

AN INTRODUCTION TO INTELLIGENT PHOTOGRAMMETRY

智能 摄影测量学导论

季顺平 著



科学出版社

智能摄影测量学导论

An Introduction to Intelligent Photogrammetry

季顺平 著

科学出版社

北京

内 容 简 介

本书系统讲述现代摄影测量学的知识体系、方法及应用。第1章为全书基础,介绍几何成像、传感器技术、摄影平台、图像处理以及最优化方法等。第2、3章分别介绍单目视觉和双目视觉,它们是摄影测量中几何概念的主体部分。第4章阐述运动平台的定位方法,并将传统的光束法平差推广为贝叶斯理论框架下的动态最优化问题。第5章介绍图像匹配,包括用于同名点提取的稀疏匹配和用于立体重建的密集匹配。第6章介绍激光雷达与光学图像的多层次融合。第7章介绍专题图和深度学习,并指出以深度学习为主的人工智能方法将为语义专题图的自动生产提供良好的解决方案。语义专题图的制作是数字摄影测量时代多年来可望不可及的目标,这预示着一个新的时代——智能摄影测量时代的到来。

本书在介绍摄影测量学基本的、必要的知识基础上,同时引入大量的国际最新研究成果,特别是计算机视觉、机器人学以及机器学习的最新进展。本书可作为摄影测量学、遥感及相关专业的高年级本科生、研究生的教材以及相关研究人员的参考阅读材料。

图书在版编目(CIP)数据

智能摄影测量学导论/季顺平著. —北京:科学出版社,2018.3

ISBN 978-7-03-057054-3

I. ①智… II. ①季… III. ①摄影测量学 IV. ①P23

中国版本图书馆 CIP 数据核字(2018)第 056750 号

责任编辑:杨光华 / 责任校对:石娟娟

责任印制:彭超 / 封面设计:苏波

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

武汉市首壹印务有限公司印刷

科学出版社发行 各地新华书店经销

*

开本: B5(720×1000)

2018年3月第 一 版 印张: 10 3/4

2018年3月第一次印刷 字数: 214 000

定价: 39.00 元

(如有印装质量问题,我社负责调换)

前 言

摄影测量学是一门通过光学像片研究被摄物体的位置、形状、大小、特性和相互关系的学科。智能摄影测量学是传统摄影测量在信息和智能时代发展的新阶段,并与计算机视觉、机器人学、人工智能等学科紧密交叉相关。正如书名所言,本书不是一本详细的操作手册,而是一份概略的导引。本书将用简洁的语言,大略描绘关于整个现代摄影测量的蓝图——透视几何、图像处理、多像解析与最优化、误差与统计、多传感器融合、智能与语义提取,以及少量物理基础。在梳理摄影测量脉络的同时,也试图为这门 20 世纪 90 年代已臻成熟的学科,引入一些新的理论基础及活力。

本书将探讨智能时代下的现代摄影测量学。它覆盖各类新型的光学传感器及距离传感器,是航空、航天、低空和地面摄影测量的综合体。回顾昨日,从 18 世纪开始,以数学家和测绘学家高斯为代表的传统地面测量,到 20 世纪初借助飞艇和飞机的航空摄影测量,再到 20 世纪 70 年代由 Landsat 系列、SPOT 系列所开辟的卫星摄影测量与遥感,然后到 2000 年左右蓬勃发展的低空无人机摄影测量,最后到现在受广泛关注的地面移动测图系统,摄影测量已经历了一个轮回。从字面上,这些前缀,即地面、航空、低空和卫星等,反映了载体的变迁;并由此带来新的问题、新的技术及新的发展和进步。同时,从 19 世纪的银版摄影,到 20 世纪的胶片相机,再到 20 世纪末兴起的数码相机和激光测距仪,也是硬件技术带动学科发展的另一个侧面。

相对于传统的航空航天摄影测量,本书所关注的现代摄影测量在以下几个方面进行拓展:相机,除线阵相机、面阵相机外,也关注曲面成像和全景成像,以及对应的传感器模型;多传感器集成,即包含 GPS、IMU、LiDAR、相机在内,考虑更细致的检校手段,以及相机和 LiDAR 的多层次融合;平台轨迹,不再是规则的航带,需要发展更为普适的定位方法,而不仅仅是静态的、全局的光束法平差;三维重建方面,经典方法多年来未曾完全解决密集匹配的挑战,需要新的思路和切入点,等等。最后,从宏观角度上,模拟摄影测量和解析摄影测量时代已经远去,数字摄影测量无法完全概括当代学科的发展,将在本书第 7 章探讨摄影测量的新时代——智能摄影测量时代。

显然,探讨现代摄影测量不可能忽略计算机视觉和机器学习。后者的发展,很

大程度上促进了前者的自动化和智能化程度以及精确性,甚至理论基础的夯实。无论是特征不变量的构建,如尺度不变特征变换 SIFT 及其变种;还是密集匹配,如半全局匹配 SGM 和图割法 Graph Cut;或同时定位与地图构建(SLAM),以及 Smart3D 等商业软件的应用,都已经深刻地改变了整个摄影测量的面貌。席卷而至的深度学习(deep learning),也将为摄影测量的后期工作——地物分类、目标识别、专题图制作等,赋予更高级的自动化水平。这些都推动着摄影测量迈入智能摄影测量时代。本书部分借鉴国际主流计算机视觉教材,一定程度上沿用了计算机视觉的数学语言;同时也保留传统摄影测量的基本术语。期望本书能够对摄影测量领域的学生进一步加深专业理解,更好地阅读计算机视觉和机器学习等相关领域的文献,以及深入理解两者间的联系起到一些帮助。

本书是第一部将计算机视觉、机器学习等学科的相关研究以一种紧密而流畅的方式编织到摄影测量中的现代化摄影测量学教材。本书引入传统摄影测量教材鲜有涉及的概念和方法,如齐次坐标、射影变换的层次、张量、变分、卡尔曼滤波、粒子滤波、相机与激光雷达的融合、深度学习等。这些概念和方法一方面将极大地简化摄影测量相关理论和方法的理解和学习;另一方面对摄影测量的研究领域进行了前所未有的扩展。除了理论和基础的夯实,本书同时涵盖了许多前沿研究,如第 2 章的曲面和全景成像、第 5 章的基于贝叶斯理论的通用运动平台定位方法,第 6 章的光学图像与激光雷达的多层次融合,以及第 7 章的深度学习和语义专题图提取。

本书采用近似口语的行文风格,一改传统中文教科书的刻板印象,以期达到娓娓道来的效果。本书在撰写过程中得到了武汉大学、东京大学、澳大利亚国立大学、中国农业科学院、首都师范大学等单位及相关领域学者的支持和帮助;同时也得到了国家自然科学基金(41471288,61403285)和 973 计划课题(2012CB719902)的资助,在此一并致谢。

限于水平,书中难免有不妥之处,恳请各位读者指正。

作 者
于格里芬湖畔

目 录

第 1 章 基础	1
1.1 物理相关	1
1.1.1 光与视觉	1
1.1.2 传感器与图像	2
1.1.3 摄影平台	4
1.2 几何变换	5
1.2.1 点、线、面、角	5
1.2.2 几何变换的层次	9
1.2.3 从 3D 到 2D 的几何变换	10
1.3 数学基础	13
1.3.1 线性时不变系统和图像处理	13
1.3.2 最小二乘平差	19
1.3.3 粗差的处理	21
1.3.4 系统误差的处理	23
1.3.5 最大似然,最大后验及贝叶斯定理	24
1.3.6 变分法	25
1.4 数学工具和平台	29
1.4.1 代数运算工具	29
1.4.2 OpenCV 库	29
1.4.3 深度学习平台	30
课后习题	31
第 2 章 单像解析	32
2.1 坐标系统	32
2.1.1 图像坐标系	32
2.1.2 传感器坐标系	34
2.1.3 平台坐标系	34
2.1.4 局部切平面坐标系	35

2.1.5 地心直角坐标系	35
2.2 普通光学相机的成像模型	35
2.2.1 平面相机	35
2.2.2 线阵相机	36
2.2.3 鱼眼相机	37
2.2.4 通用几何成像模型	39
2.3 全景相机的四种成像模型	43
2.3.1 理想的全景相机	43
2.3.2 多镜头组合式全景相机	45
2.3.3 旋转式全景相机	46
2.3.4 折反射式全景相机	47
2.4 单像解析的一些补充	48
2.4.1 空间后方交会和解法	48
2.4.2 共线条件方程的正算和反算	49
2.4.3 灭点与灭线	50
2.4.4 相机检校	51
课后习题	53
 第3章 立体几何	 54
3.1 普通核线几何	54
3.2 线阵相机的核线	57
3.3 全景相机的核线	58
3.4 三视几何	59
3.4.1 基于线的三视几何	59
3.4.2 三焦点张量*	61
3.5 核线几何的一些补充	63
3.5.1 核线立体像对	63
3.5.2 立体模型的绝对定向	64
3.5.3 前方交会和立体重建	65
课后习题	66
 第4章 运动成像平台的定位方法	 67
4.1 贝叶斯理论框架下的通用定位模型	67
4.1.1 光束法平差的缺点	67

4.1.2 通用定位模型	68
4.2 光束法区域网平差	69
4.2.1 基本方法	69
4.2.2 光束法平差的初值	73
4.2.3 光束法平差的控制条件	75
4.2.4 自检校光束法平差	77
4.3 卡尔曼滤波	78
4.3.1 离散卡尔曼滤波	79
4.3.2 扩展卡尔曼滤波	81
4.4 粒子滤波	83
4.4.1 蒙特卡罗和马尔可夫链	83
4.4.2 用带权粒子表达概率分布	84
4.4.3 粒子滤波算法	85
4.5 一些补充	87
4.5.1 从运动恢复结构	87
4.5.2 SLAM 与 VisualSLAM	87
课后习题	91
第 5 章 图像匹配	92
5.1 特征与特征提取	92
5.1.1 什么是好特征?	92
5.1.2 Harris 算子	93
5.1.3 SIFT 算子	95
5.1.4 A-SIFT 与最小二乘匹配	97
5.2 图像匹配	99
5.2.1 几种相似性测度	99
5.2.2 误匹配处理	100
5.2.3 匹配效率	102
5.3 图像密集匹配	103
5.3.1 概述	104
5.3.2 近似全局匹配方法	104
5.3.3 半全局方法	107
5.4 一些补充	110
5.4.1 视觉词袋	110

5.4.2 闭环探测	110
5.4.3 多源图像匹配	111
5.4.4 非线性光照与本征图像分解	111
课后习题	113
第6章 相机与激光雷达的融合	114
6.1 传感器静态检校	114
6.1.1 相机与 GPS、IMU 的检校	114
6.1.2 基于线的共线条件方程	115
6.1.3 激光雷达与相机的检校	115
6.2 数据级融合	117
6.2.1 看作 2D-3D 或 3D-3D 刚体配准问题	117
6.2.2 考虑激光点云中的误差	118
6.3 几何级融合	119
6.3.1 LiDAR 辅助的尺度约束	119
6.3.2 几何约束下的闭环探测	120
6.4 三维场景生成	121
6.4.1 基于图像的三维场景生成	121
6.4.2 基于点云的三维场景生成	123
6.4.3 基于图像和点云融合的三维场景生成	124
6.5 一些补充	124
6.5.1 机载激光雷达	124
6.5.2 RGB-D 相机与 Kinect	125
6.5.3 三维建模	125
课后习题	126
第7章 专题图、深度学习与智能摄影测量	127
7.1 数字地图	127
7.1.1 数字高程模型	127
7.1.2 数字正射图像	129
7.1.3 数字线划图	130
7.2 深度学习	132
7.2.1 历史	132
7.2.2 基础	133

7.2.3 深度卷积神经网络	136
7.3 专题图及深度卷积神经网络的应用	137
7.3.1 图像检索专题图	137
7.3.2 语义分割专题图	139
7.3.3 目标识别专题图	140
7.3.4 实例分割专题图	143
7.3.5 矢量提取专题图	143
7.3.6 基于深度学习的立体匹配	144
7.4 深度学习在建筑物提取中的应用	147
7.4.1 建筑物开源数据库的建立	147
7.4.2 建筑物语义分割	149
7.4.3 建筑物实例分割	150
7.4.4 建筑物矢量提取	151
7.5 展望智能摄影测量时代	152
课后习题	153
参考文献	154

第 1 章 基 础

摄影测量学是一门利用光学像片研究被摄物体的形状、位置、大小、特性及相互位置关系的学科^[1,2],它涉及图像、传感器、摄影平台、成像机理、计算和应用等方面。本章介绍有关几何成像、传感器技术、摄影平台的基本知识,以及关于图像处理、最优化和误差处理的数学方法。本章结合现代摄影测量学与计算机视觉的基础理论,为更好地理解本书的其他章节打下基础。

1.1 物 理 相 关

1.1.1 光与视觉

数亿年前,一只古老的海洋生物在浅海徘徊,它的皮肤细胞发生了基因突变,变得对光更加敏感;进化至今,成为现代人类精致的眼睛,并能感知几个光子组成的微弱的光。视网膜中遍布着称为“受光体”的两类特殊蛋白质细胞:①视椎细胞,分别对红、绿、蓝三种波段的光敏感,故将其定义为三原色——事实上人类只能看到这三种颜色,黄色和棕色等是大脑用合成的方式帮人们“看”到的;②视杆细胞,对所有的光都敏感,故人们可夜间视物,但不能区分色彩。在眼睛中两寸照片大小的面积里,共有 1.25 亿个视杆细胞和 600 万个视椎细胞——相当于 130 兆像素的数码相机,比当前普通相机的分辨率高 10 倍。不过,根据摩尔定律,人眼分辨率即将被数码相机超越。人眼的晶状体相当于透镜,汇聚光并透视成像于视网膜上。但最初,眼睛中并没有晶状体,世界浑浊不清;随着晶状体的出现和不断进化,正如镜头调焦,世界景象越加清晰。当然,进化并非总是发生,海洋中的鹦鹉螺——5 亿年前的活化石,依然没有晶状体。

光是一种电磁波,具有波粒二象性。对于均匀介质中“光子”的路径,可以用一条简单的空间直线表达,若考虑相对论效应,那是一条光的“测地线”。对于“光波”,需要用到惠更斯方程。在摄影测量中通常忽略光的波动,仅考虑三维空间的小孔成像(但遥感和天文学科的许多分支都要考虑“波”的特性)。

在不均匀介质中,如大气或从空气进入水中,光子的路径将不再是直线。严格地说,光在介质中的传播路径由费马光程定理决定:“光的行程时刻最短”。这是几何光学的“真正基础”,而光沿着直线传播只是在均匀介质中的特例。这条定理为

计算光的路径提供了准确的理论依据,而泛函分析中的变分法提供相应解算方法。然而,动态的大气密度分布通常很难精确测量,并且假定为同一密度对空基和地基的光学成像系统影响甚微。通常认为(透过大气的)光仍然沿直线传播。

1.1.2 传感器与图像

2009年,诺贝尔物理学奖授予“光纤通信和电荷耦合器件(charge coupled device, CCD)的发明者”^[3]。CCD利用电子捕获光线(背后的原理即光电效应,由赫兹1887年提出,并由爱因斯坦1905年诠释),以替代以往的胶片成像,由此带来摄影技术的彻底革命。它直接推动了医学、天文学的发展,摄影测量学自然也不例外。20世纪90年代末仍存在着关于“CCD是否将取代胶片”的争论。如今,胶片已成过往云烟,以柯达为首的胶片商业帝国也轰然崩溃。当年引以为傲的“胶片的高分辨率(由银颗粒大小所决定)”,也被量子力学支撑的光电效应抛在后面:CCD的理论分辨率甚至可以达到1个光子。

摄影测量中所使用的专业光学相机,包括大部分单反数码相机,主要成像器件就是CCD;不过,绝大部分普通数码相机,包括手机自带的相机,成像器件是互补金属氧化物半导体(complementary metal-oxide-semiconductor, CMOS)。它作为电脑等电子器件上的随机存储器广为所知,然而,CMOS也可当作感光元件使用。CMOS感光组件造价低廉,比CCD耗电量低、数据传输快,统治着中低端数码相机市场。

以下公式描述光电转换的大致过程^[4]:

$$N = \delta A \delta t \int_0^{\infty} b(\lambda) q(\lambda) d\lambda \quad (1.1)$$

式中: λ 为波长,理论上可趋于无限,但一般而言,有效成像波段仅在全色波段附近; $b(\lambda)$ 为图像上的辐射强度,即单位面积上的光子数; $q(\lambda)$ 为量子效率,是激发电子流与入射光子流之比,和感光材质有关。整个公式表示,在感光材料 δA 面积上、 δt 时间内,所激发的电子数的总和为 N 。注意到入射光与激发电子之间的线性关系(积分是一种线性运算),理论上可以通过图像来反演场景的辐射信息。

借助CCD或CMOS等感光器件,人们实现了数码成像的第一步转换。存储激发得到的电子留给另一项物理技术:模数转换(A/D)。模数转换将连续变化的模拟量(电压)转换成离散的数字(数模转换正相反,它是将数字转换成连续的电流信号)。离散化后得到的光栅矩阵,称为像素。采样定理决定了最优的采样方法。

最后需要把光栅矩阵存储为数字信号。可以先将数字信号放入内存(半导体存储单元),然后存放到外存中(软盘、硬盘、光盘等)。对于光学传感器获得的数字信号,赋予一定的文件头,并称为图像(影像、像片)。文件头定义了图像的读写格式,

常用的格式包括 jpg、bmp、tiff 等。

现在回顾整个流程:数码相机取景、调焦、拍摄,光子从物体出发,经过透光的镜头,打在 CCD 上,引起了电压的变化,这些电压经过模数转换,化为一系列 0-1 代码,被忠实记录到磁盘中。把这个图像文件拷贝到电脑上,CPU 接到显示的指令,通过数模转换,将数字信号转换为电压(或电流),显示器读出电信号,再通过光电效应发出光子,被眼睛捕获,场景再现。

然而,以上步骤再现的只是二维的场景。通过 CCD 成像,三维世界投影并压缩至二维平面,由此带来信息量的损失(主要是深度信息的损失)。从单幅二维图像恢复三维世界,通常视为病态问题(ill-posed problem)。是否有特殊的成像方式,能够从二维图像无损记录三维世界?光学全息成像就是这样一种技术。首先回顾量子力学中的“全息原理”,一个系统可以由它的边界信息(自由度)完全描述;再回顾微积分中的格林定理:平面闭区域上的“面”积分可由其边界“曲线”积分表达(有复变函数分析背景的读者也会联想到柯西积分公式:任何一个闭合区域上的全纯函数在区域内部的取值完全取决于它在区域边界上的值)。异曲同工的数理理论证明:高一维的事物,可由低一维的事物完全描述。全息原理不仅适用于光学,同样适用于 X 射线、微波、声波,甚至引力。特别是引力全息成像原理,意味着只要研究黑洞表面,就可获知黑洞的内部构造——后者可能永远无法直接观测。

具体地,在摄影测量、计算机视觉及虚拟现实,全息摄影可由干涉光完成。和 CCD 不同,干涉成像是一种主动成像模式,发射两束可见光波段的激光(因为激光的相干性最好),一束直接射向自带全息成像面,另一束射向物体后反射回成像面,从而形成干涉条纹。干涉条纹记录相位信息,可恢复物体的深度;结合物体的反射光,就恢复了真实的 3D 场景。干涉成像仪常用来获取缺乏纹理的三维物体表面,该表面通常很难由图像匹配技术(详见第 5 章)来完成。

还有一种直接从单张图像恢复三维形状的方法,称为“从明暗恢复形状”(shape from shading),一般用于已知反射率的均匀材质成像。如果人们有两张以上的图像,就可根据立体视觉原理恢复三维世界;相关内容将在 2.5 节、4.3 节和 5.3 节等详细介绍。

除了光学传感器,在摄影测量学中还涉及另一种重要传感器:激光器(laser)或激光雷达(light detection and ranging, LiDAR)。它本质上也是光学传感器,但人们更关注 LiDAR 的距离观测值,而非辐射信息。故一般将它归结为测距传感器,就像声呐(sonar)和雷达(radar)。LiDAR 扫描直接获取物体表面的点云,这些深度信息是光学图像的良好补充。

摄影测量学也涉及和位置姿态有关的传感器,这里笼统称为辅助传感器。最著名的是全球定位系统(global positioning system, GPS)。在成像平台上放置一个

GPS 接收器,可确定平台的运行轨迹及相机曝光时刻的位置。如果用 GPS 精确测量某个实物的世界坐标,这个点可以作为控制条件,辅助解算像片在坐标系中的位置。全球导航卫星系统 GNSS 是 GPS 的升级版,它包括了美国的 GPS、俄国的 GLONASS、欧盟的伽利略及中国的北斗等诸多导航定位系统。

除了位置,像片的姿态也是人们所关心的重要参数。这个值可以由惯性导航装置(inertial measurement unit,IMU)或惯性导航系统(inertial navigation system,INS)直接测得,IMU 利用了陀螺的惯性。还记得小时候玩过的陀螺吗?在惯性作用下,它能在桌面上长时间水平旋转;如果这时候将桌面倾斜,它就能测得桌面倾斜的角度。同理,将桌面换成摄影平台(如车辆),就获得了平台的姿态。如果事先获得曝光时刻像片的位置和姿态(通常称为 GNSS/INS 直接定位技术),那么空中三角测量(详见第 4 章)就不再需要。事实上高精度的 IMU 造价昂贵,目前三角测量技术仍然有效。

至于非可见光成像,有很大一部分用于现代医学。如 X 射线可穿透人体,获得人体内部图像,但一个人一年照射 X 射线的量有上限,否则会对人体造成伤害。相比 X 射线,磁共振成像(magnetic resonance imaging,MRI)是当代医学更强大的工具。这是一种非常特殊的成像方式,不是在传统的光谱空间域,而是在频率域成像。首先将生物体置于强大磁场中,体内的原子会指向同一方向;然后发射一束电子脉冲信号,原子将受激翻转;最后原子在磁场作用下又翻转回来,同时发射一个回波;通过回波,可探测原子的位置和性质,并由计算机技术重建三维图像。三维图像的像素称为“体素”(voxel)。

1.1.3 摄影平台

摄影测量有着极其丰富的摄影平台,而真正成为一门学科要从航空摄影开始。第一幅用于测绘的照片,摄于德国飞艇;然而,随着齐柏林飞艇的一声巨响,飞艇时代谢幕;能耗更高的飞机成了航空飞行器的主流(设想没有那次事故,21 世纪可能是漫天的飞艇)。用于摄影测量的专业飞机称作“航摄飞机”,我国常用的型号有“空中国王”“奖状”等。与普通飞机相比,航摄飞机具有较好的稳定性,安装伺服系统,机舱底部开有对地摄影窗口。目前,航空平台也呈现多样化的趋势,除了传统航摄飞机外,还有无人飞机、直升机、飞艇等。后者往往飞行于数百米到 1 000 米的低空,有时也被称为“低空摄影测量”。

摄影测量的一次飞跃发生在 20 世纪 70 年代 Landsat 系列卫星的发射和 Landsat TM 影像数据的普及,摄影测量被扩展为“摄影测量与遥感”。此后,分辨率越来越高的测绘、遥感卫星相继投入使用,包括法国的 SPOT 系列(1986 年至今),美国的 IKONOS、QuickBird 和 Geo-eye(1999 年,2001 年,2008 年分别发射),日本

的 ALOS(2006 年),以色列的 EROS-B(2006 年),印度的 Cartosat 系列、中国的资源系列(1999 年至今)、韩国的阿里郎三号(2012 年)等,更高光学分辨率和光谱分辨率的卫星和卫星星座仍处于持续发展中。基于卫星平台的摄影测量有时也被称为“卫星摄影测量”。

摄影测量的另外一个重要分支是“近景摄影测量”。以往,摄影测量主要为测绘服务,近景摄影测量相对冷门。事实上,近景摄影测量面向更广泛的人群,涉及更多的行业。近景摄影测量可为交通业和公众服务,提供电子地图和三维导航信息;为制造业服务,测量零件尺寸、检测器件缺陷;为文化遗产服务,制作永久性的电子仿真品;甚至可用于医学,配准异源医学图像,发现潜在疾病。近景摄影测量也是和计算机视觉、机器人学最密切相关的学科分支。这些学科的主题都是利用地面摄影平台,挖掘光学图像(可能结合距离成像)中的有用信息。地面平台的形式包括:手持、车载、机器人、固定位置(如监控镜头和工业相机)等。

深空平台和水下平台正走进人们的视野。中国嫦娥系列飞船(2007 年至今)及登陆器为人们带来了月球表面的高分辨率图像;美国的火星探测计划正逐步揭开火星的面纱。摄影测量技术可广泛应用于深空平台的自主定位和导航(没有 GPS)、星球表面的三维重建等。然而,令人尴尬的是,人们对占地球 70% 的海底世界所知甚微(仅 5% 已知),甚至比不上火星^[6]。光学相机无法在深海无光或微光环境下成像;声呐技术无法解析精细构造。这种状况正随着深海探测器的成熟和应用而逐步得到改善。

1.2 几何变换

这一部分包括两个内容:几何概念的表达式及本书涉及的几何模型。将介绍欧几里得空间的几何变换层次,从最特殊的旋转平移直到最普通的透视变换;并引出中心投影变换的几何模型“共线条件方程”,它是摄影测量学科,同时也是本书的核心之一。

1.2.1 点、线、面、角

射影几何是摄影测量的解析基础。最初,人们只关心透视变换(即小孔成像)下像点与物点之间的几何关系;随着研究的深入,线段或直线逐渐在立体匹配、平差、三维重建中得到广泛应用;同样,面在摄影测量中也是一个重要的几何实体,如建筑物立面建模、小面元微分纠正。人们需要首先建立这些几何实体恰当的数学描述。

点坐标用矢量 $\mathbf{x}$ 表示。二维欧几里得空间 R^2 中的某个点记作 $\mathbf{x} = (x, y)$, 或 $\mathbf{x} \in R^2$ 。若将其再扩充一维,写作 $\mathbf{x} = (x, y, 1)$, 两者表达同一个点。假定 $(2x, 2y, 2)$,

直到 (kx, ky, k) ,也表达同一个点。这种表达方法称为该点的“齐次坐标”。从齐次坐标恢复普通的坐标,只需要将齐次坐标的前两项除以第三项即可。有一个例外是 $(x, y, 0)$,这时普通坐标无穷大。这恰恰是人们所需要的:摄影测量中的灭点(无穷远点)、灭线等概念用齐次坐标更好表达。除此之外,引入齐次坐标将给本书涉及的大部分几何公式的书写和推导带来极大的方便。

与二维点坐标类似,三维点的齐次坐标表示为 $x = (x, y, z, 1)$ 。

线用矢量 l 表示。方程 $ax + by + c = 0$ 表达二维空间的一条直线。即:点 $(x, y, 1)$ 出现在由参数 (a, b, c) 所定义的直线上。 (a, b, c) 完全刻画了该直线的性质,令 $l = (a, b, c)$,这是线的齐次坐标(回顾初等几何,线可以只用斜率和截距两个参数表达)。此时点在线上的几何关系,可以用矢量表达为

$$x \cdot l = 0 \quad (1.2)$$

通过线的齐次坐标,可以方便地求出线与线的交点式(1.3)、两点一线式(1.4):

$$x = l_1 \times l_2 \quad (1.3)$$

$$l = x_1 \times x_2 \quad (1.4)$$

这里的“ $\cdot$ ”表示点乘,不引起歧义的前提下通常省略;“ $\times$ ”表示矢量叉积。

略作变换,可以得到二维直线的另外一种常见形式: $l = (n_x, n_y, d) = (n, d)$,其中, n 为单位法向量(与直线垂直), d 为直线与原点的距离。有一个例外是 $(0, 0, 1)$,这是一条无穷远线,它没有法向量,并且所有的无穷远点都在这条直线上。想象脚下无穷延伸的二维地平线,人们的直觉应该是一个无穷大半径的圆。在数学上,无穷大半径的圆和直线的确等价。

很遗憾,三维空间直线没有二维空间如此简单的表达形式。通常,需要额外引入一个参数来表达。设 (x_1, x_2) 为直线上已知两点的普通坐标,该直线上任意一点表达为

$$x = (1 - \lambda)x_1 + \lambda x_2 \quad (1.5)$$

式中: λ 为任意实数。

面用字母 π 表示。在二维欧几里得空间,面就是空间本身。三维空间的面,可以用 (a, b, c, d) 表达,且满足关系式:

$$x \cdot \pi = ax + by + cz + d = 0 \quad (1.6)$$

与二维直线类似,可将平面改写成法向量的形式: $\pi = (n_x, n_y, n_z, d) = (n, d)$ 。无穷远面 $(0, 0, 0, 1)$ 没有法矢。想象一个无限扩张的三维球面:它显然没有一个固定的法方向。最终这个无限的球面等价于无穷远平面,并且所有的无穷远点和无穷远线都在其上。

和面相关的几何关系除了点在面内(式(1.6)),还包括三点一面:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \pi = 0 \quad (1.7)$$

三面相交于一点:

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} x = 0 \quad (1.8)$$

平面角一般用字母 θ 表示,只需要 1 个参数。当绕逆时针旋转时,新旧坐标点的变换公式为

$$x_2 = R x_1 = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} x_1 \quad (1.9)$$

式中 R 为旋转矩阵。

三维空间的角度则复杂许多,一般有三种表达方法。最常用的是欧拉角。虽然欧拉角最为直观,但是它的缺点也相当恼人。首先是静态欧拉角,即绕固定不动的世界坐标系旋转,一般表示为 φ, ω, κ 。根据不同的旋转顺序(如按照 $y-x-z$ 轴或 $x-z-x$ 轴旋转),三个欧拉角的大小不同,由此构建的旋转矩阵(三个方向旋转矩阵的顺序乘积)也不相同。想要通过欧拉角进行计算,必须已知旋转顺序和方向。动态欧拉角同样有缺陷。动态欧拉角是绕物体的旋转,坐标系随着物体的转动而转动,通常直接用英语表示: yaw、pitch、heading。注意在摄影测量中, φ, ω, κ 表示传感器在世界坐标系中的姿态(也是要求的参数),而 yaw、pitch、heading 反映当前时刻传感器自身的旋转,如采用惯性测量装置实时测得的角度。动态欧拉角可能产生万向锁(gimble lock),丢掉一个维度。最后一个缺陷是,无论是动态还是静态,欧拉角在参数空间上不是连续的。即在绕行一个微小角度的前提下, φ, ω, κ 可能产生巨大的跳变。

即便如此,欧拉角依然在摄影测量中应用最广。以绕 $y-x-z$ 轴连续旋转为例,旋转矩阵通过绕三个轴的旋转矩阵乘积得

$$R = R_{\varphi} R_{\omega} R_{\kappa} = \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{bmatrix} \begin{bmatrix} \cos \kappa & -\sin \kappa & 0 \\ \sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1.20)$$

关于旋转矩阵 R 的其他形式,读者可以自行推导。

第二种表达方法是罗德里格斯公式(Rodriguez's formula)。试想这样一个更普遍的问题:如何绕一个任意的空间矢量 $n = (n_x, n_y, n_z)$ (而不是坐标轴),旋转某个角度 θ ,这里不加推导地给出这个通用公式:

$$R(n, \theta) = I + \sin \theta [n]_{\times} + (1 - \cos \theta) [n]_{\times}^2 \quad (1.21)$$