Seismic full-waveform inversion using truncated wavelet representations
Y. Lin, A. Abubakar, T. M. Habashy, Schlumberger-Doll Research

SUMMARY

We present a model compression scheme for solving acoustic full-waveform inversion problems using the Gauss-Newton minimization method. In this scheme, we represent P-wave velocity and mass-density distributions using wavelet basis functions. In order to reduce the number of unknown parameters in the inversion, we invert only dominant wavelet coefficients. Since the number of dominant coefficients is smaller than the number of unknown parameters in the spatial-domain, this compression scheme reduces the size of the Jacobian matrix. Hence, we reduce the memory storage of the Jacobian matrix as well as the computational time for calculating the Gauss-Newton update step. We use the Marmousi model to show that the model compression scheme can reduce the computational time and memory storage of the Gauss-Newton method without sacrificing the quality of its inversion results.

INTRODUCTION

The full-waveform inversion (FWI) method is a promising tool for obtaining high-resolution reconstructions of geologic structure from seismic data. This tool has become increasingly popular in recent years and many efforts have been made to improve its efficiency and robustness, see Pratt et al. (1998); Shipp and Singh (2002); Abubakar et al. (2003); Hustedt et al. (2004); Sirgue and Pratt (2004); Mulder and Plessix (2008); and Abubakar et al. (2009). A variety of inversion methods such as Gauss-Newton, limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS), non-linear conjugate gradient, and contrast source inversion methods are employed as the FWI optimization engine. In this work, we choose to work with the Gauss-Newton method because its convergence rate is faster than the non-linear conjugate gradient (NLCG) method (see Nocedal and Wright (2006) for generic applications, and see Pratt et al. (1998); Hu et al. (2011) for seismic applications). The NLCG-based methods are very popular because they are more memory efficient than the Gauss-Newton methods and do not require the inversion of the dense Hessian matrix to obtain the update step. However, their convergence rates may be slow. In the Gauss-Newton method, the Hessian matrix inversion is usually done by employing a conjugate-gradient least-squares (CGLS) iterative method, where the inverse of the Hessian matrix is not explicitly calculated; see for example the implementation in Hu et al. (2009).

There are several ways to reduce the computational complexity of the Gauss-Newton (or other non-linear inversion) method. The most popular approach is by reducing the number of sources and/or receivers in the inversion by using simultaneous sources schemes, see Krebs et al. (2009); Habashy et al. (2011); Ali et al. (2011). Another way is to calculate the Jacobian matrix on the fly and combine it with an Adaptive-Cross-Approximation scheme as discussed in Abubakar et al. (2011). In this work, we investigate another way to reduce the computational complexity of the Gauss-Newton method by reducing the number of unknown parameters. The idea has been investigated in the past for various electromagnetic applications, see for examples Oldenburg et al. (1993); Bucci and Isernia (1997). In these papers their main objective is to cope with ill-posedness of the problem. Hence, there is no additional regularization function is employed in the optimization process. Following this idea, we employ wavelet basis functions to represent the unknown model parameters: P-wave velocity and mass-density distributions. Since only part of the wavelet coefficients are dominant, the unknown model parameters can be approximated by only using dominant wavelet coefficients. Thus, in the inversion we invert a reduced number of unknown parameters. Hence, we reduce the memory storage of the Jacobian matrix as well as speed up the computational time for obtaining the Gauss-Newton step. We implement our acoustic FWI algorithm in the frequency-domain and we employ a finite-difference with Perfectly Matching Layer boundary conditions as our forward solver.

THEORY

The acoustic FWI problem can be considered as a discrete non-linear inverse problem described by the following equation:

\[
\mathbf{d} = s(\mathbf{m}),
\]

where \( \mathbf{d} \) is the vector of measured data containing pressure fields recorded at receiver positions and \( s \) is the simulated data obtained using the finite-difference frequency-domain forward solver.

The unknown model parameter vector \( \mathbf{m} \) contains both P-wave velocity \( V_p \) and mass-density \( \rho \) distributions. For the two-dimensional problem, usually \( V_p \) and \( \rho \) are represented by constant values in each pixel, which may also be viewed as using pulses as the basis functions. For an image with \( N_x \) by \( N_z \) pixels, \( 2N_xN_z \) pulse functions are needed to represent \( \mathbf{m} \). We note that it is also possible to use other type of basis functions, which may have a compression property.

According to some studies, wavelet is found to be very suitable for compressing an image. In addition, based on our own study using Haar and Daubechies 4 wavelets for inversions, we observe that Daubechies 4 wavelet generally provides better performance with only slightly more computational efforts than Haar wavelet. For a different application Daubechies wavelet has also been found to outperform Haar wavelet, see Mahmoud et al. (2007). Therefore, in this paper, we only focus on the use of Daubechies 4 wavelet to perform the model compression in FWI.

The one-level Daubechies 4 wavelet transform for a one-dimensional image \( s \) can be viewed as passing the image through a
low-pass filter as follows:
\[ \psi_t = h_0 s_{2t} + h_1 s_{2t-1} + h_2 s_{2t-2} + h_3 s_{2t-3} , \] (2)
where
\[ h_0 = 1 + \frac{\sqrt{3}}{4\sqrt{2}} , \quad h_1 = 3 \sqrt{\frac{3}{4\sqrt{2}}} , \quad h_2 = \frac{3 - \sqrt{3}}{4\sqrt{2}} , \quad h_3 = \frac{1 - \sqrt{3}}{4\sqrt{2}} \] (3)
and a high-pass filter as given by
\[ \phi_t = g_0 s_{2t} + g_1 s_{2t-1} + g_2 s_{2t-2} + g_3 s_{2t-3} , \] (4)
where
\[ g_0 = \frac{1 - \sqrt{3}}{4\sqrt{2}} , \quad g_1 = -\frac{3 - \sqrt{3}}{4\sqrt{2}} , \quad g_2 = \frac{3 + \sqrt{3}}{4\sqrt{2}} , \quad g_3 = -\frac{1 + \sqrt{3}}{4\sqrt{2}} \] (5)
The 1D one-level Daubechies 4 wavelet coefficients can be obtained by combining equations (2) and (4): \( \tilde{s} = \mathcal{F}(s) = [\psi, \phi] \), where \( \mathcal{F} \) is the wavelet transform operator. The higher levels of the wavelet transform coefficients can be obtained by repeating the process on the low-pass transformed representation of the previous level. Higher dimensional transforms are performed by several orthogonal 1D transforms. We employ the periodic extension for elements at the boundaries. The computational complexity of the Daubechies 4 wavelet transform is \( O(N) \), where \( N \) is the length of the image.

In the inversion, the image can be approximately reconstructed by only using the dominant wavelet coefficients through an inverse wavelet transform as follows:
\[ s \approx \mathcal{F}^{-1}(\tilde{s}_c) , \] (6)
where \( \tilde{s}_c \) is the compressed wavelet model
\[ \tilde{s}_c = \begin{cases} \tilde{s}_i , & \text{dominant coefficients} , \\ 0 , & \text{otherwise} . \end{cases} \]
The inversion is carried out on a vector of the compressed model vector \( \tilde{\mathbf{m}}^c \), which has a size of \( 2N_c \times N_c \) instead of the spatial-domain model vector \( \mathbf{m} \) with size of \( 2N_x \times N_x \). We pose the inversion problem as a minimization problem with a multiplicative cost function as designed by van den Berg et al. (1999). Hence, at the \( n \)th iteration we reconstruct the compressed model \( \tilde{\mathbf{m}}^c \) that minimizes
\[ \Phi_n(\tilde{\mathbf{m}}^c) = \Phi^d(\tilde{\mathbf{m}}^c) \times \Phi_n^m(\tilde{\mathbf{m}}^c) , \] (8)
where \( \Phi^d \) is a measure of the data misfit:
\[ \Phi^d(\tilde{\mathbf{m}}^c) = \frac{1}{2} \left\| W_d \left[ d - s \left( \mathcal{F}^{-1}[\tilde{\mathbf{m}}^c] \right) \right] \right\|^2 , \] (9)
in which \( W_d \) is a real-valued diagonal data weighting matrix.

The regularization function \( \Phi_n^m \) is a measure of the variation of the model parameters in the spatial domain. Note that we choose to regularize the model parameters in the spatial domain instead of regularizing in the transformed domain since we would like to impose a physical smoothing constraint on P-wave velocity and mass-density distributions. This approach also allows us to readily include other physical constraints such as well-log constraints (when they are available) and the depth weighting. We also note that the model compression scheme itself is a regularization approach. However, we think that an additional regularization is necessary since we are still dealing with an ill-posed problem even when we have a very limited number of unknown parameters. In order to put emphasis on the smoothness of the unknown parameter distribution we employ the following regularization function:
\[ \Phi_n^m(\tilde{\mathbf{m}}^c) = \int_D b_n^2 \left\{ \left\| \mathcal{F}^{-1}[\tilde{\mathbf{m}}^c] \right\|^2 + \delta_n^2 \right\} dx dz , \] (10)
where \( \mathcal{V} \) is the spatial differentiation operator and
\[ b_n^2(x,z) = \frac{1}{\int_D \left\| \mathcal{F}^{-1}[\tilde{\mathbf{m}}^c] \right\|^2 + \delta_n^2 dx dz} . \] (11)
The function \( \delta_n \) is a non-zero constant which is chosen to be equal to \( \delta_n^2 = \Phi^d(\tilde{\mathbf{m}}^c)/\Delta x \Delta z \), where \( \Delta x \) and \( \Delta z \) are the widths of the discretization cell in the \( x \) and \( z \) directions.

To solve equation (8) we employ a Gauss-Newton minimization framework described in Habashy and Abubakar (2004). At the \( n \)th iteration we obtain a set of linear equations for the search vector \( \tilde{\mathbf{p}}_n \):
\[ \tilde{\mathbf{H}}_n \tilde{\mathbf{p}}_n = -\tilde{\mathbf{g}}_n , \] (12)
where \( \tilde{\mathbf{H}}_n \) is the Hessian matrix with respect to \( \tilde{\mathbf{m}}^c \), and can be approximated as follows (see Abubakar et al. (2008)):
\[ \tilde{\mathbf{H}}_n \approx \bar{\mathbf{J}}_n^H \mathbf{W}_d^2 \mathbf{W}_d \bar{\mathbf{J}}_n + \Phi^d(\tilde{\mathbf{m}}^c) \mathbf{L}_n , \] (13)
where the superscript \( ^H \) denotes Hermitian. The matrix \( \bar{\mathbf{J}}_n \) is the Jacobian matrix and the matrix \( \mathbf{L}_n \) denotes the second derivative of \( \Phi_n^m \) with respect to \( \tilde{\mathbf{m}}^c \). The gradient vector \( \tilde{\mathbf{g}}_n \) is given by
\[ \tilde{\mathbf{g}}_n = \bar{\mathbf{J}}_n^H \mathbf{W}_d^2 \mathbf{W}_d ^2 \Delta - s \left( \mathcal{F}^{-1}[\tilde{\mathbf{m}}^c] \right) + \Phi^d(\tilde{\mathbf{m}}^c) \mathbf{L}_n \tilde{\mathbf{m}}^c . \] (14)

After solving equation (12) using the CGLS method to obtain the step vector \( \tilde{\mathbf{p}}_n \), we can use an inverse transform to calculate the step vector in the spatial domain as follows:
\[ \tilde{\mathbf{m}}_{n+1} = \mathcal{F}^{-1}[\tilde{\mathbf{p}}_n] . \] (15)

Then, the unknown model parameter in the spatial domain \( \mathbf{m} \) is updated using the following formula:
\[ \mathbf{m}_{n+1} = \mathbf{m}_n + \nu_n \tilde{\mathbf{m}}_{n+1} , \] (16)
where \( \nu_n \) is the step length, which is obtained through a line-search procedure described in Habashy and Abubakar (2004). The Gauss-Newton iterative process is terminated when one of the four error criteria listed in Habashy and Abubakar (2004) is satisfied.
NUMERICAL TEST

To illustrate the performance of the model compression scheme, we employ the Marmousi-2 model (Martin et al. (2002)). The P-wave velocity $V_p$ distribution of this model is shown in the Figure 1. The size of the model is 9.2 km by 2.96 km. In the computation, the model is discretized into 68,080 pixels with cell size of 20 m by 20 m. In the inversion we use 191 sources and 192 receivers. All sources and receivers are located at depth of 130 m. Sources are uniformly distributed from $x = 36$ m to $x = 9, 156$ m and receivers are uniformly distributed from $x = 12$ m to $x = 9, 180$ m. For all inversion tests we add 3% random noise. The inversions are performed on a cluster with 2.33 GHz processors and 24 cores were used for all calculations. Note that in inversions we also invert the mass-density distribution, however because of the space limitation of this proceeding paper we do not show inverted mass-density models.

We invert the multi-frequency data at 1.5, 3, 6, 10, and 15 Hz using a sequential approach (we invert single-frequency data one at a time from the lowest to the highest frequency). For initial models, we use linearly increasing velocities and density models. The inverted model of 1.5 Hz data using the Gauss-Newton method with the spatial domain unknown representation is given in Figure 2(a). The inverted model using the inversion algorithm with two-level Daubechies 4 wavelet is given in Figure 2(b). In this case, we use only low-low (LL) representation of wavelet coefficients. Hence, we invert only 6.25% of the total wavelet coefficients. We observe that there are only small differences in the two inverted models. The same conclusion applies for 3 Hz data inversions.

Next, we employ results of 3 Hz data inversions as initial models for 6 Hz data inversions. The inversion results of the Gauss-Newton method using the spatial-domain and LL coefficients of two-level Daubechies 4 wavelet representations are given in figures 3(a) and (b). We notice obvious differences and some artifacts appear in the inverted model obtained using LL coefficients of the two-level Daubechies 4. In order to improve the inversion results, we rerun the 6 Hz data inversion using LL coefficients of the one-level Daubechies 4 wavelet, which has 25% of the total wavelet coefficients (thus, we have less compression factor). The inverted model is shown in Figure 3(c), which has a similar quality as the spatial domain inversion result as shown in Figure 3(a). At this inversion frequency, we need more coefficients to capture more resolutions contained in the data. If we do not provide this flexibility, then the inversion will start to create some artifacts to compensate the additional resolution information in the data.

The models after 10 Hz data inversions are shown in Figures 4. Figure 4(b) shows the inverted model using LL coefficients of the one-level Daubechies 4 wavelet. We observe that differences are mainly appeared in the area below the depth of 2 km. This can be understood since the data are less sensitive to that area.

Finally, we invert 16 Hz data using inverted models given in...
Figure 4: The inverted $V_P$ models after 10 Hz data inversion using: (a) spatial-domain and (b) LL one-level Daubechies 4 unknown representation.

Figure 4 as initial models. The results from inversion algorithms with and without the model compression scheme are shown in Figures 5(a) and (b). Similar as the case for 6 Hz data inversions, we observe some artifacts in the inverted model obtained using the LL coefficients of one-level Daubechies 4 wavelet. Then we re-invert the 16 Hz data by employing more wavelet coefficients. The result is given in Figure 5(c), which has less oscillations than the one shown in Figure 5(b). The distribution of the wavelet coefficients is shown in Figure 5(d), which was obtained by using the well-known wavelet propagation approach.

The computational costs and final data misfits for various frequencies are shown in Table 1. We observe that inversions using the model compression scheme only obtained slightly higher data misfits than spatial-domain inversions. However, we obtained reasonable speedup factors, especially for low-frequency data inversions (more than an order of magnitude).

<table>
<thead>
<tr>
<th>Hz</th>
<th>Iteration</th>
<th>Data misfit</th>
<th>CPU time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>8</td>
<td>2.81%</td>
<td>3,077,676</td>
</tr>
<tr>
<td>1.5 D4-L2</td>
<td>8</td>
<td>2.85%</td>
<td>6,667</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>2.73%</td>
<td>1,911,988</td>
</tr>
<tr>
<td>3 D4-L2</td>
<td>8</td>
<td>2.84%</td>
<td>6,772</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>2.60%</td>
<td>263,833</td>
</tr>
<tr>
<td>6 D4-L1</td>
<td>5</td>
<td>2.73%</td>
<td>37,812</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>2.51%</td>
<td>319,583</td>
</tr>
<tr>
<td>10 D4-L1</td>
<td>6</td>
<td>2.68%</td>
<td>112,041</td>
</tr>
<tr>
<td>16</td>
<td>7</td>
<td>2.07%</td>
<td>1,034,147</td>
</tr>
<tr>
<td>16 D4-L1 Plus</td>
<td>6</td>
<td>2.50%</td>
<td>229,379</td>
</tr>
</tbody>
</table>

Table 1: Computational costs and data misfits comparison for Marmousi inversions.

Figure 5: The inverted $V_P$ models after 16 Hz data inversion using: (a) spatial-domain, (b) LL one-level Daubechies 4, c) LL one-level Daubechies 4 with additional coefficients, and d) locations of wavelet coefficients.

CONCLUSION

We investigated the use of a model compression scheme using wavelets for reducing the computational costs of the acoustic full-waveform inversions. This scheme can significantly reduce the number of unknown parameters, hence reduces the computational cost and memory requirement of the Gauss-Newton methods. We also observe that we can compress the model using higher compression factor at low-frequency data inversion than at high-frequency data inversion. We remark that this model compression scheme can straightforwardly be extended to 3D problems.

ACKNOWLEDGEMENTS

The authors would like to thank Maokun Li from Schlumberger for numerous fruitful discussions on wavelets.
EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2012 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

Abubakar, A., T. M. Habashy, V. Druskin, L. Knizhnerman, and D. Alumbaugh, 2008, 2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements: Geophysics, 73, no. 4, F165–F177.


