Image warping waveform tomography

Francesco Perrone and Paul Sava
Center for Wave Phenomena, Colorado School of Mines

ABSTRACT

Imaging the change in physical parameters in the subsurface requires an estimate of the long wavelength component of the same parameters in order to reconstruct the kinematics of the waves propagating in the subsurface. The model is unknown and must be estimated from the same data. One can try to reconstruct the method by matching the recorded data with modeled waveforms extrapolated in a trial model of the medium. Alternatively, assuming a trial model, one can obtain a set of images of the reflectors from a number of seismic experiments and match the locations of the imaged interfaces. Apparent displacements between migrated images contain information about the velocity model and can be used for velocity analysis. A number of methods are available to estimate the displacement between images; in this paper we compare penalized local correlations and image warping. We show that the image warping vector field is a more reliable tool for estimating displacements between migrated images and leads to a more robust velocity analysis procedure. Using image warping, we define an optimization problem for migration velocity analysis. We derive the expressions for the objective function using the adjoint-method and we also propose an alternative, simpler solution for the computation of the model update based on the estimated perturbation of the imaged reflectivity and demigration. The second approach has straightforward implementation and reduced computational cost compared to the adjoint-state method calculations. We also discuss the weakness of migration velocity analysis in the migrated-shot domain in the case of highly refractive media, when the Born modeling operator is far from being unitary and thus its adjoint operator (the migration operator) poorly approximates the inverse.

1 INTRODUCTION

Accurate propagation velocities in the subsurface are crucial in seismic imaging and exploration geophysics. A kinematically accurate velocity model allows us to correctly reconstruct Green functions and thus precisely image the subsurface discontinuities responsible for the data acquired at the surface or in boreholes. An accurate image of the geologic structure is key for planning drilling operations and for reducing hazards associated with drilling through faults. Seismic propagation velocities are sensitive to the stress conditions in the subsurface and carry information about the orientation of the principal stresses and overpressures (Carcione, 2007), which are both important for mitigating hazards during drilling operations and reducing risk in production.

In recent years, a main distinction has been drawn between methods that try to recover the velocity model by matching the recorded data with simulated synthetic wavefields modeled in a trial velocity model (Tarantola, 1984; Pratt, 1999; Sirgue and Pratt, 2004), and methods that focus on optimizing properties of the migrated images (Fowler, 1985; Faye and Jeannot, 1986; Al-Yahya, 1989; Chavent and Jacewitz, 1995; Biondi and Sava, 1999; Sava et al., 2005; Albertin et al., 2006). Full-waveform inversion is the name used to indicate techniques that operate in the data domain and focus on the minimization of some functional of the data residual. Migration velocity analysis, or image-domain tomography, identifies all methods that reconstruct the velocity model by analyzing the similarity between migrated images.

Full-waveform inversion is simple to implement, but it is very sensitive to every inconsistency between the recorded and modeled data. It requires an accurate parametrization of the physics of wave propagation, knowledge of the source signature, and a kinematically reliable starting model in order to converge to the global minimum of the objective function (Santosa and Symes, 1989; Kelly et al., 2010). Migration velocity analysis is not as sensitive to errors in the source signature and it is more robust against errors in the initial velocity model; it is characterized by an objective function which is convex over a wide range of model perturbations and thus intrinsically more suitable for gradient-based optimization procedures. Nonetheless, as it ignores the amplitude information in the data for inversion purposes, migration velocity analysis cannot achieve the high-resolution results that full-waveform inversion promises in the ideal scenario. Migration velocity analysis is conventionally used to produce a kinematically accurate model to initialize full-waveform inversion for refining the velocity with high-resolution details.

Image-domain tomography has advantages and disadvantages compared to the data-domain full-waveform inversion. In general, the image domain is more robust against errors in the velocity model. Full-waveform inversion needs a kinematically accurate initial guess of the model parameters in order to converge to geologically plausible models. Image-domain tomography recovers the long wavelength component of the model, which is responsible for the wave kinematics. Since it aims to match the observed data with simulated waveforms, full-waveform inversion is rather sensitive to the source signature; on the contrary, the image domain is insensitive to the source wavelet. On the other hand, the data-domain allows us to achieve higher resolution in the inversion result. Moreover, in extremely refractive media, single-shot migrated images can be contaminated by kinematic artifacts due to the fact that migration is just the adjoint, and not an inverse, of a modeling operator (Stolk and Symes, 2004). In this work, we are interested in the reconstruction of the long wavelength component of the model parameters and thus we focus on image-domain tomography.

This paper deals with the estimation of the velocity model using image warping. Perrone et al. (2012) show how the apparent displacement field estimated between two images can be used to measure velocity errors with a minimum number of migrated images. Perrone et al. (2012) also derive an approximation for the image perturbation based on the estimated displacement field and design a linearized wave-equation migration velocity analysis algorithm based on the method developed by Sava and Biondi (2004). The method presented in Perrone et al. (2012) relies on several approximations and linearizations: the wave-equation migration velocity analysis operator is obtained from the phase-shift continuation operator and several linearizations with respect to the wavefield and wavenumber (Born approximation). The entire inversion procedure stems from the particular form of the extrapolation operator that allows one to derive a linearized operator which links perturbation in the image space to perturbation in the model (slowness) parameters. Because of the strong assumption of a phase-shift continuation wave-extrapolator, this method cannot be applied to images obtained using different engines (for example, with reverse-time migration). Moreover, events that are not correctly handled by the extrapolation operator (for example, diving waves and overturning reflections) become noise for the inversion scheme.

In order to generalize the wave-equation migration velocity analysis method for nonlinear inversion, Perrone and Sava (2012) propose an image-domain wavefield tomography approach where the residual is computed from estimates of local shifts between migrated images. The shifts are obtained from penalized local correlations of images from nearby experiments. The gradient of the objective function is computed using the adjoint-state method and thus a local gradient-based nonlinear inversion scheme is implemented. The method presented in Perrone and Sava (2012) involves the solution of the full two-way wave equation and thus integrates reverse-time migration into the inversion procedure. The method shows weak points when reflectors are located near the acquisition surface (where the distance between the shot positions becomes an important illumination issue) and when several reflectors fall inside the same local window, thus biasing the computation of the local correlations through the crosstalk between the closely spaced events. In contrast, the displacement vector field appears to be a more stable and reliable measure of image variation: the displacement vectors can be computed using the algorithm developed by Hale (2007b) as an iterative procedure searching for the maximum of local correlations, thus leading to a smooth estimate of the displacement vectors.

In this paper, we extend the linearized wave-equation migration velocity analysis algorithm of Perrone et al. (2012) to incorporate a two-way extrapolation engine and to directly exploit the displacement field. The displacement vector field is used to estimate an image perturbation, then through single scattering Born modeling, the wavefield perturbation due to the image perturbation is computed. The correlation between the scattered and background wavefields represents the gradient of the objective function, which is the kernel for the model update. We compare the objective function based on the image perturbation with the objective function based on local image correlations (Perrone and Sava, 2012) and show that the displacement field leads to a robust objective function and model update. For comparison, we also derive the equations for the computation of the gradient of the objective function using the adjoint-state method and show that the greater complexity and computational cost does not produce a higher quality gradient.

2 THEORY

Perrone et al. (2012) introduce a measure of velocity error for migration velocity analysis exploiting the consistency between the orientation of the structural features in the image and the apparent displacement between two migrated images. The error measure $\mathcal{J}$ is the energy of the projection of the displacement vector $\mathbf{u}(\mathbf{x})$ onto the dip vector $\boldsymbol{\nu}(\mathbf{x})$

$$\mathcal{J}(m) = \frac{1}{2} \left\| \boldsymbol{\nu} \cdot \mathbf{u} \right\|^2, \tag{1}$$

where $m(\mathbf{x})$ represents the model parameter (slowness, slowness squared, or velocity) as a function of position $\mathbf{x}$. The dip vector $\boldsymbol{\nu}(\mathbf{x})$ is defined as the vector normal to the structure at position $\mathbf{x}$, and the displacement vector $\mathbf{u}(\mathbf{x})$ is also a function of the position vector $\mathbf{x}$. The dip vector $\boldsymbol{\nu}$ can be computed using the algorithm presented in Hale (2007a) and based on gradient square tensors (van Vliet and Verbeek, 1995). Displacement vectors $\mathbf{u}$ are obtained by iterative search of the maximum of local correlations and warping of the input images; the method, first applied to measuring the apparent shift between migrated images in 4D seismic processing, is described in detail in Hale (2007b).

Both the dip and the displacement vectors depend nonlinearly on the model parameters. The dip vector depends nonlinearly on the migrated image, since we need to compute the eigenvector associated with the largest eigenvalue of the gradient square tensor in order to obtain an estimate of the dip vector (van Vliet and Verbeek, 1995). The displacement vector

field is obtained by picking the correlation lag that maximizes the local correlation panels:

$$\mathbf{u}(\mathbf{x}) = \arg\max_{\boldsymbol{\lambda}} c(\mathbf{x}, \boldsymbol{\lambda}), \tag{2}$$

where $c(\mathbf{x}, \boldsymbol{\lambda}) = \int_{w(\mathbf{x})} R_i\left(\boldsymbol{\xi} - \frac{\boldsymbol{\lambda}}{2}\right) R_{i+1}\left(\boldsymbol{\xi} + \frac{\boldsymbol{\lambda}}{2}\right) d\boldsymbol{\xi}$ is the local correlation of two images $R_i = R_i(\mathbf{x})$ and $R_{i+1} = R_{i+1}(\mathbf{x})$, and $w(\mathbf{x})$ is are seamless overlapping windows. The $\arg\max$ operator is not differentiable and thus we cannot compute the Frechét derivative with respect to the model parameters directly.

Perrone et al. (2012) show that, for optimization purposes, it is possible to substitute the dip vector field in equation 1 with the gradient of the migrated image, the latter being at every point parallel to the dip field and oriented toward the direction of maximum increase in the image amplitude. The objective function then becomes

$$\mathcal{J}(m) = \frac{1}{2} \|\nabla R \cdot \mathbf{u}\|^2. \tag{3}$$

Substituting the dip field with the image gradient has the advantage of introducing in equation 3 quantities (the gradient of the migrated images) that we can easily differentiate with respect to the model parameters using standard perturbation theory techniques and thus enable a gradient-based optimization algorithm.

In this paper, we discuss two possible approaches for minimizing equation 3. The first solution is to use the adjoint-state method (Lions, 1972; Fichtner et al., 2006) and can be considered the brute-force approach. The second solution is based on demigration of the residual $\nabla R \cdot \mathbf{u}$, which is interpreted as the mismatch between the updates of the small wavelengths of the model parameters obtained from adjacent shots. This second approach is related to the work of Xu et al. (2012), who follow a two-step, migration/demigration procedure to better condition FWI for reflection data. Another work on a similar path is the differential waveform inversion presented by Chauris and Plessix (2012) again in the framework of FWI. Here, we set up the problem in the image domain, i.e. in the framework of migration velocity analysis. The demigration approach reduces the computational cost and complexity with respect to the adjoint-state calculations and greatly simplifies the implementation and the physical interpretation of the method. Of course, since we are approximating the image difference with our image warping procedure and we are not constructing the exact gradient for the objective function in equation 3, we are introducing a systematic error in the procedure. The exact quantification of this error requires the comparison with the exact gradient obtained via the adjoint-state method.

2.1 Direct adjoint-state calculations for vector displacements

Let us consider equation 3; we can write the gradient of the objective function as follows:

$$\begin{aligned} \nabla_m \mathcal{J} &= \frac{\partial}{\partial m} \frac{1}{2} (\nabla R \cdot \mathbf{u})^2 \\ &= (\nabla R \cdot \mathbf{u}) \left(\frac{\partial \nabla R}{\partial m} \cdot \mathbf{u} + \nabla R \cdot \frac{\partial \mathbf{u}}{\partial m} \right), \end{aligned} \tag{4}$$

in this paper $m(\mathbf{x})$ represents slowness squared. The gradient in equation 4 consists of two terms. The term $\frac{\partial \nabla R}{\partial m} \cdot \mathbf{u}$ involves a linear function of the image; the term $\nabla R \cdot \frac{\partial \mathbf{u}}{\partial m}$ involves a nonlinear function of the image, because the apparent displacement is estimated as the maximum argument of local image correlations (equation 2). Although lengthy, the adjoint-state calculations for the first part of the gradient are straightforward because of the linearity between ∇R and R; more complicated is the second term because the nonlinear and non differentiable $\arg\max$ function does not allow an immediate application of the adjoint state method. Perrone and Sava (2012) avoid the complications of computing the differential of $\arg\max$ by defining a measure of shifts through penalized local image correlations. The correlation operator is (bi)linear in the input images and allows one to derive directly the adjoint-state equation to compute the gradient with respect to the model parameters.

Using standard perturbation theory we can write

$$\begin{aligned} \mathcal{J} + \delta\mathcal{J} &= \frac{1}{2} \int (\nabla (R + \delta R) \cdot (\mathbf{u} + \delta\mathbf{u}))^2 \, d\mathbf{x} \\ &= \frac{1}{2} \int ((\nabla R + \nabla \delta R) \cdot (\mathbf{u} + \delta\mathbf{u}))^2 \, d\mathbf{x}. \end{aligned} \tag{5}$$

Expanding the products and identifying the terms that are linear in the perturbations in equation 5, we obtain

$$\begin{aligned} \delta\mathcal{J} &= \delta\mathcal{J}^{\nabla\delta R} + \delta\mathcal{J}^{\delta\mathbf{u}} \\ &= \int (\nabla R \cdot \mathbf{u}) (\nabla \delta R \cdot \mathbf{u}) \, d\mathbf{x} \\ &+ \int (\nabla R \cdot \mathbf{u}) (\nabla R \cdot \delta\mathbf{u}) \, d\mathbf{x}. \end{aligned} \tag{6}$$

The objective function perturbation in equation 6 consists of two terms: $\mathcal{J}^{\nabla\delta R}$ involves the perturbation of the image gradient, $\delta\mathcal{J}^{\delta\mathbf{u}}$ depends on the perturbation of the displacement. In the context of velocity analysis, the perturbations δR and $\delta\mathbf{u}$ depend on the the change in model parameters m. To derive an expression for the gradient of the objective function with respect to the model parameters, we need to relate the objective function perturbation to wavefield perturbations. Then, using standard adjoint-techniques, we can obtain the gradient of the objective function without actually computing the Frechét derivatives of the wavefield with respect to the model parameters (Fichtner et al., 2006). In the following, we analyze both terms $\delta\mathcal{J}^{\nabla\delta R}$ and $\delta\mathcal{J}^{\delta\mathbf{u}}$ and express them directly in terms of wavefield perturbations.

Let us start with $\delta\mathcal{J}^{\nabla\delta R}$. Since the ∇ operator is linear and the image R is a linear functional of the source and re-

ceiver wavefields, we can perturb the image with respect to the wavefields first and then apply the ∇ operator. The migrated image R is the zero-lag correlation in time of the source and receiver wavefields $u_s(\mathbf{x},t)$ and $u_r(\mathbf{x},t)$ (Claerbout, 1985):

$$R(\mathbf{x}) = \int u_s(\mathbf{x},t)\, u_r(\mathbf{x},t)\, dt. \tag{7}$$

By perturbing the source and receiver wavefields in equation 7 and collecting the terms that depend linearly on the perturbation, we obtain

$$\delta R = \int (u_s \delta u_r + \delta u_s u_r)\, dt, \tag{8}$$

where we drop the dependence on $\mathbf{x}$ and t to simplify the notation. The wavefield perturbations δu_s and δu_r in equation 8 can be easily computed using the demigration approach presented in (Xu et al., 2012) or using standard adjoint-state techniques. From the image perturbation δR is straightforward to compute the gradient by applying the operator ∇.

To link the second term in equation 6 to the perturbation of the source and receiver wavefields, we use the method of connective functions developed by Luo and Schuster (1991). Let us formally write the gradient of $\mathcal{J}^{\delta\mathbf{u}}$ with respect to the model parameters:

$$\nabla_m \mathcal{J}^{\delta\mathbf{u}} = (\nabla R \cdot \mathbf{u}) \left(\nabla R \cdot \frac{\partial \mathbf{u}}{\partial m} \right). \tag{9}$$

Notice that since we are analyzing only the changes of the apparent displacement field due to perturbations of the velocity model, we do not include in equation 9 the gradient of ∇R with respect to the model m. We want to explicitly link the displacement perturbation $\delta\mathbf{u}$ to the wavefield perturbation, then standard adjoint-state calculations allow us to avoid the computation of the Frechét derivatives of the wavefields and thus compute the gradient of the objective function in equation 3 efficiently.

By definition, the maximum of the local correlation is an extremum and thus the gradient with respect to the correlation lags must vanish for $\boldsymbol{\lambda} = \mathbf{u}$

$$\nabla_{\boldsymbol{\lambda}}\, c|_{\boldsymbol{\lambda}=\mathbf{u}} = \int_{w(\mathbf{x})} R_i(\boldsymbol{\xi})\, \nabla R_{i+1}(\boldsymbol{\xi}+\mathbf{u})\, d\boldsymbol{\xi} = 0. \tag{10}$$

Equation 10 implicitly connects the displacement vectors $\mathbf{u}$ with the model parameter m via the migrated images and the wave-equation. Thus, we define the connective function

$$F(\mathbf{x},\mathbf{u}) = \int_{w(\mathbf{x})} R_i(\boldsymbol{\xi})\, \nabla R_{i+1}(\boldsymbol{\xi}+\mathbf{u})\, d\boldsymbol{\xi} = 0 \tag{11}$$

Using equation 11 and following a derivation analogous to Luo and Schuster (1991), we rewrite the gradient of the displacement vectors with respect to the model parameters in equation 4 using the connective function $F(\mathbf{x},\mathbf{u})$

$$\nabla_m \mathcal{J}^{\delta\mathbf{u}} = (\nabla R \cdot \mathbf{u})\, \nabla R \cdot \left(\frac{\partial F}{\partial \mathbf{u}} \right)^{-1} \frac{\partial F}{\partial m}. \tag{12}$$

The derivative of the connective function with respect to the model parameters can be computed using adjoint-state techniques; the mathematical form of $F(\mathbf{x},\mathbf{u})$ involves only the migrated image and linear functions of the images, and thus, in principle, the derivative $\frac{\partial F}{\partial m}$ can be calculated using the adjoint-state method. The derivative $\frac{\partial F}{\partial \mathbf{u}}$ is the multidimensional generalization of the denominator in equation (7a) in Luo and Schuster (1991):

$$\frac{\partial F}{\partial \mathbf{u}} = \int R_i(\boldsymbol{\xi})\, \nabla\nabla R_{i+1}(\boldsymbol{\xi}+\mathbf{u})\, d\boldsymbol{\xi}, \tag{13}$$

where, in $2D$,

$$\nabla\nabla R_{i+1}(\boldsymbol{\xi}+\mathbf{u}) = \begin{bmatrix} \frac{\partial^2 R_{i+1}(\boldsymbol{\xi}+\mathbf{u})}{\partial z^2} & \frac{\partial^2 R_{i+1}(\boldsymbol{\xi}+\mathbf{u})}{\partial z \partial x} \\ \frac{\partial^2 R_{i+1}(\boldsymbol{\xi}+\mathbf{u})}{\partial x \partial z} & \frac{\partial^2 R_{i+1}(\boldsymbol{\xi}+\mathbf{u})}{\partial z^2} \end{bmatrix} \tag{14}$$

is a 2×2 tensor of second derivatives. Notice that since $\left(\frac{\partial F}{\partial \mathbf{u}}\right)^{-1}$ is a 2×2 matrix-valued function defined at every image point $\mathbf{x}$, and $\frac{\partial F}{\partial m}$ is a 2×1 vector-valued function at every point $\mathbf{x}$, their product is a 2×1 vector-valued function, and then the dot product with the spatial image gradient ∇R in equation 9 makes mathematical sense. The generalization to 3D is straightforward. This expression is similar to equations (6) and (7a) in (Luo and Schuster, 1991); the vectorial nature of the displacement field makes equation 12 tensorial.

The analysis of the differential of objective function in equation 3 with respect to the model parameters shows the mathematical complexity of implementing this solution for velocity analysis. Especially the multidimensional generalization of the procedure developed in Luo and Schuster (1991) poses a number of implementation challenges: in the time domain, we have one-dimensional shifts and equation (7a) in Luo and Schuster (1991) involves *scalar* functions; in the image domain, the shifts are vectors and we have to deal with *vector* and *matrix*-valued functions, which considerably increases the complexity of the gradient computation. Moreover, the data correlation used in Luo and Schuster (1991) is global over the entire seismic trace, intuitively that accounts for an average shift between the observed and simulated data. In the image domain, the global correlation of the migrated images has little physical sense (if any at all). The local correlation implementation increases the dimensionality of the problem and then the cost, and picking of specific location in the image is necessary for an efficient implementation.

In the next section, we show how to greatly simplify the computation of the update of the velocity model using the concept of image perturbation and the link between the image difference and the dot product $\nabla R \cdot \mathbf{u}$, as suggested in Perrone et al. (2012).

2.2 Adjoint-state method, demigration, and linearized inversion

The dot product between the gradient of the migrated image and the apparent displacement field between two images (equation 3) is a linear approximation of the image difference with respect to the displacement vector (Perrone et al., 2012). Compared to the image difference, equation 3 is less affected by amplitude differences due to the change of the shot location

and is less prone to cycle skipping problems because warping maps every event in one image to the corresponding event in the second image. The relation between image difference and the dot product of image gradient and displacement field can be used to set up a wavefield tomography procedure based on the image difference objective function, but more robust.

The adjoint-state calculations for the image difference objective function are well-known in the geophysics literature (Plessix, 2006). Although introduced as a regularization term for full-waveform inversion, the image-difference objective function represents a sensible functional for migration velocity analysis. Intuitively, all images from separate seismic experiments must be maximally similar to each-other when the velocity model is correct, and thus the energy of the first derivative along the shot dimension must be minimum. The major weak points of this approach are related to sensitivity to the acquisition geometry and cycle skipping in the image domain, as discussed in Perrone et al. (2012). Also, the kinematic artifacts due to the fact that the migration operator is the adjoint and not the inverse of the Born modeling operator limit any technique in the shot-image domain (Stolk and Symes, 2004).

The adjoint sources $g_{s_i}(\mathbf{x},t)$ and $g_{r_i}(\mathbf{x},t)$ for the image-difference objective function are straightforward to compute. They depend on the second derivative of the migrated images along the shot axis of the prestack image cube and the wavefields computed in the background model (Plessix, 2006):

$$g_{s_i} = (R_{i-1} - 2R_i + R_{i+1})\, u_{ri} \quad (15)$$

$$g_{r_i} = (R_{i-1} - 2R_i + R_{i+1})\, u_{si}. \quad (16)$$

The functions in equation 15 are the force terms for the wavefield extrapolations

$$\mathcal{L}(m)\, a_{si} = g_{s_i} \quad (17)$$

$$\mathcal{L}(m)\, a_{ri} = g_{r_i}, \quad (18)$$

where $a_{si}(\mathbf{x},t)$ and $a_{ri}(\mathbf{x},t)$ are the adjoint wavefields. The adjoint sources generate the scattered fields due to the reflectivity perturbation represented by $(R_{i-1} - 2R_i + R_{i+1})$. These wavefields can also be interpreted as the *demigrated* reflectivity perturbation $(R_{i-1} - 2R_i + R_{i+1})$. The gradient of the objective function is then the zero-lag time correlation of the background and adjoint wavefields (Fichtner et al., 2006; Plessix, 2006).

By inspecting equation 15 and 16, we can immediately see the practical issues of implementing waveform tomography by minimizing the energy of the image difference. Amplitudes in the prestack images depend on the relative position between the shots, thus even when the velocity model is correct, uneven illumination of the subsurface may lead to spurious adjoint sources and gradients of the objective function. Even more important than illumination, cycle skipping may arise when the velocity model is severely inaccurate, because different images map the same structures at inconsistent positions in the subsurface.

We propose to replace the term $(R_{i-1} - 2R_i + R_{i+1})$ in the expression of the adjoint sources with an approximation derived from image warping. The apparent displacement field absorbs the amplitude differences due to the change in the shot position and, at the same time, removes the cycle skipping problem by warping one image into its neighbor. Let us indicate with $\mathbf{u}_{i-1}$ and $\mathbf{u}_i$ the apparent displacement field that warps R_{i-1} into R_i and R_i into R_{i+1}, respectively. We can construct approximations of the image differences by computing the dot product of the image gradient with the displacement fields. In order to guarantee symmetry with respect to the ith shot, we use R_i to construct the image perturbations

$$\Delta R_{i-1} = R_i - R_{i-1} \approx \nabla R_i \mathbf{u}_{i-1} \quad (19)$$

$$\Delta R_i = R_{i+1} - R_i \approx \nabla R_i \mathbf{u}_i. \quad (20)$$

Using the expressions in equations 19-20, we can approximate the second derivative along the shot axis as follows:

$$R_{i-1} - 2R_i + R_{i+1} \approx \Delta R_i - \Delta R_{i-1}, \quad (21)$$

which is robust against cycle skipping (only one image, R_i, enters the expression for the image perturbation), and it is less sensitive to illumination patterns because the amplitude effects are absorbed in the computation of the displacement fields.

3 NUMERICAL EXAMPLES

We test image warping wavefield tomography for different depths of the reflecting interface and compare it with the direct image difference tomography approach (Plessix, 2006) and local image correlation wavefield tomography (Perrone and Sava, 2012). We show that image warping wavefield tomography is more robust with respect to the depth of the reflector (i.e. the ratio between the source receiver offset and the depth of the reflector) and lead to more stable gradients for model update. We also show how the method fails to produce sensible gradients in highly refractive media, despite the fact that the objective function is able to identify the correct model. In this later case, the displacement estimation is stable but the artifacts in the migrated image produce noisy image gradients and thus lead to spurious scattered wavefields and artifacts in the gradient of the objective function. Finally, we test our method on the Marmousi model.

3.1 Horizontal reflector

Let us consider a flat horizontal reflector in a homogeneous velocity model. In the first case, the interface is at 2 km depth (Figure 1) and it is due to a density contrast. In order to measure errors in the velocity model, we analyze two nearby shots: the shot position of the first shot is $x = 3.8$ km and the distance between the shots is 200 m. We apply different homogeneous perturbations to the velocity model, we migrate the data and construct the image perturbations for the different models from the migrated images Figure 2. Figures 2(a), 2(c), and 2(e) show the migrated imaged for high, correct, and low velocities. The image perturbations for the different models computed as the inner product between the image gradient and the

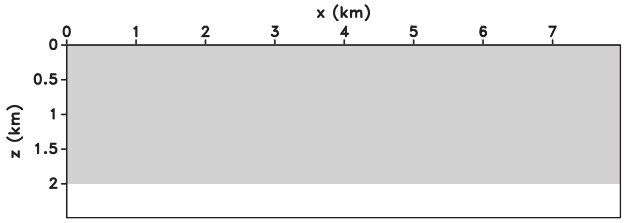


Figure 1. Horizontal reflector model. The interface is due to a density contrast in a constant velocity medium. The depth of the interface is large compared to the distance between the shot points.

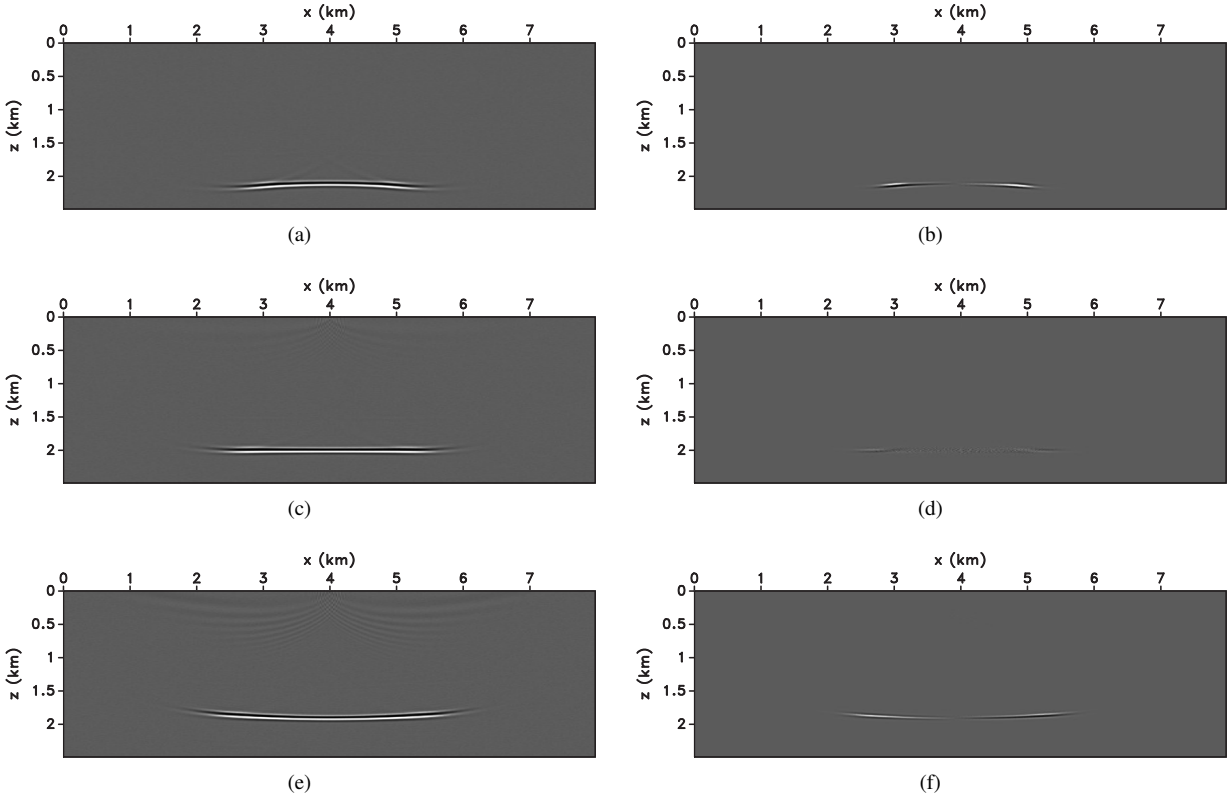


Figure 2. Migrated images using (a) high, (c) correct, and (e) low velocities. The perturbation with respect to the correct model is constant and homogeneous over the entire model. Image perturbations for (b) high , (d) correct, and (f) low velocities. The perturbation with respect to the correct model is constant and homogeneous over the entire model.

displacement vector field obtained by warping one image into the other are shown in Figures 2(b), 2(d), and2(f). Observe that the image perturbation is coherent when the velocity model is incorrect (Figures 2(b) and 2(f)), whereas it is practically weak random noise when the velocity model is correct (Figure 2(d)), except at the edges of the subsurface aperture. We obtain the sensitivity kernels and the objective function in Figures 3(a), 3(b), and 3(c) by correlating the background wavefields (used for migration) with the scattered wavefields excited by the interaction of the background wavefields with the reflectivity perturbation computed from equation 21. The gradient is smooth over the entire extent of the reflector when the velocity is incorrect and has negligible value when the model is correct. The objective function in Figure 12 is convex over the entire perturbation range and the global minimum clearly points at the correct model.

By comparison, Figures 4(a), 4(b), and 4(c) shows the gradients obtained for the same perturbations using the image-difference objective function. Notice the strong side lobes at the edges of the reflector, where we observe cycle skipping because the overlap between the images is not enough. The gradients are noisier than the ones obtained using image-warping, and also the objective function in Figure 4(d) shows a weak bias when the model is correct. Nonetheless, the gradients obtained using image warping in this example are comparable with the results reported in Perrone and Sava (2012) for the

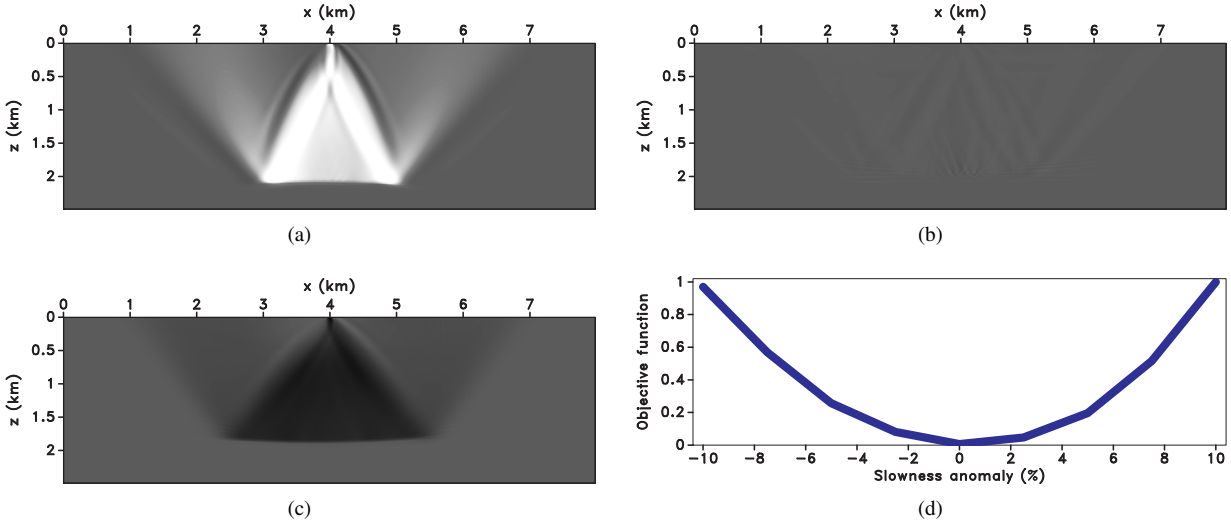


Figure 3. Sensitivity kernels obtained using *image warping* wavefield tomography for (a) high, (b) correct, and (c) low velocities. (d) Objective function for *image warping* wavefield tomography evaluated for a set of model perturbations and the *deep interface* in Figure 9(b). The minimum of the objective function indicates the correct model.

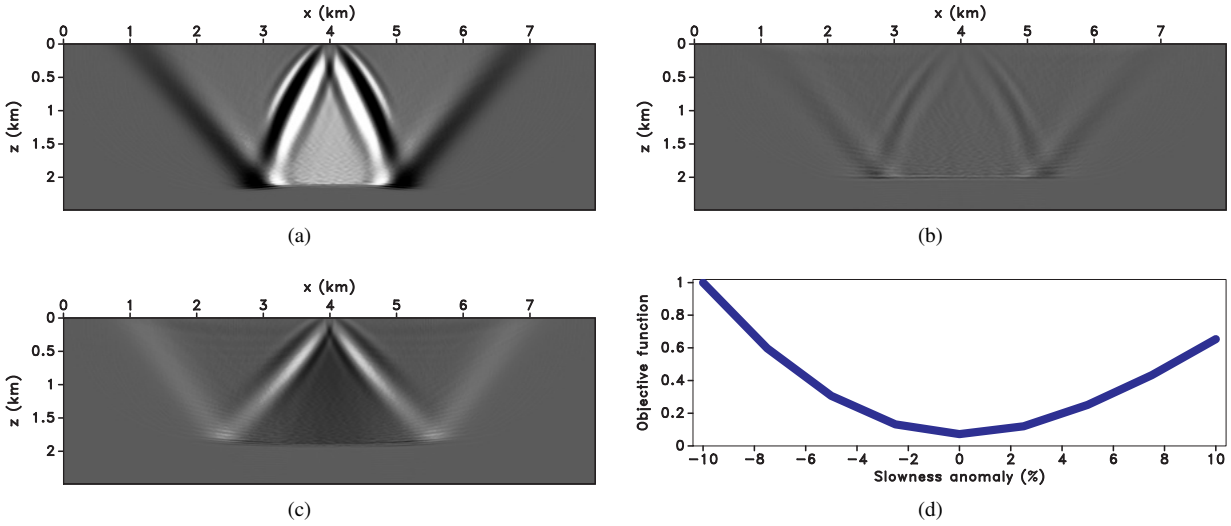


Figure 4. Sensitivity kernels obtained using *image difference* wavefield tomography for (a) high, (b) correct, and (c) low velocities. (d) Objective function for *image difference* wavefield tomography evaluated for a set of model perturbations and the *deep interface* in Figure 9(b).

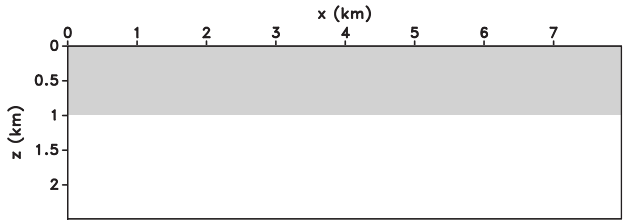


Figure 5. Horizontal reflector model. The interface is due to a density contrast in a constant velocity medium. The reflector is at half the depth of the model in Figure 1.

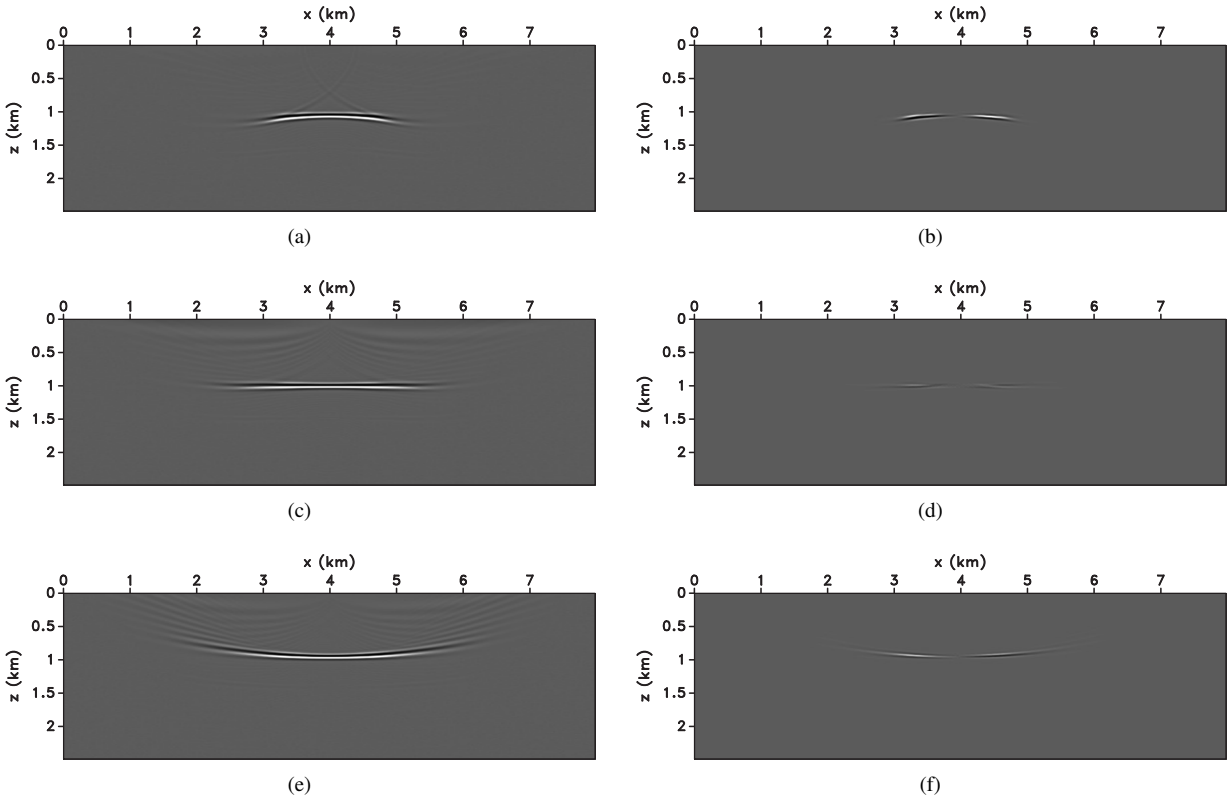


Figure 6. Migrated images of the shallow reflector using (a) high, (c) correct, and (e) low velocities. Image perturbations obtained using image warping for the shallow reflector and (b) high, (d) correct, and (f) low velocities. The perturbation with respect to the correct model is constant and homogeneous over the entire model.

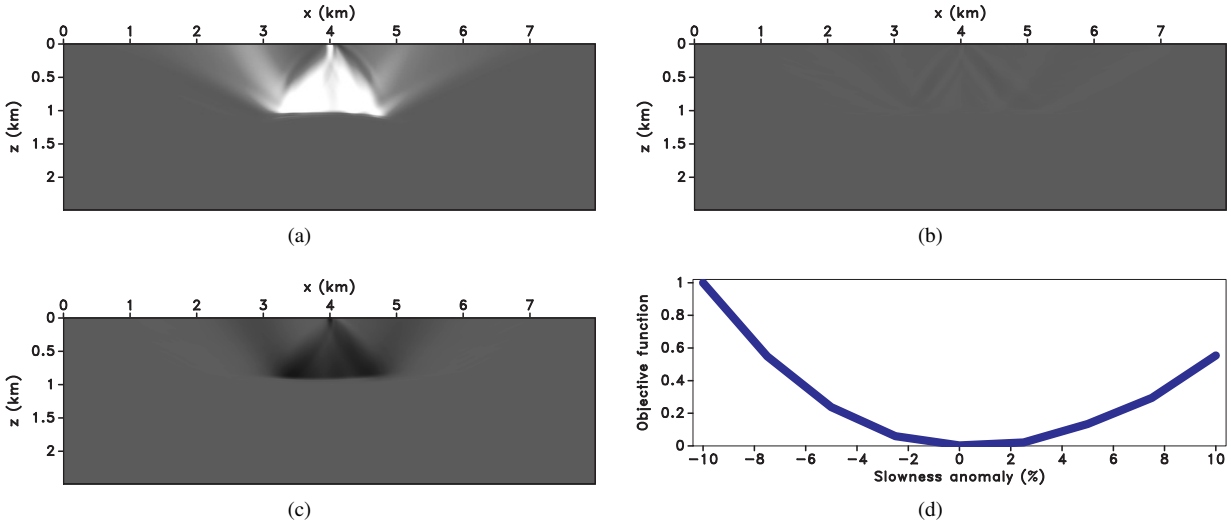


Figure 7. Sensitivity kernels obtained using *image warping* wavefield tomography for (a) high, (b) correct, and (c) low velocities. (d) Objective function for *image warping* wavefield tomography evaluated for a set of model perturbations and a *shallow interface*. As in the case of a deep reflector, the minimum of the objective function indicates the correct model.

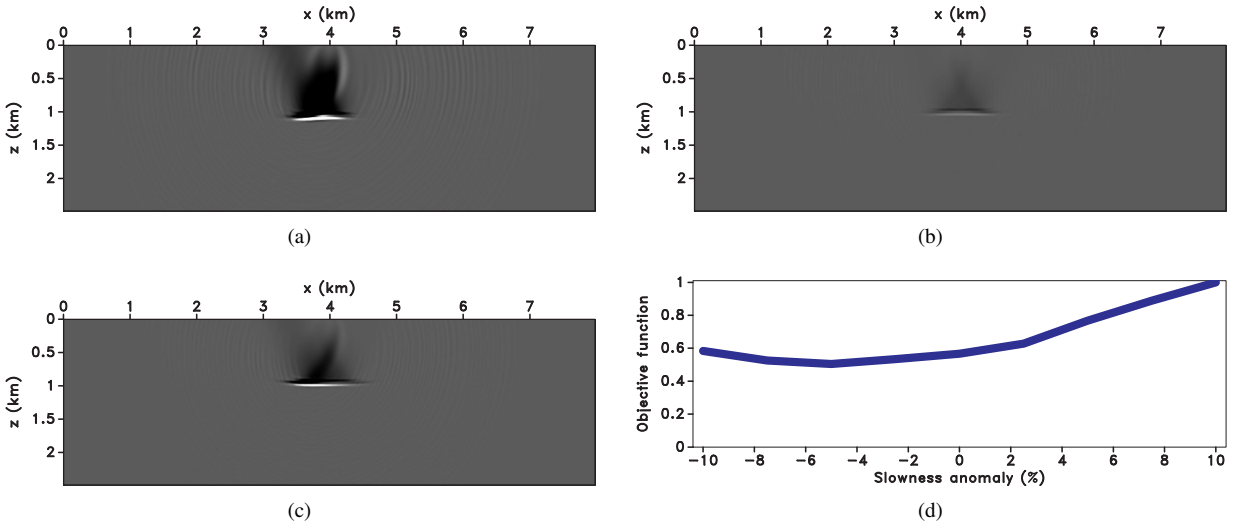


Figure 8. Sensitivity kernels obtained using *local image correlations* wavefield tomography for (a) high, (b) correct, and (c) low velocities. Notice that the kernels fail to recover the sign of the perturbation. (d) Objective function for *local image correlation* wavefield tomography evaluated for a set of model perturbations and a *shallow interface* in Figure 5. The objective function does not indicate correctly the correct model.

objective function employing local image correlations. Thus, the image warping and the penalized local image correlation are expected to perform similarly when the ratio between the depth of the reflector and the shot-point separation is large.

We also consider a shallower interface and compare the single scattering approach with the penalized local correlations. We consider the model in Figure 5 where the density interface is at 1 km depth. The model perturbations applied to the velocity field are the same used as in the previous case. The position of the first shot is again at $x = 3.8$ km and the shot separation is 200 m. The migrated images obtained using different perturbed velocity models are shown in Figures 6(a), 6(c), and 6(e). The lateral extent of the reflector that can now be imaged is smaller than in the previous case because of the shorter distance between acquisition surface and reflector. Figures 6(b), 6(d), and 6(f) shows the dot product of the migrated image with the estimated warping field in this second case. Again, the image perturbations are coherent when the migration model is incorrect (Figures 6(b) and 6(f)) and incoherent for the correct model (Figure 6(d)). Notice that the edges of the subsurface aperture impacts more severely the image perturbation when the reflector is shallow. The sensitivity kernels computed for the shallow reflector (Figures 7(a), 7(b), and 7(c)) show a good behavior, which is comparable with the case of the deep reflector. In Figure 7(d), the objective function is slightly less sensitive to positive perturbations probably because of both the vicinity of the interface to the acquisition surface and the smaller extent of the imaged reflector. Nonetheless, it maintains the convex shape seen in the preceding case and indicates the correct model (global minimum).

We also use the penalized local correlation approach presented in Perrone and Sava (2012) for this second case. The size of the local correlation window is 31 samples in both vertical and horizontal direction, which means the local window capture the overlapping part of the migrated images (the shot distance is 10 samples at the surface and the velocity model is constant). The computed gradients are shown in Figures 8(a), 8(b), and 8(c). Observe that we are not able to recover the correct sign of the perturbation. The separation between the migrated image is such that there is a poor overlap between the images, the values of the computed local correlation coefficients drop, and the estimated shift is strongly biased. The behavior of the objective function (Figure 8(d)) shows that for the case of shallow reflectors the penalized local correlations fail to measure the error in the velocity model.

3.2 Highly refractive model and migration failure

When the model is highly refractive, kinematic artifacts contaminate the single-experiment migrated images, even though the exact model is used to reconstruct the source and receiver wavefields (Stolk and Symes, 2004). This fact is due to the very nature of the migration operator, that is the adjoint of the Born modeling operator and not its inverse.

Figure 9 shows the velocity model and density model we use to study the behavior of image warping wavefield tomography in highly refractive media. The model contains a strong low velocity Gaussian anomaly, whose minimum value is 40% lower than the background 2 km/s velocity. The reflector is simulated by a density contrast. The data are generated using a 15 Hz Ricker wavelet and a staggered-grid finite-difference algorithm with absorbing boundary conditions on every side. The horizontal sampling is 20 m and the vertical sampling is 10 m.

The low velocity area in the model creates triplications of the wavefields. The shot-gather in Figure 10 shows the complexity of the wavefield recorded at the surface. The source is located at $x = 3$ km on the surface and receivers are at ev-

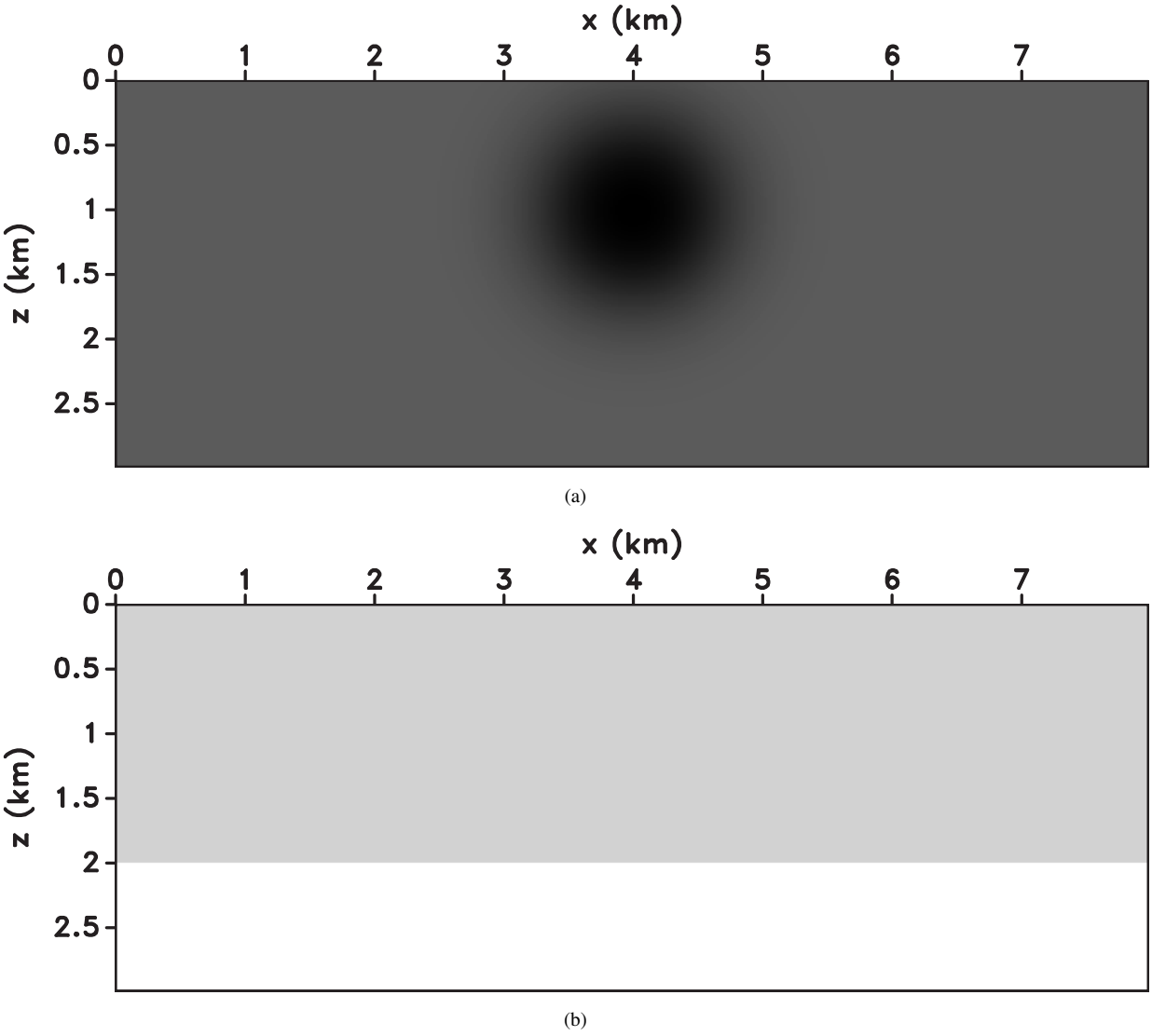


Figure 9. (a) Velocity model used to study the behavior of the inversion in highly refracting media. The low velocity anomaly is Gaussian-shaped and at its minimal value it is 40% smaller than the background 2 km/s background. (b) Density model used to generate reflections.

ery grid point at the surface. Figure 11(a) show the result of migrating the shot-gather in Figure 10 with the *exact* model. Notice that the image is contaminated at about $x = 5$ km with the kinematic artifacts described in Stolk and Symes (2004). These artifacts do not satisfy the relationship between dip field and displacement vector assumed in our inversion procedure. Moreover, since the migration operator is simply the adjoint, and not the inverse, of the Born modeling operator, the amplitudes of the image are corrupted as well. All these distortions affect the calculation of the sensitivity kernel. In Figure 11(b), we show the sensitivity kernel computed for the exact velocity model. Instead of having negligible amplitude, the sensitivity kernel is strong and totally spurious. A Possible solution to obtain kinematically correct images that are suitable for image warping is least-square migration (Nemeth et al., 1999) but would make the inversion procedure more computationally expensive.

Figure 12 shows the objective function obtained by perturbing the correct model with a Gaussian anomaly that progressively cancels out the low velocity zone. Interestingly, although the sensitivity kernel is inaccurate, the objective function is still able to measure the distance from the correct model. This indicates that although the velocity error measure is efficient, the linearizations used for computing the sensitivity kernels and model update can introduce strong errors.

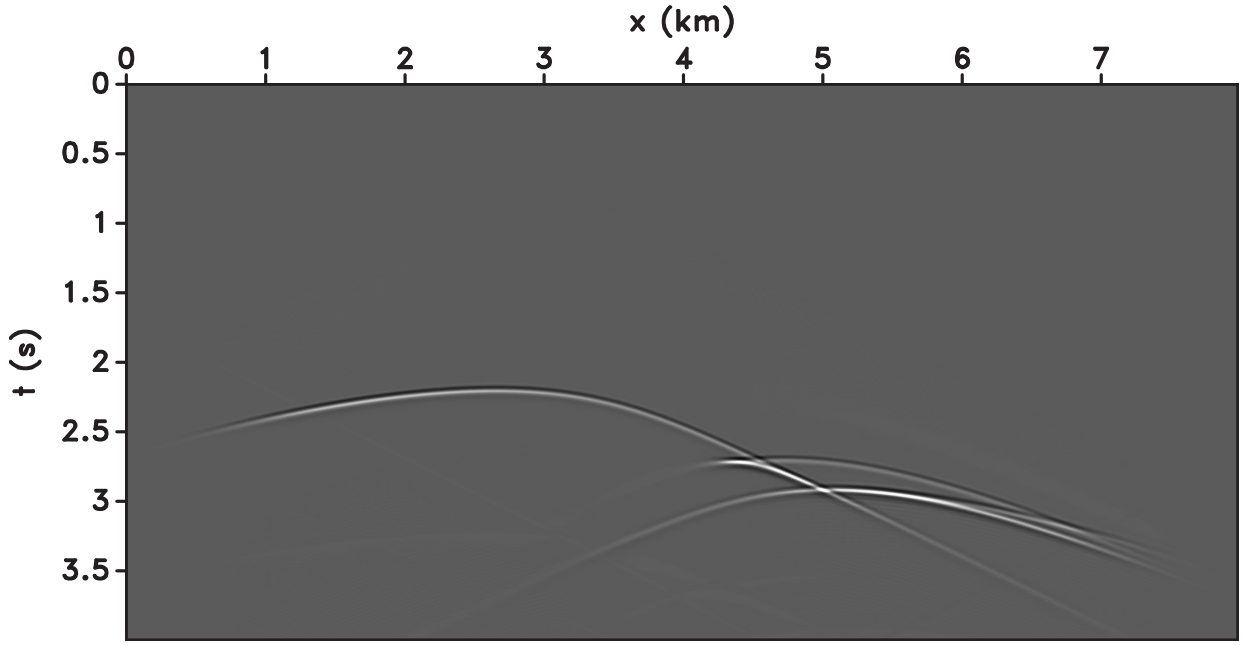


Figure 10. Shot-gather for a source at $x = 3$ km. Notice the complicated response of the medium due to the triplications of the wavefield passing through the low velocity anomaly.

3.3 Marmousi model

We also test our approach on the Marmousi model (Figure 13(a)). We model 100 shots separated by 80 m. The spatial sampling is 8 m in both the x and z direction. Receivers are located at every grid point on the surface. The data are modeled without free-surface multiples but internal multiples are present in the data. We assume that free-surface multiples can be effectively removed from the data using SRME (Verschuur, 1991). The Marmousi model represents a complex land-like scenario and it is characterized by multipathing and strong refractions. This complexity is exemplified by the raypaths calculated in a smoothed version of the correct model, Figure 13(b), for a shot at $x = 4.8$ km.

The initial model is obtained by filtering the exact Marmousi model with a triangular smoother with radius 120 samples vertically and 200 samples horizontally. The smooth model is then biased with a -0.4 km/s homogenous perturbation (Figure 14(a)). The migrated image obtained using this model is shown in Figure 14(b). The central part of the image is non interpretable and all reflectors are mispositioned and defocused; the deep and strong reflectors are shifted about 500 m upward because of the bias toward low velocities in the initial velocity model. The noise in the image is due to the strong errors in the migration model and the inconsistent imaging of the reflectors from different shots. Also, the internal multiples in the data are not imaged correctly since migration is based on single scattering (Born) assumptions.

The inversion reaches a floor after about 20 iterations and returns the velocity model in Figure 15(a) and the associated migrated image in Figure 15(b). In the final image, the reflectors are shifted toward their correct location. We improve the truncation of the sedimentary beds against the fault walls around $z = 1$ km and $x = 3.5$ km; the migration frowns in the original image, due to the discontinuities of the sediments at the fault locations, are corrected by the inversion procedure. Notice that the edges of the image are undercorrected with respect to the better illuminated part of the model. Illumination compensation using the energy of the source, receiver, or both wavefields can attenuate this problem. Because the image perturbation is related to the actual amplitude of the reflectors, the model update is naturally biased toward stronger, more coherent, and better illuminated interfaces. Illumination compensation can also help balancing the strength of the update between the sides of model, which are characterized by simple sedimentary structures, and the central part of the model, where the faults and discontinuities of the reflectors make the wavefields complicated and the wavefront less coherent.

The importance of regularization is described by Figure 15(c), which represent the image obtained after smoothing the model in Figure 15(a) using a triangular filter with radius 100 samples vertically and 400 samples horizontally. The smoothing homogenize the velocity models by spreading the update over a larger subsurface aperture, thus it effectively reduces the effects of uneven illumination. If we compare this image with Figure 13(a), we can observe that the position of the reflectors is quite accurate. The faulted portion of the model is not optimally imaged yet, but the kinematics of the wavefield has been clearly improved and, as a direct consequence, the migrated image improved as well. This result has been obtained in post-processing by smoothing the velocity

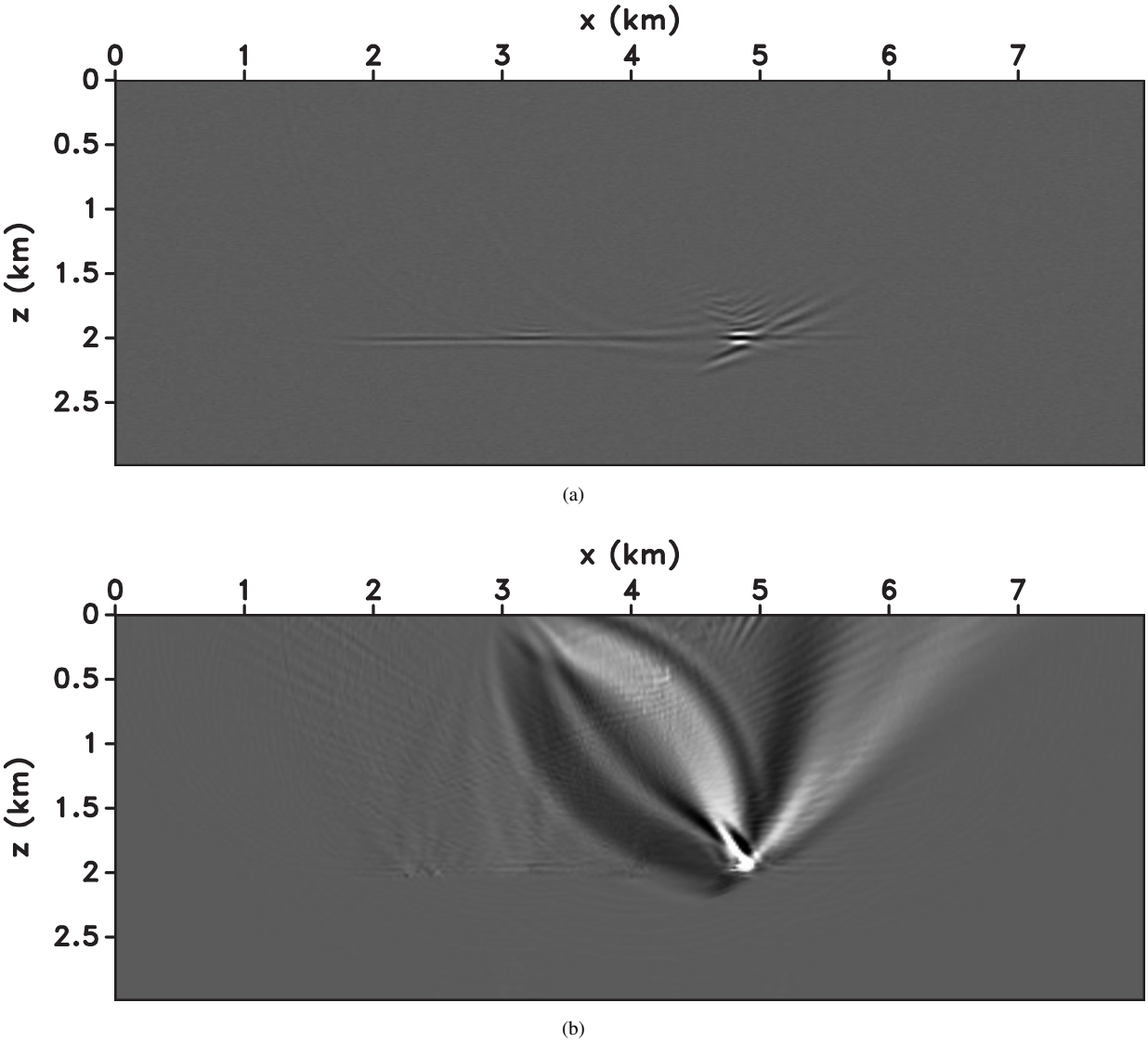


Figure 11. Migrated image obtained from the shot-gather in Figure 10 with the correct velocity model. Observe the distorted image of the reflectors and the artifacts introduced by the migration operator. (b) Sensitivity kernel obtained using image warping wavefield tomography for the correct velocity model. The kernel should be very small and incoherent because the model is actually correct; the spurious kernel is thus completely due to the artifacts in the image.

model obtained by inversion; smoothing corrects the uneven illumination and removes the abrupt changes in the velocity field that distort the image in Figure 15(b). We speculate that a careful illumination compensation and amplitude balancing in the computation of the gradient can achieve a similar result in a direct and automatic way.

4 DISCUSSION

The complete adjoint-state calculations for the objective function in equation 3 is computationally expensive and involved. We require not only the wavefields in the background medium but also their spatial derivatives. The warping field requires a dedicated procedure based on connective functions that link the warping vectors to the velocity model. From the point of the inversion, this step increases the dimensionality of the problem, as at every image point we have to consider the extra dimensions in which the warping vector is defined. Picking image point locations (Yang and Sava, 2010; Perrone and Sava, 2012) becomes crucial to reduce the dimensionality and thus the computational cost. In the method presented in this paper, no picking is necessary in order to construct the image perturbation that we use to construct the estimated scattered wavefields.

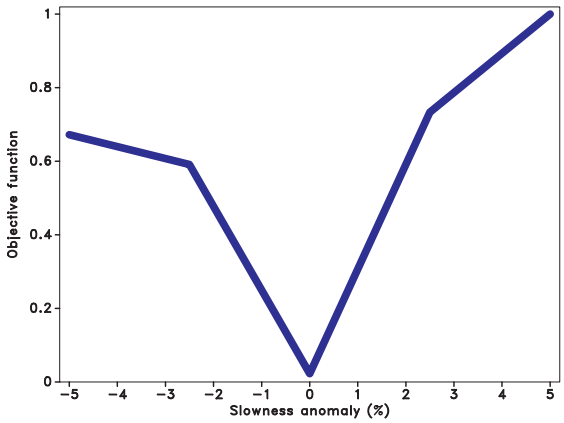


Figure 12. The objective function obtained by perturbing the model with a Gaussian anomaly having the same extent of the one in the correct model but with different amplitudes. Notice that the objective function correctly measures errors in the velocity model.

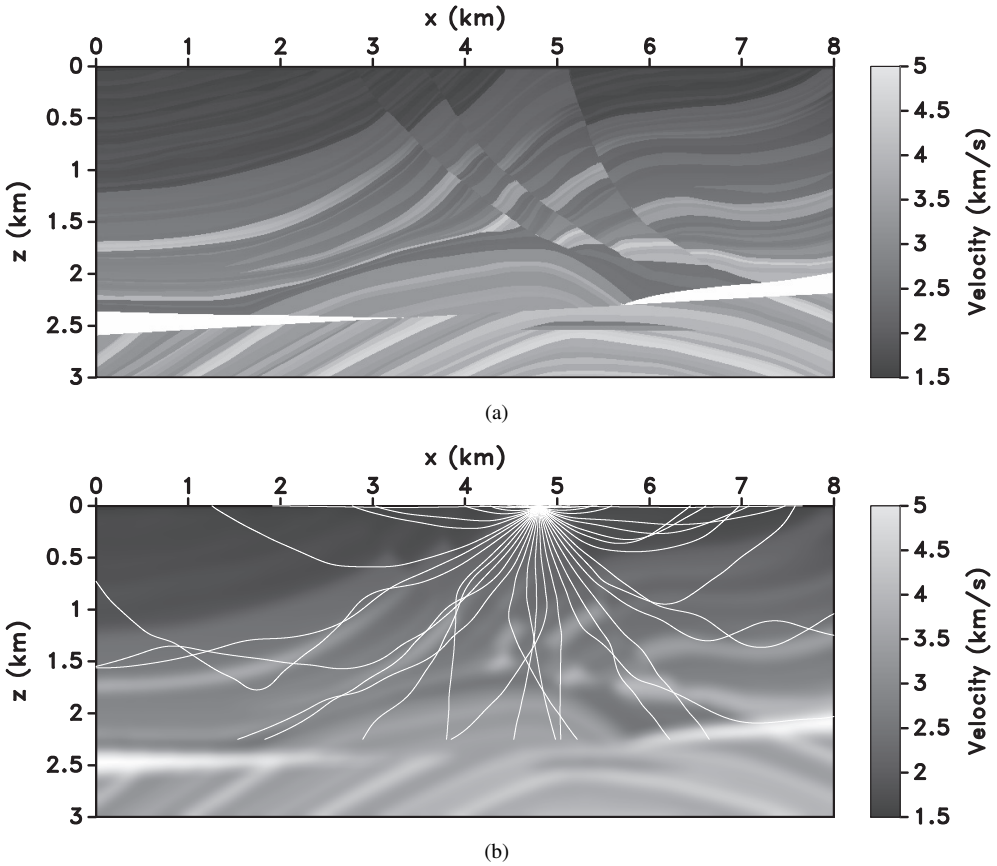


Figure 13. (a) Exact Marmousi velocity model. (b) Ray-tracing in a smooth version of the exact model for a shot at $x = 4.8$ km. The complex raypaths highlight the refractive behavior of the Marmousi model.

The image difference objective function presented in Plessix (2006) as a regularization term for full-waveform inversion is simple to implement and carries straightforward physical intuition about the action of the inversion. In media that are not highly refractive, the drawbacks related to the sensitivity to the amplitude patterns and potential cycle skipping can be effectively and efficiently handled using the apparent displacement field between migrated images and constructing robust approximations of the image difference. Estimating the displacement field and computing the image gra-

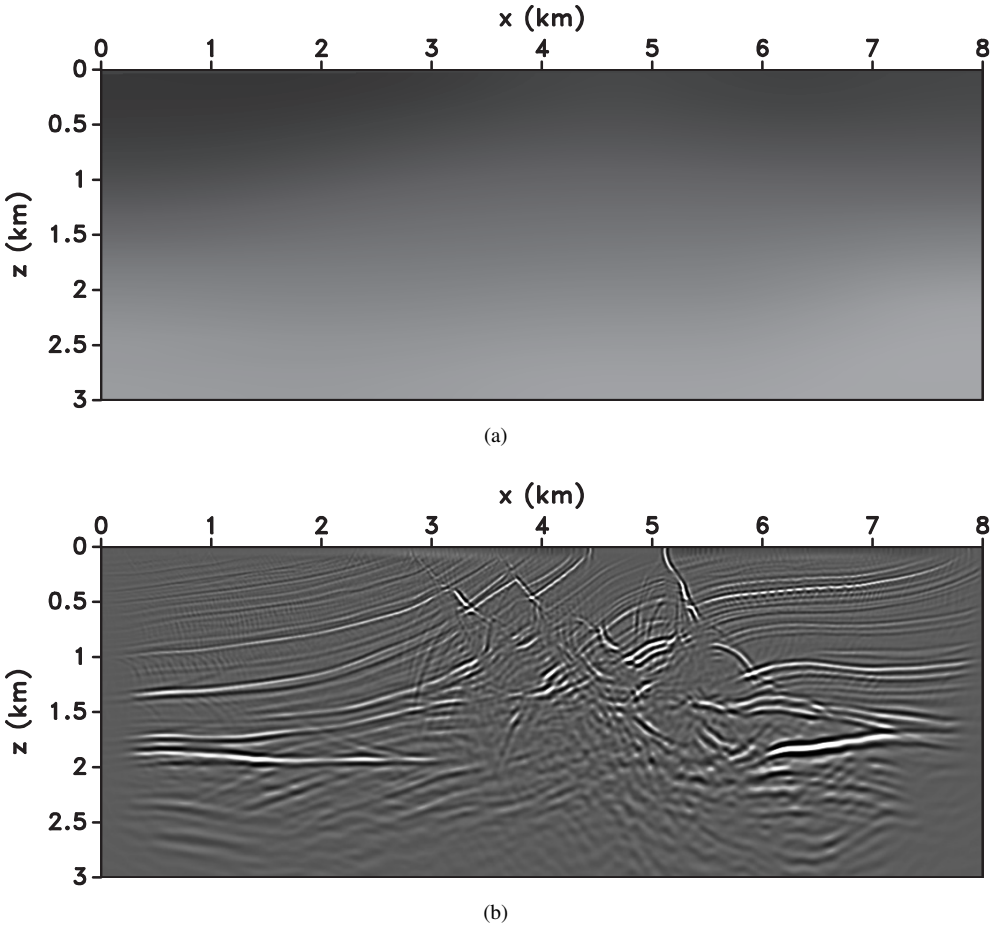


Figure 14. (a) The initial model represented by an incorrect $v(z)$ velocity. (b) The migrated image is distorted and not interpretable in the central part of the model. Observe the general upward shift of the image due to the bias in the velocity model.

dient have a negligible computational cost compared to migration. Moreover, we preserve the physical interpretation of the image perturbation that enters the expression of the adjoint sources, equation 15, which is useful for quality control and for quickly assessing the evolution of the inversion result. The residual evaluation is immediate and the calculation of the adjoint sources is straightforward compared to local correlation waveform tomography, where the computation of the adjoint-sources requires implementation of nonstationary convolutions of the input images with the penalty operators, and local scaling of the background wavefields. When compared to methods that require image-gathers or extended images, our method based on apparent displacement vectors appears easier to implement and less computationally demanding.

Computing the relative shift between two images in local seamless overlapping windows requires some degree of overlapping of the two migrated images. The quantity of interest to qualitatively characterize the distance between the two images in simple subsurface geometries is the ratio between the shot-point separation and the depth of the reflectors. The smaller this ratio, the greater the overlap between migrated images. When we estimate the shift between the migrated images using penalized local correlations, the extent of the local window constrains the domain in which the maximum of the correlation must be. When the separation between the shot is large and the local window is not wide enough to capture the position of the maximum of the local correlation, we introduce an error in the evaluation of the residuals. Notice that increasing the size of the local correlation window implies a decrease in resolution. Moreover, if other events (more reflectors) enter the local window and/or the reflector changes orientation in space, increasing the extent of the local window implies a drop in the values of the estimated correlation coefficients and thus a decrease in the quality of the estimated shifts. The displacement vector field is more robust than the penalized local correlations because of the iterative search and warping procedure used (Hale, 2007b) and thus it allows us to obtain high-quality velocity updates at shallow depths where the ratio between the shot distance and the depth of the reflector is relatively high.

The main drawback of performing inversion in the shot-migrated image domain is related to the nature of the migration operator and to the associated migration artifacts. In highly re-

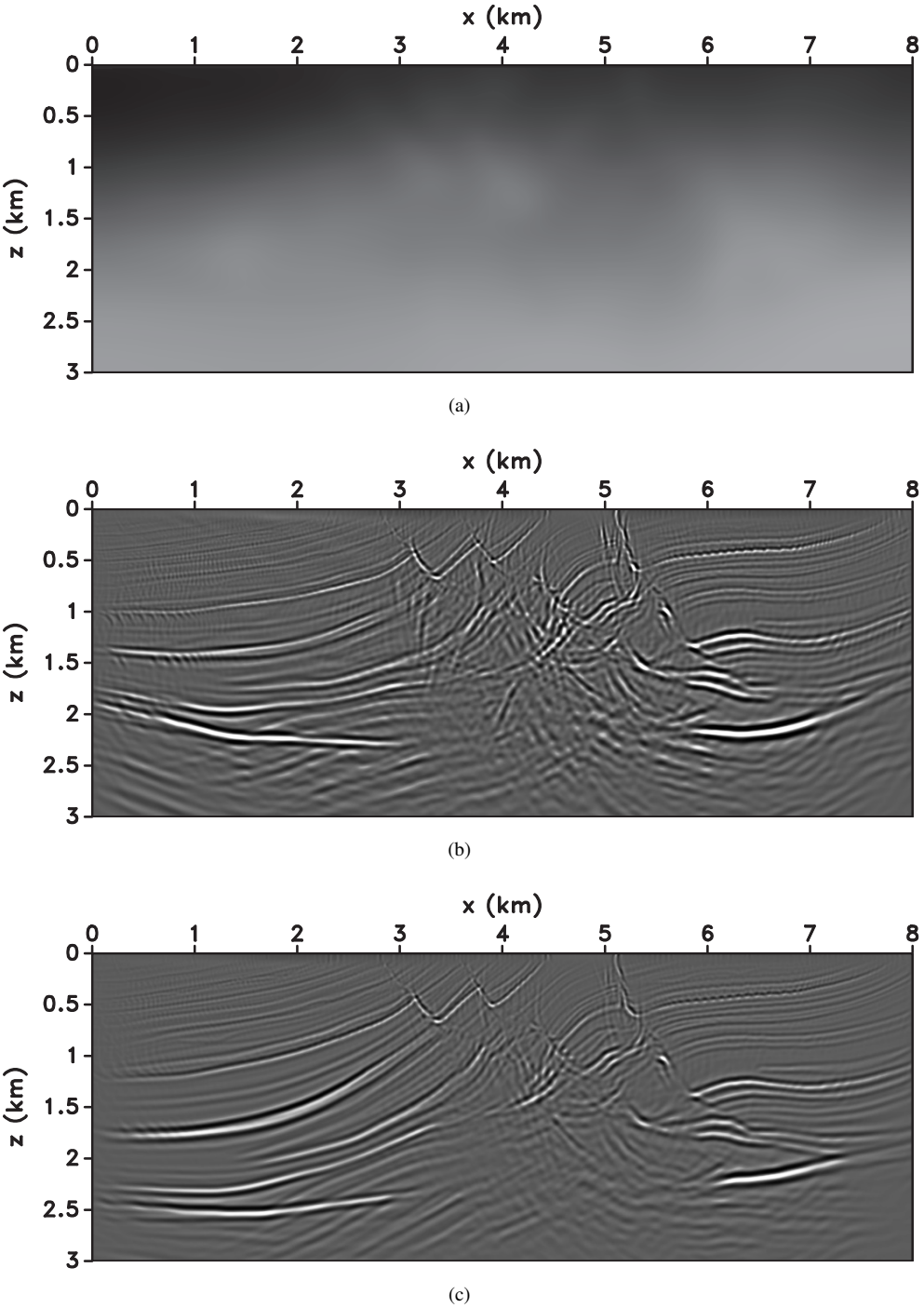


Figure 15. (a) Velocity model after 20 iteration of wavefield tomography and (b) corresponding migrated image. The reflectors shift toward their correct location, and the truncation of the beds against the fault walls improves. (c) Image obtained after smoothing the model in Figure 15(a) with a triangular filter with radius 100 vertically and 400 samples horizontally.

fractive media, triplications and multipathing between source and receiver produce artifacts in the migrated image (Stolk and Symes, 2004). The artifacts are due to the migration operator being the adjoint and not the inverse of the Born modeling operator; they have the same frequency content of the signal and, although the stack over experiments removes them from the final image, they severely contaminate the migrated images for individual shots. Since these kinematic artifacts do not satisfy the assumed relation between the directions of the image gradient and displacement vector field, they represent the principal source of noise for our method. Least-square migration may attenuate the artifacts in the reflectivity images but would increase the computational cost and time. The development of simpler and more cost effective methodologies to avoid the migration artifacts is currently a topic of research.

In complicated models, such as the Marmousi model, regularization plays an important role in balancing the strength of the update and removing the natural bias that comes from the uneven illumination and the complexity of the subsurface structures.

5 CONCLUSIONS

The image-difference objective function for wavefield tomography in the migrated-shot domain can be made more robust using image warping. Displacement fields warping a migrated image into the image obtained from a neighboring experiment absorbe amplitude differences due to the shift in source location and removes cycle skipping, because every reflector in one image is mapped into the corresponding event in the second image.

Our method is more effective than local image correlation wavefield tomography for shallow reflectors, i.e., when the ratio between the source separation and the depth of the reflector is high. However, kinematic artifacts contaminate the image in highly refractive media and destabilize the computation of the sensitivity kernel. Nonetheless, the objective function based on the warping relation between the images is able to measure the distance between different models, and to highlight the most coherent images, thus enabling local gradient-based optimization.

6 ACKNOWLEDGMENTS

This work was supported by the sponsors of the Consortium Project on Seismic Inverse Methods for Complex Structures. The reproducible numerical examples in this paper use the Madagascar open-source software package freely available from http://www.ahay.org and the Mines JTK freely available at https://github.com/dhale/jtk.

REFERENCES

Al-Yahya, K., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, P. 718–729.

Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Adjoint wave-equation velocity analysis: Presented at the 74th Ann. Internat. Mtg., SEG.

Biondi, B., and P. Sava, 1999, Wave-equation migration velocity analysis: Presented at the 69th Ann. Internat. Mtg., SEG.

Carcione, J., 2007, Wave fields in real media wave propagation in anisotropic, anelastic, porous and electromagnetic media: Elsevier. Handbook of Geophysical Exploration.

Chauris, H., and R.-E. Plessix, 2012, Investigating the differential waveform inversion: Presented at the 74th EAGE Conference & Exhibition - Workshops, EAGE.

Chavent, G., and C. A. Jacewitz, 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Publishing.

Faye, J.-P., and J.-P. Jeannot, 1986, Prestack migration velocity analysis from focusing depth imaging: Presented at the 56th Ann. Internat. Mtg., Soc. of Expl. Geophys.

Fichtner, A., H.-P. Bunge, and H.Igel, 2006, The adjoint method in seismology I. theory: Physics of the Earth and Planetary Interiors, 86–104.

Fowler, P., 1985, Migration velocity analysis by optimization: Technical Report 44, Stanford Exploration Project.

Hale, D., 2007a, Local dip filtering with directional Laplacians: Technical Report CWP-567, Center for Wave Phenomena, Colorado School of Mines.

———, 2007b, A method for estimating apparent displacement vectors from time-lapse seismic data: Technical Report CWP-566, Center for Wave Phenomena, Colorado School of Mines.

Kelly, S., J. R. Martinez, B. Tsimelzon, and S. Crawley, 2010, A comparison of inversion results for two full-waveform methods that utilize the lowest frequencies in dual-sensor recordings: Presented at the 80th Ann. Internat. Mtg., Soc. of Expl. Geophys.

Lions, J., 1972, Nonhomogeneous boundary value problems and applications: Springer Verlag, Berlin.

Luo, Y., and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

Perrone, F., and P. Sava, 2012, Wavefield tomography based on local image correlation: Presented at the 74th Conference and Exhibition, EAGE.

Perrone, F., P. Sava, C. Andreoletti, and N. Bienati, 2012, Linearized wave-equation migration velocity analysis by image warping: Submitted to Geophysics.

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

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

Santosa, F., and W. W. Symes, 1989, An analysis of least-

squares velocity inversion, *in* Geophysical Monograph no. 4: Soc. of Expl. Geophys.

Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, **52**, 593–606.

Sava, P., B. Biondi, and J. Etgen, 2005, Wave-equation migration velocity analysis by focusing diffractions and reflections: Geophysics, **70**, U19–U27.

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

Stolk, C. C., and W. W. Symes, 2004, Kinematic artifacts in prestack depth migration: Geophysics, **69**, 562–575.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 71–92.

van Vliet, L. J., and P. W. Verbeek, 1995, Estimators for orientation and anisotropy in digitalized images: Proceedings of the first annual conference of the Advanced School for Computing and Imaging, ASCI'95, 442–450.

Verschuur, D. J., 1991, Surface-related multiple elimination, an inversion approach: PhD thesis, Technische Univ., Delft (Netherlands).

Xu, S., D. Wang, F. Chen, and Y. Zhang, 2012, Full waveform inversion for reflected seismic data: Presented at the 74th Conference and Exhibition, EAGE.

Yang, T., and P. Sava, 2010, Wave-equation migration velocity analysis with extended common-image-point gathers: Presented at the 80th Ann. Internat. Mtg., Soc. of Expl. Geophys.