



计算机视觉研究中的 投影理论和方法

高满屯 曲仕茹 李西琴 著



T939.1.3

450119

G200

计算机视觉研究中的 投影理论和方法

高满屯
曲仕茹 著
李西琴

西北工业大学出版基金资助出版



3



00450119

西北工业大学出版社

1998年4月 西安

(陕) 新登字 009 号

【内容简介】 本书结合作者近 10 年的工作, 重点介绍了线性投影基本定理、基本方程、变换矩阵, 摄像机参数计算, 基于特征的刚体运动分析, 基于光流的刚体运动分析, 线图解释和形态图计算理论及方法等。

本书适用于从事计算机视觉和工程图学研究的科技工作者, 也可供信号和信息处理、计算机科学与工程、自动化等有关专业高年级学生和研究生以及有关领域的科技工作者参考。

JS 12/19

计算机视觉研究中的投影理论和方法

高满屯 曲仕茹 李西琴 著

责任编辑 季 强

责任校对 樊 力

*

© 1998 西北工业大学出版社出版发行
(710072 西安市友谊西路 127 号 电话 8493844)

全国各地新华书店经销

西北工业大学出版社印刷厂印装

ISBN 7-5612-0931-2/TP·118

*

开本: 850×1168 毫米: 1/32 印张 9.437 5 字数: 228 千字

1998 年 1 月第 1 版

1998 年 1 月第 1 次印刷

印数: 1—2 000 册

定价: 25.00 元

购买本社出版的图书, 如有缺页、错页的, 本社发行部负责调换。

前 言

平面图形（像）是记录或表现三维空间物体（场景）的重要媒介之一，如工程图样、摄影图片、电视图像等。画法几何学和计算机图形学的主要研究内容之一是如何用平面图形（像）描述已知的空间物体（场景）。计算机视觉主要研究内容之一为如何从已知的平面图形（像）识别和理解其所表示的空间物体（场景）。将空间物体（场景）变换为平面图形（像）的基本方法是投影法，其中经常采用的是线性投影法，包括透视投影法和平行投影法。

近几年，我们陆续承担了航空科学基金“机器人视觉中有关的投影理论研究”，国家自然科学基金“计算机解释立体线图的机理研究”和“复杂物体形态图计算理论和高效计算方法研究”等课题的研究工作。本书主要反映的是作者在线性投影理论和方法研究中所做的工作，同时，也编入了其他学者的一些研究成果。

全书共分为六章。第一章简要介绍线性投影研究的问题和国内外研究概况。第二章介绍线性投影基本定理、基本方程以及投影变换矩阵和摄像机参数的计算方法。第三章论述利用对应点、对应直线或对应线束进行刚体运动和结构分析的方法。第四章介绍从光流场恢复刚体运动和结构的理论和计算方法。第五章介绍线图标记结构，基于点或直线解释线图的理论和方法。最后一章介绍形态图的计算理论和确定视觉事件类型的方法。

本书由高满屯主编。曲仕茹撰写第三、五章，李西琴撰写第二、四章，高满屯撰写第一、六章。

我们的研究得到许多同行的支持和帮助，在此致以深深的谢意。

由于作者水平有限，书中难免存在缺点和不足，殷切希望广大读者批评指正。

作 者

1997年8月20日

目 录

第一章 绪论	1
1.1 线性投影研究的问题	1
1.2 国内外研究概况	3
1.2.1 从 3D-LP 求 2D	3
1.2.2 从 3D-2D 求 LP	3
1.2.3 从 2D-LP 求 3D	5
1.2.4 从 3D 求 2D-LP	12
第二章 线性投影理论基础	15
2.1 透视投影	15
2.1.1 克鲁巴第一定理	15
2.1.2 克鲁巴第二定理	19
2.1.3 克鲁巴第一定理的解析解法	26
2.1.4 透视投影基本方程	28
2.2 轴测投影	32
2.2.1 波尔凯-施瓦兹定理	33
2.2.2 高斯定理	35
2.2.3 轴测投影基本方程	37
2.3 线性投影变换矩阵	40
2.3.1 透视投影变换矩阵	40
2.3.2 轴测投影变换矩阵	45
2.3.3 算例和结论	48

2.4	摄像机参数求解	51
2.4.1	投影模型及灭点、角点的确定	51
2.4.2	确定变换矩阵及摄像机参数	60
2.5	透视投影视点坐标计算	63
2.5.1	三灭点透视视点坐标计算	64
2.5.2	两灭点透视视点坐标计算	66
2.5.3	一灭点透视视点坐标计算	67
2.6	基于面积的简比和交比	70
2.6.1	基于面积的简比	70
2.6.2	基于面积的交比	71
2.7	基于体积的简比和交比	75
2.7.1	基于体积的简比	75
2.7.2	基于体积的交比	76
第三章	刚体运动分析	79
3.1	引言	79
3.1.1	刚体性质	80
3.1.2	运动模型	80
3.2	透视投影刚体运动和结构计算	83
3.2.1	几何模型和运动方程	83
3.2.2	利用点对应计算刚体运动和结构(I)	85
3.2.3	平面状刚体	94
3.2.4	利用点对应计算刚体运动和结构(II)	105
3.3	利用直线对应计算刚体运动和结构	112
3.3.1	利用直线对应计算纯旋转运动参数	112
3.3.2	利用直线对应计算平面状刚体运动 参数和结构	116
3.4	利用线束对应计算刚体运动和结构	119

3.4.1	λ 表达式的建立	120
3.4.2	刚体运动参数计算	122
3.4.3	线把结构的恢复	131
3.5	轴测投影刚体运动和结构计算	132
3.5.1	两帧图像下的问题	132
3.5.2	两帧图像下的非唯一性	133
3.5.3	三帧图像下的问题	136
3.5.4	三帧图像下问题求解	137
3.5.5	关于平移参数和刚体结构	143
3.5.6	正轴测投影	144
第四章	光流	145
4.1	引言	145
4.1.1	光流场的计算	145
4.1.2	刚体运动方程	147
4.2	透视投影光流场分析	148
4.2.1	光流运动方程	148
4.2.2	一般刚体运动分析	150
4.2.3	平面状刚体运动分析	153
4.3	斜投影光流场分析	155
4.3.1	一般刚体运动分析	155
4.3.2	平面状刚体运动分析	158
4.4	正投影光流场分析	161
4.4.1	一般刚体运动分析	161
4.4.2	平面状刚体运动分析	168
4.5	线流场分析	174
4.5.1	基于直线的投影方程和刚体运动方程	174
4.5.2	刚体的角速度和角加速度	176

4.5.3	刚体的平移速度和加速度	179
4.5.4	确定空间直线的坐标	181
第五章	线图解释	182
5.1	引言	182
5.1.1	基本假设	183
5.1.2	线图分类	184
5.2	线图的标记	185
5.2.1	棱线的分类	185
5.2.2	节点分类与一致标记	186
5.2.3	自然线图	190
5.3	基于点的线图解释理论和方法	192
5.3.1	正投影线图解释理论和方法	192
5.3.2	透视投影线图解释理论和方法	197
5.4	基于直线的线图解释理论和方法	198
5.4.1	透视投影线图解释理论和方法	198
5.4.2	正投影线图解释理论和方法	207
5.5	关于长度和角度的约束	211
5.5.1	关于长度的约束	212
5.5.2	关于角度的约束	213
5.5.3	正交角点的解释	214
5.5.4	有两个直角的角点解释	216
5.5.5	已知三角形的解释	217
5.5.6	正交角点解释算例	219
第六章	形态图	222
6.1	物体表示方法	222
6.1.1	引例和参数表示法	222

6.1.2	单元分解和空间分割表示法	223
6.1.3	扫描表示法	225
6.1.4	实体几何构造表示法(CSG法)	226
6.1.5	边界表示法	228
6.1.6	线架表示法	230
6.1.7	形态图表示法	231
6.1.8	各种表示法的优缺点	233
6.2	基本概念	235
6.2.1	顶点	235
6.2.2	棱线	236
6.2.3	外形线和轮廓线	236
6.2.4	投影空间	237
6.2.5	特征视图	240
6.2.6	视觉事件	240
6.3	视觉事件分类	240
6.3.1	物体的边面表示	240
6.3.2	视觉事件	242
6.4	透视投影形态图计算理论	244
6.4.1	局部视觉事件	245
6.4.2	点线型视觉事件	246
6.4.3	三重点型视觉事件	248
6.4.4	相切交叉型视觉事件	252
6.4.5	尖点交叉型视觉事件	255
6.5	轴测投影形态图计算理论	258
6.5.1	局部视觉事件	259
6.5.2	点线型视觉事件	260
6.5.3	三重点型视觉事件	261
6.5.4	相切交叉型视觉事件	265

6.5.5	尖点交叉型视觉事件	267
6.6	平面立体形态图计算理论和方法	270
6.6.1	平面型视觉事件	271
6.6.2	有效 VC 型和 CCC 型视觉事件判别 准则	272
6.6.3	平面立体形态图求解方法	272
6.6.4	求解平面立体形态图举例	273
参考文献		279

第一章 绪 论

1.1 线性投影研究的问题

通过投影的方法,将高维空间信息投影到低维空间上,用低维空间信息来表示高维空间的信息,在科学研究和工程实践中有着十分重要的作用。投影的方法有多种多样,其中应用最广的一类方法是线性投影法。

设 $P(x_1, x_2, \dots, x_n)$ 为高维空间 \sum_n 中的任一点, P 在低维空间 \sum_m ($m < n$) 中的投影为 $p(X_1, X_2, \dots, X_m)$, 则定义满足关系

$$X_i = \frac{\sum_{j=1}^n a_{ij}x_j + a_{i0}}{\sum_{j=1}^n b_{ij}x_j + b_{i0}} \quad (i = 1, 2, \dots, m) \quad (1.1)$$

的投影为线性投影,其中 a_{ij}, b_{ij} ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$) 为实常数。

线性投影一般分为广义中心(透视)投影和广义平行(轴测)投影(包括正投影)。广义中心投影变换式如式(1.1)所示,广义平行投影变换式为

$$X_i = \sum_{j=1}^n a_{ij}x_j + a_{i0} \quad (i = 1, 2, \dots, m) \quad (1.2)$$

透视投影是广义中心投影的特殊情况,当式(1.2)中的实常数 a_{ij} 和 b_{ij} 满足一定的约束关系时,式(1.1)才能成为透视投影

的变换式。

轴测投影是广义平行投影的特殊情况,当式(1.2)中的实常数系数 a_{ij} 满足一定的约束关系时,式(1.2)才能成为轴测投影的变换式。

研究线性投影的方法主要有综合法、代数法、解析几何法以及矩阵法等。

本书只涉及三维空间与二维投影平面之间的若干线性投影问题。具体来说,是研究三维空间信息(3D)、线性投影条件(LP)和二维平面信息(2D)这三者之间的关系,因此,所研究的问题有:

(1) 已知三维空间信息和线性投影条件,求解二维平面信息,定名为从 3D-LP 求 2D。

(2) 已知三维空间信息和二维平面信息,求解线性投影条件,定名为从 3D-2D 求 LP。

(3) 已知二维平面信息和线性投影条件,求解三维空间信息,定名为从 2D-LP 求 3D。

(4) 已知三维空间信息,求解二维平面信息和线性投影条件,定名为从 3D 求 2D-LP。

(5) 已知二维平面信息,求解三维空间信息和线性投影条件,定名为从 2D 求 3D-LP。

(6) 已知线性投影条件,求解三维空间信息和二维平面信息,定名为从 LP 求 3D-2D。

后面两个问题,由于未知的信息太多且又没有应用背景,因此没有求解的必要性,国内外对此均没有开展研究。线性投影是一种不可逆的运算,因而使得一些线性投影问题复杂难解。

1.2 国内外研究概况

1.2.1 从 3D - LP 求 2D

已知三维空间信息和线性投影条件求解二维平面信息,是画法几何学和计算机图形学研究的主要问题之一。相对来说,这是一个比较容易解决的问题,已有相当成熟的理论和方法。例如,已知三维空间信息为一物体,则可根据已知的线性投影条件得到该物体的多面正投影图、轴测投影图和透视图等。

1.2.2 从 3D - 2D 求 LP

已知三维空间信息和二维平面信息求解线性投影条件有两方面的问题需要解决:一是,已知的三维空间信息集中是否存在某一个三维空间信息,可以恰好地线性投影为已知的二维平面信息?二是,若已知的三维空间信息可以恰好地线性投影为已知的二维平面信息,那么需要什么样的线性投影条件?这两个问题也是画法几何学研究的主要问题,即所谓的线性投影基本理论。在几何上已部分解决了这些问题,建立了一些线性投影的基本定理,如轴测投影中的波尔凯定理、波尔凯-施瓦兹定理和高斯定理等^[1~4];透视投影中的克鲁巴第一定理、克鲁巴第二定理、别斯金定理和切特维鲁新定理等^[1,5]。在代数上也已部分解决了这些问题,建立了一些投影的基本方程,如透视投影的基本方程、轴测投影的基本方程、正轴测投影的基本方程等^[6,7]。别斯金证明他提出的透视投影基本定理也采用了代数方法^[1],后来才有射影综合(几何)证明^[5]。

近年来,由于研制三维计算机视觉系统的需要,使得确定投影变换矩阵和摄像机内外方位参数变得十分必要。解析几何方法和矩阵方法在线性投影理论研究和实践中得到了广泛的应用。确定

投影变换矩阵有下列三种方式:

(1) 设置图像平面,选取视点或视线,对物体施以一系列旋转、平移和透视变换,最后进行正投影,这种方式广泛应用于计算机图形学中^[8]。

(2) 选取物体与其投影图像上若干对应几何元素,依据投影变换方程确定投影变换矩阵,这种方式广泛应用于计算机视觉中^[9]。

(3) 依据物体投影图像所提供的信息,利用物体本身固有的度量关系确定投影变换矩阵,这种方式在机器人定位中得到了广泛应用。

在求解投影变换矩阵和摄像机参数时可利用的几何元素有对应点^[10]、对应直线^[11]、对应面(积)以及这几种几何元素的组合^[12]等。但目前较多采用的是对应点的方法^[13,14],其中的一个主要问题是为了充分确定摄像机的所有内外方位参数,最少需要多少对对应点,即透视 N 点问题(PNP)。

对于可控制的环境,若预先在其中放置一个特殊的标志物,则利用单幅投影图像就可以充分地确定摄像机参数。Fukui^[15]采用的标志物是一个已知边长的正方形,该正方形的两条对角线分别处于水平位置和垂直位置,要求摄像机透镜中心与正方形中心位于同一已知高度,且摄像机光轴必须通过正方形中心。Courtney^[16]也采用正方形作为标志物,但允许摄像机透镜中心与正方形中心不在同一高度。Magee^[17]使用球体作为标志物,要求摄像机光轴通过球心,利用球体投影图的大小确定摄像机到球心的距离,利用最接近球体轮廓中心大圆的投影图上的点计算摄像机方向,摄像机位置是按球面坐标计算的。Haralick^[18]使用矩形作为标志物,要求矩形所在的平面必须平行于某一坐标面,未知矩形大小时只能确定摄像机的方向,若已知矩形的尺寸,则还可以确定摄像机的位置。Chou^[19]使用屋角(House Corner)作为标志物,利用屋角

一般都含有 3 条两两互相垂直的直线这一事实来对摄像机进行定位。Chen^[20]使用正方体作为标志物,利用从正方体投影中获得的灭点确定摄像机的位置。高^[21]提出使用长方体作为标志物,依据从其投影图像中抽取的直线确定了 7 个三角形,以三角形重心为基础拟合出 3 条直线。用透视投影基本方程对角点进行修正,证明根据已知形状的长方体的一幅三点透视投影图像,可能充分且唯一地确定摄像机的所有内外方位参数,而对长方体的位置没有任何限制。Hung^[22]提出在复杂环境中设置多个标志物,利用模式识别技术识别标志物,通过三维几何变换对机器人进行定位。Haralick^[23]利用已知大小和形状的二次参数曲线的投影求解摄像机参数,这种方法可以把六参数计算问题转化为 2 个三参数计算问题,要求光轴通过图像中心。Ganapathy^[24]在已知投影变换矩阵的基础上给出了两种计算摄像机参数的方法。在确定摄像机参数的方法中,还有采用回转体作为标志物的^[25]。另外,已经开展了一些提高摄像机参数计算精度的工作^[26]。

当确定投影变换矩阵和摄像机参数时,透视投影的灭点起着重要的作用,已经提出了一些确定灭点的方法^[27~29]。

1.2.3 从 2D-LP 求 3D

已知二维平面信息和线性投影条件,求解三维空间信息是计算机视觉中最为重要、研究工作做得最多的问题之一,已经取得了很大的进展。

1.2.3.1 从单幅 2D-LP 求 3D

这个问题是已知一幅二维平面投影图像和获得该图像的投影方法,但不知道三维空间信息;或者只知道一个三维空间信息集,但不知道该集中的哪一个信息可投影为已知的二维平面图像。前面的问题称为无模型驱动的图像理解,后面的问题称为有模型驱动(基于模型)的图像理解。

在基于模型的图像理解中,Roberts^[30]开创性的三维物体识别系统可以从二维投影图识别和描述几种已知的多面体,主要研究的是如何对二维投影图表示的三维物体进行解释,从而获得三维物体的描述。

人们看一幅投影图时,可以很容易地估计出该投影图所表示的物体的大小、三维位置和方向,这是因为人们看图时已做了一定的假设,即在人们的头脑中已存在一些约束关系,如平行性、垂直性、线段定长性和夹角定常性等。据此已进行了一些直接求解物体方位的研究工作^[31,32]。Kanatani^[33]利用对投影图进行三维正交变换的方法,建立了长度和角度的约束并将其用于解释正交角点和含有两个直角的角点,他的算式中含有较多的三角函数。高^[34]建立了统一的长度和角度约束公式,并将其应用于解释正交角点、两直角角点及三角形,该方法直观且易于理解,基本上是采用代数运算。Barnard^[35]研究了角度约束,提供了符合人类视觉习惯的“正交性假设”。Linnainmaa^[36]建立了三角形三边长与其投影的约束并将其用于解释三角形。基于完整模型知识的图像理解方法已应用到一些计算机视觉系统中^[37~39]。

透视投影和轴测投影都是不可逆的,这两种投影方法都可以把一条无限长的直线投影为图像中的一个点,因而仅依据一幅二维投影图像和已知的线性投影条件并不能推断出三维物体的结构形状和大小。但由于图像是三维物体的投影,因此,图像也反映了物体的某些三维特性。

线图的结构主要有拓扑结构、几何结构、代数结构和组合结构。拓扑结构是以线图表示物体上的基本立体、表面、线和点的从属和相对位置关系。几何结构是线图中节点及线的形式。代数结构是线图正确地表示物体的充要条件。组合结构是线图的自由度^[40]。Huffman^[41]和 Clowes^[42]独立地建立了线图的标记法。Waltz^[43]对此进行了扩充,可以进一步标记有阴影的线图。Rosen-