

# 一种用于室内虚拟环境的新表示<sup>\*</sup>

李洪举 吴恩华

(中国科学院软件研究所计算机科学实验室 北京 100080)

E-mail: lhj@ox.ios.ac.cn

**摘要** 针对室内环境的虚拟问题,提出一种用于描述封闭环境的新表示.该表示由定义在经纬度平面上的光强-距离函数组成,可以支持3个自由度的全景浏览、连续变焦以及任意路线的移动.基于这种表示,给出一个使用少数照片创建室内环境的方法,它结合了基于几何和基于图象的绘制技术,算法快速稳定,对拍摄条件和硬件设备没有苛刻要求.

**关键词** 基于图象的虚拟现实,表示,几何图象混合绘制,全景图,纹理映射.

**中图法分类号** TP391

室内环境的虚拟实现是虚拟现实,特别是近年来兴起的基于图象的虚拟现实(Image-Based Virtual Reality)研究中的一个重要方向.它在许多领域有着广泛的应用,比如虚拟的办公室、教室、博物馆、教堂、洞窟等已应用于远程办公、教育、娱乐、房地产、旅游等行业中.对于创建室内虚拟环境及其他基于图象的虚拟现实系统,构造合适的表示是至关重要的.所谓表示(Representation)是指能把场景信息表达清楚的形式化系统,包括所用的数学形式及数据结构,它是虚拟环境的构造算法和交互算法的基础. S. E. Chen<sup>[1]</sup>以及 L. McMillan 和 G. Bishop<sup>[2]</sup>最早用柱面全景图(Panoramic Image)来表示场景,从而使得观察者可以在水平方向上进行360°的环视.但在垂直方向上,柱面全景图的视野是受限的,因此,无法支持该方向上的全景浏览.

随后,一些新的表示被提出来,用于支持任意方向的全景浏览. N. Greene<sup>[3]</sup>曾采用正六面体(Cube)作为投影平面, M. Regan 和 R. Pose<sup>[4]</sup>实现了这一思想,提出了一种称为“地址重计算流水线(Address Recalculation Pipeline)”的新技术,将投影在正六面体各个面上的图象变换为正确的透视图象. J. Shade 等人<sup>[5]</sup>进一步发展了这一思想,将正六面体表示与层次方法(Hierarchical Method)结合起来,用于加速复杂环境的漫游(Walk-Through)算法.

最近, M. Levoy 和 P. Hanrahan<sup>[6]</sup>以及 S. Gortler 等人<sup>[7]</sup>分别构造了两个较为复杂的表示——light slab 和 Luminograph,通过对自由空间(即没有遮挡物的区域)进行密集采样来表示其全景信息.这一表示可以支持任意位置和角度的观察,为此付出的代价是数据量巨大,构造方法复杂,生成新视图的运算量也较大.

针对上述表示在描述室内环境时遇到的困难,本文提出一种新的表示,我们称之为 $\theta-\phi$ 综合图.与正六面体不同的是,新的表示不是将场景的信息投影在3D空间表面,而是投影在由经纬度构成的2D平面上;而且它不但表示了场景的光强分布,同时也表示了其距离分布.与其他表示相比, $\theta-\phi$ 综合图支持各种相机运动,数据量较小,交互算法简单快速,原始数据获取容易.

本文第1节给出新表示的详细定义;第2节导出基于新表示的全景浏览与漫游算法;第3节给出构造新表示的一种方法;最后是实验结果及讨论.

## 1 $\theta-\phi$ 综合图

### 1.1 定义

考虑一个由凸的封闭曲面  $S$  构成的场景,以封闭曲面内某点  $O$  为球心建立单位球面(高斯球面).曲面  $S$  上的任一点  $P$  被映射到高斯球面上的一点  $P'$ ,  $P$  点的光强度  $I$  和  $P$  到球心  $O$  的距离  $r$  被映射为  $P'$  的相应值.这样,场景就可以由高斯球面上的光强-距离分布函数来定义.由于球面的离散化比较困难,我们将高斯球面上的点进一步映射到

\* 本文研究得到国家自然科学基金资助.作者李洪举,1966年生,助理研究员,主要研究领域为虚拟现实.吴恩华,1947年生,研究员,博士生导师,主要研究领域为计算机图形学,CAD,虚拟现实.

本文通讯联系人:李洪举,北京100080,中国科学院软件研究所基础部

本文1998-02-28收到原稿,1998-04-20收到修改稿

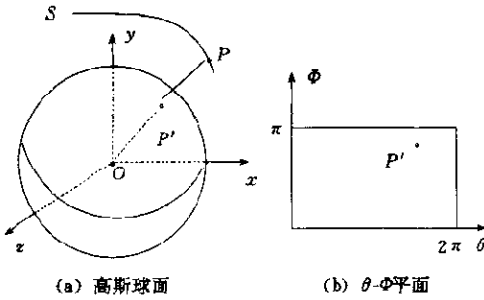


图1 θ-Φ综合图

经纬度(θ-Φ)平面上,定义在θ-Φ平面上的光强-距离分布函数称为θ-Φ综合图(如图1所示)。

### 1.2 多分辨率

同一场景由于观察或测量尺度不同,会表现出不同的几何形状和细节。在基于图象的虚拟现实系统中,一般采用多分辨率(Multiple Resolution 或 Multiresolution)技术来再现不同尺度下的场景。

用不同的抽样率对θ-Φ综合图进行离散化,就形成多分辨率综合图象。这样,对空间的抽样就转变为对视角的抽样,观察不同分辨率的综合图象,就可以获得不同层次的场景细节。多

分辨率θ-Φ综合图象用金字塔结构<sup>[8]</sup>存储。

## 2 基于θ-Φ综合图的交互观察算法

虚拟现实系统的重要特性是模拟观察者(即虚拟摄像机)的运动,包括旋转(视点位置不变,视向改变)、移动(视点位置改变)、变焦(视野改变)等方式。θ-Φ综合图可以很好地支持这些运动。下面是相机运动时生成新视图的算法。

### 2.1 相机旋转

相机旋转有3个自由度,绕水平轴(改变俯仰角)、绕垂直轴(改变方位角)和绕视向(视向即垂直于成像平面的轴)。θ-Φ综合图可支持全部3个自由度的观察。

观察过程可归结为从θ-Φ平面到成像平面的几何变换(如图2所示)。成像平面上的每个象素对应于θ-Φ综合图中的唯一一个象素,这样,相机旋转可以由地址计算来实现。设成像平面垂直于当前视向(θ<sub>0</sub>, Φ<sub>0</sub>),成像平面的水平和垂直视野(即最大视角)为2θ<sub>d</sub>和2Φ<sub>d</sub>, (u, v)为成像平面上的一点。令

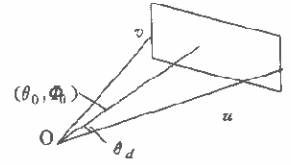


图2 投影变换

$$\begin{aligned}
 A &= \sin\theta_0 \sin\Phi_0 + u * \operatorname{tg}\theta_d \cos\theta_0 + v * \operatorname{tg}\Phi_d \sin\theta_0 \cos\Phi_0 \\
 B &= \cos\Phi_0 - v * \operatorname{tg}\Phi_d \sin\Phi_0 \\
 C &= \cos\theta_0 \sin\Phi_0 - u * \operatorname{tg}\theta_d \sin\theta_0 + v * \operatorname{tg}\Phi_d \cos\theta_0 \cos\Phi_0
 \end{aligned}
 \tag{1}$$

则其相对应θ-Φ综合图上的地址(θ, Φ)为

$$\begin{aligned}
 \theta &= \operatorname{arctg}(A/C) \\
 \Phi &= \operatorname{arctg}(\sqrt{(A * A + C * C)}/B)
 \end{aligned}
 \tag{2}$$

### 2.2 连续变焦

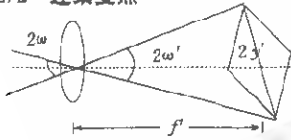


图3 zooming与视场角

连续变焦(Continuous Zooming)是指相机焦距和视野(Field of View)的改变,在电影技术中称为变焦镜头,主要用于对场景进行不同尺度的观察。

根据几何光学原理,视场角的大小取决于焦距和成像平面画幅的大小。<sup>[9]</sup>如图3所示,

$$\operatorname{tg}\omega' = \operatorname{tg}\omega = y'/f \quad (\text{当介质为空气时}), \tag{3}$$

其中2ω和2ω'分别是物方和像方的视场角,2y'为成像平面画幅对角线的大小,f

为焦距。显然,当成像平面尺寸一定时,有

$$\operatorname{tg}\omega \propto \operatorname{tg}\theta_d \propto \operatorname{tg}\Phi_d \propto 1/f. \tag{4}$$

由式(4)、(1)可知,焦距f的改变导致最大水平视角θ<sub>d</sub>和垂直视角Φ<sub>d</sub>的改变,从而改变了地址映射关系。因此,连续变焦可以通过改变最大视角来模拟。

### 2.3 视点移动

θ-Φ综合图可以支持任意路线的视点移动(Camera Movement),或称“漫游”(Walk-Through)。如图4所示,视点由原来的O(0,0,0)点移动到O'(x<sub>m</sub>, y<sub>m</sub>, z<sub>m</sub>)点,现在来计算关于O'点的光强分布图。在O'点建立坐标系x'-y'-z',各轴平行于原坐标系,对于曲面S上的任一点P,它在两个坐标系下的坐标分别为(x, y, z)和(x', y', z')。显然

$$x' = x - x_m, \quad y' = y - y_m, \quad z' = z - z_m. \tag{5}$$

表示成球坐标形式:

$$\begin{aligned} r' \sin \theta' \sin \Phi' &= r \sin \theta \sin \Phi - x_m \\ r' \cos \Phi' &= r \cos \Phi - y_m \\ r' \cos \theta' \sin \Phi' &= r \cos \theta \sin \Phi - z_m \end{aligned} \quad (6)$$

注意到

$$r = r(\theta, \Phi) \quad (7)$$

根据式(6)(7),对于  $O$  点处综合图中的每个象素,由其地址  $(\theta, \Phi)$  可以计算出该象素在  $O'$  综合图中的新地址  $(\theta', \Phi')$ ,但这样生成的关于  $O'$  的综合图会有“漏洞”,解决这一问题的方法之一是插值,而我们采用由  $(\theta', \Phi')$  反求  $(\theta, \Phi)$  的方法,可以得到更精确的图象,式(6)中消去  $r'$ , 得

$$\begin{aligned} r \sin \theta \sin \Phi - \sin \theta' \operatorname{tg} \Phi' * r \cos \Phi - x_m + y_m \sin \theta' \operatorname{tg} \Phi' &= 0 \\ r \cos \theta \sin \Phi - \cos \theta' \operatorname{tg} \Phi' * r \cos \Phi - x_m + y_m \cos \theta' \operatorname{tg} \Phi' &= 0 \end{aligned} \quad (8)$$

方程(7)(8)给出了视点移动前后综合图之间的地址变换关系,这是一个关于  $\theta, \Phi$  的二元非线性方程组,本文用 broyden 算法<sup>[11]</sup>求解  $\theta, \Phi$ .

### 3 $\theta$ - $\Phi$ 综合图的构造

室内场景的  $\theta$ - $\Phi$  综合图象可以用多种方法生成,比如,通过几何建模和绘制用计算机合成,或由照片和距离图象(Range Image)经过变换生成,本文采用了基于图象和基于几何模型的混合绘制方法来构造,用户只需要输入场景的一组照片和一些辅助数据,具体步骤如下(如图5所示). (1) 输入室内场景的原始照片(光强图象)和辅助数据(主要是面的类型、参数及关系); (2) 将每一幅原始图象变换到参数纹理空间; (3) 在参数纹理空间进行图象拼接; (4) 生成  $\theta$ - $\Phi$  综合图象.

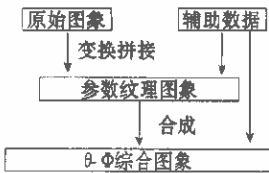


图5 工作流程

在对主要步骤进行讨论之前,先定义一个坐标系,我们把定义几何表面的参数空间和定义纹理的纹理空间合并为一个新的空间,称为参数纹理空间,即几何表面用二维参数方程定义,而其物理特性(对光源的反射)则表示为参数空间上的纹理图象,参数纹理图象在以目标为中心的坐标系中描述了目标的所有信息,是视点无关的,其坐标系称为参数纹理坐标系,简称参数坐标系.

#### 3.1 获取纹理图象

对于场景空间中的平面,由原始图象(照片)获取相应的参数纹理图象,实际上等效于从图象坐标系到参数坐标系的变换,即反投影变换(Back-Project),设原始图象坐标系为  $u$ - $v$ ,参数坐标系为  $X$ - $Y$ ,我们采用下列方程实现反投影变换.<sup>[11]</sup>

$$X = \frac{\sum_{0 \leq i+j \leq n} a_{ij} u^i v^j}{\sum_{0 \leq i+j \leq n} c_{ij} u^i v^j}, \quad Y = \frac{\sum_{0 \leq i+j \leq n} b_{ij} u^i v^j}{\sum_{0 \leq i+j \leq n} c_{ij} u^i v^j} \quad (9)$$

对于理想的针孔相机模型,  $n$  取 1. 如果考虑到变焦失真(当平面距观察者过近或过远时,变焦失真是很难避免的),  $n$  值可取 2 或更高.

对于给定的一组校正点,通过解线性方程组,可以求得方程(9)中的系数.可以看出,  $n=1$  时,只需要 4 组校正点即可.  $n=2$  时,需要 9 组.为保证较高的精度,可以选更多组校正点,然后用奇异值分解(SVD)法<sup>[10]</sup>来解决稳定的线性方程组.

获取反投影变换方程后,对于图象空间中的任意一点  $(u, v)$ ,由式(9)可以求出其在参数坐标系中的坐标  $(X, Y)$ ,然而这样取得的  $(X, Y)$  往往是不连续的,为此要把校正过程反过来进行,即由  $(X, Y)$  反求  $(u, v)$ . 此时,需要求解非线性方程组.为了避免这一问题,实际中采用的是从参数坐标系到图象坐标系的变换,在式(9)中交换  $(X, Y)$  和  $(u, v)$  即可.

#### 3.2 拼接

拼接在参数纹理空间进行.为保证所生成纹理的精度,要尽量使原始图象的抽样率(象素/单位长度)大致相同.当两幅图被反投影到同一参数纹理平面时,重叠或相邻部分会出现色彩不协调的现象,这可以用缓冲带方法<sup>[12]</sup>来消除.

在获取场景各表面的参数纹理图象后,可以通过光线追踪技术或坐标变换来绘制  $\theta$ - $\Phi$  综合图象.

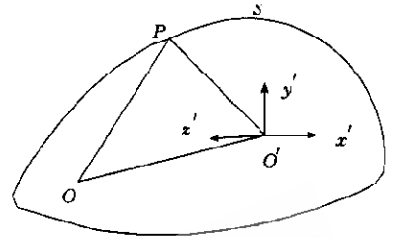
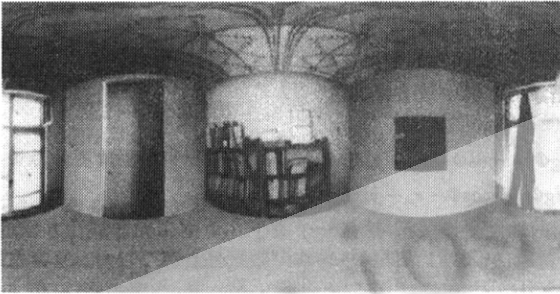


图4 视点移动

### 4 实验结果

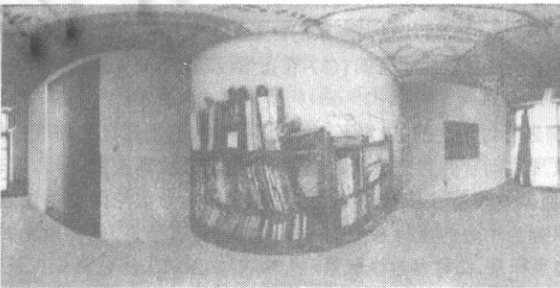
我们采用上述方法构造了一间虚拟办公室(如图6所示).图6中(a)和(b)为视点位于办公室中心时的 $\theta-\Phi$ 光强及距离图(分辨率为 $2160 \times 1080$ );(c)为视点移动后新的 $\theta-\Phi$ 光强图象;(d)为视向( $\theta=180^\circ, \Phi=90^\circ$ )处的重投影图象( $400 \times 400$ ,最大视角均为 $30^\circ$ ,zoom倍数为1.1).



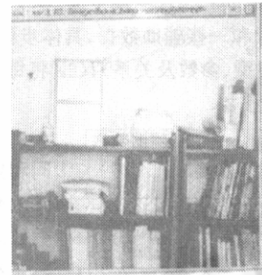
(a)  $\theta-\Phi$  光强图 ( $2160 \times 1080$ )



(b)  $\theta-\Phi$  距离图



(c) 视点移动后的  $\theta-\Phi$  光强图



(d) 重投影图象 ( $400 \times 400$ )

图6 实验结果

数据量是一个重要问题.由于人眼的最小分辨角大约为 $1' \sim 2'$ ,当两个点的视角差小于最小分辨角时,它们在视网膜上的像就会重合为一点.在 zooming 倍数不大的情况下,每隔 $2' \sim 3'$ 对方位角和俯仰角进行抽样,就能得到很好的视觉效果(在实验中,抽样间隔甚至可达 $15'$ ).图6中(a)和(b)所用的抽样率为每 $10'$ 一个像素,颜色用3个字节,距离用2个字节,总数据量约为12MB,压缩后只有1MB左右.而对于一般的场景,文献[2,6,7]往往需要数百MB乃至数GB.

$\theta-\Phi$ 综合图的生成和基于它的交互观察是两个相互分开的过程.前者不需要实时完成,所用的计算时间主要取决于室内环境的复杂程度及图象的分辨率.我们使用的平台为奔腾III/233微机,64M内存,生成图6(a)和(b)约用了10几分钟.而影响交互速度的因素主要有:视图大小、地址计算速度、图象解压速度(取决于解压程序)及数据传输速度(取决于计算机性能).其中地址计算(尤其是模拟相机的旋转和变焦)的大部分是求三角函数值,可用查表代替,这一点尤其适于硬件实现.在不进行图象压缩/解压的情况下,对 $300 \times 200$ 的视图,在模拟相机旋转时可达 $3 \sim 4$ 幅/s.

### 5 结语

本文给出了一个用于描述封闭环境的新表示: $\theta-\Phi$ 综合图.与其他表示相比, $\theta-\Phi$ 综合图有如下优点:(1)支持多自由度的全景观察与漫游;(2)存储方便,数据量较小;(3)视图生成算法简单、快速,适用于硬件实现;(4)原始数据的获取比较容易,因为光强图象和距离图象是关于场景的最原始的数据形式,对于复杂场景,获取距离图象(如用专用设备直接获取或由分析光强图象间接获取)比建立严格的几何模型要简单得多,这尤其适合基于图象的虚拟现实.

但该表示只适于凸的封闭场景。另外,为得到较精确的“移动”图象,本文所用的算法也比较复杂,这需要今后做进一步研究,以获取高质量、快速的漫游效果。

此外,本文构造  $\theta$ - $\phi$  综合图的方法不需要苛刻的拍摄条件和专门的拍摄设备,算法稳定,运算量适中。不足之处在于需要输入一些测量数据;对于非规则表面,需要用平面“逼近”,这样就会影响纹理拼接的精度,特别是当表面的曲率较大时甚至会失去意义。

### 参考文献

- 1 Chen S E. QuickTime VR—an image-based approach to virtual environment navigation. Computer Graphics, Annual Conference Series, ACM SIGGRAPH, 1995. 29~38
- 2 McMillan L, Bishop G. Plenoptic modeling: an image-based rendering system. Computer Graphics, Annual Conference Series, ACM SIGGRAPH, 1995. 39~46
- 3 Greene N. Environment mapping and other applications of world projections. Computer Graphics and Applications, 1986, 6 (11): 21~29
- 4 Regan M, Pose R. Priority rendering with a virtual reality address recalculation pipeline. Computer Graphics, Annual Conference Series, ACM SIGGRAPH, 1994. 155~162
- 5 Shade J *et al.* Hierarchical image caching for accelerated walkthroughs of complex environments. Computer Graphics, Annual Conference Series, ACM SIGGRAPH, 1996. 75~82
- 6 Levoy M, Hanrahan P. Light field rendering. Computer Graphics, Annual Conference Series, ACM SIGGRAPH, 1996. 31~42
- 7 Gortler S J *et al.* The lumigraph. Computer Graphics, Annual Conference Series, ACM SIGGRAPH, 1996. 43~54
- 8 Perlin K, Velho L. Live paint: painting with procedural multiscale textures. Computer Graphics, Annual Conference Series, ACM SIGGRAPH, 1995. 153~160
- 9 赵凯华,钟铨华. 光学(上册). 北京,北京大学出版社,1984  
(Zhao Kai-hua, Zhong Xi-hua. Optics. Beijing: Beijing University Press, 1984)
- 10 Press W H *et al.* Numerical Recipes in C: The Art of Scientific Computing. 2nd edition, Cambridge, Cambridge University Press, 1992
- 11 Wei Guo-qing, Ma Song-de. Implicit and explicit camera calibration: theory and experiments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(5): 469~480
- 12 蔡勇,刘学慧,吴恩华. 基于图象绘制的虚拟现实系统环境. 软件学报, 1997, 8(10): 721~728  
(Cai Yong, Liu Xue-hui, Wu En-hua. Image-based rendering: a technology for virtual reality system. Journal of Software, 1997, 8(10): 721~728)

## A New Representation for Indoor Virtual Environment

LI Hong-ju WU En-hua

(Laboratory of Computer Science Institute of Software The Chinese Academy of Sciences Beijing 100080)

**Abstract** In this paper, the authors give a new representation for virtual reality, which is defined as an intensity-range function on a latitude-longitude plane. It supports three-degree-freedom viewing (yaw, pitch and roll), continuous zooming and walk-through in a virtual enclosed environment. On the basis of the representation, a hybrid approach is addressed for constructing virtual indoor environment, which only requires a sparse set of still photographs and has no strait restriction on taking original photographs. Experiment results show that the approach is robust.

**Key words** Image-based virtual reality, representation, hybrid geometry and image based rendering, panoramic image, texture mapping.