

一种用于实时体绘制系统的自适应采样算法

夏放怀, 沈振康¹, 唐朝京¹, 湛海新^{1,2}
Juergen Hesser², Bernd Vettermann², Reinhard Manner²

(1. 国防科技大学电子工程学院, 湖南长沙 410073;

2. Lehrstuhl für Informatik V, Universität Mannheim, B6, 26, D 68131 Mannheim, Germany)

摘 要: 本文讨论了一种用于体绘制的、支持基于空间跳跃和自适应光线终止等优化算法的高效硬件结构及一种加速半透明物体绘制的新技术。这种硬件结构用于绘制规模为 256^3 体元的数据时可达达到 70Hz 的帧频, 但这样的帧频只是在体元的透明度都被置为 0 (完全透明) 或 1 (完全不透明) 的情况下才能取得。引入新的加速技术后, 绘制半透明物体的帧频也可接近上述数值。新的加速技术采用自适应采样步长减少体绘制过程中冗余的体元重采样, 从而使算法的复杂度和对存储器的带宽要求大大减小。新的加速技术对半透明体数据可取得高达 4.7 倍的加速比。采用新的加速技术需要在开始交互绘制前对体数据进行距离编码。优化后的距离编码对有 8 兆体元的体数据的预处理时间仅需要 8~30 秒。

关键词: 体绘制; 可视化; 算法; 自适应; 距离编码

中图分类号: TN919 **文献标识码:** A **文章编号:** 0372-2112 (2002) 03-0367-05

An Adaptive Sampling Based Algorithm for Real Time Volume Rendering System

XIA Fang-huai¹, SHEN Zhen kang¹, TANG Chao jing¹, CHEN Hai xin^{1,2}
Juergen Hesser², Bernd Vettermann², Reinhard Manner²

(1. Institute of Electronic Engineering, National Univ. Of Defense Tech., Changsha, Hunan 410073, China;

2. Lehrstuhl für Informatik V, Universität Mannheim, B6 26 D-68131 Mannheim, Germany)

Abstract: This paper presents a new volume rendering acceleration algorithm for semi-transparent volume objects. The volume rendering pipeline processor VGE provides full support for the popular ray casting acceleration techniques such as adaptive ray termination and space leaping, thus it achieves frame rate up to 70Hz. Nevertheless, such high frame rates can only be achieved in the case that the volume elements are mapped either opaque or transparent. The new acceleration algorithm is designed to relax the dependence of the frame rate on the selected opacity mapping functions. The new algorithm uses an adaptive resampling strategy to replace the equidistant resampling, hence the time complexity is reduced and the memory bandwidth requirement is lowered. Experimental results show that the new algorithm can accelerate the rendering of semi-transparent volume objects by up to 4.7 times. Like space leaping, the new algorithm needs to encode the adaptive sampling interval in a preprocessing stage. The preprocessing can be implemented in less than 30 seconds for volume data with 8 million voxels by heuristic algorithmic optimizations.

Key words: direct volume rendering; visualization; algorithm; adaptive; distance coding

1 引言

直接体绘制是显示体数据的标准技术。体数据最初的来源是计算机断层成像或核磁共振成像系统所生成的顺序堆积的图象序列。二维图象平面上的像素 (pixel) 被称为三维体数据的体元 (voxel), 因此体元实际上是三维体数据网格上的一个节点。随着应用领域的拓展, 体数据也可以用科学模拟计

算、物理采样等方法获得, 而体数据的网格既可以是规则的, 也可以是不规则的。相较于计算机图形领域里常用的基于几何模型的绘制方法, 体绘制的优势在于, 通过实时地改变透明度映射函数, 可以允许用户选择性地绘制同一数据场中具有不同物理属性的物体, 而不必烦琐地用几何建模的方式来描述和绘制。这种特性使得体绘制成为一些应用领域里不可缺少

少的技术,如交互式外科手术规划、训练系统,工业产品质量检测等。但体绘制的优点也给其实时实现带来了困难。因为在绘制过程中,从观察者发出的每条光线在数据场中都被均匀地重新采样。为避免产生锯齿效应,采样间隔通常小于或等于体元网格的尺寸。这意味着,算法的复杂程度为 $O(N^3)$, N^3 为体数据尺寸。实用的体数据尺寸通常为 $256^3 \sim 1024^3$ 。用不同的透明度映射函数,对有 256^3 个体元的中等复杂程度的体数据,用 CPU 主频为 1GHz 级的 PC 生成一帧 256×256 像素的图像通常需要几秒到几分钟,这远远不能满足实时应用的要求。

最直接的加速体绘制的途径是只绘制那些对最后的绘制结果真正有贡献的体元。空间跳跃^[1,2]和自适应光线终止^[2]是两种最常用的方法,前者忽略那些透明的“空”体元,故对场景中有大量空体元的数据很有效;后者则适合于场景中体元的不透明度很高的情形,它在当前光线的累积不透明度高于一定阈值,以至后续的体元对图象的贡献变得可以忽略时,自适应地停止对该光线的继续处理。但是对半透明物体却一直没有很有效的加速方法。

另一个加速体绘制的途径是采用专用硬件加速。第一个可以运行的用于实时体绘制的专用硬件原型系统是曼海姆大学开发的 VIRIM^[3]。VIRIM 支持虚拟场景的透视观测,能够产生真实阴影。VIRIM 使用海德堡光线跟踪模型^[4],即用两个固定位置的光源产生真实阴影。但 VIRIM 是绘制算法在硬件上的直接且固定的映射,没有支持空间跳跃和自适应光线终止等优化算法的灵活性,整个系统比较庞大。稍后,图宾根大学开发了 VIZZARD^[5]。VIZZARD 是第一个支持算法优化的体绘制加速硬件系统。VIZZARD 的主要缺点是在实时交互绘制前需要对体数据进行几分钟到几十分钟的预处理,并且绘制帧频明显依赖于数据特性和透明度映射函数的选择。目前最成功的体绘制加速硬件系统是 VolumePro^[6]。VolumePro 是三菱电器研究实验室开发的体绘制 PCI 加速板,能够以 30Hz 的帧频绘制 256^3 的体数据。其核心部件是 VG500 芯片,它建立在纽约州立大学 1982 年开始的 Cube 项目所提出的第四代 Cube^[7] 体系结构上。VG500 内部集成了 8 个并列互连的管道处理器,实时地执行 shear warp 算法。但其用于支持 shear warp 算法的特殊的存储器子系统设计使它不能以透视方式绘制场景。

最近我们提出了一种新的体绘制加速硬件结构—VGE^[8]。作为 VIRIM 的下一代,VGE 提供对算法优化的有效支持。对不透明物体,它只需要 VolumePro 硬件规模的 10% 就能达到与 VolumePro 相同的绘制帧频。但其缺点是帧频依赖于数据特性和透明度映射函数,对半透明物体的绘制速度较慢。为解决这一问题,我们将提出一种加速半透明物体绘制的新技术。

本文的下一节介绍新的体绘制加速硬件 VGE 的系统结构;第 3 节讨论解决 VGE 的帧频依赖于数据特性和透明度映射函数这一问题的新算法;第 4 节是新算法的实验结果;最后是结束语。

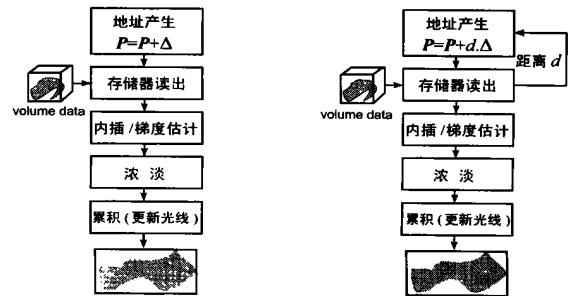
2 新的体绘制加速结构 VGE

典型的体绘制算法可以分为三个步骤。首先光线从观测

者发出,通过成像平面上的每个像素点射向虚拟场景,在沿这些光线均匀分布的采样点处,用三线性内插算法由采样点附近的体元求出采样点处的不透明度;然后将采样点处灰度的局部梯度矢量作为法向量,用 Phong 光照模型计算采样点的浓淡;最后用下面离散化的体绘制公式将各采样点处的浓淡值组合起来,得到对应像素的颜色。

$$I(x_k) = \sum_{k=0}^n b_k \prod_{j=k+1}^n \theta_j \quad (1)$$

上式中 b_k 是采样点 k 处用 Phong 模型求得的不透明度, θ_j 是采样点 j 的透明度, $I(x_k)$ 是组合完采样点 k 的贡献后光线的累积强度。



(a) 标准体绘制流程 (b) 加入空间跳跃后的体绘制流程
图 1 体绘制流程

这种体绘制算法可以很方便地映射到流水线管道硬件结构上来实现。根据计算复杂程度,上面各步骤被分配到管道的一个或多个节点上执行,以便管道各节点具有均衡的负载,减少管道阻塞。图 1(a) 是体绘制在流水线管道上的实现流程。这种结构不支持空间跳跃等算法优化技术。为了支持空间跳跃,在从内存中读取采样点附近的体元灰度信息的同时,也必须读取用距离编码^[1]预先编码好的用于空间跳跃的距离值 d ,然后用该距离值 d 产生下一个采样点的地址,如图 1(b) 所示。但是,对算法优化技术的支持导致了流水线管道中信息反馈,即后续处理依赖于当前处理的结果。在距离值 d 未取得之前,不能确定下一个采样点的位置。通常数据从存储器中读出需要几个时钟周期,在这种情况下流水线处理管道处于等待状态,这会大大降低处理器的运行效率。在流水线管道处理器结构中,称这种情况为数据或分支灾难 (data/branch hazard)。

为了既能支持算法优化,又能避免数据灾难,我们采用多线程的策略。线程是一个拥有自己独立的处理状态和寄存器状态,但和其他线程共享全局变量和地址空间的处理进程。每条光线被视为一个线程,它们按照一定的循环规则被激活,每条光线的关键数据如光线与数据场相交的起点与终点及当前采样点位置都被作为线程的状态变量保存。这样,在一条光线在等待所需的距离信息以确定下一个采样点位置时,其他已准备好的光线可以取代当前光线在流水处理管道中被处理,从而消除了管道的等待状态,避免了数据灾难。

图 2 是新的体绘制加速引擎 VGE 的系统结构框图,其与 VIRIM 结构的主要差别是在采样点地址计算单元中加入了一

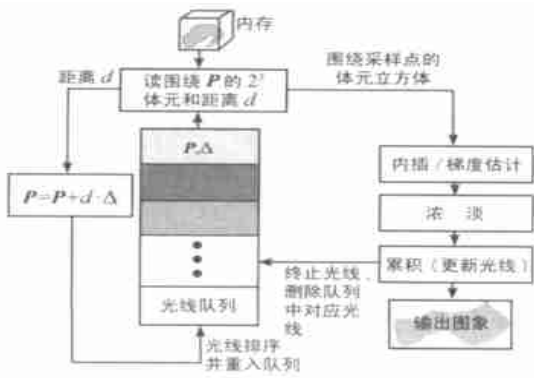


图 2 体绘制加速引擎 VGE 的系统结构

个光线队列。这个新的地址产生单元就是完成多线程切换的硬件。在光线队列中,所有与光线有关的信息,如当前采样点位置 P 以及两个采样点之间的最小距离矢量 Δ 等,都存储在对应线程的数据单元中。在管道处理器运行过程中,总是光线队列中最顶端的那条光线被弹出队列,围绕着其当前采样点 P 的 $2 \times 2 \times 2$ 体元立方体以及用于空间跳跃的距离值 d 被读出。在围绕着当前采样点 P 的 $2 \times 2 \times 2$ 体元立方体被送到管道的后续节点处理的同时,距离值 d 则被用于更新该光线的当前采样点,然后该光线被重新送回光线队列。为了增加读出数据的相关性,减少存储器换页操作,光线队列初始化时总是被填入一个 8×8 的光线束,即在图象平面上与一个有 8×8 像素的子图对应的光线束,同时,在光线的当前采样点更新后,它们也被按其当前采样点到观测者的距离重新排序。自适应光线终止由光线累积计算单元控制,当一条光线的累积不透明度超过规定的门限时,它通知光线队列删除对应的光线。模拟实验表明,这种多线程技术使得管道处理器能以平均高于 90% 的效率运行,只有在光线队列中光线数目少于读取距离值所需的时钟周期和计算光线的下一个采样点所需的时钟周期之和时管道处理器才出现等待状态。

我们用 C 模拟了上面所提出的体绘制管道处理器^[8],在模拟中我们设定每个操作的时钟周期为 10 纳秒。模拟所用的数据是有 $256^2 \times 128$ 体元的人类颌部 CT 数据。为了全面地评估新的硬件结构的性能,我们选用了不同的观测角度、投影方式和透明度映射函数。

模拟试验的结果表明,在体元被映射成不透明或完全透明的情况下,绘制全图所需的采样点数仅为全部 8 兆体元的 10% ~ 15%; 当体元被映射成半透明时,完成绘制所需的采样点数为体元数目的 25% ~ 40%。在最理想的情况下,即体元被映射成不透明或完全透明,并采用平行投影绘制的情况下,系统能取得高达 138Hz 的峰值帧频;而采用透视投影绘制时(40° 视角),由于光线发散的缘故,降低了相邻光线间的相关性,帧频降低到约为平行投影绘制的一半。值得注意的是,系统性能对数据特性和绘制参数(不透明度映射函数、场景设置等)有明显的依赖性。造成这种依赖性的原因是系统对算法优化的支持,而不支持算法优化的系统,如 VIRIM 和 Volume Pro,不存在这种情况。具体地说,我们提出的 VGE 体绘制加速管道结构提供了对空间跳跃和自适应光线终止的支持,二者

使系统有效地减少了对体数据中可忽略的完全透明和不透明体元的处理,但半透明的体元却没有被考虑。因此下一步我们将考虑如何减少绘制半透明体元所需的运算。

3 加速半透明的体元的绘制

体绘制通过模拟光线穿过物体时被物体中包含的介质衰减和反射的过程,求得成像平面上像素的强度值。成像过程可用标准的体绘制公式表示成:

$$I(x) = I(x_0)e^{-\tau(x_0, x)} + \int_{x_0}^x \eta(x')e^{-\tau(x', x)} dx' \quad (2)$$

其中 $\tau(x_0, x)$ 为光线衰减系数在区间 $[x_0, x]$ 上的积分,称为区间 $[x_0, x]$ 的光学深度。在体绘制过程中,沿每条光线体数据场被均匀地采样。由以上公式可以导出在两个相邻采样点处光线强度间的关系为:

$$I(x_k) = I(x_{k-1})e^{-\tau(x_{k-1}, x_k)} + \int_{x_{k-1}}^{x_k} \eta(x')e^{-\tau(x', x_k)} dx' \quad (3)$$

定义采样点 k 的透明度 θ_k 和明暗度 b_k 为:

$$\theta_k = e^{-\tau(x_{k-1}, x_k)} \quad (4)$$

$$b_k = \int_{x_{k-1}}^{x_k} \eta(s')e^{-\tau(x', x_k)} dx' \quad (5)$$

即得到:

$$\begin{aligned} I(x_k) &= I(x_{k-1})\theta_k + b_k = (I(x_{k-2})\theta_{k-1} + b_{k-1})\theta_k + b_k \\ &= \dots = \sum_{k=0}^n b_k \prod_{j=k+1}^n \theta_j \end{aligned} \quad (6)$$

这就是前面所提到的离散体绘制公式。这个公式里的积分通常都用矩形数值积分方法求得,且积分区间(也即采样间隔)是等距离的。但是我们可以采用自适应步长采样来减少不透明和半透明区域里采样点数,从而加速体绘制过程。如图 3 示,如果我们用梯形数值积分方法取代矩形数值积分方法,则可以大幅度减少每条光线的平均采样点数。

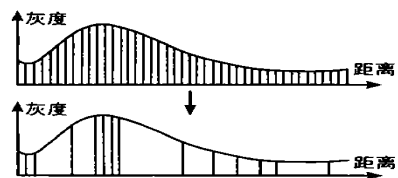


图 3 用梯形数值积分规则取代矩形数值积分规则

剩下的问题是如何选择采样步长以便能方便地控制所绘制的图象质量。我们定义在某个给定方向上从当前采样点开始的以自适应步长 d 为长度的区间上用梯形数值积分取代矩形数值积分的误差为:

$$er(d) = |rect(d) - trap(d)| \quad (7)$$

其中 $rect(d)$ 是矩形规则数值积分值, $trap(d)$ 是梯形规则数值积分值。给定允许的误差 er_{th} , 下面伪代码即可以确定沿给定方向当前采样点到下一个采样点的距离 d :

```
initialize d = 体数据边长的一半
do
```

$$\{$$

$$err(d) = |rect(d) - trap(d)|;$$

$$d-;$$

$$\text{while}(err(d) > err_{th})$$

另有一个要考虑的问题是,在交互应用中,体数据被用户随机地在场景中旋转、移动,这要求编码好的自适应采样步长独立于光线方向。如图 4 所示,在采样点 a 处,沿 ac 方向采样步长可为 7,沿其他方向,如沿 ad 方向,采样步长只能为 6。为确保沿所有的方向,误差都小于误差门限,最小采样步长(此处为 5)应作为采样点 c 所在体元的步长编码值,即步长值应满足: $\delta = \max\{||d^v|| | err(d^v) < \epsilon\}$ 。这样,编码时要求穷尽地检查所有通过当前体元的不同方向的光线,使得自适应采样步长的编码非常地耗时。例如,对有 8 兆体元的人类颌骨 CT 体数据,这种穷尽编码方法需要 4.3 小时,这是实际应用中无法接受的。

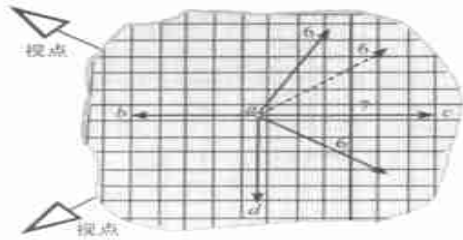


图 4 独立于视点的采样步长选择

我们采用了三个启发式的减少自适应采样步长编码预处理时间的措施。首先,编码开始时步长的初始值被设置为 9 而不是体数据边长尺寸的一半,因为对我们手头里的数据集编码结果的统计表明,最大自适应采样步长不会超过 9;其次,考虑到任何两个相邻体元的自适应采样步长之差总是小于或等于 1,步长的初始值可动态地设为当前被编码体元的 3 维邻域内已被编码的体元中最小自适应采样步长加 1;最后,我们只从 6 个平行于主轴的方向和 8 个对角方向中而不是所有可能的光线方向中选择最小采样步长。以上三个措施可以大大地压缩自适应采样步长编码时间。

4 实验结果

我们用不同数据对新的加速绘制半透明物体的算法进行了测试。表 1 为自适应采样步长编码在优化前后对不同数据的预处理时间,表 2 为采用两种自适应采样步长编码时新算法对不同数据的绘制加速比和图象误差比较。比较采用的基准算法是只用空间跳跃和自适应光线终止加速的绘制算法,测试平台是内存为 512MB 的运行 Suse Linux 6.3 的 PC, CPU 为 800MHz Athlon。

从表 1 可以看出,穷尽式的自适应采样步长编码是相当费时的。对最费时的人类颌骨 CT 数据,预处理时间长达 15512 秒,而采用启发式编码方法,预处理时间减少到约 30 秒。在交

互式应用中,少于 30 秒钟的初始化时间是可以接受的。

表 1 两种编码算法的预处理时间比较

体数据	预处理编码时间(秒)	
	穷尽式编码	启发式编码
人头(256 ³ × 113 体元)	4448.71	20.42
引擎(256 ³ × 113 体元)	3282.87	8.92
人类颌骨(256 ³ × 113 体元)	1551.73	30.77

表 2 两种编码算法对体绘制的加速比及图象误差比较

体数据	穷尽式编码		启发式编码	
	加速比	图象误差	加速比	图象误差
人头	1.661	0.422%	1.667	0.468%
引擎	2.125	0.420%	2.215	0.452%
人类颌骨(外部)	2.973	1.462%	2.999	1.483%
人类颌骨(内部)	4.276	10.190%	4.705	12.36%

表 2 显示出,采用新的自适应采样加速算法,可以对体绘制提速 1.6 到 4.7 倍。对人类颌骨 CT 数据这种有较多半透明区域的体数据,新算法有较高的加速比;对头部 CT 数据这种完全透明体元和完全不透明体元所占比例较高的体数据,新算法提供的加速比相对较低,但对这类数据,空间跳跃和自适应光线终止恰好可以有效地加速绘制过程。因此将新的自适应采样加速算法与空间跳跃和自适应光线终止集成起来,可以使得系统对不同特性的体数据及在选用不同的透明度映射函数时的绘制帧频趋近恒定。此外,表 2 表明,使用启发式自适应采样步长编码时的加速比和图象质量与穷尽式编码时的加速比和图象质量相当接近,因此没有必要使用穷尽式编码。

5 结束语

本文讨论了支持空间跳跃和自适应光线终止等优化算法的流水管道体绘制硬件系统结构 VGE。VGE 采样的支撑多线程切换的光线队列可以使流水管道体绘制硬件能以高于 90% 的效率全速运行,这使得 VGE 只需要一个流水管道处理器就可以达到和内部具有 8 个管道处理器的 VG500 相近的性能。因此 VGE 的硬件远远没有 VG500 那样复杂,而且 VGE 能同时支持平行投影和透视投影。本文提出的基于自适应采样的绘制半透明物体加速算法可以有效地减少体数据中半透明区域里的冗余采样运算,这种算法与空间跳跃和自适应光线终止算法结合可以使 VGE 能取得的帧频基本上不依赖于体数据本身的特性和透明度映射函数的选择。新的加速算法需要对自适应采样步长在预处理阶段进行编码,通过引入启发式优化措施,预处理时间被从几十分钟乃至几个小时减少到少于 30 秒,从而使新算法能够被实际的应用项目所接受。

参考文献:

- [1] K Zuiderveld, A Koning, M Viergever. Acceleration of ray casting using 3D distance transforms [A]. Proceedings of Visualization in Biomedical Computing 1992 [C], October 1992, 1808: 324-335.
- [2] M Levoy. Efficient ray tracing of volume data [J]. ACM Transactions on Graphics, 1990, 9(3): 245-261.
- [3] T Günther, C Poliwoda, C Reinhart, J Hesser, R Männer, H-P Meinz

er, H J Baur. VIRIM: A massively parallel processor for real time volume visualization in medicine [A]. Proc. Eurographics Workshop [C], Oslo, Norway, 1994: 103- 108.

- [4] H - P Meinzer, K Meetz, D Scheoelmann, U Engelmann. The Heidelbergray tracing model [J]. IEEE Computer Graphics & Applications, 1991, 11(6): 34- 43.
- [5] G Knittel, W Sträßer, A PCF-based volume rendering accelerator [A]. 10th Eurographics Workshop on Computer Graphics Hardware [C], Maastricht, 1995: 73- 82.
- [6] Hanspeter Pfister, Jan Hardenbergh, Jim Knittel, Hugh Lauer, Larry Seiler. The VolumePro real time ray-casting system [A]. SIGGRAPH'99, Proceedings of the 26th Annual Conference on Computer Graphics [C], Los Angeles, 1999: 251- 260.
- [7] Pfister H - P . Towards a scalable architecture for real time volume rendering [A]. 10th Eurographics Workshop on Computer Graphics Hardware [C], Maastricht, 1995: 123- 130.

- [8] B Vettemann, J Hesser, R Männer. Solving the hazard problem for algorithmically optimized real time volume rendering [A]. International Workshop on Volume Graphics [C], Swansea, UK, 1999: 171- 184.

作者简介:



夏放怀 女. 1973 年生于湖南桃江县. 1995 年获湖南大学电子工程专业学士学位. 目前在国防科技大学电子工程学院攻读硕士学位. 主要研究兴趣为计算机通讯、计算机图形、数字图象处理等.

沈振康 男. 1936 年 5 月生于上海. 1961 年毕业于哈尔滨军事工程学院. 现为国防科技大学电子工程学院教授, 博士生导师. 长期从事数字图象处理、目标识别、信号处理等方面的教学与科研工作.