



大数据地形可视化 建模及相关问题研究

张慧杰 / 著

Terrain Model Visualization for Big Data
and Related Research



科学出版社

大数据地形可视化建模 及相关问题研究

张慧杰 著

科学出版社

北京

内 容 简 介

本书给出基于大数据集的地形可视化建模方法,从自适应数据组织出发,阐述地形可视化多分辨率建模、骨架特征提取、地形形变以及虚拟地形合成等方法的原理和过程;并给出多分辨率外存数据组织逻辑模型,使数据载入内存的效率显著提高。在此基础上详细阐述层次模型简化技术与群智能算法相结合的模型简化方法。结合科学计算领域对高质量简化模型的需求,详细描述自适应确定误差阈值的方法,使多分辨率模型在简化比率较高的情况下,仍然保持较高的模型精度。同时阐述结合裂缝可见性的多分辨率模型裂缝修补算法。最后详细描述基于形态学的地形骨架特征提取方法,以及基于河网过程性建模的虚拟地形合成等应用。

本书可供从事计算机图形学、地形可视化等相关领域的科研人员和工程技术人员参考使用,也可以作为高等学校相关专业研究生、高年级本科生以及教师的工具书使用。

图书在版编目(CIP)数据

大数据地形可视化建模及相关问题研究/张慧杰著. —北京:科学出版社, 2016.1

ISBN 978-7-03-046218-3

I. ①大… II. ①张… III. ①影象地形图—研究 IV. ①P931
中国版本图书馆CIP数据核字(2015)第262198号

责任编辑:杨慎欣 张 震 / 责任校对:张凤琴

责任印制:赵 博 / 封面设计:无极书装

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

新科印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2016年1月第一版 开本:B5(720×1000)

2016年1月第一次印刷 印张:15 3/4 插页:1

字数:318 000

定价:96.00元

(如有印装质量问题,我社负责调换)

前 言

地形模型可视化已经在三维游戏、飞行模拟训练、战场环境仿真、地理信息系统、虚拟现实等领域有着非常广泛的应用。目前，地形模型的研究正在被拓展到航空航天、机载系统、三维地形导航系统和地质灾害预测等应用领域。近年来，随着数字测量和卫星技术的迅速发展，可利用的地形数据集越来越大，地表模型能够表现的细节也越来越多，地形场景也变得越来越复杂。因此，大规模或者超大规模地形场景的绘制必须得到外存的支持。然而，由于内外存之间存在明显的速度差异，已经成为大规模数据处理的一个瓶颈。因此，地形模型的外存管理具有很重要的研究意义，是基于大数据集的地形场景绘制的研究基础。

此外，地形模型简化不仅是解决面向大数据集的地形场景绘制问题的关键技术，而且在3D模型检索与地形导航以及地表水流模拟等科学研究领域也越来越受到关注。大部分模型简化算法很难兼顾算法的效率和简化模型的质量双重技术指标，通常根据场景绘制和科学研究两个不同的应用领域的需求算法有所侧重。前者更侧重于算法效率，满足图形硬件的绘制要求；后者更侧重于简化模型的质量，在简化比率较高的情况下要求简化模型仍具有较高的精度。然而，近年来随着科学计算可视化技术的发展，许多科学研究领域的应用也需要进行动态场景绘制。因此，寻求一种具有高自适应性的快速地形模型多分辨率建模方法有着至关重要的意义。

鉴于此，本书从大规模数据组织方法出发，阐述多分辨率建模、骨架特征提取、地形形变及合成等方法与原理，并且详细描述建模过程中可能存在的问题和相关解决方案。

全书共8章，第1章为绪论，分类介绍国内外学者在地形模型多分辨率建模领域的经典研究方法、关键技术以及研究现状。第2章系统介绍和分析地形模型的外存组织与多分辨率外存建模方法，以及支持大数据集的地形建模 out-of-core 技术。第3章在给出的面向大数据集的地形外存模型基础上，详细阐述基于聚类分析的数据调度策略：通过空间填充曲线为聚类节点排序编码、增量数据调度，提高 I/O 交互的响应速度。同时使载入内存的数据量大幅度减少，

因此通过聚类分析和增量数据调度策略使数据载入内存的效率显著提高。

第 4 章详细阐述利用群智能算法的思想求解基于特征的地形简化模型的方法、原理和过程。分析经典的逐层简化算法存在的问题，提出层次模型简化技术与群智能算法相结合的多分辨率简化模型建模方法。该方法重新构造表现简化模型的粒子和结合模型误差的粒子评价函数，通过最优粒子和粒子的最优状态作为启发信息指导简化过程，从而不再逐层细分而是根据启发信息实现模型的越层简化。因此，该方法在保证简化模型的自适应性和精度要求的同时大大提高了简化算法的效率。

第 5 章给出地形简化模型多误差阈值的自适应确定方法。分析传统算法中采用统一误差计算方法和阈值的缺点，结合科学计算领域对高质量的简化模型的需求，给出在简化比率较高的情况下，通过多误差阈值方法提高模型精度的策略。针对地势特征分割较为清晰的峡谷和高原相结合的地形，首先描述基于粒度计算的二分阈值方法，对地形进行粗分割后，有针对性地采用不同的误差计算和阈值方法，从而提高简化模型的精度。其次，针对地势特征分割不明显的起伏度较大的地形，给出基于特征域合并的自适应地形模型简化方法。其优势在于进一步细分可以提高简化模型的精度，忽略满足误差要求的节点可以提高算法效率。最后，针对地势特征分割不明显且地势变化缓慢的地形，给出结合地表起伏度的改进策略。同时重点解决多分辨率特征域合并与不规则形状的区域分割问题，加大特征域合并的粒度，因此降低特征域合并和区域分割的时间开销，模型精度和算法效率均略有提高。

第 6 章描述多分辨率简化模型的空间连续性、裂缝修复方法以及简单的纹理合成技术。分析已有的裂缝修补方法存在的问题，给出结合隐式层次结构和地形特征的裂缝修补方法。根据地势特征忽略不可见裂缝，因此减少了冗余三角形的产生，从而提高整体算法的性能。第 7 章详细阐述基于形态学的骨架特征提取方法，并与传统的剖面识别方法进行实验对比，通过引入形态学方法使提取的骨架特征更加准确，同时可以提取环路地形。第 8 章给出地形形变基本方法与原理，并且通过实验给出本书采用的插值方法应用在地形形变问题上的效果，并分析主要原因，总结不同的适用范围。同时详细描述一个基于爆炸模型的物理弹坑建模方法的应用案例。最后给出通过引入河网产生式规则进行地形过程性建模的方法，从而进一步阐述虚拟地形的合成方法。

总体看来，地形可视化研究领域仍然存在很多值得研究的科学问题，并且可利用的大数据集越来越多，同时地形可视化建模与可视分析相结合的若干应用研究将成为一个新的研究热点，也是作者未来的研究工作之一。

在此，衷心感谢东北师范大学马志强教授的支持、关心和鼓励；感谢科学出版社对于本书给予的大力支持；同时感谢一直与作者从事科研工作的刘亚鑫、何欣婷、高福建、曲德展、侯亚芳、马芸、邵亮同学。本书得到吉林省自然科学基金面上项目（项目编号：20140101179JC）、高等学校博士学科点专项科研基金项目（项目编号：20130043110016）的资助，在此表示衷心的感谢。鉴于作者的水平有限，书中难免有不足之处，恳请读者批评指正。

张慧杰

2015年7月25日于长春

目 录

前言

1 绪论	1
1.1 研究背景和意义.....	1
1.2 基于大数据集的高程数据管理策略.....	2
1.3 多分辨率地形建模方法.....	3
1.3.1 基于离散高程点的近似模型.....	3
1.3.2 基于层次的多分辨率模型简化方法现状.....	4
1.3.3 设计目标.....	7
1.4 多分辨率地形建模关键问题.....	7
1.4.1 多分辨率简化模型的误差评价.....	7
1.4.2 多分辨率模型的空间连续性.....	9
1.4.3 设计目标.....	11
1.5 骨架特征提取及地形合成.....	12
1.6 总结.....	13
2 基于超量数据集的外存模型	15
2.1 引言.....	15
2.2 问题分析及外存框架.....	16
2.2.1 金字塔外存模型.....	16
2.2.2 多分辨率外存模型框架.....	16
2.3 基于大数据集的数据组织.....	17
2.3.1 数据源.....	17
2.3.2 分块物理模型.....	18
2.4 基于特征的多分辨率外存层次结构.....	20
2.4.1 多分辨率逻辑模型.....	20
2.4.2 基于传递闭包的静态误差计算.....	22

2.4.3	FMRH 的构造算法	24
2.5	多分辨率外存模型优化及无缝拼接	26
2.5.1	误差函数	27
2.5.2	外存索引结构	29
2.5.3	建立外存索引结构算法	30
2.5.4	局部数据捡取	32
2.5.5	实验效果	37
2.6	本章小结	46
3	基于聚类分析的数据调度策略	47
3.1	引言	47
3.2	基于相近误差的聚类分析	48
3.2.1	基于相近误差的聚类定义	48
3.2.2	基于 FMRH 的聚类算法	49
3.2.3	聚类空间封闭性证明	51
3.3	基于 FMRH 的聚类层次数据块索引	52
3.3.1	基于 FMRH 的 Hilbert 空间填充曲线策略	52
3.3.2	多分辨率模型 Hilbert 空间填充曲线生成算法	54
3.3.3	多文件处理	56
3.3.4	基于大数据集的模式更新	57
3.4	基于聚类的多分辨率数据调度	58
3.4.1	目标聚类快速搜索策略	58
3.4.2	数据预取及增量调度策略	59
3.5	实验结果及分析	61
3.6	本章小结	63
4	基于 DPSO 特征选择的地形模型简化方法	64
4.1	引言	64
4.2	隐式层次结构	65
4.2.1	隐式层次结构的定义	65
4.2.2	CHS 的层次计算方法	66
4.2.3	CHS 的父子节点快速索引方法	68
4.3	简化模型的误差评价函数	70
4.3.1	简化域与邻接关系	71
4.3.2	法向矢量相似度计算方法	72

4.3.3	简化模型发散度函数	74
4.4	TMS-DPSO 算法	74
4.4.1	TMS-DPSO 算法基础	75
4.4.2	基于简化模型的粒子定义	76
4.4.3	基于简化模型的粒子评价策略	77
4.4.4	速度轨迹与状态更新	78
4.4.5	TMS-DPSO 算法描述	80
4.4.6	TMS-DPSO 算法收敛性分析与验证	82
4.5	实验结果与分析	85
4.5.1	发散度函数有效性验证	85
4.5.2	TMS-DPSO 算法的性能分析	88
4.6	本章小结	89
5	地形简化模型多误差阈值的自适应确定	91
5.1	引言	91
5.2	特征点选择	92
5.2.1	基于 CHS 的顶点集	92
5.2.2	特征点与特征域的定义	93
5.2.3	特征点选择算法	94
5.3	基于顶点粒度计算的二分阈值简化方法	96
5.3.1	基于顶点粒度的二分阈值划分	96
5.3.2	二分阈值模型简化算法	98
5.3.3	算法验证	99
5.4	基于特征域匹配度的多阈值简化方法	101
5.4.1	相似特征域合并	101
5.4.2	特征域匹配度简化算法	103
5.4.3	算法验证	104
5.5	基于区域分割的多阈值简化方法	107
5.5.1	基于起伏度的区域分割	107
5.5.2	自适应区域分割多阈值简化算法	110
5.5.3	自适应区域分割多阈值简化方法实验结果与分析	112
5.6	本章小结	115
6	多分辨率地形无缝模型及多纹理合成	117
6.1	引言	117

6.2	基于多分辨率的无缝地形简化模型	117
6.2.1	多分辨率无缝模型框架	117
6.2.2	基于特征的裂缝判定方法	118
6.2.3	基于特征的裂缝修补算法	120
6.2.4	算法实验结果与比较	123
6.3	多分辨率纹理外存模型	125
6.3.1	多分辨率纹理算法基础	125
6.3.2	基于 FMRH 的多分辨率纹理模型	126
6.4	多纹理合成方法	129
6.4.1	纹理合成整体思路	129
6.4.2	候选点选择	130
6.4.3	特征提取及纹理模板生成	131
6.4.4	基于模板的纹理合成及实验结果	133
6.5	本章小结	135
7	地形骨架特征提取方法	136
7.1	引言	136
7.2	相关工作	137
7.3	候选特征点识别与特征多边形连接	139
7.3.1	剖面识别	139
7.3.2	特征多边形连接	140
7.4	形态学特征选择算法	142
7.4.1	特征域填充	142
7.4.2	特征域细化	143
7.5	矢量化特征模型	147
7.5.1	矢量坐标定位及矢量线段生成	147
7.5.2	矢量检测优化	148
7.5.3	剪枝与环路处理	150
7.6	MEPPA 实验结果与分析	152
7.6.1	骨架特征线形态对比分析	153
7.6.2	骨架特征线提取效果	156
7.6.3	实验结果对比分析	158
7.7	M-PPA 优化算法	161
7.7.1	改进的剖面识别算法	162
7.7.2	形态简化	164

7.7.3	连接特征点	166
7.7.4	环路处理	167
7.7.5	减枝处理	172
7.8	M-PPA 实验效果与分析	176
7.8.1	效率对比	177
7.8.2	准确度分析与对比	178
7.8.3	算法稳定性对比	181
7.9	本章小结	182
8	可变地形及合成建模方法	184
8.1	引言	184
8.2	基于空间插值法的地形形变方法	185
8.2.1	薄板样条函数插值法	185
8.2.2	反距离权重法	187
8.2.3	两种形变方法比较	188
8.3	移动最小二乘法	191
8.4	地形形变方法的应用	192
8.4.1	河道地形空间插值研究	193
8.4.2	山峰高度插值研究	194
8.5	物理弹坑模型建模方法	196
8.5.1	弹坑模型相关工作	197
8.5.2	多分辨率网格表示及形变区域扩展算法	198
8.5.3	物理弹坑模型建模	201
8.5.4	构建弹坑模型物理属性表示方法	205
8.5.5	实验验证	208
8.6	基于河流水文学过程性地形建模方法	211
8.6.1	合成建模技术相关工作	212
8.6.2	基于河流水文学的河网构建	213
8.6.3	山脊的构造	216
8.6.4	地形块的嵌入以及变形处理	220
8.6.5	实验结果	222
8.7	本章小结	224
	参考文献	225

彩色图版

数字摄影、遥感、卫星等测量技术的发展,使获取高分辨率的几何数字高程模型(digital elevation model, DEM)以及纹理影像数据成为可能^[1]。因此,建立国家级乃至全球性的空间信息可视化系统成为一种必然趋势^[2]。其中,地形模型是重要的组成部分^[3]。由于地形模型的显著特点是包括海量数据,因此通常采用多分辨率简化模型代替原始模型,在数据量减少的情况下能够保证地形模型的可视化质量和精度要求^[4]。然而,多分辨率建模的相关技术仍然面临着严峻的挑战。本书的工作是在提出大数据集的外存管理及数据调度基础上,研究满足场景绘制和科学计算可视化双重需求的地形多分辨率建模方法,并解决由此产生的关键问题。

1.1 研究背景和意义

1998年1月31日,时任美国副总统戈尔在加利福尼亚科学中心提出了数字地球的概念,其具有三个特点:①可以嵌入海量空间数据;②具有多分辨率模型表示;③能够表现三维地球。2003年3月,美国有线电视新闻网(CNN)在直播美军进攻巴格达时,曾使用钥匙眼公司的三维软件,精细勾勒出美军轰炸巴格达市区的具体方位。它是一款将地球表面的卫星图像和航拍照片结合在一起,形成全景三维的世界地图的软件。这便是现在家喻户晓的Google Earth的原型。2004年10月,Google Earth收购了钥匙眼公司,经过近一年的酝酿和开发,2005年6月推出了Google Earth软件^[5]。

从Google Earth 1.0版本推出至2015年发布的Google Earth 7.1,Google Earth提供了越来越清晰的高分辨率三维影像地图,越来越多地支持三维矢量数据的模型重建,可以让用户足不出户看到世界各地建筑物和地形的三维图像。在这个大的应用背景下,许多公司正在进行为人们的工作生活提供便利服务的二次开发。在Google Earth提供的应用开发接口(API)基础上,进行二次开发的第三方网站非常多^[5]。然而,关于三维地形的可视化效果还非常不完善,并且不能满足地学等相关应用领域科学计算可视化的需求。许多专家和学者仍在致力于相关理论和关键技术的研究。

目前,很多不同分辨率的高程数据可以进行免费下载,这在全球范围内为研究者提供了良好的数据基础^[6]。常用的数据源包括:数字高程模型 GTOPO30 提供的千米分辨率覆盖全球的 DEM 数据;美国国家航空宇宙航行局(NASA)和美国国家地理空间情报局(NGA)的联合项目 SRTM 提供的 DEM 数据^[7],其中 90 米分辨率的数据全球覆盖率为 80%,30 米分辨率 DEM 数据覆盖了美国地区。SRTM 的数据总量超过 10TB 级以上。除此以外,很多科研机构还提供了小范围的 10 米高分辨率的 DEM 数据源。

面对如此庞大的数据规模,以及计算机硬件有限的数据处理能力,地形模型可视化建模正面临三个主要挑战:① 基于大数据集的 DEM 数据组织和调度^[8];② 用于动态场景绘制的多分辨率简化模型的效率问题;③ 用于科学计算的多分辨率简化模型的精度问题。很多学者已经提出了针对其中一个方面的解决方案。然而,随着地形可视化在地理信息系统(GIS)^[9]、战场环境仿真、机器人导航^[10]、地质灾害预测等领域的广泛应用,对算法的实时绘制效率和简化模型精度均有较高的要求。因此,现在面临的一个新的问题是如何高效地构建高质量的简化模型,满足场景绘制和科学计算的双重需求。

鉴于民用和科学研究领域对地形模型的需求,高分辨率 DEM 数据获取手段的多样化,以及若干尚待解决的问题,基于大数据集的地形模型多分辨率建模技术的研究具有必要的理论研究价值和很好的应用前景。

1.2 基于大数据集的高程数据管理策略

超大规模的原始数据无法一次性载入系统内存,因此通常根据具体需求分批次将数据动态载入内存。然而,内外存之间存在严重的速度差异,所以通常采取相关策略以便于减少由于内外存频繁调度导致的系统性能下降。传统的外存模型以及内外存调度方法归纳起来包括:分层分块管理数据的外存模型^[11-16]、增量数据调度^[17-20]、利用空间填充曲线建立数据索引^[21-25]、与视点相关的数据剪裁策略^[26-28],以及多种技术的混合策略^[2,29,30]。其目的都是在尽量小影响近似模型的可视化质量的情况下,有效减少载入内存的数据量,提高调度效率^[31-33]。

分块数据组织方式是将统一分辨率的数据源,按照分块原则划分为规则数据块^[12,16],其优势在于不受数据规模的限制,检索数据的效率较高。但是,由于数据块的分辨率均相同,因此载入内存的数据量相对较大,调度效率很差。鉴于此,很多文献建立了分层结构进行外存数据组织^[14,34],其原理是建立金字塔式的数据结构。自上而下将具有相同分辨率的数据存储为一层,最高层为最低分辨率的数据,然后逐层增加数据量,最底层为全分辨率数据(即分辨率最高)。此类方法实际上是以增加外存的存储容量为代价,从而实现减少载入内存的数据量的目的。

问题在于该类方法没有对多分辨率信息进一步管理,当数据规模增加时,这类外存算法需要花费大量时间查找相关数据节点,因此数据调度效率会迅速下降。

针对快速检索外存数据从而实现高效数据调度的问题,很多学者提出了离线(out-of-core)数据调度算法^[14, 17, 20]。Pajarola 描述了在大规模地形系统中采用 out-of-core 方法的思想^[17],将原始地形分解成正方形瓦片(tiles)存储在外存上,然后按照编码规则对瓦片进行排序,最后根据内存需求进行实时数据调度。其优势在于调度算法可以在短时间内检索到内存所需要的瓦片数据;缺点是数据更新困难,外存空间消耗较大。戴晨光等^[14]在此基础上提出了瓦片金字塔结构以及目标瓦片搜索算法,实现了局部地形数据的动态更新。然而其瓦片金字塔结构仍然没有解决外存空间消耗较大的问题,并且在数据调度的同时进行多分辨率目标数据搜索的策略导致了调度算法的效率下降。

除了采用 out-of-core 技术以外,也有学者采用数据压缩技术建立支持大规模数据的地形模型,典型的方法是采用基于小波的编码方案实现地形数据的压缩^[8]。由于规则格网三维 DEM 数据可以表现为一个二维矩阵,这与二维的灰度图像等价,因此可以将其看作二维离散信号,采用多进制小波技术对地形数据进行压缩和解压缩处理。该类方法的优势在于有效减少了数据规模,但是解压缩的时间消耗使系统的整体性能有所下降。

鉴于以上分析,本书的解决策略是外存模型由两部分组成:物理模型和逻辑模型。物理模型是对原始数据的等分辨率数据分块,其优势在于可以支持任何规模的数据扩展,并且不增加多分辨率外存数据的冗余存储,外存空间消耗较小。逻辑模型是在物理模型基础上建立的多分辨率的层次结构,并对外存数据进行聚类分析,可以实现减少载入内存数据量的同时降低大规模外存数据检索的时间消耗,提高调度效率。

1.3 多分辨率地形建模方法

多分辨率地形模型是由处于不同分辨率的面片构成的简化模型。在地势平坦的区域采用较低分辨率的数据,反之采用高分辨率的数据。目的在于为原始模型建立一个近似模型,使其满足两个条件:① 数据量较少;② 与原始模型的差异尽量小^[35]。

1.3.1 基于离散高程点的近似模型

在科学计算可视化领域,通常需要利用较少的高程数据构造高精度的近似模型^[36]。对于给定的有限离散点集合 $(x_i, y_i) \in \Omega, i=1, 2, \dots, n$, 且 $\Omega \in R^2$ 是二维平面有限域,对应的高程值 $z_i \in R, i=1, 2, \dots, n$, 则构造函数 $f(x_i, y_i)$, 使其满足

约束条件:

$$f(x_i, y_i) = z_i + \Delta z, \quad i = 1, 2, \dots, n \quad (1.1)$$

式中, Δz 是残差值。

根据式 (1.1), Δz 是近似曲面与原始高程值的差异, 当 Δz 趋于 0 时, 近似模型的精度达到最高^[37, 38]。这也是很多算法面对近似模型的高精度需求, 采用线性方程组进行求解的主要原因。然而, 该类方法面临的主要问题是构造支持大数据集的逼近算法, 并且提高算法的效率。显然, 基于全局样条的逼近方法^[39, 40]不适合大规模的近似模型求解, 原因在于当约束条件过多时需要解决病态约束条件问题, 所以算法效率较差, 同时耗费大量内存。

为了避免上述问题, Haber 等提出了基于任意形状大区域的逼近方法^[41], 算法步骤: ① 采用规则三角剖分方法将地形分割为多个子集; ② 在每个三角形子集上使用奇异值矩阵分解 (SVD) 计算局部最小二乘逼近; ③ 使用 Bernstein-Bézier 样条拼接各个局部曲面形成全局地表。该算法的优势在于在局部区域有效降低了方程组的规模; 缺点是在子集分割时缺少层次结构, 当某个局部子集需要更高精度的计算时, 导致全局所有的子集得重新计算。Scheib 等对上述算法进行了改进^[37], 使用三角形二叉树对原始高程数据进行自适应分层, 确保局部的修改不会影响到全局, 但是为了消除不同层次的三角形区域之间的裂缝, 算法变得相对复杂。

由此可见, 基于层次的思想是解决大数据集的有效手段^[42, 43]。Bertram 等^[44]采用了四叉树结构自适应划分原始数据为不同的聚类, 在每个聚类的局部数据点上采用最小二乘拟合方法, 最后通过聚类边界上的点把分段多项式合并成一个光滑的 B 样条曲面。显然, 该方法适用于大规模的地表逼近。Pouderoux 提出了类似的方法^[45], 具体步骤: ① 使用完全平衡二叉树细分全局表面, 使其成为相互叠加的局部子区域; ② 在每个叶节点中, 使用径向基函数 (RBF) 构造子表面; ③ 使用统一分割函数 (PoU) 将子表面合成 C_1 光滑连续的全局表面。该方法的特点是易于实现, 并且算法的时间复杂度降为 $O(N \cdot \log_2 N)$ 。

基于离散点的逼近方法属于离线地表建模方法, 虽然经过以上改进处理, 可以满足大数据集的科学计算需求^[43]。然而, 由于离散数据点不便于管理和调度, 因此不适用于需要动态更新的场景绘制。

1.3.2 基于层次的多分辨率模型简化方法现状

根据高程点分布情况不同, 地表模型主要可以分为两大类: 规则格网 (regular square grid, RSG) 模型^[46, 47]和不规则三角网 (triangulated irregular networks, TIN) 模型^[48-50]。RSG 模型的数据特点是以行列等间距的矩阵形式存储, 而 TIN 模型数据呈不规则分布。因此, 在渲染大规模地表数据集时 RSG 模型在数据管理和

多分辨率模型剖分方面更具优势,并且占用存储空间相对较小^[51]。TIN 模型数据的不规则分布通常可以体现地势特点,所以其优势在于表现的简化模型质量更高,适合科学计算方面的应用。

Hoppe 提出的渐进网格算法 (progressive meshes, PM) 是经典的基于 TIN 的简化算法^[48,49],算法步骤:① 对地形分块;② 对于每块通过边崩溃实现模型简化;③ 通过顶点分裂实现简化模型复原。该算法的优势在于可以表现任何复杂地势,但是消耗大量的内存空间,不适合大数据集的地表建模。Garland 等提出了基于二次曲面误差判断的模型简化方法,并开发了 QSlim 软件^[52],在简化模型质量和算法效率的平衡方面取得了较好的效果。在此基础上又将该方法扩展到任意维度上,解决了通用网格模型简化问题^[53]。同时开发了 GSlm 软件,相对于 QSlim 在算法性能方面没有大的提高。

近年来 Amor 等^[54]提出了一个求解 TIN 简化模型的整体框架,通过忽略冗余三角形的产生,达到提高算法效率的目的。Ben-Moshe 等^[55]以网格最简为设计目标,提出了基于距离保持评价策略的细粒度简化方法。这些算法表现的多分辨率地形的共同优势在于:① 表现的地形特征非常多样化;② 由于高程点的不规则分布简化模型的精度较高,可以满足科学计算的需求。缺点在于:① 剖分算法复杂,内存和时间消耗较大;② 不利于大规模数据管理和调度;③ 很难与规则几何纹理结合。

很多学者在 RSG 模型数据上进行大数据集的地表多分辨率建模^[20, 21, 30, 56, 57]。根据不同需求,可以归纳为两类:粗粒度的模型简化 (coarse grained simplification of model, CGSM)^[58-60]和细粒度的模型简化 (fine grained simplification of model, FGSM)^[57, 61-64]。

CGSM 方法主要是将三角形组作为基本单位对原始数据进行简化处理,实现三角形的批量绘制^[60]。其特点是:① 冗余三角形数量较大;② 剖分算法简单高效;③ 需要高性能图形硬件。Willem 等最早提出了粗粒度简化方法——Geometrical Mipmapping 算法^[65]。该方法将纹理 Mipmap 映射的思想应用到几何图形上,其整体思路:① 将地形分割成相等尺寸的正方形 Geomipmap 块;② 在每个 Geomipmap 块中,地形数据以隔行采样的方式被实时简化。细节层次以块为单位进行选择,所以算法效率较高。由于该方法简化模型冗余三角形过多,所以绘制时对图形硬件的性能依赖较大。

随着硬件可编程高级语言的出现,如 GLSL、HLSL,基于 GPU (graphic processing unit) 的可编程技术得到迅速推广^[66]。Larsen 描述了基于硬件优化的多细节实时地形绘制算法^[67]。CGSM 方法原则是减少 CPU 的计算时间,将大量三角形面片的绘制任务转移给 GPU 去实现^[68]。已有的 CGSM 方法包括 Ulrich 提出的 Chunked LOD 算法^[58]、Levenberg 提出的 Aggregated LOD 算法^[69]、Losasso 提

出的 Geometry Clipmaps 算法^[59]、Schneider 提出的 Pressive Tile LOD 算法^[70]以及 Yotam 提出的 Seamless LOD 算法^[71]。

这些算法发挥了 GPU 的硬件优势，在绘制效率方面均取得了比较理想的效果，然而，仍普遍存在着两大主要问题：① 对于模型无限扩展的真实细节，采用粗粒度的简化方法最终可能使得三角形数量超过 GPU 的处理能力，无法仅通过硬件可编程技术来解决；② 粗粒度的简化方法自适应能力差，即在三角形数量相等的条件下，对局部细节刻画的程度不够。所以，相对于 FGSM 方法，CGSM 方法的模型精度较低。因此，CGSM 方法不能胜任硬件受限的科学计算领域高精度多分辨率模型的需求。

FGSM 方法是以每个三角形为单位对地形模型进行简化。由于其细粒度简化的特性决定，该类算法在简化模型的精度、保持细节特征和自适应性等方面具有非常广阔的应用前景^[62]。但是，由于剖分粒度较小，所以相对于 CGSM 方法，FGSM 方法剖分效率较低^[61]。针对 RSG 模型数据，FGSM 方法的基本原理是将原始数据剖分成多分辨率的层次结构，从而达到降低地形模型复杂度的目的^[63]。

基于四叉树结构的 FGSM 简化算法是一类广泛采用的多分辨率建模方法^[72-74]，它们在存储结构和剖分效率方面都具有优势：① 可以使用二维数组存储四叉树结构；② 可以利用规则分布特性实现对地形数据的快速索引^[75]；③ 四叉树中非叶子节点都具有四个子节点，所以剖分的粒度较高。缺点在于不同分辨率面片之间的裂缝消除相对复杂。文献^[62]提出了实时多分辨率动态地形建模的完整解决方案，并且解决了地表网格模型的不连续问题。文献^[47]为多分辨率地形模型（MTM）建立了索引数据结构，从而在降低了简化模型的数据量的同时提高了算法效率。Pajarola 对基于四叉树的地形多分辨率建模方法进行了综述^[76]。

近年来，根据应用需求，很多学者提出了基于球体的地表多分辨率建模方法^[77, 78]。Cignoni 等提出了基于粗粒度的 TIN 和 RSG 模型相结合的算法^[34]，目的是利用 RSG 和 TIN 的优势，然而对每个数据块进行内插 TIN 的操作使算法的效率大大降低。类似地，Pajarola 提出了基于四叉树的不规则三角网算法（QuadTIN）^[79]，步骤是：① 利用四叉树进行分块；② 每个块内采用 TIN 内插。通过步骤②多分辨率简化模型的精度显著提高，但是效率有所下降。

为了利用 TIN 模型精度较高的优势，避免块内进行 TIN 数据建模带来的时间损耗，一些学者提出了正则三角形不规则网（right-triangulated irregular network, RTIN）^[64, 80]。优势在于可以利用 RSG 模型的规则剖分方法实现 TIN 模型的不规则分布特性。然而，事实上从很多文献的实验结果来看，很少有方法达到这个技术指标。更多的基于规则格网或者半规则格网的算法可参考其他综述文章^[81]。