

国家自然科学基金资助项目成果

计算机视觉检测 逆问题导论

Introduction to Inverse Problems in
Computer Vision Inspection

◎ 罗四维 罗晓月 著



北京交通大学出版社
<http://www.bjtup.com.cn>

国家自然科学基金资助项目成果

计算机视觉检测 逆问题导论

罗四维 罗晓月 著

北京交通大学出版社
· 北京 ·

内 容 简 介

本书以计算机科学技术为中心,以数学逆问题和图像处理为基础,讨论了计算机视觉检测的逆问题处理方法。本书重点介绍了用来解决数学逆问题的正则化方法及其基本原理、计算机视觉图像重建技术中的逆问题、基于前向运动视频的计算机视觉检测新思路,同时还具体描述了用于高速铁路基础设施安全检测的计算机视觉检测方法。

本书可供计算机视觉、人工智能、图像处理、视觉测量技术、高速铁路安全检测方面的研究人员和高等院校有关专业的教师、研究生等相关人员参考。

版权所有,侵权必究。

图书在版编目(CIP)数据

计算机视觉检测逆问题导论 / 罗四维, 罗晓月著. —北京: 北京交通大学出版社, 2017. 3

ISBN 978 - 7 - 5121 - 3041 - 8

I. ①计… II. ①罗… ②罗… III. ①计算机视觉 - 检测 - 研究
IV. ①TP391. 41

中国版本图书馆 CIP 数据核字 (2016) 第 213696 号

计算机视觉检测逆问题导论

JISUANJI SHIJUEJIANCE NIWENTI DAOLUN

总 策 划: 章梓茂 责任编辑: 韩 乐 助理编辑: 严慧明

出版发行: 北京交通大学出版社 电话: 010 - 51686414 [http: //www. bjtu. cn](http://www.bjtu.cn)

地 址: 北京市海淀区高粱桥斜街 44 号 邮编: 100044

印 刷 者: 北京鑫海金澳胶印有限公司

经 销: 全国新华书店

开 本: 185 mm × 260 mm 印张: 14 字数: 323 千字

版 次: 2017 年 3 月第 1 版 2017 年 3 月第 1 次印刷

书 号: ISBN 978 - 7 - 5121 - 3041 - 8 / TP · 835

定 价: 68.00 元

本书如有质量问题, 请向北京交通大学出版社质监组反映。对您的意见和批评, 我们表示欢迎和感谢。
投诉电话: 010 - 51686043, 51686008; 传真: 010 - 62225406; E-mail: press@bjtu.edu.cn。

前 言

自有人类以来，我们就用视觉支持我们的生活、工作、交流，了解世界，致使视觉成为一个古老的研究课题，但是直到 20 世纪 70 年代中期人们才对视觉计算理论进行研究。随着计算机技术的发展，计算机功能越来越强大，甚至表现出超出人们想象的能力。然而，人类轻而易举就能完成的图像识别任务却对计算机的能力构成了无情的挑战。作为一个社会整体，最智能的机器还是盲人。直面挑战，人们从图像处理、模式识别、视觉心理学、视觉神经心理学、机器人视觉、视觉智能等诸多方面展开了大量和卓有成效的研究。时至今日，计算机虽然在视觉识别能力方面远不及人类，但逐步完善的计算机视觉研究成果已经被不断推广到众多应用领域，发挥着人类视觉不可替代的作用，为数字世界带来一线“光明”。计算机视觉就是要实现机器视觉智能化，一方面可以像人类一样对可观测到的物体进行相应处理；另一方面，计算机视觉通过物理或化学的手段获取不可观测物体内部结构信息，并依据这些信息生成图像，开辟了不依靠光学系统的图像可视化或图像重建技术。图像重建技术是计算机视觉优于人类视觉的一个重要特征，仅就医学领域而言，图像重建技术开创了工程技术与医学相结合的新的篇章。

计算机视觉检测是计算机视觉领域中的一个重要应用分支，是以图像为依据的一种检测手段。计算机视觉可以精确到以像素为计算单位，这在测量范围及精度方面大大优于人类视觉功能。为此，计算机视觉检测作为一种无损检测手段已经被广泛应用到工业、农业、医学、地球物理等众多领域中。计算机视觉检测是以图像为依据的一种检测手段，图像的表现形式与检测任务息息相关，为适应不同的检测目的而采用多种手段对图像进行不同效果的处理就显得必不可少。

计算机视觉中的很多问题的理论基础是统计学理论、最优化理论及数学逆问题相关理论。逆问题本身并不是什么新的问题，它广泛存在于众多问题之中，数学逆问题只是从这些现实问题中归纳出的一类数学问题，当现实中的表现形式转化成严格的数学表述时，形式化的方法就变为可能。逆问题是相对正问题而言的，视觉自身的“视”与“觉”就形成了一对相反问题，“视”对应的正问题是成像过程，“觉”对应的逆问题是图像的认知过程。

用数学公式表示计算机视觉处理过程中的一个关键性的步骤，就是要根据处理对象的一般特性，找出有关问题的约束条件，并将它们转化成精密的模型，从而得出确凿的、经得起考验的结论。这种结论将经得起检验，因为它们是在处理对象和数学公式的

基础上得出来的。以数学为基础的计算机视觉模型的重要特征就在于它使视觉信息处理的研究变得更加严密,把视觉研究从描述水平提高到严谨的数理科学水平。

全书共 10 章。第 1 章绪论,介绍计算机视觉原理、计算机视觉检测、计算机视觉检测的任务。第 2 章相关知识,本章将这本书所涉及的一些数学方面的相关基础知识汇集在一起,便于读者在阅读本书有关内容时查阅一些数学专业术语的含义或相关基础知识。这些知识在许多教科书中都有详细解释,读者可根据自身需要进行查阅或直接跳过。第 2 章主要介绍算子方程、最小二乘估计等基础概念,变分法理论基础,抽象空间,凸优化理论,微分流形及其几何特征等相关内容。第 3~6 章可以归纳为与数学逆问题有关的内容。第 3 章逆问题,介绍逆问题涉及的问题范围,以及逆问题的数值化的几种典型方法。第 4 章正则化策略,介绍求解逆问题稳定解的正则化思路,包括 Tikhonov 正则化算子及正则参数的选择策略及方法。第 5 章其他正则化方法,介绍除 Tikhonov 正则化方法之外的迭代方法、边界约束正则化方法、最大熵正则化方法、全变分正则化方法、流形正则化方法、贝叶斯方法,另外简要讨论非线性逆问题正则化求解手段。第 6 章图像重建技术中的逆问题,以计算机断层扫描成像技术为对象介绍图像重建技术中采用的逆问题稳定解法,包括计算机视觉图像重建技术中的逆问题,计算机断层扫描图像重建技术理论基础,滤波反投影重建方法,迭代重建算法。第 7~10 章可以归纳为与计算机视觉检测有关的内容。第 7 章机器人视觉测量,主要内容包括机器人视觉概述和视觉测量。机器人视觉测量是目前计算机视觉测量在控制领域中的一类实际应用。第 8~10 章涉及以目前大量存在的前向运动视频图像为检测对象的计算机视觉检测技术。第 8 章全景图拼接,介绍全景图拼接技术,为后续两章内容作基础铺垫。第 9 章前向运动视频三维全景图,在阐述了前向运动视频的定义、应用意义的基础上,介绍近年来利用前向运动视频进行计算机视觉检测的一些新方法,着重讨论用于高速铁路基础设施安全检测的计算机视觉检测方法。第 10 章双缝采样立体全景图,介绍利用前向运动视频并采用双缝采样手段生成立体全景图的具体方法。利用前向运动视频,无论是生成场景全景图(第 9 章),还是生成场景立体全景图(第 10 章),它们最终的目标都是将传统的视频存储方式转换为图像方式,以达到减小存储空间和提高检测效率的目的。

计算机视觉检测是以图像作为对象,为此在介绍计算机视觉检测时必将涉及图像复原、增强、分割、识别等属于计算机视觉领域的重要内容。为突出重点,上述内容在本书中均未被介绍。另外,前向运动视频的运动模糊在高速运动的情况下表现得较为明显,在本书中虽未针对该问题展开讨论,但在实际使用前向运动视频图像时需要认真对待该问题。

本书强调的基本技术思想体现在,从科学层面上建立以数学为基础的详细模型,以恢复感兴趣的信息,考虑到计算的复杂性,在必要时需简化假设条件,从而使数学形式更加简洁。在工程层面上还需要对测量过程和噪声的先验可能性进行估计,并分析其产

生的不确定性，然后使用相关的优化或反演技术给出简单的描述和实现的技术，测试这些技术以了解它们的限制和故障模式。

受时间和精力所限，我们仅将最近的研究成果梳理成这本著作，有一些有价值的内容未能编入，敬请谅解，且书中一定有不妥或争议之处，恳切希望读者提出宝贵意见。

感谢所有和我一起研究、探讨问题及帮助我完成这本书的人，其中有黄雅平教授领导的视觉认知研究小组成员，以及当年从事计算机视觉研究的多位博士研究生（王亮、于欣妍、蒋欣兰、王胜春、吕国豪）。感谢国家自然科学基金委员会的资助（项目批准号：61272354，61273364），感谢北京交通大学出版社韩乐老师、严慧明老师为出版本书所付出的辛勤劳动。

作者
2017 年 2 月

目 录

第1章 绪论	1
1.1 计算机视觉	1
1.1.1 计算机视觉原理	2
1.1.2 计算机视觉研究范畴	2
1.2 计算机视觉检测	6
1.2.1 概述	7
1.2.2 计算机视觉检测应用实例	7
1.3 计算机视觉检测的任务	9
1.3.1 视觉检测	10
1.3.2 可视化检测	11
1.3.3 摄像机标定	14
参考文献	16
第2章 相关知识	17
2.1 基础概念	17
2.1.1 算子方程	17
2.1.2 最小二乘估计	18
2.1.3 范数	19
2.1.4 期望值	20
2.1.5 最大熵方法	21
2.1.6 线性和非线性	23
2.1.7 移不变	23
2.1.8 卷积	24
2.1.9 内积	25
2.1.10 欧拉方程	26
2.1.11 柯西序列	27
2.1.12 三次样条曲线原理	27
2.2 变分法理论基础	28
2.2.1 数函微分与泛函变分	28
2.2.2 变分法	29
2.3 抽象空间	32
2.3.1 对偶空间	32

2.3.2	Banach 空间与 Sobolev 空间	32
2.3.3	变指数函数空间	35
2.3.4	度量空间	35
2.3.5	Hilbert 空间	36
2.4	凸优化理论	36
2.4.1	最优化	36
2.4.2	凸函数	38
2.4.3	范数逼近问题	39
2.4.4	最速下降法	40
2.5	微分流形及其几何特征	40
第 3 章	逆问题	43
3.1	逆问题的概念	43
3.2	逆问题的求解方法	48
3.3	逆问题的数值化	51
3.3.1	基于数值近似积分方法	52
3.3.2	基于代表函数的线性组合方法	53
3.3.3	基于正交基函数的线性组合方法	53
3.3.4	Backus 和 Gilbert 的离散化方法	54
	参考文献	56
第 4 章	正则化策略	57
4.1	引言	57
4.2	理解正则化	58
4.2.1	最小二乘法	58
4.2.2	图像复原问题	60
4.2.3	正则化直观解释	62
4.3	解逆问题的思路	67
4.4	Tikhonov 正则化	69
4.4.1	构造正则化算子	69
4.4.2	正则性条件	70
4.4.3	Tikhonov 展平泛函	70
4.5	后验正则参数选择策略	72
	参考文献	83
第 5 章	其他正则化方法	85
5.1	迭代方法	85
5.1.1	概述	85
5.1.2	Landweber 迭代方法	86

5.1.3 共轭梯度方法	87
5.1.4 共轭梯度最小平方方法	90
5.2 边界约束正则化方法	92
5.3 最大熵正则化方法	92
5.4 全变分正则化方法	93
5.5 流形正则化方法	94
5.5.1 一致能量泛函正则化方法的不足	95
5.5.2 流形正则化方法的基本思想	95
5.5.3 流形空间上的几何特征	96
5.5.4 流形正则化图像复原模型	97
5.6 非线性逆问题正则化	99
5.6.1 Gauss - Newton 和 Levenberg - Marquardt 算法	99
5.6.2 非线性最小二乘问题的正则化	103
5.6.3 Occam 求逆法	104
5.7 贝叶斯方法	105
5.7.1 贝叶斯方法的解	106
5.7.2 多元正态分布情况	108
5.8 最大熵方法	109
参考文献	111
第6章 图像重建技术中的逆问题	112
6.1 重建技术	112
6.1.1 计算机断层扫描图像重建技术理论基础	112
6.1.2 计算机断层扫描的投影变换	113
6.2 一种简单的重建算法	118
6.3 滤波反投影重建算法	120
6.3.1 傅里叶中心切片定理	120
6.3.2 滤波反投影重建算法	121
6.4 迭代重建算法	122
6.4.1 概述	123
6.4.2 代数重建技术及其分析	126
6.4.3 小结	129
参考文献	130
第7章 机器人视觉测量	131
7.1 机器人视觉概述	131
7.1.1 机器人视觉	131
7.1.2 机器人视觉的作用	133
7.1.3 机器人视觉测量	133

7.1.4 摄像机标定模型	136
7.2 视觉测量	146
参考文献	151
第8章 全景图拼接	153
8.1 全景图拼接技术基础	153
8.1.1 全景图	153
8.1.2 全景图拼接基础知识	156
8.2 图像拼接技术	165
8.2.1 图像配准	165
8.2.2 单一视点全景图拼接方法	168
8.2.3 多视点全景图拼接方法	170
8.3 “条带”拼接的基本原理	174
参考文献	175
第9章 前向运动视频三维全景图	178
9.1 视频全景图拼接方法简介	178
9.1.1 基于图像配准的视频全景图拼接方法	180
9.1.2 基于流形投影的视频全景图拼接方法	181
9.1.3 基于图谱结构最优化的视频全景图拼接方法	181
9.2 基于场景几何结构的前向运动	
视频全景图拼接方法	182
9.2.1 虚拟矩形通道全景采样模型	182
9.2.2 基于消失点和摄像机运动信息的图像配准	188
9.2.3 基于虚拟矩形通道全景图的虚拟场景绘制	196
9.2.4 实例	197
参考文献	200
第10章 双缝取样立体全景图	202
10.1 立体全景图成像模型简介	202
10.2 双缝取样立体全景图成像模型	205
10.3 实验结果	208
参考文献	210

第1章 绪 论

视觉是针对眼睛观察到的目标对象所形成的反应，或者说是眼睛对所观察到的目标对象产生的一种感知。感知是有生命的主体与存在的客体之间基本关系的表达，即有生命的主体与客体的基本关系是感知的关系。人类感知器官包括视觉、听觉、嗅觉、味觉、触觉、第六感官等。主体对关系的表达就是感知，具体说就是一个关于目标物体的有效描述，如视觉感知场景中物体的形状和位置、物体的运动方式和速度等。

“感”和“知”是生命具备的基本本能，感知的能力在不同物种、不同个体间各不相同，“感”和“知”都是在本能和存在环境的影响下自然形成的。感知的能力是意识能力的基础，意识的产生以感知为条件。

因此，逻辑上客体对主体的刺激称为正向问题，而主体与客体的关系表达则称为逆向问题，逆向问题需要判断主体对存在客体的关系表达是否正确，这往往有多个解或无解。在数学上把逆向问题简称为逆问题，一般来说逆问题是病态的，解决病态问题的核心方法是正则化。

计算机视觉是多个应用领域（如工业制造、医疗、军事等）中各种智能/自主系统中不可分割的一部分。计算机视觉具有很大的重要性，因此一些先进国家把对它的研究视为一种重大挑战，并将其列为对国民经济和科学研究有广泛影响的科学和工程中的重大基本问题。计算机视觉的挑战主要表现在它是以计算机为处理中心来实现具有与人类视觉水平相当的视觉能力。计算机视觉与人类视觉密切相关，对人类视觉有一个正确的认识将对计算机视觉的研究非常有益。

作为一门学科，计算机视觉开始于20世纪60年代初，直到20世纪70年代后期，当计算机的性能提高到可以处理大规模数据时，计算机视觉才得到了正式的关注和发展。20世纪80年代后，计算机视觉取得了许多重要进展。

1.1 计算机视觉

在几亿年的进化中，视觉已经成为我们人类最重要的感知系统。在人类的大脑中有半数以上神经细胞是与视觉处理联系在一起的。人类视觉系统是迄今为止人们所知道的功能最强大和最完善的视觉系统。计算机视觉就是用各种成像系统来模拟人类“视”的功能，或者说是成像功能，然后再由计算机来代替大脑完成处理和解释。计算机视觉的最终研究目标就是使计算机能像人那样通过视觉观察世界并进一步认识世界，无疑，这是一个伟大的目标，且这一目标的实现需要一个漫长的过程，而目前计算机视觉的研究正处于这一过程之中。近年来，计算机视觉研究取得了一些进展，如建立了一种视觉

系统,该系统能依据视觉感应和对感应的反馈,达到某种程度的智能。计算机视觉典型的应用有自主车辆的视觉导航,以及具有自动避让公路上的车辆等功能的视觉辅助驾驶系统等。当然,距离实现像人那样能识别和理解任何环境并完成自主导航这一目标,还有一些未完成的工作。

1.1.1 计算机视觉原理

计算机视觉是使用计算机及相关设备对生物视觉进行的一种模拟。它的主要任务是,用各种成像系统代替视觉器官,利用计算机从二维图像中提取景物二维或三维的结构和属性描述并加以理解。计算机视觉是以视觉处理理论为核心,以图像处理、模式识别、生理学、心理学、计算机技术为基础的信息处理学科的一个重要分支。计算机视觉一方面热衷于模拟生物视觉功能,利用计算机技术对景物图像进行识别以得到对景物某些特性的解释,另一方面试图通过技术手段实现目标物体内部的可视化。显然,机器与生物之间有本质的不同,因此计算机视觉与生物视觉必然存在诸多不同之处,而这些不同之处也就决定了两者各自的优势与劣势。

数字摄影测量与计算机视觉在三维模型重建、图像匹配等问题上极其相似。虽然两者在许多方面存在共同点,但是出发点的不同最终导致它们的基本参数物理意义的设定上存在差异。如航空摄影测量中所有摄影机的空间位置坐标与图像的坐标系都是相对于该空间坐标系而言的;而计算机视觉的一个重要研究内容是怎样用计算机模拟人的眼睛以实现机器人视觉,它是以摄像机(眼睛)坐标系为准,定义的平移量是空间坐标系相对于摄像机坐标系的平移量。计算机视觉这种基于图像的重构技术,具有非接触性、实施较为简单等优点。传统的基于图像的三维重建已经形成了一套完整、严密的体系,而且有成熟的软件支持,这对计算机视觉的三维重建有很好的借鉴意义。除此之外,计算机视觉还研究物理结构的可视化,如医学领域常用的CT、MRI、PET成像模式等。

需要补充的一点是,在计算机视觉系统中,计算机扮演了人脑的角色,但显然计算机视觉与人类视觉处理问题的原理与方法不同。计算机视觉根据计算机自身的特点来进行视觉信息的处理,人类视觉处理机制的研究仅给计算机视觉的研究提供启发。在此启发下,用计算机信息处理的方法研究人类视觉的机理,建立模拟人类视觉的计算理论,这方面的研究被称为计算视觉。

1.1.2 计算机视觉研究范畴

实现计算机视觉的最终研究目标的过程是一个漫长的过程,且它需要建立在不少学科的研究成果的基础之上,包括数字图像处理,模式识别,图像识别、理解,空间形状的描述,景物分析,认知等。换句话说,上述研究内容也包含在计算机视觉的研究范畴内。

1. 数字图像处理

图像是视觉处理的对象,同时也是图像处理技术的研究内容,因此,计算机视觉与数字图像处理技术紧密相关。数字图像处理的发展过程可以分为传统的图像处理和在此

基础上发展的模式识别技术、图像理解。

数字图像处理技术是借助计算机对图像进行分析,以达到所需结果的技术。目前的数字图像主要是指用摄像机、扫描仪等设备得到的一个二维数组,该数组的元素称为像素,其数值表示图像灰度值。数字图像处理技术一般包括图像压缩,图像复原,通过处理使输出图像有较高的信噪比,通过增强处理突出图像的细节等。在计算机视觉研究中,经常会利用数字图像处理技术对图像进行预处理,包括图像压缩、图像增强、图像复原及特征抽取等。

2. 人工智能

隶属于人工智能领域的模式识别和机器学习与计算机视觉有着重要联系,因此,计算机视觉也时常被看作是人工智能与计算机科学的一个分支。模式识别技术根据从图像抽取的特征如统计特性或结构信息等对图像进行分类,典型的应用如邮件分拣、文字识别、指纹识别等。

计算机视觉、图像处理和机器视觉遇到的共同的经典问题是如何在一组图像数据中找出其中包含的某个特定内容,如图像特征或运动状态等。这一问题通常可以通过机器自动解决,但是到目前为止,在任意环境中识别任意物体还未能实现。目前的技术只能很好地解决特定目标的识别,如印刷或手写文件识别,车辆识别,有限数量的几何图形、表情、人脸识别等,而且对环境、光照、背景和姿态等有特定的要求。

识别包含鉴别和监测双重含义。鉴别是指识别、辨认单一物体本身,如某一人脸的识别,某一指纹的识别等。监测是指从图像中发现特定内容,它往往是指通过简单的图像处理发现图像中的特殊区域。

在计算机视觉中,模式识别技术经常用于对图像的某些部分(如分割区域)进行识别和分类。图像理解可完成描述和解释图像所代表的景物内容或意义这一任务。进行图像理解时,除了需要图像处理技术以外,还需要与景物成像的物理规律相关的知识与景物内容有关的知识。

3. 机器视觉

视觉是人类最主要的感知方式,可为人类提供大量信息,这就使得人类对生物视觉系统的研究倾注了极大的热情。而且自从计算机出现以后,研究者也试图将生物视觉赋予机器。生物视觉是一个非常复杂的感知系统,目前积累的生物视觉知识还存在很多不足,因此将生物视觉所有功能赋予机器还需要一个较为漫长的研究过程。

机器视觉所面临的任务是清晰的,需要的视觉功能是有限的,环境也是可控的,如机器人检测易拉罐内外侧轮廓圆度、边缘凹陷、表面划伤及压痕、拉环轮廓、表面字符及图案(见图1-1)。对于该问题,机器视觉检测效果在很大程度上取决于光照,通过对复合式光源的组合控制,能获得轮廓清晰、色彩对比鲜明的图像,并能迅速地与设定值比较,从而迅速找出那些有边缘凹陷、表面划伤、拉环变形、印刷字符及图案不清晰等缺陷的不良产品。

计算机视觉和机器视觉存在明显的重叠。计算机视觉涉及的技术是许多其他领域中自动化图像分析的核心技术。机器视觉涵盖了自动图像分析技术与其他方法和技术,可

用于自动检测及机器人指导下的工业生产过程。在许多计算机视觉生产中，基于学习的方法解决特定任务占据重要的位置。

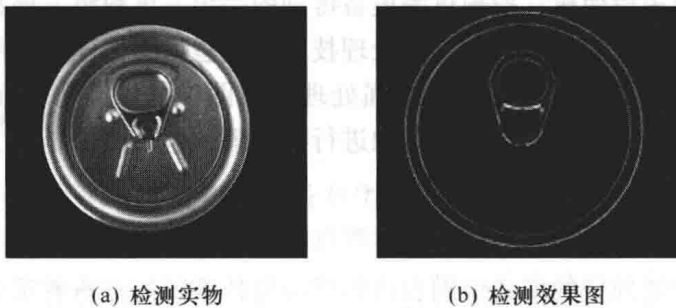


图 1-1 易拉罐检测图

机器视觉是计算机视觉在工业中的具体应用，利用视觉信息可加快制造工序的顺利完成，如在利用图像信息进行质量控制的过程中利用视觉信息来寻找产品缺陷的自动检测。机器视觉也被大量用于机器人臂位置测量、光学分拣等。

4. 成像

生成数字图像的信息来源于图像感知器，包括各类光敏摄像机、X 射线断层摄影仪、雷达、超声波接收器、遥感设备等，这些都可以作为图像感知器。数字图像成像包括二维图像、三维图组及图像序列，这取决于不同的感知器。数字图像的像素值对应于光在单一光谱或多光谱段上的强度，但也可对应于相关的各种物理数据，如声波、电磁波或核磁共振的深度、吸收度或反射度。前者对应于可见目标的成像系统，而后者对应于不可见目标的成像系统，计算机视觉研究的内容包含可见目标和不可见目标的成像。

5. 数学逆问题

在建立数学模型后，研究课题都可以被当作纯粹的数学问题来研究。计算机视觉中很多问题的数学模型在理论上与统计学、最优化理论及数学逆问题的理论知识相关。

众多问题中存在的逆问题并非是什么新问题，它只是新归纳出的一类问题。在一个通用的数学表达式 $G(z) = d$ 中，给定 z 求 d 为正问题。在解正问题时，计算 $G(z)$ 可能涉及积分、微分或矩阵运算，甚至也可能没有显式解析公式为 $G(z)$ 的计算提供一个算法。根据给定的 z 和 d 的样本来确定 G ，这类问题可归为模型识别问题。逆问题是相对正问题而言的，即给定 d 求 z 就是所谓的数学逆问题。

1) 图像处理中的逆问题

在形成图像及录制图像的过程中，目标物体（原始物体）图像将不可避免地发生退化，即图像所表示的目标物体在某种程度上与原始物体存在一定的差别。所谓形成图像的过程一般是指通过各种光学设备将目标物体映射到胶片或传感器上的过程；而录制图像的过程一般是指生成可观测和可处理图像的过程。这是目标物体变为图像的两个处理阶段。图像形成过程中发生退化常意味着图像模糊和目标物体的频带受限，而图像录制过程中发生退化常意味着噪声的加入。在此意义上可以将模糊视为一个确定过程，并

且在大多数情况下可以用一个足够准确的数学模型去描述它；而噪声是一个统计学的过程，所以它对具体图像的影响是不确定的，只能够用一个统计特性的过程来模拟它。在上述的基本依据条件下，当试图将一个退化图像恢复为一个较清晰图像时，首先应该了解退化要因，然后根据退化要因建立相应的数学模型，最后在此基础上进行图像恢复。显然，不同的退化需要不同的数学模型。

点扩散函数 (point spread function, PSF) 是一点光源经光学系统后所形成的图像分布函数，它的傅里叶变换是光学系统的传递函数。点扩散函数在空域中表征光学系统的特性，传递函数在频域中表征光学系统的特性。点扩散函数在图像处理中起着重要的作用，如若知道点扩散函数和获取的图像，就可得到原始图像，这就是图像复原处理的基本思路。

在数学表达式 $G(z) = d$ 中，若 G 表示为一个线性积分算子，则可以有

$$\int_a^b g(s, x) z(x) dx = d(s) \quad (1-1)$$

该方程可以称为是 Fredholm 方程的第一类积分方程 (integral equation of the first kind, IFK)。形如 $z(t) = f(t) + \int g(t, \tau, z(\tau)) d\tau$ 的方程称为 Fredholm 积分方程，一些数学物理问题均可以归结为解这种积分方程的问题。当未知函数只出现在积分符号内时，如 $f(t) = \int_a^b g(t, \tau) z(\tau) d\tau$ ，为第一类积分方程；当未知函数不仅仅只出现在积分符号内时，如 $f(t) = z(t) + \lambda \int_a^b g(t, \tau) z(\tau) d\tau$ ，为第二类积分方程。

对于任何线性系统，必须选择一个合适的二元函数 $g(s, x)$ 使式(1-1)成立，其中函数 $g(s, x)$ 称为核。对于一个图像系统， $z(\tau)$ 表示目标物体图像（或称原始物体图像）， $d(s)$ 为由目标物体生成的可观察的图像，利用失真图像获得目标物体图像，即从 $d(s)$ 求 $z(\tau)$ ，就是一个典型的逆问题。

若用一个一元函数来刻画线性系统，则可以简化式(1-1)，此时加入一不变约束条件，即式(1-1)可以表示为

$$\int_a^b g(s, x) z(x - \vartheta) dx = d(s - \vartheta)$$

进行变量替换，将 s 和 x 同时加上 ϑ ，得到

$$\int_a^b g(s + \vartheta, x + \vartheta) z(x) dx = d(s)$$

若 $g(s + \vartheta, x + \vartheta) = g(s, x)$ 对于所有 ϑ 均为真，这就意味着需满足当两个变量增加同样的量 ϑ 时 $g(s, x)$ 的值不变，即需满足只要 s 和 x 的差不变则 $g(s, x)$ 的函数值不变这一条件。为了与前面的表示一致，我们不妨将 $g'(s - x)$ 用 $g(s, x)$ 表示。这样，我们可以定义一个以 s 与 x 之差为自变量的函数 $g'(s - x) = g(s, x)$ 。在很多情况下式(1-1)中的核可以写为显式且依赖于 $(s - x)$ ，此时产生一个卷积方程

$$\int_{-\infty}^{\infty} g(s - x) z(x) dx = d(s) \quad (1-2)$$

这里积分区间扩展为从负无穷到正无穷，为了适应任何区间，只需对感兴趣的区间

之外令 $g(s-x)=0$ 就可以了。当一个问题具有式(1-2)的形式时，通过 $z(x)$ 求 $d(s)$ 的运算称为卷积，而通过 $d(s)$ 求 $z(x)$ 的逆运算称为反卷积，这在图像处理中有很好的用途。

2) 视觉自身的“视”与“觉”的逆问题

视觉是一个典型的逆问题，其对应的正问题是成像过程，图 1-2 为视觉过程模型图。这里举日常生活中常见的一个例子，当我们在非洲野生动物园看到一只动物且该动物几乎被树林和灌木丛遮挡时，进入眼帘的不是动物的全貌，也就是目标物体成像并非完整。此时，人们将依据观察到的动物身上的条纹特征及所处场景等信息，结合自身的经验知识，可以对观察到的动物的种类作出判断。目前让人们依然感到困惑的一个问题是，用计算机来处理这类逆问题依然是非常困难的。如同所有的逆问题一样，它是病态的，这也揭示了计算机视觉的研究还有一段很艰辛、很漫长的路要走的原因。

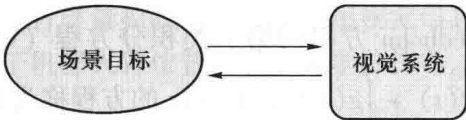


图 1-2 视觉过程模型图

6. 神经生物学

21 世纪的自然科学重心将落在生命科学上，而神经生物学和分子生物学又将是生命科学研究中的两个最重要的领域。神经生物学是一门研究人体神经系统的结构、功能、发育、衰老、遗传等规律，以及研究疾病状态下神经系统的变化过程和机制的学科。神经生物学的内容丰富，研究进展快，但距离解开智力形成之谜、老年痴呆症治疗之谜等还有一定的距离。神经生物学中的生物视觉系统研究对计算机视觉的发展具有重要意义。通过对各种动物的眼睛、神经元及与视觉刺激相关的脑部组织进行广泛研究，人类得出了一些有关“天然的”视觉系统如何运作的描述。为此，人们试图建立人工视觉系统，使之在不同复杂程度上模拟生物的视觉运作，这也是计算机视觉研究中的一个子研究领域。同时，生物机制也对计算机视觉领域中一些基于机器学习的方法的研究提供有益的参考价值。目前的深度学习算法能够帮助我们做一些视觉检测方面的工作，而这一思路来源于对大脑认识的启发。

1.2 计算机视觉检测

计算机视觉检测就是将数字摄影测量与计算机视觉技术有机结合，充分发挥计算机视觉在检测方面的优势而形成的一个跨学科的科学领域。计算机视觉检测涉及数学逆问题及计算机成像和信号处理中的滤波、复原、识别、特征提取等技术，它在计算机辅助检测、医疗诊断等众多领域中有广泛应用。

1.2.1 概述

计算机视觉检测的基础是成像技术。在成像方面,计算机视觉主要包含以下两方面的内容。①直接成像方法:直接通过摄像设备获取目标物体图像后,采用合适的计算机算法完成对图像的识别和理解工作,也可以将该方法理解为是一种模拟生物视觉功能的方法。②间接成像方法:对不能直接观察到的目标物体,采用物理、化学、生物等技术方法并通过计算机可视化方法生成图像,然后对图像进行检测,如常用的医疗设备CT等。

在某些应用场合中,视觉检测的正确性远比精确性重要。要让计算机解释图像的能力与三岁大的孩子一样,目前还难以实现,但是计算机视觉检测的精确性相对于人类的视觉检测来说要强。目前,计算机视觉无疑在诸多功能领域中还未达到人类视觉的智能程度,但基于计算机视觉的固有特性,它在检测方面有望弥补人类视觉的不足。具体来说,在对视觉检测结果的正确性要求高于精确性要求的场合,计算机视觉检测并非优于人类的视觉检测;相反,在对视觉检测结果的精确性要求高于正确性要求的场合,计算机视觉检测的优势会显现出来。一般情况下,人通过视觉本身并不能直接完成对目标物体的精确测量,通常要借助于测量工具来实现,如借助激光测距仪测量物体的长度、距离等,而计算机视觉检测则可以简单地以像素为计量单位对目标物体进行精确测量。另外,对人类无法看到的物体内部来说,人类的视觉也就无从谈起了,而计算机视觉可以借助物理、化学、生物等技术方法,通过实现对目标物体的可视化来进行检测。到目前为止,计算机视觉检测已经在医学诊断、工业检测、高空侦察、制导、航天、交通等方面得到了广泛的应用。

1.2.2 计算机视觉检测应用实例

在精确定量感知、危险场景感知、不可见物体感知的应用中,计算机视觉检测更突显其优越性。基于双目视觉的测量系统已经广泛应用于物体的尺寸检测、地形测绘、精密零部件或产品的三维外形检测等领域中。

1. 医学领域

医学领域常用的CT(computed tomography,计算机断层成像)、MRI(magnetic resonance imaging,核磁共振成像)、PET(positron emission tomography,正电子发射型计算机断层显像)的成像模式都属于间接成像。身体内部器官及大脑的三维体积测定可视化等已经成为日常病人诊断和护理程序中标准的组成部分。病理图像能进行自动测定、自动识别的前提是图像分割、去噪和增强等处理技术的成功开发。这里也包含了光学分子影像成像中的BLT(bioluminescent tomography,自发荧光断层成像)、FMT(fluorescence molecular tomography,激发荧光断层成像)^[1]。

1) 传统的影像学技术

传统的影像学技术,包括MRI、CT、FMT等,都是基于物理学(如光吸收、散射等)及生理学(如血流等)特点的图像可视化检查手段。