

华东交通大学教材建设基金资助项目

SHIJING TUXIANG PINJIE
JIQI MANYOU KONGZHI JISHU

实景图像拼接 及其漫游控制技术

傅军栋 姚孝明 编著



西南交通大学出版社
[Http://press.swjtu.edu.cn](http://press.swjtu.edu.cn)

实景图像拼接及其漫游控制技术

傅军栋 姚孝明 编著

西南交通大学出版社

· 成 都 ·

图书在版编目 (C I P) 数据

实景图像拼接及其漫游控制技术 / 傅军栋, 姚孝明
编著. —成都: 西南交通大学出版社, 2011.6
ISBN 978-7-5643-1219-0

I. ①实… II. ①傅…②姚 III. ①计算机图形学
IV. ①TP391.41

中国版本图书馆 CIP 数据核字 (2011) 第 110287 号

实景图像拼接及其漫游控制技术

傅军栋 姚孝明 编著

责任编辑	李芳芳
特邀编辑	李娟
封面设计	墨创文化
出版发行	西南交通大学出版社 (成都二环路北一段 111 号)
发行部电话	028-87600564 028-87600533
邮 编	610031
网 址	http://press.swjtu.edu.cn
印 刷	四川森林印务有限责任公司
成品尺寸	185 mm × 260 mm
印 张	19.75
字 数	506 千字
版 次	2011 年 6 月第 1 版
印 次	2011 年 6 月第 1 次
书 号	ISBN 978-7-5643-1219-0
定 价	49.00 元

图书如有印装质量问题 本社负责退换
版权所有 盗版必究 举报电话: 028-87600562

前 言

基于图像渲染的虚拟现实技术近年来飞速发展，改变了传统的以照片和视频记录场景的方式，突破了以往关于视点、视角以及拍摄区域等的局限，使得人们对于真实世界的记录方式和观察方式发生了质的飞跃。虚拟漫游技术在空间感和交互方式上有着照片和视频无法比拟的优势，在模拟训练、游戏、网上旅游、数字博览会以及影视特技制作等方面得到了广泛的应用，如谷歌街景、在线百脑汇商城等，给人们以全新的感受。

近十多年来，实景图像拼接及其虚拟漫游控制技术由于具有“照片真实感”的惊人效果而得到了迅猛的发展，相关文献浩如烟海，难以尽览；各出版社也推出了一些中文论著，以及国外有关论著的影印版；一些院校开设了相关的课程，等等，对该技术的发展发挥了良好的推动作用。然而所有的工作并没有到此为止，随着新的应用需求不断提出，实景图像拼接技术也在各种约束条件的消除中保持着其多年的热点地位。例如，原始图像采集条件已经从原来的专用设备退化到现在任意手持相机，图像内容也从原来的静态画面到现在的动态画面，有些甚至允许有较大的曝光差异，等等。

笔者从 2000 年开始相关技术的研究，积累了大量的研究资料；同时，我们也在 Julesz 视觉等价的基本思想指导下，提出了基于视觉真实的场景构建模型，参与、主持 2008 年国家 863 项目（2008AA01A0314）“基于宽带网络的实景三维城市生活”，取得了一些研究成果。遗憾的是，在多年研究与开发过程中一直因为缺乏较为系统、既有厚实理论基础又有应用实例的专著而屡感不便。因此，编写本书的主要目的并非我们个体研究的介绍，而是将各家之长做一些系统归类和分析，使其在一个相对统一的框架下各得其所，相得益彰，构成一个完整体系。

编写本书有两个目的：一是从自身的研究与开发经历入手，将该方向的相关技术与理论系统化，从而给其他研究人员提供一致性较强的参考；二是尽量从实际问题出发，精选具有代表性的算法与解决方案，力求叙述清楚，不仅有助于初学者快速入门，还可对相关研究人员起到比较好的启发作用。本书可以作为相关专业人员的技术手册，亦可用作研究生或本科生的教学参考。

本书由傅军栋和姚孝明共同担任主编，负责全书的组织和定稿。其中，姚孝明编写了第 1 章的 1.1 和 1.2 节以及第 3、8、9、10 章；其余章节由傅军栋负责，与徐征、钟化兰、肖磊、吴昊共同完成。同时，感谢相关研究人员，正是在他们科研成果的基础上，本书才能更系统地展示给广大读者。

本书为华东交通大学教材建设基金资助项目。

由于时间仓促，加上作者学术水平有限，疏漏之处，敬请批评指正！

作 者

2011 年 4 月

目 录

第 1 章 绪 论	1
1.1 虚拟现实技术	1
1.2 全景图拼接技术	5
1.3 虚拟漫游技术	10
1.4 本书创新点及章节结构安排	16
参考文献	16
第 2 章 全景图拼接的基础理论	18
2.1 全光函数理论	18
2.2 两种拼图算法	26
2.3 基于部分几何信息的图像拼接技术	36
参考文献	39
第 3 章 图像拼接误差理论模型研究	42
3.1 图像最佳重叠比例	42
3.2 对称误差导致的积累误差	47
3.3 非对称误差导致的积累误差	49
3.4 照相机标定方法	52
3.5 图像的采集及变换	55
参考文献	66
第 4 章 圆柱型全景图图像拼接	67
4.1 圆柱型全景图拼接技术概述	67
4.2 实景图像的拍摄及要求	68
4.3 圆柱体投影	69
4.4 图像配准的基本概念	74
4.5 图像变换的估计	95
4.6 图像的拼接与融合	100
4.7 圆柱型全景图实现的实例	114
参考文献	121

第 5 章 球面全景图图像拼接	123
5.1 球面全景图图像拼接技术概述.....	123
5.2 球面全景图的映射方式.....	123
5.3 球面投影变形的解决.....	135
5.4 球面全景图的拼接.....	136
5.5 球面全景图的融合.....	147
5.6 球面全景图的拼接实例.....	147
5.7 稳健的球面全景图全自动生成算法.....	150
5.8 采用分段插值的球状全景图自动拼接方法.....	155
5.9 球面图像拼接技术中还存在的问题.....	161
参考文献.....	162
第 6 章 立方体全景图图像拼接	163
6.1 立方体全景图图像拼接技术概述.....	163
6.2 立方体全景图像的映射方式.....	163
6.3 立方体表面上的图像拼接.....	178
6.4 立方体表面上全方位全景图像的生成和应用.....	184
6.5 基于纹理技术生成立方体表面全景图的算法.....	187
6.6 立方体重投影基本算法.....	191
6.7 立方体重投影的加速算法.....	193
6.8 变焦算法.....	194
6.9 反走样 (Anti-aliasing) 算法.....	195
参考文献.....	196
第 7 章 视频纹理的合成	197
7.1 纹理合成技术概述.....	198
7.2 纹理合成方法.....	199
7.3 视频纹理合成的基本方法.....	203
参考文献.....	213
第 8 章 高级动态全景图技术	214
8.1 基于动力学定常的时间控制技术.....	214
8.2 基于深度信息以及平滑迁移假设的动态全景图技术.....	231
8.3 基于分割的多层全景图技术.....	235
8.4 大运动高动态范围的全景图拼接方案.....	240
参考文献.....	248

第 9 章 虚拟漫游控制技术	252
9.1 虚拟漫游控制技术	252
9.2 Action Script 动画编程	260
参考文献	294
第 10 章 全景图虚拟漫游项目实例	295
10.1 实景虚拟漫游 NAVIRE 项目	295
10.2 海南国际旅游岛北美推广计划项目方案实例	303
参考文献	307

第1章 绪论

1.1 虚拟现实技术

1.1.1 虚拟现实技术概述

2009年好莱坞3D科幻电影《阿凡达》在为自身赢得票房价值的同时，也掀起了一股全球性3D影视热，其中最为吸引眼球的话题就是采用了虚拟现实技术。

虚拟现实（Virtual Reality, VR）技术，其目标是利用计算机系统以及相应的传感器设备（如数据手套、数据制服等）构建一种虚拟的现实环境，使人们在视觉、听觉、嗅觉甚至触觉上产生一种“身临其境”的感觉。据维基百科的定义：“虚拟现实是指一种计算机仿真的环境，既可仿真现实世界中某地方的物理存在，也可以仿真想象世界中的3D现实环境”。大多现有虚拟现实环境限于视觉体验，即感受计算机屏幕上3D场景的连续动态改变，有些需要戴上专用立体眼镜。近年来涉及医疗、游戏等应用声音、触觉等方式进行遥诊、遥控交互系统的设备也陆续面世。可见，除这些用于互动的传感器系统设备之外，通过计算机几何图形建模或者实景图像拼接与合成来构建其三维场景空间是实现其“现实”感受的重要环节、虚拟现实技术是一项综合技术，它涉及计算机图形技术、多媒体技术、传感器技术、人机交互技术、网络技术、立体显示技术以及仿真技术等多个领域，是光学、力学、图形图像学、数学等多学科相互交融的产物。

区别于其他交互应用系统，虚拟现实技术具有三个显著特征：沉浸感（Immersion）、交互性（Interactivity）及想象性（Imagination）。**沉浸感**是指用户借助交互设备可以沉浸于虚拟场景，即人们所看到的、听到的、嗅到的、触摸到的，好像与真实环境中感受到的一样。例如，人们走近一个建筑物，它就会变大，远离就会变小；玻璃在太阳下会发生反光；汽车走过会听到轰隆隆的噪声；等等，完全模拟出实际的场景。它体现了用户投入的程度，是虚拟现实系统中最基本的特性。**交互性**是指用户可以通过专门的输入、输出设备与虚拟场景中各种对象相互交流与作用。例如，用户自己打开一道门，则出现一个新的场景；又如，碰掉一个水杯，杯子会摔碎，水也会溅出，甚至溅出去的水会打湿你的衣服。交互性体现出人的控制能力及虚拟场景对信息的反馈能力。**想象性**是指虚拟现实技术通常应用于高科技领域，帮助人们解决尖端的问题或揭开不为人知的秘密，所以必须具备丰富的想象力。例如梦境漫游，或者通过虚拟现实对人的大脑工作的机制进行探索等。虚拟现实技术帮助想象变成了“现实”。

1.1.2 虚拟现实发展现状

虚拟现实的思想已经有将近 50 年的发展历程。1962 年 Morton Heilig 的一项题为“全传感仿真器 (Sensorama Simulator)”的专利获得美国专利号#3050870 的授权,它是第一个虚拟现实视频设备。Heilig 的“全传感仿真器”的主要组成部分有 3D 视频(由一对并排的 35 mm 摄像机采集)、运动、颜色、立体声音、香气、风和一个可以振动的座位。它可以仿真人骑车穿越纽约的过程,“骑车人”能感受到风,感受到路面的颠簸,当经过饭店时“骑车人”甚至能闻到食品的香味。

更早些时候,Heilig 的第一个专利号为# 2955156 的专利“头戴式电视 (Telesphere Mask)”则于 1960 年取得。正是在这一设计的基础上,Ivan Sutherland 于 1965 年提出了“The Ultimate Display”的概念。1966 年,Sutherland 在使用者眼睛前绑上两个 CRT 显示器,让使用者能看到立体的图像。大约 30 年后,现代的 HMD (Head-Mounted Display, 头盔式显示器)使用缩小的两个 CRT 或 LCD 安装在头部,原理上与 Sutherland 基本相似,只是那时的 CRT 显示器比今天的要重得多,所以 Sutherland 使用了一副机械臂来负担显示器的重量。该机械臂还具备检测用户视角的功能。大多数今天的 HMD 使用非接触式位置跟踪器(如磁跟踪器或超声波跟踪器),但后面这项技术在 20 世纪 60 年代还没有被发明出来。

Sutherland 进一步认识到可以用计算机生成 HMD 中的场景来代替摄像机拍摄的模拟图像,并开始设计这样的“场景生成器”。这便是图形加速器的先驱。早期的图形场景生成器大约于 1973 年由 Evans 和 Sutherland 研制成功,但在 1/20 s 内只能计算与显示 200~400 个简单的多边形。

这方面的研究引起了美国军方的兴趣。由于飞机模拟器可能花费数百万美元,而且是为特定的飞机型号设计的,当该飞机型号过时后,它的仿真器也被淘汰。如果仿真器能做到通用平台上,飞机型号的改变可能只需要进行少量软件上的更新。可见,使用新型技术的优点是明显的。20 世纪 70 年代和 80 年代早期,美国军方投入大量的经费,开展了大量有关“飞行头盔”和军用现代仿真器的研究,但这方面的研究成果大部分被定为机密,未能正式出版。随着国防资金的削减,这方面的研究资助大为减少。一些研究开始由军用转为民用,促使了虚拟现实技术在更广泛领域的发展。

另一个美国政府部门 NASA 则将现代仿真器用于帮助训练宇航员。由于实际生成一个与外层空间或遥远星球一致的环境很难或根本不可能,所以在 1981 年,在一笔很小经费的支撑下,他们研发了一个基于液晶显示器 (LCD) 的 HMD 原型,并命名为“虚拟显示环境显示器 (VIVED)”。NASA 科学家把 SONY 公司生产的产品“Watchman”TV 进行了简单的改装,在 LCD 前安装了特殊的光学镜头,用于把图像聚焦到眼睛相近的位置。NASA 科学家集成了 DEC PDP11/40 主计算机、由 Evans 和 Sutherland 设计的 Picture System 2 图形计算机和 Polhemus 的非接触类跟踪器,研制出第一个虚拟现实系统。跟踪器被用于测量用户的头部运动,并把测量数据传输给 PDP。PDP 主计算机再把这些数据编排后传给图形计算机,由图形计算机计算出新的图像,并立体显示在 VIVED 上。

1985 年后,Scott Fisher 加入该项目,他把新型的传感手套集成到仿真器中。手套最先是由 Thomas Zimmerman 和 Jaron Lanier 作为一个非程序员使用的虚拟编程接口开发的。到 1988 年,Fisher 和 Elizabeth Wenzel 生成了第一个能操纵 4 个 3D 虚拟声源的硬件系统。它们能生成定位

到任意空间位置的声源。这给仿真器提供了重要的扩展,原始的 VIVED 系统演变为 VIEW(Virtual Interface Environment Workstation)。以前的软件也被移植到 HP00 型功能更强的计算机上。

20 世纪 90 年代初,VR 研究热潮转移到民间的高科技企业。研究出第一套数据手套的 Jaron Lanier 成为销售 VR 产品的第一家商业公司 VPL 公司的总裁,这家公司卖的第一套传感手套叫“Data Glove”,第一套 HMD 叫“Eye Phone”。

美国虚拟技术公司提供的 VR 6 头盔式显示器 HMD,由于 HMD 佩戴在用户的头上,除 HMD 的图像外,用户看不到周围的场景,而且 HMD 中显示的内容可以随头的旋转和位移而变化,因而为用户提供了很强的沉浸感。

在其他方面,标准化的虚拟现实建模语言 VRML 为在 Internet 上构建一个可共享、可互换的虚拟环境奠定了很好的基础,VRML 本身也由早期的 VRML 1.0 发展到 VRML 2.0,再发展到 VRML 97,现在还在进一步发展中。分布式交互仿真标准化工作也由早期的 SIMNET 发展到分布式交互仿真 DIS,再发展到如今的高层体系结构 HLA。

虚拟现实系统已由单机系统发展到分布式虚拟现实系统,现在人们研究的比较热门的是支持协同工作的分布式虚拟现实系统,即协同虚拟现实系统。即由过去只支持多人共享,发展到以支持多人相互感知、协同操纵等为目的的协同感知。

虚拟场景的建模是虚拟现实系统中最关键的部件之一。过去负责实时渲染虚拟环境的虚拟现实引擎主要是依据计算机图形学理论,即首先对真实世界进行抽象,从而建立起三维几何模型,一般用多边形表示。在给定观察点和观察方向以后,利用计算机由模型实现多边形绘制、着色、消隐、光照以及投影等一系列绘制过程,产生虚拟环境。这种基于图形渲染技术所面临的主要问题是:一方面需要高性能的图形工作站,系统成本非常昂贵;另一方面需要三维建模,工作非常烦琐,而且往往需要专业人员,这部分的费用也相当高。为此,基于图像渲染技术被提出,并在近几年引起了人们的广泛关注。基于图像渲染与基于图形渲染不同,它是直接由照片图像来构造虚拟环境。这方面的研究起步较晚,还有大量的工作要做,但对虚拟现实系统的大面积推广应用影响较大。

虚拟现实开发工具的发展也很迅速,1992 年英国 SENSE 公司开发出“World Tool Kit”,该软件包提供了一系列函数,用于支持从更高的层次上开发虚拟现实应用,而且为用户屏蔽掉底层硬件上的差异,使用户只需要分析与实现虚拟现实应用本身,不需要考虑不同机型、不同操作系统、不同接口的硬件差异。World Tool Kit 使开发虚拟现实应用变得更加科学化,测试时间也大大缩短。随后,类似的软件相继推出,比较著名的有 Vega、MR、dVS 等。

虚拟现实技术应用已经非常广泛,包括军事、医学、教育、制造、航空、航天、铁道、建筑、土木、科学计算可视化、娱乐、艺术、体育等各个领域。

目前大规模的军事模拟系统已经被许多国家用作军事训练。例如,美国陆军的自动虚拟实验室 CAVE 是一个典型的虚拟现实系统;波音公司及洛克希德·马丁公司均采用 SGI 虚拟现实系统提供在虚拟环境下研究制造联合 JSF 战斗机的方案,使两家公司无需制造真正的原型试验机即可获得宝贵的数据。

汽车、建筑等设计仍大量需要虚拟现实技术的支撑,借助虚拟现实技术,设计师们能够观察他们设计的效果,包括外观效果、动力性能以及使用方便性等。特别是在设计产品投入生产前,用户需要通过虚拟现实技术考察产品设计的优劣。比如,建筑设计师为用户设计一栋大厦,如果用户在设计完成后需要“进入”大厦中漫游,以便决定大厦各房间的规划、装

饰材料的选择等是否满意,而此时大厦还未开工。在这种情况下,必须采用虚拟现实技术,根据建筑师的设计生成一个虚拟大厦,并通过虚拟现实交互手段,支持用户在其中的漫游。

科学计算可视化(Scientific Computation Visualization)主要解决如何通过虚拟现实的手段,生动地表现科学数据的内部规律与计算过程,如天气云图的运动规律、空气湍流的特性等。而信息可视化(Information Visualization)则更进一步,主要用于表现系统中信息的种类、结构、流程以及相互间的作用等。信息可视化能有效地揭示复杂系统内部的规律,解决无法定量而定性又很难准确表达的问题。

虚拟现实技术在医学领域可用于远程手术及外科手术仿真。比如医生远程做手术,医生在远地通过虚拟现实技术观察病人的身体,医生进行手术的动作通过通信技术传输到远地的病人处,由一个机械手真正地实施手术,结果再传输到医生处,并通过虚拟现实技术表现给医生。这样,医生便能够像在本地一样观察病人、实施手术,看到手术的结果,并重复上述过程动作,以完成一个完整的手术。

在教育方面,虚拟现实技术可以让学生“身临其境”地学习,不再是“纸上谈兵”。类似地,采用虚拟场景,可以完成各种危险或现实中成本高昂的工作模拟训练,如消防队员的救火训练等。本书作者之一姚孝明博士负责的国家863项目“基于宽带网络的实景三维城市生活服务”即是上海杰图公司的“城市吧”网站转换为数字电视的一个尝试,使得人们“足不出户”就可以通过遥控器享受城市生活中“看房、选房”“逛街、购物”等各种商务活动的视觉信息服务的便利。

除此之外,虚拟现实技术在创意产业、影视业、动漫游戏等诸多领域的应用不胜枚举。

1.1.3 3D 虚拟场景构建技术

据统计,人们所能感知的外部客观世界的各种信息有60%~80%是由视觉提供的,因此,虚拟现实技术的关键在于生成高质量的虚拟场景。目前,3D虚拟场景的构建技术主要有两种,一种是利用3D技术生成几何模型,再利用着色、渲染、消隐、光照、投影等手段增强真实感,产生虚拟场景。这种方法是基于图形学的方法,简称为建模法(Geometry-Based Rendering, GBR)。另外一种是通过一组自然场景的图像序列加工合成,实现仿真的效果。这些加工处理是根据用户提供的观察参数、视角和视点,对图像进行适当的变换和拼接,合成新的浏览视图展现给用户。这种方法是基于图像学的方法(Image-Based Rendering, IBR)。

GBR生成场景的效果很大程度上依赖模型的复杂程度,而模型的建立通常需要手工创建,要消耗大量的人力、物力和时间。这种方法需要进行大量的计算,对系统硬件的计算能力和图形处理能力都有很高的要求,一般只有高性能图形工作站才能完成。另外,现有的3D技术不能满足虚拟现实技术对漫游实时性的要求。

IBR具有GBR不可比拟的优点。由于它采用的是真实图像,能够真实地反映环境物体的形状和色泽的明暗、材料的纹理等特性,而不需要额外的模拟,因此其逼真度非常高。同时,IBR不需要复杂的图形建模和大量的计算,所以图像生成速度快,实时性好,硬件要求低,普通个人计算机即可以满足硬件的要求。这些特点使得IBR技术被视为一种具有很强实用性的技术,吸引了大量的学者和研究机构投入相关研究中,并得到了相应的发展。

目前,常用于IBR技术的方法主要分为以下几种:

1. 基于计算机视觉的方法

这种方法是通过多幅视图获得对象的 3D 信息，再直接或间接利用这些 3D 信息，使用视图插值、视图合成、视图变形 (Morphing) 等方法从已有的若干视图中重建其中间视图，其基础理论与方法是计算机视觉中的外极几何 (Epi-Polar Geometry) 理论。

2. 基于分层表示的方法

这种方法与图像拼接正好相反，它针对视频系列将 3D 场景分成运动独立的、由仿射运动模型描述的不同层次。每一层都产生一个 2D 图像流和一个 2D 变换流，最终组合到现实屏幕上。由于该方法可以以不同层次表现场景，所以它可以做到将场景中变换较快的前景和变换较慢的背景区分开来，或者在复杂场景的绘制中进行进一步的分解，将不同的场景对象也加以区分，赋予不同的绘制质量。在场景的压缩存储上也可进行不同的压缩编码，达到更好的压缩效果。事实上，这种方法已经成为 H.264 国际标准中场景结构的基本思想。

3. 基于全光函数的方法

全光函数描述了从空间任意点所看到的全部光线。用计算机图形学术语，它描述了给定场景中所有可能的环境映射。基于全光函数的方法视图捕获空间任意区域内的完全光流。这种方法从一些有向的离散样本中重构连续的全光函数，然后通过新的视点位置重新取样该函数来绘制新的视图。因此，若只需得到指定视点所看到的视图图像，采用全光函数就可以了。全光函数用于描述在指定视点所看到的所有光线，只要给出该视点所看到的光线，即可得到从该视点所看到的视图图像。

4. 基于全景图的方法

这种方法的思想是把同一个场景的若干幅图像通过图像配准找到图像间的重叠区域，然后将重叠部分对齐，进行平滑缝接，直至所有的图像都被加入拼接图中，组合成一幅完整的图像。该方法通常用于生成全景图像 (Panoramic Image)、图像稳定、图像分辨率增大、图像压缩以及视频扩展等多种实际应用系统中。它是基于全景图像通过拼接构建虚拟场景，因此该方法有时也被称为是基于图像拼接的方法。

基于全景图的方法相对于其他方法，有很多优点：① 图形绘制独立于场景复杂性，仅与所需生成画面的分辨率有关，因此能用于表达复杂的场景；② 预先存储的图像 (或环境映照) 既可以是计算机合成的，也可以是实际拍摄的画面，而且两者可以混合使用；③ 绘制技术所需计算量相对较少，对计算资源的要求不是很高，因而可以在普通工作站和个人计算机上实现复杂场景的实时显示。此外，基于全景图的方法还有采样简单、绘制速度快等优点。

1.2 全景图拼接技术

1.2.1 全景图的概念

全景图是指一张包含大于双眼正常有效视角 (大约水平 90° 、垂直 70°) 或双眼余光视角

(大约水平 180° 、垂直 90°) 甚至 360° 范围的场景图像。全景视图方法的基本思想是在一个给定视点采集各个方向的实景图像, 然后把这些实景图像投影到一个圆柱的内表面或者球体的内表面, 进行视图拼接, 最后利用反投影技术重新获得各个视角方向相应的场景图像。大家知道, 平时拍摄的图像往往反映局部场景 (人们朝某一个方向所看到的场景), 这种图像称为局部图像。对应的, 能反映全局场景的图像称为全景图像。全景图像是从某一个视点出发, 能够观察到该点各个方向的合适的描述形式之一, 全景图像 $P_i(\theta)$ 描述了视点在 θ 处的视点空间。

全景视图是指在一个固定的观察点, 能够提供水平方向上方位角 360° , 垂直方向上 180° 的自由浏览, 简化的全景只能提供水平方向 360° 的浏览。全景图像的获得通常有两种方法: 全景拍摄和通过图像拼接。前者需要特殊的设备, 操作起来非常方便, 简单易行, 但这种设备价格非常昂贵, 不适合普及; 后者只需要普通的相机就可以, 但是需要进行相应的图像投影和拼接。所谓图像投影是指把实景图像投影到一个统一的圆柱或者球体表面的过程, 这样可以消除图像间存在的旋转关系, 只保留平移关系, 适合拼接。图像拼接是指两幅不同视角方向具有一定重叠部分的图像合成一幅图像。显然后者的重点在于图像的拼接和融合, 一般建议每两幅图像的重叠度应该在 $30\% \sim 50\%$ 。全景图的变换示意图如图 1.1 所示, 该全景图是柱面全景图。

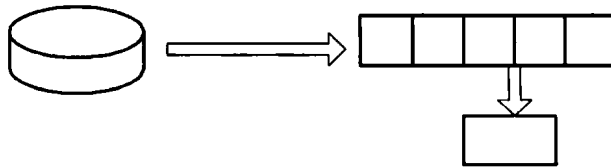


图 1.1 圆柱全景图生成示意图

获得全景图后, 用户选取任意视角进行观察, 相应视角图像的绘制过程其实质是纹理图像取景变换后在简单景物几何上重投影的过程。把平面全景图反投影回柱体或者球体, 使人们能够进行逼真的浏览。全景图拼接方法最大的特点和优点是易于实现、效率高、漫游速度快, 不需要很高的硬件配置, 以及在普通计算机上都能够实现场景的漫游。

1.2.2 全景图的分类

根据全景图投影模型的不同, 通常可将全景图分为三类: 圆柱型全景图、立方体全景图、球面全景图。

圆柱型全景图技术较为简单, 发展较为成熟, 是大多数构建全景图虚拟场景的基础。这种方式是将全景图像投影到一个以相机视点为中心的圆柱体内表面, 视线的旋转运动即转化为柱面上的坐标平移运动。这种全景可以实现水平方向 360° 连续旋转, 而垂直方向的俯仰角度则由于圆柱体的限制要小于 180° 。柱型全景有两个显著优点, 一是源图像的采集方式比其他类型全景图简单; 二是由于柱面可以展开成一个矩形图像, 因此可以直接利用二维图像的方式对其进行存储和访问。

立方体全景图由 6 个平面投影图像组成, 将全景图投影到一个立方体的内表面上。这种方式下图像的采集和相机的标定难度较大, 要求使用特殊的拍摄装置, 依次在水平、垂直方向每隔 90° 拍摄一张照片, 获得 6 张可以无缝拼接于一个立方体的 6 个面上的照片。这种方法可以

实现水平方向 360° 旋转、垂直方向 180° 俯仰的视线观察，相对于柱型全景图是一个进步。

球面全景图是指将源图像拼接投影成一个球面形状，以相机视点为球心，将图像投影到球体的内表面。与立方体全景图相似，球面全景图也可以实现水平方向 360° 旋转、垂直方向 180° 俯仰的任意视线观察。球面全景图的拼接过程及存储方式较柱型全景图大为复杂，这是因为生成球面全景图的过程中需要将平面图像投影成球面图像，而球面为不可展曲面，因此这是一个平面图像水平和垂直方向的非线性投影过程，同时也很难找到与球面对应且易于存取的数据结构来存放球面图像。事实上，在漫游体验过程中采用球面全景图常常可以觉察到明显的局部形变。图 1.2 所示为基于球面模型的全景图截图，其中在左下部分的小车形状产生了较为明显的扭曲形变。



图 1.2 基于球面模型的全景图截图

1.2.3 全景图拼接技术的应用

基于图像的绘制技术在短短几年内虽然得到了很大发展，但是在 IBR 技术的各种方法中，最成熟也最实用的还是全景图技术。这是因为全景图的获得相对比较简单，可以采用普通相机或数码相机拍摄多张实景图像通过拼接获得完整的全景图像。另一方面，作为一种场景存储和显示方式，全景图技术要求的数据量不大，绘制的实时性好，而且具有一定的交互能力，能在一定程度上模拟用户的现场视觉感受。随着网络技术的发展，人们可以通过网络浏览远处的真实场景。

全景图目前已经在很多领域得到了广泛的应用。

(1) 旅游风景区。

以优美的 360° 全景照片，显示景区内的优美景点，给旅游者以身临其境的感觉，可以用来制作风景区的介绍光盘，也可以作为旅游公司吸引游客的绝好工具。

(2) 房屋销售。

房屋开发销售公司可以利用虚拟全景浏览技术，展示楼盘的外观，房屋的结构、布局、室内设计，并可以用来制作楼盘的介绍光盘。购房者在家中通过网络即可仔细查看房屋的各个方面，提高潜在客户购买欲望。更重要的是，采用全景技术可以在楼盘建好之前将其虚拟设计出来，方便房地产开发商进行期房的销售。

(3) 宾馆、酒店。

利用网络，远程虚拟浏览宾馆的外形、大厅、客房、会议厅等各项服务场所，展现宾馆

舒适的环境、完善的服务，给客户以实在感受，促使客户预定客房。

(4) 展览会、展览馆的虚拟现实展示。

人们可以远程浏览展览会，仿佛置身于真实的场景之中。

(5) 教育领域的应用。

建筑工程学：交互性地参观还未完工的办公大楼，找寻装饰的构思；或参观房屋模型，学习建筑原理；参观世界各地经典建筑，寻找建筑设计的灵感。考古学：参观世界上你不可能到达的博物馆，研究从未对公众开放过的私人收藏的绘画或雕塑。医学教育：学生可以通过解剖一具虚拟的尸体来学习解剖学，也可以观看血细胞通过心脏的全过程；医生用从实际病人身上收集来的数据进行仿真，对手术或其他过程进行周密的计划，例如，观看胃镜的过程，了解病变组织的特征。导游培训：让学生参观世界各地虚拟的风景名胜，并学习这些名胜的历史、特点、文化内涵等。生物教育：操纵分子模型，观察不同药物的立体结构图像。历史教育：进入另一个时代，如参观商代的集市，或参加唐代的盛典。化学和物理教育：昂贵实验仪器的介绍与展示，参观你不可能进入的实验空间，如核反应堆、粒子对撞空间等。社会科学：参观世界各地的社会风情，了解各地的生活形态与习性。

(6) 智能监控的应用。

作为机器视觉，相对于固定场景来讲，生成的全景图可以“记忆”连续的场景，当场景中出現改变的图形时，可以与存储的全景图进行对比，发现其中的不同。在图像的传输方面，可以通过差别化图像，只将与原来图像不同的区域进行传输而不必整幅传输，这对于信道较小的系统来讲，可以提高网络的传输效率，达到较好的监控效果。

1.2.4 全景图拼接技术的发展趋势

利用全景图像建立虚拟环境，这个想法早在 20 世纪 70 年代就产生了。1978 年，美国麻省理工学院 MIT 的媒体实验室开发了一个称为“白杨树镇电影地图”（Aspen Movie Map）的系统，首次利用实景图像，开车穿行白杨树镇这个小城市的街道，四架摄像机同时拍摄了四个方向的景物，这些照片以超文本方式进行组织，然后装入模拟的视盘中，并在每个街道的路口加入了分支的交互手段，进行街道的漫游。播放时，用户可以通过触摸屏和游戏杆来控制自己前进的速度和去向，就像自己在开车。对小城的一些有名的场所，还可以停下来保存有关的资料（资料可以是图像、声音、文本、视频……）。同时，提供了一张鸟瞰图作为导航图，用来标志用户位置，提示关键场所和位置信息。

20 世纪 80 年代初的 LIPPMAN 制作的电影图系统应该是最早基于图像绘制的系统之一。在每一个视点上，前后左右各拍摄一幅照片，这四幅照片形成了这个点的场景，用户可以在这个点上朝各个方向浏览。当人在环境中走动时，系统就会根据人位置的远近选取合适的视点。

1991 年，Apple 公司的人机接口实验组（ATG）建立了一个基于实景的成像环境。1994 年 6 月，Apple 公司推出 QuickTime VR（QuickTime Virtual Reality）系统，该系统是全景图技术的典型杰作。它是一种基于实景图像的、在普通微机平台上能够实现的虚拟现实技术，其基本特点是能够对一个物体或空间进行 360° 的全面观察。在每一个关键点，可以实现视线方向的连续变化，通过关键点的跳转来实现场景的漫游（前后帧画面间有时会出现不连续现象，

产生跳跃感)。QuickTime VR 是一个基于图像绘制技术的虚拟现实系统，第一次使人们感受到了逼真的虚拟现实环境，利用软件把若干张有一定边缘重叠的实景图像拼接在一起，拼接成一张无缝隙的 360° 全景图像；其全景图像在压缩过程中被分成大小相同的块，存储为 QuickTime MOV 格式，通过超文本系统 HyperCard 来制作热点，将不同视点的实景图像进行相关链接。

Microsoft 公司在 1995 年 3 月开发了基于 Windows 的全景视频 Surround Video 系统。该系统直接利用硬件，通过全景照相机采集全景图像。两者的功能基本相同，但制作全景视频的方法完全不同。1995 年以后，全景技术进入迅速发展的时期，相应的软件产品越来越多，功能越来越完善，Inc 公司的 OmniView 系列软件、Interactive Picture Corp. 公司的 IPIX 系列软件，能够提供专门格式的全景图像进行浏览和制作，提供场景任意角度的浏览漫游、旋转以及缩放。IPIX 提供有在 IE 中运行的插件支持，用户目前已经可以在 IE 中打开一幅全景图像了。值得一提的是，Meta Stream 公司和 IBM 经过长期的合作并投入了大量资金，终于在网上发布了基于纹理映射技术的物体电影播放插件，通过它用户可以以任意角度对物体进行拖动观察。

近年来，基于视频的动态全景图技术在国外也有了长足的进步。例如，Alex 等人提出打破传统“亮度定常”假设，改以局部“动力学定常”假设为原则来匹配动态场景，能够较好地实现视频重组与基于时间曲线的可控性，但其以牛顿惯性定理为指导的配准原则只适合小幅匀速运动，难以处理更复杂的情况；Zhi 等人则提出以深度信息和平滑迁移假设来配准全景图，可以得到“视觉上与原视频一致”的合成动态全景图；Min 等人则结合运动与深度信息采用张量投票的分割框架，提出了基于分割的多层全景图的方法，其背景质量得到很好的改善，但动态性处理不灵活。其中，基于时间曲线的动态全景图技术能够有效地保留全景图内部运动的纹理，如瀑布、流水等，并且能够根据需要调整时间改变视频流的内容，例如，使得原本是冠军的游泳运动员在新的全景图内“看”起来却成了“落选者”。这又对视觉与真实的关系提出了新的质疑证据。

此外，近年来一些针对曝光度差异较大的图像拼接技术也迅速得到国内外广泛的关注。图 1.3 是姚孝明博士指导硕士生采用相机响应曲线估计得到的结果。其中，(a)、(b) 两幅图像为已经匹配好的具有比较大的曝光度差异的图像，(c) 图是采用相机响应曲线校正拼接的全景图像。

国内高校目前在基于图像绘制技术的虚拟现实技术研究方面也取得了较大进展。据资料调研，上海龙华博物馆与新加坡合作，利用 QuickTime VR 技术建立了一个虚拟博物馆；另外，国防科技大学的虚拟实景空间系统 HVS 是目前国内做得较为完整的一个系统，通过合适的空间模型把多幅全景图像组织成为虚拟实景空间。其他如清华大学、浙江大学、中国科学院软件研究所计算机科学实验室、中南大学、西南交通大学、重庆大学和四川师范大学也都曾经做过这方面的研究。本书作者之一姚孝明博士在其博士论文中就针对场景构成的视觉真实与物理真实做了一系列的理论与实验研究尝试，为虚拟真实场景的艺术再现提供了一些可以借鉴的思路。

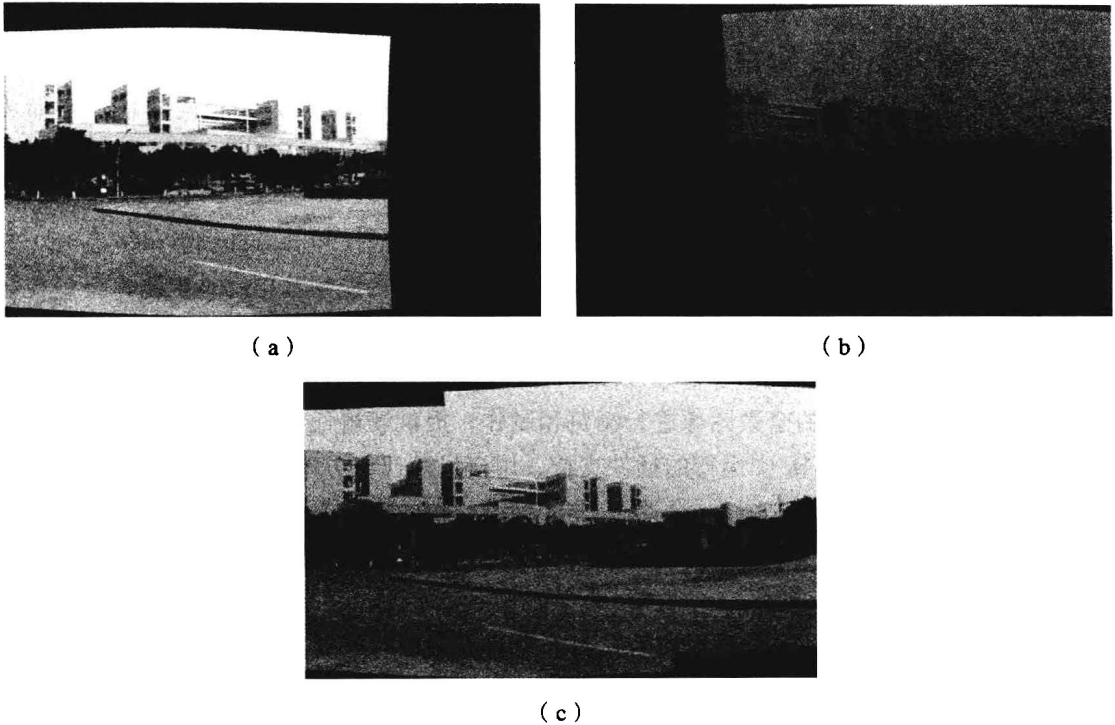


图 1.3 基于相机响应曲线参数估计的大曝光差异图像拼接效果

1.3 虚拟漫游技术

1.3.1 虚拟漫游概述

虚拟漫游技术是虚拟现实技术的重要组成部分，其目标是解决“漫游者”与虚拟场景的交互动态关系，由此产生视点、视线以及场景内容的关联变化，构成一幅幅“漫游”体验的画卷。按照 3D 场景建模方式的不同，可以将虚拟漫游分为基于图形的虚拟漫游和基于图像的虚拟漫游以及两者混合兼有的虚拟漫游三种类型。本书主要讨论基于图像的虚拟漫游。

基于图像的虚拟场景又称为虚拟实景空间。虚拟实景空间是指利用照相机采集的离散图像或摄像机采集的连续视频作为基础资料，经过图像处理生成全景图像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。

虚拟场景的漫游是指用户在已经定义好的虚拟场景中进行交互式浏览，如前进、后退、转弯等。虚拟场景的浏览动作一般可分为 3 个类型。

(1) 定点浏览。

定点浏览是指在某一个视点上浏览，浏览动作可以分为环视、仰视、俯视、斜视、变焦距、变时间浏览 6 种。

(2) 运动浏览。

运动浏览是在不同视点中浏览，这是对人在真实世界中漫步的抽象。运动浏览包括连续