

理论地震图的 F-K 算法的并行实现*

韩立波^{1,2}, 郑勇³, 倪四道^{1,2}

(1. 中国科学技术大学地球和空间科学学院, 安徽合肥 230026; 2. 蒙城地球物理国家野外科学观测研究站, 安徽蒙城 233500;
3. 中国地震局地球物理研究所, 北京 100081)

摘要:对 F-K 算法进行了并行化改进, 从而使其可以利用 PC-Cluster 或者并行计算机多节点计算的资源优势, 提高计算速度. 比较发现, 并行后的计算速度大大提高, 在计算时间较长的情况下, 并行 F-K 算法的运算速度基本与处理器的个数成正比, 为反演震源参数和地下结构提供了更为快捷的计算工具.

关键词: F-K 方法; 并行计算; 计算速度

中图分类号: P319.2 **文献标识码:** A

Parallelization of F-K method of synthetic of seismograms

HAN Li-bo^{1,2}, ZHENG Yong³, NI Si-dao^{1,2}

(1. School of Earth and Space Sciences, University of Science and Technology of China, Hefei 230026, China;
2. Mengcheng National Geophysical Observatory, Mengcheng 233500, China;
3. Geophysics Institute of CEA, Beijing 100081, China)

Abstract: In order to improve computation speed, a parallelized F-K program was built to utilize the multiple CPU PC-Clusters. By comparison it was found that the parallelized program can greatly expedite computation. For a complicated model, parallel F-K computation speed is almost proportional to the number of the processors. This method provides a more convenient tool for inversion of source parameters and underground structures.

Key words: F-K method; parallel computation; computation speed

0 引言

限于当前计算机工艺等单个 CPU 的计算能力有限, 因而面对大运算量的问题显得力不从心. 而多个 CPU 并行的 cluster 则以较低的经济成本大大提高了运算能力, 并行算法正成为一种崭新的解决问题的方法. 特别是随着现代计算机技术的发展, 并行计算机得到了越来越多的应用. 现代计算机能够将

几百个甚至几万个 CPU 组合在一起进行并行计算, 以往传统的串行程序显然不能充分发挥并行计算机的优势, 而采用高效率并行算法的并行程序则能成百上千倍的提高计算速度.

对于地下结构和地震震源机制的了解主要来自于地震图的信息. 传统上主要依靠地震波震相的走时和地震的初动解来分析地下结构和震源信息^[1]. 随着地震学的发展和计算水平的进步, 合成理论地

* 收稿日期: 2007-03-05; 修回日期: 2007-05-30

基金项目: 国家自然科学基金杰出青年基金(40425005)和中国科学院引进海外杰出人才计划资助.

作者简介: 韩立波, 男, 1979 年生, 博士生. 研究方向: 地球动力学和大规模并行计算研究. E-mail: lbhan@mail.ustc.edu.cn

通讯作者: 倪四道, 博士/教授. E-mail: sdni@ustc.edu.cn

震图方法逐渐成为研究地震学问题的主要手段. 目前, 计算理论地震图的方法主要有有限元法^[2], 有限差分法^[3], Modes 叠加法^[4], 伪谱法^[5]和频率-波数法^[6](F-K)等. 其中, F-K 方法对频率和波数进行分别积分, 采用传播矩阵计算地震的位移场分布, 由于该方法理论清晰, 计算准确, 因此在计算分层模型的理论合成地震图时, 应用非常广泛.

Haskell^[7,8], Harkrider^[9]利用传播矩阵给出了介质中点源的表面位移解. Herrmann 等^[10]对 Haskell 的方法做了一些改正, 使得其计算更为简便, 便于计算机实现. 传统上, 层状半空间中不同源产生的动态位移问题和静态变形问题是分开处理的. Zhu 等^[6]证明了在角频率趋近于 0 的情况下, 动态位移解收敛为静态变形解, 因此可以利用动态位移解处理所有频率下的表面位移解. 在此基础上, Zhu 等开发了 F-K 的计算程序, 该程序目前广泛应用于地震震源反演理论地震图拟合等研究中, 特别由于该方法能够计算各种频率下的体波和面波波形, 因此对于利用近震研究震源参数非常有效^[11].

虽然 F-K 方法得到了广泛的应用, 但其计算速度仍然显得较慢. 由于在计算位移时需要同时对频率(F)和波数(K)进行双重积分, 因此计算量相对较大. 通常, 在计算波形拟合时首先给出一个震源深度和地下速度模型, 然后计算 3 种基本断层类型的响应格林函数 (Green's Function), 再利用 Wang 等^[10]给出的一般地震图与该三种基本断层响应 Green's Function 的关系计算不同震源机制解下的理论地震图. 该方法虽然在一定程度上解决了大量震源机制解下的理论地震图的合成问题, 从而使得震源机制反演成为可能, 但在计算 Green's Function 时仍然需要大量的时间, 特别是在计算不同深度下的合成地震图时, 该方法计算速度上的缺点将更加突出. 一些近似的方法虽然能够大大提高计算速度, 例如射线理论和广义射线理论能够快速计算远震情况下的理论波形, 但由于它们都是基于高频近似这一假设的, 因此, 对于近震和低频情况处理则不太适用. 为此, 本文将传统的串行的 F-K 程序改写成并行程序, 从而达到在保证计算精度的前提下大幅度提高速度的目的.

1 理论和方法

在柱坐标系(e_r, e_θ, e_z)中垂直各向不均一性介质中的位移可以表示为 3 个正矢量的形式^[12]:

$$\mathbf{u}(r, t) = \frac{1}{2\pi} \sum_{m=0, \pm 1, \dots} \int_0^\infty e^{-i\omega t} d\omega \int_0^\infty k dk U(\omega, k) J_n(kr). \quad (1)$$

式中, $U(\omega, k)$ 为地表位移核函数, 此处没有考虑角度的变化, 只计算 Green's Function, 当考虑到震源方位角信息时, 只需利用震源辐射花样的函数卷积 Green's Function 就可以得到合成地震图, 因此计算 Green's Function 在合成地震图中占据最主要的计算量.

对于式(1)的积分, 目前有多种方法进行计算, 如 Cagniard-de Hoop 方法、最速下降法、慢度法等, 而离散 F-K 方法是目前经常使用的方法, 把 K 积分近似为离散波数域上的求和^[13~16]. ω 积分则可以通过快速傅立叶变换实现. Zhu 等^[6]在设计串行的 F-K 程序时, 采用了如下的计算公式进行 K 积分的计算:

$$\int k U(\omega, k) J_n(kr) dk = \Delta k \sum_{k=\frac{\Delta k}{2}, \frac{3\Delta k}{2}, \dots}^{k_{\max}} k U(\omega, k) J_n(kr), \quad (2)$$

而在合成位移场时, 仍旧采用式(1)的计算公式.

由 F-K 的理论可知, 波数域的积分 K 与频率 ω 是耦合的, 对其难以实现解耦操作; 而频率域的计算则与波数域无关, 因此很容易实现分离计算, 从而进行并行处理. 基于这个原因, 我们将频率积分化为有限频率的求和, 开 N 个并行的线程, 然后将频率的求和均匀分成 N 段, 每个线程计算一段的频率积分, 最后将所有线程的结果送到一个指定的线程进行组装, 这样就能实现 F-K 方法的并行操作. 设最小频率, 最大频率和频率步长分别为 ω_{\min} , ω_{\max} 和 $\Delta\omega$, 则第 i 个线程处理的频率段为

$$\begin{aligned} \omega_{\min} + \frac{\omega_{\max} - \omega_{\min}}{N} \cdot (i-1) \sim \\ \omega_{\min} + \frac{\omega_{\max} - \omega_{\min}}{N} \cdot i - \Delta\omega. \end{aligned} \quad (3)$$

为了实现并行操作, 我们采用了 MPI 方法进行并行处理. MPI 是一种消息传递编程模型, 在 MPI 上很容易移植其他的并行代码, 消息传递方式是广泛应用与多类并行机的一种模式, 特别是那些分布存储并行机. 它是一个库, MPI 库可以方便的为 Fortran77/C/Fortran90/C++ 所调用.

由于 F-K 的计算量最大的部分在于频率域的积分运算, 所以我们将频率域的离散积分根据处理

器的个数拆分,使得每个处理器处理一部分频率,这样就可以尽量利用处理器资源,大大提高了计算速度.虽然并行进程间的通信相对于串行来说,会花费一定的时间,但是相对于计算时间来说,代价仍然是微小的,特别是当计算量特别巨大的时候,这种影响更加显得无足轻重.

2 计算结果及分析

本文计算使用的服务器集群拥有 4 个 CPU, CPU 均为 32 位 Intel(R) Xeon(TM) CPU 2.40 GHz. 机器内存为 2 GB. 操作系统是 Fedora Core5, MPICH 版本为 2-1.0.2pl1.

2.1 程序准确性测试

我们首先使用一简单速度模型来对 F-K 的并行程序准确性进行验证. 该模型震源深度为 10.5 km, 震中距为 274.34 km, 参数如表 1 所示.

表 1 简单测试速度模型

Tab. 1 A simple velocity model used to bench

层厚 /km	V_S / $(\text{km} \cdot \text{s}^{-1})$	V_P / $(\text{km} \cdot \text{s}^{-1})$	ρ / $(\text{kg} \cdot \text{m}^{-3})$	Q_α	Q_β
1.400 00	2.400 0	4.920 00	2.100 0	150	50
6.200 0	3.500 0	5.950 00	2.750 0	600	250
13.800 0	3.650 00	6.210 00	2.800 0	750	300
11.100 0	4.020 00	6.840 00	2.900 0	1 200	800
0.000 0	4.720 00	8.030 0	3.330 0	1 200	800

为验证并行程序准确性,我们把 3 个 CPU 计算得到的格林函数某一分量和对应的串行程序得到的格林函数分量进行对比,如图 1 所示. 其中实线为并行程序结果,虚线为串行程序结果,可以看到二者几乎没有差别(虚线全部被实线覆盖). 图 1(b)为截取图 1(a)中一段放大后所见,同样的二者几乎没有

差别.

为了进一步测试并行程序的准确性和速度的提高效果. 我们使用了 4 个模型,即 Prem5、Prem10、Prem15 和 Prem20,进行了计算和比较. Prem5 为取 PREM 模型最上面 5 层, Prem10 为取 PREM 模型最上面 10 层,其他与此类似. 震源深度为 15 km, 震中距为 274.34 km.

基于此 4 个模型,我们分别用 3 CPU 并行, 2 CPU 并行和串行 F-K 程序对此模型进行了计算. 同样把串行和 3 CPU 并行的对应的格林函数分量画在一起进行比较. 结果如图 2 所示,实线为并行结果,虚线为串行结果. 同样的我们可以看到二者几乎无差别,只是在很高频率的拐角处出现微小的差异,这主要是因为计算式(2)时采用了近似和低通滤波处理,由于滤波器的边界效应,多段的并行计算时边界效应与单线程计算时有少量差别,从而导致这种差异. 但由于这些差异本身很小,加上主要处于高频部分,对于我们关心的频段没有影响,因此可以忽略.

2.2 计算速度的比较

并行计算的根本目的是在于提高计算速度. 计算速度可以用计算耗费的时间来衡量. 前面所述 4 个 PREM 模型计算所耗费时间如表 2 所示. 为了对

表 2 4 个 PREM 模型计算耗费时间表

Tab. 2 Elapsed time table of computation using these four PREM models

模型	3 个 CPU 计算 耗费时间/s	2 个 CPU 计算 耗费时间/s	串行计算 耗费时间/s
Prem5	87	121	170
Prem10	142	196	277
Prem15	197	275	389
Prem20	251	350	498

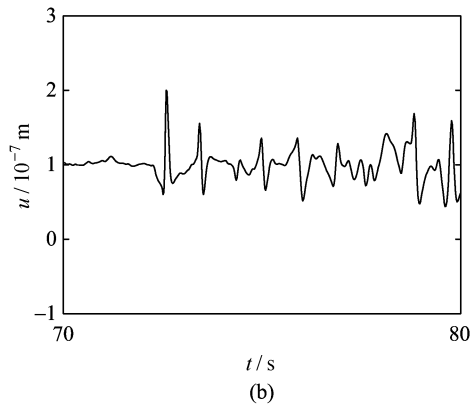
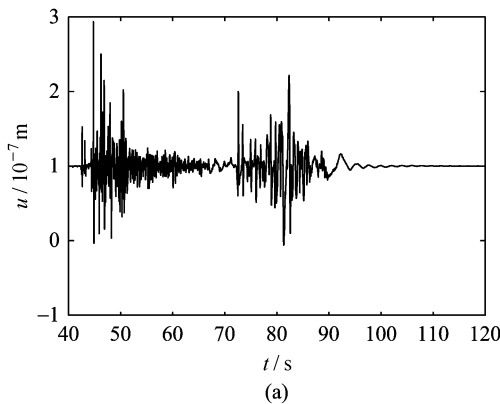


图 1 简单测试速度模型的并行和串行计算结果比较图

Fig. 1 Compare of parallel and serial result of a simple velocity model

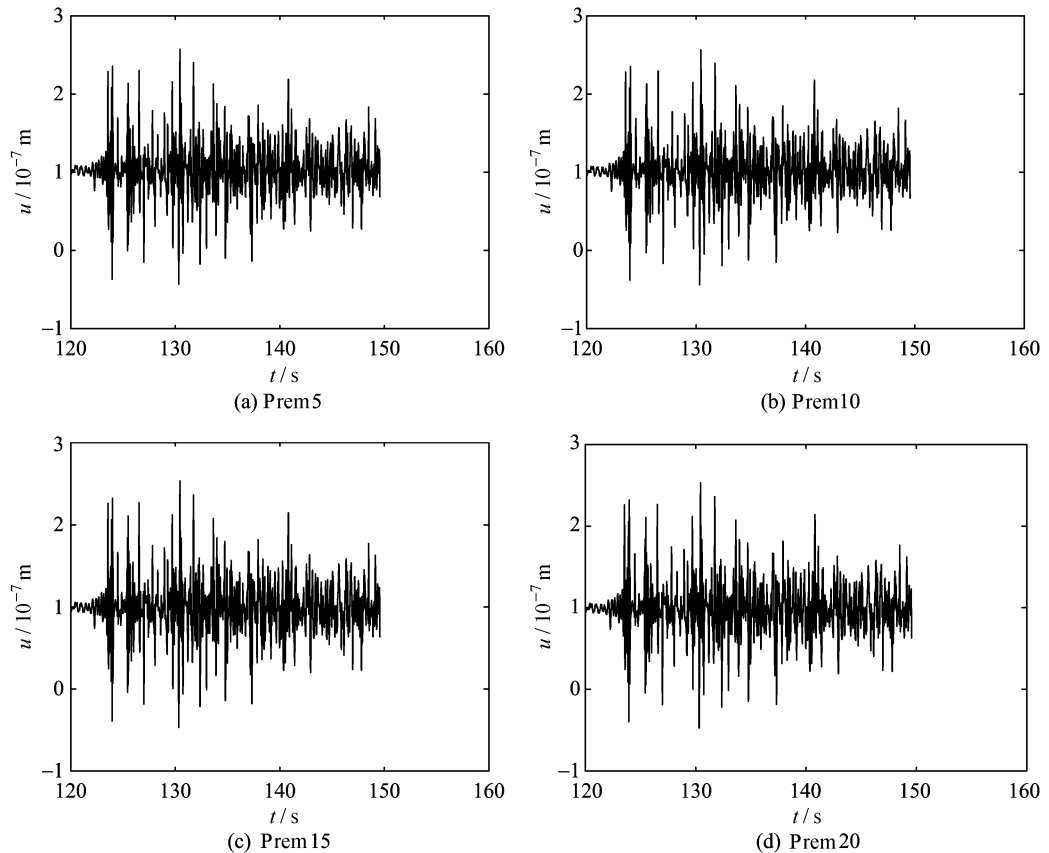


图 2 模型的并行计算和串行计算的结果比较

Fig. 2 Compare of parallel and serial result of four models

照起见,表 2 所列的耗费时间仅仅指用来处理并行部分所花费的时间.串行的耗费时间是指完成相同工作所用时间.

可以看到随着参加运算 CPU 数目的增加,耗费时间明显呈下降趋势,计算速度大幅度提高.为验证此并行速度变化趋势,我们使用了 Prem50 模型(其他参数同 Prem5),来观察此模型在 1 个 CPU(即串行)和 2,3,4 个 CPU 下计算速度的变化趋势,见图 3.图 3 中纵坐标为相对于串程序的计算速度,横坐标为使用的节点数目.我们可以看出随着 CPU 数目增加计算速度确实有明显提高,基本随 CPU 数呈近似线性增长,但随节点数目的增多 CPU 的平均效率有减小的趋势.

3 结论

我们通过对 F-K 方法的分析,在串程序的基础上建立了并行 F-K 程序.通过将频率均匀分配到多个节点上并行计算的手段,大大提高了 F-K 程序的运算速度.通过与串行计算结果比较发现,并行

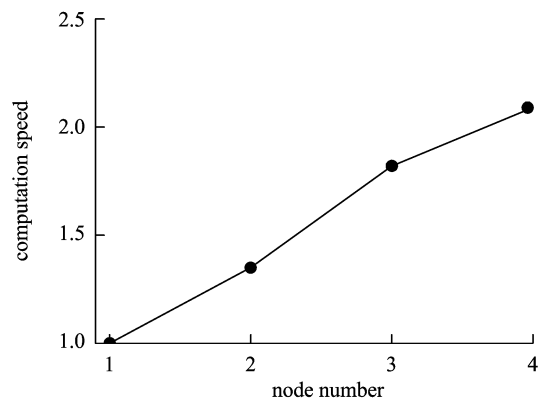


图 3 计算速度随 CPU 数目的变化图

Fig. 3 Figure of computation speed compare to CPU numbers

程序的计算结果可靠,能够满足波形拟合等研究的需要.通过不同节点数目的计算时间进行分析可知该并行程序的计算速度与节点数据呈近似线性关系增长,但在节点数目增加时计算速度增加幅度有小幅减小趋势.总体而言,该程序稳定可靠性,计算速度快,较好的解决了 F-K 方法计算耗时长的

问题.

在本次研究中,我们首要关心的是并行 F-K 方法的稳定可靠,因此做了大量和串行版本的波形对比,并行版本的单线程和串行版本的波形完全一致,而随着线程增多,两者有微小差异,我们认为这是由于程序本身的低通滤波和浮点误差造成的,在误差许可范围内.

另外关于并行的效率,我们对 F-K 的并行改进在一定范围内大幅度提高了运算速度,可以更加充分地利用处理器资源.但是通过计算发现 2 个 CPU 并行时,计算速度约为串行的 1.4 倍,平均每个 CPU 的效率为 70%;而 3 个 CPU 并行时,计算速度约为串行的 2 倍,平均每个 CPU 的效率下降为 65%.这说明 CPU 数目和进程数目的增加造成了额外的时间消耗,其消耗主要用于进程通信和数据交换等等.而且考虑到 F-K 频率域和波数域有一定相关性,即高频区间内波数计算精度也相应较高,因此频率域内均匀分割不一定是最有效的方案.这表明,程序的算法有进一步改进的空间,可以通过设计合理的进程通讯和数据传输手段,以及进一步协调进程间任务管理,更大程度地提高 F-K 程序的计算速度.

致谢 我们感谢 Zhu Lu-pei 博士提供了 F-K 的串行程序,感谢傅容珊老师的指导,感谢实验室的黄建平、万柯松、阮友谊、詹中文等人帮助测试并行程序,感谢马宏生及邵志刚在文章撰写中的指导.

参考文献(References)

- [1] Aki K, Richards P G. Quantitative Seismology: Theory and Methods [M]. New York: Freeman, 1980.
- [2] Smith W D. A nonreflecting plane boundary for wave propagation problems[J]. J Comp Phys, 1974, 15: 492-503.
- [3] Kelly K R, Ward R W, Treitel S, et al. Synthetic seismograms: A finite-difference approach [J]. Geophysics, 1976, 41: 2-27.
- [4] Knopoff L, Schwab F, Kausel E. Interpretation of Lg [J]. Geophy J R Astr Soc, 1973, 33: 983-993.
- [5] Furumura T, Kennett B L N, Furumura M. Seismic wavefield calculation for laterally heterogeneous whole earth models using the pseudospectral method [J]. Geophysical Journal International, 1998, 135 (3): 845-860.
- [6] Zhu L, Rivera L A. A note on the dynamic and static displacements from a point source in multi-layered media[J]. Geophys J Int, 2002, 148: 619-627.
- [7] Haskell N A. Radiation pattern of Rayleigh waves from a fault of arbitrary dip and direction of motion in a homogeneous medium[J]. Bull Seismol Soc Am, 1963, 53: 619-642.
- [8] Haskell N A. Radiation pattern of surface waves from point sources in a multi-layered medium [J]. Bull Seismol Soc Am, 1964, 54: 377-393.
- [9] Harkrider D G. Surface waves in multilayered elastic media I: Rayleigh and Love from buried sources in a multilayered elastic half-space [J]. Bull Seismol Soc Am, 1964, 54: 627-679.
- [10] Wang C Y, Herrmann R B. A numerical study of P, SV, and SH-wave generation in a plane layered medium [J]. Bull Seismol Soc Am, 1980, 70: 1 015-1 036.
- [11] Zhu L, Helmberger D V. Advancements in source estimation techniques using broadband regional seismograms[J]. BSSA, 1996, 86: 1 634-1 641.
- [12] Takeuchi H, Saito M. Seismic surface waves [M]// Methods in Computational Physics. New York: Academic Press, 1972, 11: 217-295.
- [13] Bouchon M, Aki K. Discrete wave number representation of seismic source wave fields [J]. Bull Seismol Soc Am, 1977, 67: 259-277.
- [14] Bouchon M. A simple method to calculate Green's function for elastic layered media [J]. Bull Seismol Soc Am, 1981, 71: 959-971.
- [15] Yao Z X, Harkrider D G. A generalized reflection transmission coefficient matrix and discrete wavenumber method for synthetic seismograms [J]. Bull Seismol Soc Am, 1983, 73: 1 685-1 699.
- [16] Saikia C K, Burdick L J. Fine-structure of P_n waves from explosion [J]. J Geophys Res, 1991, 96: 14 383-14 401.