

# 虚拟实景空间的实时漫游

孙立峰 钟力 李云浩 胡晓峰

(国防科学技术大学 703 教研室, 长沙 410073)

**摘要** 集成各种媒体信息的虚拟现实空间必须首先能够提供一种具有沉浸感的空间漫游机制, 保证用户在漫游过程中获得平滑的视觉效果, 并能维持用户视觉的一致性。文中讨论了在基于实景图象的虚拟信息空间进行漫游需要解决的图象处理问题及相关算法, 进而设计实现了一种满足实时要求的高效多线程漫游机制。

**关键词** 虚拟现实 全景图象 空间漫游 柱面投影 图象插值

## 0 引言

基于实景图象的虚拟信息空间(简称“虚拟实景空间”)是利用照相机采集的离散图象或摄像机采集的连续视频作为基础数据, 经过图象处理生成全景图象并对其进行空间关联建立起的具有空间操纵能力的虚拟环境。构建一个使用户具有身临其境的沉浸感、完善的交互能力的信息环境, 传统的方法是利用计算机图形技术, 对真实环境进行抽象从而建立其三维几何模型, 实时漫游过程中根据观察者的位置、光照、消隐信息由计算机绘制相应的视图。这种基于计算机图形学的虚拟现实系统仍然存在其局限性: 一是复杂空间的建模过程相当烦琐、工作量大; 二是不能将实时交互与高质量图形环境结合起来,

场景难以达到完全逼真; 三是对计算机软硬件要求较高, 很难满足大众化的虚拟现实应用。基于实景图象构造的虚拟信息空间避开了复杂的场景建模与绘制, 直接利用照相机或摄像机拍摄得到的实景图象(Real World Images)来构造虚拟信息空间, 这种基于实景的方法具有快速、简单、逼真的优点, 正在越来越多地应用于旅游景点、虚拟场馆介绍以及远地空间再现等方面。

虚拟实景空间在空间上的组织以视点空间以及对视点空间进行空间关联的空间链为基础, 这里对视点空间和空间链作如下定义:

**定义 1** 视点空间(ViewPoint-Space): 指用户在一个观察点处所观察到的球空间, 它由不同焦距的全景图象(如图 1 所示)按其焦距关系构成。



图 1 全景示意图

视点空间反映了观察者在虚拟环境中某一观察点所能观察到的不同细节程度的场景空间, 观察者可以在视点空间进行 360°环视、俯视、仰视以及焦距变换等多种方式的观察。

**定义 2** 空间链(SpaceLink): 是视点空间之间

的空间位置关系的描述, 其形式化描述为:

$SpaceLink ::= \langle ID \rangle \langle Source \rangle \langle Destination \rangle$   
 $\langle Orientation \rangle \langle Distance \rangle$

$\langle Source \rangle ::= ViewPoint\_Space$

$\langle Destination \rangle ::= ViewPoint\_Space + \langle Orienta-$

tion > ::= 源视点与目的视点连线的方位角(以正北方向为基准)

< Distance > ::= 源视点与目的视点之间的空间距离(实际环境中的物理距离)。

空间链是连接不同视点的运动路径,观察者可以沿空间链在视点之间进行移动,完成在虚拟环境中的游历。

虚拟实景空间参考超媒体信息组织模型,用空间链将空间上分布的视点空间连接成网状的虚拟实景世界,如图2所示。这种虚拟实景空间所要达到的基本目标是实现用户在其中“第一人称”的游历,在此基础上实现信息的获取和使用,即为用户提供

一种实景空间的漫游机制。用户在虚拟实景空间的漫游必须保证场景显示的实时性和显示画面的视觉一致性。所谓实时性是指系统根据用户的空间位置以及观察方向实时地调入对应的实景图象并进行相应的空间投影变换,将其显示于用户的观察窗口中。视觉一致性是指显示的场景必须满足一定的空间约束关系,如变焦观察时保持不同焦距全景图象间观察方向的一致性。为了实现虚拟实景空间的实时漫游,必须解决以上问题。本文针对我们研制的虚拟实景空间 HVS 系统分析了影响漫游过程中视觉感受的相关要素及相应的解决方法,并在此基础上设计实现了虚拟实景空间的实时漫游机制。

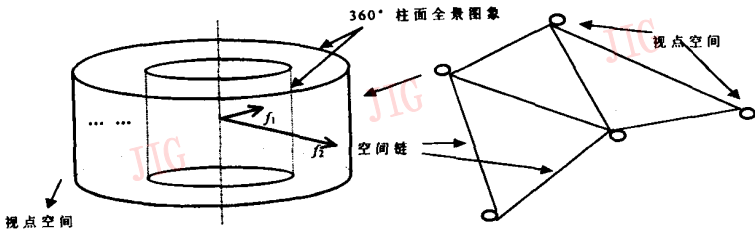


图2 由视点空间和空间链结构的虚拟实景空间

### 1 虚拟实景空间的视觉一致性

用户对虚拟实景空间的观察与感知体现在对视觉一致性有不同要求的3个方面:

- (1) 观察视点空间中某一焦距的全景,包括360°环视、仰视、俯视。
- (2) 同一视点空间中不同细节程度全景之间的切换,即变焦观察。
- (3) 不同视点之间的空间变换,即在虚拟实景空间中的运动观察。

#### 1.1 全景视觉一致性

在我们研制的虚拟实景空间 HVS 系统中,反映

特定细节程度场景的全景图象是将用数字相机采集的具有重叠边界的图象进行无缝拼接得到的。如果对重叠图象直接进行无缝拼接,将会破坏实际场景中的视觉一致性,如本来是直线的道路的拼接结果可能变成了折线。为了维持实际场景中的空间约束关系,我们必须将拍照得到的反映各自投影平面的各重叠图象映射到一个标准投影——柱面投影上,得到柱面图象后,才能进行拼接得到没有图象畸变的360°柱面全景图象。当用户实时观察全景图象时,视觉一致性要求恢复实际场景中的空间约束关系,即将柱面全景图象与用户观察方向相对应部分反变换投影到用户观察平面上,如图3所示。

柱面投影变换及其反变换算法的核心是投影变换公式,为此我们建立坐标系,如图4所示。原始图

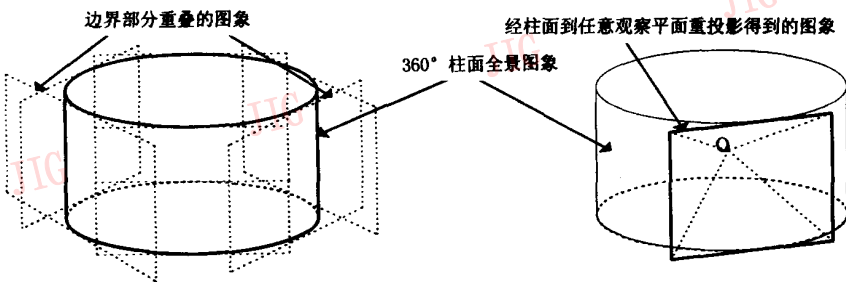


图3 重叠图象、柱面全景图象与观察图象

象  $I$ 、投影柱面图象  $J$  和柱面  $K$ , 原点为观察点  $O$  (投影中心)。设柱面半径为  $r$ , 投影宽度角为  $\theta$ , 图象宽度为  $W$ , 高度为  $H$ 。很容易的我们能得到柱面图象的宽度为  $r \times \sin\left(\frac{\theta}{2}\right)$ , 高度仍为  $H$ 。对图象  $I$  上

的任意一点  $P$ , 假设它在图象上的坐标为  $(x, y)$ ,  $P$  在柱面图象  $J$  上的对应点  $Q$ , 假设  $Q$  在  $J$  上的坐标是  $(x', y')$ 。

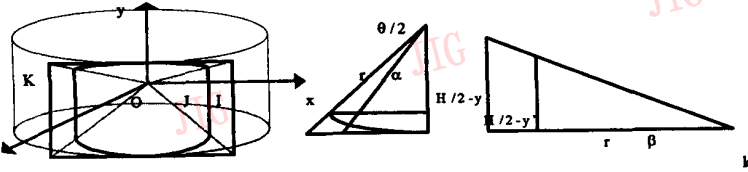


图4 柱面投影变换

我们就得到了柱面投影变换及其反变换的变换公式(推导参见参考文献[6]、[7]):

$$\begin{cases} x' = r \sin \frac{\theta}{2} + r \sin \left( \alpha \tan \left( \frac{x - \frac{W}{2}}{r} \right) \right) \\ y' = \frac{H}{2} + \frac{r \cdot (y - H/2)}{k} \end{cases}$$

柱面投影变换公式

$$\begin{cases} r = \frac{W}{2 \cdot \tan \frac{\theta}{2}} \\ k = \sqrt{r^2 + \left(\frac{W}{2} - x\right)^2} \\ \tan \alpha = \frac{\left(\frac{w}{2} - x\right)}{r}, x \leq \frac{w}{2} \\ \tan \alpha = \frac{\left(x - \frac{w}{2}\right)}{r}, x > \frac{w}{2} \end{cases}$$

$$\begin{cases} x = \frac{W}{2} + r \cdot \tan \left( \alpha \sin \left( \frac{x' - r \cdot \sin \frac{\theta}{2}}{r} \right) \right) \\ y = \frac{H}{2} + \frac{k \left( y' - \frac{H}{2} \right)}{r} \end{cases}$$

柱面投影反变换公式

其中:  $\alpha$  是  $P$  点到  $y$  轴的垂线与  $z$  轴的夹角, 由下式决定:

在这里  $\theta$  角由每张图象所占的角度唯一确定。一个全景的角度是  $360^\circ$ , 那么每张图象对应的角度就显而易见了。一般地,  $\theta$  角约为  $40^\circ$ 。图 5 是柱面变形的结果。



(a) 初始图象



(b) 柱面投影变换后的图象



(c) 反变换恢复的图象

图5 柱面投影变换及其反变换

### 1.2 变焦观察视觉一致性

变焦观察过程是模拟照相机或摄像机镜头的拉伸, 其中的视觉一致性要求保持不同焦距全景切过程中用户视线的一致性, 即不改变观察中心, 如图

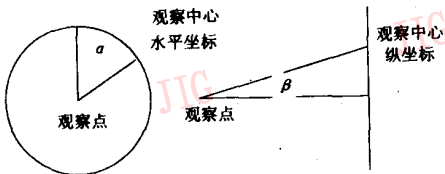
6 所示。用户在全景图象中的观察中心由用户的水平方位角  $\alpha$  和垂直方向上的仰角  $\beta$  共同决定, 如图 7 所示。  $\alpha$ 、 $\beta$  变焦过程具有水平方位角不变, 仰角根据视觉一致性相应变化的特点, 具体地讲, 对于拉近 (ZoomIn), 仰角增大, 对于拉远 (ZoomOut), 仰角减

小,因而保持变焦过程视觉一致性的关键在于确定不同焦距全景图中观察中心垂直方向坐标之间的变换关系,即处于观察中心的景物在不同焦距全景图象中成像高度之间的空间关系。假设用户从焦距为

$f_1$  的全景  $P_1$  转换到焦距为  $f_2$  的全景  $P_2$ , 处于用户观察中心的物体的物距为  $u$ , 在  $P_1$  中的像高为  $h_1$ , 在  $P_2$  中的像高为  $h_2$ , 根据透镜成像公式有:



图6 变焦观察示意图



柱面全景俯视图 柱面全景侧视图  
图7 全景图象观察中心示意图

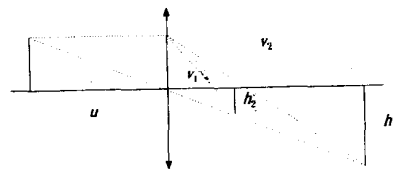


图8 透镜成像示意图

$$\frac{1}{u} + \frac{-1}{v_1} = \frac{1}{f_1}$$

$$\frac{1}{u} + \frac{1}{v_2} = \frac{1}{f_1}$$

$$\frac{v_1}{v_2} = \frac{h_1}{h_2}$$

最终结果有:

$$\frac{h_2}{h_1} = \frac{u - f_1}{u - f_2} \frac{f_2}{f_1}$$

从结果可以看出,成像高度间的关系与景物的物距也即景深有关。对于实景图象而言,从单一的照片上不可能获得其中景物的景深,为获取景深,可以根据双目立体视觉原理,拍摄景物的双目图象,从中计算其景深。但是一方面这种算法计算量大,需进行大量的预处理工作,降低了系统的简单易用性,另一方面变焦过程可能发生在360°全景图的任意观察中心,不可能对全景图上每一点计算其景深,并且有些观察方向上没有明确的景物以辅助双目图象分析,如室外的道路。因此我们在解决变焦过程视觉一致性问题时,引入了合理的简化处理。一种方法是考虑到实际的景深通常远远大于相机的焦距,等

式右端第一项的值约等于1,因而成像高度可以认为与其焦距成线性关系。另一种方法是考虑人眼的视觉特性,认为人对场景的观察符合分层景深的原则,即可将景物按其景深分为远景( $u > 10m$ )、中景( $5m < u < 10m$ )、近景( $u < 5m$ ),因此在拍摄原始场景时,我们根据场景的特点为其指定拍摄景深,如室外场景可认为是远、中景,室内场景是中、近景。之所以能够用该景深作为全景图象中各点景深的替代,是因为该景深距离上对应的景物包含了实际场景中能引起用户注意的大部分景物,对于更远的景物则因为细节程度不够而常常为观察者所忽略,更近的景物也没有必要进行变焦观察。我们对上述2种方法进行了实验比较,认为后一种的结果误差更小,在变焦过程中没有明显的画面跳动。

### 1.3 运动观察及其视觉一致性

空间移动是视点空间在空间轴上的操纵方式,是对人在现实世界中运动的一种简化和抽象,特指保持观察焦距  $f$  不变从一个视点空间转移到另一个视点空间。这种运动是通过基于源视点和目的视点的图象插值来加以实现的。我们称这种运动方向与观察方向一致的图象插值为纵向插值,在这里我们

对用户运动与观察方向进行必要的约束,规定观察方向和焦距在运动过程中保持不变,否则插值算法将极为复杂进而引起实时性的降低。

纵向插值,从实质上理解是一种观察方向上的图象仿射变换,插值的效果是实现观察方向上的对空间进行细节刻画度连续变化的观察,如图9所示。其中源视点具有最低分辨率,目的视点具有最高分辨率,从实现技术上,纵向插值的难在于插值图象的生成依赖源图象和目的图象间的信息关联:源图分辨率低却包含最多的内容信息,目的图象虽有最佳细节刻画但丢失了部分信息,这种信息关联可如图10所示。矩形2是目的图象对应的空间区域在源图象中的位置映射,矩形1是插值图象的位置映射,

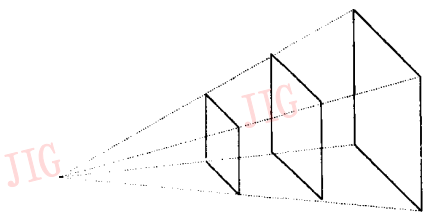


图9 全景图象纵向插值

(1) 对于室内场景,景物大都具有比较规则的几何外型,空间插值算法就是利用景物的几何信息,建立不同图象中同一景物的空间位置关联,如对应点、对应直线等,在运动过程中这些对应关系在运动步长的控制下完成初始图象到终止图象之间的变换,这种方法在图象处理中称为变形(Morph),我们对变形的控制点以手工的方式进行预先标注。

(2) 对于室外场景,由于景物形状复杂且没有规则的几何外型,中间过渡图象的生成采用基于场景渐变的思想。运动观察的特点是观察方向始终保持不变,因而起始图象、中间图象、终止图象在空间关系上可以看作是一种仿射变换,我们将中间场景的变化看作是起始图象逐渐消隐(淡出)过程与终止图象逐渐清晰(淡入)过程的叠加,起始图象在淡出过程中按仿射变换的比例关系逐渐放大,终止图象在淡入过程中由缩小显示逐渐按比例放大复原。

运动观察过程中的视觉一致性主要体现在2个方面:保持景物之间的遮挡关系和景物变化的连续性。实现正确的景物的遮挡关系是虚拟环境获得空间感与沉浸感的重要因素。不同于基于图形的虚拟现实系统,实际场景在虚拟实景空间中并没有其对应的三维空间模型,因而缺少确定景物遮挡所必需的

矩形3是目的图象在插值图象中的位置映射。阴影部分对应插值图象中仅能从源图象获取信息,源图象到插值图象的变换是一种被低分辨率高分辨率的变换,由于细节信息不够必然会导致图象的模糊,目的图象到插值图象的变换正好相反,由于插值图象分辨率是升高的趋势,矩形了的区域应逐渐主要由与进行插值,而此时阴影区域对应矩形区域0中位置的收缩,其插值结果趋向模糊,这就造成了区域3与阴影区域间的分辨率断层。图象插值算法的目的就是尽量消除上述问题,实现过渡场景的平滑连续。考虑到室内室外场景存在较大的差异,我们针对其场景特点实现了两种图象插值算法:

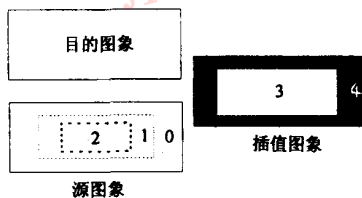


图10 纵向插值图象间空间关系

深度信息。全景图象中遮挡关系的实现是在拍摄过程中由相机的光学成像系统自动完成从三维空间到二维显示平面的投影变换,但在运动过程中,由于过渡图象是用空间插值算法生成的,无法确定景物相对于用户当前位置是否可见,因此对虚拟实景空间来说,实现运动过程中的景物遮挡更为困难。景物变化的连续性是指在运动过程中用户所看到的场景在空间上应当是连续的,不能出现场景的突变。之所以会提出这样的要求,是因为基于图象插值的运动观察过程中用户所能看到的场景依赖于起始图象和终止图象中的内容,由于遮挡的影响,实际道路上可能回有景物被遗漏,造成空间信息的丢失,导致场景变化的不连续,令观察者产生突兀的感觉。为解决上述2个问题,我们引入了参考点的概念(见图11)。参考点是运动路径上具有较大场景变化或发生显著景物遮挡的空间视点,在这些视点我们拍摄沿运动方向上的场景图象,运动路径上所有的参考点反映了实际环境中场景沿运动路径方向的变化情况。参考点的引入,一方面增加了运动路径上的空间信息,一定程度上消除了图象插值算法所固有的插值结果两端清晰中间模糊的缺陷,使场景变化更连续;另一方面参考点能够体现运动过程中出现的

显著景物遮挡情况,增强了虚拟实景空间的真实感,保证用户运动观察的视觉一致性。我们在系统实现中,将图象插值算法与参考点结合使用,明显改善了视觉效果。

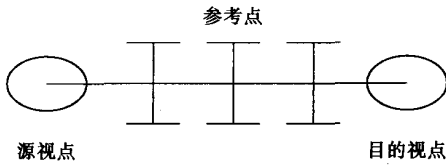


图 11 参考点示意图

## 2 多线程实时漫游机制

不同于基于图形的虚拟环境,虚拟实景空间的数据基础是实景图象,实时漫游要求系统根据用户状态(空间位置、观察方向)实时调入相应的实景图象,完成柱面全景到用户观察平面的投影变换显示,制约实时性的主要因素在于图象处理算法的性能以及图象数据存取速度。虚拟实景空间的实时漫游要求在实现图象处理快速算法的同时还必须实现灵活高效的数据缓冲机制。定点观察要求实现全景浏览迅速连续以及不同焦距全景之间过渡平滑无延迟,最简单的解决方法是将视点全景图象全部进行缓存

以适应实时浏览的要求,但是考虑到全景图象数据量巨大,一幅未压缩的全景位图其数据量达到 7M 左右,如果一个视点包含几层焦距的全景,那么仅仅缓存单个视点的数据就需占用极大的内存空间,同时大量内存的占用导致操作系统整体性能的下降,继而会降低虚拟实景空间漫游的实时性。因此必须解决局部全景的缓存问题。运动观察要求尽量降低用户从定点观察到运动观察的转换延迟,同时要求运动过程中过渡场景能够连续平滑,无明显的等待延迟,为此一方面需要设计实现快速的图象插值算法,另一方面需要对用户的运动进行预测,对相关的运动数据进行预调缓存。

我们采用多线程的体系结构实现虚拟实景空间的数据缓冲机制,如图 12 所示。我们根据实时漫游的需要,将系统中运行的线程按任务特性分为用户代理线程和系统服务线程。用户代理线程追踪用户状态,响应用户交互行为,并且负责调度系统服务线程。系统服务线程负责实现局部全景缓存(全景加载线程)、本地空间信息获取(本地环境加载线程)、运动数据预调(运动预测线程)以及完成运动观察处理(运动处理线程)。系统引入事件同步机制,对状态迁移的原因定义同步事件完成各线程间的同步与调度。

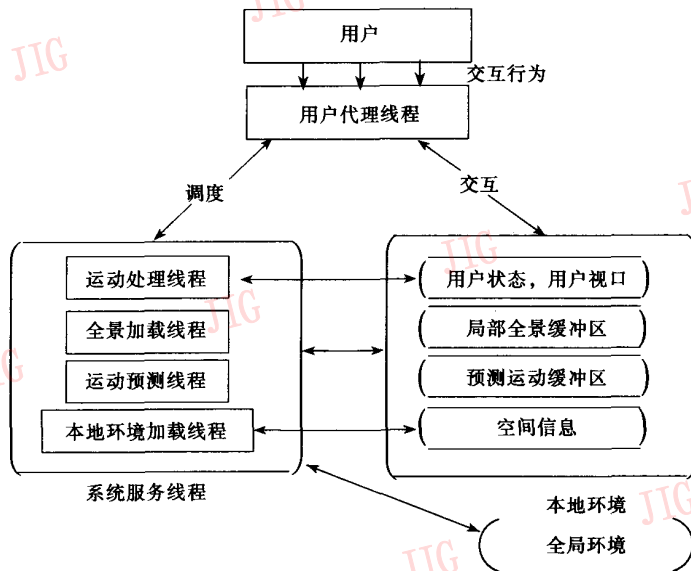


图 12 多线程漫游体系结构

## 3 结束语

本文讨论了实现虚拟实景空间实时漫游所需面临的问题及其相应的解决方法,我们在微机平台上

以 Visual C++ 开发实现了基于实景图象的虚拟空间 HVS 系统,该系统以超媒体信息模型组织全景图象,实现了以快速插值算法为核心,结合运动预测技术的多线程全景空间实时漫游机制。进一步的研究工作包括对图象插值算法进一步优化,允许用户能

在漫游过程中进行全方位的观察以及与空间对象进行操纵与交互等,目前我们正进行这方面的研究。

### 参考文献

- 1 胡晓峰,等. HVS:一个半现实全景图时空模型. 小型微型计算机系统,1994,15.
- 2 Shenchang Eric Chen. QuickTime VR: An Image-Based Approach to Virtual Environment Navigation. Computer Graphics Proceedings, Annual



**孙立峰** 1995年于国防科技大学系统工程与数学系获学士学位,现为国防科学技术大学系统工程与数学系博士研究生,研究方向为虚拟现实技术,分布虚拟环境,虚拟空间会议系统。



**李云浩** 国防科学技术大学系统工程与数学系硕士、博士连读研究生。研究方向为图象处理与计算机视觉。

Conference, 1995.

- 3 钟力等. 全景视频的信息组织和实现方法. 小型微型计算机系统,1996, 17.
- 4 A Virtual Studio for live Broadcasting: The Mona Lisa Project. Multimedia, 1996, Summer.
- 5 钟力等. 360度柱面全景图象生成算法及其实现. 小型微型计算机系统,已录用.
- 6 钟力. 虚拟实景空间的研究与实现:[学位论文]. 长沙:国防科学技术大学.

**钟力** 国防科技大学系统工程与数学系博士,主要研究方向为基于实景的虚拟空间建模技术。

**胡晓峰** 教授,博士生导师。主要研究方向为多媒体技术与应用,虚拟现实技术,综合集成研讨环境。

## Real-Time Walkthrough in Real Image - Based Virtual Space

Sun Lifeng, Zhong Li, Li Yunhao and Hu Xiaofeng

(Department 7 of National University of Defence Technology, Changsha 410073)

**Abstract** Virtual Information Space is a application environment, integrating multimedia information. The first ability of virtual information space is to give users the capability of wandering around the virtual space smoothly, and keep the consistent of user's visual sense while wandering. This article discusses the image processing problems and algorithms of walking through in panoramic based virtual space, then designs and implements a high performance multithread wandering algorithm.

**Keywords** Virtual reality, Panorama, Walk through, Cylinder projection, image morph