

# 基于采样数据集的真实感 建模技术

## Realistic Modeling Techniques Based On Real-World Sampling Dataset

(申请清华大学工学硕士学位论文)

院(系、所) : 计算机科学与技术系

专 业 : 计算机应用技术

研 究 生 : 张 竞 丹

指导教师 : 孙富春 副教授

联合指导教师 : 邓志东 教授

二零零三年六月

## 关于论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：学校有权保留学位论文的复印件，允许该论文被查阅和借阅；学校可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存该论文。

**(涉密的学位论文在解密后应遵守此规定)**

签 名：\_\_\_\_\_ 导师签名：\_\_\_\_\_ 日 期：\_\_\_\_\_

## 摘 要

随着计算机图形学、计算机视觉和人机交互技术的迅速发展，虚拟现实技术在原型产品设计与分析、影视娱乐、军事训练等领域得到了广泛的应用。传统的虚拟场景建模技术在建模效率、真实感与实时交互等方面已经无法满足要求。如何有效地从真实世界中获取有用的信息，并且把这些信息进行有效组合，用于真实感建模已经成为目前研究的热点之一。本文从基于图像建模和基于样本建模这两方面对这个问题进行了研究。

在基于图像建模方面，我们设计并实现了一个具有纹理的三维物体自动建模原型系统。这个系统运用立体视觉原理测量物体的几何信息，自动构建真实物体的三维模型；并且可以同时获取物体的颜色纹理以及所对应的纹理坐标。用这个系统得到的数据可以直接用于渲染。围绕这个反向建模系统，我们探讨了系统的标定、几何信息的采集、曲面参数化、颜色纹理的拼接和融合等方面的问题。同时我们给出了用这个系统进行建模的实例。

在基于样本建模方面，本文提出了一个将双向纹理函数(BTF, bidirectional texture function)样本合成到任意曲面的算法。BTF是一个六维的函数，能够描述物体表面随位置变化而变化的微小几何结构和光线反射特性。我们的BTF合成算法从样本BTF中抽取出纹理元信息，并用这些信息进行表面BTF合成。同时，我们提出了一个普遍适用的算法，即所谓k-连续算法来加速BTF的合成过程。用这个算法得到的模型表面的BTF不但表现出和样本BTF相似的光照特性，而且能从不同的光线和视角方向表现出模型表面的微小几何结构。

关键词：基于图像的建模，立体视觉，系统标定，曲面重构，纹理映射，双向纹理函数，三维纹理元，表面反射和阴影模型，纹理合成，多传感器融合

## Abstract

With the rapid development of computer graphics, computer visions, and HCI technologies, virtual reality has been extensively applied to many areas such as virtual prototype designs , entertainment industries, and military trainings. The traditional modeling methods for constructing 3D virtual environments are not capable of meeting the requirements of many applications due to their low efficiencies of modeling and bad fidelities. How to get useful information from real-world environments and assemble it onto a virtual world has been viewed as a hot topic in recent years. This thesis attempts to address such a challenging problem from both an image-based modeling approach and a sample-based modeling approach.

At first, we designed and implemented a prototype system to automatically reconstruct a geometric model and corresponding textured surface of 3D physical object. This system took advantage of stereo vision-based methods to acquire 3D data points through sampling on the surface of physical object. All the resulting geometric models, textures, and texture coordinates were used in a standard rendering pipeline. We discussed a collection of procedures involved in the whole system, including system calibration algorithms, 3D points acquisition algorithms, surface reconstruction algorithms, and texture registration and blending algorithms. An example was used to illustrate the proposed approaches and verify the performance of system.

Second, we developed a synthesis method for an arbitrary surface using a sample bidirectional texture function (BTF). BTF is defined as a 6-dimensional function that describes nature textures arising from both spatially-variant surface reflectance and surface mesostructure. Our approach accomplished a BTF-based synthesis using surface textons, which were extracted from the sample BTF. In order to speed up the process of such a BTF-based synthesis, we proposed a general search strategy, i.e., a k-coherence search algorithm.

Our experiments showed that the surface BTF obtained by using the above synthesis approach not only kept the same viewing/illumination properties as the sample BTF but also visually retained similar mesostructures even if we observed it at different viewing and lighting directions.

**Key Words:** Image-based modeling, Stereo vision, System calibration, Surface reconstruction , Texture mapping , Bidirectional texture function, 3D textures, Reflectance and shading models, Texture synthesis, Multi-sensor fusion

## 目 录

摘 要 .....	I
ABSTRACT .....	II
第一章 绪论 .....	1
1.1 项目背景及意义 .....	1
1.2 三维自动建模的研究现状和发展趋势 .....	2
1.2.1 几何信息的获取方式 .....	4
1.2.2 曲面重构 .....	7
1.2.3 纹理映射和纹理坐标的获取 .....	8
1.2.4 物体表面光照模型和真实感渲染 .....	12
1.3 本文的主要贡献与章节安排 .....	15
1.3.1 本文的主要贡献 .....	15
1.3.2 本文的章节安排 .....	16
第二章 三维物体反向建模原型系统的总体设计 .....	17
2.1 系统的总体设计目标 .....	17
2.1.1 系统设计需求分析 .....	17
2.1.2 系统所应具有的特点 .....	18
2.2 系统的配置 .....	18
2.2.1 硬件配置 .....	18
2.2.2 软件配置 .....	19
2.3 系统的结构和工作流程 .....	19
2.3.1 系统总体结构图和子系统划分 .....	20
2.3.2 系统工作流程 .....	21
第三章 系统标定算法的研究 .....	22
3.1 引言 .....	22
3.2 坐标系的定义 .....	22
3.3 摄像机内参数的标定 .....	25
3.3.1 线性摄像机模型 .....	25
3.3.2 标定算法分类 .....	26
3.3.3 基于多个自由平面的标定算法 .....	27
3.4 摄像机坐标系到世界坐标系的标定 .....	30
3.4.1 标定算法叙述 .....	30

3.4.2 双摄像机立体视觉原理 .....	31
3.5 机器人坐标系到世界坐标系的标定 .....	32
3.5.1 算法的基本思路 .....	32
3.5.2 机器人坐标系到世界坐标系的标定 .....	34
3.6 小结 .....	34
<b>第四章 具有真实纹理的反向建模方法的研究 .....</b>	<b>36</b>
4.1 引言 .....	36
4.2 几何模型的构造 .....	37
4.2.1 结构光标在图像上的抽取算法研究 .....	37
4.2.2 双目特征点的对应关系和空间坐标的确定 .....	39
4.2.3 曲面重构及其平滑 .....	40
4.3 纹理和纹理坐标的获取 .....	41
4.3.1 纹理和纹理坐标的获取 .....	42
4.3.2 纹理的拼接和融合 .....	43
4.4 重构结果与讨论 .....	44
<b>第五章 任意曲面上 BTF 的合成算法研究 .....</b>	<b>48</b>
5.1 引言 .....	48
5.2 系统基本结构 .....	50
5.3 三维纹理元的分析和抽取 .....	51
5.4 曲面上 BTF 的合成 .....	54
5.4.1 单尺度 BTF 合成算法 .....	54
5.4.2 多尺度 BTF 合成算法 .....	56
5.4.3 一个加速算法 .....	57
5.5 曲面 BTF 的渲染 .....	58
5.6 结果和讨论 .....	59
<b>第六章 结束语 .....</b>	<b>63</b>
6.1 结论 .....	63
6.2 进一步的工作 .....	63
<b>参考文献 .....</b>	<b>65</b>
<b>致 谢 .....</b>	<b>71</b>
<b>个人简历 .....</b>	<b>72</b>

## 第一章 绪论

### 1.1 项目背景及意义

虚拟现实 (Virtual Reality, 简称 VR) 是指利用具有强大图形处理能力的高性能计算机软硬件资源, 以及各种虚拟现实外部设备, 为用户提供一个具有自然交互能力和沉浸感的虚拟环境的综合集成技术。利用虚拟现实技术, 计算机可以产生一个三维的、基于感知信息的临场环境, 该环境对用户的控制行为做出动态反应, 并被用户的行为所控制。目前的研究主要集中在视觉、听觉、触觉方面。虚拟现实技术的主要特征可以概括为: 多感知性、沉浸感 (临境感)、实时交互性、自主性、集成性。

虚拟现实技术中一个非常重要的研究方向是在计算机图形学方面展开的, 为了给用户在视觉上的沉浸感, 系统需要实时的渲染具有真实感的三维图形。系统中三维模型渲染的逼真性在很大程度上决定了用户的沉浸感。目前构造虚拟场景的方法主要是通过人工对物体及环境进行三维建模, 其核心是计算机辅助设计(CAD)技术。运用这种方法, 工作量大, 并且需要将抽象的艺术思维通过冗长枯燥的建模过程表现出来, 这极大地限制了美术工作者的创造性。同时, 此三维模型一经建立, 就基本固定下来, 如需改动, 要有原创人员辅助, 可扩展性差。并且采用这个方法, 确定面片的纹理坐标也是一个非常复杂的问题, 需要大量的手工操作。在我们课题组为江苏科技宫开发 VR 游戏的过程中, 三维物体的真实感建模问题就一直困扰着我们, 让我们花费了大量的时间。而在现实世界中, 存在大量的可以利用的三维物体原型。因此研究从真实世界中如何快速地获取具有真实感的有效模型, 具有非常现实的意义。

三维反向建模技术在工业制造领域已经有广泛的研究, 并且技术也趋于成熟。但这样的技术并不能直接运用于虚拟现实系统的建模, 其主要原因是工业制造领域的反向建模主要考虑的是物体三维几何信息的精确表示, 一般不考虑物体表面的纹理信息和光照模型的获取, 这样得到的三维模型有大量的面片而没有表面纹理和材质数据。而在虚拟现实系统中, 是强调模型在视

觉上的真实感和实时渲染。考虑到硬件渲染速度的限制，需要三维模型有较简单的几何模型和具有真实感的纹理，而无需强调三维模型在细节上的几何准确程度。同时运用于工业的建模系统往往结构复杂，价格昂贵，这也大大限制了这样的系统运用于日益普遍的虚拟现实建模中。因此需要找到适用于虚拟现实应用的，即考虑到真实感建模的，快速、有效，且价格便宜的建模方法。

近年来，基于图像的建模和渲染技术（IBMR）引起了图形学界和计算机视觉界的广泛的兴趣。直接通过图像进行渲染（IBR）的技术因无法解决光照变化和交互性问题，而且数据量巨大，因此研究的热潮逐渐消退。而基于图像的建模技术(IBM)，包括几何模型建模、物体光照模型建模，能充分运用已经发展成熟的三维渲染技术，因此成为目前的研究热点。

基于以上考虑，传统的建模方法在建模的速度、效率、模型的真实感等方面已经无法满足要求，形成了虚拟现实应用的瓶颈。运用计算机视觉技术和基于图像的建模技术，直接从真实物体获取信息，进而建立物体的三维几何模型和表面光照模型的自动建模技术为这个问题提供了一个很好的解决方案，并且为虚拟现实技术的广泛运用铺平了道路。

## 1.2 三维自动建模的研究现状和发展趋势

三维自动建模技术的目的是利用多种传感装置获得一个真实物体的几何和色彩等信息，并将这些信息进行合理的组织，以达到渲染显示、数据计算与分析等目的。这方面的技术具有非常广阔的应有领域，包括 CAD 建模、影视娱乐、医学成像、视景仿真、空间定位等。

在工业制造领域，三维自动建模技术已经趋于成熟并且有广泛的应用。这个领域的建模系统的主要的特点是：系统的测量精度非常高；主要强调几何模型的获取而忽略物体表面纹理信息和光照模型；系统操作比较复杂，需要专业人员维护；系统的成本比较高。这样的系统主要适用于专业的集成制造厂家，不适用于一般用户建立可进行真实感渲染的模型。

在这个领域中，在国内比较有名的有华中理工大学图像识别与人工智能研究所研制成功的我国第一台三维激光彩色扫描硬件系统。在国外，成熟的

产品有 FARO 公司<sup>[76]</sup>、FOMER CimCore 公司<sup>[77]</sup>、Immersion 公司<sup>[78]</sup>的接触式测量机器人；3D Scanners Ltd 公司的 ModelMaker 系列<sup>[79]</sup>。

近几年，自动建模技术也不断的应用于其它领域，并且根据具体应用范围的不同，具有不同的特点。如在 Stanford 大学的 Digital Michelangelo 项目中<sup>[75]</sup>，自动建模技术用于获取雕像的精确几何模型。由于雕像很大并且它的位置不能移动，因此需要有比较灵活的数据获取手段，以及鲁棒的算法将雕像不同位置的几何信息进行融合。在 Berkeley 大学的数字化校园项目中<sup>[27]</sup>，自动建模技术用于获取校园建筑物的模型。因为建筑物非常庞大，不可能用一般的三维扫描技术获得，因此它采用了手工辅助建模和基于视角纹理的技术。

自动建模技术也被应用于真实感模型的构造，Cyberware 公司的激光扫描系统<sup>[80]</sup>是有代表性的产品。它主要是通过激光扫描测量物体的几何信息，然后通过摄像机获取物体表面的纹理信息，最后把这些信息进行融合。3rdTech 公司的 DeltaSphere-3000 激光全景扫描仪<sup>[81]</sup>也是有名的产品，它同样是通过激光扫描和摄像机获取物体的几何和纹理信息，主要适用于建立大场景的全景图。总体来说，此类系统的性能不错，并且价格可以被专业的模型制造厂商接受。但由于它采用了激光测量系统，因此系统的成本还是比较高。

如何用一般的设备，如工业摄像头、普通照相机，进行低成本和更加灵活的三维建模是现在研究的热点之一。此类技术大多局限于学术研究或样品阶段，还缺乏成熟的产品。比较有代表性的有：用单照相机获取人脸三维模型的系统<sup>[8]</sup>、用双摄像机获取物体几何模型的系统<sup>[1]</sup>、用一般投影仪配合摄像机获取物体模型的系统<sup>[82]</sup>、采用结构光方法的系统<sup>[6]</sup>、采用轮廓线方法获取物体动态模型的系统<sup>[83]</sup>。这些原型系统在一定程度上解决了三维建模的问题，但是都存在鲁棒性、适用性等方面的问题。例如在<sup>[8]</sup>提到的系统只能用于人脸模型的建模，而<sup>[82]</sup>中提到的系统对物体表面的反射特性有较高的要求。

近年也出现了测量物体表面光照特性的系统，这些系统主要应用于物体的真实感渲染。这方面有代表性的是测量物体表面 BRDF 和 BTF 的系统

[63,84]，测量人面部皮肤的反射特性的系统<sup>[85]</sup>，测量环境光照的系统<sup>[46]</sup>。此类系统的基本原理是：通过控制物体所处环境的光照特性和摄像机相对于物体的位置，测量物体表面的光照特性。这方面的主要研究内容是测量系统的校准问题和如何将测量的数据运用于真实感渲染。

三维自动建模技术的主要研究内容包括：三维数据的采集方式、曲面重构、模型纹理坐标的获取和不同方位纹理图像的拼接，以及物体表面光照模型的获取和真实感渲染。下面我们将分别对这几个方面进行分析和讨论：

### 1.2.1 几何信息的获取方式

自动建模技术首先要获取物体表面的三维信息。不同的三维信息采集方法对整个建模过程有较大的影响。总的说来有以下几种主要方式：

#### 1 接触法

这种技术是较早发展起来的技术，实现方式简单直观。这种方法用可以精确定位的探针去逐点接触物体的表面，测得被接触点的空间坐标。探针在物体表面扫描一遍，即可得到物体表面各点的坐标。传统的三维坐标测量机就是基于这一原理。探针可以用手工方式操作，但工作量大，花费时间长。也可以用机械方式进行辅助操作，如 FARO，Immersion 等公司将探针的伺服机构改为可以精确定位的随动式机械臂，由人牵引着装有机头的机械臂在物体表面滑动扫描，这个系统的装置简单，速度和灵活性得到很大的提高。其代表产品有 Faro 公司<sup>[76]</sup>、FOMER CimCore 公司<sup>[77]</sup>、Immersion 公司<sup>[78]</sup>的产品。图 1.1 展示了这些产品的外观图。



图 1.1 采用接触法测量的系统

总的来说，这种方法测量精度高，但装置复杂，价格昂贵，测量速度慢。测量过程由于是使用压力原理，有可能破坏表面，不能用于珍贵物体；而且测量物体不易太大。它主要运用于工业领域。

## 2 遥测法

这类方法由测距器主动向被测物体表面发射探测信号，信号遇到物体表面反射回来后，通过测量信号的传播时间或相位变化，可以推算出信号传播距离，从而得到物体表面的空间位置信息。通常用激光或超声波做为探测脉冲。基于这一原理的激光干涉仪，精度从原理上最高可以达到光波长量级。这方面技术成熟，有商品化的产品。Leica 公司<sup>[86]</sup>，Acuity 公司<sup>[87]</sup>推出了手持式和机械式的激光测距仪器，精度在毫米级；Senix<sup>[88]</sup>公司则开发了超声测距仪。



图 1.2 采用激光测距仪进行三维建模的系统

若是把激光测距仪器安装到位姿可控的平台上，并且用计算机进行自动控制，可以构成三维自动建模系统。这类系统的特点是不涉及图像处理问题，且受遮挡的影响小，测量精度高。但是扫描速度较慢，设备价格昂贵，需要配合其它设备获取物体表面的纹理信息和光照反射特性。它比较适合于专业的模型生产厂家。这方面有代表性的有 Cyberware 公司<sup>[80]</sup>、3rdTech 公司<sup>[81]</sup>、Cyra 公司<sup>[89]</sup>的产品。图 1.2 展示了这些产品的外观。

## 3 图像分析法

这类方法是以图像为获取信息方式，通过对图像的分析 and 处理得到相应

的三维空间信息。这方面还处于研究阶段，缺乏成熟的产品。但因为其数据采集部分可以采用标准的图像采集设备,如工控摄像头、数码相机、摄像机等，系统的价格低廉。并且系统比较灵活，可以同时采集物体的颜色纹理信息。因此这几年这方面的研究十分活跃，有许多的研究成果。图像分析法主要分为主动式和被动式两大类。

### (1) 主动式

主动式又称为结构光方法，是一种既利用图像又利用可控光源的测距技术。光源的几何形状有点状、条状、网格状等许多种。光源可选用可见激光光源，投影仪光源等。一般来说运用结构光有如下几种不同的方式。

一种是利用光源的几何形状信息辅助提取景物中的几何信息<sup>[4,5]</sup>。例如，利用条状光源照射在物体表面产生光条纹，根据这些光条纹的形态和间断性确定物体表面的形状信息。整个系统经过定标后，可以计算出被光照射处的三维坐标。配合机械扫描运动，可以获得物体表面各点的坐标。文献<sup>[82]</sup>是这方面的最新的研究成果，它控制投影仪在物体表面产生随时间变化的条状阴影，然后用摄像机摄取的视频计算得到物体的几何结构。

另一种是利用结构光作为标记的方式。在单目视觉系统中，可以让光源和摄像头构成一个测量系统，通过光标在图像中出现的位置来确定物体的三维信息<sup>[6,7]</sup>。在双目视差系统中，利用结构光作为左右两幅图像的对应点的标志。在文献<sup>[1]</sup>中，利用了一个激光点光源在物体上的成像作为对应点标志。

主动式的方法可以有效的控制物体的光照环境，在一定程度上控制输入图像的表现形式，从而可以减少计算的复杂性。这类方法扫描速度快，测量精度高，算法表现比较稳定，适用于物体表面对光线反射情况比较好的场合。

### (2) 被动式

这类方法不对物体的环境光照进行刻意的控制，而仅根据一般光照条件下得到的图像获取三维信息。这类方法对场景和物体的约束较小，而且硬件需求也较小，但是总的来说算法缺乏稳定性，还处于研究阶段。

当光线照在物体表面时，根据物体表面的三维结构，会造成特殊的阴影

效果。通过对这些阴影的分析，可以得到物体表面的三维结构信息<sup>[2]</sup>（Shape From Shading）。在文献<sup>[3]</sup>中，通过对真实材料表面在不同光照下的阴影进行分析，能得到这个表面上的微小几何结构。

同时光线对物体作用时，会在物体的边缘产生轮廓线。可以通过对轮廓线的检测和计算得到物体的几何信息。文献<sup>[83]</sup>中就采用了这样的方法。

运用双目立体视差技术时，被动式方法最困难的问题在于如何确定同一个三维空间点在不同视角图像中的对应点。在确定对应点时，可以考虑物体上特殊的纹理和其它约束，如双摄像机的外极线约束。在文献<sup>[9]</sup>中，物体在旋转转台上做圆周运动，算法通过跟踪特征点在不同时刻所拍摄的图像中的运动位置来决定对应点。而在文献<sup>[9]</sup>中，先计算出候选特征点，然后运用鲁棒性计算和外极线约束来决定特征点之间的对应关系，并且同时进行物体的运动参数估计。

### 1.2.2 曲面重构

曲面重构是研究如何确定空间的三维点之间的拓扑关系，并通过此拓扑关系得到物体的面片模型。三维数据点的采集是曲面重构的基础。曲面重构方法与三维数据点采集的顺序、密度有着很大的关系。一般来说采集的三维数据点集有两种类型：有结构关系的和无结构关系的。有结构关系的是指：三维数据点之间已经由采样顺序或其它约束建立了一定的拓扑结构。此时进行曲面重构算法较为简单，可以用能量最小等约束建立起点之间的完整的拓扑结构<sup>[10]</sup>。反之，无结构关系指三维数据点之间没有任何的拓扑信息，此时整个重构算法较为复杂<sup>[73]</sup>。

曲面重构算法主要基于如下几类思路：

#### (1) 基于参数描述的曲面重构

参数曲面方法用参变量的函数显式地描述曲面点集的坐标值<sup>[11]</sup>。这类方法主要通过拟合，求出最能够接近表面的参数曲面的参数。比较典型的算法有<sup>[12,13]</sup>。

#### (2) 基于隐含曲面的重构方法

隐含曲面重构的基本原理是：利用基于参数描述的势能场标量函数的等

值面来描述一个曲面。势能场函数由若干个支撑点(Support Points)对应的子势能场函数加权组成，一般而言，每个子势能函数是一个与到支撑点的距离成相反关系的函数，如负指数幂函数<sup>[14,15]</sup>。

### (3) 基于计算几何的三角剖分的方法

这种构造方法用计算几何的方法，运用约束构造三维点之间的拓扑结构，对这些三维点进行三角划分<sup>[16]</sup>。

## 1.2.3 纹理映射和纹理坐标的获取

在虚拟环境中，三维物体感渲的真实性是给用户带来沉浸感的关键因素。真实感渲染在虚拟现实的应用中受两个因素的制约：一个因素是系统的渲染速度，若渲染速度慢，则系统无法实时渲染较复杂的三维模型，真实感就无从谈起；另一个因素是三维模型本身的逼真程度，这主要是由模型的纹理决定的。一般来说，物体大的特征需要用几何模型来表达，而对于物体的细节，则需要通过纹理和纹理映射来表示<sup>[20]</sup>。

纹理映射技术最早是由 Catmull<sup>[17]</sup>提出的，他确定了以  $(u,v)$  表示的双变量实数空间（纹理空间）和以参数  $(s,t)$  表示的三维曲面之间的对应关系（映射关系）。但是在 Catmull 的算法中，被映射的值仅限制为纹理图案的彩色值。为此，Blinn 和 Newell<sup>[18]</sup>在光照模型中考虑了纹理图案的镜面反射系数和高光值，从而大大改进了经纹理映射后的图像质量。Blinn 在<sup>[19]</sup>中又提出了另一种纹理映射技术：凹凸映射，利用一个扰动函数扰动物体的表面法向量来模拟粗糙表面的纹理，实现了光滑的物体表面附上粗糙纹理的方法。在随后的发展中，纹理函数可以控制物体的透明性和其它性质。同时，也产生了能表示头发，云之类的过程纹理。

纹理映射的过程是将纹理图像映射到一个三维模型，然后在渲染阶段再映射到二维的显示设备。它涉及到三个坐标系，纹理空间坐标系  $(u,v)$ 、三维模型空间坐标系  $(x,y,z)$  以及屏幕空间坐标系  $(i,j)$ 。三个坐标的转化过程如图 1.3 所示。从纹理空间到三维模型空间的转化过程为曲面参数化，也称为纹理坐标的获取。从三维模型空间到屏幕空间的转化过程称为基于视角的投影。这个过程属于渲染过程，在本文中不予考虑。在真实感反向建模的

过程中，在纹理映射方面主要考虑两个问题：一个是纹理坐标的获取，另一个是纹理图像的获取和拼接。

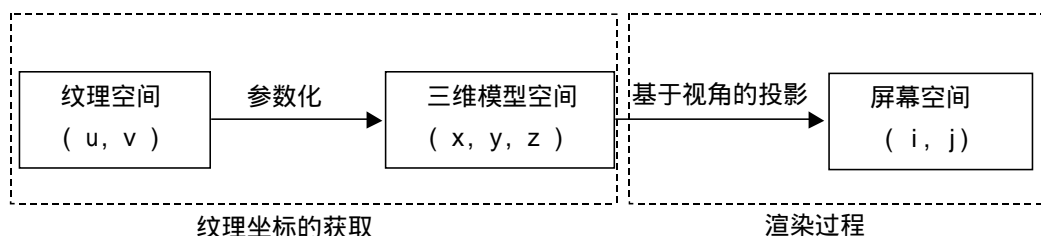


图 1.3 纹理映射过程

### 1 纹理坐标的获取

曲面参数化一直是图形学界的一大难题。因为这个过程决定了从二维平面到三维曲面的映射方式，如何解决在参数化中纹理坐标的局部拉伸和变形一直是研究的热点问题。一般来说，有如下的映射方法：

#### (1) 参数曲面的纹理映射

由于参数曲面定义了一个二维参数空间到三维空间的映射关系。因此，若将参数空间和纹理空间等同起来，参数化等价于参数曲面自身定义的逆映射。对于简单的二次曲面来说，其纹理映射函数可以解析的表达出来。而对于复杂的高次参数曲面来说，其逆映射一般无法解析表达，需采用数值求解技术来离散求得。在这方面比较有名的是 Catmull<sup>[21]</sup>的正向纹理映射技术和 Blinn<sup>[22]</sup>的反向纹理映射技术。

#### (2) 两步法纹理映射技术

对于由多边形面片组成的三维曲面，则情况要复杂得多，一般需要手工辅助来完成参数化过程。Bier 和 Sloan<sup>[23]</sup>提出了独立于物体表面的纹理映射技术，它将纹理空间到三维模型空间的映射分解为两个简单映射的复合，从而避免了物体表面的复杂的参数化过程。

这个算法的基本思想是引入一个包围物体的中介三维曲面作为中间映射媒介。其基本的过程由两个映射步骤构成：

- 1) 将二维纹理空间映射到一个简单的中介三维曲面；

2) 将上述中介三维曲面的纹理映射到三维模型的表面。

这样纹理空间到三维模型空间的纹理映射可以由上面的两步复合而成。这里面的关键问题是如何选取三维中介曲面以及确定三维中介曲面和三维模型表面的映射关系。一般来说，三维中介曲面可以选取简单的曲面，如平面、圆柱面、立方体表面和球面。本文亦采用了两步纹理映射的思想来对模型进行参数化。

## 2 纹理的获取和拼接

在真实感反向建模方面，纹理图像的获取也是一个重要的研究内容。这个问题在近几年基于图形的建模和渲染（IBMR）的研究中得到了广泛的探索<sup>[29]</sup>。在这方面，有基本不依赖于几何结构的基于光场的方法<sup>[24,25]</sup>，也有依附于一定几何信息的<sup>[26,27]</sup>。和本论文相关的有如下的技术：

### (1) 全景拼图

这也称为图像马赛克技术。当图像中的场景处于同一平面上，或者摄像机只有纯旋转运动，则两幅图像中的对应点之间满足共线或齐次约束。利用这个关系，可以将若干图像进行无缝拼接，形成环境映射全景图。图像的获取可以用专门的硬件拍摄设备，也可以对一般的拍摄设备得到的图像进行对准、融合等处理后得到。这方面，比较有名的系统是苹果公司的 Quicktime VR<sup>[28]</sup>。

全景拼图对周围的环境有一定的假设，比如假设被一个圆柱或一个球所包围。进行无缝拼接实际上就是把若干不同视角的图片参数化到这个假想的三维曲面上。本文的纹理获取也用了类似的思想。

### (2) 基于视角的纹理（view dependent texture）

物体表面存在丰富的几何结构和不同的光照反射特性，当物体在不同的视角被观测时，同一表面呈现不同的颜色。此时，用单一的纹理无法达到真实感渲染的要求。因此出现了基于视角的纹理，它的原理是同一个三维物体对应于多个纹理，当观察的角度不同时，采用不同的纹理进行渲染。

这种技术的主要研究内容是如何合理利用不同视角下的纹理进行渲染，以及如何对大量的纹理信息进行表述和压缩。在文献<sup>[26,27]</sup>，探讨了在大场

景下如何利用基于视角的纹理进行真实感渲染。最近出现的表面光流场 ( surface lightfield ) 技术<sup>[30,31]</sup>也采用了类似的思想。

### 3 纹理合成

纹理合成(texture synthesis)是另外一种获取和制作纹理的方式。它的目标是：对一个样本纹理，它能生成一个新的任意大小的纹理，新生成的纹理从肉眼看来不是样本纹理的简单重复，但是和样本纹理属于同一类纹理<sup>[52,53]</sup>。一般来说样本纹理需要满足两个性质：样本纹理是平稳的，样本纹理中的任何两个局部区域都有相似的特性；样本纹理是局部的，纹理中任何一个元素的值都是可以由其周围局部的值确定的。从数学上说，若是把图像看作二维的随机过程，则纹理样本满足马尔可夫随机场的定义，也就是说纹理样本的随机过程是局部、平稳的<sup>[54]</sup>。

纹理合成一般分为两大类方法，一类是参数化方法，这类方法是先根据样本纹理确定统计模型和参数，然后根据统计模型生成新的纹理<sup>[55,56]</sup>。这类方法计算复杂度高，而且生成的纹理的质量不好。

另外一类方法是非参数化的方法<sup>[54]</sup>，这类方法不需要确定统计模型和参数，算法简单，容易实现，而且生成的纹理具有较高的质量。这类算法的基本原理是：由于样本纹理满足局部性质，因此在生成新的纹理时，每一点的值可以由其周围局部区域的值确定。具体操作时，将待确定点邻域内的点的颜色值排列成一个高维向量 $\vec{s}$ ，然后在样本中根据相同邻域大小确定一个候选的高维向量集 $M$ ，然后从 $M$ 中选取一个和 $\vec{s}$ 距离最近的向量 $\vec{t}$ ，与 $\vec{t}$ 相对应的样本元素的颜色就是待定点的颜色值。这类方法中，最费时间的操作是从 $M$ 中选取一个和 $\vec{s}$ 距离最近的向量 $\vec{t}$ 的过程。在文章<sup>[54]</sup>中用多尺度金字塔算法来加速纹理生成的过程和提高纹理生成的质量。

纹理生成可以直接在三维模型上进行<sup>[57,58,71]</sup>。它和基于图像的区别在于：图像是以像素作为基本的运算单位，而三维模型是以模型的顶点作为基本运算单位。三维模型的顶点的分布需要致密而且均匀。通过曲面的纹理生成方法，我们可以同时得到一个模型的纹理和相对应的纹理坐标。图 1.4 说明了通过一个样本纹理生成一个新纹理图像，以及直接在模型表面生成纹

理。

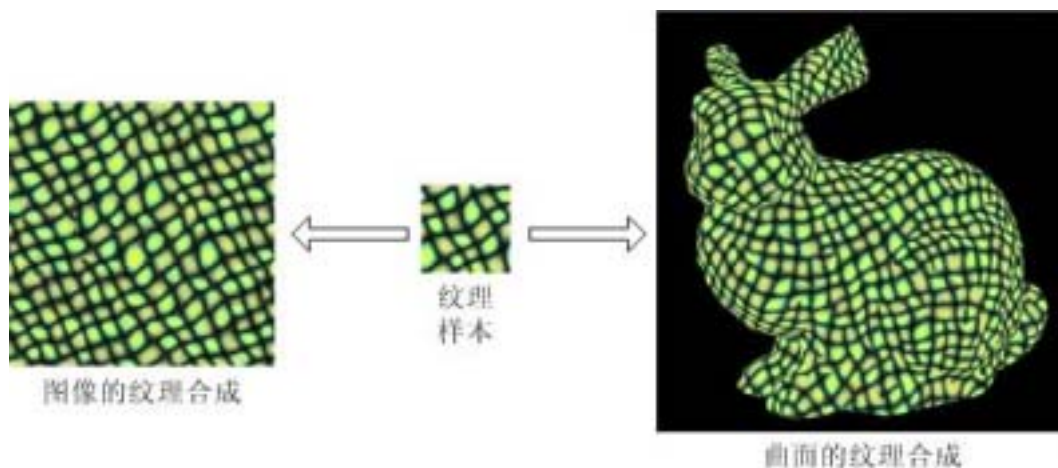


图 1.4 基于样本纹理的纹理合成

#### 1.2.4 物体表面光照模型和真实感渲染

物体表面光照模型是生成真实感图形的基础。光照模型是根据光学的有关定律，计算景物表面上任一点投向观察者眼中的光亮度的大小和色彩组成的公式<sup>[59]</sup>。光照模型一般分为局部光照模型和整体光照模型。局部光照模型仅考虑光源直接照射在物体表面所产生的光照效果，物体表面通常被假定为不透明，且具有均匀的反射率。而整体光照模型还考虑周围环境对物体表面的影响。

##### 1 若干光照模型的介绍

有一些简单，但是实用的局部光照模型。如 Lambert 漫反射模型，phong 模型等。这些模型只用很少的几个系数就可以表达多种不同材质的物体表面。比较复杂的光照模型有 Blinn 模型<sup>[60]</sup>和 Cook-Torrance 模型<sup>[61]</sup>。这些模型有较多的参数，需要通过一些测量方法确定。这类光照模型的共同特点是：用一个确定的数学表达公式来决定入射光和反射光之间的关系。但是用公式表达局部光照模型能力有限，对于一些特殊的表面，用这些光照模型无法表达。

从另一个角度来说，物体表面的反射特性可以统一用一个函数来描述，

这个函数确定了物体表面一个区域对任意方向入射光所产生的任何方向的反射光的强度。我们称这个函数为双向反射分布函数(BRDF)。可以看出 BRDF 是一个四维的函数，其中两维是入射光的方位角和极角，另外两维是出射光的方位角和极角。BRDF 有两种表示方法，一种方法是用一个四维的数值函数表示<sup>[62]</sup>，另外一种方法是用一个四维的表进行离散表达。BRDF 可以通过测量得到<sup>[45]</sup>，渲染时可以用测量得到的 BRDF 来计算反射光的强度。BRDF 的测量可以认为是物体表面光照参数的获取过程。在文献<sup>[47]</sup>中，对 BRDF 进行了推广，考虑了光在半透明物体中的发散，提出了双向次表面漫反射分布函数(BSSRDF)，并且提出了测量这个函数的方法。

## 2 BTF模型

真实世界中，物体外观在不同的光照条件下具有非常丰富的变化，同一表面从不同的角度观察也有不同的表象。这是因为物体表面不同区域对光线的反射特性不一样，同时在物体表面有非常丰富的微小几何结构(mesostructure)，这些表面的微小几何元素在光线的照射下会产生阴影(self-shadow)，从不同的视角观察会产生不同的自遮挡(self-occlusion)，光线会在这些微小的结构表面反射多次(inter-reflection)。所有的这些因素，造成了现实世界物体丰富的外观。

用传统的面片、纹理和光照模型来表示物体表面的这些特性是非常困难的。从建模的角度来说，由于物体表面具有复杂的微观几何结构，用三角面片非常难以表示，同时表面多变的光照特性也非常难以获得。从渲染的角度来说，这些微观几何结构和多变的光照特性使入射光在物体表面发生多次反射，需要模拟这些特性需要使用光线跟踪算法，这大大增加了渲染算法的复杂性。正是这些问题，使虚拟环境无法达到逼真的效果。

为了解决这个问题，Dana 中提出了双向纹理函数(BTF, Bidirectional Texture Function)的概念<sup>[63]</sup>。BTF 是一个 6 维函数，它描述了物体表面一个点对任意方向入射光所产生的任何方向的反射光的强度。它的定义形式为： $f_{BTF}(x, y, \theta_i, \phi_i, \theta_r, \phi_r)$ 。其中  $(x, y)$  是物体表面一个点的位置坐标， $(\theta_i, \phi_i)$  是入射光的方位角和极角， $(\theta_r, \phi_r)$  是反射光的方位角和极角。可以看出 BTF

是 BRDF 的一个简单的推广，它考虑了物体表面的光照性质和微小几何结构随位置变化而变化的特性。一个材料的 BTF 样本可以用图像序列进行表示，每幅图像表示这个 BTF 样本在一个特定的光照方向下，照相机从某个特定的角度拍摄到的图像。其中光照的强度和照相机的曝光程度都经过了标定。需要大量的图片才能表达一个 BTF，一般来说完整表达一个样本 BTF 需要几千幅图片和几百兆的存储空间。图 1.5 展示了一个硬塑料的 BTF 样本序列中的几幅图像。可以看出在不同方向的光照和视角条件下，同一表面呈现出完全不同的外观。

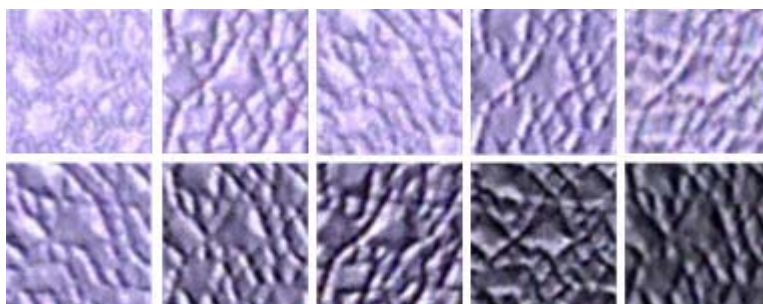


图 1.5 硬塑料的 BTF 样本序列中的几幅图像

BTF 可以通过测量的方式得到，文献<sup>[63]</sup>设计了一个装置可以得到一个样本材料的 BTF 的稀疏采样（205 个采样点）。文献<sup>[3]</sup>提出了一个根据 BTF 稀疏采样得到 BTF 致密采样的方法。

用 BTF 可以描述真实世界物体表面在各种光照条件下复杂的变化。在计算机视觉和图像领域都对它有广泛的研究。在计算机视觉领域，主要从分析的角度入手，研究的内容包括：根据 BTF 建立数学模型描述物体表面<sup>[64, 65]</sup>，根据 BTF 识别物体表面属于什么材料<sup>[66]</sup>、以及对 BTF 的数据进行压缩<sup>[67]</sup>。

在计算机图形领域，主要考虑的是如何运用 BTF 进行真实感渲染，在文献<sup>[63]</sup>中提出了把 BTF 当做一个高维的纹理函数，然后采用纹理映射的方法映射到模型的表面。但是采用一般的纹理映射方法将 BTF 映射到任意曲面会导致 BTF 的扭曲和变形，同时会导致在接缝处 BTF 所表示的微小几何结构不连续。本文在这个方面提出了一个新的解决方法。

## 1.3 本文的主要贡献与章节安排

### 1.3.1 本文的主要贡献

本文研究了基于采样数据集的真实感建模技术，主要贡献与创新之处为：

1. 设计和实现了一个三维物体快速反向建模原型系统。这个系统运行于 PC/Windows 平台，可以方便地获取三维实物的几何信息和纹理信息，所得结果可以直接利用一般的渲染工具进行渲染。

围绕该建模系统，探讨了如下的内容：

- (1) 对面向虚拟现实应用的反向建模系统进行了需求分析，并进行了系统设计。强调了对于虚拟现实应用，需要同时考虑模型的几何特征和表面纹理特征，才能达到真实感渲染的目的。
- (2) 设计了一个快速有效的系统标定过程，无需高精度标定设备的辅助，就可以方便地实现系统的标定。这方面的主要创新是：根据系统的需求和特点设计了一个标定过程；根据这个标定过程对摄像机标定算法<sup>[37]</sup>进行了推广，实现了摄像机内、外参数的标定，并且提出了标定平台转动轴的算法。
- (3) 实现了三维几何信息的采集和曲面重构算法。提出了从图像中提取结构光的步骤；采用了双目视觉原理获取物体表面的几何信息，并用外极线约束的方法确定左右两摄像机图像中特征点的对应关系；同时在曲面重构中充分考虑了空间点已经具有的拓扑关系，采用了有效的曲面重构算法。
- (4) 结合本系统的特点，借鉴全景拼图技术，实现了一个对不同视角下得到的颜色纹理进行拼接和融合的算法。

2. 提出了一种将样本 BTF 合成到任意曲面上的方法。利用该法，我们可以在任意模型表面上覆盖上和样本 BTF 光照性质相似的 BTF，增加模型的真实感。同时不会造成模型表面 BTF 的扭曲和接缝处的不连续。这个方法的基本思想是基于样本的建模，即从现实世界中获取大量有用的样本，然后用这些样本进行有目的的组合，它比基于图像的建模有了进一步的推广。

### 1.3.2 本文的章节安排

本文的章节安排如下：第一章阐述了在虚拟现实应用中从真实世界中获取信息，建立模型并进行渲染的背景、目的和意义，同时对相关的国内外研究现状和发展趋势进行了讨论；第二章对反向建模系统进行了需求分析，并进行了系统设计；第三章讨论了系统所涉及的各种坐标系，以及如何标定这些坐标系之间的相互关系；第四章给出了获取物体三维几何模型和颜色纹理的实现方法；第五章提出了在任意曲面上的 BTf 合成算法；第六章对全文的工作进行了总结，并且提出了进一步的工作设想。

## 第二章 三维物体反向建模原型系统的总体设计

随着计算机图形学和计算机视觉技术的发展，以及软件技术和硬件技术的日趋成熟，虚拟现实技术在我们生活的各方面得到了广泛的应用。为了满足视觉上的沉浸感要求，虚拟现实技术对真实物体的建模有很大的需求。为此我们将进行需求分析，并在此基础上完成对系统的设计，最后将给出系统的操作流程。

### 2.1 系统的总体设计目标

反向建模技术是利用各种传感器获得真实物体的几何和色彩信息。它在不同的领域得到了广泛的研究，特别是在工业设计领域。针对不同的应用，反向建模系统采用了不同的技术。我们将探讨在虚拟现实应用方面的需求和确定系统所要实现的功能。

#### 2.1.1 系统设计需求分析

在虚拟现实的应用中，对三维模型有其特殊的要求。总的来说有如下几点：

- (1) 模型需要具有逼真的纹理来增加物体的真实感。物体表面存在丰富的纹理细节，人们正是依赖这些纹理细节来区别各种相同形状的物体。因此，物体表面纹理细节在系统真实感渲染方面具有非常重要的作用。
- (2) 为了保证系统的实时渲染速度，构成物体模型的面片数应该尽量少。这个需求是虚拟现实应用所特有的，在实时渲染系统中，系统渲染流畅程度决定了这个系统的总体表现。同时我们也注意到，人的视觉只对几何结构的形状敏感，而对其精确性是不敏感的。因此在以渲染为目的的系统中，可以适当的减小几何模型的精确度。
- (3) 需要简单有效并且低成本的建模方法。因为虚拟现实的广泛应用前景，要有低成本的建模系统来促进应用的普及性。同时在一个虚拟环境中需要构造大量的虚拟物体，因此建模系统应该操作简单，快

速有效。

### 2.1.2 系统所应具有的特点

根据上面的需求分析，我们的原型系统有如下的特点：

- (1) 对于所构造的模型，用较少的面片表示物体的基本几何信息，用颜色纹理表示物体的局部细节信息。这里，我们采用基于图像的建模技术来达到这个目的。
- (2) 快速和简便的操作过程。这主要表现在：整个标定过程只需要简单标定设备的辅助，而且可以在数分钟内完成标定。标定完成后，系统自动获取模型的几何信息和纹理信息。
- (3) 较低廉的硬件成本。系统运用了立体视差法来获得物体的三维几何信息，采用了一般工控摄像头作为感知设备。

## 2.2 系统的配置

### 2.2.1 硬件配置

系统的硬件实物图如图 2.1 所示。该原型系统由旋转平台，数据采集设备和 PC 机等构成。旋转平台采用了大连贤科机器人技术有限公司生产的华录四自由度平面直角机器人。它有四个运动自由度：三个空间运动自由度和一个平台旋转自由度。计算机可以通过串口和此机器人通讯，完成对机器人位姿的控制。在系统实际运行时，平台空间位置固定，只控制旋转分量。因此，本系统可以用只有旋转运动的控制平台。

数据采集设备使用了两个松下公司的 WV-CP230 摄像头和 Matrix 公司的 Meteor\_II 图像采集卡。Meteor\_II 采集卡可以接入多路采集信号，使用选通的方式进行工作。摄像头放在三角架上，可随意调整两摄像头的位置。为了能较准确的测量，两摄像头相对物体的夹角应该大一些（一般 30 度左右），同时摄像头距离待测物体不要太远（一般 50 - 70 厘米），保证图像上的物体影像尽量的大一些。

结构光光源是一个自制的激光发生器。由二极管激光器发出的点状光打

到细玻璃棒发散成线状光，从而构成一个发出红色线状光的结构光光源。结构光源也是放在三脚架上，一般是放在两摄像机之间。

另外还有一些手工制作的辅助标定设备，包括一个用于标定摄像机参数的标定板和一个用于标定机器人坐标系和世界坐标关系的标定物。测量对象是模特人头。

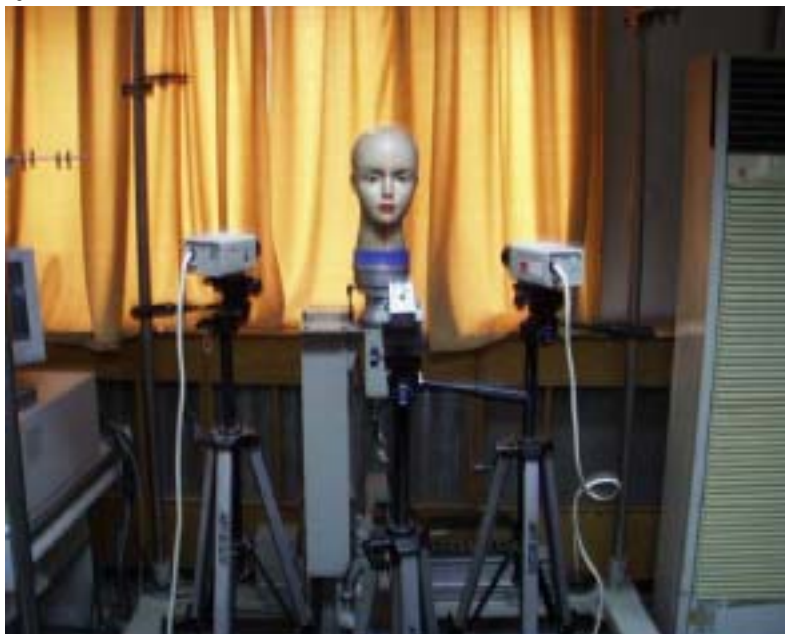


图 2.1 三维物体自动建模原型系统

### 2.2.2 软件配置

采用了 win2000 作为操作系统，使用 VC 6.0 作为主要开发工具。使用 Intel 的 Image Process Library (IPL) 作为图像处理的开发包；Intel 的 Open Source Computer Vision Library (OpenCV) 作为视差测距方面的开发包；同时使用了 KitWare 公司的 Visualization Toolkit(VTK)作为构造几何模型方面的开发包。

## 2.3 系统的结构和 workflows

根据系统的总体结构、功能描述与 workflows，我们将系统划分成若干的子系统 and 子过程，这样便于系统的实现和维护。

### 2.3.1 系统总体结构图和子系统划分

根据所要实现的功能，可以将系统分为如下三个子系统，其构成和关系如图 2.2。

#### 1 系统标定子系统

系统标定子系统的功能主要是确定各坐标系之间的变换关系，从而为后续的测量做准备。标定主要包括两摄像机内参数的标定，摄像机坐标系到世界坐标系的标定，以及机器人坐标系到世界坐标系的标定。

#### 2 几何模型获取和构造子系统

此系统的功能是获取空间点的三维信息并且进行曲面重构。包括的子过程是：结构光轮廓提取，双目特征点的对应关系和空间坐标的确定，曲面重构以及曲面平滑。

#### 3 纹理和纹理坐标获取子系统

此系统的功能是获取物体的纹理和纹理坐标。包括的子过程是：纹理图像校正，拼接和融合，以及纹理坐标的计算。

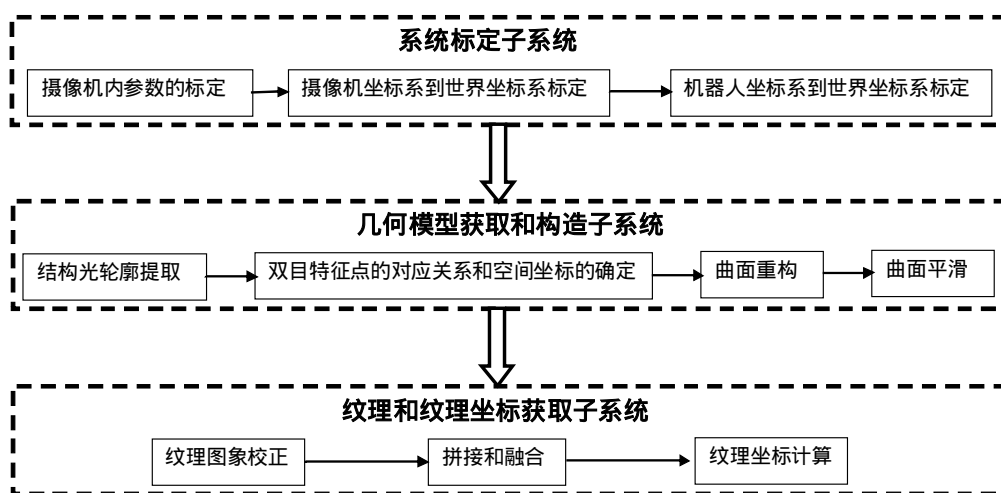


图 2.2 系统结构图

### 2.3.2 系统工作流程

系统的工作和操作的流程如下：

1．在标定阶段，将机器人转台、摄像机和光源置于合适位置。人手持标定板分别在左右摄像机前适当位置移动与旋转标定板，各拍摄大约 20 幅标定板不同位姿的图片，进行摄像机内参数的标定。然后两摄像机同时拍摄标定板同一位姿的图像，大约拍摄 20 组左右，进行摄像机坐标系到世界坐标系的标定。最后将一个标志物固定在转台上，转台旋转一周，完成机器人坐标系到世界坐标系的标定。整个标定过程（包括手工操作和计算过程）大约 5 分钟即可完成。

2．在几何模型获取和重构阶段，将对象模型置于转台上，打开结构光光源，转台旋转一周，进行空间点三维信息的获取以及曲面重构。然后手工调整参数，用曲面平滑算法对模型进行适当平滑。

3．纹理和纹理坐标获取阶段，关闭结构光光源，转台旋转一周，完成纹理和纹理坐标获取。

## 第三章 系统标定算法的研究

### 3.1 引言

计算机立体视觉研究的问题是从摄像机获取的两幅或多幅图像出发，计算三维空间中物体的几何信息，从而重构和识别物体。而三维空间某点的位置与其在图像中对应点之间的相互关系是由摄像机成像的几何模型参数和摄像机所处的位姿参数决定的。在大多数条件下这些参数必须通过实验和计算得到，这个过程称为标定过程。

对于计算机视觉系统，不同的系统有不同的精度要求，所要标定的参数也不尽相同，因此采用的标定的方式和过程也有所相同。一般来说，若需要较高的测量精度，则需采用较复杂的成像模型，并且在标定过程中需要高精度的辅助标定参照物。这样会导致标定的过程复杂、所需要的辅助标定设备价格昂贵。如在文献<sup>[32]</sup>中，为了能精确的跟踪牙科手机的位置，采用了空间坐标测量仪作为标定的辅助工具，这样导致标定的操作过程复杂，耗费之间长，而且标定操作只有专业人员才能胜任。同时系统一经标定，摄像机的相互位置难以调整，这大大限制了系统的灵活性。

因为我们的原型系统潜在的应用范围广，所针对的用户是非专业人员，因此需要标定过程简单快捷、成本低廉。同时，在虚拟现实环境中，几何模型只用于表示物体的基本结构，而不表示物体表面的细节。因此在建模过程中，我们只需要获得物体的基本几何信息，对精度的要求不高。基于这些分析，我们将摒弃昂贵复杂的标定辅助设备，而只采用一些自制的简单标定工具。同时使标定的流程尽量简单化。

### 3.2 坐标系的定义

在反向建模系统中涉及到多个坐标系，包括图像坐标系，摄像机坐标系，世界坐标系等。标定实际上就是确定这些坐标系之间的变换关系，因此我们首先需要明确给出系统涉及的各种坐标系的定义以及它们之间的相互关系。在本系统中，定义了五种不同类型的坐标系：图像坐标系、摄像机坐标

系、世界坐标系、机器人坐标系、物体坐标系。它们的定义和关系如下：

### 1 图像坐标系

图像坐标系  $O_I - UV$  是定义在摄像机所采集图像上的直角坐标系，其坐标的原点位于图像的左上角，其水平轴向右，垂直轴向下。其单位是以像素为单位，某个像素的坐标是该像素所处的列数和行数。

### 2 摄像机坐标系

摄像机坐标系  $O_C - X_C Y_C Z_C$  是定义于摄像机上的坐标系，其中  $O_C$  为摄像机的光心， $X$  轴  $Y$  轴分别和图像坐标系的  $U$  轴  $V$  轴平行。 $Z$  轴为摄像机光轴，与成像平面垂直。 $O_C - X_C Y_C Z_C$  符合右手坐标系定义。

若使用线性摄像机模型，摄像机坐标系和图像坐标系的关系如下，

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A \begin{pmatrix} X_C \\ Y_C \\ Z_C \end{pmatrix} \quad (3.1)$$

其中，图像坐标系采用了齐次坐标， $s$  为一个标量， $u$ 、 $v$  为图像坐标系的值， $X_C$ 、 $Y_C$  和  $Z_C$  为摄像机坐标系的坐标。 $A$  是一个  $3 \times 3$  的矩阵，称为摄像机的内参数矩阵。

### 3 世界坐标系

同时，我们还需要一个坐标系来描述所有物体在空间的位置，这个坐标系称为世界坐标系  $O_W - X_W Y_W Z_W$ 。可以认为这个坐标系是一个基准的绝对坐标系。摄像机所拍摄的内容还取决于摄像机在世界坐标系中的位姿，这个位姿参数就是摄像机的外参数。

因为在本系统标定过程中所采用的标定参照物没有世界坐标值，因此世界坐标系的定义和其它系统有很大的不同。我们定义世界坐标系和左摄像机的摄像机坐标系相重合。这样带来的好处是：一旦两摄像机之间的相互位置关系确定，我们就完成了摄像机外参数的标定。

从世界坐标系到右摄像机坐标系的关系如下：

$$\begin{pmatrix} X_{CR} \\ Y_{CR} \\ Z_{CR} \end{pmatrix} = R_{WCR} \begin{pmatrix} X_W \\ Y_W \\ Z_W \end{pmatrix} + t_{WCR} = R_{WCR} \begin{pmatrix} X_{CL} \\ Y_{CL} \\ Z_{CL} \end{pmatrix} + t_{WCR} \quad (3.2)$$

其中  $X_{CR}$ 、 $Y_{CR}$  和  $Z_{CR}$  为右摄像机坐标系的坐标值， $X_{CL}$ 、 $Y_{CL}$  和  $Z_{CL}$  为左摄像机坐标系的坐标值， $X_W$ 、 $Y_W$  和  $Z_W$  为世界坐标系的坐标值。 $R_{WCR}$  为旋转变换矩阵， $t_{WCR}$  为位移向量。

将公式(3.1)和(3.2)合并，可以得到世界坐标系到摄像机坐标系的变化关系：

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = P \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix} \quad (3.3)$$

其中  $P$  称为摄像机投影矩阵，且

$$P = A(R \quad t) = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{pmatrix}$$

#### 4 机器人坐标系

机器人坐标系  $O_M - X_M Y_M Z_M$  是和机器人转台相关的坐标系，它的  $Y$  轴和转台的转轴相重合，并且  $X$ 、 $Z$  轴和  $Y$  轴共同构成右手直角坐标系，其中我们对  $O_M$  的位置没有明确的要求。 $O_M - X_M Y_M Z_M$  在整个测量过程中是固定不变的。这样定义机器人坐标系的好处我们将在后面讨论。

从世界坐标系到机器人坐标系的关系如下，

$$\begin{pmatrix} X_M \\ Y_M \\ Z_M \end{pmatrix} = R_{WM} \begin{pmatrix} X_W \\ Y_W \\ Z_W \end{pmatrix} + t_{WM} \quad (3.4)$$

这里， $X_M$ 、 $Y_M$  和  $Z_M$  为机器人坐标系的坐标值， $X_W$ 、 $Y_W$  和  $Z_W$  为世界坐标系的坐标值。 $R_{WM}$  为旋转变换矩阵， $t_{WM}$  为位移向量。

## 5 物体坐标系

物体坐标系  $O_o - X_o Y_o Z_o$  是和物体相关联的坐标系，它跟随物体的运动而运动。我们定义它和机器人坐标系仅有  $Y$  轴的旋转变换关系。

设机器人的转台转动了  $\theta$  度，则物体坐标系和机器人坐标系之间的关系为

$$\begin{pmatrix} X_o \\ Y_o \\ Z_o \end{pmatrix} = R_{MO} \begin{pmatrix} X_M \\ Y_M \\ Z_M \end{pmatrix} = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \begin{pmatrix} X_M \\ Y_M \\ Z_M \end{pmatrix} \quad (3.5)$$

## 3.3 摄像机内参数的标定

摄像机内参数的标定就是要明确公式 (3.1) 中的矩阵  $A$  的值。这需要设定摄像机成像的数学模型，确定成像模型中待标定的参数，同时设计一个合适本系统要求的标定方法。

### 3.3.1 线性摄像机模型

为了避免复杂的透镜系统，我们采用针孔模型来近似成像模型。在针孔摄像机中，三维空间一个点的图像是从这个点出发的并经过针孔的光线与摄像机成像平面的交点。在理想情况下，认为从该点发出的光线是理想的直线。用针孔模型来描述一般成像系统是适合的。

在针孔模型中，摄像机内参数  $A$  的定义如下：

$$A = \begin{pmatrix} f_x & \alpha & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \quad (3.6)$$

其中  $f_x$  为  $X$  方向焦距， $f_y$  为  $Y$  方向焦距， $\alpha$  为扭曲系数，与成像栅格的方向夹角有关，当其正交时  $\alpha = 0$ 。 $(c_x, c_y)$  为主点，即摄像机光轴与成像平面的交点坐标。

实践表明，线性模型不能精确的描述成像的几何关系，在成像系统中还有非线性变换的因素存在，尤其是在使用广角镜头的时候。这些非线性变换

造成了畸变。描述非线性畸变可以用如下的公式，

$$\begin{aligned}\hat{u} &= u + \delta_{\hat{u}}(u, v) \\ \hat{v} &= v + \delta_{\hat{v}}(u, v)\end{aligned}\quad (3.7)$$

其中  $u$ 、 $v$  是理想的针孔模型投影下的值， $\hat{u}$ 、 $\hat{v}$  是实际图像下的观测值， $\delta_{\hat{u}}$  和  $\delta_{\hat{v}}$  是非线性函数。

畸变主要包括径向畸变、切向畸变和薄棱镜畸变等，其中径向畸变是最主要的畸变。我们需要根据测量精度的需求确定摄像机畸变模型。在摄像机内参数的标定过程中，畸变模型参数也要考虑在内。

### 3.3.2 标定算法分类

对于摄像机内参数的标定，根据系统需求的不同，有多种不同的标定算法<sup>[33]</sup>。根据标定过程中是否采用标定参照物，可以将标定算法分成传统标定算法和自标定算法两类。

#### 1 传统摄像机标定算法

这类算法基于一定的标定参照物。基本的原理是：基于已知形状、尺寸或空间位置的标定参照物，经过对其图像的处理，利用一系列数学变化和计算方法，求取摄像机的内参数和外参数。这类算法的优点是系统的标定精度较高，而且算法经过多年的发展，比较成熟。其缺点是标定的过程比较复杂，需要专业人员的操作。而且标定完成后，系统设置难以调整，限制了系统的灵活性。同时需要有专用的、精度较高的标定参照物，这增加了系统的复杂程度和成本。

这类算法中比较有代表性的有：

传统的摄影测量方法。比较典型的为文献<sup>[34]</sup>，它考虑了成像系统中的各种因素，精心设计了成像模型，标定能达到很高精度。但是系统的待定参数空间巨大，计算过程复杂。

直接线性变换方法(DLT)<sup>[35]</sup>。这类算法只考虑线性摄像机模型，图像坐标系和空间坐标系之间的变化关系可以直接由一组线性约束方程组求得。这类方法求解过程简单，为了提高精度，需要进一步对镜头的畸变进行补偿和

修正。

两步法<sup>[36]</sup>。它利用直接线性变换方法或者透视变换矩阵方法求解摄像机参数，再以求得的参数为初始值，考虑畸变因素，并利用最优化叠代算法进一步提高定标精度。

## 2 摄像机自标定算法

这类算法无需依赖于标定参照物，仅利用摄像机在运动过程中周围环境的图像以及图像之间的对应关系进行标定。这类算法的灵活性好，不依赖于特定的环境，标定过程方便。但是标定精度要明显差于传统的摄像机标定方法。

目前在自标定方面有如下的几类方法<sup>[33]</sup>：利用本质矩阵和基本矩阵的摄像机标定方法，利用绝对二次曲线和外极线变换性质的摄像机标定方法，利用主动系统控制摄像机做特定运动的标定方法，利用多幅图像之间的直线对应关系的摄像机标定方法。

### 3.3.3 基于多个自由平面的标定算法

考虑到系统设计的需求，我们采用了张正友提出的摄像机标定算法<sup>[37]</sup>。这种算法通过拍摄平面模板在空间中的几个不同位姿来完成标定过程。它是介于传统标定方法和自标定方法之间的一种方法，因为这种方法使用了标定参照物，但是不需要明确知道标定参照物的三维空间信息，只需要知道标定平面上标志点的二维位置信息。因为在标定过程中无需知道明确的三维信息，因此标定操作简单灵活，非专业人员易于操作；同时和完全不依赖于标定物的自标定方法相比，它的标定精度较高，鲁棒性较好。

这个方法的基本思路是：

令空间坐标系的  $X_w O Y_w$  平面为标定板所在平面， $Z_w$  垂直于该平面，则标定板平面上各点的  $Z$  均为 0，则投影方程(3.3)变为

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \cong A \begin{pmatrix} r_1 & r_2 & r_3 & t \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{pmatrix} \quad (3.8)$$

因此标定板上标志点的坐标  $(x_i, y_i)$  与其成像对应点  $(u_i, v_i)$  满足齐次变换，

$$\begin{pmatrix} u_i \\ v_i \\ 1 \end{pmatrix} \cong H \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix}, \quad H = \lambda(h_1 \quad h_2 \quad h_3) = A(r_1 \quad r_2 \quad t) \quad (3.9)$$

其中， $H$  为齐次矩阵。给定在标定板平面上的若干点以及经过  $H$  变化的对应值，可以求出  $H$  [37]。

由于  $r_1$  与  $r_2$  为旋转矩阵的两列，它们相互正交且模均为 1，则有

$$h_1^T A^{-T} A^{-1} h_2 = 0, \quad h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2 \quad (3.10)$$

根据公式(3.10)由若干个平面对应的  $H$ ，可以解得内参数矩阵  $A$ 。然后通过公式(3.9)可以得到摄像机相对于每个平板的位姿  $R$ 、 $t$ ，有如下的公式，

$$\begin{aligned} r_1 &= \lambda A^{-1} h_1 \\ r_2 &= \lambda A^{-1} h_2 \\ r_3 &= r_1 \times r_2 \\ t &= \lambda A^{-1} h_3 \end{aligned} \quad (3.11)$$

其中， $R$  并不严格遵守旋转矩阵的性质，可以采用一定方法对其进行修正 [37]。在得到  $A$ 、 $R$ 、 $t$  的初始估计后，可以计算如下的误差公式，

$$\sum_{i=1}^n \sum_{j=1}^m \|m_{ij} - \tilde{m}(A, R_i, t_i, M_{ij})\| \quad (3.12)$$

其中， $n$  是图像的总数目， $m$  是每幅图像中的标志点的个数， $m_{ij}$  和  $M_{ij}$  分别是第  $i$  幅图像的第  $j$  个点的图像坐标和在标定板上的二维位置坐标。 $R_i$ 、 $t_i$  是摄像机相对于第  $i$  个图像所对应的位姿。 $\tilde{m}$  是从标定板上的二维位置坐标到图像坐标的变换过程。

然后通过一个非线性优化过程(Levenberg-Marquardt 算法 [39])，优化  $\tilde{m}$  中的参数的值，使公式(3.12)的误差值最小化，最后得到摄像机的内参数矩阵以及外参数矩阵。

若是考虑摄像机镜头畸变对系统成像的影响，我们可以引入如下的畸变模型，即

$$\begin{aligned}\hat{u} &= u + u_r[k_1 r^2 + k_2 r^4 + 2p_1 v_r + p_2(\frac{r^2}{u_r} + 2u_r)] \\ \hat{v} &= v + v_r[k_1 r^2 + k_2 r^4 + 2p_2 u_r + p_1(\frac{r^2}{v_r} + 2v_r)]\end{aligned}\quad (3.13)$$

其中， $u$ 、 $v$ 是理想的针孔模型投影下的值， $\hat{u}$ 、 $\hat{v}$ 是实际图像中的观测值。 $k_1$ 、 $k_2$ 是径向畸变参数， $p_1$ 、 $p_2$ 是切向畸变参数。 $u_r = u - c_x$ ， $v_r = v - c_y$ ， $r^2 = u_r^2 + v_r^2$ 。

此时公式 (3.12) 可以进一步变为

$$\sum_{i=1}^n \sum_{j=1}^m \|m_{ij} - \tilde{m}'(A, k_1, k_2, p_1, p_2, R_i, t_i, M_{ij})\| \quad (3.14)$$

可以设置  $k_1$ 、 $k_2$ 、 $p_1$ 、 $p_2$ 的初值为 0，然后用非线性优化过程<sup>[39]</sup>，优化  $\tilde{m}'$  中的参数的值，得到最终解。

我们在具体实现的时候，使用了 OpenCV 中张正友的算法<sup>[38]</sup>。它的标定过程如图 3.1。其中模板由一个棋盘格纹理贴到一个平整的表面（如玻璃表面）得到，其中每个格子的长度和宽度已知，精确到 0.5 毫米。操作人员手持模板在摄像机前平移与旋转，摄像机自动拍摄不同位姿的模板图像，确定图像中的标志点。当收集到足够多的图像(20 幅左右)，采用上述方法得到摄像机的内参数矩阵和畸变参数。可以利用畸变模型对以后的拍摄图像进行校正。



图 3.1 操作人员手持模板在摄像机前平移与旋转

### 3.4 摄像机坐标系到世界坐标系的标定

双目立体视觉技术是用两个摄像机获取空间三维几何信息。在计算机视觉系统中，两个摄像机在不同的位置同时拍摄下周围景物，然后计算出周围景物的三维形状和位置。

在计算三维信息前，需要标定两个摄像机在空间的位姿，也就是摄像机的外参数。根据我们前面对坐标系的定义，我们设定世界坐标系和左摄像机坐标系相重合，因此摄像机坐标系到世界坐标系的标定就是需要求出公式(3.2)中的  $R_{WCR}$ 、 $t_{WCR}$ 。

#### 3.4.1 标定算法叙述

确定  $R_{WCR}$ 、 $t_{WCR}$ ，实际上也就是要确定两摄像机之间的相互关系。存在多种立体视觉标定算法<sup>[40]</sup>。在<sup>[41]</sup>中提出了基于直接线性变换的立体视觉标定算法，在文献<sup>[32]</sup>中推广了此算法。此类算法需要知道标定物精确的空间三维坐标。另外还有一些算法无需知道空间三维点的坐标，只需确定标志点在左右成像平面的对应关系，如在<sup>[42]</sup>中，利用计算基本矩阵和本质矩阵来确定两摄像机的关系，但此类方法所得结果不精确。

本系统对前面所用的单摄像机的标定方法进行了推广，提出了一个新的标定方法。这个标定方法的基本思路是固定标定板，分别确定左右摄像机相对于标定板的关系，然后通过此关系确定左右摄像机之间的关系。

其推导过程如下：

由某标定板位置  $i$  定义的坐标系到左、右摄像机坐标系的之间的关系为：

$$\begin{aligned} M_{CL} &= R_i^L M_i + t_i^L \\ M_{CR} &= R_i^R M_i + t_i^R \end{aligned} \quad (3.15)$$

其中， $M_i$  是由标定板位置  $i$  定义的坐标系中的点， $M_{CL}$ 、 $M_{CR}$  分别是左右摄像机坐标系的坐标， $R_i^L$ 、 $t_i^L$ 、 $R_i^R$ 、 $t_i^R$  可以分别由公式(3.12)计算得到。

由公式(3.15)可以得到：

$$\begin{aligned} R_{WCR} &= R_i^R R_i^{L-1} \\ t_{WCR} &= t_i^R - R_i^R R_i^{L-1} t_i^L \end{aligned} \quad (3.16)$$

通过 3.16 可以得到  $R_{WCR}$  ,  $t_{WCR}$  的初始值 , 然后我们得到如下的误差公式 :

$$\sum_{i=1}^n \sum_{j=1}^m \left\| m_{ij}^L - \tilde{m}^L(R_i^L, t_i^L, M_{ij}) \right\| + \left\| m_{ij}^R - \tilde{m}^R(R_{WCR}, R_i^R, t_{WCR}, t_i^R, M_{ij}) \right\| \quad (3.17)$$

这里 ,  $\tilde{m}^L$  是从标定板上的二维坐标到左摄像机图像坐标的变换过程 ,  $\tilde{m}^R$  是从标定板上的二维坐标到右摄像机图像坐标的变换过程。  $m_{ij}^L$  和  $m_{ij}^R$  分别是标定板上标志点  $M_{ij}$  所对应的左右摄像机成像点。

这个公式相当于公式(3.12)的推广 , 它在摄像机内参数已知的情况下 , 同时优化标定板和左摄像机的关系 , 以及左右摄像机之间的关系。通过和公式(3.12)相似的非线性优化过程 , 可以计算得到  $R_{WCR}$  ,  $t_{WCR}$ 。

标定过程为 , 改变标定板的位姿 , 并且左右摄像机同时拍摄下标定板的图像。在得到 20 多组图像后 , 分别对左右图像用畸变模型进行矫正。然后计算  $R_i^L$ 、  $t_i^L$ 、  $R_i^R$ 、  $t_i^R$  以及  $R_{WCR}$ 、  $t_{WCR}$  的初始值 , 最后通过公式(3.17)对  $R_{WCR}$ 、  $t_{WCR}$  进行优化。在实际标定过程中 , 单摄像机标定可以采用立体视觉系统标定过程中采集的图像 , 这样省去了单摄像机标定的操作环节。

### 3.4.2 双摄像机立体视觉原理

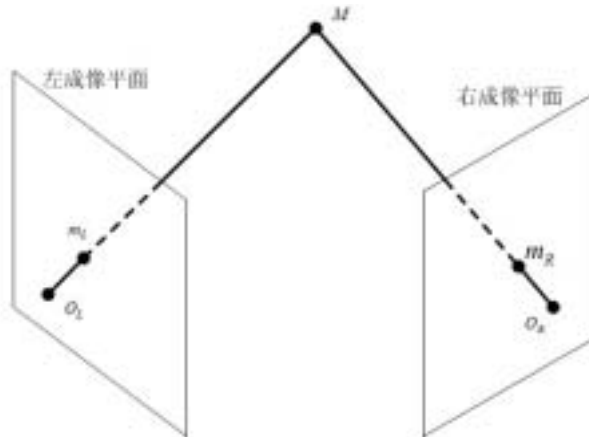


图 3.2 双摄像机成像原理图

基本原理如图 3.2。空间点  $M$  和左摄像机光心  $O_L$  构成一条直线 , 此直线在左摄像机成像平面唯一确定成像点  $m_L$ 。同时右摄像机成像平面上的像

点  $m_R$  也由  $M$  唯一确定。反之，若两摄像机位置固定，则由左右摄像机像点  $m_L$ 、 $m_R$  可以唯一确定点  $M$  的空间位置。

一个简单的想法是通过计算直线  $\overline{O_L m_L}$  和  $\overline{O_R m_R}$  的交点的方法确定  $M$ ，但是因为无法精确得到  $m_L$ 、 $m_R$ ，这种方法误差较大。较好的方法是利用投影矩阵关系计算得到  $M$ 。其推导如下：

设空间点  $M$  的坐标为  $(x, y, z)$ ，其对应的  $m_L$ 、 $m_R$  的坐标分别为  $(u_L, v_L)$ 、 $(u_R, v_R)$ 。左右摄像机的投影矩阵分别为：

$$P_L = \begin{pmatrix} P_{L1} & P_{L2} & P_{L3} & P_{L4} \\ P_{L5} & P_{L6} & P_{L7} & P_{L8} \\ P_{L9} & P_{L10} & P_{L11} & P_{L12} \end{pmatrix} \quad P_R = \begin{pmatrix} P_{R1} & P_{R2} & P_{R3} & P_{R4} \\ P_{R5} & P_{R6} & P_{R7} & P_{R8} \\ P_{R9} & P_{R10} & P_{R11} & P_{R12} \end{pmatrix}$$

根据公式(3.3)可以得到如下方程<sup>[32]</sup>：

$$CM = D \quad (3.18)$$

其中：

$$C = \begin{pmatrix} P_{L1} - P_{L9}u_L & P_{L2} - P_{L10}u_L & P_{L3} - P_{L11}u_L \\ P_{L5} - P_{L9}v_L & P_{L6} - P_{L10}v_L & P_{L7} - P_{L11}v_L \\ P_{R1} - P_{R9}u_R & P_{R2} - P_{R10}u_R & P_{R3} - P_{R11}u_R \\ P_{R5} - P_{R9}v_R & P_{R6} - P_{R10}v_R & P_{R7} - P_{R11}v_R \end{pmatrix}, D = \begin{pmatrix} u_L - P_{L4} \\ v_L - P_{L8} \\ u_R - P_{R4} \\ v_R - P_{R8} \end{pmatrix}, M = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

由最小二乘法，有：

$$M = (C^T C)^{-1} C^T D \quad (3.19)$$

### 3.5 机器人坐标系到世界坐标系的标定

在获取物体的几何信息的过程中，通过控制机器人平台旋转来测量物体各方位的结构信息。因此我们需要知道机器人平台的旋转轴在世界坐标系中的位置。也就是确定公式(3.4)中的  $R_{WM}$ 、 $t_{WM}$ ，一旦  $R_{WM}$ 、 $t_{WM}$  确定，我们就完成了标定的全过程。

#### 3.5.1 算法的基本思路

由前面的定义，机器人坐标系  $O_M - X_M Y_M Z_M$  的  $Y$  轴和转台的转轴相重合。因为在重构过程中，机器人只做旋转运动，而没有空间的位移，机器人

坐标系静止。因此通过确定机器人坐标系和世界坐标系的关系，我们可以将物体表面特征点的世界坐标系值转化到机器人坐标系，最后再转化到物体坐标系来，从而完成几何信息的测定。

同时因为我们只关心物体特征点的相对关系，对其绝对坐标值没有要求，因此系统对机器人坐标系的原点所处的位置没有要求。唯一的要求是它的 $Y$ 轴和转台的转轴相重合。因此确定转台的转轴位置是本节需要解决的问题。

机器人和摄像机的位姿不能保证绝对水平，因此不能假定其转轴和世界坐标系的 $Y$ 轴重合。同时，转轴是一个转动物体的动态特性，当物体静止时无法准确确定它的位置。因此，我们采用在转台上放置一个特征标志点，然后让转台旋转一周。通过跟踪特征点的运动轨迹(理想情况是一个圆)，确定转轴所处的位置。采用这样的标定方法的好处是操作简单，同时通过转台运动来表现转轴的位姿，能保证有较好精度和鲁棒性。

图 3.3 是从一个摄像机拍摄的特征点的图像和其运动一周的轨迹，图中黑色标志杆上的白色的小点即为特征点，红色的色块为标志点运动的轨迹。为了保证标定的准确性，标志点应该有适当的大小，太大或太小都容易导致检测标志点时的误差。同时在我们检测到图像中特征点的位置后，需要确定其中心点，作为特征点的图像坐标。在确定了特征点在左右图像的位置后，我们可以利用公式(3.19)确定特征点在世界坐标系的坐标值。下面我们讨论如何通过特征点的运动轨迹值确定转轴的位姿。



图 3.3 特征点的图像和其运动一周的轨迹

### 3.5.2 机器人坐标系到世界坐标系的标定

可以看出特征点的运动轨迹处于一个平面内，并且为一个圆。转台的转轴垂直于这个平面，并且过这个圆的圆心。因此我们只要求出这个轨迹圆所在平面的方程，以及圆的圆心，就可以完成标定过程。

设轨迹圆所在平面的方程为

$$aX_w + bY_w + cZ_w = 1 \quad (3.20)$$

利用最小二乘法，我们可以求解出  $a$ 、 $b$ 、 $c$  三个参数。

令  $\vec{n} = (a, b, c)^T / \sqrt{a^2 + b^2 + c^2}$ ，则  $\vec{n}$  为此平面的单位法向量，根据前面的定义（机器人坐标系的  $Y$  轴和转轴重合），以及公式(3.4)。我们有如下的公式，

$$R_{WM} \vec{n} = (0, 1, 0)^T \quad (3.21)$$

其中， $R_{WM}$  相当于将向量  $\vec{n}$  旋转到  $(0, 1, 0)^T$ 。可以根据四元素法计算得到  $R_{WM}$ 。

设  $\vec{N} = (n_1, n_2, n_3) = \vec{n} \times (0, 1, 0)^T$ ， $\vec{n}$  和  $(0, 1, 0)^T$  之间的夹角为  $\theta$ ， $c = \cos \theta$ ， $s = \sin \theta$ 。则有

$$R_{WM} = \begin{pmatrix} c + (n_1)^2(1-c) & n_1 n_2(1-c) - s n_3 & n_3 n_1(1-c) + s n_2 \\ n_1 n_2(1-c) + s n_3 & c + (n_2)^2(1-c) & n_3 n_2(1-c) - s n_1 \\ n_1 n_3(1-c) - s n_2 & n_2 n_3(1-c) + s n_1 & c + (n_3)^2(1-c) \end{pmatrix} \quad (3.22)$$

然后用  $R_{WM}$  对轨迹圆进行投影，则每个点的  $Y$  分量为一个恒量。对其  $X$  分量和  $Z$  分量的值进行圆拟合<sup>[43]</sup>，可以确定其圆心  $(x_o, z_o)$ 。我们可以设定机器人坐标系原点在转轴上的任意位置，因此可以认为  $t_{WM} = (-x_o, 0, -z_o)^T$ 。

## 3.6 小结

本章根据系统的需求和特点，定义了系统的各种坐标系以及相互之间的关系。详细讨论了标定流程，以及相应的算法。在摄像机标定方面，采用了<sup>[37]</sup>中所提到的方法，进行摄像机内参数的标定。这个标定算法的特点是：不需要复杂的辅助标定设备，标定的速度快，操作灵活。同时，为了对摄像机的外参数进行标定，对<sup>[37]</sup>的算法做了推广。在标定转台转动轴方面，提出了一个有效的方法，并且给出了标定算法。

对于这个测量系统的误差，可以从两个方面进行讨论。一个方面是摄像机的标定误差，这个问题的实验结果和结论可以参考<sup>[37]</sup>，本文没有做进一步的研究。另一方面是转台转动轴的标定误差。因为转台后来出现故障，没有得到准确的误差分析。但是从下面的数据我们可以看出标定精度是符合基本要求的。

转台上特征点的空间位置经过  $R_{VM}$  投影后，每个点的  $Y$  分量应该为一个恒量。在实际的数据中，每个点的  $Y$  分量的均值为：55.947mm，均方差为：0.058mm，最大误差为：0.260mm。进行圆拟合后，每个点到圆的平均误差为：0.707mm，最大误差为：1.521mm。图 3.4 显示了拟合出的圆和这些特征点的关系，可以看出特征点基本在拟合的圆上。

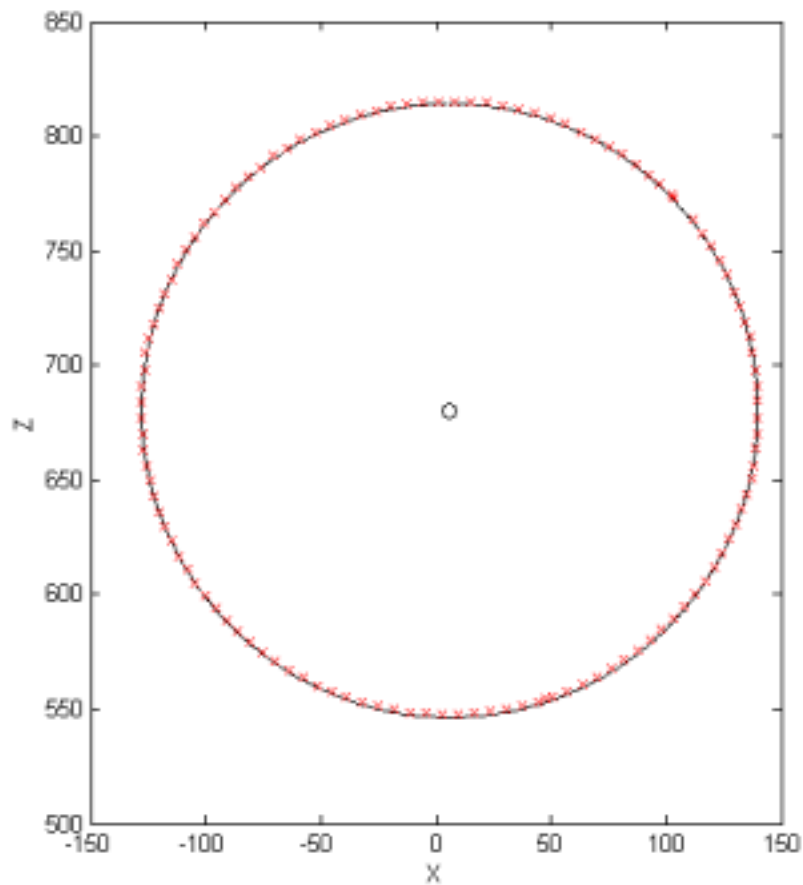


图 3.4 转台上特征点的空间位置经过  $R_{VM}$  投影后，进行圆拟合的结果。其中特征点用红色的叉表示，拟合出来的圆用黑色表示。

## 第四章 具有真实纹理的反向建模方法的研究

### 4.1 引言

渲染一个场景，需要两方面的信息：一方面是场景的几何信息，它描述了场景中物体的结构和物体之间的相互关系；另一方面是场景中光源的分布情况以及物体表面的颜色、材质等和光照模型相关的信息。场景的反向建模就是需要研究如何从真实环境中获取这两方面的信息。

不同的反向建模系统所需要获取的信息不同，他们所研究的侧重点也不一样。早期的反向建模系统主要应用于 CAD 几何模型的构造中，它们的研究重点是如何获取物体的精确几何信息。如在文献<sup>[44]</sup>中提出了利用激光束测量空间点三维信息的系统，在文献<sup>[73]</sup>中提出了如何从空间采集的三维点恢复物体的拓扑结构，在文献<sup>[74]</sup>中提出了一个利用立体视觉原理恢复物体几何信息的系统。

随着基于图像建模(IBM)的兴起，从真实世界获取物体表面光照属性和环境光照的技术得到了极大的发展。这方面比较有代表性的有：在文献<sup>[45]</sup>中提出了一个测量物体表面双向反射分布函数(BRDF)的系统，它通过控制光源和摄像机的位置，测量任意方向的光照在物体表面全方位的反射强度。在文献<sup>[46]</sup>中，提出了测量环境光照的方法。在文献<sup>[47]</sup>中，提出了测量半透明物体光照属性的方法。

由于硬件设备的限制（有两方面的硬件需求，一方面是需要能严格地控制光照环境，能对光的照射强度和方向进行标定；另一方面是能灵活地控制摄像机的位置，需要将摄像机放置在位置可控的机械设备上，可参考文献<sup>[63]</sup>），本系统无法对物体表面的各种光照属性如颜色、材质进行测量。因此我们近似认为在一定的光照条件下，物体表面的颜色等于采集到的图像的颜色，物体表面的材质由手工调节得到。在下一章，我们将专门讨论如何获取和利用物体表面光照属性的问题。

本系统的反向建模可以分为两个过程：几何信息获取过程和纹理信息获取过程。几何信息获取过程的主要工作是：控制机器人平台旋转，利用已经

标定好的双目视觉系统获取物体表面的几何结构信息。纹理信息获取过程的主要工作是：控制机器人平台旋转，获取物体不同方位的图像信息，对这些信息进行处理和融合，最后得到物体的一个全景纹理图像以及模型的纹理坐标。

## 4.2 几何模型的构造

几何模型的反向重构主要是解决两方面的问题：一方面是获取空间点的三维信息；另一方面是曲面重构，确定这些点之间的拓扑关系。不同应用要求不同精度的几何模型，它们所采用的测量系统也不同。有的系统只要求非常粗略的几何模型，如文献<sup>[26]</sup>中的从照面恢复建筑物三维模型的系统。相反，在有的系统中，需要得到非常精确的几何模型，如文献<sup>[75]</sup>。

在本系统中，我们利用结构光确定空间特征点在左右两幅图像中的对应关系，用双目视觉原理确定特征点的三维信息，然后采用一个简单的曲面重构方法进行曲面重构，最后得到物体的几何模型。

### 4.2.1 结构光标在图像上的抽取算法研究

对于利用双目视觉原理测量特征点的空间信息的系统，最大的难点在于如何确定特征点在左右两幅图像中的对应关系。若是物体表面有丰富的纹理，则比较容易采用分析图像区域相似性的方法确定对应关系，如<sup>[9]</sup>中所采用的方法。同时可以采用双目视觉中的一些约束来提高计算对应关系的准确程度<sup>[8]</sup>。也可以在物体表面贴标志物来确定对应关系<sup>[48]</sup>。但是比较鲁棒和方便的方法是采用结构光的方法来确定左右图像特征点的对应关系。

系统所采用的是红色条状光，它会在物体表面形成一个条状光标。我们需要分析摄像头所拍摄的图像，确定这条光标线在图像中的位置。由于光标线位置的准确程度直接影响到最后的测量精度，因此我们从两个方面提高抽取光标线的精度。一方面是对物体的光照环境进行控制，保证物体周围的环境光比较暗，这样结构光会在物体表面形成对比鲜明的光标；另一方面是采用鲁棒的光标抽取算法，如<sup>[49]</sup>中所用的方法。

具体的抽取过程如下：

- a. 对图像进行预处理，使得光标线在图像中变得更加明显和易于抽取。可以采用的操作有：对图像进行直方图均衡，用 gaussian 滤波去处噪音，以及用 Sobel 算子或 Prewitt 算子进行边缘加强。这里需要根据图像的实际情况由用户选定一个预处理操作流程。
- b. 根据一个设定的域值将图像中可能是光标的区域抽取出来。在这些区域中，有的面积非常小，是由噪音引起的，应当去除。
- c. 从光标区域确定光标线，光标线应该是这些光标区域的中心线。因为结构光是垂直连续的，因此我们对结构光在物体表面形成的光标有两个假设：一个是在图像的任何一行只可能有一个光标点；另一个是光标线应该具有连续性。根据这两个假设，我们可以沿图像的竖直方向确定光标线在每一行的位置，而且可以用上一行找到的光标位置约束本行的光标位置。

这个光标线抽取方法抗干扰能力强，有较好的鲁棒。图 4.1 说明了这个抽取过程。

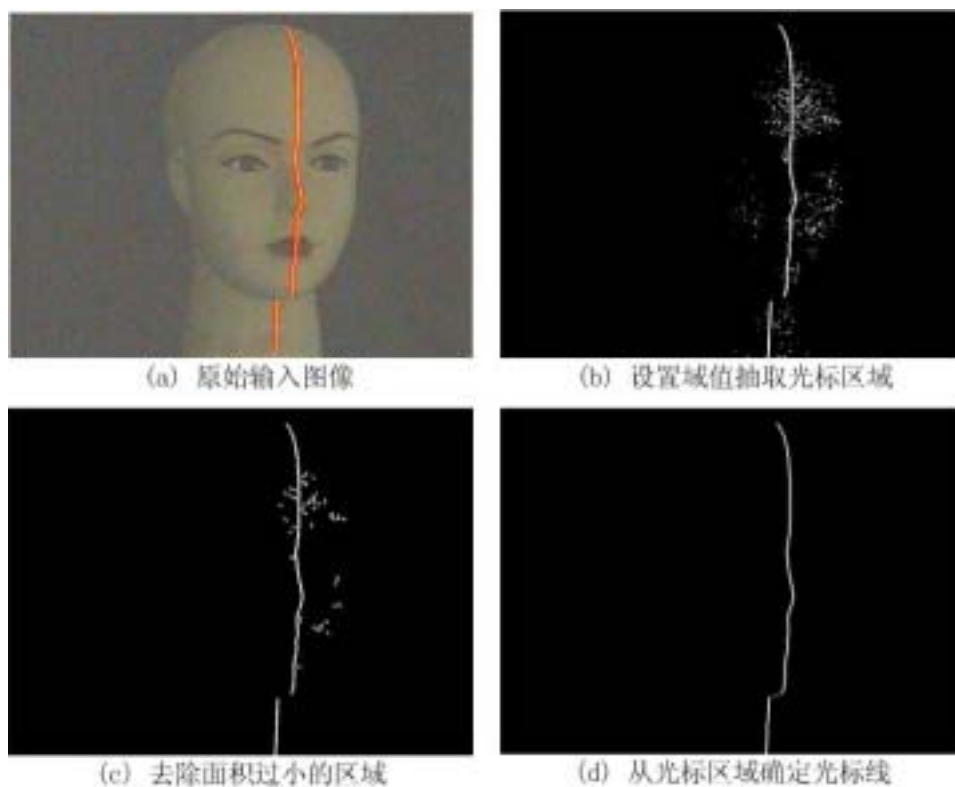


图 4.1 光标线抽取过程

## 4.2.2 双目特征点的对应关系和空间坐标的确定

在确定了图像中光标线的位置后，我们用外极线几何来确定左右两光标线的对应关系<sup>[40]</sup>。

双摄像机系统中，定义基础矩阵  $F$  为：

$$F = A_R^{-1} T_x R_{WCR} A_L^{-1} \quad (4.1)$$

其中， $A_L$ 、 $A_R$  分别为左右摄像机的内参数， $R_{WCR}$  由式(3.2)定义， $T_x$  为反对称矩阵，定义为：

$$\text{若 } t_{WCR} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}, \text{ 则 } T_x = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix} \quad (4.2)$$

这里， $t_{WCR}$  的定义如式(3.2)。

因此空间中任何一个特征点  $M$  在左右图像上的成像点  $m_L$ 、 $m_R$  满足如下关系，

$$m_R F m_L = 0 \quad (4.3)$$

通过式(4.3)，对于左图像光标线上的任意一点，可以唯一确定右图像光标线上相对应的一点。图 4.2 说明了这个对应关系。通过确定左右图像的对应关系，我们可以利用式(3.19)，(3.4)，(3.5)确定光标线上每点在物体坐标系的坐标。

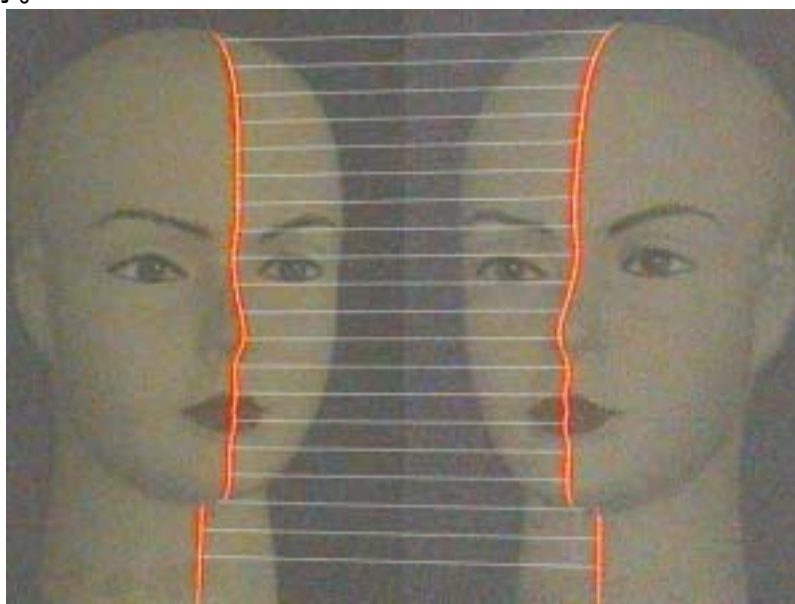


图 4.2 通过外极线约束确定光标线对应关系

### 4.2.3 曲面重构及其平滑



图 4.3 测量所得到的三维空间点集

通过对所有拍摄图像的光标线进行均匀采样和三维重构，可以得到三维空间上的一个点集，如图 4.3。但是只有空间点的信息不能表示一个模型，曲面重构的目的是需要确定这些点之间的拓扑信息。这个拓扑信息通过以这些空间点为顶点的三角面片进行表述。

曲面重构算法的难度取决于空间点集中的点是否已经具有一定的拓扑关系<sup>[10,73]</sup>。因为我们已知一条光标线上的点相互连接，因此只需要决定相邻两条光标线上的点的拓扑关系，因此算法变得相对简单。

我们一般希望三维模型中的三角形的三个角都为锐角，并且边长相差较小。因此我们可以采用类似<sup>[50]</sup>的方法，通过局部最优化的方法进行曲面三角化。具体的做法为：

- a. 将两条光标线顶端的点相连接；
- b. 以刚连接的线为边，从两个候选三角形中选择一个更符合优化准则的三角形；
- c. 反复进行 b 操作，直到完成。

因为所得到的空间点集有较好的初始拓扑信息，因此用上面的方法可以得到较好的曲面重构结果。图 4.4 为通过此曲面重构算法得到的几何模型。



图 4.4 曲面重构后的三维几何模型

可以看到模型表面还不是十分的光滑，有噪音。这主要是由于光标线抽取的误差导致的。我们可以采用一些曲面平滑算法来消除这些不必要的细节和噪音。图 4.5 为采用 VTK 中的平滑功能后得到的几何模型。

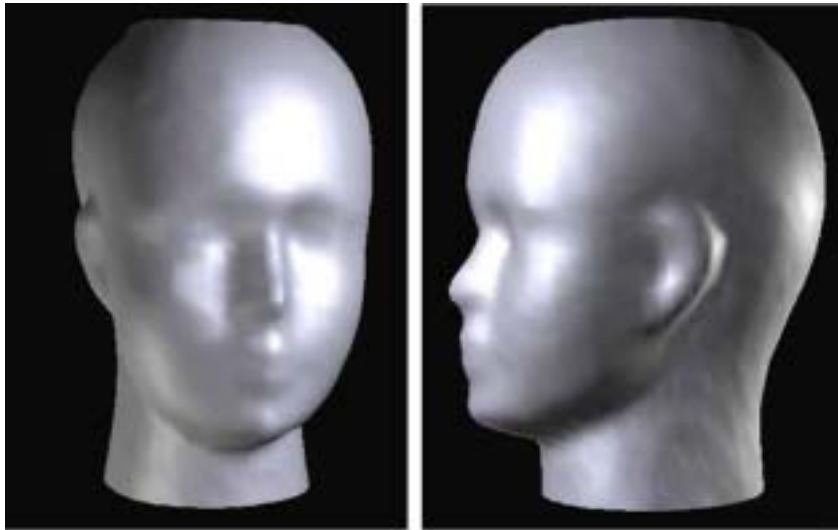


图 4.5 通过平滑后的三维几何模型

### 4.3 纹理和纹理坐标的获取

获得物体的几何结构信息后，需要进一步获取物体表面的颜色，以及材质等光照属性。这需要解决两方面的问题：一方面是确定采集到的物体表面

光照属性和几何模型之间的关系；另一方面是这些不同方位采集到的信息如何融合成一个整体。

不同的系统采用了不同的方法解决这两个问题。在<sup>[50]</sup>中采用激光扫描测量得到物体的几何模型，然后通过照相机获取物体表面的颜色信息，通过七个不变矩确定几何模型和照片中纹理的关系，但是在处理不同角度纹理的拼接时存在问题。在文献<sup>[30]</sup>中采用三维扫描仪获取物体的几何模型，通过照相机获取物体表面在不同视角下的颜色值，再利用手工标定对应点的方法确定几何模型和纹理的关系，最后采用表面光流的方法对不同视角下的纹理颜色进行统一的表达。

在本系统中，我们可以直接运用摄像机完成纹理的采集工作。控制机器人旋转一周，同时摄像机拍摄到物体全方位的纹理信息，然后把这些纹理信息进行处理，得到最后的纹理和纹理坐标。

#### 4.3.1 纹理和纹理坐标的获取

因为摄像机和物体的关系已经确定，因此物体的几何模型和图像中的纹理的关系可以通过公式(3.3)，(3.4)，(3.5)确定。我们需要重点解决的是如何将存在于不同图片中的纹理信息组织起来。

一个简单的方法是：将这些拍摄到的纹理图像直接保存起来，同时每个顶点保留包含这个顶点纹理信息的图像标号和相对应的纹理坐标，如<sup>[50]</sup>中采用的方法。但这样存在两个问题：拍摄到的图片中有大量的无用信息，全部储存会浪费许多纹理存储空间；若是相邻两个顶点用的纹理信息不是取自同一图片，则会造成物体表面不连续的现象。

因此需要将图像中有用的信息提取出来，用一张纹理图片进行表示，同时解决这些信息之间的融合拼接问题。我们采用了两步法纹理映射技术<sup>[23]</sup>来解决这个问题。基本的思路是：指定包围物体的中介三维曲面为圆柱面，然后将图像中的纹理信息投影到这个圆柱面上，三维模型的纹理坐标由模型到圆柱面的映射决定。选择圆柱面作为中介三维曲面的好处是：(1) 圆柱面上的纹理可以方便的存储在一个图像中；(2) 纹理的摄取方式是物体沿转轴转动拍摄得到，因此用圆柱面也更加自然一些。

设  $M_o = (X_o, Y_o, Z_o)^T$  为物体坐标系中三维模型的顶点，则由公式(3.3)，(3.4)，(3.5)可以得到其在图像上的坐标  $m_i = (u_i, v_i)^T$ 。令圆柱面的中轴为物体坐标系的  $Y$  轴，高  $h = \max(Y_o) - \min(Y_o)$ ， $\overrightarrow{OM_o}$  到  $X$  轴的夹角为  $\theta$ ，则  $M_o$  的圆柱面纹理坐标  $m_r = (u_r, v_r)^T$  为：

$$\begin{cases} u_r = \theta / 2\pi \\ v_r = (Y_o - \min(Y_o)) / h \end{cases} \quad (4.4)$$

则根据模型的每个顶点，可以在拍摄图像和圆柱面纹理图像之间确定一个对应关系，通过这个关系对图像进行变形(warp)，可以将图像校正到圆柱面纹理空间。

#### 4.3.2 纹理的拼接和融合

因为我们在物体的不同角度拍摄了多张图像，因此模型表面的一个三角面片可能对应多张图像中的相应区域。如何从这些图像确定这个三角面片的纹理颜色是一个值得研究的问题。文献<sup>[9]</sup>中所用的方法是从三角面片投影面积最大的图片中取纹理颜色。在文献<sup>[51]</sup>中，三角面片的纹理颜色是所有投影图像的加权平均，其中权重是由三角形在图像的投影关系决定的。

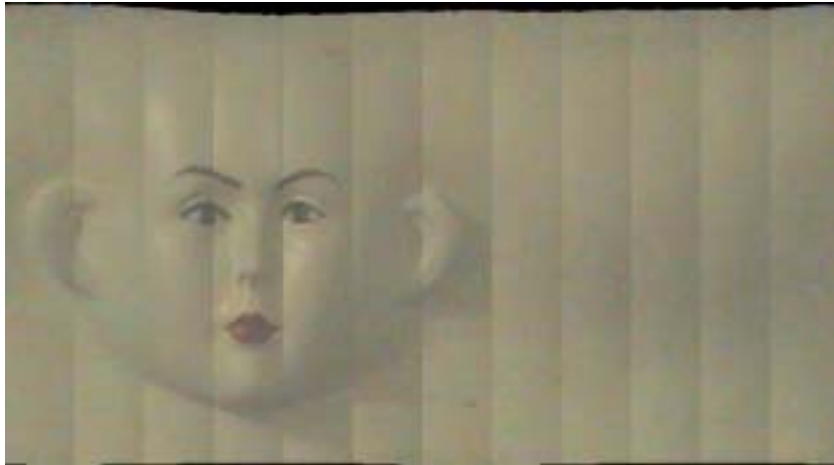


图 4.6 将不同角度图像的部分区域校正到纹理空间并拼合后的结果

我们采用了一个简单的方法解决这个问题。对每幅图像，我们用物体中轴线附近区域在图像中的投影的颜色作为此区域的纹理颜色。通过把这些图

像区域校正到纹理空间，进行拼合，可以得到整个物体的纹理图像，如图 4.6。

从图 4.6 可以明显看到纹理图像的不连续性，这是由于物体表面的光照不均匀引起的。有两个办法可以解决这个问题，一是在光照均匀的环境中拍摄物体的图像，另一个是对不同区域进行融合处理。我们采用了融合处理的方法，具体的做法是在：区域交界处的颜色是相邻两个区域的颜色加权平均。权值是到区域交界线的距离的线性函数。图 4.7 是经过融合处理的纹理图像。



图 4.7 经过融合处理的纹理图像

#### 4.4 重构结果与讨论

通过获取几何模型、纹理图像和纹理坐标，我们完成了反向重构的全过程，所得到的数据可以用标准的程序进行渲染。

重构所得到的人头三维模型采用 Wavefront Obj 格式进行存储，它存储了模型的顶点坐标、三角形连接关系、以及顶点的法向量和纹理坐标。此模型共有 4,997 个顶点和 9,506 个三角形面片。纹理的存储格式是 bmp 格式，尺寸是  $500 \times 900$ ，大小为 1.3Mb。可以采用任何一种支持面片模型和纹理的渲染程序进行渲染。本文采用了 Hemisphere, Inc.公司的 Deep exploration 进行渲染。图 4.8 展示了重构模型在不同角度的渲染效果图。

根据前面的需求分析，对于面向虚拟现实应用的系统，主要考虑渲染模

型整体的真实性，而无需特别注意几何建模精度。从所得到的结果可以看出：利用此反向建模系统，能有效地构造出真实物体的三维模型，并获取其颜色纹理，所得到的结果从肉眼观察的角度来看没有明显的误差，其精度能满足真实感渲染方面的要求。



图 4.8 重构模型渲染效果图

本系统采用了普通工业摄像头作为基本的测量设备，因此整个系统的成本较低。由于构造模型时，转台最少只需要旋转一周就可以完成几何信息和纹理信息的采集，因此以后可以考虑采用手动转台，由人根据刻度盘指示，手工控制物体的转动角度，这样可以进一步降低系统的硬件成本。

但是，这个系统也存在问题，要继续完善。因为机器人控制转台后来出现了问题，对这个系统的测量精度没有给出一个准确数值。系统目前对被测物体的拓扑结构还有限制，若是存在自遮挡，则无法准确得到物体的几何模型，而且无法构造出模型顶部的几何模型。并且可以看出渲染结果和实际物体表面的颜色有差异，这主要是物体周围的环境光控制得不好，还因为没有考虑物体表面的光照模型，直接是把采集到的图像当作物体的纹理颜色。在进行试验的时候也发现物体表面的反射特性对光标线的抽取有很大影响，当物体表面对光的吸收太强，则光标线的抽取存在问题。如何解决这些问题，让这个系统能真正实用，需要进一步研究。

另外一个值得考虑的问题是三维模型的化简问题，这对于虚拟现实的实时性应用是非常重要的。因为模型的细节用纹理表示，因此几何模型可以进行一定程度的化简，而不会导致模型的视觉效果失真。可以在进行曲面化简的时候不考虑纹理的约束，只是保留未删除顶点的纹理坐标，如文献<sup>[91]</sup>的算法。也可以在曲面化简的时候，同时考虑纹理对化简的约束，这样能得到更好的效果。如在文献<sup>[92]</sup>中提出的算法，可在曲面化简时尽量保留物体表面各种属性，包括颜色、法向量等。而在文献<sup>[93]</sup>中，探讨了具有纹理的几何模型的化简算法，这个算法考虑了在模型和纹理配合时，如何控制三维模型化简所导致的失真程度。

图 4.9 展示了采用 MS DirectX9 所实现的<sup>[91]</sup>算法，进行曲面化简的结果。这个算法的特点是化简的速度快，并且适用于模型逐步细化的渲染方式。可以看出，当模型的顶点数减小到 300 个时，还能得到较好的渲染效果。当顶点的数目较小到 180 个时，模型有较明显的变形，但是还可以勉强接受。以后可以在自动建模系统中，加入一个曲面化简的软件包。用户可以对获取的三维模型进行后处理，在不损失视觉效果的情况下，尽量减小数据的复杂度。

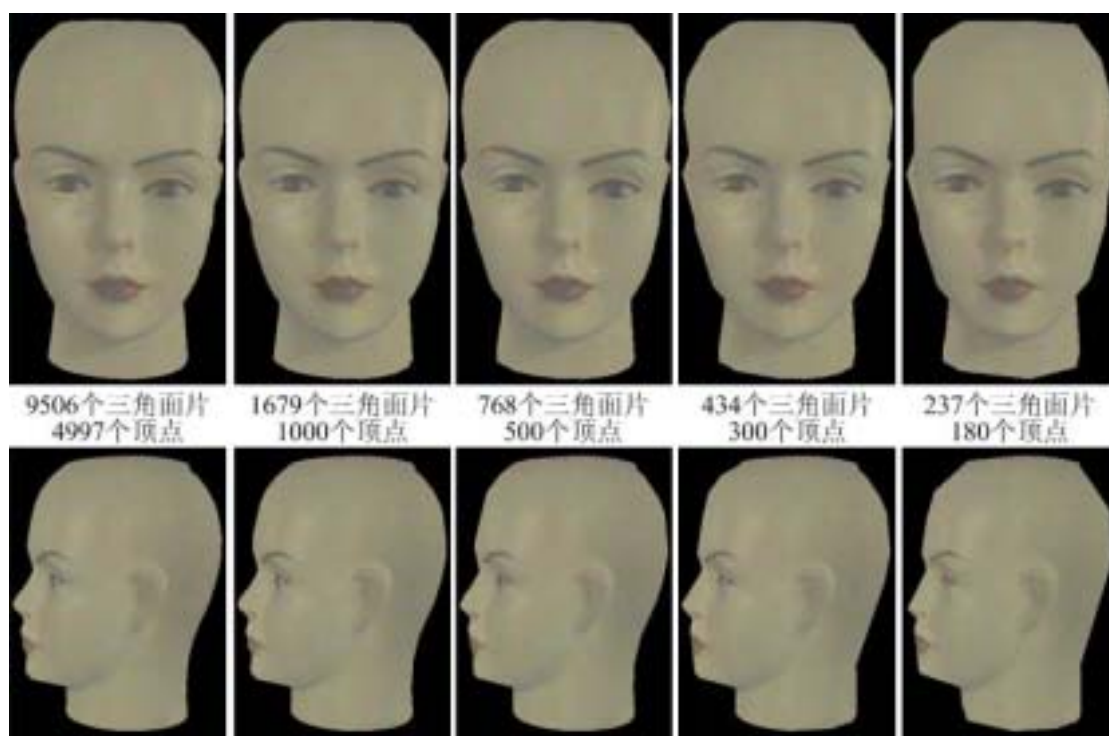


图 4.9 模型化简后的渲染效果。其中上面一行为正面渲染结果，下面一行为侧面渲染结果。

## 第五章 任意曲面上BTF的合成算法研究

### 5.1 引言

对于虚拟场景的真实感渲染问题，只考虑场景中物体的几何结构信息以及物体表面的基本颜色是不够的。例如通过前面设计与实现的建模原型系统所反向重构出的模型，虽然可以表示物体的形状和颜色纹理，但是还是能分辨出渲染的物体和真实的物体存在视觉上的差异。这些差异是因为：在模特人头的物体表面有随位置变化而变化的微小几何结构（mesostructure）和光照特性，使入射光线在物体表面产生了复杂的反射。为了进行真实感渲染，我们需要对物体表面这些复杂的反射现象进行研究。

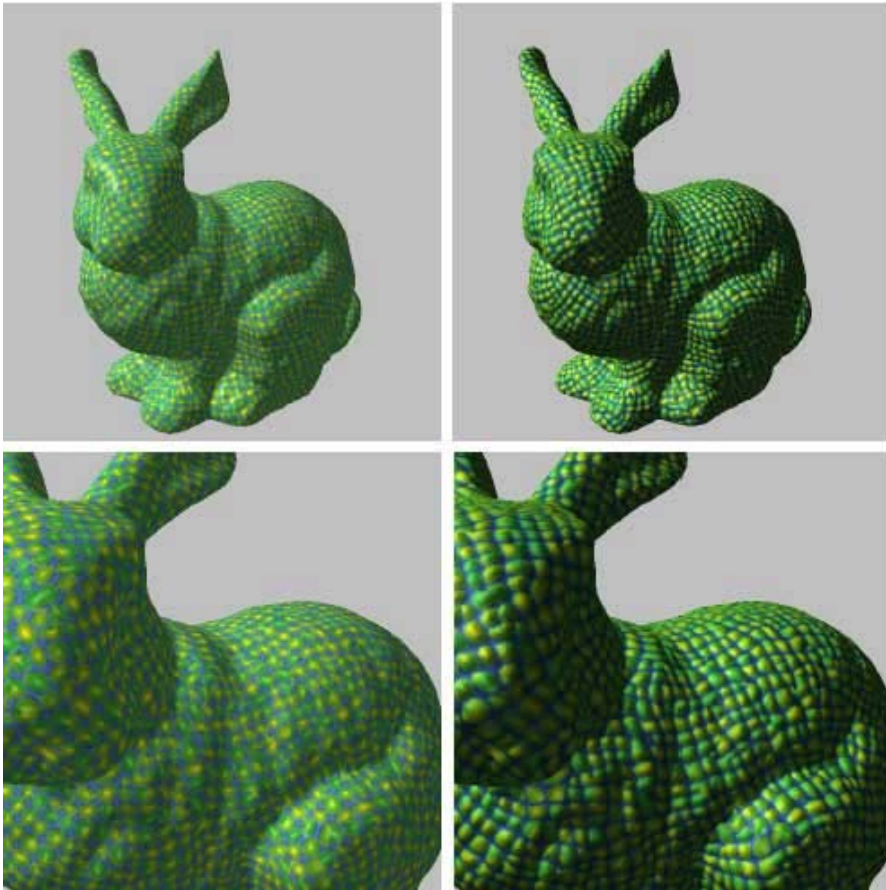


图 5.1 用颜色纹理和用 BTF 纹理进行渲染的效果对比。其中左边图片采用颜色纹理进行渲染，右边图片采用 BTF 纹理进行渲染。

用 BTF (双向纹理函数) 可以描述物体表面一个点上的任意方向入射光所产生的任何方向的反射光的强度。因此用它来实现物体表面的真实感渲染是非常适合的。图 5.1 对比了物体表面只用颜色纹理和用 BTF 进行渲染的区别, 可以看出使用了 BTF 后, 物体表面的立体感更强, 物体给人的感觉更为逼真。

BTF 是一个 6 维的函数, 要采集到一个物体表面完整的 BTF 是非常困难的, 而且表示一个物体表面的 BTF 需要大量的数据。一个自然的想法是采集一个材料的样本 BTF, 然后通过纹理映射的方法将这个 BTF 映射到曲面上。但是采用简单的纹理映射会造成 BTF 的局部扭曲, 同时因为不能保证 BTF 样本是可拼结的, 这会在物体表面造成明显的接缝。

可以采用类似<sup>[68]</sup>的方法对接缝处的 BTF 进行混合 (blend), 这虽然可以消除接缝, 但是此时在接缝处 BTF 所对应的表面的微小几何结构在不同光照条件下表现不一致, 如图 5.2 所示。因此这种方法也是不可行的。

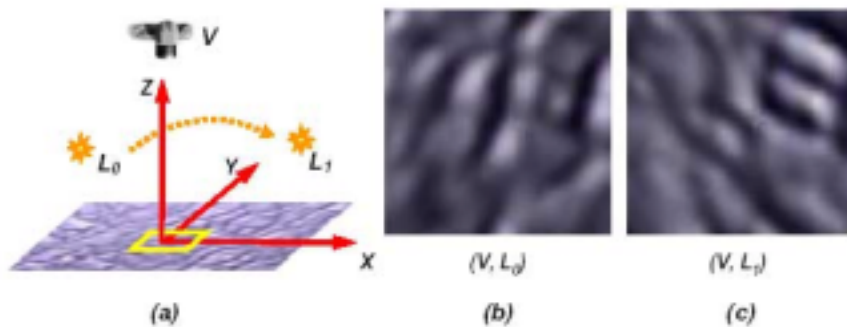


图 5.2 对接缝处的 BTF 进行混合会造成接缝处 BTF 所对应表面的微小几何结构在不同光照条件下表现不一致。(b)、(c)对应(a)中的黄框内的部分, 它们对应的视角方向都是  $V$ , 但是光照方向分别是  $L_0$  和  $L_1$ 。

纹理合成算法是做 BTF 曲面映射的理想选择。但是将纯粹的纹理合成算法用于 BTF 合成却是困难的。这是因为纹理合成算法的时间复杂度和每个像素的维数密切相关。在颜色纹理合成的过程中, 每个像素只是一个 3 维的向量, 若匹配邻域的宽度为  $n$ , 则由一个像素的邻域构成的高维向量的维数是  $3n^2$ , 通过一些加速算法<sup>[69]</sup>, 从向量集中选取匹配向量的速度还可以忍

受。在 BTF 合成的过程中，BTF 每个像素保存有各个入射和反射方向的信息，每个像素的维数都在 1000 维以上，此时计算复杂度是不可以忍受的。采用主元分析法（Principal components analysis）可以在一定程度上减少每个像素的维数，但是还是无法解决这个问题。

在文献<sup>[66]</sup>中，提出了三维纹理元（3D textons）的概念，它是基于这样的假设：在一个 BTF 的局部的区域，只可能有若干种肉眼能分辨的几何结构和反射特性。每一类局部几何结构和反射特性称为一个三维纹理元，一个 BTF 是由有限的三维纹理元构成的。这样，三维纹理元能紧致表示 BTF。但是在文献<sup>[66]</sup>中，主要讨论了如何用三维纹理元对不同种类的 BTF 进行识别和分类，不能直接将所得到的结论用到 BTF 合成中来。

本文对三维纹理元进行了推广，采用它对 BTF 数据进行分析，提取 BTF 合成算法中所需要的信息。我们从 BTF 样本中抽取出纹理元结构图（texton map），然后直接将纹理元结构图合成到任意的曲面上。同时为了进一步提高 BTF 的合成速度，我们提出了一个有效的加速算法。

## 5.2 系统基本结构

系统的流程图如图 5.3 所示。对于一个 6 维的 BTF 函数  $T(x, y, \theta_i, \phi_i, \theta_r, \phi_r)$ ，我们可以看作是一个纹理，其中每个像素  $(x, y)$  由一个四维的函数  $T_{(x,y)}(\theta_i, \phi_i, \theta_r, \phi_r)$  构成。给定一个 BTF 样本  $T$  和一个曲面  $M$ ，我们通过三维纹理元分析和合成表面纹理元结构图这两个阶段生成最后的表面 BTF  $T'$ 。其中  $T'$  和  $T$  有相同的光照性质，但是  $T'$  不是由  $T$  简单拼接构成。

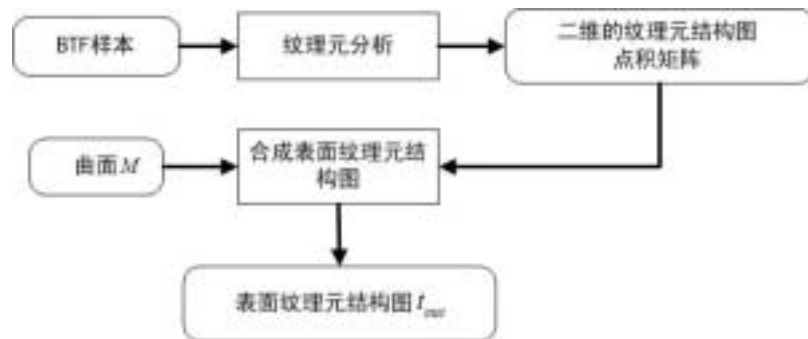


图 5.3 系统的流程图

1. 在 BTF 的分析阶段, 生成一个二维的纹理元结构图  $t_{in}$  和建立纹理元空间  $S$ 。我们用<sup>[66]</sup>所提出的方法, 从 BTF 样本  $T$  中构造一个三维纹理元集合  $V = \{t_1, \dots, t_{n_t}\}$ 。然后可以确定样本  $T$  中每个象素所属的三维纹理元的标号, 这些标号构成的二维表就是纹理元结构图  $t_{in}$ 。

纹理元空间  $S$  是用  $\{t_1, \dots, t_{n_t}\}$  作为基底张成的一个内积空间,  $S$  中的任意一点是  $\{t_1, \dots, t_{n_t}\}$  的一个线性组合, 我们称之为表面纹理元 (surface texton)。在 BTF 合成时, 需要能计算纹理元空间  $S$  中任意两个点之间的  $L_2$  距离。因为三维纹理元  $t_i$  对应于一个上千维的表征向量  $v_i$ , 直接计算  $L_2$  距离的计算量非常大, 并且保存  $\{v_1, \dots, v_{n_t}\}$  需要大量的内存空间。因此我们只存储一个点积矩阵  $\Gamma$ , 这个矩阵由集合  $\{t_1, \dots, t_{n_t}\}$  中所有元素的相互点积构成。由于  $\{t_1, \dots, t_{n_t}\}$  不是正交基, 矩阵  $\Gamma$  不是对角阵, 需要存储  $n_t \times n_t$  个元素。因为一个 BTF 样本中三维纹理元的个数是有限的, 一般为几百个, 因此点积矩阵  $\Gamma$  的数据量较少。

2. 在 BTF 的合成阶段, 生成纹理元结构图  $t_{out}$ 。 $t_{out}$  是一个表, 它的每一项与几何模型的每个顶点  $v$  相对应, 存储了这个顶点的三维纹理元的标号和这个点的纹理坐标  $p_v = (a_v, b_v)$ 。这样我们在模型表面合成了 BTF  $T'$ , 其中每个顶点的数值  $T'_v(\theta_i, \phi_i, \theta_r, \phi_r) = T(a_v, b_v, \theta_i, \phi_i, \theta_r, \phi_r)$ 。

我们把纹理元结构图  $t_{in}$  作为一个纹理样本, 然后用曲面的纹理合成算法生成  $t_{out}$ 。在生成过程中, 将逐个计算每个顶点所对应的三维纹理元和这个纹理元在样本  $T$  中的坐标。这个过程主要涉及的计算是确定纹理元空间  $S$  中任意两个点的  $L_2$  距离, 我们可以用分析阶段得到的点积矩阵  $\Gamma$  完成这个计算。

下面, 我们将详细讨论每一部分的基本原理及其相应算法。

### 5.3 三维纹理元的分析和抽取

BTF 的分析阶段由如下三个部分组成: (1) 从 BTF 样本  $T$  建立三维纹理元集合; (2) 给每个象素确定一个三维纹理元的标号, 从而得到纹理元结构图  $t_{in}$ ; (3) 建立纹理元空间  $S$  的点积矩阵  $\Gamma$ 。整个过程的原理图如图 5.4。

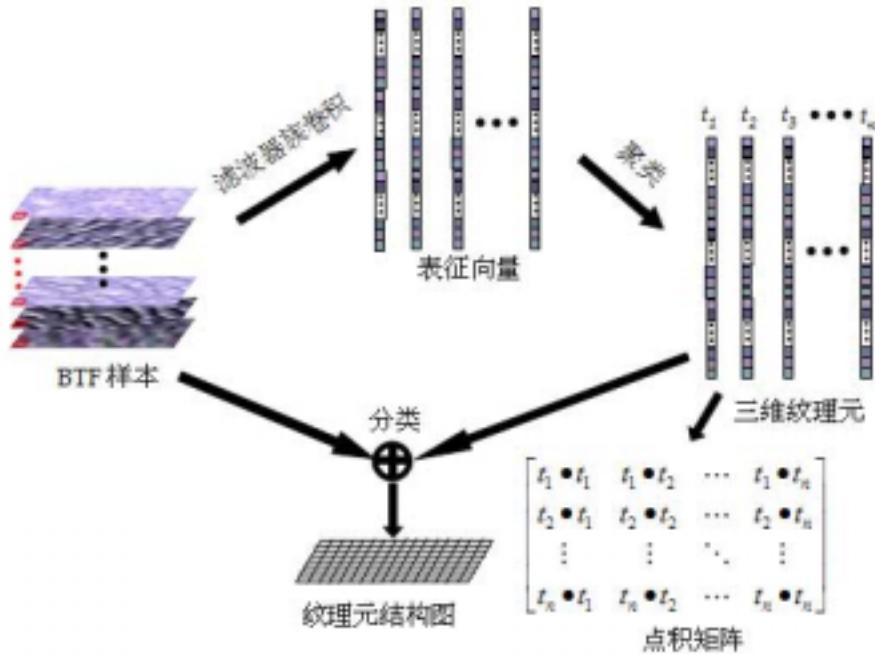


图 5.4 三维纹理元的分析过程

### 1 三维纹理元集合的建立

因为在一个 BTF 的局部区域，只可能有若干种肉眼能分辨的几何结构和反射特性。因此可用一定的方法来把每个像素进行分类，使每类中以这些像素为中心的局部区域有相似的结构和反射特性。我们采用了文献<sup>[66]</sup>中的方法来完成这个工作。

这个方法的基本步骤如下：

a. 确定一个滤波器族 (filter bank)。这个滤波器族由二维高斯函数衍生得到，基本的图形如图 5.5。从直观上来说，每个滤波器表示一个视觉系统能感应到的基本纹理元素，这个滤波器族包含了各种可能的元素。在计算中，我们选用的滤波器的个数为  $n_b = 48$ 。它包括了不同尺度，不同方向的高斯衍生滤波器。

b. 从构成 BTF 样本的几千幅图像中选取  $n_s$  幅不同光照和视角条件下的图像，要求选取的图像之间的差别尽量大一些，每幅图像具有一定的代表性。选取的方法可以是：把每幅图像作为高维空间的一个点，然后通过分类算法找出  $n_s$  个相互距离最远的点。同时，也可以采用文献<sup>[70]</sup>中提出的方

法。计算中我们设  $n_s = 400$ 。

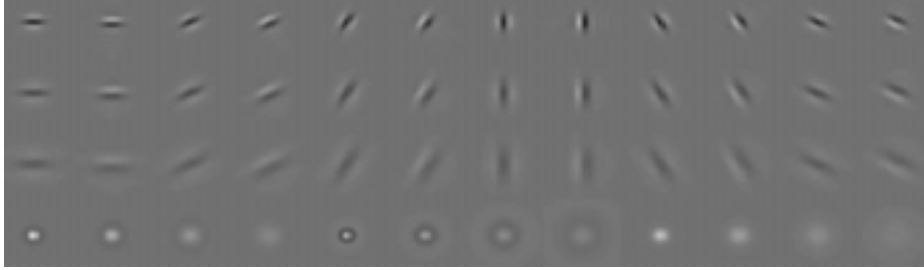


图 5.5 纹理元分析时采用的滤波器族，来自文献<sup>[66]</sup>。

c. 用滤波器族对选出的  $n_s$  幅图像进行卷积，然后把每个象素的所有卷积值组合成一个向量。这样 BTF 样本的每个象素对应于一个  $n_s \times n_b$  维的表征向量，这个向量刻画了以这个象素为中心的局部区域在不同的光照和视角条件下给视觉系统产生的响应。

d. 对这些表征向量进行  $k$  均值聚类，所形成的类中心  $\{t_1, \dots, t_{n_t}\}$  称为三维纹理元。这些类中心和  $n_s \times n_b$  维的表征向量  $\{v_1, \dots, v_{n_t}\}$  相对应。

这样我们就建立了三维纹理元集合。

## 2 确定纹理元结构图

在建立了纹理元集合  $\{t_1, \dots, t_{n_t}\}$  后，我们能根据这个集合建立 BTF 样本  $T$  的纹理元结构图  $t_{in}$ 。象素  $p$  的纹理元标号  $t_{in}(p) = \arg \min_{j=1}^{n_t} \|v(p) - v_j\|$ ，其中  $v(p)$  是象素  $p$  的表征向量， $v_j$  是纹理元  $t_j$  的表征向量。

## 3 建立纹理元空间的点积矩阵

用  $\{t_1, \dots, t_{n_t}\}$  作为基底可以构成纹理元空间  $S$ 。 $S$  中的任何一个向量为  $s = \sum_{i=1}^{n_t} a_i t_i$ ，其中  $a_1 \dots a_{n_t}$  为实数。我们称  $s$  为表面纹理元。设两个纹理元  $t_i$ 、 $t_j$  所对应的表征向量为  $v_i$ 、 $v_j$ ，则这两个纹理元的点积定义为  $t_i \bullet t_j = v_i \bullet v_j$ 。我们预计算出每一对基本向量的点积值，将其存储在  $n_t \times n_t$  的点积矩阵  $\Gamma = (\gamma_{ij})$  中，其中  $\gamma_{ij} = t_i \bullet t_j$ 。因为  $S$  是内积空间，我们可以只由  $\Gamma$  而不直接根据  $\{v_1, \dots, v_{n_t}\}$  计算  $S$  中任意两点，即任何两个表面纹理元之点积。

设  $s = \sum_{i=1}^{n_t} a_i t_i$ 、 $s' = \sum_{i=1}^{n_t} a'_i t_i$ ，则有

$$s \cdot s' = \sum_{i,j=1}^{n_i} a_i a'_j \gamma_{ij} \quad (5.1)$$

## 5.4 曲面上BTF的合成

在这一节，我们将讨论如何将纹理元结构图  $t_m$  作为一个纹理样本，用曲面的纹理合成算法把 BTF 生成到任意曲面上去。我们将先讨论单尺度的 BTF 合成算法，然后讨论多尺度的合成算法，最后提出一个算法来加速 BTF 合成的速度。

### 5.4.1 单尺度BTF合成算法

在进行曲面的 BTF 合成前，我们需要对面片模型  $M$  进行预处理<sup>[57,58]</sup>。这主要包括两个方面的工作：

一方面是需要根据初始的面片模型构造一个顶点更为密集模型。这是因为我们将表面纹理元的信息存储于每个顶点上，每个顶点对应于 BTF 样本的一个像素，因此我们需要密集的顶点来存储 BTF 合成的结果。

另一方面是需要确定每个顶点的局部坐标系  $(\vec{s}, \vec{t}, \vec{n})$ ，其中  $\vec{n}$  是这个顶点的法向， $\vec{s}$ 、 $\vec{t}$  通过模型的向量场 (vector field) 计算得到，分别表示局部坐标系向右和向上的方向。

#### 1 纹理元在曲面的合成

在单尺度的纹理元合成过程中，我们顺次遍历目标模型  $M$  的所有顶点，计算每个顶点  $v$  所对应的纹理元  $t_{out}(v)$ 。计算过程分为如下几步：首先，我们基于点  $v$  的局部坐标系  $(\vec{s}, \vec{t}, \vec{n})$  在  $(\vec{s}, \vec{t})$  平面建立一个邻域  $N(v)$ ， $N(v)$  由  $v$  周围的顶点所对应的纹理元进行重采样构成。然后，从 BTF 样本的纹理元结构图中选取一些像素构成  $t_{out}(v)$  的候选集合  $C(v)$ 。接着从候选集合  $C(v)$  找到一个像素  $p_0 = (a_0, b_0)$ ，满足  $N(v)$  和  $N(p_0)$  之间的距离最小，其中  $N(p_0)$  是由  $p_0$  周围的像素所对应的纹理元构成的邻域。最后，我们设定  $t_{out}(v)$  的纹理元标号为  $t_m(p)$ ，同时这个点的纹理坐标为  $p_v = (a_0, b_0)$ 。这个过程的伪代码如下：

```

For each vertex  $v$  on surface
    construct neighborhood textons  $N(v)$ 

     $smallest\_match = \text{BIG}$ 
    from the candidates set  $C(v)$ 
    For each pixel  $p = (a, b)$  in  $C(v)$ 
        construct neighborhood textons  $N(p)$ 
         $new\_match = \text{distance}(N(v), N(p))$ 
        if ( $new\_match < smallest\_match$ )
             $smallest\_match = new\_match$ 
             $t_0 = t_{in}(p), (a_0, b_0) = (a, b)$ 
     $t_{out}(v).texton\_label = t_0$ 
     $t_{out}(v).texton\_coordinate = (a_0, b_0)$ 
    
```

通过  $t_{out}$ ，我们可以知道点  $v$  处的 BTF 值等于其对应象素  $p_0$  点的 BTF 值。 $p_0$  的确定是通过确定  $N(v)$ 、 $N(p_0)$  之间的相似性确定的。因为纹理元是刻画 BTF 局部特性的数值量，因此  $N(v)$  和  $N(p_0)$  相似说明顶点  $v$  和象素  $p_0$  领域的 BTF 值存在相似性。这说明非参数化纹理合成算法中所用的相似性比较是合理的。 $N(v)$ 、 $N(p)$  之间的距离可以方便的通过我们预先计算的  $\Gamma$  得到。

## 2 纹理元的重采样

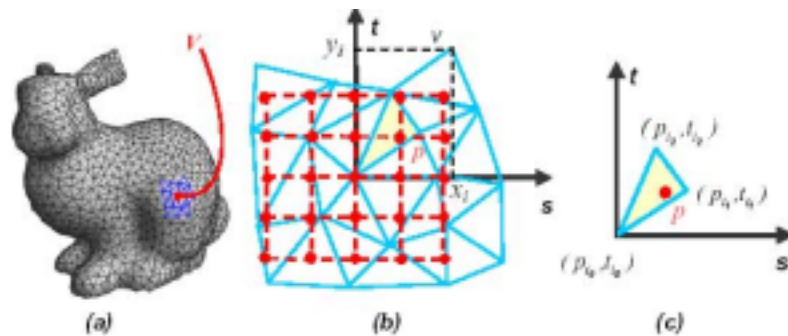


图 5.6 纹理元的重采样过程。(a)中顶点  $v$  周围的三角形进行局部的参数化，将这些三角形平铺到平面  $(\vec{s}, \vec{t})$  上构成图(b)的蓝色三角面片集  $P(v)$ 。用于进行重采样的领域模板为图(b)的红色部分。(c)说明了如何对(b)中浅黄色三角形所包含的象素  $p$  进行纹理元的重采样。

在构造邻域  $N(v)$  的时候必须对纹理元进行重采样。我们基于点  $v$  的局部坐标系  $(\vec{s}, \vec{t}, \vec{n})$  在  $(\vec{s}, \vec{t})$  平面建立一个邻域  $N(v)$ ，如图 5-6 所示。具体的过程如下：首先，对顶点  $v$  周围的三角形进行局部的参数化，将这些三角形平铺到平面  $(\vec{s}, \vec{t})$  上，构成一个三角面片集  $P(v)$ ，这个过程详细说明请查阅参考文献<sup>[58,68]</sup>。然后在平面  $(\vec{s}, \vec{t})$  上确定一个领域模板。最后对每个在邻域  $N(v)$  内的象素  $p$ ，计算所对应的表面纹理元的值  $s(p)$ ：

$$s(p) = \omega_0 t_{i_0} + \omega_1 t_{i_1} + \omega_2 t_{i_2} \quad (5.2)$$

其中， $(t_{i_0}, t_{i_1}, t_{i_2})$  是  $P(v)$  中包含象素  $p$  的三角形三个顶点所对应的纹理元， $(\omega_0, \omega_1, \omega_2)$  是  $p$  在这个三角形中的重心坐标。

### 3 距离计算

我们从候选集合  $C(v)$  找到一个象素  $p_0 = (a_0, b_0)$ ，满足  $N(v)$  和  $N(p_0)$  之间的距离最小。因此，我们需要对  $C(v)$  中的每一个象素  $p$ ，计算邻域  $N(p)$  和  $N(v)$  之间的距离。距离计算公式如下：

$$distance(N(v), N(p)) = \sum_{\lambda=1}^{n_v} \|s(p_\lambda) - t_{j\lambda}\|^2 \quad (5.3)$$

其中  $n_v$  是  $N(v)$  中象素的个数， $s(p_\lambda)$  是  $N(v)$  中一个象素的表面纹理元， $t_{j\lambda}$  是  $N(p)$  中对应象素的纹理元。

式(5.3)的每一项  $\|s(p_\lambda) - t_{j\lambda}\|^2$  可以写成两个表面纹理元的点积：

$$\|s(p_\lambda) - t_{j\lambda}\|^2 = (s(p_\lambda) - t_{j\lambda})(s(p_\lambda) - t_{j\lambda}) \quad (5.4)$$

这样可以方便的利用式 (5.1) 进行计算。

#### 5.4.2 多尺度BTF合成算法

为了能提高 BTF 合成的质量，我们设计和实现了一个和参考文献<sup>[57,58,71]</sup>类似的多尺度合成算法。在预处理阶段，建立纹理元金字塔和曲面的金字塔表述，一般我们采用四层结构。对于建立纹理元金字塔，先构造 BTF 样本图片的金字塔表述，然后在每一层，我们构造和这一层相对应的纹理元结构图和点积矩阵。对于建立曲面的金字塔，采用了<sup>[57]</sup>的算法，这个算法保证精度从高到低的面片模型的顶点数按四分之一递减。

在 BTF 合成阶段，将从金子塔的顶端开始，尺度由粗到细进行 BTF 合成。在每个尺度上进行两次纹理合成。第一次合成的目的是把上面较粗一层的 BTF 合成结果作为约束，给本层每个顶点确定一个初始的纹理元。第二次合成的目的是对本层每个顶点求精(refine)，确定其最终对应的纹理元。这个过程的原型是：可以把 BTF 看作是一个信号，可以先确定其低频部分，然后再确定它的高频部分。这样由粗到细，逐步求精的过程，可以大大提高 BTF 合成的质量。

### 5.4.3 一个加速算法

在 BTF 合成算法中，如何构造每个顶点  $v$  的候选集合  $C(v)$  是非常重要的。全搜索的方法是认为  $C(v)$  包含纹理结构图  $t_m$  的所有象素，文献<sup>[57,58]</sup>就采用了这样的方法，同时在这两个文献中还提出了针对颜色纹理图像的加速搜索方法。但是因为从  $C(v)$  中寻找符合最小距离定义的象素的时候，无法采用<sup>[57,58]</sup>中所提出的加速方法，因此 BTF 合成的速度非常的慢。

我们提出了  $k$  - 连续 ( $k$ -coherence) 算法来加速 BTF 合成的过程，这个算法通过构造一个较小  $C(v)$  来减少搜索所用的时间。可以采用这个方法的原因是纹理具有连续性<sup>[72]</sup>。连续性的意义如下：设  $v_i$  是  $v$  附近的一个点，点  $v_i$  的纹理值从样本纹理的象素  $p_i = (a_i, b_i)$  处取得，在以  $v$  做局部参数化的平面  $(s, t)$  上  $v_i$  到  $v$  的位移为  $\vec{d}_i = (dx_i, dy_i)$ ，则  $v$  的纹理值非常可能从  $p'_i = (a_i + dx_i, b_i + dy_i)$  处获得。也就是说，每个在  $v$  附近的点  $v_i$  可以推荐一个候选象素  $p'_i$ ，顶点  $v$  的候选集合  $C(v)$  由这些候选象素构成。

$k$  - 连续算法推广了这个思路。在预处理阶段，对于 BTF 样本的每一个象素  $p_i$ ，确定一个集合  $coh(p_i) = \{p_{i1}, p_{i2}, \dots, p_{ik-1}\}$ ，这个集合有  $k-1$  个元素，它们的邻域  $N(p_{ij})$  是所有的象素中和  $N(p_i)$  距离最小的  $k-1$  个。在构造  $C(v)$  的时候，对于  $v$  附近的一个点  $v_i$ ， $C(v)$  不但包括  $p'_i$ ，还包括  $p'_{i1}, \dots, p'_{ik-1}$  这  $k-1$  个象素。其中  $p'_{ij} = p_{ij} + \vec{d}_i = (a_{ij} + dx_i, b_{ij} + dy_i)$ ， $(a_{ij}, b_{ij})$  为象素  $p_{ij}$  的坐标。换句话说，每个在  $v$  附近的点  $v_i$  可以推荐  $k$  个候选象素。 $k$  的值由人工设定，取值在 1 - 12 之间。一般来说具有较多高频信息并且纹理不完全符合马尔可夫随机场定义的采用较小的  $k$  值，符合马尔可夫随机场定义的纹理采

用采用较大的  $k$  值。

采用  $k$  - 连续算法后， $C(v)$  中的候选元素大大减小，这加快了 BTF 合成的速度。图 5.7 显示了将  $k$  - 连续算法运用于平面颜色纹理合成时， $k$  由小到大变化的时候的纹理合成结果。可以看出对于这个样本纹理，当  $k$  值大于等于 5 后，能合成质量较好的纹理。

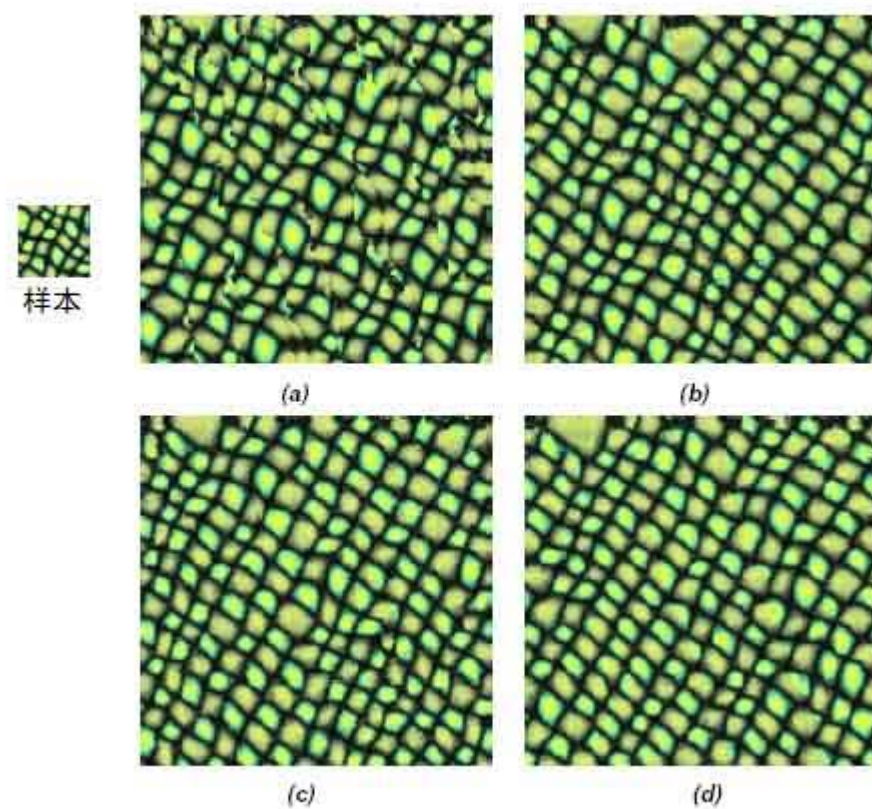


图 5.7 将  $k$  - 连续算法运用于平面颜色纹理合成时， $k$  由小到大变化时的纹理合成结果。(a)  $k=1$ 。(b)  $k=5$ 。(c)  $k=11$ 。(d) 全搜索方法。

## 5.5 曲面BTF的渲染

根据 BTF 样本  $T$  和表面纹理元  $t_{out}$  可以对目标曲面  $M$  进行渲染。具体的方法为：首先将视线和光线的方向由全局坐标系转化到每个可见顶点  $v$  的局部坐标系，计算出在局部坐标系内入射光线的方位角  $\theta_{vi}$ 、极角  $\phi_{vi}$  和视线的方位角  $\theta_{vr}$  和极角  $\phi_{vr}$ 。然后根据  $t_{out}(v)$ ，可以确定顶点  $v$  在  $T$  中的纹理坐标

$p$ 。接着再从组成 BTF 样本的图片集合中找出适当的值进行插值运算，确定顶点  $v$  对应的颜色。因为目标曲面  $M$  非常的细密，因此确定了顶点的颜色后就可以进行高质量的渲染。

因为组成 BTF 样本的图片是在  $(\theta_t, \phi_t, \theta_r, \phi_r)$  空间的一个均匀采样。因此我们可以选取采样点  $(\theta_{vi}, \phi_{vi}, \theta_{vr}, \phi_{vr})$  周围的 16 个值进行插值。可以参考文献 [27] 中的基于视角进行插值的方法。

## 5.6 结果和讨论

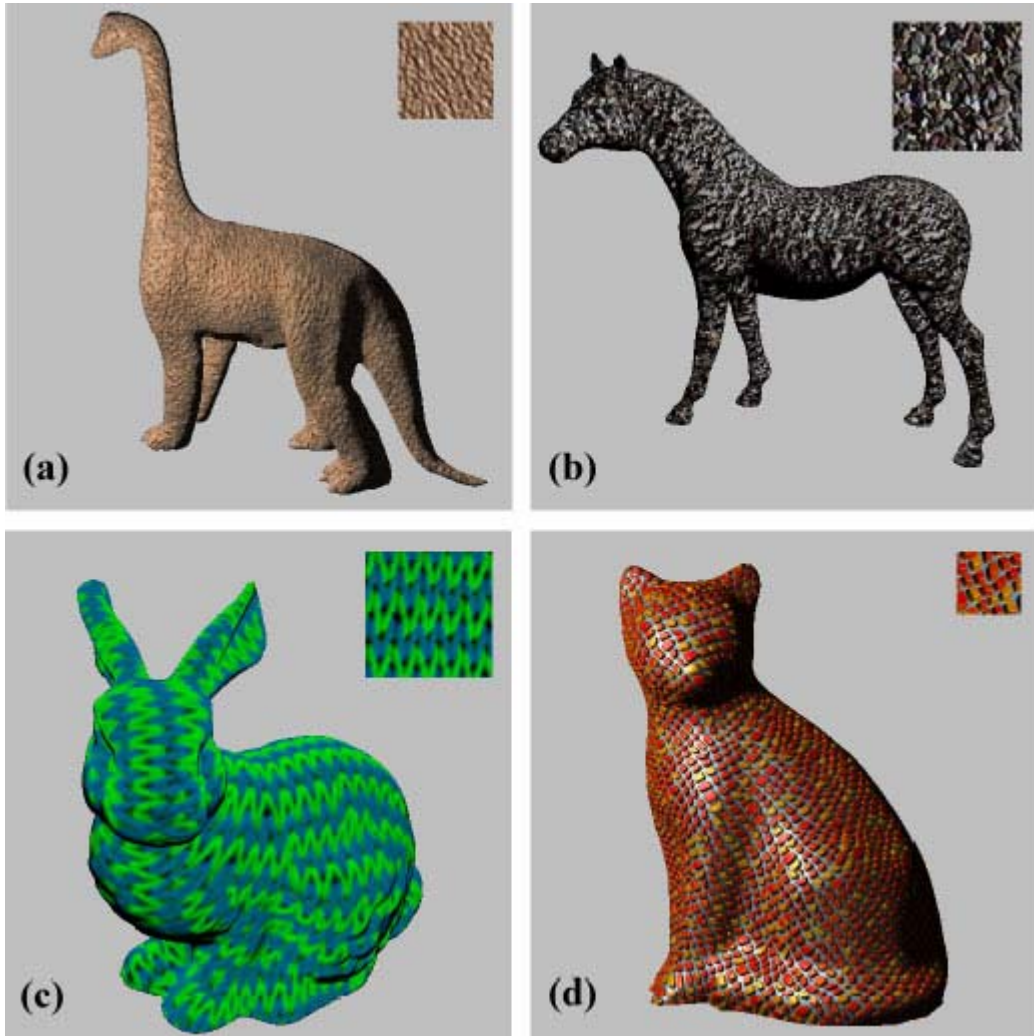


图 5.8 BTF 合成的渲染效果。其中第一行的 BTF 样本是来自 CURET 数据库<sup>[45]</sup>的真实数据，第二行的 BTF 样本是采用光线跟踪算法得到的合成数据。

我们在 PC 机上实现了上述的 BTF 合成算法。这个算法在三维纹理元的分析和抽取阶段是全自动的，不需要用户的交互；在曲面 BTF 的合成阶段，需要用户指定一个方向场，从而控制纹理在模型上的走向，同时需要根据纹理的特性确定一个  $k$  值和邻域  $N(v)$  的大小。试验数据包括来自 Curet 数据库<sup>[45]</sup>的真实数据和用光线跟踪算法得到的合成数据。

图 5.8 展示了几个 BTF 合成的结果。程序运行在 PIII 700 的机器上，计算时间如表 5.1 所示。程序的运行时间主要取决于三个参数：模型  $M$  的顶点数、 $k$  值以及邻域  $N(v)$  的大小。可以看出 BTF 样本无缝并且自然的合成到了模型的表面。并且渲染结果随着光照和视角的改变，表现出模型表面微小几何结构的阴影、相互遮挡以及高光的变化。

表 5.1 合成图 5.8 中的结果的计算时间

模型	样本尺寸	顶点数目	合成时间 (分钟)
恐龙	96 x 96	250,000	21( $k=1$ )
马	128 x 128	250,000	22( $k=1$ )
猫	64 x 64	300,000	141( $k=11$ )
兔子	128 x 128	300,000	186( $k=11$ )

表 5.2 比较了全搜索算法和  $k$  - 连续算法的计算速度。可以看出  $k$  - 连续算法的计算速度是全搜索算法的十几倍以上，其中试验数据是具有 250,000 个顶点的恐龙模型，我们设置  $k=10$ 。并且随着 BTF 样本的增大， $k$  - 连续算法的优势更为明显。图 5.9 是两种搜索算法的结果的渲染效果，可以看出用  $k$  - 连续算法不会降低 BTF 合成的质量。

表 5.2 全搜索算法和  $k$  - 连续算法的计算速度比较

样本尺寸	全搜索算法	$k$ - 连续算法
64 x 64	747 分钟	70 分钟
96 x 96	3,000 分钟	123 分钟
128 x 128	8,066 分钟	157 分钟

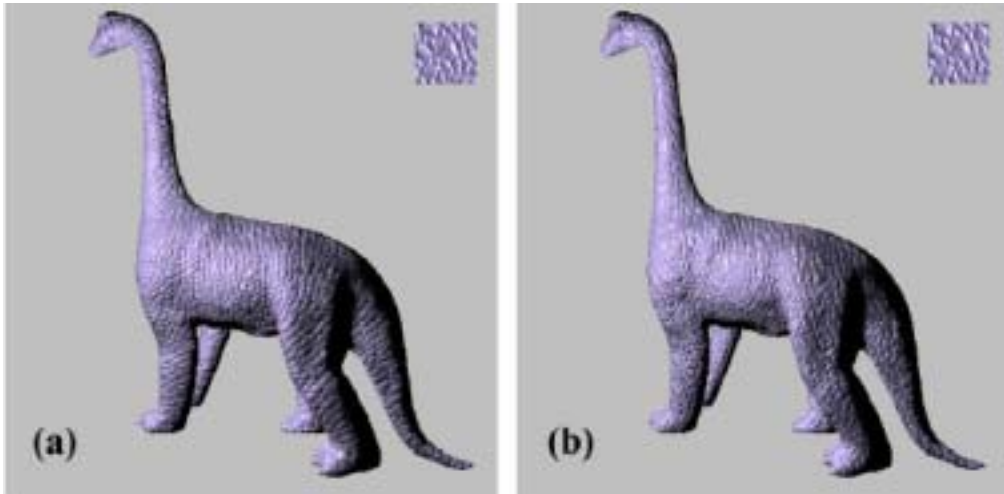


图 5.9 用  $k$  - 连续算法 (左边) 和用全搜索算法 (右边) 的质量比较

同时 BTF 纹理合成也可以看作是一个对物体表面复杂的微小几何结构和光照模型进行快速渲染的方法。我们可以离线对一个小的样本进行光线跟踪渲染, 形成样本 BTF。然后, 再对整个模型进行渲染时, 可以用查表的方式得到模型表面的颜色值。用这样的方法可以渲染出模型表面复杂的光照变化和微小几何结构的相互遮挡的效果, 而不用对整个模型进行光线跟踪渲染。图 5.10 对比了采用 BTF 进行渲染和只采用颜色纹理进行渲染的效果, 可以看出用 BTF 渲染的真实感更强一些。

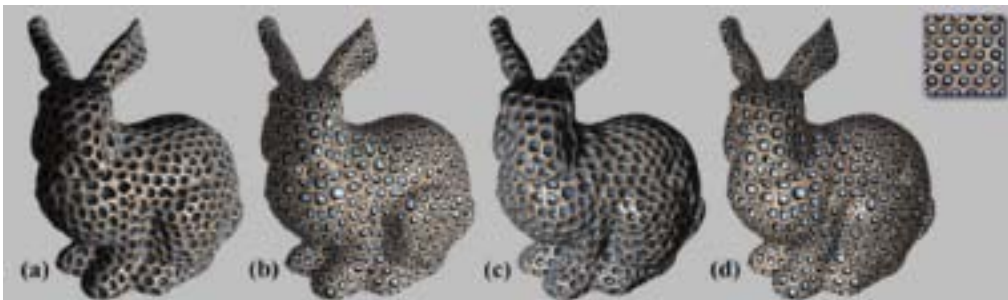


图 5.10 采用 BTF 进行渲染和采用颜色纹理进行渲染的效果对比。(a)、(c)采用了 BTF。(b)、(d)采用了颜色纹理。

本算法的计算时间是一个非常重要的因素, 它决定了整个算法的适用范围。这个问题包括两个方面。一方面是 BTF 在曲面上的合成时间。虽然这需要一个小时左右, 但是并不是整个算法的瓶颈。因为一个模型只需要进行

一次 BTF 合成，而且这个合成过程在场景构建阶段完成，可以离线计算。

另外一方面是对合成结果的渲染时间。因为本文主要研究了 BTF 合成算法，在渲染方面只实现了一个软件渲染的算法。这个软件算法渲染一帧需要一秒多，显然是不能满足虚拟现实中的实时性要求。因此，研究 BTF 合成数据的实时渲染是非常必要的。这方面已经有了一些初步的研究成果。在文献<sup>[67]</sup>中，对 BTF 数据进行了 SVD 分解，并将存储于模型顶点的 BTF 数据进行重采样，变成多个纹理，同时对面片模型进行化简。然后用 Microsoft DirectX 9.0 的 Vertex Shaders 和 Pixel Shaders 功能实现了点光源和方向光源下硬件加速的 BTF 渲染，渲染速度可以达到每秒 10 帧以上。在文献<sup>[90]</sup>中，提出了在环境光照下 BTF 的硬件加速渲染算法。随着硬件技术和软件技术的发展，BTF 渲染在虚拟现实中的应用是大有希望的。

## 第六章 结束语

### 6.1 结论

本文围绕从现实世界获取信息进行虚拟物体的真实感渲染这个主题，从基于图像建模和基于样本建模这两方面对这个问题进行了研究。

在基于图像建模方面，本文设计并实现了一个三维物体快速反向建模原型系统。这个系统运用立体视觉原理测量物体的几何信息，自动构建真实物体的三维模型；并且可以同时获取物体的颜色纹理以及所对应的纹理坐标。用这个系统得到的数据可以直接运用于渲染。基于这个原型系统，我们探讨了系统需求和设计、系统的标定、几何信息的采集、曲面参数化、颜色纹理的拼接和融合等方面的问题。

在基于样本建模方面，本文研究了运用样本 BTF 光照数据进行真实感渲染的问题。我们提出了一个将 BTF 样本数据合成到任意曲面的算法。用这个算法得到的模型表面的 BTF 不但表现出和样本 BTF 相似的光照特性，而且能从不同的光线和视角方向表现模型表面的微小几何结构。因为 BTF 可以反映出真实世界物体表面的特性，因此我们可以在虚拟物体的渲染中用 BTF 来提高物体的真实性。

### 6.2 进一步的工作

在三维物体快速反向建模原型系统中，没有给出一个准确的测量精度。这需要在以后的工作中得到准确的测量精度。同时待测物体的几何结构必须满足每点可见，若是存在遮挡关系则无法得到正确的模型。如何对存在遮挡关系的物体进行自动建模需要进一步的研究。在试验中我们也发现，结构光方法对物体表面的颜色和材质比较敏感，若是物体表面对光的吸收太强，确定图像的光标线会比较困难。为此，需要使用其它求对应点的方法。

对于 BTF 的合成，有的 BTF 数据不能用三维纹理元进行分析<sup>[66]</sup>，此时 BTF 合成算法失效。对于这类 BTF 数据如何合成需要更深入的研究。同时，BTF 样本数据非常的大，需要适当的压缩算法对其进行压缩。在渲染方

面，物体表面的 BTF 渲染还不能达到实时，研究用硬件加速的渲染算法也是非常必要的，可以参照<sup>[67, 90]</sup>的算法进行进一步的研究。

对于如何将基于图像建模和 BTF 合成这两方面的工作结合起来，本文没有做进一步的研究。这两方面的技术在思想上是相通的，都是从现实世界获取一些有用的信息，减轻人工交互建模的强度和复杂度。可以建立一个 BTF 样本的数据库，当用户通过图像建模得到一个物体的模型后，用 BTF 合成算法给物体表面加上一定的材质和光照效果。另外一个思路是构造一个系统，能同时采集物体的几何信息和表面在不同光照和视角下的反射特性。

另外一个非常有意思的研究方向是如何获取物体的动态信息。目前虚拟现实环境中物体的运动一般是通过轨迹规划实现的，因为轨迹规划需要建立物理模型，对于复杂的运动，就非常难于建模。如人在自然状态下的运动，用传统的轨迹规划方法就无法逼真模拟。文献<sup>[94]</sup>在这方面做了一些探索，这个系统通过分析传感器和摄像机获得的数据，得到人体运动时关节和表面皮肤的动态特性。如何通过摄像机捕捉真实物体的运动，实现动态特征方面的真实感建模是非常有挑战性和实用性的课题。还有一个有趣的问题是如何用一个物体的动态特征，驱动一个静态模型。用这样的技术可以实现虚拟演员，这会在影视制作方面有很好的应用前景。

## 参考文献

- [1] N.Alberto Borghese. A portable modular system for automatic acquisition of 3-D object. IEEE Transactions On Instrumentation And Measure, Vol 49, No5, October 2000.
- [2] R. Zhang, P. Tsai, J.E. Cryer, M. Shah. Analysis of shape from shading techniques. IEEE CVPR94, page 377-384.
- [3] Xinguo Liu, Yizhou yu, Heung-Yeung Shum. Synthesizing bidirectional texture functures for real-world surface. Proceedings of Siggraph 2000, page 479-488, July 2000.
- [4] 余飞鸿, 结构光三维物体面形实时测量系统, 光子学报, 23 卷, 6 期, 1994 年 12 月。
- [5] 冯华军等, 主动式三维成像技术中的光学系统设计, 光学仪器, 21 卷, 2 期, 1999 年 4 月。
- [6] 张广军等, 结构光三维视觉系统研究, 航空学报, 20 卷, 4 期, 1999 年 7 月。
- [7] 张广军等, 网条结构光三维视觉检测标定方法研究, 仪器仪表学报, 21 卷, 3 期, 2000 年 6 月。
- [8] Zicheng Liu, Zhengyou Zhang, Chuck Jacobs, Michael Cohen. Rapid modeling of animated faces from video. Technical Report MSR-TR-99-21.
- [9] W. K. Leow, Z. Huang, Y. Zhang, and R. Setiono. Rapid 3D model acquisition from images of small objects. In Proc. Geometric Modeling & Processing, 33-41, 2000.
- [10] 贾瑞玉, 汪炳权, 含病变肝脏 C T 图像三角形表面重构的实现, 微机发展, 1999 年第 2 期。
- [11] Gerald E. Farin. Curves and surface for computer aided design: A practical guide. Academic Press, 1990.
- [12] Maria-Elena Algorri. Surface reconstruction from unstructured 3d data. Computer Graphics forum Vol 15, NO.1. 47~60,1996.
- [13] R. Pfeifle, H-P. Seidel. Spherical triangular B-splines with application to data fitting. Computer Graphcs Forum, 14(3).89~96, 1995.
- [14] William E.Lorensen. Marching cubes:a high resolution 3D surface construction algorithm. Computer Graphics, V21, No 4, July 1987.
- [15] Eric Bittar. Automatic reconstruction of unstructured 3D data: combining a medial axis and implicit surface. EUROGRAPHICS '95 Vol 14, Number 3, 1994.
- [16] Robert Mencl. A Graph-based approach to surface reconstruction. Eurographcs 95, Vol 14, Number 3, 1995.

- [17] Catmull Ed. Computer display of curved surfaces. Proc. IEEE Conference on Computer Graphics, Pattern Recognition and Data Structure, 1975, 11-17.
- [18] Blinn J, Newell M. Texture and Reflection in Computer Generated Images. CACM 19, 1976, 542-547.
- [19] Blinn J. Simulation of wrinkled surfaces. Computer Graphics, 12, 1978, 286-292.
- [20] Paul Heckbert. Survey of texture mapping. IEEE Computer Graphics and Applications, Nov. 1986.
- [21] E.Catmull. Subdivision Algorithm for Computer Display of Curved Surfaces. Ph.D Thesis, Dept. of Computer Science, Univ. of Utah, Salt Lake City, Utah, 1974.
- [22] J.F.Blinn. A scan line algorithm for displaying parametrically defined surfaces. Computer Graphics, Vol.12, No.3, 1978.
- [23] E.A.Bier, K.R.Sloan. Tow-Part texture mappings. IEEE Computer Graphics Applications, Vol.6, No.9, 1986.
- [24] S.Chen et al. View interpolation for image synthesis. Siggraph 93, 1993.
- [25] M.Levoy and P.Hanrahan. Light field rendering. Siggraph 96, 1996.
- [26] P.E. Debevec, C.J.Taylor, and J.Malik. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach. Siggraph 96, 1996.
- [27] P.E. Debevec, Y. Yu and G.D. Borshukov. Efficient View-Dependent Image-Based Rendering with Projective Texture-Mapping. Eurographics Rendering Workshop 1998, 1998.
- [28] S.Chen. Quick Time VR-an image based approach to virtual environment navigation. Siggraph 95, 1995.
- [29] 李宏东等, 基于图象的真实感绘制与建模 - IBMR, 计算机应用研究, 1999 年第六期。
- [30] Daniel Wood, Daniel Azuma, Wyvern Aldinger, Brian Curlless, Tom Duchamp, David Salesin, Werner Stuetzle. Surface light fields for 3D photography. Siggraph 2000, 2000.
- [31] Wei-Chao Chen, Jean-Yves Bouguet, Michael Chu and Radek Grzeszczuk. Light field mapping: efficient representation and hardware rendering of surface light Field. Siggraph 2002, 2002.
- [32] 张景瑞, 计算机立体视觉标定算法及其在牙科手术仿真中的应用, 硕士学位论文, 清华大学计算机科学与技术系, 2001。
- [33] 邱茂林, 马颂德, 李毅, 计算机视觉中摄像机定标综述, 自动化学报, 第 26 卷, 第 1 期, 2000 年 1 月。

- [34] Faig W. Calibration of close-range photogrammetric systems: mathematical formulation. *Photogrammetric eng. Remote sensing*, 41(12):1479-1486, 1975.
- [35] Abdel-Aziz Y I, Karara H M. Directlinear transformation into objectspace coordinates in close-range photogrammetry. *Proc. Symp. Close-Range Photogrammetry*, 1-18, 1971.
- [36] Tsai R Y. An efficient and accurate camera calibration technique for 3D machine vision. *Proc. CVPR 86*, 364-374, 1986.
- [37] Z. Y. Zhang. A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, 1998.
- [38] Intel Corp. Open Source Computer Vision Library Reference Manual. <http://developer.intel.com>.
- [39] J. More. The levenberg-marquardt algorithm, implementation and theory, In G.A. Watson, editor, *Numerical Analysis, Lecture Notes in Mathematics 630*, Springer-Verlag, 1977.
- [40] 徐光祐, 《计算机视觉》, 清华大学计算机系 1999 年。
- [41] 李锦成, 袁宗保, 一种用于三维立体视觉的精密摄像机定标算法, *计算机学报*, 1990, 14(3): 194-201。
- [42] J. Salvi, X. Armangue and J. Pages. A survey addressing the fundamental matrix estimation problem. *IEEE International Conference on Image Processing v 2 Oct 7-10 2001*, page 209-212.
- [43] Andrew W. Fitzgibbon, R.B. Fisher. A Buyer's Guide to Conic Fitting, *Proc. 5th British Machine Vision Conference*, Birmingham, pp. 513-522, 1995.
- [44] M. Rioux. Laser Range Finder Based on Synchronized Scanners. *Applied Optics*, Vol. 23, No. 21, pp. 3837-3844, 1984.
- [45] K. Dana, B. van Ginneken, S. Nayar, and J. Koenderink. Reflectance and texture of real-world surfaces. *IEEE Conf. on CVPR*, p. 151, 1997.
- [46] Paul Debevec. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. *Proceedings of Siggraph 98, Computer Graphics Proceedings, Annual Conference Series*, pp. 189-198, 1998.
- [47] Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. A practical model for subsurface light transport. *Computer Graphics Proceedings, Annual Conference series*, pp. 511-518, Siggraph, 2001.
- [48] B. Guenter, C. Grimm, D. Wood, H. Malvar, and F. Pighin. Making faces. *Computer Graphics, Annual Conference Series*, pages 55--66. Siggraph, July 1998.
- [49] 张永相, 刘章文, 结构光图象的轮廓提取, *四川工业学院学报*, 第 19 卷, 第 2 期,

2000 年。

- [50] 卞学海, 具有纹理真实感的曲面重构算法及其系统实现技术的研究, 硕士学位论文, 清华大学计算机科学与技术系, 2002.
- [51] L. Wang, S. Kang, H. Shum, and R. Szeliski. Optimal texture map reconstruction from multiple views. Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR-2001), 2001.
- [52] R. Haralick. Statistical image texture analysis. In Handbook of Pattern Recognition and Image Processing, volume 86, pages 247–279. Academic Press, 1986.
- [53] H. Iversen and T. Lonnestad. An evaluation of stochastic models for analysis and synthesis of gray scale texture. Pattern Recognition Letters, 15:575–585, 1994.
- [54] Li-Yi Wei and Marc Levoy. Fast texture synthesis using tree-structured vector quantization. Proceedings of SIGGRAPH 2000, pages 479–488, July 2000.
- [55] S. Zhu, Y. Wu, and D. Mumford. Filters, random fields and maximum entropy (FRAME) - towards a unified theory for texture modeling. International Journal of Computer Vision, 27(2):107–126, 1998.
- [56] R. Paget and I. Longstaff. Texture synthesis via a noncausal nonparametric multiscale Markov random field. IEEE Transactions on Image Processing, 7(6):925–931, June 1998.
- [57] Greg Turk. Texture synthesis on surfaces. Proceedings of Siggraph 2001, pages 347–354, August 2001.
- [58] Li-Yi Wei and Marc Levoy. Texture synthesis over arbitrary manifold surfaces. Proceedings of Siggraph 2001, pages 355–360, August 2001.
- [59] 彭群生, 鲍虎军, 金小刚, 计算机真实感图形的算法基础, 科学出版社, 1999.
- [60] J.F.Blinn. Models of light reflection for computer synthesized pictures. Computer Graphics, Vol. 11, No.2, 1977.
- [61] R.L.Cook, K.E.Torrance. A reflection model for computer graphics. Computer Graphics, Vol.15, No.3, 1981.
- [62] J.J. Koenderink, A.J. van Doorn and M. Stavridi. Bidirectional reflection distribution function expressed in terms of surface scattering modes. European Conference on Computer Vision, pp. 28-39, 1996.
- [63] Kristin J. Dana, Bram van Ginneken, Shree K. Nayar, and Jan J. Koenderink. Reflectance and texture of real-world surfaces. ACM Transactions on Graphics, 18(1):1–34, January 1999.
- [64] Kristin J. Dana, Shree K. Nayar. Histogram model of 3D textures. IEEE

- Conference on Computer Vision and Pattern Recognition (CVPR), pp. 618-24, Santa Barbara, CA, June 1998.
- [65] Kristin J. Dana, Shree K. Nayar. Correlation Model for 3D Texture. Proceedings of Seventh IEEE International Conference of Computer Vision (ICCV), IEEE Computer Society, vol. 2, pp. 1061-66, Corfu, Greece, September 1999.
- [66] Thomas Leung and Jitendra Malik. Representing and recognizing the visual appearance of materials using 3d textures. International Journal of Computer Vision, 43(1):29–44, June 2001.
- [67] Xinguo Liu, Yachua Hu, Jingdan Zhang, Xin Tong, Baining Guo, Heung-Yeung Shum. Synthesis and rendering of bidirectional texture functions on arbitrary surfaces. IEEE TVCG, 2003, accepted.
- [68] Emil Praun, Adam Finkelstein, and Hugues Hoppe. Lapped textures. Proceedings of SIGGRAPH 2000, pages 465–470, July 2000.
- [69] Sunil Arya, David Mount, Nathan Netanyahu, Ruth Silverman, and Angela Wu. An optimal algorithm for approximate nearest neighbor searching. Journal of the ACM, 45:891–923, 1998.
- [70] Xin Tong, Jingdan Zhang, Ligang Liu, XiWang, Baining Guo, and Heung-Yeung Shum, Synthesis of bidirectional texture functions on arbitrary surfaces. ACM Transactions on Graphics, vol. 21, no. 3, pp. 665–672, July 2002, (Proceedings of ACM SIGGRAPH 2002).
- [71] Lexing Ying, Aaron Hertzmann, Henning Biermann, and Denis Zorin. Texture and shape synthesis on surfaces. Proceedings of 12th Eurographics Workshop on Rendering, pages 301–312, June 2001.
- [72] Michael Ashikhmin. Synthesizing natural textures. 2001 ACM Symposium on Interactive 3D Graphics, pages 217–226, March 2001.
- [73] Hugues Hoppe. Surface reconstruction from unorganized points. Siggraph 92 Chicago , 26-31 , 1992.
- [74] Nunzio Alberto Borghese. Autoscan: A flexible and portable 3D scanner. IEEE Computer Graphics and Applications June 1998.
- [75] Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., Fulk, D. The Digital Michelangelo Project: 3D scanning of large statues. Computer Graphics Proceedings, Annual Conference Series, 2000, ACM SIGGRAPH, pp. 131-144.
- [76] FARO Technologies Inc., <http://www.faro.com/Products/Products.asp>.
- [77] ROMER CimCore Inc., <http://www.romer.com/main/products.php>.

- [78] Immersion Corporation, <http://www.immersion.com/products/3d/>.
- [79] 3D Scanners Ltd, <http://www.3dscanners.com/>.
- [80] Cyberware Inc., <http://www.cyberware.com/products/index.html>
- [81] 3rdTech, Inc., <http://www.3rdtech.com/DeltaSphere.htm>
- [82] Rusinkiewicz, S., Hall-Holt, O., Levoy, M. Real-Time 3D Model Acquisition. ACM Transactions on Graphics (Proc. SIGGRAPH 2002), Vol. 21, No. 3, July, 2002, pp. 438-446.
- [83] Matusik, Wojciech, Chris Buehler, Ramesh Raskar, Steven J. Gortler, and Leonard McMillan. Image-Based Visual Hulls, Proceedings of ACM SIGGRAPH 2000, pp. 369-374.
- [84] Wojciech Matusik, Hanspeter Pfister, Matt Brand, Leonard McMillan. A Data-Driven Reflectance Model. Proceedings of ACM SIGGRAPH 2003, accepted.
- [85] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In SIGGRAPH 2000, pp. 145--156, 2000.
- [86] Leica Geosystems AG. <http://www.leica-geosystems.com/products/disto-laser-distance-meter/index.htm>.
- [87] Acuity Research, Inc. <http://www.acuityresearch.com>.
- [88] Senix Corporation. <http://www.senix.com/>.
- [89] Cyra Technologies, Inc. <http://www.cyra.com/products/products.html>.
- [90] Peter-Pike Sloan, Xinguo Liu, Heung-Yeung Shum, John Snyder. Bi-Scale Radiance Transfer. Proceedings of ACM SIGGRAPH 2003, accepted.
- [91] HOPPE, H. Progressive meshes. SIGGRAPH 1996, pp. 99-108.
- [92] COHEN, J., OLANO, M., MANOCHA, D. Appearance-preserving simplification. SIGGRAPH 1998, pp. 115-122.
- [93] M. Garland and P. Heckbert. Simplifying Surfaces with Color and Texture using Quadric Error Metrics. In Proceedings of IEEE Visualization 98.
- [94] Peter Sand, Leonard McMillan, Jovan Popovic. Continuous Capture of Skin Deformation. Proceedings of ACM SIGGRAPH 2003, accepted.

## 致 谢

三年的研究生生活很快就要结束了，在这期间，我得到了很多老师、同事和同学的关心、支持和帮助，可以说我的论文同样凝结着他们的热情和汗水。

由衷感谢导师孙富春副教授和邓志东教授对本人的悉心指导，他们的言传身教将使我终生受益。

在微软亚洲研究院的合作研究期间，得到郭百宁高级研究员和童欣研究员的精心指导与帮助，不胜感激。

只言片语难以尽数，我在这里向所有关心过我，帮助过我的人表示感谢。

---

## 声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：\_\_\_\_\_日 期：\_\_\_\_\_

## 个人简历

1996.9-2000.7 在清华大学计算机科学与技术系学习，并于 2000 年 7 月获工学学士学位。

2000 年 9 月在清华大学计算机科学与技术系计算机应用技术专业攻读硕士学位。

## 攻读硕士学位期间发表的学术论文

### A：在学术会议上发表的论文

1. 张竞丹, 邓志东, 郭百宁. 一种基于两步分割的非监督彩色图像分割方法. 第四届中国计算机图形学大会 2002, pp. 144-148.
2. Jingdan Zhang, Yongmei Wang, Baining Guo. Pyramidal Search of Maximum Coherence Direction for Biomedical Image Interpolation. IEEE International Symposium on Biomedical Imaging, pages 887-890, Washington D. C., July 2002.
3. Ke Deng, Jingdan Zhang, Lifeng Wang, Baining Guo. Texture Mapping with a Jacobian-Based Spatially-Variant Filter. IEEE Pacific Graphics, October 2002.

### B：在国外核心学术刊物上发表的论文

1. Jingdan Zhang, Kun Zhou, Luiz Velho, Baining Guo, Heung-Yeung Shum. Synthesis of Progressively-Variant Textures on Arbitrary Surfaces. ACM Transactions on Graphics(Proc. ACM SIGGRAPH) 2003, accepted.(已接受)
2. Xin Tong, Jingdan Zhang, Ligang Liu, Xi Wang, Baining Guo, Heung-Yeung Shum. Synthesis of Bidirectional Texture Functions on Arbitrary Surfaces. ACM Transactions on Graphics(Proc. ACM SIGGRAPH), Vol 21:3, pages 665-672, San Antonio, Texas, July 2002.
3. Xinguo Liu, Yaohua Hu, Jingdan Zhang, Xin Tong, Baining Guo, Heung-Yeung Shum. Synthesis and Rendering of Bidirectional Texture Functions on Arbitrary

Surfaces. IEEE Transactions on Visualization and Computer Graphics (TVCG), 2003, accepted. (已接受)

4. Yongmei Wang, Jingdan Zhang, Zhunping Zhang, Baining Guo. Directional Coherence Interpolation For Three-Dimensional Gray-Level Images. IEEE Transactions on Medical Imaging (TMI), 2003, submitted. (已提交)