

文章编号: 1002-0411(2003)01-056-05

## 基于图象的实时绘制技术

张汗灵 郝重阳 张先勇 雷方元

(西北工业大学电子与信息工程研究所 西安 710072; 西安虚拟现实工程技术研究中心 西安 710072)

**摘要:** 随着计算机技术的发展,对虚拟场景的真实感程度提出了越来越高的要求. 基于图象实时绘制技术是实现虚拟场景建模和实时绘制的新的有效方法,可以克服传统的基于几何绘制方式中建立真实模型的困难和实时绘制计算量大的缺陷. 本文对目前出现的各种基于图象的实时绘制方法进行分类介绍和综述.

**关键词:** 基于图象的建模与绘制 (IBMR); 实时图形生成; 虚拟现实

**中图分类号:** TP317

**文献标识码:** B

### REAL-TIME RENDERING APPROACH BASED ON IMAGE

ZHANG Han-ling HAO Chong-yang ZHANG Xian-yong LEI Fang-yuan

(Institute of Electronic & Information Engineering North Western Polytechnic University; Xi'an Tech. Resercher.

Center of Virtual Reality, xi'an 710072)

**Abstract:** Based on image rendering is a new powerful approach to modeling and rendering of virtual environments. But in geometry-based rendering systems, the acquisition of realistics surface models is difficult and real-time rendering of complex scenes is computationally expensive. This paper discusses the present image-based rendering methods and gives a catalogue of them

**Keywords:** image-based modeling-rendering, real-time graphics generation, virtual reality

### 1 引言 (Introduction)

随着虚拟现实、网络图形应用等技术的高速发展,需求不断激增,图形的实时绘制算法面临的挑战越来越大. 寻找在现有通用计算机平台上对复杂景物的实时真实感绘制方法,解决绘制算法在速度、质量与景物复杂度之间越来越突出的矛盾,已经是计算机图形学界面临的十分紧迫的课题. 为了解决这一技术问题,基于图象的实时绘制技术 (Image-Based Rendering 简称 IBR) 应运而生. IBR 方法已经成为近年来计算机图形学的研究热点. 它广泛应用于计算机可视化、计算机视觉、多媒体等领域. 本文结合我们在 IBR 方面的研究工作,对这一技术进行综述性介绍.

#### 1.1 什么是 IBR

传统意义上的虚拟场景基本上都是基于几何的,就是用数学意义上的曲线、曲面等数学模型预先定义好虚拟场景的几何轮廓,再采取纹理映射、光照等数学模型加以渲染. 在这种意义上,大多数虚拟现

实系统的主要部分是构造一个虚拟环境并从不同的方向进行漫游. 要到这个目标,首先是构造几何模型,其次模拟虚拟照相机在 6 个自由度 [ChenM88] 运动,并得到相应的输出画面. 现有的几何造型技术可以将极复杂的环境构造出来,存在的问题是极为繁琐. 而且在真实感程度、实时输出等方面有着难以跨越的鸿沟.

从已知的图象中合成新视图的 IBR 方法引起了计算机图形学和计算机视觉界中有关研究人员的普遍兴趣,并且在全世界范围内形成了 IBR 的研究热潮. 与传统的基于模型的绘制方法相比,IBR 方法有以下优点:

1) 计算量适中. IBR 方法所需的计算量相对较小,因此适合于工作站以及个人计算机上的虚拟现实应用.

2) 独立于场景复杂度. 交互显示的开销与场景的复杂度无关,只与分辨率有关. 因此 IBR 能用于表现非常复杂的场景.

3) 满意的绘制质量. 作为已知的源图象不仅可来自于绘制系统, 也可从真实环境中捕获, 因此可以反映更加丰富的明暗、颜色、纹理等信息, 并且计算开销都是一样的.

## 1.2 IBR 技术的运用必要性

图形学界一直有这样一种观点, 就是场景的复杂程度依赖于图形硬件的功能, 其实这是问题的一个方面. 由于对虚拟场景的复杂程度和图像质量的要求越来越高, 场景的复杂程度并不是随着图形硬件的性能提高而按比例增加, 如果说硬件加速是按照几何级数增加, 近年来的虚拟现实技术的发展促使虚拟场景的复杂程度按照指数级数增加, 这从进入 90 年代图形学的进展可以得到验证. 因此从某种意义上来说, 场景复杂程度的提高是无止境的, 按传统意义上的面片表示法衡量, 上百万个面片表示的虚拟场景已经越来越常见. 例如美国加州大学伯克利分校的虚拟现实研究组用来显示该校新建的计算机系大楼的外观及室内景观用传统的面片数衡量要超过 250 万个. 从另一方面说, 昂贵的图形工作站的价格并非大多数用户所能承受, 因而在现有的硬件环境下实现高度真实感图形的生成和绘制技术显得尤为重要, 具有广阔的应用前景和良好的实用价值.

## 2 基于图象的实时绘制硬件方法 (real-time rendering hardware approach of image-based)

实现复杂图象实时绘制的硬件途径主要有两个: 专用计算机图象生成设备 CIG 和针对虚拟现实应用而设计的通用图形工作站.

CIG (Computer Image Generator) 能根据设计者的要求实时生成特定的三维计算机图象. 它能完成的处理内容包括: 物体的坐标变换、裁剪、透视变换、隐藏面消除、明暗处理、扫描转换和表面纹理等图象显示等基本处理; 参数的预处理、运动目标处理和自然现象等特征效应的处理; 以及图形绘制过程所需的一系列复杂的计算和逻辑判断. 这些处理任务要求计算机的处理速度高且系统结构优化. 50 年代末, 美国通用电气公司建立了第一台 CIG 装置, 客观存在以显示很简单的地面模型来证明实时图象生成的可能性; 60 年代, 第二台 CIG 系统诞生, 它能在地面模型上建立简单三维模型并采用了透视技术; 70 年代末和 80 年代初的 CIG 采用多通道技术, 对产生的图形做部分光滑处理, 典型的产品是美国通用电气公司的 C-130 视觉仿真器 (5 通道, 纹理产

生器可产生 15 种不同纹理图案, 显示处理器速度为 300 万个像素) 及 Evans Sutherland 公司研制的 CT-5 系统 (6 通道, 实时显示 2500 个多边形).

通用图形工作站是以 SGI 为首的图形工作站厂商近几年所致力推广的. 以通用图形工作站为基础构建的 VR 平台在实时图形处理方面已得到了很好的应用, 与 CIG 形成了目前专用图形处理产品市场上两股抗衡的力量. SGI 的 Onyx2 是第一个从底至顶 (CPU/图形硬件/软件/外设/网络) 全面支持 VR 的超级图形计算机系统, 具有支持灵活而高速的 I/O 吞吐能力管道结构, 每秒可产生 1100 万个多边形和 8.96 亿个像素. 通用图形工作站可面向基本和高级的 VR 应用. 面向基本 VR 应用的配置 (如 Onyx2 Reality) 采用 2~4 个 CPU, 1280×1024 反走样显示器, 显示速度为 3~4M 多边形/s, 适用于医学仿真、建筑/景物设计预演、销售仿真、科学可视化和普通动画的制作. 面向高级 VR 应用的配置 (如 Onyx2 InfiniteReality) 采用 4-CPU 台式或 8-CPU 架式结构, 双 1280×1024 反走样显示器和双光栅管理器, 显示速度可以达到 6~8M 多边形/s, 是目前飞行训练模拟、外科手术仿真、体绘制、交互工程、建筑预演、流体力学可视化、虚拟电影等应用的主要平台.

CIG 和通用图形工作站都是非常昂贵的专用图形处理设备, 配有专门的图形处理软件, 称为三维图形绘制的高端系统, 通用的桌面工作站和配备三维图形绘制加速设备的微机平台则称为图形绘制低端系统. 低端系统是适用面最广、通用性最强且应用前景最为广阔的计算平台, 但是目前在低端系统上实现复杂图形的绘制还难以达到实时和真实感兼备的水平.

除了在图形软件方面寻求解决方案之外, 将成熟的图形算法硬件化也是提高低端系统绘制性能的重要途径.

## 3 基于图象的实时绘制的软件方法 (real-time rendering software approach of image-based)

基于图象的绘制技术 (Image Based Rendering 简称 IBR) 基于一些预先生成的场景画面, 对接近于视点或视线方向的画面运用变换、插值与变形等方法进行处理, 来产生不同视点的场景观察结果, 具有在通用计算平台上实现快速绘制的能力. 目前基于图象的绘制技术研究重点在如下几个方面:

### 3.1 基于全景函数的方法

一般地,全景函数由一参数方程来描述,定义了空间任一处 $(V_x, V_y, V_z)$ 、任一视角 $(\theta, \varphi)$ 、在任何时刻 $t$ 和任一波长 $\lambda$ 范围内场景中的所有可见信息. 这样我们得到全视函数的形式为:

$$P = P(\theta, \varphi, \lambda, V_x, V_y, V_z, t)$$

全景函数事实上刻划了一给定场景的所有可能的环境映照,因而以图象的方式给出了场景的精确描述. 为生成一帧给定视点沿特定方向的视图,只需将视点 $V(V_x, V_y, V_z)$ 和球面角 $\theta, \Psi$ 及时刻 $t$ 代入全景函数即可. 这其实是对全景函数的采样过程,所得视图即为全景函数的一个样本. PlenOpticModeling 是第一个基于全视函数的系统,作者将该函数描述成有三维空间位置 $(V_x, V_y, V_z)$ 和二维方向 $(\theta, \varphi)$ 决定的参数函数 $P(\theta, \varphi, V_x, V_y, V_z)$ . 这是一个柱面全景图的集合. 这种表示在所有五维上都存在大量的冗余,并且如前所述方法中还要解决立体对应问题.

### 3.2 基于图象拼合的方法

全景图拼合将同一场景的多张有重叠的图象组合成一幅较大图象的处理. 很多系统中都成功地实现了将多张有重叠的图象缝合在一起构成圆柱形和球形全景图象拼合,这也是 PlenOpticModeling 和 QuickTimeVR<sup>[8]</sup>的基本思想. 当镜头运动很小时,我们能将整合图象中的小块组合在一起形成大的全景图拼合. 如果利用全方位镜头或鱼镜头捕获全景图是非常容易的. Szeliski 和 Shum<sup>[9]</sup>利用序列图象构造了一个完全的全景图象拼合系统. 他们的拼合算法是对每一副输入图象施行一矩阵变换,而不是具体地投射所有的图象到一个共同的平面(如圆柱形). 为了减少累积的图象整合错误,对整个图象序列运用全局调整(块调整)或局部调整方法. 然而对于两幅图象有不同视点的情形就要难得多,因为除了估计决定相对相机轮廓的8个参数外,还必须估计每个象素的深度. 为解决这个问题,要么施加额外的约束,要么利用更多源图象来增加输入变量,以求获得独立于视点的场景深度表示. 一旦参数被估计出来,新的视图就能被合成,而且还可利用多个视图构造三维校正的全景拼合图象. SIGGRAPH'99 上 H. -Y. Shum and L. -W. He.<sup>[10]</sup>提出同心圆拼合. 同心圆拼合限制相机的运动在平面的同心圆圈上. 同心圆拼合由同心圆圈上不同位置上拍摄的图象组成. 同心圆拼合由三个参数来描述:圆的半径,旋转角,仰角. 同心圆拼合用一种有效的方式结合适当的

捕获来绘制新视图. 尽管在绘制的视图中存在垂直方向上的失真,但能通过深度信息来纠正. 同心圆拼合有好的空间和计算效率. 与光场或光路图相比,因为它只要求三个参数,所以同心圆拼合有非常小的文件大小. 更重要的是,同心圆拼合非常容易捕获. 除了需要更多的图象外,同心圆拼合就象捕获传统的全景图一样容易. 同心圆拼合并不要求有恢复几何、光度测量的场景的复杂的建模过程. 然而,同心圆拼合允许用户在同心圆的区域内自由地移动、观察特殊光照的变化.

### 3.3 视图插值技术

视图插值就是利用图象变形的的方法产生视点沿着一定路径变化时的中间图象,在这种情况下,虚拟照相机可以在一定的范围内移动和转动. Chen<sup>[11]</sup>用两幅或多幅图象样本进行线性插值来运算中间图象,这种方法利用相机的位置和图象深度信息,事先计算出两幅图象对应象素的偏移向量,并存储在一对变形图中. 为了提高映射速度,减少“变形图”大小,他们使用四叉树法对图象进行分块压缩“变形图”,并按块的深度值进行前后排序,产生四叉树块查找链. 绘制新图象时,根据视点位置和图象对的关系,对偏移向量进行线性插值,并按插值向量移动源图象中相应象素到新视图中,从而得到新的视图. 如果两幅图象之间的距离过大,则它们之间的相同区域就会很小,从而增加了渲染误差. 当视点平行于视平面移动的视图插值方法能够产生准确的结果.

### 3.4 基于图象变形的的方法

基于图象的变形是指利用图象上有限的几个控制点,并通过对这些控制点的移动控制,来使整幅图象产生几何变化,以达到预期变化效果的一种技术. 在最近几年,图象变形技术已经取得了很大的发展. 图象变形是一种非常有效的视觉效果,在影视、广告、娱乐、轻工 CAD 等领域已有了广泛的应用. 例如:1982年纽约理工学院的 TomBrigham 制作的由一个女人变成一只山猫;迈克尔·杰克逊的音乐录音带“黑与白”这首歌中13个不同性别和种族的人的相互渐变;电影《终结者》中机械杀手 T-1000 由液体变为金属人,由金属人变为影片中的其它角色;Exxon 公司的影视广告中,一辆银色的轿车缓缓滑行渐渐变成一只老虎等等. 图象变形的另一分支是 Morphing. 其思想是指定一个变换,该变换把第一幅图象变形成为第二幅图象,它的逆变换把第二幅图象变形为第一幅图象. 由变形而生成的图象序列中,前面部分很象第一幅原始图象,中间部分既象第

一幅原始图象又象第二幅图象,后面部分则很象第二幅图象.基于特征的图象变形技术可以把一幅照片或计算机生成的图象以一种自然流畅、戏剧化、超现实主义的方式变为另一幅图象.它的关键技术包含变换函数的生成和图象的融合两个方面,其中,变换函数的生成是指如何通过图象上有限的控制点的坐标变化来界定整幅图象的坐标点变化,以便在原始图象和变形图象之间建立起坐标映射关系;而图象融合是指通过对两幅变形图象通过插值的方法来生成中间的变形图象. D. Smythe 于 1990 年提出了基于控制点网络的图象变形,其基本思想是根据用户指定的控制点对源图象和目标图象分别生成特征网格  $M_s$  和  $M_t$ ; 利用特征网格  $M_s$  和  $M_t$  分两步完成图象变形过程,即第 1 次沿横向对图象作变形处理;第 2 次沿纵向再对图象作变形处理,从而完成整幅图象的变形过程. Beier&Neely<sup>[2]</sup>提出了一种基于特征的自然变形技术称之为 FieldMorphing,该方法是利用用户在源图象和目标图象中指定的控制线段来生成变形函数.该方法允许动画师对变形进行直观的控制,通过交互地指定图象变形的特征(线对),可以较方便地达到动画师预期的视觉效果,然而这种方法会出现“歧点”,而且其计算量随着控制源数目的增加而增加.它能很平滑地使一幅数字图象变成另一幅数字图象. Ruprecht and Muller(1995) et al 提出了径向基函数和簿板样条方法.该方法是利用用户指定的离散数据点集

Seitz<sup>[13]</sup>在这个基础上提出了基于视点的图象形态变换.从而使图象变换的含义扩展到了三维的意义上.考虑到图象象素在对应点之间可能是一条三维曲线,所以它利用重投影方法将图象样本重投影到相同图象平面或者平行的图象中,这样所有的图象样本具有相同的视点方向,对应点之间的路径将是一条直线,这时用二维线性插值就可以得到正确的插值结果,然后再把结果利用重投影技术投影到预先计算好的中间图象平面中以保持插值图象的旋转不变特性.该方法较好地解决了传统图象变换中的曲变现象,而且不需要图象的任何先验知识.

### 3.5 基于立体视觉的视图合成方法

基于立体视觉的视图合成方法<sup>[18]</sup>是指利用立体视觉技术从已知的参考图象中合成相对于新视点的视图.其关键问题是找出每对已知图象之间的对应关系(correspondencemap),即解决立体匹配问题.由对应关系可导出偏差映射,并能进一步估计出场景中可见点的深度信息.利用深度信息可以对已

有的图象进行“变形”来合成新视点的图象.也就是说,通过对应关系建立了一个基于图象的场景几何表示.这种方法将场景视图及其对应关系组成一个图结构,图中顶点表示不同物理位置的场景视图,边表示邻接视图间的对应关系.

Leveau 和 Fangeras 描述了一种从弱校正图象中构造新图象的方法,但该方法要靠人工选择 4 对对应点. McMillan 和 Bishop 的 PlenOpticModeling 系统是一个基于柱面全景图象的系统.利用偏振约束(epipolarconstraint)的柱面偏差计算出相邻两个柱面全景图的偏差映射,然后依据该偏差变形已知的全景图象,从而得到新视点的柱面全景图.另外,也可以采用多基准线的立体算法.在这种方法中,真实图象先被映射到已多边形网格化了的深度图上,然后新视图可用标准图形学中的纹理映射方法生成.

还有一类基于立体视觉的方法称为模型与图象混合的方法.这些方法据从很多图象对中收集的立体信息恢复出 3D 场景结构. Debevec 等介绍一种用少数照片构造和绘制真实建筑场景的混合方法,利用所建立的粗略模型可以计算出真实场景与模型之间的偏差,而且可以将图形重新投影从而使立体匹配能处理相距较远的图象对.

基于立体视觉的图象合成方法主要有以下优点:

- 1) 新视图可以由两幅临近的参考图及它们的对应关系合成,整体的几何模型不是必需的;
- 2) 图象变形比图象绘制快得多,而且变形时间独立于场景复杂度;
- 3) 只需知道邻接图之间的相对轮廓信息,而不需要整体的一致性,例如图象可由手持相机捕获,并标上大致的全局坐标.

## 4 分布计算的方法(distributed calculation approach)

早在 1980 年 Whitted<sup>[20]</sup>就指出:最常用的基于图象的实时绘制方法即光线跟踪算法具有潜在的并行性.每个象素颜色的计算是独立的,计算结果直接对应到输出图象的象素单元中.并行光线跟踪一度成为解决基于图象的实时绘制问题的希望.然而,并行计算对硬件的要求太高,需要设计专用的并行计算机和一套复杂的并行操作系统软件.许多研究者发现,并行计算平台的开发代价已远远抵消了并行绘制带来的方便性,并行计算平台昂贵的使用价格也使绝大多数图形应用者望而却步.

90年代,世界范围内都出现了由并行计算向分布计算的转变趋势.人们发现利用已有硬件设备间的松散互联而实现的分布化计算也能够达到可观的超级计算能力.基于Internet的分布计算表明,如果将全世界上千万台连接到Internet上的主机协调地利用起来,将比目前拥有的最大型计算机的计算能力还要高千倍以上.受这一分布化趋势的影响,并行绘制也开始向分布绘制转变.目前,全球范围的分布绘制研究力度已经占了整个并行/分布化图形绘制研究的80%以上.

分布绘制是由若干非共享主存的自治处理机通过网络消息传递来实现对同一景物的合作绘制过程.它能充分利用已有的硬件资源,具有健壮性和可扩展性等优点.分布绘制可以由同构或异构的分布环境来支持.同构分布环境中的所有计算节点的硬件性能和软件系统都要求一致,在现实应用中,这种条件是难以获得的.异构分布环境是由若干具有不同的硬件体系结构、操作系统和网络结构的自治的计算机通过底层通信机制连接起来的,能合作完成单机无法胜任的应用计算任务.现有的各种档次微机和工作站都可以用来构建异构分布平台,因而它具有比同构环境更好的通用性.

## 5 结束语(Conclusion)

综上所述,基于图象的实时绘制技术正在或即将改变我们生活的许多方面,但由于硬件处理能力和软件技术两方面的原因,基于纯粹几何模型的复杂场景的建模仍是令人望洋兴叹的事情,而基于图象的实时技术将获得人们的青睐.但毕竟研究还刚刚起步,还有待在下列问题或方面进一步研究和展:

### (1) 高性能的硬件设备

现有的多数基于图象的实时绘制硬件都是非常昂贵的专用图形处理设备.而通用性强且应用前景广的低端绘制硬件系统,实现复杂图形的绘制还难以达到实时和真实感兼备的水平.因此,除了在图形软件方面寻求解决方案之外,将成熟的图形算法硬件化也是提高基于图象的实时绘制性能的重要途径.

### (2) 有效的数据表示

为了达到基于图象的实时绘制,我们常常要在真实感和实时性之间做权衡.许多系统不得不决定采用多少几何几何和多少图象;这样为了同时提高真实感和实时性,我们必须采取有效的数据表示.例

如在绘制时采用从压缩数据直接绘制以及多分辨率表示,有效地节约了存取开销,同时又不影响绘制速度和真实感.

### (3) 动态环境的绘制

直到目前为止,大多数基于图象的实时绘制技术方法都是集中于静态环境.随着视频技术的发展,人们希望基于图象的实时绘制技术也能够应用到动态环境的绘制.这样我们必须研究:采样(应该捕获多少幅图片)和压缩(怎样减少数据量).

## 参 考 文 献(References)

- 1 E H Adelson, J Bergen. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*, Pages 3-20. MIT Press, Cambridge, MA, 1991
- 2 Levin R. Photometric Characteristics of Light Control Apparatus. *Illuminating Engineering*, 1971, 66(4)
- 3 Ashdown I. Near field Photometry: A New Approach. *Journal of the Illuminating Engineering Society*, 1993, 22(1)
- 4 Sbert A M. An Integral Geometry based Method for Forming Factor Computation. *Computer Graphics Forum*, 1993
- 5 Levoy M, Hanrahan, P. Lighting Field Rendering. *SIGGRAPH'96 Proceedings*, 1996
- 6 Gortler S J, Grzeszczuk R, Szaliski R, Cohen M F. The Lumigraph. *SIGGRAPH'96 Proceedings*, 1996
- 7 Huang Xiaohu. Virtual in Distributed & parallel Environment. Phd Thesis of Xi'an Jiaotong University, 1997
- 8 Chen S E. QuickTime VR - An Image - Based Approach to Virtual Environment Navigation. *SIGGRAPH'95 Proceedings*, 1995
- 9 R Szaliski, H Y Shum. Creating full view panoramic image mosaics and texture mapped models. *Computer Graphics (SIGGRAPH'97)*, August 1997, 251~258
- 10 H Y Shum, L W He. Rendering with concentric mosaics. In *Proc. SIGGRAPH99*, 1999, 299~306
- 11 Chen S E, Willams. View Interpolation for Image Synthesis. *SIGGRAPH'93 Proceedings*, 1993
- 12 Beier T, Neely S. Feature based Image Metamorphosis. *Proceedings of SIGGRAPH'92*, 1988
- 13 Seitz S M, Dyer C R. View Morphing". *SIGGRAPH'96 Proceedings 1996*, 21~30
- 14 S Laveau, O Faugeras. 3D Scene Representation as a Collection of Images and Fundamental Matrices, INRIA, Technical Report No. 2205, February 1994
- 15 L McMillan. An image-based approach to the three-dimensional computer graphics. Thechnical report, UNC Computer Science TR97-013, 1999
- 16 Paul E Debevec, et al. Modeling and Rendering Architecture from Photographs: A hybrid geometry- and image-based approach. *SIGGRAPH'96, Computer Graphics*, 1996, 11, 20