

知网个人查重服务报告单 (全文标明引文)

报告编号:BC202308041636143437186199

检测时间:2023-08-04 16:36:14

篇名: 计算机视觉教材

作者: 张建

检测类型: 学术出版

比对截止日期: 2023-08-04

检测结果

去除本人文献复制比: 34.6% 去除引用文献复制比: 34.4% 总文字复制比: 34.6%
单篇最大文字复制比: 4.4% (基于相机扰动校正的桥梁结构变形测量方法与应用)

重复字符数: [58624] 单篇最大重复字符数: [7405] 总字符数: [169638]

18.1%(1564)	18.1%(1564)	计算机视觉教材_第1部分 (总8621字)
11%(703)	11%(703)	计算机视觉教材_第2部分 (总6376字)
56.6%(6189)	56.6%(6189)	计算机视觉教材_第3部分 (总10935字)
50.3%(4466)	50.3%(4466)	计算机视觉教材_第4部分 (总8870字)
27.6%(2756)	27.6%(2756)	计算机视觉教材_第5部分 (总9996字)
40.4%(1620)	40.4%(1620)	计算机视觉教材_第6部分 (总4005字)
34.2%(3379)	34.2%(3379)	计算机视觉教材_第7部分 (总9883字)
65.3%(1919)	65.3%(1919)	计算机视觉教材_第8部分 (总2937字)
90.2%(1156)	90.2%(1156)	计算机视觉教材_第9部分 (总1282字)
56.5%(5646)	56.5%(5646)	计算机视觉教材_第10部分 (总10000字)
45.4%(4583)	45.4%(4583)	计算机视觉教材_第11部分 (总10103字)
38.6%(2843)	38.6%(2843)	计算机视觉教材_第12部分 (总7360字)
5%(462)	5%(462)	计算机视觉教材_第13部分 (总9223字)
5.5%(626)	5.5%(626)	计算机视觉教材_第14部分 (总11411字)
44.1%(4042)	44.1%(4042)	计算机视觉教材_第15部分 (总9158字)
36.8%(3749)	36.8%(3749)	计算机视觉教材_第16部分 (总10181字)
56.5%(1652)	56.5%(1652)	计算机视觉教材_第17部分 (总2922字)
56.1%(5171)	56.1%(5171)	计算机视觉教材_第18部分 (总9218字)
14.9%(1332)	14.9%(1332)	计算机视觉教材_第19部分 (总8925字)
45.9%(2042)	45.9%(2042)	计算机视觉教材_第20部分 (总4450字)
56.7%(1675)	56.7%(1675)	计算机视觉教材_第21部分 (总2955字)
3.8%(193)	3.8%(193)	计算机视觉教材_第22部分 (总5050字)
0%(0)	0%(0)	计算机视觉教材_第23部分 (总2606字)
65.8%(462)	65.8%(462)	计算机视觉教材_第24部分 (总702字)
21.4%(236)	21.4%(236)	计算机视觉教材_第25部分 (总1103字)
24.8%(158)	24.8%(158)	计算机视觉教材_第26部分 (总638字)
0%(0)	0%(0)	计算机视觉教材_第27部分 (总728字)

1. 计算机视觉教材_第1部分		总字符数: 8621
相似文献列表		
去除本人文献复制比: 18.1%(1564) 去除引用文献复制比: 15.8%(1364) 文字复制比: 18.1%(1564)		
1	<u>“机器视觉技术及应用”专题前言</u> 于起峰;卢荣胜;刘晓利;王程;李璋; - 《激光与光电子学进展》- 2022-07-25	3.4% (289) 是否引证: 否
2	<u>基于深度学习的计算机视觉研究新进展</u> 卢宏涛;罗沐昆; - 《数据采集与处理》- 2022-03-15	2.8% (244) 是否引证: 否
3	<u>机器视觉技术在食用菌产业中的研究应用进展与展望</u> 陈学东;马聪;张建华; - 《宁夏农林科技》- 2022-08-20	2.1% (177) 是否引证: 否
4	<u>摄像测量学简介与展望</u> 于起峰;尚洋; - 《科技导报》- 2008-12-28	2.0% (170) 是否引证: 是
5	<u>基于相机扰动校正的桥梁结构变形测量方法与应用</u> 于姗姗(导师: 张建) - 《东南大学博士论文》- 2021-04-01	2.0% (169) 是否引证: 否
6	<u>正颌外科手术机器人视觉导航与轨迹规划</u> 李倩倩(导师: 宋锐;马昕) - 《山东大学博士论文》- 2020-09-10	1.7% (145) 是否引证: 否
7	<u>机器视觉与神经网络在智能测试中的应用与展望</u> 王忠革; - 《第十九届中国航空测控技术年会论文集》- 2022-11-25	1.5% (133) 是否引证: 否
8	<u>从摄影测量到计算机视觉</u> 龚健雅;季顺平; - 《武汉大学学报(信息科学版)》- 2017-11-05	1.4% (125) 是否引证: 否
9	<u>人工智能时代测绘遥感技术的发展机遇与挑战</u> 龚健雅; - 《武汉大学学报(信息科学版)》- 2018-12-05	1.4% (120) 是否引证: 否
10	<u>人工智能时代测绘遥感技术的发展机遇与挑战</u> - 道客巴巴 - 《互联网文档资源 (https://www.doc88.co)》- 2020	1.2% (103) 是否引证: 否
11	<u>《摄影测量学》教学课程新思路研究与设计</u> 张荣春;衣雪峰;李浩;岳建平;郭献涛; - 《现代测绘》- 2019-09-25	1.1% (98) 是否引证: 否
12	<u>人工智能导向下的视觉呈现设计</u> 徐心宇(导师: 薛澄岐) - 《东南大学硕士论文》- 2019-04-01	1.0% (89) 是否引证: 否
13	<u>计算机视觉技术在果园除草机上的应用研究</u> 张雨滋; - 《农机化研究》- 2017-08-30	1.0% (85) 是否引证: 否
14	<u>基于深度学习的评论打分预测方法综述</u> 张禹;赵妍妍;司峥鸣; - 《信息技术》- 2021-10-25	0.8% (68) 是否引证: 否
15	<u>基于深度学习的视觉目标检测技术综述</u> 曹家乐;李亚利;孙汉卿;谢今;黄凯奇;庞彦伟; - 《中国图象图形学报》- 2022-06-16	0.7% (58) 是否引证: 否
16	<u>基于改进型编解码网络的眼底图像视网膜动脉硬化检测</u> 李迪(导师: 吴骏;苏龙) - 《天津工业大学硕士论文》- 2019-12-01	0.4% (34) 是否引证: 否
17	<u>基于计算机视觉的图像反求技术的研究与设计</u> 周涛; - 《科技资讯》- 2006-02-03	0.4% (32) 是否引证: 否
18	<u>气液两相流中连续液相特征参数测量与分析</u> 杨柳湘子(导师: 薛婷) - 《天津大学硕士论文》- 2015-11-01	0.3% (29) 是否引证: 否
原文内容		

目录

第1章绪论.....3

1.1计算机视觉的定义及发展历程.....3

1.2计算机视觉、机器视觉、摄影测量学的关系.....6

1.3深度学习与计算机视觉的关系.....7

1.4 计算机视觉土木工程应用及发展趋势	8
1.5 小结	12
习题	13
第2章 相机成像原理与标定	14
2.1 数字相机	14
2.2 空间坐标系变换	19
2.3 成像模型	22
2.4 相机标定	24
习题	31
第3章 数字图像预处理	32
3.1 数字图像基础	32
3.2 图像滤波	41
3.3 图像频域处理	46
3.4 边缘检测	51
习题	65
第4章 图像特征提取与匹配	67
4.1 概述	67
4.2 全局图像特征	68
4.3 角点检测	72
4.4 斑点检测	79
4.5 特征描述子	86
4.6 图像匹配方法	89
习题	100
第5章 位移测量	101
5.1 结构位移测量概述	101
5.2 数字图像相关与位移测量基本原理	103
5.3 比例系数确定方法	105
5.4 基于SURF特征匹配的位移测量	112
5.5 结构位移测量精度测试	114
5.6 位移测量精度影响因素	116
第6章 深度学习与图形处理	119
6.1 深度人工智能、机器学习与深度学习	119
6.2 深度学习与图像处理	120
6.2 深度学习网络基本组件	121
6.3 深度学习网络的迭代优化	124
6.4 深度学习的运算	126
6.5 计算机视觉中常见深度学习模型	128
6.6 深度学习的实现框架	134
习题	136
第7章 目标识别与分割	139
7.1 概述	139
7.2 目标识别原理	139
7.3 目标分类	145
7.4 目标检测	149
7.5 图像分割	149
7.6 应用拓展	155
习题	155
第8章 立体视觉与三维重建	156
8.1 三维重建方法概述	158
8.2 双目立体视觉	162
8.3 运动恢复结构	171
8.4 点云处理	176
习题	184
第9章 结构位移视觉监测应用案例	222
11.1 泗安塘大桥位移监测应用	222
11.2 南沙大桥位移监测应用	225
11.3 基于无人机的结构动位移测量	229
习题	233
第10章 结构病害智能检测应用案例	202
10.1 结构智能检测概述	202
10.2 裂缝可视化识别方法	203

10.3 智能检测设备.....	214
习题.....	221
第11章结构三维几何尺寸测量应用案例.....	185
9.1 基于双目视觉的装配式桥梁预制墩柱安装位置姿态监测.....	185
9.4 基于三维激光扫描的预制墩柱安装位置姿态监测.....	198
习题.....	201

第1章绪论

1.1计算机视觉的定义及发展历程

计算机视觉是一门研究让机器具备视觉能力的学科，即机器能对周围环境和其中的对象进行可视化分析。英国机器视觉协会（BMVA）将机器视觉定义为“对单张图像或一系列图像的有用信息进行自动提取、分析和理解”。计算机与人类的视觉系统在某些方面有相似之处，见图1-1。它通过使用各种成像系统代替人类的视觉器官，我们可以把获取的图像或视频等感知数据输入到计算机机中进行处理和分析，然后计算机输出对图像或视频中重要信息分类、分割、检测、跟踪等识别结果，进而对环境

和事件进行解释。

图1-1 计算机与人类视觉

图1-2 计算机视觉发展历史

在过去的六十年中，计算机领域的专家们一直在努力找到一种让机器能够理解视觉数据的方法。图像处理的起源可以追溯到1957年，当时美国国家标准局的Russell Kirsch和他的团队开发了一台可编程的计算机，并搭建了一台旋转鼓式扫描仪。他们利用这个设备创建了人类历史上的第一张数字图像，照片中是Kirsch三个月大儿子Walden的头肩部分。尽管这幅黑白照片尺寸只有176x176像素，与今天的数码照片相差甚远，但它奠定了数字图像技术后来的快速发展基础，因此Kirsch被誉为“像素之父”。

图1-3 人类史上第一张数字图像[1]和神经生物学家开展的动物神经元测试[2]

1959年，神经生物学家David Hubel和Torsten Wiesel发表了一篇有关动物视觉神经元活动相关的论文【】，这是在计算机视觉领域中最具影响力的论文之一。他们使用一种电生理技术记录了动物神经元的活动。他们的实验发现，当他们插入新的图像幻灯片到投影仪中，猫的一个神经元会发射信号。这表明，神经元对玻璃幻灯片锐利边缘的阴影所产生的线条运动是兴奋的。通过这些实验，研究人员确定了视觉皮层中包含处理简单和复杂任务的神经元。他们发现视觉信号在大脑传输中会被分解，而视觉处理的起点通常是定向边缘等简单结构。这意味着当我们看到一个物体时，首先由负责低级视觉处理的神经元识别边缘或确定边缘的方向。然后，复杂神经元会接收颜色、基本的二维形状。最后，更复杂的神经元进行高级视觉处理，如识别复杂特征或整合信息。这种逐层处理视觉信息的模式成为了视觉神经生理学研究的框架，也为计算机视觉学科奠定了基础。Hubel和Wiesel因为在视觉系统和视觉处理方面的突破性发现，于1981年获得诺贝尔生理学或医学奖。

上世纪60年代是人工智能学科的一个重要阶段。当时，许多研究人员对该领域的未来充满希望。其中，麻省理工学院的人工智能实验室的教授Seymour Papert发起了一个关于视觉的暑期项目，试图解决图像自动分割的问题【】。虽然这个项目并未取得成功，但它被认为是计算机视觉作为一门科学诞生的时刻。

1965年，Lawrence Roberts在MIT的博士论文《三维固体的机器感知》中描述了如何从线条图的二维照片中推导出三维几何信息，以及计算机如何从单张二维照片创建三维模型【】。书中讨论了相机投影变换、透视效应和深度感知的规则、假设等内容，是计算机视觉的先导之一。

(a) Lawrence Roberts (b)原始图 (c) 图像梯度 (d) 视角变化后渲染生成的新图[3]

图1-4 三维固体的机器感知

70年代后，人工智能领域进入了一个寒冬期。尽管当时的研究进展缓慢，但在80年代末到90年代初期仍开发出了基础神经网络。1979年，一位日本计算机科学家Kunihiko Fukushima，深受视觉神经分层处理模式的启发，构建了一个用于模式识别的神经网络模型Neocognitron【】，这是卷积神经网络的前身。1989年，LeCun结合反向传播算法与权值共享的卷积神经层，发明了卷积神经网络LeNet，并首次将卷积神经网络成功应用到美国邮局的手写字符识别系统中【】。

同一时期，关于视觉计算理论也在逐步形成，标志性事件是1982年David Marr出版的一本重要书籍《VISION》【】。Marr也被誉为现代计算机视觉之父，他提出了研究和理解视觉感知的通用框架，包含计算理论、表达和算法、算法实现三层次。计算理论着重于挖掘成像物理场景的内在属性来进行视觉计算，即利用二维图像恢复三维物体表面形状；物体表达分为物体坐标系三维形状表达，以及观测者坐标系下的2.5维表达；算法研究主要关注如何通过不同的模块（例如立体视觉、运动分析等）来获取物体的表达。这个过程可以分为三个阶段：从原始图像到2.5维图像，再到最终的三维模型表示。Marr的工作在当时是开创性的，但它非常抽象，没有涉及任何关于视觉系统中数学建模和学习过程的内容。

在20世纪90年代，多视几何理论逐渐成熟，为分层三维重建理论奠定了基础。多视几何实质上研究的是在图像中，当观察角度变化时，图像中相应点之间的关系以及空间中点和其在图像中的投影点之间的约束关系的理论和计算方法。而分层三维重建是指一种逐步分阶段进行的优化方法，用于从多张二维图像中恢复出具有三维结构的欧几里德空间。分层三维重建在计算机视觉领域非常重要，因为许多三维视觉应用，如三维地图和全景街景展示等，都是基于这个理论来实现的。

21世纪初，研究热点从三维重建转向基于特征的物体识别，如David Lowe于2004年提出的SIFT方法【】，它是一种是人工设计的特征方法，利用对旋转、位置和光照变化不敏感的局部特征进行图像配准和物体识别。2001年，Paul Viola和Michael Jones提出了第一个实时人脸检测框架【】，该算法在当时非常高效且可靠。虽然该算法是基于机器学习，但它的工作方式类似于深度学习，在处理图像时学习可以帮助人脸定位的特征。计算机视觉的重要突破发生在2012年，当时深度神经网络AlexNet在当年视觉识别挑战赛（ImageNet）中获胜。尽管卷积网络在1989年就被提出（LeNet），但AlexNet的成功不仅依赖于算法和模型本身，还依赖于高质量的数据集（如ImageNet、Pascal）和高性能的计算硬件（如NVIDIA GPU）【】。这些因素都对其成功起到了关键作用。自此之后，深度学习技术引领了计算机视觉领域的新浪潮。通过使用更深的网络结构，大规模数据集和强大的计算硬件，深度学习在图像分类、目标检测、语义分割等任务上取得了显著的性能提升，并逐渐在自动驾驶、医学图像分析、人机交互等应用领域发挥着关键作用。

我们在本书的后续章节中提到的数字图像相关(Digital image correlation, DIC)，它与计算机视觉的发展历史相互独立

，但存在一些重叠的技术和方法。DIC是实验力学领域的一项重要技术，它主要用于在实验室环境中观测物体的变形、形状和运动。通过对物体表面的图像进行比较和分析，DIC技术可以精确地测量物体在受力或加载情况下的变形和形状变化，从而为力学实验提供有价值的数据和信息。DIC算法是南卡罗来纳大学的Peters和Ranson在1982年首次提出[4]，他们使用了一种交叉相关函数与零阶形状函数相结合的图像相关性表示方法，追踪了从小型铝样本上激光散斑图案的刚体运动。接着，Sutton及其同事继续改进原始的DIC技术，使其满足工业场景中表面变形测量的需求，而DIC技术的理论基础和基本框架在2000年之前基本上已经确立[5]，商业测量系统的推广加速了DIC方法的应用领域，除了经典的实验力学、断裂力学、疲劳等测试外，DIC也可应用于监测生物组织、电子元件和基础设施受到外部载荷下的变形和应变发展。DIC技术与计算机视觉中的图像匹配在工作原理上有较高的相似度，它们都是通过检测和匹配两幅图像中的相同特征，来实现图像几何对齐。由于实验力学检测对象通常为连续体且测量精度要求高，DIC技术通常使用人工散斑图案和基于区域的匹配方法。

从工业界的角度上看，计算机视觉是目前人工智能领域中应用最广泛的技术之一。它的技术研究与产业应用正在迅速发展。我们日常生活中熟悉的人脸识别、AI修图、虚拟现实等都与计算机视觉息息相关，下面列举了其在制造业、自动驾驶、医疗健康、农业等方面的应用。

图1-5 计算机视觉应用场景（依次为制作业[6]、自动驾驶[7]、医疗健康[8]、智慧农业[9]）

制造业：计算机视觉在制造业中扮演着重要的角色，常被用于产品和质量检查、结构监测以及损伤跟踪等方面。相较于肉眼，摄像头能够更敏锐地察觉生产产品的微小缺陷，并且高效地完成任务。

随着时间的推移，工业机器人在工业生产中的应用范围越来越广泛。在工件的分拣、码垛、零件的焊接以及工件的装配等行业，工业机器人逐渐取代了人工进行各种作业。将机器视觉技术与自动化生产线中的工件分拣系统相结合，使用工业智能相机代替多种繁琐的传感器，可以大大简化软件开发的难度，提高对工件或其他物品的识别速度和精度，拥有广阔的应用前景。

自动驾驶：自动驾驶技术利用多组摄像头收集车辆周围环境的信息，包括障碍物、道路标志、交通信号灯等。通过计算机视觉算法分析图像，获取全面的环境信息，例如道路边缘定位、路标识别以及其他车辆的识别。这些信息有助于判断路况并做出决策，从而确保行车的安全。

自动驾驶是人工智能领域的重要方向，具有良好的发展前景。实时准确的目标检测与识别是保证自动驾驶汽车安全稳定运行的基础和关键。在目标检测技术的发展历程中，有许多不同的算法可供选择。其中，将YOLO算法应用于自动驾驶车辆目标的实时监测中能够取得良好的结果，它能够快速实时地识别出车辆、行人、交通标志、灯光、车道线等要素。

医疗健康：在医疗健康领域，计算机视觉技术被应用于医学影像（如X光、MRI扫描）的辅助诊断，能够自动识别出特定疾病的迹象，提高诊断的准确性和效率。举例来说，外科手术机器人可以通过光纤摄像机捕获的实时图像，自动检测异常和脆弱区域，并协助建立精细的患者档案，减轻外科医生的压力。

手术机器人的发展已经有近四十年的历史。然而，由于技术、伦理、制度等多种因素的限制，手术机器人的自主作业能力以及在临床应用领域的推广一直发展缓慢。另一方面，在医学临床领域，口腔颌面外科一直是数字化外科技术应用的先驱和领军力量，尤其是在手术方案较为量化的正外科领域，已经广泛实现了计算机辅助设计和手术导航等数字化技术在临床上的推广应用。

智慧农业：智慧农业利用计算机视觉技术结合无人机航拍图像，对田间条件、土壤湿度和作物健康信息进行评估，并预测作物产量。机器视觉也应用于农产品智能分拣机器，能够实时检测和分级农产品的大小、形状、颜色、表面缺陷等外观品质指标，实现自动高效的分拣过程。

在果园中，杂草与树木竞争营养和水分，同时也成为病虫害滋生和栖息的场所，导致果园产量减少并降低果园产品的质量。化学除草方法高效且成本低，有利于保持果园的土壤环境和微生物群落。然而，传统的大面积喷洒方式存在药液浪费和残留的问题。相反，变量喷药根据杂草的位置和密度，调节除草剂的喷洒时机和剂量，能够减少除草剂的使用量。实现变量喷药的前提是获取杂草的位置、分布和密度信息，目前常使用计算机视觉技术来实现。因此，我们可以利用计算机视觉技术来识别和获取果园中杂草的分布和密度信息，并通过处理器控制喷洒装置实现变量喷药，从而解决传统大面积喷洒带来的药液浪费和残留问题。

1.2计算机视觉、机器视觉、摄影测量学的关系

在数学、物理和生物等基础学科的支持下，以生物视觉为启发，涌现了以视觉为核心的不同领域与技术，如成像、视觉认知和机器人视觉等。这些领域之间互相影响与渗透，由此形成最常提到的三个研究领域：计算机视觉（Computer Vision, CV）、机器视觉（Machine Vision, MV）、摄影测量学（Photogrammetry）。虽然它们之间有一些区别，但也存在着一些联系。总得来说，计算机视觉为机器视觉提供理论基础，机器视觉是计算机视觉的工程实现，两者处理的核心都是图像，都被认为是人工智能的一个重要分支。另一方面，摄影测量学主要在地理测绘领域中发展，是地理学科的重要分支。虽然计算机视觉与摄影测量学的形成和发展相对独立，早期交流较少，但它们具有同样的理论基础。近十几年来，在理论、方法、算法及应用方面，这两个领域开始互相融合和交流[10]。

图1-6 研究领域关系图

机器视觉：研究如何让机器看见和理解目标场景，并快速作出决定的一种科学与技术。典型的机器视觉系统一般包含图像获取、数据传输、处理分析、决策及执行4个部分，具有“感-传-知-用”的系统特征。（图像获取是通过摄像机、传感器等设备获取图像或视频数据，将现实世界中的视觉信息转化为数字形式的图像数据。数据传输是将获取的图像数据传输到计算机或其他处理设备，以便进行后续的图像处理和分析。处理分析即利用图像处理和分析算法对图像数据进行处理，包括图像增强、滤波、分割、特征提取等，以获取图像中的关键信息和特征。决策执行即根据处理和分析得到的结果，进行决策和执行相应的操作，例如识别物体、跟踪运动、控制机器人等。）在成像技术、处理算法、算力平台和行业应用 4 个核心要素的驱动下，（成像技术：包括摄像机、传感器等设备的发展，如高分辨率图像传感器、深度传感器等，为机器视觉系统提供高质量和多样化的图像数据源。处理算法：涵盖图像处理、图像分析和机器学习等领域的算法和方法，如边缘检测、目标识别、图像分类、深度学习等，用于对图像进行特征提取、模式识别和决策推断。算力平台：随着计算机硬件的发展，特别是图形处理器（GPU）和人工智能芯片（如GPU和TPU），提供了更强大的计算能力和并行处理能力，加速了机器视觉算法的执行和实时性能。行业应用：机器视觉在工业视觉、自动驾驶、智能安防、医学影像、农业、无人机等领域有广泛的应用。它可以实现目标检测与跟踪、人脸识别、图像分割、虚拟现实等任务，为人们提供更智能、高效和安全的解决方案。）机器视觉目前已广泛应用于

工业视觉、图像解释、物体识别、虚拟现实、生命科学、智能制造、自动驾驶、人机交互等领域。例如在智能制造领域，机器视觉可以实现对产品的轮廓几何参数三维测量和表面缺陷检测；对字符、条码等的读取，对产品的分类和分组；对机械手的视觉引导与伺服控制，使其准确地执行抓取、焊接、装配、码垛和拆垛等。机器视觉可以被看作是计算机视觉理论在工厂自动化中的运用，计算机视觉理论为机器视觉的算法和方法提供了重要的基础，使机器能够从图像数据中提取特征、进行目标检测和识别，并进行高级的图像理解和分析。通过计算机视觉的技术和方法，机器能够模拟和实现人类的视觉感知和认知过程，从而实现目标场景的理解和决策。

摄影测量学：根据国际摄影测量与遥感协会的定义，它是利用非接触成像传感器和其他传感器，通过记录、量测、分析和表达来获取地球及其环境、以及其他物体的可靠性信息的科学、技术和工艺。与计算机视觉具有同样的理论基础，即小孔成像和双目视觉原理，同时处理的问题也很相似，但也有细微的区别。计算机视觉中理论的严密性要高于摄影测量，但摄影测量更强调测量精度，偏重于航测、卫测等传统遥感和测绘领域。如相机检校，摄影测量一般布设有高精度三维检校场，而计算机视觉常采用二维平面棋盘。如空中三角测量，摄影测量一般用严密的光束法区域网平差，而在计算机视觉中一般称为从运动恢复结构，除了用全局的光束法平差，也采用一些非全局解法，如增量式的局部平差、滤波等等。摄影测量学还涉及一系列重要的概念和技术，如相机校准、影像匹配、数字摄影测量、立体像对、数字高程模型等。相机校准是摄影测量中的关键步骤，通过准确确定相机的内外部参数，以保证摄影测量的精度。影像匹配是指将多个影像中相对应的特征点进行匹配，以实现立体视觉和三维测量。数字摄影测量利用数字技术对摄影测量数据进行数据的处理和计算，提高了数据的处理速度和精度。立体像对是指两个或多个在不同位置拍摄的影像对，通过对这些影像进行比对和分析，可以获取地物的三维坐标和形状信息。数字高程模型是基于摄影测量数据生成的地形模型，可以用于地形分析、地貌研究和工程设计等。摄影测量学在现代遥感技术中扮演着重要的角色。通过利用航空影像和卫星影像数据，摄影测量学可以获取地表地貌、植被覆盖、城市建筑物等信息。这些数据对于环境监测、土地利用规划、自然资源管理等领域具有重要意义。

图1-7 遥感基本原理

1.3 深度学习与计算机视觉的关系

传统计算机视觉系统的核心是从图像中提取人工设定的特征，如边缘检测、角点、颜色等。以**目标检测任务为例，通常采用人工设计的特征与浅层分类器的技术方案**，如尺度不变特征变换匹配算法(SIFT)、加速鲁棒特征算法(SURF)、二进制鲁棒独立基本特征(BRIEF)等特征提取方法，然后输入到支持向量机、AdaBoost等分类器中进行分类识别。虽然这些算法识别精准且计算复杂性低，但存在一些问题。首先这些检测算法泛化能力差，很难适应不同的检测场景。其次，为了获得准确的识别结果，这些算法依赖人工设定和参数调节。此外，由于提取到的特征质量直接影响目标检测效果，每个具体的问题领域都需要研究人员进行深入研究，构建适应性好的特征和检测器，这个过程既低效又昂贵。

自2010年以来，深度学习的出现和快速发展逐渐解决了传统计算机视觉面临的问题。深度学习属于机器学习的一个分支，使计算机能够通过经验进行学习，并以一种类似于人类思维层次结构的方式理解世界。简单来说，通过神经网络这个工具，从检测对象中提取特征信息，并通过大量的数据训练验证，使得输出结果接近我们期望的结构。相比于传统手工设计特征的方法，深度学习方法能够自动学习更具有区分力的深层特征，减少了人为设计特征造成的不完备性；同时，基于深度学习的方法将**特征提取和分类器学习统一在一个框架中，能够进行端到端的学习**。这意味着我们不再需要过多关注特征提取及分类器设计的细节，而更多地关注数据标注及增强、深度神经网络的结构设计和模型训练等方面。当然，深度学习也有一些缺点，如需要大数据支撑才能达到高精度，且算法复杂，需要强大的硬件支持。

目前，随着深度学习技术的迅猛发展和图形处理器等硬件计算设备的广泛普及，深度学习技术已经应用到计算机视觉的各个领域，如目标检测、图像分割、超分辨率重建及人脸识别等，并在图像搜索、自动驾驶、用户行为分析、文字识别、虚拟现实和激光雷达等产品中具有不可估量的商业价值和广阔的应用前景。

图像分类是为图像指定标签的任务。当图像中有只有一个类别，并且该类别在图像中清晰可见时，图像分类技术非常适用该类场景。例如，我们可以用图像分类来判断一张照片是白天或夜间拍摄的。此外，在交通领域，图像分类可用于检测汽车是否处于停车位，即停车位是否被占用。

目标检测适用于包含多个对象的图片，是一个重要的研究领域。例如，计算机视觉系统在机器人和自动驾驶汽车的应用中面对的是非常复杂的图像。毫无疑问，准确地定位和识别每一个物体是实现自动化至关重要的一步。

图像重建的目标是重建图像缺失或损坏部分。该任务可以被认为是一种没有客观评价的照片修复工具或变换器。虽然我们可以保证修复后的图像在可见属性上与原图尽可能匹配，但要求计算机重新创建没有参考的细节显然是不合理的。因此，图像重建系统存在一些限制，主要取决于可供学习的原始图像数量。有一种用于图像重建的模型被称为像素递归神经网络。它利用递归神经网络来预测图像在二维空间中缺失像素。图像重建应用的例子包括，修复老照片或者黑白电影。在自动驾驶汽车中，图像重建可以用来观察小型障碍物，比如车辆与被跟踪行人之间的路标。

计算机视觉的一个重要目标是能够识别一段时间内发生的事，即在图像或视频中跟踪特定对象。目标跟踪对几乎所有包含多个图像的计算机视觉系统都很重要。例如，在足球训练中，通过目标跟踪可以得到每个球员的时序位置信息，通过研究其体能和战术特点，进行科学的训练。

深度学习技术近年来在计算机视觉中的目标检测、图像分割、超分辨率等任务上都取得了卓越的成绩，充分证明了它的价值和潜力。然而深度学习领域仍然有不少难题无法解决，如对数据的依赖性强、模型难以在不同领域之间直接迁移与深度学习模型的可解释性不强等，如何攻克这些难题将也是计算机视觉相关任务下一阶段的发展方向。

1.4 计算机视觉土木工程应用及发展趋势

计算机视觉技术在基础设施的建造、运营、维护等各个阶段得到广泛应用。这些技术可用于检查、监测和评估基础设施的安全状况，并且能够解决传统技术所面临的耗时、费力的问题。众多学者在研究计算机视觉技术和土木工程领域的交叉融合，主要集中在下述的施工监控、病害检测、健康监测等方面。

(1) “人机料法环”智能管理(4M1E)：围绕人员(man)、机器(machine)、材料(material)、工艺方法(method)、环境(environment)等施工五大核心管理要素，利用计算机视觉、人工智能、云计算、物联网等技术手段建立智慧工地平台，已取得显著成果。

相似文献列表

去除本人文献复制比: 11%(703)

去除引用文献复制比: 11%(703)

文字复制比: 11%(703)

1	<u>单相机三维数字图像相关方法及其试验研究</u> 金杨磊(导师: 赵健) - 《北京林业大学硕士论文》 - 2019-06-07	3.8% (243) 是否引证: 否
2	<u>基于深度学习心电诊断系统的研究与实现</u> 胡文博(导师: 刘知青) - 《北京邮电大学硕士论文》 - 2019-06-05	2.3% (148) 是否引证: 否
3	<u>基于同步辐射的复合材料动力学性能与损伤实验研究</u> 别必雄(导师: 祁美兰; 罗胜年) - 《武汉理工大学博士论文》 - 2019-09-01	1.6% (102) 是否引证: 否
4	<u>基于计算机视觉的智能化建造技术研究与应用</u> 房霆宸; 杨佳林; 贺洪煜; - 《建筑施工》 - 2021-12-25	1.4% (90) 是否引证: 否
5	<u>基于条纹投影的三维形貌与形变测量技术研究进展</u> 吴周杰; 张启灿; - 《液晶与显示》 - 2023-06-12	0.8% (49) 是否引证: 否
6	<u>海上风电基础结构动力分析</u> 李炜; 郑永明; 陆飞; 罗金平; 姜贞强; 邹彩云; - 《海洋通报》 - 2012-02-15	0.8% (48) 是否引证: 否
7	<u>碳纤维复合材料电导特性和力电耦合行为研究进展</u> 韩朝锋; 薛有松; 张东生; 冯向伟; 陈莉娜; 朱晓伟; 吴海宏; 苏玉恒; - 《复合材料学报》 - 2023-01-19 18:10	0.8% (48) 是否引证: 否
8	<u>采用三点定位原理的反射镜支撑结构设计</u> 崔永鹏; 何欣; 张凯; - 《光学仪器》 - 2012-12-15	0.7% (45) 是否引证: 否
9	<u>综述: 面向工程应用的激光干涉动态测量</u> 傅愉; 陈冰; 阎可宇; 缪泓; 于起峰; - 《实验力学》 - 2021-02-15	0.7% (43) 是否引证: 否
10	<u>500kV变电支架结构模态分析与设计建议</u> 江飞; 陈寅; - 《低温建筑技术》 - 2021-09-28	0.7% (43) 是否引证: 否
11	<u>基于应力波检测的管道模态分析</u> 唐伟军; - 《广州建筑》 - 2011-04-15	0.7% (43) 是否引证: 否
12	<u>网架模型结构模态分析的试验研究</u> 王秀丽; 李晶; - 《低温建筑技术》 - 2011-09-28	0.7% (43) 是否引证: 否
13	<u>基于CAE的柴油机进气系统声学优化</u> 张袁元; 李舜酩; 刘建娅; 胡伊贤; 饶海生; - 《机械科学与技术》 - 2012-04-15	0.7% (43) 是否引证: 否
14	<u>纤维取向对三维四向编织复合材料与层合复合材料振动性能的影响</u> 裴晓园; 李嘉禄; 何玉强; - 《材料工程》 - 2013-07-20	0.7% (43) 是否引证: 否
15	<u>某型引信的结构样机与模态分析</u> 娄心豪; 王燕; 苗晋玲; - 《航空兵器》 - 2013-06-15	0.7% (43) 是否引证: 否
16	<u>声表面波温度传感器的仿真与设计</u> 吴展翔; 刘文; 卢小荣; - 《压电与声光》 - 2014-02-15	0.7% (43) 是否引证: 否
17	<u>圆柱型壳体动态分析</u> 蔡青; 李世芸; 于英林; 柴艳莉; 张亚军; - 《现代机械》 - 2006-02-25	0.7% (43) 是否引证: 否
18	<u>模态分析的损伤诊断在桅杆结构中的应用</u> 宋春; - 《山西建筑》 - 2007-06-01	0.7% (43) 是否引证: 否
19	<u>大型油罐损伤检测的模态分析方法</u> 楚孝田; - 《化学工程与装备》 - 2009-01-15	0.7% (43) 是否引证: 否
20	<u>基于模态分析的大型油罐损伤检测</u> 楚孝田; - 《石油化工腐蚀与防护》 - 2009-04-25	0.7% (43) 是否引证: 否
21	<u>考虑预应力时风扇叶片模态特征分析</u> 李云松; 任艳君; 程德蓉; - 《机械传动》 - 2010-05-15	0.7% (43) 是否引证: 否
22	<u>基于ANSYS的框架结构模态数值模拟分析</u>	0.7% (43)

	丁涛;于少春; - 《辽宁工程技术大学学报(自然科学版)》 - 2010-05-15	是否引证: 否
23	客车车身骨架模态的有限元分析 陈元华;张治龙; - 《桂林航天工业高等专科学校学报》 - 2010-09-15	0.7% (43) 是否引证: 否
24	汽车下河码头架空斜坡道结构性能的模式分析 宋成涛;刘全兴; - 《吉林水利》 - 2010-09-15	0.7% (43) 是否引证: 否
25	国家低频振动基准装置动圈模态优化 伍蒋军;陈洪军; - 《装备制造技术》 - 2015-02-15	0.7% (43) 是否引证: 否
26	长悬臂干挂石材支撑结构体系设计及其动力分析 杨明飞;史雪梅; - 《长江大学学报(自科版)》 - 2015-01-05	0.7% (43) 是否引证: 否
27	气化炉水冷壁振打除灰装置布局研究 喻九阳;郑小涛;刘裕华;汪威; - 《石油和化工设备》 - 2007-08-15	0.7% (42) 是否引证: 否
28	索-混凝土组合梁结构自振特性分析 张荣凯;陈宝军; - 《山西建筑》 - 2009-02-20	0.7% (42) 是否引证: 否
29	基于卷积神经网络的RC桥梁表观病害识别方法研究 邹俊志(导师: 杨建喜) - 《重庆交通大学硕士论文》 - 2021-04-12	0.6% (36) 是否引证: 否
原文内容		

通过分析工地现场的监控视频数据,快速智能识别工地现场存在的安全隐患并发出报警,保障建筑施工安全。静态场景中通过目标检测算法,识别出给定对象信息的感兴趣目标,如识别人员是否佩戴完整的个人防护设备。动态场景中通过目标轨迹追踪,获取人、机、料随时间变化的几何空间信息,用于评价人员是否暴露于危险区域、施工车辆停放位置是否合适等,规避潜在的事故风险。

图1-8 人机料法环智能管理的建筑工地[11]

(2)施工进度与质量控制:施工进度控制一直是建设工程项目管理的重要目标之一,传统的施工进度记录往往以表格和文字的形式保存,耗费人力和时间,且过于依赖工人经验,信息反馈不及时,影响管理者的决策。基于计算机视觉与BIM技术的进度控制方法,以无人机载相机或现场监控摄像头为数据采集手段,建立大型基础设施的稠密点云模型,利用目标检测算法识别出关键组件和重要材料位置,并借助BIM模型将实际进度可视化展示。另一个重要应用,施工质量检查与验收,如开发非接触式平直度检测技术、钢筋网扎丝绑扎质量智能验收系统等,通过计算机视觉技术中的目标检测和三维几何重建,自动获取安装质量参数,可优化工厂生产能力、劳动成本等。

a)设计模型, b)现场图像, c)图像三维重建点云模型, d)实际施工进度模型

图1-9 基于计算机视觉的施工进度可视化[12]

(3)结构病害智能检测:基于计算机视觉的结构外观病害检测技术,相比人工巡检具有检测速度快、效率高等优点,已经在裂缝发展、螺栓脱落、钢构件锈蚀、轨道扣件缺失等应用场景有取得优于人工检测的效果。利用无人机、爬壁机器人等智能检测设备,搭载高清照相机对结构关键部件进行检查,获取结构外观病害图像,再遵循识别、定位、评估三阶段完成分析。病害识别可以简化为计算机视觉里的目标检测任务,先建立高质量病害数据库,再通过深度学习算法识别出图像中的病害及类型。病害位置对评估损伤程度和监测其后期发展具有重要意义,而单张图像中识别出的病害缺少空间位置信息,因此通常采用图像拼接或三维重建方法,将病害的空间分布呈现到结构三维模型中。确定结构病害定位后,可结合规范对损伤的严重程度进行分级,也可通过有限元分析或其他计算方法对结构损伤程度进行定量分析,结合断裂力学、材料疲劳寿命等开展结构状态评估与剩余寿命预测。

图1-10 桥梁外观病害检测

(4)结构健康监测:基础设施结构的使用年限长达几十年、甚至达上百年,服役期间由于材料老化、环境侵蚀、长期荷载、疲劳与突然冲击等耦合作用,不可避免地会发生结构损伤和性能劣化。因此,结构损伤识别、服役状态和可靠性评估是结构健康诊断的重要问题。结构健康监测(Structural health monitoring, SHM),是通过测量结构关键部位的加速度、位移、应变等物理量,并从中提取损伤敏感的特征进行统计分析,来确定结构当前健康状态的技术手段。这种监测方法为结构维护和管理决策提供依据。基于计算机视觉技术的非接触式传感器,具备直接同步获取结构物多个空间位置点的变形信息功能。相比于传统接触式传感器,基于计算机视觉的非接触式传感技术具有安装方便、价格相对较低,并且支持多点测量等优点。根据应用场景不同,下面将分为全场变形及应变监测、动态位移监测及模态识别两部分介绍。

全场变形及应变监测:数字图像相关(Digital Image Correlation, DIC)是一种采用图像配准技术进行高精度、全场变形和应变测量的光学方法,应用于实验室内试件拉伸、扭转、弯曲和复合加载的静态应用。DIC测量系统一般由摄像机、照明光源及计算机组成。试验准备期间,通过喷雾罐、刷子、打印等方式在试样表面涂上散斑图案,确保加载期间散斑图案随试样表面变形一致。实验过程中,采集试件表面在加载前后的图像并存储到计算机中,通过图像分析算法获取试件表面目标全域的变形和应变数值。全局应变场演化云图可以从微观层次量化分析、总结裂纹产生和扩展规律及岩石变形损伤演化特征。

图1-11 混凝土梁四点弯曲实验:(a)梁尺寸及多相机布置;(b)表面全场裂纹分布图; 图格式不能英文[13]

动态位移监测及模态识别:模态是指结构的固有振动特性,每一个模态具有特定的固有频率、阻尼比和模态振型。模态识别,也就是从结构实测响应数据中提取出模态参数,是结构状态监测和损伤检测中的关键步骤,能帮助理解结构体如何振动以及其在外部荷载下的抵抗能力。例如,2021年赛格大厦的异常晃动,专家通过多组荷载工况下的模态测试和分析,证实该现象

系顶部桅杆风致涡激共振，也就是风吹过桅杆形成风涡引起桅杆振动，由于桅杆与楼体的某个高阶频率接近，从而带动大楼一起共振。传统的模态测试需要布设多组加速度计获取测点的加速度时程，而计算机视觉技术提供了一种非接触的简便方法，测量的结构响应是位移。测量设备由相机和计算机组成，关键测点处通常也会安装靶标作为观测对象，图像获取设备已经从固定位置拓展到无人机平台。在图像初始帧中框选出测点处的一个矩形框作为追踪目标，通过模板匹配、特征匹配、光流法等算法，在后续图像帧中定位出目标的新位置，进而获取其相对位移。监测获取的位移时程及模态识别结果，可用于修正结构有限元模型、估算斜拉桥的拉索索力等。

图1-12 在线相机监测索的异常振动文中没提到图，图可以大些

(5) 灾后快速评估：灾后第一时间准确获取建筑物、桥梁、道路等基础设施的损毁情况，能够为开展灾后应急救援、决策指挥以及灾后重建等提供技术支持和决策依据。传统的灾后基础设施损毁评估工作主要通过现场实地勘察，但效率低、且人力物力消耗大。遥感和近景航拍为灾情评估提供了丰富的数据源，而深度学习技术的不断发展显著提升了影像解析的效率，因此在灾后评估中发挥着越来越重要的作用。卫星遥感多用于城市建筑群等大尺度结构的损失评估，无人机航拍更适用于单体结构。

图1-13 基于航拍图像和深度学习技术的建筑物震害识别[14]

近年来，基于计算机视觉的土木工程结构智能建造与运维已取得显著的进展，极大改善了施工和维护阶段的效率和成本。这些进展很大程度上受益于不同学科的交叉融合。通过引入计算机视觉、人工智能等技术，我们能够实现高效、经济、自动化的基础设施建造、检测、监测。目前，一些新的理念与技术，如生物视觉、智能感知、注意力机制等，也正融入到计算机视觉系统中，进一步提升其在土木工程领域的应用水平。从技术角度而言，视觉技术的应用提升仍有很大的空间：

(1) 全面的场景理解和结构状态解析。现有的建造运维智能化解决方案中，计算机视觉技术通常被应用到同一个场景中多个独立任务中，如施工场景中的不安全行为预警、施工进度识别、质量检查等，运维场景中的结构变形响应监测、外观病害检测、车流荷载识别等。如何充分利用获取多层次多对象的视觉数据，优化施工管理体系和质量监督策略、解析结构服役性能及行为演化机制，是很具有挑战性的问题。

(2) 土木工程领域的公共数据库和泛化能力强的模式识别算法。数据是人工智能的“养料”，深度学习算法往往依赖大量的标注数据，因此研究人员往往会在简单却繁杂的标注任务上耗费了大量时间。因此，在海量的大数据中自动提取高质量、有效的特征，并通过建立的大数据库自动对比分析，融合人工智能算力算法，驱动系统主动分析识别处理，可以大幅提升计算机视觉的识别能力和解决实际工程问题的效率。

1.5本书各章节简介

本书涉及土木工程行业智能建造与计算机视觉的主要研究内容，首先以介绍计算机视觉的定义、发展历程以及与机器视觉、摄影测量学的关系为开端。随后，探讨了深度学习与计算机视觉的关系，并讨论了计算机视觉在土木工程领域的应用及发展趋势。本书按照由简到繁、从理论到实际的顺序编排各章节，涵盖了视觉模型到视觉系统、二维视觉到三维视觉、视觉算法到视觉应用、单传感器视觉到多传感器视觉的内容的顺序来安排，体现从浅入深、从理论到实例的特点，力求具有系统性、层次性和实用性。

第1章

第2-5 章基本理论。第2章详细介绍了数字相机、空间坐标系变换、成像模型以及相机标定等内容。第3章讨论了数字图像处理的基础知识，包括图像基础、滤波、频域处理、边缘检测和图像插值方法。在第4章图像特征检测与图像匹配一章中，全面探讨了图像特征的检测与描述方法，包括全局图像特征、角点检测、斑点检测和特征描述子等内容，为后续目标识别、分割和三维重建等章节的内容提供了基础。随后的第5章深度学习与图形处理一章详细介绍了深度学习与图像处理的关系，包括深度学习的基本组件、迭代优化、运算以及常见的深度学习模型和实现框架，为读者深入理解深度学习在计算机视觉中的应用奠定了基础。

第6-9章高级应用方法。第6章基于计算机视觉的位移测量原理一章，着重介绍了结构位移监测的基本原理和方法，包括比例系数确定、基于特征匹配的位移测量示例以及位移测量精度评估等内容，为读者理解计算机视觉在结构位移监测中的应用提供了实际案例。第7章目标识别与分割一章则详细讨论了目标识别方法的概述，以及图像分类、目标检测和图像分割等相关技术。接着第8章立体视觉与三维重建一章深入研究了三维重建的方法与技术，包括双目立体视觉、运动恢复结构和点云处理等内容，为读者了解三维重建的原理和实践提供了重要参考。

第9-11章典型应用。第9章基于计算机视觉的结构位移监测应用一章和结构表观病害智能检测应用一章，分别以泗安塘大桥位移监测应用和预制墩柱安装位置姿态监测为例，介绍了计算机视觉在结构位移监测和病害智能检测中的应用案例。最后，第10章以基于计算机视觉的结构三维几何尺寸的测量应用一章作为总结，详细讨论了基于双目视觉和三维激光扫描的预制墩柱安装位置姿态监测的方法和应用。

通过以上章节的安排，本书力求以层次性、系统性、先进性和实用性的特点，全面介绍了计算机视觉的主要研究内容和实际应用，为读者提供丰富的知识和实践指导。

习题

- 1-1 什么是计算机视觉，试举例说明计算机视觉相关的具体应用案例。
- 1-3 随着科技的发展，计算机视觉在越来越多的领域得到了广泛应用，你认为在未来计算机视觉还会有哪些具体的应用？
- 1-4 计算机视觉、机器视觉和摄影测量学三者之间的关系是怎样的，有怎样的联系和区别？
- 1-5 如何理解机器学习与深度学习之间的关系？深度学习的出现对计算机视觉的发展有哪些影响？
- 1-6 结构健康监测与结构智能检测有怎样的区别？与计算机视觉有怎样的联系？
- 1-7 目前计算机视觉在土木工程领域的应用有哪些？你认为计算机视觉还可以应用在土建行业的哪些场景？

参考文献（后期会调整到全书最后）

[1] <https://www.nist.gov/mathematics-statistics/first-digital-image>
[2] Hubel, David H., and Torsten N. Wiesel. "Receptive fields of single neurones in the cat's striate cortex." *The Journal of physiology* 148, no. 3 (1959): 574.
[3] Roberts, Lawrence G. "Machine perception of three-dimensional solids." PhD dissertation, Massachusetts

[4] [Peters W H and Ranson W F 1981 Digital imaging techniques in experimental stress analysis Opt. Eng. 21 427 - 31.](#)

[5] [Pan, B., 2018. Digital image correlation for surface deformation measurement: historical developments, recent advances and future goals. Measurement Science and Technology, 29\(8\), p.082001.](#)

[6] <https://qualitastech.com/image-processing/machine-vision-in-defect-detection-activities-using-ai-and-3d/>

[7] https://blog.csdn.net/weixin_39789327/article/details/110634310

[8] <https://www.c114.com.cn/m2m/2488/a1219428.html>

[9] <https://panchuang.net/2021/07/02/%e5%9f%ba%e4%ba%8e%e6%b7%b1%e5%ba%a6%e5%ad%a6%e4%b9%a0%e7%9a%84%e8%ae%a1%e7%ae%97%e6%9c%ba%e8%a7%86%e8%a7%89%e6%9d%82%e8%8d%89%e5%ae%9e%e6%97%b6%e6%a3%80%e6%b5%8b%e4%b8%8e%e5%88%86%e7%b1%bb/> 这样的文献去掉?

[10] 于起峰, 尚洋, YUQifeng, 等. 摄像测量学简介与展望[J]. 科技导报, 2008, 26(24):84-88.

[11] <https://skyengine.ai/se/skyengine-blog/82-adaptive-construction-intelligence-with-deep-learning-in-virtual-reality-using-sky-engine-ai-platform-to-improve-on-site-predictive-analytic>

[12] Braun, Alex, Sebastian Tuttas, André Borrmann, and Uwe Stilla. "Improving progress monitoring by fusing point clouds, semantic data and computer vision." Automation in Construction 116 (2020): 103210.

[13] 谷柳凝, 宫文然, 邵新星, 陈捷, 董志强, 吴刚, 何小元. 基于主应变场的混凝土全表面开裂特征实时测量与分析. 力学学报, 2021, 53(7): 1962-1970 doi: 10.6052/0459-1879-21-107.

[14] Miura, Hiroyuki, Tomohiro Aridome, and Masashi Matsuoka. 2020. "Deep Learning-Based Identification of Collapsed, Non-Collapsed and Blue Tarp-Covered Buildings from Post-Disaster Aerial Images" Remote Sensing 12, no. 12: 1924. <https://doi.org/10.3390/rs1212192>

文献格式统一

3. 计算机视觉教材_第3部分		总字符数: 10935
相似文献列表		
去除本人文献复制比: 56.6%(6189) 去除引用文献复制比: 56.6%(6189) 文字复制比: 56.6%(6189)		
1	<u>基于机器视觉的车用灯泡检测系统研究</u> 王佳炎(导师: 宋丽梅;王玉祥) - 《天津工业大学硕士论文》 - 2017-02-18	24.1% (2633) 是否引证: 否
2	<u>基于机器视觉的车用灯泡检测系统研究 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co)》 - 2020	22.8% (2492) 是否引证: 否
3	<u>双目显微立体视觉中匹配技术的研究</u> 李宗艳(导师: 宋丽梅) - 《天津工业大学硕士论文》 - 2015-12-01	13.4% (1468) 是否引证: 否
4	<u>基于图像匹配技术的桥梁位移无标记测量方法研究</u> 汪威(导师: 吴刚;吴栋) - 《东南大学硕士论文》 - 2021-05-30	4.4% (486) 是否引证: 否
5	<u>基于双目立体视觉的异物侵入检测系统设计</u> 魏强(导师: 余祖俊) - 《北京交通大学硕士论文》 - 2009-06-01	4.1% (447) 是否引证: 否
6	<u>主动视频监控中若干问题的研究</u> 杨广林(导师: 孔令富) - 《燕山大学博士论文》 - 2006-10-01	3.7% (406) 是否引证: 否
7	<u>边缘检测和摄像机标定算法的研究</u> 熊建平(导师: 吴建华;刘守义) - 《南昌大学硕士论文》 - 2005-05-01	3.3% (356) 是否引证: 否
8	<u>数字摄影测量与三维重建系统关键技术研究</u> 谢耀华(导师: 孙茂印) - 《国防科学技术大学硕士论文》 - 2004-11-01	2.8% (308) 是否引证: 否
9	<u>基于计算机视觉的车架纵梁在线检测关键技术研究</u> 王华(导师: 王龙山) - 《吉林大学博士论文》 - 2009-06-01	2.7% (299) 是否引证: 否
10	<u>织物疵点检测系统的关键技术研究</u> 姚桂国(导师: 左保齐) - 《苏州大学硕士论文》 - 2010-05-01	2.7% (299) 是否引证: 否
11	<u>[DOCIN]基于图像序列的三维虚拟城市重建关键技术研究 - 豆丁网</u> - 《互联网文档资源 (http://www.docin.com)》 - 2017	2.6% (283) 是否引证: 否
12	<u>双目视觉中的立体匹配技术研究 - 豆丁网</u> - 《互联网文档资源 (http://www.docin.com)》 - 2012	2.6% (283) 是否引证: 否

13	<u>植物根系图像监测分析系统</u>	2.5% (275)
	刘九庆(导师:李东升;汤晓华) - 《东北林业大学博士论文》 - 2005-04-01	是否引证: 否
14	<u>双目立体视觉在机器人三维重建定位中的方法研究</u>	2.5% (271)
	邱光帅(导师:晋帆) - 《昆明理工大学硕士论文》 - 2005-04-25	是否引证: 否
15	<u>全景视觉系统标定技术研究</u>	2.4% (260)
	朱睿爽(导师:原新) - 《哈尔滨工程大学硕士论文》 - 2010-09-10	是否引证: 否
16	<u>基于摄影测量系统的标定与表面数据融合算法研究</u>	2.3% (255)
	付玮(导师:吴禄慎) - 《南昌大学博士论文》 - 2017-12-17	是否引证: 否
17	<u>基于CCD摄像机的举重运动分析系统的预处理及研究</u>	2.3% (254)
	伍尤富(导师:王修信) - 《广西师范大学硕士论文》 - 2005-04-01	是否引证: 否
18	<u>基于双目立体视觉的AGV导航技术研究</u>	2.3% (254)
	林琳(导师:黄玉美;李艳) - 《西安理工大学硕士论文》 - 2007-03-01	是否引证: 否
19	线阵相机技术工作报告整理.pdf-原创力文档	2.3% (254)
	- 《互联网文档资源 (https://max.book118.)》 - 2020	是否引证: 否
20	<u>基于多尺度特征的图像匹配与目标定位研究</u>	2.1% (231)
	刘立(导师:彭复员) - 《华中科技大学博士论文》 - 2008-10-01	是否引证: 否
21	线阵相机技术报告整理下载	2.0% (223)
	- 《互联网文档资源 (https://ishare.iask.)》 - 2020	是否引证: 否
22	图像与机器视觉之相机篇-百度文库	2.0% (223)
	- 《互联网文档资源 (http://wenku.baidu.c)》 - 2012	是否引证: 否
23	线阵相机相关技术教学.pdf	2.0% (222)
	- 《互联网文档资源 (https://max.book118.)》 - 2020	是否引证: 否
24	<u>不同种植环境下作物三维重建与信息提取</u>	1.7% (191)
	王乐(导师:蒋焕煜) - 《浙江大学博士论文》 - 2021-06-30	是否引证: 否
25	计算机视觉与图像分割技术在三维几何信息测量系统中的应用研究-百度文库	1.7% (181)
	- 《互联网文档资源 (http://wenku.baidu.c)》 - 2012	是否引证: 否
26	<u>道路交通事故现场图自动绘制</u>	1.6% (177)
	花开胜(导师:王林) - 《上海交通大学硕士论文》 - 2012-01-01	是否引证: 否
27	<u>基于网络摄像机的双目立体视觉系统</u>	1.4% (153)
	吕林涛(导师:王平;赵鸿) - 《哈尔滨工程大学硕士论文》 - 2019-05-01	是否引证: 否
28	<u>基于单目图像序列的场景稠密三维重建技术研究</u>	1.2% (129)
	杨爱林(导师:贾松敏) - 《北京工业大学硕士论文》 - 2016-06-01	是否引证: 否
29	视觉检测中工业相机的参数设置_深圳科视创	1.0% (104)
	- 《网络 (http://blog.sina.com)》 - 2019	是否引证: 否
30	<u>月球车立体视觉标定与三维重建技术研究</u>	0.8% (91)
	张宏(导师:齐乃明) - 《哈尔滨工业大学硕士论文》 - 2006-06-01	是否引证: 否
31	<u>一种单相机测量三维运动轨迹的方法</u>	0.6% (69)
	张恒康;何玉明;张耿耿;朱杰兵;杨文俊; - 《固体力学学报》 - 2010-12-15	是否引证: 否
32	<u>增强现实注册跟踪技术的研究</u>	0.3% (37)
	孙可言;陈根;廖金巧; - 《科技资讯》 - 2017-02-23	是否引证: 否
33	<u>大尺寸移动机器人测量系统标定技术研究</u>	0.3% (35)
	马春雷(导师:赵延治;齐立哲) - 《燕山大学硕士论文》 - 2019-05-01	是否引证: 否
34	<u>双足机器人目标识别与跟踪方法研究</u>	0.3% (30)
	潘亚宾(导师:刘国栋) - 《江南大学硕士论文》 - 2011-03-01	是否引证: 否

原文内容

第2章相机成像与相机标定
2.1数字相机
2.1.1 传感器

工业相机主要用于对待测物体进行实际的成像，它能够将光信号转换成电信号或者数字信号，其关键部位是CCD(Charge Coupled Device, 电荷耦合器件)或CMOS(Complementary Metal Oxide Semiconductor, 互补金属氧化物半导体)图像传感器。图像像素值(即图像灰度值)是CCD或CMOS图像传感器对光强(光亮度)的测量对象。由于采集到的图像的质量将直接影响到后续图像处理的效果，所以选择一个合适的相机对计算机视觉检测系统来说非常重要。对于一般的工业相机有如下参数。

(a) CCD (b) CMOS

图 2-1 常见的传感器英文改为中文，出处？

(1) 分辨率：分辨率是工业相机最基本的参数，可以用水平分辨率(H)和垂直分辨率(V)来描述，如1920(H)×1080(V)表示每行的像元数量为1920个，共有1080行。分辨率越大，图像所占内存空间也越大，但是图片的细节越清晰。所以，在选择分辨率时，一般要首先保证图像有足够多的数据信息，能够满足系统检测的需要，其次考虑分辨率不能过大，以免浪费资源、降低效率。

(a) 低分辨率 (b) 高分辨率

图 2-2 不同分辨率下图像的表现

(2) 帧速：帧速表示相机采集图像的频率，单位为帧/秒，即每秒所能采集的画面数量，可以用来衡量采集速度。帧速是衡量一台工业相机性能的重要参数，当相机速度足够时才能满足相应的系统需求。一般而言，不同类型的机器视觉系统所需要的帧速不尽相同，帧速需要实际项目需求相匹配。目前，工业相机的帧速从十几帧/秒到几百帧/秒甚至上万帧/秒都有。

(3) 传感器像元大小：传感器芯片由许多像元阵列组成，像元尺寸指芯片像元阵列上每个像元的实际物理尺寸。像元尺寸从某种程度上反映了芯片对光的响应能力，像元尺寸越大，能够接收到的光子越多，在同样的光照条件和曝光时间内产生的电荷数量越多。通常像元尺寸为几微米，大尺寸像元感光能力更好，但是使得在相同芯片面积条件下像素分辨率更低。

图 2-3 像素需中文

(4) 动态范围：相机的动态范围表明相机探测光信号的范围，可用两种方法来界定：一种是光学动态范围，指饱和时最大光强和等价于噪声输出的光强的比值，由芯片的特性决定；另一种是电子动态范围，指饱和电压和噪声电压之间的比值。动态范围大，则相机对不同的光照强度有更强的适应能力。

2.1.2 镜头

镜头是连接待测物体所反射的光线和相机成像的通道，主要作用是实现光束的变换调制，将待测物体成像在相机图像传感器的光敏面上。工业镜头对于被测物体成像有着十分关键的作用，它的质量直接影响机器视觉应用系统的整体性能。

工业镜头的历史悠久、品类繁多，一般可以进行如下划分。

(1) 根据工业镜头的接口类型进行划分。工业镜头与工业相机间常用的接口模式有C接口、CS接口、F接口、V接口、T2接口、徕卡接口、M42接口、M50接口等。不过这些模式只是接口方式不同，并未对镜头的性能做相应的区分，并且一般情况下，为提高工业镜头的实用性和适应性，常用的接口之间也设置有相应的转接口。上述接口模式中，使用最多的为C接口和CS接口，它们的螺纹连接相同，只是后截距不同。

图 2-4 两种常见接口需中文出处

(2) 根据能否变焦进行划分。根据能否变焦可将镜头分为定焦镜头和变焦镜头。定焦镜头的焦距是固定不变的，它的焦距只有一个，即镜头只有固定的视野。定焦镜头按照等效焦距又可以划分为鱼镜头、广角镜头、标准镜头、长焦镜头、超长焦镜头。不同于变焦镜头的复杂设计，定焦镜头的内部结构更显精简。虽然变焦镜头可以适当改变焦距，但是变焦后对于物体的成像会有影响，所以定焦镜头的优势在于对焦速度快、成像质量稳定。显然，变焦镜头的优势就是焦距可变，这样便可以在物距不变时改变视场。

(3) 根据镜头光圈进行划分。镜头按照光圈的调节方式可分为手动光圈和自动光圈两种，当待测物体上的光线变化不大、较为恒定时，适合用手动光圈，而当环境光线变化较为明显时，适合用自动光圈，从而能够根据实际的环境光实时改变镜头的光圈大小。

确定具体的镜头型号前，还需要确定镜头的一些基本参数。

(1) 焦距：焦距是焦点到透镜中心点之间的距离，实际使用时镜头上面都会标注出焦距，并不需要使用者计算。

(2) 视场角：整个视觉系统所能观察到的物体实际尺寸被称为视场即视场范围(Field of View, FOV)。假设F是焦距，S是传感器宽度，FOV是视场角，那么由图 2-5 可得： $F = (S/2) / \tan(FOV/2)$ 。

图 2-5 相机成像光路图

(3) 光圈：光圈实际为可以调节孔径大小的机械部件，它通常在镜头内，利用控制镜头光孔大小来控制进入相机的光量。一般而言，当外界光较弱时，光圈应该相应开大，当外界光较强时，光圈应该相应开小。光圈大小一般用F表示，比如F1、F1.4、F2.8、F4、F5.6、F8、F11、F16、F22、F32等，这些表示中上一级正好是下一级通光量的一倍，即光圈开大了一级。

为了表示方便，光圈大小通常是焦距的分数，这个分数被称为透镜的f数，它是焦距与光圈直径的比值，记为N。如图 2-6所示是一个焦距为50mm的相机镜头，其中D代表光圈打开的大小。

图 2-6 50mm焦距时光圈打开不同大小

(4) 景深：镜头能够取得清晰图像时被测物体的前后距离范围即为景深。一般而言，改变光圈、焦距、拍摄距离时，景深会相应变化。例如，如下图所示，将光圈值调大，则景深相应变小；而将光圈值调小，则景深相应变大。对于桥梁检测而言，由于现在大跨桥梁的普遍使用，对相机景深的要求就有较高的要求，故而会在使用过程中选择小光圈，但是光圈过小会导致进光量过小，会涉及到补光等需求。

图 2-7 光圈大小对景深的影响需中文出处

(5) 畸变：理想的物体成像中，物体和成像应该完全相似，然而，在实际的成像过程中，由于镜头本身的光学结构以及成像特性，镜头会不可避免地产生畸变可以简单地理解为这是由像面上局部放大倍率不一致所导致的。选用镜头时，需要根据所需达到的目标和精度来选择不同质量的镜头。相机畸变的类型包括：

1. 径向畸变

①枕形畸变：又称鞍形形变，视野中边缘区域的放大率远大于光轴中心区域的放大率，常出现在远镜头中；

②桶形畸变，与枕形畸变相反，视野中光轴中心区域的放大率远大于边缘区域的放大率，常出现在广角镜头和鱼眼镜头中

③线性畸变：光轴与相机所拍摄的诸如建筑物类的物体的垂平面不正交，则原本应该平行的远端一侧和近端一侧，以不相同的角度汇聚产生畸变。这种畸变本质上是一种透视变换，即在某一特定角度，任何镜头都会产生相似的畸变。

(a) 正常图像 (b) 枕形畸变 (c) 桶型畸变
图2-8 径向相机畸变示意图

2. 切向畸变

切向畸变就是矢量端点沿切线方向发生的变化，也就是角度的变化 dt 。简单来说就是：假设上下两点是矢量端点，沿着切向方向变化，原本是一个正圆就变成了一个椭圆。切向畸变是由于透镜安装与“图像”平面不平行而产生的在光线的垂直面上的变形。

- 为一个机器视觉系统选择镜头时，一般可以按照下列步骤来进行。
- (1) 根据系统整体尺寸和工作距离，结合视场角，大概判断出所需镜头的焦距。
 - (2) 切换不同的光圈大小，找到最合适的值。
 - (3) 考虑镜头畸变、景深、接口等其他要求。

2.1.3 相机接口

根据需求选择匹配的相机后、系统需要将图像通过相机接口传输到相应的处理设备中，此时便要用到相机接口技术。相机接口技术可以分为模拟接口技术和数字接口技术两大类。模拟接口技术主要是将模拟数据采集卡与图像处理设备相连，其数据传输速度和精度都较差，并且随着数字化技术的发展，模拟接口技术终会消亡，但由于模拟视频设备的低价，它们在图像处理应用的低端领域还有一定的市场，其消亡还要相当长的一段时间。数字接口技术是目前相机接口的主流技术，下面介绍最常用的几种数字相机接口。

表 2-1 不同相机接口的优缺点

接口类型	吞吐量	性价比	数据线长度	标准接口	POC供电	CPU使用	I/O同步
Camera Link							
USB 2.0							
USB 3.0							
IEEE 1394							
GigE Vision							

接口类型 吞吐量 性价比 数据线长度 标准接口 POC供电 CPU使用 I/O同步
Camera Link
USB 2.0
USB 3.0
IEEE 1394
GigE Vision

(1) Camera Link接口。Camera Link是为高端应用而研发的，其数据传输速度可以达到1Gbits/s。针对不同的应用需求，例如分辨率、传输速度等，相应地有低、中、高三档格式。虽然Camera Link规范已经成为包括线阵相机、高速面阵相机在内的高速图像采集设备标准，但是此标准的缺点也很明显，那就是计算机端需要添加额外的图像采集卡以完成对数据的重构，相应地，设备的成本就增加了。

(2) IEEE 1394接口。IEEE 1394 亦被称作火线，最初的标准规定数据传输速度为93.304 Mbit/s、196.608 Mbit/s和393.216MbiVs。它需要使用6针线插件，两对双绞线用于传输信号，一根电缆用作电源，另一根用作地线。另外，此标准接口还定义了带有锁固功能的接插件，这在工业应用中是很有用的，如可以防止电缆线的意外脱开。IEEE 1394接口提供较快速度的同时，传输距离也足够远，体积相对其他接口而言较小，也能提供较高的分辨率以及帧频，适用于医学成像和实时速度要求不是特别高的应用场合。

(3) USB接口。USB作为计算机系统连接外部设备的一种串口总线标准，广泛应用于个人计算机和移动设备等通信产品。虽然最初设计的传输速度不高，但是发展较快，比如新一代的 USB 3.1传输速度为10Gbit/s, 新型Type C接口正反向都能插入设备，使用较为方便。

(4) Gigabit Ethernet接口。Gigabit Ethernet (GigE)即千兆以太网，它作为传统以太网的新型技术，在高速、远距离、大批量数据传输方面优势明显。它还具有简易、可扩展性好、安全可靠、管理维护方便等一系列优点，是未来应用的趋势。

(5) CoaXpress 接口。CoaXpress (简称CXP)接口2009年在斯图加特的VISION展会上推出，它保留了同轴电缆的优点，同时具有高速数据传输能力(高达6.25 Gbit/s)。该接口已经成为高性能相机系统实际采用的相机到计算机之间的接口之一。

2.2 空间坐标系变换

在对相机进行标定前，为确定空间物体表面上点的三维几何位置与其在二维图像中对应点之间的相互关系，首先需要对相机成像模型进行分析。在机器视觉中，相机模型通过一定的坐标映射关系，将二维图像上的点映射到三维空间。相机成像模型中涉及世界坐标系、相机坐标系、图像像素坐标系及图像物理坐标系四个坐标系间的转换。

为了更加准确地描述相机的成像过程，首先需要对上述四个坐标系进行定义。

- (1) 世界坐标系：又叫真实坐标系，是在真实环境中选择一个参考坐标系来描述物体和相机的位置。
- (2) 相机坐标系：是以相机的光心为坐标原点，z轴与光轴重合、与成像平面垂直，x轴与y轴分别与图像物理坐标系的X轴和Y轴平行的坐标系。
- (3) 图像像素坐标系：建立在图像中的平面直角坐标系，单位为像素，用来表示各像素点在像平面上的位置，其原点位于图像的左上角。
- (4) 图像物理坐标系：其原点是成像平面与光轴的交点，X轴和Y轴分别与相机坐标系的x轴与y轴平行，通常单位为mm，图像的像素位置用物理单位来表示。

图 2-9 坐标转换流程

2.2.1 世界坐标系与相机坐标系转换

图 2-10为世界坐标系与相机坐标系的转换示意图。利用旋转矩阵R与平移向量T可以实现世界坐标系中的坐标点到相机坐标系中的映射。

图 2-10 世界坐标系与相机坐标系的转换示意图

如果已知相机坐标系中的一点相对于世界坐标系的旋转矩阵与平移向量，则世界坐标系与相机坐标系的转换关系为

$$(2-1)$$

其中R为 3×3 矩阵，T为 3×1 平移向量。

2.2.2 相机坐标系与图像物理坐标系转换

如图 2-11所示，成像平面所在的平面坐标系就是图像物理坐标系。

图 2-11 相机坐标系与图像物理坐标系的转换示意图

空间中任意一点P在图像平面的投影p是光心O与P点的连接线与成像平面的交点，由透视投影，可知

$$(2-2)$$

式中， x, y 是空间点P在相机坐标系下的坐标， u, v 对应图像物理坐标系下的坐标， f 为相机的焦距。则由式（2-2）可以得到相机坐标系与图像物理坐标系间的转换关系为：

$$(2-3)$$

2.2.3 图像像素坐标系与图像物理坐标系转换

图 2-12展示了图像像素坐标系与物理坐标系之间示系的对应关系。其中， 0_{uv} 为图像像素坐标系，0点与图像左上角重合。该坐标系以像素为单位， u, v 为像素的横、纵坐标，分别对应其在图像数组中的列数和行数。为图像物理坐标系，其原点在图像像素坐标系下的坐标为 (u_0, v_0) 。 dx 与 dy 分别表示单个像素在横轴X和纵轴Y上的物理尺寸。

图 2-12 图像像素坐标系与图像物理坐标系

上述两坐标系之间的转换关系为

$$(2-4)$$

将式（4-4）转换为矩阵齐次坐标形式为

$$(2-5)$$

2.2.4 图像像素坐标系与世界坐标系转换

根据各坐标系间的转换关系，即式（2-1）、式（2-3）、式（2-5）可以得到世界坐标系与图像像素坐标系的转换关系（2-6）

式中， M 为 3×4 矩阵，被称为投影矩阵：由参数 $a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8$ 决定，这些参数只与相机的内部结构有关，因此称为相机的内部参数（内参）；被称为相机的外部参数（外参），由相机相对于世界坐标系的位置决定。确定相机内参和外参的过程即为相机的标定。

以上述四个坐标系的对应关系为基础，下文将对相机的成像模型进行分析。常用的成熟模型包括线性模型和非线性模型。

线性模型针对针孔模型或透镜模型，忽略了镜头畸变，用于视野较狭窄的相机标定。为准确地描述针孔透视成像过程，需要考虑非线性畸变，从而引入了相机的非线性模型与非线性标定。这类标定方法的优点是可以假设相机的光学成像模型非常复杂，充分考虑成像过程中的各种因素，能够得到比较高的标定精度。Faig标定方法是这一类型的代表，它充分考虑了成像过程中的各种因素，精心设计了相机模型，然后寻找在某些约束条件下的最小值，进行非线性优化求解。求解非线性优化问题可选用Levenberg-Marquart等优化算法。

2.3 成像模型

相机成像描述相机将三维场景变换到二维像平面上的过程。相机成像模型一般可分为线性模型和非线性模型。线性模型可用针孔成像模型表示；当考虑到镜头畸变等因素时，相机成像模型为非线性。

图 2-13 成像模型需中文

2.3.1 线性模型

实际应用中，相机可采用理想的小孔成像模型近似表示。小孔成像模型是各种相机模型中最简单的一种，它是相机的一个近似线性模型。在相机坐标系下，任一点在图像平面的投影点都是（即光心与点的连线）与图像平面的交点，其几何关系如图2-6所示。

图 2-14 相机小孔成像模型

小孔成像模型的投影公式为

$$(2-7)$$

简写为

$$(2-8)$$

2.3.2 非线性模型

实际成像过程中，由于相机镜头的加工误差、装备误差等原因会产生相机畸变，使成像点偏离原来应成像的位置，所以线性模型不能准确地描述相机的成像几何关系。非线性模型可用式（2-9）来描述

$$(2-9)$$

式中， x_d, y_d 是经过畸变的点， x, y 是线性模型计算出来的图像点坐标理想值， $\delta x, \delta y$ 是非线性畸变，公式为

$$(2-10)$$

式中， k_1, k_2 是径向畸变， k_3, k_4 是离心畸变， k_5, k_6, k_7, k_8 是薄棱镜畸变， $p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8$ 是畸变参数。

在相机标定过程中，通常不考虑离心畸变和薄棱镜畸变，因为对于引入的非线性畸变因素，往往需要使用附加的非线性算法对其进行优化，而大量研究表明，引入较多的非线性参数不仅对标定精度的提高作用不大，还会造成解的不稳定性，而且一般情况下只使用径向畸变就足以描述非线性畸变，则有

$$(2-11)$$

式中， r 为径向半径。式(2-11)表明，相机畸变程度与 r 有关。 r 越大，畸变越严重，位于边缘的点偏离越大。

将式(2-11)代入式(2-6)可得

(2-12)

简写为

(2-13)

2.4相机标定

相机标定是确定世界坐标到图像像素坐标之间转换关系的过程。标定技术主要依靠世界坐标系中的一组点，它们的相对坐标已知，且对应的像平面坐标也已知，通过物体表面某点的三维几何位置与其在图像对应点之间的相互关系得到相机几何模型参数，通过这种方式得到相机的内参与外参的过程称为相机标定。从广义上讲，现有的相机标定方法可以归结为两类：传统的相机标定和相机自标定。目前，传统相机标定技术研究如何有效、合理地确定非线性畸变校正模型的参数以及如何快速求解成像模型等，而相机自标定研究不需要标定参照物情况下的方法。传统的标定技术需要相机拍摄一个三维标定靶进行标定，而较新的标定技术仅仅需要一些平面靶标，如MATLAB标定工具箱等。从计算方法的角度，传统相机标定主要分为线性标定方法（透视变换矩阵和直接线性变换）、非线性标定方法、两步标定方法和平面模板方法。常见标定板包括棋盘格标定板和圆形网格标定板。

棋盘格标定板圆形网格标定板

图 2-15 两种常见的标定板图右边有小文字，版权问题

2.4.1 Tsai相机标定

直接线性变换方法或者透视变换矩阵方法利用线性方法来求取相机参数，其缺点是没有考虑镜头的非线性畸变。如果利用直接线性变换方法或透视变换矩阵方法求得相机参数，可以将求得的参数作为下一步的初始值，考虑畸变因素，利用最优化算法进一步提高标定精度，这样就形成了所谓的两步法。

两步法的第一步是解线性方程，得到部分外参的精确解。第二步再将其余外参与畸变修正系数进行迭代求解。较为典型的两步法是Tsai提出的基于径向约束的两步法。基于径向约束的相机标定方法标定过程快捷、准确，但是只考虑了径向畸变，没有考虑其他畸变。该方法所使用的大部分方程是线性方程，从而降低了参数求解的复杂性。

其标定过程是先忽略镜头的误差，利用中间变量将标定方程化为线性方程后求解出相机的外参，然后根据外参利用非线性优化的方法求取径向畸变系数 k 、有效焦距 f 以及平移分量。

径向排列约束矢量和矢量具有相同的方向。由成像模型可知，径向畸变不改变的方向。由式(4-1。。。)可得

(2-13)

(2-14)

整理可得

(2-15)

上式两边同除以，得

(2-16)

再将式(2-16)变换为矢量形式得

(2-17)

式中，行矢量是已知的，而列矢量是待求参数。(2-18)

利用最小二乘法求解这个方程组、计算有效焦距 f 、平移分量 t 和透镜畸变系数 k 时，先用线性最小二乘计算有效焦距 f 和平移向量 T 的分量，然后利用有效焦距 f 和平移向量 T 的分量值作初始值，求解非线性方程组得到 f 、 k 的准确值。

利用式(2-18)以及旋转矩阵为正交阵的特点，可以确定旋转矩阵 R 和平移分量 t 。

利用径向一致约束方法将外参分离出来，并用求解线性方程的方法求解外参。另外，可将世界坐标和相机坐标重合，这样，标定时就能只求内参，从而简化标定。

2.4.2 张正友标定

1998年，张正友提出了基于二维平面靶标的标定方法，使用相机在不同角度下拍摄多幅平面靶标的图像，比如棋盘格的图像，然后通过对手棋格的角点进行计算分析来求解相机的内外参数。

1. 对每一幅图像得到一个映射矩阵 H

一个二维点可以用表示，一个三维点可以用表示，用表示其增广矩阵，则以及。三维点与其投影图像点之间的关系为

(2-19)

式中， s 是任意标准矢量； R 、 t 为外参； A 矩阵为相机内参，可表示为

(2-20)

式中， O 是坐标系上的原点； α 和 β 是图像上 u 和 v 坐标轴的尺度因子； γ 表示图像坐标轴的垂直度。

不失一般性，假定模板平面在世界坐标系 $Z=0$ 的平面上，则由式(2-20)可得

(2-21)

式中，为标定模板平面上的齐次坐标，为模板平面上的点投影里图像平面上对应点的齐次坐标。

此时，可以得到一个 3×3 的矩阵

(2-22)

利用映射矩阵可得内参矩阵 A 的约束条件为

(2-23)

2. 利用约束条件线性求解内参矩阵 A

假设存在

(2-24)

式中， B 是对称矩阵，可以表示为六维矢量，基于绝对二次曲线原理求出 B 以后，再对 B 矩阵求逆，并从中导出内参矩阵 A ；再由 A 和映射矩阵 H 计算外参能转矩阵 R 和平移向量 t ，公式为

(2-25)

3. 最大似然估计

采用最大似然准则优化上述参数。假设图像有 n 幅，模板平面标定点有 m 个，则最大似然估计值就可以通过最小化以下公式得到：

(2-26)

式中， x_j 为第 j 个点在第 i 幅图像中的像点； R_i 为第 i 幅图像的旋转矩阵； T_i 为第 i 幅图像的平移向量； (x_j, y_j, z_j) 为第 j 个点的空间坐标；初始估计值利用上面线性求解的结果，畸变系数、初始值为0。

综上，该方法基本步骤如下所示：

①制作棋盘格标定板；

②通过移动或旋转棋盘格标定板，相机拍摄记录不同方向、角度的标定板图像；

③检测图像特征点；

④通过旋转矩阵的正交性及旋转向量模为1的性质，求解不考虑畸变的影响下线性方程组，获得相机内部参数与外部参数；

⑤通过最小二乘法求解相机径向畸变系数；

⑥计算重投影误差，利用极大似然估计，迭代优化上述步骤求解的相机参数。

由上述步骤可知，该方法在已知标定板尺寸情况下，仅需要移动标定板拍摄相关图像即可完成标定，步骤简单，对标定板尺寸精度要求不高，且标定的相机参数保持了较好的精度。

2.4.3 MATLAB相机标定实例

(1) 相机标定板制作

参照棋盘格布局在计算机上画出 10×7 (25 mm \times 25 mm)的棋盘格，并用纸打印出来粘贴到板上，如图 2-16示。

图 2-16 制作棋盘格

(2) 基于MATLAB的相机标定

利用MATLAB R2020a的相机标定工具箱对本文试验用相机进行标定计算。采用相机拍摄黑白格标定板不同姿态和距离的9幅图像。具体标定如图 2-17示。

(a) Matlab相机标定角点提取结果

(b) Matlab相机标定图像平均误差

图 2-17 基于Matlab工具箱的相机标定需中文，文中解释该图了？

相机内部参数矩阵为：内参和畸变参数是相机的固有属性，只要相机硬件参数没有变化，内参和畸变参数就不变。

该相机标定实例所用的MATLAB程序如下。

```
% Create a set of calibration images.
images = imageDatastore(fullfile(toolboxdir('vision'), 'visiondata', 'calibration', 'mono'));
imageFileNames = images.Files;
% Detect calibration pattern.
[imagePoints, boardSize] = detectCheckerboardPoints(imageFileNames);
% Generate world coordinates of the corners of the squares.
squareSize = 29; % millimeters
worldPoints = generateCheckerboardPoints(boardSize, squareSize);
% Calibrate the camera.
I = readimage(images, 1);
imageSize = [size(I, 1), size(I, 2)];
[params, ~, estimationErrors] = estimateCameraParameters(imagePoints, worldPoints, ...
'ImageSize', imageSize);
figure;
showExtrinsics(params, 'CameraCentric');
figure;
showExtrinsics(params, 'PatternCentric');
figure;
showReprojectionErrors(params);
```

然后可以利用获得的参数对原图片进行畸变校正等处理，MATLAB程序示例如下：

% 导入相机标定结果

load cameraParams.mat;

% 读入一张待校正的图像

I = imread('test_image.jpg');

% 对图像进行校正

[J, newOrigin] = undistortImage(I, cameraParams);

% 显示校正前后的图像

figure;

subplot(1,2,1);

imshow(I);

title('Original Image');

subplot(1,2,2);

imshow(J);

```

title('Corrected Image');
% 获取校正后图像中某一点的三维坐标
imagePoints = [320, 240];
worldPoints = pointsToWorld(cameraParams, [0, 0, 0], imagePoints);
% 显示校正后的点在三维坐标系中的位置
figure;
plotCamera('Location', cameraParams.Location, 'Orientation', cameraParams.Orientation, 'Size', 10);
hold on;
plot3(worldPoints(:,1), worldPoints(:,2), worldPoints(:,3), 'r*');
hold off;
xlabel('X');
ylabel('Y');
zlabel('Z');
title('Corrected Point in 3D Space');

```

习题

2-1 镜头的基本参数有哪些，这些基本参数的定义是怎样的？

2-2 相机成像模型涉及哪些坐标系？

2-3 世界坐标系与相机坐标系是如何转换的？

2-4 图像像素坐标系与世界坐标系是如何转换的？

2-5 什么是相机的标定？

2-6 简述张正友相机标定法的基本步骤。

2-7 尝试利用棋盘格，使用MATLAB相机标定工具箱对手机摄像头进行标定，获得其相关参数。

4. 计算机视觉教材_第4部分

总字符数：8870

相似文献列表

去除本人文献复制比：50.3%(4466) 去除引用文献复制比：50.3%(4466) 文字复制比：50.3%(4466)

1	<u>基于双目视觉的机械臂目标定位与抓取</u> 王朋强(导师：宋丽梅;王玉祥) - 《天津工业大学硕士论文》 - 2015-12-05	10.7% (952) 是否引证：否
2	<u>基于变分图像分解的ESPI与FPP条纹分析新方法研究</u> 朱新军(导师：唐晨) - 《天津大学博士论文》 - 2014-12-01	8.1% (722) 是否引证：否
3	<u>基于人体实体切片的图像处理及三维重建技术研究</u> 隆兴银(导师：樊庆文) - 《四川大学硕士论文》 - 2005-05-10	5.9% (520) 是否引证：否
4	<u>基于MATLAB的车牌识别研究毕业论文下载</u> - 《互联网文档资源 (https://ishare.iask.com/)》 - 2021	5.5% (488) 是否引证：否
5	<u>手腕部骨图像频域祛噪在无胶片化骨龄评价系统中的应用</u> 李宗营(导师：何文杰) - 《河北工业大学硕士论文》 - 2004-05-01	5.0% (445) 是否引证：否
6	<u>数字图像处理ppt1 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.com/)》 - 2020	4.9% (436) 是否引证：否
7	<u>基于摄像头寻迹的四旋翼飞行器设计_图文</u> - 《互联网文档资源 (http://wenku.baidu.com/)》 - 2016	4.9% (431) 是否引证：否
8	<u>摄影测量中标记点的识别与匹配方法研究</u> 苏博(导师：李占利) - 《西安科技大学硕士论文》 - 2009-04-28	4.7% (419) 是否引证：否
9	<u>LPR系统中车牌的定位与提取 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.com/)》 - 2020	4.7% (414) 是否引证：否
10	<u>基于文字特征的规则文档碎片拼接技术研究</u> 田海平(导师：杨晟院) - 《湘潭大学硕士论文》 - 2015-05-05	4.6% (407) 是否引证：否
11	<u>图像边缘检测技术的研究与应用 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.com/)》 - 2020	4.6% (404) 是否引证：否
12	<u>中值滤波与均值滤波法在条形码去噪中的应用</u> 蔡学森;戴金波;李晓宁; - 《长春师范学院学报(自然科学版)》 - 2008-08-20	4.3% (385) 是否引证：否
13	<u>基于多级分类器的人脸检测系统研究</u>	4.3% (384)

	俞晶晶(导师: 刘晓平) - 《北京邮电大学硕士论文》 - 2011-01-26	是否引证: 否
14	数字图像处理ppt1(郭文强版本)-百度文库 - 《互联网文档资源 (http://wenku.baidu.c)》 - 2012	4.2% (372) 是否引证: 否
15	基于机器视觉的轴承密封圈尺寸测量系统研究及实现 赵亚康(导师: 汤文成) - 《东南大学硕士论文》 - 2021-06-30	4.1% (368) 是否引证: 否
16	数字图像处理技术在视听资料证据中的应用 宋涛(导师: 操晓春; 姚鹏海) - 《天津大学硕士论文》 - 2010-11-01	4.0% (353) 是否引证: 否
17	基于MATLAB的车牌识别研究 - 《互联网文档资源 (https://www.doc88.co)》 - 2020	3.8% (335) 是否引证: 否
18	电力系统远程操纵定轨巡视机器人 柳洲(导师: 王少荣) - 《华中科技大学硕士论文》 - 2007-06-01	3.5% (312) 是否引证: 否
19	学术论文: SAR幅值图像相干斑抑制的变分PDE模型与算法研究(文字格式, 可编辑) - 豆丁网 - 《互联网文档资源 (http://www.docin.com)》 - 2012	3.4% (300) 是否引证: 否
20	SAR幅值图像相干斑抑制的变分PDE模型与算法研究 - 豆丁网 - 《互联网文档资源 (http://www.docin.com)》 - 2017	3.4% (300) 是否引证: 否
21	数字化图像恢复在无胶片化骨龄评价系统中的应用 孙飞(导师: 何文杰) - 《河北工业大学硕士论文》 - 2004-05-01	3.4% (298) 是否引证: 否
22	海上倒驳船舶检测系统设计与研究 原康乐(导师: 魏建国; 孙提) - 《天津大学硕士论文》 - 2019-11-01	3.3% (290) 是否引证: 否
23	SAR幅值图像相干斑抑制的变分PDE模型与算法研究 尹寿垚(导师: 肖亮) - 《南京理工大学硕士论文》 - 2009-05-01	3.3% (290) 是否引证: 否
24	数字图像处理ppt1 - 道客巴巴 - 《互联网文档资源 (https://www.doc88.co)》 - 2020	3.2% (288) 是否引证: 否
25	基于5G和工业操作系统在芯片封装领域的机器视觉检测应用 - 《自动化博览》 - 2022-04-15	3.2% (285) 是否引证: 否
26	基于视觉的精密定位系统研究及应用 刘今越(导师: 李铁军) - 《河北工业大学硕士论文》 - 2005-02-01	2.9% (255) 是否引证: 否
27	数字化图像恢复在无胶片化骨龄评价系统中的应用 - 道客巴巴 - 《互联网文档资源 (http://www.doc88.com)》 - 2020	2.7% (243) 是否引证: 否
28	基于图像获取的自动机运动规律数据处理方法研究 刘喆(导师: 郑宾) - 《中北大学硕士论文》 - 2011-06-02	2.5% (218) 是否引证: 否
29	图像处理的灰度化及特征区域的获取 张桂华; 冯艳波; 陆卫东; - 《齐齐哈尔大学学报》 - 2007-07-15	1.5% (133) 是否引证: 否
30	基于混沌系统观测器的保密通信研究 沈少捷(导师: 韩正之) - 《上海交通大学硕士论文》 - 2011-01-01	1.3% (113) 是否引证: 否
31	隧道应急通信系统设计实现及在电力隧道中的应用 肖瀚; - 《电气制造》 - 2013-05-25	1.0% (87) 是否引证: 否
32	基于多层卷积滤波与HSV颜色提取的茶轮斑病识别研究 陈钊; 王子辉; 赵玉清; 高彦玉; 时玲; - 《湖北农业科学》 - 2018-06-10	0.8% (75) 是否引证: 否
33	基于内容的图像检索在移动平台上的研究与设计 郑洁颖(导师: 周玲玲) - 《上海交通大学硕士论文》 - 2007-12-01	0.8% (72) 是否引证: 否
34	航天发动机喷雾实验系统的喷雾性能监测研究 王淑姣(导师: 朱凌云) - 《东华大学硕士论文》 - 2022-05-30	0.7% (63) 是否引证: 否
35	一种抗混叠滤波器的设计 郭红玉; - 《电子设计工程》 - 2015-02-05	0.6% (56) 是否引证: 否
36	深水多波束回波强度数据处理技术探讨 马金凤; 罗伟东; 刘胜旋; 郭军; - 《地质学刊》 - 2015-12-28	0.6% (54) 是否引证: 否
37	物联网技术在火灾预警中的应用	0.5% (45)

李振雷(导师：李兰) - 《青岛理工大学硕士论文》 - 2010-12-01	是否引证：否
38 无线胶囊内窥镜系统的研究	0.4% (34)
邱祥玲(导师：颜国正) - 《上海交通大学硕士论文》 - 2010-01-01	是否引证：否

原文内容

第3章数字图像处理

3.1数字图像基础

3.1.1图像的数字化

如图3- 1所示，一维信号处理中，模拟信号数字化（即“模数变换”）包括抽样、量化和编码三个基本过程。抽样是以相等的间隔时间来抽取模拟信号的样值，把时间连续的模拟信号转换成时间上离散、幅度连续的抽样信号；量化是把抽样信号分级“取整”，从而把时间离散、幅度连续的抽样信号转换成时间离散、幅度离散的数字信号；编码是把量化的数值用一组二进制的数码来表示，完成模拟信号到数字信号的转换。需要注意的是，抽样信号虽然是时间轴上离散的信号，但仍是模拟信号，其样值在一定的取值范围内，可有无限多个值，无法一一给出对应的数字码组，因此必须进行量化。量化后的数字信号与量化前的抽样信号相比较，会有所失真。这种量化失真在接收端还原模拟信号时表现为噪声，并被称为量化噪声。量化噪声的大小取决于把样值分级“取整”的方式，分的级数越多，即量化级差或间隔越小，量化噪声也越小。

图3- 1 模拟信号的数字化示例

与一维信号一样，二维图像在本质上是模拟量，特别是在计算机应用尚未普及之前，图片多以模拟图像形式出现。随着计算机时代的到来，数字计算成为了一种主流，相应的图像处理方式也就发生了转变，数字图像模型应运而生。

在计算机中，图形由像素组成，一幅图像被分为 $M \times N$ 个方格，每个方格即为像素。每个像素具有两个属性：坐标和幅度。坐标由行、列表示；幅度值表示该像素位置上亮暗程度的整数。

图3- 2 数字图像示意图

图像经过数字化转化为数字图像，它是将模拟图像信息转化为数字信息的过程，过程如图3-3所示。

图3- 3 模拟图像到数字图像的转换

在图像采集中，传感器输出连续变化的电压信号，得到的是模拟信号图像。一幅图像可以被定义为一个二维函数 $f(x, y)$ ，其中 (x, y) 是空间（平面）坐标，在任何坐标 (x, y) 处的幅度 f 被定义为图像在这一位置的亮度。模拟图像在 x 和 y 坐标以及在幅度变化上是连续的，要将模拟图像转换成数字形式，要求对坐标和幅度进行数字化。将坐标值数字化称为取样，将幅值数字化称为量化。因此，当 x 、 y 及幅值 f 都是有限且离散的量时，我们称图像为数字图像。

下面，重点介绍图像数字化的采样和量化两个过程。

（1）采样

采样是把一幅连续图像在空间上离散成 $M \times N$ 个网格，是实现图像数字化的重要环节，决定着图像的空间分辨率。Nyquist 采样定理规定，在离散区间采样信号时，采样频率必须不小于最大输入信号频率的两倍。

当采样频率过低时，会出现“混叠”现象。混叠是指取样信号被还原成连续信号时产生彼此交叠而失真的现象。当混叠发生时，原始信号无法从取样信号还原。以图3- 4为例，可以观察到采样频率明显低于输入信号的频率，导致采样得到的波形与输入信号有较大的差异。混叠可以通过提高采样频率，或者使用抗混叠滤波器来解决。

图3- 4 正弦波采样混叠

当采样频率过高时，会出现“过采样”现象。如图3- 5所示，当15Hz的采样频率已经能较好地表现出正弦波的形式，200Hz采样点则更加密集。在较为规律的正弦波中影响并不大，但是在实际信号采集过程中，过采样会采集到大量的噪声点，从而降低采样效率。

图3- 5 正弦波过采样

一般来说，采样间隔越大，所得图像像素数越少，空间分辨率低，质量差，严重时出现马赛克效应；采样间隔越小，所得图像像素数越多，空间分辨率高，图像质量好，但数据量大。

空间分辨率可以通过PPI(Pixels Per Inch)度量。PPI也叫像素密度，表示每英寸（1英寸=2.54厘米）的所包含的像素点数量。例如：若1英寸包含10个像素点，则图像分辨率是10PPI；若1英寸包含20个像素点，则图像分辨率就是20PPI，如图3-4所示。我们发现，随着PPI数值的增大，图像效果越来越好。

图3- 6 数字图像示意图

与一维信号一样，二维图像信号的采样也需要遵循采样定理。也就是说，二维信号的采样需要满足奈奎斯特采样定律，才能产生满足要求的数字图像。在满足采样定理的前提下，采样数量不同，图像的分辨率不同。当横向的像素数（行数）为 M ，纵向的像素数（列数）为 N 时，对同一幅图像分别采取 M 、 N 值不同的采样，其结果如图3-5所示。

图3- 7 采样示例

图3- 8 二维图像不同分辨率采样示例

同样的，二维图像在采样频率过低时也会出现混叠和过采样的现象。如图3- 8所示。

（2）量化

在时间和空间上，采样后的模拟图像被离散化为像素。但是像素灰度值依然是连续的，不便于处理，因而需要量化。所谓量化，就是将每个像素点的灰度值从模拟量转化为离散量的过程，并可选用离散量代替像素的灰度值，使像素看起来简单方便，有利于进行各种操作。量化可分为线性量化和非线性量化。

一幅数字图像中不同灰度值的个数称为灰度级。灰度级常用2的整数次幂（即 2^k ）表示（最通常是8比特量化，即 $2^8=256$ ）。当图像的采样点数一定时，采用不同量化级数的图像质量也不一样。量化等级越多，所得图像层次越丰富，灰度分辨率高，图像质量好，但数据量大；量化等级越少，图像层次欠丰富，灰度分辨率低，会出现假轮廓现象，图像质量变差，但数据量小。

如图3- 9所示，对一幅采样参数设置相同的图像，分别用不同的灰度级去量化，即将灰度级从256减小到2（ k 值从8减小到

1)，结果显而易见。伪轮廓是指出现明显的类似于等高线的不连续过渡带，从而对于画面感观质量造成影响的现象；伪轮廓通常是在以16（k值为4）或更少级数的均匀设置的灰度级显示的图像中十分明显。

图3- 9 量化示例

3.1.2 数字图像的表达

将一幅图像视为一个二维函数 $f(x, y)$ ，其中 x 和 y 是空间坐标，在 x - y 平面上任意一对空间坐标 (x, y) 的幅值 f 被称为该点图像的灰度、亮度或强度。如果 f 、 x 、 y 均为非负有限离散值，则称该图像为数字图像。用函数 $f(x, y)$ 定义数字图像仅适用于最为一般的情况，即静态的灰度图像。更严格地说，对于静态图像，用2个自变量的函数 $f(x, y)$ 表示数字图像；对于动态画面，还需要时间参数 t ，即用3个自变量的函数 $f(x, y, t)$ 表示数字图像。对于灰度图像，函数值是一个数值；而对于彩色图像，函数值是一个向量。

从线性代数和矩阵论的角度，数字图像是一个由图像信息组成的二维矩阵，矩阵的每个元素均代表对应位置上的图像亮度或彩色信息。当然，数字图像在数据和存储上可能需要一个三维矩阵来表示，这样，每个单位位置的图像信息需要三个数值来表示。因此，数字图像处理的实质是对矩阵进行各种运算和处理。也就是说，将原始图像变为目标图像的过程，实际上是由一个矩阵变为另一个矩阵的数学过程。无论是图像的点运算、几何运算、图像的统计特征还是傅里叶变换等正交变换，本质上都是基于图像矩阵的数学运算。

数字图像具有多种表达形式，常见的有三种：

- (1) 函数：二维函数 $f(x, y)$ ，即为 x - y 平面坐标系上面的点。
- (2) 矩阵：可以写成一个二维的 $M \times N$ 矩阵（其中 M 和 N 分别为图像的行数和列数），矩阵中的每个元素对应一个像素
- ：(3) 矢量：数字图像也可以用矢量来表示，各个矢量中的每个元素分别对应一个像素，一般写成如下形式：其中

， $i=1, 2, 3, \dots, N$ ， f_{Mi} 为像素值。

3.1.3 数字图像的类型

常用的数字图像按照颜色和灰度的多少可以分为二值图像、灰度图像和RGB图像三种基本类型，如图3-7所示。

(a) 二值图像 (b) 灰度图像 (c) RGB图像

图3- 10 常用数字图像类型

(1) 二值图像 (Ref: 机器视觉与机器学习)

一幅二值图像的二维矩阵仅由0、1构成，“0”代表黑色，“1”代表白色，所以计算机中二值图像的数据类型采用一个二进制位表示。二值图像通常用于文字、线条图的扫描识别和掩膜图像的存储。其优点是占用空间少，缺点是当表示人物、风景的图像时，二值图像只能描述其轮廓，不能描述细节，这时候要用更高的灰度级。

在数字图像处理中，二值图像占有非常重要的地位，在实际应用中以二值图像处理实现构成的系统是很多的。图像的二值化处理是将图像上像素的灰度值设置为0或255（或者1），让整个图像呈现出明显的黑白效果，即将256个亮度等级的灰度图像通过适当的阈值选取而获得仍然可以反映图像整体和局部特征的二值图像。

为了进行二值图像的处理与分析，首先要把灰度图像二值化。二值图像的集合性质只与灰度值为0或255的像素位置有关，而不再涉及像素的多级值，使处理变得简单。为了得到理想的二值图像，一般采用封闭、连通的边界条件定义不交叠的区域。所有灰度值大于或等于阈值的像素被判定为属于特定物体，其灰度值设为255，否则这些像素被排除在物体区域以外，灰度值设为0，表示背景或者另外的物体区域。如果某个物体内部有均匀一致的灰度值，并且其处在一个具有其它等级灰度值的均匀背景下，使用阈值法就可以得到较好的分割效果。如果物体同背景的差别不表现在灰度值上（比如纹理不同），则可以将这个差别特征转换为灰度的区别，然后利用阈值法来分割该图像。

图像二值化是图像分析与处理中最常见、最重要的处理手段之一。最常见的二值化处理方法是计算像素灰度值的平均值 T ，扫描图像的每个像素灰度值并与 T 比较，大于 T 时设为1（白色），否则设为0（黑色），即

(3-1)

公式(3-1)使用平均值作为二值化阈值可能导致部分物体像素或者背景像素丢失。为了解决此问题，可使用直方图法来寻找二值化阈值。直方图是图像的重要统计特征，用直方图法选择二值化阈值主要是先找出图像的两个最高峰，然后将阈值取为这两处之间的波谷值。

Ostu方法由日本学者大津（Ostu）提出，也叫最大类间方差法。它主要依据图像的灰度特性将图像分成背景和目标两部分。背景区域和目标区域之间的类间方差越大，其差别也越大。当部分目标被误判为背景或部分背景被误判为目标时都会导致这两部分的差别变小，因此使得类间方差最大的阈值分割意味着误判的概率最小。

(2) 灰度图像 (Ref: 基于MATLAB的数字图像处理)

灰度图像一般是指具有256级灰度值的数字图像，即8bit灰度图像。灰度图像只有灰度颜色。灰度图像矩阵元素的取值范围为 $[0, 255]$ ，数据类型为8位无符号数，“0”代表纯黑色，“255”代表纯白色，0~255之间的数字由小到大表示从纯黑色到纯白色之间的过渡色。二值图像可以看成是灰度图像的一个特例。

图像由像素构成，每个像素实际上采用二进制位（bit）来表示。以灰度图像为例，若灰度图像的每个像素值用8bit来表示，则表明该图像的每个像素有28种可能的值，习惯用0~255这256个数值表示，每个数值表示一种灰度的级别。0表示黑色，255表示白色。这样的8bit的灰度图像称为8位图像。通常所说的8位灰度图像，即为256色的灰度图像。

(3) RGB图像 (Ref: 基于MATLAB的数字图像处理)

RGB图像又称真彩色图像，是指在组成一幅彩色图像的每个像素值中，分别用红（R）、绿（G）、蓝（B）三原色的组合来表示每个像素的颜色。由于数字图像以二维矩阵表示，每个像素的颜色需由R、G、B三个分量来表示，因此RGB图像矩阵需要采用三维矩阵表示，即 $M \times N \times 3$ 矩阵， M 、 N 分别表示图像的行数和列数，3个 $M \times N$ 的二维矩阵分别表示各个像素的R、G、B颜色分量。理论上RGB图像中每个像素所表示的颜色可多达224（ $28 \times 28 \times 28$ ）种。RGB图像的数据类型一般为8位无符号整型。

综上所述，在数字图像的三种基本类型中，随着表示颜色类型的增加，数字图像所需的存储空间逐渐增加。二值图像仅能表示黑、白两种颜色，所需的存储空间最少；灰度图像可以表示由黑色到白色渐变的256个灰度级，每个像素需要一个字节存储空间；RGB图像可以表示224种颜色，相应每个像素需要3个字节的存储空间，是灰度图像的3倍。表3-1给出了数字图像三种基本类型的对比。

假设有一幅 512×512 像素的图像，256个灰度级（8个比特bit表示，即一个字节byte），则图像需要 512×512 个字节存储。如果是彩色视频（每秒30帧，每帧 512×512 ），那么一秒钟的视频数据量为 $512 \times 512 \times 1 \times 3 \times 30$ 个字节。

表3-2 数字图像三种基本类型对比

类型	二值图像	灰度图像	RGB图像
颜色数量	2	256	224
数据类型	1bit	8bit	24bit
矩阵大小	$M \times N$	$M \times N$	$M \times N \times 3$

类型二值图像灰度图像 RGB图像

颜色数量 2 256 224

数据类型 1bit 8bit 24bit

矩阵大小 $M \times N$ $M \times N$ $M \times N \times 3$

3.1.4 像素间关系（Ref:计算机视觉）（可同时参考：计算机视觉教程（微课版））

1. 像素的相邻和邻域

图像中的像素的相邻和邻域有3种，如图3- 11所示。

图3- 11 像素邻域示例

（1）4邻域

设相对于图像显示坐标系的图像中的像素P位于 (x, y) 处，则P在水平方向和垂直方向相邻的像素最多可有4个，其坐标分别为

$(x-1, y)$, $(x, y-1)$, $(x, y+1)$, $(x+1, y)$

由这4个像素组成的集合称为像素P的4邻域，记为 $N_4(P)$ 。像素的4邻域如图3-8(a)所示。

（2）4对角邻域

设相对于图像显示坐标系的图像中的像素P位于 (x, y) 处，则P在对角相邻的像素最多可有4个，其坐标分别为

$(x-1, y-1)$, $(x-1, y+1)$, $(x+1, y-1)$, $(x+1, y+1)$

由这4个像素组成的集合称为像素P的4对角邻域，记为 $N_D(P)$ 。像素的4对角邻域如图3-8(b)所示。

（3）8邻域

把像素P的4对角邻域和4邻域像素组成的集合称为像素P的8邻域，记为 $N_8(P)$ 。像素P的8邻域如图3-8(c)所示。

2. 像素间距离的度量

像素间的距离有以下几种度量方式，示例如图3- 12所示，图中的数字表示数字所处像素位置与0所在位置的距离。

图3- 12 像素距离示例

（1）欧氏距离DE

（3-2）

根据式0.1的距离度量，所有距像素点 (x, y) 的欧氏距离小于或等于DE的像素都包含在以 (x, y) 为中心，以DE为半径的圆平面中。

欧氏距离的优点是直观且显然，缺点是平方根的计算费时且数值可能不是整数。

（2）街区距离D4

（3-3）

距离表示在数字栅格（即像素）上从起点移动到终点所需的最少的基本步数。如果只允许横向和纵向的移动，则得到距离D4，D4也称为街区距离，因为它类似于在具有栅格关街道的封装房子块的城市里的两个位置的距离。

（3）棋盘距离D8

（3-4）

在数字格栅中如果允许对角线方向的移动，则得到距离D8，D8也称为棋盘距离。距离D8等于国王在棋盘上从一处移动到另一处所需的步数。

3.1.5 图像中的噪声参考文献格式，一定注意别人知识产权的保护

实际的图像会受到随机误差的影响而退化，通常称这个退化为噪声。在图像的捕获、传输或处理过程中都可能出现噪声，噪声可能依赖于图像内容，也可能与其无关。噪声一般由其概率特征来描述。

理想的噪声称为白噪声。白噪声是具有常量的功率谱，也就是说噪声在所有频率上出现且强度相同。白噪声是常用的模型，作为退化的最坏估计，不过它可以简化计算。

白噪声的一个特例是高斯噪声。服从高斯（正态）分布的随机变量具有高斯曲线型的概率密度。在一维的情况下，概率密度函数是

（3-5）

其中， μ ， σ 分别是随机变量的均值和标准差。在很多实际情况下，噪声可以很好地用高斯噪声来近似。

当图像通过信道传输时，噪声一般与出现的图像信号无关。这种独立于信号的退化被称为加性噪声，可以用下列模型来表

示
$$f(x, y)=g(x, y)+v(x, y) \quad (3-6)$$

其中，噪声v和输入图像g是相互独立的变量。

计算噪声贡献的所有平方和：

（3-7）

计算观察到的信号的所有平方和：

（3-8）

信噪比可以定义为： $SNR=S/N$ 。信噪比常用对数尺度来表示，单位为分贝：

(3-9)

3.2 图像滤波

图像滤波的目的是消除或抑制图像中的噪声，从而实现图像增强。噪声的产生有很多种方式，可能由图像传输过程中的信号干扰、相机自身的原因和拍摄过程中的抖动等造成。噪声主要分为高斯噪声、脉冲噪声（也称椒盐噪声）、散斑噪声等。一个好的图像滤波算法会在尽可能去除噪声的同时最大限度保留图像细节信息。最常见和最基本的图像滤波方法有均值滤波、中值滤波、高斯滤波、BM3D滤波（Block Matching and 3D filtering、三维块匹配滤波）和双边滤波等。

1. 均值滤波

均值滤波是线性滤波中最简单的一种，处理之后的图像像素值是根据要处理的像素领域的像素值来决定的，即每一个像素值用该像素领域中所有像素的灰度平均值来代替。均值滤波的操作可以表示为

(3-10)

式中， $b(x, y)$ 是均值滤波之后图像上的像素灰度值； $a(r, c)$ 是输入图像的像素灰度值，即要进行均值滤波的图像； m 、 n 为所用模板的大小； T_{xy} 为所使用的均值滤波模板； (r, c) 为均值滤波模板中的像素坐标。

常用的均值滤波模板有两种，第一种是计算模板内像素灰度值的平均值

(3-11)

第二种是对模板所覆盖的像素灰度值加上了权重，即每个像素值对结果的影响不一样，权重大的像素对结果的影响比较大，具体如

(3-12)

均值滤波在对图像进行平滑的同时会把阶跃变化的灰度值平滑为渐变变化，这就造成了图像细节信息的严重丢失，将导致边缘提取定位精确度的下降。

2. 中值滤波

图像中的中值滤波是另一种用来消除图像噪声以减少其对后续处理影响的操作。它使用的也是一个模板，在图像上平移模板并对模板内的像素灰度值按照大小进行排序，然后选取排在中间位置的数值，将它赋值给图像的待处理像素。如果模板由奇数个元素组成，那么中值就取排序之后中间位置元素的灰度值。如果模板有偶数个元素，中值就取排序之后中间两个灰度值的平均值。中值滤波的模板通常采用奇数个元素，这样便于计算和编程实现。对于边界的处理，可以将原图像的像素直接复制到处理之后图像的对应位置，或者将处理之后的图像边界像素灰度值直接改为0。中值滤波可以表示为

(3-13)

中值滤波在去除脉冲噪声（即椒盐噪声）方面很有效，同时又能保留图像的细节特征，如边缘信息。

3. 高斯滤波

高斯滤波属于频域滤波，它是由高斯函数的形状来选择权值的。高斯滤波对一维的高斯分布通常表示为

(3-14)

二维的分布函数为

(3-15)

式中， σ 是标准差。

对于图像处理，通常使用二维的高斯函数。图像进行高斯滤波之后的平滑程度和标准差有很大的关系：标准差越大，高斯滤波器的频带就越宽，图像就被平滑得越好。通过调节标准差可以很好地处理图像中噪声引起的欠平滑。由于二维高斯函数的旋转对称性，高斯滤波在每个方向上的平滑程度是相同的。对于一幅图像，计算机无法事先知道图像的边缘方向信息，因此高斯滤波是无法确定在哪个方向上需要做更多平滑的。高斯函数具有可分离性，使得高斯函数卷积可以分为两步来实现，首先用一维高斯函数和图像进行卷积，然后将卷积的结果与另一个方向的一维高斯函数卷积，这样可以将算法的时间复杂度从 $O(n^2)$ 降低到 $O(n)$ ，从而大大提高计算效率。

理论上，高斯分布在任何位置都是非零值，但是如果这样的话，高斯模板将是一个无限大的模板，因此在实际应用中，高斯模板的构建只要满足取值在均值的3倍标准差之内就可以。通常高斯模板的大小为 3×3 或 5×5 ，它们的权值分布为：

和

4. BM3D滤波

BM3D滤波是一种性能优越的图像滤波算法，包含了非局部和变换域两种思想。该方法通过与相邻图像块进行匹配，将若干相似的块整合为一个三维矩阵后在三维空间进行滤波，再将结果反变换融合到二维，得到滤波后的图像。作为一种非常有效的图像滤波算法，BM3D滤波成为其它新的滤波算法竞相比较的对象。此外，BM3D滤波还扩展到了图像处理的其它领域，如图像去模糊、压缩传感和超分辨率重构等。

BM3D滤波算法的实现分为两个步骤：

(1) 第一步：基础估计。

首先将图像窗口化，设定若干参考块，计算图像参考块和其它图像块之间的距离，再根据这些距离寻找若干差异最小的块作为相似块，将相似块归入对应组，形成一个三维矩阵。得到若干个参考块的三维矩阵后，首先将每个矩阵中的三维块进行二维变换编码（可采用Wavelet变换、DCT变换等）。二维变换结束后，在矩阵的第三个维度进行一维变换。通过分组和滤波得到的每一个块的估计可能是重叠的，所以需要对这些重叠的块进行加权平均，这一过程称为聚集。

(2) 第二步：最终估计。

第二步与第一步类似，但在块匹配时是用第一步的结果图即基础估计进行匹配。通过块匹配，每个参考块形成两个三维矩阵：一个是通过基础估计形成的，另一个是通过本次匹配的坐标在噪声图上整合出来的。然后两个三维矩阵都进行二维、一维变换。为了获得更好的结果，通常最终估计的二维变换采用离散余弦变换。用wiener滤波对噪声图形成的三维矩阵进行缩放，缩放系数通过基础估计的三维矩阵值以及噪声强度得出。滤波后再通过反变换将噪声图的三维矩阵变换回图像估计，之后通过与第一步类似的聚集操作复原出二维图像而形成最终估计，这样就得到了BM3D滤波后的图像。

5. 双边滤波

二维图像的双边滤波算法是指利用当前待处理像素领域内各个像素灰度值的加权平均值来代替当前像素的灰度值，加权平均过程中采用的权重因子不仅与两像素间的欧式距离有关，也与两像素的灰度值差异有关。双边滤波算法的优点是既可以对图像噪声进行抑制，又可以有效保留图像的边缘细节特征。

5. 计算机视觉教材_第5部分		总字符数：9996
相似文献列表		
去除本人文献复制比：27.6%(2756) 去除引用文献复制比：27.6%(2756) 文字复制比：27.6%(2756)		
1	<u>自行走无损检测小车视觉导航系统研究</u> 李庆伟(导师：王渝) - 《北京理工大学硕士论文》 - 2008-06-15	7.4% (744) 是否引证：否
2	<u>常用边缘检测算法的对比分析</u> - 《网络 (http://www.wendangku) 》 - 2020	6.1% (612) 是否引证：否
3	<u>基于DSP的图像边缘检测的研究</u> 王勇(导师：刘继承) - 《大庆石油学院硕士论文》 - 2009-03-15	5.9% (586) 是否引证：否
4	<u>基于DSP的图像边缘检测的研究 - 道客巴巴</u> - 《互联网文档资源 (http://www.doc88.com) 》 - 2019	5.7% (571) 是否引证：否
5	<u>常用边缘检测算法的对比分析-百度文库</u> - 《互联网文档资源 (http://wenku.baidu.c) 》 - 2012	5.6% (557) 是否引证：否
6	<u>毕业设计(论文)图像处理技术的设计 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co) 》 - 2020	5.5% (551) 是否引证：否
7	<u>毕业设计(论文)-数字图像边缘检测的设计下载</u> - 《互联网文档资源 (https://ishare.iask) 》 - 2020	5.5% (551) 是否引证：否
8	<u>毕业设计(论文)-数字图像边缘检测的设计 - 豆丁网</u> - 《互联网文档资源 (http://www.docin.com) 》 - 2013	5.4% (539) 是否引证：否
9	<u>超商货架商品分割与识别方法研究</u> 孙海侠(导师：袁夏) - 《南京理工大学硕士论文》 - 2014-01-01	5.1% (507) 是否引证：否
10	<u>图像边缘检测技术的实现及应用</u> - 《互联网文档资源 (http://www.360doc.co) 》 - 2015	5.0% (502) 是否引证：否
11	<u>基于视频分析的城市道路交通流量监测与实现</u> 诸咸斌(导师：李晓飞) - 《南京邮电大学硕士论文》 - 2013-04-01	4.7% (467) 是否引证：否
12	<u>图像边缘检测技术研究 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co) 》 - 2020	4.5% (454) 是否引证：否
13	<u>基于形态学的图像边缘检测技术研究</u> 沈阳(导师：李晓峰) - 《电子科技大学硕士论文》 - 2008-04-01	4.4% (440) 是否引证：否
14	<u>基于高拍仪的自动阅卷系统设计与实现 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co) 》 - 2020	4.3% (431) 是否引证：否
15	<u>基于高拍仪的自动阅卷系统的设计与实现</u> 徐金伟(导师：程晓荣;詹鲁安) - 《华北电力大学硕士论文》 - 2013-05-01	4.3% (431) 是否引证：否
16	<u>水下图像的目标检测与定位研究</u> 张巍(导师：彭复员) - 《华中科技大学硕士论文》 - 2007-06-01	4.1% (406) 是否引证：否
17	<u>边缘检测准则与几种典型算法</u> 李葆青;文山; - 《六盘水师范高等专科学校学报》 - 2006-06-30	3.8% (375) 是否引证：否
18	<u>基于ARM嵌入式操作系统的图像处理</u> 左春雪;吉喆明;黄小红; - 《计算机与数字工程》 - 2015-03-20	3.7% (373) 是否引证：否
19	<u>合成孔径雷达图像线特征提取算法的研究</u> 严琴(导师：蒋泽云) - 《电子科技大学硕士论文》 - 2008-04-01	3.7% (371) 是否引证：否
20	<u>图像边缘检测技术在电视节目源监控系统中的应用可行性浅论</u> 李玉贵;贾泽莘; - 《视听》 - 2012-04-15	3.5% (351) 是否引证：否

21	<u>机器视觉边缘检测技术及应用研究</u> 郭津(导师:汪仁煌) - 《广东工业大学硕士学位论文》 - 2011-05-10	3.5% (347)	是否引证: 否
22	<u>基于OpenCV的图像处理系统 课程设计报告下载</u> - 《互联网文档资源 (https://ishare.iask.)》 - 2021	3.5% (346)	是否引证: 否
23	<u>二、三维图像中一类新的边缘检测与追踪策略</u> 胡海波(导师:王利生) - 《上海交通大学硕士学位论文》 - 2008-01-01	3.4% (335)	是否引证: 否
24	<u>全自动无模板型磨边机采样系统设计</u> 刘明芹;张晓光; - 《计算机测量与控制》 - 2006-03-25	3.3% (330)	是否引证: 否
25	<u>二、三维图像中一类新的边缘检测与追踪策略</u> - 道客巴巴 - 《互联网文档资源 (http://www.doc88.com)》 - 2020	3.2% (323)	是否引证: 否
26	<u>常用边缘检测技术的对比</u> 张闯;孙兴波;陈瑶;黄祥; - 《传感器世界》 - 2013-11-25	2.7% (267)	是否引证: 否
27	<u>传统服饰纹样生成设计研究</u> 张帆;苏艺;崔强;王涛; - 《包装工程》 - 2023-02-20	2.3% (232)	是否引证: 否
28	<u>基于VC++的数字图像处理系统的设计与实现</u> 张华;展晓凯; - 《潍坊学院学报》 - 2011-04-15	2.1% (213)	是否引证: 否
29	<u>DCT特征与SVM分类在人脸检测中的应用</u> 刘玉景;程国建; - 《软件导刊》 - 2009-03-30	1.8% (180)	是否引证: 否
30	<u>基于一维傅里叶变换特征提取的纸币识别</u> 尹惠玲; - 《计算机与现代化》 - 2011-07-15	1.7% (171)	是否引证: 否
31	<u>基于傅里叶变换和连通图的聚类分析方法</u> 巨瑜芳;雷小锋;戴斌;庄伟;宋丰泰; - 《计算机应用研究》 - 2012-08-15	1.5% (154)	是否引证: 否
32	<u>采用机器视觉和傅里叶频谱特征的循环水养殖鱼类摄食状态判别</u> 陈雨琦;冯德军;桂福坤;曲晓玉; - 《农业工程学报》 - 2021-07-23	1.4% (135)	是否引证: 否
33	<u>基于位移时间序列Fourier分析的滑坡预警研究</u> 陈建胜;陈从新;赵海斌;于新华; - 《人民长江》 - 2013-11-28	1.3% (133)	是否引证: 否
34	<u>基于谱残余调幅相位的图像显著性检测方法</u> 蒋达央; - 《信息化研究》 - 2014-10-20	1.3% (126)	是否引证: 否
35	<u>傅立叶变换在仓库监控智能跟踪系统中的应用</u> 梁志勇;刘文娟; - 《微计算机信息》 - 2009-10-05	1.0% (104)	是否引证: 否
36	<u>基于频域特征提取的红外小目标跟踪算法</u> 戎怀阳;张涌; - 《激光与红外》 - 2012-08-20	0.8% (79)	是否引证: 否
37	<u>边缘检测算法在人脸表情识别中的应用与研究</u> 李斐;陈亚军;宋朝阳;李瑞; - 《计算机与数字工程》 - 2008-11-20	0.6% (61)	是否引证: 否
38	<u>基于语谱图的高频分量进行音频篡改研究</u> 王鹏;周成成;吴文琪;唐依静;孙世凡; - 《信息技术与信息化》 - 2022-09-25	0.6% (58)	是否引证: 否
39	<u>基于二维DCT的电能质量监测数据压缩方法</u> 胡志坤;何志敏;安庆;孙克辉;丁家峰; - 《中南大学学报(自然科学版)》 - 2011-04-26	0.5% (51)	是否引证: 否
40	<u>利用原子力显微镜表征页岩孔隙结构特征</u> 余昊;沈瑞;郭和坤;王国栋;邵国勇;尚祯浩; - 《科学技术与工程》 - 2022-12-28	0.5% (46)	是否引证: 否
41	<u>基于视觉伺服反馈控制的车辆尺寸参数测量系统</u> 卞晓东,张为公,郭占军 - 《测控技术》 - 2003-05-18	0.4% (38)	是否引证: 否
42	<u>基于改进MSR的锂电池X射线图像增强算法</u> 钱玉洋;魏巍;陈灯; - 《电子测量技术》 - 2022-05-20 07:56	0.3% (33)	是否引证: 否

原文内容

假设 p 是数字图像 I 中的当前待处理像素,则二维图像双边滤波算法为
(3-16)

式中, $I_b(p)$ 是 p 经过双边滤波后的像素灰度值; q 表示 p 的领域像素点; $I(q)$ 是点 q 的像素灰度值; S 是领域像素的集合

；GS(p, q)为空间邻近度因子；Gr(p, q)为灰度相似度因子。GS(p, q)和Gr(p, q)的表达式分别为

(3-17)

(3-18)

式中，(x, y)为图像像素坐标；(u, v)为中心点像素坐标； σ_s 是基于高斯函数的空间距离标准差； σ_r 是基于高斯函数的灰度标准差。

由式(3-17)和式(3-18)可见，双边滤波算法同时考虑了当前像素与周围像素的欧式距离和灰度相似性，因此邻域中与中心点距离更近、灰度更相似的像素被赋予较大的权重，反之则赋予较小的权重，这使得双边滤波算法具有距离各向异性和灰度各向异性，可以较好地保留细节特征。

通过在原始图像中添加“高斯噪声”来模拟噪声图像，如图3-10所示；均值滤波和高斯滤波后的图像如图3-11所示。

图3- 13 噪声图像模拟

图3- 14 滤波去噪示例

图3- 13和图3- 14所用的MATLAB程序如下。

```
clc,clear,close all
I=imread('GZA Bridge_RGB.jpg');
Bridge_gray=rgb2gray(I);
Bridge_gray_gn=imnoise(Bridge_gray,'gaussian',0,0.0015); % Add Gaussian noise
k=ones(3,3)/9; % Define a 3x3 mean filter
Bridge_gray_gn_m=imfilter(Bridge_gray_gn,k); % Apply mean filter to Gaussian noised image
k=fspecial('gaussian',[5 5],2); % Define a 5x5 Gaussian filter kernel
Bridge_gray_gn_g=imfilter(Bridge_gray_gn,k); % Apply Gaussian filter to Gaussian noised image
figure
subplot(2,2,1); imshow(Bridge_gray); title('原始图像')
subplot(2,2,2); imshow(Bridge_gray_gn); title('高斯噪声图像')
subplot(2,2,3); imshow(Bridge_gray_gn_m); title('均值滤波去噪')
subplot(2,2,4); imshow(Bridge_gray_gn_g); title('高斯滤波去噪')
```

3.3 图像频域处理

数字图像处理的方法有两大类：一种是空间域处理法，另一种是频域处理法。把图像信号从空域变换到变换域进行处理，便可以从另外一个角度来分析图像信号的特性。图像的频域处理最突出的特点是其运算速度快，并可采用已有的二维数字滤波技术进行所需要的各种图像处理，因此图像的频域处理得到了广泛的应用。

数字图像的频域处理主要有三种应用：利用某些频域变换可以修改图像某些频域特性，如去噪等；利用某些频域变换可以从图像中提取图像的特征，如周期纹理等；利用图像频域处理可实现图像高效压缩编码，可减小计算位数，使算术运算次数大大减少。

数字图像的频域处理最关键的是变换处理，即首先将图像从空间域变换到频域，然后再对图像进行各种各样的处理，最后将所得到的结果进行反变换，即从频域变换到空间域，从而达到图像处理的目的。频域处理一般采用的变换方法都是线性正交变换，又称酉变换。本节将重点介绍图像的离散二维傅里叶变换、离散二维余弦变换和离散二维小波变换。

数字图像处理方法主要包括空间域处理方法（空域法）和变换域处理方法（频域法）。图像变换是将图像信号从空间域变换到另一个域。图像变换的目的是根据图像在变换域中的某些性质对图像进行加工处理。虽然这些性质在空间域中很难甚至无法获取，但是在变换域中却能很好地呈现。在有些情况下，在空间域中分析图像信号很不方便，通过图像变换后再进行分析，往往变得很容易。在数字图像处理中，图像增强、图像恢复、图像编码、图像分析和描述等都需要利用图像变换作为手段。常见的图像变换有离散傅里叶变换（DFT）、离散余弦变换（DCT）和小波变换，除此之外，还有离散沃尔什-哈达玛变换（DWHT）和离散K-L变换等。

3.3.1 傅里叶变换图像处理

图像频域变换工具主要有傅里叶（Fourier）变换、余弦变换和小波变换等，其中傅里叶变换是最基本的一种，在数字图像处理中应用广泛。傅里叶变换对图像的分解可比喻成一个玻璃棱镜对光信号的分解。棱镜是可以将光分解为不同颜色的物理仪器，每个成分的颜色由波长（或频率）来决定。傅里叶变换可以看作数学上的棱镜，能将函数基于频率分解为不同的成分，从而可以更好地识别图像中的各种成分。

图像的频率是表征图像中灰度变化剧烈程度的指标，是灰度在平面空间上的梯度。如：大面积的沙漠在图像中是一片灰度变化缓慢的区域，对应的频率值很低，而地表属性变换很大的边缘区域在图像中是一片灰度变化剧烈的区域，对应的频率值较高。傅里叶变换有非常明显的物理意义，设f是一个能量有限的模拟信号，则其傅里叶变换就表示f的谱。从纯粹的数学意义上看，傅里叶变换是将一个函数转换为一系列周期函数来处理的。从物理效果看，傅里叶变换是将图像从空间域转换到频率域，其逆变换是将图像从频率域转换到空间域。换句话说，傅里叶变换的物理意义是将图像的灰度分布函数变换为图像的频率分布函数，傅里叶逆变换是将图像的频率分布函数变换为灰度分布函数。

傅里叶变换是数字图像处理中应用最广的一种变换，图像增强、图像复原和图像分析与描述等每一类处理方法都要用到图像变换，尤其是图像的傅里叶变换。傅里叶变换将时域信号分解为不同频率的正弦余弦和的形式。它是数字图像处理技术的基础，其通过在时域和频域来回切换图像，来对图像的信息特征进行提取和分析。

数字图像f(x, y)可以表达为一个M×N的矩阵。图像矩阵的二维离散傅里叶变换可按级数形式定义为

(3-19)

式中，u, x=0, 1, 2, ..., M-1; v, y=0, 1, 2, ..., N-1。

傅里叶反变换为

(3-20)

在离散傅里叶变换对中，称为离散信号的频谱，为相位角，被称为幅度谱，它们之间的关系为(3-21) 二维离散傅里叶变

换的许多性质对于数字图像分析具有非常重要的作用，几个主要性质介绍如下：

(1) 可分离性

离散二维傅里叶变换可以分解为两个一维傅里叶变换，不论是先对行还是先对列进行一维傅里叶变换都是一样的，顺序对结果并没有影响。同样，二维傅里叶反变换也可以分成两步来求。考虑 $M=N$ 时，其变换公式如下：(3-22)式中， $u, x, v, y=0, 1, 2, \dots, N-1$ 。

完成一幅 $N \times N$ 图像的二维离散傅里叶变换，需要进行 N^2 次乘法和 $N(N-1)$ 次加法运算，计算量巨大。快速傅里叶变换(FFT)解决了这一难题。根据式(3-22)，令 $u=x, v=y$ ，则有。其中，为在 y 方向 N 点的DFT；为在 x 方向 N 点的DFT。于是一个 $N \times N$ 图像的二维DFT可转换为一次 y 方向 N 点的一维DFT和一次 x 方向 N 点的一维DFT，而这两次一维DFT均可使用快速傅里叶变换(FFT)。

(2) 平移性 (Ref. 基于MATLAB的数字图像处理)

考虑图像行数和列数相等时，即 $M=N$ 时，移位性质可表述为：(3-23)

式(3-23)表明，乘以一个指数项函数后，并取其乘积的傅里叶变换，可使频率域的中心移到点。在图像处理过程中，在傅里叶变换后，往往需要将中心移动到位置上，此时有：(3-24)

表明，在空间域中，将乘以，就可将傅里叶变换的原点移动到相应 $N \times N$ 频率方阵的中心，这个过程称为图像中心化。频谱图像的中心化能够使得频域图像更加直观，并且在频域的处理也更加方便。

(3) 共轭对称性

二维离散傅里叶变换的共轭对称性可表示为：

(3-25)

二维DFT的共轭对称性给图像的频谱分析和显示带来了极大的好处，在实际应用中，只需对其一半的频谱变换结果进行计算和分析即可。

二维离散傅里叶变换还具有很多其它的性质，如周期性和旋转不变性等，这里不再一一赘述。

图像傅里叶变换与反变换效果如图3-15所示。

图3-15 图像傅里叶变换

图3-15所用的MATLAB程序如下。

```
clc,clear,close all
I=imread('GZA Bridge_RGB.jpg');
I=I(51:358,251:558,1:3);
I_Gray=rgb2gray(I);
img1 = double(I_Gray);
img1F = fft2(img1);
img1FF = fftshift(abs(img1F));
imgRec = ifft2(img1F);
figure;
subplot(2,3,1); imshow(img1,[]); title('原始图像')
subplot(2,3,2); mesh(abs(img1F)); title('频谱')
subplot(2,3,3); mesh(img1FF); title('频谱平移')
subplot(2,3,4); mesh(log(1+img1FF)); title('对数坐标频谱')
subplot(2,3,5); imshow(imgRec,[]); title('重构图像')
subplot(2,3,6); mesh(img1-imgRec); title('重构误差')
```

3.3.2 离散余弦变换

离散余弦变换(Discrete Cosine Transform, DCT)是与傅里叶变换相关的一种变换，它类似于离散傅里叶变换，但是只使用实数。离散余弦变换相当于一个长度大概是它两倍的离散傅里叶变换。离散余弦变换的类型包括I型、II型和III型等，其中最常用的是第二种类型。它的逆被称为反离散余弦变换或逆离散余弦变换(IDCT)。

I型和II型离散余弦变换公式分别为

(3-26)

式中，为的DCT变换；为图像空间坐标；为变换域坐标； N 为图像的大小。

离散余弦变换经常在信号处理和图像处理中使用，用于对信号和图像进行有损数据压缩。这是由于离散余弦变换具有很强的能量集中特性：大多数自然信号(包括声音和图像)的能量是都集中在离散余弦变换后的低频部分。离散余弦变换常被用于JPEG图像压缩。此外，它也被经常用于求解偏微分方程。

图3-16所示为图像过DCT变换，变换系数处理以及反变换重构后的结果。

图3-16 图像离散余弦变换压缩

3.4 边缘检测

3.4.1 边缘介绍

图3-17 边缘检测案例

(1) 边缘的定义：边缘是不同区域的分界线，是周围(局部)灰度值有显著变化的像素点的集合，有幅值与方向两个属性。这个不是绝对的定义，主要记住边缘是局部特征，以及周围灰度值显著变化产生边缘。

(2) 轮廓和边缘的关系：一般认为轮廓是对物体的完整边界的描述，边缘点一个个连接起来构成轮廓。边缘可以是一段边缘，而轮廓一般是完整的。人眼视觉特性，看物体时一般是先获取物体的轮廓信息，再获取物体中的细节信息。比如看到几个人站在那里，我们一眼看过去马上能知道的是每个人的高矮胖瘦，然后才获取脸和衣着等信息。

(3) 边缘的类型：图像边缘具有方向和幅度两个特征。沿边缘走向，像素的灰度值变化比较平缓，而沿垂直于边缘的走向，像素的灰度值则变化比较剧烈。这种剧烈的变化或者呈阶跃状(step edge)，或者呈屋顶状(roof edge)，分别称为阶跃状边缘和屋顶状边缘。一般常用一阶和二阶导数来描述和检测边缘。

图3- 18 图像边缘及其导数曲线规律示例

3. 4. 2边缘检测介绍

边缘检测是图像处理与计算机视觉中极为重要的一种分析图像的方法，它的目的是找到图像中亮度变化剧烈的像素点构成的集合，表现出来往往是轮廓。如果图像中的边缘能够精确地测量和定位，那么实际的物体就能够被定位和测量（包括物体的面积、物体的直径、物体的形状等就能被测量）。

图3- 19 四种边缘类型

对于来自现实世界的图像，通常在出现下面4种情况时该区域被视为一个边缘：

- 深度的不连续（物体处在不同的平面上）；
- 表面方向的不连续（如正方体的不同的两个面）；
- 物体材料不同（这样会导致光的反射系数不同）；
- 场景中光照不同（如被阴影覆盖的地面）。

例如，图3-20中的图像是图像中水平方向6个像素点的灰度值显示效果，我们很容易地判断在第3和第4个像素之间有一个边缘，因为它们之间发生了强烈的灰度跳变。在实际的边缘检测中边缘远没有图3-20这样简单明显，我们需要取对应的阈值来区分它们。

图3- 20 图像边缘示例

3. 4. 3边缘检测的基本方法

一般图像边缘检测方法主要有如下四个步骤：

第一步：图像滤波。传统边缘检测算法主要是基于图像强度的一阶和二阶导数但导数的计算对噪声很敏感，因此必须使用滤波器来改善与噪声有关的边缘检测器的性能。需要指出的是，大多数滤波器在降低噪声的同时也造成了边缘强度的损失，因此，在增强边缘和降低噪声之间需要一个折中的选择。

第二步：图像增强。增强边缘的基础是确定图像各点邻域强度的变化值。增强算法可以将邻域（或局部）强度值有显著变化的点突显出来。边缘增强一般是通过计算梯度的幅值来完成的。

第三步：图像检测。在图像中有许多点的梯度幅值比较大，而这些点在特定的应用领域中并不都是边缘，所以应该用某种方法来确定哪些点是边缘点。最简单的边缘检测判断依据是梯度幅值。

第四步：图像定位。如果某一应用场合要求确定边缘位置，则边缘的位置可用子像素分辨率来估计，边缘的方位也可以被估计出来。

3. 4. 4边缘检测算子的概念

在数学中，函数的变化率由导数来刻画，图像可以看成二维函数，其上面的像素值变化，当然也可以用导数来刻画，当然图像是离散的那我们换成像素的差分来实现。对于阶跃型边缘，图3-18中显示其一阶导数具有极大值，极大值点对应二阶导数的过零点，也就是说，准确的边缘的位置是对应于一阶导数的极大值点或者二阶导数的过零点（注意，不仅指二阶导数为0值的位置，而且指正负值过渡的零点）。故边缘检测算子的类型当然就存在一阶和二阶微分算子。

3. 4. 5常见的边缘检测算子

图像中的边缘可以通过对它们求导数来确定，而导数可利用微分算子来计算。对于数字图像来说，通常是利用差分来近似微分。

近20多年来提出了许多边缘检测算子，在这里我们仅讨论其中常见的边缘检测算子。

如图3- 21所示，常见的一阶微分边缘算子包括Roberts、Prewitt、Sobel、Kirsch和Nevitia 等将在3.6节具体讲述，常见的二阶微分边缘算子包括Laplace算子、LOG算子和Canny算子等将在3.7节中展开讲解。其中，Canny算子是最为常用的一种，也是当前被认为最优秀的边缘检测算子。

图3- 21 图像边缘检测算子示例

图3- 21所用的MATLAB程序如下。

```
clc,clear,close all
I=imread('GZA Bridge_RGB.jpg');
I=I(51:358,251:558,1:3);
Bridge_gray=rgb2gray(I);
Bridge_Ec = edge(Bridge_gray,'canny'); % Roberts edges
Bridge_Es = edge(Bridge_gray,'sobel'); % Prewitt edges
Bridge_Ep = edge(Bridge_gray,'prewitt'); % Sobel edges
Bridge_El = edge(Bridge_gray,'log'); % LOG edges
Bridge_Er = edge(Bridge_gray,'Roberts'); % Roberts edges
figure
subplot(2,3,1); imshow(Bridge_gray); title('原始图像')
subplot(2,3,2); imshow(Bridge_Ec); title('Canny边缘')
subplot(2,3,3); imshow(Bridge_Es); title('Sobel边缘')
subplot(2,3,4); imshow(Bridge_Ep); title('Prewitt边缘')
subplot(2,3,5); imshow(Bridge_El); title('LOG边缘')
subplot(2,3,6); imshow(Bridge_Er); title('Roberts边缘')
```

3. 5. 6梯度算子介绍。

（1）边缘点：对应于一阶微分幅度的最大值点以及二阶微分的零点。

（2）梯度的定义：二个曲面沿着给定方向的倾斜程度。在单变量函数中，梯度只是导数。在线性函数中梯度是线的斜率一有方向的向量。

（3）梯度算子：梯度算子属于一阶微分算子，对应一阶导数。若图像含有较小的噪声并且图像边缘的灰度值过渡较为明显

，梯度算子可以得到较好的边缘检测结果。前篇介绍的Roberts、Sobel等算子都属于梯度算子。

3. 5. 7梯度的衡量方法

对于连续函数 $f(x, y)$ ，我们计算出了它在 (x, y) 处的梯度，并且用一个矢量（沿 x 方向和沿 y 方向的两个分量）来表示，如下：现在我们需要衡量梯度的幅值可以用到以下三种范数：值得注意的是，由于使用2范数梯度要对图像中的每个像素点进行平方及开方运算，计算复杂度高，在实际应用中，通常取绝对值或最大值来近似代替该运算以实现简化，与平方及开方运算相比，取绝对值或最大值进行的边缘检测的准确度和边缘的精度差异都很小。

图3- 22 x 、 y 方向的边缘

3. 5. 8如何用梯度算子实现边缘检测

1. 原理

基于梯度算子的边缘检测大多数是基于方向导数求卷积的方法。

2. 实现过程

以 3×3 的卷积模板为例，如图3- 23所示

图3- 23 图像卷积示例

设定好卷积模板后，将模板在图像中移动，并将图像中的每个像素点与此模板进行卷积，得到每个像素点的响应 R ，用 R 来表征每个像素点的邻域灰度值变化率，即灰度梯度值，从而可将灰度图像经过与模板卷积后转化为梯度图像。模板系数 $W_i (i=1, 2, 3, \dots, 9)$ 相加的总和必须为零，以确保在灰度级不变的区域中模板的响应为零。 Z 表示像素的灰度值

(3-27)

然后我们设定一个阈值，如果卷积的结果 R 大于这个阈值那么该像素点为边缘点，输出白色；如果 R 小于这个阈值，那么该像素不为边缘点输出黑色。于是最终我们就能输出一幅黑白的梯度图像，实现边缘的检测。

由于边缘检测涉及的概念比较复杂，因此专门用整整一个小节来讲解边缘检测的内容以及如何用梯度算子去计算边缘阈值。后文将要介绍的边缘检测算法重点用于阶跃型边缘的计算，按照其一阶导数具有极大值或极大值点对应二阶导数为过零点，将边缘检测算法分为一阶微分边缘算子检测方法和二阶微分边缘算子检测方法。

3. 4. 6一阶微分边缘算子

一阶微分边缘算子也称为梯度边缘算子，它是利用图像在边缘处的阶跃性，即图像梯度在边缘取得极大值的特性进行边缘检测。

(x)

图3- 24 一节微分边缘

梯度是一个矢量，它具有方向 θ 和模 $|\Delta I|$ ：

(3-28)

梯度的方向提供了边缘的趋势信息，因为梯度方向始终是垂直于边缘方向，梯度的模值大小提供了边缘的强度信息。

在实际使用中，通常利用有限差分进行梯度近似。对于上面的公式，我们有如下的近似：

(3-29)

(1) Roberts算子

1963年，Roberts提出了这种寻找边缘的算子。Roberts边缘算子是一个 2×2 的模板，采用的是对角方向相邻的两个像素之差。从图像处理的实际效果来看，边缘定位较准，对噪声敏感。

由Roberts提出的算子是一种利用局部差分算子寻找边缘的算子，边缘的锐利程度由图像灰度的梯度决定。梯度是一个向量， ∇f 指出灰度变化的最快的方向和数量。

因此最简单的边缘检测算子是用图像的垂直和水平差分来逼近梯度算子：

(3-30)

对每一个像素计算出式（3-30）的向量，求出它的绝对值然后与阈值进行比较，利用这种思想就得到了Roberts交叉算子。

在实际使用中，Roberts算法过程非常简单。

(3-31)

选用1范数梯度计算梯度幅度： $|G(x, y)| = |G_x| + |G_y|$ 。

卷积模板如图3-19所示。

表3-3 Roberts交叉算子模板

	G _x		G _y	
	1	0	0	1
	0	-1	-1	0

G_x G_y

1 0 0 1

0 -1 -1 0

则模板运算结果：（3-32）

如果大于某一阈值，那么我们认为 (x, y) 点为边缘点。

(2) Prewitt算子

Prewitt算子是J. M. S. Prewitt于1970年提出的检测算子。不同于Roberts算子采用 2×2 大小的模板，Prewitt算法采用了 3×3 大小的卷积模板。 2×2 大小的模板在概念上很简单，但是它们对于用关于中心点对称的模板来计算边缘方向不是很有用，其最小模板大小为 3×3 。 3×3 模板考虑了中心点对段数据的性质，并携带有关于边缘方向的更多信息。

在算子的定义上，Prewitt希望通过使用“水平+垂直”的两个有向算子去逼近两个偏导数 G_x, G_y ，这样在灰度值变化很大的区域上卷积结果也同样达到极大值。

图3- 25 Prewitt边缘检测有向算子

在实际使用中，Prewitt边缘检测算子使用两个有向算子（如图3- 25(a)所示水平、如图3- 25(b)所示垂直），每一个逼近

一个偏导数，这是一种类似计算偏微分估计值的方法，x和y两个方向的近似检测算子为：

(3-33)
得出卷积模板为：记图像M，梯度幅值T，比较
(3-34)
如果最终T大于阈值threshold，那么该点为边缘点。

(3) Sobel算子
Sobel最初是1968年在一次博士生课题讨论会上提出的（“A3×3 Isotropic Gradient Operator for Image Processing”），后来正式出版发表是在1973年的一本专著（“Pattern Classification and Scene analysis”）的脚注里作为注释出现和公开的。Sobel算子和Prewitt算子都是加权平均，但是Sobel算子认为，邻域的像素对当前像素产生的影响不是等价的所以距离不同的像素具有不同的权值，对算子结果产生的影响也不同。一般来说距离越远，产生的影响越小。

将Prewitt边缘检测算子模板的中心系数增加一个权值2，不但可以突出中心像素点而且可以得到平滑的效果，这就成为索贝尔算子。

Sobel算子一种将方向差分运算与局部平均相结合的方法。该算子是在以为中心的3×3邻域上计算x和y方向的偏导数，即
(3-35)
得出卷积模板为：记图像M，梯度幅值T，比较
如果最终T大于阈值threshold，那么该点为边缘点

(4) Kirsch算子
Kirsch算子是R. Kirsch提出来的一种边缘检测算法。与前述的算法不同之处在于，Kirsch考虑到3×3的卷积模板事实上涵盖着8种方向（左上，正上，……，右下），于是Kirsch采用8个3×3的模板对图像进行卷积，这8个模板代表8个方向，并取最大值作为图像的边缘输出。

它采用下述8个模板对图像上的每一个像素点进行卷积求导数：
(3-36)
最终选取8次卷积结果的最大值作为图像的边缘输出。

(5) 小结
本小节主要介绍了常见的一阶微分边缘算子，包括Roberts算子、Prewitt算子、Sobel算子和Kirsch算子等，这些一阶微分边缘算子的核心思路是利用图像在边缘处的阶跃性，即图像梯度在边缘取得极大值的特性进行边缘检测。

6. 计算机视觉教材_第6部分		总字符数：4005
相似文献列表		
去除本人文献复制比：40.4%(1620) 去除引用文献复制比：40.4%(1620) 文字复制比：40.4%(1620)		
1	常用图像插值算法分析与比较 - 道客巴巴 - 《互联网文档资源 (https://www.doc88.co) 》 - 2020	12.7% (509) 是否引证：否
2	全景成像探测的目标拼接算法 韩卫芳(导师：高明;刘宇) - 《西安工业大学硕士论文》 - 2013-04-27	12.5% (501) 是否引证：否
3	图像插值算法-百度文库 - 《互联网文档资源 (http://wenku.baidu.c) 》 - 2012	12.4% (495) 是否引证：否
4	《数字图像几何校正》-毕业论文.doc 文档全文免费预览 - 《互联网文档资源 (https://max.book118.) 》 - 2019	11.9% (475) 是否引证：否
5	关于图像重采样插值算法 - 《互联网文档资源 (http://www.360doc.co) 》 - 2016	11.5% (460) 是否引证：否
6	屏幕软键盘自动识别技术研究 韩钢(导师：屈桢深) - 《哈尔滨工业大学硕士论文》 - 2011-06-01	10.5% (419) 是否引证：否
7	基于数学形态学的图像边缘检测方法 曹晓琳(导师：高广宏) - 《哈尔滨工业大学硕士论文》 - 2012-07-01	10.3% (412) 是否引证：否
8	关于图像重采样插值算法下载 - 《互联网文档资源 (https://ishare.iask.) 》 - 2020	9.8% (392) 是否引证：否
9	复杂成像条件下的QR码自动识别 朱新如(导师：袁晓辉) - 《东南大学硕士论文》 - 2019-05-01	9.2% (368) 是否引证：否
10	分布式并行计算在超分辨率图像增强中的应用 郑杰;包盛;杨莹; - 《电信科学》 - 2015-04-20	8.7% (347) 是否引证：否
11	超声相控阵成像算法研究及软件系统设计 许药林(导师：徐大专) - 《南京航空航天大学硕士论文》 - 2012-12-01	7.1% (283) 是否引证：否

12	<u>舱室火灾参数集成分析系统的研究与实现</u>	6.6% (264)
	严宽(导师:尹云飞) - 《重庆大学硕士论文》 - 2018-10-01	是否引证: 否
13	<u>一种改进二维码视觉精定位AGV技术研究</u>	6.4% (255)
	李照;舒志兵; - 《控制工程》 - 2019-06-20	是否引证: 否
14	<u>基于FPGA的高清视频显示控制系统的研究与设计</u>	5.7% (230)
	许恒煜(导师:杨全胜;顾洪) - 《东南大学硕士论文》 - 2021-05-28	是否引证: 否
15	<u>基于深度学习的稀疏角CT重建研究</u>	4.7% (189)
	钱雨飞(导师:李海波) - 《南京邮电大学硕士论文》 - 2020-12-16	是否引证: 否
16	<u>基于高频超声的机械扫描装置及数字扫描变换技术研究</u>	4.6% (185)
	朱国乾(导师:陶晓杰) - 《合肥工业大学硕士论文》 - 2014-04-01	是否引证: 否
17	<u>基于可调参数融合架构的Scaler方法研究与芯片实现</u>	4.2% (169)
	计慧杰(导师:冉峰) - 《上海大学博士论文》 - 2011-12-01	是否引证: 否
18	<u>基于鼻梁检测和PCA算法的眼镜识别方法</u>	4.2% (169)
	孙瑾怡;赵文静;张伟康;戴泽凯; - 《电子制作》 - 2022-04-01	是否引证: 否
19	<u>基于深度学习的机器人绘制肖像漫画技术研究与实现</u>	4.2% (169)
	季新宇(导师:朱留存) - 《扬州大学硕士论文》 - 2021-05-29	是否引证: 否
20	<u>叶片常见形态特征提取的原型系统设计与实现</u>	4.2% (167)
	陈睿(导师:武剑洁) - 《华中科技大学硕士论文》 - 2014-01-01	是否引证: 否
21	<u>移动机器人微小型高速立体视觉技术研究及应用</u>	4.1% (163)
	李琪(导师:杨华) - 《华中科技大学硕士论文》 - 2020-06-01	是否引证: 否
22	<u>基于VC++的数字图像处理软件开发</u>	3.4% (135)
	常奇峰(导师:王开福) - 《南京航空航天大学硕士论文》 - 2010-12-01	是否引证: 否
23	<u>基于改进U-Net网络的内窥镜图像烟雾净化算法</u>	3.0% (122)
	林金朝;蒋媚秋;庞宇;王慧倩; - 《中国生物医学工程学报》 - 2021-06-20	是否引证: 否
24	<u>基于边缘检测与模式识别的车脸识别算法</u>	2.7% (110)
	徐骏骅; - 《控制工程》 - 2018-02-20	是否引证: 否
25	<u>南大洋海冰影像地图投影变换与瓦片切割应用研究</u>	2.7% (108)
	田璐;艾松涛;鄂栋臣;龚洪清;沈权;徐宁;张宏洋; - 《极地研究》 - 2012-09-15	是否引证: 否
26	<u>基于随机森林模型的心脏CT图像分割算法研究</u>	2.7% (108)
	黄凡力(导师:王旭初) - 《重庆大学硕士论文》 - 2015-05-01	是否引证: 否
27	<u>不同性状粗集料对沥青混合料高温性能的影响研究</u>	0.8% (32)
	钱野(导师:肖鹏;袁峻) - 《扬州大学硕士论文》 - 2010-04-01	是否引证: 否

原文内容

3.4.7二阶微分边缘算子

学过微积分我们都知道,边缘即是图像的一阶导数局部最大值的地方,即该点的二阶导数为零。二阶微分边缘检测算子就是利用图像在边缘处的阶跃性导致图像二阶微分在边缘处出现零值这一特性进行边缘检测的。

(x)

图3- 26 二阶微分边缘

对于图像的二阶微分可以用普拉斯算子来表示:

(3-37)

我们在像素点(i, j)的3×3的邻域内,可以有如下的近似:

(3-38)

对应的二阶微分卷积核为:(1) Laplace算子

Laplace(拉普拉斯)算子是最简单的各向同性微分算子,一个二维图像函数的拉普拉斯变换是各向同性的二阶导数,式(3-39)为Laplace算子的表达式:

(3-39)

把这个表达式代入卷积模板的表达式中进行一系列推导,我们就能得到Laplace模板为:还有一种常用的卷积模板为:有时我们为了在邻域中心位置取到更大的权值,还使用如下卷积模板:实际中我们使用Laplace模板的方法如下:

1) 判断图像中心像素灰度值与它周围其他像素的灰度值,如果中心像素的灰度更高,则提升中心像素的灰度;反之降低中

心像素的灰度，从而实现图像锐化操作；

2) 在算法实现过程中，Laplacian算子通过对邻域中心像素的四方向或八方向求梯度，再将梯度相加起来判断中心像素灰度与邻域内其他像素灰度的关系；

3) 最后通过梯度运算的结果对像素灰度进行调整。

拉普拉斯算子是一种高通滤波器，是图像灰度函数在两个垂直方向二阶偏导数之和。在离散数字图像的情况下，直接用图像灰度级的二阶差分代替连续情形下的二阶偏导数，对噪声很敏感，在提取边缘时往往会出现伪边缘响应。

图3- 27 Laplace算子边缘检测

(2) LOG算子

1980年，Marr和Hildreth提出将laplace算子与高斯低通滤波相结合，提出了LOG (Laplace and Guassian) 算子，又称为马尔 (Marr) 算子。

该算子是先运用高斯滤波器平滑图像达到去除噪声的目的，然后用Laplace算子对图像边缘进行检测。这样既达到了降低噪声的效果，同时也使边缘平滑，并得到了延展。为了防止得到不必要的边缘，边缘点应该选取比某阈值高的一阶导数零交叉点。LOG算子已经成为目前对阶跃边缘用二阶导数过零点来检测的最好的算子。

LOG算子的卷积模板通常采用 5×5 的矩阵，如：

实际中我们使用LOG模板的方法如下：

(1) 遍历图像（除去边缘、防止越界），对每个像素做Gauss-Laplacian模板卷积运算。

(2) 复制到目标图像，结束。

(3) Canny算子

Canny边缘检测算法是1986年由John F. Canny开发出来一种基于图像梯度计算的边缘检测算法，同时Canny本人对计算图像边缘提取学科的发展也是做出了很多的贡献。尽管至今已经许多年过去，但是该算法仍然是图像边缘检测方法经典算法之一。

Canny根据以前的边缘检测算子以及应用归纳了如下三条准则：

(1) 信噪比准则。避免真实的边缘丢失，避免把非边缘点错判为边缘点。

(2) 定位精度准则。得到的边缘要尽量与真实边缘接近。

(3) 单一边缘响应准则。单一边缘需要具有独一无二的响应要避免出现多个响应，并最大抑制虚假响应。

以上三条准则是Canny首次明确提出并对这个问题进行完全解决的。更为重要的是，Canny同时给出了它们的数学表达式（现以一维为例），这就转化成为一个泛函优化的问题。

经典的Canny边缘检测算法通常都是从高斯模糊开始到基于双阈值实现边缘连接结束。但是在实际工程应用中，考虑到输入图像都是彩色图像，最终边缘连接之后的图像要二值化输出显示，所以完整的Canny边缘检测算法实现步骤如下：

(1) 彩色图像转换为灰度图像；

(2) 使用 5×5 高斯滤波器对图像进行高斯模糊；(3) 对平滑后的图像使用sobel算子在水平与竖直方向上计算一阶导数，得到图像梯度 (G_x 和 G_y)，根据梯度计算图像边缘幅值与角度；

(4) 根据角度对幅值进行非极大值抑制（边缘细化）：将其梯度方向近似为以下值中的一个（0, 45, 90, 135, 180, 225, 270, 315）（即上下左右和45度方向）；比较该像素点，和其梯度方向正负方向的像素点的梯度强度，如果该像素点梯度强度最大则保留，否则抑制（删除，即置为0）。

3.5 图像插值方法

常用的插值方法有三种，即最近邻插值法、双线性插值法和双三次插值法。

考虑一个针对某些参数值（如颜色或位置之类的属性或性质）进行采样的函数。所谓插值就是对这样的函数在其没有采样或测量的参数值处进行函数值估计的过程。一幅图像可以看成是一个二维函数 $I(x, y)$ 的采样，它给出每个空间位置 (x, y) 处的颜色值。通常情况下，图像采样的位置参数 x 和 y 取整数。给定 $I(x, y)$ 在整数网格点 (x, y) 处的函数值，使用插值的方法计算 $I(x, y)$ 在网格点之间并且有可能是非整数值的点 (x, y) 处的函数值。

3.5.1 最近邻插值法 (Ref. 基于MATLAB的数字图像处理)

最近邻插值法又称零阶插值法，是最简单的插值法，如图3- 28所示。首先计算 (x_0, y_0) 与其邻近的4个整数坐标点的距离，然后将最近的整数坐标点的灰度值作为其灰度值。当 (x_0, y_0) 附近各相邻像素之间的灰度变化较小时，最近邻插值法就是一种简单快速的插值法。但当 (x_0, y_0) 附近各相邻像素之间的灰度差异很大时，最近邻插值法就会产生较大的误差，甚至可能影响图像质量，出现锯齿效应。

图3- 28 最近邻插值法

3.5.2 双线性插值法 (Ref. 基于MATLAB的数字图像处理)

双线性插值法又称一阶插值法，是对最近邻插值法的一种改进。该方法采用线性内插，根据点 (x_0, y_0) 的4个相邻点的灰度值，先沿着其中的一个方向插值，然后再沿着另外一个方向插值，计算出灰度值 $f(x_0, y_0)$ ，如图3- 29所示。

双线性插值运算包含3次一维线性插值运算，具体计算过程如下：

(1) 计算 α 和 β ，其实就是 x_0 和 y_0 的小数部分，即

(3-40)

(2) 根据 $f(x, y)$ 、 $f(x+1, y)$ 线性插值求 $f(x_0, y)$ ，即

(3-41)

(3) 根据 $f(x, y+1)$ 、 $f(x+1, y+1)$ 线性插值求 $f(x_0, y+1)$ ，即

(3-42)

(4) 根据 $f(x_0, y)$ 、 $f(x_0, y+1)$ 线性插值求 $f(x_0, y_0)$ ，即

(3-43)

上述过程是先沿 x 方向，再沿 y 方向进行插值处理。其实，也可以先沿 y 方向，后沿 x 方向进行插值。两种方式的插值结果相同。

图3- 29 双线性插值法

3.5.3 双三次插值法 (Ref. 基于MATLAB的数字图像处理) 参考文献标注格式
双三次插值是一种更加复杂的插值法, 它能克服以上两种插值法的不足, 而且图像的边缘轮廓比双线性插值法更清晰。双三次插值法的计算量大, 既要考虑(x0, y0)点的直接相邻点对它的影响, 也要考虑周围16个相邻点的灰度值对它的影响, 见图2-11。

根据双三次插值, (x0, y0)点的插值公式为:
(3-44)
其中, 矩阵A、B、C表示为: s(x)为插值权重核, 表示为:
(3-45)

3.5.4、三种插值法优缺点分析整理成一个表格更好

- (1) 最近邻插值法
优点: 计算量很小, 算法也简单, 因此运算速度较快。
缺点: 仅使用离待测采样点最近的像素的灰度值作为该采样点的灰度值, 而没考虑其他相邻像素点的影响, 因而重新采样后灰度值有明显的 discontinuity, 图像质量损失较大, 精度不够, 严重失真, 会产生明显的马赛克和锯齿现象。
- (2) 双线性插值法
优点: 效果要好于最近邻插值, 但缩放后图像质量高, 基本克服了最近邻插值灰度值不连续的特点, 因为它考虑了待测采样点周围四个直接邻点对该采样点的相关性影响。
缺点: 相较于最近邻插值法计算量稍大一些, 算法复杂些, 程序运行时间也稍长些, 此方法仅考虑待测样点周围四个直接邻点灰度值的影响, 而未考虑到各邻点间灰度值变化率的影响, 因此具有低通滤波器的性质, 从而导致缩放后图像的高频分量受到损失, 图像边缘在一定程度上变得较为模糊。用此方法缩放后的输出图像与输入图像相比, 仍然存在由于插值函数设计考虑不周而产生的图像质量受损与计算精度不高的问题。
- (3) 双三次插值法
优点: 立方卷积插值不仅考虑到周围四个直接相邻像素点灰度值的影响, 还考虑到它们灰度值变化率的影响。因此克服了前两种方法的不足之处, 能够产生比双线性插值更为平滑的边缘, 计算精度很高, 处理后的图像像质损失最少, 效果是最佳的。

缺点: 计算量最大, 算法复杂, 运行速度较慢。

习题

- 3-1 维信号处理中, 模拟信号数字化的过程包括?
3-2 在计算机中, 图像的基本单位是什么?其定义是怎样的?
3-3 图像数字化过程中, 采样是指什么?采样间隔与图像分辨率的关系是怎样的?
3-4 采样定理又称奈奎斯特定理, 该定理的具体内容是?
3-5 量化的定义是什么?采用不同的量化级数对图像质量有怎样的影响?
3-6 常用的数字图像按照颜色和灰度的多少可以分为哪几种基本类型?
3-7 灰度图像一般是指具有256级灰度值的数字图像, 即8bit灰度图像。灰度图像矩阵元素的取值范围为?
3-8 常见的图像噪声有哪些?
3-9 最常见和最基本的图像滤波方法有哪些?
3-10 图像进行高斯滤波后的平滑程度与高斯滤波的标准差有怎样关系?
3-11 常用的一阶微分边缘算子有哪些?常用的二阶微分边缘算子有哪些?
3-12 常用的图像插值的方法有哪些?

7. 计算机视觉教材_第7部分			总字符数: 9883
相似文献列表			
去除本人文献复制比: 34.2%(3379)		去除引用文献复制比: 34.2%(3379)	文字复制比: 34.2%(3379)
1	基于机载LiDAR和高分辨率遥感影像的道路提取研究 高利鹏(导师: 史文中) - 《中国矿业大学硕士论文》 - 2014-06-01	9.6% (953)	是否引证: 否
2	纹理特征提取[转载]_小虫不会飞_ - 《网络 (http://blog.sina.com) 》 - 2015	9.6% (952)	是否引证: 否
3	中医舌诊辅助系统 刘龙飞(导师: 李文书) - 《浙江理工大学硕士论文》 - 2014-06-08	8.7% (864)	是否引证: 否
4	基于结构支持向量机的目标跟踪算法研究 戴诗语(导师: 王爱丽) - 《哈尔滨理工大学硕士论文》 - 2016-03-01	5.9% (588)	是否引证: 否
5	基于SIFT特征的图像匹配 - 豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2016	4.8% (477)	是否引证: 否
6	基于FPGA的图像分割提取系统研究 孙伟(导师: 黄大庆) - 《南京航空航天大学硕士论文》 - 2010-03-01	4.6% (455)	是否引证: 否
7	基于数字图像处理的圆形物体检测研究	4.6% (453)	

	张国栋(导师:郭三明) - 《河南理工大学硕士论文》 - 2009-12-15	是否引证: 否
8	(图像匹配问题研究)毕业设计论文 - 《网络 (http://www.wendangku) 》 - 2020	4.5% (446)
9	基于sift特征的图像匹配毕业论文 - 《互联网文档资源 (https://wenku.baidu .) 》 - 2018	是否引证: 否
10	基于SIFT特征的图像匹配.doc 文档全文免费预览 - 《互联网文档资源 (https://max.book118 .) 》 - 2019	4.5% (446)
11	基于视觉的服务机器人目标识别和定位研究_图文 - 《互联网文档资源 (http://wenku.baidu.c) 》 - 2016	是否引证: 否
12	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	4.3% (423)
13	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	是否引证: 否
14	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	4.3% (423)
15	基于SIFT的图像匹配毕业设计论文下载 - 《互联网文档资源 (https://ishare.iask .) 》 - 2021	是否引证: 否
16	毕业设计 (论文) -基于SIFT特征的图像匹配下载 - 《互联网文档资源 (https://ishare.iask .) 》 - 2021	4.2% (420)
17	图像分割技术分析展望 姜彬;施志刚; - 《电脑知识与技术》 - 2009-12-15	是否引证: 否
18	基于特征点的图像匹配技术研究与应-百度文库 - 《互联网文档资源 (http://wenku.baidu.c) 》 - 2012	4.2% (414)
19	沥青路面微观形貌采集系统研制 胡克波(导师:严新平) - 《武汉理工大学硕士论文》 - 2008-11-01	是否引证: 否
20	GPS/SINS/SAR组合导航系统信息融合及误差修正技术研究 黄金山(导师:刘贵喜) - 《西安电子科技大学硕士论文》 - 2010-01-01	4.0% (392)
21	(图像匹配问题研究)毕业设计论文-百度文库 - 《互联网文档资源 (http://wenku.baidu.c) 》 - 2012	是否引证: 否
22	CCTV目标定位与跟踪技术的研究 吴键川(导师:杨梅) - 《大连海事大学硕士论文》 - 2011-05-01	3.9% (390)
23	基于二维图像的准直测量技术研究 孟浩(导师:冯驰) - 《哈尔滨工程大学硕士论文》 - 2009-02-01	是否引证: 否
24	[DOCIN]基于双目视觉的图像三维重建 - 豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2017	3.9% (384)
25	双点无纺粘合衬的质量检测方法探讨 陆国琴(导师:陈庆官) - 《苏州大学硕士论文》 - 2009-04-01	是否引证: 否
26	基于角点检测的倾斜文档校正 刘铃;陈立峰;彭向前; - 《机电信息》 - 2019-09-24 1	2.2% (215)
27	基于改进FAST特征检测的ORB-SLAM方法 公维思;周绍磊;吴修振;刘刚; - 《现代电子技术》 - 2018-03-12 1	是否引证: 否
28	基于FAST和Brief算法使用tracking.js的Web人脸识别考勤系统 徐辉;陈晨; - 《现代计算机(专业版)》 - 2017-07-15	1.6% (163)
29	基于全局和角点特征的图像检索 姜雪;邵宝民;王振;李秋玲; - 《山东理工大学学报(自然科学版)》 - 2019-08-28 1	是否引证: 否
30	基于双目视觉零件特征测量技术的研究 胡炜;李长春; - 《计量与测试技术》 - 2023-02-28	1.3% (131)
		是否引证: 否

31	<u>基于特征点的影像匹配方法探讨</u> 王雅丽;厉彦玲;蔡玉林; - 《测绘科学》 - 2012-10-23 0	1.3% (126)	是否引证: 否
32	<u>眼科手术机器人双目视觉标定方法研究</u> 闫兴;曹禹;王晓楠;朱立夫;王君;何文浩; - 《工具技术》 - 2019-12-20	1.0% (98)	是否引证: 否
33	<u>平行双目视觉测量原理及实验研究</u> 孙雷; - 《电子技术与软件工程》 - 2022-07-15	1.0% (97)	是否引证: 否
34	<u>基于图像识别的辅助装配系统研究</u> 王必贤;陆培民;钱慧; - 《长江信息通信》 - 2021-05-15	1.0% (94)	是否引证: 否
35	<u>遮挡目标检测与识别技术研究</u> 蔡星艳;赵和鹏;邱鹏;朱长仁; - 《数字技术与应用》 - 2013-09-15	0.9% (85)	是否引证: 否
36	<u>基于图像处理的衣服尺寸测量方法研究</u> 徐江剑;潘玲佼;王关成; - 《工业控制计算机》 - 2021-03-25	0.8% (75)	是否引证: 否
37	<u>Internet上的图像信息检索技术</u> 熊回香 - 《情报学报》 - 2005-04-24	0.7% (73)	是否引证: 否
38	<u>基于本体的材料腐蚀图像标注方法</u> 纪钢;张菲;王平; - 《重庆理工大学学报(自然科学)》 - 2011-04-15	0.7% (71)	是否引证: 否
39	<u>一种基于局部特征索引结构的目标跟踪方法</u> 任世杰;杨小冈;齐乃新;马玛双; - 《电光与控制》 - 2018-12-24 1	0.5% (53)	是否引证: 否
40	<u>基于深度学习的行人重识别综述</u> 黄业文;许智聪;单纯; - 《广州大学学报(自然科学版)》 - 2022-06-15	0.5% (45)	是否引证: 否
41	<u>基于OCR的变电站操作票识别算法研究</u> 仲伟;邵建新; - 《电力设备管理》 - 2021-08-25	0.4% (42)	是否引证: 否
42	<u>局部特征点的鲁棒性数字稳像</u> 郭晓冉;崔少辉; - 《光电工程》 - 2013-05-15	0.3% (32)	是否引证: 否
43	<u>基于多特征多核哈希学习的大规模图像检索</u> 曾宪华;袁知洪;王国胤;杨洁; - 《中国科学:信息科学》 - 2017-08-20	0.3% (32)	是否引证: 否
44	<u>基于图像特征和深度森林的乙丙橡胶电缆绝缘老化状态识别</u> 王科;项恩新;曹伟东;徐肖伟;黄继盛;车雨轩; - 《电测与仪表》 - 2020-12-29 14:54	0.3% (30)	是否引证: 否

原文内容

第4章图像特征检测与图像匹配

4.1概述

图像领域的许多应用中，人们总是希望从分割出的区域中分辨出地物类别，例如分辨农田、森林、湖泊、沙滩等；或是希望从分割出的区域中识别出某种物体（目标），例如在河流中识别舰船；在飞机跑道上识别飞机等。进行地物分类和物体识别的第一步就是物体特征的提取和检测，然后才能根据检测和提取的图像特征对图像中可能的物体进行识别。

图像特征是用于区分一个图像内部特征的最基本的属性。图像特征按照形成方式可分成自然特征和人工特征两类。人工特征是指人们为了便于对图像进行处理和分析而人为认定的图像特征，比如图像直方图、图像频谱和图像的各种统计特征（图像的均值、图像的方差、图像的标准差、图像的熵）等。自然特征是指图像固有的特征，比如图像中的边缘、角点、纹理、形状和颜色等。

全局特征是指图像的整体属性，常见的全局特征包括颜色特征、纹理特征和形状特征，比如强度直方图等。由于是像素级的低层可视特征，因此，全局特征具有良好的不变性、计算简单、表示直观等特点，但特征维数高、计算量大是其致命弱点。此外，全局特征描述不适用于图像混叠和有遮挡的情况。局部特征则是从图像局部区域中抽取的特征，包括边缘、角点、线、曲线和特别属性的区域等。常见的局部特征包括角点类和区域类两大类描述方式。

图像特征提取是图像分析与图像识别的前提，它是将高维的图像数据进行简化表达最有效的方式，从一幅图像的M×N×3的数据矩阵中，我们看不出任何信息，所以我们必须根据这些数据提取出图像中的关键信息，一些基本元件以及它们的关系。

局部特征点是图像特征的局部表达，它只能反映图像上具有的局部特殊性，所以它只适合于对图像进行匹配、检索等应用。对于图像理解则不太适合。而后者更关心一些全局特征，如颜色分布、纹理特征、主要物体的形状等。全局特征容易受环境的干扰，光照、旋转、噪声等不利因素都会影响全局特征。相比而言，局部特征点，往往对应着图像中的一些线条交叉，在明暗变化的结构中，受到的干扰也少。

斑点与角点是两类常见的局部特征点。斑点通常是指与周围有着颜色和灰度差别的区域，如草原上的一棵树或一栋房子。它是一个区域，所以它比角点的抗噪能力强，稳定性好。而角点则是图像中一边物体的拐角或者线条之间的交叉部分。

(a)Corners角点 (b)Blobs斑点

图4- 1 两种常见的特征点出处？

除斑点与角点外，本章还会介绍一个图像局部特征——特征描述子。特征描述子（Feature Descriptors）指的是检测图像的局部特征（如边缘、角点、轮廓等），然后根据匹配目标的需要进行特征的组合、变换，以形成易于匹配、稳定性好的特征向量。

特征匹配是指通过分别提取两个或多个图像的特征（点、线、面等特征），对特征进行参数描述，然后运用所描述的参数来进行匹配的一种算法。特征就是有意义的图像区域，该区域具有独特性或易于识别性。角点与高密度区域是一个很好的特征，边缘可以将图像分为两个区域，因此可以看作很好的特征，斑点（与周围有很大区别的图像区域）也是有意义的特征。基于特征的匹配所处理的图像一般包含的特征有颜色特征、纹理特征、形状特征、空间位置特征等。基于图像特征的匹配方法可以克服利用图像灰度信息进行匹配的缺点，由于图像的特征点较像素点要少很多，大大减少了匹配过程的计算量；同时，特征点的匹配度量值对位置的变化比较敏感，可以大大提高匹配的精确程度；而且，特征点的提取过程可以减少噪声的影响，对灰度变化，图像形变以及遮挡等都有较好的适应能力。所以基于图像特征的匹配在实际中的应用越来越广泛。所使用的特征基元有点特征（明显点，角点，边缘点等），边缘线段等。特征匹配的基本流程是：首先对图像进行预处理来提取其高层次的特征，然后建立两幅图像之间特征的匹配对应关系，最后根据相似性原则对两幅图像中的特征点进行匹配。

图4- 2 特征点匹配出处？

图像配准的方法大致分为三类，一类是基于灰度和模板的，这类方法直接采用相关运算等方式计算相关值来寻求最佳匹配位置，方法简单较为死板，一般效果不会太好。第二类是基于特征的匹配方法，如sift、surf点特征，或者向量特征等等，适应性较强。第三类是基于域变换的方法，采用相位相关（傅里叶-梅林变换）或者沃尔什变换、小波等方法，在新的域下进行配准。

4. 2全局图像特征

图像特征有很多种，包括颜色特征、形状特征以及纹理特征等。至于什么是好的特征则要视具体任务而定，能够较好地解决问题的特征就是好的特征。

对于图像特征，一般是使用一个特征向量来表示。特征向量的维度根据特征的不同可以从一维到成千上万维。例如，使用图像所有像素的灰度值的均值作为图像特征，就是一个维的特征向量。将图像所有像素的灰度值连接为一个特征向量表示图像。则特征向量的维度为图像中所包含的像素的个数。

对于图像分类来说，准确的灰度或颜色值并不重要，准确的特征位置也不重要，而图像中的边缘和纹理相对重要。这是由于具体的灰度或颜色值以及特征位置容易受到光照、视角等因素的影响而发生变化，而边缘和纹理等信息与灰度和颜色相比，受光照以及视角等因素的影响较小而造成的。

图像特征是图像分析和理解的基础，获取有效的底层图像特征是图像高级语义信息提取的关键。本节主要介绍常见的底层图像特征，如颜色、纹理和形状等特征的提取方法。

4. 2. 1颜色特征

颜色特征是一种全局的特征，综合了图像区域内所有像素的颜色信息，用于描述物体表面性质。颜色直方图和颜色矩是常用的图像颜色特征描述方法。

(1) 颜色直方图

颜色直方图是常用的图像特征之一，在图像检索、图像识别等领域有着广泛的应用。颜色直方图描述的是不同颜色在整幅图像中所占的比例，而并不关心每种颜色所处的空间位置。颜色直方图可以基于不同的颜色空间来获得。灰度图像的颜色直方图也称为灰度直方图，如图4- 3所示。

(a) 灰度图像 (b) 灰度图像的灰度直方图

图4- 3 颜色直方图

计算颜色直方图需要将颜色空间划分成若干个小的颜色区间，这个过程称为颜色量化。对于灰度直方图也需要进行量化，不过一般都是将一个灰度值作为一个区间，灰度共有256级，所以对应的灰度直方图包含256个区间，然后，计算颜色落在每个小区间内的像素的数量可以得到颜色直方图。颜色量化有许多方法，常用的做法是将颜色空间的各个分量（维度）均匀地进行划分。例如，在RGB颜色空间中，将R、G和B分量分别量化为256个区间，则可以得到一个 $256 \times 256 \times 256$ 维的彩色直方图。这样的划分使得直方图的维度过高，会导致直方图中很多的区间是没有值的。例如，一幅分辨率为 1000×1000 的图像只包含100万个像素，而将每个颜色分量划分为256份得到的直方图包含超过1600万个区间，因此可以把每个颜色分量划分为较少的份数。例如，每个分量分为4份，则可以得到一个64维的直方图，也可以对三个颜色分量分别计算一个直方图，然后连接起来作为最后的彩色直方图。例如，将R、G和B分量分别量化为256个区间，将每个颜色分量的直方图进一步连接得到的直方图的维度为 256×3 。

颜色直方图是图像的统计信息，其优点是计算简单，并且对于旋转、平移等操作具有不变性。其缺点是没有考虑各种颜色的空间位置。例如，两幅完全不同的图像可以具有相同

直方图，导致颜色直方图的判别力较差。针对这个问题，可以将图像分为多个小块，计算每个小块的直方图然后连接在一起，就等于同时考虑了颜色的统计信息以及空间分布。

(2) 颜色矩

另一种简单而有效的颜色特征是由Stricker 和Oren提出的颜色矩（Color Moments）[55]。颜色矩的数学基础在于图像中任何的颜色分布都可以使用其矩来表示。由于颜色分布信息主要集中在低阶矩中，因此一般只使用颜色的一阶矩（均值 Mean）、二阶矩（方差 Variance）和三阶矩（斜度 Skewness）来表达图像的颜色分布。与颜色直方图相比，颜色矩的另一个好处在于不需要对颜色空间进行向量比，所得到的特征向量的维度较低。

颜色的一阶矩、二阶矩和三阶矩的计算公式为

(4-1)

式中，为图像中的第j个像素的第i个分量；N为图像中像素的个数。对于一幅RGB彩色图像，其每个分量可以计算得到三个颜色矩阵，可以将各个分量对应的颜色矩特征连接为一个九维的特征向量来表示该图的颜色特征。

4. 2. 2 纹理特征

一幅图像的纹理是在图像计算中经过量化的图像特征。图像纹理描述图像或其中小块区域的空间颜色分布和光强分布。纹

理特征的提取分为基于结构的方法和基于统计数据的方法。一个基于结构的纹理特征提取方法是将所要检测的纹理进行建模，在图像中搜索重复的模式。

(1) LBP特征

LBP方法 (Local binary patterns) 是一个计算机视觉中用于图像特征分类的一个方法。LBP方法在1994年首先由T. Ojala, M. Pietikäinen, 和 D. Harwood 提出[43][44], 用于纹理特征提取。后来LBP方法与HOG特征分类器联合使用, 改善了一些数据集[45]上的检测效果。

对LBP特征向量进行提取的步骤如下:

首先将检测窗口划分为 16×16 的小区域 (cell), 对于每个cell中的一个像素, 将其环形邻域内的8个点 (也可以是环形邻域多个点, 如图 34. 应用LBP算法的三个邻域示例所示) 进行顺时针或逆时针的比较, 如果中心像素值比该邻点大, 则将邻点赋值为1, 否则赋值为0, 这样每个点都会获得一个8位二进制数 (通常转换为十进制数)。然后计算每个cell的直方图, 即每个数字 (假定是十进制数) 出现的频率 (也就是一个关于每一个像素点是否比邻域内点大的一个二进制序列进行统计), 然后对该直方图进行归一化处理。最后将得到的每个cell的统计直方图进行连接, 就得到了整幅图的LBP纹理特征, 然后便可利用SVM或者其他机器学习算法进行分类了。

(2) 灰度共生矩阵

灰度共生矩阵是另一种纹理特征提取方法, 首先对于一幅图像定义一个方向 (orientation) 和一个以pixel为单位的步长 (step), 灰度共生矩阵 $T(N \times N)$, 则定义 $M(i, j)$ 为灰度级为 i 和 j 的像素同时出现在一个点和沿所定义的方向跨度步长的点上的频率。其中 N 是灰度级划分数目。由于共生矩阵有方向和步长的组合定义, 而决定频率的一个因素是对矩阵有贡献的像素数目, 而这个数目要比总共数目少, 且随着步长的增加而减少。因此所得到的共生矩阵是一个稀疏矩阵, 所以灰度级划分 N 常常减少到8级。如在水平方向上计算左右方向上像素的共生矩阵, 则为对称共生矩阵。类似的, 如果仅考虑当前像素单方向 (左或右) 上的像素, 则称为非对称共生矩阵。

4.2.3 形状特征

形状特征用来描述图像中所包含物体的形状。形状特征更接近于目标的语义特征, 包含一定的语义信息, 忽略了图像中不相关的背景或不重要的目标。通常来讲, 形状特征有以下两种表示方法:

(1) 轮廓特征, 即目标的外边界。通过检测边缘, 提取物体的轮廓, 然后计算轮廓所具有的特征。常用的轮廓特征包括链码、多边形近似、傅里叶描述子、偏心率以及边界长度等。

(2) 区域特征, 即针对整个物体区域提取特征, 是对物体区域中的所有像素集合的描述。常用的区域特征包括区域面积、几何不变矩、正交矩以及角半径变换等。

形状特征的表达是以对图像中的目标或区域的分割为基础的, 而图像分割本身就是一个非常困难的问题。此外, 用于表示图像中物体的形状特征必须满足对变换、旋转和缩放的不变性, 这也给形状相似性的计算带来了一定难度。

4.3 角点检测

4.3.1 角点介绍

在现实世界中, 角点对应于物体的拐角, 如道路的十字路口、丁字路口等。下面有两幅不同视角的图像, 通过找出对应的角点进行匹配。

图4- 4

如图4-3-1所示, 放大图像的两处角点区域。我们可以直观地概括一下角点具有的特征:

(1) 轮廓之间的交点;

(2) 对于同一场景, 即使视角发生变化, 角点具有的特征也是稳定不变的;

(3) 该点附近区域的像素点无论在梯度方向上还是在梯度幅值上都有较大变化。

从图像分析的角度来定义角点可以有以下两种定义:

(1) 角点可以是两个边缘的角点;

(2) 角点是邻域内具有两个主方向的特征点。

前者往往需要对图像边缘进行编码, 这在很大程度上依赖于图像的分割与边缘提取, 具有相当大的难度和计算量, 且一旦待检测目标局部发生变化, 很可能导致操作的失败。早期主要有Rosenfeld和Freeman等人的方法, 后期有CSS等方法。

基于图像灰度的方法通过计算点的曲率及梯度来检测角点, 避免了第一类方法存在的缺陷, 此类方法主要由Moravec算子、Forstner算子、Harris算子、SUSAN算子等。

总体来说, 对于角点检测算法而言, 基本思想是使用一个固定窗口在图像上进行任意方向上的滑动, 比较滑动前与滑动后两种情况下窗口中的像素灰度变化程度。如果任意方向上的滑动都有较大灰度变化, 那么我们可以认为该窗口中存在角点。

本部分主要介绍的角点方法包括Harris角点、FAST角点、FAST-ER角点、SUSAN角点等。

4.3.2 Harris角点

人眼对角点的识别通常是在一个局部的小区域或小窗口完成的。如果这个特定的窗口在图像各个方向在移动时, 窗口内图像的灰度没有发生变化, 那么窗口内就不存在角点; 如果窗口在某一个方向移动时, 窗口内图像的灰度发生了较大的变化, 而在另一些方向上没有发生变化, 那么窗口内的图像可能就是一条直线的线段; 如果在各个方向上移动这个特征的小窗口, 窗口内区域的灰度发生了较大的变化, 那么就认为在窗口内遇到了角点, 如图4-2所示。

图4- 5 角点的识别英文!

角点检测最原始的想法就是取某个像素的一个邻域窗口, 当这个窗口在各个方向上进行小范围移动时, 观察窗口内平均的像素灰度值的变化 (即, $E(u, v)$, Window-averaged change of intensity)。从上图可知, 我们可以将一幅图像大致分为三个区域 (‘flat’, ‘edge’, ‘corner’), 这三个区域变化是不一样的。

Harris中使用该像素点周围像素块和其周围的其它像素块的相关性刻画角点, 相关性用平方差之和进行计算 (SSD), SSD越大, 相关性差, 中心像素点越有可能成为角点。假设图像 $I(x, y)$, 将窗口 $w(x, y)$ 在 x 方向上位移 u , 在 y 方向上位移 v , 计算窗口内部的像素灰度值变化:

(4-2)

其中，。

定义角点响应函数R (corner response function)，用于角点判别。其中， λ 是M的特征值。对于每个窗口，都计算其对应的一个角点响应函数 R；然后对该函数进行阈值处理，如果 $R > \text{threshold}$ ，表示该窗口对应一个角点特征。

综上，实际使用Harris角点检测算法共需要5步：

- (1) 计算图像I(x, y)在x和y两个方向的梯度 I_x 、 I_y 。
- (2) 计算图像两个方向梯度的乘积。
- (3) 使用高斯函数对进行高斯加权（取 $\sigma=1$ ），生成矩阵M。
- (4) 计算每个像素的Harris响应值R，并将小于某一阈值t的R置为零。
- (5) 在 3×3 或 5×5 的邻域内进行非最大值抑制，局部最大值点即为图像中的角点。

图4- 6 Harris角点检测示例

Harris角点检测示例的主程序如下。

```
close all;clear
I=imread('GZA Bridge_RGB.jpg');
img=I(51:358,251:558,1:3);
I = rgb2gray(img);
[posX,posY]=harris(I);
figure;imshow(I);
hold on; plot(posX, posY, 'g*');
子程序如下：
function [posX,posY]=harris(I)
%Harris角点检测
%I:输入图像
%posX:角点X坐标
%posY:角点Y坐标
I=double(I);
[m,n]=size(I);
hx=[-1,0,1;-1,0,1;-1,0,1];
Ix=imfilter(I,hx,'replicate','same');%X方向差分图像
Iy=imfilter(I,hx,'replicate','same');%Y方向差分图像
Ix2=Ix.^2;
Iy2=Iy.^2;
Ixy=Ix.*Iy;
h=fspecial('gaussian',3,2);
Ix2=imfilter(Ix2,h,'replicate','same');%高斯滤波
Iy2=imfilter(Iy2,h,'replicate','same');
Ixy=imfilter(Ixy,h,'replicate','same');
R=zeros(m,n);
k=0.06; %建议值(0.04--0.06)
for i=1:m
    for j=1:n
        R(i,j)=(Ix2(i,j)*Iy2(i,j)-Ixy(i,j)*Ixy(i,j))-k*((Ix2(i,j)+Iy2(i,j))^2);%角点响应值
    end
end
T=0.1*max(R(:));% 阈值，可控制返回的角点个数
result=zeros(m,n);
%非极大值抑制（3*3窗口中大于阈值T的局部极大值点被认为是角点）
for i=2:m-1
    for j=2:n-1
        tmp=R(i-1:i+1,j-1:j+1);
        tmp(2,2)=0;
        if(R(i,j)>T&&R(i,j)>max(tmp(:)))
            result(i,j)=1;
        end
    end
end
[posY,posX]=find(result);
end
```

4.3.3 Fast角点

Edward Rosten和Tom Drummond于2006年在《Machine learning for high-speed corner detection》中提出了一种FAST特征点，并在2010年稍作修改后发表了《Features From Accelerated Segment Test》，简称FAST。Rosten等人将FAST角点定义为：若某像素点与其周围领域内足够多的像素点处于不同的区域，则该像素点可能为角点。也就是某些属性与众不同，考虑灰

度图像，即若该点的灰度值比其周围领域内足够多的像素点的灰度值大或者小，则该点可能为角点。注意：FAST只是一种特征点检测算法，并不涉及特征点的特征描述。

图4- 7 FAST特征点示意图

FAST角点的算法步骤如下：

1. 如图4- 7所示，一个以像素p为中心，半径为3的圆上，有16个像素点（p1、p2、...、p16）。
2. 定义一个阈值。计算p1、p9与中心p的像素差，若它们绝对值都小于阈值，则p点不可能是特征点，直接pass掉；否则，当做候选点，有待进一步考察；
3. 若p是候选点，则计算p1、p9、p5、p13与中心p的像素差，若它们的绝对值有至少3个超过阈值，则当做候选点，再进行下一步考察；否则，直接pass掉；
4. 若p是候选点，则计算p1到p16这16个点与中心p的像素差，若它们有至少9个超过阈值，则是特征点；否则，直接pass掉。
5. 对图像进行非极大值抑制：计算特征点出的FAST得分值（即score值，也即s值），判断以特征点p为中心的一个邻域（如3x3或5x5）内，计算若有多个特征点，则判断每个特征点的s值（16个点与中心差值的绝对值总和），若p是邻域所有特征点中响应值最大的，则保留；否则，抑制。若邻域内只有一个特征点（角点），则保留。得分计算公式如下（公式中用V表示得分，t表示阈值）：

(4-3)

图4- 8 FAST角点检测示例

FAST角点检测示例的主程序如下。

```
close all;clear
I=imread('GZA Bridge_RGB.jpg');
I=I(51:358,251:558,1:3);
img = rgb2gray(I);
[m n]=size(img);
score=zeros(m,n);
t=50; %阈值
for i=4:m-3
for j=4:n-3
p=img(i,j);
%步骤1，得到以p为中心的16个邻域点
pn=[img(i-3,j) img(i-3,j+1) img(i-2,j+2) img(i-1,j+3) img(i,j+3) img(i+1,j+3) img(i+2,j+2) img(i+3,j+1) ...
img(i+3,j) img(i+3,j-1) img(i+2,j-2) img(i+1,j-3) img(i,j-3) img(i-1,j-3) img(i-2,j-2) img(i-3,j-1)];
%步骤2
if abs(pn(1)-p)<t && abs(pn(9)-p)<t
continue;
end
%步骤3
p1_5_9_13=[abs(pn(1)-p)>t abs(pn(5)-p)>t abs(pn(9)-p)>t abs(pn(13)-p)>t];
if sum(p1_5_9_13)>=3
ind=find(abs(pn-p)>t);
%步骤4
if length(ind)>=9
score(i,j) = sum(abs(pn-p));
end
end
end
end
%步骤5，非极大抑制，并且画出特征点
for i=4:m-3
for j=4:n-3
if score(i,j)~=0
if max(max(score(i-2:i+2,j-2:j+2)))==score(i,j)
[img(i-3,j), img(i-3,j+1), img(i-2,j+2), img(i-1,j+3), img(i,j+3), img(i+1,j+3), img(i+2,j+2),
img(i+3,j+1), ...
img(i+3,j), img(i+3,j-1), img(i+2,j-2), img(i+1,j-3), img(i,j-3), img(i-1,j-3), img(i-2,j-2), img(i-3,j-1)]= ...
deal(255,255,255,255,255,255,255,255,255,255,255,255,255,255,255,255);
end
end
end
end
figure;
```


imshow(img);

4.3.4 FAST-ER角点

FAST-ER是FAST算法原作者在2010年提出的，它在原来算法里提高特征点检测的重复度，重复意味着第一张图片内的检测的点，也可以在第二张图片上的相应位置被检测出来，重复度可以由如下式子定义：

(4-4)

其中，指第一张图片内的检测点有多少能在第二张被检测到，而定义为有用的特征点数。这里计算的一组图像序列的总的重复度，所以和是图像序列中所以图像对的和。

由于一些形变较大因素造成的形变，很难通过简单且固定的模板将所以的角点检测出来，而原来的FAST算法其决策树的结构是固定的三层树，并不能最优的实现区分角点（实现最优的重复率）。FAST-ER就是针对这样的问题而提出的，其主要是通过模拟退火（也有通过最速下降法的）优化原先决策树的结构，从而提高重复率。

1. 引入角点检测的不变性

原先一个像素点及其附近的点送往决策树进行比较时，只需要比较两个位置的点，如果这个点被检测出是角点，但在其区域发生一定旋转、变形或强度反转（白变黑，黑变白）之后，再次重新判定，很有可能被认为是非角点。

在这种情况下，为了使角点检测具有旋转、反射、强度倒转等不变性，最简单的办法就是将所以变化后的结果都计算，即不同的变换建立不同树，只要有一个树能检测出角点，即是角点。

8. 计算机视觉教材_第8部分		总字符数：2937
相似文献列表		
去除本人文献复制比：65.3%(1919) 去除引用文献复制比：65.3%(1919) 文字复制比：65.3%(1919)		
1	<u>基于机器视觉的人体状态监测关键技术研究</u> 张宇(导师：李小霞) - 《西南科技大学硕士论文》 - 2017-06-20	20.0% (588) 是否引证：否
2	<u>基于图像局部特征的康复机器人目标识别方法研究</u> 聂海涛(导师：龙科慧) - 《中国科学院研究生院(长春光学精密机械与物理研究所)博士论文》 - 2015-04-01	16.2% (475) 是否引证：否
3	<u>基于单应性的移动机器人反馈镇定与编队跟随控制</u> 曹雨(导师：刘山) - 《浙江大学硕士论文》 - 2018-01-15	15.6% (459) 是否引证：否
4	<u>基于LOG算子的侧扫声呐图像水下小目标检测算法研究</u> 胡红波; - 《数字海洋与水下攻防》 - 2019-12-15	12.3% (361) 是否引证：否
5	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	11.5% (337) 是否引证：否
6	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	11.5% (337) 是否引证：否
7	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	11.5% (337) 是否引证：否
8	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	10.5% (308) 是否引证：否
9	<u>基于多核CPU与众核GPU车辆图像检索算法并行化研究与实现</u> 王文东(导师：周艺华) - 《北京工业大学硕士论文》 - 2016-06-01	9.3% (274) 是否引证：否
10	<u>基于改进SURF算法图像匹配方法研究</u> 储蓄(导师：林玉娥) - 《安徽理工大学硕士论文》 - 2017-06-04	9.3% (273) 是否引证：否
11	<u>运动平台视频图像去模糊算法的研究</u> 强钰琦(导师：郁文贤) - 《上海交通大学硕士论文》 - 2016-01-01	9.1% (267) 是否引证：否
12	源代码(# -*- coding: utf-8 -*- # @Time : 2017/7/24 ...) - 《源代码库 (https://raw.githubusercontent.com) 》 - 2020	7.1% (209) 是否引证：否
13	<u>放疗中患者呼吸运动与身体位移计算机监测系统的设计</u> 田菲;徐子海;朱超华;陈超敏; - 《医疗卫生装备》 - 2013-05-15	5.1% (151) 是否引证：否
14	<u>基于SIFT算法和改进最小二乘法的汽车仪表指针的识别</u> 岳晓峰;张娇; - 《机械工程师》 - 2014-12-10	5.0% (147) 是否引证：否
15	<u>遮挡目标检测与识别技术研究</u> 蔡星艳;赵和鹏;邱鹏;朱长仁; - 《数字技术与应用》 - 2013-09-15	4.8% (142) 是否引证：否

16	面向移动机器人的全方位视觉系统关键技术研究	3.6% (107)
	马朋飞(导师: 穆春阳) - 《北方民族大学硕士学位论文》 - 2017-05-18	是否引证: 否
17	平行双目视觉测量原理及实验研究	3.3% (96)
	孙雷; - 《电子技术与软件工程》 - 2022-07-15	是否引证: 否
18	基于逆投影面图像匹配的车辆外廓测量技术研究	3.3% (96)
	张斌(导师: 宋焕生) - 《长安大学硕士学位论文》 - 2018-04-10	是否引证: 否
19	视觉障碍者乘车支援MY VISION系统—公交线号检测与识别	3.0% (88)
	周叶(导师: 宋爱平; Joo Kooi Tan) - 《扬州大学硕士学位论文》 - 2021-06-01	是否引证: 否
20	基于双目视觉的混凝土布料厚度测量系统设计	2.6% (75)
	陈章华(导师: 马月辉; 李建华) - 《石家庄铁道大学硕士学位论文》 - 2020-06-01	是否引证: 否
21	一种星载平台轻量化快速影像匹配方法	2.1% (62)
	岳振宇; 范大昭; 董杨; 纪松; 李东子; - 《地球信息科学学报》 - 2022-04-29	是否引证: 否
22	运动模糊不变量研究	1.2% (35)
	蹇森; 朱剑英; - 《光电工程》 - 2007-11-15	是否引证: 否

原文内容

不过这样的话，计算量太大，为了减少计算复杂率，每次树被评估时，一般只需要应用16种变换：四个旋转方向变化（各相差90度），并结合反射（对左右对称及上下对称）同强度倒转，共16个变换操作。如果一个点能被6种变换中任一种的决策树视为角点，那么这个点就是角点。由此以来我们建立了16棵对应不同变换的决策树。

2. 决策树的结构优化

对于FAST算法来说，原来的三层决策树太过简单，不能达到最好得重复度，而重复度是关于决策树结构的非凸函数，这涉及到非凸函数的优化问题，这里许多方法，而FAST-ER则是通过模拟退火方法来优化决策树的。

4.3.4小结

本小节主要介绍了常见的角点检测方法，包括Harris角点、FAST角点和FAST-ER角点等。这些角点检测算法最核心的思想是使用一个固定窗口在图像上进行任意方向上的滑动，比较滑动前与滑动后两种情况下窗口中的像素灰度变化程度。如果任意方向上的滑动都有较大灰度变化，那么我们可以认为该窗口中存在角点。

4.4斑点检测

4.4.1斑点介绍

斑点通常是指与周围有着颜色和灰度差别的区域。在实际地图中，往往存在着大量这样的斑点，如一颗树是一个斑点，一块草地是一个斑点，一栋房子也可以是一个斑点。由于斑点代表的是一个区域，相比单纯的角点，它的稳定性要好，抗噪声能力要强，所以它在图像配准上扮演了很重要的角色。

同时有时图像中的斑点也是我们关心的区域，比如在医学与生物学领域，我们需要从一些X光照片或细胞显微照片中提取一些具有特殊意义的斑点的位置或数量。

图4- 9中建筑的洞口、蝴蝶的斑纹、花朵的花瓣和花蕊，均为各自图像中的斑点。

图4- 9 图像斑点示例

在视觉领域斑点检测的主要思路是检测出图像中比它周围像素灰度值大或比它周围像素灰度值小的区域。一般有两种方法来实现这一目标：

(1) 基于求导的微分方法，这类方法称为微分检测器；

(2) 基于局部极值的分水岭算法。

4.4.2 LOG斑点检测

利用高斯拉普拉斯（Laplace of Gaussian, LOG）算子检测图像斑点是一种十分常用的方法。对于二维高斯函数：

(4-5)

它的拉普拉斯变换为：

(4-6)

规范化的高斯拉普拉斯变换为：

(4-7)

规范化算法子在二维图像上显示是个圆对称函数，如图4- 10所示。我们可以用这个算子来检测图像中的斑点并且可以通过改变的值检测不同尺寸的二维斑点。

图4- 10 规范化算法子的二维图像显示

下面从更直观的角度解释为什么LOG算子可以检测图像中的斑点。

图像与某一个二维函数进行卷积运算实际就是求取图像与这一函数的相似性。同理，图像与高斯拉普拉斯函数的卷积实际就是求取图像与高斯拉普拉斯函数的相似性。当图像中的斑点尺寸与高斯拉普拉斯函数的形状趋近一致时，图像的拉普拉斯响应达到最大。

从概率的角度解释：假设原像是一个与位置有关的随机变量X的密度函数而LOG为随机变量Y的密度函数，则随机变量X+Y的密度分布函数即为两个函数的卷积形式。如果想让X+Y取到最大值，则X与Y能保持步调一致最好，即X上升时，Y也上升，X最大时，Y也最大。

那么LOG算子是怎么被构想出来的呢？

事实上我们知道，拉普拉斯算子可以用来检测图像中的局部极值点，但是对噪声敏感，所以在我们对图像进行拉普拉斯卷积之前，我们用一个高斯低通滤波对图像进行卷积，目标是去除图像中的噪声点。这一过程可以描述为：

先对图像 $f(x, y)$ 用方差为 σ 的高斯核进行高斯滤波，去除图像中的噪点。

$$L(x, y, \sigma) = f(x, y) * G(x, y, \sigma) \quad (4-7)$$

然后对图像作拉普拉斯变换：

$$(4-8)$$

而实际上有下面的等式

$$(4-9)$$

所以，我们可以先求高斯核的拉普拉斯算子再对图像进行卷积。也就是一开始描述的步骤。

4.4.3 DOG斑点检测

与LOG滤波核近似的高斯差分DOG滤波核，其定义为：其中， k 为两个相邻尺度间的比例因子。

如图4-11所示，DOG可以看作LOG的一个近似，但是它比LOG的效率更高。

图4-11 DOG与LOG的比较横纵坐标含义？

前面介绍的微分算子在近圆的斑点检测方面效果很好，但是这些检测算子被限定于只能检测圆形斑点，而不能估计斑点的方向，因为LOG算子都是中心对称的。如果我们定义一种二维高斯核的变形记它在 X 方向与 Y 方向上具有不同的方差，则这种算子可以用来检测带有方向的斑点。

$$(4-10)$$

其中， A 是规一性因子。

4.4.5 SIFT斑点检测

SIFT（尺度不变特征变换，Scale-Invariant Feature Transform）是计算机视觉领域中检测和描述图像中局部特征的算法该算法于1999年被David Lowe提出，并于2004年进行了补充和完善。该算法应用很广，如目标识别、自动导航、图像拼接、三维建模、手势识别、视频跟踪等。

SIFT算法所检测到的特征是局部的，而且该特征对于图像的尺度和旋转能够保持不变性。同时，这些特征对于亮度变化具有很强的鲁棒性，对于噪声和视角的微小变化也能保持一定的稳定性。SIFT特征还具有很强的可区分性，它们很容易被提取出来，并且即使在低概率的不匹配情况下也能够正确地识别出目标来。因此鲁棒性和可区分性是SIFT算法最主要的特点。

SIFT算法分为4个阶段：

（1）尺度空间极值检测。该阶段是在图像的全部尺度和全部位置上进行搜索，并通过应用高斯差分函数有效地识别出尺度不变性和旋转不变性的潜在特征点。

（2）特征点的定位。在每个候选特征点上，一个精细的模型被拟合出来用于确定特征点的位置和尺度。而特征点的最后选取依赖的是它们的稳定程度。

（3）方向角度的确定。基于图像的局部梯度方向，为每个特征点分配一个或多个方向角度。所有后续的操作都是在所确定下来的特征点的角度、尺度和位置的基础上进行的，因此特征点具有这些角度、尺度和位置的不变性。

（4）特征点的描述符。在所选定的尺度空间内，测量特征点邻域区域的局部图像梯度，将这些梯度转换成一种允许局部较大程度的形状变形和亮度变化的描述符形式。

4.4.6 SURF斑点检测

SURF（Speeded Up Robust Features）是一种具有鲁棒性的局部特征检测算法，它首先由Herbert Bay等人于2006年提出，并在2008年进行了完善。其实该算法是Herbert Bay在博士期间的研究内容，并作为博士毕业论文的一部分发表。

SURF算法的部分灵感来自于SIFT算法，但正如它的名字一样，该算法除具有重复性高的检测器和可区分性好的描述符特点外，还具有很强的鲁棒性以及更高的运算速度，如Bay所述，SURF至少比SIFT快3倍以上，综合性能要优于SIFT算法。与SIFT算法一样，SURF算法也在美国申请了专利。

之所以SURF算法有如此优异的表现，尤其是在效率上，一方面是因为该算法在保证正确性的前提下进行了适当的简化和近似，另一方面是它多次运用积分图像（Integral Image）的概念。

SURF算法包括下面几个阶段：

9. 计算机视觉教材_第9部分		总字符数：1282
相似文献列表		
去除本人文献复制比：90.2%(1156) 去除引用文献复制比：90.2%(1156) 文字复制比：90.2%(1156)		
1	SURF 算法	64.1% (822)
	- 《网络 (http://www.wendangku) 》 - 2021	是否引证：否
2	基于均值偏移算法的运动目标跟踪算法的研究	60.0% (769)
	杜凯(导师：巨永锋) - 《长安大学博士学位论文》 - 2012-10-08	是否引证：否
3	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网	58.5% (750)
	- 《互联网文档资源 (http://www.docin.com) 》 - 2012	是否引证：否
4	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网	58.5% (750)
	- 《互联网文档资源 (http://www.docin.com) 》 - 2012	是否引证：否

5	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	58.5% (750)	是否引证: 否
6	现代数字图像处理技术提高及应用案例详解 (MATLAB版) 试读 - docin.com豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2012	50.8% (651)	是否引证: 否
7	<u>无人机全景视频监控及相关技术研究</u> 于晨(导师: 王佳楠) - 《北京理工大学硕士论文》 - 2017-06-01	48.7% (624)	是否引证: 否
8	<u>基于全景视觉的移动机器人SLAM方法研究</u> 吴叶斌(导师: 朱齐丹) - 《哈尔滨工程大学博士学位论文》 - 2011-09-09	47.6% (610)	是否引证: 否
9	<u>基于低空视频的多目标检测与跟踪算法研究</u> 刘慧(导师: 李清泉) - 《武汉大学博士论文》 - 2013-10-01	43.8% (561)	是否引证: 否
10	<u>图像拼接技术研究</u> 崔汉峰(导师: 霍春宝) - 《辽宁工业大学硕士论文》 - 2013-03-01	36.2% (464)	是否引证: 否
11	<u>局部描述算子在视频序列目标检测应用中的研究</u> 潘梦涛(导师: 周宁) - 《电子科技大学硕士论文》 - 2013-05-02	21.8% (280)	是否引证: 否
12	<u>瓶盖表面排码缺陷检测与算法研究</u> 金佳(导师: 王凤文) - 《东北大学硕士论文》 - 2014-06-01	21.8% (279)	是否引证: 否
13	<u>基于移动平台增强现实系统的研究</u> 孔祥(导师: 卢朝阳) - 《西安电子科技大学硕士论文》 - 2012-01-01	20.8% (267)	是否引证: 否
14	<u>基于仿射变换的图像分块拼接方法</u> 张平梅;金立左;李久贤; - 《信息技术与信息化》 - 2020-02-10	19.0% (243)	是否引证: 否
15	算法的鲁棒性 (2020) - 《网络 (http://www.wendangku) 》 - 2021	18.9% (242)	是否引证: 否
16	<u>一种改进ORB特征匹配的半稠密三维重建ORB-SLAM算法</u> 陈文佑;章伟;史晓帆;宋芳; - 《电子科技》 - 2021-12-06	18.0% (231)	是否引证: 否
17	<u>基于ORB特征的目标检测与跟踪的研究</u> 谢成明(导师: 李小红) - 《合肥工业大学硕士论文》 - 2013-04-01	16.6% (213)	是否引证: 否
18	<u>基于词袋模型的木材分类的研究</u> 胡学晶(导师: 戴天虹) - 《东北林业大学硕士论文》 - 2017-04-01	14.0% (179)	是否引证: 否
19	<u>基于无人机遥感的智能墒情预测模型与灌溉决策系统</u> 刘晨晨(导师: 李伟) - 《江苏大学硕士论文》 - 2022-06-01	10.8% (138)	是否引证: 否
20	SURF 算法 - 《网络 (http://www.wendangku) 》 - 2021	8.6% (110)	是否引证: 否
21	<u>基于机器视觉的PCB缺陷检测系统的设计</u> 陈小勇(导师: 黄靖) - 《福建工程学院硕士论文》 - 2018-04-01	8.3% (107)	是否引证: 否
22	<u>全景视频实时拼接技术研究</u> 刘台;胡斌;田进龙;任伟嘉;高超; - 《移动通信》 - 2023-02-15	7.5% (96)	是否引证: 否
23	<u>基于RGB-NIR图像匹配的作物光谱指数特征可视化分析</u> 孙红;邢子正;张智勇;马旭颖;龙耀威;刘宁;李民赞; - 《光谱学与光谱分析》 - 2019-11-15	7.3% (94)	是否引证: 否
24	<u>基于360°全景摄像头图像识别的行车辅助系统</u> 黄思晓; - 《电子技术与软件工程》 - 2021-06-01	7.1% (91)	是否引证: 否
25	<u>基于欧氏距离变换的无人机图像拼接</u> 王鹏钊;陈宇哲;潘斐扬; - 《通信技术》 - 2021-08-10	6.7% (86)	是否引证: 否
26	学习OpenCV——Surf (特征点篇) & flann - 《互联网文档资源 (http://www.360doc.co) 》 - 2015	3.8% (49)	是否引证: 否
27	学习OpenCV——Surf (特征点篇) & flann - 《互联网文档资源 (http://www.360doc.co) 》 - 2018	3.8% (49)	是否引证: 否

原文内容

第一部分：特征点检测

(1) 基于Hessian矩阵的特征点检测：

Hessian矩阵是SURF算法的核心，构建Hessian矩阵的目的是为了生成图像稳定的边缘点（突变点），为下文的特征提取打好基础。每一个像素点都可以求出一个Hessian矩阵：

(4-11)

当Hessian矩阵的判别式取得局部极大值时，判定当前点是比周围邻域内其他点更亮或更暗的点，由此来定位关键点的位置，Hessian矩阵的判别式为：在SURF算法中，图像像素 $I(x, y)$ 即为函数值 $f(x, y)$ 。但是由于我们的特征点需要具备尺度无关性，所以在进行Hessian矩阵构造前，需要对其进行高斯滤波，选用二阶标准高斯函数作为滤波器。

(2) 尺度空间表示：

通常想要获取不同尺度的斑点，必须建立图像的尺度空间金字塔。一般的方法是通过不同 σ 的高斯函数，对图像进行平滑滤波，然后重采样图像以获得更高层的金字塔图像。SIFT特征检测算法中就是通过相邻两层图像金字塔相减得到DoG图像，然后再在DoG图像上进行斑点和边缘检测工作的。

由于采用了盒子滤波和积分图像，所以，我们并不需要像SIFT算法那样去直接建立图像金字塔，而是采用不断增大盒子滤波模板的尺寸的间接方法。通过不同尺寸盒子滤波模板与积分图像求取Hessian矩阵行列式的响应图像。然后在响应图像上采用3D非最大值抑制，求取各种不同尺度的斑点。

如前所述，我们使用 9×9 的模板对图像进行滤波，其结果作为最初尺度的尺度空间层（此时，尺度值为 $s=1.2$ ，近似 $\sigma=1.2$ 的高斯微分），后续的层将通过逐步放大滤波模板尺寸，以及放大后的模板不断与图像进行滤波得到。由于采用盒子滤波和积分图像，滤波过程并不随着滤波模板尺寸的增加而使运算工作量增加。

与SIFT算法类似，我们需要将尺度空间划分为若干组（Octaves）。一个组代表了逐步放大的滤波模板对同一输入图像进行滤波的一系列响应图。每个组又由若干固定的层组成。由于积分图像离散化的原因，两个层之间的最小尺度变化量是由高斯二阶微分滤波器在微分方向上对正负斑点响应长度 $1_{\{0\}}$ 决定的，它是盒子滤波器模板尺寸的 $1/3$ 。对于 9×9 的模板，它的 $1_{\{0\}}=3$ 。下一层的响应长度至少应该在 $1_{\{0\}}$ 的基础上增加2个像素，以保证一边一个像素，即 $1_{\{0\}}=5$ 。这样模板的尺寸就为 15×15 。以此类推，我们可以得到一个尺寸增大模板序列，它们的尺寸分别为： $9 \times 9, 15 \times 15, 21 \times 21, 27 \times 27, 33 \times 33, 39 \times 39, 45 \times 45, 51 \times 51, 57 \times 57, 63 \times 63$ ，黑色、白色区域的长度增加偶数个像素，以保证一个中心像素的存在。

(3) 特征点定位。

为了在图像及不同尺寸中定位兴趣点，我们用了 $3 \times 3 \times 3$ 邻域非最大值抑制。具体的步骤基本与SIFT一致，而且Hessian矩阵行列式的最大值在尺度和图像空间被插值。

总体来说，如果理解了SIFT算法，再来看SURF算法会发现思路非常简单。尤其是局部最大值查找方面，基本一致。关键还是一个用积分图来简化卷积的思路，以及怎么用不同的模板来近似原来尺度空间中的高斯滤波器。

10. 计算机视觉教材_第10部分		总字符数：10000
相似文献列表		
去除本人文献复制比：56.5%(5646) 去除引用文献复制比：56.5%(5646) 文字复制比：56.5%(5646)		
1	基于相机扰动校正的桥梁结构变形测量方法与应用 于姗姗(导师：张建) - 《东南大学博士论文》 - 2021-04-01	23.8% (2382) 是否引证：否
2	基于特征的图像匹配算法研究与应用 吴洁琼(导师：李建平) - 《电子科技大学硕士论文》 - 2016-03-28	8.7% (871) 是否引证：否
3	图像匹配技术综述 戴涛;朱长仁;胡树平; - 《数字技术与应用》 - 2012-03-15	4.9% (485) 是否引证：否
4	BRIEF 特征描述子详述（转载）_郑佳 - 《网络（ http://blog.sina.com ）》 - 2019	4.8% (481) 是否引证：否
5	图像匹配技术的研究 - 豆丁网 - 《互联网文档资源（ http://www.docin.com ）》 - 2012	4.5% (454) 是否引证：否
6	图像匹配技术的研究 殷伶(导师：刘贵喜) - 《西安电子科技大学硕士论文》 - 2010-01-01	4.4% (440) 是否引证：否
7	数字图像匹配技术研究 卜凡艳(导师：檀结庆) - 《合肥工业大学硕士论文》 - 2010-04-01	4.1% (414) 是否引证：否
8	基于图像的虚拟场景漫游系统关键技术的研究 刘雪香(导师：刘晓红) - 《山东理工大学硕士论文》 - 2009-04-16	4.0% (398) 是否引证：否
9	基于SIFT特征描述子的立体匹配算法研究 宰小涛(导师：赵宇明) - 《上海交通大学硕士论文》 - 2007-01-01	3.7% (367) 是否引证：否
		3.6% (358)

10	<u>物体内部三维位移场测量算法研究</u> 牛永强(导师:刘贵全;缪泓) - 《中国科学技术大学硕士论文》 - 2010-04-01	是否引证: 否
11	<u>虚拟全景空间和双目视觉三维重建的研究与实现</u> 吴忠年(导师:张建中) - 《电子科技大学硕士论文》 - 2008-04-01	3.5% (345) 是否引证: 否
12	<u>基于径向基函数的高精度变形测量方法研究</u> 代祥俊(导师:杨福俊) - 《东南大学博士论文》 - 2016-05-01	3.4% (343) 是否引证: 否
13	<u>基于特征点的图像配准技术及应用</u> 冯晓伟(导师:田裕鹏) - 《南京航空航天大学硕士论文》 - 2008-01-01	3.2% (315) 是否引证: 否
14	<u>图像相似度的匹配研究 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co)》 - 2020	3.1% (311) 是否引证: 否
15	<u>图像识别匹配技术原理要点 - 百度文库</u> - 《互联网文档资源 (https://wenku.baidu)》 - 2019	2.8% (282) 是否引证: 否
16	<u>图像识别匹配技术原理下载</u> - 《互联网文档资源 (https://ishare.iask)》 - 2020	2.8% (282) 是否引证: 否
17	<u>基于深度学习的飞行器图像自主导航算法研究</u> 刘瑞(导师:冷雪飞) - 《南京航空航天大学硕士论文》 - 2021-03-01	2.7% (272) 是否引证: 否
18	<u>图像配准 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co)》 - 2020	2.7% (269) 是否引证: 否
19	<u>高精度、实时数字图像相关变形测量</u> 邵新星(导师:何小元) - 《东南大学博士论文》 - 2018-04-01	2.7% (269) 是否引证: 否
20	<u>毕业论文(图像匹配)-百度文库</u> - 《互联网文档资源 (http://wenku.baidu.c)》 - 2012	2.7% (266) 是否引证: 否
21	<u>基于双目立体视觉的工件识别定位方法研究</u> 梁元月(导师:傅卫平) - 《西安理工大学硕士论文》 - 2008-03-01	2.6% (262) 是否引证: 否
22	<u>小波阈值去噪及其在数字图像相关中的应用研究</u> 胡斌(导师:贺玲凤) - 《华南理工大学硕士论文》 - 2011-05-01	2.6% (259) 是否引证: 否
23	<u>基于改进光流法的视觉SLAM</u> 刘冰(导师:李奇敏;李孝揆) - 《重庆大学硕士论文》 - 2018-05-01	2.5% (254) 是否引证: 否
24	<u>基于图像的可视媒体构建</u> 邢静(导师:苗振江) - 《北京交通大学硕士论文》 - 2007-12-01	2.4% (244) 是否引证: 否
25	<u>形状相似度算法研究及其在相似人脸匹配中的应用</u> 周晟毅(导师:吴飞) - 《浙江大学硕士论文》 - 2007-05-01	2.4% (244) 是否引证: 否
26	<u>基于特征点的影像匹配方法探讨</u> 王雅丽;厉彦玲;蔡玉林; - 《测绘科学》 - 2012-10-23 0	2.2% (215) 是否引证: 否
27	<u>基于点不变特征的亚像素级图像配准技术</u> 傅德友;赵秀影;梁心;商玉凤;董艳芹; - 《光电技术应用》 - 2010-06-15	2.2% (215) 是否引证: 否
28	<u>基于SIFT快速算法的单目立体视觉应用研究</u> 赵亮(导师:刘建辉) - 《辽宁工程技术大学硕士论文》 - 2010-05-01	2.2% (215) 是否引证: 否
29	<u>基于改进SIFT算法的图像搜索技术研究</u> 冯镜蒯(导师:王知衍) - 《华南理工大学硕士论文》 - 2011-05-01	2.2% (215) 是否引证: 否
30	<u>基于多标签学习的图像区域语义自动标注算法研究</u> 滕舟(导师:郭跃飞) - 《复旦大学硕士论文》 - 2011-04-20	2.1% (209) 是否引证: 否
31	<u>基于机器视觉的车用胶管表面质量检测系统研发</u> 张绍恒(导师:张牧;杨青) - 《天津工业大学硕士论文》 - 2021-01-01	1.9% (185) 是否引证: 否
32	<u>应变光学测量关键技术及在特种设备监测的应用</u> 陈光(导师:冯其波;高瞻) - 《北京交通大学博士论文》 - 2018-06-01	1.8% (180) 是否引证: 否
33	<u>特定场景大视角视频拼接技术研究</u> 林财明;卫志敏; - 《电脑与电信》 - 2017-04-10	1.6% (160) 是否引证: 否

34	改进的数字图像应变测量方法	1.4% (143)
	张樟;徐飞鸿; - 《北华航天工业学院学报》 - 2017-08-28	是否引证: 否
35	基于步态识别的跨摄像头行人再识别算法研究	1.1% (109)
	邵佳耀;宋春林; - 《信息技术与信息化》 - 2019-12-25	是否引证: 否
36	高炉出渣沟熔渣流量实时检测系统研究	1.0% (97)
	刘洪刚(导师: 徐克宝) - 《山东科技大学硕士学位论文》 - 2011-05-01	是否引证: 否
37	基于数字图像相关法的视频引伸计研究	0.7% (69)
	何晓晨(导师: 朱建国) - 《江苏大学硕士学位论文》 - 2021-06-01	是否引证: 否
38	电梯钢丝绳无损检测系统设计	0.7% (65)
	顾雯雯; - 《中国电梯》 - 2019-11-01	是否引证: 否
39	基于特征点法和直接法VSLAM的研究	0.5% (50)
	邹雄;肖长诗;文元桥;元海文; - 《计算机应用研究》 - 2019-03-12 1	是否引证: 否
40	基于LOG算子的侧扫声呐图像水下小目标检测算法研究	0.5% (47)
	胡红波; - 《数字海洋与水下攻防》 - 2019-12-15	是否引证: 否
41	基于互信息的多模医学图像配准	0.3% (30)
	陈显毅; - 《琼州学院学报》 - 2008-10-28	是否引证: 否
42	红外和可见光图像配准技术的研究	0.3% (30)
	吴文浩(导师: 蔡向东) - 《哈尔滨理工大学硕士学位论文》 - 2021-06-01	是否引证: 否

原文内容

第二部分: 特征点描述

(1) 方向角度的分配;

(2) 基于Haar小波的特征点描述符。

图4- 12 SURF斑点检测示例

SURF斑点检测示例的主程序如下。

```
clc,clear,close all
I=imread('GZA Bridge_RGB.jpg');
I=I(51:358,251:558,1:3);
I=rgb2gray(I);
points = detectSURFFeatures(I); % SURF Features
figure
imshow(I); hold on; plot(points.selectStrongest(30));
```

4.4.7小结

本小节主要介绍了常见的斑点检测方法,包括LOG斑点检测、DOG斑点检测、DOH斑点检测、SIFT斑点检测和SURF斑点检测等,斑点检测的主要思路是检测出图像中比它周围像素灰度值大或比它周围像素灰度值小的区域,目前应用较多的斑点检测算法是LOG和SIFT算法。

4.5特征描述子

4.5.1特征描述子介绍

特征描述子 (Feature Descriptors) 指的是检测图像的局部特征 (比如边缘、角点、轮廓等), 然后根据匹配目标的需要进行特征的组合变换, 以形成易于匹配、稳定性好的特征向量, 从而把图像匹配问题转化为特征的匹配问题, 进而将特征的匹配问题转化为特征空间向量的聚类问题。

4.5.2 BRIEF描述子

BRIEF (Binary Robust Independent Elementary Features) 与传统的利用图像局部邻域的灰度直方图或梯度直方图提取特征的方式不同, BRIEF是一种二进制编码的特征描述子, 既降低了存储空间的需求, 提升了特征描述子生成的速度, 也减少了特征匹配时所需的时间。

值得注意的是, 对于BRIEF, 它仅仅是一种特征描述符, 它不提供提取特征点的方法。所以, 我们必须还要使用一种特征点定位的方法, 如FAST、SIFT、SURF等。这里, 我们将使用CenSurE方法来提取关键点, 对 BRIEF来说, CenSurE的表现比SURF特征点稍好一些。

BRIEF的算法步骤如下:

它需要先平滑图像, 然后在特征点周围选择一个Patch, 在这个Patch内通过一种选定的方法挑选出nd个点对。然后对于每一个点对 (p, q), 我们来比较p, q这两个点的亮度值。如果 $I(p) > I(q)$, 则对应二值串中的值为1; 如果 $I(p) < I(q)$, 则对应二值串中的值为-1; 否则为0。所有nd个点对都进行比较, 我们就生成了一个nd长的二进制串。

对于nd的选择我们可以设置为128, 256或512, 这三种参数在OpenCV中都有提供, 但是OpenCV中默认的参数是256, 在这种情况下, 非匹配点的汉明距离呈现均值为128bit的高斯分布。一旦维数选定了, 我们就可以用汉明距离来匹配这些描述子了。

4.5.3 ORB特征提取算法

ORB特征, 从它的名字中可以看出它是对FAST特征点与BRIEF特征描述子的一种结合与改进, 这个算法是由Ethan

Ruble, Vincent Rabaud, Kurt Konolige以及Gary R. Bradski在2011年一篇名为“ORB: An Efficient Alternative to SIFT or SURF”的文章中提出。就像文章题目所写一样, ORB是除SIFT与SURF外一个很好的选择, 而且它有很高的效率, 重要的一点是它是免费的, SIFT与SURF都是有专利的, 如果在商业软件中使用, 需要购买许可。

ORB特征是将FAST特征点的检测方法与 BRIEF特征描述子结合起来, 并在它们原来的基础上做了改进与优化。

首先, 它利用FAST特征点检测的方法来检测特征点, 然后利用Harris角点的度量方法, 从FAST特征点中挑选出Harris角点响应值最大的N个特征点。其中Harris角点的响应函数定义为:

(4-12)

图4- 13 ORB特征提取示例

ORB特征提取示例的主程序如下。

```
clc,clear,close all
I=imread('GZA Bridge_RGB.jpg');
I=I(51:358,251:558,1:3);
I=rgb2gray(I);
points = detectORBFeatures(I); % SURF Features
figure
imshow(I); hold on; plot(points.selectStrongest(30));
```

4. 5. 4 BRISK特征提取算法

BRISK算法是2011年ICCV上一篇名为“BRISK: Binary Robust Invariant Scalable Keypoints”的文章中提出来的一种特征提取算法, 也是一种二进制的特征描述算子。它具有较好的旋转不变性、尺度不变性, 较好的鲁棒性等。在图像配准应用中, 速度比较如下: SIFT< SURF< BRISK< FREAK< ORB, 在对有较大模糊的图像配准时, BRISK算法表现最为出色。

BRISK算法分两步进行: 第一步进行特征点检测; 第二步进行特征点描述。

在特征点检测当中, 主要分为建立尺度空间、特征点检测、非极大值抑制和亚像素差值这四个部分; 而在特征点描述当中, 主要分为高斯滤波、局部梯度计算特征描述符和匹配方法这四个部分。

图4- 14 BRISK特征提取示例

BRISK特征提取示例的主程序如下。

```
clc,clear,close all
I=imread('GZA Bridge_RGB.jpg');
I=I(51:358,251:558,1:3);
I=rgb2gray(I);
points = detectBRISKFeatures(I);
figure
imshow(I); hold on; plot(points,'ShowScale',false)
```

4. 5. 5 FREAK特征提取算法

FREAK算法是2012年CVPR上一篇名为“FREAK: Fast Retina Keypoint”的文章中提出来的一种特征提取算法, 也是一种二进制的特征描述算子。它与BRISK算法非常相似, 主要就是在BRISK算法上的改进。FREAK依然具有尺度不变性旋转不变性对噪声的鲁棒性等。

FREAK的主要步骤包括: 采样模式、特征描述、特征方向和特征匹配。

4. 5. 6小结

本小节主要介绍了常见的特征描述子算法, 包括BRIEF描述子、ORB特征描述算法、BRISK特征描述算法和FREAK特征描述算法等特征描述子算法的主要思路是检测出图像的局部特征(比如边缘角点轮廓等), 然后根据匹配目标的需要进行特征的组合、变换, 以形成易于匹配、稳定性好的特征向量。

4. 6图像匹配方法

图像匹配(image matching)指在目标图像中搜索与给定模板图像相似图像的过程, 需在同一场景的两幅或多幅图像之间寻找对应关系, 广泛应用于目标识别、三维重构运动跟踪等领域。

图4- 15 特征匹配流程

图像匹配算法主要分为基于灰度的匹配算法和基于特征的特征匹配算法。基于灰度的匹配算法建立在整个图像的总特征上, 以两幅图像中含有的相应目标和搜索区域中的像素灰度为基础、由最佳度量值(如协方差或相关系数)判断两幅图像中的对应点。在灰度及几何畸变不大的情况下, 该算法有较好的匹配精度和鲁棒性, 抗噪能力强。但是该匹配算法运算量大, 速度慢, 抗几何畸变能力弱。归一化灰度相关匹配、最小二乘图像匹配、贝叶斯图像匹配等是常见的基于灰度的匹配算法。基于特征的特征匹配算法以图像中提取的关键特征为匹配基元, 具体包含特征提取和特征匹配两步。一幅图像经过特征提取后, 每个特征模式具有位置信息与特征描述量, 采用的匹配特征通常为角点、边缘等。在特征匹配前, 先把感兴趣的图像将任用特征提取算子检测出来, 然后在两幅图像对应的特征集中利用特征匹配算法, 将存在匹配关系的特征对选择出来, 实现两幅图像之间的快速匹配(Hollwell, 2000)。本节讨论的图像匹配是指一个模板图像与包含目标的图像之间通过一定的变换实现像素级的位置匹配的过程。作为模式识别和数字图像处理的一种基本手段, 图像匹配在目标检测、目标跟踪等计算机视觉各方面具有广泛的应用。

gfgf以与分别表示源图像与待匹配图像, 则源图像与待匹配图像配准关系的数学表达式为: , 其中代表二维几何空间变换函数, 为一维灰度变换函数。从以上图像配准关系表达式可以看出, 图像配准包括两方面含义, 一方面是实现几何空间上的对应关系, 这点可以在对函数的求解上得以体现; 另一方面是对应像素之间灰度上的一致性, 这一点可以在对函数的求解上得以体现。因此, 图像配准的主要目的可以看做是寻找配准图像间空间与灰度的最佳变换关系。在考虑畸变的前提下, 实现图像的最佳匹配。

根据匹配基元的不同, 现有的图像匹配方法可分为四类: 基于像素灰度的匹配方法、基于特征的特征匹配方法、基于变换域的

匹配方法和基于投影的匹配方法。

图4- 16 模板图像在目标图像中的移动过程

基于图像的目标追踪是指通过对影像内容、特征、结构、关系、纹理及灰度等的对应关系、相似性和一致性的分析，寻求相似影像目标的一种方法。

追踪第一步是在待匹配图像上任意选取一个像素点，并以此为中心得到一个区域（目标窗口）；第二步是在参考图像上选择同样尺寸的区域（参考窗口），并且以一定的方式移动该参考窗口，在移动前比对参考窗口中的图像与待匹配图像上的待匹配窗口内图像是否匹配，若匹配成功则停止移动并进行待匹配图像下一像素点的匹配检索，直到检索全部像素点，否则将目标窗口移动到下一位置进行匹配比对，直到遍历完整个参考图。

图像匹配算法通常由特征空间、相似度量、搜索空间、搜索策略四个要素构成。

特征空间

是指在图像匹配的过程中，可以在匹配中作为参照的，能够提高图像匹配精确度、效率的特征所构成的集合。一幅图像中有许多显著的量可以作为特征，如矩不变量、灰度值、高曲率点、线交叉点等。合理的选择特征空间能够降低搜索空间范围，提高匹配的准确率和效率，从而在匹配算法中可以使用全局特征、局部特征或者两者的结合来优化图像匹配的精确度。

相似度量

是指基于某种代价函数或者距离函数来衡量待匹配图像与参考图像之间特征向量的相似性程度，经典的度量方法有欧式空间距离、曼哈顿距离、闵可夫斯基距离、余弦相似等。

搜索空间

是指在待匹配图像与参考图像之间建立联系的所有可能变换全体构成的集合。因此，搜索空间的范围由图像畸变的类型和强度所决定，其中畸变的类型有光流场法、全局变换、局部变换。

搜索策略

是指在搜索空间中以一种高效的方式检索出一个几何变换，使得待匹配图像与参考图像具有最大的相似度。广泛使用的搜索策略的种类相当丰富，如穷举法、松弛算法、层次迭代搜索等。可是由于搜索效率的问题，在制定搜索策时，要充分考虑搜索空间的复杂性，选取高效的搜索方法。

从待匹配的研究对象来看，图像匹配算法一般分为基于灰度分布的、基于特征的和基于变换域的图像目标匹配三种。基于灰度的方法也称作相关匹配算法，用空间二维滑动模板进行图像匹配，这类方法又细分为不同算法，他们的区别主要体现在模板及相关准则的选择方面。基于特征匹配的方法，首先在原始图像中提取特征，然后再建立两幅图像之间特征的匹配对应关系。基于变换域的方法是通过建立两幅图像的相位差与图像间的平移量的关系获取目标位置变动信息的。下文对三种算法进行分类概述。

4. 6. 1 基于像素灰度的匹配

基于像素灰度的图像匹配是指利用图像的灰度信息对两幅或者多幅图像进行匹配，是一类最简单的匹配算法。该算法利用图像灰度的一些统计信息来度量图像的相似程度。

定义模板图像为 T ，待匹配的目标图像为 I 。沿目标图像中的所有点移动模板图像并在每个位置计算相似度 S 。 S 是一个相似度函数，该函数的参数包括模板中各点的灰度值以及待匹配区域移到图像当前位置时区域中的灰度值。根据这些已经得到的灰度值计算一个标量值作为相似度量。采用这个方法，在变换空间中每个点都会得到一个相似度量如下，即

(4-13)

最简单的相似性度量方法是计算模板图像与待匹配的目标图像之间差值的绝对值总和 (Sum of Absolute Differences, SAD) 或者差值的平方和 (Sum of Squared Differences, SSD)。在模板图像和待匹配的目标图像完全相同的情况下，计算后所得到的相似度量应该为 0。区别越大，差值越大。在光照情况保持不变的情况下，相似性度量方法的结果准确度比较高。但是在光照发生变化的情况下，图像中的灰度值已经不再相等，采用该方法结果会发生偏差。

基于像素灰度的图像匹配算法实现简单，在灰度及几何畸变不大的情况下有较好的估计精度和鲁棒性，抗噪性也比较强。其不足是应用范围较窄，不能直接用于校正图像的非线性形变，在最优变换的搜索过程中往往需要较大的运算量等。

GM \times NIm \times n 图像互相关算法是一种经典的基于灰度的图像匹配算法。现有一个大小为 $M \times N$ 的模板图像和大小为 $M' \times N'$ 的目标图像，模板中各点灰度值为 $T(x, y)$ ，目标图像中各点灰度值为 $I(x, y)$ 。现目标图像中有一同模板等大小的区域，且该区域相对于目标图像原点位移为 (x_0, y_0) ，则该区域与模板图像的灰度互相关的函数为

(4-14)

当模板图像在目标图像上逐行逐列遍历 (见图 12-35) 后，在每一个位置上可获得一个相关性值，如图 12-36 所示，则相关性峰值所在的位置即为匹配位置。在实际使用过程中，往往会设定一个相关性阈值，超出该阈值的峰值所在位置将被视为匹配位置，若出现多个相关性峰值，则认为有多个匹配位置。

在实际使用中，由于光照的变化，纯粹只用相关性来进行灰度匹配并不能获得稳定的匹配效果。为了解决这个问题，人们在原有灰度相关性的基础上发展了归一化灰度互相关算法。假设表示模板 G 的均值，表示目标 I 的均值，则定义相关系数

(4-15)

其中，

以归一化互相关系数作为模板图像与目标图像之间的相似性度量，能够克服任何可线性光照的干扰影响，从而达到较为良好的匹配效果。

图4- 17 NCC模板匹配结果

图4-13NCC模板匹配示例的MATLAB程序如下。

```
clc; clear; close all
%----- Read two grayscale images for use with "normxcorr2"
frame_1 = imread('GZA Bridge.png'); % import the 1st video frame
frame_1 = rgb2gray(frame_1);
%----- Select Template T from 1st video frame
```



```
T=imcrop(frame_1);
imshow(T)
imwrite(T,'template.jpg')
close all
%----- Perform cross-correlation, and display the result as a surface.
c1 = normxcorr2(T,frame_1);
figure(1), surf(c1), shading flat
%----- Find the peak in cross-correlation.
[ypeak1, xpeak1] = find(c1==max(c1(:)));
%----- Account for the padding that normxcorr2 adds.
yoffSet1 = ypeak1-size(T,1);
xoffSet1 = xpeak1-size(T,2);
%----- Display the matched area and position
figure(3)
imshow(frame_1);
imrect(gca, [xoffSet1+1, yoffSet1+1, size(T,2), size(T,1)]);
title('Template matching using NCC','FontName','Arial','fontsize', 17)
text(xoffSet1,yoffSet1-20,['x=',num2str(xoffSet1),' ', y=',num2str(yoffSet1)],'FontName','Arial','fontsize',
24,'color','r')
```

表4- 1 基于像素灰度的匹配算法

算法名称	简称	算法公式
Sum of Absolute Differences	SAD	
Zero mean Sum of AbsoluteDifferences	ZSAD	
Sum of Squared Differences	SSD	
Zero mean Sum of SquaredDifferences	ZSSD	
Normalised Cross Correlation	NCC	
Zero mean Normalised Cross Correlation	ZNCC	

注：I1代表模板图像，I2代表目标图像。

算法名称简称算法公式

- Sum of Absolute Differences SAD
- Zero mean Sum of AbsoluteDifferences ZSAD
- Sum of Squared Differences SSD
- Zero mean Sum of SquaredDifferences ZSSD
- Normalised Cross Correlation NCC
- Zero mean Normalised Cross Correlation ZNCC

注：I1代表模板图像，I2代表目标图像。

灰度图像匹配算法的基本思想：对两幅图像寻找相关性，利用统计学相关思想，判定两幅图像中的对应关系，其最常用的就是相似性度量方法：如相关函数、平方差、协方差等测度极值。基于灰度信息的图像配准方法一般不需要对图像进行复杂的预先处理，而是利用图像本身具有灰度的一些统计信息来度量图像的相似程度。该类方法的主要特点是实现简单，但应用范围较窄，不能直接用于校正图像的非线性形变，在最优变换的搜索过程中往往需要巨大的运算量。经过几十年的发展，人们提出了许多基于灰度信息的图像配准方法，大致可以分为三类：互相关法、序贯相似度检测法、交互信息法。前面两个都属于模板匹配，只不过后者是前者的提高版本。

互相关法的概念最为深入人心。1982年，Rosenfeld提出的互相关法。它像是一种相似性度量或者匹配程度的表征，而不是一种图像匹配的完整方法，但是把互相关的思想作为度量测度，在许多匹配算法里都会用到。A. Roche等人改进后的算法能够较好地解决噪声的问题，而且计算速度也提高不少。典型的相关准则包括：平均绝对差算法（Mean Absolute Differences，简称MAD算法），绝对误差和算法（SAD，Sum of Absolute Differences）。误差平方和算法（SSD，Sum of Squared Differences），也叫差方和算法。实际上，SAD算法与MAD算法和SSD算法思想几乎是完全一致的，这里不再赘述。平均误差平方和算法（MSD，Mean Square Differences），也称均方差算法。归一化积相关算法（NCC，Normalized Cross Correlation），与上面算法相似，依然是利用子图与模板图的灰度，通过归一化的相关性度量公式来计算二者之间的匹配程度。序贯相似性检测算法（SSDA，Sequential Similiarity Detection Algorithm）是1972年，Bamea等人根据传统相关方法提出的，是对传统模板匹配算法的改进，比MAD算法快几十到几百倍。这种算法在两方面有了显著的改进，一是简化计算，利用图像T和模板f之间的差值来表示变化。与相关法相比处理效果差不多，同时显著的提高了运算速度。二是改进使用了一种序列搜索的策略，由检测范围和模板大小定义了一系列窗函数和阈值，而每一个窗函数作用到图像中，当相似性超过阈值后，就进行次数累加，而后在次数最多的窗口里进行匹配，重复迭加细化直至得到所需要的结果。

基于灰度信息的变形测量方法中对具有代表性的是数字图像相关（DIC，Digital image correlation）方法，利用相机拍摄变形前后被测平面物体表面的数字散斑图像，再通过匹配变形前后数字散斑图像中的对应图像子区获得被测物体表面各点的变形（位移与应变）。DIC秉承了模板匹配的思想。但DIC又区别于模板匹配：模板匹配只考虑目标的平行移动，不适用于目标发生旋转或大小变化的情况。而DIC将形函数添加到在互相关准则中，该函数定义了被测面可能发生的形变（位移、应变），所以说，DIC丰富了互相关法的理念。由此，针对复杂的变形测量，DIC独立分支出来，形成其独特的一整套方法和研究，包括随机散斑的喷制和优化设计。总得来说，模板匹配的用途更多的是图像识别，如果用于测量变形也只适用于简单的平移，而DIC的发展基本是针对变形测量，根据变形的复杂情况可以调整形函数的表达方式。利用DIC实现高精度的变形测量，要涉及到相关函

数、形函数、初值估计、亚像素迭代4个关键问题，下文对此简述。

(1) 相关函数

在相关计算之前，需要预先定义一个评价变形前后图像子区相似程度的函数，即相关函数（Correlation creteria），它是待求变形参数的函数。数字图像相关方法就是通过寻找与变形前图像子区的相关函数为极值的目标图像子区来获得准确的变形参数估计。相关函数有多种数学表达形式，不同的相关函数将可能导致不同的计算速度和计算精度。目前常用的相关函数有互相关（Cross correlation, CC）函数和最小平方距离（Sum-Squared difference, SSD）函数两类。考虑到各种相关函数的抗干扰能力和精度，目前零均值归一化最小平方距离（Zero-mean normalized sum of squared difference, ZNSSD）相关函数和参数化最小平方距离（Parametric sum of squared difference, PSSD）相关函数是两种最为推荐使用的相关函数。

(4-16)

其中：

$f(x)$ 与 $g(x)$ 分别表示参考图和变形图中 $x=(x, y, 1)^T$ 位置的灰度；表示图像子区的平均灰度值；表示图像子区的平均灰度值；是像素点在各个子区的局部坐标；和表示形函数。

ZNSSD 相关系数实际上与通常使用的ZNCC（Zero-mean normalized cross-correlation，零均值归一化互相关）准则存在关系：

(4-17)

(4-18)

其中：分别是参考图像子区中坐标为点和目标图像子区中对应点的灰度；

为参考图像子区的灰度平均值；

和目标图像子区的灰度平均值；

p 为待求的变形参考数矢量。

(2) 形函数

由于变形后的图像子区不仅其中心位置会发生变化，而且其形状也可能改变。因此，变形前后图像子区中对应点的坐标 (x, y) ，和 (x', y') 可通过所谓的“形函数”和待定参数矢量 p 联系：

(4-18)

如果目标图像子区相对与参考图像子区只有平移，则可用零阶形函数来描述：（4-19）

上式给出的零阶形函数不允许变形后的图像子区出现刚体转动、剪切或伸缩变形。考虑到桥梁这一类大型结构的局部变形很小，尤其是在远距离图像测量时，一般选用零阶形函数足够，也就是人为目标只发生刚体平移，这种情况下对于相关系数表达式中的形函数可以表示为：

(4-20)

其中，是目标子区的预计算的位移向量。 u, v 分别表示水平和竖向的位移。

但在多数情况下，零阶形函数不能精确描述变形后的目标图像子区的形状变化。因此，一阶形函数（允许变形后的图像子区出现刚体转动、剪切或伸缩变形或其组合）更为常用：（4-21）

(3) 初值估计

相关函数是未知的变形参数矢量的非线性方程，目前广泛使用的迭代方法其收敛范围只有几个像素，只有较准确的初值估计才能使之迅速收敛，从而获得准确可靠的位移计算结果。通常目标图像子区相对参考图像子区的变形不大，因此可以通过简单的整像素位移搜索方法得到准确的位移初值估计。问题的困难在于某些情况下变形后图像相对与参考图像可能有较大的刚体转动或出现大变形，此时整像素位移相关搜索时已没有全场唯一的尖锐相关峰，相关搜索失效。

11. 计算机视觉教材_第11部分			总字符数：10103
相似文献列表			
去除本人文献复制比：45.4%(4583)		去除引用文献复制比：45.4%(4583)	文字复制比：45.4%(4583)
1	基于相机扰动校正的桥梁结构变形测量方法与应用 于姗姗(导师：张建) - 《东南大学博士论文》- 2021-04-01	27.8% (2807)	是否引证：否
2	基于图像匹配技术的桥梁位移无标记测量方法研究 汪威(导师：吴刚;吴栋) - 《东南大学硕士论文》- 2021-05-30	9.7% (982)	是否引证：否
3	基于深度学习的行车视频中的目标检测 亢福胜(导师：王海滨;杨勇) - 《西华大学硕士论文》- 2019-04-01	3.2% (325)	是否引证：否
4	基于计算机视觉的高铁桥梁结构位移测量方法研究 刘子琦(导师：李忠龙) - 《哈尔滨工业大学硕士论文》- 2020-06-01	2.9% (289)	是否引证：否
5	基于图像序列的桥梁形变位移测量方法 黄建坤(导师：龚涛) - 《西南交通大学硕士论文》- 2018-05-01	2.6% (263)	是否引证：否
6	高精度、实时数字图像相关变形测量 邵新星(导师：何小元) - 《东南大学博士论文》- 2018-04-01	2.4% (239)	是否引证：否
7	双波段野外火灾图像识别及目标定位方法研究	2.3% (230)	

	刘媛珺(导师: 赵敏) - 《南京航空航天大学硕士学位论文》 - 2009-01-01	是否引证: 否
8	<u>图像配准技术及其应用的研究</u> 宋智礼(导师: 陈雁秋) - 《复旦大学博士论文》 - 2010-04-01	2.3% (229) 是否引证: 否
9	<u>基于激光传感器的移动机器人定位和地图构建研究</u> 刘丽雯(导师: 张崇巍) - 《合肥工业大学硕士学位论文》 - 2009-04-01	2.2% (221) 是否引证: 否
10	<u>图像拼接算法研究</u> 刘冬梅(导师: 刘贵喜) - 《西安电子科技大学硕士学位论文》 - 2008-01-01	2.1% (216) 是否引证: 否
11	<u>图像匹配技术综述</u> 戴涛;朱长仁;胡树平; - 《数字技术与应用》 - 2012-03-15	2.1% (208) 是否引证: 否
12	<u>图像配准技术研究</u> 吴晨阳(导师: 牛海军;张建国) - 《西安电子科技大学硕士学位论文》 - 2013-01-01	2.0% (199) 是否引证: 否
13	<u>基于CCD成像仪多光谱图像的配准方法研究</u> 吴晓宁(导师: 夏德深) - 《南京理工大学硕士学位论文》 - 2006-06-01	1.9% (197) 是否引证: 否
14	<u>图像拼接算法研究1 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co)》 - 2020	1.9% (197) 是否引证: 否
15	<u>——关于图像配准技术研究 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co)》 - 2020	1.9% (195) 是否引证: 否
16	<u>图像配准理论及其算法研究</u> 王鑫(导师: 赵春晖) - 《哈尔滨工程大学硕士学位论文》 - 2005-01-01	1.9% (189) 是否引证: 否
17	<u>旋转弹光电成像制导图像处理关键算法研究</u> 刘忠领(导师: 周立伟) - 《北京理工大学博士论文》 - 2014-12-01	1.8% (186) 是否引证: 否
18	<u>图像拼接算法研究 - 道客巴巴</u> - 《互联网文档资源 (https://www.doc88.co)》 - 2020	1.8% (180) 是否引证: 否
19	<u>图像拼接算法研究-百度文库</u> - 《互联网文档资源 (http://wenku.baidu.c)》 - 2012	1.7% (171) 是否引证: 否
20	<u>变电站红外遥视图像配准研究</u> 张兴华(导师: 吴冬梅) - 《西安科技大学硕士学位论文》 - 2010-06-30	1.7% (170) 是否引证: 否
21	<u>基于红外成像的变电站设备安全预警关键技术研究</u> 蔺丽华(导师: 刘健) - 《西安科技大学博士论文》 - 2014-06-30	1.7% (170) 是否引证: 否
22	<u>木材CT图像配准方法研究</u> 张训华(导师: 业宁) - 《南京林业大学硕士论文》 - 2010-06-01	1.6% (157) 是否引证: 否
23	<u>融合显示红外在线监测系统的实用化研究</u> 王洪光(导师: 魏臻) - 《天津理工大学硕士论文》 - 2006-12-01	1.1% (110) 是否引证: 否
24	<u>图像拼接技术研究</u> 周剑军(导师: 欧阳宁) - 《桂林电子科技大学硕士论文》 - 2009-04-01	1.1% (110) 是否引证: 否
25	<u>多模态和大型图像配准技术研究</u> 王东峰(导师: 邹谋炎) - 《中国科学院研究生院(电子学研究所)博士论文》 - 2002-05-14	0.9% (91) 是否引证: 否
26	<u>基于计算机视觉的复杂隔震结构支座变形监测及分析</u> 杜永峰;李向雄;张超;李超;麦麦提明·图尔迪麦麦提;马振和; - 《工业建筑》 - 2022-11-01 08:54	0.4% (42) 是否引证: 否

原文内容

(4) 亚像素迭代优化

相关函数确定之后,一般通过非线性优化算法对相关函数取极值来获得亚像素精度,目前常用的算法有Bruck等提出的牛顿拉普森(Newton-Raphson, NR)算法(简称正向算法)和计算机视觉领域提出的反向组合高斯牛顿(Inverse compositional Gauss-Newton, IC-GN)算法(简称反向算法)。如图2.1所示。反向算法是由美国卡耐梅隆大学的Baker和Matthews提出,其目的是为了消除Lucas-Kanade算法中的冗余计算,即Hessian矩阵的重复计算。

是不是抄的哪个地方的?一定注意知识产权侵犯

4.6.2 基于特征点的匹配

特征点匹配最大的优点是能够将对整个图像进行的各种分析转化为对图像特征点的分析，从而大大减小了图像处理过程的运算量，对灰度变化、图像变形以及遮挡等都有较好的适应能力。特征点匹配方法一般分为三个过程：特征点提取、特征点描述子计算、特征点匹配三个步骤。

(1) 特征点提取

特征点提取是指在图像中寻找特征点，并获得特征点坐标、尺度和方向等信息的过程。因为相机拍摄距离远近、拍摄角度和光照等的变化，会导致同一目标图像的同一特征点在不同图像中有较大差别，不利于后续特征点匹配。因此图像特征点应具有以下性质：1) 可重复性 (repeatability)，即相同特征点可在不同图像中被精确找到；2) 可区别性

(distinctiveness)，即对不同特征点的描述要不同；3) 高效性 (efficiency)，即特征点的提取效率应该高；4) 局部性 (locality)，即特征点应为图像中的局部区域，以应对相机的拍摄角度变化。

(2) 特征点描述子

特征点描述子是指对特征点邻域信息进行量化数据描述后得到的特征向量，它应该能充分地反映特征点邻域图像的形状和纹理信息。一种良好的特征点描述子应该具备以下性质：1) 可区别性 (distinctiveness)，即不同特征点描述子之间的相似度应该较低；2) 鲁棒性 (robustness)，即相同特征点的描述子能够在图像仿射变换、光照条件变化等干扰下仍具有较高的相似度；3) 高效性 (efficiency)，即特征点描述子应具有较高的构建速度和相似度计算速度。

(3) 特征点匹配

特征点匹配是从两组特征子集中根据相应的特征点描述子找出距离最近的特征点对的过程。根据描述子的类型，在实际应用中可以选择不同的距离度量函数。对于基于浮点型描述子的特征，一般采用欧氏距离进行特征相似性度量；而对于基于二值描述子的特征，一般采用汉明距离进行相似性度量。

目前广泛使用的特征点匹配方法主要有：1) 暴力匹配 (Brute-Force Matcher)。暴力匹配是将待匹配特征子集中的特征点和查询特征子集中的特征点逐一进行距离比较，找出距离最近的查询点的方法。该方法采用穷举的方式进行匹配，速度较慢，精度较高，适用于描述子简短、待匹配对象较少的情况。2) KNN匹配。在匹配的时候选择K个和特征点最相似的点，如果这K个点之间的区别足够大，则选择最相似的那个点作为匹配点，通常选择K=2，也就是最近邻匹配。对每个匹配返回两个最近邻的匹配，如果第一匹配和第二匹配距离比率足够大(向量距离足够远)，则认为这是一个正确的匹配，比率的阈值通常在2左右。将不满足的最近邻的匹配之间距离比率大于设定的阈值(1/1.5)匹配剔除。在所有的机器学习算法中，KNN可能是最简单的。

3) FLANN快速最近邻搜索包 (Fast Library for Approximate Nearest Neighbors) 匹配。该机制可以根据数据本身选取合适的算法来处理数据集，FLANN比其他的最近邻搜索快10倍。FLANN的单应性匹配，单应性指的是两幅图像中的一幅出现投影畸变时，他们还能彼此匹配。

由于图像特征存在局部特性，在一定程度上限制了特征的性能，使得匹配结果中包含许多错误匹配，因此，对匹配结果进行提纯以过滤错误匹配点是非常有必要的。常见的错误匹配分为两种：一种是False-positive matches，将非对应特征点检测为匹配，这类错误匹配可以通过算法进行消除；第二种是 False-negative matches，未将匹配的特征点检测出来，这类匹配结果无法处理，因为匹配算法拒绝。为了消除False-positive matches采用如下两种方式：(1) Cross-match filter：在OpenCV中BFMatcher class已经支持交叉验证；(2) 随机样本一致性方法 (random sample consensus, RANSAC) 提纯法。RANSAC算法的核心思想为：选取4个匹配点对估算两幅图像的变换关系H；在剩余的特征点对中统计满足变换关系H的点对数量并将其称为内点；重复上述两步，具有最多内点的变换关系H可被认为是最终的模型；根据最终的模型过滤错误匹配点对。

图4- 18 提纯前的匹配结果

图4- 19 RANSAC提纯后的匹配结果示例

图4- 17和图4- 18的MATLAB程序如下。

```
clc;clear;close all
%% Step 1: Read Images
sceneImage = imread('GZA Bridge_RGB.jpg');
sceneImage = rgb2gray(sceneImage);
boxImage = imcrop(sceneImage);
figure; imshow(sceneImage); title('Image of a Cluttered Scene');
%% Step 2: Detect Feature Points
boxPoints = detectSURFFeatures(boxImage);
scenePoints = detectSURFFeatures(sceneImage);
% Visualize the strongest feature points found in the reference image.
figure; imshow(boxImage); title('100 Strongest Feature Points from Box Image');
hold on; plot(selectStrongest(boxPoints, 100));
% Visualize the strongest feature points found in the target image.
figure; imshow(sceneImage); title('300 Strongest Feature Points from Scene Image');
hold on; plot(selectStrongest(scenePoints, 300));
%% Step 3: Extract Feature Descriptors
% Extract feature descriptors at the interest points in both images.
[boxFeatures, boxPoints] = extractFeatures(boxImage, boxPoints);
[sceneFeatures, scenePoints] = extractFeatures(sceneImage, scenePoints);
%% Step 4: Find Putative Point Matches
% Match the features using their descriptors.
boxPairs = matchFeatures(boxFeatures, sceneFeatures);
% Display putatively matched features.
```

```

matchedBoxPoints = boxPoints(boxPairs(:, 1), :);
matchedScenePoints = scenePoints(boxPairs(:, 2), :);
figure;
showMatchedFeatures(boxImage, sceneImage, matchedBoxPoints, ...
matchedScenePoints, 'montage');
title('Putatively Matched Points (Including Outliers)');
%% Step 5: Locate the Object in the Scene Using Putative Matches
% "estimateGeometricTransform" calculates the transformation relating the matched points, while eliminating
outliers. This transformation allows us to localize the object in the scene.
[tform, inlierBoxPoints, inlierScenePoints] = ...
estimateGeometricTransform(matchedBoxPoints, matchedScenePoints, 'affine');
% Display the matching point pairs with the outliers removed
figure;
showMatchedFeatures(boxImage, sceneImage, inlierBoxPoints, ...
inlierScenePoints, 'montage');
title('Matched Points (Inliers Only)');
% Get the bounding polygon of the reference image.
boxPolygon = [1, 1;... % top-left
size(boxImage, 2), 1;... % top-right
size(boxImage, 2), size(boxImage, 1);... % bottom-right
1, size(boxImage, 1);... % bottom-left
1, 1]; % top-left again to close the polygon
% Transform the polygon into the coordinate system of the target image. The transformed polygon indicates
the location of the object in the scene.
newBoxPolygon = transformPointsForward(tform, boxPolygon);
% Display the detected object.
figure;
imshow(sceneImage);
hold on;
line(newBoxPolygon(:, 1), newBoxPolygon(:, 2), 'Color', 'y');
title('Detected Box');

```

4.6.3 基于变换域的匹配

基于变换域信息的图像配准方法包括:傅里叶变换、小波变换和Warsh变换等。其中常用的是傅里叶变换图像配准方法,该方法主要有以下一些优点:图像的平移、旋转、仿射等变换在傅里叶变换频域中都能反映出来,同时傅里叶变换域的方法对抵抗噪声具有一定的鲁棒性。傅氏变换由于有成熟的快速算法且易于硬件实现,因而是图像配准中常用的方法之一。但它也有相当的局限性,如傅氏变换方法只能用来配准灰度属性有线性正相关的图像,图像之间必须严格满足定义好的变换关系等。

基于傅氏变换方法的相位相关方法通过相位关系来反映偏移量,而灰度变化在频域中主要影响幅值,对相位影响很小,所以相位相关法具有很高的稳定性。

1975年,相位相关(phase correlation)的概念被应用到图像配准领域中,很好地解决了仅存在平移的图像之间的配准。相位相关方法的主要依据是傅里叶的平移性质。考虑两幅图像和存在的平移,即

(4-22)

对其进行傅氏变换,反映到频域上具有以下形式:

(4-23)

上式说明,两幅具有平移量的图像变换到频域中有相同的幅值,但有一个相位差,而这个相位差与图像间的平移量有直接的关系。根据平移定理,可知以上相位差等于两幅图像的互功率谱的相位,即

(4-24)

式中,“*”表示复共轭,式的右边部分为一个虚指数,对其进行傅里叶逆变换会得到一个冲击函数,其只有在峰值点也就是平移量处不为零,这个位置就是所需求的匹配位置。

习题

- 4-1 角点是什么?斑点是什么?
- 4-2 灰度直方图的横坐标表示?取值范围是?
- 4-3 灰度直方图中灰度值越大,则表明图像亮度()?
- 4-4 灰度直方图中灰度覆盖范围越大,则表明图像对比度()?
- 4-5 常用的角点检测算法有哪些?
- 4-6 斑点检测中DOG算法与LOG算法那个检测效率更高?
- 4-7 SIFT的全称是什么?
- 4-8 简述SIFT特征检测算法的步骤。
- 4-9 简述SURF斑点检测算法的步骤。
- 4-10 常用的图像匹配算法有哪些?

5 基于计算机视觉的位移测量原理

5.1 结构位移监测技术概述

位移作为结构性能评估和健康状况评价的一项重要指标，可以很大程度上反映结构的静动力特性。以桥梁为例，桥梁的变形主要包括主梁竖直方向的挠曲变形、桥墩竖直方向的沉降变形以及斜拉桥悬索桥中拉索的振动变形等。桥梁变形过大将一方面造成桥面上驾驶员和行人的不安全感和不舒适感，另一方面也会影响桥梁结构的安全，因此，对桥梁进行变形监测是至关重要的。

现有位移测量方法主要分为接触式和非接触式。传统的结构位移接触式测量一般会将有线或无线传感器连接在被测结构或构件上，常用的有位移计、连通管、加速度计等。其中位移计需要固定的安装支点，连通管受低温影响，加速度计属于间接测量，累积误差大。对于轻型结构而言，接触式传感器还会为结构带来附加的荷载，从而带来测量误差，同时，从长期应用的角度来讲，有线传感器需要铺设大量的电缆供电，无线传感器需要定期更换电池，因安装和维护的困难均在大型基础设施的应用上受到了一定程度的限制。全站仪是一种典型的非接触式测量设备，一般难以实现高频连续自动测量，不适用于结构动态位移监测。激光测振仪，可以提供高分辨率的传感能力，且对结构而言没有附加荷载，不需要接触测量结构。然而这些测量设备相对昂贵且需进行连续测量，耗费相当长的采集时间。GPS测量受电磁干扰大，精度和采样率较低。微波干涉雷达设备费用较高，容易受到外界水汽变化和结构周围物体反射的影响。

图5- 1 常用结构位移监测技术英文！

(a) 位移计 (b) GPS (c) 激光位移传感器 (d) 微波干涉雷达

基于计算机视觉的结构位移测量技术是一种新型的非接触式光学测量方法，它通过图像传感器采集图像，根据所采集的图像信息，利用计算机进行特征提取，实现对图像变形参数的识别和判断，具有无损、非接触、全场测量、操作简单、测量精度高、便于实现自动化等优点。目前，基于图像处理技术的桥梁动态位移测量技术是一门非常有生命力、有活力的测量技术，已经在科研领域取得很多成果，研究应用前景非常广阔。其相比于传统接触式传感器和其他非接触式传感器具有以下显著的优点：

(1) 非接触式。与接触式传感器相比，基于图像的非接触式测量无需近距离接触结构，且相机可以放置在适合测量的远距离区域，无需耗时耗力地在结构上安装传感器，也不需要结构附近固定的安装点，大大节省测量的时间和成本。对于桥梁检测不需要进行交通控制。此外，每个接触式传感器可以测量一维的位移，但单台相机可以同时测量结构二维位移。

(2) 精度较高。与需要安装在结构上（但无需固定的参考点）的非接触式GPS相比，基于视觉的传感器更精确。GPS的测量误差通常在5-10毫米的范围内，比基于图像的测量方法大一个数量级。

(3) 测量距离远。由于允许的激光功率有限，非接触式激光测振仪必须放置得非常接近测量目标，但是在使用合适的长焦镜头情况下，相机可以放置在几十甚至几百米之外，并且仍然能达到令人满意的测量精度。

(4) 与传统的位移传感器相比，单个相机可以同时跟踪多个点的结构位移。更重要的是在拍摄结构振动的视频后，可以随意改变测量点，这为获得更好的测量结果提供了独特的灵活性。

近十年间，得益于视觉传感硬件与计算机视觉算法的快速发展，基于计算机视觉的结构位移监测技术也在快速成长，作为成熟的技术方法与产品逐渐应用于工程实践。

图5- 2 基于计算机视觉的结构位移监测技术示意图

5.2 基于计算机视觉的结构位移监测基本原理

如图5-2所示，视觉位移测量系统的基本组成包括：相机、镜头、计算与分析终端等。相机与镜头组成了测量位移信号的基本传感元件，拍摄视野内的光以小孔成像的形式通过镜头汇聚并投射于相机的感光元件上，感光元件通过将光学信号转换为电信号后在计算与分析终端进行存储和处理。

图5- 3 基于计算机视觉位移监测的基本流程

如图5- 3所示，基于计算机视觉的位移测量方法一般有以下四个步骤：

(1) 相机标定：相机标定一般需要估计相机的内参数、畸变系数和外参数。内参数和畸变系数受镜头影响，具体方法已在2.4节中讲解。外参数由相机的位置和方向决定，其中相机拍摄角度一般也通过2.4节中的标定方法确定，而随相机摆放位置变化也会带来“像素尺度-物理尺度”比例因子的变化，其标定方法将在7.3节详细展开。

(2) 特征检测/目标识别：运用计算机视觉方法进行位移测量时，一般需要选取用于特征提取的区域或识别的标记点/标志物，从而可以通过重识别这些特征或跟踪这些标记达到获取结构动态位移变化的目的。图像特征检测的主要方法已在第4章中讲解，除此之外也有一些方法中应用黑白靶标或高亮LED等作为标志物附于待测点表面，通过传统的图像处理算法（如二值化分割、形态学操作等）进行检测和定位。

(3) 特征匹配/目标跟踪：对于检测到的特征或识别到的标记，通过逐帧图像进行特征或标记的再次识别和定位，得到各帧图像拍摄时测点的新位置，从而得到其在图像中的运动情况。图像匹配一般分为基于灰度的匹配、基于特征的匹配和基于变换域的匹配，这部分内容已在4.5节中讲解。

(4) 结构位移计算：通过前述步骤得到的是被测结构在图像中的位移变化，即像素位移，之后要根据相机标定计算得到的内参数矩阵、外参数矩阵或单应性矩阵，以及计算得到的比例因子，将被测结构在图像中的位移转换成现实三维世界中的实际物理位移。为了避免帧间计算误差的累计，一般选择后续帧直接与第一帧比较得到该时刻的位移。

需要指出的是，为保证位移测量的精度，可对视频图片进行预处理。对于变形测量最重要的是图像去畸变，如果采用工业相机或者在拍摄视频时将目标放置在图像中间位置，则可忽略畸变的影响。相机的畸变参数属于相机内参的一部分，在第二章已经进行了详细的说明。此外，还可以进行一些图像优化，比如去模糊、去噪、去雾等等，使得目标更清晰，目前有很多成熟的算法可以利用，相关方法在第三章中已做介绍。

基于如图5- 3中计算机视觉位移监测的基本流程，图5- 4给出了计算机视觉位移监测软件模块示意图。

图5- 4 计算机视觉位移监测软件模块示意图

5.3 比例系数确定方法

如图5- 3所示，通过比例系数将图像提取到的像素或亚像素位移换算至具有实际意义的工程位移，或者是像素或亚像素坐标换算成物理坐标，实现尺度层面标定的过程。目前广泛采用的位移换算关系标定方法可以分为以下四种：测距法、平均法、方向向量法以及单应性矩阵法。本节对此展开详细介绍。

5.3.1 测距法

测距法即是指通过测量物点到光心的距离推算放大系数的方法。

图5-5中，物点 $M(x, y, z)$ ，理想像点 $M'(x, y, z)$ ，传感器平面上的像点 $M''(x, y, z)$ ， CO 连线为光轴， f 表示焦距，也就是常说的镜头焦距参数，比如，35mm、50mm、75mm、200mm， u 表示物距， v 表示理想像距， r 为透镜半径， fp 为透镜中心到传感器平面之间的距离，即真实的像距，所谓的调焦就是调整的该参数。

图5-5 凸透镜成像原理图

首先，透镜成像存在以下几何关系：

(5-1)

一般地，物距会远大于焦距，即 $v \gg f$ ，所以， $u \approx f$ ，

(5-2)

比例系数（Scaling Factor, SF）即真实坐标与像素坐标之间的比例因子，为了用图像中目标坐标变化来反应真实桥梁振动位移而计算的转换关系，一种方法可以通过上述相机成像原理确定：

(5-3)

式中， l 为像元尺寸，也就是成像靶面上一个像素的实际尺寸，本课题组采用的工业相机的像元尺寸为 $5.5 \mu m$ 。上述法均可用于相机光轴垂直于物体表面的情况，认为此刻物体表面的上的点具有相同的景深，可以均匀的缩小到图像平面内，因此只需要一个相同的缩放因子。

但实际过程中，相机距离物体较远且不容易保证相机光轴与物体表面的垂直度，同时室外测量时，为了追踪目标，通常情况下会将相机光轴倾斜角度，也就是所谓的斜光轴测量。以竖向位移测量为例，也就需要考虑相机的仰角 α ，这种请款下的转换系数可以通过下式近似：

(5-4)

严格意义上讲，以上均属于简化的位移换算方法，由此会造成误差引入，例如相机角度测量误差、物距测量误差、可变焦镜头的焦距代入计算误差。对于利用面内已知尺寸的平均法，在测量员用鼠标定位已知尺寸端点时，至少有 ± 2 pixels的误差，对于该问题，一般可以采用多次标定取平均值的方式减少随机误差水平。而如果采用严谨的提取算法代替人工勾选，定位误差会减少至 $\pm 1/10$ pixel。采用已知尺寸参考物求解放大系数，以及基于物距和相机角度求解放大系数求解误差会随着相机倾角的增大而增大，随着镜头焦距增大而减小。

5.3.2 平均法

另外一种方法是通过目标物表面上的已知物理尺寸与其对应的像素尺寸的比值确定，可通过靶标面积或具体内部尺寸等计算，公式如下：

(5-5)

式中：表示靶标的实际尺寸，表示靶标在图像中的像素尺寸。相关试验表明，采用已知尺寸参考物求得的放大系数误差在相机角度小于 9° 时在可接受的范围。

图5-6 平均法示意图

对此方法的理论可靠性本文进行了理论推导：

如图5-6所示， L 为相机光心 O 到被测点 P 的距离，假设测点的物长和像长分别为 q 和 q' ，单位均是mm，有：

(5-6)

(5-7)

由此可知，则目标点的理论等效物象比准确值为：

(5-8)

假定已知尺寸的物体实际长度为 $2D$ ，假定已知尺寸的两端点物长分别为 $q+D$ ， $q-D$ ，则对应的像长分别为：

(5-9)

那么平均法的思想即是认为标尺中点位置处的像物比（像长与物长的比值）为：

(5-10)

由此可见，时，也就是测点为已知尺寸的中点时，基于平均法的尺度标定结果与理论值一致。由以上分析，不难看出平均法在斜光轴的实际应用时一般有以下两个局限性或者要求，可以作为工程应用的指导：

1) 尺寸方向和待测点位移方向一致，如桥梁挠度测量需要选用竖桥向的构件，比如栏杆高度，梁高或者灯杆高度等；

2) 已知尺寸越靠近测点，平均法越准确，以图5-6为例，如果测点越靠近已知尺寸的中点，平均法越可靠。如果已有的构件不能满足要求，需要人为地于待测目标附近竖立标尺。

5.3.3 方向向量法

上述方法的前提是已知物距以及相机角度已知。方向向量法的提出源于待测位移的方向与图像水平或者竖向方向不一致的考虑，这种情况下的相机倾斜角度一般也是不可测的。

方向向量法的数学基础是坐标系的旋转变换关系。如图5-7，直角坐标系旋转角度 θ 后，新旧坐标变换公式为：

(5-11)

所以二维坐标系旋转变换矩阵为：

(5-12)

图5-7 旋转矩阵推导示意图

其具体实施过程如图5-8所示。

图5-8 方向向量法示意图

X 、 Y 分别代表的是图像水平位移 u 和竖向位移 v 的方向。

去除本人文献复制比：38.6%(2843)		去除引用文献复制比：38.2%(2811)	文字复制比：38.6%(2843)
1	基于相机扰动校正的桥梁结构变形测量方法与应用 于姗姗(导师：张建) - 《东南大学博士论文》 - 2021-04-01	24.9% (1829)	是否引证：否
2	基于计算机视觉的结构位移监测综述 叶肖伟;董传智; - 《中国公路学报》 - 2019-11-15	13.1% (967)	是否引证：否
3	基于单相机的三维坐标测量及其在结构大变形测量中的应用 胡邹恒(导师：何小元) - 《东南大学硕士论文》 - 2015-04-01	2.7% (197)	是否引证：否

原文内容

4个白色点状目标布置在黑色背景靶面上，白点中心可以通过圆点定位算法得到。向量定义的是待测水平位移的方向，，定义的是待测竖向位移的方向。以方向的位移测量为例：图像位移变换至在方向后的结果为，其中，。接下来对其进行单位换算，之间的实际距离是，则基于平均法其尺度。由此得到：

(5-13)
同理，
(5-14)
其中，，。
(5-15)
其中和就是通过设置的靶标计算的位移换算的方向向量。
从推导过程不难发现，该方法也存在一定的局限性：
(1) 需要人为设置靶标，与非接触式测量的初衷有所违背。且对成像质量要求较高，不适用于远距离测量。
(2) 由于其缩放系数的求解依然是基于平均法的思路，没有考虑相机光轴不垂直于目标靶的影响。
综上所述，该方法的主要贡献只在于对待测位移方向的换算；而对基于相机倾角对斜光轴位移换算关系的校准没有实质性帮助，只适用于正面测量，或者光轴倾斜较小的情况。

5.4.4 单应性矩阵法
对于面内二维变形测量，如果相机姿态不明朗，就需要借助更多的尺度信息建立被测面与像平面之间的映射关系，即单应性矩阵H，例如在被测面设置已知尺寸的标定板，一般情况下将世界坐标系原点定位于标定板左上角的第一个特征点，x、y轴方向分别为标定板水平和竖直方向，则都等于零。那么根据H就可以得到与此共面的待测面的二维全场变形信息了。

图5- 9 单应性矩阵计算原理示意图
基于不考虑畸变的相机成像模型，在标定板上的点和它的像点之间建立了一个单应性映射关系，可以表示为：
(5-16)
其中，H为单应性矩阵，s为一未知比例系数。下面进行单应性矩阵求解，将上式写成等式形式，则有：
(5-17)
将式进行整理可以得到关于的矩阵运算等式：
(5-18)

由上可知，多个对应点的方程叠加起来，可以写成非齐次方程组形式：AX=C。则利用最小二乘法求解该方程组：。单张相片中至少需要4对已知的空间点和像点才能求解出单应性矩阵H，8个方程求解8个未知数。该方法的普遍实现方式是在被测面内放置预先设计好的已知规格的黑白棋盘格，将世界坐标系建立在棋盘格左上角，由于棋盘格间距已知，那么棋盘格角点的世界坐标已知，加之其相应的图像坐标可以通过角点监测获得。通过坐标做差即可得到测点的真实位移，不难看出单应性矩阵方法解决了相机姿态不明朗的问题，不需要测量物距也不需要测量相机各个方向的倾斜角，便可得到全场面内各点的换算关系，且属于理论真实解。

5.4.5 考虑像点移动的测距法
测距法真正应用于结构的变形测量时，遇到的情况还更复杂。一般认为相机的倾斜角度对于获得真实的物理位移有较大的影响性。比如，竖向位移测量就需要对相机的俯仰角进行关注。一般认为角度较小5°时可以忽略。目前考虑相机俯仰角最常用的方式是采用角度余弦值对直接计算到的换算关系进行近似修正。

图5- 10 考虑像点移动的测距法示意图
如图5- 10所示，以挠度测量为例，测点从P点运动到为例。建立斜光轴位移三角关系模型。对于相机仰角为α，采用测距法建立工程位移与图像位移v之间的换算关系：

(5-19)
其中为测点的初始像素坐标，近似为图像中心，v表示像素位移，表示测点到相机光心的距离，为CCD靶面的像元尺寸。其中，和α均为待测量。
斜光轴不同于正轴测量的最主要一点就是，全场的物像比不统一，而且测点运动后的成像位置会也导致物相比发生变化。但已有的测距法采用近似计算（式），这种近似的方式随着相机倾角增大所带来的换算误差也会增大。由此，本研究提出了一种考虑像点位置改变的位移换算公式。仍旧以图5-5为例，测点远离光心，即，测点从P点运动到。建立斜光轴位移三角关系模型。如果不考虑相机夹角，则光测得到的实际位移结果为

(5-20)
我们知道，PA与真实的位移长度之间存在关系：
(5-21)

其中： γ 是测点的实时像素坐标，所以 γ 是具有实时变化性的，也就是考虑了像点移动对放大系数造成的影响。不难看出，相机仰角越大，与PA的相对差距也就越大。由此，本文提出新的基于测距的位移换算公式为：

(5-22)

其中是相机主点，为了简化实施步骤，可以近似用图像中心点的像素坐标代替。

5.4 基于SURF特征匹配的结构位移监测示例

图像匹配方法的分类已在4.5节进行了基本的介绍，这一节将选取一种基于特征匹配的结构位移测量方法结合其MATLAB代码示例逐步进行讲解和分析。

图5-11为采用SURF特征的结构位移视觉测量方法流程图。依据此流程编写代码时，首先要读取视频，获取视频帧数、尺寸、帧率等，并生成各帧的时间戳。由于图像模糊、光线变化等原因，基于特征的图像匹配方法中各帧图像特征点的提取存在不确定性，所以我们通常不设置具体某一特征点进行匹配和跟踪，而是选择检测一个区域内所有的特征点，对成功配对的特征点坐标进行平均，此坐标平均值的变化即被视为该点/区域位移的变化。

因此，在结构位移测量中，我们通常选取目标测点附近区域进行检测和跟踪，在首帧中设置选取首帧图像测点周围合适区域作为后续帧匹配的模板区域，通过逐帧对其附近区域进行特征检测并与首帧特征点匹配，计算目标区域特征点的平均位移变化。通过任意两点间已知的物理尺寸与计算得到像素长度，计算比例换算系数，最终应用于位移的像素位移-物理位移换算。所述步骤均在以下MATLAB代码中体现。需要说明的是，代码以第一帧视频时结构的位置为参考零点，后续帧的位移均为与第一帧位置作差得到的相对值。

图5-11 基于SURF特征的结构位移视觉测量方法流程图

MATLAB示例：基于SURF特征的结构位移时程测量，程序如下。

```
clc; clear; close all
ti
%----- 读取视频 -----%
path='F:\Test_1.avi';
obj = VideoReader(path); % 读取视频
N = obj.NumberOfFrames; % 获取视频总帧数
h = obj.Height; % 获取视频画幅高度（单位：pixel）
w = obj.Width; % 获取视频画幅宽度（单位：pixel）
Fs = obj.FrameRate; % 获取视频帧率
dt=1/Fs;
t=0:dt:(N-1)*dt; % 生成时间序列
%----- 在首帧图像测点周围选取合适区域作为后续帧匹配的模板区域 -----%
frame1=rgb2gray(read(obj, 1)); % 将RGB图像转为灰度图像
[T,rect]=imcrop(frame1); % 在第一帧中定义模板T，和位置信息rect[左上角点横坐标（xmin），左上角点纵坐标（ymin），
col(xWidth), row(yHeight)]
[mt,nt]=size(T);
close all
%----- 尺度标定（此处使用平均法） -----%
imshow(frame1)
[x, y]=input(2); % 选择两点并获取其在图像中的像素坐标
I_known=norm(x(2)-x(1),y(2)-y(1)); % 计算两点间像素距离
D_known= input('输入已知物理尺寸（单位：mm）'); % 输入已知的两点间物理尺寸
SF=D_known/I_known; % 计算比例因子，单位：mm/pixel
%----- 定义后续帧中感兴趣区域（ROI）边界 -----%
% 为了提高检测效率，后续处理时只选取首帧选定区域附近区域做为后续帧检测的ROI
px=round(w/20);
py=round(h/20);
x_min=round(rect(1)-px);
x_max=round(rect(1)+rect(3)+px);
y_min=round(rect(2)-py);
y_max=round(rect(2)+rect(4)+py);
%----- 利用SURF方法检测ROI中的特征 -----%
Disp_P=zeros(N,2); % 初始化像素坐标
Disp_S=zeros(N,2); % 初始化物理坐标
Points_1 = detectSURFFeatures(T); % 检测第1帧ROI特征
[Feature_1, Points_1] = extractFeatures(T, Points_1); % 提取第1帧ROI特征
for i = 1:N
    frame_i=rgb2gray(read(obj, i)); % 将RGB图像转为灰度图像
    ROI_framei=frame_i(y_min:y_max, x_min:x_max,:); % 选取ROI
    Points_i = detectSURFFeatures(ROI_framei); % 检测第i帧ROI特征
    [Feature_i, Points_i] = extractFeatures(ROI_framei, Points_i); % 提取第i帧ROI特征
    indexPairs = matchFeatures(Feature_1, Feature_i); % 将第i帧ROI与模板T进行匹配
    [m,n]=size(indexPairs);
```



```
Coordinate_1=zeros(m,2); % 初始化匹配特征点坐标列表
Coordinate_i=zeros(m,2);
for j = 1:m
Coordinate_1(j,:)=Points_1.Location(indexPairs(j,1)); % 第1帧匹配特征点的坐标列表
Coordinate_i(j,:)=Points_i.Location(indexPairs(j,2)); % 第i帧匹配特征点的坐标列表
end
Disp_P(i,:)=mean(Coordinate_1)-mean(Coordinate_i); % 计算第i帧特征点相对于第1帧特征点的平均位移
Disp_S(i,:)=Disp_P(i,:)*SF; % 转换为物理位移
end
toc
save Displacement.mat Disp_S
%----- 绘制位移时程曲线 -----%
figure(1)
subplot(2,1,1)
plot(t,Disp_S(:,1),'b')
xlabel('Time (s)', 'FontName','Arial','fontsize', 12)
ylabel('Displacement(mm)', 'FontName','Arial','fontsize', 12)
title('Displacement in X direction','FontName','Arial','fontsize', 12)
subplot(2,1,2)
plot(t,Disp_S(:,2),'r')
xlabel('Time (s)', 'FontName','Arial','fontsize', 12)
ylabel('Displacement(mm)', 'FontName','Arial','fontsize', 12)
title('Displacement in Y direction','FontName','Arial','fontsize', 12)
set(gcf,'Position',[200 200 1000 600])
```

5.5 计算机视觉位移监测精度评估试验

(a) 测量系统布置(b) 采集到的目标

目标点

不动参考点

标定板

位移台

图5- 12 精度评估试验

如图5- 12所示，相机测量系统架设在学校体育馆训练场内；目标通过单向电动位移滑台控制，位移台系统由：DM542步进驱动器，步进伺服电机液晶控制器，直流开关电源三部分组成，通过编程可实现控制目标竖向的位移；目标点是预先打印好的圆环非编码标志点。主要进行了不同测距和不同相机俯仰角下的位移测量结果对比。结果如下：

图5- 13 不同测量距离下的位移结果对比和误差分析

表5-1 不同测距下的误差分析结果

测距 (m)	平均误差 (mm)	标准差 (mm)
10	0.0143	0.0028
50	-0.0469	0.0390
100	-0.1225	0.1758
200	-0.1927	0.3517

测距(m) 平均误差 (mm) 标准差 (mm)

10 0.0143 0.0028

50 -0.0469 0.0390

100 -0.1225 0.1758

200 -0.1927 0.3517

图5- 14 不同俯仰角下的位移结果对比和误差分析

表5-2 测距50米下的不同相机俯仰角的位移测量误差分析结果

仰角 (°)	平均误差	误差标准值
10	0.0155	0.0024
20	0.0781	0.0043
35	0.1105	0.0768

仰角(°) 平均误差误差标准值

10 0.0155 0.0024

20 0.0781 0.0043

35 0.1105 0.0768

5.6 位移测量精度影响因素

在采用基于计算机视觉的结构位移测量方法时，其会受到各种因素的影响从而导致测量误差。找出误差来源，并采取针对性的措施对于减小测量误差至关重要。本节将从硬件因素、算法因素和环境因素3个方面进行分析并阐述其解决方案。

5.6.1 硬件因素

由系统硬件导致的误差主要由图像采集装置引起，其中作为采集装置的相机和镜头不能保证图像正常采集进而引发测量结果误差增加。主要有以下表现：

1) 电子噪声导致的图像噪点。相机在进行图像采集时将光信号转化成电信号进行模数转换，并进行图像编码，其间会造成电子噪声。如果通过提高传感器增益进而提高传感器的敏感度，噪声会加大。在现场监测中的各种噪声会降低图像质量，最终引起位移测量结果误差。

2) 相机受热引发的图像变形。相机电子元件工作过程中会发热从而导致相机传感器受到热作用，引发图像变形，从而导致系统误差。在长期的位移监测中，现场气温也会对测量误差造成影响，这种影响不仅有日常波动模式还有累积趋势，利用小波分析可以在一定程度上分离出了温度对测量结果的影响。

3) 镜头畸变导致的图像畸变。三维世界的物体经过镜头投影转换到相机和图像坐标中可能引起畸变误差，虽然通过相机标定可以减小畸变效应，但并不能完全消除，且相机倾角过大时运用尺度因子法带来的误差不可忽略。

5.6.2 算法因素

图像处理算法因素主要是指在采用目标追踪算法进行图像处理过程中发生的误差。主要包括：

1) 亚像素估计。由于追踪引起的像素级位移估计会导致测量误差，通过采取亚像素估计可以减少误差，但是不能消除。亚像素估计方法如前所述，这里不再赘述。

2) 非刚体变形。被测结构或区域在测量过程中会发生非刚体变形，导致追踪目标在边界处的误差。

3) 图像子集处理。图像子集作为图像中被追踪目标的代表，其选取对后续追踪和测量影响误差较大。

5.6.3 环境因素

根据工程测试经验，基于计算机视觉的结构位移监测系统

在室外条件下工作时一般都会在不同程度上受到环境因素的影响，导致测量误差。这其中主要包括风和地面振动造成的相机抖动，雾气造成的视线遮挡，及空气温度场不均匀导致的光线不均匀折射等。

1) 相机自身抖动导致的画面晃动。一般基于计算机视觉的位移测量方法都是基于相机静止的假设，所以相机视角观察到的结构振动可认为是结构的真实振动，这在实验室环境下是成立的。然而室外条件下，相机将不可避免地受到风和地面振动等的影响，所以，在实际测量中一般会在相机视野内同时寻求一静止物体，采用被测物体位移减去静止物体位移消除相机自身抖动带来的影响。但这种方法仅限于相机运动平面与结构振动平面平行，如果发生面外运动，则不准确。也可以通过测量相机的加速度，并用带通滤波进行相机自身抖动的消除，但只有当相机振动的频率带宽与结构振动带宽相近时该方法才有效。目前也有利用深度学习等手段消除相机抖动的方法，并基于此方法使用无人机搭载相机进行了大型结构位移测量。

2) 相机自身抖动导致的图像模糊。相机振动或者物体运动过快会引发图像模糊，从而导致测量误差。在机器人视觉感知等方面，研究人员提出了不同解决图片模糊化的处理手段，其中不乏基于深度学习的方法，但其多是实现目标的识别而非精确的测量。由于工程测量对精度等的要求较高，因图像模糊带来的被测目标拉长、特征点丢失等问题较难解决，结合结构固有的振动频率、形状特性进行该方面的研究或许是很好的解决办法，Wang SG等人利用运动模糊图像的几何矩与运动之间的关系来表示图像中的运动模糊信息，并根据运动模糊线索估计振动参数，基于一幅运动模糊图像和一幅未模糊图像或连续两帧模糊图像来计算低频振动的参数以及高频振动的振幅和方向。该方面具有实用性的研究还相对少见。

3) 雾气遮挡。雾气在测量过程中会对被测目标造成局部遮挡，造成部分特征点或标记物检测不到。通过设置人工光源来提高被测目标的区分度，可以在一定程度上减小雾气带来的误差影响。对于雾气遮挡问题，目前还没有很好的解决办法。

4) 不均匀空气温度场影响。在相机与测量目标之间的空气可能因为局部温度变化产生不均匀空气温度场，造成光线的不均匀折射，进而导致图像质量下降和场景扭曲，影响测量。在长期监测过程中热浪影响不可忽视，有研究学者提出采用热浪滤波的方法降低误差，但目前为止这方面的研究较少。

习题

5-1 基于图像处理技术的桥梁动态位移测量技术相比于传统接触式传感器和其他非接触式传感器具有哪些显著优点？

5-2 视觉位移测量系统的基本组成包括？

5-3 基于计算机视觉的位移测量方法一般有哪四个步骤？

5-4 目前广泛采用的位移换算关系标定方法有哪些？

5-5 影响基于计算机视觉的位移测量精度的因素有哪些？

5-6 影响基于计算机视觉的位移测量精度的环境因素有很多，请举例说明。

5-7 光流法是基于计算机视觉的位移测量常用方法，请查阅相关资料并简述光流的概念。

13. 计算机视觉教材_第13部分		总字符数：9223
相似文献列表		
去除本人文献复制比： 5%(462) 去除引用文献复制比： 5%(462) 文字复制比： 5%(462)		
1	智能拆条应用系统设计和实现 陈宸;张颖风; - 《电视技术》 - 2023-03-15	2.1% (193) 是否引证： 否
2	常用大数据术语中英对照精简版-个人整理.docx - 《互联网文档资源 (https://max.book118.) 》 - 2020	2.0% (186) 是否引证： 否
3	数据分析和数据应用 - 知识科普 - 道客巴巴 - 《互联网文档资源 (https://www.doc88.co) 》 - 2020	2.0% (186) 是否引证： 否
		2.0% (183)

4	<u>基于“RPA+AI”的异常用电检测分析研究与应用</u> 庄友淳;安雅程;朱晓明;焦亚琴;李昌博;甄根;-《电力大数据》-2022-12-21	是否引证: 否
5	<u>基于需求侧的在线监测与信息处理预警系统</u> 苏龙峰;郭越;王滨滨;狄鹏;戴晓娇;唐坚;-《煤气与热力》-2023-04-15	1.8% (167) 是否引证: 否
6	<u>中小学开展人工智能教育初探</u> 易国胜;-《智力》-2022-05-11	1.8% (167) 是否引证: 否
7	<u>人工智能技术用于监舍管理的研究</u> 王晶;吴树芳;杜婉音;朱杰;-《信息与电脑(理论版)》-2012-01-15	1.6% (148) 是否引证: 否
8	<u>人工智能技术用于高校宿舍管理的研究</u> 史亚利;-《湖北函授大学学报》-2012-12-30	1.6% (148) 是否引证: 否
9	<u>智能家居领域专利分析</u> 于兰;刘珊;-《河南科技》-2020-06-25	1.6% (148) 是否引证: 否
10	<u>军事仿真应用与展望</u> 张岩;-《科技信息》-2010-12-15	1.6% (148) 是否引证: 否
11	<u>浅谈物业企业成本管控——以W企业为例</u> 袁春华;-《中国总会计师》-2021-10-15	1.6% (144) 是否引证: 否
12	<u>人工智能领域专利审查策略及检索思路</u> 徐锦超;涂亮梅;-《中国科技信息》-2022-08-29	1.6% (143) 是否引证: 否
13	<u>人工智能与计算智能在物联网方面的应用探究</u> 何昊宸;-《软件》-2022-07-15	1.6% (143) 是否引证: 否
14	<u>大数据时代人工智能在计算机网络技术中的运用</u> 屈原骏;-《信息与电脑(理论版)》-2021-02-25	1.6% (143) 是否引证: 否
15	<u>人工智能指纹识别技术在警务实战中的应用</u> 徐杰;刘哲元;霍鑫;蒋敬;戴玉阳;胡王燕;-《刑事技术》-2021-06-08	1.6% (143) 是否引证: 否
16	<u>人工智能在学校体育的应用与发展趋势</u> 庄弼;张懿;-《新体育》-2021-04-30	1.6% (143) 是否引证: 否
17	人工智能期末试题及答案完整版(最新) - 豆丁网 -《互联网文档资源(http://www.docin.com)》-2019	1.5% (141) 是否引证: 否
18	综述——浅谈模式识别、数据挖掘、机器学习、人工智能及它们与数据融合的联系与区别 - 道客巴巴 -《互联网文档资源(http://www.doc88.com)》-2019	1.5% (140) 是否引证: 否
19	<u>智能化在水泥企业内部控制中的应用与实践——以K集团为例</u> 赵永卓;罗龙;-《中国总会计师》-2021-03-15	1.5% (139) 是否引证: 否
20	<u>人工智能技术在艺术设计中的应用与探索</u> 蔺玮琪;-《大观》-2021-08-10	1.5% (139) 是否引证: 否
21	精品论文解析 利用信息理论观察自然智能和人工智能——相关文章 -《互联网文档资源(http://www.360doc.co)》-2016	1.5% (135) 是否引证: 否
22	<u>基于YOLOv5目标识别无线抓拍追踪声音报警系统</u> 黄恒一;付三丽;-《现代信息科技》-2023-05-10	1.1% (105) 是否引证: 否
23	<u>人工智能背景下应用本科高校计算机技术应用</u> 王琪;梁妍;李钰;-《数字技术与应用》-2021-02-25	1.1% (105) 是否引证: 否
24	<u>人工智能在海洋工程地质领域的应用</u> 孙永福;杜星;宋玉鹏;胡光海;-《海岸工程》-2022-12-15	1.1% (99) 是否引证: 否
25	<u>探析无人机低空遥感技术与人工智能技术融合发展</u> 车彦卓;刘寿宝;-《中国安防》-2021-04-01	1.1% (97) 是否引证: 否
26	<u>人工智能通用大模型教育应用影响探析</u> 吴砥;李环;陈旭;-《开放教育研究》-2023-03-22	1.0% (92) 是否引证: 否
27	<u>基于BP神经网络的混凝土抗压强度预测分析</u>	0.5% (49)

	旷翔文;王芳;邓洁松;付壮金;刘雅婷; - 《安徽建筑》 - 2023-01-20	是否引证: 否
28	基于Docker的人工智能实验平台搭建 王彩玲;张静; - 《信息技术与信息化》 - 2023-01-25	0.3% (29) 是否引证: 否
29	基于BP神经网络的体育赛事转播权运营风险预警模型构建研究 罗恒; - 《体育科技文献通报》 - 2023-05-20	0.3% (29) 是否引证: 否
原文内容		

第6章深度学习与图形处理

深度学习是基于人工神经网络的机器学习和人工智能的一个子集，通过人工神经网络来模拟人脑的工作方式，从而实现对大量数据的学习、分析、预测。深度学习的核心思想是通过多层的神经网络进行学习和训练，从而能够从输入数据中学习到更复杂的特征和模式。深度学习能自动高效地捕捉大数据的特征，从而减少了人工特征工程的需求，但也通常被认为是黑箱系统，算法结构复杂，可解释性弱。深度学习在图像分类、自然语言处理和语音识别等任务中已经取得了突破性的成果，成为人工智能领域的重要研究方向和技术手段。

深度学习是基于人工神经网络发展起来的，因此我们对深度学习的历史梳理需要从神经网络开始。神经网络的源头可以追溯到1943年，当时Walter Pitts和Warren McCulloch创建了一个基于人脑神经网络的计算机模型，使用了一种称为“阈值逻辑”的算法和数学组合来模拟思维过程，可以看作是神经网络模型的雏形。第一个人工神经网络算法是Rosenblatt在1957年发表的感知器（Perceptron），可以自行进行二元分类，这也激发了浅层神经网络研究的革命，持续影响多年。

卷积神经网络的最初概念是由Kunihiko Fukushima提出的，他在1979年开发了一种名为Neocognitron的人工神经网络。Neocognitron的设计灵感来自于人类视觉系统的层级处理方式，采用了分层和多层次的设计，使得计算机能够“学会”识别视觉模式；在训练过程中使用了反复激活强化策略，通过不断调整网络连接的权重来提高对特定特征的敏感性。1986年，Geoffrey Hinton和David Rumelhart等科学家首次提出了反向传播（Backpropagation）算法，用于训练多层神经网络。通过计算网络预测输出与实际输出之间的误差，并将误差沿着网络反向传播，从而更新网络中的权重，以逐步优化网络的性能和准确度，使得网络能够学习复杂的特征和模式。Yann LeCun在1989年提出了LeNet卷积神经网络，并使用反向传播算法进行训练，用于自动识别手写支票图像中的数字。自此之后，反向传播算法成为训练神经网络的核心方法，

深度学习的另一大推动力是硬件和计算力的发展，促使深度学习算法可以更高效地运行和训练，在实际应用中更加可行和有效。由于深度学习任务通常涉及大量的矩阵运算和浮点数计算，传统的中央处理器CPU在执行这些任务时效率较低。NVIDIA公司在1999年推出图形处理单元GPU，GPU具有大量的并行处理单元和专门用于高性能计算的架构，它大大加速了深度学习算法的训练速度，使得在大规模数据集上进行复杂模型的训练成为可能。

深度学习的大浪潮一般认为起于2012年Alexnet，基于卷积神经网络和强大的图形处理计算性能，在图像分类任务中取得突破性的结果，也将科学研究的重点从传统机器学习转移到深度神经网络。深度学习最大的区别在于特征提取的方式，传统方法依赖人工选择图像中检测对象的特征（如颜色、边缘、纹理等），以及微调许多特征参数。深度学习则采用“端到端学习”的概念，给定一组数据集和辅助的注释，通过神经网络自动提取对象最突出的特征[1]。

图6-1传统机器学习与深度学习的对比

6.1人工智能、机器学习与深度学习

人工智能（Artificial intelligence, AI）、机器学习（Machine learning, ML）和深度学习（Deep learning, DL）是三个相关但不同的概念。人工智能是最宽泛的，利用计算机或机器执行通常需要人类智能的任务。机器学习属于人工智能的算法，从数据中学习规律进行预测或决策。深度学习是机器学习算法中最热门的一个分支，它使用多层神经网络从数据中提取更高层次的特征来实现无人工干扰下的决策。简而言之，三者之间的关系可以表示为：人工智能>机器学习>深度学习，如图6-2所示。

图6-2 人工智能、机器学习和深度学习的关系

人工智能是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学[2]。人工智能是计算机科学的一个分支，人工智能技术是通过了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器，该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。我们通常根据智能程度，划分为狭义人工智能（Artificial Narrow Intelligence, ANI）和通用人工智能（Artificial General Intelligence, AGI）。ANI仅能胜任一项特定任务，如下棋的AlphaGo，自然语言处理工具ChatGPT。这与强人工智能AGI形成对比，AGI是一种具有广泛和通用智能的人工智能，具备人类情感、智慧和创造等所有能力，然而复制人类的智慧、情感和在未知情况下的反应能力是一项极其复杂的任务，目前AGI还只能出现在科幻电影中，如电影流浪地球2中的MOSS。

机器学习是**人工智能的一个子集，也是人工智能的核心，它是使计算机具有智能的根本途径**，通过识别数据中的模式进行决策和预测。机器学习包括数据集、模型、算法等必要元素。数据集是训练机器学习程序、找到数据模式和相关性的基础，根据任务不同，数据集可以是图像、文本、数字等，他们将用于训练、验证和测试模型。算法是在数据上执行模式识别、生成模型的计算方法，模型是算法在数据上学习的输出结果。例如对一个简单的曲线拟合任务，模型本质上是输入数据到输出数据的映射函数，其中包含未知的权重参数，算法则是通过优化过程找到模型预测误差最小化对应的权重系数。机器学习的一般流程见图6-3：从各种来源收集数据，进行特征工程的数据分析，选择机器学习算法进行模型训练，再评估算法性能，最后将优化后的模型应用到工程场景中。机器学习性能的影响参数很多，包括数据质量、模型选择、特征选择等。数据量越大、质量越高、多样性越丰富，模型通常会更加精确和可靠，但过多的数据也会增加计算资源消耗、降低计算效率。

图6-3 机器学习任务的一般流程

深度学习是机器学习的一个重要分支，是最近兴起的、重要的且具有变革性和颠覆性的技术。深度学习与传统机器学习方法的显著差异是，它使用多层神经网络来分析数据中复杂的模式。比如一个全连接的神经网络，如图6-4，左侧是输入层，右侧是输出层，中间是一个或多个隐藏层，前后两层的神经元相连，前层神经元的输出就是后层神经元的输入，各层通过一系列非

线性变换来转换输入数据，使网络能够学习输入数据的复杂表示。深度学习直接从数据中提取特征，无需人工干预，但由于模型复杂性，往往依赖强大的计算资源如GPU，训练模型的时间也更久。深度学习技术在计算机视觉、自然语言处理领域中都取得了突破性进展，如自动驾驶汽车借助图像的障碍检测和决策，服务行业兴起的智能客服答疑解惑。

图6-4 多层神经网络

目前，深度学习模型有多种，最常见的有s卷积神经网络（CNN）、循环神经网络（RNN）、生成对抗网络（GAN），见图6-5。CNN是计算机视觉领域目前最主流的网络形式，可用于图像分类、目标检测、图像分割等任务。

图6-5 常见神经网络结构[3]

为了更好地理解深度学习网络，本章第6.2从最简单的感知器模型出发，理解神经网络的基本原理和优化方法。第6.3节介绍了卷积神经网络的原理，第6.4节简要介绍了常见的其他几类网络模型，第6.5节演示了深度学习的实现方式。

6.2理解神经网络

神经网络是一种受到人脑神经网络启发的机器学习技术，旨在模拟人脑神经网络的信息处理和学习能力。人脑中的神经网络是一个高度复杂的生物神经系统，而神经元模型是对人脑中神经元工作原理进行简化的数学模型。大脑的基本计算单元是神经元，神经元通过突触相互连接，形成复杂的神经网络。图6-6展示了生物上神经元的结构形式，每个神经元从树突接收输入信号，然后经过细胞体处理后，在轴突上产生输出信号，轴突最终会分支并通过突触连接到其他神经元的树突上。人类神经系统中约有860亿个神经元，它们相互之间形成了庞大的网络，共同协调和执行各种复杂的认知和生理过程。

神经元模型是对人脑中神经元工作原理的简化数学模型。在神经元模型中，神经元类似于树突，接收来自其他神经元的输入信号。通过加权求和及激活函数的计算，神经元产生输出信号，类似于神经元的轴突传递信息。神经元可看作是一个计算与存储单元，它对输入信号进行计算、结果暂存，并传递到下一层神经元，以进行更复杂的处理。

图6-6 大脑神经元和简化的神经元数学模型[4]

感知器是一种最简单的神经网络模型，由Frank Rosenblatt于1958年提出。感知器由一个单一的神经元组成，具有多个输入和一个输出。感知器的主要功能是根据输入信号进行二元分类，将输入数据分为两个类别。随着神经网络的发展，人们提出了多层感知器结构，通过引入隐藏层，使网络能够学习更复杂的特征和模式，从而实现更广泛的机器学习任务。本节将主要介绍感知器的基本原理和训练方法。

6.2.1感知器

感知器模型是一种简单的线性分类器，只能处理线性可分的数据集。多层感知器是一种更加复杂的神经网络模型，它引入了隐藏层和非线性激活函数，使其能够处理更复杂的非线性任务。

(1) 感知器

感知器(Perceptron)即单层神经网络，是最早提出的神经网络。早期的感知器是一种线性分类模型，而且只针对二分类问题，如图6-7所示，包含输入x、输出y。输入x是一个N维向量，包含多个元素{x1, x2, x3..., xn}；输出值y通常是二进制的，用于将数据分为两类，如预测电子邮件是否为垃圾邮件的任务中，y=1表示是垃圾邮件，y=0表示非垃圾邮件。输入层和输出层通过一个线性组合器和一个二值阈值函数连接。线性组合器包含权重w和偏差b，每个输入xi都有一个对应的权重wi，表示该输入对最终输出的影响程度。输入值与各自权重相乘即加权求和，再叠加上偏差b，即为线性组合器的结果。二值阈值函数是一种特殊的激活函数，它将线性组合器的结果与阈值进行比较，若大于等于阈值，则输出1，否则输出为0。

给定数据和分类标签，感知器中的权重w和偏差b是在训练过程中确定的。权重和偏差首先被初始化为随机值，然后应用于计算输出，输出结果与标注的真实结果进行对比，通过误差最小化，对权重及偏置进行更新。然后，通过不断调整权重和偏差来最小化误差函数，以便更好地拟合数据，最终学习到最优的权重和偏差值，实现数据的精准分类。

图6-7 单层感知器结构

(2) 激活函数

激活函数(activation function)是神经网络中一个非常重要的组件，它在神经元中引入非线性性，使得神经网络可以处理复杂的非线性问题。在神经元中，输入信号经过一系列加权求和后作用于另一个函数再产生相应的输出信号，这个函数就是这里的激活函数。

神经网络中每一层的输入输出都是一个线性求和的过程，下一层的输出承接了上一层输入函数的线性变换，所以如果没有激活函数，那么无论神经网络多么复杂，最后的输出都是输入的线性组合，单纯的线性组合并不能解决复杂的实际问题。因此，激活函数的引入是神经网络能够处理复杂非线性问题的关键所在。

常见的激活函数包括：

a) Sigmoid函数：可以看作平滑的阶跃函数，将输入映射到范围为0~1的数值，用于将任意值转换为概率，并可用于二分类问题。Sigmoid函数的导数在绝对值较大的输入值时接近于零，这导致了梯度逐渐变小，称为“梯度消失”问题。在深度神经网络中，当梯度逐渐变小时，权重的更新也变得非常小，导致网络学习变得非常缓慢甚至停滞，因此Sigmoid函数目前使用较少。

b) Tanh函数：将输入值映射到范围为-1~1的值。它是sigmoid的缩放版本，平滑且可微，梯度比sigmoid更稳定，因此很少引起梯度消失问题。

c) ReLU函数：将输入值映射到0到正无穷之间的输出范围，即将负的输入映射为0，而正的输入保持原数值。ReLU函数简单且计算高效，在深度学习中广泛使用。

d) Leaky ReLU函数：它是对ReLU函数的改进，它在负数区域引入一个小的负斜率，可以解决ReLU函数在负数区域梯度为0的问题。

e) Maxout函数：在每个神经元的输出中选择输入的最大值，通常用于解决多类别分类问题，可以帮助神经网络对不同类别之间的差异进行更好的建模。

f) ELU函数：与ReLU函数类似，在x大于等于0时，它是线性的，保持输入值不变，而在x小于0时，它是指数函数。这使得ELU函数在负数范围内具有一定的平滑性，但它的计算复杂度相对较高。

图6-8 常见的激活函数

不同的激活函数在不同的情况下表现更好，选择合适的激活函数可以对神经网络的性能产生影响。在实际应用中，ReLU函数和其变种通常是首选，因为它们能够在训练过程中有效地防止梯度消失问题，并提高神经网络的训练速度和准确率。

选择适合的激活函数对深度学习网络的性能和训练效果至关重要。ReLU及其变种在图像分类、目标检测、图像分割 and 语义分割等计算机视觉任务中是最常用的激活函数。实际应用中，通常会根据具体任务和网络结构选择合适的激活函数，并结合交叉验证等方法进行调优。在图像分类任务中，常用的激活函数主要是ReLU函数

(3) 多层感知器

多层感知器 (Multi-Layer Perceptron, MLP) 是单层感知器的扩展和改进，通过引入隐藏层和非线性激活函数，使其在处理非线性问题和学习复杂模式方面更具优势。它包含一个输入层、一个或多个隐藏层、一个输出层，如图6-9。输入层接收原始数据作为输入，通常对输入数据进行预处理和归一化。输入层的神经元数量与输入数据的特征数量相等。每个隐藏层由多个神经元组成。隐藏层的神经元数量和层的个数是MLP的超参数，可以根据问题的复杂性和数据集的特点进行调整。输出层产生最终的预测结果，其神经元数量取决于任务的类别数量。例如，在分类问题中，输出层的神经元数量通常等于分类的类别数目；在回归问题中，输出层通常只有一个神经元。

图6-9 多层感知器结构

MLP的层类型是全连接，每个神经元都与前一层的所有神经元连接，并且每个连接都有一个权重。隐藏层和输出层的神经元通过激活函数处理输入，并产生输出。这些激活函数引入非线性性，使得MLP能够学习和表示复杂的非线性关系和模式。通过调整隐藏层的数量和大小以及选择适当的激活函数，MLP可以适应不同类型的数据和问题，并实现更高效和准确的学习。在深度学习中，MLP作为基础模型，还被用作其他深度学习模型的组成部分，如卷积神经网络和循环神经网络。

(4) 感知器应用算例

下面介绍一个感知器用Matlab实现的方式：选取 w 和 b ，使得误分类点到分类超平面 S 的总距离最小，即 $\min L(w, b) = -\sum y_i (w_i + b)$ 。

感知器算法原理：采取随机梯度下降法 (stochastic gradient descent)：一次随机选取一个误分类点来更改 w 和 b 的值。

- (1) 首先选取初值 $w=0, b=0$;
- (2) 在训练集中选取数据 (x_i, y_i) ;
- (3) 如果，则该点是一个误分类点，需要更新 w 和 b 的值：， λ 是学习率;
- (4) 重复步骤 (2) 和 (3) 直至没有误分类点。

代码如下：

```
function [ w, b ] = original_style( training_set, study_rate )
%选取初始值w_0, b_0
w=0;
b=0;
count=0; %每一次正确分类点个数
iteration_count=0; %迭代次数
fprintf(' 迭代次数\t误分类点\t\t权值w\t\t偏置b\t\t\t\t\t'); %输出结果标题
while count ~= size(training_set,2)
count=0;
%在训练集中选取数据 (x_i, y_i)
for i=1:size(training_set,2)
count = count+1;
%如果y_i(w*x_i+b)<=0, 则对w和b进行相应的更新
if training_set(1, i)*(w'*training_set(2:size(training_set,1), i)+b)<=0
w = w + study_rate*training_set(1, i)*training_set(2:size(training_set,1), i);
b = b + study_rate*training_set(1, i);
iteration_count=iteration_count+1;
count=count-1; %不是正确分类点，减一
fprintf(' \t\t\t\t\t', iteration_count); %输出迭代次数
fprintf(' \t\t\t\t\t', i); %输出误分类点
fprintf(' \t\t\t\t\t(%.1lg, %.1lg)'\t', w); %输出w
fprintf(' \t\t\t\t\t(%.1lg, %.1lg)\n', b); %输出b
end
end
end
end
end
```

测试代码如下：

```
training_set=[1, -1, 1; 3, 2, 5; 2, 2, 3];
study_rate=1;
[w, b]=original_style( training_set, study_rate );
```

结果如图6-10所示。

图6-10代码运行结果

6. 2. 2 训练神经网络

训练神经网络是机器学习中至关重要的一步，通过不断调整网络参数和权重，使得神经网络能够从数据中学习到的特征和模式，从而实现更准确的预测和推理任务。在训练过程中，神经网络通过反复进行前向传播、损失计算、反向传播和参数优化，不断修正和改进模型，逐渐调整网络参数，以最小化损失函数的值。

(1) 损失函数

在深度学习中，通过优化算法（如梯度下降）来调整模型的参数和权重，使得损失函数的值逐渐减小，从而提高模型的性能和预测准确性。损失函数（Loss function）是用来度量模型预测结果与实际标签之间的差异的函数。损失函数的选择在训练过程中非常重要，它会影响模型的训练效果和最终的性能。

损失函数可以根据其应用场景和任务类型进行分类。针对分类问题，常见的有0-1损失函数、绝对值损失函数、指数损失函数、Hinge损失函数、交叉熵损失函数；针对回归问题，常见的有平方回归损失、均方差损失、平均绝对误差损失、分位数损失。

在计算机视觉任务中，交叉熵损失函数（Cross-Entropy Loss Function）常用于图像分类任务，来衡量预测结果与真实标签之间的差异，特别适用于多类别分类任务。假设有一个包含N个样本的训练数据集，每个样本有C个类别。交叉熵损失函数的计算公式如下：

其中表示第i个样本的真实标签中第j类的值（0或1），表示模型对第i个样本的预测结果中第j类的概率。交叉熵损失函数的目标是最小化模型预测结果与真实标签之间的差异，使得模型能够更准确地分类样本。在深度学习中，交叉熵损失函数经常与梯度下降等优化算法结合使用，通过反向传播来更新神经网络中的权重和参数，从而使得模型不断优化和学习。

平滑L1损失函数（Smooth L1 Loss Function）常用于目标检测任务中的目标框位置回归，计算预测框和真实框之间的误差。它是均方差损失函数（L2损失）和绝对值损失函数（L1损失）的结合。平滑L1损失函数的表达式如下：

其中，x是预测值与真实值之间的差异（通常是预测框与真实框的位置偏差）。当时，损失函数是一个平方函数，这样在误差较小的情况下，损失函数的变化较平缓，有利于稳定训练过程；当时，损失函数是一个线性函数，这样在误差较大的情况下，损失函数的变化较快，有利于快速调整参数。在一些流行的目标检测算法中，如Faster R-CNN和YOLO，都采用了平滑L1损失函数来训练模型并优化预测结果。

图像分割任务中，常用的损失函数有像素级交叉熵损失函数、Dice损失函数。Dice损失函数是基于Dice系数，它用于度量两个集合之间的相似性。在图像分割任务中，可以将预测的像素分为正类和负类，然后计算Dice系数。Dice系数的计算公式如下：

其中A是预测的像素集合，B是真实标签的像素集合，表示两者交集的像素数量，和分别表示两者的像素数量。Dice损失函数定义为1减去Dice系数，即：

$$\text{Dice loss}(A, B) = 1 - \text{Dice}(A, B).$$

Dice损失函数的取值范围为0到1，其中0表示完全不匹配，1表示完全匹配。因此，最小化Dice损失函数意味着最大化Dice系数，即使预测的像素与真实标签的像素之间具有最大的相似性。Dice损失函数在处理类别不平衡的图像分割任务中表现良好，它能够减少少数类别像素数量较少导致的不平衡问题，并且更加关注像素分类的准确性。因此，Dice损失函数经常用于医学图像分割等需要处理类别不平衡问题的应用场景。

需要注意的是，对于目标检测和图像分割任务，由于模型通常需要同时进行目标分类和位置回归或像素分类，损失函数通常由多个部分组成，例如Faster R-CNN中包括平滑L1损失和交叉熵损失。

(2) 训练神经网络

反向传播（Backpropagation）是训练神经网络的关键算法之一，它通过计算损失函数对神经网络中的每个参数的梯度，从而实现对网络参数的优化。该算法的基本思想是根据损失函数的变化情况，沿着参数空间的负梯度方向对参数进行更新，使得损失函数的值逐渐减小，从而提高模型的预测准确性。反向传播算法分为两个主要步骤：首先，将输入数据通过神经网络进行前向传播，每层的神经元根据输入数据和权重进行计算，并将结果传递到下一层。前向传播的过程将输入数据映射到输出结果；然后，通过使用链式法则计算损失函数对每个参数的梯度，从输出层开始逐层向后传播梯度。一旦得到所有参数的梯度，就可以使用优化算法来更新参数，使得损失函数的值逐渐减小。这个过程使得神经网络能够通过迭代修正和改进模型，逐渐调整网络参数，从而提高模型在训练数据上的性能。

图6-11 反向传播基本原理

在神经网络训练中，优化算法决定了如何根据损失函数的梯度来更新网络参数。梯度下降法（Gradient Descent）是最基本的优化算法。它通过计算损失函数对参数的梯度方向，并按照负梯度方向更新参数，以降低损失函数的值。

14. 计算机视觉教材_第14部分		总字符数：11411
相似文献列表		
去除本人文献复制比：5.5%(626) 去除引用文献复制比：5.5%(626) 文字复制比：5.5%(626)		
1	基于深度卷积神经网络的轻量化表面缺陷检测算法研究 秦杨(导师：邢永康) - 《重庆大学硕士论文》- 2020-06-01	1.4% (158) 是否引证：否
2	基于深度学习的铁路列车轮对轴承智能故障诊断方法研究 张顺捷(导师：秦勇) - 《北京交通大学硕士论文》- 2021-06-01	0.9% (99) 是否引证：否
3	云环境下基于多重概率推理的监控视频异常检测算法研究 许振(导师：吉根林) - 《南京师范大学硕士论文》- 2021-06-30	0.7% (82) 是否引证：否
4	基于门控卷积和SENet的双判别生成对抗网络图像修复模型 傅继彬;曹玉笠; - 《计算机应用》- 2023-06-30	0.5% (55) 是否引证：否
5	基于深度卷积神经网络的ISAR图像小样本识别方法研究 曾磊(导师：白雪茹) - 《西安电子科技大学硕士论文》- 2021-06-01	0.4% (48) 是否引证：否
6	基于磁共振影像的深度学习辅助诊断肝脏局灶性病变的研究	0.4% (47)

	甄世慧(导师: 蔡秀军) - 《浙江大学博士论文》 - 2021-03-01	是否引证: 否
7	基于多尺度三维卷积神经网络的肺结节检测 阎瑞婷;强彦;王梦南;肖宁; - 《电子设计工程》 - 2022-09-05	0.4% (40) 是否引证: 否
8	用于自动驾驶汽车的深度学习技术介绍 李升波;张航; - 《建设科技》 - 2022-01-15	0.3% (35) 是否引证: 否
9	基于卷积神经网络的人脸识别和表情识别研究 董永盛(导师: 王宏) - 《东北大学硕士论文》 - 2018-12-01	0.3% (34) 是否引证: 否
10	基于深度学习的多模态磁共振图像肺部肿块分割方法研究 李家忻(导师: 陈后金) - 《北京交通大学博士论文》 - 2022-09-01	0.3% (33) 是否引证: 否
11	基于视觉识别的柔性机械爪采摘机器人设计 袁天旺;康朋飞;陈祥;董超;王向东; - 《无线互联科技》 - 2023-03-10	0.3% (33) 是否引证: 否
12	基于深度学习的驾驶员疲劳监测系统 刘钰发(导师: 郭显久) - 《大连海洋大学硕士论文》 - 2022-06-10	0.3% (32) 是否引证: 否
13	面向边缘计算的导盲场景理解智能算法研究 王子鹏(导师: 张荣芬) - 《贵州大学硕士论文》 - 2022-05-01	0.3% (31) 是否引证: 否

原文内容

学习率是梯度下降算法的一个重要超参数，它决定了在每次迭代中参数更新的步长大小。学习率越大，每次迭代参数更新的步长就越大，可能导致在损失函数的最小值附近来回震荡，甚至无法收敛；学习率越小，参数更新的步长就越小，可能导致收敛速度很慢，需要更多的迭代次数才能达到最小值。因此，学习率的选择是一个重要的平衡问题，通常，需要通过实验和手动调整来找到一个合适的值，使得梯度下降算法在训练过程中能够快速收敛到最优解。

梯度下降算法中，每次更新参数时，都会使用整个训练数据集计算损失函数关于参数的梯度，然后根据梯度方向来更新参数。这样的方式在大规模数据集上往往会因为计算消耗大而导致训练速度较慢。随机梯度下降 (Stochastic Gradient Descent, SGD) 是对传统梯度下降的一种改进变体，它在每次更新参数时只使用一个小的样本批次，而不是整个数据集。

梯度下降算法中的学习率是一个重要的超参数，它决定了在每次参数更新时参数值应该改变的幅度。然而，手动调整学习率并找到一个合适的学习率对于训练深度神经网络是一个具有挑战性的任务。自适应优化算法，可根据参数的历史梯度来自适应地调整学习率，能更好地适应不同参数的更新需求，提高训练效率。Adam (Adaptive Moment Estimation) 是计算机视觉任务中一种常用的自适应优化算法，它是由Diederik P. Kingma和Jimmy Ba在2015年提出的[5]。Adam结合了动量法和自适应学习率算法的优点，通过计算梯度的一阶矩估计（动量）和二阶矩估计来调整学习率，具有较好的收敛性能和泛化能力。

拟合 (Fitting)、欠拟合 (Underfitting) 和过拟合 (Overfitting) 是描述模型在训练数据上的性能和泛化能力的重要概念。

拟合：模型在训练数据上能够很好地逼近或匹配数据的真实分布或规律。

欠拟合：由于模型过于简单或参数过少，无法拟合数据的复杂性，导致模型在训练数据和测试数据上都表现不好。

过拟合：由于模型过于复杂、参数过多，在训练数据上表现很好，但在未见过的测试数据上表现较差，模型的泛化能力较差，无法适应新的数据。

解决欠拟合问题通常可以适当调整模型复杂度，增加模型的层数和神经元数量，提高模型的学习能力；也可以对数据进行适当的特征处理和选择，提取更有效的特征信息；通过对训练数据进行随机变换，增加数据的多样性，提高模型的泛化能力。

针对过拟合问题，可以使用合适的正则化方法来提高模型的性能和泛化能力。正则化是在模型训练过程中通过添加额外的约束或惩罚来限制模型的复杂度，以防止模型过拟合训练数据。常见的正则化方法有两种：L1正则化 (L1 regularization) 和 L2正则化 (L2 regularization)。L1正则化在损失函数中添加模型参数的L1范数（参数的绝对值之和）作为惩罚项，它倾向于使一些模型参数变为0，从而实现特征选择的效果，减少特征的数量，使模型更简化。L2正则化在损失函数中添加模型参数的L2范数（参数的平方和的开根号）作为惩罚项，它通过对模型参数进行平方惩罚，使得参数较小，模型更稳定。正则化参数是正则化方法中的一个重要超参数，用于控制正则化的强度。正则化参通常以 λ 表示，它是一个非负实数，将其与正则化项相乘添加到损失函数中。较小的 λ 值意味着正则化项的影响较小，而较大的 λ 值则意味着正则化项的影响较大，通常需要进行交叉验证等方法来确定最优的 λ 值。

6.3 卷积神经网络

上节提到的多层感知器，是否可以用来处理图像数据呢？让我们来试一试，图像可以视为M行×N列×3通道的矩阵，如一张3×3像素的黑白图像，需要展平成9列的一维矩阵输入到多层感知器中。现实中图像尺寸往往大得多，意味着接受这一长串数字，每层需要包含大量的神经元和权重参数，且二维图像展平的操作忽略了图像的空间关系，导致预测效果不佳。与多层感知器不同，卷积神经网络 (CNN或ConvNet)，通过一个小尺寸的二维空间滤波器（又称卷积核，kernel），沿图像水平和竖直方向移动，捕获图像中的空间特征，且涉及的参数量显著减少。

图6-12 多层感知器和卷积神经网络的图像处理模式

卷积神经网络是一种用于处理图像和其他二维数据的深度学习模型它在图像分类、目标检测、图像分割等任务中已取得优异的表现。CNN的设计灵感来自于人类视觉系统的工作原理，它可以有效地从图像中提取特征，并学习图像中的复杂模式和结构。CNN主要组件包括卷积层、激活函数、池化层和全连接层。卷积层用于提取图像的局部特征，激活函数引入非线性，池化层用于降低特征图的维度，而全连接层用于将特征映射到输出类别。通过多个卷积层和全连接层的堆叠，CNN可以学习复杂的特征，并实现对图像的高级理解和分析。

图6-13 卷积神经网络结构图

6.3.1 卷积层

卷积层 (Convolutional Layer) 是卷积神经网络的核心组件之一，它负责从输入图像中提取特征。卷积层使用卷积核（也称为滤波器或权重）来执行卷积操作，通过滑动卷积核在输入图像上提取特定的特征。

卷积核有尺寸 (kernel size)、步幅 (stride)、填充行/列数 (padding) 三个参数。卷积核通常是一个正方形或矩形的窗口，尺寸通常较小，如 3×3 或 5×5 。卷积核的尺寸决定了它在输入图像上感受野的大小，也就是它能够捕获的图像信息范围。步幅定义了卷积核在输入图像上滑动的步长，它决定了输出特征图的尺寸。步幅的选择会影响网络的计算量和感受野大小。填充是在输入图像的边缘上添加一些虚拟像素，使得卷积核可以更好地处理图像边缘信息。填充可以有两种方式：零填充和重复填充。零填充在边缘上添加零值像素，而重复填充是将图像的边缘像素复制多次。通过调整这三个参数，可以对卷积操作进行灵活的设计，以适应不同大小的输入图像和不同复杂度的任务。

卷积操作的过程是将卷积核与输入图像的局部区域进行逐元素相乘，并将结果相加得到输出特征图的一个元素。然后，卷积核在输入图像上移动一定的步长，继续执行卷积操作，得到整个输出特征图。如图6.4，对于一个 5×5 的图像数据，卷积核为三阶矩阵，步幅为1，填充为1。首先对图像数据进行补0，即再周围加上一圈0，这不改变数据的特征，但可以保证卷积计算之后输出仍是 5×5 的矩阵。从左到右，再从上到下，取与卷积核同样大小的矩阵进行对应位置相乘再相加，如下图浅蓝色部分演示卷积计算的过程： $0 \times 0 + 0 \times 1 + 0 \times 1 + 0 \times 1 + 1 \times 0 + 2 \times 1 + 0 \times 0 + 6 \times 1 + 7 \times 0 + 0$ (偏置) = 8。所有对应项乘完再相加，实际上就是把所有作用效果叠加起来。

图6-14 卷积运算示例

卷积层的参数共享是CNN的重要特性之一。参数共享意味着一个卷积核在整个输入图像的不同局部区域使用相同的权重进行卷积，从而减少了需要优化的参数数量。参数共享的优势在于它能够在不增加网络复杂度的情况下，增加网络对数据的统计稳定性和泛化能力。在图像处理任务中，由于图像不同位置可能有相似的边缘、纹理，通过参数共享，网络可以学习到这些局部特征的共性，对图像中不同位置的共性特征具有相似的输出值，从而提高了对图像的理解和表达能力。

卷积核的三个超参数，即卷积核尺寸、步幅和零填充行/列数，和输入图片尺寸一起决定了输出特征的尺寸。假设输入图片尺寸为 $a \times a$ ，卷积核尺寸为 $b \times b$ ，卷积步幅为 c ，填充行/列数为 d ，则输出特征尺寸为 $(a - b + d) / c + 1$ 。

在卷积神经网络的前向传播过程中，卷积层的输出首先经过激活函数，然后传递给下一层的卷积层或全连接层进行进一步的特征提取和计算。常用的激活函数是ReLU，以引入非线性，增加网络的表达能力。卷积层之后通常会加入池化层，用于降低特征图的维度，并减少计算量。

6.3.2 池化层

池化层是卷积神经网络中的另一个重要组成部分，用于对特征图进行降采样 (downsampling)。它的主要作用是减少特征图的空间尺寸，从而减少网络参数和计算量，同时保留重要的特征信息，提高网络的鲁棒性和泛化能力。需要注意的是，池化层并不引入额外的学习参数，它只是一种固定的操作，因此不会增加网络的复杂性。同时，池化层通常与卷积层交替使用，卷积层用于提取图像中的特征，而池化层用于降采样和特征的压缩。不同于卷积层里计算输和卷积核的互相关性，池化层直接计算池化覆盖区域内元素的最大值或者平均值。将池化覆盖区域所有值的平均值作为池化结果称为平均池化 (Average pooling)，将池化覆盖区域的最大值作为池化结果称为最大池化 (Max pooling)。

池化从输数据的最左上方开始，按从左往右、从上往下的顺序，依次在输数组上滑动。在池化操作中，池化窗口通常是一个固定大小的矩形区域，比如 2×2 或 3×3 的区域。在每次滑动池化窗口时，根据池化操作的类型（最大池化或平均池化），选择区域内的最大值或平均值作为输出，然后将输出值填充到输出数组的相应位置。这样，输入数据的空间尺寸被压缩，同时保留了重要的特征信息。图6-15给出池化核为 2×2 ，步长为1的最大池化和平均池化结果。

图6-15 池化运算示例

池化层通过减小特征图的空间尺寸，池化层可以大大减少网络中的参数数量，从而降低了过拟合的风险，同时减少了计算量，加速了网络训练和推理过程。同时，池化层在选择区域内的最大值或平均值时，会选择出局部区域中最显著的特征，从而保留了图像中的重要信息，有助于提高网络的鲁棒性和泛化能力。池化操作的输出不受输入图像的平移变换的影响，因为选择的最大值或平均值是局部不变的，这使得网络对图像中的平移具有一定的不变性，提高了图像识别的稳定性。

6.3.3 全连接层

全连接层是卷积神经网络的最后一部分，也称为输出层。它的主要功能是将卷积层和池化层的输出特征映射转换成最终的输出结果，如图像分类的类别预测或目标检测的位置回归。全连接层中每个神经元都与前一层中的所有神经元相连接，参数数量通常较大，可能会导致过拟合问题，因此常常需要结合正则化等技术来优化模型的性能。为了减少参数数量和提高模型的泛化能力，常常在全连接层之前使用一或多个池化层和卷积层进行特征提取和降维。

在图像分类任务中，全连接层通常包含一个或多个神经元，每个神经元对应一个类别，输出的结果是每个类别的得分或概率，然后通过softmax函数将得分转换为概率分布，确定最终的分类结果。在目标检测任务中，全连接层通常用于预测目标的位置和大小信息。全连接层的输出通常是一个向量，包含目标的坐标、宽度和高度等信息。

6.4 其他网络模型

除了卷积神经网络外，在深度学习领域还有其他比较流行的模型，如循环神经网络 (Recurrent Neural Networks, RNN)、生成对抗网络 (Generative Adversarial Networks, GAN)、注意力机制 (Attention Mechanism)。随着深度学习领域的快速发展，新的模型和算法不断涌现，为解决各种问题提供了更多的选择。每个模型都有其独特的特点和应用领域，具体使用哪种模型取决于任务需求和数据特征。由于篇幅所限，本节将简述上述三类模型，同学们想了解更多可查阅相关论文。

6.4.1 循环神经网络

传统的深度神经网络中，输入和输出之间是独立的，只允许数据单向流动，即从输入到输出，见图6-16。循环神经网络 RNN 是一类特殊的神经网络，它通过在网络中使用反馈循环（见图6-16）使信号能够双向传播，先前的输入中提取的特征被反馈到网络中，使其具备记忆能力。这些交互式网络是动态的，因为它们的状态不断变化，直到达到平衡点。RNN 主要用于处理时间序列等顺序自相关数据，它与CNN主要对比见表6-1。

图6.5 RNN与传统前馈神经网络的结构对比

表6-1 CNN和RNN主要特征对比

	CNN	RNN
模型结构	前反馈神经网络，含卷积、池化层等模块	具有循环连接，隐藏状态会被反馈回网络中
输入/输出	输入为图像或其他网格化数据；输出为分类、回归等结果和预测置信度	输入为序列数据，如时间序列、自然语言文本等；输出为完整序列或最后一个时间步的结果
应用场景	自然语言处理（如语言建模、文本生成等）、时间序列预测（如天气预测）等任务	图像分类、目标检测、图像分割等计算机视觉任务

CNN RNN

模型结构前反馈神经网络，含卷积、池化层等模块具有循环连接，隐藏状态会被反馈回网络中

输入/输出输入为图像或其他网格化数据；输出为分类、回归等结果和预测置信度输入为序列数据，如时间序列、自然语言文本等；输出为完整序列或最后一个时间步的结果

应用场景自然语言处理（如语言建模、文本生成等）、时间序列预测（如天气预测）等任务图像分类、目标检测、图像分割等计算机视觉任务

6. 4. 2生成对抗网络

生成对抗网络GAN是由生成器（Generator）和判别器（Discriminator）组成的框架，网络结构见图6-17。生成器接收一个随机噪声向量作为输入，并通过一系列的隐藏层逐步将其转化为与训练数据相似的样本；生成器的目标是生成逼真的样本，以迷惑判别器。判别器接收真实样本（来自训练数据）和生成器生成的样本，并尝试区分它们的真伪。判别器的目标是正确地区分真实样本和生成样本。

GAN的目标是使生成器生成的样本越来越逼真，同时使判别器无法区分真实样本和生成样本。生成器和判别器在训练过程中相互竞争和对抗，直到达到一种平衡状态。损失函数在GAN的训练中起着重要的作用，生成器损失函数用判别器对生成样本的输出与真实标签之间的损失来表示，判别器损失函数为真实样本的判别损失和生成样本的判别损失之和。通过交替训练生成器和判别器，使得生成器的损失函数最大化、判别器的损失函数最小化，进行参数更新，GAN逐渐收敛，并生成与真实样本相似的样本。

图6-17 生成对抗网络的基本结构

生成对抗网络（GAN）在图像领域的应用取得了重要的进展，并在图像生成、图像编辑和重建、超分辨率图像重建、图像风格迁移、图像修复和去噪、图像转换和合成等方面展现出强大的能力。通过训练生成器和判别器的对抗学习，GAN可以生成逼真的图像样本，实现图像的风格转换和合成，如将黑白图像转换为彩色图像、将日间场景转换为夜间场景、将草图转换为真实图像等。这些应用展示了GAN在图像领域的强大潜力和创造力，但GAN生成逼真的虚假图像和内容，可能被滥用于伪造信息、造假证据和传播虚假新闻等行为，造成负面的社会影响，因此信息安全和信任建立是一个人工智能时代下面临的重要挑战。

6. 4. 3注意力机制

注意力机制通过对输入序列中不同部分进行加权处理，增强了网络对重要信息的关注程度。它通常被视为一种神经网络中的组件，而不是独立的模型，可以与其他神经网络结构（如RNN、CNN等）相结合使用，以提高模型在各种任务上的性能。例如，在图像分类任务中，可以使用CNN作为特征提取器，使用注意力机制来加强网络对不同图像区域的关注。

图6-18展示了注意力通用模型，可以分为两个步骤：计算输入信息上的注意力分布(Attention weights)，再计算上下文向量(Context Vector)。让我们以一个简单的图像分类例子来演示它的流程。查询(Query)、键(Keys)和值(Value)是注意力机制中的三个关键组件。“查询”是用于表示当前关注的内容或信息的向量，通常根据任务需求生成，如使用一个向量来表示我们关注是否存在汽车的特征。“键”表示图像中不同区域的特征，每个区域都对应一个键向量，如卷积神经网络的中间层特征。“值”向量是与每个“键”相关联的信息，如图像分类任务中每个区域的特征向量或激活值。通过计算“查询”向量与“键”向量之间的相似性得分，我们可以获得注意力权重，以确定在图像分类过程中关注哪些区域的重要程度。最终，根据注意力权重，我们对“值”向量进行加权聚合，生成加权的特征表示，用于图像分类任务。这样，网络可以更加关注与“查询”相关的区域，提高图像分类的准确性和鲁棒性。

图6-18 注意力机制的通用结构[6]

Transformer模型是一个典型的使用注意力机制的模型，它使用注意力机制替代传统RNN和CNN的模型。它采用了序列到序列（Sequence-to-Sequence）的架构，由编码器和解码器组成，见图6-19。编码器由多个Transformer模块堆叠而成，每个Transformer模块结构如下：

位置编码（Positional Encoding）：为输入序列中的每个位置添加一个表示位置信息的向量。位置编码使得模型能够区分不同位置的单词或标记。

多头自注意力（Multi-head Self-Attention）：通过计算查询、键和值之间的相似性得分，生成注意力权重，然后将注意力权重与值进行加权求和，得到上下文向量表示。多头自注意力允许模型在不同的表示子空间和不同的位置上同时关注输入序列的信息。

前馈神经网络（Feed-Forward Neural Network）：通过两个线性变换和激活函数对特征进行非线性映射。前馈神经网络增加了模型的表示能力。

层归一化（Layer Normalization）：在每个子层的输入和输出之间进行归一化操作，有助于提高模型的训练稳定性。

残差连接（Residual Connection）：将输入直接添加到子层的输出，以便信息能够更容易地传递。

解码器与编码器的结构类似，但还包含一个额外的编码器-解码器注意力层，用于将解码器的每个位置与编码器的输出进行关联。

图6-19 Transformer模型结构

相对于卷积神经网络（CNN），Transformer模型具有一些优势：

(1) 捕捉长距离依赖关系：传统的卷积神经网络在处理序列数据时，由于局部感知野的限制，可能无法有效捕捉到长距离的依赖关系。而Transformer模型通过自注意力机制，可以在任意位置上关注输入序列中的其他位置，从而更好地捕捉长时间依赖关系，适用于处理具有长序列的任务。

(2) 并行计算能力：卷积神经网络的计算是基于局部感受野的滑动窗口操作，导致计算过程中存在序列上的依赖关系，难以进行并行计算。而Transformer模型中的自注意力机制是全局性的操作，允许在输入序列中的所有位置上进行并行计算，大大提高了计算效率。

(3) 捕捉全局语义信息：卷积神经网络在处理序列数据时，局部感受野的限制可能导致无法充分捕捉到全局的语义信息。相比之下，Transformer模型通过自注意力机制能够在编码器和解码器中对全局的输入信息进行关注和整合，有助于更好地捕捉序列中的上下文信息和语义关系。

总之，Transformer模型相对于卷积神经网络在处理序列数据时能够更好地捕捉长距离依赖关系、具有并行计算能力、适应不定长输入，并能够更好地捕捉全局的语义信息。这使得Transformer在自然语言处理和其他序列相关任务中取得了很好的效果。

6.5 深度学习框架

深度学习框架是一种软件工具或软件库，它们提供了一系列的高级功能和工具，帮助开发者更方便地定义、训练和部署深度学习模型。开发者可以通过简单地堆叠层来构建复杂的模型结构，使得模型的设计更加直观和灵活。同时，它也提供了可视化工具和调试器，用于可视化训练过程中的指标变化以及模型调试等；支持将训练好的模型保存为特定格式、部署到生产环境中进行实时预测。深度学习框架也提供了与GPU和其他加速硬件的交互接口和优化，利用硬件的并行计算能力，加速深度学习任务的训练和推理过程，从而提高模型的计算效率和训练速度。

主流的深度学习框架随着时间的推移可能有所变化，TensorFlow和PyTorch是目前业界最为广泛使用的两个深度学习框架，在工业界和学术界都有着广泛的用户基础和影响力。

TensorFlow是由Google开发的深度学习框架，具有完备的解决方案和丰富的生态系统。它提供了丰富的工具和库，可以支持从模型构建、训练到部署的全流程，并且具有高度的可扩展性和跨平台的特性。TensorFlow在工业界被广泛应用于各种场景，包括计算机视觉、自然语言处理、语音识别等。

PyTorch是由Facebook开发的深度学习框架，以其简洁、灵活的接口设计而受到广泛赞誉。PyTorch的设计理念注重于用户友好性和易用性，使得用户可以更快速地设计、调试和迭代网络模型。它支持动态图和静态图的混合计算，具有良好的可读性和调试能力，因此在学术界受到了广泛的欢迎。

TensorFlow和PyTorch均为开源项目，给用户提供了详细的文档和教程，帮助用户快速上手和理解框架的使用方法和原理。同学们可以参考说明文档中进行深度学习框架的安装、数据预处理、模型构建、训练和推理等实践。MATLAB软件也提供了深度学习工具箱（Deep Learning Toolbox），包含了各种深度学习模型和算法，以及用于训练、评估和部署深度学习模型的函数和工具。

下面将对通过两个MATLAB程序案例进行CNN模型构建和训练的演示。首先，从零开始训练模型将帮助你了解网络结构设计、损失函数和优化算法的选择，以及参数的调整。相比于从零开始训练模型，迁移学习通常是一种更常用和有效的选择，特别是在数据稀缺、计算资源有限或目标任务与预训练模型相似的情况下。因此，第二个案例我们将学习如何加载预训练模型并进行微调，以适应新的图像分类任务。

6.5.1 从零开始训练CNN模型

我们将构建一个简单的CNN模型，用于图像分类。

(1) 加载数字图像，并划分为训练和验证数据。

```
digitDatasetPath = fullfile(matlabroot,'toolbox','nnet','nndemos',...,'nndatasets','DigitDataset');
imds = imageDatastore(digitDatasetPath, ...
    'IncludeSubfolders',true,'LabelSource','foldernames');
figure;
perm = randperm(10000,20);
for i = 1:20
    subplot(4,5,i);
    imshow(imds.Files{perm(i)});
end
```

% 划分为训练和验证数据

```
numTrainFiles = 750;
```

```
[imdsTrain,imdsValidation] = splitEachLabel(imds,numTrainFiles,'randomize');
```

图6-20数字样本图像

(2) 定义卷积神经网络架构：包括输入层、卷积层、激活函数层、池化层、全连接层、激活函数层和分类层。

```
layers = [imageInputLayer([28 28 1]) % 输入图像尺寸为 28x28, 通道数为1
convolution2dLayer(3, 32, 'Padding', 'same') % 卷积层, 使用 3x3 的卷积核, 输出通道数为 32
reluLayer() % ReLU 激活函数
maxPooling2dLayer(2, 'Stride', 2) % 最大池化层, 使用 2x2 的池化窗口, 步幅为2
fullyConnectedLayer(10) % 全连接层, 输出大小为 10, 对应于分类的类别数
softmaxLayer() % Softmax 激活函数
classificationLayer() % 分类层
];
```

(3) 指定训练选项：使用随机梯度下降（SGD）作为优化算法，最大训练轮数为10，指定验证数据集用于模型评估。

% 设置训练选项

```
options = trainingOptions('sgdm','MaxEpochs',10,'ValidationData',validationData);
```

(4) 使用训练数据训练神经网络

% 训练模型

```
net = trainNetwork(trainData, layers, options);
(5) 对验证图像进行分类并计算分类准确率。
% 使用训练好的模型对验证数据集进行分类
YPred = classify(net, imdsValidation);
% 获取验证数据集的真实标签
YValidation = imdsValidation.Labels;
% 计算分类的准确率
accuracy = sum(YPred == YValidation)/numel(YValidation)
```

6.5.2 基于预训练模型的迁移学习

预训练模型已经通过大规模数据学习到了复杂的特征表示，包含了对常见对象和模式的抽象能力。通过迁移学习，可以直接使用这些特征提取器，无需从头训练，简化了模型的设计和实现。这里，我们将学习如何加载预训练模型并进行微调，以适应新的图像分类任务。跟6.5.1中的流程类似，主要区别在于第二步中的卷积神经网络定义。

% (1) 加载图像数据集，并划分训练集和验证集

```
imds = imageDatastore('path_to_images', 'LabelSource', 'foldernames', 'IncludeSubfolders', true);
[trainData, validationData] = splitEachLabel(imds, 0.8, 'randomized');
```

% (2) 定义CNN网络

```
net = alexnet; % 加载预训练的AlexNet网络模型
```

```
disp(net.Layers); % 显示网络结构
```

% 设置迁移学习的最后一层

```
numClasses = 10; % 新数据集的类别数量
```

```
layers = net.Layers;
```

```
layers(end-2) = fullyConnectedLayer(numClasses); % 替换原来的全连接层
```

```
layers(end) = classificationLayer; % 替换原来的分类层
```

% (3) 指定训练选项

```
options = trainingOptions('sgdm', 'MaxEpochs', 10, 'ValidationData', validationData);
```

% (4) 使用训练数据训练神经网络

```
netTransfer = trainNetwork(trainData, layers, options);
```

% (5) 对验证集进行分类并计算准确率

```
YPred = classify(netTransfer, validationData);
```

```
YValidation = validationData.Labels;
```

```
accuracy = sum(YPred == YValidation) / numel(YValidation);
```

```
disp(['Accuracy: ', num2str(accuracy)]);
```

6.6 小结

本节涵盖了深度学习的基本概念、常见的模型结构以及实现方法。我们从最简单的感知器模型出发，介绍了神经网络的基本原理和优化方法。接着，着重介绍了适用于图像处理任务的卷积神经网络，它通过卷积层、池化层和全连接层等组件进行图像特征提取和分类。除了卷积神经网络，还简要介绍了其他流行的深度学习模型，如循环神经网络和生成对抗网络。此外，还介绍了常用的深度学习框架，并MATLAB示例程序演示了如何从零开始训练CNN模型和如何利用预训练模型进行迁移学习。

总的来说，本节提供了一个基础的入门介绍，让读者对深度学习的基本概念和方法有所了解。读者可以进一步探索和应用深度学习技术，解决更复杂的问题和任务。

思考题

(1) 深度学习在计算机视觉领域的应用有哪些？请列举几个具体的任务和应用场景。

(2) 深度学习中的反向传播是什么？它在神经网络中起什么作用？

(3) 什么是损失函数？常见的损失函数有哪些，它们适用于不同的任务？

(4) 请解释卷积层和池化层在卷积神经网络中的作用和原理。

参考文献

[15] Niall O' Mahony, Sean Campbell, Anderson Carvalho, Suman Harapanahalli, Gustavo Adolfo Velasco-Hernández, Lenka Krpalkova, Daniel Riordan, Joseph Walsh. Deep Learning vs. Traditional Computer Vision. [J]. CoRR, 2019, abs/1910.13796.

[16] 蔡自兴，徐光祜. 人工智能及其应用. 第6版[M]. 清华大学出版社，2020.

[17] <https://www.asimovinstitute.org/neural-network-zoo/>

[18] <https://cs231n.github.io/neural-networks-1/>

[19] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).

[20] Zhaoyang Niu, Guoqiang Zhong, Hui Yu, A review on the attention mechanism of deep learning, Neurocomputing, Volume 452, 2021, Pages 48-62,

去除本人文献复制比: 44.1%(4042)		去除引用文献复制比: 44.1%(4042)	文字复制比: 44.1%(4042)
1	<u>传统图像分类与深度学习分类算法比较研究</u> 杨涛; - 《荆楚理工学院学报》- 2020-04-25	20.2% (1849)	是否引证: 否
2	<u>基于ISSD的铲车铲齿实时目标检测算法研究</u> 李泽(导师: 魏怡) - 《武汉理工大学硕士论文》- 2020-06-01	7.6% (695)	是否引证: 否
3	<u>基于机器视觉的四旋翼无人机目标检测与定位</u> 路宝国(导师: 吉明; 黄其涛) - 《哈尔滨工程大学硕士论文》- 2020-06-01	7.4% (675)	是否引证: 否
4	<u>基于FPGA的深度学习目标检测的实现</u> 李恒(导师: 单江东) - 《吉林大学硕士论文》- 2020-12-01	6.5% (591)	是否引证: 否
5	<u>基于智能家居摄像头的人脸表情识别应用-记录精彩一刻</u> 张阳阳(导师: 詹炜; 向汉生) - 《长江大学硕士论文》- 2021-06-30	6.4% (586)	是否引证: 否
6	<u>基于深度学习的视觉场景理解算法研究</u> 李一鸣(导师: 徐常胜) - 《合肥工业大学硕士论文》- 2020-05-24	5.5% (505)	是否引证: 否
7	<u>基于深度学习的汽车造型分析与评价</u> 王晨(导师: 李宝军) - 《大连理工大学硕士论文》- 2020-06-30	5.1% (467)	是否引证: 否
8	<u>基于SSD和Inception_resnet_v2网络的目标检测与识别算法研究</u> 谢东阳(导师: 李丽宏) - 《河北工程大学硕士论文》- 2021-06-01	4.2% (389)	是否引证: 否
9	<u>基于SVM图像分类方法的研究 - 道客巴巴</u> - 《互联网文档资源 (http://www.doc88.com)》- 2019	3.3% (304)	是否引证: 否
10	<u>三维重建中纹理映射的研究</u> 张强(导师: 唐璘) - 《中南大学硕士论文》- 2009-05-01	2.7% (248)	是否引证: 否
11	<u>三维重建中纹理映射的研究 - docin.com豆丁网</u> - 《互联网文档资源 (http://www.docin.com)》- 2012	2.7% (248)	是否引证: 否
12	<u>图像空间关系特征描述</u> 程显毅; 李小燕; 任越美; - 《江南大学学报(自然科学版)》- 2007-12-15	2.6% (238)	是否引证: 否
13	<u>图像特征提取与匹配技术研究及其在对象识别中的应用</u> 鲁小锁(导师: 李文辉) - 《吉林大学硕士论文》- 2007-04-26	2.6% (234)	是否引证: 否
14	<u>基于YOLOV3和KCF的高速公路监控视频交通事件检测</u> 郇义浩; 高志权; 张明月; 杨勇; 刘星宇; - 《中国交通信息化》- 2019-06-15	2.4% (218)	是否引证: 否
15	<u>基于内容的医学图像检索应用研究 - 豆丁网</u> - 《互联网文档资源 (http://www.docin.com)》- 2020	2.3% (215)	是否引证: 否
16	<u>图像匹配方法研究与应用</u> 肖若秀(导师: 蔡光程) - 《昆明理工大学硕士论文》- 2008-12-01	2.3% (213)	是否引证: 否
17	<u>可拓方法在图像检索的应用</u> 刘海; 朱小平; - 《现代计算机(专业版)》- 2007-11-15	2.3% (207)	是否引证: 否
18	<u>基于Radon变换和小波理论的静脉特征提取与匹配</u> 顾玲玲(导师: 崔建江) - 《东北大学硕士论文》- 2010-05-01	2.2% (197)	是否引证: 否
19	<u>常用的特征提取与匹配方法_longan</u> - 《网络 (http://blog.sina.com)》- 2018	2.1% (195)	是否引证: 否
20	<u>特征匹配_zlingh</u> - 《网络 (http://blog.sina.com)》- 2018	2.1% (195)	是否引证: 否
21	<u>学习opencv平滑处理-百度文库</u> - 《互联网文档资源 (http://wenku.baidu.c)》- 2012	2.1% (195)	是否引证: 否
22	<u>图像语义自动标注方法的研究</u> 刘钢(导师: 朱文球) - 《湖南工业大学硕士论文》- 2009-05-30	2.1% (190)	是否引证: 否
23	<u>基于层次形状特征提取模型的图像分类</u> 张盛博; 刘娜; 霍宏; 方涛; - 《高技术通讯》- 2016-01-15	1.9% (175)	是否引证: 否
24	<u>图像处理技术在超声相控阵检测中的应用</u>	1.7% (160)	

	侯林利;周小红;高晓蓉;王黎;王泽勇; - 《信息技术》 - 2010-11-25	是否引证: 否
25	图像特征特点及其常用的特征提取与匹配方法 - 豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2017	1.7% (159) 是否引证: 否
26	基于ZY-3影像的北京市六环以内城市用地信息提取及特征分析 王彩艳(导师: 王璦玲) - 《山东农业大学硕士论文》 - 2015-05-20	1.7% (154) 是否引证: 否
27	数字图书馆基于内容的多特征图像检索 赵叶玲;周媛; - 《农业图书情报学刊》 - 2009-03-05	1.3% (123) 是否引证: 否
28	基于度量矩阵特征值的图像空间关系描述 张伟;何金国;桑培罗布; - 《计算机应用研究》 - 2010-11-15	1.0% (90) 是否引证: 否
29	目标检测方法简要综述 栗佩康;袁芳芳;李航涛; - 《科技风》 - 2020-06-28	0.9% (85) 是否引证: 否
30	生物视觉启发的形状特征层次模型及其在目标识别中的应用 张盛博(导师: 方涛) - 《上海交通大学硕士论文》 - 2016-01-01	0.7% (66) 是否引证: 否
31	基于轻型网络的特征融合算法的研究 郑宇隆(导师: 马颖东) - 《内蒙古大学硕士论文》 - 2021-12-01	0.6% (58) 是否引证: 否
32	基于卷积神经网络的武夷岩茶叶片分类方法 林丽惠;魏毅;潘俊虹; - 《宁德师范学院学报(自然科学版)》 - 2021-12-28	0.3% (32) 是否引证: 否

原文内容

第7章目标识别与分割

7.1概述

在现代土木工程中，准确地了解和分析复杂的工程信息是确保项目规划、实施、运维成功的关键。然而，传统的人工方法常常耗时且容易出错。随着计算机视觉和图像处理技术的快速发展，基于目标识别与分割的方法提供了一种优化的解决方案，已经成为了土木工程领域的一个重要研究方向。

图7-1 目标分割与识别应用

目标识别与分割的目的在于通过计算机视觉算法，自动识别和分割土木工程中的目标，如建筑物、构件、设备等。在土木工程中领域得到广泛应用：在工程进度监测领域，目标识别与分割技术可以用于监测土木工程项目的进度。通过分析工程现场的图像或视频，可以自动检测和分割出已完成和未完成的工程部分。这有助于工程管理人员及时了解项目的进展情况，发现潜在的延迟和问题，并采取相应的措施进行调整和协调。在建筑物检测领域，目标识别与分割可以帮助工程师自动识别和检测出土木工程中的建筑物、构件和设备等目标。通过使用计算机视觉技术和图像处理算法，可以对现有的土木工程图纸或实地图像进行分析，准确地标记出建筑物的位置和边界。在结构损伤检测领域，土木工程中的结构损伤可能会导致安全隐患和功能问题。目标识别与分割技术可以用于识别和分割结构中的损伤区域，例如裂缝、腐蚀和破损等。通过对损伤区域进行准确的定位和分析，可以帮助工程师及时采取修复和维护措施，确保结构的安全性和可靠性。在土地利用规划领域，在城市规划和土地利用规划中，目标识别与分割可以用于自动识别和分割土地上的不同类型和用途。例如，可以识别出建筑用地、道路、绿地和水体等。这些信息可以用于评估土地利用状况、规划交通网络和城市基础设施，以及进行环境保护和资源管理等方面的决策。总之，目标识别与分割在土木工程中的应用可以提高工程效率和质量，帮助工程师更好地理解 and 处理复杂的工程信息，并支持决策和规划过程。

本章将深入探讨目标识别与分割的原理、方法和应用，帮助读者理解和掌握这一关键技术，提高土木工程的效率和质量。

7.2目标识别原理

(a) 分类: 桥 (b)检测: 桥墩、桥面 (d)实例分割

图7-2 基于深度学习的三层次分析

基于人工智能和深度学习方法的现代计算机视觉技术在过去10年里取得了显著进展。如今，它被用于图像分类、人脸识别、图像中物体的识别、视频分析和分类以及机器人和自动驾驶车辆的图像处理等应用上。

如何从图像中解析出可供计算机理解的信息，是机器视觉的中心问题。深度学习模型由于其强大的表示能力，加之数据量的积累和计算力的进步，成为机器视觉的热点研究方向。那么，如何理解一张图片，根据后续任务的需要，有三个主要的层次。

一是分类（Classification），即是将图像结构化为某一类别的信息，用事先确定好的类别(string)或实例ID来描述图片。这一任务是最简单、最基础的图像理解任务，也是深度学习模型最先取得突破和实现大规模应用的任务。其中，ImageNet是最权威的评测集，每年的ILSVRC催生了大量的优秀深度网络结构，为其他任务提供了基础。在应用领域，人脸、场景的识别等都可以归为分类任务。

二是检测（Detection）。分类任务关心整体，给出的是整张图片的内容描述，而检测则关注特定的物体目标，要求同时获得这一目标的类别信息和位置信息。相比分类，检测给出的是对图片前景和背景的理解，我们需要从背景中分离出感兴趣的目 标，并确定这一目标的描述（类别和位置），因而，检测模型的输出是一个列表，列表的每一项使用一个数据组给出检出目标的类别和位置（常用矩形检测框的坐标表示）。

三是分割（Segmentation）。分割包括语义分割（semantic segmentation）和实例分割（instance segmentation），前

者是对前背景分离的拓展，要求分离开具有不同语义的图像部分，而后者是检测任务的拓展，要求描述出目标的轮廓（相比检测框更为精细）。分割是对图像的像素级描述，它赋予每个像素类别（实例）意义，适用于理解要求较高的场景，如无人驾驶中对道路和非道路的分割。

7.3 目标分类

7.3.1 传统分类方法

传统的图像分类算法是依靠提取图像特征对图像进行表述，从而区分图像中物体的类别。传统的图像分类识别算法通常由图像预处理、特征提取、分类器设计三个步骤组成。首先图像在分类之前需要进行预处理，因原始图像包含许多干扰、噪声等，故需要对图像进行裁剪，去噪，增强等操作。经过预处理以后的图像可有效减少噪声的干扰，可显著提高图像的分类效果和准确度。其次是特征提取，其是图像分类的核心，特征提取的方式及结果将会直接影响图像分类的精度。但如果是以整张图像作为输入，运算量过大会导致分类效率下降和正确率下降等问题。因此特征提取需要对原始图像先进行降维运算，把原始图像映射到一个低维的特征空间，从而得到可以反映图像本质区别的地位样本特征。最后，将提取到的特征输入至机器学习算法中，训练得到最优的分类模型。因图像的特征提取和分类算法为图像分类的核心，故本节围绕着这两点张开叙述。

图像的特征提取将会影响检测的精度，提取的方式可以分为四大类型，包括由颜色特征、纹理特征、形状特征以及空间关系特征：

(1) 颜色特征是一种全局特征，描述了图像或者图像局部区域所对应的颜色信息。颜色特征中最常使用的颜色空间为RGB (Red Green Blue) 颜色空间和HSV (Hue, Saturation, Value) 颜色空间。描述颜色特征的方法有很多，主要分为颜色直方图、颜色集、颜色矩、颜色聚合向量及颜色相关图。

(2) 纹理特征描述了图像中重复出现的局部模式与其排列规则，反映了图像或图像区域所对应的表面性质。纹理特征的优势在于，无论怎么旋转图像，其纹理特征不会发生改变，具有较强的抗干扰能力。但是当图像的分辨率发生变化时候，纹理特征则可能发生较大的变化。此外，光照、反射也可能影响纹理特征。灰度共生矩阵 (Gray-Level Co-occurrence Matrix) 通过灰度图像得到共生矩阵，然后通过计算共生矩阵来得到矩阵部分的特征值，来分别代表图像的某些纹理特征。常用的基于灰度共生矩阵构建的纹理特征有：能量、熵、对比度、相关性。

(3) 形状特征是用于描述目标的基本特征，常见的有基于轮廓和层次性的形状描述。其中，因为视觉系统通常是基于轮廓区分形状，故基于轮廓的形状描述得到广泛的运用。其中曲率也可作为判别形状间相似性的重要轮廓特征，但是基于轮廓的形状描述也存在一些限制，比如轮廓受噪声影响较大，并且目标轮廓常常难以提取。不同于基于轮廓或区域的描述方法，层次性的形状描述可以逐层地过滤掉不匹配的图像特征，可以弥补单一特征对目标或场景分类的不足。

(4) 图像对象空间关系描述在图像分析中处于重要地位，因为从认知心理角度来讲，人们识别物体时首先关注的是图像各对象的空间关系，然后才是各个对象的颜色、纹理和形状特征。所谓空间关系，是指图像中分割出来的多个目标之间的相互的空间位置或相对方向关系[3]。通常空间位置信息可以分为两类：相对空间位置信息和绝对空间位置信息。前一种关系强调的是目标之间的相对情况，如上下左右关系等；后一种关系强调的是目标之间的距离大小以及方位。显然，由绝对空间位置可推出相对空间位置，但表达相对空间位置信息较简单。空间关系具有尺度、层次、拓扑等几个特征。现有的空间关系表示模型主要针对二维平面的简单对象、确定的现象进行建模，因而大多数空间关系模型只描述和推理二维平面对象和确定现象的空间关系。

图像分类算法可以分为无监督学习和监督学习，监督学习根据已有的数据集，知道输入和输出结果之间的关系，根据这种已知的关系，训练得到一个最优的模型。即给定样本的输入和输出标签，通过训练，分类算法可以找到输入特征与输出标签之间的联系，当给定一个未标注新样本时，分类算法可以利用已经学习到的知识输出一个类别标签。无监督学习主要是根据数据特征之间的内在关系，结构特征的差异将数据分为不同的类别。以下介绍几种常用的分类算法。

(1) K近邻算法，是一种非常简单的机器学习预测、分类算法。当给定一个测试样例时，如果离它最近的K个训练数据大多数属于同一个类别，则认为该测试数据也属于这个类别。如何衡量两个样本之间的距离是该算法的关键之处。常用的距离衡量方式主要有欧式距离、曼哈顿距离、余弦距离等等。其中欧式距离使用的最为广泛。KNN作为一个实用的机器学习分类算法，其优点在于模型简单、通俗易懂、对于简单问题分类准确率较高、训练时间复杂度较低；但缺点也很明显，该算法计算量大、占用较大的储存空间，当特征维数变多时算法效率低下。

(2) 贝叶斯分类算法，是一类利用概率统计知识进行分类的算法，该算法核心思想基于贝叶斯定理，主要用于预测未知类别隶属于各个类别的可能性，选择其中可能性最大的一个类别作为该样本的最终类别。在数据集较大的情况下能够获得较高的准确率。在使用贝叶斯算法时，为了方便建模计算，通常假定给定样本的各个属性之间相互条件独立，这种假定一定程度上会降低算法的分类精度，但是在实际的应用场景当中，能够有效降低算法的复杂性。根据贝叶斯定理，对于一个分类问题，给定具有M个特征维度的样本的情况下，且假定各个特征之间相互独立。朴素贝叶斯分类算法的优点是对于数据缺失不敏感，对待预测的样本，分类过程快速高效；缺点是朴素贝叶斯算法分类时假设特征之间相互独立，当特征之间存在关联性时，会导致分类效果不好。

(3) 支持向量机是一种监督式的学习方法，它是一种二分类的线性模型，目的是寻找一个超平面来对样本进行分割，分割的原则是正负样本之间间隔最大化，最终转化为一个凸二次规划问题来求解。SVM是一种二分类的算法，如何实现用SVM对多个类别的划分，主要有两种不同的思路：分别是一对一方法和一对多的方法，一对多的方法是指在训练时依次把某个样本归为一类，其他剩余的样本归为另一类，这样K个类别的样本就构建了K个SVM。一对一的方法，训练时对任意两类样本之间设计一个SVM，因此K个类别的样本需要设计 $k(k-1)/2$ 个SVM分类器。当对一个未知样本进行分类时，采用投票的方式，最后得票最多的类别即为该未知样本的类别。SVM不需要大量的训练样本，只需少量样本就可达到很好的分类效果。单样本在低维平面类线性不可分时，通过核函数映射到高维平面以后，样本就能够线性可分。但当训练样本过大时候，SVM算法的计算内存和时间消耗巨大，效率变慢。

传统图像分类算法通常基于手工设计的特征提取器和分类器，需要依赖领域专家的经验 and 先验知识。这种方法在处理复杂的图像任务时面临一些挑战。首先，手工设计的特征提取器往往对于图像中的多样性和复杂性不具备良好的适应性，难以捕捉到图像中的高级语义和抽象特征。其次，传统分类器在处理大规模数据时可能遇到维度灾难和计算复杂度的问题，限制了其在大规模数据集上的效果和效率。

7.3.2 基于深度学习的分类方法

近几年来，深度学习算法在图像分类方面取得了突破性进展，通过建立层级特征自动提取模型，得到更准确且接近图像高级语义的特征。深度学习模型可以端到端地学习图像中的特征表达，无需依赖人工设计的特征提取器。这使得深度学习模型能够自动地发现和利用图像中的语义和抽象特征，提高了图像分类的准确性。其次，深度学习模型可以通过增加网络的深度和参数量来适应复杂的图像任务和大规模数据集，同时利用并行计算的优势加速训练和推断过程。这使得基于深度学习的图像分类算法在处理大规模数据时具有较好的可扩展性和效率。基于深度学习的图像分类算法通过自动学习特征表示和端到端的训练，弥补了传统图像分类算法的不足之处。它们在处理复杂的图像任务和大规模数据集时表现出优秀的性能和效率，成为当今图像分类领域的主流方法。分类任务包括以下几个经典的子问题：

(1) 单标记分类。预先设定好的多个标记间彼此是互斥的，图片属于且仅属于其中一个标记。

(2) 单标记精细度分类。该问题是单标记分类的一个特例，特殊在不同类别的实体外观差异非常小，同时由于背景、拍摄角度等因素影响，同类图片外观变化又比较大。这种类间距离小、类内方差大的特性，使细粒度分类十分有挑战性。

(3) 多标记分类。输入图片可以同时被赋予多个标记。

单标签分类中通常采用准确率(Precision)、召回率(Recall)、F值(F-measure)和AUC曲线对分类结果进行评价。然而，在多标签分类中一个图片与多个标签同时关联，其复杂程度远远高于单标签分类。因此，在继承单标签分类评价指标的基础上，许多关于多标签分类的评价指标也被提出。在这里只介绍多标签分类经典的指标，有关单标签分类的指标不再赘述。

平均准确率(AP)和平均准确率均值(mAP)。同单标签分类一样，当一张图片中的所有标记均预测正确时，准确率才可以置1，否则置零。每个类别下的标签分别进行计算后，取其平均值即可获得平均准确率，对所有平均准确率取均值即可获得平均准确率均值。平均准确率可以衡量模型在每个类别的好坏程度，而平均准确率均值则衡量的是在所有类别的好坏程度。覆盖率用来度量“排序好的标签列表”平均需要移动多少步数，才能覆盖真实的相关标签集合。对预测集合Y中的所有标签{y1, y2, ..., yi, ..., yn}进行排序，并返回标签yi在排序表中的排名，排名越高，则相关性越差，反之，相关性越高。

基于深度学习的分类方法通常需要大量的标注数据进行训练。所以图像分类公开数据集是在计算机视觉领域中非常重要的资源。它们为研究人员和从业者提供了标准化的图像数据集，用于评估和比较不同的图像分类算法和模型的性能。这些数据集涵盖了各种不同的图像类别和场景，具有丰富的多样性和挑战性。通过使用这些数据集，研究人员可以进行实验和验证他们的算法在不同情况下的准确性、鲁棒性和泛化能力。

在接下来的章节中，我们将介绍一些常见的图像分类公开数据集，包括它们的特点、数据规模、类别分布等。通过对这些数据集的了解，读者将能够更好地理解图像分类领域的研究和实践，并掌握评估算法性能所需的基本工具。无论是初学者还是专业人士，熟悉这些数据集都将对他们在图像分类任务中的研究和应用产生积极的影响。

Pascal VOC数据集的主要任务是在真实场景中识别来自多个类别的目标。该数据集共有近两万张图片，共有20个类别组成。Pascal VOC官方对每张图片都进行了详细的信息标注，包括类别信息、边界框信息和语义信息，均保存在相应的xml格式文件中。通过读取xml文件中的<name>项，我们可以获取到单张图片中包含的多个物体类别信息，从而构建多标签信息集合并进行分类训练。

COCO(Common Objects in Context)数据集由微软公司赞助搭建。该数据集包含了91个类别，三十余万张图片以及近二百五十万个标签。与Pascal VOC相类似，COCO数据的标注信息均保存在图片对应的json格式文件中。通过读取json文件中的annotation字段，可以获取其中的category_id项，从而获取图片中的类别信息。同一json文件中包含多个category_id项，可以帮助我们构建多标签信息。COCO数据集的类别虽然远远大于Pascal VOC，而且每一类包含的图像更多，这也更有利于特定场景下的特征学习。

除了上述两个主流数据集，常用的还包括ImageNet数据集和NUS-WIDE数据集，以及一些全新的数据集如ML-Images等。这些数据集在计算机视觉领域发挥着重要作用，为算法评估和比较提供了标准化的图像样本集合。

基于深度学习的分类算法在土木工程领域已得到广泛应用，尤其在日常基础设施的病害检测工作中。相比传统的人工目视方法，基于深度学习的分类算法具有诸多优势。它能够处理远距离采集的图像，快速分析结构表面的损伤情况，并通过分类网络对图像进行病害类别标签的打上。基于深度学习的分类网络具有简单功能和快速推理的特点，适用于粗略的检测工作。以混凝土结构中最常见的病害类型裂缝为例，通过一个简单的基于深度学习的分类网络，可以实现快速的裂缝判别。后续提供了分类网络的简易版程序，可以了解其构建流程。下图展示了分类网络的输出效果。

图 7-3 裂缝预测结果

基于深度学习的裂缝分类实例(MATLAB)这句话的格式，全文所有类似的要注意统一，全书可按照这句话的格式

#加载裂缝图像数据集：将数据划分为训练数据集和验证数据集，以使训练集中的每个类别包含 150 个图像，并且验证集包含对应每个标签的其余图像

```
%CrackDatasetPath = 'C:\Users\admin\Desktop\classification-pytorch-main\datasets\train';
imds = imageDatastore(CrackDatasetPath, ...
'IncludeSubfolders',true, ...
'LabelSource','foldernames');
numTrainFiles = 150;
[imdsTrain,imdsValidation] = splitEachLabel(imds,numTrainFiles,'randomize');
#构建一个简单的网络架构：指定网络输入层中图像的大小以及分类层前面的全连接层中类的数量
inputSize = [227 227 3];
numClasses = 2;
layers = [
imageInputLayer(inputSize)
convolution2dLayer(5,20)
batchNormalizationLayer
reluLayer
```

```
fullyConnectedLayer(numClasses)
softmaxLayer
classificationLayer];
#训练上述构建好的网络，options内为设定的训练参数
options = trainingOptions('sgdm', ...
'MaxEpochs',50, ...
'ValidationData',imdsValidation, ...
'ValidationFrequency',30, ...
'Verbose',false, ...
'MiniBatchSize',8,...
'Plots','training-progress');
net = trainNetwork(imdsTrain, layers, options);
#测试上述训练完毕的网络精度
YPred = classify(net,imdsValidation);
YValidation = imdsValidation.Labels;
accuracy = mean(YPred == YValidation)
#使用上述训练完毕的网络对新图像进行预测
I= imread("C:\Users\admin\Desktop\classification-pytorch-main\datasets\test\crack\19001_1.jpg")
inputSize = net.Layers(1).InputSize;
I = imresize(I,inputSize(1:2));
label = classify(net,I);
figure
imshow(I)
title(string(label))
通过使用这些公开数据集和基于深度学习的分类算法，土木工程领域能够实现更准确、高效和客观的病害检测，提升工作效率和质量。
```

7.4目标检测
7.4.1传统目标检测方法

分类算法在图像处理领域取得了显著的成功，但是它们只能告诉我们图像中存在什么物体，而无法提供物体的位置信息。然而，在许多实际应用中，我们不仅需要知道物体的类别，还需要准确地定位和识别图像中的多个目标。为了解决这个问题，目标检测算法应运而生。目标检测算法不仅能够识别图像中的物体类别，还能够精确定位物体的位置，通常用边界框来表示。通过目标检测算法，我们能够更全面地理解和分析图像中的目标信息，为各种应用提供更丰富和精确的数据。目标检测算法的发展不仅扩展了图像处理的应用范围，也为实现自动化和智能化的视觉任务带来了巨大的进步。随着目标检测算法的不断发展，我们可以期待更多创新和突破，为各行业带来更多实用的解决方案。

目标检测(Object detection)在工业界有着广泛的应用，如无人驾驶、安防监控等。目标检测算法按应用场景可细分为静态图像目标检测及视频目标检测，视频目标检测除了利用关键帧的图像信息，还可以利用帧间的时序图像信息辅助检测。其通常可分别两个关键的子任务：目标分类和目标定位。目标分类主要负责输入图像或图像区域中是否有我们感兴趣的物体出现，输出为概率分数，表明我们感兴趣的类别物体出现在图像或图像区域中的可能性。目标定位，通常输出紧致包围在该物体外侧的矩形框(bounding box)四个坐标，显示目标物的图像位置信息。

传统目标检测的目的可分为检测图像中感兴趣目标的位置和对感兴趣目标进行分类。传统的目标检测与识别算法包含三个步骤：区域选择、特征提取、分类回归。首先，对输入的待检测图像利用滑动窗口进行遍历，生成大量的候选区域；然后，采用手工设计的特征提取算法对候选区域的信息进行特征提取；最后将提取到的特征信息送入分类器中进行回归与分类。目标检测比低阶的分类任务复杂，同时也是高阶图像分割的重要基础。

区域选择方法，利用区域划分算法将输入图像分成很多个小的区域，比较相邻区域之间的特征（颜色、纹理、尺寸和空间），将具有相似特征的两个小区域进行合并，直到没有相似区域为止，得到候选区域。

特征提取，Haar 特征用于人脸检测器的使用，利用简单特征分类器增强级联的方法快速计算图像中目标的 Haar 特征，对图像中像素值变化较大的区域具有较好的表达能力，使其能对图像中的边缘信息进行很好的特征表达。不过 Haar 特征针对不同的识别问题需要设计更加适应当前场景的特征，无法适用于多场景的目标检测，并且由于数据量不大，泛化能力弱。

16. 计算机视觉教材_第16部分		总字符数：10181
相似文献列表		
去除本人文献复制比：36.8%(3749)		去除引用文献复制比：36.8%(3749) 文字复制比：36.8%(3749)
1	基于计算机视觉的桥梁车流信息识别与防船撞预警方法研究 张博(导师：张建) - 《东南大学博士论文》- 2021-03-05	11.2% (1139) 是否引证：否
2	医学图像分割方法综述 任楚岚;王宁;张阳; - 《网络安全技术与应用》- 2022-02-15	10.7% (1089) 是否引证：否
3	基于SSD和Inception__resnet__v2网络的目标检测与识别算法研究	4.0% (404)

	谢东阳(导师:李丽宏) - 《河北工程大学硕士论文》 - 2021-06-01	是否引证: 否
4	<u>卷积神经网络下深度特征融合的图像语义分割方法研究</u>	2.5% (253)
	徐慧萍(导师:赵小强) - 《兰州理工大学硕士论文》 - 2021-04-03	是否引证: 否
5	<u>基于YOLOv5的电力巡检鸟窝智能检测系统</u>	2.4% (244)
	游康;谢冠鸿;黄荣;蓝如师; - 《电子技术与软件工程》 - 2022-04-15	是否引证: 否
6	<u>三维复杂环境协同感知与可视化</u>	2.3% (238)
	李聪聪(导师:刘贵喜) - 《西安电子科技大学硕士论文》 - 2020-05-01	是否引证: 否
7	<u>基于YOLOv3网络无人机影像橘子树识别</u>	2.3% (232)
	庞栋栋;刘刚;何敬;付饶; - 《测绘与空间地理信息》 - 2022-04-25	是否引证: 否
8	<u>基于计算机视觉的茶叶嫩芽识别方法研究</u>	2.0% (204)
	张雯娟(导师:毛腾跃) - 《中南民族大学硕士论文》 - 2021-03-01	是否引证: 否
9	<u>基于深度学习的脑肿瘤分割算法研究</u>	1.7% (168)
	林枫茗(导师:刘琚;吴强) - 《山东大学硕士论文》 - 2020-06-15	是否引证: 否
10	<u>图像分割技术纵览</u>	1.4% (138)
	魏弘博,吕振肃,蒋田仔,刘新艳 - 《甘肃科学学报》 - 2004-06-30	是否引证: 否
11	<u>图像背景自适应分割技术研究</u>	1.1% (113)
	王海星; - 《昆明冶金高等专科学校学报》 - 2010-09-15	是否引证: 否
12	<u>目标检测的研究进展</u>	1.1% (109)
	王彩云; - 《中国计算机用户协会网络应用分会2019年第二十三届网络新技术与应用年会论文集》 - 2019-11-07	是否引证: 否
13	<u>基于标签相关性与协同训练的卷积神经网络车辆识别算法</u>	1.0% (101)
	蒋俊钊(导师:程良伦) - 《广东工业大学硕士论文》 - 2018-05-01	是否引证: 否
14	<u>基于改进YOLOv3的道路车辆检测算法研究</u>	1.0% (100)
	马小莉(导师:郭中华) - 《宁夏大学硕士论文》 - 2022-05-01	是否引证: 否
15	<u>彩色图像分割技术研究</u>	1.0% (99)
	王泽兵,杨朝晖 - 《电视技术》 - 2005-04-17	是否引证: 否
16	<u>基于yolov3的梨实时分类抓取方法</u>	1.0% (97)
	陆泽楠;孙松丽;朱文旭; - 《价值工程》 - 2020-05-18	是否引证: 否
17	<u>基于yolov3模型的实时行人检测改进算法 网络首发 - 道客巴巴</u>	0.9% (94)
	- 《互联网文档资源 (https://www.doc88.co)》 - 2020	是否引证: 否
18	<u>基于Mask R-CNN的水面漂浮物识别方法研究</u>	0.8% (81)
	刘伟;王源楠;江山;马天;肖维; - 《人民长江》 - 2021-11-28	是否引证: 否
19	<u>论YOLO算法在机器视觉中应用原理</u>	0.8% (81)
	李茂晖;吴传平;鲍艳;房卓群; - 《教育现代化》 - 2018-10-08	是否引证: 否
20	<u>基于深度特征知识的光伏电池多尺度缺陷检测研究</u>	0.8% (81)
	张泽智(导师:陈海永) - 《河北工业大学硕士论文》 - 2021-05-01	是否引证: 否
21	<u>基于深度学习的车辆检测系统设计</u>	0.8% (81)
	黎雄;陈雨聪; - 《软件工程》 - 2022-07-05	是否引证: 否
22	<u>融合自注意力机制的安全帽佩戴检测方法</u>	0.7% (76)
	孙国栋;李超;张航; - 《计算机工程与应用》 - 2021-06-22 09:50	是否引证: 否
23	<u>图像语义分割模型及应用研究</u>	0.7% (75)
	谢晨伟(导师:吴建鑫) - 《南京大学硕士论文》 - 2018-05-31	是否引证: 否
24	<u>基于神经网络的身份证文本信息提取设计与实现</u>	0.7% (70)
	唐月标(导师:张江鑫;华惊宇) - 《浙江工业大学硕士论文》 - 2020-01-01	是否引证: 否
25	<u>腰椎神经根核磁共振图像检测算法研究</u>	0.7% (69)
	吴伯雄(导师:刘洪波) - 《大连海事大学硕士论文》 - 2020-06-01	是否引证: 否
26	<u>融合跳跃连接的全卷积汉字笔画分割算法</u>	0.6% (60)
	吴文益;李镇宇; - 《大连民族大学学报》 - 2022-09-15	是否引证: 否
27		0.6% (59)

	彩色图像的分割技术	席砺莼, 闫宏伟 - 《微机发展》- 2003-04-10	是否引证: 否
28	彩色图像分割技术综述	李永军; - 《科技情报开发与经济》- 2008-04-05	0.6% (59)
29	基于YOLOv2的视频火焰检测方法	杜晨锡; 严云洋; 刘以安; 高尚兵; - 《计算机科学》- 2019-06-15	0.4% (37)
30	基于多源遥感数据的偏远散区公安敏感目标检测	李非燕(导师: 霍宏涛) - 《中国人民公安大学博士论文》- 2020-08-07	0.3% (34)

原文内容

SIFT将图像中的关键点（角点、边缘点等）作为特征点，确定关键点的特征方向，通过比较匹配特征点，建立对应关系。SIFT 特征有较好的尺度不变性，能减少噪声、角度变化等对目标的干扰，具有一定的鲁棒性。能在海量数据的特征库中进行高速匹配，即使单个物体也能产生大量的特征向量，检测速度快，广泛应用于特征匹配中。HOG 特征用于行人检测问题并取得优秀的效果，通过计算局部区域的梯度直方图作为特征。HOG 特征将图像划分为较小的区域，统计各个区域内像素点的梯度直方图，形成图像的特征描述。利用直方图存储目标的边缘信息，具有更强的表达能力，对几何和光照变化有较好的不变性，在行人检测任务中表现优越。

分类回归，目前常用的分类器包括：支持向量机(SVM)、SoftMax 分类器、AdaBoost(Adaptive Boosting)、朴素贝叶斯分类器(NaiveBayes)、随机森林 RF(RandomForests)等。这一步骤和图像分类任务相似，是根据选定的特征和分类器，检测出有限的几种类别。

7.4.2基于深度学习的目标检测方法

传统的目标检测虽然相对成熟，但也有其固有的缺点。首先，基于滑动窗口的区域选择策略具有高计算复杂度和高窗口冗余度。其次，外观的形态多样性、光照变化的多样性以及背景的多样性使得手动设计特征变得困难。然而，深度卷积神经网络可以从低级到高级学习图像的特征，非常稳健。因此，研究人员逐渐将注意力转向深度学习的目标检测方法。

2021年至今，随着深度学习算法的快速发展，基于深度学习的目标检测算法逐渐成为主流。Pascal VOC[21]，ILSVRC[22]，MS-COCO[23]，和OID[24] 数据集是目标检测使用最多的四大公共数据集。

根据算法原理，目前基于深度学习的目标检测算法大致可以分为两类：

两阶段目标检测。基于Region Proposal的R-CNN系算法（R-CNN，Fast R-CNN，Faster R-CNN），它们是two-stage的，需要先使用启发式方法（selective search）或者CNN网络（RPN）产生Region Proposal，然后再在Region Proposal上做分类与回归。

单阶段目标检测。Yolo，SSD这类one-stage算法，其仅仅使用一个CNN网络直接预测不同目标的类别与位置。

前者是先由算法生成一系列作为样本的候选框，再通过卷积神经网络进行样本分类；后者则不用产生候选框，直接将目标边框定位的问题转化为回归问题处理。前者在检测准确率和定位精度上占优，后者在算法速度上占优。

目标检测中的评价指标mAP（mean average precision）：多个类别物体检测中，每一个类别都可以根据recall（TP/P，x轴）和precision（TP/P，y轴）绘制一条曲线，AP就是该曲线下的面积，mAP是多个类别AP的平均值。

基于深度学习的目标算法在土木工程的健康检测监测领域同样得到广泛应用。虽然基于深度学习的分类网络能够快速地判别结构是否存在病害，但是无法定位病害在结构中的位置。相比于分类网络，目标检测网络通过使用标记框在病害定位与可视化方面做了进一步的提升。同样，这里以混凝土结构中最常见的病害类型裂缝为检测目标，通过一个简单的基于深度学习的目标检测网络，实现裂缝的定位检测工作，目标检测网络的输出效果如下图所示。

图 7-4 裂缝检测结果

基于深度学习的裂缝目标检测实例（MATLAB）

#加载数据集：此示例使用包含 679个图像的裂缝数据集

#裂缝数据存储在一个包含两列的表中，其中第一列包含图像文件路径，第二列包含车辆边界框

```
data = load('crack669.mat');
```

```
T = data.gTruth696;
```

```
T(1:4,:)
```

#将数据集分成训练集、验证集和测试集。选择 60% 的数据用于训练，10% 用于验证，其余用于测试经过训练的检测器

```
rng(0);
```

```
shuffledIndices = randperm(height(T));
```

```
idx = floor(0.6 * length(shuffledIndices));
```

```
trainingIdx = 1:idx;
```

```
trainingDataTbl = T(shuffledIndices(trainingIdx),:);
```

```
... ..
```

```
trainingData = combine(imdsTrain,bldsTrain);
```

```
validationData = combine(imdsValidation,bldsValidation);
```

```
testData = combine(imdsTest,bldsTest);
```

#构建YOLO v2目标检测网络

```
inputSize = [224 224 3];
```

```
numClasses = width(T)-1;
```

```

trainingDataForEstimation = transform(trainingData,@(data)preprocessData(data,inputSize));
numAnchors = 7;
[anchorBoxes, meanIoU] = estimateAnchorBoxes(trainingDataForEstimation, numAnchors)
featureExtractionNetwork = resnet50;
featureLayer = 'activation_40_relu';
lgraph = yolov2Layers(inputSize,numClasses,anchorBoxes,featureExtractionNetwork,featureLayer);
#训练构建好的检测网络
augmentedTrainingData = transform(trainingData,@augmentData);
preprocessedTrainingData= transform(augmentedTrainingData,@(data)preprocessData(data,inputSize));
preprocessedValidationData = transform(validationData,@(data)preprocessData(data,inputSize));
options = trainingOptions('sgdm',...
'MiniBatchSize',4,...
'InitialLearnRate',1e-4,...
'MaxEpochs',100,...
'CheckpointPath',tempdir,...
'ValidationData',preprocessedValidationData,...
'Plots','training-progress');
[detector,info] = trainYOLOv2ObjectDetector(preprocessedTrainingData,lgraph,options);
#测试训练好的检测网络
preprocessedTestData = transform(testData,@(data)preprocessData(data,inputSize));
detectionResults = detect(detector, preprocessedTestData);
[ap,recall,precision] = evaluateDetectionPrecision(detectionResults, preprocessedTestData);
#使用训练好的检测网络对新图像进行检测
I = imread('C:\Users\admin\Desktop\demomo\New\.....jpg');
I = imresize(I,inputSize(1:2));
[bboxes,scores]= detect(detector,I);
I = insertObjectAnnotation(I,'rectangle',bboxes,scores);
figure
imshow(I)

```

7.5 图像分割

7.5.1 传统分割方法

对于一些应用场景，仅仅识别目标的存在并不足以满足需求，还需要准确地分割出目标的精确边界和区域。分割算法旨在将图像中的每个像素分配给特定的目标类别或背景。图像分割是计算机视觉中的一个关键过程。它包括将视觉输入分割成片段以简化图像分析。片段表示目标或目标的一部分，并由像素集或超像素组成。图像分割将像素组织成更大的部分，消除了将单个像素作为观察单位的需要。图像分割是进行目标检测和分类的基础。在分割过程本身，有两个粒度级别：

a) 语义分割：将图像中的所有像素划分为有意义的对象类。这些类是“语义上可解释的”，并对应于现实世界的类别。例如，你可以将与猫相关的所有像素分离出来，并将它们涂成绿色。这也被称为dense预测，因为它预测了每个像素的含义。

b) 实例分割：标识图像中每个对象的每个实例。它与语义分割的不同之处在于它不是对每个像素进行分类。如果一幅图像中有三辆车，语义分割将所有的车分类为一个实例，而实例分割则识别每一辆车。

传统图像分割是指将一幅图像分成若干互不重叠的子区域，使得每个子区域具有一定的相似性、而不同子区域有较为明显的差异。图像分割也可以被认为是在目标检测的基础上，对所检测到的物体和背景进行像素级分割。因此图像分割的大多数步骤和方法与目标检测类似，区别在于所使用的分割方法，下面介绍几种经典的分割算法。

(1) 阈值法

阈值法具有计算量小、实现简单、稳定性较好等优点，成为图像分割领域中运用最广泛的分割技术。其基本原理是通过设定阈值的不同，把像素点分为不同灰度级的背景区和目标区，所以它更适用于背景和背景灰度相差较大的图像。如果图像只有目标和背景两大类，就只需要设置一个阈值。当像素点所对应的函数值大于阈值，则该点属于目标，反之则不属于。如果图像中有多个目标需要分割，就要设置多个阈值，所以其分割的准确度多依赖于阈值的选择，常用的阈值选择方法有运用图像灰度直方图的峰谷法、基于过渡区法、结合连通信息的阈值方法、最大熵原则自动阈值法和最大相关性原则选择阈值。这种方法就是运用像素点灰度值的特征，并没有考虑其空间特性，此方法对噪声也比较敏感，但是目前出现了许多基于阈值法的改进算法，分割准确度有了进一步提高。

(2) 区域生长法

区域生长法是根据像素的相似性质来聚集像素点所形成的区域，这种方法是从单个像素出发，逐步合并形成所需要的区域。区域生长是从待分割区域中选取种子像素开始，作为生长起点，将种子像素邻域里符合条件的像素合并到种子像素里，形成新的种子像素，继续合并，直到找不到符合条件的像素，其关键是种子像素的选择、确定生长规则和终止条件。区域生长法计算简单，对分割具有相同特征的连通区域效果较好，但是由于噪声和灰度不均的问题，容易产生空洞和过分割。

(3) 边缘检测法

基于边缘检测的图像分割是通过检测包含不同区域的边缘来解决问题，不同区域之间边界的灰度值变化比较剧烈，一般利用一阶导数的极大值或二阶导数的过零点信息来判断边缘点。在图像领域中，最简单常用的检测方法包括：Roberts 梯度算子、Wills 算子、Prewitt 算子、Canny 算子等。不同的算子有不同的提取效果，但两大难点限制了其在图像分割中的应用，一个是不能保证边缘的连续性和封闭性；另一个是在高细节区容易出现大量碎边缘，难以形成一个大区域。

(4) 聚类法

将图像中相似灰度合并的方法称为聚类，通过聚类将图像表示为不同的区域即聚类分割方法。由于多数图像或多或少具有模糊性、图像质量低等特性，所以在分割领域最常用的聚类算法是模糊 C-均值算法（FCM），通过“隶属度”来确定像素点属于某个聚类的程度。模糊C均值聚类的基本思想是：先通过人工随机指定每个数据到各个聚类（簇）的隶属度，再根据隶属度计算每一个簇的质心，接着重新进行伪划分，直到质心不变。该算法是一种无监督算法，在一定程度上缓解了医学图像分割标签少的问题。同时聚类算可以结合其他的图像处理方法以提高分割的效果，比方说粒子群算法和遗传算法相结合的优化算法来确定初始类聚中心，再引入像素的邻域信息，克服噪声对异常值敏感的问题。

7.5.2 基于深度学习的分割方法

随着深度学习技术的快速发展，传统的图像分割算法正在逐渐过渡到基于深度学习的分割算法。传统分割算法主要基于手工设计的特征和规则，需要人工提取和选择特征，并且对于复杂的图像场景和变化较大的目标，其性能往往不理想。而基于深度学习的分割算法能够通过深度神经网络自动学习和提取图像的高级特征，从而实现更准确和鲁棒的分割结果。

基于深度学习的分割算法通常采用卷积神经网络（CNN）或全卷积神经网络（FCN）等架构。这些网络能够将图像作为输入，并通过多层卷积和池化操作来学习图像的特征表示。然后，通过上采样或反卷积等操作将特征图映射回原始图像尺寸，并生成像素级别的分割结果。通过大规模标注数据的训练，深度学习模型能够学习到更丰富的特征表示，从而提高分割算法的性能和泛化能力。

基于深度学习的分割算法在土木工程领域具有广泛的应用。例如，在基础设施检测和维护中，通过深度学习的分割算法可以更准确地分割和识别道路裂缝、桥梁损伤等结构的关键区域。这可以帮助工程师及时发现和处理问题，提高基础设施的安全性和可靠性。此外，在地质勘探和地质灾害预测中，基于深度学习的分割算法可以准确地分割和提取地质体、地层和地形等重要信息，为地质研究和灾害管理提供有力支持。

总而言之，基于深度学习的分割算法相较于传统方法具有更强的学习能力和表达能力，能够有效应对复杂的图像场景和变化多样的目标。在土木工程领域，基于深度学习的分割算法为工程师提供了一种强大的工具，可以实现精确的目标分割和区域提取，为工程决策和规划提供更可靠的信息支持。

常用的语义分割公开数据集有PASCAL VOC、Microsoft COCO、Cityscapes等。语义分割常用评价指标有：

- (1) 像素精度（pixel accuracy）：每一类像素正确分类的个数/ 每一类像素的实际个数
- (2) 均像素精度（mean pixel accuracy）：每一类像素的精度值的平均值
- (3) 平均交并比（Mean Intersection over Union）：MIoU是指求出每一类的IOU取平均值。IOU指的是两块区域相交的部分/两个部分的并集，即交集除以并集。
- (4) 权重交并比（Frequency Weight Intersection over Union）：每一类出现的频率作为权重

从本质上讲，语义分割任务是一个像素级的分类任务，但是由于任务本身的特点，在分类的同时，还要考虑其他因素，比如分割模型输出的标记个数要与原图的像素个数一致。基于深度学习的图像语义分割模型存在以下两个问题：图像中的像素数量巨大。比如一张640x480像素的图片共有约3.1万像素。因此一般基于卷积神经网络的图像语义分割模型，都会对图像或者特征图进行下采样。但这又带来另一个问题，经过下采样所得图像、特征图，由于分辨率降低而丢失很多局部细节，即使后期放大到原图分辨率，结果已经十分“模糊”。

像素点的分类彼此之间不是独立的。图片中任意一个像素，其本身并不具备足够的“语义信息”，将单独一个像素拿出来，很难甚至无法确定该像素是否来自同一类别。

基于卷积神经网络的图像分割方法，首先，通过卷积神经网络从训练样本中提取不同程度的图像特征；其次，从训练样本中学习并获取图像内容中的语义信息，构建学习样本与分割目标之间的映射关系；最后，实现图像语义分割。基于CNN的图像语义分割方法类型“日新月异”，这使得分割任务的关注点也随之丰富，依据关注点的不同可分为三类：(1)FCN扩展系列；(2)特征图分辨率的提升；(3)图像上下文信息的利用。

目标检测算法与裂缝病害的匹配适应性差，在训练样本制作阶段发现对于形态复杂延展范围大的单个裂缝需要使用多个标记框进行框选，该过程具有人工主观性，缺乏制作规则的统一性，导致在推理阶段检测算法对复杂裂缝标记框的输出数量与位置具有不确定性。基于目标检测网络的裂缝检测方法仍属于粗略的裂缝检测工作。分割网络的出现突破了目标检测网络的瓶颈，实现了像素级别的精细化裂缝提取。分割网络的输出效果如下图所示。

7-5分割网络的预测结果

基于深度学习的裂缝分割示例（MATLAB）

#加载并构建训练、测试集合

```
outputFolder = fullfile(pwd,'CamVid');
imgDir = fullfile(outputFolder,'images');
imds = imageDatastore(imgDir);
classes = ["Background"
"Crack"];
labelIDs{1,1} = [0,0,0];
labelIDs{2,1} = [255,255,255];
labelDir = fullfile(outputFolder,'mask');
pxds = pixelLabelDatastore(labelDir,classes,labelIDs);
imageFolder = fullfile(outputFolder,'imagesResized',filesep);
...
numTestingImages = numel(imdsTest.Files)
#构建分割网络：创建使用 VGG-16 权重初始化的全卷积网络层。
vgg16();
imageSize = [360 480];
numClasses = numel(classes);
```



```

lgraph = fcnLayers(imageSize,numClasses);
... ..
lgraph = removeLayers(lgraph,'pixelLabels');
lgraph = addLayers(lgraph, pxLayer);
lgraph = connectLayers(lgraph,'softmax','labels');
#训练构建好的分割网络
options = trainingOptions('adam', ...
'InitialLearnRate',1e-4, ...
'MaxEpochs',50, ...
'MiniBatchSize',4, ...
'Shuffle','every-epoch', ...
'CheckpointPath', tempdir, ...
'VerboseFrequency',2...
'Plots','training-progress');
... ..
[net, info] = trainNetwork(dsTrain, lgraph, options);
save('FCN8sCamVid.mat','net');
#使用训练好的分割网络对新图像进行分割
I = readimage(imdsTest,n); # n= 60, 63
imshow(I)
C = semanticseg(I, net);
cmap = camvidColorMap;
B = labeloverlay(I,C,'ColorMap',cmap,'Transparency',0.4);
imshow(B)

```

7.6应用拓展

目标识别与分割在土木工程领域应用广泛，并不仅仅局限于单一技术的使用，而是通过综合应用各种技术，为后续的分析任务提供服务。任务的目的是提高图像中各种信息的高效提取能力，使计算机具备一定的图像理解能力，相较于传统方法，能够更好地应对土木工程中各个应用领域中复杂的信息需求。在土木工程中，信息的复杂性很高，仅仅进行某个分类、识别或分割任务是无法满足项目需求的。因此，以下是一些具体的土木工程领域方向，展示了目标识别与分割在该领域的广泛应用：车流识别，通过目标识别与分割技术，可以准确地识别和跟踪道路上的车辆，帮助监测交通流量、分析交通状况、优化道路规划和交通控制等。这对于城市交通管理和规划非常重要。船舶识别，在港口管理和海洋工程中，利用目标识别与分割技术，可以对船舶进行自动识别和分类，监测航行状态、识别危险情况。安全帽识别：在建筑工地和工业领域中，目标识别与分割可用于识别工人是否佩戴安全帽，确保工人的安全。通过实时监测，及时发现安全隐患并采取措施，提高工作场所的安全性。除了上述例子，目标识别与分割在土木工程中还应用于许多其他领域，如建筑物损伤检测、道路设施维护、桥梁结构监测等。这些应用利用了目标识别与分割技术的复杂性和多样性，为土木工程领域提供了更精确、高效和可靠的信息提取和分析手段。通过了解这些广泛的应用领域，读者可以清楚地认识到目标识别与分割在土木工程中的重要性和复杂性，以及该领域对计算机技术的需求。这将帮助他们更好地理解和应用相关知识，为未来的土木工程项目提供更好的解决方案。

7.6.1车流识别

车辆信息的识别在交通控制、城市管理、基础设施运营维护中都有很重要的意义。特别是在当前人工智能的大背景下，如何实现高自动化、智能化的车辆信息识别是业内研究的热点。视频作为一种采集成本较低且所含信息丰富的数据形式，在车辆信息识别中被广泛应用。本节将对目标识别与分割任务的实际应用扩展，面向桥梁交通场景的基于实例分割的车辆识别方法进行介绍。提出的车辆信息识别方法基于图像实例分割网络Mask R-CNN，通过该方法可同时实现车型识别、车辆轴数识别。

本节桥面车辆信息识别依托工程为安庆长江公路大桥，该桥是安徽省境内连接池州市和安庆市的过江通道。安庆长江公路大桥主桥为连续钢箱梁斜拉桥，包含南北两座索塔。为捕获桥面车辆信息，利用南北索塔上安装的摄像机，如图所示。桥面采用双向四车道，桥面车流方向与摄像机位置如图所示。

图索塔上摄像机的安装位置

图桥面车流方向与摄像机位置示意图

Mask R-CNN是一个包含目标2D边界框定位，目标分类与实例分割的多任务网络[159]，其结构如图所示。同Faster R-CNN一样，它也分为候选框生成与目标检测两部分。在Mask R-CNN中，为了更充分地提取图像特征，采用了特征金字塔网络(FPN)[163]。本文中采用实例分割网络Mask R-CNN的原因是，由于其采用了基于候选区域的目标分割，使得在图像中一个像素可以同时属于不同的类别目标，而不同于FCNs[164]等网络对于一个像素只允许属于一个目标。

图 Mask R-CNN结构

本节中建立了一个包含轿车、客车、货车及车轮四种目标类别的小型图像数据库用于训练Mask R-CNN。数据库中场景包括桥梁、高速公路及城市街道，其中一些图像来自视频数据集[94]。三种车型如图所示。数据集中共有342张大小为1280×720及1920×1080像素的标注图像，其中243张用于训练，其余的用于测试。训练集中图像通过水平翻转进行数据增强。为了标注图像，使用了开源的标注工具LabelMe，相应的标注过程在图中展示。在图(a)中展示了原始的图像样本。根据相应车型对图像中整个车辆进行标注如图(b)所示。然后对车轮目标进行标注。图(c)中展示了车辆与车轮的标注结果。原始图像融合相应的标注结果展示在图(d)中。采用这种标注方式的目的是使属于车轮类的像素同时也归属于车类，这是计算车辆轴数的基础。

图三种车型示例

图图像标注过程

为了获得车辆轴数，有效的车轮识别方法是前提。本小节首先比较不同的车轮识别方法。第一种方法利用车轮在图像中近

似椭圆的性质，采用目标轮廓形状检测来识别车轮，相关的检测结果展示在图(a)中，其中只显示在车辆像素范围内检测到的椭圆。从图中明显发现车轮的误检率与漏检率均很高。第二种车轮识别方法基于模板匹配[169]，如图(b)所示。图中的误检情况比较明显，并且基于模板匹配的方法对于不同场景的通用性较差，需要设计针对不同场景的匹配模板。而基于Mask R-CNN实例分割的车轮识别结果展示在图3-10(c)中，在三种场景中其识别效果均为最佳。因此本文的车辆轴数计算基于Mask R-CNN的车轮识别。

17. 计算机视觉教材_第17部分		总字符数：2922
相似文献列表		
去除本人文献复制比：56.5%(1652)		去除引用文献复制比：56.5%(1652) 文字复制比：56.5%(1652)
1	基于多重深度学习网络的安全帽检测及工地人员身份识别方法研究 黄毅;包世洪;马亮;赵林;张毅; - 《建筑安全》- 2023-04-05	30.7% (898) 是否引证：否
2	基于计算机视觉的桥梁车流信息识别与防船撞预警方法研究 张博(导师：张建) - 《东南大学博士论文》- 2021-03-05	25.2% (736) 是否引证：否
原文内容		

图基于不同方法的车轮识别结果：(a) 基于目标轮廓形状检测，(b) 基于模板匹配，
(c) 基于Mask R-CNN实例分割
图车轮模板示例：(a)，(b)，(c) 中模板分别对应图3-10中三个不同场景

7.6.2船舶识别
图广东九江大桥概貌

在船桥及船间碰撞预警中，船舶实时定位是前提。然而不同于道路上的交通场景，在水面上并没有稳定的参考物，这给船舶定位带来了很大挑战。目前还缺乏有效的基于视频图像的船舶实时定位方法。为此，本节介绍一种基于视频的船舶目标识别与跟踪方法，进而获得船舶的实时像素坐标。依托工程均为广东九江大桥。该桥是沈海高速佛开段上的一座特大型桥梁，如图(a)所示，其跨越佛山南海区与鹤山市交界处的西江干流。桥梁全长1800余米，共49跨。第20~25跨采用变截面预应力连续箱梁，相应的跨径组合为50+100+160+160+100+50m，如图(b)所示，其中间两跨为船舶通航通道。

船舶的识别与跟踪是定位的基础。船舶识别采用的是基于ResNet-50主干的Mask R-CNN。为了训练该网络使其具有识别船舶的能力，建立了一个关于船舶的图像数据库，其中采用工具LabelMe对图像进行标注。数据库中的图像样本采集来自桥上的多部摄像机，部分样本如图所示。为了全天24小时的船舶识别，数据库中包含有了不同光照条件下的样本。

图船舶图像数据库中的样本示例及标注结果

图中展示了若干船舶跟踪实例，其中视频每两秒采样一帧，并且在跟踪中仅考虑处在检测区中的船舶目标，而忽略检测区外识别到的船舶目标。对于一艘已被跟踪的船舶，若超过三帧没有识别目标与其匹配，则判断该船已驶出检测区。在图5-7中，船舶的轨迹是通过连接跟踪过程中所有匹配到的船舶2D边界框的中心而生成的，该边框中心近似作为船舶的几何中心用于船舶定位。

图检测区内的船舶跟踪

7.6.3智慧工地安全帽识别

安全帽作为一种有效的劳动防护用品，能够起到缓冲减震和分散应力的作用，并且不同颜色的安全帽代表不同的等级身份，因此监测工地人员佩戴的安全帽的情况，一方面能够对未佩戴安全帽的人员做到及时提醒，另一方面能够根据安全帽颜色识别出人员身份，从而判断该身份是否符合作业要求，对于保障工地人员的安全起到关键作用。传统工地监控采用的是人工监控的方式，但工作人员面对大量的监控屏幕，难免出现漏判误判的情况。深度学习的引入使得安全帽的识别更快速准确。

在建筑工地内，工地人员、机器设备、建筑材料众多，环境复杂，通过直接识别检测安全帽的方法易受到图像中其他信息的干扰，造成误检的情况。这里介绍一种通过区分前景和背景的方法，首先将图像中人体前景区域提取出来，再在此区域的基础上进行安全帽检测，从而提高模型的识别准确率。

图基于学习网络的安全帽检测及工地人员身份识别方法

具体的检测流程：首先利用 Yolov4 网络对输入的待检测图像进行处理，识别出图像中的工地人员。在得到工地人员识别检测的结果，即图像中工地人员识别检测边界框的 (x , y , w , h) ，通过坐标变化计算得到边界框位置参数 { xleft , xright , ytop , ybottom } 。根据图像中边界框的位置参数进行图像裁剪，从而提取出图像中工地人员前景区域，作为安全帽检测网络的输入，从而避免图像中背景区域对于安全帽识别检测的干扰。

在提取出图像中工地人员前景区域后，将该前景区域图像再次输入到 Yolov4 网络中进行安全帽检测，从而判断工地人员是否佩戴安全帽以及通过安全帽颜色识别工地人员的身份，若图像中未检测到安全帽，则代表该工地人员未佩戴安全帽，反之若检测到安全帽，则根据安全帽不同的颜色判断该工地人员的身份，安全帽检测和工地人员身份识别方法的流程图如图所示。

模型训练数据库的图片来源某一施工现场。构建流程：首先利用某一工地的监控视频进行分帧，保存共 177 张图片，然后利用 labelImg 软件对每张图片上安全帽检测目标进行标注如图所示，生成对应的 xml 文件。最后建立的安全帽数据集共包含 177 张图像。在此基础上训练模型，共标注了 5 类安全帽的种类，包括红色安全帽、橙色安全帽、白色安全帽、黄色安全帽以及蓝色安全帽。根据该工地现场的规定：红色安全帽一般是管理人员佩戴；橙色和黄色安全帽一般是普通工人佩戴；白色安全帽一般是监理人员佩戴；蓝色安全帽一般则技术人员佩戴 。因此通过检测安全帽的颜色可以相应判断工人的身份。

小结

本节内容介绍了计算机视觉领域中的目标识别与分割技术，主要包括分类、检测和分割三个方面。首先，我们学习了目标分类。目标分类是指将输入的图像或视频中的物体归类到不同的类别中。我们了解到分类任务可以通过使用深度学习方法，如卷积神经网络（CNN）来实现。通过训练一个分类器，我们可以对图像中的不同物体进行准确的分类。其次，我们探讨了目标检测。目标检测是在图像或视频中定位并识别多个目标的任务。我们了解到目标检测可以分为两个主要阶段：物体定位和物体分类。常用的目标检测算法包括基于区域的方法，如候选区域提取和区域分类，以及基于锚点的方法，如单阶段检测器。这些方法结合了特征提取、感兴趣区域生成和目标分类等步骤，实现了高效准确的目标检测。最后，我们学习了目标分割。目标分割是将图像中的每个像素分配给特定的目标类别，从而实现对目标的精确边界和轮廓的提取。通过学习本节内容，我们深入了解了计算机视觉领域中目标识别与分割的关键概念和方法。掌握了这些技术，我们可以在图像处理、自动驾驶、医学图像分析等领域中应用它们。对于进一步的学习和研究，我们还可以探索更多的目标识别与分割算法和应用场景，以不断提高计算机视觉的性能和效果。

习题

1. 在土木工程领域中，目标识别与分割可以在哪些方面提供帮助？请列举具体的应用场景，并解释其价值和意义。
2. 目标识别与分割中的数据集对算法性能和泛化能力有着重要影响。请说明常用的目标识别与分割数据集，并解释其在算法评估和训练中的作用。
3. 解释监督学习和无监督学习在图像分类中的区别，并举例说明它们在实际应用中的差异和适用性。
4. 解释准确率（Accuracy）、精确率（Precision）、召回率（Recall）和F1分数（F1 Score）等评估指标，并说明它们如何衡量图像分类模型的性能。
5. 选择一种常用的目标检测算法（如Faster R-CNN、YOLO或SSD），详细描述其原理和流程，并解释它在目标检测任务中的优势和局限性。
6. 解释图像语义分割的基本思想和技术方法，并列举一些常用的语义分割算法。
7. 请解释语义分割和实例分割的区别，并举例说明它们在土木工程领域的应用。
8. 解释图像语义分割的评估指标，如IoU（Intersection over Union）和mIoU（mean Intersection over Union），并说明它们如何衡量分割结果的准确性。
9. 选择一篇最近的目标识别与分割相关的研究论文，并进行摘要、方法和实验结果的简要介绍。

18. 计算机视觉教材_第18部分		总字符数：9218
相似文献列表		
去除本人文献复制比：56.1%(5171) 去除引用文献复制比：56.1%(5171) 文字复制比：56.1%(5171)		
1	<u>摄影测量影像快速立体匹配关键技术研究</u> 李迎松(导师：郑顺义) - 《武汉大学博士论文》 - 2018-05-01	15.8% (1452) 是否引证：否
2	<u>基于机器视觉的着火点定位方法研究</u> 冯凌云(导师：杨国田) - 《华北电力大学（北京）硕士论文》 - 2010-01-01	14.4% (1330) 是否引证：否
3	<u>基于双目和结构光的弱特征SLAM研究</u> 李平伯(导师：罗立民;胡轶宁;刘庭华) - 《东南大学硕士论文》 - 2019-06-13	14.0% (1293) 是否引证：否
4	<u>双目视觉技术在四旋翼无人机障碍物识别中的研究</u> 樊建昌(导师：余粟) - 《上海工程技术大学硕士论文》 - 2020-04-01	12.4% (1139) 是否引证：否
5	<u>基于PUMA560机器人的视觉伺服控制系统的研究</u> 刘锋(导师：杨逢瑜) - 《兰州理工大学硕士论文》 - 2007-05-10	8.3% (764) 是否引证：否
6	<u>基于双目立体视觉的塑料模型逆向工程研究</u> 李育宽(导师：韩立强) - 《长春工业大学硕士论文》 - 2010-03-01	7.9% (726) 是否引证：否
7	<u>基于双目立体视觉的空间非合作目标的位姿测量</u> 董圣男(导师：王小涛) - 《南京航空航天大学硕士论文》 - 2010-01-01	7.5% (691) 是否引证：否
8	<u>双目立体视觉及管口视觉测量系统研究</u> 张浩鹏(导师：王宗义) - 《哈尔滨工程大学硕士论文》 - 2008-12-10	6.4% (590) 是否引证：否
9	<u>一种稳健性增强和精度提升的增量式运动恢复结构方法</u> 于英;张永生;薛武;王涛; - 《测绘学报》 - 2019-02-15	5.6% (520) 是否引证：否
10	<u>基于单目视觉的视觉伺服与位姿估计 - 道客巴巴</u> - 《互联网文档资源（ https://www.doc88.co ）》 - 2020	4.8% (439) 是否引证：否
11	<u>基于单目视觉的视觉伺服与位姿估计</u> 田凯(导师：孙容磊) - 《华中科技大学硕士论文》 - 2009-05-01	4.8% (439) 是否引证：否

12	基于单目视觉的视觉伺服与位姿估计 - 道客巴巴 - 《互联网文档资源 (https://www.doc88.co)》 - 2020	4.7% (436)	是否引证: 否
13	基于无人机航拍图像的三维重建SfM算法研究 韩超(导师: 李健) - 《内蒙古工业大学硕士学位论文》 - 2019-06-01	4.5% (412)	是否引证: 否
14	基于双目视觉的大尺寸工件在机三维测量 黎淑梅(导师: 全燕鸣) - 《华南理工大学硕士学位论文》 - 2011-05-27	4.4% (410)	是否引证: 否
15	结构光三维测量系统中匹配与拼合技术研究 车向前(导师: 王巍) - 《哈尔滨工程大学硕士学位论文》 - 2008-04-01	4.3% (399)	是否引证: 否
16	基于双目立体视觉的移动机器人目标定位 韩云生(导师: 刘国栋) - 《江南大学硕士学位论文》 - 2009-05-01	4.0% (368)	是否引证: 否
17	一种关于智能充电桩的系统设计 岑王洁;钱家鑫;李凯;杨嘉琦;倪嘉锋; - 《电脑与信息技术》 - 2020-10-15	3.8% (350)	是否引证: 否
18	基于机器视觉的磁性材料检测方法研究 李大伟(导师: 林明星) - 《山东大学硕士学位论文》 - 2011-05-22	3.6% (329)	是否引证: 否
19	基于极线几何的机器人视觉伺服控制系统分析 杨逢瑜;王其磊;关红艳;陈君辉;王顺;娄磊; - 《西华大学学报(自然科学版)》 - 2009-01-15	3.3% (307)	是否引证: 否
20	基于卷积神经网络和双目视觉的汽车门板焊点识别与定位系统的研究 莫志敏(导师: 谢巍;陈立定) - 《华南理工大学硕士学位论文》 - 2020-05-06	3.0% (272)	是否引证: 否
21	基于椭圆特征的双目视觉空间平面定位方法研究 杨晓俊(导师: 朱兴龙) - 《扬州大学硕士学位论文》 - 2009-05-01	2.8% (259)	是否引证: 否
22	基于激光扫描与拼接的热态大型锻件视觉测量方法研究 李亮(导师: 赵转萍) - 《南京航空航天大学硕士学位论文》 - 2009-12-01	2.8% (256)	是否引证: 否
23	双目立体视觉中图像匹配算法的研究 王月梅(导师: 曹伯燕) - 《西安电子科技大学硕士学位论文》 - 2009-01-01	2.7% (246)	是否引证: 否
24	基于标记点的三维人体步态捕获技术研究 胡琴(导师: 刘任任) - 《湘潭大学硕士学位论文》 - 2008-05-05	2.6% (244)	是否引证: 否
25	三维场景重建的理论研究和工程实现 谷永发(导师: 石光明) - 《西安电子科技大学硕士学位论文》 - 2009-01-01	2.5% (230)	是否引证: 否
26	基于运动恢复结构三维重建的应用与研究进展 王巧丽;徐增波;张玲; - 《毛纺科技》 - 2021-07-16	2.0% (187)	是否引证: 否
27	车载式路面异常检测定位系统及方法研究 董学恒(导师: 王建锋) - 《长安大学硕士学位论文》 - 2022-03-15	1.4% (126)	是否引证: 否
28	无人机影像增量式运动恢复结构研究进展 陈武;姜三;李清泉;江万寿; - 《武汉大学学报(信息科学版)》 - 2022-10-05	1.2% (110)	是否引证: 否
29	基于多视图的非增量式三维重建关键技术的研究与应用 杨静(导师: 郭向坤) - 《中国科学院大学(中国科学院沈阳计算技术研究所)硕士学位论文》 - 2018-06-01	1.0% (89)	是否引证: 否
30	双目视觉测距系统软硬件设计研究 龚文超;朱凤波;刘洁; - 《舰船电子工程》 - 2021-04-20	0.9% (84)	是否引证: 否
31	基于多视角数字图像的深基坑三维重构研究 黄身勇;蒋国平; - 《河南科技》 - 2021-06-05	0.9% (83)	是否引证: 否
32	基于双目视觉自动导引车的障碍识别与避障策略 何嘉健(导师: 吴耀华) - 《山东大学硕士学位论文》 - 2015-04-17	0.8% (76)	是否引证: 否
33	绳驱式蛇形机械臂在复杂环境下的运动控制研究 何胜波(导师: 汪平河;冯常) - 《电子科技大学硕士学位论文》 - 2021-05-11	0.7% (64)	是否引证: 否
34	面向数字孪生建筑的“信息-物理”交互策略研究 韩冬辰(导师: 宋晔皓) - 《清华大学博士学位论文》 - 2020-05-01	0.6% (53)	是否引证: 否
35	三维重建技术在轮式收割机设计中的应用研究	0.4% (35)	

尹红俊; - 《乡村科技》- 2021-12-30		是否引证: 否
36	基于变焦双目立体视觉的三维重建系统研究	0.3% (30)
李璐(导师: 钟桦;黎军) - 《西安电子科技大学硕士论文》- 2020-04-01		是否引证: 否

原文内容

第8章立体视觉与三维重建

图像三维重建 (Image-based 3D Reconstruction) 是一种利用单张或多张图像的信息, 通过计算机算法还原物体或场景的三维结构和形状的过程。它以图像为输入, 通过对图像中的特征点、纹理信息和视角关系进行分析和计算, 推导出物体的三维几何信息, 以创建物体的立体模型。三维重建可以应用于计算机视觉、计算机图形学、机器人视觉、虚拟现实、增强现实、医学影像处理等领域。在土木工程领域, 图像三维重建技术提供了结构几何形态、变形行为和性能分析的工具, 通过对比不同时间点下的图像重建模型, 我们可以进行施工进度的定量分析或土木结构的变形监测。图8- 1给出了一座拱桥的图像建模结果, 从此模型中可提取出结构当前状态下的几何线形。

图8- 1 桥梁结构三维重建

8.1概述

在三维视觉信息获取方面, 我们可以根据照明方式将其分为主动视觉测量和被动视觉测量。主动视觉测量利用光源或能量源发射到目标物表面, 并通过接收返回的光波来获取物体的深度信息, 常见的有结构光法 (Structured Light)、飞行时间法 (Time-of-Flight) 和三角测距法 (Triangulation), 工作原理见表8-1。

表8-1主动视觉测量方法

分类	工作原理	示意图	适用场景
结构光法	光学器将一定模式的光条或光斑投射于物体表面, 另一位置的相机记录下物体表面图像, 通过解调畸变的二维光条图像, 计算出物体表面的深度信息		室内环境下的近距离 (10米内) 的精确测量应用于小尺度表面几何测量、工业检测等
TOF	通过测量激光脉冲的飞行时间来获取距离信息, 使用扩散器将激光束扩散来实现空间内的同步测量		实时测量, 数十米范围内的场景, 适用于动作捕捉、机器人导航等
三角测距法	构造包含两个感测器以量测物体。这两个感测器的中心点以及对焦于物体表面的同一点, 形成一个空间上的三角。于此三角形内, 已知两感测器间的距离。借由找出两感测器投射线与基准线间的夹角, 便可用三角测量法得知两投射线交点的3D座标。		适用于较远距离的目标测量, 广泛应用于地理测量、工程测量和导航定位等领域
激光扫描法	在三角测距法基础上, 通过测量激光束从发射到返回的时间或相位差来计算物体表面的距离, 通过旋转镜或其他扫描机构的运动, 激光束可以在水平和垂直方向上进行扫描, 从而覆盖整个三维空间		毫米级精度, 覆盖数十米到千米级的测量范围, 应用于建筑测量、地形测绘、文物保护等。

分类工作原理示意图适用场景

结构光法光学器将一定模式的光条或光斑投射于物体表面, 另一位置的相机记录下物体表面图像, 通过解调畸变的二维光条图像, 计算出物体表面的深度信息室内环境下的近距离 (10米内) 的精确测量应用于小尺度表面几何测量、工业检测等

TOF 通过测量激光脉冲的飞行时间来获取距离信息, 使用扩散器将激光束扩散来实现空间内的同步测量实时测量, 数十米范围内的场景, 适用于动作捕捉、机器人导航等

三角测距法构造包含两个感测器以量测物体。这两个感测器的中心点以及对焦于物体表面的同一点, 形成一个空间上的三角。于此三角形内, 已知两感测器间的距离。借由找出两感测器投射线与基准线间的夹角, 便可用三角测量法得知两投射线交点的3D座标。 适用于较远距离的目标测量, 广泛应用于地理测量、工程测量和导航定位等领域

激光扫描法在三角测距法基础上, 通过测量激光束从发射到返回的时间或相位差来计算物体表面的距离, 通过旋转镜或其他扫描机构的运动, 激光束可以在水平和垂直方向上进行扫描, 从而覆盖整个三维空间毫米级精度, 覆盖数十米到千米级的测量范围, 应用于建筑测量、地形测绘、文物保护等。

被动视觉测量利用自然光的反射, 通过相机获取图像, 再使用特定算法计算出物体的立体空间信息, 它的优势在于成本较低、易于操作。根据摄像机数目的不同, 被动视觉测量可以分为单目视觉法、双目视觉法和多目视觉法。单目视觉法通过图像的二维特征 (明暗度、纹理、焦点、轮廓等) 推导出深度信息的方法。这一类方法设备结构和算法都较简单, 但在深度估计的精度和重建效果方面相对一般, 更适用于一些简单的场景或对深度精度要求不高的应用, 如目标检测、场景分析等。双目视觉法, 通常使用光轴平行的左右两部相机, 预先标定出双目的内参和外参 (两相机的相对位置和姿态), 利用三角测量的原理, 将匹配特征点的视差实时转换为深度信息, 重建质量、效率、自动化程度都很高。多目视觉法通过多视角的几何关系, 避免双目视觉中难以解决的假目标、边缘模糊及误匹配问题, 可以应对更复杂的场景。根据计算方法也可以分为运动恢复结构 (Structure from Motion) 和机器学习法等。运动恢复结构法是基于多个视图的图像序列, 通过匹配算法和三角测量原理, 从中恢复出场景的三维结构和相机的运动轨迹。机器学习法则无需显式的匹配过程, 通过模型训练来直接预测图像深度信息。

本节主要阐述双目视觉、运动恢复结构两种三维重建方法, 最后介绍如何对对生成的点云模型进行后处理。

8.2双目立体视觉

双目立体视觉是基于视差原理, 即两个相机看到的图像不一样的, 两幅图像的差别就构成了视差。由三角法原理进行三维信息的获取, 即由两个摄像机的图像平面和被测物体之间构成一个三角形, 预先标定两摄像机之间的位置关系, 便可以获取两摄像机公共视场内物体特征点的三维坐标。

8.2.1 测量原理与数学模型

双目视觉系统由两个相机构成，称为双目视觉系统的左右相机，左右相机可以是完全不同的两个相机，但是从双目系统性能的角度考虑的话，一般尽量做到使左右相机相同。根据两个相机安装位置的不同可分为平行式光轴双目视觉系统和汇聚式光轴双目视觉系统，如图8-2所示。

(a) 平行式光轴 (b) 汇聚式光轴

图8- 2 双目相机系统模式

典型的平行式光轴双目视觉系统中，两个摄像机相同且它们的坐标系统的各对应轴精确地平行，只是原点位置不同，如图8-3所示。两摄像机的投影中心连线的距离，即基线距为B。以左相机的坐标系为基准，两摄像头在同一时刻观看空间物体的同一特征点，分别在左相机和右相机上获取了点P的投影，图像坐标分别为，。由于两摄像机的图像在同一平面上，则特征点P的图像坐标中y坐标相等，即，则由三角几何关系得到：

(8.6)

则视差。由此可以计算出特征点P在左相机坐标系下的三维坐标为

(8.7)

图8- 3 平行式光轴双目视觉系统成像原理

采用这种模式，左摄像机像面上任意一点只要能在右摄像机像面上找到对应的匹配点，就可以确定出该点的三维坐标。这种方法是点对点的运算，基于两像面上匹配点对，就可以获取点的三维坐标。由式(8.7)可知，深度信息与视差相关联，若由于匹配点对误差引起视差d不准确，将导致距离Z产生误差。假设偏差e，则有，引起的距离误差为

(8.8)

考虑 $Z \gg \lambda$ ，将式(8.7)代入可得

(8.9)

可见，测距精度与相机焦距f、相机基线长度B、物距Z都相关。当相机焦距f越大，相机基线长度B越长，则精度越高；当物距Z越大，则精度越低。

平行式光轴为特殊情况，一般情况下对两部摄像机的摆放位置不做特别要求。设左摄像机O-XYZ位于世界坐标系的原点处且无旋转，图像坐标系为，有效焦距为；右摄像机坐标系为，图像坐标系为，有效焦距为。

图8- 4 汇聚式光轴双目视觉系统成像原理

由摄像机透视变换模型有

(8.10)

(8.11)

O-XYZ坐标系与坐标系之间的相互位置关系可通过空间转换矩阵表示。

(8.12)

其中，和分别为O-XYZ坐标系与坐标系之间的旋转矩阵和原点之间的平移变换矢量。

由式-可知，对于O-XYZ坐标系中的空间点，两摄像机像面点之间的对应关系为

(8.13)

于是，空间点三维坐标可以表示为

(8.14)

因此，已知焦距、和空间点在左、右摄像机中的图像坐标和，只要求解出两相机间的旋转矩阵R和平移矢量T，就可以得到被测物体的三维空间坐标。

8.2.2 极线几何

极线几何讨论的是两个摄像机图像平面的关系，不仅在双目立体视觉测量中两幅图像对应点匹配有着重要作用，而且在三维重建和运动分析中也具有广泛的应用。极线几何关系如下图8-5对于三维空间中一点P，由相机投影到两个不同的平面、，投影点分别为、。、为相机的光心，使P、在三维空间内构成一个平面S称为极平面。摄像机的光心和的连线构成基线。基线与摄像机图像平面的交点为极点。S与面的交线过点，称之为对应于的极线。同理S与的交线称之为对应于的极线。基线与两摄像机图像平面的交点、为极点。极线约束就是同一个点在两幅图像上的映射，已知左图映射点，那么右图映射点一定在相对于的极线上，反之亦然。极线约束给出了对应点重要地约束条件，它将对应点匹配从整幅图像寻找压缩到在一条直线上寻找对应点。因此，极大地减小了搜索地范围，对对应点匹配具有指导作用。

图8- 5 极线几何关系

如果用投影矩阵表示，空间点三维坐标可以由两个摄像机的投影模型表示，

其中：、分别为空间点在左、右摄像机中的图像坐标，为空间某点P在世界坐标系中的三维坐标，、分别为左、右摄像机的投影矩阵，将、矩阵中左面的 3×3 部分分别记作和，右边的 3×1 部分分别记作、。如果将，记作其中，则式(8.15)可写为将(8.16)消去X得由于上述中两边是三维向量，上述包含了三个等式，用这个等式消去与后，就可以得到一个与、无关的与的关系，这就是极线约束。

为了使上述消去过程更清晰，在此引入反对称矩阵：如果t为三维向量，称下列矩阵为由t定义的反对称矩阵，记作由反对称矩阵定义知，是一个不满秩的不可逆矩阵。

将式(8.17)右边的向量记作m，即将由m定义的反对称矩阵记作，将左乘(8.17)两边，则由可知将上式两边除以并记，得式(8.21)右边向量为，可见该向量与正交，将左乘式(8.21)两边，并将所得式两边除以s后得到以下重要关系：

式(8.22)的意义是，它给出了、所必须满足的关系。在给定的情况下，式(8.22)是关于的线性方程，即图像上的极线方程。反过来，在给定的情况下，式(8.22)是关于的线性方程，即图像上的极线方程。

令，则式(8.22)可以写成，F称为基本矩阵，基本矩阵是极线几何的一种代数表示，将极线约束采用基本矩阵以解析形式可以表示为：

如果已知左、右摄像机的内参数短阵、和两个摄像机之间的结构参数R、T，则极平面方程又可以表示为

其中，S为反对称矩阵，它由平移是矢量定义为。因此定义基本矩阵又可以表示为从公式(8.26)可以看出，基本矩阵实际上包括双目立体视觉测量系统的所有参数，即两个摄像机内部参数、，和视觉系统的结构参数R、t。这表明基本矩阵只与视觉测

量系统的参数(摄像机内部参数和两摄像机结构参数)参数有关,与外部场景无关,是双目立体视觉内在的一种约束关系。 F 矩阵在立体视觉与运动视觉中是一个很重要的矩阵。当摄像机内部参数、已知,可以求得特征对应点的归一化图像坐标,公式(8.26)可以写成

定义 $E=RS$ 为本质矩阵,它只与视觉测量系统的结构参数有关。

基本矩阵 F 是具有7个自由度的、秩为2的齐次矩阵,它描述了两个摄像机的相对位置。基本矩阵 F 可以分解成只差一个矩阵因子的两个摄像机的投影矩阵,因此重建结果只差一个矩阵因子 H 。所以这种重建是在射影几何意义下的重建,所计算出来的几何元素保持射影变换群下的不变量。由本质矩阵 E 可以分解成两个摄像机位置的旋转矩阵 R 和带一比例因子的平移矩阵 t ,因此在运动参数分析中具有重要作用。本质矩阵是具有5个自由度的、秩为2的矩阵。

基本矩阵 F 和本质矩阵 E 的性质简单介绍如下:

(1) 基本矩阵 F 的性质。

①具有7个自由度的、秩为2的齐次矩阵;

②如果与为对应图像点,则满足;

③为对应于的极线,为对应于的极线;

④极点为。

(2) 本质矩阵 E 的性质。

①具有5个自由度的、秩为2的矩阵;

②有一个奇异值为0,另外两个奇异值相等;

③;

④仅由平移来决定;

⑤。

8.2.3 双目立体匹配

双目立体匹配一直是双目视觉的研究热点,双目相机拍摄同一场景的左、右两幅视点图像,运用立体匹配匹配算法获取视差图,进而获取深度图。而深度图的应用范围非常广泛,由于其能够记录场景中物体距离摄像机的距离,可以用以测量、三维重建、以及虚拟视点的合成等。双目立体匹配可分为四个步骤:匹配代价计算、代价聚合、视差计算和视差优化。

1) 匹配代价计算

匹配代价计算的目的是计算参考图像上的每一个待匹配像素与目标图上每一个可能的视差范围内的候选像素之间的相关性,两个像素无论是否为同名点,都可以通过匹配代价函数计算匹配代价,代价越小则说明相关性越大,是同名点的概率也越大。每个像素在搜索同名点之前,往往会指定一个视差搜索范围 D ,视差搜索时将范围限定在 D 内,用一个大小为 $W \times H \times D$ (W 为影像宽度, H 为影像高度)的三维矩阵 C 来存储每个像素在视差范围内每个视差下的匹配代价值。矩阵 C 通常称为DSI(Disparity Space Image)。匹配代价计算的方法有很多,传统的摄影测量中,使用灰度绝对值差(AD, Absolute Differences)、灰度绝对值差之和(SAD, Sum of Absolute Differences)、归一化相关系数(NCC, Normalized Cross-correlation)等方法来计算两个像素的匹配代价;计算机视觉中,多使用互信息(MI, Mutual Information)法、Census变换(CT, Census Transform)法、Rank变换(RT, Rank Transform)法、BT(Birchfield and Tomasi)法等作为匹配代价的计算方法。不同的代价计算算法都有各自的特点,对各类数据的表现也不尽相同,选择合适的匹配代价计算函数是立体匹配中不可忽视的关键步骤。比如Census保留周边像素空间信息,对光照变化有一定鲁棒性,具有灰度不变的特性,AD边缘明显清晰,纹理重复结构效果还可以,但不适用于噪声多的情况,低纹理区域效果不好,对颜色值敏感,AD和Census结构可以取一个折中。

2) 代价聚合

代价聚合的根本目的是让代价值能够准确的反映像素之间的相关性。匹配代价的计算往往只限于局部窗口中,只会考虑局部相关性,通过两个像素邻域内一定大小的窗口内的像素信息来计算代价值,对影像噪声非常敏感,而且当影像处于弱纹理或重复纹理区域,这个代价值极有可能无法准确的反映像素之间的相关性,直接表现就是真实同名点的代价值非最小。而代价聚合则是建立邻接像素之间的联系,以一定的准则,如相邻像素应该具有连续的视差值,来对代价矩阵进行优化,这种优化往往是全局的,每个像素在某个视差下的新代价值都会根据其相邻像素在同一视差值或者附近视差值下的代价值来重新计算,得到新的DSI,用矩阵 S 来表示。实际上代价聚合类似于一种视差传播步骤,信噪比高的区域匹配效果好,初始代价能够很好的反映相关性,可以更准确的得到最优视差值,通过代价聚合传播至信噪比低、匹配效果不好的区域,最终使所有影像的代价值都能够准确反映真实相关性。

3) 视差计算

视差计算即通过代价聚合之后的代价矩阵 S 来确定每个像素的最优视差值,通常使用赢家通吃算法(WTA, Winner-Takes-All)来计算,即某个像素的所有视差下的代价值中,选择最小代价值所对应的视差作为最优视差。这一步非常简单,这意味着聚合代价矩阵 S 的值必须能够准确的反映像素之间的相关性,也表明上一步代价聚合步骤是立体匹配中极为关键的步骤,直接决定了算法的准确性。

4) 视差优化

视差优化的目的是对上一步得到的视差图进行进一步优化,改善视差图的质量,对错误匹配进行剔除、适当平滑以及子像素精度优化等步骤。错误匹配是指像素在代价聚合后得到的视差值并不是真实的最小的视差值。这种错误匹配的现象在实际中并不少见,图像的噪声、遮挡、弱纹理和重复纹理都会导致这一现象产生。一般采用左右一致性检查算法剔除错误视差,左右一致法就是将左右图像的对应点的视差值进行比较,若两个视差值小于一定的阈值,则认为匹配成功,否则认为该视差不满足唯一性将被剔除。采用剔除小连通区域算法来剔除孤立异常点;采用中值滤波、双边滤波等平滑算法对视差图进行平滑。由于WTA算法所得到的视差值是整像素精度,为了获得更高的子像素精度,需要对视差值进行进一步的子像素细化,常用的子像素细化方法是一元二次曲线拟合法,通过最优视差下的代价值以及左右两个视差下的代价值拟合一条一元二次曲线,取二次曲线的极小值点所代表的视差值为子像素视差值。

在上述框架下,立体匹配算法分为为三种:局部匹配算法、全局匹配算法、半全局匹配算法。局部匹配的步骤一般包括匹配代价计算、代价聚合和视差计算三个步骤;全局算法则包括匹配代价计算,视差计算与视差优化三个步骤;半全局算法包含

所有四个步骤。全局算法由于非常高的运算量或内存消耗，在大多数场合都无法应用，而局部算法虽然速度很快，但是鲁棒性差，匹配质量比较低。半全局的立体匹配算法SGM，该算法采用单像素互信息作为匹配代价，沿着多个方向进行一维能量最小化来近似替代二维全局能量最小化，因此被称为半全局算法。SGM的运算速度远远快于大多数全局算法，同时精度也比较高。

```
close all;clear all;clc;
% 加载stereoparameters对象
load('handshakeStereoParams.mat');
% 相机可视化
showExtrinsics(stereoParams);
% 创建视频多帧文件
videoFileLeft = 'handshake_left.avi';
videoFileRight = 'handshake_right.avi';
readerLeft = VideoReader(videoFileLeft);
readerRight = VideoReader(videoFileRight);
player = vision.VideoPlayer('Position', [20, 200, 740 560]);
% 读取并校正视频帧，校正后的图像具有水平极线，并且行对齐。
frameLeft = readFrame(readerLeft);
frameRight = readFrame(readerRight);
[frameLeftRect, frameRightRect] = ...
rectifyStereoImages(frameLeft, frameRight, stereoParams);
figure;
imshow(stereoAnaglyph(frameLeftRect, frameRightRect));
title('多帧图像');
%计算视差，在校正后的立体图像中，任何一对对应点位于同一像素行上。
frameLeftGray = rgb2gray(frameLeftRect);
frameRightGray = rgb2gray(frameRightRect);
disparityMap = disparitySGM(frameLeftGray, frameRightGray);
figure;
imshow(disparityMap, [0, 64]);
title('视差图');
colormap jet
colorbar
```

8.3运动恢复结构

运动恢复结构方法(Structure from Motion, SFM)是一种计算机视觉三维重建技术，可由影像恢复出运动相机的位置、姿态以及所摄场景的三维结构，且对参与重建的影像没有约束要求，通用性较强。运动恢复结构方法实现的策略大体上可分为增量式和全局式。增量式运动恢复结构方法(ISFM)是从2张或者3张影像的重建开始，增量式的添加新的影像后再进行重建，直到完成所有影像的重建。全局式SFM采用平移矩阵和旋转矩阵分别一次性全局求解的方法，因其省去了不断重建的步骤，一定程度上解决了增量式SFM效率低下的问题。但当影像姿态、尺度或者畸变较大时，全局式SFM会出现解算不稳定的问题，甚至导致解算失败，而增量式SFM的解算结果则更加稳健。对于全局式SFM，当重建的影像很多时，一般采用分组重建的方法，其本质上也是一种增量式SFM的思想。对于无序影像的三维重建，增量式SFM是一种更受欢迎的策略方法。目前开源的SFM软件很多，其中Visual SFM、OPENMVG、MVE以及COLMAP等是目前重建效果比较好的开源SFM软件。

8.3.1增量式SFM

增量式 SFM 的原理是首先对两视图进行特征点检测和匹配，再对图像进行几何解算。其中选取一对匹配度较好的两张种子影像进行初始化重建，通常采用随机采样一致性算法剔除错误的匹配点；然后对初始化生成的相机位姿和 3D 点集进行光束法平差优化。此后，不断添加新图像，循环进行新影像注册、前方交会、外点滤波和局部光束法平差等步骤，同时每当新影像注册量达到设定的阈值或者所有影像全部注册完毕后都会进行一次全局光束法平差，增加 SFM 的鲁棒，直到所有图像添加完毕使整个场景重建完成。

图8-6 增量式SFM流程

特征提取和匹配是三维重建过程中的起点，也是至关重要的一步，特征提取的目标是提取具有高重复性、高可区分性的特征点，并生成对应的特征描述子。高重复性可以保证所提取的特征点能够在其他重叠影像中被检测到；高可区分性则保证所提取的特征点能够在其他重叠影像中被正确匹配到。如果提取错误的特征点较多或特征点匹配效果不佳，就会导致获取的摄像机参数不准确，重建出来的三维点与实际误差较大，这些不准确的三维点连接起来并不能正确地反映场景或物体的三维结构。

增量式SFM是将初始化后选取的最优图像对作为种子图像对，种子图像对的选取通常基于两个方面：一是两视图间的特征点匹配个数；二是两视图间的基线宽度。一般情况下，两视图间匹配的特征点数目越多则越有利于问题的求解，得到的结果也越精确。同时还要考虑特征点在图像中的分布情况，分布过于集中或者分布不均都会导致求解精度下降。由于绝对尺度的缺失，基线的长度无法直接衡量，实际中可以采用视图间主光轴的夹角作为衡量尺度。重建过程是一个非凸问题，增量过程的初始参数对重建的鲁棒性、精确性等表现有着重要的影响，一个坏的初始值会导致整个重建的失败。

图像配准则是利用三角定位后的三维点信息恢复下一个相机的位姿，PnP (Perspective-n-Point)是已知场景中目标点的三维信息和对应成像平面的二维信息，估计相机的位姿的求解方法。如果两张图像中的一张特征点三维位置已知，那么至少需要3个点对就可以计算相机的运动，同时再利用一对附加的点对就可以确定唯一的值。通过重复“三角定位-PnP”这一过程，逐步递进地完成了对整个系统的相机位姿与场景三维点信息的恢复。由于误差的存在，估计结果的偏差往往会随着增量式的过程逐渐增大，因此仍然需要不断地利用光束平差法对求解结果进行优化。

图8- 7稀疏重建效果展示

8.3.2光束平差法

对于目标场景中任意三维点P，从每个视图所对应的的摄像机的光心发射出来并经过图像中P对应的像素后的光线，都将交于P这一点，对于所有三维点，则形成相当多的光束；实际过程中由于噪声、误匹配、计算精度低等问题，每条光线几乎不可能汇聚与一点，导致最终得到的三维模型面片数少、模型表面不平滑、局部细节模糊甚至空洞现象的出现。

19. 计算机视觉教材_第19部分

总字符数：8925

相似文献列表

去除本人文献复制比：14.9%(1332)

去除引用文献复制比：14.9%(1332)

文字复制比：14.9%(1332)

1	<u>基于多视图的非增量式三维重建关键技术的研究与应用</u> 杨静(导师：郭向坤) - 《中国科学院大学(中国科学院沈阳计算技术研究所) 硕士论文》 - 2018-06-01	4.4% (394)	是否引证：否
2	<u>基于深度图的三维重建算法的优化研究</u> 闫冰(导师：鲁东明;刁常宇) - 《浙江大学硕士论文》 - 2019-01-01	3.7% (331)	是否引证：否
3	<u>基于激光雷达和环境特征的水平运输车辆定位系统</u> 刘春明;王全宇;欧阳乐;黄旭东; - 《集装箱化》 - 2020-08-28	2.7% (237)	是否引证：否
4	<u>三维点云的重建与匹配(修改版) 下载</u> - 《互联网文档资源 (https://ishare.iask.com/)》 - 2021	2.0% (178)	是否引证：否
5	<u>基于旋转图像特征描述子改进的ICP算法</u> 郎萍;高媛;秦品乐;王丽芳; - 《计算机工程与设计》 - 2016-10-16	1.9% (174)	是否引证：否
6	<u>基于点云数据的砼地面平整度可视化方法研究</u> 杨小燕;陈登峰;于军琪;于光;王帅举; - 《工业仪表与自动化装置》 - 2022-04-11	1.7% (154)	是否引证：否
7	<u>基于3D Zernike矩的巡检器与空间站的相对导航算法</u> 王润;郁丰;周士兵;刘方武; - 《航空学报》 - 2020-08-10 15:25	1.6% (146)	是否引证：否
8	<u>港口多AGV智能导航解决方案设计与试验</u> 马利娜;游佳朋;熊会;于均;明新国; - 《机械设计与研究》 - 2021-12-20	1.3% (118)	是否引证：否
9	<u>基于图像语义分割的药用玻璃瓶外轮廓尺寸检测方法研究</u> 于浩;张震;张雷;刘申;赵东哲;张金辉; - 《河北省科学院学报》 - 2023-02-15	1.3% (117)	是否引证：否
10	<u>无缝钢管斜轧穿孔顶头表面缺陷非接触在线检测方法</u> 于浩;黄华贵;郑加丽;赵铁琳;周新亮; - 《中国机械工程》 - 2021-09-14 15:23	1.3% (117)	是否引证：否
11	<u>面向车间动态环境的语义地图构建技术研究</u> 谭磊(导师：杜丽) - 《电子科技大学硕士论文》 - 2022-03-20	1.2% (111)	是否引证：否
12	<u>基于激光雷达+视频的城际铁路周界入侵监测报警技术研究</u> 栗文韬;徐成伟;张万鹏;傅荟瑾;马祯;杨雪; - 《铁道运输与经济》 - 2023-07-15	1.2% (110)	是否引证：否
13	<u>室内环境下移动机器人视觉SLAM前端关键技术研究</u> 张芷晴(导师：胡为) - 《沈阳航空航天大学硕士论文》 - 2019-12-12	1.1% (100)	是否引证：否
14	<u>基于有监督降维的人脸识别方法</u> 姚明海;王娜;易玉根;栾敬钊; - 《计算机工程》 - 2014-05-15	0.9% (83)	是否引证：否
15	<u>基于NDT和ICP融合的点云配准方法</u> 张桂杨;苑壮;陶刚; - 《北京测绘》 - 2019-12-25	0.9% (79)	是否引证：否
16	<u>基于NDT与ICP结合的点云配准算法</u> 王庆闪;张军;刘元盛;张鑫晨; - 《计算机工程与应用》 - 2019-08-23 1	0.7% (61)	是否引证：否
17	<u>点云配准方法在粗配准中的比较</u> 侯彬;金尚忠;王赞;陈智慧;曹馨艺; - 《激光与光电子学进展》 - 2019-09-23 1	0.6% (52)	是否引证：否
18	<u>基于图像处理的机织物组织自动识别的研究</u> 张江丰(导师：张森林) - 《浙江大学硕士论文》 - 2013-01-01	0.5% (45)	是否引证：否
19	<u>基于点云配准的盆栽金桔果实识别与计数方法</u> 朱启兵;张梦;刘振方;黄敏;李学成; - 《农业机械学报》 - 2022-03-01 16:53	0.4% (39)	是否引证：否

20	光学薄膜损伤表面三维微观形貌的仿真与重构	0.4% (36)
	张昭琳;苏俊宏; - 《光学学报》 - 2020-09-02 13:32	是否引证: 否
21	基于先验信息采样一致性的点云柱面分割方法	0.4% (34)
	张润梅;陈佳俊;尹蕾;许维义; - 《激光与红外》 - 2023-02-20	是否引证: 否
22	分两阶段变换坐标的点云粗配准算法	0.4% (34)
	李思远;刘瑾;杨海马;刘海珊; - 《激光与光电子学进展》 - 2022-08-25	是否引证: 否
23	WEID:一种基于信息量差加权集成的Android恶意软件检测方法	0.4% (34)
	张高峰;鲍旭丹;刘敬;夏雪晗;郑利平; - 《计算机应用与软件》 - 2022-09-12	是否引证: 否
24	基于多视角数字图像的深基坑三维重构研究	0.4% (34)
	黄身勇;蒋国平; - 《河南科技》 - 2021-06-05	是否引证: 否
25	基于电塔点图关联的自主巡检导航方法研究	0.3% (31)
	李梟宇(导师: 吴华) - 《华北电力大学(北京) 硕士学位论文》 - 2020-06-30	是否引证: 否

原文内容

因此在求解过程中，需要不断对待求信息进行调整，来使得最终光线能交于点。

光束平差法(Bundle Adjustment, BA)指源自每个三维特征并汇聚到相机光心的一组光束，它在空间结构和观测参数方面进行了最佳调整，是对描述场景几何信息的三维点坐标和描述相机相对姿态的参数进行同时优化的技术。本质上是一种非线性最小二乘优化方法，该方法的原理是使反投影像点到真实像点的距离最小。即估计透视投影矩阵和三维点坐标的误差越小。

当存在多个空间三维点(范围从1到m)、多个观测相机(范围从1到n)时，所有误差之和可以用目标函数(8.28)表示：

(8.28)

其中： \mathbf{x} 和 \mathbf{p} 分别表示三维点坐标和相机参数； \mathbf{f} 是指示函数； \mathbf{p}_i 是对应的观测影像点； \mathbf{x}_i 是三维点在相机条件下的投影点。

若重建计算所得的空间点坐标和相机投影模型更加准确时，得到投影的预测位置与实际位置更加吻合，重投影误差越低，目标函数的数值结果越小。所以，光束法平差就是目标函数最小化的问题，即通过迭代优化空间三维点坐标 \mathbf{x} 的值和投影函数 \mathbf{P} 中包含的相机参数的值，不断减少目标函数的值，当目标函数的数值低于某个阈值时，得到的空间点坐标和相机内外参数就是我们优化求得的结果。

光速平方差法Matlab中可以直接调用函数bundleAdjustment返回精确的三维点和相机姿势。Matlab中给定的例子使用bundleAdjustment计算相机姿态：

```
clc; clear; close all;
data = load('sfmGlobe');
[xyzRefinedPoints,refinedPoses] = ...
bundleAdjustment(data.xyzPoints,data.pointTracks,data.cameraPoses,data.intrinsics);
pcshow(xyzRefinedPoints,'VerticalAxis','y','VerticalAxisDir','down','MarkerSize',45);
set(gcf,'color',[1,1,1]);set(gca,'color',[1,1,1]);
hold on
plotCamera(refinedPoses,'Size',0.1);
hold off
grid on
```

8.3.3多视图稠密重建

由于SFM用来做重建的点是由特征匹配提供的，导致匹配点稀疏，重建效果并不是很好，所以需要稠密点云重建(MVS)进行完善。MVS的目的是对不同图像上的同一个点进行匹配，增强重建场景的稠密性。MVS的主要任务是通过一种最佳搜索的方法，匹配不同图像的同一个点。如果需要搜索参考图像上的某个像素点与另一帧图像上哪个像素点是同一个点，只需要搜索下图中的极线即可。主要方法为逐像素判断，遍历整个搜索图，寻找到与模块图最为相似的部分。

1) 深度图计算

深度图是指将相机到场景中各点距离作为像素值的图像，在MVS中对于深度图的计算是一大重点。深度图计算采用经典的PatchMatch方法，基本思想是在自然立体对中，相对较大的像素区域可以由大致相同的平面建模。通过一个随机平面初始化每个像素来找到一个区域的平面。在这个随机初始化之后，该区域的至少一个像素带有一个接近正确的平面。本身是一种近似的密集最近邻算法。PatchMatch基本思路如下：

对于两帧的每个像素 p ，我们搜索一个平面。计算左右视差图，通过左/右一致性检查执行遮挡处理。一旦找到，可以计算 p 的视差为：(8.29)

其中、 \mathbf{p} 和 \mathbf{q} 是平面的三个参数， \mathbf{p} 表示 p 的 x 和 y 坐标。希望找到的平面是所有可能平面中最小聚合匹配成本之一：

(8.30)

其中 \mathbf{F} 表示大小为无限的所有平面的集合。这个无限的标签空间使我们无法简单地检查所有可能的标签，就像在标准局部立体匹配中所做的那样，其中标签对应于离散视差值。根据平面 f 匹配像素 p 的聚合成本计算为：(8.31)

这里， \mathbf{w} 表示以像素 p 为中心的方形窗口。在这种情况下，不再是2D，而是3D窗口，其中第三维由视频序列的前一帧和连续帧的像素形成。权重函数 $w(p,q)$ 用于克服边缘肥大问题并实现自适应支持权重思想，它通过查看像素的颜色来计算 p 和 q 位于同一平面上的可能性，即，如果颜色相似，它会返回高值：(8.32)

其中 γ 是用户定义的参数，计算RGB空间中 p 和 q 颜色的距离。

首先根据平面 f 计算 q 的视差, 并通过从 q 的 x 坐标中减去该视差得出 q 在另一个视图 q' 中的匹配点。函数 $\rho(q, q')$ 是计算 q 和 q' 之间的像素相异度 (8.33)

其中表示在 q 和 q' 处计算的灰度值梯度的绝对差。由于 q' 的 x 坐标位于连续域中, 通过线性插值推导出其颜色和梯度值。用户定义的参数 α 平衡了颜色和渐变项的影响。参数和截断遮挡区域鲁棒性的成本。

2) 纹理映射

纹理映射也称纹理贴图, 就是把二维图像的像素值映射到三维模型的对应顶点上, 即实现二维纹理到三维模型表面的映射, 给三维模型赋予颜色、几何等纹理的过程。纹理映射有两种方法: 交互式方法和自动式方法, 其中交互式方法需要人工挑选适合给一个平面贴纹理的图像, 消耗大量的时间, 而自动式方法可以直接利用点云网格化后生成的三维模型以及前期工作中产生的相关信息就可以自动对三维模型贴纹理。

由于点云网格化后得到的是能描述物体表面的三维网格模型, 当把模型中三角面片的三个顶点, 投影到同一个二维图像上时, 就可以直接提取投影区域在图像上的纹理信息, 这样纹理就生成了并贴到面片上, 直到所有的三角网格遍历完成, 纹理贴图就结束了。如果三角面片的两个顶点投影到一个图像A, 另一个顶点投影到图像B中, 则在图像A中提取纹理信息; 那么当三角面片的三个顶点投影到三个不同的图像, 此时这个三角面片就没有意义, 继续寻找下一个面片。

8.4 点云处理

图像三维重建技术能够还原物体或场景的三维结构, 它的输出结果是三维几何结构和对应的纹理信息。三维几何模型通常以点云、网格或体素表示, 见图8-8, 其中点云是最基本的表示形式, 而网格和体素是通过对点云数据进行处理而生成的更加复杂的表示形式。点云由一组离散的三维点组成, 每个点都包含了空间中的位置信息。三维网格是由一系列连接的三角形面片或四边形面片组成的表面模型, 它可以通过对点云进行三角化而得到; 网格模型更加密集, 可以更准确地表示物体的表面形状, 常用于图形渲染和可视化。三维体素模型是一种将三维空间划分为小的立方体单元 (称为体素) 的体积表示方法。类似于三维像素, 体素可以看作是三维空间的离散化表示, 它将连续的空间划分为离散的小立方体, 每个体素可以表示空间中的属性信息, 如是否占据、颜色、密度等。从点云生成体素模型的过程称为体素化处理。体素模型常用于医学图像中特定区域体积分析, 如计算心脏的体积来评估心脏功能, 也可以用于生成三维打印模型。

图8-8 三维模型的点云、体素、网格表示方式

本节将介绍点云数据预处理、点云配准和曲面重建方法。点云数据预处理是点云处理的第一步, 通过去噪、精简、法线估计、平面提取等操作, 优化原始点云数据的质量和结构, 以便后续的分析 and 处理。点云配准是将不同来源或不同时刻获取的点云数据对齐到一个共同的坐标系中。这个步骤对于融合或比较多个点云数据非常重要, 如多视角重建的图像点云与激光雷达点云的融合。曲面重建则是从离散的点云数据中恢复连续曲面, 这对于理解结构场景、进行建模和可视化等应用非常关键。本节内容主要介绍了点云处理中的基础内容, 同学们想了解更多可参考《点云智能处理》。

8.4.1 点云数据预处理

点云数据预处理是对原始点云数据进行一系列操作, 以提高数据质量、去除噪声、减少数据量等, 从而使得点云数据更易于分析和处理。点云数据预处理的具体步骤和方法取决于应用场景和数据特点, 点云预处理对于后续的点云配准、曲面重建、对象识别、分割等任务有着重要影响。本节将介绍去噪、精简、法线估计、平面提取等内容。

(1) 点云去噪

在获取结构点云数据时, 由于设备本身精度、环境影响和人员操作等的影响, 不可避免的出现噪点和离群点。这些噪声点会影响后续的处理和重建结果。去噪的目标是尽可能地消除这些噪声, 提高点云数据的质量, 使得后续处理和分析更加准确和可靠。

常见的点云去噪方法有统计滤波、半径滤波、高斯滤波、机器学习等。统计滤波是通通过点云局部点云的距离来过滤离群点, 半径滤波则是通过局部点云的密度来进行判断。

统计滤波算法 (Statistical Outlier Removal) 主要用于去除稀疏离群的噪声点, 即那些偏离主体点云或分布较为稀疏的点。对于点云中任一点 P , 定义一个局部邻域, 通常是设置一个固定的搜索半径内的 n 个最近邻点, 依次计算出点 P 与其邻域中第 i 个点的距离。假设点距离满足高斯分布, 则邻域内所有点的平均距离 μ 和标准差 σ 如下:

根据经验设置一个标准差倍数 k , 将邻域点云的距离阈值范围定义为 $(\mu - k \cdot \sigma, \mu + k \cdot \sigma)$ 。范围之外的点就视为离群点, 从原始点云中删除。该算法中的参数是邻域搜索点数量 n 、标准差倍数 k , 前者控制邻域大小, 后者控制筛选范围。统计滤波算法主要用于去除稀疏离群的噪声点, 这些离群点可能 远离主体点云, 或者分布在主体点云周围, 而主体点云则分布相对集中、密度较高。

半径滤波算法 (Radius Outlier Removal) 是根据空间点云分布来进行滤波, 分离原则是主体点云集中, 离群点云散乱。该算法主要通过设置一个固定的搜索半径来计算每个点的局部邻域内点的数量或密度, 从而判断哪些点属于稠密区, 哪些点属于噪声区。半径滤波方法的效果和性能受到搜索半径 R 和密度阈值 k 的影响, 指定半径邻域中至少应有 k 个点, 否则视为离散点, 将其删除。如图8-9所示, 对点 P 做半径 R 的圆, 设定阈值 $k=3$, 以点 P 为圆心的圆内有4个点, 判断为主体点云; 以点 P 为圆心的圆内有2个点, 视为离群点, 在滤波过程中进行删除。半径滤波方法适用于处理一些局部密度变化较大的点云数据, 尤其对于存在较均匀分布的噪声和局部密度变化的情况下, 具有良好的去噪效果。

图8-9 半径滤波示意图

(2) 点云精简

我们通过图像重建或激光扫描等途径得到的点云可能过于庞大, 不利于存储、传输和处理。点云精简 (Point Cloud Downsampling) 通过减少点云数据中的点数来降低数据量和复杂性的过程。点云精简的目的是减少 点云数量, 同时保持整体形状的特征。常见的点云精简方法有随机采样、体素滤波、基于形状特征的采样等。

随机采样不考虑点的分布、密度或形状等特征, 仅仅根据随机性对点进行采样。预设一个采样率 (如0.1), 表示想要保留点云数量比例。通过随机数生成器在原始点云数据中随机选择一些点, 以满足采样率的要求。将抽样得到的点组成新的点云数据, 作为精简后的点云数据。需要注意的是, 随机采样可能会导致采样后的点云数据丢失一些重要的细节特征, 通常会用来对点云数据进行初步处理或者快速预览时使用。

体素网格滤波 (Voxel Grid Filtering), 是将点云数据划分为规则的三维体素网格, 每个体素格内选择一个代表性的点

作为采样点，从而减少点云数据的数量。首先需要设定一个体素的大小，即规定体素格的边长。体素大小决定了体素格的密度，较小的体素会产生较密集的体素格，较大的体素会产生较稀疏的体素格。将点云数据中的所有点按照其空间坐标划分到相应的体素格中。每个点会被分配到最近的体素格内。在每个体素格内选择一个代表性的点作为采样点，可以选择体素格内所有点的中心点、重心或平均值等。将每个体素格内选择的代表性点组成新的点云数据，作为精简后的点云。体素滤波可以用于点云数据的分辨率，从而减少计算和存储资源的需求。由于体素滤波会对点云数据进行规则划分，可能会忽略掉一些细微的结构特征。

(3) 点云平面提取

点云平面提取是点云数据处理中的关键任务之一，其目标是从点云中识别和提取出表示平坦表面的点集合，可以应用于地面提取、墙面提取、物体分割等应用中。

主成分分析 (Principal Component Analysis, PCA) 是点云平面提取的一种常用的方法，用于找到点云数据中的主要方向，从而拟合出一个平面。PCA的基本思想是通过将数据降维，将原始数据从高维空间映射到低维空间，使得新的坐标轴（主成分）对应着数据的主要变化方向。给定空间拟合平面方程为

$$Ax+By+Cz=D$$

其中， $n=(A,B,C)$ 为平面的单位法向量，为坐标原点到平面的距离， D 为坐标原点到平面的目标。点云数据记为，将其构造为 $3 \times N$ 的坐标矩阵，计算这个矩阵的协方差矩阵 M 。对协方差矩阵 M 进行特征值分解，得到特征值和对应的特征向量。取最小特征值对应的特征向量 v 作为拟合平面的法向量 n ，得到拟合平面参数 $A、B、C$ 。利用点云坐标的均值，求解出参数 D ：

PCA是一种简单而有效的方法，特别适用于那些数据点分布比较规整，且包含明显平面结构的点云数据。然而，在存在噪声或离群点的情况下，PCA的效果可能会受到影响，此时可以考虑使用其他更鲁棒的方法，如随机采样一致性 (Random Sample Consensus, RANSAC)。

RANSAC方法中，首先我们从点云数据集中随机选择一小部分点，视为样本集。假设这些点属于平面，用PCA或最小二乘法来拟合平面模型，计算出平面参数 $A、B、C、D$ 。对于样本集外的点，计算它们到拟合平面的距离，并将距离小于预设内点阈值的点标记为内点，并统计样本内的内点数量。如果内点数量超过预设内点数量阈值，则认为模型是一个好的拟合，即平面模型已经收敛。如果模型没有收敛，重复以上步骤，重新随机采样一组新的样本点，并进行内点检验和收敛判断，直到找到一个满足收敛条件的平面模型。当满足收敛条件时，使用所有内点重新拟合平面模型，并得到最终的平面方程参数。RANSAC算法的优势在于它的鲁棒性，可以有效地从包含离群点和噪声的点云数据中提取出平面模型。

(4) 点云法线估计

法线向量是点云表面的一个重要几何属性，它表示垂直于表面的矢量，用于描述点云表面的朝向和几何形状。预处理中常常估计每个点的法线向量，用于表面重建、特征提取、配准、物体识别等后续分析中。

最近邻法线估计 (Nearest Neighbor Normal Estimation) 是一种简单而直接的方法。该方法假设在局部区域内，点云数据的法线方向与局部表面的法线方向一致。首先设置一个最近邻数目 K ，用于确定计算每个点的法线方向时所考虑的邻近点的数量。通常选择一个较小的正整数，如5、10。对于每个点 P ，找到其在点云数据中最近的 K 个邻域点，形成点 P 的局部邻域。采用最小二乘法或主成分分析，对邻域点进行平面拟合，获得一个拟合平面，其法线方向即为点 P 的法线方向。

最近邻法线估计是一种计算简单、效率高的法线估计方法，特别适用于处理较大规模的点云数据。然而，它对于点云密度和噪声较为敏感，在某些情况下可能产生较大的法线估计误差。对于一些对法线准确性要求较高的应用场景，可能需要考虑其他更复杂的法线估计方法，例如加权最小二乘法、法线累加法等。这些方法能够更好地处理点云数据中的噪声和密度变化，提供更精确和稳健的法线估计结果。

8.4.2 点云配准

在获取大型结构的三维几何模型时，我们通常会围绕着目标物开展多站点扫描，从不同视角获取多组点云，以便得到更全面的场景模型。由于不同视角下获取的点云数据处于不同的坐标系中，我们将多组点云合并为一个整体模型，意味着需要将它们对齐在同一个参考坐标系下，这个过程称为点云配准。

点云配准的目标是计算点云之间的几何变换参数，如平移、旋转，并将点云转换到共同的映射坐标系中。点云配准根据变换参数不同，可分别刚性配准和非刚性配准。刚性配准只考虑点云的平移和旋转，而非刚性配准则对应着仿射变换，包含了点云的平移、旋转、缩放和剪切。非刚性配准适用于复杂场景、形状变化较大或动态模型等，如材料表面变形分析、动态物体追踪、地形重建等。对于土木基础设施或结构施工场景，它们的几何形状不同视角下基本保持不变，可以采用刚性配准。对于源点云中的每个点，经过刚性配准后得到对应的目标点云中的点，转换公式如下：

其中 R 是旋转矩阵， t 是平移向量。点云配准的目标是找到最优的变换参数，使得两组点云在某种度量标准下对齐得最好，如点云中对应点之间的欧氏距离最小化、对应点的相关性最大化。点云配准任务通常会分为粗配准 (Coarse Registration)、精配准 (Fine Registration) 两个阶段。粗配准为后续精细配准提供一个良好的初始估计，改善优化过程，精细配准侧重于对点云中的细节和局部特征进行更准确的匹配，以获得更好的全局对齐效果。

根据工作原理，点云配准主要可以分为五类：

基于特征的方法：在点云中提取具有代表性的特征，并利用这些特征进行匹配和对齐，实现不同点云之间的配准。这些特征可以是鲁棒的局部区域描述，如法线、表面曲率、颜色，也可以描述整个点云数据集的属性，如点云形状直方图。

基于迭代最近点 (Iterative Closest Point, ICP) 的方法：ICP及其变体是最经典的点云配准方法之一，它通过迭代算法，将源点云中的点映射到目标点云中，然后通过最小化对应点之间的距离来估计刚性变换参数，以实现点云的对齐。ICP方法对初始输入的变换参数非常敏感，用于精配准阶段。

基于学习的方法：通过机器学习技术生成不变特征，在两个任意点云之间提供相对更加鲁棒的变换。对比人为设计的特征描述，机器学习的特征配准具有更强的自适应性和鲁棒性。

基于概率的方法：将点云配准视为一个概率估计问题，通过最大化或优化概率函数来寻找最优的配准变换参数，以使得两个点云之间的对应关系最优。这种方法可以处理不确定性和噪声，并在一定程度上提高配准的鲁棒性。

其他方法：针对特定问题或场景，会设计一些定制性的方法，比如室内环境中，可以利用墙面、地面的法向量进行快速的配准。

下面将对快速点特征直方图（FPFH）和ICP两类点云配准方法的基本原理展开介绍，它们分别可以用于粗配准、精配准两阶段。

FPFH点云配准是一种基于特征描述符的方法，它考虑了每个点与邻域点法向量之间的关系。对于每个点，搜索距离 r 范围内的 k 个邻域点，计算出该点与邻域点法向量（、）之间的几何关系。我们首先在点处定义一个局部坐标系 uvw ，

点、点对应法向量和之间的差异可以表示为一组角度特征，如下所示：

图8- 10 FPFH算法中点局部坐标系 uvw 和邻域点示意

接着，我们求解出点所有邻域点的特征三元组 α 、 ϕ 、 θ ，创建出点的简化点特征直方图(Simplified Point Feature Histogram, SPFH)。划分过程，将每个特征的范围分成 b 个子区间，并计算每个子区间中的特征值出现次数。由于三个特征都是法线之间角度的度量，它们的值可以标准化到相同区间划分。SPFH直方图中，每个区间表示法向量之间的角度差，直方值表示具有相似法向量关系的点的数量。

FPFH描述符由两部分组成，点的SPFH和 k 个邻域点距离加权的SPFH，

其中权重表示查询点与邻域点在某个给定的度量空间中的距离，如。距离权重使得距离较远的邻域点对直方图的贡献较小，而距离较近的邻域点对直方图的贡献较大。

基于FPFH特征点云配准流程见图8- 11。首先，计算出待配准点云和参考点云的FPFH特征描述符，可以获得两个点云中各个点之间的相似性度量；采用使用邻搜索或KD树等方法，找到具有相似FPFH描述符的点来进行点云匹配。针对特征匹配后的点对，利用采样一致性粗配准（Sample Consensus Initial Alignment, SAC-IA）来进行点云快速配准。从待配准点云中随机选择一定数量的点对，通过奇异值分解（SVD）计算初始刚体变换（如平移和旋转）。采用均方误差或平均点对距离作为误差函数，来衡量点云的对齐程度。在重复迭代的过程中，不断优化刚体变换参数，使得误差函数值逐渐减小。迭代过程会持续进行直到满足收敛条件或达到最大迭代次数。SAC-IA算法可以在相对较短的时间内得到点云的粗略配准结果，为后续更精确的配准算法（如ICP）提供一个良好的初始猜测。这样可以加速点云配准的过程，并提高配准的准确性和鲁棒性。

图8- 11 基于FPFH特征的点云配准流程

迭代最近点法（Iterative Closest Point, ICP）是最为经典的点云配准算法。采用最小二乘估计计算变换矩阵，原理简单且具有较好的精度。其特征在于，按照一定的约束条件求取源点云 Q 和目标点云 P 之间的对应点对，基于对应点对构造旋转平移矩阵，并利用所求矩阵，将源点云变换到目标点云的坐标系下，估计变换后源点云与目标点云的误差函数，若误差函数值大于阈值，则继续迭代上述运算直到满足给定的误差要求。其误差函数为

$E(R, t)$ (8.39)

其中 n 为最邻近点对的个数，为目标点云 P 中的一点，为源点云 Q 中与对应的最近点， R 为旋转矩阵， t 为平移向量。

迭代最近点方法的具体步骤如下：

将点云数据进行降噪、抽稀处理；

计算目标点云 P 中点集与在源点云 Q 中点集使得；

计算使上述对应点对平均距离最小的刚体变换，求得旋转矩阵 R 和平移矩阵 t ；

对使用求得的旋转矩阵 R 和平移矩阵 t 进行旋转和平移，得到新的点集；

计算与对应点集的平均距离：

；

如果 d 小于某一给定的阈值或者大于预设的最大迭代次数，则停止迭代计算。否则返回第3步，直到满足收敛条件为止。

20. 计算机视觉教材_第20部分		总字符数：4450
相似文献列表		
去除本人文献复制比：45.9%(2042)		去除引用文献复制比：45.9%(2042) 文字复制比：45.9%(2042)
1	<u>基于多视图的非增量式三维重建关键技术研究与应用</u> 杨静(导师：郭向坤) - 《中国科学院大学(中国科学院沈阳计算技术研究所) 硕士学位论文》- 2018-06-01	19.8% (881) 是否引证：否
2	<u>基于三维激光扫描技术的结构检测监测应用研究</u> 骆义(导师：张建) - 《东南大学 硕士学位论文》- 2021-05-28	16.0% (710) 是否引证：否
3	<u>激光非视域成像场景仿真与三维重建技术研究</u> 饒京京(导师：苏秀琴) - 《中国科学院大学(中国科学院西安光学精密机械研究所) 硕士学位论文》- 2018-06-01	9.1% (403) 是否引证：否
4	<u>基于Kinect的虚拟现实应用中关键问题研究</u> 陈杰(导师：梁秀霞) - 《山东财经大学 硕士学位论文》- 2017-05-01	9.1% (403) 是否引证：否
5	<u>基于点云数据的露天矿山测量验收技术研究及应用</u> 于秋宇(导师：刘光伟) - 《辽宁工程技术大学 硕士学位论文》- 2022-05-27	7.5% (335) 是否引证：否
6	<u>弱纹理物体表面鲁棒三维重建方法研究</u> 张思远(导师：乔玉晶) - 《哈尔滨理工大学 硕士学位论文》- 2020-04-01	6.9% (307) 是否引证：否
7	<u>面向机器人智能抓取的目标三维重建方法研究</u> 刘永俊(导师：华宇宁;刘意杨) - 《沈阳理工大学 硕士学位论文》- 2022-03-14	6.3% (282) 是否引证：否

8	<u>面向数字化保护的自动文物三维重建方法研究</u> 吴宁;陈佳舟;吴凯乐; - 《山西建筑》 - 2018-02-20	5.8% (259)	是否引证: 否
9	<u>基于网格模型的三维人体自动化测量方法</u> 于英;陈继华;卢学良; - 《图学学报》 - 2018-04-15	5.8% (259)	是否引证: 否
10	<u>地质景观三维重建关键技术研究</u> 王天乐(导师: 陈建平) - 《中国地质大学(北京)硕士论文》 - 2019-05-01	5.8% (259)	是否引证: 否
11	<u>基于网格模型的三维人体自动化测量方法</u> - 道客巴巴 - 《互联网文档资源 (http://www.doc88.com)》 - 2019	5.8% (259)	是否引证: 否
12	<u>基于非平行轮廓线的地质体表面重建方法研究</u> 王楚天(导师: 李国和) - 《中国石油大学(北京)硕士论文》 - 2017-08-01	5.1% (225)	是否引证: 否
13	<u>全息体视图中基于三维重建的虚实场景融合显示方法</u> 刘云鹏;刘新蕾;闫兴鹏; - 《科技传播》 - 2021-10-25	4.9% (216)	是否引证: 否
14	<u>基于泊松方程实现点云的表面重构</u> 刘金玲;唐棣; - 《计算机应用与软件》 - 2009-04-15	4.8% (214)	是否引证: 否
15	<u>基于三维点云的棉花冠层三维重构研究</u> 李春花(导师: 乔红波) - 《河南农业大学硕士论文》 - 2022-05-01	4.4% (196)	是否引证: 否
16	<u>基于多视角图像序列的三维人体重建以及骨骼提取研究</u> 沈伯伟(导师: 尹宝才) - 《北京工业大学硕士论文》 - 2017-06-08	4.0% (179)	是否引证: 否
17	<u>基于图像编辑的三维发型生成技术研究</u> 韩晓迪(导师: 张菁) - 《济南大学硕士论文》 - 2022-05-01	2.8% (124)	是否引证: 否
18	<u>基于不确定性模型与重定位技术的语义SLAM方法研究</u> 侯政华(导师: 高云峰) - 《哈尔滨工业大学硕士论文》 - 2018-06-01	2.6% (114)	是否引证: 否
19	<u>基于计算机视觉的马铃薯三维表型研究</u> 卫飞帆(导师: 李富忠) - 《山西农业大学硕士论文》 - 2020-06-01	2.4% (107)	是否引证: 否
20	<u>果树重建与果实识别方法在采摘场景中的应用</u> 熊龙焯;王卓;何宇;刘洒;杨长辉; - 《传感器与微系统》 - 2019-08-09	2.4% (105)	是否引证: 否
21	<u>利用powercrust算法对点云3D重建方法</u> 李云;郑江华;王冠生; - 《黑龙江科技信息》 - 2013-05-25	2% (89)	是否引证: 否
22	<u>逐点插入法在三维地质建模中的运用</u> 王小勇;曹谢东;魏存档;黄宏亮;曹诗咏; - 《信息技术》 - 2011-09-25	1.8% (82)	是否引证: 否
23	<u>船体曲面变形中支撑半径的确定方法及应用研究</u> 程细得;沈通;冯佰威;刘祖源; - 《中国造船》 - 2016-09-30	1.3% (56)	是否引证: 否
24	<u>基于Delaunay三角网的任意多边形三角剖分算法研究</u> 刘刚;袁纪武;李磊;王春; - 《计算机与数字工程》 - 2012-06-20	1.1% (48)	是否引证: 否
25	<u>基于Delaunay三角剖分的测头半径补偿算法</u> 赵小军; - 《制造业自动化》 - 2011-07-25	1.1% (48)	是否引证: 否
26	<u>利用Delaunary三角网优化CORS网构的方法研究</u> 台淼;王佩贤;姜东; - 《测绘科学》 - 2010-12-30 1	1.1% (48)	是否引证: 否
27	<u>SYT勘探中基于VTK的三维地质建模技术</u> 王西汉; - 《电脑知识与技术》 - 2012-02-05	1.1% (48)	是否引证: 否

原文内容

```
clear; clc; close all;
%%
mat_file = './datamat/gongqiao.mat'; % 加载数据
load(mat_file);
% xyz coordinates to point cloud
ptCloud = pointCloud(xyz,'Intensity',refl);%三维坐标加载为点云数据
```

```

ptCloud_down = pcdownsample(ptCloud,'random',0.0001);%点云抽稀便于计算
% plot
figure(1);clf;hold all;
pcshow( ptCloud_down);
xlabel(' X');ylabel(' Y');zlabel(' Z');
set(gcf,'color',[1,1,1]);set(gca,'color',[1,1,1]);
view([-10,40]);axis off;
%% 生成两部分点云
theta = 30;
rot = [ cosd(theta) sind(theta) 0; ...
-sind(theta) cosd(theta) 0; ...
0 0 1];
trans = [0 0 2];
tform = rigid3d(rot,trans);
ptCloud2=pctransform(ptCloud_down,tform);% 用以配准计算
ptCloud3=pctransform(ptCloud,tform);% 用以绘制
figure(2);clf;hold all;
pcshowpair(ptCloud_down,ptCloud2);view(0,90);
set(gcf,'color',[1,1,1]);set(gca,'color',[1,1,1]);
axis off;
%% 迭代最近点配准
thetal = -25;
rot1 = [ cosd(thetal) sind(thetal) 0; ...
-sind(thetal) cosd(thetal) 0; ...
0 0 1];
trans1 = [0 0 -1];
tform1 = rigid3d(rot1,trans1); %给予一定的初始位姿
[tf1,movingReg,rmse] = pcregistericp(ptCloud2,ptCloud_down,'InitialTransform',tform1); %迭代最近点
ptCloud4=pctransform(ptCloud3,tf1);
figure(3);clf;hold all;
pcshowpair(ptCloud,ptCloud4);view(0,90);
set(gcf,'color',[1,1,1]);set(gca,'color',[1,1,1]);
axis off;

```

图8- 12 旋转之前的目标点云与源点云

以上例子是对点云配准概念的理解，对于实际工程需要扫描多个站点，将多个站点配准在同一坐标系下。实际工程的点云配准不能简单的通过迭代最近点简单实现自动配准，实际工程扫描获得的点云数据重合部分不完整，需要先进行点云粗配准，粗配准会借助靶标或者选取特征点进行初次的配准，再利用算法进行精配准调整。精配准一般使用迭代最近点或其变体，以达到匹配效果，以下是对某一拱桥进行的配准后的结果展示。

图8- 13 配准后的桥梁点云三维模型

8. 4. 3网格重建

稠密重建得到的还是点云，点与点之间会产生空洞，不能反映出真实场景的表面细节，需要建立带有几何信息的网格模型，来近似模拟场景的表面。之后再把二维图像中的纹理信息映射到网格模型上，以增强三维模型的真实感。

网格化就是通过平面三角剖分或空间三角剖分把点云以三角形的方式连接起来形成三维拓扑结构，以还原模型的表面几何结构，如图8- 14所示。目前可以将网格化方法分为三类。①插值法：以Delaunay三角化算法为基础，对点云进行Delaunay四面体剖分，并根据Delaunay三角化规则搜索相关三角网格，直到完成整个拓扑搜索；②逼近法：根据特定规则对点云进行划分，对划分的点云区域采用多个平面近似模拟点之间的拓扑关系，需要定义一个全局调整的隐式函数，在算法过程中不断进行优化，直到整个拓扑结构完成为止，例如泊松表面重建算法；③区域增长法：以某个种子三角形为基础，然后按照某种规则扩展找到新的点，然后新点与已有的边生成新的三角形，直到遍历完所有的点。例如Power Crust算法。

图8- 14 网格化

(1) 基于Delaunay算法的三角网格建立

三维激光扫描技术获取的是离散点云数据，这些点之间是没有拓扑关系的。为了重建三维物体表面，我们需要建立邻近点之间正确的拓扑连接关系，来展示散乱数据点所表示的物体表面三维形状和拓扑结构。

Delaunay三角剖分为给定平面一组离散点，按照空圆特性、最大化最小角特性两个重要的准则构造三角剖分。其中空圆特性：在Delaunay三角形网中任一三角形的外接圆范围内不会有其它点存在，在三维情形中构造四面体剖分，使得每个四面体的外接球不包含第五个点；最大化最小角特性：在散点集可能形成的三角剖分中，Delaunay三角剖分所形成的三角形的最小角最大。此算法尽量避免出现“极瘦”的三角形，从而提高数值模拟的稳定性。

采用基于Delaunay三角生成算法主要分为两步：首先生成一个包括所有离散数据点的凸壳；再利用该凸壳生成初始的三角网。很多三维模型都是通过该算法生成，然后每个点再加上高程得到三维模型。

(2) 泊松曲面重建算法

泊松重建核心思想是点云代表了物体表面的位置，其法向量代表了内外的方向。通过拟合一个由物体派生的指示函数，可以给出一个平滑的物体表面的估计。该算法基本思路是由梯度关系得到采样点和指示函数的积分关系，根据积分关系利用划分

块的方法获得点集的向量场，计算指示函数梯度场的逼近，构成泊松方程。根据泊松方程使用矩阵迭代求出近似解，采用移动立方体算法提取等值面，对所测数据点集重构出被测物体的模型。泊松方程在边界处的误差为零，因此得到的模型不存在假的表面框。直接计算梯度场会引起向量场在表面边缘的无穷大值，因此首先用平滑滤波卷积指示函数，然后求平滑函数的梯度场。

给定一个区域M及其边界，指示函数定义为：

(8.40)

这样，把重构的问题转换为重建的问题。将点云及其法向量和联系起来的公式，图2-13非常形象地描述了这二者的联系。这个指示函数是个分段函数，定义模型内部的值大于0，外部的值小于0，而为0的部分即为等值面，提取出来就是几何目标模型的表面。

泊松表面重建流程：①构建八叉树：采用的是自适应的空间网格划分的方法（根据点云的密度调整网格的深度），根据采样点集的位置定义八叉树，然后细分八叉树使每个采样点都落在深度为D的叶节点；②设置函数空间：对八叉树的每个节点设置空间函数F，所有节点函数F的线性项和可以表示向量场V，基函数F采用了盒滤波的n维卷积；③创建向量场：均匀采样的情况下，假设划分的块是常量，通过向量场V逼近指示函数的梯度。采用三次样条插值（三线插值）；④求解泊松方程：方程的解采用拉普拉斯矩阵迭代求出；⑤提取等值面：为得到重构表面，需要选择阈值获得等值面；先估计采样点的位置，然后用其平均值进行等值面提取，然后用移动立方体算法得到等值面。

(3) Power Crust算法

Power Crust算法主要利用中心轴变换的思想，通过构建物体的骨架来生成物体的表面，即利用采样点围成的区域和中心轴进行分段的线性估计，其中中心轴是由一组离散的极点来近似模拟。给定的一组极点，生成对应的Power图，并标记出以极点为中心生成的极点球是在边界的里面还是外面，而分离这两部分极点球的Power图围成的平面（内外部边界形成的区域）就是最终的结果。该算法具有很强的理论支持，对任意散乱点云输入，它都能获取致密的表面，具有相当强的鲁棒性。该算法对具有锐利边缘的物体，采样密度不均匀的散乱点集以及高噪声散乱点集也都有较好的处理效果。其过程可用图8-16描述，具体步骤如下：

1、对点云进行采样形成样点集S，然后对S进行Delaunay四面体剖分，并连接S中所有最相邻的两点，所有由这些连线的中垂线构成的连续多边形就形成该样点集的Voronoi图。

2、计算每个采样点的极点。极点就是所有包含采样点的Voronoi单元中离采样点最远的Voronoi顶点，如果Voronoi单元不是闭合的，极点表示无穷远点，一般极点分布在中心轴附近。

3、计算极点的Power图。Power图是一种加权Voronoi图，是由与极点的Power距离最小的所有空间点组成的多边形。

4、在Power图中标记出极点是在边界的里面还是外面，找出分离内外极点的边界组成的平面，生成近似曲面即为场景的表面。

图8-15 Power Crust算法过程

8.5 小结

本章介绍了立体视觉与三维重建的内容。首先，概述了主动视觉测量和被动视觉测量两种方法，包括结构光法、飞行时间法、激光扫描法以及相机与特定算法获取立体空间信息。接着，详细介绍了双目立体视觉方法，包括测量原理与数学模型、极线几何和双目立体匹配，利用左右两个相机计算视差以获得深度信息，实现三维重建。随后，讨论了运动恢复结构方法，包括增量式SFM、光束平差法和多视图稠密重建，利用多个视图的图像序列、匹配算法和三角测量原理恢复场景的三维结构和相机的运动轨迹。最后，介绍了点云处理的相关内容，包括点云降噪方法（统计滤波算法、半径滤波算法）、点云配准方法（迭代最近点法）以及网格重建方法（基于Delaunay算法的三角网格建立、泊松曲面重建算法、Power Crust算法）。

总的来说，本章提供了立体视觉与三维重建的基础知识和方法，包括双目立体视觉和运动恢复结构方法，以及点云处理的相关技术。这些内容为读者进一步探索和应用立体视觉与三维重建提供了基础和指导。

思考题

1. 列举两种主动视觉测量方法和两种被动视觉测量方法。
2. 解释双目立体视觉方法中的视差是如何用于获取深度信息的。
3. 什么是极线几何？它在双目立体视觉中的作用是什么？
4. 简要描述增量式SFM、光束平差法和多视图稠密重建这三种运动恢复结构方法的原理或步骤。
5. 点云处理中的统计滤波算法和半径滤波算法分别是如何对点云进行降噪的？
6. 什么是点云配准？简要说明迭代最近点法在点云配准中的作用。
7. 网格重建方法中的基于Delaunay算法的三角网格建立和泊松曲面重建算法分别是如何工作的？

参考文献

- [1] 左超，张晓磊，胡岩，等. 3D真的来了吗？——三维结构光传感器漫谈[J]. 红外与激光工程，2020，49(3):45.
- [2] 刘志海，代振锐，田绍鲁，等. 非接触式三维重建技术综述[J]. 科学技术与工程，2022，22(23)：9897-9908.
- [3] Usman K, Amanullah Y, Muhammad A, et al. A Methodological Review of 3D Reconstruction Techniques in Tomographic Imaging[J]. Journal of Medical Systems, 2018, 42(10):190.
- [4] Tat J, Crawford J, Chong J, et al. Three-Dimensional (3D) Animation and Calculation for the Assessment of Engaging Hill-Sachs Lesions With Computed Tomography 3D Reconstruction[J]. Arthroscopy Sports Medicine and Rehabilitation, 2021.
- [5] Qu, Yufu, Huang, et al. Sensors, Vol. 18, Pages 225: Rapid 3D Reconstruction for Image Sequence Acquired from UAV Camera. 2018.
- [6] Wei Y, Ding Z, Huang H, et al. A non-contact measurement method of ship block using image-based 3D reconstruction technology[J]. Ocean Engineering, 2019, 178(APR.15):463-475.
- [7] Barone S, Neri P, Paoli A, et al. 3D acquisition and stereo-camera calibration by active devices: A unique structured light encoding framework[J]. Optics and Lasers in Engineering, 2020, 127(Apr.):105989.1-105989.12.

[8] Oggier T , Lehmann M , Kaufmann R , et al. An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger)[J]. International Society for Optics and Photonics, 2004.

[9] 卢荣胜, 史艳琼, 胡海兵. 机器人视觉三维成像技术综述[J]. 激光与光电子学进展, 2020, 57(4):19.

[10] 佟帅, 徐晓刚, 易成涛, 等. 基于视觉的三维重建技术综述[J]. 计算机应用研究, 2011(07):17-23.

[11] Snavely N , Seitz S M , Szeliski R . Photo tourism: Exploring photo collections in 3D[J]. ACM Transactions on Graphics (TOG), 2006, 25(3):págs. 835-846.

[12] Zhu S , Zhang R , Lei Z , et al. Very Large-Scale Global SfM by Distributed Motion Averaging[C]// IEEE/CVF Conference on Computer Vision & Pattern Recognition. IEEE, 2018.

21. 计算机视觉教材_第21部分

总字数: 2955

相似文献列表

去除本人文献复制比: 56.7%(1675)		去除引用文献复制比: 56.7%(1675)	文字复制比: 56.7%(1675)
1	基于图像匹配技术的桥梁位移无标记测量方法研究 汪威(导师: 吴刚;吴栋) - 《东南大学硕士论文》 - 2021-05-30	49.0% (1448)	是否引证: 否
2	基于相机扰动校正的桥梁结构变形测量方法与应用 于姗姗(导师: 张建) - 《东南大学博士论文》 - 2021-04-01	7.4% (218)	是否引证: 否

原文内容

第9章基于计算机视觉的结构位移监测应用

9.1 泗安塘大桥位移监测应用

9.1.1 测试方案

泗安塘大桥位于G318沪聂线长兴县境内，建成于2011年。桥梁全长477.4m，跨径为5×30m+(45+80+45)m+5×30m，桥面总宽11.8m，其中行车道宽10.8m。桥面采用沥青混凝土铺装。上部结构主桥采用变截面预应力混凝土连续箱梁，跨径布置为(45+80+45)m，为单箱单室箱形截面。引桥采用30m预应力混凝土小箱梁，先简支后结构连续。

本次现场测量采用构建S-F位移测量系统进行桥梁位移测量，采用长布道MP9005S镜头，焦距为90mm。如表11-1所示，测点为主桥中跨跨中、四分之一跨腹板底部位置及3#-4#墩间引桥小箱梁腹板底部位置。该高速桥梁在测量时间内处于正常行车状态。

表11-1泗安塘大桥测点

测点	具体位置
1	北幅（林城-李家岗方向）主桥中跨跨中
2	北幅（林城-李家岗方向）主桥中跨四分之三跨
3	北幅（李家岗-林城方向）3#-4#墩间引桥

测点具体位置

1 北幅（林城-李家岗方向）主桥中跨跨中

2 北幅（林城-李家岗方向）主桥中跨四分之三跨

3 北幅（李家岗-林城方向）3#-4#墩间引桥

现场仪器布置情况如图11-1所示。测点位置如图11-2所示。

(a) 基于图像匹配技术的桥梁位移测量系统	(b) 雷图BJQN-V2.0动静态位移检测系统
-----------------------	--------------------------

(a) 基于图像匹配技术的桥梁位移测量系统 (b) 雷图BJQN-V2.0动静态位移检测系统

图11-1仪器现场布置图

(a) 测点1	(b) 测点2	(c) 测点3
---------	---------	---------

(a) 测点1 (b) 测点2 (c) 测点3

图11-2测点示意图

相机与各测点的位置关系及比例因子见表11-2。

表11-2 相机位置与比例因子

测点		距离（m）	垂直角度（度）	比例因子（mm/pixel）
1		23.854	9.1	0.938
2		12.815	15.7	0.530
3	S1	18.858	14.7	0.773
	S2	21.558	12.5	0.867
	S3	24.286	11.1	0.967
	S4	27.037	9.7	1.067

测点距离 (m) 垂直角度 (度) 比例因子 (mm/pixel)

1	23.854	9.1	0.938
2	12.815	15.7	0.530
3	S1 18.858	14.7	0.773
S2	21.558	12.5	0.867
S3	24.286	11.1	0.967
S4	27.037	9.7	1.067

对于测点1、2，同步使用雷图BJQN-V2动静态位移检测系统（测量精度±0.02mm）进行位移测量。同时，通过无人机记录了测点1、2、3桥面车辆通行情况状况。

9.1.2 测量结果

1) 主桥位移

测点1~2位移测量结果对比见图11-3、图11-4。对于测点1、2，由位移时程曲线可以看出S-F位移测量系统的位移测量结果与BJQN-V2测量结果保持一致。		
。(a) 0~350s		
(b) 60~100s		(c) 255~265s

测点1~2位移测量结果对比见图11-3、图11-4。对于测点1、2，由位移时程曲线可以看出S-F位移测量系统的位移测量结果与BJQN-V2测量结果保持一致。(a) 0~350s

(b) 60~100s (c) 255~265s

图11-3主桥跨中位移时程曲线

(a) 0~250s		
(b) 10~30s		(c) 200~240s

(a) 0~250s

(b) 10~30s (c) 200~240s

图11-4主桥四分之三跨位移时程曲线

2) 引桥位移

如图11-5所示，展示了桥面过车时，引桥4片小箱梁协同作用下各自的位移。为方便观察，将位移数据进行了截止频率为15Hz的低通滤波。

(a) 引桥跨中位移时程曲线		
(b) 第4s引桥过车情况		(c) 第5s引桥过车情况

(a) 引桥跨中位移时程曲线

(b) 第4s引桥过车情况 (c) 第5s引桥过车情况

图11-5引桥跨中位移时程曲线与过车情况

对于测点3，如图11-5(b)、(c)所示，在测量开始后的第4s和第5s，分别有一辆车经过引桥，分别引起引桥4片小箱梁相应的位移。将此引桥4片小箱梁由南至北分别编号为S1、S2、S3、S4。由图片可知，第4s时一辆满载车由超车道经过，正如图11-5(b)所示，引起S1、S2小箱梁的较大的挠度，S3、S4小箱梁挠度依次减小。随后，第5s时，两辆空车并排经过，引起S1、S2、S3小箱梁相比于第4秒稍小的挠度，且较远的S4小箱梁挠度依旧比较小。针对引桥多片小箱梁的同步测量，本书显示了利用基于图像匹配技术的桥梁位移测量方法进行多点测量的有效性。

9.2 南沙大桥位移监测应用

9.2.1 测试方案

南沙大桥，原称虎门二桥，是中国广东省境内一座连接广州市南沙区与东莞市沙田镇的过江通道，位于珠江主流之上。该桥单跨1200米，桥塔高193.1米，为双向八车道，桥面宽40.5米，桥梁宽49.7米。在2019年3月中旬进行通车前的荷载试验时，该桥在设计阶段在基准点和挠度测量点之间安装了连通管。利用连通管原理，通过测量液面变化监测测量点相对于基准点竖直方向的位置变化（即挠度）。进行计算机视觉技术的应用时，采用型号为UI-3370CP 的IDS工业相机，其图像传感器CMOS sensor的像素大小为5.5 μm。荷载试验在夜间进行，桥面均布的LED灯是夜间测量时的理想靶标。相机被加载在岸边，以斜光轴测量的方式基于灰度平方质心法对多个截面靶标进行同步追踪。

图11-6 南沙大桥位移监测示意图

荷载试验车辆拟定为总重350kN的试验车，前轴重70kN，中轴、后轴重均为140kN，后轴间距1.4m，中轴与前轴间距3.6m，横向轮距1.8m。具体参数如图11-7所示。

图11-7 试验加载车辆示意图（单位：cm）

静载试验：以1/2截面对称加载，测量最不利荷载下的挠度。采用卡车加载，共64辆车，按四级加载。卡车的就位顺序如图11-8所示：

图11-8 静载试验布载示意图

图11-9 静载试验位移测试结果

当卡车即将上桥时，相机开始采集图片，相机采集帧频为2Hz，连续采集。该工况主要关注八分之二截面的挠度。

图11-10 行车荷载试验布载示意图

图11-11 行车荷载试验位移测试结果

9.3 基于无人机的结构动位移测量这里再讲原理就有些不统一了，建议振动台不要，无人机测位移的原理精简些和右叉悬索桥统一为一体

9.3.3 南京夹江大桥位移监测应用

在南京夹江大桥进行现场测试，对其正常车辆通行条件下跨中位置产生的竖向动位移进行了测量。如图11-16所示，该桥为

单塔斜拉桥，总长455m（35m+248m+60m+77m+35m），主跨248m，主缆锚固于P2和P5两桥墩墩顶处。测试过程中，无人机悬停于测点附近，调整无人机及相机云台角度，尽量使无人机镜头正对测点，激光灯架设于岸边扰动较小的位置，并在测点附近投射出较为清晰的光斑。测试时，在岸边架设静止相机同步进行光学测量，并以此结果与本文方法的测量结果进行对照。

图11-16 (a)无人机、静止相机拍摄视角，(b)夹江大桥构造尺寸及测点位置

图11-17显示了夹江大桥正常通车情况下垂直位移的测量结果，包括：激光点相对于无人机的位移、靶标相对于无人机的位移、使用无人机方法计算的靶标的绝对位移，以及利用固定在地面上的静止相机测量的靶标位移。结果表明，用该方法计算的位移与用静止相机测量的位移吻合较好，在截止频率为0.1 Hz的低通滤波后计算NRMSE，结果为4.394%。根据无人机拍摄的视频，当桥梁出现明显的挠度（约在第14秒、32秒、60秒、97秒、115秒和124秒）时，总有重型车辆（如卡车、大巴车等）经过，其中最显著的位移出现在60秒左右，是由一辆经过的混凝土搅拌车引起的。图11-17还展示了视频处理的中间结果，可以看出，尽管激光光斑的显示度比在室内环境中淡了很多，仍能准确识别并定位靶标和激光光斑。

图11-17 基于无人机监测的夹江大桥位移时程曲线

习题

91 结构位移监测与位移检测的区别是什么？

9-2 结合课程所学与本章案例，提出针对桥梁结构位移监测的具体方案。

22. 计算机视觉教材_第22部分		总字符数：5050
相似文献列表		
去除本人文献复制比：3.8%(193) 去除引用文献复制比：3.8%(193) 文字复制比：3.8%(193)		
1	<u>浅谈混凝土质量通病的防治措施及处理方法</u> 李增营；- 《华章》- 2010-08-30	2.0% (99) 是否引证：否
2	<u>浅析混凝土质量通病</u> 马曙光；- 《中国城市经济》- 2011-09-15	1.6% (82) 是否引证：否
3	<u>浅析混凝土质量通病防治的措施</u> 郭晏；- 《西部探矿工程》- 2006-04-15	1.6% (82) 是否引证：否
4	<u>浅谈混凝土质量通病的防治措施及处理方法</u> 李艳红;贺伟；- 《河南水利》- 2006-04-20	1.6% (81) 是否引证：否
5	<u>混凝土的几种质量通病及防治方法</u> 杨怀印；- 《科技信息(学术研究)》- 2007-08-15	1.6% (81) 是否引证：否
6	<u>水工混凝土质量通病产生的原因及防治措施</u> 李向阳;孙启凯；- 《水利技术监督》- 2007-01-30	1.6% (79) 是否引证：否
7	<u>浅析预应力混凝土空心板外观质量控制</u> 张勇;干其明；- 《科技资讯》- 2010-12-13	1.4% (71) 是否引证：否
8	<u>浅析混凝土工程的质量通病及其防治措施</u> 张会琴；- 《青海科技》- 2008-10-25	1.4% (71) 是否引证：否
9	<u>高层住宅钢筋砼楼板裂缝成因分析与处理</u> - 《网络 (http://fanwen.chazid) 》- 2016	1.0% (53) 是否引证：否
10	<u>高桩码头砼预制构件的表观质量控制</u> 于涛;林垂兴;王雪华；- 《中国水运(下半月)》- 2012-09-15	0.8% (42) 是否引证：否
11	<u>混凝土施工中常见问题及处理方法</u> 曹维江；- 《价值工程》- 2011-11-28	0.8% (41) 是否引证：否
12	<u>基于计算机视觉的大黄鱼体尺、体重性状表型测量装置开发和应用</u> 王禹莎;王家迎;辛瑞;柯巧珍;江鹏鑫;周涛;徐鹏；- 《水产学报》- 2022-11-09 08:29	0.7% (36) 是否引证：否
原文内容		

第10章结构病害智能检测典型案例

在土木工程领域，结构病害的智能检测正成为一项引人注目的研究领域。随着基础设施的日益复杂化和规模的不断扩大，及时准确地检测和评估结构病害的重要性日益凸显。结构病害不仅可能导致安全隐患和功能问题，还会对工程的可持续发展 and 使用寿命造成影响。因此，研究人员和工程师们迫切需要一种智能化的检测方法，能够实现对结构病害的精准识别、快速定位和准确评估。本章将带领读者深入探索结构病害智能检测的典型案例，该案例基于计算机视觉、深度学习和图像处理等先进技术，展示了智能检测在土木工程中的巨大潜力和应用前景。通过学习该案例，读者将了解到智能检测技术在结构病害领域的创新应用，以及其对工程安全和可持续发展的重要意义。同时，读者也将深入了解智能检测方法的原理和 workflows，为今后的研究和实践提供有益的参考。

10.1 桥梁概况
图10-1 桥梁概况

典型案例聚焦于苏淮扬高速公路上的一座大跨度桥梁，它是一座预应力混凝土斜拉桥，主跨度为370米，边跨为152米，桥梁宽度为 38.6 米。该桥的断面为双面箱梁的异形断面，因此只需对两侧箱梁的底面进行检查。该桥建于 2005 年。经过十余年的服役，桥面和桥塔上出现了一定数量的裂缝、剥落和锈蚀。以往的检查报告显示，表面病害主要分布在桥底和桥塔上，本章将结合无人机与计算机视觉、深度学习和图像处理等先进技术，对主跨的底部、两侧边跨的底部和桥塔的外表面进行检查，桥梁的总体概况和检测过程如图 3-22 所示。

10.2智能检测流程
10.2.1典型结构病害

(a)	(b)
(c)	(d)

(a) (b)
(c) (d)
图10-2 表观病害 (a) 蜂窝；(b) 麻面；(c) 锈蚀；(d) 裂缝

在执行检测任务之前，首先需要了解待检测的病害类型。土木工程中的典型表观病害主要有：蜂窝、麻面、锈蚀、裂缝等。蜂窝是指混凝土结构局部出现酥松，砂浆少、石子多，石子之间形成空隙类似蜂窝状的窟窿。麻面的现象是混凝土局部表面出现缺浆和许多小凹坑、麻点，形成粗糙面，但无钢筋外露现象。锈蚀是指钢筋表面由于暴露在潮湿、氧气和盐分等有害环境条件下，发生了化学反应而导致的腐蚀现象裂缝：裂缝是由于结构由于内外因素的作用而产生的物理结构变化，而裂缝是结构物承载能力、耐久性及防水性降低的主要原因。上述病害是一般性工程结构的通病，特殊的结构类型会存在相应的病害，例如悬索桥、斜拉桥的索套会存在破损；钢结构类型设施的螺栓会存在松动、脱落等问题，这些特殊类病害对结构同样具有危害，病害处会成为结构的薄弱点，破坏最先在这里产生，然后由局部蔓延到整体最终甚至导致结构倒塌。本章案例重点关注裂缝、剥落和钢筋锈蚀这三种病害。

10.2.2 数据采集方案

确定病害检测类型后，需要策划桥梁数据采集方案。桥梁通常位于高空或难以到达的区域，传统的检测方法需要人工爬行或搭建架设备，耗时费力且操作复杂。该典型案例将无人机作为桥梁检测工作的手段，其快速的响应能力和高效的数据采集能力，使得无人机成为桥梁检测工作的理想选择，能够克服人工检测的困难和限制，实现整个桥梁结构的轻松覆盖。同时，无人机设备可以配备高分辨率的摄像设备和先进的传感器技术获取高质量、高精度的桥梁图像和数据。通过使用无人机搭载的摄像头，可以对桥梁表面进行全方位的拍摄，捕捉细微的结构病害和损伤。同时，无人机还可以搭载其他传感器，如红外热像仪、激光扫描仪等，以获取更加全面和多样化的数据，进一步提高检测精度和可靠性。此外，无人机具备安全性和可操作性。在传统的桥梁检测中，工程师需要亲自进入高风险的环境，存在安全隐患。而使用无人机进行检测，可以避免工作人员的潜在危险，减少人为风险的发生。这些优势使得无人机成为现代土木工程领域中越来越受欢迎的检测工具，为工程师提供了一种高效、准确的方式来进行桥梁结构的智能检测与评估。

图10-3 无人机检测过程

在该典型案例中，无人机飞行时距离桥梁底部的安全距离为 2~3 米，在该距离下相机的视场宽度为 3~4 米。桥底两侧的检查区域为 8.5 米，因此每一侧的检查至少需要飞行 3 条路线。跨江大桥附近的风速普遍较高，无人机检测时的天气为晴朗但风速较大的天气，风速约为 6 米/秒。为了确保安全飞行，无人机系统由一名熟练的操作员控制，对一跨桥底的检测执行了三次飞行，总飞行时间约为 30分钟。图 10-4所示，以该结果为参考操作人员能一定程度掌握无人机的飞行路线和已检测的区域。飞行过程由人工操作完成，因此无人机航线无法维持平直规则的飞行航线。

10-4桥梁病害检测过程的无人机轨迹

10.2.3 智能识分析方法
10-5多类型病害检测网络

数据收集完成后，针对裂缝、剥落和钢筋锈蚀这三种结构病害的检测任务，本案例采用了第七章介绍的基于深度学习的目标检测网络进行多类型病害的识别。通过应用基于深度学习的目标检测网络进行多类型病害的识别，本案例取得了显著的成果。首先，针对裂缝的检测，通过训练深度学习模型，能够准确地检测出桥梁结构中各种尺寸和形态的裂缝，并对其进行分类和定位。这为工程师提供了重要的信息，使其能够及时采取修复措施，防止进一步的结构损害。其次，对于剥落问题，基于深度学习的目标检测网络能够有效地识别桥梁表面剥落的区域和程度。通过分析剥落的特征，工程师可以快速了解结构的健康状况，制定相应的修复计划，并确保桥梁的稳定性和安全性。最后，针对钢筋锈蚀这一严重的结构病害，基于深度学习的目标检测网络能够在图像中准确识别和定位锈蚀的钢筋区域。通过对锈蚀程度的评估，工程师可以及时采取防腐措施和维护工作，延长桥梁的使用寿命，确保其结构的稳定性和可靠性。

本案例为验证病害自动识别结果的准确性，从桥底检测和桥塔检测的视频中选取了分辨率为1920×1080 像素的 21 分钟视频，提1014 帧作为对比数据集，对所选取的帧采用所提出方法和人工目测识别和标记的方法标记其中的病害进行对比。结果显示，1014 幅图像的平均分类精度为 92.8%，平均重叠率为 80.7%。测试结果示例见图 10-6。

10-6测试结果示意图

10.2.4 病害定位与可视化

无人机视频的病害检测结果不仅仅显示了病害的存在，还揭示了一些重要的观察结果。分析这些结果可以提供更深入的理解和评估病害的特征和规模。观察到含有病害的图像往往以成组的形式出现，如图10-7所示。这意味着对于长度较长或面积较大的病害，例如长裂缝，单帧图像仅包含裂缝的一部分，从离散的图像中很难准确评估整个裂缝的几何参数。这样的情况使得单个图像的分析 and 判断具有一定的局限性。为了克服这个问题，本案例采用了一种病害帧聚类的方法。首先，将具有连续表面病害的相邻帧归为一组。通过将这些相邻帧进行聚类，我们可以获得一组关联度较高的图像序列，其中每个图像都涉及病害的不同部分。这种聚类方法有助于提供更全面、连续的病害信息，为后续分析和处理奠定基础。进一步地，为了生成每组病害图

像的局部病害拼接图，采用了基于 ORB 特征提取和匹配的图像拼接方法。ORB特征是一种快速而有效的特征描述符，能够对图像中的关键点进行描述和匹配。通过提取每组病害图像中的ORB特征，并进行特征匹配，我们能够将这些图像拼接成一张更大范围的局部病害拼接图。这样的拼接图提供了更全面的视角，使得病害的空间分布和形态特征更加清晰可见。综合而言，通过采用病害帧聚类 and 基于ORB特征的图像拼接方法，本案例能够从无人机视频中获得更具完整性和连续性的病害信息。这种扩充的分析方法为我们提供了更全面的了解和评估结构病害的特征，并为后续病害诊断和修复工作提供了有力支持。此外，这种基于无人机视频的病害检测和分析方法也为结构监测领域的进一步研究和发展提供了新的思路 and 方向。

10-7病害拼接

为了便于展示检测结果，并获得更全面的桥梁信息，本案例利用无人机拍摄了桥梁的视频，并采用了第8章介绍的运动恢复结构法对桥底和桥塔进行了三维重建。通过这一过程，我们能够获得桥梁的立体模型，并对其进行更加细致的分析和评估。从图10-8中可以观察到一些有趣的发现。首先，在对桥底进行检测之前，该区域已经进行了修复工作，因此检测到的表面损伤相对较少。主要存在一些局部的混凝土剥落问题，而这些病害对于整体桥梁的安全性能影响较小。这说明之前的维修工作在一定程度上保障了桥梁的结构完整性。然而，对于桥塔的侧面，长期受力的作用导致了大量的长裂缝的出现，这需要及时修复。这些长裂缝的存在可能会对桥梁的稳定性和承载能力产生潜在的影响。因此，针对这些裂缝病害，我们需要采取有效的维修和加固措施，以确保桥塔的结构安全和可靠性。通过对桥底和桥塔进行三维重建和病害检测，我们能够更全面地了解桥梁结构的状况，包括病害类型、位置和严重程度等方面的信息。这为工程师提供了重要的参考依据，以制定合理的维修计划和预防措施，确保桥梁的长期使用和安全运行。

为定量评估病害自动定位的误差，选取 760 幅桥塔侧面图像人工拼接为桥塔侧面的全景图。拼接后的图像分辨率超过 108 亿像素，按高度将图像分为 16 个子图像进行处理。由于桥塔的几何尺寸是已知的，图像的坐标系设定为以桥塔底部的水平线为 X 轴，以桥塔中间的垂直线为 Y 轴，以桥塔底部的中点为 (0, 0)。然后采用人工标注的方式，将图像中的三种损伤全部标注出来，并将标注的方框中心的坐标记录为病害的位置。将上述操作计算出的结果作为真值，与所提出的方法的定位结果对比。病害识别的误差包括真值 (TP) 和三种类型病害的准确性，病害定位的 误差包括 X 和 Y 坐标的平均绝对误差 (MAE) 和平均相对误差 (MRE)。结果显示，病害识别的平均检测精度为 91.7%，X 和 Y 坐标的定位误差在分米级。对于高度为 137 米的桥塔的顶部，由于误差的累积，平均定位误差约为 40 厘米，但靠近桥塔底部的定位误差一般小于 20 厘米。上述结果表明，所提出的方法的精度符合工程要求。

10-8桥梁底部和桥塔的检测结果

综上所述，本案例的病害智能检测工作在裂缝、剥落和钢筋锈蚀等方面取得了显著的进展。基于深度学习的目标检测网络为土木工程领域的桥梁检测工作带来了革命性的改变，提高了检测的准确性、效率和可靠性。这一案例的成功经验为类似的结构病害检测工作提供了有益的参考，并为进一步推动智能化的土木工程领域提供了宝贵的经验。

小结

本章介绍了结构病害智能检测的典型案例。随着基础设施的复杂化和规模的扩大，及时准确地检测和评估结构病害的重要性日益凸显。本章基于计算机视觉、深度学习和图像处理等先进技术，展示了智能检测在土木工程中的巨大潜力和应用前景。首先介绍了桥梁的概况，聚焦于一座预应力混凝土斜拉桥，经过十余年的使用出现了裂缝、剥落和钢筋锈蚀等病害。为了对这些病害进行检测，本案例采用了无人机作为检测工具，具有快速响应和高效数据采集的优势。无人机能够轻松覆盖整个桥梁结构，获取高质量、高精度的图像和数据。无人机还可以搭载其他传感器，如红外热像仪和激光扫描仪，进一步提高检测精度和可靠性。在数据采集完成后，本案例采用基于深度学习的目标检测网络对裂缝、剥落和钢筋锈蚀进行识别。该方法能够准确地检测和定位各种尺寸和形态的病害，为工程师提供重要的信息和修复计划。最后介绍了病害定位与可视化方法。无人机视频的检测结果不仅显示了病害的存在，还提供了更深入的理解和评估病害的特征和规模。通过对图像进行聚类和分析，可以更准确地评估整个病害的几何参数。通过学习本章的案例，读者将了解到智能检测技术在结构病害领域的应用和意义，以及智能检测方法的原理和工作流程。

习题

- 1. 什么是结构病害智能检测？为什么它在土木工程中至关重要？
- 2. 解释无人机在结构病害检测中的优势和应用。
- 3. 深度学习在结构病害智能检测中起到了什么作用？为什么它比传统方法更具优势？
- 4. 为什么病害定位和可视化对于结构病害智能检测非常重要？请提供一个例子来说明其应用。
- 5. 列举几个常见的桥梁病害类型，并描述它们的特征和可能的危害。
- 6. 如果你被要求设计一个结构病害智能检测系统，你会如何选择合适的传感器和设备？
- 7. 讨论结构病害智能检测对于桥梁维修和保养的意义。如何利用智能检测结果来制定修复计划？

23. 计算机视觉教材_第23部分		总字符数：2606
相似文献列表		
去除本人文献复制比：0% (0) 去除引用文献复制比：0% (0) 文字复制比：0% (0)		
原文内容		

第11章结构施工视觉测量应用案例

计算机视觉因其具有非接触、远距离、高精度和多点测量的优点而逐渐应用于土木工程领域，如结构缺损、变形的监测与定位、施工进度更新等。目前，基于计算机视觉技术在施工过程中的应用主要集中于施工人员安全管理和施工检测两方面，对在建结构的实时监测研究相对较少。而施工工地所拍摄的视频和图像主要用于记录及跟踪项目状态，跟进施工进度，记录结构

质量检测数据等。此外，目前施工现场采集的图像及视频需要人工检查及记录，数据挖掘不充分，自动化处理程度低。因此，在日益增长的施工监测需求和劳动力短缺的背景下，开发基于计算机技术的施工过程自动化监控和测量新技术具有重要的工程意义。

本章节提供了典型施工监测场景的案例分析：使用计算机视觉进行钢筋尺寸测量。进一步的，在拓展部分，针对预制墩柱安装位姿监测和质量控制的需求，探讨了基于双目视觉和激光扫描的两种解决方案，为施工现场预制梁构件安装提供指导。

11.2 预埋件尺寸测量

工程安装包括钢板、螺栓、接线盒等结构件，以及连接管、排水管等预埋管。工程安装的施工质量直接影响建筑工程的施工进度和结构安全，应牢牢控制。在这一阶段，各种工程安装的位置由工人在现场进行检查，然后再填充水泥，对于大型工程的施工面积，其中工程安装数量巨大，安装位置分散，如接线盒、管道等预埋件对齐复杂，因此人工检查费时费力，效率低下，并且工人在高层建筑施工中爬高是危险的。与对结构裂缝和变形(挠度和位移)的运行维护研究相比，对工程安装过程检测的研究受到的关注较少，因此有必要加强工程安装检测的研究与开发。

图11-8传统预埋件安装位置测量

对于仅使用相机的计算机视觉方法，存在无法获得深度方向坐标的缺陷；而仅使用三维激光扫描仪的方法存在扫描速度慢、点云数据量大、无法在现场快速得到测量结果等问题。此外，图像处理方法与三维激光扫描相比，具有分辨率高、图像采集快、处理速度快等优点，但无法获得深度信息。此处采用了图像和点云融合的方法，结合图像和点云各自的优点，以图像处理为主要方法，点云深度信息为辅助方法。整体工作环节如图11-9所示。图中从二维图像中快速检测出预埋件采用的是语义分割网络unet++。

图11-9激光点云与图像融合测量方法

现场测量场景如图11-10所示。

图11-10预埋件安装位置现场测量

其处理结果示意图如图11-11所示。测量误差在mm量级。

图11-11预埋件安装位置现场测量结果

11.3 预制墩柱安装位姿监测

装配式桥梁结构具有建造速度快、对周边交通干扰小等优点，因此受到推广。装配式预制墩柱吊装属于危大工程施工，在保证施工安全的同时也要严格控制安装精度质量。如图11-12所示，传统的基于全站仪和水准仪的监测方法，需预先在承台上放样出两水平轴线，用于安装观测仪器，同时在墩柱表面标注出纵向轴线，用于安装垂直度校核。

图11-12基于水准仪的预制墩柱安装监测方案

传统测量方式自动化程度低、测量前准备工作复杂，并且需要投入较多的人力、物力，故亟需提高施工现场的测量方式。双目视觉具有多点同步、三维非接触测量的特点，三维激光扫描仪具有大范围内三维精确感知、不依赖标志点等优点，两种测量方法均可实现结构六自由度位姿监测，是预制构件安装质量控制的可行方案。

双目视觉的方案通过在墩柱及地面上分别安装靶标，利用双目相机测量墩柱上多个靶标三维坐标获取墩柱三维姿态，并通过其三维姿态与地面姿态的换算关系实现墩柱的位姿监测，指导墩柱安装，其检测靶标的方法为基于深度学习目标检测的方法，使用YOLOv5检测靶标。三维激光扫描直接获取的三维点云数据，实现墩柱的位姿监测，不需要粘贴靶标。

基于双目视觉及三维激光扫描仪的测量示意图分别如图11-13、图11-14所示。

图11-13预制桥梁现场吊装施工双目监测图

图 11-14预制墩柱安装过程中现场三维激光扫描监测图

通过双目视觉与激光扫描仪获取的控制点在参考坐标系下的坐标与对应的理想坐标进行对比，当所有控制点在参考坐标系下的坐标与对应的理想坐标之间的误差均小于预设阈值时，判断完成预制构件的吊装。构件吊装监控流程如图11-15所示。

图11-15 基于构件空间姿态的吊装全过程监控

基于双目视觉的选定控制点的三维运动轨迹图如图11-16所示。测量结果显示了控制点的空间位置随时间变化，可以看到立柱在墩帽上方约 800mm 的高度开始下降。当五个控制点到500mm高度处开始调整姿态，使得五个控制点的XY平面坐标接近设计值时，使所有钢筋和套管大致对准，构件可继续下落。当构件达到指定标高时，对构件的姿态进行微调，使得立柱达到最终设计状态。可以发现，所有控制点在最后的10个调整状态逐渐靠近理想的设计状态，公差约为2mm，在允许的误差范围之内。

图11-16立柱吊装结果

基于激光扫描仪的墩柱位姿监测如图11-17所示，包含两阶段，阶段1对应承台预埋钢筋插入前，主要监测墩柱相对设计状态的水平偏差，阶段2对应墩柱竖向就位后液压千斤顶微调过程，主要监测墩柱竖向倾斜角的偏差。监测结果见下图11-17。可见，阶段1下降过程中墩柱中心点逐渐接近设计状态中心；阶段2微调后，墩柱安装的竖向倾斜角为0.02度，两水平偏差分别为-0.23 cm、0.45 cm，满足安装质量控制要求。阶段2中水平偏差变化不显著，说明有必要在阶段1墩柱下降就位前调整好水平位置，避免重复起吊安装。

图 11-17 基于扫描仪的墩柱安装位姿监测结果：(a) 阶段1墩柱下降过程中局部坐标系P相对墩柱设计坐标系的变化；(b) 阶段2墩柱微调过程中竖向倾斜角的变化；(c) 阶段2墩柱微调过程中两水平方向偏差的变化。

思考题

(5) 案例11.1中，利用了基于特征点的图像拼接方法。思考一下，还有其他的图像拼接方法可以适用于这种场景吗？

(6) 案例 11.1中，相机像素为5000*4000，而对钢筋轮廓进行语义分割提取时，输入图片为512*512，若直接将图像进行下采样会明显导致分割不精确，你有什么好的分割方法吗？

(7) 案例11.1中对钢筋分割图进行分段直线拟合，如何将钢筋进行分段拟合？

(8) 比较案例11.2中两种智能施工方式在目标获取与识别、三维重建上的异同点；

参考文献

[1] Cheng Y, Lin F, Wang W, et al. Vision-based trajectory monitoring for assembly alignment of precast concrete bridge components[J]. Automation in Construction, 2022, 140: 104350.

习题参考答案

相似文献列表

去除本人文献复制比：65.8%(462)		去除引用文献复制比：65.8%(462)	文字复制比：65.8%(462)
1	基于图像匹配技术的桥梁位移无标记测量方法研究 汪威(导师：吴刚;吴栋) - 《东南大学硕士论文》 - 2021-05-30	22.6% (159)	是否引证：否
2	基于机器视觉的车用灯泡检测系统研究 - 道客巴巴 - 《互联网文档资源 (https://www.doc88.co)》 - 2020	15.8% (111)	是否引证：否
3	基于机器视觉的电视机背板定位技术研究 谢献欢(导师：李军) - 《广东工业大学硕士论文》 - 2022-05-01	12.5% (88)	是否引证：否
4	基于BM算法的视觉匹配验证 袁娜;张在权; - 《唐山学院学报》 - 2019-11-20	11.4% (80)	是否引证：否
5	基于机器视觉技术的棉株识别系统研究 瞿端阳(导师：王维新) - 《石河子大学硕士论文》 - 2013-06-01	7.8% (55)	是否引证：否
6	基于多相机系统标定的全景拼接算法研究 林华琦;兰诚栋; - 《信息通信》 - 2018-03-15	7.4% (52)	是否引证：否
7	内窥镜三维成像关键技术研究 向静(导师：代志勇) - 《电子科技大学硕士论文》 - 2022-03-10	7.4% (52)	是否引证：否
8	基于双目视觉的车身尺寸测量系统设计 白创;谢伊伶;许百灵; - 《单片机与嵌入式系统应用》 - 2022-07-01	4.8% (34)	是否引证：否
9	基于CNN的机器视觉系统图像畸变校正算法研究 杨雪(导师：顾海军) - 《吉林大学硕士论文》 - 2022-05-01	4.8% (34)	是否引证：否
10	基于机器视觉的烟包形状识别与码垛技术的应用研究 康家旗(导师：张毅) - 《重庆邮电大学硕士论文》 - 2021-06-30	4.4% (31)	是否引证：否
11	基于视觉的芯片电容检测与分拣系统研究与实现 刘晓容(导师：邹见效) - 《电子科技大学硕士论文》 - 2021-05-01	4.1% (29)	是否引证：否
12	大口径火炮炮口振动测量方法研究 赵刚(导师：陈强) - 《南京大学硕士论文》 - 2015-05-01	4.1% (29)	是否引证：否

原文内容

第二章

2-1 镜头的基本参数有哪些，这些基本参数的定义是怎样的？

参考答案：焦距：焦距是焦点到透镜中心点之间的距离；视场角：整个视觉系统所能观察到的物体实际尺寸被称为视场即视场范围；光圈：光圈实际为可以调节孔径大小的机械部件，它通常在镜头内，利用控制镜头光孔大小来控制进入相机的光量；景深：镜头能够取得清晰图像时被测物体的前后距离范围即为景深；畸变：由像面上局部放大倍率不一致所导致的现象。

2-2 相机成像模型涉及哪些坐标系？

参考答案：相机成像模型中涉及世界坐标系、相机坐标系、图像像素坐标系及图像物理坐标系四个坐标系。

2-3 世界坐标系与相机坐标系是如何转换的？

参考答案：利用旋转矩阵与平移向量可以实现世界坐标系中的坐标点到相机坐标系中的映射。实现世界坐标系与相机坐标系的转换。

2-4 图像像素坐标系与世界坐标系是如何转换的？

参考答案：利用投影矩阵可以实现图像像素坐标系与世界坐标系的转换。

2-5 什么是相机的标定？

参考答案：确定相机内参和外参的过程即为相机的标定。

2-6 简述张正友相机标定法的基本步骤。

参考答案：①制作棋盘格标定板；②通过移动或旋转棋盘格标定板，相机拍摄记录不同方向、角度的标定板图像；③检测图像特征点；④通过旋转矩阵的正交性及旋转向量模为1的性质，求解不考虑畸变的影响下线性方程组，获得相机内部参数与外部参数；⑤通过最小二乘法求解相机径向畸变系数；⑥计算重投影误差，利用极大似然估计，迭代优化上述步骤求解的相机参数。

2-7 尝试利用棋盘格，使用MATLAB相机标定工具箱对手机摄像头进行标定，获得其相关参数。

参考答案：略。

相似文献列表

去除本人文献复制比: 21.4%(236)

去除引用文献复制比: 21.4%(236)

文字复制比: 21.4%(236)

1	<u>压缩感知理论简介</u> 宁寰宇;文亚洲; - 《电子技术》 - 2012-06-25	7.3% (80) 是否引证: 否
2	<u>航天发动机喷雾实验系统的喷雾性能监测研究</u> 王淑姣(导师: 朱凌云) - 《东华大学硕士论文》 - 2022-05-30	5.7% (63) 是否引证: 否
3	<u>地铁车辆低频设备电磁环境复合测试装置及方法</u> 黄健;陈立;刘嘉琦; - 《环境技术》 - 2022-12-25	5.0% (55) 是否引证: 否
4	<u>基于半周积分算法的变压器过励磁保护的改进算法研究</u> 刘黎;杨卉卉;秦嗣友;肖远清; - 《变压器》 - 2021-07-25	5.0% (55) 是否引证: 否
5	<u>基于DSPIC30F3011光伏发电系统数据采集卡的设计</u> 王兴贵;张会波;伊世明;徐贵生; - 《工业仪表与自动化装置》 - 2010-12-05	5.0% (55) 是否引证: 否
6	<u>基于流量控制的压缩机智能控制系统的应用</u> 魏国玲;仪维;尚勇;康爱东;曹庆峰;田勇; - 《实验技术与管理》 - 2011-12-20	5.0% (55) 是否引证: 否
7	<u>基于TMS320LF2407A的固体流量检测系统</u> 段古纳;张茜; - 《仪器仪表与分析监测》 - 2012-05-28	5.0% (55) 是否引证: 否
8	<u>频率法测量索力的研究</u> 韩宗泽; - 《天津建设科技》 - 2013-04-30	5.0% (55) 是否引证: 否
9	<u>基于稀疏快速傅里叶变换的信号压缩处理</u> 刘清华;杨桂芹;张妍妮; - 《微型机与应用》 - 2016-08-03 1	5.0% (55) 是否引证: 否
10	<u>基于压缩感知的QAR数据重构</u> 耿宏;李萍萍;刘家学;高远; - 《计算机测量与控制》 - 2013-05-25	4.8% (53) 是否引证: 否
11	<u>浮选泡沫表面动态特征提取方法研究</u> 彭井花;廖一鹏; - 《福建师范大学学报(自然科学版)》 - 2023-05-15	4.7% (52) 是否引证: 否
12	<u>一种基于人脸视频的非接触式心率测量系统</u> 张祺; - 《中国新技术新产品》 - 2022-11-25	4.6% (51) 是否引证: 否
13	<u>基于活动轮廓模型的肺结节分割方法研究</u> 陈侃(导师: 田联房) - 《华南理工大学博士论文》 - 2013-10-08	4.2% (46) 是否引证: 否
14	<u>利用微动高分辨率频率一波数谱法探测技术勘探地热资源</u> 刘杨;盛勇;贾慧涛; - 《工程地球物理学报》 - 2021-01-27	3.7% (41) 是否引证: 否
15	<u>基于二码元长度的实时ACARS信号分析算法</u> 袁树德;曹力;邓雪云;蒋兴城; - 《信息技术》 - 2012-07-25	3.7% (41) 是否引证: 否
16	<u>基于DSP的针刺手法信号采集系统的设计</u> 郭彦;郑日荣;李宗群; - 《工业控制计算机》 - 2011-07-25	3.7% (41) 是否引证: 否
17	<u>高速列车表面气动噪声偶极子声源分布数值分析</u> 郑拯宇;李人宪; - 《西南交通大学学报》 - 2011-12-15	3.7% (41) 是否引证: 否
18	<u>基于智能手机平台的环境噪声监测研究</u> 李春明;董仁才;赵景柱; - 《安全与环境学报》 - 2012-04-25	3.7% (41) 是否引证: 否
19	<u>FIR数字滤波器的Matlab实现</u> 许书云; - 《天津职业技术师范大学学报》 - 2013-06-28	3.7% (41) 是否引证: 否
20	<u>绝对式光电编码器动态误差采集系统</u> 于海;万秋华;杨守旺;赵长海;卢新然; - 《红外与激光工程》 - 2013-07-25	3.7% (41) 是否引证: 否
21	<u>点焊规范参数实时在线监测系统</u> 白志范;蔡洪丰;王文权; - 《电焊机》 - 2014-02-20	3.7% (41) 是否引证: 否
22	<u>浅析极值法提取脉搏信号特征点的方法</u>	3.7% (41)

	李丽颖;郭红静;刘灿;袁娜; - 《电子质量》 - 2014-02-20	是否引证: 否
23	焊装车间点焊设备网络化监测系统 王嘉;蔡洪峰;白志范;王文权; - 《计算机应用》 - 2014-06-15	3.7% (41) 是否引证: 否
24	故障录波器的使用及数据分析 任玉佩; - 《山西电力》 - 2007-12-15	3.7% (41) 是否引证: 否
25	一种应用Fx LMS算法的双层隔振试验装置 高伟鹏;何其伟;闫政涛;刘锦春; - 《噪声与振动控制》 - 2016-08-18	3.7% (41) 是否引证: 否
26	基于LiDAR数据开展活动断层填图的实验研究——以新疆独山子背斜-逆冲断裂带为例 魏占玉;何宏林;高伟;徐锡伟;甘卫军;卫蕾华; - 《地震地质》 - 2014-09-15	3.1% (34) 是否引证: 否
27	基于VC++的数字图像处理软件开发 常奇峰(导师:王开福) - 《南京航空航天大学硕士学位论文》 - 2010-12-01	2.9% (32) 是否引证: 否

原文内容

第三章

3-1 维信号处理中,模拟信号数字化的过程包括?

参考答案:一维信号处理中,模拟信号数字化包括抽样、量化和编码三个基本过程。

3-2 在计算机中,图像的基本单位是什么?其定义是怎样的?

参考答案:在计算机中,图像的基本单位是像素。图形由像素组成,一幅图像被分为 $M \times N$ 个方格,每个方格即为像素。

3-3 图像数字化过程中,采样是指什么?采样间隔与图像分辨率的关系是怎样的?

参考答案:采样是把一幅连续图像在空间上离散成 $M \times N$ 个网格,是实现图像数字化的重要环节,决定着图像的空间分辨率。一般来说,采样间隔越大,所得图像像素数越少,空间分辨率低,质量差,严重时出现马赛克效应;采样间隔越小,所得图像像素数越多,空间分辨率高,图像质量好,但数据量大。

3-4 采样定理又称奈奎斯特定理,该定理的具体内容是?

参考答案:在进行模拟信号的转换过程中,当采样频率大于信号中最高频率的2倍时,采样之后的数字信号完整的保留了原始信号中的信息。

3-5 量化的定义是什么?采用不同的量化级数对图像质量有怎样的影响?

参考答案:所谓量化,就是将每个像素点的灰度值从模拟量转化为离散量的过程,并可选用离散量代替像素的灰度值,使像素看起来简单方便,有利于进行各种操作。量化等级越多,所得图像层次越丰富,灰度分辨率高,图像质量好,但数据量大;量化等级越少,图像层次欠丰富,灰度分辨率低,会出现假轮廓现象,图像质量变差,但数据量小。

3-6 常用的数字图像按照颜色和灰度的多少可以分为哪几种基本类型?

参考答案:可以分为二值图像、灰度图像和RGB图像三种基本类型。

3-7 灰度图像一般是指具有256级灰度值的数字图像,即8bit灰度图像。灰度图像矩阵元素的取值范围为?

参考答案:灰度图像矩阵元素的取值范围为 $[0, 255]$ 。

3-8 常见的图像噪声有哪些?

参考答案:高斯噪声、椒盐噪声、脉冲噪声。

3-9 最常见和最基本的图像滤波方法有哪些?

参考答案:均值滤波、中值滤波、高斯滤波等。

3-10 图像进行高斯滤波后的平滑程度与高斯滤波的标准差有怎样关系?

参考答案:标准差越大,高斯滤波器的频带就越宽,图像就被平滑得越好。

3-11 常用的一阶微分边缘算子有哪些?常用的二阶微分边缘算子有哪些?

参考答案:一阶微分边缘算子:Roberts算子、Prewitt算子、Sobel算子等。二阶微分边缘算子:Laplace算子、LOG算子、Canny算子。

3-12 常用的图像插值的方法有哪些?

参考答案:常用的插值方法有三种,即最近邻插值法、双线性插值法和双三次插值法。

26. 计算机视觉教材_第26部分		总字符数: 638
相似文献列表		
去除本人文献复制比: 24.8% (158) 去除引用文献复制比: 24.8% (158) 文字复制比: 24.8% (158)		
1	高温环境下基于机器视觉的微位移测量 张睿(导师:肖世德) - 《西南交通大学硕士学位论文》 - 2020-04-01	13.0% (83) 是否引证: 否
2	一种利用特征点检测实现可交互式地图的方法 奚家泉; - 《数字通信世界》 - 2018-07-01	5.5% (35) 是否引证: 否

3	遮挡目标检测与识别技术研究	4.7% (30)
	蔡星艳;赵和鹏;邱鹏;朱长仁; - 《数字技术与应用》 - 2013-09-15	是否引证: 否
4	基于多尺度特征的视频图像拼接技术研究	4.7% (30)
	刘志刚(导师: 周石琳) - 《国防科学技术大学硕士论文》 - 2011-06-01	是否引证: 否
5	基于灰度序的描述子主方向估计方法研究	4.7% (30)
	李宏荣(导师: 李晓明) - 《山西大学硕士论文》 - 2013-06-01	是否引证: 否
6	基于仿射不变特征的无人机影像匹配研究	4.7% (30)
	史永凯(导师: 王行风) - 《中国矿业大学硕士论文》 - 2014-05-01	是否引证: 否

原文内容

第四章

4-1 角点是什么?斑点是什么?

参考答案: 在现实世界中, 角点对应于物体的拐角, 如道路的十字路口、丁字路口等; 斑点通常是指与周围有着颜色和灰度差别的区域, 如草原上的一棵树或一栋房子。

4-2 灰度直方图的横坐标表示?取值范围是?

参考答案: 灰度直方图的横坐标表示图像的灰度值, 取值范围是[0, 255]。

4-3 灰度直方图中灰度值越大, 则表明图像亮度()?

参考答案: 越大

4-4 灰度直方图中灰度覆盖范围越大, 则表明图像对比度()?

参考答案: 越大

4-5 常用的角点检测算法有哪些?

参考答案: Harris角点检测算法、FAST角点检测算法等。

4-6 斑点检测中DOG算法与LOG算法那个检测效率更高?

参考答案: DOG算法

4-7 SIFT的全称是什么?

参考答案: SIFT全称为尺度不变特征变换。

4-8 简述SIFT特征检测算法的步骤。

参考答案: SIFT算法分为4个阶段: (1) 尺度空间极值检测; (2) 特征点的定位; (3) 方向角度的确定; (4) 特征点的描述符。

4-9 简述SURF斑点检测算法的步骤。

参考答案: SURF算法包括特征点检测与特征点描述两个阶段, 特征点检测包括: (1) 基于Hessian矩阵的特征点检测; (2) 尺度空间表示; (3) 特征点定位。特征点描述包括: (1) 方向角度的分配; (2) 基于Haar小波的特征点描述符。

4-10 常用的图像匹配算法有哪些?

参考答案: SSD算法、NCC算法等。

27. 计算机视觉教材_第27部分

总字符数: 728

相似文献列表

去除本人文献复制比: 0% (0)

去除引用文献复制比: 0% (0)

文字复制比: 0% (0)

原文内容

第五章

5-1 基于图像处理技术的桥梁动态位移测量技术相比于传统接触式传感器和其他非接触式传感器具有哪些显著优点?

参考答案: 非接触式、精度较高、测量距离远、与传统的位移传感器相比, 单个相机可以同时跟踪多个点的结构位移等

5-2 视觉位移测量系统的基本组成包括?

参考答案: 相机、镜头、计算与分析终端等。

5-3 基于计算机视觉的位移测量方法一般有哪四个步骤?

参考答案: 相机标定、特征检测/目标识别、特征匹配/目标跟踪、结构位移计算。

5-4 目前广泛采用的位移换算关系标定方法有哪些?

参考答案: 测距法、平均法、方向向量法以及单应性矩阵法。

5-5 影响基于计算机视觉的位移测量精度的因素有哪些?

参考答案: 硬件因素、算法因素和环境因素等。

5-6 影响基于计算机视觉的位移测量精度的环境因素有很多, 请举例说明。

参考答案: 相机自身抖动导致的画面晃动、相机自身抖动导致的图像模糊、雾气造成的视线遮挡, 及空气温度场不均匀导致的光线不均匀折射等。

5-7 光流法是基于计算机视觉的位移测量常用方法，请查阅相关资料并简述光流的概念。

参考答案：光流是指在连续的两帧图像中由于图像中的物体移动或者摄像头的移动导致的目标像素的移动，光流是二维矢量场，表示一个点从某一帧到另一帧的位移。

第十一章

11-1 结构位移监测与位移检测的区别是什么？

参考答案：结构位移监测具有长期性，需要时时刻刻的数据，数据量巨大；位移检测具有短期性，是某一时刻的位移测量，数据量小。

11-2 结合课程所学与本章案例，提出针对桥梁结构位移监测的具体方案。

参考答案：可结合多视角转动摄像头、激光雷达、无人机、巡检机器人等进行具体作答，合理即可。

表格检测详细结果

- 说明：1. 总文字复制比:被检测文献总重复字符数在总字符数中所占的比例
2. 去除引用文献复制比:去除系统识别为引用的文献后, 计算出来的重合字符数在总字符数中所占的比例
3. 去除本人文献复制比:去除系统识别为作者本人其他文献后, 计算出来的重合字符数在总字符数中所占的比例
4. 单篇最大文字复制比:被检测文献与所有相似文献比对后, 重合字符数占总字符数比例最大的那一篇文献的文字复制比
5. 复制比按照“四舍五入”规则, 保留1位小数; 若您的文献经查重检测, 复制比结果为0, 表示未发现重复内容, 或可能存在的个别重复内容较少不足以作为判断依据
6. 红色文字表示文字复制部分; 绿色文字表示引用部分(包括系统自动识别为引用的部分); 棕灰色文字表示系统依据作者姓名识别的本人其他文献部分
7. 系统依据您选择的检测类型(或检测方式)、比对截止日期(或发表日期)等生成本报告
8. 知网个人查重唯一官方网站: <https://cx.cnki.net>