

人脸感知

Human Face Percetion: from 2D to 3D

从二维到三维

王国胤 龚 勋/著



科学出版社

本书由重庆市科学技术协会学术著作出版基金、
重庆邮电大学出版基金资助出版

人脸感知：从二维到三维

Human Face Percetion: from 2D to 3D

王国胤 龚 勋 著

科 学 出 版 社

北 京

内 容 简 介

本书围绕人脸这一研究对象,从计算机图像认知的角度,对二维人脸图像分析和三维人脸建模进行全面介绍,内容包括:人脸检测及特征点标定,人脸轮廓定位,三维人脸建模完整方案(包括人脸形状知识库的建立、人脸统计模型、基于特征点的形变算法、人脸建模具体方案、纹理映射),光照分析及姿态估计。最后,本书对一些经典算法以附录的形式加以详述,并提供源代码,以利于读者理解并尽快上手应用。

本书可作为信息处理、计算机、人工智能、模式识别、认知心理学等有关专业的高年级本科生、硕士生和博士生的学习参考用书,也可供以上领域的研究工作者参考。

图书在版编目(CIP)数据

人脸感知:从二维到三维=Human Face Percetion:from 2D to 3D/王国胤,龚勋著. —北京:科学出版社, 2011

ISBN 978-7-03-032294-4

I. 人… II. ①王…②龚… III. 计算机-信息处理-研究 IV. TP391.41

中国版本图书馆 CIP 数据核字(2011)第 182872 号

责任编辑:余 丁 杨 然 / 责任校对:何艳萍

责任印制:赵 博 / 封面设计:耕者

科 学 出 版 社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新科印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2011 年 9 月第 一 版 开本: B5(720×1000)

2011 年 9 月第一次印刷 印张: 12 1/2

印数: 1—2 500 字数: 234 000

定价: 50.00 元

(如有印装质量问题,我社负责调换)

前 言

地球上居住着 67 亿种族不同、肤色各异的人,人脸在人们日常的情感表达与交流中扮演着重要的角色,它是人类的喜、怒、哀、乐等复杂表情和语言的载体。对人脸的描述和刻画,长期以来一直得到人们的广泛关注,至今留存下了数量众多的历史人物肖像作品,其中不乏高品质的文化瑰宝。

当今信息技术日新月异,多方面的应用需求驱动着计算机领域对人脸的研究:影视中的应用首当其冲,从《阿凡达》、《指环王》和《变脸》等美国大片到国产电影《宝莲灯》等无不体现着人脸建模与动画技术的魅力,使用计算机技术生成具有真实感的虚拟人物及人脸动画具有广泛的应用前景;其次是安全领域的应用,包括视频监控、生物特征识别和身份认证等;此外,友好的人机交互需要计算机能够理解人的面部动作、行为、情感,并据此作出响应。

人脸感知可由人脸分析与人脸合成两部分构成。众所周知,人类自身对人脸非常熟悉和敏感,人类视觉系统(human visual system, HVS)具备分辨人脸复杂模式的能力,仅仅通过少量照片,就能准确、快速地在大脑中塑造其三维形状,进而实现身份辨别。然而,这种对人来说与生俱来的分析与合成能力对计算机而言却显得尤为困难,要实现计算机的自动三维人脸建模,现有技术往往需要大量的手动操作,而且效果远达不到应用要求。如何更好地借鉴人类的认知机理和相关数学理论的最新研究成果,建立新的计算模型和方法,实现从图像、视频中准确、高效地重建三维人脸模型是当前计算机视觉、人工智能等领域的一个关键问题。该问题的妥善解决,不仅对打破当前人脸识别技术瓶颈、提升视频设备的实用价值具有重要的意义,而且可为其他复杂对象的建模提供借鉴,对类似科学问题的解决具有重要的借鉴意义。

20 世纪 90 年代以来,随着信息技术的快速发展和需求的剧增,借助计算机系统强大的计算能力,并结合在认知科学、计算机视觉(computer vision)、机器学习等领域取得的研究成果,使计算机具有对人脸的感知能力,成为一个热门的研究课题,涌现了大量的研究成果,其中涉及人脸检测、人脸识别、特征提取、轮廓定位、三维重建等,研究方法和手段也逐渐丰富和成熟。近年来,已有一些较为详尽的综述及论著,但针对各个研究课题,如人脸识别、人脸检测等,这些成果都从二维图像的角度对人脸处理相关技术进行了论述。作为一个三维形体,研究人员已逐渐认识到利用人脸的三维信息是克服二维图像分析技术瓶颈(如光照、姿态变化)最有效的方法,但遗憾的是,目前还没有一部学术专著从二维平面到三维空间

对人脸图像机器感知的理论及方法进行系统的论述。因此,有必要将有关的研究资料整理成书,供相关研究人员参考。更重要的是,希望通过本书总结作者及课题组近年来在二维人脸图像分析、三维人脸建模研究中取得的成果,与国内外同行进行交流,以使人脸感知理论和技术得到更广泛的重视和研究,取得更丰硕的成果。

本书系统介绍二维人脸图像分析及三维人脸建模的相关理论、技术和方法,对理论框架和算法实现进行详尽的讨论,全书分为两部分(二维人脸图像处理及人脸三维建模,二者相辅相成),包含8章内容:

第1章对人脸感知问题进行概述,讨论人类视觉感知与机器视觉的关系,并对本书的内容及阅读的相关注意事项进行简要介绍。

接下来是本书核心内容的第一部分:二维人脸图像处理涉及的相关技术,如人脸检测、轮廓定位、特征分析等,包括第2~3章。

第2章重点介绍二维人脸图像的检测及特征点标定问题。首先概述人脸检测、特征点标定及人脸轮廓检测对人脸图像分析及建模等任务的重要作用,然后对相关技术、经典方法进行详细的介绍。

第3章讨论二维人脸识别问题。人脸识别领域已经出现了很多识别算法,该章主要对人脸识别技术的发展进行概述及分类总结,并详细介绍几种具有代表性的子空间人脸识别算法及其原理,同时阐述 Gabor 小波变换原理及其在人脸识别中的应用。

接下来是本书核心内容的第二部分:介绍基于二维图像进行人脸三维建模的关键技术,包括第4~8章。

第4章对三维人脸建模问题进行概述,对国内外当前研究现状进行分析,指出存在的问题和可行的方向;对相关三维人脸建模商业软件和三维人脸数据库进行讨论。

第5章讨论人脸形状知识库的创建,首先介绍标准化三维人脸库的概念和创建方法,然后讨论用统计学工具进行人脸形状知识的学习方法。

第6章对三维人脸形状变形相关理论进行讨论,重点介绍两种采用基于特征点的形变建模方法,具有速度快、鲁棒性强的特点,并给出人脸形状建模误差评价方法,展示形状建模效果。

第7章在上一章提出的形变算法基础上,主要讨论基于单张照片的三维人脸建模完整方案,包括形状建模和纹理映射两方面内容,最后展示通过一张照片进行人脸三维建模的效果,并进行新姿态图像合成。

第8章在人脸建模的基础上,给出人脸图像光照分析及姿态估计方法。对不同光照、不同姿态条件下的人脸识别问题,引入三维信息是从根本上解决人脸识别问题的突破口,三维人脸模型的应用使得人脸识别中的这些难点问题有了解决

的途径。该章在三维人脸模型的基础上,主要讨论单张人脸照片上人脸光照估计和补偿、人脸姿态估计等问题。

在附录中,我们对模式识别、计算机视觉领域常用的经典方法及基础工具进行详细介绍,并提供相应的源代码,这些技术包括 PCA 算法、LFA 算法、基于单高斯模型的肤色检测、混合高斯模型的参数优化及图像曲线法向量及曲率的数值计算方法。相信这些基础知识有助于初学者快速上手开展研究。

本书的顺利完成与课题组全体成员的辛勤工作密不可分,在此要首先感谢他们为本书所做的工作,邹建法、夏冉、何琨、罗飞、胡同宇、徐宁、熊黎丽、叶超、陈巍、刁显峰、任文彬、卢显利、陈丹、刘文耀等研究生以及西南交通大学的靳蕃教授、李天瑞教授、诸昌铃教授,四川大学的李志蜀教授,重庆大学的李祖枢教授、廖晓峰教授都对本课题的研究给予了大量的支持和有益的建议。在此,对他们的帮助表示衷心的感谢!本书还引用了国内外同行专家学者的一些研究成果,在此对他们表示深深的谢意!

课题组的研究工作、本书的写作完成和出版,获得了重庆市科学技术协会学术著作出版基金、重庆邮电大学出版基金的资助,此外还得到了重庆市杰出青年科学基金(No. 2008BA2041)、西南交通大学科技发展基金、西南交通大学青年教师起步项目(2009Q86)、中央高校基本科研业务费专项(2009QK17)等项目的资助;本书中出现的 BJUT-3D 三维人脸库来自北京工业大学多媒体技术与图形学实验室,CAS-PEAL 人脸数据库来自中国科学院计算技术研究所-银晨科技面像识别联合实验室,在此一并表示感谢!

由于作者水平有限,加之时间仓促,部分内容是课题组的阶段性研究成果,不妥及疏漏之处在所难免,希望读者批评指正。

作者

2011年8月

目 录

前言

第 1 章 人类视觉感知概述	1
1.1 人脸感知	1
1.2 人类视觉认知	2
1.2.1 视觉的地位	2
1.2.2 视觉的内涵	2
1.2.3 视觉的特点	4
1.2.4 人类视觉如何识别人脸	5
1.3 计算机视觉	10
1.4 机器视觉与人类视觉的联系	12
1.5 本书各章内容简介及阅读本书要注意的问题	13

第一部分 二维人脸图像处理

第 2 章 人脸及其特征的自动检测	17
2.1 概述	17
2.1.1 人脸检测	17
2.1.2 人脸特征点自动标定	18
2.1.3 人脸轮廓检测	20
2.2 图像处理基本技术	23
2.2.1 滤波	23
2.2.2 直方图	24
2.2.3 图像归一化	24
2.2.4 边缘检测	25
2.2.5 形态学运算	26
2.3 基于 Adaboost 的人脸检测	26
2.3.1 基本概念	26
2.3.2 算法实现	28
2.4 主动形状模型 ASM	30

2.4.1	特征点选择	30
2.4.2	标定点对齐	31
2.4.3	形状建模	32
2.4.4	局部灰度纹理建模	32
2.4.5	ASM 目标搜索过程	33
2.5	主动外观模型 AAM	35
2.5.1	形状、纹理建模	35
2.5.2	统计表观建模	36
2.5.3	AAM 拟合算法	37
2.5.4	ASM 与 AAM 小结	39
2.6	人脸轮廓检测	40
2.6.1	活动轮廓模型简介	40
2.6.2	Chan-Vese 模型	44
2.7	小结	48
第 3 章	人脸特征提取及识别	49
3.1	人脸识别技术的发展	50
3.2	人脸识别原理及方法	52
3.2.1	子空间识别算法	52
3.2.2	Gabor 小波	57
3.3	动态主成分分子空间的构造	61
3.3.1	基于 PCA 的人脸重建分析	61
3.3.2	动态主成分分子空间构造算法	63
3.3.3	基于 Gabor 特征的动态主成分分析算法	65
3.4	人脸图像的增强 Gabor 特征构造与应用	65
3.4.1	增强 Gabor 特征的构造	65
3.4.2	DF_LDA+EGF 算法	69
3.4.3	最近邻欧氏距离分类器	71
3.5	小结	71

第二部分 人脸三维建模

第 4 章	人脸三维建模概述	75
4.1	背景	75
4.2	生活中的三维人脸	76

4.3 人脸三维建模技术发展	77
4.3.1 基于几何数据的人脸三维建模	77
4.3.2 基于图像的人脸三维建模	79
4.3.3 人脸动画驱动模型分类	82
4.4 三维人脸数据库	85
4.5 商业三维人脸建模软件	85
4.5.1 新加坡的面部识别公司 XID Technologies 的三维人脸识别系统	86
4.5.2 美国 Bioscrypt 三维人脸识别产品 3D FastPass™ Face Reader	86
4.5.3 加拿大 Singular Inversions 公司的 FaceGen Modeller	87
4.6 小结	87
第5章 标准化三维人脸库	88
5.1 三维人脸数据库标准化基本概念	88
5.2 基于特征区域分片的三维人脸数据重采样	90
5.2.1 三维人脸区域分块	90
5.2.2 基于平面模板的网格重建	94
5.3 基于 AAM 特征点定位的三维人脸数据重采样	97
5.3.1 基本思路	97
5.3.2 平面模板	98
5.4 标准化人脸的应用	99
5.4.1 标准人脸库及其线性运算	99
5.4.2 多分辨率人脸模型	100
5.4.3 混合分辨率人脸模型	101
5.5 基于人脸形状统计模型	101
5.6 小结	104
第6章 三维人脸形状建模	105
6.1 基于图像的人脸建模概述	105
6.2 基于动态成分的形变模型	106
6.2.1 形状系数的优化求解	106
6.2.2 主成分的动力选择	107
6.3 基于 Sibson 坐标的局部特征分析	109
6.3.1 局部特征分析理论基础	109
6.3.2 基于 LFA 的局部形变模型	111
6.4 双重形变模型	113

6.5	形状建模性能分析	114
6.5.1	误差估计函数	114
6.5.2	建模实例	115
6.6	小结	116
第7章	根据单张照片进行三维人脸建模	117
7.1	引言	117
7.2	两步人脸建模方案	118
7.2.1	基本思路	118
7.2.2	特征点深度值估计	119
7.3	纹理映射	123
7.3.1	纹理映射方法分类	124
7.3.2	基于调和映射的纹理映射	125
7.3.3	基于正面人脸照片的纹理映射	128
7.4	建模效果	130
7.5	小结	131
第8章	光照分析及姿态估计	132
8.1	引言	132
8.2	基于中性人脸球面谐波模型的光照估计及补偿	132
8.2.1	球面谐波理论	133
8.2.2	光照估计	134
8.2.3	光照补偿	136
8.3	基于线性回归的人脸姿态估计	138
8.3.1	姿态估计概述	138
8.3.2	基于特征点对的线性回归模型	139
8.4	小结	143
附录A	PCA 算法及其在人脸识别中的应用	144
A.1	问题描述	144
A.1.1	K-L 变换	144
A.1.2	利用 PCA 进行人脸识别	145
A.2	PCA 的理论基础	146
A.2.1	投影	146
A.2.2	PCA 的作用及其统计特性	146

A. 2. 3 特征脸	147
A. 2. 4 图片重建	148
A. 2. 5 奇异值分解(SVD)	148
A. 2. 6 利用小矩阵计算大矩阵特征向量	149
A. 3 PCA 相关源代码(matlab 版)	150
A. 3. 1 人脸识别 FaceRec. m	150
A. 3. 2 特征人脸识别 Eigface. m	152
A. 3. 3 人脸重建 Reconstruction. m	154
附录B LFA 理论	156
B. 1 理论描述	156
B. 2 LFA 相关源代码	157
B. 2. 1 LFA 核 K 与逆核 K^{-1}	157
B. 2. 2 LFA 重建	158
B. 2. 3 LFA 的核 K 图及差 P 图	159
附录C 基于单高斯模型的肤色检测及高斯混合模型参数优化	160
C. 1 高斯混合模型概述	160
C. 1. 1 单高斯模型	160
C. 1. 2 高斯混合模型	161
C. 2 采用 EM 估计 GMM 的参数	162
C. 2. 1 初始值	162
C. 2. 2 算法流程	163
C. 3 源码	164
C. 3. 1 基于单高斯模型 SGM 的人脸肤色检测	164
C. 3. 2 高斯混合模型	167
附录D 平面曲线法向量及曲率的数值计算	172
D. 1 法向量的计算	172
D. 1. 1 封闭曲线	172
D. 1. 2 非封闭曲线	173
D. 2 曲率的计算	174
参考文献	176

第1章 人类视觉感知概述

1.1 人脸感知

人脸在人们日常交往的情感表达中扮演着重要的角色,是人类的喜怒哀乐等复杂情感的载体。长期以来对人脸的描述和刻画一直得到了人们的广泛关注,至今留存下数量不菲的历史人物肖像及雕刻,其中不乏高品质的历史文化珍品,如三星堆出土的古蜀国金面罩铜质人头像,达芬奇画笔下神秘的蒙娜丽莎,唐宋时期古典雅韵的仕女图以及一系列直逼人灵魂深处的梵高自画像等。

感知是指人通过感知器官对外部客观事物的要素及特性在头脑中所做出的反映。反映可以是分析识别,也可以是再创造。就人脸感知而言,感知器官为人眼(相机、摄像机)、外部客观事物为人脸(人脸图像)、要素和特性为人脸特征、反映为模式识别和人脸三维结构再合成。本书中所指的人脸感知包括两个方面内容,即对人脸二维图像的分析理解以及人脸三维合成(也称建模)。

众所周知,人类视觉系统(human visual system, HVS)具备感知人脸复杂模式的能力,仅仅通过一张照片就能够准确、快速地在大脑中塑造出人脸三维形状,进而实现身份辨别。此外, HVS具有超强的鲁棒性,能在很宽广的光照变化范围内识别人脸,当人脸部分遮掩或发生一定形变,甚至经过一段相当长的时间后,还能对其正确感知并进行正确识别。深入探索人类视觉认知原理对设计一个更好的机器自动识别系统具有重要意义。

目前,学者们已经对人脸识别的问题做了大量研究^[1,2],也在理论上取得了一系列的进展,但要实现计算机的自动人脸分析及建模,现有技术远达不到实际应用的要求。究其原因在于认知心理学、认知神经科学以及生物学对人类感知问题的认识还处于初级阶段,相关理论问题的研究还不成熟。而现有人脸识别系统通常根据统计理论以及图像处理的方式来实现表观的识别,对认知科学理论成果的应用还不够深入。

如果能够借鉴人类的认知机理和相关数学理论的最新研究成果,建立新的计算模型和方法,实现对面脸图像的准确分析及建模是当前计算机视觉、人工智能等领域的一个关键问题。使计算机在一定程度上具有对面脸的感知能力,更是众多研究人员孜孜以求的目标,近年来已取得一定的进展,这些技术都在本书中进行了归纳总结。

1.2 人类视觉认知

1.2.1 视觉的地位

感觉是人的大脑与周围世界联系的窗口,它的任务是识别周围的物体并判断这些物体之间的关系。人的思维活动是建立在对客观世界与环境认识的基础上,而感觉则是客观世界与人对环境认识之间的桥梁,它使我们的思维与周围世界建立某种对应关系。人有多种感觉,其中对人的智力产生影响的主要是视觉和听觉。味觉和嗅觉是丰富多样的,但平常人们很少深入地去思考它们。在视觉和听觉中,形状、色彩、运动、声音等很容易被结合成各种明确、高度复杂、多样的基于空间和时间的组织结构,所以这两种感觉就成了理智活动得以发挥作用的非常合适的媒介。就听觉而言,人要从所听到的声音中获取具有意义的信息,还需要联系其他感性材料。视觉则不同,它是一种高度清晰的媒介,所提供的关于外界世界中各种物体和事件的信息十分丰富,它是思维的一种最基本的工具。总体而言,视觉是人类最重要的感觉,它是人的主要信息来源。研究表明,人类所认识的外界信息中 80% 来源于视觉^[3]。

1.2.2 视觉的内涵

眼睛是我们的视觉器官,整个视觉系统就像一台精确的摄像机,它具有较完善的光学系统及各种使眼球转动并调节光学装置的肌肉组织。视觉过程的生理机制包括折光机制、感光机制、传导机制和中枢机制。其中折光机制和感光机制使人感知事物。当眼睛注视外物时,由物体发出的光线通过上述折光装置使物像聚焦在视网膜的中央凹,形成清晰的物像。

人类视觉的内涵不仅指视力,其基本功能包括光觉、形觉、色觉、眼球运动和双眼单视等。视觉的基础是光觉,感受外界光的刺激是视觉的最基本特征。在光觉基础上眼睛还具有形态觉,即能够分辨和认识外界物体形状的能力,又叫中央视力或视锐度。对于中央视力而言,眼球向正前方固定注视一个目标时所见的空间范围称为视野。在视野中,越趋向周边的范围对物体的精细分辨力就越差。

除了以上我们所分析的单只眼睛所具有的能力,其实人类区别于许多生物的视觉能力之一是我们还具备非常完善的双眼视觉。双眼视觉分三级,最初级为同时知觉,即双眼能够同时看到两个不同的画面;第二级为像融合功能,即双眼将两个大部分相同、但在细节上有某些差别的图像看成一个图像;第三级为立体视觉,即三维空间知觉,又称深度视觉或空间视觉。具有完善的双眼视觉,才能感知物体的立体形状及该物体与人眼的距离或两个物体相对的远近关系。立体视觉是

最高级的双眼单视功能,它是人类从事各种高级精细工作的基础。

如上所述,人类对外界的感知绝大部分通过视觉系统获得,但人的视觉并不是与生俱来的,刚出生时的新生儿的眼睛只有光觉反应。实际上完善的视觉系统所完成的功能十分复杂。人的视觉功能主要是在出生后接受了大量的视觉刺激后逐渐发育成熟完善的,所以在视觉的形成期间,大量的视觉刺激对视觉发育的作用是十分重要的。

有学者认为视觉本身就包含了思维的一切基本因素。设想你要在一个会场中寻找一位朋友,呈现在你眼前的是由参加会议的人、桌、椅、主席台等组成的复杂景物。眼睛得到这些信息以后先要对景物的各部分进行分类,然后从中选出与朋友的外表有关的特征作出判断,那么在人的眼睛视网膜上映照的景物成像是否就能直接提供判断时所需要的有关特征呢?不是的。这里需要大脑的思考。例如,虽然人在不同距离处观察同一物体时在眼睛中成像的实际大小不同,但人们在观察某人的身高时却不会因为他在近处就判断他比实际高,也不会因他处在更远的位置就判断他比实际矮。这是大脑根据被观察物体的距离与周围物体的比较,并依靠有关的知识对输入的图像信息进行处理、解释的结果。假设你在一个灯光昏暗的剧院中寻找朋友,这个问题就变得更为困难。你刚走进剧院时会感到一片漆黑什么都看不清楚,过了几分钟你的眼睛变得习惯于在黑暗中观察。这时你的视觉系统在所处环境中对微光变得更加敏感,不能对物体和背景进行有效区分。这时眼睛对许多本来可用的信息捕捉不到,也就导致许多细节难以分辨。即便这样你也能认出朋友。总之,视觉是一个复杂的感知和思维过程,视觉器官眼睛仅接受外界的刺激信息,而大脑对这些信息通过复杂的机理进行解释,使这些刺激具有明确的物理意义。

从以上分析我们还可以看到敏感(sensation)、感觉(perception)、认知(cognition)这三个概念之间的联系和差别。敏感是把外界的各种刺激转换成人体神经系统能够接受的生物电信号。它所完成的是信号的转换,并不涉及对信号的理解。例如,人眼是视觉的敏感器官,它使光信号通过视网膜转换成电信号,与摄像机的光电传感器相似。视网膜的感光细胞对光信号在平面上进行采样,产生点阵形式的电信号,所不同的是摄像机的空间采样是均匀的,而视网膜的采样是不均匀的,在中央凹附近采样分辨率高,在周围采样的分辨率低。而感觉的任务是把敏感器官的各种输入转换和处理成对外部世界的理解。例如,对视觉来说就是能说出周围世界中有什么东西和这些东西之间的空间关系。这些都是关于周围世界的概念。从输入的点阵形式的信号到形式化的对客观世界的各种概念,其中要经过复杂的信息处理和推理。而认知是以人们对周围客观世界的概念为基础的。如果没有感觉这个人与外部世界的桥梁或窗口,人的思维活动就会失去基本的依据。

1.2.3 视觉的特点

人类视觉感知大致有以下特点:具有高分辨率,能立体观察,具备优越的识别能力和灵活的推理能力,可灵活地根据各种视觉线索进行推理。

(1) 深度感知。初等生物体表面上的光敏区开始时只能提供关于光源的方向信息,例如,阴影可能表示一个捕食动物正在逼近的方向。感光区域上感知的阴影面积的增加可能意味着捕食者正在逼近,这可能是深度感知的早期形式。这样的进化过程只是一种分析,有依据说明在人类视觉系统中探测阴影逼近的机能直接与感知物体在深度上的移动有关。对人类试验者进化的心理物理学实验和对猫进行的生理学实验都支持这种机理的存在。这种类型的机理与 Marr 于 20 世纪六七十年代提出的从视网膜上的二维表象开始,经过一个或多个中间表象计算再作三维解释的机理不同。

(2) 视觉感知是个自动进行的过程。它抵制根据与其相矛盾的知识作出的修改。实验证明,如果让一个观察者先观察一个旋转着的收缩螺丝,那么在他习惯以后再去观察另一个物体,如一张人脸,就会感到人脸在膨胀。观察者可能已经知道人脸并没有膨胀,但这并不妨碍得到这种膨胀的感觉。还有许多例子可证明人会出现这种与常识或已知情况相矛盾的感觉。虽然与感觉相矛盾的知识不能改变人的感觉,但显然它可影响人对视觉刺激作出不同的反应。一个人如果根据情况已知不会有大的物体正在逼近,那么当他看到出现一个影子时不会逃跑。但当影子突然出现时,他还会不自觉地感到害怕。人类虽然已具有较高级的理解能力,但视觉系统似乎保留着对某种刺激自动作出反应的能力。从进化的观点,感知与知识相分离可能是有道理的,但对用于准确分析三维景物的视觉系统来说就是不可取的。

(3) 感知对启发式知识的应用。自然环境中的许多物体是刚体,所以在进化过程中发展起来的人类视觉系统在根据视网膜上的物体投影分析物体时假设物体是刚体以简化分析是有道理的。例如,如果视网膜上成像的大小有变化,而形状保持不变,就可认为物体的远近起了变化。但在某些特殊情况下,在分析成像的大小变化时视觉系统并不采用通常的刚性物体的假设,而是采用不同于刚性假设的其他特殊假设。

除了以上特点以外,人类视觉系统有分辨率高、识别能力强、能进行灵活推理等优点,但也存在三维形状或姿态估计错误的现象。综上所述,从建立通用的计算机视觉系统的角度来看,关键点不是机械地模仿人类视觉系统,而是通过对人类视觉系统的研究发现是什么因素使人类视觉系统的性能如此之好,并且把它结合到计算机视觉系统中去。

1.2.4 人类视觉如何识别人脸

心理学和神经科学的研究成果对于设计机器人脸感知算法有着重要的借鉴和指导意义。为了更好地设计人脸的计算机自动感知系统,我们首先应当了解人类是怎样认识人脸的。针对这一问题,国内外许多学者做了大量的工作,本节将介绍几个人脸认知心理学方面有趣的实验和较成熟的理论结果。

1. 人脸认知心理实验

1) 人脸的感知是全局还是局部特征分析的结果

在对人脸的感知和识别中,全局和局部特征的分析和处理是非常重要的。全局特征一般用来进行粗略的匹配,局部特征则可以提供更为精细的确认。但是,一个必须提及的现象是:如果存在某种奇特的局部特征,则全局的特征描述可能就用不到了,此局部特征会首先被用来确定身份。例如,人们在回忆人脸时,一般会快速地集中在那些比较奇特的特征上,如大耳朵、鹰钩鼻等。支持“人脸识别比其他物体识别需要更多的全局特征”的一个强有力的证据是“人脸倒转实验”。研究表明,倒转的人脸比正常的人脸更难识别。一个非常好的例子就是英国约克大学的彼得·托马斯教授的“撒切尔夫人倒像实验”——他将两幅面积相同的图片,放入四方格内上部,如图 1-1(a)、(b)所示,位于上方的两幅图片是撒切尔夫人的倒立头像。粗看起来,两幅头像一模一样。正立之后[图 1-1(c)、(d)]观众才发现,左边的撒切尔夫人是人们所熟知的样子,而右边的撒切尔夫人头像完全不是正常面孔的样子,她的眼睛是“倒”的,鼻子是“歪”的,嘴巴的形状也十分奇怪。这就是面孔视觉中的倒脸效应。也就是说,在一定的场景下,人们在识别一个人脸时,全局信息往往会应用较多。

2) “瘦”脸效应

2010 年度“最佳视觉错觉”颁奖大会在意大利举办,一幅名为《胖瘦脸错觉》



(a)



(b)



(c)



(d)

图 1-1 撒切尔夫人倒像实验

的作品引起了很多观众的兴趣。托马逊发现,倒立面孔可以让脸“减肥”。图 1-2 中的两张图片是一名胖乎乎的英国中年男子。他将左边的照片倒立起来,右边的照片则保持正立状态。在左边那幅倒立照片里,这个中年男子看起来似乎消瘦了很多。事实上,“倒脸效应”的产生,与人类对于面孔的认知能力息息相关。在大脑腹侧枕区的纺锤体中,存在一个面孔识别区域,这个区域一般被称为“纺锤体面孔识别区域”。它擅长对人们遇到的正立面孔进行加工,而对倒立的面孔则存在识别上的缺陷。



图 1-2 胖瘦脸错觉

3) 识别因素实验

在中国科学技术馆网站(<http://b2museum.cdstm.cn/identification/>)上,由中国科学院自动化研究所生物识别与安全技术研究设计了一系列人脸识别实验,让参与者在实验的过程中体会自己到底是如何辨别人脸的。实验设计如下:

实验共进行 4 组测试。在每组测试中要求参与者对 8 个人进行识别,每个人