Geophysical Journal International Advance Access published August 22, 2016

Efficient Edge-Guided Full Waveform Inversion by Canny Edge Detection and Bilateral Filtering Algorithms

Shiming Xiang[1,2] and Haijiang Zhang*[1,2]

1. Wantai Microseismic Lab of School of Earth and Space Sciences, University of Science and Technology of China, Hefei 230026, China

2. Laboratory of Seismology and Physics of Earth's Interior, University of Science and Technology of China, Hefei 230026, China.

*E-mail: zhang11@ustc.edu.cn

Geophysical Journal International

Revised on

August 16, 2016

SUMMARY

It is known full waveform inversion (FWI) is generally ill-conditioned and various strategies including preconditioning and regularizing the inversion system have been proposed to obtain a reliable estimation of the velocity model. Here we propose a new edge-guided strategy for FWI in frequency domain to efficiently and reliably estimate

velocity models with structures of the size similar to the seismic wavelength. The edges of the velocity model at the current iteration are first detected by the Canny edge detection algorithm that is widely used in image processing. Then the detected edges are used for guiding the calculation of FWI gradient as well as enforcing edge-preserving total variation (TV) regularization for next iteration of FWI. Bilateral filtering is further applied to remove noise but keep edges of FWI gradient. The proposed edge-guided FWI in the frequency domain with edge-guided TV regularization and bilateral filtering is designed to preserve model edges that are recovered from previous iterations as well as from lower frequency waveforms when FWI is conducted from lower to higher frequencies. The new FWI method is validated using the complex Marmousi model that contains several steeply dipping fault zones and hundreds of horizons. Compared to FWI without edge guidance, our proposed edge-guided FWI recovers velocity model anomalies and edges much better. Unlike previous image-guided FWI or edge-guided TV regularization strategies, our method does not require migrating seismic data, thus is more efficient for real applications.

1 INTRODUCTION

For oil/gas or geothermal reservoir imaging, accurately characterizing subsurface stratigraphic structures and edges is extremely important because they may provide paths for fluids or confine the reservoir boundaries (e.g., Marfurt et al., 1998; Bahorich and Farmer, 1995; Zheng et al., 2014). For subsurface exploration, great

efforts have been made on the development and application of various discontinuity detection algorithms (e.g., Bahorich and Farmer, 1995; Tingdahl and de Rooij, 2005; Lu et al., 2005). Huang et al. (2011) demonstrated that reverse-time migration has the potential to image steep fault zones for geothermal exploration, but the method requires an accurate background velocity model to start with.

Seismic full waveform inversion (FWI) has shown to be able to determine high-resolution velocity model of the reservoir with a spatial resolution on the order of the seismic wavelength and thus to better characterize the model discontinuities or model edges (e.g., Pratt et al., 1998; Plessix, 2009;Virieux and Operto, 2009). Compared with the travel-time tomography method, FWI seeks a better velocity model by matching the recorded seismic waveform with synthetic waveform, in phase and/or amplitude (Virieux and Operto, 2009). FWI usually includes two parts: forward waveform modeling to calculate synthetic waveforms based on the current model and back propagation of the waveform residuals to calculate gradient for updating the model by some inversion algorithms. The forward waveform modeling relies on the numerical solvers of the acoustic or elastic wave equations by using numerical methods such as finite element, finite difference, or integral equation. For the inversion part, nonlinear inversion algorithm such as the nonlinear conjugate gradient (NLCG) method or the Gauss-Newton method is used to update the inverted model iteratively. Because of the large amount of seismic waveform data and the large size of the velocity model, it is usually computation intensive for forward modeling and inversion in FWI.

It is well known that FWI is a highly nonlinear inverse problem and the solution could be unstable (Tarantola, 1984; Mora, 1987; Pratt et al., 1998; Guitton et al., 2012). For regularizing the inversion system, quadratic regularization such as Tikhonov regularization is computationally easy to be applied (Aster et al., 2013). However, this kind of regularization applies homogeneous smoothing to the inverted model and, consequently, it will tend to blur sharp model boundaries, or model edges. In comparison, non-quadratic regularization such as total variation (TV) can better recover model edges by enforcing sparsity on the model gradients (Aster et al., 2013). TV regularization has been proposed and successfully applied to better determine subsurface model edges in seismic imaging (Youzwishen and Sacchi, 2006; Anagaw and Sacchi, 2011).

To more reliably and accurately determine the velocity model and especially its edges, an image-guided FWI strategy was developed, in which a sparse model space consisting of far fewer parameters is inverted and the finely and uniformly sampled model that can fit the recorded waveform data is reconstructed by image-guided interpolation (Ma et al., 2012; Ma and Hale, 2012). This image-guided FWI involves first migrating seismic waveform data to get the model structure information and then uses that to guide the calculation of FWI gradient so that "the image-guided gradient preserves the structural features in the model" during velocity inversion (Ma et al., , 2012). The image-guided FWI has shown to be able to accurately recover the velocity model and partly avoid cycle skipping problem of FWI (Ma et al., 2012). Different from image-guided FWI, Lin and Huang (2013) developed an FWI method with

edge-guided TV regularization so that model grid points on the edges are not penalized by the regularization. Similar to the image-guided FWI, reverse-time migration is still needed to retrieve model edges iteratively with FWI steps, which can be computationally very expensive.

Different from the FWI strategy of Lin and Huang (2013), Guitton (2012) proposed an FWI method with blocky regularization schemes that does not require another image for guiding the regularization. Instead, velocity discontinuities are first extracted from the current model using gradient operators and then the L_1 norm or Cauchy function is used to enforce the blockiness in the model space (Guitton. 2012). By doing this, model edges can be automatically detected and enhanced along with FWI iterations. By combining the advantageous features of the above FWI methods, here we propose a new edge-guided FWI method that does not require the step of seismic migration for guiding FWI. Instead, model edges are extracted directly from the inverted models during iterations of FWI. They are then used for guiding the calculation of FWI gradient in a way different from Ma et al. (2012) and at the same time guiding the edge-preserving TV regularization following Lin and Huang (2013). Therefore, our proposed FWI strategy avoids the computational burden of seismic migration but keeps the advantages of edge-guided calculation of gradient and TV regularization. In order to extract robust edge information from the model, we adopt Canny edge detection algorithm that are widely used in image processing community (Canny, 1986). We also apply bilateral filtering (Tomasi and Manduchi, 1998; Bhonsle et al., 2012) to remove some "noise" in FWI gradient to make the inversion

more stable.

In the following sections, we will first review Canny edge detection and bilateral filtering methods. Then we introduce FWI in frequency domain and different regularization strategies including L_2 norm regularization, L_1 norm regularization and TV regularization. Edge-guided regularization and edge-guided FWI along with bilateral filtering will also be introduced. Finally synthetic data based on the complex Marmousi velocity model are used to compare different FWI strategies and to validate the improved capability of our new FWI method.

2 Canny edge detection and edge enhancement by bilateral filtering

2.1 Canny edge detection

The Canny edge detection algorithm is well known in image processing to detect edges, or discontinuities in an image (Canny, 1986). Here we employ the multi-step Canny edge detection procedure to find edges by picking local maxima of the model gradients and detect edges by two thresholds (Juneja and Sandhu, 2009). Only the strong edges and weak but connected edges are selected. Therefore, this method is not readily affected by noise interference, and is able to detect real weak edges. The Canny edge detection algorithm consists of five steps (Juneja and Sandhu, 2009):

(i) Apply Gaussian filter to smooth the model;

(ii) Search for the intensity gradients of the smoothed model;

(iii) Make use of non-maximum suppression to get rid of spurious edges;

(iv) Apply the double thresholding to determine potential edges;

(v) Track edges by the BLOB analysis for suppressing all the other edges that are not connected to strong edges or weak edges.

Next we describe these 5 steps in detail, as follows.

(i) Smoothing the model by a Gaussian filter

A two-dimensional Gaussian function is defined as:

$$\mathrm{G}(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad (1)$$

where σ is the standard deviation. For the input 2D model $m(i,j)$, the smoothed model $S(i,j)$ can be obtained by convolving it with the Gaussian smoothing filter, as follows:

$$S(i,j) = G(i,j) * m(i,j), \quad (2)$$

where the star symbol represents the convolution operator.

(ii) Finding the gradients of the smoothed model

For the smoothed model $S(i,j)$, a 2 by 2 window is used to calculate the gradient magnitude and direction by first calculating the two components $P(i, j)$ and $Q(i , j)$ of the gradient vector as follows:

$$P(i,j) \approx (S(i+1,j) - S(i,j) + S(i+1,j+1) - S(i,j+1))/2, \quad (3)$$

$$Q(i,j) \approx (S(i,j+1) - S(i,j) + S(i+1,j+1) - S(i+1,j))/2, \quad (4)$$

Then gradient amplitude and direction angle can be calculated by:

$$M(i,j) = \sqrt{P(i,j)^2 + Q(i,j)^2}, \quad (5)$$

$$\theta(i,j) = \arctan\frac{Q(i,j)}{P(i,j)}, \quad (6)$$

The gradient amplitude and the direction angle associated with a local gradient maximum give the edge strength and direction, respectively.

(iii) Suppressing non-maximum gradients

Some model gradients are non-maximum and thus some extracted edges are suspicious. For this reason, we further suppress all the gradient values to zero except for the local maximum, where the model has the sharpest change. This is done by comparing the edge strength at each model pixel with neighboring edges in both the positive and negative gradient directions. If edge strength at the current pixel is larger, its value will be preserved, otherwise it will be suppressed (Khan, 2012).

(iv) Double Thresholding

After applying non-maximum gradient suppression, the edge pixels could represent the real edges of the model. However, due to the noise in the model there are likely some erroneous edge pixels. Therefore it is essential to remove the weak gradient values and preserve the higher gradient values in the edge pixels. For this purpose, a high threshold value *A* and a low threshold value *B* are set. If the edge strength at one edge pixel is higher than *A*, it is designated as the strong edge pixel. If the edge strength at one edge pixel is smaller than *A* but greater than *B*, it is designated as the weak edge pixel. If the edge strength at one edge pixel is smaller than *B*, it will be suppressed. The double thresholding values are empirically determined for different models.

(v) Edge tracking by the BLOB analysis

Compared to strong edges that are only caused by true edges in the original model, weak edges can correspond to either true edges or noise/small model variations and only those that are connected to strong edges can be real. This is based on the assumption that some small variations like noise in the model are unable to

create strong edges. The Binary Large OBject (BLOB) analysis can be applied to implement the edge tracking (Dahab and Ghoniemy, 2012). BLOB generally refers to a group of connected pixels in a binary image, which represents only "large" objects of a certain size. Only BLOBs containing at least one strong edge pixel are preserved, while other BLOBs are suppressed. For the details of BLOB analysis, please refer to Dahab and Ghoniemy (2012).

We apply the above Canny edge detection algorithm to the Marmousi model (Figure 1a). It shows that edges in the original complex model can be well detected (Figure 1b). We use a Hanning window to smooth the original velocity model to generate the starting model for FWI (Figure 1c). By smoothing, a lot of model details are lost and only major edges are kept (Figure 1d). The detected edges are represented in a matrix with elements of 1 or 0, where element 1 indicates the location of edge pixel. In this way, it is easy to use the edge pixel matrix to extract the edges.

2.2 Enhancing model edges by edge-preserving bilateral filtering

It is inevitable that there exist noise and small variations in a model or an image. To get a reliable estimation of model edges, filtering is necessary to remove edges related to noise but to keep effective edges (Miropolsky and Fischer, 2004). Bilateral filtering is an edge-preserving smoothing scheme that has been widely used in image processing (Tomasi and Manduchi, 1998; Bhonsle et al., 2012). The basic idea of bilateral filtering is to apply traditional filtering such as Gaussian filtering in the

smooth domain of an image, but to apply range filtering around the model edges.

For an image pixel (i_0, j_0), similar to the definition in Equation (1), the weighting function related to Gaussian filtering is defined as:

$$w_s(i,j) = e^{-\frac{|x_i - x_{i0}|^2 + |x_j - y_{j0}|^2}{2s_s^2}}, \quad (7)$$

where (x_i, y_j) are coordinates of neighboring pixels, which are generally defined as those located within 1/10 of the model dimension around the pixel (i_0, j_0) (Tomasi and Manduchi, 1998). The weighting function related to range filtering is similarly defined as:

$$w_r(i,j) = e^{-\frac{|m(i_0,j_0) - m(i,j)|^2}{2s_r^2}}, \quad (8)$$

where $m(i_0, j_0)$ and $m(i, j)$ are pixel values for current and neighboring pixels. It can be seen that Gaussian filtering is related to the distance from the current pixel with higher weight for the shorter distance (Figure 2a). In comparison, range filtering is related to pixel values in that for smooth image domains filtering weights are larger while across the edges filtering weights are smaller (Figure 2b). Bilateral filtering is to combine the Gaussian filtering and the range filtering together and the related weighting function is defined as:

$$b(i,j) = w_s(i,j)w_r(i,j), \quad (9)$$

Around the sharp edge of the model, range filtering weights also have a sharp contrast across the edge thus it can be preserved (Figure 2c). For more details about bilateral filtering, please see Tomasi et al. (1998). We test bilateral filtering to a model with edges (Figure 2d). When adding random noise to the model, the edges are

blurred to some extent (Figure 2e). When applying Gaussian filtering, although noise can be removed the edges in the model are distorted (Figure 2f). In comparison, bilateral filtering not only removes noise from the model but also preserves model edges (Figure 2g).

3 Edge-guided full waveform inversion in the frequency domain

In this section, we first introduce FWI in the frequency domain. Then we introduce various regularization strategies to stabilize the FWI system. Finally we introduce the application of bilateral filtering to remove noise in the FWI gradient and the modulation of the FWI gradient and the TV regularization by detected edges.

3.1 Frequency-Domain Full-Waveform Inversion

In this study, we model the seismic wave propagation in 2D using the acoustic wave equation in the frequency domain, as follows,

$$\frac{\omega^2}{k(x,z)}p(x,z,\omega)+\frac{\partial}{\partial x}\left(\frac{1}{\rho(x,z)}\frac{\partial p(x,z,\omega)}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{1}{\rho(x,z)}\frac{\partial p(x,z,\omega)}{\partial z}\right)=-s(x,z,\omega)\ , \tag{10}$$

where $p(x, z, \omega)$ is the pressure, $\rho(x,z)$ is the mass density, ω is the angular frequency, $s(x, z, \omega)$ is the source term, and $k(x,z)$ is the complex bulk modulus. It is noted that the P-wave velocity Vp can be expressed as $Vp = \sqrt{\frac{k}{\rho}}$. Equation (10) can be further recast into a matrix form as

$$\mathbf{Ap} = \mathbf{s}\ , \tag{11}$$

where the complex-valued impedance matrix **A** depends on the frequency and medium properties, and **p** and **s** are pressure and source vectors, respectively.

We solve Equation (11) using the frequency-domain finite-difference approach with a fourth order accuracy. Equation (11) is discretized by using rectangular cells and the discrete pressure value is defined at the center of each cell. The 2-D pressure **p** and source **s** fields at one angular frequency ω are stored as vectors of dimension nx×nz with nx and nz the dimensions of the regular FD grid in X and Z directions, respectively. We used the perfectly matched layer (PML) of Berenger (1994) as the boundary condition. After discretization, for each frequency, the linear system of equations is solved by the multifrontal LU decomposition method of Davis and Duff (1997). The details of the forward modeling algorithm can be found in Pan et al. (2012).

FWI aims at minimizing the norm of the data misfit, defined as:

$$J_D(\mathbf{m}) = \frac{1}{2}\sum_{\omega}\sum_{s,r}^{ns,nr}\left\|\mathbf{d}_{s,r}^{cal}(\mathbf{m},\omega) - \mathbf{d}_{s,r}^{obs}(\omega)\right\|_2^2 \ , \tag{12}$$

where **m** is the current velocity model, ω is the angular frequency, and n_s and n_r represent the number of sources and receivers, and $\mathbf{d}^{obs}$ and $\mathbf{d}^{cal}$ are observed and calculated seismic waveforms, respectively. The data misfit norm can also take into account the relative importance for different data points, as follows:

$$J_D(\mathbf{m}) = \Delta\mathbf{d}^{*}\mathbf{W}_d\Delta\mathbf{d} \ , \tag{13}$$

where $\Delta\mathbf{d}$ is the misfit vector between the observed waveform and the calculated waveform, $\Delta\mathbf{d}^{*}$ is the conjugate of $\Delta\mathbf{d}$. $\mathbf{W}_d$ is a weighting operator that scales relative contributions of every component of $\Delta\mathbf{d}$. In the frequency domain, Sourbier et al. (2009) derived the following expression for the gradient of the data misfit objective function:

$$\nabla J_D(\mathbf{m}) = \mathrm{Re}\left\{\mathbf{p}^t[\frac{\partial \mathbf{A}^t}{\partial \mathbf{m}}]\mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*\right\}, \quad (14)$$

where **Re** denotes the real part of a complex number. The gradient of the data misfit objective function is actually computed by zero-lag cross correlation between the forward-propagated wavefield and the back-propagated residual wavefield (Pratt et al., 1998). It is noted that because of the random noise in the data and forward modeling error, the gradients of the data misfit objective function inevitably are contaminated. Therefore it is a general practice for FWI to apply smooth filtering such as Gaussian filtering to remove some random noise in the gradients (Sourbier et al., 2009). In addition, to stabilize FWI, we applied some scaling and preconditioning to the gradients of the data misfit objective function (Sourbier et al, 2009), as follows:

$$\nabla J_D(\mathbf{m}) = (\mathrm{diag}\mathbf{H}_a + \varepsilon I)^{-1}\mathrm{G}_m\mathrm{Re}\left\{\mathbf{p}^t[\frac{\partial \mathbf{A}^t}{\partial \mathbf{m}}]\mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*\right\}, \quad (15)$$

where $\mathrm{diag}\mathbf{H}_a$ denotes the diagonal elements of the weighted approximate Hessian $\mathbf{H}_a$, ε is a damping value, and G_m is a Gaussian filtering operator.

3.2 FWI with first-order Tikhonov and TV regularizations

In practice, the FWI system is highly nonlinear and unstable. To stabilize the FWI system, the objective function also includes some form of regularization in addition to the minimization of the data misfit, as follows:

$$J(\mathbf{m}) = J_D(\mathbf{m}) + \lambda J_R(\mathbf{m}), \quad (16)$$

The first term is the data misfit term and the second term is the regularization term. λ is the regularization value to balance the contributions from two terms， which can be estimated as (Guitton et al., 2012):

$$\lambda = \frac{J_D(\mathbf{m}_0)}{T J_R(\mathbf{m}_0)} \quad , \tag{17}$$

where $\mathbf{m}_0$ is the starting model, and T is an empirical value depending on the noise level in the data. The value of T is set between 1 and 10 when the noise level is low and will be greater if the noise level is high.

For the regularization term $J_R(\mathbf{m})$ in Equation (16), an often-used regularization method is known as Tikhonov regularization, which generally enforces the smoothness of the model (Aster et al., 2013). For the first-order Tikhonov regularization, the regularization term can be represented by L_2 norm of model differences along different directions. In this case, the overall objective function can be represented as follows:

$$J_{\mathrm{Tik}}(\mathbf{m}) = J_D(\mathbf{m}) + \lambda \|\mathbf{Lm}\|_2^2 \quad , \tag{18}$$

where $\mathbf{L}$ is the first-order differencing operator along each dimension of the model. Correspondingly, its gradient is

$$\nabla J_{\mathrm{Tik}}(\mathbf{m}) = (\mathrm{diag}\mathbf{H}_a + \varepsilon I)^{-1} \mathrm{G}_m \mathrm{Re}\left\{\mathbf{p}^t [\frac{\partial \mathbf{A}^t}{\partial \mathbf{m}}] \mathbf{A}^{-1} \mathbf{W}_a \Delta \mathbf{d}^*\right\} + \lambda \mathbf{L}^\mathbf{T} \mathbf{Lm} \quad , \tag{19}$$

where $\mathbf{L}^\mathbf{T}$ is the transpose of $\mathbf{L}$.

Another regularization method is known as total variation (TV), which enforces the sparseness of the model gradients by L_1 norm. The objective function with the TV regularization is defined as follows (Aster et al., 2013):

$$J_{\mathrm{Tv}}(\mathbf{m}) = J_D(\mathbf{m}) + \lambda \|\mathbf{Lm}\|_1 \quad , \tag{20}$$

where **L** is the first-order differencing operator along each dimension of the model. By defining **Y**=**Lm,** we can calculate the gradient of the objective function with the TV regularization as follows (Aster et al., 2013):

$$\nabla J_{\mathrm{Tv}}(\mathbf{m}) = (\mathrm{diag}\mathbf{H}_a + \varepsilon I)^{-1}\mathrm{G}_m \mathrm{Re}\left\{\mathbf{p}^t[\frac{\partial \mathbf{A}^t}{\partial \mathbf{m}}]\mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*\right\} + \lambda\mathbf{L}^T\mathbf{Q}\mathbf{L}\mathbf{m}\ , \tag{21}$$

where $\mathrm{Q_{i,j}} = \begin{cases} \frac{1}{|\mathrm{Y_{i,j}}|} & |\mathrm{Y_{i,j}}| \geq \beta \\ \frac{1}{\beta} & |\mathrm{Y_{i,j}}| < \beta \end{cases}$, and β is a small value to avoid the singularity at zero.

First-order Tikhonov regularization enforces the model smoothness by penalizing model variations. This means one must compromise between preserving model edges and suppressing highly oscillatory features likely arising from the inversion of noisy data. In contrast, edge-preserving TV regularization can still enforce the model to be smooth in most areas but without smearing the model edges (or discontinuities).

3.3 FWI with edge-guided TV regularization

For a 2D model, the TV regularization term in Equation (20) can be represented by

$$\|\mathbf{Lm}\|_1 = \sum_{i=1}^{\mathrm{N}_x}\sum_{j=1}^{\mathrm{N}_z}(\left|(\nabla_x\mathbf{m})_{\mathrm{i,j}}\right| + \left|(\nabla_z\mathbf{m})_{i,j}\right|), \tag{22}$$

where $(\nabla_x\mathbf{m})_{\mathrm{i,j}} = m_{i+1,j} - m_{i,j}$ and $(\nabla_y\mathbf{m})_{i,j} = m_{i,j+1} - m_{i,j}$. Following Lin and Huang (2013), we can incorporate the model edge information into the TV regularization. For the current model, at each point (i, j) we can use $e(i,j)$ to indicate whether the point is on the edge extracted by the Canny edge detection algorithm. If the point is on the edge then the value of $e(i,j)$ is 1, otherwise 0. We can then define the edge weighting factor $\mathbf{w}_e$ based on the edge information of the current model, as follows:

$$\mathrm{w}_{e_{i,j}} = \begin{cases} 0 & e(i,j) = 1 \\ 1 & e(i,j) = 0 \end{cases}, \tag{23}$$

After that, we can denote the edge-guided TV regularization as

$$\|\mathbf{W}_e\mathbf{L}\mathbf{m}\|_1 = \sum_{i=1}^{N_x}\sum_{j=1}^{N_z} \mathrm{w}_{e_{i,j}}(\left|(\nabla_x\mathbf{m})_{\mathrm{i,j}}\right| + \left|(\nabla_z\mathbf{m})_{i,j}\right|)\,, \tag{24}$$

The purpose of assigning the zero weight to the points on the edges is to free them from being penalized by the regularization. By doing this, we can keep the obtained model edges from previous iterations of FWI using lower frequencies.

The corresponding objective function for FWI with edge-guided TV regularization (denoted as ETV) is defined as:

$$J_{\mathrm{ETV}}(\mathbf{m}) = J_D(\mathbf{m}) + \lambda\|\mathbf{W}_e\mathbf{L}\mathbf{m}\|_1\ , \tag{25}$$

And the corresponding gradient of the objective function can be expressed as

$$\nabla J_{\mathrm{ETV}}(\mathbf{m}) = (\mathrm{diag}\mathbf{H}_a + \varepsilon I)^{-1}\mathrm{G}_m\mathrm{Re}\left\{\mathbf{p}^t[\frac{\partial\mathbf{A}^t}{\partial\mathbf{m}}]\mathbf{A}^{-1}\mathbf{W}_a\Delta\mathbf{d}^*\right\} + \lambda\mathbf{W}_e\mathbf{L}^{\mathbf{T}}\mathbf{Q}\mathbf{L}\mathbf{m}\ , \tag{26}$$

In the above equations for the gradients of the objective functions for FWI associated with different regularizations, Gaussian filtering ($\mathrm{G_m}$) is applied. Instead of applying the traditional Gaussian filtering to remove the noise from gradients of the data misfit term, we can apply the bilateral filtering to gradients in order to preserve their edges. This is because edges or structures in gradients are very important for guiding FWI toward faster convergence. Figure 3 shows an example of applying bilateral filtering to FWI gradient. It can be seen after bilateral filtering the gradient image has clearer structures, especially for model parameters at greater depths. For the whole model region, the edges are more enhanced and more balanced by bilateral filtering. In Equation (26) for FWI with edge-guided TV regularization, we can

replace Gaussian filtering G_m by bilateral filtering B_m to construct the gradient of the objective function for FWI with edge-guided TV regularization and bilateral filtering (denoted as ETVB), as follows:

$$\nabla J_{\mathrm{ETVB}}(\mathbf{m}) = (\mathrm{diag}\mathbf{H}_a + \varepsilon I)^{-1}\mathrm{B}_m \mathrm{Re}\left\{\mathbf{p}^t[\frac{\partial \mathbf{A}^t}{\partial \mathbf{m}}]\mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*\right\} + \lambda\mathbf{W}_e\mathbf{L}^{\mathbf{T}}\mathbf{Q}\mathbf{L}\mathbf{m}\,, \tag{27}$$

3.4 Edge-Guided FWI with edge-guided TV regularization

For each iteration of FWI, we can obtain current model edges by using the Canny edge detection algorithm. This only requires very little additional computation cost during FWI. The obtained model edges are then used to guide the calculation of the gradients of the data misfit term in the objective function. If the gradients $\nabla J_D(\mathbf{m})$ on points have the same directions as the model edge direction, it indicates that the model updates will tend to enforce the existing model edges. In this case, the corresponding gradient values conform to the related edges and thus will be more weighted. Otherwise, they will be zeroed out. In this way, we can guide FWI to enhance the detected model edges from previous iterations. As a result, model edges at different scales can be well preserved for FWI from lower to higher frequencies. This idea is similar to image-guided FWI proposed by Ma et al. (2012), but one major difference is that our edge-guided FWI uses the current model, instead of the migrated model, to guide next iteration of FWI. Here we use the angle α between the gradient $\nabla J_D(\mathbf{m})$ and the edge direction to determine whether they conform to each other. Here we simply assume that if α is not greater than 45°, then it is regarded that the two vectors conform. In the edge-guided FWI, the gradients of the data misfit term in the

objective function are thus modulated as:

$$\mathrm{E}\{g_{i,j}^{k}\} = \begin{cases} (1+\tau)g_{i,j}^{k} & e(i,j)=1, \alpha \le 45^{\circ} \\ 0 & e(i,j)=1\,, \alpha > 45^{\circ} \\ g_{i,j}^{k} & e(i,j)=0 \end{cases}, \tag{28}$$

where $g_{i,j}^{k}$ is the gradient of the data misfit term at the point (i, j) in the kth iteration, and τ is set as 0.3 to enhance the model edges that are already detected. Correspondingly, for edge-guided FWI with edge-guided TV regularization (denoted as EFWIG), the objective function and the associated gradient are expressed as:

$$J_{\mathrm{EFWIG}}(\mathbf{m}) = J_D(\boldsymbol{m}) + \lambda\|\mathbf{W}_e\mathbf{L}\mathbf{m}\|_1\,, \tag{29}$$

$$\nabla J_{\mathrm{EFWIG}}(\mathbf{m}) = \mathrm{E}\left\{(\mathrm{diag}\mathbf{H}_a + \varepsilon I)^{-1}\mathrm{G}_m \mathrm{Re}\left\{\mathbf{p}^t[\tfrac{\partial \mathbf{A}^t}{\partial \mathbf{m}}]\mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*\right\}\right\} + \lambda\mathbf{W}_e\mathbf{L}^{\mathbf{T}}\mathbf{Q}\mathbf{L}\mathbf{m}\,, \tag{30}$$

Here we stress that the edge-guided modulation of the FWI gradient in Equation (28) is different from the image-guided gradient modulation scheme proposed by Ma et al. (2012), which relies on the image-guided interpolation operator **R** and its adjoint operator $\mathbf{R}^{\mathrm{T}}$ to create the image-guided gradient $\mathbf{R}\mathbf{R}^{\mathrm{T}}\mathbf{g}$.

For our proposed edge-guided FWI in Equation (30), we can replace the traditional Gaussian filtering $\mathrm{G_m}$ by bilateral filtering $\mathrm{B_m}$. The objective function for the edge-guided FWI with edge-guided TV regularization and bilateral filtering (denoted as EFWIB) is the same as that shown in Equation (29), but for the calculation of FWI gradient the Gaussian filtering G_m is replaced by bilateral filtering B_m, as follows:

$$\nabla J_{\mathrm{EFWIB}}(\mathbf{m}) = \mathrm{E}\left\{(\mathrm{diag}\mathbf{H}_a + \varepsilon I)^{-1}\mathrm{B}_m \mathrm{Re}\left\{\mathbf{p}^t[\tfrac{\partial \mathbf{A}^t}{\partial \mathbf{m}}]\mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*\right\}\right\} + \lambda\mathbf{W}_e\mathbf{L}^{\mathbf{T}}\mathbf{Q}\mathbf{L}\mathbf{m}\,. \tag{31}$$

4 Test of different FWI strategies on the Marmousi model

For testing the edge-guided FWI method, we used the Marmousi velocity model (Figure 1a) to generate synthetic waveforms with a Ricker wavelet having the central frequency of 16 Hz. The Marmousi velocity model contains 158 horizontally layered horizons and several steeply dipping faults. By using the Canny edge detection algorithm, we can extract edges corresponding to these horizons and faults (Figure 1b). The synthetic data set consists of 199 shots and 200 receivers, with shot and receiver intervals of 100 m, respectively. A small amount of Gaussian distributed random noise was added to synthetic waveform data, and the signal to noise ratio (SNR) is 30. Figure 4 shows the comparison of frequency spectrum for one trace with and without adding the noise. This synthetic test is to see how well the edges of the original Marmousi model can be recovered by using the edge-guided FWI starting from a smooth model (Figure 1c).

For FWI in the frequency domain, 7 discrete frequencies of 3.5, 5, 7.6, 12.5, 15, 17.5, and 20.6 Hz were selected for the inversion and it is carried out sequentially from lower to higher frequencies. At each frequency, there are 30 iterations of model updating by FWI. Therefore there are 210 iterations in total. For comparison, we test six FWI strategies including (1) 1st strategy: FWI with Tikhonov regularization, (2) 2nd strategy: FWI with TV regularization, (3) 3rd strategy: FWI with edge-guided TV regularization, (4) 4th strategy: FWI with edge-guided TV regularization and bilateral filtering, (5) 5th strategy: edge-guided FWI with edge-guided TV regularization and Gaussian filtering, and (6) 6th strategy: edge-guided FWI with edge-guided TV

regularization and bilateral filtering, respectively. For the first 3 and 5th strategies, Gaussian filtering is applied to remove the noise in the FWI gradient. while the 4th and 6th strategies apply bilateral filtering to the FWI gradient. Figure 5 shows the data misfit curves along with iterations for different FWI strategies. It can be seen that FWI with TV regularization has lower data misfit values than FWI with the Tikhonov regularization (Figure 5). For the final velocity models from the two methods, the root-mean-square (RMS) data misfit values are 1.95E-004 for FWI with Tikhonov regularization and 1.77E-004 for FWI with TV regularization, respectively. Compared with FWI without edge-guided regularization, FWI with edge-guided TV regularization has a smoother data misfit curve with a lower RMS data misfit value of 0.61E-004 for the final velocity model, indicating the inversion process more monotonically converges to the final velocity model that better fits for the observed waveforms (Figure 5). When incorporating the model edge information into gradients ∇J_D, edge-guided FWI with edge-guided TV regularization and Gaussian filtering results in a final velocity model with the RMS data misfit value of 0.58E-004, slightly better compared to the case without using edge information to modulate gradients ∇J_D (Figure 5). When applying bilateral filtering instead of Gaussian filtering to FWI gradient, the 4th FWI strategy (FWI with edge-guided TV regularization and bilateral filtering) results in a final velocity model with the RMS data misfit value of 0.26E-004, having a better performance in fitting the observed data than 3rd and 5th FWI strategies. When combining bilateral filtering and edge-guided TV regularization with modulation of FWI gradient by model edges, for the final velocity model from

6th FWI strategy (edge-guided FWI with edge-guided TV regularization and bilateral filtering), the RMS data misfit value is 0.07E-004, which is approximately 95% lower than FWI with Tikhonov or TV regularization methods. For FWI methods with edge guidance, the data misfit function values decrease more quickly and monotonically along with iterations, with edge-guided FWI with edge-guided TV regularization and bilateral filtering fitting the data best (Figure 5).

By comparing the final inverted velocity models from FWI with Tikhonov or TV regularization, both methods did a good job in recovering complex horizons and faults in the Marmousi model (Figure 6a, c). They bear much similarity to the true Marmousi model and they also look very similar to each other while the TV regularization recovers model features slightly better (Figure 6b,d). We can also compare models and their edges with true ones for different methods. It is very clear that model edges are also well recovered for both FWI methods and the TV regularization recovers model edges slightly better (Figure 7). For FWI with edge-guided TV regularization, both model edges and model anomalies are better recovered (Figures 6e-f, 7). By further incorporating edge information into gradients ∇J_D, edge-guided FWI performs even better in recovering model edges and model anomalies (Figure 6i-j, 7). Especially at the lower left corner region, the model edges are recovered much better with edge-guided FWI (Figure 6j). When applying bilateral filtering instead of Gaussian filtering to gradient related to the data misfit term in the objective function, the 4th strategy (FWI with edge-guided TV regularization and bilateral filtering) can better recover model edges and model anomalies (Figure 6g-h)

than the 3rd strategy. As already indicated by the best data fitting, edge-guided FWI with edge-guided TV regularization performs the best in recovering model edges and model anomalies due to its ability of preserving the edge information in gradients ∇J_D (Figure 6k-l, 7).

To further compare models inverted using different FWI strategies with the true model, we extracted one horizontal and one vertical velocity profiles (Figure 8). It shows that around model boundaries, all FWI strategies fail to recover model edges due to the limited coverage of the data (Figure 8). However, for the areas with good data coverage, all FWI methods can recover the model anomalies well. In comparison, edge-guided FWI with edge-guided TV regularization and bilateral filtering performs the best in recovering the velocity values as well as velocity contrasts (Figure 8). For example, at depth around 2.6 km along the vertical profile, FWI with Tikhonov or TV regularization fails to recover the model edge while edge-guided FWI with edge-guided TV regularization and bilateral filtering successfully resolves small velocity changes in this place.

We further compare waveforms simulated from different models with true waveforms (Figure 9). For the initial model, waveform fitting between observed and synthetic waveforms is very poor (Figure 9a). With the final velocity models inverted from FWI with Tikhonov or TV regularization, some parts of waveforms can be matched to some extent between true and observed waveforms but still there is considerable waveform mismatch (Figure 9b, c). This reflects the fact that FWI with Tikhonov or TV regularization recovers model edges and velocity anomalies to some

extent but still there are some considerable differences between true and recovered velocity models. In comparison, for the velocity models inverted using FWI with edge-guided TV regularization or edge-guided FWI with edge-guided TV regularization and Gaussian filtering, it can be clearly seen that the waveforms are better matched in both phase and amplitude between true and simulated waveforms (Figure 9d, f). For the FWI strategy with bilateral filtering to the FWI gradient, the waveforms match better even when model edges are not used for modulating the FWI gradient (Figure 9e). The waveforms match the best for the final velocity model inverted from edge-guided FWI with edge-guided TV regularization and bilateral filtering (Figure 9g), as already suggested by the lowest data misfit and the best model fit.

4 Discussions

The test on the complex Marmousi model shows that the Canny edge detection is able to extract reliable model edges. The extracted model edge information can subsequently be used to modulate gradients ∇J_D of FWI and TV regularization. The concept is to preserve model edges that are already inverted from previous iterations of FWI. The edge-guided FWI is also designed to preserve model edges that are determined using lower frequency waveforms when the velocity model is updated by higher frequency waveforms. This can be clearly illustrated by velocity models and the associated edges inverted sequentially from lower to higher frequency seismic waveforms (Figure 10). For the velocity model inverted using seismic waveforms up

to 3.5 Hz, it is smoother and only major edges are recovered compared to the true Marmousi model (Figure 10a). When including higher frequency waveforms up to 7.6 Hz, more details of the Marmousi model are recovered and weaker model edges also show up (Figure 10b). With higher frequency waveforms up to 15 Hz, more model details and edges are inverted (Figure 10c). It is noticed that model edges recovered from edge-guided FWI using lower frequency waveforms are preserved in the model inverted using higher frequency waveforms (Figure 10).

When comparing data misfit curves for different FWI strategies, it is clear that the curves corresponding to the edge-guided FWI with edge-guided TV regularization and bilateral filtering have a more quickly and monotonically decreasing trend than other curves (Figure 5). Furthermore, the curve associated with edge-guided FWI with edge-guided TV regularization and bilateral filtering converges to a velocity model with the lowest data misfit. Without edge-guidance, the data misfit curves corresponding to FWI with Tikhonov regularization or TV regularization do not decrease as monotonically as FWI with edge-guided TV or edge-guided FWI. This is simply because some model edges resolved in previous iterations or by lower frequency waveforms could be lost in later iterations or when the model is updated by higher frequency waveforms. As a result the data misfit could become worse and the inversion could be trapped in local minima. FWI with the edge-guided TV regularization can partly mitigate this issue by better preserving model edges from lower to higher frequencies and the inversion can more quickly converge to a velocity model with a lower data misfit (Figures 5, 7). By further modulating FWI data misfit

term gradients ∇J_D with current model edges, edge-guided FWI can better preserve model edges inverted in successive iterations from lower to higher frequencies and guides the inversion to quickly converge to a better model (Figure 5). Compared to edge-guided FWI with Gaussian filtering, edge-guided FWI with bilateral filtering significantly improves the performance of waveform inversion with a quicker and more monotonic convergence towards a better model (Figure 5). This is mainly because bilateral filtering can better strengthen edge information in gradient when removing the noise from them. In this sense, edge-guided FWI with bilateral filtering is helpful for alleviating the issue of cycle skipping caused by local minima in the objective function.

From the tests of different FWI strategies on the Marmousi model, it can be seen that the three schemes of edge-guided TV regularization, bilateral filtering of the FWI gradient, and edge-guided modulation of the FWI gradient all contribute to improve the performance of FWI but at different levels. Therefore, the FWI strategy combining these three schemes perform the best in recovering velocity model anomalies and edges as well as fitting the observed data. These tests show that, for edge-guided FWI with edge-guided TV regularization and bilateral filtering, if the initial model is reasonable, then it can converge more quickly to a better model than the conventional FWI.

Compared to the image-guided FWI (Ma et al., 2012; Ma and Hale, 2012) and edge-guided TV regularization proposed by Lin and Huang (2013), our proposed edge-guided FWI with edge-guided TV regularization and bilateral filtering in this

study does not rely on the process of seismic migration to provide the model structure information for guiding FWI or TV regularization. Instead, similar to Guitton (2012) our proposed strategy uses the current model for extracting model edge information to guide the calculation of gradient ∇J_D and TV regularization. As a result, our proposed edge-guided FWI with edge-guided TV regularization method is more efficient and more applicable to the real data, where seismic migration could be challenging due to the lack of a reliable velocity model.

5 Conclusions

We have developed a novel edge-guided FWI method in the frequency domain that iteratively employs the edge information of the velocity model at current iteration to modulate FWI data misfit term gradient ∇J_D and TV regularization term for next iteration of inversion. Compared to the edge-guided strategies of Ma and Hale (2012) and Lin and Huang (2013), our edge-guided FWI does not need the process of seismic migration to provide the model edge information and therefore is more efficient and more applicable to the real data. Canny edge detection is applied to reliably extract model edges, which are then used to guide the calculation of gradient ∇J_D and TV regularization term. The purpose is to preserve edge information inverted at successive iterations from lower to higher frequency waveforms. Bilateral filtering is further employed to subdue noise while keep structures in gradient ∇J_D. We have validated the capability of our new edge-guided FWI method for accurate reconstruction of hundreds of horizons and steeply dipping reflectors in the Marmousi

model. Compared to FWI without edge guidance, edge-guided FWI converges more quickly and monotonically to a better velocity model. Therefore, our proposed edge-guided FWI method is a powerful tool for accurate velocity estimation.

ACKNOWLEDGEMENTS

This study was financially supported by the Natural Science Foundation of China under grant number 41474039. We are grateful for the constructive comments from Antoine Guitton and an anonymous reviewer, and the discussions with Wei Zhang and Qinya Liu, which are very helpful for improving the paper.

REFERENCES

1. Anagaw, A.Y. &Sacchi, M.D., 2012. Edge-preserving seismic imaging using the total variation method, Journal of Geophysics and Engineering, 9, 138.

2. Aster, R.C., Borchers, B., and Thurber, C.H. (2013). Parameter estimation and inverse problems (2nd edition), Academic Press, Waltham, MA, United States, 360 pp.

3. Bahorich, M. & Farmer, S., 1995. 3-D seismic discontinuity for faults and stratigraphic features: The coherence cube, The leading edge, 14, 1053-1058.

4. Berenger, J.-P., 1994. A perfectly matched layer for the absorption of electromagnetic waves, Journal of computational physics, 114, 185-200.

5. Bhonsle, D., Chandra, V. & Sinha, G., 2012. Medical image denoising using bilateral filter, International Journal of Image, Graphics and Signal Processing (IJIGSP), 4, 36.

6. Canny, J., 1986. A computational approach to edge detection, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 679-698.

7. Dahab, D.A., Ghoniemy, S.S. & Selim, G.M., 2012. Automated brain tumor detection and identification using image processing and probabilistic neural network techniques, International journal of image processing and visual communication, 1, 1-8.

8. Davis, T.A. & Duff, I.S., 1997. An unsymmetric-pattern multifrontal method for sparse LU factorization, SIAM Journal on Matrix Analysis and Applications, 18, 140-158.

9. Guitton, A., Ayeni, G. & Díaz, E., 2012. Constrained full-waveform inversion by model reparameterization , Geophysics, 77, R117-R127.

10. Guitton, A., 2012. Blocky regularization schemes for Full‐Waveform Inversion, Geophysical Prospecting, 60, 870-884.

11. Huang, L., Albrecht, M., Kaufman, G., Kelley, S., Rehfeldt, K. & Zhang, Z., 2011. Imaging Faults with Reverse-Time Migration for Geothermal Exploration, Geophysics, 64, 1524-1534.

12. Khan, M.S.I., 2012. Implementation of Edge & Shape Detection Techniques and their Performance Evaluation, Ryerson University Toronto, Canada.

13. Juneja, M. & Sandhu, P.S., 2009. Performance evaluation of edge detection techniques for images in spatial domain: methodology, 1, 614-621.

14. Lin Y, & Huang L. 2013. Accurate Velocity Estimation for steep Fault Zones Using Full-Waveform Inversion With Edge-Guided regularization, GRC Transactions, 37, 1031-1036.

15. Lu, W., Li, Y., Zhang, S., Xiao, H. & Li, Y., 2005. Higher-order-statistics and supertrace-based coherence-estimation algorithm, Geophysics, 70, P13-P18.

16. Ma, Y., Hale, D., Gong, B. &Meng, Z., 2012. Image-guided sparse-model full waveform inversion, Geophysics, 77, R189-R198.

17. Ma, Y. & Hale, D., 2012. Quasi-Newton full-waveform inversion with a projected Hessian matrix, Geophysics, 77, R207-R216.

18. Marfurt, K.J., Kirlin, R.L., Farmer, S.L. &Bahorich, M.S., 1998. 3-D seismic attributes using a semblance-based coherency algorithm, Geophysics, 63, 1150-1165.

19. Miropolsky, A. & Fischer, A., 2004. Reconstruction with 3D geometric bilateral filter. in Proceedings of the ninth ACM symposium on Solid modeling and applications, pp. 225-229Eurographics Association.

20. Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multi-offset seismic data, Geophysics, 52, 1211-1228.

21. Pan, G., Abubakar, A. &Habashy, T.M., 2012. An effective perfectly matched layer design for acoustic fourth-order frequency-domain finite-difference scheme, Geophysical Journal International, 188, 211-222.

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

23. Pratt, R.G., Shin, C. & Hick, G., 1998. Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion, Geophysical Journal International, 133, 341-362.

24. Sourbier, F., Operto, S., Virieux, J., Amestoy, P. & L'Excellent, J.-Y., 2009. FWT2D: A massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data—Part 1: Algorithm, Computers & Geosciences, 35, 487-495.

25. Tingdahl, K.M. & De Rooij, M., 2005. Semi-automatic detection of faults in 3D seismic data, Geophysical Prospecting, 53, 533-542.

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

27. Tomasi, C. & Manduchi, R., 1998. Bilateral filtering for gray and color images, Proceedings of the 1998 IEEE international conference on Computer Vision, Bombay, India, pp. 839-846.

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

29. Youzwishen, C. &Sacchi, M., 2006. Edge preserving imaging, Journal of Seismic Exploration, 15, 45.

30. Zheng, Z.H., Kavousi, P. & Di, H.B., 2014. Multi-Attributes and Neural Network-Based Fault Detection in 3D Seismic Interpretation, Advanced Materials Research, 838, 1497-1502.

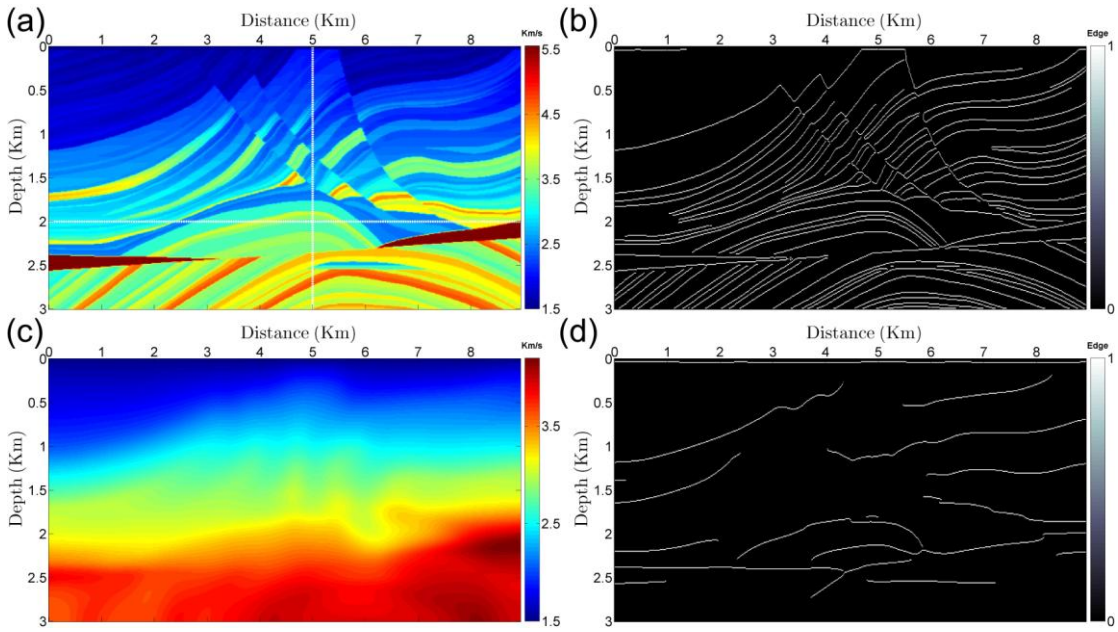


Figure 1. Original (a) and smoothed (c) Marmousi velocity models and their corresponding model edges (b and d) extracted by Canny edge detection algorithm.

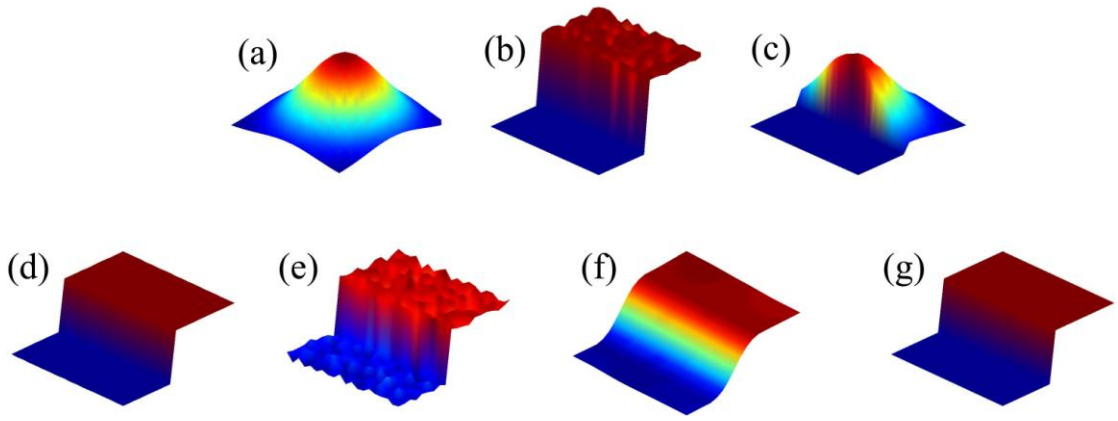


Figure 2. Test of bilateral filtering on an image with a sharp edge. (a) Gaussian filter around the edge; (b) Range filter around the edge; (c) Bilateral filter around the edge. (d) Input image with a sharp edge; (e) Image added with some random noise; (f) Filtered image with Gaussian filtering; (g) Filtered image with bilateral filtering;

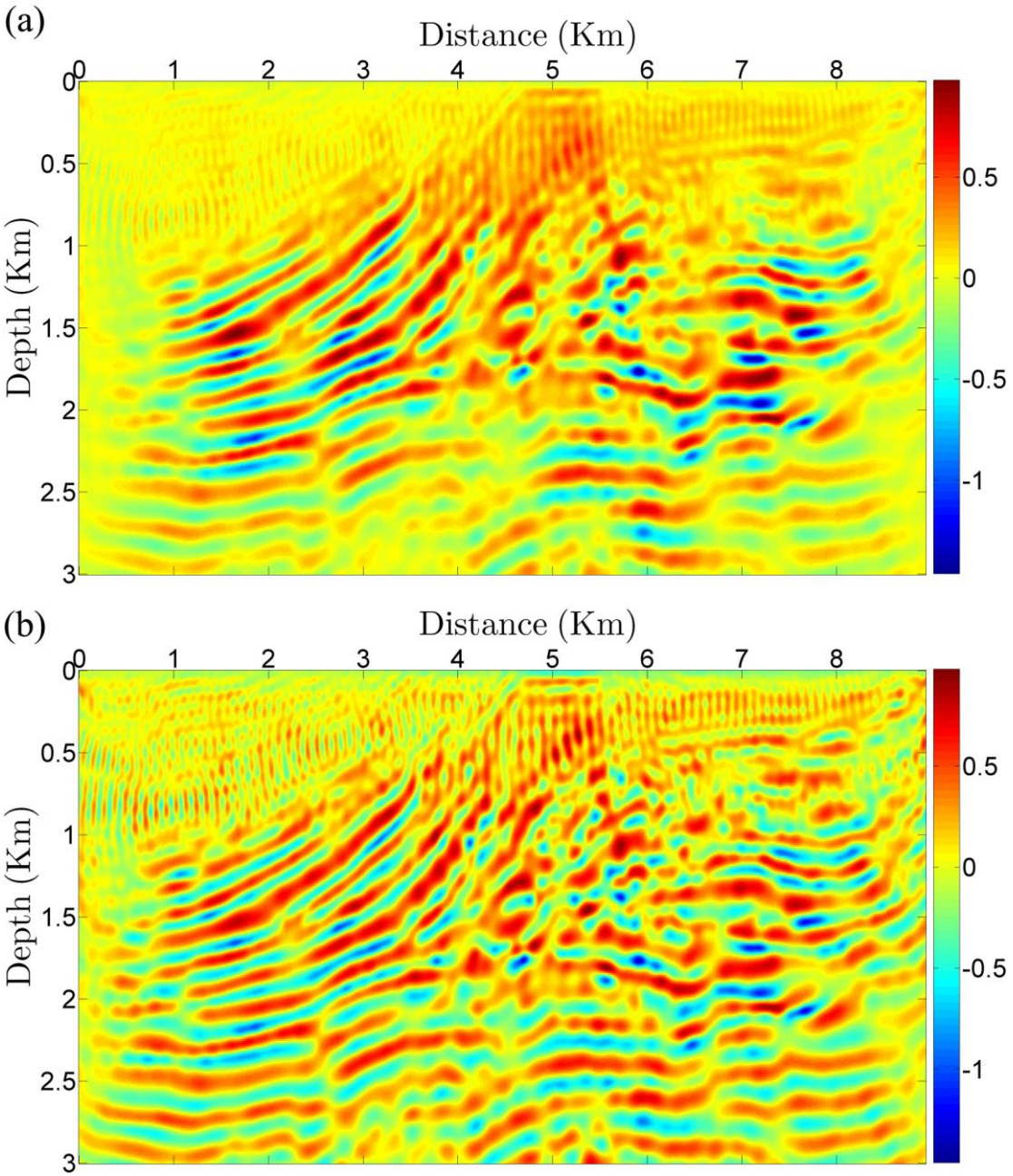


Figure 3. Comparison of (a) the FWI gradient image and (b) its filtered version by bilateral filtering.

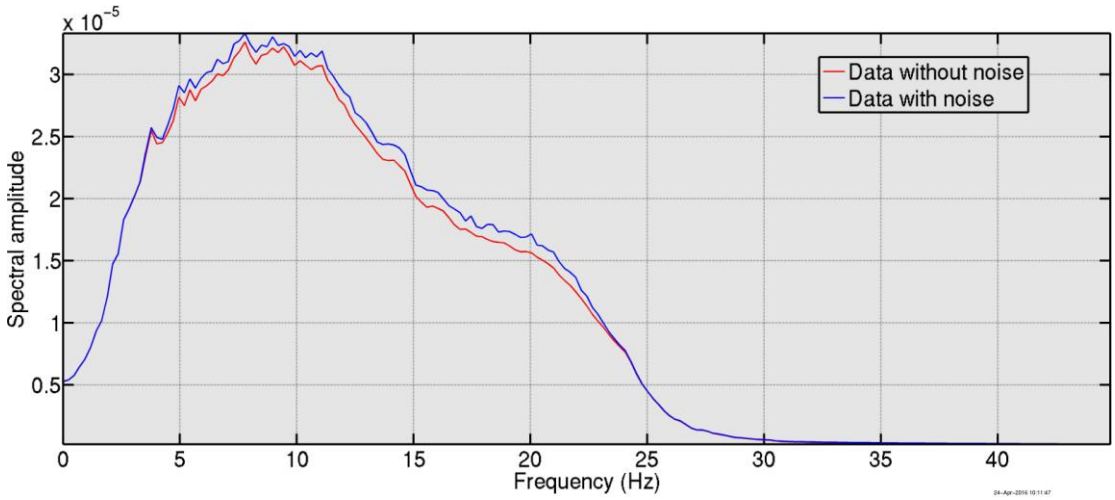


Figure 4 Frequency spectrum of the selected seismic trace from synthetic Marmousi dataset without (red) and with (blue) the random noise.

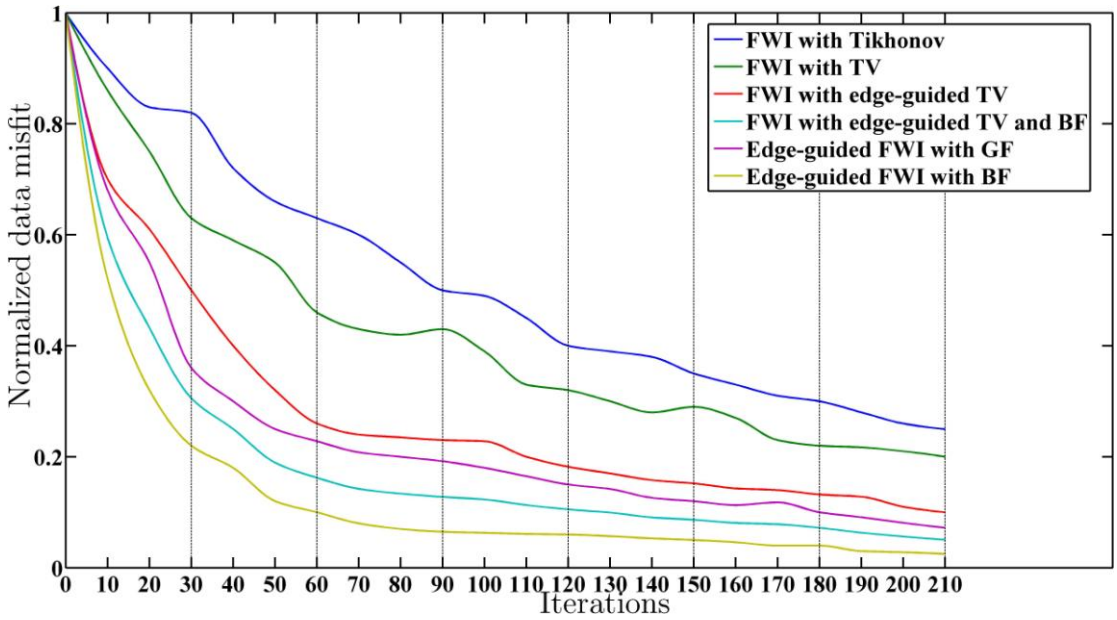
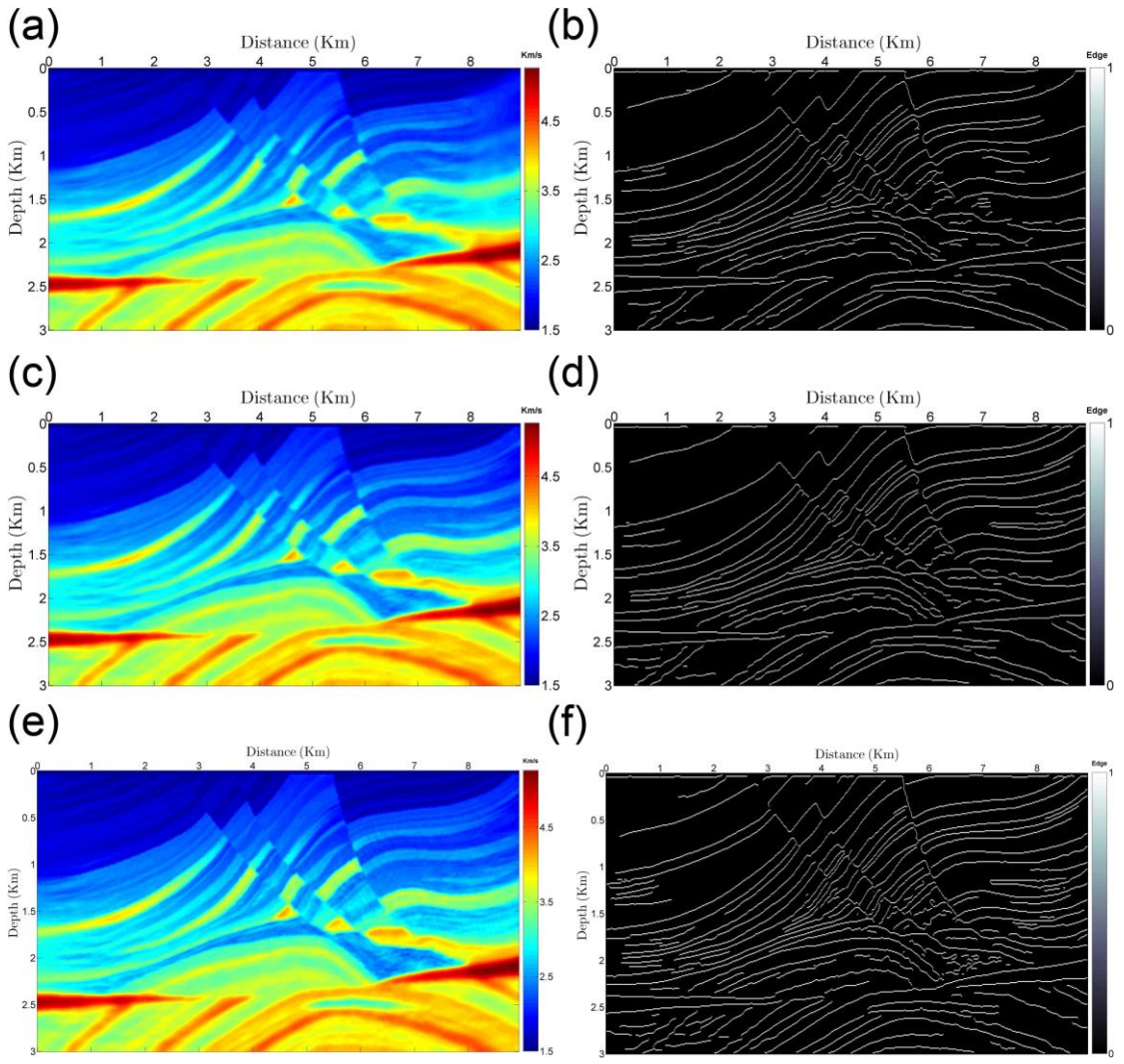


Figure 5.Variation of the normalized data misfits along iterations for (I) FWI with Tikhonov regularization, (II) FWI with TV regularization, (III) FWI with edge-guided TV regularization, (IV) FWI with edge-guided TV regularization and bilateral filtering, (V) edge-guided FWI with edge-guided TV regularization and Gaussian filtering (GF), and (VI) edge-guided FWI with edge-guided TV regularization and bilateral filtering (BF), respectively. Data misfit at each iteration is scaled with respect to the initial data misfit.



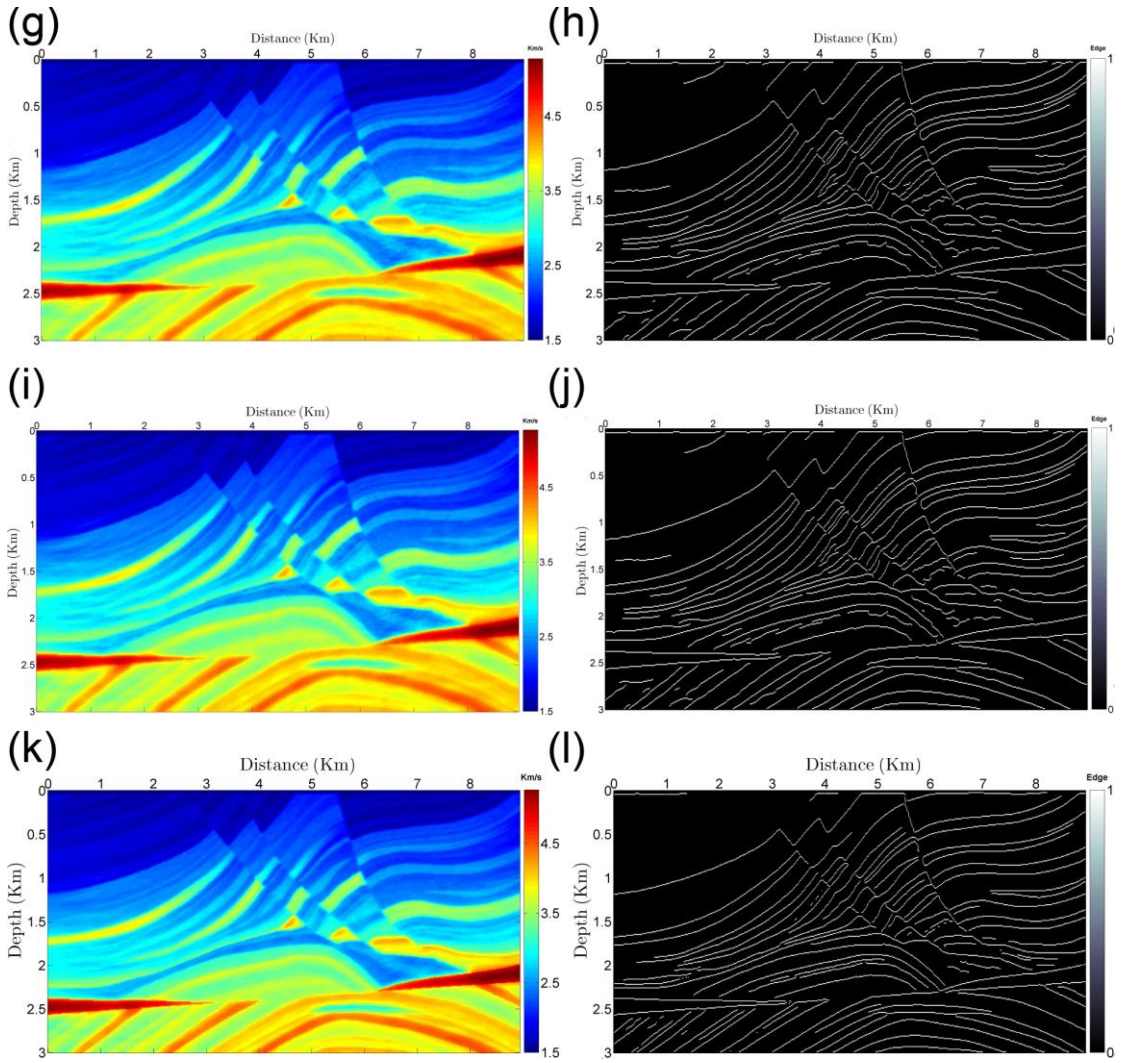


Figure 6. Comparison of final velocity models (left) and the detected model edges (right) from (a-b) FWI with Tikhonov regularization, (c-d) FWI with TV regularization, (e-f) FWI with edge-guided TV regularization, (g-h) FWI with edge-guided TV regularization and bilateral filtering; (i-j) edge-guided FWI with edge-guided TV regularization and Gaussian filtering, and (k-l) edge-guided FWI with edge-guided YV regularization and bilateral filtering.

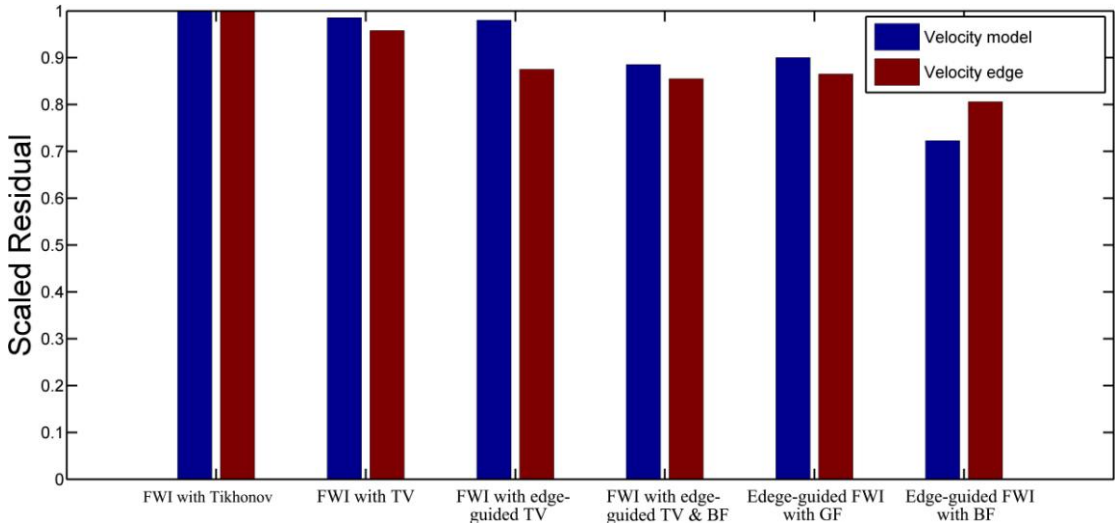


Figure 7. Comparison of scaled model (blue) and edge (red) residuals between true and inverted models for 6 different FWI strategies. RMS differences for velocity anomalies and edges are calculated between true model and recovered model from different FWI strategies. Then RMS values for all the FWI strategies are scaled by those corresponding to the FWI with Tikhonov regularization.

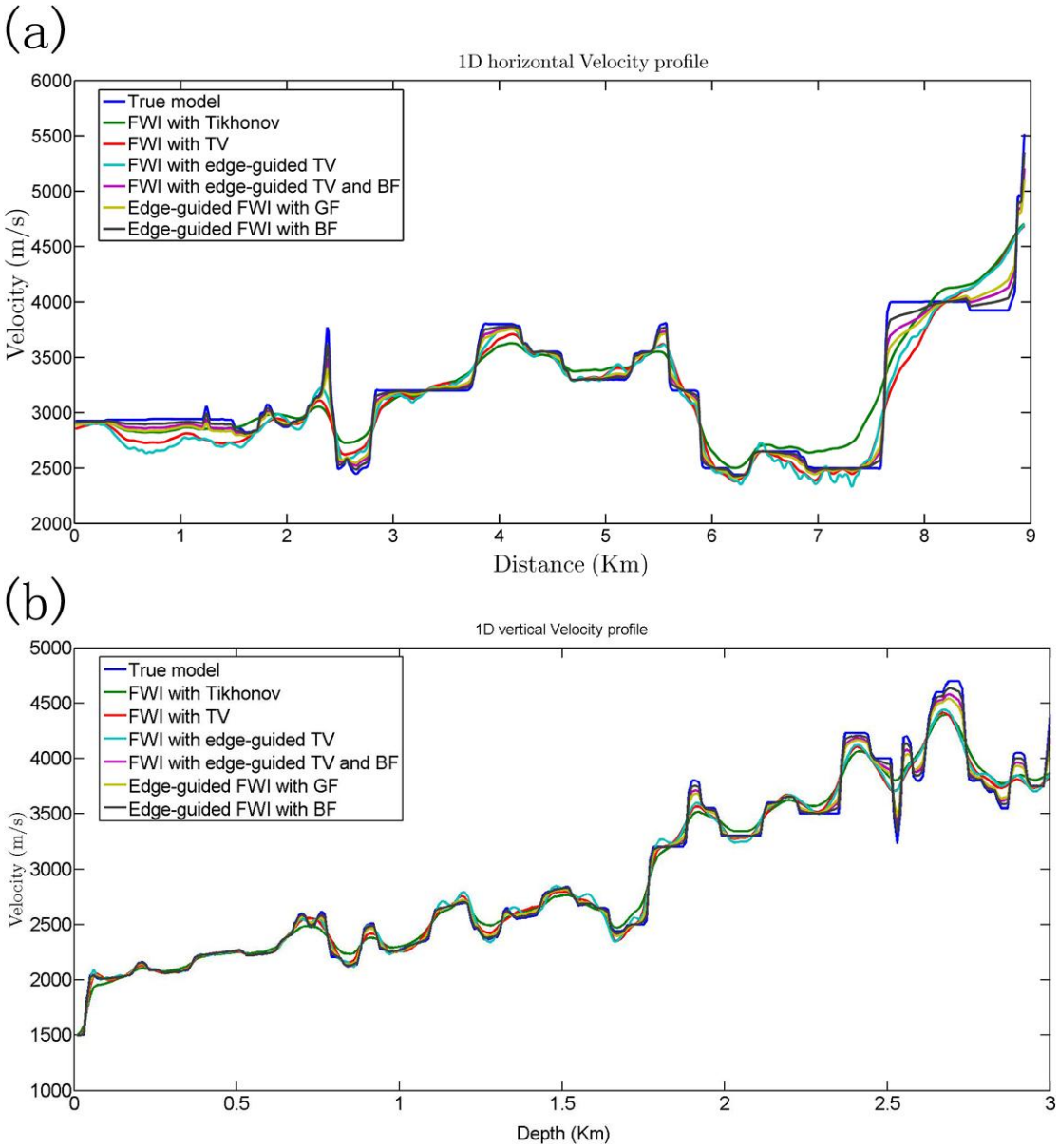


Figure 8. Comparison of velocity models along a vertical profile (a) and a horizontal profile (b) between true and inverted models from different FWI strategies. Vertical and horizontal profiles are shown as white lines in Figure 1a.

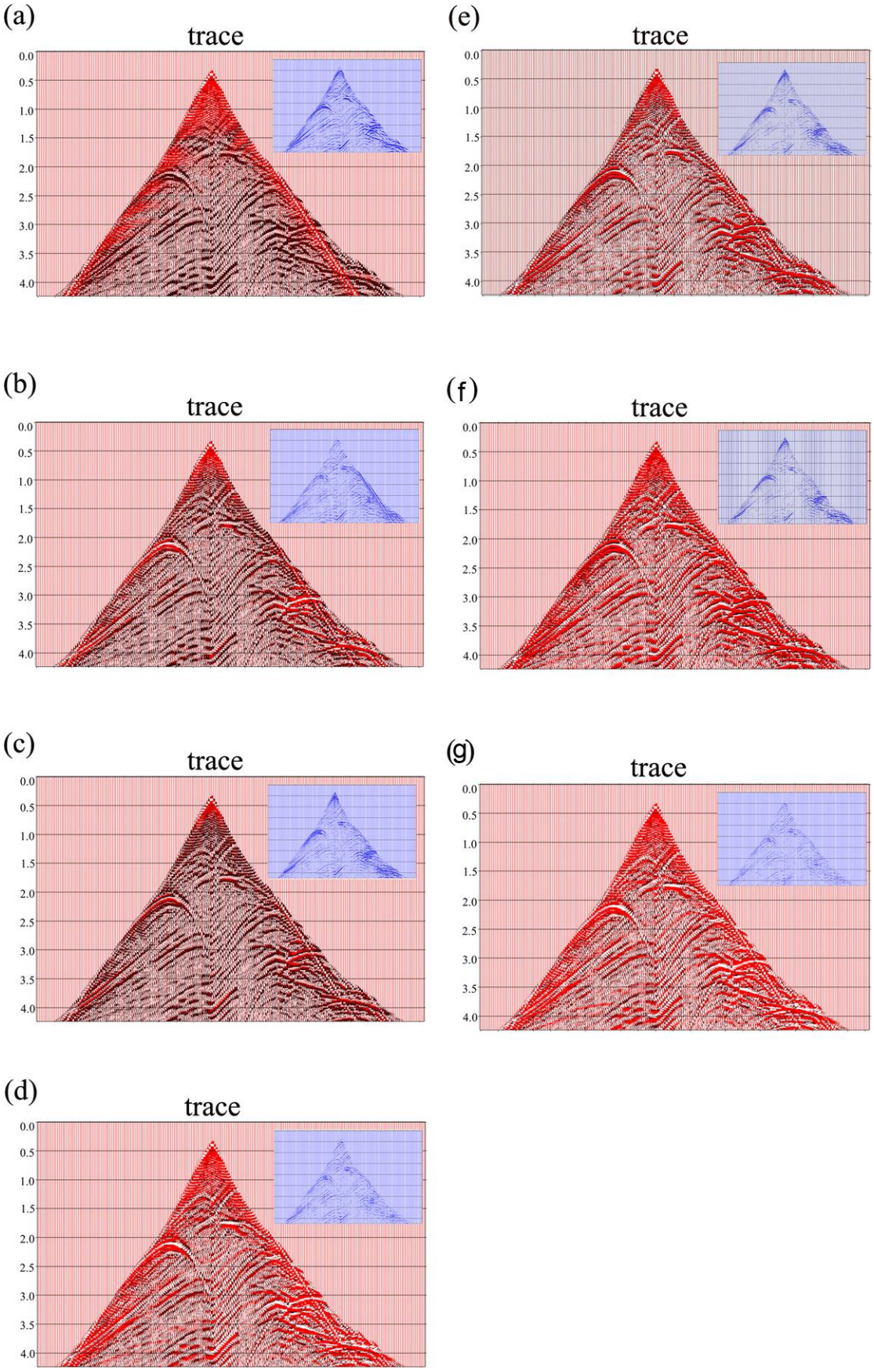


Figure 9. Comparison of one shot gather simulated using true model (black) and inverted model (red) from different FWI strategies. The waveform differences are shown in the upper right corner. (a) Initial model; (b) FWI with Tikhonov regularization; (c) FWI with TV regularization; (d) FWI with edge-guided TV regularization; (e) FWI with edge-guided TV regularization and bilateral filtering; (f) Edge-guided FWI with edge-guided TV regularization and Gaussian filtering; (g) Edge-guided FWI with edge-guided TV regularization and bilateral filtering. Note that the simulated waveforms from inverted and true models and their differences are plotted in different amplitude scales to better show different levels of waveform fitting for different FWI strategies. The amplitudes of waveform differences approximately vary from 1% to 10% of those corresponding to absolute waveforms.

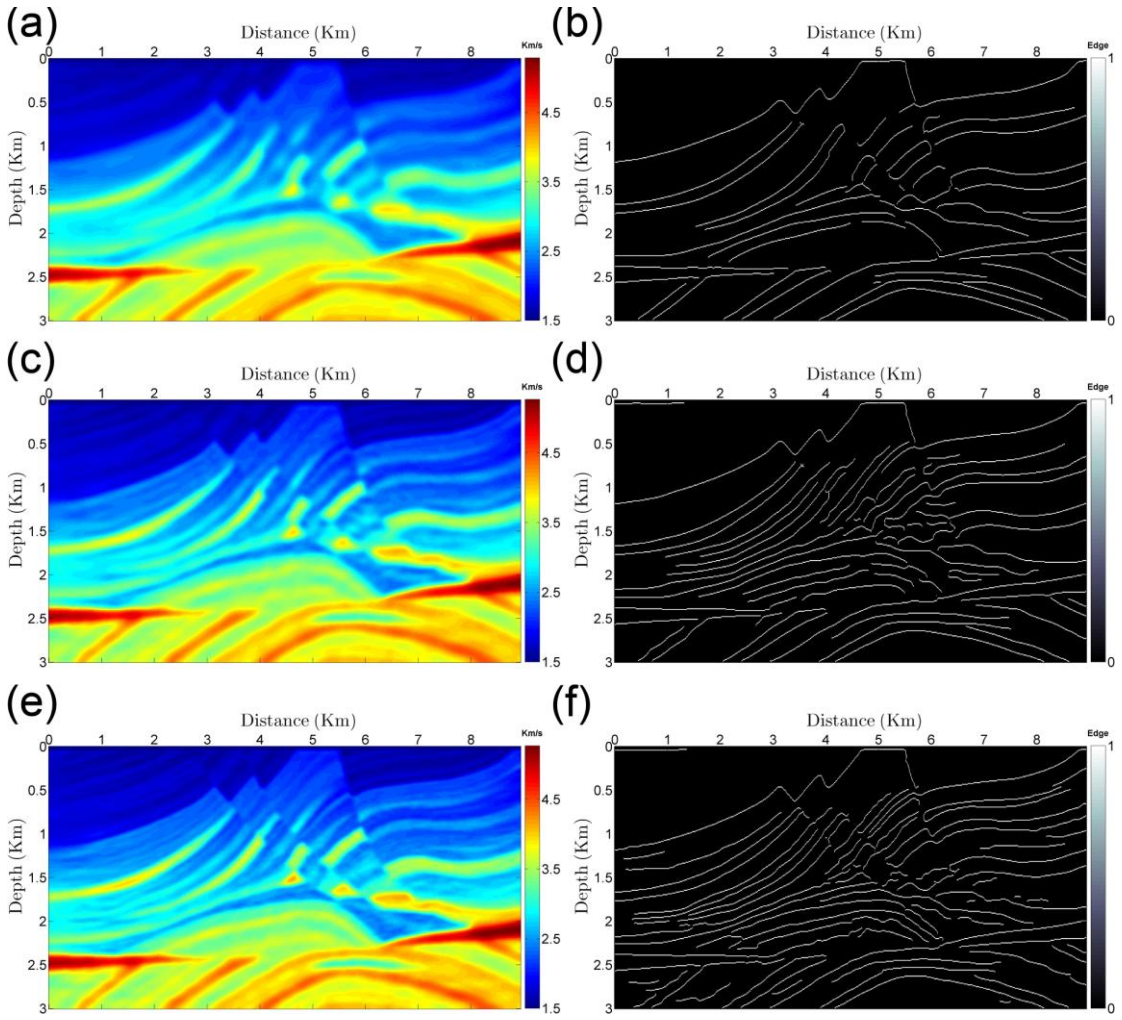


Figure 10. Inverted velocity models (left) from edge-guided FWI with edge-guided TV regularization and bilateral filtering by successively including higher frequency waveforms. (a) Up to 3.5 Hz; (c) Up to 7.6 Hz; (e) Up to 15 Hz and their corresponding model edges (right) extracted by Canny edge detection algorithm (b, d and f).