

并行群体仿真研究进展

周文平, 唐好选, 季振洲

(哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001)

摘要: 通过对比分析基于多核 CPU、GPU 和分布式系统等三种并行群体仿真体系结构, 提出了混合结构的并行群体仿真框架; 对并行环境下的按空间区域分割和按 agent 分割等多种任务分割关键算法进行了归纳总结; 对并行群体仿真的性能评估方法进行了分析。最终给出了群体仿真中并行加速算法的研究方向。

关键词: 群体仿真; 并行仿真; GPU; 任务分割; 负载均衡

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2013)01-0030-06

doi:10.3969/j.issn.1001-3695.2013.01.006

Review on parallel crowd simulation

ZHOU Wen-ping, TANG Hao-xuan, JI Zhen-zhou

(School of Computer Science & Technology, Harbin Institute of Technology, Harbin 150001, China)

Abstract: This paper compared three types of parallel architecture of crowd simulation such as multi-core CPU based architecture, GPU based architecture and distributed system architecture, and proposed a mixed structure architecture. It summed up multiple algorithms on task partition under parallel environment such as partition by space region or partition by agents. It analyzed performance evaluation methods of parallel crowd simulation. Finally, it summarized the key points and direction of using parallel computing method for crowd simulation.

Key words: crowd simulation; parallel simulation; GPU; task partition; load balance

0 引言

人群仿真是一种研究人群在各种场景环境下群体行为特征的虚拟现实方法。人群仿真技术广泛应用于社会群体安全、建筑设计、交通疏散、体操培训、军事演练、数字娱乐、历史场景重现等各种场合。在人群仿真出现之前, 人群行为是通过从物理学和社会学的角度对历史事件进行分析中总结出来的, 这对于重现事件发生的全过程是极其不方便的。近年来, 随着各种公共场合的人群安全问题屡次出现(如紧急状态下的人群踩踏事件、火灾逃逸事件), 以及人们对数字媒体中(如电影动画、电子游戏)的逼真性的需求不断提高, 各国对人群行为的研究都表现出极大的关注。人群仿真已经成为涉及物理学、社会学、心理学以及计算机科学的交叉研究领域。

用计算机仿真方式来研究群体行为最早可追溯到 1987 年 Reynolds^[1] 建立的分布式群体行为模型。由于研究目的和应用场合的不同, 形成了群体仿真研究的两大领域: a) 关注行为真实性方面; b) 关注高质量可视化方面。

在行为真实性方面, 通过对社会群体领域和社会心理学领域两方面知识的研究, 实现了从专注群体行为特征的真实性建模到逐步关注群体中个体行为的真实性和多样性建模的转变。Helbing 等人^[2] 系统地提出了恐慌状态下人群撤离的运动仿真模型; 美国宾夕法尼亚大学人体仿真实验室(HMS)的 Pelechano 等人^[3] 将心理学模型集成到群体仿真系统中, 以获得可信的紧急逃生行为; Yu 等人^[4] 将合作和协调的群体行为特征引

入群体行为模型当中; Musse 等人^[5] 提出了一种包含自治的智能体、组和分级自治的交互对象的群体行为模型; Pelechano 和 Durupinar 等人^[6,7] 将个体的个性特征加入到群体行为模型中, 增加了个体行为的多样性。

在高质量可视化方面, 对此要求比较高的场合是在动画领域和游戏领域。有说服力的视觉效果是至关重要的, 这种视觉效果的研究所关注的是渲染和动画的方法。McDonnell 等人^[8,9] 提出了群体仿真的三维模型和真实性人体动画渲染方法; Dobbyn 等人^[10] 对群体成员实现了三维模型动画生成, 并附着纹理和光照以融合仿真环境。

在大规模群体仿真的过程中, 无论是行为真实性细节的增加还是可视化质量的提高, 都会带来运算资源消耗的增长。如何提高群体仿真的运行效率, 成为群体仿真迫切需要解决的难题。为了提高群体仿真的运行效率、实现群体的实时渲染, 人们总是在群体的规模与渲染质量之间进行取舍, 已有的方法主要有 LOD 细节层次方法、基于图像的 imposter 方法和基于 GPU 的硬件加速方法三种。近年来, 有部分研究人员在并行群体仿真研究方面进行了一些尝试, 以满足运算资源的需求。

大规模群体仿真还具有如下特点: a) 环境一般比较复杂; b) 个体规模一般都比较小; c) 个体之间以及个体与环境之间交互频繁。大规模群体仿真的这些特点, 使人群仿真过程既体现出大的计算量, 又体现出可能的可并行化的特性。传统的方法是通过简化人体或人群行为的模型来达到逼真性和高效性的折中。采用并行计算的方法, 通过充分挖掘硬件性能, 在人

收稿日期: 2012-06-02; 修回日期: 2012-07-08

作者简介: 周文平(1984-), 男, 博士研究生, 主要研究方向为并行计算、虚拟现实(zhoulangtian@163.com); 唐好选(1971-), 男, 副教授, 博士, 主要研究方向为计算机图形学、虚拟现实; 季振洲(1965-), 男, 教授, 博导, 博士, 主要研究方向为计算机体系结构、并行计算。

群仿真中越来越受到专家学者的重视,并行计算方法可以应用在人群仿真的多个方面。

尽管国内外关于计算机群体仿真的研究有很多,但是关于并行群体仿真的综述性文章笔者迄今尚未发现。本文的贡献在于:a)系统性地总结了并行群体仿真的研究方向和思路;b)对比分析了三种并行群体仿真体系结构,总结出混合体系结构的优势;c)归纳了并行群体仿真的多种任务分割算法;d)提出了一种并行群体仿真的性能评估方法;e)探索了并行群体仿真新的研究方向。

1 并行群体仿真的体系结构研究

并行群体仿真的体系结构主要包括三种:a)基于CPU或多核CPU的体系结构;b)基于分布式系统的体系结构;c)基于GPU通用计算的体系结构。本文将对三种体系结构逐一介绍,同时给出优劣对比分析结果。

1.1 基于CPU或多核CPU的并行群体仿真体系结构

基于CPU的并行群体仿真主要通过采用多线程技术来实现,其中包括基于普通CPU的多线程并行仿真和基于多核CPU的并行仿真两种类型。

在基于普通CPU的并行仿真方面,Kota等人^[11]提出了一种基于普通CPU的多agent的群体仿真系统,并将其应用到救援系统当中,同时分析了该系统的可扩展性问题。在文中还提出了一种微观agent架构和一种宏观—微观结合的agent架构。其中微观agent架构的主要思想是每个个体由一个agent代替,给每个agent分配一个线程,这样,有多少个个体就在CPU上运行相应数目的线程,因此系统的规模受到系统内存资源限制的约束;宏观—微观结合的agent架构则是对微观agent架构的一种改进,通过使用数据库来避免内存局限,同时使用群组抽象方法来降低复杂度,在仿真过程中,用数据库表格维护所有个体的信息,将聚拢在一簇的个体看做是一个群组,每个群组被看做是一个单独的实体,称做群组agent(crowd agent)。

在基于多核CPU的并行仿真方面,Reynolds^[12]在包含一个PPU和七个SPU的PS3游戏机上实现了虚拟水底世界,能够处理达数千的仿真实体。系统采用交互式的粒子系统模型,通过周边相邻个体的信息来决定粒子的行为。假如相邻个体为所有其他个体,计算复杂度为 $O(n^2)$ 。PS3包含一个PPU和七个SPU,其中一个SPU保留使用,所以只使用其中六个SPU来并行更新仿真的内容。系统采用空间分割的方法,将仿真更新分解成不相互交叉的任务,使每个任务可以按任意顺序在SPU上执行。在二维动画中,对15 000个个体处理速率达到60 fps,对5 000条三维高分辨率的虚拟鱼的处理速率达到30 fps。

Berg等人^[13]通过在多核CPU上使用多线程技术实现了群体仿真的加速。他们在全局路径规划中使用预处理路径图,这种方法可以显著地提高基于多核CPU的群体仿真计算性能。该方法基于速度障碍物的避碰模型,每个agent通过判断其他agent或障碍的位置与速率信息来决定自身的运动行为。系统在16核心的Intel处理器上执行,仿真结果显示系统的性

能和系统的CPU核心个数呈线性增长关系。

1.2 基于分布式系统的群体仿真体系结构

研究分布式系统的群体仿真体系结构之前,首先需要对现有的分布式系统通信模型有个大概的了解。传统的分布式系统通信模型包括P2P模型、点对点模型、分布式对等模型、集中式模型以及服务器机群模型五种方式。分布式系统无外乎是建立在某种特定的通信模式或者通过同时使用多种通信模式的混合通信模式之上。

现有的分布式系统仿真实验常用于解决大规模复杂系统仿真的性能问题,而群体又是最典型的大规模复杂系统,通过使用分布式系统对群体的行为进行仿真,能够有效地解决群体仿真的性能问题,提高群体仿真的效率,从而满足复杂群体互动对高分辨率视觉效果实时性要求。

表1是一个基于分布式虚拟环境平台的并行群体仿真的分类。

表1 基于分布式虚拟环境的并行群体仿真分类

名称	基于分布式平台	群体行为模型	虚拟环境几何表示	相关文献
Crowd Federate	HLA + OLIVE	社会心理学模型(psychological modeling)	三维立体	文献[14]
HLA_AGENT	HLA + SIM_AGENT	漫游行为(roaming behavior)	三维立体	文献[15]
Federated Agent-Based Crowd	HLA + RePast + JMess + UnReal	认知模型(cognitive models)	三维立体	文献[16]

1.2.1 基于分布式虚拟环境平台的群体仿真体系结构

分布式虚拟环境中应用最广泛的国际标准是HLA协议标准。现有的群体仿真主要针对基于HLA协议标准的仿真平台进行仿真研究。

Petty等人^[14]实现了一种基于HLA标准的分布式仿真的联邦,能够与现有的军事仿真系统协同操作,生成的群体行为基于可信的心理学基础。

Lees等人^[15]研究了在计算机游戏和虚拟环境中使用智能agent的方法,提出一种符合HLA标准协议的HLA_AGENT构建方法,通过HLA_AGENT工具构建的富认知agent可以有效地与虚拟环境或游戏环境集成为一个整体,达到良好的互操作性。

Low等人^[16]通过整合多agent仿真系统RePast、智能推理机系统JESS、现有的商业游戏引擎Unreal以及基于HLA标准的分布式虚拟环境平台,共同构建了一个分布式的并行群体仿真架构模型。

Chen等人^[17]则将HLA标准与网格(grid)结合,实现了大规模群体基于层次化网格架构的混合建模仿真。

其优点为:

可以快速地使用现有平台搭建群体仿真框架,通过使用比较成熟的分布式虚拟环境理论,让群体仿真研究人员有更多的精力关注群体行为多样性的仿真和群体表示的渲染。

其缺点为:

a)基于HLA的群体仿真主要适合基于agent的分布式仿真系统,目前只在基于agent的群体仿真中验证成功,具有一

定的应用范围。

b) 对于一开始专注人体仿真进而转向群体仿真的研究者来说,需要花费一定的时间学习 HLA 标准的相关理论及接口规范,需要对 HLA 的体系结构有全面的了解。

1.2.2 其他的分布式群体仿真体系结构

Quinn 等人^[18]在由 11 台桌面电脑组成的 PC 群集系统上实现了行人仿真。系统采用的是 manager/worker 主从结构,manager 负责系统中其他模块之间的通信,读取建筑布局并广播给 worker,同时接收场景信息,并将每个个体的信息发送到相应的 worker 进程。

Vigueras 等人^[19]提出了一种可扩展的群体仿真架构。文中指出:群体仿真可以作为虚拟环境的一个特例,即满足虚拟环境中的 Avatars 是由智能 agents 替代用户驱动实体时的一个特例。

中国科学院计算机研究所的 Mao 等人^[20]研究了一种针对连续性群体模型(continuum crowd)的有效分布式并行仿真方法。

其他分布式群体仿真体系结构主要是不直接使用现有的分布式虚拟环境平台,而只是依据分布式虚拟环境的相关理论,通过使用并行通信接口 API 来实现的并行群体仿真系统,系统在实现上具有一定的灵活性。

1.3 基于 GPU 通用计算的并行群体仿真体系结构

GPU 作为图形运算的专用设备,包含单指令多数据(SIMD)处理单元集合,而 SIMD 处理方式又具备较好的并行特性。迄今为止,由于 GPU 图形处理单元的高性价比、硬件的可编程性以及巨大的计算能力和速度,使得 GPU 应用到了除 3D 图形渲染之外的通用计算领域。将 GPU 作为高性能通用并行流式处理器,使用 GPU 的通用计算功能来实现并行计算包括两种方式:a)在可编程像素处理阶段,使用基于纹理的 GPU 通用计算方式;b)基于 CUDA 统一计算架构的通用计算方式。CUDA 统一计算架构是由 NVidia 公司提出的一个新的基于 C 语言的通用 GPU 计算框架,在基于 GPU 的群体仿真中将逐步取代早先的通用 GPU 计算方式。

表 2 是一个关于基于 GPU 的并行群体仿真的主要文献的归类。

表 2 基于 GPU 的并行群体仿真分类

基于 GPU 的系统名称	GPU 使用方式	群体行为模型	空间几何表示	相关文献
SugarScape	GPU	SugarScape 行为模型	二维平面	文献[21]
AGGPU	GPU	导航行为 (steering behavior)	三维立体	文献[22, 23]
Flocking	CUDA	群聚行为模型 (flocking hoid)	二维平面	文献[25~27]
Virtual Marathon	CUDA	马拉松行为模型 (Marathon behavior)	三维立体	文献[28]

1.3.1 基于 GPU 的群体仿真体系结构

D'Souza 等人^[21]在 GPU 条件下实现了二维空间的基于 agent 的模型(ABM)群体仿真系统 SugarScape。在该系统中,agent 的属性变量存储在 agent 纹理中,环境的状态信息存储在

环境纹理当中,环境和 agent 状态纹理由顶点渲染器逐步更新。为了使更新持续贯穿整个仿真过程,同时采用了 ping-pong 方法。

Richmond 等人^[22]在 D'Souza 等人的基础上提出了一种在 GPU 上实现的基于 agent 模型的三维仿真方法 ABGPU。该方法能够将群体仿真的行为和群体动画的表示很好地结合在一起,利用 GPU 原生的并行能力,避免了仿真与渲染阶段之间数据的传输,对数千人口规模仿真达到了实时交互的性能。Richmond 等人^[23]继续扩展了其工作,提出了一种 AGGPU 框架,改进了三维 ABGPU 群体仿真的系统性能,同时通过提供一个类似 CPU 功能函数库的通用系统库,使研究人员可以对 agent 行为进行规格描述和脚本化配置。

1.3.2 基于 CUDA 的群体仿真体系结构

Aaby 等人^[24]提出一种基于 agent 的仿真模型。通过有效的延迟隐藏机制来适应当前异构并行计算平台的层次化组织方式,在多个网络化连接的 GPU 硬件结构中能够达到 CPU 基准执行的四个数量级的加速。

Erra 等人^[25]提出一种基于 GPU 的分布式行为模型仿真,该方法使用了最新的 GPU 编程模式 CUDA。系统使用静态网格数据结构对虚拟环境进行分区,对单元网格中的所有 agent 进行哈希索引排序,这样可以快速查找找到同一个单元网格内的所有 agent。在设定每个 agent 的近邻为七个个体、单元网格大小为 9 时,对 10 万个 boids 的仿真能够达到 60 fps 的速率。

Passos 等人^[26]提出一种基于 CUDA 的群体仿真体系结构。系统使用近邻矩阵,所有个体的全部信息组织在矩阵当中,最终映射为 GPU 的纹理数组。近邻矩阵通过使用对二维空间向量的拓扑排序方法来维持每一帧中实体位置变化后矩阵中实体的近邻关系,以保证矩阵的左上角单元内的实体 X、Y 坐标值最小,矩阵右下角单元内的实体 X、Y 坐标值最大。该排序方法名为奇-偶排序过程,一共分为四个步骤:矩阵偶数行排序;矩阵奇数行排序;矩阵偶数列排序;矩阵奇数列排序。通过该框架,可以对类似的鸟群聚集行为或连续流群体行为进行仿真,在群体规模为 65 536 个时,系统能达到 280 fps 的帧率,群体规模为 1 048 576 个时,也能达到 25 fps 的帧率,优于使用 CPU 进行群体仿真。

Husselmann 等人^[27]在鸟群行为的基础上提出一种多群落鸟群的 GPU 并行仿真方法。

Yilmaz 等人^[28]提出一种基于 CUDA 的虚拟马拉松群体仿真方法。通过在 CUDA 环境下实现基于模糊逻辑的 AI 算法,实现了更加真实的群体行为。

Scogings 等人^[29]提出一种混合基于网格和基于 agent 的近邻更新和查找方法,在 CUDA 环境下对捕食者-猎物种群模型进行了仿真模拟,其缺点是使软件和数据结构变得复杂且占用更多的内存,优点是便于空间 agent 的碰撞追踪,具有更高的运行速度和性能。

1.4 三种体系结构分析

通过以上三种并行群体仿真体系结构的分析,得出它们之间的优劣对比,如表 3 所示。

通过对比分析可以发现,GPU 和多核 CPU 都具有很好的

并行运算性能,且 GPU 的并行特性在某些方面比多核 CPU 略胜一筹,硬件结构的高内聚性使内部处理单元之间的通信具有高带宽、高速率特性,而分布式系统能够满足系统问题规模扩大对可扩展性的需求。随着硬件性能不断提升,仅仅使用其中一种体系结构还不能完全发挥现有硬件的性能,可以通过综合 GPU 的高内聚并行和分布式系统的可扩展性,或通过综合多核 CPU 和分布式体系的体系结构,有效地提高硬件的利用效率,从而解决由于系统规模增大而随之产生的性能瓶颈问题。

表3 三种并行群体仿真体系结构分类对比

体系结构类型	硬件使用方式	优势	局限性
基于 CPU 或多核 CPU 的并行群体仿真	CPU/MultiCore CPU	高通信带宽	问题规模受限于系统线程数,扩展性差
基于分布式系统的并行群体仿真	多机分布式系统	高可扩展性	低通信带宽,通信延迟对系统性能影响大
基于 GPU 的并行群体仿真	GPU、CUDA	高通信带宽、SIMD 高度并行	系统的实现受到 GPU 的运算方式和显卡所支持 CUDA 的运算能力的约束,可扩展性有限

2 并行群体仿真的关键算法研究

大规模群体的并行仿真可以借鉴现有的分布式虚拟环境和 N-Body 并行仿真方面的部分研究成果。如果把群体并行仿真问题近似看做 N-Body 问题,则对于基于 agent 的群体仿真来说,最简单地看待 agent 之间通信的方式是每个 agent 都与其他所有 agent 之间通信,也称做全配对通信(all pair),发生在 agent 之间的通信复杂度为 $O(n^2)$ 。这种方式在采用基于 CPU 的对逐个 agent 进行串行迭代运算的库和工具中应用普遍。为了达到更高的效率,通常的做法是对虚拟环境进行分区来降低问题复杂度,减小计算的规模。

虚拟环境的分区方法包括集中式虚拟环境分区方法、复制式虚拟环境分区方法、按空间区域分割的虚拟环境分区方法和按 agent 分割的虚拟环境分区方法四种。其中,集中式的分区方法可扩展性差,受到问题规模的限制;复制式虚拟环境虽然能避免问题规模的扩大带来的通信瓶颈,却带来了维持多个复制环境一致性的开销。所以现有的群体仿真环境分区方法主要是按空间区域分割和按 agent 分割两种方法。

2.1 按空间区域分割的虚拟环境分区方法

按空间区域的分割方法主要包括均匀网格的分割方法、基于矩形区域的分割方法和基于凸包的分割方法。

2.1.1 均匀网格的分割方法

均匀网格(uniform grids)空间分割方法的思想是:使用一个规则网格覆盖空间,该网格将空间分割成相同大小的空间区域或网格单元(grid cell),其空间中的每个对象都与相应的覆盖网格单元相关联。

Green^[30]使用了均匀网格的思想对粒子系统的空间进行划分,其假设网格单元的大小为粒子对象的大小(粒子直径大小),这使得每个粒子对象最多只能与有限的网格单元有交叉,如果粒子之间不重叠,三维空间中每个网格单元最多包含

四个粒子对象。

Joselli 等人^[31]提出一种近邻网格数据结构的群体仿真方法,是一种三维的均匀网格的分割方法。所有的实体信息存储在一个三维数组中,数组中每一个单元包含单个个体的所有数据信息。这样的三维数组一共需要两个,其中一个三维数组用于存储实体对象的位置坐标向量,另一个则用于存储实体对象的方向向量。

2.1.2 基于矩形区域的分割方法

基于矩形区域的分割方法的思想是:使用规则的矩形对虚拟环境进行空间划分,各矩形的大小可以不等。

Lozano 等人^[32]提出一种基于遗传算法的分布式语义数据库群体仿真模型。该方法从一组初始的 P 染色体祖先集开始,通过杂交和变异等进化过程,最终在满足作者提出的适应函数最小化要求时算法达到收敛。其中染色体为一个包含 k 维的最小包含矩形(MBR)的整型数组,每个 MBR 是一个包含 $[x_{\min}, x_{\max}, y_{\min}, y_{\max}]$ 的四元组,定义了虚拟环境的一个矩形区域。

2.1.3 基于凸包的分割方法

基于凸包的分割方法的思想是:使用对分区内 agent 的凸包多边形对虚拟环境进行空间划分。

Vigueras 等人^[33]提出一种基于凸包的虚拟环境分区方法,在系统实现中使用了 QHull 算法。该方法可以有效地减少被各个服务器节点控制的虚拟环境区域,最后通过实验结果证明了分区的形状对改进分区方法的性能具有重要影响。

2.2 按 agent 分割的虚拟环境分区方法

现有的按 agent 的分割方法主要是基于聚类的环境分区方法。Wang 等人^[34]提出一种基于 K-means 聚类的群体仿真环境分区算法。该方法的主要思想是假设每个 agent 有自己的感知区域(area of interest),无论 agent 是否属于相同的服务器节点,感知区域交叉的 agent 之间必然存在消息通信;为了最小化通信代价,对环境的分割必须满足不同服务器节点之间感知区域相互交叉的 agent 数目控制在最小化。当一次聚类达到一定程度的负载失衡时,通过重新计算一次新的聚类可以达到很好的负载均衡效果。实验证明,聚类分析方法在个体以组为行动单位的群体行为仿真中能够发挥更好的效用。K-means 聚类目标函数如式(1)所示,其中: J 为目标代价函数; n 为 agent 数目; k 为聚类服务器节点数目; c_j 为第 j 个聚类服务器节点的中心位置; P_i 为第 i 个 agent 的位置坐标,聚类的目的就是为了使 J 达到最小化。

$$J = \sum_{j=1}^k \sum_{i=1}^n \|P_i - c_j\|^2 \quad (1)$$

2.3 小结

按空间区域分割的分区方法中,均匀网格的分区方法广泛应用于 GPU 并行体系结构群体仿真中;基于矩形、凸包等方法则被应用在分布式体系结构的群体仿真中。按 agent 分割的分区方法中,K-mean 聚类方法主要应用在分布式体系结构的群体仿真中。

3 并行群体仿真性能评估

并行系统度量指标包括绝对度量指标和相对度量指标。

由于绝对度量指标很难度量系统的可扩展性和预测系统性能以及指导系统设计,因此为了满足系统设计和应用的需要,本文采用相对度量指标对系统进行度量分析。使用高性能度量指标加速比 S 对并行系统进行度量分析,同时引出并行效率 E 和扩展性 L 两个相对度量指标来评估并行系统的可扩展性^[35]。

式(2)为并行效率公式,其中 S 为加速比, n 为系统的节点数。

$$E = \frac{S}{n} \quad (2)$$

式(3)为扩展性公式,其中 E 为给定程序的系统效率; E' 为忽略了系统开销的系统效率; T_o 为包括同步和通信的系统总并行开销; T_1 为串行任务开销; T_n 为并行任务开销。

$$L = \frac{E}{E'} = \frac{T_1 / (T_n + T_o)}{T_1 / T_n} = \frac{1}{1 + T_o / T_n} \quad (3)$$

一个扩展性好的系统,并行效率随节点数的增加而保持常数,扩展性指标接近于1,可以通过这两个指标来验证系统的并行效率和扩展性。

在并行群体仿真中,环境分区的负载均衡方法能够有效地降低并行仿真系统的性能瓶颈,对提高并行系统的并行效率和扩展性具有促进作用。负载均衡的环境分区方法是群体仿真环境分区方法研究的趋势。

Lozano 等人^[32]提出了一个负载均衡的代价函数模型,如式(4)所示。

$$H(P) = \omega_1 \times \alpha(P) + \omega_2 \times \beta(P), \omega_1 + \omega_2 = 1 \quad (4)$$

其中: P 为一种分区方案; $\alpha(P)$ 为分区方案 P 中所有的感知区域与两个及两个以上的分区区域有交叉的 agent 的总和; $\beta(P)$ 为每个分区区域 agent 数与分区区域平均 agent 数之间的标准差; ω_1 、 ω_2 分别为各自的权重。

上述负载均衡模型忽略了个体与环境的交互运算代价问题。在并行群体仿真中,考虑到环境复杂度对群体碰撞检测可能带来的影响,就如何评价群体仿真负载均衡的水平,本文提出一种并行群体仿真的负载均衡模型,该模型综合了个体与个体以及个体与环境交互两方面的因素,如式(6)所示。

$$P = \frac{\sum_{i=0}^{N-1} (SS_i + SE_i)}{\sum_{i=0}^{N-1} M_i} \quad (5)$$

$$E(X) = \frac{1}{N} \sqrt{\sum_{i=0}^{N-1} \left(\frac{SS_i + SE_i}{M_i} - P \right)^2} \quad (6)$$

式(5)中的 P 为并行群体仿真系统的平均载荷率;式(6)中的 $E(x)$ 为载荷率偏差; N 为运算节点数; SS_i 为第 i 个运算节点的实体与实体的交互运算量; SE_i 为第 i 个运算节点的实体与环境的交互运算量; M_i 为第 i 个运算节点的满荷运算量。负载均衡的目标为系统的载荷率偏差 $E(x)$ 最小化。

4 趋势与展望

通过查阅文献^[28,36,37]得出以下结论,并行群体仿真未来可能取得突破的研究方向有:

a)随着硬件性能的快速提升,硬件体系结构的更新变化对如何提高群体仿真的性能带来了机遇,也提出了挑战。如何

整合现有的高性能硬件环境,实现大规模复杂群体的高分辨率实时仿真,是并行群体仿真研究的方向。群体仿真研究包括环境建模、行为建模、动画渲染等过程。在基于分布式系统与多核 GPU 的混合体系结构中,研究群体仿真各过程的算法并行化以及各并行算法在混合体系结构中相互融合的方法,对提高群体仿真性能前景可观。Vigueras 等人^[36]在此方向进行了部分尝试,将群体仿真系统通过分区方法映射到分布式系统上,同时在各主机 GPU 上实现了个体的碰撞检测运算。

b)基于 agent 的群体仿真是人群仿真的趋势,智能化的群体是未来的群体仿真方向。采用基于 agent 的群体仿真模型,给 agent 个体应用 AI 算法,达到智能化群体仿真的目的,能够展示更加丰富的群体仿真内容,构建智能的群体社区将具有更为广泛的应用价值。Yilmaz 等人^[28]将 CUDA 环境下实现的基于模糊逻辑 AI 算法应用到马拉松群体仿真中,更多智能化算法如自动推理机、遗传算法、决策网络算法等如何应用在并行群体仿真环境中是有潜力的研究方向。

c)如何有效地将高真实性的三维群体表示与群体仿真并行算法研究紧密结合在一起,而不单侧重于效率忽视群体的行为复杂度以及环境和个体的真实度表示,研究出实时高效的并行群体仿真模型是群体仿真中需要解决的问题。迄今为止,少见使用并行绘制的方式来实现群体仿真动画渲染的研究,国内外对并行绘制方法已有深入研究^[38],群体仿真过程的诸多算法如路径导航^[37]、碰撞检测、模糊逻辑 AI 算法等都有并行化实现,随着群体规模的扩大以及渲染场景和个体三维模型真实度的提高,群体的动画渲染将面临瓶颈。将并行绘制方法应用到群体仿真中能够有效解决此问题。

5 结束语

并行计算可以充分发挥当今高性能硬件的优势,有效提高群体仿真的效率,满足大规模人群实时仿真的可扩展性与计算性能需求。本文对并行群体仿真的原理、进展和方法进行了综述。在以后的工作中,将主要对混合体系结构下的并行群体仿真以及群体仿真中关键算法并行化方法进行深入研究。

参考文献:

- [1] REYNOLDS C W. Flocks, herds, and schools: a distributed behavioral model[C]//Proc of the 14th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 1987: 25-34.
- [2] HELBING D, FARKAS I, VICSEK T. Simulating dynamical features of escape panic[J]. *Nature*, 2000, 47(28): 487-490.
- [3] PELECHANO N, O'BRIEN K, SILVERMAN B, et al. Crowd simulation incorporating agent psychological models, roles and communication[C]//Proc of the 1st International Workshop on Crowd Simulation. 2005: 21-30.
- [4] YU Qin-xin, TERZOPOULOS D. A decision network framework for the behavioral animation of virtual humans [C]//Proc of ACM SIGGRAPH/Eurographics Symposium on Computer Animation. 2007: 119-128.
- [5] MUSSE S R, THALMANN D. A model of human crowd behavior: group inter-relationship and collision detection analysis[C]//Proc of

- Workshop of Computer Animation and Simulation of Eurographics. 1997:39-52.
- [6] PELECHANO N, BADLER N. Modeling crowd and trained leader behavior during building evacuation [J]. *IEEE Computer Graphics and Applications*, 2006, 26(6): 80-86.
- [7] DURUPINAR F, ALLBECK J, PELECHANO N, *et al.* Creating crowd variation with the OCEAN personality model [C]//Proc of the 7th International Joint Conference on Autonomous Agents and Multi-Agents Systems. New York: ACM Press, 2008: 1217-1220.
- [8] McDONNELL R, NEWELL F, O'SULLIVAN C. Smooth movers: perceptually guided human motion simulation [C]//Proc of ACM SIGGRAPH/Eurographics Symposium on Computer Animation. New York: ACM Press, 2007: 259-269.
- [9] McDONNELL R, LARKIN M, DOBBYN S, *et al.* Clone attack! Perception of crowd variety [J]. *ACM Trans on Graphics*, 2008, 27(3): 1-8.
- [10] DOBBYN S, HAMILL J, O'CONOR K, *et al.* Geopostors: a real-time geometry impostor crowd rendering system [C]//Proc of Symposium on Interactive 3D Graphics and Games. New York: ACM Press, 2005: 95-102.
- [11] KOTA R, BANSAL V, KARLAPALEM K. System issues in crowd simulation using massively multi-agent systems [C]//Proc of the 2nd International Workshop on Massively Multi-Agent Systems. 2006.
- [12] Van den REYNOLDS C. Big fast crowds on PS3 [C]//Proc of ACM SIGGRAPH Symposium on Videogames. New York: ACM Press, 2006: 113-121.
- [13] Van den BERG J, PATIL S, SEWALL J, *et al.* Interactive navigation of individual agents in crowded environments [C]//Proc of ACM Symposium on Interactive 3D Graphics and Games. 2008: 139-147.
- [14] PETTY M D, MCKENZIE F D, GASKINS R C, *et al.* Developing a crowd federate for military simulation [C]//Proc of the Spring Simulation Interoperability Workshop. 2004: 483-493.
- [15] LEES M, LOGAN B, THEODOROPOULOS G K. Agents, games and HLA [J]. *Simulation Modeling Practice and Theory*, 2006, 14(6): 752-767.
- [16] LOW H M Y, CAI Wen-tong, ZHOU Sui-ping. A federated agent-based crowd simulation architecture [C]//Proc of the 21st European Conference on Modeling and Simulation. 2007: 188-194.
- [17] CHEN Dan, WANG Li-zhe, WU Xiao-min, *et al.* A hybrid modelling and simulation of huge crowd over a hierarchical grid architecture [C]//Proc of Future Generation Computer Systems. 2012.
- [18] QUINN M, METOYER R, HUNTER-ZAWORSKI K. Parallel implementation of the social forces model [C]//Proc of the 2nd International Conference in Pedestrian and Evacuation Dynamics. 2003: 63-74.
- [19] VIGUERAS G, LOZANO M, PEREZ C, *et al.* A scalable architecture for crowd simulation: implementing a parallel action server [C]//Proc of the 37th International Conference on Parallel Processing. 2008: 430-437.
- [20] MAO Tian-lu, JIANG Hao, LI Jian, *et al.* Parallelizing continuum crowds [C]//Proc of the 17th ACM Symposium on Virtual Reality Software and Technology. New York: ACM Press, 2010: 231-234.
- [21] D'SOUZA R M, LYSENKO M, RAHMANI K. SugarScape on steroids: simulating over a million agents at interactive rates [C]//Proc of Agent Conference. 2007.
- [22] RICHMOND P, ROMANO D. Agent based GPU, a real-time 3D simulation and interactive visualisation framework for massive agent based modelling on the GPU [C]//Proc of International Workshop on Supervision. 2008.
- [23] RICHMOND P, Dr ROMANO D M. A high performance framework for agent based pedestrian dynamics on GPU hardware [C]//Proc of EUROESIS ESM 2008. Le Havre: Universite Heavre, 2008.
- [24] AABY B G, PERUMALLA K S, SEAL S K. Efficient simulation of agent-based models on multi-GPU and multicore clusters [C]//Proc of the 3rd International ICST Conference on Simulation Tools and Techniques. 2010: 1-10.
- [25] ERRA U, FROLA B, SCARANO V. A GPU-based method for massive simulation of distributed behavioral models with CUDA [C]//Proc of the 22nd Annual Conference on Computer Animation and Social Agents. 2009.
- [26] PASSOS E, JOSELLI M, ZAMITH M, *et al.* Supermassive crowd simulation on GPU based on emergent behavior [C]//Proc of the VII Brazilian Symposium on Computer Games and Digital Entertainment. 2008: 81-86.
- [27] HUSSELMANN A, HAWICK K. Simulating species interactions and complex emergence in multiple flocks of boids with GPUs [C]//Proc of IASTED International Conference on Parallel and Distributed Computing and Systems. 2011: 100-107.
- [28] YILMAZ E, DSILER V, YARDIMCI Y C. The virtual marathon: parallel computing supports crowd simulations [J]. *IEEE Computer Graphics & Applications*, 2009, 29(4): 26-33.
- [29] SCOGINGS C J, HAWICK K A. Optimal data structures for spatially localised agent-based automata and hybrid systems [C]//Proc of International Conference on Artificial Intelligence and Soft Computing. 2012.
- [30] GREEN S. CUDA particles [R]. 2007.
- [31] JOSELLI M, PASSOS E B, ZAMITH M, *et al.* A neighborhood grid data structure for massive 3D crowd simulation on GPU [C]//Proc of SBGAMES. 2009: 121-131.
- [32] LOZANO M, ORDUN A J M, CAVERO V. A genetic approach for distributing semantic databases of crowd simulations [C]//Proc of the 21st International Parallel and Distributed Processing Symposium. [S. l.]: IEEE Computer Society, 2007: 1-8.
- [33] VIGUERAS G, LOZANO M, ORDUN A J M, *et al.* A partitioning method based on convex hulls for crowd simulations [EB/OL]. (2009). <http://www.uv.es/grimo/publicaciones/jp2009.pdf>.
- [34] WANG Yong-wei, LEES M, CAI Wen-tong, *et al.* Cluster based partitioning for agent-based crowd simulations [C]//Proc of Winter Simulation Conference. 2009: 1047-1058.
- [35] 杨晓东, 陆松, 牟胜梅. 并行计算机体系结构技术与分析 [M]. 北京: 科学出版社, 2009.
- [36] VIGUERAS G, ORDUN A J, LOZANO M. A GPU based multi-agent system for real-time simulations [C]//Proc of the 8th International Conference on Practical Applications of Agents and Multi-Agent Systems. 2010: 15-24.
- [37] BLEIWEISS A. GPU accelerated pathfinding [C]//Proc of the 23rd ACM SIGGRAPH/EUROGRAPHICS Symposium on Graphics Hardware. 2008: 65-74.
- [38] 石教英. 分布并行图形绘制技术及其应用 [M]. 北京: 科学出版社, 2010.