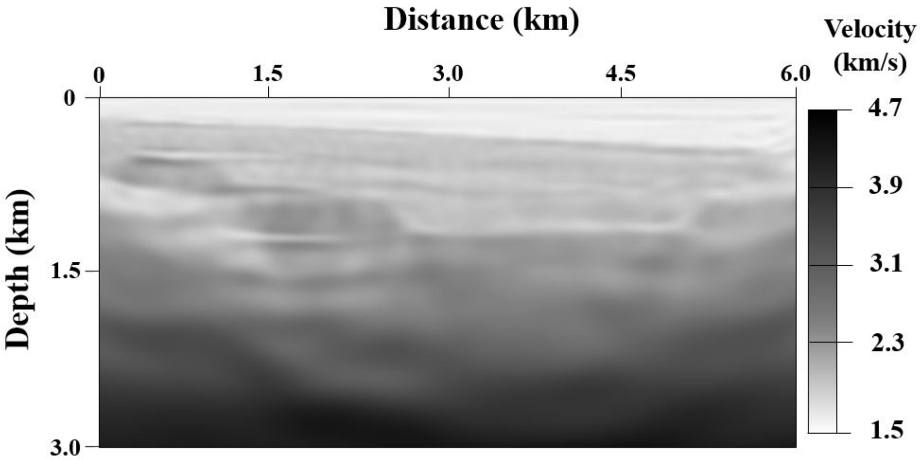
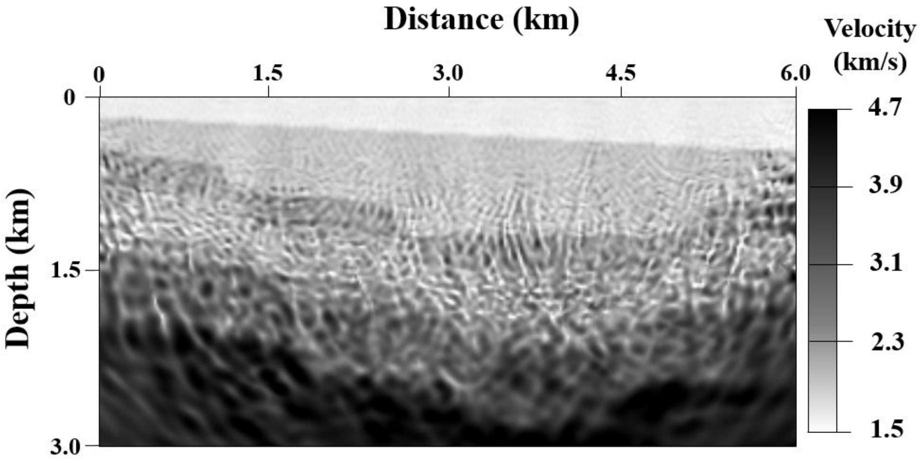


Figure 6-3 The recovered P-wave velocity structures obtained by the FWI for the *mono-frequency noise-free data* (3 Hz) with (a) the fixed damping scheme and (b) the depth scaling scheme

(a)



(b)

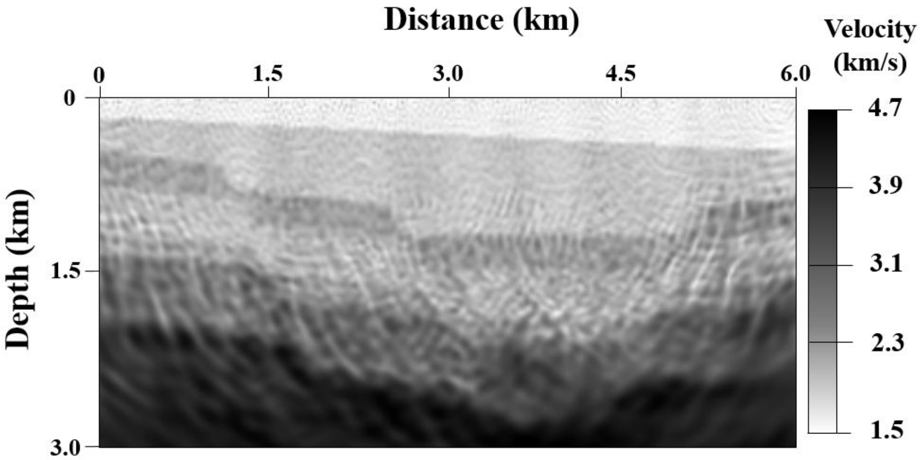


Figure 6-4 The recovered P-wave velocity structures obtained by the FWI for *the random noise included synthetic data* with (a) the fixed damping scheme and (b) the depth scaling scheme.

Chapter 7. Numerical Examples

From Chapters 3 to 6, for the multi-parametric FWI of the banded seismic data, four new methods have been introduced, those are the parameterization using the Poisson's ratio for accurate multi-parameter extraction from the seismic data, the spectral weighting method to control the spatial resolution of the gradient direction, the spectral filtering method using the denoise function to filter out noisy components of data and the depth scaling scheme using the Levenberg-Marquardt method for the stable FWI. In this chapter, as the last chapter for the new methodology, we discuss how new inversion strategy improves the inversion results for the multi-parameter FWI with the random noise-included data. To do that, the elastic FWIs for the elastic Marmousi-2 model (Figure 3-19) are performed with the same inversion setting (Table 3-3) and initial model (Figure 3-20) as those used in Chapter 3. For these numerical examples, the source-wavelet inversion is also performed during the FWI (Song et al., 1995).

Figure 7-1 shows synthetic seismograms of the observed data with random noise. The synthetic observed data are simulated using the time-domain 4^{th}-order staggered grid finite-difference method (Graves 1996). The random noises are generated in the frequency domain and designed so that vertical motions of random noise are much larger than horizontal motions. Figure 7-2a shows the spectra of the noise-free seismic signal and random noise. Based on the previous studies about the ambient ground motions (refer Chapter 5.2 and

Figure 5-3), the low-frequency components of random noise below at 2.5 Hz are made after the microseism induced by large-scale motions of the oceanic flow and the middle-frequency components around at 5 Hz are generated after high-speed motions of wind. Some cultural random noises at higher-frequency bands are also considered.

Before applying new inversion strategy, it would be important to cut off severely noise-contaminated components of the observed data in advance, because the Type-II weighting factor is still sensitive to the noise spectrum although it is designed to reduce influences of noise spectrum. Because, in addition, it is easy to identify these severely noise-contaminated components of the observed data by the human eye, there is no reason to participate these severely distorted data in the FWI. For this reason, frequency components below 2 Hz and above 7 Hz for all the shot gathers are excluded during the FWI.

Figure 7-3 shows the recovered structures obtained using the IPG-II (seismic velocities and density) without the new inversion strategy. In this example, we can notice that each parameter has different sensitivity for the random noise. The P-wave velocity structure is severely distorted by some noise-induced artifacts, because, as I mentioned in Chapter 3.3.3, the absence of P-S, S-P and S-S scattered waves in the virtual source for the P-wave velocity makes the FWI more sensitive to noises. On the other hand, the recovered S-wave velocity structure is relatively less contaminated by random noise.

As shown in Figure 7-4, we notice that the new parameterization using the Poisson's ratio (IPG-III) can improve the robustness of the FWI to the random noise, because, in the IPG-III, all the parameters generates P-P, P-S, S-P and S-S scattered waves. However, there are still lots of noise-induced artifacts in the recovered structures, because noisy components of the observed data still equally contribute to the banded descent direction. Figure 7-5 shows the recovered structures when I additionally use the Type-II weighting factors. Because the Type-II weighting factor is calculated based on the admittance spectrum (exactly, admittance spectrum over angular frequency: refer Chapter 4.3.2), that is inversely proportional to the angular frequency, the noise-contaminated low-frequency components are heavily weighted during the FWI. In addition, due to the very low S/N ratio around 5 Hz, the weighting factor provides large weights to noisy-contaminated 5 Hz component as shown in Figure 7-2b. For these reasons the recovered structures severely contaminated by low-frequency components of the random noise. This also can be a problem of the conventional frequency marching method, in which we perform the FWI from low-frequency to high-frequency components. In the frequency marching method, if we fail to select good (less noisy) starting frequency, the final inversion results will suffer from noise-induced artifacts. In addition, like this numerical example, when the noises are scattered over frequencies, it is not easy to march the FWI to higher frequency.

Figure 7-6 shows the inversion results when the denoise function is additionally applied during the FWI. For great effects of the noise reduction, the large value of the control factor, *e*, is chosen at a cost of spatial resolution. Because the denoise function reduces the contribution of noisy components, which have relatively low S/N ratio, we can notice that the recovered structures are greatly improved. However, there are still some noise-induced artifacts around the salt body located at the left side of the model.

Figure 7-7 shows the recovered structures obtained by entire new inversion strategies using the IPG-III, Type-II weighting factor, denoise function and depth scaling scheme. We can notice that, with new inversion strategy, the subsurface structures are well described with less noise-induced artifacts and all the parameters are well estimated in some extent compared with previous results. However, we notice that there are some artifacts near the surface, which originate from the inaccuracy of source-wavelet-estimation. As shown in Figure 7-8, due to noise-components of observed data, the source wavelet is poorly estimated, particularly for the phase spectrum. This result indicates that, for a successful FWI using the new inversion strategy with real field data, the new inversion strategy should be applied with robust source-estimation technique or good initial source wavelet.

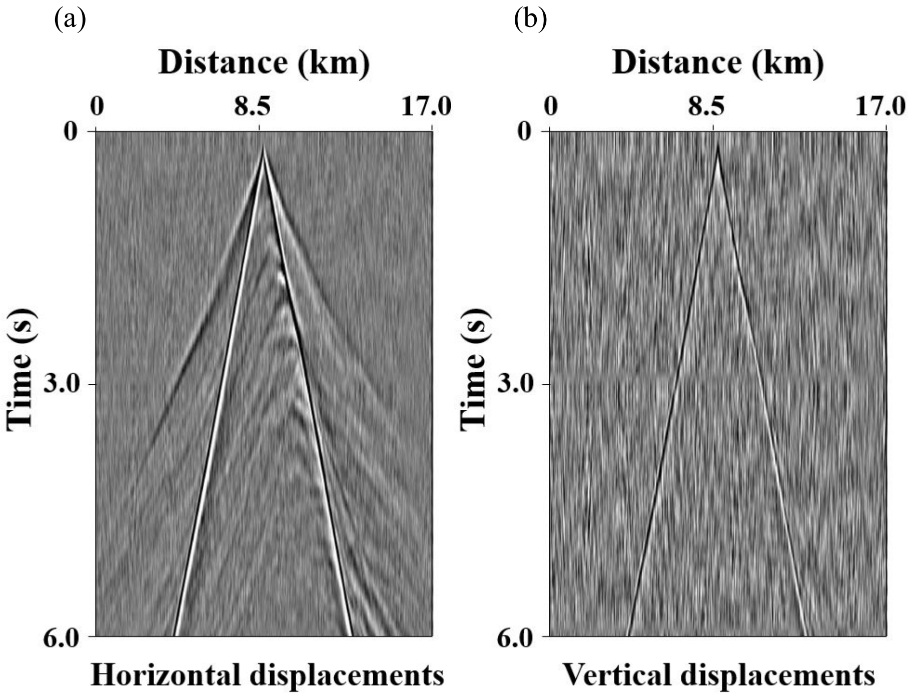
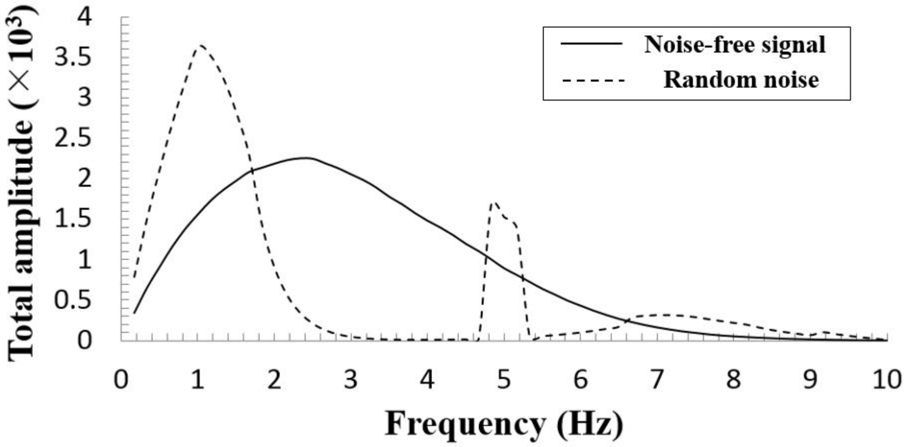


Figure 7-1 Synthetic observed data contain random noise: (a) Horizontal and (b) vertical displacements

(a)



(b)

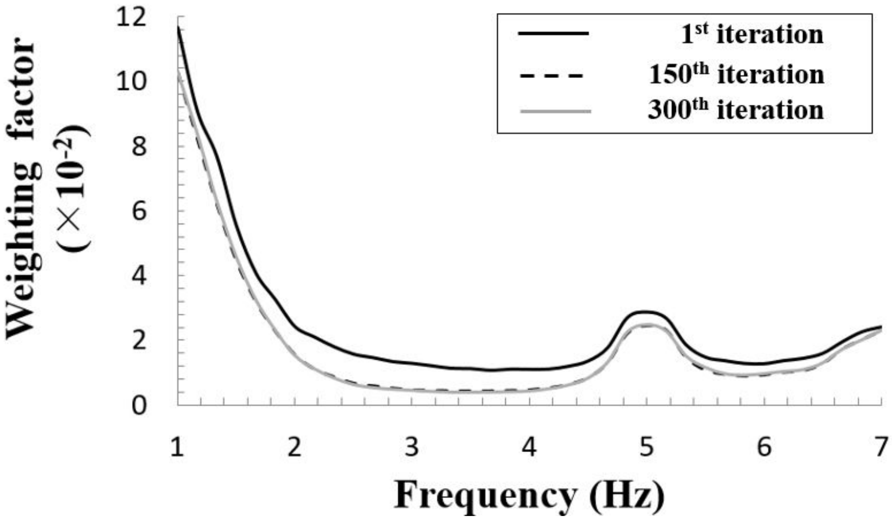
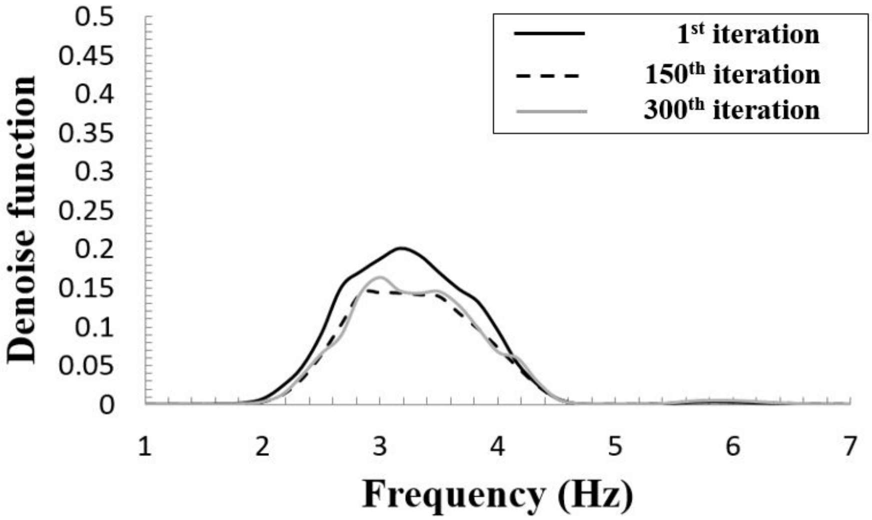


Figure 7-2 (a) The spectra of the noise-free signal and the random noise, (b) variations of the weighting factor, (c) variations of the denoise function and (d) total weighting function, which is the multiplication of the weighting factor and the denoise function

(c)



(d)

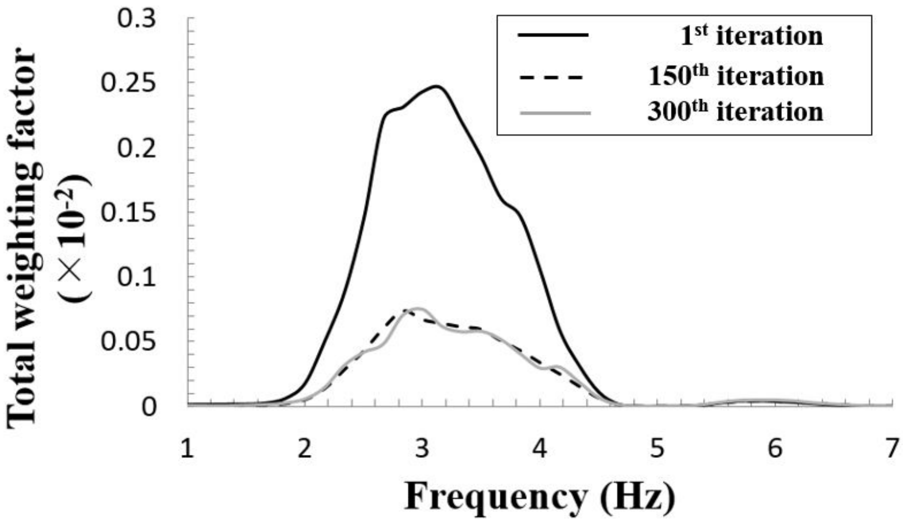


Figure 7-2 (Continued)

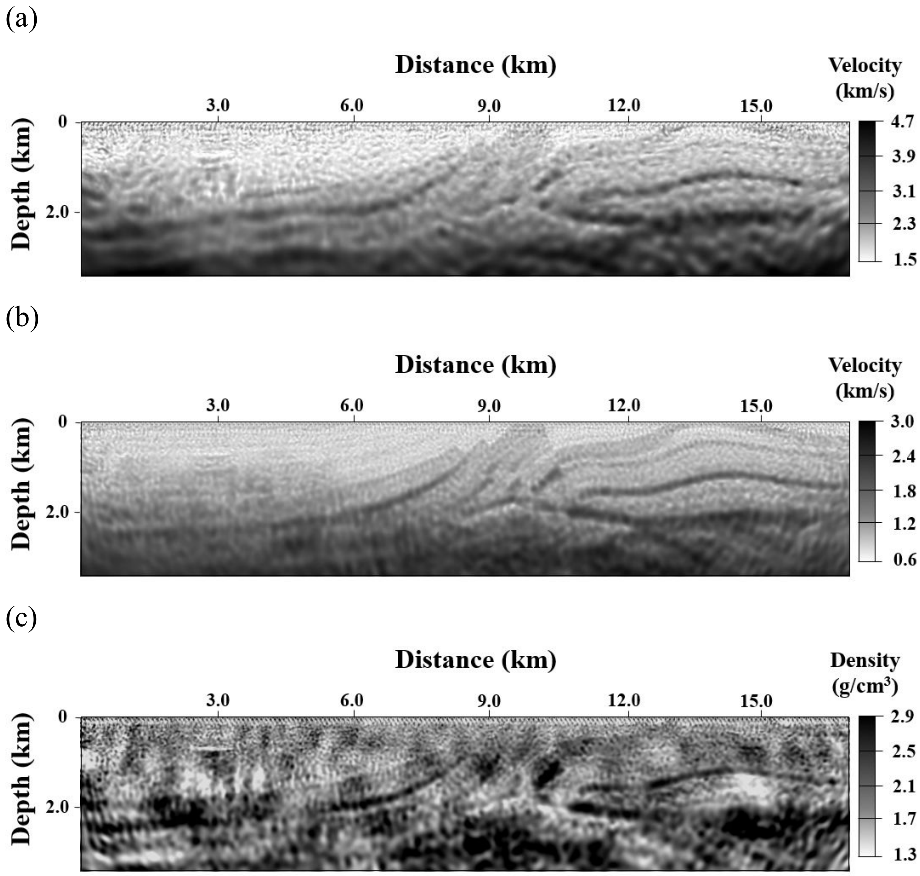


Figure 7-3 Recovered (a) P-wave velocity, (b) S-wave velocity and (c) density with the *IPG-II* and without the weighting method, denoise function and depth scaling scheme

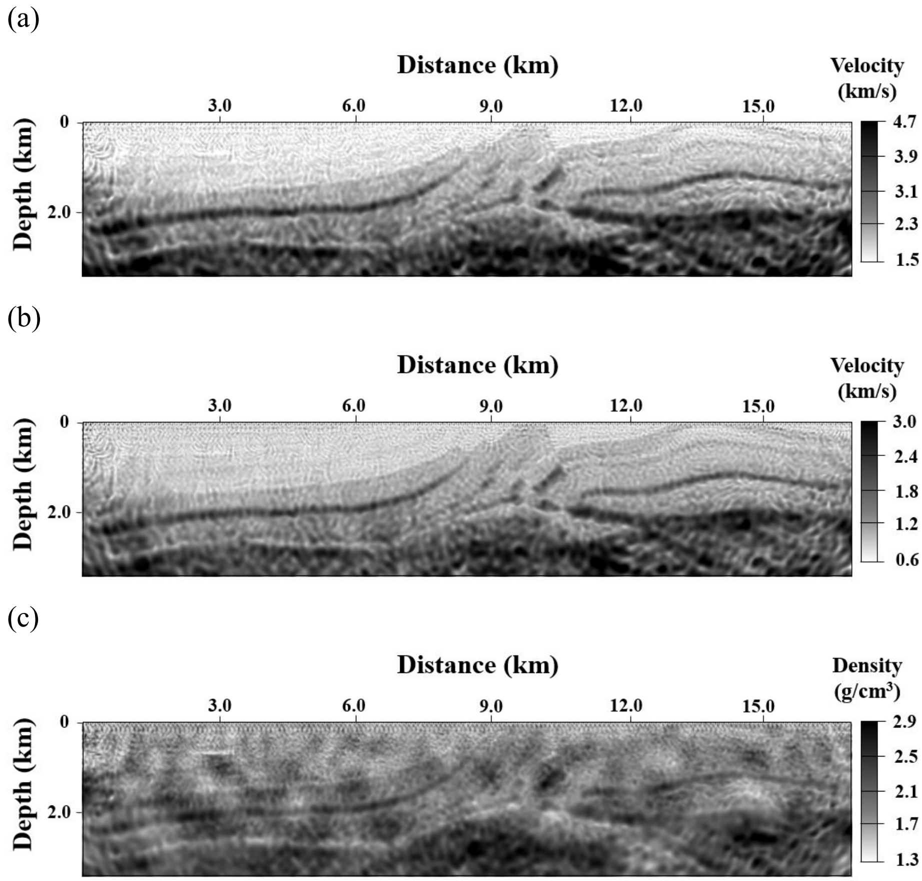


Figure 7-4 Recovered (a) P-wave velocity, (b) S-wave velocity and (c) density with the *IPG-III* and without the weighting method, denoise function and depth scaling scheme

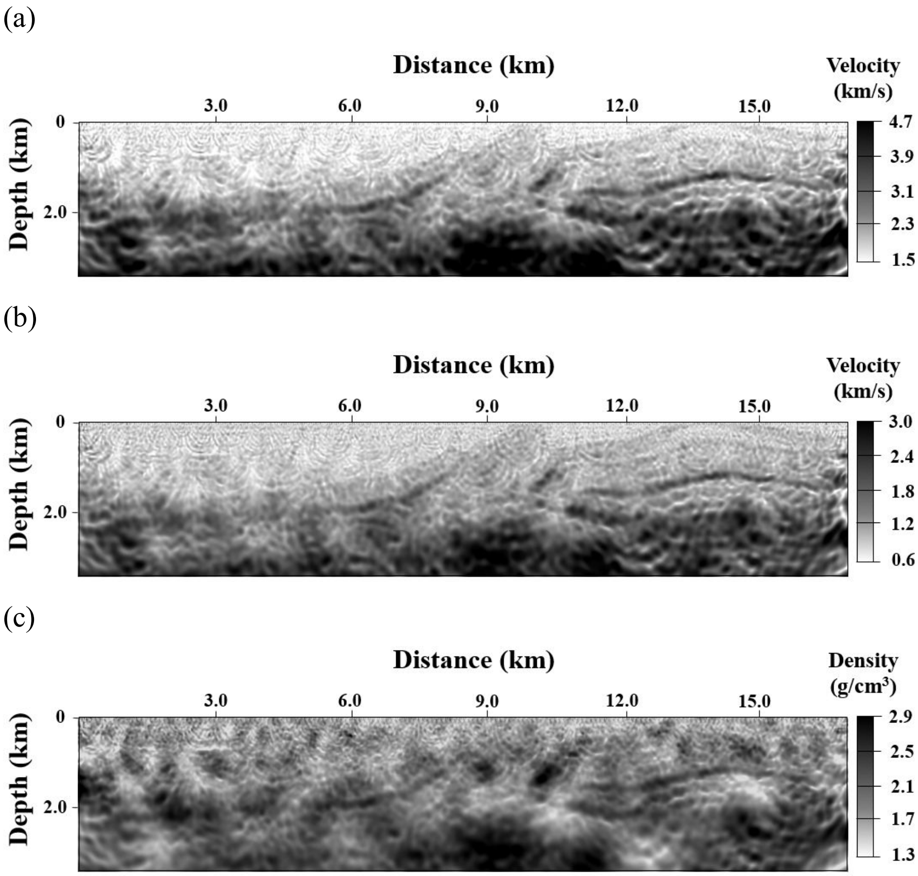
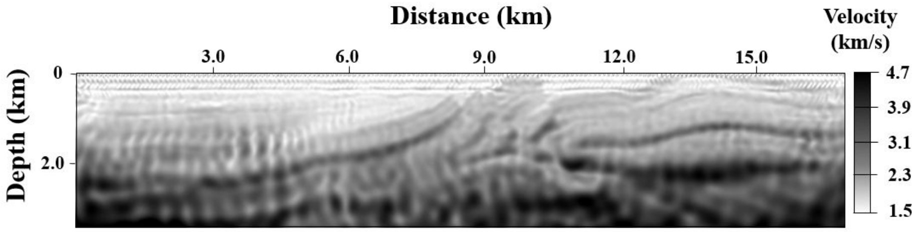
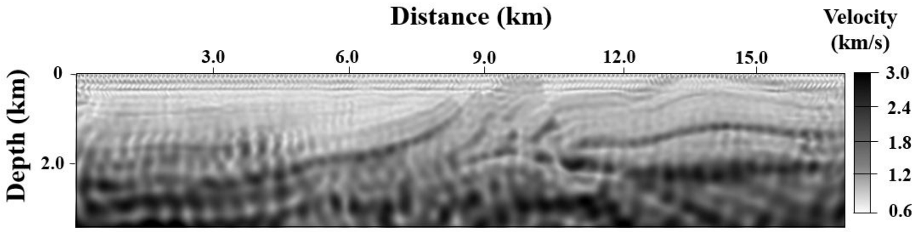


Figure 7-5 Recovered (a) P-wave velocity, (b) S-wave velocity and (c) density with the *IPG-III* and *Type-II weighting factor* without the denoise function and depth scaling scheme

(a)



(b)



(c)

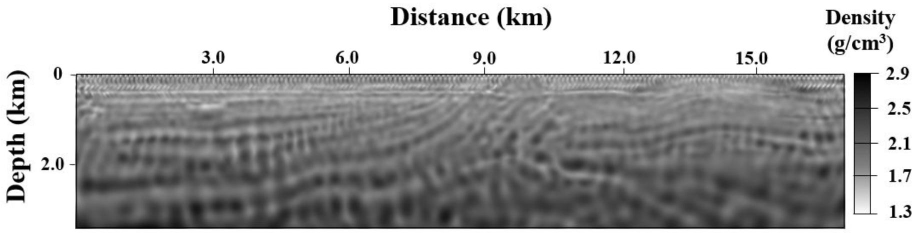
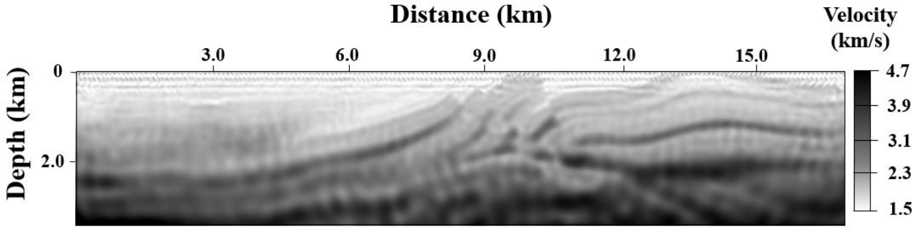
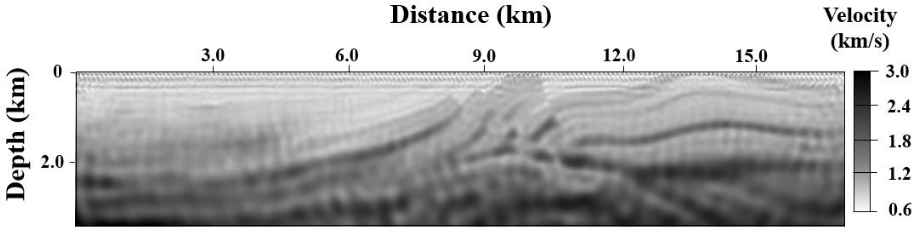


Figure 7-6 Recovered (a) P-wave velocity, (b) S-wave velocity and (c) density with the *IPG-III*, *weighting method* and *denoise function (e=5)* without the depth scaling scheme

(a)



(b)



(c)

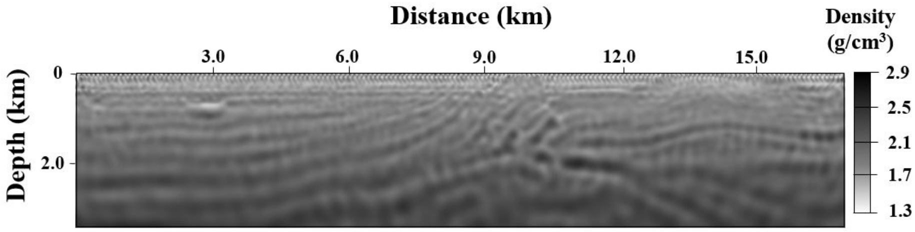
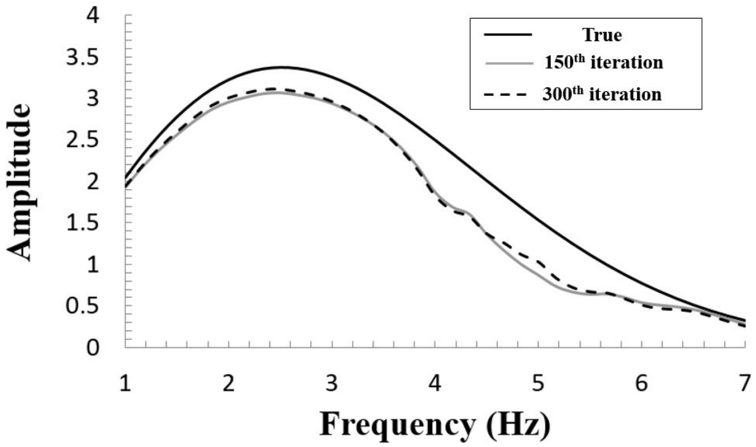


Figure 7-7 Recovered (a) P-wave velocity, (b) S-wave velocity and (c) density with the *IPG-III*, *weighting method*, *denoise function (e=5)* and *depth scaling scheme*

(a)



(b)

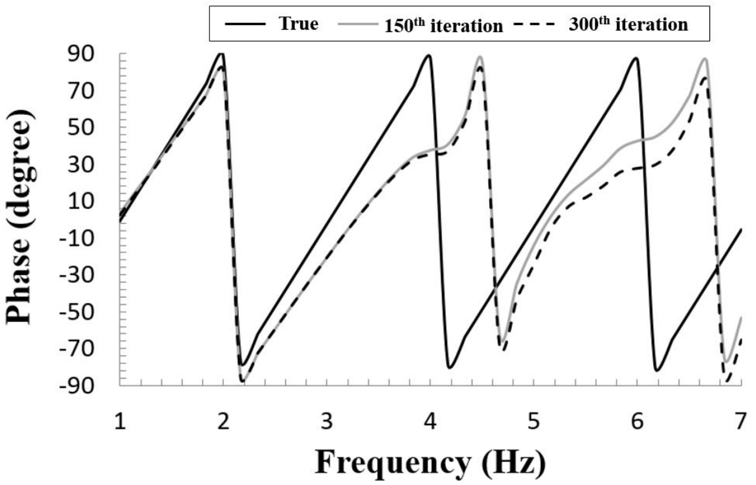


Figure 7-8 (a) The amplitude and (b) phase spectra of true source and estimated source wavelets.

Chapter 8. Discussion & Conclusions

Through the paper, the new inversion strategy is proposed to extract multi parameters for the seismic data with wide bandwidth. The new inversion strategy contains (1) the new parameterization using the Poisson's ratio, (2) the weighting factor to control the spatial resolution of the descent direction depending on the thickness of subsurface layers, (3) the denoise function which damps out severely noise-contaminated frequency components of the observed data and (4) the depth scaling scheme using the Levenberg-Marquardt method for more stable frequency-domain FWI.

To find the best parameterization, Chapter 3 described the mechanism of the virtual source with smaller concept, which is the basis virtual source. Each basis virtual source can be described by different moment tensors and partially or totally contributes to determine the radiation patterns of each model parameter. From the results of analysis, the new parameterization using the Poisson's ratio was proposed, which provides reasonable seismic velocities and density.

To overcome fixed weighting effects of conventional scaling methods, Chapter 4 introduced the weighting method in which the weighting factors are automatically calculated from the source-deconvolved backpropagated wavefields and approximately control the spatial resolution depending on the thickness of subsurface layers.

In the Chapter 5, the spectral filtering scheme for the gradient direction was suggested using the denoise function, which is also automatically calculated by summing all the shot gathers and, in Chapter 6, the depth scaling scheme using the Levenberg-Marquardt method was proposed, in which the damping factor is set to decrease when the error diverges.

The final numerical examples (Chapter 7) showed that the new inversion strategy provides reasonable inversion results for all the parameters, in the case when the observed data are severely distorted by random noise (although synthetic observed data are used), with only one step and less human interventions.

Considering that the conventional approaches require too many stages and, at end of each step, the FWI process is controlled by human interventions, the new inversion algorithm suggested in this paper will be more reasonable and efficient way because only thing we need to decide is choosing the control factor of the denoise function to control the filtering ability. However, there are still some problems to be resolved for real data application. The real data obtained by a land survey include some coherent noises and dispersive Rayleigh waves, which cannot be simulated through the numerical modeling. To overcome these limitations, the new inversion strategy should work with some coherent noise removal techniques. In addition, because the spectral weighting and filtering strategies of the new strategy are based on the exact source-spectrum, the new inversion strategy should be applied with exact source-estimation techniques or good initial source wavelet. Moreover, because the low-frequency components of real data are generally contaminated by noise, they cannot be used during the FWI. For this reason, it would be better if the new inversion strategy is applied with good initial guesses obtained from Laplace-domain FWI (Shin and Cha, 2008) or the frequency down-shifting technique (Warner et al., 2013).

Bibliography

Abma, R.L., 1995, Least-squares separation of signal and noise using multidimensional filters, Ph.D. Thesis, *Stanford University*.

Abubakar, A., Pan, G., Li, M., Zohang, L., Habashy, T.M. and van den Berg, P.M., 2011, Three-dimensional seismic full-waveform inversion using the finite-difference contrast source inversion method, *Geophysical Prospecting*, **59**, 874-888.

Aki, K. and Richards, P.G., 1980, Quantitative seismology, *Freeman*, San Francisco.

Aravkin, A., van Leeuwen, T., and Herrmann, F., 2011, Robust full-waveform inversion using the Student's t-distribution, *81st Annual International Meeting, SEG, Expanded abstracts*, 2201-2204.

Bard, P.Y., Bonnefoy-Claudet, S. and Cotton, F., 2003, SESAME: Site effects assessment using ambient excitations, *Final report, WPo8, Nature of noise wavefield, European Commission*, July 2003, http://sesame-fp5.obs.ujf-grenoble.fr/index.htm.

Bécache, E., Fauqueux, S. and Joly, P., 2003, Stability of perfectly matched layers, group velocities and anisotropic waves, *Journal of Computational Physics*, **188**, 399-433..

Berkhout, A. J., 1985, Seismic migration, imaging of acoustic energy by wave field extrapolation: volume 14A of Developments in sold earth Geophysics, *Elsevier Scientific*, New York.

Bérenger, J. P., 1994, A Perfectly Matched Layer for the absorption of electromagnetic waves, *J. Comput. Phys*., **114**, 185-200.

Bohren, C.F. and Huffman, D., 1983, Absorption and scattering of light by small particles, *John Wiley*, New York.

Brossier, R., Operto, S., and Virieux, J., 2009a, Two-dimensional seismic imaging of the Valhall model form synthetic OBC data by frequency-domain elastic full-waveform inversion, *79th Annual International Meeting, SEG, Expanded abstracts*, 2293-2297.

Brossier, R., Operto, S., and Virieux, J., 2009b, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics*, **74**, WCC63-WCC76.

Brossier, R., Operto, S. and Virieux, J., 2010, Which data residual norm for robust elastic frequency-domain full waveform inversion?, *Geophysics*,

75, R37-R46.

Bube, K.P., and Langan, R.T., 1997, Hybrid l_1/l_2 minimization with applications to tomography, *Geophysics*, **62**, 1183-1195.

Bunks, C., Saleck, F.M., Zaleski, S. and Chavent, G., 1995, Multiscale seismic waveform inversion, *Geophysics*, **60**, 1457-1473.

Burridge, R. Hoop, M.V., Miller, D. and Spencer, C., 1998, Multiparameter inversion in anisotropic elastic media, *Geophysical Journal International*, **134**, 757-777.

Butler, J.M., 1975, A study of seismic road noise, MS thesis, *Georgia Institute of Technology*.

Choi, Y., Min, D.J. and Shin, C., 2008, Frequency-domain elastic full waveform inversion using the new pseudo-Hessian matrix: Experience of elastic Marmousi-2 synthetic data, *Bulletin of the Seismological Society of America*, **98**, 2402-2415.

Choi, Y., and Min, D.J., 2012, Source-independent elastic waveform inversion using a logarithmic wavefield, *Journal of Applied Geophysics*, **76**, 13-22.

Cox, A.J., DeWeerd, A.J. and Linden, J., 2002, An experiment to measure Mie and Rayleigh total scattering cross sections, *American Journal of Physics*, **70**, 620-625.

Fichtner, A., Kennett, B.L.N., Igel, H. and Bunge, H.P., 2009, Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophysical Journal International*, **179**, 1703-1725.

Fletcher, R. and Reeves, C.M., 1964, Function minimization by conjugate gradients, *Computer Journal*, **7**, 149-154.

Gardner, G.H.F., Gardner, L.W. and Gregory, A.R., 1974, Formation velocity and density –the diagnostic basis for stratigraphic traps, *Geophysics*, **39**, 770-780.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A. and Virieux, J., 2013a, Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 1: Sensitivity and trade-off analysis, *Geophysics*, **78**, R81-R105.

Gholami, Y., Brossier, R., Operto, S., Prieux, V., Ribodetti, A. and Virieux, J., 2013b, Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 2: Synthetic and real data case studies from Valhall, **78**, R107-R124.

Graves, R.W., 1996, Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bulletin of the Seismological Society of America*, **86**, 1091-1106.

Guitton, A., 2003, Multiple attenuation with multidimensional prediction-error filters, *73rd International Meeting, SEG, Expanded Abstracts*, 57-75.

Ha, T., Chung, W. and Shin, C., 2009, Waveform inversion using a back-propagation algorithm and a Huber function norm, *Geophysics*, **74**, R15-R24.

Herrmann, F.J., 2010, Randomized sampling and sparsity: Getting more information from fewer samples, *Geophysics*, **75**, WB173-WB187.

Huber, P.J., 1973, Robust regression: Asymptotics, conjectures, and Monte Carlo, *The Annuals of Statistics*, **1**, 792-821.

Jang, U., Min, D.J. and Shin, C., 2009, Comparison of scaling methods for waveform inversion, *Geophysical Prospecting*, **57**, 49-59.

Jeong, W., Lee, H.Y. & Min, D.J., 2012, Full waveform inversion strategy for density in the frequency domain, *Geophysical. Journal International*, **188**, 1221-1242 .

Ji, J., 2012, Robust inversion using biweight norm and its application to seismic inversion: Exploration Geophysics, **43**, 70-76.

Kamath, N. and Tsvankin, L., 2013, Full-waveform inversion of multicomponent data for horizontally layered VTI media, *Geophysics*, **78**, WC113-WC121.

Kanasewich, E.R., 1990, Seismic noise attenuation, *Handbook of Geophysical Exploration, Seismic Exploration*, **7**, Pergamon press.

Kallweit, R.S. and Wood, L.C., 1982, The limits of resolution of zero-phase wavelets, Geophysics, **47**, 1035-1046.

Kim, Y., Cho, H., Min, D.J. and Shin, C., 2011, Comparison of frequency-selection strategies for 2D frequency-domain acoustic waveform inversion, *Pure and applied geophysics*, **168**, 1715-1727.

Köhn, D., Nil, D.D., Kurzmann, A., Przebindowska, A. and Bohlen, T., 2012, On the influence of model parameterization in elastic full waveform tomography, Geophysical Journal International, **191**, 325-345.

Koo, J.M., Lee, H.Y., Jeong, W.D., Min, D-.J. and Yoo, H.S., 2010, Frequency-domain elastic full waveform inversion for VTI media, *72nd Conference & Exhibition, EAGE, Extended Abstracts*, P381.

Lancaster, S. and Whitcombe, D., 2000, Fast-track 'coloured' inversion, *70th Annual International Meeting, SEG, Expanded abstracts*, 1572-1575.

Lazaratos, S. and David, R.L., 2009, Inversion by pre-migration spectral shaping, *79th Annual International Meeting, SEG, Expanded abstracts*, 2383-2387.

Lazaratos, S., Chikichev, I. and Wang, K., 2011, Improving the convergence rate of full wavefield inversion using spectral shaping, 81st *Annual International Meeting, SEG, Expanded abstracts*, 2428-2432.

Lee, H-.Y., Koo, J.M., Min, D-.J., Kwon, B-.D., and Yoo, H.S., 2010, Frequency-domain elastic full waveform inversion for VTI media, *Geophysical Journal International*, **183**, 884-904.

Levenberg, K., 1944, A method for the solution of certain non-linear problems in least squares, *Quarterly of Applied Mathematics*, **2**, 164-168.

Lines, L.R. and Treitel, S., 1984, Tutorial: A review of least-squares inversion and its application to geophysical problems, *Geophysical Prospecting*, **32**, 159-186.

Lines, L.R. and Newrick, R.T., 2004, Geophysical monograph series: number 13 Fundamentals of geophysical interpretation, *Society of Exploration Geophysicists*, Tulsa.

Marquardt, D.W., 1963, An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society for Industrial and Applied Mathematics*, **11**, 431-441.

Métivier, L., Brossier, R., Virieux, J. and Operto, S., 2012, The truncated Newton method for full waveform inversion, *Journal of Physics, conference series*, **386**, 1-6.

Min, D.J., Shin, C. and Yoo, H.S., 2004, Free-surface boundary condition in finite-difference elastic wave modeling, *Bulletin of the Seismological Society of America*, **94**, 237-250.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics*, **52**, 1211-1228.

Nocedal, J., 1980, Updating quasi-Newton matrices with limited storage: Mathematics of Computation, **35**,773-782.

Oh, J.W., and D.J. Min, 2012, Improved pseudo-Hessian for frequency-domain elastic waveform inversion: 74th Conference & Exhibition, EAGE, Extended Abstracts, P017.

Oh, J.W., and D.J. Min, 2013a, Weighting technique using back-propagated wavefields incited by deconvolved residuals for frequency-domain elastic waveform inversion: Geophysical Journal International, **194**, 322-347.

Oh, J.W., and D.J. Min, 2013b, Spectral filtering of gradient using a denoise function for l_2-norm frequency-domain elastic waveform inversion: Geophysical Journal International, **193**, 820-840

Peck, U., 2008, Overview of seismic noise and its relevance to personnel detection, *U.S. Army Corps of Engineers*, ERDC/CRREL TR-08-5.

Plessix, R., 2009, Three-dimensional frequency-domain full-waveform inversion with an iterative solver, *Geophysics*, **74**, WCC149-WCC157.

Plessix, R.E., and Cao, Q., 2011, A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium, *Geophysical Journal International*, **185**, 539-556.

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.

Prieux, V., Brossier, R., Operto, S. and Virieux, J., 2013a, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 1: imaging compressional wave speed, density and attenuation, *Geophysical Journal International*, **194**, 1640-1664.

Prieux, V., Brossier, R., Operto, S. and Virieux, J., 2013b, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 2: imaging compressive-wave and shear-wave velocities, *Geophysical Journal International*, **194**, 1665-1681.

Pyun, S., Son, W., and Shin, C., 2009, Frequency-domain waveform inversion using an l_1-norm objective function, *Exploration Geophysics*, **40**, 227-232.

Reynolds, J.M, 1997, An introduction to applied and environmental geophysics, *John Wiley & Sons*, New York.

Routh, P., Krebs, J., Lazaratos, S., Baumstein, A., Lee, S., Cha, Y.H., Chikichev, I., Downey, N., Hinkley, D. and Anderson, J., 2011, Encoded simultaneous source full-wavefield inversion for spectrally shaped marine streamer data, 81st *Annual International Meeting, SEG, Expanded abstracts*, 2433-2438.

Roden, J.A. and Gedney, S.D., 2000, Convolution PML (CPML): an efficient FDTD implementation of the CFS-PML for arbitrary media, Microwave and Optical Technology Letters, 27, 334-339.

Sheen, D.H., K. Tuncay, C.E. Baag, and P.J. Ortoleva, 2006, Time domain Gauss-Newton seismic waveform inversion in elastic media: Geophysical Journal International, **167**, 1373-1384.

Shin, C., Yoon, K., Marfurt, K.J., Park, K., Yang, D., Lim, H.Y. Chung, S. and Shin, S., 2001a, Efficient calculation of a partial-derivative wavefield using reciprocity for seismic imaging and inversion, *Geophysics*, **66**, 1856-1863.

Shin, C., Jang, S. & Min, D.J., 2001b, Improved amplitude preservation for prestack depth migration by inverse scattering theory, *Geophysical Prospecting*, **49**, 592-606.

Shin, C. and Min, D.J., 2006, Waveform inversion using a logarithmic wavefield, *Geophysics*, **71**, R31-R42.

Shin, C. and Cha, Y.H., 2008, Waveform inversion in the Laplace domain, *Geophysical Journal International*, **173**, 922-931.

Sirgue, L. and Pratt, R.G., 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies, *Geophysics*, **69**, 231-248.

Sneider, R., 2002, General theory of elastic wave scattering, Scattering and Inverse scattering in pure and applied science, *Academic Press*, San Diego, 528-542.

Song, Z.M., Williamson, P.R. and Pratt, R.G., 1995, Frequency-domain acoustic-wave modeling and inversion of crosshole data: Part II–inversion method, synthetic experiments and real-data results, *Geophysics*, **60**, 796-809.

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

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data, *Geophysics,* **51**, 1893-1903.

Tarantola, A., 1987, Inverse problem theory: Methods for data fitting and model parameter estimation, *Elsevier Scientific Publ. Co., Inc*.

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

Wang, Y. and Rao, Y., 2009, Reflection seismic waveform tomography: *Journal of Geophysical Research*, **114**, B03304.

Warner, M., Nangoo, T., Shah, N., Umpleby, A. and Morgan, J., 2013, Full-waveform inversion of cycle-skipped seismic data by frequency down-shifting, *Annual International Meeting, 2013 SEG, Expanded abstracts*, 903-907.

Widess, M., 1973, How thin is a bed?, *Geophysics*, **38**, 1176-1254.

Wu, R., 1984, Seismic wave scattering and small scale inhomogeneities in the lithosphere, *Ph.D. thesis, Dep. Earth, Atmos., Planet. Sci., Mass, Inst. Technol.*, Cambridge, Mass.

Wu, R. and Aki, K., 1985, Scattering characteristics of elastic waves by an elastic heterogeneity, *Geophysics*, **50**, 582-595.

Xu, K., Greenhalgh, S.A., and Wang, M., 2006, Comparison of source-independent methods of elastic waveform inversion, *Geophysics*, **71**, R91-R100.

Yilmaz, O., 2001, Seismic data analysis, Society of Exploration

Geophysicists.

Zienkiewicz, O.C. and Taylor, R.L., 2000, The finite element method, volume 1: The basis, Butterworth-Heinemann.

Appendix A. Particle Motion of Partial Derivative Wavefields

In Appendix A, I briefly discuss some problems, which are derived by the directional measurements, which are the displacement fields in the data acquisition system assumed in this paper.

Because, in the marine data which are generally pressure fields, the measurement is non-directional, the FWI might work well. However, it is doubt that the FWI with directional measurement like displacement fields also work well during the FWI because, particularly for the c_{13} in the VTI elastic FWI, the elastic FWI sometimes failed by parameter-update along a wrong direction of the gradient. To reveal the reason of this problem, I analyze the partial derivative wavefields, which are excited by each basis virtual source for the single perturbation model (Figure A-1).

From Figures A-2 to A-6 show the pattern of scattered wavefields generated by each basis virtual source. As we can see, all the basis virtual source generates P-P, P-S, S-P and S-S particle motions. We notice that, for some groups of basis virtual sources, P-S, S-P or S-S particle motions can be partially canceled (at certain scattering angle) or totally canceled when HB_{xxh}, HB_{xzv}, VB_{zxh} and VB_{zzv} act together. As a result of these interactions, the radiation pattern of virtual source for certain parameter depending on the parameterization can be composed as Appendix B shows.

From Figure A-8 to Figure A-17, I display the partial derivative wavefields excited by each basis virtual source and Figure A-7 shows the residual wavefields obtained by the difference between the single perturbation model and initial model (that is the same as background model). Because, in the time domain, the gradient is defined as a zero-lag cross correlation between residual and partial derivative wavefields, the seismic waves in residual (Figure A-7) and partial derivative wavefields should have the same sign to increase initial model parameters. However, we can notice that particle motions of partial derivative wavefields are very complex as results of the complicated mechanical motions for scattering patterns of each basis virtual sources.

In addition, as shown in Figure A-8, the S-P wave in partial derivative wavefields obtained by HB_{xxh} (members of full basis virtual source for the P-wave velocity in the IPG-II) is reversed. However, this is not a severe problem because other waves such as P-P, P-S and S-S waves contribute to the gradient direction in positive ways. Although, as figures show, all the partial derivative wavefields have this kind of trade-off problem, their effects are not severe. Particularly, for even shear stress-induced basis virtual sources, these negative effects can be compensated because they generally work together.

On the other hand, for the HB_{xzv}, all the S-wave related reflected waves such as P-S, S-P and S-S waves are reversed and finally the gradient direction is calculated in negative ways due to the strong S-S waves. In addition, for the VB_{zxh}, the S-S reflected wave, which has the strongest amplitude, is reversed and finally the gradient direction also can go to negative directions. For these reasons, compared to the acoustic VTI FWI, in the elastic VTI media, it is hard to invert the c_{13} when the FWI process is S-wave dominant (large amplitudes of reflected S-wave in residual wavefields) as Appendix C shows. These behavior of the HB_{xzv} and VB_{zxh}, which are members of c_{13}, can be treated as a natural phenomenon. Because a decrease of c_{13} causes an increase of S-wave velocity as shown in eq. (3-10), this behavior is faster way to reduce the error when the FWI process is S-wave dominant, in which S-waves would contribute large portion in the total error.

This reversal tendency of HB_{xzv} and VB_{zxh} also cause some problems in the isotropic elastic FWI because the FWI process can be P-wave dominant or S-wave dominant depending on the subsurface Poisson's ratio. In my experience, for the most cases when the FWI process is S-wave dominant, the conventional isotropic FWI is also severely distorted by these negative contribution of HB_{xzv} and VB_{zxh}. Further studies are required for more accurate analysis for this problem.

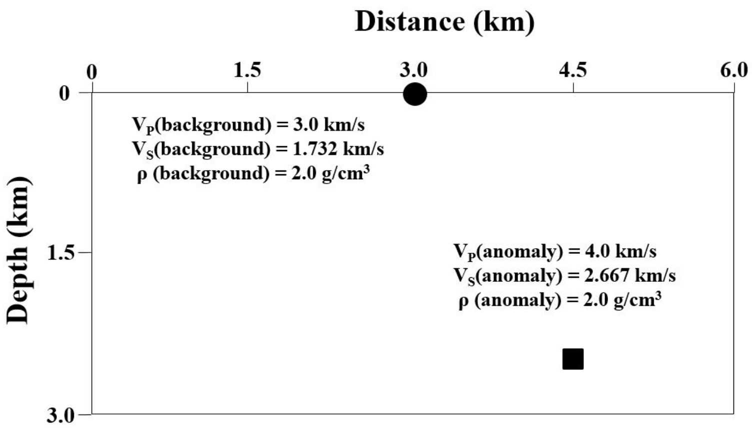


Figure A-1 Single perturbation model. Black rectangle and black circle denote the location of perturbation and the seismic source.

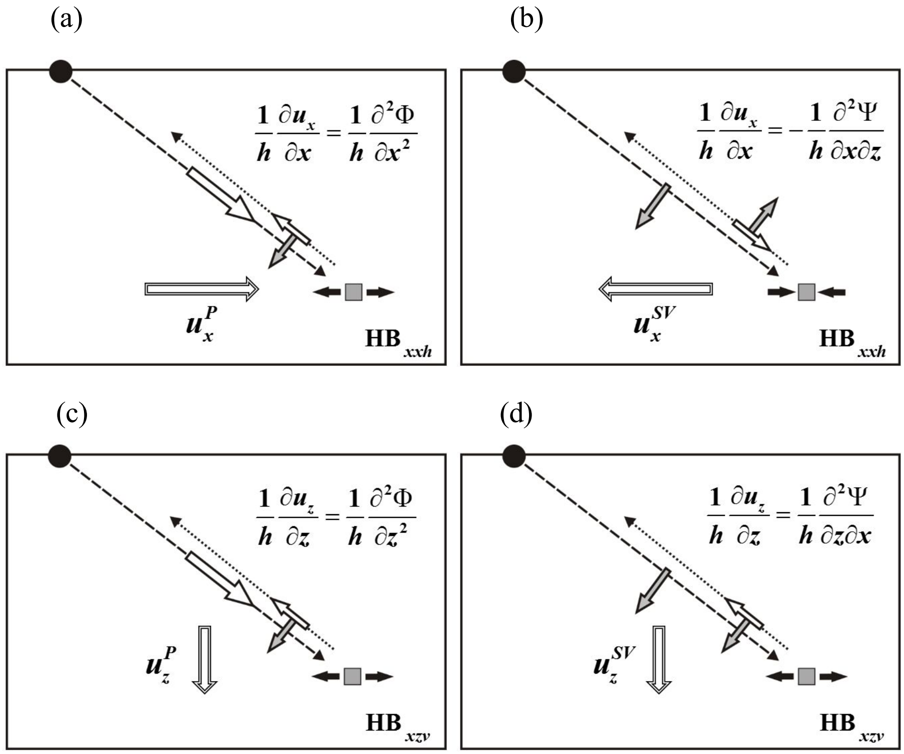


Figure A-2 The scattering mechanism of basis virtual sources, HB_{xxh} (a and b) and HB_{xzv} (c and d), induced by (a) horizontal displacements and (c) vertical displacements of incidence P-wave and (b) horizontal displacements and (d) vertical displacements of incidence SV-wave. The white and gray arrows indicate the first particle motion of P- and S-waves. The double and black arrows denote the effective incidence motion and corresponding moment tensor of basis virtual source. Black circle and gray rectangle denote the source and scatterer. The dashed and dotted lines indicate the ray path of incidence and scattered waves.

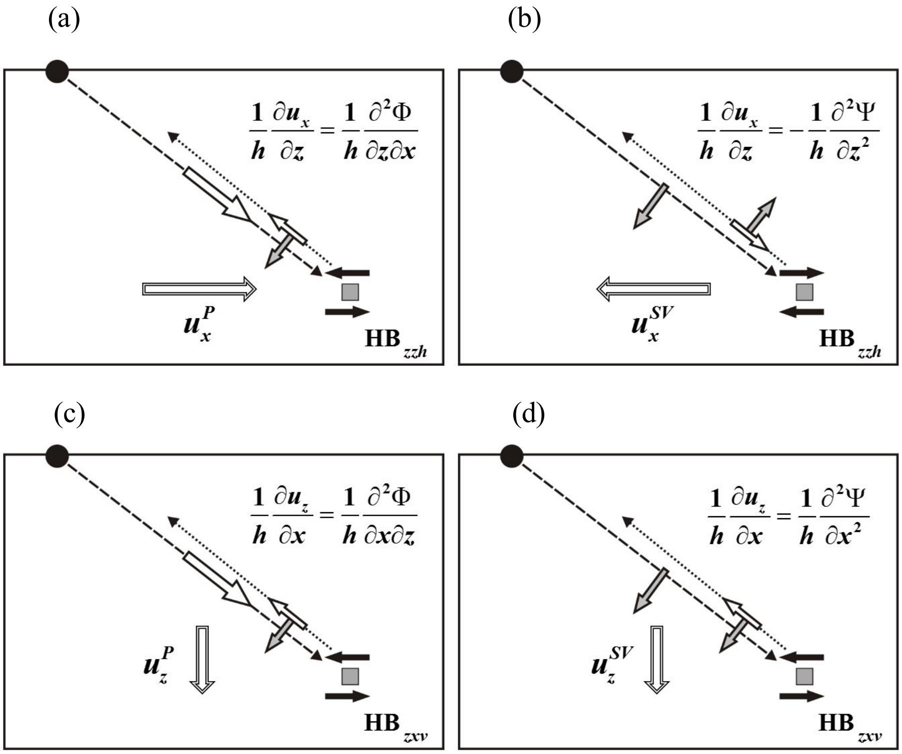


Figure A-3 The scattering mechanism of basis virtual sources, HB_{zzh} (a and b) and HB_{zxv} (c and d), induced by (a) horizontal displacements and (c) vertical displacements of incidence P-wave and (b) horizontal displacements and (d) vertical displacements of incidence SV-wave.

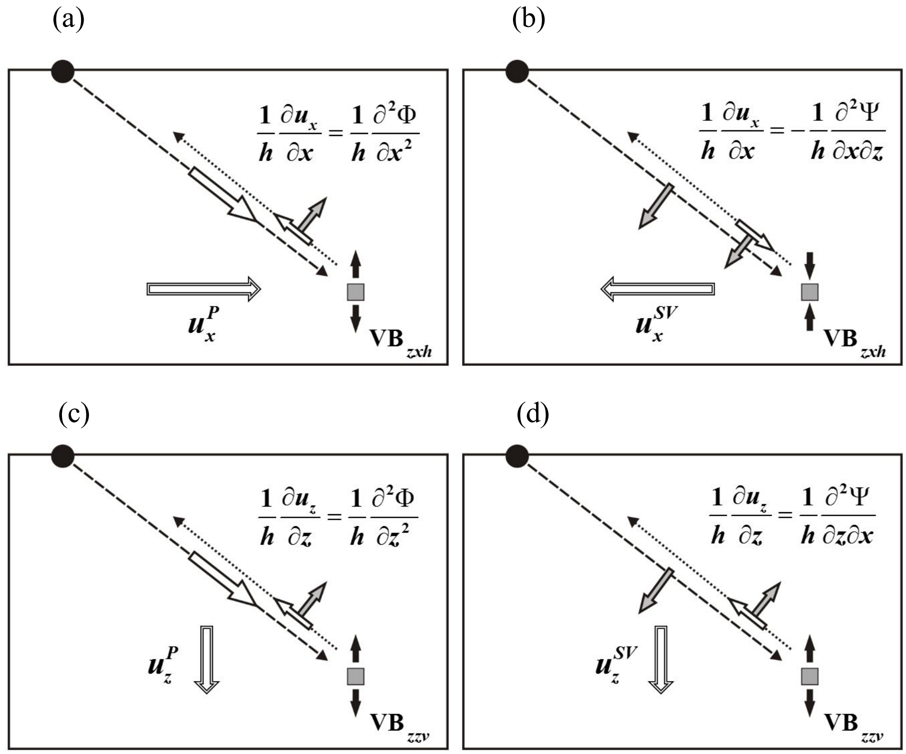


Figure A-4 The scattering mechanism of basis virtual sources, VB_{zxh} (a and b) and VB_{zzv} (c and d), induced by (a) horizontal displacements and (c) vertical displacements of incidence P-wave and (b) horizontal displacements and (d) vertical displacements of incidence SV-wave.

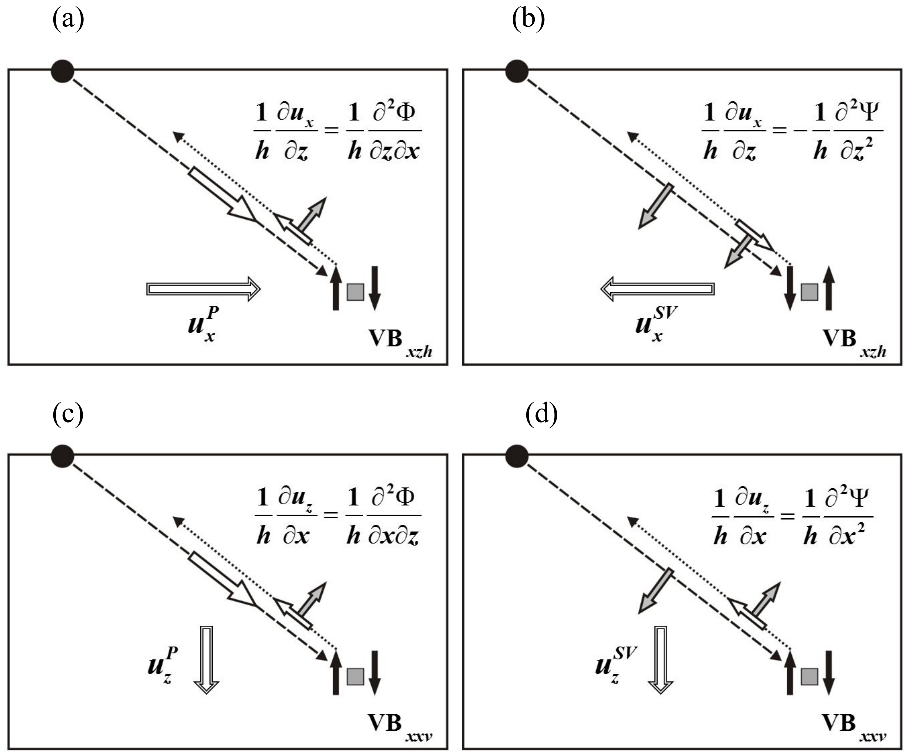


Figure A-5 The scattering mechanism of basis virtual sources, VB_{xzh} (a and b) and VB_{xzv} (c and d), induced by (a) horizontal displacements and (c) vertical displacements of incidence P-wave and (b) horizontal displacements and (d) vertical displacements of incidence SV-wave.

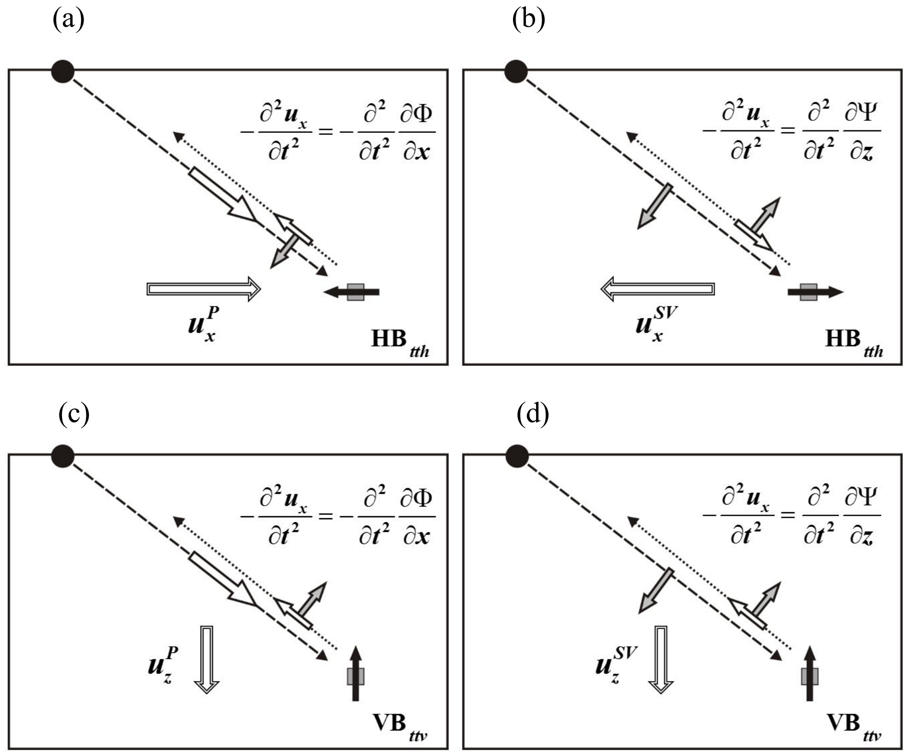


Figure A-6 The scattering mechanism of basis virtual sources, HB_{tth} (a and b) and VB_{ttv} (c and d), induced by (a) horizontal displacements and (c) vertical displacements of incidence P-wave and (b) horizontal displacements and (d) vertical displacements of incidence SV-wave.

(a) (b)

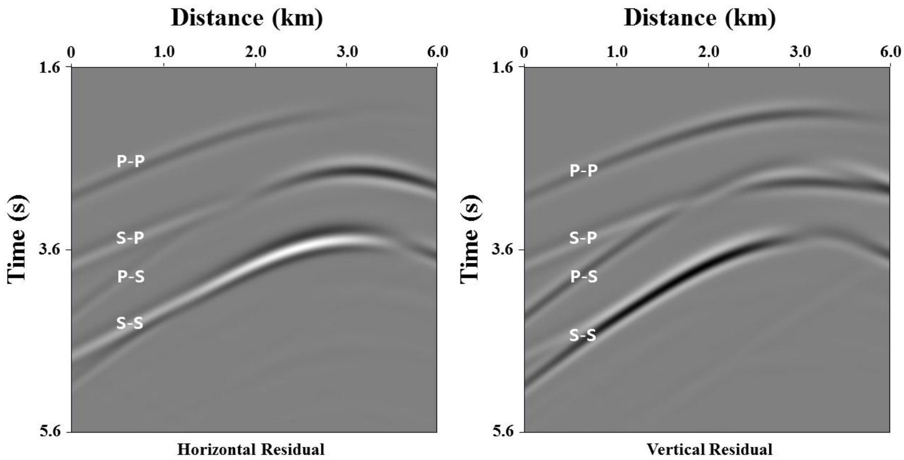


Figure A-7 Residual seismograms for the (a) horizontal and (b) vertical displacements obtained by the single perturbation and initial models (Figure A-1)

(a) (b)

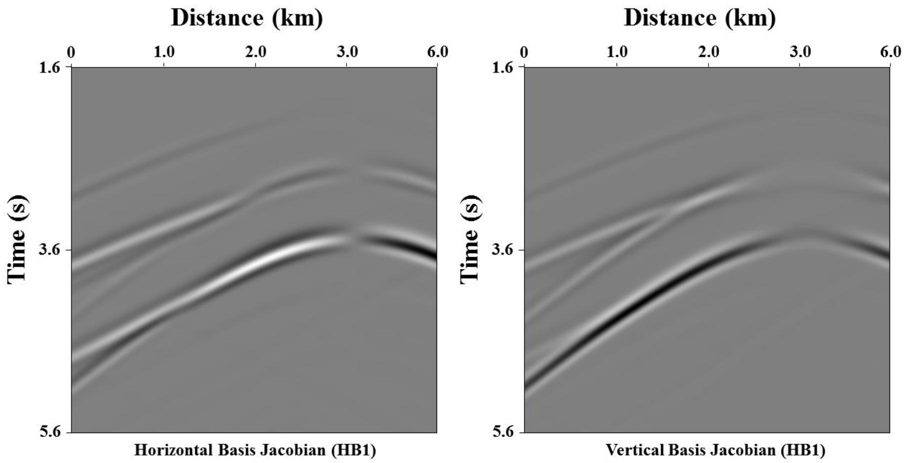


Figure A-8 Partial derivative wavefields obtained by the HB_{xxh} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

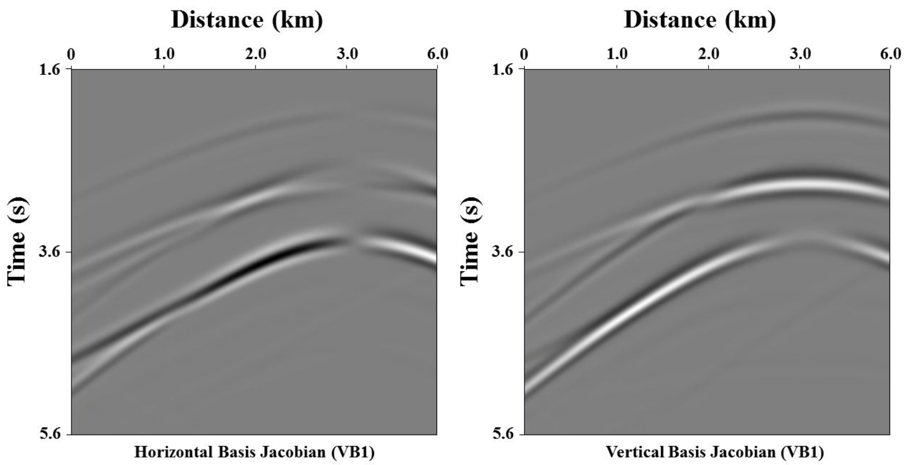


Figure A-9 Partial derivative wavefields obtained by the VB_{zxh} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

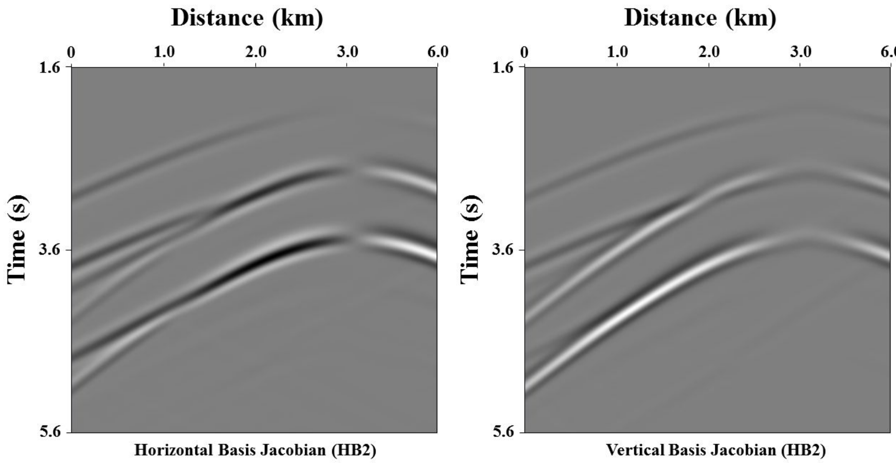


Figure A-10 Partial derivative wavefields obtained by the HB_{xzv} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

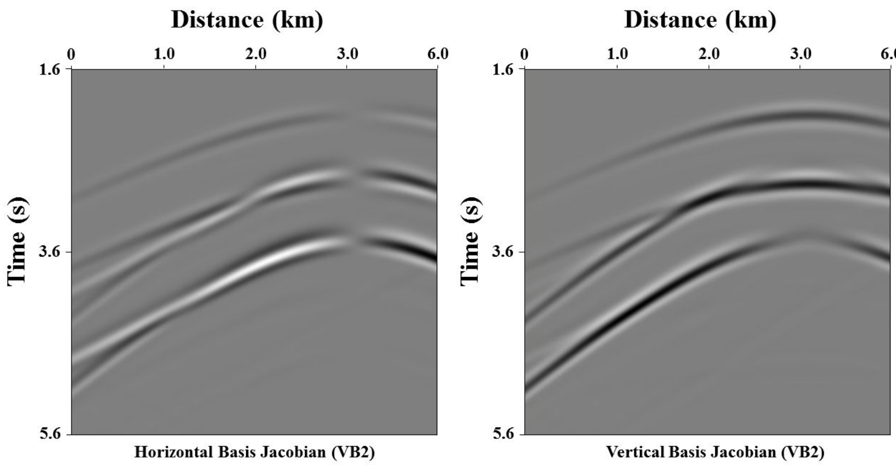


Figure A-11 Partial derivative wavefields obtained by the VB_{zzv} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

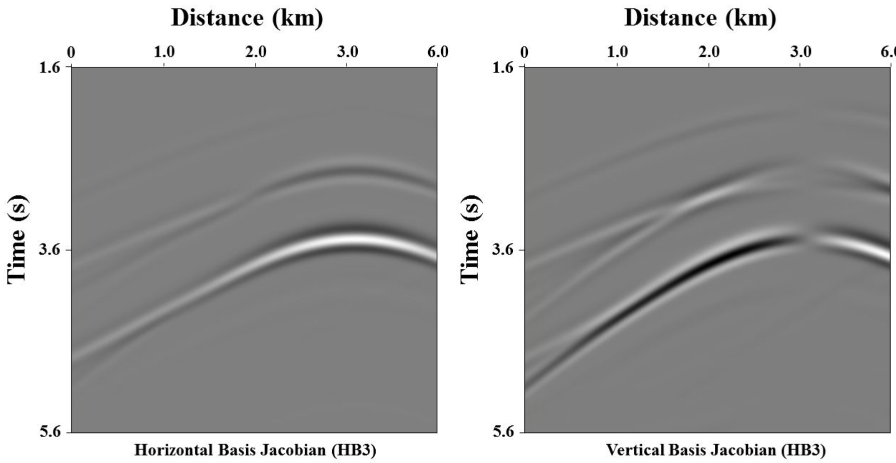


Figure A-12 Partial derivative wavefields obtained by the HB_{zzh} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

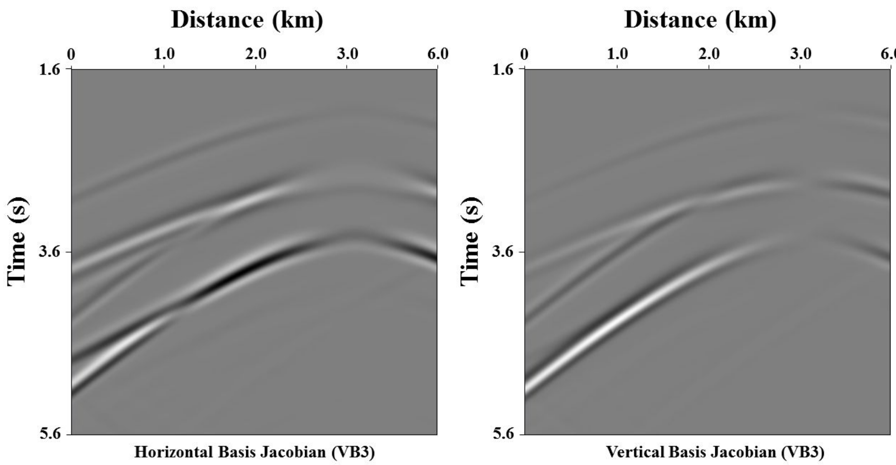


Figure A-13 Partial derivative wavefields obtained by the VB_{xzh} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

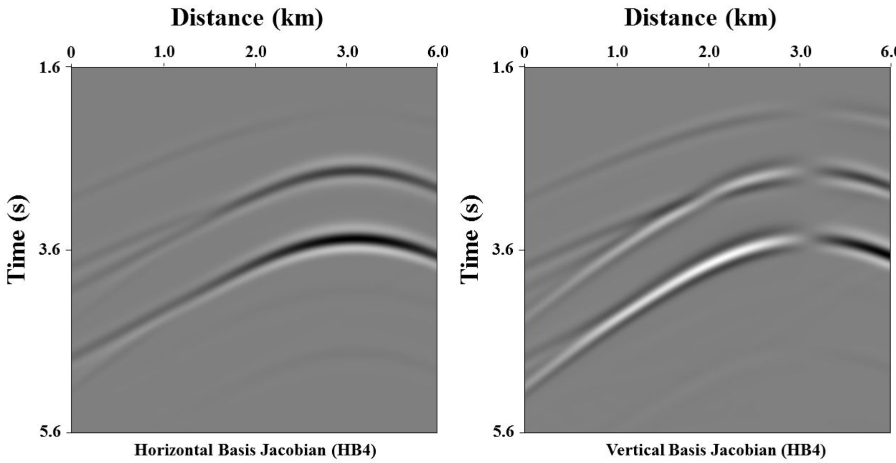


Figure A-14 Partial derivative wavefields obtained by the HB_{zxv} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

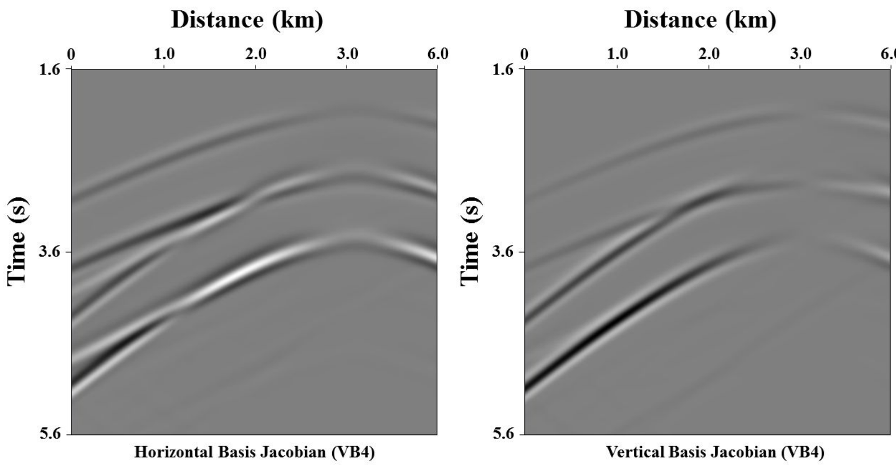


Figure A-15 Partial derivative wavefields obtained by the VB_{xxv} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

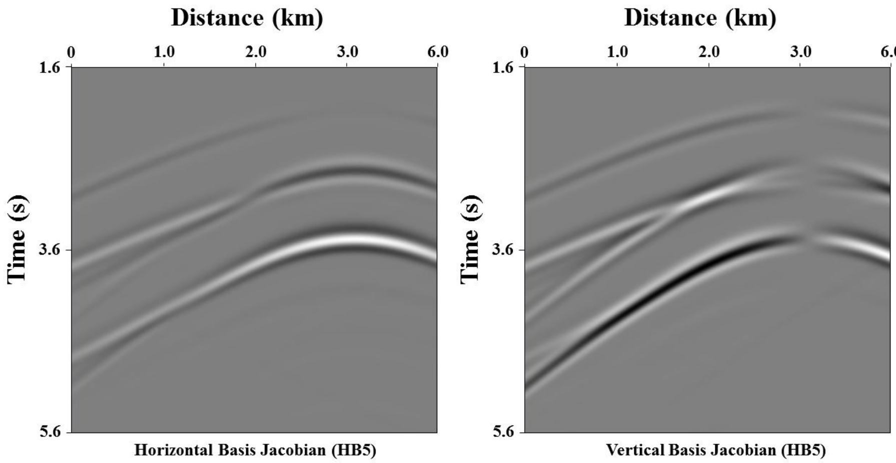


Figure A-16 Partial derivative wavefields obtained by the HB_{tth} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

(a) (b)

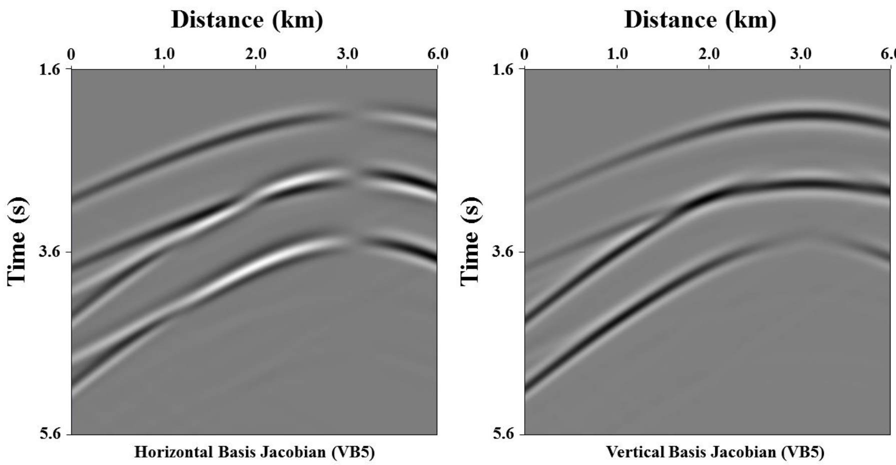


Figure A-17 Partial derivative wavefields obtained by the VB_{ttv} for the homogeneous initial model: (a) horizontal and (b) vertical displacements

Appendix B. Comparison of Radiation Pattern Depending on Parameterization

In Appendix B, we verify that the full virtual source, which is composed by linear combination of basis virtual sources, well describe the radiation patterns of certain parameters that have reported by previous studies. The partial derivative wavefields are calculated in a homogeneous medium when the source is located on the middle of the surface. The PML boundary with thickness of 1 km is applied at top boundary for visualization only when the partial derivative wavefields are computed. As I mentioned in Chapter 3.3.2, the virtual source for λ and P-wave velocity in the IPG-I and IPG-II only generates P-P scattered waves. Comparing the partial derivative wavefields from Figure B-1 to Figure B-4, with the radiation patterns in Tarantola (1986), the radiation patterns for each parameter are well described. For the new parameterization, we can notice that the virtual source for the P-wave velocity (Figures B-5 and B-6) generates all kinds of scattered waves.

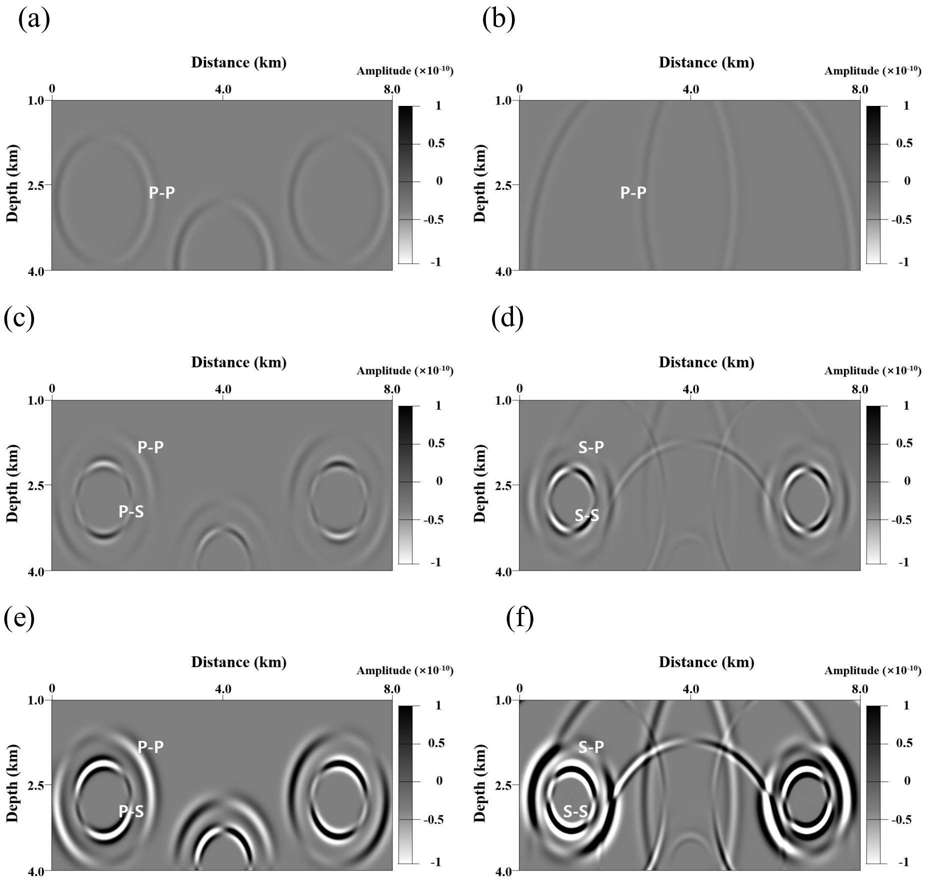


Figure B-1 The snapshots of horizontal displacements for partial derivative wavefields at 1.8s (a, c and e) and 2.7s (b, d and f) obtained by IPG-I: λ (a and b), μ (c and d) and density (e and f)

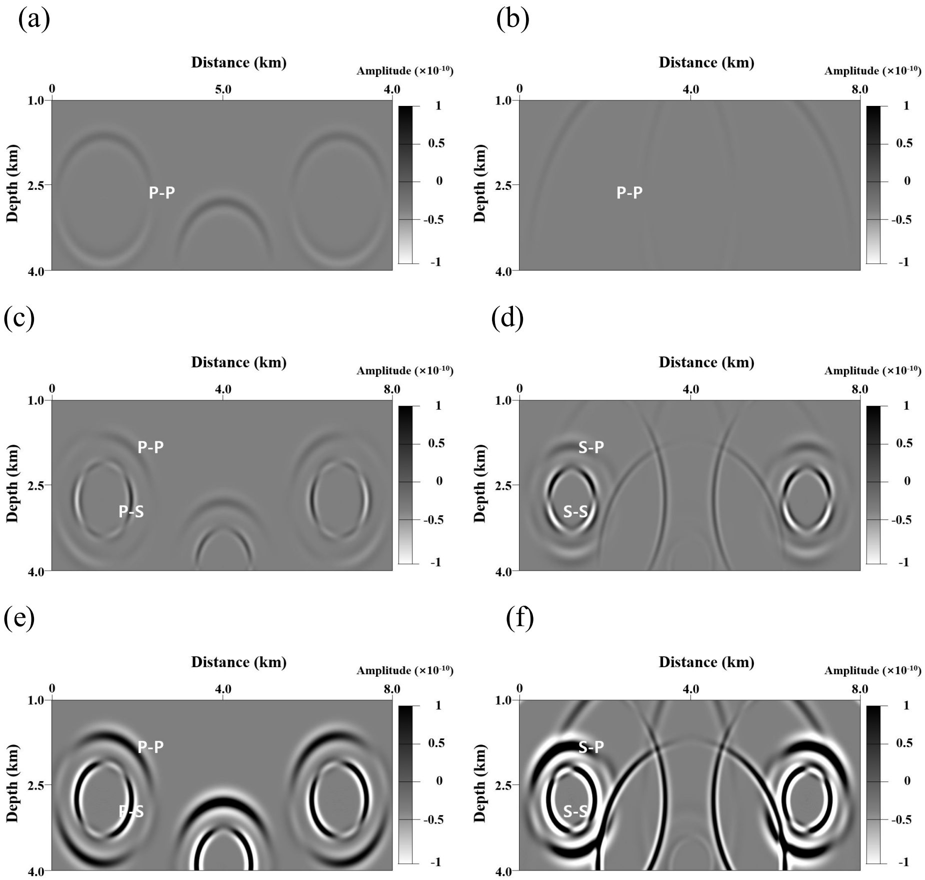


Figure B-2 The snapshots of vertical displacements for partial derivative wavefields at 1.8s (a, c and e) and 2.7s (b, d and f) obtained by IPG-I: λ (a and b), μ (c and d) and density (e and f)

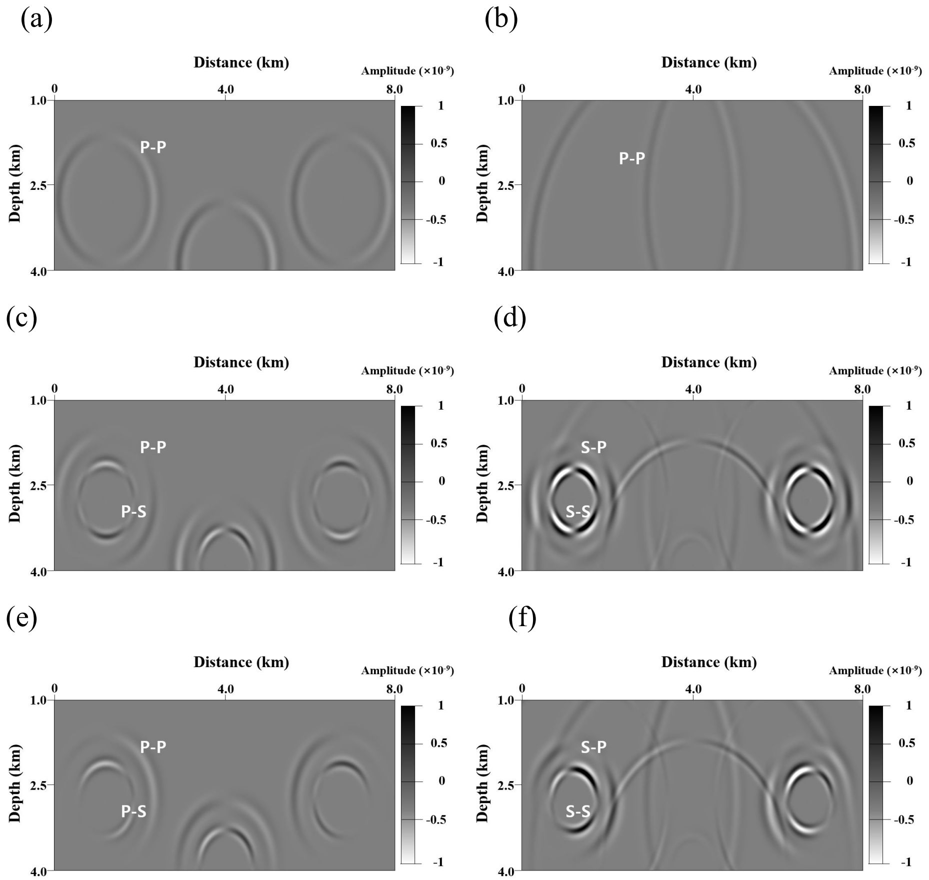


Figure B-3 The snapshots of horizontal displacements for partial derivative wavefields at 1.8s (a, c and e) and 2.7s (b, d and f) obtained by IPG-II: P-wave (a and b), S-wave (c and d) velocities and density (e and f)

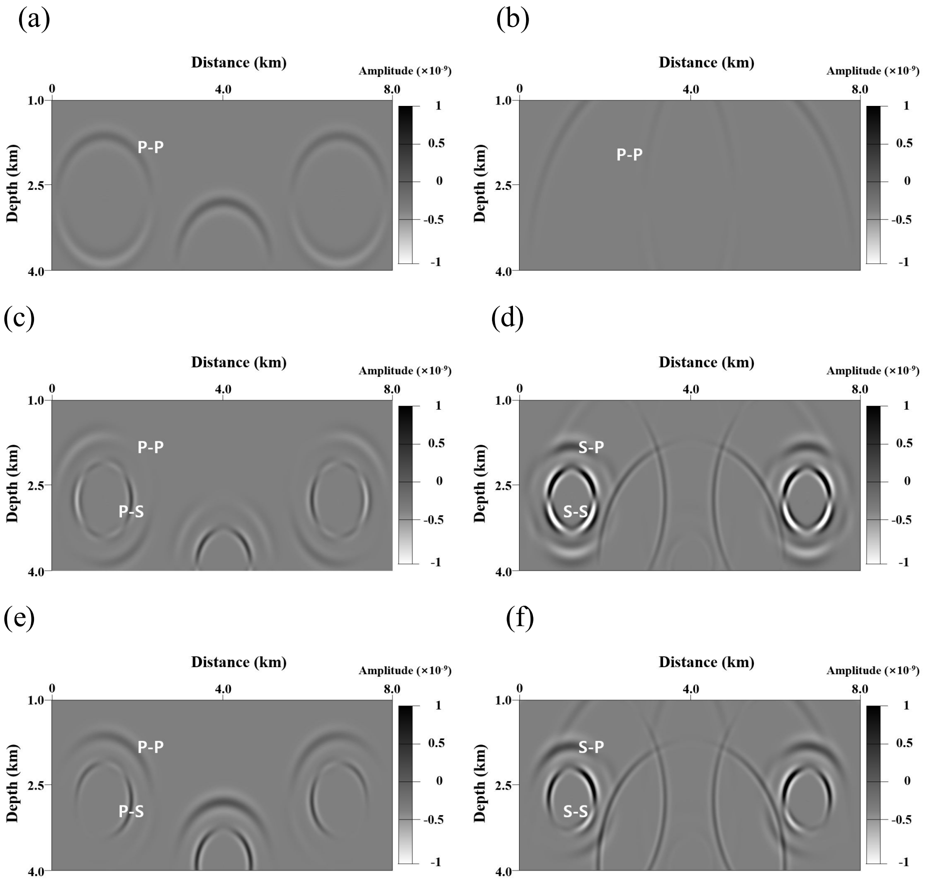


Figure B-4 The snapshots of vertical displacements for partial derivative wavefields at 1.8s (a, c and e) and 2.7s (b, d and f) obtained by IPG-II: P-wave (a and b), S-wave (c and d) velocities and density (e and f)

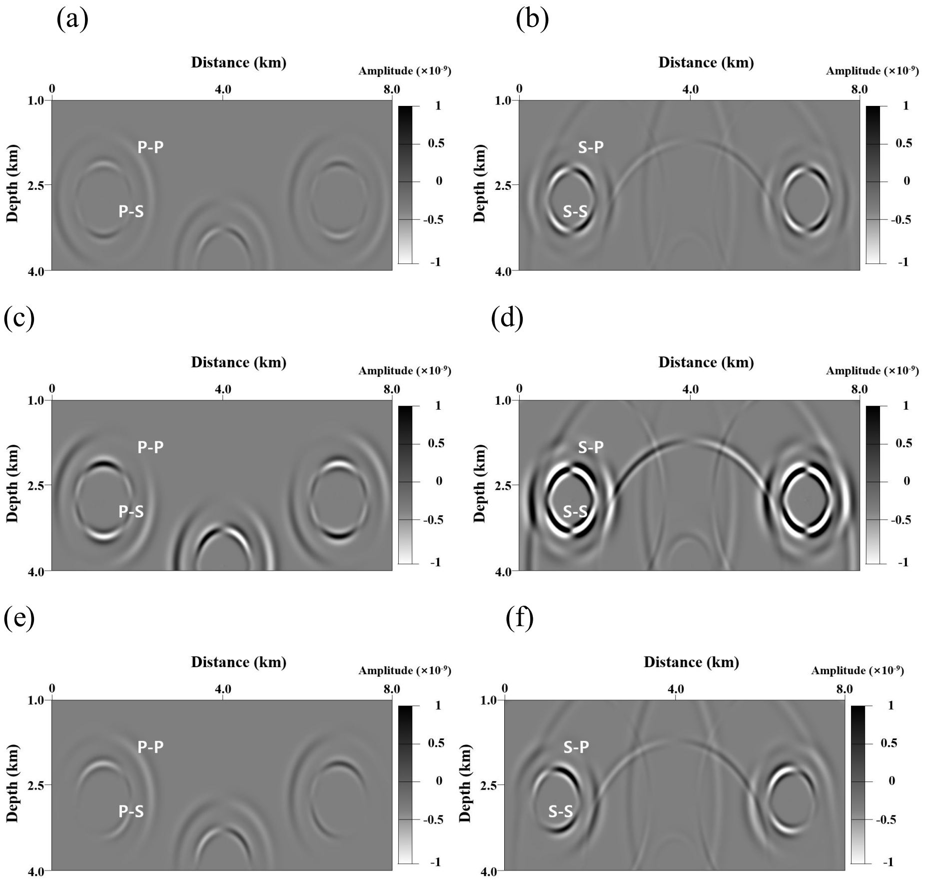


Figure B-5 The snapshots of horizontal displacements for partial derivative wavefields at 1.8s (a, c and e) and 2.7s (b, d and f) obtained by IPG-III: P-wave (a and b), Poisson's ratio (c and d) velocities and density (e and f)

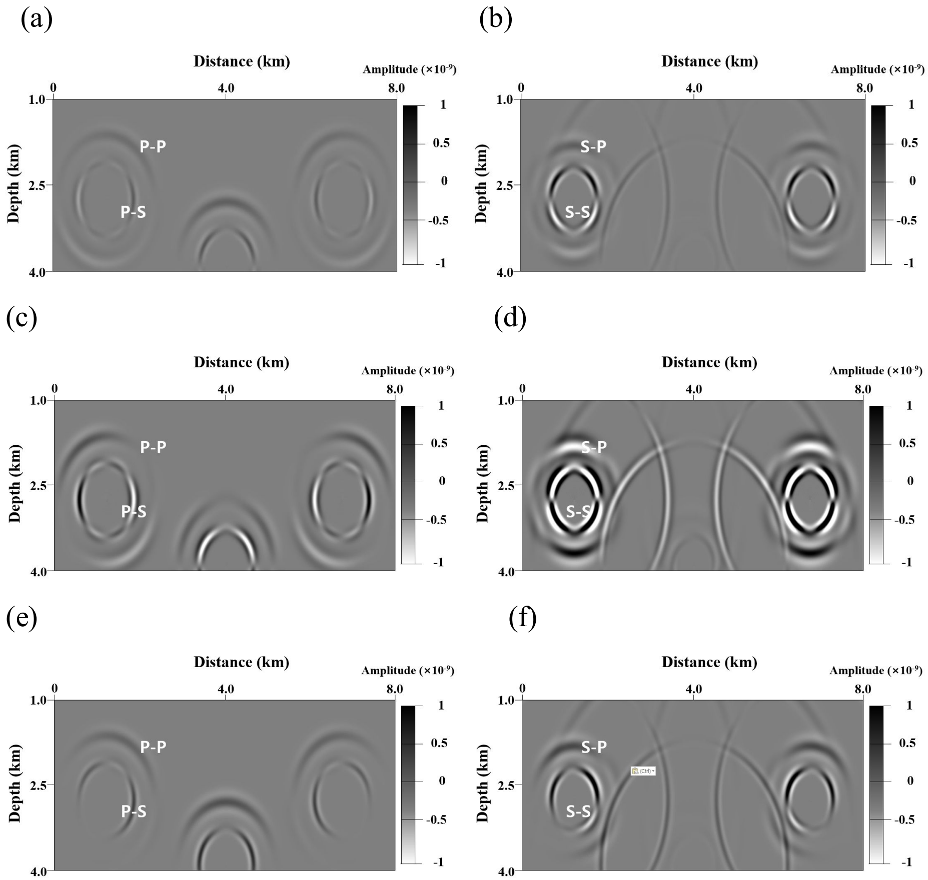


Figure B-6 The snapshots of vertical displacements for partial derivative wavefields at 1.8s (a, c and e) and 2.7s (b, d and f) obtained by IPG-III: P-wave (a and b), Poisson's ratio (c and d) velocities and density (e and f)

Appendix C. New Parameterization for VTI Elastic FWI

C.1 New Parameterization for VTI media using Poisson's ratio

The elastic wave equation to describe the elastic VTI media can be expressed by

$$\begin{aligned}
\tau_{xx} &= c_{11}\frac{\partial u_x}{\partial x} + c_{13}\frac{\partial u_z}{\partial z} \\
\tau_{zz} &= c_{13}\frac{\partial u_x}{\partial x} + c_{33}\frac{\partial u_z}{\partial z} \\
\tau_{xz} &= c_{44}\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right)
\end{aligned} \tag{C-1}$$

and

$$\begin{aligned}
\rho\frac{\partial^2 u_x}{\partial t^2} &= \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xz}}{\partial z} \\
\rho\frac{\partial^2 u_z}{\partial t^2} &= \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{zz}}{\partial z}
\end{aligned}. \tag{C-2}$$

The terms, c_{11}, c_{33}, c_{13}, c_{44} and density (ρ), denote the four elastic stiffness coefficients for the 2D VTI elastic media and the density, respectively. Other parameters, such as horizontal ($v_{p,H}$) and vertical P-wave velocity ($v_{p,V}$), S-wave velocity (v_s), Poisson's ratio (v), Thomsen's parameters (ε and δ), can be calculated as follows:

$$v_{p,H} = \sqrt{\frac{c_{11}}{\rho}} \tag{C-3}$$

$$v_{p,V} = \sqrt{\frac{c_{33}}{\rho}} \tag{C-4}$$

$$v_s = \sqrt{\frac{c_{44}}{\rho}} \tag{C-5}$$

$$\sigma = \frac{c_{33} - 2c_{44}}{2c_{33} - 2c_{44}} \tag{C-6}$$

$$\varepsilon = \frac{c_{11} - c_{33}}{2c_{33}} \tag{C-7}$$

$$\delta = \frac{\left(c_{13} + c_{44}\right)^2 - \left(c_{33} - c_{44}\right)^2}{2c_{33}\left(c_{33} - c_{44}\right)} \tag{C-8}$$

Using these basis virtual sources, the full virtual source can be generalized as a linear combination of basis virtual sources as expressed by

$$\begin{aligned} \mathbf{f}_{p_i}^{\text{full}} &= \left(\frac{\partial c_{11}}{\partial p_i}\right)\text{HB}_{xxh} + \left(\frac{\partial c_{33}}{\partial p_i}\right)\text{VB}_{zzv} + \left(\frac{\partial c_{13}}{\partial p_i}\right)\left(\text{HB}_{xzv} + \text{VB}_{zxh}\right) \\ &+ \left(\frac{\partial c_{44}}{\partial p_i}\right)\left(\text{HB}_{zzh} + \text{HB}_{zxv} + \text{VB}_{xzh} + \text{VB}_{xxv}\right) + \left(\frac{\partial \rho}{\partial p_i}\right)\left(\text{HB}_{tth} + \text{VB}_{ttv}\right) \end{aligned} . \tag{C-9}$$

If we consider the form of the elastic VTI wave equation and the isotropic elastic wave equation, we can notice that the scattering pattern induced by the basis virtual sources are totally same when the initial model is homogeneous because the scattering pattern of the basis virtual source is determined by the spatial derivatives. For this reason, in this paper, I also analyze the behavior of the VTI virtual source based on the scattering theory of the basis virtual sources as I did for the isotropic parameterization.

For the VTI elastic FWI, one of the most popular parameterization is to parameterize the VTI elastic wave equation with the displacement-based form using the c_{11}, c_{33}, c_{13}, c_{44} and density, which we call 'VTIPG-I (VTI Parameter Group-I)', as follows:

$$\begin{aligned}\rho\frac{\partial^2 u_x}{\partial t^2}&=\frac{\partial}{\partial x}\left[c_{11}\frac{\partial u_x}{\partial x}+c_{13}\frac{\partial u_z}{\partial z}\right]+\frac{\partial}{\partial z}\left[c_{44}\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right]\\ \rho\frac{\partial^2 u_z}{\partial t^2}&=\frac{\partial}{\partial z}\left[c_{13}\frac{\partial u_x}{\partial x}+c_{33}\frac{\partial u_z}{\partial z}\right]+\frac{\partial}{\partial x}\left[c_{44}\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right]\end{aligned} \qquad \text{(C-10)}$$

In this case, we invert four elastic stiffness coefficients and the other parameters, such as horizontal and vertical P-wave velocity, S-wave velocity, Poisson's ratio, Thomsen parameters can be indirectly calculated using from eqs. (C-3) to (C-8), respectively.

Using the linear combination of the basis virtual source, the full virtual sources for the c_{11}, c_{33}, c_{13}, c_{44} and density can be expressed as follows:

$$\begin{aligned}\mathbf{f}^{\text{full}}_{c_{11(i)}}&=\text{HB}_{xxh}\\ \mathbf{f}^{\text{full}}_{c_{33(i)}}&=\text{VB}_{zzv}\\ \mathbf{f}^{\text{full}}_{c_{13(i)}}&=\left(\text{HB}_{xzv}+\text{VB}_{zxh}\right)\\ \mathbf{f}^{\text{full}}_{c_{44(i)}}&=\left(\text{HB}_{zzh}+\text{HB}_{zxv}+\text{VB}_{xzh}+\text{VB}_{xxv}\right)\\ \mathbf{f}^{\text{full}}_{\rho(i)}&=\left(\text{HB}_{tth}+\text{VB}_{ttv}\right)\end{aligned} \qquad \text{(C-11)}$$

With these expressions, I notice that, the conventional parameters for the VTI media are well organized depending on their mechanical motions. In other word, the full virtual sources for the c_{11} and c_{33} only acts like the horizontal and vertical normal stress, respectively. Because the basis virtual source, HB_{xxh}, mainly use the incidence wave with high incidence angle, I can guess that the c_{11} for the deeper structures are not recovered well with the conventional full virtual source for the c_{11}. On the other hand, the basis virtual source, VB_{zzv}, dominantly use the incidence wave with low incidence angle, which means that, if the seismic sources are sparsely spaced, the conventional full virtual source for the c_{33} suffers from the lack of the data coverage. The full virtual source for the c_{13} is a combination of the horizontal and vertical normal stress. However, as I showed in Appendix A, the S-wave-related motion of these two basis virtual sources (HB_{xzv} and VB_{zxh}) partially contribute to negative direction depending on the subsurface parameters. The full virtual source for the c_{44} acts like a double coupled forces, which are good to resolve the subsurface structures located in the intermediate incidence angle. As we discussed in Chapter 3.2.3, the full virtual source for the density consists of unidirectional basis virtual sources, resulting a less sensitive Jacobian matrix. These characteristics are also easily found in the previous works of the FWI for the VTI media. Lee at al. (2010) showed that the frequency-domain FWI for VTI media with the VTIPG-I provides poor FWI results of the c_{11} and c_{13} structures and suggested the coupling method, in which the gradients for the c_{11} are indirectly calculated using the gradients for the c_{33}, and the two-stage FWI. They used modified overthrust model, whose Poisson's ratios are fixed at 0.25, and did not recover the density structures under the assumption that we know the exact density structures. However, it is still questionable whether their method also works well for the real overthrust structures with various Poisson's ratio and density because, as I mentioned in Appendix A, the FWI process can be P-wave dominant or S-wave dominant depending on the subsurface Poisson's ratio. In addition, the two-stage FWI strategy for the real field data, which include various kinds of noise, has lots of uncertainties because, in real case, it is hard to determine the exact criterion stopping the first stage of the FWI. To overcome these limitations, we try to seek a solution by finding the best parameterization for the VTI elastic media based on the scattering theory of the basis virtual source.

To enhance the gradient directions, I suggest the new parameter group for the VTI media which we call 'VTIPG-II' throughout the paper. The elastic wave equation for the VTI media can be expressed by the VTIPG-II as follows:

$$\begin{aligned}
\rho\frac{\partial^2 u_x}{\partial t^2} &= \frac{\partial}{\partial x}\left[\rho v_{p,H}^2\frac{\partial u_x}{\partial x}+c_{13}\frac{\partial u_z}{\partial z}\right] \\
&\quad+\frac{\partial}{\partial z}\left[\left(\frac{\rho v_{p,H}^2}{1+2\varepsilon}\right)\left(\frac{1-2\nu}{2-2\nu}\right)\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right] \\
\rho\frac{\partial^2 u_z}{\partial t^2} &= \frac{\partial}{\partial z}\left[c_{13}\frac{\partial u_x}{\partial x}+\left(\frac{\rho v_{p,H}^2}{1+2\varepsilon}\right)\frac{\partial u_z}{\partial z}\right] \\
&\quad+\frac{\partial}{\partial x}\left[\left(\frac{\rho v_{p,H}^2}{1+2\varepsilon}\right)\left(\frac{1-2\nu}{2-2\nu}\right)\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)\right]
\end{aligned} \tag{C-12}$$

The full virtual sources for the VTIPG-II can expressed by

$$\begin{aligned}
\mathbf{f}_{v_{p,H(i)}}^{\text{full}} &= 2\rho v_{p,H}\mathrm{HB}_{xxh}+\left(\frac{2\rho v_{p,H}}{1+2\varepsilon}\right)\mathrm{VB}_{zzv} \\
&\quad+\left[\left(\frac{2\rho v_{p,H}}{1+2\varepsilon}\right)\left(\frac{1-2\nu}{2-2\nu}\right)\right]\left(\mathrm{HB}_{zzh}+\mathrm{HB}_{zxv}+\mathrm{VB}_{xzh}+\mathrm{VB}_{xxv}\right)
\end{aligned}$$

$$
\begin{aligned}
\mathbf{f}_{v_{p,V(i)}}^{\text{full}} &= \left[\frac{-2\rho v_{p,H}^{2}}{(1+2\varepsilon)^{2}} \right] \mathrm{VB}_{zzv} \\
&\quad + \left[\frac{-2\rho v_{p,H}^{2}}{(1+2\varepsilon)^{2}} \left(\frac{1-2\nu}{2-2\nu} \right) \right] \left(\mathrm{HB}_{zzh} + \mathrm{HB}_{zxv} + \mathrm{VB}_{xzh} + \mathrm{VB}_{xxv} \right) \\
\mathbf{f}_{C_{13(i)}}^{\text{full}} &= \left(\mathrm{HB}_{xzv} + \mathrm{VB}_{zxh} \right) \\
\mathbf{f}_{\sigma_{(i)}}^{\text{full}} &= \left[\left(\frac{-2\rho v_{p,H}^{2}}{1+2\varepsilon} \right) \left(\frac{1}{2-2\nu} \right)^{2} \right] \left(\mathrm{HB}_{zzh} + \mathrm{HB}_{zxv} + \mathrm{VB}_{xzh} + \mathrm{VB}_{xxv} \right) \\
\mathbf{f}_{\rho_{(i)}}^{\text{full}} &= v_{p,H}^{2} \mathrm{HB}_{xxh} + v_{p,V}^{2} \mathrm{VB}_{zzv} + \omega^{2} \left(\mathrm{HB}_{tth} + \mathrm{VB}_{ttv} \right) \\
&\quad + \left[\left(\frac{v_{p,V}^{2}}{1+2\varepsilon} \right) \left(\frac{1-2\nu}{2-2\nu} \right) \right] \left(\mathrm{HB}_{zzh} + \mathrm{HB}_{zxv} + \mathrm{VB}_{xzh} + \mathrm{VB}_{xxv} \right)
\end{aligned}
\quad . \text{(C-13)}
$$

Comparing eq. (C-13) with eq. (C-11), we can notice that the full virtual sources for all the parameters except c_{13}, are improved and become to include double-coupled basis virtual sources. For this reason, we can expect that the coverage of gradient directions can be improved.

C.2 Numerical Example for SEG/EAGE Overthrust Model

To verify the feasibility of new VTI parameter groups (VTIPG-II) for complex geologic model, I perform the FWI for 2D SEG/EAGE overthrust model as shown in Figure B-1. Although the FWI for the 2D SEG/EAGE overthrust model have been performed by many previous works (Lee et al. 2012; Jeong et al. 2012), most works used modified version by fixing the Poisson's ratio or density and there haven't been no attempts to test the FWI for its original version. This is because, in our experiences, the SEG/EAGE overthrust model is one of difficult benchmark models, which has nearly inverted c_{13} structure and various Poisson's ratio, and inverting 5 VTI parameters simultaneously is not easy using the conventional parameterization. Another problem is the instability of the PML (Perfectly Matched Layer; Bérenger, 1994). The PML or CPML (Convolutional Perfectly Matched Layer; Roden and Gedney, 2000) is one of the most popular boundary condition due to its good efficiency. When we apply the PML to the elastic wave equation for the VTI media, we must consider three stability conditions of the PML (Béchache et al. 2003). The stability condition-I, stability condition-II and stability condition-III are expressed by

$$\left(\left(c_{13}+c_{44}\right)^2-c_{11}\left(c_{33}-c_{44}\right)\right)\left(\left(c_{13}+c_{44}\right)^2+c_{44}\left(c_{33}+c_{44}\right)\right)\le 0 \quad , \text{(C-14)}$$

$$\left(c_{13}+2c_{44}\right)^2-c_{11}c_{33}\le 0 \qquad \text{(C-15)}$$

and

$$\left(c_{13}+c_{44}\right)^2-c_{11}c_{33}-c_{44}^2\le 0\,, \qquad \text{(C-16)}$$

respectively. We can guess that, for isotropic elastic media, the stability condition-II (eq. [C-15]) always equals to zero. This means that the PML for the VTI elastic wave equation also well operates to the subsurface isotropic layer. However, because the FWI is an iterative method, the estimated parameters for subsurface layers can easily get out the stability condition-II when the layers are isotropic. In other words, when the surface layer is isotropic, the PML can be unstable if we do not perform the VTI inversion carefully. Because our inversion algorithm, in which we use the new pseudo-Hessian matrix (Choi et al. 2008), scaling method using weighting method

(Oh and Min 2013a) with the normalization (Ha et al. 2009) and fixed step length, is based on approximated Gauss-Newton method, we fixed only top layer, under the assumption that we know the exact parameters, and freeze the top layer without parameter update. Below the top layer, we assume that we know only the approximate linear vertical variations of the vertical P-wave velocity (1.6 ~ 5.5 km/s) and use isotropic initial models as shown in Figure C-2. The initial S-wave velocity and density models are estimated using the fixed Poisson's ratio (0.25) and the Gardner's equation (Gardner et al., 1974; $\rho = 1.7 \times v_{p,V}^{0.25}$), respectively. The parameters used for the FWI are listed on Table C-1. Due to the poor initial guesses, we assume that very low frequency components are available.

Figures C-3 and C-4 shows the recovered parameter structures obtained using the VTIPG-I and VTIPG-II, respectively. I notice that there are great improvements when the new VTI parameterization is applied although the c_{13} structures are still not satisfied. These numerical results also support that the interpretation tool for the radiation pattern of basis virtual source, which I suggested in Chapter 3, can be also applied to elastic VTI media.

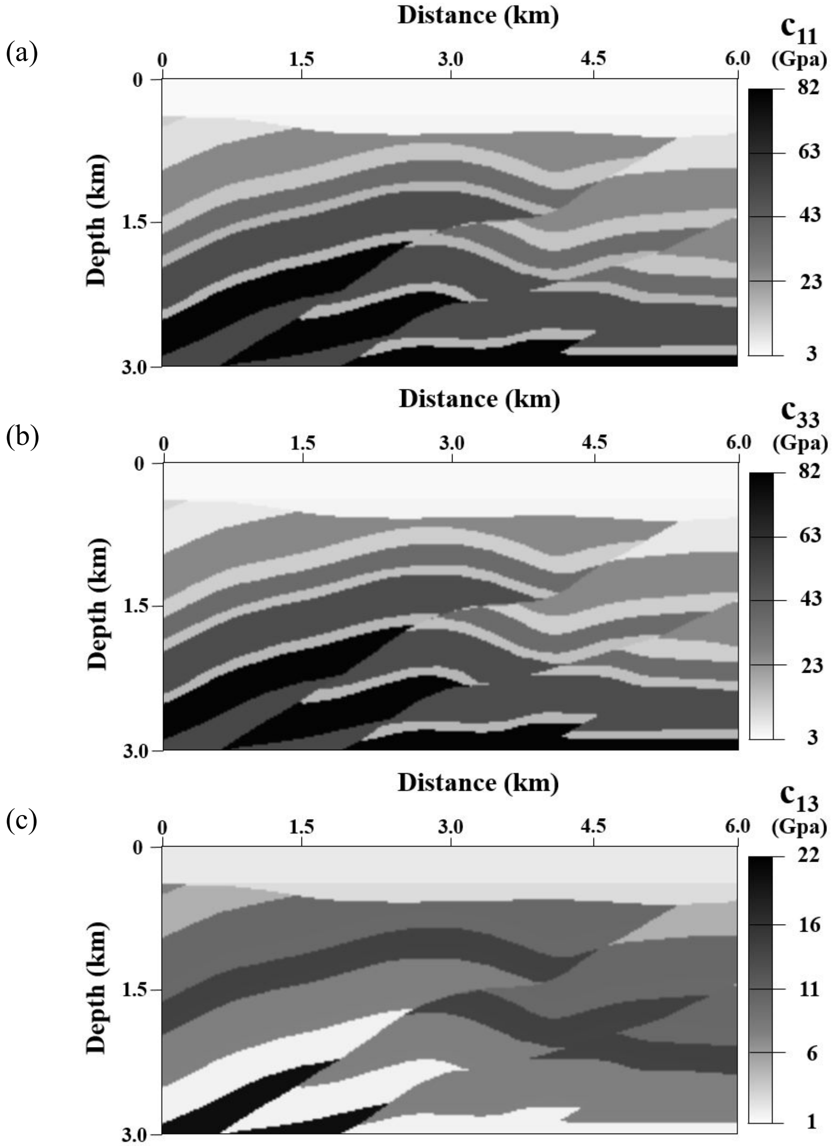
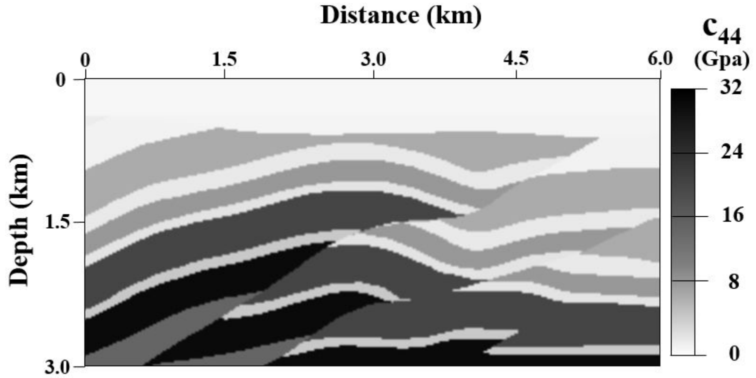


Figure C-1 2D SEG/EAGE overthrust model: (a) c_{11}, (b) c_{33}, (c) c_{13}, (d) c_{44} and (e) density

(d)



(e)

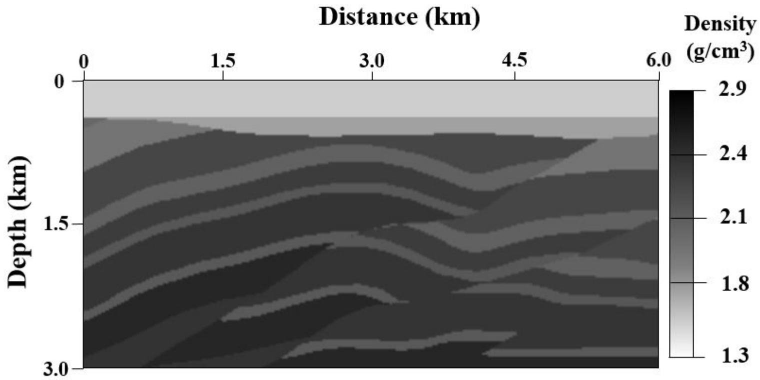


Figure C-1 (Continued)

Table C-1 Inversion parameters for the 2D SEG/EAGE overthrust model

Dimension	No. of shots	No. of receivers	Interval of shots	Interval of receivers	Recording time	Maximum Frequency	Minimum Frequency
6.0 km × 3.0 km	150	601	0.04 km	0.01 km	5 sec	10 Hz	0.2 Hz

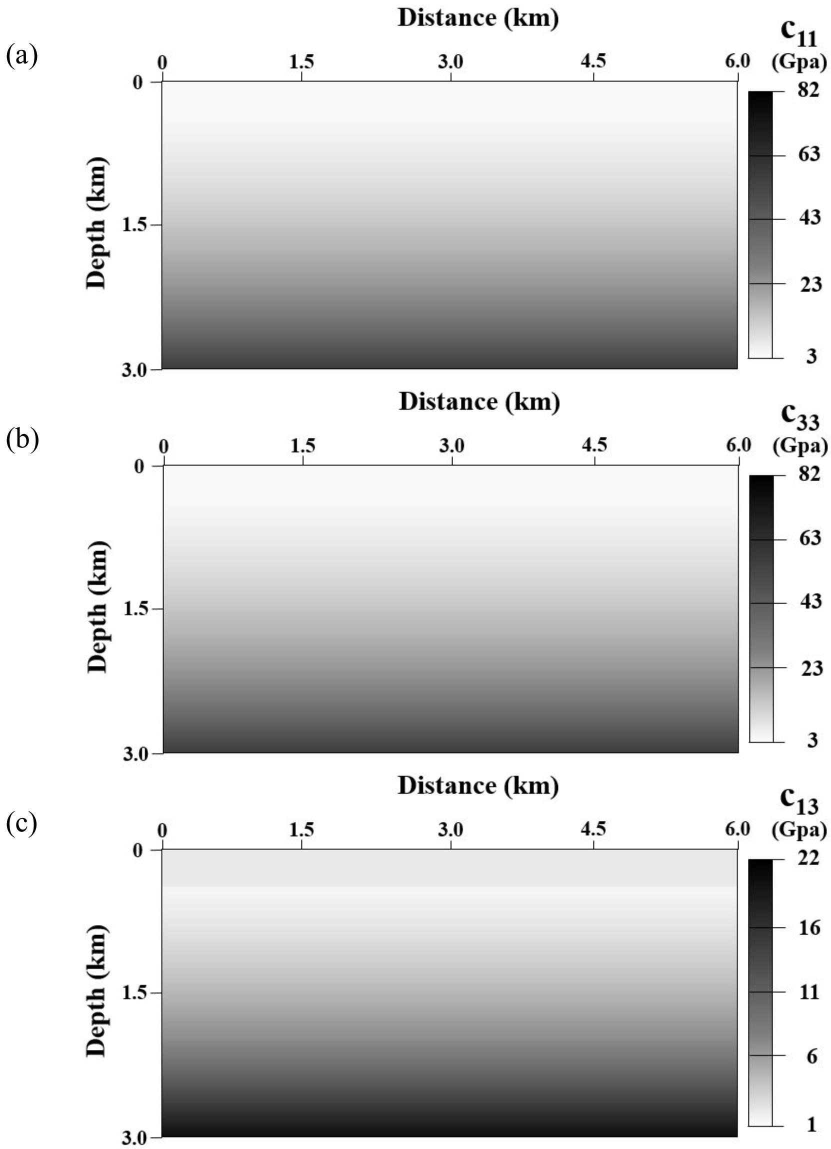


Figure C-2 Initial models for the FWI: (a) c_{11}, (b) c_{33}, (c) c_{13}, (d) c_{44} and (e) density

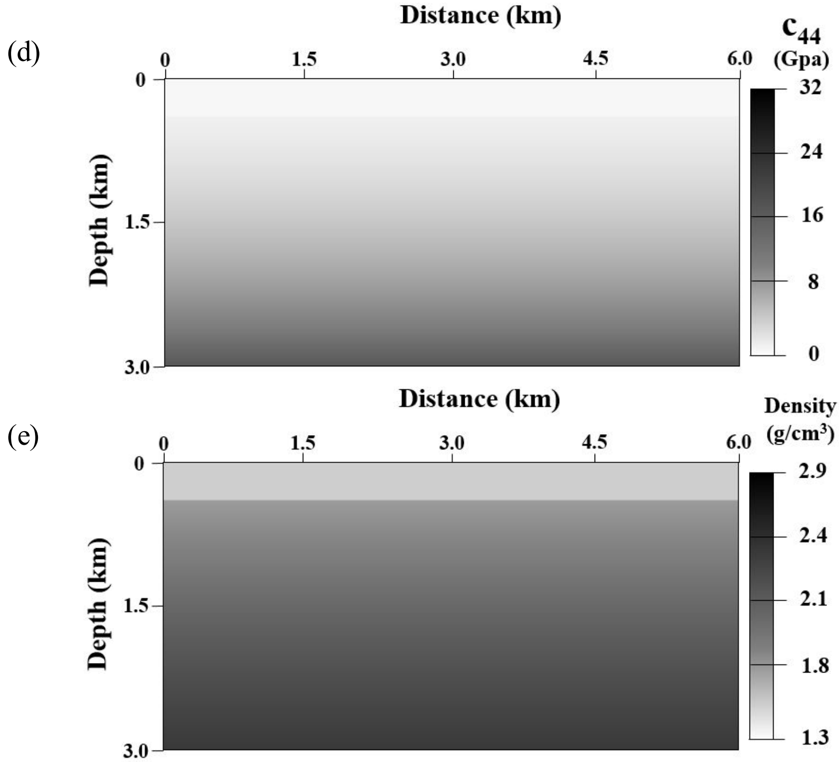


Figure C-2 (Continued)

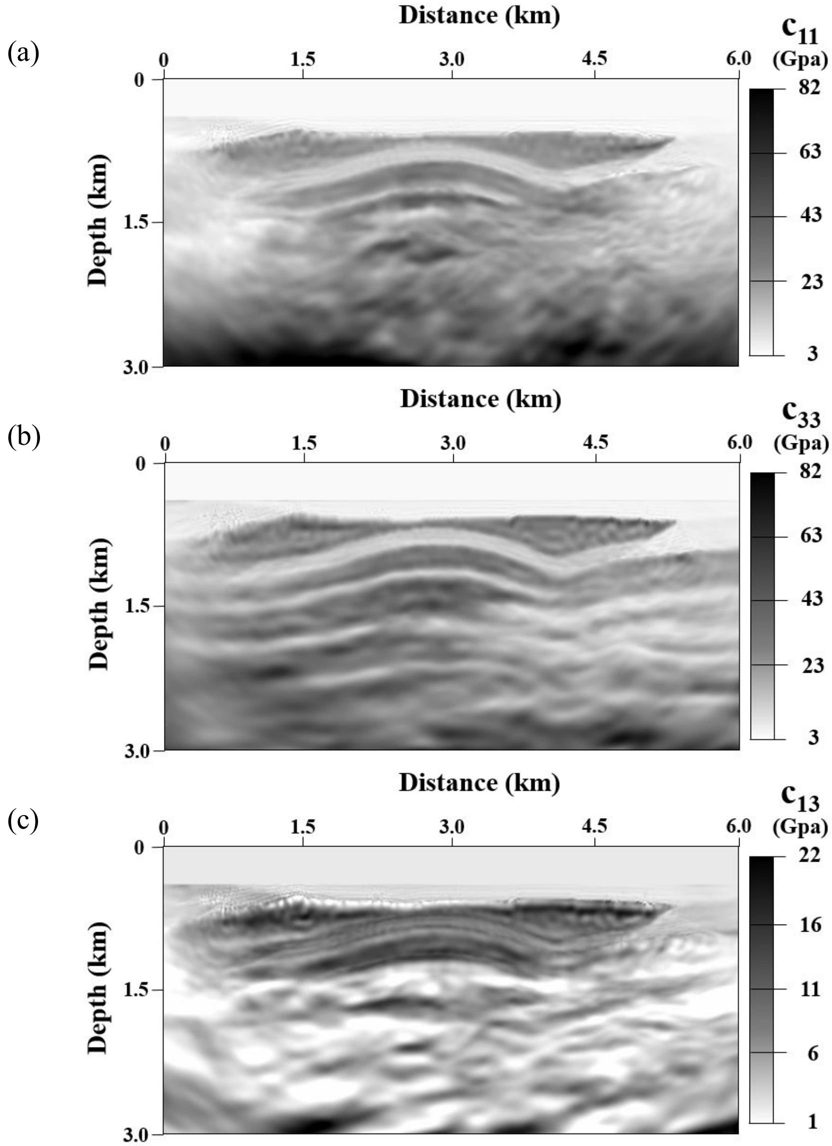


Figure C-3 The recovered structures of (a) c_{11}, (b) c_{33}, (c) c_{13}, (d) c_{44} and (e) density using the VTIPG-I.

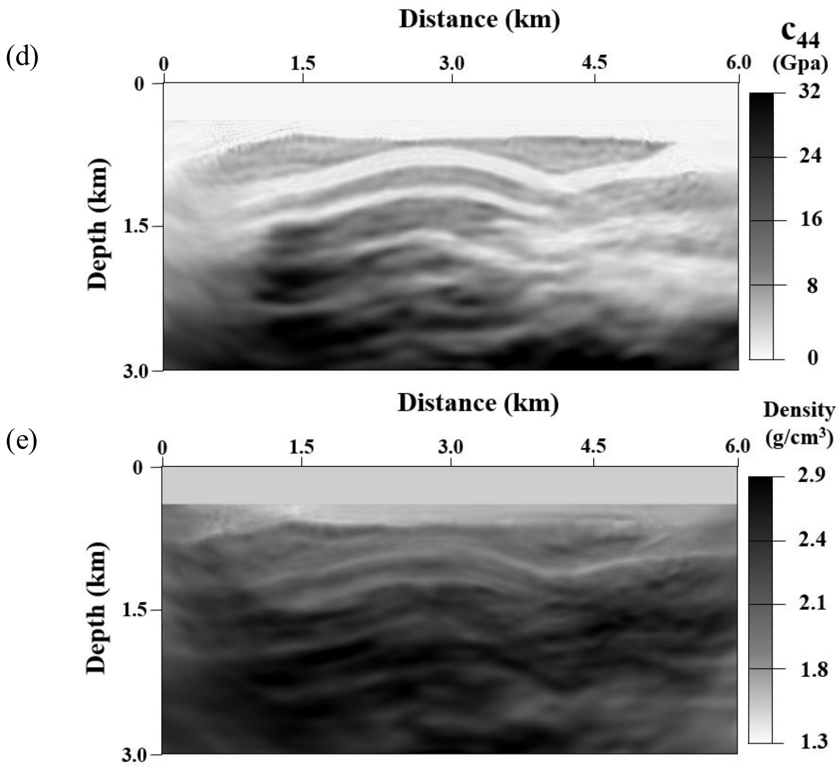


Figure C-3 (Continued)

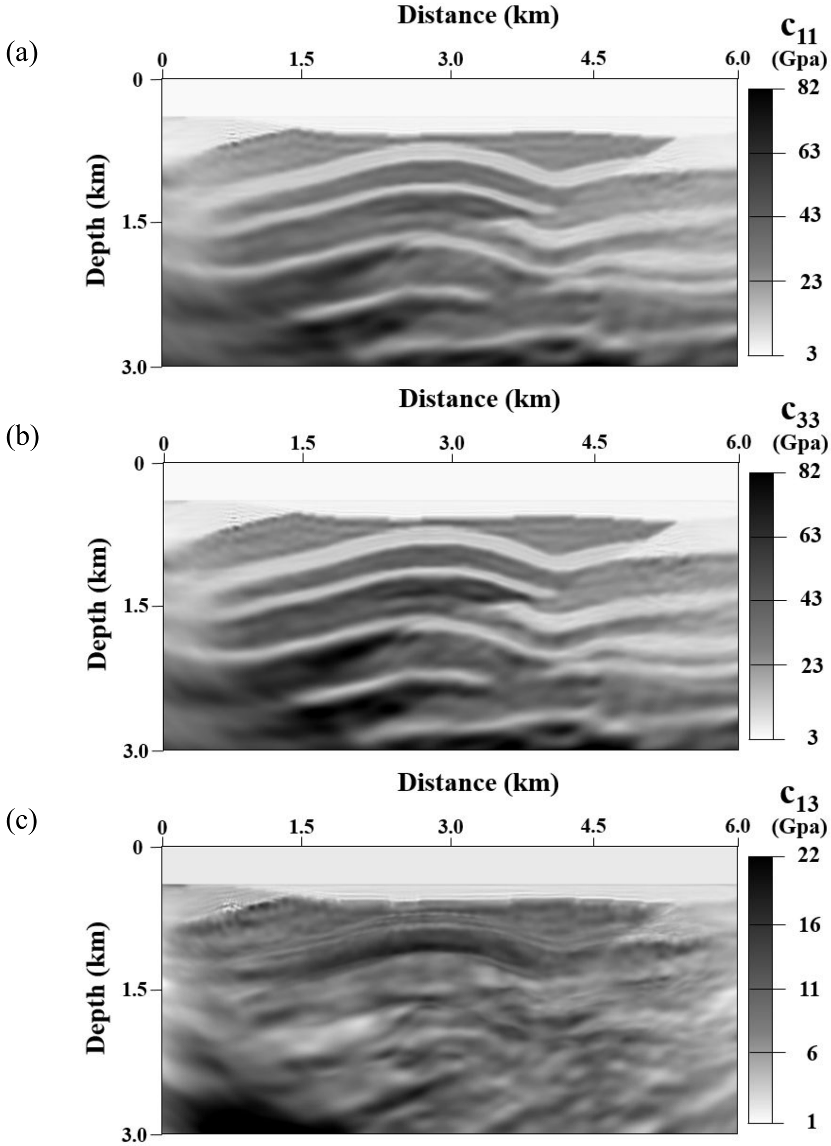


Figure C-4 The recovered structures of (a) c_{11}, (b) c_{33}, (c) c_{13}, (d) c_{44} and (e) density using the VTIPG-II.

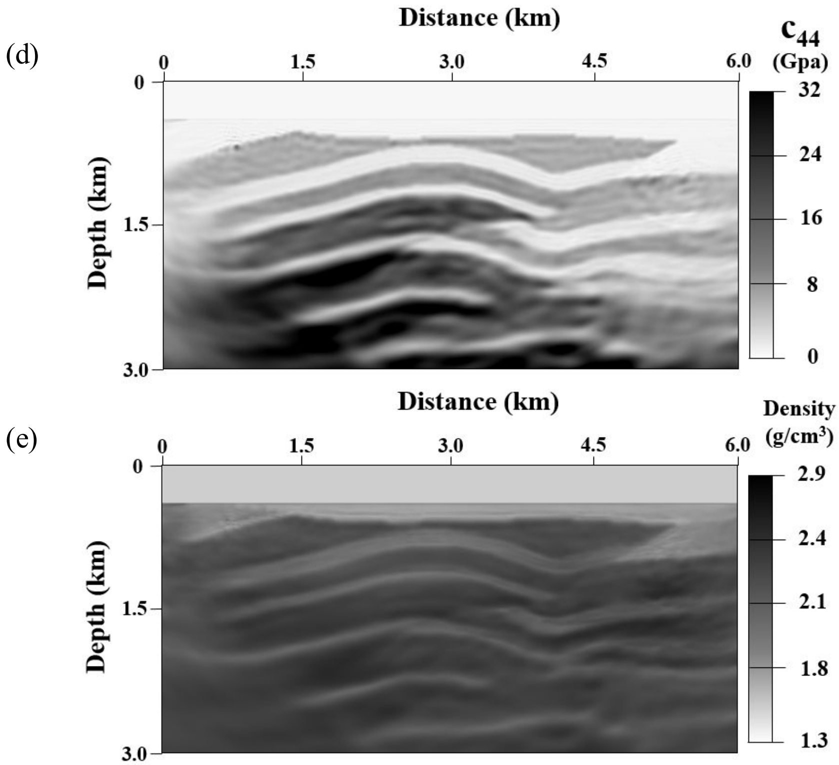


Figure C-4 (Continued)

초 록

본 연구에서는 기존의 다단계 파형역산 전략에 대한 대안으로써, 다변수 추출을 위한 다양한 주파수 대역을 갖는 탄성파 탐사 자료의 주파수 영역 파형 역산 전략을 제안한다. 먼저 가상송신원에 대한 분석을 통하여, 탄성파 완전파형역산에서 포아송비를 이용하여 지하매질을 변수화하는 것이 다변수 추출에 가장 효율적임을 확인한다. 또한 현장 자료의 스펙트럼 정보를 이용하여 각 주파수 별 최대급경사 방향에 가중치를 줌으로써 지하 지층의 두께를 대략적으로 반영하여 파형 역산의 분해능을 조절하는 가중치 기법을 제안한다. 다음으로 현장 잡음이 파형 역산에 미치는 영향을 줄이기 위하여, 송신원 별 현장 자료의 중합을 통해 현장 자료의 신호 대 잡음비를 대략적으로 추정하여 각 주파수 별 최대급경사 방향을 필터링하는 잡음 제거 함수를 제안한다. 마지막으로 기존 레벤버그-마쿼트 방법의 재해석을 통해, 주파수 영역 파형역산의 수렴성과 잡음 안정성을 증대시킬 수 있는 심도 조정 역산 전략을 제안한다. 본 논문에서 제안하는 4가지 새로운 파형역산 기법을 다양한 수치예제에 적용함으로써, 광대역 탄성파 탐사 자료로부터의 다변수 추출이 보다 효율적이고 정확함을 확인한다.

주요어: (주파수 영역 파형 역산, 변수화 방법, 가중치 기법, 잡음제거 함수, 레벤버그-마쿼트 방법)

학 번: (2010-23339)