

华中科技大学

硕士学位论文

基于肤色模型的人脸检测算法研究

姓名：柯研

申请学位级别：硕士

专业：通信与信息系统

指导教师：刘文予

20080602

## 独创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除文中已经标明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到，本声明的法律结果由本人承担。

学位论文作者签名：

日期： 年 月 日

## 学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，即：学校有权保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权华中科技大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本论文属于 ☐ 保密，在\_\_\_\_\_年解密后适用本授权书。  
☐ 不保密。

（请在以上方框内打“ ”）

学位论文作者签名：

日期： 年 月 日

指导教师签名：

日期： 年 月 日

# 摘 要

人脸检测技术在当今社会生活中扮演着相当重要的角色，如视频监控、人机交互、人脸识别、人脸数据库分类管理等业务，均用到了人脸检测这个关键技术。人脸的检测和识别技术是模式识别、计算机视觉和图像理解系统的研究热点之一，它综合了图像理解、计算机视觉、数据库和人工智能等多方面多学科的知识，应用前景广阔，受到了越来越广泛的重视，并得到了迅速的发展，并取得了很多有价值的研究成果。但是，目前的人脸检测算法一般对光照情况和背景的复杂程度的变化适应性较差，特别是在面对恶劣光照环境时候，人脸的检测率非常低。

本文主要研究了基于肤色模型的人脸检测算法的基本理论和关键技术，总结分析了基于肤色模型的人脸检测算法的优势与弱点，重点讨论了由于恶劣光照条件带来的对人脸检测结果的不利影响的问题。本文采用针对不同光照环境的光线补偿算法，克服人脸图像在复杂光照环境下的检测困难，提出了一种从粗到细、从整体到局部的基于不同颜色空间的人脸光照补偿算法。该算法不用估计环境光的方向、强度和光源的性质（如：是否点光源、是否平行光、单光源还是多光源等）。针对不同的光照环境情况（主要为光照不足或过亮、有色光源照射、由于偏光造成的局部阴影），采用不同的光线补偿策略，并配合基于混合高斯模型对肤色色度信息进行描述，并采用根据各像素点距肤色中心马式距离来计算与肤色相似度的算法，对很多图像处理都得到了比较好的结果，使得皮肤与非皮肤区域有效的分开。

理论和实验证明，本方法能够有效地解决人脸检测和识别中不同复杂光照环境影响的问题，提高肤色区域与非肤色区域的分割准确性，使在不同光照条件下的人脸检测性能得到了很大的提高。而且本方法还具有运算效率高、鲁棒性好的特点。

**关键词：**人脸检测；肤色模型；光线补偿；色彩空间；

# Abstract

Human face detection plays an important roll in applications such as video surveillance, human computer interface, face recognition, and face image database management. Face detection and recognition technology is one of the hottest spots in pattern recognition, computer vision and image understanding system; it has integrated multi-disciplinary knowledge such as image understanding, computer vision, databases and artificial intelligence, and with the rapid development it get more and more attention and has made many valuable research results. But now the face detection system can not reach our expectation, especially with the bad environment.

A face detection algorithm for color images in the presence of varying lighting conditions as well as complex backgrounds is proposed by us. Our light compensation method which is based on different color space from-wide-to-slender and from-whole-to-part instead of estimating the surround light conditions is brought out. The method does not have to estimate the direction, strength and the nature of light source (such as: whether it is point light source, parallel light, single-source or multi-source, etc.) The method with different light compensation strategy adapts different light circumstances well. With the description of skin color information based on mixed Gaussian model and the Pixel Pitch Center Ma-distance to calculate the color similarity, the algorithm can gain a good result. Based on a novel lighting compensation technique and a nonlinear color transformation, our method detects skin regions over the entire image and then generates face candidates based on the spatial arrangement of these skin patches. It can classify the face area and non-face area effectually.

It validate that the problem which is caused by light condition can be solved by the method with both the theory and experimentation. it can enhance the accuracy of

division between regional and non-skin color regional so that the face detection performance in different light conditions has been greatly improved .This method also has high efficiency operation and robustness,etc.

**Keywords : Face Detection; Skin Model; Light Compensation ; Color Space ;**

# 1 绪 论

## 1.1 论文研究背景及意义

### 1.1.1 人脸检测算法的产生背景

感觉是人的大脑联系周围世界的窗口，它的任务是认识周围的物体，并判断这些物体之间的关系。我们的思维活动是以我们对客观世界与环境的认识为基础的，而感觉则是客观世界与我们对环境的认识之间的桥梁，使我们的思维与周围世界建立某种对应关系。

人是有多重感觉的，但对人的大脑意识产生影响的主要是视觉和听觉。在视觉和听觉中形状、色彩、运动、节奏等很容易被结合成各种明确和高度复杂、多样的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行使和发挥作用的非常合适的媒体。但人的听觉要想具有意义还需要联系其它的感性材料。而视觉则不同，它是一种高度清晰的媒介，它提供关于外界世界中各种物体和事件的丰富信息，因此它是思维的一种最基本的前提。因此视觉是人类最重要的感觉，它是人的主要感觉来源，人类认识外界信息中 80% 来自视觉<sup>[1]</sup>。

视觉对正常人来说是生而有之，毫不费力的能力。但实际上视觉是一个十分复杂的感知和思维过程。心理 - 生理科学虽然早在几个世纪前就对人眼视觉开始研究，也取得了很大的进展，但至今并不完全清楚这个非常“神秘”的过程。视觉器官 - 眼睛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处理和解释，使这些刺激具有明确的物理意义。那么人脑到底是怎样完成复杂的识别任务呢？通常我们认为人类的识别系统与心理学、神经学等有关，利用各种感知器官，如视觉、听觉、触觉、嗅觉，对外界的刺激产生反映，传递到大脑进行分析、综合，最后做出判断。有时人所处的周围环境也会对识别过程起到一个辅助的作用。那么具体针对其中最重要的视觉信息，人类大脑又是怎样进行处理的？它对于人脸的识别是通过整体的分析得到结果的，还是通过局部的分析呢？哪些是它区分不同人的最主要特征呢？这些主要特征在被处理的过程中，是否不同的优先级呢？我们使用计算机怎样最好的模仿人类的识别系统呢？正是这一系列

的问题成为当前计算机视觉领域研究的重要课题<sup>[2]</sup>。

随着新的信息技术和信息媒介的普及,在人机交互方面越来越多的高效友好的方法被开发出来,这些方法不依赖于传统的设备,比如说键盘、鼠标和显示器。计算机的价格持续下降,性能不断提高,而且,近来视频设备成本下跌,预示着计算机视频系统能够在台式机和嵌入式系统中开发。人脸处理研究的快速发展是基于假设的,即关于用户身份、状态、意图的信息能够从图像中抽取出来,然后计算机做出相应响应,比方说观察一个人的面部表情。尽管心理学家、神经学学家和工程师们已经在人脸和人脸表情识别方面研究了 20 多年,但近五年中这方面已经吸引了很多人的注意。很多开发出来的科研原型系统和商业产品应用了这些方法。任何一个人脸图像处理系统的第一步是人脸在图像中的位置。

人脸识别归属与人体生物认证技术,人体的生物特征包括生理特征和行为特征两大类。人体的生理特征主要包括人脸、指纹、掌纹、掌形、虹膜、视网膜、静脉、DNA、颅骨等,这些特征是与生俱来的,是先天形成的;而行为特征包括声纹、签名、步态、耳形、按键节奏、身体气味等,这些特征是由后天的生活环境和生活习惯决定的。这些生物特征本身固有的特点决定了其在生物认证中所起的作用是不同的。人脸识别是一个活跃的研究领域,是人类视觉研究中最杰出的方面之一。虽然人脸识别的准确性要低于虹膜、指纹的识别,但由于它的无侵害性和能够给用户最自然、最直观的感觉方式,使人脸识别成为最容易被接受的生物特征识别方式。所以,人脸识别更具有公共性应用潜力。“9·11”事件是生物特征认证技术在全球发展的一个转折点。“9·11”以后,人脸识别技术的重要性被全球各国政府更加清楚地认识到。传统的身份鉴别技术在面临反对恐怖主义任务时所表现出来的缺陷,使得各国政府在研究与应用上对人脸特征识别技术开始了大规模的投资。在美国,三个相关的法案(爱国者法案、边境签证法案、航空安全法案)都要求必须采用人脸识别技术作为法律实施保证。总体上来说,在国外生物认证技术的应用已经进入了以政府应用为主的阶段。普通公众对人脸识别技术的了解也因为“9·11”的曝光率而大幅度提高,种种因素,促成了“9·11”以后全球人脸特征认证市场的加速增长。在我国,生物认证技术除了能够繁荣一个新兴的产业,更重要的是其提供的身份鉴别能力对国家安全的重要意义:首先,中国作为最耀眼的社会主义国家,国家安全问题仍然需要警惕;其次,中国人口

多、流动性大，需要有更为先进的身份鉴别技术来保证社会秩序；再次，生物认证技术将有效解决国内的假冒伪劣证件、证书问题；最后也是最实际的、需要迫切面对的问题，那就是 2008 年的奥运会和 2010 年的世界博览会，如果能将生物认证技术用于大型集会的身份鉴别中，无疑能够更好地确保大会的安全。

### 1.1.2 人脸检测算法研究的意义

人脸检测是指在输入图像中确定所有人脸的位置、大小和位姿的过程。人脸检测作为人脸信息处理中的一项关键技术，近年来已成为模式识别与计算机视觉领域内一项受到普遍重视，研究十分活跃的课题。人脸检测技术在很多方面都有重要的应用，如：

#### 1、身份鉴定

人脸身份鉴定技术是生物识别技术中的一种，以其直接性，唯一性，方便性等特点在公安，海关，交通，金融，社会保险，医疗及其它民用安全控制系统等领域具有较为广阔的发展前景和很高的社会效益<sup>[2]</sup>。

#### 2、多媒体信息检索

随着因特网技术的迅速普及，网上多媒体信息（如图像，视频）日益丰富，目前所能应用的信息资源已经非常丰富和庞大，而快速有效地寻找感兴趣的相关信息已经变得日趋困难，采用关键字和描述文本检索的方式已经无法胜任现有需求，基于内容的多媒体信息检索技术应运而生，内容为人脸的图像和视频检索是其中非常重要的部分。<sup>[3]</sup>

#### 3、智能监控

在传统的视频监控系统（如闭路电视系统）中，工作人员需要不断监控屏幕去发现可疑的事件和目标，因此工作量很大而且效率低。在采用人脸检测技术的智能监控系统中，计算机可以在分割、检测出人物目标时自动通知工作人员的介入，从而减轻工作人员的负担并提高效率；计算机还可以对这些目标的运动模式进行一些自动分析，以确定是否为可疑目标。

#### 4、人机交互

在人机交互方面，计算机把关于用户身份、状态、意图的信息从图像中抽取出来，然后做出相应响应，比方说观察一个人的面部表情，然后采用不同的方式



处理问题。

在上述几个方面,人脸识别、辨认、定位以及追踪等都与人脸检测密切相关。人脸定位的目的是确定图像中人脸的位置。假设一幅图像中只存在一张脸,则面部特征检测的目的是检测特征的存在和位置,如眼睛、鼻子、鼻孔、眉毛、嘴、嘴唇、耳朵等。人脸识别或辨认是将输入图像与数据库中的图像比对,如果存在,报告匹配结果。人脸识别的目的是检验输入图像中的个体的身份,而人脸追踪方法是实时地、连续地估计在图像序列中的人脸的位置和可能的方向。面部表情识别涉及识别人类的情感状态,如高兴、悲伤、厌恶等。

任何一个人脸图像处理系统的第一步是人脸在图像中的位置。然而,从单张图片中检测出人脸是一项具有挑战性的工作,因为人脸在大小、位置、方向、姿势方面是可变的。人脸表情、牙齿相接触的方式、光照也会改变人脸的整体外观。

## 1.2 国内外研究现状

人脸检测问题最初来源于人脸识别,它是自动人脸识别系统中的一个关键环节。人脸识别技术研究始于 60 年代末,最早的人脸识别以人脸特征点间的距离、比率等参数为特征,建立了一个半自动人脸识别系统。而早期人脸识别研究主要有两大方向:一是基于人脸几何特征的方法,人脸器官归一化的点间距离和比率以及人脸的一些特征点,如眼角、嘴角、鼻尖等部位所构成的二维拓扑结构;二是模板匹配的方法,利用计算机模板和图像灰度的自相关性来实现识别功能。Berto 在 1993 年对这两类方法作了较全面的介绍和比较后认为,模板匹配的方法优于几何特征的方法。<sup>[4]</sup>

90 年代人脸识别成为科研热点。但是人脸识别所采用的人脸库通常不大,最常用的人脸库仅包括 100 幅左右的人脸图像,如 MIT 库、Yale 库、CMU 库等人脸库均为小型库;且人脸库之间的输入条件各异,不同的识别算法很难比较。为了促进人脸识别深入研究和实用化,美国国防部发起了人脸识别技术工程,它包括一个通用人脸库和一套通用测试标准。不过该库中包括有军人的图片不能在美国以外获得,因此其他国家的研究只能采用本地的人脸库,如英国的 Manchester 人脸库。目前的人脸识别主要有两个方向:其一是基于整体特征的研究方法,其二是基于局部特征分析的方法。<sup>[2]</sup>

近几年随着电子商务等应用的发展,人脸识别成为最有潜力的生物身份验证手段,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脸检测也开始作为一个独立的课题受到研究者的重视。人脸检测问题在近十年中得到了广泛的关注和长足的发展,国内外很多研究学者提出了很多方法,在不同的领域取得了一定的成果。但是对于寻找一种能够普遍适用于各种复杂情况的准确率很高的检测算法,还有很大的距离。

国外的主要研究单位有美国的麻省理工媒体实验室(MIT Media lab)、卡耐基梅隆大学的人机交互学院(Human computer interface institute)、微软研究院的视觉技术研究组(Vision Technology Group)、英国剑桥大学工程系(Department of Engineering)等,国内的研究单位有中科院自动化所、中科院计算所、哈尔滨工业大学、清华大学、浙江大学等。

MPEG-7 标准组织也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长,如 IEEE FG(IEEE 人脸与手势自动识别国际会议)、ICIP(图像处理国际会议)、CVPR(计算机视觉与模式识别会议)等重要国际会议上每年都有大量关于人脸检测的论文,占有关人脸研究论文的近 1/3 之多。

整体来看这一方向涉及数字图像处理、计算机视觉、人工智能等多个学科领域,同时这一领域适应了现代计算机网络发展和现代通信发展的需要,无论从学术性还是从实用性来看,均具有极高的研究价值。从已有的研究来看,要完全实现一个全自动的人脸检测跟踪识别系统的难度是相当大的,但是要实现特定环境下的自动人脸检测和识别还是有可能的,随着计算机和网络技术的普及,图像、视频等多媒体内容在计算机信息中比重的加大。这一研究方向必然会得到更快的发展。<sup>[5]</sup>

### 1.3 论文的主要研究工作

人脸检测面临的挑战可以归结为以下因素:

- 1) 姿势。人脸图像会因摄像机 - 人脸的相对位置(正面,成 45 度角,侧面,上下颠倒)以及像独眼、鼻子部分或全部闭合等一些人脸特性而不同。
- 2) 组成部件的有和无。像胡子、眼镜等面部特征,有些人有,有些人没有,

而且它们在形状、颜色、和大小方面也有很大的差别。

3) 面部表情。人的面貌直接受面部表情的影响。

4) 遮挡。人脸可能被其他物体部分遮挡。在有一群人的图像里,有些人的脸会被其他人的脸部分遮挡。

5) 图像的方位。人脸图像直接受视频光轴的不同旋转角度影响。

6) 成像条件。当图像形成的时候,光照(光谱、光源分布、光密度)、摄像头特性(感应器,镜头)等因素会影响人脸的外观。<sup>[6]</sup>

目前,根据系统利用特征色彩属性的不同可以将这些方法大致分为基于肤色特征的方法和基于灰度特征的方法,以及两者的结合三类。根据人脸的色彩信息和启发式信息定位人脸,速度快但对姿态不敏感,难以有效地区分类肤色的区域;根据灰度信息进行模板匹配,计算量大,速度慢,如基于积分图像与Adaboost训练的快速人脸检测算法,虽然检测效果理想,但分类器的训练却是一个相当费时的过程<sup>[4]</sup>。因此,两者结合的检测方法已经成为人脸检测技术发展中的一个趋势,国内外在这方面也已经取得了一定的进展,但现有的算法仍然多集中在正常光照条件下人脸的检测上,对于复杂光照环境的背景下的人脸的检测上还存在很大的困难。本文针对一般彩色图像中的人脸检测问题,提出了一种综合肤色分割的方法。该方法有效地排除了不同光照环境对检测结果的影响,可以在复杂的光照环境下检测出不同的人脸,具有更高的检测率和更强的鲁棒性。

## 1.4 论文的结构安排

全文共分为五章,具体安排如下:

第一章 绪论,简要介绍了人脸检测技术产生的背景,系统地分析了其在计算机技术和网络技术普及的现在,人脸检测技术的应用范围,如:身份鉴定、视频检索、智能监控、人机交互等领域,强调了人脸检测技术研究的重要意义;简要介绍了本算法在不同光照环境下的性能优势,同时简单回顾了国内外人脸检测技术的发展概况;最后给出了本文的主要研究内容和论文的组织结构。

第二章 人脸检测技术分析,对现有的人脸检测算法进行了分类:基于知识方法、基于外观方法和模板匹配的方法。详细地分析了一些具有代表性的方法的特点,并对这三类人脸检测算法进行了性能评价和小结。

第三章 彩色图像中的肤色分割技术，包括肤色特征，色彩空间理论，肤色模型分类和建立。

第四章 基于不同光照条件下肤色模型的人脸检测算法，从三类不同光照模型入手，分析不同光照环境对人脸检测系统的肤色匹配环节影响，实验从主观和客观上验证了算法的有效性。

第五章 总结与展望，总结了全文内容和工作，并对需要进一步研究的问题进行了展望。

## 2 人脸检测技术综述

人脸检测问题在近十年中得到了广泛的关注和长足的发展,国内外很多研究学者提出了很多方法,在不同的领域取得了一定的成果。但是对于寻找一种能够普遍适用于各种复杂情况的准确率很高的检测算法,还有很大的距离。现有的人脸检测技术方法可以分为三类:基于知识的方法、模板匹配法、基于外观的方法。

### 2.1 基于知识的方法

#### 2.2.1 基于知识的方法原理

这类方法中,人脸检测是基于研究者对人脸的先验知识。提出描述人脸特征及其相互关系的单个线索并非难事。例如,出现在图像中的人脸经常包括互相对称的两只眼睛、一个鼻子、一个嘴巴。特征之间的关系可以通过相对距离和位置来表示。这类方法的步骤是:首先从输入图像中抽取出人脸特征,然后通过经编码的准则进行比较,识别候选人脸。确认过程经常被用来降低误检率。

这种方法的一个难点是怎样把人脸知识转化为定义良好的准则。如果准则过于详细,那么有些人脸将被漏检,因为其没有通过所有的准则。如果准则过于粗略,则将可能犯很多低级错误。此外,这种方法很难被扩充到检测不同姿势的人脸,因为要枚举出所有可能的情况比较困难。另一方面,这种启发式方法对于特定场景中的正面人脸检测具有很好的效果。

#### 2.2.2 自上而下的人脸特征

##### 2.2.2.1 人脸轮廓特征

Sirohey 提出了一种定位方法,把人脸从复杂背景中提取出来进行人脸验证。<sup>[7]</sup>它使用边界图(canny 检测器<sup>[8]</sup>)和启发器对边界进行移除和分类,因此,只有人脸的轮廓被保存了下来。头部区域与背景的分界线就形成了一个椭圆。在不同背景下的 48 幅图像的数据库里,这个算法的精确率达到了 80%。除了使用边

缘之外, Chetverikov 和 Lerch 提出了一种使用条纹(即相似方向的边缘)的简单人脸检测算法<sup>[9]</sup>。他们的人脸模型包括两条黑斑和三条亮斑,用来分别代表眼睛、颧骨、鼻子。这个模型使用斑纹来表示人脸、眉毛、嘴唇的轮廓,两个三角形来示意条纹间的空间关系。在进行条纹检测之前先用低通拉普拉斯算子对图像进行处理。然后扫描图像,找到特定的三角形区域,作为候选。最后通过验证候选周围的条纹来检测人脸是否存在。

Grafetal 提出了一种在灰度图像中定位人脸特征及人脸的方法<sup>[10]</sup>。图像首先经过滤波,然后用形态学方法增强图像的有形高亮度区域(例如说眼睛)。经过处理后图像的直方图有明显的峰值。基于峰值和它的宽度,选择合适的阈值,就能够将原图转化为二值图像。二值图像中互相关联的成分被认为是人脸特征的候选。将那些候选合并起来用分类器进行评价,以确定哪些是人脸以及人脸在哪个地方。他们的这种方法已经经过测试,使用的是 40 个人的头部图像和 5 个视频序列,每个序列包含 100 到 200 帧。然而,该方法没有讲清楚怎样用形态学进行操作以及怎样将候选的人脸特征联系起来定位人脸。

Leungetal 提出了一种基于局部特征探测器和随机曲线图匹配的概率方法来定位复杂背景中的人脸<sup>[11]</sup>。他们的出发点是把人脸定位问题作为一个搜索问题来阐明,其搜索的目的是找到某种最象人脸的人脸特征布局。其中用到了五个特征(即两只眼睛、两个鼻孔、鼻子和嘴唇的交接处)来描述一张典型的人脸。对于任意一对相同种类的人脸特征(比如说左右眼睛对),计算出它们的相对距离,然后用高斯分布模拟出所有图像的同种类特征的距离。人脸模版被定义为用一组多方位、多路高斯派生滤波器作用于一个数据库里的一组人脸,其输出结果的平均值。给定一张测试图像,用滤波器对每个点进行匹配,将其与已定模版向量中的相应点进行比较。然后选择可靠率排在一二位的候选特征,将其用来搜寻其他的人脸特征。因为人脸特征的位置不是随意排列的,其他特征的期望位置可以通过使用相对距离的统计模型进行估算。此外,估算值的协方差能够被计算出来。因此,人脸特征的期望位置能够被高效的估算出来。只有合适位置的候选特征组成特征群,至此我们可以确定出最象人脸的特征群。寻找最好特征群的问题就被认为是一个曲线匹配问题,曲线上的点代表人脸特征,弧长代表特征间的距离。特征群的排队是基于一个概率密度函数,一个特征群是否被判断为人脸取决于其

是否是人脸的概率。他们使用 150 幅图像进行实验,对于任意一个特征群,如果它定位到三个或三个以上人脸特征,则认为正确检测到人脸。这个系统的正确定位率达到 86%。

除了使用特征群中的特征的相对距离描述人脸特征间的关系外, Leung et al 还提出了另外一种模拟人脸的方法。该方法中,特征群的表示和排列是通过使用形状统计理论(由 Kendall、Mardia 和 Dryden 提出)完成的<sup>[12][13]</sup>。形状统计表是由  $N$  个特征点上的概率密度函数值构成的,每个特征点表示为,对于第  $i$  个特征点,假设其原始特征点符合二维高斯分布。他们使用极大可能法确定人脸位置。这些方法的一个优点是它们能够定位出部分被遮挡的人脸。然而,我们不确定这些方法能否有效检测出同一场景中的多张人脸。

Yow 和 Cipolla 提出一种基于特征的方法,该方法使用了大量从视觉图像获得的证据以其它它们的上下文证据<sup>[17][18]</sup>。第一阶段用一个伸缩长宽比为 3:1 的派生高斯滤波器,作用于原始图像。我们感兴趣的是滤波器检测到的局部最大值点,它们表示了人脸特征的可能位置。第二阶段审查兴趣点周围的边缘并把它们划到不同的区域。对边缘的分组是基于它们的方向和粗细的接近度和相似度。然后,计算出区域特征的测量值,比如说边缘长度、边缘粗细、强度大小,并把结果存在特征向量中。利用人脸特征的训练数据把人脸特征向量的平均矩阵和协方差矩阵计算出来。当相应特征向量之间的 Mahalanobis 距离小于一定的阈值,则其所在的图像区域就成为有效的人脸特征候选。按照模型知识,即特征之间的相对位置,对标记特征进一步分组。然后,使用贝耶斯网络来评价每个人脸特征及其分组。这种方法的有吸引力在于它能够检测不同方向和多姿态的人脸。以 110 幅不同大小、方位、观察点的人脸图像作为测试数据,这种方法的检测率达到了 85%<sup>[16]</sup>。然而,该方法的误检率达到 28%,并且它只对像素大于  $60 \times 60$  的人脸有效。后来,有人用活性等高模型对该方法进行了改进)。图 2-1 对基于特征的人脸检测方法进行了总结<sup>[17][18][19]</sup>。

Yow 和 Cipolla 将人脸视为一个含六个定向人脸特征(眉毛、眼睛、鼻子、嘴巴)的平面模型,每个人脸特征建立由一对有向边界组成的模型。从感兴趣点开始进行特征选择,随后进行边缘检测和连接,并用统计模型进行测试。

Takacs 和 Wechsler 提出了一种有生物根据的人脸定位方法,该方法基于一

个视网膜特征提取和眼睛细微运动的模型<sup>[20]</sup>。他们的算法当中，用一个模拟人类视觉系统中的神经节细胞的视网膜格子作用于感兴趣区域或突出区域。第一阶段通过对图像进行粗糙的扫描过滤，估计人脸的位置。每个区域包含一定数量的神经元，这些神经元都已经受高斯滤波器的作用，调整到特定方位。第二阶段通过使用更好的策略扫描图像，提炼出突出区域，将人脸局部化。用 426 幅图像（其中 200 幅取自 FERET 数据库）进行测试，该方法的错误率为 4.96%。

Han 等开发了一种基于形态学的技术来提取眼睛模式段（eye-analogue-segments），检测人脸<sup>[21]</sup>。他们认为眼睛和眉毛是最明显和最稳定的人脸特征，因此，它们对检测非常有用。他们将眼睛模式段定义为眼睛轮廓的边缘。首先，对图像进行形态学操作，如闭运算，去掉低谷，用阈值提取出亮度发生明显变化的点，这些点成为眼睛模式点。然后，对眼睛模式点执行标记处理，形成眼睛模式段。使用这些段来引导搜寻潜在人脸区域，这些区域是由眼睛、鼻子、眉毛和嘴巴按一定几何关系组合起来的。最后，用类似的神经网络法对候选人脸区域进行进一步验证。他们的方法在包含 130 张人脸的 122 幅图像的测试数据上的测试精确率为 94%。

Amit 等提出了一种用来检测形状的方法，并且将它用来检测静态强度图像中的正面人脸。检测分为两个阶段完成：即聚焦和密集分类。聚焦是根据边缘段的空间排列进行的，这些边缘段是从基于亮度差别的简单边缘检测器中提取出来的。定义出类似空间排列的 Arich 族、一定光度范围的不变量和几何变换。利用一套 300 幅人脸图像的训练集，选出象人脸的特定空间安排。同时，使用 CART 算法以训练图像为依据构造一棵分类树，并且构造从普通背景图像中识别的错误人脸(false positives)集合。给定一幅测试图像，从边缘段的空间排列中识别出感兴趣的区域。利用 CART 树把每个感兴趣区域划分为人脸或背景。他们的实验是在从 Olivetti(now AT&T)获得的 100 幅图像上做的，每 1000 个像素点中，其积极错误率为 0.2%，消极错误率为 10%。<sup>[22][23][24]</sup>

## 2.2.2.2 纹理

人脸有固有的纹理，可以用来区分于不同对象。Augusteijn 和 Skufca 开发了一种方法，通过鉴别类似于人脸的纹理推断人脸的存在<sup>[25]</sup>。通过使用从  $16 \times 16$



子图像得到的二次统计特征 ( SGLD ) 计算出纹理。考虑三种纹理：皮肤、头发、其他。他们使用一种层叠交互神经网络管理纹理分类，并使用一种 Kohonen 自组织特征图形成不同的纹理类群。为了从纹理标签推断出人脸的存在，他们建议使用头发和皮肤纹理。然而，文中只告知了纹理分类的结果，而没有讲述人脸的定位或检测。

Dai 和 Nakano 也使用了 SGLD 模型进行人脸检测<sup>[26]</sup>。颜色信息被并入人脸特征模型。他们使用人脸特征模型为彩色背景中的人脸检测设计了一个扫描计划，背景中包含人脸的橘黄色部分被增强。这种方法的优点是它能够检测非垂直人脸或包含类似胡子、眼镜之类特征的人脸。该方法在包含 60 张人脸的 30 幅图像检测集上的检测率很高。

### 2.2.2.3 肤色

人类的肤色特征已经被使用，并在从人类检测到跟踪的很多应用中被证明是有效的特征。虽然不同的人有不同的肤色，有几个研究已经表明肤色的主要差别不是在于它们的色度而主要在于其亮度<sup>[27][28]</sup>。已经有几个颜色空间被用来标记作为人类的像素，包括 RGB，标准化 RGB，HSV（或 HSI），YCrCb，YIQ，YES，CIEXYZ，和 CIE LUV。

至今，人们已经提出了很多方法来建立肤色模型。最简单的模型是使用  $Cr$ 、 $Cb$  值<sup>[29]</sup>，例如从肤色像素样本得到  $Cr$ 、 $Cb$  值  $R(Cr, Cb)$ ，来定义皮肤色调像素区域。选择合适的阈值  $[Cr1, Cr2]$  和  $[Cb1, Cb2]$ ，如果一个像素的颜色值  $(Cr, Cb)$  落入该区域，即  $Cr1 \leq Cr \leq Cr2$  及  $Cb1 \leq Cb \leq Cb2$ ，则认为该像素点有皮肤色调。在假定皮肤是由像素组成的前提下，Crowley 和 Coutaz 使用一个直方图  $h(r, g)$  获取得到精确 RGB 向量的可能性，其中  $h(r, g)$  是在标准 RGB 颜色空间中的  $(r, g)$  值。换句话说，对于一个像素，如果  $h(r, g) \geq n$ ，其中  $n$  是根据经验从直方图样本选择的阈值，则该像素被认为是皮肤。Saxe 和 Foulds 提出了一种迭代的皮肤验证方法，使用了 HSV 颜色空间中的直方图交集。用户选择一组初始肤色像素，也称为控制种子，来初始化迭代算法。为了检测肤色区域，他们每次用一组像素在图像中移动，并将控制直方图与从图像中获得的当前直方图进行比较，比较时使用的是直方图交集。如果匹配程度或相似点个数（例如交集）大于某一阈值，

则当前的那组像素被认为是肤色。Kjeldsen 和 Kender 定义了一个 HSV 空间中的颜色断言来将肤色区域和背景区分开来。不同于以上提到的无参数方法, 高斯密度函数和混合高斯函数经常被用来建立肤色模型。单峰高斯分布中的参数往往是通过极大似然法进行估计的。使用混合高斯法的原因是, 我们观察到来自不同种族背景的人的皮肤的颜色直方图不符合单峰分布, 而是一个多峰分布。混合高斯函数中的参数通常是使用 EM 算法估计出来的。最近, Jones 和 Rehg 进行了一个大规模实验, 在标准 RGB 颜色空间中, 收集了将近十亿标志肤色的像素点。将检测肤色的直方图模型与混合模型进行比较, 他们发现直方图模型在精度和计算耗费方面更优一些。

如果肤色模型能够完全应用于不同光照环境, 那么肤色信息对于鉴定人脸区域和特定的人类特征是一个有效的工具。然而, 当光源光谱有很大差别的时候, 那种肤色模型久不再有效。换句话说, 由于背景和前景光照的改变, 颜色的外观是不稳定的。虽然, 颜色的色感一致性问题已经通过基于物理的模型进行详细阐述, 但是还是有人提出了几种在不同光照条件下使用皮肤颜色的方法。McKenna 等提出了一种在不同光照条件下使用适应性颜色混合模型追踪人脸<sup>[30]</sup>。撇开基于颜色色感一致性的肤色模型不用, 他们使用一种随机模型来联机估计对象的颜色分布, 该方法能够容忍观察和光照条件的改变。初步结果表明他们的系统能够在一定范围的光照条件下追踪人脸。然而, 这种方法不能用于在单张图像中检测多张人脸。

在检测或追踪人脸的过程中, 单独使用肤色往往是不够的。最近, 开发了几种联合使用形状分析、颜色分割及运动信息的标准系统, 用来在一个图像序列中定位或追踪人头和人脸。我们将在下一部分回顾这些方法。

#### 2.2.2.4 多特征

最近, 提出了很多联合几个人脸特征来进行人脸定位或检测的方法。其中的大部分方法是利用如: 皮肤颜色、大小、形状等全局特征来找出候选人脸, 然后用局部特征验证这些候选人脸, 例如眉毛、鼻子和头发。一个典型的方法是, 开始的时候检测出像皮肤的区域(如 2.2.2.3 所描述), 然后利用组合分析或集群算法将似人脸像素分组并联系起来。如果被联系起来的区域有椭圆形或卵形, 那

么该区域就成为候选人脸。最后，利用局部特征进行验证<sup>[31]</sup>。

Yachida 等提出了一个使用模糊理论检测彩色图像中人脸的方法<sup>[32]</sup>。他们使用两个模糊模型描述 CIE XYZ 颜色空间中的皮肤和头发颜色分布，五个脑部形状模型（一个正面四个侧面）抽象出图像中人脸的外观。每个形状模型是包含  $m \times n$  个正方形单元的二维模式，每个正方形单元可能包含几个像素。给每个单元分配两个属性：皮肤比例和头发比例，它们表明了一个单元中皮肤区域或头发区域占的比例。在一个测试图像中，根据分布模型，每个像素被归类为头发、人脸、头发和人脸的边界、头发和背景的边界，因此产生了似皮肤区域和似头发区域。然后将抽取出来的似皮肤区域和似头发区域与脑部形状模型进行比较，如果相似，则检测区域就成为候选人脸。为了验证候选人脸究竟是否是人脸，使用水平边缘提取出候选人脸的眼镜、眉毛、鼻子、嘴巴等特征。

Sobottka 和 Pitas 提出了一种使用形状和颜色进行人脸定位和人脸特征提取的方法。首先，使用 HSV 空间中的颜色片断定位似人脸区域。然后通过使用粗糙的方法进行区域增长确定连续区域。对于每个连续区域使用几何矩计算出最合适的椭圆。选择最接近椭圆的连续区域作为候选人脸。接着，在作为候选的连续区域中搜寻人脸特征，验证候选是否是人脸。像眼睛、嘴巴这样的特征，通过观察，它们比人脸的其他区域黑，由此将它们抽取出来。使用高斯肤色模型对肤色像素进行分类。为了表示二值图像中像素族的特征，他们使用傅立叶变换和辐形 Mellin 变换计算出 11 个低序几何矩。利用抽取出来的几何矩训练一个神经网络，用于检测。他们的实验在 100 幅图像的测试集上检测率为 85%<sup>[33][34]</sup>。

人脸模型的对称性也被用来定位人脸<sup>[35]</sup>。在 YES 颜色空间中，使用有条件类密度函数进行皮肤或非皮肤分类，然后进行平滑操作便于产生连接区域。接下来，使用一个椭圆人脸模版，基于 Hausdorff 距离，确定肤色区域与模版的相似度<sup>[36]</sup>。最后，利用人脸中眼睛位置固有的对称性，使用几个价值函数，定位出眼睛中心。鼻尖和嘴巴中心的位置则通过它们与眼睛中心的距离来确定。这个方法的缺点是它只对单张正面人脸或两只眼睛都可见的情况有效。而后提出了一个使用颜色和局部对称性的简单方法<sup>[37]</sup>。

对比于基于像素的方法，提出了一种基于结构、颜色、几何形状的方法。首先，执行多路分割抽取出一幅图像中的相似区域。使用高斯肤色模型，抽取肤色

区域，归如椭圆组。如果在这些椭圆区域中存在像眼睛和嘴巴之类的人脸特征则人脸被检测出来。实验结果表明这种方法能够检测出包含人脸特征（如胡子、眼镜）的不同方位的人脸。

Kauth 等提出了一种块表示法，对多谱卫星图像进行周密、结果化的描述。通过将像素坐标与像素的光谱（纹理）元素串联起来，形成每个像素的特征向量，然后使用这些特征向量将像素进行聚类形成固有的连接区域，或称之为“块”。为了检测人脸，每个特征向量都包含图像坐标和标准色度，如。然后，使用一种连通性算法来增大“块”，结果得到的与规范人脸大小、形状相似的皮肤块即被认为是人脸。

Kim 等使用范围和颜色进行人脸检测。假设背景像素有相同的颜色深度，并且背景中的像素数多于前景对象中的像素数，使用差分直方图计算出差分图表，把对象从背景中分割出来。在标准 RGB 颜色空间中使用高斯分布，具有类肤色的区域则被认为是人脸。Darrell 等提出了一种类似的方法用于人脸检测和跟踪。

## 2.2 基于模版匹配的方法

在模版匹配中，通过人工事先定义或用函数参数化一个标准人脸模式（通常是正面的）。给定一张输入图像，分别计算出人脸轮廓、眼睛、鼻子、嘴巴标准模式的相关值。人脸的存在就是由这些相关值决定的。这种方法的优点是执行起来比较简单。然而，实验证明它不足以检测出人脸，因为它不能有效地处理大小、姿势、形状改变的人脸。此后，为了达到大小、形状不变的目的，提出了多方案、多路、子模版和可变形模版等概念。

据记载，早期的时候 Sakai 等就试图检测出照片中的正面人脸<sup>[38]</sup>。他们是用眼睛、鼻子、嘴巴和人脸轮廓的子模版建立的人脸模型。每个子模版是根据线段来定义的。输入图像中的线是基于梯度的最大改变值提取出来的，然后将它们与子模版中的线进行比较。首先计算出子图像与轮廓模版的相似度，检测出候选人脸位置。然后，在候选位置进行与其他子模版的比较。换句话说，第一步是确定关注区域或感兴趣区域，第二步是探测细节确定人脸的存在。在后来的人脸检测工作中使用到了关注区域和子模版的思想。

Craw 等提出了一种基于正面人脸的形状模版的定位方法<sup>[39]</sup>。首先使用 Sobel

滤波器提取出边缘。基于几个限制条件将提取出的边缘组合起来,寻找人脸模版。在定位到头部轮廓后,在不同刻度反复使用此相同过程定位出眼睛、眉毛、嘴唇之类的特征。随后,Craw 等描述了一个定位方法,使用 40 个模版用于寻找人脸特征,并用一个控制策略来引导,最后从基于模版的特征检测器中获得检测结果<sup>[40]</sup>。

Govindaraju 等提出了一个两步人脸检测法,产生人脸假定并进行正面<sup>[41][42][43]</sup>。以边缘定义的特征为依据建立人脸模型。这些特征描述了正面人脸的左右边界和头发边缘线。使用 Marr-Hildreth 边缘检测器获取输入图像的边缘图。然后用滤波器除去不可能是人脸组成部分的对象。接下来,基于它们的相邻性和相对方位将小的轮廓对联系起来。找出拐点,将轮廓分割成特征曲线。核对特征曲线的几何特性和相对位置,对其进行标记。如果特征曲线的属性兼容则用边缘线将其连接起来。将形成边界的特征对之间的比例与黄金比例进行比较,并给边缘赋予一个代价值。如果一组中的三条具有不同标记的特征曲线的代价很低,则这一组特征曲线就成为假设。当检测报纸文章中人脸时,为了选择出最好的假设,从输入图像的标题中获取间接信息,了解图像中所包含的人数。对 50 幅图片的测试集进行测试,他们的系统的检测率大概为 70%。然而,检测的人脸必须是直立的、未重叠的并且是正面的。Venkatraman 和 Govindaraju 使用小波算法延伸出相同的方法用于提取边缘。

Tsukamoto 等提出可一个人脸模式的定性模型(QMF)。在 QMF 方法中,每个样本图像被分成一些块,对每个块估算出其定性特征<sup>[44]</sup>。在这个模型中,为了参数化一个人脸模式,“光照”和“边界”被定义为特征。从而,使用该块模版计算输入图像中每个位置的人脸可能性。如果人脸可能性大于某一预先定义的阈值,则一张人脸就被检测到。

侧面影像也被用来作为人脸定位的模版。对人脸样本进行主成分分析,获取一组基本的人脸侧面影像,这些侧面影像是由一些位组成的,它们和普通的 Hough 变换一起被用来定位。提出了一种基于人脸结构的多模版定位方法。他们的方法当中为可能的人脸特征定义了大量假设。人脸存在的一组假设被定义为使用 Dempster-Shafer 理论得到的人脸结构假设<sup>[45]</sup>。给定一幅图像,特征检测器计算出人脸特征存在的信心因素。这些因素合起来就能决定相信或不相信人脸的存

在。他们的系统能在 94 幅图像中定位出 88 幅图像中的人脸。

Sinha 使用一组空间图像不变量描述人脸模式的空间特性<sup>[46]</sup>。他设计这些不变量的原因在于，当变量随着光照的改变而改变人脸不同部分（如眼睛、面颊、额头等）的亮度时，这些部分的相对亮度基本保持不变。确定一些类似区域的成对亮度比，保留这些比例的大致趋势，比如一个区域比另一个区域亮更或更暗，这些比例就给我们提高供了一个很好的不变量。因此，观察到的亮度规律被编码成一个粗糙的人脸空间比例模版，这个模版中包含适当选择的对应于主要人脸特征的子区域，如眼睛、脸颊和额头。通过子区域间的一组适当的明 - 暗关系对获取人脸特征间的亮度限制。如果一幅图像符合所有的明 - 暗限制，那么一张人脸就被定位到。这种使用局部邻接区域的亮度差别的思想后来被延伸到基于小波的行人、车辆和人脸的检测。Sinha 的方法已经被延伸，并在一个机器人视觉系统中被用于人脸定位，定义了 23 个关系的增强模版。这 23 个关系被进一步分为 11 个基本关系和 12 个有效关系（箭头线）。图像中的每条箭头线表明一个关系，箭头线的头部指示第二个区域（也就是分数的分母）。如果两个区域间的比例超过了某个阈值则其对应的关系就满足人脸模版，并且当基本关系和有效关系的个数超过某一阈值时，人脸就被定位到。

Miao 等提出用一种分层模版匹配法进行人脸检测<sup>[47]</sup>。第一步，为了便于处理旋转图像，将输入图像按 5 度的步长从 -20 度转到 20 度。形成图像分层，用拉普拉斯变换提取边缘。人脸模版包含六个人脸结构元的边缘：两个眉毛、两只眼睛、一个鼻子、一张嘴巴。最后使用启发式方法决定人脸的存在与否。他们的实验结果在含单张人脸的图像（正面或经过旋转）中优于含多张人脸的图像。

Yuille 等使用可变形模版建立人脸特征模型，该模型将优先弹性模型用于人脸特征。这个方法中用参数化模版描述人脸特征。定义了一个能量函数将输入图像中的边缘、尖峰、低谷与模版中的相应参数联系起来。通过最小化这些参数的能量函数，找到最合适的弹性模型。虽然他们的实验在跟踪非刚性特征中得到了很好的结果，但这种方法的一个缺点是可变形模版必须在感兴趣对象的附近进行初始化。开发了一种基于蛇爬模版的检测法<sup>[48]</sup>。首先，使用模糊滤波器将图像缠绕起来，然后用形态学方法增强边缘。使用经过修正的  $n$  个像素组成的蛇形寻找并除去小的曲线段。对保留下来的小蛇使 Hough 变换来寻找一个主导椭圆，

每张人脸最终被近似化为一个椭圆。因此，获取描述椭圆的四个参数集，并将其用来作为人脸定位的候选。对于每个候选，使用一种类似于可变形模版法的方法寻找细节特征。如果人脸特征数目的真实值被找到了，并且他们的比例满足基于人脸模版的比例测试，那么就认为一张人脸被检测到。Lam 和 Yan 也使用蛇形来定位脑部边界，在最小化能量函数时使用了贪婪算法<sup>[49]</sup>。

Lanitis 等描述了一种基于形状和亮度信息的人脸表示法<sup>[50]</sup>。他们首先获取一组训练图像，对图像中的样本轮廓，如眼睛边界、鼻子、下巴和面颊，进行手工标记，用样本点的向量来表示形状。他们使用点分布模型（PDM）来刻画所有人的形状向量，用类似于 Kirby 和 Sirovich 的方法表示标准形状亮度。在新图像中，使用积极形状模型（ASM）搜索，估计出人脸位置和形状参数，利用人脸形状的点分布模型定位人脸。然后将人脸块变为平均形状，提取出亮度参数。为了进行分类，形状信息和亮度信息可以一起使用。Cootes 和 Taylor 使用了类似的方法定位图像中的人脸。首先，他们在包含重要特征例子的图像中定义矩形区域。然后，对这些训练特征进行因素分析，获得函数分布。如果可能性大于某一阈值则可以决定候选特征，验证的时候使用的是 ASM 法。通过对 40 幅图像进行训练，该方法能够在 40 幅图像中定位出 35 张人脸。也有人使用 Kalman 滤波器对 ASM 方法进行了扩展，估算与形状无关的亮度参数，并跟踪图像序列中的人脸<sup>[51]</sup>。

## 2.3 基于外观的方法

模版匹配法中模版是由专家事先预定，与模版匹配法不同的是，在基于外观的方法中模版是通过对图像中的样例进行学习而得来的。总的说来，基于外观的方法主要依靠统计分析和机器学习技术来寻找人脸和非人脸的相应特征。用分布模型或判别函数表达学习到的特征，用于后面的人脸检测。同时，此类方法中，通常使用降维处理来提高计算和检测效率。

很大基于外观的方法可以通过使用概率框架来理解。令从一幅图像中获取的图像向量或特征向量为一个随机变量  $x$ ，通过条件密度函数决定  $x$  的值。然后可以使用贝叶斯分类法和最大似然法决定候选图像位置是人脸或不是人脸。不幸的事，直接使用贝叶斯分类法是不可行的，因为  $x$  的维数太高，是多峰的，并且如

果受先天性因素影响，不好理解。因此，在基于外观方法上的大部分研究工作涉及使用实际有效的参数化或非参数化近似法计算。

基于外观方法的另一个实现途径是在人脸和非人脸之间找出一个判别函数（也就是判决面，分离超平面，阈值函数）。按照惯例，将图像模式映射到低维空间，然后形成判别函数（通常基于距离）用于分类，或者用多层神经网络形成一个非线性判决面。最近，提出了支持向量机和其他内核方法。这些方法将模式映射到高维空间，然后在映射人脸和非人脸模式间构造决策面。

### 2.3.1 本征脸法

在人脸识别中使用本征向量的一个早期人物是 Kohonen，在他的方法中，示范了用一个简单的神经网络识别排列好的标准人脸图像。通过近似化计算图像自动关联矩阵的本征向量，神经网络得到关于人脸的描述。这些本征向量就是后来我们知道的<sup>[52]</sup>本征脸。

Kirby 和 Sirovich 证明，通过使用适量的基本人脸能够对人脸图像进行线性编码。其证明是基于 Karhunen-Loe`ve 变换和 Hotelling 变换，其中 Karhunen-Loe`ve 变换也有其他的名字，比如主要组件分析。Perason 在 1901 年的时候首先提出了这个想法，随后是 Hotelling 在 1933 年提出。给定一个  $n \times m$  训练图片的集合，表示成一个大小为  $m \times n$  的矢量，确定最优子空间中的基本向量，因此可以最小化将训练图像映射到子空间和原图像过程中的平均错误。他们称该最优向量集为本征图，因为这些向量仅是训练集向量化人脸图像的协方差矩阵的本征向量。用 100 幅图像对该方法进行实验，结果表明只要使用 50 幅本征图就能够对  $91 \times 50$  的人脸图像进行有效编码，而实验过程中只使用了 95% 的变量。

Turk 和 Pertland 把主要组件分析技术应用于人脸识别与检测中<sup>[52]</sup>。关于人脸图像训练集合的主要组件分析技术用于生成本征图片（这里称为本征人脸图片），它扩展了图片子空间（这里称为人脸空间）。把人脸图像映射于那些子空间，并进行归类。类似的，把非人脸训练图片映射到相同的子空间并归类。因为人脸图片在投影时并没有本质改变，因此非人脸图片看起来就大相径庭了。为了检测到一个场景中的人脸，在此图片中的所有位置，图片区域和人脸空间的距离一定要被算出来。此距离是 faceness 的一个度量指标，上面的计算结果就是一张



“人脸图”。然后，可以利用人脸图的局部极小值检测到人脸。很多关于人脸检测、人脸识别和特征提取的文章都采用了本征向量分解和聚集的思想。

### 2.3.2 基于分布的方法

Sung 和 Poggio 开发了一个以分类为基础的人脸检测系统，这个系统演示了一种物体类的图像模式分类能从正反两方面的事例（也就是图像）中得到。他们的系统由两部分组成：区分人脸/非人脸模式的分类模型和多层感知机分类器。每张人脸和非人脸样本首先被规范处理成  $19 \times 19$  像素的图像，作为一个 361 维的向量。接着，这些样本用如图 6 所示的改进的 K-means 算法分成 6 个人脸和 6 个非人脸一组的群组。每一组由高斯函数和协方差矩阵表示。

他们的方法中的测量距离量由输入图像模式和原型组群计算得出，第一个距离是测试样本和组的质心之间的规范的 Mahalanobis 距离，由组的 75 个最大的特征向量组成的低维子空间里计算得出。第二个距离是待测样本和它自己在 75 维子空间上的投影之间的欧式距离。用于记录样本区别的距离并不是由第一个距离单元测量。最后一步是多层感知机网络（MLP）用 12 对人脸和非人脸组间的距离来区分人脸样本和非人脸样本。分类器用 47316 个样本采用标准的反向传播算法进行训练，其中有 4150 个人脸样本，其余的是非人脸样本。需要说明的是，收集有代表性的人脸样本是容易的，但是要得到有代表的非人脸样本却是非常困难。这一问题可以通过 bootstrap 算法得以缓解，即在训练过程中，有选择的增加图像进行训练。开始时，先用少数的非人脸样本作为训练样本集，多层感知机用此样本进行训练。接着，运行人脸检测系统来检测一系列的随机图片并且把所有的错误判断为人脸的非人脸样本，并把这些样本作为新的非人脸样本加入到非人脸样本的训练集中来。Bootstrap 方法避免的准确收集有代表性的非人脸样本所带来的困难，并且这种方法在后来的工作中被用到<sup>[53][54]</sup>。

一个可能的基于在高维空间密度估计的视觉学习方法是由 Moghaddam 和 Pentland 开发的。主成分分析（PCA）是用来定义子空间最好的表现人脸样本集。主成分分析保护了数据的主要线性相关性和丢弃次要的。这种方法把向量空间分解成两个相互排斥和互补的子空间：主要子空间（或者特征空间）和它的直角互补。所以，目标密度被分解为两个部分：主要子空间的密度（主要成分）和

它的补角（在标准 PCA 中被放弃）。多元高斯和混合高斯被用来学习脸的静态局部特征。这些可能密度被用来对事物进行基于最大形似估计的检测。推荐的这种方法已经被应用到了人脸定位、译码和人脸识别中来。和经典的 eigenface 方法比起来，推荐的这种方法在人脸识别上表现的更好。在人脸检测上，这种技术只能在局部演示；

还有一种基于混合要素的检测方法。要素分析法[FA]是一种统计学方法，它利用少量的隐藏变量来构建高维的协方差结构。FA 方法在很多方面类似于主成分分析法（PCA）。尽管如此，和 FA 不同的是，PCA 并没有为数据定义合适的密度模型，因为译码一个数据指针和指向主成分空间的开销是相等的（也就是，密度沿着这些方向是没有规格化的）。另外，PCA 对特征数据的独立噪声没有鲁棒性，因为主成分分析法最大化了输入数据变量，所以保留了不期望的变量。合成或者真实的样本表明来自 PCA 子空间的不同类经常被涂污。在样本又特定结构的情况下，从分类的观点看，PCA 是不十分理想的。Hintonetal.已经把 FA 应用到了数字识别当中，他们比较了 PCA 和 FA 模型的性能。一种要素分析混合模型已经在应用到人脸识别当中。两个研究都表明，FA 在数字和人脸识别中的性能都优于 PCA。因为姿势、角度、表情和灯光影响着人脸的表现，所以，图像中的人脸能用多形态密度模型来更好的表示，在这种模型中，每一个形态都用来表现人脸的特定的特征。他们用一种可行的被称为混合要素分析（MFA）的方法来检测多变的人脸。在混合模型中的参数由 EM 算法来估计的。

第二种方法是用 Fisher 线性判别式（FLD）来把高空间的样本图像投影到低维特征空间。近来，在对几个数据集，包括在可变光条件下建立的 Yale 人脸数据库进行人脸识别时，使用 Fisher 方法和其他基于线性判别分析的方法已经表现出胜过广泛应用的 Eigenface。一个可能的解释是 FLD 在模式分类的时候比 PCA 提供了更好的投影，因为它的目标是找到更多的判别式的投影方向。所以，在投影子空间的分类结果可能要别其他的方法好。在第二个推荐的方法中，他们使用 Kohonen ' s Self Organizing Map(SOM)把训练的人脸和非人脸分到不同的子类。图九显示了每个人脸类的原形。在这些重新标明的例子里，内部类和中间类非别计算，所以产生了理想的基于 FLD 的投影。对每一个子类，它的密度由通过最大可能性的进行参数估计的高斯函数来模拟。为了检测出人脸，每一个输入图像

都通过计算可能的人脸区域的矩形窗口进行扫描。最大可能性规则用来决定是否人脸被检测到。两种方法用由 619 个人脸组成的 225 幅图像数据集进行测试,结果用 MFA 方法正确检测率为 92.3%,使用 FLD 方法的为 93.6%。

### 2.3.3 神经网络

人工神经网络已经成功地应用在许多模式识别问题上,比如字符识别、目标识别和自动机器人驾驶。因为人脸检测可以看作两种模式的识别问题,所以多种神经网络结构可以被应用了。使用人工神经网络的优点是系统网络训练识别复杂人脸模式的可行性。但是网络结构有一个缺点,那就是它必须不断地协调来(许多的层,许多的节点,学习效率等等)取得更好的性能。

一种早期的使用分层神经网络的方法是由 Agui 提出的。第一个阶段由两个平行的子网络组成,子网络的输入值是来自原始图像的值和一个  $3 \times 3$  的 Sobel 滤波图。第二个阶段的输入是由子网络的输出和抽取的特征值组成的,特征值包括输入样本像素的标准误差、窗口中白色像素和总的二值像素的比例、几何力矩。在第二个阶段的输出表明了输入区域人脸的存在。实验表明,如果在检测图片中所有的人脸的大小是相同的话,这种方法是能够检测出人脸的。Propp 和 Samal 开发了用于人脸检测的最早的神经网络。他们的网络由 4 层组成,包括 1024 个输入节点,256 个一层隐节点、8 个二层隐节点和两个输出节点。后来,一个相似的分层网络被提出。早期的 Soulie 的方法是通过时间延迟神经网络扫描输入图像(使用  $20 \times 25$  像素大小的图像)来检测人脸。为了应付大小的不同,输入图像通过小波变化来处理。他们称通过对 120 幅图像进行检测,错误检测为人脸的几率是 2.7%,错误检测为非人脸的几率是 0.5%。Vaillant 用回旋神经网络来检测图像中的人脸。首先制作成  $20 \times 20$  的人脸和非人脸样本,一个神经网络是以一定的比例来找到大概的人脸区域,另一个网络是以相同的比例来决定人脸的精确位置。给定一副图片,第一个网络先把那些可能存在人脸的区域确定为候选区,这些候选区再由第二个网络进行验证。Burel 和 Carel 为人脸检测提出的神经网络采用 Kohonen 的 SOM 算法,可以把大量的训练样本压缩成少量的样本。为了分类人脸和背景,多层感知机被用来学习这些样本。检测阶段要对每幅图像进行不同的扫描,对每一个确定的不同大小的窗口,要规范到标准的大小,并且要标

明平均值和变化以减小光照条件的影响。接着，每一个规范了的窗口通过 MLP 进行分类。

Feraud 和 Bernier 提出了使用自关联神经网络 ( autoassociative network ) 的方法.这一主意是建立在提出的 5 层自关联网可以处理非线性分析的观点之上的。一个自关联网被用来检测正面人脸，另一个被用来检测左右旋转 60 度范围的人脸。在整个子关联网中，用一个门限网络 ( A gating network ) 来为正面和旋转人脸检测器分配权值。这种方法也被用在了 LISTEN 和 MULTRAK 中。

Lin 提出了基于概率决策的神经网络 ( PDBNN )。PDBNN 网络结构和径向基函数网络 ( RBF ) 是很相近的，它的学习规则和概率阐明得到了改进。它并不是把整个人脸图像都转化到一个训练的值空间，而是首先依据人脸区域的值和边缘信息，包括眼眉、眼睛和鼻子来精确特征向量。这两个精确的特征向量被反馈到 PDBNN 网络，并由输出来决定分类结果。对 23 幅由 Sung 和 Poggio 提供的图像集进行人脸检测，试验结果表明，它的性能可以和其他的学习神经网络相比。

在所有使用神经网络进行人脸检测的方法中，最有意义的工作是由 Rowley 所作的。用一个多层神经网络通过人脸和非人脸图像来学习人脸和非人脸模式（也就是像素的值和空间联系），在这当中，Sung 通过距离测量来找到分类判别式用以分类人脸和非人脸模式。他们用多层神经网络和几个判断方法来提高性能，但 Burel 和 Carel 用一个单层的网络，Vaillant 也是。用两个网络来分类。这里有两个主要的组件：多重神经网络（检测人脸模式）和判决模块（通过多重神经网络给出最终的判决）。

这种方法的第一个组件是接收一个  $20 \times 20$  的像素区域，并输出  $(-1, 1)$  区间上的值。给出一个测试模式，训练网络的输出就表明了非人脸（距离 -1 近）或者人脸（距离 1 近）。为了检测一幅图片中所有的人脸，神经网络被用到了图像的所有区域。为了检测大于  $20 \times 20$  的人脸，输入图像被不断的作二次抽样，然后用网络对每一个尺寸的图像进行检测。用将近 1050 幅的不同大小、不同角度、不同位置 and 不同亮度的人脸样本来训练网络。在每幅训练图片中，眼镜、鼻子尖、嘴角和嘴的中心被人工的标定，用以把人脸规范到相同的角度和位置。这个方法的第二个组件是合并重叠的检测并在多重网络的输出中判断。为了提高性能，还应用了简单的逻辑操作 ( AND/OR )。Rowley 设计的若干个系统使用了不同的中

断策略，比 Sung 和 Poggio 的系统有更小的计算代价，在对一个 24 幅包括 144 张人脸的图像集进行检测时表现出更高的检测率。Rowley 和 Sung 提出的方法的一个限制就是：他们只能检测出正立的端正人脸。近期，Rowley 开发了一个检测旋转人脸的方法，他用一个路由器网络来处理每一个输入窗口来判断可能的人脸旋转角度并且把窗口旋转到规范的角度；神经网络的旋转窗口就像上面所描述的。尽管如此，这个新的系统比起直立人脸检测系统，对直立人脸的检测率不高。然而系统对两个大型的检测集进行检测，只能检测出 76.9% 的人脸。<sup>[4]</sup>

#### 2.3.4 支持向量机 (SVM)

支持向量机 (SVM) 是由 Osuna 第一次应用到人脸检测的。SVM 能被看作一个新的一个新的训练多项功能、神经网络或者是径向基函数分类器的范例。大部分训练分类器的方法(比如贝叶斯、神经网络和 RBF)都是基于最小化训练错误，比如经验风险最小化；SVM 使用另外一个原则，叫做结构风险最小化，他的目标是使期望风险最小。SVM 是一个线形分类器，它的分割超平面被用来最小化未见过的测试模式的期望分类错误。这一理想的超平面是由训练向量的一个小的子集联合定义的，叫做支持向量。估计这一理想的超平面等价于解一个线形二次问题。但是要计算的有时间和存储容量。在中，Osunaet 为大尺度问题提出了一个有效的方法来选练 SVM，并且应用到了人脸检测。在对一个有 10,000,000 个  $19 \times 19$  的检测集进行检测时，他们的系统错误率很低，并且运行速度大概比 Sung 和 Poggio 的系统快 30 倍。SVM 也用到了检测人脸和小波领域<sup>[55]</sup>。

#### 2.3.5 SNow

Yang 提出了使用 Snow 学习体系方法来检测不同特征、不同表情、不同姿势以及不同光照条件的人脸。他们还研究了学习原始图像并且多尺寸特征的影响。SNow(Sparse Network of Winnows)是具有线性功能的稀有网络，它利用了 Winnow 更新规则。在学习时被表明，域中隐藏的一些特征占据了决策的大部分，但是已开始并不知道。它的学习体系得特点是它有较少的连接单元、特征的分配、数据的连接、决策机构和有效的更新规则的应用。在训练 SNow 训练器时，来自 Olivetti, UMIST, Harvard, Yale, 和 FERET 数据库的 1681 幅人脸图像被用来体现人

脸模式的变化。和其他方法比起来，他们报道的结果是指用两个包括 225 幅图片 619 张人脸的数据集，错误率为 5.9%。对数据集 1 进行检测，这一技术的性能和其他方法包括使用神经网络法、熵相对信息（Kullback relative information）、朴素 Bayes 分类器（naive Bayes classifier）和支持向量机的性能一样的好，而且计算更高效。表 4 显示的和和其他检测方法的性能比较<sup>[56]</sup>。

### 2.3.6 贝叶斯判决

和使用全局人脸的模式相反，Schneiderman 和 Kanade 使用了朴素 Bayes 分类器来估计人脸模式的局部表情和位置（人脸的子区域）的联合概率。他们强调局部，因为一个物体的局部模式比其他的更具有唯一性；眼睛周围的模式比脸颊周围的更重要。有两个原因使用朴素 Bayes 分类器。第一，它提供了子区域条件密度的更好的估计；第二，朴素 Bayes 分类器提供了后验概率的函数形式用以得到物体局部表现和位置的联合概率。在每一个尺度，一张人脸图像被分解到 4 个矩形的子区域。这些子空间用 PCA 方法映射到一个低维的空间里面，并且量化到一个有限的模式集，然后这些映射样本和局部表现被用来估计投影子空间的统计。在这一规则下，他们的方法决定：当相似率比先验概率大时，判断人脸的出现。对数据集进行检测，正确率为 93%，这一方法和比较显示了相当的性能，而且能够检测一些旋转人脸和人脸轮廓。Schneiderman 和 Kanade 用这种方法和小波结合来检测人脸轮廓和汽车<sup>[57]</sup>。

一个相关的方法是由 Rickert 提出的，它使用局部特征的联合概率模型。局部通过使用多尺度和滤波器来输入到图像。特征向量（比如滤波反映）通过聚类数据进行估计，接着组成一个混合高斯模型。在模型训练和进一步的精简后，测试图像通过计算图像特征与模型的相似度来分类。他们在人脸和汽车上的实验显示了很好的结果。

### 2.3.7 隐马尔可夫模型

隐马尔可夫模型潜在的假设是：模式能够表现为随机参数的处理过程和这个过程的处理参数能够被一个精确的有良好定义的行为所估计。在用 HMM 进行模式识别时的的问题是，需要先定义许多隐状态以构成模型。接着，我们可以用代表

观测值序列的例子来训练 HMM 学习状态间变迁的概率。选练 HMM 的目的就是通过使用标准的 Viterbi 分割方法和 Baum-Welch 法则来调节 HMM 模型的参数，以最大化观测值训练值的概率。在 HMM 训练之后，观测值概率输出决定属于哪一类。

直观上人脸模式能够分成若干区域，比如前额、眼睛、鼻子、嘴巴和下巴。以适当的顺序检测这些区域以检测人脸（从上到下和从左到右）。这种方法并不是像模板匹配或者基于表现的方法一样（人脸特征，比如眼睛和鼻子需要联合来代表一个参考点）依靠各个区域的联合，它的目标是使人脸区域和 HMM 连续的特征密度相结合。基于 HMM 的方法通常把人脸看作观测向量序列，其中的每一个向量都是一个像素。在训练和检测过程中，图像以某种顺序扫描（通常是从上到下）并且观测值被当作一块像素。为了检测人脸，像素间的边界通过状态间的转化概率来表示，并且区域内的图像数据由多变量的高斯函数来模拟。每一个观测序列都是由每一块的像素值所组成的，输出状态和观测值从属的类相对应。在 HMM 训练后，由观测值的输出概率来决定它属于哪一类。HMM 已经被用到了人脸识别和人脸定位中。Samaria 显示他训练的 HMM 状态和人脸区域相对应。也就是说，一个状态负责额头的观测向量，另一个状态负责人眼睛的观测向量。为了人脸的定位，HMM 通过收集的大量人脸图像，被训练的对普通人脸有识别能力。如果获得的图像中每个矩形模式的相似度高于一开始的，那么人脸就被确定了。

Samaria 和 Young 使用了 1D 和假冒的 2D 的 HMM 来抽取人脸特征和进行人脸识别。他们的 HMM 使用了人脸的结构来加强特征变化的约束。当重要的人脸区域，比如头发、额头、眼睛、鼻子和嘴以从上到下的自然的顺序出现时，那么每个区域被分配给一维连续 HMM 的一个状态。为了选练，每个图像都要从上到下的分成 5 个状态（比如每个图像被分成 5 个不重叠的平均大小的区域）。接着这几个分割被 Viterbi 分割取代，并且 HMM 的参数使用 Baum-Welch 算法进行重新估计。设一个宽  $W$  高  $H$  的图像被分成高  $L$  宽  $W$  的重叠块。连续的块的交叠在垂直方向有  $P$  行，这些图像观察序列块和训练的 HMM 用来决定输出。Nefian 和 Haye 应用 HMM 和 KLT 来进行人脸的定位和识别。它并不是用行强度值，它的观察向量是由输入向量的计算得出的 KLT 系数组成的。他们在人脸识别上的

试验表现出了更好的识别效率。对包括 432 幅只有单人脸的 MIT 数据库进行测试, 这种 2D 的 HMM 系统得成功识别率由 90%。

Rajagopalan 提出了进行人脸检测的两种概率方法的使用多元高斯来分布人脸相反, 第一种方法是使用高阶统计量 HOS(higher order statistics)进行密度估计。相似的是, 人脸和非人脸的未知分配都是用了基于高阶统计量 HOS 的六个密度函数。用一个多层感知机来进行分类, 输入向量是由图像模式和 12 个模型间的测量的 12 个距离组成的。使用的第二种方法是用 HMM 来学习一幅图像中从人脸到非人脸和从非人脸到人脸的变化。这种方法是基于从图像产生的观测序列和与之对应的 HMM 的学习参数。要学习的观测序列首先由计算子图像到由第一种方法估计产生的 12 个人脸和非人脸中心的距离得到的。学习完成后, 理想状态序列被进一步处理。试验结果表明, HOS 和 HMM 两种方法有较高的检测效率, 但是也有更多的错误警告<sup>[58][59]</sup>。

### 2.3.8 信息理论法

人脸模式的空间性质能通过不同的方面来模拟。人脸的约束关系是很有用的, 经常被应用到纹理的分割当中。人脸模式的约束常用小邻域的像素来标明。马尔可夫随机域(Markov random field(MRF))理论提供便利和一致的方式来模拟上下文依赖实体, 比如图像像素和相互特征。这可以通过使用有条件的 MRF 刻画实体间相互的影响来得到。按照 Hammersley-Clifford 法则, MRF 能够用 Gibbs 来等价的表示, 并且参数经常最大化后来的(MAP)估计。人脸和非人脸区分能够用柱状图来估计。使用熵相对信息, 两类的最大信息辨别的马尔科夫能够用来检测。

Lew 把熵相对信息应用到了人脸检测上, 他用概率函数  $p(x)$  代表模板是人脸, 用  $q(x)$  代表模板非人脸。一个由 9 张图包含 100 个人脸的数据库用来估计人脸的分配。非人脸概率密度是用柱状图对 143000 个非人脸模板的估计得来的。在训练集中, 具有最大的信息的像素(MIP)化  $p(x)$  和  $q(x)$  之间的熵相对信息(例如给出最大的类间隔)。试验表明, MIP 主要集中在眼睛和嘴的区域, 而不是鼻子。接着 MIP 被用来获得线形分类特征, 代表方法有 ukunaga 和 Koontz。为了检测人脸, 用一个窗口在图像上移动, 计算和人脸空间(DFFS)的距离。如果 DFFS



到人脸子空间的距离比到非人脸子空间的距离要小的化，则假设窗口中没有人脸。

Colmenarez 和 Huang 也用熵相对信息来最大化人脸样本正面和反面间的信息基础分配。来自每个类训练集的图像(人脸和非人脸)被看作随机过程的观测，并且可以用两个概率函数来表示。他们用离散的马尔科夫来模拟人脸、背景样式，并且估计出概率模型。学习过程被转化到一个理想的马尔科夫处理过程，它可以最大化两个类的基本区别信息。相似度是用训练概率模型来计算的，用以检测人脸。

Qian 和 Huang 提供了一种方法，这种做法采用了以视觉和模型为基础的策略。首先，采用视觉关注算法，这种高水准的该领域的知识，以便于缩小寻找范围。这种结果可以通过把目标集中在有探测运算法（分水岭法）则得出的区域地图上，从而获取图像范围。在一选择的范围内，就可以采用模版匹配及人脸特征的相结合方法来检测，这种发发采用了分等级的马尔科夫随机域和最大化的后验可能性。

### 2.3.9 归纳学习

归纳学习算法也被用在了人脸定位和人脸检测中。Huang 应用 Quinlan 的 C4.5 算法从人脸模式的正反面例子来学习决策树。每一个训练样本都是一个  $8 \times 8$  的窗口。通过这些例子，C4.5 建立了一个个分类器作为决策树，决策树的叶子代表类，节点代表对某一个单一属性的检验。学习过的决策树用来决定输入图像中是否存在人脸。对来自 FERET 数据集的 2340 张正立人脸的进行检测，实验结果表明，定位人脸的正确率为 96%。

Duta 和 Jain 使用 Mitchell 的 Find-S 算法来学习人脸的概念。他们断定人脸模式的分布  $p(x|face)$  能够用高斯函数来估计，并且从一个人脸事例到其中一个组质心的距离应该小于这个组的任何一点到质心的距离。接着用 Find-S 算法来学习人脸和非人脸能够区分开的距离。这种方法有许多独特的方法，第一它没有用非人脸样本，但是使用了人脸和非人脸样本；第二，只有人脸的中心部分被用来训练；第三，特征向量是由 32 个值或者纹理的图像组成的，但是使用的所有尺度的值最为输入。对 CMU 的第一个数据集合测试，这种方法得到了 90% 成功率

[60]  
。

## 2.4 本章小节

本章简要介绍了现在主流的人脸检测的三类方法的具体原理以及他们的优缺点，并指出用于检测主要依据的面部特征。其中，基于知识的方法算法相对简单而且直接，但是他对图像的质量要求比较高，稳健性较差；基于模版匹配的方法优点是执行起来比较简单，但是其检测成功率与模版选取有直接关系，选取合适的模版难度较大；基于外观的方法中模版是通过对图像中的样例进行学习而得来的，基于外观的方法主要依靠统计分析和机器学习技术来寻找人脸和非人脸的相应特征，但是算法的复杂度较高。

## 3 基于肤色检测模型的图像分割技术研究

### 3.1 引言

颜色是自然界事物的一个特性,而肤色是人与其他事物相对稳定的可区别特征。近几年,人脸的初步检测技术已经从单纯的人脸图像灰度转换到图像色彩、色度、饱和度等多种手段相结合。肤色特征在人脸检测技术中越来越广泛的应用主要基于肤色在色彩空间中良好的聚类性。研究显示:尽管不同种族、不同年龄、不同性别的人的肤色看上去不同,但这种不同主要集中在亮度上,在去除亮度的色彩空间里,不同人的肤色分布具有聚类性<sup>[61]</sup>。因此,彩色图像肤色特征提取的准确性将直接影响到后续的人脸检测研究及相关系统的成功,它是整个人脸检测系统的重要组成部分,是检测系统成功的前提因素。

肤色检测就要涉及到色彩空间与肤色模型的应用,也就是将输入图像的像素值从 RGB 原始空间转换到一定的色彩空间并运用相关肤色模型运算判断该像素是否肤色像素。

所谓色彩空间,指根据某种关系式将色彩进行表达的一种三维数学空间。

所谓肤色模型,也和其他的数学建模一样,它是将色彩空间中的肤色与其他颜色区别开来的一种数学关系模型,即用一种代数的(解析的)或查找表等形式来表达哪些像素的色彩属于肤色,或者表征出某一像素的色彩与肤色的相似程度。

目前,关于彩色图像肤色检测的研究很多,应用到的色彩空间主要包括 TSL、YCbCr、HIS、YIQ、YES、r-g、CIE 等,以及各种改进的空间;肤色模型主要有统计直方图模型(Statistical Color Model)、单一高斯模型(Single Gaussian Model)、高斯混合模型(Gaussian Mixture Density Model)等<sup>[62][63][64]</sup>。这些研究工作都取得了一定的成功,但是,经过分析、计算与试验,作者发现这些研究还存在许多不足之处:几乎所有的肤色检测都对图像进行原始空间 RGB 到其他色彩空间的转换,而这并没有增强肤色像素的聚类,相反许多空间转换只是局部而不是全局线性变换,如 HSI,这样可能导致肤色与非肤色区域交叉重叠,影响聚类效果;再者,相关研究运用肤色模型,然后根据概率等特征进行肤色提取,往往会

丢失相当部分的肤色像素，导致部分人脸区域无法被检测出来。究其原因，在于色彩空间和肤色模型的选取。

## 3.2 肤色检测模型

### 3.2.1 RGB 空间模型

由于彩色图像是光谱图像的一种特殊情况，相应于人类视觉的三基色即红，绿，蓝三个波段，是对人眼的光谱量化性质的近似，因此用  $R$ 、 $G$ 、 $B$  三个分量来表征颜色是一种很自然的格式，而且多数的图像采集设备都是以 CCD 技术为核心的，直接感知色彩的  $R$ 、 $G$ 、 $B$  三分量，这使得三基色模型成为图像成像、显示、打印等设备的基础。RGB 颜色空间模型主要应用于 CRT 监视器和图形刷新设备中。尽管该颜色空间是最普遍的，但是由于  $R$ 、 $G$ 、 $B$  三色之间存在强烈的相关性，因此在大多数的肤色分割中一般没有直接利用 RGB 彩色空间，而是利用其变化后的彩色空间进行分割。

这个模型基于笛卡尔坐标系统，三个轴分别为  $R$ 、 $G$ 、 $B$ ，见图 3-1。我们感兴趣的部分是个正方体。原点对应黑色，离原点最远的顶点对应白色。在这个模型中，从黑到白的灰度值分布在从原点到离原点最远顶点间的连线上，而立方体内其余各点对应不同的颜色，可用从原点到该点的矢量表示。为了方便我们将立方体归一化为单位立方体，这样所有的  $R$ 、 $G$ 、 $B$  的值都在区间  $[0, 1]$  中。

由于 RGB 颜色空间中的向量  $(r, g, b)$  表示了一种颜色，相同方向，不同模型的向量具有相同的色度，只是亮度不同。颜色  $(r, g, b)$  的色度坐标定义为各个分量在  $R+G+B$  所占的比例，即：

$$\begin{cases} r = \frac{R}{R+G+B} \\ g = \frac{G}{R+G+B} \\ b = \frac{B}{R+G+B} \end{cases} \quad (3-1)$$

根据这个模型，每个彩色图像包括三个独立的基色平面，或者说可分解到三个平面上。反过来如果一幅图像可以被表示为三个平面，使用这个模型比较方便。

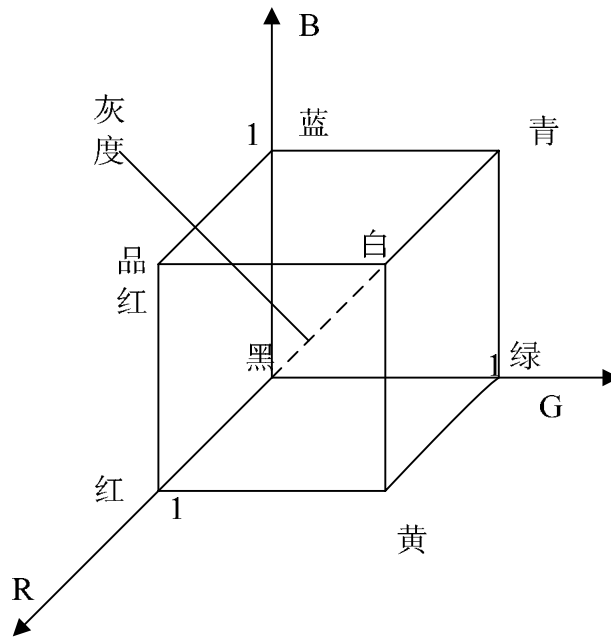


图 3-1 RGB 空间模型

### 3.2.2 HSI 空间模型

HSI 彩色空间是一个柱状彩色空间。这种格式反映了人类观察色彩的方式，同时也有利于图像处理。在对色彩信息的利用中，这种格式的优点在于它将亮度  $I$  与反映色彩本质特性的两个参数——色度  $H$  和饱和度  $S$  分开。当提取某类物体在色彩方面的特性时，经常需要了解其在某一色彩空间的聚类特性，而这一聚类特性往往体现在色彩的本质，而又经常受到光照明暗等条件的干扰影响。光照明暗给物体颜色带来的直接影响就是亮度分量  $I$ ，所以若能将亮度分量从色彩中提取出去，而只用反映色彩本质特性的色度、饱和度来进行聚类分析，会获得比较好的效果。

其变换公式为：

$$\begin{cases} I = \max(R, G, B) \\ S = \frac{\max(r, g, b) - \min(r, g, b)}{V} \\ H = h \times \frac{\pi}{6} \end{cases} \quad (3-2)$$

其中：

$$h = \begin{cases} 5 + \frac{I-b}{SI}, & \text{若 } r = \max(r, g, b) \text{ 和 } g = \min(r, g, b) \\ 1 - \frac{I-g}{SI}, & \text{若 } r = \max(r, g, b) \text{ 和 } g \neq \min(r, g, b) \\ 1 + \frac{I-r}{SI}, & \text{若 } g = \max(r, g, b) \text{ 和 } b = \min(r, g, b) \\ 3 - \frac{I-b}{SI}, & \text{若 } g = \max(r, g, b) \text{ 和 } b \neq \min(r, g, b) \\ 3 + \frac{I-g}{SI}, & \text{若 } b = \max(r, g, b) \text{ 和 } g = \min(r, g, b) \\ 5 - \frac{I-r}{SI}, & \text{其它} \end{cases} \quad (3-3)$$

由图 3-2 可以看到，原色是按 120 度分隔的，二次色与原色相隔 60 度，这意味着二次色之间也相隔 120 度，相同水平面上的任意一个彩色点，该点的色调由来自某参考点的一个角度来决定。通常与红轴的 0 度角指定为零色调，从这里开始色调逆时针增长。饱和度是从原点到该点的向量长度。<sup>[79]</sup>

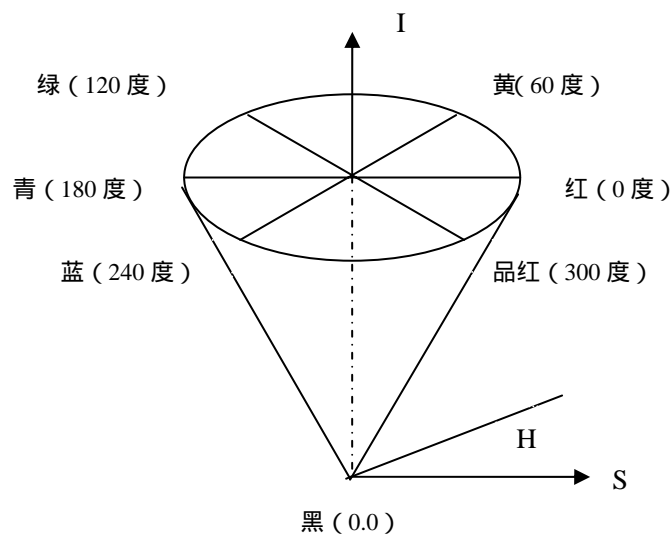


图 3-2 HIS 空间模型

### 3.2.3 YCbCr 空间模型

YCrCb 彩色模型是一种彩色传输模型，主要用于彩色电视信号传输标准方面，广泛地应用在电视色彩显示等领域中。YCrCb 彩色空间模型具有如下优点：

- (1) YCrCb 色彩空间具有与人类视觉感知过程相类似的构成原理。

- (2) YCrCb 色彩空间被广泛地应用在电视显示等领域中，也就是许多视频压缩编码，如 MPEG、JPEG 等标准中普遍采用的颜色表示空间。
- (3) YCrCb 色彩空间具有将色彩中亮度分量分离出来的优点。
- (4) 相比 HIS 等其他一些色彩空间，YCrCb 色彩空间的计算过程和空间坐标表示形式比较简单。
- (5) 实验表明在 YCrCb 色彩空间中的肤色聚类特性比较好。

RGB 空间到 YCbCr 空间的转换关系如公式 3-4 所示。<sup>[80]</sup>

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \frac{1}{256} \begin{bmatrix} 65.738 & 129.057 & 25.064 \\ -37.945 & -74.494 & 112.439 \\ 112.439 & -94.154 & -18.285 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3-4)$$

其中， $Y$  是色彩的亮度， $C_b$  和  $C_r$  分别表示蓝色和红色的色度。

### 3.2.4 TSL 空间模型

从 RGB 原始空间映射到归一化的 TSL 色彩空间是一种非线性变换，变换式为：

$$T = \begin{cases} \arctan(r' / g') / 2\pi + 0.25, g' > 0 \\ \arctan(r' / g') / 2\pi + 0.75, g' < 0 \\ 0, g' = 0 \end{cases} \quad (3-11)$$

$$S = \left[ 1.8 \times (r'^2 + g'^2) \right]^{0.5}$$

$$L = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

其中， $T$  是色调， $S$  是饱和度， $L$  是亮度，且  $T, S \in [0.0, 1.0]$ ，当  $R, G, B$  归一化时， $L \in [0.0, 1.0]$ ； $r' = r - 1/3$ ， $g' = g - 1/3$ ， $r = R / (R + G + B)$ ， $g = G / (R + G + B)$ 。

## 3.3 颜色空间的选择和区域分割

区域分割是图像分析和理解中的一项基本内容。定义区域为一个彼此相互连通的，具有一致有意义属性的像素集合。区域是一种方便的，很好的图像中层符号描述，它是目标或景物模型化以及高层理解的基础。区域分割的过程可以简单地概括为按照选定的一致性属性准则，将图像划分为互不交叠的区域集的过程。

从目的上看，区域分割是当前计算机视觉和图像理解中最困难的问题之一，虽然方法很多，但还没有尽善尽美的解决方案。对于彩色图像的肤色分割，YCbCr空间和TSL空间的肤色模型各有优缺点。TSL空间肤色模型的分割效果好，但是不能分割黄色背景。YCbCr空间肤色模型能分割黄色背景，而且不易漏检肤色，但是肤色范围较大。

我们要充分利用各种模型的优点，避免各自缺点，从而提高肤色分割效果。本文选择了在不同空间中，以肤色为一致性属性的区域分割方法，提高了基于肤色的区域分割算法的有效性。具体步骤如下：

- (1) 读入图像，将其对应在不同的空间模型中，本文选用了YCbCr空间模型和TSL空间模型。
- (2) 在YCbCr空间模型和TSL空间模型中对图像进行图像分割，用本章后述方法得到两个模型中的肤色区域  $Face_{YCbCr}$  和  $Face_{TSL}$ 。二值化图像，在图像中用1（白色区域）代表肤色匹配区域，0（黑色区域）代表非肤色的背景区域。
- (3) 取两个空间中得到的肤色区域的交集  $Face = Face_{YCbCr} \cap Face_{TSL}$ 。

在图像中表示为两幅图片中各相素相与。

在 YCrCb 空间内处理时，空间转换之后需要肤色模型进行肤色检测，根据公式将像素值导入到 YCbCr 空间。由于 YCbCr 色彩直接由 RGB 色彩通过线性变换得到，所以其亮度分量  $Y$  并不是完全独立于色度信息而存在，所以肤色的聚类区域也是随  $Y$  的不同而呈非线性变化的趋势。

在  $Y$  值不同的地方我们取  $C_b - C_r$  在平面的投影，得到的结果是不同的。所以我们必须考虑  $Y$  值不同造成的影响，根据呈现的变化规律，对 YCbCr 色彩空间进行非线性分段色彩变换。

图 3-3 则反映了肤色聚类边界上， $C_b$ 、 $C_r$  两分量随  $Y$  变化的情况。



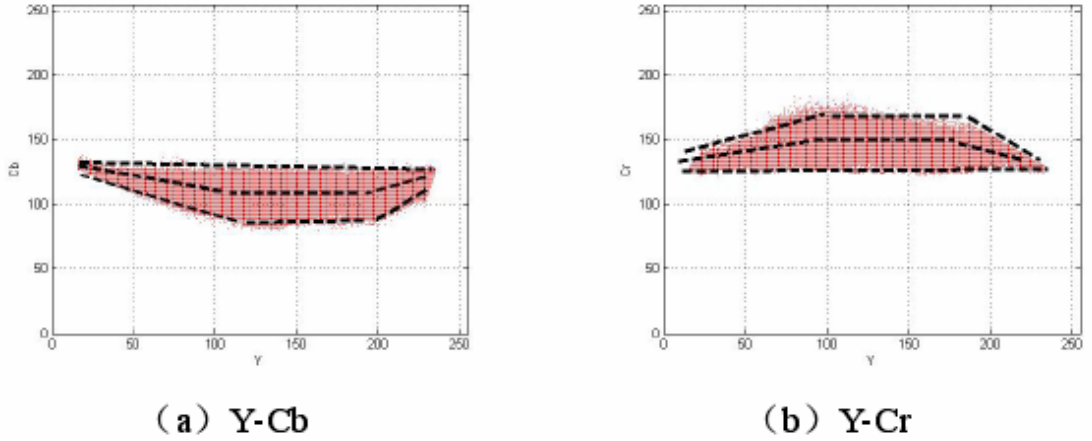


图 3-3 Cb、Cr 随 Y 的变化

用 4 个边界来限制肤色聚类区域。经过非线性分段色彩变换得到的色彩空间用  $YC_b'Cr'$  表示。 $YC_bCr$  色彩空间到  $YC_b'Cr'$  色彩空间的变换过程如下：

$$MC_b(Y) = \begin{cases} 108 + \frac{(K_i - Y) \times (118 - 108)}{K_i - Y_{\min}} \text{ if } (Y < K_i) \\ 108 + \frac{(Y - K_h) \times (118 - 108)}{Y_{\max} - K_h} \text{ if } (K_h < Y) \end{cases} \quad (3-5)$$

$$MC_r(Y) = \begin{cases} 154 + \frac{(K_i - Y) \times (154 - 114)}{K_i - Y_{\min}} \text{ if } (Y < K_i) \\ 154 + \frac{(Y - K_h) \times (154 - 114)}{Y_{\max} - K_h} \text{ if } (K_h < Y) \end{cases} \quad (3-6)$$

$C_b(Y)$  和  $C_r(Y)$  分别为  $Y-C_b$  和  $Y-C_r$  坐标系中肤色区域的中轴线，其中  $K_i = 125$ ， $K_h = 188$ ，它们是肤色聚类区域中  $Y$  分量的最小和最大值，这些数值都是根据实验得到。

$$WC_i(Y) = \begin{cases} WLC_i + \frac{(Y - Y_{\min}) \times (WC_i - WLC_i)}{K_i - Y_{\min}} \text{ if } (Y < K_i) \\ WHC_i + \frac{(Y_{\max} - Y) \times (WC_i - WHC_i)}{Y_{\max} - K_h} \text{ if } (K_h < Y) \end{cases} \quad (3-7)$$

$WC_i(Y)$  表示  $Y-C_i$  坐标系中肤色区域的宽度，其中， $i$  代表  $b$  或  $r$ ； $K_i$ ， $K_h$ ， $Y_{\min}$ ， $Y_{\max}$  同公式 (3-2)； $WC_b = 46.97$ ， $WC_r = 38.76$ ， $WLC_b = 23$ ， $WLC_r = 20$ ， $WHC_b = 14$ ， $WHC_r = 10$ 。

$$C'_i(Y) = \begin{cases} (C_i(Y) - MC_i(Y)) \times \frac{WC_i}{WC_i(Y)} + MC_i(Y) & \text{if } (Y < K_i) \text{ or } (K_h < Y) \\ C_i(Y) & \text{if } (Y \in [K_i, K_h]) \end{cases} \quad (3-8)$$

$C'_i(Y)$  表示  $C_i(Y)$  在非线性分段色彩变换后新坐标系中的值，其中， $i$  代表  $b$  或  $r$ ，其他参数同公式 (3-5) (3-6)。

经过这样的非线性分段色彩变换，YCb'Cr' 色彩空间建立。如图 3-4 所示，用一个椭圆近似这一肤色区域。

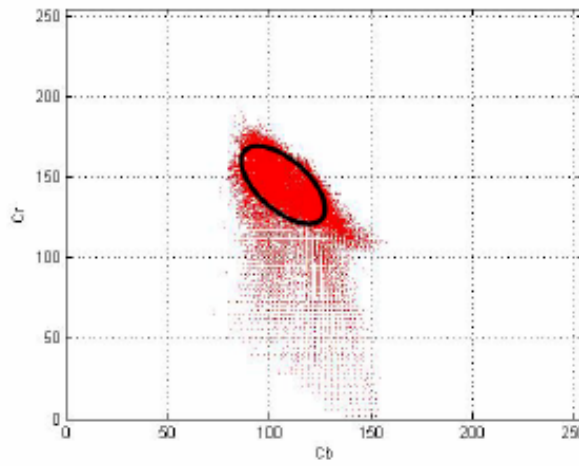


图 3-4 Cb-Cr 平面的肤色聚类模型

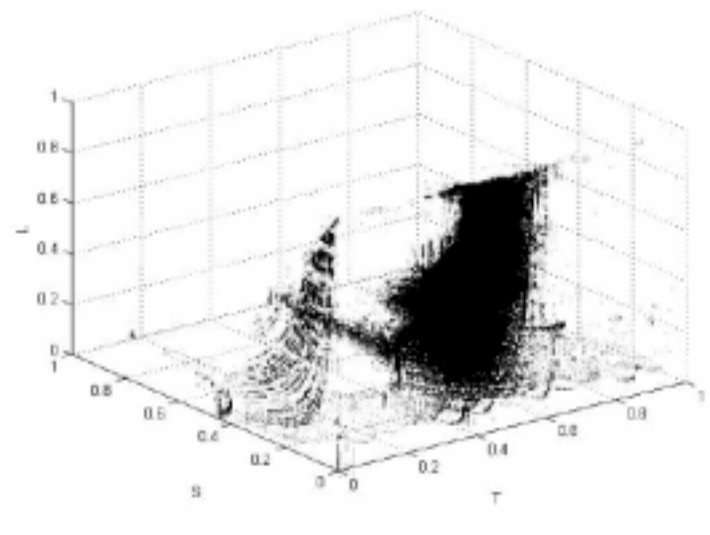
它的解析表达式为：

$$\frac{(x - Ec_x)^2}{a^2} + \frac{(y - Ec_y)^2}{b^2} \leq r^2, \quad (3-9)$$

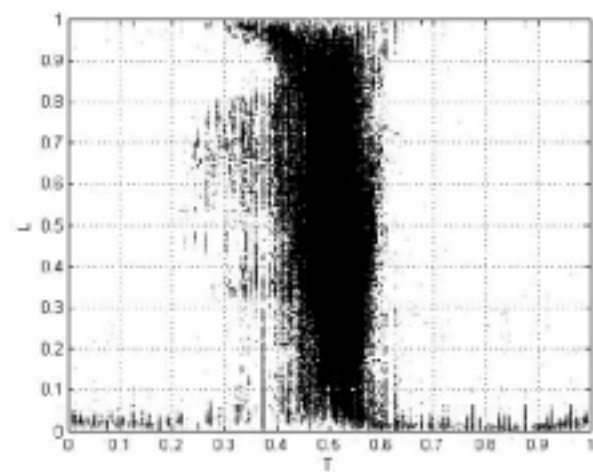
$$\text{其中} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \times \begin{pmatrix} C'_b - c_x \\ C'_r - c_y \end{pmatrix} \quad (3-10)$$

$c_x = 109.38$ ， $c_y = 152.02$ ， $\theta = 2.53$ ， $Ec_x = 1.60$ ， $Ec_y = 2.41$ ， $a=25.39$ ， $b=14.03$ ， $r=1$ 。这些数据由实验得到。

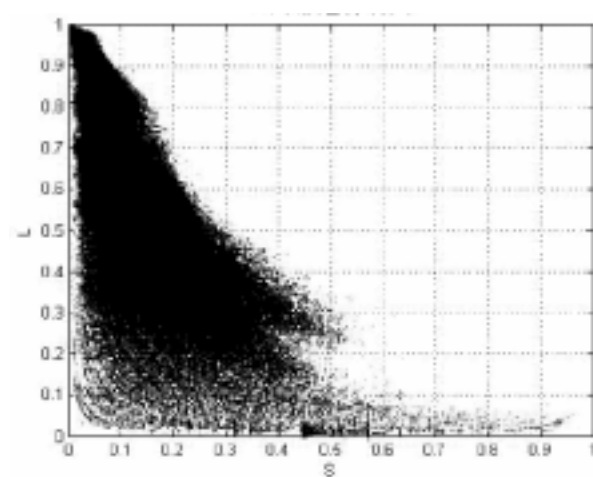
在 TSL 空间上处理时，如图 3-5，这是从挑选的 50 张人脸图像中提取的 80 万左右的人脸肤色像素，通过上述公式的转换在 TSL 空间中的分布情况。图中可以看出色调  $T$  和饱和度  $S$  坐标平面上的肤色像素分布比较集中，而且受亮度  $L$  变化的影响比较小，但仍不可避免受到亮度的影响。



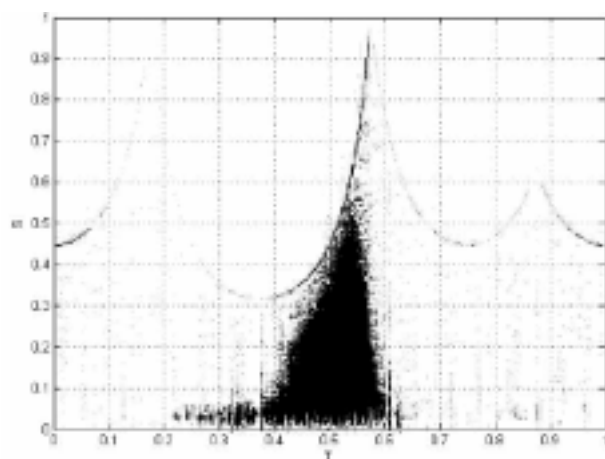
(a)TSL 空间肤色统计分布



(b)T-L 空间肤色统计分布



(c)S-L 空间肤色统计分布

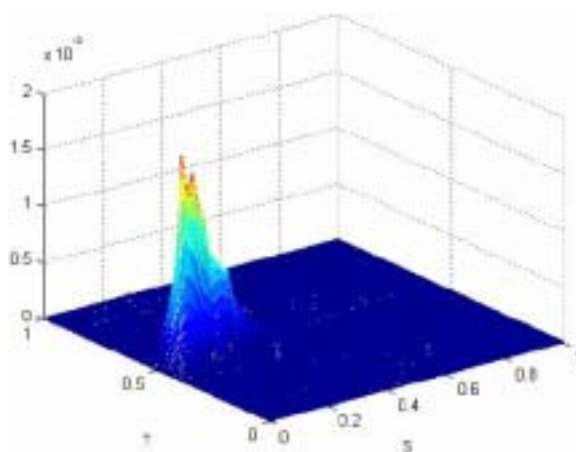


(d)T-S 空间肤色统计分布

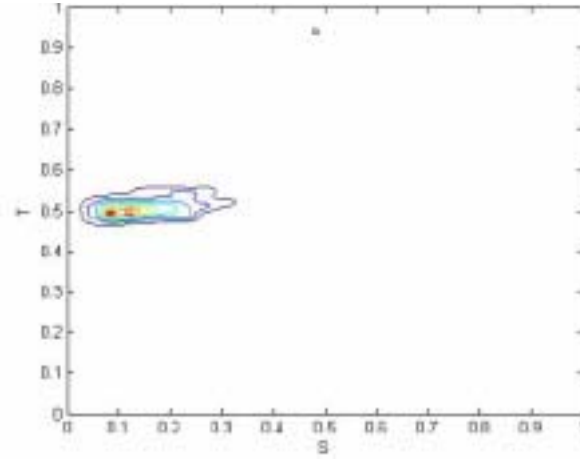
图 3-5 TSL 空间肤色统计分布

高斯模型是利用正态分布来拟合皮肤颜色的概率密度分布<sup>[65]</sup>。它的理论基础就是：

它认为人与人之间肤色的不同主要在于密度而不是颜色本身,它认为在一定的光照条件下肤色的分布是正态分布的。虽然不同人的皮肤颜色可能相差很大,但它们在色度上的差异远小于亮度上的差异,也就是说,不同人的肤色在色度上往往是很相近的,只是在亮度上有较大差异。如图 3-6,从  $T$ - $S$  平面上的概率统计反映人脸肤色的分布情况,一定程度上证实了高斯模型理论的合理性。



(a)肤色统计概率分布图



(b) 肤色统计分布等高线图

图 3-6 T-S 平面肤色统计概率分布

根据黄种人的肤色特征，采用椭圆高斯模型（pdf）表示在  $T$ - $S$  坐标系上肤色分布概率的大小，即：

$$p[x_{i,j}/W_s] = (2\pi)^{-1} |C_s|^{-0.5} \exp[-\lambda_{i,j}^2 / 2] \quad (3-12)$$

其中，向量  $x_{i,j} = [T_{i,j}, S_{i,j}]^T$ ，表示在图像中坐标为  $(i,j)$  的像素的色度的随机测量值， $W_s$  是描述皮肤的类， $C_s$  是肤色集的协方差矩阵，即：

$$C_s = \begin{pmatrix} \sigma^2 T & \sigma TS \\ \sigma TS & \sigma^2 S \end{pmatrix} \quad (3-13)$$

其中  $\lambda_{i,j}$  是从向量  $x_{i,j}$  到肤色平均向量  $v_m = [M_t, M_s]^T$  的 Mahalanobis 距离，即：

$$\lambda_{i,j}^2 = [x_{i,j} - v_m]^T C_s^{-1} [x_{i,j} - v_m] \quad (3-14)$$

由公式 (2-11) 可以看出，像素  $(i,j)$   $\lambda_{i,j}$  的值决定了  $p[x_{i,j}/W_s]$  值的大小，也就是， $\lambda_{i,j}$  越大，像素  $(i,j)$  属于肤色的可能性越低。

所以，对于肤色的判别， $\lambda_{i,j}$  需要确定一个阈值  $\lambda_{t,s}$ ，若  $\lambda_{i,j} < \lambda_{t,s}$ ，则像素  $(i,j)$  被认为属于肤色；反之，则不属于肤色。阈值大小可根据实际需要具体确定。

作为对辨别“皮肤类”和“非皮肤类”合适的折中，一个标准阈值  $\lambda_{t,s}$  可以当整个含皮肤区域的正识别  $TP$  变得与整个不含皮肤区域的负识别  $TN$  比例相等时而得到，或者相当于负误识  $FN$  的比例等于正误识  $FP$ （因为  $TP + FN = 1$  和

$TN + FP = 1$  )

为了减小搜索范围，提高人脸检测算法的效率，首先对图像进行分割，将肤色区域与非肤色区域分开，其过程如下：

(1)利用各像素点距肤色中心马式距离(Mahalanobis Distance)的幂函数  $\lambda_{i,j}$  ,  
 $\lambda_{i,j}^2 = [x_{i,j} - v_m]^T C_s^{-1} [x_{i,j} - v_m]$  ;  $\lambda_{i,j} \in [0, 1]$  , 作为相似度来计算输入图像相对于肤色中心的相似度图。

(2)对相似度图进行  $3 \times 3$  的均值滤波，消除噪声。

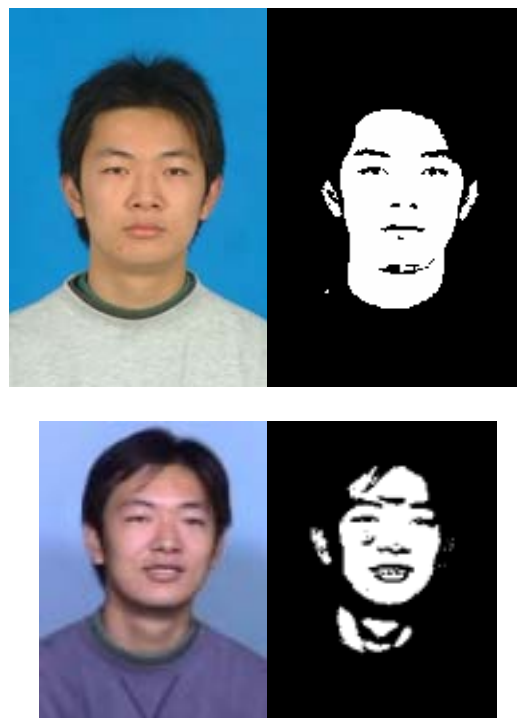
(3)与预先设置的阈值0.5作比较，二值化相似度图。

(4)利用并行的区域分割技术——连通算法来标记肤色区域。

(5)对标记的肤色区域利用先验知识进行初排除，考虑到后面匹配时所用模板的大小，删除面积小于200像素的标记区域；其次考虑大多数人脸的形状，删除高宽比小于0.8、大于2.4的区域。

算法中的各项阈值，如滤波器的大小、幂函数的系数等都是在实验中得出的，同时最终的结果也表明，使用这些系数能够取得较好的检测效果。

在Cpu为P4 2.0GHz，内存为1G的机器上，VC6.0的环境下运行，实验结果如下：



(a) 简单背景下人脸检测肤色区域划分



(b) 复杂背景下人脸检测肤色区域划分



(c) 复杂肤色情况人脸检测肤色区域划分



(d) 复杂背景下肤色区域划分

图 3-7 实验结果图

### 3.4 本章小结

从皮肤概率图像中可以看到皮肤区域比非皮肤区域的亮度高,因此皮肤区域可以通过阈值处理与非皮肤区域分开。对于不同肤色、不同人和不同亮度的各种图像,固定的阈值无法对所有的情况下的图像进行准确的分割,因此需要用模糊

分类的方法对皮肤和非皮肤进行区分。

本文采用了基于混合高斯模型对肤色色度信息进行描述,利用各像素点距肤色中心马式距离来计算与肤色相似度,对很多图像处理都得到了比较好的结果,是皮肤与非皮肤区域有效的分开。



## 4 人脸检测算法中的光照补偿策略

### 4.1 引言

与人脸检测的研究不同,人脸光照补偿的研究较少。一般的人脸检测和识别算法通常假设待处理图像是在均匀的光照条件下获得的,而实际上光照往往是不均匀的。偏色光、侧光、高光和过亮、过暗、阴影都会使人脸检测和识别率大幅度下降, FERET 测试和 FRVT 测试都表明光照变化仍是实用人脸识别系统的瓶颈之一,所以我们有必要对光照进行补偿<sup>[66][67]</sup>。

光照补偿问题一直是人脸检测和识别中的一个非常重要而难于解决的问题。Mose<sup>[68]</sup>等从理论上证明,对于光照保持不变性的函数是不存在的。Warrington 和 Taylor 的研究表明<sup>[69]</sup>,光照变化的补偿不仅需要人脑的低级视觉皮层,还需要高级视觉区域的配合。一般的,解决光照补偿问题的方法主要包括<sup>[70]</sup>:提取光照不变特征<sup>[71][72]</sup>、光照变化的建模<sup>[69][73][74][75]</sup>、光照条件标准化<sup>[76][77][78]</sup>。

第一种方法是提取人脸不随光照变化或至少对光照变化不敏感的特征。但 Mose 和 Adini ' s 等从理论上及实验结果表明<sup>[67][68]</sup>“没有任何表示本身能克服光照方向的变化”。这就告诉我们提取人脸完全不随光照变化的特征是不可能的!我们只能提取对光照变化不敏感的特征。在文献<sup>[28]</sup>中提取图像的梯度方向特征作为光照不变特征。在文献<sup>[29]</sup>中构建图像特征作为光照不变特征。尽管这些方法的思想很直接,但实践表明这种方法对光照补偿都有一定的局限性。

第二种方法是在应用光照变化的建模,主要思想是在一个合适的子空间中表示光照引起的变化。识别的过程是对待识的人脸的特征估计模型的参数,之后选择与待识人脸最相近的子空间。

第三种方法是在人脸图像被识别之前消除光照变化的影响,最后识别在标准形式上进行。一个典型的例子是 SFS ( shape-from-shading ) 方法。在文献<sup>[35]</sup>中环境的光照方向和人脸的反射率被估计出来,然后重构出人脸的形状。SFS 是一个思想很好的算法,但也有其本质的缺点:一是模型经过高度简化太过简单。二是它只能适用于单一点光源,这在应用中是很难保证的,三是非人脸图像经过其补偿后,补偿结果变成人脸图像,这导致在人脸检测中不能使用这种方法,

四是具有很高的计算复杂性。

本章中的光照补偿算法原则上属于第一种方法。通过大量的研究和比较,本文提出了一种针对不同光照情况的人脸补偿算法。实验结果表明,本算法基本上克服了复杂光照环境下无法检测的缺陷,很大程度上改善了光照补偿的效果。使不同光照条件下的人脸检测和人脸识别性能得到了很大的提高。

本文认为光照环境对人脸检测的影响主要分为:

- (1) 光照不足或过亮。
- (2) 有色光源照射。
- (3) 由于遮挡或者光线照射角度带来的“阴影”。

## 4.2 光照不足或过亮的补偿策略

针对实际应用中某些情况下图像光照条件比较恶劣,在这种情况下图像在灰度上会偏暗或偏亮,为了克服这种光照条件对图像的影响,需要对采样图像进行预处理。

其中,最直观的方法就是对灰度图像进行直方图均衡化。

由于采样图像在实际应用中,一般均过明或过暗,实际应用中,在 YCrCb 空间中先将图片进行直方图统计,将图片中最亮灰度与最暗灰度找出来,定义为“参考白”或“参考黑”,用以提高或降低整幅图像的灰度水平。

实验结果如图 4-1 所示。



图 4-1 光照不足补偿结果

### 4.3 有色光源的补偿策略

实际中,有这样的光照条件存在,图像是在有色光源下采样或图像采集设备的色差因素影响,使图片整体上偏离本质色彩而向某一方向移动,即我们常说的色彩偏冷、偏暖、偏黄、偏蓝、偏红等。这类情况比较普遍,如艺术照中图像多为有色光源光照条件下采样图像。这种光照条件会直接影响到人脸的检测效果,本文提出了对该光照条件的光照补偿策略。

在彩色空间中,色调、饱和度、亮度是 3 个基本的属性。其中色调是颜色中最重要的属性,它决定了颜色的本质特性和基本特征。在物体表面材料不变的情况下,投影光照强度的变化对物体表面色彩的饱和度和亮度有比较大的影响,但是对色调的影响较小,因而采用在以色调、饱和度、亮度为特征的 HSI 颜色空间下,进行色彩补偿。

实验原本采用对图像统计出来色调、饱和度、亮度平均值作为色彩偏移量的校正依据,但是实验结果很不理想。分析主要是由于图像中背景,局部阴影等对检测造成不利影响因素在图像中所占面积很大,造成对图像统计出来色调、饱和度、亮度平均值不足以代原图像中信息。由于根据统计图像得到的色调、饱和度、亮度平均值,往往不能很好的成为色彩偏移量的依据,所以本文引入“参考点”这个概念,用以作为校正依据。计算这些“参考点”的色调,将其与定义为参考标准色调相比较,修改整幅图像各个像素点的色调值,将其对应到 RGB 空间中,再用白平衡的方法适当修正色彩饱和度。

作为图像色彩偏移的校正依据,“参考点”提取的策略决定了图像的光照偏移修正状态,对后续的色彩平衡十分关键。

经过对大量图像的分析,我们发现,面积(占整幅图像的比例)较大的高亮区域很可能代表图像整体色调状况。据此,“参考点”提取的具体步骤如下:

(1)首先将图像按最高灰度值的 90%进行分割,初步得到若干个高亮区域,并将两个面积较大且含高亮点最多的高亮区域提取出来。

(2)将两个区域进行比较,并确定一个更好的反映色彩状况的高亮区域;然后提取其中大于该区域最高亮度值 98%的数量不多于 500 个的点作为估计光照的“参考点”。

(3)若面积较大的高亮块不存在或高亮块都不包含参考点,则提取图像中等

于最高亮度值 80% 的全部点作为估计光照的“参考点”。

在肤色分割步骤中，假定图像上有若干点  $\{x_1, x_2, x_3, x_4, \dots, x_n\}$ ，其每个点的 RGB 值经过转换后所对应的色调为  $\{H_1, H_2, H_3, H_4, \dots, H_n\}$ ，则当某标准点  $s$  的色调为  $H_s$  时，就可以将满足下式的各点视为颜色相同。

$$\{x_i \mid |H_i - H_s| \leq \alpha, i = 0, \dots, n\} \quad (4-6)$$

其中  $\alpha$  为色调阈值。

针对有色光源的补偿实验结果如下：



(a) 原图像



(b) 有色光源补偿后图像



(c) 光线补偿后肤色区域划分结果

图 4-2 有色光源补偿后实验结果

实验表明该算法在针对图片背景简单情况下，可以达到比较好的效果，但是在图片背景复杂时候，结果很差。

## 4.4 局部补偿

在检测过程中，经常会遇到由于遮挡，光照角度造成的阴影等因素的不利影响，使得阴影部分成为检测的盲区，为了克服图像中阴影部分给人脸检测带来的不利影响，本文提出了针对阴影的局部补偿算法。

在照相技术上，有一种白平衡技术。白平衡技术的核心，是通过调整由传感器得到的色彩视觉信号分量的增益，来让拍摄出来的图像颜色更接近真实的色彩。通常，白平衡方法是在色温环境下拍摄一组纯白色的物体，分析所拍摄的图像数据，对白色部分进行平均，求其 RGB 三个分量的均值，重新定义 RGB。白平衡算法的目的在于在图像进行亮度补偿后，对色彩进行简单的修正，让图像还原更为“真实”。

目前，基于白平衡技术的图像色彩均衡方法主要以 white patch 和 gray world 算法为代表。

Gray world 色彩平衡算法基于“灰度世界”假设，该假设认为对于一幅有着大量色彩变化的图像， $R$ 、 $G$ 、 $B$  三个分量各自的平均值趋于同一个灰度值。在客观世界中，一般来说物体及其周围环境的色彩变化是随机的，独立无关的，因此这一假设是合理的。该方法基本思想是首先通过图像  $R$ 、 $G$ 、 $B$  三个分量各自的平均值，确定出图像的平均灰度值，然后调整每个像素的值，使得调整后图像

的  $R$ 、 $G$ 、 $B$  三个分量各自的平均值都趋于平均灰度值。但是，对于违反“灰度世界”假设的情况，如原图像在正常光照下就偏蓝色、偏绿色或偏红色即图像色彩变化不够丰富多样，该算法则失效<sup>[80]</sup>。

White patch 色彩平衡算法认为，绝大部分的人脸图像中都包含有纯白色，如眼球外围等处，所以最大亮度的像素点的本色应该为纯白色，从而所有这样的点综合反映了图像整体光照信息。该算法基本思想是将整个图像中所有像素的亮度（是经过非线性校正后的亮度）从高到低进行排列，取前 5% 的像素，如果这些像素的数目足够多，就将他们的亮度作为“参考白”，也即将他们的色彩的  $R$ 、 $G$ 、 $B$  分量值都调整为最大的 255。整幅图像的其他像素点的色彩值也都按这一调整尺度进行变换。但是，对于亮度不高的图像，若调整最亮点为纯白亮点，则整幅图像的调整比例太大容易导致调整误差较大，该算法失效<sup>[81]</sup>。

本文认为现有的局部补偿算法并不科学，因为它们都是要先对图像中的某些信息进行统计，然后把这些统计结果作为补偿的重要依据，然而在大部分情况下统计出来的结果并不能很好的代表阴影或整幅图像的信息。而且在实际应用中，现有算法实际效果不好，所以本文提出了自己的局部补偿算法。本文局部补偿算法的关键在于准确找到需要补偿的图像块，即对阴影边缘的准确提取。局部补偿主要包括下列几个步骤：

#### （1）肤色检测

对原图像按照第三章所述方法进行肤色检测，分析在图像中划分肤色区域与非肤色区域。

#### （2）图像边缘检测

边缘是图像最基本的特征，图像的大部分信息存在于图像边缘中，主要表现为图像局部的不连续性，即灰度变化比较剧烈的地方。因此，阴影部分边缘可划分为类阶跃型和类屋顶型两种。实验采取基于梯度的边缘检测方案，由于梯度对应的为一阶导数，梯度算子是一阶导数算子，本文采用 Sobel 算子，利用像素的上、下、左、右邻域的灰度加权，根据在边缘点处达到极值这一原理进行边缘检测。根据边缘检测结果，将图像分为若干块，再根据图像块的形状，大致确定该图像块的中心位置。

### (3) 图像补偿区域的划定

根据阴影部分必须在由步骤(1)确定的肤色区域附近的原则。计算由步骤(2)中找到的图像块的中心位置到由步骤(1)确定的肤色区域距离  $d_{i,j}$  , 同时计算图片的最大距离  $d_M$  ,  $d_M$  为图像的对角线距离。

当  $d_{i,j} \leq 0.05d_M$  时, 即认为该由步骤(2)找到的图像块为所找阴影部分。

### (4) 图像亮度补偿

为了能够对阴影能够准确的进行光线补偿, 实验引入二阶微分算子, 用于参考制定补偿程度, 并结合由步骤(1)确定肤色区域的灰度以及色度信息确定补偿因子(阴影部分最暗处灰度值与肤色区域灰度值之差), 最后根据补偿因子对图像中阴影部分进行线形的亮度增加。补偿后再利用 White patch 色彩平衡算法, 对相关像素点进行色彩补偿。

实验结果如下：



(a) 原图像



(b) 补偿前肤色匹配结果



(c) 查找到的边缘



(d) 补偿后图像





(e) 补偿后匹配结果

图 4-3 局部补偿实验结果

实验结果分析如下：

实际测试时步骤（2）很难从实际图像中提取阴影的准确轮廓。由于本文用于实验的图片素材均来自互联网，并非人脸检测专用图片，很多图片没有专门的对图像去噪，而二阶导数算子对噪声有无法接受的敏感性，所以在边缘检测时候，采用的是对噪声影响相对较小的 Sobel 算子。但是，在实际应用中该方法会检测伪边缘，定位的精度不是很高。

步骤（3）中难点为确定阴影块中心位置，本文采用求取块内各点像素的水平和垂直轴上各点平均值的方式，这种方式在阴影形状较为规则时候比较有效，而在阴影块形状不好时候，定位不准确，甚至定位到阴影块外，造成定位错误。而且，在计算阴影块中心点到没有进行补偿前原图像分析确定的肤色区域距离时候，由于采用归一法，对区域内所有点求值，最后将比较找到的最小值定义为该距离，这样算法的计算量较大，可以适当优化。其中参数 0.05 为实验值，可以根据实际情况变化。

步骤（4）补偿策略的制定中，实现困难也较大，由于阴影灰度多为“屋顶型”，所以，引入二阶微分算子来辅助寻找灰度极值。但是，由于图像的光照环境不同，该极值点代表的可能是阴影区域的最暗点，也有可能是最亮点，造成光线补偿效果不是很理想，补偿时由于对阴影部分亮度增加是根据补偿因子的线形变化，会带来边缘图像突然变亮而使图像斑点化的可能。

## 4.5 本章小结

本章内容主要针对现有算法在不同光照环境下，稳健性较差的缺点，提出了对图像进行光线补偿的预处理的思路。在对人脸检测算法产生不利影响的恶劣光照环境进行分析后，本文将恶劣光照环境分为光照不足或过亮、有色光源照射、偏光造成的阴影三类。对采样的图像，分类进行光线补偿，使恶劣光照环境对图像的影响降低，从而提高算法的肤色匹配能力。

其中，针对光照不足或过亮的补偿策略已经非常成熟，但是本文作者发现针对有色光照补偿策略的研究点的研究成果，目前还没有看到很多，针对阴影的局部补偿是光线补偿研究的难点，现有算法的实际效果并不理想，本文的新算法是在分析了原有算法的优缺点的基础上提出，取得了一些成绩，但是效果还是不够理想。针对实验效果中出现的问题，本文提出了一些思路，以供今后研究完善。

## 5 总结与展望

人脸检测技术有着广阔的应用前景,这一问题的进展会给人脸识别、人脸跟踪、表情识别、姿态分析、视频监控、身份验证、多媒体检索等领域带来巨大的推动力,也会在方法论上丰富计算机视觉、模式识别等计算机科学领域的发展。

但是,由于人脸模式的复杂性和图像采集过程中不可避免的引入的各种不利因素,都使得人脸检测问题成为一个极具挑战性的研究课题。

为了在彩色图像中人脸检测领域达到更好的检测性能,本文在进行了大量的文献阅读、算法学习和实验的基础上,综合运用肤色检测、肤色分割技术,对它们存在的问题进行了较为深入的剖析,提出基于混合高斯模型对肤色色度信息进行描述,利用各像素点距肤色中心马式距离来计算与肤色相似度的算法,对很多人脸图像都得到了比较好的结果,使得皮肤与非皮肤区域有效的分开。并针对现有光照补偿策略的不足,提出针对不同光照情况下不同的补偿策略,较好的解决了因为光照原因对检测率的影响。特别是在有色光源照射和由于照射阴影造成现有算法检测率下降情况下,提出新的补偿思路。实验中,在针对有色光照射情况下,能够将图像基本“还原”,能找到更多的肤色区域;在针对阴影的局部补偿算法上,克服了现有算法的缺点,提出从根本上解决阴影的新思路,能够较好地解决由于偏光照射带来的阴影问题。

实验表明本文提出算法在一定程度上可以对人脸图像进行粗略检测,光线补偿策略也能在一定的环境下起到很好的作用。在进行肤色分割时候,如果需要对人脸区域进行准确的检测,有效地避免肤色区域的欠分割和过分割,算法还需要完善。本文算法现在仅做到将肤色与肤色区域大致分开程度,在对人脸准确定位方面本文算法还需要完善。本文提出的光线补偿算法部分能够针对一些特殊环境下情况进行处理,但是,其并非完全“自适应”的,在后续研究中应该考虑到这种需求。本文主要从肤色特征角度入手进行人脸检测算法的研究,实际操作中可以和其它的结合彩色的人脸特征角度入手进行检测,在不会明显提高算法复杂度的情况下,选用何种其它人脸特征与肤色相配合也是将来研究的方向之一。本文算法针对简单背景的图像有较高的分割效率,但是对于背景过于复杂的环境下,

检测精度受到很大的影响,去除背景对图像的影响,也是本文将来研究的方向之一。

# 致谢

两年的硕士研究生生活即将结束了，此时我内心十分激动。师生情、同学意，真切动人。首先对我的导师刘文予教授道以衷心的感谢！刘老师严谨的治学态度、锲而不舍的钻研精神、宽广博大的胸怀、睿智儒雅的学者风范，将是我一生学习的榜样和孜孜以求的目标。

在两年的学习生活中，学业上刘老师严格要求谆谆教导，老师他那敏锐的观察力和对问题的独到见解令我深深的折服且使我受益匪浅。“教书育人”是导师的风格淋漓尽致的写照。

感谢冯镇老师的热情指导和悉心帮助，感谢唐振华、姚志均博士生，肖非、唐奕斐、朱志怡硕士生等各位师兄弟无私的帮助和坦诚的学术交流！

特别感谢养育我多年的父母，正是他们的含辛茹苦、辛勤培养、无私的爱给予我无尽的战胜困难的勇气和力量，才能使我有了今天的一点成绩。

感谢王轶、李然、唐正等各位朋友，我们的意气相投让我两年的生活很是开心和快乐。

深深地感谢你们---我的老师、亲人、同学和所有的朋友们，是你们真挚的爱、不断勉励和支持使我能圆满完成学业即将顺利毕业！我会加倍努力，百尺竿头更进一步。不辜负老师的培养和亲人的厚望。再一次深深地感谢所有关心爱护我的人！

最后也由衷地感谢评阅论文的各位专家，感谢你们付出的劳动。

## 参考文献：

- [1] 夏思宇. 彩色图像序列的人脸检测、跟踪与识别研究. 东南大学博士学位论文,2006
- [2] 何坤. 人脸识别理论关键技术的研究. 四川大学博士学位论文,2006
- [3] 樊昀. 基于内容图像检索的一些关键技术研究. 国防科学技术大学博士学位论文,2001
- [4] G. Yang and T. S. Huang, Human Face Detection in Complex Background, Pattern Recognition, 1994, 27(1):53-63.
- [5] 苏宏涛. 基于统计特征的人脸识别技术研究. 西北工业大学博士学位论文,2003
- [6] Ming-Hsuan Yang ,Kriegman D J,Ahuja N. Detecting Faces in Images:A Survey. IEEE transactions on Pattern Analysis and Machine Intelligence,2002,24(1):34-58.
- [7] S.A. Sirohey, Human Face Segmentation and Identification, Technical Report CS-TR-3176, Univ. of Maryland, 1993.
- [8] O. Bernier, M. Collobert, R. Fe'raud, V. Lemarie, J.E. Viallet, and D. Collobert, MULTRAK: A System for Automatic Multiperson Localization and Tracking in Real-Time, Proc. IEEE Int'l Conf. Image Processing , 1998:136-140.
- [9] D. Chetverikov and A. Lerch, Multiresolution Face Detection, Theoretical Foundations of Computer Vision, 1993, 69:131-140.
- [10] H.P. Graf, T. Chen, E. Petajan, and E. Cosatto, Locating Faces and Facial Parts, Proc. First Int'l Workshop Automatic Face and Gesture Recognition, 1995:41-46.
- [11] T.K. Leung, M.C. Burl, and P. Perona, Finding Faces in Cluttered Scenes Using Random Labeled Graph Matching, Proc. Fifth IEEE Int'l Conf. Computer Vision,1995:637-644.
- [12] D.G. Kendall, Shape Manifolds, Procrustean Metrics, and Complex Projective Shapes, Bull. London Math. Soc, 1984 ,16: 81-121,.
- [13] K.V. Mardia and I.L. Dryden, Shape Distributions for Landmark Data, Advanced Applied Probability, 1989,21:742-755.

- [14] K.C. Yow and R. Cipolla, A Probabilistic Framework for Perceptual Grouping of Features for Human Face Detection, Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, 1996:16-21.
- [15] K.C. Yow and R. Cipolla, Feature-Based Human Face Detection, Image and Vision Computing, 1997,15(9):713-735.
- [16] K.C. Yow and R. Cipolla, Enhancing Human Face Detection Using Motion and Active Contours, Proc. Third Asian Conf. Computer Vision, 1998:515-522.
- [17] K.C. Yow and R. Cipolla, A Probabilistic Framework for Perceptual Grouping of Features for Human Face Detection, Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, 1996:16-21.
- [18] K.C. Yow and R. Cipolla, Feature-Based Human Face Detection, Image and Vision Computing, 1997,15(9):713-735.
- [19] K.C. Yow and R. Cipolla, Enhancing Human Face Detection Using Motion and Active Contours, Proc. Third Asian Conf. Computer Vision, 1998:515-522.
- [20] B. Takacs and H. Wechsler, Face Location Using a Dynamic Model of Retinal Feature Extraction, Proc. First Int'l Workshop Automatic Face and Gesture Recognition, 1995:243-247.
- [21] C.-C. Han, H.-Y.M. Liao, K.-C. Yu, and L.-H. Chen, Fast Face Detection via Morphology-Based Pre-Processing, Proc. Ninth Int'l Conf. Image Analysis and Processing, 1998:469-476.
- [22] Y. Amit, D. Geman, and B. Jedynak, Efficient Focusing and Face Detection, Face Recognition: From Theory to Applications, H. Wechsler, P.J. Phillips, V. Bruce, F. Fogelman-Soulie, and T.S. Huang, eds., 1998, 163:124-156,.
- [23] Y. Amit, D. Geman, and K. Wilder, Joint Induction of Shape Features and Tree Classifiers, IEEE Trans. Pattern Analysis and Machine Intelligence, Nov. 1997,19(11):1300-1305.
- [24] L. Breiman, J. Friedman, R. Olshen, and C. Stone, Classification and Regression Trees. Wadsworth, 1984.
- [25] M.F. Augusteijn and T.L. Skujca, Identification of Human Faces through Texture-Based Feature Recognition and Neural Network Technology, Proc. IEEE

Conf. Neural Networks, 1993:392-398.

[26] Y. Dai and Y. Nakano, Face-Texture Model Based on SGLD and Its Application in Face Detection in a Color Scene, *Pattern Recognition*, 1996, 29(6):1007-1017.

[27] H.P. Graf, T. Chen, E. Petajan, and E. Cosatto, Locating Faces and Facial Parts, *Proc. First Int'l Workshop Automatic Face and Gesture Recognition*, 1995:41-46.

[28] H.P. Graf, E. Cosatto, D. Gibbon, M. Kocheisen, and E. Petajan, Multimodal System for Locating Heads and Faces, *Proc. Second Int'l Conf. Automatic Face and Gesture Recognition*, 1996:88-93.

[29] D. Chai and K.N. Ngan, Locating Facial Region of a Head-and-Shoulders Color Image, *Proc. Third Int'l Conf. Automatic Face and Gesture Recognition*, 1998:124-129.

[30] S. McKenna, Y. Raja, and S. Gong, Tracking Colour Objects Using Adaptive Mixture Models, *Image and Vision Computing*, 1998, 17, nos. 3/4:223-229.

[31] K. Hotta, T. Kurita, and T. Mishima, Scale Invariant Face Detection Method Using Higher-Order Local Autocorrelation Features Extracted from Log-Polar Image, *Proc. Third Int'l Conf. Automatic Face and Gesture Recognition*, 1998:70-75.

[32] H. Wu, Q. Chen, and M. Yachida, Face Detection from Color Images Using a Fuzzy Pattern Matching Method, *IEEE Trans. Pattern Analysis and Machine Intelligence*, June 1999, 21(6):557-563.

[33] J.C. Terrillon, M. David, and S. Akamatsu, Automatic Detection of Human Faces in Natural Scene Images by Use of a Skin Color Model and Invariant Moments, *Proc. Third Int'l Conf. Automatic Face and Gesture Recognition*, 1998 :112-117.

[34] J.C. Terrillon, M. David, and S. Akamatsu, Detection of Human Faces in Complex Scene Images by Use of a Skin Color Model and Invariant Fourier-Mellin Moments, *Proc. Int'l Conf. Pattern Recognition*, 1998:1350-1355.

[35] E. Saber and A.M. Tekalp, "Frontal-View Face Detection and Facial Feature Extraction Using Color, Shape and Symmetry Based Cost Functions," *Pattern Recognition Letters*, 1998, 17(8):669-680.

[36] D. Huttenlocher, G. Klanderman, and W. Rucklidge, Comparing Images Using the Hausdorff Distance, *IEEE Trans. Pattern Analysis and Machine Intelligence*,



1993,15:850-863.

[37] Q.B. Sun, W.M. Huang, and J.K. Wu, Face Detection Based on Color and Local Symmetry Information, Proc. Third Int'l Conf. Automatic Face and Gesture Recognition, 1998:130-135.

[38] T. Sakai, M. Nagao, and S. Fujibayashi, Line Extraction and Pattern Detection in a Photograph, Pattern Recognition, 1969,1:233-248.

[39] I. Craw, H. Ellis, and J. Lishman, Automatic Extraction of Face Features, Pattern Recognition Letters, 1987,5:183-187.

[40] I. Craw, D. Tock, and A. Bennett, Finding Face Features, Proc. Second European Conf. Computer Vision, 1992:92-96.

[41] V. Govindaraju, Locating Human Faces in Photographs, Int'l J. Computer Vision, 1996,19(2):129-146.

[42] V. Govindaraju, D.B. Sher, R.K. Srihari, and S.N. Srihari, Locating Human Faces in Newspaper Photographs, Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1989:549-554.

[43] V. Govindaraju, S.N. Srihari, and D.B. Sher, A Computational Model for Face Location, Proc. Third IEEE Int'l Conf. Computer Vision, 1990:718-721.

[44] A. Tsukamoto, C.-W. Lee, and S. Tsuji, Detection and Tracking of Human Face with Synthesized Templates, Proc. First Asian Conf. Computer Vision, 1993:183-186.

[45] A. Dempster, A Generalization of Bayesian Theory, J. Royal Statistical Soc., 1978,30:205-247.

[46] P. Sinha, Object Recognition via Image Invariants: A Case Study, Investigative Ophthalmology and Visual Science, 1994,35(4):1735-1740.

[47] J. Miao, B. Yin, K. Wang, L. Shen, and X. Chen, A Hierarchical Multiscale and Multiangle System for Human Face Detection in a Complex Background Using Gravity-Center Template, Pattern Recognition, 1999,32(7):1237-1248.

[48] Y.H. Kwon and N. da Vitoria Lobo, Face Detection Using Templates, Proc. Int'l Conf. Pattern Recognition, 1994:764-767.

[49] K. Lam and H. Yan, Fast Algorithm for Locating Head Boundaries, J. Electronic Imaging, 1994,3(4):351-359.

- [50] A. Lanitis, C.J. Taylor, and T.F. Cootes, An Automatic Face Identification System Using Flexible Appearance Models, *Image and Vision Computing*, 1995,13(5):393-401.
- [51] G.J. Edwards, C.J. Taylor, and T. Cootes, Learning to Identify and Track Faces in Image Sequences. *Proc. Sixth IEEE Int'l Conf. Computer Vision*, 1998 :317-322.
- [52] M. Turk and A. Pentland, Eigenfaces for Recognition, *J. Cognitive Neuroscience*, 1991,3(1):71-86.
- [53] E. Osuna, R. Freund, and F. Girosi, Training Support Vector Machines: An Application to Face Detection, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1997:130-136.
- [54] H. Rowley, S. Baluja, and T. Kanade, Neural Network-Based Face Detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Jan. 1998,20(1):23-38.
- [55] E. Osuna, R. Freund, and F. Girosi, Training Support Vector Machines: An Application to Face Detection, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1997:130-136.
- [56] D. Roth, Learning to Resolve Natural Language Ambiguities: A Unified Approach, *Proc. 15th Nat'l Conf. Artificial Intelligence*, 1998:806-813.
- [57] T. Rikert, M. Jones, and P. Viola, A Cluster-Based Statistical Model for Object Detection, *Proc. Seventh IEEE Int'l Conf. Computer Vision*, 1999,2:1046-1053.
- [58] F. Samaria and S. Young, HMM Based Architecture for Face Identification, *Image and Vision Computing*, 1994,12:537-583.
- [59] F.S. Samaria, Face Recognition Using Hidden Markov Models, PhD thesis, Univ. of Cambridge, 1994.
- [60] R.J. Qian and T.S. Huang, Object Detection Using Hierarchical MRF and MAP Estimation, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1997:186-192.
- [61] Jie Yang, A. Weibel, A real-time face tracker, *Proceedings of ACV'96(Sarasota, Floreda, USA)*:142-147.
- [62] Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.). Improved Automatic Skin Detection in Color Images, *Proc. V2th Digital Image Computing: techniques and*

applications, Dec.2003:10-12.

- [63] Manoj Seshadrinathan, Jezkiel Ben-Arie. Pose Invariant Face Detection, EC-VIP-MC 2003, 4th Eurasip conference focused on Video/Image Processing and Multimedia Communications, 2-5 July 2003, Zagreb, Croatia.
- [64] Jean-Christophe Terrillon. Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images, Vision Interface'99, Trois-Rivieres, Canada, 19-32 May.
- [65] J Yang, W Lu, A Waibel, Skin-color modeling and adaptation. In: Proc ACCV, 1998:687~694
- [66] Phillips P.J., Moon H., et al. .The FERET Evaluation Methodology for Face-Recognition Algorithms. IEEE TPAMI, 2000, 22(10) :1090-1104.
- [67] Phillips P.J., Grother P., Micheals R.J, Blackburn D.M., Tabassi E., and Bone J.M. FRVT 2002: Evaluation Report, March 2003
- [68] Moses Y, Ullman S. Limitation of Non-Model-Based Recognition Schemes Springer-Verlag, [A]. Proc ECCV-92, Sandini G, ed [C]. 1992.820-828.
- [69] Warrington E, Taylor A. Two Categorical Stages of Object Recognition [J]. Perception. 1978, 7:152-164
- [70] T. Sim and T. Kanade. Combining Models and Exemplars for Face Recognition: An Illuminating Example. CVPR . 2001
- [71] H. F. Chen, P. N. Belhumeur, and D. W. Jacobs. In Search of Illumination Invariants, CVPR. 2000.
- [72] T. Riklin-Raviv and A. Shashua . The Quotient Image: Class Based Recognition and Synthesis Under Varying Illumination Conditions, CVPR. 1999.
- [73] Basri R. and Jacobs D. Lamberitian Reflectance and Linear Subspaces. Proc. Eighth IEEE Int'l Conf. Computer Vision, 2001:383-390
- [74] Georghiades A. S. Belhumeur P. N. Kriegman D. J. From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. IEEE Trans. PAMI , 2001, 23(6):643-660
- [75] Murase H. Nayar S. Visual Learning and Recognition of 3D object from appearance. IJCV, 1995, 14:5-24

- [76] Shashua A. and Riklin-Raviv T. The Quotient Image: Class-Based Re-Rendering and Recognition with Varying Illuminations. IEEE Trans. On PAMI, 2001, 23(2): 129-139
- [77] Phillips P.J., Vardi Y. Efficient Illumination Normalization of Facial Images. PRL, 1996, 17: 921-927
- [78] Zhao W. and Chellappa R. Illumination-insensitive Face Recognition Using Symmetric Shape-from-Shading. Proc. Conf. Computer Vision and Pattern Recog., 2000
- [79] 章毓晋. 图象处理和分析. 清华大学出版社. 2000
- [80] 杨小铃, 高鹏东, 刘则毅, 彭翔. 基于模糊模式匹配的多角度人脸检测, 计算机应用研究, 2005, 11: 231-232
- [81] 雷明, 张军英. 一种可变光照条件下的肤色检测算法. 计算机工程与应用, 2002, 38(24): 123-125.

# 基于肤色模型的人脸检测算法研究

作者：[柯研](#)

学位授予单位：[华中科技大学](#)

本文链接：[http://d.g.wanfangdata.com.cn/Thesis\\_D064745.aspx](http://d.g.wanfangdata.com.cn/Thesis_D064745.aspx)