

上海交通大学学术出版基金资助
上海市重点图书



FACE DETECTION

黄福珍 苏剑波 \ 著

人脸检测

上海交通大学出版社

上海交通大学学术出版基金资助
上海市重点图书

人 脸 检 测

Face Detection

黄福珍 苏剑波 著

上海 交 通 大 学 出 版 社

内 容 提 要

本书首先对人脸检测的基本问题、研究思路和方法、经典的算法和技术全方位地做了深入系统的介绍,然后着重介绍了作者在利用活动轮廓模型方法进行人脸检测方面的一些研究成果,最后指出了该领域的未来发展方向。

本书可作为高等学校有关专业的研究生、高年级本科生、研究院所和有关单位广大科技工作者和工程技术人员的参考书。

图书在版编目(CIP)数据

人脸检测/黄福珍,苏剑波著. —上海:上海交通大学出版社,2006

ISBN 7-313-04366-X

I. 人… II. ①黄… ②苏… III. 人体-面-自动检测-研究 IV. Q983

中国版本图书馆 CIP 数据核字(2006)第 023850 号

人 脸 检 测

黄福珍 苏剑波 著

上海交通大学出版社出版发行

(上海市番禺路 877 号 邮政编码 200030)

电话: 64071208 出版人: 张天蔚

立信会计出版社常熟市印刷联营厂印刷 全国新华书店经销

开本: 850mm×1168mm 1/32 印张: 6.125 字数: 157 千字

2006 年 4 月第 1 版 2006 年 4 月第 1 次印刷

印数: 1~3 050

ISBN 7-313-04366-X/Q·018 定价: 12.00 元

版权所有 侵权必究

前 言

20 世纪 90 年代以来,随着需要的剧增,人脸识别技术成为一个热门的研究课题。由于人脸识别的非侵犯性,具有直接、友好、方便的特点,它是人们最容易接受的一种身份鉴别方式。作为人脸自动识别系统的重要环节——人脸检测,也得到了越来越多研究者的重视。随着研究的深入和应用的扩大,除了应用于人脸识别外,人脸检测在图像检索、视频会议、出入口控制以及智能人机交互等领域也有着重要的应用前景,因此,人脸检测已逐渐发展成为一个相对独立的研究领域和学术方向。

人脸检测的研究具有重要的学术价值,人脸是一类具有相当复杂的细节变化的自然结构目标,对此类目标的挑战性在于:(1)人脸由于外貌、表情、肤色等不同,具有模式的可变性;(2)一般意义下的人脸上,可能存在眼镜、胡须等附属物;(3)作为三维物体的人脸影像不可避免地受由光照产生的阴影的影响。因此,如果能够找到解决这些问题的方法,成功地构造出人脸检测系统,将为解决其他类似的复杂模式的检测问题提供重要的启示。

迄今为止,涌现出了各种各样的人脸检测算法,而且还出现了一些实用产品,这方面的研究在方法和手段上已经逐渐丰富和成熟。近年来,已有了一些较为详尽的人脸检测综述,但遗憾的是,还没有一本系统地介绍有关人脸检测理论的专著。因此,有必要将有关的研究资料汇编成书,供相关研究人员参考。更重要的是,希望通过本书总结作者多年来在人脸检测方面的一些研究成果,与国内同行们沟通,以使人脸检测理论和技术得到更广泛的重视和研究,取得更丰硕的成果,获得更广泛的应用。

本书的前部分内容取材于近几年来国内外重要的研究成果,

包括期刊论文和会议文献等,作者对这些学者和专家表示由衷的感谢。其中第1章简要回顾了人脸检测技术的问题描述、研究背景和发展现状以及应用领域,并介绍了有关商用产品,第2章和第3章按基于特征类和基于图像类将现有的人脸检测算法分为两大类,具体介绍了各种重要算法的研究思路,并对其优缺点给出评点。本书的后部分内容主要总结作者多年来在人脸检测方面的科研成果,其中第4章和第5章分别介绍基于活动轮廓模型的单人脸检测方法和多人脸检测方法,第6章介绍图像序列中的人脸检测与跟踪方法。本书最后给出了人脸检测领域的未来发展方向。为保持必要的完整性和系统性,书后附录也简要介绍了研究中用到的一些相关理论和技术。

由于作者学识、时间有限,加上人脸检测方面的新理论、新技术层出不穷,本书难免存在疏漏、缺陷和错误之处,恳请广大读者给予批评指正。

本书出版得到上海交通大学出版基金资助,特此表示感谢。

作 者

2006年3月

目 录

1 人脸检测概述	1
1.1 人脸检测问题描述	1
1.2 人脸检测的研究历史与现状	4
1.3 人脸检测的应用	5
1.4 人脸检测的有关产品介绍	7
1.5 人脸检测算法的评价	8
2 基于特征的人脸检测方法	15
2.1 低层特征分析方法	16
2.2 组群特征分析方法	27
2.3 变形模型方法	30
2.4 小结	39
3 基于图像的人脸检测方法	40
3.1 线性子空间方法	41
3.2 神经网络方法	47
3.3 其他统计方法	52
3.4 小结	67
4 基于活动轮廓模型的单人脸检测方法	69
4.1 活动轮廓模型简介	70
4.2 基于梯度向量流的单人脸检测方法	76
4.3 基于 Chan-Vese 模型的单人脸检测方法	87

4.4	基于变形垂足曲线的单人脸检测方法	99
4.5	三种模型方法比较	115
4.6	小结	118
5	基于活动轮廓模型的多人脸检测方法	119
5.1	基于多相 Chan-Vese 模型的多人脸检测方法	120
5.2	人脸数目已知时的多人脸检测方法	132
5.3	小结	135
6	图像序列中的人脸检测与跟踪	137
6.1	简单背景下的人脸检测与跟踪	138
6.2	复杂背景下的人脸检测与跟踪	146
6.3	两种方法比较	153
6.4	小结	156
7	总结与展望	157
	附录	159
	A 色度空间	159
	B 水平集方法快速数值算法	164
	参考文献	170

1 人脸检测概述

人脸是一个常见而复杂的视觉模式,人脸所反映的视觉信息在人与人的交流和交往中有着重要的作用和意义,对人脸进行处理和分析在视频监控、出入口控制、视频会议以及人机交互等领域都有着广泛的应用前景,因此是模式识别和计算机视觉领域持续的研究热点。对人脸进行处理和分析包括人脸识别、人脸跟踪、姿态估计和表情识别等,早期的人脸处理和分析一般都假设已知图像中人脸的位置和大小,但对于一个人脸自动处理和分析系统而言,人脸检测是关键性的第一步,人脸检测算法的精度直接影响着整个系统的性能。

近几年来,由于人脸检测在人脸处理和分析中的地位,人脸检测已经引起了越来越多研究者的注意,并逐渐发展成为一个相对独立的研究方向,其研究方法和手段上也逐渐成熟,出现了各种各样实用的人脸检测算法。虽然人脸检测领域的研究已经取得了一些可喜的成果,但在实用要求中仍面临着许多严峻的问题,人脸的非刚体性,表情、姿态、发型和化妆的多样性以及环境的复杂性都给正确检测人脸带来了困难。要让计算机像人一样方便准确地检测出无约束环境下任意图像中的人脸,尚需不同学科研究领域的科学家不懈的共同努力。

1.1 人脸检测问题描述

人脸检测(face detection)问题一般可描述为:给定静止或动态图像,判断其中是否有人脸;若有,将所有人脸从背景中分割出来,并确定每个人脸在图像中的位置和大小^[50,169]。人脸检测问

题所包含的内容十分广泛,从不同的角度可以有多种分类方法,如表1-1所示^[221]。

表 1-1 人脸检测问题的分类

分类依据		类 别	
图像类型	图像来源	静止图像(包括数字化的照片、数码相机拍摄的图片等,目前考虑的主要问题是算法的适应性和鲁棒性,算法速度在其次)	动态图像(即视频序列,包括工作台前的人脸序列,保安监控录像、影视资料等,往往与人脸跟踪问题交织在一起,对算法的速度有很高的要求)
	颜色信息	彩色	灰度
图像前景	镜头类型	头肩部图像	半身/全身图像
	人脸姿态	正面(包括端正及平面内旋转)	侧面(包括俯仰、侧影及旋转)
	人脸数目	单人(又称为人脸定位,是人脸检测问题在已知人脸数目情况下的特例)	未知(需要判定图像中是否存在人脸,人脸的数目以及各个人脸的尺度和位置,即是完全的检测问题)
图像背景复杂程度		简单背景(指无背景或背景的特征被严格约束,在该条件下只利用人脸的轮廓、颜色、运动等少量特征,就能够进行准确检测)	复杂背景(指背景的类型和特征不受约束,某些区域可能在色彩、纹理等特征上与人脸相似,必须利用较多的人脸特征才能做到准确检测)
应用领域		人脸信息处理系统(验证、识别、表情分析等)、视频会议或远程教育系统、视觉监视与跟踪、基于内容的图像与视频检索等等	

归纳起来,根据图像来源,可以将人脸检测方法分为基于静止图像的方法和基于动态图像的方法两类,其中前者是基础与核心,后者一般与人脸跟踪(face tracking)问题交织在一起。根据利用特征的色彩属性,可以将人脸检测方法分为基于肤色特征的方法和基于灰度特征的方法两类,其中前者适用于构造快速的人脸检测算法,后者利用了人脸区别于其他物体的最为本质的特征,是人脸检测领域研究的重点。根据利用人脸知识的不同,可以将人脸检测方法分为基于特征(feature-based)的方法和基于图像(image-based)的方法两大类,其中前者主要利用人脸的明显特征,如几何特征、肤色、纹理等,将人脸图像视为一个高维向量,从

而将人脸检测问题转化为高维空间中分布信号的检测问题；而后者则把人脸检测问题视为一个广义的模式识别问题，通过训练过程将样本分为人脸和非人脸两类。由于人脸检测问题的复杂性，无论哪一类方法都无法适应所有的情况，一般都针对人脸检测领域内某个或某些特定的问题。

虽然人脸检测方法不断涌现出来，但是，现有的算法一般都只能适用于一定的环境，无约束环境下的人脸检测问题仍没有得到很好的解决。如图1-1所示，姿态、表情、发型和脸部饰物的多样性以及环境的复杂性都给正确检测人脸带来了困难。

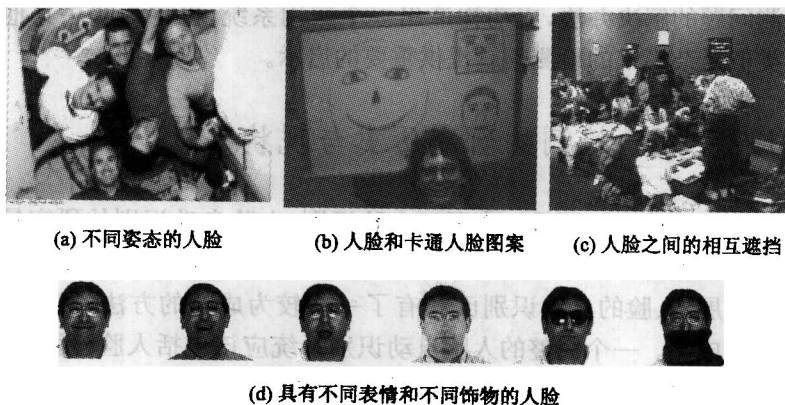


图 1-1 各种因素对人脸图像的影响

归纳起来，人脸检测不同于一般物体检测的主要困难表现在以下几个方面：

(1) 人脸构件及纹理的变化：人脸是一个非刚性物体，因而在不同的表情下，脸部器官的运动会导致人脸外观有很大的差异，如眼睛的闭合和展开，嘴部的形状等，还会有皱纹、斑点甚至化妆等带来的纹理特征，另外不同人种的人脸也有一些差异。

(2) 人脸表面的某些变化：人脸的一些具有普遍性的面部特征是部分具有的，如各种样式的胡须、头发、眼镜等结构特征，或

者不同年龄、不同性别的人其毛发也有一定差异。

(3) 无约束背景下不可预知的成像条件:图像的成像条件千变万化,如光照和拍摄视角的变化,以及具体成像的数码转换方法等,都极大地增加了人脸检测的复杂程度。

(4) 遮挡:人脸可能被图像中其他人脸或墨镜、围巾等挡住,使脸部某些重要的特征如眼睛、鼻子、嘴巴等不可见,从而造成人脸检测的困难。

所有这些未知因素都造成实际人脸检测的困难,也使人脸检测问题成为一个极富有挑战性的研究课题。如果能找到人脸检测问题的解决办法,成功构造出一个实用系统,也将为解决其他类似复杂模式的检测问题提供重要的启示。

1.2 人脸检测的研究历史与现状

人脸检测问题最初来源于人脸识别,人脸自动识别的研究最早可以追溯到 20 世纪 60~70 年代^[15,124,198,206]。经过几十年的曲折发展,人脸的自动识别已经有了一些较为成功的方法,目前正日趋成熟。一个完整的人脸自动识别系统应该包括人脸检测、特征提取以及匹配识别三大基本过程,如图 1-2 所示。

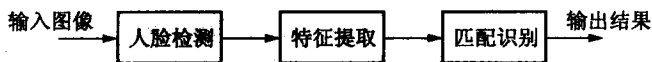


图 1-2 人脸自动识别系统构成

早期的人脸识别研究主要针对具有强约束条件的人脸图像(如无背景或背景简单的图像),往往假设人脸位置已知或很容易获得,因此人脸检测问题并未受到重视。进入 20 世纪 90 年代以后,随着网络安全和电子商务等应用需要的剧增,人脸识别成为最有潜力的生物身份鉴别方式,出现了实际的人脸识别系统和商用产品,这种应用背景要求自动人脸识别系统能够对一般环境图像具有一定的适应能力,由此所面临的一系列问题使得人们对人

脸检测的各个重要方面引起了极大的研究兴趣,并逐渐独立出来成为一个专门的研究方向。

目前,国外对人脸检测问题的研究很多,比较著名的有美国麻省理工大学(MIT)的媒体实验室和人工智能实验室^[101,136]、卡内基梅隆大学(CMU)的机器人研究所^[120,121,127]以及 Illinois 大学的 Beckman 研究所^[122,161,164-169]等;国内的清华大学^[196,197,199,200,204,206,208,210,211,217]、北京工业大学^[207,209]、南京理工大学^[203,213,214]、四川大学^[198,220]、上海交通大学^[201,224]、中国科学院计算技术研究所^[193,194,202,215,216]和中国科学院自动化研究所^[152]等都有人员从事人脸检测方面的相关研究。MPEG7 标准组织已经建立了人脸识别草案小组,人脸检测算法也是一项征集的内容。

随着人脸检测研究的深入,国际上发表的有关论文数量也大幅度增加,EI 可检索到的相关文献多达数千篇,每年的国际会议上关于这方面的专题也屡屡可见,如 IEEE 的 FG(IEEE International Conference on Automatic Face and Gesture Recognition)、CVPR(IEEE Computer Society Conference on Computer Vision and Pattern Recognition)、ICIP(IEEE International Conference on Image Processing)、ICPR(IEEE International Conference on Pattern Recognition)等重要国际会议的每次会议上都有大量关于人脸检测的论文,约占有关人脸研究论文的1/3。

1.3 人脸检测的应用

任何一种技术的发展都是由于受到了实际应用需要的激励,人脸检测技术也不例外,它的应用十分广泛,可用于人脸识别、视频会议、图像与视频检索以及智能人机交互等^[50,169,221,226,227]。下面分别介绍这几方面的应用。

1) 人脸识别

如前所述,人脸检测最初来源于人脸识别,它是一个人脸自

动识别系统中至关重要的环节,因此人脸检测技术最重要的应用是人脸识别。与指纹、视网膜、虹膜、基因、掌形等其他人体生物特征识别系统相比,人脸识别具有非侵犯性以及直接、友好和方便的特点,它是人们最容易接受的一种身份鉴别方式。由于人脸识别技术在生物特征识别技术中具有众多的优势,决定了其广泛的应用前景,如刑侦破案、证件验证、出入口控制以及视频监控等。同样,人脸检测也是人脸表情识别等其他人脸信息处理系统中关键性的第一步,人脸检测算法的精度直接影响着这些系统的性能。

2) 视频会议

可视电话和视频会议是基于内容的编码传输方法的典型应用,而他们所感兴趣的传输内容是人脸。在视频会议系统和视频监控系统中,需要将摄像机镜头始终对准与会人或被监控对象,摄像机镜头的方向取决于人脸的角度,因此有效检测到人脸成为引导摄像机的先决条件,一个鲁棒的、实时的人脸检测系统所能提供的精确的人脸区域是这些应用切实可行的前提。

3) 图像与视频检索

基于内容的图像检索(Content-based Image Retrieval,简称CBIR)尤其是视频检索的研究方兴未艾,人脸作为一种重要的、稳定的、具有一定语义的检索特征,在新闻片、故事片等各种题材类型的视频中广泛存在,可以成为图像检索的条件,因此许多CBIR系统都将人脸检测作为系统中的一个重要组成部分,利用检测到的人脸,可以有效标注、索引以及分类视频。例如,芝加哥大学开发的用于万维网上搜索图像的影像搜索引擎Webseer中就使用了基于神经网络的人脸检测作为搜索引擎的一部分^[37]。

4) 智能人机交互

智能人机交互(Human-Computer Interaction,简称HCI)是

在人机之间实现类似于人类之间的非精确的自然交互方式,它是集语音、手语、人脸、表情、唇读、头势、体势等多通道为一体的计算机智能接口系统。显然,在这种以人为中心的交互方式中,人脸检测是系统中十分重要的一环。

1.4 人脸检测的有关产品介绍

自 1995 年以来,国外一些公司看准了人脸识别系统在商贸和保安方面的广阔应用前景,动用了大量人力和物力,独立研制或与高校合作,开发了多个实用的人脸识别系统,其中大多数产品将人脸检测作为系统的一个重要组成部分,主要有如下产品。

1) 德国 Cognitec Systems 公司的 Face VACS 人脸识别系统^[185]

Face VACS(Visual Access Control System)是一个配合卡片使用的入口控制系统,它可以配合卡片和身份证使用,或者干脆代替身份证使用。在使用时,用户先把卡插入读卡机,输入正确的个人身份证号码,摄像机拍下这个用户的几张照片,并与库里存储的顾客特征信息进行比较。对于实拍图像,该系统使用一种特别训练的神经网络算法来首先搜索在图像中的人脸。即使人脸处在图像中的某个很小的区域,系统也能把它快速可靠地检测出来。

2) 德国 HumanScan 公司的 BioID 身份识别系统^[186]

BioID 是一个将面貌、声音以及嘴唇动作 3 种生物特征相结合的身份识别系统,该系统利用摄像头获取使用者的面貌长相(静态特征)、一秒钟时间长度内的声音与嘴唇的运动(动态特征),并与保存在数据库中的此人的特征进行比较,完全吻合则可以通过。BioID 采用基于模型的两阶段算法来实现人脸检测,首先采用一个二值人脸模型来进行人脸粗定位,然后再用一个人眼区域模型进行人脸精确定位。

3) 美国 Identix 公司的 FaceIt 人脸识别系统^[187]

FaceIt 是一个较有影响的人脸识别系统,其主要功能是自动检测在实况视频中出现的人脸,定位、跟踪他们并鉴别个人。该系统对通过摄像机采集来的大量图像数据进行分析,首先寻找图像中的人头形状,当一个类似人头的图像出现时,即利用各种数学拼配法确定该位置是否真有人头。该系统可以同时检测一群人的面貌并确定他们的准确位置。找出人脸后,该系统主要利用局部特征分析(local feature analysis)方法来进行人脸识别。

4) 加拿大 Imagis 公司的 ID-2000 面部识别软件^[188]

ID-2000 是一套用途广泛、功能灵活多变的面部识别软件,它采用多达 692 个面部特征来捕获和识别人脸。该系统首先利用光谱分析和 3D 建模得到人脸的 692 个特征,同时考虑人脸的俯仰和偏转以及各种光照条件,建立人脸的变形表面模型(deformable surface model),然后根据该模型来检测图像中不同姿态的人脸。

5) 美国 Viisage 公司的 FaceTools 人脸识别系统^[189]

Viisage 公司主要生产由 MIT 媒体实验室研制的基于特征脸方法的人脸识别系统,FaceTools 是一个采用特征脸(eigenface)方法来实现人脸检测和识别的软件开发工具箱。

国内关于人脸检测方面的研究是从近几年开始的,发展很快,也有一些工作已经或正在转化为商品软件,如上海银晨智能识别科技有限公司开发的人脸检测音视频压缩卡^[190],广州超越软件公司的 BioFace 系统^[191]等等。

1.5 人脸检测算法的评价

人脸检测系统的输出准确与否需要一个客观的标准进行评

价。长期以来,研究人员虽然提出了大量的人脸检测算法,但在说明他们的检测方法的性能时,往往使用不同的测试集。因此,要对人脸检测算法的性能作出精确的评价是一件很困难的事情。虽然最近出现了一些标准测试人脸图像集,但对于怎么才是一个成功的人脸检测方法,仍缺乏一个统一的标准;另一方面,对于不同的训练集和样本规模,尤其是一些基于图像的人脸检测方法,对测试结果的影响也是巨大的。

1.5.1 人脸数据库

任何一个人脸处理和分析系统的开发都需要建立一个包含人脸图像或图像系列的数据库,到目前为止,出现了很多适应不同需要的人脸图像库,如表1-2所示。

表 1-2 人脸图像库

名 称	位 置	描 述
英国 AT&T 人脸数据库	http://www.uk.research.att.com/facedatabase.html	由 40 个人的 400 幅灰度图像组成,人脸脸部表情和细节均有变化,人脸姿态也有所变化
美 国 FERET (FacE Recognition Technology) 人脸数据库	http://www.nist.gov/humanid/feret	目前最大的人脸数据库,由美国军方研究实验室提供,其中每人 8 张图像,两张正脸,3 张从右到左不同侧面角度的照片,有些人还提供了更多不同视点和不同表情的照片
欧洲 M2VTS 多模型人脸数据库	http://poseidon.csd.auth.gr/M2VTS/index.html	用于测试多模型身份鉴别方法,目前该库由 37 人组成,每人有 5 个图像系列,拍摄时间间隔一周左右,其中至少有一个系列提供合成语音
UMIST 人脸数据库	http://images.ee.umist.ac.uk/danny/database.html	由 20 个人共 564 幅人脸图像组成,每个人具有不同角度和不同姿态的多幅图像
普渡大学 AR 人脸数据库	http://rvll.ecn.purdue.edu/~aleix/aleix_face_DB.html	由 126 个人的 4000 多幅彩色人脸图像组成,人脸表情、面部装饰及光照条件均有变化

(续表)

名 称	位 置	描 述
伯尔尼大学人脸数据库	ftp://iamftp.unibe.ch/pub/Images/FaceImages/	由 30 个人的 300 幅正面人脸图像(每人 10 幅)和 150 幅侧面人脸图像(每人 5 幅)组成
哈佛大学人脸数据库	ftp://ftp.harvard.edu/pub/faces	包含大量不同光照条件下的正面人脸图像
耶鲁大学人脸数据库	http://cvc.yale.edu/	由 16 个人每人 10 幅正面人脸图像组成,人脸表情及光照条件均有变化
中国 CAS-PEAL 人脸数据库	http://www.jdl.ac.cn/peal/index.html	公开的数据库包括 1040 个人的 30871 幅图像,表情、姿态、光照以及饰物均有变化

以上这些图像库主要用来评价人脸识别算法的性能,因此一般背景比较单一,并不适合作为人脸检测测试集,但其中提供的大量人脸图像经过一些加工后,可以作为人脸训练样本。除了这些可以利用的用于人脸识别的数据库外,目前国际上已经出现了专门的评价人脸检测算法性能的数据库,比较公认的测试集有 Sung 等的 MIT 测试集和 Rowley 等的 CMU 测试集(涵盖了前者)以及 BioID 数据库等,如表 1-3 所示。

表 1-3 人脸检测测试图像库

名 称	位 置	描 述
CMU-130	http://vasc.ri.cmu.edu/idb/html/face/frontal_images/index.html	即完整的 CMU 测试集,包括 A、B、C 三个数据库,共有 130 幅灰度图像,包含 507 个不同大小的人脸
CMU-125		去掉 CMU-130 测试集中的画线图,共有 125 幅灰度图像,包含 483 个人脸
MIT-23	http://vasc.ri.cmu.edu/idb/html/face/frontal_images/index.html	即完整的 MIT 测试集,也就是 CMU-130 中 B 集数据库,共有 23 幅灰度图像,包含 149 个人脸
MIT-20		去掉 MIT-23 测试集的画线图,共有 20 幅灰度图像,包含 136 个人脸