Accelerating Nonlinear Inversion Algorithms on GPU platform for Electromagnetic Data

Maokun Li1, Xue Yang Wang1, and Aria Abubakar2
1Department of electronic engineering, Tsinghua University, Beijing, China
2Schlumberger, Houston, Texas, USA

Abstract – **Nonlinear inversion algorithms have many applications in geophysical explorations, biomedical imaging, nondestructive testing and etc. One of the bottlenecks in these algorithms is the computational efficiency. In this abstract, we investigate accelerating the contrast source inversion algorithm using graphic process unit. By taking advantage of the massively parallel computing architecture of GPU, we can significantly reduce the computational time of inversion. Numerical examples show that we can reduce the computational time from hours to only a few minutes.**

Index Terms — **Microwave imaging, nonlinear inversion, contrast source inversion, graphic process unit (GPU).**

1. Introduction

Many applications, such as geophysical prospecting, biomedical imaging, nondestructive testing, and etc., require to solve inverse problems of electromagnetic data in order to reconstruct the electrical properties of the domain of investigation. It is usually formulated as an optimization problem of the electrical properties to minimize the difference between the measured and simulated data [1]. Because measurement can never acquire complete information of the domain of investigation, electromagnetic inverse problems are usually nonlinear and ill-posed. Many algorithms have been developed to solve these problems. They usually iteratively update the electrical properties to minimize the difference between simulation and measured data. Because the computation involves many linear algebra operations of large matrices and vectors in each iteration, one of the bottlenecks of these algorithms is the computational efficiency. Many researchers have worked on improving the efficiency of inversion algorithms by either reducing the number of unknowns or developing fast forward modeling algorithms, such as the signal-subspace-based methods [2], the iterative multi-scaling method [3], the scattering approximation [4], and etc. Most of them focus on the theoretical improvement of algorithms.

In this work, we investigate the possibility of accelerating nonlinear inversion algorithms using massively parallel computing devices. We studied the multiplicative-regularized contrast source inversion algorithm (MR-CSI) [4]. Furthermore, we use compute unified device architecture (CUDA) as a tool for acceleration of the CSI algorithm because many computational procedures in CSI algorithm can be done highly paralleled. After careful implementation, we observe more than 15 times of improvements in computing speed compared with the original MR-CSI algorithm implemented on CPU.

2. Formulation

We define Domain D as the domain of investigation. Assume transmitting and receiving antennas surround Domain D from different locations around the target of investigation. We also define Domain S as the set of locations of transmitting antennas, use subscript f to denote different frequencies and subscript s for different transmitters. Variables p and q represent the position vectors in Domain S and D, respectively. The contrast can be defined as follows:

$$\chi_f(q) = \frac{\epsilon(q) - \varepsilon_b}{\varepsilon_b} + i\frac{\sigma(q)}{\omega_f \varepsilon_b}, \tag{1}$$

where ε and σ represent permittivity and conductivity that are frequency independent. ε_b represents the permittivity of the background. For 2D TM-polarization case, the field in the domain can be solved by volume integral equation as

$$\begin{aligned} u_{s,f}(p) &= u_{s,f}^{inc}(\mathrm{p}) + k_{b,f}^2 \int_D g_f(p,q)\chi_f(q)u_{s,f}(q)dv \\ &= u_{s,f}^{inc}(\mathrm{p}) + G_{D,f}w_{s,f}, \end{aligned} \tag{2}$$

where

$$g_f(p,q) = \frac{i}{4}H_0^{(1)}\left(k_{b,f}|p-q|\right), \tag{3}$$

$$k_{b,f} = \omega_f\sqrt{\epsilon_b\mu_0}, \tag{4}$$

and $w_{s,f}$ is defined as contrast sources as

$$w_{s,f} = \chi_f u_{s,f}, \tag{5}$$

with $u_{s,f}$ represents the total field in Domain D due to source s, $u^{inc}{}_{s,f}$ represents the incident field in Domain D from transmitter s all at frequency f. $k_{b,f}$ represents the wave number in the background, and $H^{(1)}{}_0(x)$ represents the first kind of Hankel function of zeroth order. The scattered field measured by receiving antennas can be estimated by the following equation:

$$u_{s,f}^{sca}(p) = k_{b,f}^2 \int_D g_f(p,q)\chi_f(q)u_{s,f}(q)dv = G_{S,f}w_{s,f}. \tag{6}$$

Based on the above definitions, the MR-CSI algorithm can be written as minimizing the following cost functional

$$C_n(w_{s,f},\chi_f) = F_n(w_{s,f},\chi_f)F_n^R(\chi_f), \tag{7}$$

where

$$F_n(w_{s,f},\chi_f) = F_n^S(w_{s,f}) + F_n^D(w_{s,f},\chi_f), \tag{11}$$

in which

$$F_n^S(w_{s,f}) = \eta^S \sum_{s,f} \left\| E_{s,f}^{sca} - G_{S,f} w_{s,f} \right\|^2, \qquad (12)$$

and

$$F_n^D(w_{s,f}, \chi_f) = \eta^D \sum_{s,f} \left\| \chi_f u_{s,f}^{inc} - w_{s,f} + \chi_f G_{D,f} w_{s,f} \right\|^2. \qquad (13)$$

η^D and η^S are normalization factors. $F_n^R(\chi_f)$ is the weighted L_2-norm regularization. For details of the above formulation, please refer to [5].

3. Algorithm parallelization on GPU

GPU platform is very good at processing data in a highly parallel fashion with minimum data transfer. We parallel the algorithm based on this principle. One of the advantages of MR-CSI algorithm is that the mapping between source and receiver can be computed paralleled. Therefore, $G_{S,f}$ and $G_{D,f}$ are partitioned into groups based on transmitter index. Based on the fact that linear algebra computation occupies a large part of computational time, we accelerate linear algebra computation first. When computing inner product $\langle a, b \rangle$, where a and b are vectors, we allocate each element in a and the corresponding element in b a thread for parallel computing because the multiplication of each element in the two vectors is unrelated. For the vector generated from paralleled multiplication, the vector elements are divided into groups with two elements inside, and each group is allocated a thread for parallel addition. Then we get a new vector for about half in size. Repeating the grouping and addition operations until we get the inner product result. In the process of computing $A \cdot b$, where A is a matrix and b is a vector, the multiplication of elements in A and b are unrelated. Thus we allocate each element-element multiplication a thread to calculate the multiplication results in parallel. Then for the multiplication of each row of matrix with the vector, we adopt the same grouping and addition operations until we get the desired result vector.

In designing the parallel algorithm on GPU, some key points should be considered in both memory management and thread allocation. For thread allocation, in order to ensure full utilization of GPU processors, the number of threads per block should be at least 32 and divisible by 32 to ensure all the threads can be fully used. This also guarantees every stream processor in one stream multiprocessor fully occupied by threads. In addition, some efficient parallel computing APIs are also used in optimization of the MR-CSI algorithm, such as cuFFT to accelerate fast Fourier transform and cuBLAS for some of the vectors and matrix operations.

4. Numerical example

We validate the GPU-paralleled MR-CSI algorithm using the 2D Fresnel dataset [6]. The dataset is measured uniformly from 2 to 10 GHz with step frequency 1 GHz. All frequencies are used in our inversion. After reconstruction, we plot the complex contrast function at the lowest frequency as shown in Figure 1, from which we can see both dielectric cylinders are properly reconstructed. We use Intel(R) Core(TM) i5-4460 CPU @ 3.20GHz and Nvidia Geforce GTX 480 for computation. The CPU has two cores with four threads and the GPU has 512 stream processors. The program for CPU is compiled using Intel Compiler and linked with Intel MKL library for computing Fast Fourier Transform. The reconstruction on CPU takes 11779 seconds, while the same computation only takes 700 seconds on GPU. The speedup is more than 16 times.

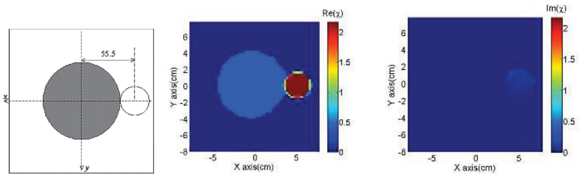

Figure 1 Reconstruction of two dielectric cylinders from Fresnel dataset

5. Conclusion

In this work, we investigate improving the computational efficiency of MR-CSI algorithm using massive parallel computing device. From the numerical tests, we observe more than 15 times reduction in computing time. This could help to alleviate the bottleneck of computing speed of nonlinear inversion algorithms. From this study, we speculate that similar acceleration strategy can be employed in other nonlinear inverse scattering approaches.

Acknowledgment

The authors would like to thank Jean-Michel Geffrin, Pierre Sabouroux and Christelle Eyraud for providing Fresnel dataset as well as Jiantao Qiu from Department of Electronic Engineering at Tsinghua University for sharing CUDA programming knowledge and tips. The authors would also like to thank the support from National Science Foundation of China under contract 61490693 and 61571264, and the National Basic Research Program of China under contract 2013CB329002.

References

[1] T. M. Habashy and A. Abubakar, "A general framework for constraint minimization for the inversion of electromagnetic measurements," *Progress in Electromagnetics Research*., vol. 46, pp. 265-312, 2004.

[2] X. Chen, "Application of signal-subspace and optimization methods in reconstructing extended scatterers," *Journal of the Optical Society of America A*, vol. 26, issue. 5, pp. 1022-1026, March. 2009.

[3] S. Caorsi, M. Donelli, D. Franceschini, and A. Massa, "A New Methodology based on an iterative multiscaling for microwave imaging," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 4, pp. 1162-1173, Apr. 2003.

[4] L.-P. Song and Q. H. Liu, "Fast three-dimensional electromagnetic nonlinear inversion in layered media with a novel scattering approximation," *Inverse Problems*, vol. 20, no. 6, pp. S171-194, Dec.

[5] P. M. van den Berg and A. Abubakar, "Contrast source inversion method: state of art," *Progress in Electromagnetics Research*, vol. 34, pp. 189-218, 2001.

[6] J.-M. Geffrin, P. Sabouroux and C. Eyraud, "Free space experimental scattering database continuation: experimental setup and measurement precision," *Inverse Problems*, vol. 21, no. 6, pp. 117-130, November. 2005.