Accelerating residual-moveout-based wave-equation migration velocity analysis with compressed-sensing

Yang Zhang, Biondo Biondi, and Robert Clapp

ABSTRACT

Residual-moveout-based wave-equation migration velocity analysis uses residual-moveout to characterize the kinematic error caused by an inaccurate velocity model and computes velocity updates using wave-equation tomographic operators. However, the per-iteration cost of this method is even more expensive than that of full waveform inversion because of the construction and back-projection of offset/angle-domain common-image gathers. In order to speed up residual-moveout-based migration velocity analysis, we examine its work flow, and propose the following acceleration scheme: 1) by using the compressed-sensing technique, we can very well reconstruct the angle-domain common image gathers from a fraction of the subsurface offset common-image gathers, therefore saving significant computation cost; 2) after we extract the residual-moveout information from the reconstructed angle-domain common image gathers, we reduce the cost of back-projection by synthesizing an approximation of the original image perturbation, which can be back-projected with much lower cost and can yield a velocity gradient with the same behavior.

INTRODUCTION

Wave-equation migration velocity analysis (WEMVA) is a reflection tomography method which uses wave-equation rather than ray-based model to retrieve the velocity model from seismic data(Chavent and Jacewitz, 1995; Biondi and Sava, 1999). The velocity information comes from the seismic data redundancy that arises because each reflector point in the subsurface is illuminated by wave energy from multiple directions. WEMVA exploits such redundancy by forming common-image gathers and then improves the velocity model by enforcing coherence among the common-image gathers. Evaluating the flatness of the subsurface angle-domain common image gathers (ADCIGs) is currently a popular choice when forming WEMVA objective functions (Biondi and Sava, 1999; Clapp and Biondi, 2000; Biondi and Symes, 2004).

The objective function is usually optimized by applying gradient-based algorithms. The computation of the gradient is performed in two steps: 1) computation of a perturbation in the migrated image domain, and 2) back-projection of the image per-

turbation into the velocity model using the image-space wave-equation tomographic (ISWET) operator (Sava and Biondi, 2004).

One flavor of the WEMVA methods is based on residual-moveout (RMO) (Zhang et al., 2012). It describes the unflatness in the ADCIGs using residual-moveout. Instead of maximizing the image-stack-power objective function directly with respect to the velocity, we link the objective function to the velocity model indirectly through an intermediate moveout parameter. By focusing on the kinematic errors in the common-image gathers, RMO-based WEMVA is more robust with respect to the cycle-skipping issue and still produces high-quality model updates.

However, for all WEMVA methods, the cost per iteration is much higher than that of full waveform inversion (FWI) because the WEMVA projection operator is more expensive compared to the projection operator in waveform inversion. In WEMVA's gradient projection, besides calculating wavefield propagation, the imaging operator and wave-equation tomographic operator also have to perform cross-correlations and convolutions that involve the subsurface offset-domain common image gathers (ODCIGs) or angle-domain common image gathers (ADCIGs). In contrast, the projection operator in FWI does not deal with the subsurface common-image gathers, therefore is much cheaper than that of WEMVA. Moreover, as the velocity error increases, the size of the ODCIGs has to grow (otherwise the velocity information will be lost), which further increases the cost of WEMVA iteration.

Due to the high cost, a computationally more efficient approach for RMO-based WEMVA is very desirable. There are two computationally significant steps in the work flow: one is the generation of ADCIGs with the current velocity model, which requires one migration with subsurface offset; the other one is the back-projection of the RMO error into the velocity model, which involves applying the image-space tomographic operator with a subsurface-offset image perturbation. For the former step, we make use of the compressed-sensing technique (Clapp, 2012), which can reconstruct the full ADCIG from a randomly subsampled ODCIG; thus the imaging cost (of such an ODCIG) is only a fraction of the original implementation. However, for the latter step, we cannot save computation by back-projecting a subsampled subsurface ODCIG perturbation, because the output gradient will not be consistent with the original gradient. Therefore we propose to approximate the image perturbation with a synthesized one, which has shorter span along the subsurface-offset axis without losing the moveout information. Back-projecting the synthesized image perturbation would cost much less, because the number of subsurface-offsets in the image perturbation is significantly smaller.

The rest of this paper is divided into three parts: first we briefly review the theory of RMO-based WEMVA; then we present the details of the modification method which becomes computationally more efficient; finally we illustrate the modified approach with some synthetic examples.

RMO-BASED WEMVA

For simplicity, we assume two dimensions in our derivation; however, extending the theory to 3-D is conceptually straightforward for this method. We denote the prestack angle-domain images as $I(z, \gamma, x)$, where x, z are the depth and horizontal axis, γ is the reflection-aperture angle, and s represents the slowness model.

The details of RMO-based WEMVA can be found in Zhang and Biondi (2011, 2013); here we simply re-state the key steps. We start from the "classical" stack power maximization objective function

$$J(s) = \frac{1}{2} \sum_x \sum_z \left[\int d\gamma \; I(z, \gamma, x; s) \right]^2, \tag{1}$$

where s is the model slowness, and $I(z, \gamma, x; s)$ is the prestack image (ADCIG) migrated with s.

RMO-based WEMVA defines an alternative objective function (maximizing the normalized RMO semblance) with respect to the RMO parameter ρ, which is a function of the slowness s:

$$J_{Sm}(\rho(s)) = \frac{1}{2} \sum_x \sum_z \frac{\int dz_w \big(\int d\gamma I(\gamma, z + z_w + \rho \tan^2 \gamma, x; s_0) \big)^2}{\int dz_w \int d\gamma I^2(\gamma, z + z_w + \rho \tan^2 \gamma, x; s_0)}, \tag{2}$$

where s_0 is the starting model and z_w is a local averaging window of length L along the depth axis. The gradient given by the objective function (2) is

$$\frac{\partial J_{S_m}}{\partial s} = \frac{\partial J_{S_m}}{\partial \rho} \frac{\partial \rho}{\partial s}, \tag{3}$$

where $\partial J_{S_m} / \partial \rho$ can be easily calculated by taking the derivative along the ρ axis of the semblance panel S_m, and $\partial \rho / \partial s$ can be derived in a way similar to the sensitivity kernel of the finite-frequency travel-time tomography method in Marquering et al. (1998).

The final expression for the gradient calculation is:

$$\frac{\partial J_{S_m}}{\partial s} = - \int dz_w \int d\gamma \sum_{z,x} \{ \frac{\partial I(z + z_w, \gamma, x; s)}{\partial s} (F_{11} \tan^2 \gamma + F_{12}) \dot{I}(z + z_w, \gamma, x; s_0) \frac{\partial J_{Sm}}{\partial \rho}(z, x) \}, \tag{4}$$

where F_{11} and F_{12} are constants computed from the initial ADCIGs, the detailed formula for which can be found in Zhang and Biondi (2013).

ACCELERATED RMO WEMVA APPROACH

The dominant cost of RMO WEMVA is the ADCIG reconstruction and the model gradient calculation (shown in eq. (4)). The following two sections describe the acceleration we propose.

Efficient reconstruction of ADCIGs

As seen in eqs. (4) and (2), the cost of computing $\partial J_{S_m}/\partial\rho$ is primarily the cost of constructing the full ADCIGs $I(z,\gamma,x;s_0)$. In our practice, this is done by constructing the subsurface-offset image $I(z,h_x,x;s_0)$ using the cross-correlation imaging condition and then transforming the offset image to the angle-domainusing the method described in Sava and Fomel (2003). Computing the subsurface-offset images (i.e. ODCIGs) is significantly more expensive than the zero-subsurface-offset image, because wavefield cross-correlation has to be performed at all subsurface-offset locations. As we move to 3-D, the I/O could also become another bottleneck because of the large size of the 5-D subsurface-offset image.

There is a significant amount of literature on compressing seismic data, recently Clapp (2012) found that the seismic images such as ADCIG and ODCIG are also highly compressible. To utilize the compressibility of seismic images, Clapp (2012) proposed to use the compressed-sensing method (Candes and Donoho, 1999; Donoho, 2006) to reduce the imaging cost. Compressed-sensing theory predicts that given certain conditions are satisfied, we can almost perfectly recover the full ADCIGs from a randomly and heavily subsampled ODCIG by solving the following optimization problem:

$$\begin{aligned} \mathbf{0} &\approx ||\mathbf{M_d}(\mathbf{d}-\mathbf{RWm})||_{\mathbf{2}}^{\mathbf{2}}, \\ \mathbf{0} &\approx ||\mathbf{m}||_{\mathbf{1}}, \end{aligned} \tag{5}$$

where $\mathbf{M_d}$ represents the sub-sampling mask for the ODCIG $\mathbf{d}$. Thus we need to compute the ODCIGs only at those locations specified by $\mathbf{M_d}$. $\mathbf{R}$ is the angle-to-offset transform operator, $\mathbf{W}$ is the multi-dimensional wavelet transform operator, $\mathbf{m}$ is the wavelet-domain coefficient of the ADCIG, and $\mathbf{m}$ is assumed to be sparse. The sparsity-promoting constraint for $\mathbf{m}$ is to minimize the ℓ_1-norm of $\mathbf{m}$, as expressed in eq. (5).

Clapp (2012) uses an ℓ_1 inversion scheme called the Stage-wise Orthogonal Matching Pursuit (StOMP) (Donoho et al., 2006), which essentially approximates the original problem by solving multiple least-square data-fitting problems. The operators $\mathbf{R}$ and $\mathbf{W}$ are much cheaper than the imaging operator itself, so this reconstruction scheme is more efficient than calculating the ADCIG directly from the full ODCIG.

Figure 1 shows a reconstruction example using the Marmousi model; here the subsampling ratio of the ODCIGs is 20%. Figure 2(a) shows the reconstructed ADCIG using the sub-sampled ODCIGs in figure 1. For comparison, figure 2(b) shows the ADCIG transformed from fully sampled ODCIGs. The comparison shows that the compressed-sensing-based reconstruction is satisfactory. Although the two are not exactly the same, the important RMO information has been well preserved on the reconstructed ADCIG.

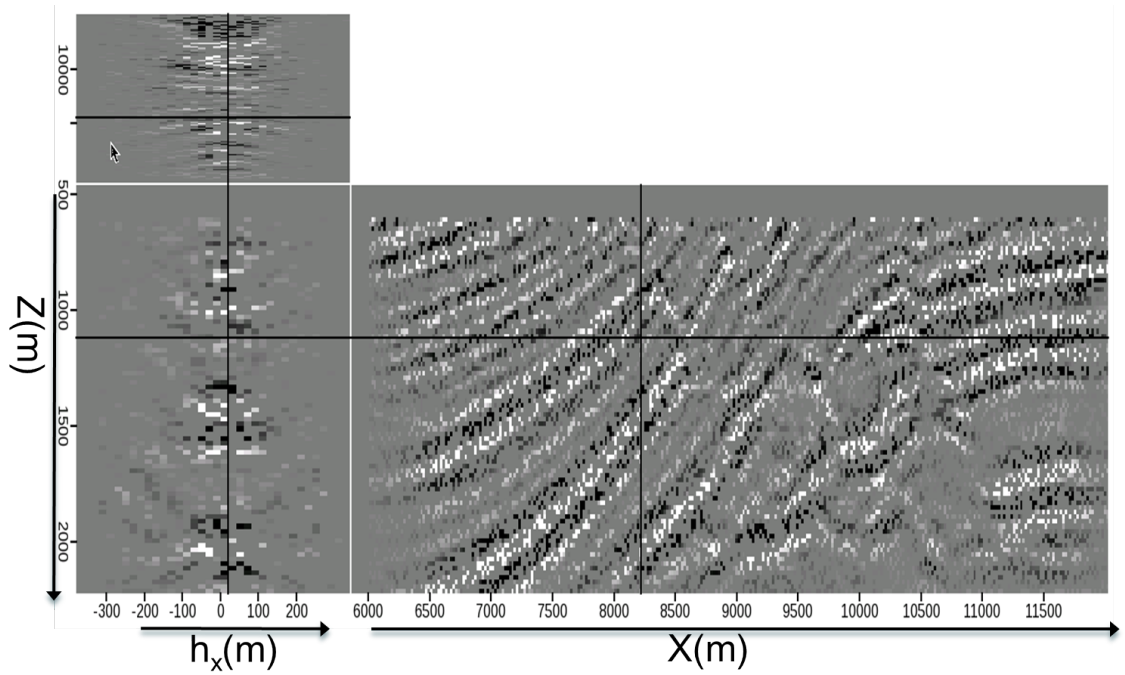

Figure 1: A Marmousi subsurface offset image after random sub-sampling. The sub-sampling ratio is 20%. [**ER**]

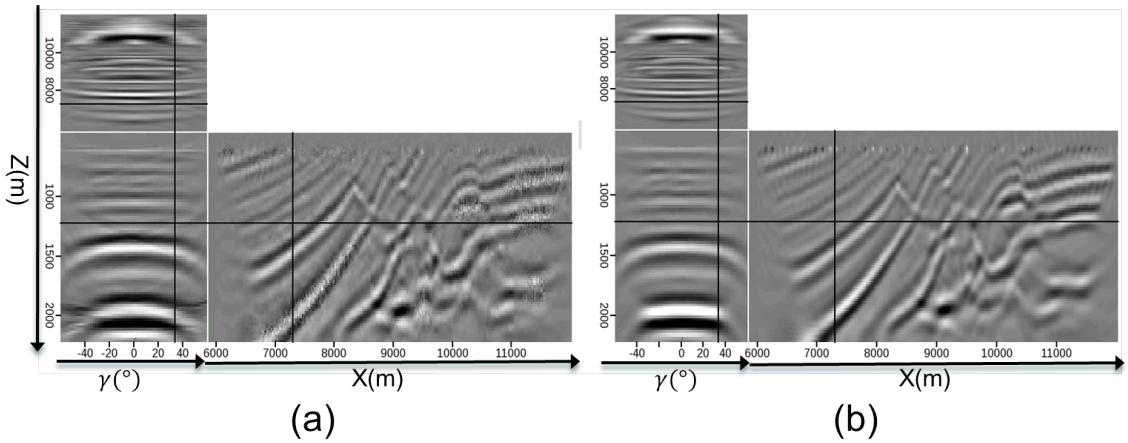

Figure 2: (a): The reconstructed ADCIGs using the sub-sampled ODCIG in figure 1. (b): The ADCIGs transformed from the fully sampled ODCIG. [**ER**]

Back-projection using synthesized image perturbation

As shown in the gradient calculation in formula (4), the other computationally expensive step is the back-projection of the image perturbation,

$$\Delta I = (F_{11}\tan^2\gamma + F_{12})\dot{I}(z+z_w,\gamma,x;s_0)\frac{\partial J_{Sm}}{\partial \rho}(z,x)$$

using the image-space wave-equation tomographic operator,

$$\mathbf{T} = \partial I(z+z_w,\gamma,x;s)/\partial s.$$

From the offset-to-angle common image gathers transform described by Sava and Fomel (2003), we know that the unflatness in the angle domain corresponds to unfocused energy in the subsurface-offset domain. The angle-domain image perturbation ΔI is the first-order z derivative of the initial ADCIG $I(s_0)$, modulated by $f(\gamma) = (F_{11}\tan^2\gamma + F_{12})\frac{\partial J_{Sm}}{\partial \rho}(z,x)$. Therefore the gather shape (curvature) in ΔI is the same as in initial ADCIG $I(s_0)$. Then the subsurface-offset domain of ΔI will have to use the same offset span as the initial ODCIG. When the velocity error is large, the offset span can be large, which increases the amount of computation for the tomographic operator.

The back-projection is expensive, because ΔI takes many subsurface offset locations to represent. Realizing that the RMO information (which determines the velocity update direction) is already embedded in the modulate function $f(\gamma)$, we can form a different image perturbation as follows:

$$\Delta\tilde{I}(z+z_w,\gamma,x) = f(\gamma)\dot{\tilde{I}}(z+z_w,\gamma,x;s_0),$$

in which $\tilde{I}(\gamma;s_0)$ is the angle-domain common image gathers for the zero-subsurface-offset image $I(h=0;s_0)$, and the dot denotes the derivative respect to depth z. The advantage of defining such $\tilde{I}(\gamma)$ is that $\tilde{I}(\gamma)$ would be flat among angles, which leads to $\Delta\tilde{I}(\gamma)$ also being flat. Then the offset-domain $\Delta\tilde{I}(h)$ will be much more focused around $h=0$, and we can use a shorter subsurface-offset axis h for $\Delta\tilde{I}(h)$. Therefore the cost of applying the tomographic operator on $\Delta\tilde{I}(h)$ will be much smaller, because $\Delta\tilde{I}(h)$ needs fewer subsurface-offset points than $\Delta I(h)$.

Besides changing $I(h)$ and $\Delta I(h)$, we need to replace the observed data d used in the tomographic operator with $\tilde{d} = \mathbf{L}I(h=0)$ correspondingly, where $\mathbf{L}$ is the Born modeling operator. By doing this, the synthesized data $\tilde{d}$ becomes kinematically consistent with the synthesized image perturbation $\Delta\tilde{I}$.

INVERSION EXAMPLE

Marmousi model

With the two modifications incorporated in our RMO WEMVA work flow, we test the new approach on the smoothed Marmousi Model. The model is 6 km in x and 1.6 km

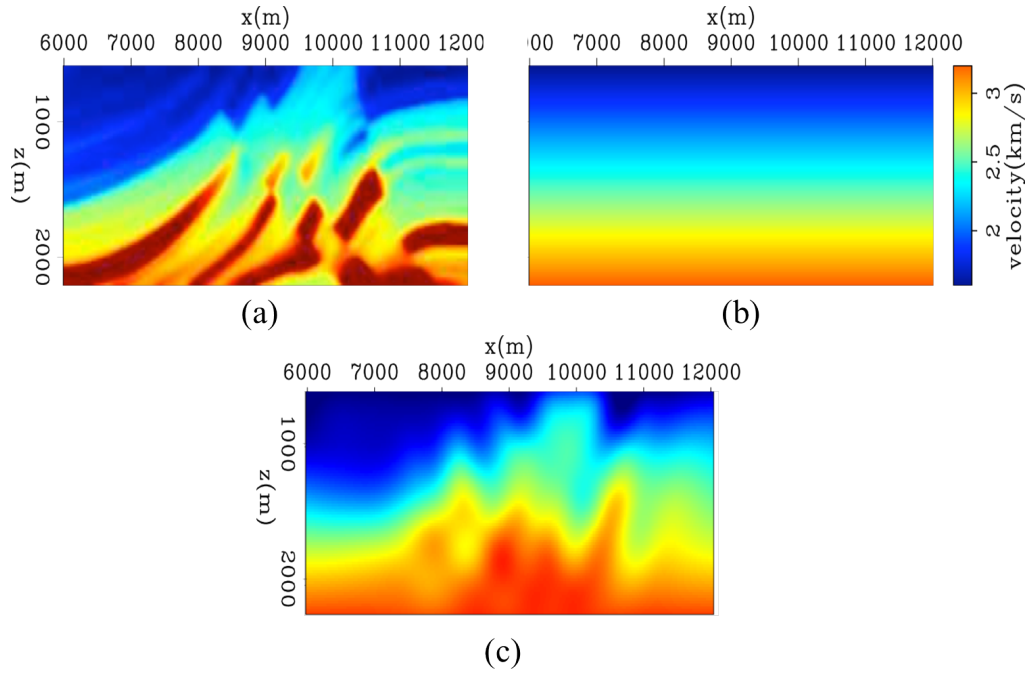

Figure 3: The slightly smoothed true velocity model of Marmousi (a), the starting velocity model (b), and the inverted velocity models using our method after 20 iterations (c). [**CR**]

in z. The spatial sampling is 20 m. The survey geometry follows the land acquisition pattern with receivers at every surface location on the top and we simulate 51 shots in total, covering the whole lateral span on the top with a spacing of 120 m. We model 64 frequencies in total using the one-way wave-equation, ranging from 5 Hz to 40 Hz.

For each iteration, we compute only 20% of the locations of the migrated ODCIGs and reconstruct the full ADCIGs using the StOMP algorithm, which could save approximately 80% of the imaging cost. After the RMO information is extracted, we back-project the synthesized image perturbation, which uses only 9 subsurface offsets rather than 33 points needed for original image perturbation, saving about 75% of the tomographic operator cost. We ran 20 nonlinear iterations for the inversion.

Figure 3(a) shows the true model (in velocity) and (b) shows the starting model, which has a vertical gradient increasing from 1600 m/s and 3200 m/s. Figure 3(c) shows the inverted velocity model using our computationally efficient approach. The result shows good convergence to the true model.

BP synthetic model

We also test this new approach on a portion of the BP synthetic model. The model is extracted from the upper-left part of the original BP model, with a size of 35 km in

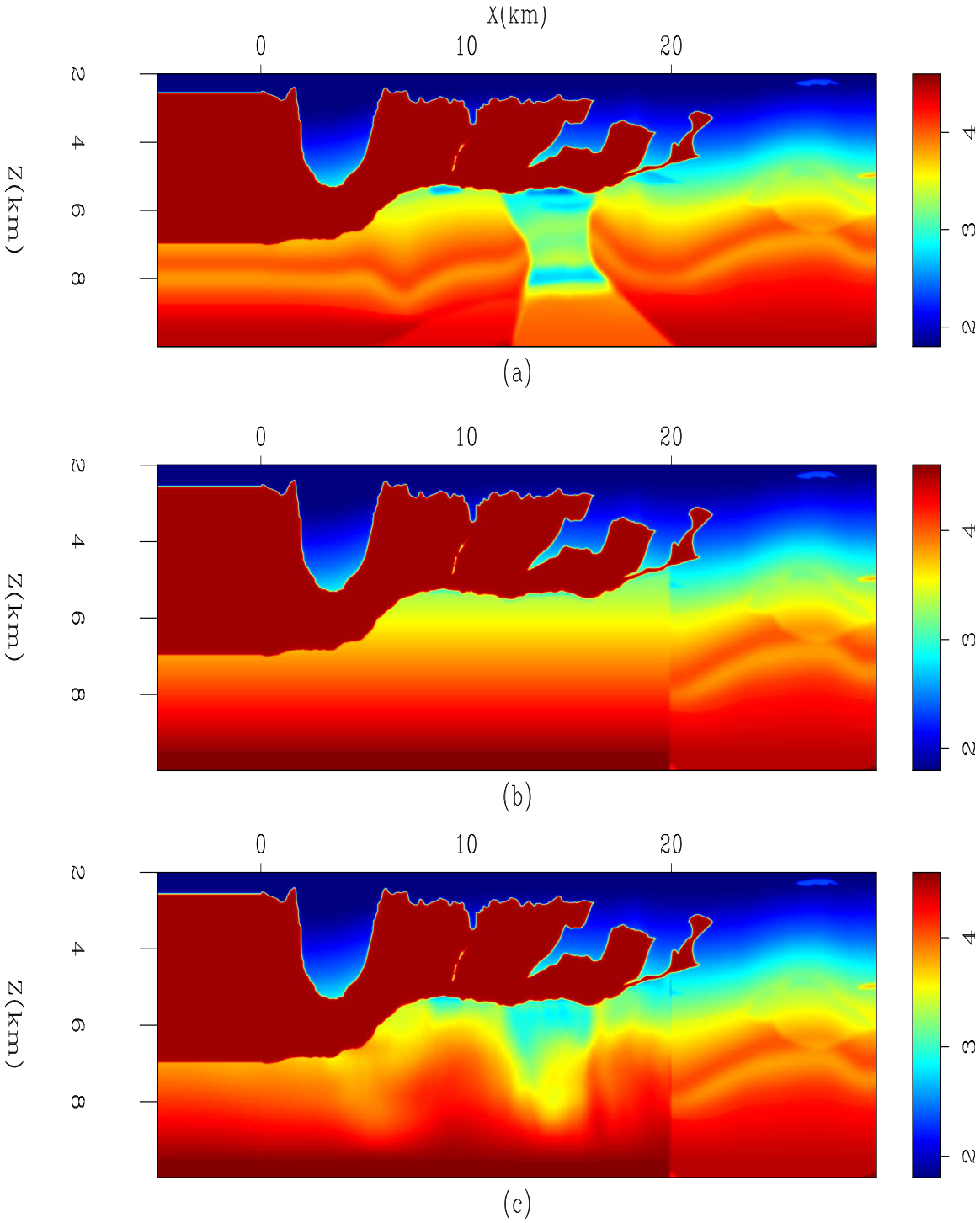

Figure 4: The true velocity model (a) of the upper-left part of the BP model, the starting velocity model (b), and the inverted velocity model (c) using our method after 40 iterations. [**CR**]

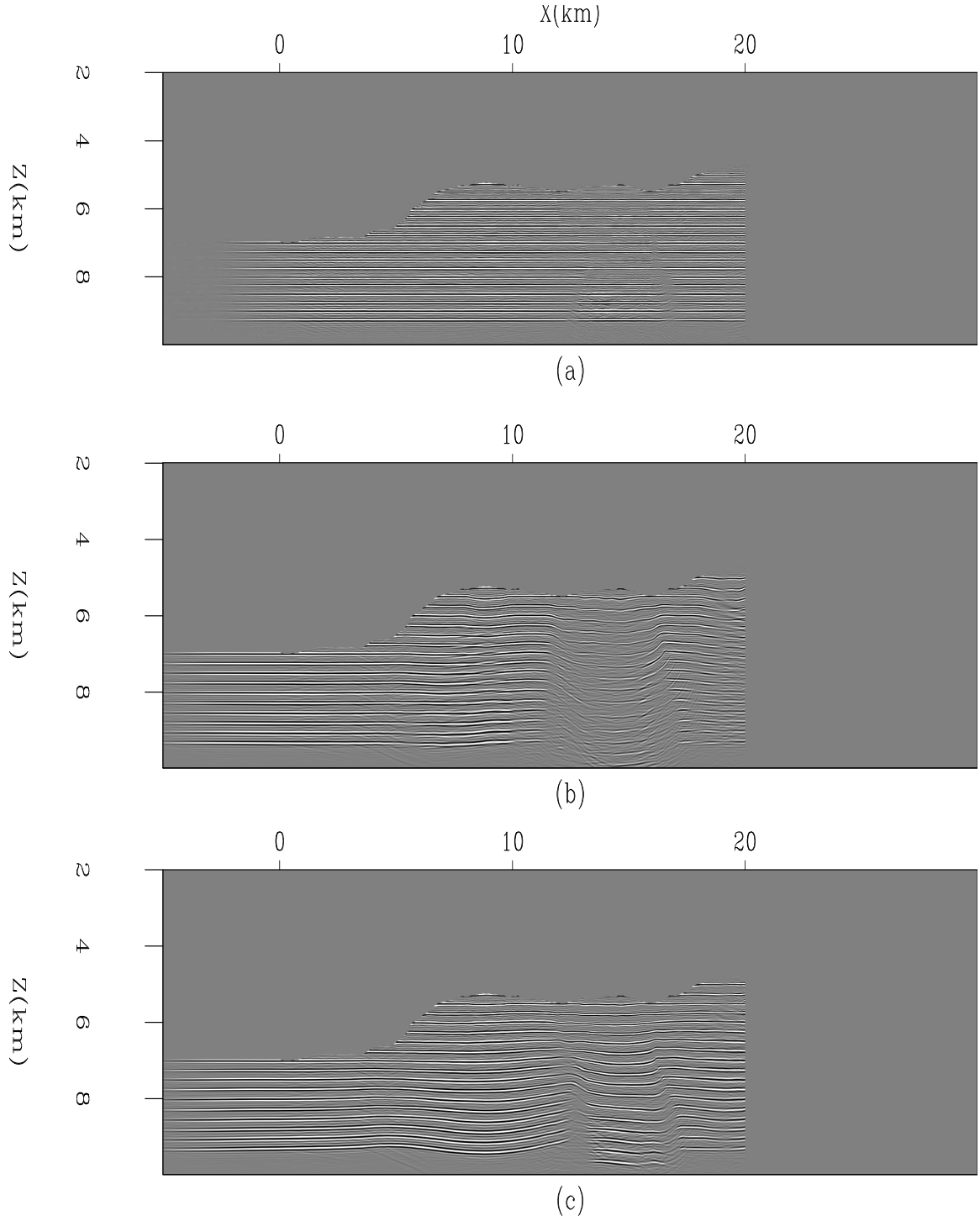

Figure 5: The migrated images using the velocity models shown in fig 4. [**CR**]

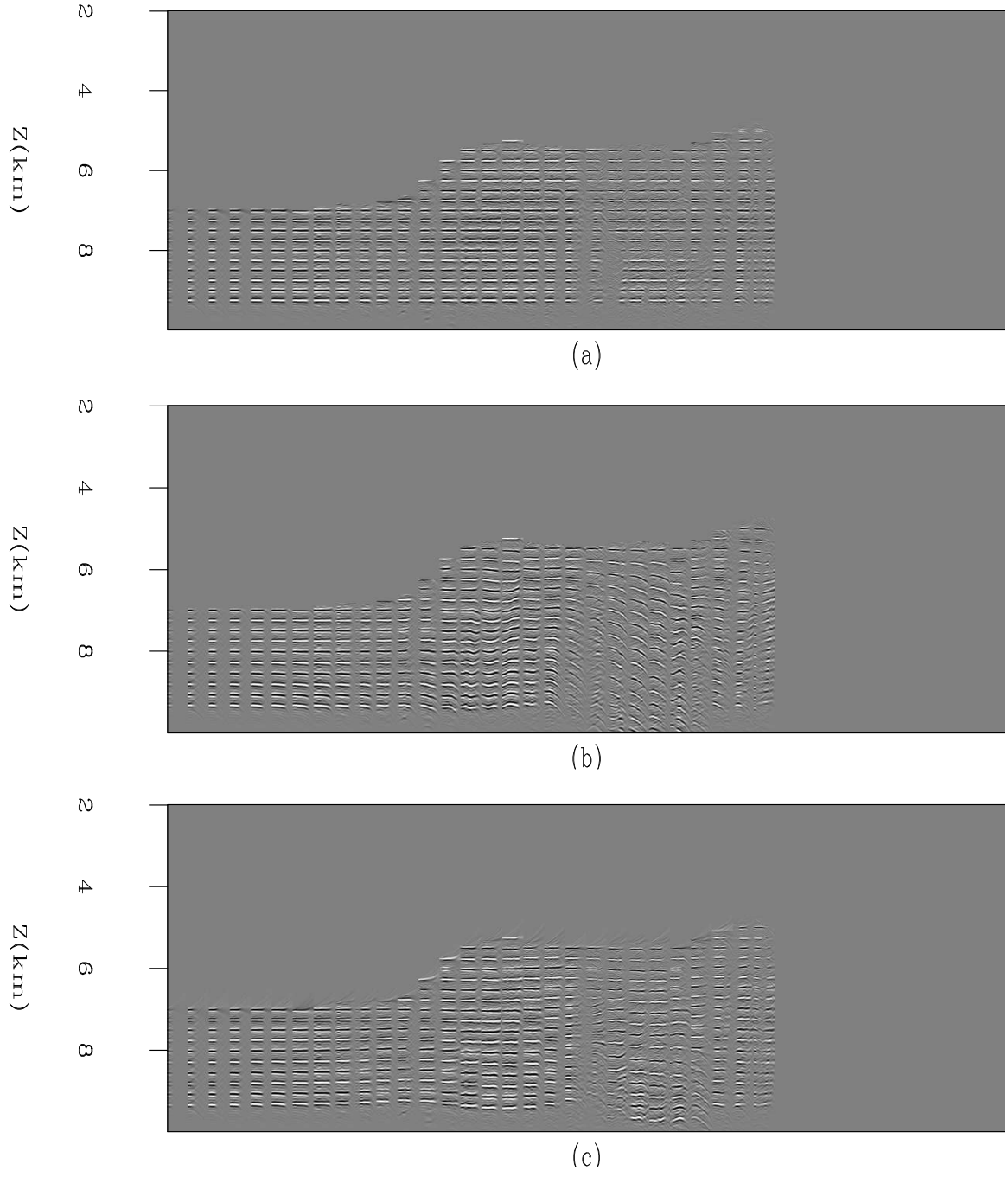

Figure 6: The angle-domain common image gathers migrated using the true model (a), using the starting velocity model (b), and using the inverted velocity model (c). We calculate the angles from 0° to 52°, with a sampling of 2°. The flatness of ADCIGs improves significantly from the initial model. [**CR**]

x and 8 km in z. The spatial sampling of the model is 25 m in x and 12.5 m in z. We synthesize several flat reflectors to simulate the seismic data. We use marine streamer geometry to simulate 200 shots on the top of the model, with receiver streamer towing from left to right. The shots start from $x = -5$ km, and the spacing is 125 m. There are 401 receivers on the 10 km long streamer with a spacing of 25 m. We model 281 frequencies in total using the one-way wave-equation, ranging from 5 Hz to 40 Hz.

Figure 4(a) shows the true model (in velocity) and (b) shows the starting model, in which we assume the complex salt overburden has been well resolved, and the sediments velocity below the salt linearly increase from 3.4 km/s to 4.4 km/s. Because we assume the velocity above the salt is accurate, we apply a mask during the inversion so that only the subsalt region of the model will be updated. Figure 4(c) shows the inverted velocity model after 40 iterations using our computationally efficient approach. Again, the result shows good convergence to the true model. Figure 5 shows the migrated images using the velocity models in fig 4. They are plotted using the same clip, so that we can easily observe the improvements of the reflector coherence by the inversion. Figure 6 show the improvement of ADCIGs flatness from the starting velocity model to the inverted velocity model.

CONCLUSION

In order to reduce the computational cost of residual-moveout based wave-equation migration velocity analysis, we propose two modifications to the workflow: first we use a compressed-sensing-based technique to reconstruct angle-domain common-image gathers without computing all of the subsurface-offset gathers; second, we synthesize an image perturbation that has more focused energy near zero subsurface offset, and back-project that image perturbation into the slowness model while keeping the model updates similar to the original case. We test this computationally efficient approach using the 2-D Marmousi model and the BP model. The examples show that adopting these modifications in RMO-based WEMVA does not degrade its inversion result; while it can increase inversion speed by five-fold. Although the increased speed we achieve currently is not overwhelming, in the 3-D case, the compressibility of the seismic images increases dramatically with the growth of dimensions, so we expect the speed up ratio to increase by additional one or two orders of magnitude.

ACKNOWLEDGEMENT

We thank the sponsors of the Stanford Exploration Project for their financial support. The first author thanks Dave Nichols for the helpful discussions and suggestions on seismic tomography.

REFERENCES

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Biondi, B. and W. W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298.

Candes, E. J. and D. L. Donoho, 1999, Curvelets — a surprisingly effective nonadaptive representation for objects with edges: Curves and Surfaces, 105–120, Vanderbilt University Press.

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

Clapp, R. G., 2012, Image gather reconstruction using StOMP: SEP-Report, **147**, 127–138.

Clapp, R. G. and B. Biondi, 2000, Tau domain migration velocity analysis using angle crp gathers and geologic constraints: SEG Technical Program Expanded Abstracts, **19**, 926–929.

Donoho, D. L., 2006, Compressed sensing: IEEE Transactions on Information Theory, **52**, 1289–1306.

Donoho, D. L., Y. Tsaig, I. Drori, and J.-L. Starck, 2006, Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit: Technical report.

Marquering, H., G. Nolet, and F. Dahlen, 1998, Three-dimensional waveform sensitivity kernels: Geophysical Journal International, **132**, 521–534.

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

Sava, P. C. and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Zhang, Y. and B. Biondi, 2011, Moveout-based wave-equation migration velocity analysis: SEP-Report, **143**, 43–58.

———, 2013, Moveout-based wave-equation migration velocity analysis: GEOPHYSICS, **78**, U31–U39.

Zhang, Y., B. Biondi, and Y. Tang, 2012, Residual moveout-based wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **31**.