



单目视觉城市建筑物 参数化三维建模

Monocular Vision Parametric
3D Modeling of Architectures

杨淼 著



中国水利水电出版社
www.waterpub.com.cn

单目视觉城市建筑物 参数化三维建模

Monocular Vision Parametric
3D Modeling of Architectures



著



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本专著是一部关于单目视觉建筑物三维重建方法的研究成果的系统专著,结合作者对图像处理、分析和三维重建等进行的研究,对从数码相机拍摄的建筑物图像中以参数化建模的方法恢复建筑物的三维几何结构进行了论述和探讨。本书内容完整、理论与实践相结合、阐述清晰、图文并茂,可供从事信号与信息处理、通信工程、信息工程、计算机科学与技术、电子科学与技术等领域的大学教师和科研工作者、研究生、大学本科高年级学生及工程技术人员参考使用。

图书在版编目(CIP)数据

单目视觉城市建筑物参数化三维建模 / 杨森著. --
北京:中国水利水电出版社, 2014. 12
ISBN 978-7-5170-2799-7

I. ①单… II. ①杨… III. ①三维—模型(建筑)—
计算机辅助设计 IV. ①TU205

中国版本图书馆CIP数据核字(2014)第305424号

书 名	单目视觉城市建筑物参数化三维建模
作 者	杨森 著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座 100038) 网址: www.waterpub.com.cn E-mail: sales@waterpub.com.cn 电话: (010) 68367658 (发行部)
经 售	北京科水图书销售中心(零售) 电话: (010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	中国水利水电出版社微机排版中心
印 刷	三河市鑫金马印装有限公司
规 格	170mm×230mm 16开本 8.25印张 148千字
版 次	2014年12月第1版 2014年12月第1次印刷
印 数	001—500册
定 价	32.00元

凡购买我社图书,如有缺页、倒页、脱页的,本社发行部负责调换

版权所有·侵权必究



前 言

P R E F A C E

随着现代计算机技术的飞速发展,计算机视觉技术越来越广泛深入地应用于工业、国防、医学、影视业、广告、旅游、经济、考古等各个领域。人们经常需要能简单、迅速地获得物体表面的三维信息并将其转变成计算机能直接处理的数据,工业界要求能快速地测量物体表面的三维坐标,影视界需要将演员道具等的立体模型输入计算机才能进行三维动画特技处理,游戏娱乐业需要在虚拟场景中放置逼真的三维模型,旅游业的发展需要能够让游客们无地域限制地随时身临其境般漫游于逼真的景点场景中,电子商务业需要更加吸引人、更加逼真的商品展示,科技工作者在地质、海底探测和三维物体识别的研究中需要快速获得大量的三维数据,所有这些都面临如何在计算机中展示三维现实世界的问题。

作为各种虚拟现实(VR, Virtual Reality)中的重要组成部分,如何由图像推导出物体的几何结构是近些年来计算机视觉研究的热点之一,其中城市景观三维可视化研究中,作为城市景观主要组成部分的三维建筑物重建是其中最为重要的内容。完成建筑物的三维重建需要从图像中获取建筑物的几何信息和相机的方位信息。由于拍摄的建筑物图像情况各不相同,而外形各异的建筑物的组成方式也不相同,因此,对建筑物的三维重建采用模型来表达建筑物的方法显然是十分复杂和困难的。

本书在分析了目前国内外已有研究成果的基础上,进行了多方面的深入研究和实验,主要研究基于单幅图像的三维参数化建模方法。全书共分7章,内容包括面向对象的点、面、体分层建筑物建

模方法，基于尺度分析建筑物图像分割方法，形态学尺度空间（形态筛）的应用，基于特征的建筑物自动分类，轮廓的自动提取等方面的算法，对平顶和非平顶城市建筑物分别介绍了实现三维参数化建模的具体方法。本书最后介绍了 Web 页面中 VRML 技术建筑物场景三维建模和漫游的应用，为各种应用环境提供有效的建模方法和三维模型，便于浏览，可移植性强。

本书的出版得到中国海洋大学软件工程所的老师 and 同学的大力支持，在此特别感谢中国海洋大学信息科学与工程学院魏志强教授对本书工作的指导。由于作者水平有限，书中难免存在不足，请各位专家、学者批评指正，共同推进我国计算机视觉研究的不断深入。

著者

2014 年 10 月

目 录

前言

第 1 章 绪论	1
1.1 研究背景和意义	1
1.2 国内外研究现状	4
1.3 本书的研究目标与主要内容	10
第 2 章 面向对象城市建筑物参数化模型研究	13
2.1 城市景观实体的分类	13
2.2 城市建筑物参数化模型	14
2.3 面向对象的建筑物建模	19
2.4 本章小结	25
第 3 章 三维重建图像预处理研究	27
3.1 图像分割	27
3.2 尺度空间方法	28
3.3 基于分水岭变换和区域融合的彩色建筑物图像分割算法	28
3.4 基于图像分析的彩色图像分割算法	35
3.5 形态学图像处理	38
3.6 本章小结	50
第 4 章 城市建筑物模型的自动分类算法研究	51
4.1 城市建筑物世界场景约束	52
4.2 城市建筑物类型判断策略	52
4.3 建筑物角点检测	53
4.4 建筑物模型的判断	56

4.5	实验	58
4.6	本章小结	60
第5章	基于单幅图像的平顶建筑物三维自动重建算法	61
5.1	建筑物轮廓提取技术	61
5.2	建筑物三维重建技术	62
5.3	投影几何与预处理	64
5.4	概率统计模型	67
5.5	轮廓提取策略	69
5.6	平顶建筑物场景参数估计	70
5.7	实验	72
5.8	本章小结	78
第6章	基于单幅图像的非平顶建筑物三维重建算法	79
6.1	单幅图像建模技术	79
6.2	消隐点几何原理	81
6.3	消隐点的计算	83
6.4	非平顶建筑物未知场景的参数估计	84
6.5	提取屋顶角点	85
6.6	求解相对深度	87
6.7	其余参数估计	88
6.8	实验	89
6.9	本章小结	92
第7章	城市建筑物三维建模应用实例	93
7.1	系统框架结构	93
7.2	图像纹理纠正算法	95
7.3	城市建筑物建模应用实现	99
7.4	本章小结	104
附录1	矩阵运算	105
附1.1	SVD分解	105
附1.2	Cholesky分解	105

附录 2 直接线性变换 (DLT) 算法	106
附 2.1 DLT 变换	106
附 2.2 归一化 DLT 变换	107
附录 3 专著中的英文缩写	109
参考文献	111
后记	122

第1章 绪 论

视觉是人类观察世界、认知世界的重要手段，人类从外界获得的信息约有75%来自视觉系统。人类视觉过程可以看做是一个复杂的从感觉到知觉的过程，最终目的从狭义上说是要对场景作出对观察者有意义的解释和描述，从广义上说是基于这些解释和描述并根据周围环境和观察者的意愿制定出行为规划。视觉是各个应用领域，如制造业、检验、文档分析、医疗诊断和军事等领域中各种智能/自主系统中不可分割的一部分。由于它的重要性，一些西方国家，例如美国把对计算机视觉的研究列为对经济和科学有广泛影响的科学和工程中的重大基本问题。

计算机视觉是各种成像系统代替视觉器官作为输入手段，由计算机来代替大脑完成处理和解释的科学和技术。计算机视觉研究的首要目标就是使计算机通过图像创建或恢复现实世界模型，然后认知现实世界。这里主要有两类方法：一类是仿生学的方法，参照人类视觉系统的结构原理建立相应的处理模块完成类似的功能和工作；另一类是工程的方法，从分析人类视觉过程的功能着手并不去刻意模仿人类视觉系统内部结构而仅考虑系统的输入和输出，并采用任何现有的可行手段实现系统功能。

计算机视觉是一门综合性的学科，它正吸引来自各个学科的研究者参加到对它的研究之中，其中包括：计算机科学、信号处理、物理学、应用数学和统计学、生物医学、生理学和认知科学等。它的发展得益于神经、生理学、心理学与认知科学对生物系统的研究，虽然由于人脑的高度复杂性对这种跨学科的研究还远远不够深入，但从事计算机视觉的研究者们已经发展起一套独立的计算理论和算法，从而对视觉信息进行分析和处理。

1.1 研究背景和意义

1.1.1 研究背景

对一个实体进行三维建模，存在着许多种方法。传统的方法是物体的三维



模型完全由人工使用几何建模软件生成,做出复杂和精确的物体模型需要有经验的建模人员花费很长的时间。现在有时也使用三维激光扫描仪对实体进行数字化处理,优点是精度高、速度快,缺点是激光扫描仪成本比较高,而且恢复出的是数字点云图,模型的数据量比较大。因此人们希望有不需专门设备和较高技术要求的简单快速的物体建模方法,基于图像的造型和绘制方法^[1](IMBR, Image-Based Modeling and Rendering)有造型过程简便、绘制速度快的优点,逐渐成为研究的热点。但现有的基于图像的建筑物三维建模系统多数都是基于2D GIS 或者是由航拍图像序列获得重建的方法,只适用于某些特定的应用,并且多数都需要人工交互选取部分特征或需要有关于模型的先验知识,因此有必要对基于陆地图像的建筑物三维自动参数化重建的相关问题进行研究。

另外,当我们无法得到目标的带有重叠度的图像,比如在拍摄时受到相机载体和目标地理位置等条件的限制无法得到目标各个角度的重叠图像,或者图像的历史性决定了不能得到它的卫星图像和激光测距结果,很可能只存有一幅照片,这时只能依靠单幅图像进行重建。从理论上讲,如果没有其他任何的先验知识,从单幅图像是无法得到物体的三维模型的。但是如果充分利用人对城市场景理解的先验知识,利用一幅图像也是可以恢复建筑物目标的三维模型的。

利用单幅图像进行三维重建并不是一个新的研究领域,以往这方面的研究都是基于使用光照、纹理、焦距以及其他类似的属性,这些方法多数都要求提供已知场景中物体的一些信息,需要做大量的控制工作,最后所得到的模型都存在一定的模糊度,而且很多方法实现起来难度较大。近几年,国内外学者使用基于射影几何的方法进行单幅图像重建方面的研究,这种方法利用了图像中的透视变形等信息,从存在这种变形的成像中得到物体的模型。这种方法处理图像中点、线、面之间的几何关系,适合几何体和代数平面的重建,因为在这些对象中更容易获得这些关系。

本书的主要内容是介绍数字城市中,数码相机拍摄的建筑物陆地图像的参数化三维建模技术。

在由图像获得建筑物三维模型的过程中,一般包含图像的预处理、建筑物特征选取、相机标定、纹理图像的校正、建筑物的三维建模等步骤。首先,图像分割是基于图像的各种三维重建方法中重要的预处理步骤,图像分割效果的好坏直接影响后续的处理。当前存在的图像分割算法适用范围不同,而且容易受到噪声的影响。我们需要一种适用于由建筑物图像完成重建过程中的有效的



分割方法,能够更利于获取关于建筑物的先验知识和纹理。其次,建筑物几何结构信息或相机的内、外参数是完成3D建筑物模型重构的必要条件,它可以为由一幅或多幅图像实现重建的各种方法提供更多的约束和场景的先验知识。在实现建筑物轮廓提取的各种方法中更多的是针对航拍、遥感图像进行处理,利用边缘检测和霍夫变换(Hough Transform)后经过复杂的后处理和各种判断的建筑物轮廓提取。而对于地面图像中建筑物的主体轮廓,由于图像背景复杂,遮挡和成像过程中引起的形变,各种边缘检测算法无法区分建筑物边缘还是其他边缘,很难实现自动提取。在研究地面图像建筑物提取的过程中,可利用建筑物的几何结构与相机的参数之间存在的联系同时实现相机参数的获取,这将是一种前景广阔的方法。另外,关于三维城市模型中建筑物实体的参数化建模方法,针对多种建筑物模型表达的特点,为城市建筑物建立面向对象的参数化模型和如何对建筑物模型进行自动分类,采用不同的参数提取方法来表达相应的建筑物等问题也有待于进一步研究。

1.1.2 研究意义

虚拟现实技术的应用极为广泛。在1993年,Helsel与Doheriy对全世界范围内已经进行的805项虚拟现实研究项目做了统计,结果表明当时在娱乐、教育及艺术方面的应用占据主流,达21.4%,其次是军事与航空方面,达12.7%,医学方面达6.13%,机器人方面达6.21%,商业方面达4.96%,另外在可视化计算、制造业等方面也有相当的比重。

从上述虚拟现实的部分应用可以看出,经过近十几年的发展,虚拟现实已经渗入到人们的日常生活中,人们可以在虚拟世界中聊天、购物、逛街、旅游及工作,如同在现实世界中一样。而且,这项技术还有更广泛的应用前景,通过与互联网技术结合,可以构造出一个更加完美的虚拟世界^[2]。

建立虚拟城市的目的是借助于计算机虚拟现实技术创造出一种崭新的人机交互手段,为用户提供模拟城市的操作环境,使其具有仿佛置身于现实城市中的临境感,形成一个逼真的,具有视觉、听觉、触觉的感官世界,同时,还可以通过感觉、语言、手势等方式进行交互式操作^[3]。用户不仅可以沉浸在虚拟城市中,而且可以进行查询分析,评价,规划和决策。在国内,虽然近年来虚拟城市的研究逐渐受到重视,但就整个世界范围来看目前虚拟城市的研究仍处于探索阶段。

基于图像的建筑物三维建模作为虚拟城市中的重要内容有许多实际应用价值,首先,三维模型比二维图像有更强的表现力,人们可以从三维立体模型获



得对目标更具体的认识,这对于历史景观的真实再现、城市规划中的模型表达、城市公共设施的管理、建筑设计的建筑物景观模拟、军事国防、旅游、游戏、电子商务等领域都有着重要的意义。其次,使用基于图像三维重建生成的模型要比由建模人员使用几何建模软件根据目测、直觉做出的模型更符合原场景,更简单。比如,在建立一个建筑物的三维模型时,通常建筑物的设计图纸这些原始资料不容易找到,而且有时也无法做到去实地测量来得到建筑物的精确数据。使用基于视频图像的三维重建技术仅需要使用照相机拍摄几张照片,然后就可以恢复出该建筑物的三维模型,得到的模型与真实建筑物的大小成比例,如果知道模型上两点间的距离对应的真实世界中的长度,就可以计算出模型上任意两点间的真实长度。

本课题研究的基于图像的城市建筑物三维自动重建参数化建模问题与那些基于2D GIS或者是由航拍图像获得城市模型的方法相比更具有普遍性,因此对推动自动建筑物参数化建模具有重要的理论意义。另外,模型建模法与恢复某视点下的三维效果的三维重建方法相比更便于网上浏览,易于操作,可移植性强。而其中研究的图像预处理算法、相机标定算法、建筑物模型自动分类法、建筑物轮廓自动提取法、建筑物三维重建方法等对推动相关领域的理论研究也具有广泛的价值。

1.2 国内外研究现状

1.2.1 基于图像的三维重建方法研究现状

以三维重建为研究目标的视觉研究方法,在三十多年的发展中,逐渐形成了各具特色的方法和理论^[4]。例如根据视点数目的不同,可分为单目视觉、双目视觉、三目视觉或多目视觉。根据获取数据的方式可分为主动获取方式和被动获取方式。前者需要人为地在物体表面做某种记号,例如条纹或者阴影,并且根据这些信息建模或由传感器获得三维信息。而后者是指利用场景中物体本身所具有的一些特征进行建模,如某些几何特征和纹理特征。被动获取的建模方式有效利用了图像中物体本身的特征信息,从而减少了获取图像前的准备工作,因此基于图像的建模通常采用被动建模的方法。

1.2.1.1 单目视觉

一般来说,用单个像机的单目视觉,即从单一图像中不可能直接得到三维信息,但是,如果已知对象的形状和性质,或作某些假设,便能够从图像的二



维特征推导出三维信息。作为图像的二维特征（统称为 X），如利用明暗度（Shading）^[5]、纹理（Texture）、调焦法（focusing）、梯度法（gradient）、轮廓法、阴影^[6, 7]等特征提取三维信息的有关理论方法，称之为由 X 恢复形状（Shape From X），基于单幅图像的重建方法都不能测定绝对距离。这些由 X 恢复形状的方法所得到的模型都存在一定的不足，而且很多方法都要求知道场景中物体的一些信息，实现起来难度较大。近几年国内外一些学者使用射影几何学进行单幅图像重建的研究，利用图像的透视变形等信息，从存在这种变形的像中得到物体模型。这种方法处理图像中的点、线、面之间的几何关系，在标定中使用几何约束，增加了结果的稳定性，具有较强的实用性。

还有一些其他的单目视觉重建方法，如基于建筑物阴影长度的重建方法，是一种通过测量建筑物阴影长度来计算建筑物高程信息的方法^[8]，并由此生成研究区建筑物的高度图像。这种方法主要应用于高分辨率卫星影像中的地震灾害识别与评估，对高层建筑物的识别精度较好，对于低层建筑物识别较差，原因在于：一是建筑物阴影影响；二是高程信息插值处理过程中产生的误差影响^[9]。还有基于图像灰度恢复物体形状的重建方法，假设只有一个已知方向的点光源，只有一幅由该光源照射下得到的图像。由单幅图像上每一点灰度值可得到一个约束条件。由灰度恢复三维物体形状的方法与其他方法比较，局限性较大：一般都需要假设光源方向已知的单个点光源，不考虑物体间的互反射，而且已知物体的反射系数等，另外，相比图像其他特性（如边缘与角点），图像灰度受噪声影响更大。

1.2.1.2 立体视觉

立体视觉也称双目视觉（Stereo Vision），是计算机被动测距方法中最为重要的距离感知方法之一。它模拟人类视觉处理景物的方式，可以在多种条件下灵活地测量景物的立体信息。立体视觉三维重建方法的一般过程为使用两个或单个摄像机对同一目标拍摄两幅或多幅图像组成立体像对，根据三角测量原理利用对应点的视差（disparity）来计算视野范围内的立体信息。主要步骤包括：图像获取、摄像机标定、特征提取、图像匹配和三维重建。适用于移动不大的相机拍摄的两幅图像建模。受匹配算法影响，对有重复纹理的模型不合适。

1.2.1.3 运动视觉

运动视觉^[10]研究的是如何从变化场景的一系列不同时刻的图像中提取出有关场景中的物体的形状位置和运动信息，在计算机视觉中，运动视觉也是受到特别重视的一部分，研究运动视觉的方法大体上可以分为两大类，一类是基



于特征的方法^[11, 12]；另一类是基于光流场的方法^[13]。基于特征的方法与立体视觉方法有许多相似之处，它也包括两个主要的步骤：一是从相继两幅或多幅不同时刻的图像，相当于立体视觉中的左右图像中抽取特征如点、线、面等，并建立起对应；二是依据这些特征之间的对应来计算物体的结构、形状、位置等和运动。基于光流场的方法不需要对图像进行处理抽取特征，而是直接对图像本身进行处理，这种方法的优点是它可以求得稠密像素处所对应物体的运动信息，而在基于特征的方法中仅能求得特征的运动信息，其他部分的运动信息要根据这些特征的信息经过一定的插值外推来求得。

1.2.1.4 基于图像序列的三维重建

基于图像序列的三维重建技术也称为多视点三维重建^[14] (Multi-view reconstruction)。同立体视觉也称双目重建 (Two-view reconstruction) 相比较由于利用了多幅图像的信息一方面降低了特征匹配难度，另一方面提高了重建精度。尽管多视点重建可以利用立体视觉中的一些技术^[15]，但更多的是建立新的理论并采用新的解决方法，其中两个比较关键的内容是图像特征对应的获取以及物体形状和运动估计，比如利用三焦点张量约束下的直线匹配完成相机的标定等，但是随着图像序列的增加会导致漂移和误差的传递积累等问题。

1.2.2 建筑物建模方法研究现状

根据文献^[16]的总结，三维建筑物建模的国内外研究综合了地理信息系统、摄影测量学与计算机科学的计算机视觉和计算机图形学三个方面的共同发展。

1.2.2.1 数字摄影测量的典型研究

(1) 从城市航空影像中自动提取建筑物，如检测二维建筑物和数字高程模型 (DEM, Digital Elevation Model) 数据^[17]、知觉组合^[18]、线条分析^[19]、使用阴影、透视几何等辅助信息^[20]、直接对建筑物或表面进行建模^[21]、基于知识的系统^[22]，以及通过影像测量并结合物体的几何知识构建出多面体对象模型的方法^[23]等。

(2) 结合二维地图矢量数据利用航空激光扫描^[24, 25]或激光高度计数据的方法^[26-28]。

(3) 利用三维深度传感器^[29]、多 CCD 相机和彩色高分辨率数字相机获取的数据实现建筑物建模。

(4) 利用虚拟现实技术实现三维 GIS 数据的可视化^[30]。

(5) 其他方法如人机交互下的半自动三维建筑物建模等^[31]。



1.2.2.2 方法分类

在计算机视觉和计算机图形学中,目前方法分为三类:基于几何造型(MBR)、基于图像绘制(IBR)以及几何、图像混合的方法(IMBR)。

(1) 基于几何造型的方法通常利用成熟的造型软件手工搭建模型,比如一种是基于CAD建模,其结果不够精确而且绘制效果也不够逼真。生成复杂的几何模型需要付出巨大的工作量,并且需要使用很多技巧。为了绘制出较好的光照效果,往往需要很昂贵的硬件开销,计算量庞大,难以在普通计算平台上实时处理,图像的真实感也没有保证。

(2) 基于图像绘制的方法,是通过一个来自多视点的原始的或合成的图片库来产生任意视点的新的虚拟图片。对于大规模场景则需要很多图像,使存储和空间深度信息的表达成为瓶颈。另外,系统能提供的场景交互手段有限,对于很多应用领域不能满足要求。

(3) 把几何造型和基于图像绘制结合起来是一种非常有效的方法,具有造型过程简便、绘制速度快的优点,最常见的例子是在几何建模的基础上,进行纹理映射。

Pollefeys 和 Koch 等人提出了由图像序列完成的自动三维重建方法^[32],该系统输入相机拍摄角度之间间隔不超过 $5^{\circ}\sim 10^{\circ}$ 的序列图像,首先由 Harris 角点检测,角点匹配,对极几何约束等预处理确定图像之间的关联,然后由初始的两幅视图完成初步射影重建,逐渐加入一幅幅图像优化并完成射影重建,根据基于绝对二次曲线的方法进行相机自标定,建立深度密集匹配,到最终的3D重建结果。该方法可完成自动的3D重建,但需要一系列高重叠度的图像,最终恢复的3D场景效果并不是具有独立意义的3D模型。

Ildiko Suveg 和 George Vosselman 提出了基于知识系统的由航拍图像序列完成的自动3D重建系统^[33],将立体像对与大尺度2D GIS地图提供的平面信息相结合,将建筑物的重建过程分为假设模型的生成和验证过程,根据2D GIS信息生成建筑物模型库中的各种建筑物模型分解假设,根据模型中线段与图像中线段的对应程序设计评价函数,用以排序各种假设的模型,按照评价函数值的高低从最高的假设开始根据提取的3D角点和线段进行匹配和重建对分解计划中的各个模型参数进行初始估计,由线框模型重投影拟合、精炼参数,模型的验证由重投影信息的匹配程度来最终确定。

Berkeley 大学的 Debevoic、Taylor、Malik 等完成了著名的建筑物重建系统 Façade 系统^[34],可以由构造一系列简单多面体类型从图像序列中恢复目标的3D模型,但是, Façade 系统是一个人工交互系统,而且建筑物的几何结构



不能太复杂,需要人工选择参数化模型,并指出图像中的特征与模型中特征的对应关系,然后由系统计算模型的相关参数以便最符合图像中标记的特征。

Dick 和 Torr 等人继 Façade 系统之后提出了一种由图像序列进行模型建模的方法^[35],文中针对一些经典的建筑物类型,结合了某一样式建筑物的建筑形状规范和纹理先验知识,比如文艺复兴时期的建筑物如果柱子高 15ft,则底部一定分 6.5 层,顶部为 5.5 层;某类型窗户的长宽比;通过学习获得的属于基元的纹理和墙面纹理先验等等。首先由角点、直线特征检测,建立几幅场景视图之间的匹配,完成相机自标定获得场景的稀疏度量重建,通过递归分割估计基本的平面,在此基础上用每个模型外观的统计数据来决定最佳的模型和模型的参数。

Franck Taillandier 提出了一种适用于大的密集城市区域的由 2D 平面地籍图和航拍图像完成的建筑物自动 3D 重建方法^[36],对每个地籍图分割得到的建筑物区域轮廓,推断可能存在的假设平面,由平面求交获得可能的 3D 模型,经过修剪,用极大似技术确定最佳的 3D 模型,根据航拍图像确定的 DEM 的中心相关决定倾斜和屋脊高度两个参数。该方法可对满足对称性结构屋顶的密集城区建筑物进行简单的 3D 重建。

Haala 和 Hahn (University of Stuttgart)^[37]研究了对影像中含有建筑物的区域,通过在与它对应的高程图(激光扫描数据)中选取得到,高程图虽然分辨率较低但是具有很高的可靠性。在这些区域内检测直线段,进行匹配和分组以得到矩形形状,然后建筑物模型便应用于这些形状。该方法对于孤立的建筑物和树木很少的区域有很好的实验结果。最近,Haala 设计了一个针对影像数据、高程数据的半自动城市建筑物建模系统^[38]。建筑物被分为多个矩形形状,针对每一个矩形,多种三维模型被测试。从而建筑物被描述为更为精确的三维盒子的组合。最后还可加入一些细小的纠正和超级结构元素以得到建筑物的表达。

Forstner, Weidner, 和 Lang (University of Bonn) 开发了多个半自动建筑物提取系统^[39, 40]。在最初的方法中,两种参变量建筑物模型应用于由立体视觉计算或激光扫描得到的三维数据,模型参数由 DEM 数据的拟合得到。棱镜模型应用于具有复杂屋顶或相连接的建筑物。在最近的研究方法中,建筑物模型同时应用于两幅或多幅影像,人工选一个模型并应用于影像中的一个建筑物区域。如果被选取的是正确的建筑物模型,则系统会自动计算它的相关参数(建筑物的尺寸、屋檐和屋脊高度等)。实验结果表明该系统可以很好的集成于半自动摄影测量生产链中,但是建筑物的形态降低了该技术应用于不同景观的



能力。

Fuageras 等利用分层重建、自标定等经典方法从图像序列中重建出建筑物。该系统考虑了建筑物的特殊性,利用建筑物上的三维坐标已知点 (Anchor points)、已知角度或平行线等物理信息标定摄影机,并将待重建场景用多面体来表示。该系统主要用于为 CAD-CMA 等提供原始的几何模型。

上述的这些研究主要是以 DEM 数据或由空中的航拍图像序列来进行建筑物、屋顶的 3D 建模,从空中观测到的密集城区环境,具有急剧的地貌变化,大量的纹理贫乏区域以及各种不同形状、大小、外貌的建筑物,不易很好的提取地面三维对象(对于树木密集的环境更是如此),难以对地表面个三维对象进行建模。而利用 2D GIS 系统(平面图,建筑立面图及平面设计图等)实现的三维城市场景建模,在 2D GIS 基础之上由另外添加信息(如房屋高度、墙面纹理、规则屋顶等)来构建三维城市模型^[41],这种方法需要 2D GIS 信息并仅能表达相对规则的建筑物,难以重构复杂的城市景观实体,同时由于缺乏纹理和准确的第三维数据以及空间关系的表达,所构建的模型真实感少且对景观信息表达不足。

目前对基于图像的建筑物三维重建的研究^[42]大多数仍集中于恢复各个点的三维坐标或由单应矩阵恢复各个平面内的度量作为手段进行重建,需要大量的人工交互,对于实现参数化的 3D 模型建模并不多见。而且关于重建过程中所涉及的各步骤中的具体算法也各有不同,缺乏应用于自动建筑物参数化建模的适用性算法。比如在实现建筑物轮廓提取的各种方法中更多的是针对航拍、遥感图像进行处理,利用边缘检测和 Hough 变换后经过复杂的后处理和各种判断的建筑物轮廓提取,缺少结合目标几何信息的可实现地面图像建筑物轮廓自动提取同时确定相机方向的算法。

而且,多数建筑物三维建模软件中建筑物的数据采集和三维模型的建立是在两个独立的系统中完成,往往在最后模型浏览的环境下才能显示出错误之处。人工建模的方法需要凭经验建模,然后在影像的辅助下重建拓扑关系,重构建筑物模型。可想而知,正确模型与仅凭经验建立的模型存在很大的差别,只凭借经验建立模型是不可靠的,并且繁琐的检查工作造成建模效率不高。另外,人工采集加上人工建立拓扑关系、人工编辑的建模思路,使得整个建模过程费时费力,效率低下。如果对有着严密几何形状的三维实体建筑物进行科学的分类,对结构简单的建筑物进行自动提取,对复杂建筑物通过引入相关的结构信息,建立一个集数据采集和三维建模为一体的拓扑数据模型,将会使建筑物建模环节大大简化。