

基于集群和 GPGPU 技术的三维地形数据生成方法研究

李 波^{1,2,3,4,5}, 赵怀慈^{1,3,4}, 花海洋^{1,3,4}, 杜 梅^{1,2,3,4}

(1 中国科学院 沈阳自动化研究所, 辽宁 沈阳 110016; 2 中国科学院大学, 北京 100039;

3 中国科学院 光电信息处理重点实验室, 辽宁 沈阳 110016; 4 辽宁省图像理解与视觉计算重点实验室,
辽宁 沈阳 110016; 5 沈阳工程学院, 辽宁 沈阳 110016)

摘 要: 针对三维地形数据生成时间过长的问题, 提出了一种基于集群和 GPGPU 技术的 HPC(High Performance Computing)高性能计算方法, 能够有效整合集群中多台计算机的 CPU 和 GPU 的处理能力. 在三维地形数据生成过程中, 该方法将地形数据进行二次细化, 分别分配给 GPU 的每个任务管线, 大幅度提高运算吞吐量, 缩短三维地形数据的生成时间.

关键词: HPC; 三维地形; 集群; GPGPU

中图分类号: TP391.9

文献标识码: A

文章编号: 1000-7180(2014)10-0001-04

The 3D Terrain Data Generation Method Based on Computer Cluster and GPGPU

LI Bo^{1,2,3,4,5}, ZHAO Huai-ci^{1,3,4}, HUA Hai-yang^{1,3,4}, DU Mei^{1,2,3,4}

(1 Department of Optical-Electronics and Information Processing, Shenyang Institute of Automation,
Chinese Academy of Science, Shenyang 110016, China; 2 University of Chinese Academy of Science,
Beijing 100039, China; 3 Key Laboratory of Optical-Electronics Information Processing,
Chinese Academy of Science, Shenyang 110016, China;

4 Key Laboratory of Image Understanding and Computer Vision,

Shenyang 110016, China; 5 Shenyang Institute of Engineering, Shenyang 110016;)

Abstract: In order to solve the problem of huge time cost in the generating process of 3D terrain data, this paper presents a HPC(High Performance Computing) method based on GPGPU and computer cluster, which can integrate the CPUs and GPUs of the computer cluster together to provide huge computing ability. When the HPC method is used to generate 3D terrain data, the terrain is divided into pieces with two levels of details first, and then assigned to GPU pipelines. In this way, the computing throughput is greatly improved and plenty of time can be saved.

Key words: HPC; 3D terrain; computer cluster; GPGPU

1 引言

随着地理信息系统、虚拟现实、战场环境仿真等应用需求的增长, 三维地形生成算法研究成为近几年计算机图形学的一个重要方向, 地形生成算法^[1]在向数据巨大化、地形精细化方向发展. 面对越来越庞大的地形数据, 目前, 单台计算机所能提供的计算

量已经远远不能满足计算需求, 而对于普通单位和个人而言, 承受不起租用超级计算机的昂贵费用, 因此有必要寻求一种成本相对较低又能提供巨大运算能力的方法以满足大数据量的三维地形生成需要. 目前留下的方法是通过高性能计算方式提供大规模地形数据运算所需的计算能力^[2]. 本文在集群和 GPGPU 技术的基础上提出了针对大规模地形数据

收稿日期: 2013-12-25; **修回日期:** 2014-02-17

基金项目: 921 载人航天领域预先研究项目(030202); 辽宁省教育厅科研项目(L2011238)

生成的高性能计算方法,成功突破了目前单台计算机的计算瓶颈,解决了大规模地形数据生成过程中对计算能力的要求。

2 传统三维地形生成方法

2.1 传统三维地形生成方法的局限性

传统的三维地形数据生成方法主要分为人工生成和计算机自动生成两大类。人工方法主要指美工人员通过地形编辑软件来制作地形,这种方法生成的地形灵活精细,可展现任意场景,但是人力耗费惊人,成本过高并且速度很慢。而通过计算机生成地形的的方法不仅速度快,而且成本相对较低,目前主要有两种计算机生成地形方式。一是基于真实地形数据的地形生成方法。真实地形数据主要来自 DEM 数据^[3](Digital Elevation Model, 数字高程模型)和 DTM(Digital Terrain Model, 数字地面模型)。这种方法可以生成和真实地形数据高度接近的地形数据,但是地形数据的获取较难,对精度也有一定要求,而且数据量异常庞大。第二种方法是通过计算机直接生成地形数据。该方法通过在空间中生成三维坐标点来构建三角形,然后将三角形连接成三角面,由三角面组成三维地形的空间框架,最后在框架上绘制上纹理形成地形。对于较小的数据量,这种方法灵活方便,可实时生成并显示地形数据;但是对于数据量较大的复杂场景,该方法所需要的数据生成时间大幅度增长,严重影响了其实用性。针对这个问题,应用了本文提出的三维地形数据生成框架,大幅度缩短了三维地形数据生成所需的时间。

2.2 三维地形构建原理

计算机中显示的三维地形实质上是一系列的三角网格上绘制上纹理数据而形成的。因此三维地形的生成时间主要受三角网格的生成时间和纹理的绘制时间影响。纹理的绘制时间主要取决于显卡的性能,其可人为优化的空间有限,因此提高地形数据生成效率的关键在于缩短三角网格的生成时间。

三角网格由众多的三角面组成,三角网格是三维地形的框架,其精细度和形状直接决定最终的三维地形的精细度和形状。三角面是由一系列的空间中的三维点通过相互连接而形成。目前构建三角网格的算法有很多,考虑到计算机生成的地形数据的顺序性和可扩展性,选用了 Delaunay 三角化算法^[4]

构建三维网格。相比于其他三角化算法, Delaunay 算法更适合通过逐点插入空间点的方式来构建空间三角面。

3 改进的三维地形生成方法

生成大规模三维地形数据需要巨大的运算量。为缩短运算时间,提高计算效率,需要通过高性能计算技术来提供超越单台计算机的运算量。

3.1 高性能计算平台的构建

通过对几种高性能计算技术进行比较,本课题最终采用集群技术来构建高性能计算框架,以 MPICH2 作为实现高性能计算的工具,并辅以 GPGPU 技术进一步提高并行运算吞吐能力。

计算机集群^[5]是高性能计算的一种常见组成方式,由多台计算机联合组成。其逻辑结构简单,价格相对低廉,具有极高的可操作性。集群可由本地的任意多台计算机通过网络连接组成,一般为星形结构,集群中的各计算机处于平等地位,通过 HPC 开发包进行任务分配和协作,进而提供超强的计算能力。相比于其他 HPC 技术,集群以相对较低的成本获得计算能力大幅度的提升,使高性能计算走向了平民化。

MPICH2 是 MPI(Message Passing Interface)的具体实现,简单易用,具有良好的可移植性,是目前最流行的 HPC 开发工具。

GPGPU 技术^[6]是近几年发展起来的通过显卡进行通用计算的新技术。进而根据 GPGPU 技术,可以通过编程为 GPU 拥有几十条到几百条管线单独分配运算任务,大幅度提高运算并行性,能够将 CPU 从高密度运算中解放出来,专注于流程控制。基于 CPU/GPU 组合的运算方式目前已经成熟^[7],并且有相对成熟的策略调配方法^[8]。大量实验数据表明,显卡参与通用运算后可以提高计算机几倍到几十倍的运算能力。本文选用 CUDA 作为 GPGPU 的实现技术。CUDA 是目前最为流行的 GPGPU 实现技术,借助于 CUDA^[9],可以使 GPU 直接参与生成地形数据的运算工作,进一步提高了该 HPC 应用程序框架的性能。

3.2 总体框架设计

用于三维地形数据生成的应用程序框架^[10]如图 1 所示,该框架中的计算机个数可随意扩张。框架

中有服务器和计算节点两种运算角色. 计算机节点仅提供运算能力,而服务器除参与运算外,还承担确定地形范围、对任务进行划分和分配、回收计算节点的运算结果等任务.

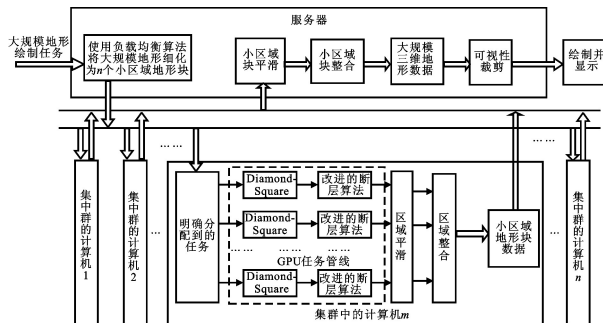


图1 应用程序框架

基于集群和 GPGPU 技术的三维地形数据生成方法描述如下:

(1)指定集群中的某台计算机作为服务器. 由于服务器负责进行任务分配和最终的图像绘制,因此需要计算能力和绘图能力较强的计算机,这里使用图形工作站作为服务器.

(2)服务器对整个工程的总工作量进行预估,根据集群中计算节点的个数将整个地形划分为多个相对小规模的区域,然后分配给集群中的各个计算节点,使各节点的负载均衡.

(3)集群中的计算节点接收到分配的任务后,由 CPU 根据 GPU 的管线情况进行任务二次划分,然后交于 GPU 的多个任务管线并行完成. 这样每个管线仅负责一小部分的地形数据的生成工作.

(4)在 GPU 的任务管线中,首先通过 Diamond-Square 算法在二维平面生成网格点数据,然后使用断层算法生成每个二维网格点的高度值,这样就生成了三维空间内的一系列点. 最后使用三角化算法将相关空间点连接在一起形成 Delaunay 三角化网格. 与此同时,由计算节点的 CPU 对各个任务管线中的小区域交界数据进行控制,以防止发生断层现象.

(5)各 GPU 管线完成各自的任务后均可形成一个小区的地形块数据,由 CPU 依据相邻地形块的信息对各个管线生成的地形区域进行边界平滑,然后组合成一个完整的区域地形数据. 至此,各计算节点完成自己的运算任务.

(6)计算节点将区域地形数据汇集到服务器进

行数据整合,再次平滑处理后,形成完整的三维地形数据. 至此,三维地形的框架数据生成完成.

(7)由服务器为地形数据补充纹理信息,然后将地形数据存储到文件中,并绘制到屏幕上.

4 实验验证

为验证应用程序框架的性能,使用 VS2010 + MPI + CUDA 对图 2 所示应用程序框架进行了实现,并进行了实验验证.



(a)1个节点运行效果 (b)2个节点运行效果 (c)3个节点运行效果

图2 第二类任务(1)实验运行效果

服务器选用惠普 Z600 图形工作站, @ 2. 13 GHz 四核 Intel Xeon E5506 CPU, 8. 0 GB 内存, 配备 NVIDIA Quadro FX580 显卡. 集群中的普通计算机节点采用 Intel Core i3-2120 (@ 3. 30 GHz 双核) CPU, 4. 0 GB 内存, 配备 GTX 650 显卡. 服务器和计算机节点的操作系统都为 Windows 7 64 位旗舰版.

通过两类任务对该应用程序框架进行验证: 第一类任务为同一负荷的任务交由不同数量的计算节点进行完成, 通过任务完成时间的差异来体现该框架的作用; 第二类任务为为每个计算节点分配相同负荷的计算任务, 从总计算负荷与完成时间的关系来体现该框架的作.

4.1 第一类任务实验

首先安排的任务为在 $3\text{ km} \times 3\text{ km}$ 范围内生成由 3 万个空间点形成的三维地形数据, 分别由一个、两个和三个计算节点来完成相关运算, 同时记录其运算耗费时间, 并计算节省的运算时间. 节省的运算时间由最终花费时间与由服务器单机完成任务所花费时间比较得来.

由实验结果可见, 本文所提出的高性能运算框架作用明显. 随着参与运算的计算节点数的增加, 生成定量地形数据所需的时间明显降低. 而且该框架的作用随着运算数据量的增大, 效果越来越明显. 当运算任务为 3 万个空间点数据时, 三个计算节点合作完成该任务所需的运算时间可以比一个节点所需

时间减少 43.2%;当运算任务为 30 万个空间点数据时,三个计算节点可以比一个运算节点节省 65.0%的运算时间.这一趋势随着计算节点的增加和运算量的增加继续增长.

4.2 第二类任务实验

首先为每个集群中的计算节点安排的计算任务为在 $3\text{ km} \times 1\text{ km}$ 范围内生成 1 万个空间点的数据,数据生成后发送到服务器,由服务器整合并绘制出来.运行效果如图 3 所示.

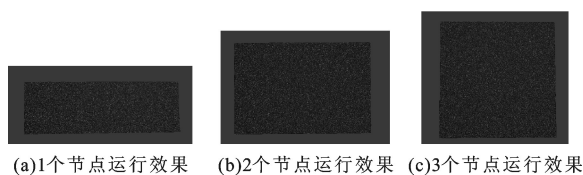


图 3 第二类任务(2)实验运行效果

由实验结果可见,在保证地形数据精度的前提下,随着计算节点的增加,整个应用程序所承担的计算负荷成倍增大,其所生成的地形范围也在不断增大,而相应增加的时间损耗却没有成比例增大.而且随着计算负荷的增大,整个应用程序所增加的时间损耗的幅度却越来越小.例如,当每个运算节点承担 1 万个空间点数据的计算负荷时,当负荷增加一倍,增加 33.5%的运算时间;当负荷增大两倍,仅增加 48.5%的运算时间.而对于每个节点 10 万个网格数据的运算量,当运算量增加一倍时,增加的时间损耗为 16.3%;当运算量增加两倍时,所增加的时间损耗仅为 26.3%.

5 结束语

由实验结果可见,本文提出的高性能计算方法能够有效发挥集群和 GPU 的并行计算能力,突破单计算机运算瓶颈,大幅度降低大规模三维地形数据生成所需的运算时间;并且随着计算量的增加,其实验效果愈加明显.该数据生成方法具有通用性,可推广至众多高密度计算领域,具有广阔的应用前景.

此外,本文提出的应用程序框架还存在着可进

一步优化的空间,例如 GPU 调度算法优化、网络传输稳定性优化、区域平滑算法优化等.

参考文献:

- [1] 许妙忠.大规模地形实时绘制地算法研究[J].武汉大学学报,2005,30(5):392-395.
- [2] 葛澎.分布式计算技术综述[J].微电子学与计算机,2012,29(5):201-203.
- [3] 毕华兴,谭秀英,李笑吟.基于 DEM 的数字地形分析[J].北京林业大学学报,2005,27(2):49-53.
- [4] 徐永安,杨钦,吴壮志,等.三维约束 Delaunay 三角化的实现[J].软件学报,2001,12(1):103-109.
- [5] Li Bo, Zhou Zhenliu, Wang Xiangfeng. A survey of HPC development[C]//2012 International Conference on Computer Science and Electronics Engineering, China, Hangzhou,2012:103-106.
- [6] 韩博,周秉锋.GPGPU 性能模型及应用实例分析[J].计算机辅助设计与图形学学报,2009,21(9):1219-1226.
- [7] 刘钢峰.基于 CPU/GPU 集群的编程的研究[J].微电子学与计算机,2013,30(2):128-131.
- [8] 袁壮,陈庆奎,曹剑玮.CPU 和 GPU 混合集群的负载均衡策略[J].微电子学与计算机,2012,29(9):47-50.
- [9] Jason Sanders, Edward Kandrot. CUDA by example: an introduction to general-purpose GPU Programming [M]. USA: Addison-Wesley Professional,2007.
- [10] 李波,王祥凤,曹福毅.大规模地形数据生成方法研究[J].沈阳工程学院:自然科学版,2013,9(3):263-265.

作者简介:

李波 男,(1980-),博士研究生.研究方向为虚拟现实与仿真技术、高性能计算及其可视化.

赵怀慈 男,(1974-),博士,研究员,博士生导师.研究方向为虚拟现实与仿真技术.

花海洋 男,(1978-),硕士,副研究员.研究方向为虚拟现实与仿真技术.

杜梅 女,(1977-),博士研究生.研究方向为通信技术.