计算机视觉中的数学方法

Mathematical Methods in Computer Vision

吴福朝　著

科学出版社
www.sciencep.com

计算机视觉中的数学方法

Mathematical Methods in Computer Vision

吴福朝　著

国家自然科学基金项目(60575019)
国家863计划项目(2006AA01Z116)　资助

科 学 出 版 社
北　京

内 容 简 介

本书由射影几何、矩阵与张量、模型估计 3 篇组成，它们是三维计算机视觉所涉及的基本数学理论与方法。射影几何学是三维计算机视觉的数学基础，本书着重介绍射影几何学及其在视觉中的应用，主要内容包括：平面与空间射影几何，摄像机几何，两视点几何，自标定技术和三维重构理论。矩阵与张量是描述和解决三维计算机视觉问题的必要数学工具，本书着重介绍与视觉有关的矩阵和张量理论及其应用，主要内容包括：矩阵分解，矩阵分析，张量代数，运动与结构，多视点张量。模型估计是三维计算机视觉的基本问题，通常涉及变换或某种数学量的估计，本书着重介绍与视觉估计有关的数学理论与方法，主要内容包括：迭代优化理论，参数估计理论，视觉估计的代数方法、几何方法、鲁棒方法和贝叶斯方法。

本书可作为大学高年级本科生、研究生的教材。相关领域的研究人员也可以参考。

图书在版编目(CIP)数据

计算机视觉中的数学方法/吴福朝著. —北京：科学出版社，2008

ISBN 978-7-03-021023-4

Ⅰ. 计…　Ⅱ. 吴…　Ⅲ. 计算机视觉-数学方法　Ⅳ. TP302.7

中国版本图书馆 CIP 数据核字(2008)第 014655 号

责任编辑：王淑兰/责任校对：刘彦妮

责任印制：吕春珉/封面设计：三函设计

科学出版社出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

骏杰印刷厂印刷

科学出版社发行　各地新华书店经销

*

2008 年 3 月第 一 版　开本：787×1092　1/16

2008 年 3 月第一次印刷　印张：24 1/4

印数：1—3 000　字数：572 000

定价：48.00 元

（如有印装质量问题，我社负责调换〈新欣〉）

销售部电话 010-62136131　编辑部电话 010-62130750

前　言

计算机视觉的研究目标是使计算机具有从二维图像认知三维现实环境的能力，这种能力不仅使计算机能感知三维物体的结构与运动信息，而且能对它们进行识别与理解。自20世纪80年代初形成的Marr计算框架，以及随后发展的其他计算理论框架，使得人们相信通过计算手段能实现从二维图像对三维现实世界的感知。为适应不同计算理论框架和增强计算机视觉系统鲁棒性的需要，研究人员通过坚持不懈的努力，引进了众多数学方法和相应的计算方法。这些数学方法涉及代数学、几何学、分析学、概率论以及众多应用数学分支，这对初学者和从事多年研究的人员都感到非常困惑，真的需要这么多复杂的数学来解决计算机视觉问题吗？是数学的无能还是我们没有发现能解决计算机视觉问题的一种“好数学”？这正是计算机视觉的魅力所在，它吸引着众多学者在该领域进行探索与研究。

如果“计算能实现从二维图像对三维现实世界的感知”是正确的话，则不论使用什么数学方法其目的都是为了做两件事情，即为计算机视觉问题建立数学模型和模型估计与选择。在建立模型过程中，需要反映计算机视觉的数学理论和需要描述模型的数学工具与方法；在模型估计与选择中，需要数学计算技术，包括数值计算与统计计算。本书从以上提出的3个需要出发，介绍三维计算机视觉所涉及的基本数学理论与方法。

本书由以下3篇构成：

第一篇：射影几何。它是三维计算机视觉的数学理论基础，主要内容包括：平面与空间射影几何，摄像机几何，两视点几何，自标定理论和三维重构理论。

第二篇：矩阵与张量。它是描述三维计算机视觉模型的基本数学工具，同时也是模型估计线性算法的数学基础，主要内容包括：矩阵分解，矩阵分析，张量代数，运动与结构，多视点张量。

第三篇：模型估计。模型估计是三维计算机视觉的基本问题。本篇描述模型估计与选择的基本数学理论与方法，主要内容包括：迭代优化理论，参数估计理论，视觉模型估计的代数方法、几何方法、鲁棒方法与统计方法，以及模型选择的数学方法。

上述3篇所涉及的数学内容是相对独立的，但计算机视觉将它们组成一个有机的整体。

本书是在作者为计算机视觉方向的研究生所开设的课程讲义基础上形成

的，目的是为初学者介绍三维计算机视觉中的基本数学理论、方法及其应用。通过阅读本书，读者能提高分析、解决有关计算机视觉问题的数学能力，为进入该领域前沿研究做准备。

本书的出版得到国家自然科学基金项目(60575019)和国家 863 计划项目(2006AA01Z116)资助，特此致谢。

本书写作过程中，参考了有关书籍和文献(见本书列出的参考文献)，在此向这些作者一并致谢。

作　者
于北京中关村

目　　录

第二篇 矩阵与张量

第三篇 模型估计

第一篇

射 影 几 何

射影几何是三维计算机
视觉的数学基础

本篇提要

本篇内容分为两个部分。第一部分由第1、2章构成，主要介绍射影几何理论，这些理论不是射影几何所涉及的全部内容，而是从事三维计算机视觉研究所必须掌握的几何知识；第二部分由第3～6章构成，主要论述射影几何在三维计算机视觉中的应用，读者将会看到射影几何的重要作用。各章具体内容如下：

第1章介绍平面射影几何。主要内容包括：点、线和二次曲线的齐次表示；二次曲线与对偶二次曲线的性质；二维射影变换与基本几何元素的射影变换规则；二维射影变换群及其子群的不变量与不变性质。

第2章介绍空间射影几何。主要内容包括：点、线、面和二次曲面的齐次表示，二次曲面与对偶二次曲面的性质；三维射影变换与基本几何元素的射影变换规则；三维射影变换群及其子群的不变量与不变性质。

第3章介绍摄像机几何。首先，对摄像机进行数学建模；然后，应用前两章的射影几何知识，给出空间基本几何元素的投影性质，以及图像平面基本几何元素的反投影性质。这些投影与反投影性质，是从图像恢复物体几何结构的基础，尤其是绝对二次曲线与绝对二次曲面的投影性质。

第4章介绍两视点几何。应用射影几何知识引进两幅图像的点、线关联关系，即所谓的极几何。基本矩阵是极几何的代数描述，同时它也给出了射影意义下的摄像机矩阵。

第5章介绍自标定理论。主要介绍正交性、圆环点、平行性、Kruppa方程、绝对二次曲线和绝对二次曲面与摄像机内参数的关系。所有自标定方法均来源于射影几何理论，它们都归结为绝对二次曲线或绝对二次曲面的投影性质。

第6章介绍三维重构理论。三维重构是计算机视觉中的核心问题，本章应用射影几何知识建立三维重构理论。不需要图像以外的知识，从图像就能获得射影重构；对于仿射重构，它与确定无穷远平面的射影坐标、确定无穷远单应是相互等价的；对于度量重构，它与确定绝对二次曲线、确定绝对二次曲面和确定摄像机内参数是相互等价的。

第 1 章　平面射影几何

1.1　射 影 平 面

1.1.1　射影平面

1. 齐次坐标

在本章中，除特别说明外，均假定在平面上建立了欧氏坐标系。平面上的点可用二维有序数组 $\tilde{\boldsymbol{p}} = (x,y)^T$ 来表示，即该点的欧氏坐标。平面上的直线方程可以表示为

$$ax + by + c = 0 \tag{1.1.1}$$

在方程(1.1.1)两边同乘以任一非零常数 t，得到下述方程

$$axt + byt + ct = 0 \tag{1.1.2}$$

方程(1.1.1)与方程(1.1.2)有相同的几何意义，它们表示同一条直线。令

$$\boldsymbol{p} = (xt, yt, t)^T, \boldsymbol{l} = (a,b,c)^T$$

则方程(1.1.2)可写成

$$\boldsymbol{l}^T\boldsymbol{p} = 0 \tag{1.1.3}$$

其中 $\boldsymbol{p}$ 是变量，表示直线上的点；$\boldsymbol{l}$ 是一个固定的向量，代表该直线。

一般地，$\boldsymbol{p} = (xt, yt, t)^T$ 称为点的齐次坐标，$\boldsymbol{l} = (a,b,c)^T$ 称为直线的齐次坐标。这里的“齐次”也可以这样理解，在这种表示下直线方程(1.1.3)关于点或直线变量都是齐次的，而方程(1.1.1)则是非齐次的。

齐次坐标可以相差任意的非零常数因子，即 $\forall s \neq 0$，$\boldsymbol{p}$ 和 $\boldsymbol{q} = s\boldsymbol{p}$ 表示同一个点，因为它们的非齐次坐标相等

$$\tilde{\boldsymbol{p}} = (x/t, y/t)^T = (sx/st, sy/st) = \tilde{\boldsymbol{q}}$$

直线的齐次坐标也可以相差任意的非零常数因子，因为方程 $(s\boldsymbol{l})^T\boldsymbol{p} = 0$ 与方程(1.1.3)确定同一条直线。

2. 射影平面

齐次坐标为 $\boldsymbol{p}_\infty = (x,y,0)^T$ 的点称为无穷远点，其中 x，y 至少有一个不为零。注意，无穷远点没有欧氏坐标，这是因为 $x/0=\infty$，$y/0=\infty$ 至少有一个成立，同时也可以看出为什么将它称为无穷远点。平面上所有无穷远点所构成的集合称为无穷远直线。由于所有无穷远点 $\boldsymbol{p}_\infty$ 都满足方程

$$0 \cdot x + 0 \cdot y + 1 \cdot 0 = 0 \tag{1.1.4}$$

所以，无穷远直线的齐次坐标为 $\boldsymbol{l}_\infty = (0,0,1)^T$。

由欧氏平面与无穷远直线的并集所形成的扩展平面称为射影平面，有时也称为二维射影空间。

1.1.2 叉积

1. 三维向量的叉积

令 $\boldsymbol{x}_1=(x_1,y_1,t_1)^T,\boldsymbol{x}_2=(x_2,y_2,t_2)^T$ 是两个三维向量，它们的叉积定义为

$$\boldsymbol{x}_1\times\boldsymbol{x}_2=\det\begin{pmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ x_1 & y_1 & t_1\\ x_2 & y_2 & t_2\end{pmatrix}=\left(\det\begin{pmatrix}y_1 & t_1\\ y_2 & t_2\end{pmatrix},-\det\begin{pmatrix}x_1 & t_1\\ x_2 & t_2\end{pmatrix},\det\begin{pmatrix}x_1 & y_1\\ x_2 & y_2\end{pmatrix}\right)^T \tag{1.1.5}$$

叉积和反对称矩阵相关联。由向量 $\boldsymbol{x}=(x,y,t)^T$，按下述方式定义反对称矩阵

$$[\boldsymbol{x}]_\times=\begin{pmatrix}0 & -t & y\\ t & 0 & -x\\ -y & x & 0\end{pmatrix} \tag{1.1.6}$$

称为由向量 $\boldsymbol{x}$ 所确定的反对称矩阵。矩阵 $[\boldsymbol{x}]_\times$ 具有下述性质：

1) 对任意非零向量 $\boldsymbol{x}$，有 $\mathrm{rank}([\boldsymbol{x}]_\times)=2$；

2) 对任意两个三维向量 $\boldsymbol{x}_1,\boldsymbol{x}_2$，有 $\boldsymbol{x}_1\times\boldsymbol{x}_2=[\boldsymbol{x}_1]_\times\boldsymbol{x}_2$；

3) $\boldsymbol{x}$ 是 $[\boldsymbol{x}]_\times$ 的右零空间，同时也是左零空间，即 $[\boldsymbol{x}]_\times\boldsymbol{x}=0,\boldsymbol{x}^T[\boldsymbol{x}]_\times=0$；

4) 对任意三维向量 $\boldsymbol{y}$，有 $\boldsymbol{y}^T[\boldsymbol{x}]_\times\boldsymbol{y}=0$。

性质 1)是明显的，3)与 4)可由性质 2)导出，下面验证性质 2)。

令 $\boldsymbol{x}_1=(x_1,y_1,t_1)^T,\boldsymbol{x}_2=(x_2,y_2,t_2)^T$，则

$$[\boldsymbol{x}_1]_\times\boldsymbol{x}_2=\begin{pmatrix}0 & -t_1 & y_1\\ t_1 & 0 & -x_1\\ -y_1 & x_1 & 0\end{pmatrix}\begin{pmatrix}x_2\\ y_2\\ t_2\end{pmatrix}=\begin{pmatrix}-t_1y_2+y_1t_2\\ t_1x_2-x_1t_2\\ -y_1x_2+x_1y_2\end{pmatrix}=\boldsymbol{x}_1\times\boldsymbol{x}_2$$

性质 2)表明，两个向量的叉积可以用其中一个向量的反对矩阵左乘另一个向量来表达。

2. 两点、两线的叉积

如果 $\boldsymbol{p}_1,\boldsymbol{p}_2$ 是射影平面上两点，则 $\boldsymbol{l}=\boldsymbol{p}_1\times\boldsymbol{p}_2$ 表示通过这两点的直线。这是因为，对直线 $\boldsymbol{l}$ 上的任一点，它的齐次坐标可以表示为 $\boldsymbol{p}=s_1\boldsymbol{p}_1+s_2\boldsymbol{p}_2$，根据反对称矩阵的性质 3)和 4)，可以得到

$$\boldsymbol{l}^T\boldsymbol{p}=\sum_{j=1}^{2}s_j([\boldsymbol{p}_1]_\times\boldsymbol{p}_2)^T\boldsymbol{p}_j=0$$

若三点 $\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3$ 共线，则必有 $\boldsymbol{p}_2^T[\boldsymbol{p}_1]_\times\boldsymbol{p}_3=-([\boldsymbol{p}_1]_\times\boldsymbol{p}_2)^T\boldsymbol{p}_3=0$。反之，若 $\boldsymbol{p}_2^T[\boldsymbol{p}_1]_\times\boldsymbol{p}_3=0$，则三点 $\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3$ 必共线。因此，有下述命题：

命题 1.1.1 1)两点 $\boldsymbol{p}_1,\boldsymbol{p}_2$ 连线的坐标是 $\boldsymbol{l}=[\boldsymbol{p}_1]_\times\boldsymbol{p}_2$；2)三点 $\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3$ 共线的充要条件是 $\boldsymbol{p}_2^T[\boldsymbol{p}_1]_\times\boldsymbol{p}_3=0$。

对偶原理 在射影平面内，点和线是一对互为对偶元素。在包含“点”和“线”元素的命题中，如果该命题成立，则将两个元素的角色互换，对应的命题也成立。

例如，命题 1.1.1 有如下对偶命题：