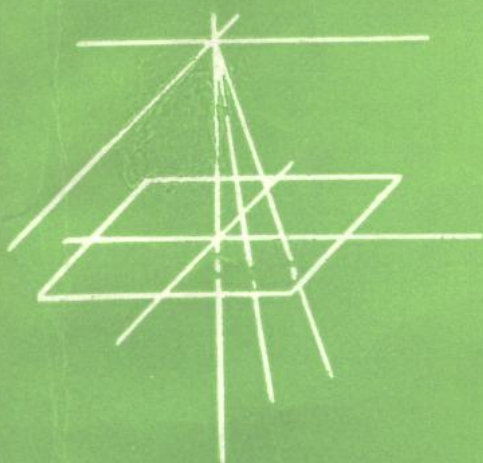




李象霖 编著

三维运动分析

72

中国科学技术大学出版社

78.72
255

三维运动分析

李象霖 编著



中国科学技术大学出版社

1994

9310015

(皖) 新登字 08 号

DE11/04

内容简介

三维运动分析(即图像序列分析)是在二维图像处理技术发展基础上发展起来的一个研究领域。主要内容是利用投影图像序列估计运动物体在三维空间中的运动参数(包括旋转与平移)。本书主要描述特征点对应、质心对应、小平面对应、直线对应及光流等运动模型,采用了张量分析及卡尔曼滤波方法估计运动参数,并采用松弛法、聚类法寻找特征匹配关系,进而提取物体的三维深度信息。书中最后介绍了常用的牛顿算法及最小二乘方算法以及一些基本的数学知识。可供从事图像处理、计算机视觉、三维景像分析与模式识别等专业的大学生、研究生、科研人员和工程技术人员参考使用。

图书在版编目(CIP)数据

三维运动分析/李象霖编. —合肥:中国
科学技术大学出版社, 1994.8
ISBN 7-312-00592-6

I. 三...

II. 李...

III. ①三维运动②图像处理 - 运动分析

IV. TN941

中国科学技术大学出版社出版
(安徽合肥金寨路 96 号 邮政编码 230026)
新华书店北京科技发行所发行各地新华书店经售
北京大兴沙窝店印刷厂印刷
850×1168 毫米 1/32 印张: 6.375 字数: 170 千
1994 年 8 月第一版 1994 年 8 月第一次印刷
印数: 1-2500 册
定价: 5.50 元

前 言

三维运动分析(即图像序列分析)是在二维图像处理技术发展基础上发展起来的一个重要分支,它是计算机视觉的重要组成部分。它研究的主要内容是利用运动物体的投影图像序列,分析与估计物体在三维空间中的运动情况,包括估计运动参数与确定物体的结构。

自 80 年代初美国伊利诺大学黄煦涛教授发表“三维运动估计”论文以来,已有十多年历史。十多年来,各国学者在建立运动模型、数学描述方法、求解运动参数、寻找特征匹配、提取深度信息及算法等方面做了很多的研究工作,发表了很多研究论文。目前,三维运动分析方面的研究工作还在发展之中,比如多物体运动分析、非刚体运动分析、动态景像分析以及人工神经网络、分数维理论、子波理论在运动分析及图像处理中的应用,正越来越引起人们的兴趣。

三维运动分析有着广泛的应用,如机器人视觉、工业过程的自动控制、运动目标的跟踪与识别、云层跟踪、高速公路上的交通监测、人体心脏的活动以及视频信号的带宽压缩等。

本书内容是在中国科学技术大学研究生院(北京)《数字图像处理》课程用讲义的第二部份《三维运动分析》的基础上增补修订而成的,主要介绍三维运动分析的基本概念、理论、方法与算法。全书共十二章。第一章绪论;第二章二维运动分析;第三章特征点对应运动分析;第四章小平面对应运动分析;第五章质心对应运动分析;第六章直线对应运动分析;第七章光流运动分析;第八章张量分析方法用于运动分析;第九章卡尔曼滤波方法用于运动分析;第十章特征匹配与深度提取;第十一章算法介绍;第十二章预备知识。

在本书编写过程中，曾得到龙文澄教授、徐元培研究员、张妙兰副教授、以及其他老师和朋友的支持和帮助，谨在此表示衷心的感谢。限于编者的水平与经验，书中缺点与错误在所难免，欢迎读者批评指正。

编者

1994 年 2 月

目 录

第一章 绪 论	(1)
1.1 数字图像处理技术的发展趋向	(1)
1.2 三维运动分析的基本问题	(2)
1.3 一些基本概念与假设	(4)
1.3.1 刚体性质	(5)
1.3.2 中心投影	(6)
1.3.3 运动参数 R 和 T	(8)
第二章 二维运动分析	(11)
2.1 二维平移	(11)
2.1.1 付里叶变换法	(12)
2.1.2 差分法	(14)
2.1.3 相关法	(15)
2.2 二维旋转	(16)
2.3 一般二维运动分析	(18)
第三章 特征点对应运动分析	(19)
3.1 一般三维运动分析	(19)
3.1.1 纯平移情况	(20)
3.1.2 纯旋转情况	(23)
3.1.3 一般三维运动方程式	(23)
3.2 旋转量小情况下的运动分析	(27)
3.3 线性算法	(32)
3.4 一般解法及其算法步骤	(38)
3.5 双摄像机情况下特征点对应模型	(44)
第四章 小平面对应运动分析	(52)
4.1 引 言	(52)
4.2 多项式方法	(52)
4.3 奇异值分解法	(58)

4.4	线性算法	(62)
第五章	质心运动模型	(70)
5.1	引 言	(70)
5.2	确定点对应关系	(71)
5.3	奇异值分解法确定运动参数	(74)
5.4	最小二乘方法确定运动参数	(77)
5.5	频域法确定运动参数	(80)
第六章	直线对应运动分析	(82)
6.1	引 言	(82)
6.2	直线对应运动模型	(82)
6.3	直线对应运动方程式	(84)
6.4	利用三幅图像求解	(89)
第七章	光流运动分析	(91)
7.1	平面光流	(91)
7.2	运动参数	(93)
7.3	光流模型运动方程式及运动参数求解 ..	(94)
7.4	一种线性算法	(98)
7.5	小平面的情况	(104)
7.6	广义光流场	(105)
第八章	张量分析方法	(109)
8.1	引 言	(109)
8.2	正交投影平面	(109)
8.3	仿射坐标变换	(112)
8.4	张量分析方法	(114)
第九章	卡尔曼滤波方法用于运动分析	(118)
9.1	引 言	(118)
9.2	二维物体运动模型	(119)

9.2.1	状态方程	(120)
9.2.2	测量方程	(121)
9.2.3	递归解	(123)
9.2.4	模拟结果	(124)
9.3	三维物体运动模型	(128)
9.3.1	运动模型与状态方程	(128)
9.3.2	测量方程	(131)
9.3.3	递归算法	(132)
9.3.4	模拟结果	(135)
第十章	特征点匹配技术	(141)
10.1	引言	(141)
10.2	特征点选择	(142)
10.3	松弛算法	(145)
10.4	聚类算法	(150)
10.4.1	特征点选择	(150)
10.4.2	特征点匹配	(152)
10.5	Epipolar 约束	(155)
10.6	视差	(156)
10.6.1	视差的概念	(156)
10.6.2	视差的几何表示	(157)
10.7	深度信息提取	(159)
第十一章	算法	(161)
11.1	引言	(161)
11.2	牛顿-瑞夫森算法	(161)
11.2.1	一个未知数一个方程情况	(161)
11.2.2	n 个未知数 n 个方程情况	(164)
11.3	最小二乘方算法	(167)
11.3.1	线性最小二乘方法	(167)
11.3.2	非线性最小二乘方法	(170)
11.4	高斯-牛顿算法	(174)
第十二章	预备知识	(175)
12.1	迹	(175)
12.2	范数	(175)

12.2.1	向量范数	(175)
12.2.2	矩阵范数	(177)
12.3	奇异值分解	(178)
12.4	张量	(180)
12.4.1	仿射向量空间	(180)
12.4.2	$r = 2$ 情况下的张量	(181)
12.4.3	坐标变换	(182)
12.4.4	图像张量	(183)
12.5	离散系统卡尔曼最优滤波方程	(184)
12.5.1	问题描述	(184)
12.5.2	离散系统模型	(185)
12.5.3	最优滤波方程	(186)
参考文献		(188)

第一章 绪 论

1.1 数字图像处理技术的发展趋向

图像处理技术已经有 30 多年的发展历史,特别是在数字计算机发展以后,使得图像处理得以迅速发展。现在,图像处理已经形成了一套比较完整的学科理论,并已被广泛地应用在国民经济的各个领域。

人们最早研究的是图像的获取、传输和存贮,那时是以电视和医学图像做为研究对象。从多年来的研究知道,图像处理技术的主要内容是图像增强、复原与图像编码。图像增强是人们对获得的图像不满意,而对图像进行像素点处理、邻域处理、平滑滤波、边缘提取等操作,处理后得到比较清晰的图像。图像复原的主要内容是,人们依据造成图像模糊的先验知识,采用代数方法、非线性数学方法对图像进行处理,去掉模糊,以期望得到接近原来的真实图像。已有维纳滤波、最小二乘方复原、奇异值(SVD)分解复原方法、最大熵复原方法、多项式扭曲方法等,是比较成熟的图像复原方法。针对图像的数据量越来越大的实际情况,怎么样用较少的比特率表示图像,经过传输后,在接收端解码以后恢复原图像,或者使恢复后的图像与原图像比较达到最小失真,这是图像编码的基本内容。早期电视图像的 DPCM 方法,现在仍是压缩编码采用的方法之一。近年来发展起来的子带编码、矢量量化编码、金字塔编码以及用于编码的神经网络方法、分数维方法等,使得编码技术的研究不断深入发展。

70 年代发展起来的 X 射线断层成像技术是图像重建,已经被广泛地应用于医学诊断及地球物理勘探工作中。雷顿(Raton)变换及付里叶变换是图像重建的理论基础。目前,热点研究课题及发展趋势是医学图像数据库及三维医学成像技术。

图像处理技术的发展是建立在数学变换方法发展的基础之上的。二维付里叶变换是图像变换的基础。人们既可以在空间域处理图像,又可以在空间频率域处理图像。早期的沃尔什变换、哈达玛变换被用于压缩图像,余弦变换也被用于压缩编码。为了消除像素之间的相关性,获得主分量图像,进行传输,使用霍特林变换是所有数学变换中的最佳变换。齐次坐标变换可被用来描述三维空间中物体的旋转,平移和比例尺寸变化。这些数学变换方法被成功地应用在数字图像处理技术中。1980年法国科学家 Morlet 在进行地震数据分析时提出了子波变换 (Wavelet Transform),它在时域和频域同时具有良好的局部化性质。而且由于对高频成分采用逐渐精细的时域和空域采样步长,从而可以聚焦到被处理对象的任意细节。这一变换已经和将要广泛用于信号处理和图像处理等科技领域中。在今后数年中,子波变换将是被经常使用的又一数学工具。总之,图像处理与分析所采用的数学方法仍处在不断发展之中。

限于二维图像平面上的处理技术与方法,已趋于成熟和完善,而且人们从单幅图像只能获得有限的信息。因此,已有的技术方法已经不能满足人们研究工作的需要。从 70 年代后期开始,人们的研究工作兴趣逐渐由二维平面转向三维空间物体的运动,进行目标的跟踪与识别、动态景像分析等。因此,相邻时刻两幅图像特征匹配与对应问题、三维空间物体上特征匹配问题,物体运动检测、特征提取与算法等,均为人们研究的热点与趋向。三维运动分析是其中主要研究课题之一。

1.2 三维运动分析的基本问题

三维运动分析是在图像处理基础上发展起来的一个重要分支,它是运动图像处理的重要组成部份,也是三维数字图像处理技术发展的必然产物。从 80 年代初美国伊利诺大学黄煦涛教授在欧洲第一届数字信号处理学术会议上发表“三维运动

估计”论文以来，在十几年的时间里，各国学者相继做了很多工作，取得了很有成效的成果。

三维运动分析主要研究空间中的物体运动情况。我们把物体在空间的飞行，称为在空间的运动。为了了解物体的运动情况，人们把 T.V. 摄像机放置在三维空间坐标的某一位置上，对运动物体进行观测，并对其拍摄一系列照片。我们把这一系列照片，称为时间序列图像，或称为运动图像。我们又知道，任何一个飞行体，都有自身的特征，如一些尖锐点、边缘直线、曲线等，被称为特征点、特征直线、特征曲线等。只要物体在空间中运动时，均是在观察者的可视范围以内，物体上的特征均可在序列图像上反映出来。我们把在某一时刻拍摄的图像，称为投影图像，可以理解为这一时刻运动物体在图像平面上的投影。因而时间序列图像是投影图像。在不同时刻拍摄得到的投影图像上，运动物体特征的投影坐标是不同的。三维运动分析研究的内容，就是通过测量与计算投影坐标在图像平面上的变化，来分析运动物体的三维结构，估计物体的运动参数。

为了分析运动物体的结构和估计物体的运动参数，需要有运动模型。目前，已经建立起来的三维运动分析模型有：特征点对应模型、直线对应模型、小平面对应模型、质心对应运动模型以及光流模型等。其中又分为单一摄像机与多摄像机情况。

这里所说的特征对应，是指运动物体上的特征与其在二维图像平面上的投影坐标的对应关系。这种对应关系，可以用运动模型的几何图形表示出来。图 1.1 是特征点对应运动模型的几何图形，也是三维运动分析的基本几何图形。我们对这个几何图形说明如下：假设在空间中有一运动物体，运动前在 t_1 时刻，物体上一特征点 p ，我们用小写字母表示空间物体上点的坐标，记做 (x, y, z) ，该点在二维图像平面上的投影点，用大写字母表示，记做 $P(X, Y)$ 。同样，运动后在 t_2 时刻，在物体上与 p 点对应的点 p' ，坐标为 (x', y', z') ，它在图像平面上的投影点为 $P'(X', Y')$ 。可以知道，同一物体从 t_1 时刻运动到 t_2 时刻，物体上同一点从 $p(x, y, z)$ 变为 $p'(x', y', z')$ ，它们在图像平面上的

投影点则从 $P(X, Y)$ 运动到 $P'(X', Y')$ 。严格地讲, $P(X, Y)$ 与 $P'(X', Y')$ 应该分别在两幅投影图像上。为了分析方便, 我们将它们画在同一图像平面上。 $(\Delta X, \Delta Y)$ 则是投影点 P 从 t_1 时刻到 t_2 时刻的图像空间的位移。如果在运动物体上有一组点, 运动前在 t_1 时刻为 $p_i(x_i, y_i, z_i), i = 1, 2, \dots$, 运动后在 t_2 时刻, 与其对应的为 $p'_i(x'_i, y'_i, z'_i), i = 1, 2, \dots$ 。 p_i 与 p'_i 均是在观察者的可视范围之内, 那么在图像平面上的投影点分别为 $P_i(X_i, Y_i)$ 与 $P'_i(X'_i, Y'_i), i = 1, 2, \dots$ 。通过使用数学方法, 可以获得这些投影点坐标或位移。因此, 利用运动前、后相邻时刻两幅图像上的投影点坐标 (X_i, Y_i) 与 $(X'_i, Y'_i), i = 1, 2, \dots$, 估计和确定三维空间中运动物体的结构与运动参数, 这就是三维运动分析的基本问题。

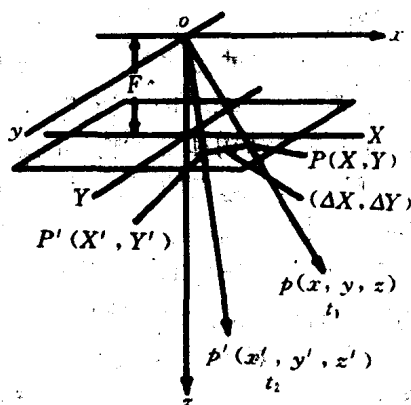


图 1.1 三维运动分析基本几何图形

1.3 一些基本概念与假设

研究观察者与运动物体之间的相对运动, 在许多场合有着重要的应用。为了了解观察者与运动物体之间的相对运动, 需

要对运动参数进行估计。只利用单帧图像是无法分析物体的运动情况,需要使用一序列图像。利用序列图像分析物体的运动,可以从中获得物体运动规律的信息。

根据三维运动分析理论与应用的发展情况,为了学习与分析方便,我们给出一些基本概念与假设。

1.3.1 刚体性质

假设一个物体,它是由一离散的点集 S 所定义 (或构成) 的,那么这个物体的运动是刚体运动,当且仅当在每一时刻 t , 它的三维速度是等投影的,也就是当且仅当满足如下关系

$$(v_1 - v_2) \cdot p_1 p_2 = 0, \quad (\forall t)(\forall p_1 \in S)(\forall p_2 \in S), \quad (1.1)$$

式中 p_1 与 p_2 为点集 S 中的两个点, v_1 与 v_2 是点 p_1 与点 p_2 的三维速度, “ $\cdot$ ” 表示点积。

上述关系式不难证明。因为 p_1 与 p_2 是 S 中的两个点,那么原点 o 与 p_1, p_2 点所形成的向量为 op_1, op_2 。见图 1.2 所示。因为是刚体,刚体上任意两点之间的距离不随时间变化,则有

$$\frac{d}{dt} \|p_1 p_2\|^2 = 0 \quad (1.2)$$

所以得到

$$\left(\frac{d}{dt} op_1 - \frac{d}{dt} op_2 \right) \cdot p_1 p_2 = 0, \quad (1.3)$$

即为

$$(v_1 - v_2) \cdot p_1 p_2 = 0. \quad (1.1)$$

因而可以得到一个定理,即:一个物体是由离散点集 S 构成,当且仅当在每个瞬时 t , 它的三维速度场的投影相等时,它的运动就是一个刚体运动。

这一假设是利用固体运动学的经典结果,它是一个刚体运动的特征描述,是基于在运动中刚体上点之间距离的时间不变性。比较精确地说,一个运动是刚体运动,当且仅当在运动物

体的每一点对之间的距离是不随时间变化的。有人称之为刚体运动的距离守恒原理。由此也可以延伸出，刚体上任意两条直线的夹角在运动中不随时间变化。

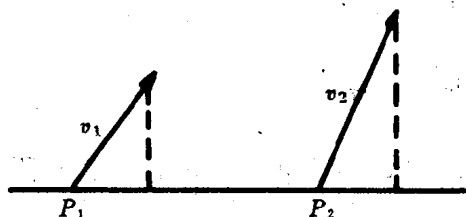


图 1.2 刚体运动定理说明

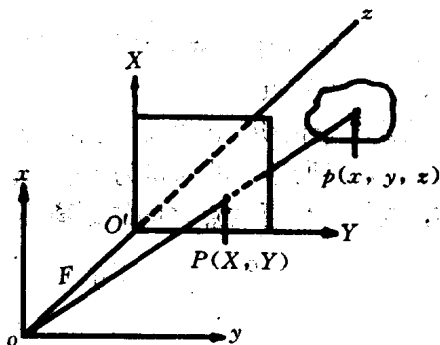


图 1.3 中心投影模型

1.3.2 中心投影

从已有的运动模型的几何图形知道，以观察者（即摄像机）为中心，观测在空间中运动的物体，可以获得物体在二维平面上的投影图像，如图 1.3 所示。图中小写字母 x, y, z 表示三维空间坐标系，原点 o 为观察点，或者叫做透视中心。大写字母 X, Y 表示成像平面。也就是二维投影图像平面。 $oo' = F$ 为焦距，表示观察者与投影图像平面之间的距离。从观察点 o 观测

空间运动物体上一特征点 $p(x, y, z)$, 在图像平面上有一投影点 $P(X, Y)$ 。观察点 o , 物体上点 $p(x, y, z)$, 与投影点 $P(X, Y)$ 在一条直线上, 它们之间满足关系

$$\frac{F}{z} = \frac{X}{x} = \frac{Y}{y}. \quad (1.4)$$

我们把满足这个关系式的结构称为中心投影, 即以观察者为中心的投影模型。

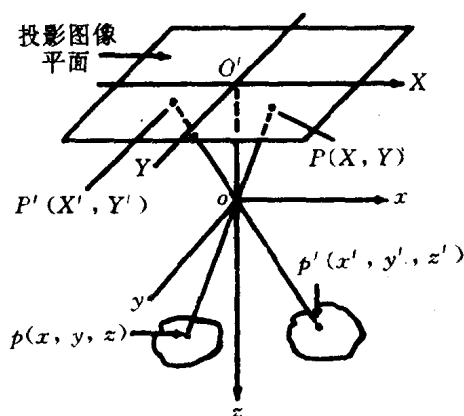


图 1.4 观察者在物体与投影平面之间的中心投影模型

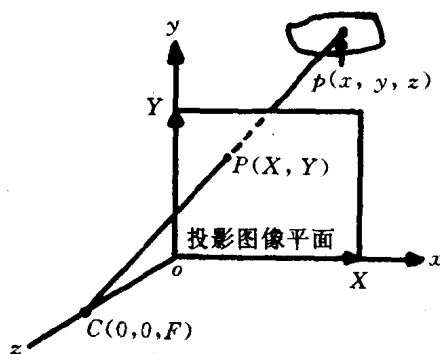


图 1.5 观察点在 z 轴上中心投影模型

图 1.1 与图 1.3 所示的中心投影模型, 观察者是在运动物体和成像平面的一侧。有时观察者也可能在运动物体与投影图像平面之间, 如图 1.4 中的原点 o ; 或者在 z 轴的某一位置上, 如图 1.5 所示的 C 点上。

1.3.3 运动参数

通常情况下, 任何一个物体在三维空间中的运动, 都可以分解为围绕通过原点的一个轴的旋转和平移。在三维空间中的物体旋转可用一个 3×3 矩阵描述, 记做 R ; 平移可以用一个 3×1 的列向量描述, 记做 T 。我们把 R 和 T 称为三维空间中运动物体的运动参数。

如图 1.1 所示, 在空间中一运动物体上有一可视特征点 p , 运动前在 t_1 时刻, 坐标为 (x, y, z) ; 运动后在 t_2 时刻, 与 p 相对应的点为 p' , 坐标为 $p'(x', y', z')$, 那么运动前、后应该满足如下关系:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix} + T, \quad (1.5)$$

其中, R 为旋转矩阵, 定义为

$$R \triangleq \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad (1.6)$$

T 为平移向量, 定义为

$$T \triangleq \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}, \quad (1.7)$$

式中 $\Delta x, \Delta y, \Delta z$ 为运动物体在 x, y, z 三个方向上的平移量。

对于复杂的运动, 可将其分解为若干时间间隔分别进行处理。如图 1.6 所示, 从 t_1 时刻到 t_4 时刻, 不能使用相同的 R 、