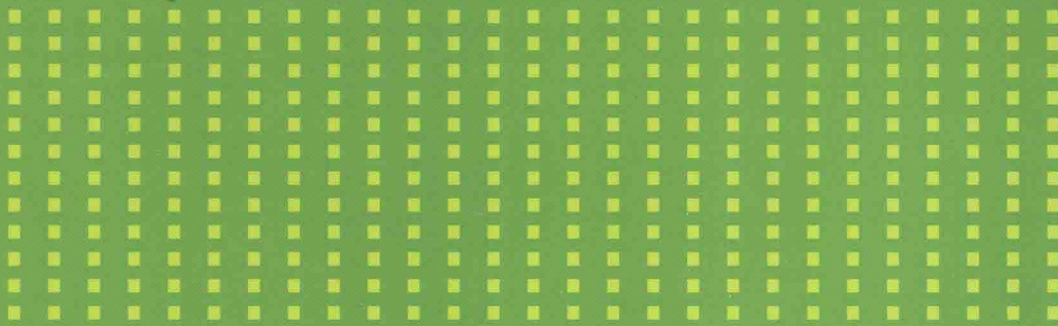




基于结构光的 计算机视觉



韩 成 杨华民 蒋振刚 范静涛 张超 杨帆 著

JIYU JIEGOU GUANG DE
JISUANJI SHIJUE



国防工业出版社
National Defense Industry Press

基于结构光的计算机视觉

韩成 杨华民 蒋振刚 范静涛 张超 杨帆 著

国防工业出版社

· 北京 ·

内 容 简 介

计算机视觉研究的主要目的是让计算机能够利用图像和图像序列来识别和认知三维世界,其最终目标是让计算机具有“视觉”功能,以满足社会对计算机高级应用的需求。本书根据研究内容把计算机视觉划分为计算理论、表达与算法和系统实现三个层次,系统地阐述了计算机视觉各个方面的相关理论和应用技术基础,包括计算机视觉的基本原理、分类、最新的国内外研究进展,重点介绍了基于结构光的三维重建与测量的方法,如摄像机和投影机的标定技术、结构光的编解码技术、点云的计算与拼接技术,最后介绍了基于结构光的三维重建与测量技术的应用实例。

本书适合作为从事计算机视觉领域科研工作者的一本基础技术参考书和最新研究成果参考文献,也可作为高等院校计算机及相关专业的高年级本科生和研究生课程教材或参考书。

图书在版编目(CIP)数据

基于结构光的计算机视觉 / 韩成等著. —北京:
国防工业出版社, 2015. 1
ISBN 978-7-118-09818-1

I. ①基… II. ①韩… III. ①计算机视觉—研究
IV. ①TP302.7

中国版本图书馆 CIP 数据核字(2014)第 293197 号

※

国防工业出版社出版发行
(北京市海淀区紫竹院南路 23 号 邮政编码 100048)
北京嘉恒彩色印刷有限责任公司
新华书店经售



*

开本 710 × 1000 1/16 印张 10 1/4 字数 190 千字
2015 年 1 月第 1 版第 1 次印刷 印数 1—2000 册 定价 40.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777
发行传真:(010)88540755

发行邮购:(010)88540776
发行业务:(010)88540717

前 言

近几十年来,随着计算机视觉、激光技术、精密光学技术、视频采集技术和图像处理技术的发展,基于结构光的主动视觉物体尺寸和三维轮廓测量法应运而生。该方法具有无接触、速度快、准确度高、自动化程度高等一系列优点,广泛应用于机器人导航、视觉监控、建筑制造、模具制造、线上品质检查等行业,在最近几年兴起的计算机虚拟现实、基于图像的绘制、三维动画等技术中也都有着广泛的应用。

本书共分为9章,系统、全面地阐述了基于结构光的计算机视觉技术,各章的主要内容如下:

第1章概要介绍计算机视觉的基本原理、特点、分类,分析了计算机视觉与相关领域的关系和国内外研究进展;

第2章介绍了基于结构光的计算机视觉技术的基本原理、关键技术、特点、分类和应用领域,分析了基于结构光的计算机视觉技术的研究进展;

第3~6章围绕基于结构光的计算机视觉中的关键技术,分别介绍了系统标定技术、编解码技术、点云拼接技术的原理,最新研究现状和比较经典的方法;

第7章介绍了基于格雷码结构光的三维重建实例,重点介绍了作者在格雷码结构光中编解码技术的研究成果;

第8章介绍了基于De Bruijn彩色结构光的三维重建实例,重点介绍了一种新的彩色结构光解码方法,并对系统的性能进行了分析;

第9章介绍了基于彩色结构光的智能投影实例,重点介绍了联合标定方法、一种新的彩色结构光编解码方法和畸变投影图像的几何校正方法。

本书的主要特色归纳如下:

- 内容广泛,深入浅出,取材新颖精练,重点突出,并以解决实际问题为目的,很多内容参考了近期的研究成果,包括本书作者申报或授权的国家发明专利以及发表的高水平学术论文;

- 书中列出了很多算法,都以函数或者过程的形式给出,读者只需用自己熟悉的编程语言稍加修改,即可实现这些算法;

- 书中的大量实例贴近生活,面向应用;

- 在每章后面列出了大量参考文献,全面覆盖各个子领域的最新研究进展和成果,这些文献不仅能够帮助读者巩固所学的内容,而且方便读者在自己感兴趣的方向上进行更深入的研究;

- 为了便于读者将理论与实现进行对照和进行自己的应用系统设计,本书还列举了大应用实例。

鉴于作者的水平所限,书中难免会有不当之处,希望广大读者批评指正。

编著者

2014 年 5 月

目 录

第 1 章 绪论	1
1.1 人类视觉系统的构成与视觉机理	1
1.1.1 人眼的结构	1
1.1.2 视觉系统	2
1.1.3 颜色	3
1.1.4 视知觉对深度的感知	4
1.1.5 CCD 摄像机	5
1.2 计算机视觉的基本原理及系统构成	6
1.3 计算机视觉研究的特点与分类	6
1.3.1 计算机视觉研究的特点	7
1.3.2 计算机视觉的分类	7
1.4 计算机视觉与相关领域的关系	8
1.5 计算机视觉的研究现状	9
参考文献	10
第 2 章 基于结构光的计算机视觉	11
2.1 基于结构光的计算机视觉的基本原理	11
2.2 基于结构光的计算机视觉中的关键技术	12
2.2.1 系统参数的标定技术	12
2.2.2 结构光的编解码技术	12
2.2.3 点云计算与拼接技术	13
2.3 基于结构光的计算机视觉技术的特点和分类	14
2.3.1 基于结构光的计算机视觉技术的特点	14
2.3.2 基于结构光的计算机视觉技术的分类	15
2.4 基于结构光的计算机视觉技术的研究进展	16
2.5 基于结构光的计算机视觉技术的应用领域	18
参考文献	22

第3章 摄像机和投影机的标定	23
3.1 空间几何变换	23
3.1.1 三维比例变换(基点为原点)	24
3.1.2 三维旋转变换	25
3.1.3 三维平移变换	26
3.2 摄像机的几何模型	26
3.2.1 常用坐标系	26
3.2.2 针孔照相机	28
3.2.3 透视投影矩阵的获得	29
3.2.4 摄像机参数的获取	31
3.2.5 径向畸变的校正	33
3.3 标定技术的研究现状	34
3.4 典型摄像机标定方法	36
3.4.1 基于平面靶标的摄像机标定	36
3.4.2 基于3D立体靶标的摄像机标定	37
3.4.3 基于径向约束的摄像机标定	37
3.4.4 摄像机自标定技术	39
3.5 典型投影机标定方法	40
3.5.1 基于相位技术的标定方法	40
3.5.2 基于反向成像的标定方法	40
3.5.3 基于交比不变性标定方法	42
参考文献	43
第4章 结构光编码策略与方法	46
4.1 结构光编码策略	46
4.2 结构光编码技术的研究现状	47
4.3 典型的结构光编码方法	51
4.3.1 RGB颜色格雷码结构光编码方法	51
4.3.2 M阵列编码方法	52
4.3.3 时间空间混合结构光编码方法	53
4.3.4 四位二进制彩色编码方法	55
参考文献	57
第5章 结构光解码技术	59
5.1 结构光解码技术基本原理	59

5.1.1	图像噪声的介绍与分析	59
5.1.2	常用的滤波算法	60
5.1.3	图像阈值分割	64
5.1.4	图像边缘检测	65
5.2	结构光解码技术的研究现状	67
5.3	典型的结构光解码方法	70
5.3.1	RGB 颜色格雷码图像的解码方法	70
5.3.2	M 阵列编码图像的解码方法	71
5.3.3	时间空间混合结构光编码图像的解码方法	71
5.3.4	四位二进制彩色编码图像的解码方法	72
	参考文献	73

第 6 章 点云的计算与拼接技术 74

6.1	三维点云计算方法	74
6.2	特征标志点的处理	75
6.2.1	图像标志点的跟踪与识别	75
6.2.2	图像标志点点心的提取	76
6.3	三维特征标志点空间搜索与识别	78
6.3.1	特征点提取与识别算法研究	78
6.3.2	空间特征点动态分层	78
6.3.3	特征点搜索与识别	79
6.4	点云拼接的国内外研究现状	81
6.4.1	基于标志点的点云拼接发展现状	82
6.4.2	基于稳健几何特征的点云拼接发展现状	83
6.5	典型的点云拼接算法	85
6.5.1	多视标签定位拼接算法	85
6.5.2	三点组合坐标变换法	90
6.5.3	最小二乘法介绍	93
6.5.4	基于三点矩阵奇异值分解的最小二乘法	94
6.5.5	基于多点的最小二乘法	95
6.6	新的点云拼接算法	96
6.6.1	基于公共标志点的三维点云拼接算法	96
6.6.2	基于标志点的三维点云自动拼接算法	102
6.6.3	基于动态分层的最近点迭代拼接算法研究	107
	参考文献	113

第7章 格雷码结构光的三维重建实例	114
7.1 概述	114
7.2 系统构成	114
7.3 摄像机和投影机的标定	115
7.4 格雷编码结构光编码方法	115
7.5 格雷编码结构光编码图像的解码方法	117
7.5.1 图像比例放大	117
7.5.2 图像灰度化	118
7.5.3 图像二值化	118
7.5.4 图像条纹边界提取	118
7.5.5 实验结果分析	119
7.6 点云的计算	120
参考文献	121
第8章 基于彩色结构光的三维模型生成方法	123
8.1 概述	123
8.2 系统参数的标定	124
8.3 基于7元3级迪布鲁英序列的编码方法	126
8.4 新的 De Bruijn 彩色结构光解码方法	127
8.4.1 编码图像的预处理	127
8.4.2 彩色结构光的颜色聚类化方法	131
8.4.3 将 De Bruijn 彩色结构光图像转化为赋权有向图模型及其简化方法	135
8.4.4 编码特征点的双梯度动态规划匹配方法	137
8.5 系统的性能分析	140
参考文献	142
第9章 基于彩色结构光曲面投影方法	143
9.1 概述	143
9.2 摄像机和投影机联合标定方法	144
9.3 具有拓扑结构和内部结构相关性的彩色结构光编码方法	145
9.4 基于波形分析的解码方法	147
9.5 结合观察视口的畸变投影图像几何校正方法	150
参考文献	155

第 1 章 绪 论

1.1 人类视觉系统的构成与视觉机理

人眼视网膜中的光感受器单元(cones)和神经节细胞(ganglion cell)的分布是非均匀的,在中央凹部分,即窝区(fovea)高度密集而周边稀疏,这种特性使视觉信息的获取是非均匀的。中央凹部分具有很高的视敏度,但稍微偏离一点,视敏度则急剧下降。离开中央凹 2° ,视敏度将下降 50%。因此,人眼能够精细地识别物体的地方,只限于视网膜中央凹附近。为此眼球需要不断地上下左右运动,将感兴趣的目标保持在中央凹附近。视网膜周边部分虽然对识别细小的物体不起作用,但对给光撤光以及物体运动等刺激是敏感的,如果外界的一部分急速发生变化,则与此对应的视网膜周边部分起反应,眼球就产生运动,使视轴朝向那个方向。可见,人眼在具有广阔视野的同时又具有局部的高分辨能力,可以使人在对感兴趣的目标保持高分辨率的同时,又能对视野的其他部分保持警戒。

1.1.1 人眼的结构

人的眼睛近似球形,眼球包括眼球壁、内容物、神经、血管等组织。眼球壁主要分为外、中、内三层。

外层由角膜、巩膜组成。前 $1/6$ 为透明的角膜,其余 $5/6$ 为白色的巩膜,俗称“眼白”。眼球外层起维持眼球形状和保护眼内组织的作用。角膜是眼球前部的透明部分,光线经此射入眼球。巩膜不透明,呈乳白色,质地坚韧。

中层具有丰富的色素和血管,包括虹膜、睫状体和脉络膜三部分。虹膜呈环圆形,位于晶状体前。不同种族人的虹膜颜色不同。中央有一 $2.5 \sim 4\text{mm}$ 的圆孔,称瞳孔。睫状体前接虹膜根部,后接脉络膜,外侧为巩膜,内侧则通过悬韧带与晶状体相连。脉络膜位于巩膜和视网膜之间。脉络膜的血循环营养视网膜外层,其含有的丰富色素起遮光暗房作用。

内层为视网膜,是一层透明的膜,也是视觉形成的神经信息传递的最敏锐的区域。视网膜所得到的视觉信息,经视神经传送到大脑。眼内容物包括房水、晶状体和玻璃体。房水由睫状突产生,有营养角膜、晶状体及玻璃体,维持眼压的作用。晶状体为富有弹性的透明体,形如双凸透镜,位于虹膜、瞳孔之后玻璃体之前。玻璃体为透明的胶质体,主要成分为水。玻璃体有屈光作用,也起支撑视

网膜的作用。

1.1.2 视觉系统

人类视觉系统只有三种视锥细胞,因此在缤纷的世界中,即使面对似锦的繁花,我们也可能犹如色盲,常常对一些色彩“视而不见”。人类对光的感知是依靠视网膜细胞。圆锥细胞负责感知光度(较强光)和色彩,杆状细胞仅能感知光度,不能感知颜色,但其对光的敏感度是圆锥细胞的一万倍。在微弱光环境下杆状细胞起主要作用,因此我们不能在暗环境中分辨颜色。一些数码相机的夜光拍摄模式也模拟了这一特性。

视网膜中三种圆锥细胞有重叠的频率响应曲线,但响应强度有所不同,它们分别对红(570nm)、绿(535nm)、蓝(445nm)光最敏感,共同决定了色彩感觉。光度正比于视网膜细胞接收到的光强度能量,但人类对相同强度不同波长的光具有不同的敏感度。可感知的波长范围 380~780nm,称为可见光,其中对绿色光产生最大的光强敏感度。

眼睛的空间分辨能力,即视力,通常用可分辨视角(degree)倒数为单位。正常人的最少可辨视觉阈值约 0.5s,最大视觉范围 200°(宽)×135°(高)。在一般室内光强下,人眼对时间频率的响应近似一个带通滤波器,对 15~20Hz 信号最敏感,有很强闪烁感,大于 75Hz 响应为 0,闪烁感消失。刚到达闪烁感消失的频率叫做临界融合频率(CFF)。在较暗的环境下,呈低通特性,且 CFF 会降低,这时对 5Hz 信号最敏感,大于 25Hz 闪烁基本消失。电影院环境很暗,放映机的刷新率为 24Hz 也不感到闪烁,这样可以减少胶卷用量和机器的转速。电脑显示器亮度较大,需要 75Hz 闪烁感才消失。闪烁消失后,亮度感知等于亮度时间平均值(塔鲁伯法则)。这种低通特性,也可以解析为视觉暂留特性,即当影像消失/变化时,大脑中的影像不会立刻消失,而是保留一个短暂时间。生活中常感受到的动态模糊,运动残像也和这个有关。

观察一个运动物体,眼球会自动跟随其运动,这种现象叫随从运动(eye pursuit movement)。这时眼球和物体的相对速度会降低,我们能更清晰地辨认物体。例如观看球类比赛(如棒球),尽管棒球的运动速度很快,由于随从运动,我们仍能够看得到球的大概样子(但会有运动模糊)。如果我们把眼睛跟着风扇转动方向转动,会发现对扇叶细节看得较清楚。眼球随从最大速度为 4~5°/s,因此我们不可能看清楚一颗子弹飞行。

定义人眼的对空间感觉的角度频率为 cpd(cycle/degree),表示眼球每转动 1°扫过的黑白条纹周期数。对给定的条纹,这个值与人眼到显示屏的距离有关,对于同样大小的屏幕,离开越远,cpd 越大。通常人眼对空间的感觉相当于一个带通滤波器。最敏感在 2~5cpd,空间截止频率为 30cpd。比如我们看油画和电视机屏幕时,当距离一定远,cpd 增大,人的眼睛就分辨不了像素点细节,便感觉

不到颗粒感了。

一般情况下,当人观察一个静止影像时,眼球不会静止一处,通常停留在一处几百毫秒完成取像后,移到别处取像,如此持续不断。这种运动称为跳跃性运动。研究表明跳跃性运动可以增大对比敏感度,但敏感度峰值却减小。

1.1.3 颜色

颜色是光波作用于人眼所产生的视觉体验。在日常生活中,颜色有广义和狭义之分。广义的颜色包括非颜色(白色、黑色和各种不同程度的灰度)和彩色;狭义的颜色就是指彩色。

不同波长的光线(电磁波)是无色的,但当它们刺激人的视觉系统时就产生了颜色视觉。因此,颜色是视觉系统接收光刺激后的产物。物体之所以显现颜色,就是因为它们反射光线进入到我们的视觉系统。光波的三种特性:振幅、波长和纯度,分别决定了人眼视觉的色调、明度和饱和度。这三个维度可以构成一个颜色立体模型,所有的颜色体验都可以从这三个维度来描述。

色调与其物理刺激的光波波长相对应。对光源来说,占优势的波长不同,色调也就不同,例如,长波占优势,光线就成红色或橘色,如果短波占优势,则光线呈蓝色或绿色。对物体表面来说,色调取决于物体表面对不同波长光线的选择性反射。如果反射光中 700nm 的波长占优势,则物体看上去是红色,如果反射光中 510nm 左右的波长占优势,则物体看上去是绿色。

明度与其物理刺激的光波强度相对应,指颜色的明暗程度。颜色的明度取决于照度和物体表面的反射系数。照明强度越强,物体表面反射系数越大,则颜色越明亮。例如,同样是 700nm 的红光,强度大的就比强度小的看上去更亮一些,同样是反射 510nm 的绿色纸,皱纹纸就不如蜡光纸看上去明亮。

饱和度与其物理刺激的光波纯度相对应,指某种颜色的纯杂度或鲜明程度。可见光谱上,各种单色光的饱和度最大,如鲜红、鲜绿等。各种单色光掺入白色就变得不再饱和,掺入的白色越多就越不饱和,如粉红、橘黄、黄褐等。黑白之间的各种不同程度的灰色是完全不饱和的颜色,它们根本没有色调,只有明度属性,因此又称为黑白系列。

不同颜色是可以相混合的。颜色混合分为加法颜色混合和减法颜色混合。在视觉系统中进行的不同波长的色光的混合,遵从“加法原则”;而不同颜色的颜料的混合不是在视觉系统中进行,是颜料自身的混合,遵从“减法原则”。

人们对颜色会产生视觉后像,即当颜色刺激移去后,仍有颜色感觉的现象。颜色视觉后像通常都是负后像,即后像的颜色与原刺激的颜色互为补色。后像持续的时间与原刺激作用的时间有关,刺激作用时间越长,产生的后像持续时间越长。对颜色后像的一种解释是,由于长时间注视一种颜色,对这种颜色的感受性大大降低,这时再注视白色背景,就相当于由白光中减去原注视颜色,因此产

生了补色的感觉。颜色视觉后像主要发生在外周视觉系统。

颜色对比是指不同颜色的物体并列或相继出现时,引起的色觉与单一颜色时不同。无彩色对比(黑白色)的结果是引起明度感觉的变化,觉得黑者更黑,白者更白;彩色对比使颜色向其背景颜色的补色变化。颜色对比与明度恒常性、颜色恒常性是相联系的。

1.1.4 视知觉对深度的感知

当外界物体反射来的光线带着物体表面的信息经过角膜、房水,由瞳孔进入眼球内部,经聚焦在视网膜上形成物像。物像刺激了视网膜上的感光细胞,这些感光细胞产生的神经冲动,沿着视神经传入到大脑皮层的视觉中枢,即大脑皮层的枕叶部位,在这里把神经冲动转换成大脑中认识的景象。这些景象的生成已经经过了加工,是“角度感”、“形象感”、“立体感”等协同工作,并把图像根据摄入的信息在大脑虚拟空间中还原,还原等于把图像往外又投了出去。虚拟位置能大致与原实物位置对准,这才是我们所见到的景物。

深度知觉又称距离知觉或立体知觉。这是个体对同一物体的凹凸或对不同物体的远近的反映。视网膜虽然是一个二维的平面,但人不仅能感知平面的物体,而且还能产生具有深度的三维空间的知觉,这主要是通过双眼视觉实现的。有关深度知觉的线索,既有双眼视差、双眼辐合、水晶体的调节、运动视差等生理的线索,也有对象的重叠、线条透视、空气透视、对象的纹理梯度、明暗和阴影以及熟悉物体的大小等客观线索。根据自己的经验和有关线索,单凭一只眼睛观察物体也可以产生深度知觉。用视觉来知觉深度,是以视觉和触摸觉在个体发展过程中形成的联系为基础的。通过大脑的整合活动就可作出深度和距离的判断。但个体在知觉对象的空间关系时,并不完全意识到上述那些主、客观条件的作用。

关于物体远近距离或深度的知觉叫深度知觉,也叫距离知觉。它对于人们判断客体间的空间关系是十分重要的。知觉深度需根据外部环境和机体内部的许多线索进行。在视觉上,这些线索可以分为单眼线索和双眼线索。具体地讲:

(1) 单眼线索。单眼线索指仅凭一只眼睛的视觉即可提供的线索。单眼线索主要是静态线索,即环境及观察对象的物理特性或现象,也包括一些单眼的运动线索。在绘画中,利用静态的单眼线索能够在二维的平面上产生三维的立体效果,因此这些单眼线索又称为图形线索。单眼线索主要包括插入、空气透视、阴影、线条透视、结构级差、相对大小、熟悉大小、运动视差、调节。

(2) 双眼线索。单眼线索能够为人们提供大量的深度信息,使人们完成许多视觉指导下的操作任务。然而,有些深度信息是需要双眼的共同作用的。双眼线索是必须由两只眼睛共同作用而提供的线索,主要包括辐合和双眼视差。

在实际生活中,各种深度线索是联合起作用的。而且,不仅是视觉,听觉、触

觉、动觉之间也相互协调,形成一定的联系,确保人们获得有效的深度知觉。

1.1.5 CCD 摄像机

电荷耦合器件(Charge Coupled Device, CCD)图像传感器,它使用一种高感光度的半导体材料制成,能把光线转变成电荷,通过模数转换器芯片转换成数字信号,数字信号经过压缩以后由相机内部的闪速存储器或内置硬盘卡保存,因而可以轻而易举地把数据传输给计算机,并借助于计算机的处理手段,根据需要和想象来修改图像。

CCD 传感器是一种新型光电转换器件,它能存储由光产生的信号电荷。当对它施加特定时序的脉冲时,其存储的信号电荷便可在 CCD 内作定向传输而实现自扫描。它主要由光敏单元、输入结构和输出结构等组成。它具有光电转换、信息存储和延时等功能,而且集成度高、功耗小,已经在摄像、信号处理和存储三大领域中得到广泛的应用,尤其是在图像传感器应用方面取得令人瞩目的发展。

一般认为,CCD 传感器有以下优点:

(1) 高解析度:像点的大小为微米级,可感测及识别精细物体,提高影像品质。

(2) 低噪声高敏感度:CCD 具有很低的读出杂信和暗电流杂信,因此提高了信噪比,同时又具高敏感度,很低光度的入射光也能侦测到,其信号不会被掩盖,使 CCD 的应用较不受天候拘束。

(3) 动态范围广:同时侦测及分辨强光和弱光,提高系统环境的使用范围,不因亮度差异大而造成信号反差现象。

(4) 良好的线性特性曲线:入射光源强度和输出信号大小成良好的正比关系,物体资讯不致损失,降低信号补偿处理成本。

(5) 高光子转换效率:很微弱的入射光照射都能被记录下来,若配合影像增强管及投光器,即使在暗夜远处的景物仍然还可以侦测得到。

(6) 大面积感光:利用半导体技术已可制造大面积的 CCD 晶片,目前与传统底片尺寸相当的 35mm 的 CCD 已经开始应用在数码相机中,成为取代专业光学相机的关键元件。

(7) 光谱响应广:能检测很宽波长范围的光,增加系统使用弹性,扩大系统应用领域。

(8) 低影像失真:使用 CCD 感测器,其影像处理不会有失真的情形,使原物体资讯真实地反应出来。

(9) 体积小、重量轻:CCD 具备体积小且重量轻的特性,因此,可容易地装置在人造卫星及各式导航系统上。

(10) 低耗电力:不受强电磁场影响。

(11) 电荷传输效率佳:该效率系数影响信噪比、解像率,若电荷传输效率不

佳,影像将变得较模糊。

(12) 可大批量生产,品质稳定,坚固,不易老化,使用方便及保养容易。

1.2 计算机视觉的基本原理及系统构成

计算机视觉就是用各种成像系统代替视觉器官作为输入敏感手段,由计算机来代替大脑完成处理和解释,如图 1.1 所示。计算机视觉的最终研究目标就是使计算机能像人那样通过视觉观察和理解世界,具有自主适应环境的能力。这是要经过长期的努力才能达到的目标。因此,在实现最终目标以前,人们努力的中期目标是建立一种视觉系统,这个系统能依据视觉敏感和反馈的某种程度的智能完成一定的任务。例如,计算机视觉的一个重要应用领域就是自主车辆的视觉导航,还没有条件实现像人那样能识别和理解任何环境,完成自主导航的系统。因此,人们努力的研究目标是实现在高速公路上具有道路跟踪能力,可避免与前方车辆碰撞的视觉辅助驾驶系统。这里要指出的一点是在计算机视觉系统中计算机起代替人脑的作用,但并不意味着计算机必须按人类视觉的方法完成视觉信息的处理。

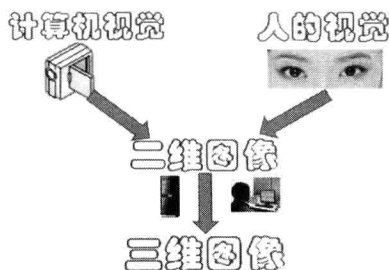


图 1.1 计算机与人的视觉

计算机视觉可以而且应该根据计算机系统的特点来进行视觉信息的处理。但是,人类视觉系统是迄今为止,人们所知道的功能最强大和完善的视觉系统。因此,用计算机信息处理的方法研究人类视觉的机理,建立人类视觉的计算理论,也是一个非常重要和引人感兴趣的研究领域。这方面的研究被称为计算视觉。计算视觉可被认为是计算机视觉中的一个研究领域。

1.3 计算机视觉研究的特点与分类

计算机视觉是一门关于如何运用照相机和计算机来获取我们所需的,被拍摄对象的数据与信息的学问。它的主要任务就是通过对采集的图片或视频进行处理以获得相应场景的三维信息,就像人类和许多其他类生物每天所做的那样。

形象地说,就是给计算机安装上眼睛(照相机)和大脑(算法),让计算机能够感知环境。我们中国人的成语“眼见为实”和西方人常说的“*One picture is worth ten thousand words*”表达了视觉对人类的重要性。

1.3.1 计算机视觉研究的特点

计算机视觉既是工程领域,也是科学领域中的一个富有挑战性重要研究领域。计算机视觉是一门综合性的学科,它已经吸引了来自各个学科的研究者参加到对它的研究之中,其中包括计算机科学和工程、信号处理、物理学、应用数学和统计学,神经生理学和认知科学等。

视觉是各个应用领域,如制造业、检验、文档分析、医疗诊断和军事等领域中各种智能/自主系统中不可分割的一部分。由于它的重要性,一些先进国家,例如美国把对计算机视觉的研究列为对经济和科学有广泛影响的科学和工程中的重大基本问题,即所谓的重大挑战。计算机视觉的挑战是要为计算机和机器人开发具有与人类水平相当的视觉能力。机器视觉需要图像信号、纹理和颜色建模、几何处理和推理,以及物体建模。一个有能力的视觉系统应该把所有这些处理都紧密地集成在一起。

1.3.2 计算机视觉的分类

如图 1.2 所示:根据测量方式的不同,计算机视觉可以分为非接触式和接触式;根据所采用光源的不同,非接触式计算机视觉可以分为主动式和被动式,其中主动式三维信息获取技术使用一个特定的光源(如结构光、激光等)对被测物体进行照明,使得被测物体在特定光源的照射下形成相应的特殊图像,该图像中包含了光源在被测物体上变形的信息,因此通过相应的解码运算可以得到被测物体的深度信息。由于这种方法具有较高的测量精度,因此大多数以深度测量为目的的三维成像系统都采用主动式测量技术。主要的主动式测量技术包括结构光法、简单三角法、激光测距法等。

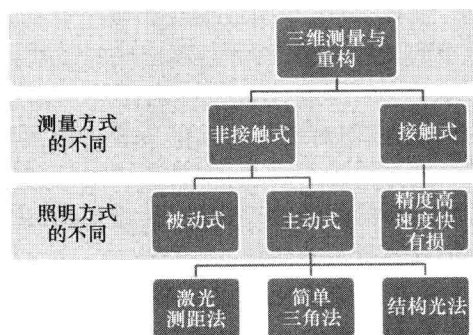


图 1.2 计算机视觉的分类

被动式三维信息获取技术不需要对场景中添加任何光照信息,被测物体仅由环境中本身的光源提供照明。被动式方法虽然不需要专门的光源,而且对成像设备要求也不高,但是对没有明显特征的图像(如无明显的边缘、角点等),计算量很大,匹配的准确度也难以保证。

1.4 计算机视觉与相关领域的关系

有不少学科的研究目标与计算机视觉相近或与此有关。这些学科中包括图像处理、模式识别或图像识别、景物分析、图像理解等,如图 1.3 所示。计算机视觉包括图像处理和模式识别,除此之外,它还包括空间形状的描述,几何建模以及认识过程。实现图像理解是计算机视觉的终极目标。

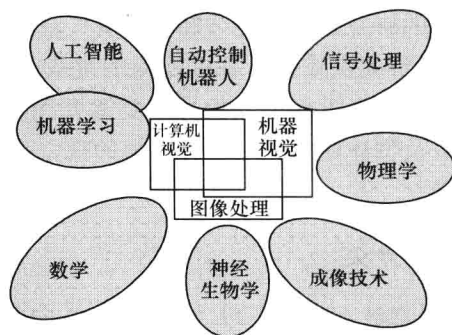


图 1.3 计算机视觉与其他领域的关系

(1) 图像处理。图像处理技术把输入图像转换成具有所希望特性的另一幅图像。例如,可通过处理使输出图像有较高的信噪比,或通过增强处理突出图像的细节,以便于操作员的检验。在计算机视觉研究中经常利用图像处理技术进行预处理和特征抽取。

(2) 模式识别。模式识别技术根据从图像抽取的统计特性或结构信息,把图像分成预定的类别。例如,文字识别或指纹识别。在计算机视觉中模式识别技术经常用于对图像中的某些部分,例如分割区域的识别和分类。

(3) 图像理解。给定一幅图像,图像理解程序不仅描述图像本身,而且描述和解释图像所代表的景物,以便对图像代表的内容作出决定。在人工智能视觉研究的初期经常使用景物分析这个术语,以强调二维图像与三维景物之间的区别。图像理解除了需要复杂的图像处理以外,还需要具有关于景物成像的物理规律的知识以及与景物内容有关的知识。

在建立计算机视觉系统时需要用到上述学科中的有关技术,但计算机视觉研究的内容要比这些学科更为广泛。计算机视觉的研究与人类视觉的研究密切相关。为实现建立与人的视觉系统相类似的通用计算机视觉系统的目标需要建