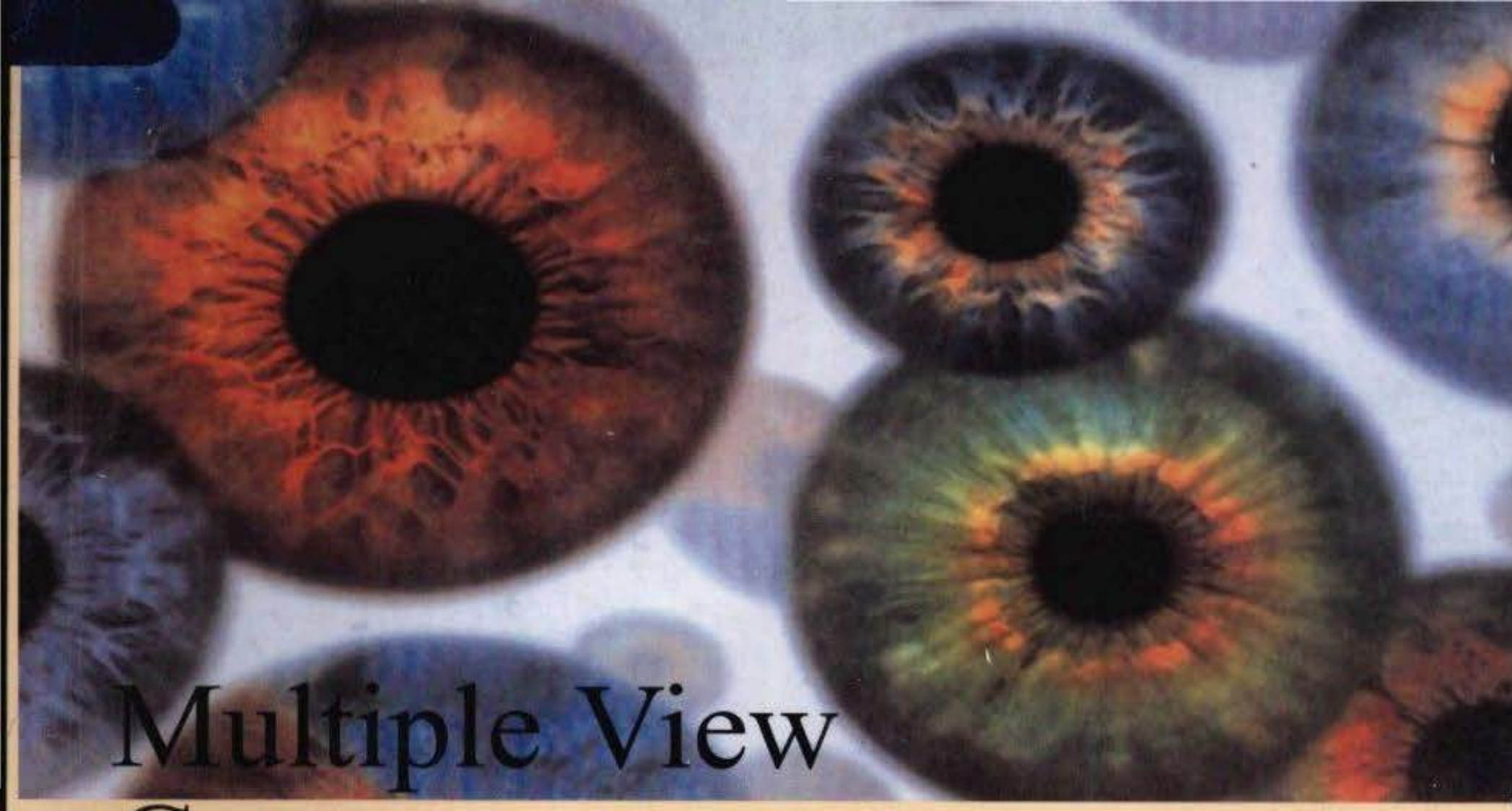


# Multiple View Geometry in Computer Vision

# 计算机视觉中的多视图几何

Richard Hartley Andrew Zisserman 著
韦穗 杨尚骏 章权兵 胡茂林 译

安徽大学出版社

TP391.41/1856
2002

Multiple View Geometry in Computer Vision

# 计算机视觉中的多视图几何

Richard Hartley Andrew Zisserman 著
韦穗 杨尚骏 章权兵 胡茂林 译

安徽大学出版社

**图书在版编目(CIP)数据**

计算机视觉中的多视图几何/ 哈特利(R. Hartley)和齐瑟曼(A. Zisserman)著;韦穗等译 .－合肥:安徽大学出版社,2002.8

ISBN 7－81052－503－4

Ⅰ.计… Ⅱ.①哈…②齐…③韦… Ⅲ.计算机视觉 Ⅳ.TP302.7

中国版本图书馆 CIP 数据核字(2002)第 094394 号

CAMBRIDGE
UNTVERSITY PRESS
剑桥大学出版社 最新引进版

Richard Hartley Andrew Zisserman 著

**计算机视觉中的多视图几何**

韦 穗 杨尚骏 章权兵 胡茂林 译

| | | | |
|---|---|---|---|
| **出版发行** | 安徽大学出版社 | **经 销** | 新华书店 |
| | (合肥市肥西路3号 邮编 230039) | **印 刷** | 合肥远东印务有限公司 |
| **联系电话** | 总编室 0551－5107719 | **开 本** | 787×1092 1/16 |
| | 发行部 0551－5107784 | **印 张** | 30.25 |
| **电子信箱** | ahdxchps@mail.hf.ah.cn | **字 数** | 710 千 |
| **特约编辑** | 章权兵 | **版 次** | 2002 年 8 月第 1 版 |
| **责任编辑** | 朱寒冬 | **印 次** | 2002 年 8 月第 1 次印刷 |
| **封面设计** | 孟献辉 | | |

ISBN 7－81052－503－4/TP·61 **定价** 50.00 元

如有影响阅读的印装质量问题,请与出版社发行部联系调换

# 内容提要

计算机视觉的基本问题是:根据若干幅世界景物的图像求得对真实世界景物结构的理解.

本书解决这个基本问题所采用的技术源于射影几何和摄影测量学. 其与众不同的特色是采用未标定的方法——不需要知道或不必计算摄像机内部参数就能得到问题的答案. 本书以一个统一的框架,对近期关于景物重构的理论和实现两方面的主要发展作了详细的介绍.

本书涵盖了摄像机投影矩阵、基本矩阵和三焦点张量的几何原理和它们的代数表达. 在讨论这些有关的理论和计算方法时都配有实际的例子,如它们在由多幅图像进行景物重构中的应用. 作者提供了综合性的背景材料,读者只要熟悉线性代数和基本的数值方法就能够理解书中给出的射影几何和估计算法,并能直接依据本书来实现有关算法.

# 序

让计算机具有视觉，科学家与工程师们，作出了近40年的不懈努力．应该说，40年努力的进展是显著的，进展主要有两个方面：

已经形成一些计算视觉的基本理论框架，如80年代初形成的以Marr为代表的视觉计算理论（有些学者称之为三维重建框架）和以后出现的基于模型的视觉（Model Based Vision）、主动视觉（Active Vision）等．现在看来，虽然我们仍然不清楚这些计算理论框架能否最终成为最理想的计算机视觉系统的基础，但有几点几乎是可以肯定的：一是迄今为止提出的各种理论框架虽然有方法论上的差异，有些甚至具有科学哲学思想的差异，但并没有本质上的相互排斥，而且是互补的．二是这些已有的视觉系统理论框架已经可以作为有一定程度视觉功能的实用视觉系统的基础．随着计算机性能价格比的指数增长，以现有视觉系统理论框架为基础的，针对特定任务的实用视觉系统，将会广泛应用于现实生活中．三是与人工智能的其他许多领域类似，真正的突破要比当初想像的要困难得多．这里，"真正的突破"是指：当我们将当前的人工智能系统与人相比时，人的智能系统具有更强的通用性、自学习能力、自适应性和对噪声的鲁棒性．

计算机视觉另一方面的重要进展是，提出了大量的计算方法．尤其是90年代以来，为适应不同计算理论框架和为改进计算机视觉系统对噪声的鲁棒性，引进了许多数学方法和与之相对应的计算方法．几乎所有的数学分支，尤其是应用数学分支都要到计算机视觉领域来一显身手，使许多初学者，甚至搞了多年研究的学者都感到困惑．人们不禁要问，难道我们真需要这么多的复杂数学分支和计算方法来解决计算机视觉问题吗？事实上，这确实反映了当前的许多数学工具还不能有效解决"更强的通用性、自学习能力、自适应性和对噪声的鲁棒性"，另一方面，现在的许多数学方法，本质上是相通的．而我们缺少既对这些方法都精通，又对计算机视觉中所面临的实际问题有深入理解的理论工作者来对各种方法加以融会贯通．

在上述视觉计算方法的研究中，基于几何的视觉计算方法，在90年代发展到了几乎是完美的程度．本书的作者既是这方面的先驱者，也在本书中作出了很好的总结与系统论述．基于几何的视觉计算方法，之所以引起很大关注是因为：

1．计算机视觉的研究目标是使计算机具有通过二维图像认知三维环境信息的能力．这种能力将不仅使机器能感知三维环境中物体的几何信息，包括它的形状、位置、姿态、运动等，而且能对它们进行识别与理解．事实上，80年代形成的Marr的计算理论框架和其他计算理论框架中，绝大部分内容都涉及利用几何方法计算环境中的三维物体的形状、位置、姿态和运动．

2．如果读者对欧几里德几何和近几百年来提出的各种几何，如本书中提到的射影几何、仿射几何等有些深入了解的话，应该理解"各种几何的本质是描述几何元素在不同变换群下的不变量"．由此，使用几何方法，不仅可以由二维图像重建（reconstruct）三维物体，还可以描述它们在摄像机变换下的不变量，从而达到识别的目的，也就是说，几何方法，可以贯

穿计算视觉理论框架下的所有部分，有人称之为基于几何的计算机视觉.

3. 90 年代以来，计算机视觉界将对应于射影几何、仿射几何、欧几里德几何的射影变换、仿射变换、欧几里德变换系统地引进到视觉计算方法中. 三种变换都构成变换群，而且，后者为前者的子群，它们所对应的几何不变量，前者为后者的子集. 这些性质比较完美地对应为视觉系统中对物体由粗到细的描述，在一些特定任务的计算机视觉系统中降低了对系统参数了解的要求(如本书中所描述的不需要对摄像机标定的三维重建)，一定条件下提高了系统对噪声的鲁棒性，而这些确实是许多实用计算机视觉系统极为需要的品质.

本书全面介绍了近 10 年来发展的基于几何的计算机视觉计算方法及其数学基础. 除了上述内容外，其中多摄像机视图几何及其计算方法，值得读者关注. 这是因为当前计算机的性能价格比大大提高，使人们有条件在视觉系统中使用更多的摄像机，以利用冗余的信息，来换取系统对噪声的鲁棒性. 系统对噪声的鲁棒性一直是实用计算机视觉系统的瓶颈问题，解决该问题的可能的办法是：提高摄像机的分辨率、多摄像机方法和近年来大量引进的统计最优化鲁棒算法(本书许多章节也有描述).

安徽大学的老师们将本书译成中文，是一件很有益的工作. 我曾长期讲授计算机视觉课程，深感我国工科大学研究生，缺乏现代几何的有关知识，对近 10 年来发展的基于几何的计算机视觉计算方法的本质接受较慢. 本书比较系统地介绍了射影几何，在各章节中也注意介绍有关数学基础，使即使缺少这方面系统知识的工科学生也能接受，应该对我国专门从事计算机视觉研究的读者有较好的参考价值. 这本书对我国从事相关数学领域研究的人士也值得一读，计算机视觉涉及的数学，量大面广，是一个典型的数学工作者可有用武之地的领域，但比起其他国家来说，我国的数学家们基本不介入一些有相当实用背景的新兴学科，学科不能交叉，创新从何而来？

由于本书是介绍计算机视觉中的一个分支的很专业的书，为了使初学者对其背景有一点了解，我对本书的内容和特点做了上述介绍，以此为中译本序，不一定准确，望读者批评指正.

**中国科学院自动化所　马颂德**

# 原序

奥利维尔·法格罗斯(Olivier Faugeras)

60 年代,在人工智能领域的带头专家眼里,使计算机具有视觉功能充其量只是属于暑期学生设计的事. 40 年以后这项任务仍然没有解决并且似乎还很艰难. 称之为计算机视觉的整个领域本身已成为一门与数学和计算机科学都有很强联系的学科,同时它与物理、感知心理学和神经科学也有一定的联系.

造成部分失败的一种可能的原因是研究者忽略了这样的事实:动物和人类的感知,特别是视觉感知比当初想像的要复杂得多. 当然没有理由要求计算机视觉算法一定要模仿生物,但事实是:

(1)生物视觉工作的方式仍有许多未知的东西,因而难以在计算机上模拟.

(2)企图忽略生物视觉而重新发明一种基于硅片的视觉并没有像当初想像的那样成功.

除了这些负面的评论外,计算机视觉方面的研究者在实践和理论两个方面都已经获得了某些显著的成功.

在实践方面,举一个例子说明,用计算机视觉技术引导汽车、卡车等交通工具在平坦的道路或崎岖的地形上行驶已成为可能,并且许多年前就在欧洲、美国和日本进行过演示. 引导车辆需要相当复杂的实时分析三维动态景物的能力. 今天,汽车制造商已慢慢地将其中的某些功能集成到他们的产品中去.

在理论方面,称之为几何计算机视觉的领域已经取得了一些显著的成就. 其中包括把从不同视点观察到的物体表观的变化描述成一个关于物体形状和摄像机参数的函数. 如果不应用相当复杂的数学技术,这样的成就是不可能得到的, 上述数学技术囊括了几何的许多领域,有古代的也有现代的. 这本书特别对世界物体的图像间存在的复杂而又美妙的几何关系加以研究. 对这些关系加以分析本身是很重要的,因为提供对视觉表观的解释是科学的目标之一, 同时研究它们的另一个重要原因是对它们的理解导致的应用范围越来越广.

这本书的作者是两位几何计算机视觉领域的开拓者和专家. 他们在具有挑战的领域取得了成功,即他们把需要理解的几何概念表达得浅显易懂,把他们以及全世界的其他学者获得的成果覆盖得很全面,分析了几何与图像测量必含有噪声这一事实之间的相互影响,把许多理论的结果表达成算法的形式,从而使它们能够很容易地被转换成计算机代码,并且给出了许多真实的例子来解释概念,展示了理论的应用范围.

回到使计算机具有视觉功能的初衷,我们也许怀疑这种工作是否是在正确方向上. 我必须让本书的读者来回答这个问题,并且我相信读者会赞同如下的断言:任何一个打算用摄像机连接计算机的系统设计者都不会忽略这项工作. 这可能是在定义使一台计算机具有视觉功能到底意味着什么这个方向上重要的一步.

# 前 言

过去十多年里，计算机视觉在多视图几何的理解和建模方向已得到迅速发展．理论和实践已达到成熟的程度，10 年前尚未解决并经常被认为无法解决的一些问题现在已经有了漂亮的结果．这些任务和算法包括：

- 给定两幅图像而不附带其他信息，计算图像之间的匹配、产生这些匹配的点的 3D 位置以及得到这些图像的摄像机．
- 给定三幅图像并不附带其他信息，类似地计算图像之间的匹配点和直线，以及这些点和直线的 3D 位置和摄像机．
- 在不需要标定物体的情况下，计算双眼装置的对极几何以及三目装置的三焦点几何．
- 由自然景物的图像序列来计算摄像机的内参数（即“在放映中”标定）．

这些算法的与众不同的特点是它们是**未标定**的——不必要知道或不必首先计算摄像机的内参数（例如焦距）．

支撑这些算法的基础是一种新的、更完整的关于多幅未标定视图的几何理论的理解：所包含的参数数目，视图中点和直线之间的约束，以及由图像对应恢复摄像机和三维空间点．例如确定一副双眼装置的对极几何仅需要指定 7 个参数，并不需要对摄像机进行标定．这些参数可以由 7 个或更多的图像点对应确定．与此非标定的路线相反，10 年前采用了预先标定的路线：每个摄像机必须首先用工程上仔细标定的并已知几何的物体的图像进行标定．标定涉及确定每一个摄像机的 11 个参数，然后由这样两组 11 个参数的数据才能计算对极几何．

该例子说明未标定（射影）方法的重要性——采用适宜的几何表达可使计算每一阶段所需要的参数更明晰．这样避免了计算那些对最后结果没有影响的参数，并得到更简单的算法．同时，在这里值得纠正一个可能产生的错误概念．在未标定框架中，实体（例如三维空间点）通常在一个准确定义的多义性下加以恢复，这种多义性并不表示点是一种不良估计．

更贴近实际来看，通常不可能对摄像机进行一次标定后就永久有效，例如摄像机被移动了（在移动的车上）或内参数改变了（具有变焦的侦察摄像机）．进一步说，在某些情况下标定信息并不能简单得到．想像如下情况：由视频序列计算摄像机的运动，或由归档的胶片卷构造虚拟现实的模型，其中运动和内标定信息都是未知的．

在多视图几何方面之所以取得成功可能是因为我们关于理论理解方面的进展，同时也是由于由图像估计数学目标的提高．第一个提高是关注的超定系统的误差必须最小化——不论它是代数的、几何的或是统计的；第二个提高是使用了鲁棒估计算法（例如 RANSAC），使得估计不受数据中“野值”的影响．同时，这些技术产生强有力的搜索和匹配算法．

我们可以说现在许多重构的问题已经解决．这些问题包括：

(1) 由图像点对应估计多焦点张量，特别是基本矩阵和三焦点张量（四焦点张量还没有得到重视）．

# 目　录

序 …… 1
原序 …… 1
前言 …… 1
第 0 篇　基础知识:射影几何、变换和估计 …… 1
本篇大纲 …… 1
第 1 章　2D 射影几何和变换 …… 2
1.1　平面几何 …… 2
1.2　2D 射影平面 …… 2
1.3　射影变换 …… 8
1.4　变换的层次 …… 11
1.5　1D 射影几何 …… 17
1.6　射影平面的拓扑 …… 18
1.7　从图像恢复仿射和度量性质 …… 20
1.8　二次曲线的其他性质 …… 28
1.9　不动点与直线 …… 30
1.10　结束语 …… 32
第 2 章　3D 射影几何和变换 …… 34
2.1　点和射影变换 …… 34
2.2　平面、直线和二次曲面的表示和变换 …… 34
2.3　三次绕线 …… 43
2.4　变换的层次 …… 44
2.5　无穷远平面 …… 46
2.6　绝对二次曲线 …… 47
2.7　绝对对偶二次曲面 …… 49
2.8　结束语 …… 50
第 3 章　估计——2D 射影变换 …… 52
3.1　直接线性变换(DLT)算法 …… 53
3.2　不同的代价函数 …… 57
3.3　统计代价函数和最大似然估计 …… 63
3.4　变换不变性和归一化 …… 65
3.5　迭代最小化方法 …… 69
3.6　算法的实验比较 …… 72

3.7 鲁棒估计 …… 73
3.8 单应的自动计算 …… 79
3.9 结束语 …… 82
**第4章 算法评价和误差分析** …… 86
4.1 性能的界定 …… 86
4.2 变换估计的协方差 …… 90
4.3 协方差估计的蒙特卡洛法 …… 98
4.4 结束语 …… 99

**第1篇 摄像机几何和单视图几何** …… 100
**本篇大纲** …… 100
**第5章 摄像机模型** …… 101
5.1 有限摄像机 …… 101
5.2 射影摄像机 …… 104
5.3 无穷远摄像机 …… 111
5.4 其他摄像机模型 …… 118
5.5 结束语 …… 119
**第6章 计算摄像机矩阵P** …… 121
6.1 基本方程 …… 121
6.2 几何误差 …… 123
6.3 受限摄像机估计 …… 126
6.4 径向失真 …… 130
6.5 结束语 …… 132
**第7章 进一步讨论单视图几何** …… 134
7.1 射影摄像机对平面、直线和二次曲线的作用 …… 134
7.2 光滑曲面的图像 …… 137
7.3 射影摄像机对二次曲面的作用 …… 138
7.4 摄像机中心的重要性 …… 140
7.5 摄像机标定与绝对二次曲线的图像 …… 145
7.6 消影点与消影线 …… 149
7.7 由消影点和消影线确定标定K …… 153
7.8 结束语 …… 155

**第2篇 两视图几何** …… 158
**本篇大纲** …… 158
**第8章 对极几何和基本矩阵** …… 159
8.1 对极几何 …… 159
8.2 基本矩阵F …… 161
8.3 由特殊运动产生的基本矩阵 …… 166

8.4 基本矩阵的几何表示 …… 168
8.5 恢复摄像机矩阵 …… 171
8.6 本质矩阵 …… 173
8.7 结束语 …… 176
**第9章 摄像机和结构的3D重构** …… 178
9.1 重构方法概述 …… 178
9.2 重构的多义性 …… 179
9.3 射影重构定理 …… 181
9.4 分层重构 …… 182
9.5 直接重构——利用地面知识 …… 188
9.6 结束语 …… 189
**第10章 计算基本矩阵F** …… 191
10.1 基本方程 …… 191
10.2 归一化8点算法 …… 193
10.3 代数最小化算法 …… 193
10.4 几何距离 …… 195
10.5 算法的实验评估 …… 198
10.6 自动计算F …… 200
10.7 F计算的特殊情形 …… 202
10.8 其他几何元素的对应 …… 203
10.9 退化 …… 204
10.10 F计算的几何解释 …… 205
10.11 对极线的包络 …… 206
10.12 图像矫正 …… 211
10.13 结束语 …… 214
**第11章 结构计算** …… 216
11.1 问题陈述 …… 216
11.2 线性三角形法 …… 217
11.3 几何误差代价函数 …… 219
11.4 Sampson近似(一阶几何矫正) …… 219
11.5 最优解 …… 220
11.6 直线重构 …… 225
11.7 计算消影点 …… 226
11.8 结束语 …… 227
**第12章 景物平面和单应** …… 228
12.1 给定平面的单应和逆问题 …… 228
12.2 给定F和图像对应求平面诱导的单应 …… 231
12.3 由平面诱导的单应来计算F …… 235
12.4 无穷单应$H_\infty$ …… 239

12.5 结束语 …… 240
第13章 仿射对极几何 …… 243
13.1 仿射对极几何 …… 243
13.2 仿射基本矩阵 …… 244
13.3 由图像点对应估计 $F_A$ …… 246
13.4 三角形法 …… 250
13.5 仿射重构 …… 251
13.6 Necker 反转和浅浮雕多义性 …… 253
13.7 计算运动 …… 254
13.8 结束语 …… 256

第3篇 三视图几何 …… 258
本篇大纲 …… 258
第14章 三焦点张量 …… 259
14.1 三焦点张量的几何基础 …… 259
14.2 三焦点张量和张量记号 …… 268
14.3 转移 …… 270
14.4 基本矩阵与三焦点张量的关系 …… 273
14.5 结束语 …… 274
第15章 三焦点张量 $\mathcal{T}$ 的计算 …… 277
15.1 基本方程组 …… 277
15.2 归一化线性算法 …… 278
15.3 代数最小化算法 …… 280
15.4 几何距离 …… 281
15.5 算法的实验评价 …… 283
15.6 $\mathcal{T}$ 的自动计算 …… 284
15.7 计算 $\mathcal{T}$ 的特殊情形 …… 287
15.8 结束语 …… 288

第4篇 $N$ 视图几何 …… 290
本篇大纲 …… 290
第16章 $N$ 线性和多视图张量 …… 291
16.1 双线性关系 …… 291
16.2 三线性关系 …… 293
16.3 四线性关系 …… 297
16.4 四张平面的交 …… 299
16.5 计数的推导 …… 300
16.6 独立方程数 …… 304
16.7 方程选取 …… 306

16.8 结束语 …… 307
第17章 N视图计算方法 …… 309
17.1 射影重构——捆集调整 …… 309
17.2 仿射重构——矩阵分解算法 …… 310
17.3 射影分解 …… 313
17.4 由平面诱导的单应重构 …… 316
17.5 利用序列重构 …… 318
17.6 结束语 …… 319
第18章 自标定 …… 323
18.1 引言 …… 323
18.2 代数框架和问题陈述 …… 324
18.3 利用绝对对偶二次曲面标定 …… 326
18.4 Kruppa 方程 …… 333
18.5 分层求解 …… 335
18.6 从旋转摄像机标定 …… 341
18.7 平面自标定 …… 343
18.8 平面运动 …… 345
18.9 双眼装置的自标定 …… 347
18.10 结束语 …… 350
第19章 对偶 …… 354
19.1 Carlsson－Weinshall 对偶 …… 354
19.2 简化重构 …… 359
19.3 结束语 …… 363
第20章 正负性 …… 364
20.1 准仿射变换 …… 364
20.2 摄像机的前面和后面 …… 367
20.3 三维点集合 …… 367
20.4 获得一个准仿射重构 …… 368
20.5 变换正负性的效果 …… 369
20.6 定向 …… 370
20.7 正负性不等式 …… 372
20.8 哪些点在第三幅视图中可见 …… 374
20.9 谁在前面 …… 375
20.10 结束语 …… 377
第21章 退化配置 …… 378
21.1 计算摄像机投影矩阵 …… 378
21.2 两视图中的退化特性 …… 383
21.3 Carlsson－Weinshall 对偶 …… 388
21.4 三视图临界配置 …… 394

21.5 七点的多义性视图 …… 398
附录 1 张量记号 …… 400
附录 2 高斯(正态)分布与卡方分布 …… 402
附录 3 数值算法 …… 404
附录 4 迭代估计方法 …… 416
附录 5 某些特殊的平面射影变换 …… 427

参考文献 …… 432
索引 …… 446
后记 …… 468

# 第 0 篇

# 基础知识:射影几何、变换和估计

## 本篇大纲

此篇的四章为本书后续章节将要用到的表示、术语和记号打基础. 射影几何的概念和表示法是多视图几何分析的核心. 例如,使用了齐次坐标就能用线性矩阵方程来表示非线性映射(例如透视投影);而且可以很自然地表示无穷远点,避免了取极限的麻烦.

第 1 章引入二维空间(以下简称 2D)的射影变换. 这些变换发生在用透视摄像机对平面摄像的时候. 该章偏重于入门介绍并为三维空间(以下简称 3D)几何铺路. 大多数的概念在 2D 中比 3D 中更容易理解和可视化. 本章介绍射影变换,包括它的特殊情况:仿射和相似变换;并把注意力主要集中在从透视图像中恢复仿射性质(例如平行线)和度量性质(例如线之间的角度).

第 2 章覆盖 3D 射影几何. 该几何的推导方式与 2D 非常相似,当然,由于维数的增加出现了新的性质. 这里,主要的新几何是无穷远平面和绝对二次曲线.

第 3 章介绍由图像测量进行几何**估计**,这是本书的主要论题之一. 我们以用点对应来估计射影变换为例来说明估计算法的要素和目的,这些算法将在整本书中被采用. 其中重要的问题是:一个代价函数应该最小化什么,比如是代数的或是几何的或是统计的测量,该问题将长篇地加以介绍. 本章还打算介绍鲁棒估计的思想,以及这样的技术在变换的自动估计中的应用.

第 4 章介绍如何评价估计算法的结果. 特别是如何计算估计的协方差.

# 第 1 章　2D 射影几何和变换

本章主要介绍理解此书内容所必须的几何概念和记号. 某些概念为大家所熟悉,例如消影点的形成和二次曲线的表示,而其余的却比较深奥,例如用虚圆点去消除图像中的射影失真. 这些概念在平面(2D)中比较容易理解,因为在 2D 中它们比较容易可视化. 以后本书将涉及的 3D 几何的内容就是平面情形的简单推广.

具体地说,本章涵盖平面的射影变换几何. 这些变换模拟透视摄像机对平面摄像时所产生的几何失真. 在透视影像下,某些几何性质被保留,例如保线性(直线被影像为直线),而有的则不被保留,例如平行线一般不被影像成平行线. 射影几何为这类影像建模而且提供适于计算的数学表达.

我们先介绍在齐次标记下点、线和二次曲线的表示,以及在射影变换下这些几何实体如何映射. 接着介绍无穷远线和虚圆点,并证明它们控制了平面的仿射和度量性质. 然后,给出平面矫正的算法,这些算法实现了由图像来计算仿射和度量性质. 最后,我们介绍射影变换下的不动点.

## 1.1　平面几何

任何学习过初等数学的人都熟悉平面几何的基本概念. 事实上,甚至因为这些概念已成为我们日常生活的部分经验,因此我们总以为它们理所当然地成立. 在初等水平上,几何研究点和线以及它们的关系.

按几何纯传统论者的观点,几何研究应该坚持"几何的"或者说与坐标无关的观点. 在这种方法中,定理的叙述和证明仅使用几何的公理而不使用代数. 经典的欧氏方法就是其中的一个例子. 但是,自笛卡尔之后,人们认识到几何可以代数化,而且,几何理论的确可以从代数的观点来推导. 在本书中,我们将使用混合的方法,有时用几何的而有时用代数的方法. 在代数方法中,几何实体用坐标和代数实体描述. 例如,一个点等同于某坐标基下的一个矢量;一条直线也等同于一个矢量;而一段圆锥截线(或简称二次曲线)用一个对称矩阵表示. 事实上,至少因为语言上的方便,我们经常采用这样的等价表示,即矢量**就是**点,对称矩阵**就是**二次曲线. 在几何中采用代数方法的显著优点是:这种方法导出的结果更容易产生算法以及实际的计算方法. 本书主要关注的是计算和算法,它将证明使用代数方法是合理的.

## 1.2　2D 射影平面

众所周知,平面上的一点可以用 $\mathbb{R}^2$ 中的一对坐标 $(x,y)$ 来表示. 因此,通常 $\mathbb{R}^2$ 等同于一张平面. 把 $\mathbb{R}^2$ 看作一个矢量空间时,坐标对 $(x,y)$ 是矢量,也就是说点等同于矢量. 本节将引入平面上点和线的**齐次**表示.

**行和列矢量** 此后，我们将考虑矢量空间之间的线性映射并把这样的映射表示成矩阵. 在通常的方式下，一个矩阵和一个矢量的积是另一个矢量，它就是该映射下的像. 由此引出"列"与"行"矢量的区别. 因为矩阵可以被列矢量右乘或被行矢量左乘，在不加说明时，几何实体用列矢量表示. 粗体符号如 $\mathbf{x}$ 总表示列矢量，它的转置是行矢量 $\mathbf{x}^\mathsf{T}$. 按此约定，平面上的点将表示为列矢量 $(x,y)^\mathsf{T}$，而不是它的转置：行矢量 $(x,y)$. 我们记 $\mathbf{x}=(x,y)^\mathsf{T}$，该方程的两边都是列矢量.

### 1.2.1 点与直线

**直线的齐次表示** 平面上的一条直线可用形如 $ax+by+c=0$ 的方程表示，$a$，$b$ 和 $c$ 的不同值给出不同的直线. 因此，一条直线也可以用矢量 $(a,b,c)^\mathsf{T}$ 表示. 直线和矢量 $(a,b,c)^\mathsf{T}$ 不是一一对应的，因为，对任何非零常数 $k$，直线 $ax+by+c=0$ 与直线 $(ka)x+(kb)y+(kc)=0$ 相同. 因此，对任何非零 $k$，矢量 $(a,b,c)^\mathsf{T}$ 与 $k(a,b,c)^\mathsf{T}$ 表示同一直线. 事实上，我们视这两个只相差一个全局缩放因子的矢量是等价的. 这种等价关系下的矢量等价类被称为**齐次矢量**. 任何具体矢量 $(a,b,c)^\mathsf{T}$ 是所属的等价类的一个代表. 在 $\mathbb{R}^3-(0,0,0)^\mathsf{T}$ 中的矢量等价类的集合组成**射影空间** $\mathbb{P}^2$. 记号：$-(0,0,0)^\mathsf{T}$ 表示矢量 $(0,0,0)^\mathsf{T}$ 不与任何直线对应，因而被排除在外.

**点的齐次表示** 点 $\mathbf{x}=(x,y)^\mathsf{T}$ 在直线 $\mathbf{l}=(a,b,c)^\mathsf{T}$ 上的充要条件是 $ax+by+c=0$. 并可用矢量内积形式把它表示为 $(x,y,1)(a,b,c)^\mathsf{T}=(x,y,1)\mathbf{l}=0$；即把"1"作为增加的最后一个坐标使 $\mathbb{R}^2$ 中的点 $(x,y)^\mathsf{T}$ 表示成 3 维矢量. 注意对任何非零 $k$ 和直线 $\mathbf{l}$，方程 $(ka,kb,kc)\mathbf{l}=0$ 的充要条件是 $(a,b,1)\mathbf{l}=0$. 因而，我们自然把由 $k$ 不为零的不同值所构成的矢量集 $(kx,ky,k)^\mathsf{T}$ 看作是 $\mathbb{R}^2$ 中点 $(x,y)^\mathsf{T}$ 的一种表示. 因此，与直线一样，点也可用齐次矢量表示. 一个点的任何齐次矢量的表示形式为 $\mathbf{x}=(x_1,x_2,x_3)^\mathsf{T}$，并表示 $\mathbb{R}^2$ 的点 $\mathbf{x}=(x_1/x_3,x_2/x_3)^\mathsf{T}$. 于是，点作为齐次矢量同样也是 $\mathbb{P}^2$ 的元素.

我们得到一个用以确定点在不在直线上的简单方程，即

**结论 1.1 点 $\mathbf{x}$ 在直线 $\mathbf{l}$ 上的充要条件是 $\mathbf{x}^\mathsf{T}\mathbf{l}=0$.**

注意：表达式 $\mathbf{x}^\mathsf{T}\mathbf{l}$ 就是两矢量 $\mathbf{x}$ 和 $\mathbf{l}$ 的内积或标量积，即 $\mathbf{x}^\mathsf{T}\mathbf{l}=\mathbf{l}^\mathsf{T}\mathbf{x}=\mathbf{x}\cdot\mathbf{l}$. 一般，我们更喜欢采用转置记号 $\mathbf{l}^\mathsf{T}\mathbf{x}$，偶尔，也用一个"·"来表示内积. 注意区分一个点的**齐次坐标**：$\mathbf{x}=(x_1,x_2,x_3)^\mathsf{T}$，它是 3 维矢量；与**非齐次坐标**：$(x,y)^\mathsf{T}$，它是 2 维矢量.

**自由度**(dof) 显然，为了指定一个点必须提供两个值，即它的 $x$ - 和 $y$ - 坐标. 同样，一条直线由两个参数指定(两个独立的比率 $\{a:b:c\}$)因而有两个自由度. 例如，在非齐次表示中，这两个参数可以取为直线的梯度和 $y$ 轴上的截距.

**直线的交点** 给定两直线 $\mathbf{l}=(a,b,c)^\mathsf{T}$ 和 $\mathbf{l}'=(a',b',c')^\mathsf{T}$，我们希望求它们的交点. 定义矢量 $\mathbf{x}=\mathbf{l}\times\mathbf{l}'$，这里 $\times$ 表示矢量积或叉积. 由三重纯量积等式：$\mathbf{l}\cdot(\mathbf{l}\times\mathbf{l}')=\mathbf{l}'\cdot(\mathbf{l}\times\mathbf{l}')=0$，推出 $\mathbf{l}^\mathsf{T}\mathbf{x}=\mathbf{l}'^\mathsf{T}\mathbf{x}=0$. 因此，如果把 $\mathbf{x}$ 视为一个点，则 $\mathbf{x}$ 同时在两条直线 $\mathbf{l}$ 和 $\mathbf{l}'$ 上，因而是两线的交点. 这表明：

**结论 1.2 两直线 $\mathbf{l}$ 和 $\mathbf{l}'$ 的交点是点 $\mathbf{x}=\mathbf{l}\times\mathbf{l}'$.**

注意：两直线交点的表示之所以这样简洁是因为采用了直线和点的齐次坐标表示.

**例 1.3** 考虑一个简单问题：求直线 $x=1$ 和 $y=1$ 的交点. 直线 $x=1$ 等价于 $-1x+1=0$，故有齐次表示 $\mathbf{l}=(-1,0,1)^{\mathsf{T}}$. 直线 $y=1$ 等价于 $-1y+1=0$，齐次表示是 $\mathbf{l}'=(0,-1,1)^{\mathsf{T}}$. 由结论 1.2 得到交点为

$$\mathbf{x}=\mathbf{l}\times\mathbf{l}'=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ -1 & 0 & 1\\ 0 & -1 & 1\end{vmatrix}=\begin{pmatrix}1\\1\\1\end{pmatrix}$$

它是非齐次点 $(1,1)^{\mathsf{T}}$，正是我们所要求的. △

**点的连线** 过两点 $\mathbf{x}$ 和 $\mathbf{x}'$ 的直线的表示式可完全类似地导出. 定义直线 $\mathbf{l}=\mathbf{x}\times\mathbf{x}'$，并不难验证点 $\mathbf{x}$ 和 $\mathbf{x}'$ 都在 $\mathbf{l}$ 上. 因此

**结论 1.4 过两点 $\mathbf{x}$ 和 $\mathbf{x}'$ 的直线是 $\mathbf{l}=\mathbf{x}\times\mathbf{x}'$.**

### 1.2.2 理想点与无穷远线

**平行线的交点** 考察两直线 $ax+by+c=0$ 和 $ax+by+c'=0$. 它们可分别用矢量 $\mathbf{l}=(a,b,c)^{\mathsf{T}}$ 和 $\mathbf{l}'=(a,b,c')^{\mathsf{T}}$ 表示，其中它们的前两个坐标是一样的. 用结论 1.2 不难算出这两条直线的交点为 $\mathbf{l}\times\mathbf{l}'=(c'-c)(b,-a,0)^{\mathsf{T}}$，忽略标量因子 $(c'-c)$，得到点 $(b,-a,0)^{\mathsf{T}}$.

现在，如果我们试图求这一点的非齐次表示，就会得到 $(b/0,\ a/0)^{\mathsf{T}}$，它只能解释为该交点有无穷大坐标. 一般地，具有齐次坐标 $(x,y,0)^{\mathsf{T}}$ 的点不与 $\mathbb{R}^2$ 中任何有限点对应. 这一观察与通常平行线交于无穷远的概念相吻合.

**例 1.5** 考察两直线 $x=1$ 和 $x=2$. 这两线平行，因而交于"无穷远点". 利用齐次表示将它们表示为 $\mathbf{l}=(-1,0,1)^{\mathsf{T}}$ 和 $\mathbf{l}'=(-1,0,2)^{\mathsf{T}}$，再由结论 1.2 求得其交点为

$$\mathbf{x}=\mathbf{l}\times\mathbf{l}'=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ -1 & 0 & 1\\ -1 & 0 & 2\end{vmatrix}=\begin{pmatrix}0\\1\\0\end{pmatrix}$$

这是在 $y$－轴方向上的无穷远点. △

**理想点与无穷远线** 当 $x_3\neq 0$ 时，齐次矢量 $\mathbf{x}=(x_1,x_2,x_3)^{\mathsf{T}}$ 对应于 $\mathbb{R}^2$ 中的有限点. 我们可以把最后坐标为 $x_3=0$ 的点加入 $\mathbb{R}^2$. 所扩展的空间是所有齐次 3 维矢量的集合，称为射影空间 $\mathbb{P}^2$. 最后坐标为 $x_3=0$ 的点被称为**理想点**，或无穷远点. 所有理想点的集合可以写成 $(x_1,x_2,0)^{\mathsf{T}}$，并由比率 $x_1:x_2$ 指定一个具体的点. 注意该集合在一条直线上，称**无穷远线**，用矢量 $\mathbf{l}_\infty=(0,0,1)^{\mathsf{T}}$ 表示. 我们可以验证 $(0,0,1)(x_1,x_2,0)^{\mathsf{T}}=0$.

由结论 1.2，我们推出直线 $\mathbf{l}=(a,b,c)^{\mathsf{T}}$ 与 $\mathbf{l}_\infty$ 交于理想点 $(b,-a,0)^{\mathsf{T}}$（因为 $(b,-a,0)\mathbf{l}=0$）. 任何一条与 $\mathbf{l}$ 平行的直线 $\mathbf{l}'=(a,b,c')^{\mathsf{T}}$ 也交 $\mathbf{l}_\infty$ 于同样的理想点 $(b,-a,0)^{\mathsf{T}}$，与 $c'$ 的取值无关. 在非齐次表示下，$(b,-a)^{\mathsf{T}}$ 是与该直线相切的矢量，与该直线的法线 $(a,b)^{\mathsf{T}}$ 相正交，因而它代表该直线的**方向**. 当直线的方向改变时，理想点 $(b,-a,0)^{\mathsf{T}}$ 沿 $\mathbf{l}_\infty$ 而变化. 基于这些理由，无穷远线可以看作是平面上所有直线方向的集合.

注意引入无穷远点概念后点与直线相交的概念得到了简化. 在射影平面 $\mathbb{P}^2$中,我们可以不加思索地说任意两条相异直线都相交于一点而任意两个相异的点都在一条直线上. 但在标准欧氏几何 $\mathbb{R}^2$中却不成立,其中平行线就构成一个特例.

研究 $\mathbb{P}^2$的几何称为射影几何. 在无坐标的纯几何研究中,射影几何的无穷远点(理想点)和普通点没有任何区别. 但在本书中,为了达到我们的目的,有时将区别理想点和非理想点. 因此,无穷远线有时被看成射影空间中的一条特殊直线.

**射影平面的模型** 一种有益的方法是把 $\mathbb{P}^2$看作 $\mathbb{R}^3$中一种射线的集合. 该集合的所有矢量 $k(x_1, x_2, x_3)^{\mathrm{T}}$当 $k$ 变化时形成过原点的射线. 这样的一条射线可以看作是 $\mathbb{P}^2$中的一个点. 在此模型中,$\mathbb{P}^2$中的直线是过原点的平面. 可以验证两相异的射线共处于一张平面上,而任何两张相异平面相交于一条射线. 这类似于两个相异的点唯一确定一条直线,而两条相异的直线总相交于一点.

点和线可以用平面 $x_3=1$ 与这些射线和平面集相交得到. 如图 1.1 所示,表示理想点的射线和表示 $\mathbf{l}_\infty$ 的平面都与平面 $x_3=1$ 平行.

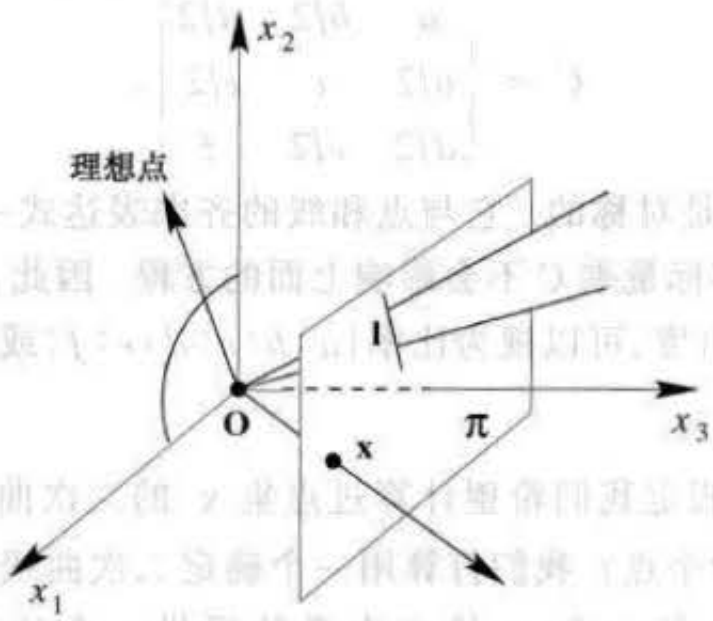



图 1.1 **射影平面的模型**. $\mathbb{P}^2$的点和线分别表示为 $\mathbb{R}^3$中过原点的射线和平面. $x_1x_2$ - 平面上的射线表示理想点,而 $x_1x_2$ - 平面表示 $\mathbf{l}_\infty$.

**对偶** 在关于线和点的性质的陈述中,读者也许已经注意到点和线的作用可以怎样互换. 特别是,直线和点的基本接合方程式 $\mathbf{l}^{\mathrm{T}}\mathbf{x}=0$ 是对称的,因为 $\mathbf{l}^{\mathrm{T}}\mathbf{x}=0$ 意味着 $\mathbf{x}^{\mathrm{T}}\mathbf{l}=0$,其中线和点的位置互相交换了. 类似地,两线相交和一线过两点的结论 1.2 和 1.4,当把点和线的作用互换后,其本质是一样的. 由此得到一个如下的一般原理:**对偶原理**:

**结论 1.6 对偶原理** **2 维射影几何中的任何定理都有一个对应的对偶定理,它可以通过互换原定理中点和线的作用而导出.**

在应用此原理时,接合的概念也必须做适当地转换. 例如,过两点的直线与过两直线的点(即两线的交点)对偶.

注意:一旦原定理已被证明,就没有必要再去证明它的对偶. 对偶定理的证明不过是原定理证明的对偶.

### 1.2.3 二次曲线与对偶二次曲线

二次曲线由平面上的二阶方程描述. 在欧氏几何中,二次曲线有三种主要类型:双曲线,椭圆和抛物线(后面将定义的所谓的退化二次曲线除外). 在经典理论中,这三类曲线是不同方向的平面与圆锥相交所产生的截线(退化的二次曲线由过锥顶的平面产生). 但是我们将了解到,在 2D 射影几何中所有非退化的二次曲线在射影变换下都等价.

在非齐次坐标中,二次曲线的方程是

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

对应于一个二阶多项式. 通过替代 $x \mapsto x_1/x_3$, $y \mapsto x_2/x_3$,"齐次化"得到

$$ax_1^2 + bx_1x_2 + cx_2^2 + dx_1x_3 + ex_2x_3 + fx_3^2 = 0 \tag{1.1}$$

或表示成矩阵形式

$$\mathbf{x}^{\mathrm{T}}\mathbf{C}\mathbf{x} = 0 \tag{1.2}$$

其中二次曲线系数矩阵 C 为

$$C = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}. \tag{1.3}$$

注意二次曲线的系数矩阵是对称的. 它与点和线的齐次表达式一样,重要的仅仅是矩阵元素的比率,因为用一个非零标量乘 C 不会影响上面的方程. 因此,C 是一条二次曲线的齐次表示. 二次曲线有五个自由度,可以视为比率 $|a:b:c:d:e:f|$ 或等价地视为对称矩阵六个元素减去一个比例因子.

**五点定义一条二次曲线** 假定我们希望计算过点集 $\mathbf{x}_i$ 的二次曲线. 在二次曲线被唯一确定之前,我们可以指定多少个点? 我们打算用一个确定二次曲线的算法来构造性地回答这个问题. 根据(1.1)可知,每一点 $\mathbf{x}_i$ 给二次系数提供一个约束,因为如果二次曲线过 $(x_i, y_i)^{\mathrm{T}}$,便有

$$ax_i^2 + b_ix_iy_i + cy_i^2 + dx_i + ey_i + f = 0$$

该约束可以重写为

$$(x_i^2 \quad x_iy_i \quad y_i^2 \quad x_i \quad y_i \quad 1)\mathbf{c} = 0$$

其中 $\mathbf{c} = (a, b, c, d, e, f)^{\mathrm{T}}$ 是把二次曲线 C 表示成一个 6 维矢量.

把五点提供的约束堆积起来,得到

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \end{bmatrix}\mathbf{c} = \mathbf{0} \tag{1.4}$$

因而二次曲线是这个 5×6 矩阵的零矢量. 它表明一条二次曲线由一般位置的五个点唯一确定(相差一个尺度因子). 这种通过求零空间来拟合一个几何实体(或关系)的方法今后将经常在本书介绍计算的章节中采用.

**二次曲线的切线** 在齐次坐标下，过二次曲线上点 $\mathbf{x}$ 的切线 $\mathbf{l}$ 有特别简单的形式：

**结论 1.7 过(非退化)二次曲线 C 上点 $\mathbf{x}$ 的切线 $\mathbf{l}$ 由 $\mathbf{l}=C\mathbf{x}$ 确定.**

**证明** 因为 $\mathbf{l}^T\mathbf{x}=\mathbf{x}^TC\mathbf{x}=0$，所以直线 $\mathbf{l}=C\mathbf{x}$ 过 $\mathbf{x}$. 如果 $\mathbf{l}$ 仅与二次曲线交于一点，那么它就是切线，我们的证明也就完成了. 否则，假定 $\mathbf{l}$ 与该二次曲线还交于另一点 $\mathbf{y}$，则有 $\mathbf{y}^TC\mathbf{y}=0$ 和 $\mathbf{x}^TC\mathbf{y}=\mathbf{l}^T\mathbf{y}=0$. 由此推出 $(\mathbf{x}+\alpha\mathbf{y})^TC(\mathbf{x}+\alpha\mathbf{y})=0$ 对所有 $\alpha$ 成立，这表明连接 $\mathbf{x}$ 和 $\mathbf{y}$ 的整条直线 $\mathbf{l}=C\mathbf{x}$ 在该二次曲线 C 上，因而 C 是退化的(见下文). □

**对偶二次曲线** 上面定义的二次曲线 C 更确切地应称为**点**二次曲线，因为它定义的是点的方程. 给出了 $\mathbb{P}^2$ 的对偶结论 1.6 以后，理所当然应该有一个由直线的方程定义的二次曲线. 这种**对偶**(或线)二次曲线也由一个 $3\times3$ 矩阵表示，我们把它记为 $C^*$. 二次曲线 C 的切线 $\mathbf{l}$ 满足 $\mathbf{l}^TC^*\mathbf{l}=0$. 其中 $C^*$ 表示 C 的伴随矩阵(伴随矩阵在附录 3($p$404)的节 A3.2($p$406)中定义). 对非奇异对称矩阵 C 有 $C^*=C^{-1}$(相差一个尺度因子).

当 C 为满秩时，对偶二次曲线方程可直接推导：根据结论 1.7，过 C 上点 $\mathbf{x}$ 的切线是 $\mathbf{l}=C\mathbf{x}$. 反之，直线 $\mathbf{l}$ 切于 C 的点 $\mathbf{x}$ 是 $\mathbf{x}=C^{-1}\mathbf{l}$. 因为 $\mathbf{x}$ 满足 $\mathbf{x}^TC\mathbf{x}=0$，我们得到，$(C^{-1}\mathbf{l})^TC(C^{-1}\mathbf{l})=\mathbf{l}^TC^{-1}\mathbf{l}=0$，因为 C 是对称的，所以 $C^{-T}=C^{-1}$，因而得到最后的结果.

对偶二次曲线也称二次曲线的包络，其理由在图 1.2 中给予说明. 对偶二次曲线有五个自由度. 与点二次曲线相类似，一般位置上的五条线定义一条对偶二次曲线.

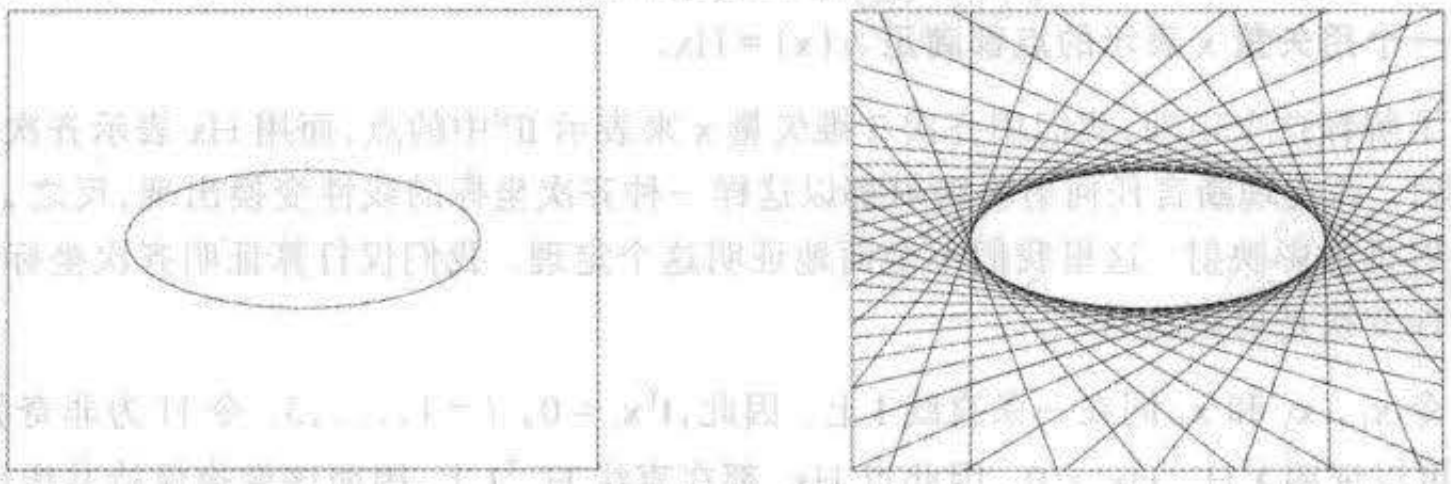


图 1.2 (a)满足 $\mathbf{x}^TC\mathbf{x}=0$ 的点 $\mathbf{x}$ 在一条点二次曲线上. (b)满足 $\mathbf{l}^TC^*\mathbf{l}=0$ 的直线 $\mathbf{l}$ 是点二次曲线 C 的切线，即二次曲线 C 是直线 $\mathbf{l}$ 的包络.

**退化二次曲线** 非满秩矩阵 C 所定义的二次曲线称作退化二次曲线. 退化的点二次曲线包含两条线(秩 2)或一条重线(秩 1).

**例 1.8** 二次曲线

$$C=\mathbf{l}\mathbf{m}^T+\mathbf{m}\mathbf{l}^T$$

由 $\mathbf{l}$ 和 $\mathbf{m}$ 两线组成. $\mathbf{l}$ 上的点满足 $\mathbf{l}^T\mathbf{x}=0$，因而在二次曲线上，因为 $\mathbf{x}^TC\mathbf{x}=(\mathbf{x}^T\mathbf{l})(\mathbf{m}^T\mathbf{x})+(\mathbf{x}^T\mathbf{m})(\mathbf{l}^T\mathbf{x})=0$. 类似地，满足 $\mathbf{m}^T\mathbf{x}=0$ 的点同样满足 $\mathbf{x}^TC\mathbf{x}=0$. 矩阵 C 是秩为 2 的对称矩阵，它的零矢量为 $\mathbf{x}=\mathbf{l}\times\mathbf{m}$，它是 $\mathbf{l}$ 和 $\mathbf{m}$ 的交点. △

退化的**线**二次曲线包含两个点(秩 2)，或一个重点(秩 1). 例如，线二次曲线 $C^*=\mathbf{x}\mathbf{y}^T+\mathbf{y}\mathbf{x}^T$ 的秩为 2 并由一切过点 $\mathbf{x}$ 或 $\mathbf{y}$ 的直线组成. 注意对非可逆矩阵而言，$(C^*)^*\neq C$.

## 1.3 射影变换

Felix Klein 在其名著："Erlangen Program"[Klein-39]中提出：几何研究的是在变换群下保持不变的性质. 根据他的观点，2D 射影几何研究的是关于射影平面 $\mathbb{P}^2$在所谓**射影映射**的变换群下保持不变的性质.

射影映射是把 $\mathbb{P}^2$的点(即齐次 3 维矢量)映射到 $\mathbb{P}^2$的点的一种可逆映射，它把直线映射到直线. 更准确地说：

**定义 1.9** **射影映射**是 $\mathbb{P}^2$到它自身的一种满足下列条件的可逆映射 $h$：三点 $\mathbf{x}_1$，$\mathbf{x}_2$ 和 $\mathbf{x}_3$ 共线当且仅当 $h(\mathbf{x}_1)$，$h(\mathbf{x}_2)$，$h(\mathbf{x}_3)$也共线.

射影映射组成一个群，因为射影映射的逆以及两个射影映射的复合也是射影映射. 射影映射也称为**保线变换**(一个有益的名字)，或**射影变换**或**单应**(***homography***)，它们是同义术语.

在定义 1.9 中，射影映射用点线关联的几何概念来定义，它与坐标无关. 基于下列定理，我们也可得到射影映射的等价的代数定义.

**定理 1.10** **映射 $h:\mathbb{P}^2\to\mathbb{P}^2$是射影映射的充要条件是：存在一个 $3\times3$ 非奇异矩阵 H，使得$\mathbb{P}^2$的任何一个用矢量 $\mathbf{x}$ 表示的点都满足 $h(\mathbf{x})=\mathrm{H}\mathbf{x}$.**

为了解释这个定理，我们用齐次 3 维矢量 $\mathbf{x}$ 来表示 $\mathbb{P}^2$中的点，而用 $\mathrm{H}\mathbf{x}$ 表示齐次坐标的线性映射. 该定理断言任何射影映射都以这样一种齐次坐标的线性变换出现，反之，任何这样的映射是射影映射. 这里我们不全面地证明这个定理. 我们仅打算证明齐次坐标的任何可逆线性变换是射影映射.

**证明** 令 $\mathbf{x}_1$，$\mathbf{x}_2$ 和 $\mathbf{x}_3$ 同在一条直线 $\mathbf{l}$ 上. 因此，$\mathbf{l}^{\mathrm{T}}\mathbf{x}_i=0$，$i=1,\ldots,3$. 令 H 为非奇异 $3\times3$ 矩阵. 可以证明 $\mathbf{l}^{\mathrm{T}}\mathrm{H}^{-1}\mathrm{H}\mathbf{x}_i=0$. 因此点 $\mathrm{H}\mathbf{x}_i$ 都在直线 $\mathrm{H}^{-\mathrm{T}}\mathbf{l}$ 上，因而该变换保持共线性.

其逆命题是每个射影映射都以这种方式出现，但相当难证明. □

根据这个定理，可以给出射影变换(或保线变换)的另一种定义.

**定义 1.11** **射影变换** 平面射影变换是关于齐次 3 维矢量的一种线性变换，并可用一个非奇异 $3\times3$ 矩阵 H 表示为：

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{1.5}$$

或更简洁地表示为 $\mathbf{x}'=\mathrm{H}\mathbf{x}$.

注意：此方程中的矩阵 H 乘以任意一个非零比例因子不会使射影变换改变. 换句话说 H 是一个**齐次**矩阵，与点的齐次表示一样，有意义的仅仅是矩阵元素的比率. 在 H 的九个元素中有八个独立比率，因此一个射影变换有八个自由度.

射影变换将每个图形投影为射影等价的图形，保持所有的射影性质不变. 在图 1.1 的射线模型中，一个射影变换就是 $\mathbb{R}^3$的一种线性变换.

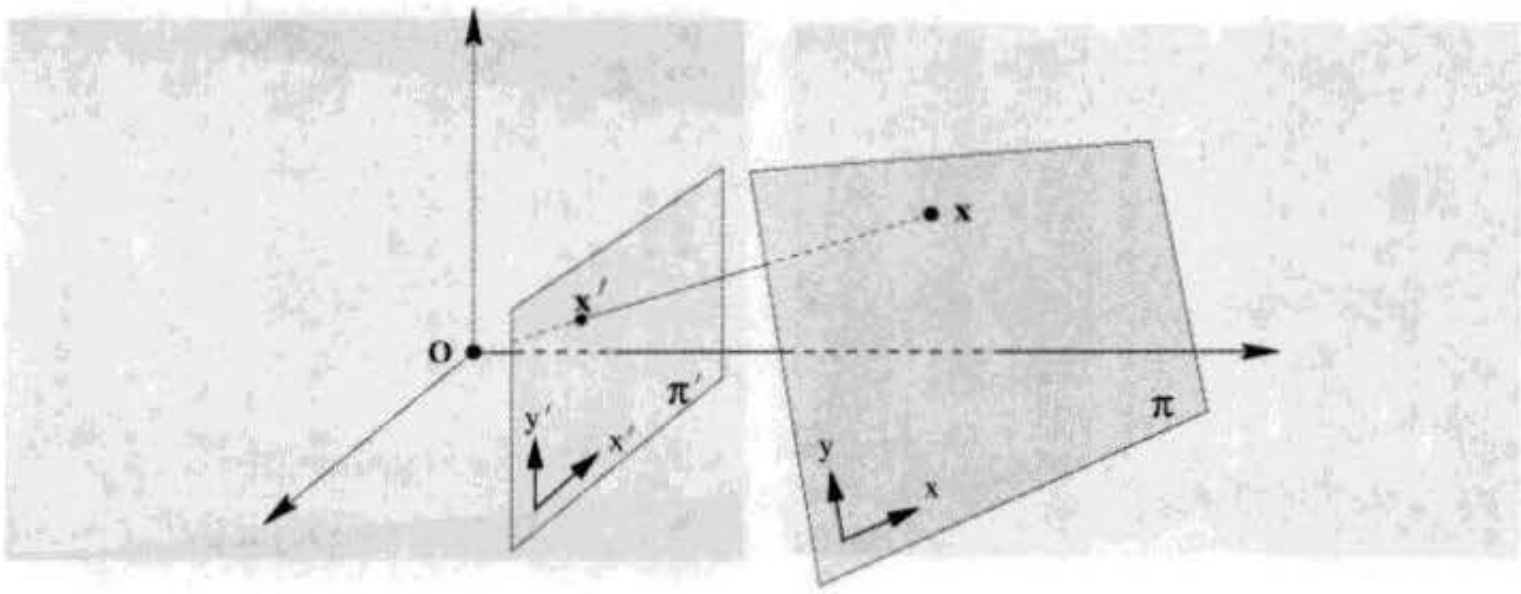



图 1.3 **中心投影把一张平面的点映射成另一张平面的点**. 过投影中心作一张与两平面 $\pi$ 和 $\pi'$ 相交的平面便不难看出中心投影也把直线映射为直线. 因为直线被映射为直线,所以中心投影是射影变换并可用齐次坐标的线性变换 $\mathbf{x}' = \mathrm{H}\mathbf{x}$ 表示.

**平面之间的映射** 图 1.3 中给出怎样应用定理 1.10 的一个例子. 沿过一个公共点(投影中心)的射线的投影定义了从一张平面到另一张平面的映射. 显然这种点到点的映射保持直线不变,其中一张平面上的直线被映射到另一张平面上的直线. 如果在每一张平面上建立坐标系并且采用齐次坐标来表示点,那么**中心投影映射可以表示成 $\mathbf{x}' = \mathrm{H}\mathbf{x}$**,其中 H 是 $3\times3$ 非奇异矩阵. 实际上,如果在两平面上建立的都是欧氏(直角)坐标系,那么这样一种由中心投影定义的映射比一般射影变换有更多的约束. 我们称它为**透视映射**而不是完全的射影映射,它可由一个六自由度的变换来表示. 我们将在节 A5.4($p$430)中再回过来讨论透视映射.

**例 1.12 消除平面透视图像的射影失真.**

在透视影像下形状会失真. 例如,虽然原有的窗子是矩形,但在图 1.4a 中的图像上的窗子不是矩形. 景物平面上的平行线在图像上一般地不平行并会聚到一个有限点. 我们已经知道平面(或部分平面)的中心投影的图像与原平面通过射影变换相关,因而该图像是原景物的一种射影失真. 通过求该射影变换的逆变换并把它应用于图像就有可能"撤消"此射影变换. 它的结果将是一幅新的合成图像,其中在此平面上的物体将显示其正确的几何形状. 我们用图 1.4a 大楼的前墙对这做解释. 注意因为地面和大楼的前墙不在同一平面上,所以用于矫正大楼前墙的射影变换必然与用于地面的射影变换不一样.

我们将在第三章详细讨论由点到点的对应来求射影变换的计算. 现在,我们仅简洁地给出一种计算该变换的方法. 首先选择世界平面与图像相对应的部分. 按图 1.3 所指示选择图像平面的 2D 局部坐标以及景物的世界坐标. 令世界与图像平面上的一对匹配点 $\mathbf{x}$ 和 $\mathbf{x}'$ 的非齐次坐标分别为 $(x, y)^{\mathrm{T}}$ 和 $(x', y')^{\mathrm{T}}$. 这里我们采用点的非齐次坐标而不是齐次坐标,因为从图像和世界平面的测量可以直接得到的是这些非齐次坐标. (1.5)式的射影变换可以写成如下非齐次形式:

$$x' = \frac{x'_1}{x'_3} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \qquad y' = \frac{x'_2}{x'_3} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}.$$

一组点对应可以提供关于 H 元素的两个方程,把它简化后得到:

$$x'(h_{31}x + h_{32}y + h_{33}) = h_{11}x + h_{12}y + h_{13}$$

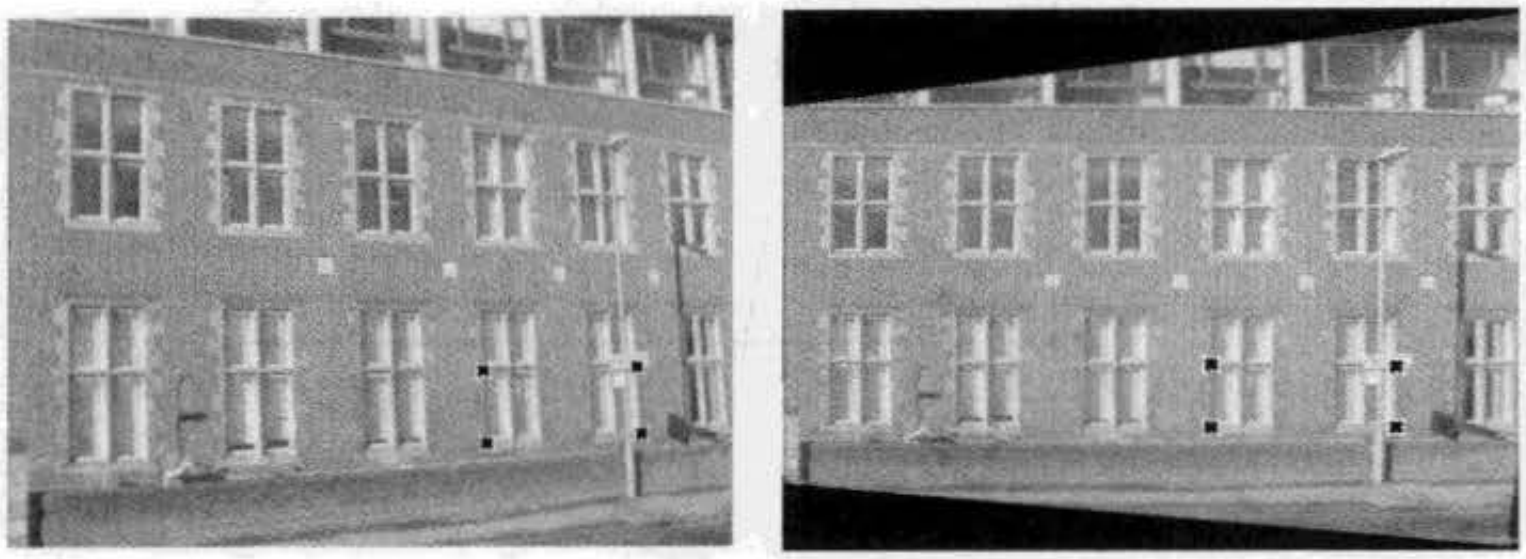


图 1.4 **消除透视失真**. (a)带有透视失真的原始图像——窗子的上下边线显然会聚(即延长后相交)于一个有限点. (b)合成得到的前墙的正视图. 墙的图像(a)与墙的真实几何通过一个射影变换相联系. 逆变换是把影像得到的四个窗角的角点映射到适当大小的矩形顶点并由此计算得到. 此变换由这四组点对应确定. 再用此变换作用于整幅图像. 注意图像中的地面部分出现进一步的射影失真. 它同样可以通过一个射影变换加以消除.

$$y'(h_{31}x + h_{32}y + h_{33}) = h_{21}x + h_{22}y + h_{23}.$$

它们是关于 H 元素的**线性**方程. 四组点对应提供八个这样的关于 H 元素的线性方程,并足以解出 H(仅相差一个不重要的乘法因子). 唯一的限制是这四点必须在"一般位置"上,即要求没有三个点共线. 用此方法计算变换 H 并求它的逆,然后作用于整幅图像,便可消除被选平面的射影失真效应. 其结果在图 1.4b 中给出. △

针对这个例子,我们给出三点说明:第一,用这种方法计算矫正变换 H 不需要知道摄像机参数或平面位置的**任何**信息;第二,为消除射影失真,并非总需要知道四个点的坐标:节 1.7将介绍另一种方法,它需要较少但不同类型的信息;第三,高级的(和推荐的)计算射影变换的方法将在第三章中阐述.

射影变换是一种重要的映射,与世界平面的透视成像相比较,它能表示更多的情形. 图 1.5中给出了其他的几个例子,我们将在本书后续的篇章中对其中的每种情形作更详尽的讨论.

### 1.3.1 直线与二次曲线的变换

**直线的变换** 定理 1.10 的证明指出:如果点 $\mathbf{x}_i$ 在直线 $\mathbf{l}$ 上,那么经射影变换后的点 $\mathbf{x}'_i = \mathbf{H}\mathbf{x}_i$ 在直线 $\mathbf{l}' = \mathbf{H}^{-\mathsf{T}}\mathbf{l}$ 上. 因为 $\mathbf{l}'^{\mathsf{T}}\mathbf{x}'_i = \mathbf{l}^{\mathsf{T}}\mathbf{H}^{-1}\mathbf{H}\mathbf{x}_i = 0$,故变换 H 保持点在直线上的性质. 它给出了关于直线的变换规则:在点变换 $\mathbf{x}'_i = \mathbf{H}\mathbf{x}_i$ 下,直线 $\mathbf{l}$ 变换为

$$\boxed{\mathbf{l}' = \mathbf{H}^{-\mathsf{T}}\mathbf{l}.} \tag{1.6}$$

我们也可以换一种写法 $\mathbf{l}'^{\mathsf{T}} = \mathbf{l}^{\mathsf{T}}\mathbf{H}^{-1}$. 注意直线和点变换的基本区别. 点变换依据 H,而直线(视为行矢量)变换则依据 $\mathbf{H}^{-1}$. 这可以用术语"协变"或"逆变"做解释. 我们称点变换为**逆变**而线变换为**协变**. 这种区别将在第 14 章讨论张量时重新提到并在附录 1($p$400)中全面地给予解释.

**二次曲线的变换** 在点变换 $\mathbf{x}' = \mathbf{H}\mathbf{x}$ 下,(1.2)变为

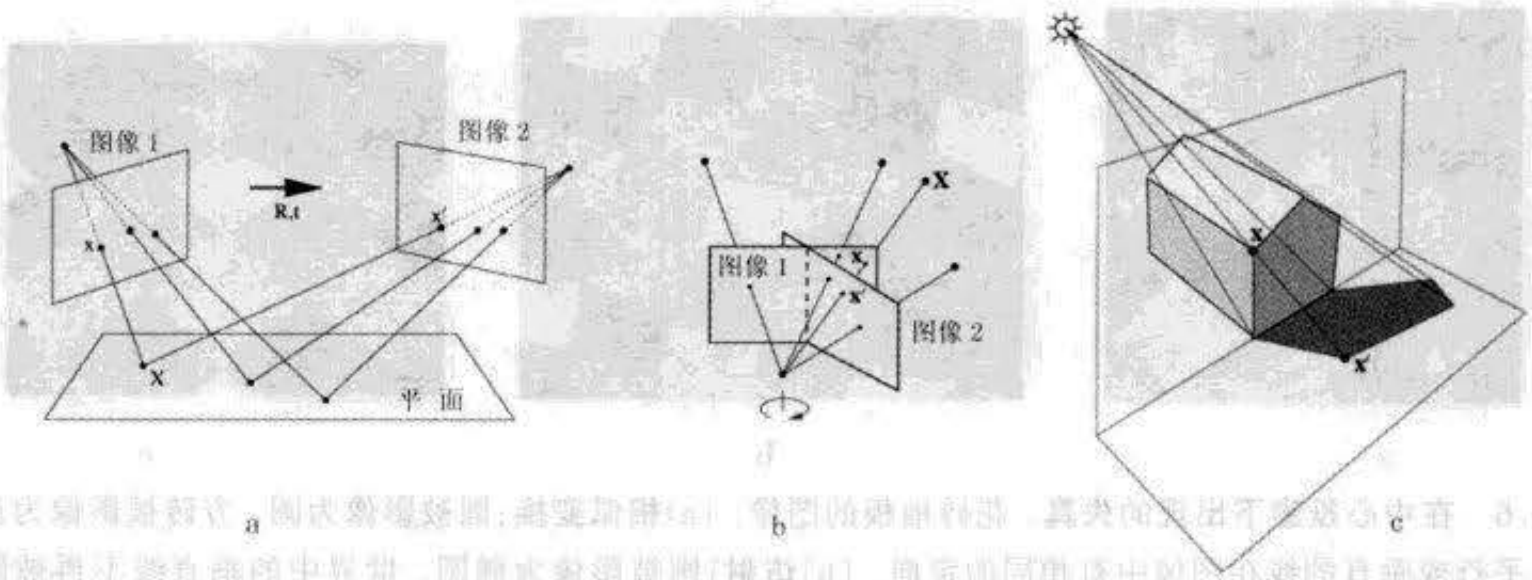



图 1.5 **出现在透视图像中的几个射影变换($\mathbf{x}'=\mathrm{H}\mathbf{x}$)的例子**. (a)由一张世界平面诱导的两幅图像之间的射影变换(两射影变换的复合是射影变换);(b)摄像中心相同的两幅图像之间的射影变换(即一个摄像机绕它的中心旋转或变化它的焦距);(c)一张平面的图像(大楼的后墙)和它的阴影在另一张平面上(地平面)的图像之间的射影变换. 图(c)取自 *Luc Van Gool*.

$$\begin{aligned}\mathbf{x}^{\mathrm{T}}\mathrm{C}\mathbf{x} &= \mathbf{x}'^{\mathrm{T}}[\mathrm{H}^{-1}]^{\mathrm{T}}\mathrm{C}\mathrm{H}^{-1}\mathbf{x}' \\ &= \mathbf{x}'^{\mathrm{T}}\mathrm{H}^{-\mathrm{T}}\mathrm{C}\mathrm{H}^{-1}\mathbf{x}'\end{aligned}$$

它是一种二次形式 $\mathbf{x}'^{\mathrm{T}}\mathrm{C}'\mathbf{x}'$,其中 $\mathrm{C}'=\mathrm{H}^{-\mathrm{T}}\mathrm{C}\mathrm{H}^{-1}$. 由此得到二次曲线变换的规则:

**结论 1.13 在点变换 $\mathbf{x}'=\mathrm{H}\mathbf{x}$ 下,二次曲线 C 变换为 $\mathrm{C}'=\mathrm{H}^{-\mathrm{T}}\mathrm{C}\mathrm{H}^{-1}$.**

因 $\mathrm{H}^{-1}$ 出现在方程中,故可以称二次曲线变换为**协变**. 对偶二次曲线的变换规则可用类似的方式导出,即:

**结论 1.14 在点变换 $\mathbf{x}'=\mathrm{H}\mathbf{x}$ 下,对偶二次曲线 $\mathrm{C}^*$ 变换为 $\mathrm{C}^{*\prime}=\mathrm{H}\mathrm{C}^*\mathrm{H}^{\mathrm{T}}$.**

## 1.4 变换的层次

我们将在本节中介绍射影变换的重要特殊情况以及它们的几何性质. 节 1.3 已经指出射影变换组成一个群. 这个群被称为**射影线性群**,我们将会看到这些特殊情况都是该群的**子群**.

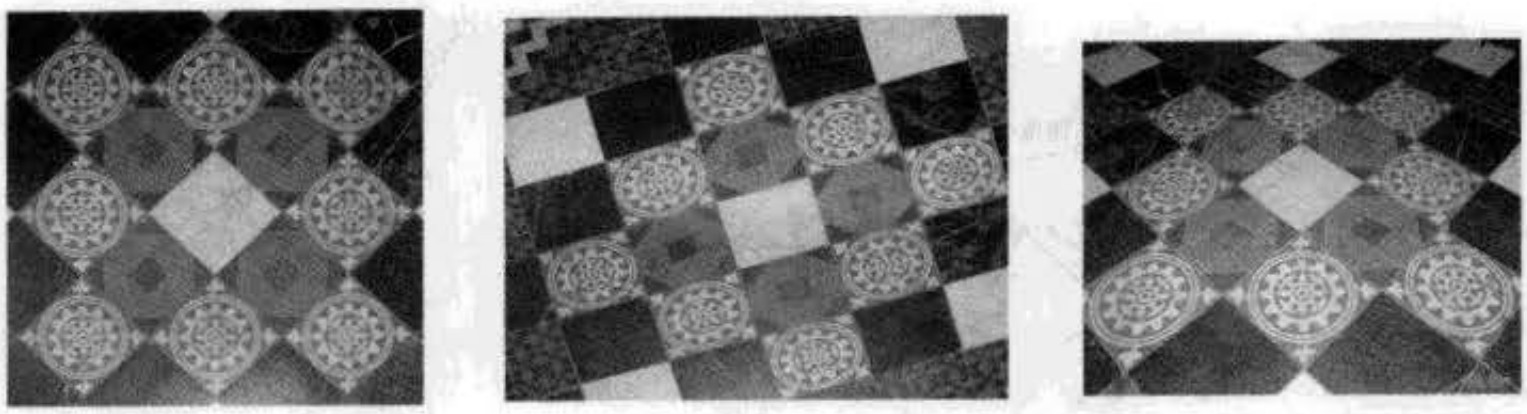
$n\times n$ 可逆实矩阵的群称为(实的)一般线性群或 $GL(n)$. 当把相差非零纯量因子的矩阵都视为等同时,便得到射影线性群,记为 $PL(n)$(它是 $GL(n)$的商群). 在平面射影变换时,$n=3$.

$PL(3)$的重要子群包括**仿射群和欧氏群**,仿射群是由 $PL(3)$中最后一行为(0,0,1)的矩阵组成的子群;欧氏群是仿射群的子群,其左上角的 $2\times2$ 矩阵是正交的. 当左上角的 $2\times2$矩阵的行列式为 1 时称为**定向欧氏群**.

我们将介绍这些变换,从最特殊的等距变换开始,并逐步推广直到射影变换. 由此定义了变换的一个**层次**. 图 1.6 给出属于不同层面上的各种变换的失真效果.

某些有趣的变换,例如透视映射,不是群(因为两个透视映射的复合是射影映射而不是透射映射.

这方面的内容在节 A5.4($p$430)中讨论.



a b c

图 1.6 **在中心投影下出现的失真**. 花砖地板的图像. (a)**相似变换**:圆被影像为圆. 方砖被影像为正方形. 平行或垂直的线在图像中有相同的定向. (b)**仿射**:圆被影像为椭圆. 世界中的垂直线不再被影像为垂直线. 但是,在世界中的方砖的平行边在图像中仍平行. (c)**射影**:平行世界线被影像为会聚线. 离摄像机近的方砖的图像比远的大.

**不变量** 介绍变换的另一种代数方法(即把变换视为作用于点或曲线的坐标的矩阵)是用被保持不变的元素或量即所谓的**不变量**来介绍. 一个几何配置的一个(标量)不变量是该配置的函数,其值在某特殊的变换下不变. 例如,两点间的距离在欧氏变换(平移和旋转)下不变,但在相似变换(即平移,旋转和均匀缩放)下则不然. 因此,距离是欧氏不变量但不是相似不变量. 两线间的夹角既是欧氏又是相似不变量.

### 1.4.1 类 I: 等距变换

等距(isometric)变换是平面 $\mathbb{R}^2$ 的变换,它保持欧氏距离不变($iso$ = 一样, $metric$ = 度量). 一个等距变换可表示为

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} \varepsilon\cos\theta & -\sin\theta & t_x \\ \varepsilon\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

其中 $\varepsilon = \pm 1$. 如果 $\varepsilon = 1$ 那么该等距变换是**保向的**并且也是**欧氏**变换(平移和旋转的复合). 如果 $\varepsilon = -1$,那么该等距变换是逆向的. 例如反射(由对角矩阵 diag(−1,1,1) 表示)与欧氏变换的复合.

欧氏变换是刚体运动的模型. 到目前为止,它们是实用中最重要的等距变换,我们将集中研究它们. 但是,保向等距变换在结构恢复时常会出现多义性.

平面欧氏变换可以用更简洁的分块形式写为

$$\mathbf{x}' = \mathrm{H}_{\mathrm{E}}\mathbf{x} = \begin{bmatrix} \mathrm{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix}\mathbf{x} \tag{1.7}$$

其中 R 是 2×2 旋转矩阵(满足 $\mathrm{R}^{\mathsf{T}}\mathrm{R} = \mathrm{R}\mathrm{R}^{\mathsf{T}} = \mathrm{I}$ 的正交矩阵),$\mathbf{t}$ 是 2 维平移矢量,而 $\mathbf{0}$ 是 2 维零矢量. 特殊的情况是纯旋转(当 $\mathbf{t} = \mathbf{0}$)和纯平移(当 $\mathrm{R} = \mathrm{I}$). 欧氏变换也称为**位移**.

平面欧氏变换有三个自由度:旋转占一个,平移占两个. 因此,确定该变换必须指定三个参量. 该变换可以由两组点对应来计算.

**不变量** 我们熟知它的不变量,例如,长度(两点的距离),角度(两线的夹角)和面积.

**群和定向** 如果等距变换左上角的 2×2 矩阵的行列式为 1,它是保向的. **保向**的等距变换形成一个群,但**逆向**的不是. 这种区别对于下面的相似和仿射变换同样如此.

### 1.4.2 类 II：相似变换

相似变换是一个等距变换与一个均匀缩放的复合．当欧氏变换(即没有反射)与均匀缩放复合时，相似变换的矩阵表示为：

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{1.8}$$

可以用更简洁的分块形式写成

$$\mathbf{x}' = \mathrm{H}_{\mathrm{S}}\mathbf{x} = \begin{bmatrix} s\mathrm{R} & \mathbf{t} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}\mathbf{x} \tag{1.9}$$

其中标量 s 表示均匀缩放．相似变换也称等形变换，因为它保持了“形状”(形式)．一个平面相似变换有四个自由度，比欧氏变换多一个缩放自由度．相似变换可由两组点对应算出．

**不变量** 在考虑缩放自由度影响的基础上，它的不变量可以由欧氏不变量推出．直线的夹角不受旋转，平移或均匀缩放的影响，因而是相似不变量．特别是平行线映射为平行线．两点间的长度不是相似不变量，但两长度的**比率**是不变量，因其缩放因子相互抵消．同样地，面积的比率是不变量，也因为缩放因子(的平方)被抵消．

**度量结构** 讨论重构(第 9 章)时常用的一个术语是**度量**．所谓**度量结构**就是确定到只相差一个相似变换的结构．

### 1.4.3 类 III：仿射变换

仿射变换是一个非奇异线性变换与一个平移变换的复合．它的矩阵表示为

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{1.10}$$

或分块形式

$$\mathbf{x}' = \mathrm{H}_{\mathrm{A}}\mathbf{x} = \begin{bmatrix} \mathrm{A} & \mathbf{t} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}\mathbf{x} \tag{1.11}$$

其中 A 是一个 $2\times2$ 的非奇异矩阵．平面仿射变换有六自由度，对应于六个矩阵元素．变换可以由三组点对应来计算．

理解仿射变换中线性成分 A 的几何效应的一个有益方法是把它看作两个基本变换——旋转和非均匀缩放的复合．仿射矩阵 A 总能分解为

$$\mathrm{A} = \mathrm{R}(\theta)\mathrm{R}(-\phi)\mathrm{D}\mathrm{R}(\phi) \tag{1.12}$$

其中 $\mathrm{R}(\theta)$ 和 $\mathrm{R}(\phi)$ 分别表示转角为 $\theta$ 和 $\phi$ 的旋转，而 D 为对角矩阵：

$$\mathrm{D} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

分解式(1.12)由 SVD(节 A3.3($p$408))(SVD 为“奇异值分解”的缩写，译者注．)直接得到：$\mathrm{A} = \mathrm{U}\mathrm{D}\mathrm{V}^{\mathrm{T}} = (\mathrm{U}\mathrm{V}^{\mathrm{T}})(\mathrm{V}\mathrm{D}\mathrm{V}^{\mathrm{T}}) = \mathrm{R}(\theta)(\mathrm{R}(-\phi)\mathrm{D}\mathrm{R}(\phi))$，其中第一个等式来自 A 的 SVD，其他等式则因为 U 和 V 是正交矩阵．

因此，仿射矩阵 A 被看成是一个旋转($\phi$)；加上在(已旋转)的 $x$ 和 $y$ 方向分别进行按比例因子 $\lambda_1$ 和 $\lambda_2$ 的缩放，再加上一个回转($-\phi$)和最后一个旋转($\theta$)的复合变换. 与相似变换相比，"新"几何仅仅是非均匀缩放. 它使仿射变换比相似变换多了两个自由度. 它们是缩放方向的角度 $\phi$ 和缩放参数比率 $\lambda_1:\lambda_2$. 仿射变换的本质是在一个特定角的两个垂直方向上进行缩放. 图 1.7 给出两个示意性例子.

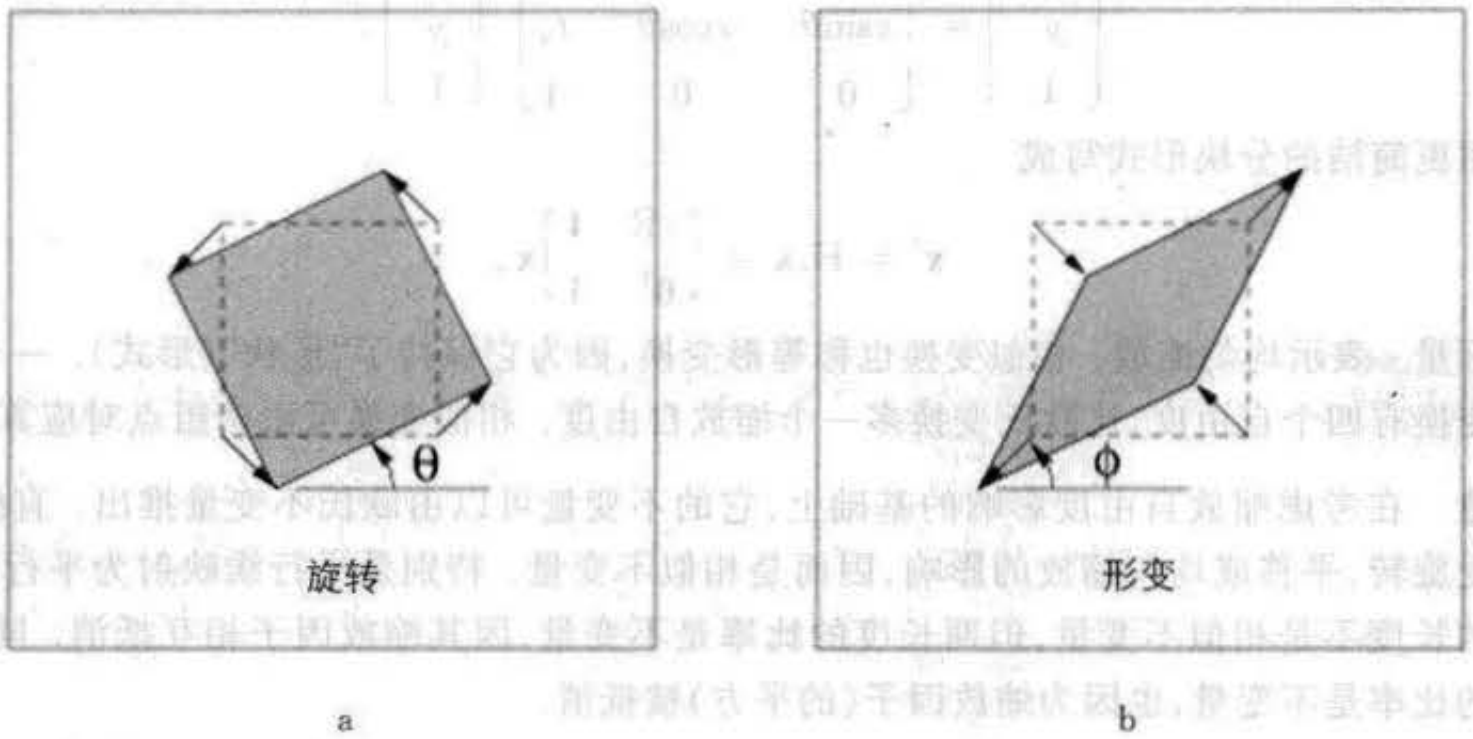



图 1.7 **平面仿射变换失真**. (a)旋转 $R(\theta)$. (b)形变 $R(-\phi)DR(\phi)$. 注意形变中缩放方向是正交的.

**不变量** 因为仿射变换包含非均匀缩放，所以长度比率和线间夹角等相似不变量在仿射变换下不再保留. 三个重要的仿射不变量是：

(1) **平行线** 考查两平行线. 它们交于某无穷远点$(x_1, x_2, 0)^T$. 在仿射变换下，该点被映射到另一个无穷远点. 因此，平行线被映射到仍然交于无穷远的直线，即它们的像仍然平行.

(2) **平行线段的长度比** 直线段的长度缩放仅与该线段方向和缩放方向之间的夹角有关. 假定该线段与正交缩放方向 $x$-轴的夹角为 $\alpha$，那么，缩放大小为 $\sqrt{\lambda_1^2\cos^2\alpha+\lambda_2^2\sin^2\alpha}$. 因该缩放因子对所有同向的直线是一样的，所以在平行线段的比率中被消去.

(3) **面积比** 可以通过分解式(1.12)直接得到该不变性. 旋转和平移不影响面积，起作用的仅仅是按 $\lambda_1$ 和 $\lambda_2$ 的缩放. 其效果是面积被缩放了 $\lambda_1\lambda_2 = \det A$ 倍. 因此，任何形状的面积都被缩放了 $\det A$ 倍，而该缩放因子在面积比中被消去. 我们将会看到这一性质在射影变换时不成立.

根据 $\det A$ 是正或负，仿射变换分别称为保向的或逆向的. 因 $\det A = \lambda_1\lambda_2$，保向性或逆向性仅与其符号有关.

### 1.4.4 类 IV：射影变换

射影变换已在(1.5)中定义. 它是**齐次**坐标的一般非奇异线性变换. 它是仿射变换的推广，仿射变换是**非齐次**坐标的一般非奇异线性变换和一个平移的复合. 我们早已见到射影变换的作用(节 1.3). 这里，我们取其分块形式

$$\mathbf{x}' = \mathrm{H}_{\mathrm{p}}\mathbf{x} = \begin{bmatrix} \mathrm{A} & \mathbf{t} \\ \mathbf{v}^{\mathrm{T}} & v \end{bmatrix}\mathbf{x} \tag{1.13}$$

其中矢量 $\mathbf{v}=(v_1, v_2)^{\mathrm{T}}$. 齐次矩阵虽有九个元素但只有它们的比率是有意义的,因此该变换由八个参数确定. 注意并不是总有可能通过对矩阵缩放而取 $v$ 为 1,因为 $v$ 可能是零. 两平面之间的射影变换可由四组点对应算出,但其中属于同一平面的三点必须不共线. 见图 1.4.

与仿射变换不同,在 $\mathbb{P}^2$ 中不能区分保向或逆向射影变换. 我们将在节 1.6 中回到这个问题.

**不变量** 最基本的射影不变量是四共线点的交比:直线上长度的比率在仿射变换下保持不变,但在射影变换下并非如此. 然而比率的比率即直线上长度的**交比**是射影不变量. 我们将在节 1.5 讨论这个不变量的性质.

### 1.4.5 小结与比较

仿射变换(6 dof)介于相似变换(4 dof)与射影变换(8 dof)之间. 仿射变换推广相似变换使得夹角不再保持不变,造成物体形状在变换后产生歪斜. 另一方面,仿射变换对平面的作用是均匀的:对于一个给定的仿射变换,平面上任何地方的物体(比如说一个正方形)的面积缩放因子 det A 都是一样的;直线变换的方向取决于它原来的方向,而与它在平面上的位置无关. 与此相反,经一个给定的射影变换后,面积的缩放随位置的改变而改变(比如在透视变换下,平面上较远的正方形比较近的正方形的图像小,如图 1.6 所示);并且直线变换的方向既取决于原线的方向又取决于它的位置(在节 7.6($p$149)中将会看到线的消影点仅取决于直线的方向,与位置无关).

射影与仿射变换的根本区别在于射影变换中矢量 $\mathbf{v}$ 不是零. 由它引起射影变换的非线性效应. 把理想点$(x_1, x_2, 0)^{\mathrm{T}}$在仿射和射影变换下的映射做一个对比. 首先看仿射变换,

$$\begin{bmatrix} \mathrm{A} & \mathbf{t} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathrm{A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ 0 \end{pmatrix}. \tag{1.14}$$

再看,射影变换

$$\begin{bmatrix} \mathrm{A} & \mathbf{t} \\ \mathbf{v}^{\mathrm{T}} & v \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathrm{A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ v_1 x_1 + v_2 x_2 \end{pmatrix}. \tag{1.15}$$

在第一种情况下,理想点仍然是理想点(在无穷远处). 在第二种情况下,理想点被映射到有限点. 正是因为具有这种能力,射影变换能对消影点建模.

### 1.4.6 射影变换的分解

射影变换可以分解为一串变换链的复合,链中的每个矩阵比它前面的一个矩阵所表示的变换层次高.

$$\mathrm{H} = \mathrm{H}_{\mathrm{S}}\mathrm{H}_{\mathrm{A}}\mathrm{H}_{\mathrm{P}} = \begin{bmatrix} s\mathrm{R} & \mathbf{t}/v \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}\begin{bmatrix} \mathrm{K} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}\begin{bmatrix} \mathrm{I} & \mathbf{0} \\ \mathbf{v}^{\mathrm{T}} & v \end{bmatrix} = \begin{bmatrix} \mathrm{A} & \mathbf{t} \\ \mathbf{v}^{\mathrm{T}} & v \end{bmatrix} \tag{1.16}$$

其中 $A = sRK + \mathbf{t}\mathbf{v}^T/v$，而 K 是满足 $\det K = 1$ 的归一化上三角矩阵. 如果 $v \neq 0$ 上述分解是有效的，而且如果 $s$ 取正值，它还是唯一的.

矩阵 $H_S$、$A_A$、$H_P$ 分别是相似变换、仿射变换和射影变换(如下标 S、A、P 所指示). 考察例 1.12 中对平面的透视图像进行矫正的过程：$H_P$(2 dof)移动无穷远直线；$H_A$(2 dof)影响仿射性质，而不移动无穷远直线；最后，$H_S$ 是一般的相似变换(4 dof)，它不影响仿射及射影性质. 变换 $H_P$ 属于节 A5.3($p$430)所介绍的一种有约束的透视变换.

**例 1.15** 射影变换

$$H = \begin{bmatrix} 1.707 & 0.586 & 1.0 \\ 2.707 & 8.242 & 2.0 \\ 1.0 & 2.0 & 1.0 \end{bmatrix}$$

可以分解为

$$H = \begin{bmatrix} 2\cos 45^\circ & -2\sin 45^\circ & 1 \\ 2\sin 45^\circ & 2\cos 45^\circ & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

△

当我们的目标只是确定部分的变换时，可以应用这样的分解. 例如，如果我们需要从平面的射影图像中测量长度比，那么仅需要确定(矫正)到相似变换. 我们将在节 1.7 回过来讨论这种方法.

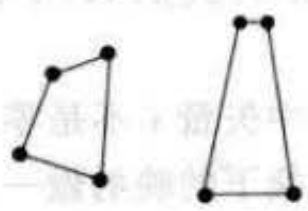
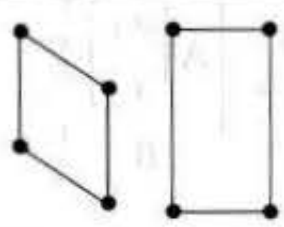
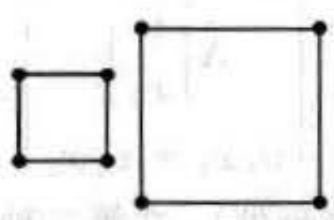
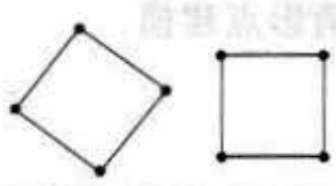
| 群 | 矩阵 | 失真 | 不变性质 |
|---|---|---|---|
| 射影<br>8 dof | $\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$ |  | 共点，共线，**接触的阶**：相交(1 阶接触)；相切(两阶接触)；拐点(与直线的 3 阶接触)；切线不连续性和歧点. 交比(长度比的比). |
| 仿射<br>6 dof | $\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$ |  | 平行，面积比，共线线段或平行线段的长度比(例如中点)，矢量的线性组合(例如形心).<br>无穷远线 $\mathbf{l}_\infty$ |
| 相似<br>4 dof | $\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$ |  | 长度比，夹角. 虚圆点 **I**,**J**(见节 1.7.3) |
| 欧氏<br>3 dof | $\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$ |  | 长度，面积 |

**表 1.1 常见平面变换的几何不变性质** $A = [a_{ij}]$ 是 $2 \times 2$ 的可逆矩阵，$R = [r_{ij}]$ 是 2D 旋转矩阵，$(t_x, t_y)^T$ 是 2D 平移矢量. 失真列给出变换对正方形所产生的典型效应. 表中高层的变换可以产生比它低层的变换的所有效应. 它们的范围从欧氏(其中仅有平移和旋转)到射影(其中正方形被变换成任意四边形(假定没有三点共线)).

### 1.4.7 不变量的数目

已知几何配置，在一种特定类型的变换下不变量的数目是多少？首先术语“数目”需要进一步精确化，如果一个量是不变量，例如欧氏变换下的长度，那么该量的任何函数也是不变量．因而，我们要寻求一个与函数无关的不变量计数法则．为了形成不变量就必须消去变换矩阵的参数数目，考虑到这点后可以导出：

**结论 1.16　与函数无关的不变量数等于或大于配置的自由度数减去变换的自由度数．**

例如，由一般位置上的四点所组成的配置有八个自由度（每点 2 个），从而有 4 个相似，2 个仿射和零个射影不变量，因为这些变换分别有 4，6，8 自由度．

表 1.1 归纳了 2D 变换群以及它们的不变性质．在表中低层的变换是它高层的变换的特殊化．在表低层的变换继承其高层变换的不变量．

## 1.5　1D 射影几何

直线射影几何 $\mathbb{P}^1$ 的推导方法与平面的几乎一样．直线上的点 $x$ 用齐次坐标表示为 $(x_1, x_2)^\mathsf{T}$，而 $(x_1, 0)^\mathsf{T}$ 是该直线的理想点．我们将用记号 $\bar{\mathbf{x}}$ 表示 2 维矢量 $(x_1, x_2)^\mathsf{T}$．直线的射影变换由一个 $2\times 2$ 的齐次矩阵来表示：

$$\bar{\mathbf{x}}' = \mathrm{H}_{2\times 2}\bar{\mathbf{x}}$$

它有 3 个自由度，即矩阵的四个元素减去一个全局缩放因子．直线的射影变换可以由 3 组对应点来确定．

**交比**　交比是 $\mathbb{P}^1$ 的基本射影不变量．给定 4 个点 $\bar{\mathbf{x}}_i$，**交比**定义为

$$\mathrm{Cross}(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \bar{\mathbf{x}}_3, \bar{\mathbf{x}}_4) = \frac{|\bar{\mathbf{x}}_1\bar{\mathbf{x}}_2||\bar{\mathbf{x}}_3\bar{\mathbf{x}}_4|}{|\bar{\mathbf{x}}_1\bar{\mathbf{x}}_3||\bar{\mathbf{x}}_2\bar{\mathbf{x}}_4|}$$

其中

$$|\bar{\mathbf{x}}_i\bar{\mathbf{x}}_j| = \det\begin{bmatrix} x_{i1} & x_{j1} \\ x_{i2} & x_{j2} \end{bmatrix}.$$

关于交比的几点注释如下：

(1) 交比的值与各点 $\bar{\mathbf{x}}_i$ 所用的具体的齐次表示无关，因为分子和分母的缩放因子在它们之间抵消．

(2) 如果每点 $\bar{\mathbf{x}}_i$ 都是有限点，并在齐次表示中均选择 $x_2 = 1$，那么 $|\bar{\mathbf{x}}_i\bar{\mathbf{x}}_j|$ 就表示由 $\bar{\mathbf{x}}_i$ 到 $\bar{\mathbf{x}}_j$ 的带符号的距离．

(3) 如果点 $\bar{\mathbf{x}}_i$ 中有一个理想点，交比的定义仍然有效．

(4) 在任何直线的射影变换下，交比的值不变：如果 $\bar{\mathbf{x}}' = \mathrm{H}_{2\times 2}\bar{\mathbf{x}}$，则

$$\mathrm{Cross}(\bar{\mathbf{x}}_1', \bar{\mathbf{x}}_2', \bar{\mathbf{x}}_3', \bar{\mathbf{x}}_4') = \mathrm{Cross}(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \bar{\mathbf{x}}_3, \bar{\mathbf{x}}_4). \tag{1.17}$$

以上四点的证明留着习题．它也可以说成是：交比对于直线的射影坐标选择是不变量．图 1.8 说明若干等交比的直线之间的射影变换．

平面射影变换下，平面上的任何直线都诱导一个 1D 射影变换．

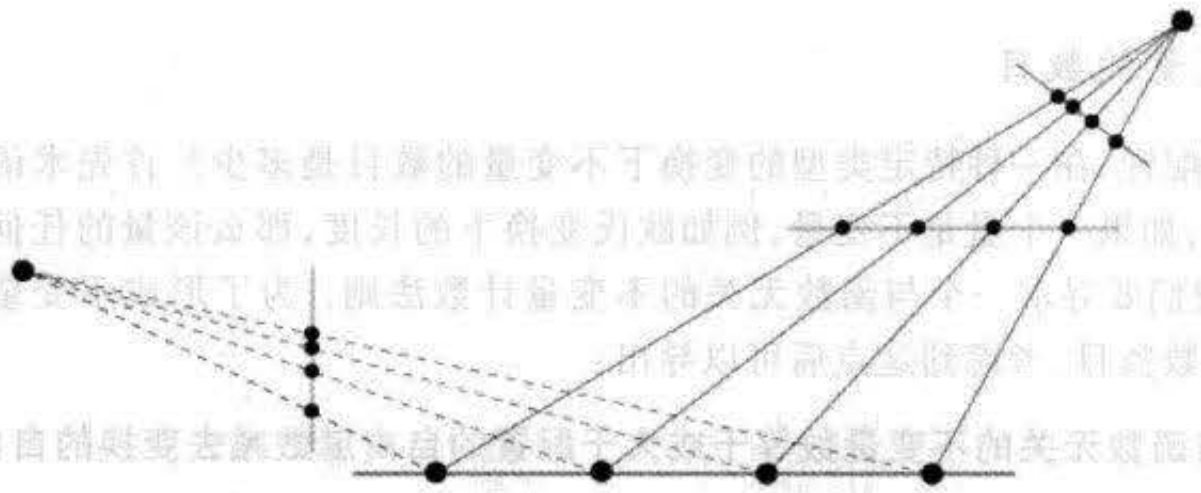


图 1.8 **直线之间的射影变换**. 图中有四组四共线点. 每组与其他组都通过直线到直线的射影变换相关联. 因为在射影变换下交比是不变量, 所以图中每一组有相同的交比值.

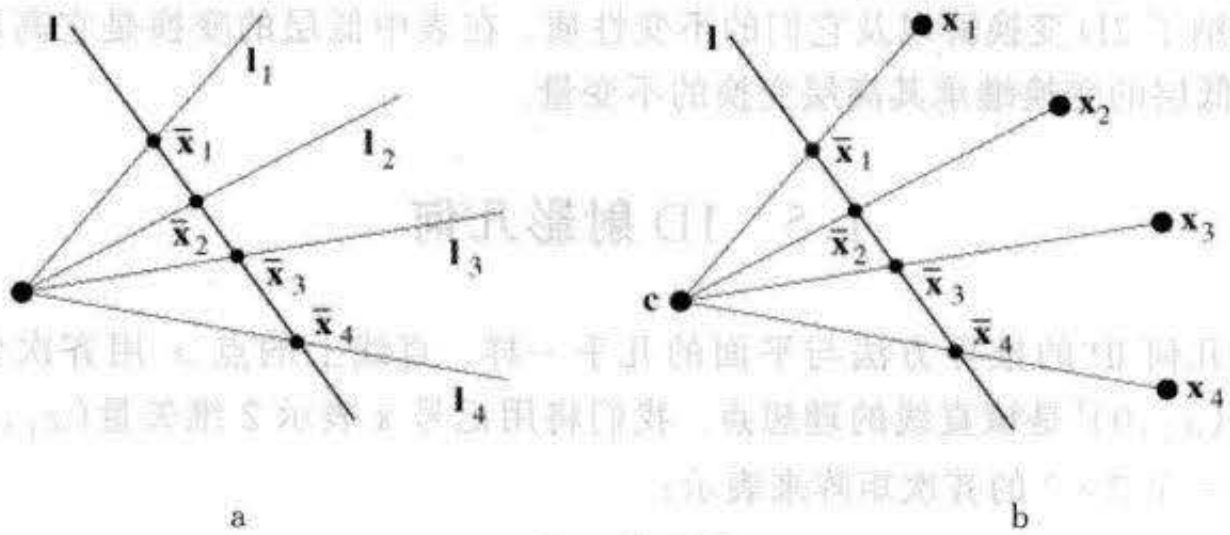



图 1.9 **共点直线**. (a)四共点线 $\mathbf{l}_i$ 与直线 $\mathbf{l}$ 交于四点 $\bar{\mathbf{x}}_i$. 这些直线的交比对平面射影变换来说是不变量. 它的值由点的交比:Cross($\bar{\mathbf{x}}_1$, $\bar{\mathbf{x}}_2$, $\bar{\mathbf{x}}_3$, $\bar{\mathbf{x}}_4$)给出. (b)共面点 $\mathbf{x}_i$ 被过中心 $\mathbf{c}$ 的投影影像到一直线 $\mathbf{l}$(也在该平面上). 像点 $\bar{\mathbf{x}}_i$ 的交比对像线 $\mathbf{l}$ 的位置来说是不变量.

**共点线** 共点线是直线上共线点的对偶. 这意味着平面上的共点线也有几何 $\mathbb{P}^1$. 特别是, 任何四条共点线都有一个确定的交比,如图 1.9a 所示.

请注意图 1.9b 如何把平面 $\mathbb{P}^2$上的点投影到 1 维图像的表示方法. 特别是,如果 $\mathbf{c}$ 表示摄像机中心,而直线 $\mathbf{l}$ 表示像直线(像平面的 1D 情况),那么点 $\bar{\mathbf{x}}_i$ 是点 $\mathbf{x}_i$ 向该像直线的投影. 点 $\bar{\mathbf{x}}_i$ 的交比刻画四个像点的射影配置. 注意,就四个像点的射影配置而论,像直线的实际位置是无关紧要的——不同像直线的选择都给出射影等价的像点配置.

共点线的射影几何对理解第 8 章中对极线的射影几何非常重要.

## 1.6 射影平面的拓扑

本节中**简短地介绍一下 $\mathbb{P}^2$的拓扑. 它并非是理解本书以后的内容所必须的**.

我们已经知道射影平面 $\mathbb{P}^2$可以视为全体齐次 3 维矢量的集合. 这种类型的矢量 $\mathbf{x}=(x_1, x_2, x_3)^{\mathrm{T}}$可以通过乘以一个非零因子使得 $x_1^2+x_2^2+x_3^2=1$ 来归一化. 这样的一个点在 $\mathbb{R}^3$的单位球面上. 但在 $\mathbb{P}^2$中,任何相差一个乘数因子 $-1$ 的矢量 $\mathbf{x}$ 和 $-\mathbf{x}$ 表示同一点. 因此,$\mathbb{R}^3$的单位球面 $S^3$ 和射影平面 $\mathbb{P}^2$之间存在一种二对一的对应. 射影平面可以画成一单位球面,其中符号相反的点视为等同. 在此表示中,$\mathbb{P}^2$上的直线被模型化为单位球面上的大圆(符号相反的正对点仍等同). 可以验证,球上任何两不同的(非正对)点恰好在一

大圆上，任何两大圆相交于一点(因为正对点等同).

按拓扑学的语言，球面 $S^2$ 是 $\mathbb{P}^2$的 2－叶复盖空间. 从而 $\mathbb{P}^2$不是**单连通**的，即在 $\mathbb{P}^2$中存在环，不能在 $\mathbb{P}^2$内收缩到一点. 更专业化地说，$\mathbb{P}^2$的基本群是阶为 2 的循环群.

在把射影平面看作正对点等同的球面的模型中，可以把球面 $S^2$ 的下半球拿掉，因为下半球的点与其在上半球的正对点一样. 在这种情形下，$\mathbb{P}^2$可以由上半球面构成并认为在赤道上正对点等同. 因为 $S^2$ 的上半球拓扑等价于一个圆盘，所以 $\mathbb{P}^2$就是一个圆盘而且边缘上正对点被视为等同或粘在一起. 这在物理上是不可能的. 通过给圆盘粘边界来构造拓扑空间是拓扑学常用的方法，并且事实上，任何 2 维流形都可以用这方法来构造，如图 1.10 所示.

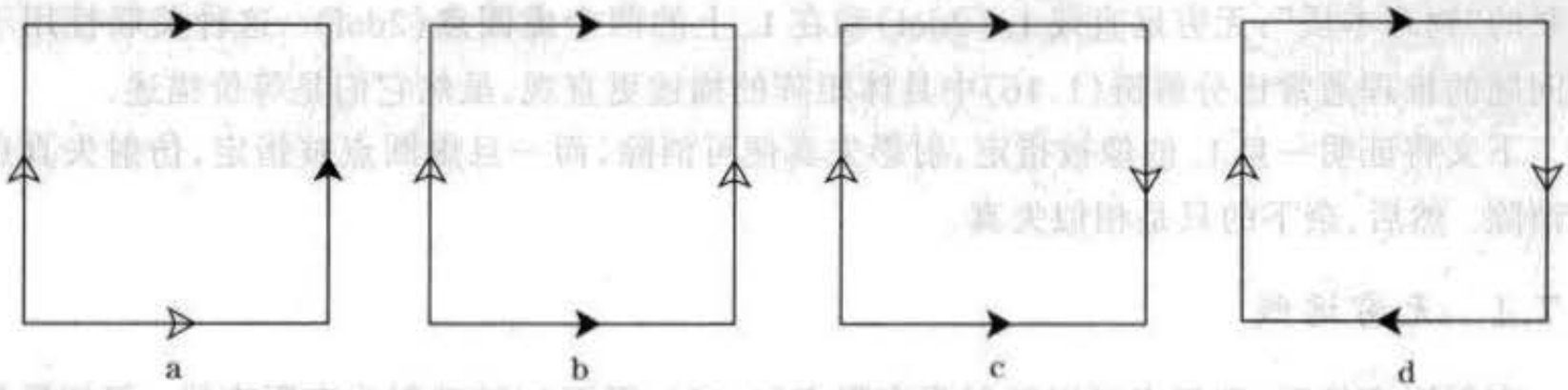



图 1.10　**表面的拓扑**. 把正方形纸片(拓扑等价于圆盘)的边适当相粘接可以构成通常的曲面. 在每种情况中，把正方形中有相同箭头的两边粘在一起并保持箭头的方向一致. 我们得到(a)球面(b)环面(c) Klein 瓶(d)射影平面. 只有球面和环面可用带箭头的纸真正实现. 球面和环面是可定向的但射影平面和 Klein 瓶则不然.

射影平面 $\mathbb{P}^2$的一个显著特点是不可定向，即不可能定义一个在整个表面上保持一致的局部方位(例如由一对有向坐标轴来表示). 图 1.11 对此给予了说明，即证明射影平面包含一条逆向的路径.

图 1.11　**曲面的定向**. 一个坐标系(图中用 L 表示)可以沿曲面上一条路径平移并最终回到它的出发点.(a)表示一个射影平面. 坐标系(由一对轴表示)沿给出的路径回到出发点时被反向了，因为正方形的边界上视为等同的两个对点有一条轴的方向调转了. 这样的路径称为逆向路径，含有逆向路径的曲面称不可定向的. (b)给出一个熟知的例子：麦比乌斯带，它由反向粘接长方形的两对边而得到.(M.C. Escher's "Moebius Strip II[Red Ants],1963©2000 Cordon Art B.V. - Baarn - Holland. 版权所有). 可以验证，绕麦比乌斯带一圈的路径是逆向的.

**$\mathbb{P}^1$的拓扑**　类似地，1 维射影线可以等同于 1 维球面 $S^1$(即圆)且正对点视为等同. 如果去掉可由上半圆复制的下半圆，那么上半圆便拓扑等价于一直线段. 因此 $\mathbb{P}^1$拓扑等价于一条

将两端视为等同的直线段，它也拓扑等价于圆 $S^1$.

## 1.7 从图像恢复仿射和度量性质

我们回到射影矫正的例子(例 1.12($p$9))，它的目的是消除平面的透视图像中的射影失真，使得原始平面的相似性质(角度，长度比)可以被测量. 在该例中，通过确认平面上四个参考点的位置(共 8 个自由度)，并显式地算出映射参考点到它们图像的变换，射影失真被完全消除. 事实上，它超定了该几何——射影变换仅比相似变换多 4 个自由度，因此，为确定度量性质仅需要指定 4 个自由度(不是 8). 在射影变换中，这 4 个自由度给出与几何对象相关联的"物理本质"：无穷远直线 $\mathbf{l}_\infty$(2dof)和在 $\mathbf{l}_\infty$ 上的两个**虚圆点**(2dof). 这种关联性用于此问题的推理通常比分解链(1.16)中具体矩阵的描述更直观，虽然它们是等价描述.

下文将证明一旦 $\mathbf{l}_\infty$ 的像被指定，射影失真便可消除，而一旦虚圆点被指定，仿射失真也可消除. 然后，余下的只是相似失真.

### 1.7.1 无穷远线

在射影变换下，理想点可以映射为有限点(1.15)，因而 $\mathbf{l}_\infty$ 被映射为有限直线. 但如果是仿射变换，$\mathbf{l}_\infty$ 不会被映射为有限直线，即仍留在无穷远处. 显然，它可直接由直线的变换(1.6 – $p$10)推出：

$$\mathbf{l}'_\infty = \mathrm{H}_\mathrm{A}^{-\mathsf{T}}\mathbf{l}_\infty = \begin{bmatrix} \mathrm{A}^{-\mathsf{T}} & \mathbf{0} \\ -\mathbf{t}^{\mathsf{T}}\mathrm{A}^{-\mathsf{T}} & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{l}_\infty$$

逆命题也是正确的，即仿射变换是保持 $\mathbf{l}_\infty$ 不变的最一般的线性变换，并可证明如下. 我们要求一个无穷远点，例如 $\mathbf{x}=(1,0,0)^{\mathsf{T}}$，被映为一个无穷远点. 该要求就决定了 $h_{31}=0$，同理可证 $h_{32}=0$，所以该变换是仿射变换. 概括起来：

**结论 1.17　在射影变换 H 下，无穷远直线 $\mathbf{l}_\infty$ 为不动直线的充要条件是 H 是仿射变换.**

但是，在仿射变换下，$\mathbf{l}_\infty$ 不是点点不动的：(1.14)表明在仿射变换下 $\mathbf{l}_\infty$ 的点(理想点)被映射为 $\mathbf{l}_\infty$ 的点，但它不是原来的点，除非 $\mathrm{A}(x_1, x_2)^{\mathsf{T}} = k(x_1, x_2)^{\mathsf{T}}$. 现在可以证明辨认 $\mathbf{l}_\infty$ 就能恢复仿射性质(平行，面积比).

### 1.7.2 由图像恢复仿射性质

在平面的像中，一旦无穷远直线的像得到辨认，就有可能对原平面进行仿射测量. 例如，如果两条直线的影像相交在 $\mathbf{l}_\infty$ 的像上，则可认定这两条直线在原平面上平行. 这是因为在欧氏平面中平行线相交在 $\mathbf{l}_\infty$ 上，又因射影变换保持交点不变，经射影变换之后直线仍然交于 $\mathbf{l}_\infty$ 的像上. 类似地，一旦 $\mathbf{l}_\infty$ 被辨认，直线上的长度比率便可由像平面上该直线上确定长度的三个点以及该直线与 $\mathbf{l}_\infty$ 的交点(它提供交比的第四点)的交比来计算，等等.

但是，一个转弯不大却更适合于计算机算法的途径是直接把已辨认的 $\mathbf{l}_\infty$ 变换到它的规范位置 $\mathbf{l}_\infty=(0,0,1)^{\mathsf{T}}$. 把实现此变换的(射影)矩阵应用于图像中的每一点以达到对图像进

行仿射矫正的目的，即变换之后，仿射测量可以直接在矫正过的图像中进行．这个基本思想在图 1.12 中加以说明．

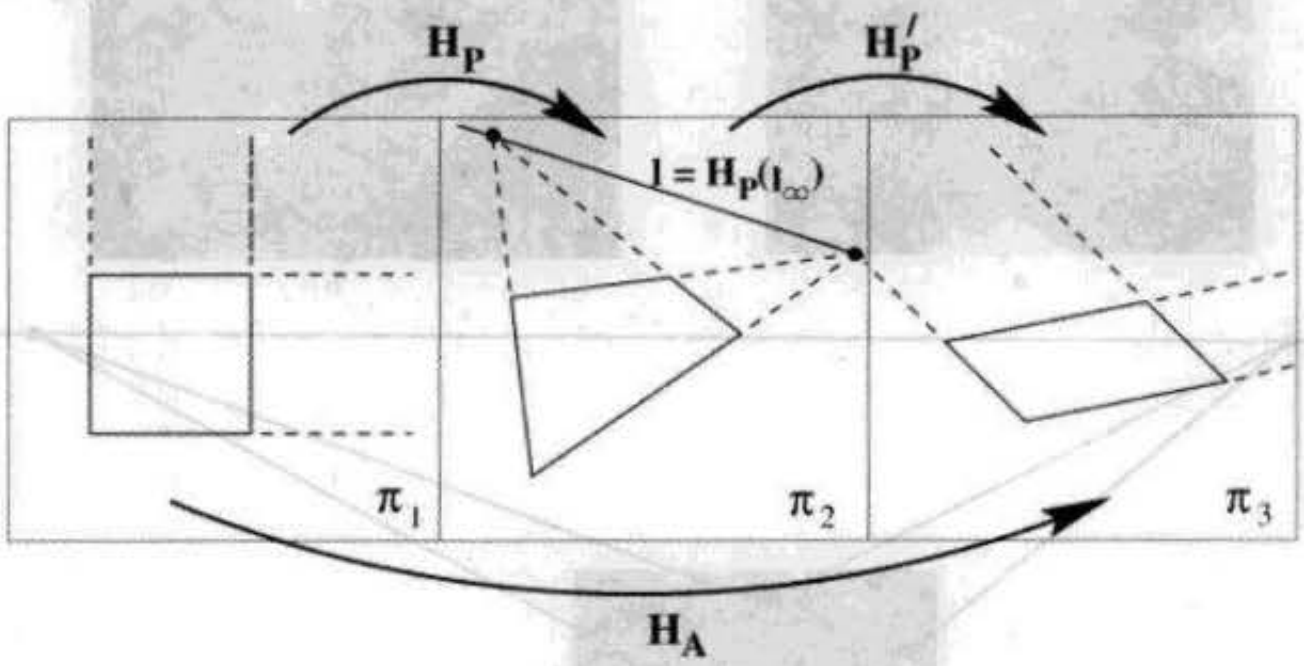



图 1.12　**仿射矫正**．射影变换把 $\mathbf{l}_\infty$ 从欧氏平面 $\pi_1$ 的 $(0,0,1)^\mathsf{T}$ 映射到平面 $\pi_2$ 的有限直线 $\mathbf{l}$．如果构造一个射影变换把 $\mathbf{l}$ 映射回 $(0,0,1)^\mathsf{T}$，那么根据结论 1.17 从第一到第三张平面的变换必定是仿射变换，因为 $\mathbf{l}_\infty$ 的标准位置保持不变．这意味着第一张平面的仿射性质可以从第三张平面上测量，即第三张平面是第一平面的仿射像．

如果无穷远直线的像是 $\mathbf{l}=(l_1,\ l_2,\ l_3)^\mathsf{T}$，假定 $l_3 \neq 0$，那么把 $\mathbf{l}$ 映射回 $\mathbf{l}_\infty=(0,\ 0,\ 1)^\mathsf{T}$ 的一个合适的射影点变换是

$$\mathrm{H}=\mathrm{H_A}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ l_1 & l_2 & l_3\end{bmatrix} \tag{1.18}$$

其中 $\mathrm{H_A}$ 可取为任何仿射变换（H 的最后一行是 $\mathbf{l}^\mathsf{T}$）．可以验证在直线变换（1.6 − $p$10）下，$\mathrm{H}^{-\mathsf{T}}(l_1,l_2,l_3)^\mathsf{T}=(0,0,1)^\mathsf{T}=\mathbf{l}_\infty$．

**例 1.18　仿射矫正**

在平面的透视图像中，世界平面上的无穷远直线被影像为该平面的消影线．第 7 章中将对它作更详细的讨论．如图 1.13 所示，消影线 $\mathbf{l}$ 可以由平行线的影像的交点来计算．然后用射影形变（1.18）作用于图像平面，使 $\mathbf{l}$ 映射到它的规范位置 $\mathbf{l}_\infty=(0,0,1)^\mathsf{T}$，从而达到图像仿射矫正的目的．△

该例表明仿射性质可以通过指定一条直线（2dof）来恢复．这等价于仅仅指定变换分解链（1.16）中的射影成分．

**直线上的距离比**　给定一条直线上有已知长度比的两个线段，该直线上的无穷远点便可以确定．典型的情况是在像直线上的三个点 $\mathbf{a}',\mathbf{b}',\mathbf{c}'$ 已被辨认．假定 $\mathbf{a},\mathbf{b},\mathbf{c}$ 是世界直线上对应的共线点，且长度比 $d(\mathbf{a},\mathbf{b}):d(\mathbf{b},\mathbf{c})=a:b$ 已知（这里 $d(\mathbf{x},\mathbf{y})$ 是点 $\mathbf{x}$ 和 $\mathbf{y}$ 之间的欧氏距离）．有可能利用交比找到消影点．其过程如下：

（1）在图像中量出距离比，$d(\mathbf{a}',\mathbf{b}'):d(\mathbf{b}',\mathbf{c}')=a':b'$．

（2）在直线 $<\mathbf{a},\mathbf{b},\mathbf{c}>$ 上建立坐标系，使点 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的坐标分别为 $0,a,a+b$．为计算方便，把这些点表示为齐次 2 维矢量 $(0,1)^\mathsf{T}$，$(a,1)^\mathsf{T}$ 和 $(\mathrm{a}+\mathrm{b},1)^\mathsf{T}$．类似地，令 $\mathbf{a}',\mathbf{b}',\mathbf{c}'$ 的坐标

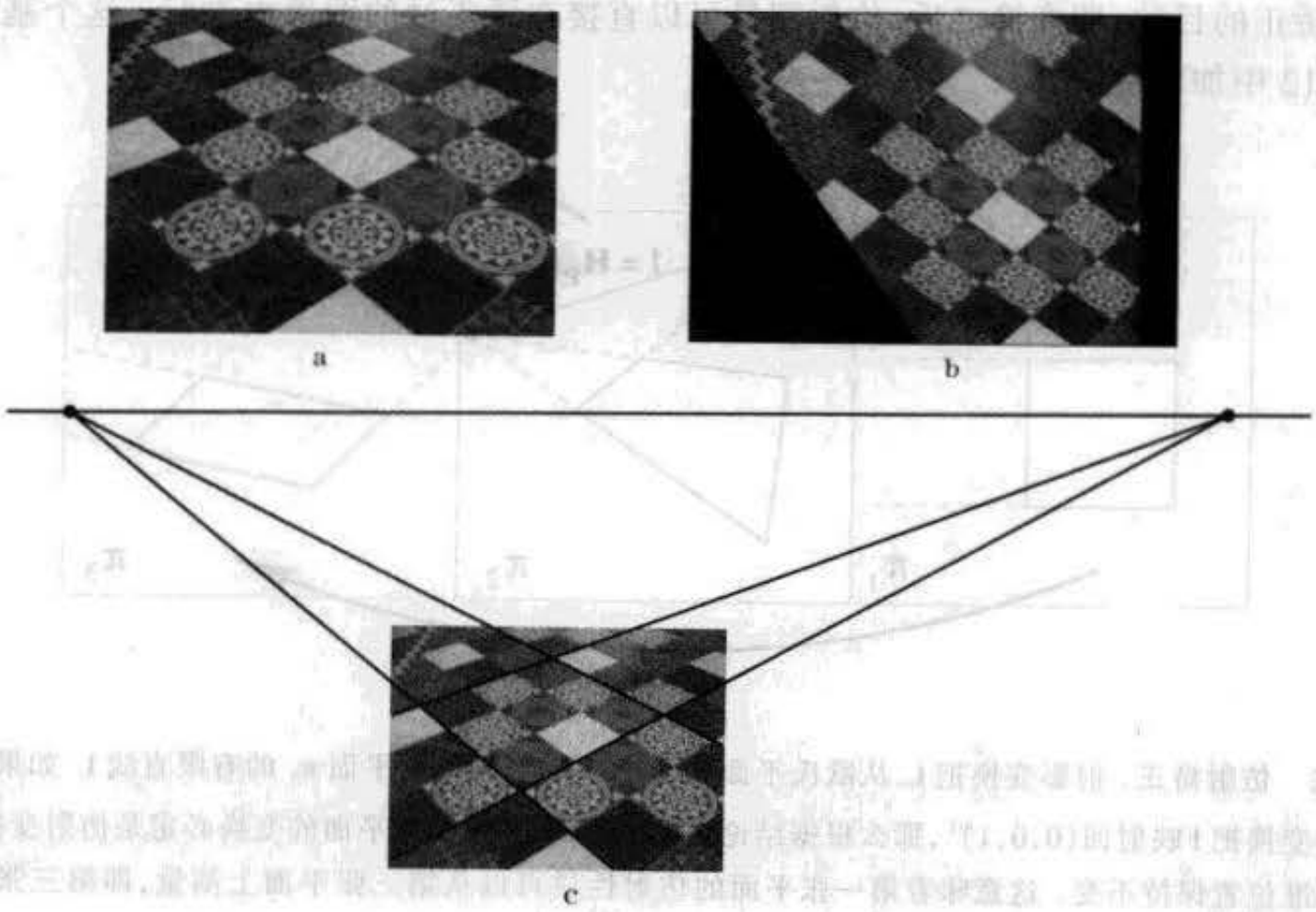



图 1.13 **通过消影线实现仿射矫正**. (a)中平面的消影线在(c)中用两组平行线的像的交点计算. 此后,图像(a)经过射影形变生成仿射矫正图像(b). 被仿射矫正的图像中平行线的像现在是平行的. 但是,夹角不再是它们在世界平面上的直角,因为它们被仿射失真了. 同时参见图 1.17.

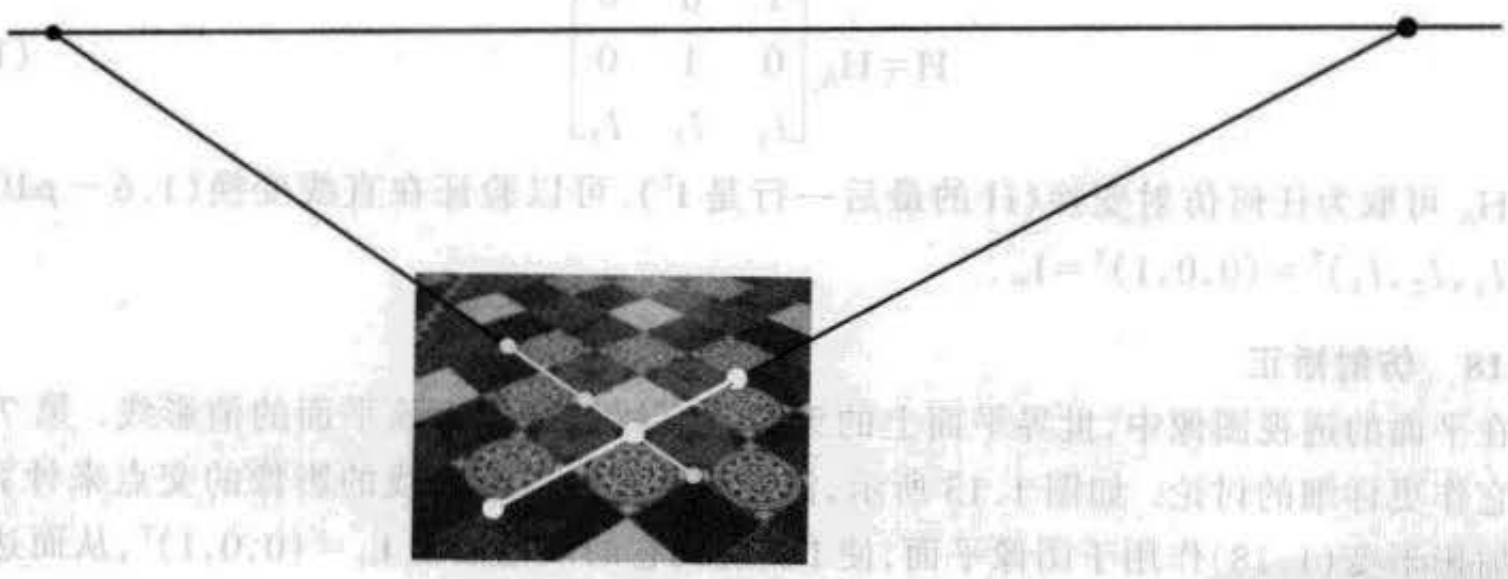


图 1.14 用直线上等距比来确定无穷远点的两个例子. 被采用的线段用细和粗的白线(以点界定)标记. 这种作图方法确定了平面的消影线. 把它与图 1.13c 作比较.

为 $0, a', a'+b'$,并且它们同样可表示为齐次矢量.

(3) 相对于这些坐标系,计算使 $\mathbf{a} \mapsto \mathbf{a}'$, $\mathbf{b} \mapsto \mathbf{b}'$, $\mathbf{c} \mapsto \mathbf{c}'$ 的 1D 射影变换 $H_{2\times 2}$.

(4) 在变换 $H_{2\times 2}$ 下无穷远点的像(坐标为 $(1,0)^T$)是直线 $<\mathbf{a}', \mathbf{b}', \mathbf{c}'>$ 的消影点. 用这种方式计算消影点的例子在图 1.14 中给出.

**例 1.19 由长度比作消影点的几何作图** 图 1.14 中给出的消影点也可以用纯几何作图的方法得到,步骤如下:

(1) 给定:图像中三共线点 $\mathbf{a}', \mathbf{b}', \mathbf{c}'$,它们与线段比是 $a:b$ 的世界共线点对应.

(2) 过 $\mathbf{a}'$ 画任意直线 $\mathbf{l}$(不与直线 $\mathbf{a}'\mathbf{c}'$ 重叠)并标注点 $\mathbf{a}=\mathbf{a}'$, $\mathbf{b}$, $\mathbf{c}$ 使线段 $<\mathbf{ab}>$, $<\mathbf{bc}>$ 的长度比为 $a:b$.

(3) 连接 $\mathbf{bb}'$ 和 $\mathbf{cc}'$,它们交于 $\mathbf{o}$.

(4) 过 $\mathbf{o}$ 作平行 $\mathbf{l}$ 的直线交直线 $\mathbf{a}'\mathbf{c}'$ 于消影点 $\mathbf{v}'$.

该作图过程在图 1.15 中说明. △

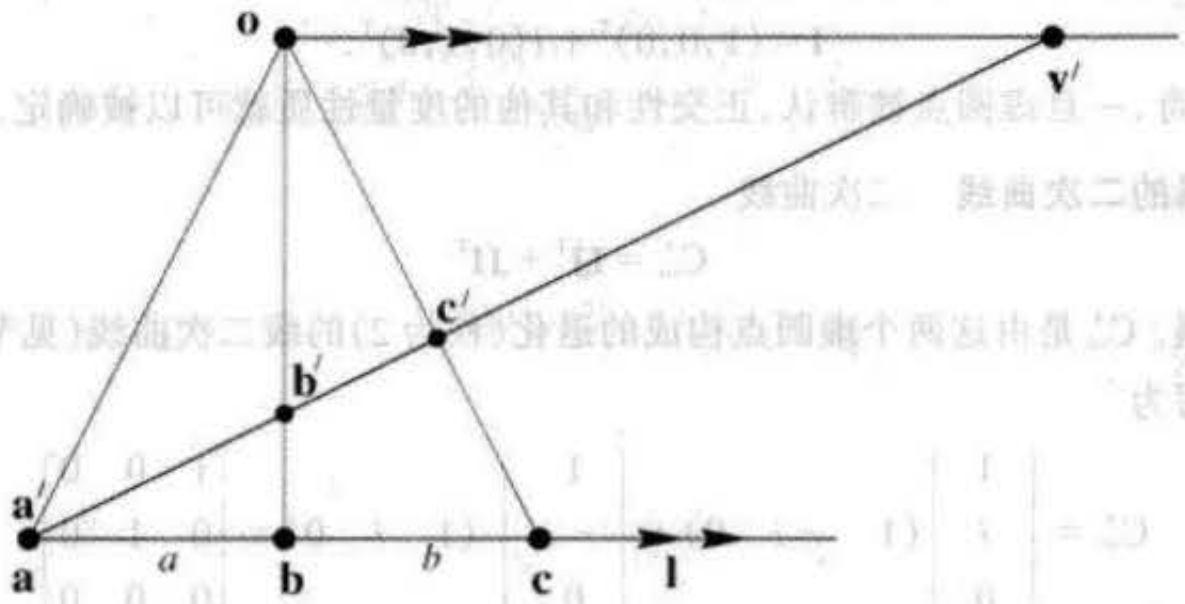



图 1.15 已知长度比,求一条直线上无穷远点的像的几何作图法. 细节在正文中给出.

### 1.7.3 虚圆点及其对偶

在相似变换下,$\mathbf{l}_\infty$ 上有两个不动点. 它们是**虚圆点**(也称**绝对点**)$\mathbf{I}$,$\mathbf{J}$,其标准坐标是

$$\mathbf{I}=\begin{pmatrix}1\\ i\\ 0\end{pmatrix}\qquad \mathbf{J}=\begin{pmatrix}1\\ -i\\ 0\end{pmatrix}.$$

这一对虚圆点是复共轭理想点. 它们在保向相似变换下不变:

$$\begin{aligned}\mathbf{I}'&=\mathbf{H}_S\mathbf{I}\\ &=\begin{bmatrix}s\cos\theta & -s\sin\theta & t_x\\ s\sin\theta & s\cos\theta & t_y\\ 0 & 0 & 1\end{bmatrix}\begin{pmatrix}1\\ i\\ 0\end{pmatrix}\\ &=se^{-i\theta}\begin{pmatrix}1\\ i\\ 0\end{pmatrix}=\mathbf{I}\end{aligned}$$

类似地,可以给出 $\mathbf{J}$ 的证明. 一个反射变换使 $\mathbf{I}$ 和 $\mathbf{J}$ 交换. 逆命题也成立,即如果虚圆点在一个线性变换下不动,那么该线性变换必是相似变换,其证明留作练习. 概括起来有:

**结论 1.20 在射影变换 H 下,虚圆点 I 和 J 为不动点的充要条件是 H 是相似变换.**

"虚圆点"的命名起源于每一圆周交 $\mathbf{l}_\infty$ 于虚圆点. 为了证明这一点,由二次曲线方程(1.1 - $p6$)开始. 在二次曲线为圆时有 $a=c$ 且 $b=0$. 则

$$x_1^2+x_2^2+dx_1x_3+ex_2x_3+fx_3^2=0$$

其中 $a$ 取为 1. 该二次曲线交 $\mathbf{l}_\infty$ 于(理想)点($x_3=0$),即

$$x_1^2+x_2^2=0$$

解得，$\mathbf{I}=(1,i,0)^{\mathsf{T}}$，$\mathbf{J}=(1,-i,0)^{\mathsf{T}}$，即任何圆都交 $\mathbf{l}_\infty$ 于虚圆点．在欧氏几何中我们知道一个圆由三点指定．虚圆点引出另一种计算．圆可以用由五个点定义的一般二次曲线的公式(1.4－$p$6)来计算，它的五个点是三个点加上两个虚圆点．节1.7.5将证明辨认虚圆点(等价地辨认它们的对偶，见下文)能够恢复相似性质(角度，长度比)．代数上，虚圆点是把欧氏几何中两正交方向$(1,0,0)^{\mathsf{T}}$，$(0,1,0)^{\mathsf{T}}$合并到一个复共轭中，即

$$\mathbf{I}=(1,0,0)^{\mathsf{T}}+i(0,1,0)^{\mathsf{T}}.$$

因此，不足为奇，一旦虚圆点被辨认，正交性和其他的度量性质就可以被确定．

**与虚圆点对偶的二次曲线**　二次曲线

$$C_\infty^*=\mathbf{I}\mathbf{J}^{\mathsf{T}}+\mathbf{J}\mathbf{I}^{\mathsf{T}} \tag{1.19}$$

与虚圆点对偶．$C_\infty^*$是由这两个虚圆点构成的退化(秩为2)的线二次曲线(见节1.2.3)．在欧氏坐标系下写为

$$C_\infty^*=\begin{pmatrix}1\\ i\\ 0\end{pmatrix}\begin{pmatrix}1 & -i & 0\end{pmatrix}+\begin{pmatrix}1\\ -i\\ 0\end{pmatrix}\begin{pmatrix}1 & i & 0\end{pmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}.$$

类似于虚圆点的不动性质，二次曲线 $C_\infty^*$ 在相似变换下不变．在某变换下，二次曲线的矩阵如果不变(相差一常数)，则称该二次曲线在此变换下不变．因为 $C_\infty^*$ 是对偶二次曲线，它的变换遵循结论 1.14($p$11)($C^{*\prime}=HC^*H^{\mathsf{T}}$)，可以验证在点相似变换 $\mathbf{x}'=H_S\mathbf{x}$ 下，

$$C_\infty^{*\prime}=H_S C_\infty^* H_S^{\mathsf{T}}=C_\infty^*.$$

其逆命题也正确，从而得到

**结论 1.21　对偶二次曲线 $C_\infty^*$ 在射影变换 H 下不变的充要条件是 H 是相似变换．**

在任何射影框架下，$C_\infty^*$所具有的一些性质：

(1) $C_\infty^*$有 4 自由度：3×3 齐次对称矩阵有 5 个自由度，但约束 $\det C_\infty^*=0$ 减去一个自由度．

(2) $\mathbf{l}_\infty$是 $C_\infty^*$的零矢量．根据定义虚圆点在 $\mathbf{l}_\infty$ 上，即 $\mathbf{I}^{\mathsf{T}}\mathbf{l}_\infty=\mathbf{J}^{\mathsf{T}}\mathbf{l}_\infty=0$，从而

$$C_\infty^*\mathbf{l}_\infty=(\mathbf{I}\mathbf{J}^{\mathsf{T}}+\mathbf{J}\mathbf{I}^{\mathsf{T}})\mathbf{l}_\infty=\mathbf{I}(\mathbf{J}^{\mathsf{T}}\mathbf{l}_\infty)+\mathbf{J}(\mathbf{I}^{\mathsf{T}}\mathbf{l}_\infty)=\mathbf{0}.$$

### 1.7.4　射影平面上的夹角

在欧氏几何中，两线间的夹角由它们法线的点乘来计算．直线 $\mathbf{l}=(l_1,\ l_2,\ l_3)^{\mathsf{T}}$ 和 $\mathbf{m}=(m_1,\ m_2,\ m_3)^{\mathsf{T}}$的法线分别平行于$(l_1,\ l_2)^{\mathsf{T}}$和$(m_1,\ m_2)^{\mathsf{T}}$，其夹角为

$$\cos\theta=\frac{l_1m_1+l_2m_2}{\sqrt{(l_1^2+l_2^2)(m_1^2+m_2^2)}}. \tag{1.20}$$

平面经仿射或射影变换后，公式(1.20)不能被使用，问题出在 $\mathbf{l}$ 和 $\mathbf{m}$ 的头两个分量在射影变换下没有良定的变换性质(它们不是张量)，但类似于(1.20)在射影变换下不变的公式为

$$\cos\theta=\frac{\mathbf{l}^{\mathsf{T}}C_\infty^*\mathbf{m}}{\sqrt{(\mathbf{l}^{\mathsf{T}}C_\infty^*\mathbf{l})(\mathbf{m}^{\mathsf{T}}C_\infty^*\mathbf{m})}} \tag{1.21}$$

其中 $C_\infty^*$是与虚圆点对偶的二次曲线．不言而喻，在欧氏坐标系下，(1.21)被简化为(1.20)．

在点变换 $\mathbf{x}' = \mathrm{H_S}\mathbf{x}$ 下，采用直线（1.6－$p$10）（$\mathbf{l}' = \mathrm{H}^{-\mathrm{T}}\mathbf{l}$）和对偶二次曲线（1.14（$p$11））（$\mathrm{C}^{*\prime} = \mathrm{HC}^{*}\mathrm{H}^{\mathrm{T}}$）的变换规则，可以证明（1.21）在射影变换下不变．例如，分子变换为：

$$\mathbf{l}^{\mathrm{T}}\mathrm{C}_{\infty}^{*}\mathbf{m} \mapsto \mathbf{l}^{\mathrm{T}}\mathrm{H}^{-1}\mathrm{HC}_{\infty}^{*}\mathrm{H}^{\mathrm{T}}\mathrm{H}^{-\mathrm{T}}\mathbf{m} = \mathbf{l}^{\mathrm{T}}\mathrm{C}_{\infty}^{*}\mathbf{m}.$$

同样可以证明齐次对象的缩放因子在分子和分母之间相消．因此（1.21）的确在射影框架下不变．所证明的结果概括为：

**结论 1.22　一旦二次曲线 $\mathrm{C}_{\infty}^{*}$ 在射影平面上被辨认，那么欧氏角可以用（1.21）来测量．**

作为它的推论有

**结论 1.23　如果 $\mathbf{l}^{\mathrm{T}}\mathrm{C}_{\infty}^{*}\mathbf{m} = 0$，则直线 $\mathbf{l}$ 和 $\mathbf{m}$ 正交．**

几何上，如果直线 $\mathbf{l}$ 和 $\mathbf{m}$ 满足 $\mathbf{l}^{\mathrm{T}}\mathrm{C}_{\infty}^{*}\mathbf{m} = 0$，则称两直线相对于二次曲线 $\mathrm{C}_{\infty}^{*}$ 共轭（见节 1.8.1）．

**长度比**　一旦 $\mathrm{C}_{\infty}^{*}$ 被辨认，长度比同样可以测量．考察图 1.16 中顶点为 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的三角形．按三角学的正弦定理，长度比 $d(\mathbf{b},\mathbf{c}):d(\mathbf{a},\mathbf{c}) = \sin\alpha:\sin\beta$，其中 $d(\mathbf{x},\mathbf{y})$ 表示点 $\mathbf{x},\mathbf{y}$ 的欧氏距离．在 $\mathrm{C}_{\infty}^{*}$ 已被辨认的射影框架下，按（1.21），$\cos\alpha$，$\cos\beta$ 可由 $\mathbf{l}' = \mathbf{a}' \times \mathbf{b}'$，$\mathbf{m}' = \mathbf{c}' \times \mathbf{a}'$ 和 $\mathbf{n}' = \mathbf{b}' \times \mathbf{c}'$ 计算出来．因此 $\sin\alpha$，$\sin\beta$ 从而比率 $d(\mathbf{a},\mathbf{b}):d(\mathbf{c},\mathbf{a})$ 可由射影映射后的点来确定．

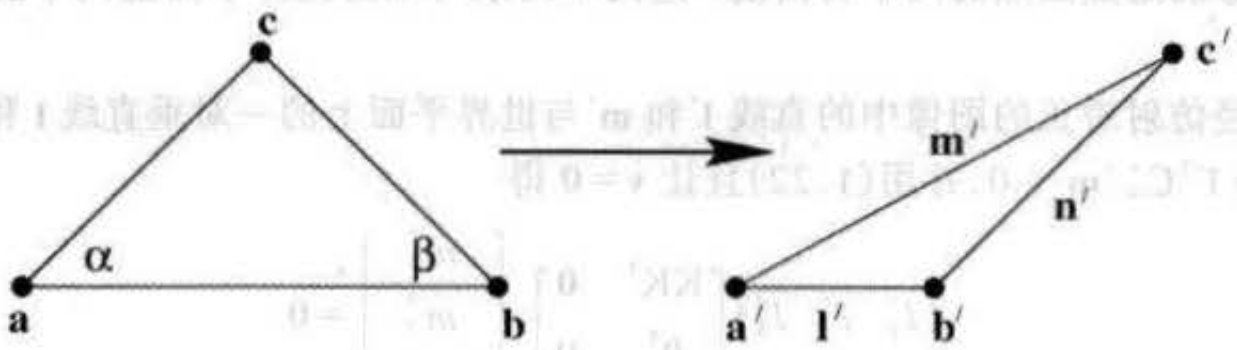



图 1.16　**长度比**．一旦 $\mathrm{C}_{\infty}^{*}$ 被辨认，欧氏长度比 $d(\mathbf{b},\mathbf{c}):d(\mathbf{a},\mathbf{c})$ 可以由射影失真的图形中测量出来．详见正文．

## 1.7.5　由图像恢复度量性质

用完全类似于节 1.7.2 和图 1.12 中通过辨认 $\mathbf{l}_{\infty}$ 来恢复仿射性质的途径，把虚圆点变换到它们的标准位置，就可以由平面的图像恢复度量性质．假定在图像上的虚圆点已被辨认，并且图像已用射影变换 H 矫正使被影像的虚圆点映回到它们在 $\mathbf{l}_{\infty}$ 上的标准位置（$(1, \pm i, 0)^{\mathrm{T}}$）．由结论 1.20 可知，世界平面和被矫正的图像之间的变换是相似变换，因为它是保持虚圆点不变的射影变换．

**用 $\mathrm{C}_{\infty}^{*}$ 进行度量矫正**　对偶二次曲线 $\mathrm{C}_{\infty}^{*}$ 几乎包含了实现度量矫正所需的全部信息．它能确定射影变换中的仿射和射影成分，而只留下相似变换的失真．这一点可以由它在射影下的变换得到证明．如果点变换是 $\mathbf{x}' = \mathrm{H}\mathbf{x}$，其中 $\mathbf{x}$ 是欧氏坐标而 $\mathbf{x}'$ 是射影坐标，$\mathrm{C}_{\infty}^{*}$ 按结论 1.14（$p$11）（$\mathrm{C}^{*\prime} = \mathrm{HC}^{*}\mathrm{H}^{\mathrm{T}}$）进行变换．由 H 的分解链（1.16－$p$15）可推出：

$$\begin{aligned}\mathrm{C}_{\infty}^{*\prime} &= (\mathrm{H_P H_A H_S})\mathrm{C}_{\infty}^{*}(\mathrm{H_P H_A H_S})^{\mathrm{T}} = (\mathrm{H_P H_A})(\mathrm{H_S C}_{\infty}^{*}\mathrm{H_S^T})(\mathrm{H_A^T H_P^T}) \\ &= (\mathrm{H_P H_A})\mathrm{C}_{\infty}^{*}(\mathrm{H_A^T H_P^T})\end{aligned}$$

在点变换 $\mathbf{x}' = \mathrm{H_S}\mathbf{x}$ 下，采用直线(1.6－$p$10)($\mathbf{l}' = \mathrm{H}^{-\mathsf{T}}\mathbf{l}$)和对偶二次曲线(1.14($p$11))($\mathrm{C}^{*\prime} = \mathrm{HC}^{*}\mathrm{H}^{\mathsf{T}}$)的变换规则，可以证明(1.21)在射影变换下不变．例如，分子变换为：

$$\mathbf{l}^{\mathsf{T}}\mathrm{C}_{\infty}^{*}\mathbf{m} \mapsto \mathbf{l}^{\mathsf{T}}\mathrm{H}^{-1}\mathrm{HC}_{\infty}^{*}\mathrm{H}^{\mathsf{T}}\mathrm{H}^{-\mathsf{T}}\mathbf{m} = \mathbf{l}^{\mathsf{T}}\mathrm{C}_{\infty}^{*}\mathbf{m}.$$

同样可以证明齐次对象的缩放因子在分子和分母之间相消．因此(1.21)的确在射影框架下不变．所证明的结果概括为：

**结论 1.22　一旦二次曲线 $\mathrm{C}_{\infty}^{*}$ 在射影平面上被辨认，那么欧氏角可以用(1.21)来测量．**

作为它的推论有

**结论 1.23　如果 $\mathbf{l}^{\mathsf{T}}\mathrm{C}_{\infty}^{*}\mathbf{m} = 0$，则直线 $\mathbf{l}$ 和 $\mathbf{m}$ 正交．**

几何上，如果直线 $\mathbf{l}$ 和 $\mathbf{m}$ 满足 $\mathbf{l}^{\mathsf{T}}\mathrm{C}_{\infty}^{*}\mathbf{m} = 0$，则称两直线相对于二次曲线 $\mathrm{C}_{\infty}^{*}$ 共轭(见节1.8.1)．

**长度比**　一旦 $\mathrm{C}_{\infty}^{*}$ 被辨认，长度比同样可以测量．考察图1.16中顶点为 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的三角形．按三角学的正弦定理，长度比 $d(\mathbf{b},\mathbf{c}):d(\mathbf{a},\mathbf{c}) = \sin\alpha:\sin\beta$，其中 $d(\mathbf{x},\mathbf{y})$ 表示点 $\mathbf{x},\mathbf{y}$ 的欧氏距离．在 $\mathrm{C}_{\infty}^{*}$ 已被辨认的射影框架下，按(1.21)，$\cos\alpha$，$\cos\beta$ 可由 $\mathbf{l}' = \mathbf{a}' \times \mathbf{b}'$，$\mathbf{m}' = \mathbf{c}' \times \mathbf{a}'$ 和 $\mathbf{n}' = \mathbf{b}' \times \mathbf{c}'$ 计算出来．因此 $\sin\alpha$，$\sin\beta$ 从而比率 $d(\mathbf{a},\mathbf{b}):d(\mathbf{c},\mathbf{a})$ 可由射影映射后的点来确定．

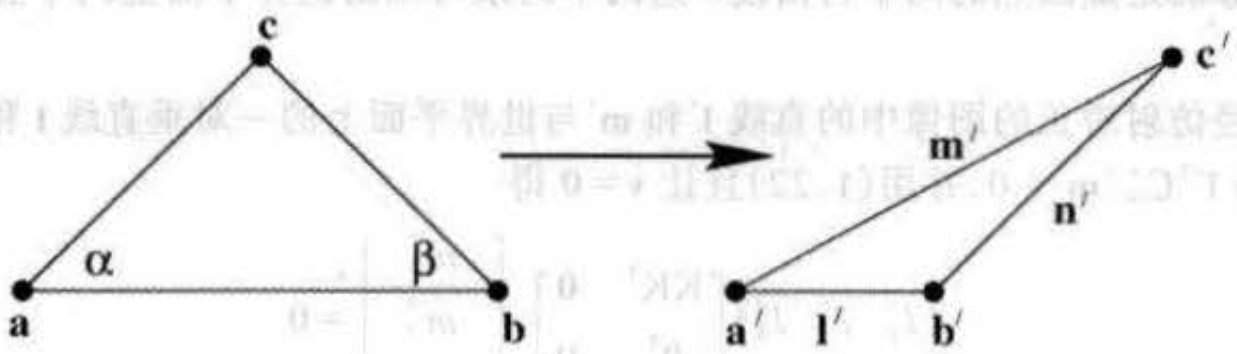



图 1.16　**长度比**．一旦 $\mathrm{C}_{\infty}^{*}$ 被辨认，欧氏长度比 $d(\mathbf{b},\mathbf{c}):d(\mathbf{a},\mathbf{c})$ 可以由射影失真的图形中测量出来．详见正文．

### 1.7.5　由图像恢复度量性质

用完全类似于节1.7.2和图1.12中通过辨认 $\mathbf{I}_{\infty}$ 来恢复仿射性质的途径，把虚圆点变换到它们的标准位置，就可以由平面的图像恢复度量性质．假定在图像上的虚圆点已被辨认，并且图像已用射影变换 H 矫正使被影像的虚圆点映回到它们在 $\mathbf{I}_{\infty}$ 上的标准位置($(1,\pm i,0)^{\mathsf{T}}$)．由结论1.20可知，世界平面和被矫正的图像之间的变换是相似变换，因为它是保持虚圆点不变的射影变换．

**用 $\mathrm{C}_{\infty}^{*}$ 进行度量矫正**　对偶二次曲线 $\mathrm{C}_{\infty}^{*}$ 几乎包含了实现度量矫正所需的全部信息．它能确定射影变换中的仿射和射影成分，而只留下相似变换的失真．这一点可以由它在射影下的变换得到证明．如果点变换是 $\mathbf{x}' = \mathrm{H}\mathbf{x}$，其中 $\mathbf{x}$ 是欧氏坐标而 $\mathbf{x}'$ 是射影坐标，$\mathrm{C}_{\infty}^{*}$ 按结论1.14($p$11)($\mathrm{C}^{*\prime} = \mathrm{HC}^{*}\mathrm{H}^{\mathsf{T}}$)进行变换．由 H 的分解链(1.16－$p$15)可推出：

$$\begin{aligned}\mathrm{C}_{\infty}^{*\prime} &= (\mathrm{H_P H_A H_S})\mathrm{C}_{\infty}^{*}(\mathrm{H_P H_A H_S})^{\mathsf{T}} = (\mathrm{H_P H_A})(\mathrm{H_S C}_{\infty}^{*}\mathrm{H_S^T})(\mathrm{H_A^T H_P^T}) \\ &= (\mathrm{H_P H_A})\mathrm{C}_{\infty}^{*}(\mathrm{H_A^T H_P^T})\end{aligned}$$

$$= \begin{bmatrix} KK^T & KK^T\mathbf{v} \\ \mathbf{v}^T KK^T & \mathbf{v}^T KK^T \mathbf{v} \end{bmatrix}. \tag{1.22}$$

显然射影成分($\mathbf{v}$)和仿射成分(K)可以直接由 $C_\infty^*$ 的像确定,但(因为根据结论 1.21,$C_\infty^*$ 在相似变换下不变)相似成分不能确定. 因而,

**结论 1.24　在射影平面上,一旦 $C_\infty^*$ 被辨认,那么射影失真可以矫正到相差一个相似变换.**

实际上,利用 SVD(节 A3.3($p$408)),可以直接从图像中已辨认的 $C_\infty^{*\prime}$ 获得所需的矫正单应变换;先将 $C_\infty^{*\prime}$ 的 SVD 写为

$$C_\infty^{*\prime} = U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^T$$

然后通过查对(1.22),求得相差一个相似变换的矫正射影变换为 $H = U$.

下面两个例子给出 $C_\infty^*$ 可能在图像中被辨认,从而能获得度量矫正的典型情形.

**例 1.25　度量矫正 I**

假定一幅图像已经仿射矫正(如上文例 1.18),那么为了确定度量矫正,我们需要两个约束以便能够确定虚圆点的两个自由度. 这两个约束可以由世界平面上两个直角的影像来获得.

假设已经仿射矫正的图像中的直线 $\mathbf{l}'$ 和 $\mathbf{m}'$ 与世界平面上的一对垂直线 $\mathbf{l}$ 和 $\mathbf{m}$ 对应. 由结论 1.23 得 $\mathbf{l}'^T C_\infty^{*\prime} \mathbf{m}' = 0$,并用(1.22)且让 $\mathbf{v} = \mathbf{0}$ 得

$$(l_1' \quad l_2' \quad l_3') \begin{bmatrix} KK^T & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \begin{pmatrix} m_1' \\ m_2' \\ m_3' \end{pmatrix} = 0$$

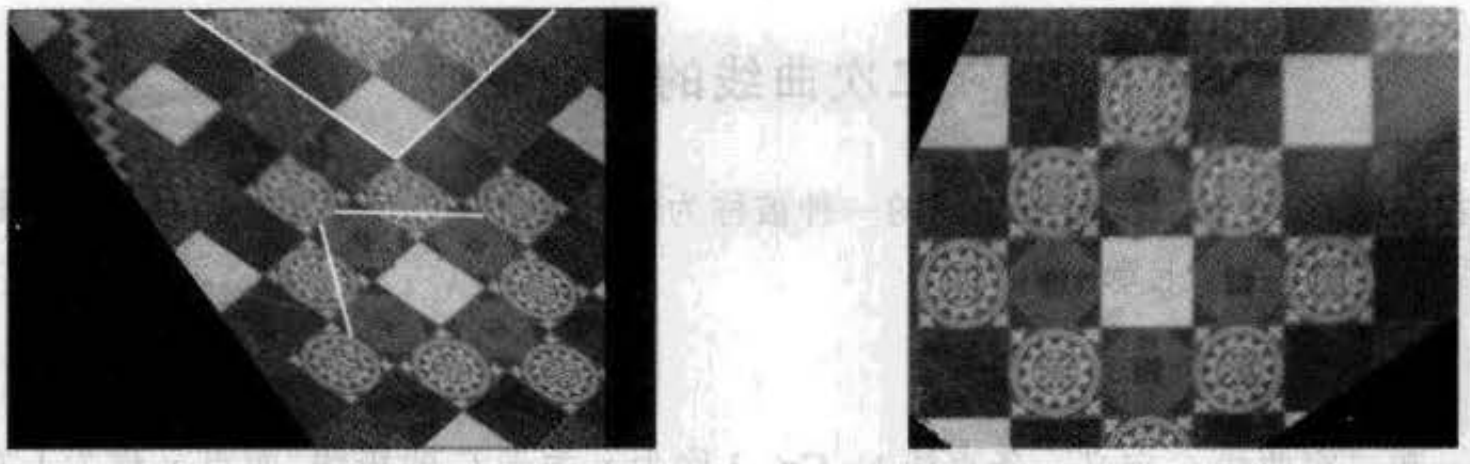
它是关于 $2\times2$ 矩阵 $S = KK^T$ 的线性约束. 矩阵 $S = KK^T$ 是对称矩阵并有三个独立元素,因而有两个自由度(因为全局的尺度因子无关紧要). 该正交条件约减为方程 $(l_1', l_2')S(m_1', m_2')^T = 0$. 并可重写为:

$$(l_1'm_1', l_1'm_2' + l_2'm_1', l_2'm_2')\mathbf{s} = 0,$$

其中 $\mathbf{s} = (s_{11}, s_{12}, s_{22})^T$ 是 S 的 3 维矢量形式. 两个这样的正交直线对就能提供两个约束,并可以联合起来给出以 $\mathbf{s}$ 为零矢量的 $2\times3$ 矩阵. 这样,在相差一个尺度因子的情况下获得 S,并进一步获得 K(利用 Cholesky 分解法:节 A3.2.1($p$407)). 图 1.17 给出一个例子,它在已进行过仿射矫正的图 1.13 上用两组正交直线对进行度量矫正. △

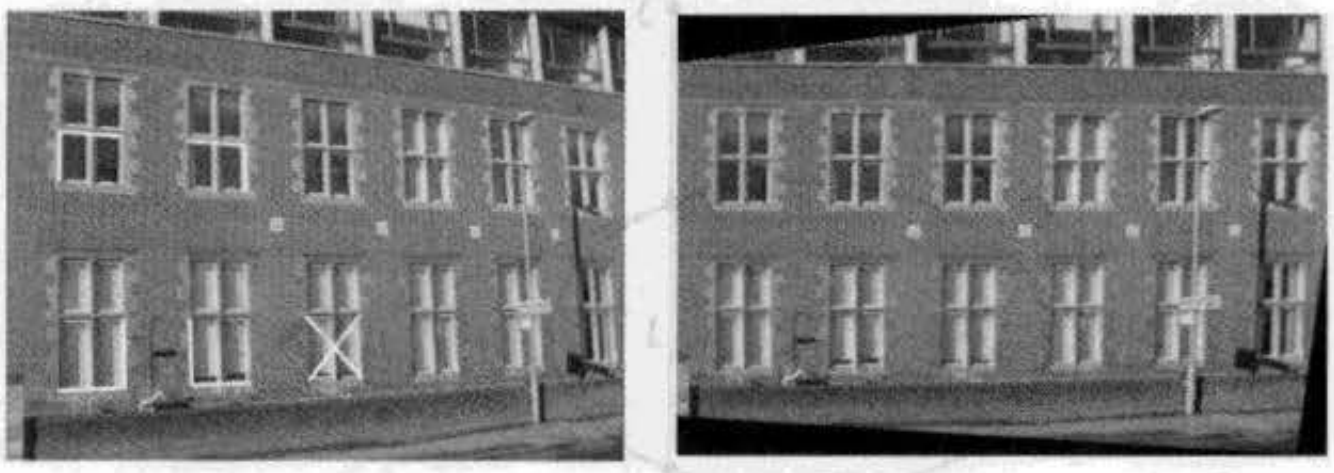
另外,度量矫正所需的两个约束可以通过一个圆的影像或两个已知的长度比来得到. 对于圆的情形,其像在仿射矫正过的图像中是椭圆,该椭圆和(已知)$\mathbf{l}_\infty$ 的交点直接确定被影像的虚圆点.

在下面的例子中,二次曲线 $C_\infty^*$ 可以用另一种方法直接在一幅透视图像中加以确定,而不用首先辨认 $\mathbf{l}_\infty$.



a　　b

图 1.17　**通过正交直线进行度量矫正 I.** 求对仿射图像进行度量矫正的仿射变换可以从被影像的正交直线中计算出来. (a)在仿射矫正过的图像(图 1.13)上,两(非平行)直线对被认定与世界平面上的正交直线相对应. (b)度量矫正图像. 注意在度量矫正过的图像中,所有在世界中正交的直线是正交的,世界正方形的长宽比为 1,而世界圆是圆的.



a　　b

图 1.18　**通过正交直线进行度量矫正 II.** (a)二次曲线 $C_\infty^*$ 用图中显示的五个正交直线对在透视图像平面(建筑物的前墙)上被确定. 二次曲线 $C_\infty^*$ 确定虚圆点,并等价于确定了度量矫正图像所需的射影变换. (b)图(a)与图 1.4($p10$)是同一个透视图像,在那里图像的透视失真是通过确定四个图像点的世界位置来消除的.

**例 1.26　度量矫正 II**

这里我们从平面的原有透视图像(不像例 1.25 中用仿射矫正过的图像)入手. 假定直线 $\mathbf{l}$ 和 $\mathbf{m}$ 是世界平面上两正交直线的图像;则按结论 1.23,$\mathbf{l}^{\mathrm{T}}C_\infty^*\mathbf{m}=0$,然后用与(1.4 − $p6$)(约束二次曲线使之过一点)类似的方式,提供关于 $C_\infty^*$ 元素的一个线性约束,即

$$(l_1m_1,\ (l_1m_2+l_2m_1)/2,\ l_2m_2,\ (l_1m_3+l_3m_1)/2,\ (l_2m_3+l_3m_2)/2,\ l_3m_3)\mathbf{c}=0$$

其中 $\mathbf{c}=(a,b,c,d,e,f)^{\mathrm{T}}$ 是 $C_\infty^*$ 的二次曲线矩阵(1.3 − $p6$)的 6 维矢量形式. 五个这样的约束联合起来,便形成一个 5×6 矩阵,使得 $\mathbf{c}$ 和 $C_\infty^*$ 作为其零矢量而求得. 它证明 $C_\infty^*$ 可以由世界平面上五个正交的直线对的图像线性地加以确定. 图 1.18 给出以这种直线对约束进行度量矫正的一个例子. △

**分层法**　注意:在例 1.26 中,仿射和射影失真通过确定 $C_\infty^*$ 一次性给予解决. 而前面的例 1.25则先消除射影失真,然后消除仿射失真. 这种两步法称为分层法. 类似的方式适用于 3D 情况,并用于第 9 章的 3D 重构和第 18 章的自标定,它们由 3D 射影重构获得度量重构.

## 1.8 二次曲线的其他性质

现在介绍点、线和二次曲线之间的一种被称为**配极**的重要几何关系. 这种(正交性表示的)关系的应用将在第七章中给出.

### 1.8.1 极点-极线关系

点 $\mathbf{x}$ 和二次曲线 C 定义一条直线 $\mathbf{l}=C\mathbf{x}$. $\mathbf{l}$ 称为 $\mathbf{x}$ 关于 C 的**极线**,而点 $\mathbf{x}$ 称为 $\mathbf{l}$ 关于 C 的**极点**.

- **点 $\mathbf{x}$ 关于二次曲线 C 的极线 $\mathbf{l}=C\mathbf{x}$ 与 C 交于两点. C 的过这两点的两条切线相交于 $\mathbf{x}$.**

这个关系在图 1.19 中给予说明.

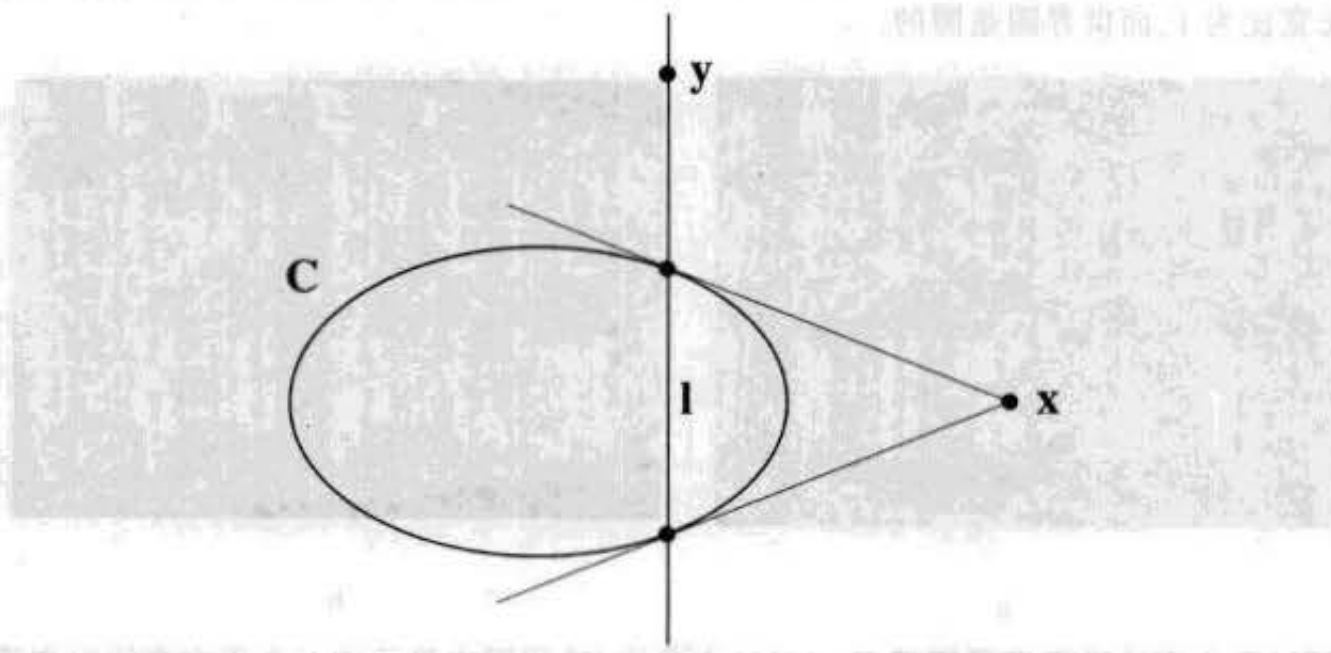



图 1.19 **极点——极线关系**. 直线 $\mathbf{l}=C\mathbf{x}$ 是点 $\mathbf{x}$ 关于二次曲线 C 的极线,而点 $\mathbf{x}=C^{-1}\mathbf{l}$ 是 $\mathbf{l}$ 关于 C 的极点. $\mathbf{x}$ 的极线与二次曲线的交点是 $\mathbf{x}$ 到二次曲线的切线的切点. 如果 $\mathbf{y}$ 在 $\mathbf{l}$ 上,则 $\mathbf{y}^T\mathbf{l}=\mathbf{y}^TC\mathbf{x}=0$. 满足 $\mathbf{y}^TC\mathbf{x}=0$的点 $\mathbf{x}$ 和 $\mathbf{y}$ 共轭.

**证明** 考察 C 上的一点 $\mathbf{y}$. C 的过 $\mathbf{y}$ 的切线是 $C\mathbf{y}$. 若满足 $\mathbf{x}^TC\mathbf{y}=0$, 则 $\mathbf{x}$ 就在此切线上. 利用 C 的对称性,条件 $\mathbf{x}^TC\mathbf{y}=(C\mathbf{x})^T\mathbf{y}=0$ 表明点 $\mathbf{y}$ 在极线 $C\mathbf{x}$ 上. 因此,极线 $C\mathbf{x}$ 与二次曲线交于 $\mathbf{y}$,$\mathbf{x}$ 在过 $\mathbf{y}$ 的切线上. □

在 $\mathbf{x}$ 趋近二次曲线的过程中,两切线变得越来越接近共线, 且它们与二次曲线上的接触点也变得越来越靠近. 在极限位置时,$\mathbf{x}$ 在 C 上,它的极线与 C 有二阶接触点 $\mathbf{x}$, 于是我们得到:

- **如果点 $\mathbf{x}$ 在 C 上,则它的极线就是二次曲线过 $\mathbf{x}$ 点的切线.**

见结论 1.7($p7$).

**例 1.27** 半径为 $r$,中心位于 $x$ 轴上的点 $x=a$ 处的圆的方程为$(x-a)^2+y^2=r^2$,并用二次曲线矩阵表示为

$$C=\begin{bmatrix}1 & 0 & -a\\ 0 & 1 & 0\\ -a & 0 & a^2-r^2\end{bmatrix}.$$

原点的极线由 $\mathbf{l}=C(0,0,1)^T=(-a,0,a^2-r^2)^T$ 给定. 这是在 $x=(a^2-r^2)/a$ 处的竖直线. 如果 $r=a$, 原点在圆周上. 在这种情况下,其极线是 $y$-轴并与圆周相切. △

显然,二次曲线诱发了 $\mathbb{P}^2$ 中点与直线之间的一个映射. 这个映射具有射影结构,因为它仅仅涉及相交和相切这两种在射影变换下不变的性质. 点与直线之间的射影变换称为**对射**(不幸的是此名字还有许多别的用处).

**定义 1.28** **对射**是 $\mathbb{P}^2$ 点到 $\mathbb{P}^2$ 线的可逆映射. 并用一个 $3\times3$ 非奇异矩阵 A 表示为 $\mathbf{l}=A\mathbf{x}$.

对射提供了点与直线关系对偶化的一种系统方法. 对射不要求一定要用对称矩阵表示,但因为讨论二次曲线,在这里我们只限于考虑对称的对射.

- **共轭点** **如果点 $\mathbf{y}$ 在极线 $\mathbf{l}=C\mathbf{x}$ 上,则 $\mathbf{y}^T\mathbf{l}=\mathbf{y}^TC\mathbf{x}=0$. 满足 $\mathbf{y}^TC\mathbf{x}=0$ 的任何两点 $\mathbf{x},\mathbf{y}$ 称为关于二次曲线 C 共轭.**

  共轭关系是对称的:

- **如果 $\mathbf{x}$ 在 $\mathbf{y}$ 的极线上,那么 $\mathbf{y}$ 也在 $\mathbf{x}$ 的极线上.**

这一点很简单,因为二次曲线矩阵是对称的——如果 $\mathbf{x}^TC\mathbf{y}=0$,则 $\mathbf{x}$ 在 $\mathbf{y}$ 的极线上,而且如果 $\mathbf{y}^TC\mathbf{x}=0$,则 $\mathbf{y}$ 在 $\mathbf{x}$ 的极线上. 因为 $\mathbf{x}^TC\mathbf{y}=\mathbf{y}^TC\mathbf{x}$,如果其中一边为零,则另一边也为零. 同时,如果 $\mathbf{l}^TC^*\mathbf{m}=0$,则存在一个关于直线 $\mathbf{l}$ 和 $\mathbf{m}$ 的对偶共轭关系.

### 1.8.2 二次曲线的分类

本节介绍二次曲线的射影和仿射分类.

**二次曲线的射影标准形式** 因为 C 是对称矩阵,所以有实特征值并可分解为乘积 $C=U^TDU$ (见节 A3.2($p$406)),其中 U 是正交矩阵,而 D 是对角矩阵. 以射影变换 U 作用于二次曲线 C,则 C 被变换成另一条二次曲线 $C'=U^{-T}CU^{-1}=U^{-T}U^TDUU^{-1}=D$. 这表明任何二次曲线都射影等价于一个由对角矩阵表示的二次曲线. 令 $D=\mathrm{diag}(\varepsilon_1d_1,\ \varepsilon_2d_2,\ \varepsilon_3d_3)$ 其中 $\varepsilon_i=\pm1$ 或 0 且 $d_i>0$. 则 D 可以写为,

$$D=\mathrm{diag}(s_1,\ s_2,\ s_3)^T\,\mathrm{diag}(\varepsilon_1,\ \varepsilon_2,\ \varepsilon_3)\mathrm{diag}(s_1,\ s_2,\ s_3)$$

其中 $s_i^2=d_i$. 注意 $\mathrm{diag}(s_1,\ s_2,\ s_3)^T=\mathrm{diag}(s_1,\ s_2,\ s_3)$. 现在用变换 $\mathrm{diag}(s_1,\ s_2,\ s_3)$ 再进行一次变换,二次曲线 D 被变为具有矩阵 $\mathrm{diag}(\varepsilon_1,\ \varepsilon_2,\ \varepsilon_3)$ 的二次曲线,其中 $\varepsilon_i=\pm1$ 或 0.

| 对角线 | 方程 | 二次曲线类型 |
|---|---|---|
| (1,1,1) | $x^2+y^2+w^2=0$ | 假二次曲线——无实点 |
| (1,1,−1) | $x^2+y^2-w^2=0$ | 圆 |
| (1,1,0) | $x^2+y^2=0$ | 单个实点 $(0,0,1)^T$ |
| (1,−1,0) | $x^2-y^2=0$ | 两条直线 $x=\pm y$ |
| (1,0,0) | $x^2=0$ | 单条直线 $x=0$ 计两次 |

表 1.2 **点二次曲线的射影分类**. 任何平面二次曲线都射影等价于表中给出的一种类型. 对某 $i$, 有 $\varepsilon_i=0$ 的那些二次曲线是退化的二次曲线,它们可以用秩小于 3 的矩阵表示. 本表的二次曲线类型栏仅介绍二次曲线的实点——例如记为复二次曲线时 $x^2+y^2=0$ 由一对直线 $x=\pm iy$ 组成.

可以用置换矩阵进一步变换，以保证值 $\varepsilon_i = 1$ 出现在值 $\varepsilon_i = -1$ 之前而后者又在值 $\varepsilon_i = 0$ 之前. 最后如果有必要，可以乘以 $-1$ 以保证 $+1$ 至少和 $-1$ 一样多. 现在可以列举各种类型的二次曲线，并在表 1.2 中给出.

**二次曲线的仿射分类** 众所周知，在欧氏几何中，(非退化或真)二次曲线可以分为双曲线、椭圆和抛物线. 如上文所示，在射影几何中这三种类型的二次曲线却都射影等价于圆. 然而在仿射几何中，上述欧氏分类仍有效，因为它仅取决于 $\mathbf{l}_\infty$ 与二次曲线的关系. 三种类型的二次曲线与 $\mathbf{l}_\infty$ 的关系在图 1.20 中给出了说明.

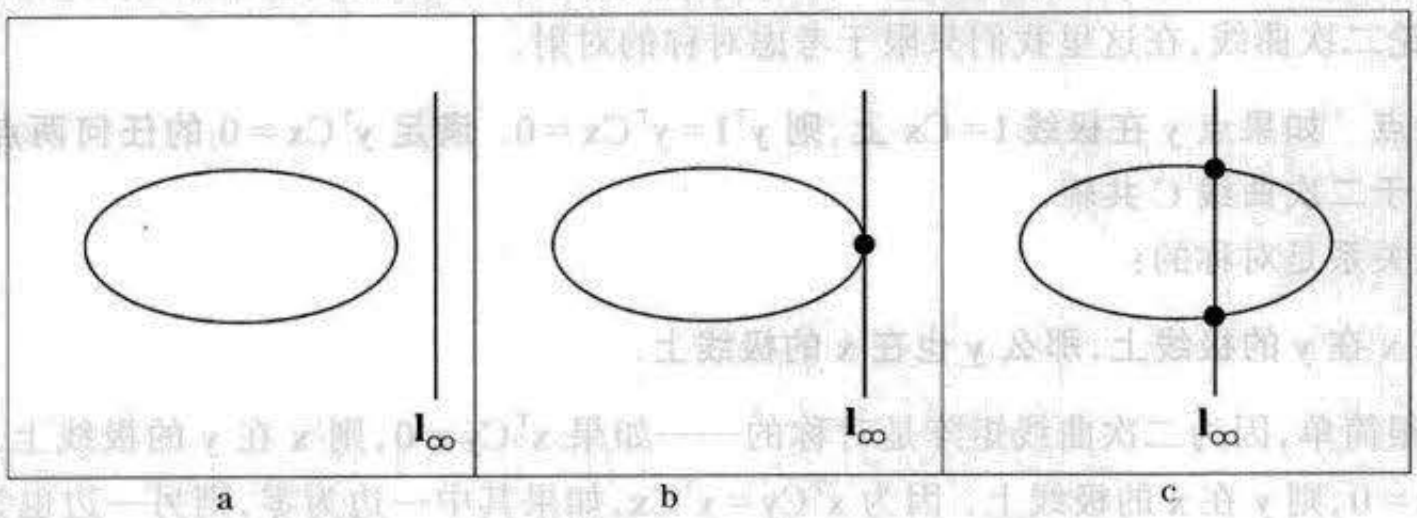



图 1.20 **点二次曲线的仿射分类**. 二次曲线是(a)椭圆，(b)抛物线，(c)双曲线；分别取决于它们与 $\mathbf{l}_\infty$ 的关系：(a)无实交点、(b)相切(2 点接触)、(c)有 2 个实交点. 在仿射变换下，$\mathbf{l}_\infty$ 是不动线而且交点保持不变. 因此，这种分类在仿射变换下不变.

## 1.9 不动点与直线

由 $\mathbf{l}_\infty$ 和虚圆点的例子，我们已经知道点和直线在射影变换下可能是不动的. 本节将对该思想作更彻底地研究.

这里，我们把源平面和目标平面视为等同(一样)，这样可把点 $\mathbf{x}$ 映射到点 $\mathbf{x}'$ 的变换在同一坐标系中进行. 关键思想是变换的一个**特征矢量**对应一个**不动点**，因为对于特征值 $\lambda$ 及其对应的特征矢量 $\mathbf{e}$ 有

$$\mathbf{H}\mathbf{e} = \lambda\mathbf{e}$$

而 $\mathbf{e}$ 和 $\lambda\mathbf{e}$ 表示同一点. 通常在计算机视觉应用中，特征矢量和特征值具有物理的或几何的重要意义.

一个 $3\times3$ 矩阵有三个特征值，如果特征值互不相同，则一个平面射影变换最多有三个不动点. 因为在此情形中特征方程是三次方程，特征值及其对应的特征矢量中有一个或三个是实的. 类似的推导可以用于不动直线，它对应于 $\mathbf{H}^{\mathrm{T}}$ 的特征矢量，因为直线的变换(1.6 $-p10$)为 $\mathbf{l}' = \mathbf{H}^{-\mathrm{T}}\mathbf{l}$.

不动点和不动直线之间的关系在图 1.21 中显示. 注意直线的不动是集合不动，不是点点不动，即该直线上的一点被映射到该直线上的另一点，这两点一般不相同. 这并不难理解：平面射影变换诱导直线上的一个 1D 射影变换. 1D 射影变换以一个 $2\times2$ 的齐次矩阵表示(节 1.5). 对应于该 $2\times2$ 矩阵的两个特征矢量，1D 射影变换有两个不动点. 这些不动点

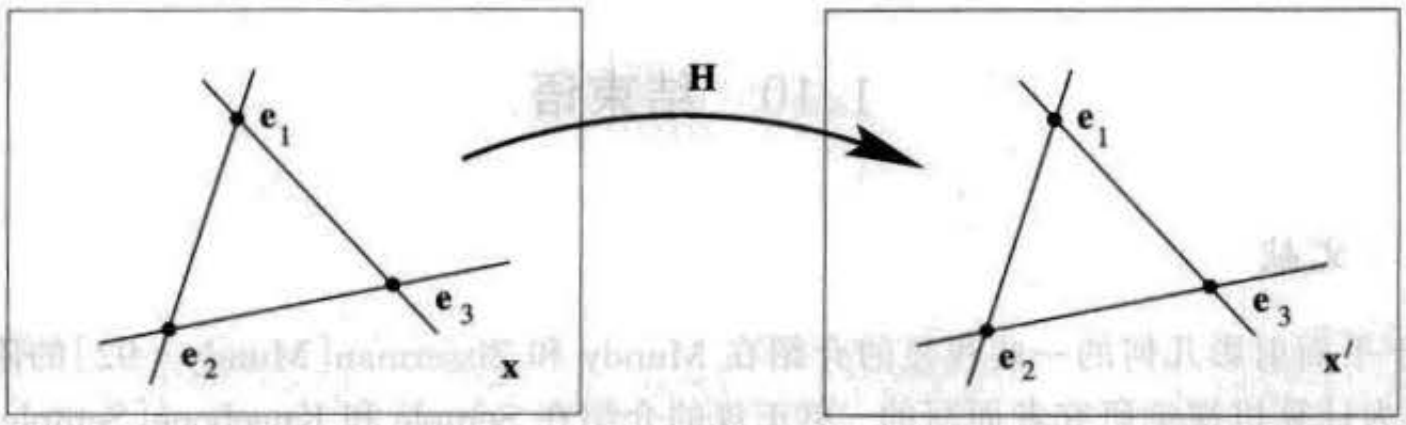



图 1.21 **平面射影变换的不动点和直线**. 这里有三个不动点和过这三点的不动直线. 不动直线和不动点可能是复的. 从代数的角度来说,不动点是点变换($\mathbf{x}' = \mathrm{H}\mathbf{x}$)的特征矢量 $\mathbf{e}_i$, 而不动直线是线变换($\mathbf{l}' = \mathrm{H}^{-\mathsf{T}}\mathbf{l}$)的特征矢量. 注意,不动直线不是点点不动:在变换下,直线上的一点被映为其上的另一点,只有不动点才映为自身.

也是 2D 射影变换的不动点.

进一步的特殊性涉及重特征值的情况. 假定两个特征值(如 $\lambda_2$, $\lambda_3$)相等,而对应于 $\lambda_2 = \lambda_3$ 存在两个不同的特征矢量($\mathbf{e}_2$, $\mathbf{e}_3$),那么包含特征矢量 $\mathbf{e}_2$、$\mathbf{e}_3$ 的直线将是点点不动的,即它是由不动点构成的直线. 假定 $\mathbf{x} = \alpha\mathbf{e}_2 + \beta\mathbf{e}_3$;则有

$$\mathrm{H}\mathbf{x} = \lambda_2\alpha\mathbf{e}_2 + \lambda_2\beta\mathbf{e}_3 = \lambda_2\mathbf{x}$$

即过两退化特征矢量的直线上的点都映射为自己(仅仅相差一比例因子). 另一种可能是 $\lambda_2 = \lambda_3$,但只有一个对应的特征矢量. 在这种情况下,特征矢量的**代数维数**为二,而**几何维数**为一. 其不动点少了一个(2 个而不是 3 个). 重特征值的各种情况将在附录 5($p$427)和 $p$33 的练习(8)中作进一步讨论.

我们现在来查看一下节 1.4 介绍的射影变换子群中的不动点和直线. 仿射变换以及比它更特殊的变换有两个特征矢量,它们都是理想点($x_3 = 0$)并且对应于左上角 $2\times2$ 矩阵的特征矢量. 第三个特征矢量通常是有限矢量.

**欧氏矩阵** 两个不动理想点是虚圆点 $\mathbf{I}, \mathbf{J}$ 组成的复共轭对,相对应的特征值是$\{e^{i\theta}, e^{-i\theta}\}$,这里 $\theta$ 是旋转角. 对应于特征值 1 的第三个特征矢量,称为**极点**. 欧氏变换等价于绕该点转 $\theta$ 角的纯旋转并且没有平移.

一种特殊的情况是纯平移(即 $\theta = 0$). 这时特征值三重退化. 无穷远线点点不动,且有一束过点$(t_x, t_y, 0)^{\mathsf{T}}$的不动直线,该点对应于平移方向. 因此平行于 $\mathbf{t}$ 的直线是不动的. 这是约束透视变换的一个例子(见节 A5.3($p$430)).

**相似矩阵** 两个不动理想点仍是虚圆点. 特征值是$\{1, se^{i\theta}, se^{-i\theta}\}$. 相似变换的作用可以理解为绕它的有限不动点的旋转和取 $s$ 为因子的均匀缩放. 注意虚圆点的特征值仍然表征旋转角.

**仿射矩阵** 两个不动理想点可以是实或复共轭的,但在任何一种情况下,过这些点的不动直线 $\mathbf{l}_\infty = (0,0,1)^{\mathsf{T}}$是实的.

## 1.10 结束语

### 1.10.1 文献

关于平面射影几何的一些浅显的介绍在 Mundy 和 Zisserman[Mundy－92]的附录中给出，它是为计算机视觉研究者而写的．更正规的介绍在 Semple 和 Kneebone[Semple－79]中给出，但[Springer－64]更容易读懂．

关于由平面的图像恢复仿射和度量景物性质的工作有：Collins 和 Beveridge[Collins－93]用消影线由卫星图像恢复仿射性质，而 Liebowitz 和 Zisserman[Liebowitz－98]利用平面的如直角一类的度量信息来恢复度量几何．

### 1.10.2 注释与练习

(1) **仿射变换**

(a) 证明仿射变换能把圆映射为椭圆，但不能把椭圆映射为双曲线或抛物线．

(b) 证明在仿射变换下平行的两线段的长度比不变，而不平行的线段的长度比则不是如此．

(2) **射影变换** 证明存在使过原点的单位圆不动(作为集合不动)的一种三参数簇的射影变换，即以原点为中心的单位圆被映射到以原点为中心的单位圆(提示，用结论 1.13($p$11)计算变换)．该簇的几何解释是什么？

(3) **各向同性** 证明两直线在相似变换下具有一个不变量；两直线和两点在射影变换下具有一个不变量．在这两种情况下，自由度的计算法则(结论 1.16($p$17))的等式情况不成立．证明在这两种情况中，相应的变换不能完全被确定，虽然它能够部分被确定．

(4) **不变量** 用点、直线和二次曲线的变换规则证明：

(a) 两条直线 $\mathbf{l}_1$、$\mathbf{l}_2$ 及不在其上的两点 $\mathbf{x}_1$、$\mathbf{x}_2$ 有不变量

$$I=\frac{(\mathbf{l}_1^{\mathrm{T}}\mathbf{x}_1)(\mathbf{l}_2^{\mathrm{T}}\mathbf{x}_2)}{(\mathbf{l}_1^{\mathrm{T}}\mathbf{x}_2)(\mathbf{l}_2^{\mathrm{T}}\mathbf{x}_1)}$$

(见前一问题)．

(b) 对于二次曲线 C 和在一般位置上的两点 $\mathbf{x}_1$，$\mathbf{x}_2$ 有不变量

$$I=\frac{(\mathbf{x}_1^{\mathrm{T}}\mathrm{C}\mathbf{x}_2)^2}{(\mathbf{x}_1^{\mathrm{T}}\mathrm{C}\mathbf{x}_1)(\mathbf{x}_2^{\mathrm{T}}\mathrm{C}\mathbf{x}_2)}.$$

(c) 证明计算夹角的射影不变量表达式(1.21)等价于 Laguerre 关于虚圆点交比的射影不变量表达式(见[Springer－64])．

(5) **交比** 证明在直线的射影变换下，四共线点的交比不变(1.17－$p$17)．提示：首先把直线上两点的变换写成 $\bar{\mathbf{x}}_i'=\lambda_i\mathrm{H}_{2\times2}\bar{\mathbf{x}}_i$ 和 $\bar{\mathbf{x}}_j'=\lambda_j\mathrm{H}_{2\times2}\bar{\mathbf{x}}_j$，这里的等式不相差一个比例因子，然后由行列式性质证明 $|\bar{\mathbf{x}}_i'\bar{\mathbf{x}}_j'|=\lambda_i\lambda_j\det\mathrm{H}_{2\times2}|\bar{\mathbf{x}}_i\bar{\mathbf{x}}_j|$，并由此继续．另一种推导方法在[Semple－79]中给出．

(6) **配极** 图 1.19 给出了椭圆外一点 $\mathbf{x}$ 的极线的几何作图．给出当点在椭圆内时极线

的几何作图. 提示:选择过 $\mathbf{x}$ 的任意直线. 该直线的极点是 $\mathbf{x}$ 极线上的一点.

(7) **对偶二次曲线** 证明矩阵 $[\mathbf{l}]_\times C[\mathbf{l}]_\times$ 表示一个秩为 2 的对偶二次曲线,它由直线 $\mathbf{l}$ 与(点)二次曲线 C 的两个交点组成. (记号 $[\mathbf{l}]_\times$ 在(A3.4 - $p$410)中定义).

(8) **不动点** 给出只有一个和两个不动点的 2D 单应变换的例子. 它们对应于几何维数小于代数维数的重特征值.

(9) **特殊射影变换** 假定一个景物平面上的点由一条直线的反射相关联,例如具有双边对称的平面物体. 证明该平面透视图像的点由满足 $H^2 = I$ 的射影变换 H 相关联. 进一步证明在 H 下有一条不动点组成的直线,它对应于反射直线的影像,并且 H 有一个不在此直线上的特征矢量,它是该反射方向的消影点(H 是一个平面调和透射,见节 A5.2($p$427)).

现在,假定点由有限对称旋转相联系:例如在一个六边形螺栓头上的点. 证明在这种情况下 $H^n = I$,这里 $n$ 是该对称旋转的阶(六边形为 6),H 的特征值确定旋转角,而对应于实特征值的特征矢量是对称旋转中心的影像.

# 第 2 章　3D 射影几何和变换

本章介绍 3 维射影空间 $\mathbb{P}^3$ 的性质和基本要素. 其中的许多内容是第一章中介绍的射影平面的直接推广. 例如,在 $\mathbb{P}^3$ 中用无穷远平面 $\boldsymbol{\pi}_\infty$ 上的理想点集把 3 维欧氏空间加以扩展,其中 $\boldsymbol{\pi}_\infty$ 类似于 $\mathbb{P}^2$ 中的 $\mathbf{l}_\infty$. 平行线以及平行平面相交在 $\boldsymbol{\pi}_\infty$ 上. 不言而喻,齐次坐标再一次起了重要的作用,这里所有的维数都增加了一维. 但是,由于维数的增加,一些额外的性质也随之出现. 例如,在射影平面上两条直线总是相交,但它们在 3 维空间中就不一定相交.

在阅读本章之前,读者应该熟悉第一章的概念和记号. 我们将不重复前一章中大量的内容,而把注意力集中于二者的差别以及由于多了一维而增加的几何性质.

## 2.1　点和射影变换

3 维空间的一点 $\mathbf{X}$ 用齐次坐标表示为一个 4 维矢量. 具体地说,齐次矢量 $\mathbf{X}=(X_1,X_2,X_3,X_4)^\mathsf{T}$ 当 $X_4\neq 0$ 时表示 $\mathbb{R}^3$ 中非齐次坐标为 $(X,Y,Z)^\mathsf{T}$ 的点,其中

$$X=X_1/X_4,\ Y=X_2/X_4,\ Z=X_3/X_4.$$

例如,$(X,Y,Z)^\mathsf{T}$ 的一种齐次表示是 $\mathbf{X}=(X,Y,Z,1)^\mathsf{T}$. $X_4=0$ 的齐次点表示无穷远点.

$\mathbb{P}^3$ 上的射影变换是由非奇异 $4\times 4$ 矩阵给出,它是关于齐次 4 维矢量的线性变换:$\mathbf{X}'=\mathrm{H}\mathbf{X}$. 变换矩阵 H 是齐次的并有 15 个自由度. 矩阵的 16 个元素扣去了一个全局尺度就是它的自由度数.

与平面射影变换的情况一样,该映射是保线变换(直线被映射到直线),它保留诸如直线与平面的交点等关联关系以及接触的阶.

## 2.2　平面、直线和二次曲面的表示和变换

在 $\mathbb{P}^3$ 中,点和平面对偶,它们的表示和推导均与 $\mathbb{P}^2$ 中点—线对偶类似. 在 $\mathbb{P}^3$ 中直线自对偶.

### 2.2.1　平面

在 3 维空间中,平面可以写成

$$\pi_1 X+\pi_2 Y+\pi_3 Z+\pi_4=0 \tag{2.1}$$

显然,这等式乘以一个非零常数仍然成立,所以只有平面方程系数的三对独立的比率 $\{\pi_1:\pi_2:\pi_3:\pi_4\}$ 是有意义的. 因此,在 3 维空间中一张平面有 3 个自由度. 平面的齐次表示是 4 维矢量 $\boldsymbol{\pi}=(\pi_1,\pi_2,\pi_3,\pi_4)^\mathsf{T}$.

用代换 $X\mapsto X_1/X_4$, $Y\mapsto X_2/X_4$, $Z\mapsto X_3/X_4$ 齐次化(2.1)得到

$$\pi_1 X_1+\pi_2 X_2+\pi_3 X_3+\pi_4 X_4=0$$

或更简洁地记成

$$\boldsymbol{\pi}^{\mathrm{T}}\mathbf{X}=0 \tag{2.2}$$

它表示点 $\mathbf{X}$ 在平面 $\boldsymbol{\pi}$ 上.

$\boldsymbol{\pi}$ 的前 3 个分量对应于欧氏几何中平面的法线——用非齐次记号,(2.2)就变成 3 维矢量形式下熟知的平面方程: $\mathbf{n}.\widetilde{\mathbf{X}}+d=0$,其中 $\mathbf{n}=(\pi_1,\pi_2,\pi_3)^{\mathrm{T}}$,$\widetilde{\mathbf{X}}=(X,Y,Z)^{\mathrm{T}}$,$X_4=1$ 而 $d=\pi_4$. 在此式中 $d/\|\mathbf{n}\|$ 是原点到平面的距离.

**联合与关联关系** 在 $\mathbb{P}^3$中,平面、点、直线之间存在许多几何关系. 例如

(1) 平面可由一般位置的三个点或一条直线与一个点的联合来唯一确定(一般位置指三点不共线或在后一种情形下指点不在直线上).

(2) 两张不同的平面相交于唯一的直线.

(3) 三张不同的平面相交于唯一的点.

这些关系有其代数表示,我们现在就来推导点和平面的表示. 含有直线关系的代数表示不再象 $\mathbb{P}^2$中用 3D 矢量代数表示那样简单(例如 $\mathbf{l}=\mathbf{x}\times\mathbf{y}$),所以我们将延迟到节 2.2.2 引入直线的表示之后.

**三点确定一张平面** 设三点 $\mathbf{X}_i$ 都在平面 $\boldsymbol{\pi}$ 上. 那么每点满足(2.2),从而 $\boldsymbol{\pi}^{\mathrm{T}}\mathbf{X}_i=0$, $i=1,2,3$. 将这些方程联合成一个矩阵方程

$$\begin{bmatrix}\mathbf{X}_1^{\mathrm{T}}\\ \mathbf{X}_2^{\mathrm{T}}\\ \mathbf{X}_3^{\mathrm{T}}\end{bmatrix}\boldsymbol{\pi}=\mathbf{0}. \tag{2.3}$$

因为一般位置上的三点 $\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3$ 线性无关,所以由它们作为行组成的 $3\times4$ 矩阵的秩为 3. 由这些点所定义的平面 $\boldsymbol{\pi}$ 作为它的 1 维(右)零空间被唯一地确定(相差一个常数因子). 如果矩阵的秩为 2,则零空间是 2 维的,那么这些点是共线的,并定义了以共线点组成的直线为轴的一个平面束.

在 $\mathbb{P}^2$中,点与线对偶,过两点 $\mathbf{x},\mathbf{y}$ 的直线 $\mathbf{l}$ 可以类似地用求以 $\mathbf{x}^{\mathrm{T}}$ 和 $\mathbf{y}^{\mathrm{T}}$ 为行组成的 $2\times3$ 矩阵的零空间而获得. 当然也可以由矢量代数直接得到一个更便利的公式 $\mathbf{l}=\mathbf{x}\times\mathbf{y}$. 在 $\mathbb{P}^3$ 中,类似的表示式由行列式和余子式的性质得到.

我们从矩阵 $M=[\mathbf{X},\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3]$开始,它由一般位置的点 $\mathbf{X}$ 和确定平面 $\boldsymbol{\pi}$ 的三点 $\mathbf{X}_i$ 组成. 当 $\mathbf{X}$ 在 $\boldsymbol{\pi}$ 上时,行列式 $\det M=0$,因为 $\mathbf{X}$ 可以表示为 $\mathbf{X}_i$, $i=1,2,3$ 的线性组合. 按 $\mathbf{X}$ 列展开行列式,得到

$$\det M=X_1D_{234}-X_2D_{134}+X_3D_{124}-X_4D_{123}$$

其中 $D_{jkl}$ 是由 $4\times3$ 矩阵$[\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3]$中第 $jkl$ 行组成的行列式. 因为 $\boldsymbol{\pi}$ 上的点满足$\det M=0$,我们可以求出平面系数而得到

$$\boldsymbol{\pi}=(D_{234},-D_{134},D_{124},-D_{123})^{\mathrm{T}}. \tag{2.4}$$

它就是方程(2.3)的解矢量(零空间).

**例 2.1** 设确定平面的三个点是

$$\mathbf{X}_1=\begin{pmatrix}\widetilde{\mathbf{X}}_1\\ 1\end{pmatrix}\quad \mathbf{X}_2=\begin{pmatrix}\widetilde{\mathbf{X}}_2\\ 1\end{pmatrix}\quad \mathbf{X}_3=\begin{pmatrix}\widetilde{\mathbf{X}}_3\\ 1\end{pmatrix}$$

其中 $\tilde{\mathbf{X}}_i=(X_i,Y_i,Z_i)^T$. 则

$$D_{234}=\begin{vmatrix} Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \\ 1 & 1 & 1 \end{vmatrix}=\begin{vmatrix} Y_1-Y_3 & Y_2-Y_3 & Y_3 \\ Z_1-Z_3 & Z_2-Z_3 & Z_3 \\ 0 & 0 & 1 \end{vmatrix}=((\tilde{\mathbf{X}}_1-\tilde{\mathbf{X}}_3)\times(\tilde{\mathbf{X}}_2-\tilde{\mathbf{X}}_3))_1$$

同理可得其他分量,即得

$$\boldsymbol{\pi}=\begin{pmatrix}(\tilde{\mathbf{X}}_1-\tilde{\mathbf{X}}_3)\times(\tilde{\mathbf{X}}_2-\tilde{\mathbf{X}}_3) \\ -\tilde{\mathbf{X}}_3^T(\tilde{\mathbf{X}}_1\times\tilde{\mathbf{X}}_2)\end{pmatrix}.$$

这是欧氏矢量几何中熟知的结果,例如,平面的法线由$(\tilde{\mathbf{X}}_1-\tilde{\mathbf{X}}_3)\times(\tilde{\mathbf{X}}_2-\tilde{\mathbf{X}}_3)$来计算. △

**三平面确定一点** 这里的推导对偶于三点确定一张平面的情形. 三张平面 $\boldsymbol{\pi}_i$ 的交点 $\mathbf{X}$ 可通过求以三张平面为行的 $3\times4$ 矩阵的(右)零空间直接计算出来:

$$\begin{bmatrix}\boldsymbol{\pi}_1^T \\ \boldsymbol{\pi}_2^T \\ \boldsymbol{\pi}_3^T\end{bmatrix}\mathbf{X}=\mathbf{0}. \tag{2.5}$$

类似于(2.4),$\mathbf{X}$ 的一个直接解可以用行列式的 $3\times3$ 子矩阵表示,而数值解的计算由算法 A3.2($p$410)获得.

下面两个结论是对应的 2D 情形的直接类推.

**射影变换** 在点变换 $\mathbf{X}'=\mathrm{H}\mathbf{X}$ 下,平面变换为

$$\boldsymbol{\pi}'=\mathrm{H}^{-T}\boldsymbol{\pi}. \tag{2.6}$$

**平面上的点的参数表示** 在平面 $\boldsymbol{\pi}$ 上的点 $\mathbf{X}$ 可以写成

$$\mathbf{X}=\mathrm{M}\mathbf{x} \tag{2.7}$$

其中 $4\times3$ 矩阵 M 的列生成 $\boldsymbol{\pi}^T$的秩为 3 的零空间,即 $\boldsymbol{\pi}^T\mathrm{M}=\mathbf{0}^T$,而 3 维矢量 $\mathbf{x}$(它是射影平面 $\mathbb{P}^2$的点)给出平面 $\boldsymbol{\pi}$ 上点的参数表示. 当然 M 不是唯一的. 设平面是 $\boldsymbol{\pi}=(a,b,c,d)^T$ 且 $a$ 非零,那么 $\mathrm{M}^T$可以写成 $\mathrm{M}^T=[\mathbf{p}|\mathrm{I}_{3\times3}]$,其中 $\mathbf{p}=(-b/a,-c/a,-d/a)^T$.

这种参数化表示就是 $\mathbb{P}^2$中的直线 $\mathbf{l}$ 在 3D 中的类推,在那里 $\mathbf{l}$ 被定义为它的 2D 零空间两个线性无关的矢量的线性组合:$\mathbf{x}=\mu\mathbf{a}+\lambda\mathbf{b}$,其中 $\mathbf{l}^T\mathbf{a}=\mathbf{l}^T\mathbf{b}=0$.

### 2.2.2 直线

两点的**连接**或两平面的相交定义一条直线. 在 3 维空间中,直线有 4 个自由度. 一种令人信服的计算直线自由度的方法如下:如图 2.1 所示,把直线看作是由它与两正交平面的相交来定义,其中每张平面上的交点由两个参数来确定,所以一条直线共有 4 个自由度.

3 维空间中的直线表示相当难处理,因为 4 个自由度的对象应该用 5 维齐次矢量表示. 但问题是 5 维齐次矢量与表示点和平面的 4 维矢量很难同时在数学表达式中使用. 为了克服这种困难,有若干种直线的表示法提出,而它们的区别在于数学复杂性不同. 我们仅考察其中三种表示. 每一种表示都提供定义直线的机制:两点的连接和对偶描述(其中直线由两平面相交定义),以及这两种定义之间的映射. 这些表示也能计算连接和关联关系,例如求直线与平面的交点.

**I. 零空间与生成子空间表示** 这种表示以直观的几何概念为基础,即直线是共线点束(单

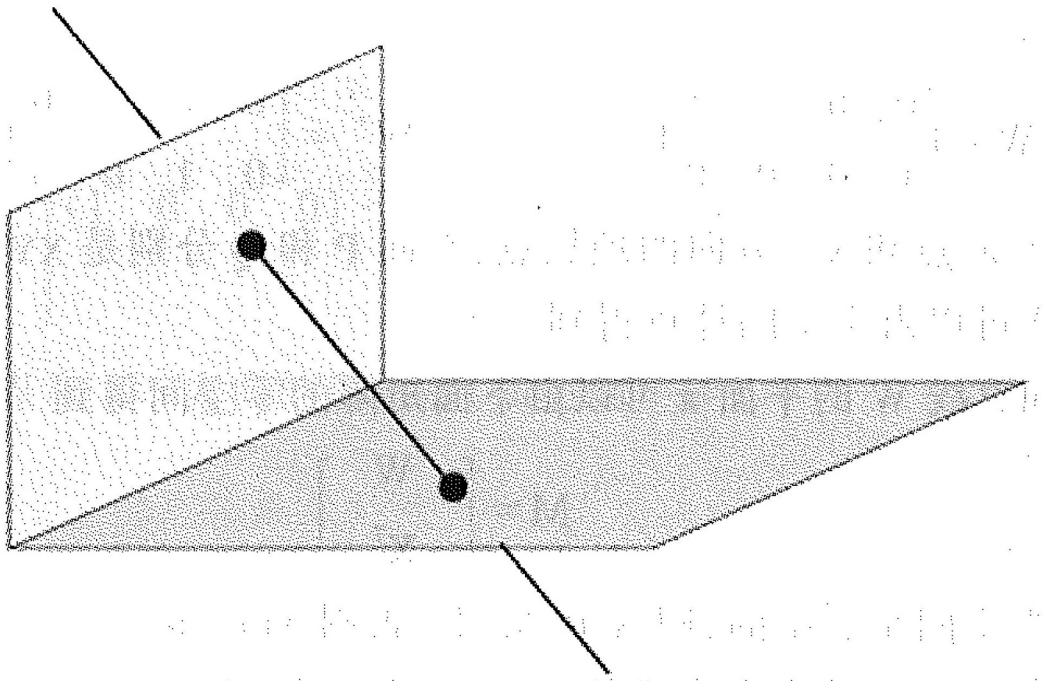


图 2.1　一条直线可以由它与两正交平面的交点来确定. 每个交点有 2 自由度,从而说明在 $\mathbb{P}^3$中的一条直线共有 4 个自由度.

参数簇)并由其中任意两点来确定. 类似地,直线是平面束的轴并且由其中的任意两平面的交线来确定. 在这两种情况中,具体的点和平面并不重要(事实上两点有 6 个自由度而且用两个 4 维矢量表示——参数用得太多). 把直线表示成两矢量的**生成子空间**,就可从数学上掌握这个概念. 假定 $\mathbf{A},\mathbf{B}$ 是两(不重合的)空间点. 那么连接这两点的直线由一个 $2\times4$ 矩阵 W 的行的生成子空间表示,W 是由 $\mathbf{A}^{\mathrm{T}}$和 $\mathbf{B}^{\mathrm{T}}$为行组成的矩阵:

$$\mathrm{W}=\begin{bmatrix}\mathbf{A}^{\mathrm{T}}\\ \mathbf{B}^{\mathrm{T}}\end{bmatrix}.$$

那么:

(1) $\mathrm{W}^{\mathrm{T}}$的生成子空间是在直线 $\lambda\mathbf{A}+\mu\mathbf{B}$ 上的点束.

(2) W 的 2 维右零空间的生成子空间是以直线为轴的平面束.

显然,在直线上的另外两点 $\mathbf{A}'^{\mathrm{T}}$和 $\mathbf{B}'^{\mathrm{T}}$产生的矩阵 $\mathrm{W}'$与 W 有相同的生成子空间,因此该生成子空间(从而直线的表示)与定义它的具体点无关.

为证明零空间性质,假定 $\mathbf{P}$ 和 $\mathbf{Q}$ 组成零空间的一组基. 那么 $\mathrm{W}\mathbf{P}=\mathbf{0}$,进而 $\mathbf{A}^{\mathrm{T}}\mathbf{P}=\mathbf{B}^{\mathrm{T}}\mathbf{P}=\mathbf{0}$,因此 $\mathbf{P}$ 是包含点 $\mathbf{A}$ 和 $\mathbf{B}$ 的平面. 同理,$\mathbf{Q}$ 是包含点 $\mathbf{A}$ 和 $\mathbf{B}$ 的另一平面. 这样一来,$\mathbf{A}$ 和 $\mathbf{B}$ 同时在(线性无关的)平面 $\mathbf{P}$ 和 $\mathbf{Q}$ 上,因而由 W 确定的直线是平面的交线. 而以该直线为轴的平面束的任何平面由生成子空间 $\lambda'\mathbf{P}+\mu'\mathbf{Q}$ 给出.

类似地,一条直线的对偶表示是两平面 $\mathbf{P},\mathbf{Q}$ 的交线. 该直线表示为以 $\mathbf{P}^{\mathrm{T}}$和 $\mathbf{Q}^{\mathrm{T}}$为行组成的一个 $2\times4$ 矩阵 $\mathrm{W}^*$

$$\mathrm{W}^*=\begin{bmatrix}\mathbf{P}^{\mathrm{T}}\\ \mathbf{Q}^{\mathrm{T}}\end{bmatrix}$$

(的行空间)的生成子空间来表示,$\mathrm{W}^*$ 具有性质:

(1) $\mathrm{W}^{*\mathrm{T}}$的生成子空间是以该直线为轴的平面束 $\lambda'\mathbf{P}+\mu'\mathbf{Q}$.

(2) $\mathrm{W}^*$ 的 2 维零空间的生成子空间是该直线上的点束.

这两种表示以 $\mathrm{W}^*\mathrm{W}^{\mathrm{T}}=\mathrm{W}\mathrm{W}^{*\mathrm{T}}=0_{2\times2}$相联系,其中 $0_{2\times2}$是 $2\times2$ 零矩阵.

**例 2.2** X－轴被表示成

$$W=\begin{bmatrix}0&0&0&1\\1&0&0&0\end{bmatrix}\qquad W^*=\begin{bmatrix}0&0&1&0\\0&1&0&0\end{bmatrix}$$

其中点 **A** 和 **B** 分别是原点和 X－方向的理想点，平面 **P** 和 **Q** 分别是 XY－和 XZ－平面. △

连接和关联关系同样可以由零空间计算得到.

(1) 包含点 **X** 和直线 W 的平面 $\boldsymbol{\pi}$ 可以由下面矩阵的零空间得到

$$M=\begin{bmatrix}W\\ \mathbf{X}^T\end{bmatrix}.$$

如果 M 的零空间是 2 维的，则 **X** 在 W 上，否则 $M\boldsymbol{\pi}=\mathbf{0}$.

(2) 由直线 W 与平面 $\boldsymbol{\pi}$ 的交点定义的点 **X** 可以通过求矩阵

$$M=\begin{bmatrix}W^*\\ \boldsymbol{\pi}^T\end{bmatrix}$$

的零空间得到. 如果 M 的零空间是 2 维的，则直线 W 在 $\boldsymbol{\pi}$ 上，否则 $M\mathbf{X}=\mathbf{0}$.

这些性质几乎可以由观察得到. 例如，第一个性质等价于三点确定一张平面(2.3).

生成子空间表示在实际的数值实现中很有用，其中的零空间可以简单地用 SVD 算法来计算(见节 A3.3 $p$408)，几乎所有的矩阵软件包都有该算法. 同样，该表示在估计问题中有用，在超参数化的估计问题中，该表示通常也不会有问题(见节 3.5($p$69)的讨论).

**II. Plücker 矩阵** 这里一条直线由 4×4 反对称齐次矩阵表示. 具体地说，连接两点 **A**、**B** 的直线由矩阵 L 表示，其元素为

$$l_{ij}=A_iB_j-B_iA_j$$

或用矢量记号等价地表示为

$$L=\mathbf{A}\mathbf{B}^T-\mathbf{B}\mathbf{A}^T \tag{2.8}$$

L 的若干主要的性质如下：

(1) L 的秩为 2. 它的 2 维零空间由以该直线为轴的平面束生成(事实上 $LW^{*T}=0$，其中 0 是一 4×2 零矩阵).

(2) 该表示具有描述一条直线所需要的 4 个自由度. 计算如下：反对称矩阵有 6 个独立的非零元素，但仅有 5 个比率是有意义的，进一步因为 $\det L=0$，所以其元素还满足一个(二次)约束(见下文). 因而净自由度数是 4.

(3) 关系 $L=\mathbf{A}\mathbf{B}^T-\mathbf{B}\mathbf{A}^T$ 是 $\mathbb{P}^2$ 中直线 **l** 的矢量积公式 $\mathbf{l}=\mathbf{x}\times\mathbf{y}$ 向 4 维空间的推广，其中确定直线的两点 **x**，**y** 都由 3 维矢量表示.

(4) 矩阵 L 与用来确定它的点 **A**，**B** 无关，因为如果用该直线上不同的点 $\mathbf{C}=\mathbf{A}+\mu\mathbf{B}$ 代替时，那么得到的矩阵是

$$\hat{L}=\mathbf{A}\mathbf{C}^T-\mathbf{C}\mathbf{A}^T=\mathbf{A}(\mathbf{A}^T+\mu\mathbf{B}^T)-(\mathbf{A}+\mu\mathbf{B})\mathbf{A}^T=\mathbf{A}\mathbf{B}^T-\mathbf{B}\mathbf{A}^T=L.$$

(5) 在点变换 $\mathbf{X}'=H\mathbf{X}$ 下，该矩阵变换为 $L'=HLH^T$，即它是一个阶为 2 的张量(见附录 1($p$400)).

**例 2.3** 根据(2.8)X－轴表示成

$$L=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}(1\quad 0\quad 0\quad 0)-\begin{pmatrix}1\\0\\0\\0\end{pmatrix}(0\quad 0\quad 0\quad 1)=\begin{bmatrix}0&0&0&-1\\0&0&0&0\\0&0&0&0\\1&0&0&0\end{bmatrix}$$

其中 **A**,**B** 分别为原点和 X－方向的理想点(如上例一样). △

由两平面 **P**,**Q** 的交线确定的直线的对偶 Plücker 表示 $L^*$ 为

$$L^*=\mathbf{P}\mathbf{Q}^{\mathrm{T}}-\mathbf{Q}\mathbf{P}^{\mathrm{T}} \tag{2.9}$$

并与 L 有相似的性质. 在点变换 $\mathbf{X}'=H\mathbf{X}$ 下,矩阵 $L^*$ 变换为 $L^{*\prime}=H^{-\mathrm{T}}L^*H^{-1}$. 矩阵 $L^*$ 可以由 L 通过简单的重写规则得到:

$$l_{12}:l_{13}:l_{14}:l_{23}:l_{42}:l_{34}=l_{34}^*:l_{42}^*:l_{23}^*:l_{14}^*:l_{13}^*:l_{12}^*. \tag{2.10}$$

其对应规则非常简单:对偶的和原来的分量的指标合在一起总包含所有的数码{1,2,3,4},因而如果原来的指标是 $ij$,那么对偶的指标是{1,2,3,4}中除去 $ij$ 后的那些数字. 例如 $12\mapsto 34$.

连接和关联性质用这些记号能表示得相当好:

(1) 由点 **X** 和直线 L 联合而确定的平面为

$$\boldsymbol{\pi}=L^*\mathbf{X}$$

并且 $L^*\mathbf{X}=\mathbf{0}$ 的充要条件是 **X** 在 L 上.

(2) 由直线 L 和平面 $\boldsymbol{\pi}$ 相交而确定的点为

$$\mathbf{X}=L\boldsymbol{\pi}$$

并且 $L\boldsymbol{\pi}=\mathbf{0}$ 的充要条件是 L 在 $\boldsymbol{\pi}$ 上.

两(或更多)条直线 $L_1,L_2,\cdots$ 的性质可以由矩阵 $M=[L_1,L_2,\ldots]$ 的零空间获得. 例如如果这些直线共面,那么 $M^{\mathrm{T}}$ 有一个 1 维零空间对应于这些直线所在的平面 $\boldsymbol{\pi}$.

**例 2.4** X 轴与平面 X=1 的交点由 $\mathbf{X}=L\boldsymbol{\pi}$ 给出:

$$\mathbf{X}=\begin{bmatrix}0&0&0&-1\\0&0&0&0\\0&0&0&0\\1&0&0&0\end{bmatrix}\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}=\begin{pmatrix}1\\0\\0\\1\end{pmatrix}$$

它的非齐次点为 $(X,Y,Z)^{\mathrm{T}}=(1,0,0)^{\mathrm{T}}$. △

**III. Plücker 直线坐标** Plücker 直线坐标是 4×4 反对称 Plücker 矩阵 L(2.8)的六个非零元素,即①

$$\mathcal{L}=\{l_{12},l_{13},l_{14},l_{23},l_{42},l_{34}\}. \tag{2.11}$$

它是 6 维齐次矢量,因而是 $\mathbb{P}^5$ 的元素. 因为 $\det L=0$,其坐标满足方程

$$l_{12}l_{34}+l_{13}l_{42}+l_{14}l_{23}=0. \tag{2.12}$$

只有当 6 维矢量 $\mathcal{L}$ 满足(2.12)式时,它才对应于一条 3 维空间的直线. 该约束的几何解释

① 习惯上用 $l_{42}$ 代替 $l_{24}$ 使它在后续的许多公式中消去了负号.

是 $\mathbb{P}^3$中的直线定义了 $\mathbb{P}^5$中一个(余维数为 1)的曲面,它称为 *Klein* **二次曲面**,之所以称为二次曲面是因为(2.12)的项是 Plücker 直线坐标的二次函数.

假定两条直线 $\mathcal{L}$,$\hat{\mathcal{L}}$分别由连接 $\mathbf{A},\mathbf{B}$ 和连接 $\hat{\mathbf{A}},\hat{\mathbf{B}}$而产生. 这些直线相交的充要条件是四点共面. 即是 $\det[\mathbf{A},\mathbf{B},\hat{\mathbf{A}},\hat{\mathbf{B}}]=0$. 可以证明此行列式可展开为

$$\begin{aligned}\det[\mathbf{A},\mathbf{B},\hat{\mathbf{A}},\hat{\mathbf{B}}] &= l_{12}\hat{l}_{34}+\hat{l}_{12}l_{34}+l_{13}\hat{l}_{42}+l_{13}\hat{l}_{42}+l_{14}\hat{l}_{23}+\hat{l}_{14}l_{23}\\ &=(\mathcal{L}\mid\hat{\mathcal{L}}).\end{aligned}\tag{2.13}$$

因为 Plücker 坐标与用来定义它们的具体点无关,双线性乘积$(\mathcal{L}\mid\hat{\mathcal{L}})$也与推导时所用的点无关而仅取决于直线 $\mathcal{L}$ 和 $\hat{\mathcal{L}}$. 这样就有

**结论 2.5 两条直线 $\mathcal{L}$ 和 $\hat{\mathcal{L}}$ 共面(因而相交)的充要条件是**$(\mathcal{L}\mid\hat{\mathcal{L}})=0$.

这个双线性乘积在许多有用的公式中出现:

(1) 如果$(\mathcal{L}\mid\mathcal{L})=0$, 则 6 维矢量 $\mathcal{L}$ 仅表示 $\mathbb{P}^3$中一条直线. 这不过是重述上面的 Klein 二次约束(2.12).

(2) 假定两条直线 $\mathcal{L}$,$\hat{\mathcal{L}}$分别是平面 $\mathbf{P},\mathbf{Q}$ 和 $\hat{\mathbf{P}},\hat{\mathbf{Q}}$ 的交线,那么

$$(\mathcal{L}\mid\hat{\mathcal{L}})=\det[\mathbf{P},\mathbf{Q},\hat{\mathbf{P}},\hat{\mathbf{Q}}]$$

同样,两直线相交的充要条件是$(\mathcal{L}\mid\hat{\mathcal{L}})=0$.

(3) 如果 $\mathcal{L}$ 是两平面 $\mathbf{P},\mathbf{Q}$ 的交线,而 $\hat{\mathcal{L}}$是两点 $\mathbf{A},\mathbf{B}$ 的连线,那么

$$(\mathcal{L}\mid\hat{\mathcal{L}})=(\mathbf{P}^{\mathrm{T}}\mathbf{A})(\mathbf{Q}^{\mathrm{T}}\mathbf{B})-(\mathbf{Q}^{\mathrm{T}}\mathbf{A})(\mathbf{P}^{\mathrm{T}}\mathbf{B}).\tag{2.14}$$

Plücker 坐标在代数推导中起作用. 在第 7 章中,它们将被用来定义 3 维空间直线到它的图像的映射.

### 2.2.3 二次曲面与对偶二次曲面

$\mathbb{P}^3$中,二次曲面由下列方程定义

$$\mathbf{X}^{\mathrm{T}}\mathbf{Q}\mathbf{X}=0\tag{2.15}$$

其中 Q 是一个 4×4 的对称矩阵. 矩阵 Q 和它定义的二次曲面经常不加区别,我们将简单地使用二次曲面 Q 的说法.

二次曲面的许多性质直接承接节 1.2.3($p6$)中的二次曲线的性质. 这里重点介绍若干性质如下:

(1) 一个二次曲面有 9 个自由度. 它们对应于 4×4 对称矩阵的 10 个独立元素再因为全局尺度的原因减去一个自由度.

(2) 一般位置上的九个点确定一个二次曲面.

(3) 如果矩阵 Q 是奇异的,那么二次曲面是**退化的**,并可以由较少的点来确定.

(4) 二次曲面定义了点和平面之间的一种配极,类似于二次曲线在点和直线之间定义的配极(节 1.8.1). 平面 $\boldsymbol{\pi}=\mathbf{Q}\mathbf{X}$ 称为是 $\mathbf{X}$ 关于 Q 的极平面. 当 Q 为非奇异并且 $\mathbf{X}$ 在二次曲面之外时,极平面由过 $\mathbf{X}$ 且与 Q 相切的射线组成的锥与 Q 相接触的点来定义. 如果 $\mathbf{X}$ 在 Q 上,那么 $\mathbf{Q}\mathbf{X}$ 是 Q 在点 $\mathbf{X}$ 的切平面.

(5) 平面 $\boldsymbol{\pi}$ 与二次曲面 Q 的交线是二次曲线 C. 计算该二次曲线可能是棘手的,因为

它要求为平面建立坐标系. (2.7)式提到一张平面的坐标系可以由 $\boldsymbol{\pi}$ 的补空间来定义即 $\mathbf{X}=M\mathbf{x}$. 如果 $\mathbf{X}^T Q\mathbf{X}=\mathbf{x}^T M^T QM\mathbf{x}=0$,则在 $\boldsymbol{\pi}$ 上的点也在 Q 上. 这些点也在一条二次曲线 C 上,因为令 $C=M^T QM$,有 $\mathbf{x}^T C\mathbf{x}=0$.

(6) 在点变换 $\mathbf{X}'=H\mathbf{X}$ 下,(点)二次曲面变换为

$$Q'=H^{-T}QH^{-1}. \tag{2.16}$$

二次曲面的对偶仍然是二次曲面. 对偶二次曲面用平面的方程定义:点二次曲面 Q 的切平面 $\boldsymbol{\pi}$ 满足 $\boldsymbol{\pi}^T Q^* \boldsymbol{\pi}=0$,其中 $Q^*$ 是 Q 的伴随矩阵,如果 Q 可逆,则 $Q^*$ 为 $Q^{-1}$. 在点变换 $\mathbf{X}'=H\mathbf{X}$ 下,对偶二次曲面变换为

$$Q^{*\prime}=HQ^* H^T. \tag{2.17}$$

对偶二次曲面的影像的代数比点二次曲面的简单得多. 这方面内容将在第 7 章中详细介绍.

### 2.2.4 二次曲面的分类

因为表示二次曲面的矩阵 Q 是对称的,它可以分解为 $Q=U^T DU$,这里 U 是实正交矩阵而 D 是实对角矩阵. 进一步地,通过将 U 的行矢量作适当的缩放可得 $Q=H^T DH$,其中 D 是对角矩阵,它的对角元素为 0,1,或 −1. 我们可更进一步地保证 D 的零元素出现在对角线的最后而 +1 出现在最前. 现在用 D 代替 $Q=H^T DH$ 等价于用矩阵 H 进行了射影变换(见(2.16)). 因此在射影等价的意义下,我们可以假定二次曲面可以由这种简单形式的矩阵 D 表示.

对角矩阵 D 的**符号差**,记为 $\sigma(D)$,定义为 D 中 +1 的个数与 −1 的个数的差值. 通过定义 $\sigma(Q)=\sigma(D)$,符号差的定义可以扩展到任意实对称矩阵 $Q=H^T DH$(其中 H 是实矩阵). 可以证明符号差是良定的,即与 H 的具体选择无关. 因为表示一个二次曲面的矩阵可以相差符号,我们可以假定其符号差是非负的. 于是二次曲面的射影类型由它的秩和符号差唯一确定. 这样我们就能列举二次曲面的不同射影等价类.

由对角矩阵 $\mathrm{diag}(d_1, d_2, d_3, d_4)$ 所表示的二次曲面对应于满足方程 $d_1X^2+d_2Y^2+d_3Z^2+d_4T^2=0$的点集. 我们可以令 $T=1$ 得出二次曲面上的非无穷远点的方程. 参看表 2.1. 二次曲面的例子在图 2.2—图 2.4 中给出.

| 秩 | σ | 对角线 | 方程 | 实现 |
|---|---|---|---|---|
| 4 | 4 | (1,1,1,1) | $X^2+Y^2+Z^2+1=0$ | 无实点 |
| | 2 | (1,1,1,−1) | $X^2+Y^2+Z^2=1$ | 球面 |
| | 0 | (1,1,−1,−1) | $X^2+Y^2=Z^2+1$ | 单叶双曲面 |
| 3 | 3 | (1,1,1,0) | $X^2+Y^2+Z^2=0$ | 一点$(0,0,0,1)^T$ |
| | 1 | (1,1,−1,0) | $X^2+Y^2=Z^2$ | 过原点的圆锥 |
| 2 | 2 | (1,1,0,0) | $X^2+Y^2=0$ | 单条直线(Z−轴) |
| | 0 | (1,−1,0,0) | $X^2=Y^2$ | 两平面 $X=\pm Y$ |
| 1 | 1 | (1,0,0,0) | $X^2=0$ | 平面 $X=0$ |

表 2.1 **点二次曲面的分类**

**直纹二次曲面** 二次曲面分为两类——直纹和非直纹二次曲面. 直纹二次曲面是包含直线的二次曲面. 更具体一点讲,非退化的直纹二次曲面(单叶双曲面)包含两个称为**母线**的直

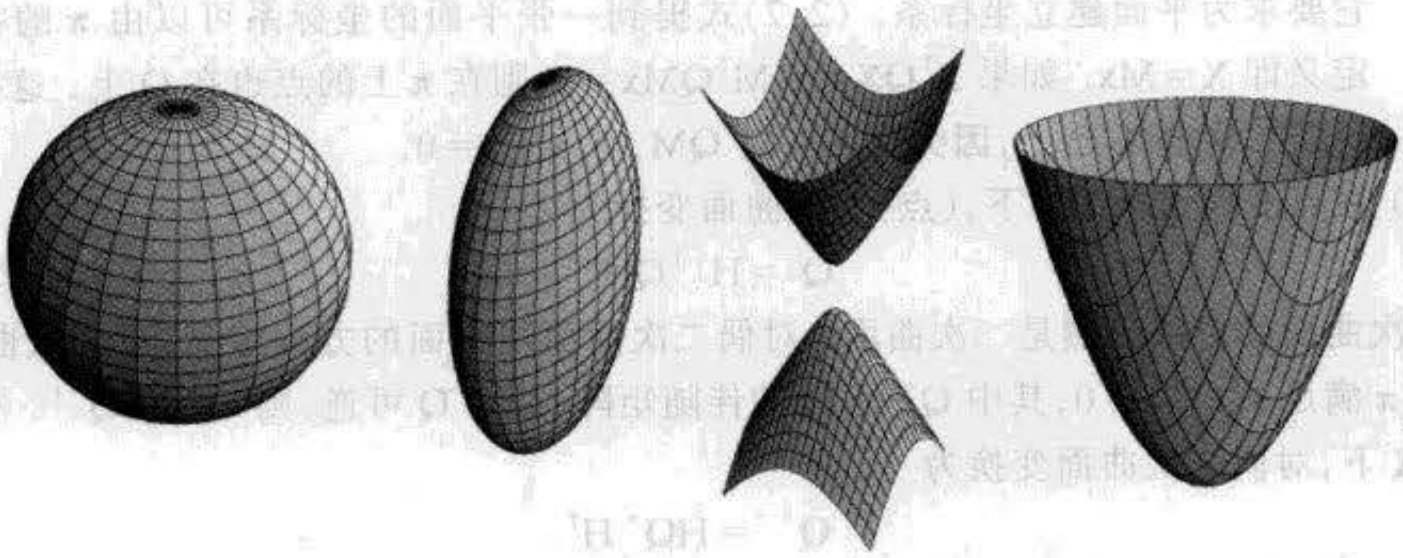


图 2.2 **非直纹二次曲面**. 这里给出了球面,椭球面,双叶双曲面和抛物面的图. 它们全是射影等价的.

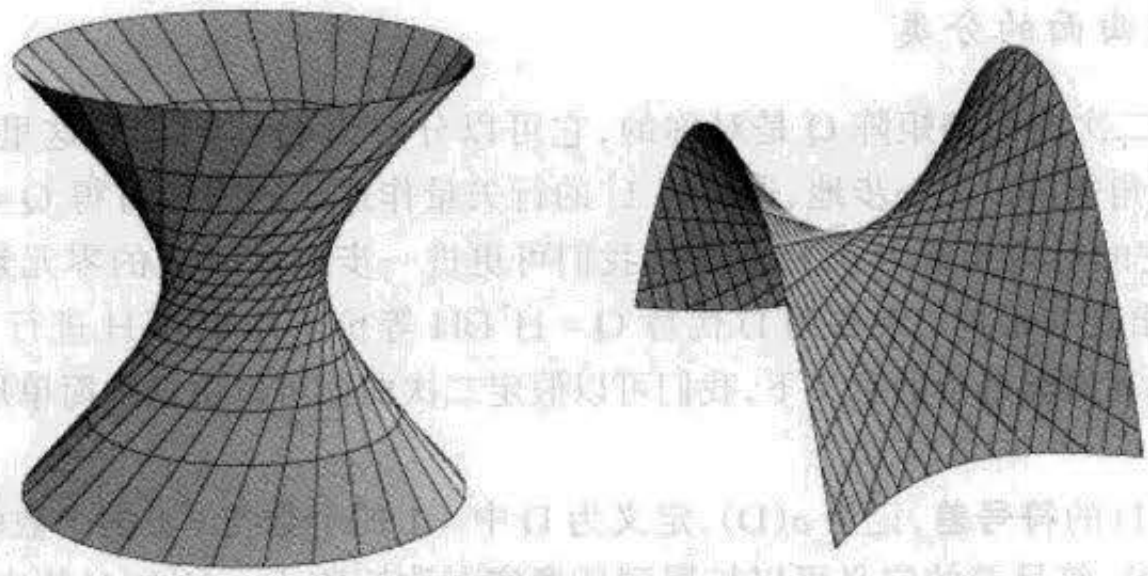


图 2.3 **直纹二次曲面**. 这里给出单叶双曲面的两个例子. 这些曲面分别由方程 $X^2+Y^2=Z^2+1$ 和 $XY=Z$ 给出,并且是射影等价的. 注意这两个曲面由两簇不相交的直线组成而且其中一簇的每根直线与另一簇的每根直线相交.

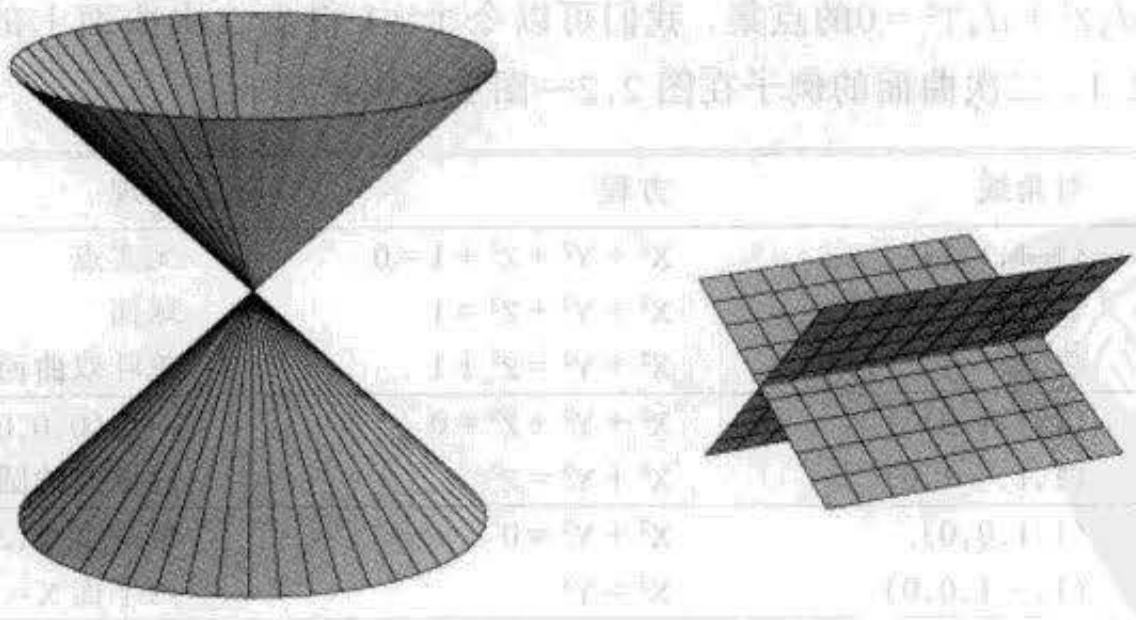


图 2.4 **退化二次曲面**. 这里给出两种最重要的退化的二次曲面:锥面和相交两平面. 这两种二次曲面都是直纹的. 表示此锥面的矩阵的秩为 3,其零矢量表示锥面的结点. 表示两(非重合)平面的矩阵秩为 2,其秩为 2 的零空间的两个生成矢量是平面交线上的两个点.

线簇. 关于直纹二次曲面的性质参看[Semple-79].

最有趣的二次曲面是秩为 4 的两种二次曲面. 注意这两种二次曲面甚至在它们的拓扑类型上也不同. 符号差为 2 的二次曲面(球面)(很显然)拓扑等价于一个球面. 另一方面,单叶双曲面却非拓扑等价(同伦)于一个球面. 事实上,它拓扑等价于环面(拓扑等价于 $S^1\times S^1$). 这显然说明二者不是射影等价的.

## 2.3 三次绕线

三次绕线可以看成 2D 二次曲线的 3 维类推(虽然从另一个角度说,二次曲面也是二次曲线的 3 维类推.)

2 维射影平面上的一条二次曲线可以由下列方程给出的一条参数曲线描述

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A\begin{pmatrix} 1 \\ \theta \\ \theta^2 \end{pmatrix} = \begin{pmatrix} a_{11}+a_{12}\theta+a_{13}\theta^2 \\ a_{21}+a_{22}\theta+a_{23}\theta^2 \\ a_{31}+a_{32}\theta+a_{33}\theta^2 \end{pmatrix} \tag{2.18}$$

其中 A 是非奇异 3×3 矩阵.

类似地,一条三次绕线定义为 $\mathbb{P}^3$中的一条曲线,它的参数形式如下

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} = A\begin{pmatrix} 1 \\ \theta \\ \theta^2 \\ \theta^3 \end{pmatrix} = \begin{pmatrix} a_{11}+a_{12}\theta+a_{13}\theta^2+a_{14}\theta^3 \\ a_{21}+a_{22}\theta+a_{23}\theta^2+a_{24}\theta^3 \\ a_{31}+a_{32}\theta+a_{33}\theta^2+a_{34}\theta^3 \\ a_{41}+a_{42}\theta+a_{43}\theta^2+a_{44}\theta^3 \end{pmatrix} \tag{2.19}$$

其中 A 是非奇异 4×4 矩阵.

三次绕线可能不为读者熟知,所以图 2.5 给出了该曲线的几种不同视图. 事实上,三次绕线是相当好(光滑)的空间曲线.

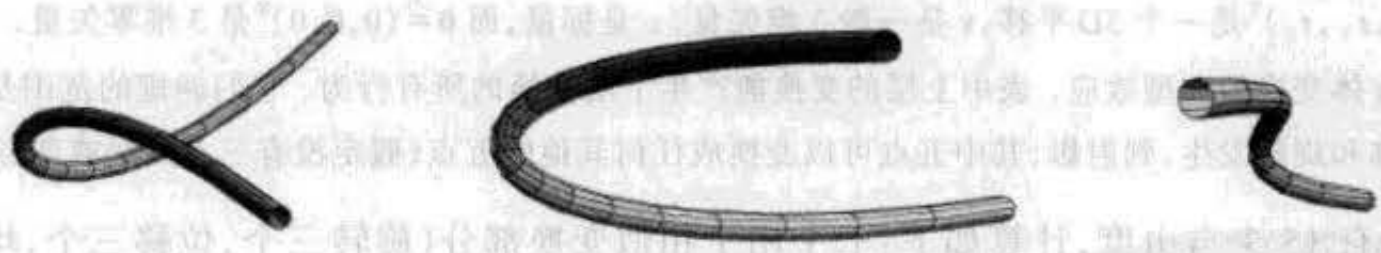


图 2.5 三次绕线$(t^3,t^2,t)^{\mathrm{T}}$的几种不同视图. 曲线加厚成管子以增强可视效果.

**三次绕线的性质** 令 **c** 为一条非奇异三次绕线. 那么 **c** 不整个地包含在 $\mathbb{P}^3$的任何一张平面中;而是与一般平面有三个不同的交点. 三次绕线有 12 个自由度(矩阵 A 有 15 个,减去 3 是因为参数 $\theta$ 的 1D 射影变换,它保持曲线不变). 要求该曲线过点 **X**,给 **c** 加了两个约束,因为 $\mathbf{X}=A(1,\theta,\theta^2,\theta^3)^{\mathrm{T}}$给出三个独立的比率,但一旦 $\theta$ 被消去仅有 2 个约束. 因此,过一般位置的六点有唯一的三次绕线 **c**. 最后,所有非退化的三次绕线都是射影等价的. 这一结论显然来自定义(2.19):射影变换 $A^{-1}$把 **c** 映为标准形式 $\mathbf{c}(\theta')=(1,\theta',\theta'^2,\theta'^3)^{\mathrm{T}}$,既然所有的三次绕线都可以映射到这条曲线,所以得出结论:所有的三次绕线都是射影等价的.

三次绕线的各种特殊情形,例如一条二次曲线和重合线的分类在[Semple-79]中给出.

在两视图几何中，三次绕线产生视在同视点(第 8 章)，而且在定义摄像机投影矩阵的退化集时担负重要角色(第 21 章).

## 2.4 变换的层次

3 维空间射影变换的几种特殊情况在本书中将经常出现. 这些特殊情况类似于节 1.4 ($p$11)中关于平面变换的层次. 每一种特殊情况是一种子群并由它的矩阵形式或不变量来刻划. 我们把它们概括在表 2.2 中. 此表仅给出 3 维空间变换比相应的 2 维空间变换**多出来**的性质——3 维空间变换同样具有相应的 2 维空间变换(罗列在表 1.1($p$16)中)的不变量.

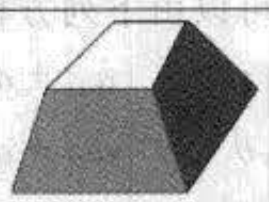
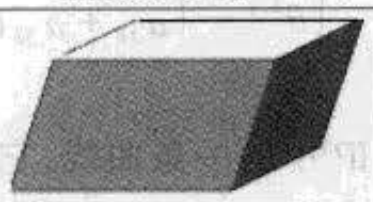
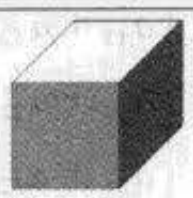
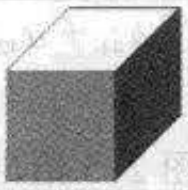
| 群 | 矩阵 | 失真 | 不变性质 |
|---|---|---|---|
| 射影<br>15dof | $\begin{bmatrix} A & \mathbf{t} \\ \mathbf{v}^T & v \end{bmatrix}$ |  | 接触表面的相交和相切，<br>高斯曲率的符号 |
| 仿射<br>12dof | $\begin{bmatrix} A & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ |  | 平面的平行性，<br>体积比，形心，<br>无穷远平面 $\pi_\infty$<br>(见节 2.5) |
| 相似<br>7dof | $\begin{bmatrix} sR & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ |  | 绝对二次曲线 $\Omega_\infty$<br>(见节 2.6) |
| 欧氏<br>6dof | $\begin{bmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ |  | 体积 |

表 2.2 **3 维空间中常发生的变换的几何不变性质**. 矩阵 A 是一个可逆的 3×3 矩阵，R 是一个 3D 旋转矩阵，$\mathbf{t}=(t_X,t_Y,t_Z)^T$ 是一个 3D 平移，$\mathbf{v}$ 是一般 3 维矢量，$v$ 是标量，而 $\mathbf{0}=(0,0,0)^T$ 是 3 维零矢量. 在失真列中给出立方体变换的典型效应. 表中上层的变换能产生下层变换的所有行为. 它们涵概的范围从欧氏，其中仅有平移和旋转发生，到射影，其中五点可以变换成任何其他的五点(假定没有三点共线或四点共面).

射影变换有 15 个自由度，计算如下：七个用于相似变换部分(旋转三个，位移三个，均匀缩放一个)，五个用于仿射变换部分，三个用于射影变换部分.

这些变换的两个重要表征是平行和角度. 例如，经仿射变换后原来的平行性保持不变，但夹角会改变，而射影变换后平行性会丢失.

下面简单介绍欧氏变换的一种分解，它将在今后讨论特殊运动中有用.

### 2.4.1 转动分解

平面欧氏变换可以看作是平移矢量 $\mathbf{t}$ 被限制在一张平面上而且转动轴垂直于该平面的一种 3 维空间欧氏变换的特殊情况. 但是，3 维空间上的欧氏运动更具一般性，因为一般情况下转动轴与平移不垂直. 转动分解能使任何欧氏运动(旋转加平移)化到几乎与 2D 的情形一样简单. 转动分解是

**结论 2.6　任何具体的平移加旋转运动都等价于绕一根转动轴的旋转加沿该转动轴的平移. 该转动轴平行于原来的旋转轴.**

平移加绕**正交**轴的旋转运动(称**平面运动**)等价于**仅仅**绕某转动轴的旋转.

**证明**　我们将勾画一个易于可视化的构造性的几何证明. 首先考虑 2D 的情形——平面上的欧氏变换. 从图 2.6 中可知对应于这种 2D 变换的转动轴显然存在. 对于 3D 的情形,将平移 $\mathbf{t}$ 分解成分别与转动轴平行和正交的两个部分 $\mathbf{t}=\mathbf{t}_{/\!/}+\mathbf{t}_{\perp}$, ($\mathbf{t}_{/\!/}=(\mathbf{t}.\mathbf{a})\mathbf{a}$, $\mathbf{t}_{\perp}=\mathbf{t}-(\mathbf{t}.\mathbf{a})\mathbf{a}$). 该欧氏运动也分解成两部分:第一部分是绕转动轴的旋转,它包含原旋转和 $\mathbf{t}_{\perp}$,第二部分则是沿转动轴的平移 $\mathbf{t}_{/\!/}$. 整个运动在图 2.7 中给予说明. □

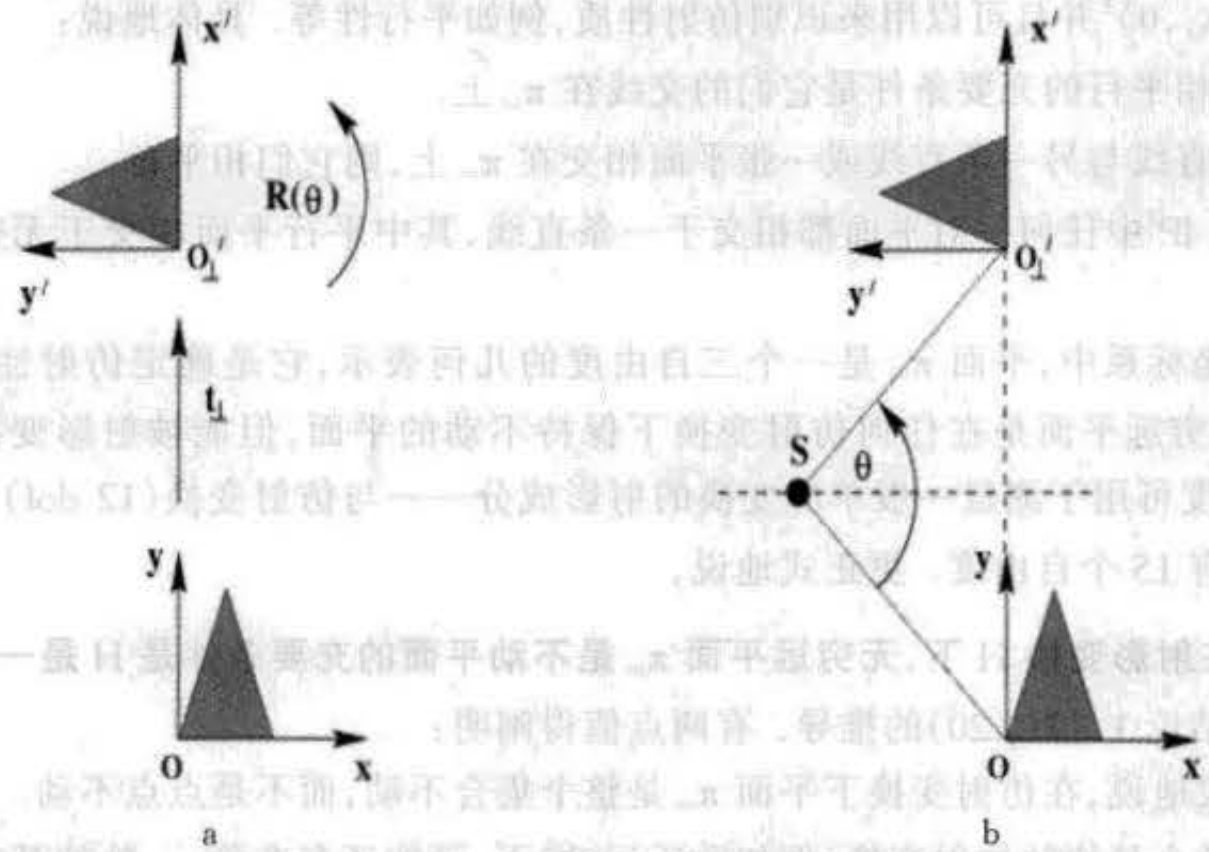



图 2.6　**2D 欧氏运动和"转动"轴.** (a)坐标系 $\{x,y\}$ 经历一次位移 $\mathbf{t}_{\perp}$ 和一次旋转 $\theta$ 后到达坐标系 $\{x',y'\}$. 运动在与转动轴垂直的平面中进行. (b)上述运动等价于绕转动轴 S 的一次旋转. 该转动轴位于对应点连线的垂直平分线上,使得连接 S 与二对应点间的直线之间的夹角为 $\theta$. 在图中对应点是两坐标系的原点而 $\theta$ 的值为 90°.

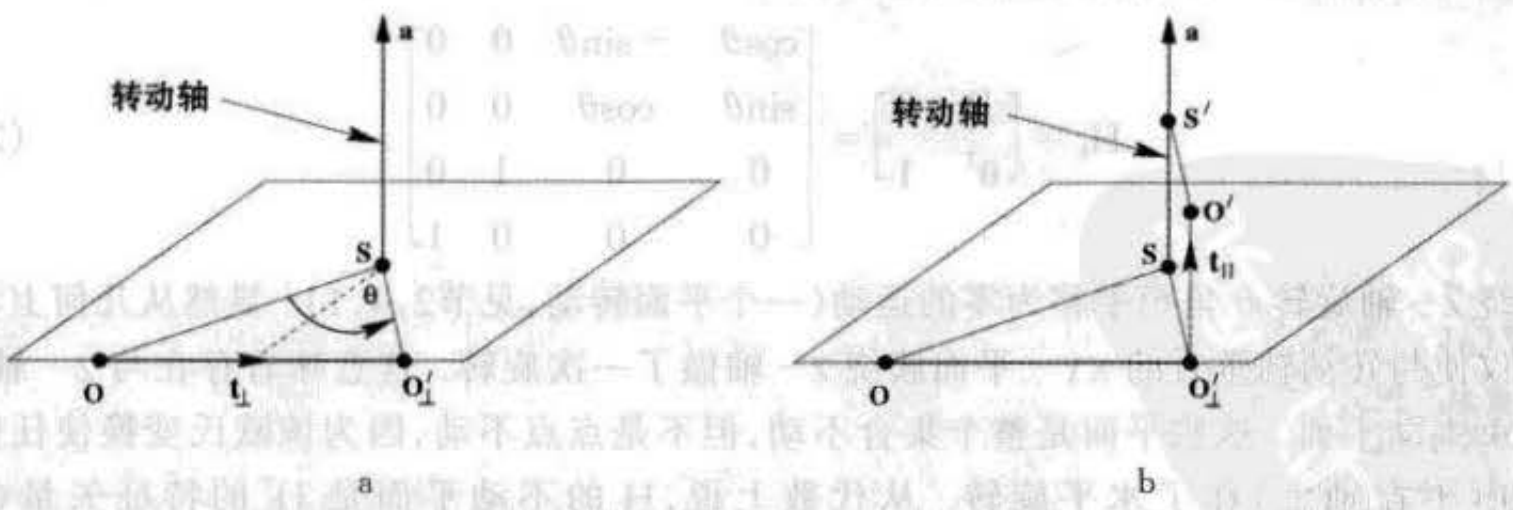



图 2.7　**3D 欧氏运动与转动分解.** 任何欧氏旋转 R 加平移 $\mathbf{t}$ 可以用如下方式得到(a)绕转动轴旋转,加(b)沿转动轴作平移 $\mathbf{t}_{/\!/}$. 其中 $\mathbf{a}$ 是转动轴的(单位)方向(使得 $R\mathbf{a}=\mathbf{a}$),而 $\mathbf{t}$ 被分解为分别与转动轴方向平行和垂直的两个矢量之和 $\mathbf{t}=\mathbf{t}_{/\!/}+\mathbf{t}_{\perp}$. S 是在转动轴上最靠近 O 的点(S 到 O 的连线与 $\mathbf{a}$ 的方向垂直). 同理,S′是在转动轴上离 O′最近的点.

转动分解可以由表示欧氏变换的 4×4 矩阵的不动点来确定. 这个想法将在本章末尾

的练习中验证.

## 2.5 无穷远平面

在平面射影几何中,辨认了无穷远线 $\mathbf{l}_\infty$ 就能测量平面的仿射性质. 辨认了 $\mathbf{l}_\infty$ 上的虚圆点就能测量其度量性质. 在 3 维空间的射影几何中,与 $\mathbf{l}_\infty$ 和虚圆点对应的几何实体是无穷远平面 $\boldsymbol{\pi}_\infty$ 和绝对二次曲线 $\Omega_\infty$.

在 3 维仿射空间中,无穷远平面的标准位置是 $\boldsymbol{\pi}_\infty = (0,0,0,1)^\mathsf{T}$. $\boldsymbol{\pi}_\infty$ 包含所有方向 $\mathbf{D} = (X_1, X_2, X_3, 0)^\mathsf{T}$ 并且可以用来识别仿射性质,例如平行性等. 具体地说:

- 两张平面相平行的充要条件是它们的交线在 $\boldsymbol{\pi}_\infty$ 上.
- 如果一条直线与另一条直线或一张平面相交在 $\boldsymbol{\pi}_\infty$ 上,则它们相平行.

因而,在 $\mathbb{P}^3$ 中任何一对平面都相交于一条直线,其中平行平面相交于无穷远平面上的直线.

在射影坐标系中,平面 $\boldsymbol{\pi}_\infty$ 是一个三自由度的几何表示,它是确定仿射性质所必要的. 通俗地说,无穷远平面是在任何仿射变换下保持不动的平面,但能被射影变换移动. 因此 $\boldsymbol{\pi}_\infty$ 的 3 自由度可用于测量一般单应变换的射影成分——与仿射变换(12 dof) 相比,这种一般变换总计有 15 个自由度. 更正式地说,

**结论 2.7 在射影变换 H 下,无穷远平面 $\boldsymbol{\pi}_\infty$ 是不动平面的充要条件是 H 是一个仿射变换.**

证明类似于结论 1.17($p$20)的推导. 有两点值得阐明:

(1) 一般地说,在仿射变换下平面 $\boldsymbol{\pi}_\infty$ 是整个集合不动,而不是点点不动.

(2) 在某个具体的仿射变换(例如欧氏运动)下,可能还存在除 $\boldsymbol{\pi}_\infty$ 外的某些平面保持不动. 但仅有 $\boldsymbol{\pi}_\infty$ 在任何仿射变换下保持不动.

下面的例子对这些要点作更详细地说明.

**例 2.8** 考察表示一个欧氏变换的下述矩阵

$$\mathrm{H}_\mathrm{E} = \begin{bmatrix} \mathrm{R} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.20}$$

这是绕 Z-轴旋转 $\theta$ 角而平移为零的运动(一个平面转动,见节2.4.1). 显然从几何上看,该变换仅使与转动轴垂直的 XY-平面簇绕 Z-轴做了一次旋转. 这意味着存在与 Z-轴垂直的一束不动平面. 这些平面是整个集合不动,但不是点点不动,因为该欧氏变换使任何(有限)点(不在轴上)作了水平旋转. 从代数上说,H 的不动平面是 $\mathrm{H}^\mathsf{T}$ 的特征矢量(参看节 1.9). 在此例中,$\mathrm{H}_\mathrm{E}^\mathsf{T}$ 的特征值是 $\{e^{i\theta}, e^{-i\theta}, 1, 1\}$,对应特征矢量是

$$\mathbf{E}_1 = \begin{pmatrix} 1 \\ i \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{E}_2 = \begin{pmatrix} 1 \\ -i \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{E}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{E}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

特征矢量 $\mathbf{E}_1$ 和 $\mathbf{E}_2$ 对应的不是实平面，故不在此作进一步的讨论. 特征矢量 $\mathbf{E}_3$ 和 $\mathbf{E}_4$ 都是退化的. 因此，存在由这些特征矢量生成的不动平面束. 该平面束包含 $\boldsymbol{\pi}_\infty$，其轴是(垂直于 Z－轴的)平面与 $\boldsymbol{\pi}_\infty$ 的交线. △

这个例子也说明射影平面 $\mathbb{P}^2$和 3 维射影空间 $\mathbb{P}^3$之间的几何联系. 平面 $\boldsymbol{\pi}$ 与 $\boldsymbol{\pi}_\infty$ 交于一条直线，它是平面 $\boldsymbol{\pi}$ 的无穷远线 $\mathbf{l}_\infty$. $\mathbb{P}^3$中的射影变换在平面 $\boldsymbol{\pi}$ 上诱导出一个**从属的**平面射影变换.

**重构的仿射性质** 在今后论述重构的各章，例如第 9 章中，我们将会看到(欧氏)景物的射影坐标可以由多视图来恢复. 在 3 维射影空间中一旦 $\boldsymbol{\pi}_\infty$ 被辨认，即已知它的射影坐标，那么就有可能确定重构的仿射性质，例如几何实体的平行性——如果它们在 $\boldsymbol{\pi}_\infty$ 上相交，那么它们平行.

从算法的角度，一种更适宜的方法是对 $\mathbb{P}^3$进行变换使已辨认的 $\boldsymbol{\pi}_\infty$ 移到它的标准位置 $\boldsymbol{\pi}_\infty=(0,0,0,1)^\mathsf{T}$. 经这样的映射所得到的是欧氏景物，其中 $\boldsymbol{\pi}_\infty$ 的坐标是$(0,0,0,1)^\mathsf{T}$，而重构与使 $\boldsymbol{\pi}_\infty$ 固定到$(0,0,0,1)^\mathsf{T}$的射影变换相关. 根据结论 2.7，可以知道景物与重构由一个仿射变换相关联. 因此仿射性质可以直接从实体的坐标中测量出来.

## 2.6 绝对二次曲线

绝对二次曲线 $\Omega_\infty$ 是在 $\boldsymbol{\pi}_\infty$ 上的一条(点)二次曲线. 在度量坐标系中 $\boldsymbol{\pi}_\infty=(0,0,0,1)^\mathsf{T}$，而在 $\Omega_\infty$ 上的点满足

$$\left.\begin{aligned} X_1^2+X_2^2+X_3^2 \\ X_4^2 \end{aligned}\right\}=0. \tag{2.21}$$

注意为了定义 $\Omega_\infty$，需要两个方程.

为确定在 $\boldsymbol{\pi}_\infty$ 上(即具有 $X_4=0$ 的)点的方向，定义 $\Omega_\infty$ 的方程可以写成

$$(X_1,X_2,X_3)\mathrm{I}(X_1,X_2,X_3)^\mathsf{T}=0$$

因而，$\Omega_\infty$ 是对应于矩阵 $C=I$ 的一条二次曲线 C. 可见它是 $\boldsymbol{\pi}_\infty$ 上由纯虚点组成的一条二次曲线.

二次曲线 $\Omega_\infty$ 的几何表示需 5 个额外自由度，这 5 个自由度是仿射坐标系中确定度量性质所需要的. $\Omega_\infty$ 的一个主要性质是它是任何相似变换下不动的二次曲线. 更正式地说，

**结论 2.9 在射影变换 H 下，绝对二次曲线 $\Omega_\infty$ 是不动二次曲线的充要条件是 H 是相似变换.**

**证明** 因为绝对二次曲线在无穷远平面上，使它不动的变换必须使无穷远平面不动，从而必须是仿射的. 这样一种变换的形式为

$$H_A=\begin{bmatrix} A & \mathbf{t} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix}.$$

限制在无穷远平面上的绝对二次曲线由矩阵 $I_{3\times3}$ 表示，既然它在 $H_A$ 作用下是不动的，那么我们可以得到 $A^{-\mathsf{T}}IA^{-1}=I$(相差一个尺度因子)，对它取逆得 $AA^\mathsf{T}=I$. 这说明 A 是正交矩阵，因而是一个带有缩放的旋转，或者是一个带有缩放的旋转加反射. 证毕. □

虽然 $\Omega_\infty$ 没有任何实点，但它仍具有任何二次曲线的性质——例如一条直线与一条二次曲线相交于两点；极点－极线的关系等等．下面给出 $\Omega_\infty$ 的几个具体性质：

(1) $\Omega_\infty$ 在一般相似变换下是集合不动，而不是点点不动的．这表明在相似变换下，$\Omega_\infty$ 上的一点可能被移动到 $\Omega_\infty$ 上的另一点，但不会被映射出该二次曲线．

(2) 所有的圆交 $\Omega_\infty$ 于两点．假定圆的支撑平面是 $\boldsymbol{\pi}$．那么 $\boldsymbol{\pi}$ 交 $\boldsymbol{\pi}_\infty$ 于一条直线，而该直线交 $\Omega_\infty$ 于两点．这两点是 $\boldsymbol{\pi}$ 的虚圆点．

(3) 所有球面交 $\boldsymbol{\pi}_\infty$ 于 $\Omega_\infty$．

**度量性质** 一旦 $\Omega_\infty$（和它的支撑平面 $\boldsymbol{\pi}_\infty$）在 3 维射影空间被辨认，那么诸如夹角和相对长度等度量性质可以被测定．

设两条直线的方向为 $\mathbf{d}_1$ 和 $\mathbf{d}_2$（3 维矢量）．在欧氏世界系中这些方向之间的夹角为

$$\cos\theta = \frac{(\mathbf{d}_1^{\mathrm{T}}\mathbf{d}_2)}{\sqrt{(\mathbf{d}_1^{\mathrm{T}}\mathbf{d}_1)(\mathbf{d}_2^{\mathrm{T}}\mathbf{d}_2)}}. \tag{2.22}$$

它可以写成

$$\cos\theta = \frac{(\mathbf{d}_1^{\mathrm{T}}\Omega_\infty\mathbf{d}_2)}{\sqrt{(\mathbf{d}_1^{\mathrm{T}}\Omega_\infty\mathbf{d}_1)(\mathbf{d}_2^{\mathrm{T}}\Omega_\infty\mathbf{d}_2)}} \tag{2.23}$$

其中 $\mathbf{d}_1$ 和 $\mathbf{d}_2$ 是直线与包含二次曲线 $\Omega_\infty$ 的平面 $\boldsymbol{\pi}_\infty$ 的交点，而 $\Omega_\infty$ 是该平面上绝对二次曲线的矩阵表示．表示式(2.23)在欧氏世界系中化简为(2.22)，其中 $\Omega_\infty = \mathbf{I}$．但该表示式在任何射影坐标系中都有效，这可以从点和二次曲线的变换性质中得到证明（见 $p32(4)(b)$）．

现在还不存在由平面的表面法线的方向来计算平面间夹角的简单公式．

**正交与配极** 基于绝对二次曲线，我们给出射影空间中正交性的几何表示．主要工具是由二次曲线诱导的点与线之间的极点—极线关系．

由(2.23)直接推出：如果 $\mathbf{d}_1^{\mathrm{T}}\Omega_\infty\mathbf{d}_2 = 0$，则 $\mathbf{d}_1$ 和 $\mathbf{d}_2$ 相垂直．因而垂直性可由关于 $\Omega_\infty$ 的**共轭性**来表征．这样做的巨大好处是共轭性是射影关系，因此在射影坐标系（由 3 维欧氏空间的射影变换得到）下，如果两方向关于 $\Omega_\infty$ 共轭，那么它们被认为相垂直（$\Omega_\infty$ 的矩阵在射影坐标系下一般不是 $\mathbf{I}$）．正交性的几何表示在图 2.8 中给出．

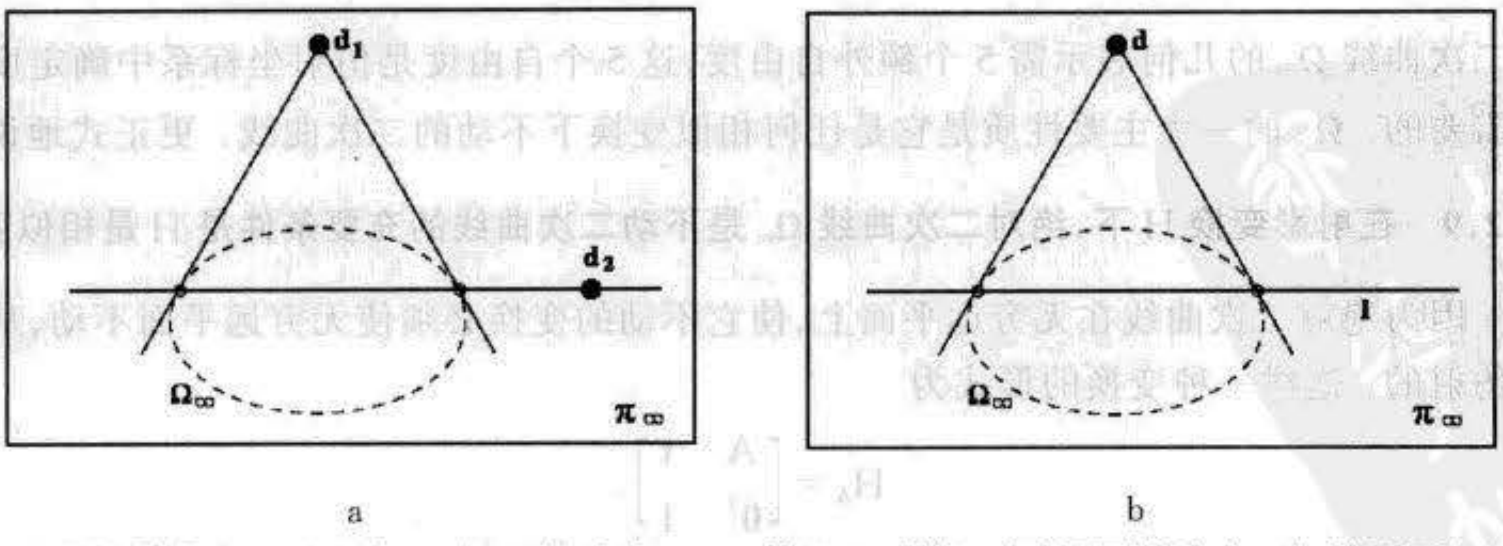



图 2.8 **正交性**与 $\Omega_\infty$．(a)在 $\boldsymbol{\pi}_\infty$ 上，正交方向 $\mathbf{d}_1, \mathbf{d}_2$ 关于 $\Omega_\infty$ 共轭．(b)平面的法向 $\mathbf{d}$ 和该平面与 $\boldsymbol{\pi}_\infty$ 的交线 $\mathbf{l}$ 是关于 $\Omega_\infty$ 的极点——极线关系．

这种表示有助于研究摄像机中射线之间的正交性，例如有助于确定过摄像机中心的平

面的法线(见节 7.6($p$149)). 如果图像点关于 $\Omega_\infty$ 的**图像**共轭,那么对应的射线相垂直.

同样,从算法的角度,一种更适宜的方法是对坐标进行射影变换,使 $\Omega_\infty$ 映射到它的标准位置(2.21),然后度量性质可以直接由坐标确定.

## 2.7 绝对对偶二次曲面

我们已经知道 $\Omega_\infty$ 由**两个**方程来定义——它是在无穷远平面上的一条二次曲线. 绝对二次曲线 $\Omega_\infty$ 的对偶是 3 维空间中一种退化的对偶**二次曲面**,称为**绝对对偶二次曲面**并记为 $Q_\infty^*$. 从几何上说,$Q_\infty^*$ 由 $\Omega_\infty$ 的切平面组成,因而 $\Omega_\infty$ 是 $Q_\infty^*$ 的"边缘". 它被称为**边二次曲面**. 想像一个椭球面的所有切平面的集合,然后把椭球面压成平饼的情况.

从代数上说,$Q_\infty^*$ 由秩 3 的 $4\times4$ 的齐次矩阵表示,它在 3 维度量空间的标准形式是

$$Q_\infty^* = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}. \tag{2.24}$$

我们将证明在对偶绝对二次曲面包络中的任何平面都与 $\Omega_\infty$ 相切,因而 $Q_\infty^*$ 真正是 $\Omega_\infty$ 的对偶. 考察由 $\boldsymbol{\pi} = (\mathbf{v}^T, k)^T$ 表示的平面. 该平面在 $Q_\infty^*$ 定义的包络上的充要条件是 $\boldsymbol{\pi}^T Q_\infty^* \boldsymbol{\pi} = 0$,根据(2.24) 的形式,等价于 $\mathbf{v}^T\mathbf{v} = 0$. 其中(见节 7.6($p$149)),$\mathbf{v}$ 表示平面 $(\mathbf{v}^T, k)^T$ 与无穷远平面的交线. 该直线与绝对二次曲面相切的充要条件是 $\mathbf{v}^T I \mathbf{v} = 0$. 因此,$Q_\infty^*$ 的包络就是由这些与绝对二次曲线相切的平面组成.

因为它很重要,我们再从另一角度来考虑它. 考虑绝对二次曲线是一系列压平了的椭球的极限,即二次曲面由矩阵 $Q = \mathrm{diag}(1,1,1,k)$ 表示. 当 $k \mapsto \infty$ 时,这些二次曲面越来越靠近无穷远平面,取极限时,它们仅包含点 $(X_1, X_2, X_3, 0)^T$,其中 $X_1^2 + X_2^2 + X_3^2 = 0$,这正是绝对二次曲线上的点. 但 Q 的对偶是二次曲面 $Q^* = Q^{-1} = \mathrm{diag}(1,1,1,k^{-1})$, 它在取极限时变成绝对二次曲线的对偶 $Q_\infty^* = \mathrm{diag}(1,1,1,0)$.

对偶二次曲面 $Q_\infty^*$ 是退化的二次曲面,有 8 个自由度(一个对称矩阵有 10 个独立元素,但与尺度无关以及行列式为零的条件各减去 1 个自由度). 这 8 自由度的几何表示是射影坐标系下确定度量性质所需的.

在代数操作上 $Q_\infty^*$ 比 $\Omega_\infty$ 有显著的优越性,因为 $\boldsymbol{\pi}_\infty$ 和 $Q_\infty^*$ 包含在单个几何对像中(而不像 $\Omega_\infty$ 需要两个方程(2.21) 来确定它)下面我们给出它的三个最重要的性质.

**结论 2.10 在射影变换 H 下,绝对二次曲面 $Q_\infty^*$ 不动的充要条件是 H 是相似变换.**

**证明** 该结论直接来自绝对二次曲线在相似变换下的不变性,以及 $Q_\infty^*$ 和 $\Omega_\infty$ 之间的平面相切关系是变换不变量. 尽管如此,我们仍给出一种与此独立的直接的证明.

因为 $Q_\infty$ 是一个对偶二次曲面,它的变换遵循(2.17 - $p$41),因而它在 H 下不动的充要条件是 $Q_\infty = H Q_\infty H^T$. 用形如

$$H = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{v}^T & k \end{bmatrix}$$

的任意一个变换代入,我们发现

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & 0 \end{bmatrix} = \begin{bmatrix} \mathrm{A} & \mathbf{t} \\ \mathbf{v}^{\mathsf{T}} & k \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{A}^{\mathsf{T}} & \mathbf{v} \\ \mathbf{t}^{\mathsf{T}} & k \end{bmatrix} = \begin{bmatrix} \mathrm{AA}^{\mathsf{T}} & \mathrm{A}\mathbf{v} \\ \mathbf{v}^{\mathsf{T}}\mathrm{A}^{\mathsf{T}} & \mathbf{v}^{\mathsf{T}}\mathbf{v} \end{bmatrix}$$

上式必须在相差一个尺度的情况下为真. 通过分析,此等式成立的充要条件是 $\mathbf{v}=\mathbf{0}$ 和 A 是带有缩放的正交矩阵(即缩放,旋转可能加上反射). 换句话说,H 是一个相似变换. □

**结论 2.11 无穷远平面 $\pi_\infty$ 是 $Q_\infty^*$ 的零矢量.**

这很容易验证,当 $Q_\infty^*$ 在度量坐标系下取其标准形式(2.24)而 $\pi_\infty=(0,0,0,1)^{\mathsf{T}}$ 时,$Q_\infty^*\pi_\infty=\mathbf{0}$. 该性质在任何坐标系下都成立,这一点可以用代数方法由平面和对偶二次曲面的变换性质推出:如果 $\mathbf{X}'=\mathrm{H}\mathbf{X}$,那么 $Q_\infty^{*\prime}=\mathrm{H}\,Q_\infty^*\mathrm{H}^{\mathsf{T}}$,$\pi_\infty'=\mathrm{H}^{-\mathsf{T}}\pi_\infty$,以及

$$Q_\infty^{*\prime}\pi'_\infty=(\mathrm{H}\,Q_\infty^*\mathrm{H}^{\mathsf{T}})\mathrm{H}^{-\mathsf{T}}\pi_\infty=\mathrm{H}\,Q_\infty^*\pi_\infty=\mathbf{0}.$$

**结论 2.12 两张平面 $\pi_1$ 和 $\pi_2$ 之间的夹角由下式给出:**

$$\cos\theta=\frac{\pi_1^{\mathsf{T}}Q_\infty^*\pi_2}{\sqrt{(\pi_1^{\mathsf{T}}Q_\infty^*\pi_1)(\pi_2^{\mathsf{T}}Q_\infty^*\pi_2)}}. \tag{2.25}$$

**证明** 设两平面的欧氏坐标为 $\pi_1=(\mathbf{n}_1^{\mathsf{T}},d_1)^{\mathsf{T}}$, $\pi_2=(\mathbf{n}_2^{\mathsf{T}},d_2)^{\mathsf{T}}$. 在欧氏坐标系下,$Q_\infty^*$ 形如(2.24),从而(2.25)化为

$$\cos\theta=\frac{\mathbf{n}_1^{\mathsf{T}}\mathbf{n}_2}{\sqrt{(\mathbf{n}_1^{\mathsf{T}}\mathbf{n}_1)(\mathbf{n}_2^{\mathsf{T}}\mathbf{n}_2)}}$$

其中平面间的夹角用它们法线的标量积来表示.

如果对平面和 $Q_\infty^*$ 进行射影变换,根据平面和对偶二次曲面的(协变)变换性质,(2.25)式将仍然能用来决定平面之间的夹角. □

证明的最后部分的详细推导将留作练习,它不过是结论 1.22($p$25)的推导在 3D 中的直接类推,在结论 1.22($p$25)的推导中,用虚圆点的对偶来计算 $\mathbb{P}^2$ 中两线夹角. 这里,$\mathbb{P}^3$ 中的平面是 $\mathbb{P}^2$ 中直线的类推,而绝对对偶二次曲面是虚圆点对偶的类推.

## 2.8 结束语

### 2.8.1 文献

第一章列举的教科书也与本章有关. 关于画法几何中透视的一般背景知识参看[Boelm-94],关于曲线和曲面许多性质的清楚解释参看 Hilbert 和 Cohn-Vossen [Hilbert-56].

$\mathbb{P}^3$ 中点、线、平面的一种重要表示:Grassmann-Cayley 代数在本章没有介绍. 在这种表示中,如连接和关联那样的几何运算被表示成基于矩阵行列式的"括号代数". 文献[Carlsson-94]对有关领域作了很好的介绍.

Faugeras 和 Maybank[Faugeras-90]把 $\Omega_\infty$ 引入计算机视觉的文献(为了确定相对定向解的多重性),而 Triggs[Triggs-97]则引入 $Q_\infty^*$ 用于自标定.

### 2.8.2 注释和练习

(1) **Plücker 坐标**

(a) 利用 Plücker 直线坐标 $\mathcal{L}$,写出关于直线与平面交点的表示式以及由一点与一直线定义的平面的表示式.

(b) 推导点在直线上,和直线在平面上的条件.

(c) 证明平行平面相交在 $\boldsymbol{\pi}_\infty$ 的一条直线上. 提示:从(2.9 - $p$39)开始去确定两平行平面的交线 $L^*$.

(d) 证明平行线交在 $\boldsymbol{\pi}_\infty$ 上.

(2) **射影变换** 证明 3 维空间的一个(实)射影变换能把一个椭球面映射到一个抛物面或双叶双曲面,但不能把一个椭球面映射到单叶双曲面(即实的直纹面).

(3) **转动分解** 证明表示欧氏变换 $\{\mathbf{R},\mathbf{t}\}$(其转动轴方向为 $\mathbf{a}$,即 $\mathrm{R}\mathbf{a}=\mathbf{a}$)的 $4\times4$ 矩阵有两个复共轭特征值和两个相等的实特征值,并有下列特征矢量结构:

(a) 如果 $\mathbf{a}$ 垂直于 $\mathbf{t}$,那么对应于实特征值的两个特征矢量不同.

(b) 否则对应于实特征值的两个特征矢量重合并在 $\boldsymbol{\pi}_\infty$ 上.

(例如选择如(2.20)那样的简单例子,另一种情况 $p$347 中给出).

在第一种情形中,对应于实特征值的两个实点(特征矢量)定义一条不动点的直线. 它是平面运动的转动轴. 在第二种情形下,也定义了转动轴的方向,但它不是不动点的直线. 试问:对应于复特征值的特征矢量表示什么?

# 第 3 章　估计——2D 射影变换

本章将讨论估计问题. 本书中,估计的含义是指在某些本质测量的基础上计算某个变换或其它数学量. 这个定义有些含糊,因此,这里我们具体给出打算考虑的若干估计问题的类型.

(1) **2D 单应**　给定 $\mathbb{P}^2$中的点集 $\mathbf{x}_i$ 和同在 $\mathbb{P}^2$中的对应点集 $\mathbf{x}'_i$,计算把每一点 $\mathbf{x}_i$ 映射到对应点$\mathbf{x}'_i$的射影变换. 在实际的情形中,点$\mathbf{x}_i$和$\mathbf{x}'_i$是两幅图像(或同一幅图像)上的点,每幅图像都视为一张射影平面 $\mathbb{P}^2$.

(2) **3D 到 2D 的摄像机投影**　给定 3D 空间的点集 $\mathbf{X}_i$ 以及一幅图像上的对应点集 $\mathbf{x}_i$,求把 $\mathbf{X}_i$ 映射到 $\mathbf{x}_i$ 的 3D 到 2D 的射影映射. 这种 3D 到 2D 的投影是由射影摄像机来实现的映射,这些将在第 5 章中讨论.

(3) **基本矩阵的计算**　给定一幅图像上的点集 $\mathbf{x}_i$ 和另一幅图像上的对应点集 $\mathbf{x}'_i$,计算与这些对应一致的基本矩阵 F. 基本矩阵将在第 8 章中讨论,它是一个对所有的 $i$ 都使$\mathbf{x}'^{\mathrm{T}}_i\mathrm{F}\mathbf{x}_i=0$成立的 $3\times3$ 奇异矩阵 F.

(4) **三焦点张量计算**　给定跨三幅图像的点对应 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i\leftrightarrow\mathbf{x}''_i$ 的集合,计算三焦点张量. 三焦点张量(将在第 14 章讨论)是与三幅视图中点或线相关的一种张量 $\mathcal{T}_i^{jk}$.

这些问题有许多共同特性,其中一个问题的研究都与其余的每个问题有关. 因此,本章将对其中的第一个问题进行详细讨论. 在解决这个问题时所学到的方法将指导我们去解决其他的问题.

2D 射影变换的估计问题除有示例的作用之外,它本身也是重要的. 我们将研究两幅图像之间对应点的集合 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i$. 我们的问题是如何计算对所有的 $i$ 都满足 $\mathrm{H}\mathbf{x}_i=\mathbf{x}'_i$ 的一个 $3\times3$矩阵 H.

**测量数**　第一个要讨论的问题是:计算射影变换 H 需要多少对应点 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i$. 在考虑了自由度的个数和约束的个数后可以给出一个下界. 一方面,矩阵 H 有 9 个元素,但仅确定到相差一个尺度因子. 因此,2D 射影变换的自由度的总数是 8. 另一方面,每一组点到点的对应提供两个约束,因为对第一幅图像上的每个点 $\mathbf{x}_i$,第二幅图像上点的两个自由度必须对应于被映射点 $\mathrm{H}\mathbf{x}_i$. 一个 2D 点有两个自由度,即 $x$ 和 $y$ 分量,二者都可以分别指定. 换一种说法,点用一个齐次 3 维矢量表示,因为尺度因子是任意的,所以它仍只有两个自由度. 总之,为了完全约束 H,需要指定四组点对应.

**近似解**　我们将会看到如果只给定四组对应,那么可以得到矩阵 H 的精确解. 这种解称为**最小配置解**. 这种解是重要的,因为它们定义了鲁棒估计算法(如节 3.7($p$73)中介绍的 RANSAC)中所需子集的大小. 但是,因为点的测量是不精确的("噪声"),如果给定多于四组的对应,那么这些对应可能不与任何射影变换完全相容,因而我们面临的任务是按给定的数据确定"最好"的变换. 通常这个任务是通过寻找最小化某个代价函数的变换 H 来完成.

本章将讨论不同的代价函数以及它们的最小化方法. 有两类主要的代价函数:基于最小化代数误差的代价函数;和基于最小化几何的或统计的图像距离的代价函数. 这两类代价函数将在节 3.2 中介绍.

**黄金标准算法** 通常存在一种最优的代价函数,其最优的含义是在一定假设下,使代价函数取最小值的 H 是变换的最好估计. 计算该代价函数最小值的算法称为"黄金标准"算法. 其他算法的结果的优劣依它们与黄金标准算法的比较来判定. 在估计两视图之间的单应时,代价函数是(3.8),关于最优性的假设在节 3.3 中给出,而黄金标准就是算法 3.3($p72$).

## 3.1 直接线性变换(DLT)算法

我们首先讨论由给定 2D 到 2D 的四组点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 确定 H 的一种简单的线性算法. 变换由方程 $\mathbf{x}'_i = \mathrm{H}\mathbf{x}_i$ 给出. 注意这是一个齐次矢量方程;因此 3 维矢量 $\mathbf{x}'_i$ 和 $\mathrm{H}\mathbf{x}_i$ 不相等,它们有相同的方向,但在大小上可能相差一个非零因子. 该等式可以用矢量叉乘:$\mathbf{x}'_i \times \mathrm{H}\mathbf{x}_i = \mathbf{0}$表示. 由该表示式可推出 H 的一个简单的线性解.

如果将矩阵 H 的第 $j$ 行记为 $\mathbf{h}^{j\mathrm{T}}$,那么

$$\mathrm{H}\mathbf{x}_i = \begin{pmatrix} \mathbf{h}^{1\mathrm{T}}\mathbf{x}_i \\ \mathbf{h}^{2\mathrm{T}}\mathbf{x}_i \\ \mathbf{h}^{3\mathrm{T}}\mathbf{x}_i \end{pmatrix}.$$

记 $\mathbf{x}'_i = (x'_i, y'_i, w'_i)^{\mathrm{T}}$,则叉积可以显式地写成:

$$\mathbf{x}'_i \times \mathrm{H}\mathbf{x}_i = \begin{pmatrix} y'_i\mathbf{h}^{3\mathrm{T}}\mathbf{x}_i - w'_i\mathbf{h}^{2\mathrm{T}}\mathbf{x}_i \\ w'_i\mathbf{h}^{1\mathrm{T}}\mathbf{x}_i - x'_i\mathbf{h}^{3\mathrm{T}}\mathbf{x}_i \\ x'_i\mathbf{h}^{2\mathrm{T}}\mathbf{x}_i - y'_i\mathbf{h}^{1\mathrm{T}}\mathbf{x}_i \end{pmatrix}.$$

因为对 $j=1,2,3$,$\mathbf{h}^{j\mathrm{T}}\mathbf{x}_i = \mathbf{x}_i^{\mathrm{T}}\mathbf{h}^j$ 皆成立,这就给出关于 H 元素的三个方程,并可以写成下列形式

$$\begin{bmatrix} \mathbf{0}^{\mathrm{T}} & -w'_i\mathbf{x}_i^{\mathrm{T}} & y'_i\mathbf{x}_i^{\mathrm{T}} \\ w'_i\mathbf{x}_i^{\mathrm{T}} & \mathbf{0}^{\mathrm{T}} & -x'_i\mathbf{x}_i^{\mathrm{T}} \\ -y'_i\mathbf{x}_i^{\mathrm{T}} & x'_i\mathbf{x}_i^{\mathrm{T}} & \mathbf{0}^{\mathrm{T}} \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0}. \tag{3.1}$$

这些方程都有 $\mathrm{A}_i\mathbf{h} = \mathbf{0}$ 的形式,其中 $\mathrm{A}_i$ 是 $3\times 9$ 的矩阵,$\mathbf{h}$ 是由矩阵 H 的元素组成的 9 维矢量,

$$\mathbf{h} = \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix}, \quad \mathrm{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \tag{3.2}$$

其中 $h_i$ 是 $\mathbf{h}$ 的第 $i$ 个元素. 下面给出关于这些方程的三个注释.

(1) $\mathrm{A}_i\mathbf{h} = \mathbf{0}$ 是未知矢量 $\mathbf{h}$ 的**线性**方程. 矩阵 $\mathrm{A}_i$ 的元素是已知点的坐标的二次多项式.

(2) 虽然在(3.1)中有三个方程,但仅有两个是线性独立的(因为第三行在相差一个比值的意义下由 $x'_i$ 乘第一行加 $y'_i$ 乘第二行而得到). 因此每组点对应给出关于 H

元素的两个方程. 在解 H 时常省去第三个等式([Sutherland-63]). 从而(为了将来参考)方程组变为

$$\begin{bmatrix} \mathbf{0}^{\mathsf{T}} & -w_i'\mathbf{x}_i^{\mathsf{T}} & y_i'\mathbf{x}_i^{\mathsf{T}} \\ w_i'\mathbf{x}_i^{\mathsf{T}} & \mathbf{0}^{\mathsf{T}} & -x_i'\mathbf{x}_i^{\mathsf{T}} \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0}. \tag{3.3}$$

它可写成

$$\mathrm{A}_i\mathbf{h} = \mathbf{0}$$

其中 $\mathrm{A}_i$ 是(3.3)的 $2\times 9$ 矩阵. 但是,其中若有一个理想点 $\mathbf{x}_i'$,则 $w_i'=0$,于是(3.3)的两个方程将退化为单个方程. 但是,方程组(3.1)仍然有两个线性独立的方程,因而在这种情况下,(3.1)的第三个方程一定不能省略(可以省略前两个中的一个). 把所有三个方程都包含进去的方程组(3.1)对于一般点对应集可给出较良态的方程组.

(3) 该方程组对 $\mathbf{x}_i'$ 的任何齐次坐标$(x_i', y_i', w_i')^{\mathsf{T}}$成立. 我们可以取 $w_i'=1$,此时$(x_i', y_i')^{\mathsf{T}}$是图像中实际测量得到的坐标. 但是,今后会看到其他的选择也行.

**求解 H**

每组点对应给出关于 H 元素的两个独立的方程. 给定四组这样的点对应,便获得方程组 $\mathrm{A}\mathbf{h}=\mathbf{0}$,其中 A 是由每组点对应产生的矩阵行 $\mathrm{A}_i$ 构成的方程组的系数矩阵,$\mathbf{h}$ 是 H 未知元素的矢量. 我们只求 $\mathbf{h}$ 的非零解,因为我们对平凡解 $\mathbf{h}=\mathbf{0}$ 毫无兴趣. 如果采用(3.1)式,A 的维数是 $12\times 9$,如果采用(3.3)式,则维数是 $8\times 9$. 但在两种情况下 A 的秩都是 8,因而有 1 维零空间,从而存在只相差一个非零尺度因子意义下确定的解 $\mathbf{h}$. 但是,变换矩阵 H 一般也仅能确定到相差一个尺度,因此解 $\mathbf{h}$ 给出所要求的 H. $\mathbf{h}$ 的非零因子可以通过对范数的要求来任意选择,例如要求 $\|\mathbf{h}\|=1$.

### 3.1.1 超定解

如果给出的点对应 $\mathbf{x}_i\leftrightarrow\mathbf{x}_i'$ 多于四组,那么由(3.3)导出的方程组 $\mathrm{A}\mathbf{h}=\mathbf{0}$ 是超定的. 如果点的位置是精确的,那么 A 的秩仍然为 8 并有一维零空间,并且存在精确解 $\mathbf{h}$. 如果图像坐标的测量是不精确的(通常称为**噪声**),情况将是两样——除零解外,超定方程组 $\mathrm{A}\mathbf{h}=\mathbf{0}$将不存在精确解. 取代求精确解,我们试图寻找一个近似解,即求使一个适当的代价函数取最小值的矢量 $\mathbf{h}$. 那么自然会产生的问题是:应该最小化什么? 显然,为避开 $\mathbf{h}=\mathbf{0}$的解需要附加一个约束. 通常附加范数条件,例如 $\|\mathbf{h}\|=1$. 范数的值是不重要的,因为 H 仅定义到相差一个尺度. 既然不存在 $\mathrm{A}\mathbf{h}=\mathbf{0}$ 的精确解,很自然会在通常约束 $\|\mathbf{h}\|=1$ 下最小化范数 $\|\mathrm{A}\mathbf{h}\|$. 这等价于求商 $\|\mathrm{A}\mathbf{h}\|/\|\mathbf{h}\|$ 的最小值问题. 如A3.4.2($p$412)所示,该解是 $\mathrm{A}^{\mathsf{T}}\mathrm{A}$ 的最小特征值的(单位)特征矢量. 也可以说,该解是 A 最小奇异值的单位奇异矢量. 由此所得到的算法称为基本 DLT 算法并概括在算法 3.1 中.

### 3.1.2 非齐次解

除把 $\mathbf{h}$ 直接作为齐次矢量来解以外,另一种方法是把等式(3.3)转成非齐次线性方程组,即给矢量 $\mathbf{h}$ 中的某个元素强加上 $h_j=1$ 的条件. 强加条件 $h_j=1$ 的道理是:既然允许解

**目标**

给定 $n(\geqslant 4)$组 2D 到 2D 点对应$\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$,确定 2D 单应矩阵 H 使得 $\mathbf{x}'_i = \mathrm{H}\mathbf{x}_i$.

**算法**

(1)根据每组对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 由(3.1)计算矩阵 $\mathrm{A}_i$. 通常仅需要使用其前两行.

(2)将 $n$ 个 $2\times 9$ 的矩阵 $\mathrm{A}_i$ 组合成一个 $2n\times 9$ 的矩阵 A.

(3)求 A 的 SVD(节 A3.3($p$408)). 对应于最小特征值的单位特征矢量便是解 $\mathbf{h}$. 具体地说,如果 $\mathrm{A}=\mathrm{UDV}^{\mathrm{T}}$且 D 为对角矩阵具有正对角元素并沿对角线按降序排列,那么 $\mathbf{h}$ 是 V 的最后一列.

(4)矩阵 H 由 $\mathbf{h}$ 按(3.2)确定.

算法 3.1 H 的基本 DLT(参看包含归一化过程的算法 3.2($p$68))

相差一个任意因子,当然可以通过选择因子值使得 $h_j=1$. 例如,如果把 $\mathbf{h}$ 的最后一个元素(它对应 $\mathrm{H}_{33}$)选为 1,那么(3.3)为

$$\begin{bmatrix} 0 & 0 & 0 & -x_iw'_i & -y_iw'_i & -w_iw'_i & x_iy'_i & y_iy'_i \\ x_iw'_i & y_iw'_i & w_iw'_i & 0 & 0 & 0 & -x_ix'_i & -y_ix'_i \end{bmatrix}\tilde{\mathbf{h}} = \begin{pmatrix} -w_iy'_i \\ -w_ix'_i \end{pmatrix}$$

其中 $\tilde{\mathbf{h}}$ 是由 $\mathbf{h}$ 的前 8 个元素组成的 8 维矢量. 把四组对应组合成形如 $\mathrm{M}\tilde{\mathbf{h}}=\mathbf{b}$ 的矩阵方程,其中 M 有 8 列而 $\mathbf{b}$ 是 8 维矢量. 该矩阵方程在 M 仅有 8 行(最小配置解)时可用解线性方程的标准方法(如高斯消去法),或在超定方程组时用最小二乘法(节 A3.3.1($p$409))来求解 $\tilde{\mathbf{h}}$.

但是,如果事实上真正的解是 $h_j=0$,那么不存在一个因子 $k$ 使 $kh_j=1$. 这意味着令 $h_j=1$得不到真解. 由于这个原因,如果被选的 $h_j$ 接近于零,则可以预见此方法会导致不稳定解. 因此,一般不提倡这种方法.

**例 3.1** 本例指出如果 H 把坐标原点映射到无穷远点,则 $h_9=\mathrm{H}_{33}$为零. 因为$(0,0,1)^{\mathrm{T}}$ 表示坐标原点 $\mathbf{x}_0$,并且$(0,0,1)^{\mathrm{T}}$ 也表示无穷远直线 $\mathbf{l}$,该条件可以写成 $\mathbf{l}^{\mathrm{T}}\mathrm{H}\mathbf{x}_0=(0,0,1)\mathrm{H}(0,0,1)^{\mathrm{T}}=0$,从而 $\mathrm{H}_{33}=0$. 在景物平面的透视图像中,无穷远线被影像为该平面的消影线(见第 7 章),例如地平线是地平面的消影线. 地平线过图像中心且坐标原点与图像中心重合的情况并不少见. 在这种情形下,图像到世界平面的映射把原点映射到无穷远直线上,从而真实解为 $\mathrm{H}_{33}=h_9=0$. 因此,在此实际情形下用 $h_9=1$ 归一化是严重失误的. △

### 3.1.3 退化配置

设用来计算单应的最小配置解的四组点对应中有三点 $\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3$ 共线. 问这种情况是否有意义? 如果对应点$\mathbf{x}'_1,\mathbf{x}'_2,\mathbf{x}'_3$也共线,那么我们可以怀疑单应不是充分约束的,而且存在把 $\mathbf{x}_i$ 映射到 $\mathbf{x}'_i$ 的一簇单应. 另一方面,如果对应点 $\mathbf{x}'_1,\mathbf{x}'_2,\mathbf{x}'_3$ 不共线,那么显然不存在把 $\mathbf{x}_i$ 变到$\mathbf{x}'_i$的变换 H,因为射影变换必须保持共线性. 但是由(3.3)导出的八个齐次方程必然有一个非零解,因而必然产生一个矩阵 H. 这样一个明显的矛盾怎样去解决呢?

方程(3.3)表示的条件是 $i=1,\ldots,4,\mathbf{x}'_i\times\mathrm{H}\mathbf{x}_i=\mathbf{0}$,因而通过解 8 个方程找到的矩阵 H 将满足这条件. 假定 $\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3$ 共线而 $\mathbf{l}$ 是它们所在的直线,即对 $i=1,\ldots,3,\mathbf{l}^{\mathrm{T}}\mathbf{x}_i=0$. 现在

定义 $H^* = \mathbf{x}_4'\mathbf{l}^T$，它是一个秩 1 的 $3\times3$ 矩阵. 可以验证对 $i=1,\ldots,3$，$H^*\mathbf{x}_i = \mathbf{x}_4'(\mathbf{l}^T\mathbf{x}_i) = \mathbf{0}$，因为 $\mathbf{l}^T\mathbf{x}_i = 0$. 另一方面，$H^*\mathbf{x}_4 = \mathbf{x}_4'(\mathbf{l}^T\mathbf{x}_4) = k\mathbf{x}_4'$. 因此，对所有 $i$，$\mathbf{x}_i' \times H^*\mathbf{x}_i = \mathbf{0}$. 注意与 $H^*$ 对应的矢量 $\mathbf{h}^*$ 由 $\mathbf{h}^{*T} = (x_4'\mathbf{l}^T, y_4'\mathbf{l}^T, w_4'\mathbf{l}^T)$ 给出，而且不难验证对所有 $i$ 该矢量都满足(3.3). 解 $H^*$ 的问题在于 $H^*$ 是秩 1 矩阵因而不能表示一个射影变换. 因此，$H^*\mathbf{x}_i = \mathbf{0}$，$i=1,2,3$ 的点是不良定义的.

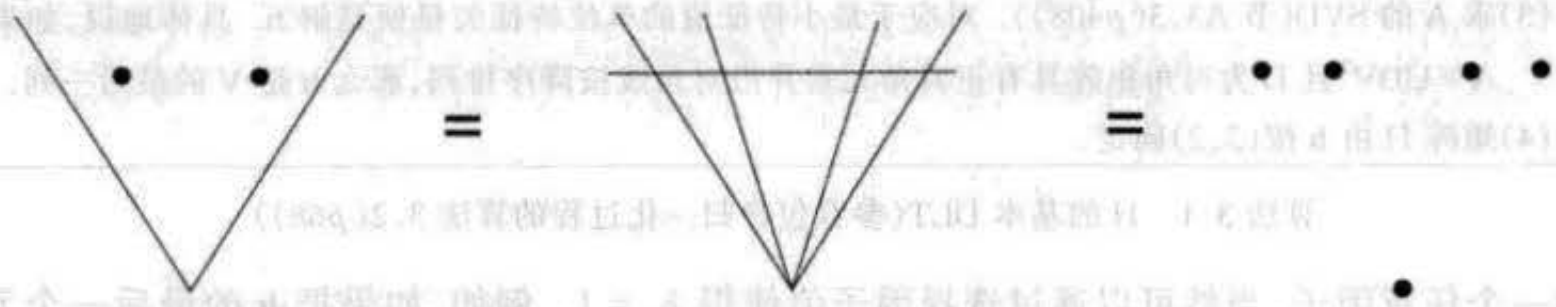


图 3.1 **点－线配置的几何等价性**. 两线与两点的配置等价于四线共点的五直线，或四点共线的五点.

我们已经证明如果 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ 共线，那么 $H^* = \mathbf{x}_4'\mathbf{l}^T$ 是(3.1)的一个解. 这里有两种情形：或者 $H^*$ 是唯一解(相差一尺度因子)，或者存在另一个解 H. 对第一种情形，因为 $H^*$ 是奇异矩阵，不存在把每个 $\mathbf{x}_i$ 变到 $\mathbf{x}_i'$ 的变换. 当 $\mathbf{x}_1,\ldots,\mathbf{x}_3$ 共线而 $\mathbf{x}_1',\ldots,\mathbf{x}_3'$ 不共线时就会产生这种情形. 对第二种情形，有另一解 H 存在，从而任何形如 $\alpha H^* + \beta H$ 的矩阵都是解. 因此有一个 2 参数的变换簇存在，进而推出(3.3)给出的 8 个方程不是独立的.

对于具体的一类变换而言，出现某种配置不能确定唯一解的情形称为**退化**. 注意退化的定义既涉及配置也涉及变换类型. 而且退化问题不仅限于最小配置解. 如果多出来的(精确的，即无误差的)点对应也共线(在 $\mathbf{l}$ 上)，那么退化问题仍没有解决.

### 3.1.4 由线和其他实体求解

到目前为止以及本章的其余部分，都只讨论由点对应来计算单应. 然而也可以平行地讨论由线对应来计算单应. 从直线变换 $\mathbf{l}_i = H^T\mathbf{l}_i'$ 出发，可以导出形如 $A\mathbf{h} = \mathbf{0}$ 的矩阵方程，最小配置解需要一般位置上的四组线对应. 类似地，单应也可以由二次曲线的对应来计算并依此类推.

于是就产生计算一个单应(或任何其他关系)需要多少组对应的问题. 一般的原则是约束数必须等于或大于变换的自由度数. 例如，2D 中每组点或线对应产生关于 H 的两个约束，3D 中每组点或平面对应产生三个约束. 因此在 2D 中四组点或线对应就足以计算出 H，因为单应的自由度数是 $4\times2=8$. 在 3D 中一个单应有 15 个自由度，则需要五组点对应或五组平面对应. 对于一个平面仿射变换(6dof)仅需要三组对应点或线对应，如此等等. 一条二次曲线为 2D 单应提供五个约束.

但采用混合类型的对应来计算 H 时必须谨慎. 例如一个 2D 单应不能由两组点对应和两组线对应唯一确定，但能由三组点和一组线或一组点和三组线来唯一确定，即使每种配置都有 8 个自由度也是如此. 三组线和一组点的情形几何上等价于四组点，因为三条线定义一个三角形而该三角形的顶点唯一地定义三点. 我们已经证明一般位置上的四组点对应能唯一确定一个单应，由此表明三组线和一组点的对应同样能唯一确定一个单应. 类似地，三组点和一组线的情形等价于四组线，同理一般位置上的四组线(即没有三条线共点)对应也唯一确定一个单应. 但是，如一个简要的示意图(图 3.1)所示，两个点和两条线的情形等价

于有四线共点的五条直线，或四点共线的五个点. 如前节所指出，这种配置是退化的，并且存在把该两点与两线配置映射到对应的配置的 1 参数单应簇.

## 3.2 不同的代价函数

我们现在来介绍为确定 H 的超定解而需最小化的若干代价函数. 最小化这些函数的方法在本章的后面介绍.

### 3.2.1 代数距离

DLT 算法最小化范数 $\|\mathbf{A}\mathbf{h}\|$. 矢量 $\boldsymbol{\varepsilon}=\mathbf{A}\mathbf{h}$ 称为残差矢量，并且要最小化的正是该误差矢量的范数. 该矢量的分量来自组成矩阵 A 的行的各点对应. 每组对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 按(3.1)或(3.3)贡献部分误差矢量 $\boldsymbol{\varepsilon}_i$ 并积累成总误差矢量 $\boldsymbol{\varepsilon}$. 矢量 $\boldsymbol{\varepsilon}_i$ 被称为关联于点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 和单应 H 的**代数误差矢量**. 该矢量的范数是一个标量，称为**代数距离**：

$$d_{\mathrm{alg}}(\mathbf{x}'_i, \mathrm{H}\mathbf{x}_i)^2 = \|\boldsymbol{\varepsilon}_i\| = \left\| \begin{bmatrix} \mathbf{0}^\mathsf{T} & -w'_i\mathbf{x}_i^\mathsf{T} & y'_i\mathbf{x}_i^\mathsf{T} \\ w'_i\mathbf{x}_i^\mathsf{T} & \mathbf{0}^\mathsf{T} & -x'_i\mathbf{x}_i^\mathsf{T} \end{bmatrix} \mathbf{h} \right\|^2. \tag{3.4}$$

对于任何两个矢量 $\mathbf{x}_1$ 和 $\mathbf{x}_2$，我们可以用更一般和简洁的写法：

$$d_{\mathrm{alg}}(\mathbf{x}_1, \mathbf{x}_2)^2 = a_1^2 + a_2^2 \quad \text{其中 } \mathbf{a} = (a_1, a_2, a_3)^\mathsf{T} = \mathbf{x}_1 \times \mathbf{x}_2.$$

代数距离与几何距离的关系将在节 3.2.4 中介绍.

给定对应的集合，量 $\boldsymbol{\varepsilon}=\mathbf{A}\mathbf{h}$ 是整个集合的代数误差矢量，由此可见：

$$\sum_i d_{\mathrm{alg}}(\mathbf{x}'_i, \mathrm{H}\mathbf{x}_i)^2 = \sum_i \|\boldsymbol{\varepsilon}_i\|^2 = \|\mathbf{A}\mathbf{h}\|^2 = \|\boldsymbol{\varepsilon}\|^2. \tag{3.5}$$

代数距离的概念起源于 Bookstein[Bookstein－79]关于二次曲线匹配的工作. 其缺点是被最小化的量没有几何或统计上的意义. 正如 Bookstein 所指出的，最小化代数距离的解也许不是直觉所期待的. 但是，通过选择适当的归一化(如节 3.4 将要讨论的)，最小化代数距离的方法能够给出非常好的结果. 其特殊的优越性在于线性(因而是唯一)解以及计算代价小. 通常代数距离的解被用作几何或统计代价函数的非线性最小化的起点. 非线性最小化给这个解一个最后的“抛光”.

### 3.2.2 几何距离

下面我们讨论另一种误差函数，它基于图像上几何距离的测量并最小化图像坐标的测量值与估计值之差.

**记号** 矢量 $\mathbf{x}$ 表示**测量**的图像坐标；$\hat{\mathbf{x}}$ 表示该点的估计值而 $\bar{\mathbf{x}}$ 表示该点的真值.

**单图像误差** 我们首先考虑第一幅图像的测量非常精确而误差仅出现在第二幅图像的情形. 显然，大多数实际图像不是如此. 这种假设比较合理的实例是估计标定模块或世界平面(其中点的测量精度非常高)与它们的图像之间的射影变换. 适宜最小化的量是**转移误差**. 它是第二幅图像上测量点 $\mathbf{x}'$ 与(从第一幅图像的对应点 $\bar{\mathbf{x}}$ 映射得到的)点 $\mathrm{H}\bar{\mathbf{x}}$ 之间的欧氏距离. 我们用标记 $d(\mathbf{x},\mathbf{y})$ 表示非齐次点 $\mathbf{x}$ 和 $\mathbf{y}$ 之间的欧氏距离. 那么对应集合的转移误差是

$$\sum_i d(\mathbf{x}'_i, \mathrm{H}\bar{\mathbf{x}}_i)^2. \tag{3.6}$$

算法要估计的单应 $\hat{\mathrm{H}}$是使误差(3.6)取最小值的单应.

**对称转移误差** 更切合实际的情形是图像测量误差在两幅图像中都发生,从而应该最小化两幅而不仅是一幅图像的误差. 一个较令人满意的误差函数构造法是同时考虑前向变换(H)和后向变换($\mathrm{H}^{-1}$)并把这两种变换对应的几何误差累加. 所得误差为

$$\sum_i d(\mathbf{x}_i, \mathrm{H}^{-1}\mathbf{x}'_i)^2 + d(\mathbf{x}'_i, \mathrm{H}\mathbf{x}_i)^2. \tag{3.7}$$

和式中第一项是第一幅图像的转移误差,第二项是第二幅图像的转移误差. 算法要估计的单应是使(3.7)取最小值的单应.

### 3.2.3 重投影误差——两幅图像

对每幅图像误差作量化的另一种方法是估计每组对应的"校正值". 有人会问为了得到图像点集的完全匹配而在每幅图像中进行测量校正有多大的必要性. 为此,应将它与单图像的几何转移误差(3.6)相对比,其中(3.6)测量的是校正值:为了得到完全的匹配点集有必要在一幅图像(第二幅图像)上对测量进行校正.

现在我们要寻找一个单应 $\hat{\mathrm{H}}$和**完全**匹配的点对$\hat{\mathrm{x}}_i$ 和$\hat{\mathrm{x}}'_i$ 以最小化总的误差函数

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 \quad \text{s.t} \quad \hat{\mathbf{x}}'_i = \hat{\mathrm{H}}\hat{\mathbf{x}}_i \quad \forall i. \tag{3.8}$$

最小化这个代价函数包括确定 $\hat{\mathrm{H}}$和附加的对应集合$\{\hat{\mathbf{x}}_i\}$和$\{\hat{\mathbf{x}}'_i\}$. 这种估计可对诸如世界平面点的图像对应点 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 的测量建模. 我们希望先由 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 估计世界平面的点 $\hat{\mathbf{X}}_i$,然后把它**重投影**到估计上认为是完全匹配的对应$\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i$ 上.

重投影误差函数与对称误差函数的比较在图 3.2 中给出. 在节 3.3 中我们将会看到该单应和对应的最大似然估计与误差函数(3.8)相关.

### 3.2.4 几何和代数距离的比较

我们回到误差仅出现在第二幅图像的情形. 令 $\mathbf{x}'_i = (x'_i, y'_i, w'_i)^\mathsf{T}$ 并定义矢量 $(\hat{x}'_i, \hat{y}'_i, \hat{w}'_i)^\mathsf{T} = \hat{\mathbf{x}}'_i = \mathrm{H}\bar{\mathbf{x}}$. 用此标记,(3.3)的左边变成

$$\mathrm{A}_i\mathbf{h} = \boldsymbol{\varepsilon}_i = \begin{bmatrix} y'_i\,\hat{w}'_i - w'_i\,\hat{y}'_i \\ w'_i\,\hat{x}'_i - x'_i\,\hat{w}'_i \end{bmatrix}.$$

该矢量是关于点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 和摄像机映射 H 的**代数误差矢量**. 因此,

$$d_{\mathrm{alg}}(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 = (y'_i\,\hat{w}'_i - w'_i\,\hat{y}'_i)^2 + (w'_i\,\hat{x}'_i - x'_i\,\hat{w}'_i)^2.$$

而点 $\mathbf{x}'_i$ 和$\hat{\mathbf{x}}'_i$ 的几何距离是

$$\begin{aligned} d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i) &= ((x'_i/w'_i - \hat{x}'_i/\hat{w}'_i)^2 + (y'_i/w'_i - \hat{y}'_i/\hat{w}'_i)^2)^{1/2} \\ &= d_{\mathrm{alg}}(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)/\hat{w}'_i w'_i. \end{aligned}$$

因此,几何距离与代数距离相关,但不相等. 然而应注意,如果$\hat{w}'_i = w'_i = 1$,那么两个距离相等.

我们总可以假定 $w_i = 1$,即把 $\mathbf{x}_i$ 表示成常用的形式 $\mathbf{x}_i = (x_i, y_i, 1)^\mathsf{T}$. 对于一类重要的

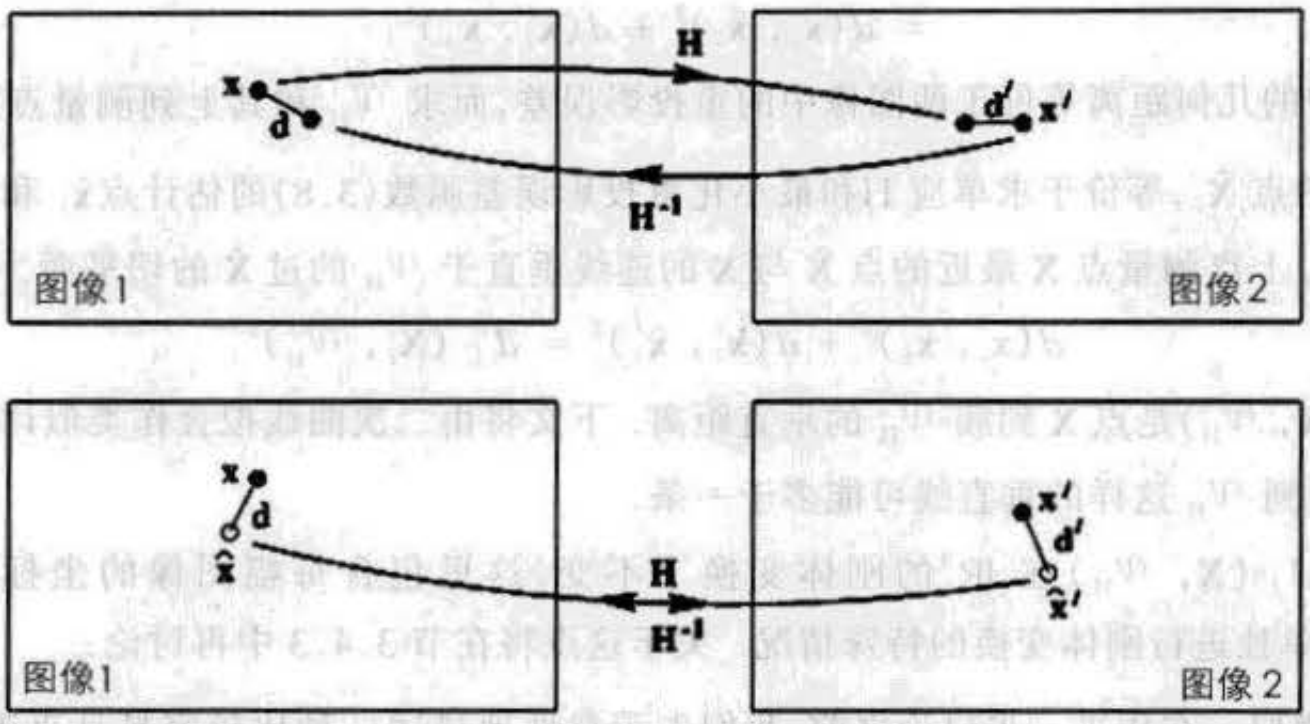



图 3.2 对称转移误差(上图)和重投影误差(下图)在估计单应时的比较. 点 $\mathbf{x}$ 和 $\mathbf{x}'$ 是测量得到(有噪声)的点. 根据估计的单应,点 $\mathbf{x}'$ 和 $\mathrm{H}\mathbf{x}$ 不完全对应(点 $\mathbf{x}$ 和 $\mathrm{H}^{-1}\mathbf{x}'$ 也不完全对应). 但估计的点 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$ 却通过 $\hat{\mathbf{x}}'=\mathrm{H}\hat{\mathbf{x}}$ 完全对应. 令 $d(\mathbf{x},\mathbf{y})$ 表示 $\mathbf{x}$ 和 $\mathbf{y}$ 的欧氏图像距离,则对称转移误差是 $d(\mathbf{x},\mathrm{H}^{-1}\mathbf{x}')^2+d(\mathbf{x}',\mathrm{H}\mathbf{x})^2$; 重投影误差是 $d(\mathbf{x},\hat{\mathbf{x}})^2+d(\mathbf{x}',\hat{\mathbf{x}}')^2$.

2D 单应, $\hat{w}_i'$ 的值也总可以是 1. 2D **仿射**变换用(1.10 − $p$13)形式的矩阵表示为:

$$\mathrm{H}_{\mathrm{A}}=\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.9}$$

我们可以从 $\hat{\mathbf{x}}_i'=\mathrm{H}_{\mathrm{A}}\bar{\mathbf{x}}_i$ 直接验证:若 $w_i=1$ 则 $\hat{w}_i'=1$. 这表明在仿射变换下,几何距离和代数距离相等. DLT 算法很容易做到强制 H 的最后一行为(0,0,1)(通过令 $h_7=h_8=0$)这个条件. 因此,对于仿射变换,最小化几何距离可以用基于代数距离的线性 DLT 算法.

### 3.2.5 重投影误差的几何解释

两平面之间的单应估计可以视为用 4D 空间 $\mathbb{R}^4$ 中的"曲面"来拟合点. 每对图像点 $\mathbf{x}$、$\mathbf{x}'$ 定义测量空间 $\mathbb{R}^4$ 的一个点,记作 $\mathbf{X}$,其坐标由 $\mathbf{x}$ 和 $\mathbf{x}'$ 的非齐次坐标拼成. 对一个给定的单应 H,满足 $\mathbf{x}'\times\mathrm{H}\mathbf{x}=\mathbf{0}$ 的图像对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 定义了 $\mathbb{R}^4$ 中的一个代数簇 $\mathcal{V}_{\mathrm{H}}$①,它是两个超二次曲面的交集. 该曲面之所以在 $\mathbb{R}^4$ 中是二次的是因为(3.1)的每一行是关于 $x$, $y$, $x'$, $y'$ 的二次多项式. H 中的元素确定多项式的每一项的系数,因而 H 定义了一个具体的二次曲面. (3.1)的两个独立的方程定义了两个这样的二次曲面.

给定 $\mathbb{R}^4$ 中的点 $\mathbf{X}_i=(x_i,\ y_i,\ x_i',\ y_i')^{\mathrm{T}}$,单应估计的任务变成寻找一个通过(或几乎通过)点 $\mathbf{X}_i$ 的簇 $\mathcal{V}_{\mathrm{H}}$. 当然,一般不可能精确地与一个簇拟合. 因此,令 $\mathcal{V}_{\mathrm{H}}$ 为对应于变换 H 的某个簇,而对每一点 $\mathbf{X}_i$,令 $\hat{\mathbf{X}}_i=(\hat{x}_i,\ \hat{y}_i,\ \hat{x}_i',\ \hat{y}_i')^{\mathrm{T}}$ 为簇 $\mathcal{V}_{\mathrm{H}}$ 上最靠近 $\mathbf{X}_i$ 的点. 可以立即看到:

$$\|\mathbf{X}_i-\hat{\mathbf{X}}_i\|^2=(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2+(x_i'-\hat{x}_i')^2+(y_i'-\hat{y}_i')^2$$

① **代数簇**是定义在 $\mathbb{R}^N$ 中的一个或多个多元多项式的公共零点集.

$$= d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2.$$

因此，$\mathbb{R}^4$中的几何距离等价于两图像中的重投影误差，而求 $\mathcal{V}_H$ 和其上到测量点 $\mathbf{X}_i$ 距离平方和最小的点 $\hat{\mathbf{X}}_i$，等价于求单应 $\hat{H}$和最小化重投影误差函数(3.8)的估计点$\hat{\mathbf{x}}_i$ 和$\hat{\mathbf{x}}'_i$.

在 $\mathcal{V}_H$ 上离测量点 $\mathbf{X}$ 最近的点 $\hat{\mathbf{X}}$ 与 $\mathbf{X}$ 的连线垂直于 $\mathcal{V}_H$ 的过 $\hat{\mathbf{X}}$ 的切平面. 因此，

$$d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 = d_\perp(\mathbf{X}_i, \mathcal{V}_H)^2$$

其中 $d_\perp(\mathbf{X}, \mathcal{V}_H)$是点 $\mathbf{X}$ 到簇 $\mathcal{V}_H$ 的垂直距离. 下文将由二次曲线拟合作类似讨论，我们将看到，从 $\mathbf{X}$ 到 $\mathcal{V}_H$ 这样的垂直线可能多于一条.

距离 $d_\perp(\mathbf{X}, \mathcal{V}_H)$在 $\mathbb{R}^4$的刚体变换下不变，这里包含每幅图像的坐标$(x, y)^T$，$(x', y')^T$单独进行刚体变换的特殊情况. 关于这点将在节 3.4.3 中再讨论.

**二次曲线类似** 在作进一步讨论之前，我们先示意性地描述一种比较容易且直观的类似估计问题. 即用二次曲线拟合 2D 点的问题，它处于直线拟合(没有曲率，太简单)和单应拟合(四维，并具有非零曲率)的中间位置.

考虑用二次曲线拟合平面上 $n>5$ 个点$(x_i, y_i)^T$ 并且使基于几何距离的误差最小. 这些点可以视为$x_i \leftrightarrow y_i$的"对应". 其转移距离和重投影(垂直)距离在图 3.3 中说明. 从图中清楚地看到 $d_\perp$ 小于或等于转移误差.

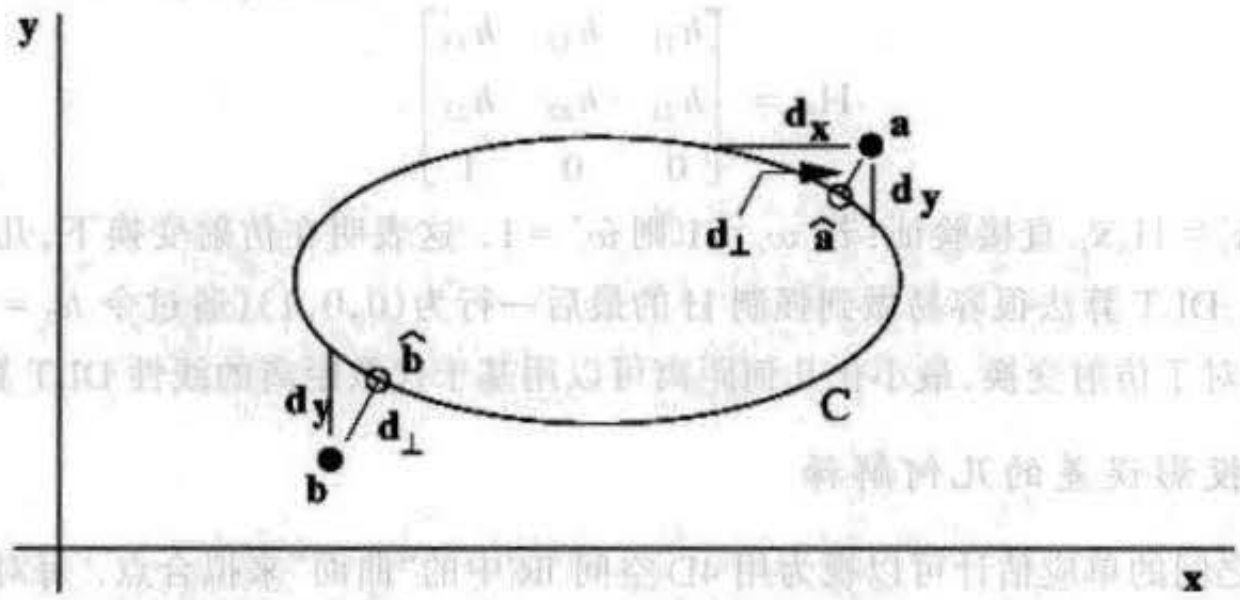



图 3.3 由一组 2D 点通过最小化"对称转移误差"$d_x^2 + d_y^2$ 或垂直距离的平方和 $d_\perp^2$ 来估计一条二次曲线. 对于转移误差须假定 $x$ 是精确的，只测量沿 $y$ 方向到二次曲线的距离 $d_y$，类似地只测量 $d_x$. 对点 **a** 而言，显然 $d_\perp \leqslant d_x$ 且 $d_\perp \leqslant d_y$. 而且 $d_\perp$ 比 $d_x$ 和 $d_y$ 更稳定，如图中点 **b** 所示，它的 $d_x$ 无法定义.

点 $\mathbf{x}$ 到二次曲线 C 的代数距离定义为 $d_{alg}(\mathbf{x}, C)^2 = \mathbf{x}^T C \mathbf{x}$. C 的一个线性解可以在 C 适当归一化后由最小化 $\sum_i d_{alg}(\mathbf{x}_i, C)^2$ 得到. 点 $\mathbf{x}$ 到 C 的垂直距离没有线性表示，因为过 $\mathbb{R}^2$ 中的每一点最多有 4 条线垂直于 C. 可以由四次方程求根得到此解. 但是，函数 $d_\perp(\mathbf{x}, C)$ 可以定义为二次曲线和点的最短距离的解. 这样，二次曲线可以通过在 C 的五个参数上最小化 $\sum_i d_\perp(\mathbf{x}, C)^2$ 来估计，虽然它不能由线性解得到. 给定二次曲线 C 和测量点 $\mathbf{x}$，求校正点$\hat{\mathbf{x}}$就是选择 C 上最近的点.

再回到单应的估计问题. 在仿射变换时，该代数簇是两张超平面的交，即它是一个 2 维线性子空间. 这是由于仿射矩阵的(3.9)形式：$\mathbf{x}' = H_A \mathbf{x}$ 给出 $x, x', y$ 之间的一个线性约束

和$x, y, y'$之间的另一个约束，每个约束定义 $\mathbb{R}^4$中的一个超平面. 与此相类似的情形是用直线拟合平面上的点. 在这两种情形中，关系(仿射变换或直线)都可以通过最小化点到该代数簇的垂直距离来估计. 在下一节中将看到这两种情形都存在封闭形式的解.

### 3.2.6 Sampson 误差

几何误差(3.8)的性质相当复杂，而且它的最小化需要同时估计单应矩阵和点$\hat{\mathbf{x}}_i$、$\hat{\mathbf{x}}'_i$. 这种非线性估计问题将在节 3.5 作进一步讨论. 其复杂性与代数误差最小化(3.4)的简单性形成鲜明对比. 节 3.2.5 对几何误差的几何解释引向另一种代价函数，其复杂性介于代数和几何代价函数之间，但非常近似于几何误差. 该代价函数称为 **Sampson 误差**，因为 Sampson[Sampson-82]曾把它用于二次曲线的拟合.

如节 3.2.5 所介绍，最小化几何误差 $\|\mathbf{X}-\hat{\mathbf{X}}\|^2$ 的矢量 $\hat{\mathbf{X}}$ 是簇 $\mathcal{V}_{\mathrm{H}}$ 上最接近测量点 $\mathbf{X}$ 的点. 由于代数簇 $\mathcal{V}_{\mathrm{H}}$ 的非线性本质，点 $\mathbf{X}$ 不能直接估计，必须通过迭代. Sampson 误差函数的思想是估计点$\hat{\mathbf{X}}$的一阶近似并假定代价函数在被估计点附近有很好的线性近似. 下面我们将直接讨论 2D 单应估计问题，但本质上可以把它不加改变地应用到本书讨论的其他估计问题.

对给定的单应 H，在 $\mathcal{V}_{\mathrm{H}}$ 上的任何点 $\mathbf{X}=(x, y, x', y')^{\mathsf{T}}$ 将满足等式(3.3-$p$54)或 $\mathbf{Ah}=\mathbf{0}$. 为了突出代价函数对 $\mathbf{X}$ 的相关性，我们把它写为 $\mathcal{C}_{\mathrm{H}}(\mathbf{X})=\mathbf{0}$，$\mathcal{C}_{\mathrm{H}}(\mathbf{X})$在此表示一个 2 维矢量. 这个代价函数可以用下列 Taylor 展开式来一阶逼近

$$\mathcal{C}_{\mathrm{H}}(\mathbf{X}+\boldsymbol{\delta}_{\mathbf{x}}) = \mathcal{C}_{\mathrm{H}}(\mathbf{X}) + \frac{\partial \mathcal{C}_{\mathrm{H}}}{\partial \mathbf{X}}\boldsymbol{\delta}_{\mathbf{x}}. \tag{3.10}$$

如果我们记 $\boldsymbol{\delta}_{\mathbf{x}}=\hat{\mathbf{X}}-\mathbf{X}$ 并希望 $\hat{\mathbf{X}}$ 在 $\mathcal{V}_{\mathrm{H}}$ 上，即 $\mathcal{C}_{\mathrm{H}}(\hat{\mathbf{X}})=\mathbf{0}$，便得到 $\mathcal{C}_{\mathrm{H}}(\mathbf{X})+(\partial\mathcal{C}_{\mathrm{H}}/\partial\mathbf{X})\boldsymbol{\delta}_{\mathbf{x}}=\mathbf{0}$，我们把它记成 $\mathrm{J}\boldsymbol{\delta}_{\mathbf{x}}=-\boldsymbol{\varepsilon}$，其中 J 是偏导数矩阵，$\boldsymbol{\varepsilon}$ 是与 $\mathbf{X}$ 相关的代价函数 $\mathcal{C}_{\mathrm{H}}(\mathbf{X})$. 我们面临的最小化问题是求满足此方程的最小 $\boldsymbol{\delta}_{\mathbf{x}}$，即

- **求在满足 $\mathrm{J}\boldsymbol{\delta}_{\mathbf{x}}=-\boldsymbol{\varepsilon}$ 条件下使 $\|\boldsymbol{\delta}_{\mathbf{x}}\|$ 取最小值的矢量 $\boldsymbol{\delta}_{\mathbf{x}}$.**

求解此类问题的标准方法是使用 Lagrange 乘子. 引入 Lagrange 乘子矢量 $\boldsymbol{\lambda}$ 后，问题就转化为最小化 $\boldsymbol{\delta}_{\mathbf{x}}^{\mathsf{T}}\boldsymbol{\delta}_{\mathbf{x}}-2\boldsymbol{\lambda}^{\mathsf{T}}(\mathrm{J}\boldsymbol{\delta}_{\mathbf{x}}+\boldsymbol{\varepsilon})$，其中添加因子 2 仅仅是为了方便. 此式先对 $\boldsymbol{\delta}_{\mathbf{x}}$ 求导并使它等于零得：

$$2\boldsymbol{\delta}_{\mathbf{x}}^{\mathsf{T}}-2\boldsymbol{\lambda}^{\mathsf{T}}\mathrm{J}=\mathbf{0}^{\mathsf{T}},$$

从而得到 $\boldsymbol{\delta}_{\mathbf{x}}=\mathrm{J}^{\mathsf{T}}\boldsymbol{\lambda}$. 再对 $\boldsymbol{\lambda}$ 求导给出 $\mathrm{J}\boldsymbol{\delta}_{\mathbf{x}}+\boldsymbol{\varepsilon}=\mathbf{0}$，即原来的约束. 消去 $\boldsymbol{\delta}_{\mathbf{x}}$ 得

$$\mathrm{J}\mathrm{J}^{\mathsf{T}}\boldsymbol{\lambda}=-\boldsymbol{\varepsilon}$$

对 $\boldsymbol{\lambda}$ 求解得 $\boldsymbol{\lambda}=-(\mathrm{J}\mathrm{J}^{\mathsf{T}})^{-1}\boldsymbol{\varepsilon}$，最后得

$$\boldsymbol{\delta}_{\mathbf{x}}=-\mathrm{J}^{\mathsf{T}}(\mathrm{J}\mathrm{J}^{\mathsf{T}})^{-1}\boldsymbol{\varepsilon}, \tag{3.11}$$

和 $\hat{\mathbf{X}}=\mathbf{X}+\boldsymbol{\delta}_{\mathbf{x}}$. 范数 $\|\boldsymbol{\delta}_{\mathbf{x}}\|^2$ 是 Sampson 误差：

$$\|\boldsymbol{\delta}_{\mathbf{x}}\|^2=\boldsymbol{\delta}_{\mathbf{x}}^{\mathsf{T}}\boldsymbol{\delta}_{\mathbf{x}}=\boldsymbol{\varepsilon}^{\mathsf{T}}(\mathrm{J}\mathrm{J}^{\mathsf{T}})^{-1}\boldsymbol{\varepsilon}. \tag{3.12}$$

**例 3.2　二次曲线的 Sampson 近似**

我们来计算图 3.3 的点 $\mathbf{x}$ 和二次曲线 C 之间的几何距离 $d_{\perp}(\mathbf{x}, \mathrm{C})$的 Sampson 近似. 在这种情况下，二次曲线的代数簇 $\mathcal{V}_{\mathrm{C}}$ 由等式 $\mathbf{x}^{\mathsf{T}}\mathrm{C}\mathbf{x}=0$ 定义，$\mathbf{X}=(x, y)^{\mathsf{T}}$ 是一个 2 维矢

量，$\boldsymbol{\varepsilon}=\mathbf{x}^{\mathsf{T}}\mathrm{C}\mathbf{x}$ 是标量，而 J 是 $1\times 2$ 矩阵：

$$\mathrm{J}=\left[\frac{\partial(\mathbf{x}^{\mathsf{T}}\mathrm{C}\mathbf{x})}{\partial x},\ \frac{\partial(\mathbf{x}^{\mathsf{T}}\mathrm{C}\mathbf{x})}{\partial y}\right].$$

这意味着 $\mathrm{JJ}^{\mathsf{T}}$ 是标量. J 的元素用复合函数求导规则计算，即

$$\frac{\partial(\mathbf{x}^{\mathsf{T}}\mathrm{C}\mathbf{x})}{\partial x}=\frac{\partial\mathbf{x}^{\mathsf{T}}}{\partial x}\frac{\partial(\mathbf{x}^{\mathsf{T}}\mathrm{C}\mathbf{x})}{\partial\mathbf{x}^{\mathsf{T}}}=2(1,0,0)\mathrm{C}\mathbf{x}=2(\mathrm{C}\mathbf{x})_1$$

其中 $(\mathrm{C}\mathbf{x})_i$ 表示 3 维矢量 $\mathrm{C}\mathbf{x}$ 的第 $i$ 个元素. 由(3.12)推出

$$d_{\perp}^2=\|\boldsymbol{\delta}_{\mathbf{x}}\|^2=\boldsymbol{\varepsilon}^{\mathsf{T}}(\mathrm{JJ}^{\mathsf{T}})^{-1}\boldsymbol{\varepsilon}=\frac{\boldsymbol{\varepsilon}^{\mathsf{T}}\boldsymbol{\varepsilon}}{\mathrm{JJ}^{\mathsf{T}}}=\frac{(\mathbf{x}^{\mathsf{T}}\mathrm{C}\mathbf{x})^2}{4((\mathrm{C}\mathbf{x})_1^2+(\mathrm{C}\mathbf{x})_2^2)}\qquad\Delta$$

几点注解：

(1) 2D 单应估计中的 $\mathbf{X}=(x,y,x',y')^{\mathsf{T}}$，其中 2D 测量是 $\mathbf{x}=(x,y,1)^{\mathsf{T}}$ 和 $\mathbf{x}'=(x',y',1)^{\mathsf{T}}$.

(2) $\boldsymbol{\varepsilon}=\mathcal{C}_{\mathrm{H}}(\mathbf{X})$ 是代数误差矢量 $\mathrm{A}_i\mathbf{h}$——一个二维矢量——$\mathrm{A}_i$ 在(3.3－$p$54)中定义.

(3) $\mathrm{J}=\partial\mathcal{C}_{\mathrm{H}}(\mathbf{X})/\partial\mathbf{X}$ 是一个 $2\times 4$ 的矩阵. 例如

$$\mathrm{J}_{11}=\partial(-w_i'\mathbf{x}_i^{\mathsf{T}}\mathbf{h}^2+y_i'\mathbf{x}_i^{\mathsf{T}}\mathbf{h}^3)/\partial x=-w_i'h_{21}+y_i'h_{31}.$$

(4) 注意(3.12)与代数误差 $\|\boldsymbol{\varepsilon}\|=\boldsymbol{\varepsilon}^{\mathsf{T}}\boldsymbol{\varepsilon}$ 的相似性. Sampson 误差可以解释成 Mahalanobis范数 $\|\boldsymbol{\varepsilon}\|_{\mathrm{JJ}^{\mathsf{T}}}$(见节 A2.1－($p$402)).

(5) 我们也可以采用(3.1－$p$53)来定义 A，其中 J 的维数是 $3\times 4$ 而 $\boldsymbol{\varepsilon}$ 是一个 3 维的矢量. 但是，Sampson 误差以及其后的解 $\boldsymbol{\delta}_{\mathbf{x}}$ 一般与使用(3.1－$p$53)或(3.3－$p$54)无关.

这里推导的 Sampson 误差(3.12)是针对一组对应点的. 把它应用于由若干点对应 $\mathbf{x}_i\leftrightarrow\mathbf{x}_i'$ 来估计一个 2D 单应 H 时，所有点对应的误差必须相加，得到：

$$\mathcal{D}_{\perp}=\sum_i\boldsymbol{\varepsilon}_i^{\mathsf{T}}(\mathrm{J}_i\mathrm{J}_i^{\mathsf{T}})^{-1}\boldsymbol{\varepsilon}_i\tag{3.13}$$

其中 $\boldsymbol{\varepsilon}$ 和 J 都与 H 有关. 为了估计 H，必须在 H 的所有值上最小化这个公式. 这是一个简单的最小化问题，其中变参数集仅由 H 的元素(或某些其他参数)组成.

Sampson 误差的推导假定每点有各向同性(圆)误差分布，在每幅图像中都如此. 适合更一般的高斯误差分布的公式在本章末的练习中给出.

**线性代价函数**

代数误差矢量 $\mathcal{C}_{\mathrm{H}}(\mathbf{X})=\mathrm{A}(\mathbf{X})\mathbf{h}$ 关于 $\mathbf{X}$ 元素是典型多重线性的. $\mathrm{A}(\mathbf{X})\mathbf{h}$ 是**线性**的对其自身很有意义. 首先在这种情况下由 Taylor 展开式(3.10)给出的几何误差的一阶近似是精确的(更高阶为零)，它表明 **Sampson 误差等同于几何误差**.

另外，由一组线性方程 $\mathcal{C}_{\mathrm{H}}(\mathbf{X})=\mathbf{0}$ 定义的代数簇 $\mathcal{V}_{\mathrm{H}}$ 是 H 的超平面. 从而，求 H 的问题变成超平面拟合的问题——在由 H 参数化的超平面中求与数据 $\mathbf{X}_i$ 拟合最好的超平面.

作为这种思想的一个例子，我们在本章末的练习中介绍了用于仿射变换的几何误差最小化(3.8)的一种线性算法.

### 3.2.7 另一种几何解释

节 3.2.5 指出求把一个点集 $\mathbf{x}_i$ 映射到另一个点集 $\mathbf{x}_i'$ 的单应等价于用一种给定类型的

代数簇去拟合 $\mathbb{R}^4$中的一个点集. 我们现在给出一种不同的解释,其中所有测量的集合表示成测量空间 $\mathbb{R}^N$ 中的一个点.

我们所考虑的估计问题都能套入一个公共的模式. 在抽像的术语中,估计问题包含下列两个成分.

- 一个由**测量矢量 X** 组成的**测量空间** $\mathbb{R}^N$,和
- 一个**模型**,在抽象术语中,该模型视为 $\mathbb{R}^N$ 中点的一个子集 $S$. 在此子集内的一个测量矢量 **X** 称为**满足此模型**. 通常,满足此模型的子空间是一个子流形或 $\mathbb{R}^N$ 中的代数簇.

现在,给定 $\mathbb{R}^N$ 中的一个测量矢量 **X**,估计问题就是求一个离 **X** 最近并满足该模型的矢量 $\hat{\mathbf{X}}$. 我们来说明 2D 单应估计问题如何套入上述模式.

**双图像误差** 令$\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$, $i=1,\ldots,n$ 为被测量匹配点的集合. 总计有 $4n$ 个测量值:两幅图像上各有 $n$ 个点,而每点有两个坐标. 那么该匹配点集表示 $\mathbb{R}^N$ 中的一个点,其中 $N=4n$. 由两幅图像上所有匹配点的坐标组成的矢量将用 **X** 来标记.

当然,并不是所有点对 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 的集合都与一个单应 H 相关联. 对一个射影变换 H 只有对所有 $i$ 都满足 $\mathbf{x}'_i = \mathrm{H}\mathbf{x}_i$ 的点对应集合$\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$才组成 $\mathbb{R}^N$ 中满足该模型的子集. 一般,该点集将组成 $\mathbb{R}^N$ 中某维数的子流形 $S$(事实上是代数簇). 该子流形维数等于用来参数化该流形的最少参数数.

我们可以在第一幅图像中任意选择 $n$ 个点$\hat{\mathbf{x}}_i$. 同时单应 H 也可以任意选择. 一旦选定,在第二幅图像中的点$\hat{\mathbf{x}}'_i$由$\hat{\mathbf{x}}'_i = \mathrm{H}\,\hat{\mathbf{x}}_i$ 来确定. 因此,可选的点由 $2n+8$ 个参数的集合来确定:点$\hat{\mathbf{x}}_i$ 的 $2n$ 个坐标,加上变换 H 的 8 个独立参数(自由度). 因此,子流形 $S \subset \mathbb{R}^N$的维数是 $2n+8$,而余维数是 $2n-8$.

给定对应于 $\mathbb{R}^N$中点 **X** 的测量点对应$\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$和 $S$ 上的估计点 $\hat{\mathbf{X}} \in \mathbb{R}^N$,容易验证

$$\|\mathbf{X}-\hat{\mathbf{X}}\|^2 = \sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2.$$

因此,在 $\mathbb{R}^N$ 中找 $S$ 上离 **X** 最近的点 $\hat{\mathbf{X}}$ 等价于最小化由(3.8)给出的代价函数. 要估计的正确点对应$\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i$是 $\mathbb{R}^N$ 中最靠近曲面点 $\hat{\mathbf{X}}$ 的点. 一旦 $\hat{\mathbf{X}}$ 已知,H 就可算出.

**单图像误差** 对于单图像误差情形其对应集合为$\{\bar{\mathbf{x}}_i \leftrightarrow \mathbf{x}'_i\}$. 点$\bar{\mathbf{x}}_i$ 假定是精确的. 用 $\mathbf{x}'_i$ 的非齐次坐标组成测量矢量 **X**. 因此,此时的测量空间的维数是 $N=2n$. 矢量$\hat{\mathbf{X}}$由精确的点的映射$\{\mathrm{H}\bar{\mathbf{x}}_1, \mathrm{H}\bar{\mathbf{x}}_2, \ldots, \mathrm{H}\bar{\mathbf{x}}_n\}$的非齐次坐标组成. 当 H 在整个单应矩阵集上变化时,满足模型的测量矢量集是集合$\hat{\mathbf{X}}$. 同样该子空间也是一个代数簇. 它的维数是 8,等于单应矩阵 H 的自由度的总数. 如前面一样,余维数是 $2n-8$. 可以验证

$$\|\mathbf{X}-\hat{\mathbf{X}}\|^2 = \sum_i d(\mathbf{x}'_i, \mathrm{H}\bar{\mathbf{x}}_i)^2.$$

因此,求 $S$ 上到测量矢量 **X** 最近的点等价于最小化代价函数(3.6).

## 3.3 统计代价函数和最大似然估计

节 3.2 考虑的各种代价函数都与一幅图像中被估计的和被测量的点之间的几何距离有关. 我们现在对使用这种代价函数的合理性作评估,然后推广到考虑一幅图像中点测量误

差的统计.

为了获得 H 的一个最好(最优)估计,有必要为测量误差(噪声)建模. 这里假定在没有测量误差时,真实的点准确满足一个单应变换,即:$\bar{\mathbf{x}}_i' = \mathrm{H}\,\bar{\mathbf{x}}_i$. 通常假定图像坐标的测量误差遵循高斯(或正态)概率分布. 当然这个假设一般没有被验证,而且也未考虑测量数据中野值(严重有误的测量)的存在. 检测并消去野值的方法将在后面节 3.7 中讨论. 一旦野值被消去,高斯误差模型的假设变得更站得住脚,尽管仍没有严格验证. 因此,从现在起我们假定图像测量误差遵循零均值各向同性高斯分布. 该分布在节 A2.1($p$402)介绍.

具体地说,我们假定每一幅图像坐标都具有零均值和统一的标准差 $\sigma$ 的高斯噪声. 这意味着 $x = \bar{x} + \Delta x$,其中 $\Delta x$ 服从方差为 $\sigma^2$ 的高斯分布. 如果进一步假设每次测量的噪声是相互独立的,那么,若点的真值是 $\bar{\mathbf{x}}$,则每个测量点 $\mathbf{x}$ 的概率密度函数(pdf)是:

$$\Pr(\mathbf{x}) = \left(\frac{1}{2\pi\sigma^2}\right) e^{-d(\mathbf{x},\,\bar{\mathbf{x}})^2/(2\sigma^2)}. \tag{3.14}$$

**单图像误差** 首先考虑误差仅出现在第二幅图像的情形. 因为假定每点的误差是独立的,点对应集 $\{\bar{\mathbf{x}}_i \leftrightarrow \mathbf{x}_i'\}$ 的概率就是它们单个 pdf 的乘积. 从而,受噪声干扰的数据的 pdf 是:

$$\Pr(\{\mathbf{x}_i'\} \mid \mathrm{H}) = \prod_i \left(\frac{1}{2\pi\sigma^2}\right) e^{-d(\mathbf{x}_i',\,\mathrm{H}\bar{\mathbf{x}}_i)^2/(2\sigma^2)}. \tag{3.15}$$

符号 $\Pr(\{\mathbf{x}_i'\} \mid \mathrm{H})$ 解释为给定真实单应 H 时获得测量 $\{\mathbf{x}_i'\}$ 的概率. 该对应集合的**对数似然**为:

$$\log\Pr(\mathbf{x}_i' \mid \mathrm{H}) = -\frac{1}{2\sigma^2}\sum_i d(\mathbf{x}_i',\,\mathrm{H}\bar{\mathbf{x}}_i)^2 + \text{常数}.$$

单应的**最大似然(ML)估计** $\hat{\mathrm{H}}$ 最大化这个对数似然,即最小化:

$$\sum_i d(\mathbf{x}_i',\,\mathrm{H}\bar{\mathbf{x}}_i)^2.$$

因此,ML 估计等价于最小化几何误差函数(3.6 - $p$58).

**双图像误差** 如果真实对应是 $\{\bar{\mathbf{x}}_i \leftrightarrow \mathrm{H}\bar{\mathbf{x}}_i = \bar{\mathbf{x}}_i'\}$,那么与上面的推导类似,受噪声干扰的数据的 pdf 是

$$\Pr(\{\mathbf{x}_i,\,\mathbf{x}_i'\} \mid \mathrm{H},\,\{\bar{\mathbf{x}}_i\}) = \prod_i \left(\frac{1}{2\pi\sigma^2}\right) e^{-(d(\mathbf{x}_i,\,\bar{\mathbf{x}}_i)^2 + d(\mathbf{x}_i',\,\mathrm{H}\bar{\mathbf{x}}_i)^2)/(2\sigma^2)}.$$

这里额外的复杂性是我们必须寻找"校正过"的图像测量,它们担任真值(上面的 $\mathrm{H}\bar{\mathbf{x}}$)的角色. 因此,射影变换 $\hat{\mathrm{H}}$ 和对应 $\{\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}_i'\}$ 的 ML 估计是求最小化

$$\sum_i d(\mathbf{x}_i,\,\hat{\mathbf{x}}_i)^2 + d(\mathbf{x}_i',\,\hat{\mathbf{x}}_i')^2$$

的单应 $\hat{\mathrm{H}}$ 和校正对应 $\{\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}_i'\}$,其中 $\hat{\mathbf{x}}_i' = \mathrm{H}\,\hat{\mathbf{x}}_i$. 注意在这种情形,ML 估计等同于最小化重投影误差函数(3.8).

**Mahalanobis 距离** 在一般高斯分布的情形,可以假定测量矢量 $\mathbf{X}$ 满足一个具有协方差矩阵 $\Sigma$ 的高斯分布函数. 上面的情况等价于协方差矩阵是单位矩阵的倍数.

最大化对数似然则等价于最小化 Mahalanobis 距离(见节 A2.1($p$402))

$$\|\mathbf{X} - \bar{\mathbf{X}}\|_{\Sigma}^2 = (\mathbf{X} - \bar{\mathbf{X}})^{\mathrm{T}}\Sigma^{-1}(\mathbf{X} - \bar{\mathbf{X}}).$$

当每幅图像都有误差并假定一幅图像中的误差与另一幅图像中的误差是独立的时，合适的代价函数是

$$\|\mathbf{X}-\overline{\mathbf{X}}\|_{\Sigma}^{2}+\|\mathbf{X}'-\overline{\mathbf{X}}'\|_{\Sigma'}^{2}$$

其中$\Sigma$和$\Sigma'$是两幅图像的测量的协方差矩阵.

最后，如果我们假定所有点 $\mathbf{x}_i$ 和 $\mathbf{x}'_i$ 的误差是独立的，并分别具有协方差矩阵 $\Sigma_i$ 和 $\Sigma'_i$，那么上面的表达式展开为

$$\sum\|\mathbf{x}_i-\overline{\mathbf{x}}_i\|_{\Sigma_i}^{2}+\sum\|\mathbf{x}'_i-\overline{\mathbf{x}}'_i\|_{\Sigma'_i}^{2} \tag{3.16}$$

这个方程允许用于非各向同性的协方差矩阵，在用两非垂直线的相交来计算点的位置时会出现这种情况. 在已知一幅图像中的点是精确的而误差仅发生在另一幅图像时，(3.16)的两个累加项中的一项消失.

## 3.4 变换不变性和归一化

我们现在来讨论节 3.1 介绍的 DLT 算法的性质和性能并把它与最小化几何误差算法作比较. 第一个议题是算法对图像坐标的不同选择的不变性问题. 显然，我们希望算法的结果不依赖于图像坐标系的原点、尺度甚至定向的选择.

### 3.4.1 关于图像坐标变换的不变性

图像坐标的原点有时设在图像的左上角，有时设在中心，它是否会使变换的计算结果产生差别呢？类似地，如果图像坐标轴上的尺度乘上某个常数因子，是否也会改变算法的结果呢？更一般地，在什么程度上用于估计单应的最小化代价函数的算法结果依赖于图像中坐标的选择呢？例如，假定在算法运行之前，图像坐标由某个相似、仿射甚至射影变换变换了，其结果会不会有实质性改变呢？

正式地说，假定一幅图像的坐标 $\mathbf{x}$ 被 $\tilde{\mathbf{x}}=\mathrm{T}\mathbf{x}$ 替代，而另一幅图像的坐标 $\mathbf{x}'$ 被 $\tilde{\mathbf{x}}'=\mathrm{T}'\mathbf{x}'$ 替代，其中 T 和 T′是 3×3 单应. 把它们代入等式 $\mathbf{x}'=\mathrm{H}\mathbf{x}$，我们得到 $\tilde{\mathbf{x}}'=\mathrm{T}'\mathrm{H}\mathrm{T}^{-1}\tilde{\mathbf{x}}$. 由此推出$\widetilde{\mathrm{H}}=\mathrm{T}'\mathrm{H}\mathrm{T}^{-1}$是点对应 $\tilde{\mathbf{x}}\leftrightarrow\tilde{\mathbf{x}}'$的变换矩阵. 因此，求把 $\mathbf{x}_i$ 映射到 $\mathbf{x}'_i$ 的变换的另一种方法是：

(1) 根据公式 $\tilde{\mathbf{x}}_i=\mathrm{T}\mathbf{x}_i$ 和 $\tilde{\mathbf{x}}'_i=\mathrm{T}'\mathbf{x}'_i$变换图像坐标.

(2) 由对应 $\tilde{\mathbf{x}}_i\leftrightarrow\tilde{\mathbf{x}}'_i$ 求变换$\widetilde{\mathrm{H}}$.

(3) 令 $\mathrm{H}=\mathrm{T}'^{-1}\widetilde{\mathrm{H}}\mathrm{T}$.

把由此得到的变换矩阵 H 应用于原始未变换的点对应 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i$. 如何选择变换 T 和 T′暂不作具体说明. 现在要决定的是该算法给出的结果是否与所用的变换 T 和 T′无关. 理想的情形是 T 和 T′应该至少是相似变换，因为图像坐标系的尺度、定向或原点的不同选择应该本质上不影响算法的结果.

下面我们将证明最小化几何误差的算法关于相似变换不变. 但不幸的是节 3.1 的 DLT 算法的结果却不是相似变换不变的. 补救办法是在应用 DLT 算法之前对数据进行归一化变换. 归一化变换将消除由任意选取图像坐标系的原点和尺度所产生的影响，这就意味着

组合算法对图像的相似变换不变. 合适的归一化变换将在今后讨论.

### 3.4.2 DLT 算法的非不变性

考虑对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 的集合和由 DLT 应用于该对应点集合所求得的矩阵 H. 再考虑相关对应 $\tilde{\mathbf{x}} \leftrightarrow \tilde{\mathbf{x}}'$ 的集合,其中 $\tilde{\mathbf{x}}_i = \mathrm{T}\mathbf{x}_i$ 和 $\tilde{\mathbf{x}}'_i = \mathrm{T}'\tilde{\mathbf{x}}'_i$,并定义 $\tilde{\mathrm{H}} = \mathrm{T}'\mathrm{H}\mathrm{T}^{-1}$. 依据节 3.4.1,这里要决定的问题是:

- **将 DLT 算法应用于对应集 $\tilde{\mathbf{x}}'_i \leftrightarrow \tilde{\mathbf{x}}'_i$ 会得到变换 $\tilde{\mathrm{H}}$ 吗?**

我们将使用如下记号:矩阵 $\mathrm{A}_i$ 是由点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 导出的 DLT 方程的矩阵 (3.3 – $p$54),而 A 是由 $\mathrm{A}_i$ 组成的 $2n \times 9$ 的矩阵. 类似地,由对应 $\tilde{\mathbf{x}}_i \leftrightarrow \tilde{\mathbf{x}}'_i$ 定义矩阵 $\tilde{\mathrm{A}}_i$,其中,$\tilde{\mathbf{x}}_i = \mathrm{T}\mathbf{x}_i$, $\tilde{\mathbf{x}}'_i = \mathrm{T}'\mathbf{x}'_i$, T 和 T′ 为射影变换.

**结论 3.3** **令 T′ 为具有缩放因子 $s$ 的相似变换,T 为任意的射影变换. 此外,假设 H 是任何 2D 单应并定义 $\tilde{\mathrm{H}} = \mathrm{T}'\mathrm{H}\mathrm{T}^{-1}$. 那么 $\|\tilde{\mathrm{A}}\tilde{\mathbf{h}}\| = s\|\mathrm{A}\mathbf{h}\|$,其中 $\mathbf{h}$ 和 $\tilde{\mathbf{h}}$ 为 H 和 $\tilde{\mathrm{H}}$ 的元素组成的矢量.**

**证明** 定义矢量 $\boldsymbol{\varepsilon}_i = \mathbf{x}'_i \times \mathrm{H}\mathbf{x}_i$. 注意 $\mathrm{A}_i\mathbf{h}$ 是由 $\boldsymbol{\varepsilon}_i$ 前两个元素组成的矢量. 类似地,$\tilde{\boldsymbol{\varepsilon}}_i$ 定义为变换量 $\tilde{\boldsymbol{\varepsilon}}_i = \tilde{\mathbf{x}}'_i \times \tilde{\mathrm{H}}\tilde{\mathbf{x}}_i$. 作如下计算:

$$
\begin{aligned}
\tilde{\boldsymbol{\varepsilon}}_i &= \tilde{\mathbf{x}}'_i \times \tilde{\mathrm{H}}\tilde{\mathbf{x}}_i = \mathrm{T}'\mathbf{x}'_i \times (\mathrm{T}'\mathrm{H}\mathrm{T}^{-1})\mathrm{T}\mathbf{x}_i \\
&= \mathrm{T}'\mathbf{x}'_i \times \mathrm{T}'\mathrm{H}\mathbf{x}_i = \mathrm{T}'^{*}(\mathbf{x}'_i \times \mathrm{H}\mathbf{x}_i) \\
&= \mathrm{T}'^{*}\boldsymbol{\varepsilon}_i
\end{aligned}
$$

其中 $\mathrm{T}'^{*}$ 是 T′ 的余因子矩阵,而倒数第二个等式由引理 A3.2($p$406)得到. 对于一个一般的变换 T,误差矢量 $\mathrm{A}_i\mathbf{h}$ 和 $\tilde{\mathrm{A}}_i\tilde{\mathbf{h}}$(即 $\boldsymbol{\varepsilon}_i$ 和 $\tilde{\boldsymbol{\varepsilon}}_i$ 前两个元素)不是简单的关联. 在 T′ 是相似变换的特殊情形,我们得到 $\mathrm{T}' = \begin{bmatrix} s\mathrm{R} & \mathbf{t} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}$,其中 R 是旋转矩阵,$\mathbf{t}$ 是一个位移,$s$ 是缩放因子. 在此情形下,我们看到 $\mathrm{T}'^{*} = s\begin{bmatrix} \mathrm{R} & \mathbf{0} \\ -\mathbf{t}^{\mathrm{T}}\mathrm{R} & s \end{bmatrix}$. 将 $\mathrm{T}'^{*}$ 作用于 $\boldsymbol{\varepsilon}_i$ 的前两个分量,可以看到:

$$
\tilde{\mathrm{A}}_i\tilde{\mathbf{h}} = (\tilde{\varepsilon}_{i1}, \tilde{\varepsilon}_{i2})^{\mathrm{T}} = s\mathrm{R}(\varepsilon_{i1}, \varepsilon_{i2})^{\mathrm{T}} = s\mathrm{R}\mathrm{A}_i\mathbf{h}.
$$

因为旋转不影响矢量范数,故 $\|\tilde{\mathrm{A}}\tilde{\mathbf{h}}\| = s\|\mathrm{A}\mathbf{h}\|$,得证. 该结论可以用代数误差表示成:

$$
d_{\mathrm{alg}}(\tilde{\mathbf{x}}'_i, \tilde{\mathrm{H}}\tilde{\mathbf{x}}_i) = s d_{\mathrm{alg}}(\mathbf{x}'_i, \mathrm{H}\mathbf{x}_i). \qquad \square
$$

这样一来,不计常数因子,H 和 $\tilde{\mathrm{H}}$ 之间存在给出相同误差的一一对应. 因此,表面上看,最小化代数误差的矩阵 H 和 $\tilde{\mathrm{H}}$ 由公式 $\tilde{\mathrm{H}} = \mathrm{T}'\mathrm{H}\mathrm{T}^{-1}$ 相关联,从而 H 可以由乘积 $\mathrm{T}'^{-1}\tilde{\mathrm{H}}\mathrm{T}$ 恢复. 然而这结论是**错误的**. 理由是:虽然这样定义的 H 和 $\tilde{\mathrm{H}}$ 给出同样的误差 $\varepsilon$,但是对解施加的约束条件 $\|\mathrm{H}\| = 1$ 不等价于条件 $\|\tilde{\mathrm{H}}\| = 1$. $\|\mathrm{H}\|$ 和 $\|\tilde{\mathrm{H}}\|$ 并不以任何简单的方式相关联. 因此,H 和 $\tilde{\mathrm{H}}$ 之间不存在给出同样的误差 $\varepsilon$ 并同时满足 $\|\mathrm{H}\| = \|\tilde{\mathrm{H}}\| = 1$ 的一一

对应. 具体地说,

$$\text{最小化} \sum_i d_{\text{alg}}(\mathbf{x}'_i, \mathrm{H}\mathbf{x}_i)^2 \text{ 并满足 } \|\mathrm{H}\| = 1$$

$$\Leftrightarrow \text{最小化} \sum_i d_{\text{alg}}(\tilde{\mathbf{x}}'_i, \widetilde{\mathrm{H}}\tilde{\mathbf{x}}_i)^2 \text{ 并满足 } \|\mathrm{H}\| = 1$$

$$\not\Leftrightarrow \text{最小化} \sum_i d_{\text{alg}}(\tilde{\mathbf{x}}'_i, \widetilde{\mathrm{H}}\tilde{\mathbf{x}}_i)^2 \text{ 并满足 } \|\widetilde{\mathrm{H}}\| = 1.$$

因此,该变换方法给出的变换矩阵的解不同. 这是 DLT 算法相当不遂人愿的特性,它的结果随坐标甚至仅仅坐标原点的改变而改变. 但是,如果使范数 $\|\mathrm{A}\mathbf{h}\|$ 最小化的约束在变换下不变,那么被计算的矩阵 H 和$\widetilde{\mathrm{H}}$可以按正确方式相关联. 使 H 为变换不变量的最小化条件的例子在本章末的练习中给出.

### 3.4.3 几何误差的不变性

现在来证明为求 H 而最小化的几何误差在相似(尺度变化的欧氏)变换下是不变的. 如前面一样,考虑一组点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$和一个变换矩阵 H. 同时定义相关的对应集合 $\tilde{\mathbf{x}}\leftrightarrow\tilde{\mathbf{x}}'$,其中 $\tilde{\mathbf{x}}=\mathrm{T}\mathbf{x}$ 和 $\tilde{\mathbf{x}}'=\mathrm{T}'\mathbf{x}'$,并令$\widetilde{\mathrm{H}}=\mathrm{T}'\mathrm{H}\mathrm{T}^{-1}$. 假定 T 和 T′表示 $\mathbb{P}^2$中的欧氏变换. 可以验证

$$d(\tilde{\mathbf{x}}', \widetilde{\mathrm{H}}\tilde{\mathbf{x}}) = d(\mathrm{T}'\mathbf{x}', \mathrm{T}'\mathrm{H}\mathrm{T}^{-1}\mathrm{T}\mathbf{x}) = d(\mathrm{T}'\mathbf{x}', \mathrm{T}'\mathrm{H}\mathbf{x}) = d(\mathbf{x}', \mathrm{H}\mathbf{x}),$$

最后一个等式成立的理由是欧氏距离在如 T′一样的欧氏变换下不变. 这表明如果 H 最小化对应集的几何误差,那么$\widetilde{\mathrm{H}}$最小化变换后的对应集的几何误差,因此,在欧氏变换下最小化几何误差不变.

在相似变换下,几何误差被乘以变换的缩放因子,从而最小化变换的对应方式与欧氏变换相同. 所以最小化几何误差在相似变换下不变.

### 3.4.4 归一化变换

如节 3.4.2 所指出,计算 2D 单应的 DLT 算法的结果与点的坐标系有关. 事实上其结果在图像的相似变换下不是不变的. 这就提出一个问题:在计算 2D 单应时是否有某些坐标系以某种方式优于其他的坐标系呢? 对此问题的回答是绝对**肯定**的. 本节将介绍一种数据归一化的方法,它包括图像坐标的平移和尺度缩放. 归一化必须在实施 DLT 之前进行. 之后再对结果进行适当的校正就能得到关于原坐标系的 H.

数据归一化不仅提高了结果的精度,它还提供第二个好处,即对初始数据归一化的算法将对任何尺度缩放和坐标原点的选择不变,原因是归一化步骤通过为测量数据选择有效的标准坐标系,预先消除了坐标变换的影响. 因此,由于代数最小化在一个固定的标准坐标系中进行,因而使 DLT 算法实际上关于相似变换不变.

**各向同性缩放** 归一化的第一步是对每幅图像中的坐标进行平移(每幅图像的平移不同)使点集的形心移至原点. 其次对坐标系进行缩放使得点 $\mathbf{x}=(x,y,w)^{\mathrm{T}}$ 中的 $x,y,w$ 总体上有一样的平均值. 对坐标方向,我们选择向各同性的,而不是不同的缩放因子,使一个点的 $x$ 和 $y$ 坐标等量缩放. 最后,我们选择缩放因子使点 $\mathbf{x}$ 到原点的平均距离等于$\sqrt{2}$. 这意味着"平均"点为$(1,1,1)^{\mathrm{T}}$. 概括起来,变换如下进行:

(1) 对点进行平移使其形心位于原点.

(2) 对点进行缩放使它们到原点的平均距离等于$\sqrt{2}$.

(3) 对两幅图像独立进行上述变换.

具有数据归一化的 DLT 算法的推荐版本在算法 3.2 中给出. 该算法带有数据归一化应该比基本的 DLT 算法 3.1($p55$)优先使用.

数据归一化 DLT 算法对结果的影响在图 3.4 中用图示的方法给出. 由此得出结论:数据归一化的结果明显较好. 选择这个例子的目的是为了使结果更直观. 然而,在点对应数目更多并且点分布更广时这种明显的优势仍然保留. 为此我们强调:

- **数据归一化在 DLT 算法中是实质性的,一定不要视它为可有可无的.**

数据归一化对于不太良定的问题——例如基本矩阵和三焦点张量的 DLT 计算显得尤为重要. 这些将在后续的章节中讨论.

目标

给定 $n\geqslant 4$ 组 2D 到 2D 的点对应$\{\mathbf{x}_i\leftrightarrow\mathbf{x}'_i\}$,确定 2D 单应矩阵 H 使得 $\mathbf{x}'_i=\mathrm{H}\mathbf{x}_i$.

算法

(1) **归一化 x:** 计算一个只包括位移和缩放的相似变换 T,将点 $\mathbf{x}_i$ 变到新的点集 $\tilde{\mathbf{x}}_i$,使得点 $\tilde{\mathbf{x}}_i$ 的形心位于原点$(0,0)^{\mathrm{T}}$,并且它们到原点的平均距离是$\sqrt{2}$.

(2) **归一化 x′:** 针对第二幅图像上的点,类似地计算一个相似变换 T′,将点 $\mathbf{x}'_i$ 变换到 $\tilde{\mathbf{x}}'_i$.

(3) **DLT:** 将算法 3.1($p55$)应用于对应 $\tilde{\mathbf{x}}_i\leftrightarrow\tilde{\mathbf{x}}'_i$ 求得单应 $\tilde{\mathrm{H}}$.

(4) **解除归一化:** 令 $\mathrm{H}=\mathrm{T}'^{-1}\tilde{\mathrm{H}}\mathrm{T}$.

算法 3.2 求 2D 单应的归一化 DLT

图 3.4 Monte Carlo 模拟(见节 4.3($p98$)2D 单应的计算)的结果. 5 点集(用大十字架表示)用于计算 2D 单应. 5 个点都分别(在无噪声的情况中)映射到具有同样坐标的点,因此 H 是恒等变换. 现在做 100 次试验,每次试验让每幅图像的每点加 0.1 像素的高斯噪声. (为便于参照,大的十字架跨了 4 像素)用 DLT 算法算出的映射 H 被用来把一个较远的点变换到第二幅图像. 该点的 100 次投射用小十字架给出并且把它们的散射矩阵计算出来的 95% 的结果的椭圆也给出. (a)数据没有归一化的结果,(b)归一化的结果. 最左和最右的参考点(未归一化)的坐标为(130,108)和(170,108).

**归一化与条件数** 归一化的影响与 DLT 方程组的条件数或准确地说与方程组矩阵 A 的第一个和倒数第二个奇异值的比率 $d_1/d_{n-1}$ 有关. 这一点在[Hartley-97a]中有更详细地研究. 现在,有充分理由说在精确数据和无限精度的算术运算条件下,结果与归一化变换无关. 但是,在有噪声存在时解将偏离其正确结果. 大条件数会放大这种偏离. 甚至在无限精确算术运算下也是如此——这不是舍入误差的原因.

**非各向同性缩放** 可以采用其他的缩放方法. 在非各向同性缩放中,点的形心和前面一样平移到原点. 经平移后点在原点附近形成云. 然后进行坐标尺度缩放使该点集的两主矩都等于1. 这样使该点集形成以原点为中心,半径为1的近似的对称圆云. [Hartley－97a]中的实验结果指出非各向同性缩放增加了额外的开销,结果并不比各向同性缩放有显著提高.

**无穷远附近的点的缩放** 考虑无穷远平面和图像间的单应估计. 如果投视方向足够倾斜,那么平面上很远的点在图像上仍可看到——如果地平线可视,即使点在无穷远(消影点)也可以看到. 此时取形心为原点来归一化无穷远平面点的坐标是无意义的,因为该形心的坐标可能很大或无定义. 对于这种情形的一种归一化方法在页83的练习(2)中加以考虑.

## 3.5 迭代最小化方法

本节介绍节3.2和节3.3中给出的各种几何代价函数的最小化方法. 最小化这些代价函数需要使用迭代技术. 这是不幸的,因为迭代技术与诸如归一化DLT算法3.2的线性算法相比有某些缺点:

(1) 它们比较慢.

(2) 它们一般在开始迭代时需要一个初始估计.

(3) 它们有不收敛、或收敛到局部最小而不是全局最小的风险.

(4) 选择迭代停止的标准可能是非常技巧性的.

总而言之,迭代技术通常需要更仔细的操作. 迭代最小化技术一般由五步组成:

(1) **代价函数** 代价函数是最小化的基础. 各种可能的代价函数已在节3.2中讨论.

(2) **参数化** 把要计算的变换(或其他实体)表示成有限数目的参数. 一般并不要求它是最小参数集,事实上超参数化常有优越性.(见下面的讨论.)

(3) **函数确定** 必须确定一个用参数集描述的代价函数.

(4) **初始化** 计算一个适当的初始参数估计. 一般将由一个线性算法(例如DLT算法)来实现.

(5) **迭代** 由初始解开始,在迭代中逐步修正参数以达到最小化代价函数的目的.

**关于参数化**

一个给定的代价函数通常有若干种参数化选择. 指导参数化的一般策略是选择能覆盖最小化的整个空间并能用一种方便的方式来计算代价函数. 例如H可以用9个参数来参数化——它是超参数化的,因为实际上它仅有8个自由度(全局尺度无关紧要). 一个**最小**参数化(即等于自由度的参数数)仅包括8个参数.

一般这种类型的最小化问题采用超参数化似乎不会产生坏影响,只要对于所有选择的参数,对应的目标是所要求的类型. 特别对于齐次目标(如这里的3×3的射影矩阵)通常没有必要或不鼓励追求消除尺度因子的多义性而采用最小参数化.

理由如下:**不必要**采用最小参数化是因为一个性能好的非线性最小化算法会"注意"不在冗余方向(例如矩阵缩放方向)移动. Gill和Murray[Gill－78]中所介绍的算法是Gauss－Newton方法的改进,它有丢弃参数冗余组合的有效策略. 类似地,Levenberg－Marquardt算法(见节A4.2($p$417))也能轻松地处理冗余参数化问题. 我们**不推荐**使用最小参

数化方法，经验证明使用最小化参数会使代价函数曲面变得更复杂. 因而，程序在局部最小处停止不前的可能性更大.

选择参数化的另一个考虑是限制变换为一个具体的类型. 例如，假定已知 H 是透射，那么如节 A5.2($p$427)所介绍它可以参数化为

$$\mathrm{H}=\mathrm{I}+(\mu-1)\frac{\mathbf{v}\mathbf{a}^{\mathsf{T}}}{\mathbf{v}^{\mathsf{T}}\mathbf{a}}$$

其中 $\mu$ 是标量，而 $\mathbf{v}$ 和 $\mathbf{a}$ 是 3 维矢量. 一个透射有 5 个自由度，对应于标量 $\mu$ 以及 $\mathbf{v}$ 和 $\mathbf{a}$ 的方向. 如果用 H 的 9 个矩阵元素来参数化，那么估计出来的 H 不太可能正好是一个透射. 但是如果 H 由 $\mu,\mathbf{v},\mathbf{a}$ 来参数化(共 7 个参数)，那么估计出来的 H 保证是一个透射. 该参数化与透射**一致**(它也是超参数化的). 在以后的章节中我们将会回过来讨论一致、局部、最小和超参数化问题.

**函数确定**

在节 3.2.7 中我们看到一般类型的估计问题与包含模型曲面 $S$ 的测量空间 $\mathbb{R}^N$ 有关. 给定了一个测量 $\mathbf{X}\in\mathbb{R}^N$，估计的任务是求在 $S$ 上离 $\mathbf{X}$ 最近的点. 在 $\mathbb{R}^N$ 上加上非均匀高斯误差分布时，**"最近"**这个词解释成按 Mahalanobis 距离最近. 迭代最小化方法将用这个估计模型来介绍. 在参数拟合的迭代估计中，模型曲面 $S$ 被局部参数化，而且允许参数变化以便最小化与测量点的距离. 更具体地说，

(1) 有一个协方差矩阵为 $\Sigma$ 的**测量矢量** $\mathbf{X}\in\mathbb{R}^N$.

(2) 一组参数被表达成一个矢量 $\mathbf{P}\in\mathbb{R}^M$.

(3) 定义一个映射 $f$: $\mathbb{R}^M\rightarrow\mathbb{R}^N$. 其值域(至少局部地)是 $\mathbb{R}^N$ 中表示容许测量集的模型曲面 $S$.

(4) 最小化的代价函数是 Mahalanobis 距离的平方

$$\|\mathbf{X}-f(\mathbf{P})\|_{\Sigma}^{2}=(\mathbf{X}-f(\mathbf{P}))^{\mathsf{T}}\Sigma^{-1}(\mathbf{X}-f(\mathbf{P})).$$

事实上，我们希望寻找一组参数 $\mathbf{P}$ 使 $f(\mathbf{P})=\mathbf{X}$，或(不成功时)使 $f(\mathbf{P})$ 在 Mahalanobis 距离意义下尽可能接近 $\mathbf{X}$. 当最小化代价函数属于此类型时，Levenberg－Marquardt 算法是最小化迭代的一般工具. 现在来讨论如何把本章介绍的各种不同类型的代价函数规范为上述模式.

**单图像误差** 这里，我们固定第一幅图像中点 $\mathbf{x}_i$ 的坐标，并变化 H 以最小化代价函数(3.6－$p$58)，即

$$\sum_i d(\mathbf{x}_i',\ \mathrm{H}\tilde{\mathbf{x}}_i)^2.$$

测量矢量 $\mathbf{X}$ 由点 $\mathbf{x}_i'$ 的 $2n$ 个非齐次坐标组成. 我们可以选单应矩阵 H 的元素组成的矢量 $\mathbf{h}$ 作为参数. 函数 $f$ 定义为

$$f:\mathbf{h}\mapsto(\mathrm{H}\mathbf{x}_1,\mathrm{H}\mathbf{x}_2,\ldots,\mathrm{H}\mathbf{x}_n).$$

这里和下面方程中的 $\mathrm{H}\mathbf{x}_i$ 都是非齐次坐标. 可以验证 $\|\mathbf{X}-f(\mathbf{h})\|^2$ 等于(3.6－$p$58).

**对称传递误差** 对称代价函数(3.7－$p$58)

$$\sum_i d(\mathbf{x}_i,\ \mathrm{H}^{-1}\mathbf{x}_i')^2+d(\mathbf{x}_i',\ \mathrm{H}\mathbf{x}_i)^2$$

中测量矢量 $\mathbf{X}$ 为由点 $\mathbf{x}_i$ 和点 $\mathbf{x}_i'$ 的非齐次坐标组成的 $4n$ 维矢量. 与上面一样，参数矢量是

由 H 的元素组成的矢量 $\mathbf{h}$. 函数 $f$ 由下式定义

$$f:\mathbf{h}\mapsto(\mathrm{H}^{-1}\mathbf{x}_1',\ldots,\mathrm{H}^{-1}\mathbf{x}_n',\ \mathrm{H}\mathbf{x}_1,\ldots,\mathrm{H}\mathbf{x}_n).$$

如前面一样,我们发现 $\|\mathbf{X}-f(\mathbf{h})\|^2$ 等于(3.7 - $p$58).

**重投影误差** 最小化代价函数(3.8 - $p$58)比较复杂. 其困难在于要求同时最小化点 $\hat{\mathbf{x}}_i$ 的所有选择和变换矩阵 H 的元素. 如果点对应太多,它会变成一个非常庞大的最小化问题. 该问题可以用点 $\hat{\mathbf{x}}_i$ 坐标和 $\hat{\mathrm{H}}$矩阵的元素——总共 $2n+9$ 个参数来参数化. 对 $\hat{\mathbf{x}}_i'$ 的坐标可不作要求,因为它们可通过 $\hat{\mathbf{x}}_i'=\hat{\mathrm{H}}\hat{\mathbf{x}}_i$ 由其他参数求出. 因此,参数矢量是 $\mathbf{P}=(\mathbf{h},\ \hat{\mathbf{x}}_1,\ldots,\ \hat{\mathbf{x}}_n)$. 测量矢量包含所有点 $\mathbf{x}_i$ 和 $\mathbf{x}_i'$ 的非齐次坐标. 函数 $f$ 由下式定义

$$f:(\mathbf{h},\ \hat{\mathbf{x}}_1,\ldots,\ \hat{\mathbf{x}}_n)\mapsto(\hat{\mathbf{x}}_1,\hat{\mathbf{x}}_1',\ldots,\hat{\mathbf{x}}_n,\ \hat{\mathbf{x}}_n')$$

其中 $\hat{\mathbf{x}}_i'=\hat{\mathrm{H}}\hat{\mathbf{x}}_i$. 可以验证 $\|\mathbf{X}-f(\mathbf{P})\|^2$(其中 $\mathbf{X}$ 是 $4n$ 维矢量)等于代价函数(3.8 - $p$58). 该代价函数必须在所有 $2n+9$ 个参数上最小化.

**Sampson 近似** 与重投影误差有 $2n+9$ 个参数不同,它在单图像误差(3.6 - $p$58)或对称转移误差(3.7 - $p$58)的误差最小化中仅需要在矩阵 H 的 9 个元素上最小化——通常是一个比较容易处理的问题. 重投影误差的 Sampson 近似使投影误差也能仅对 9 个参数最小化.

这是一个很重要的思路,因为一个 $m$ 参数的非线性最小化问题的迭代解,在使用诸如 Levenberg - Marquardt 方法时,每一次迭代涉及求一个 $m\times m$ 的线性方程组的解. 这是具有 $O(m^3)$复杂度的问题. 因此,应取小规模的 $m$.

Sampson 误差避免了重投影误差在 $2n+9$ 参数上最小化,因为它对每个 $\mathbf{h}$ 的具体选择能有效地确定 $2n$ 个变量$\{\hat{\mathbf{x}}_i\}$. 因而,仅需要最小化 $\mathbf{h}$ 的 9 个参数. 实际上,这种近似在误差相对于测量值比较小时给出相当好的结果.

**初始化**

可以用线性技术来求参数化的初始估计. 例如归一化的 DLT 算法 3.2 可直接提供 H 并由此求得 9 维矢量 $\mathbf{h}$,它可用来参数化迭代最小化. 一般如果有 $n\geqslant4$ 组对应,那么所有的对应都将用于线性解. 但是,在节 3.7 关于鲁棒估计的讨论中我们将看到,如果这些对应中包含野值时,应该采用通过慎重选择的对应作最小配置集(即四组对应). 线性技术或最小配置解是本书推荐的两个初始化方法.

有时还采用的另一种方法(例如见[Horn - 90,Horn - 91])是在参数空间进行足够密的采样,从每一个采样初始点开始迭代并保留最好结果. 这种方法仅在参数空间的维数足够小时才可采用. 参数空间的采样可以是随机的,或遵循某种模式. 另一个初始化方法是完全不要任何有效的初始化,在参数空间中的一个已知固定点开始迭代. 这种方法不常有效. 迭代很可能落入虚假的最小或不收敛. 即使在最好的情况中,初始点离最终解越远,所需的迭代数将越大. 鉴于此,使用好的初始化方法是最好的选择.

**迭代方法**

对选定的代价函数有各种最小化的迭代方法,其中最常用的有 Newton 迭代和 Levenberg - Marquardt方法. 这些方法在附录 4($p$416)中介绍. 其他适用于最小化代价函数的一般方法如 Powell 方法和单纯形方法都在[Press - 88]中介绍.

**小结** 本节的中心思想在算法 3.3 中汇总，它介绍用于估计(两幅图像的点对应之间的)单应映射的黄金标准和 Sampson 方法.

目标

给定 $n>4$ 组图像点对应 $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$，确定两图像之间单应的最大似然估计 $\hat{\mathrm{H}}$.

MLE 还包括最小化

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2$$

的附属点集 $\{\hat{\mathbf{x}}_i\}$，其中 $\hat{\mathbf{x}}'_i = \hat{\mathrm{H}}\hat{\mathbf{x}}_i$

算法

(1) **初始化**：计算一个初始估计 $\hat{\mathrm{H}}$ 以提供几何最小化的一个初始点. 例如，由四组点对应用线性归一化的 DLT 算法 3.2 或 RANSAC(节 3.7.1)计算 $\hat{\mathrm{H}}$.

(2) **几何最小化——或用 Sampson 误差**

- 最小化几何误差的 Sampsom 近似(3.12 - $p$61).
- 代价函数的最小化可以在 $\hat{\mathrm{H}}$ 的适当参数化下采用 Newton 算法(节 A4.1($p$416))或 Levenberg - Marquardt 算法(节 A4.2($p$417)). 例如该矩阵可以用它的 9 个元素来参数化.

**或用黄金标准误差：**

- 用测量点 $\{\mathbf{x}_i\}$ 或(最好)用这些点的 Sampson 矫正(3.11 - $p$61)来计算附属变量 $\{\hat{\mathbf{x}}_i\}$ 的初始估计.
- 在 $\hat{\mathrm{H}}$ 和 $\hat{\mathbf{x}}_i$ ($i=1,\ldots,n$)上最小化代价函数

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2$$

用 Levenberg - Marquardt 算法在 $2n+9$ 个变量上最小化代价函数，其中 $2n$ 对应于 $n$ 个 2D 点，而 9 对应于单应矩阵 $\hat{\mathrm{H}}$.

- 如果点的数目过大，那么建议采用节 A4.4($p$422)中的稀疏方法来最小化该代价函数.

**算法 3.3** 从图像对应估计 H 的黄金标准算法及其变化. 在 2D 单应计算中黄金标准算法比 Sampson 方法更可取.

## 3.6 算法的实验比较

我们用图 3.5 中的图像做算法比较. 表 3.1 给出本章介绍的几种算法测试的结果. 它给出了两对图像的剩余误差. 除个别例外，使用的方法一看自明. 方法"仿射"是指企图用一个最佳仿射映射拟合射影变换. "最佳"是指 ML 估计并假定噪声水平是一个像素.

第一对图像是图 3.5 的(a)和(b)，具有 55 组点对应. 看来几乎所有方法表现得一样好(仿射方法除外). 理论最佳残差大于所得到的结果，因为噪声水平(未知)小于一个像素.

图 3.5 的(c)由(a)重采样合成，第二对图是(a)和(c)取了 20 组点对应. 如表 3.1 所示，在这种情形下，几乎所有的方法都表现最佳. 例外的是仿射方法(预料会差，因为原本不是一个仿射变换)和非归一化的线性方法. 非归一化的方法预料会差(虽然也许不这么糟糕). 尚不了解为什么它在第一对图像中表现好而在第二对图像中表现很坏. 在任何情况下，最好都避免用这种方法，而应采用归一化线性或黄金标准方法.

a b c

图 3.5 用来比较由图像对应点计算射影变换方法的一张平面的三幅图像.

| 方法 | 第一对<br>图 3.5a&b | 第二对<br>图 3.5a&c |
|---|---|---|
| 线性归一化 | 0.4078 | 0.6602 |
| 黄金标准 | 0.4078 | 0.6602 |
| 线性非归一化 | 0.4080 | 26.2056 |
| 齐次缩放 | 0.5708 | 0.7421 |
| Sampson | 0.4077 | 0.6602 |
| 单视图误差 | 0.4077 | 0.6602 |
| 仿射 | 6.0095 | 2.8481 |
| 理论最佳 | 0.5477 | 0.6582 |

表 3.1 各种算法的以像素为单位的残差.

进一步的评估在图 3.6 中给出. 要估计的变换是把图中显示的棋盘图像映射到与该轴排成一线的方格栅. 如图所示,与方格栅相比图像严重失真. 为了试验,随机选择图像中的点与方格栅上的对应点匹配. 图 3.6 把(归一化的)DLT 和黄金算法与理论最小值或残差(见第 4 章)相比较. 注意当噪声达 5 个像素时,DLT 算法表现还好. 但当噪声达 10 个像素时就失败了. 注意在一个 200 像素的图像有 10 个像素的误差是非常高的了. 对接近于恒等映射的单映,DLT 的表现几乎与黄金标准算法一样好.

## 3.7 鲁棒估计

迄今,我们假定给出的对应集$\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$的唯一误差源来自点的位置测量,它服从高斯分布. 在许多实际运行中,因为点被错配而使这种假设无效. 错配点对高斯误差分布而言是**野值**. 这些野值会严重干扰单应的估计,从而应该加以识别. 因此,我们的目的是从给出的"对应"中确定**内点集**,使单应得以用上节介绍的算法在这些内点中用最佳方式来估计. 这就是**鲁棒估计**,因为这种估计对野值(测量服从一个不同的并可能非模型的误差分布)是鲁棒的(容忍的).

### 3.7.1 RANSAC

我们先给出一个容易可视化的简单例子——估计一组 2 维点的直线拟合. 它可以认为是对两直线上的对应点作 1 维仿射变换($x' = ax + b$)的估计.

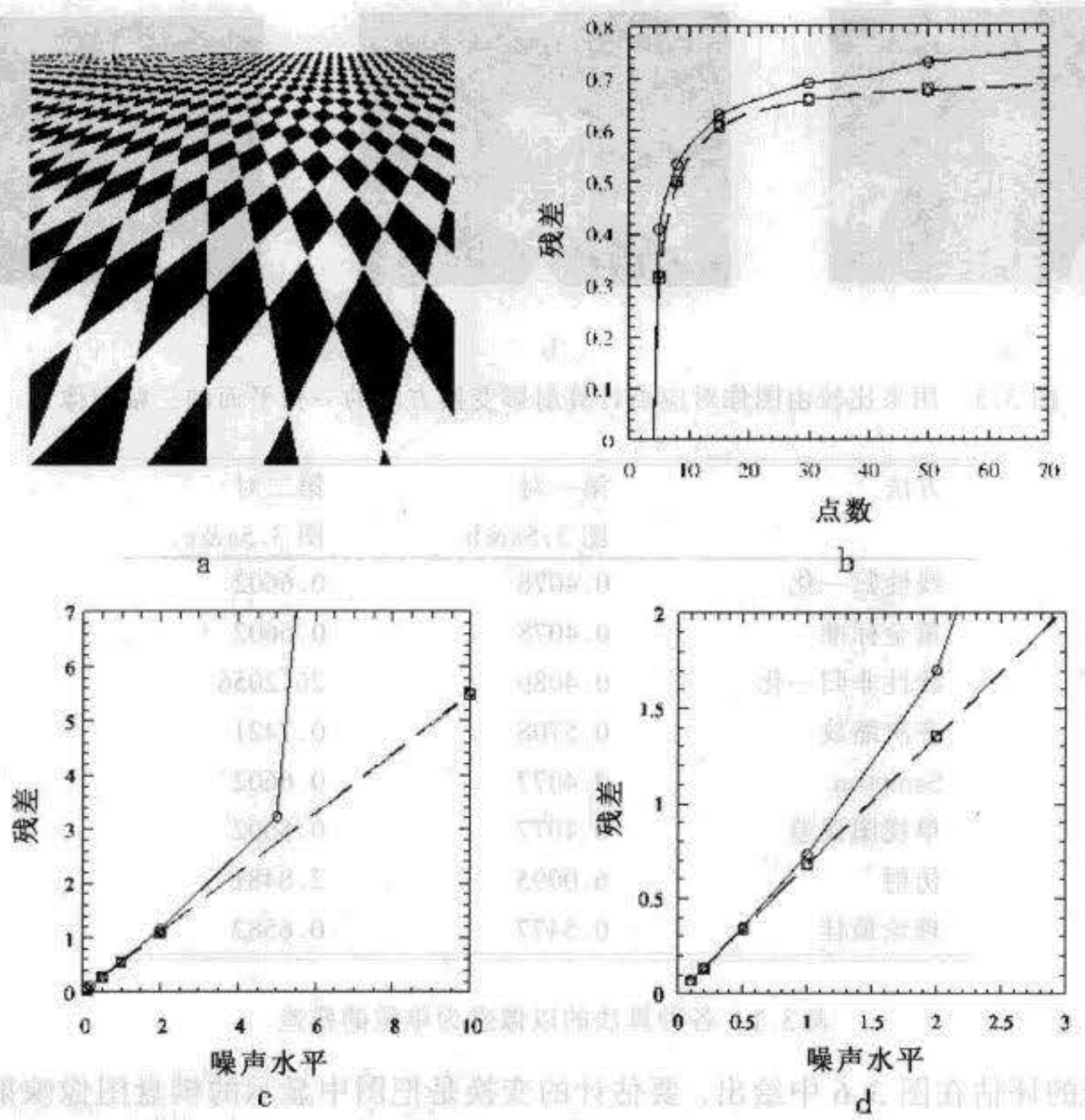



图 3.6 DLT 和黄金标准算法与理论最佳残差的比较. (a)在棋盘和它的图像之间计算单应. 在所有三幅图中,黄金标准算法的结果重合并与理论最小值没有区别. (b)作为点数函数的残差. (c)(d)分别给出噪声范围变化为 10 个点和 50 个点时的效应.

这个问题(见图 3.7a 的说明)可表述为:给定一组 2D 点的数据,寻找一条直线,它能最小化点到该直线的垂直距离的平方和(正交回归),并且所有有效点偏离该线的距离小于 $t$ 单位. 这实际是两个问题:用一条直线拟合数据并将数据分成内点(有效点)和野值. 阈值 $t$ 根据测量噪声设置(例如 $t=3\sigma$)并将在下文中讨论. 有许多类型的鲁棒算法,而具体如何选择在某种程度上依赖于野值所占的比例. 例如,如果已知只有一个野值,那么可以轮流删去每个点而用余下的点估计直线. 这里我们详细介绍一个通用的且非常成功的 Fischler 和 Bolles 的鲁棒估计算法——随机抽样一致算法 RANdom SAmple Consensus(RANSAC)[Fischler-81]. RANSAC 算法能够应付大比例的野值.

该算法思想非常简单:随机选择两点;这两点定义一条线. 该线的**支集**由在一定距离阈值内的点数来度量. 令这样的随机选择重复多次,具有最大支集的线就认为是鲁棒拟合. 在距离阈值内的点称为内点(并组成**一致集**). 直觉上,如果直线中某一点是野值那么该线将不会赢得大的支集,见图 3.7b.

对线的支集打分的另一个好处是有利于选取更好的拟合. 例如,图 3.7b 中的线 $\langle \mathbf{a},\mathbf{b}\rangle$ 的支集有 10 个点,而线 $\langle \mathbf{a},\mathbf{d}\rangle$(此时样本点相邻)的支集仅有 4 个点. 因此,虽然

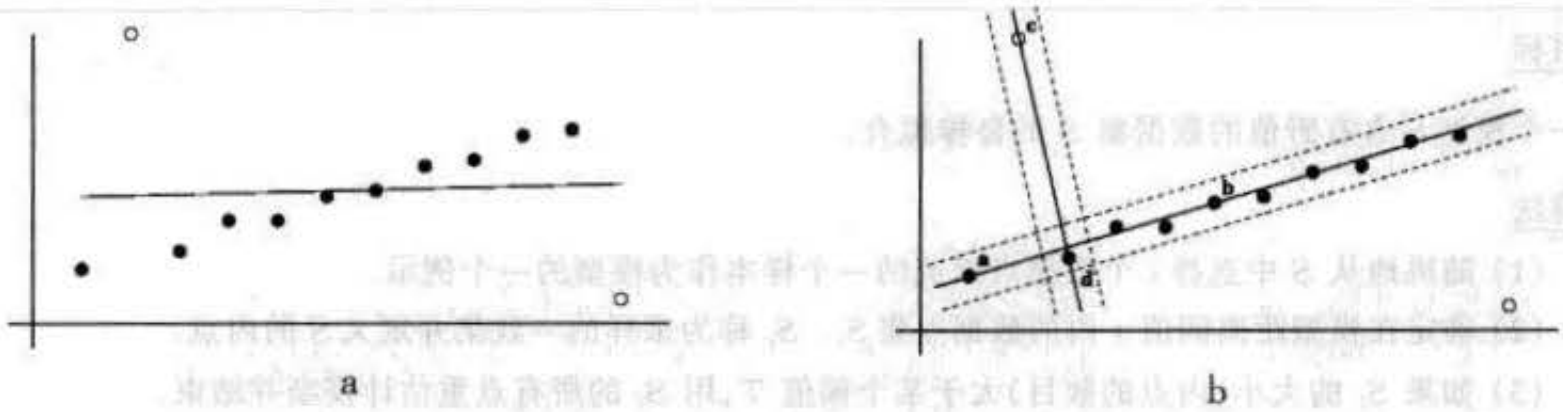



图 3.7 **线的鲁棒估计**. 实点表示内点,空点表示野值. (a)点数据的最小二乘(正交回归)拟合受野值严重影响. (b)在 RANSAC 算法中随机选择的点对组成的直线的支集由该线两侧不超过距离阈值内的点数来测量. 线<a,b>的支集是 10(点 a 和 b 都是内点),线<c,d>的支集是 2,其中点 c 是野值.

两个采样都不包括野值,按打分线<a,b>将被选中.

更一般地说,我们希望用**模型**(在这里是一条线)去拟合数据,随机**样本**包含足以确定模型的最小数据集(在这里是两点). 如果模型是平面单应,而数据是 2D 点对应集,那么最小子集包含四组对应. 应用 RANSAC 来估计单应在下文中介绍.

如 Fischler 和 Bolles[Fischler - 81]所阐述:"RANSAC 过程与通常的平滑技术相反:不是用尽可能多的点去获得一个初始解并在以后消除无效点,RANSAC 使用满足可行条件的尽量少的初始数据集并在可能时用一致性数据集扩大它.

RANSAC 算法概括在算法 3.4 中. 随之产生的三个问题是:

1. **什么是距离阈值**? 我们希望选择的距离阈值 $t$ 使点为内点的概率是 $\alpha$. 该计算需要知道内点到模型的距离的概率分布. 在实际中距离阈值通常靠经验选取. 但是,如果假定测量误差为零均值和标准方差 $\sigma$ 的高斯分布,那么 $t$ 的值可以算出. 因此,点的距离的平方 $d_\perp^2$ 是高斯变量的平方和并服从一个自由度 $m$ 的 $\chi_m^2$ 分布,其中 $m$ 等于模型的余维度. 对于直线,余维度是 1——仅测量到直线的垂直距离. 如果模型是一个点,余维度是 2,距离的平方是 $x$ 和 $y$ 测量误差的平方和. 随机变量 $\chi_m^2$ 的值小于 $k^2$ 的概率由累积 $\chi^2$ 分布 $F_m(k^2) = \int_0^{k^2} \chi_m^2(\zeta)d\zeta$ 给出. 这两个分布在节 A2.2($p$403)中介绍. 由该累积分布可知

$$\begin{cases} \text{内点} & d_\perp^2 < t^2 \\ \text{野值} & d_\perp^2 \geqslant t^2 \end{cases} \quad \text{且 } t^2 = F_m^{-1}(\alpha)\sigma^2 \tag{3.17}$$

通常 $\alpha$ 取 0.95,即点为内点的概率是 95%. 它表明内点被错误排斥的概率仅是次数的 5%. 本书研究的各有关模型在取 $\alpha = 0.95$ 时,$t$ 的值在表 3.2 中列出.

| 余维度 $m$ | 模型 | $t^2$ |
|---|---|---|
| 1 | 直线,基本矩阵 | $3.84\sigma^2$ |
| 2 | 单应,摄像机矩阵 | $5.99\sigma^2$ |
| 3 | 三焦点张量 | $7.81\sigma^2$ |

表 3.2 点(对应)是内点的概率为 $\alpha = 0.95$ 时,距离阈值 $t^2 = F_m^{-1}(\alpha)\sigma^2$.

目标

一个模型与含有野值的数据集 $S$ 的鲁棒拟合.

算法

(1) 随机地从 $S$ 中选择 $s$ 个数据点组成的一个样本作为模型的一个例示.

(2) 确定在模型距离阈值 $t$ 内的数据点集 $S_i$. $S_i$ 称为采样的一致集并定义 $S$ 的内点.

(3) 如果 $S_i$ 的大小(内点的数目)大于某个阈值 $T$,用 $S_i$ 的所有点重估计模型并结束.

(4) 如果 $S_i$ 的大小小于 $T$,选择一个新的子集并重复上面的过程.

(5) 经过 $N$ 次试验选择最大一致集 $S_i$,并用 $S_i$ 的所有点重估计模型.

算法 3.4 RANSAC 鲁棒估计算法,取自[Fischler-81]. 以最小所需要的 $s$ 个数据点作为模型自由参数的示例. 算法中三个阈值 $t$、$T$ 和 $N$ 在正文中讨论.

**2. 采样多少次为宜?** 尝试每个可能的样本通常在计算上不可行也不必要. 其实只要采样次数 $N$ 足够大,以保证由 $s$ 个点组成的随机样本中至少有一次没有野值的概率为 $p$. 通常 $p$ 取为 0.99. 假定 $w$ 是任意选择的数据点为内点的概率,那么 $\varepsilon = 1 - w$ 是其为野值的概率. 那么至少需要 $N$ 次选择(每次 $s$ 个点),其中 $(1-w^s)^N = 1-p$,从而

$$N = \log(1-p)/\log(1-(1-\varepsilon)^S). \tag{3.18}$$

给定 $s$ 和 $\varepsilon$,表 3.3 给出当 $p=0.99$ 时的 $N$ 值.

| 样本大小 | 野值 $\varepsilon$ 的比例 | | | | | | |
|---|---|---|---|---|---|---|---|
| $s$ | 5% | 10% | 20% | 25% | 30% | 40% | 50% |
| 2 | 2 | 3 | 5 | 6 | 7 | 11 | 17 |
| 3 | 3 | 4 | 7 | 9 | 11 | 19 | 35 |
| 4 | 3 | 5 | 9 | 13 | 17 | 34 | 72 |
| 5 | 4 | 6 | 12 | 17 | 26 | 57 | 146 |
| 6 | 4 | 7 | 16 | 24 | 37 | 97 | 293 |
| 7 | 4 | 8 | 20 | 33 | 54 | 163 | 588 |
| 8 | 5 | 9 | 26 | 44 | 78 | 272 | 1177 |

表 3.3 给定样本大小 $s$ 和野值的比例 $\varepsilon$ 并保证至少有一次没有野值的概率是 $p=0.99$ 时所需的采样数 $N$.

**例 3.4** 在图 3.7 的直线拟合问题中共有 $n=12$ 个数据点,其中两个是野值,从而 $\varepsilon = 2/12 = 1/6$. 由表 3.3 可知,对 $s=2$ 的最小子集,至少需要 $N=5$ 次采样. 如果把它与穷尽每一点对的工作量做比较,后者需要 $\binom{12}{2}=66$ 个样本(记号 $\binom{n}{2}$ 表示由 $n$ 中取 2 个的组合数,$\binom{n}{2} = n(n-1)/2$). △

**注意**

(1) 采样次数与野值所占比例而不与其数目相关. 这表明需要的采样次数可能小于野值数. 因此采样的计算代价即使在野值数目大时也能接受.

(2) (对于给定的 $\varepsilon$ 和 $p$)样本数随最小子集的增大而增大. 也许有人认为采用大于最小子集的子集会有好处,例如在直线拟合时取三个或更多的点,因为这样会获得关

于直线的更好的估计，并且测量的支集会更精确地反映真正的支集．然而由于增加采样数而导致增加的计算代价一般会大大超过给测量支集带来的可能的好处．

3．**一致集多大为宜**？根据经验，在给定野值的假定比例后，如果一致集大小接近期望属于该数据集的内点数时迭代就停止，即对 $n$ 个数据，$T=(1-\varepsilon)n$．对于图 3.7 中直线拟合的例子，$\varepsilon$ 的保守估计是 $\varepsilon=0.2$，因此 $T=(1.0-0.2)12=10$．

**自适应地决定采样次数**　通常，数据中野值所占比例 $\varepsilon$ 是未知的．对此情形，算法从 $\varepsilon$ 的最坏估计开始，当发现更大的一致集时就把原估计更新．例如，如果最坏的估计是 $\varepsilon=0.5$，但一旦发现内点的一致集占数据的 80%，那么估计更新为 $\varepsilon=0.2$．

通过一致集"探察"数据的思想可以重复应用以便自适应地确定采样次数 $N$．仍以上面的例子来说明，根据(3.18)，由 $\varepsilon=0.5$ 的最坏估计确定一个初始 $N$．当发现一个一致性集包含大于 50% 的数据时，那么我们知道内点至少是这个比例．由(3.18) 用更新后的 $\varepsilon$ 把 $N$ 变小．对于每个样本，一旦发现其一致性集的 $\varepsilon$ 低于当前的估计就重复更新过程，$N$ 就再一次减少．一旦完成 $N$ 次采样，算法就终止．可能会出现这样的一个样本，由它的 $\varepsilon$ 所确定的 $N$ 小于已经执行的采样次数，这说明已经执行了足够多次的采样，因而算法可以终止．自适应地计算 $N$ 的伪码在算法 3.5 中给出．

这种自适应方法效果相当好并实际解决了采样次数和终止算法的问题．$\varepsilon$ 的初始值可以取为 1.0，此时 $N$ 的初始值是无穷大．聪明的做法是在(3.18)中取一个保守的概率 $p$，例如 0.99．第 82 页上的表 3.4 给出计算单应时所用的 $\varepsilon$ 和 $N$ 的例子．

- $N=\infty$，sample _ count = 0
- 当 $N>$sample _ count 重复
  - 选取一个样本并计算内点数
  - 令 $\varepsilon=1-$(内点数)/(总点数)
  - 取 $p=0.99$ 并由 $\varepsilon$ 及(3.18)求 $N$
  - sample _ count 加 1
- 终止

算法 3.5　确定 RANSAC 样本次数的自适应算法．

## 3.7.2　鲁棒最大似然估计

RANSAC 算法把数据划分为内点(最大一致性集)和野值(数据集的余下部分)两个不相交子集，同时给出该模型的估计 $M_0$，它由具有最大支集的最小集算出．RANSAC 算法的最后一步是用所有的内点重新估计模型．该重新估计应该是最优的并涉及节 3.3 所介绍的最小化一个 ML 代价函数．在直线的情形，ML 估计等价于正交回归并存在封闭形式的解．但是 LM 估计一般涉及迭代最小化，并以最小集估计的 $M_0$ 为初始点．

这个常被采用的过程的唯一缺点是内点－野值的分类是不能取消的．当模型已经与一致集最佳拟合后，如果把距离阈值应用于该新模型，很有可能又有一些点成为内点．例如图 3.8 中的$\langle \mathbf{A},\mathbf{B}\rangle$由 RANSAC 得到．该线的支集由都是内点的四点组成．当最佳拟合这四点后，有 10 个点应该重新划为内点．这两步是：最佳地拟合内点；用(3.17)重新分类内点；然后才可以再迭代直到内点数收敛．根据内点到模型的距离来加权的最小二乘的拟合方法

经常在此步骤中使用.

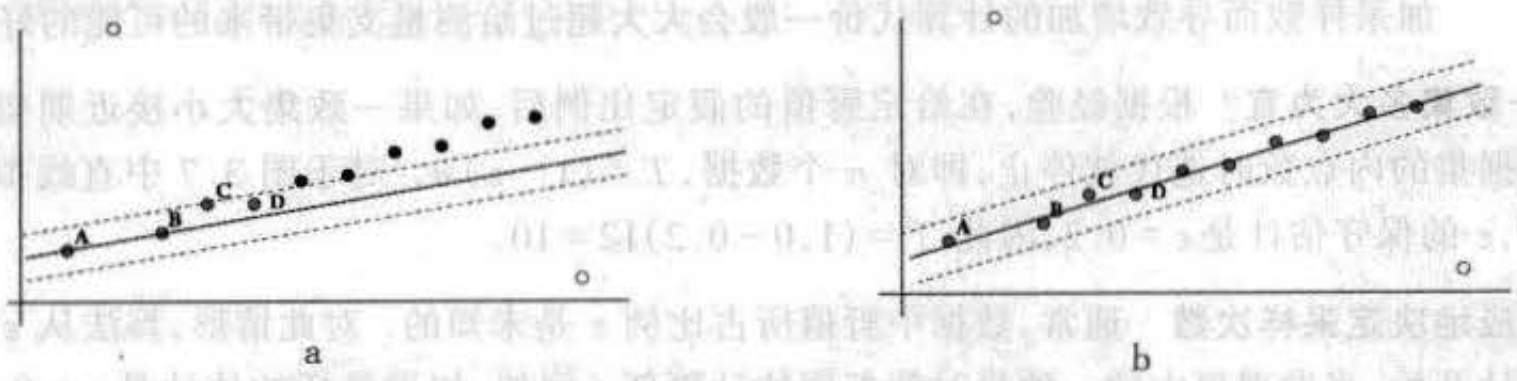



图 3.8 **鲁棒 ML 估计**. 灰色点被分类为线的内点. (a)由点〈**A**,**B**〉定义的直线的支集由 4 点组成({**A**,**B**,**C**,**D**}). (b)ML 直线拟合(正交最小二乘)这四点. 该拟合比由〈**A**,**B**〉定义的直线有很大的改进,它使 10 点被分类为内点.

**鲁棒代价函数** 不同于仅在内点上最小化 $\mathcal{C}=\sum_i d_{\perp i}^2$,另一种改进的鲁棒方法是在所有数据上最小化. 一个合适的鲁棒代价函数是

$$\mathcal{D}=\sum_i \gamma(d_{\perp i}) \text{ 且 } \gamma(e)=\begin{cases} e^2 & e^2<t^2 \quad \text{内点} \\ t^2 & e^2 \geqslant t^2 \quad \text{野值} \end{cases} \tag{3.19}$$

这里 $d_{\perp i}$ 是点的误差而 $\gamma(e)$ 是一个鲁棒代价函数[Huber-81],其中给野值赋一个固定的代价. 同(3.17)一样,阈值用 $\chi^2$ 确定,其中 $t^2$ 是定义的. 如节 3.3 所介绍由高斯误差模型产生内点的平方代价. 在鲁棒代价函数中,赋野值为常数的理由是假定野值服从一个扩散或均匀分布,这些分布的对数似然是一个常数. 也许有人会认为野值可以通过直接对 $d_{\perp i}$ 取阈值而从代价函数中除去. 这样单独取阈值所产生的问题是它将导致最后的结果仅包括野值,因为它们没有代价.

代价函数 $\mathcal{D}$ 能使最小化考虑到所有的点,不论它们是野值或内点. 在最小化迭代开始时 $\mathcal{D}$ 与 $\mathcal{C}$ 仅用一个常数相区别(用野值数的 4 倍赋值). 但是,随着最小化的进行,野值可能重新被指定为内点,而且这种情况通常会在实际中发生.

### 3.7.3 其他鲁棒算法

在 RANSAC 中,由最小集推测的模型根据在阈值距离内的点数来记分. 另一种模型记分的方法是根据数据中所有点的距离中值. 选择具有最小中值的模型. 这就是最小中值平方(LMS)估计,这里,如 RANSAC 一样,随机选择的最小样本子集由(3.18)获得采样次数. LMS 的优点在于它**不需要**阈值或误差方差的**先验知识**. LMS 的缺点是如果多于一半的数据是野值那么它要失败,因为距离的中值可能是一个野值. 解决的办法是用野值的比例来确定所选的距离. 例如有 50%野值时可取低于中值(统计学的说法)的阈值.

RANSAC 和 LMS 算法都能处理野值占大比例的情况. 如果野值的数量少,那么其他鲁棒方法也许会更有效. 其中包括删点方法,即每点轮流被删除而用模型对剩下的数据拟合;和加权最小二乘方迭代,其中一个数据点对拟合的影响以它的残差来反加权. 我们一般不推荐这些方法. Torr[Torr-95a],Xu 及 Zhang[Xu-96]介绍并比较了各种估计基本矩阵的鲁棒算法.

## 3.8 单应的自动计算

本节介绍自动计算两幅图像间的单应的算法. 算法的输入仅仅是图像,不需要其他先验信息;而输出是单应的估计以及一组对应的兴趣点. 该算法可以应用于诸如一张平坦表面的两幅图像或由摄像机绕其光心旋转得到的两幅图像.

算法的第一步是计算每幅图像上的兴趣点. 我们面临一个类似"先有鸡或先有蛋"的问题:一旦兴趣点间对应被建立,单应便可算出;反之给定单应,兴趣点间的对应便不难建立. 这个问题将用鲁棒估计解决,这里用 RANSAC 作为"搜索引擎". 该思想是先用某种方法获得假设的点对应. 可以预料到这些对应的一部分是错配. RANSAC 就是针对处理这种情形设计的——估计单应以及与该估计一致的一组内点(真对应)和野值(错配).

该算法归纳在算法 3.6 中,并在图 3.9 给出应用的例子,而且下文对它的步骤还有更详细的介绍. 可用本质上相同的算法直接由两幅或三幅输入图像自动地计算基本矩阵和三焦点张量. 第 10 章和第 15 章将对这种计算作介绍.

目标

计算两幅图像间的 2D 单应.

算法

(1)**兴趣点**:在每一幅图像上计算兴趣点.

(2)**假设对应**:根据兴趣点灰度邻域的接近和相似,计算它们的匹配集.

(3) RANSAC **鲁棒估计**:重复 $N$ 次采样,这里 $N$ 按算法 3.5 用自适应方法确定:

(a) 选择由四组对应组成的一个随机样本并计算单应 H.

(b) 对假设的每组对应,计算距离 $d_\perp$.

(c) 根据 $d_\perp < t = \sqrt{5.99}\sigma$ 像素确定对应数,进而计算与 H 一致的内点数.

选择具有最大内点数的 H. 当数目相等时选择内点的标准方差最低的 H.

(4)**最优估计**:由划定为内点的所有对应重新估计 H,其中用节 A4.2($p$417)中的 Levenberg - Marquardt 算法来最小化 ML 代价函数(3.8 - $p$58).

(5) **引导匹配**:用估计的 H 去定义转移点位置附近的搜索区域,进一步确定兴趣点的对应.

最后两步可以重复直到对应的数目稳定为止.

算法 3.6 用 RANSAC 自动估计两幅图像之间的单应

**确定假设对应** 算法的目标是在不知道单应的条件下提供初始的点对应集. 对应的大部分应该是正确的,但算法的目标并不是完全匹配,因为此后用 RANSAC 可消除错配. 这些对应可以想象成"种子"对应. 这些假设对应由在每一幅图像上独立地检测兴趣点而获得,然后用邻域灰度的近似度和相似度的组合来匹配这些兴趣点. 为简洁起见,兴趣点将称为'角点'. 但是,这些角点不一定是景物中物理角点的图像. 角点定义为图像自相关函数的最小值.

对图像 1 中的每个角点$(x,y)^\mathsf{T}$,在图像 2 中以$(x,y)^\mathsf{T}$为中心的方形区域中搜索具有最高邻域互相关的匹配. 对称地,对图像 2 中的每个角点在图像 1 中搜索其匹配. 偶尔会存在冲突,一幅图像上的一个角点被另一幅上不止一个角点所"认定". 在此情形采用"赢者取

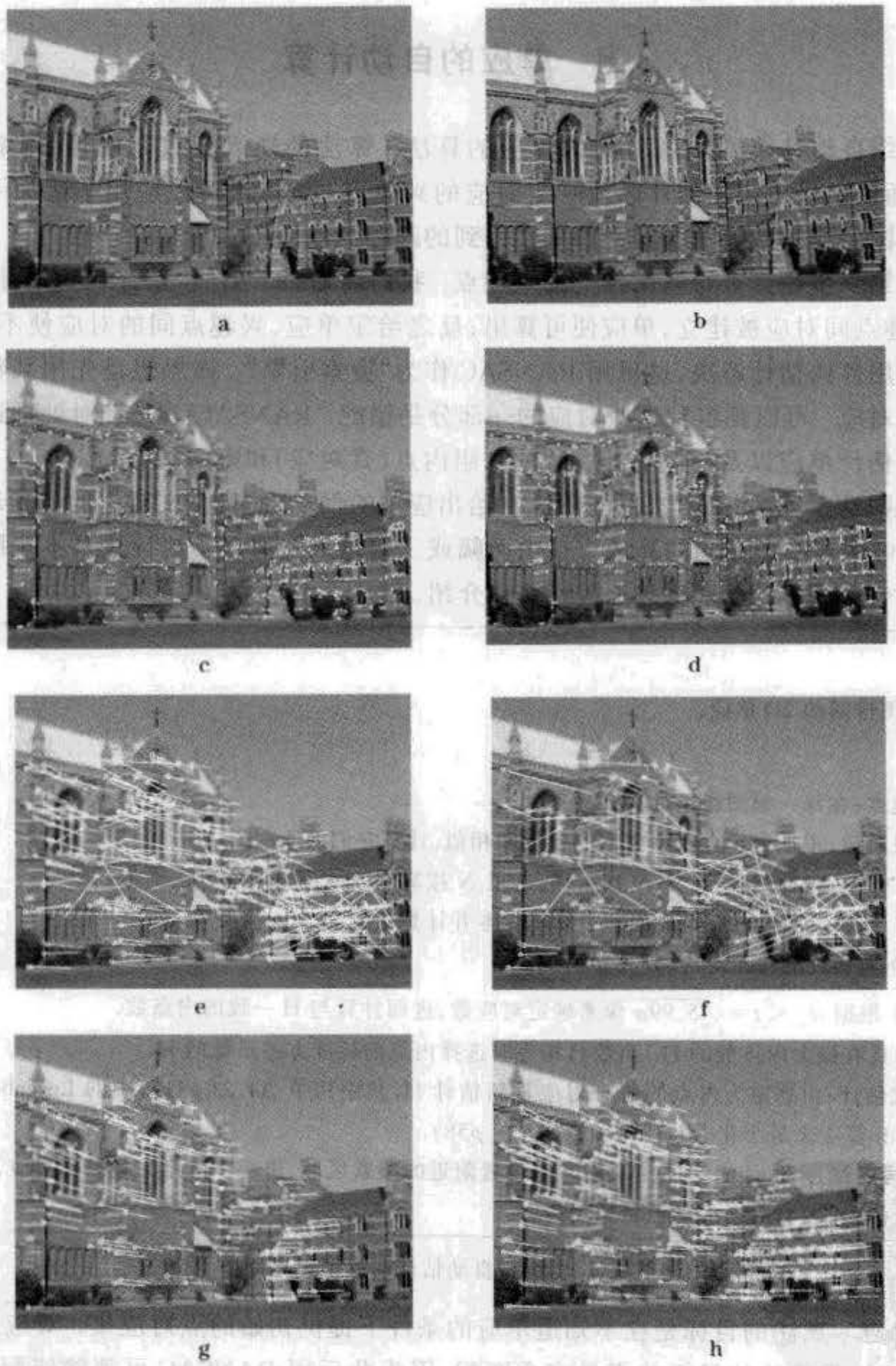



图 3.9 **用 RANSAC 自动计算两幅图像间的单应**. 视图间的运动是绕摄像机中心旋转,因此图像之间精确地由一个单应相联系. (a)(b)牛津大学 Keble 学院的左图和右图. 图像是 640×480 像素. (c)(d)将检测到的角点叠加在图像上. 每幅图像上有大约 500 个角点. 下面的结果仅仅叠加在左图上. (e) 268 假设匹配用连接角点的线表示,注意明显的错配;(f) 野值——假设匹配中有 117 个;(g) 内点——与估计 H 一致的有 151 组对应;(h)经过引导匹配和 MLE 后,最后的 262 组对应的集合.

1 败者取 0"策略因而仅保留最高互相关的匹配.

相似测量的另一种方法是用灰度差的平方和(简记为 SSD)代替(归一化的)互相关(简

记为 CC). CC 对灰度值的仿射映射不变(即 $I \mapsto \alpha I + \beta$ 缩放加偏置),实际中仿射映射经常在图像之间发生. 然而 SSD 对此映射不是不变的. 但在图像之间的灰度值变化不大时,通常倾向于用 SSD,因为它的测量比 CC 灵敏而且计算代价较小.

**用 RANSAC 求单应** 把 RANSAC 算法应用于假设对应集,以求得单应估计和与此估计相一致的(内点)对应. 样本大小是 4,因为四组对应确定一个单应. 如算法 3.5 所介绍,采样次数由每个一致集中野值所占比例而自适应性地设置.

这里有两个要讨论的问题:此时"距离"是什么? 以及怎样选择样本?

(1) **距离测量**:通过单应 H 估计一组对应的误差的最简单方法是采用对称转移误差,即 $d^2_{\text{transfer}} = d(\mathbf{x}, \mathrm{H}^{-1}\mathbf{x}')^2 + d(\mathbf{x}', \mathrm{H}\mathbf{x})^2$,其中 $\mathbf{x} \leftrightarrow \mathbf{x}'$ 是点对应. 一个更好但开销更大的距离测量是重投影误差 $d^2_{\perp} = d(\mathbf{x}, \hat{\mathbf{x}})^2 + d(\mathbf{x}', \hat{\mathbf{x}}')^2$,其中 $\hat{\mathbf{x}}' = \mathrm{H}\hat{\mathbf{x}}$ 是完全对应. 这种测量开销大的原因是必须计算 $\hat{\mathbf{x}}$. 另一种替代方法是采用 Sampson 误差.

(2) **样本选择**:这里有两个议题. 第一,退化的样本应该丢弃. 例如,如果四点中有三点共线那么单应就得不到;第二,组成样本的点应该在整个图像上有合理的空间分布. 这是由于外插问题的缘故——在计算点所跨越的区域内被估计的单应将准确地映射,但一般,准确性将随离该区域距离加大而变坏(想像四点位于图像上方拐角的情形). 空间的分布采样可以这样来实施:划分图像并通过随机采样的适当加权来保证属于不同小区的点比同属于一个小区的点有较大可能性进入样本.

**鲁棒 ML 估计与引导匹配** 最后一步的目标是双重的:第一,用所有估计得到的内点(而不仅仅用样本的四点)来改进单应估计;第二,从假设对应集获得更多的内点匹配,因为求得了更准确的单应. 再通过最小化 ML 代价函数从内点中计算进一步改进的单应估计. 最后一步可以用两种方法实现. 一种方法是在内点上执行一个 ML 估计,然后用新的估计 H 重新计算内点,并且重复这一循环直到内点数目收敛. ML 代价函数的最小化采用节 A4.2 ($p$418)介绍的 Levengerg - Marquart 算法. 另一种方法是采用节 3.7.2 介绍的最小化(3.19)的鲁棒 ML 代价函数同时估计单应和内点. 同时估计的方法的缺点是在最小化代价函数的计算方面要花很大功夫. 因为这样的原因,循环的方法通常更具有吸引力.

### 3.8.1 应用域

通常要求算法能够在图像上相当均匀地恢复兴趣点,那么反过来则要求景物和分辨率支持这一要求. 景物应该有纹理——空白墙的图像是不理想的.

为搜索窗口而设置的近似约束给出了视图之间角点图像运动的上限(**视差**). 但是,如果不用此约束,算法不会失败,事实上近似约束的主要作用是降低计算复杂性,因为搜索窗口越小要评估的角点匹配越少.

最终,算法应用范围受限于角点邻域相似测量(SSD 或 CC)的成功,即提供对应之间的非二义性. 失败一般由于缺乏空间不变性:测量仅对图像位移不变,对超出此范围的变换会产生严重失真,例如图像旋转或透视压缩等. 一种解决办法是对图像之间的单应映射采用更大不变性的测量,例如采用旋转不变的测量. 另一种解决办法是用得到的初始估计的单应在灰度邻域之间建立映射. 详细内容已超出我们讨论的范围,但在[Pritchett - 98, Schmid - 98]中有讨论. 采用鲁棒估计在相当程度上增强了对非相关运动、阴影变化、部分遮挡等

的免疫能力.

### 3.8.2 实现与运行细节

兴趣点用 Harris[Harris－88]角点检测器获得. 检测器把角点定位到亚像素精度,并且经验发现对应误差通常小于一个像素[Schmid－98].

在算法的假设对应阶段,在求种子对应时,用于邻域相似性测量的阈值应该尽量保守,以便最小化不正确匹配(SSD 阈值为 20). 在引导匹配阶段,该阈值应放松(加倍)使更多的假设对应有效.

在图 3.9 的例中,图像是 640×480 像素,搜索窗口是±320 像素即整个图像. 当然在给定点的实际误差后可以采用一个小得多的窗口. 通常在视频系列中±40 像素的搜索窗口就足够了(即中心在当前位置而边长为 80 的方块),内点阈值是 $t=1.25$ 像素.

此例总共需要 43 次采样,采样运行如表 3.4 所示. 引导匹配需要 MLE－内点分类循环的两次迭代. $d_\perp$ 像素误差的 RMS 值在 MLE 之前是 0.23 而之后是 0.19. Levenberg－Marquardt算法需要 10 次迭代.

| 内点数 | $1-\epsilon$ | 自适应的 $N$ |
|---|---|---|
| 6 | 2% | 20,028,244 |
| 10 | 3% | 2,595,658 |
| 44 | 16% | 6,922 |
| 58 | 21% | 2,291 |
| 73 | 26% | 911 |
| 151 | 56% | 43 |

表 3.4 用 RANSAC 的自适应算法 3.5 计算图 3.9 的单应的结果. $N$ 是算法运行时所需采样的总数,条件是样本中没有野值的概率是 $p=0.99$. 算法在 43 次抽样后终止.

## 3.9 结束语

本章已对在估计表征多视图关系的张量中的有关问题和技术作了说明. 这些思想将在本书的所有计算章节中重复出现. 每次都涉及所需最小对应数、应该避免的退化几何配置、当多于最小对应数时可用来最小化的代数和几何误差以及对张量施加内部约束的参数化等问题.

### 3.9.1 文献

DLT 算法至少可追述到 Sutherland[Sutherland－63]. Sampson 关于二次曲线的拟合(是经典的 Bookstein 算法的改进)的经典文章出现在[Sampson－82]. 归一化由 Hartley[Hartley－97a]发表在计算机视觉文献中.

关于数值方法可以在 **Numerical Recipes in** C[Press－88]中找到,而关于迭代最小化则见 Gill 和 Murray 的[Gill－78].

Fischler 和 Boll[Fischler－81]的 RANSAC 是最早出现的一种鲁棒算法,并且事实上是为了解决计算机视觉的问题(由 3 点求姿态)而推导的. 该原始论文论证得非常清楚因而非

常值得一读. 其他关于鲁棒估计的背景材料可以在 Rousseeuw[Rousseeuw－87]找到. 鲁棒估计在计算机视觉中主要的应用是估计基本矩阵(第 10 章),其中 Torr 和 Murray[Torr－93]用 RANSAC 而 Zhang 等[Zhang－95]用 LMS. 单应的自动 ML 估计由 Torr 和 Zisserman [Torr－98]给出.

### 3.9.2 注释和练习

(1) DLT **的变换不变性** 我们讨论在满足各种约束下最小化代数误差 $\|\mathbf{Ah}\|$ (见(3.5－$p$57))的 2D 单应计算. 证明如下情况:

(a) 如果在约束 $h_9 = H_{33} = 1$ 下, $\|\mathbf{Ah}\|$ 被最小化,那么其结果在缩放变化下不变但在坐标平移时**变化**.

(b) 如果约束改为 $H_{31}^2 + H_{32}^2 = 1$ 那么结果在相似变换下不变.

(c) **仿射情形**:$H_{31} = H_{32} = 0$;$H_{33} = 1$ 的约束仍然使上述结果成立.

(2) **缩放无界点集** 当点位于或接近平面的无穷远时,用本章给出的各向同性(或非各向同性)的缩放方式来归一化坐标是既无道理又不可行的,因为形心坐标和缩放因子为无穷大或接近无穷大. 一个似乎会给出好结果的方法是归一化点集 $\mathbf{x}_i = (x_i, y_i, w_i)^\mathsf{T}$ 使得

$$\sum_i x_i = \sum_i y_i = 0;\ \sum_i x_i^2 + y_i^2 = 2\sum_i w_i^2;\ x_i^2 + y_i^2 + w_i^2 = 1, \forall i,$$

注意这里出现的 $x_i$ 和 $y_i$ 是齐次坐标,并且此条件不再蕴含形心在原点. 研究实现这种归一化的方法,并评价它的性质.

(3) **图像坐标微分的表示式** 对映射 $\mathbf{x}' = (x', y', w')^\mathsf{T} = \mathrm{H}\mathbf{x}$,导出如下表示式(其中 $\tilde{\mathbf{x}}' = (\tilde{x}', \tilde{y}')^\mathsf{T} = (x'/w', y'/w')^\mathsf{T}$ 是图像点的非齐次坐标):

(a) 对 $\mathbf{x}$ 求导

$$\partial\tilde{\mathbf{x}}'/\partial\mathbf{x} = \frac{1}{w'}\begin{bmatrix}\mathbf{h}^{1\mathsf{T}} - \tilde{x}'\mathbf{h}^{3\mathsf{T}} \\ \mathbf{h}^{2\mathsf{T}} - \tilde{y}'\mathbf{h}^{3\mathsf{T}}\end{bmatrix} \tag{3.20}$$

其中 $\mathbf{h}^{j\mathsf{T}}$ 是 H 的第 $j$ 行.

(b) 对 H 求导

$$\partial\tilde{\mathbf{x}}'/\partial\mathbf{h} = \frac{1}{w'}\begin{bmatrix}\mathbf{x}^\mathsf{T} & 0 & -\tilde{x}'\mathbf{x}^\mathsf{T} \\ 0 & \mathbf{x}^\mathsf{T} & -\tilde{y}'\mathbf{x}^\mathsf{T}\end{bmatrix} \tag{3.21}$$

其中 $\mathbf{h}$ 由(3.2－$p$53)定义.

(4) **具有非各向同性误差分布的 Sampson 误差** 节 3.2.6($p$61)推导 Sampson 误差时假定被测量点 $\mathbf{X}$ 具有圆误差分布. 当点 $\mathbf{X} = (x, y, x', y')^\mathsf{T}$ 用协方差矩阵 $\Sigma_\mathbf{x}$ 测量时,就应该改用最小化 Mahalanobis 范数 $\|\boldsymbol{\delta}_\mathbf{x}\|^2_{\Sigma_\mathbf{x}} = \boldsymbol{\delta}_\mathbf{x}^\mathsf{T}\Sigma_\mathbf{x}^{-1}\boldsymbol{\delta}_\mathbf{x}$. 试证在这种情形下,对应于(3.11－$p$61)和(3.12－$p$61)的公式是

$$\boldsymbol{\delta}_\mathbf{x} = -\Sigma_\mathbf{x}\mathrm{J}^\mathsf{T}(\mathrm{J}\Sigma_\mathbf{x}\mathrm{J}^\mathsf{T})^{-1}\boldsymbol{\varepsilon} \tag{3.22}$$

和

$$\|\boldsymbol{\delta}_\mathbf{x}\|^2_{\Sigma_\mathbf{x}} = \boldsymbol{\varepsilon}^\mathsf{T}(\mathrm{J}\Sigma_\mathbf{x}\mathrm{J}^\mathsf{T})^{-1}\boldsymbol{\varepsilon}. \tag{3.23}$$

注意如果测量误差在两图像中是独立的,那么协方差矩阵 $\Sigma_\mathbf{x}$ 将是对应于两幅图像

的两个 $2\times2$ 的对角块组成的分块对角矩阵.

(5) Sampson **误差编程的提示** 在 2D 单应估计和事实上在本书考虑的其他类似的问题中,节 3.2.6($p61$)的代价函数 $\mathcal{C}_H(\mathbf{X})=A(\mathbf{X})\mathbf{h}$ 是 $\mathbf{X}$ 坐标的多重线性函数. 因此计算偏导数$\partial\mathcal{C}_H(\mathbf{X})/\partial\mathbf{X}$ 可以非常简单. 例如,求导公式

$$\partial\mathcal{C}_H(x,\ y,\ x',\ y')/\partial x=\mathcal{C}_H(x+1,\ y,\ x',\ y')-\mathcal{C}_H(x,\ y,\ x',\ y')$$

是**确切的**,而不是有限差分近似. 这使编程大为简便,即不必编写求导数的特殊程序——而只用计算 $\mathcal{C}_H(\mathbf{X})$的程序就够了. 记 $\mathbf{E}_i$ 为在第 $i$ 位置为 1 其余位置全为 0 的矢量,可以看到$\partial\mathcal{C}_H(\mathbf{X})/\partial X_i=\mathcal{C}_H(\mathbf{X}+\mathbf{E}_i)-\mathcal{C}_H(\mathbf{X})$并且:

$$JJ^T=\sum_i(\mathcal{C}_H(\mathbf{X}+\mathbf{E}_i)-\mathcal{C}_H(\mathbf{X}))(\mathcal{C}_H(\mathbf{X}+\mathbf{E}_i)-\mathcal{C}_H(\mathbf{X}))^T.$$

同样注意从公式 $JJ^T\boldsymbol{\lambda}=-\boldsymbol{\varepsilon}$ 直接求解 $\boldsymbol{\lambda}$ 比先求逆后用公式 $\boldsymbol{\lambda}=-(JJ^T)^{-1}\boldsymbol{\varepsilon}$ 求解的计算效率高.

(6) **最小化仿射变换的几何误差** 给定一组对应$(x_i,\ y_i)^T\leftrightarrow(x_i',\ y_i')^T$,求最小化几何误差 (3.8 - $p58$) 的仿射变换 $H_A$. 我们将着重推导基于 Sampson 近似(此时它是精确的而非近似的)的一个线性算法. 完整的方法概括在算法 3.7 中.

目标

给定 $n\geqslant4$ 组图像点对应$\{\mathbf{x}_i\leftrightarrow\mathbf{x}_i'\}$,确定最小化两幅图像的重投影误差(3.8 - $p58$)的仿射单应 $H_A$.

算法

(a) 把点表示为 2 维非齐次矢量. 对点 $\mathbf{x}_i$ 作平移 $\mathbf{t}$ 使它们的形心为原点. 按同样方式对 $\mathbf{x}_i'$ 作平移 $\mathbf{t}'$. 后续工作均在平移后的坐标上进行.

(b) 形成 $n\times4$ 矩阵 A,它的行矢量是

$$\mathbf{X}_i^T=(\mathbf{x}_i^T,\ \mathbf{x}_i'^T)=(x_i,y_i,x_i',y_i').$$

(c) 令 $\mathbf{V}_1$ 和 $\mathbf{V}_2$ 是 A 的对应于最大两个 (sic)奇异值的右奇异矢量.

(d) 令 $H_{2\times2}=CB^{-1}$,其中 B 和 C 是 $2\times2$ 分块矩阵,满足

$$[\mathbf{V}_1\quad\mathbf{V}_2]=\begin{bmatrix}B\\C\end{bmatrix}.$$

(e) 所求的单应为

$$H_A=\begin{bmatrix}H_{2\times2} & H_{2\times2}\mathbf{t}-\mathbf{t}'\\ \mathbf{0}^T & 1\end{bmatrix}.$$

算法 3.7 从图像对应估计仿射单应 $H_A$ 的黄金标准算法.

(a) 证明最佳仿射变换把 $\mathbf{x}_i$ 的形心映射到 $\mathbf{x}_i'$ 的形心,因此通过点的平移使它们的形心平移到原点,则变换的平移部分就确定了. 从而仅需要确定变换的线性部分,即 $H_A$ 左上角的 $2\times2$ 的子阵 $H_{2\times2}$.

(b) 点 $\mathbf{X}_i=(\mathbf{x}_i^T,\ \mathbf{x}_i'^T)^T$ 在 $\mathcal{V}_H$ 上的充要条件是$[H_{2\times2}\mid-I_{2\times2}]\mathbf{X}=\mathbf{0}$ 因此 $\mathcal{V}_H$ 是 $\mathbb{R}^4$中余维度为 2 的子空间.

(c) 任何余维度为 2 的子空间可以通过适当的 $H_{2\times2}$表示成 $[H_{2\times2}\mid-I_{2\times2}]\mathbf{X}=\mathbf{0}$. 因此,给定测量 $\mathbf{X}_i$,估计任务就等于找拟合最好的余维度为 2 的子空间.

(d) 给定行为 $\mathbf{X}_i^{\mathrm{T}}$ 的矩阵 M，$\mathbf{X}_i$ 的最好拟合子空间由对应 M 的两个最大奇异值的特征矢量 $\mathbf{V}_1$ 和 $\mathbf{V}_2$ 生成.

(e) 对应于由 $\mathbf{V}_1$ 和 $\mathbf{V}_2$ 生成子空间的 $\mathrm{H}_{2\times 2}$，可以通过解方程 $[\mathrm{H}_{2\times 2} \mid -\mathrm{I}][\mathbf{V}_1 \quad \mathbf{V}_2]=\mathbf{0}$ 而求得.

# 第4章 算法评价和误差分析

本章介绍如何评价和量化估计算法的结果. 通常仅有一个变量或变换的估计是不够的. 还需要进行置信度或不可靠性度的测量.

本章将概要地给出计算这种不确定性(协方差)的两种方法. 第一种方法采用线性逼近并涉及到各种雅可比式的毗连. 第二种方法比较容易,采用蒙特卡洛法.

## 4.1 性能的界定

一旦开发了某类变换的估计算法,就应该测试它的性能. 可以用实际数据或合成数据对它进行测试. 本节将采用合成数据来测试,并概要介绍测试的方法论.

我们重提一下记号的规定:

- 测量得到的量,如 $\mathbf{x}$ 表示图像点.
- 被估计的量用加帽子来表示,如 $\hat{\mathbf{x}}$ 或 $\hat{\mathrm{H}}$.
- 量的真值用加横杠表示,如 $\bar{\mathbf{x}}$ 或 $\bar{\mathrm{H}}$.

通常,测试的第一步是用合成方法产生两幅图像之间的图像对应集合 $\bar{\mathbf{x}}_i \leftrightarrow \bar{\mathbf{x}}'_i$. 对应的数目可变化. 对应点按一个给定的固定射影变换 $\bar{\mathbf{H}}$ 来选择,该对应是精确的,也就是说 $\bar{\mathbf{x}}'_i = \bar{\mathrm{H}}\bar{\mathbf{x}}_i$ 精确到机器精度.

下一步是在图像测量上人为地加上高斯噪声,即用一个已知方差的零均值的高斯随机变量对点的 $x$ 和 $y$ 坐标进行干扰. 产生的噪声点记为 $\mathbf{x}_i$ 和 $\mathbf{x}'_i$. 一种适宜的高斯随机数发生器在[Press-88]中给出. 然后,运行估计算法并求出被估计的量. 以第三章中讨论的2D射影变换为例,就是要估计射影变换本身,也许还要估计原始无噪声的正确的图像点. 然后,依据计算得到的模型与(有噪声的)输入数据的匹配程度,或估计得到的模型与原先无噪声的数据吻合的程度来评价该算法. 为了获得统计上有意义的性能评估必须用不同的噪声(即随机数产生器采用不同的种子,即使每次的噪声方差一样)进行许多次这样的过程.

### 4.1.1 单图像误差

为了说明问题,我们继续研究2D单应的估计问题. 为了简单起见,我们考虑仅给第二幅图像的坐标加噪声的情形. 于是,对所有 $i$ 有 $\mathbf{x}_i = \bar{\mathbf{x}}_i$. 令 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 为两幅图像间加噪匹配点的集合,它们由对完全匹配的数据中的第二幅图像(加撇号)的两个坐标注入方差为 $\sigma^2$ 的高斯噪声而生成. 设有 $n$ 组这样的匹配点. 根据这些数据,用第三章介绍的任何一种算法估计射影变换 $\hat{\mathrm{H}}$. 显然,估计得到的变换 $\hat{\mathrm{H}}$ 一般不会精确地将 $\mathbf{x}_i$ 映射到 $\mathbf{x}'_i$,也不会精确地将 $\bar{\mathbf{x}}_i$ 映射到 $\bar{\mathbf{x}}'_i$,因为在坐标 $\mathbf{x}'_i$ 中已注入了噪声. RMS(均方根)残差

$$\varepsilon_{\mathrm{res}} = \left(\frac{1}{2n}\sum_{i=1}^{n} d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2\right)^{1/2} \tag{4.1}$$

测量加噪声的输入数据($\mathbf{x}'_i$)和估计得到的点$\hat{\mathbf{x}}'_i = \hat{\mathrm{H}}\bar{\mathbf{x}}_i$ 之间的平均误差. 因此自然地称它为**残差**. 它测量计算得到的变换与输入数据的匹配程度,因而是估计该过程的一种适宜的质量指标.

残差的值本身并**不是**解的质量的一个绝对度量. 例如,考虑输入数据仅有 4 组匹配点的 2D 射影情形. 因为一个射影变换由 4 组点对应唯一并完全地确定,任何合理的算法都将算出一个$\hat{\mathrm{H}}$并与这些点完全匹配,即$\mathbf{x}'_i = \hat{\mathrm{H}}\mathbf{x}_i$. 这意味着残差是零. 我们还能指望一种算法有比这更好的性能吗?

注意$\hat{\mathrm{H}}$把投影点与输入数据 $\mathbf{x}'_i$ 匹配,而不是把它们与原来无噪声的数据 $\bar{\mathbf{x}}'_i$ 匹配. 事实上,因为无噪声和有噪声坐标之间的差的方差为 $\sigma^2$,在最小集 4 点的情形中投影点$\hat{\mathrm{H}}\mathbf{x}_i$与无噪声数据 $\bar{\mathbf{x}}'_i$ 之间的残差也有方差$\sigma^2$. 因此,在 4 点时,虽然模型完全匹配噪声输入点(即残差是零),但并没有真正给出无噪声值的一个很好的近似.

在多于 4 点的匹配中,残差值将增加. 直觉上,我们期望随测量(匹配点)数的增加,估计的模型将越来越接近无噪声的真值. 在渐近情况下方差应该反比例于匹配点的数目而减小. 与此同时,残差将增加.

### 4.1.2 双图像误差

对双图像误差情形,残差是

$$\varepsilon_{\mathrm{res}} = \frac{1}{\sqrt{4n}}\Big(\sum_{i=1}^{n} d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + \sum_{i=1}^{n} d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2\Big)^{1/2} \tag{4.2}$$

其中$\hat{\mathbf{x}}_i$ 和$\hat{\mathbf{x}}'_i$ 是估计的点并满足$\hat{\mathbf{x}}_i = \hat{\mathrm{H}}\hat{\mathbf{x}}'_i$.

### 4.1.3 最优估计算法(MLE)

我们先在一般框架中考虑评估性能的界,然后具体到误差在一幅和两幅图像的情形. 我们的目的是推导最大似然估计(MLE)的残差的期望公式. 如前所述,几何误差的最小化等于 MLE,因此任何实现几何误差最小化的算法的目标应该是达到 MLE 给出的理论界. 最小化其他不同代价函数(例如代数误差)的算法可以根据它与 MLE 所给出的界的接近程度来作性能判断.

如节 3.2.7($p62$)所介绍,一般的估计问题关系到一个由 $\mathbb{R}^M$到 $\mathbb{R}^N$ 的函数$f$,其中$\mathbb{R}^M$是参数空间,而 $\mathbb{R}^N$ 是测量空间. 现在令点 $\bar{\mathbf{X}} \in \mathbb{R}^N$,且存在一个参数矢量 $\bar{\mathbf{P}} \in \mathbb{R}^M$使得$f(\bar{\mathbf{P}}) = \bar{\mathbf{X}}$(点 $\bar{\mathbf{X}}$ 在 $f$ 的值域中并以 $\bar{\mathrm{P}}$ 为其前象). 在仅需对第二幅图像进行测量的 2D 射影变换估计的问题中,第二幅图像对应于无噪声的点集:$\bar{\mathbf{x}}'_i = \bar{\mathrm{H}}\,\bar{\mathbf{x}}_i$. $n$ 个点 $\bar{\mathbf{x}}'_i$,($i = 1, \ldots, n$)的 $x$ 和 $y$ 分量组成 $N$ 维矢量 $\bar{\mathbf{X}}$,这里 $N = 2n$,而单应的参数组成矢量 $\bar{\mathbf{P}}$,根据 $\bar{\mathrm{H}}$ 的参数化取法,$\mathbf{P}$ 可能是 8 或 9 维矢量.

令 $\mathbf{X}$ 是根据各向同性高斯分布选取的测量矢量,其均值为测量真值 $\bar{\mathbf{X}}$ 而方差为 $N\sigma^2$(此记号表示每个 $N$ 分量都有方差$\sigma^2$). 如图 4.1 所说明,当参数矢量 $\mathbf{P}$ 的值在点 $\bar{\mathbf{P}}$ 的邻域变化时,函数 $f(\mathbf{P})$的值形成 $\mathbb{R}^N$ 中过点 $\bar{\mathbf{X}}$ 的曲面 $S_M$. 曲面 $S_M$ 由 $f$ 的值域给出. 曲面 $S_M$ 是 $\mathbb{R}^N$ 的子流形,其维数等于 $d$,而 $d$ 正是本质参数的数目(即自由度的数目,或参数的最低

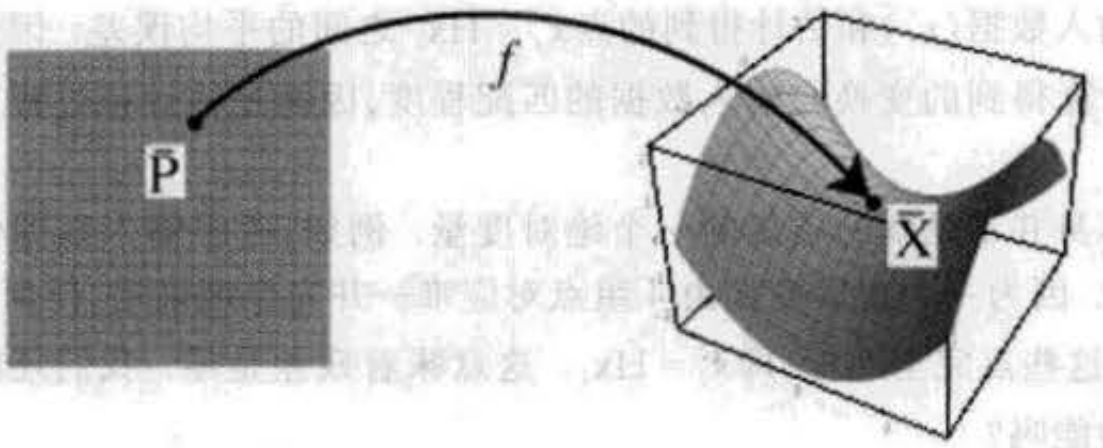



图 4.1 随参数矢量 **P** 值的变化对应函数的象形成过真值 $\bar{\mathbf{X}}$ 的曲面 $S_M$.

限度数目). 在单图像误差的情形,它等于 8,因为由矩阵 H 所确定的映射与尺度因子无关.

给定测量矢量 **X**,最大似然(ML)估计 $\hat{\mathbf{X}}$ 是 $S_M$ 上的最接近 **X** 的点. ML 估计算法就是返回该曲面上离 **X** 最近的点的算法. 把这个 ML 估计记为 $\hat{\mathbf{X}}$.

假定在 $\bar{\mathbf{X}}$ 的邻域曲面基本上是平面,即切平面可作为它的一个很好的近似——至少在 $\bar{\mathbf{X}}$ 的周围处于噪声方差数量级的邻域内是如此. 基于这个线性近似的假设,ML 估计 $\hat{\mathbf{X}}$ 是 **X** 到切平面上的垂足. 残差则是点 **X** 到估计值 $\hat{\mathbf{X}}$ 的距离. 而 $\hat{\mathbf{X}}$ 到(未知的)$\bar{\mathbf{X}}$ 的距离是最佳估计值到真值的距离,如图 4.2 所示. 我们的任务就是计算这些误差的期望值.

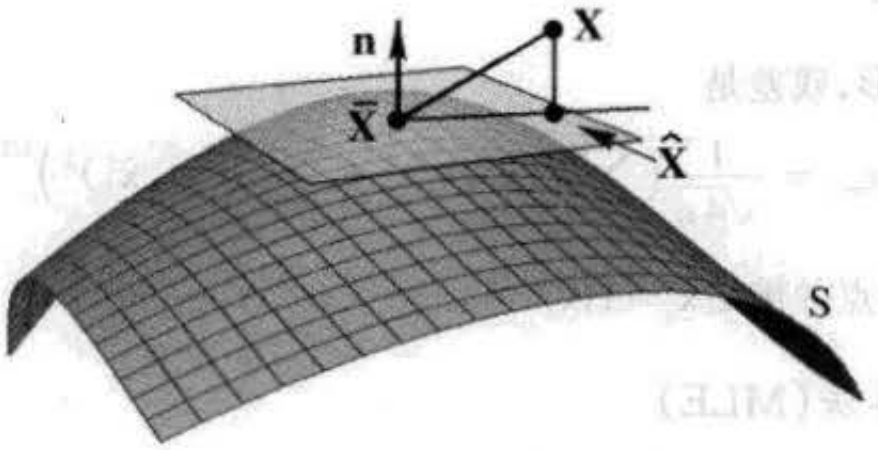



图 4.2 在误差测量空间中用切面逼近 $S_M$ 的几何. 估计点 $\hat{\mathbf{X}}$ 是 $S_M$ 上离被测量点 **X** 最近的点. 残差是测量点 **X** 和 $\hat{\mathbf{X}}$ 之间的距离. 估计误差是由 $\hat{\mathbf{X}}$ 到真值点 $\bar{\mathbf{X}}$ 的距离.

计算 ML 残差的期望值现在可以抽象为如下的几何问题. $N$ 维高斯分布的**总方差**是协方差矩阵的迹,即在每一轴方向的方差的和. 它在坐标系的正交变化下不变. 每一维变量都具有独立方差 $\sigma^2$ 的 $N$ 维各向同性高斯分布的总方差是 $N\sigma^2$. 现在,给定定义在 $\mathbb{R}^N$ 上的总方差是 $N\sigma^2$ 而均值是真值点 $\bar{\mathbf{X}}$ 的各向同性高斯随机变量,我们希望计算该随机变量到过 $\bar{\mathbf{X}}$ 的超平面(维数为 $d$)的距离的期望值. $\mathbb{R}^N$ 上的该高斯随机变量到 $d$ 维切平面的投影给出了**估计误差**(估计值与真值之差)的分布. 而到该切平面的($N-d$)维的法曲面的投影给出残差的分布.

我们可以不失一般性地假定切平面与前 $d$ 个坐标轴重合(必要时旋转坐标轴). 在余下的轴方向取积分给出如下结论.

**结论 4.1** **$\mathbb{R}^N$ 上总方差为 $N\sigma^2$ 的各向同性高斯分布向一个 $s$ 维子空间的投影是总方差为 $s\sigma^2$ 的各向同性高斯分布.**

证明是直接的,因而省略. 我们把它应用到 $s=d$ 和 $s=N-d$ 两种情形得出如下结论.

**结论 4.2** 考虑一个估计问题，其中 $N$ 个测量由依赖于 $d$ 个本质参数集的函数模型化．假定每个测量变量有标准差 $\sigma$ 的独立高斯噪声．

(1) ML 估计算法的 RMS 残差(测量值到估计值的距离)是

$$\varepsilon_{res} = E[\,\|\hat{\mathbf{X}} - \mathbf{X}\|/\mathbf{N}]^{1/2} = \sigma(1 - d/N)^{1/2}.$$

(2) ML 估计算法的 RMS 估计误差(估计值到真值的距离)是

$$\varepsilon_{est} = E[\,\|\hat{\mathbf{X}} - \bar{\mathbf{X}}\|/\mathbf{N}]^{1/2} = \sigma(d/N)^{1/2}$$

其中 $\mathbf{X}, \hat{\mathbf{X}}, \bar{\mathbf{X}}$ 分别是测量矢量的测量值、估计值和真值．

结论 4.2 直接由结论 4.1 得到，把它除以 $N$ 求得每个测量的方差，然后取平方根得到标准差，并用它取代方差．

这些值给出评价一个具体估计算法的残差的下界．

**2D 单应——单图像误差** 对本章所考虑的 2D 射影变换的估计，假定误差仅出现在第二幅图像中，我们有 $d = 8$ 和 $N = 2n$，其中 $n$ 是匹配点的数目．于是，对这个问题我们得到

$$\varepsilon_{res} = \sigma(1 - 4/n)^{1/2}$$
$$\varepsilon_{est} = \sigma(4/n)^{1/2}. \tag{4.3}$$

当 $n$ 变化时，这些误差的图在图 4.3 中给出．

**双图像误差** 在此情形，$N = 4n$ 和 $d = 2n + 8$．如上面一样，假定在测量矢量的真值 $\hat{\mathbf{X}}$ 的邻域有切平面的线性近似，结论 4.2 中给出如下误差期望值．

$$\varepsilon_{res} = \sigma\left(\frac{n-4}{2n}\right)^{1/2}$$
$$\varepsilon_{est} = \sigma\left(\frac{n+4}{2n}\right)^{1/2}. \tag{4.4}$$

当 $n$ 变化时，这些误差的图也在图 4.3 中给出．

从图中可以看到一个有趣的现象：此时，真值的渐进误差是 $\sigma/\sqrt{2}$ 而不是单图像误差时的 0．该结果是意料之中的，因为事实上，每点的位置有两次测量，每幅图像上各一次并与射影变换相关．一个点进行两次测量，估计点的位置方差降低了 $\sqrt{2}$ 倍．相反地，在前一种情形误差仅发生在一幅图像中，因而第一幅图像上的每点有准确的测量．因此，当变换 H 估计的精确度越来越大时，在第二幅图像中点的精确位置变得已知，其不确定性渐近地趋向 0．

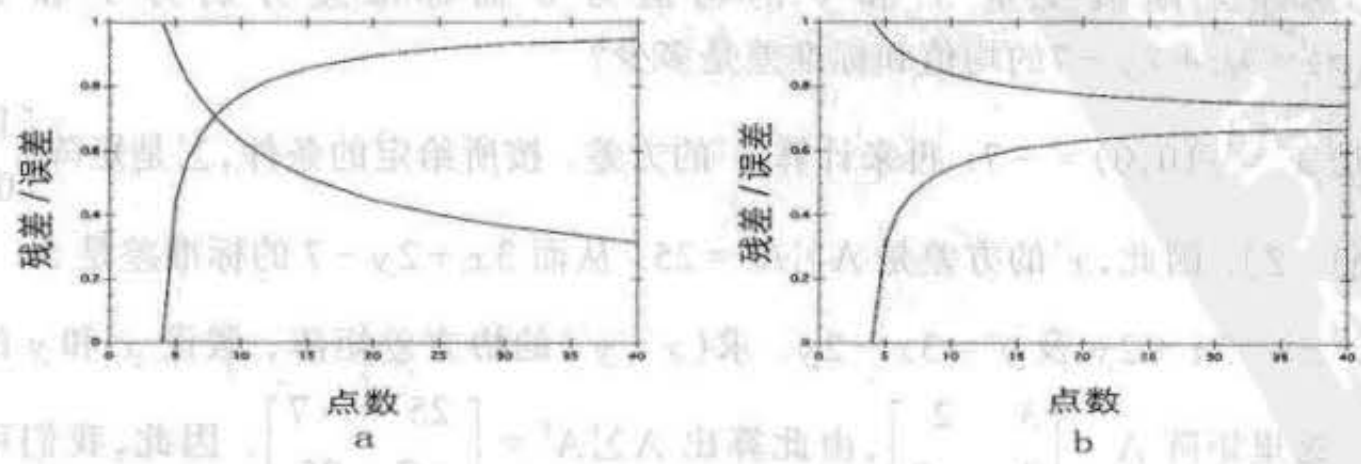



图 4.3 当点的数目变化时，噪声在(a)一幅图像和(b)两幅图像所得到的最优误差．假定误差水平是 1 个像素．下降曲线表示估计误差 $\varepsilon_{est}$，而上升曲线表示残差 $\varepsilon_{res}$

**Mahalanobis 距离** 上面的公式是在测量空间的误差分布服从各向同性高斯分布的假设下推导的，这意味着每个坐标的误差是独立的. 该假设不是本质的. 我们可以假设误差服从具有协方差矩阵$\Sigma$的任何高斯分布，把 $\epsilon$ 换成 Mahalanobis 距离的期望值 $E[\|\hat{\mathbf{X}}-\mathbf{X}\|_{\Sigma}]$ 时结论 4.2 的公式仍然成立. 标准差 $\sigma$ 也不再出现，由取 Mahalanobis 距离取代.

在测量空间 $\mathbb{R}^N$ 上用一个坐标变换可使协方差矩阵变为单位阵. 在这新的坐标系中，Mahalanobis 距离与欧氏距离一样，因而就可证明上面的论述是正确的.

## 4.2 变换估计的协方差

在上节中，我们考虑了 ML 估计以及如何计算平均误差的期望值. 把一个算法所得到的残差或估计误差与 ML 误差相比较是评价一个具体的估计算法的性能的一种好方法，因为它把算法的结果与在没有任何先验信息的情况下能取得的最好结果(最优估计)比较.

但是，我们主要关心的是变换本身的准确度如何. 变换估计的不可靠性取决于许多因素，它包括用于计算它的点数、给定的匹配点的准确度以及点的配置. 配置是重要的，假定用于计算变换的点接近于一种退化配置，那么求得的变换不可能有好的准确性. 例如，如果用一条直线附近的点来计算变换，那么变换在垂直于直线的维度上的作用不能被准确地确定. 因此，尽管残差和估计误差看上去仅取决于对应点的数目及其精度，但是求得的变换的准确性却依赖于具体的点. 求得的变换的不可靠性通常由变换的**协方差矩阵**获取. 因为 H 是 9 元素的矩阵，它的协方差矩阵是一个 $9\times9$ 矩阵. 本节将介绍如何来计算这个协方差矩阵.

### 4.2.1 协方差的前向传播

协方差矩阵在仿射变换下有简单性质，如下面的定理所述.

**结论 4.3** **令 $\mathbf{v}$ 是 $\mathbb{R}^M$ 中的一个具有均值 $\bar{\mathbf{v}}$ 和协方差矩阵 $\Sigma$ 的随机矢量，假定 $f:\mathbb{R}^M\to\mathbb{R}^N$ 是一个仿射映射：定义为 $f(\mathbf{v})=f(\bar{\mathbf{v}})+\mathrm{A}(\mathbf{v}-\bar{\mathbf{v}})$. 那么 $f(\mathbf{v})$ 是一个具有均值 $f(\bar{\mathbf{v}})$ 和协方差矩阵 $\mathrm{A}\Sigma\mathrm{A}^{\mathrm{T}}$ 的随机变量.**

注意：我们没有假定 A 是方阵. 我们现在以举例说明来代替该结论的证明.

**例 4.4** 令独立随机变量 $x$ 和 $y$ 的均值为 0 而标准差分别为 1 和 2. 那么 $x'=f(x,y)=3x+2y-7$ 的均值和标准差是多少？

均值是 $\overline{x'}=f(0,0)=-7$. 再来计算 $x'$ 的方差. 按所给定的条件，$\Sigma$ 是矩阵 $\begin{bmatrix}1 & 0\\0 & 4\end{bmatrix}$，而 A 是矩阵 $[3\quad 2]$. 因此，$x'$ 的方差是 $\mathrm{A}\Sigma\mathrm{A}^{\mathrm{T}}=25$. 从而 $3x+2y-7$ 的标准差是 5. △

**例 4.5** 令 $x'=3x+2y$ 及 $y'=3x-2y$. 求 $(x',y')$ 的协方差矩阵，假设 $x$ 和 $y$ 的分布与上例一样. 这里矩阵 $\mathrm{A}=\begin{bmatrix}3 & 2\\3 & -2\end{bmatrix}$，由此算出 $\mathrm{A}\Sigma\mathrm{A}^{\mathrm{T}}=\begin{bmatrix}25 & -7\\-7 & 25\end{bmatrix}$. 因此，我们可以看到 $x'$ 和 $y'$ 的方差是 25(标准差是 5)，而 $x'$ 和 $y'$ 是负相关的，其协方差是 $E[x'y']=-7$. △

**非线性传播** 如果 $\mathbf{v}$ 是 $\mathbb{R}^M$ 中的一个随机矢量，而 $f:\mathbb{R}^M\to\mathbb{R}^N$ 是一个作用于 $\mathbf{v}$ 上的非线性

函数,那么若假设 $f$ 在 $\mathbf{v}$ 的分布的均值附近近似于仿射变换,我们就可以计算 $f(\mathbf{v})$ 的均值和协方差的近似值. $f$ 的仿射逼近是 $f(\mathbf{v})\approx f(\bar{\mathbf{v}})+\mathrm{J}(\mathbf{v}-\bar{\mathbf{v}})$,这里 J 是偏导数(雅可比)矩阵 $\partial f/\partial\mathbf{v}$ 在 $\bar{\mathbf{v}}$ 的值. 注意 J 的维数是 $N\times M$. 那么我们有如下结论.

**结论 4.6** **令 $\mathbf{v}$ 是 $\mathbb{R}^M$ 中一个具有均值 $\bar{\mathbf{v}}$ 和协方差矩阵 $\Sigma$ 的随机矢量,令 $f:\mathbb{R}^M\to\mathbb{R}^N$ 在 $\bar{\mathbf{v}}$ 的邻域可微. 那么在精确到一阶近似的程度下, $f(\mathbf{v})$ 是一个具有均值 $f(\bar{\mathbf{v}})$ 和协方差矩阵 $\mathrm{J}\Sigma\mathrm{J}^{\mathrm{T}}$ 的随机变量,其中 J 是 $f$ 的雅可比矩阵在 $\bar{\mathbf{v}}$ 的值.**

这个结果与 $f(\bar{\mathbf{v}})$ 的实际均值和方差近似的程度取决于在处于 $\bar{\mathbf{v}}$ 周围与 $\mathbf{v}$ 概率分布的支集相当大小的区域上,函数 $f$ 近似于线性函数的程度.

**例 4.7** 令 $\mathbf{x}=(x,y)^{\mathrm{T}}$ 是一个具有均值 $(0,0)^{\mathrm{T}}$ 和协方差矩阵 $\sigma^2\mathrm{diag}(1,4)$ 的高斯随机矢量. $x'=f(x,y)=x^2+3x-2y+5$. 那么我们可以根据下面的公式计算 $f(x,y)$ 的均值和标准差的真值

$$\bar{x}'=\iint_{-\infty}^{\infty}P(x,y)f(x,y)dxdy$$

$$\sigma_{x'}^2=\iint_{-\infty}^{\infty}P(x,y)(f(x,y)-\bar{x}')^2dxdy$$

其中

$$P(x,y)=\frac{1}{4\pi\sigma^2}e^{-(x^2+y^2/4)/2\sigma^2}$$

是高斯概率分布(A2.1 - $p$402). 我们得到

$$\bar{x}'=5+\sigma^2$$

$$\sigma_{x'}^2=25\sigma^2+2\sigma^4.$$

应用结论 4.6 给出的近似值,并注意到 $\mathrm{J}=[3\quad -2]$,我们求得其估计值为

$$\bar{x}'=5$$

$$\sigma_{x'}^2=\sigma^2[3\quad -2]\begin{bmatrix}1 & \\ & 4\end{bmatrix}[3\quad -2]^{\mathrm{T}}=25\sigma^2.$$

因此,只要 $\sigma$ 足够小,它们就是 $x'$ 的均值和方差的正确值的一个好的近似. 下面的表中分别对 $\sigma$ 的两种不同取值给出 $f(x,y)$ 的均值和标准差的真值和近似值.

| | $\sigma=0.25$ | | $\sigma=0.5$ | |
|---|---|---|---|---|
| | $\bar{x}'$ | $\sigma_{x'}$ | $\bar{x}'$ | $\sigma_{x'}$ |
| 估计值 | 5.0000 | 1.25000 | 5.00 | 2.5000 |
| 真值 | 5.0625 | 1.25312 | 5.25 | 2.5249 |

作为参照,在 $\sigma=0.25$ 的情形,我们看到只要 $|x|<2\sigma$(大约占总分布的 95%). $f(x,y)=x^2+3x-2y+5$ 的值与它的线性近似的差别不大于 $x^2<0.25$ △

**例 4.8** 更一般地,假定 $x$ 和 $y$ 为独立、零均值的高斯随机变量,函数 $f(x,y)=ax^2+bxy+cy^2+dx+ey+f$,我们可以算出此函数的

$$\text{均值}=a\sigma_x^2+c\sigma_y^2+f$$

$$\text{方差}=2a^2\sigma_x^4+b^2\sigma_x^2\sigma_y^2+2c^2\sigma_y^4+d^2\sigma_x^2+e^2\sigma_y^2$$

只要 $\sigma_x$ 和 $\sigma_y$ 足够小，它们就接近估计的均值 $= f$ 和方差 $= d^2\sigma_x^2 + e^2\sigma_y^2$. △

### 4.2.2 协方差的反向传播

本节和下节 4.2.3 的内容比较高深. 节 4.2.4 的例子给出这几节中的结论的直接应用，可以先阅读.

考虑由"参数空间"$\mathbb{R}^M$到"测量空间"$\mathbb{R}^N$的一个可微映射 $f$，并令一个协方差矩阵为 $\Sigma$ 的高斯概率分布定义在 $\mathbb{R}^N$ 上. 令 $S_M$ 为映射 $f$ 的值域. 我们假定 $M < N$ 且 $S_M$ 与参数空间 $\mathbb{R}^M$ 有相同的维数 $M$. 我们暂时不考虑超参数化的情况. $\mathbb{R}^M$ 的矢量 $\mathbf{P}$ 是 $S_M$ 上的点 $f(\mathbf{P})$ 的一种参数化. 在 Mahalanobis 距离的意义下，求在 $S_M$ 上最接近 $\mathbb{R}^N$ 上给定点 $\mathbf{X}$ 的点定义了由 $\mathbb{R}^N$ 到曲面 $S_M$ 的一个映射. 我们称它为 $\eta: \mathbb{R}^n \to S_M$. 现在假定 $f$ 在曲面 $S_M$ 上可逆，那么定义 $f^{-1}: S_M \to \mathbb{R}^M$ 为其逆函数.

通过映射 $\eta: \mathbb{R}^N \to S_M$ 和 $f^{-1}: S_M \to \mathbb{R}^M$ 的复合，我们得到一个映射 $f^{-1} \circ \eta: \mathbb{R}^N \to \mathbb{R}^M$. 这个映射把对应于 ML 估计 $\hat{\mathbf{X}}$ 的参数矢量 $\mathbf{P}$ 指定给测量矢量 $\mathbf{X}$. 原则上，我们可以把测量空间 $\mathbb{R}^N$ 的概率分布的协方差传播给对应于 ML 估计的这组参数 $\mathbf{P}$ 的协方差矩阵的计算. 我们的目标是应用结论 4.3 或结论 4.6.

我们首先考虑的情况是映射 $f$ 是由 $\mathbb{R}^M$ 到 $\mathbb{R}^N$ 的仿射映射. 下一步我们将证明 $f^{-1} \circ \eta$ 同样是仿射变换并将给出 $f^{-1} \circ \eta$ 的一种具体形式，由此，我们能应用结论 4.3 去计算被估计的参数 $\hat{\mathbf{P}} = f^{-1} \circ \eta(\mathbf{X})$ 的协方差.

因为 $f$ 是仿射，我们可以记 $f(\mathbf{P}) = f(\bar{\mathbf{P}}) + \mathrm{J}(\mathbf{P} - \bar{\mathbf{P}})$，其中 $f(\bar{\mathbf{P}}) = \bar{\mathbf{X}}$ 是 $\mathbb{R}^N$ 上概率分布的均值. 因为我们假定曲面 $S_M = f(\mathbb{R}^M)$ 的维数是 $M$，J 的秩等于它的列矢量的维数. 给定一个测量矢量 $\mathbf{X}$，ML 估计 $\hat{\mathbf{X}}$ 最小化 $\|\mathbf{X} - \hat{\mathbf{X}}\|_\Sigma = \|\mathbf{X} - f(\hat{\mathbf{P}})\|_\Sigma$. 因此，我们寻求最小化后一个量的 $\hat{\mathbf{P}}$. 然而，

$$\|\mathbf{X} - f(\hat{\mathbf{P}})\|_\Sigma = \|(\mathbf{X} - \bar{\mathbf{X}}) - \mathrm{J}(\hat{\mathbf{P}} - \bar{\mathbf{P}})\|_\Sigma$$

在 $(\hat{\mathbf{P}} - \bar{\mathbf{P}}) = (\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1}\mathrm{J}^\mathrm{T}\Sigma^{-1}(\mathbf{X} - \bar{\mathbf{X}})$ 时被最小化（见 A3.4.1（$p$411）的（A3.9 − $p$411））.

记 $\bar{\mathbf{P}} = f^{-1}\bar{\mathbf{X}}$ 和 $\hat{\mathbf{P}} = f^{-1}\hat{\mathbf{X}}$，我们看到

$$\begin{aligned} f^{-1} \circ \eta(\mathbf{X}) &= \hat{\mathbf{P}} \\ &= (\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1}\mathrm{J}^\mathrm{T}\Sigma^{-1}(\mathbf{X} - \bar{\mathbf{X}}) + f^{-1}(\bar{\mathbf{X}}) \\ &= (\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1}\mathrm{J}^\mathrm{T}\Sigma^{-1}(\mathbf{X} - \bar{\mathbf{X}}) + f^{-1} \circ \eta(\bar{\mathbf{X}}). \end{aligned}$$

这表明 $f^{-1} \circ \eta$ 是仿射，而 $(\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1}\mathrm{J}^\mathrm{T}\Sigma^{-1}$ 是它的线性部分. 应用结论 4.3，我们得到 $\hat{\mathbf{P}}$ 的协方差矩阵是

$$\begin{aligned} &[(\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1}\mathrm{J}^\mathrm{T}\Sigma^{-1}]\Sigma[(\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1}\mathrm{J}^\mathrm{T}\Sigma^{-1}]^\mathrm{T} \\ &= (\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1}\mathrm{J}^\mathrm{T}\Sigma^{-1}\Sigma\Sigma^{-1}\mathrm{J}(\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1} = (\mathrm{J}^\mathrm{T}\Sigma^{-1}\mathrm{J})^{-1}, \end{aligned}$$

其中用了 $\Sigma$ 的对称性. 以上证明了如下定理.

**结论 4.9 协方差的反向输送——仿射情形.** 令 $f: \mathbb{R}^M \to \mathbb{R}^N$ 是形为 $f(\mathbf{P}) = f(\bar{\mathbf{P}}) + \mathrm{J}(\mathbf{P} - \bar{\mathbf{P}})$ 的仿射映射，其中 J 的秩等于 $M$. 令 $\mathbf{X}$ 是 $\mathbb{R}^N$ 中的一个具有均值 $\bar{\mathbf{X}} = f(\bar{\mathbf{P}})$ 和协方差矩阵 $\Sigma$ 的随机变量. 令 $f^{-1} \circ \eta: \mathbb{R}^N \to \mathbb{R}^M$ 是一个映射，它把测量矢量 $\mathbf{X}$ 映射到对应于 ML 估计 $\hat{\mathbf{X}}$ 的参数矢量 $\mathbf{P}$. 那么 $\hat{\mathbf{P}} = f^{-1} \circ \eta(\mathbf{X})$ 是一个具有均值 $\bar{\mathbf{P}}$ 的随机变量，其协方差矩阵是

$$\Sigma_P = (J^T \Sigma_X^{-1} J)^{-1}. \tag{4.5}$$

当 $f$ 不是仿射映射时，可以通过通常的途径用一个仿射函数逼近 $f$ 来获得均值和协方差的近似，下文将对此给予介绍.

**结论 4.10 协方差的反向输送——非线性情形.** 令 $f: \mathbb{R}^M \to \mathbb{R}^N$ 是一个可微映射，而 J 是它在点 $\bar{\mathbf{P}}$ 处的雅可比矩阵. 假定 J 的秩为 $M$，则 $f$ 在 $\bar{\mathbf{P}}$ 的邻域是一一对应的. 令 $\mathbf{X}$ 是 $\mathbb{R}^N$ 中的一个具有均值 $\bar{\mathbf{X}} = f(\bar{\mathbf{P}})$ 和协方差矩阵 $\Sigma_X$ 的随机变量. 令映射 $f^{-1} \circ \eta: \mathbb{R}^N \to \mathbb{R}^M$，把测量矢量 $\mathbf{X}$ 映射到对应于 $ML$ 估计 $\hat{\mathbf{X}}$ 的参数矢量 $\mathbf{P}$. 那么在一阶精度下，$\hat{\mathbf{P}} = f^{-1} \circ \eta(\mathbf{X})$ 是一个具有均值 $\bar{\mathbf{P}}$ 和协方差矩阵 $(J^T \Sigma_X^{-1} J)^{-1}$ 的随机变量.

### 4.2.3 超参数化

我们可以把结论 4.9 和结论 4.10 推广到冗余参数集——超参数化的情形. 其中，参数空间 $\mathbb{R}^M$ 到测量空间 $\mathbb{R}^N$ 的映射 $f$ 局域上不是一一对应的. 例如，如节 3.5($p69$)所述，在估计 2D 单应时，存在一个映射 $f(\mathbf{P})$，其中 $\mathbf{P}$ 是表示单应矩阵 H 的元素的 9 维矢量. 因为单应仅有 8 个自由度，故映射 $f$ 不是一一对应的. 具体地说，对任何常数 $k$，矩阵 $k$H 表示同样的映射，因而图像坐标矢量 $f(\mathbf{P})$ 和 $f(k\mathbf{P})$ 相等.

在一般情形，一个映射 $f: \mathbb{R}^M \to \mathbb{R}^N$ 的雅可比矩阵 J 不是满秩 $M$，而是有一个较小的秩 $d < M$. 该秩 $d$ 称为**本质参数数**. 此时，矩阵 $J^T \Sigma_X^{-1} J$ 的维数是 $M$ 但秩为 $d < M$. 公式(4.5)：$\Sigma_P = (J^T \Sigma_X^{-1} J)^{-1}$ 显然不成立，因为该式右边的矩阵不可逆.

事实上，很清楚如果没有进一步的约束，被估计的矢量 $\hat{\mathbf{P}}$ 的元素显然可以通过乘以任意非零常数 $k$ 而无界地变化. 从而，该元素具有无穷方差. 通常应通过某种约束来限制被估计的单应矩阵 H 或更一般地限制参数矢量 $\mathbf{P}$. 通用的约束是 $\|\mathbf{P}\| = 1$，虽然其他的约束也是可能的，例如要求最后一个参数等于 1(见节 3.4.2($p66$)). 因此参数矢量 $\mathbf{P}$ 被限制在参数空间 $\mathbb{R}^9$ 或更一般地在 $\mathbb{R}^M$ 的一个曲面上. 对第一种约束，曲面 $\|\mathbf{P}\| = 1$ 是 $\mathbb{R}^M$ 中的一个单位球面. 第二种约束 $P_m = 1$ 则表示 $\mathbb{R}^M$ 中的一张平面. 一般，我们可以假定被估计的矢量 $\mathbf{P}$ 在 $\mathbb{R}^M$ 的某个子流形上，如下面定理所示.

**结论 4.11 协方差的反向输送——超参数化情形.** 令 $f: \mathbb{R}^M \to \mathbb{R}^N$ 是一个可微映射，它把一组参数 $\bar{\mathbf{P}}$ 映射到测量矢量 $\bar{\mathbf{X}}$. 令 $S_P$ 是嵌入 $\mathbb{R}^M$ 中的过点 $\bar{\mathbf{P}}$ 的 $d$ 维光滑流形并使得映射 $f$ 在流形 $S_P$ 上 $\bar{\mathbf{P}}$ 的一个邻域内是一一对应的，$f$ 把 $S_P$ 局域地映射到 $\mathbb{R}^N$ 上的流形 $f(S_P)$. 函数 $f$ 有一个局部逆函数，记为 $f^{-1}$，它限制在曲面 $f(S_P)$ 上 $\bar{\mathbf{X}}$ 的一个邻域内. 定义 $\mathbb{R}^N$ 上的一个具有均值 $\bar{\mathbf{X}}$ 和协方差 $\Sigma_X$ 的高斯分布，并令 $\eta: \mathbb{R}^N \to f(S_P)$ 把 $\mathbb{R}^N$ 上的点映射到 $f(S_P)$ 上并在 Mahalanobis 范数 $\|\cdot\|_{\Sigma_X}$ 意义下最近的点. $\mathbb{R}^N$ 上具有协方差矩阵 $\Sigma_X$ 的概率分布通过 $f^{-1} \circ \eta$ 诱导 $\mathbb{R}^M$ 上的概率分布，它在一阶精度下的协方差矩阵是

$$\Sigma_P = (J^T \Sigma_X^{-1} J)^{+A} = A(A^T J^T \Sigma_X^{-1} J A)^{-1} A^T \tag{4.6}$$

其中 A 是任意 $m \times d$ 矩阵，它的列矢量生成 $S_P$ 的过点 $\bar{\mathbf{P}}$ 的切空间.

图 4.4 给出这方面的说明. 由(4.6)定义的记号 $(J^T \Sigma_X^{-1} J)^{+A}$ 将在节 A3.4($p410$)中作进一步讨论.

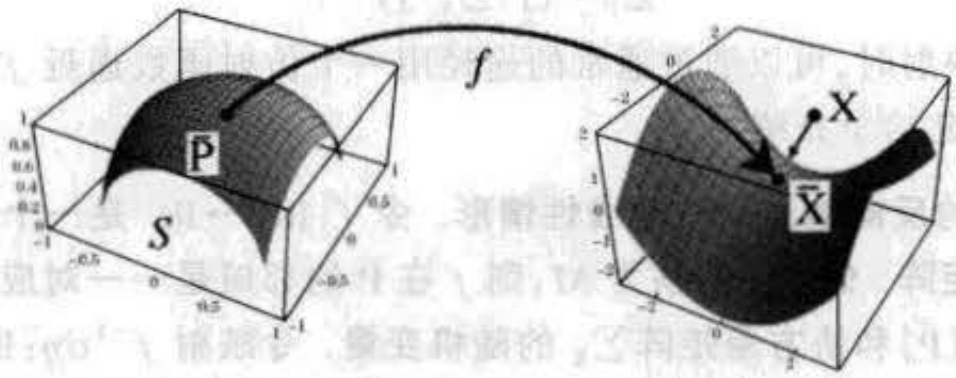



图 4.4 **反向输送(超参数)**. 映射 $f$ 把约束参数曲面映射到测量空间. 一个测量 $\mathbf{X}$(通过映射 $\eta$)映射到 $f(S_{\mathbf{P}})$上最近的点,然后通过 $f^{-1}$ 返回参数空间,提供参数的 $ML$ 估计. $\mathbf{X}$ 的协方差通过 $f^{-1}o\eta$ 传递到参数的协方差.

**证明** 结论 4.11 的证明是容易的. 令 $d$ 为本质参数数. 定义映射 $g:\mathbb{R}^d\to\mathbb{R}^M$把$\mathbb{R}^d$中的开邻域 $U$ 映射到包含点 $\bar{\mathbf{P}}$ 的开集 $S_{\mathbf{P}}$. 那么复合映射 $fog:\mathbb{R}^d\to\mathbb{R}^N$在邻域 $U$ 上是一一对应的. 让我们用 J 和 A 分别表示 $f$ 和 $g$ 的偏导数矩阵. 那么 $fog$ 的偏导数矩阵为 JA. 现在应用结论 4.10,我们看到 $\mathbb{R}^N$上具有协方差矩阵 $\Sigma$ 的概率分布函数可以反向输送到$\mathbb{R}^d$上的具有协方差矩阵$(A^TJ^T\Sigma^{-1}JA)^{-1}$的概率分布函数. 应用结论 4.6,把它再正向输送到$\mathbb{R}^M$可以得到 $S_{\mathbf{P}}$ 上的协方差矩阵 $A(A^TJ^T\Sigma^{-1}JA)^{-1}A^T$. 把这个矩阵记为$(J^T\Sigma^{-1}J)^{+A}$,它与节 A3.4($p$410)中所定义的$(J^T\Sigma^{-1}J)$的伪逆有关. 表达式(4.6)不依赖于矩阵 A 的具体选择,只要 A 的列矢量生成子空间不变. 具体地说,对任意可逆的 $d\times d$ 矩阵 B,如果 A 被 BA 代替,那么(4.6)的值不变. 因此,任何矩阵 A 只要它的列矢量生成 $S_{\mathbf{P}}$ 的过点 $\bar{\mathbf{P}}$ 的切空间就行. □

注意上述证明给出了计算生成切空间矩阵 A——即 $g$ 的雅可比矩阵的一种特殊方法. 如我们将要看到的,在许多情况下存在求 A 的更方便的方法. 注意协方差矩阵(4.6)是奇异的. 具体地说,它的维数是 $M$ 而秩是$d<M$. 这是因为被估计的参数矢量的方差在与约束曲面 $S_{\mathbf{P}}$ 正交的方向上是零——即在正交方向上没有变化. 注意虽然 $J^T\Sigma^{-1}J$ 是不可逆的,$d\times d$ 矩阵 $A^TJ^T\Sigma^{-1}JA$ 的秩是 $d$,并且是可逆的.

一种重要的情形是约束曲面局部地正交于雅可比矩阵的零空间. 记 $N_L(X)$是矩阵 X 的左零空间,即是由所有使 $\mathbf{x}^TX=\mathbf{0}^T$的矢量 $\mathbf{x}$ 组成的空间. 那么如节 A3.4($p$410)所示,**伪逆** $X^+$ 为

$$X^+=X^{+A}=A(A^TXA)^{-1}A^T$$

的充要条件是 $N_L(A)=N_L(X)$. 下面的结论可以从结论 4.11 中直接得到.

**结论 4.12** **令可微映射 $f:\mathbb{R}^M\to\mathbb{R}^N$ 把 $\bar{\mathbf{P}}$ 映射到 $\bar{\mathbf{X}}$,并令 J 为 $f$ 的雅可比矩阵. 设 $\mathbb{R}^N$ 上一个具有协方差矩阵 $\Sigma_{\mathbf{X}}$ 的高斯分布定义在 $\bar{\mathbf{X}}$,同时如结论 4.11,令 $f^{-1}o\eta:\mathbb{R}^N\to\mathbb{R}^M$是把一个测量 $\mathbf{X}$ 映到约束在局部正交于 J 的零空间的曲面 $S_{\mathbf{P}}$ 上的 MLE 参数矢量 $\mathbf{P}$ 的映射. 那么 $f^{-1}o\eta$ 诱导在 $\mathbb{R}^M$上的一个分布,它的协方差矩阵在一阶精度下是**

$$\Sigma_{\mathbf{P}}=(J^T\Sigma_{\mathbf{X}}^{-1}J)^+. \tag{4.7}$$

注意 $\mathbf{P}$ 限制在一个局部正交于 J 的零空间的曲面上的约束在许多情况中是自然的约束. 例如,如果 $\mathbf{P}$ 是一个**齐次**参数矢量(如一个齐次矩阵的元素),则约束满足通常的限制

$\|\mathbf{P}\|=1$. 在这种情形,约束曲面是单位球面,且任何一点的切面垂直于该参数矢量. 另一方面,因为 $\mathbf{P}$ 是齐次矢量,函数 $f(\mathbf{P})$ 不随比例因子而改变,因而 J 在径向方向有一个零矢量,因此垂直于约束曲面.

对于其他情形,在计算参数的协方差矩阵时对参数矢量上加什么限制通常不是关键. 另外,因为伪逆运算就是它自身的求逆,我们可以根据 $J^T\Sigma_X^{-1}J=\Sigma_P^+$ 从它的伪逆求得原始矩阵,然后,我们可以根据下式计算对应任何其他的子空间的协方差矩阵:

$$(J^T\Sigma_X^{-1}J)^{+A}=(\Sigma_P^+)^{+A}$$

其中 A 的列矢量生成参数空间的约束子空间.

### 4.2.4 应用与举例

**单图像误差** 让我们把上述理论用于求被估计的 2D 单应 H 的协方差. 首先,我们来看误差限制在第二幅图像中的情形. 3×3 矩阵 H 被表示成 9 维参数矢量的形式,并用记号 $\mathbf{h}$ 代替 $\mathbf{P}$ 以便提示我们它由 H 的元素组成. 被估计的 $\hat{\mathbf{h}}$ 的协方差是 9×9 的对称矩阵. 匹配点 $\bar{\mathbf{x}}_i\leftrightarrow\mathbf{x}'_i$ 的集是给定的. 点 $\bar{\mathbf{x}}_i$ 是固定的真值,而点 $\mathbf{x}'_i$ 是随机变量,它的每个分量受协方差为 $\sigma^2$ 或(如果有必要的话)更一般协方差的高斯噪声影响. 函数 $f:\mathbb{R}^9\rightarrow\mathbb{R}^{2n}$ 定义一个映射,它把表示矩阵 H 的 9 维矢量 $\mathbf{h}$ 映射到由点 $\mathbf{x}'_i=H\bar{\mathbf{x}}_i$ 的坐标组成的 $2n$ 维矢量上. $\mathbf{x}'_i$ 的坐标组成 $\mathbb{R}^N$ 中的一个组合矢量,我们把它记为 $\mathbf{X}'$. 如我们所看到的,当 $\mathbf{h}$ 变化时,点 $f(\mathbf{h})$ 画出 $\mathbb{R}^{2n}$ 上的一个 8 维曲面 $S_P$. 在曲面上的每一点 $\mathbf{X}'$ 表示一组与 第一幅图像上的点 $\bar{\mathbf{x}}_i$ 相关的点 $\mathbf{x}'_i$. 给定一个测量矢量 $\mathbf{X}'$,我们可以在 Mahalanobis 距离的意义下选择在曲面 $S_P$ 上与之最靠近的点 $\hat{\mathbf{X}}'$. 满足约束条件 $\|\mathbf{h}\|=1$ 的前象 $\mathbf{h}=f^{-1}(\hat{\mathbf{X}}')$ 表示用 ML 估计算法得到的估计单应矩阵 $\hat{H}$. 由 $\mathbf{X}'$ 值的概率分布,我们希望导出被估计的 $\hat{\mathbf{h}}$ 的分布. 其协方差矩阵 $\Sigma_{\mathbf{h}}$ 由结论4.12给出. 该协方差矩阵对应于约束 $\|\mathbf{h}\|=1$.

因此,计算一个被估计的变换的协方差矩阵的过程如下:

(1) 由给定数据估计变换 $\hat{H}$.

(2) 计算雅可比矩阵 $J_f=\partial\mathbf{X}'/\partial\mathbf{h}$ 在 $\hat{\mathbf{h}}$ 处的值.

(3) 估计 $\mathbf{h}$ 的协方差矩阵由(4.7)给出:$\Sigma_{\mathbf{h}}=(J_f^T\Sigma_{X'}^{-1}J_f)^+$.

我们将略微详细地研究该过程的最后两步.

**导数矩阵的计算** 首先考虑雅可比矩阵 $J=\partial\mathbf{X}'/\partial\mathbf{h}$. 该矩阵可以自然分解成分块形式,即 $J=(J_1^T,J_2^T,\cdots,J_i^T,\cdots,J_n^T)^T$,其中 $J_i=\partial\mathbf{x}'_i/\partial\mathbf{h}$. $\partial\mathbf{x}'_i/\partial\mathbf{h}$ 的公式在(3.21 – $p$83)中给过:

$$J_i=\partial\mathbf{x}'_i/\partial\mathbf{h}=\frac{1}{w'_i}\begin{bmatrix}\tilde{\mathbf{x}}_i^T & \mathbf{0}^T & -x'_i\tilde{\mathbf{x}}_i^T\\ \mathbf{0}^T & \tilde{\mathbf{x}}_i^T & -y'_i\tilde{\mathbf{x}}_i^T\end{bmatrix}\tag{4.8}$$

其中 $\tilde{\mathbf{x}}_i^T$ 表示矢量 $(x_i,y_i,1)$.

把所有点 $\mathbf{x}_i$ 的矩阵叠合起来给出导数矩阵 $\partial\mathbf{X}'/\partial\mathbf{h}$. 一种重要情形是图像测量 $\mathbf{x}'_i$ 是独立随机矢量. 此时有 $\Sigma=\mathrm{diag}(\Sigma_1,\ldots,\Sigma_n)$,其中每个 $\Sigma_i$ 是第 $i$ 个测量点 $\mathbf{x}'_i$ 的 2×2 协方差矩阵. 于是算得

$$\Sigma_{\mathbf{h}} = (J^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}J)^{+} = \left(\sum J_i^{\mathsf{T}}\Sigma_i^{-1}J_i\right)^{+}. \tag{4.9}$$

**例 4.13** 我们考虑一个简单的数值例子，它包含下面的 4 组点对应

$$\begin{aligned}
\mathbf{x}_1 = (1,0)^{\mathsf{T}} &\leftrightarrow (1,0)^{\mathsf{T}} = \mathbf{x}_1' \\
\mathbf{x}_2 = (0,1)^{\mathsf{T}} &\leftrightarrow (0,1)^{\mathsf{T}} = \mathbf{x}_2' \\
\mathbf{x}_3 = (-1,0)^{\mathsf{T}} &\leftrightarrow (-1,0)^{\mathsf{T}} = \mathbf{x}_3' \\
\mathbf{x}_4 = (0,-1)^{\mathsf{T}} &\leftrightarrow (0,-1)^{\mathsf{T}} = \mathbf{x}_4'
\end{aligned}$$

即在一个射影基点上的恒等映射. 我们假定点 $\mathbf{x}_i$ 准确地给定，而点 $\mathbf{x}_i'$ 在每个坐标方向有一个像素的标准差. 这意味着协方差矩阵 $\Sigma_{\mathbf{X}_i'}$ 是单位矩阵.

显然，被计算的单应将是恒等映射. 为简单起见，我们将它归一化(改变尺度)使它真的是恒等矩阵，因而令 $\|H\|^2=3$ 而不采用通常的归一化 $\|H\|=1$. 在这种情形，(4.8)中所有的 $w_i'$ 都等于 1. 由(4.8)不难得到矩阵 J：

$$J = \left[\begin{array}{ccc|ccc|ccc}
1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & -1 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
\hline
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & -1 \\
\hline
-1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 1 \\
0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\
\hline
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 & 1
\end{array}\right].$$

那么

$$J^{\mathsf{T}}J = \left[\begin{array}{ccc|ccc|ccc}
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \\
0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 4 & 0 & 0 & 0 & -2 & 0 & 0 \\
\hline
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & -2 \\
0 & 0 & 0 & 0 & 0 & 4 & 0 & -2 & 0 \\
\hline
0 & 0 & -2 & 0 & 0 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -2 & 0 & 2 & 0 \\
-2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 4
\end{array}\right]. \tag{4.10}$$

我们可以用(4.6)求该矩阵的伪逆，其中矩阵 A 的列生成约束曲面的切平面. 因为 H 在一个表示超球面的约束条件 $\|H\|^2=3$ 下被计算，故约束曲面与求得的单应 H 的矢量 $\mathbf{h}$ 相垂直. 对应矢量 $\mathbf{h}$ 的一个 Householder 矩阵 A(见节 A3.1.2($p$405)) 的性质是 $A\mathbf{h} = (0,\ldots,0,1)^{\mathsf{T}}$，因此 A 的前 8 列(记作 $A_1$)垂直于 $\mathbf{h}$，这正是我们希望得到的. 这样可以精确计算伪逆而不必采用 SVD. 应用(4.6)计算该伪逆得

$$\Sigma_{\mathbf{h}} = (J^T J)^{+A_1} = A_1 (A_1^T (J^T J) A_1)^{-1} A_1^T = \frac{1}{18}\left[\begin{array}{ccc|ccc|ccc} 5 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & -1 \\ 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 & 0 & 9 & 0 & 0 \\ \hline 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 9 & 0 & 9 & 0 \\ \hline 0 & 0 & 9 & 0 & 0 & 0 & 18 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 9 & 0 & 18 & 0 \\ -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 2 \end{array}\right] \tag{4.11}$$

对角线给出 H 的每个元素的方差. △

上面计算得到的方差可用来评价例 4.14 中点转移的准确性.

### 4.2.5 双图像误差

对误差在两幅图像的情形,变换的协方差的计算要复杂一点. 如在节 3.2.7($p62$)中所见,我们可以定义 $2n+8$ 个参数组成的集合,其中 8 个参数表示变换矩阵而 $2n$ 个参数 $\hat{\mathbf{x}}_i$ 表示在第一幅图像中点的估计. 我们可以更方便地用超参数化方法,以 9 个参数表示变换 H. 雅可比矩阵自然地分成形如 $J=[A|B]$ 的两部分,其中 A 和 B 分别是对摄像机参数和点 $\mathbf{x}_i$ 的导数. 应用(4.7)算得

$$J^T \Sigma_{\mathbf{x}}^{-1} J = \begin{bmatrix} A^T \Sigma_x^{-1} A & A^T \Sigma_x^{-1} B \\ B^T \Sigma_x^{-1} A & B^T \Sigma_x^{-1} B \end{bmatrix}.$$

该矩阵的伪逆是参数集的协方差,而该伪逆的左上角方块是 H 的元素的协方差. 关于它的详细讨论在节 A4.4.1($p422$)中给出,那里还说明如何利用雅可比的块结构来简化计算.

在前节由 4 点估计 H 的协方差的例 4.13 中,协方差矩阵的最终结果是 $\Sigma_{\mathbf{h}} = 2(J^T \Sigma_{\mathbf{x}}^{-1} J)^+$,即为单图像误差时的协方差矩阵的两倍. 这里假定两幅图像中所测量的点具有相同的协方差. 单图像和双图像的协方差之间的这种简单的关系一般不成立.

### 4.2.6 在点转移中应用协方差矩阵

一旦得到协方差矩阵,我们就可以计算一个给定点的转移的不可靠性. 考虑在第一幅图像中的一个没有被用于计算变换 H 的新点 $\mathbf{x}$. 它在第二幅图像中的对应点是 $\mathbf{x}' = H\mathbf{x}$. 但是,由于在估计 H 中存在不可靠性,$\mathbf{x}'$ 的正确位置也有相应的不可靠性. 我们可以由 H 的协方差矩阵来计算这个不可靠性.

点 $\mathbf{x}'$ 的协方差矩阵由下面的公式给出:

$$\Sigma_{\mathbf{x}'} = J_{\mathbf{h}} \Sigma_{\mathbf{h}} J_{\mathbf{h}}^T \tag{4.12}$$

其中 $J_{\mathbf{h}} = \partial \mathbf{x}' / \partial \mathbf{h}$. $\partial \mathbf{x}' / \partial \mathbf{h}$ 的公式在(3.21 – $p83$)中给出.

如果点 $\mathbf{x}$ 本身的测量也具有某种不确定性,那么在 $\mathbf{x}$ 和 $\mathbf{h}$ 间没有互相关的假设下,公式(4.12)可以用下面的公式代替:

$$\Sigma_{\mathbf{x}'} = J_{\mathbf{h}} \Sigma_{\mathbf{h}} J_{\mathbf{h}}^T + J_{\mathbf{x}} \Sigma_{\mathbf{x}} J_{\mathbf{x}}^T. \tag{4.13}$$

上面的假设是合理的,因为点 $\mathbf{x}$ 是一个没被用来计算变换 H 的新点. 关于雅可比矩阵的公

式 $J_x = \partial \mathbf{x}'/\partial \mathbf{x}$ 在(3.20 - $p$83)中给出.

(4.12)中的协方差矩阵 $\Sigma_{x'}$ 用变换 H 的协方差矩阵 $\Sigma_h$ 表示. 由(4.6)可知协方差矩阵 $\Sigma_h$ 依赖于估计 H 时所用的具体约束. 因此似乎 $\Sigma_{x'}$ 同样依赖于用于约束 H 的具体方法. 但是,可以验证这些公式与用来计算协方差矩阵 $\Sigma_P = (J^T \Sigma_X^{-1} J)^{+\wedge}$ 的具体约束 A 无关.

**例 4.14** 我们继续使用例 4.13,令 2D 单应 H 为有(4.11)中的方差 $\Sigma_h$ 的恒等矩阵. 考虑一个任意点 $(\mathbf{x},\mathbf{y})^T$ 被映射到点 $\mathbf{x}' = H\mathbf{x}$. 在此情形,协方差矩阵 $\Sigma_{x'} = J_h \Sigma_h J_h^T$ 可以用符号计算求得

$$\Sigma_{x'} = \begin{bmatrix} \sigma_{x'x'} & \sigma_{x'y'} \\ \sigma_{x'y'} & \sigma_{y'y'} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 2 - x^2 + x^4 + y^2 + x^2y^2 & xy(x^2 + y^2 - 2) \\ xy(x^2 + y^2 - 2) & 2 - y^2 + y^4 + x^2 + x^2y^2 \end{bmatrix}.$$

注意 $\sigma_{x'x'}$ 和 $\sigma_{y'y'}$ 是 $x$ 和 $y$ 的偶函数,而 $\sigma_{x'y'}$ 是奇函数. 这是因为计算 H 的点集具有关于 $x$ 和 $y$ 轴的对称性. 同时注意 $\sigma_{x'x'}$ 和 $\sigma_{y'y'}$ 的差别仅在于交换 $x$ 和 $y$,这也是因为定义点集的对称性.

可以看到变量 $\sigma_{x'x'}$ 以 $x$ 的 4 次方量级变化,从而标准差是 $x$ 的平方量级. 这说明如果外插所用的变换点 $\mathbf{x}' = H\mathbf{x}$ 离用于计算 H 的点集太远,那么这个值是不可靠的. 更具体地说,$\mathbf{x}'$ 位置的 RMS 的不可靠性等于 $(\sigma_{x'x'} + \sigma_{y'y'})^{1/2} = \sqrt{\mathrm{trace}(\Sigma_{x'})}$,可以求得它等于 $(1 + (x^2 + y^2)^2)^{1/2} = (1 + r^4)^{1/2}$,其中 $r$ 是离原点的径向距离. 注意一个有趣的现象:RMS 的误差仅取决于径向距离. 事实上可以证明点 $\mathbf{x}'$ 的概率分布仅取决于 $\mathbf{x}'$ 的径向距离,虽然它有两个主轴:指向径向和切向. 图 4.5 给出了 $\mathbf{x}'$ 的 RMS 误差作为 $r$ 的函数的图形. △

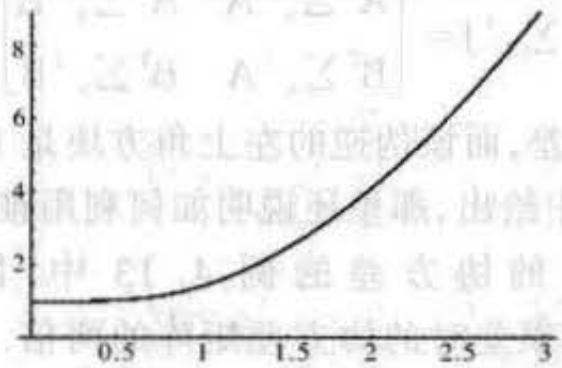



图 4.5 投影点 $\mathbf{x}'$ 位置的 RMS 误差是点 $\mathbf{x}'$ 到原点的径向距离的函数. 单应 H 由围绕原点的单位圆上等距的 4 点计算得到而且误差仅在第二幅图像上. RMS 误差与计算 H 时所采用的点的假设误差成比例,垂直轴根据这个假设误差来标定.

这个例子计算了在 4 组点对应的最小配置时一个转移点的协方差. 在多于 4 组对应时,情形并没有实质性的不同. 外插越过计算单应时所用的点集的范围是不可靠的. 事实上,我们可以证明如果 H 由单位圆周上等距的 $n$ 点(而不象上例用 4 点)算得,那么 RMS 误差的平方等于 $\sigma_{x'x'} + \sigma_{y'y'} = 4(1 + r^4)/n$,因此误差同样按二次方增长.

## 4.3 协方差估计的蒙特卡洛法

前几节所讨论的协方差估计方法都建立在线性假定之上. 换句话说,假定曲面 $f(\mathbf{h})$ 在被估计的点的附近是局部平的,至少在与噪声分布的近似程度相对应的区域上如此. 同时也假设变换 H 的估计方法是最大似然估计. 如果曲面不完全平,那么方差的估计可能不正确. 另外,一个具体估计方法的性能可能低于 ML 估计,因而给被估计的变换 H 引入额外的

不可靠性.

取得协方差估计的一般(虽然昂贵)方法是穷尽模拟. 假定噪声来自一个给定的噪声分布,我们从精确对应于给定变换的点匹配集开始. 然后对点加噪并用选定的估计过程计算相应的变换. 然后,按所假设的噪声分布进行多次试验,用统计方法计算变换 H 或进一步计算转移点的协方差. 图 4.6 对恒等映射的情况给予说明.

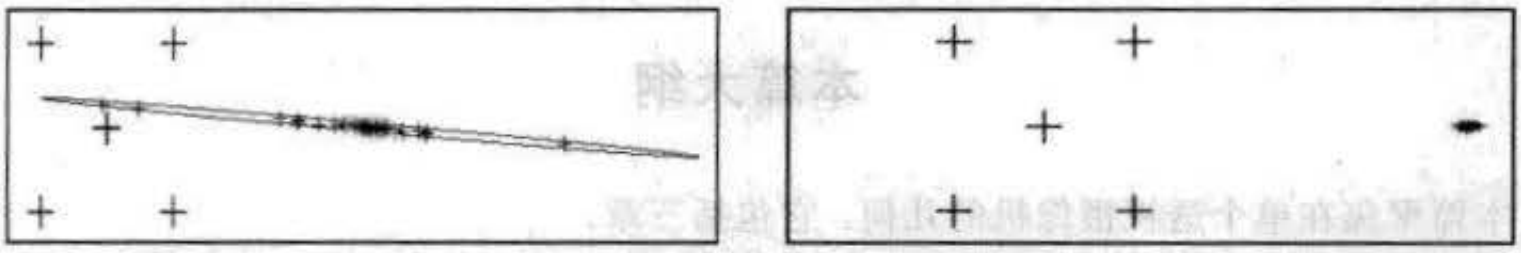


图 4.6 在归一和不归一化 DLT 算法中,点在恒等映射下的转移,同时参考图 3.4($p$68)以得到进一步的解释.

当我们不知道 H 的真值时,估计变换 H 的协方差的解析法和蒙特卡洛法都可以用于实际数据的协方差估计. 由给定的数据计算 H 的估计以及点 $\mathbf{x}_i'$ 和 $\mathbf{x}_i$ 对应的真值. 然后,把估计值当作匹配数据点和变换的真值来计算协方差. 得到的协方差矩阵就认为是真值变换的协方差. 这种等同的做法是基于假定数据点的真值与被估计的值非常接近,足以使协方差矩阵本质上不受影响.

## 4.4 结束语

### 4.4.1 文献

由于仅采用一阶泰勒展开和高斯误差分布的假定使本章的所有推导大为简化. 类似的思想(ML,协方差…)可以通过采用 Fisher 信息矩阵推广到其他的误差分布. 相关的文献可以在 Kanatani[Kanatani - 96]、Press 等[Press - 88]和其他统计书中找到.

Criminisi 等[Criminisi - 99]给出当确定单应的对应在数量和位置上变化时,计算点转移协方差的许多例子.

### 4.4.2 注释和练习

(1) 考虑用正交回归计算与平面上一组 2D 点最佳拟合的直线的问题. 假定 $N$ 个点的每一个坐标的测量具有独立标准差 $\sigma$. 每点到拟合直线的距离的 RMS 期望值是什么? **答案**:$\sigma((n-2)/n)^{1/2}$.

(2)(较难)节 17.4($p$316)给出一种由跨 $m$ 幅视图的 $n+4$ 组点对应计算射影重构的方法,其中 4 组点对应假定来自一张平面. 假定这 4 组平面点的对应精确地给定,并且其他的 $n$ 个图像点在测量时带有 1 个像素的误差(每幅图像的每个坐标). 问残差 $\|\mathbf{x}_j^i - \hat{\mathbf{P}}^i\mathbf{X}_j\|$ 的期望值是什么?

# 第 1 篇

# 摄像机几何和单视图几何

## 本篇大纲

本篇聚焦在单个透视摄像机的几何. 它包括三章.

第 5 章介绍 3D 景物空间到 2D 图像平面的投影. 摄像机映射由一个矩阵表示,在点映射的情况下它是一个 $3\times4$ 的矩阵 P,它把 3 维空间的一个世界点的齐次坐标映射到图像平面上影像点的齐次坐标. 这个矩阵一般有 11 个自由度,并且摄像机的一些性质,例如它的中心和焦距等,可以从中抽取. 具体地说,摄像机内部参数,例如焦距和长宽比,包含在一个由 P 经过简单分解得到的 $3\times3$ 矩阵 K 中. 有两类特别重要的摄像机矩阵:有限摄像机和中心在无穷远的摄像机(例如表示平行投影的**仿射摄像机**).

第 6 章介绍给定一组世界和图像的对应点的坐标求摄像机矩阵 P 的估计问题. 该章同时介绍如何把摄像机的约束有效地引入估计的方法,以及镜头径向失真的矫正方法.

第 7 章有三个主要议题. 第一,它包含摄像机对几何物体而不是有限点的作用. 这些几何物体包括直线、二次曲线、二次曲面和无穷远点. 无穷远点/线的图像是消影点/线. 第二个议题是摄像机标定,即在不求出整个矩阵 P 的情况下计算摄像机矩阵的内参数 K. 特别介绍内参数与绝对二次曲线的图像间的关系,以及由消影点和消影线来标定摄像机的方法. 最后的议题是**标定二次曲线**,这是实现可视化摄像机标定的一种简单的几何手段.

# 第5章 摄像机模型

摄像机是3D世界(物体空间)和2D图像之间的一种映射. 本书主要关心的摄像机是**中心投影**. 本章将推导若干种摄像机**模型**,它们是表征摄像机映射的具体性质的矩阵.

我们将看到所有中心投影的摄像机是**一般射影摄像机**的特殊情况. 我们将用射影几何的工具来研究这种最一般的摄像机模型的构造. 我们将会看到摄像机的几何元素(如投影中心和图像平面)可以非常简单地根据它的矩阵表示计算得到. 一般射影摄像机的内在性质(例如几何性质)可以用同样的代数表示式来计算.

具体模型主要分成两类——有限中心的摄像机模型和"无穷远"中心的摄像机模型. 在无穷远摄像机中,**仿射摄像机**特别重要,因为它是平行投影的自然推广.

本章主要讨论点的投影. 摄像机对其他几何元素(例如直线)的作用将推迟到第7章中讨论.

## 5.1 有限摄像机

本节,我们从最具体和最简单的摄像机模型即基本的针孔摄像机开始,然后通过一系列的升级逐步把这个模型一般化.

我们推导的模型主要针对CCD类型的传感器,但同样适用于其他摄像机,例如X-射线图像、扫描负片、扫描放大负片等.

**基本针孔模型** 我们考虑空间点到一张平面上的中心投影. 令投影中心位于一个欧氏坐标系的原点,而平面 $Z=f$ 被称为**图像平面**或**聚焦平面**. 在针孔摄像机模型下,空间坐标为 $\mathbf{X}=(X,Y,Z)^{\mathsf{T}}$ 的点 $\mathbf{X}$ 被映射到图像平面上的一点,该点是连接点 $\mathbf{X}$ 与投影中心的直线与图像平面的交点. 图5.1对此作了说明. 根据相似三角形,可以很快地算出点 $(X,Y,Z)^{\mathsf{T}}$ 被映射到图像平面上点 $(fX/Z,fY/Z,f)^{\mathsf{T}}$. 略去最后一个图像坐标之后,从世界坐标到图像坐标的中心投影是

$$(X,Y,Z)^{\mathsf{T}}\mapsto(fX/Z,fY/Z)^{\mathsf{T}} \tag{5.1}$$

这是从3维欧氏空间 $\mathbb{R}^3$ 到2维欧氏空间 $\mathbb{R}^2$ 的一个映射.

投影中心称为**摄像机中心**,也称为**光心**. 摄像机中心到图像平面的垂线称为摄像机的**主轴**或**主射线**,而主轴与图像平面的交点称为**主点**. 过摄像机中心平行于图像平面的平面称为摄像机的**主平面**.

**用齐次坐标表示中心投影** 如果用齐次矢量表示世界和图像点,那么中心投影可以非常简单地表示成齐次坐标之间的线性映射. 具体地说,(5.1)可以写成如下矩阵乘积形式

$$\begin{pmatrix}X\\Y\\Z\\1\end{pmatrix}\mapsto\begin{pmatrix}fX\\fY\\Z\end{pmatrix}=\begin{bmatrix}f&&&0\\&f&&0\\&&1&0\end{bmatrix}\begin{pmatrix}X\\Y\\Z\\1\end{pmatrix}. \tag{5.2}$$

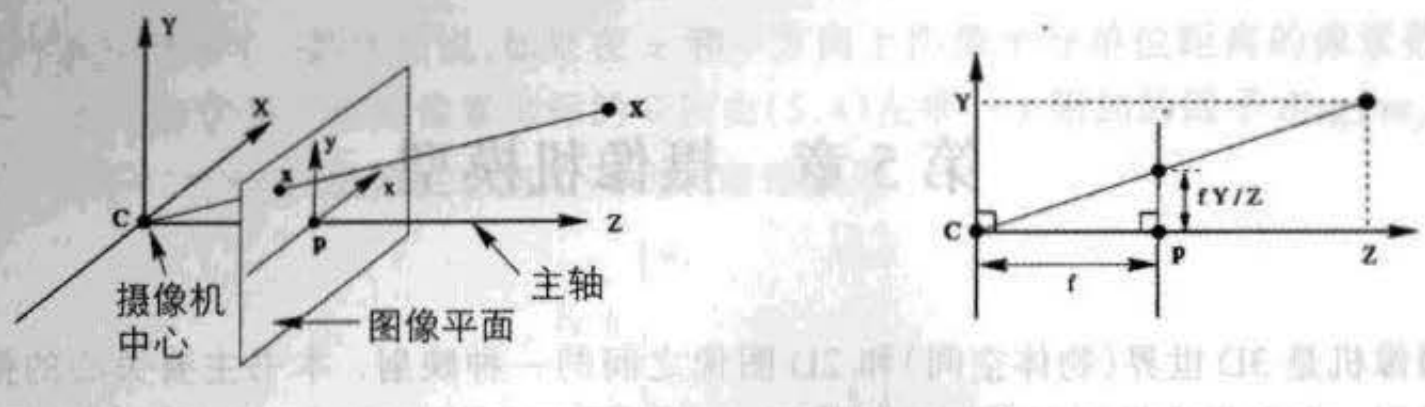



图 5.1 **针孔摄像机几何**. C 是摄像机中心而 p 是主点. 这里的摄像机中心位于坐标原点. 注意图像平面位于摄像机中心之前.

该表达式中的矩阵可以写成 $\mathrm{diag}(f,f,1)[\mathrm{I}|\mathbf{0}]$,其中 $\mathrm{diag}(f,f,1)$是对角矩阵,而$[\mathrm{I}|\mathbf{0}]$表示矩阵分块成一个 $3\times3$ 恒等矩阵加上一个零列矢量.

我们现在引入如下记号:世界点 $\mathbf{X}$ 用 4 维齐次矢量$(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}$ 表示;图像点 $\mathbf{x}$ 被表示成 3 维齐次矢量的形式;P 表示 $3\times4$ 齐次摄像机投影矩阵. 这样一来(5.2)可以紧凑地写为

$$\mathbf{x}=\mathrm{P}\mathbf{X}$$

它定义了中心投影的针孔模型的摄像机矩阵为

$$\mathrm{P}=\mathrm{diag}(f,f,1)[\mathrm{I}|\mathbf{0}].$$

**主点偏置** (5.1)式假定图像平面的坐标原点在主点上. 实际情况可能不是这样,因此一般情形的映射为

$$(\mathrm{X},\mathrm{Y},\mathrm{Z})^{\mathrm{T}}\mapsto(f\mathrm{X}/\mathrm{Z}+p_x,f\mathrm{Y}/\mathrm{Z}+p_y)^{\mathrm{T}}$$

其中$(p_x,p_y)^{\mathrm{T}}$ 是主点的坐标. 参看图 5.2. 该方程用齐次坐标可以表示成

$$\begin{pmatrix}\mathrm{X}\\\mathrm{Y}\\\mathrm{Z}\\1\end{pmatrix}\mapsto\begin{pmatrix}f\mathrm{X}+\mathrm{Z}p_x\\f\mathrm{Y}+\mathrm{Z}p_y\\\mathrm{Z}\end{pmatrix}=\begin{bmatrix}f&&p_x&0\\&f&p_y&0\\&&1&0\end{bmatrix}\begin{pmatrix}\mathrm{X}\\\mathrm{Y}\\\mathrm{Z}\\1\end{pmatrix}. \tag{5.3}$$

若记

$$\mathrm{K}=\begin{bmatrix}f&&p_x\\&f&p_y\\&&1\end{bmatrix} \tag{5.4}$$

则(5.3)有一个简洁的形式

$$\mathbf{x}=\mathrm{K}[\mathrm{I}|\mathbf{0}]\mathbf{X}_{\mathrm{cam}}. \tag{5.5}$$

矩阵 K 称为**摄像机标定矩阵**. 在(5.5)中我们记$(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}$ 为 $\mathbf{X}_{\mathrm{cam}}$是为了强调摄像机被设定在一个欧氏坐标系的原点且主轴沿着 Z 轴的指向,而点 $\mathbf{X}_{\mathrm{cam}}$按此坐标系表示. 这样的坐标系可以称为**摄像机坐标系**.

**摄像机旋转与位移** 一般,空间点采用不同的欧氏坐标系表示,称为**世界坐标系**. 两个坐标系通过旋转和平移相联系. 参看图 5.3. 如果 $\widetilde{\mathbf{X}}$是一个 3 维非齐次矢量,表示世界坐标系中一点的坐标,而 $\widetilde{\mathbf{X}}_{\mathrm{cam}}$是以摄像机坐标系来表示的同一点,那么我们可以记$\widetilde{\mathbf{X}}_{\mathrm{cam}}=\mathrm{R}(\widetilde{\mathbf{X}}-\widetilde{\mathbf{C}})$,其中 $\widetilde{\mathbf{C}}$表示摄像机中心在世界坐标系中的坐标,R 是一个 $3\times3$ 的旋转矩阵,表示摄像机坐

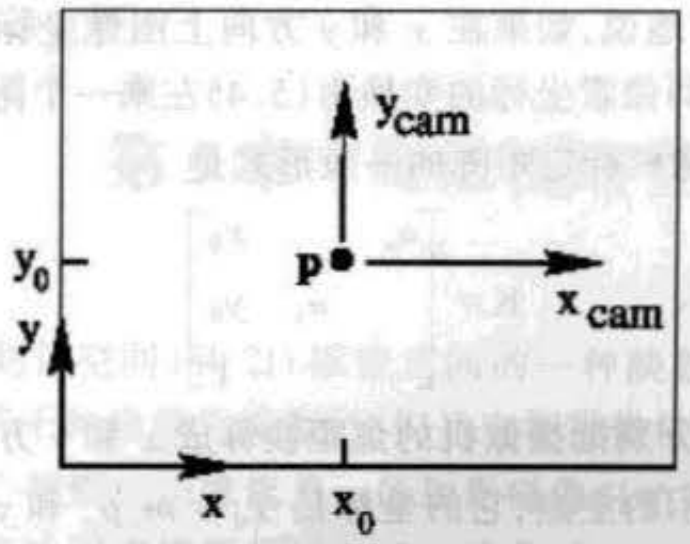



图 5.2 图像坐标系$(x,y)^T$和摄像机坐标系$(x_{cam},y_{cam})^T$.

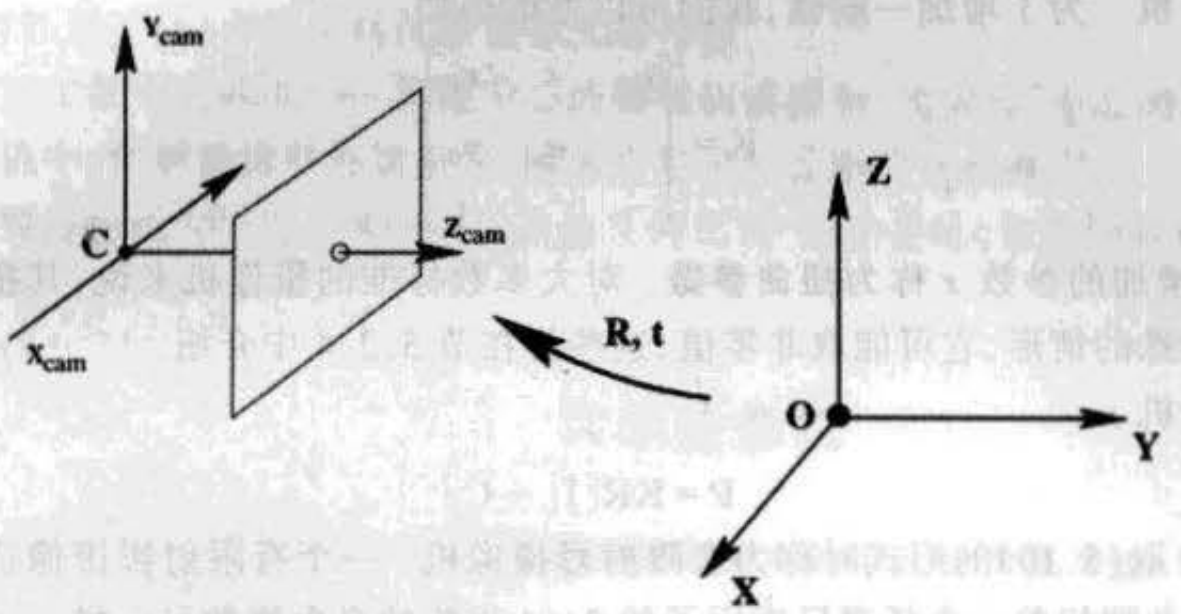



图 5.3 世界和摄像机坐标架之间的欧氏变换.

标系的方位. 这个方程在齐次坐标下可以写成

$$\mathbf{X}_{cam}=\begin{bmatrix}\mathrm{R} & -\mathrm{R}\widetilde{\mathbf{C}}\\ \mathbf{0}^T & 1\end{bmatrix}\begin{pmatrix}X\\ Y\\ Z\\ 1\end{pmatrix}=\begin{bmatrix}\mathrm{R} & -\mathrm{R}\widetilde{\mathbf{C}}\\ \mathbf{0}^T & 1\end{bmatrix}\mathbf{X}. \tag{5.6}$$

把它与(5.5)结合起来形成公式

$$\mathbf{x}=\mathrm{KR}[\mathrm{I}|-\widetilde{\mathbf{C}}]\mathbf{X} \tag{5.7}$$

其中 $\mathbf{X}$ 用世界坐标系表示. 这是由一个针孔模型给出的一般映射. 我们看到一个一般的针孔摄像机 $\mathrm{P}=\mathrm{KR}[\mathrm{I}|-\widetilde{\mathbf{C}}]$有 9 个自由度:3 个来自 K(元素 $f,p_x,p_y$),3 个来自 R,3 个来自 $\widetilde{\mathbf{C}}$. 包含在 K 中的参数称为摄像机**内部参数**或摄像机的**内部校准**. 包含在 R 和 $\widetilde{\mathbf{C}}$中的参数与摄像机在世界坐标系中的方位和位置有关并称为**外部参数**或**外部校准**.

为方便起见,通常摄像机中心不明显标出,而把世界到图像的变换表示成 $\widetilde{\mathbf{X}}_{cam}=\mathrm{R}\,\widetilde{\mathbf{X}}+\mathbf{t}$. 在此情形时摄像机矩阵简化成

$$\mathrm{P}=\mathrm{K}[\mathrm{R}|\mathbf{t}] \tag{5.8}$$

其中根据(5.7),$\mathbf{t}=-\mathrm{R}\,\widetilde{\mathbf{C}}$.

**CCD 摄像机** 刚才推导的针孔模型假定图像坐标在两个轴向上有等尺度的欧氏坐标. 但 CCD 摄像机的像素可能不是正方形. 如果图像坐标以像素来测量,那么需要在每个方向上

引入非等量尺度因子. 具体地说,如果在 $x$ 和 $y$ 方向上图像坐标单位距离的像素数分别是 $m_x$ 和 $m_y$,那么由世界坐标到像素坐标的变换由(5.4)左乘一个附加的因子 $\mathrm{diag}(m_x, m_y, 1)$ 而得到. 因此一个 CCD 摄像机标定矩阵的一般形式是

$$\mathrm{K}=\begin{bmatrix}\alpha_x & & x_0\\ & \alpha_y & y_0\\ & & 1\end{bmatrix} \tag{5.9}$$

其中 $\alpha_x = fm_x$ 和 $\alpha_y = fm_y$ 分别把摄像机的焦距换算成 $x$ 和 $y$ 方向的像素量纲. 同理,$\tilde{\mathbf{x}}_0 = (x_0, y_0)^{\mathrm{T}}$ 是用像素量纲表示的主点,它的坐标是 $x_0 = m_x p_x$ 和 $y_0 = m_y p_y$. 因此,一个 CCD 摄像机有 10 个自由度.

**有限射影摄像机** 为了增加一般性,我们可以考虑形如

$$\mathrm{K}=\begin{bmatrix}\alpha_x & s & x_0\\ & \alpha_y & y_0\\ & & 1\end{bmatrix} \tag{5.10}$$

的标定矩阵,增加的参数 $s$ 称为**扭曲参数**. 对大多数标准的摄像机来说,其扭曲参数为零. 但是在某些特殊的情形,它可能取非零值,这些将在节 5.2.4 中介绍.

一个摄像机

$$\mathrm{P}=\mathrm{KR}[\mathrm{I}|-\tilde{\mathbf{C}}] \tag{5.11}$$

的标定矩阵 K 取(5.10)的形式时称为**有限射影摄像机**. 一个有限射影摄像机有 11 个自由度. 与一个定义到相差一个任意尺度因子的 $3\times4$ 矩阵的自由度数目一样.

注意 P 左边的 $3\times3$ 子矩阵等于 KR,且是非奇异的. 因此,如果一个 $3\times4$ 的矩阵 P 的左边 $3\times3$ 子矩阵是非奇异的,则它是某个有限射影摄像机的摄像机矩阵,因为 P 可以分解成 $\mathrm{P}=\mathrm{KR}[\mathrm{I}|-\tilde{\mathbf{C}}]$. 事实上,令 M 为 P 的左边 $3\times3$ 子矩阵,我们可以将 M 分解为乘积 $\mathrm{M}=\mathrm{KR}$,其中 K 是形如(5.10)的上三角矩阵而 R 是旋转矩阵. 该分解本质上是节A3.1.1 ($p$404)中介绍的 RQ 矩阵分解,其中更多的内容将在节 5.2.4 中说明. 于是,矩阵 P 可以写成 $\mathrm{P}=\mathrm{M}[\mathrm{I}|\mathrm{M}^{-1}\mathbf{p}_4]=\mathrm{KR}[\mathrm{I}|-\tilde{\mathbf{C}}]$,其中 $\mathbf{p}_4$ 是 P 的最后一列. 简而言之:

- **把有限射影摄像机矩阵与左边为非奇异的 $3\times3$ 子矩阵的 $3\times4$ 齐次矩阵等同.**

**一般射影摄像机** 在射影摄像机层次化的最后一步是移去加在左边 $3\times3$ 子矩阵的非奇异性约束. **一般射影**摄像机由秩为 3 的任意 $3\times4$ 齐次矩阵表示. 它有 11 个自由度. 之所以要求秩 3 是因为如果秩小于 3,那么矩阵映射的值域将是一条直线或一个点而不是整张平面,换句话说不是 2D 图像.

## 5.2 射影摄像机

一般射影摄像机 P 按公式 $\mathbf{x}=\mathrm{P}\mathbf{X}$ 把世界点 $\mathbf{X}$ 映射到图像点 $\mathbf{x}$. 在这个映射的基础上,我们将分解该摄像机模型以揭示诸如摄像机中心一类的几何元素是如何在此矩阵中编码的. 在我们考虑的性质中,某些性质仅适用于有限射影摄像机和它们的特殊情况,而其余的则将适用于一般摄像机. 它们在上下文中很容易区分. 摄像机的有关性质概括在表 5.1 中.

| | |
|---|---|
| **摄像机中心** | 摄像机中心 $\mathbf{C}$ 是 P 的一维右零空间，即 $\mathrm{P}\mathbf{C}=\mathbf{0}$. |
| ◇ **有限摄像机**(M 非奇异) | $\mathbf{C}=\begin{pmatrix} -\mathrm{M}^{-1}\mathbf{p}_4 \\ 1 \end{pmatrix}$ |
| ◇ **无穷远摄像机**(M 奇异) | $\mathbf{C}=\begin{pmatrix} \mathbf{d} \\ 0 \end{pmatrix}$ 其中，$\mathbf{d}$ 是 M 的 3 维零矢量，即 $\mathrm{M}\mathbf{d}=\mathbf{0}$. |
| **列点** | 对于 $i=1,\cdots,3$，列矢量 $\mathbf{p}_i$ 分别对应于 X，Y，Z 轴在图像上的消影点. 列 $\mathbf{p}_4$ 是坐标原点的图像. |
| **主平面** | 摄像机的主平面是 P 的最后一行 $\mathbf{P}^3$. |
| **轴平面** | 平面 $\mathbf{P}^1$ 和 $\mathbf{P}^2$(P 的第一和第二行)表示空间中过摄像机中心的平面，分别对应于映射到图像上直线 $x=0$ 和 $y=0$ 的点. |
| **主点** | 图像点 $\mathbf{x}_0=\mathrm{M}\mathbf{m}^3$ 是摄像机的主点，其中 $\mathbf{m}^{3\mathrm{T}}$ 是 M 的第三行. |
| **主射线** | 摄像机的主射线(主轴)是过摄像机中心 $\mathbf{C}$ 而方向矢量为 $\mathbf{m}^{3\mathrm{T}}$ 的射线. 主轴矢量 $\mathbf{v}=\det(\mathrm{M})\mathbf{m}^3$ 指向摄像机的前方. |

表 5.1　射影摄像机 P 性质小结. 该矩阵表示为分块形式 $\mathrm{P}=[\mathrm{M}|\mathbf{p}_4]$.

### 5.2.1　摄像机构造

一般射影摄像机可以按 $\mathrm{P}=[\mathrm{M}|\mathbf{p}_4]$ 分块，其中 M 是 $3\times3$ 矩阵. 我们将看到：如果 M 是非奇异的，那么它是有限摄像机；反之则不然.

**摄像机中心**　矩阵 P 有一个 1 维右零空间，因为它有 4 列而秩是 3. 假定该零空间由 4 维矢量 $\mathbf{C}$ 生成，即 $\mathrm{P}\mathbf{C}=\mathbf{0}$. 我们将证明 $\mathbf{C}$ 是用齐次 4 维矢量表示的摄像机中心.

考察包含 $\mathbf{C}$ 和 3 维空间中任何一点 $\mathbf{A}$ 的直线. 该直线上的点可以表示为

$$\mathbf{X}(\lambda)=\lambda\mathbf{A}+(1-\lambda)\mathbf{C}.$$

在映射 $\mathbf{x}=\mathrm{P}\mathbf{X}$ 下，此线上的点被投影到

$$\mathbf{x}=\mathrm{P}\mathbf{X}(\lambda)=\lambda\mathrm{P}\mathbf{A}+(1-\lambda)\mathrm{P}\mathbf{C}=\lambda\mathrm{P}\mathbf{A}$$

之所以得到最后一步是因为 $\mathrm{P}\mathbf{C}=\mathbf{0}$. 上式表明该直线的所有点都映射到同一个图像点 $\mathrm{P}\mathbf{A}$，因而该直线必然是过摄像机中心的一条射线. 由此推出，$\mathbf{C}$ 是摄像机中心的齐次表示，因为对于 $\mathbf{A}$ 的所有选择，直线 $\mathbf{X}(\lambda)$ 都是通过摄像机中心的一条射线.

因为图像点 $(0,0,0)^{\mathrm{T}}=\mathrm{P}\mathbf{C}$ 无定义，因而摄像机中心是空间中唯一的图像没有定义的点，但这并不影响我们求到摄象机中心. 在有限摄像机的情形下，结果可以直接建立，因为 $\mathbf{C}=(\tilde{\mathbf{C}}^{\mathrm{T}},1)^{\mathrm{T}}$ 显然是 $\mathrm{P}=\mathrm{KR}[\mathrm{I}|-\tilde{\mathbf{C}}]$ 的零矢量. 这结论甚至在 P 的第一个 $3\times3$ 子矩阵 M 是奇异时也正确. 当然在奇异的情形，零矢量的形式为 $\mathbf{C}=(\mathbf{d}^{\mathrm{T}},0)^{\mathrm{T}}$，其中 $\mathrm{M}\mathbf{d}=\mathbf{0}$. 此时摄像机中心在无穷远点. 这类摄像机模型将在节 5.3 中讨论.

**列矢量**　射影摄像机的列是 3 维矢量，它们的几何涵义是特殊的图像点. 记 P 的列为 $\mathbf{p}_i$，$i=1,\ldots,4$，那么 $\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3$ 分别是世界坐标 X，Y 和 Z 轴的消影点，因为这些点是轴方向的图像. 例如 X－轴的方向 $\mathbf{D}=(1,0,0,0)^{\mathrm{T}}$，被映像到 $\mathbf{p}_1=\mathrm{P}\mathbf{D}$. 参见图 5.4. 列 $\mathbf{p}_4$ 是世界原点的图像.

**行矢量**　射影摄像机(5.12)的行是 4 维矢量，在几何上解释成特殊的世界平面. 这些平面将在下文讨论. 我们引入 P 的行记号 $\mathbf{P}^{i\mathrm{T}}$，因而

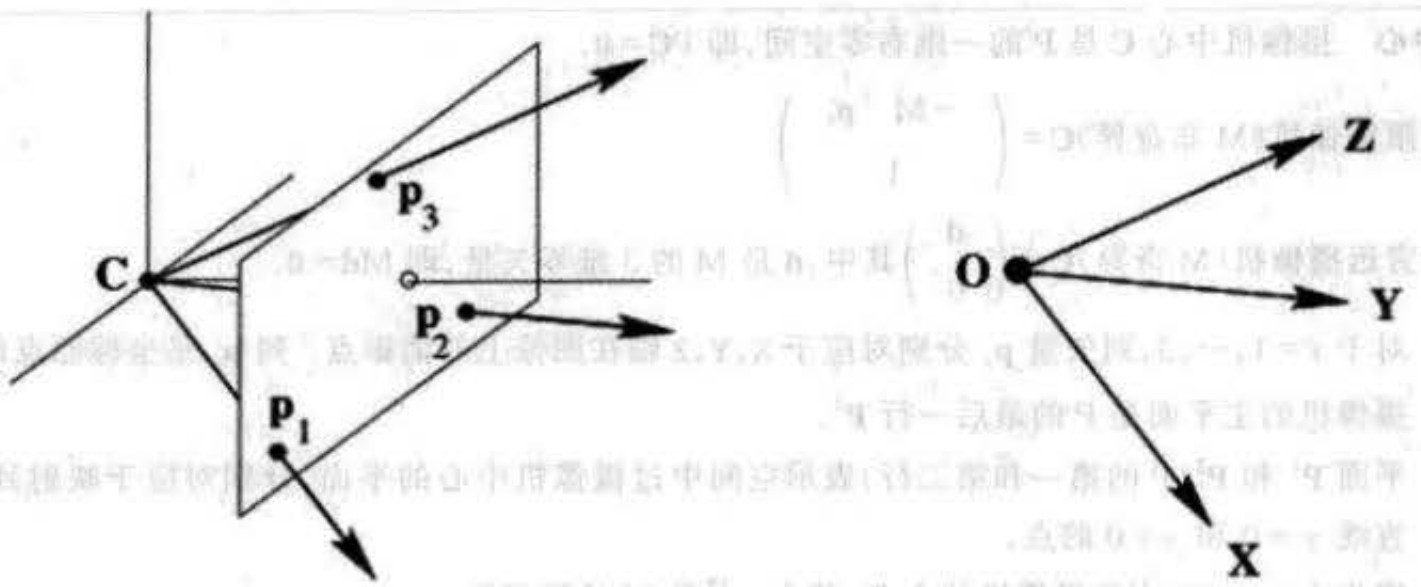



图 5.4　射影矩阵的列 $\mathbf{p}_i, i=1,\cdots,3$ 定义的三个图像点是世界坐标轴的方向的消影点.

$$\mathrm{P}=\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}=\begin{bmatrix} \mathbf{P}^{1\mathrm{T}} \\ \mathbf{P}^{2\mathrm{T}} \\ \mathbf{P}^{3\mathrm{T}} \end{bmatrix}. \tag{5.12}$$

**主平面**　主平面是过摄像机中心并平行于图像平面的平面. 它由被影像到图像上无穷远直线的点集 $\mathbf{X}$ 组成. 更明确的表示就是 $\mathrm{P}\mathbf{X}=(x,y,0)^{\mathrm{T}}$. 因此一个点在摄像机主平面上的充要条件是 $\mathbf{P}^{3\mathrm{T}}\mathbf{X}=0$. 换句话说, $\mathbf{P}^3$ 是摄像机主平面的矢量表示. 如果 $\mathbf{C}$ 是摄像机中心, 那么 $\mathrm{P}\mathbf{C}=\mathbf{0}$, 当然有 $\mathbf{P}^{3\mathrm{T}}\mathbf{C}=0$. 因而 $\mathbf{C}$ 在摄像机的主平面上.

**轴平面**　考察在平面 $\mathbf{P}^1$ 上的点集 $\mathbf{X}$. 该集合满足 $\mathbf{P}^{1\mathrm{T}}\mathbf{X}=0$, 因此被影像到图像 $\mathrm{P}\mathbf{X}=(0,y,w)^{\mathrm{T}}$ 处, 它们是图像 $y$ - 轴上的点. 此外由 $\mathrm{P}\mathbf{C}=\mathbf{0}$ 得 $\mathbf{P}^{1\mathrm{T}}\mathbf{C}=0$, 因而 $\mathbf{C}$ 也在平面 $\mathbf{P}^1$ 上. 其结果是平面 $\mathbf{P}^1$ 由摄像机中心和图像中的直线 $x=0$ 来定义. 类似地, 平面 $\mathbf{P}^2$ 由摄像机中心和直线 $y=0$ 来定义.

与主平面 $\mathbf{P}^3$ 不同, 轴平面 $\mathbf{P}^1$ 和 $\mathbf{P}^2$ 依赖于图像的 $x$ - 和 $y$ - 轴, 即与图像坐标系的选择有关. 因而, 它们与摄像机几何的本质耦合不如主平面那样紧密. 具体地说, 平面 $\mathbf{P}^1$ 和 $\mathbf{P}^2$ 的交线是一条连接摄像机中心和图像原点的直线, 即图像原点的反向投影. 该直线一般不与摄像机的主轴重合. 由 $\mathbf{P}^i$ 产生的平面在图 5.5 中说明.

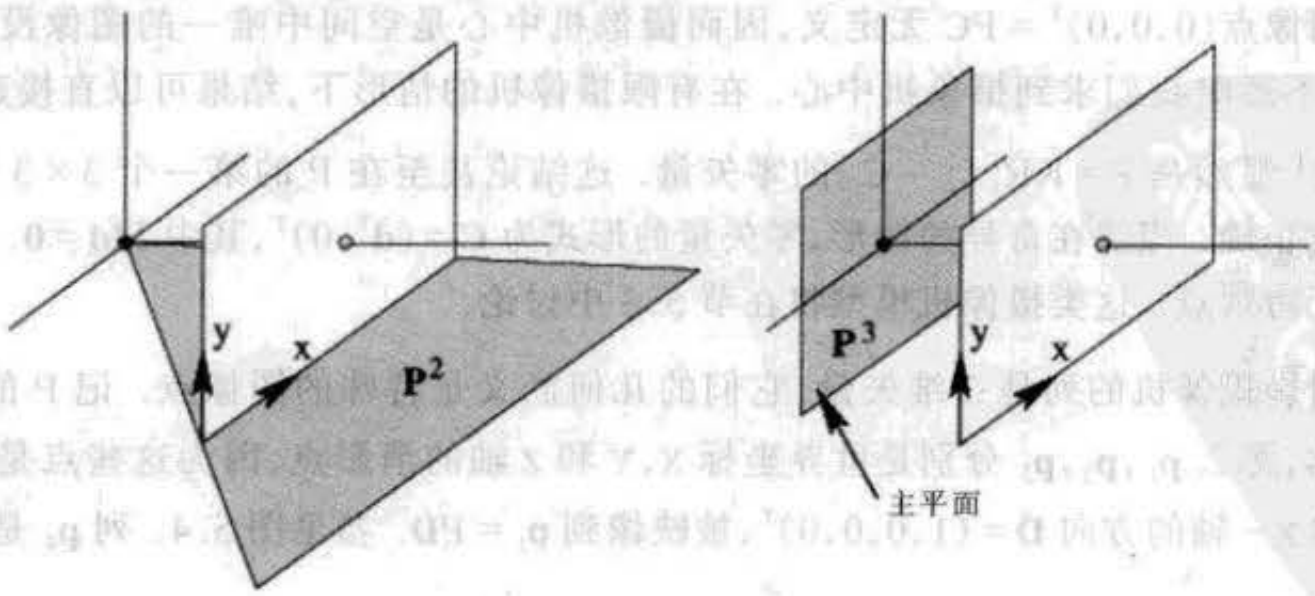



图 5.5　由射影矩阵的行定义的三张平面中的两张平面.

摄像机中心 $\mathbf{C}$ 位于所有三张平面上, 又由于这些平面两两不同(因为 $\mathbf{P}$ 矩阵的秩为 3),

故它必定是三平面的交点. 从代数上说,中心在所有三张平面上的条件是 $\mathbf{PC}=\mathbf{0}$,这是上面给出的关于摄像机中心的原始方程.

**主点** 主轴是过摄像机中心 $\mathbf{C}$ 并且方向垂直于主平面 $\mathbf{P}^3$ 的直线. 主轴与图像平面交于主点. 我们可以用下面的方法来确定主点. 一般来说平面 $\boldsymbol{\pi}=(\pi_1,\pi_2,\pi_3,\pi_4)^\mathsf{T}$ 的法线是一个矢量$(\pi_1,\pi_2,\pi_3)^\mathsf{T}$. 它也可以表示为无穷远平面上的一个点$(\pi_1,\pi_2,\pi_3,0)^\mathsf{T}$. 在摄像机的主平面为 $\mathbf{P}^3$ 时,该点是$(p_{31},p_{32},p_{33},0)^\mathsf{T}$,我们把它记为 $\hat{\mathbf{P}}^3$. 该点被摄像机矩阵 P 投影到摄像机的主点 $\mathrm{P}\hat{\mathbf{P}}^3$. 注意该公式仅仅涉及 $\mathrm{P}=[\mathrm{M}|\mathbf{p}_4]$左边的 $3\times3$ 子矩阵. 事实上主点由 $\mathbf{x}_0=\mathrm{M}\mathbf{m}^3$来计算,其中 $\mathbf{m}^{3\mathsf{T}}$是 M 的第三行.

**主轴矢量** 根据 $\mathbf{x}=\mathrm{P}\mathbf{X}$,虽然任何不在主平面上的点 $\mathbf{X}$ 都可以被映射为一个图像点,实际上仅在摄像机前面的点即空间中一半的点能在图像上见到. 令 $\mathrm{P}=[\mathrm{M}|\mathbf{p}_4]$. 我们刚才说过矢量 $\mathbf{m}^3$ 指向主轴方向. 我们习惯于定义该矢量指向摄像机的前方(正方向). 但是,注意 P 的定义可以相差一个符号. 这就产生了多义问题:是 $\mathbf{m}^3$ 还是 $-\mathbf{m}^3$ 指向正向. 我们现在就来解决这个多义问题.

我们从考察摄像机坐标系下的坐标着手. 根据(5.5),一个 3D 点到图像上一个点的投影方程是 $\mathbf{x}=\mathrm{P}_{\mathrm{cam}}\mathbf{X}_{\mathrm{cam}}=\mathrm{K}[\mathrm{I}|\mathbf{0}]\mathbf{X}_{\mathrm{cam}}$,这里 $\mathbf{X}_{\mathrm{cam}}$是用摄像机坐标表示的 3D 点. 在这种情形,可以看到矢量 $\mathbf{v}=\det(\mathrm{M})\mathbf{m}^3=(0,0,1)^\mathsf{T}$ 在主轴方向上并**指向摄像机的前方**,而不受 $\mathrm{P}_{\mathrm{cam}}$的尺度因子符号的影响. 例如,如果 $\mathrm{P}_{\mathrm{cam}}\rightarrow k\mathrm{P}_{\mathrm{cam}}$那么 $\mathbf{v}\rightarrow k^4\mathbf{v}$ 方向相同.

如果 3D 点用世界坐标表示,那么 $\mathrm{P}=k\mathrm{K}[\mathrm{R}|-\mathrm{R}\tilde{\mathbf{C}}]=[\mathrm{M}|\mathbf{p}_4]$,其中 $\mathrm{M}=k\mathrm{KR}$. 因为 $\det(\mathrm{R})>0$,矢量 $\mathbf{v}=\det(\mathrm{M})\mathbf{m}^3$ 同样不受尺度因子的影响. 概括起来:

- $\mathbf{v}=\det(\mathrm{M})\mathbf{m}^3$ **是在主轴方向上指向摄像机前方的矢量.**

### 5.2.2 射影摄像机对点的作用

**正向投影** 我们已经知道,一般的射影摄像机根据映射 $\mathbf{x}=\mathrm{P}\mathbf{X}$ 把空间的一个点 $\mathbf{X}$ 映射到一个图像点. 在无穷远平面上的点 $\mathbf{D}=(\mathbf{d}^\mathsf{T},\mathbf{0})^\mathsf{T}$ 表示消影点. 这些点映射到

$$\mathbf{x}=\mathrm{P}\mathbf{D}=[\mathrm{M}|\mathbf{p}_4]\mathbf{D}=\mathrm{M}\mathbf{d}$$

因而仅仅受 P 的前 $3\times3$ 子矩阵 M 的影响.

**点到射线的反向投影** 给定图像中的一个点 $\mathbf{x}$,我们来确定是空间的哪些点被映射到该点. 这些点将组成过摄像机中心的一条空间射线. 该射线的形式有若干种表示方法,与我们如何表示 3 维空间的一条直线有关. 其中的 Plücker 表示要推迟到节7.1.2($p$135)中去介绍. 而这里,直线被表示成两点的连接.

射线上有两个点是已知. 它们是摄像机中心 $\mathbf{C}(\mathrm{P}\mathbf{C}=\mathbf{0})$和点 $\mathrm{P}^+\mathbf{x}$,其中 $\mathrm{P}^+$ 是 P 的伪逆. P 的伪逆是矩阵 $\mathrm{P}^+=\mathrm{P}^\mathsf{T}(\mathrm{P}\mathrm{P}^\mathsf{T})^{-1}$并满足 $\mathrm{P}\mathrm{P}^+=\mathrm{I}$(参见节 A3.4($p$410)). 点 $\mathrm{P}^+\mathbf{x}$ 在射线上,因为它投影到 $\mathbf{x}$,即 $\mathrm{P}(\mathrm{P}^+\mathbf{x})=\mathrm{I}\mathbf{x}=\mathbf{x}$. 那么射线由这两点的连接形成

$$\mathbf{X}(\lambda)=\mathrm{P}^+\mathbf{x}+\lambda\mathbf{C}. \tag{5.13}$$

在有限摄像机情形可以导出另一种表示式. 记 $\mathrm{P}=[\mathrm{M}|\mathbf{p}_4]$,摄像机中心为 $\tilde{\mathbf{C}}=-\mathrm{M}^{-1}\mathbf{p}_4$. 一个图像点 $\mathbf{x}$ 反向投影的射线交无穷远平面于点 $\mathbf{D}=((\mathrm{M}^{-1}\mathbf{x})^\mathsf{T},0)^\mathsf{T}$,因而 $\mathbf{D}$ 提供了射线上

的第二个点. 同样将射线写成其上的两点的连接

$$\mathbf{X}(\mu)=\mu\begin{bmatrix}\mathrm{M}^{-1}\mathbf{x}\\0\end{bmatrix}+\begin{bmatrix}-\mathrm{M}^{-1}\mathbf{p}_4\\1\end{bmatrix}=\begin{bmatrix}\mathrm{M}^{-1}(\mu\mathbf{x}-\mathbf{p}_4)\\1\end{bmatrix}. \tag{5.14}$$

### 5.2.3 点的深度

下面,我们考虑在摄像机主平面前或后的一个点离主平面的距离. 设摄像机矩阵 $\mathrm{P}=[\mathrm{M}|\mathbf{p}_4]$,把 3 维空间的点 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}=(\widetilde{\mathbf{X}}^{\mathrm{T}},1)^{\mathrm{T}}$ 投影到图像点 $\mathbf{x}=w(x,y,1)^{\mathrm{T}}=\mathrm{P}\mathbf{X}$. 令 $\mathbf{C}=(\widetilde{\mathbf{C}},1)^{\mathrm{T}}$ 为摄像机中心. 那么 $w=\mathbf{P}^{3\mathrm{T}}\mathbf{X}=\mathbf{P}^{3\mathrm{T}}(\mathbf{X}-\mathbf{C})$(因为 $\mathrm{P}\mathbf{C}=\mathbf{0}$),其中 $\mathbf{C}$ 为摄像机中心. 但是 $\mathbf{P}^{3\mathrm{T}}(\mathbf{X}-\mathbf{C})=\mathbf{m}^{3\mathrm{T}}(\widetilde{\mathbf{X}}-\widetilde{\mathbf{C}})$,其中 $\mathbf{m}^3$ 是主射线方向,因此 $w=\mathbf{m}^{3\mathrm{T}}(\widetilde{\mathbf{X}}-\widetilde{\mathbf{C}})$可以解释成过摄像机中心和点 $\mathbf{X}$ 的射线与主射线方向的点乘. 如果摄像机矩阵已被归一化使得$\det\mathrm{M}>0$ 且 $\|\mathbf{m}^3\|=1$,那么 $\mathbf{m}^3$是指向正轴向的单位矢量. 则 $w$ 可以解释成:从摄像机中心 $\mathbf{C}$ 到点 $\mathbf{X}$ 在主轴方向上的深度. 图 5.6 对此给予说明.

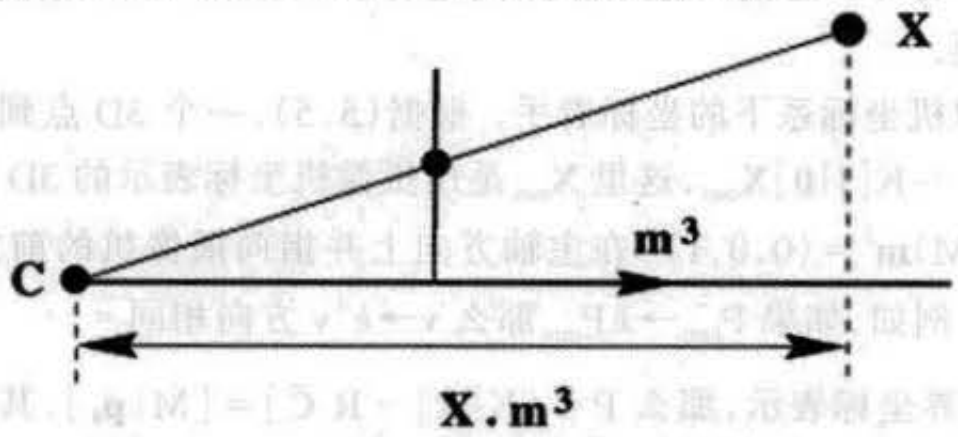



图 5.6 如果对摄像机矩阵 $\mathrm{P}=[\mathrm{M}|\mathbf{p}_4]$进行归一化,使得 $\|\mathbf{m}^3\|=1$ 和 $\det\mathrm{M}>0$,并且 $\mathbf{x}=w(x,y,1)^{\mathrm{T}}=\mathrm{P}\mathbf{X}$,其中 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}$,那么 $w$ 是点 $\mathbf{X}$ 到摄像机中心在摄像机主轴方向上的深度.

任何摄像机矩阵可以通过乘以一个适当的因子而归一化. 但是,在不作摄像机矩阵归一化时,点的深度可以用下列方法计算:

**结论 5.1** **令 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},\mathrm{T})^{\mathrm{T}}$ 是一个 3D 点而 $\mathrm{P}=[\mathrm{M}|\mathbf{p}_4]$是一个有限摄像机的摄像机矩阵,假定 $\mathrm{P}(\mathrm{X},\mathrm{Y},\mathrm{Z},\mathrm{T})^{\mathrm{T}}=w(x,y,1)^{\mathrm{T}}$. 那么**

$$\mathrm{depth}(\mathbf{X};\mathrm{P})=\frac{\mathrm{sign}(\det\mathrm{M})\,w}{\mathrm{T}\,\|\mathbf{m}^3\|} \tag{5.15}$$

**是在摄像机主平面前方的点 $\mathbf{X}$ 的深度.**

该公式是确定点 $\mathbf{X}$ 是否在摄像机前方的一种有效方式. 可以验证无论点 $\mathbf{X}$ 或摄像机矩阵 P 乘一个常数因子 $k$,深度 $\mathrm{depth}(\mathbf{X};\mathrm{P})$的值都不变. 因而,$\mathrm{depth}(\mathbf{X};\mathrm{P})$与 $\mathbf{X}$ 和 P 的具体的齐次表示无关.

### 5.2.4 摄像机矩阵的分解

若 P 为一般射影摄像机的摄像机矩阵. 我们希望由 P 找到摄像机的中心、方位和内部参数.

**求摄像机中心** 摄像机中心 $\mathbf{C}$ 是使 $\mathrm{P}\mathbf{C}=\mathbf{0}$ 的点. 数值计算上,这个右零矢量可以由 $\mathbf{P}$ 的 SVD 得到,见节 A3.3($p$408). 从代数上说,中心 $\mathbf{C}=(\mathrm{X},\mathrm{Y},\mathrm{Z},\mathrm{T})^{\mathrm{T}}$ 可以用下面方法得到

(见 2.5($p$36))

$$X=\det([\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4]) \qquad Y=-\det([\mathbf{p}_1,\mathbf{p}_3,\mathbf{p}_4])$$
$$Z=\det([\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_4]) \qquad T=-\det([\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3]).$$

**求摄像机定向和内部参数** 在有限摄像机的情形,根据(5.11),

$$P=[M|-M\widetilde{\mathbf{C}}]=K[R|-R\widetilde{\mathbf{C}}].$$

采用 RQ 分解把 M 分解成 M=KR,就能轻松地找到 K 和 R. 这种分解为上三角矩阵和正交矩阵的乘积的方法在节 A3.1.1($p$404)中介绍. 矩阵 R 给出摄像机的定向,而 K 是摄像机的标定矩阵. 分解的多义性可以通过要求 K 有正对角元素来解决.

矩阵 K 有形式(5.10)

$$K=\begin{bmatrix}\alpha_x & s & x_0\\ 0 & \alpha_y & y_0\\ 0 & 0 & 1\end{bmatrix},$$

其中

- $\alpha_x$ 是 $x$ - 坐标方向的比例因子,
- $\alpha_y$ 是 $y$ - 坐标方向的比例因子,
- $s$ 是扭曲参数,
- $(x_0,y_0)^{\mathrm{T}}$ 是主点的坐标,

**像素长宽比**是 $\alpha_x/\alpha_y$.

**例 5.2** 给定摄像机矩阵

$$P=\begin{bmatrix}3.53553e+2 & 3.39645e+2 & 2.77744e+2 & -1.44946e+6\\ -1.03528e+2 & 2.33212e+1 & 4.59607e+2 & -6.32525e+5\\ 7.07107e-1 & -3.53553e-1 & 6.12372e-1 & -9.18559e+2\end{bmatrix}$$

其中 $P=[M|-M\widetilde{\mathbf{C}}]$,则中心 $\widetilde{\mathbf{C}}=(1000.0,2000.0,1500.0)^{\mathrm{T}}$,并且矩阵 M 分解为

$$M=KR=\begin{bmatrix}468.2 & 91.2 & 300.0\\ & 427.2 & 200.0\\ & & 1.0\end{bmatrix}\begin{bmatrix}0.41380 & 0.90915 & 0.04708\\ -0.57338 & 0.22011 & 0.78917\\ 0.70711 & -0.35355 & 0.61237\end{bmatrix}.$$

△

**什么时候 $s\neq 0$?** 如节 5.1 所述,一般 $s=0$,故一个真正的 CCD 摄像机仅有 4 个摄像机内参数. 如果 $s\neq 0$,那么可以解释为 CCD 阵列的像素元素产生扭曲使得 $x$ - 轴和 $y$ - 轴不垂直. 显然这是不大可能发生的.

在现实中,非零扭曲可能在对图像再次取像(例如一幅照片再次重拍,或底片放大)时发生. 例如用一个针孔摄像机(例如普通的胶片摄像机)放大一幅图像时放大镜的轴与胶片平面或被放大的图像平面不垂直.

在这种"由图片取像"过程中可能发生的最严重的失真是一个平面单应. 假定原(有限)摄像机用矩阵 P 表示,那么表示这种由图片取像的摄像机矩阵是 HP,其中 H 是一个单应矩阵. 因为 H 是非奇异的,所以 HP 左边的 3×3 子矩阵是非奇异的因而可以分解成乘积 KR——并且不要求 K 的 $s=0$. 但是注意 K 和 R 不再是原摄像机的标定矩阵和定向.

另一方面，可以证明由图片取像的过程不改变摄像机的视在中心. 的确，因为 H 是非奇异的，所以 $\mathbf{HPC}=\mathbf{0}$ 的充要条件是 $\mathbf{PC}=\mathbf{0}$.

**什么场合需要分解？** 如果摄像机矩阵 P 由(5.11)给定，那么其参数是已知的，分解显然没有必要. 由此产生这样一个问题——在什么场合下我们会得到一个摄像机并且它的分解是未知的？事实上，在整本书中将用种种方法来计算摄像机，而分解一个未知的摄像机在实践中经常用到. 例如摄像机可以通过**标定**来直接计算——其中摄像机由世界点到图像点的对应集来计算(第 6 章)——或通过多视图关系(例如基本矩阵或三焦点张量)非直接计算，然后由这关系来计算投影矩阵.

**关于坐标定向的注释** 在推导摄像机模型和它的参数化(5.10)时，我们假定图像和 3D 世界的坐标系都是右手系，如图 5.1($p$102)所示. 但是，在通常图像坐标的测量中 $y$－坐标沿向下的方向增加，因而定义了一个与图 5.1($p$102)相反的左手坐标系. 此时，一个切实可行的建议是把图像点的 $y$－坐标取负，从而使坐标系变为右手系. 不过，即使图像坐标系是左手系，后果也不严重. 世界和图像坐标的关系仍然由一个 $3\times4$ 的摄像机矩阵表示. 它仍可能按(5.11)分解摄像机矩阵使 K 取(5.10)的形式并使 $\alpha_x$ 和 $\alpha_y$ 为正. 所不同的是 R 现在表示的摄像机定向是相对于负 Z 轴的. 另外，由(5.15)给出的点的深度在摄像机前的点是负而不是正的. 如果记住了这一点，那么在图像中使用左手坐标系也是允许的.

### 5.2.5 欧氏与射影空间

在前几节的推导中都隐含地假定世界和图像坐标系是欧氏的. 从射影几何中借用了一些概念(如点的方向在 $\boldsymbol{\pi}_\infty$ 上)以及齐次坐标的方便记号使得用线性来表示中心投影成为可能.

在本书的后续章中我们将更多地采用射影坐标系. 把世界坐标系假定为射影的是容易办到的，此时摄像机和世界坐标系之间的变换(5.6－$p$103)仍然用一个 $4\times4$ 的齐次矩阵 H 表示：$\mathbf{X}_{\mathrm{cam}}=\mathbf{HX}$；而且由 3 维射影空间 $\mathbb{P}^3$ 到图像的映射仍然用一个秩为 3 的 $3\times4$ 矩阵 P 表示. 事实上，最一般的射影摄像机是一个从 $\mathbb{P}^3$ 到 $\mathbb{P}^2$ 的映射，它包括了一个 3 维空间的射影变换、一个从 3 维空间到图像的投影和一个图像的射影变换的复合. 可以简单地用矩阵的毗接来表示这些映射：

$$\mathbf{P}=[3\times3\ \text{单应}]\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\end{bmatrix}[4\times4\ \text{单应}]$$

其结果是一个 $3\times4$ 的矩阵.

但是，记住摄像机是欧氏装置，这点是重要的而且不能简单的因为我们有一个摄像机的投影模型就可以避开欧氏几何的概念.

**欧氏与仿射解释** 虽然一个(有限)$3\times4$ 矩阵总能如节 5.2.4 所示进行分解以获得一个旋转矩阵、一个标定矩阵 K 等等. 这样获得的参数的欧氏解释仅当图像和空间坐标在适当的坐标系下才有意义. 该分解要求图像和 3 维空间都是欧氏坐标系. 另一方面，即便两个坐标系都是射影的，P 的零矢量是摄像机中心的解释仍有效——这个解释仅要求射影概念中的保线性. $\mathbf{P}^3$ 解释成主平面至少要求图像和 3 维空间是仿射坐标系. 最后，把 $\mathbf{m}^3$ 解释成主射线则要求一个仿射图像坐标系和一个欧氏世界坐标系，这样才能使垂直(于主平面)的概念

有意义.

## 5.3 无穷远摄像机

我们现在来考虑中心在无穷远平面上的摄像机. 它意味着摄像机矩阵 P 的左边 $3\times3$ 子矩阵是奇异的. 与有限摄像机一样,摄像机中心可由 $\mathbf{PC}=\mathbf{0}$ 求得.

无穷远摄像机可以大致分为两种不同的类型:**仿射摄像机**和**非仿射摄像机**. 我们首先考虑在实用中最重要的仿射摄像机.

**定义 5.3** **仿射**摄像机是矩阵 P 的最后一行 $\mathbf{P}^{3\mathrm{T}}$ 形如(0,0,0,1)的摄像机.

之所以称它为仿射摄像机是因为无穷远点被它映射为无穷远点.

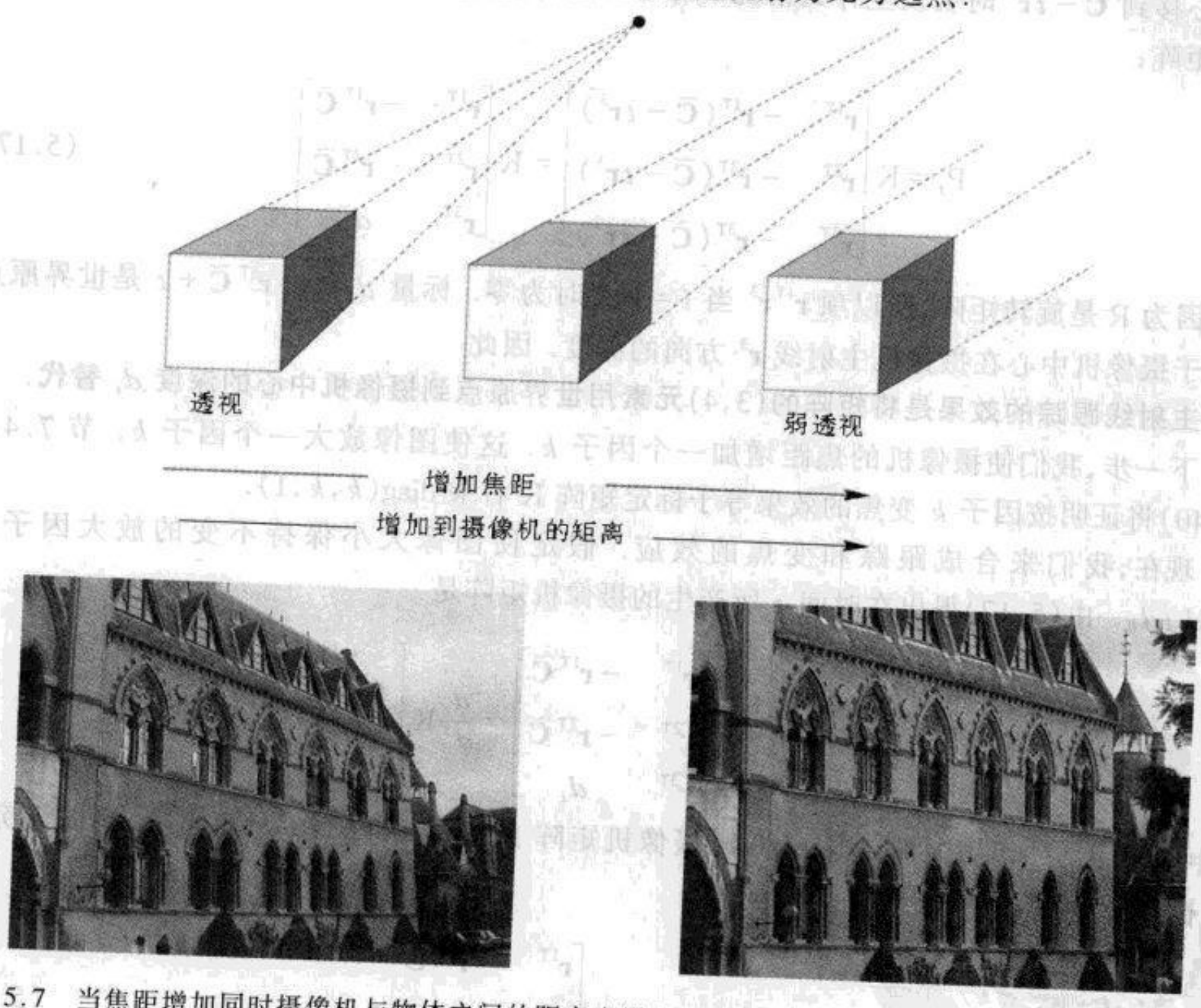



图 5.7 当焦距增加同时摄像机与物体之间的距离也增加时,图像保持同样大小,但透视效应消失.

### 5.3.1 仿射摄像机

设想当我们采用后退并放大的电影摄影技术使感兴趣的物体始终保持同样大小的图像时会发生什么①? 这种情况在图 5.7 中说明. 我们将用同时增加焦距和摄像机的主轴与物体之间的距离并取极限的方法对此过程建模.

① 参见'Vertigo'(Dir. Hitchcock,1958)和"Mishima'(Dir. Schrader,1985).

我们先从有限射影摄像机(5.11 - $p$104)开始. 该摄像机矩阵可以写成

$$P_0 = KR[I| - \widetilde{C}] = K\begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\widetilde{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\widetilde{C} \\ \mathbf{r}^{3T} & -\mathbf{r}^{3T}\widetilde{C} \end{bmatrix} \tag{5.16}$$

其中 $\mathbf{r}^{iT}$ 是旋转矩阵的第 $i$ 行. 摄像机定位在 $\widetilde{C}$ 并由矩阵 R 定向,而内部参数矩阵 K 的形式由(5.10 - $p$104)给出. 按节 5.2.1,该摄像机主射线与矢量 $\mathbf{r}^3$ 同方向,而数值 $d_0 = -\mathbf{r}^{3T}\widetilde{C}$ 是世界原点到摄像机中心在主射线方向上的距离.

现在,我们来考虑如果摄像机在一段时间 $t$ 内以单位速度沿主射线向后移动,并使摄像机中心移到 $\widetilde{C} - t\mathbf{r}^3$ 时会发生什么情况. 用 $\widetilde{C} - t\mathbf{r}^3$ 代替(5.16)中的 $\widetilde{C}$,给出在 $t$ 时刻的摄像机矩阵:

$$P_t = K\begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}(\widetilde{C} - t\mathbf{r}^3) \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}(\widetilde{C} - t\mathbf{r}^3) \\ \mathbf{r}^{3T} & -\mathbf{r}^{3T}(\widetilde{C} - t\mathbf{r}^3) \end{bmatrix} = K\begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\widetilde{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\widetilde{C} \\ \mathbf{r}^{3T} & d_t \end{bmatrix} \tag{5.17}$$

其中因为 R 是旋转矩阵,所以项 $\mathbf{r}^{iT}\mathbf{r}^3$ 当 $i = 1,2$ 时为零. 标量 $d_t = -\mathbf{r}^{3T}\widetilde{C} + t$ 是世界原点相对于摄像机中心在摄像机主射线 $\mathbf{r}^3$ 方向的深度. 因此

- **沿主射线跟踪的效果是将矩阵的(3,4)元素用世界原点到摄像机中心的深度 $d_t$ 替代.**

下一步,我们使摄像机的焦距增加一个因子 $k$. 这使图像放大一个因子 $k$. 节 7.4.1 ($p$140)将证明按因子 $k$ 变焦的效果等于标定矩阵 K 右乘 diag($k$,$k$,1).

现在,我们来合成跟踪和变焦的效应. 假定使图像大小保持不变的放大因子是 $k = d_t/d_0$. 由(5.17)得出在时间 $t$ 所产生的摄像机矩阵是

$$P_t = K\begin{bmatrix} d_t/d_0 & & \\ & d_t/d_0 & \\ & & 1 \end{bmatrix}\begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\widetilde{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\widetilde{C} \\ \mathbf{r}^{3T} & d_t \end{bmatrix} = \frac{d_t}{d_0}K\begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\widetilde{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\widetilde{C} \\ \mathbf{r}^{3T}d_0/d_t & d_0 \end{bmatrix}$$

其中因子 $d_t/d_0$ 可以忽略. 当 $t = 0$ 时摄像机矩阵 $P_t$ 对应于(5.16). 而当 $d_t$ 趋向无穷大时,该矩阵变成

$$P_\infty = \lim_{t\to\infty} P_t = K\begin{bmatrix} \mathbf{r}^{1T} & -\mathbf{r}^{1T}\widetilde{C} \\ \mathbf{r}^{2T} & -\mathbf{r}^{2T}\widetilde{C} \\ \mathbf{0}^T & d_0 \end{bmatrix} \tag{5.18}$$

它是把原来的摄像机矩阵(5.16)最后一行的前三个元素取零. 根据定义 5.3,$P_\infty$ 是仿射摄像机.

### 5.3.2 应用仿射摄像机的误差

可以看到:在过世界原点并垂直于主轴方向 $\mathbf{r}^3$ 的平面上的任何点的图像在这种变焦和运动的复合作用下保持不变. 事实是,这样的一个点可以记为

$$\mathbf{X}=\begin{pmatrix}\alpha\mathbf{r}^1+\beta\mathbf{r}^2\\ 1\end{pmatrix}.$$

因为 $\mathbf{r}^{3\mathrm{T}}(\alpha\mathbf{r}^1+\beta\mathbf{r}^2)=0$，故 $\mathrm{P}_0\mathbf{X}=\mathrm{P}_t\mathbf{X}=\mathrm{P}_\infty\mathbf{X}$ 对所有 $t$ 成立.

对于不在该平面上的点，由 $\mathrm{P}_0$ 和 $\mathrm{P}_\infty$ 映射所产生的图像是不同的，我们将研究相应的误差范围. 设一点 $\mathbf{X}$ 离该平面的垂直距离是 $\Delta$. 这个 3D 点可以表示成

$$\mathbf{X}=\begin{pmatrix}\alpha\mathbf{r}^1+\beta\mathbf{r}^2+\Delta\mathbf{r}^3\\ 1\end{pmatrix}$$

并分别被摄像机 $\mathrm{P}_0$ 和 $\mathrm{P}_\infty$ 影像到

$$\mathbf{x}_{\mathrm{proj}}=\mathrm{P}_0\mathbf{X}=\mathrm{K}\begin{pmatrix}\tilde{x}\\ \tilde{y}\\ d_0+\Delta\end{pmatrix}\qquad \mathbf{x}_{\mathrm{affine}}=\mathrm{P}_\infty\mathbf{X}=\mathrm{K}\begin{pmatrix}\tilde{x}\\ \tilde{y}\\ d_0\end{pmatrix}$$

其中 $\tilde{x}=\alpha-\mathbf{r}^{1\mathrm{T}}\tilde{\mathbf{C}}$，$\tilde{y}=\beta-\mathbf{r}^{2\mathrm{T}}\tilde{\mathbf{C}}$. 现在记标定矩阵为

$$\mathrm{K}=\begin{bmatrix}\mathrm{K}_{2\times2} & \tilde{\mathbf{x}}_0\\ \tilde{\mathbf{0}}^{\mathrm{T}} & 1\end{bmatrix},$$

其中 $\mathrm{K}_{2\times2}$ 是 $2\times2$ 上三角矩阵，得到

$$\mathbf{x}_{\mathrm{proj}}=\begin{pmatrix}\mathrm{K}_{2\times2}\tilde{\mathbf{x}}+(d_0+\Delta)\tilde{\mathbf{x}}_0\\ d_0+\Delta\end{pmatrix}\qquad \mathbf{x}_{\mathrm{affine}}=\begin{pmatrix}\mathrm{K}_{2\times2}\tilde{\mathbf{x}}+d_0\,\tilde{\mathbf{x}}_0\\ d_0\end{pmatrix}$$

$\mathrm{P}_0$ 的图像点通过除以第三个元素消去齐次化而得到：$\tilde{\mathbf{x}}_{\mathrm{proj}}=\tilde{\mathbf{x}}_0+\mathrm{K}_{2\times2}\tilde{\mathbf{x}}/(d_0+\Delta)$，而对于 $\mathrm{P}_\infty$ 的非齐次图像点是 $\tilde{\mathbf{x}}_{\mathrm{affine}}=\tilde{\mathbf{x}}_0+\mathrm{K}_{2\times2}\tilde{\mathbf{x}}/d_0$. 因此两点之间的关系是

$$\tilde{\mathbf{x}}_{\mathrm{affine}}-\tilde{\mathbf{x}}_0=\frac{d_0+\Delta}{d_0}(\tilde{\mathbf{x}}_{\mathrm{proj}}-\tilde{\mathbf{x}}_0)$$

它表明

- **用 $\mathrm{P}_\infty$ 仿射近似真实摄像机矩阵 $\mathrm{P}_0$ 的效应是把点 $\mathbf{X}$ 的图像径向地靠近或离开主点 $\tilde{\mathbf{x}}_0$，其加权因子是 $(d_0+\Delta)/d_0=1+\Delta/d_0$.**

图 5.8 对此给予了说明.

**仿射影像条件** 由 $\tilde{\mathbf{x}}_{\mathrm{proj}}$ 和 $\tilde{\mathbf{x}}_{\mathrm{affine}}$ 的表达式可以推得

$$\tilde{\mathbf{x}}_{\mathrm{affine}}-\tilde{\mathbf{x}}_{\mathrm{proj}}=\frac{\Delta}{d_0}(\tilde{\mathbf{x}}_{\mathrm{proj}}-\tilde{\mathbf{x}}_0)\tag{5.19}$$

它表明真实透视图像的位置与用仿射摄像机 $\mathrm{P}_\infty$ 逼近的位置之间的距离相差不大的条件是：

(1) 深度的起伏($\Delta$)相对平均深度($d_0$)较小，以及

(2) 点离主轴的距离较小.

后一个条件在小视场时满足. 一般，长焦距镜头容易满足这些条件，因为视场和景深都小于相同 CCD 阵列的短焦距镜头.

对于具有许多深度不同的点的景物，仿射摄像机不是一个好的近似. 例如，当景物中既含有前景又有背景的物体时，就不宜使用仿射摄像机模型. 但是，此时可对不同区域采用不同的仿射模型.

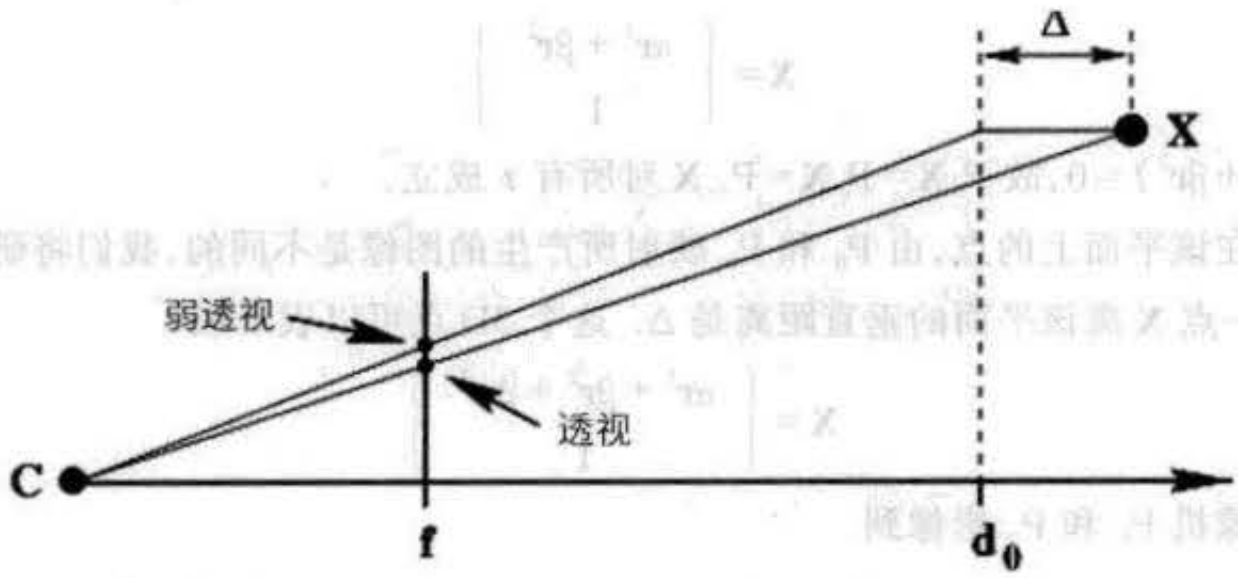



图 5.8 **透视与弱透视投影**. 弱透视摄像机的行为等价于先正交投影到一张平面(在 $Z=d_0$)上,接着从该平面进行透视投影. 透视和弱透视投影的图像点之间的区别依赖于点 X 离该平面的距离 Δ,以及该点离主轴的距离.

### 5.3.3 $P_\infty$ 的分解

摄像机矩阵(5.18)可以写成

$$P_\infty = \begin{bmatrix} K_{2\times2} & \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^\top & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} \\ \mathbf{0}^\top & d_0 \end{bmatrix}$$

其中 $\hat{R}$由旋转矩阵的前两行组成,$\hat{\mathbf{t}}$ 是矢量$(-\mathbf{r}^{1\top}\tilde{\mathbf{C}}, -\mathbf{r}^{2\top}\tilde{\mathbf{C}})^\top$,而 $\hat{\mathbf{0}}$是矢量$(0,0)^\top$,$2\times2$ 矩阵 $K_{2\times2}$是上三角矩阵. 我们可以很快地验证

$$P_\infty = \begin{bmatrix} K_{2\times2} & \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^\top & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} \\ \mathbf{0}^\top & d_0 \end{bmatrix} = \begin{bmatrix} d_0^{-1}K_{2\times2} & \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^\top & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

所以我们可以用 $d_0^{-1}K_{2\times2}$代替 $K_{2\times2}$并假定 $d_0=1$. 将此积乘起来有

$$P_\infty = \begin{bmatrix} K_{2\times2}\hat{R} & K_{2\times2}\hat{\mathbf{t}}+\tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^\top & 1 \end{bmatrix} = \begin{bmatrix} K_{2\times2} & \hat{\mathbf{0}} \\ \hat{\mathbf{0}}^\top & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}}+K_{2\times2}^{-1}\tilde{\mathbf{x}}_0 \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

$$= \begin{bmatrix} K_{2\times2} & K_{2\times2}\hat{\mathbf{t}}+\tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^\top & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{0}} \\ \mathbf{0}^\top & 1 \end{bmatrix}.$$

因此,只要对 $\hat{\mathbf{t}}$ 或$\tilde{\mathbf{x}}_0$ 作适当的替代就可以把仿射摄像机矩阵写成两种形式中的一种:

$$P_\infty = \begin{bmatrix} K_{2\times2} & \hat{\mathbf{0}} \\ \hat{\mathbf{0}}^\top & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} \\ \mathbf{0}^\top & 1 \end{bmatrix} = \begin{bmatrix} K_{2\times2} & \tilde{\mathbf{x}}_0 \\ \hat{\mathbf{0}}^\top & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{0}} \\ \mathbf{0}^\top & 1 \end{bmatrix}. \tag{5.20}$$

这样一来,摄像机 $P_\infty$ 可以根据上面的分解以两种方式(即取$\tilde{\mathbf{x}}_0=\mathbf{0}$ 或 $\hat{\mathbf{t}}=\hat{\mathbf{0}}$)中的一种加以解释. 用(5.20)的第二种分解时,世界坐标原点的图像是 $P_\infty(0,0,0,1)^\top=(\tilde{\mathbf{x}}_0^\top,1)^\top$. 因此,$\tilde{\mathbf{x}}_0$ 的值依赖于世界坐标的具体选择,从而不是摄像机本身的内部性质. 这意味着摄像机矩阵 $P_\infty$ 没有主点. 因此,我们倾向于用(5.20)中的 $P_\infty$ 的第一种分解,并记为

$$P_\infty = \begin{bmatrix} K_{2\times 2} & \hat{\mathbf{0}} \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{\mathbf{t}} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{5.21}$$

其右边的两个矩阵分别表示 $P_\infty$ 的摄像机内参数矩阵和外参数矩阵.

**平行投影** 概括起来,$P_\infty$ 和有限摄像机的本质区别是

- **平行投影矩阵** $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ **代替了有限摄像机的规范射影矩阵**$[I|\mathbf{0}]$(5.5 − $p$102).
- **标定矩阵** $\begin{bmatrix} K_{2\times 2} & \hat{\mathbf{0}} \\ \hat{\mathbf{0}}^T & 1 \end{bmatrix}$ **代替了有限摄像机的** K(5.10 − $p$104).
- **主点无定义**.

### 5.3.4 仿射摄像机的层次

类似于节 5.1 中推导有限摄像机分类的方法,我们从平行投影的基本操作开始建立一个摄像机模型的层次,它逐步地表示平行投影的更一般情形.

**正投影** 考虑沿 Z 轴的投影,它由下面的矩阵形式表示

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.22}$$

该映射把点$(X,Y,Z,1)^T$ 映射到图像点$(X,Y,1)^T$,即去掉 Z 坐标.

为得到一般正投影映射,我们在此映射之前加一个形如

$$H = \begin{bmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

的 3D 欧氏坐标变化,记 $\mathbf{t} = (t_1, t_2, t_3)^T$,我们看到一般正投影摄像机的形式是

$$P = \begin{bmatrix} \mathbf{r}^{1T} & t_1 \\ \mathbf{r}^{2T} & t_2 \\ \mathbf{0}^T & 1 \end{bmatrix}. \tag{5.23}$$

正投影摄像机有 5 个自由度,即描述旋转矩阵 R 的三个参数,加上两个偏置参数 $t_1$ 和 $t_2$. 正投影矩阵 $P = [M|\mathbf{t}]$的特征是矩阵 M 最后一行是零而前两行正交并有单位范数,以及 $t_3 = 1$.

**缩放正投影** 缩放正投影是正投影紧接一个均匀缩放. 因此,它的矩阵一般可以写成如下的形式

$$P = \begin{bmatrix} k & & \\ & k & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}^{1T} & t_1 \\ \mathbf{r}^{2T} & t_2 \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}^{1T} & t_1 \\ \mathbf{r}^{2T} & t_2 \\ \mathbf{0}^T & 1/k \end{bmatrix}. \tag{5.24}$$

它有六个自由度. 缩放正投影矩阵 $P = [M|\mathbf{t}]$的特征是矩阵 M 最后一行是零而前两行正交

并有相等的范数.

**弱透视投影** 与有限 CCD 摄像机类似,我们可以考虑两根图像轴缩放因子不一样的无穷远摄像机. 这样的摄像机的投影矩阵的形式是

$$P=\begin{bmatrix}\alpha_x & & \\ & \alpha_y & \\ & & 1\end{bmatrix}\begin{bmatrix}\mathbf{r}^{1T} & t_1\\ \mathbf{r}^{2T} & t_2\\ \mathbf{0}^T & 1\end{bmatrix}. \tag{5.25}$$

它有七个自由度. 弱透视投影矩阵 $P=[M|\mathbf{t}]$的特征是矩阵 M 最后一行是零而前两行正交(与缩放正投影不一样,它们不要求有相等的范数). 该摄像机的几何行为在图 5.8 中说明.

**仿射摄像机** $P_A$ 如 $P_\infty$一样,如果对矩阵的元素不加限制的话,一般摄像机矩阵的仿射形式,可以分解成

$$P_A=\begin{bmatrix}\alpha_x & s & \\ & \alpha_y & \\ & & 1\end{bmatrix}\begin{bmatrix}\mathbf{r}^{1T} & t_1\\ \mathbf{r}^{2T} & t_2\\ \mathbf{0}^T & 1\end{bmatrix}.$$

它有八个自由度,并可以认为是有限射影摄像机的平行投影(5.11 - $p$104).

仿射摄像机的最一般的形式是

$$P_A=\begin{bmatrix}m_{11} & m_{12} & m_{13} & t_1\\ m_{21} & m_{22} & m_{23} & t_2\\ 0 & 0 & 0 & 1\end{bmatrix}.$$

它有八个自由度,对应于八个非零非单位(1)的矩阵元素. 我们记左上角 $2\times3$ 子矩阵为 $M_{2\times3}$. 对仿射摄像机的唯一限制是 $M_{2\times3}$的秩是 2. 它出自对 P 的秩是 3 的要求.

这个仿射摄像机包括三个变换的复合效应,它们是一个 3 维空间的仿射变换、一个从 3 维空间到图像的正投影以及一个图像平面的仿射变换. 可以简单地把这些映射的矩阵毗接起来表示为

$$P_A=[3\times3\ 仿射]\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\end{bmatrix}[4\times4\ 仿射]$$

其结果是一个 $3\times4$ 仿射矩阵.

仿射摄像机下的投影在**非齐次**坐标下是一个线性映射加上一个平移:

$$\begin{pmatrix}x\\y\end{pmatrix}=\begin{bmatrix}m_{11} & m_{12} & m_{13}\\ m_{21} & m_{22} & m_{23}\end{bmatrix}\begin{pmatrix}X\\Y\\Z\end{pmatrix}+\begin{pmatrix}t_1\\t_2\end{pmatrix}$$

可以更紧凑地写成

$$\tilde{\mathbf{x}}=M_{2\times3}\tilde{\mathbf{X}}+\tilde{\mathbf{t}}. \tag{5.26}$$

点$\tilde{\mathbf{t}}=(t_1,t_2)^T$ 是世界坐标原点的图像.

本节的摄像机模型可以看作是满足附加约束的仿射摄像机,因此仿射摄像机是这个层次的一个抽象. 例如弱透视摄像机的 $M_{2\times3}$的行是经缩放的旋转矩阵的行,因而是正交的.

### 5.3.5 仿射摄像机的其他性质

空间中的无穷远平面被映射到图像中的无穷远点. 这不难通过计算 $P_A(X,Y,Z,0)^T=(X,Y,0)^T$ 得到证实. 同时也把它解释成摄像机的主平面是无穷远平面,这种解释扩展了有限射影摄像机的术语. 因为光心在主平面上,所以它必然也在无穷远平面上. 由此,我们得到

(1) 反过来说,主平面是无穷远平面的任何射影摄像机矩阵是仿射摄像机矩阵.

(2) 平行的世界直线被投影成平行的图像直线. 这是因为平行世界直线相交于无穷远平面,而该交点又被映射到图像中的无穷远点. 因此图像直线是平行的.

(3) 满足 $M_{2\times3}\mathbf{d}=\mathbf{0}$ 的矢量 $\mathbf{d}$ 是平行投影的方向,而 $(\mathbf{d}^T,0)^T$ 是摄像机中心,因为 $P_A\begin{pmatrix}\mathbf{d}\\0\end{pmatrix}=\mathbf{0}$.

由平行投影和仿射变换(或是空间的,或是图像的)的复合效应组成的任何摄像机都将具有仿射形式. 例如,**平行-透视投影**就是由这样的两个映射组成,其中第一个是平行投影到过形心并与图像平面平行的平面 $\pi$ 上. 平行投影的方向是连接形心与摄像机中心的射线方向. 平行投影之后是 $\pi$ 与图像之间的一个仿射变换(实际上是一个相似变换). 因此,平行透视摄像机是仿射摄像机.

### 5.3.6 一般无穷远摄像机

仿射摄像机是主平面为无穷远平面的摄像机. 由此,它的摄像机中心在无穷远平面上. 但是,也有摄像机的中心在无穷远平面而整个主平面不在无穷远平面的可能.

如果 $P=[M|\mathbf{p}_4]$ 并且 M 是奇异矩阵,那么这个摄像机的中心在无穷远平面上. 显然,这比要求 M 的最后一行是零(如仿射摄像机的情形)的条件更弱. 如果 M 是奇异的但 M 的最后一行不是零,那么摄像机不是仿射的,但也不是有限射影摄像机. 这样的摄像机是相当奇怪的,故不在本书中做详细的研究. 我们可以比较仿射和非仿射无穷远摄像机的性质如下:

| | 仿射摄像机 | 非仿射摄像机 |
|---|---|---|
| 摄像机中心在 $\pi_\infty$ 上 | 是 | 是 |
| 主平面是 $\pi_\infty$ | 是 | 不是 |
| $\pi_\infty$ 的点的图像在 $\mathbf{l}_\infty$ 上 | 是 | 一般不是 |

在这两种情形,摄像机中心都是投影的方向. 而且,在仿射摄像机的情形,所有非无穷远点都在摄像机的前面. 对非仿射摄像机来说,空间被主平面划分成两组点集.

一般的无穷远摄像机可以由仿射摄像机产生的图像的透视图像产生. 该影像过程可用一个表示平面单应的一般 3×3 矩阵左乘仿射摄像机矩阵来描述. 所得的 3×4 矩阵仍然是无穷远摄像机,但它没有仿射形式,因为世界平行线在图像中一般表现为会聚线.

## 5.4 其他摄像机模型

### 5.4.1 推扫式摄像机

线阵推扫式(LP)摄像机是卫星常用的一类传感器(例如 SPOT 传感器). 在这样的摄像机中,用一个线阵传感器一次获取一条影像线. 当传感器移动时,传感器平面扫过空间的一个区域(因而命名为推扫)并每次取一根图像线. 图像的第二维由传感器的运动提供. 线性推扫式模型假定传感器沿着一条直线相对地面作匀速运动. 此外,我们假定传感器阵列的方位相对于移动的方向是恒定的. 在传感器方向上,图像实际上是透视图像,而在传感器运动方向上它是正投影. LP 摄像机的几何在图 5.9 中说明. 它表明由物体空间到图像的映射可以用一个 3×4 的摄像机矩阵描述,与一般射影摄像机情形完全相同. 但是,应用这个矩阵的方式略有不同.

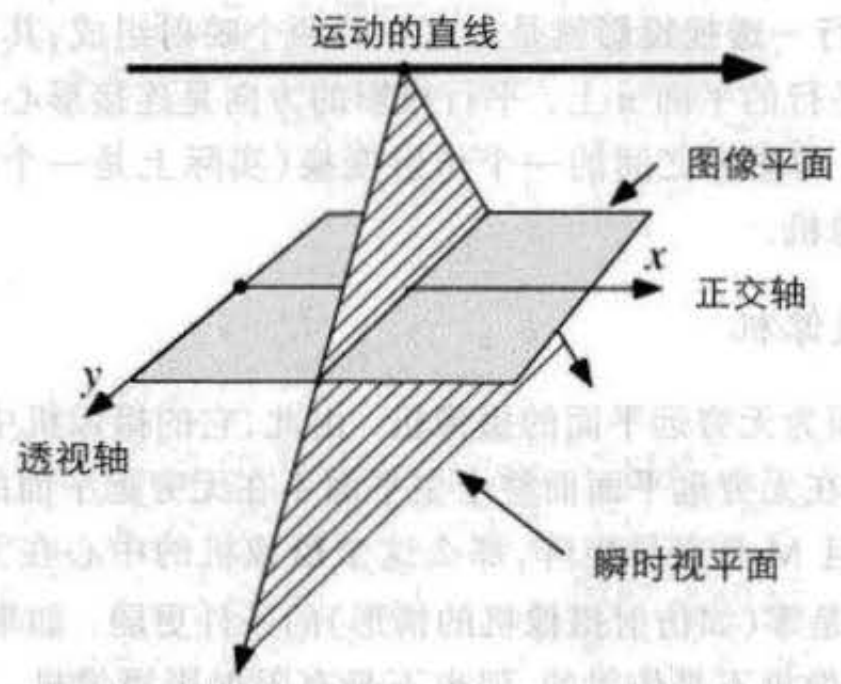



图 5.9 推扫式摄像机的摄像几何.

- **令 $\mathbf{X}=(X,Y,Z,1)^{\mathsf{T}}$ 是一个物体点,而 P 是 LP 摄像机的摄像机矩阵. 假定 $P\mathbf{X}=(x,y,w)^{\mathsf{T}}$. 那么对应的图像点(表示成一个非齐次 2 维矢量)是 $(x,y/w)^{\mathsf{T}}$.**

我们应该把它与射影摄像机映射作比较. 后一种映射用 $(x,y,w)^{\mathsf{T}}$ 表示点 $(x/w,y/w)^{\mathsf{T}}$. 注意其差别是:在 LP 的情形,为得到图像坐标,坐标 $x$ 不被比例因子 $w$ 除. 在此公式中,图像中的 $x$ 轴是传感器运动的方向,而 $y$ 轴是线性传感器阵列的方向. 此摄像机有 11 个自由度. LP 投影公式的另一种写法是

$$\tilde{x}=x=\mathbf{P}^{1\mathsf{T}}\mathbf{X} \qquad \tilde{y}=y/z=\frac{\mathbf{P}^{2\mathsf{T}}\mathbf{X}}{\mathbf{P}^{3\mathsf{T}}\mathbf{X}} \tag{5.27}$$

其中 $(\tilde{x},\tilde{y})^{\mathsf{T}}$ 是图像点.

注意 $\tilde{y}$ 坐标是射影行为,而 $\tilde{x}$ 由点 $\mathbf{X}$ 在垂直于平面 $\mathbf{P}^1$ 的方向上的正投影得到. 矢量 $\mathbf{P}^1$ 表示在时刻 $t=0$(即坐标为 $\tilde{x}=0$ 的直线被获取时)摄像机扫过的平面.

**直线的映射** LP 摄像机的一个新颖的特性是空间的直线不映射为图像中的直线(它们在射影摄像机的情形被映射为直线——见节 7.1.2). 在 3D 直线上的点集 $\mathbf{X}$ 可以写成 $\mathbf{X}_0+\alpha\mathbf{D}$,

其中 $\mathbf{X}_0=(X,Y,Z,1)^T$ 是直线上的一点，而 $\mathbf{D}=(D_X,D_Y,D_Z,0)^T$ 是该直线与无穷远平面的交. 在此情形，由(5.27)算出

$$\tilde{x}=\mathbf{P}^{1T}(\mathbf{X}_0+t\mathbf{D})$$

$$\tilde{y}=\frac{\mathbf{P}^{2T}(\mathbf{X}_0+t\mathbf{D})}{\mathbf{P}^{3T}(\mathbf{X}_0+t\mathbf{D})}.$$

这可以写成一对方程 $\tilde{x}=a+bt$ 和 $(c+dt)\tilde{y}=e+ft$. 从这些方程中消去 $t$ 得到一个形如 $\alpha\tilde{x}\tilde{y}+\beta\tilde{x}+\gamma\tilde{y}+\delta=0$ 的方程，它是图像平面上的双曲线方程，它在一个方向渐进于直线 $\alpha\tilde{x}+\gamma=0$，而在另一个方向渐进于直线 $\alpha\tilde{y}+\beta=0$. 双曲线由两条曲线组成. 但是，其中仅有一条曲线作为直线的图像实际出现在图像中——该双曲线的另一条曲线所对应的点在摄像机的后面.

### 5.4.2 线阵摄像机

本章已经研究了 3 维空间到 2D 图像的中心投影. 类似的推导可以用于一张平面到一条 1D 直线的中心投影. 见图 21.1($p$379). 这种几何的摄像机模型是

$$\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}=\mathrm{P}_{2\times3}\mathbf{x}$$

它是从平面的齐次表示到直线的齐次表示的线性映射. 该摄像机有 5 个自由度. 投影矩阵 $\mathrm{P}_{2\times3}$ 的零空间 $\mathbf{c}$ 仍然是摄像机中心，并且用类似于有限射影摄像机的方法(5.11 - $p$104)可以把矩阵分解为：

$$\mathrm{P}_{2\times3}=\mathrm{K}_{2\times2}\mathrm{R}_{2\times2}[\mathrm{I}_{2\times2}\,|-\tilde{\mathbf{c}}]$$

其中 $\tilde{\mathbf{c}}$ 是表示中心的非齐次 2 维矢量(2dof)，$\mathrm{R}_{2\times2}$ 是旋转矩阵(1dof)，而

$$\mathrm{K}_{2\times2}=\begin{bmatrix} \alpha_x & x_0 \\ & 1 \end{bmatrix}$$

是内标定矩阵(2dof).

## 5.5 结束语

本章讨论了摄像机模型、分类和构造. 后续的几章将讨论由世界到图像的对应来估计摄像机以及摄像机在各种几何物体(例如直线和二次曲面)上的作用. 消影点和消影线也将在第 7 章中作更详细的介绍.

### 5.5.1 文献

[Aloimonos - 90]定义了包括平行一透视在内的一个摄像机模型的层次. Mundy 和 Zisserman[Mundy - 92]把它推广到仿射摄像机. Faugeras 在他的书[Faugeras - 93]中推导了射影摄像机的性质. 关于线阵推扫式摄像机更多的细节在[Gupta - 97]中给出，而 2D 摄像机在[Quan - 97]中给出.

### 5.5.2 注释和练习

(1) 若 $I_0$ 是射影图像，而 $I_1$ 是 $I_0$ 的图像(图像的图像). 令其合成图像为 $I'$. 证明 $I'$

的视在摄像机中心与 $I_0$ 的一样. 思考如何用它来解释为什么肖像的眼睛"当你在房间走动时总是看着你". 另一方面验证 $I'$ 和 $I_0$ 的其他参数可能不同.

(2) **仿射摄像机**

(a) 证明仿射摄像机是把平行世界直线映射到平行图像直线的最一般的齐次坐标线性映射. 为证明这一点,考虑在 $\boldsymbol{\pi}_\infty$ 上的点的投影,并证明只有当 P 具有仿射形式时,它们才映射到图像的无穷远点.

(b) 证明被仿射摄像机映射的平行直线上的线段的长度比是不变的. 在仿射摄像机下还有什么其他的不变量?

(3) **有理多项式摄像机**是一种一般的摄像机模型,广泛地用于卫星侦察领域. 其图像坐标由下列比值定义:

$$x = N_x(\mathbf{X})/D_x(\mathbf{X}) \qquad y = N_y(\mathbf{X})/D_y(\mathbf{X})$$

其中函数 $N_x, D_x, N_y, D_y$ 是 3 维空间点 $\mathbf{X}$ 的三次齐次多项式. 每一个三次多项式有 20 个系数,因此整个摄像机有 78 个自由度. 本章研究的所有摄像机(射影,仿射,推扫)都是有理多项式摄像机的特殊情形. 但把这种一般的摄像机模型用于本章所研究的摄像机就过分地超参数化了. 更多的细节在 Hartley 和 Saxena[Hartley-97]中给出.

(4) 有限射影摄像机(5.11 - $p$104)P 可以通过右乘一个 4×4 的单应 H 变换成正投影摄像机(5.22)

$$\mathrm{PH} = \mathrm{KR}[\mathrm{I}|-\tilde{\mathbf{C}}]\mathrm{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathrm{P}_{\mathrm{orthog}}.$$

(选择 H 的最后一行使 H 的秩为 4). 于是因为

$$\mathbf{x} = \mathrm{P}(\mathrm{HH}^{-1})\mathbf{X} = (\mathrm{PH})(\mathrm{H}^{-1}\mathbf{X}) = \mathrm{P}_{\mathrm{orthog}}\mathbf{X}'$$

故在 P 下的影像等价于首先将 3 维空间的点 $\mathbf{X}$ 变为 $\mathbf{X}' = \mathrm{H}^{-1}\mathbf{X}$,然后对它进行一个正投影. 这样一来,任何摄像机的作用都可以认为是先作一个 3 维空间的射影变换,接着作一个正投影.

# 第 6 章　计算摄像机矩阵 P

本章介绍由 3 维空间与图像平面元素间的对应来估计摄像机投影矩阵的数值方法. 摄像机矩阵的这种计算称为摄像机参数确定. 最简单的对应是 3D 点 $\mathbf{X}$ 与它被未知摄像机映射的影像点 $\mathbf{x}$ 之间的对应. 只要知道足够多的对应 $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$,摄像机矩阵 P 便可以被确定. 同理,P 也可以由足够多的世界与图像的直线对应来确定.

如果对矩阵 P 施以附加约束(例如像素是正方形),那么符合这些约束的一个**受限**的摄像机矩阵可以由世界与图像的对应来估计.

在整本书中我们都假定 3 维空间到图像的映射是线性的. 如果镜头存在失真,那么该假定无效. 镜头径向失真的矫正问题将在本章处理.

摄像机的内参数可以利用节 5.2.4 的分解方法由矩阵 P 求得. 另外,用第 7 章的方法可以直接计算内参数而不必估计 P.

## 6.1　基本方程

我们假定 3D 点 $\mathbf{X}_i$ 和 2D 图像点 $\mathbf{x}_i$ 之间的若干组点对应 $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$ 已经给定. 问题是求一个摄像机矩阵 P——即一个 $3\times4$ 的矩阵使对所有 $i$ 都有 $\mathbf{x}_i = \mathrm{P}\mathbf{X}_i$. 这个问题显然与第 3 章中计算 2D 射影变换 H 的问题相似. 唯一的区别是维数不同. 在 2D 时,矩阵 H 的维数是 $3\times3$,而现在 P 的维数是 $3\times4$. 可以预料,第 3 章的许多材料几乎可以不加改变地适用于现在的情形.

如节 3.1($p$53)一样,对于每组对应 $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$,我们可以导出一个关系式:

$$\begin{bmatrix} \mathbf{0}^\mathrm{T} & -w_i\mathbf{X}_i^\mathrm{T} & y_i\mathbf{X}_i^\mathrm{T} \\ w_i\mathbf{X}_i^\mathrm{T} & \mathbf{0}^\mathrm{T} & -x_i\mathbf{X}_i^\mathrm{T} \\ -y_i\mathbf{X}_i^\mathrm{T} & x_i\mathbf{X}_i^\mathrm{T} & \mathbf{0}^\mathrm{T} \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = \mathbf{0} \tag{6.1}$$

其中 $\mathbf{P}^{i\mathrm{T}}$ 是 P 的第 $i$ 行,它是一个 4 维矢量. 另外,因为(6.1)的三个方程是线性相关的,我们可以选择只用其前两个方程:

$$\begin{bmatrix} \mathbf{0}^\mathrm{T} & -w_i\mathbf{X}_i^\mathrm{T} & y_i\mathbf{X}_i^\mathrm{T} \\ w_i\mathbf{X}_i^\mathrm{T} & \mathbf{0}^\mathrm{T} & -x_i\mathbf{X}_i^\mathrm{T} \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = \mathbf{0}. \tag{6.2}$$

利用 $n$ 组点对应并把每组对应的方程(6.2)并排起来,便得到一个 $2n\times12$ 的矩阵 A. 投影矩阵 P 通过解方程组 $\mathrm{A}\mathbf{p}=\mathbf{0}$ 算出,其中 $\mathbf{p}$ 是由矩阵 P 的元素组成的矢量.

**最小配置解**　因为矩阵 P 有 12 个元素和(忽略缩放因子)11 个自由度,所以解 P 需要 11 个方程. 由于每组点对应可以导出两个方程,为了求解 P 至少需要 $5\frac{1}{2}$ 组这样的对应. $\frac{1}{2}$ 表示仅需要用第六组点对应的一个方程,所以我们仅需要知道第六组图像点的 $x$ 坐标或 $y$ 坐

标.

给定这个最小数目的对应时,解是精确的,即空间的点准确地投影到它们被测量的图像上. 该解由解方程组 $A\mathbf{p}=\mathbf{0}$ 得到,其中 A 是 $11\times12$ 的矩阵. A 的秩一般为 11,而解矢量 $\mathbf{p}$ 是 A 的 1 维右零空间.

**超定解** 如果由于点坐标的噪声导致数据不精确并且给定 $n\geq6$ 组点对应,那么 $A\mathbf{p}=\mathbf{0}$ 将不存在精确解. 与单应的估计一样,P 的解可以通过最小化一个代数或几何误差来获得.

代数误差所用的方法是在满足某个归一化的约束下,求 $\|A\mathbf{p}\|$ 的最小值. 可能的约束是

(1) $\|\mathbf{p}\|=1$;

(2) $\|\hat{\mathbf{p}}^3\|=1$,其中 $\hat{\mathbf{p}}^3$ 是由 P 最后一行的前三个元素组成的矢量 $(p_{31},p_{32},p_{33})^{\mathsf{T}}$.

第一个约束适于常规使用,我们现在就使用它. 第二个归一化约束将在节 6.2.1 中讨论. 每种情形的残差 $A\mathbf{p}$ 都称为**代数误差**. 用这些方程计算摄像机矩阵 P 的完整的 DLT 算法与计算 H 的算法 3.1($p$55)过程一样.

**退化配置** 分析估计 P 的退化配置比 2D 单应情形要复杂很多. 有两种类型的配置都存在 P 的多义性解. 这些配置将在第 21 章中详细研究. 最重要的临界配置是

(1) 摄像机中心和点都在一条三次绕线上.

(2) 这些点都在一张平面和包含摄像机中心的一条直线的并集上.

对于这样的配置,摄像机不能由点的图像唯一地得到. 它可以沿该三次绕线或该直线任意地移动. 如果数据接近于一种退化配置,那么得到的是 P 的粗劣估计. 例如,如果摄像机远离一个地势起伏小的景物,如取接近天边区域的视图,那么这种情形接近平面性退化.

**数据归一化** 如 2D 单应估计的情形一样,对数据实行某种类型的归一化是重要的. 用与前面同样的方法对图像的点作适当的归一化. 也就是说,点应该作适当平移使它们的形心在原点,并通过缩放使它们到原点的 RMS(均方根)距离等于$\sqrt{2}$. 在选择 3D 点 $\mathbf{X}_i$ 的归一化方法时会遇到一些困难. 当点到摄像机的深度变化相对比较小时,采用同样类型的归一化是有道理的. 因此,把点的形心平移到原点,并对它们的坐标进行缩放使它们到原点的 RMS(均方根)距离等于$\sqrt{3}$(使"平均"点的坐标是$(1,1,1,1)^{\mathsf{T}}$). 这种方法适用于点紧致分布的情形,如图 6.1 中标定物体上点的分布.

如果某些点距离摄像机很远,前面的归一化技术效果不好. 例如,有些点接近摄像机,同时又有些点在无穷远(它们被影像成消影点)或接近于无穷远(如在地形的斜视图上可能出现的那样),那么通过点的平移使它们的形心移到原点是不可能的或不合理的. 此时采用第 83 页练习(2)中的归一化方法可能更合适,尽管它还没有经过彻底的验证. 经适当的归一化后,估计 P 的算法与计算 H 的算法 3.2($p$68)过程相同.

**直线对应** 扩展 DLT 算法使之也适用于直线对应的情形是一件简单的事. 3D 中的直线可以用它通过的两点 $\mathbf{X}_0$ 和 $\mathbf{X}_1$ 来表示. 这里,根据结论 7.1($p$135),由图像直线 $\mathbf{l}$ 反向投影得到的平面为 $P^{\mathsf{T}}\mathbf{l}$. 那么点 $\mathbf{X}_j$ 在该平面上的条件是

$$\mathbf{l}^{\mathrm{T}}\mathrm{P}\mathbf{X}_j = 0 \quad 其中\ j = 0,1. \tag{6.3}$$

$j$ 的每个选择给出关于矩阵 P 的元素的一个线性方程，因此由每个 3D 到 2D 的直线对应得到两个方程. 因为这些方程关于 P 的元素是线性的，它们可以加到由点对应得到的方程(6.1)中，然后计算该组合方程组的解.

## 6.2 几何误差

可以与 2D 单应(第 3 章)一样来定义几何误差. 首先假定世界点 $\mathbf{X}_i$ 比图像测量点要准确得多. 例如点 $\mathbf{X}_i$ 可能来自一个准确加工的标定物体. 那么图像中的几何误差是

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2$$

其中 $\mathbf{x}_i$ 是被测量的点，$\hat{\mathbf{x}}_i$ 是点 $\mathrm{P}\mathbf{X}_i$，即 $\mathbf{X}_i$ 在 P 作用下的精确的图像点. 如果测量误差满足高斯分布，那么

$$\min_{\mathrm{P}} \sum_i d(\mathbf{x}_i, \mathrm{P}\mathbf{X}_i)^2 \tag{6.4}$$

的解是 P 的最大似然估计.

如 2D 单应一样，最小化几何误差需要应用诸如 Levenberg－Marquardt 的迭代技术. P 的参数化可由矩阵元素的矢量 $\mathbf{p}$ 提供. DLT 解或最小配置解可以用作迭代最小化的初始点. 完整的黄金标准算法概括在算法 6.1 中.

目标

给定 $n \geq 6$ 组由世界到图像的点对应 $\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i\}$，确定摄像机投影矩阵 P 的最大似然估计，即求使 $\sum_i d(\mathbf{x}_i, \mathrm{P}\mathbf{X}_i)^2$ 取最小值的 P.

算法

(1) **线性解** 用诸如算法 3.2($p$68)的线性算法计算 P 的一个初始估计：

(a) **归一化**：用一个相似变换 T 去归一化图像点，而用第二个相似变换 U 去归一化空间点. 设归一化后的图像点是 $\tilde{\mathbf{x}}_i = \mathrm{T}\mathbf{x}_i$，归一化后的空间点是 $\tilde{\mathbf{X}}_i = \mathrm{U}\mathbf{X}_i$.

(b) **DLT**：把每组对应 $\tilde{\mathbf{X}}_i \leftrightarrow \tilde{\mathbf{x}}_i$ 产生的方程(6.2)并列起来形成一个 $2n \times 12$ 矩阵 A. 记 $\mathbf{p}$ 为矩阵 $\tilde{\mathrm{P}}$ 的元素的矢量. 满足 $\|\mathbf{p}\| = 1$ 的 $\mathrm{A}\mathbf{p} = \mathbf{0}$ 的解是对应于 A 的最小奇异值的单位奇异矢量.

(2) **最小化几何误差** 以线性估计为初始点，在 $\tilde{\mathrm{P}}$ 上用诸如 Levenberg－Marquardt 的迭代算法最小化几何误差(6.4)：

$$\sum_i d(\tilde{\mathbf{x}}_i, \tilde{\mathrm{P}}\tilde{\mathbf{X}}_i)^2.$$

(3) **解除归一化** 由 $\tilde{\mathrm{P}}$ 求原(未归一化)坐标系下的摄像机矩阵，即

$$\mathrm{P} = \mathrm{T}^{-1}\tilde{\mathrm{P}}\mathrm{U}.$$

算法 6.1 在世界点精确已知的条件下，由世界到图像的点对应估计 P 的黄金标准算法.

### 例 6.1 由标定物体估计摄像机

我们利用来自图 6.1 显示的标定物体的数据，将 DLT 算法与黄金标准算法 6.1 作一个比较. 图像点 $\mathbf{x}_i$ 由下列步骤从标定物体中得到：

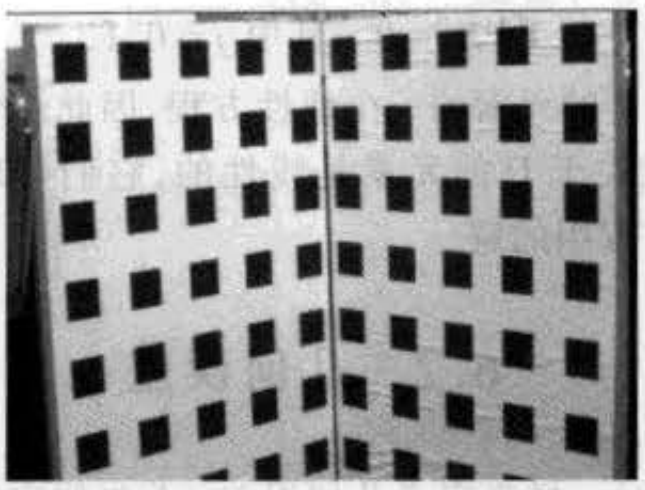


图 6.1 一种典型标定物体的图像. 黑白棋盘模式("Tsai 栅格")的设计是为了使正方形的角点的影像获得高的位置精度. 总共识别了 197 个点并用于标定本章例子中的摄像机.

| | $f_y$ | $f_x/f_y$ | 扭曲 | $x_0$ | $y_0$ | 残差 |
|---|---|---|---|---|---|---|
| 线性 | 1673.3 | 1.0063 | 1.39 | 379.96 | 305.78 | 0.365 |
| 迭代 | 1675.5 | 1.0063 | 1.43 | 379.79 | 305.25 | 0.364 |

表 6.1 DLT 和黄金标准标定

(1) Canny 边缘检测[Canny-86].

(2) 对检测过并连接的边缘进行直线拟合.

(3) 求直线的交点以便得到角点的影像.

如果足够细心,那么获取的点 $\mathbf{x}_i$ 的定位精度可以远小于 1/10 像素. 根据经验,一个好的估计所需的约束(点的测量)数目应该超过未知数(11 个摄像机参数)的五倍,也就是说至少应该采用 28 个点.

表 6.1 给出用线性 DLT 算法和黄金标准算法得到的标定结果. 注意黄金标准算法得到的结果改进非常小. 千分之一像素的残差区别可忽略不计. △

**世界点有误差**

世界点的测量可能出现不具有"无限"精度的情形. 此时,我们可以选择最小化 3D 几何误差或最小化图像几何误差或同时采用两种方法来估计 P.

如果仅考虑世界点的误差,那么 3D 几何误差定义为

$$\sum_i d(\mathbf{X}_i, \hat{\mathbf{X}}_i)^2$$

其中 $\hat{\mathbf{X}}_i$ 是空间最接近 $\mathbf{X}_i$ 并通过 $\mathbf{x}_i = \mathrm{P}\hat{\mathbf{X}}_i$ 准确地映射到 $\mathbf{x}_i$ 的点.

更一般地,如果世界和图像点的误差都考虑,那么应对世界和图像误差的加权累加进行最小化. 如 2D 单应一样,这时需要增加被估计的 3D 点 $\hat{\mathbf{X}}_i$ 的参数以扩展参数集. 我们需要最小化

$$\sum_{i=1}^{n} d_{\mathrm{Mah}}(\mathbf{x}_i, \mathrm{P}\hat{\mathbf{X}}_i)^2 + d_{\mathrm{Mah}}(\mathbf{X}_i, \hat{\mathbf{X}}_i)^2$$

其中 $d_{\mathrm{Mah}}$ 表示测量 $\mathbf{x}_i$ 和 $\mathbf{X}_i$ 的误差协方差矩阵的 Mahalanobis 距离. 最简单的 Mahalanobis 距离仅是一种加权的几何距离,其中加权值的选择反映图像和 3D 点的相对测量精度,同时因为图像和世界点通常采用不同的测量单位.

### 6.2.1 代数误差的几何解释

假定所有的点 $\mathbf{X}_i$ 在 DLT 算法中已经归一化使得 $\mathbf{X}_i=(X_i,Y_i,Z_i,1)^T$ 以及 $\mathbf{x}_i=(x_i,y_i,1)^T$. 节 3.2.4($p58$)指出,此时 DLT 算法要最小化的量是 $\sum_i(\hat{w}_i d(\mathbf{x}_i,\hat{\mathbf{x}}_i))^2$,其中 $\hat{w}_i(\hat{x}_i,\hat{y}_i,1)^T=\mathrm{P}\mathbf{X}_i$. 然而,根据(5.15 - $p108$)

$$\hat{w}_i=\pm\|\hat{\mathbf{p}}^3\|\,\mathrm{depth}(\mathbf{X};\mathrm{P})$$

因此 $\hat{w}_i$ 可以解释成点 $\mathbf{X}_i$ 沿主轴方向到摄像机的深度,这里假定已经对摄像机进行了归一化,使得 $\|\hat{\mathbf{p}}^3\|^2=p_{31}^2+p_{32}^2+p_{33}^2=1$. 由图 6.2 可以看到 $\hat{w}_i d(\mathbf{x}_i,\hat{\mathbf{x}}_i)$ 与 $fd(\mathbf{X}',\mathbf{X})$ 成比例,其中 $f$ 是焦距而 $\mathbf{X}_i'$ 是被映射到 $\mathbf{x}_i$ 的点,它在过 $\mathbf{X}_i$ 并平行于摄像机主平面的平面上. 因此,要最小化的代数误差等于 $f\sum_i d(\mathbf{X}_i,\mathbf{X}_i')^2$.

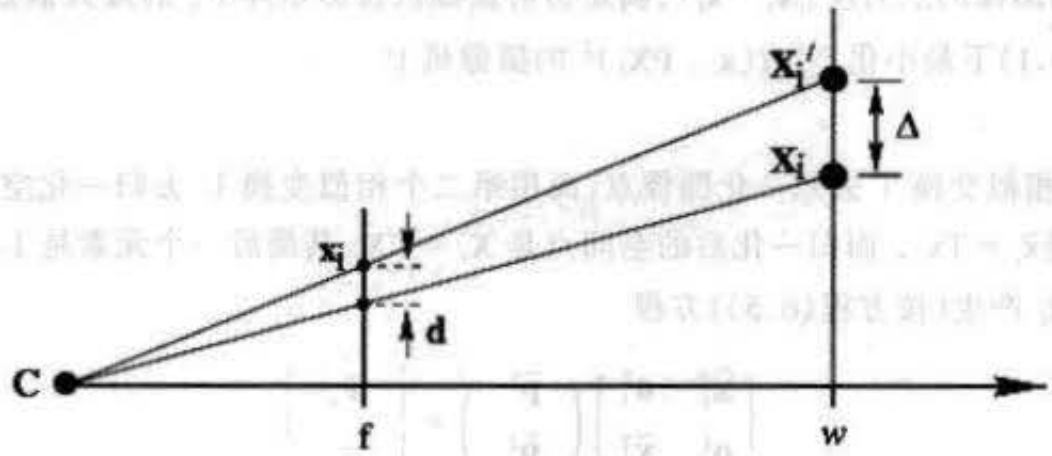



图 6.2 DLT 算法最小化点 $\mathbf{X}_i$ 和 $\mathbf{X}_i'$ 之间几何距离 Δ 的平方和,$\mathbf{X}_i'$ 在过 $\mathbf{X}_i$ 与摄像机主平面平行的平面上并精确地映射到 $\mathbf{x}_i$. 由一个简单的计算可知 $wd=f\Delta$.

为了能与被测量的图像点 $\mathbf{x}_i$ 精确地对应,需要用距离 $d(\mathbf{X}_i,\mathbf{X}_i')$ 对测量的 3D 点进行矫正. 限制条件是矫正必须在垂直于摄像机主轴的方向上进行. 由于这个限制,点 $\mathbf{X}_i'$ 和点 $\hat{\mathbf{X}}_i$ 不一样,$\hat{\mathbf{X}}_i$ 是映射到图像 $\mathbf{x}_i$ 并与 $\mathbf{X}_i$ 最接近的点. 但当 $\mathbf{X}_i$ 离摄像机主轴不太远时,距离 $d(\mathbf{X}_i,\mathbf{X}_i')$ 是 $d(\mathbf{X}_i,\hat{\mathbf{X}}_i)$ 的合理近似. 但是由于 DLT 最小化 $d(\mathbf{X}_i,\mathbf{X}_i')$(稍大于 $d(\mathbf{X}_i,\hat{\mathbf{X}}_i)$)的平方和,从而造成对远离摄像机主轴的点略有偏重. 另外,由于代数误差的表达式中出现焦距 $f$,DLT 算法将倾向于最小化焦距而使 3D 几何误差略有增加.

**变换不变性** 我们刚才看到在约束 $\|\hat{\mathbf{p}}^3\|=1$ 下最小化 $\|\mathbf{Ap}\|$ 可以解释成最小化 3D 几何距离. 这样的解释既不受 3D 空间也不受图像空间的相似变换的影响. 因此,我们可以料到对图像或 3D 中的点坐标的数据进行平移和缩放不会对解产生任何影响. 我们可以用节 3.4.2($p66$)的论点来证明这个事实.

### 6.2.2 仿射摄像机的估计

上面有关射影摄像机推导的方法可以直接用于仿射摄像机. 仿射摄像机定义为射影矩阵的最后一行是(0,0,0,1)的摄像机. 仿射摄像机的 DLT 估计是在 P 的最后一行满足上述条件下最小化 $\|\mathrm{A}\mathbf{p}\|$. 如计算 2D 仿射变换一样,仿射摄像机的代数误差和几何图像误差相等. 它表示图像的几何距离可以用一个线性算法来最小化.

如前面一样,假定所有的点 $\mathbf{X}_i$ 已经归一化使得 $\mathbf{X}_i=(X_i,Y_i,Z_i,1)^T$,$\mathbf{x}_i=(x_i,y_i,1)^T$,

并且 P 的最后一行具有仿射形式，那么，关于单组对应的(6.2)式简化为

$$\begin{bmatrix} \mathbf{0}^{\mathrm{T}} & -\mathbf{X}_i^{\mathrm{T}} \\ \mathbf{X}_i^{\mathrm{T}} & \mathbf{0}^{\mathrm{T}} \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \end{pmatrix} + \begin{pmatrix} y_i \\ -x_i \end{pmatrix} = \mathbf{0} \tag{6.5}$$

它表明对于这种情形代数误差的平方等于几何误差的平方

$$\| \mathrm{A}\mathbf{p} \|^2 = \sum_i (x_i - \mathbf{P}^{1\mathrm{T}}\mathbf{X}_i)^2 + (y_i - \mathbf{P}^{2\mathrm{T}}\mathbf{X}_i)^2 = \sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 .$$

该结果可以通过对图 5.8($p$114)和图 6.2 比较从几何上加以理解.

仿射摄像机的几何误差最小化的一个线性估计算法在算法 6.2 中给出. 在高斯测量误差的假定下，它就是 $\mathrm{P}_{\mathrm{A}}$ 的最大似然估计.

目标

给定 $n \geq 4$ 组世界到图像的点对应 $\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i\}$，确定仿射摄像机投影矩阵 $\mathrm{P}_{\mathrm{A}}$ 的最大似然估计，即求在仿射约束 $\mathbf{P}^{3\mathrm{T}} = (0,0,0,1)$ 下最小化 $\sum_i d(\mathbf{x}_i, \mathrm{P}\mathbf{X}_i)^2$ 的摄像机 P.

算法

(1) **归一化**：用一个相似变换 T 去归一化图像点，而用第二个相似变换 U 去归一化空间点. 假定归一化后的图像点是 $\tilde{\mathbf{x}}_i = \mathrm{T}\mathbf{x}_i$，而归一化后的空间点是 $\tilde{\mathbf{X}}_i = \mathrm{U}\mathbf{X}_i$ 其最后一个元素是 1.

(2) 每组对应 $\tilde{\mathbf{X}}_i \leftrightarrow \tilde{\mathbf{x}}_i$ 产生(按方程(6.5))方程

$$\begin{bmatrix} \tilde{\mathbf{X}}_i^{\mathrm{T}} & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0}^{\mathrm{T}} & \tilde{\mathbf{X}}_i^{\mathrm{T}} \end{bmatrix} \begin{pmatrix} \tilde{\mathbf{P}}^1 \\ \tilde{\mathbf{P}}^2 \end{pmatrix} = \begin{pmatrix} \tilde{x}_i \\ \tilde{y}_i \end{pmatrix}$$

把它们全起来形成一个 $2n \times 8$ 的矩阵方程 $\mathrm{A}_8\mathbf{p}_8 = \mathbf{b}$，其中 $\mathbf{p}_8$ 为包含矩阵 $\tilde{\mathrm{P}}_{\mathrm{A}}$ 前两行的 8 维矢量.

(3) 解由 $\mathrm{A}_8$ 的伪逆(见节 A3.4 ($p$410))给出：

$$\mathbf{p}_8 = \mathrm{A}_8^{+}\mathbf{b}$$

并且 $\tilde{\mathbf{p}}^{3\mathrm{T}} = (0,0,0,1)$.

(4) **解除归一化**. 原(未归一化)坐标系下的摄像机矩阵由 $\tilde{\mathrm{P}}_{\mathrm{A}}$ 求得，即

$$\mathrm{P}_{\mathrm{A}} = \mathrm{T}^{-1}\tilde{\mathrm{P}}_{\mathrm{A}}\mathrm{U}$$

算法 6.2　利用从世界到图像的点对应来估计仿射摄像机矩阵 $\mathrm{P}_{\mathrm{A}}$ 的黄金标准算法.

## 6.3　受限摄像机估计

如迄今为止所介绍的那样，DLT 算法通过 3D 到 2D 的点对应集合计算一般射影摄像机的矩阵 P. 中心在有限点的矩阵 P 可以分解成 $\mathrm{P} = \mathrm{K}[\mathrm{R} \mid -\mathrm{R}\tilde{\mathbf{C}}]$，其中 R 是 $3 \times 3$ 的旋转矩阵而 K 具有(5.10 - $p$104)的形式：

$$\mathrm{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}. \tag{6.6}$$

作为 P 的内标定参数，K 中的非零元素是有几何意义的量. 我们可以在关于摄像机参数的限制条件下寻求一个最适配的摄像机矩阵 P. 通常的限制是

(1) 扭曲 $s$ 为零.

(2) 像素是正方形：$\alpha_x = \alpha_y$.

(3) 主点$(x_0, y_0)^T$已知.

(4) 整个摄像机标定矩阵 K 已知.

在某些情形下,一个受限的摄像机矩阵有可能用线性算法来估计(见本章末的练习).

我们用求一个适配于测量点集的最理想的针孔摄像机模型(即满足 $s=0$ 和 $\alpha_x=\alpha_y$ 的射影摄像机)来作为受限估计的一个例子. 如下文所讨论,我们可以用最小化几何误差或最小化代数误差来解这个问题.

**最小化几何误差** 为最小化几何误差,我们在摄像机矩阵中选择一组将要计算的参数. 例如,假定我们强调约束 $s=0$ 和 $\alpha_x=\alpha_y$,我们可以用余下的 9 个参数来参数化摄像机矩阵. 它们是 $x_0$、$y_0$、$\alpha$ 和表示摄像机定向 R 和位置 $\tilde{\mathbf{C}}$ 的 6 个参数. 记该参数集为 $\mathbf{q}$. 那么摄像机矩阵 P 可以用这些参数显式地算出.

几何误差可以用迭代最小化方法(例如 Levenberg－Marquardt)相对于这组参数来最小化. 注意当仅对图像误差最小化时,问题的大小是 $9\times 2n$(假定摄像机有 9 个未知参数). 换句话说,LM 最小化算法是最小化函数 $f: \mathbb{R}^9 \to \mathbb{R}^{2n}$. 当对 3D 和 2D 的误差都最小化时,函数 $f$ 是从 $\mathbb{R}^{3n+9} \to \mathbb{R}^{5n}$ 的映射,原因是测量点必须包含 3D 点并且最小化中还包括估计 3D 点的真实位置.

**最小化代数误差** 这个问题有可能用代数误差最小化来解,在这种情形下迭代最小化问题会变得非常小,这将在下文中解释. 考虑把参数集 $\mathbf{q}$ 映射到摄像机矩阵 $\mathrm{P}=\mathrm{K}[\mathrm{R}\mid -\mathrm{R}\tilde{\mathbf{C}}]$ 的参数化映射 $g$. 确切地说,我们有一个映射 $\mathbf{p}=g(\mathbf{q})$,其中 $\mathbf{p}$ 是矩阵 P 的元素的矢量. 最小化所有点匹配的代数误差等价于最小化 $\|\mathrm{A}g(\mathbf{q})\|$.

**简化的测量矩阵** 一般,$2n\times 12$ 的矩阵 A 可能有很多行. 但可用一个 $12\times 12$ 的矩阵 $\hat{\mathrm{A}}$ 代替 A,使得对任何矢量 $\mathbf{p}$ 有 $\|\mathrm{A}\mathbf{p}\| = \mathbf{p}^T\mathrm{A}^T\mathrm{A}\mathbf{p} = \|\hat{\mathrm{A}}\mathbf{p}\|$. 这样的一个矩阵称**简化的测量矩阵**. 计算它的一种方法是利用奇异值分解(SVD). 令 $\mathrm{A}=\mathrm{UDV}^T$ 是 A 的 SVD,并定义 $\hat{\mathrm{A}}=\mathrm{DV}^T$. 那么

$$\mathrm{A}^T\mathrm{A}=(\mathrm{VDU}^T)(\mathrm{UDV}^T)=(\mathrm{VD})(\mathrm{DV}^T)=\hat{\mathrm{A}}^T\hat{\mathrm{A}}$$

符合我们的要求. 获取 $\hat{\mathrm{A}}$ 的另一种途径是利用 QR 分解 $\mathrm{A}=\mathrm{Q}\hat{\mathrm{A}}$,其中 Q 具有正交列而 $\hat{\mathrm{A}}$ 是上三角方阵.

注意映射 $\mathbf{q}\mapsto \hat{\mathrm{A}}g(\mathbf{q})$ 是从 $\mathbb{R}^9$ 到 $\mathbb{R}^{12}$ 的映射. 这是一个简单的参数最小化问题,可以用 Levenberg－Marquardt 方法求解. 须注意的要点是:

- **给定 $n$ 组世界到图像的点对应集:$\mathbf{X}_i \leftrightarrow \mathbf{x}_i$,求最小化代数距离 $\sum_i d_{alg}(\mathbf{x}_i, \mathrm{P}\mathbf{X}_i)^2$ 的约束摄像机矩阵 P 的问题简化为与点对应数目 $n$ 无关的函数 $\mathbb{R}^9 \to \mathbb{R}^{12}$ 的最小化问题.**

$\|\hat{\mathrm{A}}g(\mathbf{q})\|$ 在参数 $\mathbf{q}$ 的所有值上最小化. 注意如果 $\mathrm{P}=\mathrm{K}[\mathrm{R}\mid -\mathrm{R}\tilde{\mathbf{C}}]$ 中的 K 如(6.6)所示,那么 P 满足条件 $p_{31}^2+p_{32}^2+p_{33}^2=1$,因为这些元素与旋转矩阵 R 的最后一行相同. 因此,最小化 $\mathrm{A}g(\mathbf{q})$ 将给出一个满足约束 $s=0$ 和 $\alpha_x=\alpha_y$ 并通过缩放使 $p_{31}^2+p_{32}^2+p_{33}^2=1$ 的矩阵 P,而且它还最小化所有点对应的代数误差.

**初始化** 求摄像机初始参数的一种途径是:

(1) 用诸如 DLT 的线性算法求出一个初始的摄像机矩阵.

(2) 把固定参数强制到所希望的取值范围(例如,令 $s=0$ 并且设 $\alpha_x=\alpha_y$ 并等于 DLT 算法求得的平均值).

(3) 把摄像机矩阵分解(见节 5.2.4)所获得的初始值赋给参数变量.

在理想情形下固定参数的假定值应接近 DLT 所得到的值. 但是,实际上不总是如此. 因此,把有关参数用它们的希望值替代会得到不正确的初始摄像机矩阵,并可能导致大的残差和难于收敛. 实践中效果更好的方法是采用**软**约束,即在代价函数中额外添加一些项. 例如,在 $s=0$ 和 $\alpha_x=\alpha_y$ 的情形,我们把 $ws^2+w(\alpha_x-\alpha_y)^2$ 作为额外项加到代价函数中. 因此,在图像几何误差的情形,代价函数变成

$$\sum_i d(\mathbf{x}_i,\ \mathbf{PX}_i)^2+ws^2+w(\alpha_x-\alpha_y)^2.$$

我们首先用由 DLT 算法估计得到的参数值. 在估计过程中,加权值从一个小的值开始并在每次迭代中逐渐增加. 这样一来,$s$ 的值和长宽比将逐渐地接近它们的预期值. 从而有可能在最后的估计中把它们最终固定到所预期的值.

**外部校准** 假定摄像机所有的内参数已知,那么剩下来就是要确定摄像机的位置和定向(姿态). 这是"外部校准"问题,它在已标定系统的分析中有重要意义.

为了计算外部校准,需对世界坐标位置准确已知的一个配置进行影像. 之后求摄像机的姿态. 在机器人系统的手眼标定中求摄像机位置就是这样的情形;还有在采用配准技术的基于模型的识别中,需要知道物体相对摄像机的位置.

这里必须确定六个参数,三个用于定向、三个用于位置. 因为世界到图像的每组对应产生两个约束,所以有三点就足够了. 事实的确如此,而且求得的非线性方程一般有四个解.

**实验评估**

对例 6.1 中的标定栅格进行受限估计的结果在表 6.2 中给出.

| | $f_y$ | $f_x/f_y$ | 扭曲 | $x_0$ | $y_0$ | 残差 |
|---|---|---|---|---|---|---|
| 代数 | 1633.4 | 1.0 | 0.0 | 371.21 | 293.63 | 0.601 |
| 几何 | 1637.2 | 1.0 | 0.0 | 371.32 | 293.69 | 0.601 |

表 6.2 受限摄像机矩阵的标定

代数和几何最小化都是在 9 个参数上的迭代最小化. 但是,代数误差方法快得多,因为它只要最小化 12 个误差,而几何最小化却是 $2n=396$ 个. 注意把扭曲和长宽比固定会造成其他参数值的改变(比较表 6.1)并会增大残差值.

**协方差估计** 协方差估计和误差在图像中的传播技术可以象 2D 单应(第 4 章)一样处理. 类似地,计算最小残差期望值的技术可按结论 4.2($p$89)进行. 假定所有的误差仅发生在图像测量中, ML 残差期望值等于

$$\varepsilon_{res}=\sigma(1-d/2n)^{1/2}$$

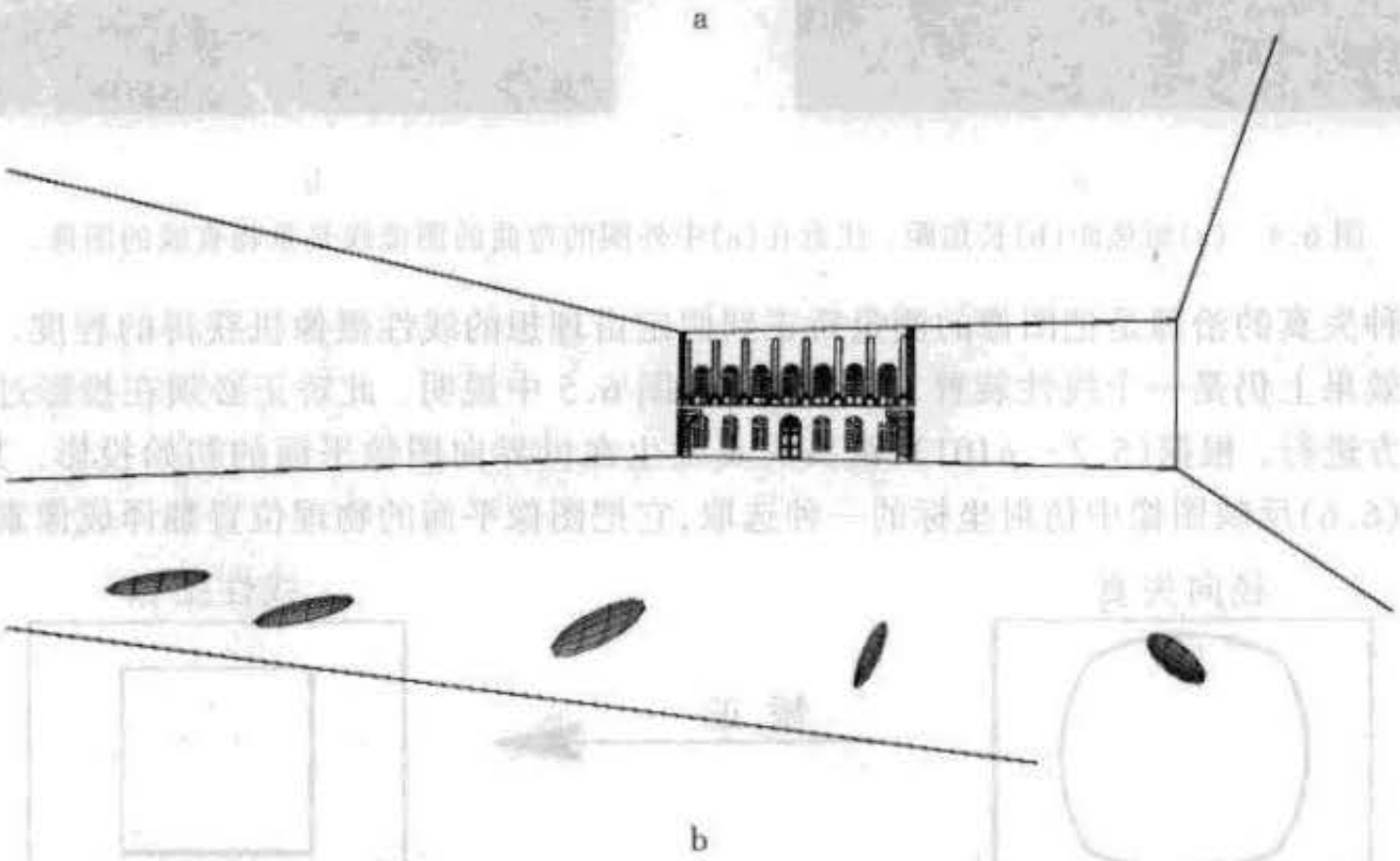
其中 $d$ 是要拟合的摄像机参数数目(对完整的针孔摄像机模型是 11). 给定一个残差,该公式也可以用来估计点测量的准确性. 例 6.1 中,当 $n=197$ 和 $\varepsilon_{res}=0.365$ 时,得到的值是

$\sigma = 0.37$. 这个值比预期的大. 其原因出自摄像机模型——我们忽略了径向失真.

**例 6.2　摄像机估计的方差椭球**

假定摄像机矩阵的估计采用了在一组摄像机参数上最优化的最大似然(黄金标准)方法. 根据结论 4.10($p$93),通过后向传播可以把测量点的估计的协方差用来计算摄像机模型的协方差. 由此给出 $\Sigma_{\text{camera}} = (J^T \Sigma_{\text{points}}^{-1} J)^{-1}$,其中 J 是以摄像机参数表示的测量点的 Jacobian矩阵. 3D 世界点的不确定性也用同样方法考虑. 如果摄像机用有意义的参数(例如摄像机的位置)来参数化,那么每个参数的方差可以从协方差矩阵的对角元素直接得到.

a



b

图 6.3　**摄像机中心的协方差椭球**. (a)Stanislas 广场(法国南锡)的 5 幅图像,其 3D 标定点已知. (b)对应于每一幅图像的摄像机中心的方差椭球,它们由估计的摄像机计算得到,摄像机通过标定点的图像来估计. 注意椭球为典型的雪茄形状并指向景物数据. 此图的采用得到 Vincent Lepetit, Marie - Odile Berger 和 Gilles Simon 的允许.

已知摄像机参数的协方差,误差界或方差椭球就能算出. 例如由所有参数的协方差矩阵,我们可以取出表示摄像机位置的 3×3 协方差矩阵的子块 $\Sigma_C$. 那么,摄像机中心的置信度椭球可以定义为

$$(\mathbf{C}-\bar{\mathbf{C}})^T \Sigma_C^{-1} (\mathbf{C}-\bar{\mathbf{C}})^T = k^2$$

其中 $k^2$ 是置信度期望为 $\alpha$ 时累积分布 $\chi_n^2$ 的逆,即 $k^2 = F_n^{-1}(\alpha)$(见图 A2.1($p$403)). 这里 $n$ 是变量数——对摄像机中心而言是 3. 若摄像机中心所选的置信度为 $\alpha$,它将在此椭球中.

图 6.3 中的例子给出计算摄像机中心的不确定性的椭球区域. 若摄像机的协方差估计矩阵已得到,节 4.2.6($p$97)的技术可以用来计算 3D 世界点在图像中位置的不确定性. △

## 6.4 径向失真

在前几章中，我们总假定线性模型是影像过程的精确模型．因此世界点、图像点和光心共线并且世界直线被影像为直线．但实际的(非针孔)镜头不符合这种假设．通常最重要的偏差是径向失真．而且当镜头的焦距(和价格)减少时该误差会更显著．见图 6.4．

a　　b

图 6.4　(a)短焦距(b)长焦距．注意在(a)中外围的弯曲的图像线是景物直线的图像．

这种失真的治理是把图像的测量矫正到假定由理想的线性摄像机获得的程度．这样摄像机在效果上仍是一个线性装置．这个过程在图 6.5 中说明．此矫正必须在投影过程的恰当的地方进行．根据(5.2－$p$101)，镜头失真发生在世界向图像平面的初始投影．其后，标定矩阵(6.6)反映图像中仿射坐标的一种选取，它把图像平面的物理位置翻译成像素坐标．

图 6.5　一个正方形的严重径向失真的图像被矫正成假定由理想的线性镜头获得的图像．

我们用$(\tilde{x}, \tilde{y})^{\mathsf{T}}$标记在理想(非失真)针孔投影下点以焦距为测量单位的坐标．那么，对一点 $\mathbf{X}$ 我们有(见(5.5－$p$102))

$$(\tilde{x}, \tilde{y}, 1)^{\mathsf{T}} = [\mathbf{I}|\mathbf{0}]\mathbf{X}_{\text{cam}}$$

其中 $\mathbf{X}_{\text{cam}}$ 是摄像机坐标下的 3D 点，它与世界坐标的关系由(5.6－$p$103)给出．实际的投影点通过一个径向位移与理想点关联．因此，径向(透镜)失真的模型是

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = L(\tilde{r}) \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \tag{6.7}$$

其中

- $(\tilde{x}, \tilde{y})^{\mathsf{T}}$是理想图像位置(遵循线性投影)．
- $(x_d, y_d)^{\mathsf{T}}$经径向失真后的实际图像的位置．
- $\tilde{r}$为到径向失真中心的径向距离$\sqrt{\tilde{x}^2 + \tilde{y}^2}$．

• $L(\tilde{r})$是一个失真因子，它仅仅是半径$\tilde{r}$的函数.

**失真矫正** 在像素坐标中，失真矫正记为

$$\hat{x} = x_c + L(r)(x - x_c) \qquad \hat{y} = y_c + L(r)(y - y_c)$$

其中$(x, y)^T$是测量的坐标，$(\hat{x}, \hat{y})^T$是矫正后的坐标，而$(x_c, y_c)^T$是径向失真的中心，并且$r^2 = (x - x_c)^2 + (y - y_c)^2$. 注意如果长宽比不是1，那么在计算$r$时必须对它进行矫正. 通过这样的矫正，坐标$(\hat{x}, \hat{y})^T$与3D世界点的坐标由一个线性投影摄像机矩阵相关联.

**失真函数和中心的选择** 函数$L(r)$仅当$r$为正值时有定义并且$L(0) = 1$. 一个任意函数$L(r)$的逼近可以由泰勒展开式$L(r) = 1 + \kappa_1 r + \kappa_2 r^2 + \kappa_3 r^3 + \cdots$给出. 径向矫正的系数$\{\kappa_1, \kappa_2, \kappa_3, \cdots, x_c, y_c\}$是摄像机内标定要考虑的一部分. 主点经常被用作径向失真的中心，虽然它们未必完全重合. 此矫正和摄像机标定矩阵一起规定了一个图像点到摄像机坐标系中的一条射线的映射.

**计算失真函数** 函数$L(r)$可以通过最小化一个基于映射的线性偏差的代价函数来计算. 例如，算法6.1($p$123)通过最小化标定物体(如图6.1($p$124)的Tsai栅格)的几何图像误差来估计P. 失真函数可以作为影像过程的一部分，把参数$\kappa_i$与P一起在几何误差的最小化迭代中计算. 类似地，失真函数可以在估计单个Tsai栅格和其图像的单应时得到.

一种简单且更一般的确定$L(r)$的方法是要求景物直线的图像必须是直线. 代价函数定义在通过$L(r)$矫正映射后的图像直线上(例如图像直线的端点与中点之间的距离). 此代价函数以迭代方式在失真函数参数$\kappa_i$和径向失真中心上进行最小化. 这对城市景物图像是一种非常实用的方法，因为其中通常存在大量的直线. 它的优越性是不需要特殊的标定模块，因为景物提供了标定元素.

**例6.3 径向矫正** 通过最小化一个基于被影像景物直线的直线程度的代价函数来计算图6.6a中图像的函数$L(r)$. 图像大小是640×480像素，而计算得到的矫正参数和中心是$\kappa_1 = 0.103689$. $\kappa^2 = 0.00487908$. $\kappa^3 = 0.00116894$. $\kappa^4 = 0.000841614$. $x_c = 321.87$, $y_c = 241.18$像素，其中像素由图像一半大小的像素的平均归一化. 图像外围矫正达30个像素. 图像插补的结果在图6.6b显示. △

**例6.4** 我们用图6.1($p$124)中的标定栅格的例子继续例6.1($p$123)的讨论. 用直线方法把径向失真移去，然后用本章介绍的方法标定摄像机. 结果在表6.3中给出.

注意径向矫正后，残差明显变小. 由此残差推出的点测量误差的估计值是$\sigma = 0.18$像素. 因为径向失真涉及到图像有选择性的拉伸，如表中所示图像的有效焦距变化了. △

在径向失真的矫正中，通常不需要实际地对图像进行形变插补. 测量可以在原图中进行(如角点的位置)，而只须把测量按(6.7)映射. 关于特征**应该**在什么地方测量的问题没有一个肯定的回答. 形变插补图像会使噪声模型失真(由于取平均)而且很可能引入混频效应. 因为这个理由，通常首选在未插补的图像上进行特征检测. 但是，特征群组合，例如连接边界点形成直线元素，最好在插补后进行，因为原图的线性度阈值很可能会过大.

a　　b

**图 6.6　径向失真的矫正.** (a)原始图像,在此图像中世界直线的像是弯曲的. 其中的若干条线用虚线标注. (b)对图像进行形变插补以消除径向失真. 注意现在图像外围的线是直的了,但图像的边界却弯曲了.

| | $f_y$ | $f_x/f_y$ | 扭曲 | $x_0$ | $y_0$ | 残差 |
|---|---|---|---|---|---|---|
| 线性 | 1580.5 | 1.0044 | 0.75 | 377.53 | 299.12 | 0.179 |
| 迭代 | 1580.7 | 1.0044 | 0.70 | 377.42 | 299.02 | 0.179 |
| 代数 | 1556.0 | 1.0000 | 0.00 | 372.42 | 291.86 | 0.381 |
| 迭代 | 1556.6 | 1.0000 | 0.00 | 372.41 | 291.86 | 0.380 |
| 线性 | 1673.3 | 1.0063 | 1.39 | 379.96 | 305.78 | 0.365 |
| 迭代 | 1675.5 | 1.0063 | 1.43 | 379.79 | 305.25 | 0.364 |
| 代数 | 1633.4 | 1.0000 | 0.00 | 371.21 | 293.63 | 0.601 |
| 迭代 | 1637.2 | 1.0000 | 0.00 | 371.32 | 293.69 | 0.601 |

**表 6.3　通过和不通过径向矫正的标定**　在中隔线以上的结果是径向矫正过的——以下的结果是没有经过径向矫正的(它引自前面的表,为了比较在这里给出). 两种情形中,上面两种方法用于一般摄像机模型,而下面两种仅用于具有正方形像素的受限模型.

## 6.5　结束语

### 6.5.1　文献

在[Sutherland－63]中,DLT 最初的应用是计算摄像机. 由几何误差的迭代最小化来进行估计是摄影测量术的标准过程,例如见[Slama－80].

标定摄像机的最小配置解(由 3 点的图像得到的姿态)作为一个原创问题由 Fischler 和 bolles[Fischler－81]在他们的 RANSAC 的文章中给予研究. 这个问题的解经常重新出现在文献中;[Wolfe－91]和[Haralick－91]给出了一个好的论述. 准线性解在[Quan－98]和[Triggs－99a]中给出,它比最少的点对应 $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$ 数目多一组.

这里未涉及的另一类方法是由仿射摄像机开始对射影摄像机进行迭代估计. Dememthon和 Davis[Dementhon－95]的算法"基于 25 行编码姿态的模型"采用了这个思想. 一个类似的方法被[Christy－96]采用.

Devernay 和 Faugeras[Devernay－95]把计算径向失真的直线方法引入计算视觉的文献．在摄影测量学中该方法称为“铅锤线矫正”，见[Brown－71]．

### 6.5.2 注释和练习

(1) 给定 5 组世界到图像的点对应 $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$，证明对具有零扭曲的摄像机矩阵 P 一般有四个解，它们都精确地把世界点映射到图像点．

(2) 给定 3 组世界到图像的点对应 $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$，证明对具有已知标定 K 的摄像机矩阵 P 一般有四个解，它们都精确地把世界点映射到图像点．

(3) 在下面的每种条件下，求计算摄像机矩阵 P 的线性算法：

(a) 摄像机的位置(但不是定向)已知．

(b) 摄像机主轴的定向已知．

(c) 摄像机的位置和摄像机主轴的定向已知．

(d) 摄像机的位置和摄像机的整个定向已知．

(e) 摄像机的位置和定向已知，同时摄像机内参数($\alpha_x, \alpha_y, s, x_0$ 和 $y_0$)的某些子集也已知．

(4) **焦距和在主轴上的位置的综合考虑**．比较一个深度为 $d$ 的点在摄像机焦距增加 $\Delta f$ 或摄像机沿主轴后退一个位移 $\Delta t_3$ 之前和之后被影像的位置．令 $(x, y)^{\mathrm{T}}$ 和 $(x', y')^{\mathrm{T}}$ 为变化前和后的点的图像坐标．遵循与(5.19－$p$113)类似的推导，证明

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + k \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

其中 $k^f = \Delta f / f$ 表示焦距的变化，或 $k^{t_3} = -\Delta t_3 / d$ 表示位移(这里扭曲 $s = 0$，而 $\alpha_x = \alpha_y = f$)．

对标定点 $\mathbf{X}_i$ 的集合，当它们的深度变化($\Delta_i$)比平均深度 $d_0$ 小时，有

$$k_i^{t_3} = -\Delta t_3 / d_i = -\Delta t_3 / (d_0 + \Delta_i) \approx -\Delta t_3 / d_0$$

即 $k_i^{t_3}$ 在整个集合上近似于常数．由此推出由这样的集合来标定时，变化焦距 $k^f$ 或移动摄像机 $k^{t_3}$ 得到相似的图像残差．因此，被估计的参数：焦距和在主轴上的位置相互关联．

(5) **推扫式摄像机的计算**．在节 5.4.1 中介绍的推扫式摄像机也可以用 DLT 方法来计算．投影矩阵的 $x$(正投影)部分有 4 个自由度，它们可以由四个或更多的点对应 $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$ 来确定；投影矩阵的 $y$(透视)部分有 7 个自由度，它们可以由 7 个点对应来确定．因此，一个最小配置解需要 7 个点．细节在[Gupta－97]中给出．

# 第 7 章　进一步讨论单视图几何

第 5 章引入投影矩阵作为摄像机对点作用的模型. 本章将介绍在透视投影下其他的 3D 几何实体与其 2D 图像之间的联系. 这些实体包括平面、直线、二次曲线、二次曲面;我们也将推导它们的正向和反向投影的性质.

我们将进一步地剖析摄像机,并简化到它的中心点和图像平面. 最终建立两种性质:由同一中心的摄像机所获取的图像以一个平面射影变换相关联;无穷远平面 $\boldsymbol{\pi}_\infty$ 上的几何实体的图像与摄像机的位置无关,仅取决于摄像机的旋转和摄像机的内参数 K.

无穷远平面 $\boldsymbol{\pi}_\infty$ 上的几何实体(点、线、二次曲线)的图像特别重要. 我们将看到 $\boldsymbol{\pi}_\infty$ 上点的图像是消影点,而 $\boldsymbol{\pi}_\infty$ 上直线的图像是消影线,它们的图像仅与 K 和摄像机旋转有关. 但是,绝对二次曲线的图像 $\omega$ 仅与 K 有关,而不受摄像机旋转的影响. 二次曲线 $\omega$ 与摄像机标定 K 密切关联并有关系 $\omega=(KK^T)^{-1}$. 因而,由图像点反向投影产生的射线之间的角度可根据 $\omega$ 来确定.

有了这些性质,就可以用消影点来计算摄像机的相对旋转而无需知道摄像机的位置. 进一步地说,因为利用 K 就能由图像点来计算射线之间的角度,那么反过来,K 也可以通过射线之间的角度来计算. 特别是 K 可以由景物的正交方向所对应的消影点来确定. 这说明摄像机可以由景物特征来标定而不需要知道具体的世界坐标.

本章介绍的最后一个几何实体是标定二次曲线,它用几何方式使 K 可视化.

## 7.1　射影摄像机对平面、直线和二次曲线的作用

本节(实际上本书的大部分内容)在确定投影矩阵 P 的作用时通常重要的只是它的 3×4**形式**和秩. 而该矩阵元素的具体性质和关系常是相对地无关紧要.

### 7.1.1　对平面的作用

点的影像方程 $\mathbf{x}=P\mathbf{X}$ 是世界坐标系点到图像坐标点的一个映射. 我们可以自由地选择世界坐标系. 假定选择 XY－平面与景物中的平面 $\boldsymbol{\pi}$ 对应,使得在此景物平面上的点的 Z 坐标为零,如图 7.1 所示(设摄像机中心不在景物平面上). 那么,如果用 $\mathbf{p}_i$ 表示 P 的列,则 $\boldsymbol{\pi}$ 上的点的图像为

$$\mathbf{x}=P\mathbf{X}=[\mathbf{p}_1\quad \mathbf{p}_2\quad \mathbf{p}_3\quad \mathbf{p}_4]\begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix}=[\mathbf{p}_1\quad \mathbf{p}_2\quad \mathbf{p}_4]\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}.$$

因此,$\boldsymbol{\pi}$ 上的点 $\mathbf{x}_\pi=(X,Y,1)^T$ 与它们的图像 $\mathbf{x}$ 之间的映射是一个一般的平面单应(平面到平面的射影变换):$\mathbf{x}=H\mathbf{x}_\pi$,其中 H 是一个秩为 3 的 3×3 矩阵. 它表明:

• **在透视影像下，一张景物平面与一张图像平面之间最一般的变换是平面射影变换.**

如果是仿射摄像机，那么，类似的推导表明景物与图像平面由一个仿射变换相关联.

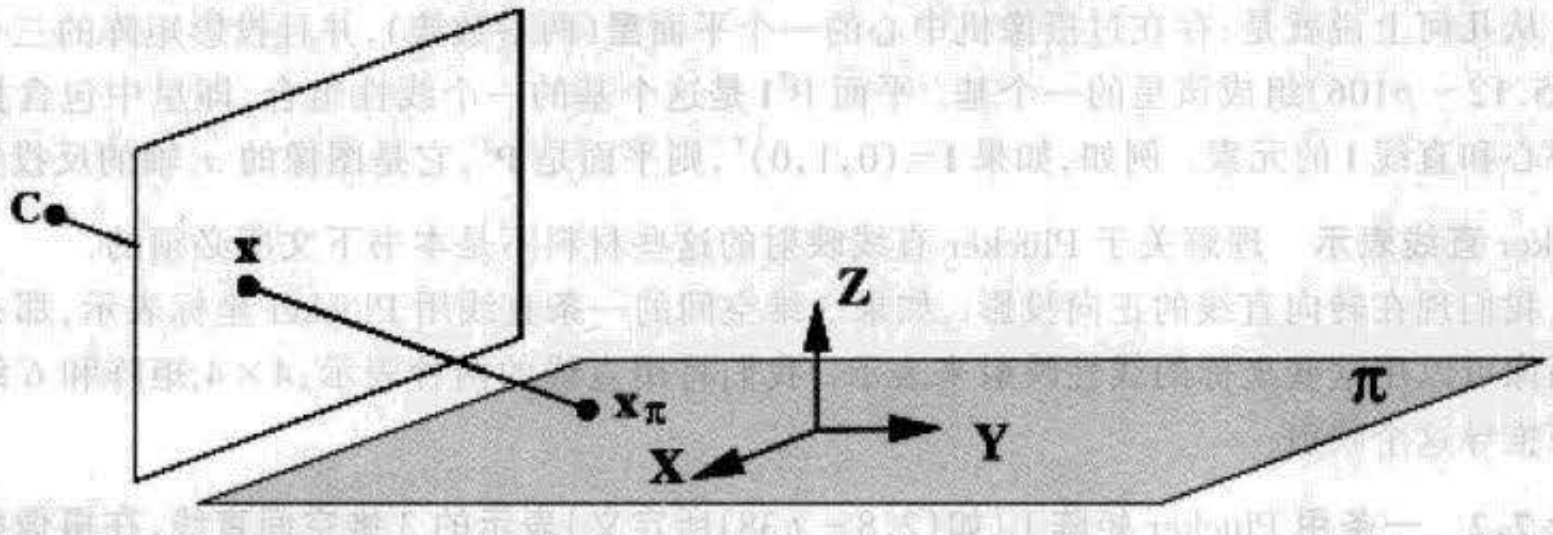



图 7.1 **平面上点的透视图像**. 世界坐标系的 XY－平面与平面 $\pi$ 一致. 图像和景物平面的点由一个平面射影变换相关联.

### 7.1.2 对直线的作用

**正向投影** 3 维空间的一条直线投影到图像平面上的一条直线. 从几何上不难看出该直线与摄像机中心确定一张平面，而图像是该平面和图像平面的交(图 7.2)，也可从代数上加以证明：设 **A** 和 **B** 是 3 维空间的点，而 **a**,**b** 是它们在 P 作用下的图像，那么连接 3 维空间点 **A** 和 **B** 的直线上的点 $\mathbf{X}(\mu)=\mathbf{A}+\mu\mathbf{B}$ 投影到连接 **a** 和 **b** 的直线上的点：

$$\begin{aligned}\mathbf{x}(\mu)=\mathrm{P}(\mathbf{A}+\mu\mathbf{B})&=\mathrm{P}\mathbf{A}+\mu\mathrm{P}\mathbf{B}\\&=\mathbf{a}+\mu\mathbf{b}\end{aligned}$$

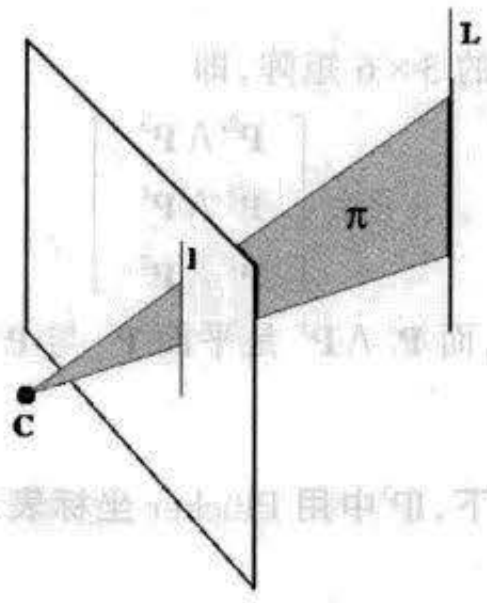



图 7.2 **直线投影**. 3 维空间的直线 **L** 由透视摄像机影像成直线 **l**. 图像直线 **l** 是平面 $\pi$ 和图像平面的交，其中平面 $\pi$ 由 **L** 和摄像机中心 **C** 确定. 反之，图像直线 **l** 反向投影成 3 维空间的平面 $\pi$. 该平面是图像直线的“反拉伸”.

**直线的反向投影** 空间中映射到图像上是一条直线的点集是由摄像机中心和图像直线确定的一张平面，如图 7.2 所示. 从代数上说，就是

**结论 7.1 经摄像机矩阵 P 映射成一条直线 l 的空间的点集是平面 $\mathrm{P}^{\mathrm{T}}\mathbf{l}$.**

**证明** 点 **x** 在 **l** 上的充要条件是 $\mathbf{x}^{\mathrm{T}}\mathbf{l}=0$. 空间点 **X** 的图像点 P**X** 在直线 **l** 上的充要条件是

$\mathbf{X}^{\mathsf{T}}\mathrm{P}^{\mathsf{T}}\mathbf{l}=0$. 因此如果令 $\mathrm{P}^{\mathsf{T}}\mathbf{l}$ 表示一张平面,那么 $\mathbf{X}$ 在此平面上的充要条件是 $\mathbf{X}$ 映射到 $\mathbf{l}$ 上的一个点. 换句话说,$\mathrm{P}^{\mathsf{T}}\mathbf{l}$ 是 $\mathbf{l}$ 的反向投影. □

从几何上说就是:存在过摄像机中心的一个平面星(两参数簇),并且投影矩阵的三个行 $\mathbf{P}^{i\mathsf{T}}$(5.12 - $p$106)组成该星的一个基. 平面 $\mathrm{P}^{\mathsf{T}}\mathbf{l}$ 是这个基的一个线性组合,即星中包含摄像机中心和直线 $\mathbf{l}$ 的元素. 例如,如果 $\mathbf{l}=(0,1,0)^{\mathsf{T}}$,则平面是 $\mathbf{P}^2$,它是图像的 $x$ 轴的反投影.

**Plücker 直线表示** 理解关于 Plücker 直线映射的这些材料不是本书下文所必须的.

我们现在转向直线的正向投影. 如果 3 维空间的一条直线用 Plücker 坐标表示,那么它的图像可以用这些坐标的线性映射来表示. 我们将用直线的两种表示:4×4 矩阵和 6 维矢量来推导这个映射.

**结论 7.2 一条用 Plücker 矩阵 L(如(2.8 - $p$38)所定义)表示的 3 维空间直线,在摄像机映射 P 作用下被映射成满足**

$$[\mathbf{l}]_\times = \mathrm{PLP}^{\mathsf{T}} \tag{7.1}$$

**的直线 $\mathbf{l}$,其中 $[\mathbf{l}]_\times$ 在(A3.4 - $p$410)中定义.**

**证明** 假定 $\mathbf{a}=\mathrm{PA}$,$\mathbf{b}=\mathrm{PB}$. 过 $\mathbf{A}$,$\mathbf{B}$ 的 3 维空间直线的 Plücker 矩阵 L 是 $\mathrm{L}=\mathbf{AB}^{\mathsf{T}}-\mathbf{BA}^{\mathsf{T}}$. 那么矩阵 $\mathrm{M}=\mathrm{PLP}^{\mathsf{T}}=\mathbf{ab}^{\mathsf{T}}-\mathbf{ba}^{\mathsf{T}}$ 是 3×3 反对称矩阵,并有零空间 $\mathbf{a}\times\mathbf{b}$. 因而 $\mathrm{M}=[\mathbf{a}\times\mathbf{b}]_\times$ 又因为过图像点的直线由 $\mathbf{l}=\mathbf{a}\times\mathbf{b}$ 给出,由此可以完成证明. □

显然,由(7.1)的形式可知在图像直线坐标 $l_i$ 与世界直线坐标 $L_{jk}$ 之间存在一种线性关系,但该关系的系数关于点投影矩阵 P 的元素是二次的. 因此,可以重排(7.1)使得 Plücker 线坐标 $\mathcal{L}$(6 维矢量)和图像的线坐标 $\mathbf{l}$(3 维矢量)之间的映射表示成一个 3×6 的矩阵. 可以证明

**定义 7.3 线投影矩阵 $\mathcal{P}$ 是秩 3 的 3×6 矩阵,即**

$$\mathcal{P}=\begin{bmatrix}\mathbf{P}^2\wedge\mathbf{P}^3\\ \mathbf{P}^3\wedge\mathbf{P}^1\\ \mathbf{P}^1\wedge\mathbf{P}^2\end{bmatrix} \tag{7.2}$$

其中 $\mathbf{P}^{i\mathsf{T}}$ 是点摄像机矩阵 P 的行,而 $\mathbf{P}^i\wedge\mathbf{P}^j$ 是平面 $\mathbf{P}^i$ 与 $\mathbf{P}^j$ 的交线的 Plücker 坐标.

于是,正向直线投影为

**结论 7.4 在线投影矩阵 $\mathcal{P}$ 作用下,$\mathbb{P}^3$ 中用 Plücker 坐标表示的直线 $\mathcal{L}$(如(2.11 - $p$39)所定义)被映射到图像直线**

$$\mathbf{l}=\mathcal{P}\mathcal{L}=\begin{bmatrix}(\mathbf{P}^2\wedge\mathbf{P}^3\mid\mathcal{L})\\ (\mathbf{P}^3\wedge\mathbf{P}^1\mid\mathcal{L})\\ (\mathbf{P}^1\wedge\mathbf{P}^2\mid\mathcal{L})\end{bmatrix} \tag{7.3}$$

**其中积($\mathcal{L}\mid\hat{\mathcal{L}}$)在(2.13 - $p$40)中定义.**

**证明** 假定该 3 维空间直线是点 $\mathbf{A}$ 和 $\mathbf{B}$ 的连线,而且这两点分别投影到 $\mathbf{a}=\mathrm{PA}$,$\mathbf{b}=\mathrm{PB}$. 那么其图像直线为 $\mathbf{l}=\mathbf{a}\times\mathbf{b}=(\mathrm{PA})\times(\mathrm{PB})$. 考虑 $\mathbf{l}$ 的第一个分量

$$l_1=(\mathbf{P}^{2\mathsf{T}}\mathbf{A})(\mathbf{P}^{3\mathsf{T}}\mathbf{B})-(\mathbf{P}^{2\mathsf{T}}\mathbf{B})(\mathbf{P}^{3\mathsf{T}}\mathbf{A})$$

$$= (\mathbf{P}^2 \wedge \mathbf{P}^3 \mid \mathcal{L})$$

其中第二个等式由(2.14 - $p$40)得到. 其他的分量可由类似的方式得到. □

线投影矩阵 $\mathcal{P}$ 对直线的作用如同 P 对点的作用. 在节 5.2.1($p$105)中把点摄像机矩阵 P 的行解释成平面,这里 $\mathcal{P}$ 的行在几何上解释成**直线**. P 的行 $\mathbf{P}^{i\mathsf{T}}$ 是摄像机的主平面和轴平面. $\mathcal{P}$ 的行是这些摄像机平面对的交. 例如 $\mathcal{P}$ 的第一行是 $\mathbf{P}^2 \wedge \mathbf{P}^3$,这是 $y=0$ 的轴平面 $\mathbf{P}^2$ 与主平面 $\mathbf{P}^3$ 的交线以 Plücker 线表示的 6 维矢量. 这三条线与 $\mathcal{P}$ 的三个行对应并且交于摄像机中心. 考虑 3 维空间满足 $\mathcal{P}\mathcal{L}=\mathbf{0}$ 的直线 $\mathcal{L}$. 这些直线在 $\mathcal{P}$ 的零空间中. 因为 $\mathcal{P}$ 的每行是一条直线,并且根据结论 2.5($p$40),如果两直线 $\mathcal{L}_1$,$\mathcal{L}_2$ 相交,则乘积($\mathcal{L}_1 \mid \mathcal{L}_2$)=0,所以 $\mathcal{L}$ 与 $\mathcal{P}$ 的行所表示的每条直线都相交. 这些直线是摄像机平面的交线,而摄像机中心是唯一在摄像机所有三张平面上的点. 因此我们得到

- **$\mathbb{P}^3$ 中满足 $\mathcal{P}\mathcal{L}=\mathbf{0}$ 的直线 $\mathcal{L}$ 必过摄像机中心.**

3×6 矩阵 $\mathcal{P}$ 有三维零空间. 在允许有齐次尺度因子的条件下,这个零空间是包含摄像机中心的两参数直线族. 这是可以预料的,因为 $\mathbb{P}^3$ 中存在共点的直线星(两参数簇).

### 7.1.3 对二次曲线的作用

**二次曲线的反向投影** 一条二次曲线 C 反向投影成一个锥面. 锥面是一种退化的二次曲面,即表示该二次曲面的 4×4 矩阵不满秩. 锥面的顶点(现在就是摄像机中心)是该二次曲面矩阵的零空间.

**结论 7.5 在摄像机 P 作用下,二次曲线 C 反向投影成锥面**

$$\mathrm{Q}_{\mathrm{CO}} = \mathrm{P}^{\mathsf{T}}\mathrm{C}\mathrm{P}.$$

**证明** 点 $\mathbf{x}$ 在 C 上的充要条件是 $\mathbf{x}^{\mathsf{T}}\mathrm{C}\mathbf{x}=0$. 空间点 $\mathbf{X}$ 映射到点 $\mathrm{P}\mathbf{X}$,它在二次曲线上的充要条件是 $\mathbf{X}^{\mathsf{T}}\mathrm{P}^{\mathsf{T}}\mathrm{C}\mathrm{P}\mathbf{X}=0$. 因此如果用 $\mathrm{Q}_{\mathrm{CO}}=\mathrm{P}^{\mathsf{T}}\mathrm{C}\mathrm{P}$ 来表示一个二次曲面,则 $\mathbf{X}$ 在此二次曲面上的充要条件是 $\mathbf{X}$ 映射为二次曲线 C 上的一个点. 换句话说,$\mathrm{Q}_{\mathrm{CO}}$ 是二次曲线 C 的反向投影. □

注意摄像机中心 $\mathbf{C}$ 是退化二次曲面的顶点,因为 $\mathrm{Q}_{\mathrm{CO}}\mathbf{C}=\mathrm{P}^{\mathsf{T}}\mathrm{C}(\mathrm{P}\,\mathbf{C})=\mathbf{0}$.

**例 7.6** 假定 $\mathrm{P}=\mathrm{K}[\mathrm{I}|\mathbf{0}]$;那么二次曲线 C 反向投影为锥面

$$\mathrm{Q}_{\mathrm{CO}} = \begin{bmatrix} \mathrm{K}^{\mathsf{T}} \\ \mathbf{0}^{\mathsf{T}} \end{bmatrix} \mathrm{C}[\mathrm{K}|\mathbf{0}] = \begin{bmatrix} \mathrm{K}^{\mathsf{T}}\mathrm{C}\mathrm{K} & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & 0 \end{bmatrix}.$$

矩阵 $\mathrm{Q}_{\mathrm{CO}}$ 的秩为 3. 它的零矢量是摄像机中心 $\mathbf{C}=(0,0,0,1)^{\mathsf{T}}$. △

## 7.2 光滑曲面的图像

一个光滑曲面 $S$ 的图像的外形线由影像射线与曲面的**切点**产生,如图 7.3 所示. 类似地,外形线的切线反向投影成曲面的**切面**.

**定义 7.7 轮廓生成元 $\Gamma$** 定义为影像射线与 $S$ 的所有切点 $\mathbf{X}$ 的集合. 在图像上,与它对应

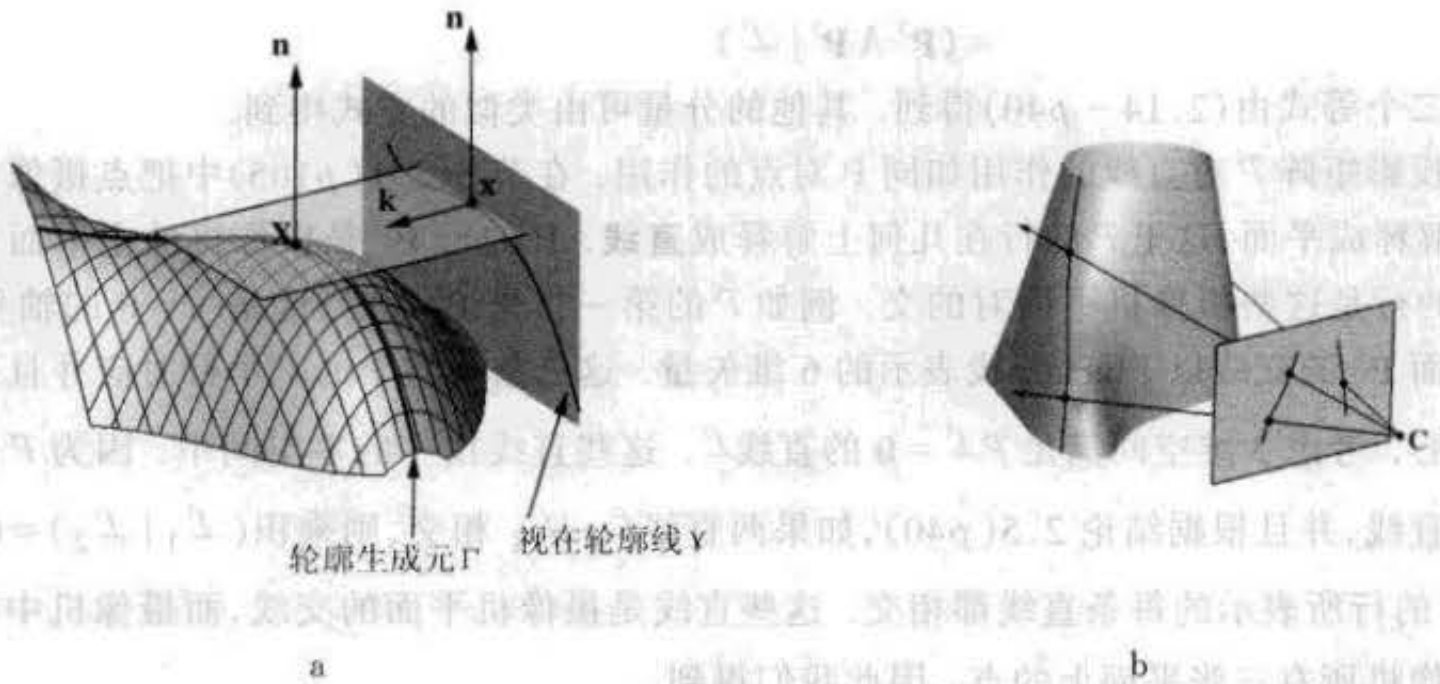



**图 7.3 轮廓生成元和视在轮廓线.** (a)平行投影情形 (b)中心投影情形. 由摄像机中心到 **x** 的射线与曲面相切于 **X**. 此类切点 **X** 的集合定义了轮廓生成元,而它们的图像定义视在轮廓线. 轮廓生成元一般是一条空间曲线. 此图的引用得到 Roberto Cipolla 和 Peter Giblin 的允许.

的**视在轮廓线** $\boldsymbol{\gamma}$ 是 **X** 的图像点 **x** 的集合,即 $\boldsymbol{\gamma}$ 是 $\boldsymbol{\Gamma}$ 的图像.

视在轮廓线也称"外形线"或"轮廓". 如果沿摄像机中心到 **X** 的方向去观察曲面,那么曲面看上去受到折迭,或有一个边界线或封闭轮廓.

显然,轮廓生成元 $\boldsymbol{\Gamma}$ 仅取决于摄像机中心与曲面的相对位置,而与图像平面无关. 但是,视在轮廓线 $\boldsymbol{\gamma}$ 由图像平面与轮廓生成元的射线的交确定,因此与图像平面的位置有关.

考虑沿方向 **k** 的平行投影的情形,所有与 **k** 平行的射线与 $S$ 相切,见图 7.3a. 这些射线形成一个切射线的"柱面",该柱面与 S 的切点组成的曲线就是轮廓生成元 $\boldsymbol{\Gamma}$. 该柱面与图像平面相交的曲线就是视在轮廓线 $\boldsymbol{\gamma}$. 注意 $\boldsymbol{\Gamma}$ 和 $\boldsymbol{\gamma}$ 本质上都依赖于 **k**. 当方向 **k** 变化时,集合 $\boldsymbol{\Gamma}$ 在曲面上滑动. 例如,当 $S$ 是一个球时,$\boldsymbol{\Gamma}$ 是正交于 **k** 的大圆. 在此情形下,$\boldsymbol{\Gamma}$ 是一条平面曲线,但 $\boldsymbol{\Gamma}$ 一般是一条空间曲线.

下面,我们介绍二次曲面的投影性质. 这类曲面轮廓生成元和视在轮廓线可用代数表达式来推导.

## 7.3 射影摄像机对二次曲面的作用

二次曲面是一种光滑的曲面,因此它的外形曲线由反向投影射线与曲面相切的点产生,如图 7.4 所示.

假如二次曲面是一个球面,那么摄像机中心与二次曲面之间的射线锥面是一个正圆锥面,即轮廓生成元是圆,该圆所在平面与摄像机中心和球心的连线垂直. 这一点可以由该几何绕这条直线是旋转对称的性质推知. 球面的图像由该圆锥面与图像平面的交得到. 显然,这是一个经典的圆锥截线,所以一个球的视在轮廓线是一条二次曲线. 特别是如果球心在摄像机的主轴(Z 轴)上,那么该二次曲线是圆周.

现在考虑上述几何实体的 3 维射影变换. 在射影映射下,球面变换成二次曲面而视在轮廓线变换成一条二次曲线. 但是,因为相交和相切性被保留,所以轮廓生成元是一条(平面)二次曲线. 因此,一般二次曲面的视在轮廓线是一条二次曲线而且轮廓生成元也是一条

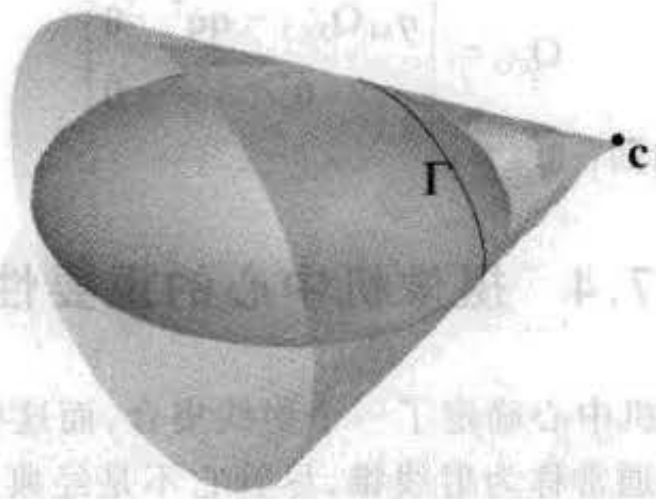



图 7.4 **二次曲面的射线锥**. 锥的顶点是摄像机中心. 二次曲面的轮廓生成元是平面曲线(二次曲线), 它是二次曲面与摄像机中心 C 的极平面的交线.

二次曲线. 我们现在给出这些几何结果的代数表达.

**二次曲面的正向投影** 既然外形线由相切性产生,那么对偶二次曲面 $Q^*$ 自然会在这里起重要作用,因为二次曲面 Q 的切平面由它定义.

**结论 7.8 在摄像机矩阵 P 的作用下,二次曲面 Q 的外形线是下式给定的一条二次曲线 C:**

$$C^* = PQ^*P^T. \tag{7.4}$$

**证明** 该表达式可以由如下事实直接导出:二次曲线 C 的切线 $\mathbf{l}$ 满足 $\mathbf{l}^T C^* \mathbf{l} = 0$. 这些直线反向投影到与该二次曲面相切的平面 $\boldsymbol{\pi} = P^T\mathbf{l}$ 并且满足 $\boldsymbol{\pi}^T Q^* \boldsymbol{\pi} = 0$. 事实上对每条直线有

$$\begin{aligned}\boldsymbol{\pi}^T Q^* \boldsymbol{\pi} &= \mathbf{l}^T PQ^* P^T \mathbf{l} \\ &= \mathbf{l}^T C^* \mathbf{l} = 0\end{aligned}$$

而且,由于它对 C 的所有切线都成立,从而这个结论得到证明. □

注意(7.4)与由 Plücker 矩阵(7.1)表示的直线投影的相似性. 点二次曲面 Q 的投影的表达式可以由(7.4)导出,但相当复杂. 不过轮廓生成元的平面却可以容易地用 Q 表示:

- **与二次曲面 Q 和中心为 C 的摄像机相对应的轮廓生成元 Γ 的平面是 $\boldsymbol{\pi}_\Gamma = QC$.**

这个结论直接来自节(2.2.3)($p40$)关于点和二次曲面的极点和极线的关系. 它的证明留作练习. 注意,二次曲面与平面的交是二次曲线. 因此如上文所述,Γ 是二次曲线而且它的图像 $\gamma$(即视在轮廓线)也是二次曲线.

我们同样可以导出由摄像机中心与二次曲面形成的射线锥面的表达式. 这个锥面是秩为 3 的退化的二次曲面.

**结论 7.9 顶点为 V 并与二次曲面 Q 相切的锥面是一个退化的二次曲面**

$$Q_{CO} = (\mathbf{V}^T Q \mathbf{V})Q - (Q\mathbf{V})(Q\mathbf{V})^T.$$

注意 $Q_{CO}\mathbf{V} = \mathbf{0}$. 因而 V 正是所要求的锥面的顶点. 具体证明省略.

**例 7.10** 我们用分块形式记该二次曲面为

$$Q = \begin{bmatrix} Q_{3\times 3} & \mathbf{q} \\ \mathbf{q}^T & q_{44} \end{bmatrix}.$$

如果对应于锥面的顶点 $\mathbf{V} = (0,0,0,1)^T$ 是世界坐标系的中心,那么

$$Q_{CO}=\begin{bmatrix} q_{44}Q_{3\times 3}-\mathbf{q}\mathbf{q}^{\mathsf{T}} & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & 0 \end{bmatrix}$$

显然这是一个退化的二次曲面. △

## 7.4 摄像机中心的重要性

3维空间的物体和摄像机中心确定了一个射线集合,而这些射线与一张平面的交就产生该物体的图像. 这个集合通常称为射线锥,尽管它不是经典意义上的圆锥. 如图7.5所示,假如射线锥与两张平面相交,那么所得两幅图像 $I$ 和 $I'$ 显然以一个透视映射相关联. 这表明由同一摄像机中心所获得的图像可以通过一个平面射影变换而相互映射,换句话说,它们是射影等价的,因而有相同的射影性质. 因此,摄像机可以看成一种射影影像的装置——度量以摄像机中心为顶点的射线锥的射影性质.

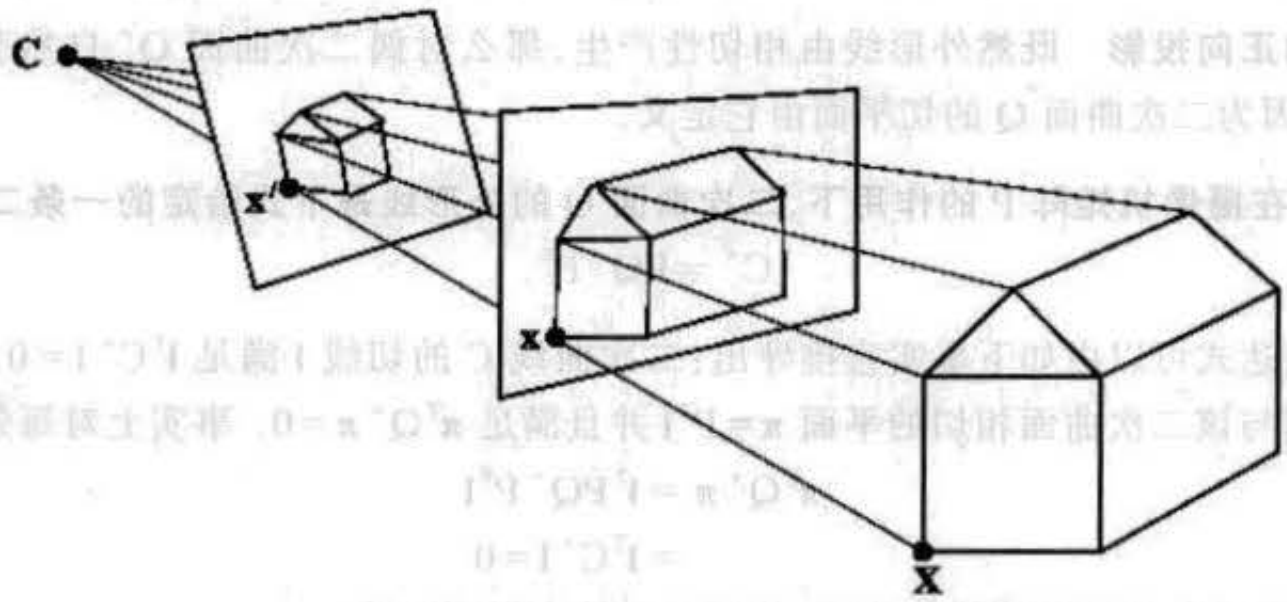



图7.5 **以摄像机中心为顶点的射线锥**. 这个锥与一张平面的交产生一幅图像. 3维空间点 $\mathbf{X}$ 和摄像机中心 $\mathbf{C}$ 之间的射线在图像点 $\mathbf{x}$ 和 $\mathbf{x}'$ 处穿过该平面. 所有这样的像点以一个平面单应 $\mathbf{x}'=\mathrm{H}\mathbf{x}$ 相关联.

现在用代数方法来推导两幅图像 $I$ 与 $I'$ 以一个单应相关联的结论并求这个单应的公式. 设两个摄像机

$$\mathrm{P}=\mathrm{KR}[\mathrm{I}|-\tilde{\mathbf{C}}],\quad \mathrm{P}'=\mathrm{K}'\mathrm{R}'[\mathrm{I}|-\tilde{\mathbf{C}}]$$

具有共同的中心. 注意因为摄像机有共同的中心,故它们之间存在一个简单的关系,即 $\mathrm{P}'=(\mathrm{K}'\mathrm{R}')(\mathrm{KR})^{-1}\mathrm{P}$. 从而由这两个摄像机产生的3维空间点 $\mathbf{X}$ 的图像点之间的关系如下

$$\mathbf{x}'=\mathrm{P}'\mathbf{X}=(\mathrm{K}'\mathrm{R}')(\mathrm{KR})^{-1}\mathrm{P}\mathbf{X}=(\mathrm{K}'\mathrm{R}')(\mathrm{KR})^{-1}\mathbf{x}.$$

也就是说,对应的图像点以一个形如 $\mathbf{x}'=\mathrm{H}\mathbf{x}$ 的平面单应($3\times 3$ 矩阵)相关联,其中 $\mathrm{H}=(\mathrm{K}'\mathrm{R}')(\mathrm{KR})^{-1}$.

我们现在来研究固定摄像机中心而移动图像平面的几种情形. 为简单起见,选择世界坐标系与摄像机坐标系相一致,因而 $\mathrm{P}=\mathrm{K}[\mathrm{I}|\mathbf{0}]$(并假定图像平面不包含中心,否则图像将是退化的).

### 7.4.1 移动图像平面

首先考虑焦距增加的情形. 在一阶近似的精度下,它相当于图像平面沿主轴移动. 而在图像上的效果是简单的放大. 这仅是一阶近似,因为组合透镜的变焦会同时引起主点和

摄像机的有效中心的摄动. 从代数上说,如果点 $\mathbf{X}$ 在变焦前后的图像分别是 $\mathbf{x}$ 与 $\mathbf{x}'$,那么

$$\mathbf{x}=\mathrm{K}[\mathrm{I}|\mathbf{0}]\mathbf{X}$$

$$\mathbf{x}'=\mathrm{K}'[\mathrm{I}|\mathbf{0}]\mathbf{X}=\mathrm{K}'\mathrm{K}^{-1}(\mathrm{K}[\mathrm{I}|\mathbf{0}]\mathbf{X})=\mathrm{K}'\mathrm{K}^{-1}\mathbf{x}$$

因而 $\mathbf{x}'=\mathrm{H}\mathbf{x}$,其中 $\mathrm{H}=\mathrm{K}'\mathrm{K}^{-1}$. 如果 K 和 K′之间仅是焦距的不同,那么由简单计算可知

$$\mathrm{K}'\mathrm{K}^{-1}=\begin{bmatrix} k\mathrm{I} & (1-k)\tilde{\mathbf{x}}_0 \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}.$$

其中$\tilde{\mathbf{x}}_0$ 是非齐次主点,而 $k=f'/f$ 是放大因子. 这个结果可以从相似三角形直接得到:变焦因子 $k$ 的效果是使图像点$\tilde{\mathbf{x}}$沿由主点$\tilde{\mathbf{x}}_0$ 到点$\tilde{\mathbf{x}}'=k\,\tilde{\mathbf{x}}+(1-k)\tilde{\mathbf{x}}_0$ 的射线移动. 从代数上说,利用标定矩阵 K 的最一般形式(5.10 – $p$104),我们可以写出

$$\mathrm{K}'=\begin{bmatrix} k'\mathrm{I} & (1-k)\tilde{\mathbf{x}}_0 \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}\mathrm{K}=\begin{bmatrix} k\mathrm{I} & (1-k)\tilde{\mathbf{x}}_0 \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}\begin{bmatrix} \mathrm{A} & \tilde{\mathbf{x}}_0 \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}$$

$$=\begin{bmatrix} k\mathrm{A} & \tilde{\mathbf{x}}_0 \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}=\mathrm{K}\begin{bmatrix} k\mathrm{I} & \\ & 1 \end{bmatrix}.$$

它表明

- **因子 $k$ 的变焦效果等于用 diag($k$,$k$,1)右乘摄像机标定矩阵 K.**

### 7.4.2 摄像机旋转

第二个常用的例子是摄像机在不改变它的内部参数的情况下绕它的中心旋转. 这种"纯"旋转的例子在图 7.6 和图 7.9 中给出. 代数上说,如果点 $\mathbf{X}$ 在纯旋转前和后的图像是 $\mathbf{x}$ 和 $\mathbf{x}'$,那么

$$\mathbf{x}=\mathrm{K}[\mathrm{I}|\mathbf{0}]\mathbf{X}$$

$$\mathbf{x}'=\mathrm{K}[\mathrm{R}|\mathbf{0}]\mathbf{X}=\mathrm{K}\mathrm{R}\mathrm{K}^{-1}\mathrm{K}[\mathrm{I}|\mathbf{0}]\mathbf{X}=\mathrm{K}\mathrm{R}\mathrm{K}^{-1}\mathbf{x}$$

因而 $\mathbf{x}'=\mathrm{H}\mathbf{x}$,其中 $\mathrm{H}=\mathrm{K}\mathrm{R}\mathrm{K}^{-1}$. 这种单应是一个**共轭旋转**,并将在节 A5.1($p$427)中作进一步讨论. 现在我们通过举例来给出它的若干性质.

a b c

图 7.6 在图像(a)与(b)之间,摄像机绕它的中心作了旋转. 对应点(即同一 3D 点的图像)以一个平面射影变换相关联. 注意不同深度的 3D 点,例如杯子的口和六脚器,在(a)中是重合点,在(b)中也是重合点,因此其中不存在运动视差. 但是,在图像(a)和(c)之间摄像机既绕中心旋转又作平移. 在这种一般运动之下,在(a)中不同深度的重合点在(c)中却被影像成为不同的点,因此存在运动视差.

**例 7.11 共轭旋转的性质**

单应 $\mathrm{H}=\mathrm{K}\mathrm{R}\mathrm{K}^{-1}$ 与旋转矩阵 R 有相同的特征值$\{\mu,\mu e^{i\theta},\mu e^{-i\theta}\}$(相差一个常数因子),

其中 $\mu$ 是未知的尺度因子(如果 H 经尺度变换使得 $\det(H)=1$,那么 $\mu=1$). 因此,视图之间的旋转角可以直接由 H 的复特征值的相位算出. 类似地,可以证明(见练习)与 H 的实特征值对应的特征矢量是旋转轴的消影点.

例如,图 7.6 的图像(a)与(b)之间有一个摄像机的纯旋转. 单应 H 由算法 3.6($p$79)计算,算得的旋转角度是 4.66°,而消影点的轴是$(-0.0088,1,0.0001)^{\mathrm{T}}$,它实质上是 $y$ 方向的无穷远点. 因此旋转轴几乎与 $y$-轴平行. △

变换 $H=KRK^{-1}$是**无穷单应**映射 $H_\infty$ 的一个例子,此种映射将在本书中多次出现并在节 12.4($p$239)中给出定义. 第 18 章把此共轭性质用于自标定.

### 7.4.3 应用与举例

具有同一摄像机中心的图像之间的单应关系能以若干方式加以利用. 一种应用是通过射影形变插补来合成图像. 另一个应用是拼图,由旋转摄像机获得的视图可以用平面单应来"缝"成一个全景图像.

**例 7.12 视图合成**

可以对现有的图像进行平面单应性的形变插补来产生新的图像,新的图像对应于不同的摄像机定向(但摄像机中心相同).

在正视图中,长方形被映射成长方形,并且世界与图像的长方形的长宽比相同. 反过来,一个正视图可以通过对图像进行单应性的形变插补来合成,该单应把四边形图像映射回具有正确的长宽比的矩形. 其算法是:

(1) 计算把图像中的四边形映射回具有正确长宽比矩形的单应 H.

(2) 用这个单应对原图像进行射影形变插补.

相应的例子在图 7.7 中给出. △

a b c

图 7.7 **合成视图**. (a)原始图像. (b)走廊地板的正视图,它由(a)产生,用地板块的四个拐角计算有关单应. (c)走廊墙的正视图,它由(a)产生,用门框的四个拐角计算有关单应.

**例 7.13 平面全景拼图**

摄像机绕其中心旋转所得到的一组图像之间以一个平面单应相关联. 这样的一组图像可以通过射影形变插补与其中一幅图像的平面配准,如图 7.8 所示.

其算法要点是

(1) 选择图像集合中的一幅作为参考图像.

(2) 在其余的图像中选一幅图像,计算把它映射到参考图像的单应 H.

(3) 用得到的单应对这幅图像进行射影形变插补,并且用插补图像与参考图像的非重叠部分来扩大参考图像.

(4) 对图像集合中余下的图像重复上述最后两步过程.

所需单应可以通过辨认四组对应点,或采用算法 3.6($p$79) 的自动方法来算出. 图 7.9 给出拼图的一个例子. △

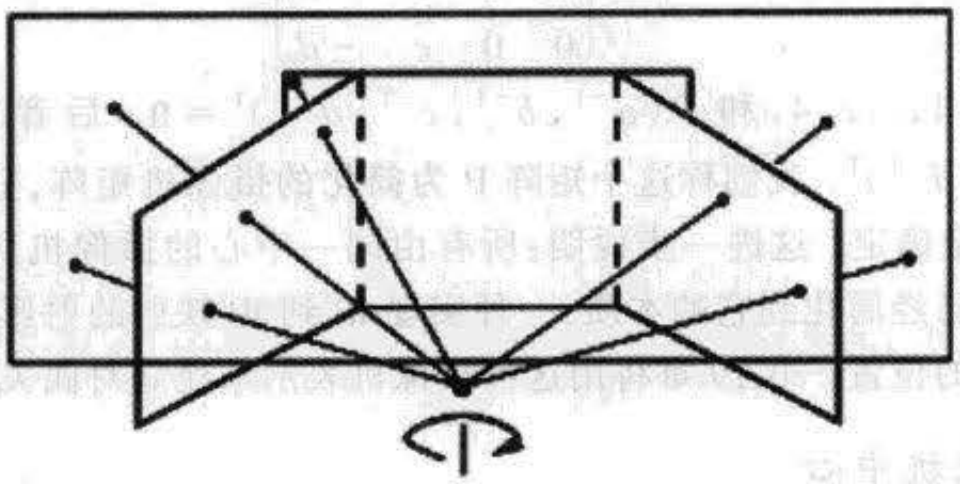


图 7.8. 对旋转摄像机获得的三幅图像进行配准,如图所示,通过射影形变插补将外边的两幅图像与中间的图像配准.

图 7.9 **平面全景拼图**. 由便携式摄像机绕其中心拍摄的(三十幅中的)八幅图像. 这三十幅图像用平面单应配准并形成一幅全景拼图. 注意"蝴蝶领结"式的特征形状,这是因为与图像序列的中间一幅图像配准的缘故.

### 7.4.4 (简化的)射影记号

在第 19 章中我们将会看到如果世界和图像点都选择规范的射影坐标,即

$$\mathbf{X}_1=(1,0,0,0)^{\mathsf{T}},\quad \mathbf{X}_2=(0,1,0,0)^{\mathsf{T}},\quad \mathbf{X}_3=(0,0,1,0)^{\mathsf{T}},\quad \mathbf{X}_4=(0,0,0,1)^{\mathsf{T}},$$

和

$$\mathbf{x}_1=(1,0,0)^{\mathsf{T}},\quad \mathbf{x}_2=(0,1,0)^{\mathsf{T}},\quad \mathbf{x}_3=(0,0,1)^{\mathsf{T}},\quad \mathbf{x}_4=(1,1,1)^{\mathsf{T}},$$

那么摄像机矩阵

$$\mathrm{P}=\begin{bmatrix} a & 0 & 0 & -d \\ 0 & b & 0 & -d \\ 0 & 0 & c & -d \end{bmatrix} \tag{7.5}$$

满足 $\mathbf{x}_i=\mathrm{P}\mathbf{X}_i$, $i=1,\ldots,4$, 和 $\mathrm{P}(a^{-1},b^{-1},c^{-1},d^{-1})^{\mathsf{T}}=\mathbf{0}$, 后者表示摄像机中心是 $\mathbf{C}=(a^{-1},b^{-1},c^{-1},d^{-1})^{\mathsf{T}}$. 我们称这个矩阵 P 为**简化的**摄像机矩阵,显然它由摄像机中心 $\mathbf{C}$ 的三个自由度完全确定. 这进一步说明:所有由同一中心的摄像机所获取的图像是射影等价的——摄像机已经简化到它的本质:一种实现 $\mathbb{P}^3$到 $\mathbb{P}^2$映射的射影装置,唯一影响其结果的是摄像机中心的位置. 第 19 章将用这种摄像机表示来建立对偶关系.

### 7.4.5 移动摄像机中心

摄像机变焦和旋转的情况说明:移动图像平面但同时保持摄像机中心不动所诱导的图像之间的变换仅与图像平面的运动有关而与 3 维空间的结构**无关**. 因而,不能从这个行动中获得 3 维空间的任何信息. 但是,如果移动了摄像机中心,那么图像对应点之间的映射便与其 3 维空间结构有关,并且可用来(部分地)确定该空间结构. 这正是本书后面许多内容将涉及的主题.

我们如何只从图像来判断摄像机中心是否移动呢? 考虑在第一幅视图中重合的两个 3 维空间点,即在同一射线上的点. 如果摄像机中心被移动了(没有沿着这条射线),那么其图像就不再重合了. 原来是重合的图像点之间的相对位移称为**视差**,图 7.6 给出过说明,图 7.10更以图解方式加以说明. 如果景物是静止的而在两幅视图中有明显的运动视差,那么摄像机中心被移动了. 其实,在摄像机纯旋转时,一种获得它的运动参数的便利方法就是消除所有的视差.

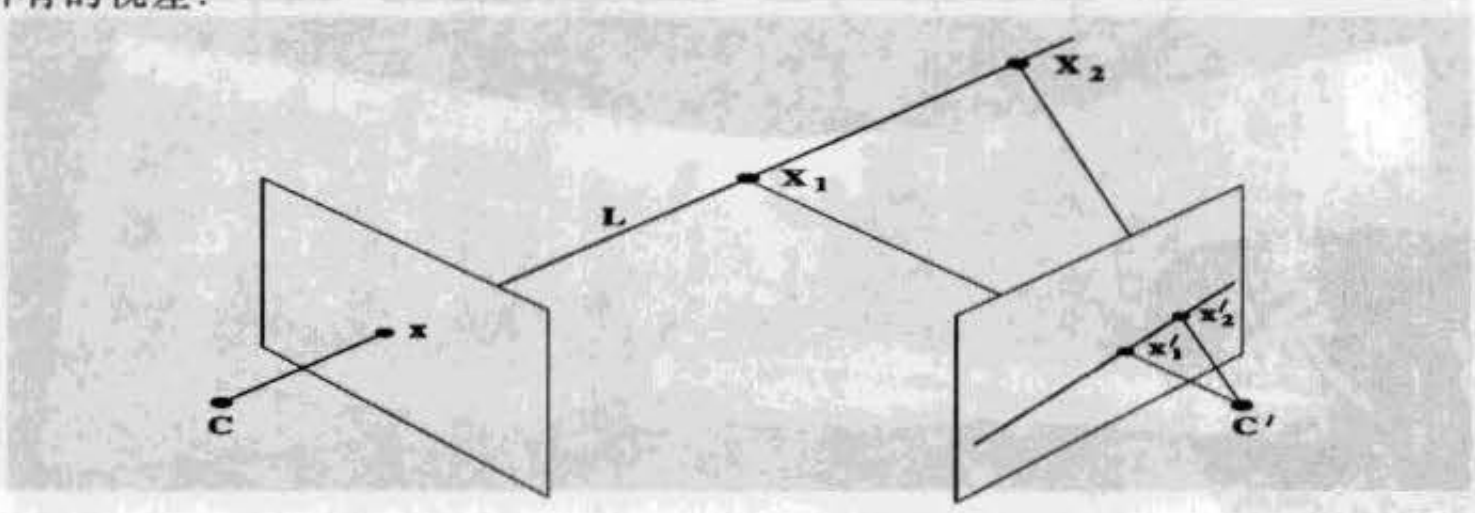



图 7.10 **运动视差**. 从以 $\mathbf{C}$ 为中心的摄像机观看时,空间点 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 的图像是重合的. 而从以 $\mathbf{C}'$ 为中心的摄像机观看时,只要 $\mathbf{C}'$不在过 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 的直线 $\mathbf{L}$ 上,两空间点的图像必然不重合. 事实上,过图像点 $\mathbf{x}_1'$和 $\mathbf{x}_2'$的直线正是射线 $\mathbf{L}$ 的影像,在第 8 章中我们将看到该直线是一条**对极线**. 点 $\mathbf{x}_1'$和 $\mathbf{x}_2'$之间的矢量是视差.

3 维空间结构的一种重要的特殊情形是所有的景物点都共面. 对于这种情形,即使移

动摄像机中心，图像的对应点仍以平面单应相关联. 在这种情形下，图像之间的映射将在关于平面问题的第 12 章作详细讨论. 特别是消影点(平面 $\pi_\infty$ 上点的图像点)在任何摄像机运动下都由一个平面单应相关联. 我们将在节 7.6 中再来讨论这个问题.

## 7.5 摄像机标定与绝对二次曲线的图像

节 7.4 证明了射影摄像机是一种射影成像装置. 图像的射影性质仅与摄像机中心的位置有关，因为具有相同摄像机中心的所有图像在相差一个平面射影变换的意义下等价. 这里，我们讨论使摄像机成为能测量两条射线间夹角的一种欧氏影像装置所需的要素.

**标定给出了什么？** 图像中的一个点 $\mathbf{x}$ 的反向投影是由该点 $\mathbf{x}$ 和摄像机中心确定的一条射线. 标定把此图像点与此射线的**方向**相关联. 在摄像机的欧氏坐标系中，假定射线上的点被记为 $\tilde{\mathbf{X}}=\lambda\mathbf{d}$，那么这些点映射为点 $\mathbf{x}=\mathrm{K}[\mathrm{I}|\mathbf{0}](\lambda\mathbf{d}^\mathsf{T},1)^\mathsf{T}=\mathrm{K}\mathbf{d}$. 因此，

**结论 7.14 摄像机标定矩阵 K 是 $\mathbf{x}$ 与在摄像机的欧氏坐标系下测量的射线方向 $\mathbf{d}=\mathrm{K}^{-1}\mathbf{x}$ 之间的一个(仿射)变换.**

因此若不计常数因子的差别，方向 $\mathbf{d}$ 由图像点 $\mathbf{x}$ 确定，即 $\mathbf{d}=\mathrm{K}^{-1}\mathbf{x}$. 注意 $\mathbf{d}=\mathrm{K}^{-1}\mathbf{x}$ 一般不是单位矢量.

令两条射线的方向 $\mathbf{d}_1$，$\mathbf{d}_2$ 分别与图像点 $\mathbf{x}_1$，$\mathbf{x}_2$ 对应，它们之间的角度可以由我们熟悉的两矢量夹角的余弦公式得到：

$$\cos\theta=\frac{\mathbf{d}_1^\mathsf{T}\mathbf{d}_2}{\sqrt{\mathbf{d}_1^\mathsf{T}\mathbf{d}_1}\sqrt{\mathbf{d}_2^\mathsf{T}\mathbf{d}_2}}=\frac{(\mathrm{K}^{-1}\mathbf{x}_1)^\mathsf{T}(\mathrm{K}^{-1}\mathbf{x}_2)}{\sqrt{(\mathrm{K}^{-1}\mathbf{x}_1)^\mathsf{T}(\mathrm{K}^{-1}\mathbf{x}_1)}\sqrt{(\mathrm{K}^{-1}\mathbf{x}_2)^\mathsf{T}(\mathrm{K}^{-1}\mathbf{x}_2)}}$$

$$=\frac{\mathbf{x}_1^\mathsf{T}(\mathrm{K}^{-\mathsf{T}}\mathrm{K}^{-1})\mathbf{x}_2}{\sqrt{\mathbf{x}_1^\mathsf{T}(\mathrm{K}^{-\mathsf{T}}\mathrm{K}^{-1})\mathbf{x}_1}\sqrt{\mathbf{x}_2^\mathsf{T}(\mathrm{K}^{-\mathsf{T}}\mathrm{K}^{-1})\mathbf{x}_2}}. \tag{7.6}$$

公式(7.6) 表明如果 K，从而矩阵 $\mathrm{K}^{-\mathsf{T}}\mathrm{K}^{-1}$ 已知，那么射线之间的夹角可以从它们的对应图像点得到测量. 已知 K 的摄像机称为**已标定**的摄像机. 一个已标定的摄像机是一个能够测量射线方向的**方向传感器**——像一个 2D 量角器一样.

标定矩阵 K 也提供图像直线和景物平面之间的关系：

**结论 7.15 一条图像直线 $\mathbf{l}$ 确定一张过摄像机中心的平面，在摄像机的欧氏坐标系下，该平面的法线方向为 $\mathbf{n}=\mathrm{K}^\mathsf{T}\mathbf{l}$.**

注意此法线 $\mathbf{n}$ 一般不是单位矢量.

**证明** 直线 $\mathbf{l}$ 上的点 $\mathbf{x}$ 反向投影到与平面的法线垂直的方向 $\mathbf{d}=\mathrm{K}^{-1}\mathbf{x}$ 上，即 $\mathbf{d}^\mathsf{T}\mathbf{n}=\mathbf{x}^\mathsf{T}\mathrm{K}^{-\mathsf{T}}\mathbf{n}=0$. 因为在 $\mathbf{l}$ 上的点满足 $\mathbf{x}^\mathsf{T}\mathbf{l}=0$，从而得到 $\mathbf{l}=\mathrm{K}^{-\mathsf{T}}\mathbf{n}$，因此 $\mathbf{n}=\mathrm{K}^\mathsf{T}\mathbf{l}$. □

这个结果今后将多次出现. 现在，我们来考虑 $\pi_\infty$ 上几何实体的图像.

### 7.5.1 绝对二次曲线的图像

我们先来推导一个把标定矩阵 K 与绝对二次曲线的图像 $\omega$ 相关联的非常重要的结果. 我们必须首先确定无穷远平面 $\pi_\infty$ 和摄像机图像平面之间的映射. $\pi_\infty$ 上的点可以写为 $\mathbf{X}_\infty=(\mathbf{d}^\mathsf{T},0)^\mathsf{T}$，并被一般摄像机 $\mathrm{P}=\mathrm{KR}[\mathrm{I}|-\tilde{\mathbf{C}}]$ 影像为

$$\mathbf{x}=\mathrm{P}\mathbf{X}_\infty=\mathrm{KR}[\mathrm{I}\mid-\tilde{\mathbf{C}}]\begin{pmatrix}\mathbf{d}\\0\end{pmatrix}=\mathrm{KR}\mathbf{d}.$$

它表明

- **$\pi_\infty$ 与图像平面之间的映射由平面单应 $\mathbf{x}=\mathrm{H}\mathbf{d}$ 给出，其中**

$$\mathrm{H}=\mathrm{KR} \tag{7.7}$$

注意这个映射与摄像机的位置 $\mathbf{C}$ 无关，而仅取决于摄像机的内标定及其相对于世界坐标系的定向.

既然绝对二次曲线 $\Omega_\infty$（节 2.6（$p47$））在 $\pi_\infty$ 上，我们就可以计算它在 H 变换下的图像，并得到

**结论 7.16　绝对二次曲线的图像（简称 IAC）是二次曲线 $\omega=(\mathrm{KK}^\mathrm{T})^{-1}=\mathrm{K}^{-\mathrm{T}}\mathrm{K}^{-1}$.**

**证明**　根据结论 1.13（$p11$），在点单应变换 $\mathbf{x}\mapsto\mathrm{H}\mathbf{x}$ 下，二次曲线 C 映射为 $\mathrm{C}\mapsto\mathrm{H}^{-\mathrm{T}}\mathrm{C}\mathrm{H}^{-1}$. 由此推出 $\Omega_\infty$（它是 $\pi_\infty$ 上的二次曲线 $\mathrm{C}=\Omega_\infty=\mathrm{I}$）映射为 $\omega=(\mathrm{KR})^{-\mathrm{T}}\mathrm{I}(\mathrm{KR})^{-1}=\mathrm{K}^{-\mathrm{T}}\mathrm{RR}^{-1}\mathrm{K}^{-1}=(\mathrm{KK}^\mathrm{T})^{-1}$. 所以 IAC 是 $\omega=(\mathrm{KK}^\mathrm{T})^{-1}$. □

与 $\Omega_\infty$ 一样，二次曲线 $\omega$ 是虚（没有实点）的点二次曲线. 现在我们可以把它看作一种方便的代数工具，在本章后面的计算以及第 18 章摄像机的自标定中都会用到它. 这里提出几点注意：

(1) 绝对二次曲线的图像 $\omega$ 仅与矩阵 P 的内参数 K 有关；而与摄像机的定向或位置无关.

(2) 按(7.6)可以用下列简单的表达式给出两条射线的夹角：

$$\cos\theta=\frac{\mathbf{x}_1^\mathrm{T}(\mathrm{K}^{-\mathrm{T}}\mathrm{K}^{-1})\mathbf{x}_2}{\sqrt{\mathbf{x}_1^\mathrm{T}(\mathrm{K}^{-\mathrm{T}}\mathrm{K}^{-1})\mathbf{x}_1}\sqrt{\mathbf{x}_2^\mathrm{T}(\mathrm{K}^{-\mathrm{T}}\mathrm{K}^{-1})\mathbf{x}_2}}=\frac{\mathbf{x}_1^\mathrm{T}\omega\mathbf{x}_2}{\sqrt{\mathbf{x}_1^\mathrm{T}\omega\mathbf{x}_1}\sqrt{\mathbf{x}_2^\mathrm{T}\omega\mathbf{x}_2}}. \tag{7.8}$$

这个表达式与图像的射影坐标系无关，也就是说，在图像的射影变换下保持不变. 为了看清这一点，考虑任何 2D 射影变换 H. 点 $\mathbf{x}_i$ 变为 $\mathrm{H}\mathbf{x}_i$ 而 $\omega$（作为任何图像二次曲线）变为 $\mathrm{H}^{-\mathrm{T}}\omega\mathrm{H}^{-1}$. 因此，(7.8)并没有变化，所以对图像中的任何射影坐标系都成立.

(3) 我们也可以定义绝对二次曲线的对偶图像（简称 DIAC）为

$$\omega^*=\omega^{-1}=\mathrm{KK}^\mathrm{T}. \tag{7.9}$$

这是一个对偶（线）二次曲线，而 $\omega$ 是点二次曲线（虽然它不含实点）. 二次曲线 $\omega^*$ 是 $\Omega_\infty^*$ 的图像.

(4) 结论 7.16 说明一旦 $\omega$（或等价地 $\omega^*$）在图像中被辨认，那么 K 就被确定，因为一个对称矩阵通过 Cholesky 分解（见节 A3.2.1（$p407$））可以唯一地分解成一个上三角矩阵及其转置的乘积，即 $\omega^*=\mathrm{KK}^\mathrm{T}$.

(5) 在第二章中我们看到一张平面 $\pi$ 与 $\pi_\infty$ 相交于一条直线，而这条直线与 $\Omega_\infty$ 相交于 $\pi$ 上两个虚圆点. 虚圆点的图像是 $\omega$ 上的点，它们是平面 $\pi$ 的消影直线与 $\omega$ 的交点.

下面的例子将说明：$\omega$ 的最后两个性质是标定算法的基础.

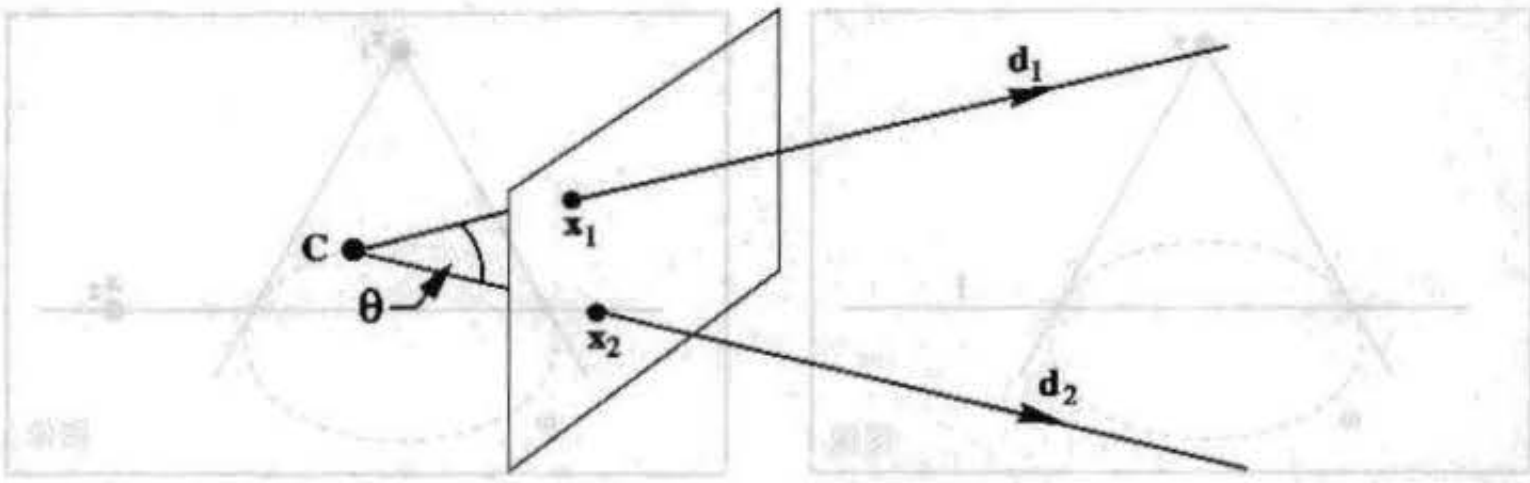



图 7.11　两射线之间的夹角 $\theta$.

**例 7.17　一个简单的标定装置**

三个正方形(它们所在平面是不平行的,但也不必正交)的图像提供计算 K 的足够约束. 考虑其中的一个正方形. 它的四个角点和它们的图像之间的对应确定正方形的平面 $\pi$ 和图像之间的一个单应 H. 把这个单应作用于 $\pi$ 上的虚圆点得到它们的图像是 $H(1,\pm i,0)^T$. 从而给出(目前还未知的)$\omega$ 的两个点. 把类似的过程应用于其他的正方形总共给出 $\omega$ 的六个点,由它们便计算出 $\omega$(因为确定一条二次曲线需要五个点). 概括起来算法的步骤如下:

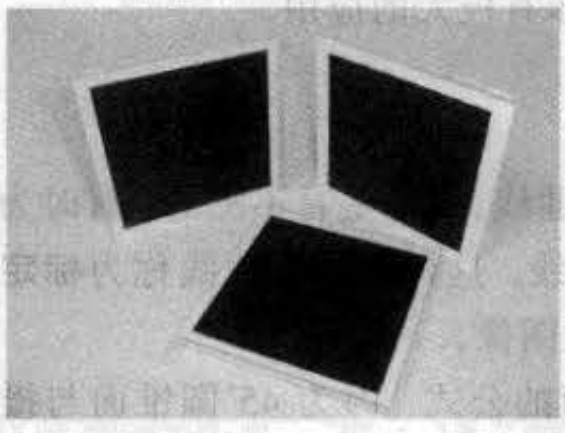
(1) 对每个正方形,计算把它的角点$(0,0)^T$、$(1,0)^T$、$(0,1)^T$、$(1,1)^T$映射到相应的图像点的单应 H.

(2) 计算该正方形所在平面虚圆点的图像,即$H(1,\pm i,0)^T$.

(3) 用(1.4 - $p$6)设计的矩阵,由这六个虚圆点的图像拟合出一条二次曲线 $\omega$.

(4) 用 Cholesky 分解由 $\omega=(KK^T)^{-1}$计算标定 K.

图 7.12 给出印有正方形的三张平面组成的标定物体及算得的矩阵 K. 就计算内标定的目的而言,正方形比标准标定物体(例如图 6.1($p$124))优越,它不需要测量 3D 坐标.



a

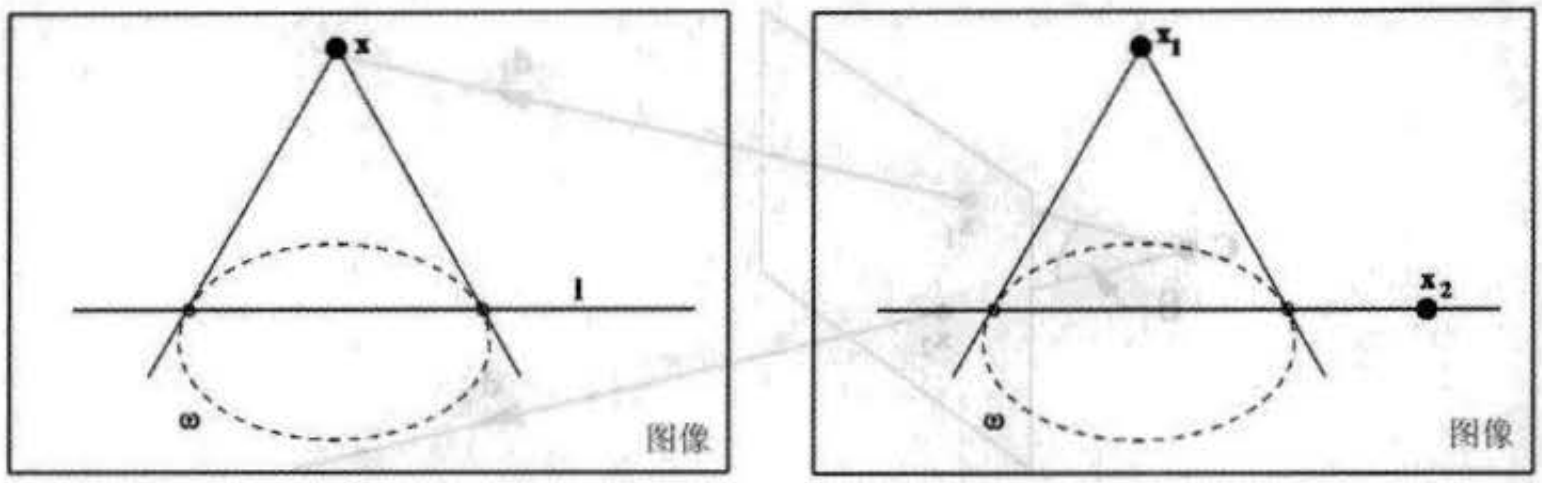
$$K=\begin{bmatrix}1108.3 & -9.8 & 525.8\\ 0 & 1097.8 & 395.9\\ 0 & 0 & 1\end{bmatrix}$$

b

图 7.12　**由度量平面进行标定**. (a)三个正方形提供一种简单的标定物体. 这些平面不必正交. (b)用例 7.17的算法计算标定矩阵. 图像的大小是 1024×768 像素.

**图像中的正交性**　由(7.8)可知,如果图像点关于 $\omega$ 共轭,即如果 $\mathbf{x}_1^T\omega\mathbf{x}_2=0$,那么这两个图像点对应于正交的方向,其中点 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 是反向投影为这两条正交射线的点. 类似地,关于 $\omega$ 是极点 - 极线的点和直线反向投影为(通过摄像机中心的)相互正交的射线和平面. 图 7.13给出这些关系的图示. 作为配极包含正交信息的例子,我们可以重新推证结论 7.15.

- **假定直线 l 反向投影到法线方向为 n 的平面 $\pi$. 那么法线的影像为点 Kn,而且直线 l 是**




a　　b

图 7.13　**由极点—极线关系表示正交性**. (a)如果点 $\mathbf{x}$ 和直线 $\mathbf{l}$ 关于 $\omega$ 为极点－极线关系，即 $\mathbf{l}=\omega\mathbf{x}$，那么 $\mathbf{x}$ 和 $\mathbf{l}$ 反向投影所得的射线和平面相互正交. 例如(参看节 7.6)一个平面法线方向的消影点是该平面消影线(又称极线)的极点. (b)如果图像点 $\mathbf{x}_1$，$\mathbf{x}_2$ 关于 $\omega$ 共轭，即 $\mathbf{x}_1^{\mathsf{T}}\omega\mathbf{x}_2=0$，那么它们反向投影为正交的射线.

**该点的极线，因此 $\mathbf{l}=\omega K\mathbf{n}=(KK^{\mathsf{T}})^{-1}K\mathbf{n}=K^{-\mathsf{T}}\mathbf{n}$. 简言之，$\pi$ 的法线是 $\mathbf{n}=K^{\mathsf{T}}\mathbf{l}$.**

由图像点测量两射线夹角的射影表示(7.8)就是由这两条射线与 $\pi_\infty$ 的交来测量它们的夹角的射影表示(2.23－$p$48)的图像，即

$$\cos\theta=\frac{\mathbf{d}_1^{\mathsf{T}}\Omega_\infty\mathbf{d}_2}{\sqrt{\mathbf{d}_1^{\mathsf{T}}\Omega_\infty\mathbf{d}_1}\sqrt{\mathbf{d}_2^{\mathsf{T}}\Omega_\infty\mathbf{d}_2}}.$$

在映射(7.7)$H=KR$(它是世界坐标系的平面 $\pi_\infty$ 和图像平面之间的单应)作用下，$\Omega_\infty\mapsto H^{\mathsf{T}}\omega H=(KR)^{\mathsf{T}}\omega(KR)$ 和 $\mathbf{d}_i=H^{-1}\mathbf{x}_i=(KR)^{-1}\mathbf{x}_i$；把这些关系代入(2.23－$p$48)就得到(7.8). 类似地，图像中体现正交性的共轭和极点—极线关系就是 $\pi_\infty$ 上这些关系的直接影像，可以通过比较图 2.8($p$48)和图7.13来理解.

在实际中，这些正交结果在消影点和消影线的情形有极大的应用.

### 7.5.2　标定二次曲线

绝对二次曲线的图像(IAC)是图像中一条虚二次曲线，因而是看不见的. 有时为了可视化的目的，考虑与摄像机标定紧密相关的另外一种曲线. 这样的二次曲线称为**标定二次曲线**，它是一个顶角为 45°而轴为摄像机主轴的圆锥面的图像，.

我们希望用摄像机的标定矩阵来计算这个圆锥面的公式. 因为 45°圆锥面与摄像机一起移动，它的图像显然与摄像机的定向和位置无关. 因此我们可以假定摄像机定位在原点且定向为 Z 轴的正方向. 因而，可令摄像机矩阵为 $P=K[I|\mathbf{0}]$. 那么，45°圆锥面的任何一点满足 $X^2+Y^2=Z^2$. 该圆锥面的点映射为二次曲线

$$C=K^{-\mathsf{T}}\begin{bmatrix}1&&\\&1&\\&&-1\end{bmatrix}K^{-1}\tag{7.10}$$

上的点，这不难由结论 7.5($p$137)得到验证. 这条二次曲线称为摄像机的标定二次曲线. 对于一个具有单位标定矩阵 $K=I$ 的摄像机，标定二次曲线是中心在原点(主轴点的图像)的单位圆. 二次曲线(7.10)就是根据结论 1.13($p$11)的二次曲线变换规则($C\mapsto H^{-\mathsf{T}}CH^{-1}$)对

这个单位圆进行一个仿射变换得到. 因此,摄像机(标定矩阵为 K)的标定二次曲线是中心在原点的单位圆经过矩阵 K 的仿射变换后得到的曲线.

标定参数可从标定二次曲线轻易求得. 如图 7.14 所示,主点是二次曲线的中心,而且能轻易地辨认比例因子和扭曲因子. 在零扭曲的情形,标定二次曲线的主轴与图像坐标轴方向一致. 一个实际图像的例子在图 7.21 中给出.

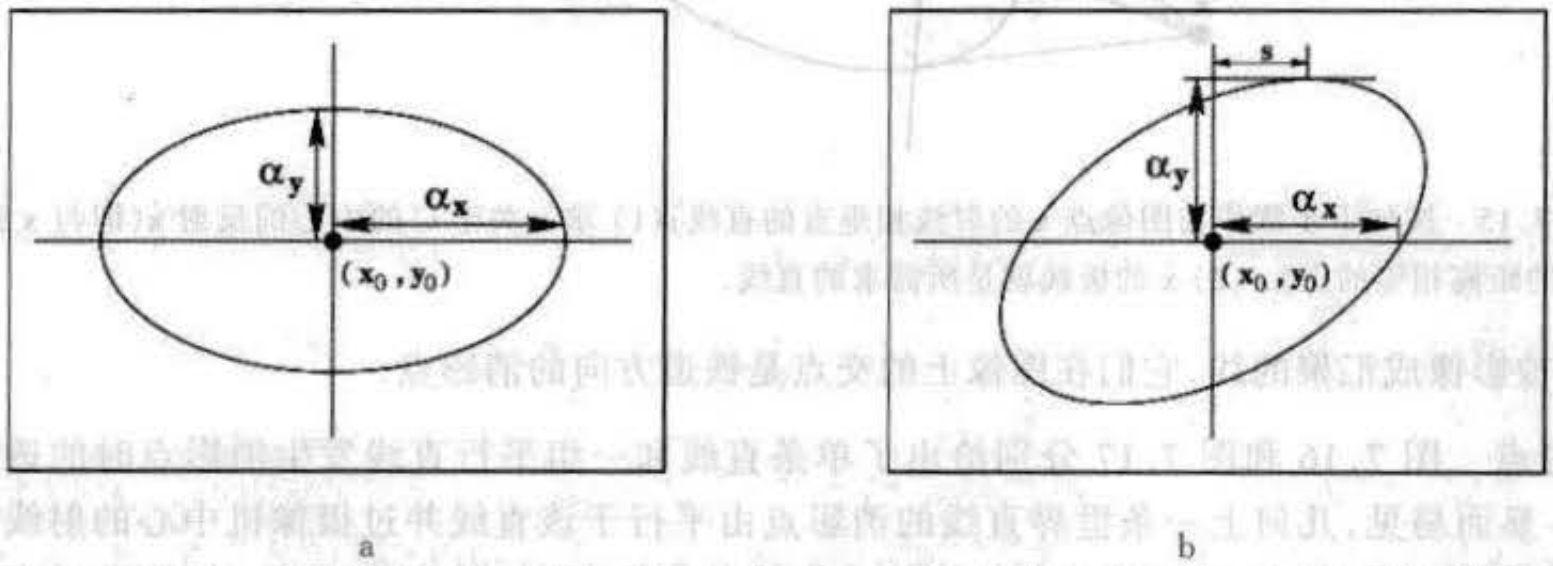



图 7.14 **由标定二次曲线读出摄像机的内参数** K. (a)扭曲 $s$ 为零. (b)扭曲 $s$ 不为零. K 的扭曲参数(参见(5.10 - $p$104))由二次曲线的最高点的 $x$ 坐标给出.

**例 7.18** 假定一个摄像机的标定矩阵 $K = \mathrm{diag}(f, f, 1)$,它的焦距为 $f$ 个像素、零扭曲、正方形像素并且图像原点与主点重合. 那么由(7.10)得到标定二次曲线是 $C = \mathrm{diag}(1, 1, -f^2)$,它是半径为 $f$,中心为主点的圆.

**正交性和标定二次曲线**

(7.8)给出对应于两个图像点的射线间夹角的公式. 特别是当 $\mathbf{x}'^{\mathrm{T}}\omega\mathbf{x} = 0$ 时,对应于图像点 $\mathbf{x}$ 和 $\mathbf{x}'$ 的两条射线正交. 如图 7.13 所示,它可以解释成点 $\mathbf{x}'$ 在直线 $\omega\mathbf{x}$ 上,而该直线是 $\mathbf{x}$ 关于 IAC 的极线.

我们希望用标定二次曲线来作类似的分析. 记 $C = K^{-\mathrm{T}}DK^{-1}$, 其中 $D = \mathrm{diag}(1, 1, -1)$, 我们有

$$C = (K^{-\mathrm{T}}K^{-1})(KDK^{-1}) = \omega S$$

其中 $S = KDK^{-1}$. 然而对任何点 $\mathbf{x}$,乘积 $S\mathbf{x}$ 表示点 $\mathbf{x}$ 关于二次曲线 C 的中心(即摄像机主点)的反射. 用 $\dot{\mathbf{x}}$ 表示反射点,我们得到

$$\mathbf{x}'^{\mathrm{T}}\omega\mathbf{x} = \mathbf{x}'^{\mathrm{T}}C\dot{\mathbf{x}}.$$

由此推出下面的几何结论:

**结论 7.19 如果图像平面上的一条直线对应于垂直于图像点 $\mathbf{x}$ 的射线的一个平面,那么,这条直线是 $\mathbf{x}$ 关于标定二次曲线的反射点 $\dot{\mathbf{x}}$ 的极线 $C\dot{\mathbf{x}}$.**

图 7.15 对它的作图方法给予说明.

## 7.6 消影点与消影线

透视投影的一个显著特征是延伸至无穷远的物体的图像可能出现在有限范围内. 例如,一条无穷直线被影像成终止在**消影点**的一条线段. 类似地,平行的世界直线,例如铁道

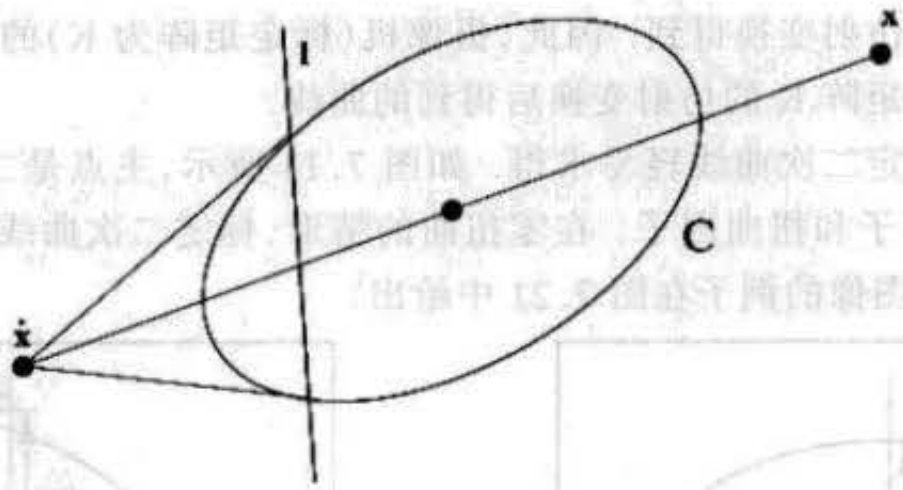
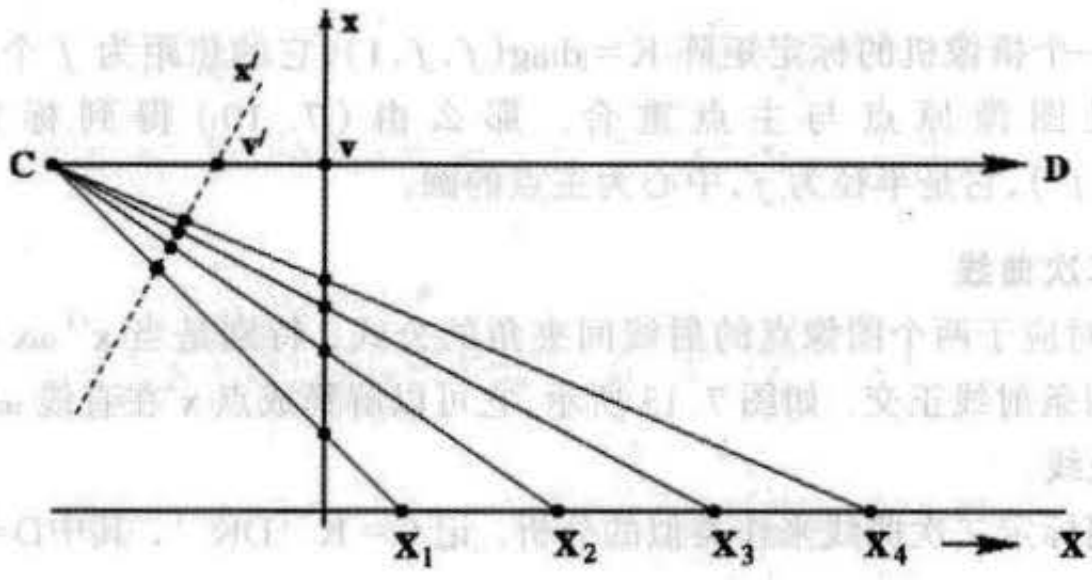


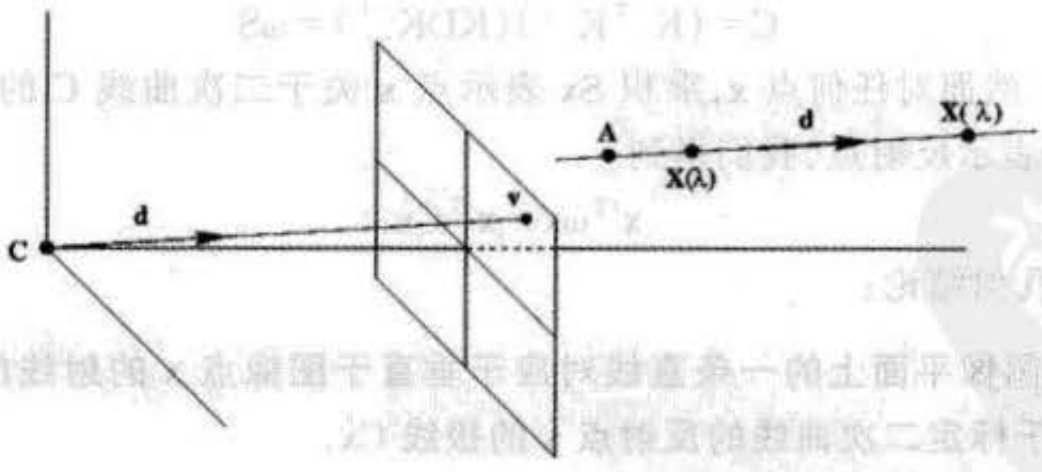

图 7.15　按如下步骤作与图像点 $\mathbf{x}$ 的射线相垂直的直线：(1) 求 $\mathbf{x}$ 关于 C 的中心的反射 $\dot{\mathbf{x}}$(即与 $\mathbf{x}$ 到中心的距离相等的点). (2) $\dot{\mathbf{x}}$ 的极线就是所需求的直线.

线，被影像成汇聚的线，它们在图像上的交点是铁道方向的消影点.

**消影点**　图 7.16 和图 7.17 分别给出了单条直线和一组平行直线发生消影点时的透视几何. 显而易见，几何上一条世界直线的消影点由平行于该直线并过摄像机中心的射线与图像平面的交点得到. 因此消影点仅依赖于直线的**方向**，而与其位置无关. 如果世界直线平行于图像平面，那么消影点在图像的无穷远处.




a




b

图 7.16　**消影点的形成**. (a)平面到线阵摄像机. 点 $\mathbf{X}_i$, $i=1,\cdots,4$ 在世界直线上是等间距的，但它们在图像直线上的间距是单调下降的. 当 $\mathbf{X}\to\infty$ 时，该世界点被影像到垂直图像线的 $\mathbf{x}=\mathbf{v}$ 和倾斜图像线的 $\mathbf{x}'=\mathbf{v}'$ 上. 因此，世界直线的消影点由过摄像机中心 C 并平行于它的射线与图像平面的交产生. (b)3 维空间到面阵摄像机. 方向为 $\mathbf{d}$ 的一条直线的消影点 $\mathbf{v}$ 由平行于 $\mathbf{d}$ 并过摄像机中心 C 的射线与图像平面的交产生. 该世界直线可以用 $\mathbf{X}(\lambda)=\mathbf{A}+\lambda\mathbf{D}$ 来参数化，其中 $\mathbf{A}$ 是直线上的点，而 $\mathbf{D}=(\mathbf{d}^{\mathrm{T}},0)^{\mathrm{T}}$.

代数上，消影点可以作为极限点而求得，方法如下：3 维空间中过点 $\mathbf{A}$ 且方向为 $\mathbf{D}=(\mathbf{d}^{\mathrm{T}},0)^{\mathrm{T}}$ 的一条直线上的点可以记为 $\mathbf{X}(\lambda)=\mathbf{A}+\lambda\mathbf{D}$，参看图 7.16b. 当参数 $\lambda$ 由 0 变到 $\infty$，点 $\mathbf{X}(\lambda)$ 由有限点 $\mathbf{A}$ 变到无穷远点 $\mathbf{D}$. 在射影摄像机 $\mathrm{P}=\mathrm{K}[\mathrm{I}|\mathbf{0}]$ 作用下，点 $\mathbf{X}(\lambda)$ 被影像为

$$\mathbf{x}(\lambda)=\mathrm{P}\mathbf{X}(\lambda)=\mathrm{P}\mathbf{A}+\lambda\mathrm{P}\mathbf{D}=\mathbf{a}+\lambda\mathrm{K}\mathbf{d}$$

其中 $\mathbf{a}$ 是 $\mathbf{A}$ 的图像. 从而该直线的消影点 $\mathbf{v}$ 通过取极限得到

$$\mathbf{v}=\lim_{\lambda\to\infty}\mathbf{x}(\lambda)=\lim_{\lambda\to\infty}(\mathbf{a}+\lambda\mathrm{K}\mathbf{d})=\mathrm{K}\mathbf{d}.$$

根据结论 7.14，$\mathbf{v}=\mathrm{K}\mathbf{d}$ 表示消影点，$\mathbf{v}$ 反向投影成方向为 $\mathbf{d}$ 的射线. 注意消影点 $\mathbf{v}$ 仅依赖于该直线的方向 $\mathbf{d}$，而与 $\mathbf{A}$ 的位置无关.

用射影几何的语言，这个结果可以直接得到：在 3 维射影空间中，无穷远平面 $\boldsymbol{\pi}_\infty$ 是直线方向的平面，而且所有具有相同方向的直线交 $\boldsymbol{\pi}_\infty$ 于同一点（参见第二章）. 消影点就是这个交点的图像. 因此如果一条直线的方向是 $\mathbf{d}$，则它与 $\boldsymbol{\pi}_\infty$ 交点是 $\mathbf{X}_\infty=(\mathbf{d}^{\mathrm{T}},0)^{\mathrm{T}}$. 而 $\mathbf{v}$ 是 $\mathbf{X}_\infty$ 的图像

$$\mathbf{v}=\mathrm{P}\mathbf{X}_\infty=\mathrm{K}[\mathrm{I}|\mathbf{0}]\begin{pmatrix}\mathbf{d}\\0\end{pmatrix}=\mathrm{K}\mathbf{d}.$$

概括起来

**结论 7.20　方向为 $\mathbf{d}$ 的 3 维空间直线的消影点是过摄像机中心且方向为 $\mathbf{d}$ 的射线与图像平面的交点 $\mathbf{v}$，即 $\mathbf{v}=\mathrm{K}\mathbf{d}$.**

注意与图像平面平行的直线被影像成平行直线，因为 $\mathbf{v}$ 在图像的无穷远处. 但是，反过来——图像中的平行直线是景物中平行直线的图像——不成立，因为主平面上相交的直线被影像成平行直线.

**例 7.21　由消影点求摄像机旋转**

消影点是无穷远点的图像，它提供定向（姿态）信息的方式与不动星提供的方式类似. 考虑由定向和**位置**不同的两个标定摄像机获取的一个景物的两幅图像. 无穷远点作为景物的一部分与摄像机无关. 它们的图像，即消影点，不受摄像机位置变化的影响，但要受摄像机旋转的影响. 假定两个摄像机的标定矩阵都是 K，但在摄取上述两幅图像时，它们之间作了旋转 R.

令一条景物直线在第一幅视图中的消影点是 $\mathbf{v}_i$，在第二幅视图中的是 $\mathbf{v}_i'$. 消影点 $\mathbf{v}_i$ 在第一个摄像机的欧氏坐标系中测量得到的方向是 $\mathbf{d}_i$，而对应的消影点 $\mathbf{v}_i'$ 在第二个摄像机的欧氏坐标系中测量得到的方向是 $\mathbf{d}_i'$. 这些方向都可以由消影点算得，例如，$\mathbf{d}_i=\mathrm{K}^{-1}\mathbf{v}_i/\|\mathrm{K}^{-1}\mathbf{v}_i\|$，其中加进归一化因子 $\|\mathrm{K}^{-1}\mathbf{v}_i\|$ 是为了保证 $\mathbf{d}_i$ 是单位矢量. 方向 $\mathbf{d}_i$ 和 $\mathbf{d}_i'$ 通过摄像机的旋转 $\mathbf{d}_i'=\mathrm{R}\mathbf{d}_i$ 相关联，它给出关于 R 的两个独立约束. 因此旋转矩阵 R 可以由这样的两组对应方向算出. △

**消影线**　3 维空间的平行平面与 $\boldsymbol{\pi}_\infty$ 交于一条公共的直线，而这条直线的图像就是平面的消影线. 如图 7.17 所示，几何上，消影线由平行于景物平面并过摄像机中心的一张平面与图像的交线得到. 显然，消影线仅与景物平面的**定向**有关，而与它的位置无关.

根据结论 7.15，过摄像机中心并且法线方向为 $\mathbf{n}$ 的平面与图像平面的交线是 $\mathbf{l}=\mathrm{K}^{-\mathrm{T}}\mathbf{n}$. 因此，$\mathbf{l}$ 是与 $\mathbf{n}$ 垂直的平面的消影线. 简言之

• **在摄像机欧氏坐标系下,垂直于方向 n 的平面集的消影线是 $\mathbf{l}=\mathbf{K}^{-T}\mathbf{n}$.**

因为直线与平行于它们的平面交于 $\pi_\infty$,因此不难看出与平面平行的直线的消影点必然在该平面的消影线上. 这样的一个例子在图 7.18 中给出.

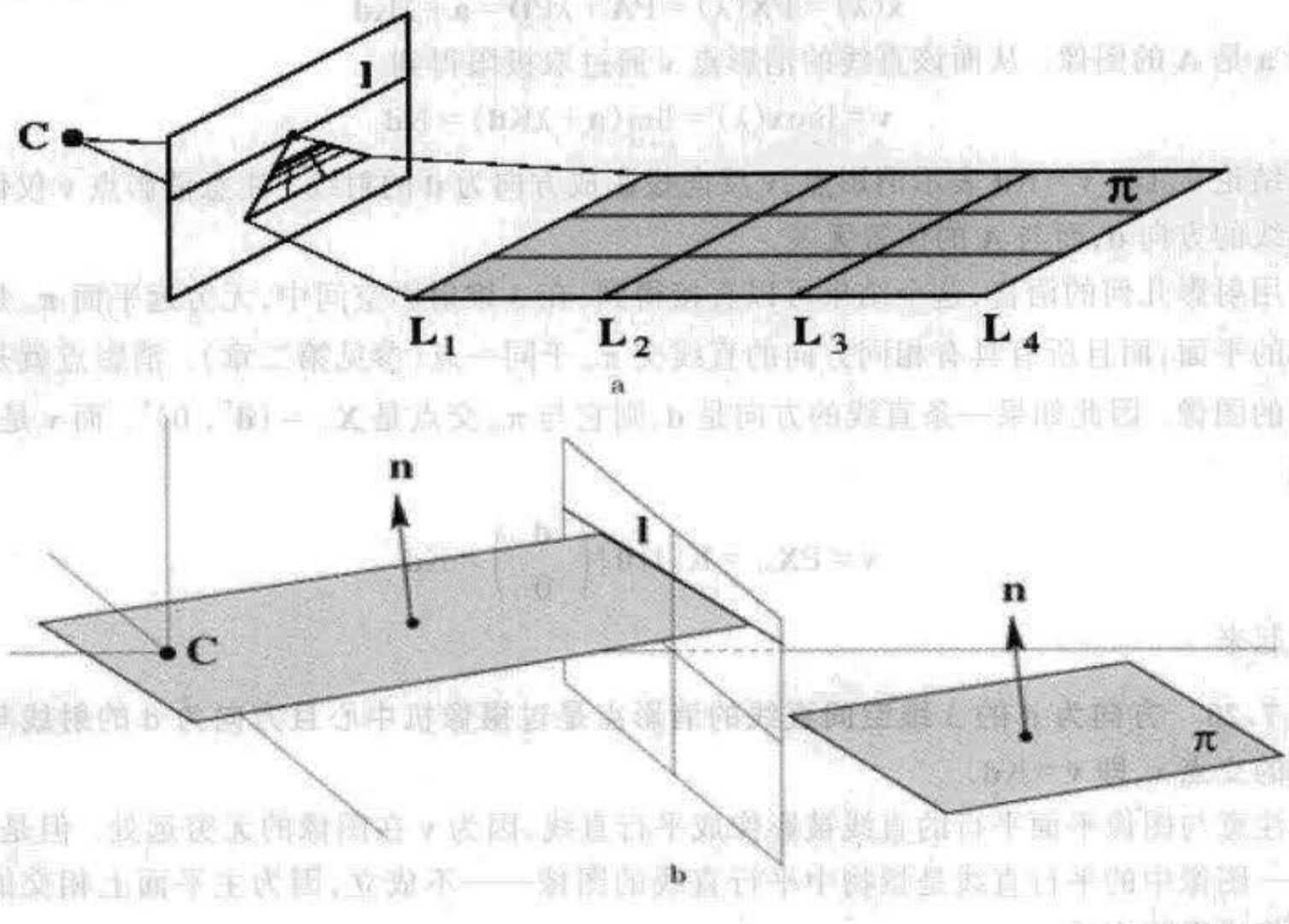



图 7.17 **消影线的形成**. (a)直线 $\mathbf{L}_i$, $i=1,\cdots,4$ 在景物平面 $\pi$ 上是等间距的,但它们的图像间距是单调下降的. 在平面上的点离摄像机无限远的极限情形,它们的图像在平面的消影线 l 上. 平面上平行直线集合的消影点在平面的消影线 l 上. 地平线是地平面的消影线. (b)平面 $\pi$ 的消影线由与 $\pi$ 平行并过摄像机中心 C 的平面与图像平面的交获得.

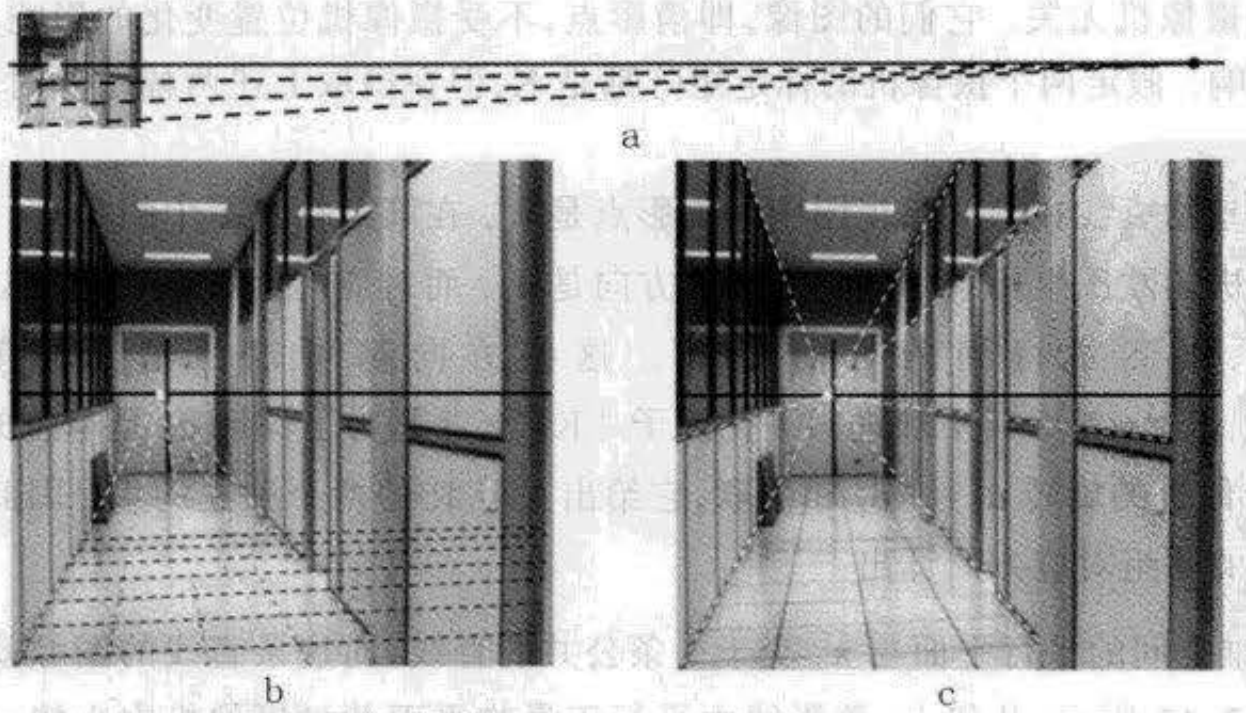



图 7.18 **消影点和消影线**. 走廊地平面的消影线可以由该平面上的两组平行线获得. (a)几乎平行于图像平面的直线的消影点离有限(实际)的图像很远. (b)注意:对应地板砖对边的等间距平行线的图像的间距单调下降. (c)平行于一张平面(这里是地平面)的直线的消影点在该平面的消影线上.

**两条景物直线之间的夹角** 我们已经知道一条景物直线的消影点反向投影成平行于该景物直线的射线. 因此,确定由图像点反向投影所得的射线之间的夹角的公式(7.8)使通过两条景物直线的消影点来测量它们方向间的夹角成为可能.

**结论 7.22 令 $\mathbf{v}_1$, $\mathbf{v}_2$ 是一幅图像中两条直线的消影点,而令 $\omega$ 为图像中绝对二次曲线的图像. 如果 $\theta$ 是两直线方向间的夹角,那么**

$$\cos\theta = \frac{\mathbf{v}_1^{\mathrm{T}}\omega\mathbf{v}_2}{\sqrt{\mathbf{v}_1^{\mathrm{T}}\omega\mathbf{v}_1}\sqrt{\mathbf{v}_2^{\mathrm{T}}\omega\mathbf{v}_2}}. \tag{7.12}$$

由消影线和 $\omega^*$ 计算两平面法线之间的角度,和一个直线方向与平面法线之间的角度也有类似的表示式. 这些将留作练习.

**正交性关系** 在实际中常出现的情况是产生消影点的直线和平面正交. 此时它们的消影点和消影线与 $\omega$ 有特别简单的关系,并且这些关系可以进一步用来(部分地)确定 $\omega$ 并进而确定摄像机的标定 K.

根据(7.12),两条垂直世界直线的消影点 $\mathbf{v}_1$, $\mathbf{v}_2$ 满足 $\mathbf{v}_1^{\mathrm{T}}\omega\mathbf{v}_2 = 0$. 这表示消影点关于 $\omega$ 共轭,如图 7.13 所说明的那样. 类似的结论对与平面正交的直线也成立. 这些图像关系概括如下:

(1) 具有垂直方向的直线的消影点满足

$$\mathbf{v}_1^{\mathrm{T}}\omega\mathbf{v}_2 = 0. \tag{7.13}$$

(2) 一张平面的法线方向的消影点 $\mathbf{v}$ 可以按公式 $\mathbf{v} = \omega^*\mathbf{l}$ 由平面的消影线 $\mathbf{l}$ 获得;反之有

$$\mathbf{l} = \omega\mathbf{v}. \tag{7.14}$$

(3) 两垂直平面的消影线满足 $\mathbf{l}_1^{\mathrm{T}}\omega^*\mathbf{l}_2 = 0$.

## 7.7 由消影点和消影线确定标定 K

两条直线之间的每个已知的角度给出关于 $\omega$ 的形如(7.12)的约束. 不幸的是,对任何的角度而言,$\mathbf{v}_1$ 和 $\mathbf{v}_2$ 给出的是关于 $\omega$ 元素的二次约束. 但是,如果直线相互垂直,(7.12)化简为(7.13),即 $\mathbf{v}_1^{\mathrm{T}}\omega\mathbf{v}_2 = 0$,从而对 $\omega$ 的约束是线性的. 给定足够数目的这种约束,$\omega$ 便可以算出,从而标定矩阵 K 也随之得到,因为根据结论(7.16),$\omega = (\mathrm{KK}^{\mathrm{T}})^{-1}$.

给定五对垂直线,我们可以用类似于由五点确定一条二次曲线的方式(1.4 - $p$6) 来解出 $\omega$ 的元素:二次曲线表示为一个齐次 6 维矢量,该矢量作为由线性约束组成的 $5\times6$ 矩阵的零矢量而得到.

$\omega$ 的线性约束同样可由对应于正交的直线与平面的消影点和消影线得到. 常见的一个例子是铅垂方向和水平平面. 根据上面的(7.14),该约束是 $\mathbf{l} = \omega\mathbf{v}$. 把它写成 $\mathbf{l}\times(\omega\mathbf{v}) = \mathbf{0}$ 以除去齐次比例因子从而得到关于 $\omega$ 元素的三组齐次线性方程. 它们给出 $\omega$ 的两个独立的约束.

最后,当标定矩阵 K 具有比(5.10 - $p$104)更特殊的形式时,需要由这样的景物约束来确定的 $\omega$ 的数目可以减少. 当已知的 K 为零扭曲或正方形长宽比($\alpha_x = \alpha_y$)时,我们可以巧妙利用这些条件来协助寻找 $\omega$. 具体地说,通过直接计算可以很快地验证:

**结论 7.23** **如果** $s = K_{12} = 0$ **并且** $[\omega_{ij}] = \omega = K^{-T}K^{-1}$，**那么** $\omega_{12} = \omega_{21} = 0$. **如果还满足** $\alpha_x = K_{11} = K_{22} = \alpha_y$，**那么** $\omega_{11} = \omega_{22}$.

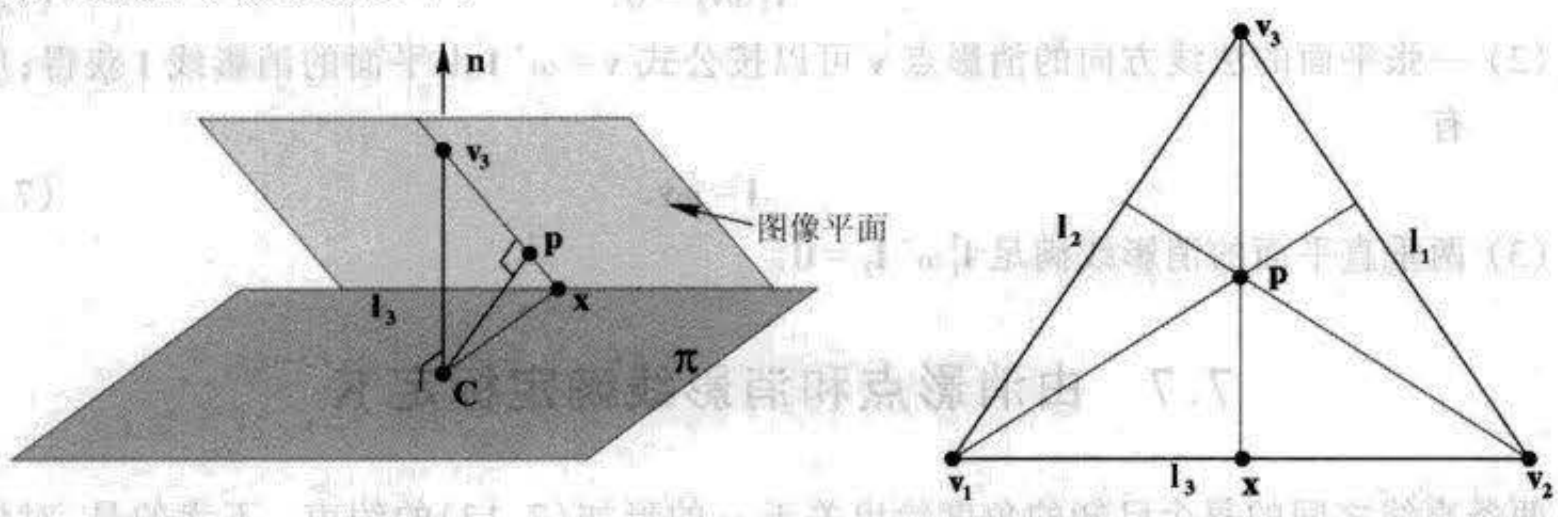
因此，在求绝对二次曲线时，如果我们确信摄像机存在零扭曲或正方形长宽比约束，则可以方便地利用这些约束. 但可以证明在 K 的元素与 $\omega^* = KK^T$ 的元素之间没有类似于结论 7.23 这样简单的关系存在.

零扭曲约束即 $s=0$，是图像的一个正交性约束：它表示 $x$ 和 $y$ 轴正交. 我们已经知道图像上的正交性由关于 $\omega$ 共轭来表征，那么零扭曲可以简单地用 $x$ 方向 $(1,0,0)^T$ 和 $y$ 方向 $(1,0,0)^T$ 关于 $\omega$ 共轭这个约束公式化，即 $(1,0,0)\omega(0,1,0)^T = 0$. 此式还蕴涵 $\omega_{12} = \omega_{21} = 0$，如结论 7.23 中所示. 因此，零扭曲可以直接作为补充(7.13)的另一个线性正交约束. 在此情形下，仅需要由直线方向再提供四个正交约束. 零扭曲约束经常用在自标定算法中，它是第 18 章讨论的主题.

**例 7.24 由三个正交消影点来标定**

假定摄像机是零扭曲并且像素是正方形(或其长宽比已知). 那么 $\omega$，进而 K，可以直接由三组正交方向的消影点来计算.

几何上，主点是以消影点为顶点的三角形的垂心. 图 7.19 显示主点在三角形的一个顶点到它对边的垂线上. 对其他两边也作类似的作图从而可以证明主点在垂心上. 这个结论的一种代数推导放在练习中.




a　　b

图 7.19 (a)过摄像机中心 C 的平面 $\pi$ 的法线 **n** 与主轴定义了一个交图像于直线 $\mathbf{l} = \langle \mathbf{v}_3, \mathbf{x} \rangle$ 的平面. 直线 $\mathbf{l}_3$ 是 $\pi$ 与图像平面的交，同时也是 $\pi$ 的消影线. 点 $\mathbf{v}_3$ 是法线与图像平面的交也是法线消影点. 显然主点在 **l** 上，而 **l** 和 $\mathbf{l}_3$ 在图像平面上相互垂直. (b)主点可以由作为三角形垂心的这样的三个约束来确定.

图 7.20 给出一个例子. 其中的图像是 768×1024 像素，而要计算的标定矩阵是

$$K = \begin{bmatrix} 1162.9 & 0 & 548.27 \\ 0 & 1162.9 & 404.25 \\ 0 & 0 & 1 \end{bmatrix}.$$

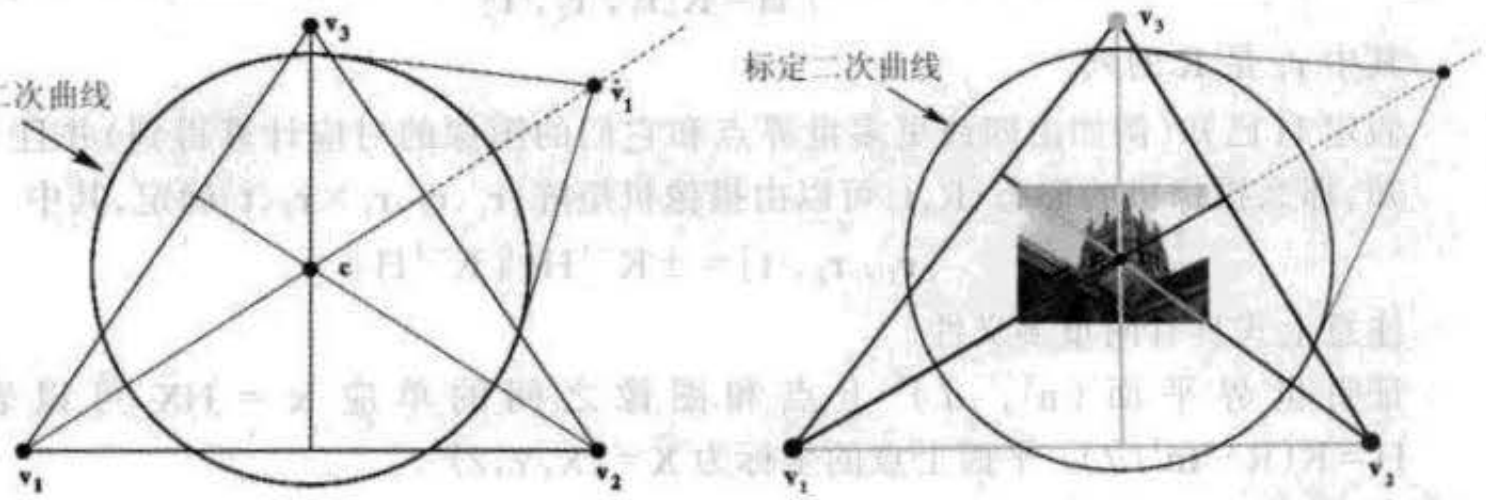
△

**例 7.25 由三个正交消影点给定的标定二次曲线**

图 7.20 中例图的标定二次曲线可以直接画出. 如果仍然假定零扭曲和正方形像素，那么标定二次曲线是一个圆. 若给定三组相互垂直直线的消影点，我们可以通过直接的几何作图(如图 7.21 所示)找到标定二次曲线.

a　　b　　c

图 7.20　在图像的扭曲是零而长宽比是 1 时，主点是三组正交方向的消影点的垂心．(a)原图像．(b)景物中的三组平行线，其中每两组的方向相互正交．(c)主点是以消影点为顶点的三角形的垂心．




a　　b

图 7.21　由正交消影点求得的标定二次曲线．(a)几何作图　(b)图 7.20 的图像的标定二次曲线．

(1) 首先构作顶点为三个消影点 $\mathbf{v}_1$，$\mathbf{v}_2$，$\mathbf{v}_3$ 的三角形．

(2) C 的中心是三角形的垂心．

(3) 关于中心反射其中的一个消影点(例如 $\mathbf{v}_1$)得到$\dot{\mathbf{v}}_1$．

(4) 根据条件：$\dot{\mathbf{v}}_1$ 的极线过 $\mathbf{v}_2$，$\mathbf{v}_3$ 来确定 C 的半径．

## 7.8　结束语

### 7.8.1 文献

Faugeras 与 Mourrain[Faugeras－95c]，和 Faugeras 与 Papadopoulo[Faugeras－97]用 Plücker 坐标推导了直线的投影．Koenderink [Koenderink－84，Koenderink－90]和 Giblin 与 Weiss[Giblin－87]给出轮廓生成元和视在轮廓线的许多性质，以及它们与曲面微分几何的关系．

[Kanatani－92]给出标定条件下的消影点和线的另一种处理方法以及有相同中心的摄像机获取的图像之间是以一个平面单应相关联的结论．Mundy 与 Zisserman [Mundy－92]用几何手段证明了这个结论；[Hartley－94d]基于摄像机投影矩阵给出此结论一个简单的代数推导．[Faugeras－92a] 引入了射影(简化)摄像机矩阵．绝对二次曲线的图像与摄像机的标定之间的联系由[Faugeras－92b]首先给出．

由已知其度量结构(例如正方形)的多个景物平面来计算 K 的结论在[Liebowitz－98]

中给出. 它的算法在[Liebowitz－99a、Sturm－99a、Zhang－99] 中给出. 当加了零扭曲约束后,用 $\omega$ 而不是 $\omega^*$ 的优越性首先由[Armstrong－96b]提出. 用相互正交方向的三个消影点进行内标定的方法由 Caprile 与 Torre[Caprile－90]给出,虽然这个结论在较早的摄影测量术的文献[Gracie－68]中曾被引用过.

### 7.8.2 注释和练习

(1) **世界平面的单应**. 遵循节 7.1.1 的推导,证明对一个已标定的摄像机(5.8－$p$103) $P=K[R|\mathbf{t}]$,世界平面 $Z=0$ 和图像之间的单应是

$$H=K[\mathbf{r}_1,\ \mathbf{r}_2,\ \mathbf{t}]$$

其中 $\mathbf{r}_i$ 是 R 的列.

假定 H 已知(例如由四或更多世界点和它们的图像的对应计算得到)并且 K 也已知,那么摄像机的姿态$\{R,\mathbf{t}\}$可以由摄像机矩阵$[\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_1\times\mathbf{r}_2,\mathbf{t}]$确定,其中

$$[\mathbf{r}_1,\ \mathbf{r}_2,\ \mathbf{t}]=\pm K^{-1}H/\|K^{-1}H\|$$

注意上式具有两重多义性.

证明世界平面$(\mathbf{n}^\mathsf{T},\ d)^\mathsf{T}$上点和图像之间的单应 $\mathbf{x}=H\tilde{\mathbf{X}}$ 可以表示成 $H=K(R-\mathbf{t}\mathbf{n}^\mathsf{T}/d)$. 平面上点的坐标为 $\tilde{\mathbf{X}}=(X,Y,Z)^\mathsf{T}$.

(2) **直线投影**

(a) 证明包含摄像机中心的任何直线在映射(7.1－$p$136)的零空间中,即它被投影到直线 $\mathbf{l}=\mathbf{0}$.

(b) 证明在 $\mathbb{P}^3$中直线 $\mathcal{L}=P^\mathsf{T}\mathbf{x}$ 是过图像点 $\mathbf{x}$ 和摄像机中心的射线. 提示:从结论 2.5($p$40)开始,并证明摄像机中心 $\mathbf{C}$ 在 $\mathcal{L}$ 上.

(c) $P$ 的列的几何解释是什么?

(3) 如果摄像机具有正方形像素,那么中心在主轴上的一个球的视在轮廓线是一个圆. 如果此球作平行于图像平面的平移,那么视在轮廓线由圆变形为主点在它的长轴上的一个椭圆.

(a) 如何把以上观察用作内参数标定的方法.

(b) 从几何上证明该椭圆的长短轴比与球到摄像机的距离无关.

如果该球作平行于主轴的平移,视在轮廓线可变形为双曲线,但该双曲线仅有一支被影像. 为什么?

(4) **二次曲面的轮廓生成元**. 二次曲面的轮廓生成元 $\boldsymbol{\Gamma}$ 由 Q 上所有这样的点 $\mathbf{X}$ 组成:过 $\mathbf{X}$ 的切平面包含摄像机中心 $\mathbf{C}$. Q 的过点 $\mathbf{X}$ 的切平面由 $\boldsymbol{\pi}=Q\mathbf{X}$ 给定,而 $\mathbf{C}$ 在 $\boldsymbol{\pi}$ 上的条件是 $\mathbf{C}^\mathsf{T}\boldsymbol{\pi}=\mathbf{C}^\mathsf{T}Q\mathbf{X}=0$. 于是 $\boldsymbol{\Gamma}$ 上的点 $\mathbf{X}$ 满足 $\mathbf{C}^\mathsf{T}Q\mathbf{X}=0$,从而在平面 $\boldsymbol{\pi}_\Gamma=Q\mathbf{C}$ 上,因为 $\boldsymbol{\pi}_\Gamma^\mathsf{T}\mathbf{X}=\mathbf{C}^\mathsf{T}Q\mathbf{X}=0$. 这表明二次曲面的轮廓生成元是一条平面曲线,而且,因为 $\boldsymbol{\pi}_\Gamma=Q\mathbf{C}$,所以 $\boldsymbol{\Gamma}$ 的平面是摄像机中心关于此二次曲面的极平面.

(5) **代数曲面的视在轮廓线**. 证明一个 $n$ 次的齐次代数曲面的视在轮廓线是 $n(n-1)$ 次的曲线. 例如,如果 $n=2$ 那么曲面是二次的并且视在轮廓线也是二次曲线. 提示:记曲面为 $F(X,Y,Z,W)=0$,那么切平面包含摄像机中心 $\mathbf{C}$ 的条件是

$$C_X\frac{\partial F}{\partial X}+C_Y\frac{\partial F}{\partial Y}+C_Z\frac{\partial F}{\partial Z}+C_W\frac{\partial F}{\partial W}=0$$

这是$(n-1)$次的曲面.

(6) $H=KRK^{-1}$**的旋转轴的消影点**. 一个共轭旋转的单应 $H=KRK^{-1}$有一个特征矢量 $K\mathbf{a}$,其中 $\mathbf{a}$ 是旋转轴的方向,因为 $HK\mathbf{a}=KR\mathbf{a}=K\mathbf{a}$. 最后的等式来自 $R\mathbf{a}=1\mathbf{a}$,即 $\mathbf{a}$ 是 R 的单位特征矢量. 进一步,我们得到(a)$K\mathbf{a}$ 在单应 H 下是不动点;以及(b)由结论 7.20 得到 $\mathbf{v}=K\mathbf{a}$ 是旋转轴的消影点.

(7) **合成旋转**. 如例 7.11($p$141)所示,假定用绕摄像机中心作纯旋转的两幅图像来估计一个单应. 那么被估计的单应将是一个共轭旋转,因而 $H=KR(\theta)K^{-1}$(虽然 K 和 R 未知). 但是,把 $H^2$ 应用于第一幅图像所产生的图像与摄像机绕同一轴旋转两倍角所获得的图像一样,因为 $H^2=KR^2K^{-1}=KR(2\theta)K^{-1}$. 更一般地,H 可以分解为 $H(\theta)=U\mathrm{diag}(1,e^{i\theta},e^{-i\theta})U^{-1}$,其中 U 的列是 H 的归一化特征矢量,$\theta$ 和 U 都可以由已估计的 H 算得. 如果在此分解中用 $\varnothing$ 代替 $\theta$,那么 $H(\varnothing)$对应于旋转 $\varnothing$ 的单应. 通过原始图像之间的内插(如果 $0<\varnothing<\theta$)或外插(如果 $\varnothing>\theta$)就可以产生合成图像.

(8) 证明具有消影线 $\mathbf{l}_1$ 和 $\mathbf{l}_2$ 的两平面之间的夹角是

$$\cos\theta=\frac{\mathbf{l}_1^{\mathrm{T}}\omega^*\mathbf{l}_2}{\sqrt{\mathbf{l}_1^{\mathrm{T}}\omega^*\mathbf{l}_1}\sqrt{\mathbf{l}_2^{\mathrm{T}}\omega^*\mathbf{l}_2}}.$$

(9) 当消影点由三个正交方向产生和图像是正方形像素时,用代数方法证明主点是以消影点为顶点的三角形的垂心. 提示:假定在三角形的一个顶点上的消影点是 $\mathbf{v}$ 而它对边的一条直线(通过另外两个消影点)是 $\mathbf{l}$. 那么由(7.14)得到 $\mathbf{v}=\omega^*\mathbf{l}$,因为 $\mathbf{v}$ 和 $\mathbf{l}$ 分别来自正交直线和平面. 证明主点在从 $\mathbf{v}$ 到 $\mathbf{l}$ 的直线上并且 $\mathbf{v}$ 与 $\mathbf{l}$ **在图像上**互相垂直. 因为这个结论对任何顶点成立,所以主点是三角形的垂心.

(10) 证明在零扭曲的情形,$\omega$ 是二次曲线

$$\left(\frac{x-x_0}{\alpha_x}\right)^2+\left(\frac{y-y_0}{\alpha_y}\right)^2+1=0.$$

它可以解释成一个中心在主点的椭圆,它的轴分别与 $x$ 和 $y$ 方向一致而且轴的长度分别是 $i\alpha_x$ 和 $i\alpha_y$.

(11) 证明三组相互正交的方向的消影点是关于 $\omega$ 的自极线三角形[Springer-64]的顶点.

(12) 例 7.17($p$147)说明一个度量平面(即能够算出虚圆点的平面)的图像提供在 IAC($\omega$)上的两个点(约束). 基于这个结论证明如果一个摄像机的长宽比和主点已知,其焦距和姿态可以由一张平面的单个图像加以计算. 关于这种计算的退化配置是什么?

(13) 证明如果下列的任何一种图形在该平面上,那么一个透视影像平面的虚圆点可以算出:(a)正方形方格栅;(b)两个矩形,一个矩形的边与另一个矩形的边不平行;(c)半径相等的两个圆;(d)半径不等的两个圆.

# 第 2 篇　两视图几何

## 本篇大纲

本篇涵盖了两幅透视视图的几何. 这些视图可以由一个双眼装置同时摄取,或者比如,由一个相对于景物运动的摄像机相继地摄取. 这两种情形在几何上是等价的,并且在这里将不加以区别. 每幅视图有相应的摄像机矩阵 P、P′,其中"′"标记第二幅视图的元素,而且 3 维空间的一点 $\mathbf{X}$ 在第一幅视图的影像是 $\mathbf{x}=\mathrm{P}\mathbf{X}$,在第二幅视图的是 $\mathbf{x}'=\mathrm{P}'\mathbf{X}$. 像点 $\mathbf{x}$ 与 $\mathbf{x}'$ **对应**,因为它们是同一个 3 维空间点的像. 我们打算着重讨论三个问题:

(1) **对应几何**　给定第一幅视图的像点 $\mathbf{x}$,它怎样约束第二幅视图的对应点 $\mathbf{x}'$ 的位置?

(2) **摄像机几何(运动)**　给定对应点集 $\{\mathbf{x}_i\leftrightarrow\mathbf{x}'_i\}$, $i=1,2,\ldots,n$,对应于这两幅视图的摄像机矩阵 P 和 P′是什么?

(3) **景物几何(结构)**　给定对应像点 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 和摄像机矩阵 P、P′,它们在 3 维空间(的原象)的位置 $\mathbf{X}$ 是什么?

第 8 章介绍两视图的**对极几何**并直接回答第一个问题:一幅视图中的点定义了另一幅视图中对应点所在的对极线. 对极几何只依赖于摄像机——其相对位置及其内参数,与景物结构则毫不相干. 对极几何由 3×3 矩阵 F 表示,称为**基本矩阵**. 本章对基本矩阵中的元素的功能进行介绍并在两个摄像机矩阵 P 和 P′已知时给出求基本矩阵的计算方法. 然后证明 P 和 P′在相差一个 3 维空间射影变换的意义下可以由 F 计算得到.

第 9 章介绍未标定多视图几何的一个最重要结果——仅由图像对应点,不需要其它的信息,就可以计算摄像机和景物结构的**重构**. 此结果同时回答了第二和第三个问题. 仅根据点对应得到的重构将相差一个 3 维的射影变换,但是这种不确定性可由摄像机或景物的良定的额外信息来解决. 用这种方法,可以由未标定的图像计算一个仿射或度量的重构. 接下来的两章将介绍有关细节和计算这个重构的数值算法.

第 10 章介绍由一组对应图像点 $\{\mathbf{x}_i\leftrightarrow\mathbf{x}'_i\}$ 计算 F 的方法,该方法甚至不要求知道这些点的结构(3D 原象 $\mathbf{X}_i$)和摄像机矩阵. 摄像机矩阵 P 和 P′在相差一个射影变换的意义下可以由所算得的 F 来确定.

第 11 章介绍在给定摄像机和对应像点后,计算景物结构的**三角形法**——由被相应的摄像机 P、P′反向投影的对应点 $\mathbf{x}$ 和 $\mathbf{x}'$ 的射线的交点来计算 3 维空间点 $\mathbf{X}$. 类似地,其它几何元素(如直线或二次曲线)的 3D 位置也可以由给定的它们的图像的对应来计算.

第 12 章涵盖平面的两视图几何. 它给出第一个问题的另一种回答:如果景物在一张平面上,则一旦这张平面的有关几何被计算出来之后,则一个点在一幅图像上的像 $\mathbf{x}$ 决定该点在另一幅图像上的像 $\mathbf{x}'$ 的位置. $\mathbf{x}$、$\mathbf{x}'$ 由一个平面射影变换相联系. 本章还介绍了视图之间的一个特别重要的射影变换—**无穷单应**,这是由无穷远平面所引起的变换.

第 13 章介绍两个仿射摄像机矩阵 P 和 P′的两视图几何. 这种特殊情形相对于一般射影的情形有许多简化,并且能为许多实际情形提供一个很好的近似.

# 第 8 章　对极几何和基本矩阵

对极几何是两幅视图之间内在的射影几何．它独立于景物结构，只依赖于摄像机的内参数和相对姿态．

基本矩阵 F 集中了这个内在几何的精华．它是一个秩 2 的 3×3 矩阵．如果一个 3 维空间点 **X** 在第一、第二幅视图中的像分别为 **x**、**x**′，则这两个图像点满足关系 $\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0$．

我们将首先介绍对极几何并推导基本矩阵．然后，针对摄像机在视图之间的一般运动和一些经常发生的特殊运动，阐述基本矩阵的性质．接下来证明在相差一个 3 维空间射影变换下由 F 可恢复摄像机矩阵．这个结论是第 9 章射影重构理论的基础．最后，在已知摄像机内参数的条件下，我们将证明在存在有限个解的意义下，由基本矩阵可算出摄取这些视图的摄像机的欧氏运动．

基本矩阵虽独立于景物结构．然而它可以仅依据景物的像点的对应计算得到，并不要求知道摄像机的内参数或相对姿态．这种计算将在第 10 章中介绍．

## 8.1　对极几何

本质上，两幅视图之间的对极几何是图像平面与以基线（基线是连接两摄像机中心的直线）为轴的平面束的交的几何．这种几何通常由双眼匹配中搜索对应点的问题推动，我们就将从这个目的开始．

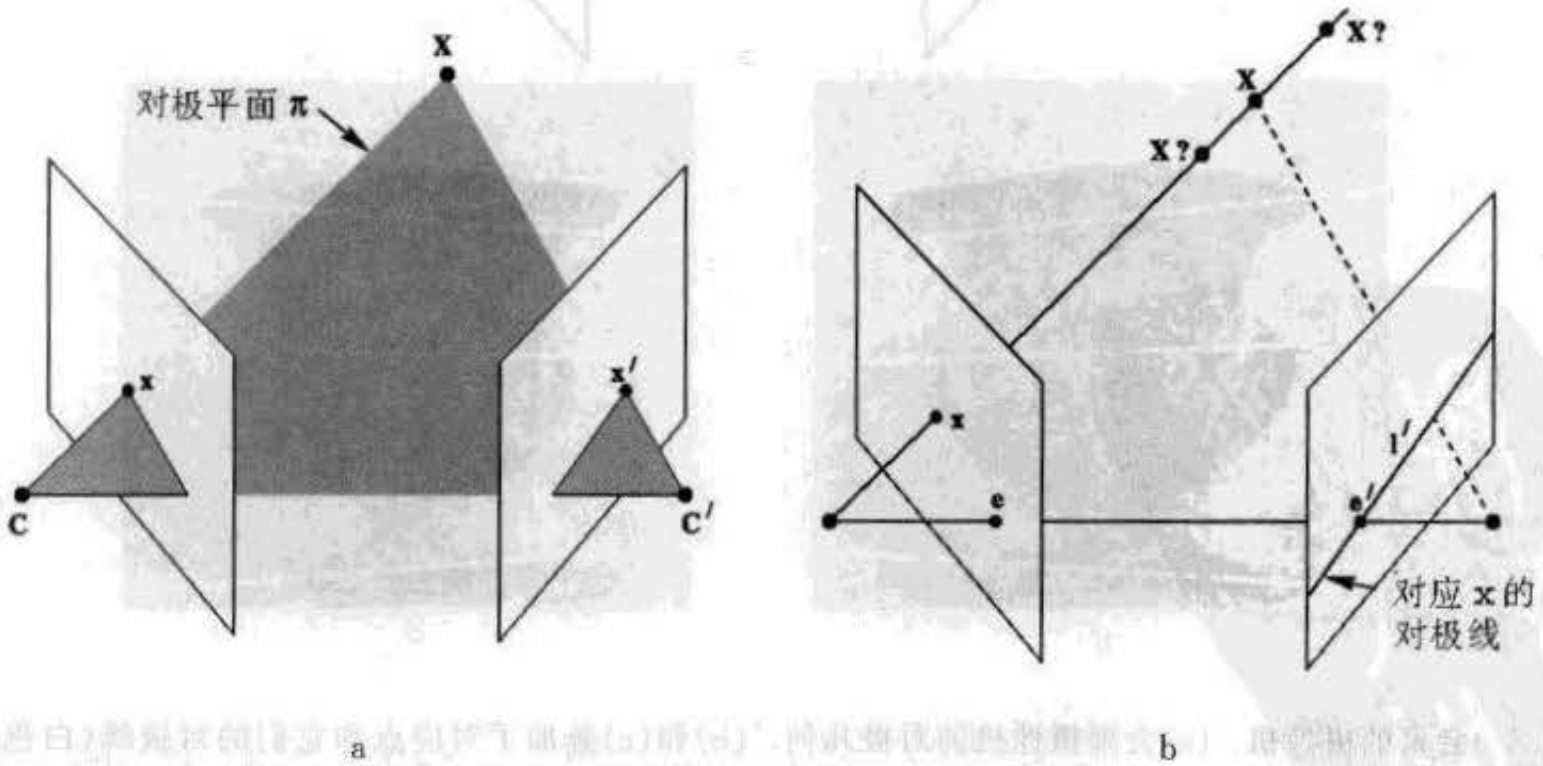



图 8.1　**对应点的几何**．(a)两摄像机由它们的中心 **C** 和 **C**′以及它们的图像平面表示．摄像机中心、3 维空间点 **X** 和它的图像 **x** 和 **x**′都处在一张公共平面 π 上．(b) 图像点 **x** 反向投影成 3 维空间的一条射线，它由第一个摄像机中心 **C** 和 **x** 确定．这条射线在第二幅视图中被影象成一条直线 **l**′．投影到 **x** 的 3 维空间点 **X** 必然在这条射线上，因此 **X** 在第二幅视图上的图像 **x**′必然在 **l**′上．

假定 3 维空间中的一个点 **X** 在两幅视图中成像，它在第一、二幅视图上的像分别为 **x**、**x**′. 那么在对应的图像点 **x** 和 **x**′之间存在什么关系呢？如图 8.1a 所示，像点 **x** 和 **x**′、空间点 **X** 和摄像机中心是共面的. 记这张平面为 $\pi$. 显然，从 **x** 和 **x**′反向投影的射线相交于 **X**，因而这两条射线共面并在 $\pi$ 上. 在搜索点对应中，这后一种性质有非常重要的意义.

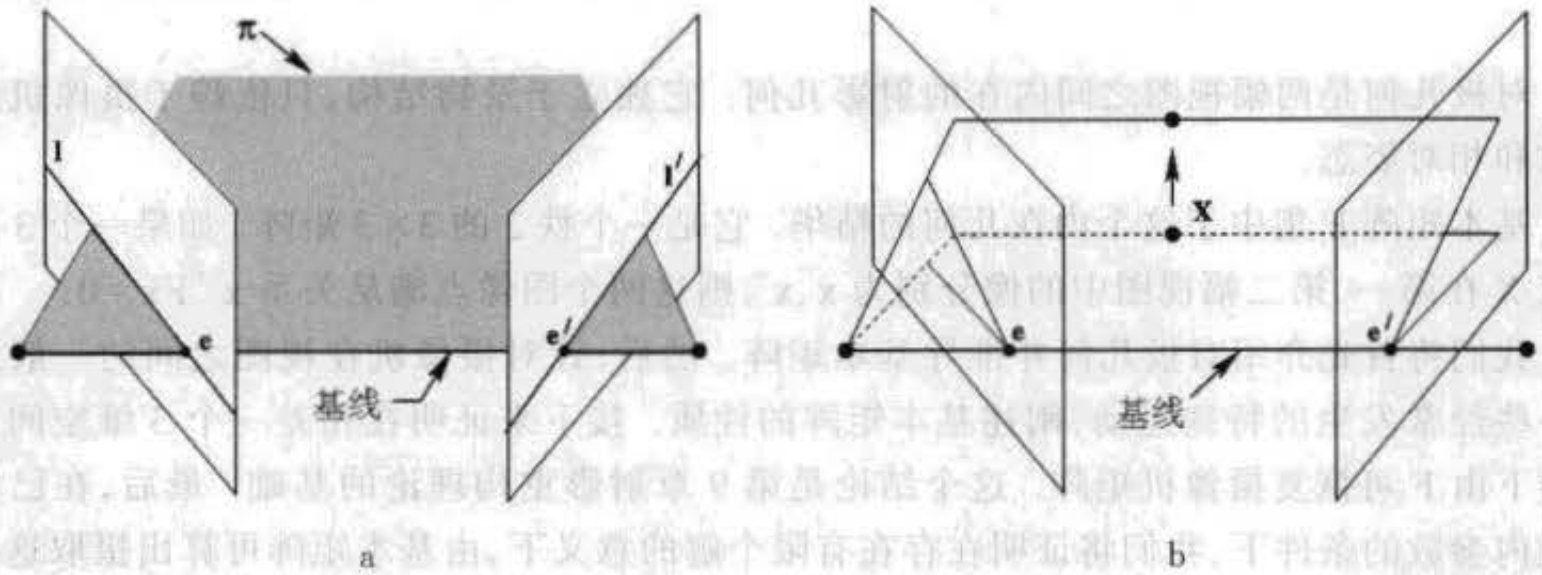



图 8.2 **对极几何**. (a)摄象机基线与每幅图像平面交于对极点 **e** 和 **e**′. 任何包含基线的平面 $\pi$ 是一张对极平面，并且与图像平面交于相应的对极线 **l** 和 **l**′. (b)当 3D 空间点 **X** 位置变化时，该对极平面绕基线"旋转". 这个平面簇称为对极平面束. 所有对极线相交于对极点.

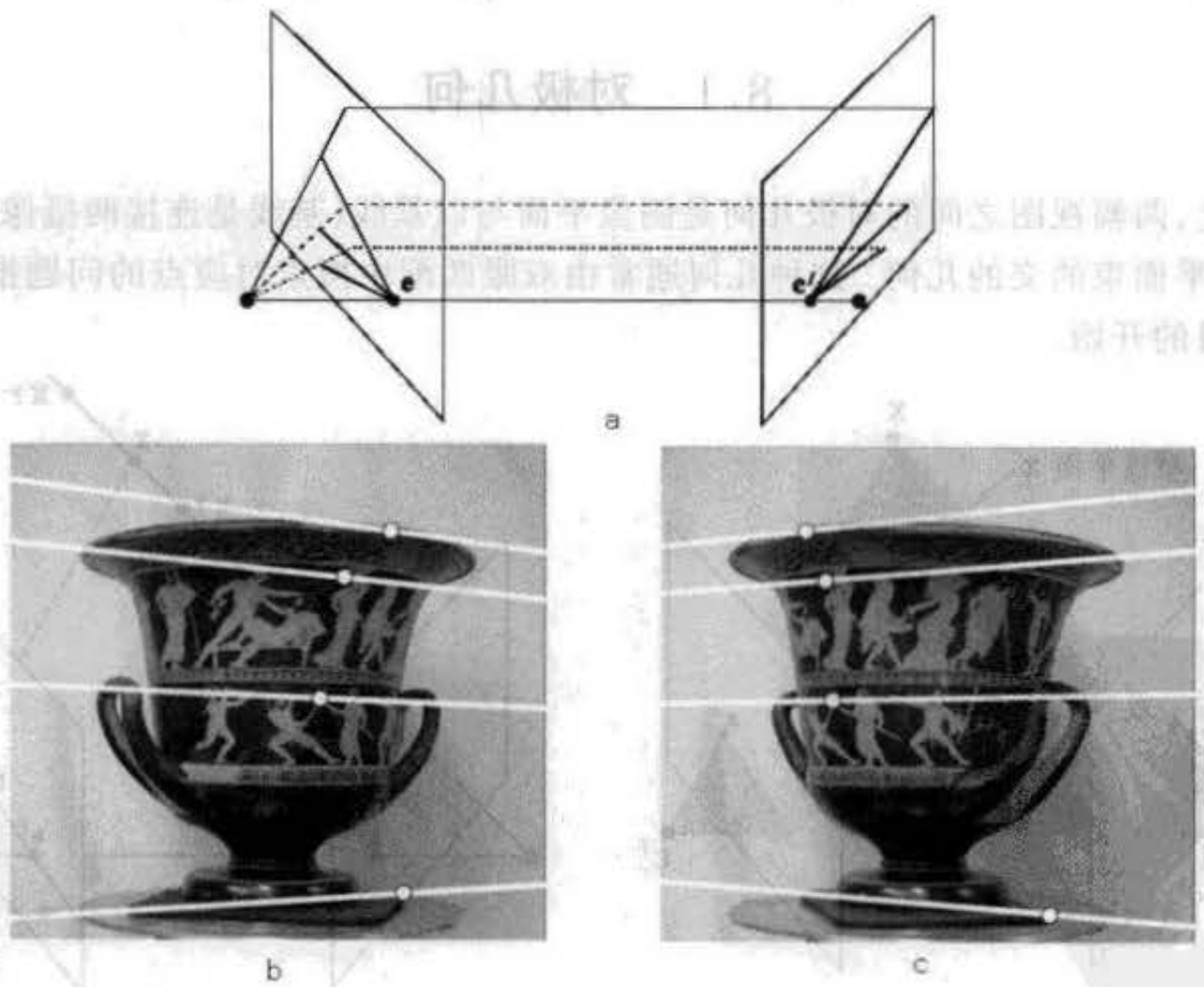



图 8.3 **会聚的摄像机**. (a)会聚摄像机的对极几何. (b)和(c)叠加了对应点和它们的对极线(白色)的一对图像. 视图之间的运动是一个平移加旋转. 在每幅图像中，另一个摄像机的方向可以由对极线束的交点来推算. 在此例中，两个对极点都在可视图像的外面.

现在假定我们只知道 **x**，我们会问其对应点 **x**′是如何被约束的. 此时平面 $\pi$ 被基线与由 **x** 定义的射线所确定. 从上面讨论知道对应于(未知)点 **x**′的射线在 $\pi$ 上，因此点 **x**′在平面 $\pi$ 与第二幅像平面的交线 **l**′上，直线 **l**′是从 **x** 反向投影的射线在第二幅视图上的像. 从双眼对

应算法的角度来看，它的好处在于无须在整幅像平面上搜索对应于 $\mathbf{x}$ 的点，只要限制在直线 $\mathbf{l}'$ 上即可.

对极几何中所涉及到的几何元素在图 8.2 中说明，其中有关术语是：

- **对极点**是连接两摄像机中心的直线(基线)与像平面的交点. 等价地，对极点是在一幅视图中另一个摄像机中心的像. 它也是基线(平移)方向的消影点.
- **对极平面**是一张包含基线的平面. 存在着对极平面的一个单参数簇(束).
- **对极线**是对极平面与图像平面的交线. 所有对极线相交于对极点. 一张对极平面与左或右像平面相交于对极线，并定义了对极线之间的对应.

在图 8.3 和图 8.4 中给出了对极几何的例子. 这些图像对的对极几何以及本章的所有例子都按节 10.6($p$200)所介绍的方法直接由图像计算得到.

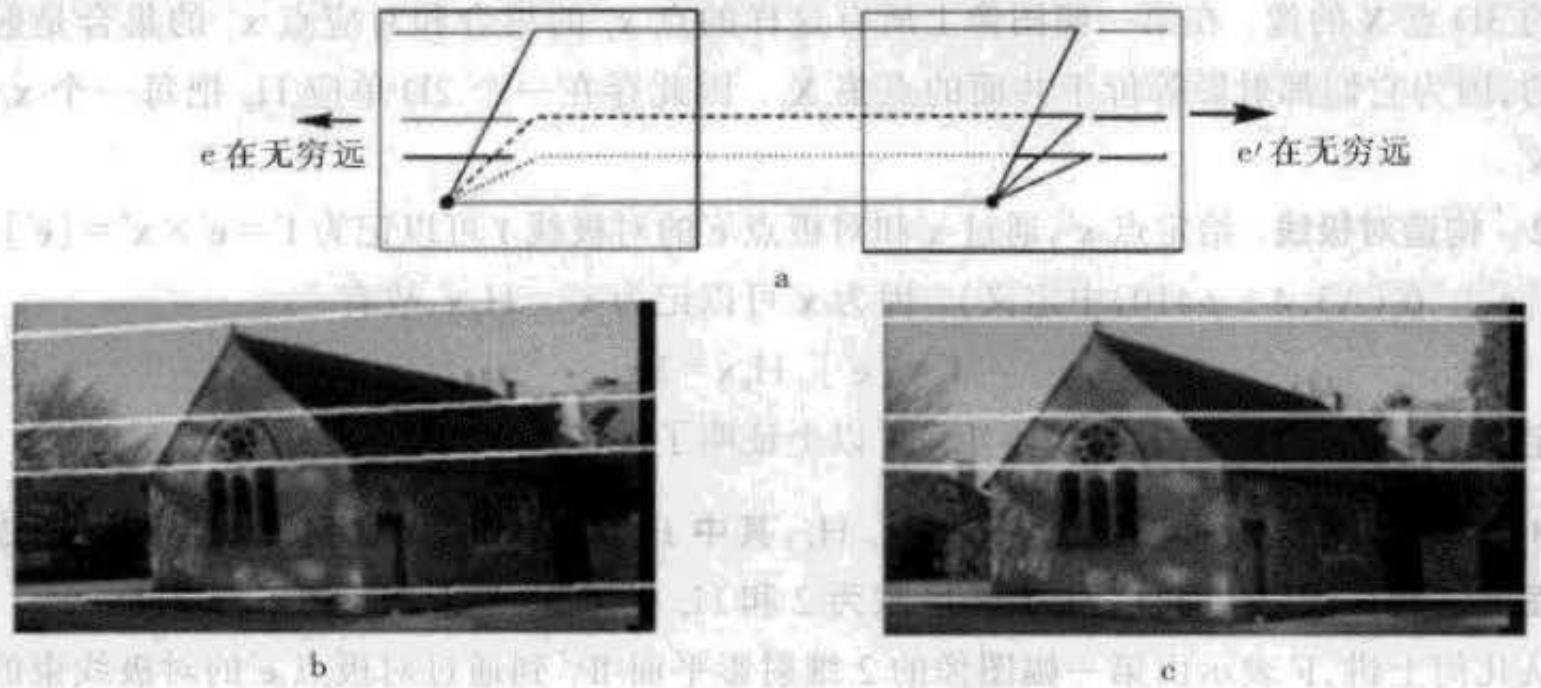



图 8.4 **平行于图像平面的运动**. 对于平行于图像平面作平移以及旋转轴垂直图像平面的特殊运动情形，基线与图像平面的交位于无穷远. 因此，对极点是无穷远点而对极线是平行线. (a)平行于图像平面作平移的对极几何. (b)和(c)一对图像：视图之间的运动(大致)是平行于 $x$ 轴的平移而没有旋转. 四条对应的对极线用白色线叠加在图像上. 注意对应点在对应的对极线上.

## 8.2 基本矩阵 F

基本矩阵是对极几何的代数表示. 下面，我们从一个点与它的对极线之间的映射来推导基本矩阵，然后详细说明这个矩阵的性质.

给定一对图像，我们已在图 8.1 中看到对于一幅图像上的每点 $\mathbf{x}$，在另一幅图像中存在一条对应的对极线 $\mathbf{l}'$. 在第二幅图像上，任何与该点 $\mathbf{x}$ 匹配的点 $\mathbf{x}'$ 必然在对极线 $\mathbf{l}'$ 上. 该对极线是过点 $\mathbf{x}$ 与第一个摄像机中心 $\mathbf{C}$ 的射线在第二幅图像上的投影. 因此，存在一个从一幅图像上的点到另一幅图像上与之对应的对极线的映射

$$\mathbf{x} \mapsto \mathbf{l}'$$

我们要探索的正是这个映射的本质. 结论是这个映射是一个(奇异)**对射**，即是由称为基本矩阵的矩阵 F 表示的从点到直线的射影映射.

### 8.2.1 几何推导

我们从基本矩阵的几何推导开始. 一幅图像上的一个点到另一幅图像上与之对应的对极线的映射可以分解为两步. 第一步,把点 $\mathbf{x}$ 映射到在另一幅图像上它的对极线 $\mathbf{l}'$ 上的某点 $\mathbf{x}'$. 这个点 $\mathbf{x}'$ 是点 $\mathbf{x}$ 的一个潜在的匹配点. 第二步,连结 $\mathbf{x}'$ 与对极点 $\mathbf{e}'$ 所得的直线就是对极线 $\mathbf{l}'$.

**步骤 1　点通过平面的转移**. 参考图 8.5,考虑空间中不通过任何两个摄像机中心的平面 $\pi$. 过第一个摄像机中心和 $\mathbf{x}$ 的射线与平面 $\pi$ 相交于一点 $\mathbf{X}$. 这个点 $\mathbf{X}$ 再投影到第二幅图像上的一点 $\mathbf{x}'$. 这个过程称为通过平面 $\pi$ 的转移. 如图 8.1b 所说明的,因为 $\mathbf{X}$ 在对应于 $\mathbf{x}$ 的射线上,它的投影点 $\mathbf{x}'$ 必然在对应于这条射线的图像即对极线 $\mathbf{l}'$ 上. 点 $\mathbf{x}$ 和 $\mathbf{x}'$ 都是在一张平面上的 3D 点 $\mathbf{X}$ 的像. 在第一幅图像上所有这样的点 $\mathbf{x}_i$ 的集合和对应点 $\mathbf{x}'_i$ 的集合是射影等价的,因为它们都射影等价于共面的点集 $\mathbf{X}_i$. 因此存在一个 2D 单应 $\mathrm{H}_\pi$ 把每一个 $\mathbf{x}_i$ 映射到 $\mathbf{x}'_i$.

**步骤 2　构造对极线**. 给定点 $\mathbf{x}'$,通过 $\mathbf{x}'$ 和对极点 $\mathbf{e}'$ 的对极线 $\mathbf{l}'$ 可以记为 $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = [\mathbf{e}']_\times \mathbf{x}'$ (记号 $[\mathbf{e}']_\times$ 在(A3.4 – $p$410)中定义). 因为 $\mathbf{x}'$ 可以记为 $\mathbf{x}' = \mathrm{H}_\pi \mathbf{x}$,故有

$$\mathbf{l}' = [\mathbf{e}']_\times \mathrm{H}_\pi \mathbf{x} = \mathrm{F}\mathbf{x}$$

其中定义 $\mathrm{F} = [\mathbf{e}']_\times \mathrm{H}_\pi$,它就是基本矩阵. 以上证明了

**结论 8.1　基本矩阵 F 可以记为 $\mathrm{F} = [\mathbf{e}']_\times \mathrm{H}_\pi$,其中 $\mathrm{H}_\pi$ 是从一幅图像到另一幅图像通过任意平面 $\pi$ 的转移映射. 而且,因为 $[\mathbf{e}']_\times$ 秩为 2 和 $\mathrm{H}_\pi$ 秩为 3,所以 F 是秩 2 的矩阵.**

从几何上讲,F 表示由第一幅图像的 2 维射影平面 $\mathbb{P}^2$ 到通过对极点 $\mathbf{e}'$ 的对极线束的映射. 因此,它表示从一个 2 维到 1 维的射影空间的映射,因此秩必须为 2.

**注**　在上面的几何推导中用到景物平面 $\pi$,但是该平面对 F 的存在**不是**必需的. 在这里,该平面仅仅用作定义从一幅图像到另一幅图像的点映射的手段. 基本矩阵与一幅图像到另一幅图像的点通过一张平面转移之间的联系将在第 12 章中作较深入地讨论.

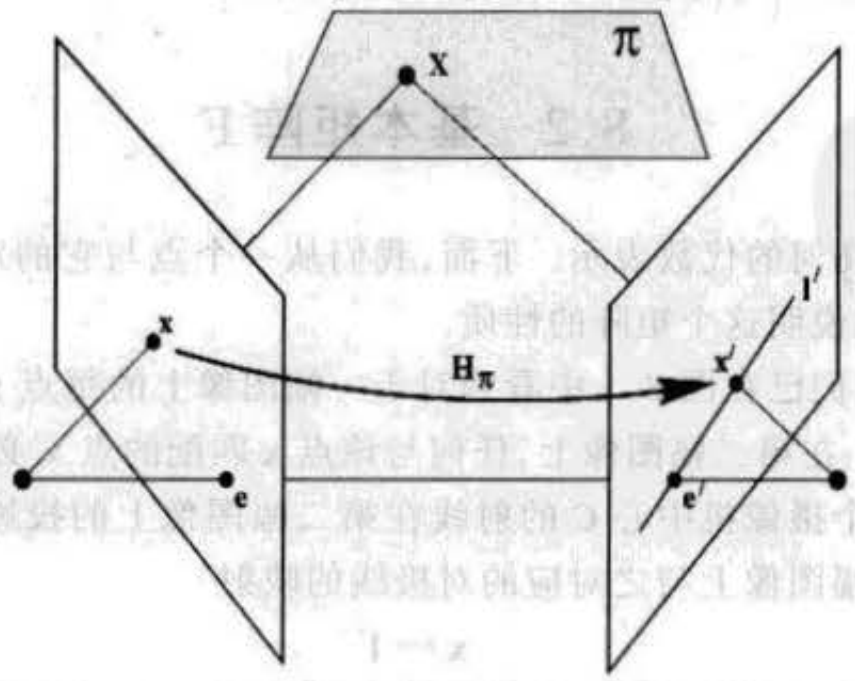



图 8.5　一幅图像的点 $\mathbf{x}$ 通过平面 $\pi$ 转移到第二幅图像的一个匹配点 $\mathbf{x}'$. 过 $\mathbf{x}'$ 的对极线由连接 $\mathbf{x}'$ 与 $\mathbf{e}'$ 得到. 我们可以用符号写成 $\mathbf{x}' = \mathrm{H}_\pi \mathbf{x}$ 和 $\mathbf{l}' = [\mathbf{e}']_\times \mathbf{x}' = [\mathbf{e}']_\times \mathrm{H}_\pi \mathbf{x} = \mathrm{F}\mathbf{x}$,其中 $\mathrm{F} = [\mathbf{e}']_\times \mathrm{H}_\pi$ 是基本矩阵.

### 8.2.2 代数推导

可以用代数方法由两个摄像机的射影矩阵 P、P′来推导基本矩阵的形式. 下面的公式来自 Xu 和 Zhang[Xu－96].

在 P 作用下,从 $\mathbf{x}$ 反向投影的射线通过求解方程 $P\mathbf{X}=\mathbf{x}$ 得到. (5.13－$p$107)给出它的单参数簇解的形式为

$$\mathbf{X}(\lambda) = P^{+}\mathbf{x} + \lambda\mathbf{C}$$

其中 $P^{+}$ 是 P 的伪逆,即 $PP^{+}=I$,而 $\mathbf{C}$ 为摄像机中心,即 P 的零矢量并定义为 $P\mathbf{C}=\mathbf{0}$. 这条射线由标量 $\lambda$ 参数化. 该射线上特殊的两点是 $P^{+}\mathbf{x}$(当 $\lambda=0$),和第一个摄像机的中心 $\mathbf{C}$(当 $\lambda=\infty$). 这两个点被第二个摄像机 P′分别影像到第二幅视图上的点 $P'P^{+}\mathbf{x}$ 和 $P'\mathbf{C}$. 对极线就是连接这两个投影点的直线,即 $\mathbf{l}'=(P'\mathbf{C})\times(P'P^{+}\mathbf{x})$. 点 $P'\mathbf{C}$ 是第二幅图像的对极点,即第一个摄像机中心的投影,且可以记为 $\mathbf{e}'$. 这样一来,$\mathbf{l}'=[\mathbf{e}']_{\times}(P'P^{+})\mathbf{x}=F\mathbf{x}$,其中 F 是基本矩阵

$$F = [\mathbf{e}']_{\times} P'P^{+}. \tag{8.1}$$

这个公式与前节推导基本矩阵时所得到的公式本质上相同,单应 $H_{\pi}$ 可用两个摄像机矩阵表示为:$H_{\pi}=P'P^{+}$. 注意上面的推导在两个摄像机中心相同时不能采用. 因为,如果 $\mathbf{C}$ 是 P 和 P′两个摄像机共同的中心,则 $P'\mathbf{C}=\mathbf{0}$. 从而(8.1)定义的 F 是零矩阵.

**例 8.2** 假设摄像机矩阵是一个已标定的双眼装置且世界原点在第一个摄像机上:

$$P = K[I \mid \mathbf{0}] \qquad P' = K'[R \mid \mathbf{t}].$$

则

$$P^{+} = \begin{bmatrix} K^{-1} \\ \mathbf{0}^{T} \end{bmatrix} \qquad \mathbf{C} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

且 $F = [P'\mathbf{C}]_{\times}P'P^{+} = [K'\mathbf{t}]_{\times}K'RK^{-1}$

$$= K'^{-T}[\mathbf{t}]_{\times}RK^{-1} = K'^{-T}R[R^{T}\mathbf{t}]_{\times}K^{-1} = K'^{-T}RK^{T}[KR^{T}\mathbf{t}]_{\times} \tag{8.2}$$

其中的各种变化形式由结论 A3.3($p$407)得到. 注意对极点(定义为另一个摄像机中心的图像)是

$$\mathbf{e} = P\begin{pmatrix} -R^{T}\mathbf{t} \\ 1 \end{pmatrix} = KR^{T}\mathbf{t} \qquad \mathbf{e}' = P'\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = K'\mathbf{t}. \tag{8.3}$$

则我们可以记(8.2)为

$$F = [\mathbf{e}']_{\times}K'RK^{-1} = K'^{-T}[\mathbf{t}]_{\times}RK^{-1} = K'^{-T}R[R^{T}\mathbf{t}]_{\times}K^{-1} = K'^{-T}RK^{T}[\mathbf{e}]_{\times}. \tag{8.4}$$

△

基本矩阵的表达式可以用许多方法导出,而且事实上在本书今后还要多次给予推导. 特别在(16.3－$p$292)中,将用每幅视图的摄像机矩阵的行组成的 4×4 行列式来表示 F.

### 8.2.3 对应条件

到目前为止我们考虑了由 F 定义的映射 $\mathbf{x}\mapsto\mathbf{l}'$. 我们现在可以叙述基本矩阵最基本的性质.

**结论 8.3 对两幅图像中任何一对对应点 $\mathbf{x}\leftrightarrow\mathbf{x}'$,基本矩阵都满足条件**

$$\mathbf{x}'^{\mathsf{T}}\mathrm{F}\mathbf{x}=0.$$

这的确是成立的，因为如果点 $\mathbf{x}$ 和 $\mathbf{x}'$ 对应，则 $\mathbf{x}'$ 在对应于点 $\mathbf{x}$ 的对极线 $\mathbf{l}'=\mathrm{F}\mathbf{x}$ 上. 也就是说，$0=\mathbf{x}'^{\mathsf{T}}\mathbf{l}'=\mathbf{x}'^{\mathsf{T}}\mathrm{F}\mathbf{x}$. 反之，如果图像点满足关系 $\mathbf{x}'^{\mathsf{T}}\mathrm{F}\mathbf{x}=0$，则由这两点定义的射线是共面的. 这是点对应的一个必要条件.

结论 8.3 给出的关系的重要性在于可以仅从对应图像点而不用参考摄像机矩阵，就能给出一种刻画基本矩阵的方式。从而可以根据图像点的对应来计算 F. 由(8.1)可知 F 可以从两个摄像机矩阵 P、P′计算出来，具体地说，F 可在相差一个整体尺度因子的意义下由摄像机唯一确定. 当然，我们会问由 $\mathbf{x}'^{\mathsf{T}}\mathrm{F}\mathbf{x}=0$ 计算 F 需要多少对应？而且在什么条件下该矩阵可由这些对应唯一地确定. 关于这方面的细节将在第 10 章介绍，其中将说明在一般情况下计算 F 至少需要 7 组对应.

### 8.2.4 基本矩阵的性质

**定义 8.4** 假设两幅图像由中心不重合的摄像机获得，则**基本矩阵** F 为对所有的对应点 $\mathbf{x}\leftrightarrow\mathbf{x}'$，都满足

$$\mathbf{x}'^{\mathsf{T}}\mathrm{F}\mathbf{x}=0 \tag{8.5}$$

的秩 2 的唯一的 $3\times3$ 齐次矩阵.

- F 是秩 2、自由度 7 的齐次矩阵.
- **点对应**：如果 $\mathbf{x}$ 和 $\mathbf{x}'$ 是对应的图像点，那么 $\mathbf{x}'^{\mathsf{T}}\mathrm{F}\mathbf{x}=\mathbf{0}$.
- **对极线**：
  - ◇ $\mathbf{l}'=\mathrm{F}\mathbf{x}$ 是对应于 $\mathbf{x}$ 的对极线.
  - ◇ $\mathbf{l}=\mathrm{F}^{\mathsf{T}}\mathbf{x}'$ 是对应于 $\mathbf{x}'$ 的对极线.
- **对极点**：
  - ◇ $\mathrm{F}\mathbf{e}=\mathbf{0}$.
  - ◇ $\mathrm{F}^{\mathsf{T}}\mathbf{e}'=\mathbf{0}$.
- **由摄像机矩阵 P、P′进行计算**：
  - ◇ 一般摄像机.
    $\mathrm{F}=[\mathbf{e}']_\times\mathrm{P}'\mathrm{P}^+$，其中 $\mathrm{P}^+$ 是 P 的伪逆，$\mathbf{e}'=\mathrm{P}'\mathbf{C}$ 且 $\mathrm{P}\mathbf{C}=\mathbf{0}$.
  - ◇ 规范摄像机，$\mathrm{P}=[\mathrm{I}|\mathbf{0}]$，$\mathrm{P}'=[\mathrm{M}|\mathbf{m}]$，
    $\mathrm{F}=[\mathbf{e}']_\times\mathrm{M}=\mathrm{M}^{-\mathsf{T}}[\mathbf{e}]_\times$，其中 $\mathbf{e}'=\mathbf{m}$ 且 $\mathbf{e}=\mathrm{M}^{-1}\mathbf{m}$.
  - ◇ 非无穷远摄像机 $\mathrm{P}=\mathrm{K}[\mathrm{I}|\mathbf{0}]$，$\mathrm{P}'=\mathrm{K}'[\mathrm{R}|\mathbf{t}]$，
    $\mathrm{F}=\mathrm{K}'^{-\mathsf{T}}[\mathbf{t}]_\times\mathrm{R}\mathrm{K}^{-1}=[\mathrm{K}'\mathbf{t}]_\times\mathrm{K}'\mathrm{R}\mathrm{K}^{-1}=\mathrm{K}'^{-\mathsf{T}}\mathrm{R}\mathrm{K}^{\mathsf{T}}[\mathrm{K}\mathrm{R}^{\mathsf{T}}\mathbf{t}]_\times$.

表 8.1 基本矩阵性质摘要.

我们现在简要地列举基本矩阵的一些性质. 其中最重要的性质在表 8.1 中作了摘要.

(1) **转置** 如果 F 是摄像机对(P,P′)的基本矩阵，则 $\mathrm{F}^{\mathsf{T}}$ 是其反序对(P′,P)的基本矩阵.

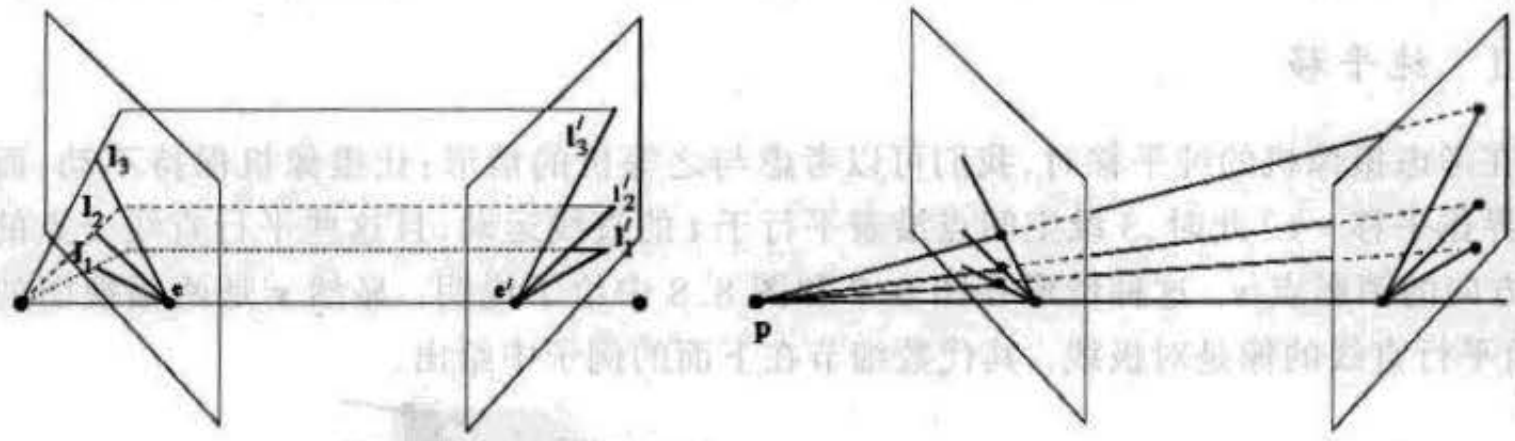
(2) **对极线** 与第一幅图像上任意一点 $\mathbf{x}$ 对应的对极线是 $\mathbf{l}'=\mathrm{F}\mathbf{x}$. 类似地，$\mathbf{l}=\mathrm{F}^{\mathsf{T}}\mathbf{x}'$ 表示对应于第二幅图像上点 $\mathbf{x}'$ 的对极线.

(3) **对极点** 任何(不同于 $\mathbf{e}$ 的)点 $\mathbf{x}$ 的对极线 $\mathbf{l}'=\mathrm{F}\mathbf{x}$ 包含对极点 $\mathbf{e}'$. 因此对所有 $\mathbf{x}$，$\mathbf{e}'$ 都满足 $\mathbf{e}'^{\mathsf{T}}(\mathrm{F}\mathbf{x})=(\mathbf{e}'^{\mathsf{T}}\mathrm{F})\mathbf{x}=0$. 从而推出 $\mathbf{e}'^{\mathsf{T}}\mathrm{F}=\mathbf{0}$，即 $\mathbf{e}'$ 是 F 的左零矢量. 同理可得

$\mathbf{Fe}=\mathbf{0}$，即 $\mathbf{e}$ 是 F 的右零矢量.

(4) F 有 7 个自由度：一个 3×3 齐次矩阵有 8 个独立的比率（矩阵有 9 个元素，但公共的因子不影响自由度）；此外，F 还满足约束 $\det F=0$，因此再减去一个自由度.

(5) F 是一种**对射**，一种把点映射到直线的射影映射（见定义 1.28（$p$29））. 在此情形下，第一幅图像上的点 $\mathbf{x}$ 确定了第二幅图像上的一条直线 $\mathbf{l}'=F\mathbf{x}$，即 $\mathbf{x}$ 的对极线. 如果 $\mathbf{l}$ 和 $\mathbf{l}'$ 是对应的对极线（见图 8.6a），则在 $\mathbf{l}$ 上的任何点都映射到同一直线 $\mathbf{l}'$. 这意味着不存在逆映射，从而 F 不是满秩的. 由于这个原因，F 并不是真对射（真对射应该是可逆的）.




a　　b

图 8.6　**对极线的单应**. (a)在每幅图像中存在以对极点为中心的对极线束. 以基线为轴的平面束确定对极线 $\mathbf{l}_i\leftrightarrow\mathbf{l}'_i$ 之间的对应. (b)可由基线上任何点 $\mathbf{p}$ 为中心的一个透视变换来确定对应线之间的对应关系. 由此可见对极线束之间的这种对应是一种 1D 单应.

### 8.2.5　对极线单应

每幅图像的对极线集合组成过对极点的直线束. 这样的直线束可以视为一个 1 维射影空间. 从图 8.6b 可清楚地看出：对应的对极线是透视相关的，即在第一幅视图中以 $\mathbf{e}$ 为中心的对极线束与在第二幅视图中以 $\mathbf{e}'$ 为中心的对极线束之间存在着一种单应. 这种 1 维射影空间之间的单应有 3 个自由度.

因此基本矩阵的自由度可计算如下：$\mathbf{e}$ 有 2 个，$\mathbf{e}'$ 有 2 个，把过 $\mathbf{e}$ 的直线映射到过 $\mathbf{e}'$ 直线的对极线单应有 3 个（共 7 个）. 这种单应的几何表示在节 8.4 中给出. 现在来推导这种映射的一个显式公式.

**结论 8.5　假设 $\mathbf{l}$ 和 $\mathbf{l}'$ 是对应的对极线，且 $\mathbf{k}$ 是不过对极点 $\mathbf{e}$ 的任何直线，则 $\mathbf{l}$ 和 $\mathbf{l}'$ 间的关系是 $\mathbf{l}'=F[\mathbf{k}]_\times\mathbf{l}$. 对称地有 $\mathbf{l}=F^T[\mathbf{k}']_\times\mathbf{l}'$.**

**证明**　表达式 $[\mathbf{k}]_\times\mathbf{l}=\mathbf{k}\times\mathbf{l}$ 给出两条直线 $\mathbf{k}$ 和 $\mathbf{l}$ 的交点，从而是对极线 $\mathbf{l}$ 上的一点——记之为 $\mathbf{x}$. 因此，$F[\mathbf{k}]_\times\mathbf{l}=F\mathbf{x}$ 是对应于点 $\mathbf{x}$ 的对极线，即直线 $\mathbf{l}'$. □

进一步，$\mathbf{k}$ 的一种方便的选择是直线 $\mathbf{e}$，因为 $\mathbf{k}^T\mathbf{e}=\mathbf{e}^T\mathbf{e}\neq0$. 于是直线 $\mathbf{e}$ 不过点 $\mathbf{e}$，满足所需要求. 选择 $\mathbf{k}'=\mathbf{e}'$ 时类似推理也成立. 这样一来，对极线单应可以记为

$$\mathbf{l}'=F[\mathbf{e}]_\times\mathbf{l}\qquad \mathbf{l}=F^T[\mathbf{e}']_\times\mathbf{l}'.$$

## 8.3 由特殊运动产生的基本矩阵

平移方向 **t** 和转轴方向 **a** 之间的一种特殊关系产生一种特殊运动. 我们将要讨论两种情形:没有旋转的**纯平移**和 **t** 正交于 **a** 的**纯平面运动**(平面运动的意义在节 2.4.1($p$44)中介绍过). "纯"表示内参数没有改变. 这样的情形是重要的,首先是因为它们在实际中发生,比如一个摄像机拍摄正在转盘上旋转的物体等价于这一对视图的平面运动;其次是因为其基本矩阵有特殊的形式,因而有一些特殊的性质.

### 8.3.1 纯平移

在考虑摄像机的纯平移时,我们可以考虑与之等价的情形:让摄像机保持不动,而使对象世界作平移 $-\mathbf{t}$. 此时,3 维空间点沿着平行于 **t** 的直线运动,且这些平行直线交点的图像是 **t** 方向的消影点 **v**. 这种情形在图 8.7 和图 8.8 中给予说明. 显然 **v** 是两幅视图的对极点,而平行直线的像是对极线. 其代数细节在下面的例子中给出.

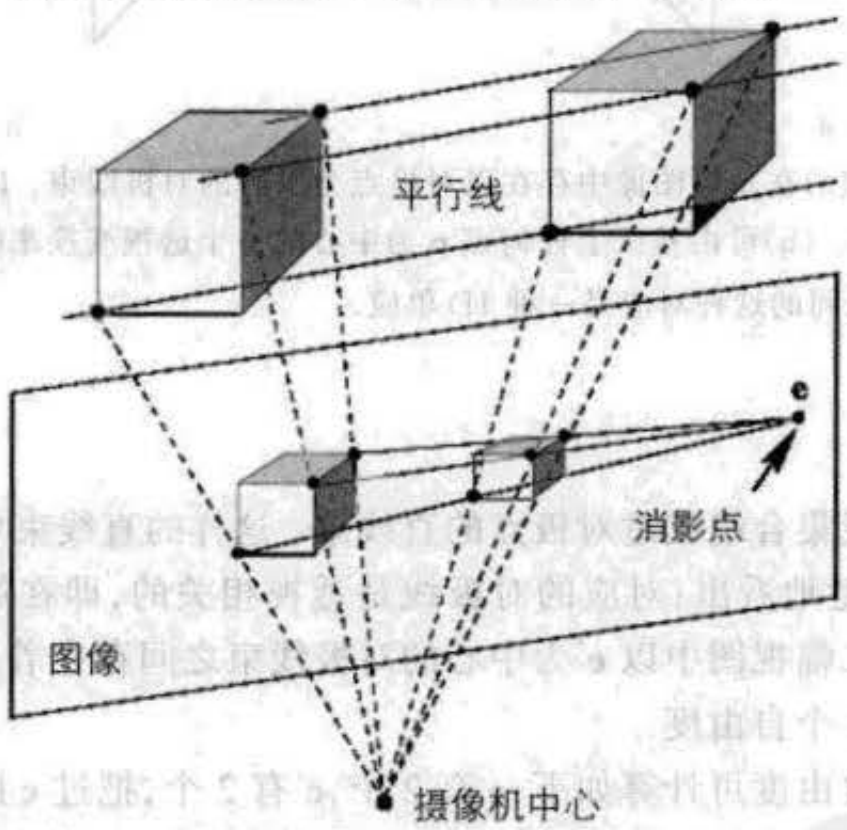



图 8.7 当摄像机作纯平移运动时,3D 点沿平行轨道滑动. 这些平行线的像相交于平移方向的消影点. 对极点 **e** 是消影点.

**例 8.6** 假设摄像机的运动是没有旋转并且内参数不变的纯平移. 可以假设两个摄像机矩阵分别是 $P=K[I|\mathbf{0}]$ 和 $P'=K[I|\mathbf{t}]$,则根据(8.4)(利用 $R=I$ 和 $K=K'$)得

$$F=[\mathbf{e}']_\times KK^{-1}=[\mathbf{e}']_\times .$$

如果摄像机平行于 $x$ - 轴平移,则 $\mathbf{e}'=(1,0,0)^\mathrm{T}$,因此

$$F=\begin{bmatrix}0&0&0\\0&0&-1\\0&1&0\end{bmatrix}$$

对应点之间的关系:$\mathbf{x}'^\mathrm{T}F\mathbf{x}=0$ 简化为 $y=y'$,即对极线是对应的光栅. 这正是节 10.12($p$211)中介绍的图像矫正中所要寻求的情形. △

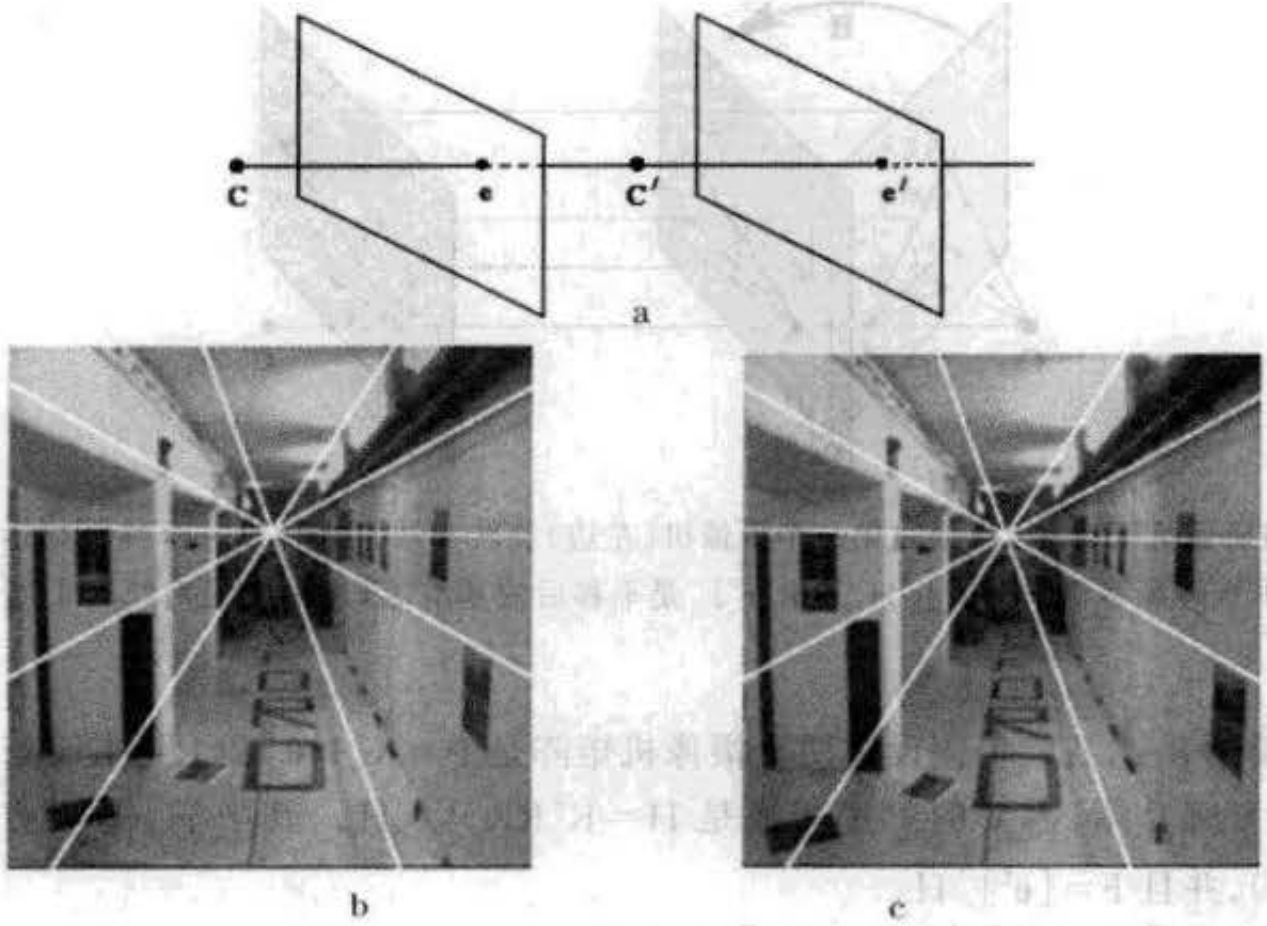



图 8.8 **纯平移运动**.(a)在这种运动下,对极点是不动点,即对极点在两幅图像中有相同的坐标,点好象是沿着由该对极点发出的射线移动. 在此情形中对极点被称为展开焦点(FOE),(b)和(c)两幅图像都叠加了一样的对极线. 注意墙上柱子的运动,它们沿着对极线滑动.

事实上,如果图像点 $\mathbf{x}$ 归一化为 $\mathbf{x}=(x,y,1)^{\mathsf{T}}$,则由 $\mathbf{x}=\mathrm{P}\mathbf{X}=\mathrm{K}[\mathrm{I}|\mathbf{0}]\mathbf{X}$ 可以得到空间点的(非齐次)坐标是 $(\mathrm{X},\mathrm{Y},\mathrm{Z})^{\mathsf{T}}=\mathrm{Z}\mathrm{K}^{-1}\mathbf{X}$,其中 Z 是点 $\mathbf{X}$ 的深度(在第一个摄像机的主轴上测量从 $\mathbf{X}$ 到摄像机中心的距离). 那么由 $\mathbf{x}'=\mathrm{P}'\mathbf{X}=\mathrm{K}[\mathrm{I}|\mathbf{t}]\mathbf{X}$ 得到从像点 $\mathbf{x}$ 到像点 $\mathbf{x}'$ 的映射是

$$\mathbf{x}'=\mathbf{x}+\mathrm{K}\mathbf{t}/\mathrm{Z}. \tag{8.6}$$

式(8.6)给出的运动 $\mathbf{x}'=\mathbf{x}+\mathrm{K}\mathbf{t}/\mathrm{Z}$ 指出像点从 $\mathbf{x}$ 点"开始"沿着由 $\mathbf{x}$ 与对极点 $\mathbf{e}=\mathbf{e}'=\mathbf{v}$ 定义的直线移动. 这个运动的范围与平移矢量 $\mathbf{t}$(在这里它不是一个齐次矢量)的大小成正比而与深度 Z 成反比,因此靠近摄像机的点看上去比那些离摄像机远的点移动得快——与人们由火车的窗户看出去时的体验相同.

注意在纯平移情形,$\mathrm{F}=[\mathbf{e}']_\times$ 是反对称的而且只有 2 个自由度,它们对应于对极点的位置. $\mathbf{x}$ 的对极线是 $\mathbf{l}'=\mathrm{F}\mathbf{x}=[\mathbf{e}]_\times\mathbf{x}$,且 $\mathbf{x}$ 在这条直线上,因为 $\mathbf{x}^{\mathsf{T}}[\mathbf{e}]_\times\mathbf{x}=0$,即 $\mathbf{x},\mathbf{x}'$ 与 $\mathbf{e}=\mathbf{e}'$ 是共线的(假设把两幅图像叠起来). 这种共线性质称为**自对极**,并且对于一般运动不成立.

**一般运动** 纯平移给出了关于一般运动更深入的洞察. 给定两个任意摄像机,我们可以把第一幅图像的摄像机旋转使它与第二个摄像机平行. 该旋转可以通过对第一幅图像施加一个射影变换来仿真. 针对两幅图像的标定矩阵的差别还可对第一幅图像施加进一步的矫正. 这两次矫正的结果等价于对第一幅图像作射影变换 H. 假设这些矫正已经完成,那么这两个摄像机之间的有效关系是一个纯平移. 因此矫正后的第一幅和第二幅图像的基本矩阵的形式是 $\hat{\mathrm{F}}=[\mathbf{e}']_\times$,满足 $\mathbf{x}'^{\mathsf{T}}\hat{\mathrm{F}}\,\hat{\mathbf{x}}=0$,其中 $\hat{\mathbf{x}}=\mathrm{H}\mathbf{x}$ 是第一幅图像上矫正后的点. 由此推导出 $\mathbf{x}'^{\mathsf{T}}[\mathbf{e}']_\times\mathrm{H}\mathbf{x}=0$,所以与原始点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 相对应的基本矩阵应是 $\mathrm{F}=[\mathbf{e}']_\times\mathrm{H}$,图 8.9 对此给予了说明.

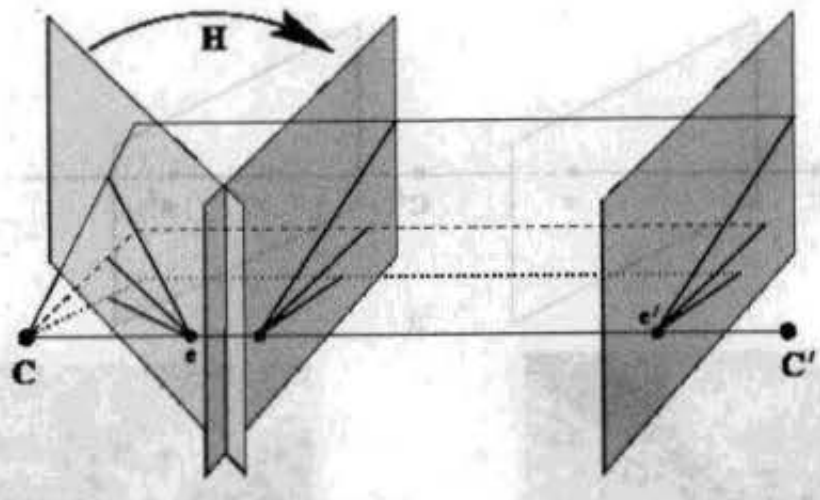



图 8.9 **一般摄像机运动**. 可以通过第一个摄像机(左边)旋转和矫正来仿真一个纯平移运动. 原始摄像机对的基本矩阵是乘积 $F=[\mathbf{e}']_\times H$,其中$[\mathbf{e}']_\times$是平移后的基本矩阵,而 H 是对第一个摄像机作矫正的射影变换.

**例 8.7** 继续例 8.2,我们仍假设这两个摄像机矩阵是 $P=K[I|\mathbf{0}]$和 $P'=K'[R|\mathbf{t}]$. 则如节 7.4.2($p$141) 所介绍,需要的射影变换是 $H=K'RK^{-1}=H_\infty$,其中 $H_\infty$ 是无穷单应(见节 12.4($p$239)),并且 $F=[\mathbf{e}']_\times H_\infty$.

与例 8.6 一样,如果图像点 $\mathbf{x}$ 归一化为 $\mathbf{x}=(x,y,1)^\mathsf{T}$,则$(X,Y,Z)^\mathsf{T}=ZK^{-1}\mathbf{x}$,且从 $\mathbf{x}=P'\mathbf{X}=K'[R|\mathbf{t}]\mathbf{X}$推出图像点 $\mathbf{x}$ 到图像点 $\mathbf{x}'$的映射是

$$\mathbf{x}'=K'RK^{-1}\mathbf{x}+K'\mathbf{t}/Z. \tag{8.7}$$

这个映射分为两部分:第一项只依赖于图像位置,即 $\mathbf{x}$ 而与点的深度 Z 无关,第一项还考虑了摄像机的旋转和内参数的变化;第二项依赖于深度,与图像位置 $\mathbf{x}$ 无关,并考虑了摄像机的平移. 在纯平移情形($R=I,K=K'$)式(8.7)简化为式(8.6). △

### 8.3.2 纯平面运动

在纯平面运动时旋转轴与平移方向正交. 正交性给这种运动增加了一个约束,在本章末的习题中证明了如果 $K'=K$,则 F 的对称部分:$F_s$ 在这种平面运动情形时的秩为 2(注:对于一般运动 F 的对称部分为满秩). 因此条件 $\det F_s=0$ 是加在 F 的一个附加约束并且使它的自由度个数从一般运动的 7 个减少到纯平面运动的 6 个.

## 8.4 基本矩阵的几何表示

本节对于初次阅读不是必须的,读者可以选择跳过此节而直接阅读节 8.5.

本节把基本矩阵分解为对称和反对称两部分,并且给出每一部分的几何表示. 基本矩阵的对称和反对称部分分别是

$$F_s=(F+F^\mathsf{T})/2 \qquad F_a=(F-F^\mathsf{T})/2$$

满足 $F=F_s+F_a$.

为解释这样分解的动机,考虑在两幅图像上映射为相同点的 3 维空间点 $\mathbf{X}$. 这样的图像点在摄像机运动时不变,即 $\mathbf{x}=\mathbf{x}'$. 这两点显然是对应点,从而满足对应点的必要条件 $\mathbf{x}^\mathsf{T}F\mathbf{x}=0$. 我们知道任何反对称矩阵 A 的二次形式 $\mathbf{x}^\mathsf{T}A\mathbf{x}$ 恒等于零. 因此只有 F 的对称部分对 $\mathbf{x}^\mathsf{T}F\mathbf{x}=0$ 起作用,从而导出 $\mathbf{x}^\mathsf{T}F_s\mathbf{x}=0$. 下面我们将会看到,矩阵 $F_s$ 可以视为图像平面上

的二次曲线.

从几何上讲,该二次曲线是这样产生的:3 维空间中满足 $\mathbf{x}=\mathbf{x}'$ 的点 $\mathbf{X}$ 的轨迹称为**同视点**曲线. 一般来说,此曲线是 3 维空间中过两个摄像机中心[Maybank-93]的一条三次绕线(见节 2.3($p$43)). 同视点的像是由 $F_s$ 定义的二次曲线. 我们将在第 21 章中再回过来介绍同视点.

**对称部分** 矩阵 $F_s$ 是对称的并且一般地秩为 3,它有 5 个自由度并等同于一条点二次曲线,称为 *Steiner* **二次曲线**(下面将解释这个名字). 对极点 $\mathbf{e}$ 和 $\mathbf{e}'$ 在二次曲线 $F_s$ 上. 为说明对极点在二次曲线上,即 $\mathbf{e}^T F_s \mathbf{e}=0$. 先从 $F\mathbf{e}=0$ 开始,然后有 $\mathbf{e}^T F\mathbf{e}=0$,从而有 $\mathbf{e}^T F_s \mathbf{e}+\mathbf{e}^T F_a \mathbf{e}=0$. 但因对任何反对称矩阵 S 都有 $\mathbf{x}^T S\mathbf{x}=0$,故 $\mathbf{e}^T F_a \mathbf{e}=0$,于是得到 $\mathbf{e}^T F_s \mathbf{e}=0$. 对 $\mathbf{e}'$ 的推导方法与此类似.

**反对称部分** 矩阵 $F_a$ 是反对称的并可以记为 $F_a=[\mathbf{x}_a]_\times$,其中 $\mathbf{x}_a$ 是 $F_a$ 的零矢量. 反对称部分有 2 个自由度且等同于点 $\mathbf{x}_a$.

点 $\mathbf{x}_a$ 与二次曲线 $F_s$ 之间的关系在图 8.10a 中说明. $\mathbf{x}_a$ 的极线与 Steiner 二次曲线 $F_s$ 相交于对极点 $\mathbf{e}$ 和 $\mathbf{e}'$(极点—极线关系在节 1.2.3($p$6)中介绍). 这个结果的证明留作习题.

**对极线对应** 根据 Steiner 引入的射影几何的经典定理[Semple-77],由一个单应关联的两个直线束,它们对应直线的交点的轨迹是一条二次曲线. 这正是我们目前面临的情形. 这两个束都是对极线束,一个过 $\mathbf{e}$ 而另一个过 $\mathbf{e}'$. 如节 8.2.5 所述,这些对极线由一个 1D 单应相关联. 其交点的轨迹是二次曲线 $F_s$.

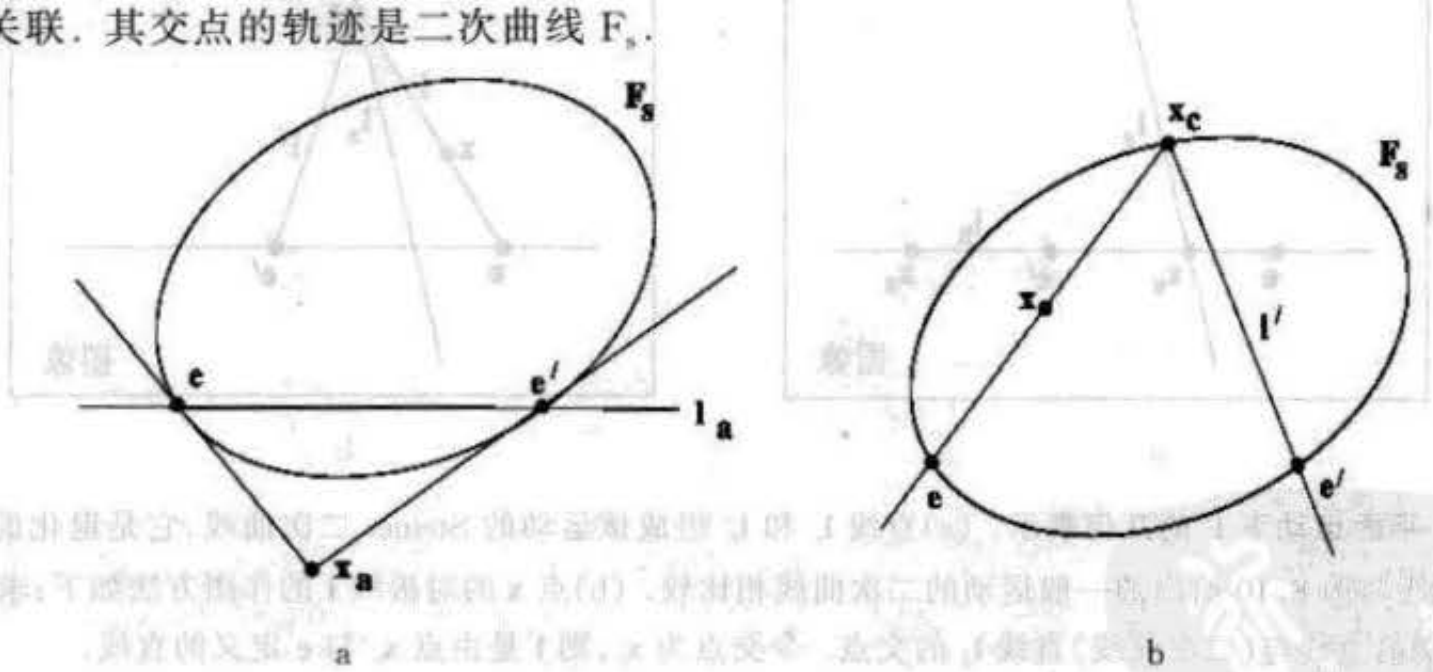



图 8.10 **F 的几何表示**. (a)二次曲线 $F_s$ 表示 F 的对称部分,而点 $\mathbf{x}_a$ 表示其反对称部分. 假定把两幅图像叠在一起,那么二次曲线 $F_s$ 是对应的对极线交点的轨迹. 它是同视点曲线的图像. 直线 $\mathbf{l}_a$ 是 $\mathbf{x}_a$ 相对于二次曲线 $F_s$ 的极线. 它与二次曲线交于对极点 $\mathbf{e}$ 和 $\mathbf{e}'$. (b)点 $\mathbf{x}$ 的对极线 $\mathbf{l}'$ 由下面的方法画图:作由点 $\mathbf{e}$ 和 $\mathbf{x}$ 定义的直线并与二次曲线相交. 令交点为 $\mathbf{x}_c$. 那么 $\mathbf{l}'$ 是由点 $\mathbf{x}_c$ 和 $\mathbf{e}'$ 定义的直线.

图 8.10b 说明通过几何作图可由二次曲线和对极点来确定对应的对极线. 这种作图是基于 Steiner 二次曲线 $F_s$ 的不动点性质. 第一幅视图的对极线 $\mathbf{l}=\mathbf{x}\times\mathbf{e}$ 在 3 维空间中定义了一张对极平面,它与同视点曲线相交于一点,我们记为 $\mathbf{X}_c$. 点 $\mathbf{X}_c$ 在第一幅视图的像为 $\mathbf{x}_c$,它是 $\mathbf{l}$ 与二次曲线 $F_s$ 的交点(因为 $F_s$ 是同视点的像). 由于同视点的不动点性质,$\mathbf{X}_c$ 在第二幅视图的像也是 $\mathbf{x}_c$. 所以 $\mathbf{x}_c$ 是点 $\mathbf{x}$ 的对极平面上的一点在第二幅视图上的像. 由此推出 $\mathbf{x}_c$

在 $\mathbf{x}$ 的对极线 $\mathbf{l}'$ 上，从而可以算出 $\mathbf{l}'$，即 $\mathbf{l}'=\mathbf{x}_c\times\mathbf{e}'$.

二次曲线 $F_s$ 和在这个二次曲线上的两点解释了 F 的 7 个自由度：二次曲线有 5 个自由度加上在此二次曲线上的对极点各有 1 个自由度. 给定了 F，则二次曲线 $F_s$、对极点 $\mathbf{e}$、$\mathbf{e}'$ 和反对称点 $\mathbf{x}_a$ 都能唯一地确定，然而，$F_s$ 和 $\mathbf{x}_a$ 不能唯一地确定 F，因为两个对极点还不能唯一地确定，即 $\mathbf{x}_a$ 的极线可确定对极点，但是并没有确定哪一个是 e 和哪一个是 e′.

**纯平面运动** 我们回到在上节 8.3.2 讨论的平面运动情形，其中 $F_s$ 的秩为 2. 显然此时，Steiner 二次曲线是退化的而且根据节 1.2.3($p8$)，它等价于两条不重合直线.

$$F_s=\mathbf{l}_h\mathbf{l}_s^{\mathsf{T}}+\mathbf{l}_s\mathbf{l}_h^{\mathsf{T}}$$

如图 8.11a 所示. 对应于节 8.4 中点 $\mathbf{x}$ 的对极线 $\mathbf{l}'$ 的几何作图在目前情形下有一个简单的代数表示. 与一般运动一样，它分为三个步骤，如图 8.11b 说明：第一，计算连接 $\mathbf{e}$ 和 $\mathbf{x}$ 的直线 $\mathbf{l}=\mathbf{e}\times\mathbf{x}$；第二，确定它与该"二次曲线"的交点 $\mathbf{x}_c=\mathbf{l}_s\times\mathbf{l}$；第三，对极线 $\mathbf{l}'=\mathbf{e}'\times\mathbf{x}_c$ 是 $\mathbf{x}_c$ 和 $\mathbf{e}'$ 的连线. 综合这些步骤得到

$$\mathbf{l}'=\mathbf{e}'\times[\mathbf{l}_s\times(\mathbf{e}\times\mathbf{x})]=[\mathbf{e}']_\times[\mathbf{l}_s]_\times[\mathbf{e}]_\times\mathbf{x}.$$

从而推出 F 可以记为

$$F=[\mathbf{e}']_\times[\mathbf{l}_s]_\times[\mathbf{e}]_\times. \tag{8.8}$$

F 的 6 个自由度解释为：两个对极点的分别有 2 个自由度和直线 $\mathbf{l}_s$ 有 2 个自由度.

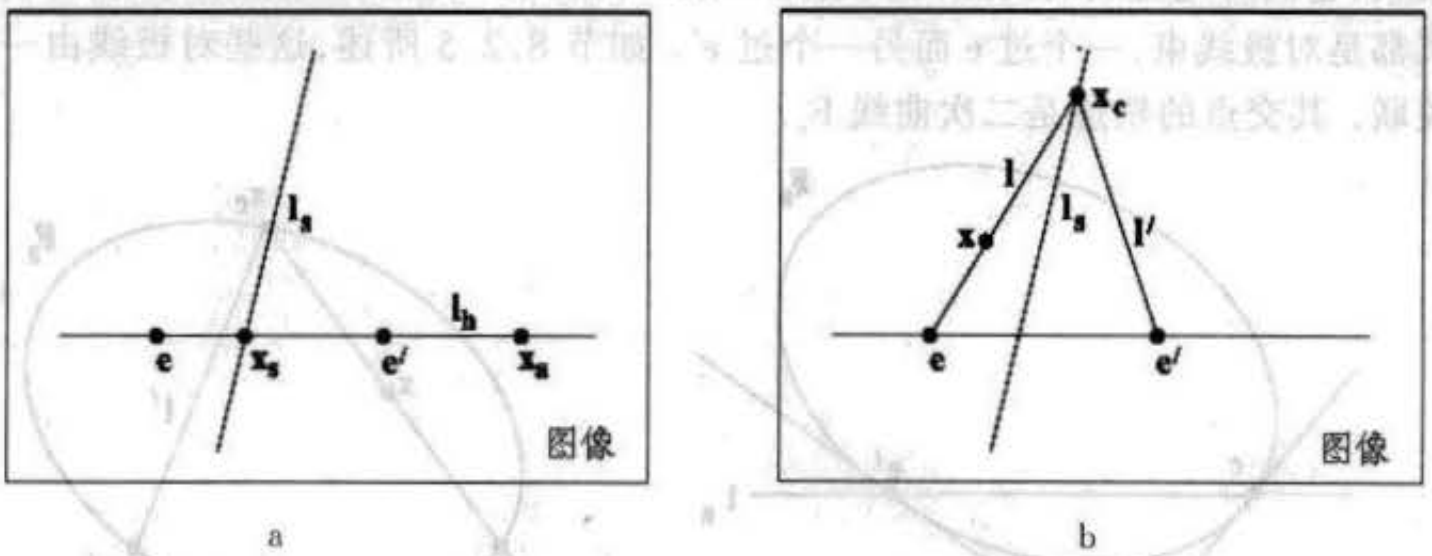



图 8.11 **平面运动下 F 的几何表示**. (a)直线 $\mathbf{l}_s$ 和 $\mathbf{l}_h$ 组成该运动的 Steiner 二次曲线，它是退化的. 将此二次曲线与图 8.10 给出的一般运动的二次曲线相比较. (b)点 $\mathbf{x}$ 的对极线 $\mathbf{l}'$ 的作图方法如下：求由点 $\mathbf{e}$ 和 $\mathbf{x}$ 定义的直线与(二次曲线)直线 $\mathbf{l}_s$ 的交点. 令交点为 $\mathbf{x}_c$，则 $\mathbf{l}'$ 是由点 $\mathbf{x}_c$ 与 $\mathbf{e}'$ 定义的直线.

这种情形的几何能够方便地可视化：这个运动所对应的同视点是一条退化三次绕线，由在这张运动平面上的一个圆(这个平面同旋转轴正交且包含两个摄像机中心)和一个平行于旋转轴并且与圆相交的直线组成，这条直线是转动轴(见节 2.4.1($p44$)). 这个运动等价于平移为零绕该转动轴的一个旋转. 在这个运动下，转动轴上的点是固定的，因此它们的图像是固定的. 直线 $\mathbf{l}_s$ 是转动轴的像. 直线 $\mathbf{l}_h$ 是图像与运动平面的交线. 第 18 章将用这个几何进行自标定.

## 8.5 恢复摄像机矩阵

到目前为止，我们已经研究了 F 的性质以及点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 的图像关系. 现在我们转向 F 的一个最重要的性质，即该矩阵可以用来确定两视图的摄像机矩阵.

### 8.5.1 射影不变性和规范摄像机

显然，根据节 8.2 的推导，映射 $\mathbf{l}'=\mathrm{F}\mathbf{x}$ 和对应条件 $\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0$ 都是**射影**关系：因为这个推导只涉及诸如直线和平面的相交等射影几何关系，并且有关代数方面也不过是世界和图像点之间射影摄像机的线性映射. 因此这些关系只依赖于图像上的射影坐标，并不依赖于诸如射线的夹角等欧氏度量. 换句话说，图像关系是射影不变的：即在图像坐标的射影变换 $\hat{\mathbf{x}}=\mathrm{H}\mathbf{x}$，$\hat{\mathbf{x}}'=\mathrm{H}'\mathbf{x}'$ 之下，存在一个对应映射 $\hat{\mathbf{l}}'=\hat{\mathrm{F}}\,\hat{\mathbf{x}}$，其中 $\hat{\mathrm{F}}=\mathrm{H}'^{-\mathrm{T}}\mathrm{F}\mathrm{H}^{-1}$ 是对应的秩 2 的基本矩阵.

类似地，F 也仅依赖于摄像机矩阵 P、P′ 的射影性质. 摄像机矩阵把 3 维测量与图像的测量相关联，因此它们既依赖于图像的坐标系，也依赖于世界坐标系的选取. 然而 F 却不依赖于世界坐标系的选取，例如，世界坐标系的一个旋转改变 P、P′，但不改变 F. 事实上，基本矩阵在 3 维空间的射影变换下总是不变的，更准确地说，

**结论 8.8　如果 H 是表示 3 维射影变换的一个 4×4 矩阵，那么对应于摄像机矩阵(P,P′)和(PH,P′H)的基本矩阵是相同的.**

**证明**　注意到 $\mathrm{P}\mathbf{X}=(\mathrm{PH})(\mathrm{H}^{-1}\mathbf{X})$ 并且 P′ 也有类似关系式，如果 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 是对应于 3D 点 $\mathbf{X}$ 的关于摄像机对(P,P′)的匹配点，则它们也是关于摄像机对(PH,P′H)的匹配点并且对应于点 $\mathrm{H}^{-1}\mathbf{X}$. □

因此，虽然(从(8.1－$p$163)可知)摄像机矩阵对(P,P′)唯一决定基本矩阵 F，但反过来不成立. 由基本矩阵来确定摄像机矩阵最好的结果也要相差一个右乘 3D 射影变换. 下面将看到这也是多义性的全部，即在相差一个射影变换的意义下摄像机矩阵可以由基本矩阵确定.

**摄像机矩阵的规范形式**　由于这样的多义性，通常要对给定的基本矩阵定义摄像机矩阵对的一种特殊的**规范形式**，即规定第一个矩阵取简单的形式$[\mathrm{I}|\mathbf{0}]$，其中 I 是 3×3 单位矩阵而 $\mathbf{0}$ 为 3 维零矢量. 该规定总是可能的，即让 P 增加一行使之形成 4×4 非奇异矩阵，并记为 $\mathrm{P}^*$. 现在取 $\mathrm{H}=\mathrm{P}^{*-1}$，可以验证 $\mathrm{PH}=[\mathrm{I}|\mathbf{0}]$ 正如所求.

下面的结论将经常被用到.

**结论 8.9　对应于摄像机矩阵对 $\mathrm{P}=[\mathrm{I}|0]$ 和 $\mathrm{P}'=[\mathrm{M}|\mathbf{m}]$ 的基本矩阵等于 $[\mathbf{m}]_\times\mathrm{M}$.**

它是(8.1－$p$163)的特殊情形因而不难导出.

### 8.5.2 给定 F 后的摄像机射影多义性

我们已经知道一对摄像机矩阵唯一地确定一个基本矩阵. 但是这个映射不是单射(一对一)，因为相差一个射影变换的摄像机矩阵对都定义同一个基本矩阵. 我们将证明仅存在

这样的多义性，即在相差右乘一个射影变换的意义下，一个给定的基本矩阵唯一确定一对摄像机矩阵. 因此基本矩阵主要刻画了两个摄像机的射影关系.

**定理 8.10　令 F 为基本矩阵而$(P,P')$和$(\widetilde{P},\widetilde{P}')$都是与基本矩阵 F 对应的两组摄像机矩阵对，则存在一个非奇异的 $4\times4$ 矩阵 H 使得 $\widetilde{P}=PH$ 和 $\widetilde{P}'=P'H$.**

**证明**　假设给定的 F 对应于两组不同的摄像机矩阵对$(P,P')$和$(\widetilde{P},\widetilde{P}')$. 首先，为了简化问题我们可以假定这两组摄像机矩阵对都具有规范形式 $P=\widetilde{P}=[I|\mathbf{0}]$，因为如果需要的话可以对每组摄像机对进行适当的射影变换来做到这一点. 因此可假设 $P=\widetilde{P}=[I|\mathbf{0}]$，$P'=[A|\mathbf{a}]$和 $\widetilde{P}'=[\widetilde{A}|\widetilde{\mathbf{a}}]$，根据结论 8.9，这个基本矩阵可写为 $F=[\mathbf{a}]_\times A=[\widetilde{\mathbf{a}}]_\times\widetilde{A}$.

我们需要下面的引理：

**引理 8.11　若秩 2 的矩阵 F 可以分解为两个不同的形式：$F=[\mathbf{a}]_\times A$ 和 $F=[\widetilde{\mathbf{a}}]_\times\widetilde{A}$，则对某非零常数 $k$ 和 3 维矢量 $\mathbf{v}$ 有 $\widetilde{\mathbf{a}}=k\mathbf{a}$ 和 $\widetilde{A}=k^{-1}(A+\mathbf{a}\mathbf{v}^\mathsf{T})$.**

**证明**　首先，注意到 $\mathbf{a}^\mathsf{T}F=\mathbf{a}^\mathsf{T}[a]_\times A=\mathbf{0}$，同理 $\widetilde{\mathbf{a}}^\mathsf{T}F=\mathbf{0}$，由于 F 的秩为 2，因此得到所需要的结果 $\widetilde{\mathbf{a}}=k\mathbf{a}$. 其次，从$[\mathbf{a}]_\times A=[\widetilde{\mathbf{a}}]_\times\widetilde{A}$ 得到$[\mathbf{a}]_\times(k\widetilde{A}-A)=\mathbf{0}$，因而对某个 $\mathbf{v}$，有 $k\widetilde{A}-A=\mathbf{a}\mathbf{v}^\mathsf{T}$. 由此又得到所需要的另一结果 $\widetilde{A}=k^{-1}(A+\mathbf{a}\mathbf{v}^\mathsf{T})$. □

对两个摄像机矩阵 $P'$和 $\widetilde{P}'$运用引理 8.11 表明：如果它们产生同一个 F，那么 $P'=[A|\mathbf{a}]$和 $\widetilde{P}'=[k^{-1}(A+\mathbf{a}\mathbf{v}^\mathsf{T})|k\mathbf{a}]$. 现在只剩下证明这两个摄像机对是射影相关的. 令矩阵 $H=\begin{bmatrix}k^{-1}I & \mathbf{0}\\ k^{-1}\mathbf{v}^\mathsf{T} & k\end{bmatrix}$. 则可验证 $PH=k^{-1}[I|\mathbf{0}]=k^{-1}\widetilde{P}$，并且进一步有：

$$P'H=[A|\mathbf{a}]H=[k^{-1}(A+\mathbf{a}\mathbf{v}^\mathsf{T})|k\mathbf{a}]=[\widetilde{A}|\widetilde{\mathbf{a}}]=\widetilde{P}'$$

因此，矩阵对 $P,P'$和 $\widetilde{P},\widetilde{P}'$的确是射影相关的. □

这一结论也可以用来精确计算自由度数：摄像机矩阵 P 和 $P'$各有 11 个自由度，共有 22 个自由度. 确定世界射影坐标系，需要 15 个自由度(2.1 节($p$34))，因此当把世界坐标系的自由度从两个摄像机自由度中取走时，剩下 $22-15=7$ 个自由度——它正好对应基本矩阵的 7 个自由度.

### 8.5.3　给定 F 求规范摄像机对

我们已经证明了在相差一个 3D 的射影变换下，F 唯一确定摄像机对. 给定 F，我们将推导出具有规范形式的摄像机对的具体公式，我们将利用对应于摄像机矩阵对的基本矩阵 F 的下述特征.

**结论 8.12　一个非零矩阵 F 是对应于一对摄像机矩阵 P 和 $P'$的基本矩阵的充要条件是 $P'^\mathsf{T}FP$ 是反对称矩阵.**

**证明**　$P'^\mathsf{T}FP$ 是反对称矩阵的条件等价于对所有 $\mathbf{X}$ 有 $\mathbf{X}^\mathsf{T}P'^\mathsf{T}FP\mathbf{X}=0$ 成立. 令 $\mathbf{x}'=P'\mathbf{X}$，$\mathbf{x}=P\mathbf{X}$，它就等价于基本矩阵的定义方程：$\mathbf{x}'^\mathsf{T}F\mathbf{x}=0$. □

对应于一个基本矩阵，我们可以写出在规范形式下摄像机矩阵对的具体解：

**结论 8.13　令 F 是基本矩阵，S 是任意反对称矩阵. 定义摄像机矩阵对为**

$$P=[I|\mathbf{0}]\text{和 }P'=[SF|\mathbf{e}'],$$

**其中 $\mathbf{e}'$ 是满足 $\mathbf{e}'^{\mathsf{T}}F=\mathbf{0}$ 的对极点，并假定所定义的 $P'$ 是有效的摄像机矩阵(秩为 3)，则 F 是对应于$(P,P')$的基本矩阵.**

为证明它，我们利用结论 8.12 并且直接验证

$$[SF \mid \mathbf{e}']^{\mathsf{T}}F[I \mid \mathbf{0}] = \begin{bmatrix} F^{\mathsf{T}}S^{\mathsf{T}}F & \mathbf{0} \\ \mathbf{e}'^{\mathsf{T}}F & 0 \end{bmatrix} = \begin{bmatrix} F^{\mathsf{T}}S^{\mathsf{T}}F & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & 0 \end{bmatrix} \tag{8.9}$$

是一个反对称矩阵.

反对称矩阵 S 可用它的零矢量表为 $S=[\mathbf{s}]_\times$. 按下文的论证，如果 $\mathbf{s}^{\mathsf{T}}\mathbf{e}'\neq 0$，则$[[\mathbf{s}]_\times F|\mathbf{e}']$的秩为 3. 因 $\mathbf{e}'^{\mathsf{T}}F=\mathbf{0}$，故 F 的列空间(由列生成的空间)垂直于 $\mathbf{e}'$. 但是如果 $\mathbf{s}^{\mathsf{T}}\mathbf{e}'\neq 0$，则 $\mathbf{s}$ 不垂直于 $\mathbf{e}'$，因而不在 F 的列空间中. 现在，$[\mathbf{s}]_\times F$ 的列空间是由 $\mathbf{s}$ 与 F 列矢量的叉积生成，因而等于垂直于 $\mathbf{s}$ 的平面，所以$[\mathbf{s}]_\times F$ 的秩为 2. 又因 $\mathbf{e}'$ 不垂直于 $\mathbf{s}$，故 $\mathbf{e}'$ 不在这个平面上，所以$[[\mathbf{s}]_\times F|\mathbf{e}']$的秩为 3，正如所求.

Luong 和 Vieville[Luong－96]曾建议，S 的一个适当的选择是 $S=[\mathbf{e}']_\times$，因为$\mathbf{e}'^{\mathsf{T}}\mathbf{e}'\neq 0$，由此导出下列有用结论.

**结论 8.14　基本矩阵 F 的摄像机矩阵可以选择为 $P=[I|\mathbf{0}]$和 $P'=[[\mathbf{e}']_\times F|\mathbf{e}']$.**

注意摄像机矩阵 $P'$左边的 $3\times 3$ 子矩阵$[\mathbf{e}']_\times F$ 的秩为 2，它对应于中心在 $\boldsymbol{\pi}_\infty$ 上的摄像机. 不过我们也没有特别的理由要回避这种情形.

定理 8.10 的证明说明了摄像机对在规范形式 $\tilde{P}=[I|0]$和 $\tilde{P}'=[A+\mathbf{a}\mathbf{v}^{\mathsf{T}}|k^2\mathbf{a}]$下的 4 参数簇与规范形式对 $P=[I|0]$、$P'=[A|\mathbf{a}]$有相同的基本矩阵，故它是最一般的解，概括起来有：

**结论 8.15　对应于基本矩阵 F，一对规范形式的摄像机矩阵的一般公式是**

$$P=[I \mid \mathbf{0}] \qquad P'=[[\mathbf{e}']_\times F+\mathbf{e}'\mathbf{v}^{\mathsf{T}} \mid \lambda\mathbf{e}'] \tag{8.10}$$

**其中 $\mathbf{v}$ 是任何 3 维矢量，$\lambda$ 是一个正标量.**

## 8.6　本质矩阵

本质矩阵是归一化图像(见下面)坐标下的基本矩阵的特殊形式. 历史上本质矩阵的引入(由 Louguet－Higgins)比基本矩阵还早，并且基本矩阵可以看作是本质矩阵的推广，基本矩阵把要求标定摄像机这样一个(非本质的)假设去掉了. 与基本矩阵相比较，本质矩阵有较少的自由度但增加了一些性质. 这些性质在下文中介绍.

**归一化坐标**　考虑分解为 $P=K[R|\mathbf{t}]$的摄像机矩阵，并令 $\mathbf{x}=P\mathbf{X}$ 为图像上的一点. 如果知道标定矩阵 K，用它的逆矩阵作用于点 $\mathbf{x}$ 得到点 $\hat{\mathbf{x}}=K^{-1}\mathbf{x}$. 则 $\hat{\mathbf{x}}=[R|\mathbf{t}]\mathbf{X}$，其中 $\hat{\mathbf{x}}$是图像上的点在**归一化坐标**下的表示. 它可以被视为空间点 $\mathbf{X}$ 在摄像机$[R|\mathbf{t}]$的标定矩阵等于单位矩阵 I 情形下的像. 摄像机矩阵 $K^{-1}P=[R|\mathbf{t}]$称为**归一化摄像机矩阵**，对于它，标定矩阵的影响已经去掉了.

现在，考虑一对归一化的摄像机矩阵 $P=[I|\mathbf{0}]$和 $P'=[R|\mathbf{t}]$. 与归一化摄像机矩阵对应的基本矩阵按惯例称为**本质矩阵**，根据(8.2－$p$163)，它具有如下形式

$$E=[\mathbf{t}]_\times R=R[R^{\mathsf{T}}\mathbf{t}]_\times.$$

**定义 8.16** 用归一化图像坐标表示对应点 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 时，本质矩阵的定义方程是

$$\hat{\mathbf{x}}'^{\mathsf{T}}\mathrm{E}\hat{\mathbf{x}} = 0. \tag{8.11}$$

以 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$ 的表示式代入(8.11)得 $\mathbf{x}'^{\mathsf{T}}\mathrm{K}'^{-\mathsf{T}}\mathrm{E}\mathrm{K}^{-1}\mathbf{x}=0$，把它与基本矩阵关系式 $\mathbf{x}'^{\mathsf{T}}\mathrm{F}\mathbf{x}=0$ 比较便可推出基本矩阵和本质矩阵之间的关系是

$$\mathrm{E}=\mathrm{K}'^{\mathsf{T}}\mathrm{F}\mathrm{K}. \tag{8.12}$$

### 8.6.1 本质矩阵的性质

本质矩阵 $\mathrm{E}=[\mathbf{t}]_\times\mathrm{R}$ 只有 5 个自由度：旋转矩阵 R 和平移矢量 $\mathbf{t}$ 各有 3 个自由度，但是有一个整体尺度因子的多义性——与基本矩阵一样本质矩阵也是一个齐次量.

本质矩阵减少的自由度数转变为比基本矩阵满足更多的约束，我们就来探讨到底是一些什么样的约束.

**结论 8.17 一个 $3\times3$ 矩阵是本质矩阵的充要条件是它的奇异值中有两个相等而第三个是 0.**

**证明** 这个不难从 E 的分解式 $[\mathbf{t}]_\times\mathrm{R}=\mathrm{SR}$ 推出，其中 S 是反对称矩阵. 我们将利用下列矩阵

$$\mathrm{W}=\begin{bmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix},\quad \mathrm{Z}=\begin{bmatrix}0 & 1 & 0\\ -1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}. \tag{8.13}$$

可以验证 W 为正交矩阵而 Z 为反对称矩阵. 根据结论 A3.1($p$406)给出的一般反对称矩阵的块分解方法，$3\times3$ 反对称矩阵 S 可以写为 $\mathrm{S}=k\mathrm{U}\mathrm{Z}\mathrm{U}^{\mathsf{T}}$，其中 U 正交. 注意在相差一个符号的意义下，$\mathrm{Z}=\mathrm{diag}(1,1,0)\mathrm{W}$，故在相差一个常数的意义下，$\mathrm{S}=\mathrm{U}\mathrm{diag}(1,1,0)\mathrm{W}\mathrm{U}^{\mathsf{T}}$ 和 $\mathrm{E}=\mathrm{SR}=\mathrm{U}\mathrm{diag}(1,1,0)(\mathrm{W}\mathrm{U}^{\mathsf{T}}\mathrm{R})$. 这正是 E 的奇异值分解并如所需要证明的具有两个相等的奇异值. 反之，一个具有两个相同奇异值的矩阵可以用同样方法分解为 SR. □

既然 $\mathrm{E}=\mathrm{U}\mathrm{diag}(1,1,0)\mathrm{V}^{\mathsf{T}}$，乍看似乎 E 有 6 个而不是 5 个自由度，因为 U、V 都有 3 个自由度. 然而由于它的两个奇异值相等，此 SVD 分解并不是唯一的——事实上，有一个关于E 的单参数簇的 SVD 分解. 的确，另一种 SVD 可选择为 $\mathrm{E}=(\mathrm{U}\mathrm{diag}(\mathrm{R}_{2\times2},1))\mathrm{diag}(1,1,0)(\mathrm{diag}(\mathrm{R}_{2\times2}^{\mathsf{T}},1))\mathrm{V}^{\mathsf{T}}$，其中 R 为任意 $2\times2$ 旋转矩阵.

### 8.6.2 由本质矩阵恢复摄像机矩阵

本质矩阵可以利用归一化图像坐标从(8.11)直接计算，或者利用(8.12)由基本矩阵来计算.（计算基本矩阵的方法推迟到第 10 章再介绍）. 一旦本质矩阵已知，如下文所述，摄像机矩阵便可由 E 得到恢复. 与基本矩阵具有射影多义性的情形不同，在相差一个尺度因子和一个 4 重多义性下，可以从本质矩阵恢复摄像机矩阵. 这就是说，除整体尺度因子外，不能被确定的仅是 4 个可能解.

可以假定第一个摄像机矩阵是 $\mathrm{P}=[\mathrm{I}|\mathbf{0}]$. 为了计算第二个摄像机矩阵 P′，必须把 E 分解成反对称矩阵和旋转矩阵的乘积 SR.

**结论 8.18 若 E 的 SVD 分解为 $\mathrm{U}\mathrm{diag}(1,1,0)\mathrm{V}^{\mathsf{T}}$，利用(8.13)的记号，$\mathrm{E}=\mathrm{SR}$ 有如下两种可能的分解(忽略符号)：**

$$\mathrm{S}=\mathrm{U}\mathrm{Z}\mathrm{U}^{\mathsf{T}}\quad \mathrm{R}=\mathrm{U}\mathrm{W}\mathrm{V}^{\mathsf{T}}\ \text{或者}\ \mathrm{U}\mathrm{W}^{\mathsf{T}}\mathrm{V}^{\mathsf{T}}. \tag{8.14}$$

**证明** 不难看出这种分解是有效的. 下面证明不存在其它的分解. 假定 $E=SR$, $S$ 的形式是根据它的左零矢量也是 $E$ 的左零矢量这样的事实来确定的. 故有 $S=UZU^T$, 旋转 $R$ 可记为 $UXV^T$, 其中 $X$ 是某个旋转矩阵, 则

$$U\mathrm{diag}(1,1,0)V^T=E=SR=(UZU^T)(UXV^T)=U(ZX)V^T$$

由此推出 $ZX=\mathrm{diag}(1,1,0)$. 因 $X$ 是一个旋转矩阵, 故又推出 $X=W$ 或 $X=W^T$, 即得证所需要的结论. □

分解式(8.14)在相差尺度因子的意义下由 $S=[\mathbf{t}]_\times$ 确定了摄像机矩阵 $P'$ 的 $\mathbf{t}$ 部分. 然而, $S=UZU^T$ 的 Frobenius 范数平方是 2 意味着如果 $S=[\mathbf{t}]_\times$ **(包含尺度因子)**, 则 $\|\mathbf{t}\|=1$. 这是对两个摄像机矩阵基线的一种常用的归一化. 因为 $S\mathbf{t}=\mathbf{0}$, 所以 $\mathbf{t}=U(0,0,1)^T=\mathbf{u}_3$, 即等于 $U$ 的最后一列. 然而, $E$ 的符号进而 $\mathbf{t}$ 的符号仍不能决定. 因此, 对应于给定的本质矩阵, 基于 $R$ 的两种可能选择和 $\mathbf{t}$ 的两种可能符号, 摄像机矩阵 $P'$ 有四种可能的选择. 归结为:

**结论 8.19 已知本质矩阵 $E=U\mathrm{diag}(1,1,0)V^T$ 和第一个摄像机矩阵 $P=[I|\mathbf{0}]$, 那么第二个摄像机矩阵 $P'$ 有下列四种可能的选择:**

$$P'=[UWV^T|\mathbf{u}_3];\quad [UWV^T|-\mathbf{u}_3];\quad [UW^TV^T|\mathbf{u}_3];\quad [UW^TV^T|-\mathbf{u}_3].$$

### 8.6.3 四个解的几何解释

显然前面二个解之间的差别就是第一个摄像机到第二个的平移矢量是反向的.

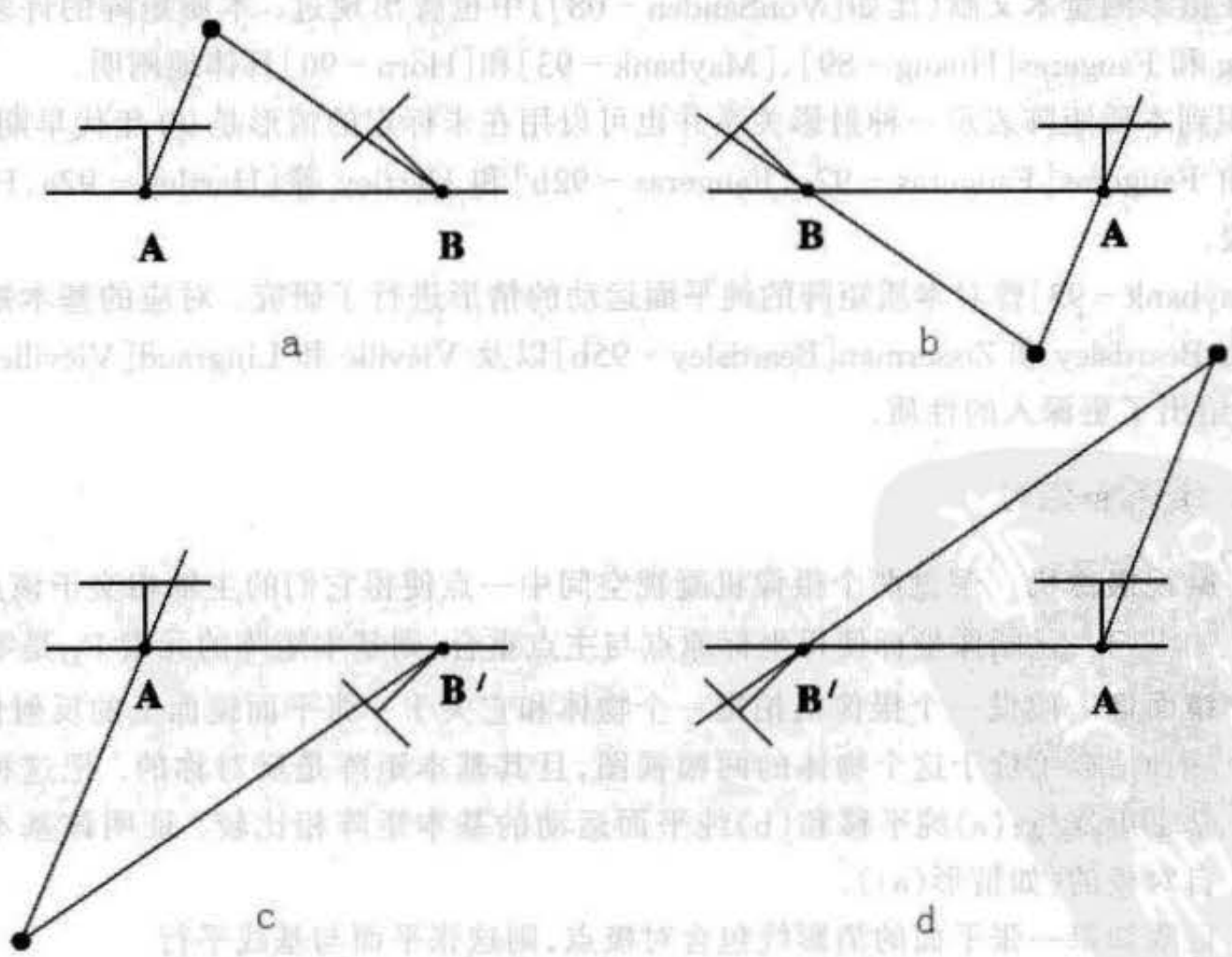



**图 8.12 由 E 作标定重构的 4 个可能解.** 左和右两边的差别是基线倒置. 上下两行的差别是摄像机 B 绕基线旋转了 180°. 注意仅在(a)时重构点同时在两个摄像机的前面.

结论 8.19 中的第一与第三个解的关系更复杂一点. 然而可以验证

$$[UW^TV^T \mid \mathbf{u}_3] = [UWV^T \mid \mathbf{u}_3]\begin{bmatrix} VW^TW^TV^T & \\ & 1 \end{bmatrix}$$

且 $VW^TW^TV^T = V\mathrm{diag}(-1,-1,1)V^T$ 是绕两个摄像机中心的连线作 180°的旋转. 用这种方式联系起来的两个解称为"扭转对".

图 8.12 对四个解做了说明,其中证明了重构的点 $\mathbf{X}$ 同时在两个摄像机的前面的情形在四个可能解中仅为一个. 这样一来只要用一个点作测试,验证它是否在两个摄像机前面就足以从四个不同解中确定一个作为摄像机矩阵 $P'$.

**注** 这里所采用的观点是本质矩阵是一个齐次量. 另一个可选择的观点是本质矩阵是由方程 $E=[\mathbf{t}]_\times R$ 精确地确定(即包含尺度因子),并在相差一个不确定的尺度因子的意义下,由方程 $\mathbf{x}'^TE\mathbf{x}=0$ 确定. 观点的选择依赖于你考虑用这两个方程中的哪一个来定义本质矩阵的性质.

## 8.7 结束语

### 8.7.1 文献

本质矩阵由 Longuet-Higgens[Longuet Higgens-81]引入到计算机视觉界,与 E 类似的矩阵在摄影测量术文献(比如[VonSanden-08])中也曾出现过. 本质矩阵的许多性质曾由 Huang 和 Faugeras[Huang-89]、[Maybank-93]和[Horn-90]具体地阐明.

认识到本质矩阵表示一种射影关系并也可以用在未标定的情形是 90 年代早期的成果,它同时由 Faugeras[Faugeras-92a, Faugeras-92b]和 Hartley 等[Hartley-92a, Hartley-92c]发表.

[Maybank-93]曾对本质矩阵的纯平面运动的情形进行了研究. 对应的基本矩阵的情形则是由 Beardsley 和 Zisserman[Beardsley-95b]以及 Viéville 和 Lingraud[Viéville-95]研究,其中给出了更深入的性质.

### 8.7.2 注释和练习

(1) **凝视摄像机** 假设两个摄像机凝视空间中一点使得它们的主轴相交于该点. 证明如果归一化图像坐标使得坐标原点与主点重合,则基本矩阵的元素 $F_{33}$ 是零.

(2) **镜面像** 假设一个摄像机拍摄一个物体和它关于一张平面镜面上的反射像. 证明这种情形等价于这个物体的两幅视图,且其基本矩阵是反对称的. 把这种配置的基本矩阵与:(a)纯平移和(b)纯平面运动的基本矩阵相比较. 证明该基本矩阵是自对极的(如情形(a)).

(3) 证明如果一张平面的消影线包含对极点,则这张平面与基线平行.

(4) Steiner **二次曲线** 证明 $\mathbf{x}_a$ 的极线与 Steiner 二次曲线 $\mathbf{F}_a$ 相交于对极点(图 8.10a). 提示,从 $F\mathbf{e}=F_a\mathbf{e}+F_a\mathbf{e}=\mathbf{0}$ 开始,因为 $\mathbf{e}$ 在二次曲线 $F_a$ 上,则 $\mathbf{l}_1=F_a\mathbf{e}$ 是过 $\mathbf{e}$ 点的切线,且 $\mathbf{l}_2=F_a\mathbf{e}=[\mathbf{x}_a]_\times\mathbf{e}=\mathbf{x}_a\times\mathbf{e}$ 是通过 $\mathbf{x}_a$ 和 $\mathbf{e}$ 的直线.

(5) **平面运动** [Maybank－93]证明，如果转轴方向是正交或平行于平移方向，则本质矩阵的对称部分的秩为 2. 这里我们假设 $K=K'$，则根据(8.12)有 $F=K^{-T}EK^{-1}$. 因而

$$F_s=(F+F^T)/2=K^{-T}(E+E^T)K^{-T}/2=E^{-T}E_sK^{-1}.$$

再由 $\det(F_s)=\det(K^{-1})^2\det(E_s)$ 推出 F 的对称部分也是奇异的. 如果 $K\neq K'$，这个结果成立吗？

(6) 任何秩 2 的矩阵 F 是对应于某摄像机对$(P,P')$的基本矩阵. 这可直接由结论 8.14 推出，因为规范摄像机的解只依赖于 F 的秩 2 性质.

(7) 证明由 E 得到重构的多义性中的一种重构所确定的 3D 点与由另一种重构所确定的对应 3D 点的联系或者是(1)通过第二个摄像机中心的一个倒置；或(2)3 维空间的一个调和透射(见节 A5.2($p$427))，其中透射平面与基线垂直并且通过第二个摄像机的中心，而顶点是第一个摄像机的中心.

(8) 仿照节 8.2.2 的推导，推导两个线阵推扫式摄像机的基本矩阵的形式. 这个矩阵的细节在[Gupta－97]中给出，该文证明了由一对图像进行仿射重构是可能的.

# 第 9 章　摄像机和结构的 3D 重构

本章介绍由两幅视图可以怎样或在什么程度上恢复景物的具体外观和摄像机矩阵. 假设给定了图像之间的点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 的集合. 但假定该对应集来自一组未知的 3D 点 $\mathbf{X}_i$. 同样地,摄像机的位置、朝向和标定也未知. 重构的任务是求摄像机矩阵 P 和 P′,以及 3D 中的点 $\mathbf{X}_i$,使得对所有 $i$,有

$$\mathbf{x}_i = \mathrm{P}\mathbf{X}_i \qquad \mathbf{x}'_i = \mathrm{P}'\mathbf{X}_i$$

若给的点太少,这个任务就不可能完成. 然而,如果有足够多的点对应,就可能唯一地计算出基本矩阵,进而在相差一个射影变换的多义性下,该景物可以重构出来. 这是一个非常瞩目的结果,是未标定方法的主要成就之一.

如果提供有关摄像机或景物的额外信息,重构的多义性可以减少. 我们介绍一种两步近似方法,首先把多义性降至仿射,然后降至度量,每一步都要求提供适当类型的信息.

## 9.1　重构方法概述

下面,我们介绍由两视图重构的一种方法:

(1) 由点对应计算基本矩阵.

(2) 由基本矩阵计算摄像机矩阵.

(3) 针对每组对应点 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$,计算空间中映射到这两个图像点的点.

这种方法可能有多种版本. 比如,若摄像机是标定的,则可以用计算本质矩阵代替基本矩阵. 进一步,可以利用摄像机的运动,景物约束或摄像机部分标定的信息来获得有关重构的改善.

在下面几个段落中,我们将对这种重构方法的每一步作简要讨论. 所介绍的方法只不过是对重构的概念性描述. 我们要提醒读者仅仅基于本节所给出的概念性描述还不能实现一个重构,因为真实图像的测量是有"噪声"的,第 10 章和第 11 章将在本章一般概述的基础上介绍适用的重构方法.

**计算基本矩阵**　给定两幅图像中的点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 集合,基本矩阵 F 对所有的 $i$ 满足条件 $\mathbf{x}'_i{}^{\mathrm{T}}\mathrm{F}\mathbf{x}_i = 0$. 当 $\mathbf{x}_i$ 和 $\mathbf{x}'_i$ 已知时,这个方程是矩阵 F 的(未知)元素的线性方程. 事实上,每组点对应产生关于 F 元素的一个线性方程. 给定至少 8 组点对应,在相差一个尺度因子的意义下,可以线性求解 F 的元素(在 7 组点对应时有非线性求解方法). 当超过 8 个方程时,则求其最小二乘解. 这就是计算基本矩阵方法的一般原则.

由点对应集合计算基本矩阵的一种推荐的方法将在第 10 章中介绍.

**计算摄像机矩阵**　直接利用结论 8.14 的公式很容易计算出与基本矩阵 F 相对应的一对摄像机矩阵 P 和 P′.

**三角形法** 给定摄像机 P 和 P′，令 $\mathbf{x}$ 和 $\mathbf{x}'$ 为两幅图像上满足对极约束 $\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0$ 的两个点. 第 8 章指出这种约束可以借助对应于两个图像点的空间射线几何加以解释，具体地说，它表示点 $\mathbf{x}'$ 在对极线 $\mathrm{F}\mathbf{x}$ 上. 它进而又表示从图像点 $\mathbf{x}$ 和 $\mathbf{x}'$ 反向投影的射线在一个公共的对极平面上，即在一张过两个摄像机中心的平面上. 因为这两条射线在一张平面上，它们将相交于某点. 该交点 $\mathbf{X}$ 通过两个摄像机分别投影到两幅图像上的点 $\mathbf{x}$ 和 $\mathbf{x}'$，图 9.1 对此作了说明.

3D 空间中只有在介于两个摄像机之间的基线上的点不能由它们的图像确定. 在这种情形下，反向投影的射线共线(两条线都等同于基线)且沿着整个长度相交. 因此，点 $\mathbf{X}$ 不能唯一地确定. 基线上的点都投影到两幅图像上的对极点.

通过 $\mathbf{x}$ 和 $\mathbf{x}'$ 的两条反向投影的射线相交来实际确定 $\mathbf{X}$ 的稳定性的数值方法将在第 11 章介绍.

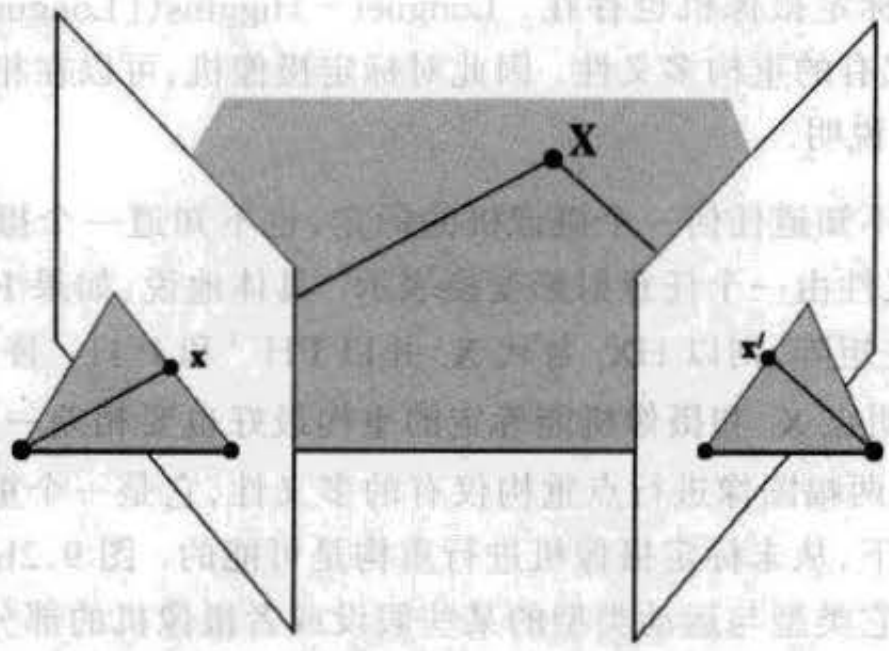



图 9.1 **三角形法**. 图像点 $\mathbf{x}$ 和 $\mathbf{x}'$ 反向投影成射线. 如果满足对极几何约束 $\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0$，那么这两条射线共面，从而相交于一个 3D 点 $\mathbf{X}$.

## 9.2 重构的多义性

在本节中，我们将讨论从点对应求景物重构所固有的多义性. 这个课题将在一般的背景下作讨论，不限定实现重构的具体方法.

若没有景物在某个 3D 坐标系下位置的信息，一般不可能从一对视图(或事实上从任意数目的视图)重构该景物的绝对位置和朝向. 这一结论的成立与任何有关摄像机内参数和它们的相对位置的信息毫不相干. 举一个例子，我们不能准确地计算图 8.8($p$167)中景物(或任何景物)的纬度和经度，也不能决定这个走廊的走向是南北还是东西. 可以这样说，景物重构的最好结果也要与世界坐标系相差一个欧氏变换(旋转和平移).

还有一个几乎是显然的事实是景物的整体尺寸是不能确定的. 再次以图 8.8($p$167)为例，仅仅根据图像不可能决定这个走廊的宽度. 它可以是 2 米、1 米. 它甚至可能是一个木偶小屋的像，其走廊的宽度是 10 厘米. 我们凭生活经验可以猜测从地板到天花板大约 3 米，在此基础上允许我们想象该场景的真正尺寸. 这个额外信息就是景物的辅助知识的一个例子，它不能从图像测量中得到. 所以，没有这类辅助知识，仅由图像来确定景物就要相

差一个相似变换(旋转、平移和缩放).

为了从数学上证明上述由观察得到的结论,令 $\mathbf{X}_i$ 为空间点集,P、P′为把 $\mathbf{X}_i$ 投影到图像点 $\mathbf{x}$ 和 $\mathbf{x}'$ 的一对摄像机. 该点集 $\mathbf{X}_i$ 和该摄像机对是从图像对应求景物重构的要素. 现在令

$$H_s = \begin{bmatrix} R & \mathbf{t} \\ \mathbf{0}^T & \lambda \end{bmatrix}$$

为任意的相似变换:R 是旋转,$\mathbf{t}$ 是平移,而 $\lambda^{-1}$ 表示整体尺度. 以 $H_s\mathbf{X}_i$ 替代 $\mathbf{X}_i$ 且分别以 $PH_s^{-1}$,$P'H_s^{-1}$ 替代摄像机 P,P′,这样做不会改变被观察的图像点,因为 $P\mathbf{X}_i = (PH_s^{-1})(H_s\mathbf{X}_i)$. 进一步,如果把 P 分解为 $P = K[R_P \mid \mathbf{t}_P]$,则可算出

$$PH_s^{-1} = K[R_P R^{-1} \mid \mathbf{t}'],$$

其中的 $\mathbf{t}'$ 并不需要准确地算出. 这个结果证明 P 乘以 $H_s^{-1}$ 不会改变它的标定矩阵. 因此这种重构多义性即使对标定摄像机也存在. Longuet - Higgins([LonguetHiggins - 81])证明,对标定摄像机,这是仅有的重构多义性. 因此对标定摄像机,可以在相差一个**相似变换**下重构. 图 9.2a 对此作了说明.

**射影多义性** 如果既不知道任何一个摄像机的标定,也不知道一个摄像机与另一个的相对位置,那么重构的多义性由一个任意射影变换表示. 具体地说,如果 H 是表示 $\mathbb{P}^3$ 空间射影变换的任意 4×4 可逆矩阵,则以 $H\mathbf{X}_i$ 替代 $\mathbf{X}_i$ 并以 $PH^{-1}$ 和 $P'H^{-1}$ 替代 P 和 P′(同前一段)不改变图像点. 这证明点 $\mathbf{X}_i$ 和摄像机能确定的重构最好也要相差一个射影变换. 本章(9.3 节)将证明这也是从两幅图像进行点重构仅有的多义性,它是一个重要的结论. 这样一来在相差一个**射影变换**下,从未标定摄像机进行重构是可能的. 图 9.2b 对此作了说明.

重构多义性的其它类型与运动类型的某些假设或者摄像机的部分知识有关,比如:

(1) 如果两个摄像机之间有平移运动,但标定不变,则可在相差一个仿射变换下重构.

(2) 如果两个摄像机除了焦距外都已标定,则仍然可在相差一个相似变换下重构.

这两种情况将分别在后面的节 9.4.1 和例 18.8($p$334)中讨论.

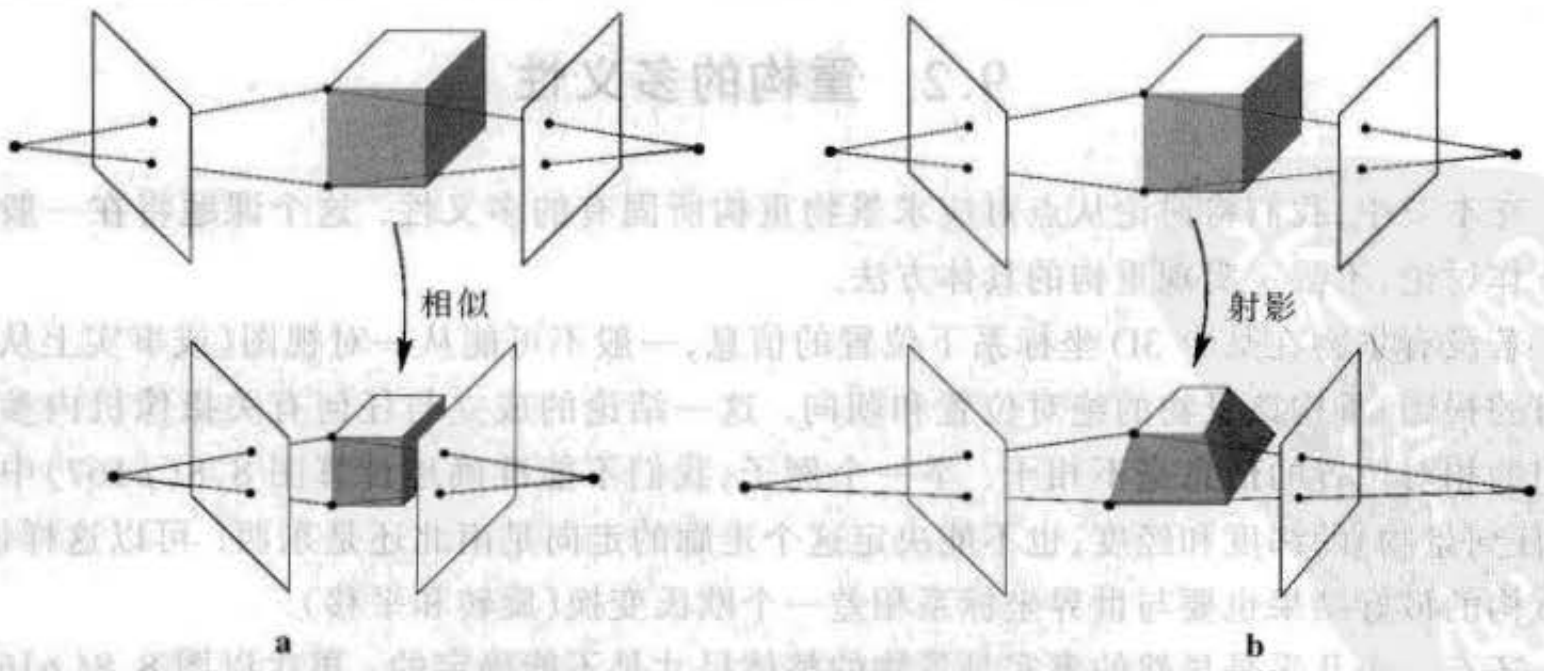



图 9.2 **重构多义性**. (a) 如果摄像机已被标定,那么任何重构必须不改变图像中被测量的射线的夹角. 结构和摄像机位置的相似变换不改变被测量的角度. 射线和基线(对极点)之间的角度同样不变. (b) 如果摄像机是未标定的,那么重构只须不改变图像点(射线和图像平面的交). 结构和摄像机位置的射影变换不改变被测量的点,但可改变射线之间的角度. 对极点(与基线的交点)也不改变.

**术语** 在根据真实数据(由点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 组成)推导的任何重构问题中,存在一个**真正**的重构,它由实际的点 $\overline{\mathbf{X}}_i$ 和产生观察测量点的实际摄像机 $\overline{\mathrm{P}}$、$\overline{\mathrm{P}}'$ 组成. 重构得到的点集 $\mathbf{X}_i$ 和摄像机与真正的重构相差一个给定的类或群(例如,一个相似、射影或者仿射变换)的变换. 我们用**射影重构**、**仿射重构**、**相似重构**等等来表示涉及到的变换类型. 然而,通常倾向于正式地用术语**度量重构**来代替定义上相同的**相似重构**. 这个术语表示诸如直线之间的角度和长度比等的度量性质可以在重构上量出并且具有它们的真值(因为它们是相似不变量). 另外,在出版物中经常用到术语**欧氏重构**,其实也表示与相似或度量重构同样的东西,因为在没有额外信息时真正的欧氏重构(包括确定整体尺度)是不可能的.

## 9.3 射影重构定理

本节将证明由未标定摄像机进行射影重构的基本定理. 这个定理可以非正式地陈述如下:

- **如果两幅视图的一个点对应集唯一地确定了基本矩阵,则景物和摄像机可以仅由这些对应重构,而且由这些对应产生的任何两个重构都是射影等价的.**

在两个摄像机中心连线上的点必须除外,因为即使摄像机矩阵确定后这些点也不能唯一地重构. 结论的正式陈述如下:

**定理 9.1(射影重构定理)** **假定 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 是两幅图像点之间对应的集合而且基本矩阵 F 由条件 $\mathbf{x}_i'^{\mathrm{T}}\mathrm{F}\mathbf{x}_i = 0$(对所有 $i$ 成立)唯一地确定. 令 $(\mathrm{P}_1, \mathrm{P}_1', \{\mathbf{X}_{1i}\})$ 和 $(\mathrm{P}_2, \mathrm{P}_2', \{\mathbf{X}_{2i}\})$ 为对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 的两个重构,则存在一个非奇异矩阵 H 使得 $\mathrm{P}_2 = \mathrm{P}_1\mathrm{H}^{-1}$, $\mathrm{P}_2' = \mathrm{P}_1'\mathrm{H}^{-1}$ 并且除了使得 $\mathrm{F}\mathbf{x}_i = \mathrm{F}^{\mathrm{T}}\mathbf{x}'_i = \mathbf{0}$ 的那些 $i$ 外,对其余每个 $i$ 都有 $\mathbf{X}_{2i} = \mathrm{H}\mathbf{X}_{1i}$.**

**证明** 既然基本矩阵由点对应唯一地确定,可以推知 F 是既对应于摄像机对 $(\mathrm{P}_1, \mathrm{P}_1')$ 也对应于摄像机对 $(\mathrm{P}_2, \mathrm{P}_2')$ 的基本矩阵. 根据定理 8.10($p$172),存在着一个射影变换 H 使得 $\mathrm{P}_2 = \mathrm{P}_1\mathrm{H}^{-1}$ 和 $\mathrm{P}_2' = \mathrm{P}_1'\mathrm{H}^{-1}$,这正是所需要证明的.

对点而言,我们观察到 $\mathrm{P}_2(\mathrm{H}\mathbf{X}_{1i}) = \mathrm{P}_1\mathrm{H}^{-1}\mathrm{H}\mathbf{X}_{1i} = \mathrm{P}_1\mathbf{X}_{1i} = \mathbf{x}_i$. 另一方面 $\mathrm{P}_2\mathbf{X}_{2i} = \mathbf{x}_i$,因此 $\mathrm{P}_2(\mathrm{H}\mathbf{X}_{1i}) = \mathrm{P}_2\mathbf{X}_{2i}$. 所以 $\mathrm{H}\mathbf{X}_{1i}$ 和 $\mathbf{X}_{2i}$ 都被摄像机矩阵 $\mathrm{P}_2$ 映射为同一点 $\mathbf{x}_i$,由此推出 $\mathrm{H}\mathbf{X}_{1i}$ 和 $\mathbf{X}_{2i}$ 在过摄像机 $\mathrm{P}_2$ 中心的同一射线上. 同理,可以推出这两点也在过摄像机 $\mathrm{P}_2'$ 中心的射线上. 这里存在两种可能:要么如所需要证明的一样 $\mathbf{X}_{2i} = \mathrm{H}\mathbf{X}_{1i}$,要么它们是两个摄像机中心连线上的不同点. 在后一种情形,图像点 $\mathbf{x}_i$ 和 $\mathbf{x}'_i$ 与两幅图像上的对极点重合,因此 $\mathrm{F}\mathbf{x}_i = \mathrm{F}^{\mathrm{T}}\mathbf{x}'_i = \mathbf{0}$. □

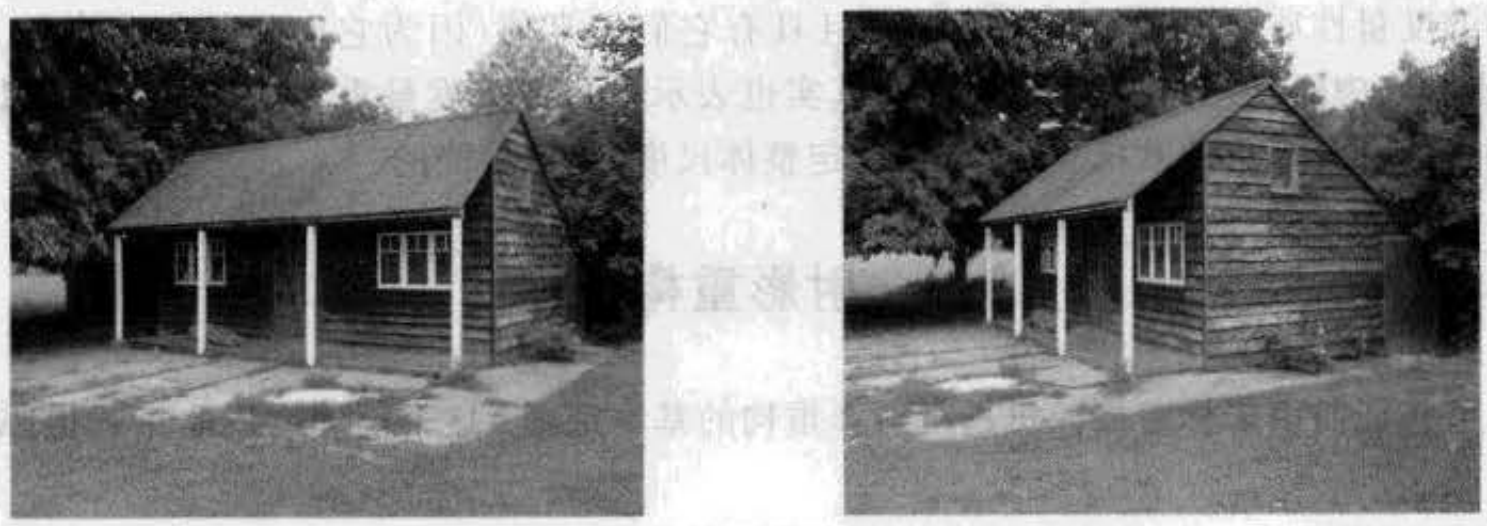
这是一个意义非常重大的结果,因为它指出仅仅根据图像的对应便可以从两幅视图计算出景物的一个射影重构,无需知道任何有关两个摄像机的标定和姿态的信息. 特别是真实的重构在射影重构的一个射影变换之中. 图 9.3 给出了从两幅图像的部分射影重构计算 3D 结构的一个例子.

更为具体地说,我们假设真正的欧氏重构是 $(\mathrm{P}_{\mathrm{E}}, \mathrm{P}_{\mathrm{E}}', \{\mathbf{X}_{\mathrm{E}i}\})$ 而射影重构是 $(\mathrm{P}, \mathrm{P}', \{\mathbf{X}_i\})$,则这些重构由一个非奇异矩阵 H 联系,使得

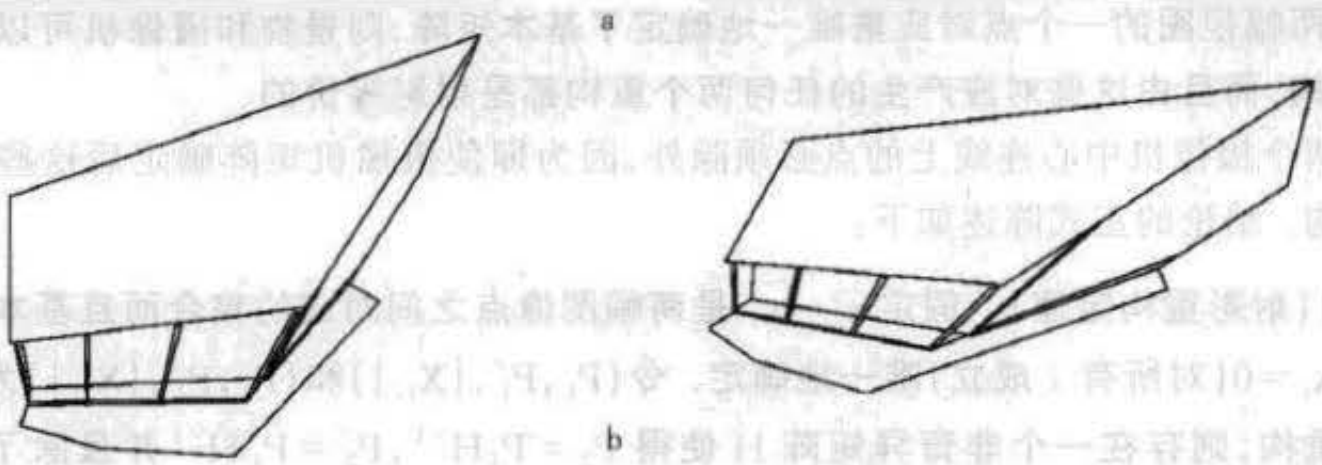
$$\mathrm{P}_{\mathrm{E}} = \mathrm{P}\mathrm{H}^{-1}, \mathrm{P}_{\mathrm{E}}' = \mathrm{P}'\mathrm{H}^{-1}, \text{和 } \mathbf{X}_{\mathrm{E}i} = \mathrm{H}\mathbf{X}_i \tag{9.1}$$

其中 H 是一个 4×4 未知的单应矩阵并对所有点都一样.

对某些应用来说,射影重构就是全部的目的. 例如,诸如"直线和平面相交于什么点?","由特殊的曲面,例如平面或二次曲面,诱导的两幅视图之间的映射是什么?"等一类问题都可以直接由射影重构来处理. 后面还将进一步看到射影重构是迈向仿射或度量重构的第一步.



a



b

图 9.3 **射影重构**. (a)原图像对. (b)该景物的一个 3D 重构的 2 幅视图. 重构不需要摄像机矩阵或景物几何的信息. 基本矩阵 F 由图像之间的点对应计算得到,摄像机矩阵则由 F 恢复,然后 3D 点用三角形法根据对应计算得到. 线框图的线把计算得到的 3D 点连起来.

## 9.4 分层重构

分层重构方法始于射影重构,然后如有可能,再逐步地将它改善到仿射或最终到度量重构. 当然,如刚才讲过的那样,没有关于景物、运动或者摄像机标定的进一步信息,仿射和度量重构是不可能的.

### 9.4.1 步入仿射重构

仿射重构的本质是用某些方法定位无穷远平面,因为定位无穷远平面等价于仿射重构. 这种等价性的 2D 情形在 1.7 节($p20$)解释过. 为了了解重构的这种等价性,假设我们已经确定了一个景物的射影重构,它由三元组$(\mathrm{P},\mathrm{P}',\{\mathbf{X}_i\})$组成. 进一步假设借助某些办法某一平面 $\boldsymbol{\pi}$ 被确定为真实的无穷远平面. 在射影重构的坐标系下平面 $\boldsymbol{\pi}$ 表示为一个 4 维矢量. 在真实的重构下 $\boldsymbol{\pi}$ 坐标是$(0,0,0,1)^{\mathrm{T}}$,我们可以寻找使 $\boldsymbol{\pi}$ 映射到$(0,0,0,1)^{\mathrm{T}}$ 的射影变换. 考虑到射影变换对平面作用的方式,我们需要求 H 使得 $\mathrm{H}^{-\mathrm{T}}\boldsymbol{\pi}=(0,0,0,1)^{\mathrm{T}}$. 这样的变换就

是：

$$H = \begin{bmatrix} I & \mathbf{0} \\ \boldsymbol{\pi}^T & \end{bmatrix}. \tag{9.2}$$

事实上，可以直接验证 $H^T(0,0,0,1)^T = \boldsymbol{\pi}$，因此 $H^{-T}\boldsymbol{\pi} = (0,0,0,1)^T$ 就是我们所需的．现在把这个变换 H 作用于所有点和两个摄像机上．然而应注意，如果 $\boldsymbol{\pi}^T$ 的最后一个坐标是 0，则这个公式无效．在这种情形下，可以把 $H^{-T}$ 作为 Householder 矩阵来计算（A3．2－$p$406）并由此计算出使得 $H^{-T}\boldsymbol{\pi} = (0,0,0,1)^T$ 的合适的 H．．

到目前为止，我们得到的重构不一定是真正重构——我们仅知道无穷远平面已被正确定位．现有的重构与真实的重构相差一个以无穷远平面为不动平面的射影变换．然而，根据结论 2.7（$p$46），一个固定无穷远平面的射影变换是仿射变换．因此这个重构与真实重构相差一个仿射变换——即它是一个**仿射重构**．

对某些应用来说，仿射重构已经足够了．例如，现在可以计算两个点的中点和一个点集的形心，可以构造与其他直线和平面平行的直线．显然这些计算不可能从射影重构中得到．

我们讲过，除非增加额外信息否则无穷远平面不可能被辨认．我们将给出若干例子说明能够辨认无穷远平面的信息类型．

**平移运动**

摄像机作纯平移运动时，可以由两幅视图实现仿射重构．理解这一事实的一个简单办法是以下容易观察到的事实：无穷远平面上一点 $\mathbf{X}$ 在由摄像机平移所产生的两幅图像上的像是同一点．这也是常识，当一个人沿一条直线运动（比如在一条直路上行驶的汽车中）时，距离很远的物体（比如月亮）好象固定不动——而只有附近的物体在视野中快速运动．因此，可以发现许多匹配点 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$，它们在一幅图像上的点与另一幅图像上的同点对应．注意没有必要在两幅图像中实际寻找这样的对应点——任何点及其在另一幅图像上的相同点就行了．给定一个射影重构，可以重构对应于匹配 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$ 的点 $\mathbf{X}_i$．点 $\mathbf{X}_i$ 将在无穷远平面上．从三组这样的点我们可以得到在无穷远平面上的三个点——足够唯一地把它确定．

虽然以上讨论给出了通过平移摄像机作仿射重构的一个构造性证明．这并不意味着在数值上这是最好的实施办法．事实上，平移运动的假定给出基本矩阵一个非常限制性的形式——如节 8.3.1 所指出，它是一个反对称矩阵．求解基本矩阵时，应该考虑这一特殊形式．

**结论 9.2　假设摄像机的运动是没有旋转且内参数保持不变的一个纯平移运动，则如例** 8.6（$p$166）**所示，**$F = [\mathbf{e}]_\times = [\mathbf{e}']_\times$**，并且作为仿射重构可选的两个摄像机为** $P = [I|\mathbf{0}]$ **和** $P' = [I|\mathbf{e}']$．

**景物约束**　景物约束或条件也可以用来获得仿射重构，只要能够辨认在无穷远平面上的三个点，该平面就可以辨认，并且该重构可以变换到仿射重构．

**平行直线**　最明显的条件是知道某些 3D 直线实际是平行的．空间两条平行直线的交点给出无穷远平面上的一个点，该交点的图像是这对平行线的消影点，而且是两条像直线的交点．假设在景物中能够确定三对平行直线．每对相交于无穷远平面上的一点，如果每对有

不同的方向，则三点将不同．因为三点确定一张平面，故这个信息足够确定平面 $\pi$.

实际地来计算空间直线的交点是个棘手的问题，由于噪声的存在，本应该相交的直线很少相交．这将在第 11 章中作较详细地讨论，计算无穷远平面的正确的数值方法将在第 12 章给出．图 9.4 给出从三组平行直线计算仿射重构的例子．

注意没有必要在两幅图像上同时求消影点．假定从第一幅图像上的平行直线计算出消影点 $\mathbf{v}$，且 $\mathbf{l}'$ 是它在第二幅图像上对应的直线．消影点满足对极约束，因此在第二幅图像上的消影点 $\mathbf{v}'$ 可以由 $\mathbf{l}'$ 和 $\mathbf{v}$ 的对极线 $\mathrm{F}\mathbf{v}$ 的交点来计算．3D 点 $\mathbf{X}$ 重构的简洁的代数表示为方程 $([\mathbf{v}]_\times \mathrm{P})\mathbf{X}=\mathbf{0}$ 和 $(\mathbf{l}'^{\mathrm{T}}\mathrm{P}')\mathbf{X}=\mathbf{0}$ 的解．这些方程表示 $\mathbf{X}$ 映射到第一幅图像上的点 $\mathbf{v}$ 和第二幅图像的 $\mathbf{l}'$ 上的一点．

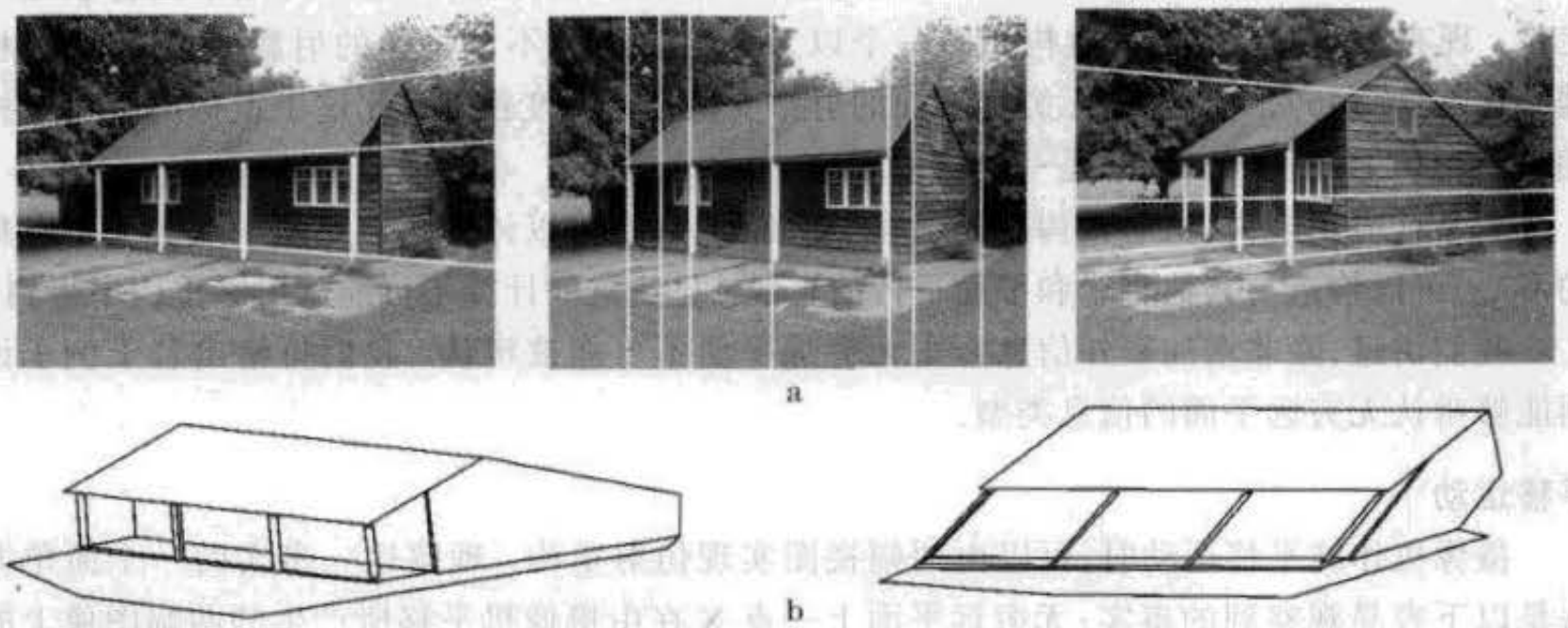



图 9.4 **仿射重构**．图 9.3 的射影重构可以用平行的景物直线升级为仿射重构．(a)在景物中有 3 组平行线，每组有不同的方向．这 3 组平行线可以使无穷远平面 $\pi_\infty$ 的位置从射影重构中计算出来．图 9.3 线框图给出的射影重构用单应(9.2)矫正为仿射重构．(b)显示线框图仿射重构的两个正交视图．注意平行景物线在重构中是平行的，但垂直景物线在重构中不垂直．

**直线上的距离比率** 除了用图像上平行直线的交点计算消影点外，另一个方法是运用景物中仿射长度比率不变的知识．例如，在直线上给定两个已知长度比的线段，就可以决定该直线上的无穷远点．这意味着从一条已知世界长度比率(如三个等距离点)的直线的像，就可以决定其消影点．这种计算和其它计算消影点和消影直线的方法，曾在 1.7 节($p20$)中给出．

**无穷单应**

一旦无穷远平面被定位，我们就能得到仿射重构，而且我们同时得到一种称作“无穷单应”的图像到图像的映射，这种 2D 单应映射将在第 12 章作更详细地介绍．简要地说，它是一种映射，按如下方式把点通过无穷远平面从 P 图像转移到 P′图像：延长过点 $\mathbf{x}$ 的射线与无穷远平面相交于点 $\mathbf{X}$；该点投影到另一幅图像上的一点 $\mathbf{x}'$，从 $\mathbf{x}$ 到 $\mathbf{x}'$ 的单应记作 $\mathbf{x}'=\mathrm{H}_\infty\mathbf{x}$.

现在，我们证明得到了仿射重构就等于知道了无穷单应．已知两个摄像机 $\mathrm{P}=[\mathrm{M}|\mathbf{m}]$ 和 $\mathrm{P}'=[\mathrm{M}'|\mathbf{m}']$ 的一个仿射重构，无穷单应便由 $\mathrm{H}_\infty=\mathrm{M}'\mathrm{M}^{-1}$ 给出．这是因为在无穷远平面上的点 $\mathbf{X}=(\tilde{\mathbf{X}}^{\mathrm{T}},0)^{\mathrm{T}}$，在一幅图像中映射成 $\mathbf{x}=\mathrm{M}\tilde{\mathbf{X}}$，在另一幅图像中映射成 $\mathbf{x}'=\mathrm{M}'\tilde{\mathbf{X}}$．因此对

$\pi_\infty$ 上的点有 $\mathbf{x}' = M'M^{-1}\mathbf{x}$. 进一步可以验证它在摄像机的 3D 仿射变换下不变. 因此,无穷单应可以从仿射重构显式地计算出来并且反过来也成立:

**结论 9.3 如果一个仿射重构已得到并且其中的摄像机矩阵是 $P=[I|\mathbf{0}]$ 和 $P'=[M'|\mathbf{e}']$,则无穷单应是 $H_\infty = M'$. 反过来,如果已得到无穷单应 $H_\infty$,则仿射重构的摄像机可以取为 $P=[I|\mathbf{0}]$ 和 $P'=[H_\infty|\mathbf{e}']$.**

无穷单应可以直接从图像对应元素而不必间接地由仿射重构来计算. 例如,$H_\infty$ 可以由对应的三个消影点和 F 来计算或者由一条消影直线与消影点对应和 F 来计算. 第 12 章将给出这些计算的正确的数值过程. 然而,这种直接计算完全等价于在射影重构中确定 $\pi_\infty$.

**其中一个是仿射摄像机的情形**

仿射重构的另一个重要情形是已知其中一个摄像机(按节 5.3.1 ($p$111)的定义)是仿射摄像机. 为了说明在这种情形下仿射重构一定是可能的,可参考节 5.3.5 ($p$117),那里证明了仿射摄像机的主平面是无穷远平面. 因而为了将射影重构变换为仿射重构,只需要找到仿射摄像机的主平面并且把它映射到平面 $(0,0,0,1)^T$. 回忆(节 5.2 ($p$104))摄像机的主平面是相应摄像机矩阵的第三行. 例如,考虑摄像机矩阵为 $P=[I|\mathbf{0}]$ 和 $P'$ 的一个射影重构并设第一个为仿射摄像机. 为了把 P 的第三行映射到 $(0,0,0,1)^T$,只要交换这两个摄像机矩阵最后两列,并同时交换每一个 $\mathbf{X}_i$ 的第三和第四个坐标即可. 这是对应于一个置换矩阵 H 的射影变换,它表明:

**结论 9.4 若 $(P,P',\{\mathbf{X}_i\})$ 是从一点对应集得到的射影重构且 $P=[I|\mathbf{0}]$. 若已知 P 事实上是一个仿射摄像机,则可以通过交换 P 和 $P'$ 的最后两个列和每一个 $\mathbf{X}_i$ 的最后两个坐标而得到一个仿射重构.**

注意一个摄像机是仿射的条件不对基本矩阵产生约束,因为任何一对规范摄像机 $P=[I|\mathbf{0}]$ 和 $P'$ 都可以变换为 P 是仿射的摄像机对. 如果已知两个摄像机都是仿射,则我们将会看到基本矩阵具有(13.1 - $p$245)给出的约束形式. 对于这种情形,为了数值计算的稳定性,求解基本矩阵时必须强制性要求基本矩阵取这个特殊形式.

当然真实的仿射摄像机是不存在的——仿射摄像机模型不过是一种近似,只是当图像上看到的点的深度变化相对于它到摄像机的距离比较小时才有效. 尽管如此,仿射摄像机的假定有利于对从射影到仿射重构施加重要的约束.

### 9.4.2 步入度量重构

仿射重构的关键是辨认无穷远平面,而度量重构的关键是辨认绝对二次曲线(节 2.6 ($p$47)). 因为绝对二次曲线 $\Omega_\infty$ 是无穷远平面上的一条平面二次曲线,确定绝对二次曲线就意味着确定无穷远平面.

按照分层的方法,我们的过程是从射影重构到仿射重构再到度量重构. 因此在求绝对二次曲线之前我们已经知道了无穷远平面. 假设在无穷远平面上的绝对二次曲线已得到辨认. 则下一步主要是对该仿射重构运用仿射变换使得被辨认的绝对二次曲线映射到标准欧氏坐标系下的绝对二次曲线(在 $\pi_\infty$ 上,它的方程是 $X_1^2+X_2^2+X_3^2=0$). 所求重构与真实的重构之间相差一个使绝对二次曲线不变的射影变换. 由结论 2.9($p$47)可知该射影变换是一个

相似变换，我们因此而达到了度量重构的目的.

实际上，实现这个过程的最简单的方法是考虑其中一幅图像中的绝对二次曲线的像. 绝对二次曲线(如同任意二次曲线)的像是图像上的二次曲线. 二次曲线的反向投影是一个锥面，该锥面与无穷远平面相交于一条二次曲线，由此定义了绝对二次曲线. 记住绝对二次曲线的像是图像本身的性质，与任何图像点、直线和其它特征一样，是不依赖于任何具体的重构，因而在重构的 3D 变换下是不变的.

假设在仿射重构下，由矩阵为 $P=[M|\mathbf{m}]$ 的摄像机所摄取的绝对二次曲线的像是二次曲线 $\omega$. 我们将说明怎样用 $\omega$ 来定义把仿射重构变换到度量重构的单应 H.

**结论 9.5　若已知绝对二次曲线在某幅图像中的像是 $\omega$，且已知仿射重构中的摄像机矩阵是 $P=[M|\mathbf{m}]$，则用形如**

$$H=\begin{bmatrix} A^{-1} & \\ & 1 \end{bmatrix}$$

**的 3D 变换，可以把仿射重构变换到度量重构，其中 A 由方程 $AA^T=(M^T\omega M)^{-1}$ 的 Cholesky 分解得到.**

**证明**　在变换 H 下，摄像机矩阵 P 被变换为矩阵 $P_M=PH^{-1}=[M_M|\mathbf{m}_M]$，如果 $H^{-1}$ 形如

$$H^{-1}=\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix},$$

则 $M_M=MA$. 然而，绝对二次曲面的像与欧氏坐标系下摄像机矩阵 $P_M$ 的关系是

$$\omega^*=M_M M_M^T.$$

这是因为摄像机矩阵可以分解为 $M_M=KR$，且从(7.9－$p$146)可知 $\omega^*=\omega^{-1}=KK^T$，此式与 $M_M=MA$ 联合推出 $\omega^{-1}=MAA^TM^T$，整理后得到 $AA^T=(M^T\omega M)^{-1}$. A 满足这个关系的一个具体值可以通过 $(M^T\omega M)^{-1}$ 的 Cholesky 分解得到. 后一个矩阵要求是正定的(见结论 A3.5($p$407))，否则，不存在这样的矩阵 A，从而度量重构将不可能. □

这种度量重构的方法依赖于辨认绝对二次曲线的像. 实现它有多种方式，下面将就此进行讨论. 我们给出三种绝对二次曲线的图像的约束源，实践中常把这些约束组合起来使用.

**1. 景物正交性的约束**　景物的正交直线的一对消影点 $\mathbf{v}_1,\mathbf{v}_2$ 对 $\omega$ 提供一个线性约束：

$$\mathbf{v}_1^T\omega\mathbf{v}_2=0.$$

类似地，来自一个方向的消影点 $\mathbf{v}$ 和与之垂直的平面的消影直线 $\mathbf{l}$ 对 $\omega$ 提供了两个约束

$$\mathbf{l}=\omega\mathbf{v}.$$

一个常用的例子是垂直方向的消影点和水平地平面的消影直线.

**2. 已知内参数的约束**　如果摄像机矩阵的标定矩阵等于 K，则绝对二次曲线的像是 $\omega=K^{-T}K^{-1}$. 因此包括在 K 中的摄像机内参数信息可以用来约束或确定 $\omega$ 的元素. 在已知 K 的扭曲因子为零($s=0$)时有

$$\omega_{12}=\omega_{21}=0$$

若像素还是正方形(零扭曲参数且 $\alpha_x=\alpha_y$)，则

$$\omega_{11} = \omega_{22}.$$

节 7.7($p$153)针对单幅视图的标定问题详细讨论了上面的两类约束,它还提供了仅仅利用这个信息来标定一个摄像机的例子. 下面还有来自多幅视图的另一类约束源.

**3. 同一摄像机在多幅图像中的约束** 绝对二次曲线的一个性质是它在图像上的投影只依赖于摄像机的标定矩阵,而不依赖于摄像机的位置和方向. 当两个摄像机具有相同的标定矩阵时(通常意味着同一摄像机在不同位置拍摄的两幅图像)则有 $\omega = \omega'$,即绝对二次曲线在两幅图像上的像是相同的. 给定足够多的图像,可以利用这个性质从仿射重构得到度量重构. 这种度量重构的方法及其在摄像机自标定中的应用,将在第 18 章中作更详细地讨论. 目前我们只给出一般原理.

因为绝对二次曲线在无穷远平面上,它的像可以通过无穷单应从一幅视图转移到另一幅视图. 由此得到方程(见结论 1.13($p$11))

$$\omega' = H_\infty^{-T}\omega H_\infty^{-1} \tag{9.3}$$

其中 $\omega$ 和 $\omega'$ 是 $\Omega_\infty$ 在两幅视图中的像. 形成这些方程的必要条件是已经得到仿射重构,因为必须已知无穷单应. 如果 $\omega = \omega'$,则(9.3)给出一组关于 $\omega$ 元素的方程. 一般情况下,这一组线性方程对 $\omega$ 提供了 4 个约束,因 $\omega$ 有 5 个自由度,故还不能完全被确定. 然而,把这些线性方程和由景物正交性或者已知内参数提供的约束结合起来,就可以唯一地确定 $\omega$. 事实上(9.3)可以用来把 $\omega$ 的约束转移到 $\omega'$ 的约束. 图 9.5 显示了以组合约束的方式得到的度量重构的一个例子.

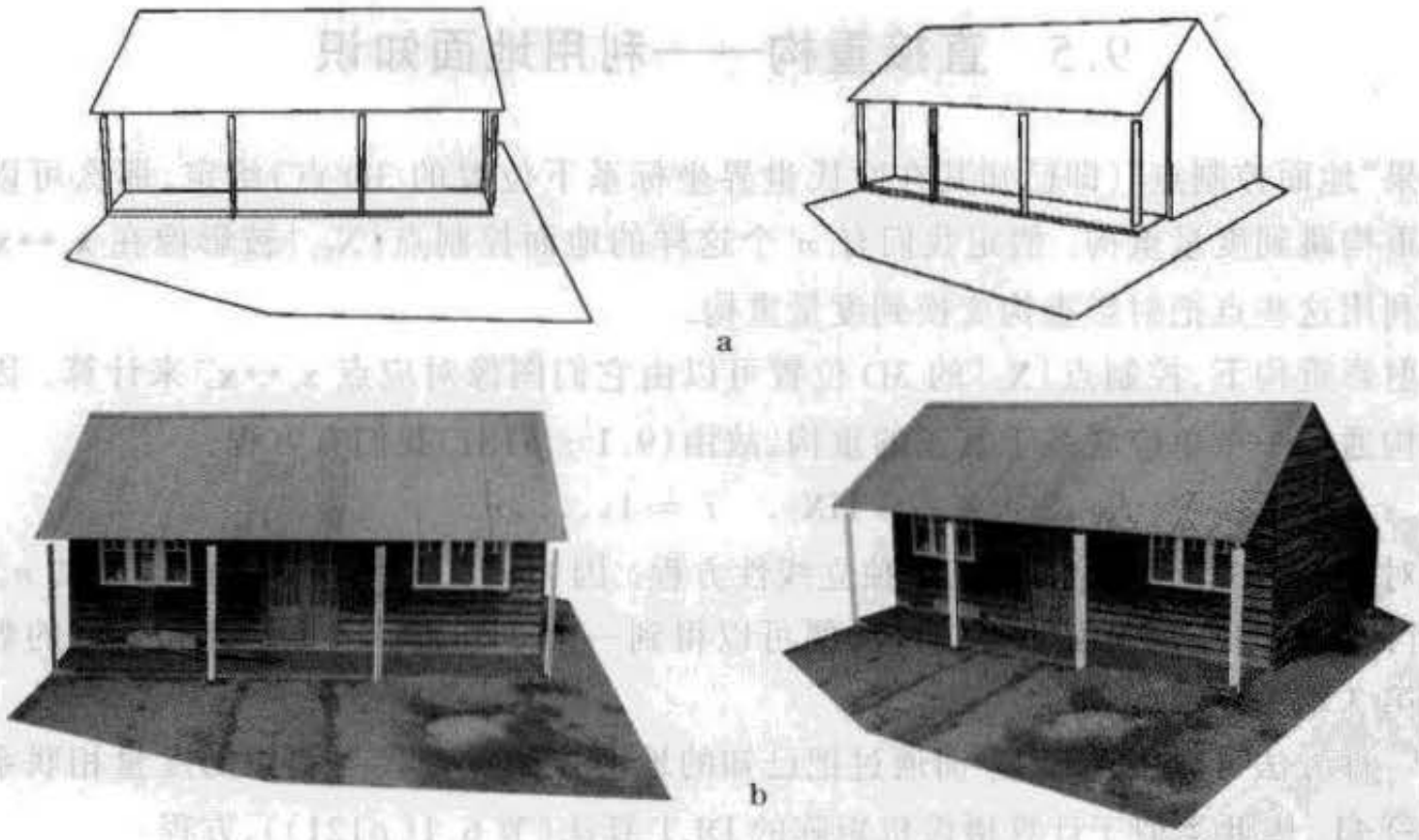



图 9.5 **度量重构**. 用绝对二次曲线的像使图 9.4 的仿射重构提升为度量重构. 采用的信息是图 9.4 显示的平行线集合的方向的正交性以及两幅图像具有正方形象素. $H_\infty$ 把正方形象素的约束由一幅图像传递到另一幅图像. (a)度量重构的两幅视图. 景物中正交的直线在重构中仍正交而且房子的边缘的长宽比与真实的一样. (b)由线框视图通过分片平面模型的纹理映射得到的两幅视图.

### 9.4.3 利用 $\omega$ 直接度量重构

前面的讨论给出如何利用图像上的绝对二次曲线(IAC)的信息把仿射重构变换为度量重构. 然而,知道 $\omega$ 并至少给定两幅视图时,也可以直接进行度量重构. 至少有两种不同的方法来实现这种重构. 最明显的方法是利用 IAC 来计算每个摄像机的标定,然后再进行已知标定下的重构.

这种方法根据 $\omega$ 和标定矩阵 K 的关系,即 $\omega=(KK^T)^{-1}$. 因此可以从 $\omega$ 计算 K,即先对它取逆,然后作 Cholesky 分解求 K. 如果知道每幅图像中的 IAC,则两个摄像机都可以用这个方法来标定. 下一步对标定的摄像机可以利用本质矩阵来计算景物的度量重构,如节 8.6 所示. 注意可能出现四个解,其中两个解只是镜像图像,而另外两个解不同并组成一个扭转对(但是考虑了点在摄像机前后位置,除一个解外其他解都可以排除.).

度量重构的另一个更概念性的方法是运用 IAC 的信息直接确定无穷远平面和绝对二次曲线. 如果已知在射影坐标系中的摄像机矩阵 P 和 P′和在每幅图像上的一条二次曲线(特别是绝对二次曲线的像),则 3D 空间中的 $\Omega_\infty$ 可以显式地算出. 它可以通过把二次曲线反向投影成锥面而得到,该锥面必然与绝对二次曲线相交. 因此,可以确定 $\Omega_\infty$ 和它的支撑平面 $\boldsymbol{\pi}_\infty$(有关的代数解见 $p242$ 的习题(10)). 然而,在一般情况下两个锥相交于属于不同的支撑平面的两条不同的平面二次曲线. 这样一来绝对二次曲面有两个可能的解,我们可以认为它属于组成扭转多义性对的两种不同的重构.

## 9.5 直接重构——利用地面知识

如果"地面控制点"(即已知其在欧氏世界坐标系下位置的 3D 点)给定,那么可以直接从射影重构跳到度量重构. 假定我们有 $n$ 个这样的地面控制点 $\{X_{Ei}\}$ 被影像在 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$. 我们希望利用这些点把射影重构变换到度量重构.

在射影重构下,控制点 $\{\mathbf{X}_i\}$ 的 3D 位置可以由它们图像对应点 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 来计算. 因为该射影重构通过一个单应联系于真正的重构,故由(9.1 - $p181$)我们有方程:

$$\mathbf{X}_{Ei} = H\mathbf{X}_i, \quad i = 1, \ldots, n.$$

每组点对应提供关于 H 元素的三个独立线性方程. 因 H 有 15 个自由度,故若提供 $n \geq 5$ 组(且没有四个控制点是共面的)点对应,便可以得到一个线性解. 这类计算和典型的数值方法已在第 3 章中介绍.

另一种方法可以不计算 $\mathbf{X}_i$ 而通过把已知的地面控制点直接与图像的度量相联系的途径来计算 H. 因此类似于计算摄像机矩阵的 DLT 算法(节 6.1($p121$)),方程

$$\mathbf{x}_i = PH^{-1}\mathbf{X}_{Ei}$$

提供了未知的 $H^{-1}$ 的元素的两个线性独立方程,而所有其它的量都是已知的. 如果 $\mathbf{x}'_i$ 已知,类似的方程可以由另一幅图像推出. 地面控制点不必在两幅图像中都可见. 仍需注意即使对某个给定点 $\mathbf{X}_{Ei}$,$\mathbf{x}_i$ 和 $\mathbf{x}'_i$ 都可见,但由于 $\mathbf{x}$ 和 $\mathbf{x}'$ 共面的约束,用这种方法产生的 4 个方程只有 3 个是独立的.

一旦 H 被计算出来,就可以把射影重构的摄像机矩阵 P、P′变换为其真实欧氏重构的摄

像机矩阵. 利用这种直接计算度量结构的一个例子在图 9.6 给出.

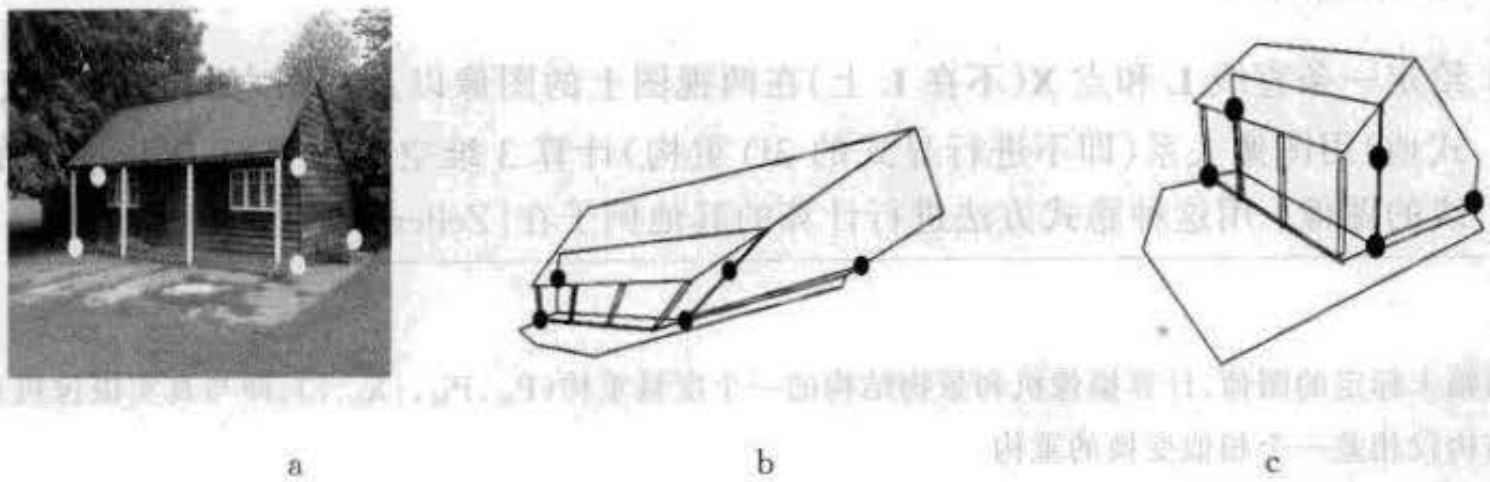



图 9.6 **直接重构**. 图 9.3 的射影重构可以通过给出 5 个(或更多)世界点的位置提升为度量重构:(a)被采用的 5 个点;(b)在 图 9.3 的射影重构上对应的点. (c)把 5 个点映射到它们的世界位置后的重构.

## 9.6 结束语

在本章,我们纵览了由两幅图像产生一个度量重构所必需的步骤并概括在算法 9.1 中,而这些步骤的计算过程将在下面几章中介绍. 一般的讨论仍主要限制在点上,但有关概念(三角形法,多义性,分层)同样适用于其他图像几何元素,例如直线、二次曲线等.

我们已经看到为了度量重构必须辨认射影坐标系下的两个几何元素,它们是无穷远平面 $\pi_\infty$(用于仿射)和绝对二次曲线 $\Omega_\infty$(用于度量). 反之,给定了 F 和一对标定过的摄像机,可以在 3 维空间中显式地计算 $\pi_\infty$ 和 $\Omega_\infty$. 这两个元素都有一个基于图像的对应元素:辨认无穷单应 $H_\infty$ 等价于辨认 3 维空间的 $\pi_\infty$;而辨认在每幅图像的绝对二次曲线的图像 $\omega$ 则等价于辨认 3 维空间的 $\pi_\infty$ 和 $\Omega_\infty$. 这种等价性在表 9.1 中概括.

最后,值得注意的是如果对度量精度没有严格要求,那么若能猜测出一个逼近正确值的内部参数值则一个可接受的度量重构一般地可直接从射影获得. 这样的一种"准欧氏重构"经常适用于可视化的目的.

### 9.6.1 文献

Koenderink 和 van Doorn[Koenderink－91]对仿射摄像机分层的有关讨论非常精彩. [Faugeras－95a]把它扩展到透视并附上 Luong 和 Vieville[Luong－94,Luong－96]的进展. 给定 F 后射影重构的可能性在[Faugeras－95b]和 Hartley 等的工作[Hartley－92c]中给出.

由纯平移计算仿射重构的方法首先出现于 Moons 等的工作[Moons－94]. 在多幅视图上把景物和内部参数结合起来的介绍在[Faugeras－95b,Liebowitz－99b]中给出.

| 给定的图像信息 | 视图关系和射影物体 | 3 维空间物体 | 重构多义性 |
|---|---|---|---|
| 点对应 | F | | 射影 |
| 包括消影点的点对应 | $F, H_\infty$ | $\pi_\infty$ | 仿射 |
| 点对应和摄像机内参数标定 | $F, H_\infty, \omega, \omega'$ | $\pi_\infty, \Omega_\infty$ | 度量 |

表 9.1 各种类型的重构多义性中的两视图关系、图像元素及其 3 维空间的对应物体.

### 9.6.2 注释和练习

(1) 给定一条直线 **L** 和点 **X**(不在 **L** 上)在两视图上的图像以及视图之间的 $H_\infty$,仅(隐式地)用图像关系(即不进行显式的 3D 重构)计算 3 维空间中平行 **L** 并过 **X** 的直线的图像. 用这种隐式方法进行计算的其他例子在[Zeller-96]中给出.

目标

给定两幅未标定的图像,计算摄像机和景物结构的一个度量重构($P_M, P'_M, \{\mathbf{X}_{M_i}\}$),即与真实摄像机和景物结构仅相差一个相似变换的重构.

算法

(1) **计算射影重构**($P, P', \{\mathbf{X}_i\}$):

(a) 由两幅图像之间的点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ **计算基本矩阵**.

(b) **摄像机恢复**:由基本矩阵计算摄像机矩阵 $P, P'$.

(c) **三角形法**:对每组点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 计算投影到这两个图像点的空间点 $\mathbf{X}_i$.

(2) **把射影重构矫正到度量重构**:

- 或用**直接的方法**:由已知欧氏位置的 5 个或更多的地面控制点 $\mathbf{X}_{E_i}$ 来计算满足 $\mathbf{X}_{E_i} = H\mathbf{X}_i$ 的单应 H,则度量重构是

$$P_M = PH^{-1}, P'_M = P'H^{-1}, \mathbf{X}_{M_i} = H\mathbf{X}_i.$$

- 或用**分层方法**:

(a) **仿射重构**:按节 9.4.1 介绍的方法计算无穷远平面 $\boldsymbol{\pi}_\infty$,然后用单应

$$H = \begin{bmatrix} I & 0 \\ \boldsymbol{\pi}_\infty^T & \end{bmatrix}$$

把射影重构提升到仿射重构

(b) **度量重构**:按节 9.4.2 介绍的方法计算绝对二次曲线的像 $\omega$,然后用单应

$$H = \begin{bmatrix} A^{-1} & \\ & 1 \end{bmatrix}$$

把仿射重构提升到度量重构,其中 A 由方程 $AA^T = (M^T\omega M)^{-1}$ 的 Cholesky 分解得到,而 M 是仿射重构中的摄像机矩阵左上角的 $3\times3$ 子矩阵,计算 $\omega$ 的目的是为了求 A.

**算法 9.1** 由两幅未标定的图像计算度量重构.

# 第 10 章　计算基本矩阵 F

本章介绍利用两幅图像之间的点对应集来估计基本矩阵的数值方法. 我们先介绍由两幅图像的点对应产生 F 的方程,以及它们的最小解. 接下来几节给出利用代数距离估计 F 的线性方法、各种几何代价函数并包括 MLE(“黄金标准”)算法和 Sampson 距离在内的各种解法.

然后介绍自动获得点对应的一种算法,使得 F 可以从两幅图像直接得到估计. 我们还讨论特殊摄像机运动下的 F 的估计.

本章还涵盖用所算得的 F 来进行图像矫正的方法.

## 10.1　基本方程

基本矩阵由下面方程定义:

$$\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0 \tag{10.1}$$

其中 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 是两幅图像的任意一对匹配点. 给定足够多的匹配点 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i$(至少 7 对),方程(10.1)可用来计算未知的矩阵 F. 进一步地说,记 $\mathbf{x}=(x,y,1)^{\mathrm{T}}$ 和 $\mathbf{x}'=(x',y',1)^{\mathrm{T}}$,每一组点匹配提供关于 F 的未知元素的一个线性方程. 其系数可以很容易地用已知点 $\mathbf{x}$ 和 $\mathbf{x}'$ 的坐标来表示. 具体地说,对应于一对点 $(x,y,1)^{\mathrm{T}}$ 和 $(x',y',1)^{\mathrm{T}}$ 的方程是

$$x'xf_{11}+x'yf_{12}+x'f_{13}+y'xf_{21}+y'yf_{22}+y'f_{23}+xf_{31}+yf_{32}+f_{33}=0. \tag{10.2}$$

用矢量 $\mathbf{f}$ 表示由 F 的元素组成并按行优先顺序排列的 9 维矢量. 那么(10.2)可以表示为一个矢量内积

$$(x'x,x'y,x',y'x,y'y,y',x,y,1)\mathbf{f}=0$$

从 $n$ 组点匹配的集合,我们得到如下线性方程组:

$$\mathrm{A}\mathbf{f}=\begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix}\mathbf{f}=\mathbf{0} \tag{10.3}$$

这是一个齐次方程组,即 $\mathbf{f}$ 可以被确定到只差常数因子. 如果存在一个解,矩阵 A 的秩必然最多为 8,如果秩正好为 8,则解是唯一的(只差尺度因子),而且可以用线性算法求解——此解是 A 的右零空间的生成元.

如果因点的坐标有噪声而导致数据不准确,则矩阵 A 的秩可能大于 8(事实上等于 9,因为 A 有 9 列). 此时要求最小二乘解. 除了方程形式不同(比较(10.3)和(3.3 – $p$54))之外,这个问题和节 3.11 中的估计问题本质上是相同的. 参考 3.1($p$55)的算法. $\mathbf{f}$ 的最小二乘解是对应于 A 的最小奇异值的奇异矢量,即是 A 的 SVD 分解 $\mathrm{A}=\mathrm{UDV}^{\mathrm{T}}$ 中矩阵 V 的最后一列矢量. 用这种方法得到的解矢量 $\mathbf{f}$ 在条件 $\|\mathbf{f}\|=1$ 下取 $\|\mathrm{A}\mathbf{f}\|$ 的最小值. 以上介绍的算法是计算基本矩阵算法的基本方法,称为 8 点算法.

### 10.1.1 奇异性约束

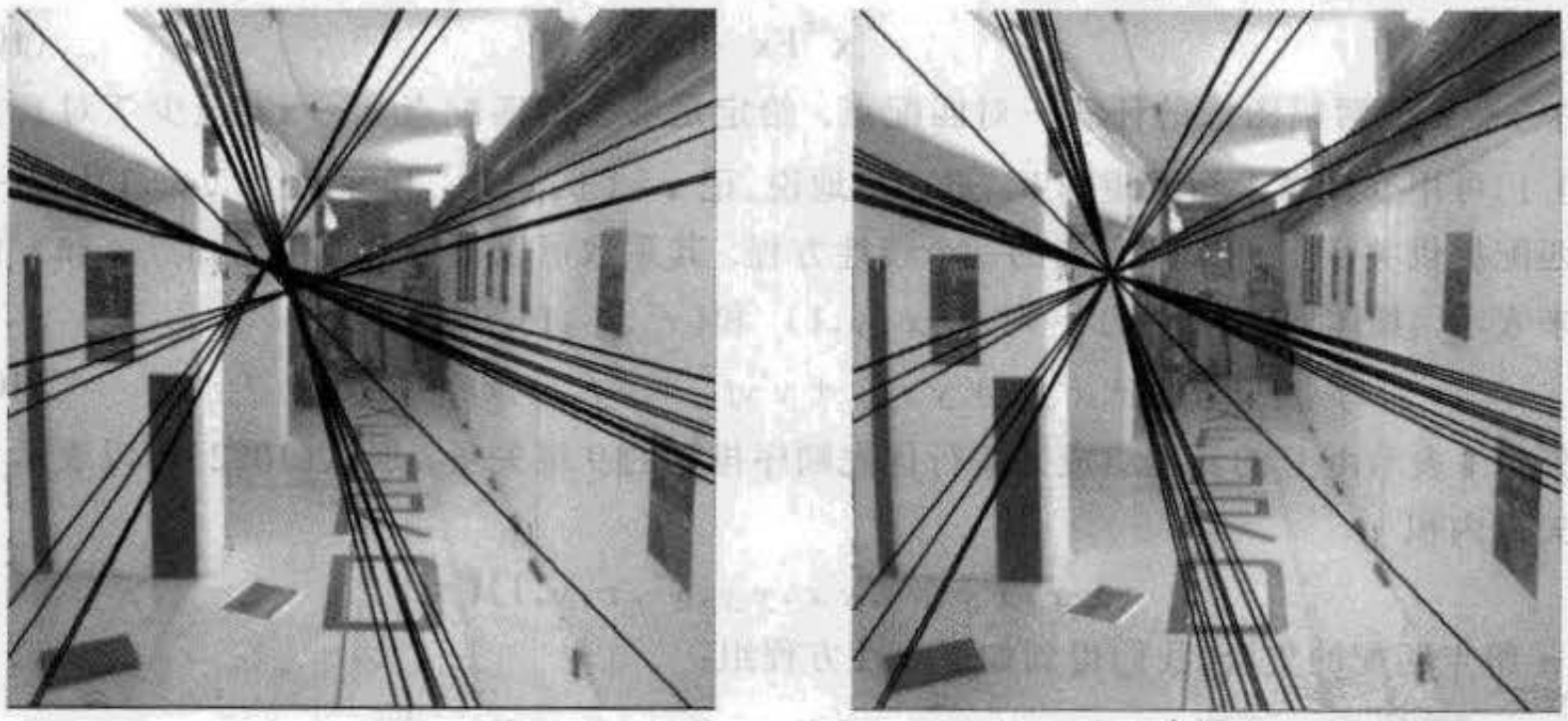
基本矩阵的一个重要性质是奇异性,表现为秩为 2. 进一步说,F 的左和右零空间由两幅图像的两个对极点(在齐次坐标系下)的矢量生成. 基本矩阵的大多数应用与它的秩为 2 这个事实有关. 例如,如果基本矩阵非奇异,则所计算的对极线将不重合,如图 10.1 所示. 由解线性方程组(10.3)所得到的矩阵 F 在一般情况秩不是 2,我们应该采取步骤来强迫这种约束. 实现这个步骤的最方便方法是修正由 A 的 SVD 分解得到的矩阵 F. 矩阵 F 用矩阵 F′来代替并在约束 $\det F'=0$ 下使 Frobenius 范数 $\|F-F'\|$ 达到最小. 实现它的一种简便方法还是用 SVD 分解. 具体地说,若 $F=UDV^T$ 是 F 的 SVD,其中 D 是对角阵,$D=\mathrm{diag}(r,s,t)$ 满足 $r\geq s\geq t$. 则 $F'=U\mathrm{diag}(r,s,0)V^T$ 最小化 $F-F'$的 Frobenius 范数.

因此,计算基本矩阵的 8 点算法可以明确表达为下面的二个步骤:

(1) **线性解** 由对应于 A 的最小奇异值的矢量 $\mathbf{f}$ 求得解 F,其中 A 在(10.3)中定义.

(2) **强迫约束** 在 Frobenius 范数下用最接近 F 的奇异矩阵 F′代替 F. 仍采用 SVD 分解作修正.

假定有适宜的现成线性代数程序,这里陈述的算法非常简单并且便于实现. 同样按惯例必须进行归一化,我们将在节 10.2 中回到这个问题.



a b

图 10.1 **对极线**. (a)非奇异基本矩阵的效应. 对极线根据 $\mathbf{l}'=F\mathbf{x}$ 计算,当 $\mathbf{x}$ 变化时不交于一个公共的对极点. (b)用本节介绍的强迫奇异性的 SVD 方法的效果.

### 10.1.2 最小配置解的情形——7 组点对应

方程 $\mathbf{x}_i'^T F\mathbf{x}_i=0$ 给出形如 $A\mathbf{f}=\mathbf{0}$ 的方程组. 如果 A 的秩为 8,则可以在不计尺度因子下求解 $\mathbf{f}$. 当矩阵 A 的秩为 7 时,仍然可以利用奇异性约束来求解基本矩阵. 仅知道 7 组点对应的情形是一种最重要的情形(其他的情形将在节 10.9 中讨论),由它得到一个 $7\times 9$ 的矩阵 A 的秩一般为 7.

在此情形下,方程 $A\mathbf{f}=\mathbf{0}$ 的解是形如 $\alpha F_1+(1-\alpha)F_2$ 的 2 维空间,其中 $\alpha$ 是纯量变数. 矩阵 $F_1$ 和 $F_2$ 是对应于 A 的右零空间生成元 $\mathbf{f}_1$ 和 $\mathbf{f}_2$ 的矩阵. 现在我们施加约束 $\det F=0$,它可以写成 $\det(\alpha F_1+(1-\alpha)F_2)=0$. 因为 $F_1$ 和 $F_2$ 是已知的,此式给出 $\alpha$ 的一个三次多项

式方程. 可以通过解此多项式方程来求 $\alpha$ 的值. 解的结果可能是一个或三个实解(复解被舍去[Hartley-94b]). 把解代回到方程 $F=\alpha F_1+(1-\alpha)F_2$ 得到基本矩阵的一个或三个可能的解.

这种在最少点数(7)下计算一个或三个基本矩阵的方法将在节 10.6 的鲁棒性算法中采用. 在节 10.9 中,我们还要讨论解的个数问题.

## 10.2 归一化 8 点算法

8 点算法是计算基本矩阵的最简单的方法,它仅涉及构造并(最小二乘)解一个线性方程组. 如果小心的话,它可以执行得非常好. 最初的算法归功于 Longue-Higgins[Longue-Higgins-81]. 8 点算法成功的关键是在构造解的方程之前应对输入的数据认真进行适当的归一化. 对输入的数据进行归一化在本书的许多算法中都采用,并在节 3.4($p65$)中对一般情形做了交代. 在形成 8 点算法的线性方程组之前,图像点的一个简单变换(平移或变尺度)将使这个问题的条件极大地改善,从而提高结果的稳定性. 而且进行这种变换所增加的计算复杂性并不显著.

值得推荐的一种归一化方法是对每一幅图像作平移和缩放使得参考点的形心在坐标原点并且点到原点的 RMS(均方根)距离等于$\sqrt{2}$. 其理由基本上与第 3 章一样. 最基本的方法与算法 3.2($p68$)类似且总结在算法 10.1 中.

注意:奇异性条件约束应该在解除归一化之前强迫实行. 它的合理性请参考[Hartley-97d].

目标

给定 $n\geq 8$ 组点对应 $\{\mathbf{x}_i\leftrightarrow\mathbf{x}'_i\}$,确定基本矩阵 F 使得 $\mathbf{x}'^{\mathsf{T}}_i F\mathbf{x}_i=0$.

算法

(1) **归一化:** 根据 $\hat{\mathbf{x}}_i=T\mathbf{x}_i$ 和 $\hat{\mathbf{x}}'_i=T'\mathbf{x}'_i$ 变换图像坐标,其中 T 和 T′是归一化变换,由平移和缩放组成.

(2) 求对应于匹配 $\hat{\mathbf{x}}_i\leftrightarrow\hat{\mathbf{x}}'_i$ 的基本矩阵 $\hat{F}'$的方法如下

(a) **线性解:** 由 $\hat{A}$的最小奇异值的奇异矢量确定 $\hat{F}$,其中 $\hat{A}$由匹配 $\hat{\mathbf{x}}_i\leftrightarrow\hat{\mathbf{x}}'_i$ 形成,如(10.3)所定义.

(b) **强迫约束:** 用 SVD 并以 $\hat{F}'$代替 $\hat{F}$使得 $\det\hat{F}'=0$(见节 10.11).

(3) **解除归一化:** 令 $F=T'^{\mathsf{T}}\hat{F}'T$. 矩阵 F 是对应于原始数据 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i$ 的基本矩阵.

算法 10.1 F 的归一化 8 点算法.

## 10.3 代数最小化算法

归一化 8 点算法包括了对基本矩阵强迫奇异性约束的方法. 初始估计的 F 由使 $\|F-F'\|$ 达到最小的奇异矩阵 F′替代. 该过程可以通过 SVD 分解来完成,好处是简单而

快速.

然而,这个方法不是最优的数值方法,因为 F 中元素的重要性并不相同,点对应的数据对有些元素的约束实际上比其他元素更严格. 较为正确的过程应根据输入数据计算 F 元素的协方差矩阵,然后根据协方差矩阵,在 Mahalanobis 距离下求最接近 F 的奇异矩阵 F′. 不幸的是,对一般的协方差矩阵 $\Sigma$,最小化 Mahalanobis 距离 $\|F-F'\|_{\Sigma}$ 不是线性的,因此这个方法没有吸引力.

另一种方法是直接求奇异矩阵 F′. 因此,如在 $\|\mathbf{f}\|=1$ 的约束下最小化范数 $\|A\mathbf{f}\|$ 来计算 F 一样,我们应该把目标放在求矩阵 F′使 $\|A\mathbf{f}'\|$ 在 $\|\mathbf{f}'\|=1$ 的约束下为最小. 该过程不可能用线性的非迭代法实现,主要是因为 $\det F'=0$ 是三次约束而不是线性约束. 然而我们将看到存在一个有效的简单迭代方法.

任何一个 3×3 奇异矩阵,比如基本矩阵 F,都可以写为乘积 $F=M[\mathbf{e}]_\times$,其中 M 是一个非奇异矩阵,而 $[\mathbf{e}]_\times$ 是任何反对称矩阵,$\mathbf{e}$ 对应于在第一幅图像上的对极点.

假设我们希望计算 $F=M[\mathbf{e}]_\times$ 并在约束条件 $\|\mathbf{f}\|=1$ 下最小化代数误差 $\|A\mathbf{f}\|$. 现在先假定极点 $\mathbf{e}$ 已知. 以后我们将让 $\mathbf{e}$ 变化,但目前它是固定的. 用由 F 和 M 的元素组成的矢量 $\mathbf{f}$ 和 $\mathbf{m}$ 把方程 $F=M[\mathbf{e}]_\times$ 记为方程 $\mathbf{f}=E\mathbf{m}$,其中 E 是一个 9×9 的矩阵. 假设 $\mathbf{f}$ 和 $\mathbf{m}$ 把相应矩阵的元素按行优先排序,则可以验证 E 的形式是

$$E=\begin{bmatrix}[\mathbf{e}]_\times & & \\ & [\mathbf{e}]_\times & \\ & & [\mathbf{e}]_\times\end{bmatrix} \tag{10.4}$$

因为 $\mathbf{f}=E\mathbf{m}$,该最小化问题变成:[①]

$$\text{在约束条件 } \|E\mathbf{m}\|=1 \text{ 下,最小化 } \|AE\mathbf{m}\| \tag{10.5}$$

这个最小化问题可用算法 A3.7(*p*414)求解. 可以看到 E 的秩为 6,因为每个对角块的秩都为 2.

<u>目标</u>

求最小化 $\|A\mathbf{f}\|$ 的基本矩阵 F 并满足 $\|\mathbf{f}\|=1$ 和 $\det F=0$.

<u>算法</u>

(1) 用归一化 8 点算法(10.1)求基本矩阵的第一个近似 $F_0$,接着求 $F_0$ 的右零矢量 $\mathbf{e}_0$.

(2) 从对极点的估计 $\mathbf{e}_i=\mathbf{e}_0$ 开始,按照(10.4)计算矩阵 $E_i$,接着求最小化 $\|A\mathbf{f}_i\|$ 并满足 $\|\mathbf{f}_i\|=1$的矢量 $\mathbf{f}_i=E_i\mathbf{m}_i$. 它用算法 A3.7(*p*414)完成.

(3) 计算代数误差 $\boldsymbol{\varepsilon}_i=A\mathbf{f}_i$. 因为 $\mathbf{f}_i$ 从而 $\boldsymbol{\varepsilon}_i$ 在相差一个符号的意义下定义,修正 $\boldsymbol{\varepsilon}_i$ 的符号(必要时乘负 1)使得当 $i>0$ 时 $\mathbf{e}_i^T\mathbf{e}_{i-1}>0$. 这样做是为了保证作为 $\mathbf{e}_i$ 的函数的 $\boldsymbol{\varepsilon}_i$ 平滑变化.

(4) 上两步定义了 $\mathbb{R}^3\to\mathbb{R}^9$ 的一个映射:$\mathbf{e}_i\mapsto\boldsymbol{\varepsilon}_i$. 现在用 Levenberg - Marquardt 算法(节 A4.2(*p*417))迭代地变化 $\mathbf{e}_i$ 以便最小化 $\|\boldsymbol{\varepsilon}_i\|$

(5) 收敛时,$\mathbf{f}_i$ 代表所求的基本矩阵.

算法 10.2 迭代最小化代数误差来计算满足 $\det F=0$ 的 F.

① 它不是在条件 $\|\mathbf{m}\|=1$ 下最小化 $\|AE\mathbf{m}\|$,因为当 $\mathbf{m}$ 是 E 的右零空间的单位矢量时是它的一个解. 在这种情形下,$E\mathbf{m}=\mathbf{0}$,因而 $\|AE\mathbf{m}\|=\mathbf{0}$.

### 10.3.1 迭代估计

(10.5)给出了在对极点 $\mathbf{e}$ 的值已知时计算代数误差矢量 $\mathbf{Af}$ 的一种最小化方法. 映射 $\mathbf{e} \mapsto \mathbf{Af}$是一个从 $\mathbb{R}^3$ 到 $\mathbb{R}^9$ 的映射. 注意 $\mathbf{Af}$ 的值不受 $\mathbf{e}$ 的尺度变化的影响. 由 F 的初始估计导出右零空间的生成元 $\mathbf{e}$ 的一个估计值,然后可以通过迭代最终求出最小化代数误差的 F. 可以通过 8 点算法或任何其他的简单算法来得到 F 的初始估计. 这种计算 F 的完整的算法在算法 10.2 中给出.

注意这种计算 F 的方法的长处是算法的迭代部分仅涉及三个参数的估计($\mathbf{e}$ 的齐次坐标)的一个非常小的参数最小化问题. 这个算法找到的基本矩阵在所有匹配点上最小化代数误差. 尽管这些匹配点本身并不进入最后的迭代估计.

## 10.4 几何距离

本节介绍三种最小化图像几何距离的算法. 我们推荐的算法是黄金标准方法,但在实现时需花最多的代价. 其他的算法也产生非常好的结果并且较易实现,但缺点是在图像误差是高斯噪声的假设下不是最优的. 每个算法都涉及两个重要问题:非线性最小化的初始化和代价函数的参数化. 这些算法一般用前一节中的某种线性算法作初始化. 在自动算法中采用另一种方法,它选择 7 组对应点并由此产生关于 F 的一个或 3 个解. 各种参数化方法在节 10.4.2 中讨论. 对于所有的算法,我们都建议用平移和缩放使图像点归一化. 这种归一化不改变噪声的特性,因此不影响下面介绍的黄金标准算法的最优性.

### 10.4.1 黄金标准方法

基本矩阵的最大似然估计与误差模型的假定有关. 我们假定图像点的测量噪声服从高斯分布. 因此,ML 估计就是一种最小化几何距离(重投影误差)

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 \tag{10.6}$$

的估计,其中 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 是测量得到的对应;而 $\hat{\mathbf{x}}_i$ 和 $\hat{\mathbf{x}}'_i$ 是待估计的"真"对应,它们准确地满足 $\hat{\mathbf{x}}'^{\mathsf{T}}_i \mathrm{F} \hat{\mathbf{x}}_i = 0$,其中 F 为待估计的秩为 2 的基本矩阵.

误差函数可以按下面的方式最小化. 定义一对摄像机矩阵 $\mathrm{P} = [\mathrm{I} | \mathbf{0}]$和 $\mathrm{P}' = [\mathrm{M} | \mathbf{t}]$. 另外定义 3D 中的点 $\mathbf{X}_i$. 现在令 $\hat{\mathbf{x}}_i = \mathrm{P}\mathbf{X}_i$ 和 $\hat{\mathbf{x}}'_i = \mathrm{P}'\mathbf{X}_i$,变化 P′和点 $\mathbf{X}_i$ 使这个误差表达式最小化. 接下来 F 可由 $\mathrm{F} = [\mathbf{t}]_\times \mathrm{M}$ 计算. 矢量 $\hat{\mathbf{x}}_i$ 和 $\hat{\mathbf{x}}'_i$将满足 $\hat{\mathbf{x}}'^{\mathsf{T}}_i \mathrm{F} \hat{\mathbf{x}}_i = 0$. 这个误差的最小化可用 A4.2($p$417)介绍的 Levenberg - Marquardt 算法来实现. 参数的初始估计用归一化的 8 点算法计算,然后按第 11 章的介绍进行射影重构. 因此,用这种方法估计基本矩阵实际上等价于射影重构. 这个算法的步骤概括在算法 10.3 中.

用这种方法计算基本矩阵看上去计算代价很大. 然而,采用稀疏的 LM 技术后结果并不比其他迭代技术的开销大很多,有关细节在 A4.5($p$423)中给出.

### 10.4.2 秩 2 矩阵的参数化

几何距离代价函数的非线性最小化需要在具有秩 2 的强迫条件下对基本矩阵参数化.

我们介绍三种这样的参数化方法.

**超参数化** 参数化 F 的一种方法是记 $F=[\mathbf{t}]_\times M$,其中 M 是任何一个 $3\times3$ 矩阵. 此方法保证 F 是奇异的,因为$[\mathbf{t}]_\times$是奇异的. 这样一来,F 被 M 的 9 个元素和 $\mathbf{t}$ 的三个元素参数化——总共有 12 个参数,它超过了最小参数个数:7. 一般,这样做不会发生大的问题.

目标

给定 $n\geqslant8$ 组图像点对应$\{\mathbf{x}_i\leftrightarrow\mathbf{x}'_i\}$,确定基本矩阵的最大似然估计(MLE)$\hat{F}$.

MLE 同时包括求满足 $\hat{\mathbf{x}}'^{\mathrm{T}}_i\hat{F}\hat{\mathbf{x}}_i=0$ 的点对应集合$\{\hat{\mathbf{x}}_i\leftrightarrow\hat{\mathbf{x}}'_i\}$,并且最小化

$$\sum_i d(\mathbf{x}_i,\hat{\mathbf{x}}_i)^2+d(\mathbf{x}'_i,\hat{\mathbf{x}}'_i)^2.$$

算法

(1) 用线性算法(例如算法 10.1)来计算 $\hat{F}$的一个秩为 2 的初始估计.

(2) 计算附属变量$\{\hat{\mathbf{x}}_i,\hat{\mathbf{x}}'_i\}$的初始估计如下:

(a) 选择摄像机矩阵 $P=[I|\mathbf{0}]$和 $P'=[[\mathbf{e}']_\times\hat{F}|\mathbf{e}']$,其中 $\mathbf{e}'$由 $\hat{F}$得到.

(b) 用第 11 章的三角形法,从对应 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i$ 和 $\hat{F}$确定 $\hat{\mathbf{X}}_i$ 的一个估计.

(c) 与 $\hat{F}$相容的对应是 $\hat{\mathbf{x}}_i=P\hat{\mathbf{X}}_i$, $\hat{\mathbf{x}}'_i=P'\hat{\mathbf{X}}_i$.

(3) 在 $\hat{F}$和 $\hat{\mathbf{x}}_i$, $i=1,\ldots,n$,上最小化代价函数

$$\sum_i d(\mathbf{x}_i,\hat{\mathbf{x}}_i)^2+d(\mathbf{x}'_i,\hat{\mathbf{x}}'_i)^2.$$

用 Levenberg-Marquardt 算法来最小化这个代价,它有 $3n+12$ 个变量:$3n$ 来自 $n$ 个 3D 点 $\hat{\mathbf{X}}_i$,而 12 来自摄像机矩阵 $P'=[M|\mathbf{t}]$,其中 $\hat{F}=[\mathbf{t}]_\times M$ 和 $\hat{\mathbf{x}}_i=P\hat{\mathbf{X}}_i$, $\hat{\mathbf{x}}'_i=P'\hat{\mathbf{X}}_i$.

算法 10.3 从图像对应估计 F 的黄金标准算法.

**对极参数化** 另一种参数化 F 的方法是指定 F 的前两列和两个乘子 $\alpha$ 和 $\beta$ 使得其第三列可记为前两列的线性组合 $\mathbf{f}_3=\alpha\mathbf{f}_1+\beta\mathbf{f}_2$. 于是,基本矩阵可以参数化为

$$F=\begin{bmatrix} a & b & \alpha a+\beta b \\ c & d & \alpha c+\beta d \\ e & f & \alpha e+\beta f \end{bmatrix} \tag{10.7}$$

它总共有 8 个参数. 为了得到最小参数化集合,其中一个元素,比如 $f$,可以设为 1. 实际上,常令 $a,\ldots,f$ 中绝对值最大的参数为 1. 这个方法只用最少参数就保证得到一个奇异性矩阵 F. 其主要缺点是它有一个奇性——当 F 的前两列线性相关时就无效了,因为此时不可能根据前两列求得第三列. 这个问题是有意义的,因为当右极点为无穷远点时会发生此种情形. 此时 $F\mathbf{e}=F(e_1,e_2,0)^{\mathrm{T}}=0$,从而 F 的前两列线性相关. 尽管如此,这种参数化方法仍被广泛地运用而且如果采取措施回避了这个奇性效果还相当好. 当对极点在一个坐标轴上导致奇性发生时,可以取另外的两列代替前两列作为基. 在实际运算中,这种奇性在最小化过程中可被检测到,一旦发现,可把此参数化方法转为另一种参数化方法来处理.

注意$(\alpha,\beta,-1)^{\mathrm{T}}$ 是这个基本矩阵的右对极点——对极点坐标显式地出现在参数化中. 为了得到最好的结果,参数化应该选取该对极点中绝对值最大的元素为 1.

注意：基本矩阵的完整流形可能不由单个参数化所覆盖，而由最小参数化块的集合覆盖. 在参数最小化的过程中，在流形上会划出一条路径，在穿过两块边界时它从一块转到另一块. 在现在的情形下，实际上一共有 18 个不同的参数块，与 a,...,$f$ 中哪一个最大的和哪两列被取做基有关.

**两个对极点都作为参数** 上面的参数化用一个对极点作为参数化的一部分. 为对称起见，我们可以用两个对极点作为参数. 所得 F 的形式为

$$\mathrm{F}=\begin{bmatrix} a & b & \alpha a+\beta b \\ c & d & \alpha c+\beta d \\ \alpha' a+\beta' c & \alpha' b+\beta' d & \alpha'\alpha a+\alpha'\beta b+\beta'\alpha c+\beta'\beta d \end{bmatrix} \tag{10.8}$$

两个对极点为$(\alpha,\beta,-1)^{\mathsf{T}}$和$(\alpha',\beta',-1)^{\mathsf{T}}$. 如上所述，可以把 $a,b,c,d$ 中的一个取 1. 为避免奇点，我们必须对取作基的两行和两列的不同选择进行权衡. 加上 $a,b,c,d$ 之一取为 1 的四个选择，覆盖基本矩阵的完整流形总共有 36 个参数化块.

### 10.4.3 一阶几何误差(Sampson 距离)

Sampson 距离的概念在节 3.2.6($p61$)中作了长篇的讨论. 这里 Sampson 近似用于由 $\mathbf{x}'^{\mathsf{T}}\mathrm{F}\mathbf{x}=0$ 定义的簇，以便提供一阶几何误差的近似.

(3.13 - $p62$)给出了 Sampson 代价函数的一般公式. 用于基本矩阵的估计时，这个公式更为简单. 因为每组点对应只有一个方程(同时参见例 3.2($p61$)). 从而偏导数矩阵 J 只有一行，因此 $\mathrm{JJ}^{\mathsf{T}}$ 是一个纯量且(3.12 - $p61$)变为

$$\frac{\boldsymbol{\varepsilon}^{\mathsf{T}}\boldsymbol{\varepsilon}}{\mathrm{JJ}^{\mathsf{T}}}=\frac{(\mathbf{x}_i'^{\mathsf{T}}\mathrm{F}\mathbf{x}_i)^2}{\mathrm{JJ}^{\mathsf{T}}}.$$

从 J 的定义和(10.2)的左边给出的显式 $\mathrm{A}_i=\mathbf{x}_i'^{\mathsf{T}}\mathrm{F}\mathbf{x}_i$ 得到

$$\mathrm{JJ}^{\mathsf{T}}=(\mathrm{F}\mathbf{x}_i)_1^2+(\mathrm{F}\mathbf{x}_i)_2^2+(\mathrm{F}^{\mathsf{T}}\mathbf{x}_i')_1^2+(\mathrm{F}^{\mathsf{T}}\mathbf{x}_i')_2^2$$

其中$(\mathrm{F}\mathbf{x}_i)_j^2$ 表示矢量 $\mathrm{F}\mathbf{x}_i$ 第 $j$ 个元素的平方. 因此，该代价函数为

$$\sum_i \frac{(\mathbf{x}_i'^{\mathsf{T}}\mathrm{F}\mathbf{x}_i)^2}{(\mathrm{F}\mathbf{x}_i)_1^2+(\mathrm{F}\mathbf{x}_i)_2^2+(\mathrm{F}^{\mathsf{T}}\mathbf{x}_i')_1^2+(\mathrm{F}^{\mathsf{T}}\mathbf{x}_i')_2^2}. \tag{10.9}$$

此式给出几何误差的一阶逼近. 如果高阶项比一阶项相对地小，那么就可指望它会给出好的结果. 这个近似被[Torr - 97, Toor - 98, Zhang - 98]成功地应用在估计算法中. 注意这个近似在由这两个对极点确定的 $\mathbb{R}^4$ 空间的点上没有定义，因为 $\mathrm{JJ}^{\mathsf{T}}$ 为 0. 在数值实现时应该避免这一点.

用这种方法逼近几何误差的突出优点是得到的代价函数只涉及到 F 的参数，这意味着在精确到一阶条件下最小化黄金标准代价函数**不需要**引入辅助变量集合，即不需要 $n$ 个空间点 $\mathbf{X}_i$ 的坐标. 从而把一个 $7+3n$ 个自由度的最小化问题减少为一个只有 7 个自由度的问题.

**对称对极点距离** 方程(10.9)在形式上类似于另一种代价函数

$$\sum_i d(\mathbf{x}_i',\mathrm{F}\mathbf{x}_i)^2+d(\mathbf{x}_i,\mathrm{F}^{\mathsf{T}}\mathbf{x}_i')^2=\sum_i(\mathbf{x}_i'^{\mathsf{T}},\mathrm{F}\mathbf{x}_i)^2\left(\frac{1}{(\mathrm{F}\mathbf{x}_i)_1^2+(\mathrm{F}\mathbf{x}_i)_2^2}+\frac{1}{(\mathrm{F}^{\mathsf{T}}\mathbf{x}_i')_1^2+(\mathrm{F}^{\mathsf{T}}\mathbf{x}_i')_2^2}\right) \tag{10.10}$$

它在每幅图像上最小化点到它的投影对极线的距离. 但是,这个代价函数给出的结果似乎比(10.9) 稍差一点(见[Zhang-98]),因此不作进一步讨论.

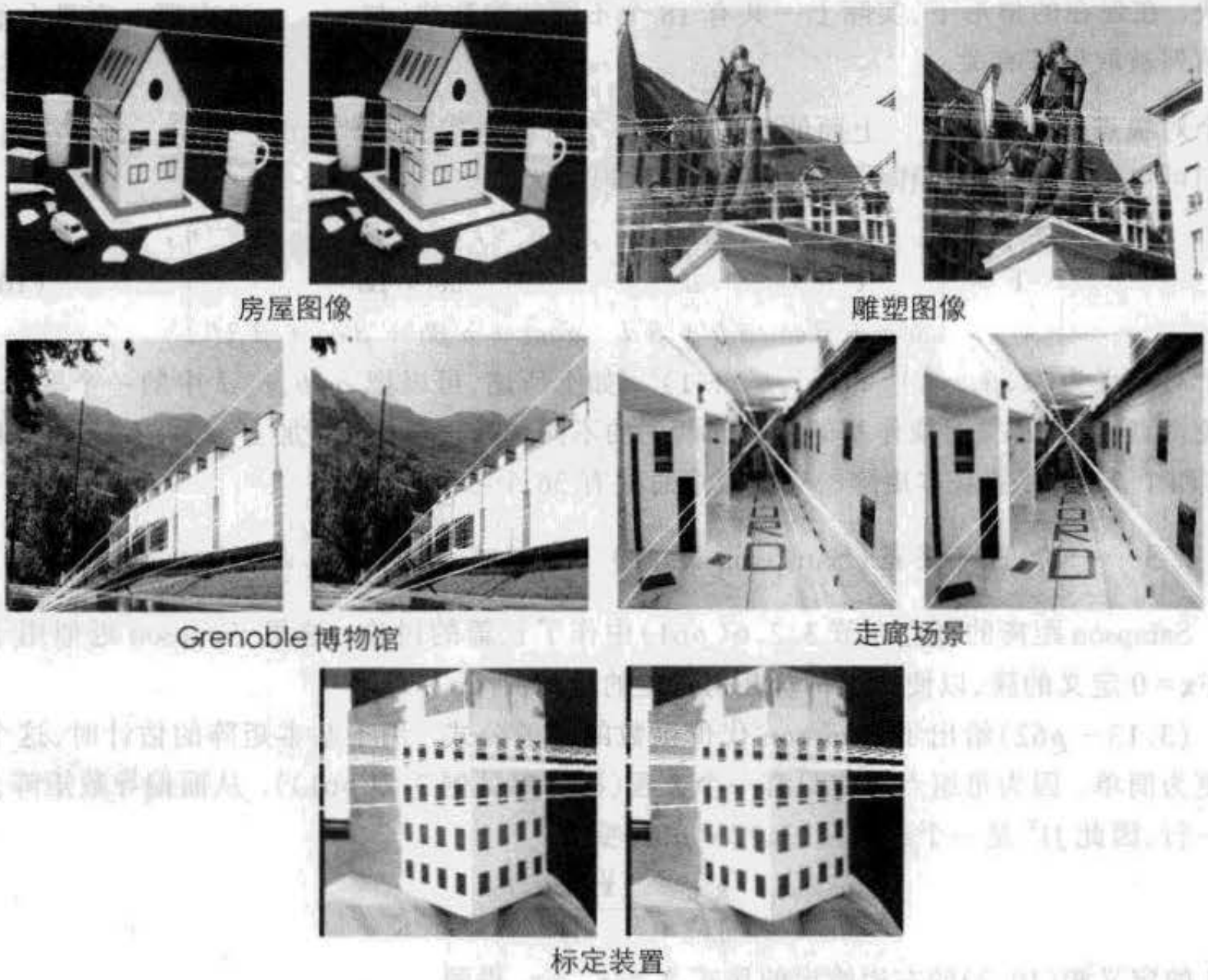



图 10.2 **用于算法比较的图像对**. 在顶部的两对图像中,对极点远离图像中心. 在中间的两对图像中,博物馆的对极点相距很近而走廊的对极点在图像上. 对于标定图像,匹配点可以极其精确地得到.

## 10.5 算法的实验评估

我们用若干组图像对的点对应来估计 F 以便比较前一节的三种算法. 这些算法是:

(1) 归一化 8 点算法(算法 10.1)

(2) 代数误差的最小化同时强迫奇异约束(算法 10.2)

(3) 黄金标准几何算法(算法 10.3)

实验过程如下:对每一对图像,从匹配中随机地选取 $n$ 组匹配点,并估计基本矩阵和计算残差(见下面). 这个实验对每个 $n$ 值和每一对图像重复 100 次,然后画出对应于 $n$ 的平均残差的图. 它给出了当点数增加时不同算法的性能概况,实验中使用的点数 $n$ 从 8 增加到所有匹配点的 3/4.

**残差** 该误差定义为

$$\frac{1}{N}\sum_{i}^{N} d(\mathbf{x}'_i, \mathrm{F}\mathbf{x}_i)^2 + d(\mathbf{x}_i, \mathrm{F}^{\mathsf{T}}\mathbf{x}'_i)^2$$

其中 $d(\mathbf{x},\mathbf{l})$是点 $\mathbf{x}$ 到直线 $\mathbf{l}$ 的距离(以像素为单位). 该误差等于所有 $N$ 组匹配中每一点的对极线与该点在另一幅图像上的匹配点之间的距离(对匹配中的两点都计算)平方和的平均值. 注意这个误差对**所有** $N$ **对**匹配点进行评估,而不仅仅对用于计算 F 的 $n$ 组匹配. 对应于对极点距离的残差在(10.10)中定义. 注意这里评估的任何算法**没有**直接对这个误差最小化.

各种算法用 5 个不同的图像对作试验. 这些图像在图 10.2 中给出,它体现了图像类型和对极点位置的多样性. 图像上显示了若干条对极线. 对极线束的交点是对极点. 虽然错匹配在预处理中已去掉,不同图像上的匹配点的准确性仍然区别很大.

**结果**

图 10.3 显示和解释了这些实验的结果. 它们表明最小化代数误差和最小化几何误差的结果没有本质的区别.

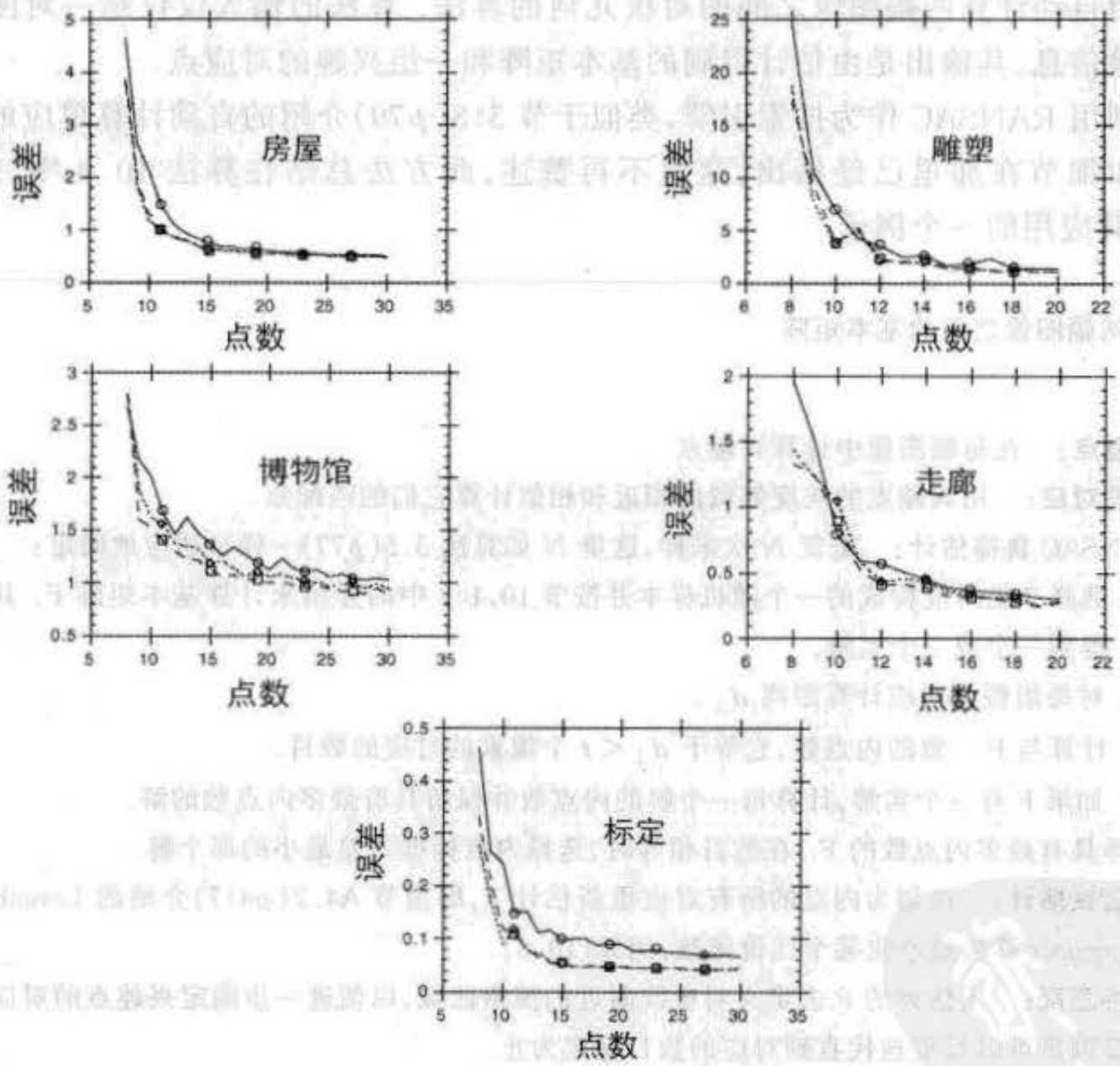



图 10.3 **算法实验评估的结果**. 对每一种情形比较计算 F 的 3 种方法. 残差与用于计算 F 的点数间的关系用作图显示. 在每一幅图中,上部曲线(实线)显示归一化 8 点算法的结果. 同时显示了最小化几何误差的结果(长划线)以及在行列式约束下的迭代最小化代数误差的结果(短划线). 在大多数情况下,迭代最小化代数误差与最小化几何误差的结果几乎没有区别. 这两种算法都明显优于未迭代的归一化 8 点算法,尽管后者的结果也不错.

## 10.5.1 建议

本章讨论了计算基本矩阵的若干方法,我们觉得有必要给出方法使用的某些要点. 简要地说,我们的建议是:

- 不要使用未归一化的 8 点算法.
- 快速并易于实行的方法是归一化的 8 点算法 10.1,它常能给出足够好的结果,也适宜作为其他算法的第一步.
- 如果需要更高的准确性,采用代数最小化方法,用或者不用关于对极点位置的迭代均可.
- 另一种能给出很好结果的方法是运用最小化 Sampson 代价函数(10.9)的迭代最小化方法. 它给出与迭代代数方法相似的结果.
- 若高斯噪声假设成立,采用黄金标准算法肯定能得到最好的结果.

## 10.6 自动计算 F

本节介绍自动计算两幅图像之间的对极几何的算法. 算法的输入仅仅是一对图像,不需其他的**先验**信息,其输出是由估计得到的基本矩阵和一组兴趣的对应点.

该算法利用 RANSAC 作为搜索引擎,类似于节 3.8($p$79)介绍的自动计算单应的方法. 算法的思想和细节在那里已经给出,这里不再赘述,此方法总结在算法 10.4 中,并在图 10.4 中给出其应用的一个例子.

目标　计算两幅图像之间的基本矩阵

算法

(1) **兴趣点:** 在每幅图像中计算兴趣点

(2) **假设对应:** 用兴趣点的灰度邻域的相近和相似计算它们的匹配集.

(3) RANSAC **鲁棒估计:** 重复 $N$ 次采样,这里 $N$ 如算法 3.5($p$77)一样自适应地确定:

(a) 选择 7 组对应构成的一个随机样本并按节 10.1.2 中的介绍来计算基本矩阵 F. 其结果得到一个或三个实解.

(b) 对每组假设对应计算距离 $d_\perp$.

(c) 计算与 F 一致的内点数,它等于 $d_\perp < t$ 个像素的对应的数目.

(d) 如果 F 有三个实解,计算每一个解的内点数并保留具有最多内点数的解.

选择具有最多内点数的 F. 在数目相等时,选择内点标准方差最小的那个解.

(4) **非线性估计:** 由划为内点的所有对应重新估计 F,即用节 A4.2($p$417)介绍的 Levenberg-Marquardt 算法最小化某个代价函数,例如(10.6).

(5) **引导匹配:** 用估计的 F 去定义对极线附近的搜索区域,以便进一步确定兴趣点的对应.

最后两步可以反复迭代直到对应的数目稳定为止.

算法 10.4　用 RANSAC 自动估计两幅图像之间的基本矩阵的算法.

有关这种方法的一些注意事项是:

(1) **RANSAC 样本**　仅用 7 组点对应来估计 F. 它的好处是必定生成一个秩 2 的矩阵,从而无需(如线性算法那样)强迫矩阵的秩为 2.

不采用线性算法的 8 组对应而采用 7 组对应的第二个理由是:保证一个高概率的没有野值的结果所需的采样次数是样本集大小的指数函数. 例如,从表 3.3($p$76)可知:8 组对应达 99% 没有野值的置性度(从一个包含 50% 野值的集合采样)时所需的采样数是 7 组对应所需采样数的两倍.

采用 7 组对应略为不利的方面是它可能给出 F 的 3 个实数解，且所有 3 个解都要通过检验来选择.

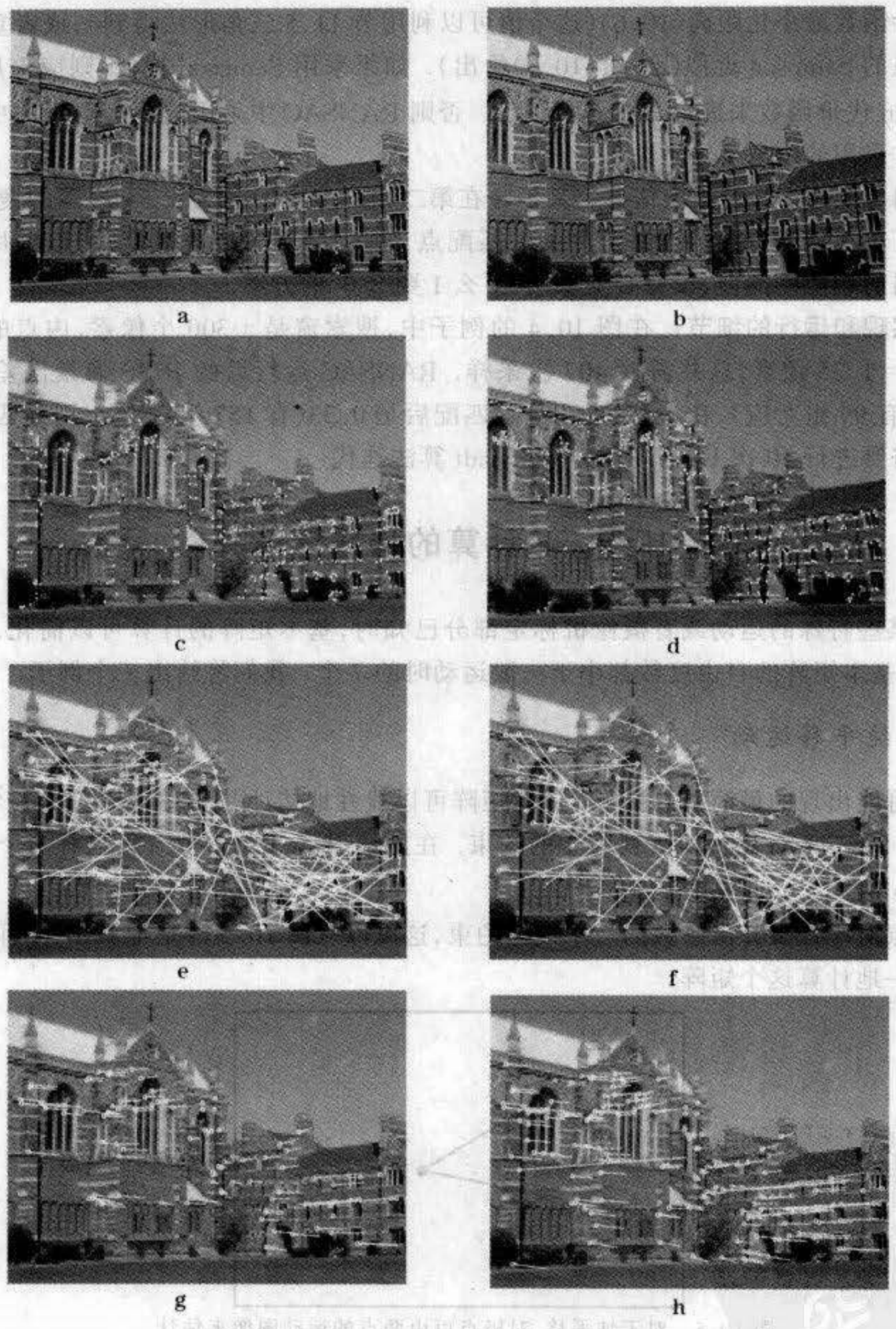



图 10.4 **用 RANSAC 自动计算两幅图像之间的基本矩阵**. (a)(b) 牛津大学 Keble 学院的左和右图像. 视图之间的运动是一个平移加一个旋转. 图像像素是 640×480. (c)(d)把检测到的角点叠加在图像上. 每幅图像上大约有 500 个角点. 下面的结果皆叠加在左图像上:(e)188 对假设匹配用连接角点的线显示,注意明显的误匹配;(f)野值——假设匹配中有 89 对. (g)99 组对应与估计的 F 一致;(h)通过引导匹配和 MLE 后,最后得到 157 组对应. 仍然存在若干误匹配的现象. 例如图左边的长线.

(2) **距离测量** (从 RANSAC 样本)给出 F 的一个当前的估计,用距离 $d_\perp$ 测量一对匹配点满足对极几何的接近程度. 关于 $d_\perp$ 有两种明显的选择:重投影误差,即用代价函数最小化距离(10.6)(这个值可以利用节 11.5 三角形法得到);或者重投影误差的 Sampson 近似($d_\perp^2$ 由(10.9)给出). 如果采用 Sampson 近似,则应该用 Sampson 代价函数来进行 F 的迭代估计. 否则 RANSAC 中采用的距离与算法中其他地方采用的距离不一致.

(3) **引导匹配** F 的当前估计定义了 $\mathbf{x}$ 在第二幅图像中的对极线 $\mathbf{Fx}$ 周围的搜索范围. 可以在这个范围内搜索角点 $\mathbf{x}$ 的匹配点. 因为限制了寻找区域,可以采用一个较弱的相似性阈值,没有必要采用"要么 1 要么 0"的方案.

(4) **实现和运行的细节** 在图 10.4 的例子中,搜索窗是 ±300 个像素,内点的阈值是 $t=1.25$ 像素,总共需要 407 次采样. RANSAC 运行后的 RMS 像素误差是 0.34(有 99 组对应),在 MLE 和引导性匹配后是 0.33(有 157 组对应). 引导匹配 MLE 需要进行 10 次 Levenberg-Marquadt 算法迭代.

## 10.7 F 计算的特殊情形

对于某些特殊的运动或者摄像机标定部分已知时,基本矩阵的计算可以简化. 在每一种情形下,基本矩阵的自由度数都小于一般运动时的 7 个. 我们将给出三个例子.

### 10.7.1 纯平移运动

这是可能出现的最简单的情形. 此时矩阵可以线性地估计并同时满足它是一个反对称矩阵(见节 8.3.1($p$166))即其秩为 2 的约束. 在此情形下,$\mathbf{F}=[\mathbf{e}']_\times$,且只有两个自由度,它可以由 $\mathbf{e}'$ 的三个元素来参数化.

每组点对应提供齐次参数的一个线性约束,这可以从图 10.5 清楚地看出. 从两组点对应可以唯一地计算这个矩阵.

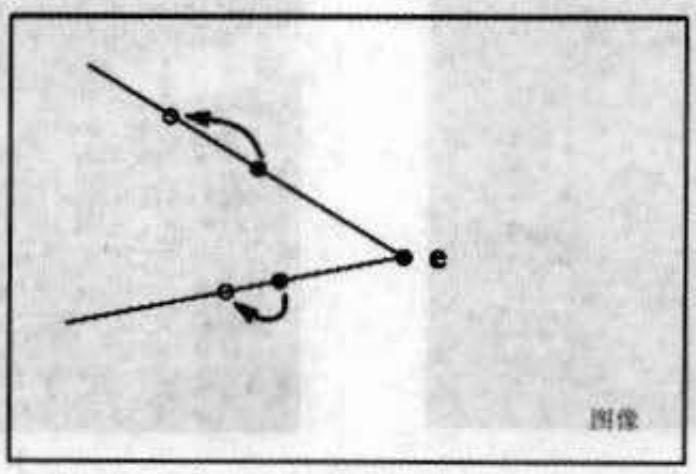



图 10.5 对于纯平移,对极点可由两点的运动图像来估计.

注意:对于一般的运动情形,如果所有 3D 点共面,则它是一种退化结构(见节 10.9),基本矩阵不能由图像对应唯一地确定. 但是,对于纯平移运动,则没有问题(两个 3D 点永远是共面的). 唯一的退化结构是两个 3D 点与两个摄像机中心共面.

这个特殊形式也简化了黄金标准估计以及结构恢复的三角形法.

### 10.7.2 平面运动

在节 8.3.2($p$168)描述的平面运动中,除了矩阵 F 秩为 2 的标准条件外,我们要求 F 的对称部分的秩也为 2. 可以验证(8.8 - $p$170)的参数化,即 $F=[\mathbf{e}']_\times[\mathbf{l}_\infty]_\times[\mathbf{e}]_\times$ 同时满足这两个条件. 如果用不加约束的 3 维矢量来表示 $\mathbf{e}'$,$\mathbf{l}_\infty$ 和 $\mathbf{e}$ 则用了 9 个参数,然而在平面运动时基本矩阵只有 6 个自由度. 和通常一样,这样超参数化不成问题.

### 10.7.3 标定的情形

对于摄像机已标定的情形,可以用归一化图像坐标并用计算本质矩阵代替求基本矩阵. 和基本矩阵一样,本质矩阵可以由 8 点或更多的点通过线性方法来计算,因为对应点满足定义方程 $\mathbf{x}_i'^{\mathsf{T}}E\mathbf{x}_i=0$.

而该方法与计算基本矩阵不同之处是约束的强制. 我们要求基本矩阵满足 $\det F=0$,而要求本质矩阵除此之外还要满足其两个奇异值相等的附加条件. 这个约束可以由下面仅给出而不证明的结论来处理.

**结论 10.1 若 E 是一个 $3\times3$ 矩阵,其 SVD 分解为 $E=UDV^{\mathsf{T}}$,其中 $D=\mathrm{diag}(a,b,c)$ 且 $a\geq b\geq c$,则在 Frobenius 范数下最接近 E 的本质矩阵是 $\hat{E}=U\hat{D}V^{\mathsf{T}}$,其中 $\hat{D}=\mathrm{diag}((a+b)/2,(a+b)/2,0)$.**

如果重构的目标是计算两个归一化的摄像机矩阵 P 和 P′,则其实没有必要通过乘积 $\hat{E}=U\hat{D}V^{\mathsf{T}}$ 来计算 $\hat{E}$. 根据结论 8.19($p$175),矩阵 P′可直接由 SVD 分解计算得到. 在 4 个可能的解中选取 P′的原则是:能观察到的点必须在两个摄像机前面. 详见节 8.6.3($p$175)的解释.

## 10.8 其他几何元素的对应

迄今为止,本章的计算仅采用了点对应,自然会产生的问题是:F 能否由图像上其他几何元素的对应来计算? 答案是肯定的. 但不包括所有类型的元素,这里讨论一些普通的例子.

**直线** 视图之间的图像直线的对应对 F 完全**没有**约束. 这里的直线是指无限直线,不是直线段. 考虑图像点对应的的情形:每幅图像上的点反向投影成过摄像机中心的射线,这些射线相交于一个 3D 点. 一般情况下两条 3D 直线是偏斜的(它们不相交);因此射线相交的条件为对极几何提供了一个约束. 相反在图像直线对应的情形,每一视图的直线反向投影为平面. 但是,在 3D 空间中两个平面永远相交,因此,不为对极几何提供约束(在 3 视图时,存在一个约束).

但平行直线的消影点的对应能提供关于 F 的一个约束. 其实,一个消影点同任何有限点的地位相同,即它也提供一个约束.

**空间曲线和曲面** 如图 10.6 所示,在对极平面与空间曲线相切的点上,影像曲线与对应的对极线相切. 这一事实为两视图几何提供了一个约束,即在一幅视图中如果一条对极线与一条像曲线相切,则其对应的对极线必须与另一视图中的像曲线相切. 类似地,在曲面情

形，在对极平面与曲面相切的点上，影像轮廓线也相切于对应的对极线．对极线切点与点对应作用一样并且可以如[Porrill－91]所介绍的那样把它包括在估计算法中．

二次曲线和二次曲面特别重要，因为它们都是代数对象，从而可以导出代数解．这样的例子在本章末尾的注释和练习中给出．

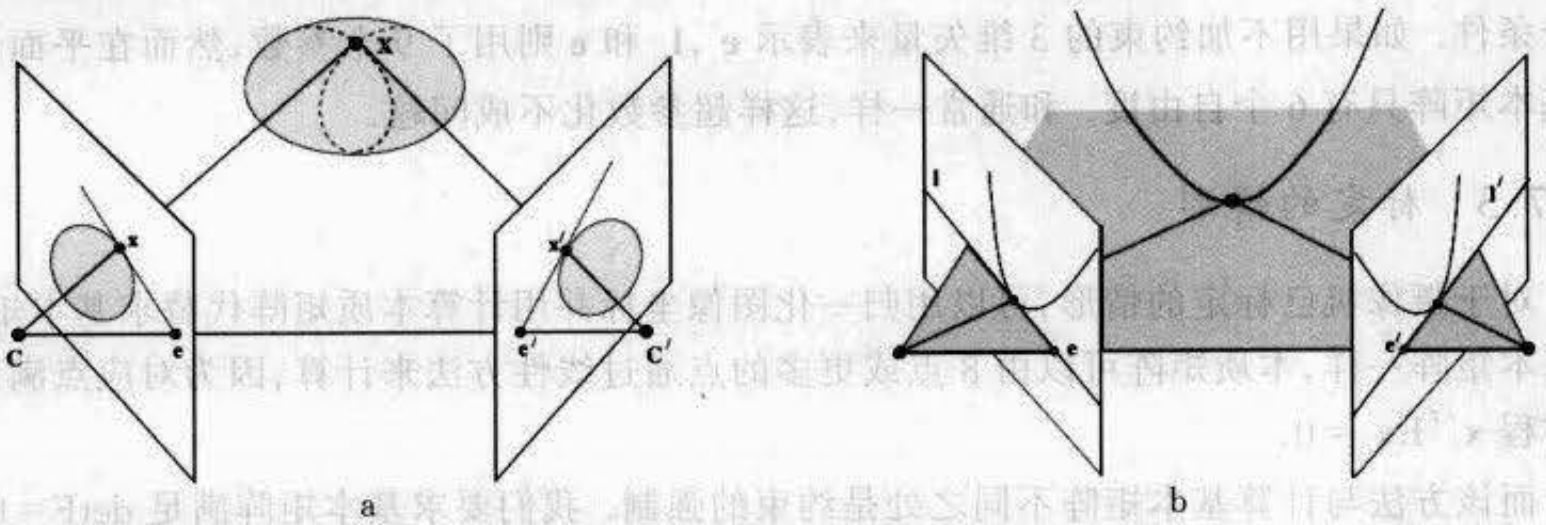



图 10.6　**对极几何的相切性**．(a)曲面的情形；(b)空间曲线的情况——图像取自 Porrill 和 Pollard[Porrill－91]．在(a)中对极平面 $\mathbf{CC'X}$ 与曲面相切于 $\mathbf{X}$．在两视图中影像轮廓线与对极线在 $\mathbf{x}$ 和 $\mathbf{x}'$ 处相切．曲面上的虚线是边界生成元．在(b)中对极平面与该空间曲线相切．对应的对极线 $\mathbf{l}\leftrightarrow\mathbf{l}'$ 与该曲线的影像相切．

## 10.9　退化

如果点对应集 $\{\mathbf{x}_i\leftrightarrow\mathbf{x}_i', i=1,\ldots,n\}$ 不能唯一地确定极几何，或等价地说，如果存在线性独立的秩为 2 的矩阵 $\mathrm{F}_j, j=1,2$，使得

$$\mathbf{x}_i'^{\mathrm{T}}\mathrm{F}_1\mathbf{x}_i = 0 \quad 和 \quad \mathbf{x}_i'^{\mathrm{T}}\mathrm{F}_2\mathbf{x}_i = 0 \qquad (1\leq i\leq n).$$

则称它在几何上相对于 F 是退化的．

第 21 章将对退化主题作详细地讨论．这里仅对景物点在直纹二次曲面或平面上的两种重要情形作一个扼要的介绍．

如果两个摄像机中心不重合，那么对极几何可唯一确定．它可以采用诸如(8.1－$p$163)的方法由摄像机矩阵 P，P′计算得到．这里主要介绍对极几何不能从点对应计算出来的那些结构．了解估计算法中的退化结构很重要，因为“接近于”退化的结构将有可能导致数值上的病态估计．表 10.1 概括了这些退化情形．

| |
|---|
| dim($N$)＝1：　唯一解——非退化情形． |
| 发生于一般位置的 $n\geq 8$ 组点对应．如果 $n>8$，那么点对应必须是完全精确的(即无噪声)． |
| dim($N$)＝2：　1 或 3 个解． |
| 发生于 7 组对应的情形，同时也发生于 $n>7$ 完全精确的点对应的情形，其中 3D 点和摄像机中心在一个称为临界曲面的直纹二次曲面上．二次曲面可以是非退化的(单叶双曲面)或退化的． |
| dim($N$)＝3：　两参数的解簇． |
| 发生于与一个单应 $\mathbf{x}_i'=\mathrm{H}\mathbf{x}_i$ 相关联的 $n\geq 6$ 组完全精确的点对应． |
| • 绕摄像机中心旋转(一种退化的运动)． |
| • 所有的世界点在一个平面上(一种退化的结构)． |

表 10.1　由点对应估计 F 时的退化情形，它们按矩阵 A 的零空间 $N$ 的维数分类．

### 10.9.1 点在直纹二次曲面上

第21章将证明如果两个摄像机中心和所有3D点都在一个(直纹)二次曲面上(称为**临界曲面**[Maybank-93]),则退化将发生. 直纹二次曲面可以是非退化的(单叶双曲面—冷却塔)或者退化的(如两平面,圆锥面和圆柱面)——参见节2.2.4($p$41);但是临界曲面不可能是椭圆面或者双叶双曲面. 一个临界面配置可能有三个基本矩阵.

注意如果刚好是7组点对应,那么加上两个摄像机中心则共有9个点. 一般的二次曲面有9个自由度,并且过9点我们总可以构造一个二次曲面. 当二次曲面是直纹二次曲面时,它是一个临界曲面并且F可能有三个解. 当二次曲面不是直纹面时F仅有一个实解.

### 10.9.2 点在平面上

8点算法的一个重要失败模式是所有点在一个平面上. 注意在这种退化情形中,所有点和两个摄像机中心都在一个直纹二次曲面上,即在由过这些点的平面和过两个摄像机中心的平面组成的退化二次曲面上.

平面点集的两幅视图通过一个2D射影变换H相联系. 因此,假定给定对应点集$\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$使得$\mathbf{x}_i' = \mathrm{H}\mathbf{x}_i$. 就可以给定任意多个点$\mathbf{x}_i$及其对应点$\mathbf{x}_i' = \mathrm{H}\mathbf{x}_i$. 这一对摄像机的基本矩阵满足方程$\mathbf{x}_i'^{\mathsf{T}}\mathrm{F}\mathbf{x}_i = \mathbf{x}_i'^{\mathsf{T}}(\mathrm{FH}^{-1})\mathbf{x}_i' = 0$. 只要$\mathrm{FH}^{-1}$是反对称矩阵,这方程组就成立. 因此,F的解是形为F=SH的任何矩阵,其中S是反对称矩阵. 一个3×3反对称矩阵S可以记成$\mathrm{S}=[\mathbf{t}]_\times$的形式,矢量$\mathbf{t}$是任意3维矢量. 因此,S有三个自由度,从而F也如此. 更准确地说,这些点对应$\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$给出基本矩阵F的一个三参数簇(注:一个参数为矩阵的尺度因子,因此仅仅是一个两参数簇的齐次矩阵). 这样一来,由这个对应集得到的方程组的系数矩阵A的秩必然不超过6.

由F=SH的分解和结论8.9($p$171)得到的对应于基本矩阵F的摄像机矩阵对是$[\mathrm{I}|\mathbf{0}]$和$[\mathrm{H}|\mathbf{t}]$. 这里,矢量$\mathbf{t}$可取任意值. 如果点$\mathbf{x}_i = (x_i, y_i, 1)^{\mathsf{T}}$, $\mathbf{x}_i' = \mathrm{H}\mathbf{x}_i$,则可以验证点$\mathbf{X}_i = (x_i, y_i, 1, 0)^{\mathsf{T}}$通过这两个摄像机矩阵映射到$\mathbf{x}_i$和$\mathbf{x}_i'$. 因而,点$\mathbf{X}_i$构成了景物的一个重构.

### 10.9.3 无平移情形

如果两个摄像机的中心重合,则对极几何没有定义. 此外,诸如结论8.9($p$171)的公式都赋给基本矩阵一个0值. 在这种情形下,两幅图像通过一个2D单应(见节7.4.2$p$141)相联系.

如果仍试图寻求这个基本矩阵,则如上所述,F的解至少是一个2参数的簇. 即使摄像机没有做平移运动,用来计算基本矩阵的方法如8点算法仍将产生满足$\mathbf{x}_i'^{\mathsf{T}}\mathrm{F}\mathbf{x}_i = 0$的一个矩阵F,其形式为F=SH,其中H是联系这些点的单应,而S是本质上任意的反对称矩阵. 由H关联的点$\mathbf{x}_i$和$\mathbf{x}_i'$将满足这个关系.

## 10.10 F计算的几何解释

由图像点对应集$\{\mathbf{x}_i \leftrightarrow \mathbf{x}_i'\}$估计F与由一组2D点$\{x_i, y_i\}$估计二次曲线(或由一组3D点

估计二次曲面)的问题有许多相似之处.

方程 $\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0$ 给出关于 $x,y,x',y'$ 的一个约束,因此在 $\mathbb{R}^4$ 空间定义了余维数为 1(维数为 3)的一个曲面(簇)$\mathcal{V}$. 该曲面是二次曲面,因为该方程关于 $\mathbb{R}^4$ 中的坐标 $x,y,x',y'$ 是二次的. 有一个从 3 维射影空间到簇 $\mathcal{V}$ 的自然映射,把任何 3D 点映射为两幅视图上图像对应点的四元组 $(x,y,x',y')^{\mathrm{T}}$. 如果将 $\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0$ 写成如下形式,

$$(x \quad y \quad x' \quad y' \quad 1)\begin{bmatrix} 0 & 0 & f_{11} & f_{21} & f_{31} \\ 0 & 0 & f_{12} & f_{22} & f_{32} \\ f_{11} & f_{12} & 0 & 0 & f_{13} \\ f_{21} & f_{22} & 0 & 0 & f_{23} \\ f_{31} & f_{32} & f_{13} & f_{23} & 2f_{33} \end{bmatrix}\begin{pmatrix} x \\ y \\ x' \\ y' \\ 1 \end{pmatrix} = 0.$$

则它显然是二次形式. 二次曲线拟合是估计 F 的一个适宜的(低维)模型. 为指出这两种估计问题的相似性:如节 1.2.3($p6$)所介绍,一个点 $(x_i,y_i)^{\mathrm{T}}$ 对二次曲线的 5 个自由度提供了一个约束:

$$ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i + f = 0.$$

类似地,如(10.2,$p191$),一组点对应 $(x_i,y_i,x_i',y_i')^{\mathrm{T}}$ 为 F 的 8 个自由度提供下列约束:

$$x_i'x_if_{11} + x_i'y_if_{12} + x_i'f_{13} + y_i'x_if_{21} + y_i'y_if_{22} + y_i'f_{23} + x_if_{31} + y_if_{32} + f_{33} = 0.$$

但两者不完全类似,因为由基本矩阵定义的关系仅出现两个集合的指标的双线性项,这一点也可从上面的二次型矩阵的零元素明显看出,然而二次曲线的方程是任意的二次方程. 同时,由 F 定义的曲面还必须满足由 $\det(\mathrm{F})=0$ 产生的一个约束,而在二次曲线拟合中却没有这样的约束.

我们知道当数据只能与二次曲线的一小部分拟合时会出现外插问题,基本矩阵的拟合也有类似的问题. 在不少情况下数据足以精确地确定二次曲线的一条切线,但不足以确定二次曲线本身,见图 10.7. 在基本矩阵的情形,$\mathbb{R}^4$ 中二次曲面的切平面是仿射基本矩阵(第 13 章),当透视效果比较小时,这种近似是适宜的.

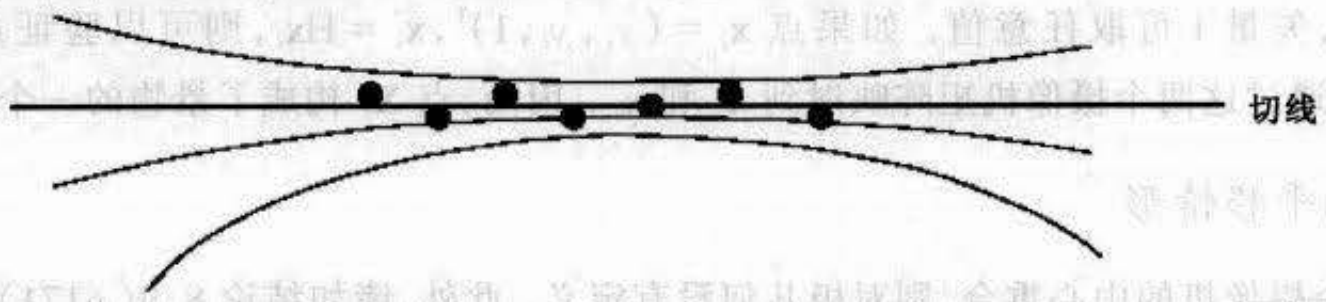



图 10.7 由点数据(表示为 •)估计一条二次曲线的条件可能不足. 所有显示的二次曲线的残差都在点的误差分布之内. 即便如此,二次曲线的估计仍存在多义性,但切线是良定的,并且可以根据这些点计算得到.

## 10.11 对极线的包络

基本矩阵的一个应用是在第二幅图像上确定与第一幅图像上的点对应的对极线. 例如,如果我们正在两幅图像上搜索匹配点,第一幅图像上的给定点 $\mathbf{x}$ 的匹配点可以在第二幅图像上沿着对极线 $\mathrm{F}\mathbf{x}$ 搜索. 当然在噪声存在时,匹配点不一定就在 $\mathrm{F}\mathbf{x}$ 上,因为我们仅知道基本矩阵在由协方差矩阵表示的某个范围内. 在一般情况下,不能只沿对极线 $\mathrm{F}\mathbf{x}$ 搜索,而有必要在对极线 $\mathrm{F}\mathbf{x}$ 的两边的一个区域内搜索. 我们现在来考虑怎样利用基本矩阵的协方

差矩阵来确定这个搜索区域.

令 $\mathbf{x}$ 是一个点而 F 是一个基本矩阵并已求得它的协方差矩阵 $\Sigma_F$. 点 $\mathbf{x}$ 与一条对极线 $\mathbf{l}=F\mathbf{x}$ 对应,并且可以根据结论 4.6($p91$)把协方差矩阵 $\Sigma_F$ 转移为协方差矩阵 $\Sigma_l$. 再根据结论 4.6($p91$),由 $\bar{\mathbf{l}}=\bar{F}\mathbf{x}$ 给出对极线的平均值. 为避免出现奇异的情形,把表示对极线的矢量 $\mathbf{l}$ 归一化为 $\|\mathbf{l}\|=1$. 则映射 $\mathbf{x}\mapsto\mathbf{l}$ 由 $\mathbf{l}=(F\mathbf{x})/\|F\mathbf{x}\|$ 给出. 如果 J 是该映射关于 F 的元素的 Jacobian 矩阵. 则 J 是一个 $3\times9$ 矩阵,且 $\Sigma_l=J\Sigma_F J^T$.

虽然约束 $\|\mathbf{l}\|=1$ 是最方便的约束,但下面的分析适用于限制表示对极线的矢量在 $\mathbb{R}^3$ 的一个 2维曲面上变化的任何约束. 这种情形下的协方差矩阵 $\Sigma_l$ 是奇异的,并且秩为 2,因为在约束曲面的法线方向上不允许变化. 举 $\mathbf{l}$ 的一个具体例子,均值的偏差 $(\bar{\mathbf{l}}-\mathbf{l})$ 必须沿着约束曲面,从而(在线性逼近中)与 $\Sigma_l$ 的零空间相垂直.

在剩下的推导中,为了避免标记的混乱,我们用 $\mathbf{m}$ 来表示对极线的均值矢量 $\bar{\mathbf{l}}$. 现在,假设对极线矢量 $\mathbf{l}$ 服从 Gauss 分布,所有直线集合的似然由下列方程给定:

$$(\mathbf{l}-\mathbf{m})^T\Sigma_l^+(\mathbf{l}-\mathbf{m})=k^2 \tag{10.11}$$

其中 $k$ 是某个常数. 为作进一步分析,我们运用坐标的正交变换将 $\Sigma_l$ 变为对角矩阵. 则可以写成

$$U\Sigma_l U^T=\Sigma_l'=\begin{bmatrix}\tilde{\Sigma}_l' & \mathbf{0}\\ \mathbf{0}^T & 0\end{bmatrix}$$

其中 $\tilde{\Sigma}_l'$ 是一个 $2\times2$ 非奇异对角矩阵. 对直线运用同一变换,可定义 2 维矢量 $\mathbf{m}'=U\mathbf{m}$ 和 $\mathbf{l}'=U\mathbf{l}$. 因 $\mathbf{l}'-\mathbf{m}'$ 与 $\Sigma_l$ 的零空间 $(0,0,1)^T$ 正交,故 $\mathbf{m}'$ 和 $\mathbf{l}'$ 的第三个坐标相同. 我们可以假定这个坐标是 1,如果必要的话,把 U 乘以一个常数,因此,对某两个 2 维矢量 $\tilde{\mathbf{l}}'$ 和 $\tilde{\mathbf{m}}'$ 我们可以记 $\mathbf{l}'=(\tilde{\mathbf{l}}'^T,1)^T$ 和 $\mathbf{m}'=(\tilde{\mathbf{m}}'^T,1)^T$. 可以验证

$$\begin{aligned}k^2&=(\mathbf{l}-\mathbf{m})^T\Sigma_l^+(\mathbf{l}-\mathbf{m})\\&=(\mathbf{l}'-\mathbf{m}')^T\Sigma_l'^+(\mathbf{l}'-\mathbf{m}')\\&=(\tilde{\mathbf{l}}'-\tilde{\mathbf{m}}')^T\tilde{\Sigma}_l'^{-1}(\tilde{\mathbf{l}}'-\tilde{\mathbf{m}}')\end{aligned}$$

将此方程展开:

$$\tilde{\mathbf{l}}'^T\tilde{\Sigma}_l'^{-1}\tilde{\mathbf{l}}'-\tilde{\mathbf{m}}'^T\tilde{\Sigma}_l'^{-1}\tilde{\mathbf{l}}'-\tilde{\mathbf{l}}'^T\tilde{\Sigma}_l'^{-1}\tilde{\mathbf{m}}'+\tilde{\mathbf{m}}'^T\tilde{\Sigma}_l'^{-1}\tilde{\mathbf{m}}'-k^2=0$$

并可以记为:

$$(\tilde{\mathbf{l}}'^T\ \ 1)\begin{bmatrix}\tilde{\Sigma}_l'^{-1} & -\tilde{\Sigma}_l'^{-1}\tilde{\mathbf{m}}'\\ -\tilde{\mathbf{m}}'^T\tilde{\Sigma}_l'^{-1} & \tilde{\mathbf{m}}'^T\tilde{\Sigma}_l'^{-1}\tilde{\mathbf{m}}'-k^2\end{bmatrix}\begin{pmatrix}\tilde{\mathbf{l}}'\\1\end{pmatrix}=0$$

或等价于(可以验证):

$$(\tilde{\mathbf{l}}'^T\ \ 1)\begin{bmatrix}\tilde{\mathbf{m}}'\tilde{\mathbf{m}}'^T-k^2\tilde{\Sigma}_l' & \tilde{\mathbf{m}}'\\ \tilde{\mathbf{m}}'^T & 1\end{bmatrix}^{-1}\begin{pmatrix}\tilde{\mathbf{l}}'\\1\end{pmatrix}=0. \tag{10.12}$$

最后,上式又等价于

$$\mathbf{l}'^T[\mathbf{m}'\mathbf{m}'^T-k^2\Sigma_l']^{-1}\mathbf{l}'=0. \tag{10.13}$$

以上证明满足(10.11)的直线构成由矩阵 $(\mathbf{m}'\mathbf{m}'^T-k^2\Sigma_l')^{-1}$ 定义的线二次曲线. 这些线的包络是由矩阵 $\mathbf{m}'\mathbf{m}'^T-k^2\Sigma_l'$ 定义的点二次曲线. 现在可以再变回原坐标系来确定在原坐

标系下这些直线的包络. 变换后的二次曲线是:

$$C = U^T(\mathbf{m}'\mathbf{m}'^T - k^2\Sigma_l')U = \mathbf{m}\mathbf{m}^T - k^2\Sigma_l. \tag{10.14}$$

注意:当 $k=0$ 时,二次曲线 C 退化为 $\mathbf{m}\mathbf{m}^T$,表示在直线 $\mathbf{m}$ 上的点集合. 当 $k$ 增大时,二次曲线变为双曲线,其两个分支分别在直线 $\mathbf{m}$ 的两边.

假如我们想通过选取 $k$ 使得极线的某个比例 $\alpha$ 在这个双叶双曲线围成的区域内. (10.11)中的值 $k^2 = (\mathbf{l}-\mathbf{m})^T\Sigma_l^+(\mathbf{l}-\mathbf{m})$ 服从 $\chi_n^2$ 分布,累积 $\chi_n^2$ 分布 $F_n(k^2) = \int_0^{k^2}\chi_n^2(\zeta)d\zeta$ 表示一个 $\chi_n^2$ 随机变量的值小于 $k^2$ 的概率($\chi_n^2$ 和 $F_n$ 分布在节 A2.2($p$403)中定义). 把它运用于一条随机直线 $\mathbf{l}$,我们可以看出为了保证直线的某个比例 $\alpha$ 在由(10.14)定义的双曲线围成的区域中间. 必须选取 $k^2$ 使得 $F_2(k^2)=\alpha$($n=2$,因为协方差矩阵 $\Sigma_l$ 的秩为 2). 因此 $k^2 = F_2^{-1}(\alpha)$,例如当 $\alpha=0.95$ 时,可求得 $k^2=5.9915$. 由(10.14)给出的对应的双曲线是 $C = \mathbf{m}\mathbf{m}^T - 5.9915\Sigma_l$. 总结以上讨论得:

**结论 10.2　如果 $\mathbf{l}$ 是服从高斯分布的随机直线并具有均值 $\bar{\mathbf{l}}$ 和秩 2 的协方差矩阵 $\Sigma_l$,则平面二次曲线**

$$C = \bar{\mathbf{l}}\bar{\mathbf{l}}^T - k^2\Sigma_l \tag{10.15}$$

**表示一个等似然边界,它把直线 $\mathbf{l}$ 所有事件中占某个比例的事件围住. 如果 $F_2^{-1}(k^2)$ 表示累积 $\chi_2^2$ 分布,且选择的 $k^2$ 满足 $F_2^{-1}(k^2)=\alpha$,则占所有直线的 $\alpha$ 比例部分在由 C 所围的区域内. 也就是说这些直线落在这个区域的概率是 $\alpha$.**

运用这个公式时必须清楚它仅表示一个近似,因为对极线不是正态分布. 我们总假定分布可以利用 Jacobian 正确地变换,但这是一种线性的假设. 这个假设对于方差小且接近于均值的分布是最为合理的. 这里,我们应用它寻找 95% 样本所在的区域,即几乎是误差分布的全体. 在这种情况下,误差是高斯分布的假设是欠合理的.

### 10.11.1　对极线协方差的验证

我们现在提供对极几何包络的一些例子以证实和说明上面推导的理论. 但在此之前,我们将直接验证关于对极线的协方差矩阵的理论. 因为一条直线的 $3\times3$ 协方差矩阵在数量上不太容易理解,我们考虑对极线方向的方差. 给定一条直线 $\mathbf{l}=(l_1,l_2,l_3)^T$,表示其方向的角度由 $\theta=\arctan(-l_1/l_2)$ 给出. 令 J 等于映射 $\mathbf{l}\mapsto\theta$ 的 $1\times3$ 的 Jacobian 矩阵,可以发现角度 $\theta$ 的方差是 $\sigma_\theta^2 = J\Sigma_l J^T$. 这个结论可以通过如下的模拟来验证.

设一对图像中的点对应已经确定. 由这些点对应计算基本矩阵,再把这些点矫正,使得它们在对极几何映射下精确地对应(如在节 11.3 所描述). 再用 $n$ 组这样矫正过的对应来计算基本矩阵 F 的协方差矩阵. 然后,用第一幅图像上另外一组矫正过的"测试"对应点 $\mathbf{x}_i$ 算出对应对极线 $\mathbf{l}_i' = F\mathbf{x}_i$ 的均值和协方差,从而算出该对极线方向角的均值和方差. 由此给出了这些量的理论值.

下一步,进行 Moute Carlo 模拟,对用来计算 F 的点的坐标加入高斯噪声. 用计算得到的 F 求每一个测试点所对应的对极线和它们相应的角度以及角度与均值的偏差. 如此进行多次,便可算出该角度的标准偏差,最后将它同理论值相比较. 对图 10.2($p$198)中的雕塑的图像对计算的结果在图 10.8 中给出.

**雕塑图像的对极线包络** 图 10.2($p$198)的雕塑图像对很有意义,因为整个图像的深度变化很大. 在图像上出现近点(在雕像上)和远点(在后面的建筑物上)非常靠近的现象. 基本矩阵已由若干点计算得到. 选定在第一幅图像上的一个点(见图 10.9)且用 Moute Carlo 模拟来计算在每组匹配点坐标的噪声在 0.5 个像素的水平时的几条可能的对极线. 为了检验这个理论,我们在理论上计算了对极线的均值和协方差. 算出 95% 对极线包络并在第二幅图像上画出. 图 10.10 显示了用不同点数计算 F 的结果. 当 $n=15$ 时,该 95% 包络非常接近于模拟的直线包络.

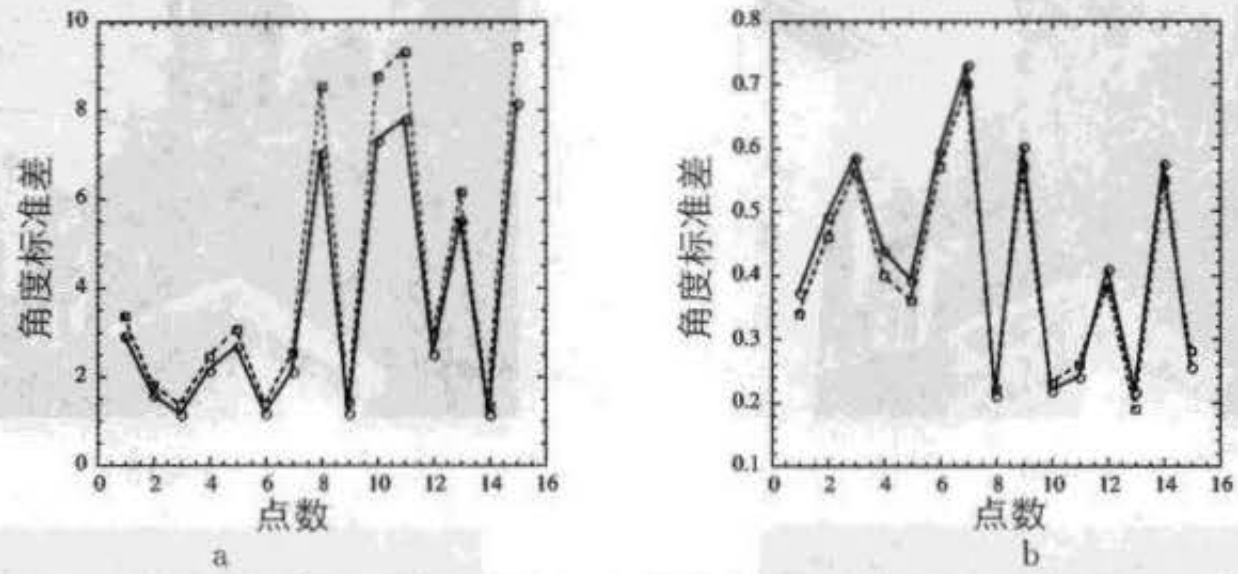



图 10.8 对图 10.2($p$198)中的雕塑图像对中的 15 个测试点的对极线的方向角进行理论值和 Monte Carlo 模拟值的比较. 水平轴表示点数(1 到 15)而垂直轴表示角度标准差. (a)对极几何结构(基本矩阵)由 15 组点匹配算出时的结果. (b)由 50 组点匹配算出时的结果. **注意**:这些图的水平轴表示 1 到 15 个离散点. 这些图用连续曲线表示仅仅为了视觉上更清晰.

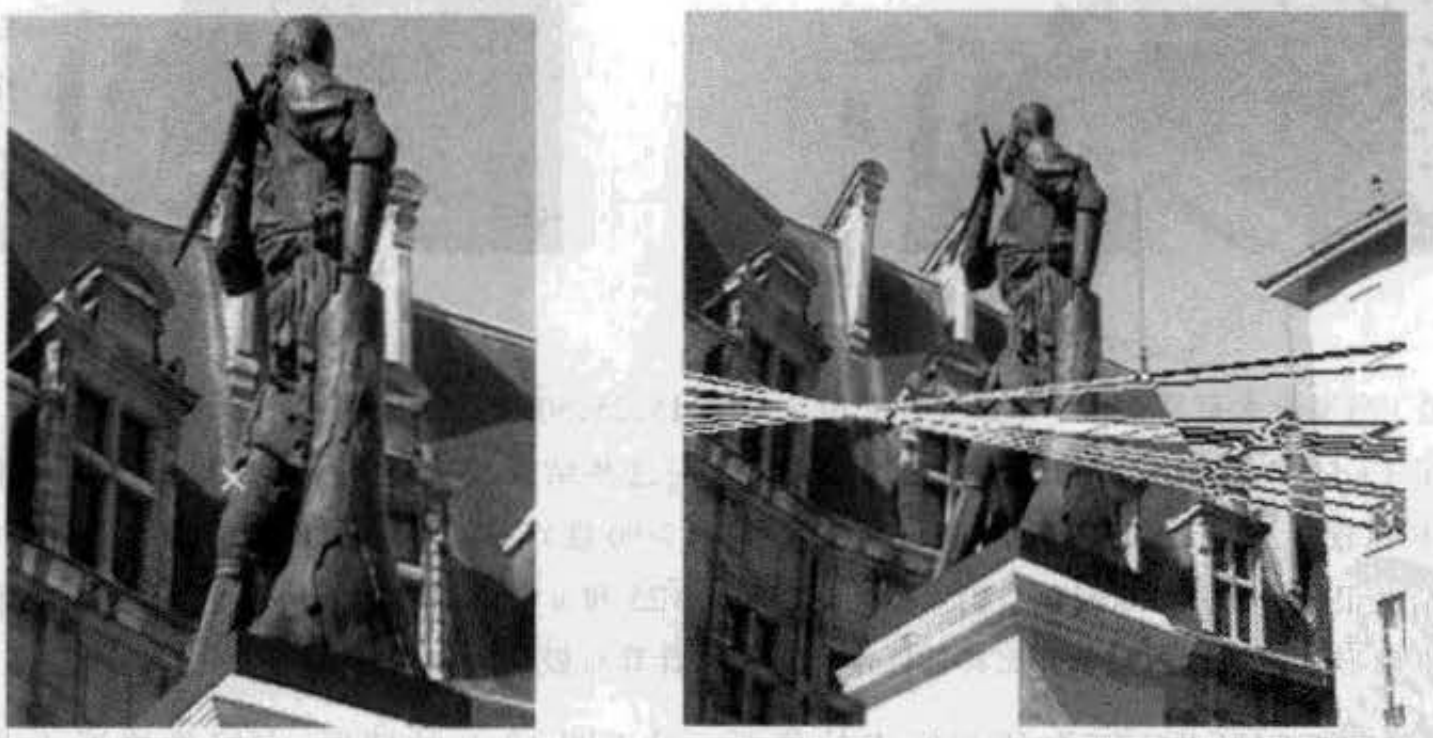



图 10.9 (a)第一幅图像上用于计算第二幅图像的对极线包络的点. 注意在第二幅图上找出的是哪个点出现了多义性. 所标记的点可以表示雕塑腿上的点(前景)或雕塑后面的大楼上的点(背景). 在第二幅图像上,这两点分得很开,并且对极线必须都通过它们. (b)由 $n=15$ 组点匹配计算得到的对极线. 不同的直线对应于匹配点注入不同噪声的事件. 在计算点所对应的对极线之前,给理想匹配点位置的每一坐标加了 0.5 像素的高斯噪声. 用 ML 估计算法(黄金标准算法)计算 F. 这个实验证明由少数的点来计算对极线基本上具有不稳定性. 为寻找左边图像上被选定的点的匹配点,我们需要搜索这些对极线所覆盖的区域.

图 10.10 显示的结果表明在点的匹配中计算对极线包络有重要实际意义. 假定人们试

图寻找图 10.9 中前景点的对应. 如果仅用 10 组匹配点来计算对极线,那么在给定的包络宽度上作对极线搜索不大可能成功. 即使 $n=15$ 时,在正确匹配的水平上的包络的宽度还有几十个像素. 当 $n=25$,情形就比较满意. 注意这种不稳定属于这个问题的内在性质,而不是任何计算 F 的特定算法的后果.

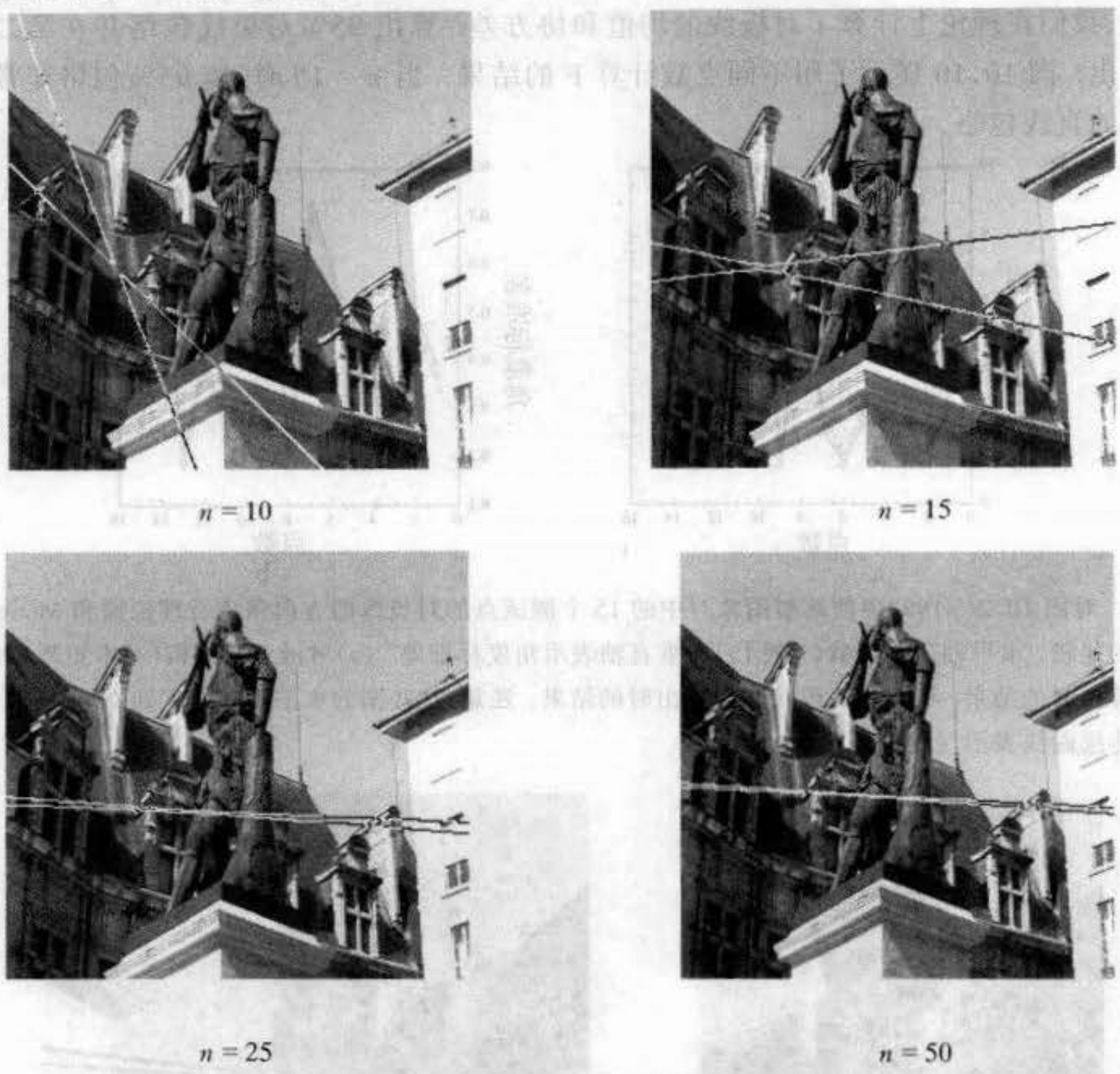



图 10.10 当噪声水平是 0.5 个像素并且 F 由 $n=10,15,25,50$ 个点计算时对极线的 95% 包络. 对每种情形,Monte Carlo 模拟的结果(虽然没有显示在这里)与这些结果很接近. 可把这里的 $n=15$ 的情形与图 10.9相比较. 注意当 $n=10$ 时,对极线包络非常宽(>90 度),表明对这种情形由 10 个点计算得到的对极线置信度很小. 当 $n=15$ 时包络仍然相当宽. 当 $n=25$ 和 $n=50$ 时对极线可以相当准确地得到. 当然,包络的准确形状强烈地依赖于什么样的匹配点被用来计算对极几何结构.

一个有趣问题是关于包络最窄点的位置. 对于图 10.9 的情形,它好像接近于背景点的正确匹配位置. 前景点(雕像的脚)的匹配远离这个包络的最窄点. 虽然我们对包络窄点的准确位置不能充分理解,好像此情形是因为用来计算 F 的多数点在背景的建筑物上. 这就给其他接近建筑物的平面上的匹配点增加更多的筹码. 对于深度很不同的点的匹配准确性要差.

**靠近对极点的匹配点——走廊场景** 当要匹配的点靠近对极点时,对极线的确定更不稳定,因为任何关于对极点位置的不确定性都导致对极线斜率的不确定性. 另外,当接近这个不稳定位置时,(10.14)推导的线性近似变得无法接受. 特别是对极线的分布将偏离正态分布.

## 10.12 图像矫正

本节给出一种图像矫正的方法，它是对视点差别很大的双眼图像对进行重新采样的过程，目的是为了产生一对“匹配对极线的投影”. 这些投影的对极线平行于 $x$ 轴并且在视图之间相匹配，因此，图像之间的视差仅发生在 $x$ 方向，即 $y$ 方向没有视差.

这个方法建立在基本矩阵之上. 为了使对极线匹配，对两幅图像施加两个 2D 射影变换. 业已证明可以通过选择两个变换使得匹配点几乎有相同的 $x$ 坐标. 用这种方式使两幅图像(如果重叠在一起)尽可能地对应，并且任何视差都平行于 $x$ 轴. 因为施加任意的 2D 射影变换可能造成图像极大的失真，寻找变换对的方法要满足使图像失真最小的条件.

事实上，对两幅图像进行适当的射影变换的问题可化简成由一对并排放置且主轴相互平行的相同摄像机所产生的对极几何问题. 在以前的文献中所介绍的许多双眼匹配算法都假定了这个几何. 经过这样的矫正后，根据简单的对极结构和两幅图像的相邻对应，匹配点的搜索也极大地简化了，它可以作为复杂图像匹配的预备步骤.

### 10.12.1 映射对极点到无穷远点

本节我们讨论求把一幅图像的对极点映射到无穷远点的射影变换 H 的问题. 事实上，如果对极线被变换为平行于 $x$ 轴的直线，则对极点就被映射到特定的无穷远点 $(1,0,0)^{\mathsf{T}}$. 这个问题给 H 敞开了许多自由度(实际是 4 个)，而且如果选择了不恰当的 H，图像将发生严重的射影失真. 为了使重新采样的图像与原始图像看起来一样，我们可以对 H 的选取施加严格的限制.

一个导致好结果的条件是要求变换 H 对图像给定的选择点 $\mathbf{x}_0$ 的邻域的作用应该尽可能地是一个刚性变换. 这意味着在一阶近似程度上，$\mathbf{x}_0$ 的邻域只被旋转和平移，因而原图像和重新取样的图像看起来是相同的. 取点 $\mathbf{x}_0$ 为图像的中心是一种适当的取法. 例如，在航空摄影中如果已知视点不是过分倾斜，这会是一种好的选择.

现在，假定 $\mathbf{x}_0$ 是原点且对极点 $\mathbf{e}=(f,0,1)^{\mathsf{T}}$ 在 $x$ 轴上. 下面的变换：

$$G=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -1/f & 0 & 1\end{bmatrix} \tag{10.16}$$

把对极点 $(f,0,1)^{\mathsf{T}}$ 变到无穷远点 $(f,0,0)^{\mathsf{T}}$，这正是我们所需要的. 点 $(x,y,1)^{\mathsf{T}}$ 被 G 映射到点 $(\hat{x},\hat{y},1)^{\mathsf{T}}=(x,y,1-x/f)^{\mathsf{T}}$. 若 $|x/f|<1$，则：

$$(\hat{x},\hat{y},1)^{\mathsf{T}}=(x,y,1-x/f)^{\mathsf{T}}=(x(1+x/f+\ldots),y(1+x/f+\ldots),1)^{\mathsf{T}}.$$

其 Jacobi 矩阵是

$$\frac{\partial(\hat{x},\hat{y})}{\partial(x,y)}=\begin{bmatrix}1+2x/f & 0\\ y/f & 1+x/f\end{bmatrix}$$

外加 $x$ 和 $y$ 的高阶项. 现在，若 $x=y=0$，则它是一个恒等映射. 也就是说，G 在原点近似(一阶近似)于恒等映射.

对于任意位置上的感兴趣的点 $\mathbf{x}_0$ 和对极点 $\mathbf{e}$，所需要的映射 H 是乘积 $H=GRT$，其中

T 是把点 $\mathbf{x}_0$ 映到原点的平移，R 是一个绕原点的旋转使对极点 $\mathbf{e}'$ 转到 $x$ 轴上点 $(f,0,1)^\mathsf{T}$，而 G 是刚讨论过的把 $(f,0,1)^\mathsf{T}$ 变到无穷远点的映射. 这个复合映射是 $\mathbf{x}_0$ 邻域内的一个刚性变换的一阶近似.

### 10.12.2 匹配变换

上节给出了如何把一幅图像上的对极点映射到无穷远点的方法. 接下来，我们将看到怎样把一个映射施加到另一幅图像上使得对极线互相匹配. 我们考虑两幅图像 $J$ 和 $J'$. 我们的意图是对这两幅图像重新采样，使之产生作用于 $J$ 的变换 H 和作用于 $J'$ 的变换 H′. 重采样的目的是使在 $J$ 中的对极线与在 $J'$ 中对应的对极线匹配. 更具体地说，如果 $\mathbf{l}$ 和 $\mathbf{l}'$ 是两幅图像中任何一对对应的对极线，则 $\mathrm{H}^{-\mathsf{T}}\mathbf{l} = \mathrm{H}'^{-\mathsf{T}}\mathbf{l}'$（回忆 $\mathrm{H}^{-\mathsf{T}}$ 是对应于点映射 H 的线映射）. 满足这个条件的任何变换对将称为一个变换**匹配对**.

我们选择变换匹配对的策略是：如上节所述，先选择某个变换 H′ 把对极点 $\mathbf{e}'$ 映到无穷远点. 然后寻求最小化距离平方和

$$\sum_i d(\mathrm{H}\mathbf{x}_i, \mathrm{H}'\mathbf{x}_i')^2 \tag{10.17}$$

的匹配变换 H. 首先要解决的问题是如何寻找一个匹配 H′ 的变换. 这个问题由下面的结论回答.

**结论 10.3** **令 J 和 J′ 是具有基本矩阵 $\mathrm{F} = [\mathbf{e}']_\times \mathrm{M}$ 的两幅图像，而 H′ 是 J′ 的一个射影变换. 则 J 的一个射影变换 H 与 H′ 匹配的充要条件是对某个矢量 $\mathbf{a}$，H 具有形式**

$$\mathrm{H} = (\mathrm{I} + \mathrm{H}'\mathbf{e}'\mathbf{a}^\mathsf{T})\mathrm{H}'\mathrm{M}. \tag{10.18}$$

**证明** 如果 $\mathbf{x}$ 是 J 上的一个点，则 $\mathbf{e}\times\mathbf{x}$ 是在第一幅图像上的对极线，而 $\mathrm{F}\mathbf{x}$ 是在第二幅图像上的对极线. 变换 H 和 H′ 是一个匹配对的充要条件是 $\mathrm{H}^{-\mathsf{T}}(\mathbf{e}\times\mathbf{x}) = \mathrm{H}'^{-\mathsf{T}}\mathrm{F}\mathbf{x}$. 因为该式必须对所有的点 $\mathbf{x}$ 成立，故可等价地写为 $\mathrm{H}^{-\mathsf{T}}[\mathbf{e}]_\times = \mathrm{H}'^{-\mathsf{T}}\mathrm{F} = \mathrm{H}'^{-\mathsf{T}}[\mathbf{e}']_\times \mathrm{M}$，或者应用结论 A3.3（$p$407），得

$$[\mathrm{H}\mathbf{e}]_\times \mathrm{H} = [\mathrm{H}'\mathbf{e}']_\times \mathrm{H}'\mathrm{M} \tag{10.19}$$

根据引理 8.11（$p$172），由此推出所需要的 $\mathrm{H} = (\mathrm{I} + \mathrm{H}'\mathbf{e}'\mathbf{a}^\mathsf{T})\mathrm{H}'\mathrm{M}$.

下面证明逆命题，如果(10.18)成立，则

$$\begin{aligned}\mathrm{H}\mathbf{e} &= (\mathrm{I} + \mathrm{H}'\mathbf{e}'\mathbf{a}^\mathsf{T})\mathrm{H}'\mathrm{M}\mathbf{e} = (\mathrm{I} + \mathrm{H}'\mathbf{e}'\mathbf{a}^\mathsf{T})\mathrm{H}'\mathbf{e}' \\ &= (1 + \mathbf{a}^\mathsf{T}\mathrm{H}'\mathbf{e}')\mathrm{H}'\mathbf{e}' = \mathrm{H}'\mathbf{e}'.\end{aligned}$$

它和(10.18)一起足以证明(10.19)成立，因此 H 和 H′ 是匹配变换. □

我们对把 $\mathbf{e}'$ 变到无穷远点 $(1,0,0)^\mathsf{T}$ 的变换 H′ 特别感兴趣. 在这种情形，矩阵 $\mathrm{I} + \mathrm{H}'\mathbf{e}'\mathbf{a}^\mathsf{T} = \mathrm{I} + (1,0,0)^\mathsf{T}\mathbf{a}^\mathsf{T}$ 有如下形式

$$\mathrm{H}_\mathrm{A} = \begin{bmatrix} a & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{10.20}$$

它表示一个仿射变换. 因而，结论 10.3 的一种特殊情形是：

**推论 10.4** **令 J 和 J′ 是具有基本矩阵 $\mathrm{F} = [\mathbf{e}']_\times \mathrm{M}$ 的图像，且 H′ 是 J′ 的把对极点 $\mathbf{e}'$ 映到无穷远点 $(1,0,0)^\mathsf{T}$ 的一个射影变换. 则 J 的一个变换 H 与 H′ 匹配的充要条件是 H 的形式为 $\mathrm{H} = \mathrm{H}_\mathrm{A}\mathrm{H}_0$，其中 $\mathrm{H}_0 = \mathrm{H}'\mathrm{M}$，而 $\mathrm{H}_\mathrm{A}$ 是形如(10.20)的一个仿射变换.**

给定把对极点映射到无穷远点的 H′，我们可以利用这个推论选择使视差最小化的一个匹配变换 H. 记 $\hat{\mathbf{x}}_i' = \mathrm{H}'\mathbf{x}_i'$ 和 $\hat{\mathbf{x}}_i = \mathrm{H}_0\mathbf{x}_i$. 则最小化问题(10.17)要求形如(10.20)的 $\mathrm{H}_\mathrm{A}$ 使得

$$\sum_i d(\mathrm{H}_\mathrm{A}\hat{\mathbf{x}}_i, \hat{\mathbf{x}}_i')^2 \tag{10.21}$$

最小化.

具体地说，令 $\hat{\mathbf{x}}_i = (\hat{x}_i, \hat{y}_i, 1)^\mathrm{T}$，$\hat{\mathbf{x}}_i' = (\hat{x}_i', \hat{y}_i', 1)^\mathrm{T}$. 因为 H′和 M 已知，这些矢量可以从匹配点 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$ 算出. 于是使(10.21)最小化的量可以表示为

$$\sum_i (a\hat{x}_i + b\hat{y}_i + c - \hat{x}_i')^2 + (\hat{y}_i - \hat{y}_i')^2$$

因 $(\hat{y}_i - \hat{y}_i')^2$ 是常数，它又等价于最小化

$$\sum_i (a\hat{x}_i + b\hat{y}_i + c - \hat{x}_i')^2$$

这是一个简单的线性最小二乘参数最小化问题，利用线性方法(见 A3.3.1($p$409))可以容易地求得 $a, b, c$. 则 $\mathrm{H}_\mathrm{A}$ 可由(10.20)算出，而 H 由(10.18)算出. 注意一个线性解是可能得到的，因为 $\mathrm{H}_\mathrm{A}$ 是一个仿射变换. 如果它只是一个射影变换，那就不是线性问题了.

### 10.12.3 算法概述

现在总结重采样算法. 其输入是有共同重叠区域的一对图像. 其输出是重采样的一对图像，使得这两幅图像中的对极线都是水平的(与 $x$ 轴平行)，并且两幅图像上的对应点尽可能相互接近. 两个匹配点之间的任何视差都在水平极线上. 该算法的一个最高层的概述如下.

(1) 在两幅图像之间确定从图像到图像匹配的种子集合 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$. 最少需要 7 个点，当然越多越好. 可以通过自动的方法来发现这样的匹配.

(2) 计算基本矩阵 F 并且在两幅图像上找到对极点 $\mathbf{e}$ 和 $\mathbf{e}'$.

(3) 选择一个把对极点 $\mathbf{e}'$ 映射到无穷远点 $(1,0,0)^\mathrm{T}$ 的射影变换 H′. 节 10.12.1 的方法常能得到好结果.

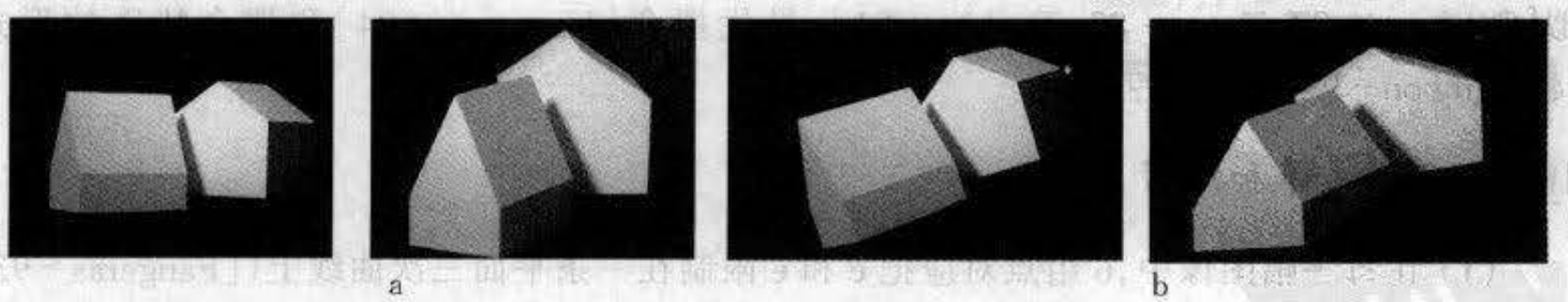
(4) 求最小化最小二乘距离

$$\sum_i d(\mathrm{H}\mathbf{x}_i, \mathrm{H}'\mathbf{x}_i') \tag{10.22}$$

的匹配射影变换 H，采用的方法是节 10.12.2 所介绍的线性方法.

(5) 根据射影变换 H 和射影变换 H′分别重采样第一幅和第二幅图像.

**例 10.5 模型房屋图像**



a b

图 10.11 **图像矫正**. (a)一座房屋的一对图像. (b)由(a)计算得到的重采样图像，它采用由 F 计算得到的一个射影变换. 注意在(b)上的对应点沿水平方向匹配.

图 10.11(a)显示某个木块房屋的一对图像. 在这两幅图像中自动地抽取边缘和顶点并且用手工匹配少数公共的顶点. 然后，按照这里介绍的方法重采样这两幅图像. 其结果在

图 10.11(b)中显示. 由于视点和物体的三维形状有很大差别,这两幅图像即使重采样后看起来也相当不同. 然而,这就是第一幅图像的任一点与第二幅图像上具有相同 $y$ 坐标的点匹配的真实情形. 因此,为了进一步求得两幅图像之间的匹配点仅需进行一维搜索. △

### 10.12.4 仿射矫正

本节所讨论的理论可同样适用于仿射重采样. 如果两个摄像机可以由仿射摄像机来近似,那么我们可以仅用仿射变换对之进行矫正. 具体实施时我们可以用仿射基本矩阵(见节 13.2($p$244))代替一般的基本矩阵. 上面的方法只要稍加改动就可用来计算一对匹配的仿射变换. 图 10.12 给出用仿射变换矫正后的一对图像.

a b

图 10.12 **用仿射变换作图像矫正**. (a)一对原始图像(b)用仿射变换矫正后的图像细节. 在 512×512 图像上,矫正后 $y$ 方向的平均视差在 3 个像素的量级上. (对正确矫正的图像来说,$y$ 方向的视差应该是零.)

## 10.13 结束语

### 10.13.1 文献

计算基本矩阵的基本思想由[LonguetHiggins-81]给出,它非常值得一读. 虽然它仅针对标定矩阵的情形,其原则也适用于未标定的情形. 未标定情形的一篇好的参考文献是[Zhang-98],它考虑了绝大多数最好的方法. 另外,该文献也考虑了极线包络的不确定性. 对未标定时的 8 点算法的详细研究由[Hartley-97]给出. Weng 等在[Weng-89]中把 Sampson 逼近用于基本矩阵的代价函数. 强迫被估计的 F 具有秩 2 的 SVD 方法由 Tsai 和 Huang[Tsai-84]建议.

有大量的关于二次曲线匹配的文献——最小化代数距离[Bookstein-79];几何距离近似[Sampson-82, Pratt-87, Taubin-91];最优拟合[Kanatani-94]和拟合特殊的形式[Fitzgibbon-99].

### 10.13.2 注释和练习

(1) 在每一幅图像中,6 组点对应把 $\mathbf{e}$ 和 $\mathbf{e}'$ 限制在一条平面三次曲线上([Faugeras-93] $p$298). 该三次曲线在每一幅图像中也通过这 6 点. 下面是这些结果的推导概要.

给定 6 组点对应,(10.3 - $p$191)中的 A 的零空间是 3 维. 从而其解是 $F=\alpha_1F_1+\alpha_2F_2+\alpha_3F_3$,其中 $F_i$ 表示对应于生成该零空间的诸矢量的矩阵. 对极点满足 $F\mathbf{e}=\mathbf{0}$,因此 $[(F_1\mathbf{e}),(F_2\mathbf{e}),(F_3\mathbf{e})](\alpha_1,\alpha_2,\alpha_3)^T=\mathbf{0}$. 因为该方程有一个非零解,

由此推出 $\det[(F_1\mathbf{e}),(F_2\mathbf{e}),(F_3\mathbf{e})]=0$，它是 $\mathbf{e}$ 的三次多项式.

(2) 证明 4 个共面点的图像对应和一个二次曲面的轮廓线可以确定基本矩阵到具有两重多义性(提示：参见算法 12.2($p$237)).

(3) 证明一条(平面)二次曲线的图像对应等价于关于 F 的两个约束. 详见[Kahl-98b].

(4) 假定一组双眼图像对是通过一个摄像机沿它的主轴向前平移而得到. 节 10.12 所介绍的图像矫正的几何能否适用于这种情况？

# 第 11 章　结构计算

本章介绍当 3D 点在两幅视图上的像和这些视图的摄像机矩阵给定时，如何计算该点的位置. 假定误差仅在图像坐标测量中发生，而投影矩阵 P，P′没有误差.

在上述条件下，由测量图像点的反向投影射线构成的简单三角形法将不再适用，因这两条射线一般不相交. 因此有必要为 3D 空间点**估计**求一个最佳解.

一个最佳解需要定义并最小化一个适当的代价函数. 这个问题在仿射和射影重构中特别重要，因为它们不存在物体空间意义上的度量信息. 我们希望寻找一种在空间射影变换下不变的三角形法.

在下面各节中，我们将给出 $\mathbf{X}$ 和它的协方差的估计. 推导关于该点的最优(MLE)估计算法，并且证明不需要数值最小化就能得到一个解.

注意，上述情形是**先知** F 而后求 $\mathbf{X}$. 另一种情形是从图像点对应 $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$ 同时估计 F 和 $\{\mathbf{X}_i\}$，但后一种情形不在本章考虑的范围之内. 这个问题可以利用本章所介绍的算法作为一个初始的估计并用节 10.4.1 中的黄金标准算法来解.

## 11.1　问题陈述

假定摄像机矩阵已给定，从而基本矩阵也就给定；或者假定基本矩阵已给定，从而可以构作一对相容的摄像机矩阵(如节 8.5($p$171)所述). 在任一种情况下都假定这些已知矩阵是准确的，或者与两幅图像上的一对匹配点相比具有极大的准确性.

因为**被测量**点 $\mathbf{x}$ 和 $\mathbf{x}'$ 有误差，所以从点反向投影的射线不共面，这意味着**不存在**一个点 $\mathbf{X}$ 准确地满足 $\mathbf{x}=\mathrm{P}\mathbf{X}, \mathbf{x}'=\mathrm{P}'\mathbf{X}$；并且图像点**不**满足对极几何约束 $\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0$. 这两种说法是等价的，因为对应于一对匹配点 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 的射线在空间中相交的充分必要条件是这两点满足对极几何约束. 见图 11.1.

三角形法的一个理想特性应是其在重构的某种变换下不变——若摄像机矩阵相差一个仿射(或射影)变换，则显然希望采用仿射(或射影)不变的三角形法来计算 3D 空间点. 因此，定义 $\tau$ 是由对应点 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 以及一对摄像机矩阵 P 和 P′来计算 3D 空间点 $\mathbf{X}$ 的一种三角形法. 我们记

$$\mathbf{X} = \tau(\mathbf{x}, \mathbf{x}', \mathrm{P}, \mathrm{P}')$$

该三角形称为在变换 H 下是不变的，如果

$$\tau(\mathbf{x}, \mathbf{x}', \mathrm{P}, \mathrm{P}') = \mathrm{H}^{-1}\tau(\mathbf{x}, \mathbf{x}', \mathrm{P}\mathrm{H}^{-1}, \mathrm{P}'\mathrm{H}^{-1}).$$

这意味着利用变换了的摄像机作三角形法得出变换的点.

显然，特别对射影重构，在 3D 射影空间 $\mathbb{P}^3$ 中对误差进行最小化是不适宜的. 例如，求空间中两条射线的公垂线的中点对射影重构是不合适的，因为诸如距离和垂直的概念在射影几何的背景下无效. 事实上，在射影重构下，依赖具体采用的射影重构，这个方法将给出不同的结果——即这个方法不是射影不变的.

这里我们将给出一种射影不变的三角形法. 其关键的思想是估计一个 3D 点 $\hat{\mathbf{X}}$,使它精确地满足给定的摄像机几何,从而它的投影是

$$\hat{\mathbf{x}} = \mathrm{P}\hat{\mathbf{X}} \qquad \hat{\mathbf{x}}' = \mathrm{P}'\hat{\mathbf{X}}$$

其目的是由图像测量 $\mathbf{x}$ 和 $\mathbf{x}'$ 来估计 $\hat{\mathbf{X}}$. 如节 11.3 所述,在高斯噪声下,最大似然估计由最小化重投影误差——$\hat{\mathbf{X}}$的投影到图像测量点的(平方和)距离得到.

这样的三角形法是射影不变的,因为只是最小化了图像的距离,且 $\hat{\mathbf{X}}$的投影点 $\hat{\mathbf{x}}$和 $\hat{\mathbf{x}}'$不依赖于 $\hat{\mathbf{X}}$所定义的射影坐标系,即不同的射影重构将投影到同样的点.

下面,我们将给出简单的线性三角形法. 然后定义 MLE,且证明通过求一个六次多项式方程的根能得到一个最优解,因而避免采用非线性最小化代价函数.

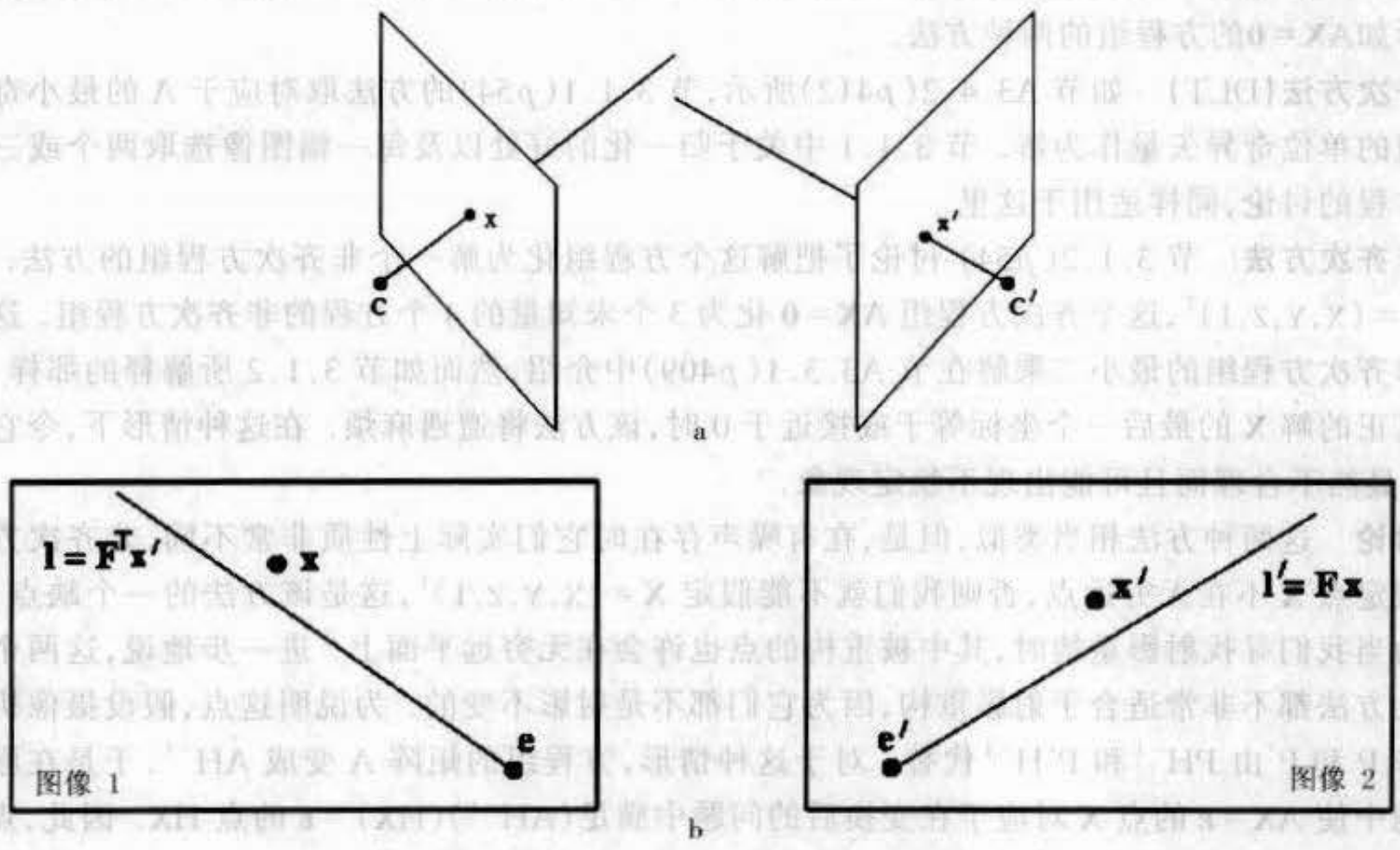



图 11.1 (a) 从非理想测量点 $\mathbf{x}$ 和 $\mathbf{x}'$反向投影的射线在 3 维空间中一般不共面. (b)$\mathbf{x}$ 和 $\mathbf{x}'$的对极几何. 测量点不满足对极几何约束. 对极线 $\mathbf{l}' = \mathrm{F}\mathbf{x}$ 是过 $\mathbf{x}$ 的射线的像,而对极线 $\mathbf{l} = \mathrm{F}^{\mathrm{T}}\mathbf{x}'$是过 $\mathbf{x}'$的射线的像. 因为射线不相交,$\mathbf{x}'$不在 $\mathbf{l}'$上,并且 $\mathbf{x}$ 也不在 $\mathbf{l}$ 上.

## 11.2 线性三角形法

本节将介绍简单的线性三角形法,通常,被估计的点不能准确地满足几何关系,因而不是一个最优估计.

线性三角形法是直接模仿节 3.1($p$53) 介绍的 DLT 方法. 在每一幅图像上分别有测量 $\mathbf{x} = \mathrm{P}\mathbf{X}$,$\mathbf{x}' = \mathrm{P}'\mathbf{X}$,且这些方程可以组合成 $\mathrm{A}\mathbf{X} = \mathbf{0}$ 的形式,它是关于 X 的线性方程.

首先通过叉乘消去齐次纯量因子,使每一个图像点给出三个方程,其中两个是线性独立的,例如,对第一幅图像有 $\mathbf{x} \times (\mathrm{P}\mathbf{X}) = \mathbf{0}$ 并把它展开得

$$x(\mathbf{P}^{3\mathrm{T}}\mathbf{X}) - (\mathbf{P}^{1\mathrm{T}}\mathbf{X}) = 0$$
$$y(\mathbf{P}^{3\mathrm{T}}\mathbf{X}) - (\mathbf{P}^{2\mathrm{T}}\mathbf{X}) = 0$$
$$x(\mathbf{P}^{2\mathrm{T}}\mathbf{X}) - y(\mathbf{P}^{1\mathrm{T}}\mathbf{X}) = 0$$

其中 $\mathbf{P}'^{\mathrm{T}}$ 是 P 的行,这些方程关于 $\mathbf{X}$ 的分量是线性的.

可以组成形如 $\mathrm{A}\mathbf{X}=\mathbf{0}$ 的方程,其中

$$\mathrm{A} = \begin{bmatrix} x\mathbf{P}^{3\mathrm{T}} - \mathbf{P}^{1\mathrm{T}} \\ y\mathbf{P}^{3\mathrm{T}} - \mathbf{P}^{2\mathrm{T}} \\ x'\mathbf{P}'^{3\mathrm{T}} - \mathbf{P}'^{1\mathrm{T}} \\ y'\mathbf{P}'^{3\mathrm{T}} - \mathbf{P}'^{2\mathrm{T}} \end{bmatrix}.$$

这里从每一像点取两个方程,总共给出 4 个齐次未知量的 4 个方程. 这是一个冗余方程组,因为解在相差一个尺度因子的意义下被确定. 这里将再次考虑节 3.1($p$53) 讨论过的求解形如$\mathrm{A}\mathbf{X}=\mathbf{0}$的方程组的两种方法.

**齐次方法**(DLT) 如节 A3.4.2($p$412)所示,节 3.1.1($p$54)的方法取对应于 A 的最小奇异值的单位奇异矢量作为解. 节 3.1.1 中关于归一化的好处以及每一幅图像选取两个或三个方程的讨论,同样适用于这里.

**非齐次方法** 节 3.1.2($p$54) 讨论了把解这个方程组化为解一个非齐次方程组的方法. 令 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}$,这个齐次方程组 $\mathrm{A}\mathbf{X}=\mathbf{0}$ 化为 3 个未知量的 4 个方程的非齐次方程组. 这个非齐次方程组的最小二乘解在节 A3.3.1($p$409)中介绍,然而如节 3.1.2 所解释的那样,当真正的解 $\mathbf{X}$ 的最后一个坐标等于或接近于 0 时,该方法将遭遇麻烦. 在这种情形下,令它为 1 显然不合理而且可能出现不稳定现象.

**讨论** 这两种方法相当类似,但是,在有噪声存在时它们实际上性质非常不同,非齐次方法假定点 $\mathbf{X}$ 不在无穷远点,否则我们就不能假定 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}$,这是该方法的一个缺点,因为当我们寻找射影重构时,其中被重构的点也许会在无穷远平面上. 进一步地说,这两个线性方法都不非常适合于射影重构,因为它们都不是射影不变的. 为说明这点,假设摄像机矩阵 P 和 P′由 $\mathrm{PH}^{-1}$和 $\mathrm{P}'\mathrm{H}^{-1}$代替. 对于这种情形,方程组的矩阵 A 变成 $\mathrm{AH}^{-1}$. 于是在原问题中使 $\mathrm{A}\mathbf{X}=\boldsymbol{\varepsilon}$ 的点 $\mathbf{X}$ 对应于在变换后的问题中满足$(\mathrm{AH}^{-1})(\mathrm{H}\mathbf{X})=\boldsymbol{\varepsilon}$ 的点 $\mathrm{H}\mathbf{X}$. 因此,点 $\mathbf{X}$ 和 $\mathrm{H}\mathbf{X}$ 之间存在给出相同误差的一一对应. 然而,在应用了射影变换 H 后,齐次方法的条件 $\|\mathbf{X}\|=1$和非齐次方法的条件 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}$ 都不是不变的. 因此原问题的解 $\mathbf{X}$ 一般不对应于变换后的解 $\mathrm{H}\mathbf{X}$.

但是对仿射变换而言情形则不同. 在仿射变换下条件下 $\|\mathbf{X}\|=1$ 虽然不一定保持不变,但条件 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}$ 一定保持不变,因为对仿射变换 H 有 $\mathrm{H}(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}=(\mathrm{X}',\mathrm{Y}',\mathrm{Z}',1)^{\mathrm{T}}$. 这意味着使 $\mathrm{A}(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}=\boldsymbol{\varepsilon}$ 的向量 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},1)^{\mathrm{T}}$ 与使$(\mathrm{AH}^{-1})(\mathrm{X}',\mathrm{Y}',\mathrm{Z}',1)^{\mathrm{T}}=\boldsymbol{\varepsilon}$ 的向量 $\mathrm{H}\mathbf{X}=(\mathrm{X}',\mathrm{Y}',\mathrm{Z}',1)^{\mathrm{T}}$ 之间存在着一一对应的关系. 对应点的误差相同. 这样一来,使误差 $\|\boldsymbol{\varepsilon}\|$ 最小化的点也是对应的,因此,非齐次方法是仿射不变的,而齐次方法却不是这样.

在本章余下的部分中,我们将介绍对于摄像机的射影坐标系不变并且最小化几何图像误差的一种三角形法. 它是推荐的三角形法,然而,上面介绍的齐次线性方法也常常提供可接受的结果. 而且,它的优点是容易把三角形法推广到视图多于两幅的情形.

## 11.3 几何误差代价函数

观察由一般不满足对极几何约束的一组噪声点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 形成的典型情形. 实际上,对应图像点的正确值应该是在测量点 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 附近的点 $\bar{\mathbf{x}}\leftrightarrow\bar{\mathbf{x}}'$,且准确满足对极几何约束 $\bar{\mathbf{x}}'^{\mathsf{T}}\mathrm{F}\bar{\mathbf{x}}=0$.

我们寻求点 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$,它们最小化函数

$$\mathcal{C}(\mathbf{x},\mathbf{x}') = d(\mathbf{x},\hat{\mathbf{x}})^2 + d(\mathbf{x}',\hat{\mathbf{x}}')^2 \text{ 满足 } \hat{\mathbf{x}}'^{\mathsf{T}}\mathrm{F}\hat{\mathbf{x}} = 0 \tag{11.1}$$

其中 $d(*,*)$ 是点之间的欧氏距离. 该问题等价于最小化点 $\hat{\mathbf{X}}$ 的重投影误差,$\hat{\mathbf{X}}$ 由与 F 相容的投影矩阵映射到 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$,如图 11.2 所示.

正如节 3.3($p63$)所解释的,在高斯误差分布的假设下,$\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$ 是关于真实的图像对应点的最大似然估计(MLE). 一旦求得 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$,点 $\hat{\mathbf{X}}$ 便可以由任何三角形法求得,因为相对应的射线在空间相交.

当然,代价函数也可以利用数值最小化方法,比如 Levenberg - Marquardt 方法来最小化(见节 A4.2($p417$)). 最小值的一个逼近也可以用几何代价函数的一阶导数,即 Sampson 误差来得到,这将在下一节中介绍. 此外,节 11.5 将证明可以通过解一个 6 次多项式的非迭代方法来求该最小值.

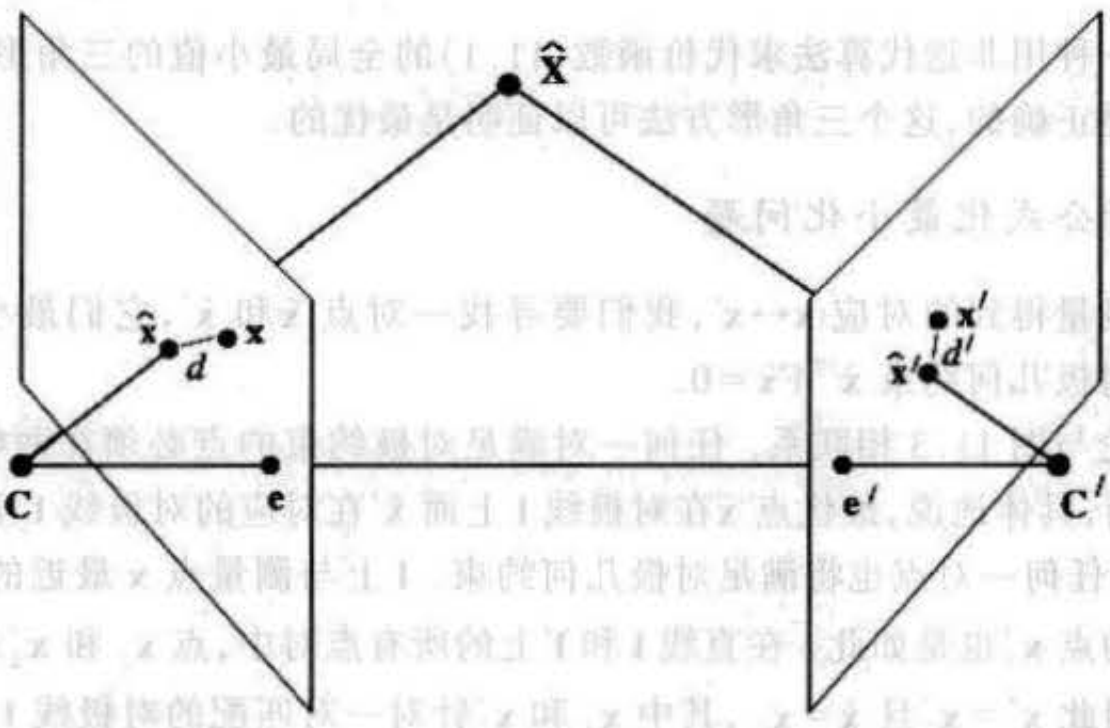



图 11.2 **最小化几何误差**. 估计的 3 维空间点 $\hat{\mathbf{X}}$ 投影到两幅图像的 $\hat{\mathbf{x}},\hat{\mathbf{x}}'$. 与测量点 $\mathbf{x}$ 和 $\mathbf{x}'$ 不同,图像对应点 $\hat{\mathbf{x}},\hat{\mathbf{x}}'$ 满足对极几何约束. 点 $\hat{\mathbf{X}}$ 的选择标准是使重投影误差 $d^2+d'^2$ 最小化.

## 11.4 Sampson 近似(一阶几何矫正)

在推导多项式的精确解之前,我们先推导 Sampson 近似,该方法在测量误差相对于测量值很小时是有效的. 对于基本矩阵情形,几何代价函数的 Sampson 近似已在节 10.4.3 中讨论. 这里我们所关心的是对测量点进行矫正的计算.

测量点 $\mathbf{X}=(x,y,x',y')^{\mathsf{T}}$(注,在本节中,$\mathbf{X}$ 不表示齐次 3D 空间点)的 Sampson 矫正 $\boldsymbol{\delta}_{\mathbf{x}}$ 在节 3.2.6($p61$)中被证明是(3.11 - $p61$)

$$\boldsymbol{\delta}_{\mathbf{X}} = -\mathrm{J}^{\mathrm{T}}(\mathrm{J}\mathrm{J}^{\mathrm{T}})^{-1}\boldsymbol{\varepsilon},$$

而矫正后的点是

$$\hat{\mathbf{X}} = \mathbf{X} + \boldsymbol{\delta}_{\mathbf{X}} = \mathbf{X} - \mathrm{J}^{\mathrm{T}}(\mathrm{J}\mathrm{J}^{\mathrm{T}})^{-1}\boldsymbol{\varepsilon}.$$

节 10.4.3($p$269)指出:对于由 $\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}=0$ 所定义的簇,该误差是 $\boldsymbol{\varepsilon}=\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}$,其Jacobian是

$$J = \partial\boldsymbol{\varepsilon}/\partial x = [(\mathrm{F}^{\mathrm{T}}\mathbf{x}')_1, (\mathrm{F}^{\mathrm{T}}\mathbf{x}')_2, (\mathrm{F}\mathbf{x})_1, (\mathrm{F}\mathbf{x})_2]$$

其中$(\mathrm{F}^{\mathrm{T}}\mathbf{x}')_1 = f_{11}x' + f_{21}y' + f_{31}$等,于是矫正点的一阶近似就是

$$\begin{bmatrix}\hat{x}\\ \hat{y}\\ \hat{x}'\\ \hat{y}'\end{bmatrix} = \begin{bmatrix}x\\ y\\ x'\\ y'\end{bmatrix} - \frac{\mathbf{x}'^{\mathrm{T}}\mathrm{F}\mathbf{x}}{(\mathrm{F}\mathbf{x})_1^2 + (\mathrm{F}\mathbf{x})_2^2 + (\mathrm{F}^{\mathrm{T}}\mathbf{x}')_1^2 + (\mathrm{F}^{\mathrm{T}}\mathbf{x}')_2^2}\begin{bmatrix}(\mathrm{F}^{\mathrm{T}}\mathbf{x}')_1\\ (\mathrm{F}^{\mathrm{T}}\mathbf{x}')_2\\ (\mathrm{F}\mathbf{x})_1\\ (\mathrm{F}\mathbf{x})_2\end{bmatrix}.$$

如果在每一幅图像上的矫正量很小(小于一个象素),则该近似是准确的,且计算量小,但是注意:矫正的点不会准确地满足对极几何关系 $\hat{\mathbf{x}}'^{\mathrm{T}}\mathrm{F}\hat{\mathbf{x}}=0$,用下一节的方法计算的点 $\hat{\mathbf{x}}$,$\hat{\mathbf{x}}'$将准确地满足对极几何约束,但是需要更大的计算量.

## 11.5 最优解

本节介绍一种用非迭代算法求代价函数(11.1)的全局最小值的三角形法,如果高斯噪声模型的假定是正确的,这个三角形方法可以证明是最优的.

### 11.5.1 重新公式化最小化问题

给定一组测量得到的对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$,我们要寻找一对点 $\hat{\mathbf{x}}$和 $\hat{\mathbf{x}}'$,它们最小化距离平方和(11.1)并满足对极几何约束 $\hat{\mathbf{x}}'^{\mathrm{T}}\mathrm{F}\hat{\mathbf{x}}=0$.

下面的讨论与图 11.3 相联系. 任何一对满足对极约束的点必须在两幅图像的对应的对极线上. 因而,具体地说,最优点 $\hat{\mathbf{x}}$在对极线 $\mathbf{l}$ 上而 $\hat{\mathbf{x}}'$在对应的对极线 $\mathbf{l}'$上. 另一方面,在直线 $\mathbf{l}$ 和 $\mathbf{l}'$上的任何一对点也将满足对极几何约束. $\mathbf{l}$ 上与测量点 $\mathbf{x}$ 最近的点 $\mathbf{x}_\perp$ 以及相应地定义在 $\mathbf{l}'$上的点 $\mathbf{x}'_\perp$也是如此. 在直线 $\mathbf{l}$ 和 $\mathbf{l}'$上的所有点对中,点 $\mathbf{x}_\perp$ 和 $\mathbf{x}'_\perp$最小化(11.1)的距离平方和. 因此 $\hat{\mathbf{x}}'=\mathbf{x}'_\perp$且 $\hat{\mathbf{x}}=\mathbf{x}_\perp$,其中 $\mathbf{x}_\perp$ 和 $\mathbf{x}'_\perp$针对一对匹配的对极线 $\mathbf{l}$ 和 $\mathbf{l}'$而定义. 因此我们可以记 $d(\mathbf{x},\hat{\mathbf{x}})=d(\mathbf{x},\mathbf{l})$,其中 $d(\mathbf{x},\mathbf{l})$表示从点 $\mathbf{x}$ 到直线 $\mathbf{l}$ 的垂直距离. 类似的表达式对$d(\mathbf{x}',\hat{\mathbf{x}}')$成立.

**图 11.3** 一个估计得到的 3D 点 $\hat{\mathbf{X}}$的投影 $\hat{\mathbf{x}}$和 $\hat{\mathbf{x}}'$在一组对应的对极线上. 最佳 $\hat{\mathbf{x}}$和 $\hat{\mathbf{x}}'$在测量点 $\mathbf{x}$ 和 $\mathbf{x}'$到对极线的垂足上. 把对应的对极线参数化成单参数簇,$\hat{\mathbf{X}}$的最佳估计转化成在对应的对极线上的单参数搜索,使得垂直距离的平方和 $d^2+d'^2$ 最小化.

根据上一段的分析，我们可以用如下的不同方式公式表述最小化问题. 我们要最小化

$$d(\mathbf{x},\mathbf{l})^2 + d(\mathbf{x}',\mathbf{l}')^2 \tag{11.2}$$

其中 $\mathbf{l}$ 和 $\mathbf{l}'$ 取遍所有对应的对极线. 因此，点 $\hat{\mathbf{x}}$ 是在直线 $\mathbf{l}$ 上与点 $\mathbf{x}$ 最靠近的点，$\hat{\mathbf{x}}'$ 也类似定义.

最小化(11.2)的策略如下：

(1) 用参数 $t$ 参数化第一幅图像的对极线束，并记第一幅图像上的对极线为 $\mathbf{l}(t)$.

(2) 利用基本矩阵 F，计算第二幅图像上的对应的对极线 $\mathbf{l}'(t)$.

(3) 把距离函数 $d(\mathbf{x},\mathbf{l}(t))^2 + d(\mathbf{x}',\mathbf{l}'(t))^2$ 显式地表示为 $t$ 的函数.

(4) 求最小化这个函数的 $t$ 值.

在这种方式下，问题化为求单变量 $t$ 的函数的最小值，即

$$\min_{\hat{\mathbf{x}}} \mathcal{C} = d(\mathbf{x},\hat{\mathbf{x}})^2 + d(\mathbf{x}',\hat{\mathbf{x}}')^2 = \min_t \mathcal{C} = d(\mathbf{x},\mathbf{l}(t))^2 + d(\mathbf{x}',\mathbf{l}'(t))^2.$$

因此，通过对极线束的一个适当的参数化，该距离函数将转化成 $t$ 的一个有理多项式函数. 利用初等微积分的方法，使该最小化问题转化为求一个 6 次多项式的实根的问题.

### 11.5.2 最小化的细节

如果一个空间点的两个图像点都为对极点，则该点在连接两个摄像机中心的直线上. 对于这种情形，该空间点的位置不能确定. 如果对应点中只有一个是对极点，则我们可以推知该空间点必然与另外一个摄像机的中心重合. 因此，我们假定两个图像点 $\mathbf{x}$ 和 $\mathbf{x}'$ 都不是对极点.

为简化分析，我们对每一幅图像都作刚体变换使得两个点 $\mathbf{x}$ 和 $\mathbf{x}'$ 都是齐次坐标系的原点 $(0,0,1)^\mathsf{T}$. 更进一步使对极点分别置于 $x$ 轴上的点 $(1,0,f)^\mathsf{T}$ 和 $(1,0,f')^\mathsf{T}$. $f$ 的值为 0 意味着对极点在无穷远. 这两个刚体变换对于(11.1)中的平方和距离函数没有影响，从而不改变最小化问题.

因此，我们假设在齐次坐标下，$\mathbf{x} = \mathbf{x}' = (0,0,1)^\mathsf{T}$，而两个对极点为点 $(1,0,f)^\mathsf{T}$ 和 $(1,0,f')^\mathsf{T}$. 在此情形，由于 $\mathrm{F}(1,0,f)^\mathsf{T} = (1,0,f')^\mathsf{T}\mathrm{F} = \mathbf{0}$，该基本矩阵应有下列特殊形式

$$\mathrm{F} = \begin{bmatrix} ff'd & -f'c & -f'd \\ -fb & a & b \\ -fd & c & d \end{bmatrix}. \tag{11.3}$$

考虑第一幅图像中通过点 $(0,t,1)^\mathsf{T}$（仍然在齐次坐标下）和对极点 $(1,0,f)^\mathsf{T}$ 的对极线. 我们标记这条对极线为 $\mathbf{l}(t)$. 这条对极线的矢量表示由叉积 $(0,t,1)^\mathsf{T} \times (1,0,f)^\mathsf{T} = (tf,1,-t)^\mathsf{T}$ 给出，所以由原点到该直线的距离的平方是

$$d(\mathbf{x},\mathbf{l}(t))^2 = \frac{t^2}{1+(tf)^2}.$$

在另一幅图像上，利用基本矩阵寻求对应的对极线得

$$\mathbf{l}'(t) = \mathrm{F}(0,t,1)^\mathsf{T} = (-f'(ct+d), at+b, ct+d)^\mathsf{T}. \tag{11.4}$$

这是直线 $\mathbf{l}'(t)$ 的齐次矢量表示. 从原点到该直线的距离的平方等于

$$d(\mathbf{x}',\mathbf{l}'(t))^2 = \frac{(ct+d)^2}{(at+b)^2 + f'^2(ct+d)^2}.$$

因此总的距离的平方是

$$s(t) = \frac{t^2}{1+f^2t^2} + \frac{(ct+d)^2}{(at+b)^2 + f'^2(ct+d)^2}. \tag{11.5}$$

我们的任务就是求这个函数的最小值.

我们可以利用如下的初等微积分的方法来求这个最小值,先求它的导数

$$s'(t)=\frac{2t}{(1+f^2t^2)^2}-\frac{2(ad-bc)(at+b)(ct+d)}{((at+b)^2+f'^2(ct+d)^2)^2}. \tag{11.6}$$

$s(t)$的最大值和最小值将在 $s'(t)=0$ 处出现. 把 $s'(t)$中的两项在一个公分母上合并且使分子等于0,给出条件

$$\begin{aligned} g(t) &= t((at+b)^2+f'^2(ct+d)^2)^2 \\ &\quad -(ad-bc)(1+f^2t^2)^2(at+b)(ct+d) \\ &= 0. \end{aligned} \tag{11.7}$$

$s(t)$的最大值和最小值将在这个多项式的根处出现. 这是一个6次多项式,它最多有六个实根,对应于函数 $s(t)$的3个极小值和3个极大值. 函数 $s(t)$的绝对最小值可以通过求 $g(t)$的根以及由(11.5)给定的函数 $s(t)$在每一个实根处的值而找到. 更简单地,可以只检查 $s(t)$在 $g(t)$每一个根(复或实)的实部的值以省去确定一个根是实或是复的麻烦. 我们也应该检查当 $t\to\infty$,$s(t)$的渐进值来确定最小距离是否发生在 $t=\infty$的情形,它对应于在第一幅图像中的一条对极线 $fx=1$.

整个方法总结在算法11.1中.

**目标**

已知一组测量得到的点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$和一个基本矩阵 F,计算在对极几何约束 $\hat{\mathbf{x}}'^{\mathsf{T}}\mathrm{F}\hat{\mathbf{x}}=0$ 下最小化几何误差(11.1)的矫正对应 $\hat{\mathbf{x}}\leftrightarrow\hat{\mathbf{x}}'$.

**算法**

(1) 定义变换矩阵

$$\mathrm{T}=\begin{bmatrix}1 & & -x\\ & 1 & -y\\ & & 1\end{bmatrix} \text{ 和 } \mathrm{T}'=\begin{bmatrix}1 & & -x'\\ & 1 & -y'\\ & & 1\end{bmatrix}$$

这些变换把 $\mathbf{x}=(x,y,1)^{\mathsf{T}}$ 和 $\mathbf{x}'=(x',y',1)^{\mathsf{T}}$ 平移到坐标原点.

(2) 用 $\mathrm{T}'^{-\mathsf{T}}\mathrm{F}\mathrm{T}^{-1}$代替 F. 新的 F 对应于平移后的坐标系.

(3) 计算右、左极点 $\mathbf{e}=(e_1,e_2,e_3)^{\mathsf{T}}$ 和 $\mathbf{e}'=(e_1',e_2',e_3')^{\mathsf{T}}$,使得 $\mathbf{e}'^{\mathsf{T}}\mathrm{F}=\mathbf{0}$ 和 $\mathrm{F}\mathbf{e}=0$. 归一化 $\mathbf{e}$(乘一个标量)使得 $e_1^2+e_2^2=1$,同时对 $\mathbf{e}'$作同样的处理.

(4) 构造矩阵

$$\mathrm{R}=\begin{bmatrix}e_1 & e_2 & \\ -e_2 & e_1 & \\ & & 1\end{bmatrix} \text{ 和 } \mathrm{R}'=\begin{bmatrix}e_1' & e_2' & \\ -e_2' & e_1' & \\ & & 1\end{bmatrix}$$

并注意 R 和 R′是旋转矩阵,满足 $\mathrm{R}\mathbf{e}=(1,0,e_3)^{\mathsf{T}}$ 和 $\mathrm{R}'\mathbf{e}'=(1,0,e_3')^{\mathsf{T}}$.

(5) 用 $\mathrm{R}'\mathrm{F}\mathrm{R}^{\mathsf{T}}$ 代替 F. 新的 F 必然具有 (11.3) 的形式.

(6) 令 $f=e_3$,$f'=e_3'$,$a=F_{22}$,$b=F_{23}$,$c=F_{32}$,$d=F_{33}$.

(7) 根据(11.7)形成多项式 $g(t)$,它是 $t$ 的多项式. 解 $t$ 得到6个根.

(8) 用 $g(t)$的每一个根的实部计算代价函数(11.5)(或仅用 $g(t)$的实根来计算). 同时求当 $t=\infty$时(11.1)的渐近值,即 $1/f^2+c^2/(a^2+f'^2c^2)$. 选择使代价函数取最小值的 $t$ 值 $t_{\min}$.

(9) 在 $t_{\min}$处计算两条直线 $\mathbf{l}=(tf,1,-t)^{\mathsf{T}}$ 和由(11.4)给出的 $\mathbf{l}'$,并求出在这些直线上最接近于原点的点 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$. 对于一般的直线 $(\lambda,\mu,\upsilon)^{\mathsf{T}}$,在直线上最接近原点的点的公式是 $(-\lambda\upsilon,-\mu\upsilon,\lambda^2+\mu^2)^{\mathsf{T}}$.

(10) 用 $\mathrm{T}^{-1}\mathrm{R}^{\mathsf{T}}\hat{\mathbf{x}}$替换 $\hat{\mathbf{x}}$并用 $\mathrm{T}'^{-1}\mathrm{R}'^{\mathsf{T}}\hat{\mathbf{x}}'$替换 $\hat{\mathbf{x}}'$变回到原来的坐标.

(11) 3维空间的点 $\hat{\mathbf{X}}$可以由节11.2的齐次方法获得.

算法 11.1 最佳三角形法.

### 11.5.3 局部极小

(11.7)中的 $g(t)$ 的次数为 6 意味着 $s(t)$ 最多可能有 3 个极小值. 如下面例子所示,这样的情形是可能出现的. 令 $f=f'=1$ 则

$$\mathbf{F}=\begin{pmatrix} 4 & -3 & -4 \\ -3 & 2 & 3 \\ -4 & 3 & 4 \end{pmatrix}$$

给出一个函数

$$s(t)=\frac{t^2}{1+t^2}+\frac{(3t+4)^2}{(2t+3)^2+(3t+4)^2}$$

其图形在图 11.4a[①]中显示. 其中清楚可见三个极小值.

作为第二个例子,我们考虑 $f=f'=1$ 以及.

$$\mathbf{F}=\begin{pmatrix} 0 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

对此情形,函数 $s(t)$ 由下式给定

$$s(t)=\frac{t^2}{t^2+1}+\frac{t^2}{t^2+(2t-1)^2}$$

而且当 $t=0$ 值时,代价函数的两项都为零,其意义是说对应点 $\mathbf{x}$ 和 $\mathbf{x}'$ 完全地满足对极约束. 这可由计算来证实 $\mathbf{x}'^{\mathrm{T}}\mathbf{F}\mathbf{x}=0$. 所以这两个点准确地匹配. 该代价函数 $s(t)$ 的图形在图 11.4b中显示. 除了在 $t=0$ 时有绝对最小值外,在 $t=1$ 时也有一个局部极小值. 因此,即使在准确匹配的情形下也可能出现局部极小值. 这个例子说明试图最小化(11.1)或者(11.2)中的代价函数时,从任意初始值开始的迭代搜索算法都有陷入求到局部极小值的危险,即使准确匹配的情形也不例外.

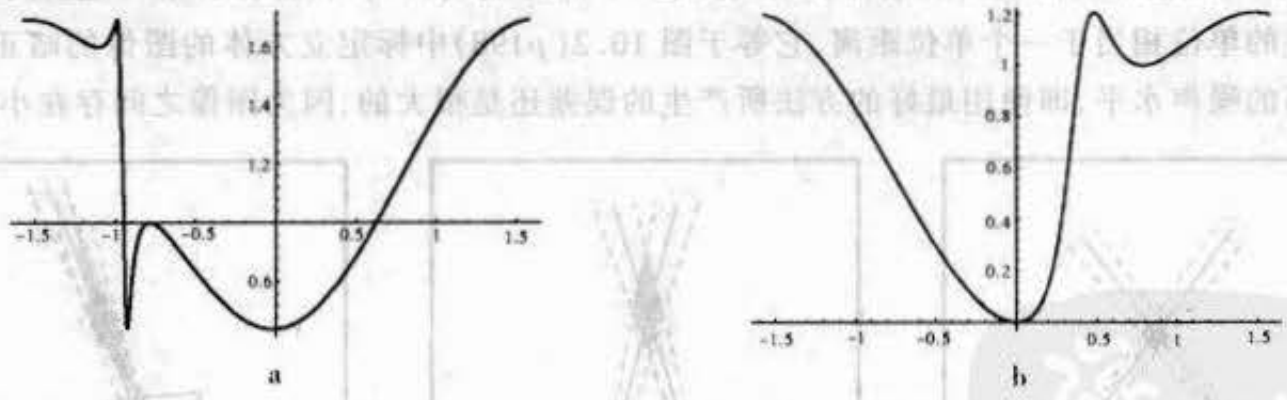



图 11.4 (a)具有 3 个极小值的代价函数的例子. (b)这是一个准确点匹配的代价函数,但它仍有两个极小值.

### 11.5.4 在实际图像上的计算

我们利用图 10.2($p$198)所示的标定立方体图像来进行试验,目的是确定三角形法怎样影响重构的准确性. 该立方体的一个欧氏模型,用来作为基本的真值依据,利用准确的图像测量来估计和修正. 测量像素的位置被矫正到准确地对应于该欧式模型,所需要的坐标矫

① 在此图和图 11.4b 中,我们作了替换 $t=\tan(\theta)$,并且在范围 $-\pi/2\leq\theta\leq\pi/2$ 之内按值 $\theta$ 绘制图形,以便显示 $t$ 的整个无限区域.

正平均在 0.02 像素.

因此,我们取得了一个模型和准确地对应于该模型的一组匹配点. 下一步,计算这些点的射影重构并计算使射影重构同该欧式模型相一致的一个射影变换 H. 把零均值的高斯噪声加到点的坐标上,在射影坐标系下采用两种三角形法并应用变换 H,最终在欧式坐标系下测量出每种方法的误差. 图 11.5 显示了两种三角形法所作实验的结果. 该图显示了在每次选取的一个噪声水平上分别作 10 次实验的所有点平均重构误差. 它清楚地表明最佳方法的重构结果较好.

在这组图像对中,两个对极点都远离对应图像. 对于两个对极点都接近图像的情形,合成图像上的结果进一步地肯定了多项式方法的优越性.

**被估计的 3D 点的协方差** 图 11.6 给出了重构点的协方差的一个说明. 根据经验,两条射线之间的角度可用来确定重构的准确性. 这样的经验比简单地考虑基线(它更常用于测量)要好.

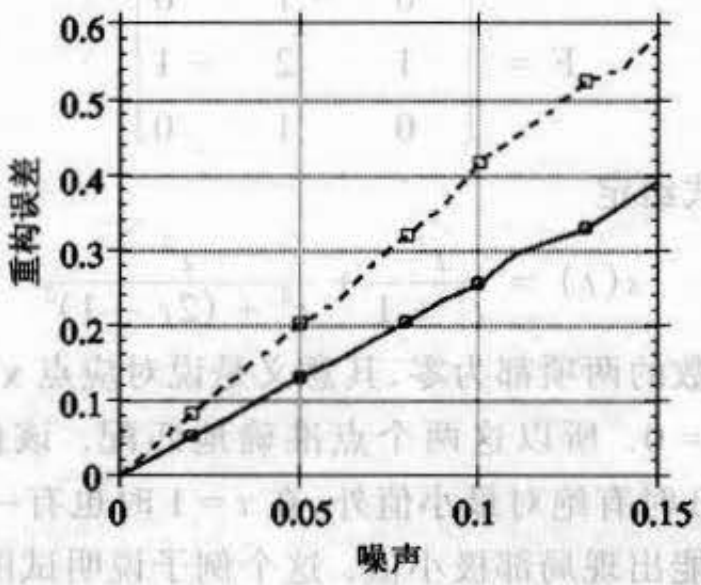



图 11.5 **三角形法重构误差的比较**. 此图显示由两种三角形法得到的重构误差:(1)在射影坐标下选择射线的公垂线的中点(上部的曲线)(2)最佳多项式方法(下部的曲线). 水平轴表示噪声,垂直轴表示重构误差. 重构误差的单位相当于一个单位距离,它等于图 10.2($p$198)中标定立方体的图像的暗正方形的一条边. 对于较高的噪声水平,即便用最好的方法所产生的误差还是很大的,因为图像之间存在小的运动.

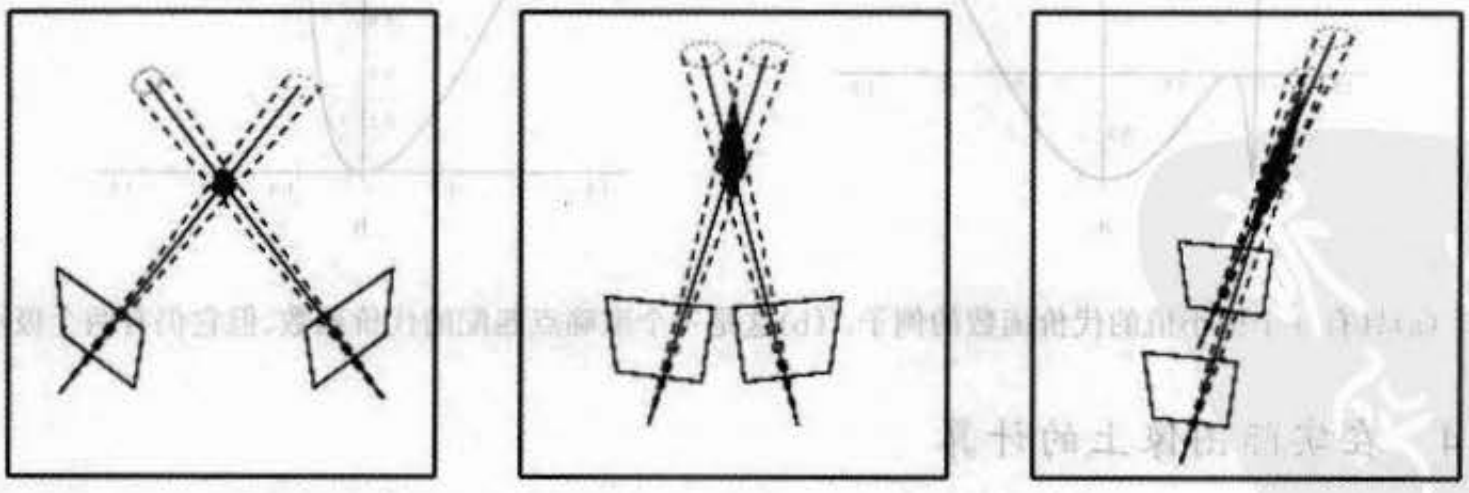


图 11.6 **重构的不确定性**. 对每一种情形所勾画出来的阴影区域说明不确定性区域的形状,它与射线之间的角度有关. 当射线越接近平行时点的定位精确度越差. 特别是向前运动会得到很差的重构,因为射线在许多视场中几乎是平行的.

## 11.6 直线重构

假定 3D 空间中一条直线投影为两幅视图上的直线 $\mathbf{l}$ 和 $\mathbf{l}'$. 该 3D 空间直线可由 $\mathbf{l}$ 和 $\mathbf{l}'$ 反向投影给出的两张 3D 空间平面的交来重构.

由直线定义的平面是 $\boldsymbol{\pi}=\mathrm{P}^{\mathsf{T}}\mathbf{l}$ 和 $\boldsymbol{\pi}'=\mathrm{P}'^{\mathsf{T}}\mathbf{l}'$. 在实际中,比较方便的做法是由图像直线定义的两平面来参数化 3D 空间中的直线,即如节 2.2.2($p$36)介绍的子空间生成表示那样,把该直线表示为 $2\times4$ 矩阵

$$\mathrm{L}=\begin{bmatrix}\mathbf{l}^{\mathsf{T}}\mathrm{P}\\ \mathbf{l}'^{\mathsf{T}}\mathrm{P}'\end{bmatrix}$$

因此若 $\mathrm{L}\mathbf{X}=\mathbf{0}$,则点 $\mathbf{X}$ 在该直线上.

在对应点的情形,前象(即投影到图像点的 3D 空间点)是超定的,因 3 自由度的 3D 空间点有 4 个测量值. 相反,在直线的情形,前像是准确地确定的,因 3D 空间中一条直线有 4 个自由度,而每幅视图的图像直线提供两个测量值. 注意,这里我们把直线看作无限长,并不用它们的端点.

**退化** 图 11.7 指出在对极平面上的 3D 空间直线 $\mathbf{l}$ 不能从它的两幅视图上的图像来确定. 这样的直线与摄像机的基线相交. 实际上,当测量有误差时,与基线几乎相交的直线在重构中定位也非常差.

直线的退化远比点的退化严重:在点的情形只有在对极线上的单参数点簇不能被恢复. 对于直线的情形,存在一个三参数直线簇:一个参数是在基线上的位置,而另外两簇是过基线上每点的直线的星.

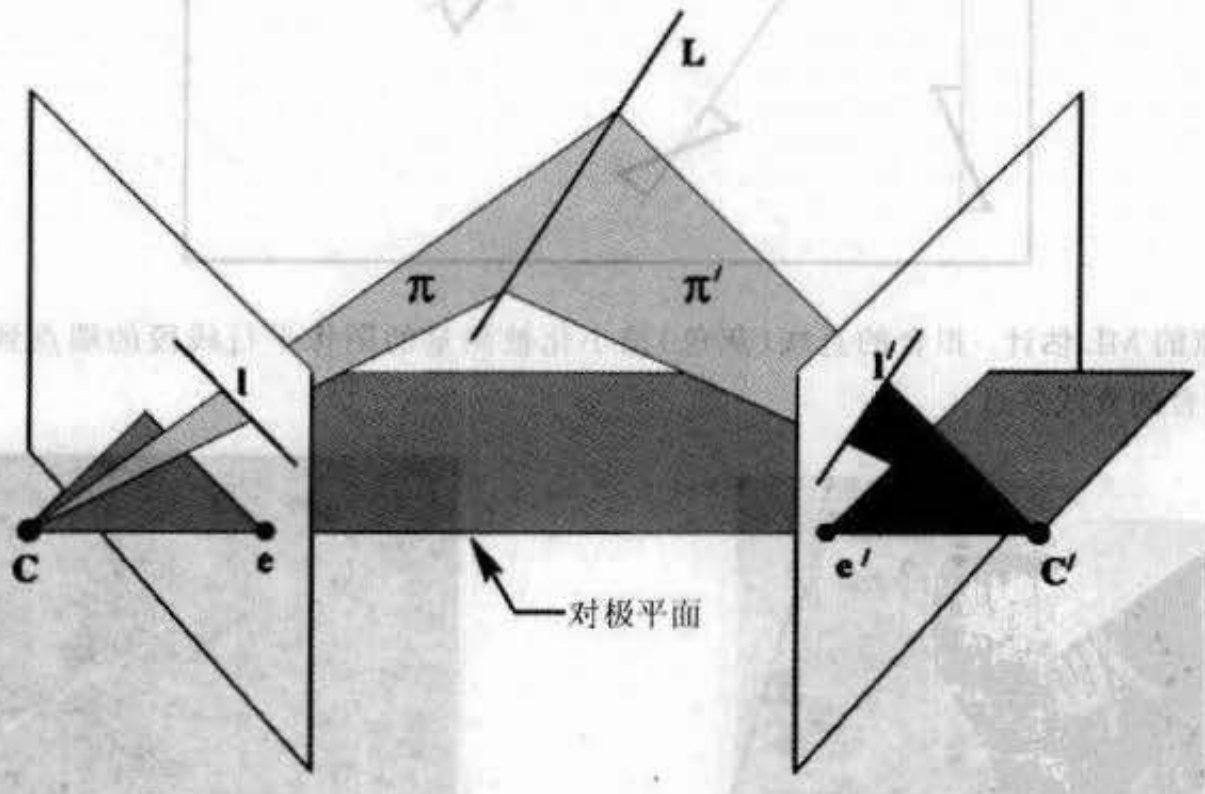



图 11.7 **直线重构**. 图像直线 $\mathbf{l},\mathbf{l}'$ 分别反向投影成平面 $\boldsymbol{\pi},\boldsymbol{\pi}'$. 此两平面的交确定 3 维空间的直线 $\mathbf{L}$. 如果 3 维空间的直线在对极平面上,那么它在 3 维空间的位置不能由它的图像确定. 此时对极点在图像直线上.

### 多于两张平面的相交

在下文的各章节中(特别在第 14 章)我们将考虑由 3 幅或更多视图的重构. 为了重构由几张平面相交而成的直线,可按如下步骤进行. 用 4 维矢量表示每一张平面 $\boldsymbol{\pi}_i$,并且对 $n$ 张平面形成一个行矢量为 $\boldsymbol{\pi}_i^{\mathsf{T}}$ 的 $n\times4$ 的矩阵 A. 令 $\mathrm{A}=\mathrm{UDV}^{\mathsf{T}}$ 为奇异值分解. V 中对应于

两个最大奇异值的两个列矢量生成 A 的秩为 2 的最好近似，并且可以用来定义这些平面的交线. 如果这些平面由图像直线反向投影来定义，则 3D 空间直线 **L** 的最大似然估计可以通过最小化几何图像距离得到，该距离是指 **L** 到每一幅图像上的投影和在那幅图像上的测量直线之间的距离. 这将在节 15.4.1($p$281)中讨论.

## 11.7 计算消影点

本节介绍由图像线段估计消影点. 图像线段来自场景中平行直线. 目标是最优地估计它们共同的图像交点——景物中平行直线方向的消影点.

这里给出的方法同样可以应用于其它多条直线相交的问题，例如，计算一个建筑物三个面的结合点处三条直线的交.

由于测量噪声，图像线段一般不相交于唯一的点. 因此，通常消影点是与所有被测量的直线最接近的点. 然而，这不是最优的结果. 消影点和线段的最大似然估计(MLE)由真正相交于一点的直线集合来确定，如图 11.8 所示，它最小化从测量线段的两端点到直线集合的垂直距离的平方和. 该最小化采用 Levenberg－Marquardt 算法(节 A4.2($p$417))来计算.

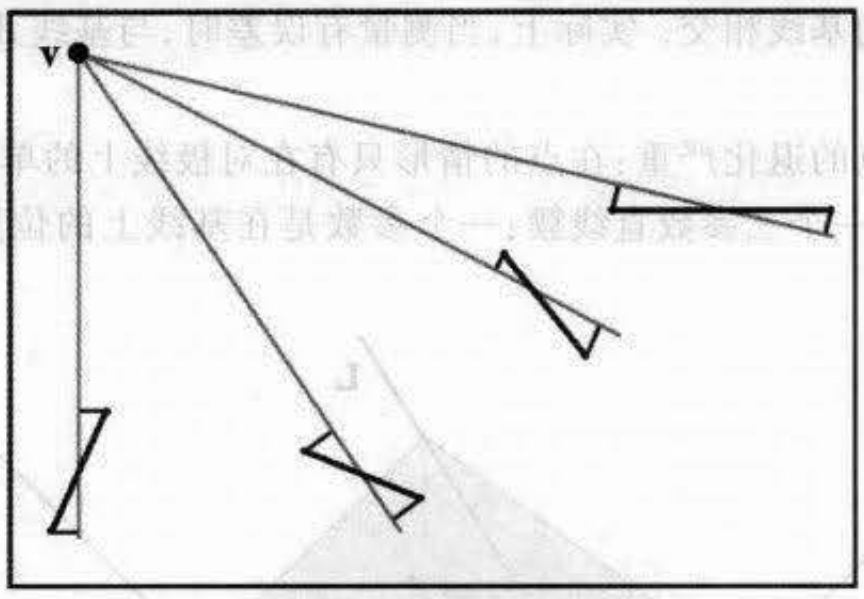
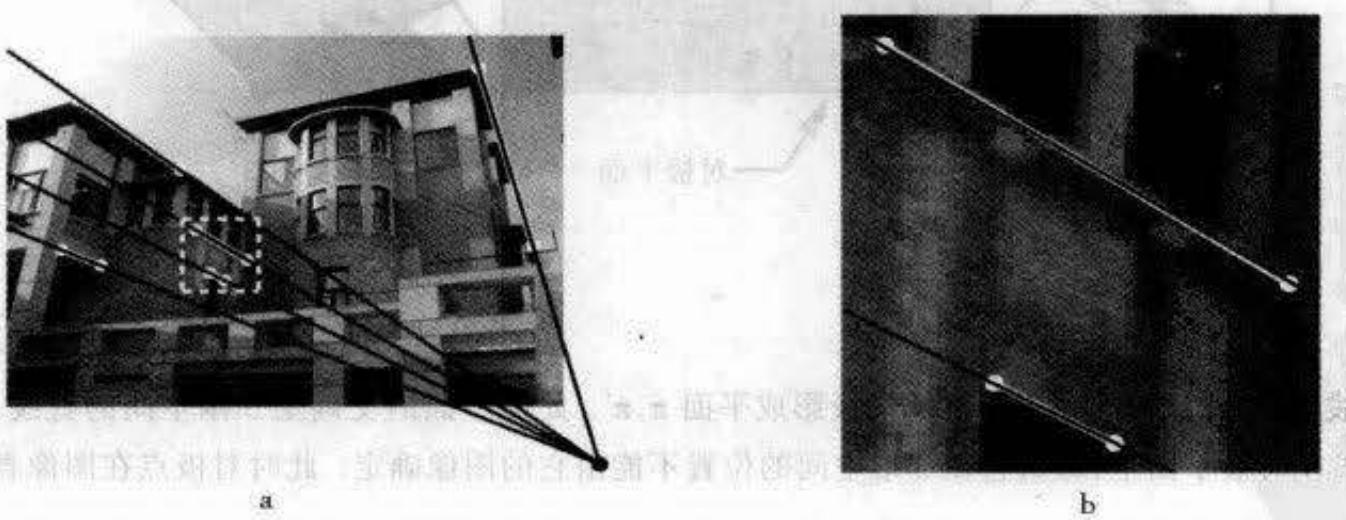



图 11.8 **消影点的 ML 估计**. 拟合的直线(灰色)最小化被测量的图像平行线段的端点到它的垂直距离(黑色). 所有拟合的直线共点.



a b

图 11.9 **由景物中平行直线的图像求消影点的 ML 估计**. (a)被测量的线段用白色表示，拟合的线段用黑色表示. 所有拟合的直线在消影点处共点. (b)(a)中虚线方框的放大. 注意测量直线和拟合直线之间有非常小的角度区别.

注意，如果这些直线不仅由它们的端点，而且由许多点的拟合来确定，则可以利用节15.7.2($p$287)介绍的方法将每一条直线变成相应的加权的端点对之后运用本算法.

图 11.9 给出用这种方式计算消影点的例子. 显然被测量的直线和拟合的直线之间的残差很小.

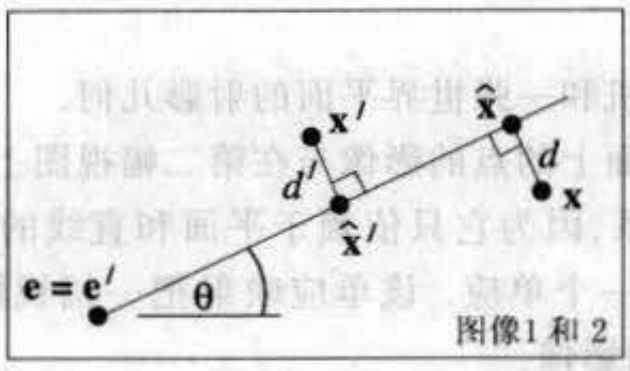



图 11.10 **纯平移的对极几何** 在这种情形中，对应的对极线是相同的(见节 10.7.1($p$202)). 最小化 $d^2+d'^2$ 的对极线(由 $\theta$ 来参数化)可以直接算出.

## 11.8 结束语

如何把三角形法的多项式方法推广到 3 幅或更多幅视图还不明朗. 然而，线性方法的推广是明显的. 更有趣的是 Sampson 方法也可以推广到 3 幅或更多幅视图，如[Torr－97]所介绍. 其缺点是在多视图时计算的代价(同时包括对编码的要求)明显增加.

### 11.8.1 文献

最佳三角形法是由 Hartley 和 Sturm[Hartley－95c]给出，而计算消影点的 ML 方法是由 Liebowitz 和 Zisserman[Liebowitz－98]给出.

### 11.8.2 注释和练习

(1) 针对于摄像机作纯平移运动的情形推导一个三角形法. 提示，参见图 11.10. 给出关于参数的一个封闭形式的解是可能的，这个方法被用在[Armstrong－94].

(2) 把多项式三角形法用于一对仿射摄像机(或更一般地，用于具有相同主平面的摄像机). 对于这种情形，基本矩阵有一个简单的形式(13.1－$p$245)，并且该方法被简化成一个线性方法.

(3) 证明线性三角形法在仿射摄像机时给出一个最优解.

(4) 证明 Sampson 方法(节 11.4)在图像的欧氏坐标变换(和 F 的相应改变)下是不变的.

(5) 针对于平面单应的情形，推导类似的三角形法的多项式解，即由已知一组测量对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 计算在约束 $\hat{\mathbf{x}}'=\mathrm{H}\hat{\mathbf{x}}$ 下最小化下面函数的点 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$：

$$\mathcal{C}(\mathbf{x},\mathbf{x}') = d(\mathbf{x},\hat{\mathbf{x}})^2 + d(\mathbf{x}',\hat{\mathbf{x}}')^2.$$

参看[Sturm－97b]，该文证明了该解是单变量的一个 8 次多项式.

# 第12章 景物平面和单应

本章介绍关于两个摄像机和一张世界平面的射影几何.

如图12.1所示,一张平面上的点的影像与在第二幅视图上的对应影像由一个(平面)单应相关联. 这是一个射影关系,因为它只依赖于平面和直线的相交. 这一事实被说成是该平面**诱导**了两幅视图之间的一个单应. 该单应映射把一幅视图中的点**转移**到另一幅视图中,好象它们是该平面上点的影像.

两幅视图之间存在两种关系:第一种,通过对极几何一幅视图中的一点可确定另一视图中的一条直线,它是过该点的射线的影像;第二种,通过上述单应一幅视图中的点可确定另一视图中的一个点,它是射线和一个平面交点的影像. 本章把两视图几何的这两种关系紧密地结合在一起.

这里还介绍另外两个重要的概念:相对于一张平面的视差和无穷单应.

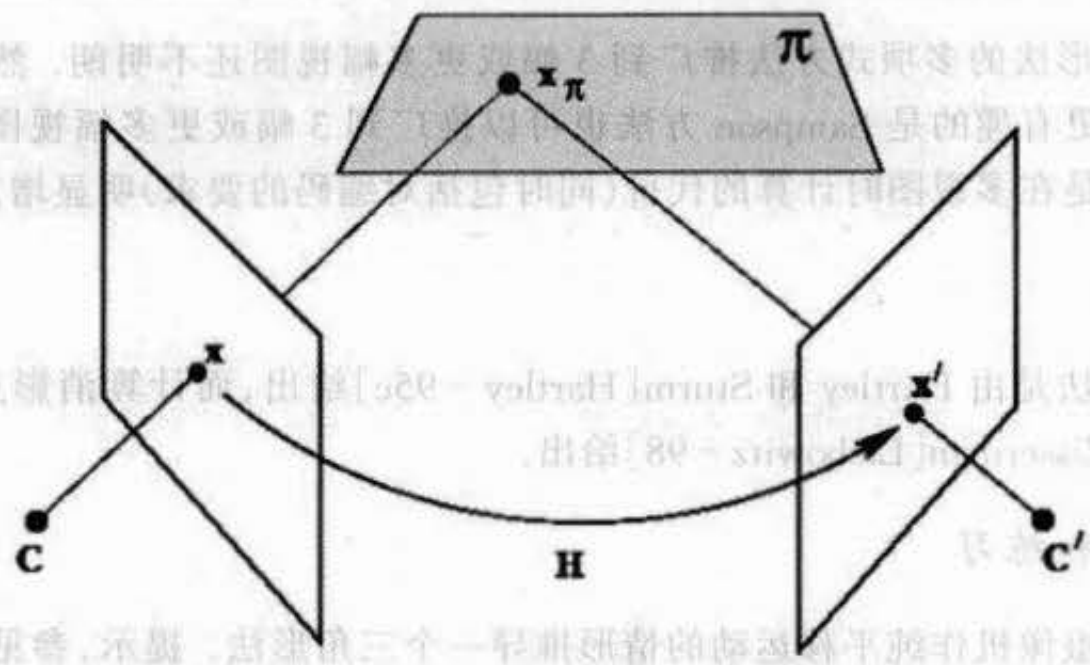



图12.1 **由平面诱导的单应**. 把点 $\mathbf{x}$ 的射线延长并与平面 $\pi$ 交于点 $\mathbf{x}_\pi$;该点被投影到另一图像上的点 $\mathbf{x}'$. 从 $\mathbf{x}$ 到 $\mathbf{x}'$ 的映射是由平面 $\pi$ 诱导的单应. 在世界平面 $\pi$ 和第一幅图像平面之间存在一个透视变换 $\mathbf{x}=\mathrm{H}_{1\pi}\mathbf{x}_\pi$;而在世界平面 $\pi$ 和第二幅图像平面之间存在透视变换 $\mathbf{x}'=\mathrm{H}_{2\pi}\mathbf{x}_\pi$. 这两个透视变换的复合是两个图像平面之间的一个单应变换 $\mathbf{x}'=\mathrm{H}_{2\pi}\mathrm{H}_{1\pi}^{-1}\mathbf{x}=\mathrm{H}\mathbf{x}$.

## 12.1 给定平面的单应和逆问题

我们首先证明对一般位置的平面,上述单应由该平面唯一地确定,反之也成立. 这里一般位置是指该平面不包含任何一个摄像机中心. 如果该平面包含了一个摄像机中心,则所诱导的单应是退化的.

假定在3D空间中一张平面 $\pi$ 由它在世界坐标系中的坐标指定. 我们首先推导诱导单应的一个显式表达式.

**结论12.1 给定两幅视图的投影矩阵**

$$P = [I \mid \mathbf{0}] \qquad P' = [A \mid \mathbf{a}]$$

**和由 $\boldsymbol{\pi}^{\mathsf{T}}\mathbf{X}=0$ 定义的一张平面，其中 $\boldsymbol{\pi}=(\mathbf{v}^{\mathsf{T}},1)^{\mathsf{T}}$，则由这个平面诱导的单应是 $\mathbf{x}'=H\mathbf{x}$，其中**

$$H = A - \mathbf{a}\mathbf{v}^{\mathsf{T}}. \tag{12.1}$$

由于平面不通过第一个摄像机的中心 $(0,0,0,1)^{\mathsf{T}}$，我们可以假定 $\pi_4=1$.

注意，3D 空间平面是一个三参数簇，因此由 3D 空间中平面诱导的两幅视图之间的一个三参数单应簇，这三参数是由矢量 $\mathbf{v}$ 的元素指定，这里 $\mathbf{v}$ **不是**齐次 3 维矢量.

**证明** 为计算 H，把第一幅视图上的点 $\mathbf{x}$ 反向投影且确定该射线和平面 $\boldsymbol{\pi}$ 的交点 $\mathbf{X}$，然后再把这个 3D 点 $\mathbf{X}$ 投影到第二幅视图上.

对第一幅视图有 $\mathbf{x}=P\mathbf{X}=[I|\mathbf{0}]\mathbf{X}$，因此该射线上的任何点 $\mathbf{X}=(\mathbf{x}^{\mathsf{T}},\rho)^{\mathsf{T}}$ 都投影到 $\mathbf{x}$，其中 $\rho$ 是确定该射线上点的参数. 因 3D 点 $\mathbf{X}$ 在平面 $\boldsymbol{\pi}$ 上，故它满足 $\boldsymbol{\pi}^{\mathsf{T}}\mathbf{X}=0$. 此式确定了 $\rho$，从而 $\mathbf{X}=(\mathbf{x}^{\mathsf{T}},-\mathbf{v}^{\mathsf{T}}\mathbf{x})^{\mathsf{T}}$. 3D 点 $\mathbf{X}$ 在第二幅图像上的投影是

$$\begin{aligned}\mathbf{x}' &= P'\mathbf{X} = [A \mid \mathbf{a}]\mathbf{X} \\ &= A\mathbf{x} - \mathbf{a}\mathbf{v}^{\mathsf{T}}\mathbf{x} = (A - \mathbf{a}\mathbf{v}^{\mathsf{T}})\mathbf{x}.\end{aligned}$$

□

**例 12.2 一个标定的双眼装置**

假定标定的双眼装置的摄像机矩阵分别为(取世界原点在第一个摄像机中心)：

$$P_E = K[I \mid \mathbf{0}] \qquad P'_E = K'[R \mid \mathbf{t}],$$

并且世界平面 $\boldsymbol{\pi}_E$ 的坐标是 $\boldsymbol{\pi}_E=(\mathbf{n}^{\mathsf{T}},d)^{\mathsf{T}}$，即该平面上的点满足 $\mathbf{n}^{\mathsf{T}}\widetilde{\mathbf{X}}+d=0$. 我们希望计算由这个平面诱导的单应的一个表达式.

根据结论 12.1 及 $\mathbf{v}=\mathbf{n}/d$，关于摄像机 $P=[I|\mathbf{0}]$，$P'=[R|\mathbf{t}]$ 的单应是

$$H = R - \mathbf{t}\mathbf{n}^{\mathsf{T}}/d.$$

对图像施加变换 K 和 K′，我们得到摄像机 $P_E=K[I|\mathbf{0}]$，$P'_E=K'[R|\mathbf{t}]$ 及其诱导单应是

$$H = K'(R - \mathbf{t}\mathbf{n}^{\mathsf{T}}/d)K^{-1}. \tag{12.2}$$

这是一个三参数单应簇，并由 $\mathbf{n}/d$ 参数化. 整个簇由该平面、摄像机内参数和两摄像机相对关系的外参数确定.

### 12.1.1 与对极几何兼容的单应

假定在一张景物平面上选定 4 个点 $\mathbf{X}_i$，则这些点在两幅视图之间的图像对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 所确定的单应 H 就是由该平面诱导的单应. 这些图像对应也满足对极几何约束，即 $\mathbf{x}'^{\mathsf{T}}_i F\mathbf{x}_i=0$，因为它们由场景点的图像得到. 对**任何** $\mathbf{x}$，对应 $\mathbf{x}\leftrightarrow\mathbf{x}'=H\mathbf{x}$ 都满足对极几何约束，同样因为 $\mathbf{x}$ 和 $\mathbf{x}'$ 是一个景物点的图像；此时景物点由景物平面和 $\mathbf{x}$ 反向投影的射线的交点给出，这样的单应 H 称为与 F 相容或**兼容**.

现在假定在第一幅视图上**任意**选择 4 个图像点，同时在第二幅视图任意选择 4 个图像点；这样也可算得一个单应 $\widetilde{H}$，它把一组点映射到另一组点(如果在任何一幅视图中没有 3 个点共线). 但是，点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'=\widetilde{H}\mathbf{x}$ 可能**不**满足对极几何约束. 而如果对应 $\mathbf{x}\leftrightarrow\mathbf{x}'=\widetilde{H}\mathbf{x}$ 不满足对极几何约束，则不存在诱导单应为 $\widetilde{H}$ 的任何景物平面.

对极几何决定了两幅视图之间的射影几何，并且可以用作确定单应由实际景物平面诱导的条件. 图 12.2 说明了对极几何和景物平面之间的若干关系，它们可以用来定义上述的条件. 例如，如果 H 由一个平面诱导，那么对应 $\mathbf{x}\leftrightarrow\mathbf{x}'=H\mathbf{x}$ 满足对极几何约束，从而由 $\mathbf{x}'^{\mathsf{T}}F\mathbf{x}=0$

得到

$$(\mathbf{Hx})^{\mathsf{T}}\mathbf{Fx} = \mathbf{x}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\mathbf{Fx} = 0,$$

并对所有点 $\mathbf{x}$ 成立. 因此:

- **一个单应 H 与一个基本矩阵 F 相容的充要条件是矩阵 $H^{\mathsf{T}}F$ 是反对称的:**

$$H^{\mathsf{T}}F + F^{\mathsf{T}}H = 0. \tag{12.3}$$

上面的论证已证明了该条件的必要性. Luong 和 Viéville[Luong-96]证明了该条件的充分性. 下面我们来计算自由度,(12.3)对 H 的 8 个自由度施加了 6 个齐次(5 个非齐次)约束. 因此还给 H 留了 8-5=3 个自由度:这三个自由度对应于 3D 空间平面的三参数簇.

相容约束(12.3)是关于 H 和 F 的一个隐式方程. 我们现在推导在给定 F 下,由平面诱导的单应 H 的显式表达式,它是更适合于计算的算法.

**结论 12.3 给定两幅视图之间的基本矩阵 F,由一个世界平面诱导的 3 参数单应簇是:**

$$H = A - \mathbf{e}'\mathbf{v}^{\mathsf{T}} \tag{12.4}$$

**其中 $[\mathbf{e}']_\times A = F$ 是基本矩阵的任意分解.**

**证明** 结论 12.1 已经证明在给定一对视图的摄像机矩阵 $P=[I|\mathbf{0}]$, $P'=[A|\mathbf{a}]$后,平面 $\boldsymbol{\pi}$ 诱导单应 $H=A-\mathbf{a}\mathbf{v}^{\mathsf{T}}$,其中 $\boldsymbol{\pi}=(\mathbf{v}^{\mathsf{T}},1)^{\mathsf{T}}$. 然而,根据结论 8.9($p$171),对于基本矩阵 $F=[\mathbf{e}']_\times A$,我们能选择两个摄像机为$[I|\mathbf{0}]$和$[A|\mathbf{e}']$. □

**注释** 上面的推导是基于一张平面上点的投影,保证了单应与极几何的相容性. 从代数上分析,单应(12.4)必与基本矩阵相容,因为它满足充分必要条件(12.3),即 $F^{\mathsf{T}}H$ 是反对称的. 这由下式得到:

$$F^{\mathsf{T}}H = A^{\mathsf{T}}[\mathbf{e}']_\times(A - \mathbf{e}'\mathbf{v}^{\mathsf{T}}) = A^{\mathsf{T}}[\mathbf{e}']_\times A$$

(其中利用了$[\mathbf{e}']_\times\mathbf{e}'=\mathbf{0}$),显然 $A^{\mathsf{T}}[\mathbf{e}']_\times A$ 是反对称的.

把(12.4)与在引理 8.11($p$172)或者(8.10-$p$173)所给出的基本矩阵的一般分解式比较,显然它们为同一个公式(除了符号可能不同之外),事实上,基本矩阵的分解式(在引理 8.11 中相差一个多义性比例因子 $k$)和由世界平面诱导的单应之间存在着一一对应,正如下面推论所述:

**推论 12.4 一个变换 H 是由某世界平面诱导的两幅图像之间单应的充要条件是这两幅图像的基本矩阵 F 有一个分解式:$F=[\mathbf{e}']_\times H$.**

这个分解式的选择就是对应于射影世界坐标系的选择. 事实上,在 $P=[I|\mathbf{0}]$和 $P'=[H|\mathbf{e}']$的重构中,H 是与坐标为$(0,0,0,1)^{\mathsf{T}}$的平面相对应的变换.

给定一对摄像机矩阵,求诱导已知单应的平面是一件简单的事,即:

**结论 12.5 给定在标准形式下的摄像机 $P=[I|\mathbf{0}]$, $P'=[A|\mathbf{a}]$,则诱导这两幅视图之间的给定单应 H 的平面 $\boldsymbol{\pi}$ 的坐标是 $\boldsymbol{\pi}=(\mathbf{v}^{\mathsf{T}},1)^{\mathsf{T}}$,其中 $\mathbf{v}$ 可以由线性求解方程组 $\lambda H = A - \mathbf{a}\mathbf{v}^{\mathsf{T}}$ 得到,该方程组关于 $\mathbf{v}$ 和 $\lambda$ 的元素是线性的.**

注意,只有 H 满足与 F 的相容性约束(12.3),这些方程才有一个精确解. 使用噪声数据由数值方法计算得到的单应一般不成立,并且该线性系统是超定的.

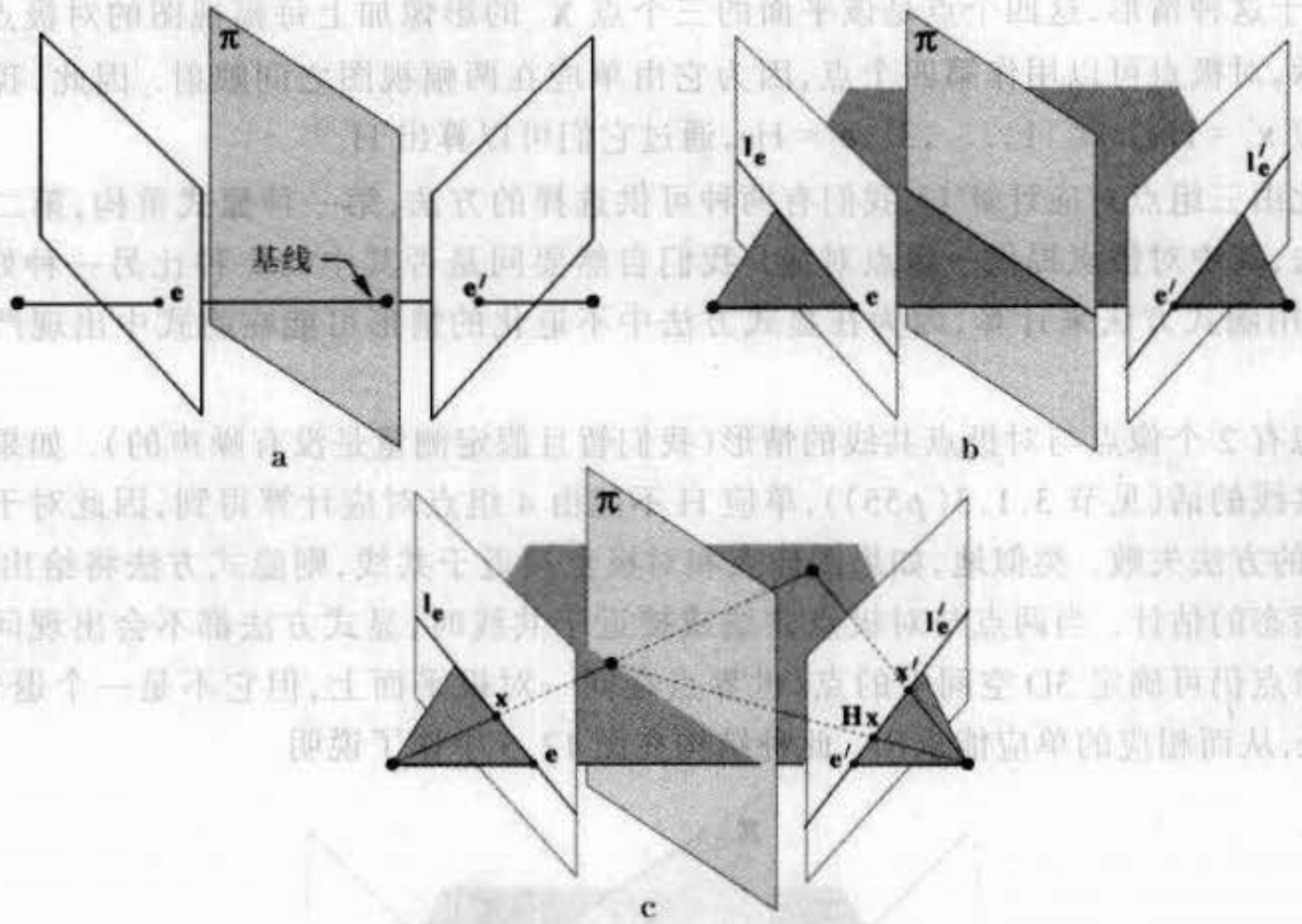



图 12.2 **相容性约束**. 由一个平面诱导的单应与对极几何相关联并满足若干约束. (a)对极点由该单应映射,即 $\mathbf{e}'=\mathbf{H}\mathbf{e}$,这是因为对极点是平面 $\pi$ 与基线的交点的图像. (b)对极线由该单应映射为 $\mathbf{H}^{\mathsf{T}}\mathbf{l}'_e=\mathbf{l}_e$. (c)任何点 $\mathbf{x}$ 由该单应映射到它对应的对极线 $\mathbf{l}'_e$ 上,因此 $\mathbf{l}'_e=\mathbf{F}\mathbf{x}=\mathbf{x}'\times(\mathbf{H}\mathbf{x})$.

## 12.2 给定 F 和图像对应求平面诱导的单应

3D 空间的一张平面可以由三个点确定,或由一条直线和一个点确定等等. 接下来这些 3D 元素可以由图像对应来确定. 在节 12.1 中单应由平面的坐标算出. 在下文中,该单应将直接由确定这个平面的对应图像元素来算出. 这是应用中自然采用的一种方法.

我们将考虑两种情形:(1)三个点;(2)一直线和一点. 每一种情形中,对应元素都足以唯一确定 3D 空间的一张平面. 在每一种情形中我们将会看到:

(1) 对应的图像元素必须满足对极几何的相容性约束.

(2) 存在 3D 元素和摄像机的**退化配置**,此时单应没有定义. 这样的退化来源于 3D 元素和极几何的共线与共面性. 同时,求解方法也可能产生退化,但这种退化是可以避免的.

我们打算对三点的情形介绍得详细些.

### 12.2.1 三点情形

假设三个(非共线)点 $\mathbf{X}_i$ 在两幅视图中的影像和基本矩阵 F 都已知. 由这三点所在平面诱导的单应 H 原则上可通过两种途径来计算:

第一种,先在射影重构下恢复点 $\mathbf{X}_i$ 的位置(第 11 章),然后确定通过这些点的平面(2.3 – $p$35),再根据结论 12.1 由这个平面算出该单应. 第二种,可以由四组点对应计算该

单应，对于这种情形，这四个点是该平面的三个点 $\mathbf{X}_i$ 的影像加上每幅视图的对极点. 如图12.2所示，对极点可以用作第四个点，因为它由单应在两幅视图之间映射. 因此，我们有四组点对应 $\mathbf{x}'_i = \mathrm{H}\mathbf{x}_i$，$i \in \{1,\ldots,3\}$，$\mathbf{e}' = \mathrm{H}\mathbf{e}$，通过它们可以算出 H.

因此由三组点对应计算 H，我们有两种可供选择的方法，第一种**显式**重构，第二种则是**隐式**方法，其中对极点提供一组点对应. 我们自然要问是否其中的一种比另一种好. 答案是：**不宜**用隐式方法来计算，因为在显式方法中不退化的情形可能在隐式中出现严重的退化.

考虑有 2 个像点与对极点共线的情形（我们暂且假定测量是没有噪声的）. 如果其中的三个点共线的话（见节 3.1.3（$p$55）），单应 H 不能由 4 组点对应计算得到，因此对于这种情形，隐式的方法失败. 类似地，如果图像点和对极点接近于共线，则隐式方法将给出 H 的一个非常病态的估计. 当两点与对极点共线或接近于共线时，显式方法都不会出现问题——对应图像点仍可确定 3D 空间中的点（世界点在同一对极平面上，但它不是一个退化情形）和平面 $\boldsymbol{\pi}$，从而相应的单应能求出. 此种结构在图 12.3 中作了说明.

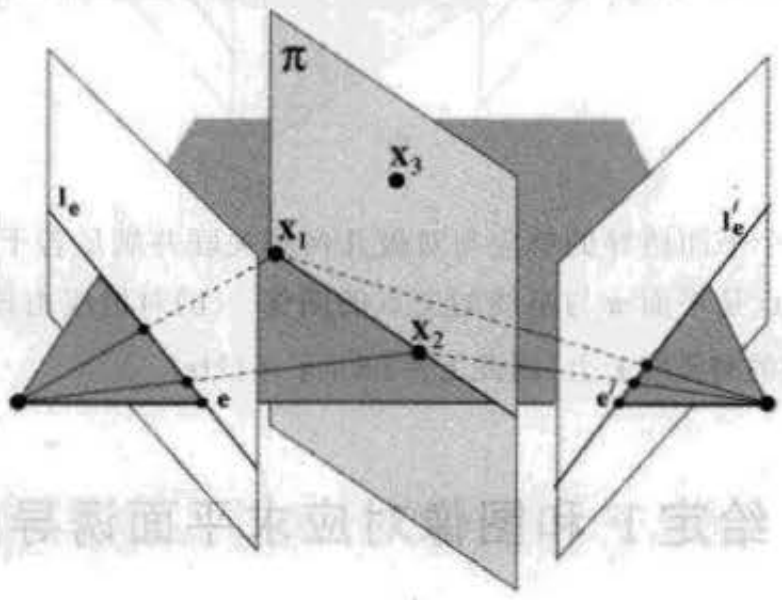



图 12.3 **用隐式方法计算单应时的退化几何**. 由点 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 定义的直线在一个极平面上，因而与基线相交. $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 的影像与对极点共线，因而 H 不能由对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$，$i \in \{1,\ldots,3\}$，$\mathbf{e} \leftrightarrow \mathbf{e}'$ 唯一地算出. 这一结构对显式方法不是退化的.

我们现在对显式方法作更详细的代数推导. 其中未必真要算出 $\mathbf{X}_i$ 的坐标，重要的是它们加在（与 F 兼容的）三参数单应簇的约束条件(12.4)：$\mathrm{H} = \mathrm{A} - \mathbf{e}'\mathbf{v}^{\mathrm{T}}$ 由 $\mathbf{v}$ 参数化. 于是问题简化为由三组点对应求解 $\mathbf{v}$，其解法如下：

**结论 12.6 已知 F 和三组图像点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$，则 3D 点所在平面诱导的单应是：**

$$\mathrm{H} = \mathrm{A} - \mathbf{e}'(\mathrm{M}^{-1}\mathbf{b})^{\mathrm{T}},$$

**其中 $\mathrm{A} = [\mathbf{e}']_\times \mathrm{F}$，$\mathbf{b}$ 是一个 3 维矢量，其元素为**

$$b_i = (\mathbf{x}'_i \times (\mathrm{A}\mathbf{x}_i))^{\mathrm{T}}(\mathbf{x}'_i \times \mathbf{e}')/\|\mathbf{x}'_i \times \mathbf{e}'\|^2,$$

**而 M 是一个行矢量为 $\mathbf{x}_i^{\mathrm{T}}$ 的 $3\times3$ 矩阵.**

**证明** 根据结论 8.14（$p$173），F 可以分解为 $\mathrm{F} = [\mathbf{e}']_\times \mathrm{A}$. 然后由(12.4)给出 $\mathrm{H} = \mathrm{A} - \mathbf{e}'\mathbf{v}^{\mathrm{T}}$，并且每一组对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 产生关于 $\mathbf{v}$ 的一个线性约束：

$$\mathbf{x}'_i = \mathrm{H}\mathbf{x}_i = \mathrm{A}\mathbf{x}_i - \mathbf{e}'(\mathbf{v}^{\mathrm{T}}\mathbf{x}_i), \quad i = 1,\ldots,3. \tag{12.5}$$

从(12.5)可知,矢量 $\mathbf{x}'_i$ 和 $\mathrm{A}\mathbf{x}_i - \mathbf{e}'(\mathbf{v}^\mathsf{T}\mathbf{x}_i)$ 是平行的,因此它们的矢量积是零:

$$\mathbf{x}'_i \times (\mathrm{A}\mathbf{x}_i - \mathbf{e}'(\mathbf{v}^\mathsf{T}\mathbf{x}_i)) = (\mathbf{x}'_i \times \mathrm{A}\mathbf{x}_i) - (\mathbf{x}'_i \times \mathbf{e}')(\mathbf{v}^\mathsf{T}\mathbf{x}_i) = \mathbf{0}.$$

再把它与矢量 $\mathbf{x}'_i \times \mathbf{e}'$ 进行数量积给出

$$\mathbf{x}_i^\mathsf{T}\mathbf{v} = \frac{(\mathbf{x}'_i \times (\mathrm{A}\mathbf{x}_i))^\mathsf{T}(\mathbf{x}'_i \times \mathbf{e}')}{(\mathbf{x}'_i \times \mathbf{e}')^\mathsf{T}(\mathbf{x}'_i \times \mathbf{e}')} = b_i \tag{12.6}$$

它关于 $\mathbf{v}$ 是线性的. 注意,方程与 $\mathbf{x}'$ 的尺度无关,因为 $\mathbf{x}'$ 在分子和分母上出现的次数相同. 每组对应产生一个方程 $\mathbf{x}_i^\mathsf{T}\mathbf{v} = b_i$,把它们联合在一起就得到 $\mathrm{M}\mathbf{v} = \mathbf{b}$. □

注意:如果 $\mathrm{M}^\mathsf{T} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$ 不是满秩的,则 $\mathbf{v}$ 无解. 从代数上说,若三个图像点 $\mathbf{x}_i$ 共线,则 $\det\mathrm{M} = 0$. 从几何上说,三个共线的图像点由共线或共面(该平面包含第一个摄像机的中心)的世界点产生. 总之,上述的任一种情形都不能定义一个满秩的单应.

**相容性条件** 方程(12.5)等价于 6 个约束,因为每组点对应向该单应施加两个约束. 但确定 $\mathbf{v}$ 只需三个约束,故一个有效解还必须满足剩下的三个约束. 这些约束可以由(12.5)与 $\mathbf{e}$ 的叉积得到,即

$$\mathbf{e}' \times \mathbf{x}'_i = \mathbf{e}' \times \mathrm{A}\mathbf{x}_i = \mathrm{F}\mathbf{x}_i.$$

方程 $\mathbf{e}' \times \mathbf{x}'_i = \mathrm{F}\mathbf{x}_i$ 是 $\mathbf{x}_i$ 和 $\mathbf{x}'_i$ 之间的一个相容约束,因为它与 $\mathbf{v}$ 无关. 这是一个关于对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 的(隐含的)对极几何约束:其左边是过 $\mathbf{x}'_i$ 的对极线,而右边是 $\mathbf{x}_i$ 在第二幅图像上对应的极线 $\mathrm{F}\mathbf{x}_i$,即该方程强迫 $\mathbf{x}'_i$ 在 $\mathbf{x}_i$ 的对极线上,因此该点对应与对极几何相容.

**由噪声点估计** 确定平面和单应的三组点对应必须满足对极几何的相容约束. 经测量得到的对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 一般不会准确地满足这个约束. 因此我们需要一个最优矫正测量点的过程,使得估计的点 $\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i$ 满足该对极几何约束. 庆幸的是,这样的过程在三角形算法 11.1 ($p$222)中已经给出,它可以在这里直接地采用. 这样,在高斯图像噪声的假设下,我们有一个关于 H 和 3D 点的最大似然估计. 这个方法概括在算法 12.1 中.

目标

给定 F 和三组点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$,它们是 3D 点 $\mathbf{X}_i$ 的影像,计算 $\mathbf{X}_i$ 所在平面诱导的单应 $\mathbf{x}' = \mathrm{H}\mathbf{x}$.

算法

(1) 用算法 11.1($p$222)计算每组对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 的矫正对应 $\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i$.

(2) 按结论 12.6,选择 $\mathrm{A} = [\mathbf{e}']_\times \mathrm{F}$,并解线性方程组 $\mathrm{M}\mathbf{v} = \mathbf{b}$ 求出 $\mathbf{v}$.

(3) 最后得 $\mathrm{H} = \mathrm{A} - \mathbf{e}'\mathbf{v}^\mathsf{T}$.

算法 12.1 由 3 点确定的平面所诱导的单应的最佳估计.

### 12.2.2 一点和一直线

本节为一点和一直线的对应所确定的平面推导单应的一个表达式. 我们先仅考虑直线对应并证明它把与 F 兼容的 3 参数单应簇(12.4)转化成 1 参数簇. 然后再证明点对应唯一地确定这张平面和对应的单应.

两条图像直线的对应确定 3D 空间的一条直线并且 3D 直线在一个单参数平面簇(束)上,参图 12.4. 这个平面束诱导两幅图像之间的一个单应束,而且束中的任何一个成员都使两条对应的直线相互映射.

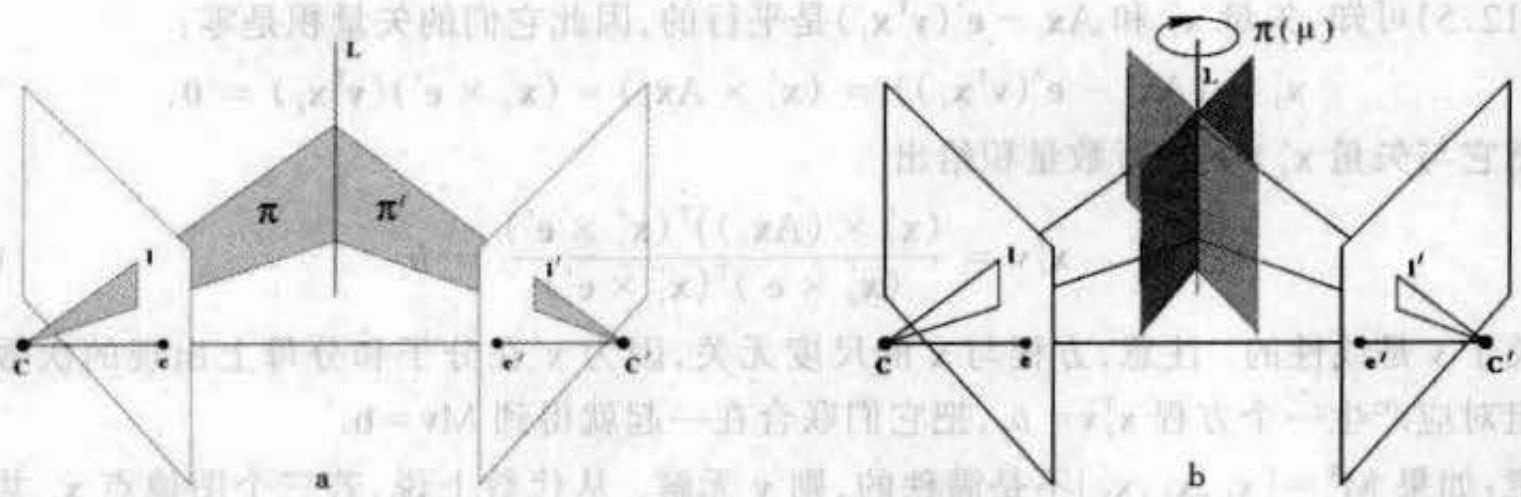



图 12.4 (a)图像直线 $\mathbf{l}$ 和 $\mathbf{l}'$ 分别确定平面 $\pi$ 和 $\pi'$. 这两个平面的交定义了 3 维空间的直线 $\mathbf{L}$. (b) 3 维空间的直线 $\mathbf{L}$ 包含在单参数平面簇 $\pi(\mu)$ 中. 该平面簇诱导图像之间的一个单参数单应簇.

**结论 12.7 如果 $\mathbf{l}'^{\mathsf{T}}\mathbf{e}'\neq 0$,则由直线对应 $\mathbf{l}\leftrightarrow\mathbf{l}'$ 定义的平面束诱导的单应是**

$$\mathrm{H}(\mu) = [\mathbf{l}']_\times \mathrm{F} + \mu\mathbf{e}'\mathbf{l}^{\mathsf{T}} \tag{12.7}$$

**其中 $\mu$ 是射影参数.**

**证明** 根据结论 7.1($p$135),直线 $\mathbf{l}$ 的反向投影为过第一个摄像机中心的一张平面 $\mathrm{P}^{\mathsf{T}}\mathbf{l}$,而 $\mathbf{l}'$ 则反向投影为过第二个摄像机中心的一张平面 $\mathrm{P}'^{\mathsf{T}}\mathbf{l}'$,如图 12.4a 所示. 这两张平面是由 $\mu$ 参数化的平面束的基. 如结论 12.3 的证明中所述,我们可以选 $\mathrm{P}=[\mathrm{I}|\mathbf{0}]$,$\mathrm{P}'=[\mathrm{A}|\mathbf{e}']$,那么该平面束是

$$\begin{aligned}\pi(\mu) &= \mu\mathrm{P}^{\mathsf{T}}\mathbf{l} + \mathrm{P}'^{\mathsf{T}}\mathbf{l}' \\ &= \mu\begin{pmatrix}\mathbf{l}\\ 0\end{pmatrix} + \begin{bmatrix}\mathrm{A}^{\mathsf{T}}\mathbf{l}'\\ \mathbf{e}'^{\mathsf{T}}\mathbf{l}'\end{bmatrix}.\end{aligned}$$

根据结论 12.1,诱导单应是 $\mathrm{H}(\mu)=\mathrm{A}-\mathbf{e}'\mathbf{v}(\mu)^{\mathsf{T}}$,其中

$$\mathbf{v}(\mu) = (\mu\mathbf{l} + \mathrm{A}^{\mathsf{T}}\mathbf{l}')/(\mathbf{e}'^{\mathsf{T}}\mathbf{l}'). \tag{12.8}$$

利用分解式 $\mathrm{A}=[\mathbf{e}']_\times\mathrm{F}$,我们得到

$$\begin{aligned}\mathrm{H} &= ((\mathbf{e}'^{\mathsf{T}}\mathbf{l}'\mathrm{I} - \mathbf{e}\mathbf{l}'^{\mathsf{T}})[\mathbf{e}']_\times \mathrm{F} - \mu\mathbf{e}'\mathbf{l}^{\mathsf{T}})/(\mathbf{e}'^{\mathsf{T}}\mathbf{l}') = -([\mathbf{l}']_\times[\mathbf{e}']_\times[\mathbf{e}']_\times \mathrm{F} + \mu\mathbf{e}'\mathbf{l}^{\mathsf{T}})/(\mathbf{e}'^{\mathsf{T}}\mathbf{l}') \\ &= -([\mathbf{l}']_\times \mathrm{F} + \mu\mathbf{e}'\mathbf{l}'^{\mathsf{T}})/(\mathbf{e}'^{\mathsf{T}}\mathbf{l}')\end{aligned}$$

其中最后一个等式由结论 A3.4($p$410)的公式 $[\mathbf{e}']_\times[\mathbf{e}']_\times\mathrm{F}=\mathrm{F}$ 得出. 此式在相差一个常数的意义下等于(12.7). □

**对应点和直线确定的单应** 由直线对应我们得到 $\mathrm{H}(\mu)=[\mathbf{l}']_\times\mathrm{F}+\mu\mathbf{e}'\mathbf{l}^{\mathsf{T}}$,现在利用点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 来求 $\mu$.

**结论 12.8 已知 F 及对应点 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 和对应直线 $\mathbf{l}\leftrightarrow\mathbf{l}'$,由相应 3D 点和直线所在平面所诱导的单应是**

$$\mathrm{H} = [\mathbf{l}']_\times \mathrm{F} + \frac{(\mathbf{x}'\times\mathbf{e}')^{\mathsf{T}}(\mathbf{x}'\times((\mathrm{F}\mathbf{x})\times\mathbf{l}'))}{\|\mathbf{x}'\times\mathbf{e}'\|^2(\mathbf{l}^{\mathsf{T}}\mathbf{x})}\mathbf{e}'\mathbf{l}^{\mathsf{T}}.$$

它的推导与结论 12.6 类似. 与三点情形一样,图像点的对应必须与对极几何相容. 这意味着在使用结论 12.8 之前,测量(噪声)点必须用算法 11.1($p$222)矫正. 但是目前还没有关于直线的相容性约束也没有现成的矫正方法.

**点映射 $\mathrm{H}(\mu)$ 的几何解释** 值得对映射 $\mathrm{H}(\mu)$ 作进一步探讨. 因为 $\mathrm{H}(\mu)$ 与对极几何相容,第

一幅视图的点 $\mathbf{x}$ 被映射到第二幅视图上与 $\mathbf{x}$ 对应的对极线 $\mathbf{Fx}$ 上的一点 $\mathbf{x}'=\mathrm{H}(\mu)\mathbf{x}$. 一般来说,点 $\mathbf{x}'=\mathrm{H}(\mu)\mathbf{x}$ 在对极线上的位置随 $\mu$ 而变. 然而,如果点 $\mathbf{x}$ 在直线 $\mathbf{l}$ 上(因此 $\mathbf{l}^{\mathrm{T}}\mathbf{x}=0$),则

$$\mathbf{x}' = \mathrm{H}(\mu)\mathbf{x} = ([\mathbf{l}']_\times \mathrm{F} + \mu\mathbf{e}'\mathbf{l}^{\mathrm{T}})\mathbf{x} = [\mathbf{l}']_\times \mathrm{F}\mathbf{x}$$

便与 $\mu$ 的值无关,仅依赖于 F. 因此,如图 12.5 所示,对极几何为该直线上的点定义了一个点到点的映射.

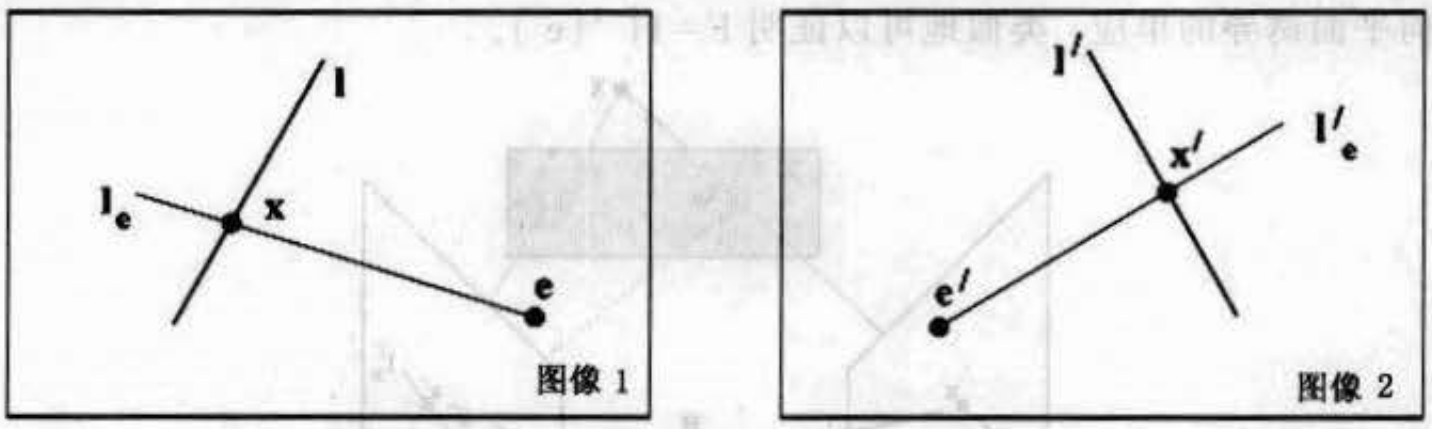



图 12.5 对极几何诱导对应直线 $\mathbf{l}\leftrightarrow\mathbf{l}'$ 之间的一个单应,该对应直线是 3 维空间中的直线 L 的影像. $\mathbf{l}$ 上的点映射到 $\mathbf{l}'$ 上的点,即 $\mathbf{x}'=[\mathbf{l}']_\times\mathrm{F}\mathbf{x}$,其中 $\mathbf{x}$ 与 $\mathbf{x}'$ 是 L 与 $\mathbf{l}_e$ 和 $\mathbf{l}'_e$ 组成的对极平面的交点的影像.

**退化单应** 我们讲过,如果该世界平面包含一个摄影机中心,则所诱导的单应是退化的. 表示该单应的矩阵不满秩,并且平面上的点被映射成一条直线(如果 H 的秩为 2)或者一个点(如果 H 的秩为 1). 尽管如此,由(12.7)仍能得到退化单应的一个显式表达式. 退化(奇异)单应在该单应束中的 $\mu=\infty$ 和 $\mu=0$ 处. 它们分别对应于过第一和第二摄像机中心的平面. 图 12.6 显示了该平面包含第二摄像机中心并与图像平面相交于直线 $\mathbf{l}'$ 的情形. 第一幅视图上的点 $\mathbf{x}$ 被映射到 $\mathbf{l}'$ 上的点 $\mathbf{x}'$,其中

$$\mathbf{x}' = \mathbf{l}' \times \mathrm{F}\mathbf{x} = [\mathbf{l}']_\times \mathrm{F}\mathbf{x}.$$

因此,该单应是 $\mathrm{H}=[\mathbf{l}']_\times\mathrm{F}$,它是一个秩为 2 的矩阵.

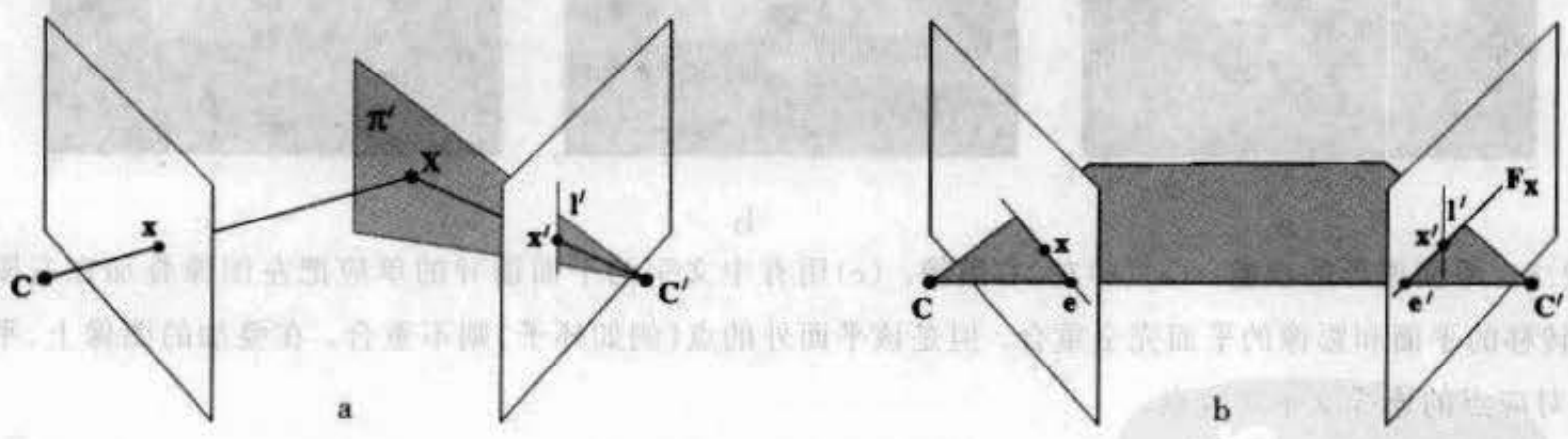



图 12.6 **一种退化单应**. (a)由过第二个摄像机中心的平面诱导的映射是一个退化的单应 $\mathrm{H}=[\mathbf{l}']_\times\mathrm{F}$. 平面 $\pi'$ 与第二张像平面交于直线 $\mathbf{l}'$. 第一幅视图上的所有点都被映射到第二幅视图的直线 $\mathbf{l}'$ 上的点. (b) 第一幅视图上的点 $\mathbf{x}$ 被映射到 $\mathbf{x}$ 的对极线 $\mathbf{Fx}$ 与 $\mathbf{l}'$ 的交点 $\mathbf{x}'$,因此 $\mathbf{x}'=\mathbf{l}'\times\mathbf{Fx}$.

## 12.3 由平面诱导的单应来计算 F

到目前为止,我们都假定 F 已知,并且在提供了不同的附加信息后,实现计算 H 的目标,现在我们反过来做,并且证明如果 H 给定,则在提供了附加信息后也可以计算 F. 我们先介绍一个重要的几何概念:相对某张平面的视差,它将使代数推导更直接.

**平面诱导的视差** 如示意图 12.7 和图 12.8 给出的例子所示，由平面诱导的单应产生了一个虚拟视差(见节 7.4.5($p$144)). 它的重要之处在于 3D 点 $\mathbf{X}$ 在第二幅视图上的影像 $\mathbf{x}'$ 和由该单应映射的点 $\tilde{\mathbf{x}}' = \mathrm{H}\mathbf{x}$ 都在 $\mathbf{x}$ 的对极线上：因为这两点都是过 $\mathbf{x}$ 的射线上的点的影像. 因此，直线 $\mathbf{x}' \times (\mathrm{H}\mathbf{x})$ 是第二视图上的对极线并且为对极点的位置提供了一个约束. 一旦对极点被确定(这样的两个约束就足够了)，则如结论 8.1($p$162)所证明的 $\mathrm{F} = [\mathbf{e}']_\times \mathrm{H}$，其中 H 是由任何平面诱导的单应. 类似地可以证明 $\mathrm{F} = \mathrm{H}^{-\mathsf{T}}[\mathbf{e}']_\times$.

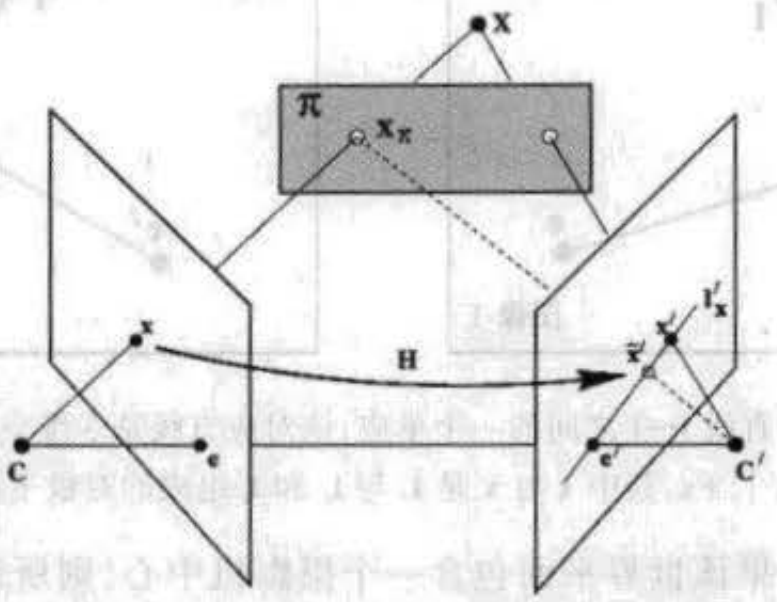



图 12.7 **平面诱导的视差**. 过 $\mathbf{X}$ 的射线与平面 $\pi$ 交于点 $\mathbf{X}_\pi$. $\mathbf{X}$ 和 $\mathbf{X}_\pi$ 的影像在第一视图重合于 $\mathbf{x}$；在第二视图的影像分别是点 $\mathbf{x}'$ 和 $\tilde{\mathbf{x}}' = \mathrm{H}\mathbf{x}$. 这两点不重合(除非 $\mathbf{X}$ 在 $\pi$ 上)，但都在 $\mathbf{x}$ 的对极线 $\mathbf{l}'_{\mathbf{x}}$ 上. 点 $\mathbf{x}'$ 和 $\tilde{\mathbf{x}}'$ 之间的矢量是相对于由平面 $\pi$ 诱导的单应的视差. 注意如果 $\mathbf{X}$ 在平面 $\pi$ 的另一边，那么 $\tilde{\mathbf{x}}'$ 将在 $\mathbf{x}'$ 的另一边.

a b c

图 12.8 **平面诱导的视差**. (a)(b)左、右图像. (c)用有中文字的平面诱导的单应把左图像叠加在右图像上. 转移的平面和影像的平面完全重合. 但是该平面外的点(例如杯子)则不重合. 在叠加的图像上，平面外的对应点的连线交于对极点.

算法 12.2 给出虚拟视差的一个应用，它指出 F 可以由 6 点的影像唯一地算出，其中 4 点共面，其余两点不在该平面上. 4 个共面点的影像确定单应，而两个不在该平面上的点提供足以确定对极点的约束. 这个 6 点结论非常令人吃惊，因为对于一般位置上的 7 个点存在 F 的 3 个解(见节 10.1.2($p$192)).

**射影深度** 一个世界点 $\mathbf{X} = (\mathbf{x}^\mathsf{T}, \rho)^\mathsf{T}$ 被影像在第一视图上的 $\mathbf{x}$ 和在第二视图上的

$$\mathbf{x}' = \mathrm{H}\mathbf{x} + \rho\mathbf{e}'. \tag{12.9}$$

注意 $\mathbf{x}'$，$\mathbf{e}'$，$\mathrm{H}\mathbf{x}$ 共线. 标量 $\rho$ 是**相对**单应 H 的视差，并且可以解释为相对平面 $\pi$ 的"深度". 若 $\rho = 0$，则 3D 点 $\mathbf{X}$ 在平面上，否则 $\rho$ 的"符号"将表示点 $\mathbf{X}$ 在平面 $\pi$ 的哪"一边"(见图 12.7 和图 12.8). 陈述这些必须小心，因为在非定向射影几何下，一个齐次对象的符号和一个平

目标

已知 6 组点对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$，它们是 3D 点 $\mathbf{X}_i$ 的影像，其中前 4 个 3D 点 $i \in \{1,\ldots,4\}$ 共面，计算基本矩阵 F.

算法

(1) 计算单应 H，使得 $\mathbf{x}_i' = \mathrm{H}\mathbf{x}_i$，$i \in \{1,\ldots,4\}$.

(2) 确定作为直线 $(\mathrm{H}\mathbf{x}_5) \times \mathbf{x}_5'$ 和 $(\mathrm{H}\mathbf{x}_6) \times \mathbf{x}_6'$ 的交点的对极点 $\mathbf{e}'$.

(3) $\mathrm{F} = [\mathbf{e}']_\times \mathrm{H}$.

参看图 12.9.

算法 12.2　给定 6 点对应(其中 4 点共面)计算 F.

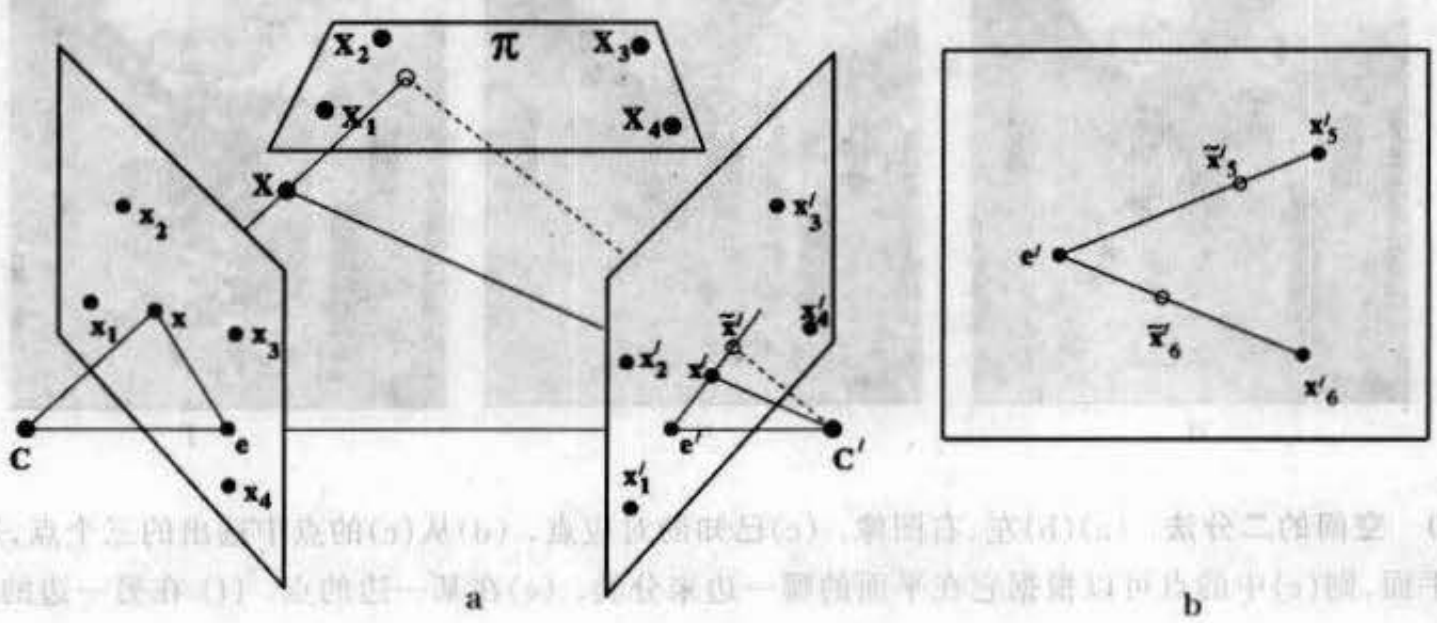



图 12.9　基本矩阵可以由其中 4 点共面的 6 个 3D 点的影像唯一确定. (a)点 $\mathbf{X}$ 的视差. (b)对极点由两条视差直线：$\tilde{\mathbf{x}}_5' = \mathrm{H}\mathbf{x}_5$ 到 $\mathbf{x}_5'$ 的连线和 $\tilde{\mathbf{x}}_6' = \mathrm{H}\mathbf{x}_6$ 到 $\mathbf{x}_6'$ 的连线的交点确定.

面的两边没有意义.

**例 12.9　空间的二分法** 虚拟视差的符号(sign(ρ))可以用来计算由平面 $\boldsymbol{\pi}$ 产生的 3D 空间的一个划分. 假设给定 F 和由其对应的图像点确定的三个空间点. 那么由这三个点定义的平面可以用来把所有其它对应点划分为在这个平面的两边(或在这个平面上)的集合. 图 12.10 给出了一个例子. 注意，这三个点没有必要是实际物理点的影像，因此这个方法可用于虚拟平面. 联合几个平面，可以辨认 3D 的一个区域. △

**两个平面**　假定在景物空间中有两个平面 $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2$，它们分别诱导了单应 $\mathrm{H}_1, \mathrm{H}_2$. 显然每一个平面提供了不在另一个平面上的信息，加上视差的概念，这两个单应应该足以确定 F. 事实上在这个配置下 F 是超定的，它意味着两个单应还必须满足相容性约束.

考虑图 12.11. 单应 $\mathrm{H} = \mathrm{H}_2^{-1}\mathrm{H}_1$ 是一个从第一幅图像到它自身的映射. 在这个映射下，对极点 $\mathbf{e}$ 是不动点，即 $\mathrm{H}\mathbf{e} = \mathbf{e}$，因此可以从 H 的(非退化)特征矢量确定. 那么可以根据结论 8.1($p$162)计算基本矩阵 $\mathrm{F} = [\mathbf{e}']_\times \mathrm{H}_i$，其中 $\mathbf{e}' = \mathrm{H}_i\mathbf{e}$($i = 1$ 或 2). 从图 12.11 可以看出映射 H 另外的一些性质. 该映射有一条不动点的直线和不在该直线上的一个不动点(关于不动点和不动直线见节 1.9($p$30)). 这就是说 H 的两个特征值相等. 事实上 H 是一个平面透射(见节 A5.2($p$427)). 反过来，$\mathrm{H} = \mathrm{H}_2^{-1}\mathrm{H}_1$ 的这些性质定义了关于 $\mathrm{H}_1$ 和 $\mathrm{H}_2$ 的相容性约束，

使得它们的复合有所述性质.

到目前为止,本章的结论完全是关于射影的. 现在来介绍关于仿射的结论.

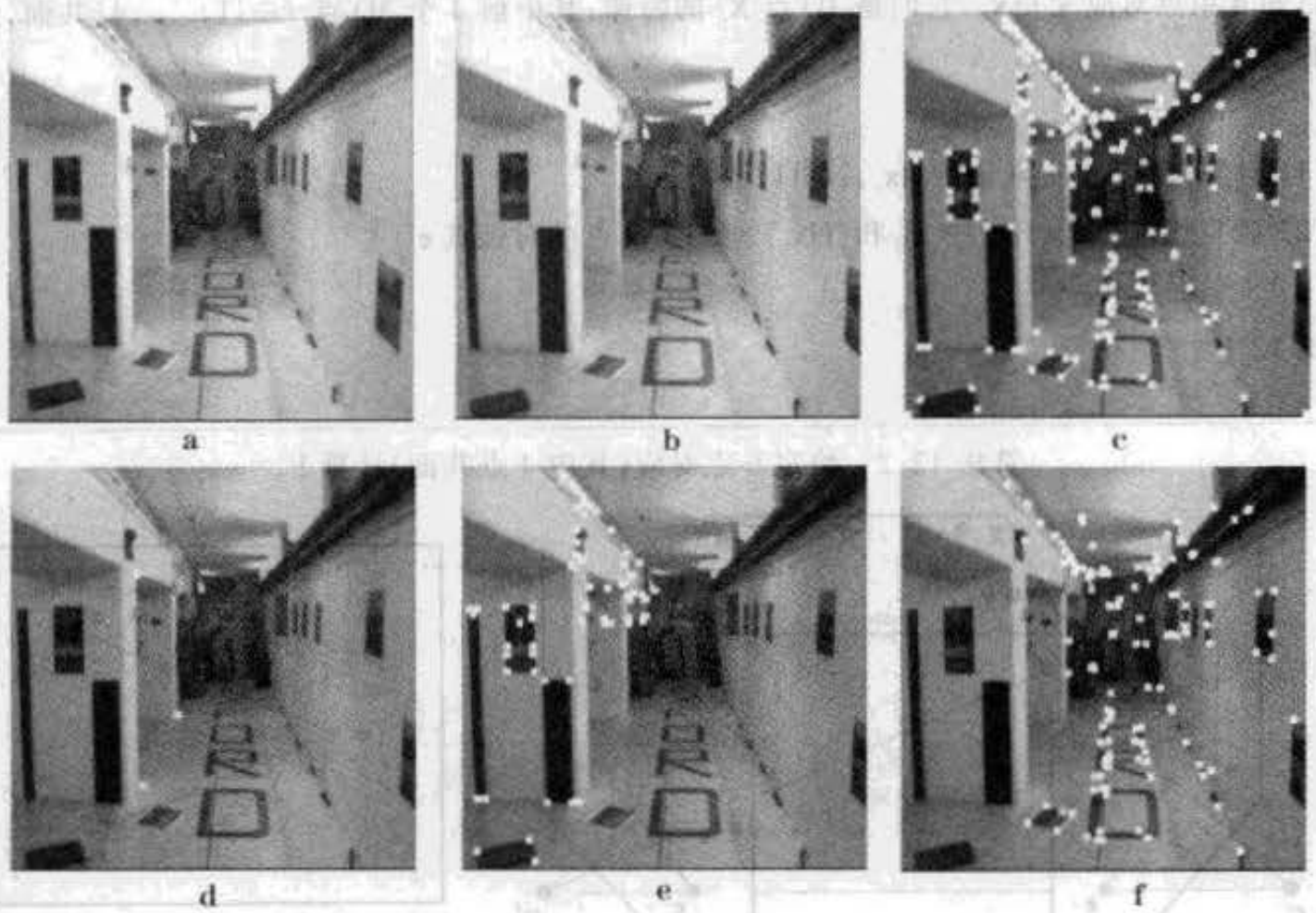



图 12.10 **空间的二分法**. (a)(b)左、右图像. (c)已知的对应点. (d)从(c)的点中选出的三个点,这三点定义一个平面,则(c)中的点可以根据它在平面的哪一边来分类. (e)在某一边的点. (f) 在另一边的点.

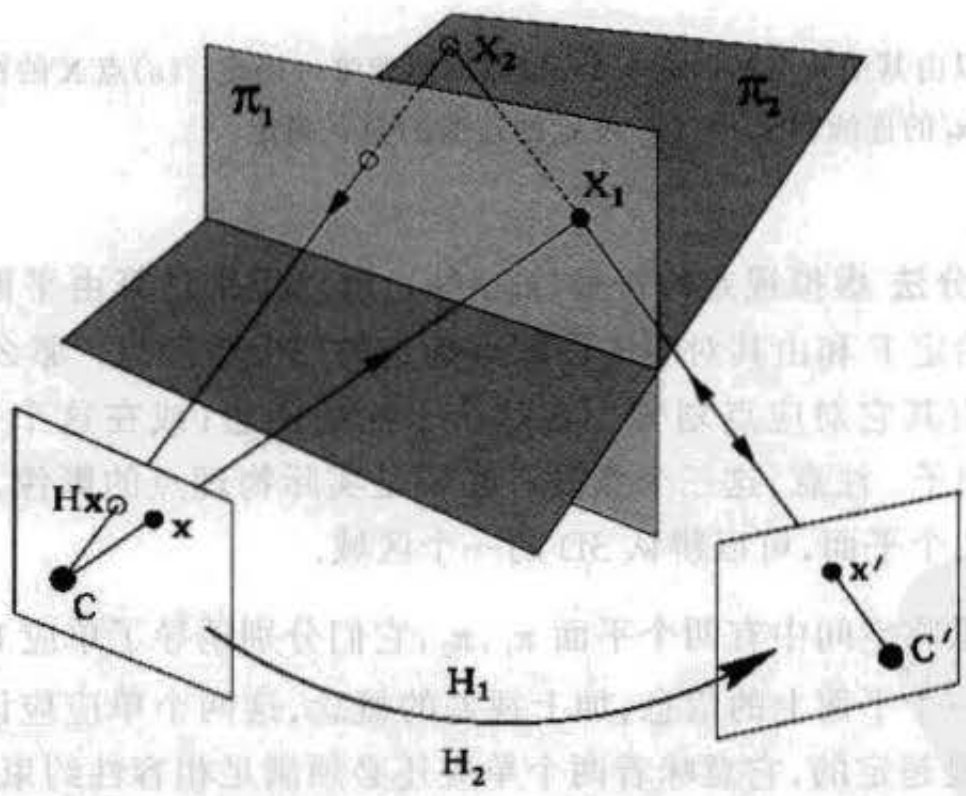



图 12.11 映射 $H=H_2^{-1}H_1$ 对第一幅图像中点 **x** 的作用是首先把它作为 3D 点 $\mathbf{X}_1$ 的影像而转移到 **x**′,然后把它作为 3D 点 $\mathbf{X}_2$ 的影像再映射回第一幅图像. 第一幅视图中,位于这两平面交线的影像直线上的点将被映射到它们自身,即是此映射的不动点. 对极点 **e** 也是此映射的一个不动点.

## 12.4 无穷单应 $H_\infty$

无穷远平面是一个特别重要的平面，由该平面诱导的单应该用特殊名字加以区别：

**定义 12.10** 由无穷远平面 $\boldsymbol{\pi}_\infty$ 诱导的单应称为无穷单应 $H_\infty$.

该单应的形式可以用一个极限过程来推导，根据(12.2－$p$229)$H=K'(R-\mathbf{t}\mathbf{n}^T/d)K^{-1}$，其中 $d$ 是第一个摄像机到该平面的垂直距离，取极限得

$$H_\infty = \lim_{d\to\infty} H = K'RK^{-1}.$$

此式意味着 $H_\infty$ 与两幅视图之间的平移无关，而仅依赖于旋转和摄像机的内参数. 另一种推导是根据(8.7－$p$168)，对应的图像点有下列关系式：

$$\mathbf{x}' = K'RK^{-1}\mathbf{x} + K'\mathbf{t}/Z = H_\infty\mathbf{x} + K'\mathbf{t}/Z, \tag{12.10}$$

其中 Z 是从第一摄像机测量的深度. 我们可以再一次看到被 $H_\infty$ 映射的是无穷远点($Z=\infty$). 注意若在(12.10)中平移 $\mathbf{t}$ 为 $\mathbf{0}$，也可得到 $H_\infty$，它表示摄像机绕其中心的旋转，因此若摄像机绕其中心旋转，则 $H_\infty$ 是与**任何**深度的图像点都对应的单应(见节 7.4($p$140)).

因 $\mathbf{e}'=K'\mathbf{t}$，故(12.10)可写为 $\mathbf{x}'=H_\infty\mathbf{x}+\mathbf{e}'/Z$，对照(12.9)表明$(1/Z)$扮演了 $\rho$ 的角色. 这就是说欧氏深度的倒数可以理解为相对于 $\boldsymbol{\pi}_\infty$ 的视差.

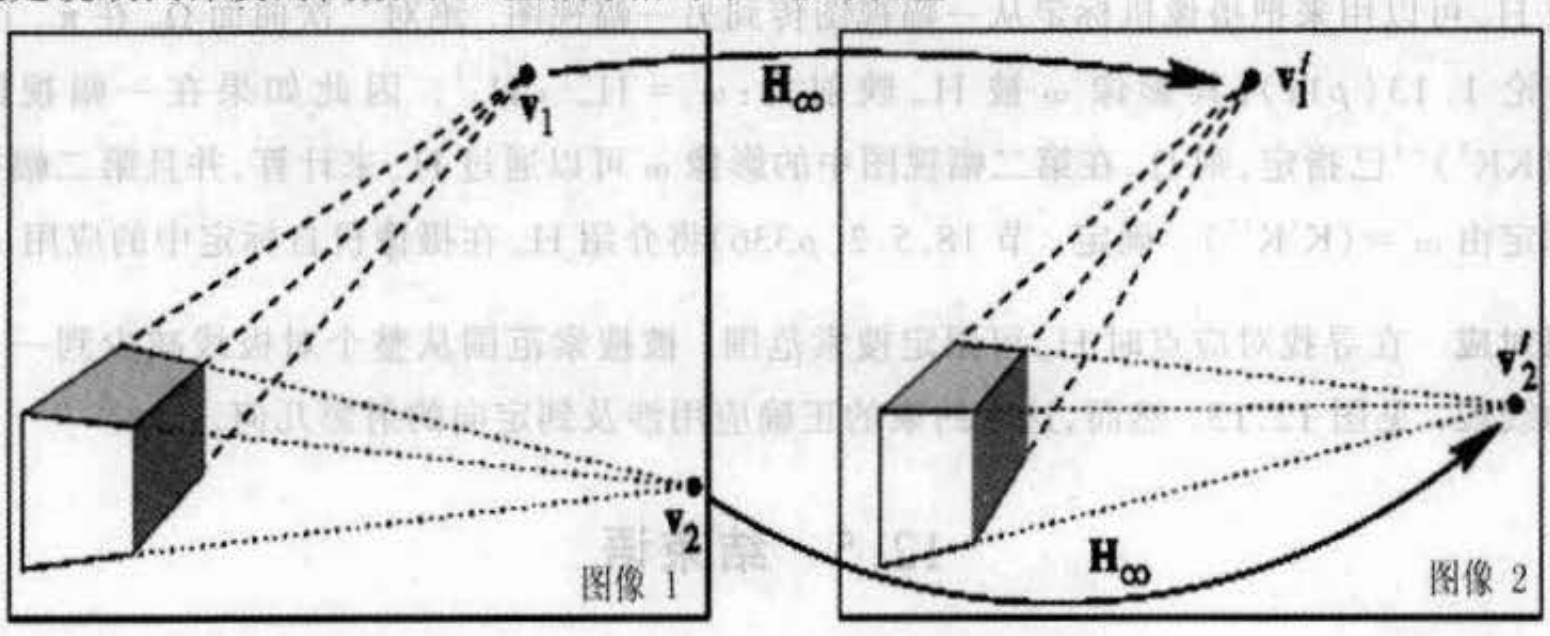



图 12.12 无穷单应 $H_\infty$ 映射两幅图像之间消影点.

**消影点与消影直线** $\boldsymbol{\pi}_\infty$ 上的点的影像由 $H_\infty$ 映射. 这些影像是消影点，因此 $H_\infty$ 映射两幅图像之间的消影点，即 $\mathbf{v}'=H_\infty\mathbf{v}$，其中 $\mathbf{v}'$ 和 $\mathbf{v}$ 为对应的消影点，见图 12.12. 因此，利用结论 12.6，$H_\infty$ 可由三个(非共线)消影点的对应和 F 算出. 或者如节 12.2.2 所介绍的，$H_\infty$ 也可由一组消影直线对应和一组消影点对应(不在该直线上)和 F 一起算得.

**仿射与度量重构** 第 9 章指出辨认了 $\boldsymbol{\pi}_\infty$ 可以使一个射影重构升级到仿射重构. 毫不奇怪，由于 $H_\infty$ 与 $\boldsymbol{\pi}_\infty$ 相联系，它自然会在该矫正过程中出现. 事实上，如果选取的摄像机矩阵为 $P=[I|\mathbf{0}]$ 和 $P'=[H_\infty|\lambda\mathbf{e}']$，则重构是仿射的.

反过来，假定世界坐标系是仿射的(即 $\boldsymbol{\pi}_\infty$ 取其标准位置 $\boldsymbol{\pi}_\infty=(0,0,0,1)^T$)：则 $H_\infty$ 可以由摄像机投影矩阵直接确定. 假设 $M,M'$ 分别是 P 和 $P'$ 左边的 $3\times3$ 子矩阵. 则在 $\boldsymbol{\pi}_\infty$ 上的点 $\mathbf{X}=(\mathbf{x}_\infty^T,0)^T$ 在两视图中的影像为 $\mathbf{x}=P\mathbf{X}=M\mathbf{x}_\infty$ 和 $\mathbf{x}'=P'\mathbf{X}=M'\mathbf{x}_\infty$. 因此 $\mathbf{x}'=M'M^{-1}\mathbf{x}$,

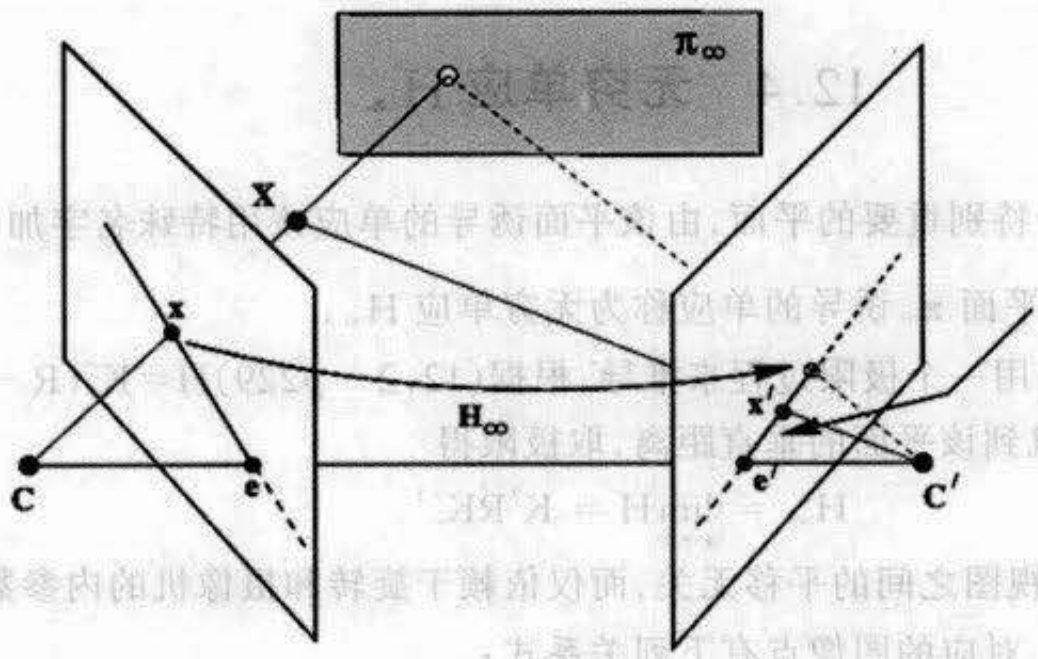



图 12.13 **用 $H_\infty$ 减少搜索范围**. 3 维空间的点不会比 $\pi_\infty$"更"远. $H_\infty$ 实现这一约束把对极线上的搜索限制在一个方向上. 摄像机之间的基线把每一个极平面分成两部分. 左图中在对极线上一"边"的点将被映射到右图中的对极线上的对应"边"(图中用实线指示). 因此对极点阻止了在另一方向的搜索.

从而

$$H_\infty = M'M^{-1}. \tag{12.11}$$

单应 $H_\infty$ 可以用来把摄像机标定从一幅视图传到另一幅视图. 绝对二次曲面 $\Omega_\infty$ 在 $\pi_\infty$ 上,根据结论 1.13($p$11),其影像 $\omega$ 被 $H_\infty$ 映射为:$\omega' = H_\infty^{-T}\omega H_\infty^{-1}$. 因此如果在一幅视图中 $\omega=(KK^T)^{-1}$ 已指定,则 $\Omega_\infty$ 在第二幅视图中的影像 $\omega'$ 可以通过 $H_\infty$ 来计算,并且第二幅视图的标定由 $\omega'=(K'K'^T)^{-1}$ 确定. 节 18.5.2($p$336)将介绍 $H_\infty$ 在摄像机自标定中的应用.

**双眼对应** 在寻找对应点时 $H_\infty$ 可限定搜索范围. 被搜索范围从整个对极线减少到一个有界的线段. 见图 12.13. 然而,这个约束的正确应用涉及到定向的射影几何.

## 12.5 结束语

本章对平面的射影技术做了大量说明,这些方法可以应用于其它许多曲面. 平面就是一个 3 个自由度的参数化曲面. 可以为其它曲面给出非常类似的推导,其中的自由度由曲面上点的影像来确定. 比如在二次曲面的情形,该曲面可由其上点的影像确定,和/或(平面不可能有此推广)由每一幅视图中它的轮廓线来确定[Cross-98, Shashua-97]. 本章讨论的有关概念,如曲面诱导的转移、当曲面不能完全由其像确定时的曲面簇、曲面诱导的视差、相容性约束、隐式计算、退化几何,都可移植到其它曲面的情形.

### 12.5.1 文献

对极几何和诱导单应的兼容性由 Luong & Viéville[Luong-96]做了彻底的研究. F 的六点算法出现在 Beardsley 等[Beardsley-92]和[Mohr-92]中. 给定两个平面求 F 的算法出现在 Sinclair[Sinclair-92]之中. [Zeller-96]给出了仅从对极几何和它们的图像投影来确定其性质的许多配置的例子,他同时对退化的情况进行了归类.

### 12.5.2 注释和练习

(1) **由平面诱导的单应**(12.1 - $p$229).

(a) 单应 H 的逆由下式给出

$$H^{-1} = A^{-1}\left(I + \frac{\mathbf{a}\mathbf{v}^{\mathsf{T}}A^{-1}}{1 - \mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{a}}\right)$$

假设 $A^{-1}$存在它. 有时被称为 Sherman - Morrison 公式.

(b) 证明:如果平面包含第二摄像机的中心,则单应 H 是退化的. 提示,在这种情形,$\mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{a}=1$ 且注意 $H=A(I-A^{-1}\mathbf{a}\mathbf{v}^{\mathsf{T}})$.

(2) 证明:如果摄像机作平面运动,其平移平行于该平面而旋转平行于该平面的法线,那么由这个平面诱导的单应共轭于一个平面欧氏变换. 证明:该单应的不动点是这个平面虚圆点的影像.

(3) 利用(12.2 - $p$229)证明:如果摄像机作一个纯平移,则由这个平面诱导的单应是一个平面透射(如节 A5.2($p$427)所定义),并且有一条不动点的直线对应于这个平面的消影线. 进一步证明:如果平移平行于这个平面,则该单应是一个约束透视变换(如节 A5.3($p$430)所定义).

(4) 证明两条空间直线共面的一个必要但不充分的条件是$(\mathbf{l}_1'\times\mathbf{l}_2')^{\mathsf{T}}F(\mathbf{l}_1\times\mathbf{l}_2)=0$,为什么它不是一个充分条件?

(5) **直线和平面的交点** 通过勾画配置(假定是一般位置)的草图,验证下面的每一个结论. 对于每一种情形,确定使这个结论无效的退化配置.

(a) 假设 3D 直线 $\mathbf{L}$ 被影像为 $\mathbf{l}$ 和 $\mathbf{l}'$,而平面 $\boldsymbol{\pi}$ 诱导单应 $\mathbf{x}'=H\mathbf{x}$. 则 $\mathbf{L}$ 和 $\boldsymbol{\pi}$ 的交点在第一幅图像被影像为 $\mathbf{x}=\mathbf{l}\times(H^{\mathsf{T}}\mathbf{l}')$,在第二幅图像为 $\mathbf{x}'=\mathbf{l}'\times(H^{\mathsf{T}}\mathbf{l})$.

(b) 无穷单应可以用来寻找在两幅图像上都可视的直线的消影点. 若 $\mathbf{l}$ 和 $\mathbf{l}'$是在两幅图像上的对应直线,而 $\mathbf{v},\mathbf{v}'$ 分别是它们在每幅图像上的消影点,则 $\mathbf{v}=\mathbf{l}\times(H_\infty^{\mathsf{T}}\mathbf{l}')$,$\mathbf{v}'=\mathbf{l}'\times(H_\infty^{-\mathsf{T}}\mathbf{l})$

(c) 若平面 $\boldsymbol{\pi}_1$ 和 $\boldsymbol{\pi}_2$ 诱导的单应分别是 $\mathbf{x}'=H_1\mathbf{x}$ 和 $\mathbf{x}'=H_2\mathbf{x}$,则 $\boldsymbol{\pi}_1$ 和 $\boldsymbol{\pi}_2$ 的交线在第一幅图像上的像满足 $H_1^{\mathsf{T}}H_2^{-\mathsf{T}}\mathbf{l}=\mathbf{l}$,并且可以由平面透射 $H_1^{\mathsf{T}}H_2^{-\mathsf{T}}$的实特征向量确定(见图 12.11).

(6) **四点的共面性** 设 F 和四组图像对应点 $\mathbf{x}_i\leftrightarrow\mathbf{x}_i'$ 已知. 怎样确定这四点的前象是否共面? 一种可能是通过结论 12.6($p$232)利用其中三个点确定一个单应,然后测量第四个点的转移误差. 第二种可能是计算连接图像点的直线,并确定直线的交点是否满足对极几何约束(见[Faugeras - 92a]). 第三种可能是计算从对极点到图像点的四条直线的交比——如果这四个景物点是共面的,那么该交比在两幅图像上的值相同. 因此这个等式是共面的必要条件,但它也是一个充分条件吗? 当存在测量误差(噪声)时应该采用什么统计检验?

(7) 证明:对极几何能唯一地从四条共面直线和这些直线的平面外的两点的影像计算得到. 如果直线中的两条直线用点来代替,此对极几何仍能被计算出来吗?

(8) 由摄像机矩阵 $P=[M|\mathbf{m}]$,$P'=[M'|\mathbf{m}']$出发,证明由一个平面 $\boldsymbol{\pi}=(\hat{\boldsymbol{\pi}}^{\mathsf{T}},\pi_4)^{\mathsf{T}}$ 诱导的单应 $\mathbf{x}'=H\mathbf{x}$ 由下式给出:

$H = M'(I - \mathbf{t}\mathbf{v}^T)M^{-1}$，其中 $\mathbf{t} = (M'^{-1}\mathbf{m}' - M^{-1}\mathbf{m})$，$\mathbf{v} = \hat{\boldsymbol{\pi}}/(\pi_4 - \hat{\boldsymbol{\pi}}^T M^{-1}\mathbf{m})$.

(9) 证明：按结论 12.6($p$232)计算的单应与 F 的尺度无关. 选择 F 的任意一个固定尺度，使得 F 不再是一个齐次量，从一个具有固定尺度的矩阵 $\tilde{F}$ 开始，证明如果 $H = [\mathbf{e}']_\times \tilde{F} - \mathbf{e}'(M^{-1}\tilde{\mathbf{b}})^T$，其中 $\tilde{b}_i = \mathbf{c}_i'^T(\tilde{F}\mathbf{x}_i)$，则用 $\lambda\tilde{F}$ 代替 $\tilde{F}$ 等价于把 H 用 $\lambda$ 做比例缩放.

(10) 给定一条(平面)二次曲线的两幅透视图像和视图之间的基本矩阵，则二次曲线的平面(进而由此平面诱导的单应)在两重歧义下被定义. 假设该二次曲线的图像是 C 和 C′，则诱导的单应是 $H(\mu) = [C'\mathbf{e}']_\times F - \mu\mathbf{e}'(C\mathbf{e})^T$，其中 $\mu$ 的两个值从下式得到

$$\mu^2[(\mathbf{e}^T C\mathbf{e})C - (C\mathbf{e})(C\mathbf{e})^T](\mathbf{e}'^T C'\mathbf{e}') = -F^T[C'\mathbf{e}']_\times C'[C'\mathbf{e}']_\times F.$$

细节在[Schmid－98]中给出.

(a) 从几何角度证明为了与对极几何兼容，二次曲线必须满足这样一个相容性约束：对极切线是对应的对极线(见图 10.6($p$204)). 现在由上面的 $H(\mu)$ 出发，用代数方法推导这个结论.

(b) 若对极点在这条二次曲线上，则这个代数表达式是无效的(因为 $\mathbf{e}^T C\mathbf{e} = \mathbf{e}'^T C'\mathbf{e}' = 0$). 这种退化性是几何的或者仅仅是该表达式的？

(11) **估计** 假设 $n > 3$ 个点 $\mathbf{X}_i$ 在 3 维空间的一个平面上，我们希望在给定 F 和它们的图像对应的条件下最优估计由该平面诱导的单应. 那么该单应的 ML 估计(如通常一样，假定测量噪声是独立高斯分布)由估计平面 $\hat{\boldsymbol{\pi}}$(3 自由度)和 $n$ 个点 $\hat{\mathbf{X}}_i$(因为点在平面上，每点有 2 个自由度)得到，该估计应最小化 $n$ 个点的重投影误差.

# 第 13 章　仿射对极几何

本章重叙前一章关于两视图几何的有关内容与推导，但是要用仿射摄像机代替那里的射影摄像机．仿射摄像机非常有用并且是许多实际情形的良态的近似．其最大优点在于它的线性特征，因而许多最优算法可以用线性代数（矩阵求逆，SVD 等）实现，而针对射影摄像机的解法或者涉及到高阶多项式（比如三角形法）或者仅能用数值最小化（例如 F 的黄金标准估计）才能实现．

我们首先介绍两个仿射摄像机的对极几何性质和根据点对应的最优计算．其次介绍三角形法和仿射重构．最后扼要地介绍一下由平行射影重构引起的多义性，以及由对极几何计算非多义的运动参数．

## 13.1　仿射对极几何

两个仿射摄像机的对极几何在许多方面与两个透视摄像机的对极几何相同，例如，一幅视图上的一个点在另一幅视图上定义了一条对极线，并且这样的对极线束相交于对极点．所不同的是由于摄像机是仿射的，其中心在无穷远处，而且是从景物到图像的平行投影．这一事实导致仿射对极几何的某些简化：

**对极线**　考虑第一幅视图中的两点 $\mathbf{x}_1, \mathbf{x}_2$．这些点反向投影为 3D 空间中的平行射线，因为所有投影射线是平行的．在第二幅视图中的对极线是反向投影的射线的图像．这两条射线在第二视图中的影像也是平行的，因为一个仿射摄像机把平行景物直线映射到平行图像直线．因此，所有对极线平行，对极平面也是如此．

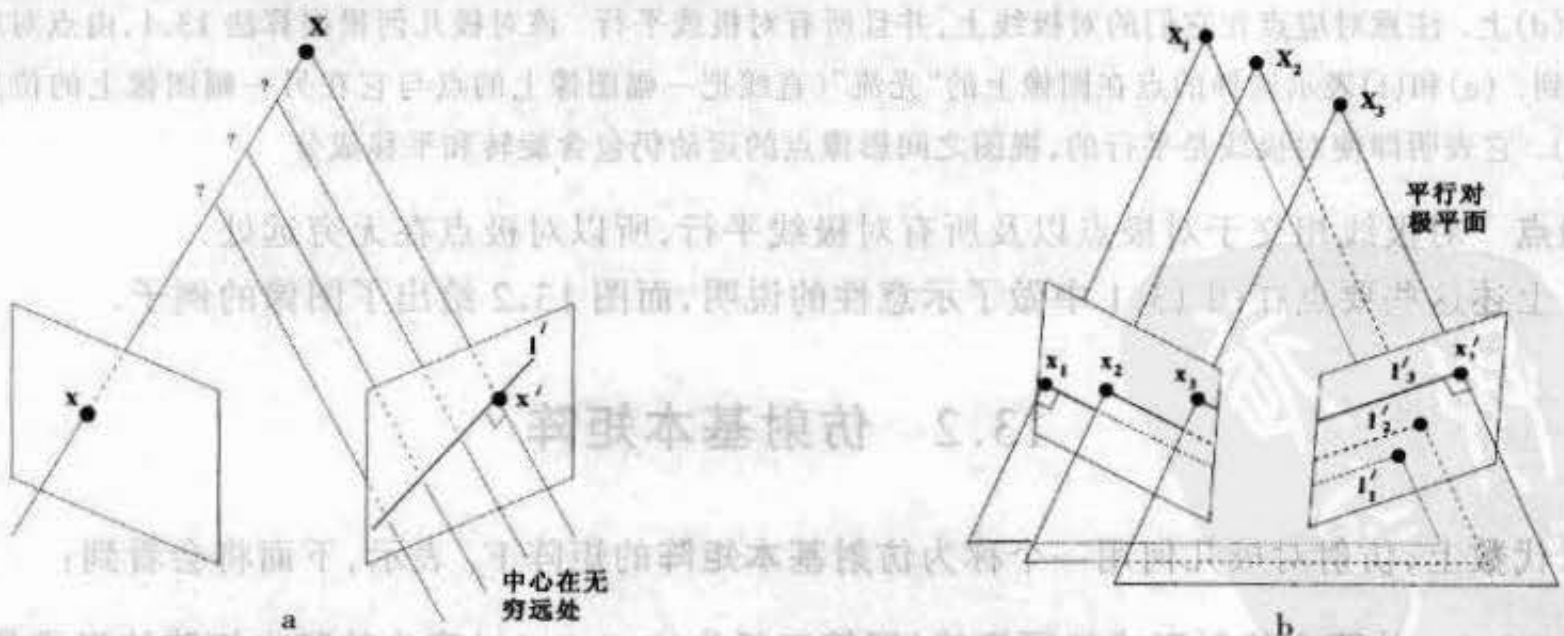



图 13.1　**仿射对极几何**．(a)对应几何：投影射线是平行的并相交于无穷远点．点 $\mathbf{x}$ 反向投影成 3 维空间中由第一个摄像机中心（在无穷远）与 $\mathbf{x}$ 定义的一条射线．这条射线在第二视图上影像成一条直线 $\mathbf{l}'$．投影到 $\mathbf{x}$ 的 3 维空间点 $\mathbf{X}$ 在此射线上，所以 $\mathbf{X}$ 在第二视图上的影像在 $\mathbf{l}'$ 上．(b)对极线和对极平面都是相互平行的．

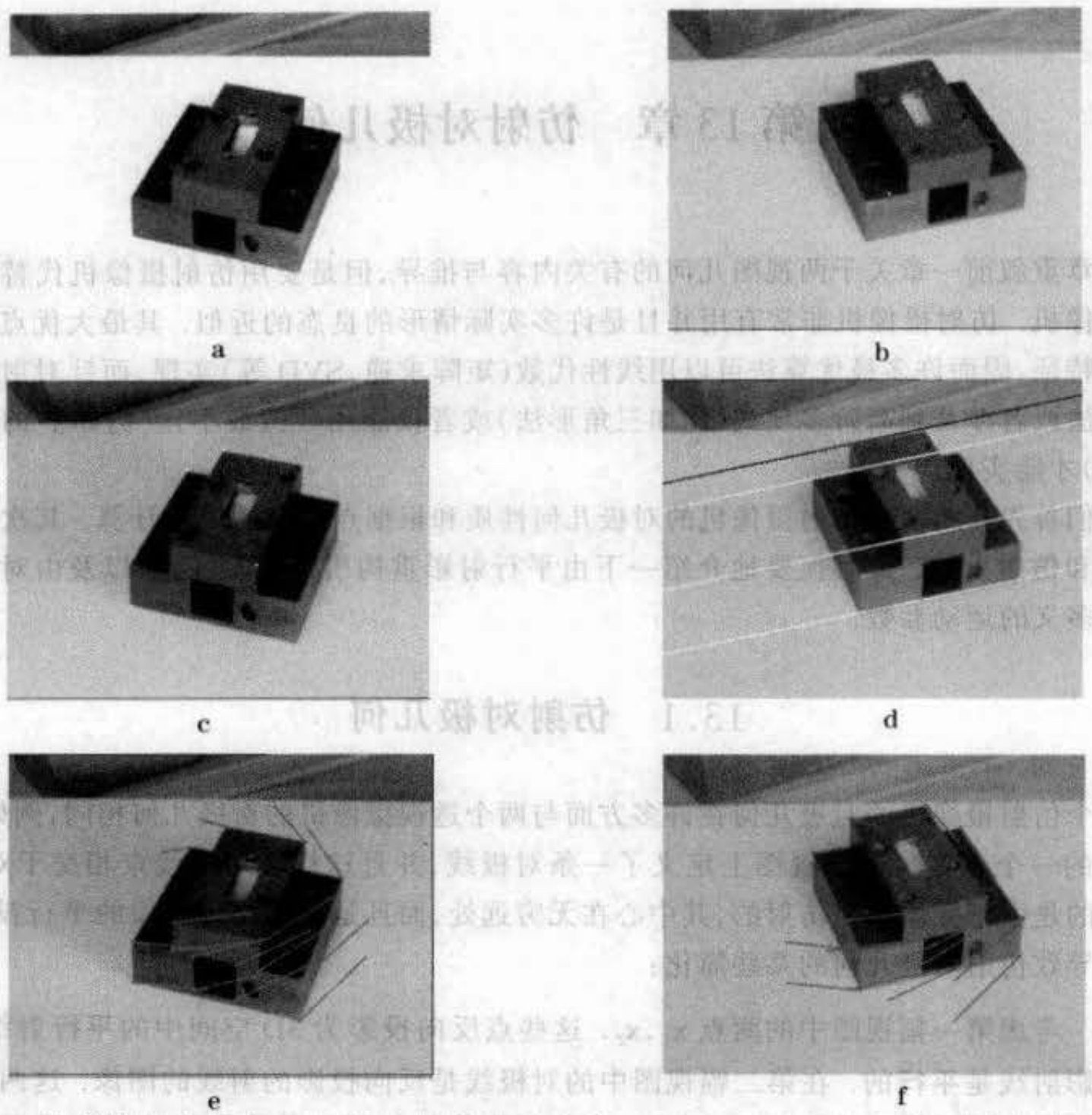



图 13.2 **仿射对极线**. (a),(b)打孔机在仿射影像条件下的两幅视图. 对应(c)中所标记的点的对极线叠加在(d)上. 注意对应点在它们的对极线上,并且所有对极线平行. 该对极几何根据算法 13.1,由点对应计算得到. (e)和(f)表示选择的点在图像上的"光流"(直线把一幅图像上的点与它在另一幅图像上的位置连起来). 它表明即使对极线是平行的,视图之间影像点的运动仍包含旋转和平移成分.

**对极点** 对极线相交于对极点以及所有对极线平行,所以对极点在无穷远处.

上述这些要点在图 13.1 中做了示意性的说明,而图 13.2 给出了图像的例子.

## 13.2 仿射基本矩阵

代数上,仿射对极几何用一个称为**仿射基本矩阵**的矩阵 $F_A$ 表示,下面将会看到:

**结论 13.1 从两个仿射形式的摄像机(即第三行为(0,0,0,1))产生的基本矩阵的形式是**

$$F_A = \begin{bmatrix} 0 & 0 & * \\ 0 & 0 & * \\ * & * & * \end{bmatrix}$$

**其中 * 表示非零元素.**

为了方便起见把五个非零元素记为

$$F_A = \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{bmatrix}. \tag{13.1}$$

注意 $F_A$ 的秩一般为 2

### 13.2.1 推导

**几何推导** 这个推导类似于节 8.2.1($p$162)中关于一对射影摄像机的推导. 如图 13.5($p$250)所示,从一幅图像上的点到另一幅图像对应的对极线的映射可以分解为两步.

(1) **点通过一个平面 $\pi$ 转移** 因为两个摄像机都是仿射的,点在图像和景物平面之间通过平行投影来映射. 因此 $\pi$ 和图像之间的映射是一个平面仿射变换;第一视图与 $\pi$ 以及 $\pi$ 与第二视图之间的仿射变换的合成也是一个仿射变换,即 $\mathbf{x}' = H_A\mathbf{x}$.

(2) **构作对极线** 对极线由过 $\mathbf{x}'$ 和对极点 $\mathbf{e}'$ 的直线得到,即 $\mathbf{l}' = \mathbf{e}' \times H_A\mathbf{x} = F_A\mathbf{x}$,因此 $F_A = [\mathbf{e}']_\times H_A$.

现在利用仿射矩阵 $H_A$ 的特殊形式以及当 $\mathbf{e}'$ 在无穷远处(因此最后一个元素为 0)时是反对称矩阵 $[\mathbf{e}']_\times$ 的特殊形式推出:

$$F_A = [\mathbf{e}']_\times H_A = \begin{bmatrix} 0 & 0 & * \\ 0 & 0 & * \\ * & * & 0 \end{bmatrix} \begin{bmatrix} * & * & * \\ * & * & * \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & * \\ 0 & 0 & * \\ * & * & * \end{bmatrix}. \tag{13.2}$$

其中 * 表示非零元素. 这里仅用了摄像机中心在无穷远平面的几何性质推导出了 F 的仿射形式.

**代数推导** 当两个摄像机都是仿射摄像机时,基本矩阵的仿射形式可以直接从 F 的表达式(8.1 - $p$163)得到,即用伪逆表示 $F = [\mathbf{e}']_\times P'P^+$,其中 $\mathbf{e}' = P'\mathbf{C}$,而 $\mathbf{C}$ 是摄像机中心且是 P 的零矢量. 证明细节留作练习. 节 16.1.2($p$293)用仿射摄像机矩阵的行矢量所形成的行列式给出了 $F_A$ 的一个巧妙的推导.

### 13.2.2 性质

仿射基本矩阵是有五个非零元素的齐次矩阵,因此有 4 个自由度. 具体计算如下:两个对极点各一个自由度(对极点在 $\mathbf{l}_\infty$ 上,故只需要指定它们的方向);加上从一幅视图的对极线束映到另一幅的对极线束的 1D 仿射变换要两个自由度.

几何元素(对极点等)在 $F_A$ 中的编码与在 F 中的编码方式一样. 然而,通常这些表达式非常简单,因此能显式地给出.

**对极点** 第一幅视图的对极点是 $F_A$ 的右零矢量,即 $F_A\mathbf{e} = \mathbf{0}$. 由此确定 $\mathbf{e} = (-d, c, 0)^\mathsf{T}$,它是 $\mathbf{l}_\infty$ 上的一个点(方向). 因为所有的对极线都相交于此对极点,这证明了所有对极线是平行的.

**对极线** 在第二幅视图上,对应于第一幅视图的 $\mathbf{x}$ 的对极线是 $\mathbf{l}' = F_A\mathbf{x} = (a, b, cx + dy + e)^\mathsf{T}$. 它再次证明所有对极线相互平行,因为直线的定向是 $(a, b)^\mathsf{T}$,与 $(x, y)^\mathsf{T}$ 无关.

上述性质以及其他的一些性质概括在表 13.1 中.

## 13.3 由图像点对应估计 $F_A$

对于两幅图像上的任何一对匹配点 $\mathbf{x}\leftrightarrow\mathbf{x}'$,基本矩阵由方程 $\mathbf{x}'^{\mathsf{T}}F_A\mathbf{x}=0$ 定义. 给定足够多的点匹配 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i$,这个方程可以用来计算未知的矩阵 $F_A$. 具体地说,记 $\mathbf{x}_i=(x_i,y_i,1)^{\mathsf{T}}$ 和 $\mathbf{x}'_i=(x'_i,y'_i,1)^{\mathsf{T}}$,每一组点匹配生成关于 $F_A$ 的未知元素 $\{a,b,c,d,e\}$ 的一个线性方程

$$ax'_i+by'_i+cx_i+dy+e=0. \tag{13.3}$$

### 13.3.1 线性算法

按通常的方式,可以把(13.3)重写为:

$$(x'_i,y'_i,x_i,y_i,1)\mathbf{f}=0$$

其中 $\mathbf{f}=(a,b,c,d,e)^{\mathsf{T}}$,由此可求得到 $F_A$ 的解. 从 $n$ 组点匹配,我们得到形如 $A\mathbf{f}=\mathbf{0}$ 的线性方程组,其中 A 是 $n\times5$ 的矩阵:

$$\begin{bmatrix} x'_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n & y'_n & x_n & y_n & 1 \end{bmatrix}\mathbf{f}=\mathbf{0}.$$

当有 $n=4$ 组点对应时得到一个最小配置解,它是 $4\times5$ 矩阵 A 的右零空间. 这样一来,如果这些 3D 空间点处于一般位置上,那么 $F_A$ 可以仅由 4 组对应点唯一地确定. 一般位置的条件在下面的节 13.3.3 中介绍.

如果超过 4 组对应且数据不准确,那么 A 的秩可能超过 4. 在这种情形,可以用与节 3.1.1($p$54)基本相同的方式求一个满足约束 $\|\mathbf{f}\|=1$ 的最小二乘解,即求对应于 A 的最小奇异值的奇异矢量. 其细节请参考算法 3.2($p$68). 这种线性解法等价于计算一般基本矩阵的 8 点算法 10.1($p$193). 我们不推荐用这个方法来估计 $F_A$,因为下面介绍的黄金标准算法在计算实现上不比它难但一般比它性能更优.

**奇异性约束** $F_A$ 的形式(13.1)保证了该矩阵的秩不超过 2. 因此,如果用上述的线性方法估计 $F_A$ 就不再需要强加奇异性约束. 这一点比用线性 8 点算法求一般 F 的估计优越,对于后者被估计的矩阵不保证秩为 2,因而必须接着进行修正.

**几何解释** 在本书的好几处我们都看到,由点对应计算两视图关系等价于 $\mathbb{R}^4$ 中求拟合点 $x,y,x',y'$ 的一个曲面(簇). 对于方程 $\mathbf{x}'^{\mathsf{T}}F_A\mathbf{x}=0$ 的情形,该关系 $ax'_i+by'_i+cx_i+dy_i+e=0$关于坐标是**线性的**,因此由仿射基本矩阵定义的簇 $\mathcal{V}_{F_A}$ 是一个超平面.

由此得到两个简化:第一,求 $F_A$ 的最优估计可以形式化为(熟悉的)平面拟合问题;第二,Sampson 误差等于几何误差,而在一般(非仿射)基本矩阵情形(10.9 - $p$197),它只是一个一阶近似. 如节 3.2.6($p$61)所讨论的那样,后一个性质一般只发生于仿射(线性)关系,因为 Sampson 近似的切平面等于该曲面.

- $F_A$ 是秩 2 和自由度 4 的齐次矩阵，它的形式为

$$F_A = \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{bmatrix}.$$

- **点对应** 如果 $\mathbf{x}$ 和 $\mathbf{x}'$ 是仿射摄像机下的对应图像点，那么 $\mathbf{x}'^{\mathsf{T}}F_A\mathbf{x}=0$. 对于有限点有

$$ax' + by' + cx + dy + e = 0$$

- **对极线**
  - ◇ $\mathbf{l}' = F_A\mathbf{x} = (a, b, cx + dy + e)^{\mathsf{T}}$ 是对应于 $\mathbf{x}$ 的对极线.
  - ◇ $\mathbf{l} = F_A^{\mathsf{T}}\mathbf{x}' = (c, d, ax + by + e)^{\mathsf{T}}$ 是对应于 $\mathbf{x}'$ 的对极线.
- **对极点**
  - ◇ 由 $F_A\mathbf{e} = \mathbf{0}$ 求得 $\mathbf{e} = (-d, c, \mathbf{0})^{\mathsf{T}}$.
  - ◇ 由 $F_A^{\mathsf{T}}\mathbf{e}' = \mathbf{0}$ 求得 $\mathbf{e}' = (-b, a, 0)^{\mathsf{T}}$.
- **根据摄像机矩阵 $P_A, P'_A$ 计算：**
  - ◇ 一般摄像机

    $F_A = [\mathbf{e}']_\times P'_A P_A^+$，其中 $P_A^+$ 是 $P_A$ 的伪逆，而 $\mathbf{e}'$ 是由 $\mathbf{e}' = P'_A\mathbf{C}$ 所确定的对极点，其中 $\mathbf{C}$ 是第一摄像机的中心.
  - ◇ 标准摄像机，

$$P_A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad P'_A = \left[\begin{array}{ccc|c} & M_{2\times 3} & & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$$a = m_{23}, \quad b = -m_{13}, \quad c = m_{13}m_{21} - m_{11}m_{23}$$

$$d = m_{13}m_{22} - m_{12}m_{23}, \quad e = m_{13}t_2 - m_{23}t_1$$

表 13.1 仿射基本矩阵性质摘要.

### 13.3.2 黄金标准算法

给定 $n$ 组图像点对应 $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$ 的一个集合，我们要在图像测量噪声服从各向同性齐次高斯分布的假设下求 $F_A$ 的最大似然估计. 这个估计通过最小化几何图像距离的代价函数：

$$\min_{\{F_A, \hat{\mathbf{x}}_i, \hat{\mathbf{x}}'_i\}} \sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 \tag{13.4}$$

而得到，其中 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ 是测量得到的对应；而 $\hat{\mathbf{x}}_i$ 和 $\hat{\mathbf{x}}'_i$ 是被估计的“真实”对应，对于被估计的仿射基本矩阵，它们精确地满足 $\hat{\mathbf{x}}'^{\mathsf{T}}_i F_A \hat{\mathbf{x}}_i = 0$. 该几何距离在图 13.3 中说明. 这里的真实对应也是必须加以估计的辅助变量.

正如上面以及节 3.2.5($p59$)所讨论的，最小化代价函数(13.4)等价于用一张超平面拟合 $\mathbb{R}^4$ 中一组点 $\mathbf{X}_i = (x'_i, y'_i, x_i, y_i)^{\mathsf{T}}$. 估计的点 $\hat{\mathbf{X}}_i = (\hat{\mathbf{x}}'_i, \hat{\mathbf{y}}'_i, \hat{\mathbf{x}}'_i, \hat{\mathbf{y}}'_i)^{\mathsf{T}}$ 满足方程 $\hat{\mathbf{x}}'^{\mathsf{T}}_i F_A \hat{\mathbf{x}}_i = 0$，它可记为 $(\hat{\mathbf{X}}_i^{\mathsf{T}}, 1)\mathbf{f} = 0$，其中 $\mathbf{f} = (a, b, c, d, e)^{\mathsf{T}}$. 这是 $\mathbb{R}^4$ 中点在平面 $\mathbf{f}$ 上的方程. 我们要求平面 $\mathbf{f}$ 最小化到测量点和估计点之间距离平方，也就是最小化到点 $\mathbf{X}_i = (x'_i, y'_i, x_i, y_i)^{\mathsf{T}}$ 的垂直距离的平方和.

从几何上看，这个解非常简单，图 13.4 用 2D 中的直线拟合作了模拟性说明. 点

$\mathbf{X}_i=(x_i,y_i,x_i',y_i')^{\mathrm{T}}$到平面 $\mathbf{f}$ 的垂直距离为

$$d_{\perp}(\mathbf{X}_i,\mathbf{f})=\frac{ax_i'+by_i'+cx_i+dy_i+e}{\sqrt{a^2+b^2+c^2+d^2}}.$$

则最小化(13.4)的矩阵 $\mathrm{F}_{\mathrm{A}}$ 由在 $\mathbf{f}$ 的 5 个参数$\{a,b,c,d,e\}$上最小化下列代价函数得到

$$\mathcal{C}=\sum_i d_{\perp}(\mathbf{X}_i,\mathbf{f})^2=\frac{1}{a^2+b^2+c^2+d^2}\sum_i(ax_i'+by_i'+cx_i+dy_i+e)^2. \tag{13.5}$$

把该超平面的法矢量记为 $\mathbf{N}=(a,b,c,d)^{\mathrm{T}}$,则

$$\mathcal{C}=\frac{1}{\|\mathbf{N}\|^2}\sum_i(\mathbf{N}^{\mathrm{T}}\mathbf{X}_i+e)^2.$$

用一个经典的平面正交回归问题的简单线性算法就能使这个代价函数最小化. 该算法有两个步骤:

第一步在参数 $e$ 上最小化 $\mathcal{C}$. 我们得到

$$\frac{\partial\mathcal{C}}{\partial e}=\frac{1}{\|\mathbf{N}\|^2}\sum_i 2(\mathbf{N}^{\mathrm{T}}\mathbf{X}_i+e)=0$$

从而

$$e=-\frac{1}{n}\sum_i(\mathbf{N}^{\mathrm{T}}\mathbf{X}_i)=-\mathbf{N}^{\mathrm{T}}\bar{\mathbf{X}}$$

因此这个超平面解通过数据的形心 $\bar{\mathbf{X}}$. 把 $e$ 的上述表达式代入代价函数中,使 $\mathcal{C}$ 化成

$$\mathcal{C}=\frac{1}{\|\mathbf{N}\|^2}\sum_i(\mathbf{N}^{\mathrm{T}}\Delta\mathbf{X}_i)^2$$

其中 $\Delta\mathbf{X}_i=\mathbf{X}_i-\bar{\mathbf{X}}$ 是 $\mathbf{X}_i$ 相对于数据形心 $\bar{\mathbf{X}}$ 的矢量.

第二步是在 $\mathbf{N}$ 上最小化这个简化的代价函数,令 A 表示行矢量为 $\Delta\mathbf{X}_i^{\mathrm{T}}$ 的矩阵,显然有

$$\mathcal{C}=\sum_i\|\mathrm{A}\mathbf{N}\|^2/(\|\mathbf{N}\|^2).$$

最小化这个表达式等价于在约束 $\|\mathbf{N}\|=1$ 下最小化 $\|\mathrm{A}\mathbf{N}\|$,这是通常的由 SVD 求解的齐次最小化问题. 这些步骤概括在算法 13.1 中.

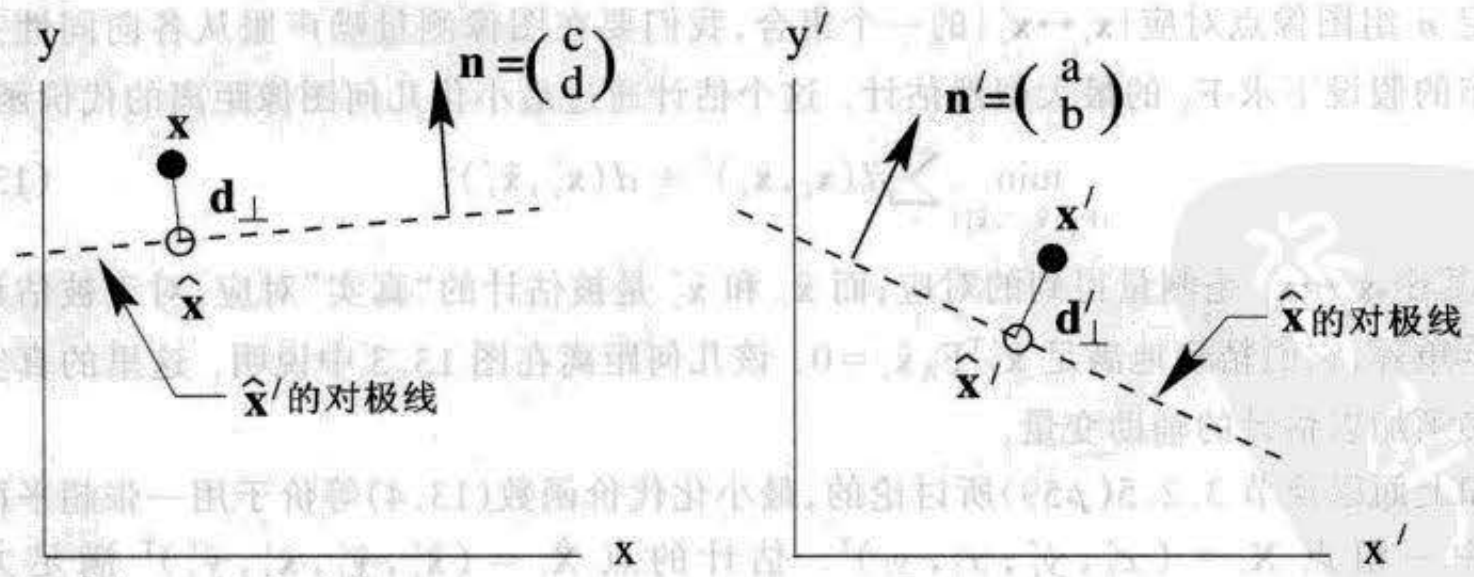



图 13.3 由测量得到的对应点集合$\{\mathbf{x}_i\leftrightarrow\mathbf{x}_i'\}$估计 $\mathrm{F}_{\mathrm{A}}$ 的 MLE 涉及估计 5 个参数 $a,b,c,d,e$ 及精确满足 $\hat{\mathbf{x}}_i'^{\mathrm{T}}\mathrm{F}_{\mathrm{A}}\hat{\mathbf{x}}=0$ 的一组点对应$\{\hat{\mathbf{x}}\leftrightarrow\hat{\mathbf{x}}_i'\}$. 这个问题有一个线性解.

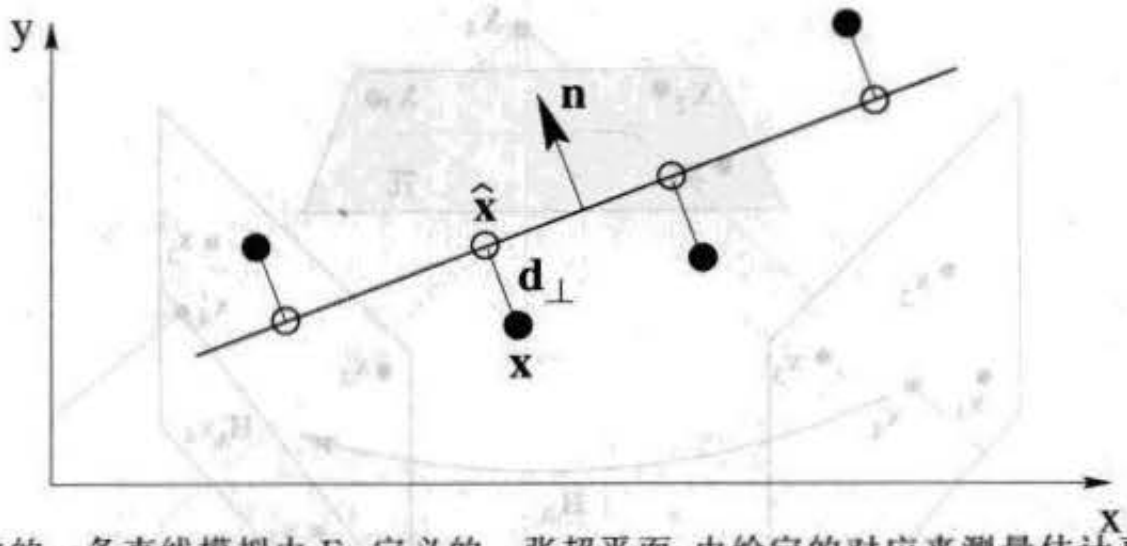



图 13.4　用 2D 中的一条直线模拟由 $F_A$ 定义的一张超平面，由给定的对应来测量估计真对应的问题可用确定直线 $ax+by+c=0$ 上离测量点 $(x,y)^T$ 最近点 $(\hat{x},\hat{y})^T$ 的问题来模拟．该直线的法线方向是 $(a,b)^T$，而点 $(x,y)^T$ 到该直线的垂直距离是 $d_\perp=(ax+by+c)/\sqrt{a^2+b^2}$，因此 $(\hat{x},\hat{y})^T=(x,y)^T-d_\perp\hat{\mathbf{n}}$，其中 $\hat{\mathbf{n}}=(a,b)/\sqrt{a^2+b^2}$．

目标

给定 $n\geq 4$ 组图像点对应 $\{\mathbf{x}_i\leftrightarrow\mathbf{x}_i'\}$，$i=1,\ldots,n$，确定仿射摄像机矩阵的最大似然估计 $F_A$．

算法

把对应表示为 $\mathbf{X}_i=(x_i',y_i',x_i,y_i)^T$．

(1) 计算形心 $\bar{\mathbf{X}}=\frac{1}{n}\sum_i\mathbf{X}_i$，并把各矢量聚中 $\Delta\mathbf{X}_i=\mathbf{X}_i-\bar{\mathbf{X}}_i$．

(2) 计算行矢量是 $\Delta\mathbf{X}_i^T$ 的 $n\times 4$ 矩阵 A．

(3) $\mathbf{N}=(a,b,c,d)^T$ 是对应 A 的最小奇异值的奇异矢量，而 $\mathbf{e}=-\mathbf{N}^T\bar{\mathbf{X}}$．矩阵 $F_A$ 的形式是(13.1)．

算法 13.1　由图像对应来估计 $F_A$ 的黄金标准算法．

值得注意的是黄金标准算法提供一个与矩阵分解算法 17.1($p$312)(由 $n$ 组点对应得到的仿射重构)完全相同的 $F_A$ 估计．

### 13.3.3　最小配置

我们回到估计 $F_A$ 的最小配置，即 3D 空间中一般位置的 4 个点的对应图像，这种配置的 $F_A$ 的几何计算方法在算法 13.2 中介绍．该最小配置解在鲁棒性算法(比如 RANSAC)中有用而这里将用于说明退化配置．注意由这个最小配置，能得到 $F_A$ 的一个准确解，并且节 13.3.1 的线性算法黄金标准算法 13.1 和最小算法 13.2 都给出同样的结果．

目标

给定 4 组图像点对应 $\{\mathbf{x}_i\leftrightarrow\mathbf{x}_i'\}$，$i=1,\ldots,4$，计算仿射基本矩阵．

算法

前 3 个 3D 点 $\mathbf{X}_i$，$i=1,\ldots,3$ 定义一个平面 $\boldsymbol{\pi}$．见图 13.5．

(1) 计算仿射变换矩阵 $H_A$，使得 $\mathbf{x}_i'=H_A\mathbf{x}_i$，$i=1,\ldots,3$．

(2) 根据 $\mathbf{l}'=(H_A\mathbf{x}_4)\times\mathbf{x}_4'$ 确定在第二视图上的对极线．由此得对极点 $\mathbf{e}'=(-l_2',l_1',0)^T$．

(3) 任意点 $\mathbf{x}$ 在第二视图上的对极线是 $\mathbf{e}'\times(H_A\mathbf{x})=F_A\mathbf{x}$．因此 $F_A=[(-l_2',l_1',0)^T]_\times H_A$．

算法 13.2　按 4 组点对应的最小配置计算 $F_A$．

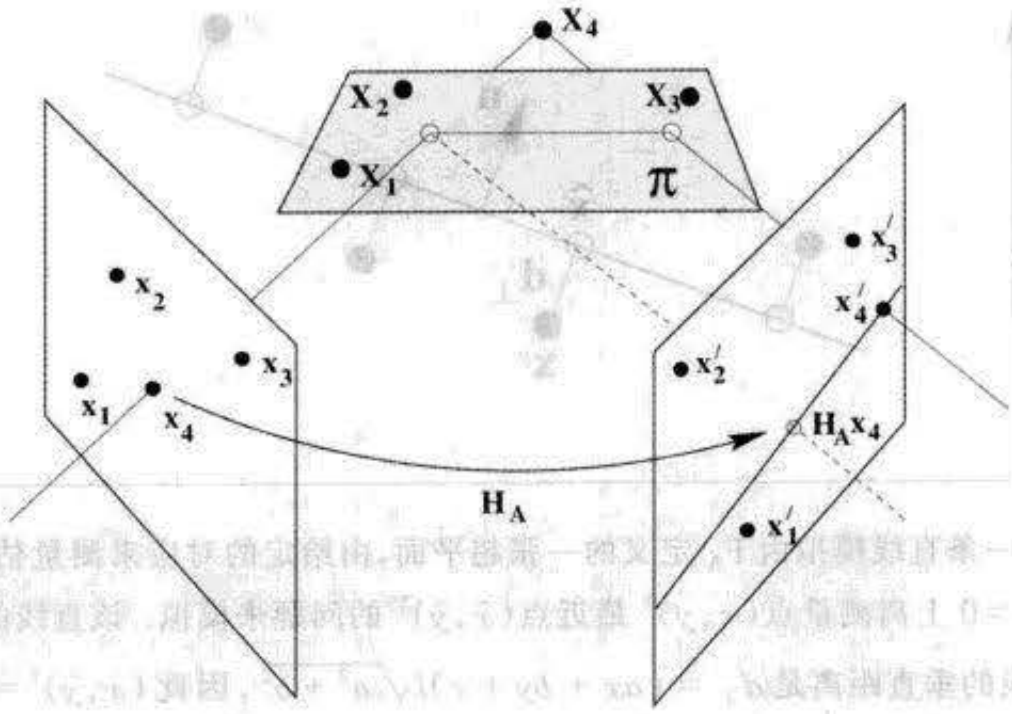



图 13.5 **由 4 点的一个最小配置计算仿射对极线**. 该对极线由平面 $\pi$ 所诱导的虚拟视差来计算. 请与图 12.7($p$236)相比较.

**一般位置** 图 13.5 给出的四点配置说明了给定的 3D 点位于一般位置是计算 $F_A$ 的必要条件. 不能算出 $F_A$ 的配置称为退化的,它们分成两类:第一类,退化配置只依赖于结构,例如如果四点是共面的(因此没有视差),或者如果前三个点是共线的(使得 $H_A$ 不能被算出);第二类,只依赖于摄像机的退化,例如如果两个摄像机有相同的摄像方向(因此在无穷远平面上有共同的中心).

这里再次强调视差的重要性——在图 13.5 中,当点 $\mathbf{X}_4$ 趋近于由另外三点定义的平面时,定义对极线方向的视差矢量的长度单调地减少. 其结果是对极线方向的准确性也相应地降低. 上述结论对最小配置的黄金标准算法 13.1 也适用;当起伏减少到零时,即当点集合接近于共面时,被估计的 $F_A$ 的方差将增加.

## 13.4 三角形法

假设我们有一组测量对应 $(x,y)^\top \leftrightarrow (x',y')^\top$ 和仿射基本矩阵 $F_A$. 我们希望在通常的假设(即图像测量的误差满足高斯分布)下,确定真对应 $(\hat{x},\hat{y})^\top \leftrightarrow (\hat{x}',\hat{y}')^\top$ 的最大似然估计. 该 3D 点可以从 ML 估计的对应中确定.

我们已在第 11 章了解到,MLE 可确定精确地满足仿射对极几何,即 $(\hat{x},\hat{y},1)F_A(\hat{x}',\hat{y}',1)^\top=0$,并且也最小化到测量点的图像距离

$$(x-\hat{x})^2+(y-\hat{y})^2+(x'-\hat{x}')^2+(y'-\hat{y}')^2$$

的"真"对应. 从几何上看,这个解非常简单,图 13.4 对 2D 情形作了说明. 我们在由 $F_A$ 定义的超平面上寻找与 $\mathbb{R}^4$ 中的测量对应 $\mathbf{X}=(x,y,x',y')^\top$ 最接近的点. 同样,此时 Sampson 矫正(3.11—$p$61)是精确的. 从代数上看,超平面的法线方向是 $\mathbf{N}=(a,b,c,d)^\top$,并且点 $\mathbf{X}$ 到该超平面的垂直距离由 $d_\perp=(\mathbf{N}^\top\mathbf{X}+e)/\|\mathbf{N}\|$ 给出,因此,

$$\hat{\mathbf{X}}=\mathbf{X}-d_\perp\frac{\mathbf{N}}{\|\mathbf{N}\|}.$$

或者详细地写成

$$\begin{bmatrix}\hat{x}'\\ \hat{y}'\\ \hat{x}\\ \hat{y}\end{bmatrix}=\begin{bmatrix}x'\\ y'\\ x\\ y\end{bmatrix}-\frac{(ax'+by'+cx+dy+e)}{(a^2+b^2+c^2+d^2)}\begin{bmatrix}a\\ b\\ c\\ d\end{bmatrix}$$

## 13.5 仿射重构

假设我们有 $n\geq 4$ 组图像点对应 $\mathbf{x}_i\leftrightarrow\mathbf{x}'_i$，$i=0,\ldots,n-1$，并假设没有噪声，那么我们可以计算 3D 点和摄像机的重构．射影摄像机的(具有 $n\geq 7$ 个点)重构是射影的．毫不奇怪，在仿射情形，重构是**仿射**的．现在我们给出这个结论的一个简单的构造性推导．

3D 空间仿射坐标系可以由四个有限的非共面的基点 $\mathbf{X}_i$，$i=0,\ldots,3$．确定，如图 13.6 所示，其中一个点 $\mathbf{X}_0$ 取为原点，而其他三个点定义基矢量 $\widetilde{\mathbf{E}}_i=\widetilde{\mathbf{X}}_i-\widetilde{\mathbf{X}}_0$，$i=1,\ldots,3$，其中 $\widetilde{\mathbf{X}}_i$ 是对应于 $\mathbf{X}_i$ 的非齐次 3 维矢量．一个点 $\mathbf{X}$ 的位置于是可以通过简单矢量加法定义为

$$\widetilde{\mathbf{X}}=\widetilde{\mathbf{X}}_0+\mathrm{X}\widetilde{\mathbf{E}}_1+\mathrm{Y}\widetilde{\mathbf{E}}_2+\mathrm{Z}\widetilde{\mathbf{E}}_3$$

并且$(\mathrm{X},\mathrm{Y},\mathrm{Z})^\mathsf{T}$ 是 $\widetilde{\mathbf{X}}$ 对应于这个基的仿射坐标．也就是说该基点 $\mathbf{X}_i^\mathsf{T}$ 的标准坐标$(\mathrm{X},\mathrm{Y},\mathrm{Z})^\mathsf{T}$ 是

$$\widetilde{\mathbf{X}}_0=\begin{bmatrix}0\\0\\0\end{bmatrix}\quad\widetilde{\mathbf{X}}_1=\begin{bmatrix}1\\0\\0\end{bmatrix}\quad\widetilde{\mathbf{X}}_2=\begin{bmatrix}0\\1\\0\end{bmatrix}\quad\widetilde{\mathbf{X}}_3=\begin{bmatrix}0\\0\\1\end{bmatrix}.\tag{13.6}$$

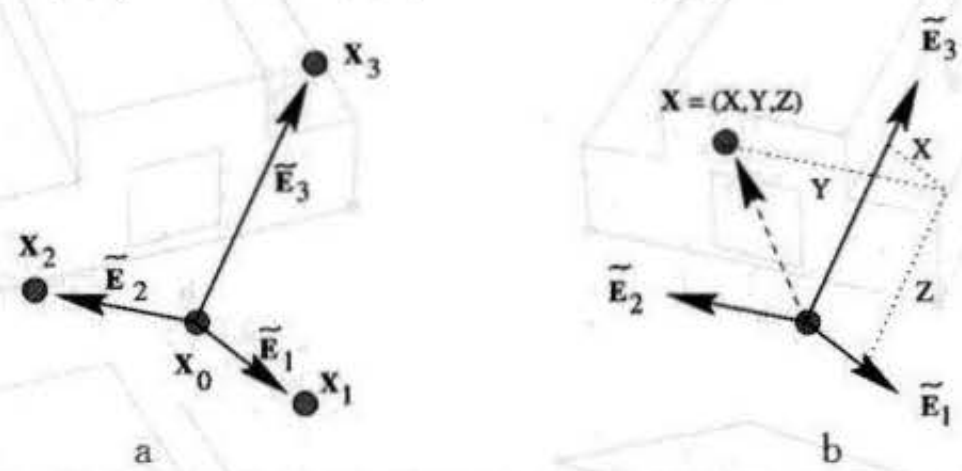



图 13.6 **仿射坐标**．(a)3 维空间的 4 个非共面的点($\mathbf{X}_1$，$\mathbf{X}_2$，$\mathbf{X}_3$ 以及原点 $\mathbf{X}_0$)定义一组坐标轴，使得其它的点 $\mathbf{X}$ 可以按这些轴来指定相应的仿射坐标$(\mathrm{X},\mathrm{Y},\mathrm{Z})^\mathsf{T}$．(b)每一个仿射坐标由平行方向的长度比率定义(它是仿射不变量)．例如，X 可由下面的两步算出：第一步，把 $\mathbf{X}$ 沿平行于 $\widetilde{\mathbf{E}}_2$ 方向投影到 $\widetilde{\mathbf{E}}_1$ 和 $\widetilde{\mathbf{E}}_3$ 所生成的平面上．第二步，把投影得到的点平行于 $\widetilde{\mathbf{E}}_3$ 方向投影到轴 $\widetilde{\mathbf{E}}_1$ 上．则坐标 X 的值是最后投影得到的点到原点的长度和 $\widetilde{\mathbf{E}}_1$ 的长度的比率．

在两幅视图中给定四个基点的仿射投影，其他任何点的 3D 仿射坐标可以直接地由图像恢复，如下面正要解释的那样(见图 13.7)．

仿射摄像机的投影可以表示为(5.26 − $p$116)．

$$\widetilde{\mathbf{x}}=\mathrm{M}_{2\times 3}\widetilde{\mathbf{X}}+\mathbf{t}$$

其中 $\widetilde{\mathbf{x}}=(x,y)^\mathsf{T}$ 是对应于 $\mathbf{x}$ 的 2 维非齐次矢量．矢量的差可消去 $\mathbf{t}$．例如，基矢量的投影是 $\widetilde{\mathbf{e}}_i=\mathrm{M}_{2\times 3}\widetilde{\mathbf{E}}_i$，$i=1,\ldots,3$．因此对任何点 $\mathbf{X}$，在第一幅视图中的像是

$$\widetilde{\mathbf{x}}-\widetilde{\mathbf{x}}_0=\mathrm{X}\widetilde{\mathbf{e}}_1+\mathrm{Y}\widetilde{\mathbf{e}}_2+\mathrm{Z}\widetilde{\mathbf{e}}_3.\tag{13.7}$$

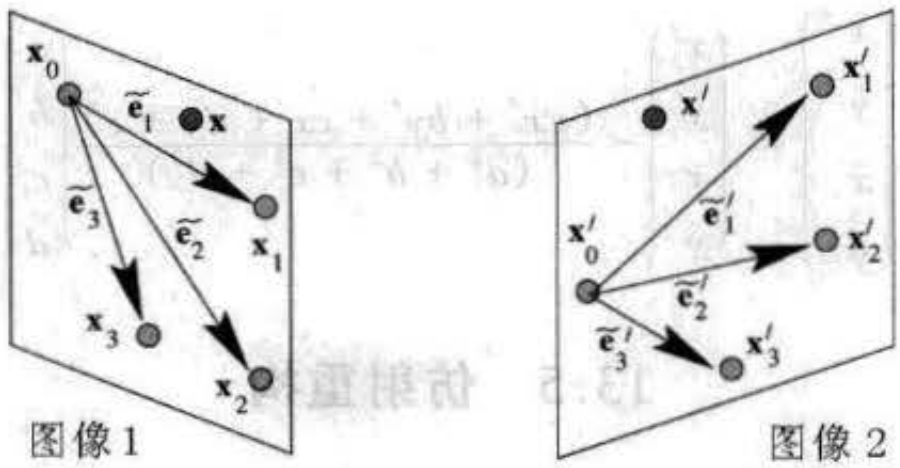



图 13.7 **由两幅图像作重构**. 在两幅视图中的图像分别为 $\mathbf{x},\mathbf{x}'$ 的 3D 点 $\mathbf{X}$ 的仿射坐标可以由图 13.6 中的基点 $\mathbf{x}_i$ 和基矢量 $\tilde{\mathbf{e}}_i$ 的投影线性地算出.

类似地,在第二幅视图中的像($\tilde{\mathbf{x}}' = \mathrm{M}'_{2\times 3}\tilde{\mathbf{X}} + \tilde{\mathbf{t}}'$)是

$$\tilde{\mathbf{x}}' = \tilde{\mathbf{x}}'_0 = \mathrm{X}\tilde{\mathbf{e}}'_1 + \mathrm{Y}\tilde{\mathbf{e}}'_2 + \mathrm{Z}\tilde{\mathbf{e}}'_3. \tag{13.8}$$

每一个方程(13.7)和(13.8)为空间点 $\mathbf{X}$ 的未知仿射坐标 X,Y,Z 提供了两个线性约束. 该方程中的所有其他项由图像测量给定(例如图像的基矢量 $\tilde{\mathbf{e}}_i$, $\tilde{\mathbf{e}}'_i$ 可以从四个基点 $\tilde{\mathbf{X}}_i, i=0,\ldots,3$的投影计算出来). 因此,存在关于三个未知量 X,Y,Z 的四个线性联立方程,并且它的解能直接求出. 这就说明了点 $\mathbf{X}$ 的仿射坐标可以从它在两幅视图中的像计算得到.

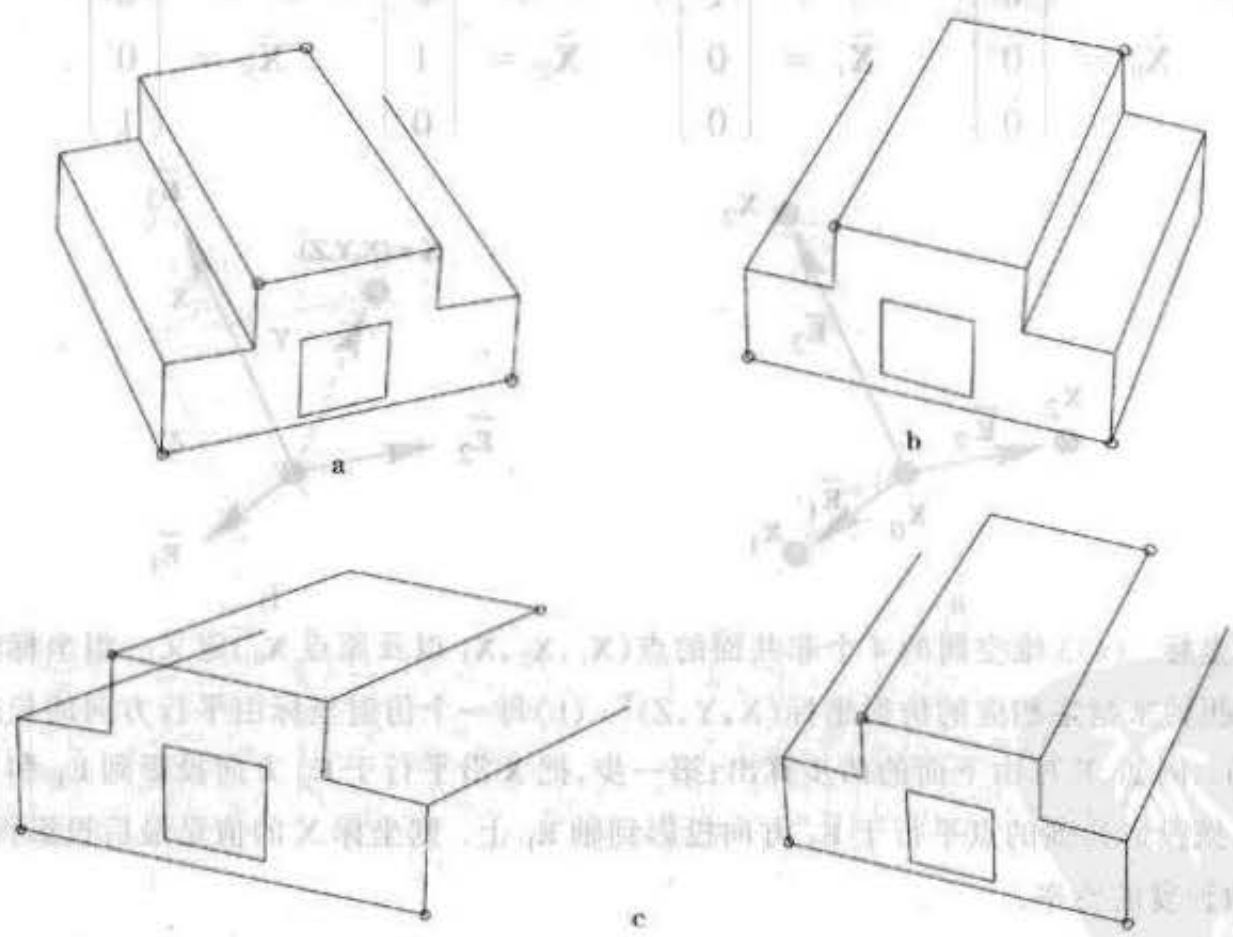



图 13.8 **仿射重构**. (a)(b)由图 13.2($p$244)的两幅图像得到的打孔机的线框轮廓. 圆圈表示选为仿射基的点. 这些连线仅仅为了可视化目的. (c)由线框顶点计算得到的 3D 仿射结构的两幅视图.

两幅视图的摄像机矩阵 $\mathrm{P}_\mathrm{A}, \mathrm{P}'_\mathrm{A}$ 可以从 3 维空间点 $\tilde{\mathbf{X}}_i$(它们的坐标由(13.6)给出)与它们的测量图像之间的对应计算得到. 例如,$\mathrm{P}_\mathrm{A}$ 从对应 $\tilde{\mathbf{x}}_i \leftrightarrow \tilde{\mathbf{X}}_i, i=0,\ldots,3$ 计算得到.

以上的推导不是最优的,因为基点被看作是精确的,且所有的测量误差都集中于第 5 个点 $\mathbf{X}$. 对于仿射情形,在所有点上最小化重构误差的最优重构算法是非常简单的. 然而,它的介绍将推迟到节 17.2($p$310)中给出,因为在那里介绍的分解算法可适用于任意数目的视图.

**例 13.2. 仿射重构**

在图 13.2($p$244)打孔机的图像中通过选择四个图像点作为仿射基,然后由上面的线性算法依次地计算剩下每一个点的仿射坐标,从而得出它的一个 3D 重构. 所得重构的两幅视图在图 13.8 显示. 不过请注意,我们不推荐这种 5 点算法,其实应该采用最优仿射重构算法 17.1($p$312).

## 13.6 Necker 反转和浅浮雕多义性

在前一节中,我们已经看到一个仿射重构可以在没有任何标定信息时仅根据点对应计算得到. 本节将指出即使摄像机的标定已知,在两幅视图的情形仍存在不能解决的重构的多义性簇.

仿射情形与透视投影的情形不同,透视投影中一旦确定了内标定,摄影机的运动只有有限数目的多义性(来自本质矩阵,见节 8.6($p$173)). 对于平行投影还有另外两个重要的多义性:有限反射多义性(Necker 反转);和一个单参数旋转多义性(浅浮雕多义性)簇.

**Necker 反转多义性** 这种多义性的背景是一个物体旋转 $\rho$ 和其镜像旋转 $-\rho$ 在平行射影下产生同样的图像,见图 13.9(a). 因此,结构只能恢复到与前置平面相差一个反射. 这种多义性在透视情形不存在,因为这些点在这两种情形下有不同的深度,因此不投射到相重合的图像点.

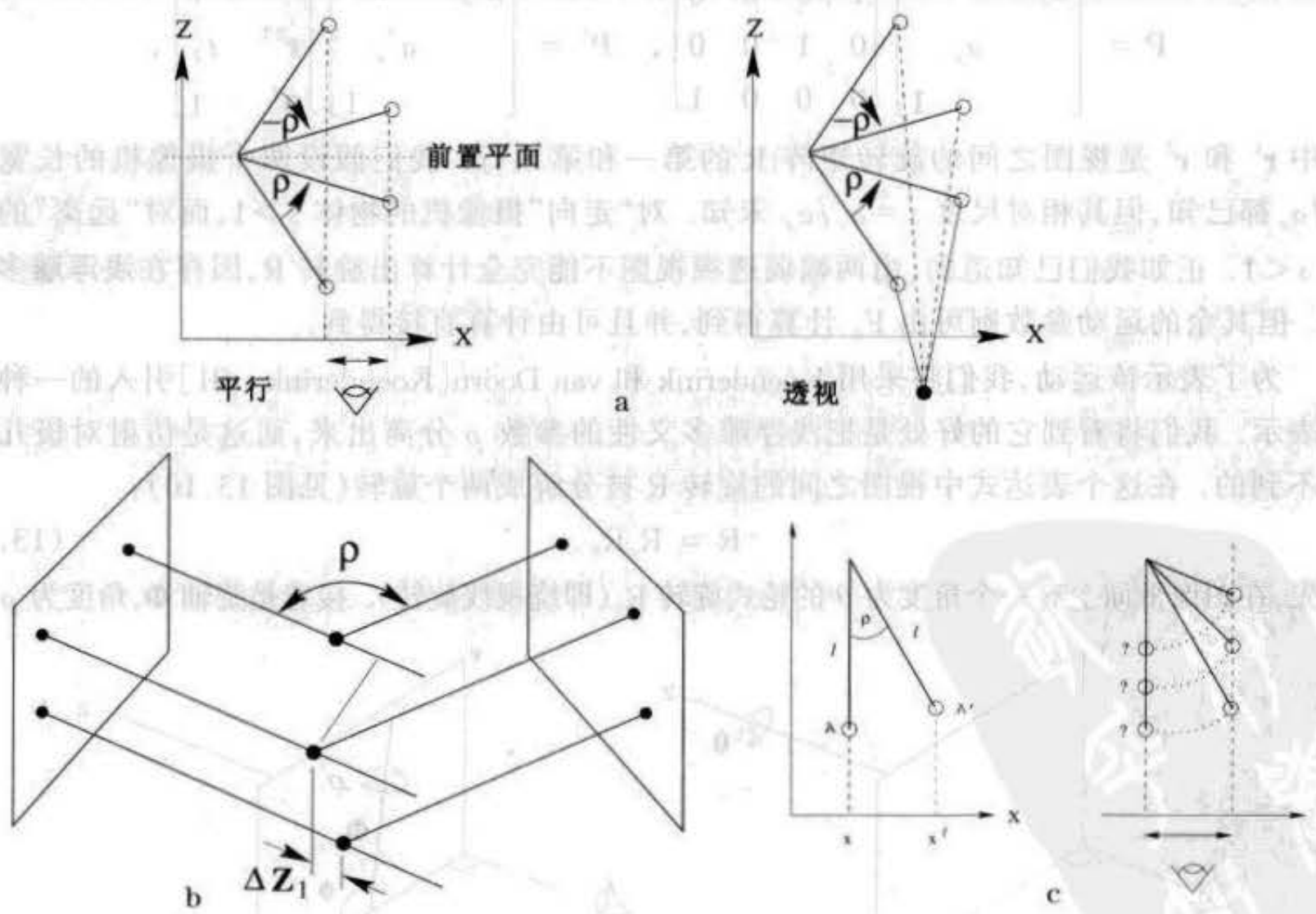



图 13.9 **平行投影下的运动多义性**. (a)Necker 反转:一个旋转的物体产生的图像与它的镜象反方向旋转时产生的图像一样. 但在透视投影下两图像是不同的. (b)摄像机可以旋转($\rho$)而仍然保持射线相交. 透视摄像机则不会发生. (c)浅浮雕多义性:考虑一条长度为 $l$ 并旋转了角度 $\rho$ 的杆子,则 $x - x' = l\sin\rho$. 这种浅浮雕(或深度 - 旋转)多义性之所以这样命名是因为一个浅的物体经历一个大的转动(即小 $l$ 而大 $\rho$)与一个深的物体经历小的转动(即大 $l$ 而小 $\rho$)产生同样的图像.

**浅浮雕多义性** 这种多义性在图 13.9(b)中作了说明. 想象发自一个摄像机的一组平行射线,并调整从第二个摄像机发出的一组平行射线直到每条射线与其对应的射线相交. 这些射线在一簇平行极平面上,而且一个摄像机仍可绕这些平面的法线自由旋转而保持射线相交性. 这种所谓**浅浮雕**(或**深度－转角**)多义性是关于旋转角度和深度的解的单参数簇. 深度参数 $\Delta Z$ 和旋转 $\sin\rho$ 混淆在一起并且不能独立地确定——只能计算它们的乘积. 因此,一个浅的物体经历了一个大的转动(即小 $\Delta Z$ 而大 $\rho$)与一个深的物体经历了小的转动(即大 $\Delta Z$ 而小 $\rho$)产生同样的图像. 该术语来自浅浮雕雕刻. 固定深度或者角度就能唯一地确定结构或者运动. 额外增加点不能解决这个多义性,但是附加一幅视图(即 3 幅视图)一般会解决这个问题.

这种多义性阐明了由两个透视摄像机作重构的稳定性:当影像条件近似于仿射时,旋转角的估计将非常差,但是旋转角和深度的乘积将是稳定的.

## 13.7 计算运动

本节将对两个弱透视摄像机的情形(节 5.3.4($p$115))给出由 $F_A$ 计算摄像机运动的表达式. 这些摄像机矩阵可以取为

$$P = \begin{bmatrix} \alpha_x & & \\ & \alpha_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad P' = \begin{bmatrix} \alpha'_x & & \\ & \alpha'_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}^{1\mathrm{T}} & t_1 \\ \mathbf{r}^{2\mathrm{T}} & t_2 \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix},$$

其中 $\mathbf{r}^1$ 和 $\mathbf{r}^2$ 是视图之间的旋转矩阵 R 的第一和第二行. 我们假设两个摄像机的长宽比 $\alpha_y/\alpha_x$ 都已知,但其相对尺度 $s=\alpha'_x/\alpha_x$ 未知. 对"走向"摄像机的物体 $s>1$,而对"远离"的物体 $s<1$. 正如我们已知道的,由两幅弱透视视图不能完全计算出旋转 R,因存在**浅浮雕**多义性. 但其余的运动参数则可由 $F_A$ 计算得到,并且可由计算直接得到.

为了表示该运动,我们将采用 Koenderink 和 van Doorn[Koenderink－91]引入的一种旋转表示. 我们将看到它的好处是把浅浮雕多义性的参数 $\rho$ 分离出来,而这是仿射对极几何办不到的. 在这个表达式中视图之间的旋转 R 被分解成两个旋转(见图 13.10):

$$R = R_\rho R_\theta. \tag{13.9}$$

首先,在图像平面上有一个角度为 $\theta$ 的轮式旋转 $R_\theta$(即绕视线旋转). 接着是绕轴 $\mathbf{\Phi}$,角度为 $\rho$ 的

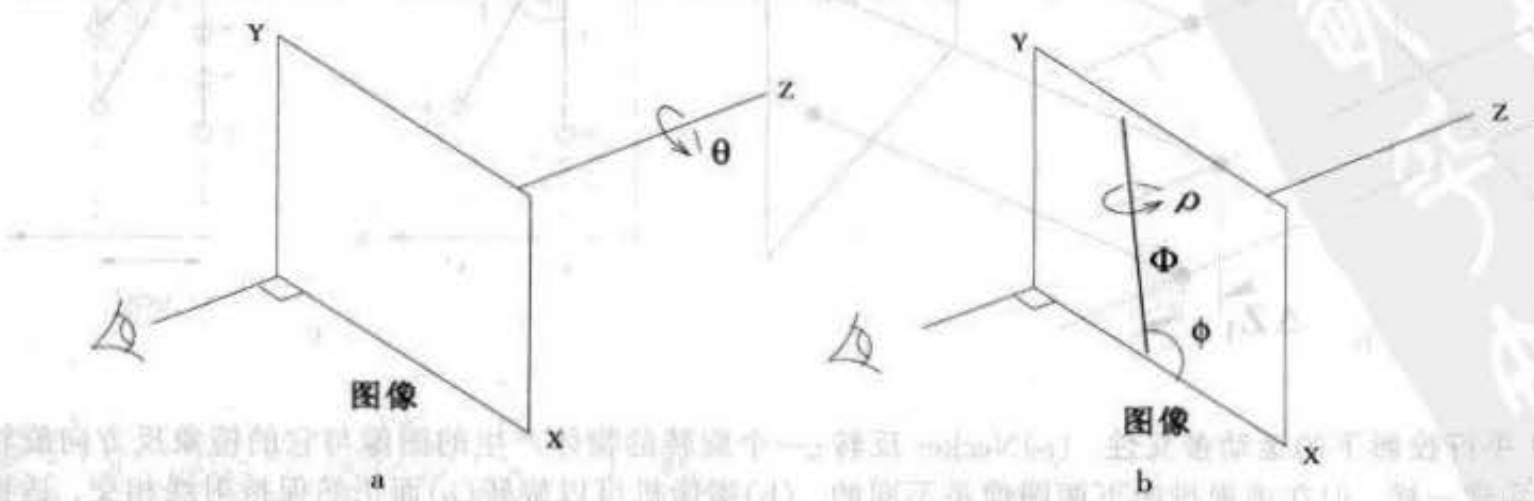



图 13.10 **旋转表示**. ($a$)绕 Z 轴旋转 $\theta$;(b)接着绕轴 $\mathbf{\Phi}$ 旋转 $\rho$,轴 $\mathbf{\Phi}$ 与图像平面平行并与 X 轴的角度是 $\phi$. 轴 $\mathbf{\Phi}$ 的分量是$(\cos\phi,\sin\phi,0)^{\mathrm{T}}$.

旋转 $R_\rho$，$\mathbf{\Phi}$ 的方向平行于图像平面并且与正 X 轴的角度是 $\varnothing$，即一个**转出**图像平面的纯旋转.

**解** $s$，$\varnothing$ 和 $\theta$　我们现在证明尺度因子($s$)，旋转($\varnothing$)的轴的投影以及轮式旋转角($\theta$)都可以直接从仿射极几何计算得到. 在求解之前，我们先从几何上解释对极线是如何与未知运动参数相关联的.

考虑一个绕平行于图像平面的轴 $\mathbf{\Phi}$ 旋转的摄像机(图 13.11(a)). 对极平面 $\boldsymbol{\pi}$ 与此轴和两幅视图都垂直，并与图像交于对极线 $\mathbf{l}$ 和 $\mathbf{l}'$. 因此：

- **旋转轴 $\mathbf{\Phi}$ 的投影垂直于对极线.**

如果图像平面附加一个轮式旋转 $\theta$，上述关系仍成立(图 13.11(b))；轴 $\mathbf{\Phi}$ 和交线 $\mathbf{l}'$ 在空间上保持不变，对极线和轴的投影保持正交，仅仅要在图像上的一个新角度才观察到. 因此在两幅图像上的对极线的方向相差 $\theta$. 重要的是：改变旋转角 $\rho$ 的大小不会以任何方式改变对极几何(图 13.12). 因此，从两幅视图不能确定这个角，此为浅浮雕多义性的推论. 图13.13说明尺度的效应. 设一个 3D 物体被切成一些平行的对极平面，每一平面约束物体的一个特定的切片如何运动. 改变该物体的有效大小(例如向它靠近)仅仅改变了相邻对极平面的相对距离.

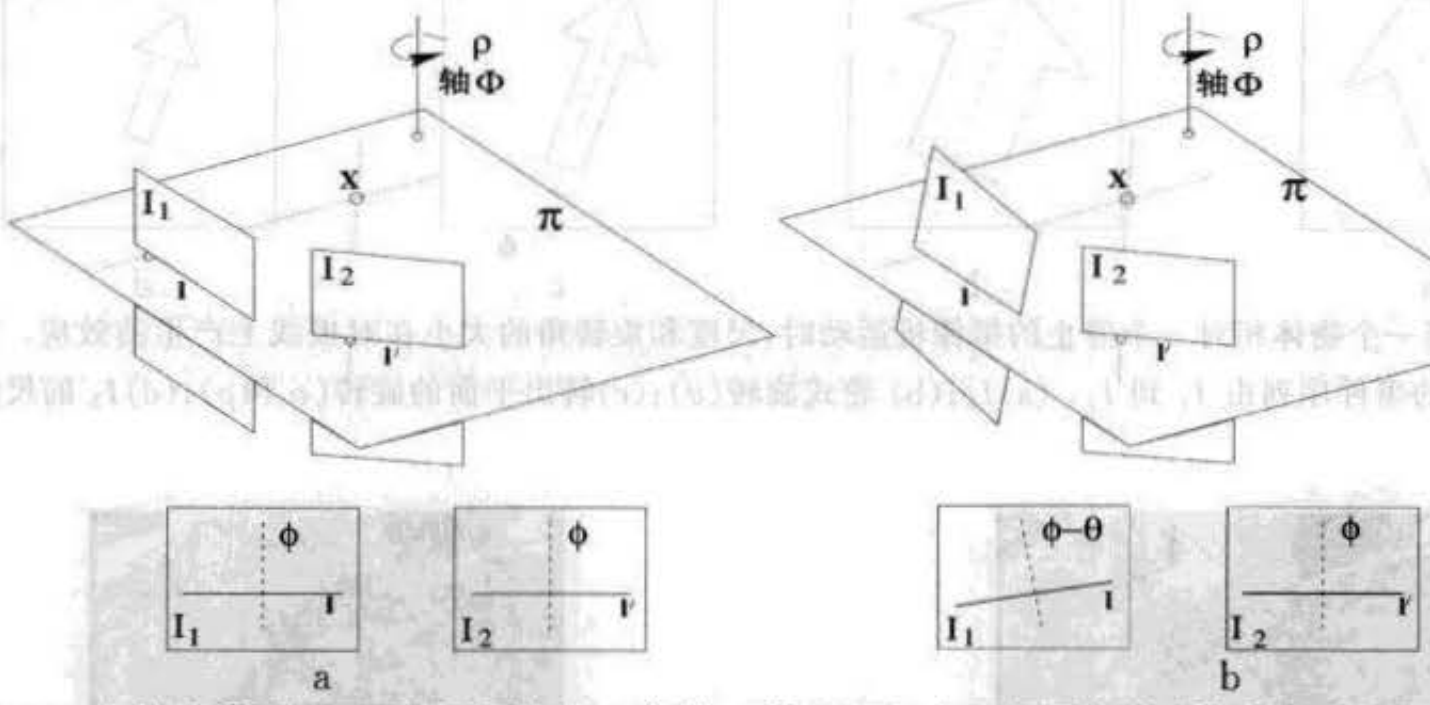



图 13.11　**摄像机绕平行于图像平面的轴 $\mathbf{\Phi}$ 旋转**. 对极平面 $\boldsymbol{\pi}$ 与两幅图像的交线给出对极线 $\mathbf{l}$ 和 $\mathbf{l}'$. 而轴 $\mathbf{\Phi}$ 在该图像的投影与这些对极线正交：(a)没有轮式旋转发生($\theta=0°$)；(b)在 $I_1$ 上摄像机逆时针旋转了 $\theta$ 使得对极线的方向变化了 $\theta$.

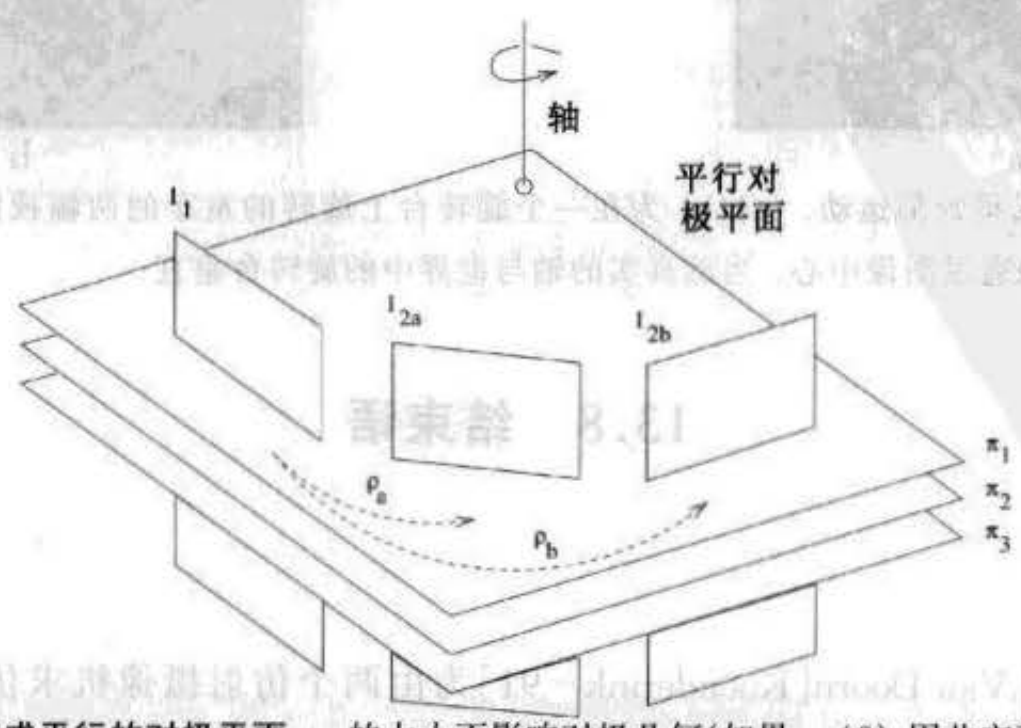



图 13.12　**景物可以切成平行的对极平面**. $\rho$ 的大小不影响对极几何(如果 $\rho\neq0$)，因此它不能由两视图确定.

概括来说，轮式旋转仅仅旋转对极线，转出平面的旋转会造成沿对极线按照远近法缩小（正交于 $\mathbf{\Phi}$），而尺度的变化将一致性地改变对极线的间距（图 13.13）.

可以证明（留着练习）$s$，$\varnothing$ 和 $\theta$ 可以由仿射极几何直接计算如下：

$$\tan\varnothing = \frac{b}{a}, \quad \tan(\varnothing - \theta) = \frac{d}{c} \text{ 和 } s^2 = \frac{c^2 + d^2}{a^2 + b^2}, \tag{13.10}$$

其中 $s>0$（根据定义）. 注意 $\varnothing$ 是转出平面的旋转轴在 $I_2$ 上的投影角，而 $(\varnothing - \theta)$ 是它在 $I_1$ 上的投影角.

**例 13.3. 由仿射基本矩阵计算运动**

图 13.14 给出一个童车在一个旋转台上旋转时的两幅图像. 该图像大小为 256×256 像素而长宽比是 0.65. 仿射基本矩阵用算法 13.1 算得，而运动参数用上面的（13.10）由 $F_A$ 算得. 所算出的旋转轴叠加在该图像上. △

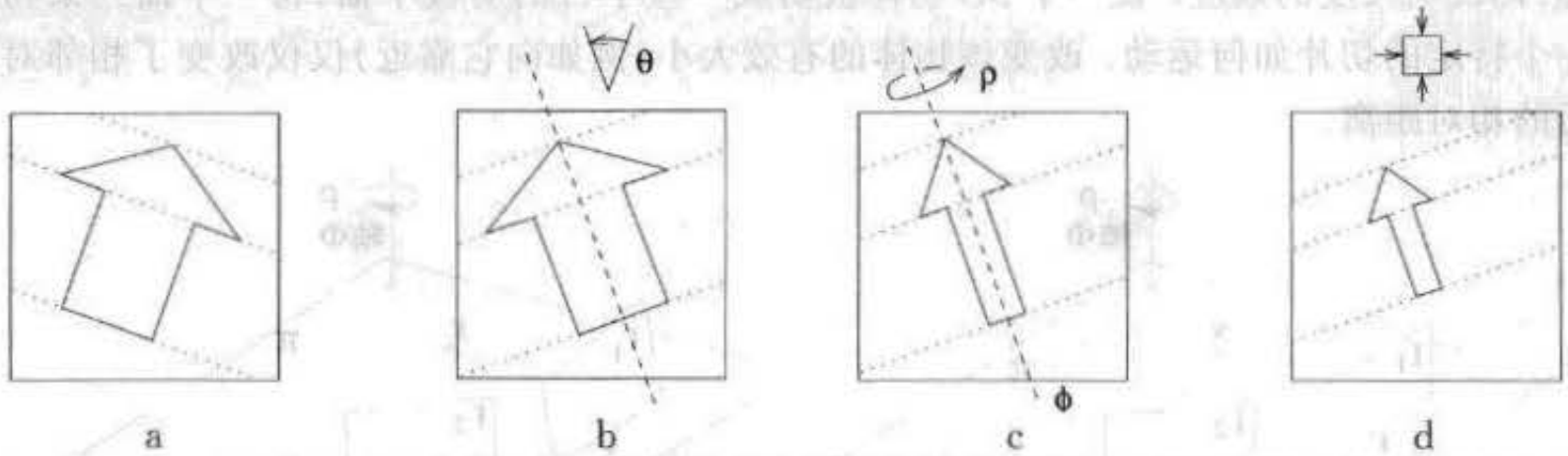



图 13.13 当一个物体相对一个静止的摄像机运动时，尺度和旋转角的大小在对极线上产生的效应. 它同时说明假定的事件序列由 $I_1$ 到 $I_2$. (a) $I_1$；(b) 轮式旋转（$\theta$）；(c) 转出平面的旋转（$\varnothing$ 和 $\rho$）；(d) $I_2$ 的尺度变化.

a b

图 13.14 **由仿射极几何计算运动**. (a)(b) 为在一个旋转台上旋转的童车的两幅视图. 所算出的旋转轴叠加在该图像上，该轴通过图像中心. 当然真实的轴与世界中的旋转台垂直.

## 13.8 结束语

### 13.8.1 文献

Koenderink 和 Van Doorn[Koenderink－91]为由两个仿射摄像机求仿射重构奠定了基础. 这篇论文应该整篇阅读. 仿射基本矩阵首先在[Zisserman－92]中定义，由 $F_A$ 计算运动

参数是在 Shapiro 等[Shapiro－95]中介绍，并且具体地给出第三幅视图不能解决浅浮雕多义性的情形．[Szeliski－96]给出了一种有益的多义性的特征矢量分析方法．三视图仿射运动情形在[Shimshoni－99]中作了很巧妙的处理．

### 13.8.2 注释和练习

(1) 一张景物平面诱导两个仿射摄像机之间的一个**仿射**变换．在 $\mathbb{R}^3$ 中存在着由平面的三参数簇定义的此种仿射变换的三参数簇．给定 $F_A$，这个仿射变换簇可以写为(结论 12.3($p$230))$H_A=[\mathbf{e}']_\times F_A+\mathbf{e}'\mathbf{v}^T$，其中 $F_A^T\mathbf{e}'=\mathbf{0}$，且此平面簇由 3 维矢量 $\mathbf{v}$ 参数化，反过来，证明给定由一张景物平面诱导的单应 $H_A$，则 $F_A$ 在相差一个单参数多义性下被确定．

(2) 考虑透视摄像机，即它的矩阵不具有仿射形式．证明：如果摄像机运动包含平行于像平面的一个平移以及绕主轴的一个旋转，则 F 有仿射形式．这一事实说明具有仿射形式的基本矩阵不蕴涵其影像条件必是仿射的．是否存在其他的摄像机运动也产生仿射形式的基本矩阵？

(3) 根据(8.1－$p$163)，两个仿射摄像机 $P_A$，$P'_A$ 唯一地定义了一个仿射基本矩阵 $F_A$．证明：若这两个摄像机矩阵被一个公共的仿射变换右乘，即 $P_A\mapsto P_AH_A$，$P'_A\mapsto P'_AH_A$，则变换后的摄像机仍定义原先的 $F_A$，这一事实证明了仿射基本矩阵对世界坐标的仿射变换是不变的．

(4) 假设两个摄像机之一是仿射的，而另一个是透视的．证明：对于这种情形两幅视图的对极点一般都是有限的．

(5) 4×4 置换单应

$$H=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}$$

把一个有限射影摄像机的标准矩阵 $P=[I|\mathbf{0}]$ 映射到平行投影 $P_A$ 的标准矩阵：

$$P_A=[I\mid\mathbf{0}]H=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\end{bmatrix}.$$

把这个变换应用于一对有限射影摄像机矩阵，证明本章的结论(如在表 13.1($p$247)所列出来的那些性质)能从前一章它们的对应非仿射部分直接地得到．具体地推导与 $F_A$ 相容的仿射摄像机对 $P_A$ 和 $P'_A$ 的表达式．

# 第 3 篇　三视图几何

## 本篇大纲

本篇的两章讨论三视图几何. 景物可由三目装置的三个摄像机同时拍摄,或者由一个移动摄像机相继地摄像.

第 14 章介绍一个新的多视图对象——三焦点张量. 它与两视图几何的基本矩阵有类似的性质:它仅依赖于摄像机间的(射影)关系,而与景物结构无关. 摄像机矩阵在相差一个 3 维空间的公共射影变换下可以从三焦点张量恢复,视图对之间的基本矩阵也可被唯一地确定.

与两视图情形相比较,这个新几何具有从两幅视图**转移**到第三幅视图的能力:给定其中两幅视图的一组点对应,该点在第三幅视图上的位置就可以确定;类似地,给定两幅视图的一组**直线**对应,该直线在第三幅视图上的位置就可以确定. 这个转移性质对建立多视图的对应有很大的用处.

如果说两幅视图的对极线约束的本质是从对应点反向投影的射线共面的话,那么三视图的三焦点约束的本质是由 3 维空间中直线上的一点的图像引起的一种点—线—线对应的几何:两幅视图中的对应直线反向投影成平面并在 3 维空间中交于一条直线,并且从第三幅视图的对应点反向投影的射线必然与这条直线相交.

第 15 章介绍由三幅视图的点和**线**对应求三焦点张量的计算. 给定张量和由它恢复的摄像机矩阵,射影重构可以由多视图的对应来算得. 如果以两视图同样的方式提供额外信息,这个重构可以升级为相似或度量重构.

在重构中,三视图比两视图几何多一个好处. 给定摄像机,两视图的每组点对应为 3 维空间中 3 个自由度点(位置)提供 4 个测量. 在三视图时,为这 3 个自由度的点却能提供 6 个测量. 然而,对直线有更大的好处. 在两视图时,测量的数目等于 3 维空间中直线的自由度数,即 4 个. 从而无法消除测量误差的影响. 然而,在三视图时,对 4 个自由度,能提供的景物直线的 6 个测量,因此景物直线是超定的,并且可以通过对测量误差适当的最小化来估计.

# 第 14 章　三焦点张量

三焦点张量在三视图中的作用类似于基本矩阵在两视图中的作用，它囊括了三幅视图间不依赖于景物结构的所有(射影)几何关系.

本章首先简要介绍三焦点张量的主要几何和代数性质. 三焦点张量及其性质的正式推导涉及张量记号的使用. 然而，为方便起见，我们先使用标准的矢量和矩阵符号，从而可不必使用一些(可能)不熟悉的符号就能先获得关于三焦点张量的某些几何洞察. 张量记号的使用在节 14.2 介绍.

节 14.1 介绍张量的三个主要的几何性质：通过第三幅视图中的直线反向投影得到的平面所诱导的两幅视图的单应性、由 3 维空间中的关联关系产生的点和线的图像对应关系以及由张量恢复基本矩阵和摄像机矩阵.

张量可以被用来把两幅视图的对应点转移到第三幅视图的对应点，它还可应用于直线：直线在一幅视图中的影像可由其在另两幅视图上对应的影像算出. 节 14.3 将介绍这种转移.

张量仅依赖于视图间的运动和摄像机内参数，且可由这些视图的摄像机矩阵唯一确定. 然而，它也可以仅由图像对应获得，而无需知道有关运动和标定的信息. 这种计算在第 15 章中讲述.

## 14.1　三焦点张量的几何基础

可以以多种方式引入三焦点张量，但本节的出发点是三条对应直线的关联关系.

**直线的关联关系**　如图 14.1 所示，假设三维空间中一条直线被影像到三幅视图中，那么在这些对应的像直线之间存在什么约束呢？由每一幅视图上的直线反向投影所得的平面必交于空间中的一条直线——投影到三幅图像中的匹配直线的 3D 直线. 因为空间中任意三张平面一般不交于一条直线，所以这个几何**关联**条件就对这组对应直线提供一个真正约束. 我们现在把对三条直线的这个几何约束翻译为代数约束.

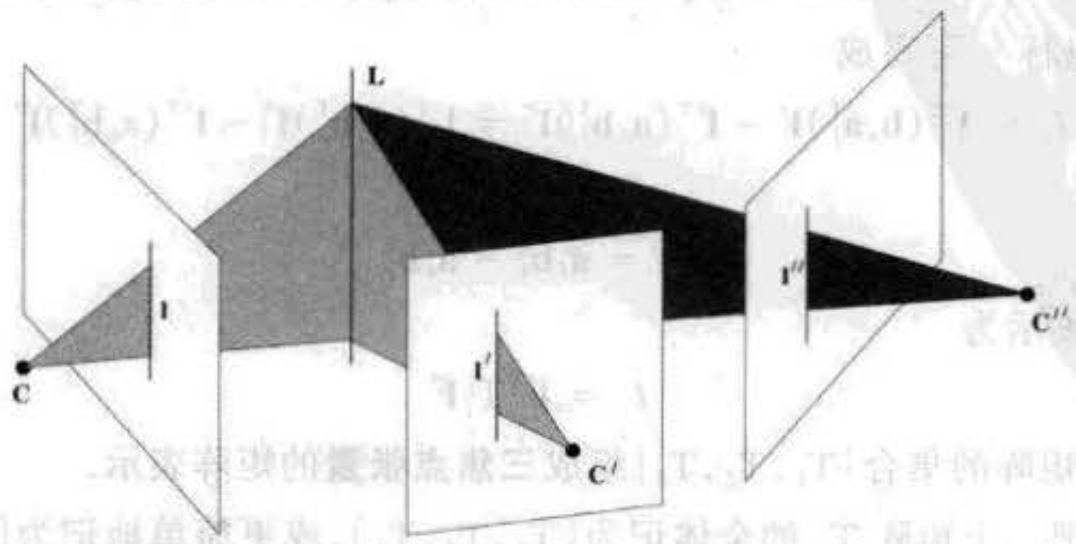



图 14.1　3 维空间中的直线 L 被影像成在由中心 C,C′,C″和像平面为标记的三幅视图中，形成对应的三线组 l↔l′↔l″. 反过来，对应直线通过第一、第二和第三幅图像的反向投影都交于空间的一条 3D 直线.

我们把一组对应的直线记为 $\mathbf{l}_i \leftrightarrow \mathbf{l}'_i \leftrightarrow \mathbf{l}''_i$. 令三幅视图的摄像机矩阵分别为 $P=[I|\mathbf{0}]$, $P'=[A|\mathbf{a}_4]$, $P''=[B|\mathbf{b}_4]$,其中 A 与 B 是 $3\times3$ 矩阵,矢量 $\mathbf{a}_i$ 和 $\mathbf{b}_i$ 是对应摄像机的第 $i$ 列, $i=1,\ldots,4$.

- $\mathbf{a}_4$ 和 $\mathbf{b}_4$ 分别为第二和第三幅视图上对应于第一摄像机的对极点. 本章将用 $\mathbf{e}'$ 和 $\mathbf{e}''$ 来记这些极点,即 $\mathbf{e}'=P'\mathbf{C}$, $\mathbf{e}''=P''\mathbf{C}$,其中 $\mathbf{C}$ 是第一摄像机的中心(大多数情形我们不考虑第二幅视图和第三幅视图之间的对极点).
- A 和 B 分别是第一个摄像机到第二个和第三个摄像机的无穷单应.

如在第 8 章所指出:在空间射影变换下任何三个摄像机的集合等价于包含有 $P=[I|\mathbf{0}]$ 的摄像机集合. 本章考虑的是在 3D 射影变换下不变的性质(例如:图像坐标和关联关系),所以可以放心地选取这种形式的摄像机集合.

如图 14.1 所示,每一条像直线反向投影为一张平面. 由结论 7.1($p$135),这三张平面是:

$$\boldsymbol{\pi} = P^{\mathsf{T}}\mathbf{l} = \begin{pmatrix}\mathbf{l}\\ 0\end{pmatrix} \quad \boldsymbol{\pi}' = P'^{\mathsf{T}}\mathbf{l}' = \begin{bmatrix}A^{\mathsf{T}}\mathbf{l}'\\ \mathbf{a}_4^{\mathsf{T}}\mathbf{l}'\end{bmatrix} \quad \boldsymbol{\pi}'' = P''^{\mathsf{T}}\mathbf{l}'' = \begin{bmatrix}B^{\mathsf{T}}\mathbf{l}''\\ \mathbf{b}_4^{\mathsf{T}}\mathbf{l}''\end{bmatrix}.$$

因为三条像直线由空间同一条直线产生,所以这三张平面并不是相互无关的,而必须交于这条公共的空间直线. 这个相交约束可以代数地描述为:$4\times3$ 矩阵 $M=[\boldsymbol{\pi},\boldsymbol{\pi}',\boldsymbol{\pi}'']$ 的秩为 2. 这一结论可以推导如下:该交线上的点可以表示为 $\mathbf{X}=\alpha\mathbf{X}_1+\beta\mathbf{X}_2$,其中 $\mathbf{X}_1$, $\mathbf{X}_2$ 线性无关. 这样的点在所有三张平面上,因此 $\boldsymbol{\pi}^{\mathsf{T}}\mathbf{X}=\boldsymbol{\pi}'^{\mathsf{T}}\mathbf{X}=\boldsymbol{\pi}''^{\mathsf{T}}\mathbf{X}=0$. 由此推出 $M^{\mathsf{T}}\mathbf{X}=\mathbf{0}$,因为 $M^{\mathsf{T}}\mathbf{X}_1=\mathbf{0}$ 和 $M^{\mathsf{T}}\mathbf{X}_2=\mathbf{0}$,所以 M 有 2 维的零空间.

这个相交约束将诱导出像直线 $\mathbf{l},\mathbf{l}',\mathbf{l}''$ 之间的一个关系. 因 M 的秩为 2,故其列 $\mathbf{m}_i$ 之间存在一个线性相关关系. 记

$$M = [\mathbf{m}_1,\mathbf{m}_2,\mathbf{m}_3] = \begin{bmatrix}\mathbf{l} & A^{\mathsf{T}}\mathbf{l}' & B^{\mathsf{T}}\mathbf{l}''\\ \mathbf{0} & \mathbf{a}_4^{\mathsf{T}}\mathbf{l}' & \mathbf{b}_4^{\mathsf{T}}\mathbf{l}''\end{bmatrix}$$

则该线性相关性可写成 $\mathbf{m}_1=\alpha\mathbf{m}_2+\beta\mathbf{m}_3$. 再注意到 $\mathbf{M}$ 的左下角元素为 0,从而可推出 $\alpha=k(\mathbf{b}_4^{\mathsf{T}}\mathbf{l}'')$, $\beta=-k(\mathbf{a}_4^{\mathsf{T}}\mathbf{l}')$,其中 $k$ 为尺度因子. 将此式应用于 M 各列顶部的三维矢量,得到(相差一个齐次尺度因子)

$$\mathbf{l} = (\mathbf{b}_4^{\mathsf{T}}\mathbf{l}'')A^{\mathsf{T}}\mathbf{l}' - (\mathbf{a}_4^{\mathsf{T}}\mathbf{l}')B^{\mathsf{T}}\mathbf{l}'' = (\mathbf{l}''^{\mathsf{T}}\mathbf{b}_4)A^{\mathsf{T}}\mathbf{l}' - (\mathbf{l}'^{\mathsf{T}}\mathbf{a}_4)B^{\mathsf{T}}\mathbf{l}''$$

因此 $\mathbf{l}$ 的第 $i$ 个坐标 $l_i$ 可写成

$$l_i = \mathbf{l}''^{\mathsf{T}}(\mathbf{b}_4\mathbf{a}_i^{\mathsf{T}})\mathbf{l}' - \mathbf{l}'^{\mathsf{T}}(\mathbf{a}_4\mathbf{b}_i^{\mathsf{T}})\mathbf{l}'' = \mathbf{l}'^{\mathsf{T}}(\mathbf{a}_i\mathbf{b}_4^{\mathsf{T}})\mathbf{l}'' - \mathbf{l}'^{\mathsf{T}}(\mathbf{a}_4\mathbf{b}_i^{\mathsf{T}})\mathbf{l}''$$

引入记号

$$T_i = \mathbf{a}_i\mathbf{b}_4^{\mathsf{T}} - \mathbf{a}_4\mathbf{b}_i^{\mathsf{T}} \tag{14.1}$$

则该关联关系可表示为

$$l_i = \mathbf{l}'^{\mathsf{T}}T_i\mathbf{l}'' \tag{14.2}$$

**定义 14.1** 三个矩阵的集合 $\{T_1,T_2,T_3\}$ 组成**三焦点张量**的矩阵表示.

我们进一步把三个矩阵 $T_i$ 的全体记为 $[T_1,T_2,T_3]$,或更简单地记为 $[T_i]$①,则上面最

① 此记号仍嫌冗赘,其义也不能自明. 基于这个理由,14.2 节将介绍张量记号.

后一个关系式可改写为

$$\mathbf{l}^{\mathsf{T}} = \mathbf{l}'^{\mathsf{T}}[\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3]\mathbf{l}'' \tag{14.3}$$

其中 $\mathbf{l}'^{\mathsf{T}}[\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3]\mathbf{l}''$ 理解为矢量 $(\mathbf{l}'^{\mathsf{T}}\mathrm{T}_1\mathbf{l}'', \mathbf{l}'^{\mathsf{T}}\mathrm{T}_2\mathbf{l}'', \mathbf{l}'^{\mathsf{T}}\mathrm{T}_3\mathbf{l}'')$.

显然三幅视图之间不存在本质区别，因此由(14.3)类推还可得到类似的关系 $\mathbf{l}'^{\mathsf{T}} = \mathbf{l}^{\mathsf{T}}[\mathrm{T}'_i]\mathbf{l}''$ 和 $\mathbf{l}''^{\mathsf{T}} = \mathbf{l}^{\mathsf{T}}[\mathrm{T}''_i]\mathbf{l}'$. 这三个张量 $[\mathrm{T}_i]$, $[\mathrm{T}'_i]$, $[\mathrm{T}''_i]$ 都存在但有区别. 事实上，尽管这三个张量都可以分别由其中任何一个计算得到，但它们之间不存在简单的关系. 因此，对于给定的三幅视图，事实上存在着三个三焦点张量. 通常人们满足于只考虑其中的一个. 不过，练习(7)(*p*275)概述了一种由给定张量 $[\mathrm{T}_i]$ 计算另两个张量 $[\mathrm{T}'_i]$、$[\mathrm{T}''_i]$ 的方法.

注意：(14.3)仅仅是图像坐标之间的一种关系，并不涉及 3D 坐标，因此(如同前面所注释过的)，尽管它是在标准摄像机集(即 $\mathrm{P} = [\mathrm{I}|\mathbf{0}]$)的假设下推出的，但矩阵元素 $[\mathrm{T}_i]$ 的值并不依赖于摄像机的形式. 摄像机矩阵给定的三焦点张量的特别简单的公式(14.1)只适用于 $\mathrm{P} = [\mathrm{I}|\mathbf{0}]$，我们今后将推导对应于任意三个摄像机的三焦点张量的一般公式(16.12 – *p*294).

**自由度** 三焦点张量由三个 3×3 矩阵组成，因此有 27 个元素，所以除去整个矩阵的一个全局尺度后有 26 个独立的比率. 然而，该张量仅有 18 个独立的自由度. 换句话说，一旦指定了其中 18 个参数，在相差一个公共尺度下所有 27 个元素都可以确定. 自由度的数目可计算如下：三个摄像机矩阵分别都有 11 个自由度，共计有 33 个. 然而，考虑了世界射影坐标系须减去 15 个自由度，从而留下 18 个自由度. 所以张量满足 26 – 18 = 8 个独立的代数约束. 我们将在第 15 章再回到这个问题.

### 14.1.1 由平面诱导的单应

三焦点张量固有的一个基本几何性质是存在由第二幅图像上的一条直线诱导的第一幅图像和第三幅图像之间的单应，这一事实在图 14.2 和 14.3 中作了解释. 第二幅图像上的一条直线由反向投影定义了 3 维空间中的一张平面，同时这张平面诱导了第一幅图像和第三幅之间的一个单应.

我们现在利用三焦点张量来推导这个几何性质的代数表示. 图 14.2 和 14.3 中由平面 $\boldsymbol{\pi}'$ 定义的第一幅和第三幅图像之间的单应映射可以分别写成 $\mathbf{x}'' = \mathrm{H}\mathbf{x}$ 和 $\mathbf{l} = \mathrm{H}^{\mathsf{T}}\mathbf{l}''$(1.6 – *p*10). 注意，图 14.3 中三条直线 $\mathbf{l}$、$\mathbf{l}'$ 和 $\mathbf{l}''$ 是一个对应三线组，它们是 3D 直线 $\mathbf{L}$ 的投影. 因此它们满足(14.2)给出的线的关联关系 $\mathbf{l}_i = \mathbf{l}'^{\mathsf{T}}\mathrm{T}_i\mathbf{l}''$. 比较这个公式和 $\mathbf{l} = \mathrm{H}^{\mathsf{T}}\mathbf{l}''$ 证明了

$$\mathrm{H} = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3], \text{其中 } \mathbf{h}_i = \mathrm{T}_i^{\mathsf{T}}\mathbf{l}'.$$

因此由上面公式定义的 H 表示由第二幅视图的直线 $\mathbf{l}'$ 所确定的第一幅视图和第三幅之间的(点)单应 $\mathrm{H}_{13}$.

第二幅和第三幅视图起着类似的作用，由第三幅视图的一条直线确定的第一幅视图和第二幅之间的单应可以用类似的方法得到. 这些思想形式化成下列结论：

**结论 14.2 由第二幅图像的直线 $\mathbf{l}'$ 诱导的第一幅图像到第三幅图像上的单应(见图 14.2)由 $\mathbf{x}'' = \mathrm{H}_{13}(\mathbf{l}')\mathbf{x}$ 给出，其中**

$$\mathrm{H}_{13}(\mathbf{l}') = [\mathrm{T}_1^{\mathsf{T}}, \mathrm{T}_2^{\mathsf{T}}, \mathrm{T}_3^{\mathsf{T}}]\mathbf{l}'.$$

**类似地，第三幅视图的直线 $\mathbf{l}''$ 定义了第一幅图像到第二幅上的一个单应 $\mathbf{x}' = \mathrm{H}_{12}(\mathbf{l}'')\mathbf{x}$，其中**

$$H_{12}(\mathbf{l}'') = [T_1, T_2, T_3]\mathbf{l}''.$$

一旦理解了这种映射，张量的代数性质就一目了然而且可以很容易地得到. 下一节中我们将基于(14.3)和结论 14.2 来推导点和线之间存在的若干关联关系.

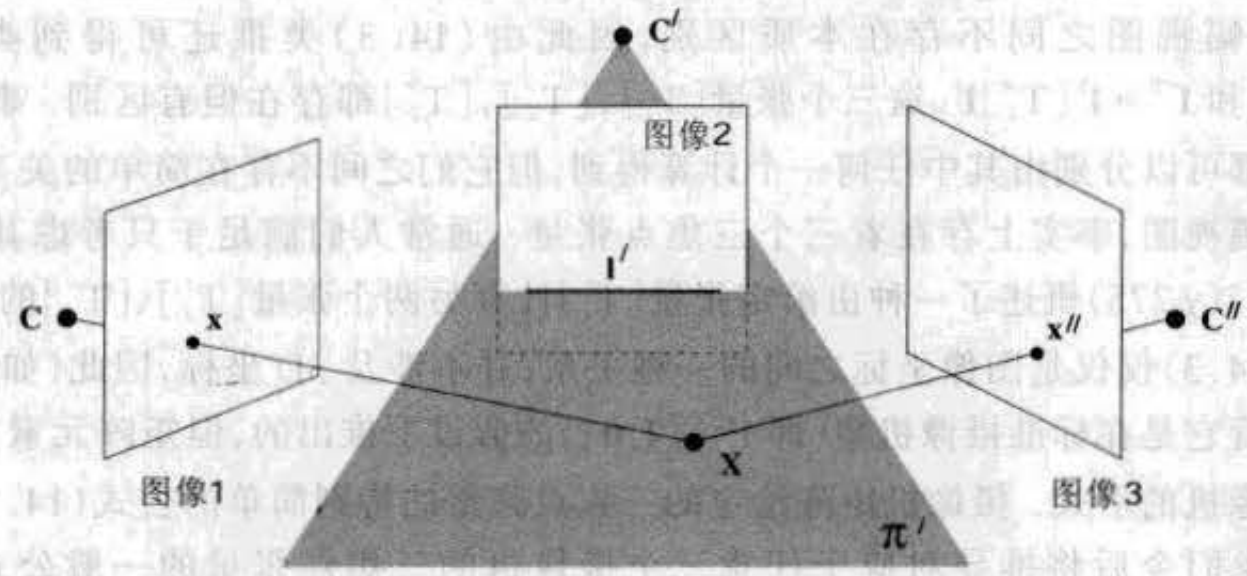



图 14.2 **点转移**. 在第二幅视图的直线 $\mathbf{l}'$ 反向投影为 3 维空间的一张平面 $\pi'$. 第一幅视图的点 $\mathbf{x}$ 定义了 3 维空间的一条射线，它与 $\pi'$ 的交点是 $\mathbf{X}$. 再把点 $\mathbf{X}$ 影像成第三幅视图的点 $\mathbf{x}''$. 因此，任何直线 $\mathbf{l}'$ 诱导的第一和第三幅视图间的单应由其反向投影平面 $\pi'$ 所定义.

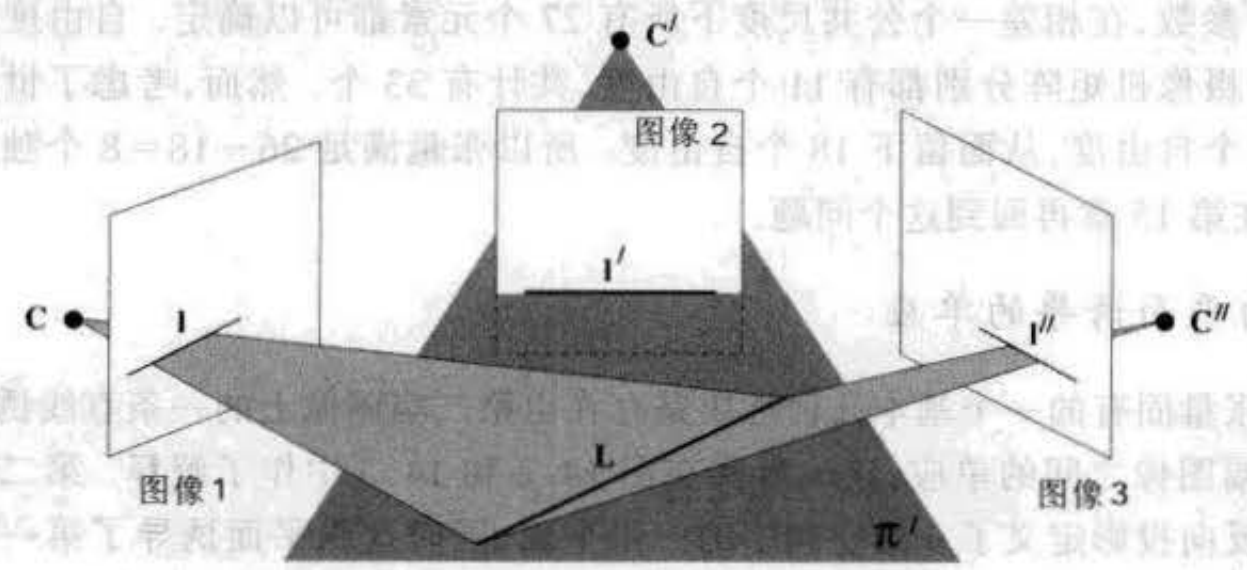



图 14.3 **线转移**. 图 14.2 所定义的单应在直线上的作用同样可以用几何方式可视化. 在第一幅图像的直线 $\mathbf{l}$ 定义了 3 维空间的一张平面，它与 $\pi'$ 的交线是 $\mathbf{L}$. 然后直线 $\mathbf{L}$ 又被影像成第三幅视图的直线 $\mathbf{l}''$.

### 14.1.2 点与线关联关系

利用三焦点张量很容易推导出三幅图像上线和点的各种线性关系. 我们已经见过一个这样的关系，即(14.3). 这个关系在相差一个尺度的条件下成立，因为它涉及到齐次量. 我们可以通过对此式两边矢量进行叉积(它必然为 $\mathbf{0}$)来消除这个尺度因子. 由此导出下列公式

$$(\mathbf{l}'^{\mathsf{T}}[T_1, T_2, T_3]\mathbf{l}'')[\mathbf{l}]_\times = \mathbf{0}^{\mathsf{T}} \tag{14.4}$$

这里我们用矩阵 $[\mathbf{l}]_\times$ 来表示叉积(见(A3.4 – $p$406))，或更简洁地记为 $(\mathbf{l}'^{\mathsf{T}}[T_i]\mathbf{l}'')[\mathbf{l}]_\times = \mathbf{0}^{\mathsf{T}}$. 考虑到 $\mathbf{l}'$ 和 $\mathbf{l}''$ 的对称性——对换这两条直线的位置，相当于把每一 $T_i$ 转置，则得到关系 $(\mathbf{l}''^{\mathsf{T}}[T_i^{\mathsf{T}}]\mathbf{l}')[\mathbf{l}]_\times = \mathbf{0}^{\mathsf{T}}$.

再考虑图 14.3，直线 $\mathbf{l}$ 上一点 $\mathbf{x}$ 必满足 $\mathbf{x}^{\mathsf{T}}\mathbf{l} = \sum_i x^i l_i = 0$(用上标表示点坐标，预示张量符号的使用). 因为 $l_i = \mathbf{l}'^{\mathsf{T}}T_i\mathbf{l}''$，上式可写为

$$\mathbf{l}'^{\mathrm{T}}\left(\sum_i x^i \mathrm{T}_i\right)\mathbf{l}'' = 0 \tag{14.5}$$

(注意:$(\sum_i x^i \mathrm{T}_i)$是一个 3×3 矩阵). 这是在第一幅图像上的一个关联关系:该关系对于一组点一线一线对应成立,此种对应就是当某条 3D 直线 **L** 映射到第二幅和第三幅图像的 $\mathbf{l}'$ 和 $\mathbf{l}''$ 及第一幅图像中通过点 **x** 的一条直线. 使(14.5)成立的点一线一线对应的重要等价定义产生于 3 **维空间的一种关联关系**——存在一个 3D 点 **X** 映到第一幅图像的 **x** 及第二幅和第三幅图像的直线 $\mathbf{l}'$ 和 $\mathbf{l}''$ 上的点(如图 14.4(a)所示).

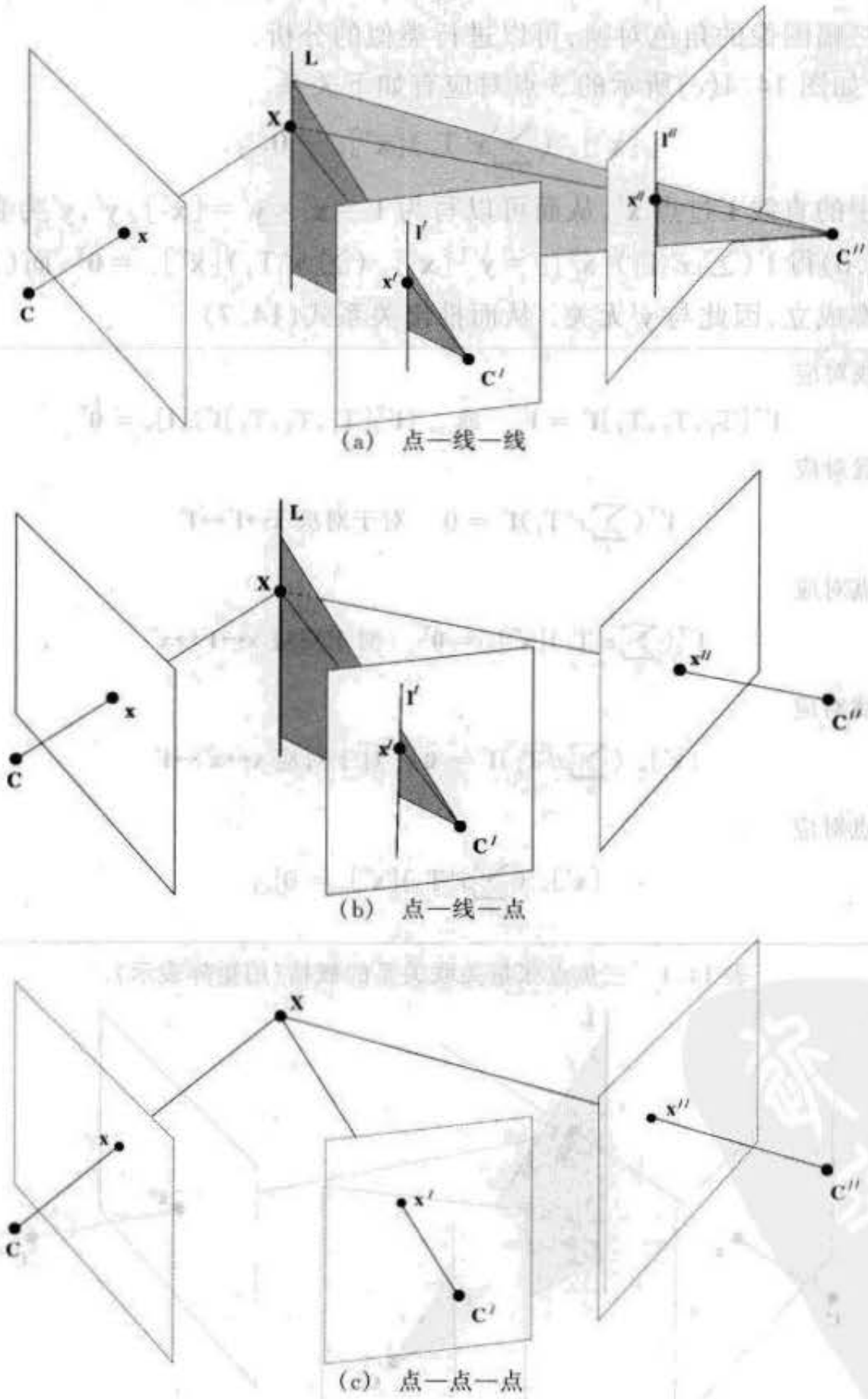



图 14.4 **关联关系**. (a)考虑 3 幅视图的一组点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'\leftrightarrow\mathbf{x}''$. 若 $\mathbf{l}'$ 和 $\mathbf{l}''$ 分别为过 $\mathbf{x}'$ 和 $\mathbf{x}''$ 的**任意**直线,则 $\mathbf{x}\leftrightarrow\mathbf{l}'\leftrightarrow\mathbf{l}''$ 组成对应于一条3D 直线 **L** 的一组点一线一线对应. 因此,(14.5)对过 $\mathbf{x}'$ 的直线 $\mathbf{l}'$ 和过 $\mathbf{x}''$ 的直线 $\mathbf{l}''$ 的任何选择都成立. (b)空间点 **X** 与空间直线 **L** 相关联. 它定义了所在图像的一组关联关系 $\mathbf{x}\leftrightarrow\mathbf{l}'\leftrightarrow\mathbf{x}''$. (c)由空间点 **X** 的图像产生的对应 $\mathbf{x}\leftrightarrow\mathbf{x}'\leftrightarrow\mathbf{x}''$.

根据结论 14.2,我们可以得到包含第二幅和第三幅图像的点 $\mathbf{x}'$ 和 $\mathbf{x}''$ 的关系. 参考图 14.4(b)所示的点—线—点对应得到

$$\mathbf{x}'' = \mathrm{H}_{13}(\mathbf{l}')\mathbf{x} = [\mathrm{T}_1^{\mathsf{T}}\mathbf{l}', \mathrm{T}_2^{\mathsf{T}}\mathbf{l}', \mathrm{T}_3^{\mathsf{T}}\mathbf{l}']\mathbf{x} = (\sum_i x^i \mathrm{T}_i^{\mathsf{T}})\mathbf{l}'$$

此式对第二幅图像上任何一条过点 $\mathbf{x}'$ 的直线 $\mathbf{l}'$ 都成立. 把此式两边的转置同时(右)乘 $[\mathbf{x}'']_\times$ 可以消除齐次尺度因子而得到

$$\mathbf{x}''^{\mathsf{T}}[\mathbf{x}'']_\times = \mathbf{l}'^{\mathsf{T}}(\sum_i x^i \mathrm{T}_i)[\mathbf{x}'']_\times = \mathbf{0}^{\mathsf{T}} \tag{14.6}$$

把第二幅和第三幅图像的角色对换,可以进行类似的分析.

最后,对于如图 14.4(c)所示的 3 点对应有如下关系

$$[\mathbf{x}']_\times (\sum_i x^i \mathrm{T}_i)[\mathbf{x}'']_\times = 0_{3\times 3}. \tag{14.7}$$

**证明** (14.6)中的直线 $\mathbf{l}'$ 过点 $\mathbf{x}'$,从而可以写为 $\mathbf{l}' = \mathbf{x}' \times \mathbf{y}' = [\mathbf{x}']_\times \mathbf{y}'$,$\mathbf{y}'$ 为直线 $\mathbf{l}'$ 上的某一点. 于是由(14.6)得 $\mathbf{l}'(\sum_i x^i \mathrm{T}_i)[\mathbf{x}'']_\times = \mathbf{y}'^{\mathsf{T}}[\mathbf{x}']_\times(\sum_i x^i \mathrm{T}_i)[\mathbf{x}'']_\times = \mathbf{0}^{\mathsf{T}}$. 而(14.6)对所有通过 $\mathbf{x}'$ 的直线 $\mathbf{l}'$ 都成立,因此与 $\mathbf{y}'$ 无关. 从而推出关系式(14.7) □

(1) 线—线—线对应

$$\mathbf{l}'^{\mathsf{T}}[\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3]\mathbf{l}'' = \mathbf{l}^{\mathsf{T}} \quad 或 \quad (\mathbf{l}'^{\mathsf{T}}[\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3]\mathbf{l}'')[\mathbf{l}]_\times = \mathbf{0}^{\mathsf{T}}$$

(2) 点—线—线对应

$$\mathbf{l}'^{\mathsf{T}}(\sum_i x^i \mathrm{T}_i)\mathbf{l}'' = 0 \quad 对于对应\ \mathbf{x} \leftrightarrow \mathbf{l}' \leftrightarrow \mathbf{l}''$$

(3) 点—线—点对应

$$\mathbf{l}'^{\mathsf{T}}(\sum_i x^i \mathrm{T}_i)[\mathbf{x}'']_\times = \mathbf{0}^{\mathsf{T}} \quad 对于对应\ \mathbf{x} \leftrightarrow \mathbf{l}' \leftrightarrow \mathbf{x}''$$

(4) 点—点—线对应

$$[\mathbf{x}']_\times (\sum_i x^i \mathrm{T}_i)\mathbf{l}'' = \mathbf{0} \quad 对于对应\ \mathbf{x} \leftrightarrow \mathbf{x}' \leftrightarrow \mathbf{l}''$$

(5) 点—点—点对应

$$[\mathbf{x}']_\times (\sum_i x^i \mathrm{T}_i)[\mathbf{x}'']_\times = 0_{3\times 3}$$

表 14.1 三焦点张量关联关系的概括(用矩阵表示).

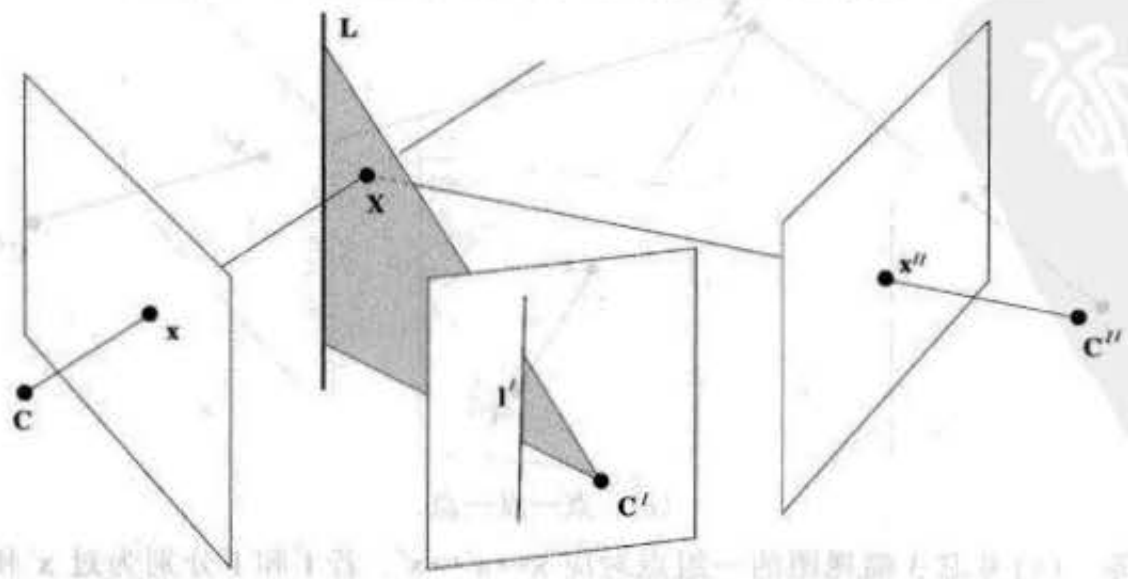



图 14.5 **非关联结构**. 这个结构的像点和像直线满足表 14.1 中的点—线—点关联关系. 但是空间点 **X** 与直线 **L** 不关联. 请与图 14.4 比较.

三幅视图中线和点之间的各种关系概括在表 14.1 中，它们的性质在引入张量符号后会在节 14.2.1 中作进一步研究. 注意：这里没有列出点在第二幅或第三幅视图上的点—线—线对应关系，当第一幅视图是特殊视图时，这种以三焦点张量表示的简单关系不成立. 还值得注意的是如图 14.5 所示满足图像的关联关系并不保证 3 维空间中的关联关系.

我们现在由三焦点张量求两视图几何、对极点和基本矩阵.

### 14.1.3 对极线

当由直线 $\mathbf{l}'$ 反向投影得到的平面 $\boldsymbol{\pi}'$ 是前两个摄像机的对极平面(从而通过第一个摄像机的中心 $\mathbf{C}$)时，就产生一种点—线—线对应的特殊情形. 设 $\mathbf{X}$ 是平面 $\boldsymbol{\pi}'$ 的一个点，那么由 $\mathbf{X}$ 和 $\mathbf{C}$ 定义的射线在这张平面上，并且 $\mathbf{l}'$ 是对应于 $\mathbf{X}$ 的图像点 $\mathbf{x}$ 的对极线，如图 14.6 所示.

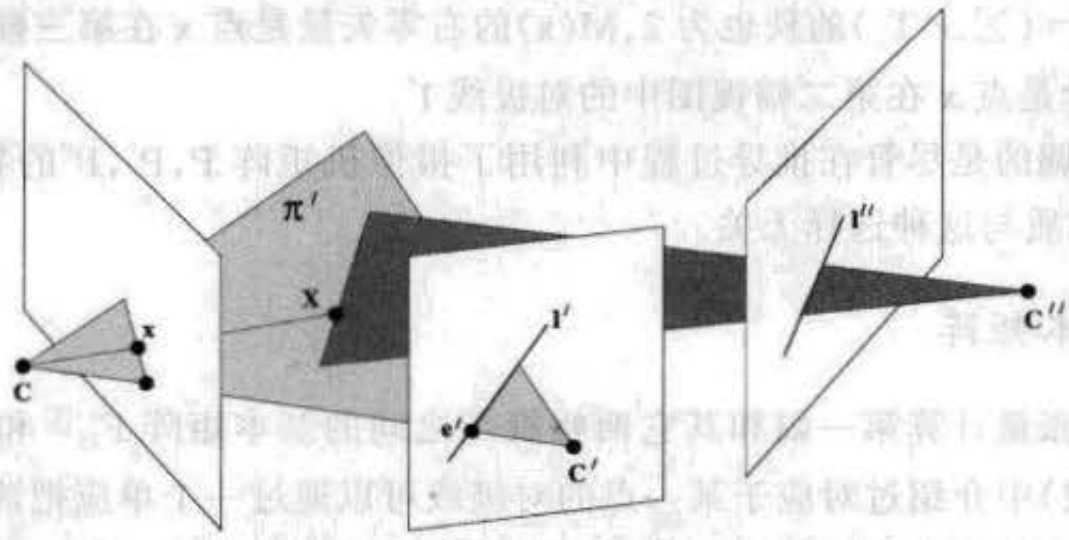



图 14.6 若由 $\mathbf{l}'$ 确定的平面 $\boldsymbol{\pi}'$ 是前两幅视图的一张对极平面，则在第三幅视图上的任意直线 $\mathbf{l}''$ 都给出一个点—线—线关联.

由第三幅图像上直线 $\mathbf{l}''$ 反向投影得到的平面 $\boldsymbol{\pi}''$ 交平面 $\boldsymbol{\pi}'$ 于直线 $\mathbf{L}$. 此外，因对应于 $\mathbf{x}$ 的射线完全在平面 $\boldsymbol{\pi}'$ 上，故它必与 $\mathbf{L}$ 相交. 这就给出了由点 $\mathbf{x}$ 和直线 $\mathbf{l}'$，$\mathbf{l}''$ 分别反向投影得到的射线和两幅平面的 3 路交点，从而它们构成了满足 $\mathbf{l}'^{\mathsf{T}}(\sum_i x^i \mathrm{T}_i)\mathbf{l}''=0$ 的一组点—线—线对应. 这里至关重要的一点是：上式对**任意**的直线 $\mathbf{l}''$ 都成立，从而推出 $\mathbf{l}'^{\mathsf{T}}(\sum_i x^i \mathrm{T}_i)=\mathbf{0}^{\mathsf{T}}$. 当 $\mathbf{l}'$ 和 $\mathbf{l}''$ 的角色对换时同样的讨论也成立. 归纳如下：

**结论 14.3 如果 $\mathbf{x}$ 是一个点，$\mathbf{l}'$ 和 $\mathbf{l}''$ 分别是第二幅和第三幅图像上对应于它的对极线，那么**

$$\mathbf{l}'^{\mathsf{T}}\left(\sum_i x^i \mathrm{T}_i\right) = \mathbf{0}^{\mathsf{T}} \text{ 和} \left(\sum_i x^i \mathrm{T}_i\right)\mathbf{l}'' = \mathbf{0}$$

**从而，点 $\mathbf{x}$ 对应的对极线 $\mathbf{l}'$ 和 $\mathbf{l}''$ 可以作为矩阵 $\sum_i x^i \mathrm{T}_i$ 的左、右零矢量而算得.**

当点 $\mathbf{x}$ 变化时，对应的对极线也随之而变，但一幅图像上的所有对极线都通过对极点. 这样一来可以通过计算不同 $\mathbf{x}$ 所对应的对极线的交点来求该对极点. 点 $\mathbf{x}$ 的三个方便的选择是由齐次坐标：$(1,0,0)^{\mathsf{T}}$，$(0,1,0)^{\mathsf{T}}$，$(0,0,1)^{\mathsf{T}}$ 表示的 3 个点，对 $\mathbf{x}$ 的这三种选择，$\sum_i x^i \mathrm{T}_i$ 分别等于 $\mathrm{T}_1$，$\mathrm{T}_2$ 和 $\mathrm{T}_3$. 由此推导出下列重要结论：

**结论 14.4 第二幅图像上的对极点 $\mathbf{e}'$ 是由矩阵 $\mathrm{T}_i\,(i=1,\ldots,3)$ 的左零矢量所表示的对极线的公共交点. 类似地对极点 $\mathbf{e}''$ 是由矩阵 $\mathrm{T}_i$ 的右零矢量所表示的对极线的公共交点.**

注意这里的对极点是第二幅和第三幅图像中对应于第一幅图像中心 $\mathbf{C}$ 的对极点.

目前这个结论的用处还不是很明显，但下面将会看到，它是由三焦点张量计算摄像机矩

阵和在第 15 章中有关张量精确计算的一个重要步骤.

**矩阵 $T_i$ 的代数性质** 本节给出了矩阵 $T_i$ 的一些代数性质,我们将其归纳如下:

- 每一矩阵 $T_i$ 的秩为 2,可由(14.1)推出,因为 $T_i = \mathbf{a}_i\mathbf{e}''^{\mathsf{T}} - \mathbf{e}'\mathbf{b}_i^{\mathsf{T}}$ 是两个外积的和.
- $T_i$ 的右零矢量是 $\mathbf{l}_i'' = \mathbf{e}'' \times \mathbf{b}_i^{\mathsf{T}}$, $i$ 分别取 1,2,3 时,它分别是点 $\mathbf{x} = (1,0,0)^{\mathsf{T}}$, $(0,1,0)^{\mathsf{T}}$、$(0,0,1)^{\mathsf{T}}$在第三幅视图上的对极线.
- 对极点 $\mathbf{e}''$是对极线 $\mathbf{l}_i''(i=1,2,3)$的公共交点.
- $T_i$ 的左零失量是 $\mathbf{l}_i' = \mathbf{e}' \times \mathbf{a}_i^{\mathsf{T}}$, $i$ 分别取 1,2,3 时,它分别是点 $\mathbf{x} = (1,0,0)^{\mathsf{T}}$, $(0,1,0)^{\mathsf{T}}$,$(0,0,1)^{\mathsf{T}}$在第二幅视图上的对极线.
- 对极点 $\mathbf{e}'$是对极线 $\mathbf{l}_i'(i=1,2,3)$的公共交点.
- 矩阵和 $M(\mathbf{x}) = (\sum_i x^i T_i)$的秩也为 2,$M(\mathbf{x})$的右零矢量是点 $\mathbf{x}$ 在第三幅视图中的对极线 $\mathbf{l}''$,其左零矢量是点 $\mathbf{x}$ 在第二幅视图中的对极线 $\mathbf{l}'$.

值得再次强调的是尽管在推导过程中利用了摄像机矩阵 $P,P',P''$的特殊标准形式,但 $T_i$ 矩阵的对极性质与这种选择无关.

### 14.1.4 求基本矩阵

利用三焦点张量计算第一幅和其它两幅视图之间的基本矩阵 $F_{21}$[①] 和 $F_{31}$ 是很简单的. 在节 8.2.1($p$162)中介绍过对应于某一点的对极线可以通过一个单应把该点转移到另一幅视图上再由连接该转移点和对极点而得到. 考虑第一幅视图上的一点 $\mathbf{x}$. 根据图 14.2 和结论14.2,$\mathbf{x}' = ([T_1,T_2,T_3]\mathbf{l}'')\mathbf{x}$ 给出了由第三幅视图上的直线 $\mathbf{l}''$诱导的从第一幅到第二幅视图的单应. 然后对应于 $\mathbf{x}$ 的对极线由连接 $\mathbf{x}'$和极点 $\mathbf{e}'$的直线 $\mathbf{l}' = [\mathbf{e}']_\times([T_1,T_2,T_3]\mathbf{l}'')\mathbf{x}$ 给出. 由此得到

$$F_{21} = [\mathbf{e}']_\times [T_1,T_2,T_3]\mathbf{l}''$$

这个公式对任意矢量 $\mathbf{l}''$都成立;但要避免 $\mathbf{l}''$处于任一 $T_i$ 的零空间中的退化条件. 一个好的选择是 $\mathbf{e}''$,因为我们已经知道 $\mathbf{e}''$垂直于每一个 $T_i$ 的右零空间. 这就给出下面的公式

$$F_{21} = [\mathbf{e}']_\times [T_1,T_2,T_3]\mathbf{e}'' \tag{14.8}$$

类似地,公式 $F_{31} = [\mathbf{e}'']_\times [T_1^{\mathsf{T}},T_2^{\mathsf{T}},T_3^{\mathsf{T}}]\mathbf{e}'$成立.

### 14.1.5 恢复摄像机矩阵

我们强调过由于三焦点张量仅表达图像元素之间的关系,故它与 3D 射影变换无关. 反之,这一事实说明由三焦点张量可以在只差一个射影多义性的意义下算得摄像机矩阵. 现在就来说明如何来实现.

如两幅图像重构的情形一样,由于射影多义性,第一个摄像机矩阵可以取为$P = [I|\mathbf{0}]$. 现在因为 $F_{21}$ 已知(由(14.8)),我们可以利用结论 8.9($p$171)推出第二个摄像机矩阵形为

$$P' = [[T_1,T_2,T_3]\mathbf{e}''|\mathbf{e}']$$

从而摄像机对$\{P,P'\}$具有基本矩阵 $F_{21}$. 也许有人会认为第三个摄像机可以类似地取为

① 对于对应点 $\mathbf{x}\leftrightarrow\mathbf{x}'$,基本矩阵 $F_{21}$ 满足 $\mathbf{x}'^{\mathsf{T}}F_{21}\mathbf{x}=0$. 下标记号参见图 14.8.

$P'' = [[T_1^T, T_2^T, T_3^T]\mathbf{e}' | \mathbf{e}'']$，但**这是错误的**. 其原因是两个摄像机对$\{P, P'\}$和$\{P, P''\}$并不一定定义在同一个射影世界坐标系中；尽管每一对就其本身是正确的，但整个三元组$\{P, P', P''\}$是不相容的.

第三个摄像机的选择不能独立于前两个摄像机的射影坐标系. 为解释这一点，设摄像机对$\{P, P'\}$已选定并且点$\mathbf{X}_i$由其图像对应$\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$重构得到，那么$\mathbf{X}_i$的坐标由$\{P, P'\}$定义的射影坐标系确定，并且一个相容的摄像机$P''$可以由对应$\mathbf{X}_i \leftrightarrow \mathbf{x}_i''$而算得. 显然，$P''$依赖于由$\{P, P'\}$定义的坐标系. 然而，并无必要显式地重构出三维结构，因为一个相容的摄像机三元组可以直接由三焦点张量来得到.

$P = [I | \mathbf{0}]$和$P' = [[T_1, T_2, T_3]\mathbf{e}'' | \mathbf{e}']$并不是与给定基本矩阵$F_{21}$相容的唯一的摄像机矩阵对. 根据(8.10 - $p$173)，$P'$的最一般的形式是$P' = [[T_1, T_2, T_3]\mathbf{e}'' + \mathbf{e}'\mathbf{v}^T | \lambda\mathbf{e}']$，其中$\mathbf{v}$是某个矢量，$\lambda$是尺度因子；对$P''$类似的选择也成立. 为了求得一组与该三焦点张量相容的摄像机矩阵三元组，我们需要从这两矩阵簇中找到与三焦点张量的形式(14.1)相容的$P'$和$P''$的正确值.

由于射影多义性，我们可自由选取$P' = [[T_1, T_2, T_3]\mathbf{e}'' | \mathbf{e}']$，因此$\mathbf{a}_i = T_i\mathbf{e}''$，这个选取固定了该射影世界坐标系，使$P''$被唯一确定(相差一个尺度因子). 然后代入到(14.1)(注意$\mathbf{a}_4 = \mathbf{e}'$和$\mathbf{b}_4 = \mathbf{e}''$)

$$T_i = T_i\mathbf{e}''\mathbf{e}''^T - \mathbf{e}'\mathbf{b}_i^T$$

由此得到$\mathbf{e}'\mathbf{b}_i^T = T_i(\mathbf{e}''\mathbf{e}''^T - I)$. 因为可以通过选择尺度因子使得$\|\mathbf{e}'\| = \mathbf{e}'^T\mathbf{e}' = 1$，上式两边用$\mathbf{e}'^T$左乘并转置就得到$\mathbf{b}_i = (\mathbf{e}''\mathbf{e}''^T - I)T_i^T\mathbf{e}'$.

因此$P'' = [(\mathbf{e}''\mathbf{e}''^T - I)[T_1^T, T_2^T, T_3^T]\mathbf{e}' | \mathbf{e}'']$. 由三焦点张量求摄像机矩阵的步骤归纳在算法14.1中.

给定用矩阵$[T_1, T_2, T_3]$表示的三焦点张量

(1) **求对极点$\mathbf{e}', \mathbf{e}''$**

令$\mathbf{u}_i$和$\mathbf{v}_i$分别为$T_i$的左和右零矢量，即$\mathbf{u}_i^T T_i = \mathbf{0}^T$，$T_i\mathbf{v}_i = \mathbf{0}$. 则此对极点可以通过求下面3×3矩阵的零矢量而获得：

$$\mathbf{e}'^T[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] = \mathbf{0}^T \text{ 和 } \mathbf{e}''^T[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = \mathbf{0}^T$$

(2) **求基本矩阵$F_{21}, F_{31}$**

$$F_{21} = [\mathbf{e}']_\times [T_1^T, T_2^T, T_3^T]\mathbf{e}'' \text{ 和 } F_{31} = [\mathbf{e}'']_\times [T_1^T, T_2^T, T_3^T]\mathbf{e}'$$

(3) **求摄像机矩阵$P', P''$($P = [I | \mathbf{0}]$)**

把对极点归一化为单位范数. 那么

$$P' = [[T_1, T_2, T_3]\mathbf{e}'' | \mathbf{e}'] \text{ 和 } P'' = [(\mathbf{e}''\mathbf{e}''^T - I)[T_1^T, T_2^T, T_3^T]\mathbf{e}' | \mathbf{e}'']$$

算法 14.1 **由三焦点张量求F和P的算法概括**. 注意$F_{21}$和$F_{31}$是唯一确定的. 但是$P'$和$P''$只确定到相差一个3维空间的公共的射影变换的程度.

我们已经知道三焦点张量可以由三个摄像机矩阵计算得到；反之，三个摄像机矩阵可以在相差一个射影等价的意义下由三焦点张量计算得到，因此，三焦点张量在相差一个射影等价下完全掌握了三个摄像机.

## 14.2 三焦点张量和张量记号

到现在为止，我们所用的三焦点张量的记号都是由标准的矩阵、矢量记号推导出来的. 因为矩阵仅有两个指标，所以可以利用矩阵的转置及左乘或右乘来区分这两个指标，从而在处理矩阵和矢量时，可以不必明确写出它们的指标. 与矩阵只有两个指标不同，三焦点张量有三个指标，所以再沿用这种记号模式将非常冗赘. 从现在开始，我们要转到使用标准的张量符号去处理三焦点张量. 附录 1($p$400)为不熟悉张量记号的读者提供一个简要介绍，在学习本章之前应先阅读该附录.

图像上的点和直线分别用齐次的**列**和**行** 3 维矢量表示，即 $\mathbf{x}=(x^1,x^2,x^3)^{\mathrm{T}}$ 和 $\mathbf{l}=(l_1,l_2,l_3)$，矩阵 $A$ 的$(i,j)$元记为 $a^i_j$，其中指标 $i$ 是逆变(行)指标，$j$ 是协变(列)指标. 我们遵守惯例：在逆变和协变位置上重复出现的指标表示对该指标在范围(1,...,3)内求和. 例如方程 $\mathbf{x}'=\mathrm{A}\mathbf{x}$ 等价于 $x'^i=\sum_j a^i_j x^j$，可记为 $x'^i=a^i_j x^j$.

我们从(14.1－$p$260)给出的三焦点张量的定义出发，利用张量记号，它就变为

$$\mathcal{T}_i^{jk}=a^j_i b^k_4-a^j_4 b^k_i. \tag{14.9}$$

$\mathcal{T}_i^{jk}$ 中指标的位置(两个逆变，一个协变)由该方程右边表达式中指标的位置确定. 这样一来，三焦点张量是一个逆变—协变混合张量. 利用张量记号，基本关联关系(14.3－$p$261)变为

$$l_i=l'_j l''_k \mathcal{T}_i^{jk}. \tag{14.10}$$

注意：与标准的矩阵记号不一样，乘张量时元素的顺序并不紧要. 例如上式的右边可写为：

$$l'_j l''_k \mathcal{T}_i^{jk}=\sum_{j,k} l'_j l''_k \mathcal{T}_i^{jk}=\sum_{j,k} l'_j \mathcal{T}_i^{jk} l''_k=l'_j \mathcal{T}_i^{jk} l''_k.$$

图 14.2 和图 14.3 的单应映射现在可以由关联关系(14.10)推得. 当平面由直线 $\mathbf{l}'$ 反向投影定义时

$$l_i=l'_j l''_k \mathcal{T}_i^{jk}=l''_k(l'_j \mathcal{T}_i^{jk})=l''_k h^k_i,\text{其中 } h^k_i=l'_j \mathcal{T}_i^{jk}$$

$h^k_i$ 是单应矩阵 H 的元素. 这个单应把第一幅视图和第三幅视图之间的点映射为

$$x''^k=h^k_i x^i.$$

注意：这个单应由该张量被一条直线缩并得到，即对张量的一个逆变(上)指标和该直线的协变(下)指标求和从该张量中提取一个 3×3 矩阵——把张量看成是一个算子，它用一条直线产生一个单应矩阵. 表 14.2 概括了三焦点张量的定义和转移性质.

A1.1($p$400)中定义了一对特别重要的张量：$\varepsilon_{ijk}$ 和其逆变对应式 $\varepsilon^{ijk}$. 这个张量用来表示矢量积. 例如，连接两点 $x^i$，$y^j$ 的直线为叉积 $x^i y^j \varepsilon_{ijk}=l_k$，反对称矩阵$[\mathbf{x}]_\times$ 用张量记号写为 $x^i \varepsilon_{ijs}$. 现在可相对简洁地写出表 14.1 给出的三焦点张量的基本的关联的结论，结论概括在表 14.3 中，符号 $\mathbf{0}_r$ 表示零矢量.

如果注意到 $\mathcal{T}_i^{jk}$ 的三个指标 $i,j,k$ 分别对应着第一、第二和第三幅视图，那么表 14.3 各关系的形式就更容易理解了. 因此，某些表达式诸如 $l''_j \mathcal{T}_i^{jk}$ 不可能出现，因为指标 $j$ 属于第二幅视图，从而它不可能属于第三幅视图上的直线 $l''$. 重复的指标(表示求和)必须作为逆变(上)指标和协变(下)指标各出现一次. 因此我们不能写 $x'^j \mathcal{T}_i^{jk}$，因为指标 $j$ 在上面位

置出现两次. 可以考虑用张量 $\epsilon$ 提升或下降指标,例如用 $x^i\epsilon_{ijk}$ 代替 $l'_j$. 然而,练习4($p$275)指出,我们并不能任意这样做.

**定义** 三焦点张量 $\mathcal{T}$ 是一个阶为3的张量 $\mathcal{T}_i^{jk}$,有两个逆变和一个协变指标. 它用一个齐次的 $3\times3\times3$ 阵列(即27个元素)来表示,其自由度为18.

**由摄像机矩阵来计算** 如果标准的 $3\times4$ 摄像机矩阵是

$$\mathrm{P}=[\mathrm{I}\mid\mathbf{0}],\quad \mathrm{P}'=[a_j^i],\quad \mathrm{P}''=[b_j^i]$$

那么

$$\mathcal{T}_i^{jk}=a_i^jb_4^k-a_4^jb_i^k$$

关于由一般摄像机矩阵的计算请参看(16.12 - $p$294).

**由第二和第三视图的对应直线到第一视图的直线转移**

$$l_i=l'_jl''_k\mathcal{T}_i^{jk}$$

**由一个单应的转移.**

(1) **通过第二幅视图的一条直线从第一幅视图向第三幅视图的点转移**

缩并 $l'_j\mathcal{T}_i^{jk}$ 是通过由第二幅视图中直线 $\mathbf{l}'$ 反向投影所确定的平面来诱导的第一和第三幅视图之间的一个单应映射:

$$x''^k=h_i^kx^i \text{ 其中 } h_i^k=l'_j\mathcal{T}_i^{jk}$$

(2) **通过第三幅视图的一条直线从第一幅视图向第二幅视图的点转移**

缩并 $l''_k\mathcal{T}_i^{jk}$ 是通过由第三幅视图中直线 $\mathbf{l}''$ 反向投影所确定的平面来诱导的第一和第二幅视图之间的一个单应映射:

$$x'^j=h_i^jx^i,\text{其中 } h_i^j=l''_k\mathcal{T}_i^{jk}$$

表14.2 三焦张量的定义和转移性质.

### 14.2.1 三重线性

表14.3中的关联关系都是图像元素(点和线)坐标的**三线性**关系或**三重线性**. '**三**'是指该关系中每个单项都包含了每幅的元素坐标,'**线性**'是指对每个代数元素(即张量的三个"自变量")都是线性的. 例如在点一点一点关系中,$x^i(x'^j\epsilon_{jpr})(x''^k\epsilon_{kqs})\mathcal{T}_i^{pq}=0_{rs}$,设 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 都满足这个关系,那么 $\mathbf{x}=\alpha\mathbf{x}_1+\beta\mathbf{x}_2$ 也满足,即该关系对它的第一个自变量是线性的. 类似地,该关系对第二个和第三个自变量也是线性的. 这个多重线性是张量的一个标准性质,并且直接由该张量关于所有三个指标(自变量)的缩并形式 $x^il'_jl''_k\mathcal{T}_i^{jk}=0$ 推出.

(1) 线—线—线对应

$$(l_r\epsilon^{rs})l'_jl''_k\mathcal{T}_i^{jk}=\mathbf{0}^s$$

(2)点—线—线对应

$$x^il'_jl''_k\mathcal{T}_i^{jk}=0$$

(3)点—线—点对应

$$x^il'_j(x''^k\epsilon_{kqs})\mathcal{T}_i^{pq}=\mathbf{0}_s$$

(4)点—点—线对应

$$x^i(x'^j\epsilon_{jpr})l''_k\mathcal{T}_i^{pk}=\mathbf{0}_r$$

(5)点—点—点对应

$$x^i(x'^j\epsilon_{jpr})(x''^k\epsilon_{kqs})\mathcal{T}_i^{pq}=0_{rs}$$

表14.3 三焦点张量关联关系的概括——三线性等式.

现在我们来详细地介绍点—点—点三线性等式. 由 $r$ 和 $s$ 的三种选择可得到 9 个这样的三线性关系. 从几何上看,这些三线性关系是由点—线—线关系在第二幅和第三幅图像上选择特殊直线而得到(见图 14.4(a)). 选择 $r=1,2,3$ 分别对应于图像上平行于 $x$ - 轴、$y$ - 轴和通过图像原点(点$(0,0,1)^{\mathsf{T}}$)的直线. 例如,选择 $r=1$,并展开 $x'^j\varepsilon_{jpr}$ 得

$$l'_p = x'^j\varepsilon_{jp1} = (0, -x'^3, x'^2)$$

它是第二幅视图上过点 $\mathbf{x}'$ 的一条水平直线(因为对任意 $\lambda$,形如 $\mathbf{y}' = (x'^1+\lambda, x'^2, x'^3)^{\mathsf{T}}$ 的点都满足 $\mathbf{y}'^{\mathsf{T}}\mathbf{l}'=0$). 在第三幅图像上选择 $s=2$ 类似地得到过 $\mathbf{x}''$ 的铅垂线

$$l''_q = x''^k\varepsilon_{kq2} = (x''^3, 0, -x''^1)$$

从而该点的三线性关系展开为

$$\begin{aligned} 0 &= x^i x'^j x''^k \varepsilon_{jp1}\varepsilon_{kq2}\mathcal{T}_i^{pq} \\ &= x^i[-x'^3(x''^3\mathcal{T}_i^{21} - x''^1\mathcal{T}_i^{23}) + x'^2(x''^3\mathcal{T}_i^{31} - x''^1\mathcal{T}_i^{33})]. \end{aligned}$$

在这 9 个三重线性等式中四个是独立的. 即所有 9 个三线性等式可由 4 个三线性等式构成的一组基通过线性组合得到. 四个自由度可以回溯到点—线—线关系 $x^i l'_j l''_k \mathcal{T}_i^{jk}=0$,并计算如下:在第三幅视图中有过 $\mathbf{x}''$ 的单参数直线簇,如果 $\mathbf{m}''$ 和 $\mathbf{n}''$ 是该直线簇中两个元素,则任一条过 $\mathbf{x}''$ 的其他直线都可以由它们的线性组合得到

$$\mathbf{l}'' = \alpha\mathbf{m}'' + \beta\mathbf{n}''$$

因该关联关系关于 $\mathbf{l}''$ 是线性的,故又给出

$$\begin{aligned} l'_j m''_k \mathcal{T}_i^{jk} x^i &= 0 \\ l'_j n''_k \mathcal{T}_i^{jk} x^i &= 0 \end{aligned}$$

于是对其他任意直线 $\mathbf{l}''$ 的关联关系可以由这两式的线性组合得到. 所以,对 $\mathbf{l}''$ 来说仅有两个独立的关联关系. 类似地,过 $\mathbf{x}'$ 有一个单参数的直线簇,并且该关联关系关于过 $\mathbf{x}'$ 的直线也是线性的. 因此在第一幅视图上的点和第二、第三幅视图上的直线之间总共有四个独立的关联关系.

这些三线性等式的主要优点在于它们是线性的,另外,如下一节中所述的,它们的性质通常包含在转移中.

## 14.3 转移

给定一个景物的三幅视图和其中两幅上的一对匹配点,我们希望确定该点在第三幅视图上的位置. 给定足够的关于摄像机位置的信息,通常可以在不涉及图像内容的条件下确定该点在第三幅视图上的位置. 这就是所谓的点转移问题,对直线也有类似的转移问题.

这个问题在给定三幅视图的摄像机矩阵后原则上可以解决. 由第一幅和第二幅图像上的对应点反向投影的射线相交,从而确定了这个 3D 点. 通过把这个 3D 点投影到第三幅图像可以计算出在该视图上对应点的位置. 类似地,由第一幅和第二幅图像上对应直线反向投影的平面相交于 3D 直线,把这条 3D 直线投影到第三幅图像可以确定其图像的位置.

### 14.3.1 利用基本矩阵的点转移

仅利用基本矩阵的知识就可以解决点转移问题. 为此,假设我们已经知道联系三幅视

图的三个基本矩阵 $F_{21}$，$F_{31}$，$F_{32}$，并设前两幅视图上的点 $\mathbf{x}$ 和 $\mathbf{x}'$ 是匹配点，我们希望在第三幅视图上求其对应点 $\mathbf{x}''$.

所求点 $\mathbf{x}''$ 与第一幅图像的点 $\mathbf{x}$ 匹配，因此它必在对应于 $\mathbf{x}$ 的对极线上. 由于我们已经知道 $F_{31}$，这条对极线可以算出并且等于 $F_{31}\mathbf{x}$. 利用类似的推理，$\mathbf{x}''$ 必定在对极线 $F_{32}\mathbf{x}'$ 上. 求这两条直线的交点(见图 14.7(a))得到

$$\mathbf{x}'' = (F_{31}\mathbf{x}) \times (F_{32}\mathbf{x}')$$

注意：这个表达式中并没有用到基本矩阵 $F_{21}$. 一个自然产生的问题是：我们能否由 $F_{21}$ 获得什么好处？答案是肯定的. 当存在噪声时，点对 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 不会是一对精确匹配点，即它们并不会正好满足方程 $\mathbf{x}'F_{21}\mathbf{x}=0$. 给定了 $F_{21}$，如同算法 11.1($p$222)，我们可以利用算法 11.1($p$222)的最优三角形法来矫正 $\mathbf{x}$ 和 $\mathbf{x}'$，得到精确满足该方程的一对点 $\hat{\mathbf{x}}\leftrightarrow\hat{\mathbf{x}}'$. 从而该转移点可以由 $\mathbf{x}''=(F_{32}\hat{\mathbf{x}})\times(F_{32}\hat{\mathbf{x}}')$ 来计算. 这种利用基本矩阵的点转移称为**对极转移**.

尽管对极变换曾经被用于点转移，但它有一个严重缺陷使得它没被实际计算所采用. 这个缺陷产生的退化情形可通过图 14.7(b)来理解：当两条极线在第三幅图像上重合时对极变换失效(直线越"接近"重合就越病态). 此退化条件为 $\mathbf{x}''$，$\mathbf{e}_{31}$ 和 $\mathbf{e}_{32}$ 在第三幅图像上共线，它意味着摄像机中心 $\mathbf{C}$，$\mathbf{C}'$ 和 3D 点 $\mathbf{X}$ 在过第三个摄像机中心 $\mathbf{C}''$ 的某平面上；从而 $\mathbf{X}$ 在三个摄像机中心确定的三焦点平面上(见图 14.8). 对极转移对三焦点平面上的点将失效，并且对邻近三焦点平面的点将很不精确. 注意，当这三个摄像机中心共线时，三焦点平面不唯一确定，同时对极变换对所有的点都失效，此时 $\mathbf{e}_{31}=\mathbf{e}_{32}$.

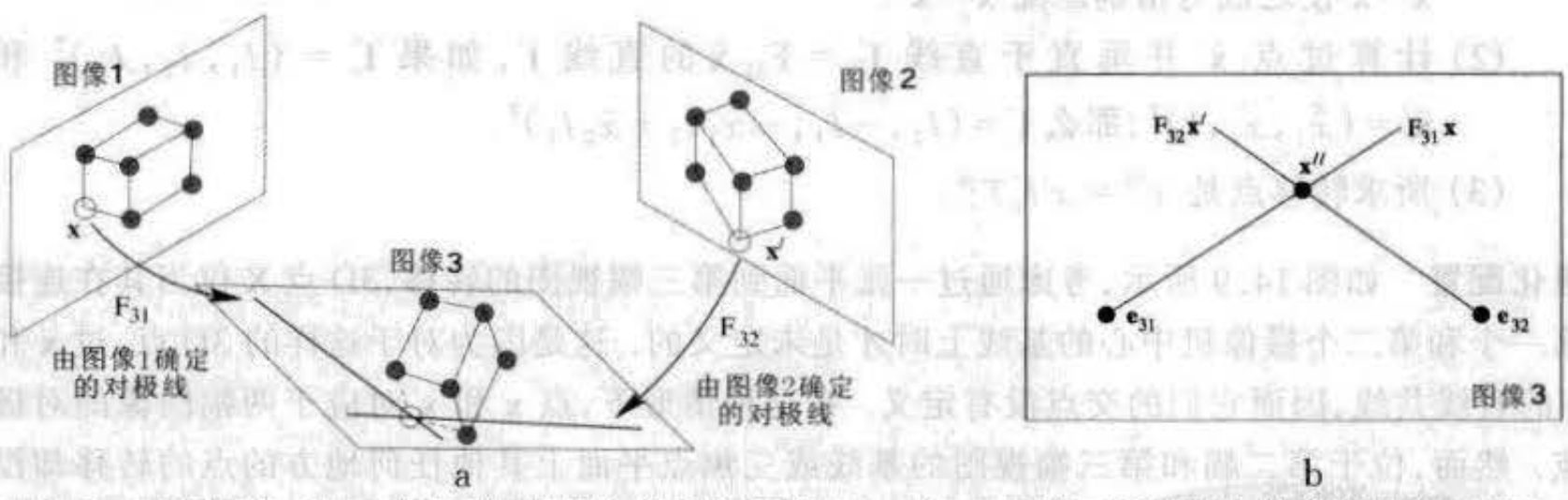



图 14.7 **对极转移**. (a)$\mathbf{X}$ 在前两幅视图上的影像是对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$. $\mathbf{X}$ 在第三幅视图上的影像是对极线 $F_{31}\mathbf{x}$ 与 $F_{32}\mathbf{x}'$ 的交点. (b)在第三幅图像上看到的对极点和转移点 $\mathbf{x}''$ 的配置. 点 $\mathbf{x}''$ 可由分别过两对极点 $\mathbf{e}_{31}$ 和 $\mathbf{e}_{32}$ 的对极线的交点确定. 但若 $\mathbf{x}''$ 在过这两个对极点的直线上，则其位置不能确定. 接近于该直线的点的估计将很不精确.

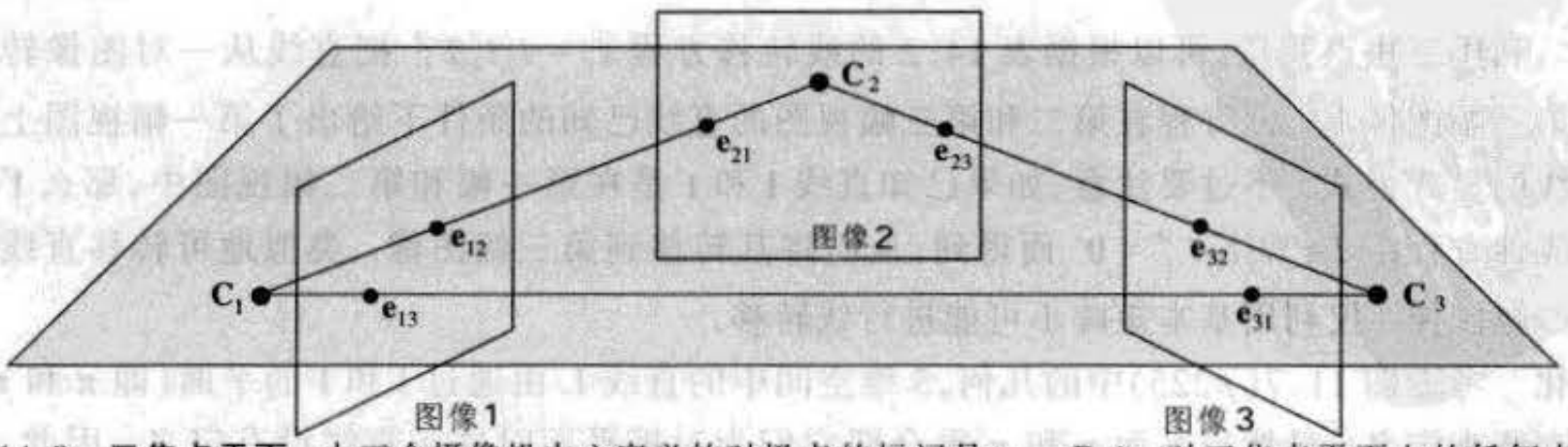



图 14.8 **三焦点平面**. 由三个摄像机中心定义的对极点的标记是 $\mathbf{e}_{ij}=P_i\mathbf{C}_j$. 对三焦点平面上的任何点 $\mathbf{X}$ 来说对极转移将无效. 如果三个摄像机的中心共线，那么存在包含三个中心的单参数簇平面.

### 14.3.2 利用三焦点张量的点转移

利用三焦点张量可以避免对极转移的退化. 考虑对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$,如果通过点 $\mathbf{x}'$ 的直线 $\mathbf{l}'$ 选在第二幅视图上,那么根据表 14.2,对应点 $\mathbf{x}''$ 可以由通过第一幅上的点 $\mathbf{x}$ 到第三幅上的转移 $x''^k = x^i l'_j \mathcal{T}_i^{jk}$ 得到. 由图 14.4(b)($p$263)显而易见该转移对于在三焦点平面上的一般点 $\mathbf{X}$ 不是退化的.

然而由结论 14.3($p$265)和图 14.6 指出:如果 $\mathbf{l}'$ 是对应 $\mathbf{x}$ 的对极线,那么 $x^i l'_j \mathcal{T}_i^{jk} = \mathbf{0}^k$,因此 $\mathbf{x}''$ 是没有定义的. 所以,直线 $\mathbf{l}'$ 的选择很重要. 为了避免只选到一条对极线,一种可能是利用通过 $\mathbf{x}'$ 的两条或三条不同的直线,即 $l'_{jp} = x'^r \varepsilon_{rjp}$,选取 $p=1,2,3$. 对每一条这种直线,计算出 $\mathbf{x}''$ 的值并保留其中最大范数(即最不可能是 0)的一个. 寻找 $\mathbf{x}''$ 的另一方法是求线性方程组 $x^i(x'^j\varepsilon_{jpr})(x''^k\varepsilon_{kqs})\mathcal{T}_i^{pq}=0_{rs}$ 的最小二乘解 $\mathbf{x}''$,但这种方法可能是杀鸡用了牛刀.

我们推荐的方法如下:在试图计算由点对 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 转移的点 $\mathbf{x}''$ 之前,首先按上面在对极转移中介绍的方法,利用基本矩阵 $\mathrm{F}_{21}$ 矫正该点对,如果 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$ 是精确匹配,那么转移点 $x''^k = \hat{x}^i l'_j \mathcal{T}_i^{jk}$ 将不依赖于过点 $\hat{\mathbf{x}}'$ 的直线 $\mathbf{l}'$ 的选择(只要它不是对极线). 这可以参考图 14.2($p$262)从几何上验证. 一个好的选择是总取垂直于 $\mathrm{F}_{21}\hat{\mathbf{x}}$ 的直线.

求测量对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 的转移方法概括为下列步骤:

(1) 由三焦点张量计算 $\mathrm{F}_{21}$(利用算法 14.1 所给的方法),并利用算法 11.1($p$222)矫正 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 使之成为精确匹配 $\hat{\mathbf{x}}\leftrightarrow\hat{\mathbf{x}}'$.

(2) 计算过点 $\hat{\mathbf{x}}'$ 并垂直于直线 $\mathbf{l}'_e = \mathrm{F}_{21}\hat{\mathbf{x}}$ 的直线 $\mathbf{l}'$,如果 $\mathbf{l}'_e = (l_1, l_2, l_3)^\mathsf{T}$ 和 $\hat{\mathbf{x}}' = (\hat{x}_1, \hat{x}_2, 1)^\mathsf{T}$,那么 $\mathbf{l}' = (l_2, -l_1, -\hat{x}_1 l_2 + \hat{x}_2 l_1)^\mathsf{T}$.

(3) 所求转移点是 $x''^k = \hat{x}^i l'_j \mathcal{T}_i^{jk}$.

**退化配置** 如图 14.9 所示,考虑通过一张平面到第三幅视图的转移,3D 点 $\mathbf{X}$ 仅当其在连接第一个和第二个摄像机中心的基线上时才是未定义的. 这是因为对于这样的 3D 点,过 $\mathbf{x}$ 和 $\mathbf{x}'$ 的直线共线,因而它们的交点没有定义. 在这种情形下,点 $\mathbf{x}$ 和 $\mathbf{x}'$ 对应于两幅图像的对极点. 然而,位于第二幅和第三幅视图的基线或三焦点平面上其他任何地方的点的转移却没有问题. 这是对极转移和三焦点张量的转移之间的关键性差别. 前者对三焦点平面上的任何点没有定义.

### 14.3.3 利用三焦点张量的线转移

利用三焦点张量,可以根据表 14.2 的线转移方程 $l_i = l'_j l''_k \mathcal{T}_i^{jk}$ 把直线从一对图像转移到第三幅图像上. 该方程在第二和第三幅视图的直线已知的条件下给出了第一幅视图上的直线的显式公式. 不过要注意,如果已知直线 $\mathbf{l}$ 和 $\mathbf{l}'$ 是在第一幅和第二幅视图中,那么 $\mathbf{l}''$ 由解线性方程组 $l_r\varepsilon^{rjs}l'_j l''_k \mathcal{T}_i^{jk} = \mathbf{0}^s$ 而得到,从而将其转移到第三幅图像. 类似地可转移直线到第二幅图像. 仅利用基本矩阵不可能进行线转移.

**退化** 考虑图 11.7($p$225)中的几何,3 维空间中的直线 $\mathbf{L}$ 由通过 $\mathbf{l}$ 和 $\mathbf{l}'$ 的平面(即 $\boldsymbol{\pi}$ 和 $\boldsymbol{\pi}'$)的交线来定义. 显然当平面 $\boldsymbol{\pi}$ 和 $\boldsymbol{\pi}'$ 重合即它们为对极平面时这条直线没有定义. 因此,如果 $\mathbf{l}$ 和 $\mathbf{l}'$ 对应于第一幅和第二幅视图的对极线,那么不能在第一幅和第三幅之间进行线变

换. 从代数上看,线转移方程给出 $l_i = l'_j l''_k T_i^{jk} = 0$,用来解 $\mathbf{l}''$ 的方程的矩阵 $(l_r \varepsilon^{rts}) l'_j T_i^{jk}$ 变为零. 直线接近于对极线的情形时有发生,并且它们的转移不准确,因此要经常检测这个条件. 存在由第三幅视图上直线定义的第一幅和第二幅之间的线转移的一个等价退化. 当第一幅和第三幅视图上的直线对应于这两幅视图的对极线时,会再次发生这种退化情形.

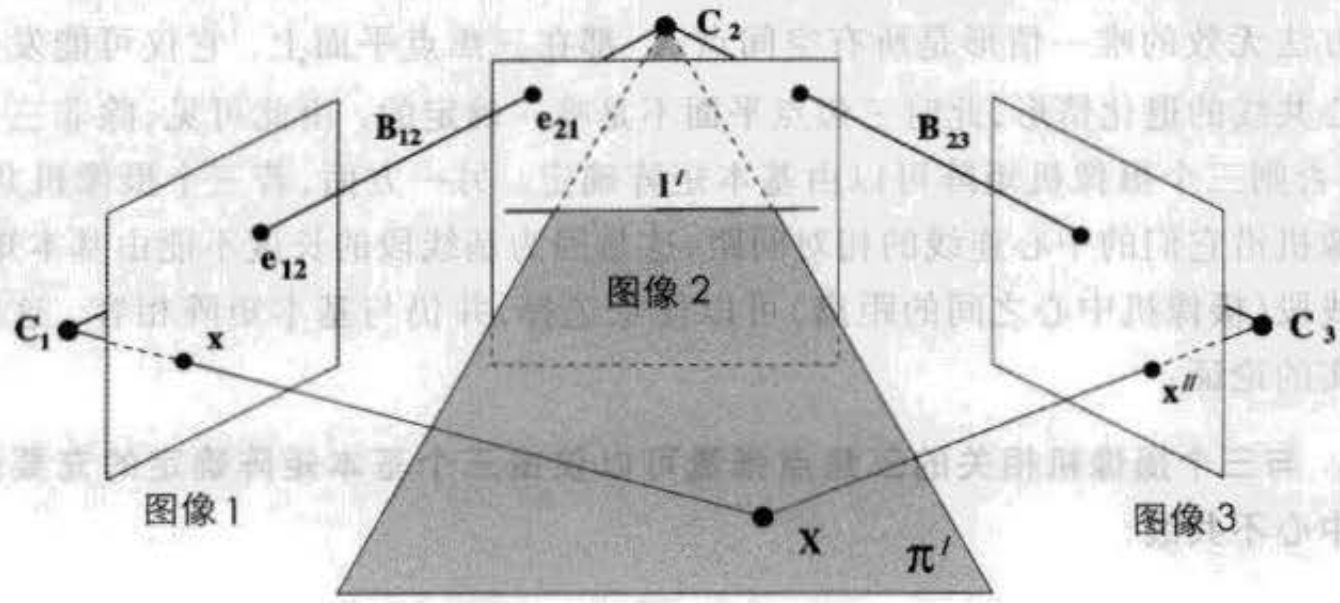



图 14.9 **用三焦张量计算点转移的退化性**. 3D 点 $\mathbf{X}$ 由过 $\mathbf{x}$ 的射线与平面 $\pi'$ 的交点来确定. 在第一与第二视图之间的基线段 $B_{12}$ 上的点 $\mathbf{X}$ 不能用这种方式确定. 因此,在线段 $B_{12}$ 上的 3D 点不能由第二幅视图上的一条直线定义的单应转移到第三幅视图. 注意在直线 $B_{12}$ 上的点投影到第一幅视图的 $\mathbf{e}_{12}$ 和第二幅视图的 $\mathbf{e}_{21}$. 除直线 $B_{12}$ 之外的任何点都可以转移. 特别是对介于视图二和三之间的基线段 $B_{23}$ 上的点或三焦平面上其他任何点都不存在退化问题.

一般来说,第一幅和第二幅之间的对极几何与第一幅和第三幅之间的对极几何会有所不同. 例如,由第二幅视图引起的在第一幅的极点 $\mathbf{e}_{12}$ 与由第三幅视图引起的第一幅的极点 $\mathbf{e}_{13}$ 会不一样. 因此在第一幅视图中对应于第一和第二两幅视图的对极线与对应于第一和第三两幅视图的对极线不同. 从而当直线变换到第三幅视图是退化时,变换到第二幅视图一般不会是退化的. 然而,对于三焦点平面上的直线,转移总是退化(即未定义)的.

## 14.4 基本矩阵与三焦点张量的关系

因为三个摄像机的射影结构可以显式地由三焦点张量计算出来,所以对于三组视图对的所有三对基本矩阵都可以由该三焦点张量确定. 事实上,算法 14.1 给出了两个基本矩阵 $F_{21}$ 和 $F_{31}$ 的简单的公式.

问题是能否由给定的三个基本矩阵来求出三焦点张量. 我们将证明,大多数情况下答案是肯定的. 首先注意到三焦点张量由三个摄像机矩阵确定. 事实上,摄像机矩阵只确定到相差一个射影变换就足够了. 前两个摄像机矩阵 P 和 P′ 可以用关于两视图的技术(第 8 章)由基本矩阵来确定,余下的是在相同的射影坐标系下确定第三个摄像机矩阵 P″. 原则上可按下列步骤进行:

(1) 在前两幅图像中选择一组匹配 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$,满足 $\mathbf{x}'^{\mathrm{T}}_i F_{21} \mathbf{x}_i = 0$,并利用三角形法确定对应的 3D 点 $\mathbf{X}_i$.

(2) 利用基本矩阵 $F_{31}$,$F_{32}$,由对极转移确定第三幅图像上的对应点 $\mathbf{x}''$.

(3) 由这组 3D—2D 对应 $\mathbf{X}_i \leftrightarrow \mathbf{x}''_i$，求解摄像机矩阵 $\mathrm{P}''$.

当点 $\mathbf{X}_i$ 在三焦点平面上时，这个算法的第二步将失效. 这样的点 $\mathbf{X}_i$ 很容易被发现并舍去，因为它投到第一幅图像上两个对极点 $\mathbf{e}_{12}$ 和 $\mathbf{e}_{13}$ 的连线上的点 $\mathbf{x}_i$. 因为有无限多个可能的匹配点，我们可以计算足够多的点以求得 $\mathrm{P}''$.

这种方法无效的唯一情形是所有空间点 $\mathbf{X}_i$ 都在三焦点平面上. 它仅可能发生在三个摄像机中心共线的退化情形，此时三焦点平面不是唯一确定的. 由此可见，除非三个摄像机中心共线，否则三个摄像机矩阵可以由基本矩阵确定. 另一方面，若三个摄像机共线，则无法确定摄像机沿它们的中心连线的相对间距，这是因为基线段的长度不能由基本矩阵确定. 这三条基线段(摄像机中心之间的距离)可以任意选择，并仍与基本矩阵相容. 这样就完成了下述事实的论证：

**结论 14.5　与三个摄像机相关的三焦点张量可以仅由三个基本矩阵确定的充要条件是三个摄像机中心不共线.**

## 14.5　结束语

三视图几何的推导与本书第二篇中的两视图几何相类似. 三焦点张量可以由三幅视图的图像对应计算得到，随之可以得到摄像机和 3D 景物的射影重构. 三焦点张量的计算在第 15 章中讲述. 采用与第 9 章中同样的方式，通过提供景物或摄像机的附加信息，射影多义性可以化为仿射或度量多义性. 用与第 12 章中类似的推导，可以给出由景物平面和三焦点张量诱导的单应之间的关系.

### 14.5.1　文献

三焦点张量的发现可以追溯到[Spetsakis－91]和[Weng－88]，它们在摄像机已标定的情况下，用三焦点张量由直线对景物进行重构. 后来[Hartley－94c]证明它同样可以用于未标定情形下的景物的射影重构. 在此之前都使用矩阵记号，然而[Viéville－93]对此问题采用了张量符号.

与此同时，Shashua 独立地引入了关于未标定摄像机的三幅视图中的对应点坐标之间的三线性条件[Shashua－94，Shashua－95a]. 其后[Hartley－95a，Hartley－97c]证明了：由直线求点和景物重构的 Shashua 关系式来自共同的张量，因此该三焦点张量被正式确认.

在随后的工作中，张量的性质被[Shashua－95b]等进行了研究. 特别是[Triggs－95]介绍了指标的混合协变—逆变性质，[Zisserman－96]描述了由张量表示的单应的几何性质. Faugeras 和 Mourrain[Faugeras－95c]创造性地给出了三焦点张量方程新的推导，并在涉及到多视图的一般线性约束的范畴内考虑了三焦点张量. 这方面内容将在第 16 章中讨论. 张量的更多的几何性质由 Faugeras 和 Papadopoulo[Faugeras－97]给出.

对极点转移在 [Barrett－92，Faugeras－94]中给出了介绍，[Zisserman]等指出了它的不足之处.

三焦点张量已经被广泛应用于各个应用领域，包括建立图像序列中对应[Beardsley－96]、独立运动检测[Torr－95b]和摄像机自标定[Armstrong－96a]等.

### 14.5.2 注释和练习

(1) 三焦点张量在 3D 射影变换下不变. 显式地证明:如果 $H_{4\times4}$ 是使第一个摄像机矩阵变为 $P=[I|\mathbf{0}]$ 的变换,那么由(14.1 - $p$260)定义的张量不变.

(2) 本章中三焦点张量的推导是由三条对应直线的关联性质开始的. 证明:另一个出发点可以是由平面诱导的单应.

推理要点:选定标准的摄像机组 $P=[I|\mathbf{0}]$,$P'=[A|\mathbf{a}_4]$,$P''=[B|\mathbf{b}_4]$,并由被平面 $\boldsymbol{\pi}'$ 诱导的第一幅和第三幅图像之间的单应 $H_{13}$ 出发. 根据结论 12.1($p$228),这个单应矩阵可以写成 $H_{13}=B-\mathbf{b}_4\mathbf{v}^{\mathsf{T}}$,其中 $\boldsymbol{\pi}'^{\mathsf{T}}=(\mathbf{v}^{\mathsf{T}},1)$. 此时由第二幅视图上的直线 $\mathbf{l}'$ 定义的平面为 $\boldsymbol{\pi}'=P'^{\mathsf{T}}\mathbf{l}'$. 然后证明结论 14.2($p$261).

(3) 第一幅视图的单应可以用三焦点张量 $\mathcal{T}_i^{jk}$ 表示成结论 14.2($p$261)给出的简单形式. 试研讨由第一幅图像上的直线 $\mathbf{l}$ 诱导的第二幅到第三幅视图的单应 $H_{23}$ 是否有简单的公式.

(4) 缩并 $x^i\mathcal{T}_i^{jk}$ 是一个 $3\times3$ 矩阵. 证明它可以被解释为由第一幅视图上的点 $\mathbf{x}$ 反投影的直线诱导的第二幅和第三幅视图之间的对射(见定义 1.28($p$29)).

(5) 当三个摄像机有两个中心相同时,三焦点张量可以有更简单的形式. 这里有两种情形:

(a) 如果第二个和第三个摄像机有相同的中心,那么 $\mathcal{T}_i^{jk}=F_{ri}H_s^k\epsilon^{rjs}$,其中 $F_{ri}$ 是前两幅视图的基本矩阵,H 是第二幅到第三幅视图由其中心相同这一事实诱导出的单应.

(b) 如果第一幅和第二幅视图有相同的中心,那么 $\mathcal{T}_i^{jk}=H_i^j\mathbf{e}''^k$,其中 H 是第一幅到第二幅视图的单应,$\mathbf{e}''$ 是第三幅图像的对极点.

利用第 16 章中的方法证明这些关系.

(6) 考虑摄像机间短基线的情形并推导该三焦点张量的一种不同形式,见[Aström - 98, Triggs - 99b].

(7) 三幅视图其实有三个不同的三焦点张量,它们取决于三个摄像机中的哪一个对应于协变指标. 给定一个这样的张量 $[T_i]$,证明张量 $[T_i']$ 可以通过下列几步算得:

(a) 由三焦点张量求出三个摄像机矩阵 $P=[I|\mathbf{0}]$,$P'$ 和 $P''$.

(b) 求出使得 $P'H=[I|\mathbf{0}]$ 的 3D 射影变换 H,并同样应用于 $P'$ 和 $P''$.

(c) 用(14.1 - $p$260)计算张量 $[T_i']$.

(8) 考虑在节 8.3($p$166)中(对基本矩阵)所描述的特殊运动(纯旋转、平面运动)下三焦点张量的形式和性质(如矩阵 $T_i$ 的秩).

(9) 比较表 14.3 中的关联关系式说明我们可以用表达式 $\epsilon_{jrs}x_r'$ 来替换直线 $l_j'$,并可类似地用于 $l_k''$. 另外,用 $\epsilon^{irs}l_i$ 来替换 $x^i$ 我们也可以得到一个三视图直线方程. 这两种运算是否可以同时进行而得到下面的方程

$$(\epsilon^{irs}l_i)(\epsilon_{jsu}x'^j)(\epsilon_{ktw}x''^k)\mathcal{T}_r^{st}=0_{vrw}^u\,?$$

为什么行或为什么不行?

(10) **仿射三焦点张量** 如果三个摄像机 P,$P'$ 和 $P''$ 都是仿射(定义 5.3($p$111))的,那

么对应的张量 $\mathcal{T}_A$ 称为**仿射三焦点张量**. 张量的这种仿射特殊化形式有 12 个自由度和 16 个非零元素,它首先由[Torr－95a]定义.

(11) **基本矩阵和对极点之间的关系** 三个基本矩阵 $F_{12}, F_{23}, F_{31}$ 不是独立的,而要满足三个(二次)关系式:

$$F_{12}\mathbf{e}_{13} = \mathbf{e}_{23} \times \mathbf{e}_{21} \quad F_{23}\mathbf{e}_{21} = \mathbf{e}_{31} \times \mathbf{e}_{32} \quad F_{31}\mathbf{e}_{32} = \mathbf{e}_{12} \times \mathbf{e}_{13}.$$

这可以由图 14.8 的三焦点平面的几何推导出,例如第一个关系式表示第一幅图像的对极点 $\mathbf{e}_{13}$(它是第一幅和第三幅图像中心之间的基线的像)在第二幅图像上的对极线是三焦点平面与第二幅视图的交线 $\mathbf{e}_{23} \times \mathbf{e}_{21}$. 三焦点张量有 18 个自由度,这可以由从基本矩阵的 $3 \times 7 = 21$ 个自由度中减去二次关系的三个自由度来计算. 由张量计算的基本矩阵将自动满足这些关系.

# 第 15 章　三焦点张量 $\mathcal{T}$ 的计算

本章介绍由给定的三幅视图的点对应和线对应的一个集合来估计三焦点张量的数值方法. 其推导过程非常类似于对基本矩阵的计算,并且使用与第 10 章非常相同的技术. 我们将具体讨论 5 种方法:

(1) 基于直接解线性方程组的线性算法(经过适当的数据归一化)(节 15.2).

(2) 在满足关于该张量的所有适当约束下最小化代数误差的迭代算法(节 15.3).

(3) 最小化几何误差的迭代算法("黄金标准"算法)(节 15.4.1).

(4) 最小化几何误差的 Sampson 近似的迭代算法(节 15.4.3).

(5) 基于 RANSAC 的鲁棒估计(节 15.6).

## 15.1　基本方程组

三焦点张量的(三重)线性方程组的一个完整集合在表 15.1 给出,所有这些方程关于三焦点张量 $\mathcal{T}$ 的元素都是线性的.

| 对应 | 关系 | 方程数 |
|---|---|---|
| 三点 | $x^i x'^j x''^k \varepsilon_{jps} \varepsilon_{krt} \mathcal{T}_i^{qr} = 0_{st}$ | 4 |
| 两点、一直线 | $x^i x'^j l''_r \varepsilon_{jps} \mathcal{T}_i^{qr} = 0_s$ | 2 |
| 一点、两直线 | $x^i l'_q l''_r \mathcal{T}_i^{qr} = 0$ | 1 |
| 三直线 | $l_p l'_q l''_r \varepsilon^{piw} \mathcal{T}_i^{qr} = 0^w$ | 2 |

表 15.1　在三视图中的点和直线坐标之间的三线性关系. 最后一列表示线性独立的方程数. 记号 $0_{st}$ 表示全部元素为零的 2 维张量. 因此本表的第一行对应于 9 个方程组,每个方程对应于 $s$ 和 $t$ 的每一种选择. 但是,在 9 个方程中,仅有 4 个是线性独立的.

用给定的三幅图像之间的若干点对应或线对应可产生的方程组的形式为 $\mathbf{At}=\mathbf{0}$,其中 $\mathbf{t}$ 是由三焦点张量元素构成的 27 维向量,由这些方程可以解出三焦点张量的元素. 注意:包含点的方程可以与包含线的方程放在一起——一般情况下表 15.1 中所有现成的方程都可以同时使用. 因为 $\mathcal{T}$ 有 27 个元素,所以在相差一个尺度因子下解 $\mathbf{t}$ 需要 26 个方程. 当多于 26 个方程时,则求最小二乘解,如求基本矩阵一样,我们利用算法 A3.5($p$413),在满足约束 $\|\mathbf{t}\|=1$ 下最小化 $\|\mathbf{At}\|$.

以上给出了计算三焦点张量的线性算法的一个简要概述. 然而为了由此建立一个实用的算法,还需要注意其它的一些问题,例如归一化. 具体地说,张量估计必须满足各种约束,我们将在下文考虑这些问题.

### 15.1.1　内在约束

三焦点张量与基本矩阵最显著的区别是三焦点张量的约束数目较大. 基本矩阵仅有一

个约束，即 $\det(F)=0$，不计尺度因子，它有 7 个自由度．而三焦点张量有 27 个元素，在相差一个射影变换意义下仅需 18 个参数就可确定等价的摄像机配置．因此三焦点张量的元素须满足 8 个独立的代数约束．这个条件通常叙述如下：

**定义 15.1** 一个三焦点张量 $\mathcal{T}_i^{jk}$ 称为"几何有效"或"满足所有内在约束"，如果存在三个摄像机矩阵 $P=[I|\mathbf{0}]$，$P'$ 和 $P''$，使得 $\mathcal{T}_i^{jk}$ 按照(14.9－$p$268)对应于这三个摄像机矩阵．

如基本矩阵一样，重要的是以某种方式强置这些约束而使所考虑的三焦点张量为几何有效三焦点张量．如果该张量不满足这些约束，便会出现类似于基本矩阵的秩不为 2 的后果——对于不同的 $\mathbf{x}$，由 $F\mathbf{x}$ 计算的对极线不交于同一点(见图 10.1($p$192))．例如，如果该张量不满足这些内在约束并且如 14.3 节所介绍的把它用来在给定两幅视图的对应下转移一个点到第三幅视图，那么转移点的位置将随使用表 15.1 中的方程组的不同而不同．因此我们的目标始终是估计几何有效的三焦点张量．

三焦点张量元素满足的约束不能简单地显式表示(如 $\det=0$)，有些人曾认为这是精确计算三焦点张量的一个障碍．然而，实际上，为了处理和计算三焦点张量，并不需要显式地表示这些约束——而是通过三焦点张量的适当参数化来隐式地强置它们，并且很少发生麻烦．我们将在 15.3 节和 15.4.2 节中回到这个参数化问题．

### 15.1.2 最小配置——6 点对应

一个几何有效的三焦点张量可以由 6 点配置的图像来算得，只要这些景物点处在一般位置上即可．它有一个或三个实解．该张量由节 19.2.4($p$360)中介绍的算法 19.1($p$361)获得的三个摄像机矩阵来计算．这种最小配置 6 点解将用于节 15.6 的鲁棒算法．

## 15.2 归一化线性算法

用表 15.1 中关于 $\mathcal{T}$ 的方程组来构造矩阵方程 $A\mathbf{t}=\mathbf{0}$ 时，不必使用从每组对应推出的所有方程，因为不是所有这些方程都是线性无关的．例如在点—点—点对应情形(表 15.1 中的第一行)，$s$ 和 $t$ 的所有选择导致 9 个方程，但这些方程中仅有 4 个是线性无关的，这 4 个方程可以由 $s$ 和 $t$ 各选 2 个值(例如 1 和 2)而得到．节 16.7($p$306)将对此作更详细地讨论．

读者可以验证对于 $s$ 和 $t$ 的一个给定值，由表 15.1 得到的三点方程可以展开为

$$x^k(x'^i x''^m \mathcal{T}_k^{jl} - x'^j x''^m \mathcal{T}_k^{il} - x'^i x''^l \mathcal{T}_k^{jm} + x'^j x''^l \mathcal{T}_k^{im}) = 0^{ijlm}. \tag{15.1}$$

其中 $i,j\neq s$ 且 $l,m\neq t$．方程(15.1)在 $i=j$ 或 $l=m$ 时变为 $0=0$，故不起作用，并且对换 $i$ 和 $j$(或 $l$ 和 $m$)仅改变方程的符号．令 $j=m=3$ 且让 $i$ 和 $l$ 自由变动，可以得到 4 个独立方程组中的一种选取法．可以令坐标 $x^3$，$x'^3$ 和 $x''^3$ 为 1 以获得被观测的图像坐标之间的关系．则方程(15.1)变为

$$x^k(x'^i x''^l \mathcal{T}_k^{33} - x''^l \mathcal{T}_k^{i3} - x'^i \mathcal{T}_k^{3l} + \mathcal{T}_k^{il}) = 0. \tag{15.2}$$

$i,l=1,2$ 的四种不同选择给出被观测的图像坐标的 4 个不同的方程．

**如何表示直线**

表 15.1 中三条直线对应的方程可以写为

$$l_i = l'_j l''_k \mathcal{T}_i^{jk}$$

与通常的齐次元素一样，这里的等式可以相差一个尺度因子．当噪声存在时，这个关系仅被测量直线 $\mathbf{l}, \mathbf{l}'$ 和 $\mathbf{l}''$ 近似地满足，但被靠近这三条测量直线的三条直线 $\hat{\mathbf{l}}, \hat{\mathbf{l}}'$ 和 $\hat{\mathbf{l}}''$ 精确地满足．

现在的问题是有微小数量差别的两个齐次坐标集是否描述某种几何意义下互相靠近的直线．考虑两个向量 $\mathbf{l}_1 = (0.01, 0, 1)^{\mathrm{T}}$ 和 $\mathbf{l}_2 = (0, 0.01, 1)^{\mathrm{T}}$，显然作为矢量，它们没有多大区别，事实上 $\| \mathbf{l}_1 - \mathbf{l}_2 \|$ 很小．但是 $\mathbf{l}_1$ 表示直线 $x = 100$，$\mathbf{l}_2$ 表示直线 $y = 100$，因此在几何意义下，这两直线完全不同．注意，这个问题可以通过尺度缩放得到缓解．如果坐标用因子 0.01 进行缩放，那么直线坐标变为 $\mathbf{l}_1 = (1, 0, 1)^{\mathrm{T}}$，$\mathbf{l}_2 = (0, 1, 1)^{\mathrm{T}}$ 显然它们很不相同．

上面的观察表明在表示直线时需要谨慎．设已经给定三条直线 $\mathbf{l}, \mathbf{l}'$ 和 $\mathbf{l}''$ 的一组对应，并选定 $\mathbf{l}$ 上两点 $\mathbf{x}_1$ 和 $\mathbf{x}_2$．这两点中的每点给出三幅视图的一个对应 $\mathbf{x}_s \leftrightarrow \mathbf{l}' \leftrightarrow \mathbf{l}''$，$s = 1, 2$，也就是存在一条 3D 直线映到第二幅和第三幅图像上的 $\mathbf{l}'$ 和 $\mathbf{l}''$，并且映到第一幅图像上通过 $\mathbf{x}_s$ 的一条直线（即 $\mathbf{l}$）．从这两个对应得到形如 $x_s^i l_j' l_k'' \mathcal{T}_i^{jk} = 0_s$，$s = 1, 2$ 的两个方程，用此办法避免了在第一幅图像上使用直线，尽管不能在其它图像上都避免．图像上的直线经常很自然地由一对点，可能是线段的两个端点来定义．如节 15.7.2 将要介绍的，甚至可以把图像上由一组边缘点的最好拟合所定义的直线当作由两点定义的直线．

**归一化**

如这种类型的所有算法一样，在建立和求解线性方程组之前有必要对输入数据预先进行归一化．然后必须在归一化基础上进行矫正以找到原数据的三焦点张量．值得推荐的归一化方法与基本矩阵的计算中所采用方法非常相同．对每幅图像进行一个平移使点的形心位于原点，然后乘以一个比例因子使点到原点的平均（RMS）距离为 $\sqrt{2}$．对直线的情形，该变换须考虑每条直线由两个端点（或在图像直线上可视的某些典型的点）来定义．三焦点张量在这些归一化变换下的变换规则在节 A1.2（$p$401）中给出．计算 $\mathcal{T}$ 的归一化线性算法概括在算法 15.1 中．

目标

已知 3 幅图像的 $n \geq 7$ 组图像点对应，或者至少 13 组直线对应或者点对应和直线对应的混合，计算三焦点张量．

算法

(1) 求适用于这三幅图像的变换矩阵 $\mathrm{H}, \mathrm{H}', \mathrm{H}''$．

(2) 按照 $x_i^s \mapsto \hat{x}_i^s = \mathrm{H}_j^s x_i^s$ 对点进行变换，按照 $l_i \mapsto \hat{l}_i = (\mathrm{H}^{-1})_i^j l_j$ 对直线进行变换．第二幅和第三幅图像上的点和直线按相同方式变换．

(3) 利用表 15.1 中的方程，采用算法 A3.5（$p$413）解形如 $\mathbf{A}\mathbf{t} = \mathbf{0}$ 的方程组，线性地计算关于变换后的点和直线的三焦点张量．

(4) 按照 $\mathcal{T}_i^{jk} = \mathrm{H}_i^r (\mathrm{H}'^{-1})_s^j (\mathrm{H}''^{-1})_t^k \hat{\mathcal{T}}_r^{st}$ 计算对应于原数据的三焦点张量．

算法 15.1　计算 $\mathcal{T}$ 的归一化线性算法

这个算法没有考虑节 15.1.1 中讨论的须应用于 $\mathcal{T}$ 的约束条件．这些约束必须加在上述算法的消除归一化（最后一步）之前强迫实行．强迫这些约束的方法将在下面考虑．

## 15.3 代数最小化算法

如节 15.1.1 中所讨论的,线性算法 15.1 给出的张量不一定对应于任何几何配置. 下一个任务就是矫正这个张量使其满足所要求的约束.

我们的任务是从一组图像对应来计算一个几何有效的三焦点张量$\mathcal{T}_i^{jk}$. 计算得到的张量将最小化相应输入数据的代数误差. 也就是说,我们在满足 $\|\hat{\mathbf{t}}\|=1$ 下最小化 $\|\mathrm{A}\hat{\mathbf{t}}\|$,其中 $\hat{\mathbf{t}}$是由几何有效的三焦点张量的元素构成的矢量. 这个算法非常类似于求基本矩阵的代数算法(节 10.3($p$193)). 与基本矩阵一样,第一步也是计算对极点.

**恢复对极点**

设 $\mathbf{e}'$和 $\mathbf{e}''$是第二幅和第三幅图像对应于第一个摄像机中心的对极点(即是该中心的像). 由结论 14.4($p$365)可知这两个对极点 $\mathbf{e}'$(和 $\mathbf{e}''$)同时正交于三个矩阵 $\mathrm{T}_i$ 的左(对应地右)零矢量. 从而这两个对极点原则上可以用算法 14.1($p$267)中概述的方法由三焦点张量而算得. 然而,在有噪声时,可以转换为基于算法 A3.5($p$413)的 4 次应用来求对极点的算法.

(1) 对每个 $i=1,2,3$,求最小化 $\|\mathrm{T}_i\mathbf{v}_i\|$ 的单位矢量 $\mathbf{v}_i$,其中 $\mathrm{T}_i=\mathcal{T}_i^{\cdot\cdot}$,以 $\mathbf{v}_i^{\mathrm{T}}$ 作为第 $i$ 行构造矩阵 V.

(2) 把使 $\|\mathrm{V}\mathbf{e}''\|$ 最小的单位矢量作为对极点 $\mathbf{e}''$.

用 $\mathrm{T}_i^{\mathrm{T}}$ 代替 $\mathrm{T}_i$ 可以类似地计算对极点 $\mathbf{e}'$.

**代数最小化**

对极点求出之后,下一步就是确定可用来计算三焦点张量的摄像机矩阵 $\mathrm{P}'$,$\mathrm{P}''$中其余的元素. 这一步计算是线性的.

从三焦点张量的形式(14.9 - $p$268)可以看到一旦对极点 $\mathbf{e}'^j=a_4^j$ 和 $\mathbf{e}''^k=b_4^k$ 已知,那么该三焦点张量可以用摄像机矩阵的其余元素 $a_i^j$ 和 $b_i^k$ 线性地表示出来. 这个关系可以被线性地写为 $\mathbf{t}=\mathrm{E}\mathbf{a}$,其中 $\mathbf{a}$ 是由余下元素 $a_i^j$ 和 $b_i^k$ 构成的矢量,$\mathbf{t}$ 是该三焦点张量的元素构成的向量,E 是由(14.9 - $p$268)表示的线性关系. 我们希望对所有满足 $\|\mathbf{t}\|=1$,即 $\|\mathrm{E}\mathbf{a}\|=1$ 的 $\mathbf{a}$,最小化代数误差 $\|\mathrm{A}\mathbf{t}\|=\|\mathrm{AE}\mathbf{a}\|$. 这个最小化问题由算法 A3.7($p$414)解决. 解 $\mathbf{t}=\mathrm{E}\mathbf{a}$ 给出了一个满足所有约束并且在给定对极点下最小化代数误差的三焦点张量.

**迭代方法**

用来计算几何有效的三焦点张量 $\mathcal{T}_i^{jk}$ 的两个对极点由 $\mathcal{T}_i^{jk}$ 的估计值确定,而 $\mathcal{T}_i^{jk}$ 由线性算法得到. 类似于基本矩阵的情形,映射$(\mathbf{e}',\mathbf{e}'')\mapsto\mathrm{AE}\mathbf{a}$ 是一个 $\mathbb{R}^6\mapsto\mathbb{R}^{27}$的映射. 应用 Levenberg - Marquardt算法优化对极点的选择将得到三焦点张量的一个(关于代数误差的)最优估计. 注意这个迭代问题的规模是适度的,因为仅有 6 个参数(即对极点的齐次坐标)包括在迭代问题中.

这个方法与后面将要考虑的用几何误差求最优三焦点张量的迭代估计形成鲜明对比. 后一问题需要估计三个摄像机的参数,加上所有点的坐标,它是一个大型估计问题.

估计三焦点张量的完整的代数算法概括在算法 15.2 中.

目标

已知三幅视图中的点对应和直线对应的集合，计算三焦点张量.

算法

(1) 从点和直线对应的集合，利用表 15.1 所给出的关系计算形如 $\mathbf{At}=\mathbf{0}$ 的方程组.

(2) 利用算法 A3.5($p$413)解这些方程，求得三焦点张量 $\mathcal{T}_i^{jk}$ 的一个初始估计.

(3) 从 $\mathcal{T}_i^{jk}$ 求两个对极点 $\mathbf{e}'$ 和 $\mathbf{e}''$ 的方法是使对极点 $\mathbf{e}'(\mathbf{e}'')$ 同时垂直于三个 $\mathrm{T}_i$ 的左(右)零矢量.

(4) 构作使得 $\mathbf{t}=\mathrm{E}\mathbf{a}$ 的 $27\times 18$ 矩阵 E，其中 $\mathbf{t}$ 是 $\mathcal{T}_i^{jk}$ 的元素组成的矢量，$\mathbf{a}$ 是表示 $a_i^j$ 和 $b_i^k$ 的元素的矢量，而 E 表示线性关系 $\mathcal{T}_i^{jk}=a_i^j e''^k-e'^j b_i^k$.

(5) 解最小化问题：利用算法 A3.7($p$414)，在满足 $\|\mathrm{E}\mathbf{a}\|=1$ 下最小化 $\|\mathrm{AE}\mathbf{a}\|$. 计算误差矢量 $\boldsymbol{\varepsilon}=\mathrm{AE}\mathbf{a}$.

(6) **迭代**：映射 $(\mathbf{e}',\mathbf{e}'')\mapsto\boldsymbol{\varepsilon}$ 是一个从 $\mathbb{R}^6$ 到 $\mathbb{R}^{27}$ 的映射. 利用 Levenberg－Marquardt 算法在上面最后两步中改变 $\mathbf{e}'$ 和 $\mathbf{e}''$ 的值进行迭代以找到最优的 $\mathbf{e}',\mathbf{e}''$. 由此找到包含 $\mathcal{T}_i^{jk}$ 的元素的最优解 $\mathbf{t}=\mathrm{E}\mathbf{a}$.

算法 15.2 **最小化代数误差的三焦点张量计算**. 同算法 15.1 一样，须对归一化的数据进行计算. 为简单起见归一化和反归一化步骤予以省略. 本算法求最小化代数误差的几何有效的三焦点张量. 倘若以稍微差一点的解为代价，上面最后的迭代步骤可省略，从而提供一个快速的非迭代算法.

## 15.4 几何距离

### 15.4.1 三焦点张量的黄金标准算法

如基本矩阵的计算一样，最好的结果可以期望从最大似然估计(或“黄金标准”)中得到. 因为该解在基本矩阵计算中已经充分介绍，对三视图情形没有多少需要补充.

给定三幅视图中的点对应集合 $\{\mathbf{x}_i\leftrightarrow\mathbf{x}_i'\leftrightarrow\mathbf{x}_i''\}$，要最小化的代价函数是

$$\sum_i d(\mathbf{x}_i,\hat{\mathbf{x}}_i)^2+d(\mathbf{x}_i',\hat{\mathbf{x}}_i')^2+d(\mathbf{x}_i'',\hat{\mathbf{x}}_i'')^2 \tag{15.3}$$

其中点 $\hat{\mathbf{x}}_i,\hat{\mathbf{x}}_i',\hat{\mathbf{x}}_i''$ 准确地满足待估计的三焦点张量的三焦点约束(如表 15.1). 如基本矩阵一样，我们需要进一步引入对应于 3D 点 $\mathbf{X}_i$ 的变量并用矩阵 P′和 P″的元素参数化该三焦点张量(见下文). 然后代价函数在 3D 点 $\mathbf{X}_i$ 的位置和满足 $\hat{\mathbf{x}}_i=[\mathrm{I}|\mathbf{0}]\mathbf{X}_i,\hat{\mathbf{x}}_i'=\mathrm{P}'\mathbf{X}_i,\hat{\mathbf{x}}_i''=\mathrm{P}''\mathbf{X}_i$ 的两个摄像机矩阵 P′和 P″上最小化. 本质上，这是对三幅视图进行捆集调整. 这需要用到 A4.3($p$418)节中的稀疏矩阵技术.

一种好的初始估计的方法是代数算法 15.2，当然最后的迭代步骤可忽略. 这个算法给出了 P′和 P″元素的一个直接估计. 3D 点 $\mathbf{X}_i$ 的初始估计可以由节 11.2($p$217)的线性三角形法而获得. 本算法的步骤概括在算法 15.3 中.

这个技术可以推广到包含线对应的情形. 为此，需要找到便于计算的 3D 直线的表示. 给定一组三视图直线对应 $\mathbf{l}\leftrightarrow\mathbf{l}'\leftrightarrow\mathbf{l}''$，每幅图像中的这些直线可由它们的端点定义. 在 LM 参数最小化中描述 3D 直线的一个非常方便的办法是利用它在第二幅和第三幅视图中的投影 $\hat{\mathbf{l}}'$ 和 $\hat{\mathbf{l}}''$. 给定一个侯选三焦点张量，利用线变换方程 $\hat{l}_j=\hat{l}_j'\hat{l}_k''\mathcal{T}_i^{jk}$ 可以很容易计算 3D 直

目标

已知 $n \geq 7$ 组图像点对应 $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i \leftrightarrow \mathbf{x}''_i\}$，确定三焦点张量的最大似然估计.

MLE 还包含求一组辅助点对应 $\{\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i \leftrightarrow \hat{\mathbf{x}}''_i\}$ 的解，它们准确地满足所估计的三焦点张量的三线性关系并且最小化

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 + d(\mathbf{x}''_i, \hat{\mathbf{x}}''_i)^2.$$

算法

(1) 利用诸如算法 15.2 的一个线性算法计算 $\mathcal{T}$ 的一个初始的几何有效估计.

(2) 计算辅助变量 $\{\hat{\mathbf{x}}_i, \hat{\mathbf{x}}'_i, \hat{\mathbf{x}}''_i\}$ 的一个初始估计：

(a) 由 $\mathcal{T}$ 恢复摄像机矩阵 $\mathrm{P}'$ 和 $\mathrm{P}''$.

(b) 利用第 11 章的三角形法，由对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i \leftrightarrow \mathbf{x}''_i$ 和 $\mathrm{P} = [\mathrm{I}|\mathbf{0}]$，$\mathrm{P}'$，$\mathrm{P}''$ 确定 $\hat{\mathbf{X}}_i$ 的一个估计

(c) 获得与 $\mathcal{T}$ 相容的对应为

$$\hat{\mathbf{x}}_i = \mathrm{P}\hat{\mathbf{X}}_i, \hat{\mathbf{x}}'_i = \mathrm{P}'\hat{\mathbf{X}}_i, \hat{\mathbf{x}}''_i = \mathrm{P}''\hat{\mathbf{X}}_i$$

(3) 在 $\mathcal{T}$ 和 $\hat{\mathbf{X}}_i$，$i = 1, \ldots, n$ 上最小化代价函数

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 + d(\mathbf{x}''_i, \hat{\mathbf{x}}''_i)^2$$

利用 Levenberg－Marquardt 算法在 $3n + 24$ 个变量（$n$ 个 3D 点 $\hat{\mathbf{X}}_i$ 有 $3n$ 个变量而摄像机矩阵 $\mathrm{P}'$，$\mathrm{P}''$ 的元素有 24 个变量）上对此代价函数进行最小化.

算法 15.3　从图像对应估计 $\mathcal{T}$ 的黄金标准算法

线到第一幅视图上的投影. 然后依据测量直线和估计直线之间的距离 $d(\mathbf{l}_i, \hat{\mathbf{l}}_i)$ 的某个适当解释，最小化直线距离的平方和

$$\sum_i d(\mathbf{l}_i, \hat{\mathbf{l}}_i)^2 + d(\mathbf{l}'_i, \hat{\mathbf{l}}'_i)^2 + d(\mathbf{l}''_i, \hat{\mathbf{l}}''_i)^2$$

如果测量直线由其端点确定，那么可用的距离度量是从测量端点到估计直线的距离. 一般情形下可以使用 Mahalanobis 距离.

### 15.4.2　三焦点张量的参数化

如果三焦点张量仅用其 27 个元素参数化，那么所估计的张量将不满足内在约束. 一个确保张量满足其内在约束（从而确保其为几何有效）的参数化称为是**相容的**.

由定义 15.1 可知：若一个张量是由（14.9－$p$268）产生的三个摄像机矩阵 $\mathrm{P} = [\mathrm{I}|\mathbf{0}]$，$\mathrm{P}'$，$\mathrm{P}''$，则它是几何有效的，由此推出这三个摄像机矩阵总给出一个相容的参数化. 注意：这是一种**超**参数化方法，因为它需要确定 24 个参数，即矩阵 $\mathrm{P}' = [\mathrm{A}|\mathbf{a}_4]$ 和 $\mathrm{P}'' = [\mathrm{B}|\mathbf{b}_4]$ 的各 12 个元素. 没有必要试图去定义一个最小（即 18）参数集，那将是一个困难的任务. 摄像机的任意选择都是一个相容参数化，具体的射影重构对三焦点张量没有影响.

另一个相容参数化方法是节 19.2（$p$359）介绍的由跨三幅视图的 6 组点对应计算三焦点张量的方法. 每幅图像上点的位置用作参数化——共 6（点）×2（对 $x, y$）×3（图像）＝36 个参数. 然而，在最小化过程中仅需改变这些点的一个子集，或者这些点的运动可以限制为正交于该三焦点张量簇.

### 15.4.3 一阶几何误差(Sampson 距离)

三焦点张量可以利用基于 Sampson 近似的几何代价函数,其方法完全类似于计算基本矩阵的 Sampson 方法(节 10.4.3($p$197). 同样,其优点是不必引入一些辅助变量,因为这个一阶几何误差仅需要在张量的参数化(例如,如果使用上文给出的 P′,P″,仅 24 个参数)上进行最小化. 最小化可以用一个简单的 Levenberg - Marquardt 迭代算法实现,并且这种方法由迭代代数算法 15.2 初始化.

此 Sampson 代价函数与基本矩阵的相应的代价函数(10.9 - $p$197)相比,计算上要稍微复杂一点,因为每组点对应给出 4 个方程,而在基本矩阵时仅为一个. 更一般的情形已在节 3.2.6($p$61)中讨论过. 在当前情形下误差函数(3.13 - $p$62)是

$$\sum_i \boldsymbol{\varepsilon}_i^{\mathrm{T}}(\mathrm{J}_i\mathrm{J}_i^{\mathrm{T}})^{-1}\boldsymbol{\varepsilon}_i \tag{15.4}$$

其中 $\boldsymbol{\varepsilon}_i$ 是单个三视图对应的代数误差向量 $\mathrm{A}_i\mathbf{t}$(当每点有 4 个方程时,它是一个 4 维矢量),J 是 $\boldsymbol{\varepsilon}$ 对每一组对应点 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i' \leftrightarrow \mathbf{x}_i''$ 的坐标求偏导数得到的 $4\times 6$ 矩阵. 如在 84 页的练习(5)的程序提示中给出的一样,偏导数矩阵 J 的计算可因代价函数是点 $\mathbf{x}_i$,$\mathbf{x}_i'$,$\mathbf{x}_i''$的坐标的多线性而简化.

Sampson 误差方法有许多优点:

- 它利用一个相对简单的迭代算法,给出了实际几何误差的一个很好的近似(最佳的).
- 与实际几何误差一样,对这些点中每一点可以指定非各向同性和不相等误差分布而不会使算法的复杂度显著加大,请参看第三章练习.

## 15.5 算法的实验评价

现在对计算三焦点张量的(迭代)代数算法 15.2 与黄金标准算法 15.3 的结果作一个简要的比较. 在对噪声水平进行控制的情况下,算法对合成的数据进行计算. 这样可以与理论上最优的 ML 算法结果作比较,并确定这些算法与由最优 ML 算法达到的残差理论下界的接近程度.

实验用计算机生成的 10、15、20 个点的数据集合来测试算法,摄像机在点群周围以随机角度放置. 摄像机参数选为近似于标准的 35mm 摄像机,并选择一个尺度因子使得图像的大小为 600×600 像素.

对图像测量中添加的高斯噪声的一个给定水平,可以根据结论 4.2($p$89)算出由 ML 算法产生的期望残差. 在这种情形,若点的数目是 $n$,则测量的数目是 $N=6n$,并且拟合中的自由度数是 $d=18+3n$,其中 18 表示三个摄像机的自由度数目(3×11 减去射影多义的 15 个),$3n$ 表示 $n$ 个空间点的自由度数目. 因此该 ML 残差是

$$\varepsilon_{rms} = \sigma(1-d/N)^{1/2} = \sigma\left(\frac{n-6}{2n}\right)^{1/2}$$

### 15.5.1 结果和建议

结果显示在图 15.1 中,从这个结果我们获悉两件事情:代数误差最小化产生的残差大约

在最优值的 15% 内，以及使用这个估计作为最小化几何误差的初始点得到了一个实际最优估计.

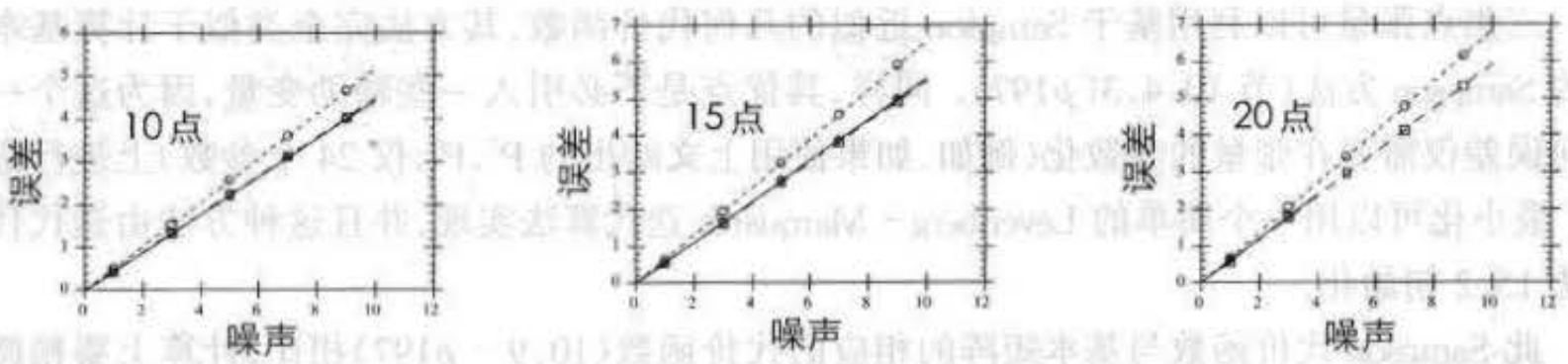



图 15.1 **三焦点张量估计算法的比较**. 图中画出利用 10、15、20 个点对三焦点张量的计算运行 100 次所得残差的 RMS 平均值与噪声水平的关系. 每幅图包含三条曲线. 最上面的曲线是代数误差最小化算法的结果，而下面两条曲线(它们在图中实际上无法区分)表示理论最小误差和利用代数最小化作为初始点由黄金标准算法得到的误差. 注意：残差几乎正比于添加的噪声，它们也应该如此.

除节 15.1 的线性算法外，上面推导的所有算法都强置了关于三焦点张量的内在约束. 建议不要单独使用线性算法，仅在必要时用作其它大多数方法的初始化. 与估计基本矩阵时一样，我们建议使用迭代代数算法 15.2，或者节 15.4.3 的 Sampson 几何近似法，两个算法都给出了极好的结果. 若高斯噪声的假设是可行的，则实施黄金标准算法 15.3 必能得到最好结果.

## 15.6 $\mathcal{T}$ 的自动计算

本节介绍自动计算三幅图像之间的三焦点几何的算法. 算法的输入仅仅是图像三元组，无需其他的**先验**信息，而输出是估计的三焦点张量和一个跨三幅图像的对应点集.

给定一个点在两幅视图中的像的位置后，三焦点张量可以用来确定该点在第三幅视图中像的精确位置，这个事实意味着三幅视图比两视图产生的误匹配少. 在两视图中仅能利用关于对极线的较弱的几何约束来检验可能的匹配.

三视图算法类似于节 3.8($p79$)介绍的单应的自动计算方式，使用 RANSAC 作为搜索引擎. 算法的思想和细节在那里已经给出，这里就不再重复了. 该方法概括在算法 15.4 中，一个应用它的例子在图 15.2 中给出. 这些步骤的补充解释在下文给出. 图 15.3 给出自动计算线匹配的第二个例子.

**距离测量——重投影误差** 给定匹配 $\mathbf{x}\leftrightarrow\mathbf{x}'\leftrightarrow\mathbf{x}''$ 和 $\mathcal{T}$ 的当前估计，我们需要确定重投影误差——$d_\perp^2 = d^2(\mathbf{x},\hat{\mathbf{x}}) + d^2(\mathbf{x}',\hat{\mathbf{x}}') + d^2(\mathbf{x}'',\hat{\mathbf{x}}'')$ 的最小值，其中图像点 $\hat{\mathbf{x}},\hat{\mathbf{x}}',\hat{\mathbf{x}}''$ 与 $\mathcal{T}$ 相容. 相容的图像点照例可由 3 维空间点 $\hat{\mathbf{X}}$ 的投影

$$\hat{\mathbf{x}} = [\mathrm{I} \mid \mathbf{0}]\hat{\mathbf{X}} \qquad \hat{\mathbf{x}}' = \mathrm{P}'\hat{\mathbf{X}} \qquad \hat{\mathbf{x}}'' = \mathrm{P}''\hat{\mathbf{X}}$$

获得，其中摄像机矩阵 $\mathrm{P}',\mathrm{P}''$ 由 $\mathcal{T}$ 算得. 然后距离 $d_\perp^2$ 可以由已确定点 $\hat{\mathbf{X}}$ 而获得，点 $\hat{\mathbf{X}}$ 最小化测量点 $\mathbf{x},\mathbf{x}',\mathbf{x}''$ 和其投影点之间的图像距离.

获得这一距离的另一方法是利用 Sampson 误差(15.4)，这个误差是几何误差的一阶近似. 然而，实际上直接用非线性最小二乘迭代(一个小型的 Levenberg－Marquardt 问题)来估计该误差更快. 从 $\hat{\mathbf{X}}$ 的一个初始估计出发，通过迭代改变 $\hat{\mathbf{X}}$ 的坐标使重投影误差最小化.

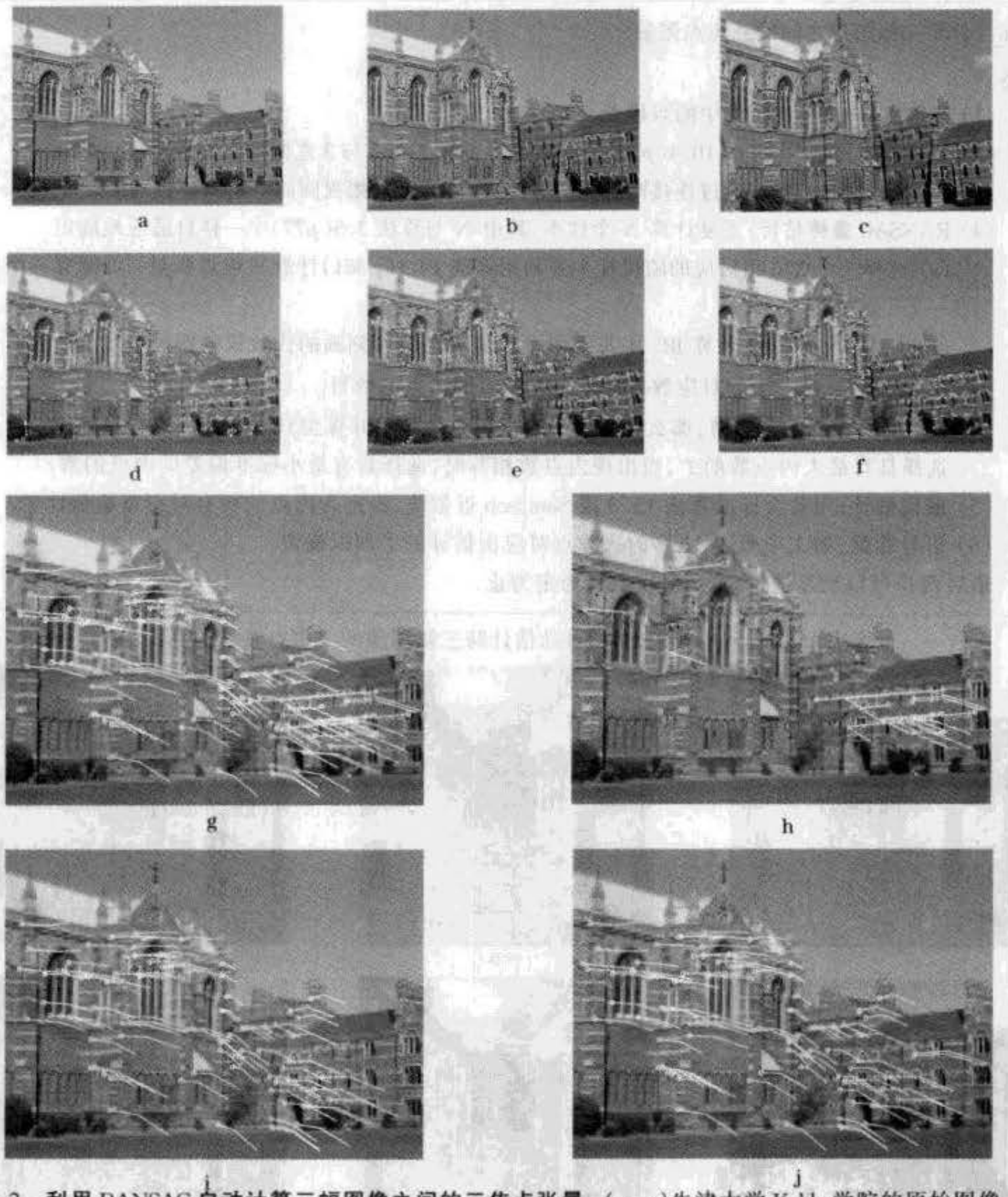



图 15.2 **利用 RANSAC 自动计算三幅图像之间的三焦点张量**. (a－c)牛津大学 Keble 学院的原始图像. 视图之间的运动包括平移和旋转. 图像是 640×480 像素. (d－f) 检测到的角点叠加在图像上,每幅图像上大约有 500 个角点. 下面的结果皆叠加在图像(a)上:(g) 106 对假设匹配,用连接角点的直线来显示,注意明显的误匹配:(h)野值——假设匹配中的 18 个,(i)内点——与估计的 $\mathcal{T}$ 相容的 88 组对应;(j)在引导匹配和 MLE 之后的最终 95 组对应,其中没有误匹配.

**引导匹配** 有了一个 $\mathcal{T}$ 的初始估计之后就可用其产生和判断三视图中更多的点对应. 第一步是从 $\mathcal{T}$ 算出视图 1 和 2 之间的基本矩阵 $F_{12}$. 然后在两视图的引导匹配中采用宽松的匹配阈值来计算. 每个两视图的匹配用 $F_{12}$ 矫正,从而给出与 $F_{12}$ 相容的 $\hat{\mathbf{x}}, \hat{\mathbf{x}}'$. 矫正的两视图匹配(与 $\mathcal{T}$ 一起)在第三幅视图中定义了一个小搜索窗,可以在其中搜索对应点. 如上所述,任何三视图点对应由计算 $d_\perp$ 来判断. 如果 $d_\perp$ 小于阈值 $t$,就接受该匹配. 注意:RANSAC 和引导匹配应采用相同的阈值来进行内点检测.

目标　计算三幅图像之间的三焦点张量.

算法

(1) **兴趣点**:计算每幅图像中的兴趣点.

(2) **两视图对应**:利用算法 10.4($p$200),计算视图 1 与 2 和 2 与 3 之间的兴趣点对应(和 F).

(3) **假定的三视图对应**:通过连接两视图的匹配集,计算跨三幅视图的一组兴趣点对应.

(4) RANSAC **鲁棒估计**:重复计算 $N$ 个样本,其中 $N$ 与算法 3.5($p$77)中一样自适应地确定.

(a) 选择一个含 6 组对应的随机样本并利用算法 19.1($p$361)计算三焦点张量. 可能有一个或三个实解.

(b) 如节 15.6 所述,计算 $\mathbb{R}^6$ 中每组假设对应到由 $\mathcal{T}$ 所刻画的代数簇的距离 $d_\perp$.

(c) 由满足 $d_\perp < t$ 的对应数目计算与 $\mathcal{T}$ 相容的内点的数目.

(d) 如果 $\mathcal{T}$ 有三个实解,那么计算每个解的内点的数目并保留具有最大内点数目的解.

选择具有最大内点数的 $\mathcal{T}$,当出现内点数相等时,选择具有最小标准偏差的内点的解.

(5) **最优估计**:用黄金标准算法 15.3 或 Sampson 近似法,由定为内点的所有对应重新估计 $\mathcal{T}$.

(6) **引导匹配**:如文中所述,更多的兴趣点对应由估计的 $\mathcal{T}$ 加以确定.

最后两步可以反复迭代,直到对应数目稳定为止.

算法 15.4　利用 RANSAC 自动估计跨三幅图像的三焦点张量的算法

图 15.3　**图像三元组匹配**. 利用算法 15.4,三焦点张量由我们感兴趣的点自动计算得到,然后用它来匹配跨视图的直线特征. (a)一个走廊的三幅图像,(b)自动匹配直线段. 本匹配算法在[Schmid－97]中介绍.

实践发现这里的引导匹配比单应估计时更为重要,它能生成更多的对应.

**实施与运行细节**　对图 15.2 的例,搜索窗是 ±300 像素. 内点阈值是 $t = 1.25$ 像素. 总共需要 26 次采样. RMS 像素误差在 RANSAC 之后是 0.43(对 88 组对应),在 MLE 之后是 0.23(对 88 组对应),在 MLE 和引导匹配之后是 0.19(对 95 组对应). MLE 需要用 Levenberg－Marquardt算法迭代 10 次.

注意:这里的 RANSAC 所需的工作比算法 10.4($p$200)估计 F 和它的对应所需的工作

少得多,因为在三幅视图的假设对应产生之前两视图算法已经删去了大量野值.

## 15.7 计算 $\mathcal{T}$ 的特殊情形

### 15.7.1 由一张平面加视差来计算 $\mathcal{T}_i^{jk}$

这里,我们介绍从一张世界平面(从它可以计算视图间的单应)和平面外两点组成的特殊配置的图像来计算 $\mathcal{T}_i^{jk}$. 当然并不需要这张平面真正存在. 它可以是虚拟的,或者可以仅由 4 个共面点或 4 条共面直线的图像来规定该单应. 这个方法类似于基本矩阵的算法 12.2 ($p$237).

这个问题的解法可以是首先构造三个摄像机矩阵(相差一个共同的三维空间的射影变换),然后按(14.9 - $p$268)由这些矩阵算出此三焦点张量. 由世界(参考)平面诱导的第一幅和第二幅视图之间的单应是 $H_{12}$,第一幅和第三幅视图之间的单应是 $H_{13}$. 节 12.3($p$235)证明:对第一幅和第二幅视图而言,对极点 $\mathbf{e}'$ 可以由平面外的两组点对应直接得到,摄像机矩阵选为 $P=[I|\mathbf{0}]$ 和 $P'=[H_{12}|\mu\mathbf{e}']$,其中 $\mu$ 是一个标量. 注意:这里 $H_{12}$ 和 $\mathbf{e}'$ 的尺度因子被视为是固定的,从而它们不再是齐次量. 类似地,对第一幅和第三幅视图而言,$\mathbf{e}''$ 也可以由两组点对应确定,并且摄像机矩阵选为 $P=[I|\mathbf{0}]$,$P''=[H_{13}|\lambda\mathbf{e}'']$,其中 $\lambda$ 是一个标量.

容易证明与该三视图相容的一个摄像机集合(见 $p$267 关于相容摄像机三元组的讨论)由

$$P=[I\mid\mathbf{0}],\quad P'=[H_{12}\mid\mathbf{e}'],\quad P''=[H_{13}\mid\lambda\mathbf{e}''] \tag{15.5}$$

给出,其中 $\mu$ 已经被设为 1. $\lambda$ 的值由跨三幅视图的点对应之一确定,其证明留作练习. 关于平面加视差重构的更多信息见节 17.4($p$316).

注意:对这个配置而言,三焦点张量的估计是超定的. 对于两视图的基本矩阵情形,该单应确定了除两个自由度(对极点)之外的所有自由度,并且每组点对应提供一个约束,因此约束的数目等于矩阵的自由度的数目. 在三焦点张量时,单应决定了除 5 个自由度(两个对极点和它们的相对比例)之外的所有自由度. 然而,每组点对应给出三个约束(6 个坐标测量减去该 3D 点的位置所需的 3 个),从而 5 个自由度有 6 个约束. 此时因为测量数比自由度多,张量应由最小化一个基于几何误差的代价函数来估计.

### 15.7.2 由多点确定的直线

在介绍按直线的重构算法中,我们已考虑了直线由它们的两端点定义的情形. 在一幅图像中确定直线的另一个常用方法是取与多点最佳拟合的直线. 现在将证明如何把后一种情形简化为由两端点定义直线的情形. 考虑一幅图像中点的集合,把它们归一化使得其第三个坐标等于 1. 令 $\mathbf{l}=(l_1,l_2,l_3)^\mathsf{T}$ 是一条直线,并假设它已被归一化使得 $l_1^2+l_2^2=1$. 在这种情形下,点 $\mathbf{x}_i$ 到直线 $\mathbf{l}$ 的距离等于 $\mathbf{x}_i^\mathsf{T}\mathbf{l}$. 该距离的平方可以写为 $d^2=\mathbf{l}^\mathsf{T}\mathbf{x}_i\mathbf{x}_i^\mathsf{T}\mathbf{l}$,而所有距离的平方和是

$$\sum_i \mathbf{l}^\mathsf{T}\mathbf{x}_i\mathbf{x}_i^\mathsf{T}\mathbf{l}=\mathbf{l}^\mathsf{T}\Big(\sum_i \mathbf{x}_i\mathbf{x}_i^\mathsf{T}\Big)\mathbf{l}.$$

矩阵 $E=(\sum_i \mathbf{x}_i\mathbf{x}_i^\mathsf{T})$ 是正定对称的.

**引理 15.2** **矩阵$(E-\varepsilon_0 J)$是半正定的，其中 J 是对角矩阵 diag(1,1,0)，$\varepsilon_0$ 是方程 $\det(E-\varepsilon J)=0$的最小解.**

**证明** 我们首先在条件 $x_1^2+x_2^2=1$ 下计算最小化 $\mathbf{x}^T E\mathbf{x}$ 的矢量 $\mathbf{x}=(x_1,x_2,x_3)^T$. 在使用拉格朗日乘数法后它便归结为最小化 $\mathbf{x}^T E\mathbf{x}-\zeta(x_1^2+x_2^2)$，其中 $\zeta$ 记拉格朗日系数. 对 $\mathbf{x}$ 求微分并令其为 $\mathbf{0}$，我们得到 $2E\mathbf{x}-\zeta(2x_1,2x_2,0)^T=\mathbf{0}$，此式可以写为$(E-\zeta J)\mathbf{x}=\mathbf{0}$，由此推出 $\zeta$ 是方程 $\det(E-\zeta J)=0$ 的一个根，$\mathbf{x}$ 是 $E-\zeta J$ 的零空间的生成元. 因为 $\mathbf{x}^T E\mathbf{x}=\zeta\mathbf{x}^T J\mathbf{x}=\zeta(x_1^2+x_2^2)=\zeta$，所以为了最小化 $\mathbf{x}^T E\mathbf{x}$ 必须选择 $\zeta$ 为方程 $\det(E-\zeta J)=0$ 的最小根 $\zeta_0$. 在此情形下，对最小化矢量 $\mathbf{x}_0$ 有 $\mathbf{x}_0^T E\mathbf{x}_0-\zeta_0=0$；对其他任何不一定是最小化的矢量 $\mathbf{x}$，我们有 $\mathbf{x}^T E\mathbf{x}-\zeta_0\geq 0$，从而有 $\mathbf{x}^T(E-\zeta_0 J)\mathbf{x}=\mathbf{x}^T E\mathbf{x}-\zeta_0\geq 0$，所以 $E-\zeta_0 J$ 是半正定的. □

由于矩阵 $E-\zeta_0 J$ 是对称的，它可以写成 $E-\zeta_0 J=V\mathrm{diag}(r,s,0)V^T$ 的形式，其中 V 是正交矩阵，$r$ 和 $s$ 是正数，由此推出

$$\begin{aligned} E-\zeta_0 J &= V\mathrm{diag}(r,0,0)V^T+V\mathrm{diag}(0,s,0)V^T \\ &= r\mathbf{v}_1\mathbf{v}_1^T+s\mathbf{v}_2\mathbf{v}_2^T \end{aligned}$$

其中 $\mathbf{v}_i$ 是 V 的第 $i$ 列. 所以 $E=\zeta_0 J+r\mathbf{v}_1\mathbf{v}_1^T+s\mathbf{v}_2\mathbf{v}_2^T$，那么对满足 $l_1^2+l_2^2=1$ 的任意直线 $\mathbf{l}$ 我们有

$$\begin{aligned} \sum_i(\mathbf{x}_i^T\mathbf{l})^2 &= \mathbf{l}^T E\mathbf{l} \\ &= \zeta_0+r(\mathbf{v}_1^T\mathbf{l})^2+s(\mathbf{v}_2^T\mathbf{l})^2 \end{aligned}$$

因此，我们用一个不可被最小化的常值 $\zeta_0$，加上到两点 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 的距离的加权平方和来取代若干点的(距离)平方和. 概括起来：当建立涉及由点 $\mathbf{x}_i$ 定义的直线的三焦点张量方程时，可用权值分别为$\sqrt{r}$和$\sqrt{s}$的点 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 表示的两个点的方程.

## 15.8 结束语

### 15.8.1 文献

计算三焦点张量的线性算法首先由[Hartley-97c]给出，其中报告了用实际数据的点和线对应作估计的进一步实验结果. [Hartley-98d]中给出了估计相容张量的迭代代数算法.

Torr 和 Zisserman[Torr-97]推导了由三幅图像估计相容张量 $\mathcal{T}$ 的自动算法，该文章还对迭代最小化的几个参数化法进行了比较. Faugeras 和 Papadopoulo[Faugeras-97]给出了多种方法来表达和引入对张量的约束.

### 15.8.2 注释和练习

(1) 考虑给定三焦点张量求最小化图像测量点 $\mathbf{x},\mathbf{x}',\mathbf{x}''$的重投影误差的三维空间点 $\mathbf{X}$ 的估计问题. 它类似于第 11 章中的三角形问题. 证明：对一般运动，第 11 章中推导的对极线的单参数簇参数化方法不能从两视图推广到三视图. 然而，当三个摄像机中心共线时，两视图参数化方法可以被推广到三视图.

(2) 仿射三焦点张量可以从一般位置上的 4 点的最小配置算得. 具体的计算类似于算

法 13.2($p$249),并且所算得的张量满足仿射三焦点张量的内在约束. 仿射情形有多少个约束?

如果在估计中用的点多于 4,那么利用 17.2($p$310)的分解算法可估计出一个几何有效张量.

(3) 张量的变换规则是 $\mathcal{T}_i^{jk} = A_i^r (B^{-1})_s^j (C^{-1})_t^k \hat{\mathcal{T}}_r^{st}$,这可以容易地计算如下:

```
Binv = B.inverse();
Cinv = C.inverse();
for (i=1; i<=3; i++) for (j=1; j<=3; j++) for (k=1; k<=3; k++)
{
  T[i][j][k]=0.0;
  for (r=1; r<=3; r++) for (s=1; s<=3; s++) for (t=1; t<=3; t++)
  T[i][j][k] += A[r][i] * Binv[j][s] * Cinv[k][t] * T_hat[r][s][t];
}
```

这一计算包含多少乘法和循环迭代? 寻找计算这个变换的一个更好的方法.

(4) 在利用平面加视差计算三焦点张量(节 15.7.1)过程中证明:若 $\rho$ 是该平面外一点的射影深度(即 $\mathbf{x}' = H_{12}\mathbf{x} + \rho\mathbf{e}'$,见(12.9 - $p$236)),则(15.5)中的标量 $\lambda$ 可以由方程 $\mathbf{x}'' = H_{13}\mathbf{x} + \rho\lambda\mathbf{e}''$ 计算.

平面加两点配置的三幅视图中存在丰富的几何,设(参考)平面外的点为 $\mathbf{X}$ 和 $\mathbf{Y}$. 分别从三个摄像机中心将点 $\mathbf{X}$ 投影到该参考平面上形成三角形 $\mathbf{x},\mathbf{x}',\mathbf{x}''$,类似地投影点 $\mathbf{Y}$ 形成三角形 $\mathbf{y},\mathbf{y}',\mathbf{y}''$. 于是这两个三角形形成一个 Desargues 配置,并由一个平面单应相联系(见 A5.2 节($p$427)). 用一个简单的草图就可表明连接两三角形对应顶点的直线$(\mathbf{x},\mathbf{y})$,$(\mathbf{x}',\mathbf{y}')$,$(\mathbf{x}'',\mathbf{y}'')$共点,并且它们的交点正是 $\mathbf{X}$ 和 $\mathbf{Y}$ 的连线与参考平面的交点. 类似地,两三角形对应边的交点共线,并且由这样形成的直线正是摄像机的三焦点平面与参考平面的交线. 进一步的细节在[Criminisi - 98, Irani - 98]中给出.

# 第4篇 $N$ 视图几何

## 本篇大纲

本篇的部分内容是择要重述，部分内容是新的材料.

第16章是择要重述. 我们重新回到两和三视图几何，但在更一般的框架之下加以叙述使之能自然地推广到四和 $n$ 视图. 多视图上的基本射影关系来自直线(点的反向投影)和平面(直线的反向投影)的相交. 这些相交性质由视图摄像机形成的某些为零的行列式来表征. 基本矩阵、三焦点张量以及四视图的新的张量——**四焦点张量**——分别作为两、三和四视图的多视图张量由这些行列式自然产生. 去掉3D结构和摄像机矩阵的非本质部分，剩下来的就是张量. 这些张量仅介绍到四视图.

每组视图的张量是唯一的，并且产生关于图像测量坐标的多重线性关系. 这些张量能从图像对应的集合中计算得到，然后每幅视图的摄像机矩阵又可以从张量计算得到. 最后，3D结构可以由恢复的摄像机和图像对应计算得到.

第17章覆盖从多视图进行重构的计算. 特别给出了由视图仿射重构的重要分解算法. 这个算法之所以重要是因为它是最优的，同时又是非迭代的.

第18章介绍了摄像机的自标定. 它们是基于多幅图像上的约束来计算摄像机内参数的各种方法. 与第6章所介绍的传统标定方法不同，它们不采用专门的景物标定物体，而只有简单的约束，比如多幅图像的摄像机内参数是一致的，或者摄像机绕着其中心旋转且不改变长宽比等.

第19章强调点和摄像机的对偶性，以及它如何与贯穿于本书给出的各种配置和算法相联系. 本章包含根据六点在三幅视图中的影像来计算重构的一种算法.

第20章研究点是否在一个或多个摄像机的前面或后面的问题. 这是一个超出了贯穿本书的齐次表示的问题，因为齐次表示不区别一条射线的方向.

第21章指出哪些配置使本书中所介绍的估计算法失败. 这个重要议题的一个例子是关于摄像机投影矩阵的计算：如果所有的3D点和摄像机中心在一条三次绕线上，则摄像机矩阵不能被计算得到.

# 第 16 章　$N$ 线性和多视图张量

本章介绍四幅视图之间的四焦点张量 $Q^{ijkl}$，它类似于两幅视图的基本矩阵和三幅视图的三焦点张量. 四焦点张量全面地包含了四幅视图中的点和直线之间的关系.

已经证明多视图关系可以直接并一致地由直线和点的反向投影的相交性质推导出来. 通过这个分析，基本矩阵 F、三焦点张量 $\mathcal{T}_i^{jk}$ 和四焦点张量 $Q^{ijkl}$ 都在矩阵行列式的共同框架下出现. 可以用摄像机矩阵给出上述每个张量的具体公式.

我们还将推导计算张量自由度以及计算这些张量所需要的点和直线对应的数目的一般方法. 我们给出了处于一般位置的配置和当四个或多个元素共面这样一种重要的特殊情形时的有关结果.

## 16.1　双线性关系

我们首先考虑在两幅视图中都能看到的点的坐标之间存在的关系. 令 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 是一组对应点，它们是在两幅视图中都能看到的空间同一点 $\mathbf{X}$ 的像. 为清晰起见，用矩阵 A 和 B 代替常用的记号 P 和 P′来表示这两个摄像机矩阵. 从空间到图像的投影可以表示为 $k\mathbf{x}=\mathrm{A}\mathbf{X}$ 和 $k'\mathbf{x}'=\mathrm{B}\mathbf{X}$，其中 $k$ 和 $k'$ 是两个未定常数. 这一对方程可以写为一个方程：

$$\begin{bmatrix} \mathrm{A} & \mathbf{x} & \mathbf{0} \\ \mathrm{B} & \mathbf{0} & \mathbf{x}' \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -k \\ -k' \end{pmatrix} = \mathbf{0}$$

并且容易验证上式等价于之前提到的两个方程. 记矩阵 A 的第 $i$ 行为 $\mathbf{a}^i$，类似地记矩阵 B 的第 $i$ 行为 $\mathbf{b}^i$，以上方程可以写为更详细的形式. 我们同时记 $\mathbf{x}=(x^1,x^2,x^3)^\mathsf{T}$ 和 $\mathbf{x}'=(x'^1,x'^2,x'^3)^\mathsf{T}$. 则该方程组是：

$$\begin{bmatrix} \mathbf{a}^1 & x^1 & \\ \mathbf{a}^2 & x^2 & \\ \mathbf{a}^3 & x^3 & \\ \hline \mathbf{b}^1 & & x'^1 \\ \mathbf{b}^2 & & x'^2 \\ \mathbf{b}^3 & & x'^3 \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -k \\ -k' \end{pmatrix} = \mathbf{0}. \tag{16.1}$$

这是一个 $6\times6$ 方程组，由假设它有一个非零解向量 $(\mathbf{X}^\mathsf{T},-k,-k')^\mathsf{T}$. 由此推出(16.1)的系数矩阵的行列式的值必须为零. 我们将看到这个条件引导出由基本矩阵 F 表示的向量 $\mathbf{x}$ 和 $\mathbf{x}'$ 元素之间的一个双线性关系. 我们将具体地研究这个关系的形式.

考虑(16.1)中出现的矩阵. 把它记为 X，X 的行列式可以写成量 $x^i$ 和 $x'^i$ 的一个表达式. 注意元素 $x^i$ 和 $x'^i$ 只出现在 X 的两个列中. 这表明 X 的行列式可以表示为 $x^i$ 和 $x'^i$ 的

二次式. 事实上,由于所有元素 $x^i$ 都出现在同一列,不可能有形如 $x^i x^j$ 和 $x'^i x'^j$ 的项,简而言之,用 $x^i$ 和 $x'^i$ 表示时,X 的行列式是一个双线性表达式. 这个行列式为零可以写为一个方程

$$(x'^1, x'^2, x'^3)\mathrm{F}(x^1, x^2, x^3)^{\mathrm{T}} = x^j x'^i F_{ij} = 0 \tag{16.2}$$

其中 F 是一个 3×3 矩阵,即基本矩阵①.

我们为矩阵 F 的元素推导一个具体的计算公式. F 的元素 $F_{ij}$ 是 X 的行列式展开式中项 $x'^i x^j$ 的系数. 为了求这个系数,我们必须消去矩阵中包含 $x'^i$ 和 $x^j$ 的行和列,再对所得到的矩阵取行列式并且适当地乘以 ±1. 比如 $x'^1 x^1$ 的系数可以通过消去矩阵 X 的两行和最后两列而得到. 所余下的矩阵是

$$\begin{bmatrix} \mathbf{a}^2 \\ \mathbf{a}^3 \\ \mathbf{b}^2 \\ \mathbf{b}^3 \end{bmatrix}$$

因而 $x'^1 x^1$ 的系数等于这个 4×4 矩阵的行列式. 一般,我们可以写为

$$F_{ji} = (-1)^{i+j} \det \begin{bmatrix} \sim \mathbf{a}^i \\ \sim \mathbf{b}^j \end{bmatrix} \tag{16.3}$$

在这个表达式中,记号 $\sim \mathbf{a}^i$ 用来标记从矩阵 A 中**略去**行 $\mathbf{a}^i$ 而得到的矩阵. 这样一来符号 ~ 可以读作**略去**,而且 $\sim \mathbf{a}^i$ 表示 A 的两行. 因此在(16.3)的右边出现的行列式是一个 4×4 行列式.

利用张量 $\varepsilon_{rst}$(定义在节 A1.1($p$400)中)可以将 $F_{ji}$ 用另一种方法表示为:②

$$F_{ji} = \left(\frac{1}{4}\right) \varepsilon_{ipq} \varepsilon_{jrs} \det \begin{bmatrix} \mathbf{a}^p \\ \mathbf{a}^q \\ \mathbf{b}^r \\ \mathbf{b}^s \end{bmatrix}. \tag{16.4}$$

为了证实这一点,注意(16.4)中定义的 $F_{ji}$ 是 $p$、$q$、$r$ 和 $s$ 的所有可能值所对应的行列式的累加. 然而,对给定 $i$ 值,除非 $p$,$q$ 不同于 $i$ 且相互不同,否则张量 $\varepsilon_{ipq}$ 为零. 这令 $p$ 和 $q$ 只留下了两种选择(例如若 $i=1$,则我们可以选择 $p=2$, $q=3$ 或者 $p=3$, $q=2$). 类似地,对于 $r$ 和 $s$ 也只有两种不同的选择给出非零项. 这样一来这个和式只包含四个非零项. 而且,出现于这四项中的行列式都包含矩阵 A 和 B 的同样的四行,从而除了符号外它们具有相等的值. 然而,$\varepsilon_{ipq}\varepsilon_{jrs}$ 的值使得这四项都有相同的符号且相等. 因此,(16.4)中的和等于(16.3)出现的单项.

### 16.1.1 张量形式的对极点

基本矩阵的表达式(16.3)中的矩阵涉及从 A 和 B 各取两行所得的行列式. 若我们考虑用一个矩阵的三行和另一个矩阵的一行组成的矩阵的行列式,则所得到的行列式便表示对

① 似乎用公式 $x^i x'^j F_{ij} = 0$ 来定义 $F_{ij}$ 更合理,但是公式(16.2)是习惯用法.

② 当然因子 1/4 是非本质的,因为 F 在相差一个比例因子下被定义. 它包含在这里仅仅是为了表明与(16.3)式的联系.

极点. 具体地说,我们有

$$e^i = \det\begin{bmatrix}\mathbf{a}^i \\ \mathrm{B}\end{bmatrix} \qquad e'^j = \det\begin{bmatrix}\mathrm{A} \\ \mathbf{b}^j\end{bmatrix} \tag{16.5}$$

其中 $\mathbf{e}$ 和 $\mathbf{e}'$ 是两幅图像上的对极点. 为了理解它,注意对极点的定义为 $e^i = \mathbf{a}^i\mathbf{C}'$,其中 $\mathbf{C}'$ 是第二个摄像机的中心,由 $\mathrm{B}\mathbf{C}' = \mathbf{0}$ 定义. 公式(16.5)现在可以通过把行列式按第一行展开而得到(与推导(2.4 - $p$35)相类似的方法).

### 16.1.2 仿射特例

当两个摄像机都是仿射摄像机时,其基本矩阵具有特别简单的形式. 仿射摄像机矩阵是最后一行为(0,0,0,1)的矩阵. 现在注意:若(16.3)中 $i,j$ 都不等于 3,则在 $F_{ij}$ 的表达式中 A 和 B 的第三行都出现. 这个行列式便有两个相同的行,因而为零. 因此 F 的形式是

$$\mathrm{F}_{\mathrm{A}} = \begin{bmatrix} & & a \\ & & b \\ c & d & e \end{bmatrix},$$

其中所有未标记的其他元素为零. 这样一来仿射基本矩阵只有 5 个非零元素,因而有 4 个自由度. 它的性质在节 13.2($p$244)中描述.

注意这个论证仅仅依赖于两个摄像机都有相同的第三行这一事实. 由于一个摄像机矩阵的第三行表示摄像机的主平面(见节 5.2.1($p$105)),由此推出两个有相同主平面的摄像机的基本矩阵具有上述形式.

## 16.2 三线性关系

推导基本矩阵的行列式方法可以用来推导在三幅视图中都能看到的点的坐标之间的关系. 这一分析导出了三焦点张量的一个公式. 与基本矩阵不同,三焦点张量与三幅图像中的直线和点都有联系. 我们从介绍对应点的关系开始.

### 16.2.1 三焦点中的点关系

考虑一组跨三视图的点对应:$\mathbf{x}\leftrightarrow\mathbf{x}'\leftrightarrow\mathbf{x}''$. 令第三个摄像机矩阵为 C,且 $\mathbf{c}^i$ 为它的第 $i$ 行. 类似于(16.1),我们能用一个方程描述点 $\mathbf{X}$ 在三幅图像上的投影为

$$\begin{bmatrix}\mathrm{A} & \mathbf{x} & & \\ \mathrm{B} & & \mathbf{x}' & \\ \mathrm{C} & & & \mathbf{x}''\end{bmatrix}\begin{pmatrix}\mathbf{X} \\ -k \\ -k' \\ -k''\end{pmatrix} = \mathbf{0}. \tag{16.6}$$

如前面一样,把这个矩阵记为 X,它有 9 行和 7 列. 根据这个方程组有一个(非零)解,我们可以推出它的秩最多为 6. 因此它的任何 7×7 的子式都为零. 这个事实确立了点 $\mathbf{x}$,$\mathbf{x}'$ 和 $\mathbf{x}''$ 的坐标之间存在的三线性关系.

本质上存在着 X 的两种不同类型的 7×7 子式. 在选择 X 的 7 行方面,我们可以有下列两种选择

(1) 从两个摄像机矩阵中各选三行而从第三个中选一行,或者

(2) 从一个摄像机矩阵中选三行而从另外两个中各选两行.

让我们考虑第一种类型. X 的一个典型的 7×7 子矩阵的形式是

$$\begin{bmatrix} \mathrm{A} & \mathbf{x} & \\ \mathrm{B} & & \mathbf{x}' \\ \mathbf{c}^i & & x''^i \end{bmatrix}. \tag{16.7}$$

注意这个矩阵最后一列只包括一个元素即 $x''^i$. 按最后一列展开这个行列式推出这个行列式等于

$$x''^i \det\begin{bmatrix} \mathrm{A} & \mathbf{x} \\ \mathrm{B} & \mathbf{x}' \end{bmatrix}.$$

除因子 $x''^i$ 不同外,这正是节 16.1 中所讨论的由基本矩阵表示的双线性关系.

7×7 余子式的另外一种类型更有趣. 这种行列式的一个例子形如

$$\det\begin{bmatrix} \mathrm{A} & \mathbf{x} & & \\ \mathbf{b}^j & & x'^j & \\ \mathbf{b}^l & & x'^l & \\ \mathbf{c}^k & & & x''^k \\ \mathbf{c}^m & & & x''^m \end{bmatrix}. \tag{16.8}$$

采用在论证双线性关系时所用的同样的推理,令行列式为零给出形如 $f(\mathbf{x},\mathbf{x}',\mathbf{x}'')=0$ 的三线性关系. 按包含 $x^i$ 的列展开这个行列式,我们可以得到如下的特殊公式.

$$\det \mathrm{X}_{uv} = -\frac{1}{2} x^i x'^j x''^k \varepsilon_{ilm}\varepsilon_{jqu}\varepsilon_{krv} \det\begin{bmatrix} \mathbf{a}^l \\ \mathbf{a}^m \\ \mathbf{b}^q \\ \mathbf{c}^r \end{bmatrix} = 0_{uv}, \tag{16.9}$$

其中 $u$ 和 $v$ 对应于为了产生(16.8)式从矩阵 B 和 C 中略去的行的自由指标. 我们引入张量

$$\mathcal{T}_i^{qr} = \frac{1}{2}\varepsilon_{ilm} \det\begin{bmatrix} \mathbf{a}^l \\ \mathbf{a}^m \\ \mathbf{b}^q \\ \mathbf{c}^r \end{bmatrix}, \tag{16.10}$$

则三线性关系(16.9)可以写为

$$x^i x'^j x''^k \varepsilon_{jqu}\varepsilon_{krv}\ \mathcal{T}_i^{qr} = 0_{uv}. \tag{16.11}$$

如在节 14.2.1($p$269)所讨论的张量 $\mathcal{T}_i^{qr}$ 正是三焦点张量,(16.11)是一个三线性关系. 指标 $u$ 和 $v$ 是自由指标,且 $u$ 和 $v$ 的每一种选择会导致一种不同的三线性关系.

与基本矩阵情形一样,可以用稍微不同的方式把张量 $\mathcal{T}_i^{qr}$ 的公式写为

$$\mathcal{T}_i^{qr} = (-1)^{i+1} \det\begin{bmatrix} \sim \mathbf{a}^i \\ \mathbf{b}^q \\ \mathbf{c}^r \end{bmatrix}. \tag{16.12}$$

如节 16.1 一样,表达式 $\sim\mathbf{a}^i$ 表示略去第 $i$ 行的矩阵 A. 注意我们从第一个摄像机矩阵略去第 $i$ 行,同时包含另外两个摄像机矩阵中的第 $q$ 行和第 $r$ 行.

对于常用的情形，第一个摄像机矩阵 A 具有标准形式$[\mathrm{I}|\mathbf{0}]$，从而三焦点张量表达式(16.12)可以简单写为

$$\mathcal{T}_i^{qr} = b_i^q c_4^r - b_4^q c_i^r. \tag{16.13}$$

注意事实上按这种方式可能形成 27 种三线性关系(参考(16.8))．具体地说，从一个摄像机取三行并从另两个摄像机中各取两行产生每一个关系．它给出下面的计算．

- 从第一个摄像机取所有 3 行有 3 种方式可选择．
- 从第二个摄像机略去特定的行有 3 种方式可选择．
- 从第三个摄像机略去特定的行有 3 种方式可选择．

这总共给出 27 种三线性关系．然而，从第二个和第三个摄像机矩阵的选择两行的 9 种方式中，只有四种是线性独立的(我们将在节 16.6 中回到这个问题)．这意味着总共有 12 个线性独立的三线性关系．

然而区别三线性关系的数目与不同的三焦点张量的数目是非常重要的，如(16.11)所示，几个不同的三线性关系可以只用一个三焦点张量来表示．在(16.11)中自由指标 $u$ 和 $v$ 的每种不同选择将产生一个不同的三线性关系，所有这些关系都是用同一个三焦点张量 $\mathcal{T}_i^{qr}$ 来表示．另一方面，在(16.10)给出的三焦点张量定义中，摄像机矩阵 A 的处理与另外两个不同：在定义任何给定 $\mathcal{T}_i^{qr}$ 元素的行列式中，A 提供了两行(在略去行 $i$ 后)，然而另外两个摄像机矩阵只提供了一行．这意味着实际上只有三种不同三焦点张量，与选择三个摄像机矩阵中的哪一个提供两行有关．

### 16.2.2 三焦点直线关系

一幅图像上的一条直线用一个协变矢量 $l_i$ 表示，而点 $\mathbf{x}$ 在该直线上的条件是 $l_i x^i = 0$．令 $\mathrm{X}^j$ 表示一个空间点 $\mathbf{X}$，而 $a_j^i$ 表示一个摄像机矩阵 A．该 3D 点 $\mathrm{X}^j$ 映射为图像上的点 $x^i = a_j^i \mathrm{X}^j$．因此点 $\mathrm{X}^j$ 的投影为直线 $l_i$ 上的点的条件是 $l_i a_j^i \mathrm{X}^j = 0$．从另一个角度来看，方程 $l_i a_j^i$ 表示由所有投影到直线 $l_i$ 上的点组成的平面．

考虑一个点 $\mathrm{X}^j$ 映射到一幅图像上的点 $x^i$，同时映射到另两幅图像的直线 $l'_q$ 和 $l''_r$ 上某点的情形．这可以表示为方程

$$x^i = k a_j^i \mathrm{X}^j \qquad l'_q b_j^q \mathrm{X}^j = 0 \qquad l''_r c_j^r \mathrm{X}^j = 0.$$

这三个方程又可以写为如下形式的单个矩阵方程

$$\begin{bmatrix} \mathrm{A} & \mathbf{x} \\ l'_q \mathbf{b}^q & 0 \\ l''_r \mathbf{c}^r & 0 \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -k \end{pmatrix} = \mathbf{0}. \tag{16.14}$$

由这个方程组有一个(非零)解推出 $\det \mathrm{X} = 0$，其中 X 是该方程左边的矩阵．按最后一列展开这个行列式给出

$$0 = -\det \mathrm{X} = \frac{1}{2} x^i \varepsilon_{ilm} \det \begin{bmatrix} \mathbf{a}^l \\ \mathbf{a}^m \\ l'_q \mathbf{b}^q \\ l''_r \mathbf{c}^r \end{bmatrix} = \frac{1}{2} x^i l'_q l''_r \varepsilon_{ilm} \det \begin{bmatrix} \mathbf{a}^l \\ \mathbf{a}^m \\ \mathbf{b}^q \\ \mathbf{c}^r \end{bmatrix}$$

$$= x^i l'_q l''_r \mathcal{T}_i^{qr}. \tag{16.15}$$

此式表明了三焦点张量和这组直线的关系. 这两条直线 $l'_q$ 和 $l''_r$ 反向投影所得的平面在空间相交于一条直线. 这条直线在第一幅图像上的影像是一条直线,它可以表示为 $l_i$. 对该直线的任一点 $x^i$,关系(16.15)都成立. 从而推出 $l'_q l''_r \mathcal{T}_i^{qr}$ 是直线 $l_i$ 的表达式. 因此,对于在三幅图像中的三条对应直线有

$$l_p = l'_q l''_r \mathcal{T}_p^{qr}. \tag{16.16}$$

当然,其两边在相差一个尺度因子的意义下相等. 因为关系(16.16)的两边是矢量,它可以解释为两边的矢量积是零. 用张量 $\epsilon^{ijk}$ 表示这个矢量乘积,我们得到一个方程

$$l_p l'_q l''_r \epsilon^{ipw} \mathcal{T}_i^{qr} = 0^w. \tag{16.17}$$

用推导(16.11)和(16.15)的类似方法,我们可以推导在两幅图像上的对应点和第三幅图像上一条直线之间的关系. 具体地说,如果空间点 $X^j$ 映射到前两幅图像上的点 $x^i$ 和 $x'^j$,和在第三幅图像直线 $l''_r$ 上的某个点,则关系是

$$x^i x'^j l''_r \epsilon_{jqu} \mathcal{T}_i^{qr} = 0_u. \tag{16.18}$$

在这个关系中,$u$ 是自由指标,并且对于 $u = 1,\ldots,3$ 的每一种选择都有一个这样的关系,其中两个是线性独立的.

表 16.1 总结了这一节的结果,其中最后一列表示线性独立方程的数目.

注意不同方程集合是如何相互联系的. 比如,在表中的第二行可以从第一行用直线 $l''_r$ 代替 $x''^k \epsilon_{krv}$ 且消去自由指标 $v$ 而推导出.

| 对应 | 关系 | 方程数 |
|---|---|---|
| 三点 | $x^i x'^j x''^k \epsilon_{jqu} \epsilon_{krv} \mathcal{T}_i^{qr} = 0_{uv}$ | 4 |
| 两点,一直线 | $x^i x'^j l''_r \epsilon_{jqu} \mathcal{T}_i^{qr} = 0_u$ | 2 |
| 一点,两直线 | $x^i l'_q l''_r \mathcal{T}_i^{qr} = 0$ | 1 |
| 三线 | $l_p l'_q l''_r \epsilon^{piw} \mathcal{T}_i^{qr} = 0^w$ | 2 |

表 16.1 三线性关系(同时见表 15.1($p$277)).

### 16.2.3 两视图和三焦点张量之间的关系

到目前为止,我们已经考虑了跨三视图的对应和三焦点张量. 这里我们介绍当对应只跨两幅视图时所产生的约束. 鉴于两视图的点对应能约束基本矩阵,我们期望它也能对 $\mathcal{T}$ 有所约束.

考虑对应点 $x'^j$ 和 $x''^k$ 在第二和第三幅图像上的情形. 它意味着有一个空间点 **X** 映射到点 $x'^j$ 和 $x''^k$. 点 **X** 也映射到第一幅图像上某点 $x^i$,但是 $x^i$ 还不知道. 尽管如此,在这些点之间存在一个线性关系: $x^i x'^j x''^k \epsilon_{jqu} \epsilon_{krv} \mathcal{T}_i^{qr} = 0_{uv}$. 对 $u$ 和 $v$ 的每种选择,记 $A_{i,uv} = x'^j x''^k \epsilon_{jqu} \epsilon_{krv} \mathcal{T}_i^{qr}$. $A_{i,uv}$ 的元素是 $\mathcal{T}_i^{qr}$ 的元素的线性表达式,它可以由已知点 $x'^j$ 和 $x''^k$ 显式地确定. 存在一点 **x** 使得 $x^i A_{i,uv} = 0$. 对 $u$,$v$ 的每种选择,我们可以考虑 $A_{i,uv}$ 为一个以 $i$ 标记的三维矢量,并且对于 $u$ 和 $v$ 的不同选择,存在 4 个线性独立的表达式. 因此,A 可以看作一个$3\times4$的矩阵. 条件 $x^i A_{i,uv} = 0$ 表示 $A_{i,uv}$ 的秩为 2. 这意味着 A 的每个 $3\times3$

的子式是零，它导致三焦点张量元素的一个三次约束. 由于几何原因，方程 $x^i A_{i,uv}$ 对于 $u$ 和 $v$ 的四种选择不是代数独立的. 因此我们从两视图的一组点对应得到关于 $\mathcal{T}_i^{jk}$ 的一个三次约束. 细节留给读者思考.

对于点的对应在第一和第二(或者第三)视图之间的情形，有关分析稍微不同. 然而，每种情形的结论是：虽然跨两视图的一组点对应导致关于三焦点张量的一个约束，但是这个约束不是一个线性约束，与跨三幅视图的点对应时不一样.

### 16.2.4 仿射三焦点张量

当三个摄像机全是仿射时，三焦点张量将满足某些约束. 一个摄像机矩阵若其最后一行为(0,0,0,1)便是仿射的. 由此推出若(16.12)中的矩阵有两行是这种形式，则 $\mathcal{T}_i^{jk}$ 对应的元素是零. 元素 $\mathcal{T}_1^{j3}$，$\mathcal{T}_2^{j3}$，$\mathcal{T}_1^{3k}$，$\mathcal{T}_2^{3k}$ 和 $\mathcal{T}_3^{33}$ 属于这种情形——共有 11 个元素. 因此，在确定到相差一个尺度因子的意义下，三焦点张量包含 16 个非零元素. 如在仿射基本矩阵时一样，这个分析对摄像机共享同一主平面的情形也有效.

## 16.3 四线性关系

对于四幅视图的情形，类似的论证也有效. 考虑跨四幅视图的一组点对应：$\mathbf{x}\leftrightarrow\mathbf{x}'\leftrightarrow\mathbf{x}''\leftrightarrow\mathbf{x}'''$. 当摄像机矩阵为 A,B,C 和 D 时，投影方程可以写为

$$\begin{bmatrix} A & \mathbf{x} & & & \\ B & & \mathbf{x}' & & \\ C & & & \mathbf{x}'' & \\ D & & & & \mathbf{x}''' \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ -k \\ -k' \\ -k'' \\ -k''' \end{bmatrix} = \mathbf{0}. \tag{16.19}$$

由于这个方程有一个(非零)解，其左边的矩阵 X 的秩最多为 7，因此所有 8×8 行列式为零. 与三线性的情形一样，任何一个行列式如果只包含某个摄像机矩阵的一行，则产生关于剩下的视图之间的一个三线性或双线性关系，当 8×8 行列式由每个摄像机矩阵中的两行组成时，出现一种不同的情况. 这样的行列式导致一种新的四线性关系

$$x^i x'^j x''^k x'''^l \varepsilon_{ipw}\varepsilon_{jqx}\varepsilon_{kry}\varepsilon_{lsz} Q^{pqrs} = 0_{wxyz}, \tag{16.20}$$

其中自由变量 $w$，$x$，$y$ 和 $z$ 的每一种选择给出一个不同的方程，并且该四维**四焦点张量** $Q^{pqrs}$ 定义为

$$Q^{pqrs} = \det \begin{bmatrix} \mathbf{a}^p \\ \mathbf{b}^q \\ \mathbf{c}^r \\ \mathbf{d}^s \end{bmatrix}. \tag{16.21}$$

注意这个四视图张量的四个指标是逆变的，并且没有一个视图处于特殊地位，这一点与三焦点张量不一样. 对于给定的四幅视图只有一个四视图张量，并且这个张量产生 81 个不同的四线性关系，其中 16 个是线性独立的(见节 16.6).

如三焦点张量一样，四视图张量也存在直线和点之间的关系. 事实上点的方程只是直

线关系的特例. 然而对于 4 线对应的情形,将产生某些差异,我们现在就来解释这件事. 四直线组和四焦点张量之间关系由如下公式给出

$$l_p l'_q l''_r l'''_s Q^{pqrs} = 0. \tag{16.22}$$

上式适用于任何一组对应直线:$l_p$、$l'_q$、$l''_r$ 和 $l'''_s$. 然而,此式的推导表明:只要存在一个空间点投射到四条图像直线上,这个条件就成立. 它并不需要四条图像直线是对应的(即指它们是一条共同空间直线的像). 这种配置在图 16.1a 中作了说明.

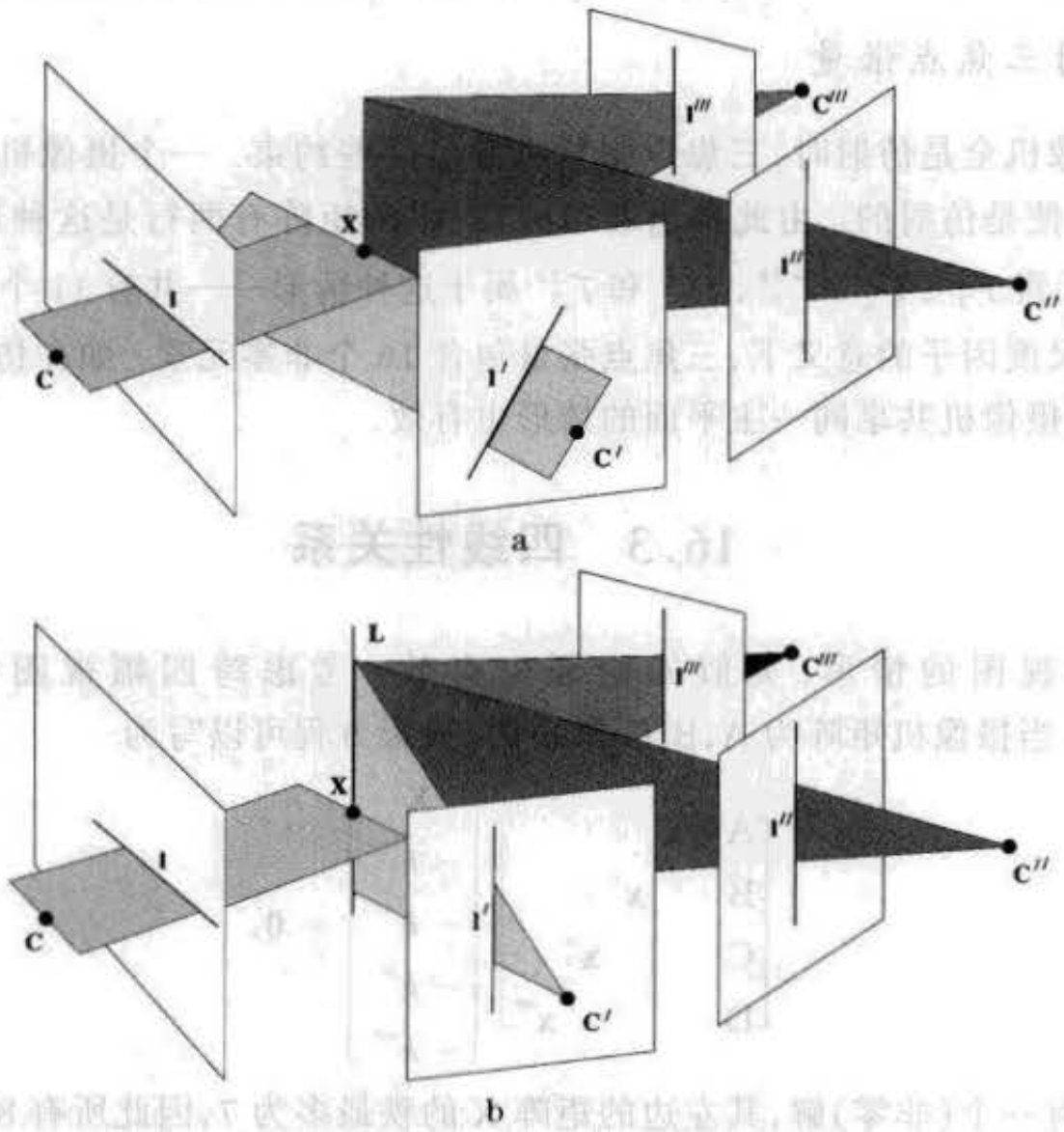



图 16.1 **四直线"对应"**. 四条直线 $\mathbf{l}$、$\mathbf{l}'$、$\mathbf{l}''$、$\mathbf{l}'''$满足四线性关系(16.22),因为它们的反向投影交于一个公共点 $\mathbf{X}$. (a)没有三张平面交于一条公共直线. (b)三条直线 $\mathbf{l}' \leftrightarrow \mathbf{l}'' \leftrightarrow \mathbf{l}'''$是 3 维空间中同一条直线 $\mathbf{L}$ 的影像.

现在考虑从单条 3D 直线得到的三条直线对应的情形(比如 $l'_q$, $l''_r$ 和 $l'''_s$)(图 16.1b). 现在令 $l_p$ 为第一幅图像上的**任意**一条直线,这条直线的反向投影是一张平面,它将与这条 3D 直线相交于一点 $\mathbf{X}$,并且所给出的条件使(16.22)成立. 由于它对任何直线 $l_p$ 都成立,因此必然有 $l'_q l''_r l'''_s Q^{pqrs} = 0^p$. 它给出了涉及 $l'_q$, $l''_r$和 $l'''_s$的三个线性独立方程. 然而,如上文,在四幅图像上给出对应直线的一个集合,我们可以选择它的一个三直线的子集,并且对每个三直线子集用这种方式得到三个方程. 因为存在着四种三直线子集的选择,因此方程的总数是 12.

然而,这些方程中只有 9 个是独立的,具体推导如下:假定直线为 $\mathbf{l} = \mathbf{l}' = \mathbf{l}'' = \mathbf{l}''' = (1,0,0)^{\mathrm{T}}$,这等价于一般情形,因为分别对四幅图像作射影变换,可以把任何一条直线对应变换成这种情形. 现在方程 $l'_q l''_r l'''_s Q^{pqrs} = 0^p$ 意味着任何元素 $Q^{p111} = 0$. 把这个推导运用于三幅视图中的所有四个三直线子集,我们得到当至少有三个指标为 1 时 $Q^{pqrs}$

$=0$. 总共有 9 个这样的元素. 因为由对应直线产生的方程组等价于取这些元素的每一个为零,所以在总共 12 个方程中只有 9 个是独立的.

表 16.2 总结了四视图关系. 这里对于三条直线和一点的情况没有给出方程,因为对张量而言,它并没有比仅是三条直线对应的情形给出更多的限制.

| 对应 | 关系 | 方程数 |
|---|---|---|
| 四点 | $x^i x'^j x''^k x'''^l \varepsilon_{ipw}\varepsilon_{jqx}\varepsilon_{kry}\varepsilon_{lsz} Q^{pqrs} = 0_{wxyz}$ | 16 |
| 三点,一直线 | $x^i x'^j x''^k l'''_s \varepsilon_{ipw}\varepsilon_{jqx}\varepsilon_{kry} Q^{pqrs} = 0_{wxy}$ | 8 |
| 两点,两直线 | $x^i x'^j l''_r l'''_s \varepsilon_{ipw}\varepsilon_{jqx} Q^{pqrs} = 0_{wx}$ | 4 |
| 三条直线 | $l_p l'_q l''_r Q^{pqrs} = 0^s$ | 3 |
| 四条直线 | $l_p l'_q l''_r Q^{pqrs} = 0^s, l_p l'_q l'''_s Q^{pqrs} = 0^r, \cdots$ | 9 |

表 16.2 四线性关系.

## 16.4 四张平面的交

多视图张量可以由另一种不同的推导给出,这种推导使它们的意义显示得更加清楚. 在这个解释中,最基本的几何性质是四张平面的相交. 空间的四张平面一般不交于一公共点. 使得它们交于一点的一个充要条件是由表示这些平面的矢量组成的 $4\times4$ 矩阵的行列式等于零.

**记号** 我们用 $\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d}$ 来表示由行矢量为 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$、$\mathbf{d}$ 组成的一个 $4\times4$ 矩阵的行列式,但仅限于在本节使用. 在更一般范围内,符号 $\wedge$ 表示在双代数中的相交算子(见本章文献). 然而,为了当前的需要读者只需把它看成行列式的缩写.

我们从四焦点张量开始,因为对它的推导最简单. 考虑由摄像机矩阵为 A、B、C 和 D 的四个摄像机形成的图像上四条直线 $\mathbf{l}$、$\mathbf{l}'$、$\mathbf{l}''$ 和 $\mathbf{l}'''$. 直线 $\mathbf{l}$ 通过摄像机 A 的反向投影记为平面 $l_i\mathbf{a}^i$,其记号如(16.14). 使这四张平面相交于一点的条件可以记为

$$(l_p\mathbf{a}^p)\wedge(l'_q\mathbf{b}^q)\wedge(l''_r\mathbf{c}^r)\wedge(l'''_s\mathbf{d}^s)=0.$$

然而,由于行列式关于每一行是线性的,此式又可以写为

$$0 = l_p l'_q l''_r l'''_s(\mathbf{a}^p\wedge\mathbf{b}^q\wedge\mathbf{c}^r\wedge\mathbf{d}^s) \stackrel{\text{def}}{=} l_p l'_q l''_r l'''_s Q^{pqrs}. \tag{16.23}$$

这对应于定义(16.21)和关于四焦点张量的直线关系(16.22). 其基本的几何性质是在空间中这四张平面相交.

**三焦点张量的推导** 现在来考虑三视图的点—直线—直线关系 $x^i\leftrightarrow l'_j\leftrightarrow l''_k$,并且令 $l^1_p$ 和 $l^2_q$ 为在第一幅图像上过像点 $\mathbf{x}$ 的两条直线. 通过这四条直线反向投影的平面相交于一点(见图 16.2). 因此我们可以写出

$$l^1_l l^2_m l'_q l''_r(\mathbf{a}^l\wedge\mathbf{a}^m\wedge\mathbf{b}^q\wedge\mathbf{c}^r)=0.$$

下面一步是一个代数技巧——把这个方程乘以 $\varepsilon^{ilm}\varepsilon_{ilm}$,这是一个纯量值(事实上等于 6,即 $(ilm)$ 的排列数). 重新组合后的结果是

$$(l^1_l l^2_m\varepsilon^{ilm})l'_q l''_r\varepsilon_{ilm}(\mathbf{a}^l\wedge\mathbf{a}^m\wedge\mathbf{b}^q\wedge\mathbf{c}^r)=0.$$

现在注意：表达式 $l_l^1 l_m^2 \varepsilon^{ilm}$ 不过是两条直线 $l_1$ 和 $l_m$ 的叉积，换句话说，是两直线的交点 $x^i$，因此最后我们可以写为

$$0 = x^i l'_q l''_r (\varepsilon_{ilm}(\mathbf{a}^l \wedge \mathbf{a}^m \wedge \mathbf{b}^q \wedge \mathbf{c}^r)) \stackrel{\text{def}}{=} x^i l'_q l''_r T_i^{qr} \tag{16.24}$$

它正是定义(16.10)和三焦点张量的基本关联关系(16.15).

**基本矩阵** 我们能用同样的方法推导基本矩阵，给定一组点对应 $\mathbf{x} \leftrightarrow \mathbf{x}'$，选过 $\mathbf{x}$ 的一对直线 $l_p^1$ 和 $l_q^2$ 以及过 $\mathbf{x}'$ 的 $l_r'^1$ 和 $l_s'^2$. 这些直线反向投影所得的平面都相交于一点，因此我们有

$$l_p^1 l_q^2 l_r'^1 l_s'^2 (\mathbf{a}^p \wedge \mathbf{a}^q \wedge \mathbf{b}^r \wedge \mathbf{b}^s) = 0.$$

乘以$(\varepsilon_{ipq}\varepsilon^{ipq})(\varepsilon_{jrs}\varepsilon^{jrs})$且按上面的推导过程得到共面约束

$$0 = x^i x'^j (\varepsilon_{ipq}\varepsilon_{jrs}(\mathbf{a}^p \wedge \mathbf{a}^q \wedge \mathbf{b}^r \wedge \mathbf{b}^s)) \stackrel{\text{def}}{=} x^i x'^j F_{ji} \tag{16.25}$$

请把它与(16.4 - $p$292)相比较.

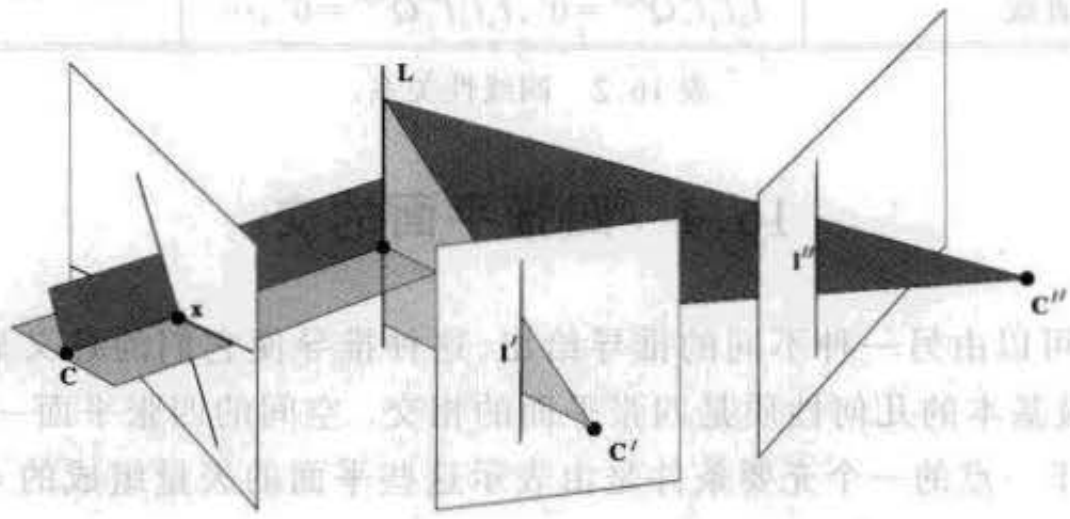



图 16.2 涉及三幅图像的点—直线—直线对应 $\mathbf{x} \leftrightarrow \mathbf{l}' \leftrightarrow \mathbf{l}''$ 可以解释如下：在第一幅图上选择过点 $\mathbf{x}$ 的任意两条直线，然后把四条直线反向投影成在空间交于一点的平面.

## 16.5 计数的推导

本节我们将确定从若干视图实现重构所需要的点或直线的数目. 这个分析与计算相应的张量的自由度数有关. 然而，在进行此分析时，有必要把当张量看作是一个没有约束的代数物体时的自由度数和由摄像机及其矩阵组成的一个配置所产生的自由度数加以区分.

比如，考虑基本矩阵时，在一个层面上，基本矩阵可以看作是一个 $3 \times 3$ 齐次矩阵，因此有 8 个自由度(9 减去 1，因为尺度因子不确定). 另一方面，根据(16.3 - $p$292)从一对摄像机矩阵产生的基本矩阵必须满足附加的约束 $\det F = 0$. 因此，这样的基本矩阵只有 7 个自由度. 由于摄像机矩阵可以在相差一个 3D 射影变换下由基本矩阵确定(且反过来也成立)，我们可以通过计算摄像机矩阵的自由度来计算基本矩阵的自由度数. 两个摄像机矩阵共有 22 个自由度(两个齐次 $3 \times 4$ 矩阵). 一个 3D 单应可以用一个 $4 \times 4$ 齐次矩阵表示，从而有 15 个自由度. 这给出了两个摄像机配置在相差一个射影变换下共有 $7 = 22 - 15$ 个自由度，这与基本矩阵的 7 个自由度相符合，从而提供了关于前一个计算的一种核查.

类似地，三焦点张量蕴藏了三个摄像机矩阵的射影结构，从而有 $18 = 3 * 11 - 15$ 个自由度. 同理四焦点张量有 $29 = 4 * 11 - 15$ 个自由度. 一般地，对于 $m$ 个摄像机，我们有

$$\# \text{dof} = 11m - 15.$$

由于三焦点和四焦点张量作为齐次代数阵列考虑时分别有 26 和 80 个自由度，它们必须满

足由摄像机几何推出的一组附加的约束——对三焦点张量有 8 个约束，对四焦点张量有 51 个约束.

在计算几何结构时，我们可以利用基于多线性求解约束方程来估计多焦点张量的线性代数方法. 所需要的对应数目由每组点或直线对应产生的方程数目决定. **线性方法**没有考虑几何加在张量上的约束.

另一方面，所需要的对应数目可以由计算每一组对应给出的几何约束的数目并且与整个系统的自由度的总数比较来确定. 考虑由 $m$ 幅视图中的 $n$ 点组成的一个配置. 这个系统的自由度总数是 $11m-15+3n$，因为每个 3D 点有 3 个自由度. 为了估计射影结构，可用的数据是在 $m$ 幅图像上的 $n$ 个点，总共有 $2mn$ 个测量值(每个 2D 点有两个坐标). 这样一来，为了使重构成为可能，我们要求 $2mn \geq 11m-15+3n$ 或 $(2m-3)n \geq 11m-15$. 因此所需要的点的个数是

$$n \geq \frac{11m-15}{2m-3} = 5 + \frac{m}{2m-3}.$$

我们也可以这样来理解上式，每一组对应点对摄像机贡献 $2m-3$ 个约束，即每幅视图上点的坐标是 $2m$，减去 3D 点的 3 个自由度.

类似的论证也适用于直线对应. 在 3D 空间中直线有 4 个自由度，而它的图像直线由 2 个自由度描述，因此每组直线对应提供 $2m-4$ 个约束，从而所要求的直线的数目是

$$l \geq \frac{11m-15}{2m-4}.$$

对于上述任一种情形，如果约束(方程)的数目等于摄像机和点或直线配置的(未知的)自由度数，则一般我们能得到多个解，除非是线性的情况. 然而，如果方程数多于未知数，这个系统是超定的，并且一般存在一个解.

| # 视图 | 张量 | # 元素 | # 自由度 | 线性 | | 非线性 | |
|---|---|---|---|---|---|---|---|
| | | | | # 点 | # 线 | # 点 | # 线 |
| 2 | $F$ | 9 | 7 | 8 | — | 7* | — |
| 3 | $\mathcal{T}$ | 27 | 18 | 7 | 13 | 6* | 9*? |
| 4 | Q | 81 | 29 | 6 | 9 | 6 | 8? |

表 16.3 **射影自由度和约束** "线性"列给出线性求解张量(精确到尺度因子)所需的跨所有视图的对应的最小数目. "非线性"列给出所需的对应的最小数目. 星号表示有多个解，而问号表示没有已知的实用重构算法.

表 16.3 总结了重构所需的对应数. 星号(*)表示可能有多解. 对非线性解的情形，实际求解的方法并不都知道，除非采用硬做的方法以及解联立多项式方程组. 对不知是否有简单方法的情形，用一个问号作标记. 注意关于直线对应的非线性解法知道的还不多. 具体的非线性算法是：

(1) 2 视图的 7 点算法：见节 10.1.2($p$192)，可能有三个解.

(2) 3 视图的 6 点算法：见节 19.2.4($p$360)，可能有三个解. 这是前一个问题的对偶(见第 19 章).

(3) 4 视图的 6 点算法. 先用 3 视图的 6 点算法求解，然后应用求摄像机矩阵的 DLT

算法(节 6.1($p$121))求第四幅视图,并且只保留一个解而消去其他的解. 然而,不象前两种情形,4 视图的 6 点算法是超定的,并且这个解只是对完全正确的数据有效. 对噪声数据的估计在第 19 章讨论.

### 16.5.1 仿射摄像机

仿射摄像机重构所需要的对应较少. 在重构中无穷远平面可以被辨认,它是所有摄像机的主平面,且重构的多义性是仿射而不是射影的. 当有 $m$ 个仿射摄像机时,自由度的数目为

$$\# \text{ dof} = 8m - 12.$$

对一个新的 $3\times4$ 仿射摄像机矩阵,每幅视图增加 8 个自由度,我们减去 12 是针对一个 3D 仿射变换的自由度.

和射影摄像机一样,一组点对应提供 $2m-3$ 个约束,而一组直线对应提供 $2m-4$ 个约束. 与前面一样,所需要的点的个数可以计算为 $n(2m-3)\geq 8m-12$,即

$$n \geq \frac{8m-12}{2m-3} = 4.$$

所需的直线的条数则是

$$l \geq \frac{8m-12}{2m-4} = 4 + \frac{2}{m-2}.$$

与线性算法情形一样,张量中元素的个数是 $3^m$. 在仿射情形,张量在相差一个尺度因子的意义下元素的个数不变,但是却有许多元素是零,如同我们在节 16.1.2 和节 16.2.4 所看到的那样. 这一性质减少了所需要的对应数. 计数结果在表 16.4 中给出. 注意对于点对应,线性算法适用于由上面方程给定的最小对应数. 因而,非线性算法同线性算法一样.

| # 视图 | 张量 | # 非零元素 | # 自由度 | 线性 | | 非线性 | |
|---|---|---|---|---|---|---|---|
| | | | | #点 | #线 | #点 | #线 |
| 2 | $F_A$ | 5 | 4 | 4 | — | 4 | — |
| 3 | $\mathcal{T}_A$ | 16 | 12 | 4 | 8 | 4 | 6*? |
| 4 | $Q_A$ | 48 | 20 | 4 | 6 | 4 | 5*? |

表 16.4 **仿射自由度和约束**. 每幅视图的摄象机是仿射的. 细节请看表 16.3 的说明.

### 16.5.2 已知四个共面点——平面加视差

前面关于计数的论证是针对在一般位置上的点和直线的情形. 现在我们考虑一种重要情形,即四个或更多个 3D 点共面的情形. 我们将看到张量以及射影结构的计算被显著地简化. 这个讨论基于图像上四个共面点已知. 但是,这里真正重要的是由这个平面诱导的图像之间的单应应该已知.

已知四个共面点计算基本矩阵在节 12.3($p$235)中考虑过(见算法 12.2($p$237)),而关于三焦点张量的计算则在节 15.7.1 中考虑过. 现在我们考虑四幅视图的情形. 利用 3D 中一张平面上的点推导出 4(或更多)组点对应之后,我们能计算由该平面诱导的从第一幅视图分别到第二、第三和第四幅视图的单应性 $H'$,$H''$和 $H'''$. 在射影重构中,我们选择包含这

些点的平面为无穷远平面的情形，此时 $H'$，$H''$ 和 $H'''$ 是无穷单应．在进一步假设第一幅图像过世界坐标原点的条件下，这四个摄像机矩阵现在可以记为

$$A = [I \mid \mathbf{0}] \quad B = [H' \mid \mathbf{t}'] \quad C = [H'' \mid \mathbf{t}''] \quad D = [H''' \mid \mathbf{t}'''].$$

向量 $\mathbf{t}'$，$\mathbf{t}''$，$\mathbf{t}'''$ 在相差一个公共尺度因子的意义下可以被确定．

由于这些摄像机矩阵的左边 $3\times3$ 子矩阵块现在是知道的，从(16.21)不难看出 $Q^{pqrs}$ 的元素关于摄像机矩阵剩下的元素 $\mathbf{t}'$，$\mathbf{t}''$，$\mathbf{t}'''$ 是线性的．利用(16.21)我们可以写出用 $\mathbf{t}'$，$\mathbf{t}''$，$\mathbf{t}'''$ 的元素表示 $Q^{pqrs}$ 的元素的一个表示式．事实上，我们可以把它显式地记为 $\mathbf{q}=\mathbf{M}\mathbf{t}$，其中 M 是一个 $81\times9$ 矩阵，$\mathbf{q}$ 和 $\mathbf{t}$ 是表示 $Q$ 和 $\mathbf{t}'$，$\mathbf{t}''$，$\mathbf{t}'''$ 的元素的矢量．因此，该四焦点张量可以用 9 个齐次坐标线性地参数化，从而有 8 个自由度．

给定了一组方程 $E\mathbf{q}=\mathbf{0}$(根据表 16.2 由对应推得)，通过替换 $\mathbf{q}=\mathbf{M}\mathbf{t}$，它们可以根据最小参数集 $\mathbf{t}$ 重写为 $E\mathbf{M}\mathbf{t}=\mathbf{0}$．这个方程可给出关于 $\mathbf{t}$ 的线性解，并由此得到摄像机矩阵和(如果希望)张量 $\mathbf{q}=\mathbf{M}\mathbf{t}$ 的线性解．注意这里突出的优点是这样得到的张量自动地对应于一组摄像机矩阵，从而满足所有几何约束——这样一来，要求四焦点张量满足 51 个累赘的约束便消失殆尽了．

以上分析是针对计算四焦点张量的，但也同样能应用于基本矩阵和三焦点张量．

**计数推导** 我们回到 $m$ 幅视图的一般情形并从几何的观点来考虑它．参数化张量的自由度数等于摄像机矩阵中余下的几何自由度数，即 $3m-4$，其中 $m$ 是视图数．这由除了第一个摄像机外每个摄象机矩阵的最后一列的自由度总数 $3(m-1)$ 减去一个公共的尺度因子得到．

然而，计算由点和直线对应提供的约束则需一点技巧，并必须小心对待计数推理，切莫忽视隐藏的相关性．首先，考虑跨三幅视图的直线对应 $\mathbf{l}\leftrightarrow\mathbf{l}'\leftrightarrow\mathbf{l}''$——这个推导对于 $m\geq3$ 视图一般将成立．需要解决的问题是从图像直线的测量能推导出多少信息．令人吃惊的是知道图像之间的平面单应会减少由一条直线对应所提供的信息．为简化论证，我们可以假定对图像已经施加了平面单应使得四个共面参考点映射到每一幅图像的相同点．其结果是在参考平面上任何其他的点也将映射到每一幅图像的同一个点．所考虑的直线对应的 3D 直线 $\mathbf{L}$ 必然与参考平面相交于一点 $\mathbf{X}$．由于 $\mathbf{X}$ 在参考平面上，$\mathbf{X}$ 投影到所有图的同一点 $\mathbf{x}$，并且该 $\mathbf{x}$ 必然在所有三条像直线 $\mathbf{l}$，$\mathbf{l}'$，$\mathbf{l}''$ 上．因此，在三幅视图中的对应直线不可能是任意的——它们必须都通过一个公共的图像点．一般在 $m$ 视图上，测量直线的自由度数是 $m+2$．即确定点 $\mathbf{x}$ 需要 2 个自由度，而通过这个点的每条直线只剩下一个自由度(它的朝向)．减去空间直线的 4 个自由度，我们得到

- **跨 $m$ 幅视图的每一组直线对应对摄像机余下的自由度产生 $m-2$ 个约束．**

请注意图像上直线必须相交于一点的条件是怎样限制直线对应产生的方程数．没有观察到这个条件时，人们会指望由每条直线得到 $2m-4$ 个约束．然而，在数据完全正确时，$2m-4$ 个方程组的秩将只有 $m-2$．对于有噪声的数据，图像直线将不会完全相容，因而这个系统可能是满秩的．但是这完全是由于噪声所致，并且这个系统的最小奇异值在本质上是随机的．尽管如此，在解这个系统时，我们应该包含所有可用的方程，因为它会减少噪声影响．

关于点对应的推导是类似的．通过任何两个像点的直线是空间直线的影像．这条 3D

直线的投影的约束在前面已经讨论过并且给匹配点施加了约束. 测量的自由度是 $3m+2$:在每一幅视图中,2 个用于直线与平面的交点,1 个用于直线的定向;而其中的 2 对应于在直线上两个点各自的位置. 减去两个 3D 空间点的 6 个自由度,结果是

- **跨 $m$ 幅视图的两组点对应为摄像机余下的自由度产生 $3m-4$ 个约束.**

因为此数与摄像机的自由度数是一样的,因此 2 个点足以计算结构.

由于对摄像机的几何约束数与张量的自由度数相同,因此对张量的约束与对摄像机的几何约束没有区别. 这样一来,点和直线对应可以用来产生对张量的线性约束,没有必要用非线性方法. 所需要对应数总结在表 16.5 中.

| # 视图 | 张量 | # 自由度 | 点 | | 线 | |
|---|---|---|---|---|---|---|
| | | | #点 | #约束 | #线 | #约束 |
| 2 | F | 2 | 2 | 2 | — | |
| 3 | $T$ | 5 | 2 | 5 | 5 | 5 |
| 4 | Q | 8 | 2 | 8 | 4 | 8 |

表 16.5 给定视图之间由一个平面诱导 2D 的单应后,为计算射影结构还需的对应数. 该单应可以由共面的四点或更多点的匹配或者通过任何其他手段计算得到. 用于计算该单应的点没有统计在表中.

## 16.6 独立方程数

表 16.2 断言跨四视图的每一组点对应给出关于四焦点张量元素的 16 个线性独立方程. 我们现在来更仔细地讨论它.

给定跨四幅视图的足够多的点匹配,我们便可以求解张量 $Q^{pqrs}$. 而一旦张量 $Q$ 已知就可以求解摄像机矩阵,以及计算射影结构,这在[Heyden-95a, Heyden-97c, Hartley-98b]中给出了证明,但在本书中未作进一步讨论. 然后当有人计算它所必需的点匹配数时,一个奇怪的现象发生了. 正如上面指出的:每组点匹配给出关于张量 $Q^{pqrs}$ 元素的 16 个线性独立方程,并且似乎从两个完全不相关的点对应集合推导的方程不可能有任何相关性. 因此好像从五组点对应就可以得到 80 个方程,并足以在相差一个尺度因子的意义下求解 $Q^{pqrs}$ 的 81 个元素. 根据这个推论,似乎仅用跨四视图的五组点匹配便可以解出该张量,从而,在相差一个射影多义性下可以解出摄像机矩阵,但是这个结论与下面的观察相矛盾.

- **不可能从五点的图像确定四个(或任何数目的)摄像机的位置.**

它由节 16.5 的计数论证得到. 显而易见上面的计数出了错. 原理包括在下文中.

**结论 16.1 由跨四幅视图的单组点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'\leftrightarrow\mathbf{x}''\leftrightarrow\mathbf{x}'''$ 导出的含有 81 个方程的线性方程组(16.20-$p$297)包含关于 $Q^{pqrs}$ 的 16 个独立约束. 而且,若此方程记为 $\mathrm{A}\mathbf{q}=\mathbf{0}$,其中 A 是一个 $81\times81$ 矩阵,而 $\mathbf{q}$ 是一个包含 $Q^{pqrs}$ 元素的矢量. 则 A 的 16 个非零的奇异值都相等.**

这个结论所说的是:从一组点对应的确可以得到 16 个线性独立的方程,但可以通过正交变换(在方程 A 的左边乘以正交矩阵 U)使这个方程组事实上简化为 16 个正交方程的集合. 这个证明在本节末给出. 然而令人吃惊的事实是对应于两个不关联的点对应的方程组有相关性,如下面的结论所述.

**结论 16.2　由跨 4 幅视图 $n$ 组一般点对应集合导出的方程组(16.20 - $p$297)的秩为 $16n-\binom{n}{2}$,其中 $n\leq 5$.**

记号$\binom{n}{2}$表示从 $n$ 个取 2 个的组合数. 具体地说,$\binom{n}{2}=n(n-1)/2$. 因此对于 5 个点只有 70 个独立方程,不够求解 $Q^{pqrs}$. 对于 $n=6$ 个点,$16n-\binom{n}{2}=81$,我们有足够的方程求解 $Q^{pqrs}$ 的 81 个元素.

我们现在证明上面的两个结论. 证明结论 16.1 的关键是涉及到反对称矩阵的奇异值(见结论 $A3.1(p406)$).

**结论 16.3　一个 3×3 反对称矩阵有两个相等的非零奇异值.**

关于结论 16.1 余下的证明是非常直接的,只要不在记号中迷失.

**证明**(结论 16.1)　从单组点对应导出的整共 81 个方程的形式是 $x^i\varepsilon_{ipw}x'^j\varepsilon_{jqx}x''^k\varepsilon_{kry}x'''^l\varepsilon_{lsz}Q^{pqrs}=0_{wxyz}$. 由 $w,x,y,z$ 在 $1,\ldots,3$ 的范围内变化总共产生 81 个方程. 因此,方程矩阵 A 可以写为

$$\mathrm{A}_{(wxyz)(pqrs)}=x^i\varepsilon_{ipw}x'^j\varepsilon_{jqx}x''^k\varepsilon_{kry}x'''^l\varepsilon_{lsz} \tag{16.26}$$

其中指标$(wxyz)$指示 A 的行而$(pqrs)$指示 A 的列. 我们把一组指标,比如在此情形下的$(wxyz)$视为一个矩阵的行或者列的单个指标. 这些指标将用括号括起来,并且称它们为一个**组合指标**.

现在考虑表达式 $x^i\varepsilon_{ipw}$,此式可以视为由自由指标 $p$ 和 $w$ 所表示的矩阵. 进一步,因为 $x^i\varepsilon_{ipw}=-x^i\varepsilon_{iwp}$,可知它是一个反对称矩阵,从而有相同的奇异值. 我们标记这个矩阵为 $\mathrm{S}_{wp}$. 利用张量把结论 16.3 记为

$$\mathrm{U}_a^w\mathrm{S}_{wp}\mathrm{V}_e^p=k\mathrm{D}_{ae} \tag{16.27}$$

其中对角矩阵 D 具有结论 16.3 所述的性质. (16.26)中的矩阵 A 可以记为 $\mathrm{A}_{(wxyz)(pqrs)}=\mathrm{S}_{wp}\mathrm{S}'_{xq}\mathrm{S}''_{yr}\mathrm{S}'''_{zs}$. 因此,利用(16.27)我们有

$$\mathrm{U}_a^w\mathrm{U}_b'^x\mathrm{U}_c''^y\mathrm{U}_d'''^z\mathrm{A}_{(wxyz)(pqrs)}\mathrm{V}_e^p\mathrm{V}_f'^q\mathrm{V}_g''^r\mathrm{V}_h'''^s=kk'k''k'''\mathrm{D}_{ae}\mathrm{D}_{bf}\mathrm{D}_{cg}\mathrm{D}_{dh}. \tag{16.28}$$

现在,令

$$\hat{\mathrm{U}}^{(wxyz)}_{(abcd)}=\mathrm{U}_a^w\mathrm{U}_b'^x\mathrm{U}_c''^y\mathrm{U}_d'''^z\qquad \hat{\mathrm{V}}^{(pqrs)}_{(efgh)}=\mathrm{V}_e^p\mathrm{V}_f'^q\mathrm{V}_g''^r\mathrm{V}_h'''^s\qquad \hat{\mathrm{D}}_{(abcd)(efgh)}=\mathrm{D}_{ae}\mathrm{D}_{bf}\mathrm{D}_{cg}\mathrm{D}_{dh}$$

和 $\hat{k}=kk'k''k'''$,则(16.28)可以写为

$$\hat{\mathrm{U}}^{(wxyz)}_{(abcd)}\mathrm{A}_{(wxyz)(pqrs)}\hat{\mathrm{V}}^{(pqrs)}_{(efgh)}=\hat{k}\,\hat{\mathrm{D}}_{(abcd)(efgh)}. \tag{16.29}$$

矩阵 $\hat{\mathrm{D}}_{(abcd)(efgh)}$ 是有 16 个都等于 1 的非零对角元素的对角矩阵. 证明了(16.29)是矩阵 $A_{(wxyz)(pqrs)}$ 的 SVD 分解,就可完成证明,现在只剩下证明 $\hat{\mathrm{U}}^{(wxyz)}_{(abcd)}$ 和 $\hat{\mathrm{V}}^{(pqrs)}_{(efgh)}$ 是正交矩阵. 而这只要证明其列是标准正交的即可. 这证明是一目了然的,留作一个习题. □

**证明**(结论 16.2)　考虑跨四幅视图的两组点对应:$x^i\leftrightarrow x'^i\leftrightarrow x''^i\leftrightarrow x'''^i$ 和 $y^i\leftrightarrow y'^i\leftrightarrow y''^i\leftrightarrow y'''^i$. 它们产生形如(16.26)的两个方程组 $\mathrm{A}^x\mathbf{q}=\mathbf{0}$ 和 $\mathrm{A}^y\mathbf{q}=\mathbf{0}$. 每个系数矩阵的秩为 16,若 $\mathrm{A}^x$ 的行矢量与 $\mathrm{A}^y$ 的行矢量是无关的,则组合的方程组的秩是 32. 然而,若 $\mathrm{A}^x$ 和 $\mathrm{A}^y$ 的行矢量之

间是线性相关的，则它们组合的秩最多是 31.

我们定义具有组合指标($pqrs$)的矢量 $\mathbf{s}_x$ 为 $\mathbf{s}_x^{(pqrs)} = x^p x'^q x''^r x'''^s$. 类似地定义矢量 $\mathbf{s}_y$. 我们将证明 $\mathbf{s}_y^{\mathsf{T}}\mathrm{A}^x = \mathbf{s}_x^{\mathsf{T}}\mathrm{A}^y$，此式意味着 $\mathrm{A}^x$ 和 $\mathrm{A}^y$ 的行是线性相关的. 将 $\mathbf{s}_y^{\mathsf{T}}\mathrm{A}^x$ 展开得

$$
\begin{aligned}
\mathbf{s}_y{}^{\mathsf{T}}\mathrm{A}^x &= \mathbf{s}_y{}^{(wxyz)}\mathrm{A}^x_{(wxyz)(pqrs)} \\
&= (y^w y''^x y'''^y y''''^z)x'^j\varepsilon_{ipw}x^i\varepsilon_{jqx}x''^k\varepsilon_{kry}x'''^l\varepsilon_{lsz} \\
&= (x^i x''^j x'''^k x''''^l)y^w\varepsilon_{wpi}y'^x\varepsilon_{xqj}y''^y\varepsilon_{yrk}y'''^z\varepsilon_{zsl} \\
&= \mathbf{s}_x^{\mathsf{T}}\mathrm{A}^y.
\end{aligned}
$$

该式说明 $\mathrm{A}^x$ 和 $\mathrm{A}^y$ 的行矢量是相关的，从而它们组合的秩最多为 31. 我们现在考虑该组合的秩比 31 小的可能性. 若该秩小于 31，则矩阵$[\mathrm{A}^x, \mathrm{A}^y]$的所有 $31\times 31$ 子行列式必须为零. 这些子行列式可以表示为关于点 $\mathbf{x}, \mathbf{x}', \mathbf{x}'', \mathbf{x}''', \mathbf{y}, \mathbf{y}', \mathbf{y}''$和 $\mathbf{y}'''$的系数多项式. 这 24 个系数一起生成一个 24 维空间. 因此，对某个 $N$(等于这样的 $31\times 31$ 子行列式的个数)存在一个函数 $f: \mathbb{R}^{24}\to\mathbb{R}^N$，使得该方程矩阵只在函数 $f$ 的零点集上的秩小于 31. 任意选择的例子(略去)都可以用来证明函数 $f$ 不恒等于零. 由此推出使方程组的秩小于 31 的点对应的集合是 $\mathbb{R}^{24}$中的一个代数簇，因而是无处稠密的. 因此，由跨四视图的一般位置的一对点的对应所产生的方程组的秩为 31.

我们现在回到跨所有四幅视图的 $n$ 组点对应的一般情形. 注意对两组点对应成立的两个线性关系不是一般性的，而是依赖于这一对点对应. 因此，一般给定 $n$ 组点对应时，将有 $\binom{n}{2}$这样的关系. 这就把由该方程组生成的空间的维数减少到 $16n-\binom{n}{2}$，这正是我们需要的. □

**三视图情况** 对三视图也有类似的推导，用同样方法可以证明由一组匹配点产生的 9 个方程只有 4 个是独立的. 我们把它的证明留作习题(见页 308).

## 16.7 方程选取

节 16.6 证明：由四点方程推导的整个方程组的奇异值都相等. 这个证明可以容易地用于三视图的情形(见习题，$p$308). 推导的关键点是 $3\times 3$ 反对称矩阵的两个非零的奇异值相等. 显然这个证明也可以推广到节 16.2 和节 16.3 给出的由点或直线对应导出的任何其他方程组.

我们仍考虑四视图情形. 有关奇异值的结论表明一般应该包含由这组对应所导出的全部 81 个方程，而不是只选择 16 个独立方程. 这样做将避免在接近奇异情形时产生困难. 这一点被实验观察所支持. 数值的例子确实证明由几组点对应推导的方程组在包含了每组点对应的全部方程时条件数大为改善. 这里方程组的条件数是指第一个(最大的)与第 $n$ 个奇异值的比值，其中 $n$ 是线性独立方程的数目.

包含全部 81 个方程而不仅是 16 个，意味着方程组比较大，导致解的复杂性增加. 这可以按如下方法进行补救. 奇异值相等的基础是 $\mathrm{S}_{wp} = x^i\varepsilon_{ipw}$，以及其他类似定义的项是反对称矩阵，从而有相同的奇异值. 对具有相等奇异值的其他任何矩阵 S 也有同样的效应. 我们只要求 S 的列矢量必须表示通过点 $\mathbf{x}$ 的直线(或者表为 $\mathrm{S}_{wp}x^w = 0_p$). 如果矩阵的 S 列矢

量是标准正交的，则它有相等的奇异值. 当 S 是一个 $3\times2$ 矩阵时就可以得到这些条件. 如果对每幅视图中的点都这样做了，方程的总数将从 $3^4=81$ 减少到 $2^4=16$，而且这 16 个方程将是正交的. 如下文所示，选择 S 的一种方便方法是利用 Householder 矩阵(见节 A3.1.2 ($p$405)).

这个讨论也适用于三焦点张量，即允许把方程的个数从 9 减少到 4，同时保留相等的奇异值. 我们针对三焦点张量总结这个讨论：

- **给定跨三幅视图的点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'\leftrightarrow\mathbf{x}''$，产生如下形式的方程组**

$$x^i\hat{l}'_{qx}\hat{l}''_{ry}\mathcal{T}_i^{qr}=0_{xy}\ \text{对}\ x,y=1,2.$$

**其中 $\hat{l}'_{q1}$ 和 $\hat{l}'_{q2}$ 是由标准正交向量表示的过 $\mathbf{x}'$ 的两条直线(对 $\hat{k}''_{ry}$ 类似). 求 $\hat{l}'_{qx}$ 和 $\hat{l}''_{ry}$ 的一种方便的方法如下：**

(1) **寻找** Householder **矩阵** $h'_{qx}$ **和** $h''_{ry}$，**使得** $x'^q h'_{qx}=\delta_{3x}$ **和** $x''^r h''_{ry}=\delta_{3y}$.

(2) **对** $x,y=1,2$ **取** $\hat{l}'_{qx}=h'_{qx}$ **和** $\hat{l}''_{ry}=h''_{ry}$.

显而易见这个方法本质上适用于在表 16.1($p$296)和表 16.2($p$299)中所总结的全部类型的方程.

在三视图时本章推荐最基本的关系是点—直线—直线对应方程 $x^i l'_q l''_r \mathcal{T}_i^{qr}=0$，而在四视图时是直线对应方程 $l_p l'_q l''_r l'_s Q^{pqrs}=0$. 的确，通过细心地选择直线，把其他的对应化简到这种形式的对应可以提高数值计算的鲁棒性.

## 16.8 结束语

### 16.8.1 文献

虽然用了稍微不同的途径，本章在推导对应图像坐标之间的多线性关系时总结了[Triggs－95]和 Faugeras 与 Mourrain[Faugeras－95c]先前的结果. 点和直线混合对应关系的公式则是[Hartley－95a, Hartley－97c]的结果的扩展.

表 16.1($p$296)和表 16.2($p$299)关于完整的多线性关系的枚举；对多视图张量的公式以及从点对应分析独立方程的个数都取自[Hartley－95b]. 多视图张量的类似分析也在[Heyden－98]中出现过.

四焦点张量也许首先被[Triggs－95]发现. 四线性约束和它们相关的张量在几篇论文[Triggs－95, Faugeras－95c, Shashua－95b, Heyden－95a, Heyden－97c]中介绍过.

双(或 Grassmann－Cayley)代数是在文献[Carlsson－93]中引入到计算机视觉文献，进一步应用可参看[Faugeras－95c, Faugeras－97].

计算四焦点张量的算法和基于简化张量的重构算法在[Heyden－95a, Heyden－97c]中给出. 之后的论文[Hartley－98b]改进了该算法.

### 16.8.2 注释和练习

(1) 确定**仿射四焦点张量**(即从仿射摄像机矩阵计算的四焦点张量)的性质. 具体地

说，用该张量的行列式定义(16.21 - $p$297)来验证在表 16.4 中给出的非零元素的数目.

(2) 证明从跨三视图的单组点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'\leftrightarrow\mathbf{x}''$ 推导的 9 个线性方程(16.11 - $p$294)包含 4 个线性独立方程. 进一步令该方程组写为 $\mathbf{A}\mathbf{t}=\mathbf{0}$，其中 A 是一个 $9\times27$ 的矩阵. 那么 A 的 4 个非零奇异值相等. 与四焦点张量不同，由不同的点匹配推导的方程是线性独立的，因此 $n$ 个点匹配产生 $4n$ 个独立方程.

(3) 如果选择图像坐标的标准仿射基使得三组对应点在每一幅视图中的坐标分别为 $(0,0)^{\mathsf{T}}$，$(1,0)^{\mathsf{T}}$，$(0,1)^{\mathsf{T}}$，则所得的张量具有一个简单的形式. 这些"简化的张量"比一般形式的张量有更多的零元素. 比如，在这种情形，简化的基本矩阵的对角元素是零，简化的张量只有 15 个非零元素. 同时张量可以用更少的参数表示，例如，在简化的基本矩阵时为四个，因为基点有效地确定了其他参数. 进一步的细节在[Heyden - 95a，Heyden - 95b]中给出.

(4) 证明如果 4 个摄像机的中心共面，则四焦点张量有 28 个几何自由度.

# 第 17 章 $N$ 视图计算方法

本章介绍从图像的集合估计射影或仿射重构的计算方法——特别是对于视图数目多的情形.

我们从最一般情形即捆集调整开始. 然后介绍仿射摄像机和景物中含有平面的特殊情形. 最后我们介绍从一个图像序列获得点对应以及由这些对应导出射影重构的方法.

## 17.1 射影重构——捆集调整

设 3D 中的一组点 $\mathbf{X}_j$ 被矩阵为 $\mathrm{P}^i$ 的一组摄像机所拍摄. 我们用 $\mathbf{x}_j^i$ 标记第 $j$ 个点在第 $i$ 个摄像机像平面上的坐标. 我们希望求解下面的重构问题:已知图像坐标 $\mathbf{x}_j^i$ 的集合,求摄像机矩阵 $\mathrm{P}^i$ 和点 $\mathbf{X}_j$ 使得 $\mathrm{P}^i\mathbf{X}_j=\mathbf{x}_j^i$. 如果对 $\mathrm{P}^i$ 或者 $\mathbf{X}_j$ 不加进一步约束,这样的重构是一个射影重构,因为点 $\mathbf{X}_j$ 与真实的重构可能相差一个任意的 3D 射影变换.

**捆集调整** 如果图像测量有噪声则方程 $\mathbf{x}_j^i=\mathrm{P}^i\mathbf{X}_j$ 将不会完全地满足. 在这种情形下,我们将寻求最大似然(ML)解并假设测量噪声满足 Gauss 分布;我们希望估计射影矩阵 $\hat{\mathrm{P}}^i$ 和真正地投射到图像点 $\hat{\mathbf{x}}_j^i$ 的 3D 点 $\hat{\mathbf{X}}_j$,即 $\hat{\mathbf{x}}_j^i=\hat{\mathrm{P}}^i\hat{\mathbf{X}}_j$,并且在这些 3D 点出现的每幅视图中最小化重投影点和被检测(测量)的图像点 $\mathbf{x}_j^i$ 之间的图像距离,即

$$\min_{\hat{\mathrm{P}}^i,\mathrm{w}\hat{\mathrm{x}}_j}\sum_{ij}d(\hat{\mathrm{P}}^i\hat{\mathbf{X}}_j,\mathbf{x}_j^i)^2 \tag{17.1}$$

其中 $d(\mathbf{x},\mathbf{y})$是齐次点 $\mathbf{x}$ 和 $\mathbf{y}$ 之间的几何图像距离. 这种涉及到最小化重投影误差的估计称为**捆集调整**——它涉及到调整每个摄像机中心和这些 3D 点(或等价地每个 3D 点和摄像机中心集合)之间的射线丛.

捆集调整一般应该在任何重构算法的最后一步使用. 这个方法的好处是能够容忍数据的丢失并提供真正的 ML 估计. 同时它允许对每一个测量值指派单个协方差并且可以扩展到先验估计和摄像机参数或点的位置的约束. 简而言之,它称得上一个理想的算法,但有以下不足:(1)它需要提供一个好的初始值 (2)由于涉及到大量参数,它可能成为一个非常大的最小化问题. 我们将简要地讨论这两个问题.

**迭代最小化** 由于每个摄像机有 11 个自由度而每个 3 维空间点有 3 个自由度,一个涉及跨 $m$ 幅视图的 $n$ 个点的重构需要最小化 $3n+11m$ 个以上的参数. 事实上,因为元素常常是超参数化(例如对于齐次 P 矩阵用 12 个参数),此数是一个下界. 如果用 Levenberg-Marquardt 算法来最小化(17.1),则必须要分解(或者有时须求逆)维数为$(3n+11m)\times(3n+11m)$的矩阵. 当 $m$ 和 $n$ 增加时,它的开销变得非常大,甚至最终不可行. 对这个问题有下面几种解决办法:

**(1)减少 $n$ 和/或 $m$** 不要包括所有的视图或所有的点,后来用摄像机参数估计或三角

形法把这些点分别补上;或者,将数据分成几个集合,分别对每个集合进行捆集调整之后再融合. 这些策略将在节 17.5 中作进一步讨论.

(2)**交替方法** 交替进行通过改变摄像机最小化重投影误差和通过改变点最小化重投影误差. 因为每个点是在给定的固定摄像机下独立地被估计,并且类似地每个摄像机也由固定点独立地被估计,必须求逆的最大矩阵是用来估计一个摄像机的$11\times11$矩阵,可采用捆集调整一样的代价函数交替进行最小化,如果它收敛到一个全局最小,则应该可以得到相同的解,但也有可能收敛会更慢一些. 在[Triggs-2000]中把交替方法与捆集调整进行了比较.

(3)**稀疏方法** 这些方法在附录 4($p$416)中描述.

**初始解** 接下来的几节将介绍若干种初始化方法. 若在仿射摄像机下重构,则只要点被影像到每一幅视图,矩阵分解方法(节 17.2)可给出封闭形式的最优解. 即使在射影摄像机下,只要点被影像到每一幅视图,则也有一个适用的(迭代)矩阵分解方法(节 17.3). 如果数据带有更多的适用信息,比如这些点部分共面,则封闭形式的解仍然是可能的(节 17.4). 最后,针对点不是在每幅视图都可看到的情形,节 17.5 介绍了可供使用的分层方法.

## 17.2 仿射重构——矩阵分解算法

本节介绍由图像点对应集来重构,其图像为仿射摄像机所拍摄. 正如节 16.5.1 所介绍的,这是仿射重构的情形.

下面将介绍 Tomasi 和 Kanade[Tomasi-92]的矩阵分解算法,它总结在算法 17.1.1 中并具有下面的性质:

- **假设噪声服从各向同性零均值的 Gauss 分布并且每一个测量点的噪声是独立的且服从相同的分布,则矩阵分解算法可达到一个最大似然仿射重构.**

这个事实首先由 Reid 和 Murray[Reid-96]指出. 然而,这个方法要求每一点在所有视图中都有测量值. 这实际是一个限制,因为匹配点可能不在某些视图中出现.

一个仿射摄像机可以用它的最后一行等于(0,0,0,1)来表征. 然而,在本节中我们将用稍许不同的方法来标记它,即把摄像机映射的平移和纯线性变换分隔开. 因此我们有

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathrm{M}\begin{bmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \end{bmatrix} + \mathbf{t}$$

其中 M 是一个 $2\times3$ 矩阵而 $\mathbf{t}$ 是一个 2 维矢量. 从现在起为了阅读方便,$\mathbf{x}$ 表示一个非齐次图像点 $\mathbf{x}=(x,y)^{\mathsf{T}}$,而 $\mathbf{X}$ 表示非齐次世界点 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z})^{\mathsf{T}}$.

我们的目的是求使图像坐标的测量值的几何误差最小的重构. 这就是,我们希望估计摄像机$[\mathrm{M}^i,\mathbf{t}^i]$和 3D 点 $\mathbf{X}_j$,使得估计得到的图像点 $\hat{\mathbf{x}}_j^i=\mathrm{M}^i\mathbf{X}_j+\mathbf{t}^i$ 和测量的图像点之间的距离最小:

$$\min_{\mathrm{M}^i,\mathbf{t}^i,\mathbf{X}_j}\sum_{ij}(\mathbf{x}_j^i-\hat{\mathbf{x}}_j^i)^2=\min_{\mathrm{M}^i,\mathbf{t}^i,\mathbf{X}_j}\sum_{ij}(\mathbf{x}_j^i-(\mathrm{M}^i\mathbf{X}_j+\mathbf{t}^i))^2. \tag{17.2}$$

在这类最小化问题中,通常通过选择这些点的形心为坐标系的原点,平移矢量 $\mathbf{t}^i$ 可以

事前被消去. 这是因为,几何上仿射摄像机把 3D 点的形心映射到它们投影的形心. 因此,如果这些 3D 点的形心和其在每幅视图上图像点集的形心都被放置在坐标原点,那么我们可以假设 $\mathbf{t}^i=\mathbf{0}$. 这一步同样也要求 $n$ 个点被影像在所有的视图上,即没有哪一幅视图上某一点的图像坐标是未知的. 这个结论的一种解析推导如下. 对 $\mathbf{t}^i$ 求最小化要求

$$\frac{\partial}{\partial \mathbf{t}^i}\sum_{kj}(\mathbf{x}_j^k-(\mathrm{M}^k\mathbf{X}_j+\mathbf{t}^k))^2=\mathbf{0}.$$

简单的计算后上式简化为 $\mathbf{t}^i=\langle\mathbf{x}^i\rangle-\mathrm{M}^i\langle\mathbf{X}\rangle$,其中 $\langle\mathbf{x}^i\rangle=\frac{1}{n}\sum_j\mathbf{x}_j^i$ 和 $\langle\mathbf{X}\rangle=\frac{1}{n}\sum_j\mathbf{X}_j$ 是形心. 3D 坐标系的原点是任意定的,因此可以选它与形心 $\langle\mathbf{X}\rangle$ 相重合,此时 $\langle\mathbf{X}\rangle=\mathbf{0}$ 并且得到

$$\mathbf{t}^i=\langle\mathbf{x}^i\rangle. \tag{17.3}$$

于是,如果用被投影点的形心作为坐标系的原点来测量图像坐标,那么 $\mathbf{t}^i=\mathbf{0}$. 这样一来,我们用 $\Delta\mathbf{x}_j^i=\mathbf{x}_j^i-\langle\mathbf{x}^i\rangle$ 代替每一个 $\mathbf{x}_j^i$. 今后我们将假设这项工作已经完成,并且在这种中心化了的坐标下计算. 在这些新的坐标系下 $\mathbf{t}^i=\mathbf{0}$,从而(17.2)简化成

$$\min_{\mathrm{M}^i,\mathbf{X}_j}\sum_{ij}(\mathbf{x}_j^i-\hat{\mathbf{x}}_j^i)^2=\min_{\mathrm{M}^i,\mathbf{X}_j}\sum_{ij}(\mathbf{x}_j^i-\mathrm{M}^i\mathbf{X}_j)^2 \tag{17.4}$$

当写成一个矩阵时,这个最小化问题有一个非常简单的形式. 测量矩阵 W 由测量图像点的中心化坐标组成,是一个 $2m\times n$ 矩阵.

$$\mathrm{W}=\begin{bmatrix}\mathbf{x}_1^1 & \mathbf{x}_2^1 & \cdots & \mathbf{x}_n^1\\ \mathbf{x}_1^2 & \mathbf{x}_2^2 & \cdots & \mathbf{x}_n^2\\ \vdots & \vdots & \cdots & \vdots\\ \mathbf{x}_1^m & \mathbf{x}_2^m & \cdots & \mathbf{x}_n^m\end{bmatrix} \tag{17.5}$$

因为每个 $\mathbf{x}_j^i=\mathrm{M}^i\mathbf{X}_j$,完整的方程组可以写为

$$\mathrm{W}=[\mathrm{M}^{1\mathrm{T}},\mathrm{M}^{2\mathrm{T}},\cdots,\mathrm{M}^{m\mathrm{T}}]^{\mathrm{T}}[\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_n]$$

当噪声存在时,这个方程不被完全地满足,于是我们转而求在 Frobenius 范数下尽可能接近 W 的一个矩阵 $\hat{\mathrm{W}}$,使得矩阵 $\hat{\mathrm{W}}$ 可以分解为

$$\hat{\mathrm{W}}=\begin{bmatrix}\hat{\mathbf{x}}_1^1 & \hat{\mathbf{x}}_2^1 & \cdots & \hat{\mathbf{x}}_n^1\\ \hat{\mathbf{x}}_1^2 & \hat{\mathbf{x}}_2^2 & \cdots & \hat{\mathbf{x}}_n^2\\ \vdots & \vdots & \cdots & \vdots\\ \hat{\mathbf{x}}_1^m & \hat{\mathbf{x}}_2^m & \cdots & \hat{\mathbf{x}}_n^m\end{bmatrix}=\begin{bmatrix}\mathrm{M}^1\\ \mathrm{M}^2\\ \vdots\\ \mathrm{M}^m\end{bmatrix}[\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_n] \tag{17.6}$$

在此情况下,可以验证:

$$\|\mathrm{W}-\hat{\mathrm{W}}\|_{\mathrm{F}}^2=\sum_{ij}(\mathrm{W}_{ij}-\hat{\mathrm{W}}_{ij})^2=\sum_{ij}(\mathbf{x}_j^i-\hat{\mathbf{x}}_j^i)^2=\sum_{ij}(\mathbf{x}_j^i-\mathrm{M}^i\mathbf{X}_j)^2$$

把它与(17.4)相比较,我们发现此几何误差的最小化等价于在 Frobenius 范数下寻找尽可能接近 W 的一个矩阵 $\hat{\mathrm{W}}$.

注意满足(17.6)的一个矩阵 $\hat{\mathrm{W}}$ 是一个 $2m\times 3$ 的**运动矩阵** $\hat{\mathrm{M}}$ 和一个 $3\times n$ 的**结构矩阵** $\hat{\mathrm{X}}$ 的乘积:因此 $\hat{\mathrm{W}}=\hat{\mathrm{M}}\hat{\mathrm{X}}$ 的秩为 3. 换句话说,我们在 Frobenius 范数下求最接近 W 的一个秩为 3 的矩阵. 这样的一个矩阵可以通过 W 的 SVD 分解并截短到秩为 3 得到. 更详细地讲,若 SVD 分解是 $\mathrm{W}=\mathrm{UDV}^{\mathrm{T}}$,则 $\hat{\mathrm{W}}=\mathrm{U}_{2m\times 3}\mathrm{D}_{3\times 3}\mathrm{V}_{n\times 3}^{\mathrm{T}}$ 是在 Frobenius 范数下最接近 W 的秩为 3

的矩阵，其中 $U_{2m\times 3}$ 由 U 的前 3 列组成，$V^{T}_{n\times 3}$ 由 V 的前三列组成，而是 $D_{3\times 3}$ 是包含前三个奇异值的对角矩阵，$D_{3\times 3}=\mathrm{diag}(\sigma_1,\sigma_2,\sigma_3)$.

注意 $\hat{M}$ 和 $\hat{X}$ 的选择不是唯一的. 比如可以选取 $\hat{M}=U_{2m\times 3}D_{3\times 3}$ 和 $\hat{X}=\hat{V}^{T}_{n\times 3}$，或者 $\hat{M}=U_{2m\times 3}$ 和 $\hat{X}=D^{T}_{3\times 3}V_{n\times 3}$，因为在每一种情况下都有 $\hat{W}=\hat{M}\hat{X}=U_{2m\times 3}D_{3\times 3}V^{T}_{n\times 3}$

**仿射多义性** 事实上对任何这样的选择都存在附加的多义性，因为任意一个秩为 3 的 $3\times 3$ 矩阵 A 可以这样插入到分解中去，因为 $\hat{W}=\hat{M}AA^{-1}\hat{X}=(\hat{M}A)(A^{-1}\hat{X})$. 这意味着从 $\hat{M}$ 得到的摄像机矩阵 $M^i$ 和从 $\hat{X}$ 得到的 3D 点 $\mathbf{X}_j$ 在相差一个共同的矩阵 A 意义下确定. 换句话说，MLE 是仿射重构.

通过提供景物的度量信息(如节 9.4.2($p$185)所描述)，或利用第 18 章描述的自标定方法，这个仿射重构可以升级到度量重构. 注意在仿射摄像机情形，只需要确定三个内参数(与射影摄像机的五个相比较)因而它的自标定任务相应地比较简单.

目标

已知跨 $m$ 幅视图的 $n\geqslant 4$ 个图像点对应 $\mathbf{x}^i_j$，$j=1,\ldots,n;i=1,\ldots,m$，确定仿射摄像机矩阵 $\{M^i,\mathbf{t}^i\}$ 和 3D 点 $\{\mathbf{X}_j\}$ 使得重投影误差

$$\sum_{ij}(\mathbf{x}^i_j-(M^i\mathbf{X}_j+\mathbf{t}^i))^2$$

在 $\{M^i,\mathbf{t}^i,\mathbf{X}_j\}$ 上最小化，其中 $M^i$ 是一个 $2\times 3$ 的矩阵，$\mathbf{X}_j$ 是一个 3 维矢量，$\mathbf{x}^i_j=(x^i_j,y^i_j)^T$ 和 $\mathbf{t}^i$ 是 2 维矢量.

算法

(1)**计算平移** 每一平移 $\mathbf{t}^i$ 当作第 $i$ 幅图像上点的形心来计算，即

$$\mathbf{t}^i=\langle\mathbf{x}^i\rangle=\frac{1}{n}\sum_j\mathbf{x}^i_j.$$

(2)**中心化数据** 把每一图像上的点用它们相对形心的坐标表示使这些点中心化：

$$\mathbf{x}^i_j\leftarrow\mathbf{x}^i_j-\langle\mathbf{x}^i\rangle.$$

此后，对这些中心化的坐标进行操作.

(3)按(17.5)中的定义由中心化的数据构造 $2m\times n$ **测量矩阵** W，并计算它的 SVD 分解 $W=UDV^T$.

(4)然后矩阵 $M^i$ 由 U 的前三列乘以奇异值得到：

$$\begin{bmatrix}M^1\\M^2\\\vdots\\M^m\end{bmatrix}=[\sigma_1\mathbf{u}_1\quad\sigma_1\mathbf{u}_2\quad\sigma_3\mathbf{u}_3].$$

矢量 $\mathbf{t}^i$ 计算如第(1)步而 3D 结构从 V 的前三列读出

$$[\mathbf{X}_1\quad\mathbf{X}_2\quad\cdots\quad\mathbf{X}_n]=[\mathbf{v}_1\quad\mathbf{v}_2\quad\mathbf{v}_3]^T.$$

算法 17.1 由跨 $m$ 幅视图的 $n$ 个图像点对应确定一个仿射重构的 MLE 的矩阵分解算法(在 Gauss 图像噪声的假定下).

### 17.2.1 仿射多视图张量

矩阵分解算法提供了由图像点对应计算仿射多视图张量的最一种优方法. 这些张量有

仿射基本矩阵、仿射三焦点张量和仿射四焦点张量. 对每一种情形,这个算法在相差一个全局仿射多义性的意义下确定摄像机矩阵,然后这些张量可直接由摄像机矩阵算出(如用第16章的方法). 计算这些张量与3D空间的仿射多义性无关,因为它们不受3维空间的仿射变换影响. 事实上没有必要计算整个W的SVD分解,因为只需分解U的分量. 如果与视图的数目相比,点数 $n$ 很大时,则用不计算V的SVD计算(见表A3.1 ($p$409))可以带来很大的节省.

另一个代替SVD的方法是应用 $WW^T$ 的特征值分解,因为 $WW^T=(UDV^T)(UDV^T)^T=UD^2U^T$. 在三幅视图的情形(三焦点张量的计算)矩阵 $WW^T$ 的维数只是 $9\times9$. 因此这个方法有显著的节省. 然而,从数值计算角度考虑它是下策,因为 $WW^T$ 矩阵的条件数变为原来的平方(见在[Golub-89]中SVD的讨论). 由于我们只需要最大的三个特征矢量,因此,也许它不会出现问题. 然而,这个方法的节省与避免计算V的SVD分解比较并不那么显著.

矩阵分解方法可以用来计算多视图的任何仿射张量. 然而在计算仿射基本矩阵时,第13章介绍的算法13.1($p$249)更直接. 这两种方法的结果完全相同.

## 17.3 射影分解

仿射分解的方法不能直接应用于射影重构. 然而文献[Sturm-96]观察到:如果知道每一个点的"射影深度"则结构和摄像机参数可以用一个简单的矩阵分解算法估计,该算法在风格上类似于仿射矩阵分解算法.

考虑一组图像点 $\mathbf{x}_j^i=P^i\mathbf{X}_j$. 这个表示射影映射的方程必须解释为在相差一个常数因子下才成立. 明显地写出这些常数因子,我们有 $\lambda_j^i\mathbf{x}_j^i=P^i\mathbf{X}_j$. 在这个方程中,以及今后在射影分解算法的描述中,符号 $\mathbf{x}_j^i$ 总表示图像点的三维矢量 $(x_j^i,\ y_j^i,1)^T$. 这样一来,第三个坐标等于1并且 $x_j^i$ 和 $y_j^i$ 是实际测量得到的图像坐标. 如果每个点在每幅视图中都是能观察到的,因此对所有 $i,j,\mathbf{x}_j^i$ 是已知的,则完整的方程组可以写成如下的单个矩阵方程:

$$\begin{bmatrix}\lambda_1^1\mathbf{x}_1^1 & \lambda_2^1\mathbf{x}_2^1 & \cdots & \lambda_n^1\mathbf{x}_n^2\\ \lambda_1^2\mathbf{x}_1^2 & \lambda_2^2\mathbf{x}_2^2 & \cdots & \lambda_n^2\mathbf{x}_n^2\\ \vdots & \vdots & \cdots & \vdots\\ \lambda_1^m\mathbf{x}_1^m & \lambda_2^m\mathbf{x}_2^m & \cdots & \lambda_n^m\mathbf{x}_n^m\end{bmatrix}=\begin{bmatrix}P^1\\ P^2\\ \vdots\\ P^m\end{bmatrix}[\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_n]. \tag{17.7}$$

只有对每一个测量点 $\mathbf{x}_j^i$ 加上正确的权因子 $\lambda_j^i$ 后这个方程才成立. 我们暂时假定这些深度是已知的. 如分析仿射矩阵分解算法时一样,我们希望左边的矩阵——标记它为W——的秩为4,因为它分别是4列和4行的两个矩阵的乘积. 实际的测量矩阵可以利用SVD分解矫正到秩为4. 这样一来,如果 $W=UDV^T$,除了D的前4个对角元素外所有的元素取零,得到矩阵 $\hat{D}$. 矫正的测量矩阵便是 $\hat{W}=U\hat{D}V^T$. 摄像机矩阵由 $[P_1^T,P_2^T,\cdots,P_m^T]^T=U\hat{D}$ 恢复而点则由 $[\mathbf{X}_1,\mathbf{X}_2\cdots,\mathbf{X}_n]=V^T$ 恢复. 注意这个分解不是唯一的(说明了重构有射影多义性),事实上在(17.7)右边的两个矩阵之间我们可以插入一个任意的 $4\times4$ 射影变换H和它的逆.

射影分解方法的步骤总结在算法17.2中.

**目标**

给定在 $m$ 幅视图中观察到的 $n$ 个图像点对应集

$$\mathbf{x}_j^i; i=1,\ldots,m; j=1,\ldots,n$$

计算一个射影重构.

**算法**

(1) 与节 3.4.4($p67$)一样用均匀缩放来归一化图像数据

(2) 从射影深度 $\lambda_j^i$ 的一个初始估计开始. 它可以由诸如一个初始的射影重构技术得到,或置全部 $\lambda_j^i=1$.

(3) 用常数因子乘行和列来归一化深度 $\lambda_j^i$. 一种办法是先令所有行的范数为 1,然后令所有列的范数为 1.

(4) 形成(17.7)左边的 $3m\times n$ 测量矩阵,用 SVD 找最接近于它的秩为 4 的逼近并通过分解求摄像机矩阵和 3D 点.

(5) **迭代选择项** 把这些点重投影到每一图像中以便得到深度的新的估计,然后从第(2)步起作循环.

算法 17.2 通过矩阵分解求射影重构.

### 17.3.1 选取深度

加权因子 $\lambda_j^i$ 称为点的**射影深度**. 之所以用这个术语是因为在一个欧氏坐标系下当摄像机矩阵已知时这些 $\lambda_j^i$ 与实际深度有关系,参考节 5.2.3($p108$),特别是图 5.6($p108$). 这个射影矩阵分解算法的主要困难在于必须事先知道这些射影深度,而我们又对其一无所知. 存在各种估计深度的技术.

(1)由利用其他方法(如在下面节 17.5 中讨论的)获得的初始射影重构开始,然后通过重投影 3D 点计算 $\lambda_j^i$.

(2)由最初的深度都等于 1 开始,计算这个重构再重投影得到的新的深度估计. 为了得到改进的估计,这个步骤可以反复进行. 然而,不保证该过程一定收敛到一个全局最小.

[Sturm-96]首次给出了一个计算深度的方法,它把通过基本矩阵或三焦点张量得到的深度估计一对对地串起来. 这个方法非常类似于把三幅图像组串起来获得初始射影重构(见节 17.5),同时保证对同一射影重构有一致的尺度因子.

### 17.3.2 最小化什么?

在有噪声或者 $\lambda_j^i$ 不正确时,方程(17.7)不被完全满足. 我们确定一个校正的测量矩阵 $\hat{\mathbf{W}}$在 Frobenius 范数意义下最接近于 $\mathbf{W}$,并且满足秩为 4 的条件. 记这个矩阵的元素为 $\hat{\lambda}_j^i\hat{\mathbf{x}}_j^i$,则解应最小化表达式

$$\|\mathbf{W}-\hat{\mathbf{W}}\|^2=\sum_{ij}\|\lambda_j^i\mathbf{x}_j^i-\hat{\lambda}_j^i\hat{\mathbf{x}}_j^i\|^2=\sum_{ij}(\lambda_j^i x_j^i-\hat{\lambda}_j^i\hat{x}_j^i)^2+(\lambda_j^i y_j^i-\hat{\lambda}_j^i\hat{y}_j^i)^2+(\lambda_j^i-\hat{\lambda}_j^i)^2. \tag{17.8}$$

由于最后一项,在取最小值时 $\hat{\lambda}_j^i$ 必须接近 $\lambda_j^i$. 假定二者相等,(17.8)便简化为

$\sum_{ij}(\lambda_j^i)^2 \| \mathbf{x}_j^i - \hat{\mathbf{x}}_j^i \|^2$. 注意 $\| \mathbf{x}_j^i - \hat{\mathbf{x}}_j^i \|$ 是测量点和估计点之间的几何距离，而被最小化的是加权几何距离的平方和，其中每一点被 $\lambda_j^i$ 加权. 如果所有的几何深度 $\lambda_j^i$ 接近相等，则矩阵分解方法最小化的是具有共同缩放值 $\lambda_j^i$ 的几何距离的一个逼近.

### 17.3.3 深度的归一化

这里定义的射影深度不是唯一的. 事实上假设 $\lambda_j^i \mathbf{x}_j^i = \mathrm{P}^i \mathbf{X}_j$. 如果用 $\alpha^i \mathrm{P}^i$ 代替 $\mathrm{P}^i$ 并用 $\beta_j \mathbf{X}_j$ 代替 $\mathbf{X}_j$，则有

$$(\alpha^i \beta_j \lambda_j^i)\mathbf{x}_j^i = (\alpha^i \mathrm{P}^i)(\beta_j \mathbf{X}_j)$$

换句话说，射影深度 $\lambda_j^i$ 可以用一个因子 $\alpha^i$ 乘(17.7)的第 $i$ 行和用一个因子 $\beta_j$ 乘第 $j$ 列来代替. 根据上一段推导，所有的 $\lambda_j^i$ 越接近于 1，表示几何距离的误差表达式越准确. 因此，通过用常数因子 $\alpha^i$ 和 $\beta^j$ 乘以测量矩阵的行和列来归一化 $\lambda_j^i$ 的值，使得它们尽可能接近 1 是有好处的. 一个简单的启发式的做法是用因子 $\alpha^i$ 乘以每一行使得它的范数为 1，接下来类似地对列归一化. 行和列的归一化可以迭代进行.

### 17.3.4 图像坐标的归一化

如本书介绍的大多数涉及图像坐标齐次表示的数值算法一样，归一化图像坐标是非常重要的. 一个合理的方案是节 3.4.4($p67$)介绍的各向同性归一化方法. 我们能从下面的情形中很清楚地看到归一化的必要性. 考虑两个图像点 $\mathbf{x} = (200, 300, 1)^\mathrm{T}$ 和 $\hat{\mathbf{x}} = (250, 375, 1)^\mathrm{T}$. 显然从几何意义上讲，这两点相距很远. 然而，误差表示式(17.8)度量的不是几何误差，而是齐次矢量之间的距离 $\|\lambda\mathbf{x} - \hat{\lambda}\hat{\mathbf{x}}\|$. 选取 $\lambda = 1.25$ 和 $\hat{\lambda} = 1.0$，该误差是 $\|(200, 300, 1.25)^\mathrm{T} - (200, 300, 1)^\mathrm{T}\|$，它在比例上非常小. 换一个角度来看，点 $\mathbf{x} = (200, 300, 1)^\mathrm{T}$ 和 $\hat{\mathbf{x}} = (199, 301, 1)^\mathrm{T}$ 在几何上非常接近，通过选择 $\lambda$ 和 $\hat{\lambda}$(除了很小的值)可以使它们的距离不是很小. 读者可以观察到如果这些点用一个因子 200 缩小，则这些不正常的情形不再发生. 简而言之，采用了归一化坐标，该误差更逼近几何误差.

### 17.3.5 何时假设 $\lambda_j^i = 1$ 是合理的?

根据结论 5.1($p108$)，如果把摄像机矩阵归一化使得 $p_{31}^2 + p_{32}^2 + p_{33}^3 = 1$，且把 3D 点归一化使得最后的坐标 $\mathrm{T} = 1$，则由 $\lambda_j^i (x_j^i, y_j^i, 1)^\mathrm{T} = \mathrm{P}^i \mathbf{X}_j$ 定义的 $\lambda_j^i$ 是在欧氏坐标系下摄像机的真正的深度. 如果所有点在整个图像序列中与摄像机是等距离的，则我们可以合理地假设每一个 $\lambda_j^i = 1$，因为(17.7)至少有一个解 $\mathrm{P}^i$ 和 $\mathbf{X}_j$ 是以上述方式归一化了的真正的摄像机和点. 更一般地，假设点位于不同的深度，但是在整个过程中，每点仍与摄像机保持近似于相同的深度 $d_j$. 在此情形下将存在一个解具有 $\lambda_j^i = 1$ 和 $\mathbf{X}_j = d_j^{-1}(\mathrm{X}_j, \mathrm{Y}_j, \mathrm{Z}_j, 1)^\mathrm{T}$. 类似地，允许摄像机矩阵乘以一个因子，我们有

- **如果不同的 3D 点 $\mathbf{X}_j$ 的真正的深度比率在一个图像序列中几乎保持常数，则假设 $\lambda_j^i = 1$ 可作为射影深度的合理的首次近似.**

一个航空图像摄像机在恒常数高度上直接向下拍摄的情形就是一个例子.

## 17.4 由平面诱导的单应重构

由节 16.5.2 可知：如果已知四点共面并在每幅视图上都可能观察到，则与图像点相关的多焦点张量的计算将非常简单．最主要的好处在于满足所有约束的张量可以用一种线性算法来计算．在计算三焦点张量或四焦点张量时尤为有利，因为它们必须满足由该几何结构施加的很多约束(分别为 8 和 51)．这里将证明这个方法可以扩展为同时估计所有的摄像机矩阵．

从共面点推导的四组图像对应的条件等价于已知由空间一张平面诱导的图像之间的单应，因为一个单应可以由四个点计算得到．在下面的方法中只有这些单应是最重要的．这些单应可以从四组或更多组点对应或直线对应计算得到，或者利用直接对射方法由图像来直接估计．

现在来推广节 16.5.2 的讨论，如果 $H^i$ 是平面诱导的单应，这个单应把第一幅图像上的点映射到第 $i$ 幅图像上它的匹配点，则摄像机矩阵的集合可以假设有形式 $P^i=[H^i|\mathbf{t}^i]$，其中 $H^i$ 已知且尺度固定，但最后一列 $\mathbf{t}^i$ 则不是．我们可以假设 $P^1=[I|\mathbf{0}]$，使得 $\mathbf{t}^1=\mathbf{0}$．剩下所有 $\mathbf{t}^i$ 的集合有 $3m-4$ 个自由度，因为 $\mathbf{t}^i$ 只确定到相差一个共同的尺度因子．现在假设跨两幅或更多视图的若干点或直线对应是已知的(直线对应时需要三幅视图)．这些对应必须由不在(用来计算 $H^i$ 的)参考平面上的 3D 点或直线推导出来．跨两幅视图的每组点对应给出关于基本矩阵元素的一个线性方程．类似地，跨三幅或四幅视图的对应点或直线给出关于三焦点或四焦点张量元素的线性方程．

关键之处(如节 16.5.2 中所解释的)是我们可以用矢量 $\mathbf{t}^i$ 的元素线性地表示基本矩阵(或三焦点或四焦点张量)的元素．这样一来由一个点或直线对应诱导的每一个线性关系可以反过来联系到由 $\mathbf{t}^i$ 的元素表示的一组线性关系．因此，比如，跨视图 $i,j$ 和 $k$ 的一个对应产生用三个矢量 $\mathbf{t}^i,\mathbf{t}^j$，和 $\mathbf{t}^k$ 元素表示的一组线性方程．跨许多视图的一组对应可以分解成跨连续视图集合的对应．因此，比如，跨 $m>4$ 视图的单个点对应给出如下形式的一组方程：

$$\begin{bmatrix} \mathsf{X} & \mathsf{X} & \mathsf{X} & \mathsf{X} & & & & & & & \\ & \mathsf{X} & \mathsf{X} & \mathsf{X} & \mathsf{X} & & & & & & \\ & & \mathsf{X} & \mathsf{X} & \mathsf{X} & \mathsf{X} & & & & & \\ & & & & & & \ddots & & & & \\ & & & & & & & \mathsf{X} & \mathsf{X} & \mathsf{X} & \mathsf{X} \\ \mathsf{X} & & & & & & & & \mathsf{X} & \mathsf{X} & \mathsf{X} \\ \mathsf{X} & \mathsf{X} & & & & & & & & \mathsf{X} & \mathsf{X} \\ \mathsf{X} & \mathsf{X} & \mathsf{X} & & & & & & & & \mathsf{X} \end{bmatrix}$$

其中每一行表示由四焦点张量关系推出的一组方程．每一个 X 表示 3 列组成的块并对应于 $\mathbf{t}^i$ 的一个矢量．在上面的表中，从最后一幅视图到第一幅视图，方程进行了卷绕以增加刚性．否则 $\mathbf{t}^i$ 的值可能产生第一幅视图到最后一幅视图的偏移．选择视图组合的其他方案也是可能的，没有必要限制为相继的视图．

线性关系也可以在足够多的图像的任何子集(子集元素为 2,3 或 4，依赖于用什么张量来产生方程)之间产生．我们必须在增加解的稳定性和由于增加更多的方程而导致增加计

算开销之间作权衡. 也可以用两焦、三焦点和四焦点的混和约束来产生所有方程组,并且也没有必要所有点在所有视图中都被观察到.

### 17.4.1 产生的方程数

令视图的总数为 $m$. 考虑 $s$ 幅视图的一个子集($s=2,3$ 或 4)并且设这些视图之间给定了 $n$ 组点对应. 我们简要地考虑从 $s$ 幅视图组成的子集单独地进行重构的问题. 从这些点对应我们能生成这 $s$ 幅视图的元素 $\mathbf{t}'$ 之间的一个方程组 $\mathbf{A}\mathbf{t}'$,然后再估计 $\mathbf{t}'$ 的 $3s$ 个元素的值. 在操作时,我们可以假设第一幅视图有 $\mathbf{t}=\mathbf{0}$. 且剩下 $s-1$ 幅视图的矢量 $\mathbf{t}$ 仅确定到相差一个共同尺度. 因此在方程组 $\mathbf{A}\mathbf{t}'$ 中的 $\mathbf{A}$ 的右零空间维数至少为 4(对应于解的 4 个自由度). 一般有:

**结论 17.1 忽略噪声数据的影响,由 $s$ 幅视图上 $n\geqslant 2$ 组点对应产生的方程组的总秩数是 $3s-4$. 此数与用来产生它们的点(或直线)对应的数目无关.**

准确地说,上面的推导说明秩最多是 $3s-4$. 对于跨 2 视图、3 视图和 4 视图的对应,它分别等于 2、5 或 8. 然而,只要有两组对应就能达到这个最大秩,因为正如节 16.5.2 计数的推导所证明的:当 $s=2,3$ 或 4 时,对视图重构而言两点是足够的.

现在考虑 $m$ 幅视图的全部集合. 所有 $\mathbf{t}^i$ 的可恢复参数的总数是 $3m-4$. 因此,对一个可能的解,方程的个数必须超过 $3m-4$,这给出了下面的结论:

**结论 17.2 如果从 $m$ 幅视图中选出每个子集由 $s_k$ 幅视图组成 $S$ 个子集,则为了求解摄像机矩阵最后一列的所有矢量 $\mathbf{t}^i$,必须满足**

$$\sum_{k=1}^{S}(3s_k-4)\geqslant 3m-4.$$

可以验证若用跨 2 视图的对应,涉及到从基本矩阵的约束推导的方程,则在一个形如

$$\begin{bmatrix} \mathsf{X} & \mathsf{X} & & & & \\ & \mathsf{X} & \mathsf{X} & & & \\ & & \mathsf{X} & \mathsf{X} & & \\ & & & \ddots & & \\ & & & & \mathsf{X} & \mathsf{X} \\ \mathsf{X} & & & & & \mathsf{X} \end{bmatrix}$$

的配置中只用一对相继视图是不充分的. 对这种情形,所生成的方程总数是 $m(3s-4)=m(3\cdot 2-4)=2m$,而所需要的方程总数 $3m-4$. 因此在 $m>4$ 时没有足够多的方程. 这是因为相继视图的基本矩阵不足以确定该视图序列的结构. 它必须加进由非相继视图得到的另外的约束. 比如:

$$\begin{bmatrix} \mathsf{X} & & \mathsf{X} & & & \\ & \mathsf{X} & & \mathsf{X} & & \\ & & \mathsf{X} & & \mathsf{X} & \\ & & & \ddots & & \\ \mathsf{X} & & & & \mathsf{X} & \\ & \mathsf{X} & & & & \mathsf{X} \end{bmatrix}$$

当然应该注意,节 14.4($p273$)的讨论中建议在跨三幅或四幅视图时应选用三焦点或四

焦点约束.

## 17.5 利用序列重构

在本章的最后一节中,我们把本书前面的若干思想融会在一起. 我们的目标是从视频提供的帧序列中计算重构. 解这个问题有三个步骤:(1)在整个序列中计算对应的特征;(2)计算一个可以用作第三步起点的初始重构;(3)捆集调整(如节 17.1 所描述).

这里我们考虑的特征是兴趣点,虽然其他的比如直线也同样能采用. 对应问题是令人烦恼的问题,因为一个兴趣点特征一般不出现在所有图像中,且常常从相继图像中消失. 然而,捆集调整不受消失的对应所影响.

一个视频序列较之一个任意图像集合的几个优点是:(1)图像有序;(2)在相继两帧的摄像机中心之间的距离(基线)小. 小的基线是重要的,因为它使相继图像之间可能的特征匹配更容易地获得和评定. 因为视图之间图像点没有移动"很远",因此可以使用最近的区域来搜索,所以匹配点更容易得到;因为邻近图像表观上是相似的,所以匹配更容易评定(如它们是否由三维空间中同一点产生),小基线的缺点是由它得到 3D 结构估计效果差. 然而,这个缺点可以通过越过序列中的许多视图使得有效基线变大而弥补.

该方法概括在算法 17.3 中,有若干种策略可以用来获得初始重构,虽然这个领域在某种程度上仍需改进. 三种可能的策略如下:

1. **延伸基线** 假设在整个序列中可观察到的景物点的数量是合理的. 利用图像对的匹配(从 F),或三元组匹配点(从 $\mathcal{T}$)可以从第一帧到最后一帧寻找对应. 事实上,如果相继帧之间的基线很小(相对于结构深度),则图像对的匹配可以利用单应的计算(算法 3.6($p$79))得到——它提供一个比 F(点到直线)更强的匹配约束(点到点).

然后一个三焦点张量能从序列中第一帧、中间一帧(比如)和最后一帧的对应点来估计. 这个张量确定了对应于这些点和帧的一个射影重构. 对应于中间帧的摄像机则可以通过计算摄像机矩阵的方法来估计,而整个序列中不是都被观察到的景物点可通过三角形法来估计.

2. **子序列的分层合并** 这里的思想是把图像序列划分为便于操作的子序列(可以有若干层次的划分). 然后对每个子序列计算射影重构并把这些重构"压缩"(合并)到一起.

考虑两个三元组视图(其中两幅视图重迭)合并的问题. 在这里视图上扩展对应是一个简单问题:一个跨三元组 1-2-3 同时跨三元组 2-3-4 的对应可以扩展到跨帧 1-2-3-4,因为图像对 2-3 在这两个三元组是重叠的. 摄像机矩阵和 3D 结构则由帧 1-2-3-4 计算得到,比如首先计算摄像机矩阵然后用捆集调整. 通过合并邻近的帧群,这个过程可一直扩展到整个序列的摄像机矩阵和对应点得到确定为止. 用这个方法,误差能够均匀地分布在整个序列中.

3. **增量捆集调整** 每当新的一帧的对应加入之后就进行一次新的捆集调整. 这个方法的缺点是计算开销大并且有可能产生系统的积累误差.

当然这三种方法可以结合起来. 比如,图像序列可以分解成子序列,其中公共点都被观察到,并且利用扩展基线方法来为子序列建立重构. 然后再分层合并这些子序列.

用这个方法,结构和摄像机可以自动地从包含数百帧的序列中计算出来. 这些重构可以奠定诸如导航(确定摄像机/车辆的位置)和虚拟模型生成等任务的基础. 通常有必要利用第 9 章和第 18 章介绍的方法首先从射影重构计算度量重构. 度量重构和虚拟模型的产生在下面的例子中说明.

目标

给定一个视频图像序列,计算对应并重构景物结构和每帧图像的摄像机矩阵

算法

(1) **兴趣点**:计算每幅图像中的兴趣点.

(2) 2 **视图对应**:计算兴趣点的对应并用算法 10.4($p$200)计算相继帧之间的 F(如果基线运动太小,一些帧可以略去).

(3) 3 **视图对应**:计算兴趣点的对应以及用算法 15.4($p$286)计算所有相继图像三元组之间的 $T$.

(4) **初始重构**:见正文.

(5) **捆集调整**整个序列的摄像机和 3D 结构.

(6) **自标定**:见第 18 章(选择项)

算法 17.3 由一个图像序列求重构的总结.

**例 17.3 走廊序列**

一个摄像机装在一个移动的车上以获得这个图像序列. 这车沿地板运动并向左转. 在这一图像序列中向前的平移使结构恢复非常困难,原因是对三角形法来说基线太小. 在这种情形,利用序列中所有帧的好处是显著的. 图 17.1 显示了恢复的结构. △

**例 17.4 "Wilshire"序列**

这是直升机拍摄的洛杉机 Wilshire 大道. 对于这种情形,重构遇到的阻力是场景中重复的结构——许多特征点(比如在摩天大楼窗户上的那些特征点)有非常类似的灰度邻域,因此基于相关关系的追踪产生许多假的候选点. 然而,鲁棒的几何指导的匹配成功地排除了不正确的对应. 图 17.2 显示了该结构. △

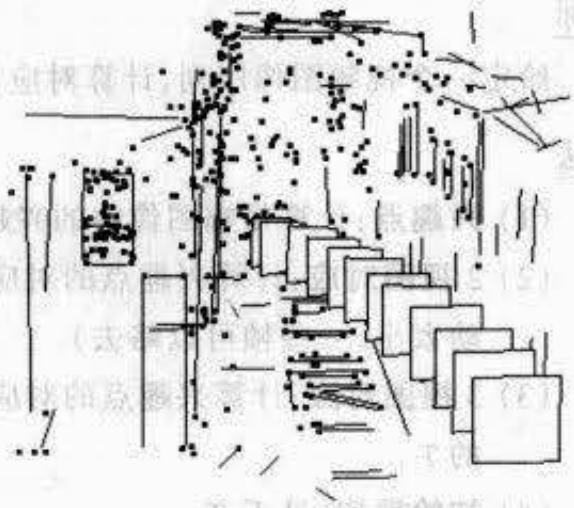
**例 17.5 恐龙序列**

模型恐龙在一个转盘上旋转使得摄像机有效地处于这个目标的周围. 特征选取在彩色信号的亮度成分上进行. 对于这种情形,运用了已知运动是圆周作为[Fitzgibbon - 98a]附加约束,其结果提高了结构保真性. 图 17.3 显示了恢复的点结构和摄像机. △

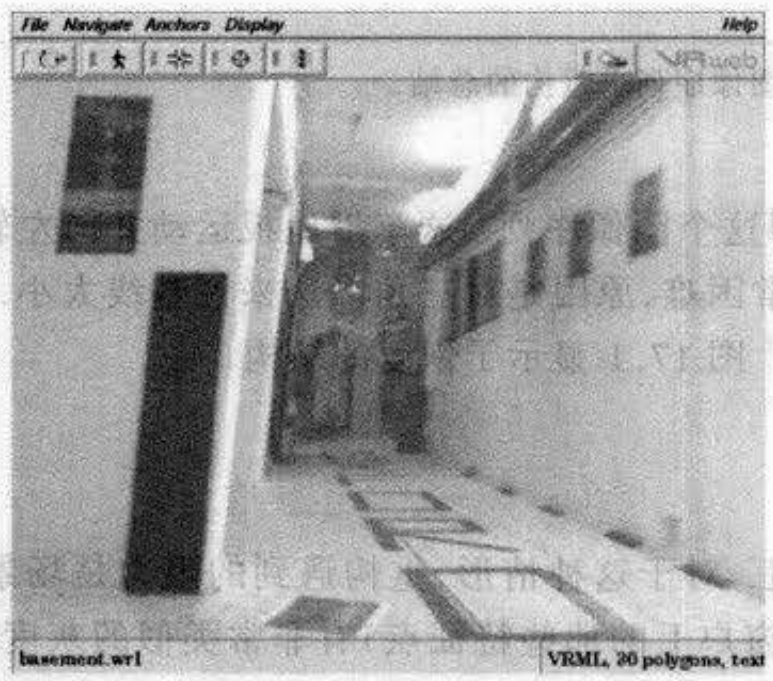
## 17.6 结束语

可以坦城地说,还不知道一个完全满意的技术可用于从一个射影图像序列中重构. 许多临时应对的方法被运用并取得一定的成功. 4 视图是基于多视图张量的封闭形式解的上限. 当视图数比较大时,对于这个问题还没有一个整齐的数学公式. 它的一个例外是基于

a



b




c

d

图 17.1 **走廊序列**.(a)该景物中点和直线的 3 维重构,(b)由图像自动计算得到的摄像机(由它们的图像平面表示). 然后用[Baillard－99]的方法自动构造由三角面片的图形模型映射的一种纹理. (c)由一个新视点绘制的景物,与序列中的任何一幅都不同. (d)该景物的 VRML 模型,其中摄像机由它们的图像平面表示(纹理由序列中原始图像的映射得到).

对偶(见第 19 章)的 $m$ 视图技术,但是这个技术只限于 6 到 8 个点并取决于采用的对偶张量(基本矩阵、三焦点张量或四焦点张量). 大多数图像序列包含比它多得多的匹配点.

### 17.6.1 文献

Tomasi－Kanade 算法首先针对正交投影提出[Tomasi－92],但是后来推广到平行透视[Poelman－94]. 它现已推广到直线和二次曲线,比如[Kahl－98a],但是 MLE 的性质不再适用,而且仍不清楚在仿射重构中什么被最小化.

矩阵分解法推广到射影摄像机归功于 Sturm 和 Triggs[Sturm－96]. 利用矩阵分解的

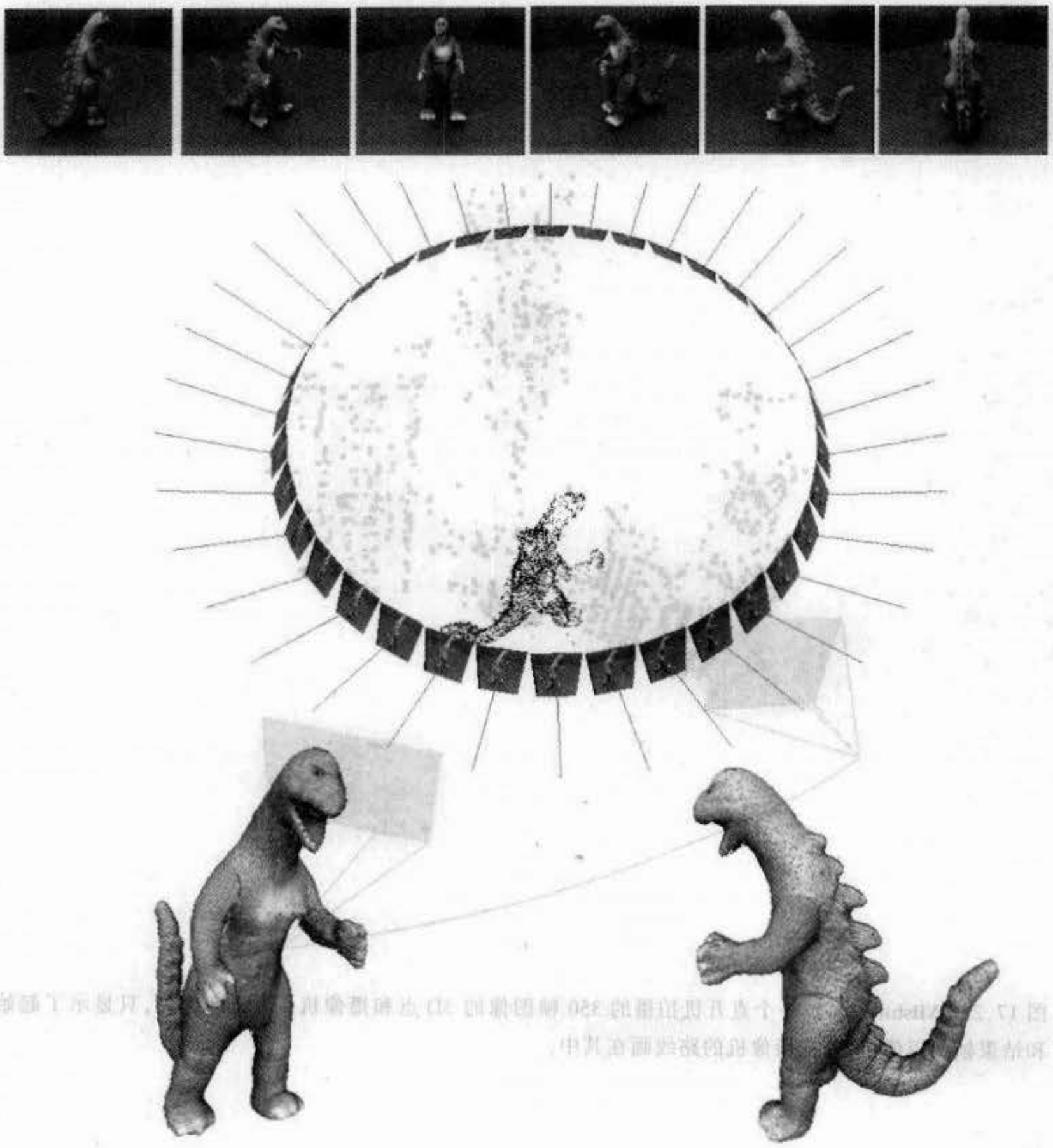


图 17.3 **恐龙**.该恐龙序列的 3D 点(大约 5000 点)和摄像机位置. 这个序列的 3D 图形模型的自动计算在[Fitzgibbon－98a]中描述. 图像序列取自 Hannover 大学.

迭代方法曾由[Heyden－97b, Triggs－96]提出.

基于平面单应来计算多摄像机的方法在[Cross－99]中被使用,它用平面自标定来初始化 $\mathbf{t}^i$ 矢量.

由一个图像序列得到初始射影重构的方法在[Avidan－98,Beardsley－94,Beardsley－96,Fitzgibbon－98b,Laveau－96a,Sturm－97b]中介绍. [Torr－99]讨论了在序列重构中可能出现景物和运动退化这样一个重要的问题.

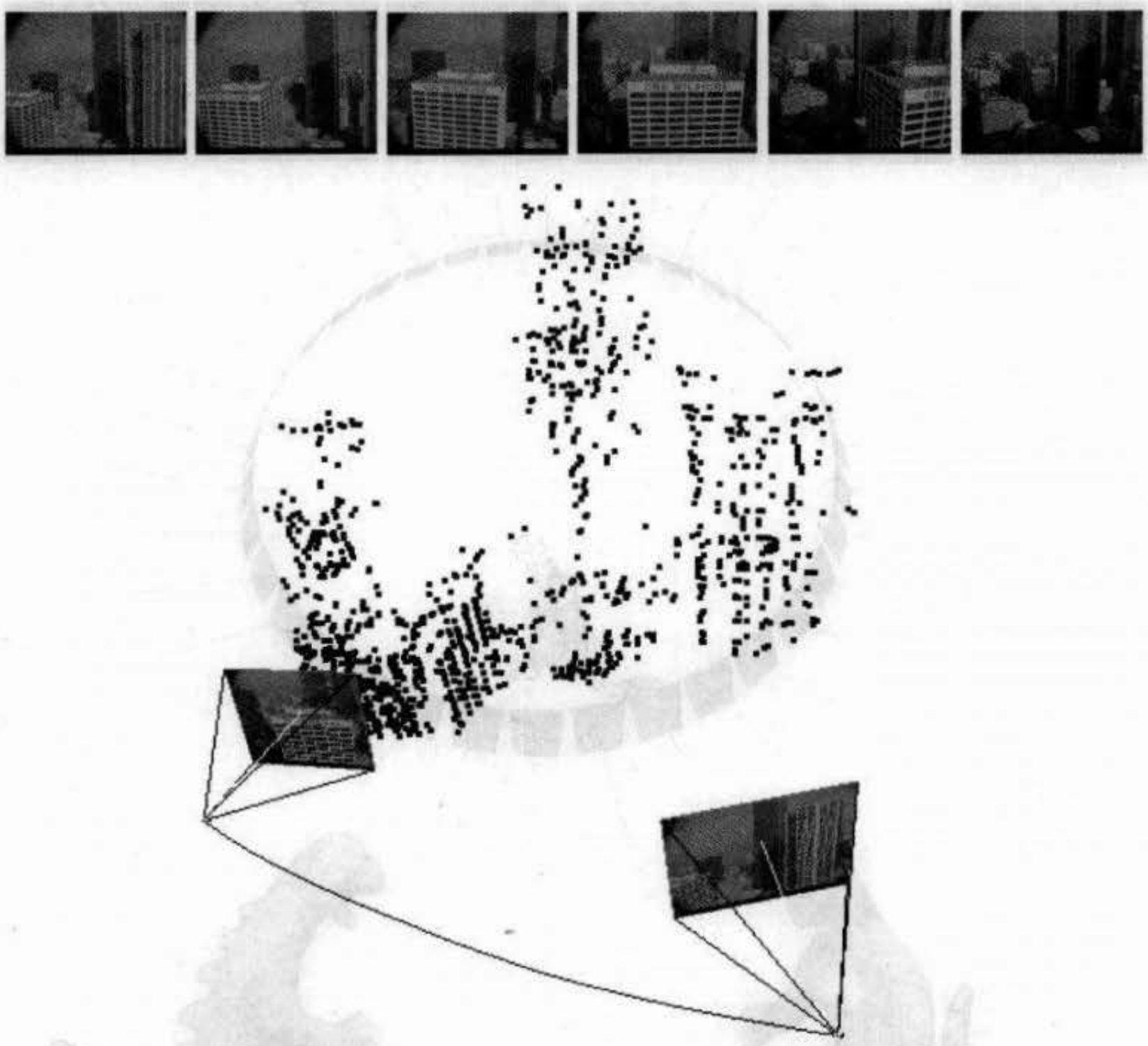


图 17.2　**Wilshire 大道**. 一个直升机拍摄的 350 帧图像的 3D 点和摄像机. 为清晰起见,只显示了起始和结束帧的摄像机,并把摄像机的路线画在其中.

# 第18章 自标定

自标定指直接由未标定的多幅图像来确定摄像机内参数的过程. 自标定一旦完成，就可由这些图像来计算度量重构. 自标定避免了用特殊的标定物体来标定摄像机的繁琐任务. 因而有极大的灵活性，比如，一个摄像机可以直接从图像序列中标定，尽管运动和某些内参数的变化未知.

这种方法的根源是摄像机作刚体运动时绝对二次曲线在此运动下保持不变. 反过来，如果能从图像上用某种方式确定3维空间中的一条唯一不动的二次曲线，则它就是 $\Omega_\infty$. 如前面各章所介绍：一旦 $\Omega_\infty$ 被辨认，度量几何便能计算. 目前针对确定 $\Omega_\infty$ 的任务，已有大量的自标定方法.

本章分四个主要部分. 首先我们给出自标定问题的代数结构，并且指出怎样从内参数或外参数约束中产生自标定方程. 第二，我们介绍自标定的若干种**直接**方法，它们涉及到计算绝对二次曲线或它的图像. 这些方法包括跨多幅视图来估计绝对对偶二次曲面或从视图对中估计 Kruppa 方程. 第三是一些**分层**自标定方法，它们包括两步——首先求解无穷远平面，然后利用它求解绝对二次曲线. 第四部分考虑一些特殊的配置：包括摄像机绕其中心旋转、摄像机平面运动和双眼装置的运动.

## 18.1 引言

自标定指由一组未标定图像来计算摄像机和/或景物的度量性质. 它不同于习惯上的标定，习惯上摄像机的标定矩阵 K 是从已知标定栅(第6章)或者景物的性质(比如正交方向的消影点(第7章))来确定. 而自标定中的度量性质是直接根据内和/或外参数的约束来确定.

例如，假设我们用内参数固定的摄像机获取了一组图像，并由跨这个图像集合的点对应来计算射影重构. 该重构计算每幅视图的射影摄像机矩阵 $P^i$. 我们的约束是真实的摄像机内参数矩阵 K 对于每幅视图都是相同的(但未知). 现在，每个摄像机 $P^i$ 的射影重构可以分解为 $P^i = K^i[R^i|\mathbf{t}^i]$，但一般每幅视图的标定矩阵 $K^i$ 将不相同. 这样一来，此射影重构不满足上述约束.

然而，我们可以用一个单应 H 通过摄像机矩阵变换来随意地改变射影重构. 因为真实摄像机是内参数固定的，所以存在一个单应(或一簇单应)使得变换后的矩阵 $P^iH$ 一定可以分解为 $P^iH = KR^i[I|\mathbf{t}^i]$，使每个摄像机的标定矩阵相同，因而使重构与上面的约束相一致. 如果有足够多的视图且视图之间是一般运动（见后面），则可用上面的一致性对 H 进行约束，使得被 H 变换后的重构至多与真实摄像机和景物相差一个相似变换，即我们得到了一个度量重构.

虽然用于取得度量重构的具体约束可以不同，但是下面的例子仍然说明了一般的过程：

(1) 获得一个射影重构$\{P^i, \mathbf{X}_j\}$

(2) 由自标定约束确定一个矫正单应 H，并由此变换到一个度量重构$\{P^iH, H^{-1}\mathbf{X}_j\}$.

下面的几节中将包括各具特色的自标定方法. 它们的差别在于所用的约束 以及确定单应 H 的方法. 这些方法可以分成两类:直接确定 H 和分层的方法,后一种方法首先确定 H 的射影然后再确定仿射成分. 它的好处是一旦得到仿射重构,即已知 $\boldsymbol{\pi}_\infty$，度量重构的解便是线性的.

如果我们的目的是摄像机标定而不是景物度量重构,则不一定总要计算一个显式的射影重构，这样有时摄像机标定可以更直接地计算而不用经过一个矫正变换. 例如摄像机绕其中心旋转而没有平移时就属于这种情形，它将在节 18.6 中讨论.

## 18.2 代数框架和问题陈述

假设我们得到了一个射影重构$\{P^i, \mathbf{X}_j\}$;然后基于摄像机内参数或运动的约束，我们希望确定一个矫正单应 H 使得$\{P^iH, H^{-1}\mathbf{X}_j\}$是一个度量重构.

我们从具有标定的摄像机和在欧式坐标系下表示的结构的真正度量实况开始. 因此，我们实际有 $m$ 个摄像机 $P_M^i$,它们把一个 3D 点 $\mathbf{X}_M$ 投影到每幅视图上的图像点$\mathbf{x}^i = P_M^i\mathbf{X}_M$，其中下标 M 表示摄像机是标定的并且世界坐标系是欧式的. 这些摄像机可以写为 $P_M^i = K^i[R^i|\mathbf{t}^i]$,其中 $i=1,\ldots,m$.

在一个射影重构中，我们得到摄像机 $P^i$，它与 $P_M^i$ 的关系是

$$P_M^i = P^iH \quad i=1,\ldots,m \tag{18.1}$$

其中 H 是三维空间的一个未知 4×4 单应. 我们的目标是确定 H.

准确地说,我们不打算考虑重构中的绝对旋转、平移和缩放,我们现在就来除去这个相似成份. 我们取世界坐标系与第一个摄像机重合，使得 $R^1 = I$ 和 $\mathbf{t}^1 = \mathbf{0}$. 然后 $R^i$ 和 $\mathbf{t}^i$ 指出的是第 $i$ 个摄像机和第一个摄像机之间的欧式变换，且 $P_M^1 = K^1[I|\mathbf{0}]$. 类似地，在该射影重构中,我们选取第一幅视图为通常的标准摄像机，使得 $P^1 = [I|\mathbf{0}]$. 令 H 为

$$H = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{v}^T & k \end{bmatrix}$$

则由(18.1)推出的关系 $P_M^1 = P^1H$ 变成$[K^1|\mathbf{0}] = [I|\mathbf{0}]H$,即 $A = K^1$ 和 $\mathbf{t} = \mathbf{0}$. 另外，由于 H 是非奇异的，$k$ 必须不是零，因此我们可以假设 $k=1$(这固定了重构的尺度). 因而可证明 H 具有如下形式:

$$H = \begin{bmatrix} K^1 & \mathbf{0} \\ \mathbf{v}^T & 1 \end{bmatrix}.$$

这样做就除去了相似成分.

矢量 $\mathbf{v}$ 和 $K^1$ 共同确定了该射影重构的无穷远平面，因为 $\boldsymbol{\pi}_\infty$ 的坐标是

$$\boldsymbol{\pi}_\infty = H^{-T}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \begin{bmatrix} (K^1)^{-T} & -(K^1)^{-T}\mathbf{v} \\ \mathbf{0} & 1 \end{bmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \begin{pmatrix} -(K^1)^{-T}\mathbf{v} \\ 1 \end{pmatrix}.$$

我们记 $\boldsymbol{\pi}_\infty=(\mathbf{p}^\mathsf{T},1)^\mathsf{T}$，其中 $\mathbf{p}=-(\mathrm{K}^1)^{-\mathsf{T}}\mathbf{v}$. 把迄今我们已经证明的总结为：

**结论 18.1　一个射影重构 $\{\mathrm{P}^i,\mathbf{X}_j\}$，其中 $\mathrm{P}^1=[\mathrm{I}|\mathbf{0}]$，能用如下形式的矩阵 H 变换到度量重构 $\{\mathrm{P}^i\mathrm{H},\mathrm{H}^{-1}\mathbf{X}_j\}$：**

$$\mathrm{H}=\begin{bmatrix}\mathrm{K} & \mathbf{0}\\ -\mathbf{p}^\mathsf{T}\mathrm{K} & 1\end{bmatrix} \tag{18.2}$$

**其中 K 是一个上三角矩阵. 而且有：**

**(1) $\mathrm{K}=\mathrm{K}^1$ 是第一个摄像机的标定矩阵.**

**(2) 该射影重构的无穷远平面的坐标由 $\boldsymbol{\pi}_\infty=(\mathbf{p}^\mathsf{T},1)^\mathsf{T}$ 给出.**

**反过来，如果在射影坐标系下的无穷远平面和第一个摄像机的标定矩阵已知，则把射影重构变换到度量重构的变换 H 由(18.2)给出.**

由这个结论可知为了把一个射影重构变换到度量重构，只要指定 8 个参数——$\mathbf{p}$ 的 3 个元素和 $\mathrm{K}^1$ 的 5 个元素就足够了. 这与几何的计数推导相符. 寻找度量重构等价于指定无穷远平面和绝对二次曲线(分别是 3 个和 5 个自由度). 在一个度量重构中，每个摄像机的标定 $\mathrm{K}^i$，和它相对于第一个摄像机的旋转 $\mathrm{R}^i$，以及在相差一个共同尺度的意义下相对于第一个摄像机的平移 $\mathbf{t}^i$，即 $\mathbf{t}^i\mapsto s\mathbf{t}^i$，都被确定.

我们现在来推导基本的自标定方程. 我们把射影重构的摄像机记为 $\mathrm{P}^i=[\mathrm{A}^i|\mathbf{a}^i]$. 乘出(18.1)并利用(18.2)给出

$$\mathrm{K}^i\mathrm{R}^i=(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})\mathrm{K}^1 \qquad i=2,\ldots,m. \tag{18.3}$$

由此又推出 $\mathrm{R}^i=(\mathrm{K}^i)^{-1}(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})\mathrm{K}^1$，$i=2,\ldots,m$. 最后，可以用 $\mathrm{RR}^\mathsf{T}=\mathrm{I}$ 消去旋转 $\mathrm{R}^i$ 得出

$$\mathrm{K}^i\mathrm{K}^{i\mathsf{T}}=(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})\mathrm{K}^1\mathrm{K}^{1\mathsf{T}}(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})^\mathsf{T}.$$

注意 $\mathrm{K}^i\mathrm{K}^{i\mathsf{T}}=\omega^*$ 是绝对二次曲线的对偶图像(或记为 DIAC)——见(7.9 - $p$146). 做了这个替换后给出自标定的基本方程

$$\begin{aligned}\omega^{*i}&=(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})\,\omega^{*1}(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})^\mathsf{T}\\ \omega^i&=(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})^{-\mathsf{T}}\,\omega^1(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})^{-1},\end{aligned} \tag{18.4}$$

第二个方程是第一个的逆，其中 $\omega$ 是绝对二次曲线的像(或记为 IAC). 这些方程把 $\omega^{*i}$ 或 $\omega^i$，$i=1,\ldots,m$ 的未知元素以及未知参数 $\mathbf{p}$ 与射影摄像机的**已知**参数 $\mathrm{A}^i,\mathbf{a}^i$ 相联系.

自标定的艺术在于利用关于 $\mathrm{K}^i$ 的约束，比如 $\mathrm{K}^i$ 的某一个元素是零，从(18.4)中产生关于 $\mathbf{p}$ 和 $\mathrm{K}^1$ 的八个参数的方程组. **所有的**自标定方法都是在求解这些方程上变化花样，在下面的几节中我们将介绍其中的若干种方法. 其过程一般是首先计算 $\omega^i$ 或 $\omega^{*i}$ 并从中抽取标定矩阵 $\mathrm{K}^i$ 的值——虽然迭代方法(例如捆集调整)可以直接由 $\mathrm{K}^i$ 参数化. 方程(18.4)可在几何上解释为在绝对二次曲线上的映射，我们将在节 18.3 和节 18.5.2 回来讨论它.

我们从一个简单的例子开始，这个例子解释如何由(18.4)产生关于上面提到的 8 个参数的方程.

**例 18.2　有相同 $\mathrm{K}^i$ 的自标定方程**

假设所有的摄像机的内参数相同，即 $\mathrm{K}^i=\mathrm{K}$，则(18.4)变成

$$\mathrm{KK}^\mathsf{T}=(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})\mathrm{KK}^\mathsf{T}(\mathrm{A}^i-\mathbf{a}^i\mathbf{p}^\mathsf{T})^\mathsf{T}, \qquad i=2,\ldots,m \tag{18.5}$$

每幅视图 $i=2,\ldots,m$ 提供一个方程，(原则上)我们能计算为确定 8 个未知参数所需的视图数. 除了第一幅视图外，每幅视图给出 5 个约束，因为这个方程每边是一个 $3\times3$ 的对称矩阵(即 6 个独立元素)，并且是齐次的. 假设每幅视图给出的约束是独立的，则只要 $5(m-1)\geqslant8$便可以得到一个解. 因此，只要 $m\geqslant3$，就能得到一个解，至少原则上是如此. 显然，如果 $m$ 远大于 3，则未知的 K 和 $\mathbf{p}$ 是非常超定的. △

人们可能会想到利用(18.5)作为直接估计矫正变换 H 的一个基础. 它可以构架成一个参数最小化问题，使(18.2)中 8 个参数变化的目标是最小化一个代价函数，它被用于刻划(18.4)被满足的程度或测量与度量结构接近的程度. 当然，还需要一种获得初始解的方法. 初始解和迭代最小化，这两个步骤是下面几节主要研究的内容——虽然它们所受约束比相同内参数的情形为少.

## 18.3 利用绝对对偶二次曲面标定

绝对对偶二次曲面 $Q_\infty^*$ 是一个退化的对偶(即平面)二次曲面并由一个秩为 3 的 $4\times4$ 齐次矩阵表示. $Q_\infty^*$ 的重要性在于以非常简明的方式同时包含了 $\boldsymbol{\pi}_\infty$ 和 $\Omega_\infty$，例如 $\boldsymbol{\pi}_\infty$ 是 $Q_\infty^*$ 的零矢量，并且它有一个代数上简单的图像投影：

$$\omega^* = PQ_\infty^* P^T. \tag{18.6}$$

它就是一个(对偶)二次曲面的投影方程(7.4—$p$139). 用文字表达就是 $Q_\infty^*$ 投影为绝对二次曲线的对偶图像 $\omega^* = KK^T$.

基于 $Q_\infty^*$ 的自标定思想是利用(18.6)通过(已知)摄像机矩阵 $P^i$ 把 $\omega^*$ 上的约束转移成 $Q_\infty^*$ 上的约束. 下文将说明这个方法在射影重构中可以根据 $K^i$ 的约束来确定表示 $Q_\infty^*$ 的矩阵. 其实在[Triggs－97]中引入 $Q_\infty^*$ 就是为了方便自标定的表示.

下文将证明一旦 $Q_\infty^*$ 被确定，矫正单应 H(18.2)也就被确定了. 因此我们得到一个自标定的一般框架：由特定 $K^i$ 的约束来确定 $Q_\infty^*$，然后由 $Q_\infty^*$ 确定 H. 这种一般的方法总结在算法 18.1 中. 在节 18.3.1 中，我们将集中介绍这个算法的第二步：$Q_\infty^*$ 的估计. 我们首先补充一些细节.

目标

给定跨若干视图的一组匹配点和有关标定矩阵 $K^i$ 的约束，计算点和摄像机的度量重构.

算法

(1) 由一组视图计算射影重构，得出摄像机矩阵 $P^i$ 和点 $\mathbf{X}_j$.

(2) 用(18.6)和由 $K^i$ 产生的加在形式 $\omega^{*i}$ 上的约束来估计 $Q_\infty^*$.

(3) 把 $Q_\infty^*$ 分解为 $H\tilde{I}H^T$，其中 $\tilde{I}$ 是矩阵 diag(1,1,1,0).

(4) 把 $H^{-1}$ 作用于点并把 H 作用于摄像机以得到度量重构.

(5) 用迭代的最小二乘最小化来改善解(见节 18.3.3)

或者，每一摄像机的标定矩阵可以直接计算如下：

(1) 用(18.6)对所有 $i$ 计算 $\omega^{*i}$.

(2) 用 Cholesky 分解由方程 $\omega^* = KK^T$ 计算标定矩阵 $K^i$.

算法 18.1 基于 $Q_\infty^*$ 的自标定.

**绝对对偶二次曲面的简单性质** 节 2.7($p$49)对 $Q_\infty^*$ 作了完整描述. 为方便自标定，在这里概述其重要的性质. 在欧式坐标系中 $Q_\infty^*$ 的标准形式是

$$\tilde{I}=\begin{bmatrix} I_{3\times 3} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}, \tag{18.7}$$

在射影坐标系下 $Q_\infty^*$ 的形式是 $Q_\infty^*=H\tilde{I}H^T$，其中 $\tilde{I}$ 是 (18.7)中的矩阵. 这是根据对偶二次曲面的射影变换规则 $Q_\infty^* \mapsto HQ_\infty^* H^T$(2.17 - $p$41)得到的. 因此有：

**结论 18.3 在一个任意射影坐标系下，对偶绝对二次曲面由一个对称 4×4 矩阵表示并具有下面的性质：**

**(1) 它是一个秩为 3 的奇异矩阵，因为 $Q_\infty^*$ 是一个退化锥面.**

**(2) 它的零空间是表示无穷远平面的矢量，因为 $Q_\infty^*\boldsymbol{\pi}_\infty=\mathbf{0}$.**

**(3) 它是半正定的(或半负定的——取决于齐次比例因子).**

在欧式坐标系中,这些性质可以从 $Q_\infty^*$ 的标准形式直接得到，并且很容易推广到任意坐标系.

**由 $Q_\infty^*$ 求矫正单应** 在射影坐标系下已知一个估计得到的 $Q_\infty^*$，我们希望确定单应 H. 而求 H 就是按如下方法对表达式进行分解.

**结论 18.4 如果 $Q_\infty^*$ 分解为 $Q_\infty^*=H\tilde{I}H^T$(见上文的记号)，则 $H^{-1}$ 是把射影坐标系变成欧式坐标系的一个 3D(点)单应.**

注意作用于摄像机的变换是作用于点的变换的逆，因此 H 是作用于摄像机使得 $P_M=PH$的矫正矩阵. 这样一来,H 是作用于摄像机的矫正变换. 可以容易地从 $Q_\infty^*$ 的特征值分解中把其形如 $H\tilde{I}H^T$ 的分解式计算出来(见 A3.2 节($p$406)中关于它的 Jacobi 算法).

**自标定方程的等价方程** 描述 $Q_\infty^*$ 的图像投影的方程(18.6)就是自标定方程(18.4)的一个几何表示,现在来说明这一点.

我们已知道在一个射影坐标系下 $Q_\infty^*$ 的形式是 $H\tilde{I}H^T$. 该射影重构是通过(18.2)与度量重构联系，更详细地写出得

$$Q_\infty^*=H\tilde{I}H^T=\begin{bmatrix} K^1K^{1T} & -K^1K^{1T}\mathbf{p} \\ -\mathbf{p}^TK^1K^{1T} & \mathbf{p}^TK^1K^{1T}\mathbf{p} \end{bmatrix}=\begin{bmatrix} \omega^{*1} & -\omega^{*1}\mathbf{p} \\ -\mathbf{p}^T\omega^{*1} & \mathbf{p}^T\omega^{*1}\mathbf{p} \end{bmatrix} \tag{18.8}$$

应用(18. 6)并取 $P^i=[A^i|\mathbf{a}^i]$,我们再一次得到自标定方程(18.4).

$$\omega^{*i}=P^iQ_\infty^*P^{iT}=(A^i-\mathbf{a}^i\mathbf{p}^T)\omega^{*1}(A^i-\mathbf{a}^i\mathbf{p}^T)^T.$$

这是(18.4)的几何解释——$Q_\infty^*$ 是在摄像机的欧式运动下的不动二次曲面，并且 DIAC $\omega^{*i}$ 是 $Q_\infty^*$ 在每一幅视图中的影像.

### 18.3.1 从一组图像线性求解 $Q_\infty^*$

这里的目标是在射影坐标系下直接由关于内参数的约束来估计 $Q_\infty^*$. 我们首先介绍三种线性解的情形. 在这里自然应该对 DIAC 和 IAC 的形式作一个总结. 请参考表18.1.

对于有形如(5.10－$p$104)的标定矩阵 K 的摄像机，$\omega=(KK^T)^{-1}$ 和 $\omega^*=\omega^{-1}=KK^T$ 的形式为

$$\omega^*=\begin{bmatrix}\alpha_x^2+s^2+x_0^2 & s\alpha_y+x_0y_0 & x_0\\ s\alpha_y+x_0y_0 & \alpha_y^2+y_0^2 & y_0\\ x_0 & y_0 & 1\end{bmatrix} \tag{18.9}$$

和

$$\omega=\begin{bmatrix}\alpha_y^2 & -s\alpha_y & -x_0\alpha_y^2+y_0s\alpha_y\\ -s\alpha_y & \alpha_x^2+s^2 & \alpha_ysx_0-\alpha_x^2y_0-s^2y_0\\ -x_0\alpha_y^2+y_0s\alpha_y & \alpha_ysx_0-\alpha_x^2y_0-s^2y_0 & \alpha_x^2\alpha_y^2+\alpha_x^2y_0^2+(\alpha_yx_0-sy_0)^2\end{bmatrix} \tag{18.10}$$

如果扭曲参数是零，即 $s=0$，那么表达式简化成

$$\omega^*=\begin{bmatrix}\alpha_x^2+x_0^2 & x_0y_0 & x_0\\ x_0y_0 & \alpha_y^2+y_0^2 & y_0\\ x_0 & y_0 & 1\end{bmatrix} \tag{18.11}$$

和

$$\omega=\frac{1}{\alpha_x^2\alpha_y^2}\begin{bmatrix}\alpha_y^2 & 0 & -\alpha_y^2x_0\\ 0 & \alpha_x^2 & -\alpha_x^2y_0\\ -\alpha_y^2x_0 & -\alpha_x^2y_0 & \alpha_x^2\alpha_y^2+\alpha_y^2x_0^2+\alpha_x^2y_0^2.\end{bmatrix} \tag{18.12}$$

表 18.1　用摄像机的内参数表示绝对二次曲线的图像 $\omega$ 和绝对二次曲线的对偶图像 $\omega^*$.

**指定关于 $Q_\infty^*$ 的线性约束**　如果主点已知，则可以得到 $Q_\infty^*$ 的线性约束. 假设主点是已知的，则可以使图像坐标系原点与主点相重合. 从而 $x_0=0, y_0=0$，并且根据表 18.1，DIAC 变成

$$\omega^*=\begin{bmatrix}\alpha_x^2+s^2 & s\alpha_y & 0\\ s\alpha_y & \alpha_y^2 & 0\\ 0 & 0 & 1\end{bmatrix}. \tag{18.13}$$

对每幅视图 $i$ 运用射影方程(18.6)$\omega^*=PQ_\infty^*P^T$，则关于 $Q_\infty^*$ 的线性方程由(18.13)中的零元素产生. 例如两个方程

$$(P^iQ_\infty^*P^{iT})_{13}=0 \quad 和 \quad (P^iQ_\infty^*P^{iT})_{23}=0 \tag{18.14}$$

由 $\omega_{13}^{*i}=\omega_{23}^{*i}=0$ 直接得到.

如果还有关于 $K^i$ 的导致 $\omega^*$ 元素之间进一步关系的约束，那么可以提供附加的线性方程. 比如，假设扭曲参数是零，则(18.13)中(1,2)元素为零，从而提供了关于 $Q_\infty^*$ 元素的类似于(18.14)的另一个线性方程. 知道长宽比也可提供一个进一步的约束. 表 18.2 总结了可能用得上的一些约束.

**线性解**　因为 $Q_\infty^*$ 是对称的，它有 10 个齐次线性参数，即位于对角线及其以上的 10 个元素. 这 10 个元素可以用一个 10 维矢量 $\mathbf{x}$ 表示. 按通常方法，关于 $Q_\infty^*$ 的线性方程可以组合成一个形如 $A\mathbf{x}=\mathbf{0}$ 的矩阵方程，并且通过 SVD 分解得到 $\mathbf{x}$ 的最小二乘解. 例如，从每幅视图，(18.14)的两个方程提供这个矩阵的两行. 从五幅图像总共得到 10 个方程(只假设主点

是已知的)，并且可能存在一个线性解．从四幅图像产生 8 个方程．如由 7 个点计算基本矩阵的方法一样，存在含 2 个参数的解簇．条件 $\det Q_\infty^* = 0$ 可给出一个四次方程，从而关于 $Q_\infty^*$ 至多有四个解．

| 条件 | 约束 | 类型 | #约束 |
|---|---|---|---|
| 零扭曲 | $\omega_{12}^*\omega_{33}^* = \omega_{13}^*\omega_{23}^*$ | 二次 | $m$ |
| 主点(p.p.)在原点 | $\omega_{13}^* = \omega_{23}^* = 0$ | 线性 | $2m$ |
| 零扭曲(p.p.在原点) | $\omega_{12}^* = 0$ | 线性 | $m$ |
| 固定(未知)长宽比<br>(零扭曲和 p.p.在原点) | $\omega_{11}^{*i}/\omega_{22}^{*i} = \omega_{11}^{*j}/\omega_{22}^{*j}$ | 二次 | $m-1$ |
| 已知长宽比 $r = \alpha_y/\alpha_x$<br>(零扭曲和 p.p.在原点) | $r^2\omega_{11}^* = \omega_{22}^*$ | 线性 | $m$ |

表 18.2 **由 DIAC 导出的自标定约束** "#约束"列中给出的数目是假定约束在每幅视图都为真时由 $m$ 幅视图给出的约束的总数．每项附加信息产生附加方程．例如如果主点已知并且扭曲参数是零，那么每幅视图有 3 个约束．

**例 18.5 可变焦距的线性解**

假设除了焦距外摄像机是标定的——主点已知，长宽比是 1(如果不是，则可变换该方程使得它从已知值变成 1)，以及扭曲参数是零——焦距未知而且可以在视图之间变化．从表 18.2 可知在这种情况下每幅视图产生 4 个关于 $Q_\infty^*$ 的线性约束．两幅视图就产生 8 个线性约束且利用条件 $\det Q_\infty^* = 0$ 最多得到 4 个解．如果 $m \geq 3$，则存在唯一的线性解． △

例 18.8 将对这种最小配置的焦距的确定进行讨论．

### 18.3.2 非线性求解 $Q_\infty^*$

我们现在来介绍从形式(18.6)得到的各种非线性方程．我们已知道 $\omega^{*i} = P^i Q_\infty^* P^{iT}$ 的每一个元素可以用 $Q_\infty^*$ 的参数表示为一个线性表达式．其结果是各种 $\omega^{*i}$ 的元素之间的任何关系变成关于 $Q_\infty^*$ 的元素的方程．具体地说，$\omega^{*i}$ 元素间的线性或二次关系分别产生 $Q_\infty^*$ 元素间的线性或二次关系．给定足够多的这样的方程，我们可以求解 $Q_\infty^*$．

**恒定内参数** 如果所有摄像机的内参数是相同的，则对于所有的 $i$ 和 $j$，$\omega^{*i} = \omega^{*j}$，它展开为 $P^i Q_\infty^* P^{iT} = P^j Q_\infty^* P^{jT}$．然而，由于它们都是齐次量，这些等式在相差一个未知尺度下成立．它们产生由五个方程组成的方程组：

$$\omega_{11}^{*i}/\omega_{11}^{*j} = \omega_{12}^{*i}/\omega_{12}^{*j} = \omega_{13}^{*i}/\omega_{13}^{*j} = \omega_{22}^{*i}/\omega_{22}^{*j} = \omega_{23}^{*i}/\omega_{23}^{*j} = \omega_{33}^{*i}/\omega_{33}^{*j}.$$

它给出 $Q_\infty^*$ 元素的二次方程组．给定三幅视图，总共产生可以用来求 $Q_\infty^*$ 的 10 个方程．

**扭曲参数为零时的标定** 在每个摄像机都为零扭曲的假设下，DIAC 简化为(18.11)那样的形式．具体地说，在零扭曲情形时，我们得到 $\omega^*$ 元素之间的如下约束

$$\omega_{12}^*\omega_{33}^* = \omega_{13}^*\omega_{23}^*. \tag{18.15}$$

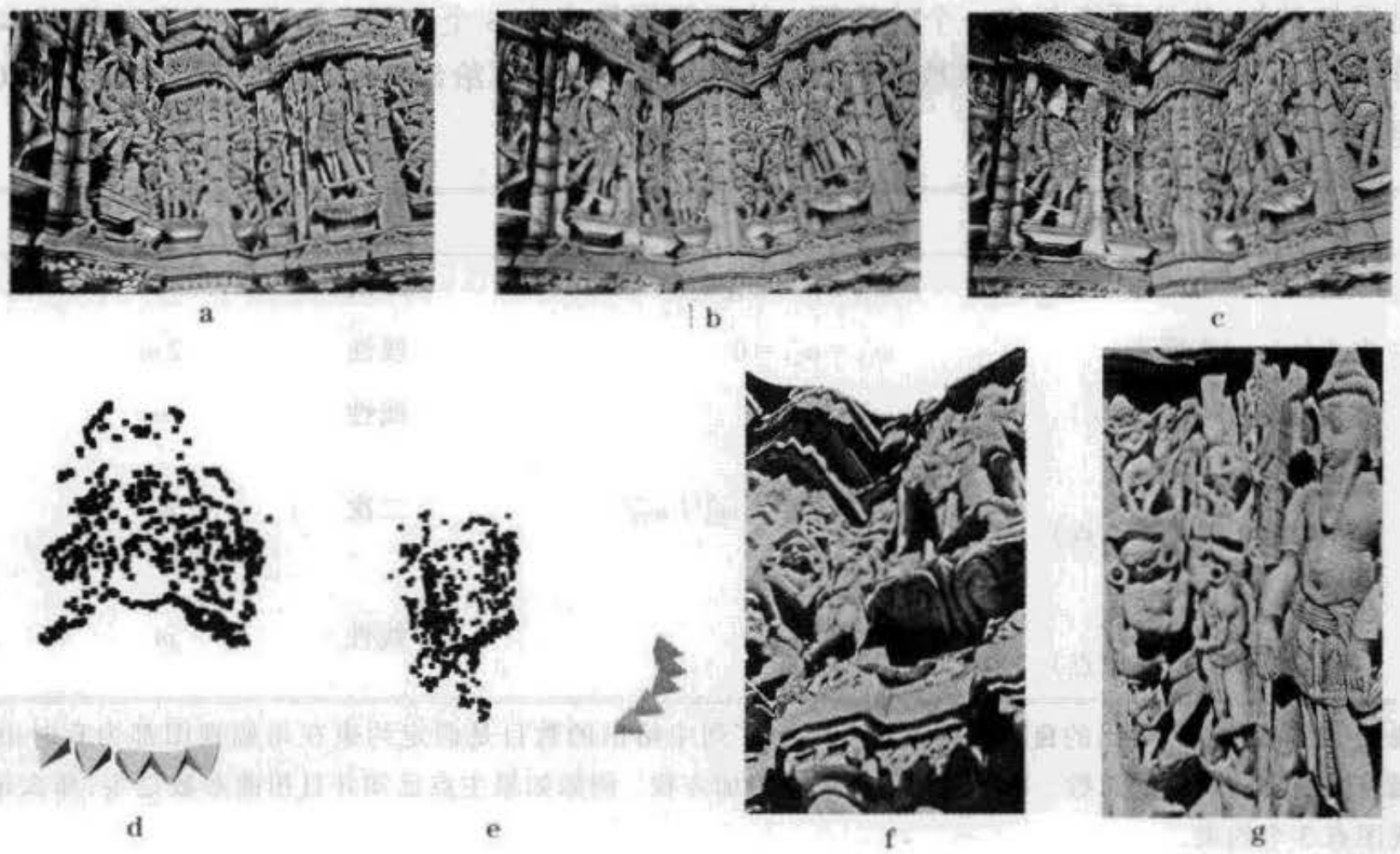



图 18.1 **一般运动的欧氏重构**. ($a$)—($c$)由手持摄像机获得的(5 幅中的)3 幅视图. ($d$)和($e$)由跨五幅视图的兴趣点的匹配算得的距离重构获得的两幅视图. 摄像机用顶点为所算得摄像机中心的角锥表示. ($f$)和($g$)利用基于双眼算法的区域由原图像和重构的摄像机算得的 3D 模型的纹理图的两幅视图.

上式给出了用 $Q_\infty^*$ 元素表示的一个二次方程. 从 $m$ 幅视图中,我们得到 $m$ 个二次方程. 而且由绝对对偶二次曲面是退化的这一事实可推导出另一个方程 $\det Q_\infty^* = 0$. 由于 $Q_\infty^*$ 有 10 个齐次线性参数,它可以从 8 幅视图 (至少原则上)中计算出来.

这些不同的标定约束也总结在表 18.2 中.

### 18.3.3 迭代方法

正如我们所看到的,在整本书中有许多场合都存在着最小化代数或几何误差的选择. 对于当前情形,一个合适的代数误差由(18.4)提供. 在从前的一些情形, 比如(3.1—$p$53), 未知的尺度因子通过叉乘来消去. 而这里的尺度因子可以用矩阵的范数来消去. 该代价函数是

$$\sum_i \| K^i K^{iT} - P^i Q_\infty^* P^{iT} \|_F^2, \tag{18.16}$$

其中 $\|M\|_F$ 是 M 的 Frobenius 范数, 且 $K^i K^{iT}$ 和 $P^i Q_\infty^* P^{iT}$ 都被归一化使其 Frobenius 范数为 1. 代价函数由 $Q_\infty^*$ 的(至多有 8 个)未知元素和所有 $\omega^{*i} = K^i K^{iT}$ 的未知元素参数化. 可以利用展开式(18.8)来参数化绝对对偶二次曲面. 比如, 在例 18.5 中,每幅视图的焦距是唯一未知量, (18.16)将在 $m+3$ 个参数上最小化. 这些参数包括每幅视图的焦距 $f^i$ 和 $\mathbf{p}$ 的三个分量. 注意这种参数化保证在整个最小化过程中 $Q_\infty^*$ 的秩为 3.

因为上面的代价函数没有特别的几何意义, 建议在它之后进行一个完整的捆集调整. 事实上, 给定一个好的初始线性估计, 我们可以直接地进行捆集调整. 正如节 17.1 所介绍

的，把标定参数的假设与完整的捆集调整合并不困难.

**例 18.6　一般运动的度量重构**

图 18.1(a－c)给出了用手持摄像机拍摄的印度古庙的视图. 如节 17.5 所介绍，一个射影重构由图像点对应计算得到，并且在恒定摄像机参数和已知主点的约束下利用算法 18.1 得到了度量重构. 在图 18.1(d)和(e)中显示了所算得的摄像机和 3D 点群. △

### 18.3.4　计数推导

我们已见到的约束有两种类型：参数为已知值；或参数在视图序列中固定但值未知. 实际使用的约束依赖于摄像机获取、数字化和剪裁以及最终形成图像的整个成像过程的物理环境. 例如，对于镜头变焦的图像序列来说，扭曲参数和长宽比可能是固定的（但未知），但是焦距和主点在整个序列中是变化的. 通常遇到的情况是像素为正方形的或者具有已知的长宽比，因此扭曲参数(它是零)和长宽比是已知的.

我们现在考虑完全确定一个度量重构所需的约束数.

为完成标定所必须计算的参数数是 8. 它等于绝对对偶二次曲面的基本参数数，包括尺度多义性和秩 3 约束. 考虑 $m$ 幅视图并假设在所有视图中 $k$ 个内参数已知，而 $f$ 个内参数在视图序列中是固定但未知的(其中 $k+f\leq 5$). 一个固定并已知的标定参数通过条件 $\omega^{*i}=P^iQ_\infty^*P^{iT}$ 为每幅视图提供一个约束，总共有 $mk$ 个约束. 一个固定但未知的标定参数提供的约束要少一个，因为只是未知参数的值不知道. 因此固定参数 $f$ 提供总共 $f(m-1)$ 个约束. 那么完成标定的要求是

$$mk+(m-1)f\geq 8.$$

表 18.3 给出若干种约束组合时 $m$ 的值. 记住退化结构中有些约束是相关的. 因此会增加所需的视图数.

| 条件 | 固定 $f$ | 已知 $k$ | 视图 $m$ |
|---|---|---|---|
| 恒常内参数 | 5 | 0 | 3 |
| 长宽比和扭曲参数已知<br>焦距和主点变化 | 0 | 2 | 4* |
| 长宽比和扭曲参数恒定<br>焦距和主点变化 | 2 | 0 | 5* |
| 扭曲参数为零，所有其他参数变化 | 0 | 1 | 8* |
| 主点已知，所有其他参数变化 | 0 | 2 | 4*，5(线性) |
| 主点已知，扭曲参数为零 | 0 | 3 | 3(线性) |
| 主点、扭曲参数和长宽比已知 | 0 | 4 | 2，3(线性) |

表 18.3　为了得到足以进行自标定的约束，在各种条件下所需的视图 $m$. 加星号表示可能有多个解，甚至在视图之间作一般运动时也如此.

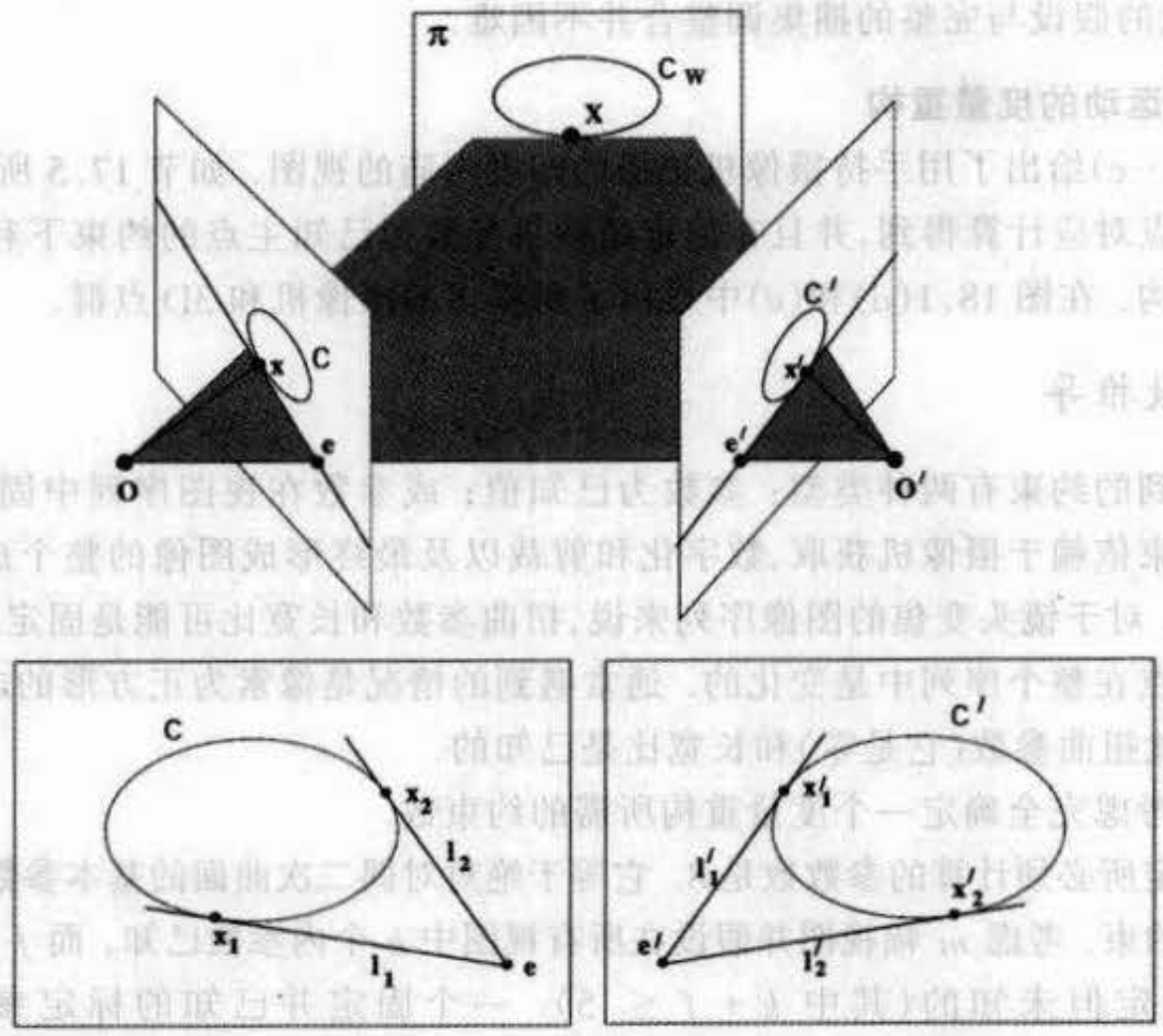



**图 18.2　二次曲线的对极切线**. 世界平面 $\pi$ 上的一条二次曲线 $C_w$ 影像到两视图上对应的二次曲线 $C \leftrightarrow C'$. 像二次曲线与视图对的对极几何相一致. 上图:世界平面二次曲线 $C_w$ 的切平面定义了对应的对极线,它们与像二次曲线相切. 下图:在第一幅视图中与像二次曲线相切的对极线 $\mathbf{l}_1, \mathbf{l}_2$ 分别对应于第二幅视图中与像二次曲线相切的对极线 $\mathbf{l}'_1, \mathbf{l}'_2$.

### 18.3.5　用绝对二次曲面方法标定的局限性

下面的考虑适用于采用这种方法的标定.

**最小二乘代数解的局限性**　例如,线性求解 $\omega^*$ 中的 $\mathbf{AX}=\mathbf{0}$ 的最小二乘解达到了最小化但并不受约束强制,所以得到的解不能准确地满足所要求的条件. 这种现象在超约束情形也出现. 如例 18.5 估计焦距的情形,元素 $\omega_{11}^{*i}$ 和 $\omega_{22}^{*i}$ 不是所需要的比率,并且对角线以外的元素也不准确地为零. 这意味着 $K^i$ 不能准确地表达所希望的形式.

用线性方法计算得到的绝对对偶二次曲面的秩一般不是 3,因为线性方程不施加这样的约束. $Q_\infty^*$ 的秩 3 矩阵可以在利用结论 18.1 求矫正变换时令最小特征值为零而得到. 另外一个方法是通过 SVD 分解来计算并用类似于在 8 点算法中(见节 10.1.1($p$192))计算秩 2 矩阵 F 的方法. 然后可以直接分解这个秩 3 矩阵以便利用结论 18.4 得到矫正单应(18.2). 另外,如节 18.3.3 所介绍,这个秩 3 矩阵可为迭代最小化提供初始值.

**正定条件**　这种方法最棘手的麻烦是强制 $Q_\infty^*$ 是半正定(或者如符号相反是半负定的). 这与条件:$\omega^* = PQ_\infty^*P^T$ 必须正定有关. 如果 $\omega^*$ 不是正定的,则不能用 Cholesky 分解来计算标定矩阵. 这是基于 IAC 或 DIAC 来估计自标定方法的一个反复出现的问题. 如果数据有噪声,则这个问题可能发生并表现为数据与度量重构不一致. 如果这种情况发生,则寻找最接近的半正定解是不恰当的,因为一般这将是导致谬误标定的一个边界情况.

## 18.4 kruppa 方程

另一种自标定方法涉及到运用 Kruppa 方程，它由 Faugeras、Luong 和 Maybank [Faugeras－92b]首先引入到计算机视觉，并且历史上被看作是第一个自标定方法．它们是两视图约束且只要求 F 已知，这些约束由 $\omega^*$ 元素的两个独立的二次方程组成．

Kruppa 方程是描述二次曲线的对极切线的对应的代数表示，这种对应的几何在图 18.2 中说明．假设世界平面上一条二次曲线 $C_w$ 在第一和第二幅视图上的影像分别是二次曲线 C 和 C′，而 $C^*$ 和 $C^{*\prime}$ 是它们的对偶．在第一幅视图中两条对极切线 $\mathbf{l}_1$ 和 $\mathbf{l}_2$ 可以组成一个退化的点二次曲线(见例 1.8($p7$))：$C_t=[\mathbf{e}]_\times C^*[\mathbf{e}]_\times$．(可以验证任何在直线 $\mathbf{l}_1$ 和 $\mathbf{l}_2$ 上的点 $\mathbf{x}$ 都满足 $\mathbf{x}^\mathsf{T}C_t\mathbf{x}=0$)．类似地，在第二幅视图中对应的对极线 $\mathbf{l}_1'$ 和 $\mathbf{l}_2'$ 可以被记为 $C_t'=[\mathbf{e}']_\times C^{*\prime}[\mathbf{e}']_\times$．这些对极切线在由任意世界平面 $\boldsymbol{\pi}$ 诱导的单应 H 下互相对应．因为 $C_t$ 是一个点二次曲线，根据结论 1.13($p11$)它的变换是 $C_t'=H^{-\mathsf{T}}C_tH^{-1}$，而这些直线的对应要求

$$[\mathbf{e}']_\times C^{*\prime}[\mathbf{e}']_\times=H^{-\mathsf{T}}[\mathbf{e}]_\times C^*[\mathbf{e}]_\times H^{-1}=FC^*F^\mathsf{T}. \tag{18.17}$$

最后一个等式出自 $F=H^{-\mathsf{T}}[\mathbf{e}]_\times$(见 $p236$)．注意，这个方程没有强制极切线各自映射到它们对应直线的条件，只是强制它们的对称乘积映射到对应的对称乘积．

至止，上面的推导适用于任何二次曲线．然而，在这里感兴趣的是世界二次曲线在无穷远平面上的绝对二次曲线的情形：$C^*=\omega^*$，$C^{*\prime}=\omega^{*\prime}$(以及 $H=H_\infty$)，则(18.17)具体化为

$$[\mathbf{e}']_\times\omega^{*\prime}[\mathbf{e}']_\times=F\omega^*F^\mathsf{T}. \tag{18.18}$$

如果内参数跨视图时总保持恒定，则 $\omega^{*\prime}=\omega^*$，从而 $[\mathbf{e}']_\times\omega^*[\mathbf{e}']_\times=F\omega^*F^\mathsf{T}$，这是最初由 Viéville[Viévill－95]给出的 Kruppa 方程的形式．消去齐次比例因子，我们得到 $\omega^*$ 元素的二次方程．

虽然(18.18)简洁地表示了 Kruppa 方程，但它不是一种易于使用的形式．我们现在给出 Kruppa 方程的一种简单而易用的形式．我们证明 $[\mathbf{e}']_\times$ 的零空间(它在(18.18)的两边都是相同的)能被消去并剩下一个由两个 3 维矢量组成的方程．

**结论 18.7　Kruppa 方程(18.18)等价于**

$$\begin{pmatrix}\mathbf{u}_2^\mathsf{T}\omega^{*\prime}\mathbf{u}_2\\ -\mathbf{u}_1^\mathsf{T}\omega^{*\prime}\mathbf{u}_2\\ \mathbf{u}_1^\mathsf{T}\omega^{*\prime}\mathbf{u}_1\end{pmatrix}\times\begin{pmatrix}\sigma_1^2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_1\\ \sigma_1\sigma_2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_2\\ \sigma_2^2\mathbf{v}_2^\mathsf{T}\omega^*\mathbf{v}_2\end{pmatrix}=0, \tag{18.19}$$

**其中 $\mathbf{u}_i$，$\mathbf{v}_i$ 和 $\sigma_i$ 是 F 的 SVD 的列矢量和奇异值．它提供了关于 $\omega^*$ 元素 $\omega_{ij}^*$ 的三个二次方程，其中两个是独立的．**

**证明**　基本矩阵的秩为 2，从而其 SVD 分解式为

$$F=UDV^\mathsf{T}=U\begin{bmatrix}\sigma_1 & & \\ & \sigma_2 & \\ & & 0\end{bmatrix}V^\mathsf{T}$$

其中零矢量是 $F^\mathsf{T}\mathbf{u}_3=\mathbf{0}$ 和 $F\mathbf{v}_3=\mathbf{0}$，这意味着对极点是 $\mathbf{e}=\mathbf{v}_3$ 和 $\mathbf{e}'=\mathbf{u}_3$．把上述表达式代入(18.18)得到

$$[\mathbf{u}_3]_\times \omega^{*\prime}[\mathbf{u}_3]_\times = \mathrm{UDV}^\mathsf{T}\omega^* \mathrm{VDU}^\mathsf{T}. \tag{18.20}$$

我们现在利用 U 是正交矩阵的性质. 对(18.20)左乘 $\mathrm{U}^\mathsf{T}$ 和右乘 U，则其左边变成

$$\mathrm{U}^\mathsf{T}[\mathbf{u}_3]_\times \omega^{*\prime}[\mathbf{u}_3]_\times \mathrm{U} = [\mathbf{u}_2 \quad -\mathbf{u}_1 \quad \mathbf{0}]^\mathsf{T}\omega^{*\prime}[\mathbf{u}_2 \quad -\mathbf{u}_1 \quad \mathbf{0}]$$

$$= \begin{bmatrix} \mathbf{u}_2^\mathsf{T}\omega^{*\prime}\mathbf{u}_2 & -\mathbf{u}_2^\mathsf{T}\omega^{*\prime}\mathbf{u}_1 & 0 \\ -\mathbf{u}_1^\mathsf{T}\omega^{*\prime}\mathbf{u}_2 & \mathbf{u}_1^\mathsf{T}\omega^{*\prime}\mathbf{u}_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

而(18.20)的右边变成

$$\mathrm{DV}^\mathsf{T}\omega^* \mathrm{VD} = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{bmatrix} \mathrm{V}^\mathsf{T}\omega^* \mathrm{V} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{bmatrix} = \begin{bmatrix} \sigma_1^2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_1 & \sigma_1\sigma_2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_2 & 0 \\ \sigma_1\sigma_2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_2 & \sigma_2^2\mathbf{v}_2^\mathsf{T}\omega^*\mathbf{v}_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

显然两边都化简为只有三个独立元素的对称矩阵. 两边的三个元素可以表示为一个齐次 3 维矢量：

$$\mathbf{x}_{\mathrm{LHS}}^\mathsf{T} = (\mathbf{u}_2^\mathsf{T}\omega^{*\prime}\mathbf{u}_2, \quad -\mathbf{u}_1^\mathsf{T}\omega^{*\prime}\mathbf{u}_2, \quad \mathbf{u}_1^\mathsf{T}\omega^{*\prime}\mathbf{u}_1)$$

$$\mathbf{x}_{\mathrm{RHS}}^\mathsf{T} = (\sigma_1^2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_1, \quad \sigma_1\sigma_2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_2, \quad \sigma_2^2\mathbf{v}_2^\mathsf{T}\omega^*\mathbf{v}_2).$$

两边在只相差一个尺度因子的意义下相等. 于是，按通常方法用矢量叉乘得出所要求的等式：$\mathbf{x}_{\mathrm{LHS}} \times \mathbf{x}_{\mathrm{LHS}} = \mathbf{0}$. 另一种推导在[Hartley－97b]中给出. □

注意，Kruppa 方程涉及到的不是 IAC 而是 DIAC，因为该方程由切线约束产生，而直线约束利用对偶(线)二次曲线更容易表示.

我们现在来讨论 Kruppa 方程的解，我们先从一个简单的例子开始，此例中除了焦距外其余所有的内参数都已知. 求解这个例子的另一种方法是利用绝对对偶二次曲线并已在例 18.5 中给出.

**例 18.8　一对视图的焦距**

假设两个摄像机是零扭曲并且主点和长宽比已知，但是焦距未知并且两个焦距不同(如例 18.5). 则从(18.11)通过适当的坐标系变换，它们的 DIAC 可以记为，

$$\omega^* = \mathrm{diag}(\alpha^2, \alpha^2, 1), \quad \omega^{*\prime} = \mathrm{diag}(\alpha'^2, \alpha'^2, 1)$$

其中 $\alpha, \alpha'$ 分别是第一和第二幅视图的未知的焦距. 记 Kruppa 方程(18.19)为

$$\frac{\mathbf{u}_2^\mathsf{T}\omega^{*\prime}\mathbf{u}_2}{\sigma_1^2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_1} = -\frac{\mathbf{u}_1^\mathsf{T}\omega^{*\prime}\mathbf{u}_2}{\sigma_1\sigma_2\mathbf{v}_1^\mathsf{T}\omega^*\mathbf{v}_2} = \frac{\mathbf{u}_1^\mathsf{T}\omega^{*\prime}\mathbf{u}_1}{\sigma_2^2\mathbf{v}_2^\mathsf{T}\omega^*\mathbf{v}_2}.$$

显然每个分子关于 $\alpha'^2$ 是线性的，每一个分母关于 $\alpha^2$ 是线性的，交叉相乘提供了关于 $\alpha^2$ 和 $\alpha'^2$ 的两个简单的二次方程，它们是容易求解的. $\alpha$ 和 $\alpha'$ 的值由开平方得到. 注意，如果两幅视图的内参数相同（即 $\alpha = \alpha'$），那么结论 18.7 的每一个方程提供只有一个未知量 $\alpha^2$ 的一个二次方程. △

**把 Kruppa 方程推广到多视图**　当内参数跨视图恒常但未知时，Kruppa 方程提供关于 5 个未知参数的两个独立约束. 因此给定三幅视图并已知每对视图之间的 F 时，一般有 6 个二次约束，它仍足以确定 $\omega^*$. 利用其中的任何 5 个方程产生关于 5 个未知量的五个二次式，总共有 $2^5$ 种可能的解. 解这组方程不是特别有希望的方法，虽然通过同伦连续[Luong－92]和最小化序列图像中每对视图的代数残差等方法已得到过解[Zeller－96].

**多义性** 如果视图之间没有旋转则 Kruppa 方程不对 $\omega^*$ 产生约束. 这一点可以从(18.18)看出:在纯平移情形, 该式化简成$[\mathbf{e}']_\times \omega^* [\mathbf{e}']_\times = [\mathbf{e}']_\times \omega^* [\mathbf{e}']_\times$, 因为 $F=[\mathbf{e}']_\times$.

Kruppa 方程与在无穷单应下由 IAC 转移所提供的标定约束紧密相关, 这方面的内容将在后面的节 18.5.2 中讨论. 由后面的讨论可知:对于一对视图的情形,Kruppa 方程加在 $\omega^*$ 上的约束比无穷单应矩阵所加的约束(18.25)弱. 因此,由 Kruppa 方程引进的 $\omega^*$ 的多义性是由 (18.25)引进的多义性的超集.

把 Kruppa 方程用于三幅或更多幅视图时,提供的约束较之其他方法,比如模约束(节 18.5.1)或绝对对偶二次曲线(节 18.3.1)更弱. 这是因为 Kruppa 约束是关于一对视图**二次曲线**的约束,它们由 3D(对偶)的**二次曲面**投影产生. 它们不强制该(对偶)二次曲面是退化的, 或者等价地不强制 $\Omega_\infty$ 跨多视图由公共平面支持. 因此,如[Sturm-97b]所介绍的会产生附加的多义性解.

虽然 Kruppa 方程应用于自标定是文献记录上的第一例,但是由于它们在求解方面的困难性和多义性问题,所以面对更具有吸引力的方法(比如对偶二次曲面公式),它们已失去了宠爱. 然而, 如果仅给定两幅视图时,那么**仅** Kruppa 方程能提供关于 $\omega^*$ 的约束.

## 18.5 分层求解

确定一个度量重构涉及到同时得到摄像机标定矩阵 K 和无穷远平面 $\boldsymbol{\pi}_\infty$. 另一种途径是用某些方法首先求得 $\boldsymbol{\pi}_\infty$, 或等价地求得一个仿射重构. 接下来 K 的确定则相对比较简单,因为存在一个线性解. 我们现在就来介绍这种途径,从确定 $\boldsymbol{\pi}_\infty$ 的方法开始,然后求得 $H_\infty$,接着由已知 $H_\infty$ 计算 K 的方法.

### 18.5.1 仿射重构——确定 $\boldsymbol{\pi}_\infty$

对于一般运动和恒常内参数的情形,可以通过重新安排 (18.3 - $p$325)来提供仅关于 $\boldsymbol{\pi}_\infty$ 的一个约束, 称为**模约束**. 它允许直接求解 $\boldsymbol{\pi}_\infty$ 的坐标 $\mathbf{p}$, 具体介绍如下.

**模约束**

模约束是关于 $\boldsymbol{\pi}_\infty$ 坐标的一个多项式方程. 假设内参数恒定; 则式(18.3 - $p$326)中 $K'=K$,并有

$$A - \mathbf{a}\mathbf{p}^\mathsf{T} = \mu KRK^{-1} \tag{18.21}$$

其中比例因子 $\mu$ 被显式地给出, 而且为了简洁上标被省略了. 由于 $KRK^{-1}$ 共轭于一个旋转, 它的特征值为$\{1, e^{i\theta}, e^{-i\theta}\}$. 从而,$(A-\mathbf{a}\mathbf{p}^\mathsf{T})$的特征值是$\{\mu, \mu e^{i\theta}, \mu e^{-i\theta}\}$, 因此有相同的模, 这就是关于无穷远平面坐标 $\mathbf{p}$ 的模约束.

为推导这个约束, 进一步考虑 $A-\mathbf{a}\mathbf{p}^\mathsf{T}$ 的特征多项式:

$$\begin{aligned}\det(\lambda I - A + \mathbf{a}\mathbf{p}^\mathsf{T}) &= (\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)\\ &= \lambda^3 - f_1\lambda^2 + f_2\lambda - f_3\end{aligned}$$

其中 $\lambda_i$ 是三个特征值, 且

$$f_1 = \lambda_1+\lambda_2+\lambda_3 = \mu(1+2\cos\theta)$$

$$f_2 = \lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3 = \mu^2(1+2\cos\theta)$$

$$f_3 = \lambda_1\lambda_2\lambda_3 = \mu^3$$

消去尺度 $\mu$ 和角度 $\theta$ 我们得到

$$f_3 f_1^3 = f_2^3.$$

更仔细地观察这个特征多项式，我们发现 $\mathbf{p}$ 只是秩为 1 的项 $\mathbf{a}\mathbf{p}^\mathsf{T}$ 的一部分. 这意味着 $\mathbf{p}$ 的元素线性地出现在行列式 $\det(\lambda \mathrm{I} - \mathrm{A} + \mathbf{a}\mathbf{p}^\mathsf{T})$ 中，从而在 $f_1$，$f_2$，$f_3$ 中都是线性出现. 因此模约束可以写为 $\mathbf{p}$ 的三个元素 $p_i$ 的一个**四次**多项式. 这个多项式方程只是特征值模相等的必要条件，而不是充分条件.

每对视图产生关于 $\boldsymbol{\pi}_\infty$ 的坐标的一个四次方程. 因此，原则上，三幅视图确定 $\boldsymbol{\pi}_\infty$，它是 3 个变量的四次多项式的交点——有 $4^3$ 个可能的解. 模约束还可以与景物信息结合起来. 例如，如果在两幅视图中有一组对应的消影直线，则 $\boldsymbol{\pi}_\infty$ 在单参数多义性下被确定. 运用此模约束重解这个多义性，得到一个单变量的四次方程.

模约束可以看作是 Kruppa 方程的堂兄弟：Kruppa 方程是关于 $\omega^*$ 的方程但不涉及 $\boldsymbol{\pi}_\infty$；反过来，模约束是关于 $\boldsymbol{\pi}_\infty$ 的方程而不涉及 $\omega^*$. 一旦 $\omega^*$ 或 $\boldsymbol{\pi}_\infty$ 中的一个已知，则另外一个跟着就被确定.

**求 $\boldsymbol{\pi}_\infty$ 的其他方法**

因为存在同时求解四次方程组的问题，模约束不是求无穷远平面的一种很满意的实际方法. 事实上，求无穷远平面是自标定最困难的部分而且是人们最可能陷入困境的地方.

无穷远平面可以用各种其他方法来辨认. 其中的若干方法在第 9 章介绍过. 一个直接的方法(它超出纯自标定的范围)是利用场景几何的性质. 例如，两幅视图之间的一组消影点对应确定 $\boldsymbol{\pi}_\infty$ 上的一个点，在射影重构中三个这样的点确定 $\boldsymbol{\pi}_\infty$. 事实上，$\boldsymbol{\pi}_\infty$ 的一个合理近似可以从远景中的三个点的对应来得到. 第二种方法是运用两幅视图之间的纯平移运动，即让摄像机平移但不旋转或者不改变内参数；通过这样的运动能唯一确定 $\boldsymbol{\pi}_\infty$.

如结论 18.3($p$327)所示，无穷远平面也可以从绝对对偶二次曲面计算得到，而且如果主点已知，则这个方法是非常吸引人的. 利用正负性不等式来界定 $\boldsymbol{\pi}_\infty$ 的位置的方法将在第 20 章介绍. 它们利用有关点在摄像机前面的信息来得到接近于仿射重构的**准仿射**重构. 从这个初始准仿射重构开始，通过迭代求无穷远平面的一种方法在[Hartley－99]中介绍. 最近，[Hartley－99]用正负性来界定 3D 参数空间的一个矩形区域，而且无穷远平面矢量 $\mathbf{p}$ 必然在矩形区域中. 然后在这个区域内，采用一种穷尽搜索来寻找难以捕获的无穷远平面.

### 18.5.2 仿射到度量的转换——给定 $\boldsymbol{\pi}_\infty$ 确定 K

一旦无穷远平面被确定，一个仿射重构实际就已知了. 剩下的步骤是把仿射重构变换成度量重构. 与从射影变换到仿射的步骤相比，它是一个非常简单的步骤. 事实上，已有基于 IAC 或它的对偶的变换的线性算法.

**无穷单应** 无穷单应 $\mathrm{H}_\infty$ 是由无穷远平面 $\boldsymbol{\pi}_\infty$ 诱导的两幅图像之间的平面射影变换(见第 12 章). 如果在任何射影坐标下，无穷远平面 $\boldsymbol{\pi}_\infty = (\mathbf{p}^\mathsf{T}, 1)^\mathsf{T}$ 和摄像机矩阵 $[\mathrm{A}^i \mid \mathbf{a}^i]$ 已知，则关于无穷单应的一个显示公式可以由结论 12.1($p$228)推导出来：

$$\mathrm{H}_\infty^i = \mathrm{A}^i - \mathbf{a}^i\mathbf{p}^\mathsf{T} \tag{18.22}$$

其中 $H_\infty^i$ 表示从摄像机$[I|\mathbf{0}]$到摄像机$[A^i|\mathbf{a}^i]$的单应. 因此一旦无穷远平面已知,$H_\infty^i$ 可以从射影重构计算出来.

如果第一个摄像机不是标准形式$[I|\mathbf{0}]$, 则仍然能用下式计算第一幅到第 $i$ 幅图像的单应 $H_\infty^i$

$$H_\infty^i = (A^i - \mathbf{a}^i\mathbf{p}^T)(A^1 - \mathbf{a}^1\mathbf{p}^T)^{-1}. \tag{18.23}$$

然而这不是绝对必要的, 因为人们可以**创造**一幅新视图使其摄像机矩阵**取**标准形式$[I|\mathbf{0}]$, 并相对于这幅视图表示各个无穷单应. 在下面的讨论中, 我们用 K 和 $\omega$(没有上标)表示取标准形式参考视图或第一幅视图.

绝对二次曲线在 $\boldsymbol{\pi}_\infty$ 上,因此它在视图之间的像由 $H_\infty^i$ 映射. 在点变换 $\mathbf{x}^i = H_\infty^i\mathbf{x}$ 下(其中 $H_\infty^i$ 是参考视图与视图 $i$ 之间的无穷单应), 由对偶二次曲线(结论 1.14($p$11))和点二次曲线(结论 1.13(p11))的变换规则导出关系式:

$$\omega^{*i} = H_\infty^i\omega^* H_\infty^{iT} \text{和} \omega^i = (H_\infty^i)^{-T}\omega(H_\infty^i)^{-1} \tag{18.24}$$

其中 $\omega^i$ 是第 $i$ 幅视图中的 IAC. 可以验证这些方程就是自标定方程(18.4 - $p$325), 而且是自标定方程的另一种几何解释.

上面的关系属于自标定关系中最重要的关系. 它们是由仿射重构获得度量重构的基础, 并且如我们将在后面的第 18.6 节所看到的,它们也可以用来标定非平移摄像机. 这个关系对自标定的重要性在于如果 $H_\infty^i$ 已知, 则它们表示 $\omega^i$ 和 $\omega$ 之间(并且对 $\omega^*$ 也类似)的**线性**关系. 这意味着 $\omega^i$ 在一幅视图中的约束能容易地转移到另一幅视图,因而用这个方法可以组合充分多的约束足以用线性方法确定 $\omega$. 一旦 $\omega$ 被确定, 则可由 Cholesky 分解得到 K. 我们将以固定内参数的例子来说明这种方法.

**内参数恒定时的解法** 如果 $m$ 幅视图的内参数是恒定的, 则对 $i=1,\ldots m$, $K^i = K$ 且 $\omega^{*i} = \omega^*$,那么(18.24)的 $\omega^*$ 方程变为

$$\omega^* = H_\infty^i\omega^* H_\infty^{iT}. \tag{18.25}$$

这里方程(18.25) 的尺度因子非常关键.

- **虽然(18.25)表示齐次量之间的一个关系, 如果归一化 $H_\infty^i$ 使得 $\det H_\infty^i = 1$, 则在齐次方程中的尺度因子可以选为 1.**

因此可以得到产生关于对称矩阵 $\omega^*$ 的独立元素的 6 个方程. 然后(18.25)可以写成齐次线性形式

$$A\mathbf{c} = \mathbf{0} \tag{18.26}$$

其中 A 是由 $H_\infty^i$ 的元素组成的一个 $6\times6$ 矩阵, 而 $\mathbf{c}$ 是记作 6 维矢量的二次曲线 $\omega^*$. 如下面所讨论的, $\mathbf{c}$ 不能由一个这样的方程唯一确定,因为 A 的秩最多为 4. 然而, 如果把 $m \geq 2$对视图的线性方程(18.26)组合起来, 使 A 是一个 $6m\times6$ 的矩阵, 而且如果视图之间绕不同轴旋转, 则一般 $\mathbf{c}$ 可以被唯一地确定.

与唯一性问题相联系的是数值稳定性问题. 在单个运动下,由 $H_\infty^i$ 线性地计算 K 对 $H_\infty^i$ 的准确性极其敏感. 如果 $H_\infty^i$ 不准确则可能得不到正定矩阵 $\omega$(或者 $\omega^*$)以及(运用 Cholerky 分解)K. 如果有更多的运动以及 $\omega$ 是由若干 $H_\infty^i$ 组合的约束得到,那么这种敏感性会减小.

**采用 IAC 的好处** 类似于线性求解 $\omega^*$ 的方法，由(18.24)也可以得到 $\omega$ 的线性解. 事实上采用涉及 IAC 的方程是有诱感力的，其理由如下. 对于零扭曲的情形，(18.12—$p$328)中的 IAC 比对应公式(18.11——$p$328)中的 DIAC 更简单且更清楚地反映了每个标定参数的作用. 为了用 DIAC 方程(18.11)得到线性方程，必须假设主点已知. 而由 IAC 推导方程时这个假设不是必要的. 零扭曲的假设是非常自然的并且对大多数成像条件而言也是一个安全的假设. 然而，主点已知的假设就很站不住脚，正是由于这一点，通常自标定中优先采用(18.24)的 IAC 约束而不是 DIAC 约束.

**其他的标定约束** 刚才介绍的算法针对内参数恒定但任意的. 如果对 K 知道得更多，比如知道长宽比的值或零扭曲，对应的约束就可直接地加入到关于 $\omega$(或 $\omega^*$)的方程组中. 表 18.2($p$329)给出由 DIAC 导出的约束（在节 18.3 中相同的约束被用来计算 $Q_\infty^*$），而由 IAC 导出的约束在表 18.4 中.

| 条件 | 约束 | 类型 | #约束 |
|---|---|---|---|
| 零扭曲 | $\omega_{12}=0$ | 线性 | $m$ |
| 主点在原点 | $\omega_{13}=\omega_{23}=0$ | 线性 | $2m$ |
| 已知长宽比 $r=\alpha_y/\alpha_x$(假定零扭曲) | $\omega_{11}=r^2\omega_{22}$ | 线性 | $m$ |
| 固定(未知)长宽比(假定零扭曲) | $\omega_{11}^i/\omega_{22}^i=\omega_{22}^j/\omega_{22}^j$ | 二次 | $m-1$ |

表 18.4 **由 IAC 导出的自标定约束** 这些约束直接由(18.10－$p$328)和(18.12－$p$328)的形式推导出来. 假定该约束对每一视图都为真时，"#约束"列给出在 $m$ 幅视图上的约束总数.

如上文所提到的，从 IAC 导出的约束一般是线性的. 然而从 DIAC 导出的约束，只有在主点是已知(并在原点)的假设下才是线性的.

与绝对对偶二次曲面的方法一样，可以允许摄像机内参数变化，只要有足够多的约束. 当摄像机的内参数恒常时，$m$ 幅视图的标定参数提供总共 $5(m-1)$个约束. 如果让某些参数变化，我们就可以用更少的约束来达到目的. 这种具有内参数变量的标定方法非常类似于在绝对对偶二次曲面中所采用的方法. 根据(18.24)，$\omega^i$ 的每个元素可以表示为 $\omega$ 元素的线性表达式. 由此，在 $\omega^i$ 的某些元素上的一个线性约束可反向映射到在 $\omega$ 元素上的线性约束.

**注意**：为了避免对第一幅图像作不同的处理，任何加在第一个摄像机的约束，比如 $\omega_{12}^1=0$(摄像机 1 有零扭曲)，都应该把它并到整个方程组来处理，而不是减少描述 $\omega^1$ 的参数数. 后一种方法会导致零扭曲约束仅强制在第一幅图像上(一个硬约束)，但是它在其他图像上将是一个软约束.

算法 18.2 概括了参数恒定和变化时的分层算法.

我们可以直接把这个算法装在机器人上的摄像机中：为了确定 $\pi_\infty$，摄像机首先作一个纯平移运动；其后，摄像机可以既平移又旋转直到积累充分多的旋转以便达到唯一地确定 K 的目的.

目标

给定一个射影重构$\{P^i, \mathbf{X}_j\}$，其中$P^i = [A^i | \mathbf{a}^i]$，通过一个中介的仿射重构确定度量重构.

算法

(1)**仿射矫正**：用节 18.5.1 中介绍的某种方法确定定义$\pi_\infty$的矢量$\mathbf{p}$. 因此，可以得到一个仿射重构$\{P^i H_p, H_p^{-1}\mathbf{X}_j\}$，其中

$$H_p = \begin{bmatrix} I & \mathbf{0} \\ -\mathbf{p}^T & 1 \end{bmatrix}.$$

(2)**无穷单应**：按下式计算参考视图和其他视图之间的无穷单应

$$H_\infty^i = (A^i - \mathbf{a}^i \mathbf{p}^T).$$

归一化该矩阵使得$\det H_\infty^i = 1$.

(3)**计算**$\omega$：

- 对于恒常标定的情形：把方程$\omega = (H_\infty^i)^{-T}\omega(H_\infty^i)^{-1}$，$i = 1, \ldots, m$ 记为$A\mathbf{c} = \mathbf{0}$，其中 A 是一个$6m \times 6$的矩阵，而$\mathbf{c}$是二次曲线$\omega$的元素排列成的一个 6 维矢量，或
- 对于有参数变量的标定：用方程$\omega^i = (H_\infty^i)^{-T}\omega(H_\infty^i)^{-1}$把对$\omega^i$的元素的线性约束(例如零扭曲)表示成$\omega$元素的线性方程.

(4)通过 SVD 得到$A\mathbf{c} = \mathbf{0}$的一个最小二乘解.

(5)**度量矫正**：由 Cholesky 分解$\omega = (KK^T)^{-1}$确定摄像机矩阵 K. 从而得到一个度量重构：$\{P^i H_p H_A, (H_p H_A)^{-1}\mathbf{X}_j\}$，其中

$$H_A = \begin{bmatrix} K & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}.$$

(6) 用迭代的最小二乘最小化方法来改善所求得的解(见节 18.3.3)

算法 18.2. 用 IAC 约束的分层自标定算法.

### 18.5.3 应用无穷单应关系时的多义性

在本节中，我们介绍如果只有一个旋转轴时，(18.25)确定内参数的多义性. 我们假设内参数未知但恒定.

旋转矩阵 R 有一个对应于特征值 1 的特征矢量$\mathbf{d}_r$，$R\mathbf{d}_r = 1\mathbf{d}_r$，其中$\mathbf{d}_r$是该旋转轴的方向. 因此，矩阵$H_\infty^i = KR^iK^{-1}$也有一个对应于特征值 1 的特征矢量(假定$H_\infty^i$已归一化使得$\det H_\infty^i = 1$). 这个特征矢量是$\mathbf{v}_r = K\mathbf{d}_r$，并且图像点$\mathbf{v}_r$是旋转轴方向的消影点. 假设$\omega_{true}^*$是真正的$\omega^*$；从而可以验证如果$\omega_{true}^*$满足(18.25)且$H_\infty^i = KR^iK^{-1}$，那么(对偶)二次曲线的单参数簇也满足

$$\omega^*(\mu) = \omega_{true}^* + \mu \mathbf{v}_r \mathbf{v}_r^T \tag{18.27}$$

其中$\mu$是这个簇的参数. 同样(18.24)的 IAC 方程也存在着一个单参簇(束)的解. 这个论证表示尽管无穷单应约束好象为$\omega^*$的 5 个自由度提供了 6 个约束，但是这些约束中只有 4 个是线性独立的.

**消除多义性** 单参数的多义性可以用若干种方式解决. 第一，如果有绕与$\mathbf{d}_r$不同方向的轴

旋转的另一幅视图，则这两组约束的组合将不存在这种多义性. 用(18.26)的方法很容易得到一个线性解. 因此最低限度有 3 幅视图(即多于一个旋转)时一般可以得到唯一解. 解决这个多义性的第二种方法是对摄像机的内参数给予假设:例如零扭曲的假设(见表18.4). 强制零扭曲的方程作为硬约束可加到要求解的方程组中去.

另一种(但等价)的方法是在下面的方法中加入一个**后验**约束. 求解 $\mathbf{c}$ 的多义性来自线性方程组 $\mathbf{Ac}=\mathbf{0}$，它发生在当 A 有一个二维(或更大)的右零空间时. 对于这种情形,在求解 $\omega$ 时将得到一簇下面形式的解

$$\omega(\alpha)=\omega_1+\alpha\omega_2.$$

这里 $\omega_1$ 和 $\omega_2$ 由零空间的生成元得到，而 $\alpha$ 则必须要确定. 剩下的任务只是寻找 $\alpha$ 的值,以获得满足表 18.4 中所选择的约束条件的解. 这可以用线性方法求解. 也可以用求解 DIAC 途径求解，但是约束条件将是三次的(见表 18.2($p$329))，且其中一个解会是假的.

在某些情形，这些附加的约束不能解决多义性问题. 例如，如果绕图像的 $x$ 或 $y$ 轴旋转，则零扭曲不能解决它的多义性问题. 这样的一些例外在[Zisserman－98]中有详细的介绍，我们现在给出几个经常发生的例子.

**典型多义性** (18.27)给出 $\omega^*(\mu)$的单参数解簇对应由从 $\omega^*(\mu)$产生的单参数簇标定矩阵,因为 $\omega^*(\mu)=\mathrm{K}(\mu)\mathrm{K}(\mu)^\mathsf{T}$. 为简单起见,我们假设真正的摄像机 K(它是该簇中的元素)有零扭曲，因此 K 有 4 个未知参数.

如果旋转轴是平行于摄像机 X－轴，则 $\mathbf{d}_r=(1,0,0)^\mathsf{T}$ 且 $\mathbf{v}_r=\mathrm{K}\mathbf{d}_r=\alpha_x(1,0,0)^\mathsf{T}$. 根据 $\omega^*$ 零扭曲的形式(18.11－$p$328),簇(18.27)是

$$\omega^*(\mu)=\omega^*_{\text{true}}+\mu\mathbf{v}_r\mathbf{v}_r^\mathsf{T}=\begin{bmatrix}\alpha_x^2(1+\mu)+x_0^2 & x_0y_0 & x_0\\ x_0y_0 & \alpha_y^2+y_0^2 & y_0\\ x_0 & y_0 & 1\end{bmatrix} \tag{18.28}$$

注意因为整个簇为零扭曲，所以此时只有元素 $\omega^*_{11}$ 会变化. 这意味着主点和 $\alpha_y$ 的确定无多义性——因为它们可以从不受多义性影响的元素中读出. 然而，$\alpha_x$ 显然不能被确定，因为它只在变化的元素 $\omega^*_{11}(\mu)$中出现. 把它和两个其他的典型情形总结起来得:

- **在假设零扭曲的摄像机条件下,如果 K 由无穷单应关系(18.25)计算得到，则对某些运动,只剩下一个不确定的标定参数，对于绕各种轴的旋转,多义性的情况如下:**
  **(1) 绕 X－轴: $\alpha_x$ 不确定;**
  **(2) 绕 Y－轴: $\alpha_y$ 不确定;**
  **(3) 绕 Z－轴(主轴): $\alpha_x$ 和 $\alpha_y$ 不确定，但是它们的比率 $\alpha_y/\alpha_x$ 是确定的.**

**几何注释** 这些多义性不仅局限于一对视图标定的情形，而且适用于整个的图像序列. 例如，如果一个摄像机总是绕摄像机的 X－轴旋转，则将存在重构的多义性，同样的结论对绕 Y－轴的旋转也成立. 我们可以从几何的角度给予如下解释. 设度量重构所采用的图像序列是由绕 Y－轴旋转的摄像机拍摄. 我们可以定义一个坐标系使世界坐标系的 Y－轴与摄像机的 Y－轴方向平行. 现在,设我们来"挤压"整个重构(点和摄像机的位置)使得它们的 Y 坐标都乘某个因子 $k$. 根据成像几何不难看出它的影响是把任何图像点的 Y 坐标乘以相同因子 $k$，但是不影响 X－坐标. 然而，这种影响可以通过用逆因子 $k^{-1}$乘以图像坐标的尺度

因子 $\alpha_y$ 而撤消，使得图像坐标不被改变. 这证明了 $\alpha_y$ 的确定是多义性的，事实上，它是非约束的. 总结起来，在度量重构中有一个平行于旋转轴的单参数多义性并使得内参数中也有单参数的多义性. 这个论证表明多义性问题是运动本身决定的，与任何具体的自标定算法无关.

**与 Kruppa 方程的关系** 记两视图的 (18.24)为 $\omega^{*\prime} = H_\infty \omega^* H_\infty^T$ 并用矩阵 $[\mathbf{e}']_\times$ 左乘和右乘得到

$$[\mathbf{e}']_\times \omega^{*\prime} [\mathbf{e}']_\times = [\mathbf{e}']_\times H_\infty \omega^* H_\infty^T [\mathbf{e}']_\times = F\omega^* F^T$$

因为 $F = [\mathbf{e}']_\times H_\infty$. 这就是 Kruppa 方程(18.18－p445)，这表明该方程可直接由无穷远单应约束推出. 因为 $[\mathbf{e}']_\times$ 不可逆，所以反方向走不通，即不能从 Kruppa 方程推导出无穷远单应约束. 因此 Kruppa 方程是一个较弱的约束.

然而，它们的差别是在应用(18.24)时，人们需要知道无穷远平面(从而仿射结构)，因为它仅对无穷单应而不是对任何 H 成立. 但另一方面，Kruppa 方程不涉及景物仿射结构的任何知识. 不过，这个关系表示序列图像在无穷单应关系下的任何标定多义性也是 Kruppa 方程的多义性.

## 18.6 从旋转摄像机标定

本节，我们开始考虑特殊成像条件下的标定. 在这里考虑的情形是摄像机绕着它的中心旋转但不作平移. 我们将考虑内参数固定和某些参数已知且固定而其余的未知且变化的两种情况.

上述情形经常发生. 有关的例子包括：水平与上下转动并且变焦的监视摄像机；用来报道体育事件的摄像机，它的位置几乎不变但可自由地旋转和变焦；又例如手持摄像机，它常常是在单个视点上水平转动. 即使这样的旋转也不是准确地围绕着中心，但在实际上平移相对于景物点的距离，一般可以忽略，因而，固定中心是一个极佳的逼近.

如节 18.5.2 中所给出的，旋转摄像机的标定问题在数学上等同于分层重构中从仿射到度量的标定步骤. 仅仅由一个非平移的摄像机不可能得到一个仿射(或任何)重构，因为没有办法求解深度. 然而，我们可以计算图像之间的无穷单应，它可以完全满足摄像机标定的要求.

如先前(节 7.4($p$140))所证明的，具有相同中心的两个摄像机的图像由一个平面射影变换相联系. 事实上，如果 $\mathbf{x}'$ 和 $\mathbf{x}$ 是对应的图像点，则它们由 $\mathbf{x}' = H'\mathbf{x}$ 相联系，其中 $H' = K'R'(K)^{-1}$，且 $R'$ 是视图 $i$ 和参考视图之间的旋转，进一步，因为这个映射与成像于 $\mathbf{x}$ 的点的深度无关，它也适用于在无穷远的点，因此，如节 12.4($p$239)所证明，

$$H' = H'_\infty = K'R'(K)^{-1}$$

从而便得到了直接由图像来测量 $H_\infty$ 的一种方便的方法.

给定 $H_\infty$，由旋转摄像机获得的所有图像的标定矩阵 $K'$ 的解可以按节 18.5.2 的介绍得到. 这种方法既可以应用于固定的也可以用于变化的内参数并且总结在算法 18.3 中. 我们将通过若干例子来说明它.

目标

给定 $m \geq 2$ 幅视图(它们由绕其中心旋转的摄像机拍摄,摄像机的内参数或是固定或是变化的),计算每一个摄像机的参数. 假定旋转不总是绕着同一轴.

算法

(1)**图像之间的单应**:例如,用算法 3.6($p$79)计算每一视图 $i$ 与参考视图之间的单应 $H^i$ 使得 $\mathbf{x}^i = H^i\mathbf{x}$. 归一化这些矩阵使得 $\det H^i = 1$.

(2)**计算** $\omega$:

- 对于恒定标定的情形:把方程 $\omega^i = (H^i)^{-T}\omega(H^i)^{-1}$, $i = 1, \ldots, m$ 重写成 $A\mathbf{c} = \mathbf{0}$,其中 A 是一个 $6m \times 6$ 的矩阵,而 $\mathbf{c}$ 是二次曲线 $\omega$ 的元素重新排列成的一个 6 维矢量,或
- 对于变化参数的标定:用方程 $\omega^i = (H^i)^{-T}\omega(H^i)^{-1}$ 把在表 18.4 中的关于 $\omega^i$ 的元素的线性约束(例如长宽比为 1)表示成关于 $\omega$ 元素的线性方程.

(3)**计算** K:确定 $\omega$ 的 Cholesky 分解 $\omega = UU^T$,从而得到 $K = U^{-T}$.

(4)**迭代改进**:(可选)改进 K 的线性估计的方法是在 K 和 $R^i$ 上最小化

$$\sum_{i=2,m;j=1,n} d(\mathbf{x}_j^i, KR^iK^{-1}\mathbf{x}_j)^2$$

其中 $\mathbf{x}_j$, $\mathbf{x}_j^i$ 分别是在第一和第 $i$ 幅图像上测量得到的第 $j$ 个点的位置. 该最小化的初始估计由 K 和 $R^i = K^{-1}H^iK$ 得到.

算法 18.3 摄像机绕它的中心旋转时的标定.

**例 18.9 固定内参数的绕中心旋转**

图 18.3 中的图像是由 35mm 摄像机在普通的黑白胶卷上产生的负片. 摄像机被手持着,且没有特别小心地保持摄像机中心不动.

用平板扫描仪把这些负片的放大照片数字化. 如果负片和相纸不精确地平行,放大的过程会导致一个非零的 $s$ 值而且 $\alpha_x$ 和 $\alpha_y$ 的值不等. 获得的图像尺寸是 $776 \times 536$ 像素.

这里所用的约束是内参数恒常. 用算法 18.3 介绍的方法计算得到的摄像机矩阵是:

$$K_{\text{linear}} = \begin{bmatrix} 964.4 & -4.9 & 392.8 \\ & 966.4 & 282.0 \\ & & 1 \end{bmatrix} \quad K_{\text{iterative}} = \begin{bmatrix} 956.8 & -6.4 & 392.0 \\ & 959.3 & 281.4 \\ & & 1 \end{bmatrix}$$

线性和迭代估计之间几乎没有差别. 而且计算得到的长宽比(实质上为 1)和主点是非常合理的. △

**例 18.10 变内参数的绕中心旋转**

图 7.9($p$143)中的图像由近似地绕它的中心水平旋转的摄像机拍摄. 摄像机没有变焦,但是由于自动对焦,焦距和主点也许稍微有变化.

此例中,所用的约束是像素是正方形的:即扭曲是零并且长宽比是 1;但是焦距和主点是未知而且不是固定的. 那么根据表 18.4,由每一幅视图可得到关于 $\omega$ 的两个线性约束. 从零扭曲得到 $(H_\infty^{i-T}\omega H_\infty^{i-1})_{12} = 0$,和从长宽比为 1 得到 $(H_\infty^{i-T}\omega H_\infty^{i-1})_{11} = (H_\infty^{i-T}\omega H_\infty^{i-1})_{22}$. 这些约束组合起来给出了关于 $\omega$ 的一个如算法 18.3 所介绍的线性方程组.

计算得到的每幅视图的摄像机矩阵的内参数在图 18.4 中给出. 显然,恢复的焦距和主点相当恒定(即使不被约束强制),而且对手持序列而言水平和上下旋转的角度也是很合理的.

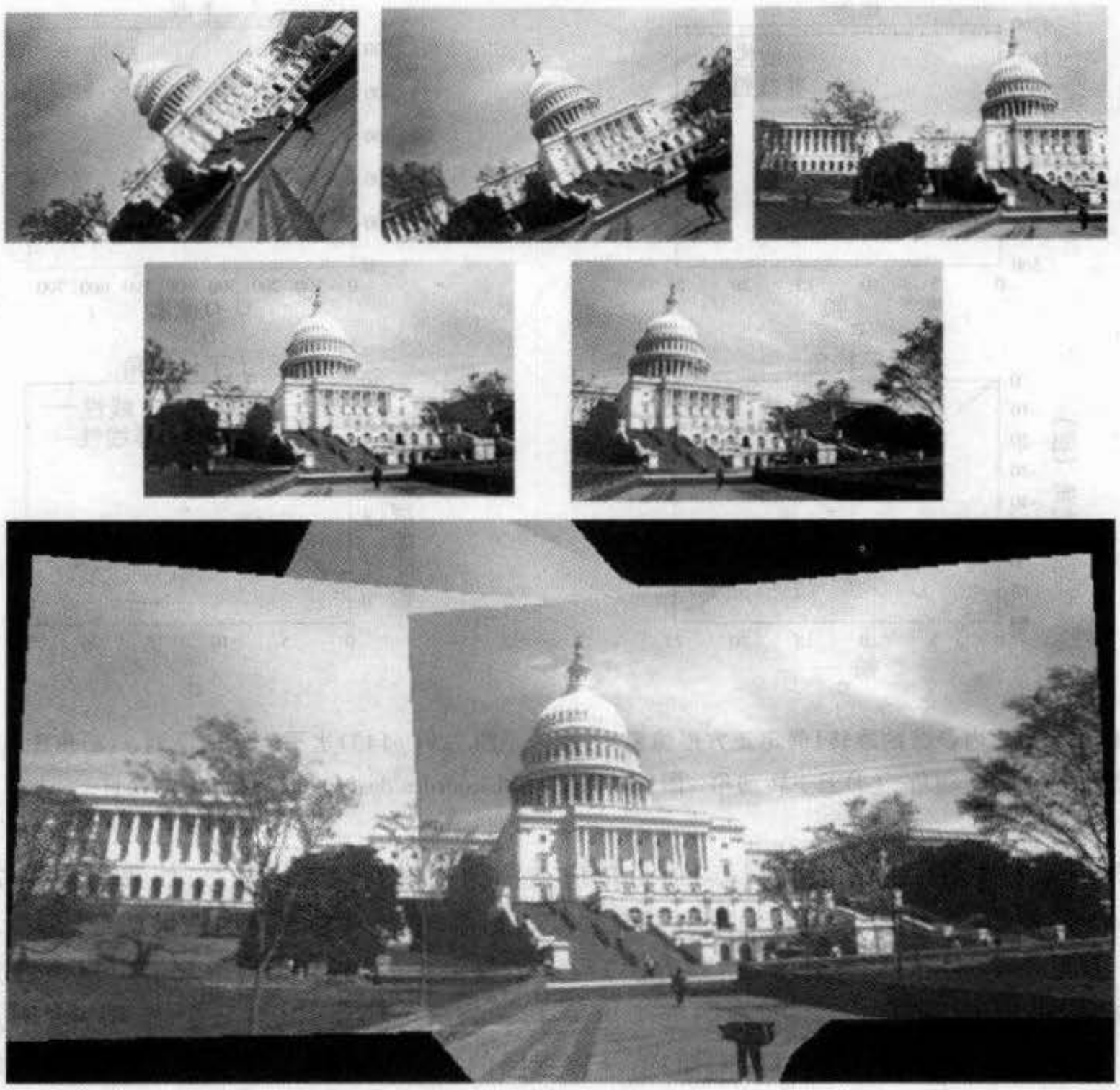


图 18.3 **标定一个绕它自己中心旋转的摄像机.** (上图)用近似绕其中心旋转的摄像机拍摄的 5 幅美国国会大厦的图像. (下图)由这 5 幅图像构成的拼图(见例 7.13($p$142)). 拼图非常清楚地显示了图像之间无穷单应的失真效果. 这种失真的分析提供自标定算法的基础. 标定的计算如算法 18.3 所介绍.

## 18.7 平面自标定

对于平面景物的一组图像,先估计一个射影重构,再计算使它变换到度量重构的矫正变换的两步算法不再有效. 这是因为没有深度起伏就不可能确定摄像机. 如节 16.5.2($p$302)所示,最少需要不在这个平面上的两个点. 然而,由景物平面来进行自标定是可能的. 这一点被[Triggs-98]证明,该文对恒定内参数的情形给出一个解. 从潜在应用的角度来看,这个方法是非常有趣的. 在人造的环境中包含平面的景物是非常普遍的,比如,地平面. 进一步,在由高空的飞机或卫星拍摄的区域图像中,与图像的广度相比场景的深度起伏是很小的,景物可以近似逼近于一个平面并且此自标定方法是适用的.

首先,算法求由景物平面诱导的一组图像到图像的单应. 这些图像与第一幅图像相联系并提供一组单应 $H^i$. 从几何上看,平面自标定是两种思想的联姻. 第一,平面上的虚圆点

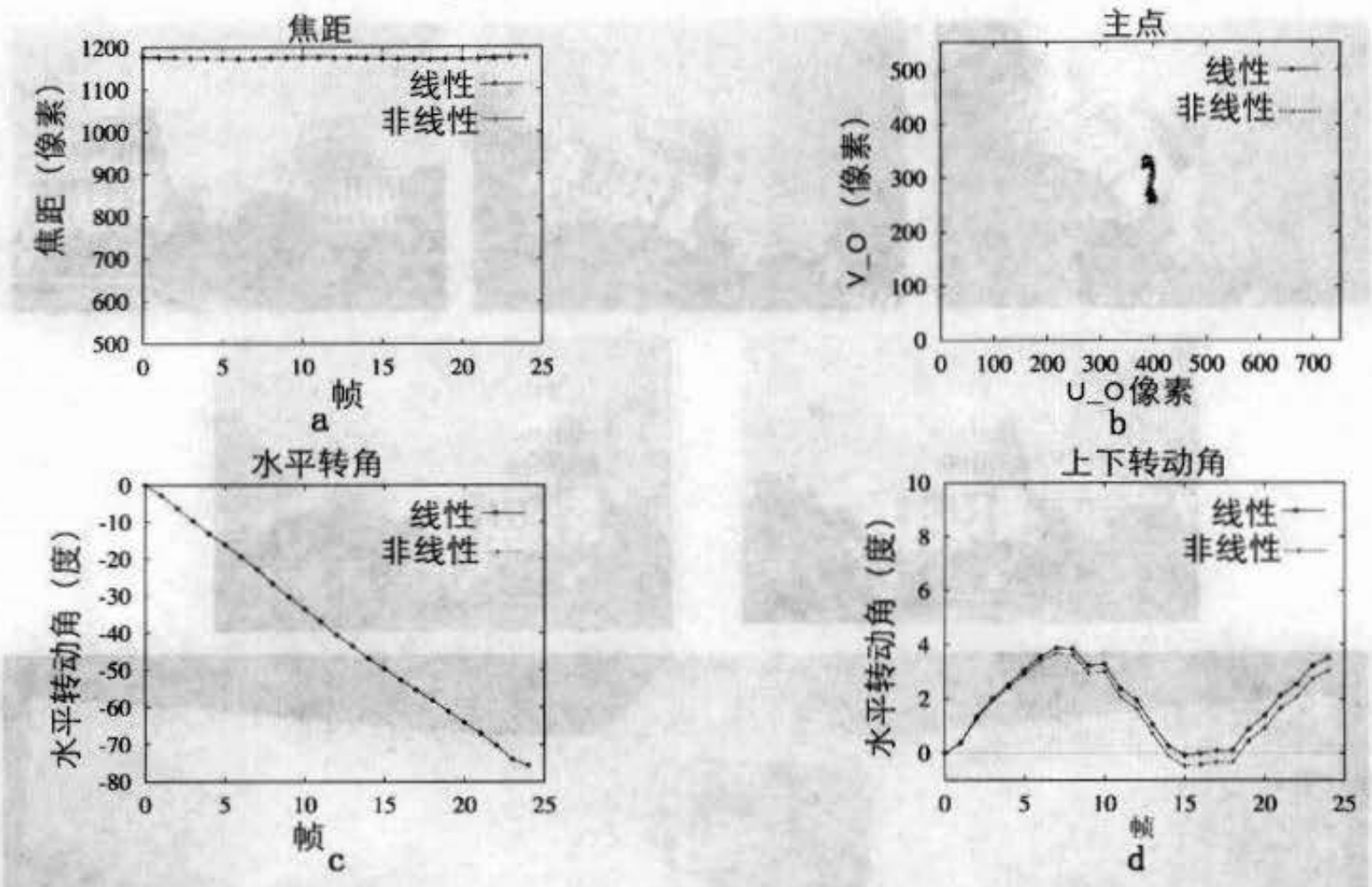



图 18.4 **变化内参数的旋转(假定正方形像素)** 这些是图 7.9($p$143)水平旋转的序列.(a)焦距.(b)主点.(c)水平转动角.(d)上下转动角.图像的采用得到 Lourdes de Agapito Vicente 的允许.

(它是绝对二次曲线和平面的交点)通过单应从图像映射到图像.第二,如我们在例 7.17($p$147)所见,标定矩阵 K 可以由平面的虚圆点的影像确定(每一幅图像提供两个约束).

因此假设在第一幅图像上(用某方法)确定了虚圆点的像(4dof),则它们可以由已知的 $H^i$ 转移到其他的视图上.每幅视图则给出关于 $\omega$ 的两个约束,因为 $\omega$ 上有两个虚圆点的像.详细地说,如果我们记(目前未知)第一幅视图的虚圆点的像为 $\mathbf{c}_j$, $j=1,2$,则自标定方程是

$$(H^i\mathbf{c}_j)^T\omega^i(H^i\mathbf{c}_j)=0,\ i=1,\ldots,m,\ j=1,2 \tag{18.29}$$

其中 $H^1=I$.在求解这些方程时,未知量是第一幅图像上的虚圆点的坐标和某些未标定参数.虽然虚圆点是复点,但是它们相互复共轭,因此 4 个参数就足够描述它们.此外如果所有 $m$ 幅视图的内参数 $K^i$ 中共有 $v$ 个未知量,那么只要 $2m\geq v+4$,就是可解的,因为每幅视图提供两个方程.

根据表 18.4,加在摄像机内参数的限制会导致进一步的代数约束,而且这些可用来补充由(18.29)所产生的约束.表 18.5 考虑了各种情形.对于大多数情形,平面标定是一个非线性问题,而且关于怎样寻求一个解的计算方法这里没有更多可说.最小化某个代价函数的迭代方法是必要的.关于最小化方法的细节见(Triggs-98).

**实现** 在实现上,这个方法的一个相当大的好处是它需要的只是平面之间的单应,而不是由 3 维空间点产生的点对应,后者一般需要估计多视图的张量,比如基本矩阵.平面间的匹配变换要简单、稳定和准确得多,它由点到点的图像间变换的约束本质所确定.算法 3.6($p$79)的方法可以用来估计两幅图像间的这个变换.另外,也可采用基于相关的算法,根据图像灰度直接地估计参数化的单应.

| 条件 | 自由度 | 视图 |
|---|---|---|
| 未知但恒常内参数 | 5 | 5 |
| 已知恒常扭曲和长宽比，未知但恒常主点和焦距 | 3 | 4 |
| 除变焦外，所有内参数已知 | $m$ | 4 |
| 焦距变化，其他内参数固定但未知 | $m+4$ | 8 |

表 18.5 **在不同条件下由平面标定所需的视图数**. 如果 $2m \geqslant v+4$，标定(原则上)是可能的。

## 18.8 平面运动

具有某种实用意义的情形是摄像机在一张平面上运动并且绕垂直于该平面的轴旋转. 例如一部在地面上行走并带有固定摄像机的车. 此时，摄像机必然在与(水平)地面平行的平面上移动，并且当车转动时，摄像机将绕一个垂直轴旋转. 这里**没有**假设摄像机的指向是水平，或是相对于车的任何其他特殊的方向. 然而，我们假设摄像机是恒定内参数标定. 这种运动约束在本质上使标定任务大大地简化.

可以证明给定平面运动序列的 3 幅或更多幅图像，可以计算一个仿射重构. 为此，我们需要确定无穷远平面. 它可通过辨认无穷远平面上的三个点来实现. 这些点就是图像序列中的不动点.

**平面运动下的不动像点** 根据节 2.4.1($p$44)，任何刚体运动(例如摄像机运动)可以解释成绕转动轴的一个旋转加上沿轴方向的一个平移. 在平面运动时，转动轴垂直于运动的平面，而且转动运动的平移部分是零. 想象一辆车水平地绕转动轴打转. 相对于摄像机，转动轴的位置不变，因此它将构成图像上由不动点组成的一条直线. 如果第二个运动绕不同轴打转，则第二根轴的像在第二对图像上是不动的. 这两根轴的像一般不同，但交于两根转轴的交点的像. 由于这两根轴是垂直的，它们必然平行，从而在无穷远平面上它们的共同的方向处相交. 这个方向在图像上的投影为转轴交点的像，它是转动轴方向的消影点. 这个像点称作顶点. 现在，我们得到一个跨视图之间的不动点，在射影重构中它将用作确定无穷远平面上的一个点.

如我们已经看到的，旋转轴的图像是一对图像中不动点组成的一条直线. 另外还存在一条不动直线，它可以从一对图像按如下方式辨认. 因为摄像机运动所在的平面(它将称为地平面)是固定的，平面上的这些点对所有图像都映射到同一条直线. 这条直线称为水平线并且是地平面的消影线. 由于每一个摄像机在地平面上，它们的对极点必然都在水平线上. 与转动轴的像不同，水平线是一条不动的直线，但不是一条不动点的直线.

虽然水平线不是点点不动的，但是它包括图像对中的两个不动点，即地平面上两个虚圆点的像. 这些虚圆点是绝对二次曲线和地平面的交点. 由于在刚体运动下绝对二次曲线的像是不动的，以及地平面的像是不动的，两个虚圆点的像必然不动. 事实上，它们在所有平面运动序列的图像中不变. 这在图 18.5 中说明.

至此，我们已描述了这种运动序列的不动点. 计算这些不动点等价于仿射重构，因为我们能通过这些不动像点的反向投影来寻找对应的无穷远平面上的 3D 点. 虽然其中的顶点可以从 2 幅视图计算得到，但是需要 3 幅视图来计算虚圆点的影像.

**计算不动点** 如果空间中的点映射到两幅图像的像点相同，则称这样的点的集合为同视点. 一般同视点是一条三次绕线，但在平面运动时，它退化到一条直线(转轴)和在地平面上的一条二次曲线. 同视点的像是由 $F+F^T$ 所定义的二次曲线，其中 F 是基本矩阵(见节 8.4 ($p$168)). 对于平面运动的情形，它们是由两条直线组成的退化二次曲线，即旋转轴和地平线的像(见图 8.11($p$170)). 通过该二次曲线的分解可以确定这两条直线. 根据 3 幅视图，我们可以计算由它们组成的三对视图的任何一对视图的同视点的像，因此得到 3 组地平线和旋转轴的像. 水平线是它们的公共成分，而其他成分(旋转轴的像)在图像上相交于顶点.

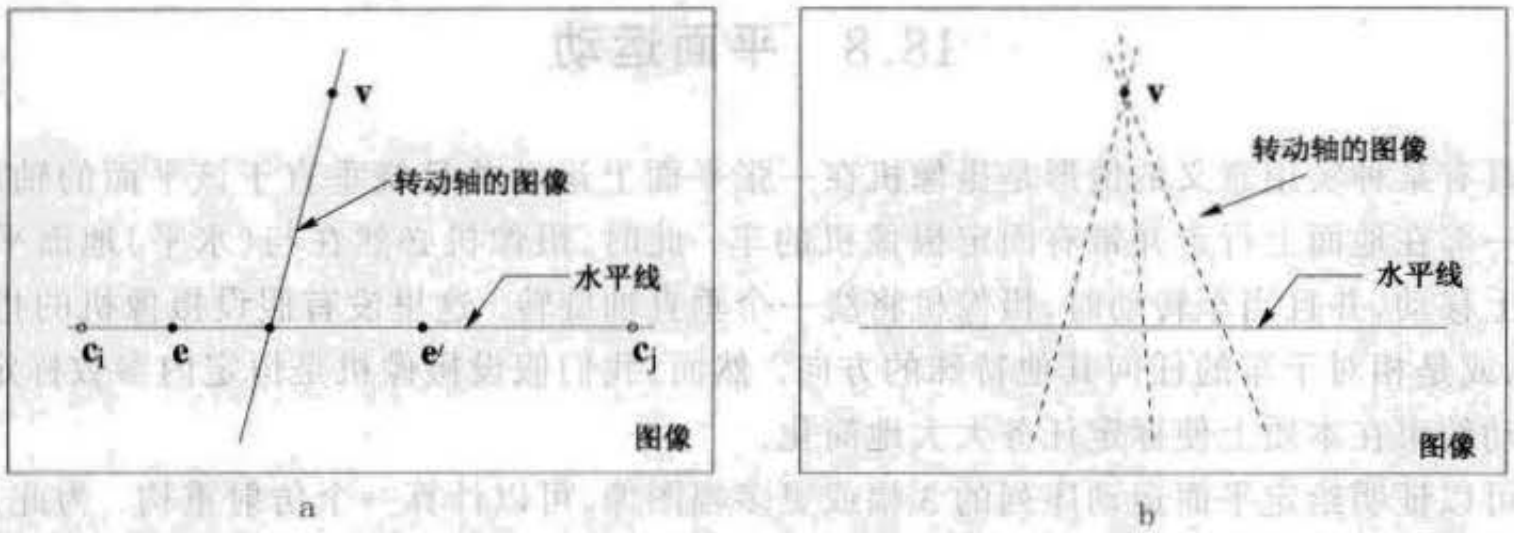



图 18.5 **平面运动中不动图像实体** (a)在两视图中，转动轴的图像是运动图像中不动点的直线. 地平线是运动图像的一条不动直线. 对极点 e,e′和地平面的虚圆点的影像 $c_i$,$c_j$ 都在地平线上. (b)由平面运动的 3 幅图像配对得到的不动直线之间的关系. 每对图像中水平面的影像重合，而每对图像中转动轴的图像交于顶点 v. 所有对极点在地平线上.

现在我们回到计算虚圆点. 理解对应于两组图像对的一组同视点的几何将有助于计算. 令 $C^{12}$ 为描述在视图 1 和 2 中地平面上的同视点的一条二次曲线. 这条二次曲线通过两个虚圆点、两个摄像机的中心和转动轴与地平面的交点. 由于二次曲线包含虚圆点，它是一个圆. 令 $C^{23}$ 为由视图 2 和 3 定义的对应的圆. 这两个虚圆点和第二个摄像机中心将同时在这两个圆上. 由于一般两条二次曲线相交于四个点，一定还有另外一个(实的)交点. 然而，它可以被舍去，因为感兴趣是两个复交点，即在地平面上的虚圆点.

在实现时，作者们采用了不同的方法来寻找这两个虚圆点. [Armstrong－96]的方法利用三焦点张量寻找在三幅视图中通过顶点的不动直线. [Faugeras－98,Sturm－97b]的方法涉及计算一个把 2D 地平面映射到 1D 图像的 1D 摄像机的三焦点张量. 在上述两种情形，虚圆点在地平线上的位置都由解一个单变量的三次方程得到.

这种仿射标定方法的主要步骤总结在算法 18.4 中.

**度量重构** 一旦找到顶点和两个虚圆点，就可以算出无穷平面和图像间的无穷单应. 标定和度量重构现在就可以按一般的方式进行. 然而必须注意：运动的约束本质表明存在关于标定的一个单参数的解簇，因为所有的摄像机绕同一轴旋转. 我们已在节 18.5.3 把标定的多义性进行了分类. 为了寻求唯一的结果，有必要给内参数一个假设. 但我们已经知道在零扭曲的约束下，如果摄像机的 $y$ 轴与旋转轴平行(这在实际情况中可能是真的)则约束是不充分的. 最好的方案是强制零扭曲并已知长宽比约束(例如，如果像素是正方形的).

目标

给定 3(或更多)幅图像,它们由在平面运动下的有恒常内参数的摄像机得到,计算一个仿射重构

算法

(1) **计算一个射影重构**. 它由 3 视图的三焦点张量 $\mathcal{T}$ 算得. 三焦点张量可以用诸如算法 15.4($p$286)来计算.

(2) **由 $\mathcal{T}$ 计算每对图像的基本矩阵**. 见算法 14.1($p$267). 把每个基本矩阵的对称部分分解成两条直线,一条水平线和转动轴的图像.

(3) **计算顶点**. 由转动轴的三幅影像确定顶点 $\mathbf{v}$.

(4) **计算三元组的地平线**. 由三个基本矩阵得到 6 个对极点,通过这些对极点的正交回归拟合来确定地平线.

(5) **计算虚圆点的影像**. 计算虚圆点在地平线上的位置 $\mathbf{c}_i$, $\mathbf{c}_j$(见正文).

(6) **计算无穷远平面**. 由对应图像点 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 计算无穷远平面上的三角形的顶点,这里对每个 $\mathbf{v}$, $\mathbf{c}_i$, $\mathbf{c}_j$ 有 $\mathbf{x}=\mathbf{x}'$. 它确定了 $\boldsymbol{\pi}_\infty$ 上的三个点,因而确定该平面.

(7) **计算仿射重构**. 按算法 18.2($p$340)利用计算得到的 $\boldsymbol{\pi}_\infty$ 矫正射影重构.

算法 18.4 平面运动下的仿射标定.

**例 18.11 平面运动下的度量重构**

图 18.6(a)显示了一个平面运动序列的 7 幅图像中的 4 幅. 拍摄这个序列时摄像机的仰角近似为 20°. 用所有可能的视图对计算得到的转动轴的影像和地平线在图 18.6(b)中显示,图中还给出估计得到的顶点与地平线. 虚圆点的像的位置的估计值为 $x=104\pm362i$, $y=-86\mp2i$. 假定长宽比为 1:1,计算得到标定矩阵 K 的内参数的结果是 $\alpha_x=330$, $\alpha_y=363$, $x_0=123$, $y_0=50$. 度量重构的准确性由测量度量不变量来判断. 典型的结果显示在图 18.6(c)中. △

## 18.9 双眼装置的自标定

本节介绍一种用于标定两个摄像机"固定"的双眼装置的分层方法. 这里"固定"意味着装置中的摄像机的相对方向在运动中不变,而且每只摄像机的内参数也不变. 我们将证明由这个装置的单个运动就能唯一地确定无穷远平面.

假设一个固定的双眼装置作了一个一般的运动. 运动前后的景物的射影结构分别表示 $\mathbf{X}$ 和 $\mathbf{X}'$. 因为 $\mathbf{X}$ 和 $\mathbf{X}'$ 是同一景物的两个射影重构,它们由一个 $4\times4$ 射影变换 $\mathrm{H}_\mathrm{P}$ 相关联,即

$$\mathbf{X}'=\mathrm{H}_\mathrm{P}\mathbf{X}$$

然而,因为该装置的实际运动是欧式的,因此(见下文)这个单应 $\mathrm{H}_\mathrm{P}$ 共轭于表示该运动的欧氏变换. 共轭性是一个关键的结论,因为在共轭关系下不动元素映射到不动元素. 因此欧式运动的不动元素(特别地无穷远平面)能从由 $\mathrm{H}_\mathrm{P}$ 表示的射影运动的不动元素得到.

**共轭关系** 假设 $\mathbf{X}_\mathrm{E}$ 表示以该装置上的欧式坐标系表示的 3 维空间中的一点,而 $\mathbf{X}_\mathrm{E}'$ 表示装置运动后的同一点. 则这些点的联系为

$$\mathbf{X}_\mathrm{E}'=\mathrm{H}_\mathrm{E}\mathbf{X}_\mathrm{E} \tag{18.30}$$

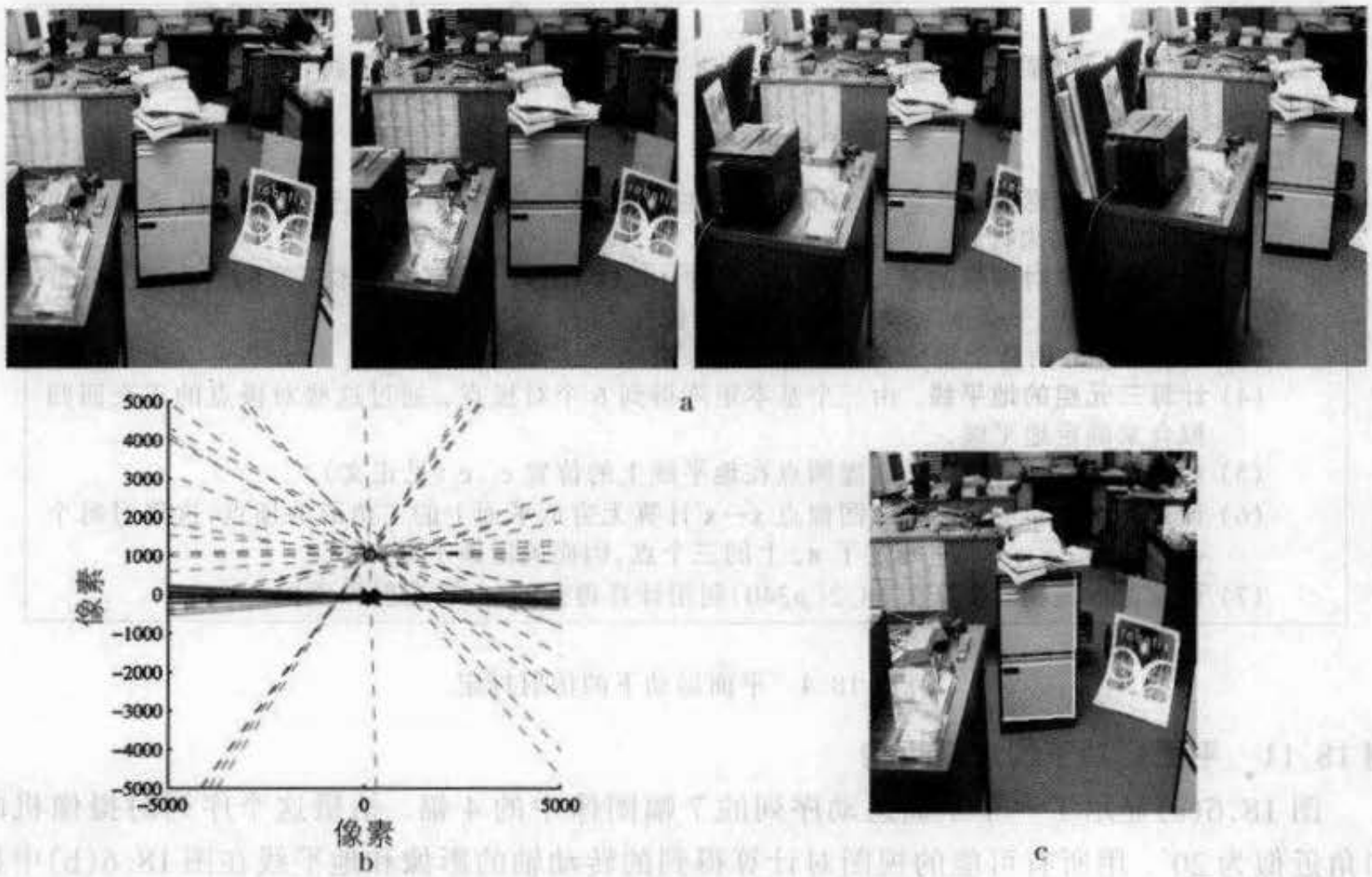



图 18.6 **平面运动序列** (a)(采用的 7 幅中的)4 幅图像,由安置在正在作平面运动的车上的摄像机拍摄. (b)由所有图像对计算得到的对极点(×),地平线(灰实线)以及转动轴的影像(灰虚线). (c)在度量重构中测量欧氏不变量:直角的测量值为 89°,非平行长度比率测量值为 0.61(真实值大约是 0.65).

其中 $H_E$ 是一个非奇异 4×4 欧式变换矩阵,它表示该装置的旋转和平移. 假设该点也用装置上的射影坐标系表示(它由射影重构得到);那么

$$\mathbf{X}_E = H_{EP}\mathbf{X} \quad \mathbf{X}_E' = H_{EP}\mathbf{X}', \tag{18.31}$$

其中 $H_{EP}$ 是一个非奇异 4×4 矩阵,由它把射影与度量结构联系起来. 这里要注意的基本要点是运动之前和之后的两个射影重构中的摄像机必须在同一个射影坐标系中,也就是说,运动之前和之后所用的摄像机矩阵对相同.

从(18.30)和(18.31)得出

$$H_P = H_{EP}^{-1} H_E H_{EP}. \tag{18.32}$$

因此 $H_P$ 共轭于一个欧式变换. 这种共轭关系有两个重要的性质:

(1)$H_P$ 和 $H_E$ 有相同的特征值.

(2)如果 $\mathbf{E}$ 是 $H_E$ 一个特征矢量,则 $H_P$ 与之对应的相同特征值的特征矢量是($H_{EP}^{-1}\mathbf{E}$);即 $H_E$ 的特征矢量被点变换(18.31)映射到 $H_P$ 的特征矢量. 它由(18.32)得到,因为如果 $H_E\mathbf{E} = \lambda\mathbf{E}$,则 $H_{EP}H_P H_{EP}^{-1}\mathbf{E} = \lambda\mathbf{E}$,而在左乘 $H_{EP}^{-1}$ 后就能得到需要的结果.

**欧式变换的不动点** 考虑由矩阵

$$H_E = \begin{bmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

表示的欧式变换，它是绕 Z 轴旋转 $\theta$ 并沿着 Z 轴平移一个单位（它是一个一般的转动）. 在此变换下，$H_E$ 的特征矢量是不动点（参考节 1.9（$p30$））. 对此情形，特征值是 $\{e^{i\theta}, e^{-i\theta}, 1, 1\}$ 而 $H_E$ 对应的特征矢量是

$$\mathbf{E}_1 = \begin{pmatrix} 1 \\ i \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{E}_2 = \begin{pmatrix} 1 \\ -i \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{E}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{E}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

所有特征矢量都在 $\boldsymbol{\pi}_\infty$ 上. 这意味着 $\boldsymbol{\pi}_\infty$ 作为一个集合是不动的，但**不是**不动点的平面. 特征矢量 $\mathbf{E}_1$ 和 $\mathbf{E}_2$ 是垂直于 Z（旋转）轴的平面上的虚圆点. 另外两个（相同的）特征矢量 $\mathbf{E}_3$ 和 $\mathbf{E}_4$ 是旋转轴的方向.

**计算 $\boldsymbol{\pi}_\infty$** 如果点变换矩阵是 $H_E$，则根据（2.6 - $p36$），平面变换矩阵是 $H_E^{-T}$. $H_E^{-T}$ 的特征矢量是该运动下的不动平面. 矩阵 $H_E^{-T}$ 也有两个相等的单位特征值和一个对应于它们的特征矢量，可以容易地验证它就是平面 $\boldsymbol{\pi}_\infty$. 如上面描述的 $H_E$ 的特征矢量映射到 $H_P$ 的特征矢量一样，$H_E^{-T}$ 的特征矢量映射到 $H_P^{-T}$ 的特征矢量. 因此，在射影重构中 $\boldsymbol{\pi}_\infty$ 是对应于 $H_P^{-T}$（双重）实特征值的特征矢量. 因此，

- **$\boldsymbol{\pi}_\infty$ 可以作为 $H_P^{-T}$ 的实特征矢量或等价地且更简单地，作为 $H_P^T$ 的实特征矢量唯一地被计算出来.**

这里我们观察到虽然实特征值的代数重数为 2，它的几何重数（在非平面运动情形）是 1. 这就是我们寻找无穷远平面的原因. 有关的仿射标定的过程总结在算法 18.5 中.

目标

给定两（或多）对双眼图像，它们由一个固定的双眼装置在一般运动（即 R 和 **t** 都非零，并且 **t** 不垂直于 R 的轴）下获取，计算一个仿射变换.

算法

(1) **计算一个初始射影重构 X**：按第 9 章的介绍，用第一对双眼计算一个射影重构 $(P^L, P^R, \{\mathbf{X}_j\})$. 它涉及到计算基本矩阵以及第一对图像之间的点对应 $x_j^L \leftrightarrow x_j^R$，这可以用诸如算法 10.4（$p200$）来实现.

(2) **计算运动后的射影重构 X′**：计算第二对双眼图像之间的对应 $x_j'^L \leftrightarrow x_j'^R$. 由于摄像机的内参数和相对的外参数是固定的，第二对双眼与第一对双眼的基本矩阵 F 相同. 利用相同的矩阵 $P^L, P^R$，由在第二对双眼计算得到的对应 $x_j'^L \leftrightarrow x_j'^R$ 来计算 3 维空间的三角形顶点 $\mathbf{X}_j'$.

(3) **计算把 X 关联到 X′ 的 4×4 矩阵 $H_P$**：计算两个双眼对中左边图像之间的对应 $x_j^L \leftrightarrow x_j'^L$（例如再用算法 10.4（$p200$））. 它建立空间点 $\mathbf{X}_j$ 和 $\mathbf{X}_j'$ 之间的对应. 单应 $H_P$ 可以由 5 组或更多组这样的 3 维空间的点对应线性地估计出来，然后该估计可以用在 $H_P$ 上最小化一个适当的代价函数来改进. 例如最小化 $\sum_j (d(x_j^L, P^L H\mathbf{X}_j')^2 + d(x_j^R, P^R H\mathbf{X}_j')^2)$ 使测量与重投影的图像点之间的距离最小化.

(4) **仿射重构**：由 $H_P^T$ 的实特征矢量计算 $\boldsymbol{\pi}_\infty$，并由此得到一个仿射重构.

算法 18.5 一个固定双眼装置的仿射标定.

**度量标定和多义性** 一旦 $\pi_\infty$ 被辨认，度量重构可以按分层算法 18.2（$p$340）进行．由于该装置是固定的，左右两边摄像机的参数在运动中不变．由单个运动，每个摄像机的内参数在相差一个单参数簇的意义下确定，如节 18.5.3 所介绍，该单参数簇产生于单个旋转．

通常，单个运动所产生的多义性可由另外的运动或通过提供附加约束（例如像素是正方形）来消除．如果还有另外的运动，则可以用来改善对 $\pi_\infty$ 的估计．该度量重构的结果给出该装置的完全标定（即摄像机之间的相对朝向以及它们的内参数）．

**平面运动** 当平移与旋转轴的方向正交这样一种特殊的正交（平面）运动时，对应于重（实）特征值的特征矢量空间是二维的，因此，$\pi_\infty$ 在相差一个单参数簇的意义下给予确定．因此我们不能唯一地找到无穷远平面（这在例 2.8（$p$46）中详细地研究过）．这种多义性可以通过绕不同于第一个方向轴的第二个正交运动来解决．

**例 18.12 两对双眼的自标定．**

图 18.7(a)(b)显示了两对双眼，它们被用于双眼装置的仿射标定而具体的过程按算法 18.5 进行．标定的准确性是通过计算右边图像的消影点的两种方式进行评价：第一种用平行直线影像的交点；第二种是确定在左边图像上对应的消影点（由同一平行线的图像），并用由 $H_P^T$ 的特征矢量算出的无穷单应把个消影点映射到右边的图像．把消影点之间的差异作为 $H_\infty$ 计算准确性的一种测量．度量标定用了零扭曲约束来解决单参数的多义性．度量重构的角度的精度在 1°以内． △

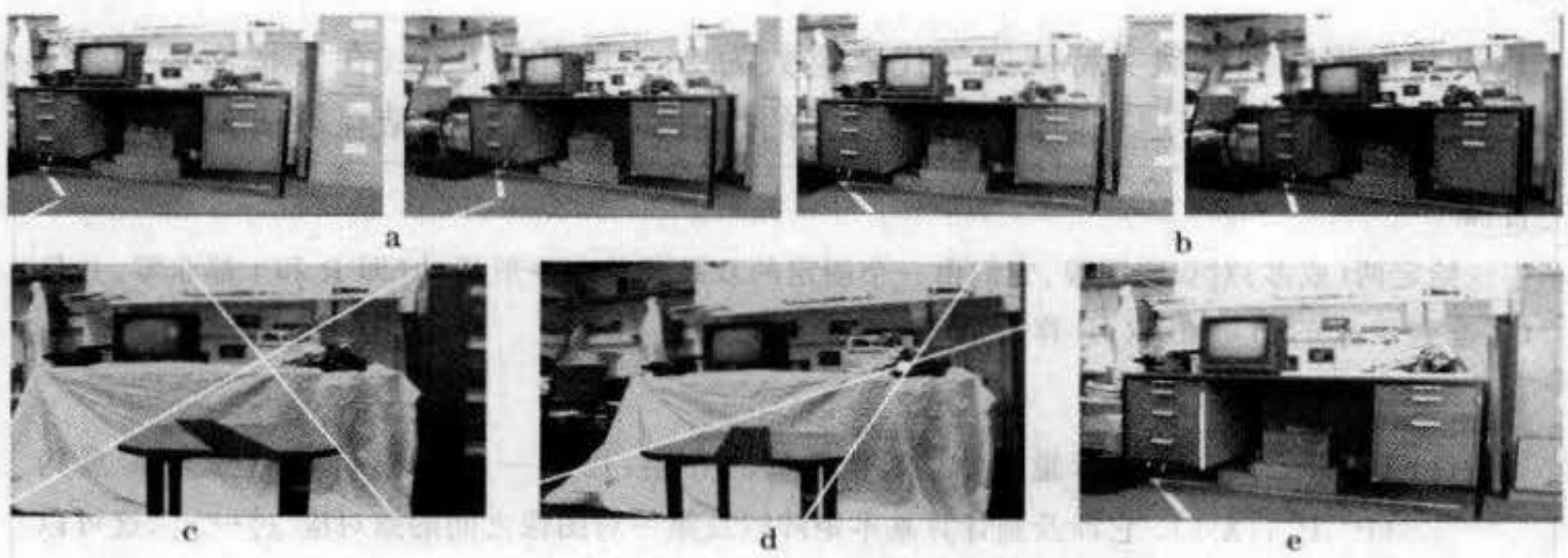



图 18.7 **双眼装置的自标定**．双眼对的输入：该装置运动之前（a）；之后（b）．双眼装置向左移动大约 20cm，水平旋转大约 10°，而仰角变化大约 10°．计算得到的 $H_\infty$ 的准确性判断在由同一个装置获取的另一对双眼中以如下方式进行：在（c）中，（双眼的）左图上通过桌边（它们在景物中是平行的）的图像交点计算一个消影点．在（d）中，（双眼的）右图上计算对应的消影点．白方块（在直线交点的附近）是用计算得到的 $H_\infty$ 把左图映射到右图得到的消影点．在没有误差时这两点应该是一个点．（e）按照度量标定，由 3D 重构得到的书桌边之间的角度（用白色显示）是 90.7°，与垂直值非常一致．

## 18.10 结束语

自标定领域的研究仍然非常活跃，而且比本章介绍的方法更好的方法也可能仍在开发．目前仍然缺乏由多视图张量求解的一种封闭形式解以及自动检测临界运动序列（见下面）的

算法.

**临界运动序列** 在本章中,我们已经了解到根据某些类型的运动不可能完全地确定矫正单应H. 所产生的重构处于度量和射影之间的某个层次上. 例如,在平面运动加恒常内参数的情形,存在一个平行于旋转轴的单尺度参数的多义性;而在恒常内参数和纯平移时,重构是仿射的. 出现多义性的摄像机运动序列称为"临界运动序列",Sturm[Sturm - 97a, Sturm - 97b]针对恒常内参数的情形进行了系统分类. 这种分类已经推广到更一般的标定约束,比如变焦距[Pollefeys - 99, Sturm - 99b]. 关于最近的工作见[Kahl - 99, Ma - 99].

**建议** 看上去自标定似乎为度量重构提供了一个完整的解答. 摄像机的标定是不必要的且我们能用诸如摄像机零扭曲这样的弱约束来实现. 不幸的是对这种完全信赖于自标定的做法还必须谨慎. 自标定在正确的环境下能正常工作,但是如果使用轻率的话就会失败. 下面是几条具体的建议.

(1) 小心地避免多义性运动序列. 我们已经了解如果运动过度受限,例如,仅绕单个轴旋转,将发生标定退化. 运动不应该太小,或者覆盖太小的视场. 自标定常常意味着估计无穷单应,在小视场时,它的效应不明显.

(2)尽可能多地利用所能掌握的信息. 虽然依据最少的信息(例如零扭曲)可以标定,但如果有其他信息可利用时,应该尽量利用. 例如,如果已知长宽比的约束就应该采用,已知主点的知识也应该用. 即使已知的这些值不太准确,但这些信息可以用一个方程加入到线性自标定算法中,并给予它低的权重.

(3)这条建议是关于捆集调整的. 一般最好在结束之前采用捆集调整. 在操作时,建议不要让摄像机的内参数在无界区域内游动. 例如,通常总能知道主点在某个合理的区域内(例如它不接近于无穷远)即使知道得不准确. 类似地,长宽比一般在 0.5 和 3 之间. 这些知识应该通过对代价函数的进一步约束而结合到捆集调整中,必要时可给予小的权重(标准偏差). 当自标定是病态(因此不稳定)时,这样做能使结果有显著的改善,它阻止代价函数为了一个小的并且不太显著的改善而使解在参数空间的一个大区域内徘徊.

(4) 采用限制运动的方法通常比允许一般运动的方法更可靠. 例如用于光旋转不平移的摄像机的方法通常比一般运动的更加可靠. 仅依据一个平移的仿射重构也是如此.

### 18.10.1 文献

摄像机自标定的思想起源于 Faugeras 等[Faugeras - 92b],其中使用了 Kruppa 方程. 早期的文章考虑了恒常内参数的情况. [Hartley - 94a]和 Mohr 等[Mohr - 93]针对多于 2 幅视图的情况研究了类似捆集调整的方法.

纯平移的仿射重构解法由 Moons 等给出[Moons - 94],并由 Armstrong 等扩展到纯平移紧跟着一个旋转的组合并用于完全的度量重构[Armstrong - 94]. 绕中心旋转的摄像机的自标定情况由[Hartley - 94d]给出. 模约束首先由 Pollefeys 等发表[Pollefeys - 96].

最早的双眼装置的自标定方法由 Zisserman 等给出[Zisserman - 95a],采用另一种参数化的方法由 Devernay 和 Faugeras[Devernay - 96]以及 Horaud 和 Csurka[Horaud - 98]给出. 双眼装置平面运动的特殊情形包含在[Beardsley - 95a, Csurka - 98]中. 最早的单个摄像机平面运动方法由 Armstrong 等发表[Armstrong - 96b],另一种数值解法是由 Faugeras 等给

出[Faugeras－98].

最近的论文中有研究比恒常内参数更少的约束限制. 许多"存在证明"已经给出:Heyden 和 Åström[Heyden－97a]证明了在仅知道扭曲和长宽比时度量重构是可能的,而[Pollefeys－98,Heyden－98]证明了有零扭曲的条件就是充分的.

Triggs[Triggs－97]引入了绝对(对偶)二次曲面作为数值工具使自标定问题公式化,并且用线性方法和序列二次规划解 $Q_\infty^*$. Pollefeys 等[Pollefeys－98]证明了基于 $Q_\infty^*$ 的计算可以用在一般运动下计算变焦距的真实图像序列的度量重构.

对于旋转摄像机的情形,de Agapito 等[deAgapito－98]对变内参数给出了基于 DIAC 的非线性求解方法. [deAgapito－99]又将此法修改为基于 IAC 的线性方法.

### 18.10.2 注释和练习

(1)[Hartley－92a] 首先给出从基本矩阵求焦距的一种解法,但所给出的算法不实用. 一个简捷的公式在[Bougnoux－98]中给出:

$$\alpha^2 = \frac{\mathbf{p}'^{\mathsf{T}}[\mathbf{e}']_\times \tilde{\mathrm{I}}\mathrm{F}\mathbf{p}\mathbf{p}^{\mathsf{T}}\mathrm{F}^{\mathsf{T}}\mathbf{p}'}{\mathbf{p}'^{\mathsf{T}}[\mathbf{e}']_\times \tilde{\mathrm{I}}\mathrm{F}\,\tilde{\mathrm{I}}\mathrm{F}^{\mathsf{T}}\mathbf{p}'} \tag{18.33}$$

其中 $\tilde{\mathrm{I}}$ 是矩阵 diag(1,1,0),而 $\mathbf{p}$ 和 $\mathbf{p}'$ 是在这两幅图像上的主点. 假设长宽比为 1 和扭曲为零. $\alpha'^2$ 的公式通过对换这两个图像的角色(和转置 F)给出. 注意这个算法的最后一步是取平方根. **在给定可靠的数据和主点的合理猜想下,**应该保证 $\alpha^2$ 和 $\alpha'^2$ 在用(18.33)计算时是正的. 然而在实际中它并不总是保持,而且负值可能产生. 这与节 18.3.5中提到的问题完全一样. 另外,如[Newsam－96]所提示,当两个摄像机的主射线在空间中相交时,这个方法本质上是退化的,此时不可能独立地计算焦距. 这种情形在两个摄像机瞄准同一点时发生,因而会常发生.

当基线和一个摄像机的主轴定义的平面与基线和另一个摄像机的主轴定义的平面相垂直时,退化将进一步发生. 一般地说,我们认为用这种方法计算焦距所得到的值是有疑问的.

(2)证明如果是恒常内参数,则从两幅视图得到的关于 $Q_\infty^*$ 的约束(18.6－$p$326)等价于 Kruppa 方程(18.18－$p$333). 提示,根据(8.10－$p$173)摄像机可以选择为 $\mathrm{P}^1=[\mathrm{I}|\mathbf{0}]$,$\mathrm{P}^2=[[\mathbf{e}']_\times\mathrm{F}|\mathbf{e}']$.

(3)由(18.21－$p$335)证明一个具有恒常内参数的摄像机作平移而没有旋转的运动时,无穷远平面可以直接从射影重构中计算出来.

(4)由节 12.4($p$239)中定义的:$\mathrm{H}_\infty^{ij}=\mathrm{K}^i\mathrm{R}^{ij}(\mathrm{K}^j)^{-1}$ 出发,只用两行就可推导无穷单应关系(18.24－$p$337). 此式可以重新安排成 $\mathrm{H}_\infty^{ij}\mathrm{K}^j=\mathrm{K}^i\mathrm{R}^{ij}$,利用正交性:$\mathrm{R}^{ij}\mathrm{R}^{ij\mathsf{T}}=\mathrm{I}$ 消去旋转给出 $\mathrm{H}_\infty^{ij}(\mathrm{K}^j\mathrm{K}^{j\mathsf{T}})\mathrm{H}_\infty^{ij\mathsf{T}}\doteq\mathrm{K}^i\mathrm{K}^{i\mathsf{T}}$.

(5)在 $\mathrm{H}_\mathrm{E}$ 下,$\boldsymbol{\pi}_\infty$ 上的点(即 $\mathrm{X}_4=0$)由一个 3×3 单应 $\mathbf{x}_\infty\mapsto\mathrm{R}\mathbf{x}_\infty$ 映射到 $\boldsymbol{\pi}_\infty$ 上的点. 在这个点变换下,根据结论 1.13($p$11)$\boldsymbol{\pi}_\infty$ 上的二次曲线映射为 $\mathrm{C}\mapsto\mathrm{R}^{-\mathsf{T}}\mathrm{C}\mathrm{R}^{-1}=\mathrm{R}\mathrm{C}\mathrm{R}^{\mathsf{T}}$. 由于 $\mathrm{R}\mathrm{I}\mathrm{R}^{\mathsf{T}}=\mathrm{I}$,绝对二次曲线 $\Omega_\infty$ 是不动的. 现在,用 $\mathbf{a}$ 定义旋转轴的方向,使得 $\mathrm{R}\mathbf{a}=\mathrm{I}\mathbf{a}$. 这条(退化的)点二次曲线 $\mathbf{a}\mathbf{a}^{\mathsf{T}}$ 也是不动的. 由此推出在这个映射下存在着不动的二次曲线束 $\mathrm{C}_\infty(\mu)=\mathrm{I}+\mu\mathbf{a}\mathbf{a}^{\mathsf{T}}$,因为:

$$R(\Omega_\infty + \mu \mathbf{a}\mathbf{a}^T)R^T = RIR^T + \mu R\mathbf{a}\mathbf{a}^T R^T = \Omega_\infty + \mu \mathbf{a}\mathbf{a}^T$$

标量 $\mu$ 是该束的参数. 这表明在一种特殊的相似变换下 $\pi_\infty$ 上存在一条不动的二次曲线的单参数簇. 尽管如此,只有 $\Omega_\infty$ 是唯一的在任何相似变换下都不动的二次曲线.

(6)标定多义性还存在于一种常见类型的机器人头,即水平旋转(或上下转动)的头. [deAgapito-98]证明了因为摄像机可以绕它的 X 或 Y 轴旋转,它的方向集合只形成一个 2 参数簇,而不是一般旋转的 3 参数簇. 这种局限导致了摄像机的长宽比率 $\alpha_x$ 的多义性,从而也导致 $x_0$ 的多义性.

# 第 19 章　对偶

由于 Carlsson[Carlsson－95]和 Weinshall 等[Weinshall－95]的工作，我们知道存在一个对偶化原则：允许把被若干摄像机所拍摄的点与摄像机的中心作角色交换．原则上这蕴含了对偶化射影重构算法以获得新算法的可能性。本章将发展这个主题以便勾画对偶化任何射影重构算法的一种显式方法．然而，在实际实现的层次上仍存在着需要克服的困难，才能使这种对偶方法产生可执行的算法．

## 19.1　Carlsson－Weinshall 对偶

令 $\mathbf{E}_1=(1,0,0,0)^{\mathrm{T}}$，$\mathbf{E}_2=(0,1,0,0)^{\mathrm{T}}$，$\mathbf{E}_3=(0,0,1,0,)^{\mathrm{T}}$ 和 $\mathbf{E}_4=(0,0,0,1)^{\mathrm{T}}$ 为 $\mathbb{P}^3$中一个射影基的组成部分．令 $\mathbf{e}_1=(1,0,0)^{\mathrm{T}}$，$\mathbf{e}_2=(0,1,0)^{\mathrm{T}}$，$\mathbf{e}_3=(0,0,1)^{\mathrm{T}}$ 和 $\mathbf{e}_4=(1,1,1)^{\mathrm{T}}$ 是射影图像平面 $\mathbb{P}^2$ 的一个射影基．

现在考虑矩阵为 P 的一个摄像机．我们假定摄像机中心 $\mathbf{C}$ 不在任何坐标面上，即 $\mathbf{C}$ 的四个坐标没有一个是零．对于这种情形，在图像上的四点 $\mathrm{P}\mathbf{E}_i$，$i=1,\cdots,4$，没有三个是共线的．因此，可以对该图像施加射影变换 H 使得 $\mathbf{e}_i=\mathrm{HP}\mathbf{E}_i$．我们假设 H 已经施加过，并把 HP 简记为 P．因为 $\mathrm{P}\mathbf{E}_i=\mathbf{e}_i$．矩阵 P 的形式是

$$\mathrm{P}=\begin{bmatrix} a & & & d \\ & b & & d \\ & & c & d \end{bmatrix}.$$

**定义 19.1**　摄像机矩阵 P 称为**简化**摄像机矩阵，如果对每个 $i=1,\ldots,4$，它把 $\mathbf{E}_i$ 映射到 $\mathbf{e}_i$，也就是说，$\mathbf{e}_i=\mathrm{P}\mathbf{E}_i$．

现在，对任何点 $\mathbf{X}=(\mathrm{X},\mathrm{Y},\mathrm{Z},\mathrm{T})^{\mathrm{T}}$ 可以验证

$$\mathrm{P}=\begin{bmatrix} a & & & d \\ & b & & d \\ & & c & d \end{bmatrix}\begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ \mathrm{T} \end{pmatrix}=\begin{pmatrix} a\mathrm{X}+d\mathrm{T} \\ b\mathrm{Y}+d\mathrm{T} \\ c\mathrm{Z}+d\mathrm{T} \end{pmatrix}. \tag{19.1}$$

注意在这个方程中摄像机矩阵的元素和点的坐标之间的对称性．它们可以进行如下交换

$$\begin{bmatrix} a & & & d \\ & b & & d \\ & & c & d \end{bmatrix}\begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ \mathrm{T} \end{pmatrix}=\begin{bmatrix} \mathrm{X} & & & \mathrm{T} \\ & \mathrm{Y} & & \mathrm{T} \\ & & \mathrm{Z} & \mathrm{T} \end{bmatrix}\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}. \tag{19.2}$$

摄像机和点的角色互换可以理解为对偶的一种形式．将它称为 Carlsson－Weinshall 对偶，或更简略地 Carlsson 对偶，这个对偶的内容将在本章余下部分进行探讨．

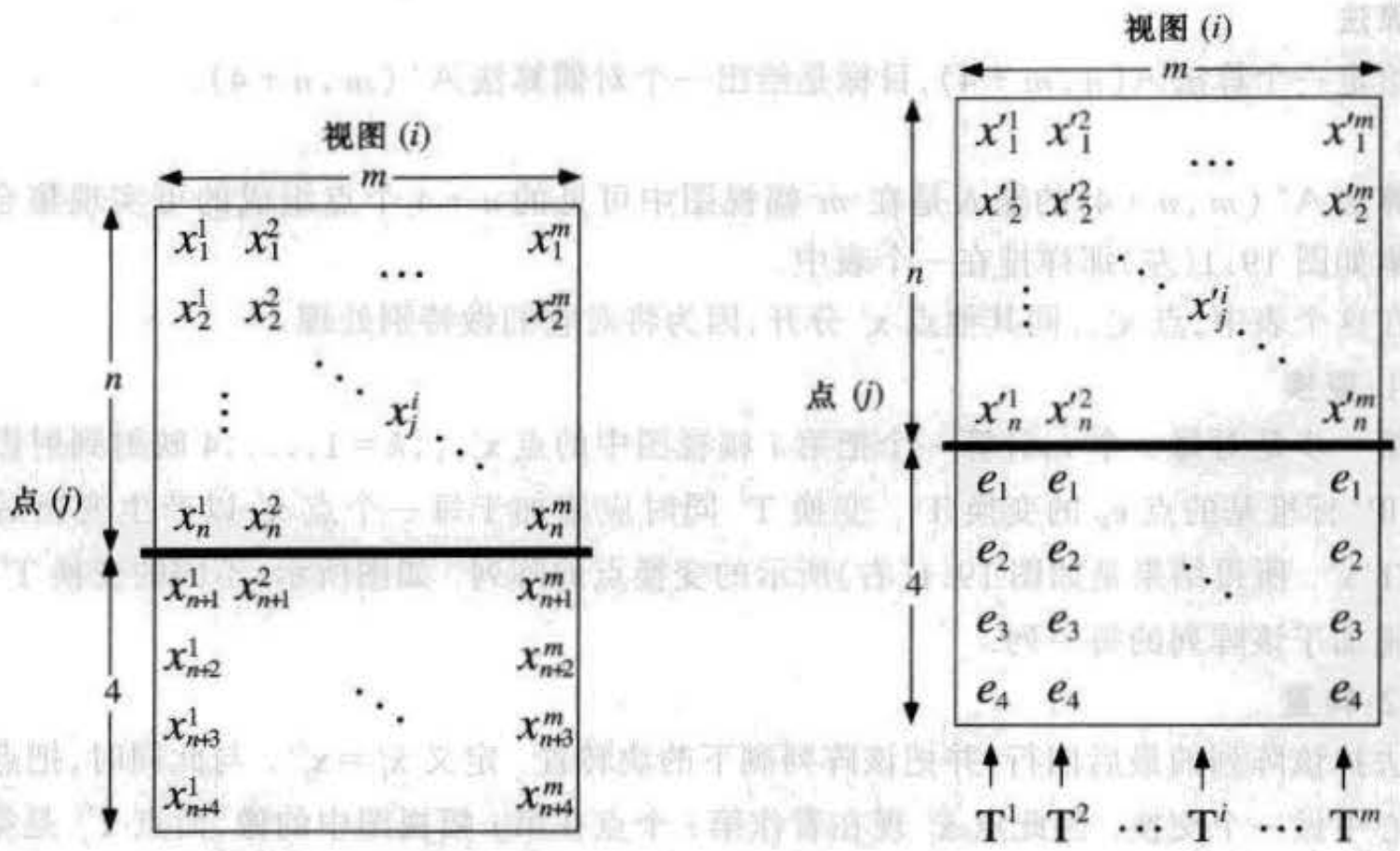



图 19.1 左:算法 $A^*(m,n+4)$的输入. 右:变换后的输入数据.

### 19.1.1 对偶算法

首先,我们将用 Carlsson 对偶从一个给定的射影重构算法推导出一个对偶算法. 具体地说,如果有一个从 $m+4$ 个点的 $n$ 幅视图进行射影重构的算法,那么可以证明存在一个从 $n+4$ 个点的 $m$ 幅视图进行射影重构的算法. 这个由 Carlsson[Carlsson-95]观察到的结论,将通过描述对偶算法步骤而具体化.

我们考虑一个射影重构问题,并将它记为 $\mathcal{P}(m,n)$. 它是从 $n$ 个点的 $m$ 幅视图进行射影重构的问题. 定义图像点为 $\mathbf{x}_j^i$,它表示第 $j$ 个物体空间点在第 $i$ 幅视图上的像. 即上标表示视图的编号,下标表示点的编号. 如果存在一组摄像机矩阵 $\mathrm{P}^i$ 和 3D 点 $\mathbf{X}_j$ 的集合使得 $\mathbf{x}_j^i=\mathrm{P}^i\mathbf{X}_j$,这样的点集$\{\mathbf{x}_j^i\}$便称为**可实现**的. 射影重构问题 $\mathcal{P}(m,n)$就是对给定的 $n$ 个点和 $m$ 幅视图的可实现集合$\{\mathbf{x}_j^i\}$求这样的摄像机矩阵 $\mathrm{P}^i$ 和点 $\mathbf{X}_j$. 这组摄像机和 3D 点一起称为该点对应集合的一个**实现**(**或射影实现**).

令 $\mathcal{A}(n,m+4)$表示求解射影重构问题 $\mathcal{P}(n,m+4)$的一个算法,现在将介绍一个求解射影重构问题 $\mathcal{P}(m,n+4)$的算法. 这个算法将被记为 $\mathcal{A}^*(m,n+4)$,意味着算法 $\mathcal{A}(n,m+4)$的对偶.

作为起步,我们将不加证明地给出算法. 另外,难点将掩饰过去以便阐述总的思想而不至陷入细节之中. 在描述这个算法时,重要的是跟踪指标的范围和弄清楚它们是指摄像机还是点. 因此,下面的说明可能有助于保持跟踪.

- 上标表示视图的编号.
- 下标表示点的编号.
- $i$ 从 1 到 $m$.
- $j$ 从 1 到 $n$.
- $k$ 从 1 到 4.

**对偶算法**

给定一个算法$A(n,m+4)$,目标是给出一个对偶算法$A^*(m,n+4)$.

**输入:**

算法$A^*(m,n+4)$的输入是在 $m$ 幅视图中可见的 $n+4$ 个点组成的可实现集合. 这个点集如图 19.1(左)那样排在一个表中.

在这个表中,点 $\mathbf{x}_{n+k}^i$ 同其他点 $\mathbf{x}_j^i$ 分开,因为将对它们做特别处理.

**步骤 1:变换**

第一步是对每一个 $i$,计算一个把第 $i$ 幅视图中的点 $\mathbf{x}_{n+k}^i, k=1,\ldots,4$ 映射到射影二维空间 $\mathbb{P}^2$ 标准基的点 $\mathbf{e}_k$ 的变换 $\mathrm{T}^i$. 变换 $\mathrm{T}^i$ 同时应施加于每一个点 $\mathbf{x}_j^i$ 以产生变换后的点 $\mathbf{x}'^i_j=\mathrm{T}^i\mathbf{x}_j^i$. 所得结果是如图 19.1(右)所示的变换点的阵列. 如图所示,不同的变换 $\mathrm{T}^i$ 被算出并施加于该阵列的每一列.

**步骤 2:转置**

去掉该阵列的最后四行,并把该阵列剩下的块转置. 定义 $\hat{\mathbf{x}}_i^j=\mathbf{x}_j'^i$. 与此同时,把点和视图在心中做一个交换. 因此点 $\hat{\mathbf{x}}_i^j$ 现在看作第 $i$ 个点在第 $j$ 幅视图中的像,而点 $\mathbf{x}'^i_j$ 是第 $j$ 个点在第 $i$ 幅视图中的像. 这里实质上发生了点和摄像机角色的互换,这正是(19.2)表示的 Carlsson 对偶的基本概念. 这样得到的转置阵列在图 19.2(左)中显示.

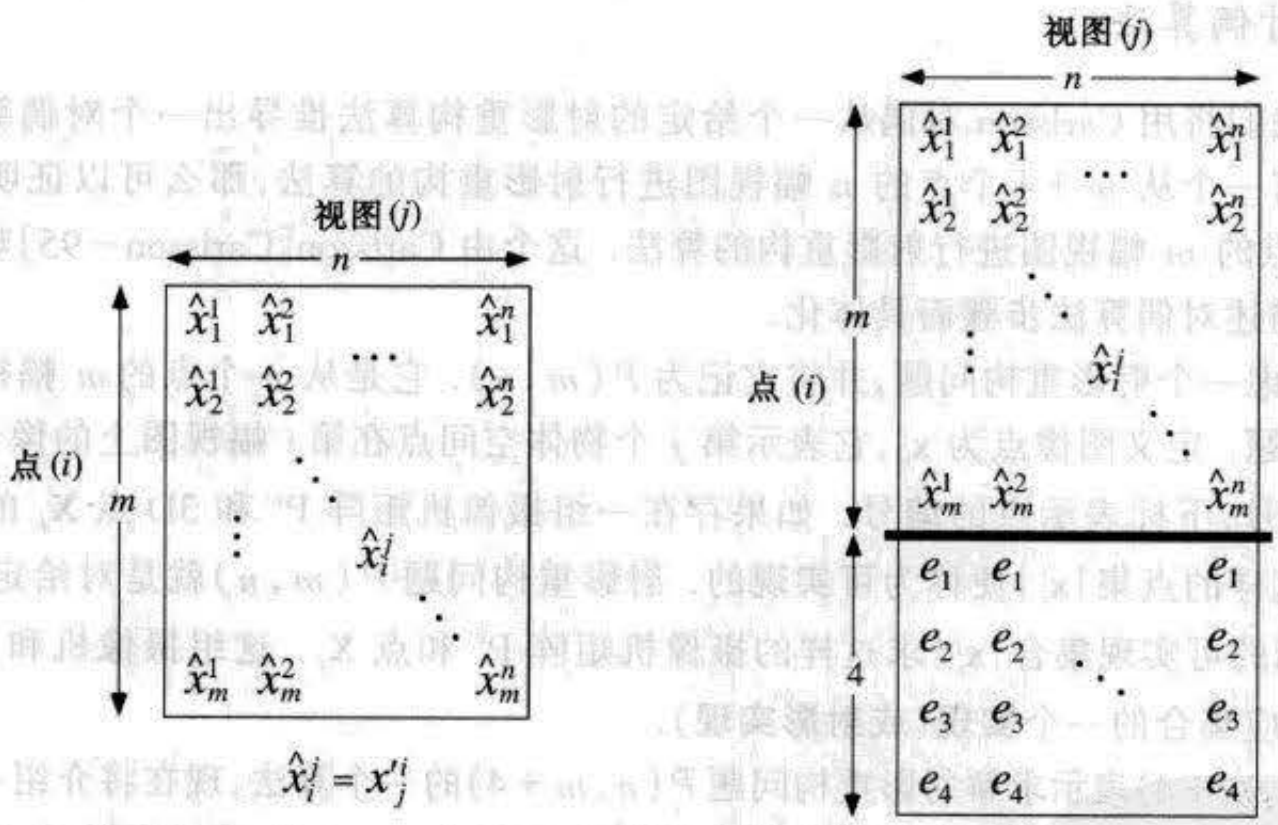



图 19.2　**左:**转置的数据. **右:**由附加额外点扩展的转置数据.

**步骤 3:扩展**

现在把上面所得阵列扩展 4 行:点 $\mathbf{e}_k$ 出现在该阵列中的第$(m+k)$行的所有位置上,如图 19.2(右)所示. 扩展的目的将在节 19.1.2 中解释.

**步骤 4:求解**

上一步得到的点阵有 $m+4$ 行和 $n$ 列,且可以看作是在 $n$ 幅视图中可见的 $m+4$ 个点的位置. 它正是算法$A(n,m+4)$(我们假设该算法已经给出)求得的一个候选解. 这里的本质是这个阵列中的点形成了一个点对应的可实现集合. 其正确性论证暂时推迟. 算法

$A(n, m+4)$的结果是满足 $\hat{\mathbf{x}}_i^j = \hat{P}^j \hat{\mathbf{X}}_i$ 的一组摄像机 $\hat{P}^j$ 和点 $\hat{\mathbf{X}}_i$. 另外,对应于该阵列的最后四行,存在点 $\hat{\mathbf{X}}_{m+k}$,使对所有 $j$ 有 $\mathbf{e}_k = \hat{P}^j \hat{\mathbf{X}}_{m+k}$.

**步骤 5:3D 变换**

由于上一步得到的是射影重构,故可以作变换(或者说,选择一个射影坐标系)使得点 $\hat{\mathbf{X}}_{m+k}$ 成为 $\mathbb{P}^3$ 标准基组成部分即四点 $\mathbf{E}_k$. 唯一的要求是由射影重构得到的点 $\hat{\mathbf{X}}_{m+k}$ 不共面. 这样假设的原因将在后面说明.

至此,我们看到 $\mathbf{e}_k = \hat{P}^j \hat{\mathbf{X}}_{m+k} = \hat{P}^j \mathbf{E}_k$. 由此推出 $\hat{P}^j$ 有特别的形式

$$\hat{P}^j = \begin{bmatrix} a^j & & & d^j \\ & b^j & & d^j \\ & & c^j & d^j \end{bmatrix}. \tag{19.3}$$

**步骤 6:对偶**

令 $\hat{\mathbf{X}}_i = (X_i, Y_i, Z_i, T_i)^\top$,且 $\hat{P}^j$ 如(19.3)所给定. 现在定义点 $\mathbf{X}_j = (a^j, b^j, c^j, d^j)^\top$ 和摄像机

$$P'^i = \begin{bmatrix} X_i & & & T_i \\ & Y_i & & T_i \\ & & Z_i & T_i \end{bmatrix}.$$

则可以验证

$$\begin{aligned} P'^i \mathbf{X}_j &= (X_i a^j + T_i d^j, Y_i b^j + T_i d^j, Z_i c^j + T_i d^j)^\top \\ &= \hat{P}^j \hat{\mathbf{X}}_i \\ &= \hat{\mathbf{x}}_i^j \\ &= \mathbf{x}_j'^i. \end{aligned}$$

另外,如果定义 $\mathbf{X}_{n+k} = \mathbf{E}_k, k = 1, \ldots, 4$,则 $P'^i \mathbf{X}_{n+k} = \mathbf{e}_k$. 于是,摄像机 $P'^i$ 和点 $\mathbf{X}_j$ 与 $\mathbf{X}_{n+k}$ 显然形成由该算法第一步得到的变换数据阵列的一个射影实现.

**步骤 7:反变换**

最后,定义 $P^i = (T^i)^{-1} P'^i$,并加上上一步得到的点 $\mathbf{X}_j$ 和 $\mathbf{X}_{n+k}$,我们便得到原始数据的一个射影实现. 可以验证

$$P^i \mathbf{X}_j = (T^i)^{-1} P'^i \mathbf{X}_j = (T^i)^{-1} \mathbf{x}_j'^i = \mathbf{x}_j^i.$$

至此,完成了该算法的描述. 可以看到它采取以下几个步骤.

(1) 在第 1 步中,通过选择四个特殊点,数据被变换到标准的图像参考坐标系中.

(2) 在第 2 和第 3 步中,问题被映射到对偶域,产生一个对偶问题 $\mathcal{P}(n, m+4)$.

(3) 在第 4 和第 5 步中,求解对偶问题.

(4) 第 6 步把所得解映射回原始定义域.

(5) 第 7 步撤消初始变换的效应.

### 19.1.2 算法合理性的证明

为了证明算法的合理性,我们必须确认在第 4 步中变换后的问题确实存在一个解. 在

考虑这一点之前，有必要先解释第 3 步以及第 5 步的目的，前者通过增加包含图像点 $\mathbf{e}_k$ 的行来扩展数据；后者把任意的射影解变换到其中四个点等于 3D 基点 $\mathbf{E}_k$ 的一个解.

这两步的目的是保证对偶重构问题得到一个解，其中的 $\hat{\mathrm{P}}^j$ 具有(19.3)给出的特殊形式，即摄像机矩阵只用 4 个值来参数化. 把对偶算法这样来描述是为了使它对无论什么算法 $A(n, m+4)$ 都有效. 然而，如果已知的算法 $A(n, m+4)$ 具有直接对摄像机施加这个约束的能力，则第 3 和第 5 两步都可以省去. 正如即将见到的，基于基本矩阵、三焦点或四焦点张量的算法都可以容易地按这种方式修改.

同时，由于具有(19.3)形式的 $\hat{\mathrm{P}}^j$ 称为**简化的摄像机矩阵**. 我们称一个图像对应集合的任何重构为**简化重构**，只要它的每个摄像机矩阵具有这种形式. 但是，并非所有可实现的点对应的集合允许一个简化实现，下面的结论刻画了具有这种性质的点对应集合.

**结论 19.2　一个图像点的集合 $\{\mathbf{x}_j^i : i=1,\ldots,m; j=1,\ldots,n\}$ 有一个简化实现的充要条件是它可以通过增加补充对应 $\mathbf{x}_{n+k}^i = \mathbf{e}_k, k=1,\ldots,4$，使得**

**(1) 图像对应的整个集合是可实现的，并且**

**(2) 对应于补充的图像对应的重构点 $\mathbf{X}_{n+k}$ 不共面.**

**证明**　证明是很直截了当的. 假设这个对应集合允许一个简化实现，并令 $\mathrm{P}^i$ 是简化摄像机矩阵集合，令点 $\mathbf{X}_{n+k} = \mathbf{E}_k, k=1,\ldots,4$，被投影到 $m$ 幅图像上. 对所有的 $i$，这些投影是 $\mathbf{x}_{n+k}^i = \mathrm{P}^i\mathbf{X}_{n+k} = \mathrm{P}^i\mathbf{E}_k = \mathbf{e}_k$.

反过来，假设增加的点集是可实现的，并且点 $\mathbf{X}_{n+k}$ 不共面. 此时，可以选择一个射影基使得 $\mathbf{X}_{n+k} = \mathbf{E}_k$. 于是对每一幅视图及所有 $k$ 有 $\mathbf{e}_k = \mathrm{P}^i\mathbf{E}_k$. 由此得出：每一个 $\mathrm{P}^i$ 有所希望的形式(19.3). □

在证明这个算法的正确性之前，必须再做一个注释.

**结论 19.3　如果一个图像点集 $\{\mathbf{x}_j^i : i=1,\ldots,m; j=1,\ldots,n\}$ 允许一个简化实现，那么转置集合 $\{\hat{\mathbf{x}}_i^j : j=1,\ldots,n; i=1,\ldots,m\}$ 也如此，其中对所有 $i$ 和 $j$，有 $\hat{\mathbf{x}}_i^j = \mathbf{x}_j^i$.**

这是基本的对偶性质，上面算法的第 6 步给出了有效地构造性证明. 现在可以证明算法的正确性了.

**结论 19.4　若 $\mathbf{x}_j^i$ 和 $\mathbf{x}_{n+k}^i$(如图 19.1(左)所示)是一组可实现的图像点对应，并且假设**

**(1) 对每个 $i$，4 个点 $\mathbf{x}_{n+k}^i (k=1,\ldots,4)$ 不包括三个共线点.**

**(2) 在一个射影重构中，四个点 $\mathbf{X}_{n+k}$ 不共面.**

**那么，节 19.1.1 中的算法将会成功.**

**证明**　由于第一个条件，对每一个 $i$ 把输入数据变换到图 19.1(右)显示的数据的变换 $\mathrm{T}^i$ 存在. 变换后的数据也是可实现的，因为变换后的数据与原始数据仅相差一个图像射影变换.

现在，把结论 19.2 应用于图 19.1(右)时，对应的 $\mathbf{x}_j'^i$ 允许一个简化实现. 根据结论 19.3，转置的数据(图 19.2(左))也允许一个简化实现. 再用一次结论 19.2 证明扩展的数据(图 19.2(右))是可实现的. 而且，点 $\hat{\mathbf{X}}_{m+k}$ 不共面，因此第 5 步有效. 接下来进行第 6 和第 7

步就没有问题了. □

第一个条件可以从图像对应 $\mathbf{x}_j'^i$ 来验证. 也许有人会认为验证第二个条件需要实现重构. 然而,无需进行重构就可以验证重构的点是否共面. 这留给读者作为一个习题($p241$).

## 19.2 简化重构

本节,我们主要考虑并重新评价上节介绍的算法中的第 3－5 步. 扼要地说,这些步骤的目的是为了从一组图像对应得到一个简化的重构. 因此,输入的是允许简化实现的一组图像对应 $\hat{\mathbf{x}}_i^j$(见图 19.2(左)). 而输出的则是一组使得对所有 $i,j$,有 $\hat{\mathrm{P}}^j\hat{\mathbf{X}}_i = \hat{\mathbf{x}}_i^j$ 的**简化的**摄像机矩阵 $\hat{\mathrm{P}}^j$ 和点 $\hat{\mathbf{X}}_i$.

如我们所知道的,进行这个给定算法的一种方法(如步骤 3－5)是把另外 4 个合成的点对应 $\hat{\mathbf{x}}_{m+k}^j$ 加入该点集,实行射影重构,然后应用 3D 单应使得 3D 点 $\hat{\mathbf{X}}_{m+k}$ 映射到 $\mathbb{P}^3$ 的一个射影基的基点 $\mathbf{E}_k$. 它的问题在于有噪声存在时,射影重构是不准确的. 因此,用这个方法得到的摄像机矩阵将把点 $\mathbf{E}_k$ 映到接近但不等于 $\mathbf{e}_k$ 的点. 这意味着摄像机矩阵不是准确的简化形式. 因此,我们现在要考虑计算点对应方法的实现,使其摄像机是精确简化的.

### 19.2.1 简化的基本矩阵

这些对偶的方法最显然的应用是关于基本矩阵和三焦点张量的对偶重构算法. 它将引出跨 $N$ 幅视图(分别)对 6 或 7 点的重构算法. 本节考虑由 6 点的重构. 重构问题 $\mathcal{P}(N,6)$ 的对偶是问题 $\mathcal{P}(2,N+4)$,即在两幅视图中由 $N+4$ 点的重构. 第 9 章中有关基本矩阵的方法是解决这个问题的标准方法.

为此,我们定义简化基本矩阵.

**定义** 19.5. 基本矩阵 $\hat{\mathrm{F}}$ 称为**简化**基本矩阵,如果它满足条件 $\mathbf{e}_i^{\mathsf{T}}\hat{\mathrm{F}}\,\mathbf{e}_i = 0, i = 1,\ldots,4$.

由于简化基本矩阵已经满足从四组点对应推导的约束,它显然可以从少量的附加点计算出来,事实上,线性计算需四个点,而非线性计算需三个点.

### 19.2.2 简化基本矩阵的计算

对一个简化基本矩阵,条件 $\mathbf{e}_i^{\mathsf{T}}\hat{\mathrm{F}}\,\mathbf{e}_i = 0\,(i = 1,\ldots,3)$ 给出 $\hat{\mathrm{F}}$ 的对角元素是零. 要求 $(1,1,1)\hat{\mathrm{F}}(1,1,1)^{\mathsf{T}} = 0$ 又给出另外一个条件:$\hat{\mathrm{F}}$ 的元素和为零. 因此可以把 $\hat{\mathrm{F}}$ 的形式写为

$$\hat{\mathrm{F}} = \begin{bmatrix} 0 & p & q \\ r & 0 & s \\ t & -(p+q+r+s+t) & 0 \end{bmatrix} \tag{19.4}$$

从而,使这个基本矩阵参数化并满足所有四个线性约束(但不是条件 $\det\hat{\mathrm{F}} = 0$). 现在,不难看出满足 $\mathbf{x}'^{\mathsf{T}}\hat{\mathrm{F}}\,\mathbf{x} = 0$ 的另外的一组点对应 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 将提供 $\hat{\mathrm{F}}$ 的参数 $p,\ldots,t$ 的一个线性方程. 给定最少四组这样的对应,可以在相差一个不连贯的比例因子下解出这些参数. 只给定三个这样的对应,则额外约束 $\det\hat{\mathrm{F}} = 0$ 可以用来提供确定 $\hat{\mathrm{F}}$ 所必要的额外约束. 可能存在一

个或三个解. 这个计算类似于节 10.1.2($p$192)从七组点对应计算基本矩阵所采用的方法. 在给定四组或更多对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$ 时,可以求最小二乘解.

### 19.2.3 恢复简化摄像机矩阵

计算与简化基本矩阵对应的一对简化摄像机矩阵却惊人地棘手. 我们不可以按通常的射影摄像机的情形假设第一个摄像机是$[\mathrm{I}|\mathbf{0}]$,因为它是非一般的,其摄像机的中心对应于基点 $\mathbf{E}_4=(0,0,0,1)^{\mathrm{T}}$. 然而,我们可以另假定这一对摄像机有如下形式

$$\mathrm{P}=\begin{bmatrix}1 & & & 1\\ & 1 & & 1\\ & & 1 & 1\end{bmatrix} \text{和 } \mathrm{P}'=\begin{bmatrix}a & & & d\\ & b & & d\\ & & c & d\end{bmatrix}. \tag{19.5}$$

因为第一个摄像机的中心是$(1,1,1,1)^{\mathrm{T}}$,它相对基点 $\mathbf{E}_1,\ldots,\mathbf{E}_4$ 是一个一般点. 进一步,如果 $d\neq 0$,则我们可以假设 $d=1$,但是我们不主张这样做.

对应于这对摄像机的简化基本矩阵是

$$\hat{\mathrm{F}}=\begin{bmatrix}0 & b(d-c) & -c(d-b)\\ -a(d-c) & 0 & c(d-a)\\ a(d-b) & -b(d-a) & 0\end{bmatrix}. \tag{19.6}$$

读者可以验证它满足四个线性约束以及零行列式的条件. 目前的任务是在给定基本矩阵条件下恢复$(a,b,c,d)$的值. 它看上去好像是要联立求解二次方程组,但是存在如下的一个线性求解方法.

(1)通过解齐次线性方程组:

$$\begin{bmatrix}f_{12} & f_{21} & 0\\ f_{13} & 0 & f_{31}\\ 0 & f_{23} & f_{32}\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}=\mathbf{0} \tag{19.7}$$

可以求得比率 $a:b:c$,其中 $f_{ij}$ 是 $\hat{\mathrm{F}}$ 的第 $ij$ 元素. 这里出现的矩阵显然与 $\hat{\mathrm{F}}$ 有相同的秩(即 2),因此这个方程组存在唯一解(相差尺度因子). 解 $a:b:c=A:B:C$ 提供了关于 $a$,$b$ 和 $c$ 的一组齐次方程,即 $Ba=Ab$,$Ca=Ac$ 和 $Cb=Bc$,其中的两个是线性独立的.

(2)类似地,通过解方程组$(d-a,d-b,b-c)\hat{\mathrm{F}}=\mathbf{0}$ 可以求出比率 $d-a:d-b:d-c$. 这个解也是唯一的. 这又提供了关于 $a,b,c,d$ 的另外两个线性方程.

(3)由该方程组可以在相差一个尺度因子的意义下求解$(a,b,c,d)$,并可以根据(19.5)重构第二个摄像机矩阵.

### 19.2.4 由三视图中六点求解

简化基本矩阵计算的最小配置是三个点,此时可能存在三个解. 根据对偶化规则,它可转化为用跨三视图中的六个点来解的一个重构问题. 它在三焦点张量作野值检测的应用已在算法 15.4($p$286)中介绍. 出于对最小配置解本身的兴趣,并由于它的实际应用,这里把它作为算法 19.1 显式地给出. 这个算法基本上是把前面介绍过的内容放在一起,然而还是

有若干小的改动.

在算法 19.1 中,摄像机矩阵的最后估计在原先点测量的定义域中进行. 另一种做法是如节 19.1.1 中的基本算法那样在对偶的定义域中运用 DLT 算法,. 然而,现在的方法似乎更简单. 它的优点是不需要应用逆变换 T,T′,T″. 更主要的是最后利用原始数据来计算摄像机矩阵. 这是重要的,因为变换可能严重扭曲数据的噪声特征.

这个算法和下节介绍的 $n$ 视图的算法的基础是对偶基本矩阵 $\hat{F}$. 注意对偶基本矩阵如何表示同一图像中点之间的关系. 事实上,用来解 $\hat{F}$的方程是由同一图像中的点构造的. 它与标准的基本矩阵不同,后者所编码的关系是不同图像中看到的点之间的关系.

目标

给定跨 3 视图的 6 组点对应 $\mathbf{x}_j \leftrightarrow \mathbf{x}_j' \leftrightarrow \mathbf{x}_j''$,由这些点计算一个重构,它包括 3 个摄像机矩阵 P, P′和 P″以及 6 个 3D 点 $\mathbf{X}_1, \ldots, \mathbf{X}_6$.

算法

(1)选择 4 个点,其中没有 3 个点在 3 幅视图的任一幅中共线. 令它们是对应 $\mathbf{x}_{2+k} \leftrightarrow \mathbf{x}_{2+k}' \leftrightarrow \mathbf{x}_{2+k}''$, $k=1,\ldots,4$.

(2)对第一视图求一个射影变换 T,它把每个 $\mathbf{x}_{2+k}$映射到 $\mathbf{e}_k$. 把 T 作用于另外两点 $\mathbf{x}_1$ 和 $\mathbf{x}_2$,产生点 $\hat{\mathbf{x}}_1 = T\mathbf{x}_1$ 和 $\hat{\mathbf{x}}_2 = T\mathbf{x}_2$.

(3)如(19.4)一样,由**对偶对应** $\hat{\mathbf{x}}_1 \leftrightarrow \hat{\mathbf{x}}_2$,得出一个关于简化基本矩阵 $\hat{F}$的元素 $p,q,r,s,t$ 的方程. 该方程由关系 $\hat{\mathbf{x}}_2^T \hat{F} \hat{\mathbf{x}}_1 = 0$ 导出.

(4)用与上面两步一样的方法,由另外两幅视图的点形成另外两个方程. 它产生 5 个未知量的 3 个齐次方程. 因而存在简化基本矩阵的 2 参数解簇,它由两个独立的解 $\hat{F}_1$ 和 $\hat{F}_2$ 生成.

(5)一般解是 $\hat{F} = \lambda\hat{F} + \mu\hat{F}$. 根据要求:$\det \hat{F} = 0$,我们可以导出关于$(\lambda,\mu)$的一个齐次三次方程,由它解出$(\lambda,\mu)$并因此得到 $\hat{F}$. 将有 1 个或 3 个实解. 下面将对每个解轮流应用以下步骤.

(6)用节 19.2.3 的方法求在(19.5)中定义的第二个简化摄像机矩阵 $\hat{P}'$的参数$(a,b,c,d)^T$.

(7)我们在原始测量区域中完成重构. 由$(a,b,c,d)$定义的摄像机 $\hat{P}'$的对偶是点 $\mathbf{X}_2 = (a,b,c,d)^T$. 因此 6 个 3D 点是 $\mathbf{X}_1 = (1,1,1,1)^T$, $\mathbf{X}_2 = (a,b,c,d)^T$ 和 $\mathbf{X}_{2+k} = \mathbf{E}_k$, $k=1,\ldots,4$. 这就给出重构景物的结构. 我们然后可以用节 6.1($p$121)介绍的摄像机标定的 DLT 算法来计算 3 个摄像机矩阵. 由于这里我们需要相对于原摄像机坐标定义的摄像机矩阵,我们是用原始坐标求解 P,P′和 P″,使得 $P\mathbf{X}_j = \mathbf{x}_j$ 等. 因为是准确解,DLT 解将是充分的.

算法 19.1　由三视图中的六点计算一个射影重构的算法

### 19.2.5　$n$ 视图中的六点

对用于三视图六点的方法作很少修改就能应用于多视图六个点的情形,主要的差别是简化的基本矩阵 $\hat{F}$将由数据唯一地确定. 具体地说,在节 19.2.4 算法的第四步,每幅视图贡献一个方程,对于四幅或更多的视图,将足够确定 $\hat{F}$.

对这种数据冗余的情形,必须当心噪声的影响. 如我们已经指出的那样,显然应该用原始的未变换的点进行该算法的最后一步. 这能减轻因采用变换的点而可能造成噪声畸变的效应.

### 19.2.6 $n$ 视图中的七点

$n$ 视图中的七点问题与三视图中的 $n+4$ 点的情形相对偶，并且用 $n$ 组点对应计算简化三焦点张量的方法求解.

**定义 19.6** 三焦点张量 $T$ 被称为**简化三焦点张量**，如果它满足由合成点对应 $\mathbf{e}_k \leftrightarrow \mathbf{e}_k' \leftrightarrow \mathbf{e}_k''$，$k=1,\ldots,4$，所产生的线性约束.

由 7 点重构的一般方法类似于在 $n$ 视图中的六点方法，除了用简化的三焦点张量代替简化基本矩阵外，仍存在着一些小的差别.

在计算简化三焦点张量时，对应于合成的对应 $\mathbf{e}_j \leftrightarrow \mathbf{e}_j \leftrightarrow \mathbf{e}_j$ 的约束应该完全准确地满足，而用来计算张量的其他对应是有噪声的，故仅仅是近似地满足. 若不然，则计算的张量将没有准确的简化形式. 对于简化的基本矩阵的情形，通过简化基本矩阵的一种特殊的参数化来处理，也就是说，利用参数化使得由合成对应产生的约束自动地满足(见(19.4)). 对于三焦点张量情形，这样一种便利的参数化的可能性不明显. 该合成约束的形式是

$$e^i e^p e^q \varepsilon_{jpr} \varepsilon_{kqs} T_i^{jk} = 0_{rs}.$$

它比简化基本矩阵的线性约束要复杂得多. 我们可以按下面给出的另一方法进行.

在通常计算三焦点张量的线性方法中，我们必须求解形如 $\mathrm{A}\mathbf{t}=\mathbf{0}$ 的一组线性方程组，或者更准确地说，寻求满足 $\|\mathbf{t}\|=1$ 并使 $\|\mathrm{A}\mathbf{t}\|$ 最小的矢量 $\mathbf{t}$. 在求解简化三焦点张量时，矩阵 A 可以分成两个部分，一部分对应于应该准确地满足由合成对应产生的约束，另一部分对应于由实际对应产生的约束，它必须在最小二乘的意义下满足. 第一种约束集合的形式是 $\mathrm{C}\mathbf{t}=\mathbf{0}$，而第二种集合可以记为 $\hat{\mathrm{A}}\mathbf{t}=\mathbf{0}$. 问题变成：求在约束 $\|\mathbf{t}\|=1$ 和 $\mathrm{C}\mathbf{t}=\mathbf{0}$ 下最小化 $\|\hat{\mathrm{A}}\mathbf{t}\|$ 的 $\mathbf{t}$. 求解这个问题的一个算法由算法 A3.6($p$413)给出.

对于从简化张量求三个摄像机矩阵的问题，好像还没有一种类似于节 19.2.3 对基本矩阵所介绍的那样简单方法存在. 我们可以用在一般对偶算法第 5 和第 6 步中所介绍的方法来代替(见 $p$357).

这种类型的最小配置是两视图的 7 点，对于这种情形，最好直接地用节 10.1.2($p$192)的方法求解，而不是采用 6 视图 3 点的对偶化方法.

### 19.2.7 与性能有关的问题

基于简化基本矩阵和三焦点张量的对偶重构方法已被实现和测试. 这些测试的结果在一个学生的报告中，由 Gilles Debunne 于 1996 年 8 月提供. 由于这个报告实际上已经不复存在，我们把该结果总结在这里.

最严重的困难是由于对图像进行了射影变换 $\mathrm{T}^i$ 而使噪声分布的失真. 对图像数据进行射影变换会使任何加在数据上的噪声分布失真. 这个问题与在任何图像上必须选择四个非共线点有关. 如果在任一图像上这些点接近共线，那么用于图像的射影变换(在该算法的第 1 步中)可以使该图像产生极大的失真. 这种失真可以严重降低算法的性能.

如果不对噪声失真给予特别关注，则算法的性能一般不令人满意，尽管在算法的第 4－6 步($p$356)极小心地最小化由于噪声产生的误差，当第 7 步用了逆射影变换后，平均误差变得非常

大. 虽然某些点保持非常小的误差,而在那些失真非常大的图像中却产生相当大的误差.

用节 3.4.4($p67$)意义上的归一化也有问题. 业已证明了对数据进行归一化对于线性重构算法的性能是很关键的. 然而,对图 19.1($p355$)(右)已变换过的数据应该采用哪一类归一化使它在几何上与实际图像的测量无关仍是一个迷.

为得到好的结果,好像有必要把算法第 1 步假定的误差分布传递下去以便得到变换过的数据图 19.1($p355$)(右)的假定的误差分布,然后在重构过程中最小化与这个传递误差分布有关的残差. 或者说,在重构过程中要最小化的代价函数必须回过来与原始图像点的测量误差相联系. [Hartley - 2000b, Schaffalitzky - 2000]中报告的最近研究工作证明它的确使结果显著提高.

## 19.3 结束语

### 19.3.1 文献

Carlsson - Weinshall 对偶的基本思想首先在两篇同时出现的论文:[Carlsson - 95]和 Weinshall[Weinshall - 95],接着在一篇合著的论文[Carlsson - 98]中被介绍. 这里所给出的对偶一个算法的一般方法的处理是由[Hartley - 98c]给出,它由这些早期论文推导得来. 研究噪声传递方法的细节在[Hartley - 2000]中给出,它建立在 Gillles Debunne(1996 年 8 月)的一篇现已不存在的报告的基础上. 由 $n$ 视图六点重构的另一种方法在[Schaffalitzky - 2000]中给出.

由三视图六点的重构问题也许首先在一个技术报告[Hartley - 92b]中研究(后来发表为[Hartley - 94b]),其中证明了最多存在八个解. [Quan - 94] 给出了这个问题的一个完整的解答,其中证明了只有三个可能的解,这一点也在[Quan - 94]中指出. 论文[Carlsson - 95] 指出这个问题是两视图七点问题的对偶并且它的解是已知的. 这使本章给出的方法有可能系统地讲述. [Torr - 97]用最小的六点配置来对三焦点张量进行鲁棒性估计.

### 19.3.2 注释和练习

(1)节 19.1.1 的对偶算法曾提到这个方法只有在定义图像变换的 4 个点不共面时才有效. 不过请注意,在这种情形,节 17.4($p316$)的算法将线性地计算出一个射影重构.

(2)如果所选择的 4 点是共面的,则单应 $T^i$ 将把这个平面映到一个坐标系. 然后变换得到的点 $\mathbf{x}_j'^i$ 将满足节 16.5.2($p302$)的条件,即连接任何一对点 $\mathbf{x}_j'^i$ 和 $\mathbf{x}_k'^i$($j$ 和 $k$ 是固定的,而每一个 $i$ 对应不同的直线)的直线将相交于一个共同点. 这种情形的对偶在[Criminisi - 98, Irani - 98]中介绍.

(3)仍然是所选择的 4 点为共面的情形:在施加 $T^i$ 之后,这 4 点所在平面的任何点将映射到所有图像中的同一点. 因此,与这组点对应相一致的基本矩阵将是反对称的.

# 第 20 章 正负性

当我们用一组点对应来进行景物的射影重构时，通常非常重要的一条信息被忽略了——如果点在图像上可见，那么它们必然在摄像机的前面. 一般，射影重构的景物与欧氏坐标系解释的实际景物不能保证非常相似. 景物常常在无穷远平面处被切开，图 20.1 用二维的情形给予说明. 把这个简单的约束考虑进来后，就可能使重构至少非常接近于景物的仿射重构. 这样产生的重构称为"准仿射"，它介于射影重构和仿射重构之间. 景物对象在无穷远平面处不再分开，虽然它们也许仍然存在严重的射影失真.

如果我们不介意摄像机而仅要求景物有正确的准仿射形式，那么把射影重构转变到准仿射是非常简单的——事实上它能在大约两行编程中完成(见推论 20.9). 如果还要求摄像机则需要解一个线性规划问题.

## 20.1 准仿射变换

设 $B$ 为 $\mathbb{R}^n$ 的子集，如果连接 $B$ 中任何两点的线段完全在 $B$ 内，则称 $B$ 为凸集. 包含 $B$ 的最小凸集称为 $B$ 的凸包，记为 $\overline{B}$. 我们主要关心 3 维点集，因此 $n=3$. 我们视 $\mathbb{R}^3$ 是 $\mathbb{P}^3$ 的一个子集，它由所有非无穷远点组成. 无穷远点构成无穷远平面，记为 $\boldsymbol{\pi}_\infty$. 这样一来，$\mathbb{P}^3=\mathbb{R}^3\cup\boldsymbol{\pi}_\infty$. $\mathbb{P}^3$ 的子集是凸集的充要条件是它包含在 $\mathbb{R}^3$ 中而且是 $\mathbb{R}^3$ 中的凸集. 因此，根据这个定义，一个凸集不包含任何无穷远点.

**定义 20.1** 考虑一点集 $\{\mathbf{X}_i\}\subset\mathbb{R}^3\subset\mathbb{P}^3$. 一个射影映射 $h:\mathbb{P}^3\to\mathbb{P}^3$ 称为是保持了点 $\{\mathbf{X}_i\}$ 的凸包的，如果

(1) 对所有 $i$，$h(\mathbf{X}_i)$ 是一个有限点，并且

(2) $h$ 把点 $\{\mathbf{X}_i\}$ 的凸包双射到点 $\{h(\mathbf{X}_i)\}$ 的凸包.

在图 20.1 中显示的例子可以帮助理解这个定义. 这个例子是关于 2D 点集的，但是原理是一样的. 图中显示了一把梳子的图像以及根据一个射影映射重取样的图像. 然而，该射影映射不保持梳子的凸包. 大部分人会认同这个重取样的图像与摄像机或人类的眼睛所见到的一把梳子的任何视图不一样.

如下面给出的定理所示，保持点集的凸包的性质可以用各种不同的方式来表征. 为了叙述这个定理，我们引入一个新的记号.

**记号** 符号 $\hat{\mathbf{X}}$ 是点 $\mathbf{X}$ 的最后一个坐标等于 1 的齐次表示.

在本章中，我们一般感兴趣的是表示成齐次矢量的 3D 点之间的准确等式(不是相差一个常数的等式). 这样一来，例如，如果 H 是一个射影变换，我们可以用 $\mathrm{H}\hat{\mathbf{X}}=\tau'\hat{\mathbf{X}}'$ 来表示把点 $\hat{\mathbf{X}}$ 映射到点 $\hat{\mathbf{X}}'$ 的变换，常数因子 $\tau'$ 要求能使这个等式准确地成立.

现在来叙述该定理.

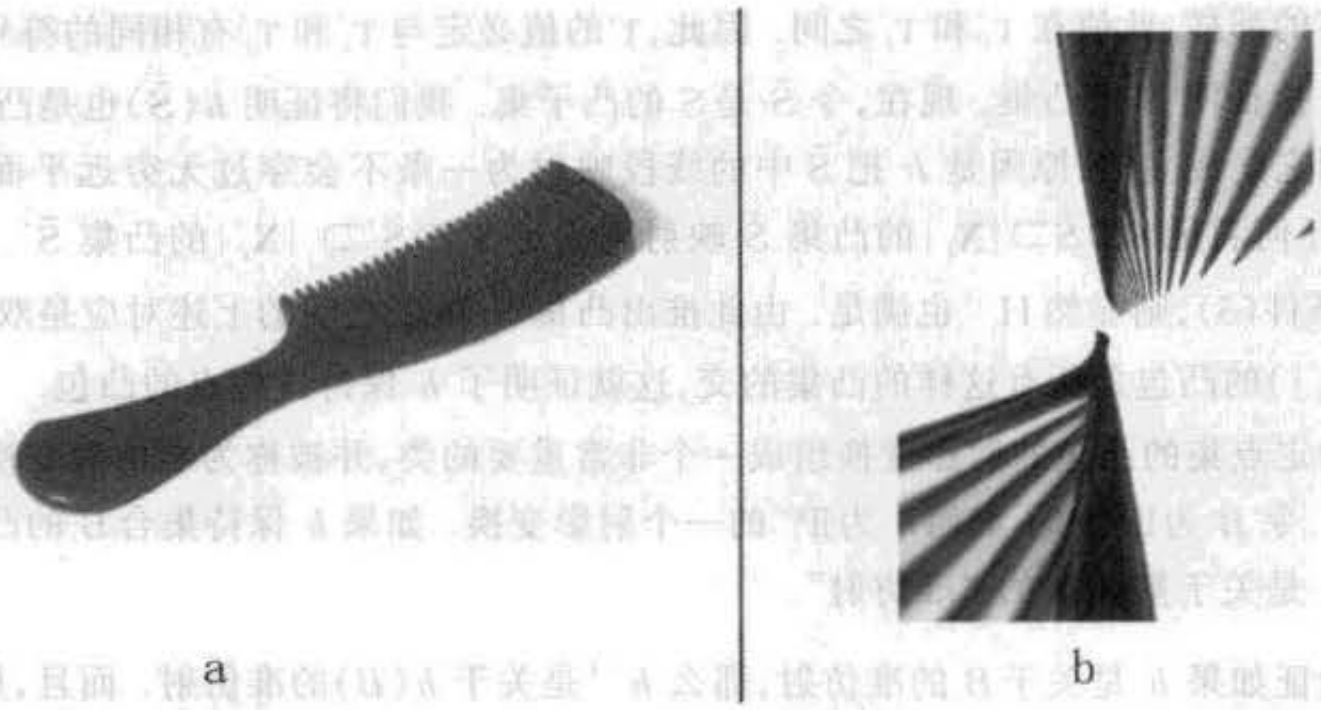



图 20.1 (a)一把梳子的图像,(b)对该图像施加一个射影变换的结果. 这个射影变换不保持由梳子的点组成的凸包. 在原始图像中,梳子的凸包是包含在可视图像内的有限集合. 但是,此凸包内的某些点被该变换映射到无穷远.

**定理 20.2 考虑一个射影变换 $h: \mathbb{P}^3 \to \mathbb{P}^3$ 和一个点集合 $\{\mathbf{X}_i\}$, $\boldsymbol{\pi}_\infty$ 是由 $h$ 映射到无穷远的平面. 下面的陈述是等价的.**

**(1) $h$ 保持点集 $\{\mathbf{X}_i\}$ 的凸包.**

**(2) 对点集 $\{\mathbf{X}_i\}$ 的凸包中的任何点 $\tilde{\mathbf{X}}$ 有 $\tilde{\mathbf{X}} \cap \boldsymbol{\pi}_\infty = \emptyset$ 成立.**

**(3) 令 H 是表示变换 $h$ 的一个矩阵,并假设 $\mathrm{H}\hat{\mathbf{X}}_i = \mathrm{T}_i'\hat{\mathbf{X}}_i'$,则所有的常数 $\mathrm{T}_i'$ 都有相同的符号.**

**证明** 我们将按(1)⇒(2)⇒(3)⇒(1)来证明.

(1)⇒(2) 如果 $h$ 保持点的凸包,那么对于 $\mathbf{X}_i$ 凸包中的任何点 $\tilde{\mathbf{X}}$, $h(\tilde{\mathbf{X}})$ 都是有限点. 因此 $\tilde{\mathbf{X}} \cap \boldsymbol{\pi}_\infty = \emptyset$.

(2)⇒(3) 考虑连接两点 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 的弦并假设 $\mathrm{T}_i'$ 和 $\mathrm{T}_j'$(如在定理的第三部分中)有相反的符号. 因为 $\mathrm{T}_i'$ 是 $\mathbf{X}_i$ 坐标的连续函数,所以在从 $\mathbf{X}_i$ 到 $\mathbf{X}_j$ 的弦上一定存在一个点使得 $\mathrm{T}'$ 等于零. 这意味着 $\mathrm{H}\tilde{\mathbf{X}} = (\tilde{\mathrm{X}}', \tilde{\mathrm{Y}}', \tilde{\mathrm{Z}}', 0)^\mathsf{T}$. 由于 $\tilde{\mathbf{X}}$ 在点 $\{\mathbf{X}_i\}$ 的凸包中,这与(2)相悖.

(3)⇒(1) 我们假设使 $\mathrm{H}\hat{\mathbf{X}}_i = \mathrm{T}_i'\hat{\mathbf{X}}_i'$ 成立的常数 $\mathrm{T}_i'$ 有相同符号. 令 $S$ 是 $\mathbb{P}^n$ 的一个子集,它包含所有满足条件 $\mathrm{H}\hat{\mathbf{X}} = \mathrm{T}'\hat{\mathbf{X}}'$ 的点 $\mathbf{X}$,其中 $\mathrm{T}'$ 与 $\mathrm{T}_i'$ 有相同符号. 集合 $S$ 包含 $\{\mathbf{X}_i\}$. 我们将证明 $S$ 是凸集. 如果 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 是 $S$ 的两点,其对应常数是 $\mathrm{T}_i'$ 和 $\mathrm{T}_j'$,则任何介于 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 之间的中间点 $\mathbf{X}$ 必有介于 $\mathrm{T}_i'$ 和 $\mathrm{T}_j'$ 之间的 $\mathrm{T}'$ 的值. 为说明这一点,考虑一个点 $\hat{\mathbf{X}} = \alpha\hat{\mathbf{X}}_i + (1-\alpha)\hat{\mathbf{X}}_j$,其中 $0 \leqslant \alpha \leqslant 1$. 此点在 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 之间. 用 $\mathbf{h}_4^\mathsf{T}$ 记 H 的最后一行. 则,

$$
\begin{aligned}
\mathrm{T}' &= \mathbf{h}_4^\mathsf{T}\hat{\mathbf{X}} \\
&= \mathbf{h}_4^\mathsf{T}(\alpha\hat{\mathbf{X}}_i + (1-\alpha)\hat{\mathbf{X}}_j) \\
&= \alpha\mathbf{h}_4^\mathsf{T}\hat{\mathbf{X}}_i + (1-\alpha)\mathbf{h}_4^\mathsf{T}\hat{\mathbf{X}}_j \\
&= \alpha\mathrm{T}_i' + (1-\alpha)\mathrm{T}_j'
\end{aligned}
$$

如断言的那样，此值在 $T_i'$ 和 $T_j'$ 之间. 因此，$T'$ 的值必定与 $T_i'$ 和 $T_j'$ 有相同的符号，从而 $\mathbf{X}$ 也在 $S$ 中. 这证明 $S$ 是凸集. 现在，令 $\tilde{S}$ 是 $S$ 的凸子集. 我们将证明 $h(\tilde{S})$ 也是凸集. 可以容易地看到它是成立的，原因是 $h$ 把 $\tilde{S}$ 中的线段映射为一条不会穿过无穷远平面的线段. 因此，$h$ 把任何满足 $S\supset\tilde{S}\supset\{\mathbf{X}_i\}$ 的凸集 $\tilde{S}$ 映射到满足 $S'\supset\tilde{S}'\supset\{\mathbf{X}_i'\}$ 的凸集 $\tilde{S}'$. 而且，如果 H 满足条件(3)，则显然 $H^{-1}$ 也满足. 由此推出凸集 $\tilde{S}$ 和 $\tilde{S}'$ 之间的上述对应是双射. 因为 $\{\mathbf{X}_i\}$(或 $\{\mathbf{X}_i'\}$)的凸包是所有这样的凸集的交，这就证明了 $h$ 保持这些点的凸包. □

保持给定点集的凸包的射影变换组成一个非常重要的类，并被称为**准仿射**变换.

**定义 20.3** 令 $B$ 为 $\mathbb{R}^n$ 的子集而 $h$ 为 $\mathbb{P}^n$ 的一个射影变换. 如果 $h$ 保持集合 $B$ 的凸包，则称射影变换 $h$ 是关于集合 $B$ 的"准仿射".

可以验证如果 $h$ 是关于 $B$ 的准仿射，那么 $h^{-1}$ 是关于 $h(B)$ 的准仿射. 而且，如果 $h$ 关于 $B$ 是准仿射而 $g$ 关于 $h(B)$ 是准仿射，则 $g\circ h$ 关于 $B$ 是准仿射. 因此，准仿射射影变换按这个方式可以复合. 然而，严格地说，关于给定的固定点集合的准仿射射影变换不构成一个群.

我们只打算考虑通过一个射影变换相互对应的点集 $\{\mathbf{X}_i\}$ 和 $\{\mathbf{X}_i'\}$. 当我们提到这个射影变换是**准仿射**的，总意味着与集合 $\{\mathbf{X}_i\}$ 有关.

**二维准仿射映射**

二维准仿射映射是 3D 空间中平面点集与它们在射影摄像机映射下的图像之间的变换，正式地陈述如下.

**定理 20.4 如果 $B$ 是 $\mathbb{R}^3$ 中一个平面("物平面")上的点集，并且 $B$ 完全位于一个射影摄像机的前面，那么由该摄像机定义的从物平面到像平面的映射是关于 $B$ 的准仿射.**

**证明** 存在一个把物平面映射到像平面的射影变换 $h$ 是众所周知的. 这里需要证明的是这个射影是关于 $B$ 的准仿射. 令 $\mathbf{L}$ 是摄像机主平面与物平面的交线. 因为 $B$ 完全在摄像机的前面，所以 $\mathbf{L}$ 与 $B$ 的凸包不相交. 然而，根据主平面的定义 $h(\mathbf{L})=\mathbf{L}_\infty$，其中 $\mathbf{L}_\infty$ 是该像平面上的无穷远直线. 从而，可以推出 $h(\overline{B})\cap\mathbf{L}_\infty=\emptyset$，进而根据定理 20.2 可知这个变换是关于 $B$ 的准仿射. □

注意如果点 $\mathbf{x}_i$ 在一幅图像中可见，则对应的物点必须在摄像机的前面，把定理 20.4 应用于一序列的成像操作(例如，一幅图像的图像的图像等)，其结果是在这个序列的最先和最后图像由一个平面射影变换相关联，该变换是关于物平面上在最后一幅图像上可见的任意点集的准仿射.

类似地，如果两幅图像取自一个平面上的点集 $\{\mathbf{X}_i\}$，$\{\mathbf{x}_i\}$ 和 $\{\mathbf{x}_i'\}$ 是两幅图像上的对应点，那么存在一个准仿射映射(关于 $\mathbf{x}_i$)把每个 $\mathbf{x}_i$ 映射到 $\mathbf{x}_i'$，因此应用定理 20.2 可得下面的结论：

**结论 20.5 如果 $\{\mathbf{x}_i\}$ 和 $\{\mathbf{x}_i'\}$ 是两幅视图中的对应点，它们是一个平面上物体点集合 $\{\mathbf{X}_i\}$ 的影像，那么存在一个表示平面射影变换的矩阵 H，使得 $H\hat{\mathbf{x}}_i=w_i\hat{\mathbf{x}}_i'$，并且全部 $w_i$ 有相同的符号.**

## 20.2 摄像机的前面和后面

一个点 $\mathbf{X}=(X,Y,Z,T)^T$ 相对摄像机的深度在(5.15 - $p$108)中被证明是

$$\mathrm{depth}(\mathbf{X};P)=\frac{\mathrm{sign}(\det M)\,w}{T\,\|\,\mathbf{m}^3\,\|},\tag{20.1}$$

其中 M 是 P 的左边 3×3 子矩阵，$\mathbf{m}^3$ 是 M 的第三行，并且 $P\mathbf{X}=w\hat{\mathbf{x}}$. 这个表示式与 $\mathbf{X}$ 或 M 的具体的齐次表示无关，即乘以一个非零尺度因子它不会改变. 这个深度的定义被用来确定一个点在摄像机的前面与否.

**结论 20.6　点 X 在摄像机 P 前面的充要条件是** $\mathrm{depth}(\mathbf{X};P)>0$.

事实上，摄像机前面的点深度为正，摄像机后面的点为负，无穷远平面上的点为无穷大，而摄像机主平面上的点为零. 如果摄像机中心或点 $\mathbf{X}$ 在无穷远，则深度没有定义.

在这一节我们通常只关心深度的符号而不是它的大小，于是我们可以写为

$$\mathrm{depth}(\mathbf{X};P)\doteq w\,T\det M\tag{20.2}$$

其中记号≐表示符号相等. 量 $\mathrm{sign}(\mathrm{depth}(\mathbf{X};P))$将称为点 $\mathbf{X}$ 相对于摄像机 P 的**正负性**. 一个点如果其正负性从 1 变为 -1，或者从 -1 变为 1，则称正负性被一个变换反置.

## 20.3 三维点集合

本节将解释点关于一个摄像机的正负性与点集的凸包之间的联系. 现在来陈述主要结论.

**定理 20.7　令 $P^E$ 和 $P'^E$ 为两个摄像机. $\mathbf{X}_i^E$ 为位于两个摄像机前面的点集合. 而$\{\mathbf{x}_i\}$和$\{\mathbf{x}_i'\}$为对应的图像点.（上标 E 表示欧式）**

**(1)令$(P,P',\{\mathbf{X}_i\})$是从图像对应 $\mathbf{x}_i\leftrightarrow\mathbf{x}_i'$得到的任何射影重构，并且 $P\mathbf{X}_i=w\hat{\mathbf{x}}_i$ 和 $P'\mathbf{X}_i=w_i'\hat{\mathbf{x}}_i'$，则对于所有 $i$，$w_iw_i'$的符号相同.**

**(2)如果每个 $\mathbf{X}_i$ 都为有限点，以及对所有 $i$，有 $P\hat{\mathbf{X}}_i=w_i\hat{\mathbf{x}}_i$ 且 $w_i$ 有相同符号，则存在一个准仿射变换 H 把每个 $\mathbf{X}_i$ 变到 $\mathbf{X}_i^E$.**

根据定理 9.1($p$181)，一定存在一个把每个 $\mathbf{X}_i$ 映到 $\mathbf{X}_i^E$ 的射影变换. 当前的定理给出了额外的信息，即该变换是准仿射，从而存在一个准仿射重构.

注意所有 $w_iw_i'$的符号相同的条件不受 P，P′或任何点 $\mathbf{X}_i$ 乘以一个常数因子的影响，从而不因这些量的齐次表示的选择而改变. 具体地说，如果 P 乘了一个负常数，那么对所有的 $i$，$w_i$ 也乘了一个负常数. 因此对每个 $i$，$w_iw_i'$的符号反置了，但是它们仍保持相同的符号. 类似地，如果一个点 $\mathbf{X}_i$ 乘了一个负常数，那么 $w_i$ 和 $w_i'$两者都改变了符号，因而 $w_iw_i'$的符号没有改变. 同样，对所有 $i$，每个 $w_i$(在定理的第(2)部分)符号相同的条件也不受摄像机矩阵乘以一个负(当然或者正的)常数的影响.

**证明**　点 $\mathbf{X}_i^E$ 在摄像机 $P^E$ 和 $P'^E$ 的前面，从而相对于这些摄像机有正的深度. 根据(20.2)

$$\mathrm{depth}(\mathbf{X}_i^E;\,P^E)\doteq\det(M^E)\,w_i\,T_i^E.$$

因此，对所有 $i$，$\det(M^E)w_i T_i^E>0$. 类似地，对于第二个摄像机有 $\det(M'^E)w_i' T_i^E>0$. 把这些表达式乘在一起，并把 $T_i^E$ 消去(因为它出现两次)，给出 $w_i w_i' \det(M^E)\det(M'^E)>0$. 由于 $\det(M^E)\det(M'^E)$ 是常数，这证明了 $w_i w_i'$ 符号不变.

以上是根据真实配置来证明的. 然而，注意对任何 H，有 $w_i\hat{\mathbf{x}}_i = P^E\mathbf{X}_i^E = (P^E H^{-1})(H\mathbf{X}_i^E)$，因而 $w_i w_i'$ 对于射影重构 $(P^E H^{-1}, P'^E H^{-1}, \{H\mathbf{X}_i^E\})$ 也符号相同. 因为任何射影重构都是这种形式(除一个齐次常数因子外)，而且 $w_i w_i'$ 符号相同的条件与 $\mathbf{X}_i^E, P^E, P'^E$ 的齐次表示的选择无关，由此推出在任何射影重构中，对所有的 $i$，$w_i w_i'$ 符号相同. 这证明了定理的第一部分.

为证明第二部分，假设在射影重构 $w_i\hat{\mathbf{x}}_i = P\hat{\mathbf{X}}$ 中的 $w_i$ 对所有的 $i$ 符号相同. 由于这是一个射影重构，存在着一个由 H 表示的变换使得对某些常数 $\eta_i$ 和 $\varepsilon$ 有 $H\hat{\mathbf{X}}_i = \eta_i\hat{\mathbf{X}}_i^E$ 和 $PH^{-1} = \varepsilon P^E$. 则

$$\omega_i\hat{\mathbf{x}}_i = P\hat{\mathbf{X}}_i = (PH^{-1})(H\hat{\mathbf{X}}_i) = (\varepsilon P^E)(\eta_i\hat{\mathbf{X}}_i^E)$$

因此对所有 $i$ 有

$$P^E\hat{\mathbf{X}}_i^E = (\omega_i/\varepsilon\eta_i)\hat{\mathbf{x}}_i.$$

然而，由于对所有 $i$ 有 $\mathrm{depth}(\mathbf{X}_i^E, P^E)>0$. 我们得到对所有 $i$ 有 $\det(M^E)w^i/\varepsilon\eta_i>0$. 又由于 $\det(M^E)/\varepsilon$ 是常数，并且根据假设对所有 $i$，$w_i$ 符号相同. 这样一来，对所有的 $i$，$\eta_i$ 符号相同，因此根据定理 20.2，满足 $H\hat{\mathbf{X}}_i = \eta_i\hat{\mathbf{X}}_i^E$ 的映射 H 是关于点 $\hat{\mathbf{X}}_i$ 的一个准仿射.

注意对所有的 $i$，$w_i$ 符号相同的条件只要对其中的一个摄像机验证即可. 根据定理的第(1)部分定义 $P'\hat{\mathbf{X}}_i = w_i'\hat{\mathbf{x}}_i'$，则对所有 $i$，$w_i w_i'$ 符号相同. 因此，如果所有的 $w_i$ 符号相同，则 $w_i'$ 也如此. □

## 20.4 获得一个准仿射重构

根据定理 20.7，任何使得 $P\hat{\mathbf{X}}_i = w_i\hat{\mathbf{x}}_i$ 且对所有 $i$，$w_i$ 符号相同的射影重构是一个准仿射重构. 与一个任意的射影变换相比，一个准仿射重构的优势在于它给出真实物体形状的一个更接近的近似. 正如[Hartley-94a]中所说的，它可以用作景物度量重构过程中的一个垫脚石. 另外我们用它可以恢复物体的凸包或者确定诸如两点是否在一个平面同一边等问题.

结果表明在已知一个射影重构条件下，准仿射重构是非常简单的，正如下面的定理所示

**定理 20.8 其中一个摄像机是仿射摄像机的任何射影重构是一个准仿射重构.**

**证明** 一个仿射摄像机是最后一行形如(0,0,0,1)的摄像机. 对于这种情形，记 $w_i\hat{\mathbf{x}}_i = P\hat{\mathbf{X}}_i$ 可以立即验证对所有 $i$，都有 $w_i=1$，因而它们都有相同的符号. 根据定理 20.2，这意味着这个重构与真实的景物相差一个准仿射变换. □

由它立即得出下面的结论

**推论 20.9 若 $(P, P', \{\mathbf{X}_i\})$ 是一个景物的射影重构，其中 $P=[I|\mathbf{0}]$. 则通过交换 P 和 P′ 的最后两列，以及每个 $\mathbf{X}_i$ 的最后两个坐标会得到景物的一个准仿射重构.**

这类似于结论 9.4($p$185),在那里证明了如果已知摄像机 P 实际上是仿射摄像机,则上面的步骤提供了一个仿射重构.

## 20.5 变换正负性的效果

我们来推导与(20.2)中定义的深度公式稍微不同的一种形式. 令 P 是一个摄像机矩阵. P 的中心是使 $\mathbf{PC}=\mathbf{0}$ 的唯一的点 $\mathbf{C}$. 则可写出 $\mathbf{C}$ 的显式公式如下.

**定义 20.10** 给定一个摄像机矩阵 P,我们定义 $\mathbf{C}_{\mathrm{P}}^{\mathrm{T}}$ 为矢量$(c_1,c_2,c_3,c_4)$,其中

$$c_i=(-1)^i\det\hat{\mathrm{P}}^{(i)}.$$

而 $\hat{\mathrm{P}}^{(i)}$ 是 P 去掉第 $i$ 列所得到的矩阵.

我们用$[\mathrm{P}/\mathbf{V}^{\mathrm{T}}]$记由 $3\times4$ 的摄像机矩阵 P 增加最后一行 $\mathbf{V}^{\mathrm{T}}$ 而组成的 $4\times4$ 矩阵. 由定义 20.10 可导出 $\det[\mathrm{P}/\mathbf{V}^{\mathrm{T}}]$的一个简单公式. 按最后一行的余子式展开这个行列式给出对任何行矢量 $\mathbf{V}^{\mathrm{T}}$ 有 $\det[\mathrm{P}/\mathbf{V}^{\mathrm{T}}]=\mathbf{V}^{\mathrm{T}}\mathbf{C}_{\mathrm{P}}$ 成立. 作为它的一个特例,如果 $\mathbf{p}_i^{\mathrm{T}}$ 是 P 的第 $i$ 行,则

$$\mathbf{p}_i^{\mathrm{T}}\mathbf{C}_{\mathrm{P}}=\det[\mathrm{P}/\mathbf{p}_i^{\mathrm{T}}]=0.$$

上面的等式成立是因为这个矩阵有重复的行. 由于该式对所有 $i$ 都成立,从而推出 $\mathbf{PC}_{\mathrm{P}}=\mathbf{0}$,因此 $\mathbf{C}_{\mathrm{P}}$ 是摄像机的中心,正是上面所断言的.

注意子矩阵 $\hat{\mathrm{P}}^{(4)}$ 与分解式 $\mathrm{P}=[\mathrm{M}|\mathbf{v}]$中的矩阵 M 一样,因此有 $\det\mathrm{M}=c_4$,这允许我们重新公式化(20.2)如下:

$$\mathrm{depth}(\mathbf{X};\mathrm{P})\doteq w(\mathbf{E}_4^{\mathrm{T}}\mathbf{X})(\mathbf{E}_4^{\mathrm{T}}\mathbf{C}_{\mathrm{P}}). \tag{20.3}$$

其中 $\mathbf{E}_4^{\mathrm{T}}$ 是矢量(0,0,0,1). 这里值得注意的是 $\mathbf{E}_4$ 是表示无穷远平面的矢量——点 $\mathbf{X}$ 在无穷远平面上的充要条件是 $\mathbf{E}_4^{\mathrm{T}}\mathbf{X}=0$.

我们现在考虑由矩阵 H 表示的一个射影变换. 如果 $\mathrm{P}'=\mathrm{PH}^{-1}$ 和 $\mathbf{X}'=\mathrm{H}\mathbf{X}$,则图像对应被这个变换保持. 当说到一个射影变换被应用到点集和一个摄像机时其含意是指点 $\mathbf{X}$ 被变换到 $\mathrm{H}\mathbf{X}$ 且摄像机矩阵被变换到 $\mathrm{PH}^{-1}$.

本节将考虑这样的射影变换及其在点相对于一个摄像机的正负性上的影响. 首先,我们希望确定当 P 变换到 $\mathrm{PH}^{-1}$时,$\mathbf{C}_{\mathrm{P}}$ 会发生什么变化. 为回答这个问题,考虑任何 4 维矢量 $\mathbf{V}$. 我们有

$$\mathbf{V}^{\mathrm{T}}\mathrm{H}^{-1}\mathbf{C}_{\mathrm{PH}^{-1}}=\det(\mathrm{PH}^{-1}/\mathbf{V}^{\mathrm{T}}\mathrm{H}^{-1})=\det[\mathrm{P}/\mathbf{V}^{\mathrm{T}}]\det\mathrm{H}^{-1}=\mathbf{V}^{\mathrm{T}}\mathbf{C}_{\mathrm{P}}\det\mathrm{H}^{-1}.$$

因为上式对所有矢量 $\mathbf{V}$ 成立,故由它推出 $\mathrm{H}^{-1}\mathbf{C}_{\mathrm{PH}^{-1}}=\mathbf{C}_{\mathrm{P}}\det\mathrm{H}^{-1}$,或者

$$\mathbf{C}_{\mathrm{PH}^{-1}}=\mathrm{H}\mathbf{C}_{\mathrm{P}}\det\mathrm{H}^{-1}. \tag{20.4}$$

在一个层次上,这个公式表明变换 H 把摄像机中心 $\mathbf{C}=\mathbf{C}_{\mathrm{P}}$ 变到新位置 $\mathbf{C}_{\mathrm{PH}^{-1}}\approx\mathrm{H}\mathbf{C}$ 上. 然而,我们对 $\mathbf{C}_{\mathrm{PH}^{-1}}$ 的准确坐标,特别是出现在(20.3)中最后一个坐标 $c_4$ 的符号感兴趣,这样一来,因子 $\det\mathrm{H}^{-1}$是主要的.

现在,把(20.4)应用于(20.3)给出

$$\begin{aligned}\mathrm{depth}(\mathrm{H}\mathbf{X};\mathrm{PH}^{-1})&\doteq w(\mathbf{E}_4^{\mathrm{T}}\mathrm{H}\mathbf{X})(\mathbf{E}_4^{\mathrm{T}}\mathbf{C}_{\mathrm{PH}^{-1}})\\&\doteq w(\mathbf{E}_4^{\mathrm{T}}\mathrm{H}\mathbf{X})(\mathbf{E}_4^{\mathrm{T}}\mathrm{H}\mathbf{C}_{\mathrm{P}})\det\mathrm{H}^{-1}\end{aligned}$$

可以把表达式 $\mathbf{E}_4^{\mathrm{T}}\mathrm{H}$ 解释为被 H 映射到无穷远的平面 $\boldsymbol{\pi}_\infty$. 这是因为点 $\mathbf{X}$ 在 $\boldsymbol{\pi}_\infty$ 上的充要条件是 $\mathrm{H}\mathbf{X}$ 的最后一个坐标是零,即 $\mathbf{E}_4^{\mathrm{T}}\mathrm{H}\mathbf{X}=0$. 另一方面,$\mathbf{X}$ 在 $\boldsymbol{\pi}_\infty$ 上的充要条件是 $\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{X}=0$. 最后,把变换矩阵 H 的第 4 行记为 $\mathbf{h}_4^{\mathrm{T}}$,并把 sign(detH)记为 $\delta$,我们得到

**结论 20.11　如果 $\boldsymbol{\pi}_\infty$ 是被一个射影变换 H 映射到无穷远的平面并且 $\delta=\mathrm{sign}(\det\mathrm{H})$,则**

$$\mathrm{depth}(\mathrm{H}\mathbf{X};\mathrm{P}\mathrm{H}^{-1})\doteq w(\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{X})(\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{C}_{\mathrm{p}})\delta$$

这个方程将被广泛地应用,它可被视为(20.3)的一种推广. 在下一节中我们将看到 $\delta=\mathrm{sign}(\det\mathrm{H})$是判断 H 是反向或者保向变换的一个指标. 这样一来,变换 H 对正负性的影响仅由被映射到无穷远平面 $\boldsymbol{\pi}_\infty$ 的位置以及 H 是保向或反向来确定.

我们现在考虑不同的变换在点相对于一个摄像机的正负性上的影响. 首先考虑的是仿射变换的影响.

**结论 20.12　具有正行列式的仿射变换保持任何点相对一个摄像机的正负性. 具有负行列式的仿射变换则使正负性反向.**

**证明**　仿射变换保持无穷远平面不变,因此 $\boldsymbol{\pi}_\infty=\mathbf{E}_4$. 然后比较(20.3)和结论 20.11 即得出此结果. □

我们现在来确定一个任意的射影变换如何影响正负性.

**结论 20.13　令 H 表示具有正行列式的一个射影变换,并且 $\boldsymbol{\pi}_\infty$ 是被 H 映射到无穷远的空间平面. 点 $\mathbf{X}$ 的正负性被 H 保持的充要条件是 $\mathbf{X}$ 与摄像机中心在平面 $\boldsymbol{\pi}_\infty$ 的同一边.**

**证明**　由于 $\det\mathrm{H}>0$,从(20.3)和结论 20.11 得出 $\mathrm{depth}(\mathbf{X};\mathrm{P})\doteq\mathrm{depth}(\mathrm{H}\mathbf{X};\mathrm{P}\mathrm{H}^{-1})$的充要条件是$(\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{X})(\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{C})\doteq(\mathbf{E}_4^{\mathrm{T}}\mathbf{X})(\mathbf{E}_4^{\mathrm{T}}\mathbf{C})$. 假设点 $\mathbf{X}$ 和摄像机 P 位于有限点使得其正负性是良定的,并且通过变尺度使得 $\mathbf{X}$ 和 $\mathbf{C}$ 的最后坐标都等于 1. 在这种情形下,$(\mathbf{E}_4^{\mathrm{T}}\mathbf{X})(\mathbf{E}_4^{\mathrm{T}}\mathbf{C})=1$ 并由此看出保持正负性的充要条件是 $(\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{X})(\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{C})\doteq1$ 或改写为 $\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{X}\doteq\boldsymbol{\pi}_\infty^{\mathrm{T}}\mathbf{C}$. 这个条件可以解释为点 $\mathbf{C}$ 和 $\mathbf{X}$ 都位于平面 $\boldsymbol{\pi}_\infty$ 的同一边. 因此,点 $\mathbf{X}$ 的正负性被变换 H 保持的充要条件是它与摄像机中心在平面 $\boldsymbol{\pi}_\infty$ 的同一边. □

## 20.6　定向

我们现在考虑图像定向问题. 如果一个 $\mathbb{R}^n$ 到它自身的映射 $h$ 的 Jacobian(偏导数矩阵的行列式)在点 $\mathbf{X}$ 是正的,则称 $h$ 在点 $\mathbf{X}$ 处是保向的;如果该 Jacobian 值是负的,则称 $h$ 为反向的. $\mathbb{R}^n$ 中的点相对于一个超平面的反射(镜面影像)是反向映射的一个例子. $\mathbb{R}^n$ 到自身的一个射影变换 $h$ 限制为 $\mathbb{R}^n-\boldsymbol{\pi}_\infty$ 到 $\mathbb{R}^n$ 的一个映射,其中 $\boldsymbol{\pi}_\infty$ 是由 H 映射到无穷远的超平面(直线,平面). 考虑 $n=3$ 的情形并且令 H 是表示射影变换 $h$ 的一个 $4\times4$ 矩阵. 我们希望确定在 $\mathbb{R}^n-\boldsymbol{\pi}_\infty$ 中,映射 $h$ 在哪些点 $\mathbf{X}$ 上是保向的. 可以验证(利用 Mathematica [Mathematica-92]非常容易):如果 $\mathrm{H}\hat{\mathbf{X}}=w\hat{\mathbf{X}}'$ 而 J 是 $h$ 的 Jacobian 矩阵在 $\mathbf{X}$ 处的值,则 $\det(\mathrm{J})=\det(\mathrm{H})/w^4$. 由此推出下面的结论.

**结论 20.14　由矩阵 H 表示的 $\mathbb{P}^3$ 中的一个射影变换 $h$ 在 $\mathbb{R}^n-\boldsymbol{\pi}_\infty$ 中的任何点上是保向的充要条件是 $\det(\mathrm{H})>0$.**

当然,可定向的概念可以扩展到整个$\mathbb{P}^3$上,并且可以证明 $h$ 在整个$\mathbb{P}^3$上是保向的充要条件是 $\det(H)>0$. 这里最基本的特征是作为一个拓扑流形,$\mathbb{P}^3$是可定向的.

通过准仿射变换对应的两个点集$\{\mathbf{X}_i\}$和$\{\overline{\mathbf{X}}_i\}$称为**反向**,如果该变换是反向的. 作为一个例子,考虑由对角矩阵 $H=\mathrm{diag}(1,1,-1,1)$给出的变换. 该变换的行列式为负,因而是反向的. 另一方面,它是仿射的,因此是准仿射的. 所以总是能够构造景物的一个反向的准仿射重构. 因此看上去好像景物的方向不能从一对图像来确定. 虽然有时这是对的,但有时有可能排除场景的另一方向准仿射重构,从而确定景物真正的定向.

这里,常识提供一些线索. 具体地说是一副双眼可以看作给一个眼睛提供一幅图像而对另一个眼睛提供另一幅图像. 如果送得正确,则头脑感知景物的一个 3D 重构. 然而如果把这两幅图像交换并且提供给相对的眼睛. 则透视将被反过来——山峰变成山谷而反之也然. 实际上,头脑是能够计算图像对的两个相反定向的重构. 因此,似乎在某种环境下一个图像对的两种相反定向的实现是存在的. 这在图 20.2 中说明.

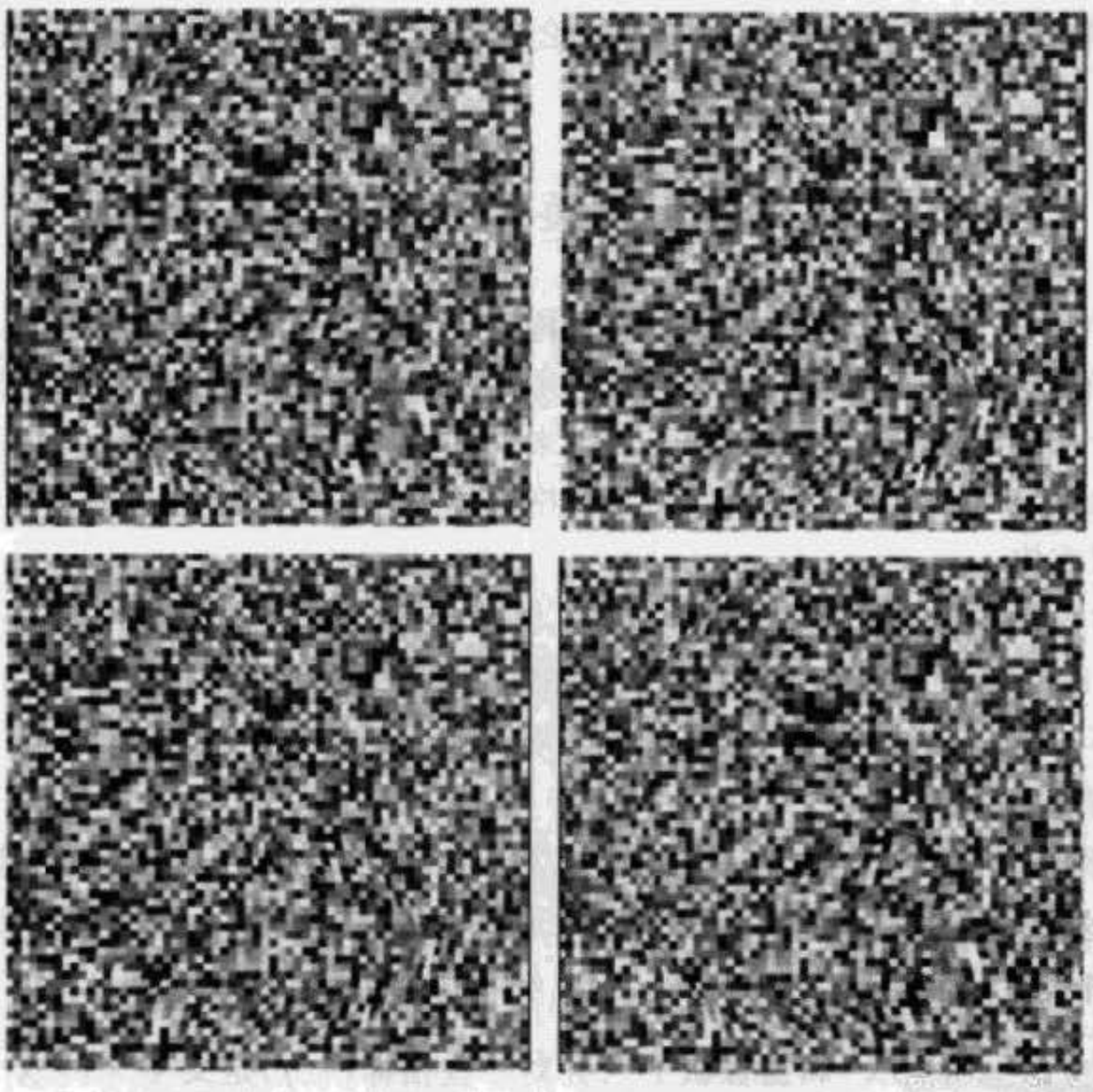


图 20.2 用交叉融合看见的双眼图像对(双眼交叉让左眼看右图而右眼看左图). 底部的两幅图像与上部的一对图像是一样的,只是它们的位置被交换. 在上部的图像对中,我们看到一个 L 型的区域浮在平面背景上. 在下部图像对中,L 型区域看上去是一个凹痕. 两个"重构"的区别是对背景平面作了反射. 它表明同样一对图像可以给出两种不同定向的射影重构.

但是如下面定理所证明的:这种情况不总是存在. 如果重构的 3D 点 $\mathbf{X}_i$ 在所有摄像机的前面,那么该点对应集合的射影实现被称为**强实现**,这个定义将在下面的定理和本章其他地方用到.

**定理 20.15** **令$(P,P',\{\mathbf{X}_i\})$是点对应的唯一可实现集合的一个强实现. 存在另一个反向的强实现$(\overline{P},\overline{P}',\{\overline{\mathbf{X}}_i\})$的充要条件是在$\mathbb{R}^3$中存在着一个平面使得两个摄像机 P 和 P′的透视中心在这个平面的一边,而点 $\mathbf{X}_i$ 在另一边.**

**证明** 考虑该结构的一个强实现. 根据定义,所有的点都在两个摄像机的前面. 假设存在一个平面把两摄像机中心和这些点分开. 令 G 是把给定平面映射到无穷远的一个射影变换,而 A 是一个仿射变换. 进一步假设 $\det G>0$ 和 $\det A<0$. 令 H 是复合变换 $H=AG$. 根据结论 20.13,变换 G 对这些点来说是反置正负性的,因为这些点和两摄像机中心各在这个平面的一边. 根据结论 20.12,A 也是反置正负性的,因为 $\det A<0$. 因此复合变换 H 一定是保持正负性的,并且把一个强结构变换到另一个强结构. 然而因为 H 的行列式为负,它是反向的,因此这个两个强实现有相反的定向.

反之,假设存在两个反向的强实现,并且令 H 是把一个映射到另一个的变换. 因为 H 是反向的,故 $\det H<0$. 由定义,相对于两个摄像机,映射 H 在所有点上是保持正负性的. 如果 $\boldsymbol{\pi}_\infty$ 是被 H 映射到无穷远的平面,则根据结论 20.13,点 $\mathbf{X}$ 必须在平面 $\boldsymbol{\pi}_\infty$ 不同于摄像机中心的另一边. □

## 20.7 正负性不等式

节 20.4 给出了一种可以直接地从射影重构得到景物的一个准仿射重构的非常简单的方法. 然而,在那里得到的重构没有注意点必须在所有摄像机前面的条件. 事实上,在这个构造中第一个摄像机是一个仿射摄像机,对它来说,前面和后面不是良定的. 充分利用可见点必须在摄像机前面这个事实的好处,可更严格地约束这个重构以导致更加逼近景物的一个真正的仿射重构.

我们将给出由多幅图像推导重构的方法. 给定一组图像点$\{\mathbf{x}_i^j\}$,其中 $\mathbf{x}_i^j$ 是第 $i$ 个点在第 $j$ 幅图像上的投影. 不是所有的点在每一幅图像上都可见,因此对某些$(i,j)$,点 $\mathbf{x}_i^j$ 没有给出,此时,不知道第 $i$ 个点是否在第 $j$ 个摄像机前面. 另一方面,一个像点 $\mathbf{x}_i^j$ 的存在表示这点是在摄像机的前面.

开始时,我们假设已知景物的一个射影重构,它包含一组 3D 点 $\mathbf{X}_i$ 和摄像机 $P^j$ 使得 $\mathbf{x}_i^j\approx P^j\mathbf{X}_i$. 把这个方程中隐含的常数因子显式地给出得到 $w_i^j\hat{\mathbf{x}}_i^j=P^j\mathbf{X}_i$. 该方程中的 $P^j$ 和 $\mathbf{X}_i$ 分别是矩阵或矢量的任意选择的齐次表示. 结合定理 20.7,对多视图可以陈述为:

**结论 20.16** **考虑一组点 $\mathbf{X}_i^E$ 和摄像机 $P^{jE}$,并对一些指标$(i,j)$定义 $\mathbf{x}_i^j=P^{jE}\mathbf{X}_i^E$ 为点 $\mathbf{X}_i^E$ 在摄像机 $P^{jE}$ 前面. 令$(P^j;\mathbf{X}_i)$是由 $\mathbf{x}_i^j$ 得到的射影重构. 则存在摄像机矩阵 $\widetilde{P}^j=\pm P^j$ 和 $\widetilde{\mathbf{X}}^i=\pm\mathbf{X}_i$ 使得对每个 $\mathbf{x}_i^j$ 有定义的$(i,j)$有**

$$\widetilde{P}^j\widetilde{\mathbf{X}}_i=w_i^j\hat{\mathbf{x}}_i^j \quad \text{并且} \quad w_i^j>0$$

简要地说,我们总可以在必要时用 $-1$ 乘以摄像机矩阵和点来调整射影重构,使得只要图像点 $\mathbf{x}_i^j$ 存在,$w_i^j$ 就是正的. 该证明因简单而略去. 为寻求矩阵 $\widetilde{P}^j$ 和点 $\widetilde{\mathbf{X}}_i$,可以假设其中一个摄像机 $\widetilde{P}_1=P_1$,不然的话所有点和摄像机都乘以 $-1$. 对所有使得 $\mathbf{x}_i^1$ 有定义的 $i$,$P_1\widetilde{\mathbf{X}}_i=w_i^1\mathbf{x}_i^1$ 且 $w_i^1>0$ 的条件决定是选择 $\widetilde{\mathbf{X}}_i=\mathbf{X}_i$ 还是 $-\mathbf{X}_i$. 为了使得 $\mathbf{x}_i^j$ 有定义,对所有的 $j$,由每

个已知的 $\tilde{\mathbf{X}}_i$ 确定 $\tilde{\mathrm{P}}^j$. 依此类推,我们可以容易地求出需要用于 $\mathrm{P}^j$ 和 $\mathbf{X}_i$ 以便得到 $\tilde{\mathrm{P}}^j$ 和 $\tilde{\mathbf{X}}_i$ 的因子 $\pm 1$. 假设这个过程已完成,并用每个 $\tilde{\mathrm{P}}^j$ 代替 $\mathrm{P}^j$ 和 $\tilde{\mathbf{X}}_i$ 代替 $\mathbf{X}_i$. 下面,我们将省去波浪号并且对修正过的 $\mathrm{P}^j$ 和 $\mathbf{X}_i$ 继续进行运算. 目前我们知道只要图像点 $\mathbf{x}_i^j$ 给定便有 $w_i^j > 0$.

现在,我们来求变换 H,它把射影重构变换到一个准仿射重构并且按要求使所有的点在摄像机的前面. 用 4 维矢量 $\mathbf{v}$ 表示被 H 映射到无穷远的平面 $\boldsymbol{\pi}_\infty$,这个条件可以写为(见结论 20.11):

$$\mathrm{depth}(\mathbf{X}_i;\mathrm{P}^j) \doteq (\mathbf{v}^\mathsf{T}\mathbf{X}_i)(\mathbf{v}^\mathsf{T}\mathbf{C})\delta > 0$$

其中 $\delta = \mathrm{sign}(\det \mathrm{H})$. 这个条件对给定 $\mathbf{x}_i^j$ 的所有 $(i,j)$ 成立.

因为如果必要我们可以自由地用 $-1$ 乘以矢量 $\mathbf{v}$,故我们可以假设摄像机 $\mathrm{P}^1$ 的中心满足 $(\mathbf{v}^\mathsf{T}\mathbf{C}^1)\delta > 0$. 现在可以不困难地推出下面的不等式

$$\begin{aligned} &\mathbf{X}_i^\mathsf{T}\mathbf{v} > 0 \text{ 对所有 } i \text{ 成立} \\ &\delta \mathbf{C}'^\mathsf{T}\mathbf{v} > 0 \text{ 对所有 } j \text{ 成立} \end{aligned} \tag{20.5}$$

这些方程(20.5)可以称为**正负性不等式**. 由于每一个 $\mathbf{X}_i$、$\mathbf{C}$ 和 $\mathbf{C}'$ 的值都是知道的,它们组成了关于 $\mathbf{v}$ 的元素的不等式组. $\delta$ 的值**预先**不知道,因此有必要对 $\delta = 1$ 和 $\delta = -1$ 两种情形求解.

为求所需要的变换 H,我们首先对 $\delta = 1$ 或 $\delta = -1$ 解正负性不等式来求 $\mathbf{v}$ 值. 所求的矩阵 H 是任何以 $\mathbf{v}^\mathsf{T}$ 作为其最后一行并满足条件 $\det \mathrm{H} \doteq \delta$ 的矩阵. 如果 $\mathbf{v}$ 的最后一个元素非零,则 H 能选择简单的形式,即其前 3 行的形式是 $\pm[\mathrm{I}|\mathbf{0}]$.

如果一个欧式重构(或更具体地说,一个准仿射重构)是可能的,则必然对 $\delta = 1$ 或者 $\delta = -1$ 存在解. 在某些情形对 $\delta = 1$ 和 $\delta = -1$,该正负性不等式组都存在解. 这意味着存在两个反向的强实现. 可能发生这种情形的条件在节 20.6 中讨论过.

**求解正负性不等式**

自然,正负性不等式可以用线性规划的技术来求解. 然而如它们所表示的,如果 $\mathbf{v}$ 是一个解,则对任何正因子 $\alpha$,$\alpha\mathbf{v}$ 也是解. 为了限制在有界解的范围内,可以加进另外的不等式. 例如,如果 $\mathbf{v} = (v_1, v_2, v_3, v_4)^\mathsf{T}$,则不等式 $-1 < v_i < 1$ 可以用来把解的范围限制在有界的多面体内.

为得到唯一解,我们需要指定一个要线性化的目标函数. 一个合适的策略是寻求最大化每个不等式要满足的范围. 为做到这一点,我们引入另一个变量 $d$. 形式(20.5)中每一个不等式 $\mathbf{a}^\mathsf{T}\mathbf{v}$ 都用一个不等式 $\mathbf{a}^\mathsf{T}\mathbf{v} > d$ 代替. 我们求满足所有这些不等式的最大化的 $d$. 这是一个标准的线性规划问题. 对它存在许多求解方法,比如单纯形法([Press-88]①. 如果对某个 $d > 0$ 找到一个解,那么它是一个所希望的解.

**算法的总结**

现在,我们给出采用正负性不等式来计算景物的准仿射重构的完整算法. 这个算法在上文概述时仅对两视图情形作了讨论. 现在将给出任意视图数的算法. 对多视图的推广是直接的.

① 在[Press-88]给出的单纯形算法不适合作为标准使用,因为它做了所有变量为非负的不必要的假定. 用于此问题时它需要做修正.

目标

给定一个由 3D 点 $\mathbf{X}_i$ 和摄像机矩阵 $P^j$ 组成的集合，它们组成由图像点集得到的一个射影重构，计算一个把该射影重构变换到准仿射重构的射影变换 H.

算法

(1)对每一对$(i,j)$，当点 $\mathbf{x}_i^j$ 给定时令 $P^j\mathbf{X}_i = w_i^j\hat{\mathbf{x}}_i^j$.

(2)用 $-P^j$ 代替某些摄像机 $P^j$，和用 $-\mathbf{X}_i$ 代替某些点 $\mathbf{X}_i$，以便能按要求保证每个 $w_i^j>0$.

(3)形成正负性不等式(20.5)，其中 $\mathbf{C}^j = \mathbf{C}_{P^j}$ 由定义 20.10 定义.

(4)对 $\delta=\pm1$ 的每个值，为这组正负性不等式选择一个解(如果存在). 令该解为 $\mathbf{v}_\delta$. 至少对于 $\delta$ 的一个值，解一定存在，有时对 $\delta$ 的两个值解都存在.

(5)定义一个矩阵 $H_\delta$，它的最后一行等于 $\mathbf{v}_\delta$ 并满足 $\det(H)\doteq\delta$. 矩阵 $H_\delta$ 就是所要求的变换矩阵. 如果 $H_+$ 和 $H_-$ 都存在，那么它们导出两个反向的准仿射重构.

算法 20.1 计算一个准仿射重构.

**界定无穷远平面**

当然一个准仿射重构不是唯一的，它被确定到相差一个关于点和摄像机中心的准仿射变换. 然而，一旦求得了一个，便可以界定无穷远平面坐标的范围. 因此，令 $P^j$ 和 $\mathbf{X}_i$ 组成一个景物的准仿射重构. 可以通过选择 $P^j$ 和 $\mathbf{X}_i$ 的符号使得 $\mathbf{X}_i$ 的最后一个坐标和每个 $M^j$ 的行列式为正. 可以对点和摄像机进行平移使得坐标原点在这些点和摄像机中心的凸包中. 为简单起见，可令这些点的形心放置在原点.

可以施加另一个准仿射变换 H 得到另一个重构. 令 $\boldsymbol{\pi}_\infty$ 是被 H 映到无穷远的平面. 我们的兴趣限于保向的变换，且希望寻找使 H 是准仿射时有关 $\boldsymbol{\pi}_\infty$ 坐标的约束. 有这种性质的平面 $\boldsymbol{\pi}_\infty$ 的充要条件是它整个位于点和摄像机中心的凸包之外. 由于这个平面 $\boldsymbol{\pi}_\infty$ 不穿过凸包，它不可能通过原点. 用矢量 $\mathbf{v}$ 表示 $\boldsymbol{\pi}_\infty$，那么 $\mathbf{v}$ 的最后一个坐标非零. 因此可写为 $\mathbf{v}=(v_1,v_2,v_3,1)^\mathsf{T}$. 因为原点与所有的点在这个平面的同一边，正负性不等式变成

$$\mathbf{X}_i^\mathsf{T}\mathbf{v}>0 \text{ 对所有 } i \text{ 成立}$$
$$\mathbf{C}^{j\mathsf{T}}\mathbf{v}>0 \text{ 对所有 } j \text{ 成立} \tag{20.6}$$

其中 $\mathbf{v}=(v_1,v_2,v_3,1)^\mathsf{T}$. 通过求解在这些约束下最大化 $v_i$ 或 $-v_i$ 的线性规划问题可以找到每一个 $v_i$ 的上界和下界. 没有一个 $v_i$ 可以是无界的，因为否则的话，由矢量 $\mathbf{v}$ 表示的平面 $\boldsymbol{\pi}_\infty$ 可以任意地接近原点.

在求解该系统之前，好的实际做法是用一个仿射变换去归一化这组点和摄像机中心，使得它们的形心在原点并且它们的主矩都等于 1

完整的计算无穷远平面位置的界限的算法在算法 20.2 中给出.

## 20.8 哪些点在第三幅视图中可见

考虑由两视图得到的一个景物重构. 现在来考虑哪些点在第三幅视图中可见的问题. 这样的问题会产生在给定两幅未标定的景物视图并要求合成第三幅视图的时候. 它可由前两幅视图进行景物的射影重构而后投影到第三幅视图来完成. 此时，非常重要的事是要确定一个点是否是在第三个摄像机的前面，即确定它可见或不可见.

目标

给定景物的准仿射重构，确立无穷远平面坐标的界限.

算法

(1)归一化点 $\mathbf{X}_i=(X_i,Y_i,Z_i,T_i)^T$ 使得 $T_i=1$，并归一化摄像机 $P^j=[M_j\mid \mathbf{t}^j]$ 使得 $\det M^j=1$.

(2)进一步作如下的归一化：用 $H^{-1}\mathbf{X}_i$ 替代 $\mathbf{X}_i$，用 $P^jH$ 替代 $P^j$，其中 H 是一个仿射变换，它把形心移动到原点并在主轴方向上进行缩放使得主轴相等.

(3) 令 $\mathbf{v}=(v_1,v_2,v_3,1)^T$，形成正负性不等式(20.6). 任何保向并把该重构映射到图像的一个仿射重构的变换 H 必然有形式

$$H=\begin{bmatrix} I & & & \mathbf{0} \\ v_1 & v_2 & v_3 & 1 \end{bmatrix}$$

其中矢量 $\mathbf{v}$ 满足这些正负性不等式.

(4)每个 $v_i$ 的上下界可以通过运行 6 次线性规划问题而求到. 所希望的变换 H 的坐标必然在这些界所定义的长方体内.

算法 20.2　确立无穷远平面的界限.

如果只给出了第三幅视图相对于某给定重构的参考坐标系下的摄像机矩阵，则不可能确定点在真实景物中是在第三个摄像机的前面还是后面. 其基本的多义性在图 20.3 中说明. 然而，如下面的结论所指出的，知道一个点在第三幅视图中可见的信息可以用来消除这个多义性. 把定理 20.7($p$367)应用于第一和第三幅视图时得到下面的准则.

**结论 20.17　令点$(P^1,P^2,\{\mathbf{X}_i\})$是一组对应 $\mathbf{x}_i^1\leftrightarrow\mathbf{x}_i^2$ 的一个实现. 令 $P^3$ 是第三幅视图的摄像机矩阵且假设 $w_j^i\hat{\mathbf{x}}_i=P^i\mathbf{X}_j$，$i=1,\ldots3$，则 $w_j^1w_j^3$ 对所有在第三幅视图可见的点 $\mathbf{X}_j$ 都有相同的符号.**

在实际中，常常是至少知道一个点 $\mathbf{X}_0$ 在第三幅视图中可见. 用它来定义 $w_0^1w_0^3$ 的符号后，任何其他点 $\mathbf{X}_j$ 在摄像机 $P^3$ 的前面的充要条件是 $w_j^1w_j^3=w_0^1w_0^3$.

我们来举一个例子，如果一旦利用两视图的一个射影重构已经作出，第三个摄像机的摄像机矩阵可以从已知在它前面的六个或更多点的像确定下来，方法是对给定的对应 $\mathbf{x}_i^3=P^3\mathbf{X}_i$(其中点 $\mathbf{X}_i$ 是重构的点)直接求解摄像机矩阵 $P^3$. 然后我们可以无多义性地确定其余的哪些点在 $P^3$ 的前面.

## 20.9　谁在前面

当我们试图用两个或更多未标定视图重构的景物来合成一幅新视图时，有时需要考虑点被其他点遮挡的可能性. 这样就导致如下问题：给定投射到新视图上同一点的两个点，哪个点更靠近摄像机并遮挡了另一点？在可能出现相反定向的准仿射重构时，仍可能还无法确定一对点中的哪一点更靠近新的摄像机. 这在图 20.4 中说明. 如果存在一张平面，它把点集与摄像机中心分开，则存在两个相反定向的重构，从而不能确定谁在前面. 在图 20.4 显示的这个多义性类型只能发生于如下的情形中，即有一张平面 $\boldsymbol{\pi}_\infty$ 把所有可见点的集合与

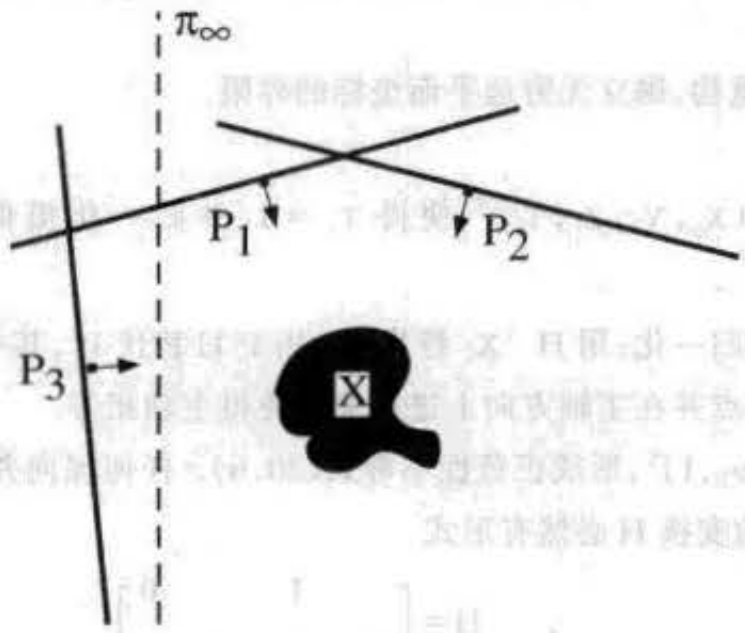



图 20.3　如图所示，点集 **X** 在所有 3 个摄像机的前面．但是若用一个保向的射影变换 H 把 $\pi_\infty$ 映射到无穷远，那么点集将在摄像机 $P^1$ 和 $P^2$ 的前面而在摄像机 $P^3$ 的后面．因此假设点集 **X** 是由摄像机 $P^1$ 和 $P^2$ 所获取的图像重构而得，并且令 $P^3$ 为任何其他的摄像机矩阵．如果存在一个平面把摄像机 $P^3$ 的中心与其他摄像机中心分开且不与点集 **X** 的凸包相交，则不能确定点是否在 $P^3$ 的前面．

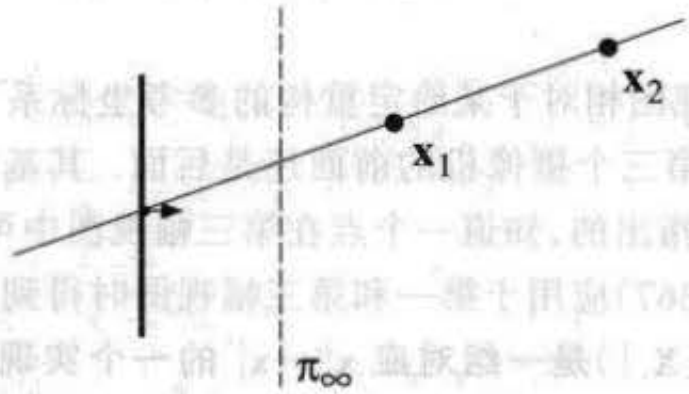



图 20.4　如图所示，点 $\mathbf{x}_1$ 比 $\mathbf{x}_2$ 更靠近摄像机．但是如果用一个反向的射影变换，把 $\pi_\infty$ 映射到无穷远，那么 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 仍然在摄像机前面，但是 $\mathbf{x}_2$ 比 $\mathbf{x}_1$ 更靠近摄像机．

摄像机中心分开．如果不是这种情形，则我们能算得一个准仿射重构而使这个问题容易求解．然而，为了避免计算准仿射重构的开销，我们倾向于只利用景物的射影重构来解这个问题．下文将解释如何来做这件事．

我们可以求(20.1—$p$367)的逆，得到 $\mathrm{depth}^{-1}(\mathbf{X};P)=1/\mathrm{depth}(\mathbf{X};P)$的一个表达式．这个逆深度函数在摄像机主平面上是无穷大，在无穷远平面上是零，在摄像机前面的点为正而在摄像机后面的点为负．为了记号简单起见，我们用 $\chi(\mathbf{X};P)$ 代替 $\mathrm{depth}^{-1}(\mathbf{X};P)$．

对通过摄像机中心的射线上的点 **X**，$\chi(\mathbf{X};P)$的值沿着这条射线单调地减小，它在摄像机中心处为正无穷大，通过正值减小到在无穷远平面上时为零，然后再继续减小，通过负值减小到在摄像机中心时为 $-\infty$．一个点$\mathbf{X}_1$ 比 $\mathbf{X}_2$更靠近摄像机前面 的充要条件是 $\chi(\mathbf{X}_1)>\chi(\mathbf{X}_2)$．这在图 20.5 中说明．

现在，如果该配置经受一个把平面 $\pi_\infty$ 变换到无穷远的保向变换 H，则参数 $\chi$ 将被一个新的参数 $\chi'$代替，定义为 $\chi'(\mathbf{X})=\chi(H\mathbf{X};PH^{-1})$．$\chi'$的值也必然沿着射线单调地变化．因为 H 是保向的．根据结论 20.13，在摄像机中心前面的点变换后仍将在摄像机的前面，因此 $\chi$ 和 $\chi'$都沿着相同射线方向单调减小．如果 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 是该直线上的两点，那么 $\chi'(\mathbf{X}_1)>\chi'(\mathbf{X}_2)$

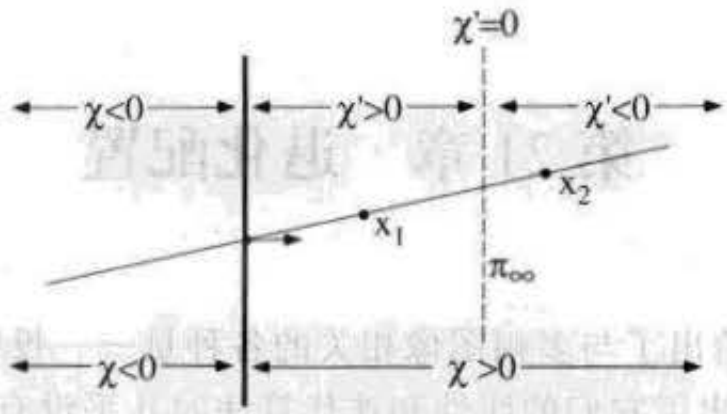



图 20.5 参数 $\chi$

的充分必要条件是 $\chi(\mathbf{X}_1)>\chi(\mathbf{X}_2)$.

当射影变换有负行列式时,在摄像机的前面和后面将局部地反置. 在这种情形中,参数 $\chi'$增加的方向也将被反置. 因此,$\chi'(\mathbf{X}_1)>\chi'(\mathbf{X}_2)$的充分必要条件是 $\chi(\mathbf{X}_1)<\chi(\mathbf{X}_2)$.

当射影变换把景物变换到"真实"景物时,如果有两点被投影到图像上同一点,则有较大 $\chi'$值的点更靠近摄像机. 这导致下面一个能确定在任意一个射影重构中摄像机前面的两个点中谁更靠近摄像机的结论.

**结论 20.18 假设两点 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 映射到一幅图像上的同一点. 考虑这个景物的一个射影重构,并且令参数 $\chi$ 在射影重构的坐标系中(由公式(20.3—$p$369)定义. 如果这个射影重构与真实的场景有相同定向,则在真实场景中离摄像机前面较近的点有比较大的 $\chi$ 值. 另一方面,如果射影变换有相反定向,则 $\chi$ 值比较小的点在真实场景中离摄像机的前面较近.**

如前面已经提到的,除非存在一个平面把用来重构的摄像机与点集分开,否则景物的定向是唯一确定的,并且我们能确定结论 20.18 的射影变换的行列式是正或负. 然而,它可能需要用节 20.7 介绍的线性规划方法来计算该结构的一个强实现. 如果有不同定向的强实现存在,则如图 20.4 所说明的,存在一个本质的多义性. 然而,这个多义性可以由一个点对与摄像机的相对距离的知识来解决.

## 20.10 结束语

### 20.10.1 文献

本章的议题涉足到有向射影几何,它在一本标准和可读的教科书[Stolfi-91]中讨论. Laveau 和 Faugeras 在[Laveau-96b]中应用了有向射影几何的思想. [Robert-93] 也用摄像机的前面和后面的概念来计算射影重构中的凸包. 本章引自论文[Hartley-98a],它也讨论了诸如准仿射映射不变量和允许任意对应有准仿射重构的条件等议题.

[Hartley-94a,Hartley-99]指出正负性、特别是正负性不等式在确定准仿射重构(它是过渡到仿射和度量重构的中间步骤)中有用.

# 第21章 退化配置

在前面各章中，我们给出了与多幅图像相关的各种量——投影矩阵、基本矩阵和三焦点张量的估计算法. 但在给出了它们的线性和迭代算法时几乎没有考虑这些算法失败的可能性. 我们现在就来考虑这种可能性在什么条件下会发生.

通常，如果给定足够多的对应点并在某种类型的"一般位置"上，估计中所求的量将被唯一确定，同时我们给出的算法也将是成功的. 但是，如果给出的对应点太少或所有的点在某种临界配置上，那么唯一解将不存在. 有时会是有限个不同解，而有时会是整个的解簇.

本章集中了本书中我们所碰到的三个主要估计问题：计算摄像机投影矩阵、由两视图重构和由三视图重构. 这里给出的某些结果是经典的，如计算摄像机矩阵和两视图临界曲面问题. 其他的则是相当近期的结果. 我们将依次考虑不同的估计问题.

## 21.1 计算摄像机投影矩阵

我们先考虑给定空间的一组点以及它们在图像上的对应点集来计算摄像机投影矩阵的问题. 这就是说，我们已知空间的点集 $\mathbf{X}_i$ 和它们被投影矩阵为 P 的摄像机映射到图像上的点 $\mathbf{x}_i$，即空间和图像中的点的坐标已知，要求计算矩阵 P. 这个问题已在第 6 章中考虑过. 在考虑这个问题的临界配置之前，我们先离题来研究一下摄像机投影的抽象概念.

**把摄像机看作点** 假定存在对应点的集合 $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$. 让我们假定存在一个唯一的摄像机矩阵 P 使得 $\mathbf{x}_i = \mathrm{P}\mathbf{X}_i$. 现令 H 为该图像的一个射影变换矩阵，并令 $\mathbf{x}_i' = \mathrm{H}\mathbf{x}_i$ 为变换后的图像坐标集合. 那么显然存在一个唯一的摄像机矩阵 P′ 使 $\mathbf{x}_i' = \mathrm{P}'\mathbf{X}_i$，即摄像机矩阵 $\mathrm{P}' = \mathrm{HP}$. 反过来，如果把 $\mathbf{X}_i$ 映射到 $\mathbf{x}_i$ 的摄像机矩阵 P 不止一个，那么把 $\mathbf{X}_i$ 映射到 $\mathbf{x}_i'$ 的摄像机矩阵 P′ 也不止一个. 因此，投影矩阵 P 是否存在唯一解的问题**在相差一个射影变换** H 的意义下由图像点 $\mathbf{x}_i$ 确定.

其次，注意对摄像机矩阵 P 进行射影变换 H 不会改变摄像机中心. 具体地说，点 $\mathbf{C}$ 是摄像机中心的充要条件是 $\mathrm{P}\mathbf{C} = \mathbf{0}$. 但 $\mathrm{P}\mathbf{C} = \mathbf{0}$ 的充要条件是 $\mathrm{HP}\mathbf{C} = \mathbf{0}$. 因此，图像的射影变换保持摄像机中心不变. 下一步，我们将证明这基本上是摄像机保留的仅有的性质.

**结论 21.1 令两个摄像机矩阵 P 和 P′ 有同一中心，那么存在一个非奇异矩阵 H 的图像射影变换使得 $\mathrm{P}' = \mathrm{HP}$ 成立.**

**证明** 如果中心 $\mathbf{C}$ 不在无穷远，那么摄像机矩阵的形式为 $\mathrm{P} = [\mathrm{M} \mid -\mathrm{M}\mathbf{c}]$ 和 $\mathrm{P}' = [\mathrm{M}' \mid -\mathrm{M}'\mathbf{c}]$，其中 $\mathbf{c}$ 是表示摄像机中心的一个非齐次 3 维矢量. 于是显然有 $\mathrm{P}' = \mathrm{M}'\mathrm{M}^{-1}\mathrm{P}$. 如果 $\mathbf{C}$ 是一个无穷远点，那么选择一个 3D 射影变换 G 使得 $\mathrm{G}\mathbf{C} = \hat{\mathbf{C}}$ 为有限点. 此时，两个摄像机矩阵 $\mathrm{PG}^{-1}$ 和 $\mathrm{P}'\mathrm{G}^{-1}$ 有同样的中心，即 $\hat{\mathbf{C}}$，由此推出对某个 H 有 $\mathrm{P}'\mathrm{G} = \mathrm{HPG}$. 消去 G 即得 $\mathrm{P}' = \mathrm{HP}$. □

该结论可以解释成图像在相差一个射影变换的意义下仅仅由摄像机中心确定. 由此看

出在考虑摄像机矩阵的唯一性问题时，可以忽略除摄像机中心以外的其他所有参数，因为只有它能确定图像的射影变换类型，从而确定解是否唯一.

**把图像看作射线的等价类**

为了深入了解计算摄像机投影矩阵的临界配置，我们首先来考虑把 $\mathbb{P}^2$ 映射到 $\mathbb{P}^1$ 的 2 维摄像机. 考虑一个摄像机中心 $\mathbf{c}$ 和空间点的集合 $\mathbf{x}_i$. 射线 $\overline{\mathbf{cx}}_i$ 与一条线阵图像 $\mathbf{l}$ 交于一组点 $\overline{\mathbf{x}}_i$；因此点 $\overline{\mathbf{x}}_i$ 是点 $\mathbf{x}_i$ 的图像. 在 1D 图像中，点 $\mathbf{x}_i$ 到点 $\overline{\mathbf{x}}_i$ 的投影可以用节 5.4.2（$p$119）中所介绍的 $2\times3$ 的射影矩阵来描述.

如第 1 章所证明的，射线集 $\overline{\mathbf{cx}}_i$ 与图像点 $\overline{\mathbf{x}}_i$ 的射影等价类相同. 图 21.1 对此给予了说明. 因此，除了可以把图像看成线阵图像上的点集外，还可把图像看成由摄像机中心出发过每个图像点所组成的射线集合的射影等价类. 在正好有 4 个图像点时，点 $\overline{\mathbf{x}}_i$（或等价地，射线）的交比是该射影等价类的表征. 我们记 $<\mathbf{c};\mathbf{x}_1,\ldots,\mathbf{x}_n>$ 为射线集合 $\overline{\mathbf{cx}}_i$ 的射影等价类.

这些说明对 3D 点到 2 维图像的投影同样有效. 我们也可以扩展上面的标记，即用 $<\mathbf{C},\mathbf{X}_1,\ldots,\mathbf{X}_n>$ 来表示所有射线 $\overline{\mathbf{CX}}_i$ 的等价类. 正如 2 维的情形一样，它是点 $\mathbf{X}_i$ 相对于中心为 $\mathbf{C}$ 的摄像机的投影的一个抽象.

我们将考虑摄像机中心和 3D 点的配置，并采用 $\{\mathbf{C}_1,\ldots\mathbf{C}_m;\mathbf{X}_1,\ldots\mathbf{X}_n\}$ 或其变化形式来标记. 在分号之前的符号表示摄像机中心，而分号之后的表示 3D 点. 为了使推导出的结论叙述简洁，我们定义**图像等价**的概念.

**定义** 21.2　两个配置

$$\{\mathbf{C}_1,\ldots,\mathbf{C}_m;\mathbf{X}_1,\ldots,\mathbf{X}_n\}\text{ 和 }\{\mathbf{C}_1',\ldots,\mathbf{C}_m';\mathbf{X}_1',\ldots,\mathbf{X}_n'\}$$

称为**图像等价**，如果对所有 $i=1,\ldots,m$，都有 $\{\mathbf{C}_i;\mathbf{X}_1,\ldots,\mathbf{X}_n\}=\{\mathbf{C}_i';\mathbf{X}_1',\ldots,\mathbf{X}_n'\}$.

图像等价的概念与其中点和摄像机中心集合的射影等价概念不同. 实际上与重构多义性相关的概念是如果一个配置 $\{\mathbf{C}_1,\ldots\mathbf{C}_m;\mathbf{X}_1,\ldots,\mathbf{X}_n\}$ 与另一个射影不等价的配置有图像等价集，那么就等于存在射影重构问题的多义性，因为点和摄像机的射影结构不能由图像集合唯一确定. 在这种情形，我们说配置 $\{\mathbf{C}_1,\ldots\mathbf{C}_m;\mathbf{X}_1,\ldots,\mathbf{X}_n\}$ **允许另一种重构**.

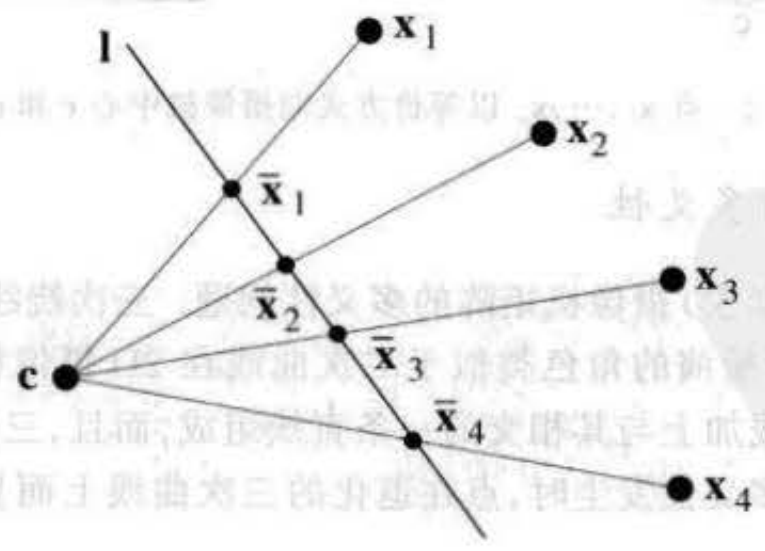



图 21.1　用一个 2D 摄像机产生的平面上点的投影. 1D 图像由射线 $\mathbf{l}_i=\overline{\mathbf{cx}}_i$ 与线阵图像 $\mathbf{l}$ 的交点形成. 图像点集合 $\{\overline{\mathbf{x}}_i\}$ 射影等价于射线集合 $\{\mathbf{l}_i\}$. 图像的射影等价类由 4 点的交比确定.

### 21.1.1 2D 中的多义性——Chasles 定理

在考虑常用的 3D 摄像机之前，我们先讨论较简单的 2D 摄像机. 2D 摄像机投影的唯一性分析涉及到平面二次曲线. 由已知的投影点集 $\mathbf{x}_i$ 确定摄像机中心的多义性意味着由两个不同的中心 $\mathbf{c},\mathbf{c}'$ 所产生的投影点相同. 问题是对于什么样的点配置这种情形会发生. 这个问题的答案由 Chasles 定理给出.

**定理 21.3 Chasles 定理** **令 $\mathbf{x}_i$ 为 $n$ 个点的集合，而 $\mathbf{c}$ 和 $\mathbf{c}'$ 为两个摄像机中心，它们都在同一平面上. 那么**

$$<\mathbf{c};\mathbf{x}_1,\ldots,\mathbf{x}_n>=<\mathbf{c}';\mathbf{x}_1,\ldots,\mathbf{x}_n>$$

**的充要条件是**

**(1) 点 $\mathbf{c},\mathbf{c}'$ 和所有 $\mathbf{x}_i$ 都在一条非退化的二次曲线上，或**

**(2) 所有点在两条直线(一种退化的二次曲线)的并集上，而且两个摄像机中心同在其中的一条直线上.**

这两种配置都在图 21.2 中给出.

注意该定理的一个简单推论为：如果 $\mathbf{c}''$ 为任何其他的摄像机中心并与 $\mathbf{c},\mathbf{c}'$ 和 $\mathbf{x}_i$ 一样也在一条二次曲线(退化或不退化的)的同一分支上，那么点到 $\mathbf{c}''$ 的投影等价于它们到原来中心的投影. 而且，对于增加的任意数目的点 $\mathbf{x}_j$，只要它们仍在同一条二次曲线上就不会摆脱多义性.

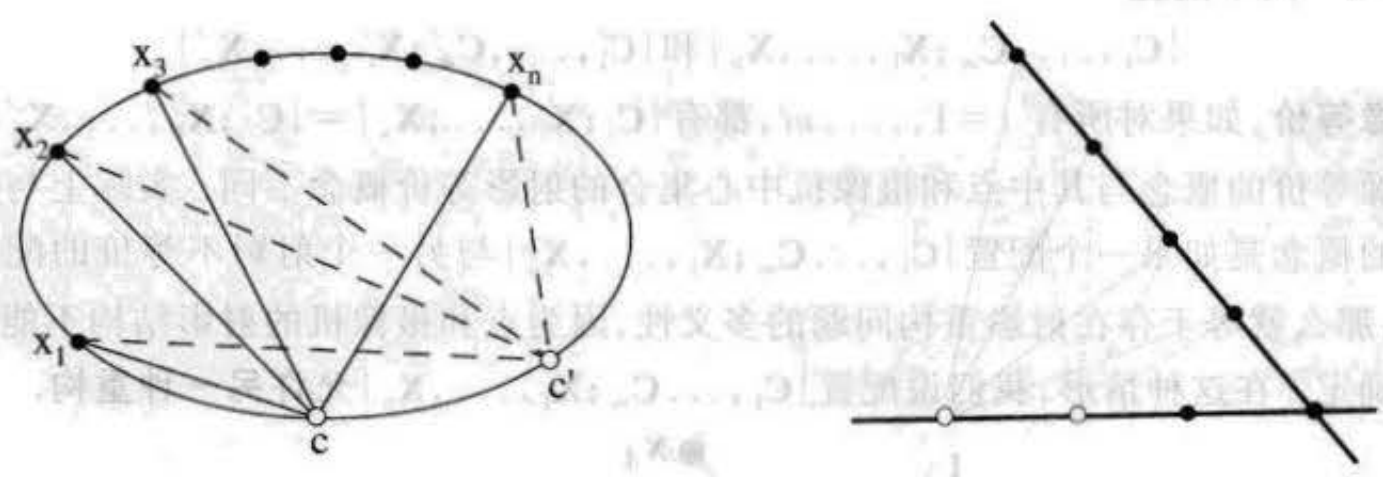



图 21.2 点 $\mathbf{x}_1,\cdots,\mathbf{x}_n$ 以等价方式向摄像机中心 $\mathbf{c}$ 和 $\mathbf{c}'$ 投影.

### 21.1.2 3D 摄像机的多义性

现在我们来讨论计算 3D 摄像机矩阵的多义性问题. 三次绕线(在节 2.3($p$43)中介绍)在 3D 情形的多义性中所扮演的角色类似于二次曲线在 2D 摄像机中的角色. 三次绕线的退化形式由一条二次曲线加上与其相交的一条直线组成，而且，三条直线也是退化的三次曲线. 如 2D 的情形一样，多义性发生时，点在退化的三次曲线上而且摄像机的中心必须在同一支上.

图 21.3 以及下面的定义给出导致摄像机投影矩阵计算多义性的点和摄像机配置的完整分类. 我们现在介绍其几何配置. 与摄像机配置多义性确切相关的问题将在后面给出.

**定义 21.4**

计算摄像机矩阵的**临界集合**由两部分组成：

(1) 包含摄像机中心以及任何数目的 3D 点的一条代数曲线 $\mathcal{C}$. 这条曲线可以是

(a) 一条非退化的三次绕线(自由度 3),

(b) 一条平面二次曲线(自由度 2),

(c) 一条直线(自由度 1),或

(d) 一个点(自由度 0).

(2) 包含任何数目的 3D 点的线性子空间 L(线或平面)的并集.

曲线 $\mathcal{C}$ 和线性子空间满足三个条件:

(1) 每一个线性子空间必须与曲线 $\mathcal{C}$ 相交.

(2) 曲线 $\mathcal{C}$ 的自由度与线性子空间的维数之和最多为 3.

(3) 除 $\mathcal{C}$ 是单点的情形外,摄像机都不在 $\mathcal{C}$ 和线性子空间的交点上.

图 21.3 给出各种可能性,很容易验证它们完全列举了与定义 21.4 相符的所有配置.

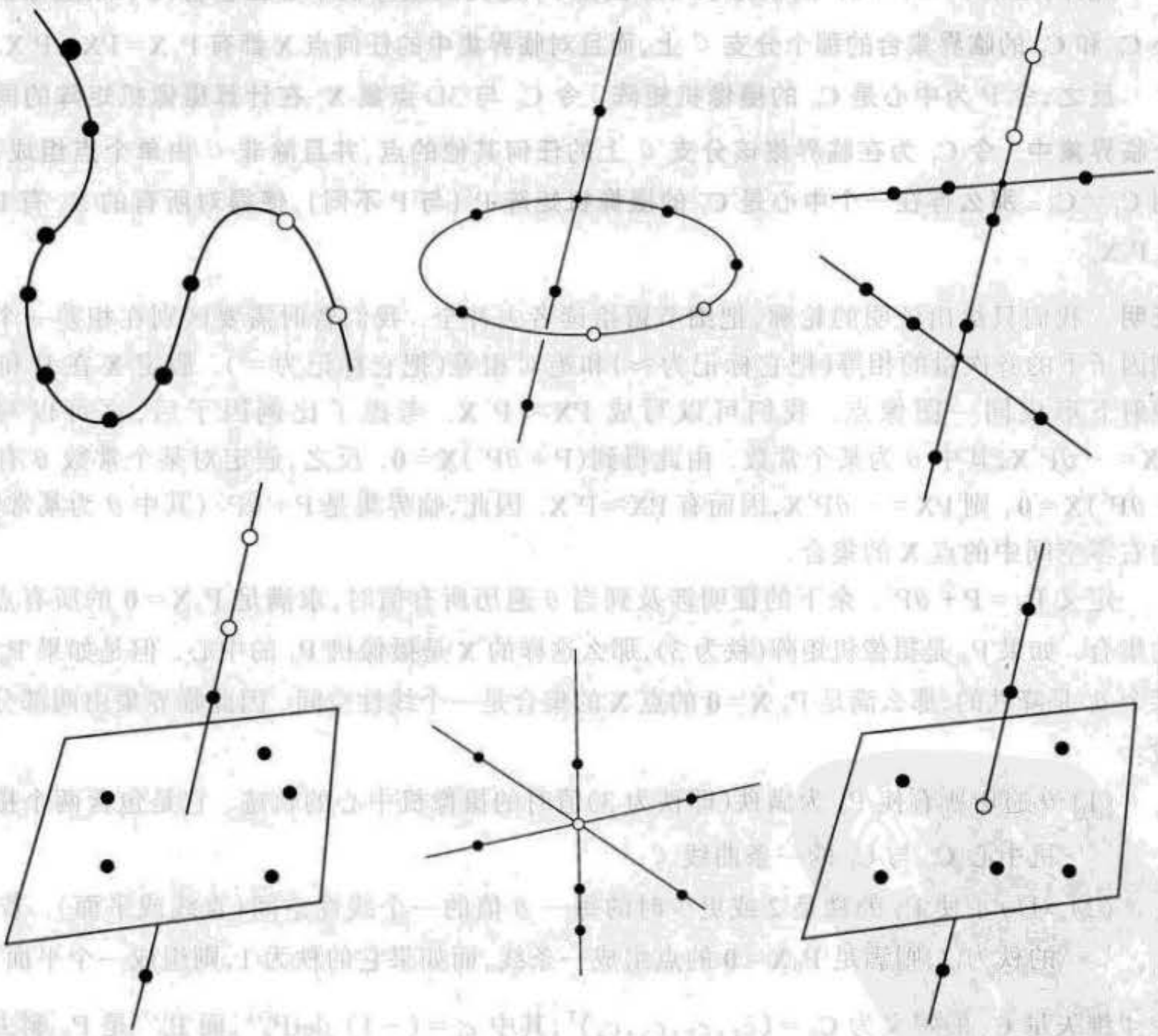


图 21.3 由单幅视图计算摄像机矩阵的不同临界配置. 空心圈表示投影中心而实心圈表示点. 每一种情形都由一条包含摄像机中心的代数曲线(或一个点)$\mathcal{C}$ 外加一个线性子空间(直线或平面)的集合构成.

**结论 21.5** 计算摄像机投影矩阵时可能出现的各种临界集是

(1) 一条非退化的三次饶线 $\mathcal{C}$(自由度 3).

(2) 一条平面二次曲线 $\mathcal{C}$(自由度 2)加上与之相交的一条直线(维数 1).

(3) 一条直线 $\mathcal{C}$(自由度 1)加上最多两条与第一条直线相交的其他直线(总维数 2).

(4) 一条直线 $\mathcal{C}$(自由度 1)加上一张平面(维数 2).

(5) 一个点 $\mathcal{C}$(自由度 0)加上最多三条过该点的直线(总维数 3).

(6) 一个点 $\mathcal{C}$(自由度 0)和都过该点的一条直线和一张平面.

这些临界集合与计算摄像机矩阵的多义性的确切关系由下面的结论给出.

**结论 21.6** 令 P 和 P′为中心是 $\mathbf{C}_0$ 和 $\mathbf{C}_1$ 的两个不同的摄像机矩阵. 那么满足 $\mathrm{P}\mathbf{X}_i = \mathrm{P}'\mathbf{X}_i$ 的两个摄像机中心和点 $\mathbf{X}_i$ 都在定义 21.4 的临界集合上.

此外,如果 $\mathrm{P}_\theta = \mathrm{P} + \theta\mathrm{P}'$的秩为 3[①],那么由 $\mathrm{P}_\theta$ 定义的摄像机中心在包含两个原摄像机中心 $\mathbf{C}_0$ 和 $\mathbf{C}_1$ 的临界集合的那个分支 $\mathcal{C}$ 上,而且对临界集中的任何点 $\mathbf{X}$ 都有 $\mathrm{P}_\theta\mathbf{X} = \mathrm{P}\mathbf{X} = \mathrm{P}'\mathbf{X}$.

反之,令 P 为中心是 $\mathbf{C}_0$ 的摄像机矩阵. 令 $\mathbf{C}_0$ 与 3D 点集 $\mathbf{X}_i$ 在计算摄像机矩阵的同一个临界集中. 令 $\mathbf{C}_1$ 为在临界集该分支 $\mathcal{C}$ 上的任何其他的点,并且除非 $\mathcal{C}$ 由单个点组成,否则 $\mathbf{C}_1 \neq \mathbf{C}_0$. 那么存在一个中心是 $\mathbf{C}_1$ 的摄像机矩阵 P′(与 P 不同),使得对所有的 $\mathbf{X}_i$ 有 $\mathrm{P}\mathbf{X}_i = \mathrm{P}'\mathbf{X}_i$.

**证明** 我们只给出证明的轮廓,把细节留给读者去补全. 我们暂时需要区别在相差一个比例因子下的齐次量的相等(把它标记为≈)和绝对相等(把它标记为 =). 假定 $\mathbf{X}$ 在 P 和 P′映射下形成同一图像点. 我们可以写成 $\mathrm{P}\mathbf{X} \approx \mathrm{P}'\mathbf{X}$. 考虑了比例因子后,这可以写成 $\mathrm{P}\mathbf{X} = -\theta\mathrm{P}'\mathbf{X}$,其中 $\theta$ 为某个常数. 由此得到$(\mathrm{P} + \theta\mathrm{P}')\mathbf{X} = \mathbf{0}$. 反之,假定对某个常数 $\theta$ 有$(\mathrm{P} + \theta\mathrm{P}')\mathbf{X} = \mathbf{0}$, 则 $\mathrm{P}\mathbf{X} = -\theta\mathrm{P}'\mathbf{X}$,因而有 $\mathrm{P}\mathbf{X} \approx \mathrm{P}'\mathbf{X}$. 因此,临界集是 $\mathrm{P} + \theta\mathrm{P}'$ (其中 $\theta$ 为某常数)的右零空间中的点 $\mathbf{X}$ 的集合.

定义 $\mathrm{P}_\theta = \mathrm{P} + \theta\mathrm{P}'$. 余下的证明涉及到当 $\theta$ 遍历所有值时,求满足 $\mathrm{P}_\theta\mathbf{X} = \mathbf{0}$ 的所有点 $\mathbf{X}$ 的集合. 如果 $\mathrm{P}_\theta$ 是摄像机矩阵(秩为 3),那么这样的 $\mathbf{X}$ 是摄像机 $\mathrm{P}_\theta$ 的中心. 但是如果 $\mathrm{P}_\theta$ 对某个 $\theta_i$ 是降秩的,那么满足 $\mathrm{P}_{\theta_i}\mathbf{X} = \mathbf{0}$ 的点 $\mathbf{X}$ 的集合是一个线性空间. 因此临界集由两部分组成:

(1) $\theta$ 遍历所有使 $\mathrm{P}_\theta$ 为满秩(即秩为 3)值时的摄像机中心的轨迹. 它是包含两个摄像机中心 $\mathbf{C}_0$ 与 $\mathbf{C}_1$ 的一条曲线 $\mathcal{C}$.

(2) 对应于使 $\mathrm{P}_\theta$ 的秩是 2 或更少时的每一 $\theta$ 值的一个线性空间(直线或平面). 若 $\mathrm{P}_\theta$ 的秩为 2,则满足 $\mathrm{P}_\theta\mathbf{X} = \mathbf{0}$ 的点组成一条线,而如果它的秩为 1,则组成一个平面.

令 4 维矢量 $\mathbf{C}_\theta$ 的定义为 $\mathbf{C}_\theta = (c_1, c_2, c_3, c_4)^\mathsf{T}$,其中 $c_i = (-1)^i \det \mathrm{P}_\theta^{(i)}$,而 $\mathrm{P}_\theta^{(i)}$ 是 $\mathrm{P}_\theta$ 删去了第 $i$ 列后的矩阵. 因为每个 $\mathrm{P}_\theta^{(i)}$ 是一个 $3 \times 3$ 的矩阵并且 $\mathrm{P}_\theta$ 的元素是 $\theta$ 的线性函数,所以每一个 $c_i = (-1)^i \det \mathrm{P}_\theta^{(i)}$ 是 $\theta$ 的三次多项式. 根据定义 20.10($p$369)之后的那一段讨论,

① 包括 $\mathrm{P}_\infty = \mathrm{P}'$的情况

$P_\theta C_\theta = 0$,从而如果 $C_\theta \neq 0$,那么它是 $P_\theta$ 的摄像机中心,并且当 $\theta$ 变化时点 $C_\theta$ 的轨迹就是曲线 $\mathcal{C}$. 因为 $C_\theta$ 的坐标是三次多项式,所以它一般是一条三次绕线. 但是如果 $C_\theta$ 的四个元素有重根 $\theta_i$,那么曲线 $\mathcal{C}_\theta$ 的自由度减少. 在这种情形,$P_{\theta_i}$ 不是满秩,并且存在使 $P_{\theta_i}X = 0$ 的点 $X$ 的线性空间. 这样一来,临界集合的线性子空间的元素对应于使 $C_\theta$ 为零的 $\theta$ 值. 显然最多有三个这样的值. 进一步的细节留给读者去补全.

定理的最后部分是一个反命题——也就是说如果点和一个摄像机中心在同一个临界配置上,那么存在另一个摄像机投影矩阵使得它所对应的摄像机在 $\mathcal{C}$ 的任何位置上. 如上所看到的:在这里摄像机矩阵 P 的确切形式并不重要,唯一重要的是摄像机的中心. 对图21.3中的大多数配置而言,从几何上就足以清楚地看到当摄像机沿轨迹 $\mathcal{C}$ 移动时,临界集的图像不变(相差一个射影变换的意义下). 当 $\mathcal{C}$ 是平面二次曲线时,可以轻易地从 1D 摄像机的情形(定理 21.3)类推得到有关结论. 但三次绕线的情形例外. 图 21.4 用图示的方法对此作了说明. 我们现暂不作证明而在以后的内容中(结论 21.25($p$393))再回过来讨论这个问题. □

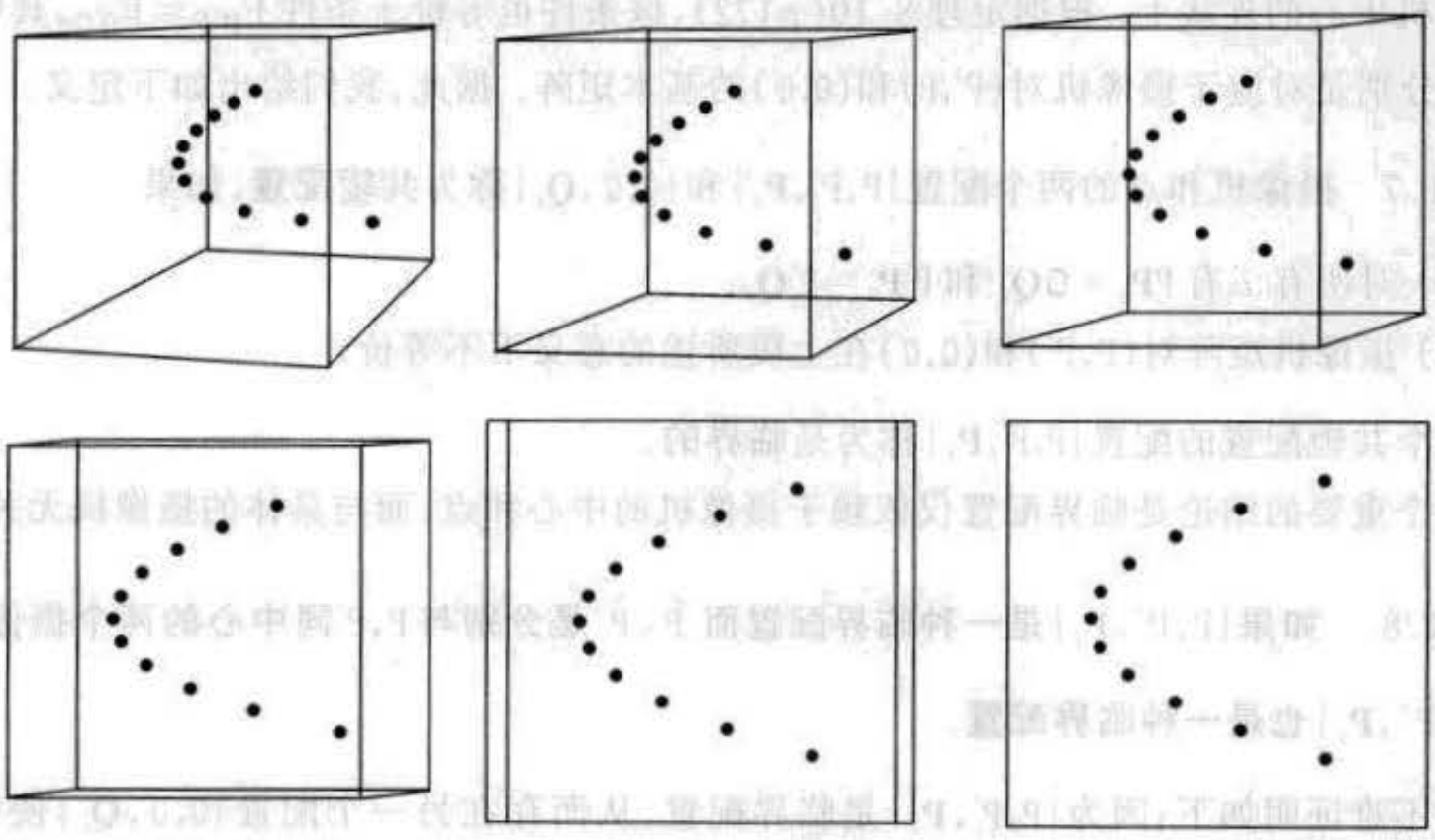


图 21.4 由一个投影中心观看三次绕线 $(t^3, t^2, t, 1)^\top$ 上点集的不同的视图. 看到的点是当 $t = 3, 4, 5, 10, 50, 1000$ 时所观察到的图像并且对图像做了适当的放大以防止点集变得太小. 从图像大致可以看到,这些点集相差一个射影变换. 从三次绕线的观点来看,三次绕线看上去象一条二次曲线,对于现在的具体情形,它是一条抛物线.

## 21.2 两视图中的退化特性

**记号** 在本章的余下部分,用 P 和 Q 表示摄像机矩阵,用 **P** 和 **Q** 表示 3D 点,而与它们对应的 2D 点用 $\mathbf{p} = P\mathbf{P}$ 和 $\mathbf{q} = Q\mathbf{Q}$ 表示. 因此摄像机和 3D 点仅根据它们的字形来区分. 这样做也许显得有点混淆,但采用下标或加撇等其他方法更容易混淆. 在由图像坐标重构的多义性讨论中,我们用 P 和 **P** 表示一个重构,而 Q 和 **Q** 表示另一个重构.

现在我们转向一个物体有两幅视图的情形. 给定足够多的处于"一般位置"的点,我们可以确定两个摄像机的位置并在相差一个射影变换的意义下使该点集得到重构. 这可以用第九章中介绍的某个射影重构算法来完成. 我们现在希望确定在什么情况下该技术会失败.

因此,我们来考虑一组图像对应 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$. 这组对应由一对摄像机矩阵 P 和 P′以及对所有 $i$ 满足 $\mathbf{x}_i = \mathrm{P}\mathbf{P}_i$ 和 $\mathbf{x}_i' = \mathrm{P}'\mathbf{P}_i$ 的一组 3D 点 $\mathbf{P}_i$ 组成. 对 $\{\mathrm{P},\mathrm{P}',\mathbf{P}_i\}$ 和 $\{\mathrm{Q},\mathrm{Q}',\mathbf{Q}_i\}$ 而言,如果有一个用矩阵 H 表示的 3D 射影变换使得对所有 $i$ 都有 $\mathrm{Q} = \mathrm{P}\mathrm{H}^{-1}$, $\mathrm{Q}' = \mathrm{P}'\mathrm{H}^{-1}$ 和 $\mathbf{Q}_i = \mathrm{H}\mathbf{P}_i$,则它们射影等价.

由于问题的特殊性,这种等价的定义不完全适用当前的讨论. 由射影重构定理 9.1 (p181)可知,我们不能确定在两摄像机中心连线上的点的位置. 因此,如果某些点在摄像机中心的连线上,那么非射影等价的重构总是存在. 这种类型的重构多义性没有很大意义,因此我们将修改等价性的概念:两个重构等价的条件是存在一个 H 使 $\mathrm{Q} = \mathrm{P}\mathrm{H}^{-1}$ 以及 $\mathrm{Q}' = \mathrm{P}'\mathrm{H}^{-1}$. 如同证明射影重构定理一样,这样的 H 也把 $\mathbf{Q}_i$ 映射到 $\mathbf{P}_i$,可能的例外是重构点 $\mathbf{Q}_i$ 在摄像机中心的连线上. 根据定理 8.10($p$172),该条件也等价于条件 $\mathrm{F}_{\mathrm{P}'\mathrm{P}} = \mathrm{F}_{\mathrm{Q}'\mathrm{Q}}$,其中 $\mathrm{F}_{\mathrm{P}'\mathrm{P}}$ 和 $\mathrm{F}_{\mathrm{Q}'\mathrm{Q}}$ 分别是对应于摄像机对(P′,P)和(Q,Q′)的基本矩阵. 据此,我们给出如下定义.

**定义 21.7** 摄像机和点的两个配置 $\{\mathrm{P},\mathrm{P}',\mathbf{P}_i\}$ 和 $\{\mathrm{Q},\mathrm{Q}',\mathbf{Q}_i\}$ 称为**共轭配置**,如果

(1) 对所有 $i$ 有 $\mathrm{P}\mathbf{P}_i = \mathrm{Q}\mathbf{Q}_i$ 和 $\mathrm{P}'\mathbf{P}_i = \mathrm{Q}'\mathbf{Q}_i$.

(2) 摄像机矩阵对(P,P′)和(Q,Q′)在上段所述的意义下不等价.

拥有一个共轭配置的配置 $\{\mathrm{P},\mathrm{P}',\mathbf{P}_i\}$ 称为是**临界的**.

一个重要的结论是临界配置仅依赖于摄像机的中心和点,而与具体的摄像机无关.

**结论 21.8** **如果 $\{\mathrm{P},\mathrm{P}',\mathbf{P}_i\}$ 是一种临界配置而 $\hat{\mathrm{P}}$、$\hat{\mathrm{P}}'$ 是分别与 P,P′同中心的两个摄像机,那么 $\{\hat{\mathrm{P}},\hat{\mathrm{P}}',\mathbf{P}_i\}$ 也是一种临界配置.**

**证明** 不难证明如下:因为 $\{\mathrm{P},\mathrm{P}',\mathbf{P}_i\}$ 是临界配置,从而存在另一个配置 $\{\mathrm{Q},\mathrm{Q}',\mathbf{Q}_i\}$ 使得对所有 $i$ 有 $\mathrm{P}\mathbf{P}_i = \mathrm{Q}\mathbf{Q}_i$ 和 $\mathrm{P}'\mathbf{P}_i = \mathrm{Q}'\mathbf{Q}_i$. 但是因为 P、$\hat{\mathrm{P}}$ 有同样的摄像机中心,根据结论 21.1 有 $\hat{\mathrm{P}} = \mathrm{H}\mathrm{P}$,类似有 $\hat{\mathrm{P}}' = \mathrm{H}'\mathrm{P}'$. 因此

$$\hat{\mathrm{P}}\mathbf{P}_i = \mathrm{H}\mathrm{P}\mathbf{P}_i = \mathrm{H}\mathrm{Q}\,\mathbf{Q}_i \text{ 和}$$

$$\hat{\mathrm{P}}'\mathbf{P}_i = \mathrm{H}'\mathrm{P}'\mathbf{P}_i = \mathrm{H}'\mathrm{Q}'\mathbf{Q}_i.$$

因此,$\{\mathrm{H}\mathrm{Q},\mathrm{H}'\mathrm{Q}',\mathbf{Q}_i\}$ 是 $\{\hat{\mathrm{P}},\hat{\mathrm{P}}',\mathbf{P}_i\}$ 的一种替代配置,因而是临界的. □

本节的目的是确定在什么条件下点对应集的非等价实现会发生,这个问题可以由下面的定理得到完全地解决,该定理将逐步地给予证明.

**定理 21.9** **(1) 摄像机和点的共轭配置是三元组. 因此一种临界配置 $\{\mathrm{P},\mathrm{P}',\mathbf{P}_i\}$ 有两个共轭配置 $\{\mathrm{Q},\mathrm{Q}',\mathbf{Q}_i\}$ 和 $\{\mathrm{R},\mathrm{R}',\mathbf{R}_i\}$.**

(2) **如果$\{P,P',\mathbf{P}_i\}$是一种临界配置，那么所有的点$\mathbf{P}_i$和两摄像机中心$\mathbf{C}_P$和$\mathbf{C}_{P'}$在一个直纹二次曲面$S_P$上.**

(3) **反过来，假定摄像机$P,P'$的中心和3D点集$\mathbf{P}_i$在一个直纹二次曲面上**(不包括 $p527$ 的退化二次曲面(5)和(8))；**那么$\{P,P',\mathbf{P}_i\}$是一种临界配置.**

上文中"直纹二次曲面"表示任何(不包括 $p527$ 的退化二次曲面(5)和(8))包含直线的二次曲面. 我们将会看到它包含各种退化的二次曲面. 二次曲面的一般讨论和分类在节 2.2.4($p41$)中给出. 一个**二次曲面**通常定义为满足 $\mathbf{X}^T S\mathbf{X}=0$ 的点集 $\mathbf{X}$，其中 S 是一个 $4\times4$ 的对称矩阵. 但是，如果 S 是任意的一个$4\times4$矩阵，即不一定是对称的，那么我们可以看到对任何点 $\mathbf{X}$，我们有 $\mathbf{X}^T S\mathbf{X}=(\mathbf{X}^T S_{sym}\mathbf{X})$，其中 $S_{sym}=(S+S^T)/2$ 是 S 的对称部分. 这样一来，$\mathbf{X}^T S\mathbf{X}=0$ 的充要条件是$(\mathbf{X}^T S_{sym}\mathbf{X})=0$，因而 S 与 $S_{sym}$ 定义同样的二次曲面. 在研究重构多义性时，通常用非对称的矩阵 S 表示二次曲面比较方便.

**证明** 我们先证明定理的第一部分. 令 F 和 F′ 为两个不同的基本矩阵，对所有对应 $\mathbf{x}_i'\leftrightarrow\mathbf{x}_i$ 满足 $\mathbf{x}_i'^T F\mathbf{x}_i=\mathbf{x}_i'^T F'\mathbf{x}_i=0$. 定义 $F_\theta=F+\theta F'$. 容易验证 $\mathbf{x}_i'^T F_\theta\mathbf{x}_i=0$. 但 $F_\theta$ 仅当 $\det F_\theta=0$ 时为基本矩阵. 因为 $\det F_\theta$ 是 $\theta$ 的 3 次多项式. 该多项式的根 $\theta=0$ 和 $\theta=1$ 分别对应于 F 和 F′. 第三个(实)根对应于第三个基本矩阵，因而有三个非等价重构. □

定理的第二部分由证明下面的引理得到.

**引理 21.10** **考虑对应不同基本矩阵 $F_{P'P}$和 $F_{Q'Q}$的两对摄像机$(P,P')$和$(Q,Q')$. 定义二次曲面 $S_P=P'^T F_{Q'Q}P$ 和 $S_Q=Q'^T F_{P'P}Q$.**

(1) **二次曲面 $S_P$包含 P 和 P′的摄像机中心. 同样，$S_Q$包含 Q 和 Q′的摄像机中心.**

(2) **$S_P$是直纹二次曲面.**

(3) **如果 $\mathbf{P}$ 和 $\mathbf{Q}$ 是满足 $P\mathbf{P}_i=Q\mathbf{Q}_i$ 和 $P'\mathbf{P}_i=Q'\mathbf{Q}_i$ 的 3D 点，那么 $\mathbf{P}$ 在二次曲面 $S_P$上，而 $\mathbf{Q}$ 在 $S_Q$上.**

(4) **反之，如果 $\mathbf{P}$ 在二次曲面 $S_P$上，那么存在一个在 $S_Q$上的点 $\mathbf{Q}$ 满足 $P\mathbf{P}=Q\mathbf{Q}$ 和 $P'\mathbf{P}=Q'\mathbf{Q}$.**

**证明** 根据结论 8.12($p172$)，矩阵 F 是一对摄像机$(P,P')$的基本矩阵的充要条件是 $P'^T FP$ 为反对称矩阵. 但是，由于 $F_{P'P}\neq F_{Q'Q}$，在这里定义的矩阵 $S_P$ 和 $S_Q$ 不是反对称的，从而表示非平凡二次曲面.

我们约定矩阵 P 或 P′ 所表示的摄像机的中心用 $\mathbf{C}_P$ 或 $\mathbf{C}_{P'}$ 表示.

(1) P 的摄像机中心满足 $P\mathbf{C}_P=\mathbf{0}$. 那么因为 $P\mathbf{C}_P=\mathbf{0}$，有

$$\mathbf{C}_P^T S_P\mathbf{C}_P=\mathbf{C}_P^T(P'^T F_{Q'Q}P)\mathbf{C}_P=\mathbf{C}_P^T(P'^T F_{Q'Q})P\mathbf{C}_P=0,$$

因此 $\mathbf{C}_P$在二次曲面 $S_P$上. 同样，$\mathbf{C}_{P'}$ 在 $S_P$上.

(2) 令 $\mathbf{e}_Q$ 为由 $F_{Q'Q}\mathbf{e}_Q=\mathbf{0}$ 定义的对极点，并考虑过 $\mathbf{C}_P$ 的射线，它包含所有使 $\mathbf{e}_Q=P\mathbf{X}$ 的点. 因此，对任何这样的点，可以验证 $S_P\mathbf{X}=P'^T F_{Q'Q}P\mathbf{X}=P'^T F_{Q'Q}\mathbf{e}_Q=0$. 这样一来，该射线在 $S_P$上，从而 $S_P$是直纹二次曲面.

(3)在给定条件下,我们可以看到

$$\mathbf{P}^{\mathsf{T}}S_{\mathrm{P}}\mathbf{P}=\mathbf{P}^{\mathsf{T}}\mathrm{P}'^{\mathsf{T}}F_{Q'Q}\mathrm{P}\mathbf{P}=\mathbf{Q}^{\mathsf{T}}(Q'^{\mathsf{T}}F_{Q'Q}Q)\mathbf{Q}=0,$$

因为$Q'^{\mathsf{T}}F_{Q'Q}Q$是反对称的. 因此$\mathbf{P}$在二次曲面$S_{\mathrm{P}}$上. 同理$\mathbf{Q}$在$S_Q$上.

(4)令$\mathbf{P}$在二次曲面$S_{\mathrm{P}}$上并定义$\mathbf{x}=\mathrm{P}\mathbf{P}$和$\mathbf{x}'=\mathrm{P}'\mathbf{P}$. 那么,由$\mathbf{P}^{\mathsf{T}}S_{\mathrm{P}}\mathbf{P}=0$,我们推得$0=\mathbf{P}^{\mathsf{T}}\mathrm{P}'^{\mathsf{T}}F_{Q'Q}\mathrm{P}\mathbf{P}=\mathbf{x}'^{\mathsf{T}}F_{Q'Q}\mathbf{x}$,因而$\mathbf{x}\leftrightarrow\mathbf{x}'$是相对$F_{Q'Q}$的对应点对. 因此存在一个点$\mathbf{Q}$使得$Q\mathbf{Q}=\mathbf{x}=\mathrm{P}\mathbf{P}$和$Q'\mathbf{Q}=\mathbf{x}'=\mathrm{P}'\mathbf{P}$. 由该引理的(3)可知$\mathbf{Q}$必然在$S_Q$上. □

这个引理全面描述了产生图像对应多义性的3D点集. 注意只要世界点在给定的二次曲面上,任何两个任意选择的摄像机对都会给出多义性的图像对应.

### 21.2.1 多义性的例子

至此,定理21.9的逆仍有待证明. 这需要考虑所有类型的直纹二次曲面以及两个摄像机中心在二次曲面的任意放置. 直纹二次曲面的不同类型(包括退化情形)是(见节2.2.4($p41$)):单叶双曲面、锥面、两个(相交)平面、单个平面、单条直线. 完全地枚举两个摄像机中心在一个直纹二次曲面的放置的类型在下列结论中给出.

**结论21.11 两个不同点(具体是摄像机中心)在一个直纹二次曲面的可能配置枚举如下:**

**(1) 单叶双曲面,两点不在同一条母线上**

**(2) 单叶双曲面,两点在同一条母线上**

**(3) 锥面,一点在顶点而另一点不在**

**(4) 锥面,点既不在顶点也不在同一条母线上**

**(5) 锥面,没有一个点在顶点上,但两点都在同一条母线上**

**(6) 一对平面,两点在两平面的交线上**

**(7) 一对平面,一点在交线上而另一点不在**

**(8) 一对平面,两点都不在交线上,但在不同的平面上**

**(9) 一对平面,两点都不在交线上,但在同一平面上**

**(10) 单个平面和两个点**

**(11) 单条线和两个点**

**在同一类中的任何两个二次曲面/点对是射影等价的.**

显然,通过情形枚举可知上面结论的列举是完全的. 很明显,在同一类型中任何两个配置是射影等价的,唯一不明显的是单叶双曲面的非退化情形. 这个事实的证明将留作本章末尾($p398$)的练习(1).

现在来考虑一种临界配置$\{\mathrm{P},\mathrm{P}',\mathbf{P}_i\}$的例子,其中所有$\mathbf{P}_i$在一个二次曲面$S_{\mathrm{P}}$上,$S_{\mathrm{P}}$同时还包含两个摄像机中心. 该二次曲面和两个摄像机中心属于结论21.11中的一种类型.

令两个新摄像机$(\hat{\mathrm{P}},\hat{\mathrm{P}}')$的中心以及一组点$\hat{\mathbf{P}}_i$在一个二次曲面$\hat{S}_{\mathrm{P}}$上. 假定$\hat{S}_{\mathrm{P}}$和两个摄像机中心与给定的例子在同一类型中. 因为在同一类型中的两个配置是射影等价的,我们可以假定$\hat{S}_{\mathrm{P}}=S_{\mathrm{P}}$且$\mathrm{P}$和$\hat{\mathrm{P}}$的中心一样,同样$\mathrm{P}'$和$\hat{\mathrm{P}}'$的中心也一样. 因为点$\hat{\mathbf{P}}_i$在$S_{\mathrm{P}}$上,因而$\{\mathrm{P},\mathrm{P}',\hat{\mathbf{P}}_i\}$是临界配置,因此,根据结论21.8,$\{\hat{\mathrm{P}},\hat{\mathrm{P}}',\hat{\mathbf{P}}_i\}$也是.

这表明为了证明定理 21.9 的逆命题，只要对属于结论 21.11 的每一类给出一个临界配置的例子就足够了. 下面将给出若干类型的例子但不是全部. 剩下的情形留给感兴趣的读者.

**临界配置的例子**

考虑 $P=Q=[I|\mathbf{0}]$ 和 $P'=[I|\mathbf{t}_0]$ 的情形. 对此情形，我们看到

$$S_P=\begin{bmatrix} I \\ \mathbf{t}_0^T \end{bmatrix} F_{Q'Q}[I|\mathbf{0}] = \begin{bmatrix} I \\ \mathbf{t}_0^T \end{bmatrix}[F_{Q'Q}|\mathbf{0}],$$

因此

$$S_{P\mathrm{sym}}=\frac{1}{2}\begin{bmatrix} F_{Q'Q}+F_{Q'Q}^T & F_{Q'Q}^T\mathbf{t}_0 \\ \mathbf{t}_0^T F_{Q'Q} & 0 \end{bmatrix}.$$

给定秩为 2 的基本矩阵 $F_{Q'Q}$ 和摄像机矩阵 $Q=[I|\mathbf{0}]$，我们可以轻易地求出另一个摄像机矩阵 $Q'$. 这可以由结论 8.13($p$172)的公式得到. 我们现在来考虑结论 21.11 中不同类型的临界配置的几个例子.

**例 21.12** 单叶双曲面——两中心不在同一条母线上. 我们选择 $F_{Q'Q}=\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$ 和 $\mathbf{t}_0=(0,2,0)^T$，那么 $\mathbf{t}_0^T F_{Q'Q}=(0,2,0)$，并且我们求得

$$S_{\mathrm{sym}}=\begin{bmatrix} 1 & & & \\ & 1 & & 1 \\ & & -1 & \\ & 1 & & 0 \end{bmatrix}.$$

这是方程为 $X^2+Y^2+2Y-Z^2=0$ 或 $X^2+(Y+1)^2-Z^2=0$ 的二次曲面. 它是单叶双曲面. 注意此时，摄像机中心为 $\mathbf{C}_P=\mathbf{C}_Q=(0,0,0)^T$(在非齐次坐标下). 摄像机中心 $\mathbf{C}_{P'}=(0,2,0)^T$. 我们注意到由 $\mathbf{C}_P$ 到 $\mathbf{C}_{P'}$ 的直线不在此二次曲面上. 摄像机中心 $\mathbf{C}_{Q'}$ 并不是由早已给定的信息唯一确定，因为基本矩阵不唯一确定这两个摄像机. 但是，满足 $F_{Q'Q}\mathbf{e}=\mathbf{0}$ 的对极点 $\mathbf{e}$ 是 $\mathbf{e}=(1,0,-1)^T$. 由于 $\mathbf{e}=Q\mathbf{C}_{Q'}$ 我们得到 $\mathbf{C}_{Q'}=(1,0,-1,k^{-1})^T$. 在非齐次坐标下，对任何 $k$ 有 $\mathbf{C}_{Q'}=k(1,0,-1)^T$. 我们证明了由 $\mathbf{C}_P$ 到 $\mathbf{C}_{Q'}$ 的直线完全在该二次曲面上，但 $\mathbf{C}_{Q'}$ 可能在这条直线的任何地方. △

**例 21.13** 单叶双曲面，两中心在同一条母线上. $F_{Q'Q}$ 的选择和上例一样而 $\mathbf{t}_0=(3,4,5)^T$，那么 $\mathbf{t}_0^T F_{Q'Q}=(-2,4,-2)$ 并且我们求得

$$S_{\mathrm{sym}}=\begin{bmatrix} 1 & & & -1 \\ & 1 & & 2 \\ & & -1 & -1 \\ -1 & 2 & -1 & 0 \end{bmatrix}.$$

这是方程为 $(X-1)^2+(Y+2)^2-(Z+1)^2=4$ 的二次曲面. 它也是单叶双曲面. 可以验证直线 $(X,Y,Z)=(3t,4t,5t)$ 完全在二次曲面上并且包含两个摄像机中心 $(0,0,0)^T$ 和

$(3,4,5)^T$. △

**例 21.14** 锥面——两个中心在该锥面的顶点上. $F_{Q'Q}$的选择和上例一样而 $\mathbf{t}_0=(1,0,1)^T$. 对此情形我们看到 $\mathbf{t}_0^T F_{Q'Q}=\mathbf{0}^T$ 并且求得 $S_{sym}=\mathrm{diag}(1,1,-1,0)$. 这是一个锥面的例子. P 和 Q 的摄像机中心都在该锥面的顶点,即点 $\mathbf{C}_P=(0,0,0,1)^T$ 上. △

**例 21.15** 锥面——两个中心都不在锥的顶点上. 选择 $F_{Q'Q}=\mathrm{diag}(1,1,0)$ 和 $\mathbf{t}_0=(0,2,0)^T$. 对此情形我们求得

$$S_{sym}=\begin{bmatrix}1&&&\\&1&&1\\&&0&\\&1&&0\end{bmatrix}.$$

这是方程为 $X^2+Y^2+2Y=0$ 或 $X^2+(Y+1)^2=1$ 的二次曲面. 它是一个平行于 Z 轴的圆柱面,从而射影等价于顶点在无穷远点$(0,0,1,0)^T$ 的锥面. 没有一个摄像机中心在该锥的顶点上. △

**例 21.16** 两张平面. 我们选择 $F_{Q'Q}=\mathrm{diag}(1,-1,0)$ ,$\mathbf{t}_0=(0,0,1)^T$,使得 $\mathbf{t}_0^T F_{Q'Q}=\mathbf{0}^T$. 对于这种情形,我们看到 $S_P=S_{sym}=\mathrm{diag}(1,-1,0,0)$,它表示一对平面. 对于这种情形,摄像机中心在两平面的交线上. △

**例 21.17** 单张平面. 我们选择 $F_{Q'Q}=\begin{bmatrix}1&1&0\\-1&0&0\\0&0&0\end{bmatrix}$ 和 $\mathbf{t}_0=(0,0,1)^T$. 那么 $S_{sym}=\mathrm{diag}(1,0,0,0)$,它表示单张平面 $x=0$. △

**例 21.18** 单条直线. 我们选择 $F_{Q'Q}=\mathrm{diag}(1,1,0)$和 $\mathbf{t}_0=(0,0,1)^T$. 对于这种情形,我们看到 $S_{sym}=\mathrm{diag}(1,1,0,0)$,它表示单条直线——Z 轴. 所有点和两个摄像机中心都在这条直线上. △

我们已经给出两个摄像机中心在直纹二次曲面的不同配置上的几个例子. 这些例子没有穷尽各种可能性,但可以通过例子验证所有可能的配置都可能发生.

**最小配置——七点** 如在节 10.1.1($p$192)中所看到的,由一般位置的点进行射影重构的最少点数是七. 该方法归结为解一个三次方程,它有一个或者有三个实解. 从临界曲面的观点来看,在一个配置中的七点和两个摄像机中心必须在一个二次曲面上(因而 9 点在一个二次曲面上). 如果该二次曲面是直纹的,那么将会有三个共轭解. 另一方面,如果它不是直纹曲面(例如一个椭球面)那么将仅有一个解. 这表示三次方程有一个和三个解的差别产生于点和摄像机中心所在二次曲面是直纹或非直纹的差别——代数和几何之间一种令人满意的关系.

## 21.3 Carlsson - Weinshall 对偶

第 19 章研究的摄像机和点之间的对偶性可以用来对偶化退化情形的结论,这将在本节

作解释. 这里我们将对 Carlsson－Weinshall 对偶做一个更正式的处理.

Carlsson－Weinshall 对偶的基础是方程

$$\begin{bmatrix} a & & & -d \\ & b & & -d \\ & & c & -d \end{bmatrix}\begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix} = \begin{bmatrix} X & & & -T \\ & Y & & -T \\ & & Z & -T \end{bmatrix}\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix},$$

左边的摄像机矩阵对应于中心为$(a^{-1},b^{-1},c^{-1},d^{-1})^{\mathsf{T}}$的摄像机. 我们感兴趣的是用摄像机的中心来描述摄像机. 这里表明如把摄像机中心与 3D 点对调,我们必须使摄像机和点的坐标互为倒数. 例如,点$(X,Y,Z,T)^{\mathsf{T}}$与中心为$(X^{-1},Y^{-1},Z^{-1},T^{-1})^{\mathsf{T}}$的摄像机对偶.

我们把中心为$\mathbf{C}=(a,b,c,d)^{\mathsf{T}}$的简化摄像机矩阵

$$P=\begin{bmatrix} a^{-1} & & & -d^{-1} \\ & b^{-1} & & -d^{-1} \\ & & c^{-1} & -d^{-1} \end{bmatrix} \tag{21.1}$$

记为$P_{\mathbf{C}}$. 现在,定义$\overline{\mathbf{X}}=(X^{-1},Y^{-1},Z^{-1},T^{-1})^{\mathsf{T}}$和$\overline{\mathbf{C}}=(a^{-1},b^{-1},c^{-1},d^{-1})^{\mathsf{T}}$,我们立刻可以验证下式成立

$$P_{\mathbf{C}}\mathbf{X}=P_{\overline{\mathbf{X}}}\overline{\mathbf{C}} \tag{21.2}$$

这样一来,这个变换交换了 3D 点和摄像机中心的结果. 因此,中心为 $\mathbf{C}$ 的摄像机作用于 $\mathbf{X}$ 与中心为 $\overline{\mathbf{X}}$ 作用于 $\overline{\mathbf{C}}$ 给出一样的结果.

这个观察导致下面的定义

**定义 21.19** 由

$$(X,Y,Z,T)^{\mathsf{T}}\rightarrow(YZT,\ ZTX,TXY,\ XYZ)^{\mathsf{T}}$$

给出的 $\mathbb{P}^3$ 到自身的映射称为 Carlsson-Weinshall 映射,并用 $\Gamma$ 表示. 在 $\Gamma$ 作用下点 $\mathbf{X}$ 的象记为 $\overline{\mathbf{X}}$. 在 $\Gamma$ 作用下的一个物体的象有时称为**对偶物体**.

Carlsson－Weinshall 映射是 **Cremona** 变换的一个例子. 关于 Cremona 变换更详细的信息,读者可以参考 Semple 和 Kneebone([Semple－79]).

**注释** 如果 $\mathbf{X}$ 的坐标中没有一个为零,那么我们可以用 XYZT 除以 $\overline{\mathbf{X}}$. 因此 $\Gamma$ 等价于$(X,Y,Z,T)^{\mathsf{T}}\mapsto(X^{-1},Y^{-1},Z^{-1},T^{-1})^{\mathsf{T}}$. 这是我们对该映射将常采用的形式. 当 $\mathbf{X}$ 中有一个坐标为零时,那么这个映射将按定义解释. 注意任何点$(0,Y,Z,T)^{\mathsf{T}}$(假定没有其他的坐标为零)被 $\Gamma$ 映射到点$(1,0,0,0)^{\mathsf{T}}$. 因此,这个映射不是一一对应的.

如果 $\mathbf{X}$ 中有两个坐标为零,那么$\overline{\mathbf{X}}=(0,0,0,0)^{\mathsf{T}}$,它是一个没有定义的点. 因此,$\Gamma$ 不是在所有点上有定义. 事实上. 没有一个方法可把 $\Gamma$ 连续地扩展到这些点.

定义一个**参考四面体**为顶点是 $\mathbf{E}_1=(1,0,0,0)^{\mathsf{T}}$,$\mathbf{E}_2=(0,1,0,0)^{\mathsf{T}}$,$\mathbf{E}_3=(0,0,1,0)^{\mathsf{T}}$,$\mathbf{E}_4=(0,0,0,1)^{\mathsf{T}}$的四面体. 如我们刚才看到的那样,除非在参考四面体的面上,$\Gamma$ 是一一对应的. $\Gamma$ 把参考四面体的一个面映射到它对面的顶点,而且在参考四面体的边缘上无定义. 下面我们研究 $\Gamma$ 作用于其他几何物体的方式.

**定理 21.20** Carlsson－Weinshall **映射 $\Gamma$ 以下列方式作用:**

**(1) 它把过一般位置上的两点 $\mathbf{X}_0$ 和 $\mathbf{X}_1$ 的一条直线映射成过 $\overline{\mathbf{X}}_0$ 和 $\overline{\mathbf{X}}_1$ 和四个参考顶点**

$\mathbf{E}_1,\cdots,\mathbf{E}_4$ 的三次绕线.

(2) 它把过任何 $\mathbf{E}_i$ 的一条直线映射为过同一点 $\mathbf{E}_i$ 的直线,但不包括在参考四面体的面上的直线,因为这样的直线将被映射到一个点.

(3) 它把过四个点 $\mathbf{E}_i$, $i=1,\cdots,4$ 的二次曲面 S 映射到过同样四个点的二次曲面(表示成 $\overline{S}$). 如果 S 是直纹二次曲面,那么 $\overline{S}$ 也是. 如果 S 是退化的,那么 $\overline{S}$ 也是.

**证明**

(1) 一条直线的参数方程是 $(X_0+a\theta, Y_0+b\theta, Z_0+c\theta, T_0+d\theta)^T$,而在这条直线上的点由 Carlsson-Weinshall 映射到点

$$((Y_0+b\theta)(Z_0+c\theta)(T_0+d\theta),\cdots,(X_0+a\theta)(Y_0+b\theta)(Z_0+c\theta))^T.$$

因此,该矢量的元素是 $\theta$ 的三次函数,从而该曲线是一条三次绕线. 现在,令 $\theta=-X_0/a$ 使项 $X_0+a\theta$ 变为 0,而对应的对偶点是 $((Y_0+b\theta)(Z_0+c\theta)(T_0+d\theta),0,0,0)^T\approx(1,0,0,0)^T$. 第一项是唯一不含 $(X_0+a\theta)$ 的项,因此是唯一没有消去的项. 这表示参考顶点 $\mathbf{E}_1=(1,0,0,0)^T$ 在三次绕线上. 类似的推理证明其他的点 $\mathbf{E}_2,\cdots,\mathbf{E}_4$ 也在三次绕线上. 注意一条三次绕线由六个点确定,而这条三次绕线由在它上面的已知六点:$\mathbf{E}_i$, $\overline{\mathbf{X}}_0$, $\overline{\mathbf{X}}_1$ 确定,其中 $\mathbf{X}_0$ 和 $\mathbf{X}_1$ 是定义该直线的任何两点.

(2) 我们仅对过点 $\mathbf{E}_1=(1,0,0,0)^T$ 的直线证明. 而对其他的 $\mathbf{E}_i$,类似的证明成立. 选择该直线上的另一点 $\mathbf{X}=(X,Y,Z,T)^T$,并使 $\mathbf{X}$ 不在参考四面体的任何面上. 则 $\mathbf{X}$ 没有零坐标. 在过 $(1,0,0,0)^T$ 和 $\mathbf{X}=(X,Y,Z,T)^T$ 的直线上的所有点的形式为 $(\alpha,Y,Z,T)^T$,其中 $\alpha$ 的值可变. 这些点由变换映射到 $(\alpha^{-1},Y^{-1},Z^{-1},T^{-1})^T$. 它表示过两点 $(1,0,0,0)^T$ 和 $\overline{\mathbf{X}}=(X^{-1},Y^{-1},Z^{-1},T^{-1})^T$ 的一条直线.

(3) 由于二次曲面 S 过所有点 $\mathbf{E}_i$, S 的所有对角元素必然都是零. 这表明该二次曲面的方程中没有坐标的平方项(例如 $X^2$). 因此该二次曲面的方程仅包含混合项(例如 XY, YZ 或 XT). 因此二次曲面 S 可以由方程 $aXY+bXZ+cXT+dYZ+eYT+fZT=0$ 来定义. 用 XYZT 除以该方程,我们得到 $aZ^{-1}T^{-1}+bY^{-1}T^{-1}+cY^{-1}Z^{-1}+dX^{-1}T^{-1}+eX^{-1}Z^{-1}+fX^{-1}Y^{-1}=0$. 由于 $\overline{\mathbf{X}}=(X^{-1},Y^{-1},Z^{-1},T^{-1})^T$,这是 $\overline{\mathbf{X}}$ 的元素的一个二次方程. 这样一来,$\Gamma$ 把二次曲面映射到二次曲面. 具体地说,设 S 用矩阵形式表示为

$$S=\begin{bmatrix}0 & a & b & c\\ a & 0 & d & e\\ b & d & 0 & f\\ c & e & f & 0\end{bmatrix},\text{ 那么 }\overline{S}=\begin{bmatrix}0 & f & e & d\\ f & 0 & c & b\\ e & c & 0 & a\\ d & b & a & 0\end{bmatrix}$$

并且 $\mathbf{X}^TS\mathbf{X}=0$ 蕴含着 $\overline{\mathbf{X}}^T\overline{S}\,\overline{\mathbf{X}}=0$. 如果 S 是直纹的,那么它包含过任一点(特别是过每一个 $\mathbf{E}_i$)的两条母线. 根据定理的第(2)部分,它们被映射到必然在 $\overline{S}$ 上的直线. 因此 $\overline{S}$ 是直纹的. 我们可以进一步证明 $\det S=\det\overline{S}$,由此推出:若 S 是非退化的($\det S\neq 0$),则 $\overline{S}$ 也是如此. 在非退化的情形,如果 S 是单叶双曲面,那么 $\det S>0$,由此推出 $\det\overline{S}>0$. 因此 $\overline{S}$ 也是单叶双曲面. □

$\Gamma$ 在其他几何实体的作用在练习(*p*398)中研究.

我们希望用与坐标无关的方式来解释对偶方程(21.2). 根据定义,矩阵 $P_C$ 的形式在

(21.1)给出，并把 $\mathbf{E}_i$ 映射到 $\mathbf{e}_i$ ($i=1,\cdots,4$). 从而图像 $P_C\mathbf{X}$ 可以看成是 $\mathbf{X}$ 相对于图像中射影基 $\mathbf{e}_i$ 的投影表示. 或者，$P_C\mathbf{X}$ 表示五条射线 $\overline{\mathbf{CE}_1},\cdots,\overline{\mathbf{CE}_4}$，$\overline{\mathbf{CX}}$ 集合的射影等价类. 因此，$P_C\mathbf{X}=P_{C'}\mathbf{X}'$ 的充要条件是由 $\mathbf{C}$ 到 $\mathbf{X}$ 和参考四面体的四个顶点的射线集合与由 $\mathbf{C}'$ 到 $\mathbf{X}'$ 和四个顶点射影的射线集合等价. 用先前引入的标记，我们可以用不同的形式把(21.2)记为

$$\langle \mathbf{C};\mathbf{E}_1,\ldots\mathbf{E}_4,\mathbf{X}\rangle=\langle \overline{\mathbf{X}};\mathbf{E}_1,\ldots,\mathbf{E}_4,\overline{\mathbf{C}}\rangle \tag{21.3}$$

**对偶性原则**

对偶的基础是(21.2)：$P_C\mathbf{X}=P_{\overline{\mathbf{X}}}\overline{\mathbf{C}}$，其中所用的标记与(21.2)一样.

记号 $P_C\mathbf{X}$ 表示 $\mathbf{X}$ 在规范图像坐标系下的射影坐标，该坐标系由参考四面体的角点的投影定义. 等价地，$P_C$ 可以看成表示五个被投影点 $P_C\mathbf{E}_i$ 和 $P_C\mathbf{X}_i$ 的射影等价类. 用本章的记号就是 $\langle \mathbf{C};\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}\rangle$. 因此，对偶关系可以写为

$$\langle \mathbf{C};\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}\rangle=\langle \overline{\mathbf{X}};\mathbf{E}_1,\cdots,\mathbf{E}_4,\overline{\mathbf{C}}\rangle \tag{21.4}$$

其中横扛表示 Carlsson－Weinshall 映射.

虽然 $P_C$ 由规范的射影基定义，但是用 $\mathbf{E}_1,\cdots,\mathbf{E}_4$ 作为参考四面体的顶点没有什么特别之处，它们只是不共面而已. 给定任何不共面的四个点，我们可以定义一个射影坐标系，其中这四个点就是形成射影坐标基一部分的 $\mathbf{E}_i$. 然后，Carlsson－Weinshall 映射可以在这个坐标系下定义. 所产生的映射称为 Carlsson－Weinshall 相对于给定参考四面体的映射.

为表述得更准确一点，应该看到在 $\mathbb{P}^3$ 中由五个点(而不是四个)定义一个射影坐标系. 事实上，使四个非共面点的坐标为 $\mathbf{E}_i$ 的射影坐标系不止一个(实际上是一个三参数的簇). 因此，Carlsson－Weinshall 映射相对于给定参考四面体的映射不是唯一的. 但是，相对于任何这样的坐标系，由定义 21.19 给出的映射都可以使用.

给定关于点集相对一个或更多的投影中心的投影命题或定理，我们可以引出一个对偶命题. 我们要求在被投影的四个点中，有可以形成一个参考四面体的四个非共面的点. 相对于这个参考四面体，在一般对偶映射下

(1) 点(不属于参考四面体)被映射到投影中心.

(2) 投影中心被映射到点.

(3) 直线被映射到三次绕线.

(4) 包含参考四面体的直纹二次曲面被映射到包含参考四面体的直纹二次曲面.

应该避免使用在参考四面体边界上的点，因为 Carlsson-Weinshall 映射对这些点没有定义. 把它当作一种转移表，我们可以把已经存在的关于点投影的定理对偶化，给出不需要再进行证明的新定理.

**注意** 仅仅是那些不属于参考四面体的点可以通过对偶映射到摄像机中心，这一点是重要的. 参考四面体的顶点仍然属于点集. 在实际应用对偶原理时，我们可以选择任何四点形成一个参考四面体，只要这四点不共面. 一般，在下一节叙述的定理中将假定(不总是明显地表述)：**所考虑的点的集合中有四点不共面**，它们可以视为参考四面体.

### 21.3.1 单视图的多义性

本节将说明如何应用对偶性由已知或明显的几何命题直接推导各种多义性重构的结

论.

**由五点计算摄像机投影矩阵**

对射影摄像机来说,用五组 3D－2D 的对应来计算它的投影矩阵是不充分的. 但是有趣的问题是:由五组点对应能确定什么?

为说明由 Carlsson 对偶性能推导出什么,考虑下面一个简单问题:什么时候两个点投影到图像中的同一点? 显然,答案是当两个点同在过摄像机中心的一条射线上时. 对偶化这样一个简单的现象,图 21.5 表明:被这五组 3D－2D 对应所约束的摄像机的中心必然在过这 5 个 3D 点的一条三次绕线上.

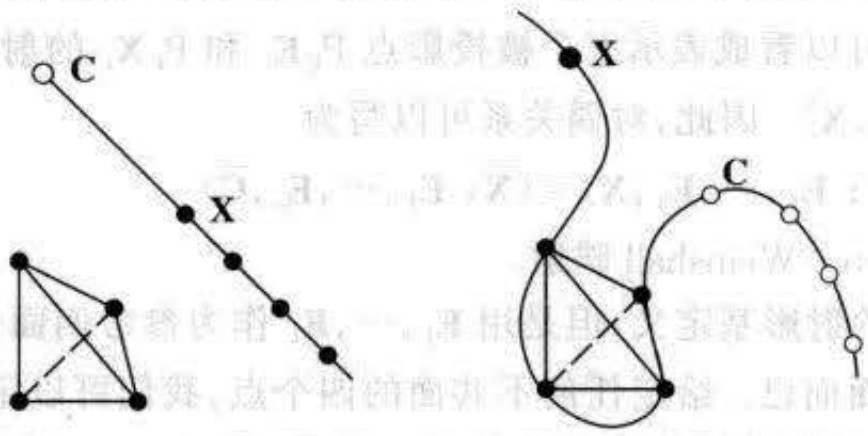



图 21.5　左:过 **C** 和 **X** 的直线上的任何点被投影中心 **C** 射影到同一点. 右:对偶命题——如果投影中心 **C** 在过 **X** 和参考四面体的顶点的三次绕线的任何位置上,那么该 5 点以同样的方式被投影(在射影等价的意义下). 因此摄像机被它的 5 个已知点的图像约束在一条三次绕线上.

**双眼单视界**

用类似的方式,我们可以计算由两个摄像机确定的双眼单视界的形式. 双眼单视界是指映射到两幅图像上的点是同一点的空间点集. 论据在图 21.6 中给出,并从关于直线的一个简单观察开始.

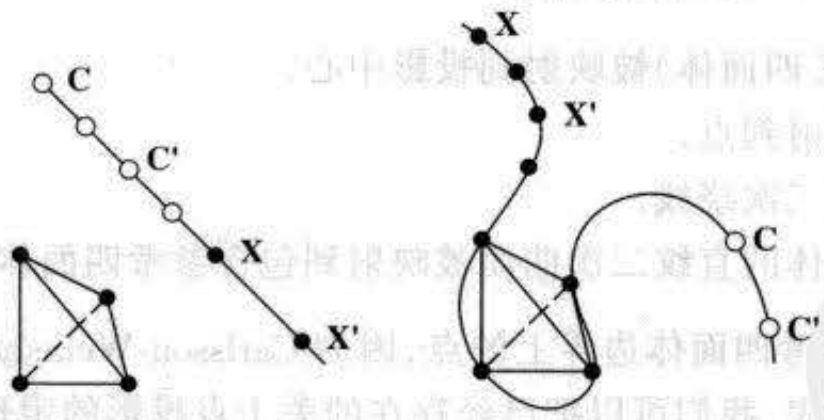



图 21.6　左:在过 **X** 和 **X**′的直线上的任何射影中心 **C**,**C**′,...将点 **X** 和 **X**′投射到同一射线上. 也就是说,对在直线上的所有 **C** 都有 $\langle \mathbf{C};\mathbf{E}_i,\mathbf{X}\rangle = \langle \mathbf{C};\mathbf{E}_i,\mathbf{X}'\rangle$. 右:对偶陈述——如果所有点在过 **C** 和 **C**′以及参考四面体的顶点的三次绕线上,那么相对于两个射影中心,它们被投影的方式一样. 也就是说,对在三次绕线上的所有 **X** 有 $\langle \mathbf{C};\mathbf{E}_i,\mathbf{X}\rangle = \langle \mathbf{C}';\mathbf{E}_i,\mathbf{X}\rangle$. 这条曲线称为两个投影中心的双眼单视界.

**结论 21.21　给定点 X 和 X′,满足**

$$\langle \mathbf{C};\ \mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}\rangle = \langle \mathbf{C};\ \mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}'\rangle$$

**的摄像机中心 C 的轨迹是过 X 和 X′的直线.**

这在图 21.6(左)中说明. 该陈述的对偶确定一对摄像机的双眼单视界(见图 21.6(右)).

**结论 21.22　给定投影中心 C 和 C′,它们与一个参考四面体的四个点 $\mathbf{E}_i$ 都不共线,则满足 $\langle \mathbf{C};\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}\rangle=\langle \mathbf{C}';\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}\rangle$ 的点集 X 是过 $\mathbf{E}_1,\cdots,\mathbf{E}_4$ 和两个射影中心 C 和 C′ 的一条三次绕线.**

**计算摄像机投影矩阵中的多义性**

最后,我们来考虑计算摄像机投影矩阵中的多义性. 它与双眼单视界有密切联系. 为了使它可视化,读者可以再一次参考图 21.6,虽然它不完全贴近现在的情况.

**结论 21.23　考虑一组摄像机中心 $\mathbf{C}_1,\cdots,\mathbf{C}_m$ 和一个点 $\mathbf{X}_0$ 都在一条直线上,并且令 $\mathbf{E}_i$($i=1,\cdots,4$)为一个参考四面体的顶点. 令 X 为另一个点. 那么两个配置**

$$\{\mathbf{C}_1,\cdots,\mathbf{C}_m,\mathbf{E}_1,\cdots,\mathbf{E}_4;\mathbf{X}\}\text{和}\{\mathbf{C}_1,\cdots,\mathbf{C}_m;\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}_0\}$$

**是图像等价配置的充要条件是 X 与 $\mathbf{X}_0$ 和摄像机在同一条直线上.**

根据定理 21.20,过渡到对偶命题时,直线变成过参考四面体四个顶点的一条三次绕线. 这样一来,结论 21.23 的对偶命题是:

**结论 21.24　点集 $\mathbf{X}_i$ 和一个摄像机中心 $\mathbf{C}_0$ 都在一条三次绕线上,同时这条绕线也过参考四面体的四个顶点 $\mathbf{E}_i$. 令 C 为任何其他的摄像机中心. 那么配置**

$$\{\mathbf{C};\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}_1,\cdots,\mathbf{X}_m\}\text{和}\{\mathbf{C}_0;\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}_1,\cdots,\mathbf{X}_m\}$$

**是图像等价的充要条件是 C 在同一条三次绕线上.**

因为 $\mathbf{E}_i$ 可以是任何不共面的四个点,而一条三次绕线不可能包含 4 个共面的点,我们可以将上一结论用下列形式叙述为:

**结论 21.25　令 $\mathbf{X}_1,\cdots,\mathbf{X}_m$ 为一组点而 $\mathbf{C}_0$ 为摄像机中心,它们都在一条三次绕线上. 那么对任何其他的摄像机中心 C ,配置**

$$\{\mathbf{C};\mathbf{X}_1,\cdots,\mathbf{X}_m\}\text{和}\{\mathbf{C}_0;\mathbf{X}_1,\cdots,\mathbf{X}_m\}$$

**是图像等价的充要条件是 C 在同一条三次绕线上.**

这在图 21.6(右)中说明. 它表明只要所有的点和一个摄像机中心在一条三次绕线上,摄像机的姿态就不能唯一确定. 这也给出了结论 21.6($p$382)的一个独立的证明,它也解决了以前尚未证明的情形.

用类似的方法,我们可以证明它是仅可能的两种多义性情形中的一种. 多义性发生的另一种情形是当所有的点和两个摄像机中心都在一个平面和一条直线的并集上. 这种情形的对偶发生在当这条直线过摄像机中心并与参考四面体的一个顶点相交时. 对这种情形,这条直线的对偶也是过同一参考顶点的一条直线(见定理 21.20),并且所有的点必须在这条直线或它所对的参考四面体的面上.

注意在这两个例子中,对偶性如何利用关于共线点投影这样一种直觉上显而易见的陈述推导出点在三次绕线上这样一个并不明显的结论.

### 21.3.2　两视图中的多义性

关于两视图中临界曲面的基本结论([Maybank-90])可以叙述如下.

**定理 21.26** **由两个摄像机中心和 $n$ 个点构成的一个配置 $\{\mathbf{C}_1,\mathbf{C}_2;\mathbf{X}_1,\cdots,\mathbf{X}_n\}$ 允许另一个重构存在的充要条件是两个摄像机中心 $\mathbf{C}_1,\mathbf{C}_2$ 和所有的点 $\mathbf{X}_j$ 都在一个直纹二次曲面上. 而且当另一个重构存在时,总存在第三个与它们都不相同的重构.**

我们可以直接写出如下的对偶陈述.

**定理 21.27** **由任何数目的摄像机和 6 点组成的一个配置 $\{\mathbf{C}_1,\cdots,\mathbf{C}_n;\mathbf{X}_1,\cdots,\mathbf{X}_6\}$ 有另一种重构存在的充要条件是所有摄像机中心 $\mathbf{C}_1,\cdots,\mathbf{C}_n$ 和所有点 $\mathbf{X}_1,\cdots,\mathbf{X}_6$ 都在一个直纹二次曲面上①. 而且,当另一个重构存在时,总存在第三个与之不同的重构.**

这个定理最初在[Maybank-98]中被证明,这里将给出定理 21.27 的一种证明仅为了强调怎样进行对偶性证明.

**证明** 考虑配置 $\{\mathbf{C}_1,\cdots,\mathbf{C}_n;\mathbf{X}_1,\cdots,\mathbf{X}_6\}$. 我们把点重新编号为 $\{\mathbf{C}_1,\cdots,\mathbf{C}_n;\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}_1,\mathbf{X}_2\}$,其中 $\mathbf{E}_1,\cdots,\mathbf{E}_4$ 是四个非共线点,并组成一个参考四面体的顶点. 如果这个配置还有另一个重构,那么存在另一个配置 $\{\mathbf{C}_1',\cdots,\mathbf{C}_n';\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}_1',\mathbf{X}_2'\}$ 使得对所有 $i=1,\cdots,n$ 和 $j=1,2$,有 $\langle\mathbf{C}_i;\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}_j\rangle=\langle\mathbf{C}_i';\mathbf{E}_1,\cdots,\mathbf{E}_4,\mathbf{X}_j'\rangle$. 用(21.3)把它对偶化得到:对所有 $i=1,\cdots,n$ 和 $j=1,2$ 有

$$\langle\overline{\mathbf{X}}_j;\mathbf{E}_1,\cdots,\mathbf{E}_4,\overline{\mathbf{C}}_i\rangle=\langle\overline{\mathbf{X}}_j';\mathbf{E}_1,\cdots,\mathbf{E}_4,\overline{\mathbf{C}}_i'\rangle.$$

现在应用定理 21.26 推导出两个摄像机中心 $\overline{\mathbf{X}}_j$,参考顶点 $\mathbf{E}_1,\cdots,\mathbf{E}_4$ 和点 $\overline{\mathbf{C}}_i$ 都在一个直纹二次曲面 $\overline{\mathrm{S}}$ 上. 应用定理 21.20(3)的逆对偶,我们可以看到点 $\mathbf{X}_1,\mathbf{X}_2$ 和摄像机中心 $\mathbf{C}_i$ 都在一个直纹二次曲面 S 上. 这就证明了本定理的必要性. 以类似的方式可证其充分性. □

根据节 19.2.4($p360$)的研究,定理 21.27 中令我们感兴趣的最小配置是 $n=3$. 在这种情形下总共有九个点(三个摄像机和六个点). 过这九个点,我们可以构造一个二次曲面(一个二次曲面由九个点确定). 如果该二次曲面是直纹二次曲面(在非退化情形为单叶双曲面),那么可能存在三种不同的重构. 否则,重构是唯一的. 与两视图中由七点重构类似,三视图中六点算法 19.1($p361$)需要解一个三次方程. 与七个点时一样,三次方程有一个实解还是三个实解可以由二次曲面是否为直纹来区别.

## 21.4 三视图临界配置

我们现在转向考虑三视图中可能产生的多义性配置.

在本节中我们用上标代替撇号来区别三个摄像机. 令 $\mathrm{P}^0,\mathrm{P}^1,\mathrm{P}^2$ 为三个摄像机而 $\{\mathbf{P}_i\}$ 为一个点集. 我们问在什么条件下存在另一个配置,它包含另外三个摄像机矩阵 $\mathrm{Q}^0,\mathrm{Q}^1,\mathrm{Q}^2$ 和点 $\{\mathbf{Q}_i\}$,使得对所有的 $i,j$ 有 $\mathrm{P}^j\mathbf{P}_i=\mathrm{Q}^j\mathbf{Q}_i$. 我们要求这两个配置是射影不等价的.

有各种特殊的多义性配置存在.

**点在一个平面上**

如果所有点都在一个平面上并且对所有 $i$ 有 $\mathbf{P}_i=\mathbf{Q}_i$,那么其中任何摄像机都可在不改变投影点的射影等价类条件下移动. 可以把 $\mathrm{P}^j$ 和 $\mathrm{Q}^j$ 的中心选择在预先指定的任何两个位

① 在这个陈述中假定 $\mathbf{X}_1,\cdots,\mathbf{X}_6$ 包括形成一个参考四面体的四个非共面的点并且其他两个点 $\mathbf{X}_j$ 和摄像机中心 $\mathbf{C}_i$ 都不在这个四面体上. 这个条件是否是本质的还没有解决.

置上并满足 $P^i\mathbf{P}_i = Q^i\mathbf{Q}_i$

**点在一条三次绕线上**

当所有的点以及其中一个摄像机，比如 $P^2$，在一条三次绕线上时，类似的多义性的情形会发生．对这种情形，我们可以选取 $Q^0 = P^0$，$Q^1 = P^1$ 和对所有的 $i$，$\mathbf{Q}_i = \mathbf{P}_i$．那么根据点在一条三次绕线时计算摄像机投影矩阵的多义性（节 21.1.2），对于三次绕线上的任何点 $\mathbf{C}_Q^2$，可以选取的中心在 $\mathbf{C}_Q^2$ 的摄像机矩阵 $Q^2$，使得对所有的 $i$ 有 $P^2\mathbf{P}_i = Q^2\mathbf{Q}_i$．

这些多义性的例子并不很有趣，因为它们不过是单视图中计算摄像机投影矩阵多义性的扩展．在以上例子的每种情形，点 $\mathbf{Q}_i$ 和 $\mathbf{P}_i$ 一样，多义性仅由摄像机相对于点所放置的位置而定．更有趣的多义性也会发生，如我们下面所要考虑的．

**三视图的一般多义性**

假定摄像机矩阵 $(P^0, P^1, P^2)$ 和 $(Q^0, Q^1, Q^2)$ 是固定的，我们希望寻找当 $i = 0,1,2$ 时使 $P^i\mathbf{P} = Q^i\mathbf{Q}$ 成立的所有点集．注意，这里我们试图复制两视图中两个摄像机矩阵集合都事先选定的情形．以后，我们将转向较少的约束，即事先仅选定一个摄像机集合的情形．

一个简单的观察是三视图的临界配置也是其中每对视图的临界配置的集合．因此人们自然地会想到临界配置 $\{P^0, P^1, P^2, \mathbf{P}_i\}$ 的点集就是每一个临界配置 $\{P^0, P^1, \mathbf{P}_i\}$，$\{P^1, P^2, \mathbf{P}_i\}$ 和 $\{P^0, P^2, \mathbf{P}_i\}$ 的交集．因为根据引理 21.10（$p$385），每一个这样的点集都是一个直纹二次曲面，人们可能会由此推断三视图的临界点集仅仅是这三个二次曲面的交．虽然它离事实不太远，但是这样推理有点混乱．这个推理遗漏的关键点是三对摄像机中的每一对摄像机所对应的共轭点也许不一样．

更精确地说，对应于临界配置 $\{P^0, P^1, \mathbf{P}_i\}$，存在一个共轭配置 $\{Q^0, Q^1, \mathbf{Q}_i^{01}\}$ 使得 $j = 0,1$ 时有 $P^j\mathbf{P}_i = Q^j\mathbf{Q}_i^{01}$．类似地，对应于临界配置 $\{P^0, P^2, \mathbf{P}_i\}$，存在一个共轭配置 $\{Q^0, Q^2, \mathbf{Q}_i^{02}\}$ 使得 $j = 0,2$ 时有 $P^j\mathbf{P}_i = Q^j\mathbf{Q}_i^{02}$．但是，点 $\mathbf{Q}_i^{02}$ 未必与 $\mathbf{Q}_i^{01}$ 一样．因此我们不能认定存在一个点 $\mathbf{Q}_i$ 使得对所有的 $i$ 和 $j = 0,1,2$ 有 $P^j\mathbf{P}_i = Q^j\mathbf{Q}_i$——至少不能直接认定．

现在，我们更贴近地来考虑这个问题．先考虑头一对摄像机 $\{P^0, P^1\}$ 和 $\{Q^0, Q^1\}$，引理 21.10（$p$385）告诉我们：若 $\mathbf{P}, \mathbf{Q}$ 是使 $P^j\mathbf{P} = Q^j\mathbf{Q}$ 的点，则 $\mathbf{P}$ 一定在由这些摄像机矩阵确定的一个二次曲面 $S_P^{01}$ 上．类似地点 $\mathbf{Q}$ 在 $S_Q^{01}$ 上．依此类推，考虑摄像机对 $\{P^0, P^2\}$ 和 $\{Q^0, Q^2\}$，我们发现点 $\mathbf{P}$ 一定在由这两个摄像机对确定的第二个二次曲面 $S_P^{02}$ 上．类似地，存在由摄像机对 $\{P^1, P^2\}$ 和 $\{Q^1, Q^2\}$ 定义的另一个二次曲面并且点 $\mathbf{P}$ 必定在这个二次曲面之上．这样一来，对 $j = 0,1,2$ 满足 $P^j\mathbf{P} = Q^j\mathbf{Q}$ 的点 $\mathbf{P}$ 和 $\mathbf{Q}$ 存在的必要条件是 $\mathbf{P}$ 必须在三个二次曲面的交上：$\mathbf{P} \in S_P^{01} \cap S_P^{02} \cap S_P^{12}$．我们就会看到这几乎是一个充分必要条件．

**结论 21.28** **令 $(P^0, P^1, P^2)$ 和 $(Q^0, Q^1, Q^2)$ 是摄像机矩阵的两个三元组并假定 $P^0 = Q^0$．对每一对 $(i, j) = (0,1), (0,2)$ 和 $(1,2)$，如在引理 21.10（$p$385）中一样，令 $S_P^{ij}$ 和 $S_Q^{ij}$ 为由摄像机矩阵对 $(P^i, P^j)$ 和 $(Q^i, Q^j)$ 定义的直纹二次临界曲面，则有**

**(1) 如果对 $j = 0,1,2$，存在使 $P^j\mathbf{P} = Q^j\mathbf{Q}$ 的点 $\mathbf{P}$ 和 $\mathbf{Q}$，那么 $\mathbf{P}$ 必须在交集 $S_P^{01} \cap S_P^{02} \cap S_P^{12}$ 上而 $\mathbf{Q}$ 必须在 $S_Q^{01} \cap S_Q^{02} \cap S_Q^{12}$ 上．**

**(2) 反之，如果 $\mathbf{P}$ 是二次曲面的交 $S_P^{01} \cap S_P^{02} \cap S_P^{12}$ 上的点，但不在包含三个摄像机中心**

$\mathbf{C}_Q^0$,$\mathbf{C}_Q^1$和 $\mathbf{C}_Q^2$ 的平面上，那么，存在在 $S_Q^{01} \cap S_Q^{02} \cap S_Q^{12}$ 上的一个点 $\mathbf{Q}$，使得对 $i=0,1,2$ 有 $P^i\mathbf{P} = Q^i\mathbf{Q}$.

注意条件 $P^0 = Q^0$ 不会对一般性产生任何限制，因为两个配置的射影坐标系($P^0, P^1, P^2$)和($Q^0, Q^1, Q^2$)是独立的. 我们可以容易地为第二个配置选择一个射影坐标系使这个条件成立. 做这个假设也仅仅是为了使我们可以在第二个摄像机集合的射影坐标系下考虑点 $\mathbf{P}$.

点 $\mathbf{P}$ 不在摄像机中心 $\mathbf{C}_Q^i$ 所在平面上的附加条件是必要的，读者可参考[Hartley－2000a]来了解该断言的理由. 但是，注意这种情形通常并不发生，因为三个二次曲面与三焦点平面的交将是空集，或在特殊情形为有限个点. 它真发生的可能性是三个摄像机中心 $\mathbf{C}_Q^0$,$\mathbf{C}_Q^1$和 $\mathbf{C}_Q^2$ 共线，在此情形，任何其他点与这三个摄像机中心共面.

**证明** 对第一部分，点 $\mathbf{P}$ 和 $\mathbf{Q}$ 在三个二次曲面的交上的事实由把引理 21.10($p$385)依次应用于每一对摄像机而得到(如定理陈述之前所指出的那样).

为证明第二部分，假定 $\mathbf{P}$ 在三个二次曲面的交上. 然后，把引理 21.10($p$386)应用于三个二次曲面 $S_P^{ij}$ 的每一个，推出存在点 $\mathbf{Q}^{ij}$ 使下列条件成立：

$$P^0\mathbf{P} = Q^0\mathbf{Q}^{01} \qquad P^1\mathbf{P} = Q^1\mathbf{Q}^{01}$$
$$P^0\mathbf{P} = Q^0\mathbf{Q}^{02} \qquad P^2\mathbf{P} = Q^2\mathbf{Q}^{02}$$
$$P^1\mathbf{P} = Q^1\mathbf{Q}^{12} \qquad P^2\mathbf{P} = Q^2\mathbf{Q}^{12}$$

这里容易被上标弄糊涂，但是每一行是把引理 21.10($p$385) 分别应用于三对摄像机矩阵所得的结论. 这些方程可以重新排成

$$P^0\mathbf{P} = Q^0\mathbf{Q}^{01} = Q^0\mathbf{Q}^{02}$$
$$P^1\mathbf{P} = Q^1\mathbf{Q}^{01} = Q^1\mathbf{Q}^{12}$$
$$P^2\mathbf{P} = Q^2\mathbf{Q}^{02} = Q^2\mathbf{Q}^{12}$$

条件 $Q^1\mathbf{Q}^{01} = Q^1\mathbf{Q}^{12}$ 意味着点 $\mathbf{Q}^{01}$ 和 $\mathbf{Q}^{12}$ 与 $Q^1$ 的摄像机中心 $\mathbf{C}_Q^1$ 共线. 这样一来，假定 $\mathbf{Q}^{ij}$ 两两不同，则它们必然在如图 21.7 所示的一种配置上. 从图中我们可以看到如果其中有两点相同，那么第三个点与其他两个点相同. 如果三个点不同，那么三个点 $\mathbf{Q}^{ij}$ 和三个摄像机中心 $\mathbf{C}_Q^i$ 共面，因为它们都在由 $\mathbf{Q}^{01}$ 与连接 $\mathbf{Q}^{02}$ 和 $\mathbf{Q}^{12}$ 的直线所确定的平面上. 因此三个点都在摄像机中心 $\mathbf{C}_Q^i$ 所属的平面上. 但是，由 $P^0\mathbf{P} = Q^0\mathbf{Q}^{01} = Q^0\mathbf{Q}^{02}$ 推出 $\mathbf{P}$ 必然与 $Q^{01}$ 和 $Q^{02}$ 在同一直线上，因而必然与摄像机中心 $\mathbf{C}_Q^i$ 在同一平面上. □

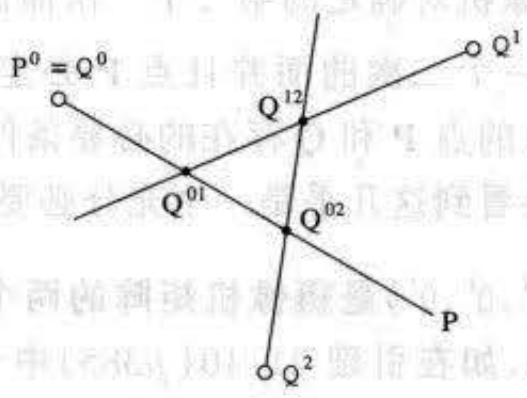



图 21.7 由三个摄像机中心和三个多义性点组成的配置. 如果 $\mathbf{Q}^{ij}$ 是不同的三个点，那么它们都在由摄像机中心 $\mathbf{C}_Q^i$ 组成的平面上.

一般，三个二次曲面的交集由八点组成. 在此情形，对应于两个三元组的摄像机矩阵的

临界集合仅由这 8 点组成. 但是,有时可以通过选择摄像机矩阵使三个二次曲面相交于一条曲线. 如果这三个二次曲面 $S_P^i$ 线性相关,那么这种情形将发生. 例如如果 $S_P^{12}=\alpha S_P^{01}+\beta S_P^{02}$,那么,任何满足 $\mathbf{P}^T S_P^{01}\mathbf{P}=0$ 和 $\mathbf{P}^T S_P^{02}\mathbf{P}=0$ 的 $\mathbf{P}$ 也将满足 $\mathbf{P}^T S_P^{12}\mathbf{P}=0$. 因此,这三个二次曲面的交集与它们中两个二次曲面的交集一样,它通常将是一条四次的空间曲线.

**一个例子**

作为多义性的一个具体的例子,考虑下面的配置. 令

$$P^0=[I|\mathbf{0}] \qquad Q^0=[I|\mathbf{0}]$$

$$P^1=\begin{bmatrix}1 & & & -1\\ & 1 & & -1\\ & & 1 & -1\end{bmatrix} \qquad Q^1=\begin{bmatrix}1 & & & 1\\ & -1 & & 0\\ & & -1 & 1\end{bmatrix}$$

$$P^2=\begin{bmatrix}1 & & & 1\\ & 1 & & 1\\ & & 1 & -1\end{bmatrix} \qquad Q^2=\begin{bmatrix}1 & & & 0\\ & 1 & & -1\\ & & -1 & 1\end{bmatrix}$$

对于这种情形,我们可以验证 $S_P^{01}=S_P^{02}$,它们都表示二次曲面 $XY=Z$. 该二次曲面与 $S_P^{12}$ 的交是一条曲线. 事实上,对任何 $t$,令 $Y(t)=1-t^2\pm\sqrt{1-t^2+t^4}$ 并且

$$\mathbf{P}_t=(t,Y(t),tY(t),1)^T$$

$$\mathbf{Q}_t=(t,Y(t),tY(t),(Y(t)-t)/(1+t))^T \tag{21.5}$$

那么我们可以验证对所有的 $i$ 和 $t$ 有 $P^i\mathbf{P}_t=Q^i\mathbf{Q}_t$. 我们同时可以验证所有三个摄像机中心 $\mathbf{C}_P^0=(0,0,0,1)^T$,$\mathbf{C}_P^1=(1,1,1,1)^T$,$\mathbf{C}_P^2=(-1,-1,1,1)^T$ 都在曲线 $\mathbf{P}_t$ 上.

注意这个例子看上去可能有一点特殊,因为两个二次曲面相等. 但是,这种情形的特殊性仅因为我们事先选择了所有六个摄像机矩阵. 用这个例子我们现在可以对三个摄像机任何配置描述一种临界集合.

发现这个例子的方法是由摄像机矩阵 $P^i$ 开始,然后计算所需要的基本矩阵 $F_{QQ}^{10}$ 和 $F_{QQ}^{20}$ 使它们保证二次曲面 $S_P^{01}$ 和 $S_P^{02}$ 具有所要求的形式. 由基本矩阵我们可以用标准方法计算矩阵 $Q^i$.

**定理 21.29** **给定中心不共线的三个摄像机($P^0,P^1,P^2$),存在(至少)一条由两个直纹二次曲面的交组成的四次曲线 $\mathbf{P}_t$,该曲线包含三个摄像机的中心并且这三个摄像机中心不能由它们产生的投影来唯一地重构. 具体地说,存在另外三个摄像机 $Q^i$ 和另外一条四次曲线 $\mathbf{Q}_t$ 使得对所有的 $i$ 和 $t$,**

$$P^i\mathbf{P}_t=Q^i\mathbf{Q}_t$$

**并使两个配置($P^0,P^1,P^2,\mathbf{P}_t$)和非($Q^0,Q^1,Q^2,\mathbf{Q}_t$)非射影等价.**

**证明** 证明相当简单. 因为这三个摄像机中心不共线,我们可以用一个射影变换(如果必要)将它们变换成前例中的摄像机中心 $\mathbf{C}_P^0=(0,0,0,1)^T$,$\mathbf{C}_P^1=(1,1,1,1)^T$,$\mathbf{C}_P^2=(-1,-1,1,1)^T$. 现在,应用结论 21.8($p$384)我们可以假定这三个摄像机等价于例子中的三个摄像机 $P^i$. 仿照前例,选择 $Q^i$,$\mathbf{P}_t$ 和 $\mathbf{Q}_t$ 即给出所要求的重构多义性. □

值得注意的是定理 21.29 中的三个特殊摄像机的临界曲线甚至对于摄像机矩阵固定的

情形也是不唯一的——存在一条这样的六参数簇的曲线，因为任何把这三个摄像机中心映射为它们自己的射影变换将把这条临界曲线映射成另一条临界曲线. 总计起来，我们已经辨认了下面的几种临界配置

(1) 包含三个摄像机中心的一个六参数的四次曲线簇

(2) 任何不含摄像机中心的平面

(3) 过三个摄像机中心之一的任何三次绕线.

## 21.5 七点的多义性视图

按下面做法可以把定理 21.29 对偶化，

- 三个不共线的摄像机中心换成不在一条绕线的七个点.
- 两个直纹二次曲面的交仍然是两个直纹二次曲面的交.
- 七个点中至少包含起着参考四面体作用的四个非共面点的集合.

**定理 21.30** **给定不在一条三次绕线的七个非共面的点 $\mathbf{P}_j$，存在一条由两个二次曲面的交组成的曲线 $\gamma$，使得中心 $\mathbf{C}_{\mathrm{P}'}$ 在曲线 $\gamma$ 上的任何数目的摄像机 $\mathrm{P}'$ 给出的 $\mathbf{P}_j$ 的投影都不足以唯一地确定点 $\mathbf{P}_j$ 的射影结构. 具体地说，存在一个不等价的点集 $\mathbf{Q}_j$ 和摄像机集 $\mathrm{Q}'$ 使得对所有 $i$ 和 $j$ 有 $\mathrm{P}'\mathbf{P}_j = \mathrm{Q}'\mathbf{Q}_j$.**

简言之，这个定理给出了七点集合的临界视点位置的描述. 如果摄像机沿一条特殊的临界曲线移动，它给出的图像不能确定七点集合的射影结构. 当然这里假定不存在对摄像机内参数标定的约束.

### 21.5.1 文献

[Buchanan－88]使计算机视觉界注意到三次绕线是摄像机投影矩阵的临界曲线. 读者可以参考[Maybank－90]和书[Maybank－93]得到更多的关于两视图的临界集合的信息. 在这两种情形中临界点被发现的时间还要早. 事实上[Buchanan－88]就提示读者参考德文摄影测量术的文献[Krames－42，Rinner－72].

这里没有提到三幅或更多幅视图中有关直线的临界配置问题，但这个论题已在[Buchanan－92]中作过研究. 另外，由直线进行线性重构的临界配置(线性线丛)已由[Stein－99]指出.

六点集的临界摄像机位置的问题由[Maybank－90]用直接论证的方法作过研究. 未发表的报告[Shashua－96]研究过三视图中的临界配置. 本章关于三视图临界配置的处理在[Hartley－2000a]中有更详细的介绍.

### 21.5.2 注释和练习

(1) 填充下面的证明梗概的细节，要证明的是：由一个单叶双曲面和在该双曲面上的两个点组成的任何两个配置是射影等价的(通过 $\mathbb{P}^3$ 的射影变换)，只要这两对点都在或都不在同一条母线上.

由于任何单叶双曲面都与 $X^2+Y^2-Z^2=1$ 射影等价，任何两个单叶双曲面都相互

射影等价，也与由 $Z=XY$ 给定的双曲面射影等价．定义一个 1D 射影变换 $h_X(X)=X'=(aX+b)/(cX+d)$．可算得

$$\begin{bmatrix} a & & & b \\ & d & c & \\ & b & a & \\ c & & & d \end{bmatrix}\begin{pmatrix} X \\ Y \\ XY \\ 1 \end{pmatrix}=\begin{pmatrix} X' \\ Y \\ X'Y \\ 1 \end{pmatrix}$$

这是一个把曲面 $Z=XY$ 变为自己的 3D 射影变换．把它与 Y 的一个相似变换复合，我们找到把 $(X,Y,XY,1)^T$ 变成 $(X',Y',X'Y',1)^T$ 同时使二次曲面 $Z=XY$ 保持不变的一个射影变换．由于 $h_X$ 和 $h_Y$ 是任意的 1D 射影变换，这给出了足够的自由度把任何两点映射到其他两点．

(2)证明 Γ 把一条与参考四面体一条边相交的直线映射成一条二次曲线．

(3)证明 Γ 把一条与参考四面体两条对边相交的直线映射成与同样两条边相交的一条直线．

(4)这些配置如何与计算摄像机投影矩阵的退化配置相关联，如图 21.3($p$381)所示．

(5)当所有点和两个摄像机都在一个直纹二次曲面上时，出现两视图的退化．给定在欧氏立方体上的八个角点和两个摄像机中心，证明这十个点总在一个二次曲面上．如果它是一个直纹二次曲面，那么该配置是退化的，因而不可能由 8 点得到重构．研究在什么条件下二次曲面是直纹的．**提示**：过立方体的顶点存在一个两参数的二次曲面簇．这个两参数簇看上去像什么？

# 附录 1　张量记号

在计算机视觉中，张量记号不常使用，所以在此简要地介绍一下它的应用．读者若查阅[Triggs－95]可得到更多的细节．为简单起见，这些概念将在低维射影空间（而不在它们的最一般形式）的范围内加以解释．但是，这些概念适用于任意维矢量空间．

考虑 2 维射影空间 $\mathbb{P}^2$ 的一组基矢量 $\mathbf{e}_i(i=1,\cdots,3)$．我们将把指标写成下标，其理由在下文中将会自明．在这组基下，$\mathbb{P}^2$ 中的点可以用一组坐标 $x^i$ 表示为 $\sum_{i=1}^{3} x^i \mathbf{e}_i$．我们用上标来表示该坐标．令 $\mathbf{x}$ 表示该坐标的三元组：$\mathbf{x}=(x^1,x^2,x^3)^{\mathrm{T}}$．

现在，考虑坐标轴的变化，即由新的基矢量 $\hat{\mathbf{e}}_j$ 代替基矢量 $\mathbf{e}_i$，其中 $\hat{\mathbf{e}}_j=\sum_i H_j^i \mathbf{e}_i$，而 H 是元素为 $H_j^i$ 的基变换矩阵．如果 $\hat{\mathbf{x}}=(\hat{x}^1,\hat{x}^2,\hat{x}^3)^{\mathrm{T}}$ 是该矢量在新基下的坐标，那么，我们可以验证 $\hat{\mathbf{x}}=\mathrm{H}^{-1}\mathbf{x}$．因此，如果基矢量按 H 变换，那么点的坐标按其逆变换 $\mathrm{H}^{-1}$ 变换．

下一步，考虑 $\mathbb{P}^2$ 的一条直线，它在原基下的坐标用 $\mathbf{l}$ 表示．可以验证，在新基下坐标为 $\hat{\mathbf{l}}=\mathrm{H}^{\mathrm{T}}\mathbf{l}$．因此，直线的坐标按 $\mathrm{H}^{\mathrm{T}}$ 变换．

再举一个例子，令 P 是表示射影（或矢量）空间之间的映射的一个矩阵．如果 G 和 H 分别表示定义域和值域空间的基变换．那么，在新基下，映射由一个新的矩阵 $\hat{\mathrm{P}}=\mathrm{H}^{-1}\mathrm{PG}$ 表示．注意在这些例子中变换有时用 H 或 $\mathrm{H}^{\mathrm{T}}$ 表示，而有时用 $\mathrm{H}^{-1}$ 表示．

上述三个坐标变换的例子可以显式地表示如下：

$$\hat{x}^i=(H^{-1})^i_j x^j \qquad \hat{l}_i=H_i^j l_j \qquad \hat{P}^i_j=(H^{-1})^i_k G^l_j P^k_l,$$

其中我们采用了张量求和约定，即在一个乘积中，当上标和下标重复出现时，表示对该指标在整个范围内求和．注意那些写成上标的指标按 $\mathrm{H}^{-1}$ 变换，而那些写成下标的指标按 H（或 G）变换．并且在张量记号中，按 H 变换的指标和按 $\mathrm{H}^{\mathrm{T}}$ 变换的指标之间没有区别．一般，张量的指标或者按 H 或者按 $\mathrm{H}^{-1}$ 来变换——事实上，这是张量的特征．按 H 变换的指标称为**协变**指标并写成下标；按 $\mathrm{H}^{-1}$ 变换的指标则称为**逆变**指标并写成上标．指标的数目称该张量的**阶**．在指标上进行累加，例如 $H_i^j l_j$，称为**收缩**，在此例中称张量 $H_i^j$ 按直线 $l_j$ 收缩．

## A1.1　张量 $\varepsilon_{rst}$

对 $r,s,t=1,\cdots,3$，张量 $\varepsilon_{rst}$ 定义如下：

$$\varepsilon_{rst}=\begin{cases} 0 & \text{除非 } r,s,t \text{ 互不相同} \\ +1 & \text{如果 } rst \text{ 是 123 的偶置换} \\ -1 & \text{如果 } rst \text{ 是 123 的奇置换} \end{cases}$$

张量 $\varepsilon_{ijk}$（或与它对应的逆变张量 $\varepsilon^{ijk}$）与两个矢量的叉乘有关．如果 $\mathbf{a}$ 和 $\mathbf{b}$ 是两个矢量，而 $\mathbf{c}=\mathbf{a}\times\mathbf{b}$ 是它们的叉乘，那么下面的公式可以轻易地得到验证．

$$c_i=(\boldsymbol{a}\times\boldsymbol{b})_i=\varepsilon_{ijk}a^j b^k.$$

与此相关的是用于表示反对称矩阵 $[\mathbf{a}]_\times$ 的表达式（A3.4）．用张量的记号我们可以将它记成

$$([\mathbf{a}]_\times)_{ik} = \varepsilon_{ijk} a^j .$$

因此，我们看到，如果 **a** 是一个逆变矢量，那么 $[\mathbf{a}]_\times$ 是一个具有两个协变指标的矩阵. 对 $[\mathbf{v}]_\times$ 也有类似的公式成立，其中 **v** 是协变量.

最后，张量 $\varepsilon_{ijk}$ 与行列式有关：对三个逆变张量 $a^i$，$b^j$ 和 $c^k$ 来说，我们可以验证 $a^i b^j c^k \varepsilon_{ijk}$ 是以 $a^i$，$b^j$ 和 $c^k$ 为行矢量的一个 $3\times3$ 矩阵的行列式.

## A1.2 三焦点张量

三焦点张量 $\mathcal{T}_i^{jk}$ 有一个协变和两个逆变指标. 对于矢量及矩阵，例如 $x^i$，$l_i$ 和 $P_j^i$，可以用标准的线性代数记号来书写变换规则，例如 $\mathbf{x}' = \mathrm{H}\mathbf{x}$. 但是，对有三个或更多指标的张量而言，这样做就不方便了. 在处理三焦点张量时，除采用张量记号外实在没有其它的选择.

**变换规则** 三焦点张量的指标安排蕴含下列变换规则：

$$\hat{\mathcal{T}}_i^{jk} = F_i^r (G^{-1})_s^j (H^{-1})_t^k \, \mathcal{T}_r^{st}, \tag{A1.1}$$

它对应于三幅图像中基的变化. 值得指出的是这里可能产生一个混淆. 变换规则(A1.1)表明当三幅图像中基发生变换时张量如何变换. 通常我们关心的是点坐标的变换. 因此，如果 F′，G′和 H′表示图像中的坐标变换，即 $\hat{x}^j = F_i'^j x^i$ 并且 G′和 H′在其它图像上也类似定义，那么变换规则可以写成

$$\hat{\mathcal{T}}_i^{jk} = (F'^{-1})_i^r G'^j_s H'^k_t \, \mathcal{T}_r^{st} .$$

**张量图** 一个矢量 **x** 可以视为排列成一列或一行的一组数，而一个矩阵 H 可以视为 2 维数组. 类似地，有三个指标的张量可以视为 3 维数组. 特别地，三焦点张量可视为一个 $3\times3\times3$ 单元立方体，如图 A1.1 所示.

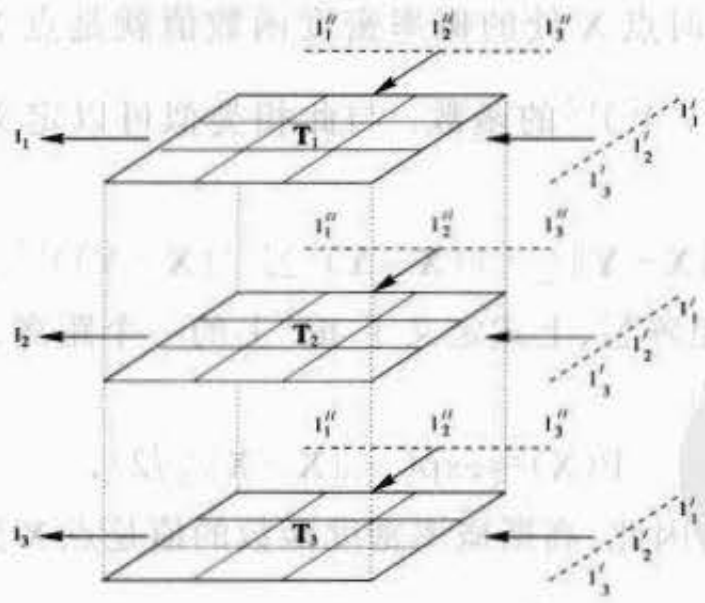



图 A1.1 三焦点张量的 3 维表示. ——图取自 Faugeras 和 Papadopoulo[Faugeras－97]. 图中给出 $l_i = l_j' l_k'' \mathcal{T}_i^{jk}$，它是该张量对 $\mathbf{l}'$ 和 $\mathbf{l}''$ 的收缩并产生一条直线 $\mathbf{l}$. 在伪矩阵记号中它可以写为 $l_i = \mathbf{l}'^{\mathrm{T}} \mathrm{T}_i \mathbf{l}''^{\mathrm{T}}$，其中 $(\mathrm{T}_i)_{jk} = \mathcal{T}_i^{jk}$.

# 附录 2　高斯(正态)分布与卡方分布

## A2.1　高斯概率分布

给定随机变量 $x_i(i=1,\cdots,N)$ 的矢量 $\mathbf{X}$,它的均值为 $\overline{\mathbf{X}}=E[\mathbf{X}]$,其中 $E[\cdot]$ 表示期望值,而 $\Delta\mathbf{X}=\mathbf{X}-\overline{\mathbf{X}}$,其协方差矩阵 $\Sigma$ 为如下 $N\times N$ 的矩阵

$$\Sigma=E[\Delta\mathbf{X}\Delta\mathbf{X}^{\mathrm{T}}],$$

其中 $\Sigma_{ij}=E[\Delta x_i\Delta x_j]$. 矩阵 $\Sigma$ 的对角元是单个变量 $x_i$ 的方差,而非对角元是交叉协方差的值.

如果 $\mathbf{X}$ 的概率分布密度形如

$$P(\overline{\mathbf{X}}+\Delta\mathbf{X})=(2\pi)^{-N/2}\det(\Sigma^{-1})^{1/2}\exp(-(\Delta\mathbf{X})^{\mathrm{T}}\Sigma^{-1}(\Delta\mathbf{X})/2),\qquad (A2.1)$$

其中 $\Sigma^{-1}$ 是某个半正定矩阵,那么,这变量 $x_i$ 称为遵循一个联合高斯(Gauss)分布. 可以验证 $\overline{\mathbf{X}}$ 和 $\Sigma$ 是该分布的均值和协方差. 高斯分布由它的均值和协方差唯一确定. 因子 $(2\pi)^{-N/2}\det(\Sigma^{-1})^{1/2}$ 仅仅是使分布的总积分等于 1 的归一化因子.

在 $\Sigma$ 为纯量矩阵 $\Sigma=\sigma^2\mathrm{I}$ 的特殊情形,高斯概率密度函数取下面的简单形式

$$\mathrm{P}(\mathbf{X})=(\sqrt{2\pi}\sigma)^{-N}\exp(-\sum_{i=1}^{N}(x_i-\overline{x}_i)^2/2\sigma^2),$$

其中 $\mathbf{X}=(x_1,x_2,\cdots,x_N)^{\mathrm{T}}$. 这种分布称为**各向同性高斯分布**.

**Mahalanobis 距离**　注意此时点 $\mathbf{X}$ 处的概率密度函数值就是点 $\mathbf{X}$ 到均值 $\overline{\mathbf{X}}=(\overline{x}_1,\overline{x}_2,\cdots,\overline{x}_N)^{\mathrm{T}}$ 的欧氏距离 $(\sum_{i=1}^{N}(x_i-\overline{x}_i)^2)^{1/2}$ 的函数. 与此相类似可以定义两个矢量 $\mathbf{X}$ 和 $\mathbf{Y}$ 之间的 **Mahalanobis 距离**为

$$\|\mathbf{X}-\mathbf{Y}\|_{\Sigma}=((\mathbf{X}-\mathbf{Y})^{\mathrm{T}}\Sigma^{-1}(\mathbf{X}-\mathbf{Y}))^{1/2}.$$

我们可以验证对一个正定矩阵 $\Sigma$,上式定义了 $\mathbb{R}^N$ 上的一个距离. 用这个记号,高斯概率密度函数的一般形式可以写为

$$\mathrm{P}(\mathbf{X})\approx\exp(-\|\mathbf{X}-\overline{\mathbf{X}}\|_{\Sigma}^2/2),$$

其中归一化因子已被省略. 因此,高斯概率密度函数的值是点 $\mathbf{X}$ 到均值 $\overline{\mathbf{X}}$ 的 Mahalanobis 距离的一个函数.

**坐标变化**　因为 $\Sigma$ 是对称正定的,它可以写为 $\Sigma=\mathrm{U}^{\mathrm{T}}\mathrm{D}\mathrm{U}$,其中 U 是一个正交矩阵,$\mathrm{D}=(\sigma_1^2,\sigma_2^2,\ldots,\sigma_N^2)$ 是对角矩阵. 记 $\mathbf{X}'=\mathrm{U}\mathbf{X}$ 和 $\overline{\mathbf{X}}'=\mathrm{U}\overline{\mathbf{X}}$,并代入(A2.1)得到

$$\begin{aligned}\exp(-(\mathbf{X}-\overline{\mathbf{X}})^{\mathrm{T}}\Sigma^{-1}(\mathbf{X}-\overline{\mathbf{X}})/2)&=\exp(-(\mathbf{X}'-\overline{\mathbf{X}}')^{\mathrm{T}}\mathrm{U}\Sigma^{-1}\mathrm{U}^{\mathrm{T}}(\mathbf{X}'-\overline{\mathbf{X}}')/2)\\&=\exp(-(\mathbf{X}'-\overline{\mathbf{X}}')^{\mathrm{T}}\mathrm{D}^{-1}(\mathbf{X}'-\overline{\mathbf{X}}')/2).\end{aligned}$$

因此,由 $\mathbf{X}$ 到 $\mathbf{X}'=\mathrm{U}\mathbf{X}$ 的正交坐标变换把一个一般的高斯概率密度函数变换为具有对角协方差矩阵的高斯概率密度函数. 在每一方向上再用 $\sigma_i$ 进行缩放可以进一步将它变换为一

个各向同性的高斯分布. 等价地说,可以应用坐标变换把 *Mahalanobis* 距离变换成普通的欧氏距离.

## A2.2 卡方分布

$\chi_n^2$(卡方)分布是 $n$ 个独立高斯随机变量的平方和的分布. 当应用于有非奇异协方差矩阵$\Sigma$的高斯随机矢量 $\mathbf{v}$ 时,$(\mathbf{v}-\bar{\mathbf{v}})^{\mathrm{T}}\Sigma^{-1}(\mathbf{v}-\bar{\mathbf{v}})$的值满足 $\chi_n^2$ 分布,其中 $n$ 是 $\mathbf{v}$ 的维数. 如果该协方差矩阵$\Sigma$是奇异的,那么我们必须用伪逆$\Sigma^{+}$替代$\Sigma^{-1}$. 此时

**• 如果 $\mathbf{v}$ 是均值为 $\bar{\mathbf{v}}$ 并且协方差矩阵为$\Sigma$的高斯随机矢量,那么值$(\mathbf{v}-\bar{\mathbf{v}})^{\mathrm{T}}\Sigma^{+}(\mathbf{v}-\bar{\mathbf{v}})$满足 $\chi_r^2$ 分布,其中 $r=\mathrm{rank}\Sigma$.**

累积 $\chi^2$ 分布定义为 $F_n(k^2)=\int_0^{k^2}\chi_n^2(\xi)d\xi$. 它表示随机变量 $\chi_n^2$ 的值小于$k^2$ 的概率. 当 $n=1,\cdots,4$ 时,$\chi_n^2$ 分布和累积 $\chi_n^2$ 逆分布的图在图 A2.1 中给出. 计算累积 $\chi_n^2$ 分布 $F_n(k^2)$的程序在[Press-88]中给出. 因为它是单调增函数,我们可以通过任何简单的技术(例如细分)来计算逆函数,它们的值在表 A2.1 中列出(与图 A2.1 相比较).

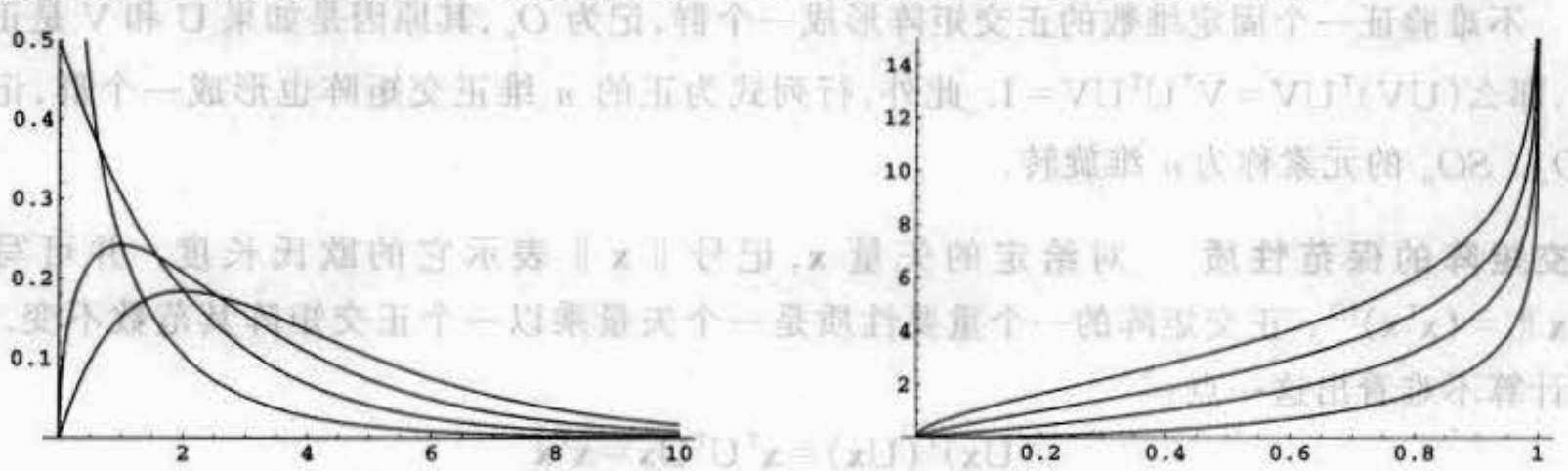



**图 A2.1 $n=1,\cdots,4$ 时的 $\chi_n^2$ 分布(左)和累积 $\chi_n^2$ 逆分布 $F_n^{-1}$(右). 两种情形都由下至上给出 $n=1,\cdots,4$ 时的图(在水平轴的中点).**

| $n$ | $\alpha=0.95$ | $\alpha=0.99$ |
|---|---|---|
| 1 | 3.84 | 6.63 |
| 2 | 5.99 | 9.21 |
| 3 | 7.81 | 11.34 |
| 4 | 9.49 | 13.28 |

**表 A2.1 自由度为 $n$ 的累积$\chi^2$ 分布 $F_n(k^2)$等于 $\alpha$ 时所对应的$k^2$ 的值,即 $k^2=F_n^{-1}(\alpha)$,其中 $\alpha$ 是概率.**

# 附录 3　数值算法

本附录讨论在本书中所用到的各种数值算法.

## A3.1　正交矩阵

方阵 U 称为**正交**的,如果它的转置矩阵等于它的逆矩阵——即 $U^TU=I$,其中 I 是恒等矩阵. 它意味着 U 的列矢量都有单位范数并且是正交的. 它也可以记为 $\mathbf{u}_i^T\mathbf{u}_j=\delta_{ij}$. 由条件 $U^TU=I$,我们不难推出 $UU^T=I$. 所以 U 的行矢量也都有单位范数并且是正交的. 我们再仔细考虑方程 $U^TU=I$. 两边取行列式得到方程 $(\det U)^2=1$,因为 $\det U=\det U^T$. 因此,如果 U 是正交的,那么 $\det U=\pm 1$.

不难验证一个固定维数的正交矩阵形成一个群,记为 $O_n$,其原因是如果 U 和 V 是正交的,那么 $(UV)^TUV=V^TU^TUV=I$. 此外,行列式为正的 $n$ 维正交矩阵也形成一个群,记为 $SO_n$. $SO_n$ 的元素称为 $n$ 维旋转.

**正交矩阵的保范性质**　对给定的矢量 $\mathbf{x}$,记号 $\|\mathbf{x}\|$ 表示它的欧氏长度. 并可写为 $\|\mathbf{x}\|=(\mathbf{x}^T\mathbf{x})^{1/2}$. 正交矩阵的一个重要性质是一个矢量乘以一个正交矩阵其范数不变. 通过计算不难看出这一点:

$$(U\mathbf{x})^T(U\mathbf{x})=\mathbf{x}^TU^TU\mathbf{x}=\mathbf{x}^T\mathbf{x}.$$

一个矩阵的 QR 分解通常指将一个矩阵 A 分解为乘积 A = QR,其中 Q 是正交矩阵,R 是上三角矩阵. 字母 R 表示**右**,指上三角矩阵. 与 QR 分解类似的有 QL,LQ,RQ 分解,其中 L 表示**左**或下三角矩阵. 这里只讨论矩阵的 RQ 分解,因本书实际上最常用到它. 最重要的情形是 3×3 矩阵的分解,我们将在下节中集中讨论它.

### A3.1.1　Givens 旋转和 RQ 分解

一个 3 维 Givens 旋转是绕三个坐标轴中的一个轴进行的旋转. 这三个 Givens 旋转是

$$\mathsf{Q}_x=\begin{bmatrix}1 & & \\ & c & -s\\ & s & c\end{bmatrix}\quad \mathsf{Q}_y=\begin{bmatrix}c & & s\\ & 1 & \\ -s & & c\end{bmatrix}\quad \mathsf{Q}_z=\begin{bmatrix}c & -s & \\ s & c & \\ & & 1\end{bmatrix},\tag{A3.1}$$

其中 $c=\cos(\theta)$ 而 $s=\sin(\theta)$,$\theta$ 为某角度,而空白表示零元素.

在一个 3×3 矩阵 A 右边乘以 $\mathsf{Q}_z$ 的效果是保持 A 的最后一列不变,而其前两列用原来的两列的线性组合代替. 通过角度 $\theta$ 的选择可使前两列中的任何一个指定元素变为零.

例如,为了使元素 $A_{21}$ 为零,我们需要解方程 $ca_{21}+sa_{22}=0$. 此方程的解是 $c=-a_{22}/(a_{22}^2+a_{21}^2)^{1/2}$ 和 $s=a_{21}/(a_{22}^2+a_{21}^2)^{1/2}$. 因为 $c=\cos(\theta)$,$s=\sin(\theta)$,故要求 $c^2+s^2=1$,而上面给出的 $c$ 和 $s$ 的值满足这个要求.

目标

用 Givens 旋转将一个 $3\times3$ 矩阵 A 进行 RQ 分解.

算法

(1)乘 $Q_x$ 使 $A_{32}$ 为零.

(2)乘 $Q_y$ 使 $A_{31}$ 为零. 这一乘法不改变 A 的第二列,因此 $A_{32}$ 保持为零.

(3)乘 $Q_z$ 使 $A_{21}$ 为零. 这一乘法使 A 的前两列由它们的线性组合代替. 因此 $A_{31}$ 和 $A_{32}$ 保持为零.

算法 A3.1　$3\times3$ 矩阵的 RQ 分解.

RQ 算法的策略是通过乘 Givens 旋转每次使矩阵的下半部分的一个元素变为 0. 考虑把 $3\times3$ 矩阵 A 分解成 $A=RQ$,其中 R 是上三角阵而 Q 是一个旋转阵. 这可以分三步进行. 每一步包括右边乘以一个 Givens 旋转使矩阵 A 的一个选定的元素变为零. 乘法次序的选择必须使已经变为零的元素不受干扰. RQ 分解的一种实现在算法 A3.1 中给出.

可以选择其它的 Givens 旋转序列并给出同样的结果. 作为运算的结果我们发现 $AQ_xQ_yQ_z=R$,其中 R 是上三角矩阵. 因此,$A=RQ_z^TQ_y^TQ_x^T$,进而有 $A=RQ$,其中 $Q=Q_z^TQ_y^TQ_x^T$ 是一个旋转. 另外,与三个 Givens 旋转相关联的角度 $\theta_x$,$\theta_y$ 和 $\theta_z$ 提供一种由三个欧拉角描述的旋转参数化,它们也被称为滚动角、旋转角和逸出角.

由分解算法的描述,我们应该清楚 QR,QL 和 LQ 矩阵分解是如何相似以及应该如何进行. 不仅如此,这个算法也能轻易地推广到更高的维数.

### A3.1.2　Householder 矩阵和 QR 分解

对于更大维数的矩阵来说,用 Householder 矩阵进行 QR 分解较为有效. 对称矩阵

$$H_v=I-2\mathbf{v}\mathbf{v}^T/\mathbf{v}^T\mathbf{v} \tag{A3.2}$$

满足 $H_v^TH_v=I$,所以 $H_v$ 是正交矩阵.

令 $\mathbf{e}_1$ 为矢量$(1,0,\cdots,0)^T$,$\mathbf{x}$ 为任何矢量. 令 $\mathbf{v}=\mathbf{x}\pm\|\mathbf{x}\|\mathbf{e}_1$. 不难验证 $H_v\mathbf{x}=\mp\|\mathbf{x}\|\mathbf{e}_1$;因此 $H_v$ 是一个正交阵并把矢量 $\mathbf{x}$ 变换成 $\mathbf{e}_1$ 的一个倍数. 几何上说,$H_v$ 是关于由 $\mathbf{v}$ 定义的一个"平面"的反射,并且 $\mathbf{v}=\mathbf{x}\pm\|\mathbf{x}\|\mathbf{e}_1$ 是把矢量 $\mathbf{x}$ 和 $\|\mathbf{x}\|\mathbf{e}_1$ 角平分的一个适当的矢量. 为稳定起见,$\mathbf{v}$ 中定义的符号多义性应该通过令

$$\mathbf{v}=\mathbf{x}+\mathrm{sign}(x_1)\|\mathbf{x}\|\mathbf{e}_1 \tag{A3.3}$$

给予解决. 如果 A 是一个矩阵,$\mathbf{x}$ 是 A 的第一列并由(A3.3)定义 $\mathbf{v}$,那么作乘积 $H_vA$ 将使矩阵 A 的第一列对角元以下的元素变为 0,并代之以$(\|\mathbf{x}\|,0,0,\cdots,0)^T$. 我们继续用 Householder 矩阵左乘使 A 的对角线以下的部分变为 0. 按此方法进行,我们最终得到 $QA=R$;其中 Q 是正交矩阵的积而 R 是一个上三角矩阵. 因此,我们有 $A=Q^TR$. 这正是矩阵 A 的 QR 分解.

采用 Householder 矩阵作乘法时,若用一般的矩阵乘法程序则会产生错误并浪费计算机时间. 我们应该利用 Householder 矩阵的特殊形式来加速计算. 事实上,不必显式地形成 Householder 矩阵. 关于这方面的细节以及 Householder 矩阵和 QR 分解的其它性质,请读者参考[Golub-89].

**注意** 在 QR 和 RQ 分解中，R 是指上三角矩阵而 Q 是指正交矩阵. 在本书的其它地方所用的记号中，R 通常是指旋转(因而正交)矩阵.

## A3.2 对称和反对称矩阵

对称和反对称矩阵在本书中扮演重要角色. 矩阵 A 称为**对称的**，如果 $A^T = A$；称为**反对称的**，如果 $A^T = -A$. 这类矩阵的特征值分解概括于下面的结论中.

**结论 A3.1 特征值分解**

**(1)如果 A 是一个实对称矩阵，那么 A 可以分解为 $A = UDU^T$，其中 U 是正交矩阵，D 是实对角矩阵. 因此一个实对称矩阵有实特征值，其特征矢量两两正交.**

**(2)如果 S 是实的反对称矩阵，那么 $S = UBU^T$，其中 B 为形如 $\mathrm{diag}(a_1 Z, a_2 Z, \cdots, a_m Z, 0, \cdots, 0)^T$ 的分块对角矩阵，$Z = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. S 的特征矢量都是纯虚数并且奇数阶的反对称矩阵必是奇异的.**

该结论的一种证明在[Golub-89]中给出.

**Jacobi 方法** 一般来说，求任意矩阵的特征值是一个困难的数值问题. 但对于实对称矩阵，存在一个非常稳定的方法：Jacobi 方法. 这个方法的实现在[Press-88]中给出.

**叉乘**

我们对 3×3 的反对称矩阵有特殊的兴趣. 对于 3 维矢量 $\mathbf{a} = (a_1, a_2, a_3)^T$，我们定义一个对应的反对称矩阵如下：

$$[\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \tag{A3.4}$$

注意任何 3×3 的反对称矩阵都可以用一个适当的矢量 **a** 写为 $[\mathbf{a}]_\times$ 的形式. 矩阵 $[\mathbf{a}]_\times$ 是奇异的，并以 **a** 为其零矢量(右或左). 因此一个 3×3 的反对称矩阵由它的零矢量确定到只差纯量因子的程度.

两个 3 维矢量的叉乘(或矢量积) $\mathbf{a} \times \mathbf{b}$(有时记为 $\mathbf{a} \wedge \mathbf{b}$)是矢量 $(a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)^T$. 叉乘与反对称矩阵的关系是

$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_\times \mathbf{b} = (\mathbf{a}^T [\mathbf{b}]_\times)^T. \tag{A3.5}$$

**余因子和伴随矩阵** 令 M 是一个方阵. 用 $M^*$ 表示 M 的余因子矩阵. 即矩阵 $M^*$ 的 $(ij)$ 元素等于 $(-1)^{i+j} \det \hat{M}_{ij}$，其中 $\hat{M}_{ij}$ 是把 M 划去第 $i$ 行和第 $j$ 列所得到的矩阵. 余因子矩阵 $M^*$ 的转置矩阵称为 M 的**伴随矩阵**，记为 adj(M).

如所周知，若 M 可逆，则

$$M^* = \det(M) M^{-T}, \tag{A3.6}$$

其中 $M^{-T}$ 是 M 的转置的逆. 该公式对非可逆矩阵不成立，但 $\mathrm{adj}(M)M = M\mathrm{adj}(M) = \det(M)I$ 总是有效.

余因子矩阵与矩阵相对于叉乘的分配律有关.

**引理 A3.2　如果 M 是任意 3×3 矩阵(可逆或不可逆),而 x 和 y 是任意列矢量,那么**

$$(\mathbf{Mx})\times(\mathbf{My})=\mathbf{M}^*(\mathbf{x}\times\mathbf{y}). \tag{A3.7}$$

删去非本质的矢量 y,这个方程可以写为$[\mathbf{Mx}]_\times \mathbf{M}=\mathbf{M}^*[\mathbf{x}]_\times$. 现在,令 $\mathbf{t}=\mathbf{Mx}$ 并假定 M 可逆,我们得到计算反对称矩阵$[\mathbf{t}]_\times$与任何非奇异矩阵 M 乘积的公式. 我们可以将(A3.7)写成

**结论 A3.3　对任何矢量 t 和非奇异矩阵 M,我们有**

$$[\mathbf{t}]_\times \mathbf{M}=\mathbf{M}^{-\mathsf{T}}[\mathbf{M}^{-1}\mathbf{t}]_\times$$

注意(见结论 8.9($p$171))$[\mathbf{t}]_\times \mathbf{M}$ 是两个摄像机矩阵 $\mathbf{P}=[\mathbf{I}|\mathbf{0}]$和 $\mathbf{P}'=[\mathbf{M}|\mathbf{t}]$的基本矩阵. 结论 A3.3 的公式被用来推导基本矩阵的另一种形式(8.2－$p$163).

3×3 的反对称矩阵的一个奇妙的性质是在相差一个尺度因子的情况下有$[\mathbf{a}]_\times=[\mathbf{a}]_\times[\mathbf{a}]_\times[\mathbf{a}]_\times$(包含尺度则$[\mathbf{a}]_\times^3=-\|\mathbf{a}\|^2[\mathbf{a}]_\times$),这是容易验证的,因为右边的矩阵显然是反对称的且它的零空间生成元是 a. 由此可直接得到下面的结论:

**结论 A3.4　如果 $\mathbf{F}=[\mathbf{e}']_\times \mathbf{M}$ 是一个基本矩阵(3×3 奇异矩阵),那么$[\mathbf{e}']_\times[\mathbf{e}']_\times \mathbf{F}=\mathbf{F}$(相差一个尺度因子). 因此,F 可分解为 $\mathbf{F}=[\mathbf{e}']_\times \mathbf{M}$,其中 $\mathbf{M}=[\mathbf{e}']_\times \mathbf{F}$.**

### A3.2.1　正定对称矩阵

具有正的实特征值的特殊实对称矩阵称为**正定**对称矩阵. 我们列出实正定对称矩阵的一些重要性质如下.

**结论 A3.5　正定实对称矩阵**

**(1)一个对称矩阵 A 是正定的充要条件是对任何非零矢量 x 有 $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}>0$.**

**(2)一个正定对称矩阵 A 可以唯一地分解为 $\mathbf{A}=\mathbf{K}\mathbf{K}^\mathsf{T}$,其中 K 是对角元素全为正的上三角实矩阵.**

**证明**　本结论的第一部分几乎直接由分解式 $\mathbf{A}=\mathbf{U}\mathbf{D}\mathbf{U}^\mathsf{T}$ 推出. 对第二部分而言,因为 A 是对称正定的,它可以写成 $\mathbf{A}=\mathbf{U}\mathbf{D}\mathbf{U}^\mathsf{T}$,其中 D 是正的实对角阵而 U 是正交阵. 我们可以取 D 的平方根,并记为 $\mathbf{D}=\mathbf{E}\mathbf{E}^\mathsf{T}$,其中 E 是对角矩阵. 于是 $\mathbf{A}=\mathbf{V}\mathbf{V}^\mathsf{T}$,其中 $\mathbf{V}=\mathbf{U}\mathbf{E}$. 矩阵 V 不是上三角阵. 但我们可应用 RQ 分解(节 A3.1.1)将它写为 $\mathbf{V}=\mathbf{K}\mathbf{Q}$,其中 K 是上三角阵而 Q 是正交阵. 则 $\mathbf{A}=\mathbf{V}\mathbf{V}^\mathsf{T}=\mathbf{K}\mathbf{Q}\mathbf{Q}^\mathsf{T}\mathbf{K}^\mathsf{T}=\mathbf{K}\mathbf{K}^\mathsf{T}$. 这是 A 的 Cholesky 分解. 用对角元素为±1 的一个对角阵乘 K,我们可以保证的 K 对角元素都是正的. 这个做法不会改变 $\mathbf{K}\mathbf{K}^\mathsf{T}$ 的积.

现在我们来证明分解的唯一性. 具体地说,如果 $\mathbf{K}_1$ 和 $\mathbf{K}_2$ 是两个满足 $\mathbf{K}_1\mathbf{K}_1^\mathsf{T}=\mathbf{K}_2\mathbf{K}_2^\mathsf{T}$ 的上三角矩阵,那么 $\mathbf{K}_2^{-1}\mathbf{K}_1=\mathbf{K}_2^\mathsf{T}\mathbf{K}_1^{-\mathsf{T}}$. 因为该方程的左边是上三角矩阵而右边是下三角矩阵,所以它们必须同时是对角阵. 因此 $\mathbf{D}=\mathbf{K}_2^{-1}\mathbf{K}_1=\mathbf{K}_2^\mathsf{T}\mathbf{K}_1^{-\mathsf{T}}$. 但是 $\mathbf{K}_2^{-1}\mathbf{K}_1$ 是 $\mathbf{K}_2^\mathsf{T}\mathbf{K}_1^{-\mathsf{T}}$ 转置的逆,所以 D 等于它本身转置的逆,从而是对角元素为±1 的一个对角阵. 如果 $\mathbf{K}_1$ 和 $\mathbf{K}_2$ 都具有正的对角元素,那么 $\mathbf{D}=\mathbf{I}$ 和 $\mathbf{K}_1=\mathbf{K}_2$. □

以上证明给出了计算 Cholesky 分解一种构造性的方法. 但是,还有更简单和更有效的直接计算的方法,参看[Press－88].

## A3.3 奇异值分解

奇异值分解(SVD)是最有用的矩阵分解方法中的一种,特别是对数值计算而言. 给定一个方阵 A,SVD 把 A 分解为 $A=UDV^T$,其中 U 和 V 是正交矩阵而 D 是一个元素为非负的对角矩阵. 注意在此分解中习惯用 $V^T$ 代替 V. 该分解可以使 D 的对角元素以降序排列并且我们假定总是这样做的. 因此,诸如"V 的对应于最小特征值的列"这样的绕口的话就代之以"V 的最后一列".

非方阵 A 的 SVD 也存在. 最有趣的情形是 A 的行多于列. 具体地说,令 A 是 $m\times n$ 矩阵且 $m\geqslant n$. 在这种情形,A 可以分解为 $A=UDV^T$,其中 U 是具有正交列的 $m\times n$ 矩阵,D 是 $n\times n$ 对角矩阵而 V 是一个 $n\times n$ 正交矩阵. U 具有正交列意味着 $U^TU=I_{n\times n}$. 不仅如此,U 还具有保范性质,即(可实际验证)对任何矢量 $\mathbf{x}$ 都有 $\|U\mathbf{x}\|=\|\mathbf{x}\|$. 另一方面,$UU^T$ 一般不是单位矩阵除非 $m=n$.

我们同样可以定义列多于行的矩阵的奇异值分解,但一般地我们对它不感兴趣. 在本书,我们偶而也要求 $m<n$ 的矩阵 A 的奇异值分解,我们在实际做法上先用全是零元素的行来把 A 扩展成方阵,再求所得矩阵的 SVD. 我们都将这样做且不作特别说明.

通常 SVD 的实现,如[Press-88]中的那样,都假定 $m\geqslant n$. 因为在这种情形,矩阵 U 具有与输入矩阵 A 一样的维数 $m\times n$,矩阵 A 可以被输出矩阵 U 覆盖.

**SVD 的实现** 本书不给出奇异值分解算法的介绍或它的存在性的证明. 关于算法如何运行的介绍,读者可以参考[Golub-89]. SVD 的一个具体实现在[Press-88]中给出. 但是在"Numerical Recipes in C"第一版给出的 SVD 实现的算法有时会给出不正确的结果. "Numerical Recipes in C"第二版[Press-88]给出的算法纠正了先前版本的错误.

**奇异值和特征值** 在 SVD 中的矩阵 D 的对角元素是非负的. 这些元素称为矩阵 A 的奇异值. 它们与特征值不是同一件事情. 为了解 A 的奇异值与特征值的联系,我们由 $A=UDV^T$ 得到 $A^TA=VDU^TUDV^T=VD^2V^T$. 因为 V 是正交的,$V^T=V^{-1}$,因而 $A^TA=VD^2V^{-1}$. 这是特征值的定义方程,它表明 $D^2$ 的元素是 $A^TA$ 的特征值而 V 的列是 $A^TA$ 的特征矢量. 简单地说,A 的奇异值是 $A^TA$ 的特征值的平方根.

注意 $A^TA$ 是对称半正定的(见上面的节 A3.2.1),所以它的特征值是实非负的. 因此,奇异值是实非负的.

**SVD 的计算复杂度**

SVD 的计算复杂度取决于需要返回信息的多少. 例如在后面要考虑的算法 A3.5 中,要求的解是 SVD 中矩阵 V 的最后一列. 矩阵 U 没有用到,因而不需要计算 U. 另一方面,节 A3.3.1 的算法 A3.2 需要计算整个 SVD. 对于行数过多于列数的方程组,计算矩阵 U 所需的额外开销是不可小视的.

对计算一个 $m\times n$ 矩阵的 SVD 大约所需的浮点运算(flops)数在[Golub-89]中给出. 计算矩阵 U,V 和 D 总共需要 $4m^2n+8mn^2+9n^3$ flops. 但是,如果只计算矩阵 V 和 D,那么仅需要 $4mn^2+8n^3$ flops. 这是一个重要的区别,因为后一个表达式不包含 $m^2$ 的项. 具体地说,计算 U 所需的运算数随行数 $m$ 的平方而变化. 但计算 V 和 D 的复杂度都是 $m$ 的

线性函数. 因此,当行数过多于列数时,(如果计算时间是关注的问题)除非有必要,否则应避免计算矩阵 U. 为了说明这一点,我们来考虑第 6 章介绍的计算摄像机投影矩阵的 DLT 算法. 在此算法中由 $n$ 个 3D 到 2D 的点匹配来计算一个 $3\times4$ 的摄像机矩阵. 该算法涉及到用算法 A3.5 求解方程组 $\mathbf{Ap}=\mathbf{0}$,其中 A 是 $2n\times12$ 的矩阵. 解矢量 $\mathbf{p}$ 是 A 的 SVD 分解 $A=UDV^T$ 中的矩阵 V 的最后一列. 因而可不需要矩阵 U. 表 A3.1 给出对 6(最小配置数),100,1000 组点对应进行 SVD 所需的 flops 总数.

| 点数 | 方程数<br>每点 | 方程数<br>($m$) | 未知数<br>($n$) | 运算数<br>不计算 U | 运算数<br>计算 U |
|---|---|---|---|---|---|
| 6 | 2 | 12 | 12 | 20,736 | 36,288 |
| 100 | 2 | 200 | 12 | 129,024 | 2,165,952 |
| 1000 | 2 | 2000 | 12 | 1,165,824 | 194,319,552 |

表 A3.1 计算 SVD 所需的 flops 数的比较,矩阵的大小是 $m\times n$,其中 $m$ 取不同值而 $n=12$. 注意不计算 U 时,计算复杂度与方程数成次线性增加. 另一方面,计算 U 的额外计算负担非常大,尤其对方程数目大的情形.

### A3.3.1 线性方程组的解

奇异值分解最经常的应用是求解超定方程组. 我们考虑几个适宜于采用 SVD 求解的问题. 考虑形如 $\mathbf{Ax}=\mathbf{b}$ 的方程组. 令 A 为 $m\times n$ 矩阵. 这里有三种可能性:

(1)如果 $m<n$,那么未知数大于方程数. 在这种情形,解不唯一而且有一个解矢量空间.

(2) 如果 $m=n$,那么只要 A 可逆便有唯一解.

(3)如果 $m>n$,那么方程数大于未知数. 方程组一般没有解,除非 $\mathbf{b}$ 属于由 A 的列所生成的子空间.

**最小二乘解:满秩情形** 我们考虑 $m\geqslant n$ 并且 A 的秩为 $n$ 的情形. 如果解不存在,那么对许多情形,找一个最接近于所给方程组 $\mathbf{Ax}=\mathbf{b}$ 的一个解矢量 $\mathbf{x}$ 仍然是有意义的. 换句话说,我们寻求一个向量 $\mathbf{x}$ 使 $\|\mathbf{Ax}-\mathbf{b}\|$ 最小,其中 $\|\cdot\|$ 表示矢量范数. 这样的 $\mathbf{x}$ 称为该超定方程组的一个**最小二乘解**. 用 SVD 能方便地求最小二乘解,其方法如下所述.

我们寻求使 $\|\mathbf{Ax}-\mathbf{b}\|=\|UDV^T\mathbf{x}-\mathbf{b}\|$ 取最小值的矢量 $\mathbf{x}$. 由于正交变换的保范性质,我们要最小化的量是 $\|UDV^T\mathbf{x}-\mathbf{b}\|=\|DV^T\mathbf{x}-U^Tb\|$. 记 $\mathbf{y}=V^T\mathbf{x}$ 和 $\mathbf{b}'=U^T\mathbf{b}$,问题变成最小化 $\|D\mathbf{y}-\mathbf{b}'\|$,其中 D 为 $m\times n$ 矩阵并且对角线以外的元素为零. 该方程组的形式是

$$\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \\ \hline & & 0 & \end{bmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}=\begin{pmatrix} b_1' \\ b_2' \\ \vdots \\ b_n' \\ \hline b_{n+1} \\ \vdots \\ b_m \end{pmatrix}.$$

显然,离 $\mathbf{b}'$ 最近的 $D\mathbf{y}$ 是矢量 $(b_1',b_2',\cdots,b_n',0,\cdots,0)^T$,并通过令 $y_i=b_i'/d_i\ (i=1,\cdots,n)$ 而得到. 注意假定 A 的秩为 $n$ 保证了 $d_i\neq0$. 最后由 $\mathbf{x}=V\mathbf{y}$ 求出 $\mathbf{x}$. 整个算法是

目标

求 $m\times n$ 方程组 $A\mathbf{x}=\mathbf{b}$ 的最小二乘解，其中 $m>n$，并且 $\mathrm{rank}A=n$.

算法

(1)求 SVD: $A=UDV^T$.

(2)令 $\mathbf{b}'=U^T\mathbf{b}$.

(3)求由 $y_i=b_i'/d_i$ 定义的 $\mathbf{y}$，其中 $d_i$ 是 D 的第 $i$ 个对角元素.

(4)所求解为 $\mathbf{x}=V\mathbf{y}$.

算法 A3.2　满秩的超定线性方程组的线性最小二乘解.

**降秩方程组**　有时，我们需要求解矩阵不是列满秩的方程组. 令 $r=\mathrm{rank}A<n$，其中 $n$ 是 A 的列数. 由于噪声的影响，矩阵 A 的秩可能实际上大于 $r$，但是，出于被考虑的特定问题理论分析的要求，我们希望强制执行秩为 $r$ 的约束. 在这种情形，该方程组的解将是 $(n-r)$ 个参数的簇，其中 $r=\mathrm{rank}A<n$. 这一簇解也适宜用 SVD 来求，具体如下：

目标

求方程组 $A\mathbf{x}=\mathbf{b}$ 的通解，其中 A 是 $m\times n$ 矩阵并且 A 的秩 $r<n$.

算法

(1) 求 SVD: $A=UDV^T$，其中 D 的对角元素 $d_i$ 按递降顺序排列.

(2) 令 $\mathbf{b}'=U^T\mathbf{b}$.

(3) 求 $\mathbf{y}$，其定义是：当 $i=1,\cdots r$ 时，$y_i=b_i'/d_i$，否则 $y_i=0$.

(4) 有最小范数 $\|\mathbf{x}\|$ 的解 $\mathbf{x}$ 是 $V\mathbf{y}$.

(5) 通解是 $\mathbf{x}=V\mathbf{y}+\lambda_{r+1}\mathbf{v}_{r+1}+\cdots+\lambda_n\mathbf{v}_n$，其中 $\mathbf{v}_{r+1}\cdots,\mathbf{v}_n$ 是 V 的最后 $n-r$ 列.

算法 A3.3　降秩方程组的通解.

这个算法给出降秩方程组的最小二乘解的 $(n-r)$ 参数簇（由不定值 $\lambda_i$ 参数化）. 该算法的正确性的证明类似于满秩方程组最小二乘解的算法 A3.2.

**秩未知的方程组**　在本书中遇到的大多数线性方程组在求解之前已从理论上知道其矩阵的秩. 如果不知道方程组的秩，那么我们必须猜它的秩. 在这种情形，适宜的做法是把与最大的奇异值相比相对很小的奇异值设为零. 因此如果 $d_i/d_0<\delta$（其中 $\delta$ 是相当于机器精度[①]数量级的一个小常数），那么设 $y_i=0$. 如前，最小二乘解由 $\mathbf{x}=V\mathbf{y}$ 给出.

## A3.4　伪逆

给定对角方阵 D，我们定义它的**伪逆** $D^+$ 是对角矩阵满足

$$D_{ii}^+=\begin{cases}0 & \text{如果 } D_{ii}=0\\ D_{ii}^{-1} & \text{其他}\end{cases}.$$

现在考虑 $m\times n$ 矩阵 A，其中 $m\geqslant n$. 令 A 的 SVD 为 $A=UDV^T$. 我们定义 A 的**伪逆**为矩阵

① 机器精度是使 $1.0+\varepsilon=1.0$ 的最大的浮点数值 $\varepsilon$.

$$A^+ = VD^+U^T. \tag{A3.8}$$

通过简单的验证可知在算法 A3.2 或 A3.3 中的矢量 $\mathbf{y}$ 不过就是 $D^+\mathbf{b}'$，其中 $\mathbf{b}' = U^T\mathbf{b}$.

**结论 A3.6　秩为 $n$ 的 $m \times n$ 方程组 $A\mathbf{x} = \mathbf{b}$ 的最小二乘解由 $\mathbf{x} = A^+\mathbf{b}$ 给出. 在降秩方程组情形，$\mathbf{x} = A^+\mathbf{b}$ 是使 $\|\mathbf{x}\|$ 最小化的解.**

如在讨论 SVD 时所提到的：如果 A 的行比列少，那么通过添加元素全为零的行扩充 A 为方阵之后就可以采用此结论.

**对称矩阵**　我们可以用如下方法推广对称矩阵的伪逆. 这种推广已在第 4 章节 4.2.3 ($p$93)讨论奇异协方差矩阵时被用到. 如果 A 是非可逆对称矩阵而 $X^TAX$ 是可逆的，那么我们记 $A^{+X} \xlongequal{\text{def}} X(X^TAX)^{-1}X^T$. 我们可以看到它仅依赖于由 X 的列所生成的子空间. 换句话说，如果 X 由 XB 代替(其中 B 为任何可逆矩阵)，那么 $A^{+X} = A^{+XB}$. 也就是说，$A^{+X}$ 仅依赖于 X 的(左)零空间，即垂直于 X 所有列的矢量空间. 定义零空间 $N_L X = \{\mathbf{x}^T \mid \mathbf{x}^T X = \mathbf{0}\}$. 我们发现在一个简单的条件下，$A^{+X}$ 是 A 的伪逆.

**结论 A3.7　令 A 为对称矩阵，那么 $A^{+X} \xlongequal{\text{def}} X(X^TAX)^{-1}X^T = A^+$ 的充要条件是 $N_L(X) = N_L(A)$.**

我们仅给出证明的大意. 必要性是明显的，因为 $N_L(X)$ 和 $N_L(A)$ 分别是方程的左和右边的零空间. 为证明充分性，我们可以假定 X 的列是正交的，因为上面已经证明，仅有 X 的零空间是重要的. 于是，X 可以通过增加列 X′ 扩充为正交矩阵 $U = [X \mid X']$. 由此可见，$X'^T$ 的行矢量生成 X 的零空间(根据假设也是 A 的零空间). 现在，通过比较定义 $A^{+X} \xlongequal{\text{def}} X(X^TAX)^{-1}X^T$ 和伪逆的定义(A3.8)可以在几行内完成所需的证明.

### A3.4.1　用正规方程解线性最小二乘问题

线性最小二乘问题也可以用一个涉及所谓**正规方程**的方法来解. 我们再来考虑线性方程组 $A\mathbf{x} = \mathbf{b}$，其中 A 为 $m \times n$ 矩阵并且 $m > n$. 这个方程组一般不存在解 $\mathbf{x}$. 因此，我们的任务是求最小化范数 $\|A\mathbf{x} - \mathbf{b}\|$ 的矢量 $\mathbf{x}$. 当 $\mathbf{x}$ 取遍所有的值时，积 $A\mathbf{x}$ 将遍历 A 的整个列空间，即由 A 的列生成的一个 $\mathbb{R}^m$ 的子空间. 因此我们的任务是在 A 的列空间中寻求最接近 $\mathbf{b}$ 的那个矢量，接近程度由矢量的范数定义. 令 $\mathbf{x}$ 是这个问题的解，因此 $A\mathbf{x}$ 是最接近 $\mathbf{b}$ 的点. 在这种情形，差 $A\mathbf{x} - \mathbf{b}$ 必然是与 A 的列空间垂直的矢量. 更明确地说，$A\mathbf{x} - \mathbf{b}$ 垂直于 A 的每一列，因此 $A^T(A\mathbf{x} - \mathbf{b}) = \mathbf{0}$. 把它乘出来并分项得到方程

$$(A^TA)\mathbf{x} = A^T\mathbf{b}. \tag{A3.9}$$

这是一个 $n \times n$ 的线性方程组，称为**正规方程**. 该方程组有解并可用来求问题 $A\mathbf{x} = \mathbf{b}$ 的最小二乘解. 即使 A 不满秩(秩为 $n$)，该方程组也应该有解，因为 $A^T\mathbf{b}$ 在 $A^TA$ 的列空间中. 当 A 的秩为 $n$ 时，矩阵 $A^TA$ 可逆，从而 $\mathbf{x}$ 可以通过 $\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b}$ 求得. 因为 $\mathbf{x} = A^+\mathbf{b}$，故由上式不难得到下面的结论：

**结论 A3.8　如果 $m \times n$ 矩阵 A 的秩为 $n$，那么 $A^+ = (A^TA)^{-1}A^T$.**

目标

求使$\|A\mathbf{x}-\mathbf{b}\|$最小化的$\mathbf{x}$.

算法

(1)解正规方程$A^{\mathsf{T}}A\mathbf{x}=A^{\mathsf{T}}\mathbf{b}$.

(2)如果$A^{\mathsf{T}}A$可逆,那么解是$\mathbf{x}=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}\mathbf{b}$.

算法 A3.4　利用正规方程求线性最小二乘解的算法.

这个结论不仅在理论分析上有用,而且当$n$相对于$m$很小时,也给出比用SVD计算伪逆方法更为简单的方法(因此计算$(A^{\mathsf{T}}A)$的逆比计算A的SVD要节省).

**矢量空间的范数**　有时我们希望相对于矢量空间$\mathbb{R}^n$的不同的范数来最小化$A\mathbf{x}-\mathbf{b}$. 矢量空间$\mathbb{R}^n$中常用范数由常用内积给出. 对$\mathbb{R}^n$中的两个矢量$\mathbf{a}$和$\mathbf{b}$,我们可以定义常用内积为$\mathbf{a}\cdot\mathbf{b}=\mathbf{a}^{\mathsf{T}}\mathbf{b}$. 于是矢量$\mathbf{a}$的常用范数定义为$\|\mathbf{a}\|=(\mathbf{a}\cdot\mathbf{a})^{1/2}=(\mathbf{a}^{\mathsf{T}}\mathbf{a})^{1/2}$. 注意下列性质:

(1) 内积是$\mathbb{R}^n$上的一个对称双线性形式.

(2) 对所有的非零矢量$\mathbf{a}\in\mathbb{R}^n$有$\|\mathbf{a}\|>0$.

我们说内积是正定对称双线性形式. 可以在矢量空间$\mathbb{R}^n$上定义其它的内积. 令C是一个实对称正定矩阵,并定义一个新的内积为$C(\mathbf{a},\mathbf{b})=\mathbf{a}^{\mathsf{T}}C\mathbf{b}$. 内积的对称性由C的对称性推出. 用$\|\mathbf{a}\|_C=(\mathbf{a}^{\mathsf{T}}C\mathbf{a})^{1/2}$可以定义一个范数. 它是正定的,因为假定C是正定矩阵.

**加权线性最小二乘问题**　有时我们希望解形如$A\mathbf{x}-\mathbf{b}=\mathbf{0}$的加权线性最小二乘问题,即最小化误差的C-范数$\|A\mathbf{x}-\mathbf{b}\|_C$. 这里C是在$\mathbb{R}^n$上用以定义内积和范数$\|\cdot\|_C$的一个对称正定矩阵. 如前,我们可以说最小化误差矢量$A\mathbf{x}-\mathbf{b}$必须(在C定义的内积意义下)垂直于A的列空间. 由此推出$A^{\mathsf{T}}C(A\mathbf{x}-\mathbf{b})=\mathbf{0}$. 将它重新排列便得到加权正规方程

$$(A^{\mathsf{T}}CA)\mathbf{x}=A^{\mathsf{T}}C\mathbf{b}. \tag{A3.10}$$

最常用的加权是其中的C为对角矩阵,对应于$\mathbb{R}^n$的每一坐标轴方向独立加权. 但是也可以采用一般的加权矩阵C.

### A3.4.2　齐次方程组的最小二乘解

与前一个问题类似的问题是求形如$A\mathbf{x}=\mathbf{0}$的方程组的解. 这个问题常出自于一些重构问题. 我们考虑方程数多于未知元素的情形——超定方程组. 因平凡解$\mathbf{x}=\mathbf{0}$是我们不感兴趣的,故我们寻求该方程组的非零解. 注意如果$\mathbf{x}$是这方程组的一个解,那么对任何标量$k$,$k\mathbf{x}$也是解. 一个合理的约束是只求$\|\mathbf{x}\|=1$的解.

这样的方程组一般不存在精确解. 假定A的维数是$m\times n$,那么存在精确解的充要条件是$\mathrm{rank}(A)<n$——矩阵A不是满列秩. 当没有精确解时,我们通常求它的一个最小二乘解. 问题可以叙述为

- **求使$\|A\mathbf{x}\|$最小化并满足$\|\mathbf{x}\|=1$的$\mathbf{x}$.**

这个问题可按如下方法来解. 令$A=UDV^{\mathsf{T}}$. 那么问题变成最小化$\|UDV^{\mathsf{T}}\mathbf{x}\|$. 但是$\|UDV^{\mathsf{T}}\mathbf{x}\|=\|DV^{\mathsf{T}}\mathbf{x}\|$和$\|\mathbf{x}\|=\|V^{\mathsf{T}}\mathbf{x}\|$. 因此,我们需要在条件$\|V^{\mathsf{T}}\mathbf{x}\|=1$下最小化$\|DV^{\mathsf{T}}\mathbf{x}\|$. 令

$\mathbf{y}=V^T\mathbf{x}$,则问题变为在条件$\|\mathbf{y}\|=1$下最小化$\|D\mathbf{y}\|$. 现在,D是对角元素按降序排列的一个对角矩阵. 由此推出该问题的解是$\mathbf{y}=(0,0,\cdots,0,1)^T$,它具有一个非零元素1并在最后的位置上. 最后注意$\mathbf{x}=V\mathbf{y}$就是V的最后一列. 这方法概括在算法A3.5中.

如节A3.3所述,V的最后一列也可以描述为对应于$A^TA$最小特征值的特征矢量.

目标

给定一个行数不少于列数的矩阵A,求最小化$\|A\mathbf{x}\|$并满足$\|\mathbf{x}\|=1$的$\mathbf{x}$.

解

$\mathbf{x}$是V的最后一列,其中$A=UDV^T$是A的SVD.

算法A3.5 齐次线性方程组的最小二乘解.

### A3.4.3 带约束方程组的最小二乘解

在前一节中,我们考虑了形如$A\mathbf{x}=\mathbf{0}$的方程组的最小二乘解的方法. 这样的问题可能出自对图像特征集合进行测量的情形. 对准确的测量和准确的影像模型,该数学模型给出了该方程组的一个准确的解. 在非准确的图像测量或有噪声存在时,精确解将不存在. 此时,求最小二乘解是有意义的.

但是在其它一些场合,矩阵A的某些行所表示的方程是由数学约束精确地导出的,而且要求精确地满足. 这样的一组约束可以用矩阵方程$C\mathbf{x}=\mathbf{0}$来描述,并要求它被精确地满足. 其它的方程产生于图像测量并且受到噪声的干扰. 这导致下列的问题

- **求$\mathbf{x}$,它最小化$\|A\mathbf{x}\|$并满足$\|\mathbf{x}\|=1$和$C\mathbf{x}=\mathbf{0}$.**

这个问题可按如下方式来解. 要求$\mathbf{x}$满足条件$C\mathbf{x}=\mathbf{0}$意味着$\mathbf{x}$垂直于C的每一行. 所有这种$\mathbf{x}$的集合是一个矢量空间,称为C的行空间的正交补. 我们希望求这个正交补.

首先,如果C的行数少于列数,那么加若干零行将它扩展成方阵. 这样做不会对约束集合$C\mathbf{x}=\mathbf{0}$产生影响. 现在,作C的奇异值分解:$C=UDV^T$,其中D是有$r$个非零对角元素的对角矩阵. 在此情形C的秩为$r$并且C的行空间由$V^T$的前$r$行生成. C的行空间的正交补由$V^T$的余下的行生成. 以$C^\perp$记矩阵V消去前$r$列得到的矩阵,则$CC^\perp=\mathbf{0}$,因此,满足$C\mathbf{x}=\mathbf{0}$的矢量$\mathbf{x}$的集合由$C^\perp$的列生成,从而,可以把任何这样的$\mathbf{x}$记为$\mathbf{x}=C^\perp\mathbf{x}'$,其中$\mathbf{x}'$为适当的矢量. 因为$C^\perp$具有正交列,故有$\|\mathbf{x}\|=\|C^\perp\mathbf{x}'\|=\|\mathbf{x}'\|$. 这样一来,上述最小化问题化为

- **求$\mathbf{x}'$,它最小化$\|AC^\perp\mathbf{x}'\|$并满足$\|\mathbf{x}'\|=1$.**

这正是节A3.4.2讨论过可用算法A3.5求解的问题的一个例子. 求解约束最小化问题的完整算法在算法A3.6中给出.

目标

给定一个$m\times n$矩阵A,$m\geqslant n$,求矢量$\mathbf{x}$,它最小化$\|A\mathbf{x}\|$并且满足$\|\mathbf{x}\|=1$和$C\mathbf{x}=\mathbf{0}$.

算法

(1)如果C的行数少于列数,那么加若干零行将C扩展成方阵. 计算C的SVD:$C=UDV^T$,其中D的对角元非零元排在零元前面. 令$C^\perp$是矩阵V消去前$r$列得到的矩阵,其中$r$是D中非零元素的数目(C的秩).

(2)用算法A3.5求最小化问题$AC^\perp\mathbf{x}'=\mathbf{0}$的解. 此解由$\mathbf{x}=C^\perp\mathbf{x}'$给出.

算法A3.6 约束最小化算法.

### A3.4.4 进一步讨论约束最小化

约束最小化的进一步问题产生于本书中经常用到的代数估计方法——例如，计算基本矩阵(节 10.3($p$193))或三焦点张量(节 15.3($p$280))．

这个问题是

• **在条件 $\|\mathbf{x}\|=1$ 和 $\mathbf{x}=G\hat{\mathbf{x}}$ 下最小化 $\|A\mathbf{x}\|$，其中 G 是一个给定的矩阵而 $\hat{\mathbf{x}}$ 是某个未知的矢量．**

注意这个最小化问题与节 A3.4.3 的非常类似，因为如果现在的问题中的矩阵 G 有正交列那就简化成 A3.4.3 中问题的形式．对某个 $\hat{\mathbf{x}}$ 有 $\mathbf{x}=G\hat{\mathbf{x}}$ 成立的条件不过是要求 $\mathbf{x}$ 属于 G 的列所生成的空间．因此为了利用算法 A3.6 来解现在的问题，我们仅需要用与 G 具有同样列空间(即由这些列生成的空间)但其列为标准正交的矩阵替代 G．如果 $G=UDV^{\mathsf{T}}$，其中 D 有 $r$ 个非零元素(即 G 的秩为 $r$)，那么令 $U'$ 为 U 前 $r$ 列组成的矩阵．则 G 和 $U'$ 有同样的列空间．按节 A3.4.3 方法求使 $\|AU'\mathbf{x}'\|$ 最小化的单位矢量 $\mathbf{x}'$，然后令 $\mathbf{x}=U'\mathbf{x}'$，即可求解 $\mathbf{x}$．

如果还需要求 $\hat{\mathbf{x}}$，那么它可以通过解 $G\hat{\mathbf{x}}=\mathbf{x}=U'\mathbf{x}'$ 得到．所需要的解用伪逆表示(节 A3.4)为 $\hat{\mathbf{x}}=G^{+}\mathbf{x}=G^{+}U'\mathbf{x}'$；如果 G 不是列满秩，则解可能不是唯一的．因为 $G^{+}=VD^{+}U^{\mathsf{T}}$，我们有 $\hat{\mathbf{x}}=VD^{+}U^{\mathsf{T}}U'\mathbf{x}'$，它可简化为 $\hat{\mathbf{x}}=V'D'^{-1}\mathbf{x}'$，其中 $V'$ 由 V 的前 $r$ 列组成而 $D'$ 是 D 左上部的 $r\times r$ 子矩阵．

整个方法概括在算法 A3.7 中．

<u>目标</u>

求矢量 $\mathbf{x}$，它最小化 $\|A\mathbf{x}\|$ 并满足 $\|\mathbf{x}\|=1$ 和 $\mathbf{x}=G\hat{\mathbf{x}}$，其中 G 的秩为 $r$．

<u>算法</u>

(1)计算 G 的 SVD：$G=UDV^{\mathsf{T}}$，其中 D 的对角线的非零值出现在零值前面．

(2)令 $U'$ 是由 U 前 $r$ 列组成的矩阵．

(3)用算法 A3.5 求使 $\|AU'\mathbf{x}'\|$ 最小化的单位矢量 $\mathbf{x}'$．

(4)所要求的解是 $\mathbf{x}=U'\mathbf{x}'$．

(5)如果有要求，我们可以用 $V'D'^{-1}\mathbf{x}'$ 计算 $\hat{\mathbf{x}}$，其中 $V'$ 由 V 的前 $r$ 列组成而 $D'$ 是 D 左上部的 $r\times r$ 子矩阵．

算法 A3.7 受一个生成空间约束的最小化算法．

### A3.4.5 另一个最小化问题

一个非常类似的问题是

• **在条件 $\|C\mathbf{x}\|=1$ 下最小化 $\|A\mathbf{x}\|$．**

这个问题来自诸如解 DLT 摄像机标定问题(节 6.3(p126))．一般情况下 $\mathrm{rank}C<n$，其中 $n$ 是矢量 $\mathbf{x}$ 的维数．这个问题的几何解释是求(由 $\mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}A\mathbf{x}$ 确定)二次曲面的“最低”点，该点必须满足在(非齐次)“二次曲面” $\mathbf{x}^{\mathsf{T}}C^{\mathsf{T}}C\mathbf{x}=1$ 上的约束．

我们首先取矩阵 C 的 SVD，得到 $C=UDV^{\mathsf{T}}$．因条件 $\|UDV^{\mathsf{T}}\mathbf{x}\|=1$ 等价于 $\|DV^{\mathsf{T}}\mathbf{x}\|=1$，故不需要显式地计算 U．记 $\mathbf{x}'=V^{\mathsf{T}}\mathbf{x}$，则问题变成：在条件 $\|D\mathbf{x}'\|=1$ 下最小化 $\|AV\mathbf{x}'\|$．记

$A' = AV$,问题变成:最小化$\|A'\mathbf{x}'\|$并满足条件$\|D\mathbf{x}'\| = 1$. 因此我们已经简化到约束矩阵是对角矩阵 D 的情形.

我们假定 D 的对角元素有 $r$ 个为非零和 $s$ 个为零($r + s = n$),并且非零元素排在零元素之前. 那么 $\mathbf{x}'$的元素 $x_i'$在 $i > r$ 时不影响$\|D\mathbf{x}'\|$的值,因为 D 中对应的对角元素是零. 因此,对于 $x_i'(i = 1, \ldots, r)$的一个具体取值,应选取其它的 $x_i'(i = r + 1, \ldots, n)$使$\|A'\mathbf{x}'\|$值最小. 我们记 $A' = [A_1' | A_2']$,其中 $A_1'$由 $A'$的前 $r$ 列组成而 $A_2'$由余下的 $s$ 列组成. 类似地,令 $\mathbf{x}_1'$为由 $\mathbf{x}'$的前 $r$ 个元素组成的 $r$ 维矢量,而 $\mathbf{x}_2'$由 $\mathbf{x}'$余下的 $s$ 个元素组成. 此外,令 $D_1$ 为由 D 的前 $r$ 个对角元素组成的 $r \times r$ 对角矩阵. 于是,$A'\mathbf{x}' = A_1'\mathbf{x}_1' + A_2'\mathbf{x}_2'$,并且所考虑的最小化问题就是在条件$\|D_1\mathbf{x}_1'\| = 1$下最小化

$$\|A_1'\mathbf{x}_1' + A_2'\mathbf{x}_2'\|. \tag{A3.11}$$

现在,暂时固定 $\mathbf{x}_1'$,则(A3.11)取节 A3.3.1 讨论过的那种最小二乘最小化的类型的形式. 根据结论 A3.6,最小化(A3.11)的值 $\mathbf{x}_2'$是 $\mathbf{x}_2' = -A_2'^{+}A_1'\mathbf{x}_1'$. 把它代入(A3.11)得到$\|(A_2'A_2'^{+} - I)A_1'\mathbf{x}_1'\|$,它是我们要在条件$\|D_1\mathbf{x}_1'\| = 1$下最小化的方程. 最后记 $\mathbf{x}'' = D_1\mathbf{x}_1'$,本问题最终化为我们所熟悉的算法 A3.5 的最小化问题的形式:

- **在条件$\|\mathbf{x}''\| = 1$下最小化$\|(A_2'A_2'^{+} - I)A_1'D_1^{-1}\mathbf{x}_1''\|$.**

我们现在将算法概括如下

<u>目标</u>

在$\|C\mathbf{x}\| = 1$下最小化$\|A\mathbf{x}\|$.

<u>算法</u>

(1)计算 C 的 SVD:$C = UDV^{T}$,并记 $A' = AV$.

(2)假定 $\mathrm{rank}D = r$ 并令 $A' = [A_1' | A_2']$,其中 $A_1'$由 $A'$的前 $r$ 列组成而 $A_2'$由余下的 $s$ 列组成.

(3) 令 $D_1$ 为 D 左上部的 $r \times r$ 子阵.

(4) 计算 $A'' = (A_2'A_2'^{+} - I)A_1'D_1^{-1}$. 这是一个 $n \times r$ 的矩阵.

(5) 用算法 A3.5 最小化$\|A''\mathbf{x}'\|$并满足条件$\|\mathbf{x}''\| = 1$.

(6) 计算 $\mathbf{x}_1' = D_1^{-1}\mathbf{x}''$和 $\mathbf{x}_2' = -A_2'^{+}A_1'\mathbf{x}_1'$. 令 $\mathbf{x}' = \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \end{pmatrix}$.

(7) 解由 $\mathbf{x} = V\mathbf{x}'$ 给出.

算法 A3.8　在约束$\|C\mathbf{x}\| = 1$下求齐次线性方程组的最小二乘解.

# 附录 4　迭代估计方法

本附录将介绍两种最常用的参数最小化的迭代方法. 它们是 Newton 迭代和 Levenberg - Marquardt 迭代. 作为求单变量函数零点方法的 Newton 迭代的一般概念为学习数值方法的大多数学生所熟悉. 它可以相当直接地推广到多变量情形并可用于求方程组的最小二乘解(不是精确解).

Levenberg - Marquardt 方法是 Newton 迭代的一种简单的改进方法,旨在为超参数化问题提供更快的收敛和正则化. 该方法也可视为 Newton 迭代和下降法的混合.

就本书所考虑的问题类型而言,计算的复杂度的实质性减少是通过把参数集合分成两部分而得到. 一般,这两部分由表示摄像机矩阵或单应的参数集和表示点的参数集组成. 其结果会导致一种稀疏结构,下面将就这个问题进行介绍.

## A4.1　Newton 迭代

给定一种假设性的函数关系 $\mathbf{X}=f(\mathbf{P})$,其中 $\mathbf{X}$ 是**测量矢量**,$\mathbf{P}$ 是**参数矢量**,分别属于欧氏空间 $\mathbb{R}^N$和 $\mathbb{R}^M$. 假定已知逼近于真值 $\overline{\mathbf{X}}$ 的测量值 $\mathbf{X}$,希望求出矢量 $\hat{\mathbf{P}}$能最近似地满足这个函数关系. 更准确地说,我们要求满足 $\mathbf{X}=f(\hat{\mathbf{P}})+\boldsymbol{\varepsilon}$ 的矢量 $\hat{\mathbf{P}}$并使 $\|\boldsymbol{\varepsilon}\|$ 最小化. 注意:节 A3.3.1中所考虑的线性最小二乘问题正好属于这种类型,所定义的函数 $f$ 是一个线性函数 $f(\mathbf{P})=\mathrm{A}\mathbf{P}$.

为了求 $f$ 不是一个线性函数时的解,Newton 迭代方法从一个初始估计值 $\mathbf{P}_0$ 开始,并在函数 $f$ 是局部线性的假设下逐次对这个估计进行改进. 令 $\boldsymbol{\varepsilon}_0$ 由 $\mathbf{X}=f(\mathbf{P}_0)+\boldsymbol{\varepsilon}_0$ 定义. 我们假定函数 $f$ 在点 $\mathbf{P}_0$ 的值由 $f(\mathbf{P}_0+\boldsymbol{\Delta})=f(\mathbf{P}_0)+\mathrm{J}\boldsymbol{\Delta}$ 逼近,其中 J 是由 Jacobi 矩阵 $\mathrm{J}=\partial\mathbf{X}/\partial\mathbf{P}$ 表示的线性映射. 我们寻找一个点 $f(\mathbf{P}_1)$(其中 $\mathbf{P}_1=\mathbf{P}_0+\boldsymbol{\Delta}$),最小化 $\mathbf{X}-f(\mathbf{P}_1)=\mathbf{X}-f(\mathbf{P}_0)-\mathrm{J}\boldsymbol{\Delta}=\boldsymbol{\varepsilon}_0-\mathrm{J}\boldsymbol{\Delta}$. 因此,它要求关于 $\boldsymbol{\Delta}$ 最小化 $\|\boldsymbol{\varepsilon}_0-\mathrm{J}\boldsymbol{\Delta}\|$,这是一个线性最小化问题. 该最小化问题的解矢量 $\boldsymbol{\Delta}$ 由解**正规**方程组 $\mathrm{J}^{\mathrm{T}}\mathrm{J}\boldsymbol{\Delta}=\mathrm{J}^{\mathrm{T}}\boldsymbol{\varepsilon}_0$(见(A3.9))或用伪逆 $\boldsymbol{\Delta}=\mathrm{J}^{+}\boldsymbol{\varepsilon}_0$ 而求得. 因此,解矢量 $\hat{\mathbf{P}}$ 由一个估计 $\mathbf{P}_0$ 开始并按下面的公式计算其逐次逼近

$$\mathbf{P}_{i+1}=\mathbf{P}_i+\boldsymbol{\Delta}_i,$$

其中 $\boldsymbol{\Delta}_i$ 是线性最小二乘问题的解

$$\mathrm{J}\boldsymbol{\Delta}_i=\boldsymbol{\varepsilon}_i.$$

矩阵 J 是取值在 $\mathbf{P}_i$ 上的 Jacobian 矩阵 $\partial\mathbf{X}/\partial\mathbf{P}$ 而 $\boldsymbol{\varepsilon}_i=\mathbf{X}-f(\mathbf{P}_i)$. 我们希望这个算法能收敛于所要求的最小二乘解 $\hat{\mathbf{P}}$. 不幸的是这个迭代过程有可能只收敛到一个局部最小或根本不收敛. 迭代算法的行为强烈地依赖于初始估计 $\mathbf{P}_0$.

**加权 Newton 迭代**　在 Newton 迭代中所有因变量都等量加权,除此之外,也可以提供一个

加权的矩阵规定因变量 $\mathbf{X}$ 的加权值. 更具体地说,我们可以假定测量矢量 $\mathbf{X}$ 满足一个协方差矩阵为 $\Sigma_X$ 的高斯分布,并希望最小化 Mahalanobis 距离 $\|\mathbf{X}-f(\hat{\mathbf{P}})\|_\Sigma$. 这协方差矩阵可以是对角的,说明 $\mathbf{X}$ 的每个坐标是独立的,或可以是更一般的任意正定对称矩阵. 在这种情形,正规方程变为 $J^T\Sigma^{-1}J\mathbf{\Delta}_i=J^T\Sigma^{-1}\boldsymbol{\varepsilon}_i$. 算法的其余部分保持不变.

## A4.2 Levenberg - Marquardt 迭代

Levenberg - Marquardt(简称 LM)迭代方法是对 Newton 迭代方法略作了一点改变. 正规方程 $N\mathbf{\Delta}=J^TJ\mathbf{\Delta}=J^T\boldsymbol{\varepsilon}$ 被**增量正规方程** $N'\mathbf{\Delta}=J^T\boldsymbol{\varepsilon}$ 替代,其中 $N'_{ii}=(1+\lambda)N_{ii}$ 而且当 $i\neq j$ 时 $N'_{ij}=N_{ij}$[①]. 开始时值 $\lambda$ 被设定为某个值,通常 $\lambda=10^{-3}$. 如果通过解增量正规方程得到的 $\mathbf{\Delta}$ 值导致误差减少,那么该增量被接受并在下一次迭代之前将 $\lambda$ 除以 10. 反之,如果 $\mathbf{\Delta}$ 值导致误差增加,那么将 $\lambda$ 乘以 10 并重解增量正规方程. 继续这一过程直到求出一个使误差下降的 $\mathbf{\Delta}$ 为止. 对不同的 $\lambda$ 重复地解增量正规方程直到求出一个可接受的 $\mathbf{\Delta}$,整个过程组成 LM 算法的一次迭代. LM 算法的实现在[Press - 88]中给出.

**LM 的合理性** 为了理解该方法的合理性,考虑对于大小不同的 $\lambda$ 会产生什么现象. 当 $\lambda$ 非常小时,该方法与 Newton 迭代本质上没有不同. 如果误差函数 $\|\mathbf{X}-f(\mathbf{P})\|$ 接近于 $\mathbf{P}$ 的二次函数,那么 Newton 迭代将很快地收敛到最小值. 另一方面,当 $\lambda$ 为一个大的值时(本质上大于 1),$J^TJ$ 的非对角元素相对于对角元素而言变得不重要. $J^TJ$ 的第 $i$ 个对角元素是 $\sum_{j=1}^{N}(\partial x_j/\partial p_i)^2$,其中 $x_j$ 是 $\mathbf{X}$ 的第 $j$ 个坐标而 $p_i$ 是第 $i$ 个参数. 令 $\mathbf{V}_i$ 表示矢量 $\partial\mathbf{X}/\partial p_i$,则 $J^TJ$ 的第 $i$ 个对角元素是 $\mathbf{V}_i^T\mathbf{V}_i$. 忽略非对角元素,正规方程组 $J^TJ\mathbf{\Delta}=J^T\boldsymbol{\varepsilon}$ 分解成一组简单的方程 $(\lambda+1)\mathbf{V}_i^T\mathbf{V}_i\delta p_i=\mathbf{V}_i^T\boldsymbol{\varepsilon}$,其中 $\delta p_i$ 是 $\mathbf{\Delta}$ 的第 $i$ 个元素. 其解是 $\delta p_i=(1+\lambda)^{-1}(\mathbf{V}_i^T\boldsymbol{\varepsilon})/\|\mathbf{V}_i\|^2$.

现在,矢量 $\boldsymbol{\varepsilon}$ 是测量矢量所要寻求的变化,即由当前估计 $\hat{\mathbf{X}}_i=f(\mathbf{P}_i)$ 指向被测量点 $\mathbf{X}$ 的矢量. 这正是我们希望移动的方向. 但是,通过参数 $p_i$ 的增量,事实上是在方向 $\mathbf{V}_i=\partial\mathbf{X}/\partial p_i$ 上移动. 商 $(\mathbf{V}_i^T\boldsymbol{\varepsilon})/\|\mathbf{V}_i\|^2$ 是一种度量,表示在方向 $\mathbf{V}_i$ 移动一个单位距离时在方向 $\boldsymbol{\varepsilon}$ 移动了多远. 这个方向上的一个充分小的移动可能使指向目标 $\mathbf{X}$ 的距离几乎不变. 因此,用 $p_i+(1+\lambda)^{-1}\delta p_i$ 代替 $p_i$ 希望足够大的 $\lambda$ 将使指向目标的距离递减. 每个参数也同时增加,与它对到 $\mathbf{X}$ 的距离的影响成正比同时与 $\|\mathbf{V}\|$ 的大小成反比.

因此,LM 在 Newton 迭代和下降方法之间无缝地移动,前者将造成在解的邻域附近迅速收敛,而后者在运行困难时保证代价函数是下降的.

**LM 的实现** 在执行 Levenberg - Marquardt 最小化的最简单形式时,仅需要提供一个例行程序来计算被最小化的函数、被观察的目标矢量 $\hat{\mathbf{X}}$ 或该函数的期望值和一个初始估计 $P_0$. Jacobi矩阵 J 的计算可以用数值方法或提供一个用户程序来实现.

数值微分可以按如下方法实现. 每个独立变量 $x_i$ 依次增加为 $x_i+\delta$,用计算 $f$ 的程序计算函数值并用比率计算微分. 通过设 $\delta$ 为 $|10^{-4}\times x_i|$ 和 $10^{-6}$ 中的最大值已经找到了好的

① 这里的分析反映了[Press - 88]中的算法实现. 其他的专业人士则喜欢用相加而不是相乘的方法对 N 进行摄动.

结果. 这种选择似乎能给出微分的一种好的逼近. 在实际中,我们发现用数值微分几乎没有什么缺点. 但是对简单的函数 $f$,我们倾向于提供一种计算 J 的程序,其原因部分为了完美、部分为了有可能略微加速收敛以及部分为了速度.

## A4.3 稀疏 Levenberg－Marquardt 算法

节 A4.2 中介绍的 LM 算法对参数数目少的最小化问题是相当适宜的. 因此,在 2D 单应估计(见第 3 章)的简单的代价函数(3.6－$p$58)和(3.7－$p$58)(它们仅对单应矩阵 H 的元素最小化)等时,LM 算法行之有效. 但是当要最小化的代价函数包含很大数目的参数时,单用 LM 算法就不很合适. 这是因为 LM 的核心步骤(解正规方程(A3.9))的复杂度是 $N^3$,$N$ 为参数数目,并且这一步要重复许多次. 然而,在本书所考虑的估计问题的类型中,正规方程组有某种稀疏分块结构,我们可以利用它来节约大量计算时间.

应用这种方法的一个例子是通过最小化代价函数(3.8－$p$58)来估计两视图之间的 2D 单应并假定误差在两幅图像中都存在. 这个问题可以用刻划 2D 单应的参数集合(可能是单应矩阵的 9 个元素),加上第一幅视图的 $n$ 个点的参数,总共有 $2n+9$ 个参数来参数化.

应用这种方法的另一个例子是图像的对应跨两幅或更多(例如 $m$)视图并且要求估计所有摄像机的参数以及所有点的 3D 位置的重构问题. 我们可以假定摄像机是任意的射影摄像机或者是完全或部分标定的摄像机. 进一步,为了去掉一些非主要的自由度我们可以固定一个摄像机. 例如射影重构问题,可以用所有摄像机矩阵的元素(总共 $12m$ 或 $11m$ 个参数,这取决于摄像机如何参数化)加上 3D 点的坐标的 $3n$ 个来参数化.

稀疏 LM 算法的实现通常被认为是复杂和困难的. 为了有助于克服它,采用菜单方式给出这些算法. 如果有一个适当的矩阵运算标准程序库,实现该算法应该不困难.

• **记号**:如果 $\mathbf{a}_1,\mathbf{a}_2,\cdots\mathbf{a}_n$ 是矢量,那么将它们一个接着一个地排成一个列矢量,所得到的矢量用$(\mathbf{a}_1^\mathsf{T},\mathbf{a}_2^\mathsf{T},\cdots,\mathbf{a}_n^\mathsf{T})^\mathsf{T}$ 表示. 关于矩阵也有类似的记号.

### A4.3.1 在 LM 方法中划分参数

我们将主要结合重构问题对稀疏 LM 方法进行介绍,因为重构问题是与这种方法相关的典型问题. 首先我们将用一般术语来处理估计问题,因为它可以说明一般方法且不涉及更多细节. 在此抽象的层次上,这种方法的好处并不明显,但在 A4.3.3 中它将会变得更加清楚. 我们首先把问题的参数分成两个集合:一个是描述摄像机的参数集,另一个是描述点的参数集. 更正规地说,"参数矢量"$\mathbf{P}\in\mathbb{R}^M$ 可以划分成参数矢量 $\mathbf{a}$ 和 $\mathbf{b}$,使得 $\mathbf{P}=(\mathbf{a}^\mathsf{T},\mathbf{b}^\mathsf{T})^\mathsf{T}$. 我们已知的是用空间 $\mathbb{R}^N$ 的 $\mathbf{X}$ 表示的一个"测量矢量". 在重构问题中,它是所有图像点的坐标所组成的矢量. 另外令 $\Sigma_\mathbf{X}$ 为测量矩阵的协方差矩阵①. 我们考虑把参数矢量 $\mathbf{P}$ 映射到测量矢量的估计 $\hat{\mathbf{X}}=f(\mathbf{P})$的 一般函数 $f$: $\mathbb{R}^M\rightarrow\mathbb{R}^N$. 记 $\boldsymbol{\varepsilon}$ 为测量值和估计值之间的差 $\mathbf{X}-\hat{\mathbf{X}}$,我们要求使 Mahalanobis 距离的平方$\|\boldsymbol{\varepsilon}\|^2_{\Sigma_\mathbf{X}}=\boldsymbol{\varepsilon}^\mathsf{T}\Sigma_\mathbf{X}^{-1}\boldsymbol{\varepsilon}$ 最小化的一组参数.

对应于参数 $\mathbf{P}=(\mathbf{a}^\mathsf{T},\mathbf{b}^\mathsf{T})^\mathsf{T}$ 的划分,Jacobi 矩阵 $\mathrm{J}=[\partial\hat{\mathbf{X}}/\partial\mathbf{P}]$具有形如 $\mathrm{J}=[\mathrm{A}|\mathrm{B}]$的分块结

① 在没有其它信息时,我们通常假定$\Sigma_\mathbf{X}$ 是单位矩阵

构,其中的 Jacobi 子矩阵定义为

$$A=[\partial\hat{\mathbf{X}}/\partial\mathbf{a}]$$

和

$$B=[\partial\hat{\mathbf{X}}/\partial\mathbf{b}].$$

解方程组 $J\boldsymbol{\delta}=\boldsymbol{\varepsilon}$ 是 LM 算法的中心步骤(见节 A4.2)现在具有形式

$$J\boldsymbol{\delta}=[A|B]\begin{pmatrix}\boldsymbol{\delta}_a\\ \boldsymbol{\delta}_b\end{pmatrix}=\boldsymbol{\varepsilon}. \tag{A4.1}$$

于是,在 LM 算法中每一步要解的正规方程 $J^T\Sigma_X^{-1}J\boldsymbol{\delta}=J^T\Sigma_X^{-1}\boldsymbol{\varepsilon}$ 具有形式

$$\left[\begin{array}{c|c}A^T\Sigma_X^{-1}A & A^T\Sigma_X^{-1}B\\ \hline B^T\Sigma_X^{-1}A & B^T\Sigma_X^{-1}B\end{array}\right]\begin{pmatrix}\boldsymbol{\delta}_a\\ \boldsymbol{\delta}_b\end{pmatrix}=\begin{pmatrix}A^T\Sigma_X^{-1}\boldsymbol{\varepsilon}\\ B^T\Sigma_X^{-1}\boldsymbol{\varepsilon}\end{pmatrix}. \tag{A4.2}$$

在 LM 算法的这一点上,这个矩阵的对角块通过用一个因子 $1+\lambda$ 乘它们的对角元素而增加了参量,其中 $\lambda$ 是变化的参数. 这个增量改变了矩阵 $A^T\Sigma_X^{-1}A$ 和 $B^T\Sigma_X^{-1}B$. 所得矩阵记为 $(A^T\Sigma_X^{-1}A)^*$ 和 $(B^T\Sigma_X^{-1}B)^*$.

方程组(A4.2)现在可以写成分块形式

$$\begin{bmatrix}U^* & W\\ W^T & V^*\end{bmatrix}\begin{pmatrix}\boldsymbol{\delta}_a\\ \boldsymbol{\delta}_b\end{pmatrix}=\begin{pmatrix}\boldsymbol{\varepsilon}_A\\ \boldsymbol{\varepsilon}_B\end{pmatrix}. \tag{A4.3}$$

作为解此方程组的第一步,将两边左乘 $\begin{bmatrix}I & -WV^{*-1}\\ 0 & I\end{bmatrix}$ 得到

$$\begin{bmatrix}U^*-WV^{*-1}W^T & 0\\ W^T & V^*\end{bmatrix}\begin{pmatrix}\boldsymbol{\delta}_a\\ \boldsymbol{\delta}_b\end{pmatrix}=\begin{pmatrix}\boldsymbol{\varepsilon}_A-WV^{*-1}\boldsymbol{\varepsilon}_B\\ \boldsymbol{\varepsilon}_B\end{pmatrix}. \tag{A4.4}$$

这样做的结果使右上角的子块变为 0. 从而该方程组的上半部分成为

$$(U^*-WV^{*-1}W^T)\boldsymbol{\delta}_a=\boldsymbol{\varepsilon}_A-WV^{*-1}\boldsymbol{\varepsilon}_B. \tag{A4.5}$$

解这些方程求得 $\boldsymbol{\delta}_a$. 然后,$\boldsymbol{\delta}_b$ 的值可以通过回代求得,即

$$V^*\boldsymbol{\delta}_b=\boldsymbol{\varepsilon}_B-W^T\boldsymbol{\delta}_a. \tag{A4.6}$$

如节 A4.2 所述,如果参数矢量 $\mathbf{P}=((\mathbf{a}+\boldsymbol{\delta}_a)^T,(\mathbf{b}+\boldsymbol{\delta}_b)^T)^T$ 的新计算值使误差函数的值减少,那么便接受该新参数矢量 $\mathbf{P}$,并把 $\lambda$ 值减少为十分之一后再进行下一次迭代. 另一方面,若误差的值增加,则拒绝新的 $\mathbf{P}$ 并用一个增加为十倍的新 $\lambda$ 再作试探.

一个完整的 Levenberg - Marquardt 划分参数算法在算法 A4.1 中给出.

虽然在此方法中,我们先求解 $\boldsymbol{\delta}_a$,然后基于新的 $\mathbf{a}$ 值求解 $\boldsymbol{\delta}_b$,但千万别认为这方法不过就是关于 $\mathbf{a}$ 和 $\mathbf{b}$ 的独立迭代. 如果我们要求保持 $\mathbf{b}$ 为常数条件下求解 $\mathbf{a}$,那么用来解 $\boldsymbol{\delta}_a$ 的正规方程将具有比 A4.5 更简单的形式:$U\boldsymbol{\delta}_a=\boldsymbol{\varepsilon}_A$. 但是我们不推荐交替解 $\boldsymbol{\delta}_a$ 和 $\boldsymbol{\delta}_b$ 的方法,因为潜伏着收敛很慢的问题.

### A4.3.2 协方差

在结论 4.12($p$94)中,我们看到估计参数的协方差矩阵为

$$\Sigma_P=(J^T\Sigma_X^{-1}J)^+. \tag{A4.7}$$

在超参数化的情形,由(A4.7)给出的协方差矩阵 $\Sigma_P$ 是奇异的,特别是在垂直于约束曲面的方向上不允许参数有变化——在这些方向上方差为零.

在参数集合被划分成 $\mathbf{P}=(\mathbf{a}^{\mathsf{T}},\mathbf{b}^{\mathsf{T}})^{\mathsf{T}}$ 的情形，矩阵 $(J^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}J)$ 具有由(A4.2)和(A4.3)给出的分块形式(但不是带星的增量形式). 我们有

$$J^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}J=\begin{bmatrix}A^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}A & A^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}B\\ B^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}A & B^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}B\end{bmatrix}=\begin{bmatrix}U & W\\ W^{\mathsf{T}} & V\end{bmatrix}. \tag{A4.8}$$

协方差矩阵 $\Sigma_{\mathbf{P}}$ 是这个矩阵的伪逆. 在 V 是可逆的假定下，重定义 $Y=WV^{-1}$. 那么矩阵可以按下式对角化

$$J^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}J=\begin{bmatrix}U & W\\ W^{\mathsf{T}} & V\end{bmatrix}=\begin{bmatrix}I & Y\\ 0 & I\end{bmatrix}\begin{bmatrix}U-WV^{-1}W^{\mathsf{T}} & 0\\ 0 & V\end{bmatrix}\begin{bmatrix}I & 0\\ Y^{\mathsf{T}} & I\end{bmatrix}. \tag{A4.9}$$

我们假定：对于可逆矩阵 G 和矩阵 H，恒等式

$$(GHG^{\mathsf{T}})^{+}=G^{-\mathsf{T}}H^{+}G^{-1}.$$

成立. 在本附录末的练习(1)提出的条件下该恒等式成立. 把它用于(A4.9)并乘开来即得该伪逆的公式

$$\Sigma_{\mathbf{P}}=(J^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}J)^{+}=\begin{bmatrix}X & -XY\\ -Y^{\mathsf{T}}X & Y^{\mathsf{T}}XY+V^{-1}\end{bmatrix}, \tag{A4.10}$$

其中 $X=(U-WV^{-1}W^{\mathsf{T}})^{+}$.

使此式成立的条件是 $\mathrm{Span}(A)\cap\mathrm{Span}(B)=\varnothing$，其中 $\mathrm{Span}(\cdot)$ 表示该矩阵列的生成空间. 这里 A 和 B 和(A4.8)中的一样. 这一事实的证明以及 $\mathrm{Span}(A)\cap\mathrm{Span}(B)=\varnothing$ 的条件的解释在 *p*425 的练习(1)中有扼要说明.

把矩阵 $J^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}J$ 按(A4.8)划分为块与(A4.2)中将 **P** 划分为 **a** 和 **b** 相对应. 对参数矢量 **P** 的协方差矩阵的截取分别得到参数 **a** 和 **b** 的协方差矩阵. 结论概括在算法 A4.2 中.

已知　测量矢量 **X** 以及它的协方差矩阵 $\Sigma_{\mathbf{X}}$，参数集合 $\mathbf{P}=(\mathbf{a}^{\mathsf{T}},\mathbf{b}^{\mathsf{T}})^{\mathsf{T}}$ 的一个初始估计和将参数矢量映射到测量矢量估计的函数 $f:\mathbf{P}\mapsto\hat{\mathbf{X}}$.

目标　求最小化 $\boldsymbol{\varepsilon}^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}\boldsymbol{\varepsilon}$ 的参数集合 **P**，其中 $\boldsymbol{\varepsilon}=\mathbf{X}-\hat{\mathbf{X}}$.

算法

(1) 初始化一个常数 $\lambda=0.001$(典型值).

(2) 计算微分矩阵 $A=[\partial\hat{\mathbf{X}}/\partial\mathbf{a}]$ 和 $B=[\partial\hat{\mathbf{X}}/\partial\mathbf{b}]$ 以及误差矢量 $\boldsymbol{\varepsilon}$.

(3) 计算中间表达式

$$U=A^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}A \qquad V=B^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}B \qquad W=A^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}B \qquad \boldsymbol{\varepsilon}_A=A^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}\boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon}_B=B^{\mathsf{T}}\Sigma_{\mathbf{X}}^{-1}\boldsymbol{\varepsilon}.$$

(4) 用 $1+\lambda$ 乘 U 和 V 的对角元素.

(5) 计算逆矩阵 $V^{*-1}$ 并定义 $Y=WV^{*-1}$. 此逆矩阵可以覆盖 $V^{*}$ 的值因后者已不再需要了.

(6) 通过解 $(U^{*}-YW^{\mathsf{T}})\,\boldsymbol{\delta}_a=\boldsymbol{\varepsilon}_A-Y\boldsymbol{\varepsilon}_B$ 求 $\boldsymbol{\delta}_a$.

(7) 通过回代求得 $\boldsymbol{\delta}_b=V^{*-1}(\boldsymbol{\varepsilon}_B-W^{\mathsf{T}}\boldsymbol{\delta}_a)$.

(8) 通过增加增量矢量 $(\boldsymbol{\delta}_a^{\mathsf{T}},\boldsymbol{\delta}_b^{\mathsf{T}})^{\mathsf{T}}$ 更新参数矢量并计算新的误差矢量.

(9) 如果新的误差矢量小于老的误差，则接受这个新的参数值，把 $\lambda$ 减少至十分之一并且返回步骤(2)重新开始，否则终止运行.

(10) 如果新的误差矢量大于老的误差，则回到老参数值 $\lambda$ 并将其增加至十倍再返回步骤(4)继续进行.

算法 A4.1　划分参数的 Levenberg－Marquardt 算法.

**目标**

计算用算法 A4.1 估计的参数 **a** 和 **b** 的协方差.

**算法**

(1) 与算法 A4.1 一样地计算 U,V 和 W,并令 $Y = WV^{-1}$.

(2) $\Sigma_{\mathbf{a}} = (U - WV^{-1}W^{\mathsf{T}})^{+}$.

(3) $\Sigma_{\mathbf{b}} = Y^{\mathsf{T}}\Sigma_{\mathbf{a}}Y + V^{-1}$.

(4) 交叉协方差 $\Sigma_{\mathbf{ab}} = -\Sigma_{\mathbf{a}}Y$.

算法 A4.2 计算 LM 参数的协方差矩阵.

### A4.3.3 稀疏 LM 的一般方法

在前几页中,我们介绍了在参数向量可以划分为两个子矢量 **a** 和 **b** 的条件下实现 LM 迭代并计算解的协方差的一种方法. 由前面的讨论我们还不清楚所介绍的方法在一般情形是否真有任何计算上的优越性. 但是正如下面所述,这种方法在 Jacobi 矩阵满足一定的稀疏条件时会变得很重要.

我们假定"测量矢量"$\mathbf{X} \in \mathbb{R}^N$可以划分成若干段:$\mathbf{X} = (\mathbf{X}_1^{\mathsf{T}}, \mathbf{X}_2^{\mathsf{T}}, \ldots, \mathbf{X}_n^{\mathsf{T}})^{\mathsf{T}}$. 类似地,假定"参数矢量"$\mathbf{P} \in \mathbb{R}^M$也可相应地划分成 $\mathbf{P} = (\mathbf{a}^{\mathsf{T}}, \mathbf{b}_1^{\mathsf{T}}, \mathbf{b}_2^{\mathsf{T}}, \ldots, \mathbf{b}_n^{\mathsf{T}})^{\mathsf{T}}$. 对应于参数的一个给定的指派,$\mathbf{X}_i$ 的估计值将由 $\hat{\mathbf{X}}_i$ 表示. 我们作下列**稀疏假定**:每个 $\hat{\mathbf{X}}_i$ 仅由 **a** 和 $\mathbf{b}_i$ 决定,而与其它的参数 $\mathbf{b}_j$ 无关. 在此假设下,当 $i \neq j$ 时 $\partial\hat{\mathbf{X}}_i/\partial\mathbf{b}_j = 0$. 而对 $\partial\hat{\mathbf{X}}_i/\partial\mathbf{a}$ 没有做假设. 这种情形出现在本讨论开始所描述的重构问题中,其中 $\mathbf{b}_i$ 是第 $i$ 点的参数矢量而 $\hat{\mathbf{X}}_i$ 是这个点在所有视图的图像中的测量矢量. 对于这种问题,因为一个点的图像不依赖于任何其它的点,所以正如我们所希望的有 $\partial\hat{\mathbf{X}}_i/\partial\mathbf{b}_j = 0$,除非 $i = j$.

对应于这样的划分,Jacobi 矩阵 $J = [\partial\hat{\mathbf{X}}/\partial\mathbf{P}]$有一个稀疏分块结构. 我们定义 Jacobi 矩阵

$$A_i = [\partial\hat{\mathbf{X}}_i/\partial\mathbf{a}] \qquad B_i = [\partial\hat{\mathbf{X}}_i/\partial\mathbf{b}_i].$$

对给定形如 $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1^{\mathsf{T}}, \ldots, \boldsymbol{\varepsilon}_n^{\mathsf{T}})^{\mathsf{T}} = \mathbf{X} - \hat{\mathbf{X}}$ 的误差矢量,方程组 $J\boldsymbol{\delta} = \boldsymbol{\varepsilon}$ 现在有形式

$$J\boldsymbol{\delta} = \left[\begin{array}{c|cccc} A_1 & B_1 & & & \\ A_2 & & B_2 & & \\ \vdots & & & \ddots & \\ A_n & & & & B_n \end{array}\right] \left(\begin{array}{c} \boldsymbol{\delta}_{\mathbf{a}} \\ \hline \boldsymbol{\delta}_{\mathbf{b}_1} \\ \vdots \\ \boldsymbol{\delta}_{\mathbf{b}_n} \end{array}\right) = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \vdots \\ \boldsymbol{\varepsilon}_n \end{pmatrix}. \tag{A4.11}$$

我们进一步假定所有测量 $\mathbf{X}_i$ 的协方程矩阵 $\Sigma_{\mathbf{X}_i}$ 相互独立. 因此整个测量矢量的协方差矩阵 $\Sigma_{\mathbf{X}}$ 具有对角形式 $\Sigma_{\mathbf{X}} = \mathrm{diag}(\Sigma_{\mathbf{X}_1}, \ldots, \Sigma_{\mathbf{X}_n})$.

根据算法 A4.1 的记号,我们得到

$$\begin{aligned} A &= [A_1^{\mathsf{T}}, A_2^{\mathsf{T}}, \ldots, A_n^{\mathsf{T}}]^{\mathsf{T}} \\ B &= \mathrm{diag}(B_1, B_2, \ldots, B_n) \\ \Sigma_{\mathbf{X}} &= \mathrm{diag}(\Sigma_{\mathbf{X}_1}, \ldots, \Sigma_{\mathbf{X}_n}) \\ \boldsymbol{\delta}_{\mathbf{b}} &= (\boldsymbol{\delta}_{\mathbf{b}_1}^{\mathsf{T}}, \boldsymbol{\delta}_{\mathbf{b}_2}^{\mathsf{T}}, \ldots, \boldsymbol{\delta}_{\mathbf{b}_n}^{\mathsf{T}})^{\mathsf{T}} \\ \boldsymbol{\varepsilon} &= (\boldsymbol{\varepsilon}_1^{\mathsf{T}}, \boldsymbol{\varepsilon}_2^{\mathsf{T}}, \ldots, \boldsymbol{\varepsilon}_n^{\mathsf{T}})^{\mathsf{T}}. \end{aligned}$$

现在，将这些公式直接代入算法 A4.1. 结果在算法 A4.3 中给出，它表示 LM 算法其中一步的计算方法. 重要的事情是在此形式中，算法的每一步所需的计算时间为 $n$ 的线性函数. 如果不利用稀疏结构的特点(例如盲目地用算法 A4.1)，计算复杂度会是 $n^3$.

## A4.4 用稀疏 LM 算法估计 2D 单应

给定两幅图像中的对应图像点集 $\mathbf{x}_i \leftrightarrow \mathbf{x}_i'$，我们应用前面的讨论来估计 2D 单应 H. 每幅图像中的点都受到噪声的影响，并且我们的目标是最小化代价函数(3.8 - $p$58). 定义一个测量矢量 $\mathbf{X}_i = (\mathbf{x}_i^\mathsf{T}, \mathbf{x}_i'^\mathsf{T})^\mathsf{T}$. 在当前情形下，参数矢量 $\mathbf{P}$ 可以划分成 $\mathbf{P} = (\mathbf{h}^\mathsf{T}, \hat{\mathbf{x}}_1^\mathsf{T}, \hat{\mathbf{x}}_2^\mathsf{T}, \ldots, \hat{\mathbf{x}}_n^\mathsf{T})^\mathsf{T}$，其中值 $\hat{\mathbf{x}}_i$ 是第一幅图像中图像点的估计值，而 $\mathbf{h}$ 是单应 H 的元素组成的矢量. 因此，我们必须同时估计单应 H 和第一幅图像的每个点的参数. 函数 $f$ 把 $\mathbf{P}$ 映射到 $(\hat{\mathbf{X}}_1^\mathsf{T}, \hat{\mathbf{X}}_2^\mathsf{T}, \ldots, \hat{\mathbf{X}}_n^\mathsf{T})^\mathsf{T}$，其中每一个 $\hat{\mathbf{X}}_i = (\hat{\mathbf{x}}_i^\mathsf{T}, \mathrm{H}\hat{\mathbf{x}}_i^\mathsf{T})^\mathsf{T} = (\hat{\mathbf{x}}_i^\mathsf{T}, \hat{\mathbf{x}}_i'^\mathsf{T})^\mathsf{T}$. 然后直接应用算法 A4.3.

此时，Jacobi 矩阵有其特殊形式，首先因为 $\hat{\mathbf{x}}_i$ 与 $\mathbf{h}$ 无关，故有

$$\mathrm{A}_i = \partial\hat{\mathbf{X}}_i/\partial\mathbf{h} = \begin{bmatrix} 0 \\ \partial\hat{\mathbf{x}}_i'/\partial\mathbf{h} \end{bmatrix},$$

其次因为 $\hat{\mathbf{X}}_i = (\hat{\mathbf{x}}_i^\mathsf{T}, \hat{\mathbf{x}}_i'^\mathsf{T})^\mathsf{T}$，又有

$$\mathrm{B}_i = \partial\hat{\mathbf{X}}_i/\partial\hat{\mathbf{x}}_i = \begin{bmatrix} \mathrm{I} \\ \partial\hat{\mathbf{x}}_i'/\partial\hat{\mathbf{x}}_i \end{bmatrix}.$$

### A4.4.1 协方差的计算

作为一个例子，我们来计算协方差，考虑和节 4.2.4($p$95)一样的问题，其中单应由点对应来估计. 同样，我们仅考虑被估计的单应实际上是恒等映射 H = I. 就这个例子的目的而言，点的数目和它们的分布是不重要的. 但仍假定所有点测量的误差互相独立. 回忆误差仅出现在第二幅图像的情形，公式(4.9 - $p$96)给出 $\Sigma_{\mathbf{h}} = (\Sigma_i \mathrm{J}_i^\mathsf{T} \Sigma_i^{-1} \mathrm{J}_i)^+$，其中 $\mathrm{J}_i = [\partial\hat{\mathbf{x}}_i'/\partial\mathbf{h}]$.

我们现在来计算摄像机参数矢量 $\mathbf{h}$ 在第一幅图像的点也有噪声时的协方差. 我们进一步假设 $\Sigma_{\mathbf{x}_i}^{-1} = \Sigma_{\mathbf{x}_i'}^{-1}$，并用 $\mathrm{S}_i$ 表示. 在这种情形下，协方差的逆矩阵 $\Sigma_{\mathbf{X}}^{-1}$ 是分块对角阵，$\Sigma_{\mathbf{X}}^{-1} = \mathrm{diag}(\Sigma_{\mathbf{x}_i}^{-1}, \Sigma_{\mathbf{x}_i'}^{-1})$. 那么，用 A4.3 算法的步骤来计算 $\mathbf{h}$ 的协方差矩阵如下：

$$\begin{aligned}
\mathrm{A}_i &= [0^\mathsf{T}, \mathrm{J}_i^\mathsf{T}]^\mathsf{T} \\
\mathrm{B}_i &= [\mathrm{I}^\mathsf{T}, \mathrm{I}^\mathsf{T}]^\mathsf{T} (\because \mathrm{H} = \mathrm{I}) \\
\mathrm{U} &= \sum_i \mathrm{A}_i^\mathsf{T} \mathrm{diag}(\mathrm{S}_i, \mathrm{S}_i) \mathrm{A}_i = \sum_i \mathrm{J}_i^\mathsf{T} \mathrm{S}_i \mathrm{J}_i \\
\mathrm{V}_i &= \mathrm{B}_i^\mathsf{T} \mathrm{diag}(\mathrm{S}_i, \mathrm{S}_i) \mathrm{B}_i = 2\mathrm{S}_i \\
\mathrm{W}_i &= \mathrm{A}_i^\mathsf{T} \mathrm{diag}(\mathrm{S}_i, \mathrm{S}_i) \mathrm{B}_i = \mathrm{J}_i^\mathsf{T} \mathrm{S}_i \\
\mathrm{U} - \sum_i \mathrm{W}_i \mathrm{V}_i^{-1} \mathrm{W}_i^\mathsf{T} &= \sum_i \mathrm{J}_i^\mathsf{T} (\mathrm{S}_i - \mathrm{S}_i/2) \mathrm{J}_i = \sum_i \mathrm{J}_i^\mathsf{T} \mathrm{S}_i \mathrm{J}_i/2 \\
\Sigma_{\mathbf{h}} &= 2(\sum_i \mathrm{J}_i^\mathsf{T} \mathrm{S}_i \mathrm{J}_i)^+.
\end{aligned}$$

因此 $\mathbf{h}$ 的协方差矩阵是误差仅在一幅图像中出现时的协方差矩阵的两倍值. 这是由于 H 是

恒等映射的结果，但一般它不成立．作为练习，可以验证如下结论：

- **如果 H 表示具有因子 $s$ 的缩放($H=sI$)，那么** $\Sigma_{\mathbf{b}}=(s^2+1)(\sum_i J_i^T S_i J_i)^+$.
- **如果 H 是一个仿射变换，而 D 是 H 左上部的 2×2 部分(非平移部分)，并且如果对所有 $i$ 有 $S_i=I$(各向同性和噪声独立)，那么** $\Sigma_{\mathbf{b}}=(\sum_i J_i^T(I-D(I+D^TD)^{-1}D^T)J_i)^+$.

<u>目标</u>　参数矢量分块成 $\mathbf{P}=(\mathbf{a}^T,\mathbf{b}_1^T,\cdots,\mathbf{b}_n^T)^T$，测量矢量分块成 $\mathbf{X}=(\mathbf{X}_1^T,\cdots,\mathbf{X}_n^T)^T$并使得对 $\forall\, i\neq j,\partial\hat{\mathbf{X}}_i/\partial\mathbf{b}_j=0$时的 LM 算法．

<u>算法</u>　算法 A4.1 中的第(2)到第(7)步变成：

(1)计算微分矩阵 $A_i=[\partial\hat{\mathbf{X}}_i/\partial\mathbf{a}]$和 $B_i=[\partial\hat{\mathbf{X}}_i/\partial\mathbf{b}_i]$和误差矢量 $\boldsymbol{\varepsilon}_i=\mathbf{X}_i-\hat{\mathbf{X}}_i$.

(2)计算中间值

$$U=\sum_i A_i^T\Sigma_{\mathbf{X}_i}^{-1}A_i$$

$$V=\mathrm{diag}(V_1,\cdots,V_n),\text{其中 } V_i=B_i^T\Sigma_{\mathbf{X}_i}^{-1}B_i$$

$$W=[W_1,W_2,\cdots,W_n],\text{其中 } W_i=A_i^T\Sigma_{\mathbf{X}_i}^{-1}B_i$$

$$\boldsymbol{\varepsilon}_A=\sum_i A_i^T\Sigma_{\mathbf{X}_i}^{-1}\boldsymbol{\varepsilon}_i$$

$$\boldsymbol{\varepsilon}_B=(\boldsymbol{\varepsilon}_{B_1}^T,\boldsymbol{\varepsilon}_{B_2}^T,\cdots,\boldsymbol{\varepsilon}_{B_n}^T)^T,\text{其中 } \boldsymbol{\varepsilon}_{B_i}=B_i^T\Sigma_{\mathbf{X}_i}^{-1}\boldsymbol{\varepsilon}_i$$

$$Y_i=W_iV_i^{*-1}.$$

(3)由方程

$$(U^*-\sum_i Y_iW_i^T)\boldsymbol{\delta}_{\mathbf{a}}=\boldsymbol{\varepsilon}_A-\sum_i Y_i\boldsymbol{\varepsilon}_{B_i}$$

计算 $\boldsymbol{\delta}_{\mathbf{a}}$.

(4)由方程 $\boldsymbol{\delta}_{\mathbf{b}_i}=V_i^{*-1}(\boldsymbol{\varepsilon}_{B_i}-W_i^T\boldsymbol{\delta}_{\mathbf{a}})$计算每个 $\boldsymbol{\delta}_{\mathbf{b}_i}$.

<u>协方差</u>

(1)重新定义 $Y_i=W_iV_i^{-1}$

(2)$\Sigma_{\mathbf{a}}=(U-\sum_i Y_iW_i^T)^*$

(3)$\Sigma_{\mathbf{b}_i\mathbf{b}_j}=Y_i^T\Sigma_{\mathbf{a}}Y_j+\delta_{ij}V_i^{-1}$

(4)$\Sigma_{\mathbf{a}\mathbf{b}_i}=-\Sigma_{\mathbf{a}}Y_i$

算法 A4.3　稀疏 Levenberg—Marquardt 算法.

## A4.5　用稀疏 LM 算法估计基本矩阵

在估计基本矩阵和 3D 点集时，节 A4.3.3 介绍的算法是有效的，只要将估计 2D 单应的稀疏 LM 算法稍做修改就能适用于现在的目的．这个与 2D 单应估计问题类似的方法是：在 2D 单应估计中，我们有一个映射 H 把点 $\mathbf{x}_i$ 对应于点 $\mathbf{x}_i'$；对于当前的问题，映射要由一对摄像机矩阵 P 和 P′ 表示，它们将一个 3D 点映射到一对对应点($\mathbf{x}_i,\mathbf{x}_i'$).

为方便起见，这里仍使用节 A4.3.3 的记号．具体地说，在 A4.3.3 和这里 **X** 都表示总的测量矢量(当前情形为($\mathbf{x}_1,\mathbf{x}_1',\ldots,\mathbf{x}_n,\mathbf{x}_n'$))而不是单个的 3D 点．同样，要仔细区别参数矢量 **P** 和摄像机矩阵 P.

将参数矢量 $\mathbf{P}$ 划分为 $\mathbf{P}=(\mathbf{a}^{\mathrm{T}},\mathbf{b}_1{}^{\mathrm{T}},\ldots,\mathbf{b}_n{}^{\mathrm{T}})^{\mathrm{T}}$，其中

(1)$\mathbf{a}=\mathbf{p}'$ 由摄像机矩阵 P′的的元素组成，而

(2)$\mathbf{b}_i=(\mathrm{X}_i,\mathrm{Y}_i,\mathrm{T}_i)^{\mathrm{T}}$ 是 3 维矢量，是第 $i$ 个 3D 点$(\mathrm{X}_i,\mathrm{Y}_i,1,\mathrm{T}_i)^{\mathrm{T}}$的参数化.

这样一来，总共有 $3n+12$ 个参数，其中 $n$ 是点的数目. 参数矢量 $\mathbf{a}$ 提供摄像机 P′的一个参数化，而另一个摄像机 P 取为$[\mathrm{I}|\mathbf{0}]$. 注意为了方便这里把 3D 点的第三个坐标取为 1，而且这样做也是允许的. 因为点$(\mathrm{X}_i,\mathrm{Y}_i,0,\mathrm{T}_i)^{\mathrm{T}}$映射到无穷远点$(\mathrm{X}_i,\mathrm{Y}_i,0)^{\mathrm{T}}$，它不会是被测量点$(x_i,y_i,1)^{\mathrm{T}}$ 的邻近点.

测量矢量 $\mathbf{X}$ 被划分为 $\mathbf{X}=(\mathbf{X}_1{}^{\mathrm{T}},\mathbf{X}_2{}^{\mathrm{T}},\cdots,\mathbf{X}_n{}^{\mathrm{T}})^{\mathrm{T}}$，其中 $\mathbf{X}_i=(\mathbf{x}_i{}^{\mathrm{T}},\mathbf{x}_i'^{\mathrm{T}})^{\mathrm{T}}$ 是第 $i$ 点的图像测量.

现在，可以算出 Jacobi 矩阵 $\mathrm{A}_i=\partial\hat{\mathbf{X}}_i/\partial\mathbf{a}$ 和 $\mathrm{B}_i=\partial\hat{\mathbf{X}}_i/\partial\mathbf{b}_i$ 并用算法 A4.3 估计参数，从而求出 P′并由它可以算出 F.

**偏导数矩阵**

因为 $\hat{\mathbf{X}}_i=(\hat{\mathbf{x}}_i^{\mathrm{T}},\hat{\mathbf{x}}_i'^{\mathrm{T}})^{\mathrm{T}}$，那么 $\mathrm{A}_i$ 和 $\mathrm{B}_i$ 具有与节 A4.4 中的 Jacobi 矩阵的类似形式：

$$\mathrm{A}_i=\begin{bmatrix}0\\ \partial\hat{\mathbf{x}}_i'/\partial\mathbf{a}\end{bmatrix}\qquad \mathrm{B}_i=\begin{bmatrix}\mathrm{I}_{2\times2}|\mathbf{0}\\ \partial\hat{\mathbf{x}}_i'/\partial\mathbf{b}_i\end{bmatrix}.$$

协方差矩阵 $\Sigma_{\mathbf{X}_i}$ 为对角分块矩阵 $\mathrm{diag}(\mathrm{S}_i,\mathrm{S}_i')$，其中 $\mathrm{S}_i=\Sigma_{\mathbf{x}_i}^{-1}$ 和 $\mathrm{S}_i'=\Sigma_{\mathbf{x}_i'}^{-1}$. 现在，计算算法 A4.3 第二步的中间表达式得到

$$\mathrm{V}_i=\mathrm{B}_i^{\mathrm{T}}\mathrm{diag}(\mathrm{S}_i,\mathrm{S}_i')\mathrm{B}_i=[\mathrm{I}_{2\times2}|\mathbf{0}]^{\mathrm{T}}\mathrm{S}_i[\mathrm{I}_{2\times2}|\mathbf{0}]+(\partial\hat{\mathbf{x}}_i'/\partial\mathbf{b}_i)^{\mathrm{T}}\mathrm{S}_i'(\partial\hat{\mathbf{x}}_i'/\partial\mathbf{b}_i).\quad(\mathrm{A4.12})$$

$\mathrm{A}_i^{\mathrm{T}}\Sigma_{\mathbf{X}_i}^{-1}\mathrm{A}_i$ 的抽象形式与 2D 单应的情形一样，而其它的表达式 $\mathrm{W}_i=\mathrm{A}_i^{\mathrm{T}}\Sigma_{\mathbf{X}_i}^{-1}\mathrm{B}_i$，$\boldsymbol{\varepsilon}_{\mathrm{B}_i}=\mathrm{B}_i^{\mathrm{T}}\Sigma_{\mathbf{X}_i}^{-1}\boldsymbol{\varepsilon}_i$，$\boldsymbol{\varepsilon}_{\mathrm{A}_i}=\mathrm{A}_i^{\mathrm{T}}\Sigma_{\mathbf{X}_i}^{-1}\boldsymbol{\varepsilon}_i$ 可容易计算. 除此之外，其余估计过程完全与算法 A4.3 一样.

**F 的协方差**

根据节 A4.3.3 的讨论，特别是根据算法 A4.3，摄像机参数(即 P′的元素)的协方差矩阵，由

$$\Sigma_{\mathrm{P}'}=(\mathrm{U}-\sum_i\mathrm{W}_i\mathrm{V}_i^{-1}\mathrm{W}_i^{\mathrm{T}})^{+}\qquad(\mathrm{A4.13})$$

给出，其中的记号和算法 A4.3 一样.

在计算该伪逆时，知道 $\Sigma_{\mathrm{P}'}$ 的秩的期望是有益的. 此时它的秩为 7，因为涉及两个摄像机和 $n$ 个点匹配解的总自由度是 $3n+7$. 用另一种途径来看，P′不是唯一确定的，因为如果 $\mathrm{P}'=[\mathrm{M}|\mathbf{m}]$，那么任何其它的矩阵$[\mathrm{M}+\mathbf{t}\mathbf{m}^{\mathrm{T}}|\alpha\mathbf{m}]$也确定同样的基本矩阵. 因此，在计算(A4.13)右边的伪逆时，我们应该令 5 个奇异值为零.

前面的讨论给出如何计算 P′的元素的协方差矩阵. 我们希望计算 F 的元素的协方差. 我们知道，存在一个用 $\mathrm{P}'=[\mathrm{M}|\mathbf{m}]$的元素来表示 F 元素的简单公式：$\mathrm{F}=[\mathbf{m}]_\times\mathrm{M}$. 如果我们希望计算归一化的 F(即$\|\mathrm{F}\|=1$)的协方差矩阵，那么令 $\mathrm{F}=[\mathbf{m}]_\times\mathrm{M}/(\|[\mathbf{m}]_\times\mathrm{M}\|)$. 因此，我们可以将 F 的元素表示成 P′的元素的简单函数. 令 J 为这函数的 Jacobi 矩阵. 那么 F 的协方差可以利用结论 4.6($p91$)由传播 P′的协方差来计算，即

$$\Sigma_{\mathrm{F}}=\mathrm{J}\Sigma_{\mathrm{P}'}\mathrm{J}^{\mathrm{T}}=\mathrm{J}(\mathrm{U}-\sum_i\mathrm{W}_i\mathrm{V}_i^{-1}\mathrm{W}_i^{\mathrm{T}})^{+}\mathrm{J}^{\mathrm{T}},\qquad(\mathrm{A4.14})$$

其中$\Sigma_{P'}$由(A4.13)给出. 这就是由给定点对应并**根据 ML 算法估计得到**的基本矩阵的协方差.

## A4.6 用稀疏 LM 算法作多幅图像的捆集调整

前一节考虑了用稀疏 Levenberg－Marquardt 算法计算基本矩阵. 它本质上就是由两幅视图进行重构的问题. 它如何可以容易地推广到计算三焦点张量和四焦点张量应该是清楚的. 更一般地,我们可以用它来同时估计多个摄像机和对应点集,从而计算射影结构,或在给定的适当约束下计算仿射或度量结构. 这个技术称为**捆集调整**.

对于多个摄像机的情形,我们可以利用不同摄像机的参数之间没有交互作用的有利条件,这正是我们现在要给出的. 在下面的讨论中,为了标记上的方便,我们将假定每一点在所有视图中可视. 这个假定并不是必要的,因为测量点一般仅可能在有效视图的某个子集中可视.

我们使用与节 A4.3.3 一样的记号. 具体地说,测量数据可以表示为一个矢量 $\mathbf{X}$,它又划分成表示某个 3D 点在所有视图中测量得到的图像坐标的分段 $\mathbf{X}_i$. 我们可以进一步划分 $\mathbf{X}_i$ 为 $\mathbf{X}_i=(\mathbf{x}_{i1}^{\mathsf{T}},\mathbf{x}_{i2}^{\mathsf{T}},\ldots,\mathbf{x}_{im}^{\mathsf{T}})^{\mathsf{T}}$,其中 $\mathbf{x}_{ij}$是第 $i$ 点在第 $j$ 幅图像中的图像. 参数矢量 $\mathbf{a}$(摄像机参数)可以对应地划分成 $\mathbf{a}=(\mathbf{a}_1{}^{\mathsf{T}},\mathbf{a}_2{}^{\mathsf{T}},\ldots,\mathbf{a}_m{}^{\mathsf{T}})^{\mathsf{T}}$,其中 $\mathbf{a}_j$ 是第 $j$ 个摄像机的参数. 因为图像点 $\mathbf{x}_{ij}$仅与第 $j$ 个摄像机的参数有关,而与任何其它的摄像机无关,故除非 $j=k$, 否则 $\partial\hat{\mathbf{x}}_{ij}/\partial\mathbf{a}_k=0$. 类似地,除非 $i=k$, 否则参数 $\mathbf{b}_k$ 对第 $k$ 个 3D 点的导数为 $\partial\hat{\mathbf{x}}_{ij}/\partial\mathbf{b}_k=0$.

参考算法 A4.3 中 Jacobi 矩阵的定义,我们看到 $A_i=[\partial\hat{\mathbf{X}}_i/\partial\mathbf{a}]$是一个对角分块矩阵:$A_i=\mathrm{diag}(A_{i1},\ldots,A_{im})$,其中 $A_{ij}=\partial\hat{\mathbf{x}}_{ij}/\partial\mathbf{a}_j$. 类似地,矩阵 $B_i=[\partial\hat{\mathbf{X}}_i/\partial\mathbf{b}_i]$ 分解为 $B_i=[B_{i1}^{\mathsf{T}},\ldots,B_{im}^{\mathsf{T}}]^{\mathsf{T}}$,其中 $B_{ij}=\partial\hat{\mathbf{x}}_{ij}/\partial\mathbf{b}_i$. 在标准情况下也可以假定$\Sigma_{\mathbf{X}_i}$有对角结构$\Sigma_{\mathbf{X}_i}=\mathrm{diag}(\Sigma_{\mathbf{x}_{i1}},\ldots,\Sigma_{\mathbf{x}_{im}})$,其含义是被投影的点在各个图像中的测量是独立的(或更准确地说是不相关的). 有了这些假设,我们可以容易地利用算法 A4.3,并在算法 A4.4 中给出(留给读者去证明).

## A4.7 练习

(1)我们用下列各步来证明在(A4.10)中给出的分块矩阵的伪逆的形式.

(a)回顾 $H^+=G(G^{\mathsf{T}}HG)^{-1}G^{\mathsf{T}}$ 的充要条件是为 $N_L(G)=N_L(H)$(见节 A3.4).

(b)令 G 可逆. 那么$(GHG^{\mathsf{T}})^+=G^{-\mathsf{T}}H^+G^{-1}$的充要条件 $N_L(H)G^{\mathsf{T}}=N_L(H)G^{-1}$.

(c)把此条件应用于(A4.9)得使(A4.10)成立的充要条件是 $N_L(U-WV^{-1}W^{\mathsf{T}})\subseteq N_L(Y)=N_L(W)$.

(d)与(A4.8)一样地,用 A 和 B 来定义 U,V 和 W,则上面的充要条件等价于条件:$\mathrm{span}(A)\cap\mathrm{span}(B)=\emptyset$.

(2)研究在什么条件下 $\mathrm{span}(A)\cap\mathrm{span}(B)=\emptyset$ 成立. 它可以解释成变换参数 $\mathbf{a}$(例如摄像机参数)的效应和变换 $\mathbf{b}$(点参数)的效应不能互补. 显然(例如)非约束射影重构就不属于这种情况,其中摄像机和点都可以变化而不影响测量. 在这种情形下,参

数 a 和 b 在方向 $\boldsymbol{\delta}_{\mathbf{a}}$ 和 $\boldsymbol{\delta}_{\mathbf{b}}$ 上的变化是无穷的，其中 $\boldsymbol{\delta}_{\mathbf{a}}, \boldsymbol{\delta}_{\mathbf{b}}$ 满足 $A\boldsymbol{\delta}_{\mathbf{a}} = B\boldsymbol{\delta}_{\mathbf{b}}$.

(1)计算导数矩阵 $A_{ij} = [\partial \hat{\mathbf{x}}_{ij}/\partial \mathbf{a}_j]$ 和 $B_{ij} = [\partial \hat{\mathbf{x}}_{ij}/\partial \mathbf{b}_i]$ 以及误差矢量 $\boldsymbol{\varepsilon}_{ij} = \mathbf{x}_{ij} - \hat{\mathbf{x}}_{ij}$.

(2)计算中间值

$$U_j = \sum_i A_{ij}^{\mathrm{T}} \Sigma_{\mathbf{x}_{ij}}^{-1} A_{ij} \qquad V_i = \sum_j B_{ij}^{\mathrm{T}} \Sigma_{\mathbf{x}_{ij}}^{-1} B_{ij} \qquad W_{ij} = A_{ij}^{\mathrm{T}} \Sigma_{\mathbf{x}_{ij}}^{-1} B_{ij}$$

$$\boldsymbol{\varepsilon}_{\mathbf{a}_j} = \sum_i A_{ij}^{\mathrm{T}} \Sigma_{\mathbf{x}_{ij}}^{-1} \boldsymbol{\varepsilon}_{ij} \qquad \boldsymbol{\varepsilon}_{\mathbf{b}_i} = \sum_j B_{ij}^{\mathrm{T}} \Sigma_{\mathbf{x}_{ij}}^{-1} \boldsymbol{\varepsilon}_{ij} \qquad Y_{ij} = W_{ij} V_i^{*-1}$$

其中 $i = 1, \ldots, n$ 和 $j = 1, \ldots, m$.

(3)由下列方程计算 $\boldsymbol{\delta}_{\mathbf{a}} = (\boldsymbol{\delta}_{\mathbf{a}_1}^{\mathrm{T}}, \ldots, \boldsymbol{\delta}_{\mathbf{a}_m}^{\mathrm{T}})^{\mathrm{T}}$:

$$S\boldsymbol{\delta}_{\mathbf{a}} = (\mathbf{e}_1^{\mathrm{T}}, \ldots, \mathbf{e}_m^{\mathrm{T}})^{\mathrm{T}},$$

其中 S 是 $m \times m$ 分块矩阵，它的块 $S_{jk}$ 定义为

$$S_{jk} = \delta_{jk} U_j^* - \sum_i Y_{ij} W_{ik}^{\mathrm{T}}$$

和

$$\mathbf{e}_j = \boldsymbol{\varepsilon}_{\mathbf{a}_j} - \sum_i Y_{ij} \boldsymbol{\varepsilon}_{\mathbf{b}_i}.$$

(4)依次由方程

$$\boldsymbol{\delta}_{\mathbf{b}_i} = V_i^{*-1} (\boldsymbol{\varepsilon}_{\mathbf{b}_i} - \sum_j W_{ij}^{\mathrm{T}} \boldsymbol{\delta}_{\mathbf{a}_j})$$

计算每个 $\boldsymbol{\delta}_{\mathbf{b}_i}$.

<u>协方差</u>

(1) 重新定义 $Y_{ij} = W_{ij} V_i^{-1}$.

(2) $\Sigma_{\mathbf{a}} = S^{+}$，其中 S 的定义如上且没有用 * 表示的扩展.

(3) $\Sigma_{\mathbf{b}_i \mathbf{b}_j} = Y_i^{\mathrm{T}} \Sigma_{\mathbf{a}} Y_j + \delta_{ij} V_i^{-1}$.

(4) $\Sigma_{\mathbf{a}\mathbf{b}_i} = -\Sigma_{\mathbf{a}} Y_i$.

**算法 A4.4　一般的稀疏 Levenberg－Marquardt 算法.**

# 附录 5　某些特殊的平面射影变换

射影变换(单应变换)可以根据它们的特征值的代数重数和几何重数进行分类. 一个特征值的代数重数是指它作为特征方程根的重复次数. 几何重数可以由矩阵$(H-\lambda I)$的秩来确定,其中 H 是单应变换而 $\lambda$ 是特征值. 一种完全的分类在射影几何的教科书中给出,例如[Springer－64]. 这里我们仅提到在实用中重要且在本书中多处出现的若干特殊情况. 这里将讨论平面变换(H 是 $3\times3$ 矩阵),但可直接向 3 维空间推广.

这些特殊形式之所以重要是由于 H 还满足若干关系(回忆一般射影变换的唯一约束是满秩). 因 H 满足一些约束故它有较少的自由度,从而它可以用比一般的射影变换少的对应来计算. 同时这种特殊变换比一般的变换有更丰富的几何性质和不变量.

注意:与第 1 章中讨论的能构成子群的那些特殊形式(如仿射)不一样,下面介绍的特殊射影变换一般不构成群,因为它们在乘法下一般不封闭. 只有当其中所有的元素均有重合的不动点和不动线时(即它们仅是特征值不同时),它们才构成一个子群.

## A5.1　共轭旋转

一个旋转矩阵 R 有特征值$\{1,e^{i\theta},e^{-i\theta}\}$,分别对应于特征矢量$\{\mathbf{a},\mathbf{I},\mathbf{J}\}$,其中 $\mathbf{a}$ 是旋转轴,即 $R\mathbf{a}=\mathbf{a}$,$\theta$ 是绕此轴的旋转角,而 $\mathbf{I}$ 和 $\mathbf{J}$(它们互为复共轭)是与 $\mathbf{a}$ 正交的平面的虚圆点. 假定两平面之间的一个射影变换的形式是

$$H=TRT^{-1},$$

其中 T 是一般射影变换;那么 H 是一个**共轭旋转**. 特征值在共轭关系①下被保持,因此射影变换 H 的特征值在相差一个公共因子的意义下也是$\{1,e^{i\theta},e^{-i\theta}\}$.

考虑由摄像机绕它的中心旋转得到的两幅图像(如图 1.5($p$11)$b$);那么如节 7.4.2($p$141)所证明了的,这两幅图像由一个共轭旋转相关联. 在此情形下,其复特征值确定摄像机旋转角度 $\theta$,而对应于实特征值的特征矢量是旋转轴的消影点. 注意 $\theta$——度量不变量——可以直接由射影变换测量到.

## A5.2　平面透射

一个平面射影变换 H 称为平面透射,如果它有一条由不动点组成的直线(称为轴)和不在该直线上的一个不动点(称为顶点)(见图 A5.1). 从代数上说,该矩阵有两个相等和另一个不同的特征值并且对应于相等特征值的特征空间是 2 维的. 其轴是过生成这 2 维特征空间的两个特征矢量(即两点)的直线. 顶点对应于另一个特征矢量. 不相同的特征值与重复特征值的比率是透射的特征不变量 $\mu$(即不计比例因子的差别,其特征值是$\{\mu,1,1\}$).

① 共轭也称"相似变换". 这个"相似"的意思与本书所用的相似变换为带缩放的等矩变换的含义毫无关系.

平面透射的性质包括

- 连接对应点的直线相交于顶点，对应直线（即通过两组对应点的直线）相交于轴．这是 Desargues 定理的一个例子，见图 A5.2a.
- 由顶点、一组对应点和连接这些点的直线与不动点的直线的交点所定义的交比对所有与该透射相关的点都是一样的．见图 A5.2b.
- 对曲线及它的平面透射的曲线图像而言，其对应的切线（定义为切点与邻点连线的极限）相交于轴.
- 顶点(2dof)，轴(2dof)和不变量交比(1dof)足以完全定义透射．因此平面透射有 5 个自由度.
- 3 组匹配点足以计算一个平面透射．这些点匹配的 6 个自由度超约束了 5 个自由度的透射.

平面透射自然发生于由 3 维空间的透视变换相关联的两个平面的图像中（即连接两个平面的对应点的直线交于一点）．这种变换的例子是一个平面物体的图像和它在一个平面上的阴影的图像之间的变换．在此情形下，轴是两平面交线的图像，而顶点是光源的图像，见图 1.5($p$11)c.

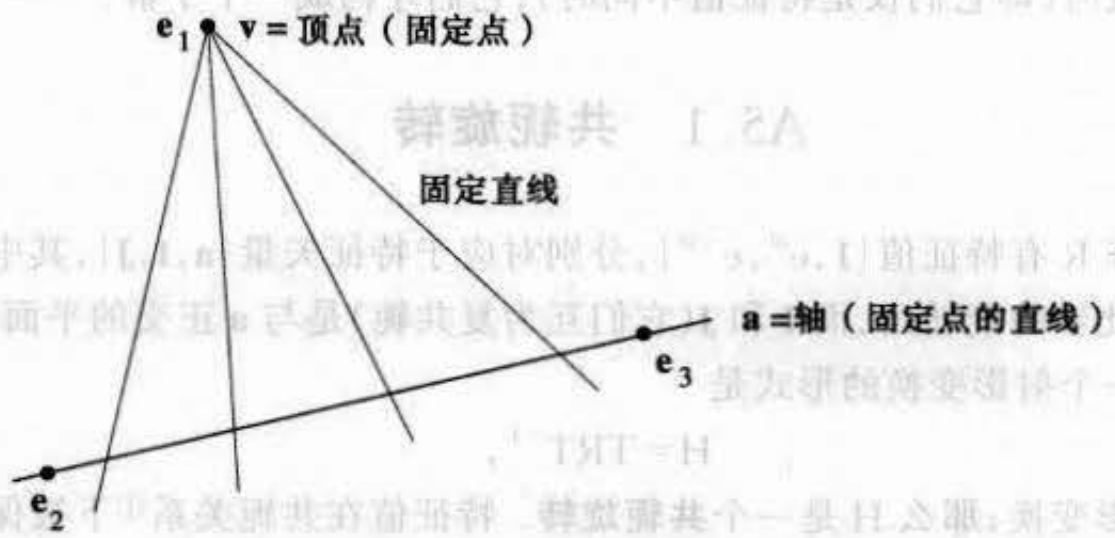



图 A5.1 **平面透射**．平面透射是一种平面射影变换：它有一条由不动点组成的直线（称为轴）a 和不在该直线上的一个不动点 v（称为透射的中心或顶点）．存在过顶点的一束不动线．从代数上说，变换矩阵的两个特征值相等（图中重复特征值是 $\lambda_2$ 和 $\lambda_3$），而不动直线与矩阵的 2D 不变空间对应.

**参数化** 表示透射的射影变换可以直接用表示轴 a 和顶点 v 的 3 维矢量以及特征比率 $\mu$ 来参数化，即

$$H = I + (\mu - 1)\frac{\mathbf{v}\mathbf{a}^T}{\mathbf{v}^T\mathbf{a}},$$

其中 I 是单位矩阵．可以验证它的逆变换由下式给出

$$H^{-1} = I + (\frac{1}{\mu} - 1)\frac{\mathbf{v}\mathbf{a}^T}{\mathbf{v}^T\mathbf{a}}.$$

其特征矢量是

$$\{\mathbf{e}_1 = \mathbf{v}, \mathbf{e}_2 = \mathbf{a}_1^{\perp}, \mathbf{e}_3 = \mathbf{a}_2^{\perp}\},$$

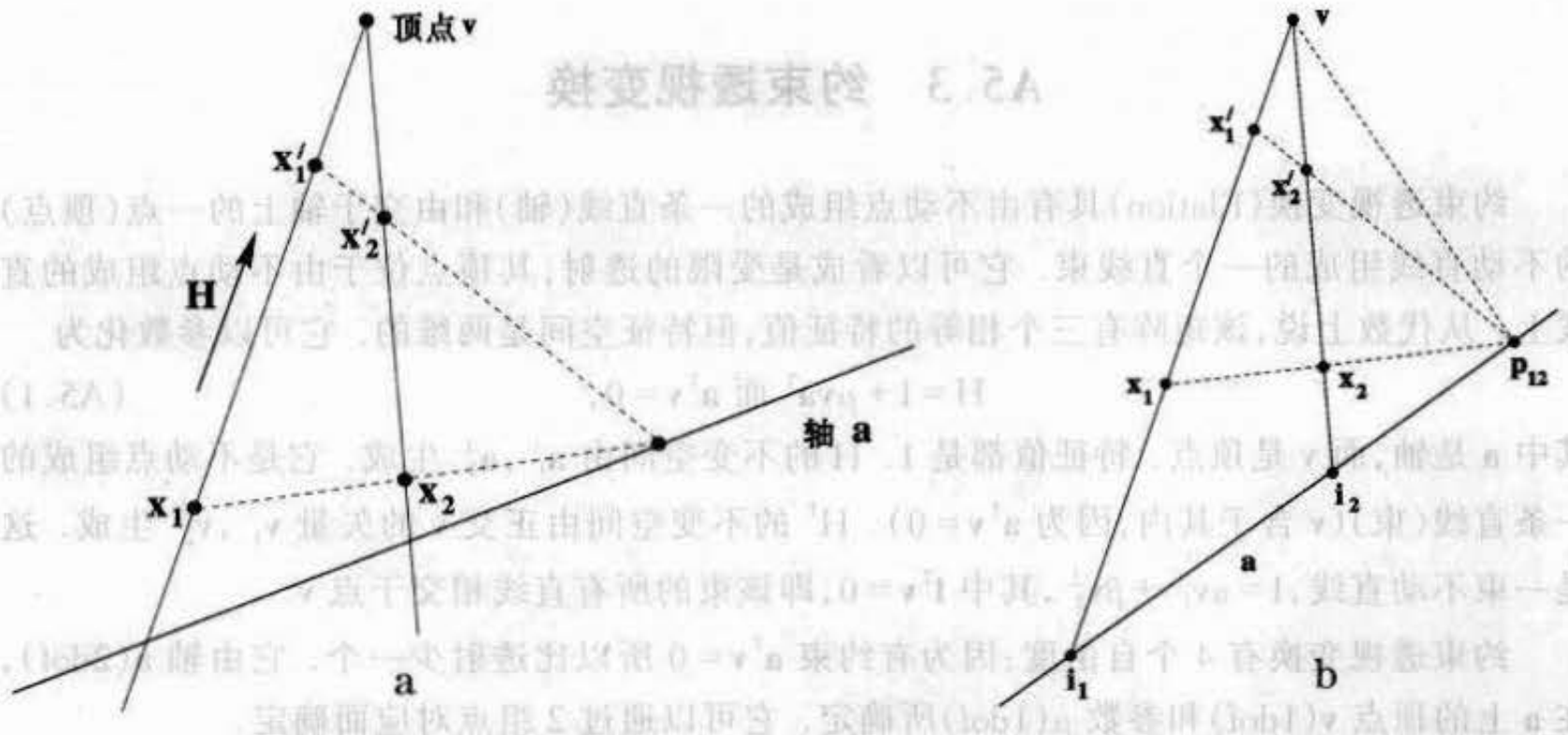



图 A5.2 **透射变换**. (a)在此变换下,轴上的点被映射为自身;不在轴上的每个点都位于过 $\mathbf{v}$ 并与 $\mathbf{a}$ 相交的一条不动直线上且被映射为此直线上的另一个点. 因此,对应点对 $\mathbf{x}\leftrightarrow\mathbf{x}'$ 和透射的顶点共线. 对应线——即过对应点对的直线——相交于轴上:例如直线 $<\mathbf{x}_1,\mathbf{x}_2>$ 和 $<\mathbf{x}_1',\mathbf{x}_2'>$. (b)由顶点 $\mathbf{v}$、对应点 $\mathbf{x}$、$\mathbf{x}'$ 和它们与轴的交点 $\mathbf{i}$ 定义的交比是透射的特征不变量,并且所有对应点取值相同. 例如,四点 $\{\mathbf{v},\mathbf{x}_1',\mathbf{x}_1,\mathbf{i}_1\}$ 和四点 $\{\mathbf{v},\mathbf{x}_2',\mathbf{x}_2,\mathbf{i}_2\}$ 的交比相等,因为它们通过共点于 $\mathbf{p}_{12}$ 的直线透视相关. 由此推出与透射相关的所有点的交比都相同.

对应的特征值为

$$\{\lambda_1=\mu,\lambda_2=1,\lambda_3=1\},$$

其中 $\mathbf{a}_i^{\perp}$ 是生成与 3 维矢量 $\mathbf{a}$ 正交的空间的两个矢量,即 $\mathbf{a}^{\mathrm{T}}\mathbf{a}_i^{\perp}=0$ 和 $\mathbf{a}=\mathbf{a}_1^{\perp}\times\mathbf{a}_2^{\perp}$.

如果轴或顶点在无穷远,那么透射是一个仿射变换. 从代数上说,如果 $\mathbf{a}=(0,0,1)^{\mathrm{T}}$,那么轴在无穷远;或如果 $\mathbf{v}=(v_1,v_2,0)^{\mathrm{T}}$,那么顶点在无穷远;并且在这两种情形下变换矩阵 H 的最后一行都是(0,0,1).

**平面调和透射** 平面透射的一种特殊情形是交比是调和的($\mu=-1$). 这样的平面透射称为平面调和透射,并因交比不变量已知故其自由度为 4. 变换矩阵 H 满足 $\mathrm{H}^2=\mathrm{I}$,即变换是恒等矩阵的平方根,它称为对合(也称周期为 2 的保线变换). 在相差一个公共尺度因子的情况下,其特征值是 $\{-1,1,1\}$. 两组点对应决定 H.

在具有双侧对称的透视图像中,图像中的对应点依照一个平面调和透射相关联. 该透射的轴是对称轴的图像. 从代数上说,H 是一个共轭反射,其中共轭元素是一个平面射影变换. 在仿射图像中(由仿射摄像机产生),所产生的变换是一个**反对称**,而且共轭元素是一个平面仿射变换. 反对称的顶点为无穷远点,连接对应点的直线互相平行.

调和透射可以参数化为

$$\mathrm{H}=\mathrm{H}^{-1}=\mathrm{I}-2\frac{\mathbf{v}\mathbf{a}^{\mathrm{T}}}{\mathbf{v}^{\mathrm{T}}\mathbf{a}}.$$

同样,如果轴或顶点为无穷远点,那么该变换是仿射的.

## A5.3 约束透视变换

约束透视变换(Elation)具有由不动点组成的一条直线(轴)和由交于轴上的一点(顶点)的不动直线组成的一个直线束. 它可以看成是受限的透射,其顶点位于由不动点组成的直线上. 从代数上说,该矩阵有三个相等的特征值,但特征空间是两维的. 它可以参数化为

$$H = I + \mu \mathbf{v}\mathbf{a}^T \text{ 而 } \mathbf{a}^T\mathbf{v} = 0, \tag{A5.1}$$

其中 $\mathbf{a}$ 是轴,而 $\mathbf{v}$ 是顶点. 特征值都是 1. H 的不变空间由 $\mathbf{a}_1^\perp, \mathbf{a}_2^\perp$ 生成. 它是不动点组成的一条直线(束)($\mathbf{v}$ 含于其内,因为 $\mathbf{a}^T\mathbf{v}=0$). $H^T$ 的不变空间由正交 $\mathbf{v}$ 的矢量 $\mathbf{v}_1^\perp, \mathbf{v}_2^\perp$ 生成. 这是一束不动直线,$\mathbf{l} = \alpha\mathbf{v}_1^\perp + \beta\mathbf{v}_2^\perp$,其中 $\mathbf{l}^T\mathbf{v}=0$,即该束的所有直线相交于点 $\mathbf{v}$.

约束透视变换有 4 个自由度:因为有约束 $\mathbf{a}^T\mathbf{v}=0$ 所以比透射少一个. 它由轴 $\mathbf{a}$(2dof),在 $\mathbf{a}$ 上的顶点 $\mathbf{v}$(1dof)和参数 $\mu$(1dof)所确定. 它可以通过 2 组点对应而确定.

在实际中约束透视变换通常作为共轭平移出现. 考虑在平面上实施重复平移的一种模式,例如在建筑物的墙上完全相同的窗子. 对于墙平面的作用可以表示为

$$H_E = \begin{bmatrix} I & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix},$$

其中 $\mathbf{t} = (t_x, t_y, 0)^T$ 是重复窗的平移方向,而 $\mathbf{a} = (0,0,1)^T$ 是无穷远直线. 图像中的窗由共轭位移 $H = TH_ET^{-1}$ 相互关联,其中 T 是把墙平面映射到该图像的射影变换. 对于这种情形,H 的顶点是位移方向的消影点,而轴是墙平面的消影线.

## A5.4 透视变换

射影变换的另一个特殊情形是**透视变换**,图 A5.3 给出平面上的一个 1D 射影变换. 透视变换的最重要的一个性质是连接对应点的直线共点. 透视变换和射影变换之间的区别可以通过考虑两个透视变换的合成来弄清楚. 如图 A5.4 所示,两个透视变换的复合一般**不是**透视变换. 但是,该复合变换**是**射影变换,**因为透视变换是射影变换且射影变换构成群(封闭). 所以两个射影变换的复合是射影变换. 概括起来:**

- **两(或更多)个透视变换的复合是射影变换,但一般不是透视变换.**

一个世界平面的中心投影图像(如图 1.3($p$9)所示)是不同平面之间的 2D 透视变换的一个例子. 注意辨认一个射影变换是一个透视变换需要将有关平面嵌入 3 维空间.

最后,试想象图 1.3($p$9)的两个平面和摄像机中心都被(另一个透视变换)映射到二平面之一. 那么这透视变换是在同一平面上的点之间的一个新的映射,并被看作是一个平面透射(节 A5.2).

**进一步的阅读参考** [Springer-64]对射影变换进行了分类并讨论了一些特殊情况,例如平面透射. 平面透射以许多不同面貌出现于文献中:在[VanGool-98]中是阴影图像关系的建模;在[Zisserman-95b]中是挤压曲面图像的建模;在[Basri-99]中是平面姿态恢复的关系的建模. 平面透射的参数化在 Viéville 和 Lingrand[Viéville-95]中给出. 约束透视变换出现在平面上重复模式的图像分组中[Schaffalitzky-99]而在 3 维空间中它们出现在广义的

浅浮雕的多义性中[Kriegman－98].

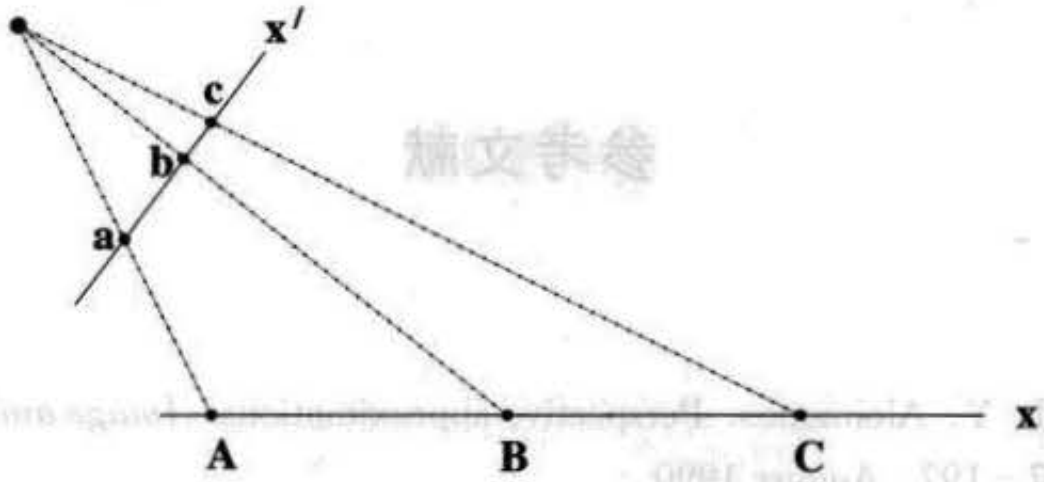



图 A5.3 **直线的透视变换**. 连接对应点(**a**,**A** 等)的直线共点. 请与图 A5.4 比较.

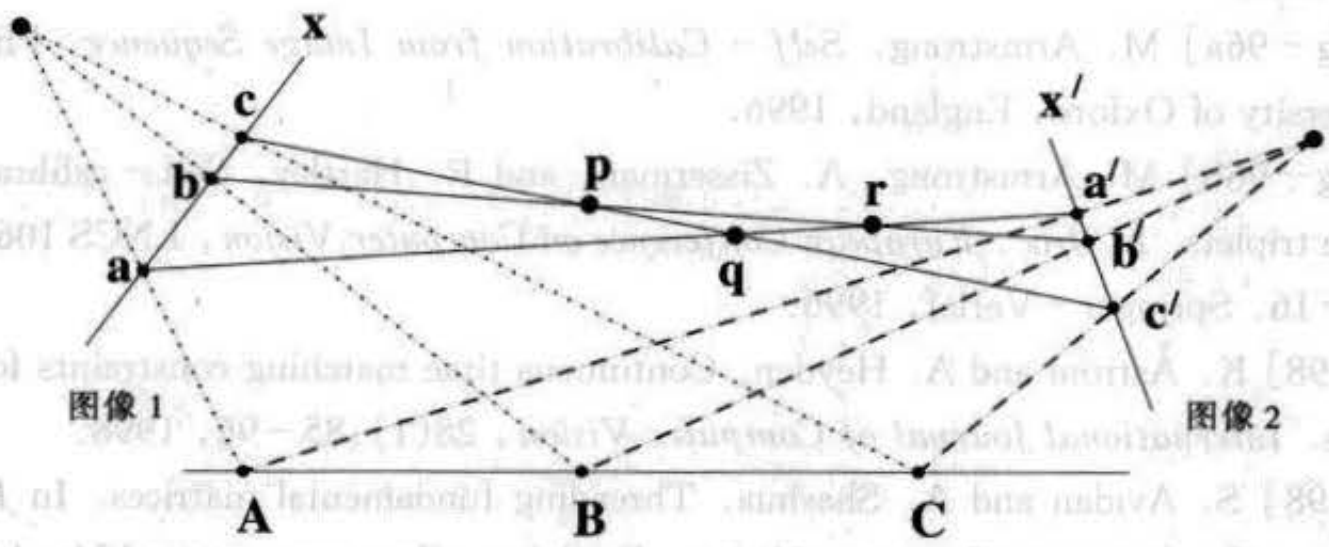



图 A5.4 **直线的射影变换**. 点{**a**,**b**,**c**}与点{**A**,**B**,**C**}通过直线到直线的透视变换相关联. 点{**a′**,**b′**,**c′**}与点{**A**,**B**,**C**}也通过一个透视变换相关联. 但是,点{**a**,**b**,**c**}与点{**a′**,**b′**,**c′**}仅通过一个射影变换相关联;它们不是由一个透视变换相关联,因为连接对应点的直线不再共点. 事实上,每两对直线之间的交点产生三个不同的点{**p**,**q**,**r**}.

# 参考文献


[Aloimonos－90] J. Y. Aloimonos. Perspective approximations. *Image and Vision Computing*, 8(3):177－192, August 1990.

[Armstrong－94] M. Armstrong, A. Zisserman, and P. Beardsley. Euclidean reconstruction from uncalibrated images. In *Proc*, *British Machine Vision Conference*, pages 509－518, 1994.

[Armstrong－96a] M. Armstrong. *Self－Calibration from Image Sequence*. PhD thesis, University of Oxford, England, 1996.

[Armstrong－96b] M. Armstrong, A. Zisserman, and R. Hartley. Self－calibration from image triplets. In *Proc*. *European Conference on Computer Vision*, LNCS 1064/5, pages 3－16. Springer－Verlaf, 1996.

[Åström－98] K. Åström and A. Heyden. Continuous time matching constraints for image-streams. *International Journal of Computer Vision*, 28(1):85－96, 1998.

[Avidan－98] S. Avidan and A. Shashua. Threading fundamental matrices. In *Proc*. *5th European Conference on Computer Vision*, *Freiburg*, *Germany*, pages 124－140, 1998.

[Baillard－99] C. Baillard and A. Zisserman. Automatic reconstruction of piecewise planar models from multiple views. In *Proc*. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 559－565, June 1999.

[Barrett－92] E. B. Barrett, M. H. Brill, N. N. Haag, and P. M. Payton. Linear resection, intersection, and perspective－independent model matching in photogrammetry: Theory. In J. L. Mundy and A. Zisserman, editors, *Geometric invariance in computer vision*. MIT Press, Cambridge, 1992.

[Basri－99] R. Basri and D. Jacobs. Projective alignment with regions. In *Proc*. *7th International Conference on Computer Vision*, *Kerkyra*, *Greece*. pages 1158－1164, 1999.

[Beardsley－92] P. A. Beardsley, D. Sinclair, and A. Zisserman. Ego－motion from six points. Insight meeting, Catholic University Leuven. February 1992.

[Beardsley－94] P. A. Beardsley, A. Zisserman. and D. W. Murray. Navigation using affine structure and motion. In *Proc*. *European Conference on Computer Vision*, LNCS 800/801, pages 85－96. Springer－Verlag, 1994.

[Beardsley－95a] P. A. Beardsley, I. D. Reid, A. Zisserman. and D. W. Murray. Active visual navigation using non－metric structure. In *Proc*. *International Conference on Computer Vision*, pages 58－64, 1995.

[Beardsley－95b] P. A. Beardsley and A. Zisserman. Affine calibration of mobile vehicles. In

*Europe – China workshop on Geometrical Modelling and Invariants for Computer Vision*. pages 214 – 221. Xidan University Press. Xi'an, China. 1995.

[Beardsley – 96] P. A. Beardsley, P. H. S. Torr, and A. Zisserman. 3D model aquisition from extended image sequences. In *Proc. 4th European Conference on Conputer Vision. LNCS* 1065, *Cambridge*, pages 683 – 695, 1996.

[Boehm – 94] W. Boehm and H. Prautzsch. *Geometric Concepts for Geometric Design*. A. K. Peters, 1994.

[Bookstein – 79] F. Bookstein. Fitting conic sections to scattered data. *Computer, Graphics and Image Processing*, 9:56 – 71, 1979.

[Bougnoux-98]S. Bougnoux. From Projective to Euclidean space under any practical situation, a criticism of self-calibration. In *Proc. 6th International Conference on Computer Vision, Bombay, India*, pages 790 – 796, January 1998.

[Brown-71] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37 (8):855 – 866, 1971.

[Buchanan-88] T. Buchanan. The twisted cubic and camera calibration. *Computer Vision. Graphics and Processing*, 42:130 – 132, 1988.

[Buchanan-92] T. Buchanan. Critical sets for 3D reconstruction using lines. In *Proc. European Conference on Computer Vision*, LNCS 588. page 730 – 738. Spinger-Verlag, 1992.

[Canny-86] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679 – 698, 1986.

[Caprile-90] B. Caprile and V. Torre. Using vanishing points for camera calibration. *International Journal of Computer Vision*, 4:127 – 140, 1990

[Carlsson-93] S. Carlsson. Multiple image invariance using the double algebra. In *Applications of Invariance in Computer Vision*, volume SLN Comp. Science vol 825, pages 335 – 350, 1993.

[Carlsson-94] S. Carlsson. Multiple image invariance using the double algebra. In J. Mundy, A. Zisserman, and D. Forsyth, editors, *Applications of Invariance in Computer Vision LNCS* 825. Springer-Verlag, 1994.

[Carlsson-95] S. Carlsson. Duality of reconstruction and positioning from projective views. In *IEEE Workshop on Representation of Visual Scenes*, Boston, 1995.

[Carlsson-98] S. Carlsson and D. Weinshall. Dual computation of projective shape and camera positions from multiple images. *International Journal of Computer Vision*, 1998. in Press.

[Christy-96] S. Christy and R. HoraudEuclidean shape and motion from multiple perspective views by affine iteration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1098 – 1104, November 1996.

[Collins-93] R. T. Collins and J. R. Beveridge. Matching perspective views of coplanar structures using projective unwarping and similarity matching. In *Proc. IEEE. Conference on Computer Vision and Pattern Recognition*, 1993.

[Criminisi-98] A. Criminisi, I. Reid, and A. Zisserman. Duality, rigidity and planar parallax. In *Proc. European Conference on Computer Vision*. page 846 – 861. Springer-Verlag. June 1998.

[Criminisi-99] A. Criminisi, I. Reid, and A. Zisserman. A plane measuring device. *Image and Vision Computing*, 17(8):625 – 634. 1999.

[Cross-98] G. Gross and A. Zisserman. Quadric surface reconstruction from dual-geometry. In *Proc. 6th International Conference on Computer Vision. Bombay. India*. pages 25 – 31, January 1998.

[Cross-99] G. Cross, A. W. Fitzgibbon, and A. Zisserman. Palallax geometry of smooth surfaces in multiple views. In *Proc. 7th International Conference on Computer Vision. Kerkyra, Greece*, pages 323 – 329. September 1999.

[Csurka-98] G. Csurka, D. Demirdjian, A. Ruf, and R. Horand. Closed-form solutions for the euclidean calibration of a stereo rig. In *Proc. 5th European Conference on Computer Vision, Freiburg, Germany*, pages 426 – 442. June 1998.

[deAgapito-98] L. de Agapito, E. Hayman, and I. Reid. Self-calibration of a rotating camera with varying intrinsic parameters. In *Proc. 9th British Machine Vision Conference, Southampton*, 1998.

[deAgapito-99] L. de Agapito, R. I. Hartley, and E. Hayman. Linear self-calibration of a rotating and zooming camera. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 15 – 21, 1999.

[Dementhon-95] D. Dementhon and L. Davis. Model basded pose in 25 lines of code. *International Journal of Computer Vision*, 15(1/2):123 – 141, 1995.

[Devernay-95] F. Devernay and O. D. Faugeras. Automatic calibration and removal of distortion from scenes of stuuctured environments. In *SPIE*, volume 2567, San Diego, CA, July 1995

[Devernay-96] F. Devernay and O. D. Faugeras. From projective to euclidean reconstruction. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 264 – 269, 1996.

[Faugeras-90] O. D. Faugeras and S. J. Maybank. Motion from point matches: Multiplicity of solutions. *International Journal of Computer Vision*, 4:225 – 246, 1990.

[Faugeras-92a] O. D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. European Conference on Computer Vision*, LNCS 588, pages 563 – 578. Spinger-Verlag, 1992.

[Faugeras-92b] O. D. Faugeras, Q. Luong, and S. Maybank. Camera self-calibration: Theory and experiments. In *Proc. European Conference on Computer Vision*, LNCS 588, pages 321 – 334, Springer-Verlag, 1992.

[Faugeras-93] O. D. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIF Press, 1993.

[Faugeras-94] O. D. Faugeras and L. Robert. What can two images tell us about a third one.

In J. O. Eckland, editor, *Proc. 3rd European Conference on Computer Vision, Stockholm*, pages 485 – 492. Springer-Verlag, 1994.

[Faugeras-95a] O. D. Faugeras. Stratification of three-dimensional vision: projevtive, affine, and metric representation. *Journal of the Optical Society of America*, A12:465 – 484, 1995.

[Faugeras-95b] O. D. Faugeras, S. Laveau, L. Robert, G. Csurka, and C. Zeller. 3-D reconstruction of urban scenes from sequences of images. Technical report, INRIA. 1995.

[Faugeras-95c] O. D. Faugeras and B. Mourrain. On the geometry and algebra of point and line correspondences between N images. In *Proc. International Confererce on Computer vision*, pages 951 – 962, 1995.

[Faugeras-97] O. D. Faugeras and T. Papadopoulo. Grassmann-Cayley algebra for modeling systems of cameras and the algebraic equations of the manifold of trifocal tensors. Technical Report 3225, INRIA. Sophia-Antipolis, France. 1997.

[Faugeras-98] O. D. Faugeras, L. Quan, and P. Sturm. Self-calibration of a 1D projective camera and its application to the self-calibration of a 2D projective camera. In *Proc. European Conference on Computer Vision*, pages 36 – 52, 1998.

[Fischler-81] M. A. Fischler and R. C. Bolles. Random smple consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm, Assoc. Comp. Mach.* 24(6):381 – 395, 1981.

[Fitzgibbon-98a] A. W. Fitzgibbon, G. Cross, and A. Zisserman. Automatic 3D model construction for turn-table sequences. In R. Koch and L. Van Gool, editors. *3D Structure from Multiple Images of Large-Scale Environments*, LNCS 1506, pages 155 – 170. Springer – Verlag, June 1998.

[Fitzgibbon-98b] A. W. Fitzgibbon and A. Zisserman Automatic camera recovery for closed or open image sequences. In *Proc. European Conference on Computer Vision*. pages 311 – 326, Spinger-Verlag. June 1998.

[Fitzgibbon-99] A. W. Fitzgibbon, M. Pin, and R. B. Fisher. Direct least-squares fitting of ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):476 – 480, May 1999.

[Giblin – 87] P. Giblin and R. Weiss. Reconstruction of surfaces from profiles. In *Proc. 1st International Conference on Computer Vision, London*, pages 136 – 144, London, 1987.

[Gill – 78] P. E. Gill and W. Murray. Algorithms for the solution of the nonlinear least – squares problem. *SIAM J Num Anal*, 15(5):977 – 992, 1978.

[Golub – 89] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, second edition, 1989.

[Gracie – 68] G. Gracie. Analytical photogrammetry applied to single terrestrial photograph mensuration. In *XIth International Conference of Photogrammetry, Lausanne, Switzerland*, July 1968.

[Gupta – 97] R. Gupta and R. I. Hartley. Linear pushbroom cameras. *IEEE Transactions on*

*Pattern Analysis and Machine Intelligence*, September 1997.

[Haralick-91] R. M. Haralick, C. Lee, K. Ottenberg, and M. Nlle. Analysis and solutions of the three point perspective pose estimation problem. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, page 592-598, 1991.

[Harris-88] C. J. Harris and M. Stephens. A combined corner and edge detector. In *Proc. 4th Alvey Vision Conference, Manchester*, pages 147-151, 1988.

[Hartley-92a] R. I. Hartley. Estimation of relative camera positions for uncalibrated cameras. In *Proc. European Conference on Computer Vision*, LNCS 588, pages 579-587. Springer Verlag, 1992.

[Hartley-92b] R. I. Hartley. Invariants of points seen in multiple images. GE internal report, GE CRD, Schenectady, NY 12301, USA, May 1992.

[Hartley-92c] R. I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1992.

[Hartley-94a] R. I. Hartley. Euclidean reconstruction from uncalibrated views. In J. Mundy, A. Zisserman, and D. Forsyth, editors, *Applications of Invariance in Computer Vision*, LNCS 825, pages 237-256. Springer-Verlag, 1994.

[Hartley-94b] R. I. Hartley. Projective reconstruction and invariants from multiple images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16:1036-1041, October 1994.

[Hartley-94c] R. I. Hartley. Projective reconstruction from line correspondence. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1994.

[Hartley-94d] R. I. Hartley. Self-calibration from multiple views with a rotationg camera In *Proc. European Conference on Computer Vision*, LNCS 800/801, pages 471-478. Springer-Verlag, 1994.

[Hartley-95a] R. I. Hartley. A linear method for reconstruction from lines and points. In *Proc. Internatinal Conference on Computer Vision*, pages 882-887. 1995.

[Hartley-95b] R. I. Hartley. Multilinear relationships between coordinates of corresponding image points and lines. In *Proceedings of the Sophus Lie Symposium. Nordfjordeid, Norway* (not published yet). 1995.

[Hartley-95c] R. I. Hartley and P. Sturm. Triangulation. In *Proc. Conference Computer Analysis of Images and Patterns*. Prague, Czech Republic, 1995.

[Hartley-97a] R. I. Hartley. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 19(6):580-593. October 1997.

[Hartley-97b] R. I. Hartley. Kruppa's equations derived from the fundamental matrix. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 19(2):133-135. 1997.

[Hartley-97c] R. I. Hartley. Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision*, 22(2):125-140. 1997.

[Hartley-97d] R. I. Hartley and T. Saxena. The cubic rational polynomial camera model. In

*Proc. DARPA Image Undrstanding Workshop*, pages 649 – 653, 1997.

[Hartley – 98a] R. I. Hartley. Chirality. *International Journal of Computer Viision*, 26 (1):41 – 61, 1998.

[Hartley – 98b] R. I. Hartley. Computation of the quadrifocal tensor. In *Proc. European Conference on Computer Vision*, LNCS 1406, pages 20 – 35. Springer – Verlag, 1998.

[Hartley – 98c] R. I. Hartley. Dualizing scene reconstruction algorithms. In R. Koch and L. Van Gool, editors, *3D Structure from Multiple Images of Large – Scale Environments*, *LNCS* 1506, pages 14 – 31. Springer – Verlag, June, 1998.

[Hartley – 98d] R. I. Hartley. Minimizing algebraic error in geometric estimation problems. In *Proc. International Conference on Computer Vision*, pages 469 – 475, 1998.

[Hartley – 99] R. I. Hartley, L. de Agapito, E. Hayman, and I. Reid. Camera calibration and the search for infinity. In Proc. *7th International Conference on Computer Vision*, *Kerkyra*, *Greece*, Pages 510 – 517, September 1999.

[Hartley – 2000a] R. I. Hartley. Ambiguous configurations for 3 – view projective reconstruction. In *Proc. European Conference on Computer Vision*, 2000.

[Hartley – 2000b] R. I. Hartley and N. Y. Dano. Reconstruction from six – point sequences. In *Proc. IEEE Conference on Computer Vision and Pattern recognition*, 2000.

[Heyden – 95a] A. Heyden. *Geometry and Algebra of Multiple Projective Transformations*. PhD thesis, Department of Mathematics, Lund University, Sweden, December1995.

[Heyden – 95b] A. Heyden. Reconstruction from image sequences by means of relative depths. In E. Grimson, editor, *Proc. 5th International Conference on Computer Vision*, *Boston*, Cambridge, MA, June 1995.

[Heyden – 97a] A. Heyden and K. Astrom. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1997.

[Heyden – 97b] A. Heyden Projective structure and motion from image sequences using subspace methods. In *Scandinavian Conference on Image Analysis*, *Lappenraanta*, pages 963 – 968, 1997.

[Heyden – 97c] A. Heyden . Reconstruction from multiple images by means of using relative depths. *International Journal of Computer Vision*, 24(2):155 – 161, 1997.

[Heyden – 98] A. Heyden. Algebraic varieties in multiple view geometry. In *Proc. 5th European Conference on Computer Vision*, *Freiburg. Germany*, pages 3 – 19, 1998.

[Hilbert – 56] D. Hilbert and S. Cohn – Vossen. *Geometry and the Imagination*. Chelsea, NY, 1956.

[Horaud – 98] R. Horaud and G. Csurka. Self – calibration and Euclidean reconstruction using motions of a stereo rig. In *Proc. 6th International Conference on Computer Vision*, *Bombay*, *India*, pages 96 – 103, January 1998.

[Horn – 90] B. K. P. Horn. Relative orientation. *International Journal of Computer Vision*, 4:59 – 78, 1990.

[Horn-91] B. K. P. Horn. Relative orientation revisited. *Journal of the Optical Society of America*, 8(10):1630-1638, 1991.

[Huang-89] T. S. Huang and O. D. Faugeras. Some properties of the E-matrix in two-view motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 11:1310-1312, 1989.

[Huber-81] P. J. Huber. *Robust Statistics*. John Wiley and Sons. 1981.

[Irani-98] M. Irani. P. Anandan, and D. Weinshall. From reference frames to reference planes: Multi-view parallax geometry and applications. In *Proc. European Conference on Computer Vision*. 1998.

[Kahl-98a] F. Kahl and A. Heyden. Structure and motion from points lines and conics with affine cameras. In *Proc. European Conference on Computer Vision*. pages 327-341. 1998.

[Kahl-98b] F. Kahl and A. Heyden. Using conic correspondences in two images to estimate epipolar geometry. In *Proc. 6 th International Conference on Computer Vision, Bombay, India*, pages 761-766, 1998.

[Kahl-99] F. Kahl. Critical motions and ambiguous euclidean reconstructions in auto-cabibration. In *Proc. 7th International Conference on Computer Vision, kerkyra, Greece*, pages 469-475, 1999.

[Kanatani-92] K. Kanatani. *Geometric computation for machine vision*. Oxford University Press. Oxford, 1992.

[Kanatani-94] K. Kanatani. Statistical bias of conic fitting and renormalization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(3):320-326, 1994.

[Kanatani-96] K. Kanatani. *Statistical Optimization for Geometric Computation: Theory and Practice*. Elsevier Science, Amsterdam, 1996.

[Klein-39] F. Klein. *Elementary Mathematics from an advanced Standpoint*. Macmillan, New York, 1939.

[Koenderink-84] J. J. Loenderink. What does the occluding contour tell us about solid shape? *Perception*, 13:321-330, 1984.

[Koenderink-90] J. J. Loenderink. *Solid Shape*, MIR Press, 1990.

[Koenderink-91] J. J. Loenderink and A. J. van Doorn, Affine structure from motion. *Jounal of the Optical Society of America*, 8(2):377-385, 1991.

[Krames-42] J. Krames. Über die bei der Hauptaufgabe der Luftphotogrammetrie auftretenden "*gefährlichen*" *Flächen*. *Bildmessung und Luftbildwesen* (*Beilage zur Allg. Vermessungs-Nachr.*), 17, Heft 1/2:1-18, 1942.

[Kriegman-98] D. J. Kriegman and P. Belhumeur. What shadows reveal about object structure. In *Proc. European Conference on Computer Vision*, pages 399-414, 1998.

[Laveau-96a] S. Kaveau. *Géométrie d'un système de N caméras. Théorie, estimation et applications*. PhD thesis, INRIA, 1996.

[Laveau-96b] S. Kaveau and O. D. Faugeras. Oriented projective geometry in computer vi-

sion. In *Proc*. 4 *th European Conference on Computer Vision*. LNCS 1065. Cambridge, pages 147 – 156, Springer – Verlag, 1996. Buxton B. and Cipolla R.

[Liebowitz – 98] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In *Proc*. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 482 – 488. June 1998.

[Liebowitz – 99a] D. Liebowitz, A. Criminisi, and A. Zisserman. Creating architectural models from images. In *Proc*. *EuroGraphics*, *volume* 18, pages 39 – 50, September 1999.

[Liebowitz – 99b] D. Liebowitz and A. Zisserman. Combining scene and auto – calibration constraints. In *Proc* 7 *th International Conference on Computer Vision*. *Kerkyra*, *Greece*. September 1999.

[LonguetHiggins – 81] H. C. Longuet – Higgins. A computer algorithm for reconstructing a scence from two projections. *Nature*, 293:133 – 135. September 1981.

[Luong – 92] Q. Luong. *Matrice Fondementale et Autocalibration en Vision par Ordinateur*. PhD thesis. Université de Paris – Sud, France. 1992.

[Luong – 94] Q. T. Luong and T. Viéville. Canonic representations for the geometrics of multiple projective views. In *Proc*. 3 *th European Conference on Computer Vision*. *Stockholm*. pages 589 – 599. May 1994.

[Luong – 96] Q. T. Luong and T. Viéville. Canonical representations for the geometrics of multiple projective views. *Computer Vision and Image Understanding*, 64(2):193 – 229. September 1996.

[Ma – 99] Y. Ma. S. Soatto, J. Kosecka, and S. Sastry. Euclidean reconstruction and reprojection up to subgroups. In *Proc*. 7 *th International Conference on Computer Vision*. *Kerkyra*, *Greece*, pages 773 – 780. 1999.

[Mathematica – 92] S. Wolfram. *Mathematica A System for Doing Mathematics by Computer second edition*. Addison – Wesley, 1992.

[Maybank – 90] S. J. Maybank. The projective geometry of ambiguous surfaces. *Philosophical Transactions of the Royal Society of London*, *SERIES A*, A 332:1 – 47, 1990.

[Maybank – 93] S. J. Maybank. *Theory of reconstruction from image motion*. *Springer – Verlag*, Berlin, 1993.

[Maybank – 98] S. J. Maybank and A. Shashua. Ambiguity in reconstruction from images of six points. In *Proc*. 6 *th International Conference on Computer Vision*, *Bombay*, *India*, pages 703 – 708, 1998.

[Mohr – 92] R. Mohr. Projective geometry and computer vision. In C. H. Chen, L. F. Pau, and P. S. P. Wang, editors, *Handbook of Pattern Recognition and Computer Vision*. World Scientific, 1992.

[Mohr – 93] R. Mohr, F. Veillon, and L. Quan. Relative 3D reconstruction using multiple uncalibrated images. In *Proc*. *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 543 – 548, 1993.

[Moons – 94] T. Moons, L. Van Gool, M. Van Diest, and E. Pauwels. Affine reconstruction

from perspective images pairs. In J. Mundy, A. Zisserman, and D. Forsyth, editors, *Applications of Invariance in Computer Vision*, LNCS 825. Springer-Verlag. 1994.

[Mundy-92] J. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision*. MIT Press. 1992.

[Newsam-96] G. Newsam, D. Q. Huynh, M. Brooks, and H. P. Pan. Recovering unknown focal lengths in self-calibration: An essentially linear algorithm and degenerate configurations. In *Int. Arch. Photogrammetry & Remote Sensing*, Volume XXXI-B3, pages 575-580, Vienna, 1996.

[Poelman-94] C. Poelman and T. Kanade. A paraperspective factorization method for shape and motion recovery. In *Proc. 3 rd European Conference on Computer Vision. Stockholm, volume* 2, pages 97-108, 1994.

[Pollefeys-96] M. Pollefeys, L. Van Gool, annd A. Oosterlinck. The modulus constraint: a new constraint for self-calibration. In *Proc. International Conference on Pattern Recognition*, pages 31-42, 1996.

[Pollefeys-98] M. Pollefeys, R. Koch, and L. Van Gool. Self calibration and metric reconstrution in spite of varying and unknown internal camera parameters. In *Proc. 6 th International Conference on Computer Vision, Bombay, India*, pages 90-96, 1998.

[Pollefeys-99] M. Pollefeys. *Self-calibration and metric 3D reconstruction from uncalibrated image sequences*. PhD thesis, ESAT-PSI, K.U.Leuven, 1999.

[Ponce-94] J. Ponce, D. H. Marimont, and T. A. Casis. Analytical methods for uncalibrated stereo and motion measurement. In *Proc. 3 rd European Conference on Computer Vision. Stockholm, volume* 1. pages 463-470, 1994.

[Porrill-91] J. Porrill and S. B. Pollard. Curve matching and stereo calibration. *Image and Vision Computing*, 9(1):45-50, 1991.

[Pratt-87] V. Pratt. Direct least-squares fitting of algebraic surfaces. *Computer Graphics*. 21(4):145-151, 1987.

[Press-88] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1988.

[Pritchett-98] P. Pritchett and A. Zisserman. Wide baseline stereo matching. In *Proc. 6 th International Conference on Computer Vision. Bombay, India*, pages 754-760, Jaunary 1998.

[Quan-94] L. Quan. Invariants of 6 points from 3 uncalibrated images. In J. O. Eckland, editor. *Proc. 3rd European Conference on Computer Vision. Stockholm*, pages 459-469, Springer-Verlag, 1994.

[Quan-97] L. Quan. Uncalibrated 1D projective camera and 3D affine reconstruction of lines. In *Proc. IEEE Conderence on Computer Vision and Pattern Recognition*, pages 60-65, 1997.

[Quan-98] L. Quan and Z. Lan. Linear $n \geqslant 4$-point pose determination. In *Proc. 6 th International Conference on Comptter Vision, Bombay, India*, pages 778-783, 1998.

[Reid-96] I, D, Reid and D. W. Murray. Active tracking of foveated feature clusters using affine structure, *International Journal of Computer Vision*, 18(1):41-60, 1996.

[Rinner-72] K. Rinner and R. Burkhardt. Photogrammetrie. In *Handbuch der Vermessungskunde*, volume Band Ⅲ a/3, Jordan, Eggert, Kneissel, Stuttgart: J. B. Metzlersche Verlagsbuchhandlung, 1972.

[Robert-93] L. Robert and O. D. Faugeras. Relative 3D positioning and 3D convex hull computation from a weakly calibrated stereo pair. In *Proc*. 4 *th International Conference on Computer Vision*, *Berlin*, pages 540-544, 1993.

[Rousseeuw-87] P. J. Rousseeuw and Annick M. Leroy. *Robust Regression and Outlier Detection*. Wiley, New York, 1987.

[Sampson-82] P. D. Sampson. Fitting conic sections to 'very scattered' data: An iterative refinement of the Bookstein algorithm. *Computer Vision*, *Graphics*, *and Image Processing*, 18:97-108, 1982.

[Schaffalitzky-99] F. Schaffalitzky and A. Zisserman. Geometric grouping of repeated elements within images, In D. A. Forsyth, j. L. Munsy, V. Di Gesu, and R. Cipolla, editors, *Shape*, *Contour and Grouping in Computer Viaion*, LNCS 1681, pages 165-181. Springer-Verlag, 1999.

[Schaffalitzky-2000] F. Schaffalitzky, A. Zisserman, R. I. Hartley, and P. H. S. Torr. A six point solution for structure and motion. In *Proc*. *European Conference on Computer Vision*, 2000.

[Schmid-97] C. Schmid and A. Zisserman. Automatic line matching across views. In *Proc*, *IEEE Conference on Computer Vision and Pattern Recognition*, pages 666-671, 1997.

[Schmid-98] C. Schmid and A. Zisserman. The geometry and matching of curves in multiple views. In *Proc*. *European Conference on Computer Vision*, pages 394-409, Springer-Verlag, June 1998.

[Semple-79] J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1979.

[Shapiro-95] L. S. Shapiro, A. Zisserman, and M. Brady. 3D motion recovery via affine epipolar geometry. *International Journal of Computer Vision*, 16(2):147-182, 1995.

[Shashua-94] A. Shashua. Trilinearity in visual recognition by alignment. In *Proc*. 3 *rd European Conference on Computer Vision*, *Stockholm*, *volume* 1. pages 479-484, May 1994.

[Shashua-95a] A. Shashua. Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):779-789, August 1995.

[Shashua-95b] A. Shashua and M. Werman. On the trilinear tensor of three perspective views and tis underlying geometry. In *Proc*. 5 *th International Conference on Computer Vision*, *Boston*, 1995.

[Shashua-96] A. Shashua and S. J. Maybank. Degenerate N-point configurations of three views: Do critical surfaces exist? Technical Report TR 96-19, Hebrew University,

Computer Science, November 1996.

[Shashua-97] A. Shashua and S. Toelg, The quadric reference surface: Theory and applications. *International Journal of Computer Vision*, 23(2): 185-198, 1997.

[Shimshoni-99] I. Shimshoni, R. Basri, and E. Rivlin. A geometric interpretation of weak-perspective motion. Technical report. Technion, 1999.

[Sinclair-92] D. A. Sinclair. *Experiments in Motion and Correspondence*. PhD thesis, Unniversity of Oxford, 1992.

[Slama-80] C. Slama. *Manual of Photogrammetry*. American Society of Photogrammetry, Falls Church, VA, USA, 4 th edition, 1980.

[Spetsakis-91] M. E. Spetsakis and J. Aloimonos. A multi-frame approach to visual motion perception. *International Journal of Computer Vision*, 16(3):245-255, 1991.

[Springer-64] C. E. Springer. *Geometry and Analysis of Projective Spaces*. Freeman, 1964.

[Stein-99] G. Stein and A. Shashua. On degeneracy of linear reconstruction from three views: Linear line complex and applications. *IEEE Transactions on Patteen Analysis and Machine Intelligence*, 21(3):244-251, 1999.

[Stolfi-91] J. Stolfi. *Oriented Projective Geometry*. Academic Press, 1991.

[Sturum-96] P. Sturm and W. Triggs. A factorization based algorithm for multi-image projective structure and motion. In *Proc. European Conference on Computer Vision*, pages 709-720, 1996.

[Sturum-97a] P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction. In *Proc. IEEE Conference on computer Vision and Pattern Recongnition, Puerto Rico*, pages 1100-1105, June 1997.

[Sturum-97b] P. Strum. *Vision 3D non calibrée: Contributions à la reconstruction projective ét etude des mouvements critiques pour l' auto calibrage*. PhD thesis, INRIA Rh one-Alpes, 1997.

[Sturum-99a] P. Sturm and S. Maybank. On plane based camera calibration: A general algorithm, singularities, application. In *pro. IEEE conference on Computer Vision and Pattern Recognitions*, pages 432-437, June 1999.

[Sturum-99b] P. Strum. Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. In *Proc. 10th British machine Vision Conference, Nottingham*, pages 63-72, 1999.

[Sutherland-63] I. E. Sutherland. Sketchpad: Anman-machine graphical communications system. Technical Report 296. MIT Lincoln Laboratories, 1963. Also published by Garland Publishing, New York, 1980.

[Szeliski-96] R. Szeliski and S. B. Kang. Shape ambiguties in structure from motion. In B. Buxton and Cipolla R., editors, *Proc. 4th European Conference on Computer Vision, LNCS* 1064, *Cambridge* pages 709-721. Springer-Verlag, 1996.

[Taubin-91] G. Taubin. Estimation of planar curves, surface, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. PA-

MI, 13(11):1115-1138, 1991.

[Tomasi-92] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization approach. *International Journal of Compumter Vision*, 9 (2):137-154, November 1992.

[Torr-93] P.H.S. Torr and D.W. Murray. Outlier detection and motion segmentation. In *Proc SPIE Sensor Fusion VI*, pages 432-443. Boston, September 1993.

[Torr-95a] P.H.S. Torr. *Motion segmentation and outlier detection*. PhD thesis, Dept. of Engineering Science, University of Oxford. 1995.

[Torr-95b] P.H.S. Torr, A. Zisserman, and D.W. Murray. Motion clustering using the trilinear constraint over three view. In R. Mohr and C. Wu, editors, *Europe-China Workshop on Geometrical Modelling and Invariants for Computer Vision*. pages 118-125. Xidan University Press, 1995.

[Torr-97] P.H.S. Torr and Zisserman. Robust parameterization and computation of the trifocal tensor. *Image and Vision Computing*. 15:591-605. 1997.

[Torr-98] P.H.S. Torr and A. Zisserman. Robust computation and parameterization of multiple view relation. In *proc. 6th. International Conference on Computer Vision. Bombay, India*, pages 727-732. January 1998.

[Torr-99] P.H.S. Torr, A.W. Fitzgibbon, and Zisserman. The problem of degeneracy in strcture and motion recovery from uncablibtated image sequences. *International Journal of Computer Vision*. 32(1):27-44. August 1999.

[Triggs-95] W. Triggs. The geometry of projective reconstruction i: Matching constraints and the joint image. In *Proc. International Conference on Computer Vision*, pages 338-343, 1995.

[Triggs-96] W. Triggs. Factorization methods for projective structure and motion. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 845-851, 1996.

[Triggs-97] W. Triggs. Auto-calibration and the absolute quadric. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 609-614, 1997.

[Triggs-98] W. Triggs. Autocalition from planar scenes. In *Proc. 5th European Conference on Computer Vision, Freiburg, Germany*. 1998.

[Triggs-99a] W. Triggs. Camera pose and calibration from 4 or 5 known 3D points. In *Proc. International Conference on Computer Vision*, pages 278-284, 1999.

[Triggs-99b] W. Triggs. Differential matching constraints. In *Proc. International Conference on Computer Vision*, pages 370-376, 1999.

[Triggs-2000] W. Triggs, P.F. McLauchlan, R.I. Hartley, and A. Fitzgibbon. Bundle adjustment for structure from motion. In *Vision Algorithms: Theory and Practice*. Springer-Verlag, 2000.

[Tsai-84] R.Y. Tsai and T.S. Huang. The perspective view of three points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:13-27, 1984.

[VanGool-98] L. Van Gool, M. Proesmans, and A. Zisserman. Planar homologies as a basis

for grouping and recognition. *Image and Vision Computing*, 16:21 – 26, January 1998.

[Viéville – 93] T. Viéville and Q. Luong. Motion of points and lines in the uncalibrated case. Technical Report 2054, I.N.R.I.A., 1993.

[Viéville – 95] T. Viéville and D. Lingrand. Using singular displacements for uncalibrated monocular vision systems. Technical Report 2678. I.N.R.I.A., 1995.

[VonSanden – 08] H. von Sanden. *Die Bestimmung der Kernpunkte in der Photogrammetrie*. PhD thesis, Univ. Göttingen, December 1908.

[Weinshall – 95] D. Weinshall, M. Werman, and A. Shashua. Shape descriptors: Bilinear, trilinear and quadrilinear relations for multi – point geometry and linear projective reconstruction algorithms. In *IEEE Workshop on Representation of Visual Scenes*, *Boston*, pages 58 – 65, 1995.

[Weng – 88] J. Weng, N. Ahuja, and T. S. Huang. Closed – form solution and maximum likelihood: A robust approach to motion and structure estimation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1988.

[Weng – 89] J. Weng, T. S. Huang, and N. Ahuja. Motion and structure from two perspective views: algorithms, error analysis and error estimation. *IEEE Transactions on Pattern analysis and Machine Intelligence*, 11(5):451 – 476. 1989.

[Wolfe – 91] W. J. Wolfe, D. Mathis. C. Weber Sklair, and M. Magee. The perspective view of three points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 13 (1):66 – 73. Junuary 1991.

[Xu – 96] G. Xu and Z. Zhang. *Epipolar Geometry in Stereo, Motion and Object Recognition*. Kluwer Acadenic Publishers. 1996.

[Zeller – 96] C. Zeller. *Projectibe, Affine and Euclidean Calibration in Computer Vision and the Application of Three Dimensional Perception*. PhD thesis. Robot Vis Group. INRIA Sophia – Antipolis, 1996.

[Zhang – 95] Z. Zhang, R. Deriche. O. D. Faugeras, and Q. Luoug. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78:87 – 119. 1995.

[Zhang – 98] Z. Zhang. Determining the epipolar geometry and tis uncertainty a review. *International Journal of Computer Vision*. 27(2): 161 – 195. March 1998.

[Zhang – 99] Z. Zhang. A flexible new technique for camera calibration. In *Proc. 7th International Conference on Computer Vision*, *Kerkyra*, *Greece*, September 1999.

[Zisserman – 92] A. Zisserman. Notes on geometric invariance in vision. Tutorial, British Machine Vision Conference, 1992.

[Zisserman – 94] A. Zisserman and S. Maybank. A case aganist epipolar geometry. In J. Mundy, A. Zisserman, and D. Forsyth, editors. *Applications of Invariance in Computer Vision LNCS* 825. Springer – Verlag, 1994.

[Zisserman – 95a] A. Zisserman, P. Beardsley, and I. Reid. Metric calibration of a stereo rig. In *IEEE Workshop on Represention of Visual Scenes*, *Boston*, pages 93 – 100, 1995.

[Zisserman-95b] A. Zisserman, J. Mundy, D. Forsyth, J. Liu, N. Pillow, C. Rothwell, and S. Utcke. Class-based grouping in perspective images. In *Proc. International Conference on Computer Vision*, 1995.

[Zisserman-96] A. Zisserman. A users guide to the trifocal tensor. Dept. of Engineering Science, University of Oxford, 1996.

[Zisserman-98] A. Zisserman, D. Liebowitz, and M. Armstrong. Resolving ambiguities in auto-calibration. *Philosophical Transactions of the Royal Society of London*, *SERIES A*, 356(1740):1193-1211, 1998.

# 索 引


$IP^1$ 的拓扑 (topology of $IP^1$)

**B**

伴随矩阵 (adjoint matrix) 7

保持凸包 (preserve the convex hull)

定义 364

保线性 (collineation)——也见射影变换 (projective transformation) 8,34,429

本质矩阵 (essential matrix) 173-176,178,187-188,203,252

计算 203

定义 173

求摄像机 174

四个解 175

性质 174

编程提示 (programming hint) 84,283

变换(transformation)

二次曲线的 10

直线的 10

变换不变量 (transformation invariance) 83

DLT 的 125

变换规则 (transformation rule) 10-11,24-25,32,148,279,289,337,400-401

标定(calibration)

内参数 119

它给出了什么? 145

标定二次曲线 (calibrating conic) 100,134,148

正交性 149

标定矩阵 (calibration matrix) 102,108-110,111-112,115,130,141,147,147-149,331,334,347,323,325,326,328,173,180,186-187,151,153-155

标定约束 (calibration constraints)

已知内参数 186

所有图像采用相同摄像机 187

不变量 (invariants) 13-14,16-17,32,44,119,181,347-348,377,427

变换的 12-13,15

不变量数 (the number of invariants) 16

不动点计算 (computing the fix points) 346

C

参数化（parametrization） 43,63,69－70,71,72,82,87,92,109,109,123,127,195－196,197,203,221－222,278,283,288,331,262,405,431
参数空间（parameter space） 71,87,92－95,336,351
Carlsson－Weinshall 映射（Carlsson－Weinshall map） 389,391
　定义 389
叉乘（cross product） 406
长度比（length rations） 13,16,26,27,184
场景约束（scene constraints）
　距离比
　　直线上 183,22
　平行线 183
　场景正交性 186
超参数化（over－parametrization） 38,69－70,92－95,120,309,416,419
Chasles 定理（Chasles theorem） 380
Cholesky 分解（Cholesky factorization） 146,186－187,190,331,408
重采样（resampling） 72,211,213－214
重构(reconstruction)
　仿射——见仿射重构
　仿射和度量 239
　仿射
　　由仿射摄像机 185
　　由场景约束 183
　　由平移 183
　欧氏 181,373
　度量——也见自标定 158,182,185－189,312,318,323－324,325,327,331－332,335,337,340,339,346－348,349,351－352,368,377
　射影 19,159,181－182,184－185,190,216,217,223,228,239,258,274,282,309,314,319,323,323－325,326－327,332,337,340,344,347,348－349,354－355,357,358－359,361,363－364,367－369,372－374,376－377,383,383－384,418,325
　准仿射 336,364,367－369,370,372－375,377
重投影误差（reprojection error）
　两幅图像 58
　几何解释 59
初始化（initialization） 69,71－72,277,309
　摄像机参数估计情形 127
纯平移运动（pure translational motion） 202

**D**

代价函数（cost function） 1,52,55,56－58,61,63－65,69－72,77－79,91,83,87,128,130,191,195,197－198,198,200,202,214－215,217,219,220,220－225,247－248,281,282－283,287,310,326,329－331,343,349,351,363,417－418,422

代数距离（algebraic distance） 57－58,59－60,127,191,214

与几何距离的比较 58

代数误差（algebraic error）——见误差代数（error algebraic）

代数最小化（algebraic minimization） 67,199,277

带有视差的平面重构（plane plus parallax reonstruction）

自变量数 302

单应（homography）

2D 54,56,61－62,66,68,72,79,82－84,86,93,95,122－123,124,128,162,182,195,418,422,424

3D 301,293

由平面诱导 261

单应（homographies）

与极几何兼容 229

单应估计（homography estimation）

近似解 53

二次曲线模拟 60

退化配置 55

DLT 54

两幅图像中均带有误差 87,97

一幅图像中带有误差 57,63,70－71,86,95

世界点带有误差 125

实验评价 127

由噪声点 233

函数描述 70

黄金标准算法 52

非齐次解 55

迭代方法 72

线性代价函数 63

超定解 86

参数化 69

对称转移误差 58,70

利用 RANSAC 79

导数矩阵（derivative matrix）

计算 95
等距变换 (isometries)
DIAC——也见绝对二次曲线的对偶图像点 (point)
由三个平面定义 36
点的深度 (depth of points) 108
点方程 (point equations)
推荐使用的方法 307
点转移 (point transfer)
利用基本矩阵 270
利用三焦张量 272
迭代方法 (iterative method) 72,326,344
迭代算法 (iterative algorithm) 378
迭代最小化 (iterative minimization) 69-71,77-79,83,123,126-128,131,132,288,326,332
DLT——也见单应计算、变换不变性、三角形法 54,57,59-60,65-69,72-73,83,99,121-123,124-126,127,132-133,188,218,301,360-361,409,414
DLT 算法 (DLT algorithm)
非不变性 66
利用直线对应 122
度量标定 (metric-calibration)
双眼装置的 348
度量重构 (metric reconstruction) 473
通过无穷远单应 336
度量矫正 (metric rectification) 25-27,339
度量结构 (metric structure) 13
度量性质 (metric properties) 1-2,24-25,46,47-49,181,323-324
重构的 48
从图像恢复 25
对称部分 (symmetric part) 168-170,176,203,347,284
对称对极点距离 (symmetric epipolar distance) 197
对称矩阵 (symmetric matrices) 412
对极点 (epipoles) 159-161,163-164,168-170,176,178,180,181,192,197-198,213,221-222,223,225,231,244-246,257,260,265,264,270-272,273,276,279-281,287,292,333,346,347-348
作为参数 194
作为张量 292
仿射摄像机情形 244
从三焦张量中恢复 280

恢复 280
对极点距离（epipolar distance） 271
对极线（epipolar line） 160,161,162,164－165,167,169,220－222,227,240－242,244－245,256,265－266,270－271,277,333
仿射摄像机 243
对极线包络（epipolar envelope） 208－209,211
对极线单应（epipolar line homography） 165
对极线对应（epipolar line correspondence） 168
对极转移（epipolar tranfer）
退化 271
对偶（dual）
绝对二次曲线的——见绝对对偶二次曲面
虚圆点的 23
对偶重构算法（dual reconstruction algorithm） 355,356－357
理由 357
对偶二次曲面（dual quadric） 40,408
对偶二次曲线（dual conic） 6,7,11,24－26,32,337
对偶绝对二次曲面（dual absolute quadric）
计算
零扭曲情形 329
局限性 330
带有约束参数的 329
与自标定的等价性 327
线性解 329
具体的线性约束 328
对偶性（duality） 4－6,7,34,143,290,291,355,356,358,362,389,390－391,393－394
对偶虚圆点（dual circular points） 23
对偶原理（duality principle） 391
对射（correlation）——定义 28
对应（correspondence）
曲线与曲面间的 204
直线间的 204
多重多义性（ambiguities） 335,340,339
自标定 329
多视图张量（multiview tensor）
仿射 312
多义性（ambiguity）
仿射 312
浅浮雕 253－354,256,431

例子 386
2D 摄像机情形 380
3D 摄像机情形 381
Necker 反转 253
自标定 335,339
序列情形 339
射影 158,274
单视图 291
两视图 393

**E**

二次曲面 (quadric) 39－43,49－50,59,100,102,103,134,137－140,156,181,203－206,214,240,325－327,329,335,352,384－391,393－394,395－398,399
对偶——见对偶二次曲面
规则的 386,391,398
定义 41
分类 40
二次曲线 (conic) 6
仿射分类 30
分类 29
由 5 点定义 6
退化 7
对偶——也见对偶二次曲线 7
射影归一化形式 28
切线 7
变换 10

**F**

反向投影 (back－projection)
二次曲线的 137
直线的 135
点的 107
方差 (variance) 63,78,86－90,90－91,93－95,96,98,128－129,208－209,402,419,425
仿射变换 (affine transformation) 13
仿射重构 (affine reconstruction) 181,182－185,187－188,189－190,239,249,252－253,256－257,309,310－311,312,320,323,335,336－337,340,345,346－347,349－350,351,364,368,372,375－376
仿射多视图张量 (affine multiple view tensors)

计算 312
仿射基本矩阵（affine fundamental matrix） 206,214,244,245－247,249－250,256－257,292,312
代数推导 247
推导 245
几何推导 245
几何解释 246
黄金标准算法 247
线性算法 246
最小配置 249
性质 245
奇异约束 246
仿射矩阵（affine matrix） 13,31,60,247
仿射摄像机（affine camera） 111,113－115,116－117,119,126－127,185,243,245,251,292,302,308,310－311,368
自变量数 302
分解 113
定义 111
应用误差 113
估计 126
分层 115
更多性质 117
仿射性质（affine properties） 1,16,20－21,25,32,46－47
重构的 47
从图像恢复 20
仿射影像模型（affine imaging model） 113
仿射影像条件（affine imaging conditions） 245
非各向同性缩放（non－isotropic scaling）——见缩放——非各向同性
非零扭曲（non－zero skew） 109
非齐次方法（inhomogeneous method） 218－219
分层（stratification） 27,189－190
分层重构（stratified reconstruction）
仿射 182
度量 185－186
分层合并（hierarchical merging） 318
G
高斯（Gaussian）
分布 64－65,89,90,89－94,207－208,402,417

各向同性 87,89
误差 62,64,70,73,78,99,219
噪声 68,86-87,89,95,200,208-209,220,224,283-284,310
刚体运动的转动分解 (screw decomposition of a rigid motion) 44,45
Givens 旋转 (Givens rotation) 405
共点直线 (concurrent lines) 18
共轭 (conjugacy) 29,44,45,48,148,154,347-348,427
共轭配置 (conjugate configurations)——定义 384
共面约束 (coplanarity constraint) 188,300
关联关系 (incidence relations) 34,37-38,259-261,262-265,268-269,269-370,276,300
直线情形 259
关系 (relation)
双重线性 297
四重线性 298
三重线性 278,281,293-295,296
归一化 (normalization) 54-55,57,60,67-69,82-83,95,121-122,123,126,175,193,195,218,277-278,279,281,315,363
摄像机参数估计情形 122
三焦点张量情形 279
归一化变换 (normalizing transformation) 65,67,68,193,279
归一化坐标 (normalized coordinates) 173,315
已标定摄像机情形 173

**H**

行和列矢量 (row and column vectors) 3
行矢量 (row vector) 3,105,369,404
Householder 矩阵 (Householder matrix)
黄金标准算法 (Gold Standard algorithm)——也见基本矩阵一计算,仿射基本矩阵,三焦点张量一计算,单应估计 53,72-73,84,123-124,127,195-196,197-199,202,209,216,243,246,249,282,283-284
仿射基本矩阵情形 247
基本矩阵情形 195
三焦点张量情形 281
恢复摄像机矩阵 (retrieving the camera matrices) 266

**I**

IAC—见绝对二次曲线的图像

J

Jacobi 方法（Jacobi's method） 406
基本矩阵（fundamental matrix）
仿射的——见仿射基本矩阵
代数推导 163
算法结果和建议 283
反对称部分 169
以两个对极点为参数 197
计算 178
由七点计算 192
与二次曲面的关系 388
由两个平面计算 237
黄金标准算法 195
最大似然 195
点对应的条件 163
协方差 424
定义 164
退化
无平移 204
点在平面上 394
点在规则二次曲面上 205
由四个平面推导 299
对极参数化 196
由标准摄像机提取 172
平面运动情形
形式 168
纯平面运动情形 170
平移情形
计算 166
形式 166
几何推导 162
迭代估计 195
超参数化 196
参数化 195
给定 F 的射影多义性 171
射影不变性 171
性质 164
简化的 308,359－363
计算 359

估计的 Sampson 方法 197,283
奇异约束 192
对称部分 169
基本矩阵的提取 (extracting the fundamental matrices) 266
极点 - 极线关系 (pole - polar relationship) 288
几何表示 (geometric representation) 46,47 - 49,165,168,169 - 170,327
几何解释 (geometric interpretation) 32,40,61,156,326,327,337
估计的 62
几何距离 (geometric distance) 57
几何误差 (geometric error)——也见误差几何 (error geometric)
摄像机参数估计 126
坐标变换的不变性 37
几何最小化 (geometric minimization) 72,128
简化的测量矩阵 (reduced measurement matrix)
简化的基本矩阵 (reduced fundamental matrix)——见基本矩阵——简化的
简化的三焦点张量 (reduced trifocal tensor)——见三焦点张量——简化的
简化的摄像机矩阵 (reduced camera matrix) 144
定义 354
建议 (recommendations) 199
降秩方程组 (deficient - rank systems) 410
交 (intersection)
直线的 3,170,183,265,290
平行线的 4
交比 (cross ratio) 15,16 - 18,20,21,32,379,428 - 429
焦距 (focal length) 10,100,104,111 - 112,113,125,129 - 130,131,133,140,149,157,180,328 - 332,334,342,345,351,352 - 353
角 (angle)
图像上两条直线之间的 153
结构 (structure)
仿射 253,341
度量 155,188,325,424
射影 373,300,302 - 303,347,379,398 - 399
结果 (results) 199
径向失真 (radial distortion) 128 - 132
估计 131
距离比 (distance ratios)
直线上 21 - 22,184
绝对对偶二次曲面 (absolute dual quadric) 49 - 50,323,326,327,331,332,336,338
提取矫正变换 327

性质 326
绝对二次曲线 (absolute conic) 1,45,46,47-50,134,145-146,154,185-188,189-190,240,323,324-325,328,333,337,344,346,352
与正交性的关系 48
绝对二次曲线的对偶图像 (dual image of the absolute conic) 327-328,329,338
定义 146
绝对二次曲线的图像 (image of the absolute conic) 100,146,148,153,155,185-188,332,335,336,338,339-340,346
和图像正交性 148
与标定的关系 145

**K**

卡方分布 ($\chi^2$ - distribution) 75-76,77,208,403
可实现点 (realizable points)
定义 355
Klein
Erlangen Program 8
Klein 二次曲面 (Klein quadric) 40
Klein 瓶颈 (Klein bottle) 18
Kruppa 方程 (Kruppa equations) 323,333-335,336,341,351,352
与无穷远单应的关系 341
捆集调整 (bundle adjustment) 281,309-310,318-319,326,331,351,425
增量 319

**L**

Levenberg-Marquardt 127,195,282,417-419,421-422
协方差 419
实现 417
判断 417
稀疏 418,421
理想点 (ideal points) 4-5
两平面 (two planes) 4,8-9,15,36-38,39-40,41-42,46,48,49-50,59-60,140,156,203,205,224,234,237,240,386,388,427,428
两视图与三焦点张量间的关系 (relations between two views and the trifocal tensor) 296
列矢量 (column vector) 3,93,102,105,404,406-407,418
临界集 (critical set) 382
关于 2D 摄像机 380
定义 380
临界配置 (critical configurations) 122,378-379,382,382,387,395,398-399

例子 386
临界曲面 (critical surface)
三视图情形 395
例子 397
零空间 (null - space) 35 - 39,42,54,94 - 95,104 - 105,119,121,136 - 137,156,164,175,191 - 192,193 - 195,204,207,214,246,260,266,287,316,333,339,382,407,444
零扭曲 (zero skew) 109,130,148,154,157,186,329 - 330,334,337 - 339,342,350,351
零矢量 (null - vector) 245
鲁棒代价函数 (robust cost function) 78
鲁棒方法 (robust methods) 1,52,71,73 - 76,77 - 79,81,82 - 83,193,200,249,277,282,320,363
鲁棒算法 (robust algorithms) 78
轮廓生成元 (contour generator)——定义 137

**M**

Mahalanobis
距离 64,70,90,92,124,194,282,402 - 403,416,418
范数 61,83,93
MLE——见最大似然估计 87
模约束 (modulus constraint) 335 - 336,352

**N**

Necker 反转多义性 (Necker reversal ambiguity) 353
内参数 (internal parameters) 100,109,111,121,128,134,141,146,158 - 159,166 - 167,168,179,183,186 - 187,189,239,259,290,321,323 - 324,326,327 - 328,329 - 330,331,333 - 345,347,349,350 - 353
摄像机的
求解它们 109
牛顿迭代 (Newton iteration) 69,72,416 - 418
加权 417
扭转对 (twisted parir) 175,188

**O**

欧氏变换 (Euclidean transformation)
不动点 348
欧氏变换的不动点 (fixed points of a Euclidean transformation) 348
欧氏矩阵 (Euclidean matrix) 31
欧氏与仿射解释 (Euclidean and affine interpretation) 110

**P**

平面 (plane)

由三点定义 35

齐次矢量表示 34

平面计算 (computing a plane)

从点与直线的对应 233

从三点的对应 231

平面上的点 (points on a plane) 205

平面射影变换 (planar projective transformation)

定义 8

平面调和透射 (planar harmonic homology) 33,429

平面透射 (planar homology)

参数化 428

平面诱导的视差 (plane induced parallax) 236

平面运动 (planar motion) 44,51,166,168,169 - 170,176,203,241,276,232,345 - 247,349,350,352

不动图像点 345

平行投影 (parallel projection) 100 - 101,115,116 - 117,137 - 138,243,244,253,257

平行透视 (para - perspective) 117,120

平行线 (parallel lines) 1 - 2,3 - 5,9,13 - 14,16,20 - 21,34,50,117,119,149 - 153,154,166 - 167,183 - 184,203,226,350

Plücker

坐标 38,50

直线坐标 38,136 - 137,155

直线表示 136

矩阵 38,136

**Q**

齐次表示 (homogeneous representation) 4,6,8,34,105,108,119,220,315,364,367

直线的 2

点的 3

齐次矢量 (homogeneous vector) 3 - 5,22,34,53,54,95,101,167,221,315

齐次缩放 (homogeneous scaling) 72,154

奇异值 (singular values)

与特征值的关系 408

奇异值分解 (Singular Value Decomposition) 13,25,38,54,108,127,174 - 175,191 - 193,203,214,225,243,305,311 - 312,314,332,333,408 - 412,413 - 415

实现 312,408
奇异性约束 (singularity constraint) 192,194,198,246
浅浮雕多义性 (bas-relief ambiguity) 253
求 $\pi_\infty$ 的其他方法(other methods of finding $\pi_\infty$) 336
QR 分解 (QR decomposition) 127,404-405

**R**

RANSAC 52,73-76,77-81,80-83,132,200-202,249,277,284-286
自适应 77
应用领域 81
距离阈值 75
2D 单应情形 81
样本数 75
运行细节 82
一致集的大小 77
弱透视 (weak perspective) 115-117,254

**S**

Sampson 近似 (Sampson approximation) 61,71-72,197,202,214,219,246,277,282,286
Sampson 距离 (Sampson distance) 191,197
Sampson 误差 (Sampson error) 61
三次绕线 (twisted cubic) 43-44,122,169-170,290,346,380,382-383,289,391-393,394,398-399
性质 43
三重线性 (trilinearities)——也见关系——三重线性 268
三焦点张量 (trifocal tensor) 52,68,76,79,109,258-259,260-261,262,264,266-269,270,272-287,288-289,290-291,293-299,300-302,303,307,312,214,316,318,321,346-347,359,360,360-363,378,400-401,414,424
仿射 288,297
代数性质 266
由摄像机矩阵计算 269
计算
由六点 278
黄金标准算法 281
迭代 280
重投影误差 284
结果和建议 283
Sampson 距离 197,283

特定的直线 287
利用代数最小化 280
约束 278
定义 260,269
由四个平面推导 300
几何有效 278
直线关系 295
参数化 282
图片 401
点和直线的关系 262
点关系 293
简化的 307,362
变换规则 401
三角形法（triangulation） 158,178－179,182,190,196,202,216－217,219,220,222,223－224,227,233,243,271,273,282,309,318
DLT 方法 218
在真实图像上评估 224
仿射摄像机情形 250
非齐次方法 217
局部最小化 223
最小化代价函数 221
问题公式化 220
摄像机（camera）
仿射——见仿射摄像机
在无穷远处 117
基本针孔模型 101
已标定的 132,145,148
CCD 103
有限 115
有限射影 103
一般射影 104
线阵 119
射影 101,103－105,107,110,117－119,126－127,132,145,149,171,257,323,360,366
作用于点 107
推仓式 117
摄像机参数估计多义性 392
摄像机（cameras）
被看作点 378

摄像机标定（camera calibration） 100,102,126,128,134,145,153,155,178,182,189,202,240,253,323－324,335,341,361,414
摄像机参数（camera parameters） 97,100,103,109,114,123,126,127,128,133,149,283,309,313,323,331,418,424－425
摄像机参数计算（resection,or camera resectioning）
由5点 391
协方差 128
退化配置 122
最小解 121
超定解 122
摄像机的外部校准（exterior orientation of camera） 128
摄像机方位（camera orientation） 102,142,146
求解 109
摄像机构造（camera anatomy） 104
摄像机矩阵（camera matrix）
标准形式 171
由F计算 178
分解 108
在射影空间中 110
简化的——见简化摄像机矩阵
摄像机平移（camera translation） 103,166,168
摄像机旋转（camera rotation） 103,134,141,144,151,168,347
摄像机中心（camera center） 104
求解 108
移动 144
射影变换（projective transformation）
分解 15
定义 8
射影变换群（groups of projectivities） 12
射影重构（projective reconstuction）
对偶算法 355
由简化的摄像机矩阵 360
由n视图中的七点 362
由六点 360
由n视图中的六点 361
已知一个平面 316
射影多义性（projective ambiguity） 179
射影分解（projective factorization）
归一化的图像坐标 315

归一化加权 315
什么被最小化 314
射影平面模型 (projective plane model)4,18
射影摄像机 (projective camera)
作用于二次曲线 137
作用于直线 135
作用于平面 135
定义 104
射影深度 (projective depth) 236,290,314-316
射影实现 (projective realization)——见实现—射影
深度选择 (choosing the depths) 314
生成子空间表示 (span representation) 38,225
失真函数 (distortion function) 130
失真函数和中心的选择 (choice of the distortion function and centre) 130
失真矫正 (distortion correction) 130
实施和运行细节 (implementation and run details) 286
实现 (realization) 354,354-355,355,375,386
定义 354
简化的 354-355
强 270-272
等价
定义 383
非等价 384
反向 371,372-373
视差 (parallax)
矢量几何 (vector geometry) 36
矢量空间范数 (vector space norms) 412
术语 (terminology) 1,118,160,314
双曲面 (hyperboloid)
双眼对应 (stereo correspondence) 159,240
双眼矫正 (stereo rectification)
仿射 214
算法概要 214
将对极点映射到无穷远 212
匹配变换 213
双眼装置 (stereo-rig)
度量标定 349
四焦点张量 (quadrifocal tensor)
从四个平面中推导

缩放（scaling）
各向同性 12－13,31,44,67－68,114
非各向同 13,68
接近无穷远的点 68
缩放正投影（scaled orthographic projection） 68－69
SVD——见奇异值分解
SVD 的计算复杂性（computational complexity of the svd） 408

T
特征值提取（eigenvalue extraction）
Jacobi 方法 406
条件数（condition number） 68,312
同视点（horopter） 43,168－169,346,392－393
投影（projection）
正交 114
平行 113
缩放的正交 114,254
弱透视 114
图像等价配置（image－equivalent configurations）
定义 380
图像平面（image plane）
移动 140
图像中的正交性（orthogonality in the image） 148,153
图像坐标变换的不变性（invariance to image coordinate transformations）
退化（degeneracy）
三角形法 225
退化单应（degenerate homographies） 236
退化配置（degenerate configurations）——也见摄像机参数估计、转移、2D 单应估计
2D 单应

W
无穷单应（infinite homography）
定义 239
无穷单应约束（infinite homography constraint）
与 Kruppa 方程的关系 341
无穷远平面（plane at infinity） 45,47,48,49－50,106－107
无穷远直线（line at infinity） 2,5,15－16,20－21,31,46－47,55,106,366,430
误差（error）
代数的 56－57,58,61－62,66,121,125－126,128,194－195,198,199,280－281,283

-284,330
摄像机参数估计 128
几何解释 124
几何的 61,62,64,65,67,72,84,87,123-125,126-127,197,222,227,283-284,310,311
在两幅图像中 63,64,87,89,97
在一幅图像中 57,63-64,70-71,86,95
代数最小化 83,127,198,199
几何最小化 67,84,123,200,283
重投影 58-60,65,70-71,84,202,212,218-219,242,253,309-310,312
残差 72-73,87-91,99,129,198,200,284,362
转移 57-59,61,240
误差估计 (estimation error) 88-89

**X**

线性方程 (linear equations)
未知秩方程组 410
相同内参数的解法 (sketch solution for identical internal parameters) 337
相似变换 (similarity transformations) 13
相似矩阵 (similarity matrix) 31
消影点 (vanishing points) 15,22-23,107,186
消影点和直线 (vanishing points and lines) 239
消影线 (vanishing lines) 100,120,148,151,152,184
协方差 (covariance)
反向传播 92
超参数情形 93
估计 128
正向传播 90
矩阵 64,70,83,90-91,92-95,97-98,129-130,194,206-208,402-403,417,419-421,422-423
3D 点的 225
对极线的 209
用于点转移 97
协方差估计 (covariance estimation) 422
摄像机参数估计 128
虚圆点 (circular points) 2,17,21,23-27,30-31,32,46,47,50,146,157,242,326,327-328,427

**Y**

延伸基线 (extending the baselines) 318
一般三视图多义性 (general three－view ambiguity) 394
一般位置 (general position) 6－7,9,17,34－35,41,43,56－57,204,228,235,240,247,249,278,288,291,302,378,383,388
一般运动 (general motion) 141,159,167－168,170,171,202,288,331－332,335,347,350,351－352
已知主点 (known principal point) 331,334,337
引导匹配 (guided matching) 79,80－81,201－202,281－283,321
余因子和伴随矩阵 (cofactor and adjoint matrices) 407
圆 (circle) 346
约束透视变换 (elation) 15,31,241,425－426

Z

自变量数 (counting argument)
　平面的 303
自标定 (auto－calibration)——也见度量重构 (metric reconstruction)
　多义性 329
　临界移动序列 351
　求无穷远平面 335
　迭代方法 330
　绝对对偶二次曲面方法的局限性 332
　由仿射到度量 336
　需要的视图数 331
　双眼装置 348
　$Q_\infty^*$非正定时的问题
　消除多义性 339
　求解 $Q_\infty^*$
　　线性 327
　　非线性 330
　典型多义性 339
　利用 $Q_\infty^*$ 327
　变化的参数 338
自对极 (auto－epipolar) 167,176
自由度 (degrees of freedom) 3
　三焦点张量的 261
正定对称矩阵 (positive－definite symmetric matrices) 407
正定矩阵 (positive definite matrix) 417
正负性不等式 (cheiral inequalities)
　求解 373

正规方程 (normal equations) 411
正交关系 (orthogonality relations) 152
正交回归 (orthogonal regression) 249,347
正交矩阵 (orthogonal matrices)
保范性质 404
正交投影 (orthographic projection) 114-115,118,120,321
正向投影 (forward projection) 107,135
直线的 135
二次曲面的 139
直线 (lines)
表示 36,279
变换 10
直线重构 (line reconstruction) 225
直线的映射 (mapping of lines) 119
直线对应 (line correspondences) 56,122,277,280,281-282,288,291,301,304,307,316
直线估计 (line estimation) 74
直线投影矩阵 (line projection matrix)
定义 136
直线转移 (line transfer) 269,272
直纹二次曲面 (ruled quadric) 393
中心投影 (central projection)
轴平面 (axis planes) 105-106,137
摄像机的 106
主点 (principal point) 101-102,103,105,106,108,114-115,126,131,141,148-149,154-155,156-157,176,328-329,331,332,336,337-338,339,342-343,345,351-352
主点偏置 (principal point offset) 103
主平面 (principal plane) 101,105-106,107,111,116-117,125,137,151,185,227,293,297,302,366-367,376
摄像机的 105
主轴矢量 (principal axis vector) 105,106
转动轴 (screw axis) 44-45,51,171,346
转移 (transfer) 57,59,60,68,71,79,80,94,96-99,162,163,184,187,188,222,228,235,238,240,245,258-259,261,266,269,270-274,277,282,326,335,337,344
退化配置 271-272
准仿射重构 (quasi-affine reconstruction)
算法 374
无穷远平面上 $\pi_\infty$ 的边界 374

准仿射映射（quasi－affine mapping）
　定义　366
　二维　366
走廊场景（corridor scene）　199
最大似然估计(MLE)(Maximum likelihood estimation)
　鲁棒　77,81
最小二乘（least－squares）　415
最小二乘解（least－squares solution）　412,416
　约束方程组的　413－414
　齐次方程的　412
　利用正规方程　411
　加权　412
　满秩情形　410
最小对应数（minimum number of correspondences）　121,301－302
最小解（minimal solution）　55－56,69－71,123,132－133,191,246,249
坐标变化（change of coordinates）　66,90,207,333,403
坐标定向（coordinate orientation）　109

# 后 记

由剑桥大学出版社 2000 年出版的 Richard Hartley 和 Andrew Zisserman 的专著《Multiple View Geometry in Computer Vision》全面深入地总结了基于几何的计算机视觉近 10 年来的最新研究成果，对有关的概念、结果及其意义和应用都作了严谨而系统的论述. 尤其难能可贵的是从数学和计算数学方面作了详细而严格的分析与论证，使具有一定数学基础的读者不需阅读参考书就能基本明了. 我们尽微薄之力及时地把它译成中文以飨我国广大同行.

本书译者感谢英国剑桥大学出版社、原著作者 Richard Hartley 和 Sman 在翻译出版过程中提供的帮助. Hartley 教授还特别就译文中的许多细节与我们进行了切磋，他的认真负责精神值得我们学习. 我们感谢马颂德教授在百忙中为本书写了序言. 我们也要感谢安徽大学出版社的大力支持和研究生刘润、王海贤、鲍文霞、周婷婷等协助校对与修改工作.

由于我们水平有限，翻译中错误和不当之处在所难免，恳请广大读者批评、指正。

译 者

2002 年 8 月

CAMBRIDGE
UNIVERSITY PRESS

剑桥大学出版社　最新引进版

特约编辑 章权兵
责任编辑 朱寒冬
封面设计 孟献辉

# 计算机视觉中的多视图几何

# Multiple View Geometry in Computer Vision

ISBN 7-81052-503-4
9 787810 525039 >

ISBN 7-81052-503-4/TP·61
定价 50.00元