Journal of Applied Geophysics 144 (2017) 18–27

Contents lists available at ScienceDirect

Journal of Applied Geophysics

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

Multi-GPU parallel algorithm design and analysis for improved inversion of probability tomography with gravity gradiometry data

Zhenlong Hou [a,*], Danian Huang [b]

[a] Key Laboratory of Ministry of Education on Safe Mining of Deep Metal Mines, School of Resources and Civil Engineering, Northeastern University, Shenyang 110819, PR China
[b] College of Geo-exploration Science and Technology, Jilin University, Changchun 130026, PR China

ARTICLE INFO

Article history:
Received 14 November 2016
Received in revised form 30 April 2017
Accepted 22 June 2017
Available online 3 July 2017

Keywords:
Gravity gradiometry data
Inversion of probability tomography
Parallel computing
OpenMP
CUDA
GPU

ABSTRACT

In this paper, we make a study on the inversion of probability tomography (IPT) with gravity gradiometry data at first. The space resolution of the results is improved by multi-tensor joint inversion, depth weighting matrix and the other methods. Aiming at solving the problems brought by the big data in the exploration, we present the parallel algorithm and the performance analysis combining Compute Unified Device Architecture (CUDA) with Open Multi-Processing (OpenMP) based on Graphics Processing Unit (GPU) accelerating. In the test of the synthetic model and real data from Vinton Dome, we get the improved results. It is also proved that the improved inversion algorithm is effective and feasible. The performance of parallel algorithm we designed is better than the other ones with CUDA. The maximum speedup could be more than 200. In the performance analysis, multi-GPU speedup and multi-GPU efficiency are applied to analyze the scalability of the multi-GPU programs. The designed parallel algorithm is demonstrated to be able to process larger scale of data and the new analysis method is practical.

1. Introduction

The method of probability tomography is first proposed by Patella and applied for the self-potential data (Patella, 1997a, 1997b). Then it is extended by many scholars and used for geo-electric, geo-electromagnetic and resistivity anomaly data (Mauriello et al., 1998; Mauriello and Patella, 1999a, 1999b), also for gravity, magnetic, and their vertical gradient data (Mauriello and Patella, 2001a, 2001b, 2005, 2008; Guo et al., 2011a, 2011b). Meng et al. (2012) and Liu et al. (2014) implemented IPT, a kind of fast inversion, with magnetic and gravity anomaly, and the good results are obtained.

However, the amount of data increases significantly in geophysical explorations at present. The increasing amount of data brings some computing problems for data processing and interpretation. Even the fast algorithm, such as the IPT, could not finish the calculations in short time. The large scale of data could also consume too much memory. That forces geo-explorers to combine new techniques to process big geophysical data. Some scholars have considered the parallel computing methods and implemented for the forward modeling (Commer, 2011; Moorkamp et al., 2011; Roberts et al., 2016). Computed Unified Device Architecture (CUDA) is a parallel computing platform and programming model presented by Nvidia, specially designed for GPUs of Nvidia. CUDA has been applied for geophysical explorations in recent years. Moorkamp et al. (2010) presented an effective parallel approach with CUDA for forward modeling of scalar and tensor gravity. Chen et al. (2012a, 2012b) used CUDA for the forwarding, inversion and probability tomography with gravity and gradiometry data. Liu et al. (2012) implemented IPT with three dimensional magnetic data. The parallel algorithm of reweighted focusing inversion is also implemented with CUDA by Hou et al. (2015). When the algorithm is based on the data parallel model, CUDA could improve its efficiency significantly. It is often used with single GPU, but the memory of single GPU is not enough to deal with large scale of data. The research and application of parallel inversion algorithms with multiple GPUs are very important for successfully developing of fast processing bigger data in the future. Martin et al. (2013) have implemented 3-D gravity inversion with GPU cluster, a significant acceleration was achieved. Čuma and Zhdanov (2014) have made a research of massively parallel inversion on GPUs. The hybrid parallel algorithm with multiple GPUs for the inversion, aiming at the gravity gradiometry data, is designed by Hou et al. (2016a), getting the expected accelerating effect.

According to the research of Hou (2016), the improved IPT with multiple gravity gradiometry tensors is described in this paper. It is proved by the synthetic model and real data tests that the improved algorithm has a higher space resolution than the original one. CUDA and Open Multi-Processing (OpenMP) are combined to be a hybrid parallel algorithm based on OpenMP and CUDA (OpenMP-CUDA algorithm), which could schedule multiple GPUs for inversion more efficiently. In the performance analysis, it is demonstrated that the parallel algorithm has a high speedup. Compared with the algorithm with single GPU, with multiple GPUs only using CUDA (CUDA algorithm) or

* Corresponding author.
E-mail addresses: houzlatjlu@163.com (Z. Hou), dnhuang@jlu.edu.cn (D. Huang).

http://dx.doi.org/10.1016/j.jappgeo.2017.06.009
0926-9851/

combined with Message Passing Interface (MPI) (MPI-CUDA algorithm), the performance of OpenMP-CUDA algorithm is better than them in varying degrees. Some metrics are used to judge the scalability of the multi-GPU program. It is confirmed that the parallel algorithm in this paper has the ability of processing big data.

2. Theory of improved IPT and data test

2.1. Inversion theory

We set the study area to the Cartesian coordinate system, whose x-y plane is horizontal and z axis points vertically downwards. According to Chen et al. (2012b), the underground sub-space would be divided into a series of small uniform prisms, and the gravity gradiometry $g_{\alpha\beta,j}$, which is caused by the prism j, at the surface observation i (x_i, y_i, z_i) is

$$g_{\alpha\beta,j}(x_i,y_i,z_i) = \gamma\Delta\rho_j v_j B_{\alpha\beta,j}(x_i,y_i,z_i), \tag{1}$$

$$\alpha,\beta = x,y,z, \tag{2}$$

where γ is the universal gravitational constant; $\Delta\rho_j$ and v_j are residual density and volume of the prism j, respectively; $B_{\alpha\beta,j}$ is the geometrical function of gravity gradiometry at the observation i caused by the prism j, we can compute it by the equations below:

$$B_{xx,j}(x_i,y_i,z_i) = \frac{2(x_i-x_j)^2-\left(y_i-y_j\right)^2-(z_i-z_j)^2}{\left[(x_i-x_j)^2+\left(y_i-y_j\right)^2+(z_i-z_j)^2\right]^{5/2}}, \tag{3}$$

$$B_{xy,j}(x_i,y_i,z_i) = B_{yx,j}(x_i,y_i,z_i) = \frac{3(x_i-x_j)\left(y_i-y_j\right)}{\left[(x_i-x_j)^2+\left(y_i-y_j\right)^2+(z_i-z_j)^2\right]^{5/2}}, \tag{4}$$

$$B_{xz,j}(x_i,y_i,z_i) = B_{zx,j}(x_i,y_i,z_i) = \frac{3(x_i-x_j)(z_i-z_j)}{\left[(x_i-x_j)^2+\left(y_i-y_j\right)^2+(z_i-z_j)^2\right]^{5/2}}, \tag{5}$$

$$B_{yy,j}(x_i,y_i,z_i) = \frac{2\left(y_i-y_j\right)^2-(x_i-x_j)^2-(z_i-z_j)^2}{\left[(x_i-x_j)^2+\left(y_i-y_j\right)^2+(z_i-z_j)^2\right]^{5/2}}, \tag{6}$$

$$B_{yz,j}(x_i,y_i,z_i) = B_{zy,j}(x_i,y_i,z_i) = \frac{3\left(y_i-y_j\right)(z_i-z_j)}{\left[(x_i-x_j)^2+\left(y_i-y_j\right)^2+(z_i-z_j)^2\right]^{5/2}}, \tag{7}$$

$$B_{zz,j}(x_i,y_i,z_i) = \frac{2(z_i-z_j)^2-(x_i-x_j)^2-\left(y_i-y_j\right)^2}{\left[(x_i-x_j)^2+\left(y_i-y_j\right)^2+(z_i-z_j)^2\right]^{5/2}}. \tag{8}$$

Then the correlation coefficient of gravity gradiometry is given as (Chen et al., 2012b):

$$C_{\alpha\beta,j} = \frac{\sum_{i=1}^{N} g_{\alpha\beta}(x_i,y_i,z_i)B_{\alpha\beta,j}(x_i,y_i,z_i)}{\sqrt{\sum_{i=1}^{N} g_{\alpha\beta}^2(x_i,y_i,z_i)\sum_{i=1}^{N} B_{\alpha\beta,j}^2(x_i,y_i,z_i)}}, \tag{9}$$

where N is the total number of the observations. It is concluded that the range of correlation coefficient is $[-1,1]$ according to Cauchy-Schwarz inequality. When correlation coefficient is positive, it indicates density excess; if negative, that means density defect. The joint correlation coefficient of n kinds of tensors is (Hou et al., 2016a):

$$C_{joint,j} = C_{1,j} \times C_{2,j} \times \cdots \times C_{n,j}, \tag{10}$$

where $C_{1,j}, C_{2,j}, \ldots, C_{n,j}$ means correlation coefficients of different gravity gradiometry data. The correlation coefficient could be used for probability tomography; the results show the presence of anomalous bodies (Liu et al., 2014).

According to the research of Hou (2016), the steps of the joint IPT with multiple gravity gradiometry tensors are:

(1) Compute the correlation coefficients of the underground prisms from the observations;
(2) Compute $\Delta\rho$: according to Liu et al. (2014), when using two dimensional data, if $\Delta\rho$ is given properly, the fitting data $\mathbf{d}_{fit}$ would be close to the observations; set the initial densities are zero in the inversion, the fitting data would be:

$$\mathbf{d}_{fit} = \mathbf{G}\boldsymbol{\rho} = \mathbf{GC}\Delta\rho, \tag{11}$$

where $\mathbf{G}$ and $\mathbf{C}$ represent the forwarding matrix and correlation coefficient vector, respectively. The value of $\Delta\rho$ could be computed according to Eq. (11); by trial and error, it could be computed more reasonably as:

$$\Delta\rho = \mathbf{d}_{\max}/\max[\mathbf{G}\times\mathbf{C}], \tag{12}$$

where $\mathbf{d}_{\max}$ and $\max[\mathbf{G}\times\mathbf{C}]$ are the maximum of the observed vector $\mathbf{d}$ and the result vector of $\mathbf{G}\times\mathbf{C}$, respectively; as $\Delta\rho$ could be computed by many tensors, it may have different values, and the minimum is chosen so that the density would not increase too fast in the iterations;
(3) Convert $C_{joint,j}$ to density by multiplying $\Delta\rho$, and use the results for the forwarding,

$$\rho_{j,iter+1} = \rho_{j,iter} + C_{joint,j} \times \Delta\rho, \tag{13}$$

where *iter* is the count of iteration, $\rho_{j,iter}$ is the residual density of the prism j in the *iter* iteration; a relationship has been proved (Liu et al., 2014) that when correlation coefficients are multiplied by a certain $\Delta\rho$, forward modeling data has the same trend as the real data; there must be a $\Delta\rho$ minimize the difference between forward modeling and real data; from that, Eq. (13) could be used to compute density;
(4) Subtract the fitting data in step (3) from the observations, and use the difference to generate new correlation coefficients, then compute new densities;
(5) Repeat step (3) and (4) until the iterations stop and final density results would be obtained.

To make a further improvement of the vertical space resolution, we introduce the depth weighting matrix $\mathbf{W}$, a diagonal matrix proposed by Li and Oldenburg (2003). Its form could be:

$$W_{j,j} = \sqrt[4]{\sum_{i}^{N} G_{ij}^2}, \tag{14}$$

where $W_{j,j}$ is the diagonal element of the matrix; G_{ij} is the element of the forwarding matrix. The weighted density equation is:

$$\rho_{j,iter+1} = \rho_{j,iter} + C_{joint,j} \times W_{j,j} \times \Delta\rho. \tag{15}$$