Computers & Geosciences 93 (2016) 45–54

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

Reducing disk storage of full-3D seismic waveform tomography (F3DT) through lossy online compression

Peter Lindstrom [a], Po Chen [b,*], En-Jui Lee [c]

[a] Lawrence Livermore National Laboratory, USA
[b] Department of Geology and Geophysics, University of Wyoming, USA
[c] Department of Earth Sciences, National Cheng Kung University, Taiwan

A R T I C L E I N F O

Article history:
Received 30 October 2015
Received in revised form
16 April 2016
Accepted 18 April 2016
Available online 5 May 2016

Keywords:
Seismic tomography
Full-3D tomography
Scattering-integral method
Waveform tomography
Full-waveform
Full-wave
Compression
Lossy compression
Online compression

A B S T R A C T

Full-3D seismic waveform tomography (F3DT) is the latest seismic tomography technique that can assimilate broadband, multi-component seismic waveform observations into high-resolution 3D subsurface seismic structure models. The main drawback in the current F3DT implementation, in particular the scattering-integral implementation (F3DT-SI), is the high disk storage cost and the associated I/O overhead of archiving the 4D space-time wavefields of the receiver- or source-side strain tensors. The strain tensor fields are needed for computing the data sensitivity kernels, which are used for constructing the Jacobian matrix in the Gauss–Newton optimization algorithm. In this study, we have successfully integrated a lossy compression algorithm into our F3DT-SI workflow to significantly reduce the disk space for storing the strain tensor fields. The compressor supports a user-specified tolerance for bounding the error, and can be integrated into our finite-difference wave-propagation simulation code used for computing the strain fields. The decompressor can be integrated into the kernel calculation code that reads the strain fields from the disk and compute the data sensitivity kernels. During the wave-propagation simulations, we compress the strain fields before writing them to the disk. To compute the data sensitivity kernels, we read the compressed strain fields from the disk and decompress them before using them in kernel calculations. Experiments using a realistic dataset in our California statewide F3DT project have shown that we can reduce the strain-field disk storage by at least an order of magnitude with acceptable loss, and also improve the overall I/O performance of the entire F3DT-SI workflow significantly. The integration of the lossy online compressor may potentially open up the possibilities of the wide adoption of F3DT-SI in routine seismic tomography practices in the near future.

1. Introduction

Seismic tomography has been one of the most effective means for imaging the internal structure of the Earth in the past few decades (e.g., Anderson and Dziewonski, 1984; Nolet, 1987a, 1987b; Iyer, 1993; Nolet, 2008, 2012). The techniques used in seismic tomography have been constantly improving. Recent advances in computing technology has drastically reduced the computational cost for solving the 3D (visco)elastic seismic wave equation, which has enabled full-3D tomography (F3DT) (e.g., Chen et al., 2007a, 2007b; Fichtner et al., 2009; Tape et al., 2010; Lee et al., 2014a, 2014b; Chen and Lee, 2015). In F3DT, the starting seismic structural model can be fully three-dimensional and the Fréchet (sensitivity) kernels are computed by numerically solving the inhomogeneous equations of motion for a heterogeneous, (an) elastic solid. It accounts for the nonlinearity of waveform inversions through iterated cycles of wave-propagation simulations, misfit measurements, sensitivity kernel calculations and inversions.

There are two complementary F3DT implementations: the adjoint-wave-field method (F3DT-AW), which constructs the gradient of the objective function using the adjoint method and solves the optimization problem using gradient-based algorithms (e.g., Fichtner et al., 2009; Tape et al., 2010), and the scattering-integral method (F3DT-SI), which sets up the Jacobian of the objective function by calculating and storing the data sensitivity (Fréchet) kernel for each misfit measurement and solves the optimization problem using the Gauss–Newton algorithm (e.g., Chen et al., 2007a; Lee et al., 2014a). These two types of implementations are based on the same physics, but their computational requirements can be highly different (Chen et al., 2007b). For tomography problems involving a large number of seismic sources F3DT-SI may significantly reduce the total amount of computing time at the expense of substantially higher disk storage cost. For

* Corresponding author.
E-mail address: pchen@uwyo.edu (P. Chen).

http://dx.doi.org/10.1016/j.cageo.2016.04.009

the F3DT inversion in Southern California (Lee et al., 2014a), the peak disk storage for F3DT-SI was about 39 TB, while the peak disk storage for F3DT-AW was only about 200 GB, a nearly 200 times difference. The high disk storage cost of F3DT-SI is becoming the main obstacle to the wide adoption of F3DT-SI in routine seismic tomography, especially on small to medium-sized shared computer clusters without large amounts of high-speed disk storage.

In this paper, we describe a potential solution for significantly reducing the disk storage of F3DT-SI through lossy but error-bounded online compression. Our compression algorithm, named *zfp*, provides high compression ratios with minimal CPU overhead and can work inside the wave-propagation simulation code and the sensitivity kernel calculation code in a streaming setting during the I/O stage (Lindstrom, 2014). Whereas *zfp* was originally designed for fixed-rate compression in order to support random access, we use its fixed-accuracy (variable-rate) mode in order to limit compression-induced errors. Although this mode sacrifices the ability to perform constant-time random access, we require only sequential reads and writes of entire strain fields. Moreover, relaxing the fixed-rate constraint can significantly improve the quality per bit of compressed storage. Preliminary experiments using realistic simulations in our California statewide F3DT-SI inversion show highly promising results. On average, the disk space for storing the 4D strain tensor fields can be reduced by at least an order of magnitude with much improved I/O performance. The kernels computed from the compressed strain fields have negligible differences from those computed using the raw strain fields. By integrating *zfp*, we expect to make F3DT-SI much more affordable on small clusters.

2. Disk storage cost of F3DT-SI

F3DT is often implemented using gradient- or Hessian-based iterative optimization algorithms. The discretized earth structural model $\mathbf{m}$ is iteratively updated through a finite series of perturbations,

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \Delta\mathbf{m}_k, \quad k = 0, 1, 2, \ldots, K \tag{1}$$

where k is the iteration index. The perturbation for the kth iteration, $\Delta\mathbf{m}_k$, can be obtained by minimizing an objective function,

$$\chi^2(\mathbf{m}, \mathbf{m}_k) = \mathbf{d}^{\mathrm{T}}(\mathbf{m}, \mathbf{m}_k)\mathbf{C}_d^{-1}\mathbf{d}(\mathbf{m}, \mathbf{m}_k) + (\mathbf{m} - \mathbf{m}_k)^{\mathrm{T}}\mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_k) \tag{2}$$

where $\mathbf{m}$ is the "target" structural model, $\mathbf{d}(\mathbf{m}, \mathbf{m}_k)$ is a column-vector composed of misfit measurements that quantify the discrepancies between the ith-component observed seismogram generated by the sth seismic source and recorded at the rth receiver, $\bar{u}_i^s(\mathbf{x}_r, t)$, and the corresponding synthetic seismogram $u_i^s(\mathbf{x}_r, t)$ computed using the latest structural model $\mathbf{m}_k$, $\mathbf{C}_d$ and $\mathbf{C}_m$ are respectively the data and model covariance matrices.

In F3DT-SI, the objective function in Eq. (2) is minimized using the Gauss–Newton algorithm, which requires the solution of the Gauss–Newton normal equation

$$\begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{A}_k \\ \mathbf{C}_m^{-1/2} \end{bmatrix} \Delta\mathbf{m}_k = \begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{d}_k \\ \mathbf{0} \end{bmatrix} \tag{3}$$

where $\mathbf{A}_k = \partial\mathbf{d}_k/\partial\mathbf{m}_k$ is the Jacobian matrix for the kth iteration. In F3DT-SI, the Jacobian matrix is explicitly constructed and Eq. (3) is solved using the scalable parallel LSQR algorithm (Lee et al., 2013).

Each row of the Jacobian is a discretized data sensitivity kernel, which can be computed using the 4D strain fields from the source and those from the receiver. Equations for constructing the data sensitivity kernels using the receiver-side strain Green's tensors (RSGTs) have been given in (e.g., Zhao et al., 2005, 2006; Chen et al., 2007a, 2007b; Chen and Lee, 2015). The calculation involves temporal convolution between the strain field from the source and the RSGT for the corresponding receiver, which can be computed by placing a point impulsive source at the receiver location (Zhao et al., 2006). To construct the kernels for all misfit measurements, we need to store either the RSGTs or the source-side strain fields. When seismic sources outnumber receivers, it is more economical to store the RSGTs.

In practice, the kernels are usually smoother than the strain fields and we often regularize the inverted model perturbation $\Delta\mathbf{m}_k$ through smoothness damping. Therefore we can sample the kernels on a mesh that is sparser than the mesh used for the wave-propagation simulations. The accuracy of the temporal convolution is usually sufficient if we have 10 time samples per dominant period. Because of these considerations, the disk space for storing strain fields can be reduced significantly through decimation in space and time. But even after decimation, the disk storage for all strain fields used in a realistic inversion can still be significant.

3. Lossy online compression of strain fields

Previous studies on seismic data compression mainly focused on the compression of observed active-source seismic data in the space-time domain in an off-line setting (e.g., Wood, 1974; Jonsson and Spanias, 1990; Mandyam et al., 1996; Villasenor et al., 1996; Wang and Wu, 2000; Averbuch et al., 2001). Blind application of traditional lossless compression algorithms on observed active-source seismic wavefields can only provide low compression ratios of around 2 (e.g., Villasenor et al., 1996), while applications of lossy compression algorithms were able to achieve compression ratios ranging from ~20 to over 100 with acceptable losses of useful seismic information (e.g., Villasenor et al., 1996; Wang and Wu, 2000; Averbuch et al., 2001). For the 4D synthetic RSGTs from the California statewide inversion considered in this study, perfectly lossless compression using a state-of-the-art floating-point lossless compressor *FPZIP* (Lindstrom and Isenburg, 2006) provided a compression ratio of merely 1.55. For F3DT purposes, a lossless compression of synthetic strain fields is both unnecessary and inefficient. A more desirable compression scheme is a lossy algorithm that can work in an online I/O setting from within the wave-propagation simulation code with minimal CPU overhead and can achieve significant compression ratios without introducing significant artifacts into the data sensitivity kernels.

Lossy compression of observed active-source seismic data has been extensively studied in the past two decades (e.g., Lervik et al., 1996; Villasenor et al., 1996; Vassiliou and Wickerhouser, 1997; Wang and Wu, 2000; Averbuch et al., 2001; Al-Moohimeed, 2004; Wang et al., 2004; Aparna and David, 2006; Wu et al., 2006; Xie and Qin, 2009; Aqrawi and Elster, 2011; Zheng and Liu, 2012; Fajardo et al., 2015). The majority of the compression algorithms usually follow a 3-stage process: de-correlating transformation, quantization and coding. In the transformation stage, suitable basis functions can lead to a much sparser representation of the original data in the transformed domain. Wavelets, wavelet packets, (adaptive) local trigonometric functions and various combinations of the above have been widely used in previous studies (e.g., Villasenor et al., 1996; Al-Moohimeed, 2004; Wang and Wu, 2000; Wu et al., 2006; Zheng and Liu, 2012). The floating-point transform coefficients are then mapped to a set of integers in the quantization stage. The majority of previous studies adopted uniform quantization schemes (e.g., Lervik et al., 1996; Al-Moohimeed, 2004; Aparna and David, 2006; Wu et al., 2006; Fajardo et al., 2015). In general, as the number of quantization bits decreases, the compression ratio, as well as the loss of useful