

Fresnel 层析成像并行算法研究

张建中 杨国辉 林文 蔡骏

(厦门大学信息科学与技术学院 厦门 361005)

(zhangjz@xmu.edu.cn)

A Parallel Algorithm for Fresnel Tomography

Zhang Jianzhong, Yang Guohui, Lin Wen, and Cai Jun

(School of Information Science and Technology, Xiamen University, Xiamen 361005)

Abstract In contrast with ray-based travelttime tomography, Fresnel tomography accounts for the band-limited nature of seismic waves and gives the higher resolution tomograms. Because Fresnel tomography demands much computer memory and much running time, a parallel algorithm for it is proposed. The tomographic inversion is transformed to resolving respectively a series of single equation in light of backprojection principle, each equation corresponding to a Fresnel zone. The forward and inverse computation concerning a Fresnel zone is allocated to one process and is independent of other processes. Then the storage and calculation of the large-scale matrix in the tomography are avoided. No message delivers between the slave processes, and only a little of data delivers between a master process and the slave ones. By using the portable message passing interface standard (MPI) for the communication, the computing code of the algorithm is implemented on Linux system, which allows to distribute the work on several PCs connected via standard Ethernet in an in-house network, and greatly expands the applicability of Fresnel tomography. The tests on the synthetic and observed seismic travel time data show that this parallel algorithm has a good performance on Linux PCs.

Key words tomography; Fresnel zone; travelttime; parallel computing; MPI

摘要 与射线层析成像相比, Fresnel 层析成像考虑波频率的影响, 具有较高的分辨率, 但所需的存储空间和计算量更大, 因此提出了 Fresnel 层析成像的并行算法。把大型层析反演方程组的求解, 转化成对其中的各个方程进行相互独立的计算, 避免了大型系数矩阵的存储问题; 把一个 Fresnel 带的正演和反演计算放在一个进程, 不同 Fresnel 带的计算相互独立进行, 不需要信息传递, 达到了极高的并行度; 从进程之间没有通信, 仅当从进程计算结束后, 在主进程与各从进程之间有少量的数据传递, 使通信开销达到了极小的程度。应用 MPI 在 Linux PC 集群环境下实现了该算法, 实际测试表明, 该算法具有较高的并行度和加速比。

关键词 层析成像; Fresnel 带; 旅行时间; 并行计算; MPI

中图法分类号 TP301.6; TP316.4

层析成像技术是重构地下介质分布的有效手段, 被广泛应用于地球内部结构探测、资源和能源勘探与开发、工程勘查与检测等领域。地震波频率是

影响地震层析成像分辨率的主要因素之一, 有限频率地震波的层析成像是当前关注的研究热点^[1-4]。与常规的射线层析成像相比, Fresnel 层析成像考虑

收稿日期: 2007-02-25; 修回日期: 2007-07-30

基金项目: 国家自然科学基金项目(40774065); 福建省自然科学基金项目(2006J0044)

地震波频率的影响和第一 Fresnel 带内所有介质对地震波传播时间的效应,不但更符合波传播规律,而且能减小层析成像矩阵的零空间维数,减轻层析反演问题的不适定性,提高成像精度和分辨率,目前已取得了一些研究和实验结果^[4-5].

地震层析成像是一个非线性逆问题,一般采用逐步线性化迭代方法来求解.在每次迭代中,不但要计算波传播路径和旅行时间,而且涉及到求解大型矩阵方程.在实际应用中,矩阵方程的阶数往往很高,所需的存储和计算量很大.特别是三维层析成像一直是一个非常耗计算机资源的大型作业.而 Fresnel 层析成像不仅要确定变速介质的第一 Fresnel 带,而且对每个旅行时的正演和反演都要考虑相应的 Fresnel 带内的所有介质,比射线层析成像要求的存储量和计算量更大.单台微机或工作站难以完成这样的任务,从而限制了 Fresnel 层析成像的实际应用.

并行计算技术是解决层析成像大存储量和大计算量问题的有效途径.由于层析成像中的正演计算部分,即各对激发点和接收点之间的射线路径和旅行时的计算,具有相对的独立性,层析成像并行算法的关键是反演问题,即大型矩阵方程组的存储与求解.现有的射线层析成像并行算法主要是通过对大型层析矩阵进行分块,采用分布式存储和计算方式进行并行处理^[6-8].对于存储和计算量更大的 Fresnel 层析成像,我们基于反投影原理,在反演过程的每次迭代中,把大型方程组的求解分成独立的两步进行:第 1 步对一个旅行时直接进行反投影计算,获取该旅行时对应 Fresnel 带内每个离散单元的速度;第 2 步对同一单元对应的所有 Fresnel 带得到的速度求概率均值,得到该单元的速度(或速度修正量).其中在第 1 步,把大型方程组的求解转化成若干单个方程的独立计算,各方程的计算具有充分的并行性,且不需存储层析矩阵.在此反演算法的基础上,结合 Fresnel 层析成像的特点,实现了基于 MPI 的 PC 机群的 Fresnel 层析成像的并行计算,使其达到了实用的程度.

1 方法原理

按照 Fresnel 带的定义,对于频率为 f 、周期为 T 的地震波,激发点 S 和接收点 R 之间的第一 Fresnel 带为满足下面条件的所有点 F 的集合:

$$|t(S, F) + t(F, R) - t(S, R)| \leq T/2, \quad (1)$$

其中, $t(S, F)$ 和 $t(F, R)$ 分别是 S 和 R 到 F 点的最小旅行时; $t(S, R)$ 是从 S 到 R 的最小旅行时,即沿 S 和 R 之间的常规射线的旅行时. F 点的集合在空间内形成一个三维 Fresnel 带,在均匀介质中为以 S 和 R 作为 2 个焦点的扁椭球体.射线处于三维 Fresnel 带的轴线.

由于地震能量主要在第一 Fresnel 带内传播,从激发点 S 到接收点 R 地震波的传播时间与第一 Fresnel 带内的所有介质有关,可表示成下列积分形式^[9]:

$$t(S, R) = \int_L [R(l)]^{-1} s da dl, \quad (2)$$

其中, s 表示模型慢度,是三维空间位置的函数; L 表示从激发点到接收点的三维 Fresnel 带轴线; A 表示与该轴线垂直的 Fresnel 带截面,在均匀介质中截面为圆; dl 表示沿 L 的积分线元; da 表示 A 上的积分面元.

把介质模型用规则网格离散成长方体单元,并统一编号.设第 i 条三维 Fresnel 带轴线在模型第 j 列单元中的长度为 l_{ij} ,对应的 Fresnel 带包含 K 个单元,总体积为 V ,其中第 k 个单元在 Fresnel 带内的体积为 v_k ,根据式(2),对每对激发点和接收点可获得下列离散方程^[10]:

$$\sum_{j=1}^J \sum_{k=1}^K \frac{v_k}{V} l_{ij} s_k = t_i, \quad (3)$$

其中, t_i 是第 i 对激发点和接收点对应的旅行时残差(测量旅行时与当前模型计算旅行时之差); s_k 是第 k 个单元的慢度增量; k 是反映各单元能流密度大小的权系数,在 Fresnel 带轴线穿过的单元权系数最大, Fresnel 带边界上的单元权系数为零,且 $\sum_{k=1}^K k = 1$.

三维层析成像的激发点和接收点的数目很多,即式(3)的方程个数很多,且离散模型单元数一般也很多,这样通过式(3)就形成了一个大型线性方程组,不仅计算量大,而且所需存储空间也很大.层析成像就是逐次迭代地建立和求解该方程组的过程.

2 并行算法

与射线层析成像相比, Fresnel 层析成像具有更大的计算量和存储量.一方面, Fresnel 层析成像需要确定每一对激发点和接收点对应的三维 Fresnel 带.式(1)表示了确定 Fresnel 带的精确的和直接的

方法. 其中不仅要计算从激发点到模型各处的旅行时^[11-12], 而且也要计算从接收点到模型各处的旅行时, 计算量特别大, 所需要的计算机存储空间也非常大, 这是与常规射线层析成像区别最大的地方. 另一方面, 反演方程组中的每一个方程包含比射线方法更多的未知数, 大型层析矩阵的存储问题也非常严峻. 单台 PC 机或工作站是远远满足不了解决实际问题的需要, 因此, 我们根据 Fresnel 层析成像的特点研究了并行算法, 着重考虑了有关的存储问题、多进程计算及其通信问题.

1) 反演算法的并行化. 在层析成像非线性反演的每次迭代中, 需要求解由式(3)构成的大型线性方程组, 直接求解将占用很大内存. 我们构造出按行计算的算法, 以实现并行计算. 把式(3)简记为

$$\sum_{j=1}^M ij \cdot sj = ti, \tag{4}$$

其中, M 表示第 i 个 Fresnel 带内的所有模型单元数. 利用反投影算法原理^[13], 由第 i 个 Fresnel 带对应的旅行时残差 ti 得到的第 j 个单元的慢度修正量可表示为

$$s_{ij} = \frac{sj \cdot ti}{ti} \cdot \frac{ij}{M}, \tag{5}$$

式中, sj 是当前模型第 j 个单元的慢度, ti 是计算的当前模型第 i 个 Fresnel 带对应的旅行时. 该式相当于按不同单元对旅行时贡献的大小进行反投影. 对于不同的 Fresnel 带旅行时残差 ti , 求得的一个单元上的 s_{ij} 一般是不同的, 可以把 s_{ij} 看成是一随机变量, 求其均值, 可获得表示为第 j 个单元的慢度增量, 即

$$sj = \frac{1}{N} \sum_{i=1}^N s_{ij}, \tag{6}$$

其中 N 为覆盖第 j 个单元的 Fresnel 带个数.

式(5)仅涉及方程组中一个方程的运算, 而式(6)只与未知量有关, 而与方程系数无关. 这样, 就把大型方程组的求解转化成若干单个方程的独立计算, 避免了大型方程系数矩阵的存储, 而单个方程的存储和计算可分给不同的子进程独立进行, 从而实现了大型方程组的并行求解.

2) 计算任务的并行处理. 在 Fresnel 层析成像中, 一个离散单元一般要处在很多条不同的 Fresnel 带中, 它的慢度要由这些所有的 Fresnel 带对应的旅行时同时决定. 由于从激发点或接收点到模型各处的旅行时的计算是相互独立的, 使得不同激发点和接收点对应的 Fresnel 带的计算也是相互独立的.

反演算法通过上述的并行化处理, 利用式(5), 由各个旅行时残差计算节点的慢度修正量也是相互独立的. 这样, 我们就可以把 Fresnel 层析成像的整体问题转化成非常充分的并行处理问题. 即先分别独立地逐个进行每一对激发点和接收点对应的正演和反演计算, 再利用式(6)简单快速地获得模型各单元的慢度增量 sj . 所以, 在理论上, 层析成像有多少对激发点和接收点, 即多少条 Fresnel 带就可以使用多少个进程进行独立计算.

当然, 进程数一般是远远少于激发点和接收点对的数目, 所以需要根据进程数对整个计算任务进行分解. 考虑到接收点数一般多于激发点数, 在对计算任务进行分解时, 先按进程数对总接收点数进行平均, 把接收点平均分给各个进程, 再把每个进程中的接收点涉及的激发点也分给该进程. 这样, 各个进程可以独立进行相应的正演和反演计算. 需要指出的是, 这种划分使一个激发点可能放在不同的进程中, 并在各个进程中对该激发点的波前时间进行了重复计算. 这完全是出于尽量减少各进程之间通信量的考虑.

在每个进程中, 对应于每个 Fresnel 带(激发点 S 和接收点 R)的计算包括: 计算式(1)中的各量^[11-12]; 由式(1)确定 Fresnel 带; 计算式(4)中的系数; 由式(5)计算 s_{ij} . 每条 Fresnel 带的这些计算可以独立进行. 而且, 式(5)中的系数 ij 只对应一条 Fresnel 带, 而与其他 Fresnel 带无关, 这样就不需要存储这些系数, 避免了层析成像中大型方程组系数矩阵的存储问题.

3) 主从进程之间的通信. 由于在数据和任务分解时, 把同一条 Fresnel 带的激发点和接收点及其数据都放在同一个进程中, 各个从进程使用的数据和进行的正演反演计算完全是独立的, 所以从进程之间不存在数据的传递, 没有通信开销. 式(6)的计算要用到各进程独立计算出的 s_{ij} , 这就需要各从进程把计算的这些 s_{ij} 发给主进程. 这是主从进程之间也是整个算法的主要通信量. 与现有层析成像并行算法相比极大地减少了通信量. 这一点对加快实际生产任务的进度非常重要.

3 程序实现

MPI是目前最为流行的消息传递并行环境之一, 得到了广泛的应用. 我们在 Linux 集群环境下, 采用 MPICH-1.2.5 并行库, 实现了上述 Fresnel 层



析成像的并行算法. 程序在结构上采用主从与对等的混合模式, 主进程也承担一个从进程的计算工作, 以增大并行粒度. 在通信方面, 采用标准通信模式, 调用 MPI 消息传递库函数 MPI-SEND() 和 MPI-

RECV() 分别实现对数据的发送和接收. 定义了 MPI 的结构数据体变量, 把需要传输的不同的计算数据存入一个数据体数组进行发送, 以减少通信次数, 降低通信开销. 程序流程如图 1 所示:

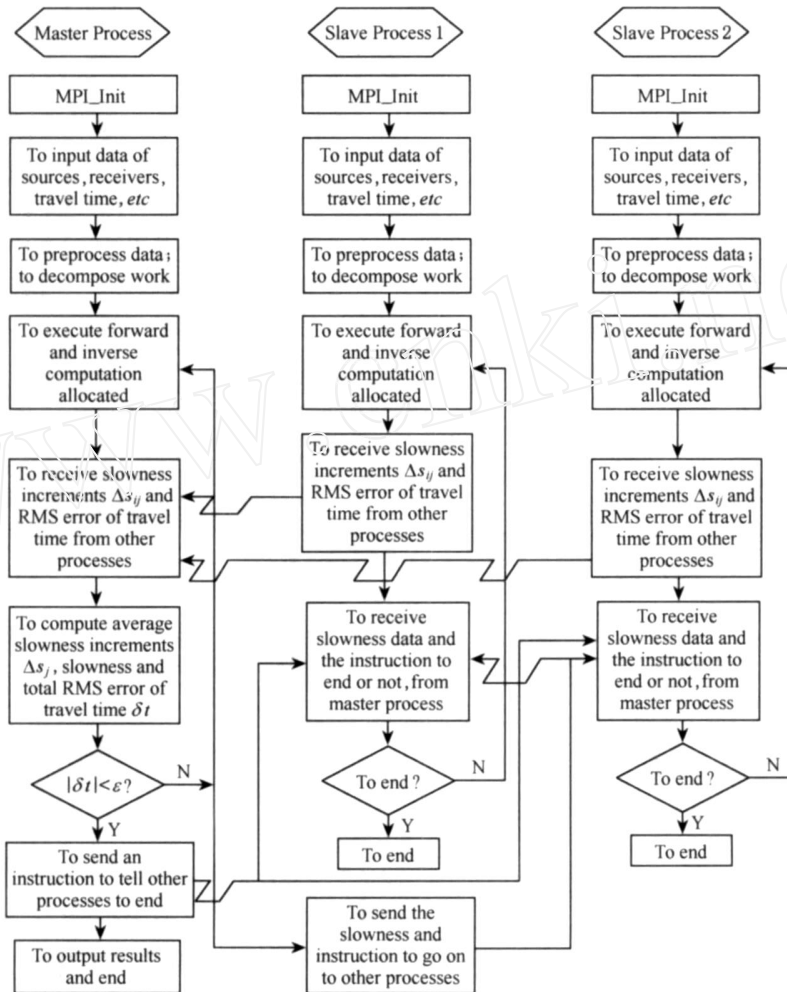


Fig. 1 Flow chart of the parallel algorithm.

图 1 并行算法流程图

4 算法和程序测试

为了检验本文算法的正确性, 我们用三维理论模型的合成数据进行了实验. 模型由上下两层介质组成, 上层介质速度 2200m/s, 下层介质速度 4500m/s, 在上层介质浅部有一个速度为 500m/s 的低速体. 理论模型在 $y = 0$ 处的切片如图 2(a) 所示, 本文算法结果的对应切片如图 2(b) 所示. 可以看出本文算法的正确性和良好效果.

为了测试算法的并行性能, 对实际地震勘探的初至时间资料进行了处理. 该资料激发点 386 个, 接收点 1089 个, 时间数据(即 Fresnel 带) 60000 多

个, 离散模型正演和反演单元数 10000 多个. 分别用 1, 8, 16, 24 个进程进行了计算, 取 3 次迭代的运行时间进行平均, 计算了加速比(串行运行时间/并行运行时间)、效率(加速比/进程数)和成本(运行时间 × 进程数)等并行性能度量参数^[14], 一次迭代计算的测试结果如表 1 所示. 对于所选的这个数据量较小的资料, 反演迭代 30 多次, 若使用单进程, 约需 7 个多小时, 若使用 24 个进程, 所需时间只有半个多小时, 多进程大大缩短了工作时间. 测试结果表明, 该算法具有较高的并行加速比, 而且随着进程数的增加所用时间会大大减少, 并行加速比不断增大. 同时, 随着进程数的增加, 并行效率在不断降低, 运行成本也不断增高. 这是因为进程数增加将使主从

进程之间的通信量增加,从而增加运行时间;另一方面,为了使具有较大计算量和数据量的 Fresnel 带的计算只在各进程独立进行,避免从进程之间较重的通信量,我们在分解观测物理点时,按照接收点数目进行分配,并根据接收点放置激发点,这样,同一个激发点可能会分给几个不同的进程,使同一个激发点波前时间的计算在不同进程中重复几次,从而降低了并行效率. 但这样做大大减少了通信量,提高了加速比. 经过实验比较,对于实际应用,这种做法是可取的,也是必要的.

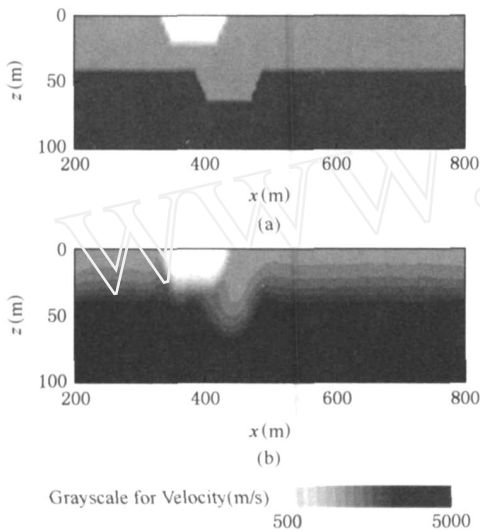


Fig. 2 The slice of tomographic velocity model at $y = 0$, using synthetic data. (a) Theoretical model and (b) Tomographic model.

图 2 利用合成数据的层析成像速度模型在 $y = 0$ 处的切片图. (a) 理论模型; (b) 层析成像模型

Table 1 Test Results of Parallel Computing

Performance for One Iteration Step

表 1 一次迭代计算的并行性能测试结果

Process Number	Running Time(s)	Accelerate Ratio	Parallel Efficiency(%)	Cost (s)
1	777	1	100	777
8	191.1	4.07	51	1529
16	120	6.48	41	1920
24	93.5	8.31	35	2244

5 结 论

Fresnel 层析成像具有较高的分辨率,但所需要更大的计算量和更大的存储空间制约了它的实际应用,并行计算解决该问题的有效途径. 本文提出和实现了 Fresnel 层析成像的并行算法,且具有下列

特点: 把大型层析反演方程组的求解,转化成对其中的各个方程进行相互独立的计算,避免了大型系数矩阵的存储问题; 不同 Fresnel 带对应的有关正演和反演计算都可以放在不同的进程独立进行,这些计算之间不需要信息传递,使算法达到了极高的并行度; 从进程之间没有通信,仅当从进程计算结束后,在主进程与各从进程之间有少量的数据传递,使整个算法的通信开销达到了极小的程度;

一些激发点的部分正演计算在不同进程中有一些重复,使得随着进程数的增加运行成本增加,并行效率降低. 这种设计是为了避免从进程之间的通信,以便在实际工作中可以尽量使用更多的进程,加快整个作业的进度. 实际资料测试结果表明,该算法设计合理,具有较高的并行度和加速比,使 Fresnel 层析成像的运行时间达到了实际生产可以接受的程度.

致谢 本研究还得到中国石油东方地球物理公司物探技术研究中心和厦门大学 985 二期信息创新平台项目(0000-X07204)的资助,赵波博士和陈宝孚高级工程师在技术研究和资料处理等方面给予了大力帮助,在此一并表示感谢!

参 考 文 献

- [1] De Nicola S, Fedele R, Man'ko M A, et al. Fresnel tomography: A novel approach to wave-function reconstruction based on the Fresnel representation of tomograms [J]. Theoretical and Mathematical Physics, 2005, 144(2): 1206-1213
- [2] J Sheng, G T Schuster. Finite-frequency resolution limits of wave path travelttime tomography for smoothly varying velocity models [J]. Geophysical Journal International, 2003, 152(3): 669-676
- [3] T C Johnson, P S Routh, M D Knoll. Fresnel volume georadar attenuation-difference tomography [J]. Geophysical Journal International, 2005, 162(1): 9-24
- [4] J Spetzler, R Snieder. The Fresnel volume and transmitted waves [J]. Geophysics, 2004, 69(3): 653-663
- [5] T C Johnson, P S Routh, W Barrash, et al. A field comparison of Fresnel zone and ray-based GPR attenuation-difference tomography for time-lapse imaging of electrically anomalous tracer or contaminant plumes [J]. Geophysics, 2007, 72(2): G21-G29
- [6] Luo Shengxian, Li Luming. Parallel computation collective communication based on MPI and its application [J]. Computer Applications, 2003, 23(6): 51-53 (in Chinese) (罗省贤, 李录明. 基于 MPI 的并行计算集群通信及应用 [J]. 计算机应用, 2003, 23(6): 51-53)

- [7] Liu Jinsong, Liu Futian, Liu Jun, *et al.* The optimization and parallelization of sequential orthogonal triangularization algorithms used in seismic tomography [J]. *Progress in Geophysics*, 2005, 20(4): 911-915 (in Chinese)
(刘劲松, 刘福田, 刘俊, 等. 天然地震走时反演矩阵顺序三角化算法的优化和并行化[J]. *地球物理学进展*, 2005, 20(4): 911-915)
- [8] Liu Jinsong, Liu Futian, Liu Jun, *et al.* The parallelization of LSQR algorithms used in seismic tomography [J]. *Super Computation Letter*, 2005, 3(2): 5-11 (in Chinese)
(刘劲松, 刘福田, 刘俊, 等. 地震层析成像 LSQR 算法的并行化[J]. *超级计算通信*, 2005, 3(2): 5-11)
- [9] D W Vasco, J E Peterson, E L Jr Majer. Beyond ray tomography: Wavepaths and Fresnel volumes [J]. *Geophysics*, 1995, 60(6): 1790-1804
- [10] Wang Hongwei. Traveltime seismic tomography with Fresnel volume ray tracing [J]. *Oil Geophysical Prospecting*, 1996, 31(3): 309-326 (in Chinese)
(王宏伟. 菲涅耳体层析成像[J]. *石油地球物理勘探*, 1996, 31(3): 309-326)
- [11] Zhang Jianzhong, Chen Shijun, Xu Chuwei. A method of shortest path raytracing with dynamic networks [J]. *Chinese Journal of Geophysics*, 2004, 47(5): 899-904 (in Chinese)
(张建中, 陈世军, 徐初伟. 动态网络最短路径射线追踪[J]. *地球物理学报*, 2004, 47(5): 899-904)
- [12] Zhang Jianzhong, Chen Shijun. Numerical modeling of seismic first break in complex media [J]. *Chinese Journal of Computational Physics*, 2003, 20(5): 429-433 (in Chinese)
(张建中, 陈世军. 复杂介质地震初至波数值模拟[J]. *计算物理*, 2003, 20(5): 429-433)
- [13] J A Hole. Nonlinear high-resolution three-dimensional seismic travel time tomography [J]. *Journal of Geophysical Research*, 1992, 97(B5): 6553-6562
- [14] Zhang Linbo, Chi Xuebin, Mo Zeyao, *et al.* Introduction to Parallel Computing [M]. Beijing: Tsinghua University Press, 2006. 210-215 (in Chinese)
(张林波, 迟学斌, 莫则尧, 等. 并行计算导论[M]. 清华大学出版社, 2006. 210-215)

Research Background

It is common to use a geometric ray theory in seismic tomography. This is partly due to the limitations in computer power and memory because ray theory is computationally efficient and easy to implement. According to ray theory it is assumed that waves propagate along infinitely narrow lines through space that join the source and receiver. In reality, recorded waves have finite-frequency content. The band limitation of waves implies that the propagation of waves is extended to a 3-D Fresnel zone. Consequently Fresnel tomography can reach higher resolution than the ray-based, but requires more computation and more memory. In recent years Fresnel tomography has been investigated actively with the development of high performance computing technique. It should be a better way to use clusters of workstations or PCs for the application of Fresnel tomography. Therefore, we research and implement the parallel algorithm for Fresnel tomography on PCs using MPI. Our work is supported by the Natural Science Foundation of Fujian Province of China (2006J0044), the Geophysical Technique Center, BGP, CNPC, and the 985 Innovation Project on Information Technique of Xiamen University (0000-X07204).



Zhang Jianzhong, born in 1963. Received his Ph. D. degree in 2001, and has been a professor since 2002. Ph. D. supervisor. His main research interests include numerical computation and digital imaging for electromagnetic wave and elastic wave, and signal and information processing.

张建中, 1963年生, 博士, 教授, 博士生导师, 主要研究方向为电磁波和弹性波的数值计算与数字成像、信号与信息处理等。



Yang Guohui, born in 1983. He now is a postgraduate for master degree in communication engineering. His current research interests are signal and information processing for communication engineering.

杨国辉, 1983年生, 硕士研究生, 主要研究方向为通信信号处理。



Lin Wen, born in 1984. He now is a postgraduate for master degree in communication engineering. His current research interests are signal and information processing for communication networks.

林文, 1984年生, 硕士研究生, 主要研究方向为通信网信息处理。



Cai Jun, born in 1966. Received his doctor degree of engineering in 1996 and now is an associate professor of Xiamen University. His main research interests are speech recognition, artificial intelligence, machine

learning, and design of operating systems.

蔡骏, 1966年生, 博士, 副教授, 主要研究方向为语音识别、人工智能、机器学习、操作系统设计。