Localized image-difference wave-equation tomography

Musa Maharramov[*] and Uwe Albertin[†].
[*]*High Performance Computing Group, Exploration and Production Digital and Communications Technology, BP, Houston, TX 77079,* [†]*Advanced Imaging, Exploration and Production Technology Group, BP, Houston, TX 77079*

SUMMARY

Current wave-equation tomography techniques based on migrated image differences, such as those observed in 4D data sets, use the image difference as a measure of velocity misfit. Computation of the objective function gradient is accomplished by the adjoint application of the derivative of the imaging operator to this image difference. In all techniques developed to date this process is carried out by computing the gradient over a relatively large number of depth steps, and then optimizing the objective function globally over the entire range. In this abstract, we extend this concept to compute the gradient and objective function locally, within several or even one depth step at a time. In principle, for objective functions that are sharply peaked around the global minimum, and have other minima elsewhere, this localization should reduce the possibility of falling into a false minimum, and significantly reduce the number of iterations required in the optimization. In addition, since the velocity is optimized in depth as the extrapolation proceeds, the method is significantly more immune to cycle-skipping at higher frequencies than global methods.

INTRODUCTION

One key aspect of reservoir monitoring is the successful tracking of hydrocarbon movement as the reservoir is depleted. In order to accomplish this, most current methods rely on the conversion of picked time shifts between migrated sections to local impedence changes as well as reflector movement. While quite successful, this technique relies on a significant amount of interpretation and quality control in the conversion process. More recently, however, wave-equation image-difference tomography has been proposed as a more automatic technique to recover velocity perturbations associated with reservoir depletion (Albertin et al, 2006).

Wave-equation image-difference tomography is based on an optimization that tries to minimize the difference between migrated images. In its simplest form, the method relies on a relation between an image perturbation and the slowness c according to

$$\delta I = \frac{\delta I}{\delta c}\delta c. \quad (1)$$

Solving this equation for δc in the adjoint sense gives an update to the slowness according to

$$\delta c^* = \left(\frac{\delta I}{\delta c}\right)^* \delta I \quad (2)$$

Recovery of the slowness perturbation associated with the image difference is then accomplished using the following steps. First, a downward extrapolation is done on both data sets to obtain an image difference, using one of the data sets as a reference. Second, an upward continuation is done in which the image difference is correlated against the downward-extrapolated reference field to recover the slowness perturbation. Since the procedure does not use the exact inverse to $\frac{\delta I}{\delta c}$, an iterative optimization is used to converge to a final velocity update; the objective function corresponding to the square of the norm $|\delta I|^2$is checked at each iteration to make sure the image difference is actually decreasing, and the new model is used to obtain a new set of images and a new image difference. Note that at each iteration of the tomography the upward and downward extrapolations are done over the entire model.

This procedure is effective in recovering velocity changes associated with reservoir depletion to high resolution, but it is a nonlinear inversion, and hence the issue of local minima becomes significant when solving for the update. Unlike wave-equation tomography based on differential semblance of migrated gathers (Shen et al, 2003), a tomography based on a direct image-difference has an objective function that is much steeper near the global minimum. In addition, it is easy to see that if the velocity perturbation is sufficiently large, and spread over a large enough depth window, the difference in the images will be larger than the dominant wavelength in the image, and will cycle-skip. Recovery of the proper velocity in such cases will become quite difficult.

To circumvent this issue, one possible method is to slowly increase the frequency content of the image, and to do the tomography in stages, from lower frequency to higher frequencies. Such a method would be directly analogous to methods used in waveform inversion to avoid cycle skipping (Sirgue and Pratt, 2004). The inversion is initially carried out at a sufficiently low frequency that the image difference does not cycle skip. Further iterations at successively higher and higher frequencies use the model obtained in the previous inversion as a starting model. Such a technique has been used successfully in both waveform inversion and wave-equation tomography when objective functions with narrow basins of attraction are utilized.

In this abstract we employ a different technique to mitigate the effects of cycle-skipping, which makes particular use of the fact that the inversions we are doing utilize one-way depth-extrapolation to propagate the wavefield. The central idea goes back to the concept of layer-stripping in standard tomography. Layer-stripping in standard reflection tomography is a technique where the velocity update is initially localized to the top-most layer of a model. This significantly reduces the degrees of freedom in the backprojection, thereby stabilizing it. The velocity in this layer is then held fixed while a new tomography is done in the next lower layer. This process is continued until the end of the model is reached. The primary drawback to such a scheme occurs if velocities are not sufficiently corrected in early layers. In this case, significant velocity error can build up as the layer stripping proceeds.

It is simple to extend the idea of layer-stripping to the wave-equation context as follows. Instead of downward and upward propagating the wavefields over the entire model to obtain the image difference, gradient, and new objective-function value, we downward continue only a few steps in the model. We then stop and get a gradient in this small depth window, and then compute an objective function value over this small window as well. The procedure is then iterated in the small depth window to minimize the objective function before proceeding to the next depth window. The starting model for the next depth window is taken to be velocity obtained in the very last depth step of the previous window, and the entire procedure is then repeated for this new window. The size of this depth window can be reduced all the way to a single depth step, if so desired.

There are several advantages to such a scheme. First, by reducing the depth window over which the optimization is carried out, we significantly reduce the number of degrees of freedom in the backprojection. This leads to an objective function with far fewer minima, and rapid convergence. Second, by employing the velocity from the previous depth window as a starting model for the new window, the velocity difference between the incorrect starting model and the correct model never becomes that large. In this way, depth errors between the two images never become large, and the method thereby eliminates to a large extent the possibility of cycle skipping.

A DESCRIPTION OF THE METHOD

Let $p(x,h,\omega)$denote the acquired wavefield in the frequency domain in midpoint-offset coordinates. Let $E_n = e^{i\Delta z\phi_n}$be the downward extrapolation operator so that $p_{n+1} = E_n p_n$ gives the wavefield at the depth step $n+1$ from that at n. Suppose also that p_0 denotes a second reference wavefield that we are trying to match in the seismic history. Let I denote the migated image corresponding to p, and let Γ denote summation over frequency, with the offset set to zero so that $I = \Gamma p$. Assuming that the velocity has already been reconstructed down to the depth layer n, we take as a misfit function in the depth window $k = n+1 \ldots n+L$ the sum of the squares of the differences between the images $I - I_0$ according to

$$J(c) = \frac{1}{2}\sum_{k=n+1}^{n+L} \left(\Gamma\left(p_k - p_{k,0}\right), \Gamma\left(p_k - p_{k,0}\right)\right)_x. \quad (3)$$

or

$$J(c) = \frac{1}{2}\sum_{k=n+1}^{n+L} \left(\Gamma E_n^k(c)\left(p_n - p_{n,0}\right), \Gamma E_n^k(c)\left(p_n - p_{n,0}\right)\right)_x. \quad (4)$$

Here L is the size of the depth window, $E_n^k(c) = E_n E_{n+1} \ldots E_k$ is the extrapolator from depth layer n down to depth layer k, $(a,b)_x$ is an inner product that means multiply b by the complex conjugate of a pointwise and sum over midpoint. Differentiating this misfit function with respect to slowness leads to

$$\frac{\delta J}{\delta c} = \sum_{k=n+1}^{n+L} \left(\Gamma \frac{\delta E_n^k(c)}{\delta c}\left(p_n - p_{n,0}\right), \Gamma E_n^k(c)\left(p_n - p_{n,0}\right)\right)_x. \quad (5)$$

Up to this point, the treatment generally follows that of Shen et al (2003) except for the use of a narrow depth-window L, an objective function based on an image difference instead of differential semblance and expressing the objective function gradient via that of the extrapolator. The procedure for computing the velocity update globally – i.e., across the entire seismic section – is then the following steps:

1. Downward continue over the section to obtain the images I and I_0 and subtract these to obtain the residual field $\delta I = \Gamma \delta p$.

2. Compute the value of the objective function in eqn.(3) with L equal to the depth of the seimic section, setting $n = 0$, and compare with its previous value. Invoke line search if necessary and repeat step (1).

3. Compute the gradient according to eqn. 4 with L equal to the depth of the seimic section and $n = 0$. Multiply this gradient by a suitable scalar (or use a suitable nonlinear solver) to update the model, and then repeat the procedure, proceeding to step 1.

In our new method, we replace the global objective function $J(c)$ with the local function at a single depth step (i.e., setting $L = 1$)

$$J_n(c) = \frac{1}{2}\left(\Gamma\left(p_{n+1} - p_{n+1,0}\right), \Gamma\left(p_{n+1} - p_{n+1,0}\right)\right)_x. \quad (6)$$

The gradient is also computed at a single depth step according to

$$\frac{\delta J_n}{\delta c} = \left(\Gamma \frac{\delta E_n^{n+1}(c)}{\delta c}\left(p_n - p_{n,0}\right), \Gamma E_n^{n+1}(c)\left(p_n - p_{n,0}\right)\right)_x. \quad (7)$$

The new procedure is as follows:

1. Downward continue a single depth step to obtain I and I_0 and subtract the images at this depth step to obtain a residual field.

2. Compute the value of the objective function via eqn. 6 for this single depth step. Invoke line search if necessary and repeat step (1).

3. Compute the gradient for a single depth step according to eqn. 7. Multiply this gradient by a suitable scalar (or use a suitable nonlinear solver) to update the model, and then repeat the procedure, proceeding to step 1.

4. After the velocity has been optimized for this step, move to the next depth step. Use the velocity from the previous step as a starting model for this new step.

This algorithm can be suitably generalized to do a number of depth steps in each optimization window according to eqns 3 – 5. The primary advantage of this scheme is that the images are matched within each depth window as the optimization proceeds, so that the number of degrees of freedom in the optimization is relatively small, and the chance of cycle-skipping is significantly reduced.

AN EXAMPLE OF THE METHOD

The method described above was tested on the Marmousi model data set. The goal of the optimization was to recover the Marmousi velocity model from an image difference. Since two direct data sets were not available, the method was tested by migrating with the correct model, and with an incorrect model, and using the resulting difference as a residual field for the velocity backprojection. Although such an optimization could be done globally, substantial difficulties would normally occur if the two velocity models used to create the two images are signficantly different, because the two resulting images would be significantly out of phase at nominal seismic frequencies by the bottom of the section. We instead used the method of the previous section to recover the velocity. At each depth step, the starting velocity at that step is initially taken to be the velocity from the preceeding step. The starting velocity at the first step was taken to be constant. The downward continued wavefield at the new step was subtracted from the image with the correct model at that depth step to produce a residual field, which was then used to produce a gradient to the objective function. A narrow-aperture explicit extrapolator was used in eqn. 7, making the gradient evaluation especially computationally efficient, as the coefficients of the extrapolator gradient had been precomputed. A nonlinear LBFGS solver was then used to update the velocity. The optimization was typically run for two to three steps before moving to the next depth step.

The results of this algorithm are shown in Figures 1 and 2. Figure 1 shows the true Marmousi velocity model, while Figure 2 shows the recovered model. Overall, the model has been recovered quite well, except for several areas where noticable 'streaking' occurs. In these areas, the optimization did not properly recover the velocity because the difference in images was not sufficient to produce a strong gradient. Since the starting velocity at each step was taken as that in the previous step, velocity errors propagate downward, and appear as streaks. This is entirely analogous to the phenomenon of velocity error propagation when using layer stripping in standard tomography. Any unresolved velocity errors that are not properly accounted for shallow in the section will produce errors deeper down.

SUMMARY

In conclusion, we have presented a method for velocity recovery from direct image differences. The method is directly applicable to 4D data sets where image differences from reservoir depletion are available. The primary advantage of the method is that it optimizes and recovers the velocity locally in one or several depth steps at a time. The resulting optimization has far fewer degrees of freedom to solve for in each optimization window, and because image differences are solved for locally, the possiblilty of cycle skipping in the inversion is greatly reduced. Initial results on the Marmousi data set indicate that the method is quite effective in recovering velocities even when the model is quite complex.

ACKNOWLEDGEMENTS

The authors would like to thank Paul Sava and Bill Symes for a number of useful discussions. They would also like to thank Laurent Sirgue, Sverre Brandesberg-Dahl, and John Etgen for numerous helpful discussions, and BP for permission to publish this work.

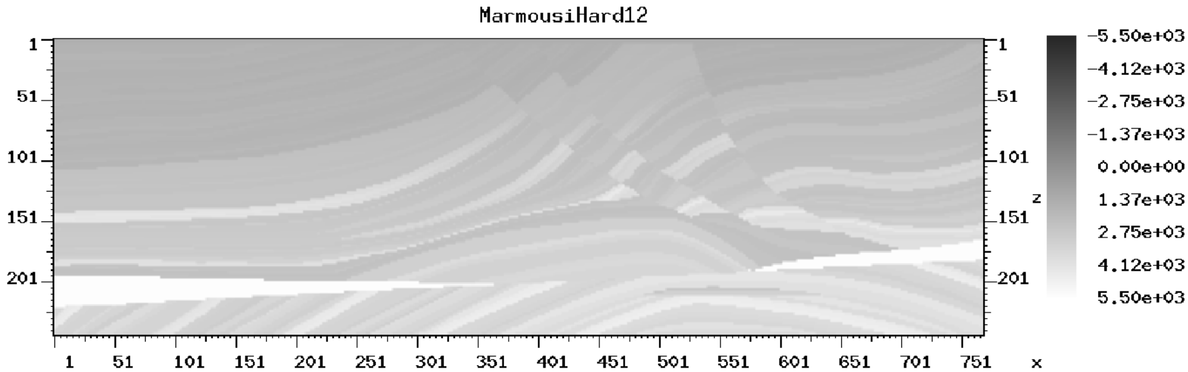

Figure 1: The exact Marmousi velocity model.

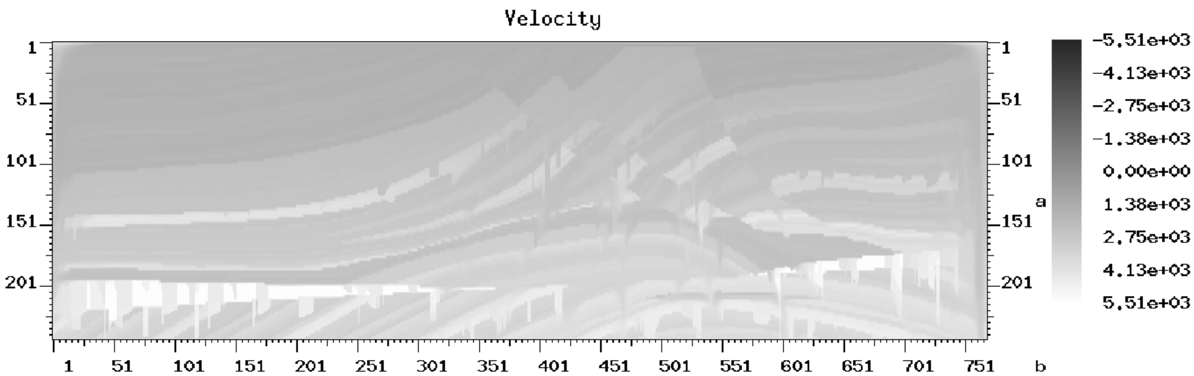

Figure 2: The reconstructed model using localized image-difference tomography. In general the model has been very well reconstructed except for some noticable streaking in the model.

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2007 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Adjoint wave-equation velocity analysis: 76th Annual International Meeting, SEG, Expanded Abstracts, 3345–3349.

Shen, P., C. Stolk, and W. Symes, 2003, Automatic velocity analysis by differential semblance optimization: 73rd Annual International Meeting, SEG, Expanded Abstracts, 893–896.

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