DOI:10.14075/j.jgg.2019.10.009 文章编号:1671-5942(2019)10-1033-04

基于 OpenMP 多核并行算法的垂线偏差快速计算

黄 炎[1] 王庆宾[1] 冯进凯[1] 谭勖立[1]

1 信息工程大学地理空间信息学院,郑州市科学大道 62 号,450001

摘 要:针对利用超高阶地球重力场模型计算大范围、高分辨率区域垂线偏差效率低的问题,提出基于 OpenMP 多核并行技术的数组升维和分区计算方法。实验表明,该方法计算垂线偏差的加速比最高达到 5.6 倍,显著提高了超高阶垂线偏差的计算效率,也为解决重力场数据处理过程中类似的快速计算问题提供了思路。

关键词:OpenMP;超高阶重力场模型;垂线偏差;并行计算;数组升维

中图分类号:P223 **文献标识码**:A

利用重力场位系数模型计算垂线偏差的耗时会随着截断阶次、计算范围、分辨率的增大而增加,难以满足一些工程的需求[1-4]。尽管利用交换积分次序等方法可以减少计算量从而提高计算效率,但随着计算区域和分辨率的增加,上述方法计算耗时显著增加,因为常规串行程序运行过程中仅使用单核计算,极大地浪费运算能力[5]。OpenMP 是一个并行运算库,它可用于共享内存并行系统多处理器程序设计编译处理[6],但 OpenMP 在对复杂程序进行整体并行时,常会造成内存读写冲突,导致并行化失败。本文提出利用 OpenMP 并行计算方法计算超高阶垂线偏差,并通过数组升维和分区计算的方法解决并行化过程中内存读写冲突的问题,能显著提升计算效率。

1 Belikov 递推法与交换积分次序法

1.1 缔合勒让德函数 Belikov 递推法

缔合勒让德函数 Belikov 递推法公式为:

$$\begin{cases}\widetilde{P}_{n0}(\cos\theta)=t\widetilde{P}_{n-1,0}(\cos\theta)-\dfrac{u}{2}\widetilde{P}_{n-1,1}(\cos\theta), & m=0\\ \widetilde{P}_{nm}(\cos\theta)=t\widetilde{P}_{n-1,m}(\cos\theta)-u\left[\dfrac{1}{4}\widetilde{P}_{n-1,m+1}(\cos\theta)-\widetilde{P}_{n-1,m-1}(\cos\theta)\right], & m>0\end{cases} \tag{1}$$

将 $\widetilde{P}_{nm}(\cos\theta)$ 转化为完全正常化勒让德函数:

$$\overline{P}_{nm}(\cos\theta)=\sqrt{2n+1}\widetilde{N}_{nm}\widetilde{P}_{nm}(\cos\theta) \tag{2}$$

其中,

$$\begin{cases}\widetilde{N}_{nm}=\sqrt{(2-\delta_0^m)\dfrac{(n+m)!}{2^m n!}\dfrac{(n-m)!}{2^m n!}}\\ \delta_0^m=\begin{cases}1, & m=0\\ 0, & m\neq 0\end{cases}\end{cases} \tag{3}$$

1.2 交换积分次序编程法

由位系数模型计算模型垂线偏差的计算公式为:

$$\begin{cases}\xi_M^{ij}=\sum\limits_{m=0}^{N}(xE_m^i\cos m\lambda_j+xF_m^i\sin m\lambda_j)\\ xE_m^i=-\sum\limits_{n=0}^{N}\left(\dfrac{R}{r_i}\right)^{n+2}\overline{C}_{nm}^{*}\dfrac{d}{\mathrm{d}\varphi}\overline{P}_{nm}(\sin\varphi_i)\\ xF_m^i=-\sum\limits_{n=0}^{N}\left(\dfrac{R}{r_i}\right)^{n+2}\overline{S}_{nm}\dfrac{d}{\mathrm{d}\varphi}\overline{P}_{nm}(\sin\varphi_i)\end{cases} \tag{4}$$

$$\begin{cases}\eta_M^{ij}=\sum\limits_{m=0}^{N}(yE_m^i\cos m\lambda_j+yF_m^i\sin m\lambda_j)\\ yE_m^i=-\sum\limits_{n=0}^{N}\left(\dfrac{R}{r_i}\right)^{n+2}m\overline{S}_{nm}\dfrac{\overline{P}_{nm}(\sin\varphi_i)}{\cos\varphi_i}\\ yF_m^i=\sum\limits_{n=0}^{N}\left(\dfrac{R}{r_i}\right)^{n+2}m\overline{C}_{nm}^{*}\dfrac{\overline{P}_{nm}(\sin\varphi_i)}{\cos\varphi_i}\end{cases} \tag{5}$$

式中,$(r_i,\lambda_i,\varphi_j)$ 为第 i 个纬度圈第 j 个网格中点处地心向径和经纬度(为简单起见,实验中可假定 $r_i=R$)。这样,同一纬度圈只需递推勒让德函数及其导数便可一次求得系数 xE_m^i、xF_m^i、yE_m^i 和 yF_m^i,从而加快其计算速度。而经度的正余弦函数值亦可递推求得。

2 OpenMP 多核并行算法

单核 CPU 运算时,串行程序依照编写顺序依次进行计算。多核 CPU 进行并行运算时,多

收稿日期:2018-11-07

第一作者简介:黄炎,硕士生,主要从事物理大地测量与并行计算等研究,E-mail:781367531@qq.com。

个CPU可同时对程序各进程进行运算，从而提高程序的运行效率。OpenMP支持Fortran、C/C^{++}的共享存储并行编程。它基于fork-join的并行执行模型，将程序主体划分为并行区和串行区，不同线程间可通过共享变量实现数据交换[6]。本文采用OpenMP多核并行算法，实现依托于超高阶重力场模型的高精度垂线偏差并行运算。

2.1 并行计算的需求分析

缔合勒让德函数值在同一纬度上相同，因此在计算格网模型垂线偏差时将纬度作为外层循环、经度作为内层循环来减少缔合勒让德函数的计算次数。考虑到缔合勒让德函数的求解涉及到递推算法，因此无法直接利用并行手段对该过程进行直接改化。本文主要考虑依据纬度的不同对相应格网点进行分组，进行多纬度同时计算以提高计算效率。

2.2 并行方案

使用同构并行面临两个问题。一是并行手段的选择问题。在共享存储器环境下，OpenMP多线程技术以其良好的简洁性和可移植性成为共享存储系统并行编程的工业标准[7]。OpenMP可直接在高级语言(如C语言)的串行代码上通过编写并行制导语句实现程序的并行化，而不需要作太大的修改。但如果需要并行化的程序比较复杂，则需要改变程序内部结构以防止并行后内存读写冲突，这一过程相对较为复杂。二是怎样更为有效地提高并行效率，这个问题将在下节进行讨论。

2.3 并行算法设计

本节主要分析讨论利用OpenMP技术进行并行设计的几个关键问题：一是如何将原来复杂的串行程序改为并行程序，避免内存读写冲突；二是如何更为有效地利用计算机的运算能力，提高并行效率。

2.3.1 *并行程序实现*

经过前面的分析可知，计算不同纬度圈上的格网点模型垂线偏差有着天然的可并行性，而缔合勒让德函数因其递推计算的性质不利于并行化，在并行过程中会造成数组和变量的内存读写冲突，即数组和变量在被不同线程调用读写的过程中会产生内存读写冲突导致程序中断，因此需要对它们进行一定的处理以消除内存读写冲突。

使用OpenMP中private数据子句解决简单变量的内存读写冲突问题。private子句可将简单变量声明为本线程的私有变量，每个线程都存有变量的副本，其他线程无法访问。即使在并行区域外有同名的共享变量，此共享变量也不会对并行区域造成影响，且并行区域内部计算也不会改变外部共享变量的值[7]。因此在并行起始语句中加入private(m，f，n，…)即可解决简单变量的内存读写冲突问题。

然而对于数组变量，尤其是二维数组变量，在普通便携式计算机上通过private子句进行变量私有会占用大量的内存空间，针对以上问题本文提出数组升维和区域分组方法。

数组升维是指将一维数组升至二维数组(如将$xE[f]$、$xF[f]$、$yE[f]$、$yF[f]$升维至$xE[v][f]$、$xF[v][f]$、$yE[v][f]$、$yF[v][f]$)；将二维数组升至三维数组(如将$P[e][f]$和$dP[e][f]$升维至$P[v][e][f]$和$dP[v][e][f]$)。这样，就可通过循环v实现不同线程在高维数组中调用不同的低维数组。当v等于网格行数时，就可通过一条OpenMP并行语句实现格网点模型垂线偏差的并行计算。

在计算截断阶数为j的超高阶垂线偏差时，如果令v=gw2(gw2为格网行数)，则需要开辟多个$j\times j\times$gw2的三维数组。假定截断阶数j=2 160，计算区域4°×8°(纬度×经度)，分辨率2′×2′，则gw2=120。一个double型数据需要占用8字节内存，则开辟一个double型三维数组需要占用4.17 G内存，这显然不适用于普通计算机，如何使用有限的内存来达到并行效果成为本文另一个研究重点。

为解决以上问题，以8线程为例，对计算格网进行区域分组，将格网以行为单位，8行为1组进行计算。在使用OpenMP进行并行计算时，在每组内进行并行计算，这样只需要开辟2 160×2 160×8的三维数组，此时内存占用仅为之前的1/15。

2.3.2 *并行程序改进*

OpenMP是以fork-join(分叉-合并)的形式来执行并行指令，其中fork代表开启新线程，join代表多个线程会合。fork-join模型在运行前只有一个线程存在，称之为“主线程”；在运行过程中，遇到并行开始代码，计算机则调用多个线程来执行并行任务；在并行代码执行结束后，调用的程被释放，不再工作[8](图1)。

为了减少线程调度的时间消耗，本文提出一种优化算法：当待计算格网行数为gw2、使用线程数为p时，将格网平均分为p组，若gw2与p之比不为整数，余数为q，则将余下q行分配至前q个线程，以保证线程利用率最高。

分别取7线程和8线程为例。首先要对计算

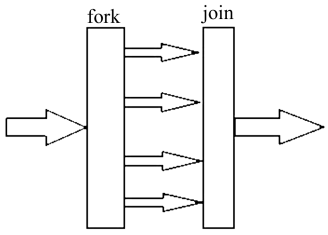


图 1　线程调度示意图[6]

Fig. 1　Thread scheduling diagram[6]

区域进行分区,以 4°×8°(分辨率为 2′×2′)范围格网为例,格网行数 gw2 为 120。7 线程时,将区域分为 7 组,第 1 组分配 18 行,后 6 组分配 17 行;8 线程时,将区域分为 8 组,每组分配 15 行,将每组的计算分配给独立的线程。

3　实验计算分析

为验证本文所提方法的有效性,利用便携式个人计算机,基于 VS2013 平台(C 语言)对上述方法进行实验。实验所用计算机处理器为 Intel ®Core™ i7-4700MQ CPU@ 2.40 GHz,内存为 8.00 GB,操作系统为 64 位 Windows 10 专业版操作系统。实验采用超高阶地球重力场模型 EIGEN6C4[4],拓展阶次为 2 160 阶,空间分辨率为 5′×5′。

3.1　并行程序改进前后对比分析

计算范围为 4°×8°(纬度×经度),分辨率分别取 2′×2′和 1′×1′,统计不同线程数下该区域格网模型垂线偏差的求解时间,计算耗时见表 1。

表 1　程序改进前后格网计算耗时

Tab. 1　Time-consuming grid computing before and after program improvement

线程数	2′×2′		1′×1′	
	改进前耗时/s	改进后耗时/s	改进前耗时/s	改进后耗时/s
1	54.13	54.28	122.79	122.28
2	29.94	28.10	62.06	61.60
3	22.08	20.25	49.23	44.42
4	16.83	15.80	37.10	34.38
5	16.69	13.62	37.85	29.77
6	16.62	12.25	36.29	26.08
7	15.15	11.95	34.75	25.28
8	12.82	10.92	27.26	22.92

由实验结果可知,多线程任务模式下,改进后的算法(多线程分区计算法)计算时间更短(在计算范围 4°×8°、分辨率 1′×1′时,改进后算法耗时减少 4.34 s),充分证明改进后的算法能减少 OpenMP 线程调度时间,提高代码运行效率。

因此在接下来的实验中,均采用改进后的分区计算方法进行任务测试。

3.2　不同线程数加速效果分析

利用上述实验结果,统计改进后程序耗时的加速比(不同线程数与串行程序效率比),结果见表 2 和图 2。

表 2　改进后程序计算的加速比

Tab. 2　Using the improved program to calculate the acceleration ratio

线程数	分辨率	
	2′×2′	1′×1′
1	1.000	1.000
2	1.932	1.985
3	2.680	2.753
4	3.436	3.557
5	3.985	4.108
6	4.432	4.689
7	4.541	4.836
8	4.971	5.335

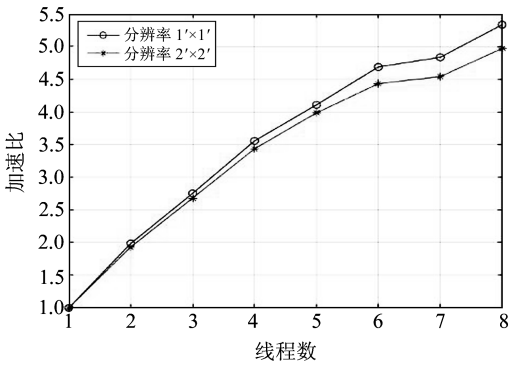


图 2　不同线程数下计算的加速比

Fig. 2　Acceleration ratios computied with different threads

由测试结果可知:

1)随着线程数的增加,计算的加速比在提高,加速比在 8 线程下可达到 5.3 倍。

2)加速比会小于对应的线程数(如 8 线程时加速比为 5.3)。因为在程序运算过程中,OpenMP 调度线程和同步各个线程计算过程都需要消耗一定的时间,改进后的程序尽管减少了线程调度时间,但没有减少同步线程计算的时间。

3)由图 2 可知,线程数为 7 时,加速比提高效果不明显,主要原因在于格网数不能整除线程数。按照上述线程分配原则,0 号线程分配 18 行格网、1～6 线程分配 17 行格网,计算耗时以任务量最大的 0 号线程计算结束为准,因此加速效果不明显。

4)格网分辨率提高,计算量变大,同等线程数下加速比会提高。

3.3　不同范围计算效果对比

在上述实验的基础上,基于 8 线程组织实验,实验范围分别为实验区 1(2°×4°)、实验区 2(4°×8°)、实验区 3(8°×16°),对应分辨率分别为 2′×2′和 1′×1′。统计不同线程数下各个实验区格网模型垂线偏差赋值时间,计算耗时和加速比(表 3)。

由测试结果可知,计算效率会随着计算范围

的增大而提高，计算范围 8°×16°、分辨率为 1′×1′的格网串行运算时间为 307 s，并行加速后计算耗时仅为 54 s，加速比达到 5.6 倍，计算效率获得大幅提升。

表 3 不同范围计算的耗时与加速比

Tab. 3 Time-consuming and acceleration ratio in different ranges

实验区	2′×2′			1′×1′		
	串行耗时/s	并行耗时/s	加速比	串行耗时/s	并行耗时/s	加速比
1	27.94	6.43	4.344	54.12	10.97	4.936
2	54.28	10.92	4.971	122.28	22.92	5.335
3	122.86	23.24	5.288	307.07	54.27	5.659

4 结 语

本文对计算超高阶垂线偏差的并行化问题进行探讨。根据实际计算需求和设备的特点，引入 OpenMP 并行算法。通过实例阐述并行化过程的关键技术，即数组升维和分区计算技术，并进行数值验证。结果表明，并行计算充分利用现有计算设备的计算能力，在不增加任何多余成本的基础上，提高了模型垂线偏差的解算速度，加速比最高可达 5.6 倍，为大范围乃至全球计算提供了保障。

本文提出的数组升维和分区计算方法，能有效解决 OpenMP 算法处理大型复杂程序整体并行过程中的内存读写冲突问题，并通过实验验证这一技术的可行性，为解决重力场数据计算和处理过程中各种复杂程序整体并行问题提供了解决思路，同时也是提高数据处理计算能力的有效途径之一。

参考文献

[1] 范昊鹏，李姗姗. 局部区域模型重力异常快速算法研究[J]. 大地测量与地球动力学，2013，33(6)：28-30(Fan Haopeng, Li Shanshan. Study on a Fast Algorithm for Model Gravity Anomalies in Local Areas[J]. Journal of Geodesy and Geodynamics, 2013, 33(6): 28-30)

[2] 曲政豪，曾添，周强. 超高阶地球重力场模型的高程异常优化算法[J]. 海洋测绘，2014，34(4)：5-8(Qu Zhenghao, Zeng Tian, Zhou Qiang. Optimization Algorithm for Height Anomaly of Ultra-High Degree and Order Earth Gravity Field Model[J]. Hydrographic Surveying and Charting, 2014, 34(4): 5-8)

[3] 章传银，郭春喜，陈俊勇，等. EGM2008 地球重力场模型在中国大陆适用性分析[J]. 测绘学报，2009，38(4)：283-298(Zhang Chuanyin, Guo Chunxi, Chen Junyong, et al. EGM 2008 and Its Application Analysis in Chinese Mainland[J]. Acta Geodaeticaet Cartographica Sinica, 2009, 38(4): 283-298)

[4] 欧阳明达，张敏利，于亮. 采用 Belikov 列推和跨阶次递推方法计算超高阶缔合勒让德函数[J]. 测绘工程，2017，26(7)：12-15(Ouyang Mingda, Zhang Minli, Yu Liang. Calculating the Ultra-High-Order Associated Legendre Functions by Belikov Column Method and Recursive Method Between Every Other Order and Degree[J]. Engineering of Surveying and Mapping, 2017, 26(7): 12-15)

[5] 黄佳喜. 扰动重力场快速(并行)计算方法研究[D]. 郑州：信息工程大学，2017(Huang Jiaxi. Research on Fast (Parallel) Computing of Disturbing Gravity[D]. Zhengzhou: Information Engineering University, 2017)

[6] 刘文志. 并行算法设计与性能优化[M]. 北京：机械工业出版社，2015(Liu Wenzhi. Parallel Computing and Performance Optimization[M]. Beijing: China Machine Press, 2015)

[7] 郁志辉. 高性能计算并行编程技术——MPI 并行程序设计[M]. 北京：清华大学出版社，2001(Yu Zhihui. High Performance Parallel Programming with MPI [M]. Beijing: Tsinghua University Press, 2001)

[8] 刘胜飞，张云泉，孙相征. 一种改进的 OpenMP 指导调度策略研究[J]. 计算机研究与发展，2010，47 (4)：687-694(Liu Shengfei, Zhang Yunquan, Sun Xiangzheng. An Improved Guided Loop Scheduling Algorithm for OpenMP[J]. Journal of Computer Research and Development, 2010, 47 (4): 687-694)

Fast Calculation of Plumb Line Deviation Based on OpenMP Multi-Core Parallel Algorithm

HUANG Yan[1] *WANG Qingbin*[1] *FENG Jinkai*[1] *TAN Xuli*[1]

1 School of Surveying and Mapping, Information Engineering University, 62 Kexue Road, Zhengzhou 450001, China

Abstract: In order to solve the problem of low efficiency in calculating vertical deviation in large-scale and high-resolution regions by using the super-high-order earth gravity field model, this paper presents an array dimension-lifting and zoning methods based on OpenMP multi-core parallel technology. Experiments show that the acceleration ratio of this method to calculate vertical deviation is up to 5.6 times, which greatly improves the calculation efficiency of super-high order vertical deviation, and provides ideas for solving similar fast calculation problems in data processing of gravity field.

Key words: OpenMP; ultra-high order gravity field model; plumb line deviation; parallel computation; array ascending dimension

About the first author: HUANG Yan, postgraduate, majors in physical geodesy and parallel compution, E-mail: 781367531@qq.com.