



GPU-CA 模型及大尺度土地利用变化模拟

李丹^①, 黎夏^{①*}, 刘小平^①, 陈逸敏^①, 李少英^①, 刘凯^②, 乔纪纲^③, 郑奕钟^①, 张亦汉^①, 劳春华^①

① 中山大学地理科学与规划学院, 广东省城市化与地理环境空间模拟重点实验室, 广州 510275;

② 广州地理研究所, 广州 510070;

③ 广东商学院资源与环境学院, 广州 501320

* 联系人, E-mail: lixia@mail.sysu.edu.cn

2011-12-05 收稿, 2012-01-16 接受

国家自然科学基金重点项目(40830532)和国家自然科学基金(40901187)资助

摘要 在全球变化研究中, 对省和国家尺度的土地利用变化过程进行模拟具有重要的意义. 但大尺度的土地利用变化模拟存在数据量大、计算复杂、计算时间长等难点, 由于计算能力和传统基于 CPU 计算模式的制约, 在普通计算机中难以执行这样的模拟. 把最新发展的 GPU 高性能计算技术与元胞自动机(CA)模型相结合, 提出了 GPU-CA 大尺度土地利用变化模拟模型, 并将其应用到省级尺度的土地利用变化模拟中. 介绍了模型的问题映射、需解决的关键技术问题和计算的流程, 并分析了 GPU-CA 模型在计算性能上的优势. 通过全广东省的土地利用变化模拟, 对 GPU-CA 模型进行了验证. 实验表明, GPU-CA 模型可以将原有一般 CA 模型的运行效率提高 30 倍以上, 能够有效地应用于省和国家级的土地利用变化模拟, 为全球变化模型提供重要的地表景观变化信息.

关键词

GPU

元胞自动机

土地利用变化模拟

大尺度

全球变化

近年来, 许多学者利用元胞自动机(cellular automata, CA)开展了城市扩展、疾病扩散、火灾蔓延、人口迁移、经济发展、沙漠化、洪水淹没等格局和过程变化的模拟, 取得了卓有成效的研究成果^[1,2]. Batty 和 Xie^[3]模拟了美国纽约州 Buffalo 地区的城市扩张, Clarke 等人^[4]利用 SLEUTH 模型模拟了美国旧金山湾区的城市发展过程, Wu 和 Webster^[5]模拟了广州市的城市用地变化, 黎夏和刘小平^[6]也模拟了珠江三角洲地区的城市扩张. 黎夏等人^[7,8]还提出了地理模拟与优化系统(GeoSOS)理论, 将土地利用变化模拟与空间优化耦合起来, 为复杂资源环境下的土地利用变化模拟和优化提供了强有力的过程分析工具.

但在进行更大尺度(如省级行政区、全国尺度), 特别是对全球环境变化起关键作用的全球尺度土地利用变化过程研究时, 迫切需要进行大尺度的土地利用变化模拟. 大尺度土地利用变化模拟通过仿真模型模拟和再现了地球表层真实的土地利用变化格

局和过程, 对全球气候变化、粮食安全、碳循环的研究都具有重要的意义.

进行大尺度的土地利用变化模拟不仅需要了解和掌握大尺度区域内的土地利用变化规律, 而且需要有合适的数据、实验方法和技术条件支持. 通常选取栅格类型的空间数据进行模拟实验, 此时确定采用何种空间精度的数据将是非常重要的. 如果使用低空间分辨率的数据, 其单个像元包含了多种土地利用类型, 多用百分比来表示类型间的组成比例^[9], 较低的空间清晰程度必然影响模型分析的效果. 而高分辨率数据的每个像元可以代表其所包含空间范围的主导土地利用类型. 同时, 低分辨率数据会丧失局部的地理格局信息和非线性特征, 需要使用高空间分辨率的数据来保证微观尺度的相关信息不会被遗漏. 元胞自动机(CA)模型使用自下而上的方式来研究复杂系统的涌现行为, 在较细粒度的解析水平下才能更好地体现土地利用变化过程的真实规律.

英文版见: Li D, Li X, Liu X P, et al. GPU-CA model and large-scale land-usechange simulation. Chin Sci Bull, 2012, 57, doi: 10.1007/s11434-012-5085-3

因此,在大尺度土地利用变化模拟中使用高空间分辨率的数据是非常必要的。

然而使用高分辨率数据进行大尺度土地利用变化模拟,其数据量非常大、运算过程复杂、运算时间超长,如所涉及空间数据的存储大小通常达到数百 M 乃至数 G,普通计算机及基于 CPU 的串行计算模式受计算能力的制约几乎无法进行相关的模拟实验。因此有必要将并行计算模式引入到土地利用变化模拟中以解决计算能力的限制。目前已有一些基于并行计算的土地利用变化模拟研究,如基于负载均衡的并行元胞自动机模拟^[10],基于网格计算的元胞自动机模拟等,但以上方式存在着计算成本高、部署环境复杂、计算加速性能不高等缺点,需要寻找一种低成本、配置简单、加速性能好的并行计算模式来解决进行大尺度土地利用变化模拟时遇到的计算瓶颈问题。

为此,本文提出把 GPU 高性能计算技术和 CA 模型结合起来进行大尺度土地利用变化模拟的新方法。GPU 计算技术是使用计算机显卡的图形处理单元(graphic processing unit, GPU)进行通用计算的高性能计算技术,它是近年来发展的一种典型的并行计算模式,具有低成本、高度并行化、可编程、伸缩性好等优点^[11],当前“CPU + GPU”的计算模式已成为今后高性能计算发展的趋势^[12]。NVIDIA 公司于 2007 年推出 CUDA 计算平台^[13],提供了一种基于 GPU,可以使用类 C 语言开发的通用计算环境和软件体系,能够在任何支持 CUDA 的计算机显卡上执行通用计算任务,其计算成本较低。目前基于 GPU 的高性能计算已被广泛地应用于物理模拟^[14]、图像处理^[15]、三维地形生成、信号处理、人工智能等领域,极大地提高了原本基于 CPU 环境程序的运行效率。由于地理问题十分复杂,因此将 GPU 计算技术应用于大尺度、高精度的土地利用变化模拟等地理过程研究领域将具有很大的意义,国外有部分 GPU 计算技术与 CA 结合的研究^[16,17],但都没有应用在地球科学领域,国内目前也没有相关研究的报道。

1 GPU-CA 计算模型框架

GPU-CA 计算模型是基于 GPU 通用计算技术实现的元胞自动机(CA)土地利用变化模型,利用 GPU 计算技术的高性能计算属性来加速 CA 模拟的效率,从而解决在进行大尺度土地利用变化模拟时遇到的

计算瓶颈问题。其计算模型框架包括模型问题映射、需解决的关键技术问题及计算流程设计等三方面。GPU-CA 模型可以采用逻辑回归 CA^[18]、神经网络 CA^[2]、决策树 CA^[19]等成熟的 CA 土地利用变化模型。本研究主要关注研究区域内城市用地与非城市用地的转化过程,因此采用基于逻辑回归 CA 模型的数据并行实现方法。

1.1 模型问题映射

基于逻辑回归 CA 的城市土地利用变化模型原理如下:

根据逻辑回归模型计算的元胞土地利用类型转变概率为

$$P_g(S_{ij} = \text{urban}) = \frac{\exp(z)}{1 + \exp(z)} = \frac{1}{1 + \exp(-z)}, \quad (1)$$

$P_g(S_{ij} = \text{urban})$ 表示栅格空间中位于 ij 位置的元胞转变为城市用地状态的概率,取值范围为 0~1。 z 描述了影响土地利用变化的各类因子所产生的影响,其构成如下:

$$z = \alpha + \sum_k b_k x_k, \quad (2)$$

其中 α 为回归常数, b_k 为逻辑回归模型的系数, x_k 代表了 ij 位置元胞第 k 个影响因子的值。

逻辑回归 CA 模型中 t 时刻 ij 位置元胞的土地利用类型转变概率为

$$p'_{ij} = (1 + (-\ln \gamma)^\alpha) \times \frac{1}{1 + \exp(-z'_{ij})} \times \text{con}(s'_{ij}) \times \Omega'_{ij}, \quad (3)$$

$(1 + (-\ln \gamma)^\alpha)$ 为随机干扰项,代表土地利用变化中的随机影响。其中 γ 为 0~1 之间的随机数, α 的取值为 1~10 之间的整数。 $\text{con}(s'_{ij})$ 为 ij 位置元胞发生土地利用类型转变的约束条件。 Ω'_{ij} 则表示 CA 模型中的邻域影响,代表邻域窗口中的土地利用情况。

最终元胞是否发生土地利用类型转变则由下面的条件决定

$$\begin{cases} p'_{ij} \geq p_{\text{threshold}} & \text{转变为城市用地,} \\ p'_{ij} < p_{\text{threshold}} & \text{不转变为城市用地,} \end{cases} \quad (4)$$

如果 p'_{ij} 的值大于或等于预设的土地利用转变阈值 $p_{\text{threshold}}$,且 ij 位置元胞的土地利用类型为非城市用地,则该元胞的土地利用类型转变为城市用地,否则不发生转变。当元胞空间中的所有元胞都进行相应计算后,一次迭代计算过程结束。通过多次迭代计算过

程,完成整个土地利用变化过程的模拟。

GPU-CA 计算模型将上述计算问题映射到 GPU 编程模型和存储模型上,其模型的结构如图 1 所示。

GPU 通用计算在逻辑上划分为 3 个计算层次,其中最小的计算单元为线程(Thread),多个线程可以组成一个线程块(Block),一个线程块拥有一块共享的存储单元(Shared memory)。一个或多个线程块可以组成一个计算网格(Grid)。

GPU-CA 计算模型采用数据并行的计算方式,将一个线程映射为元胞自动机中的一个元胞,负责该元胞的计算操作。显卡从计算机主机端(Host)将模拟数据、相关的空间变量数据等信息读入到全局存储单元(Global memory)中,随后各种数据被分块拷贝到每个线程块的共享存储单元中,由每个线程根据索引读取其对应元胞的数据。由于元胞自动机模型需要获取元胞的邻域信息,因此每个线程单元同时读取其对应元胞的邻域数据,这里采用 3×3 的 Moore 邻域。当数据读取完成后,多个线程单元通过并行计算模式完成一次模拟迭代计算,同时将结果数据写回全局存储单元中的相应存储区域,供下一次运算使用。

1.2 GPU-CA 模型需解决的关键技术问题

(i) 数据映射。元胞自动机模型采用栅格数据结构,与 GPU 的线程模型有天然的相似性,因此 GPU-CA 计算模型将空间数据单元映射到 GPU 的线程计算单元上,采用数据并行的模式进行计算。这时

需要考虑数据索引映射方式、GPU 运算单元数量和存储单元大小等相关因素。索引映射方式可以采用一维数组或二维数组等方式,将栅格数据的位置索引信息与线程模型的位置索引信息关联,保证映射关系的正确性,两种索引映射方式分别如图 2(a)和(b)所示。其中 T1, T2, ..., Tn 分别代表 Block 中的 Thread。从图上可以看出,一维数组索引映射方式具有关系简单、映射直接的特点,二维数组索引映射方式则具有表达直观的优点,但其索引映射关系较一维数组复杂。本研究采用了二维数组索引映射方式。同时, GPU 运算单元的数量和存储单元的大小决定了 GPU 内核函数(Kernel)每次运算可以处理的数据量大小,与显卡的硬件属性相关。如果每次运算可以处理的数据量小于需要计算的总数据量,则不能一次将全部数据读入显卡显存内,需要多次分块读取相关数据,并多次运行内核函数进行计算。

(ii) 数据竞写。在 CA 模拟中,常常需要掌握一些全局信息,如每次迭代后发生状态变化的元胞总数。这类数据的特点是与每个元胞的信息都相关,每个元胞都有可能改变该类数据的值。如图 3 所示,在一次计算迭代中每个 Block 内的每个 Thread 都可能改写 Global memory 中某一物理地址所代表的全局属性信息的值。在串行计算模式中,每个元胞可以依次改变全局变量的值,相互之间不会发生冲突。而在并行计算环境下,如果每个计算单元同时执行数据修改操作,则会造成竞写问题。如果发生竞写现象,将不能保证全局信息的正确性。因此需要引入 CUDA 环境中的原子操作函数来防止竞写现象的发生。原子操作函数所提供的机制能够在多个线程同时向一个存储地址写入数值时保证该数值的正确性,从而有效地避免了竞写现象的发生,解决了大尺度

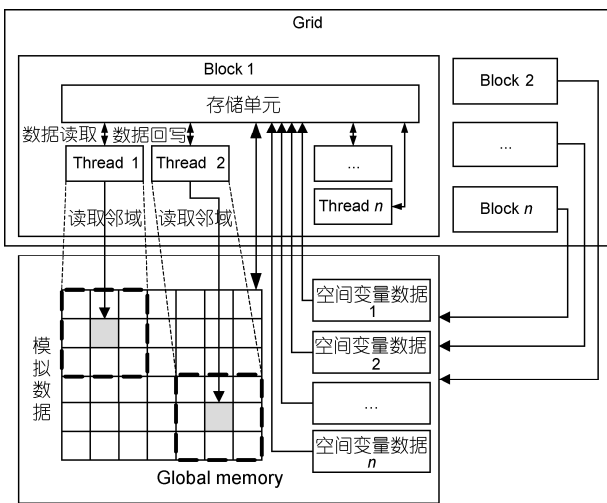


图 1 GPU-CA 计算模型结构图

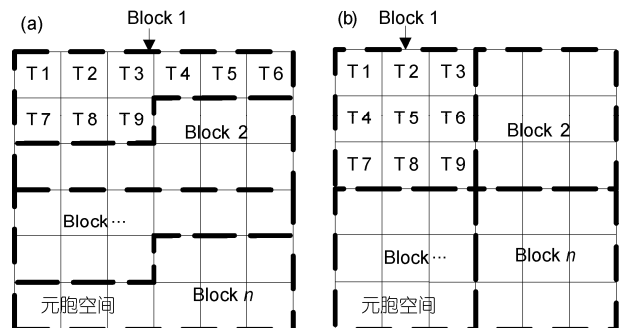


图 2 数据索引映射方式

(a) 一维数据索引映射方式; (b) 二维数据索引映射方式

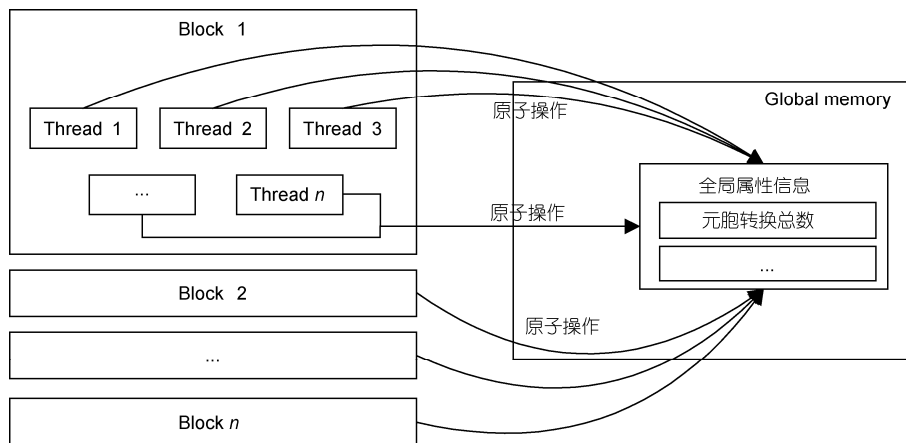


图 3 GPU-CA 模型所解决的数据竞写现象

土地利用变化模拟中遇到的相关问题。

1.3 GPU-CA 模型计算流程

GPU 通用计算的流程一般包括初始化计算环境, 数据传输, 使用内核函数并行计算, 退出计算环境等步骤。下面是 GPU-CA 模型计算流程的伪代码:

1. 定义计算所使用的 Grid 和 Block 大小及相应的存储空间大小。
2. 启动 CUDA 环境。
3. 读取数据到 Host 端内存中。
4. 在 GPU 端分配显存并将数据从 Host 端内存拷贝到显存中。

5. 使用内核函数计算整个元胞空间的 P_g 值。

6. 使用内核函数进行多次模拟迭代运算。

While (尚未达到模拟终止的条件)

6.1 对于每个 Block, 根据索引将其对应的所有元胞的土地利用类型代码值和 P_g 值从 Global Memory 读入 Shared Memory 中。

6.2 对于每个 Thread, 根据索引从 Global Memory 中读取其所对应元胞 3×3 邻域的数据。

6.3 对于每个 Thread, 根据计算公式计算其概率值。

If(达到转变成城市用地的条件)

转变成城市用地

End if

6.4 对于每个 Thread, 如果所对应元胞的状态转变为城市用地, 则对该 Thread 使用原子操作更新本次迭代所转换的城市用地总量。

6.5 对于每个 Block, 将所对应元胞的土地利用类型代码值从 Shared Memory 拷贝到 Global Memory 中。

End While

7. 对模拟结果进行模拟精度统计。

8. 将模拟结果从 GPU 显存拷贝到 Host 端内存中, 并输出模拟结果文件。

9. 释放 GPU 端和 Host 端的存储空间, 退出 CUDA 环境。

2 实例应用与分析

研究选取广东省为研究区域, 陆地面积为 $17.98 \times 10^4 \text{ km}^2$ 。分别获取了广东省 2000、2005 和 2006 年的土地利用分类数据和相关空间变量数据。采用 ArcGIS ASCII 格式的栅格数据文件, 空间分辨率为 80 m, 栅格大小为 9792×7376 , 单个数据文件的大小约为 500 M。土地利用分类数据包括 6 种用地类型, 分别是耕地、林地、草地、水域、城乡建设用地和未利用地。空间变量包括各栅格单元到城镇中心的距离, 到铁路的距离, 到高速公路的距离和到公路的距离。研究采用逻辑回归 CA 模型, 首先使用 2000 和 2005 年的土地利用分类数据和空间变量数据获取了广东省全省的城乡建设用地发展规律, 然后使用 2005 年土地利用数据作为初始数据, 模拟了 2005~2006 年间的城乡建设用地扩展情况。得到模拟结果后, 使用 2006 年的土地利用数据通过肉眼比较和点对点比较的方式来验证模拟结果的有效性^[1]。同时利用地理模拟优化系统软件(GeoSOS, <http://www.geosimulation.cn/>)所提供的一般 CA 进行了同样的模拟, 以分析 GPU-CA 对计算效率的改善程度。研究随后还进行了广东省 2010 和 2015 年城乡建设用地变化情况的预测。图 4 是 GPU-CA 模拟结果图。

为了与基于 CPU 串行计算模式的一般 CA 进行对比, 本研究采用了配置较高的硬件运行平台, 具体为 $2 \times \text{Intel Xeon E5620 } 2.5 \text{ GHz CPU}$, 24 G 主机内存,

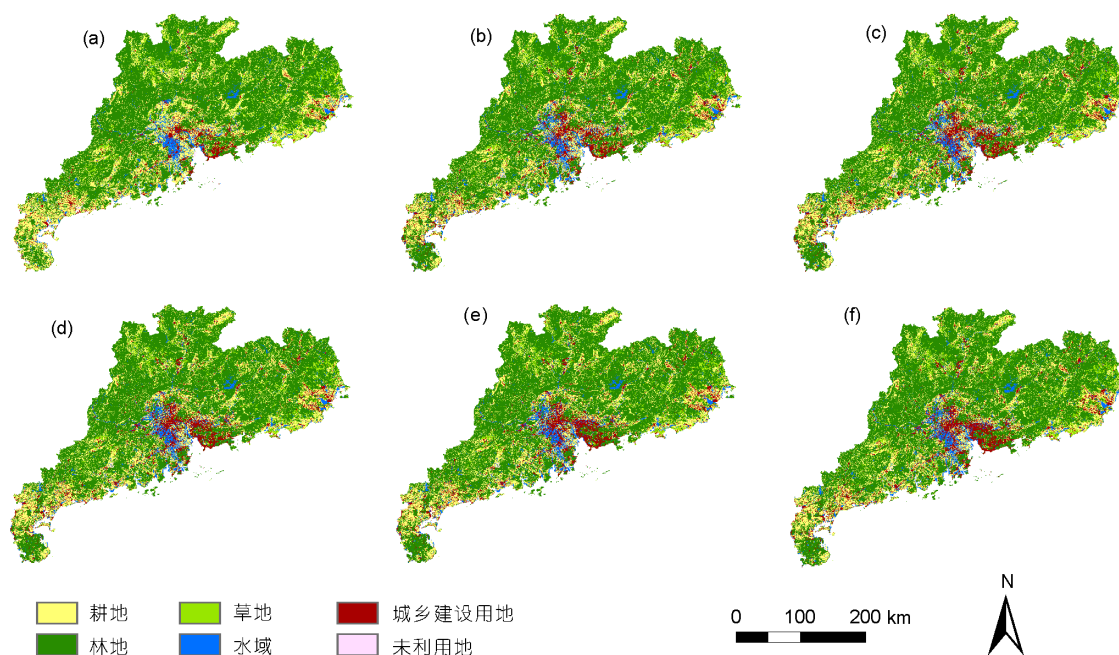


图4 GPU-CA 模拟结果图

(a)~(c)分别列出了广东省 2000, 2005 和 2006 年的土地利用现状图; (d)为使用从 2000 和 2005 年数据提取的土地利用转换规则由 2005 年模拟到 2006 年的结果; (e)和(f)分别是由 2005 年模拟到 2010 和 2015 年的结果

显卡采用 NVIDIA 公司的 Tesla 1060C, 显存为 4 G, 核心频率 1.30 GHz, 操作系统为 Windows 7 64 位专业版. GPU 计算环境为 CUDA 3.2, 开发环境为 Microsoft Visual C++ 2008.

研究通过图像一致性比较中常用的点对点比较方式得到混淆矩阵, 然后进行模拟精度验证表明, GPU-CA 模拟的总体精度为 82.9%, 说明 GPU-CA 的模拟结果是有效的, GPU-CA 模型可以应用于大尺度的土地利用变化模拟研究.

研究根据实验数据和模拟及预测结果, 分别统计了 2000~2006, 2000~2010 和 2000~2015 年间广东省各市级行政区城乡建设用地占该市土地面积的百分比以及该市城乡建设用地的扩张情况, 如表 1 所示. 从表中可以看出, 广东省各市级行政区在这些时期内都经历了城乡建设用地的扩张过程, 2000 年城乡建设用地占广东省土地总面积的 4.57%, 2006 年该比例为 7.39%, 预计 2010 年该比例将达到 7.70%, 2015 年将达到 7.89%. 2000~2006 年城乡建设用地的扩张比例为 61.76%, 预计 2010 年该比例为 68.60%, 2015 年为 72.55%. 可见整个广东省有大量土地转变为城乡建设用地, 土地资源紧张的局面将持续出现. 但各市级行政区的扩张速度并不相同, 区域间差距较大.

根据 2000~2015 年间各市级行政区城乡建设用地面积占总面积比例的不同情况, 可以将广东省的市级行政区分为 3 种类型, 如图 5 所示. 其中广州市、深圳市、东莞市、佛山市、中山市和珠海市, 其城乡建设用地占比均超过了 20%. 江门市、阳江市、茂名市、湛江市、惠州市、汕头市、汕尾市、潮州市和揭阳市, 其城乡建设用地占比在 5%~20%之间. 韶关市、肇庆市、清远市、云浮市、河源市和梅州市, 其城乡建设用地占比均小于 5%. 可以看出城乡建设用地占比超过 20%的 6 个城市均分布在珠江三角洲地区, 随着 2000~2006 年间区域经济的快速增长, 有大量的土地转变为城乡建设用地. 特别是东莞市, 城乡建设用地占比在 2006 年已经达到 43.63%, 预计在 2010 年达到 51.63%, 在 2015 年达到 52.95%, 如图 6 所示. 城乡建设用地占比在 5%~20%之间的城市分布在广东省的沿海地区, 受沿海开放政策和区位的影响, 经济也得到了较快发展, 需要一定的城乡建设用地来满足发展的需要. 而城乡建设用地占比小于 5%城市主要分布在粤西北、粤东北山区和原经济不发达地区, 受经济条件制约, 城乡建设用地占比较小.

同时研究还根据 2000~2015 年间城乡建设用地扩张速度的不同将各市级行政区划分为 3 种类型, 如

表 1 基于 GPU-CA 的广东省各市级行政区 2000~2015 年城乡建设用地面积(km²)统计及城乡建设用地占比和扩张情况

市级行政区	土地面积	2000 年	2005 年	2006 年	2010 年	2015 年	2006 年城乡 建设用地 占比(%)	2010 年城乡 建设用地 占比(%)	2015 年城乡 建设用地 占比(%)	2000~2006 年 城乡建设用地 扩张比例(%)	2000~2010 年 城乡建设用地 扩张比例(%)	2000~2015 年 城乡建设用地 扩张比例(%)
广州市	7226.27	844.04	1328.59	1374.23	1373.40	1447.62	19.02	19.62	20.03	62.82	67.97	71.51
深圳市	1863.31	656.09	799.53	842.96	843.62	876.98	45.24	46.43	47.07	28.48	31.87	33.67
珠海市	1449.95	180.24	261.00	283.33	283.48	295.94	19.54	19.95	20.41	57.19	60.48	64.19
汕头市	2152.27	361.65	390.48	411.41	410.77	427.90	19.12	19.44	19.88	13.76	15.67	18.32
佛山市	3845.65	556.83	929.53	1085.32	1085.50	1147.26	28.22	28.87	29.83	94.91	99.42	106.04
韶关市	18439.42	225.71	343.51	362.85	362.75	376.33	1.97	2.01	2.04	60.76	64.26	66.73
河源市	15659.39	73.29	285.03	318.22	318.97	339.50	2.03	2.09	2.17	334.22	345.90	363.25
梅州市	15903.24	164.61	435.06	461.54	461.56	491.22	2.90	3.01	3.09	180.38	191.17	198.41
惠州市	11312.79	520.26	746.44	802.98	804.15	868.60	7.10	7.43	7.68	54.34	61.48	66.95
汕尾市	4800.22	134.78	275.97	281.63	282.04	288.22	5.87	5.95	6.00	108.95	111.97	113.84
东莞市	2426.23	696.84	935.00	1058.54	1058.96	1284.62	43.63	51.63	52.95	51.90	79.76	84.35
中山市	1692.76	201.15	428.58	445.82	445.97	467.48	26.34	27.14	27.62	121.64	128.39	132.41
江门市	9281.44	582.55	693.81	756.23	756.48	793.60	8.15	8.31	8.55	29.81	32.33	36.23
阳江市	7788.38	243.83	423.92	438.88	438.52	461.08	5.64	5.80	5.92	79.99	85.36	89.09
湛江市	12177.76	862.58	1255.70	1286.78	1286.99	1342.91	10.57	10.86	11.03	49.18	53.28	55.69
茂名市	11337.11	663.62	680.32	701.15	701.70	749.65	6.18	6.47	6.61	5.66	10.59	12.96
肇庆市	14880.81	216.28	467.16	500.59	499.48	538.30	3.36	3.50	3.62	131.46	140.95	148.90
清远市	18977.04	242.02	679.04	723.35	724.09	775.98	3.81	3.96	4.09	198.89	210.62	220.63
揭阳市	3101.45	144.19	220.13	228.01	227.84	233.96	7.35	7.43	7.54	58.13	59.86	62.26
潮州市	5238.42	326.03	461.95	462.67	462.64	483.73	8.83	9.15	9.23	41.91	46.98	48.37
云浮市	7788.36	207.70	264.13	283.35	283.75	294.34	3.64	3.68	3.78	36.42	38.12	41.71

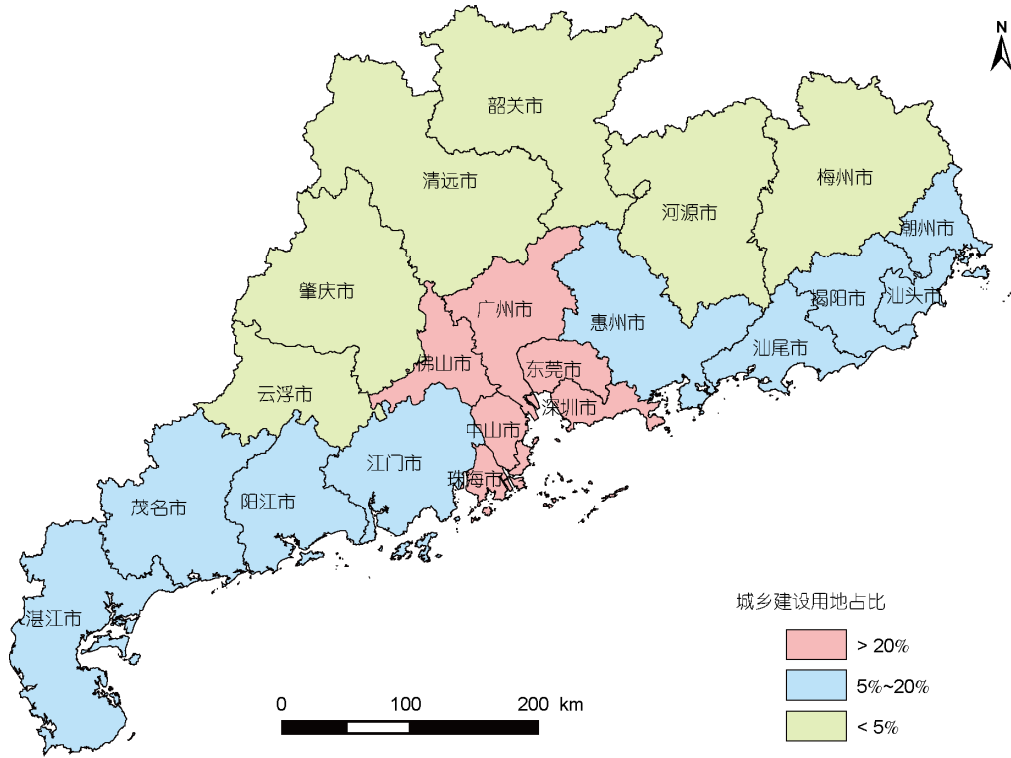


图 5 广东省各市级行政区 2000~2015 年城乡建设用地面积占城市总面积比例

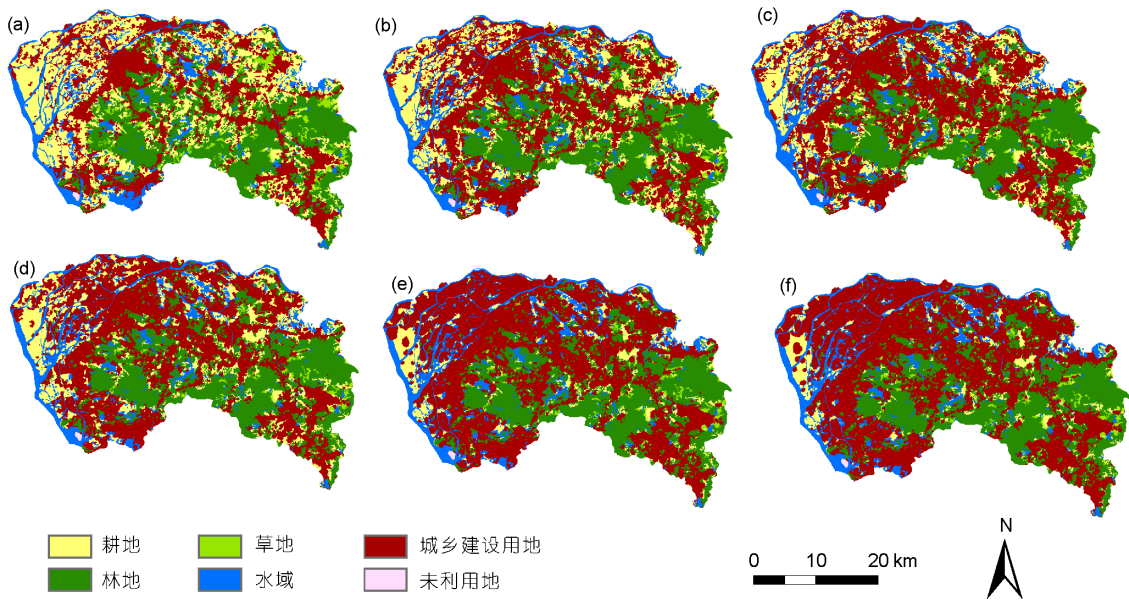


图 6 东莞市 2000~2006 年城乡建设用地变化及 GPU-CA 模拟结果

(a) 东莞市 2000 年土地利用分类图; (b) 东莞市 2005 年土地利用分类图; (c) 东莞市 2006 年土地利用分类图; (d) GPU-CA 东莞市 2006 年模拟结果; (e) GPU-CA 东莞市 2010 年模拟结果; (f) GPU-CA 东莞市 2015 年模拟结果

图 7 所示. 其中中山市、肇庆市、清远市、梅州市、河源市、汕尾市, 其扩张比例均超过 100%, 典型的如中山市, 2000~2015 年的扩张比例达到 132.41%,

有大量土地转变为城乡建设用地, 如图 8 所示. 广州市、珠海市、佛山市、韶关市、东莞市、惠州市、阳江市、湛江市、潮州市的扩张比例在 50%~100%之间,

典型的如惠州市, 扩张比例为 66.95%, 属于中速的扩张, 如图 9 所示. 深圳市、汕头市、江门市、茂名市、揭阳市、云浮市的扩张比例为 50% 以下, 典型的如云浮市, 其扩张比例为 41.71%, 有较少的土地转变为城乡建设用地, 如图 10 所示.

在以上多种城市扩张类型的模拟中, GPU-CA 均能准确的模拟该类型城乡建设用地的变化, 说明 GPU-CA 模型能够有效地应用于土地利用变化格局和演变过程的研究.

3 GPU-CA 模型的性能分析

并行计算的性能分析通常采用加速比的概念^[20,21], 即通过并行计算时间与 CPU 串行计算时间的比值来体现并行计算的加速性能. 因此在分析 GPU-CA 的计算性能时, 分别记录了 GPU-CA 模拟所耗费的时间及相同条件下 CPU 串行运算的时间, 通过计算其加速比来了解 GPU 计算对 CA 模拟过程的效率提升程度. 为了更好地对比 GPU-CA 的加速性能, 本研究依次选取了东莞市、珠江三角洲地区和广

东省作为模拟的区域, 使用 GPU-CA 和 GeoSOS 软件中的一般 CA 分别进行了运算, 比较了其模拟过程的运算性能, 如表 2 所示. 为了保证相同的对比条件, GeoSOS 软件在模拟过程中不进行界面刷新操作, 只进行后台运算. 同时由于两者的数据读取与输出过程基本相同, 因此主要通过对比 CA 模拟计算所耗费的时间, 比较 GeoSOS 软件串行计算方式与 GPU-CA 并行计算方式之间运算效率的差别.

从表 2 中可以看到, GPU-CA 对 CA 模拟过程的加速是非常明显的, 时间加速比最高可以达到 33.97, 比使用 CPU 运算时的执行效率有了非常大的提升, 说明 GPU-CA 与一般 CA 相比具有十分显著的加速性能.

图 11 表示了 GPU-CA 运算规模与时间加速比的关系, 可以看到随着运算规模的增加, 时间加速比也在提升. 主要原因是在运算数据量较小情况下, 数据传输的时间可能大于数据计算的时间, 因此并行计算的加速效果不明显. 随着运算规模的增加, 数据传输所花费的时间在总运算时间中所占的比重不断下降, 因此 GPU-CA 在大数据量情况下能获得更好的

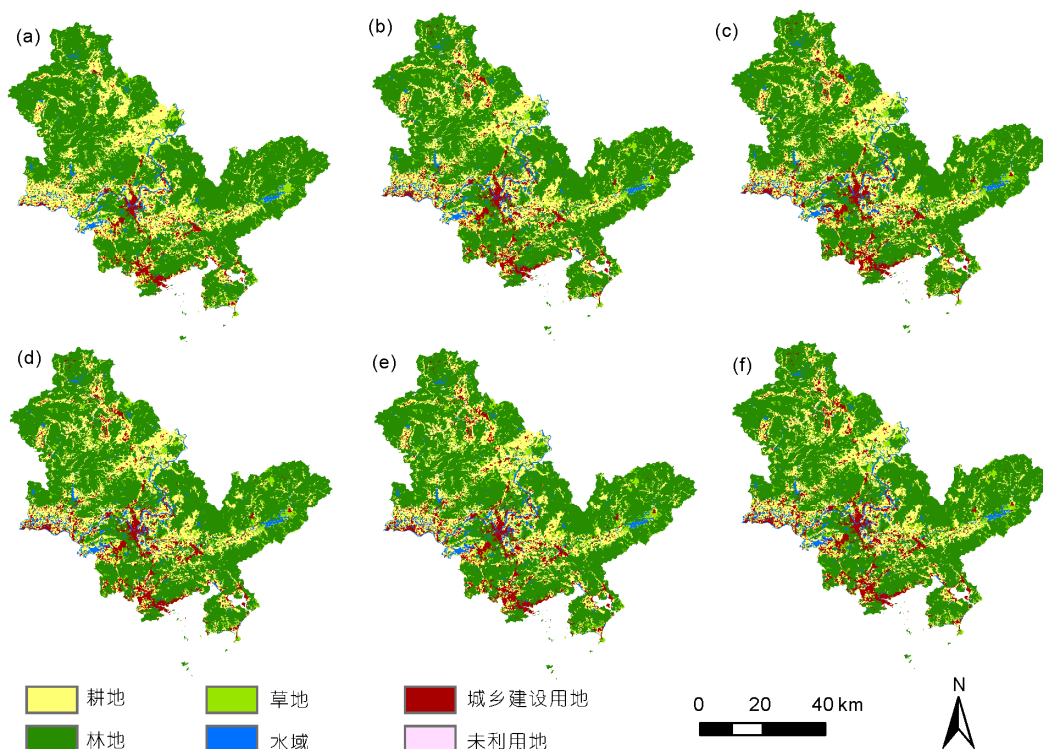


图 9 惠州市 2000~2006 年城乡建设用地变化及 GPU-CA 模拟结果图

(a) 惠州市 2000 年土地利用分类图; (b) 惠州市 2005 年土地利用分类图; (c) 惠州市 2006 年土地利用分类图; (d) GPU-CA 惠州市 2006 年模拟结果; (e) GPU-CA 惠州市 2010 年模拟结果; (f) GPU-CA 惠州市 2015 年模拟结果

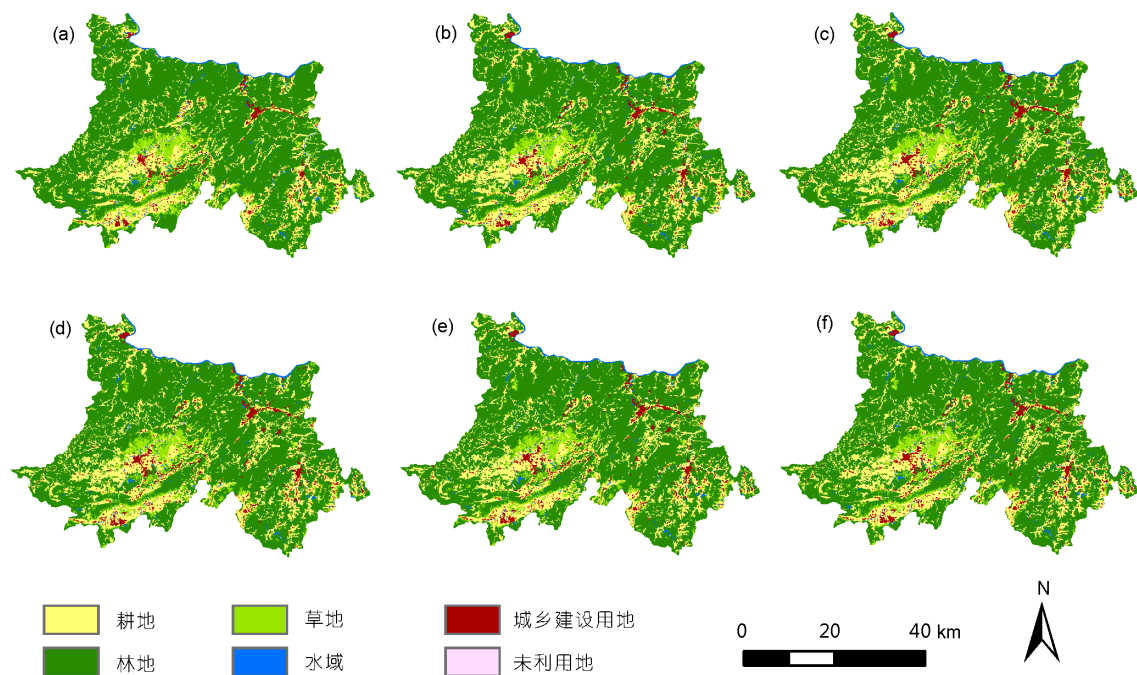


图 10 云浮市 2000~2006 年城乡建设用地变化及 GPU-CA 模拟结果图

(a) 云浮市 2000 年土地利用分类图; (b) 云浮市 2005 年土地利用分类图; (c) 云浮市 2006 年土地利用分类图; (d) GPU-CA 云浮市 2006 年模拟结果; (e) GPU-CA 云浮市 2010 年模拟结果; (f) GPU-CA 云浮市 2015 年模拟结果

表 2 CPU-CA 与一般 CA 模拟运算时间对比

研究区域	区域大小(栅格个数)	GPU-CA 计算时间(s)	一般 CA 计算时间(s)	时间加速比
东莞市	768×544	1.059	17.700	16.71
珠江三角洲地区	2912×2096	14.715	258.876	17.59
广东省	9792×7376	150.729	5102.147	33.97

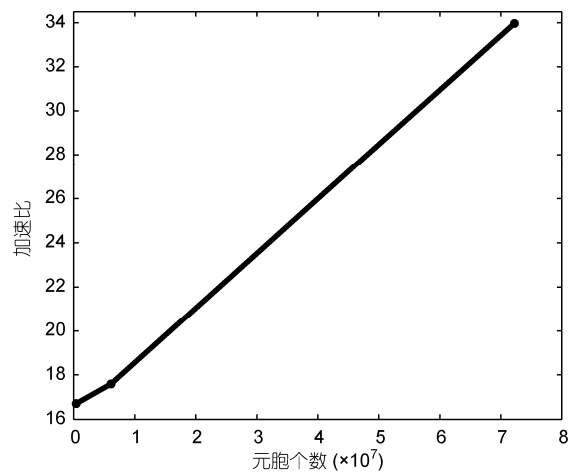


图 11 GPU-CA 运算规模与时间加速比关系图

并行加速效果. 这说明 GPU-CA 在处理大数据量问题时具有更加明显的优势, 同时也体现了 GPU 高性

能计算所具有的并行性好、计算密度高的优点, 实验结果与并行计算加速性能的理论分析是一致的^[20]. 因此 GPU-CA 计算模型非常适合应用于大尺度、高精度的土地利用变化模拟研究.

同时将 GPU-CA 的加速性能与基于负载平衡的并行元胞自动机模型 Parallel-CA 进行了比较^[10]. Parallel-CA 采用 3×3 邻域进行珠江三角洲地区模拟的计算时间为 1200~1600 s 左右, 加速比为 2 左右. 而 GPU-CA 的模拟时间为 15 s 左右, 加速比接近 18, GPU-CA 在计算性能上比 Parallel-CA 有很大的提高. 同时, Parallel-CA 需要使用 8 台计算机进行并行操作, 而 GPU-CA 只需在一台计算机中使用支持 CUDA 环境的显卡, 其计算成本与 Parallel-CA 相比也具有很大的优势, 这充分体现了 GPU-CA 计算成本低、计算密度高的特点.

4 结论

将 GPU 通用计算技术应用于大尺度土地利用变化模拟中构建 GPU-CA 计算模型,能够在保证模拟结果有效性的同时极大地提高模拟计算的效率,使原有 CPU 串行计算模式所面临的无法有效处理海量空间数据、计算时间过长等问题得以解决,是今后进行大尺度地理模拟和地理计算的有效计算模式。与其他并行计算模式相比, GPU 通用计算具有计算成本低、计算密度高、安装部署容易、计算加速比随计算规模提升等优点,更适合完成地理模拟和地理计算中的密集计算任务。本文采用广东省作为大尺度土地利用变

化模拟的研究区域,实际上 GPU-CA 计算模型的应用范围可以很容易地推广到全国和全球尺度。

本研究主要采用逻辑回归 CA 模型进行土地利用变化模拟,适合模拟非城市用地到城市用地的转变,而较难揭示多种土地利用类型之间的变化规律,今后可以引入神经网络 CA、决策树 CA 等多地类土地利用变化模型,以更好地解释和模拟研究区域内的土地利用变化过程。同时, GPU 通用计算可以通过程序优化等手段进一步缩短运算的时间,也可以使用 openMP 等并行计算协议来构建多 GPU 显卡的计算集群,以更有效地提升计算的效率和扩大处理数据的规模,这也是今后研究的方向。

参考文献

- 黎夏, 叶嘉安, 刘小平, 等. 地理模拟系统: 元胞自动机与多智能体. 北京: 科学出版社, 2007. 250-258
- Li X, Yeh A G. Neural-network-based cellular automata for simulating multiple land use changes using GIS. *Int J Geogr Inf Sci*, 2002, 16: 323-343
- Batty M, Xie Y. From cells to cities. *Environ Plan B-Plan Design*, 1994, 21: 531-548
- Clarke K C, Hoppen S, Gaydos L. A self-modifying cellular automaton of historical urbanization in the San Francisco Bay area. *Environ Plan B-Plan Design*, 1997, 24: 247-261
- Wu F, Webster C J. Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environ Plan B-Plan Design*, 2004, 5: 103-126
- 黎夏, 刘小平. 基于案例推理的元胞自动机及大区域城市演变模拟. *地理学报*, 2007, 62: 1097-1109
- 黎夏, 刘小平, 何晋强, 等. 基于耦合的地理模拟优化系统. *地理学报*, 2009, 64: 1009-1018
- 黎夏, 李丹, 刘小平, 等. 地理模拟优化系统 GeoSOS 软件构建与应用. *中山大学学报*, 2010, 49: 1-5
- Veldkamp A, Fresco L O. CLUE-CR: An integrated multi-scale model to simulate land use change scenarios in Costa Rica. *Ecol Model*, 1996, 91: 231-248
- Li X, Zhang X H, Yeh A G, et al. Parallel cellular automata for large-scale urban simulation using load-balancing techniques. *Int J Geogr Inf Sci*, 2010, 24: 803-820
- 吴恩华, 柳有权. 基于图形处理器(GPU)的通用计算. *计算机辅助设计与图形学学报*, 2004, 16: 601-612
- Owens J D, Houston M, Luebke D, et al. GPU computing. *Proc IEEE*, 2008, 96: 879-899
- NVIDIA. NVIDIA CUDA Compute Unified Device Architecture Programming Guide Version 2.0. 2008
- Chen F G, Ge W, Li J H. Molecular dynamics simulation of complex multiphase flow on a computer cluster with GPUs. *Sci China Ser B-Chem*, 2009, 52: 372-380
- 杨靖宇, 张永生, 董广军. 基于 GPU 的遥感影像 SAM 分类算法并行化研究. *测绘科学*, 2010, 35: 9-11
- Gobron S, Devillard F, Heit B. Retina simulation using cellular automata and GPU programming. *Mach Vision Appl*, 2007, 18: 331-342
- Ferrando N, Gosálvez M A, Cerda J. Octree-based, GPU implementation of a continuous cellular automaton for the simulation of complex, evolving surfaces. *Comput Phys Commun*, 2011, 182: 628-640
- Wu F. Calibration of stochastic cellular automata: The application to rural-urban land conversions. *Int J Geogr Inf Sci*, 2002, 16: 795-818
- Li X. Data mining of cellular automata's transition rules. *Int J Geogr Inform Sci*, 2004, 18: 723-744
- 刘德才, 王鼎兴, 沈美明, 等. 数据并行的性能分析. *软件学报*, 1994, 5: 8-15
- 杨庚, 王绍隶, 沈金龙. 基于曙光并行机的超大规模非线性方程组并行算法研究. *计算机学报*, 2002, 25: 397-402