

摘要

人脸检测是在输入图像中确定所有人脸的位置、大小、位姿的图像处理过程，最初作为自动人脸识别系统的定位环节被提出，近年来由于其在安全访问控制、视觉监测、基于内容的检索和新一代人机界面等领域的应用价值，开始作为一个独立的课题受到研究者的普遍重视。

本文主要研究彩色静止图像中的人脸检测。利用彩色图像中肤色特征缩小搜索区域，然后结合器官分布特征、积分投影、边缘图像、方差投影等信息实现人脸特征提取，并最终定位出人脸。

在肤色分割中，使用色彩均衡技术消除由光照引起的色彩偏移，并提出一种新的肤色模型（H_SI_I 模型），它对光亮度具有强鲁棒性。最后采用模糊聚类方法进行肤色再分割，以进一步去掉伪肤色像素点。

在特征提取算法中，首先根据器官分布特征，在肤色分割的基础上直接对孔洞匹配定位双眼。若眼睛定位失败，则采用水平灰度积分投影估计人眼水平位置，然后利用 Canny 算子得到边缘图像，再用加窗处理快速定位人眼。若眼睛定位再次失败，则在利用积分投影估计出人眼水平位置基础上，利用两次方差投影运算准确定位出眼睛。

从实验效果可以看出，本文提出的人脸检测算法能够实现彩色图像中复杂背景下正面的以及有一定侧面旋转角度和表情变化的人脸检测，具有较高的检测性能。

主题词：人脸检测 肤色模型 模糊聚类 积分投影 边缘检测 方差投影

ABSTRACT

Face detection is to locate the faces and determine their sizes in an image. Face detection problem, originated in face recognition research as its first step to locate faces in the images, has recently been intensively researched as an independent problem because of its applications in many different fields including security access control, visual surveillance, content-based information retrieval advanced human and computer interaction.

In this paper, the face detection of color still images is discussed. Skin characteristic is used to reduce the search range firstly. Then the facial features extraction comes into being by integrating some information, such as the physical structure features of facial organs, the integral projection, edge images, and the variance projection. The exact face location is detected by the facial features extraction finally.

In skin color segmentation, a color balance technique is used to remove the color bias produced by light condition. Meanwhile, a new skin model (H_SI_I skin model) which has the robustness for variety of brightness is built. At last, a cluster-based skin model to remove background points which is wrongly retained by the H_SI_I skin model is proposed.

In the facial features extraction, the “holes” are directly matched to locate eyes basing on the physical structure features of facial organs firstly. If eyes can not be gained accurately, facial region’s horizontal gray-level projection is used to estimate the horizontal position of eyes, then the edge features are checked by Canny operator, the proper position of eyes is quickly obtained by dealing with “windows” finally. If eyes can not be found still, basing on the horizontal position of eyes, the exact position of eyes is extracted by using the variance projection two times.

Experiments show that the proposed face detection algorithm can detect human faces up right or rotated with little angle in color images with complex background, and expression change is tolerated, so the algorithm is feasible and effective.

Keyword: face detection; skin model; fuzzy cluster; integral projection; edge detection; variance projection

表 目 录

表 4.1 人脸检测结果统计	37
----------------------	----

图 目 录

图 1.1 人脸检测在人脸图像应用系统中的应用	1
图 1.2 人脸检测系统流程框图	8
图 2.1 RGB颜色模型	10
图 2.2 HSI颜色圆和HSI颜色空间	13
图 2.3 色彩均衡前后肤色分割结果的对比	15
图 2.4 不同亮度下肤色样本在H_S平面和H_SI平面聚类结果的对比	17
图 2.5 不同亮度下同一幅图片的肤色分割结果	17
图 2.6 肤色模型初次分割和模糊聚类后二次分割的对比	20
图 2.7 形态学处理前后的二值化图像	22
图 2.8 区域归并和几何筛选后得到的候选人脸区域	24
图 2.9 肤色分割算法流程图	25
图 3.1 特征提取算法的流程图	26
图 3.2 嘴区与双眼关系图	29
图 3.3 基于器官分布特征的人脸定位方法流程图	28
图 3.4 水平积分投影估计人眼水平位置的过程	30
图 3.5 边缘检测和加窗处理定位人眼的过程	33
图 3.6 基于积分投影和边缘图像的人脸定位方法流程图	34
图 3.7 积分投影和方差投影的比较	35
图 3.8 左眼方差投影图	35
图 3.9 方差投影定位人眼的过程	36
图 3.10 基于积分投影和方差投影的人脸定位方法流程图	36
图 4.1 基于肤色分割算法和方法 1 检测得到的相关图像	37
图 4.2 基于肤色分割算法和方法 2 检测得到的相关图像	38
图 4.3 基于肤色分割算法和方法 3 检测得到的相关图像	38
图 4.4 部分图像的人脸检测结果	39
图 4.5 双眼匹配失败造成人脸漏检的实例	40
图 4.6 人眼水平位置估计失败造成人脸漏检的实例	41
图 4.7 方差投影错误定位人眼成人脸误检的实例	41

独 创 性 声 明

本人声明所呈交的学位论文是我本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，除文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表和撰写过的研究成果，也不包含为获得国防科学技术大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文题目：彩色图像中复杂背景下的人脸检测

学位论文作者签名：张书真 日期：2006年11月5日

学 位 论 文 版 权 使用 授 权 书

本人完全了解国防科学技术大学有关保留、使用学位论文的规定。本人授权国防科学技术大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档，允许论文被查阅和借阅；可以将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、汇编学位论文。

(保密学位论文在解密后适用本授权书。)

学位论文题目：彩色图像中复杂背景下的人脸检测

学位论文作者签名：张书真 日期：2006年11月5日

作者指导老师签名：杨卫平 日期：2006年11月6日

第一章 绪论

1.1 人脸检测的研究背景与意义

随着社会的发展，进行快速、有效、自动的人身辨别的要求日益迫切。生物特征是人的内在属性，具有很强的自身稳定性和个体差异，因此是身份辨别的主要依据。相比利用指纹、视网膜、声音等其它人体生物特征的人身鉴别方法，人脸识别具有蕴涵信息量大、直接、友好、方便等特点，易于为用户所接受。人脸是个信息极丰富的模式集合，是人类互相判别、认识、记忆的主要标志，通过对人脸的观察，我们可以判断一个人的性别、种族，甚至身份和性格。因此，对人脸的研究不仅局限在工业、科学上，在哲学、心理学、医学、人文等领域也有着广泛的研究意义。

人脸检测（face detection）是指在输入图像中确定所有人脸（如果存在）的位置、大小和姿态的过程。人脸检测系统输入的是可能包含人脸的图像，输出的是关于图像中是否存在人脸的数目、位置、尺度等信息的参数化描述。

人脸检测问题最初来源于人脸识别（face recognition）。人脸识别的研究可以追溯到 20 世纪 60—70 年代，经过几十年的曲折发展该技术已日趋成熟。人脸检测是自动人脸识别系统的一个关键环节，但早期的人脸识别研究主要针对具有较强约束条件的人脸图像（如无背景的图像），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没有受到重视。近些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开始作为一个独立的课题受到重视。今天，人脸检测的背景应用已经远远超出了人脸识别系统的范畴，在出入安全检查、视频监视、智能人机接口、基于内容的图像检索和视频编码中都有广泛的应用价值。

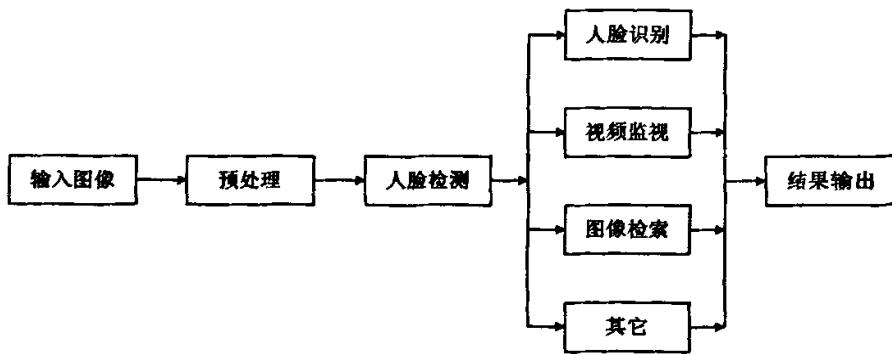


图 1.1 人脸检测在人脸图像应用系统中的应用

人脸检测在人脸图像应用系统中的作用如图 1.1 所示，从中我们可以看出，人脸检测的好坏直接关系到人脸应用系统的性能，因此人脸检测技术对整个系统的性能起着关键性

作用。

人脸检测系统的主要性能指标包括：

- (1) 检测准确性，即检测是否含有人脸的结果要准确，尽量减少误检率和漏检率。
- (2) 定位精确性，即检测得到的人脸的位置要精确，尽量减少结果中非人脸区域，特别是根据人脸的主要特征对图像进行定位时，定位精度要在像素级，这样可以为人脸识别的高识别率打下基础。
- (3) 检测的快速性，如果系统是实时的，即对检测的速度要求较高，这就要求在算法上进行优化，但是一般情况下，提高速度往往是以降低准确率为代价，所以对检测速度的要求，不同的识别系统有所不同。

目前，国外对人脸检测问题的研究很多，比较著名的有 MIT、CMU、USC 等；国内的清华大学、北京工业大学、亚洲微软研究院、中科院计算机技术研究所、中科院自动化研究所等都有人员从事人脸检测方面的相关研究。而且，MPEG7 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加，IEEE 的 PMAI 还于 1997 年 7 月出版了人脸识别专辑。IEEE 的 FG (IEEE International Conference on Automatic Face and Gesture Recognition)、ICIP (International Conference on Image Processing)、CVPR (Conference on Computer Vision and Pattern Recognition) 等重要国际会议上每年都有大量关于人脸检测的论文。

1.2 人脸检测的研究难点

人脸是一类具有复杂细节变化的自然结构目标，人脸检测涉及到模式识别、图像处理、人工智能等多方面知识，因此人脸检测技术虽然在近十年中得到广泛的关注和长足的发展，国内外很多研究人员提出许多方法，在不同领域取得一定的成果，但是对于寻找一种普遍适用于各种复杂情况的准确率很高的检测算法，仍有很大的探索空间。人脸检测的困难主要同以下几个因素有关：

- (1) 人脸数目和尺寸。图像中人脸的数目和尺寸具有不确定性。
- (2) 头部旋转。头部在图像中按不同角度的旋转，将给人脸特征的提取和匹配造成一定困难。
- (3) 模式可变性。人脸是个包含五官、毛发等极不规律的复杂待测目标，不同的人脸在形状、大小、颜色、质地等方面都有很大的变化。同时，一般意义的人脸存在着诸如眼镜、胡须等附属物，它在人脸检测过程中必然形成干扰。
- (4) 脸部表情。脸部表情的变化经常被看作是一种模板变形，它经常会影响到人脸的结构信息，使得人脸检测变得困难。
- (5) 遮挡。人脸可能被其它物体遮挡，或者在一幅有多张人脸的图像中，其中的一些脸可能被其它人脸遮挡，这样的人脸很难检测出来。
- (6) 成像条件。光照条件（光谱、光源的分布和亮度）、摄像机的性能（感应器反应、

镜头)等因素直接影响图像的信息结构,使图像原始特征的提取和匹配变得困难。

上述因素同人脸检测有着非常密切的联系,通过上述问题的解决,成功构造出人脸检测系统,将为解决其它类似复杂模式的检测问题提供重要的启示。

1.3 人脸检测的研究课题

人脸检测技术已经吸引了很多研究人员的注意,目前,人脸检测的研究课题主要集中在以下五个方面:

- (1) 怎样正确地描述一张典型的人脸。
- (2) 怎样处理不同尺寸的人脸。
- (3) 怎样加快人脸检测的速度。
- (4) 怎样准确地定位人脸。
- (5) 怎样将检测的结果联系起来。

1.4 人脸检测方法综述

人脸图像所包含的模式特征非常丰富,如何选择合适的模式特征,是人脸检测研究的一个关键性问题,总的来说人脸模式的主要特征包括肤色特征和灰度特征,前者适用于构建快速的人脸检测算法,后者利用了人脸区别于其它物体更为本质的特征,是人脸检测领域研究的重点。人脸模式具有复杂而细致的变化,因此一般需要采用多种模式特征综合的方法。

近几年,人脸检测技术发展迅猛,研究人员提出很多人脸检测的方法,但是由于人脸检测问题的复杂性,实现通用的人脸检测方法还不实际,因此解决特定约束条件下或某种应用背景下的人脸检测问题仍将是该领域研究的主要课题。

从目前来看,人脸检测方法大致可分为四类^[1,2]:

1.4.1 基于知识的方法

这种方法是利用对人脸的先验知识给出检测规则来进行人脸检测。人脸局部特征的分布总是存在着一定的规则,如图像中出现的脸通常会有一个鼻子、一张嘴和一双彼此对称的眼睛。这些特征之间的联系能够被它们之间的相对距离和位置表示出来。在一幅输入的图像中,脸部特征将首先被提取出来,然后基于检测规则将候选的脸部区域辨别出来。

- (1) 器官分布法:虽然人脸在外观上变化很大,但遵循一些普遍适用的规则,如五官的空间分布大致符合“三停五眼”的标准等,检测图像中是否有人脸即是测试该图像中是否存在满足这些规则的图像块。这种方法一般有两种思路:一种思路是“从上至下”,其中最为简单有效的是 Yang 和 Huang^[3]提出的 Mosaic 方法,它给出了基于人脸区域灰度分布的规则,依据这些规则对图像从粗分辨率

到高分辨率进行筛选，以满足这些规则的程度作为检测的判据。这种方法的吸引之处在于使用从粗到细的策略减少了计算量。另一种思路则是“从下至上”，先直接检测几个器官可能分布的位置，然后将这些位置点分别组合，利用器官分布的几何关系准则进行筛选，找到可能存在的人脸^[4]。

- (2) 投影法^[5,6]：这是一种利用图像的投影快速定位人脸边界和脸部特征的方法。在处理一幅输入图像时，首先获得其水平投影，然后通过检测其图像的急剧变化处得到两个局部最小值，这两个值对应头部的左右轮廓线。类似地，通过检测垂直投影的局部最小值，检测出人脸的嘴唇、鼻尖和眼睛。这种方法的缺陷在于无法有效检测出有多张人脸的图像以及在复杂背景中的人脸。
- (3) 对称法：人脸具有一定的轴对称性，各个器官也具有很强的点对称性。Zabrodsky^[7]提出连续对称性检测方法，检测一个圆形区域的对称性，从而确定是否为人脸；Riesfield^[8]提出广义对称变换方法检测局部对称性强的点来进行人脸器官定位；卢春雨^[9]则定义方向对称变换，分别在不同方向上考察对称性，不仅能用来寻找强对称点，而且还可以描述强对称性物体的形状信息，这样的变换对人脸偏转、脸部表情变换、光照等因素都不敏感，使人脸器官定位更为有效。

1.4.2 基于特征的方法

基于特征的方法与基于知识的方法相对应，研究人员试图先找出人脸的稳定特征，再利用其它信息进一步验证人脸的存在。人脸的稳定特征是指在光照变化和姿势变化时相对不变的特征，包括人的脸部特征、纹理、肤色以及特征综合等。

- (1) 脸部特征：研究人员提出了许多方法都是首先检测脸部特征，然后推断是否有人脸存在。人的脸部特征，如眼睛、眉毛、嘴巴、鼻子等通常使用边缘检测来获得。基于这些特征，建立起描述特征之间的统计模型，然后去校验是否有人脸存在。这种方法的缺陷在于表情、光照、噪声等因素会造成脸部特征边界弱化，这样就影响到人脸检测结果的准确性。
- (2) 纹理特征：人脸的纹理特征分为三类，即皮肤、毛发和其它类，这些独特的纹理可以被用来区别其它物体。Augusteijn 和 Skufca^[10]利用是否类似人脸纹理来推断图像中是否有人脸存在。Dai 和 Nakano^[11]利用空间灰度共生矩阵（Space Gray-Level Dependence matrix, SGLD）纹理图信息作为特征进行低分辨率的人脸检测。这种方法的优点是在人脸不正直，或有胡须、眼镜等附属物存在时，检测结果仍然比较准确。
- (3) 肤色特征：研究^[12]表明，人类肤色在去除亮度的色度空间具有聚类性，颜色信息在一定程度上可以将人脸同大部分背景区分开来，因此可以利用肤色模型有效地检测人脸。利用肤色信息检测人脸的方法具有计算量小，方法相对简单，

易于实时性处理及适合复杂背景下的人脸检测定位等特点。Lee 等^[13]设计了肤色模型表征人脸颜色，利用感知模型进行复杂背景下人脸及器官的检测与分割。Saber 等^[14]将颜色、形状等结合起来进行人脸检测。

1.4.3 基于模板匹配的方法

基于模板匹配的方法可以用来检测人脸，也可以用来做人脸特征定位。首先给定一些不同标准的模板来描述人脸的整体特征或局部特征，然后根据模板与目标图像的相似性进行匹配检测。

- (1) 预定模板匹配法：根据人脸的先验知识确定出人脸轮廓模板以及各个器官特征的子模板，先通过计算图像区域和人脸轮廓模板的相关值检测出候选人脸区域，然后利用器官特征子模板验证上一步检测出的候选人脸区域是否为真的人脸。Sakai^[15]等人早期对照片中的正面人脸进行检测，他们使用了眼睛、鼻子、嘴巴、脸部轮廓的子模板建立起人脸模型。Craw^[16]等人首先使用一个 Sobel 滤波器来提取边缘，然后将这些边缘集合起来搜索基于几个约束条件的人脸模板。梁路宏等^[17]使用双眼模板首先进行粗检测，然后使用不同长宽比的人脸模板进行匹配，最后使用马赛克规则进行验证。预定模板匹配法的缺点在于不能有效处理尺度、姿态和形状的变化。
- (2) 变形模板法^[18]：其主要思想是定义一个可变形的参数模板和一个能量函数来描述特征，通过一个非线性最优化方法求得能使能量函数最小的参数模板，此模板即被认为是所求特征的描述。这种方法充分考虑到人脸是变形体的特点，稳定可靠，而且与姿态和光照无关，但仍然存在能量函数的系数难以适应一般情况和计算量巨大的问题。并且该方法必须是在感兴趣的目标的附近，变形模板必须被初始化。
- (3) 动态轮廓法：又称 Snakes 模型法。Kass 等^[19]人提出在不需要更多先验知识或高层处理结果指导下实现自追迹以得到目标的闭合、光滑、连续的轮廓线，该方法具有较强的抗噪能力。Snakes 模型通过对一条初始给定的连续闭合曲线计算初始能量函数，然后通过变分方法求取能量函数的最小化，以得到人脸的轮廓曲线。

1.4.4 基于统计模型的方法

由于人脸图像的复杂性，显式地描述人脸特征具有一定的困难，因此另一类方法——基于统计模型的方法越来越受到重视。此类方法将人脸区域看作一类模式，即模板特征，使用大量的“人脸”与“非人脸”样本训练、构造分类器，通过判别图像中所有可能区域属于哪类模式的方法实现人脸检测。实际上，人脸检测问题被转化为统计模式识别的二分类问题，即判断区域是否为人脸。

- (1) 特征空间法：将人脸图像变换到某一特征空间，根据特征空间的分布规律划分“人脸”与“非人脸”两类模式。主分量分析^[20] (Principal-Component Analysis, PCA) 是一种常用的方法。它根据图像的统计特征进行正交变换 (KL 变换)，以消除原有向量各个分量间的相关性。变换得到对应特征值依次递减的特征向量，即特征脸 (Eigen face)。利用前若干张特征脸将人脸向量投影到主元子空间和与其正交的补空间，相应的距离度量分别称为 DIFS (Distance In Feature Space) 和 DFFS (Distance From Feature Space)。对于人脸检测问题，由于没有考虑“非人脸”样本的分布，需要同时使用 DIFS 和 DFFS 才能取得较好的效果。属于特征空间方法的还有因子分解法 (Factor Analysis, FA) 和 Fisher 准则法 (Fisher Linear Discriminant, FLD)。Yang 等^[21] 在混合线性子空间中对“人脸”和“非人脸”样本进行建模，分别使用基于 EM 算法的扩展 FA 方法和基于自组织映射 (Self-Organizing Map, SOM) 的 FLD 方法构造检测器。
- (2) 人工神经网络法：人工神经网络方法 (Artificial Neural Network, ANN) 通过对人脸样本集和非人脸样本集进行学习产生分类器。该方法对于人脸这类复杂的、难以显性描述的模式，具有独特的优势。G. Burel 和 D. Carel^[22] 在多分辨率分析的基础上采用多层感知器 (Multi-Layer, MLP) 进行人脸检测。J. Vincent 等^[23] 利用多个 MLP 构成分层人脸特征检测系统 (Hierarchical Perception Feature Location System, HPFLS)。类似地，P. Juell 和 R. Marsh^[24] 也提出了一种分层的神经网络，包括子层的检测眼睛、嘴巴、鼻子的三个 BP 网络以及父层检测全脸的一个 BP 网络等。
- (3) 支持向量机法：支持向量机 (Support Vector Machines, SVM) 是 Vapnik 等^{[25],[26]} 提出的基于结构风险最小化原理 (Structural Risk Minimization Principle, SRM) 的统计学理论，用于分类和回归问题。SRM 使 VC (Vapnik-Chervonenkis) 维数的上限最小化，这使得 SVM 方法比基于经验风险最小化 (Empirical Risk Minimization Principle, ERM) 的人工神经网络方法具有更好的泛化能力。但是由于 SVM 的训练需要求解计算复杂度极高的二次规划问题，SVM 对人脸研究的改进主要在于减少支撑向量的数量和训练方法的改进。如 Osuna 等^[27] 在 SVM 的训练使用了大量人脸样本和“自举”方法^[28] 收集“非人脸”样本，并使用逼近优化的方法^[29] 降低支撑向量的数量；Platt 等^[30] 提出 SMO (Sequential Minimal Optimization) 算法解决了 SVM 训练的困难。
- (4) 概率模型法：基于概率模型的方法常见的有两种思路。一种思路是利用贝叶斯原理计算输入图像区域 region 属于人脸区域 object 的后验概率 $P(\text{object}/\text{region})$ ，据此对所有可能的图像窗口进行判别。CMU 的 Schneiderman 等^{[31],[32]} 提出一种基于后验概率估计的人脸检测方法。另一种思路是应用隐马尔可夫模型 (Hidden Markov Models, HMM)，将人脸模式看作参数化的随机过程，

把人面部的额头、眼睛、鼻子、嘴巴和下巴等器官所在部位看作随机过程的状态，通过对符合人脸各器官分布的状态的随机过程的检测来实现对人脸检测，该方法一般只使用“人脸”样本进行训练，主要针对用于人脸识别的头肩部图像^{[33][34]}。

1.5 论文的结构和主要工作

如前所述，在人脸检测定位方面，前人已经提出了一些解决方案，但在定位方法的简便性、算法的稳定性、定位结果的准确性以及适用条件等方面依然存在局限：有些算法限定检测图片只包含简单背景（有些特定的应用的确可限定为简单的背景）；有些算法能比较准确地检测出入脸，但时间复杂度太高，难以达到整个系统的实时性要求；有些算法速度虽快，但误检率和漏检率较高。

随着图像采集技术的飞速发展，随着计算机速度的提高，以及随着计算机色彩理论的深入研究，彩色图像处理逐渐成为数字图像处理的核心内容。在这种情况下，对于人脸检测技术而言，实现对彩色图像的人脸检测成为该技术发展的必然趋势。目前，越来越多的研究者开始利用将肤色信息与人脸的灰度信息相结合的方法来解决人脸检测与定位问题，本文的检测算法正是在这种思想的指引下完成的。

文中首先利用肤色聚类特征，通过自建的 H_SI_I 肤色模型和模糊聚类算法得到肤色区域，然后对肤色区域进行数学形态学处理和几何筛选，获得候选人脸区域。在肤色分割的基础上，针对候选人脸区域，将器官分布特征、积分投影、边缘图像、差分投影等信息相结合，实现对人脸特征的提取。最后，在特征提取的基础上定位出人脸。

论文共分五章，各章主要内容概述如下：

第一章：绪论。概述了人脸检测技术的研究内容和应用范围，重点阐述了人脸检测技术的主要方法。

第二章：肤色分割和候选区域筛选。首先利用色彩均衡技术对彩色图像进行预处理，以增强肤色分割算法的稳健性。然后，根据光照理论，分析了 HSI 颜色空间对光亮度变化敏感的原因，建立起一个改进的颜色空间（H_SI_I 空间），通过实验证明，在该空间的肤色模型对于光亮度变化具有强鲁棒性。在肤色初分割的基础上，再采用模糊 C 聚类算法进行肤色二次分割，以进一步去掉伪肤色像素点，提高肤色分割的准确性。在候选区域筛选阶段，采用形态学处理方法和基于面积占有率及长宽比的几何筛选方法得到候选人脸区域。

第三章：特征提取和人脸定位。对于特征提取，本文分三步来实现。首先根据器官分布特征，在肤色分割的基础上直接对孔洞匹配定位人眼。若眼睛定位失败，则采用水平灰度积分投影估计人眼水平位置，然后利用 Canny 算子得到边缘图像，再用加窗处理快速定位人眼。若眼睛定位再次失败，则在积分投影所估计出人眼水平位置的基础上，利用两次方差投影运算准确定位出眼睛。双眼一旦提取出来，即可利用“三停五眼”规则输出人脸

区域。

第四章：实验结果及分析。在本章中给出了实验测试结果及部分检测实例，并通过实验结果分析检测算法的性能。

第五章：总结与展望。对全文进行了总结，对今后的工作提出了进一步要求。

本文主要研究了彩色静止图像中的正面人脸图像的检测，系统流程框图如图 1.2 所示。

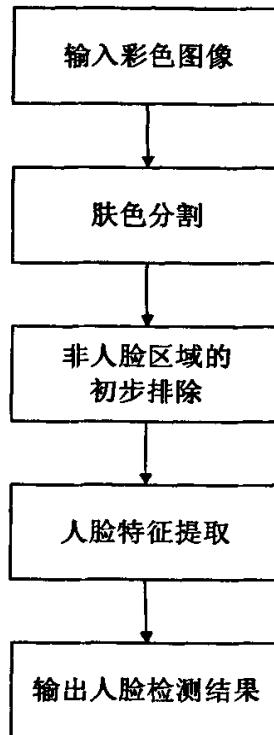


图 1.2 人脸检测系统流程框图

第二章 肤色分割和候选区域筛选

肤色是人脸的重要信息，不依赖于脸部的细节特征，对于旋转、表情等变化情况都能适用，具有相对的稳定性，并且和大多数背景物体的颜色相区别。肤色较灰度信息包含更多的信息，是人体表面最为显著的特征之一。简单环境下的基于肤色的人脸检测、人脸跟踪、手势识别等，可以利用肤色分割直接实现，对于复杂环境下的应用，利用人体肤色特征可以大幅度减小搜索空间，提高检测速度。

目前，基于统计思想的肤色分割器性能相对较好，其它方法就是基于肤色空间聚类的方法，其主要思想是用大量肤色像素得到肤色在颜色空间的聚类，利用闭合曲线或曲面进行拟合，检测时判断像素是否落在拟合的曲线或曲面内，若落到则判断该像素为肤色像素。该方法主要需要解决两个问题：第一是颜色空间的建立；第二是如何拟合出肤色在所建立颜色空间的闭合曲线或曲面，即如何构建出实用的肤色聚类模型。

肤色分割后仍然存在大量的噪声，且裸露的手臂、脖子等非人脸部分以及背景中的类肤色区域同时也会检测出来，因此需要对分割后的类肤色区域进行去噪处理和初步的人脸筛选，从而得到候选人脸区域。

2.1 颜色空间

场景或物体的颜色是由照射光源的光谱成分、光线在物体上反射和吸收的情况决定。可见光的波长分布在 380nm 到 780nm 之间，人的颜色感觉是不同波长的可见光刺激人的视觉器官的结果。根据光度学和色度学原理，所有颜色都可以用相互独立的三种基本颜色混合得到，这三种颜色被称为三基色，由此构成了颜色空间。著名的格拉斯曼定律表明视觉对颜色的反应取决于红、绿、蓝三种颜色在输入量中的代数和。格拉斯曼定律包括如下四个方面的内容：

- (1) 所有颜色都可以用互相独立的三基色混合而得到。
- (2) 若加入三基色的混合比相等，则色调 (Hue) 和饱和度 (Saturation) 也相等。
- (3) 任意两种颜色混合产生的新颜色与采用三基色分别合成这两种颜色的各自成分再混合得到的结果相同。
- (4) 混合色的光亮度 (Luminance) 是原来各分量光亮度的总和。

这里的色调、饱和度、亮度是表示色觉程度的。色调表示各种颜色，饱和度表示颜色的深浅。

根据计算机色彩理论，对于同一种颜色而言，在计算机中有不同的表达方式，这样就形成了不同的颜色空间。每一种颜色空间都有其各自的产生背景及应用领域。下面介绍几种常用的颜色空间^[35-37]。

2.1.1 RGB 颜色空间

RGB 颜色空间也称为红、绿、蓝基色模型。为标准化起见, CIE (国际照明委员会) 在 1931 年选择红色 (波长为 700.0nm), 绿色 (波长为 546.1nm), 蓝色 (波长为 435.8nm) 三种单色光作为表色系统的三基色, 这就是 RGB 颜色空间。

RGB 颜色空间可以用图 2.1 所示的 RGB 直角坐标定义的单位立方体来说明。坐标系原点 (0, 0, 0) 表示黑色, 坐标点 (1, 1, 1,) 表示白色, 坐标轴上的三个顶点表示三个基色。因此颜色空间是三维的线性空间, 任意一种具有一定亮度的颜色光都可以用空间的一个点或一个矢量来表示。

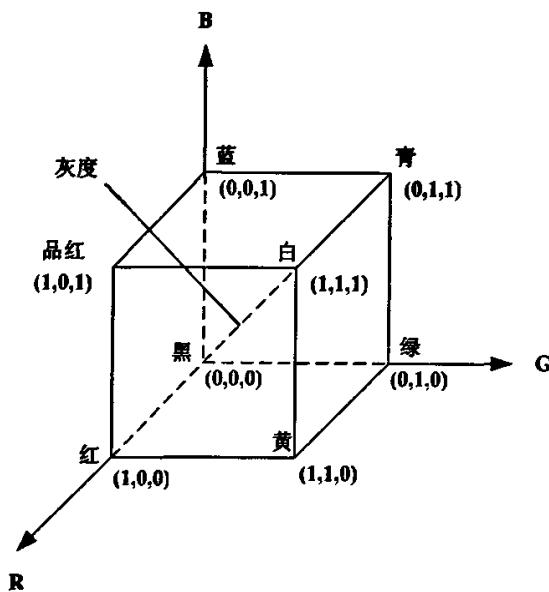


图 2.1 RGB 颜色模型

RGB 适合用来进行显示, 广泛用于视频监视器显示和彩色摄像机中。但是 R、G、B 分量之间有高的相关性, 所以它不适合用来进行场景分割和分析。高相关性是指如果光强变化了, 色调、饱和度和灰度三个分量都会发生改变。因此, 大部分情况下, 是以 RGB 颜色空间为基础来描述其它类型的颜色空间, 将其它颜色空间的基色描述为 RGB 三基色的线性或非线性函数。

2.1.2 RGB 线性转换的颜色空间

(1) CIE 颜色空间

使用 RGB 模型生成颜色时, 三基色比例系数会出现负值, 使用起来十分不便。同时, 不同研究者所用的三基色和标准白色不同, 使得研究结果很难比较。因此, 1931 年国际照明委员会 CIE (Commission Internationale de l'Eclairage—the International Commission

on Illumination) 规定一种新的颜色系统, 定义为 CIE-XYZ 颜色模型, 具体表达式为:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2.7689 & 1.7517 & 1.1302 \\ 1.0000 & 4.5907 & 0.0601 \\ 0.0000 & 0.0565 & 5.5943 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.1)$$

由此可导出一系列的 CIE 颜色空间, 比如 CIE-Lab、CIE-Luv 等。

CIE 颜色空间能够更独立地控制颜色和灰度信息, 并且比 RGB 三基色要简单。在颜色空间中基于几何学的分离可直接进行颜色比较。因此, 这种模型用于测量小的颜色差别很有效。

(2) CMY 颜色空间

青 (Cyan)、品红 (Magenta)、黄色 (Yellow) 是光的二次色, 换句话说, 它们是颜料的原色。例如, 当被染成青色的平面被白光照射时, 从该平面反射的白光中将减去红光, 白光本身是等量的红、绿、蓝光的组合。

大多数彩色印刷或彩色复印都要求输入 CMY 数据或在机器内部将 RGB 转化为 CMY, 这种变换为

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.2)$$

在这里假设所有的颜色值都被归一化到 [0, 1] 区间。(2.2) 式表明了青色表面不反射红光, 品红色表面不反射绿光, 黄色表面不反射蓝光。

(3) YUV 颜色空间

YUV 是一个基本的颜色空间, 被欧洲的电视系统所采用。Y 指明视度, 即亮度 (实际上代表灰度值), U 和 V 指色调。由 RGB 空间线性变换到 YUV 空间的公式如下:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.3)$$

(4) YIQ 颜色空间

YIQ 是从 YUV 颜色空间衍生而来的, 主要应用于美国的电视系统。这种颜色空间的主要优点就是灰度信息可以从彩色数据中分离得到。因此, 相同的信号既可以用于彩色电视机也可以用于黑白电视机。在该颜色空间系统中, Y 代表颜色的亮度, I 和 Q 是两个相互正交的分量表示色度信息。RGB 转换到 YIQ 的线性变换如下:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & -0.312 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.4)$$

当 $0 \leq R \leq 1$, $0 \leq G \leq 1$, $0 \leq B \leq 1$.

(5) YCbCr 颜色空间

YCbCr 颜色空间被广泛用于数字视频中, JPEG、MPEG 等均采用此格式。在这个颜色系统中, 亮度信息被一个独立成分 Y 所表示, 颜色信息被保存为两种色差成分, Cb 和 Cr。Cb 是蓝色成分和一个参考值的差别, Cr 是红色成分和一个参考值的差别。RGB 空间转换为 YCbCr 空间的线性变换如下:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.996 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.5)$$

2.1.3 RGB 非线性转换的颜色空间

(1) 归一化的 RGB (Nrgb) 颜色空间

在彩色图像的分割中, 为达到好的分割效果, 常常希望颜色不依赖于光强的变化, 于是人们提出了归一化的 RGB。此颜色空间可以表示为:

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} \frac{1}{(R+G+B)} & 0 & 0 \\ 0 & \frac{1}{(R+G+B)} & 0 \\ 0 & 0 & \frac{1}{(R+G+B)} \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.6)$$

其中 [r, g, b] 表示归一化的三个颜色分量, 从 (2.4) 中可以看出 $r+g+b=1$, 因此, 归一化的 RGB 空间仅由 r 和 g 的值表示, 蓝色分量为冗余信息, $b=1-r-g$ 。

Nrgb 空间可以适应一定的光照条件变化。因为 RGB 空间中若有两个像素点 $[R_1, G_1, B_1]$ 和 $[R_2, G_2, B_2]$, 存在 $\frac{R_1}{R_2} = \frac{G_1}{G_2} = \frac{B_1}{B_2}$, 则在 Nrgb 空间中这两个像素点的 r, g, b 值相同, 表明两个有着相同色度, 不同亮度的像素点汇聚在 Nrgb 空间上的同一个位置上。但 Nrgb 空间中的 r, g, b 只去除了 R, G, B 中的相对亮度成分, r, g, b 中仍然存在亮度信息, 因此若用于彩色图像分割, 其亮度适应性不太好。

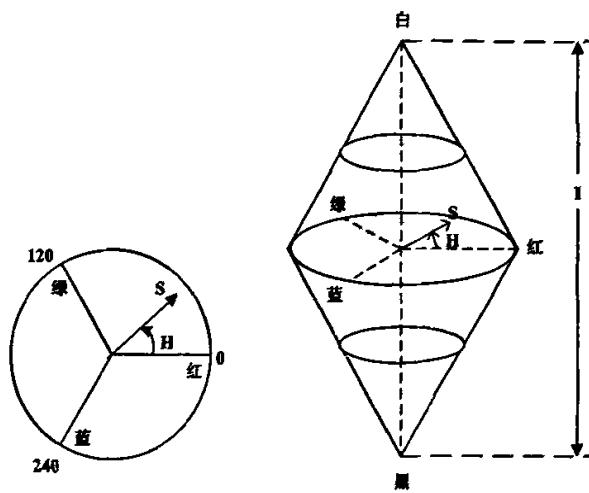
(2) HSI 颜色空间

HSI 颜色空间也称为色调 H、饱和度 S、亮度 I 模型。在模型中 I 分量与反映色彩本质特性的 H 分量和 S 分量分离开。因此 HSI 颜色空间符合人类的感知特性, 非常适合基于彩色描述的图像处理方法。

HSI 颜色空间可以用图 2.2 来说明。颜色的色调 H 是指颜色刺激的主波长, 用角度表示。当色调为 0 时为红色, 120 时为绿色, 240 时为蓝色。色调在 0 至 240 之间时代表可见频谱光, 240 至 360 之间的是非频谱光, 如紫色。饱和度 S 相当于颜色刺激的纯度, 即通常所说的颜色的深浅。饱和度用图中的点与颜色圆圆心之间的距离表示, 圆周上饱和度参数为 1, 圆心处饱和度参数为 0。亮度 I 是指颜色刺激的强弱, 在图中沿轴线方向从 0

到 1 变化，0 对应黑色，1 对应白色，是个与饱和度正交的量。

从图示中可以看出 HSI 颜色空间有如下两个缺点：第一，在 R, G, B 取值相近的情况下，色调是不稳定的，即 R, G, B 值有微小的变化，色调值可能变化很大，这种现象称为奇异值问题；第二，当亮度很高或很低时，色调分量也是不稳定的，所以对于图像中强烈的高光和阴影部分，还要进行一些特殊处理，才能更好地利用色调分量。



(a) HSI 颜色圆 (b) HSI 颜色空间

图 2.2 HSI 颜色圆和 HSI 颜色空间

由 RGB 空间转换到 HSI 空间的变换公式如下：

$$I = \frac{1}{3}(R + G + B) \quad (2.7)$$

$$S = 1 - \frac{3}{R+G+B} (RGB)_{\min} \quad (2.8)$$

$$H = \begin{cases} \theta & \text{若 } B \leq G \\ 2\pi - \theta & \text{若 } B > G \end{cases} \quad (2.9)$$

其中 $\theta = \arccos \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{\frac{1}{2}}} \right\}$, 而 $(RGB)_{\min}$ 指 R、G、B 三个分量中最小的

分量值。当饱和度 $S=0$ 时，色度 H 没有意义；当灰度 $I=0$ 时，饱和度 S 没有意义，灰度图像的饱和度都为 0。

与 HSI 同系列的颜色空间还有 HSV、LHSV 等，这里不再赘述。

2.2 彩色图像预处理

肤色等色彩信息经常受到光源颜色、图像采集设备的色彩偏差等因素的影响，在整体上偏离本质色彩，向某个方向移动，即通常所说的色彩偏冷、偏暖、照片偏黄、偏蓝等，

这种现象在艺术照片中更为常见。因此在肤色分割前，首先对彩色图像进行色彩平衡预处理，消除可能有的色光干扰。

在色彩平衡前应先判断图像是否有色光干扰，否则色彩平衡的结果可能会使得图像中人脸的颜色发生较大改变，反而影响肤色模型的检测效果。根据双色反射模型理论，由于物体表面存在镜面反射或界面反射，物体常常产生高光，在图像中，高光部分常对应高亮度区域（接近最高亮度值）。物体表面的镜面反射主要反射了照射光源的实际颜色，而非物体表面的颜色。基于这些事实，本文设定色光干扰判定的具体步骤如下：

- (1) 将整个图像中所有象素的亮度从高到低进行排列，取前 5% 的象素，如果这些象素数目足够多（如大于 500），则计算这些象素 R、G、B 分量的各自平均值 AR、AG、AB；
- (2) 计算 $\max(AR, AG, AB)/\min(AR, AG, AB)$ 的值，若大于 1.2 则判定该图像有色光干扰。

确认有色光干扰后，开始对图像进行色彩平衡。根据双色反射模型理论，图像的颜色由照射光源的光谱成分以及图像对入射光谱的反射系数来决定，因此当照射光源为色光时，彩色图像将出现整体上的色彩偏移。同时，彩色图像的高亮度区域主要反射光源实际颜色，因此色彩平衡的最终效果就是使高亮度区域的 R、G、B 三个分量的平均值趋于同一个灰度值。根据上述理论，本文对彩色图像实现色彩平衡的具体步骤如下：

- (1) 计算高亮度区域的平均灰度值 avgGray。avgGray 由前 5% 象素 R、G、B 分量的平均值 AR、AG、AB，得到：

$$\text{avgGray} = (AR + AG + AB)/3 \quad (2.10)$$

- (2) 对于图像的每一个象素 P，调整其 R、G、B 分量值。P(R)、P(G)、P(B) 的计算公式分别为：

$$P(R) = P(R) * \text{avgGray} / AR \quad (2.11)$$

$$P(G) = P(G) * \text{avgGray} / AG \quad (2.12)$$

$$P(B) = P(B) * \text{avgGray} / AB \quad (2.13)$$

- (3) 归一化所有象素的各个分量的值，让它们都在 [0, 255] 范围内。设 PR_{\max} 、 PB_{\max} 、 PG_{\max} 分别为图像中所有象素 R、G、B 分量的最大值，则 $\text{factor} = \max(PR_{\max}, PG_{\max}, PB_{\max})$ 为调整后象素分量的最大值，若 $\text{factor} > 255$ ，则将各分量调整为：

$$P(R) = 255 * P(R) / \text{factor} \quad (2.14)$$

$$P(G) = 255 * P(G) / \text{factor} \quad (2.15)$$

$$P(B) = 255 * P(B) / \text{factor} \quad (2.16)$$

图 2.3 给出一个色彩平衡的实例。原图收集于互联网。其中，2.3(a)是在偏黄光源下采集的，经后面阐述的肤色模型检测，得到的分割结果发现偏差很大，2.3(b)图是色彩均衡后肤色分割的效果，可看到各皮肤块基本检测出来，因此该平衡方案是有效的。



图 2.3 色彩均衡前后肤色分割结果的对比

- (a) 色彩均衡前的图片及肤色分割结果
 (b) 色彩均衡后的图片及肤色分割结果

2.3 基于颜色空间的肤色初分割

2.3.1 颜色空间的选取

研究表明^[39]: 尽管不同种族、不同年龄、不同性别的人，肤色看上去不同，但这种不同主要集中在亮度上，在去除亮度的颜色空间，不同人的肤色分布具有聚类性。因此肤色分割时，能否有效地将亮度分量和色度分量分离，是选取颜色空间的重要依据。本文采用了HSI颜色空间，它不但与人类的视觉感知相接近，而且色调值比较稳定，对一定范围光照的变化不太敏感，适用于彩色图像的分割。

正如上节所提到的，HSI颜色空间存在两个问题，一是在R, G, B取值相近的情况下，色调是不稳定的；二是当亮度很高或很低时，色调分量是不稳定的。对于第一个问题，因为R, G, B取值相近时，相当于灰度图像，这时，如果在正常光照范围内，就表明整幅图像都接近灰度图像，这不大可能，因此本文忽略该问题。对于第二个问题，本文提出一种改进的H_SI_I颜色空间，实验证明，在此空间建立的肤色模型能够适用于光照亮度变化范围较大条件下的肤色检测。

2.3.2 H_SI_I颜色空间

仔细分析HSI颜色空间的三个分量，通过公式(2.7)、(2.8)可以发现亮度I与饱和度S相互关联。由光照理论可知，图像的颜色由照射光源的光谱成分以及图像对入射光谱的

反射系数来决定。图像亮度的增加可以认为是向图像中增加了白光成分，同时白光成分的增加必然导致图像饱和度分量 S 相应地降低，因此 HSI 空间内建立肤色模型对光亮度变化敏感。

光亮度的增加意味着增加了白光成分，而白光中 R、G、B 分量应该相等（否则为色光）。设若图像的像素对三种基色光的反射系数相差不大，那么亮度改变引起的 R、G、B 分量的改变值 ΔR 、 ΔG 、 ΔB 近似相等，即 $\Delta R = \Delta G = \Delta B$ 。令 $R' = R + \Delta R$ ， $G' = G + \Delta G$ ， $B' = B + \Delta B$ ，则有 $R - G = R' - G'$ ，同理(R-B)、(R-G)、(R-B)、(G-B)的值都不改变，将上述结果代入 (2.9) 式可得色调分量 H 的值不随光亮度的变化而改变。

同时，将 (2.8) 式作一个变换，写成如下形式：

$$S = \frac{[R - (RGB)_{\min}] + [G - (RGB)_{\min}] + [B - (RGB)_{\min}]}{R + G + B} \quad (2.17)$$

然后再将 (2.7) 式与 (2.17) 式相乘，得：

$$SI = \frac{[R - (RGB)_{\min}] + [G - (RGB)_{\min}] + [B - (RGB)_{\min}]}{3} \quad (2.18)$$

光亮度改变后，因为 $\Delta R = \Delta G = \Delta B$ ，(2.18) 式中 $[R - (RGB)_{\min}] = [(R' - (R'G'B')_{\min})]$ ，同理 $[G - (RGB)_{\min}]$ ， $[B - (RGB)_{\min}]$ 的值都不改变，即 SI 的值不随光亮度的变化而改变。

通过上述分析，本文建立起一个改进的 H_SI_I 颜色空间。在该空间中，实际颜色相同的两个像素，在光亮度变化的条件下，仍然聚集到相同的空间位置上，在其二维投影 H_SI 平面上可以构造出适应光亮度变化范围较大的肤色模型。

2.3.3 H_SI_I 空间中的肤色模型

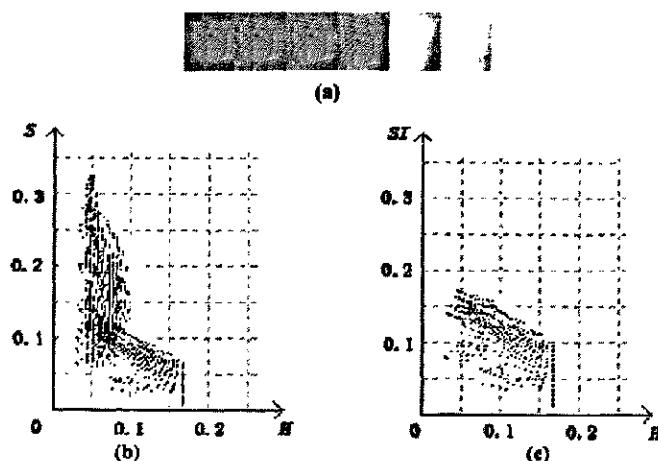
本文在建立肤色模型前，通过不同光亮度下肤色点分布的实验证明了上述理论推导的正确性。图 2.4 表示了同一块皮肤样本在不同光亮度下分别在 HS 平面和 H_SI 平面上的聚类结果，图中各坐标分量都做了归一化处理。从图中可以看到，在两个平面内，肤色点 H 的分布范围是一致的，为 $H \in [0.03, 0.17]$ ，在 H_SI 平面上 SI 的范围是 $SI \in [0.03, 0.18]$ ，而 HS 平面上 S 的取值范围扩展到 $S \in [0.02, 0.33]$ 。实验中还选用了其它多个不同人脸的肤色样本，在光照变化条件下，都有类似的结论，即在光亮度变化的条件下 H_SI_I 空间比 HSI 空间具有更好的聚类效果。

进行 H_SI_I 空间的肤色模型训练时，本文选取了不同光亮度、不同背景、不同年龄等因素的皮肤样本共 40 块，通过统计分析，给出一个阈值分割的肤色模型：

$$H \in [0.00, 0.17] \text{ 且 } SI \in [0.02, 0.37]$$

或是

$$H \in [0.97, 0.99] \text{ 且 } SI \in [0.02, 0.05]$$

图 2.4 不同亮度下肤色样本在 H_S 平面和 H_SI 平面聚类结果的对比

(a) 同一肤色样本在不同光强下的图像

(b) 在 H_S 平面内肤色点的分布(c) 在 H_SI 平面内肤色点的分布

2.3.4 基于 H_SI_I 肤色模型的肤色分割实验

利用上面给出的肤色模型，对不同光强下同一幅彩色图像进行二值化操作，实验结果如图 2.5 所示。

通过实验可以看到，图像亮度变化范围虽然较大，但肤色分割的结果大致相同，肤色区域几乎都被检测出来。因此，实验证明了在 H_SI_I 空间建立的肤色模型是有效可行的，它对光强的变化具有强鲁棒性。

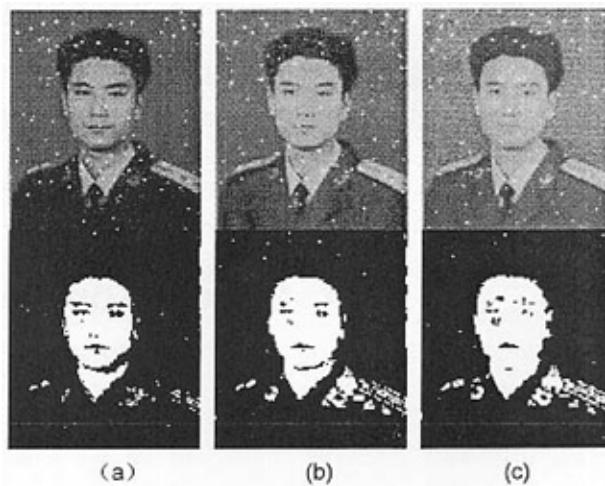


图 2.5 不同亮度下同一幅图片的肤色分割结果

(a) 亮度较低的图像及其检测结果

(b) 正常亮度的图像及其检测结果

(c) 亮度较高的图像及其检测结果

2.4 基于模糊聚类算法的肤色再分割

采用肤色模型进行肤色分割，通常只利用到模型分布的先验信息。因此，通用的肤色模型虽然具有普适性，但肤色分割的结果在保留了肤色像素的同时，常会保留下许多背景点。考虑到一幅特定的彩色图像中肤色像素点在颜色空间分布具有一定的稳定性，而其它点分布相对散乱，本文在肤色模型进行肤色初分割后，进一步采用模糊聚类的方法排除掉伪肤色干扰点，从而提高肤色分割的准确性。

2.4.1 模糊 C 均值算法^[40]

动态聚类方法是一种普遍采用的方法，它有以下三个要点：

- (1) 选择某种距离量度作为样本间的相似度量；
- (2) 确定某个评价聚类结果质量的准则函数；
- (3) 给定某个初始分类，然后用迭代算法找出使准则函数取极值的最好聚类结果。

C 聚类算法是一种基本的动态聚类方法，其目的是把 n 个样本划分到 c 个类别中的一个中，使得各个样本与其所在类均值的误差平方和最小。本节中采用的模糊 C 均值算法则是对 C 聚类算法的一种改进，它将硬分类变为模糊分类，使得聚类结果对预先确定的聚类数目不敏感，具有更好的鲁棒性。

设 $\{x_i, i=1, 2, \dots, n\}$ 是 n 个样本组成的样本集合，c 为预定的类别数目， $\{m_j, j=1, 2, \dots, c\}$ 为每个聚类的中心， $\mu_j(x_i)$ 是第 i 个样本对第 j 类的隶属度函数。

用隶属度函数定义的聚类损失函数可写为：

$$J_f = \sum_{j=1}^c \sum_{i=1}^n [\mu_j(x_i)]^b \|x_i - m_j\|^2 \quad (2.19)$$

其中， $b > 1$ 是一个可以控制聚类结果的模糊程度的常数。

在模糊 c 均值算法中，一个样本对各聚类的隶属度之和为 1，即：

$$\sum_{j=1}^c \mu_j(x_i) = 1 \quad (2.20)$$

式中 $i = 1, 2, \dots, n$ 。

在(2.20)式的约束下求(2.19)式的极小值，可得：

$$m_j = \frac{\sum_{i=1}^n [\mu_j(x_i)]^b x_i}{\sum_{i=1}^n [\mu_j(x_i)]^b} \quad (2.21)$$

式中 $j = 1, 2, \dots, c$ 。

$$\mu_j(x_i) = \frac{(1/\|x_i - m_j\|^2)^{1/(b-1)}}{\sum_{k=1}^c (1/\|x_i - m_k\|^2)^{1/(b-1)}} \quad (2.22)$$

式中 $i=1, 2, \dots, n$, 而 $j=1, 2, \dots, c$ 。

根据(2.21)、(2.22)式, 可通过迭代的方式使聚类中心由初始位置向目标结果位置靠近。迭代具体步骤如下:

- (1) 设定聚类数目 c 和参数 b ;
- (2) 初始化各聚类中心 m_i ;
- (3) 用当前的聚类中心根据(2.22)式计算隶属度函数 $\mu_j(x_i)$;
- (4) 用当前的隶属度函数根据(2.21)式更新各聚类中心;
- (5) 若算法收敛则停止, 否则返回步骤 (3)。

2.4.2 肤色聚类区域的判决

模糊聚类后, 需要通过一些判决以确定哪些聚类区域是肤色聚类区域。为了获取判决条件, 在进行肤色模型训练时, 对颜色空间各位置出现肤色点的次数进行统计, 然后做归一化处理, 得到在 H_SI 平面上肤色分布区各颜色点密度分布的概率统计值。利用颜色空间中各位置点的密度概率统计结果, 设定聚类区域的判决法则:

判决一: 先验信息判决。若聚类后, 某聚类中心及其大部分点均出现在肤色点分布的高密度概率区, 那么可以认为该聚类区域就是肤色区域, 肤色分割维持初始分割结果, 不需要二次分割;

判决二: 聚类中各点概率方差判决。对于一幅具体的图像, 其肤色区域在颜色空间各位置出现的概率较为均匀, 背景干扰点的概率值相对凌乱, 因此我们利用每个聚类区域中各点出现的概率方差作为判决法则。若某聚类的方差大于给定阈值, 则判定为背景干扰区域, 反之, 方差小于给定阈值的聚类区域, 则判定为肤色区域。

2.4.3 基于模糊聚类的二次肤色分割实验

实验选用了从 Internet 上下载和数码相机拍摄的彩色图像多幅, 每幅图像中有一个或一个以上的人脸, 且图中人脸大小不一、背景复杂、光照变化大。图 2.6 给出了模糊聚类后, 肤色分割改进较为明显的部分检测图像。实验结果表明采用模糊聚类的二次分割能有效提高肤色分割的准确性。

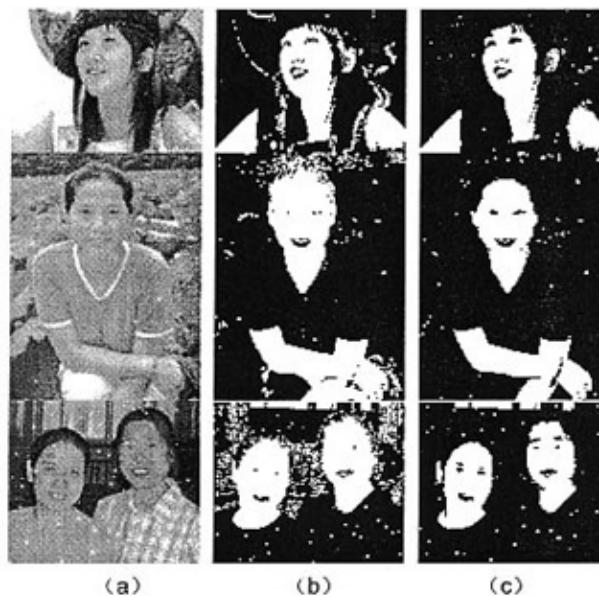


图 2.6 肤色模型初次分割和模糊聚类后二次分割的对比

- (a) 原彩色图像
- (b) 肤色模型初次分割的结果
- (c) 模糊聚类后二次分割的结果

2.5 数学形态学处理

数学形态学的基本思想是用具有一定形态的结构元素去量度和提取图像中的对应形状以达到对图像分析和识别的目的。数学形态学的数学基础和所用语言是集合论。数学形态学的应用可以简化图像数据，保持它们基本的形状特征，并除去不相干的结构。

数学形态学的基本运算有四个：膨胀、腐蚀、开启和闭合^[4]。

2.5.1 膨胀

膨胀的运算符为 \oplus ，图像集合 A 用结构元素 B 来膨胀，记做 $A \oplus B$ ，其定义为：

$$A \oplus B = \{x | [(\hat{B})_x \cap A] \neq \emptyset\} \quad (2.23)$$

其中， \hat{B} 表示 B 的映射，式(2.16)表明用 B 膨胀 A 的过程如下：先对 B 做关于原点的映射，再将其映象平移 x ，当 A 与 B 映象的交集不为空时， B 的原点就是膨胀集合的象素。也就是说，用 B 来膨胀 A 得到的集合是 \hat{B} 的位移与 A 至少有一个非零元素相交时 B 的原点位置的集合。因而，(2.23)式也可以表示为：

$$A \oplus B = \{x | [(\hat{B})_x \cap A] \subseteq A\} \quad (2.24)$$

如果将 B 看成一个卷积模板，膨胀就是对 B 作关于原点的映射，再将映象连续地在 A 上移动而实现的。

2.5.2 腐蚀

腐蚀的运算符为 \ominus , A 用 B 来腐蚀记做 $A \ominus B$, 其定义为:

$$A \ominus B = \{x | (B)_x \subseteq A\} \quad (2.25)$$

上式表明 B 腐蚀 A 的集合是 B 完全包含在 A 中时 B 的原点位置的集合。

以上讨论中都假设原点是位于结构元素之中的, 对于膨胀运算总有 $A \subseteq A \oplus B$, 对于腐蚀运算总有 $A \ominus B \subseteq A$ 。当原点不包含在结构元素中时, 相应的结果会有所不同, 对于膨胀运算总有 $A \not\subseteq A \oplus B$, 但对于腐蚀运算则有两种可能, 一种是 $A \ominus B \subseteq A$, 另一种是 $A \ominus B \not\subseteq A$ 。

2.5.3 开启和闭合

膨胀和腐蚀并不是互为逆运算, 它们可以级连结合使用。使用同一个结构元素对图像进行先腐蚀后膨胀的运算为开启运算, 进行先膨胀后腐蚀的运算为闭合运算。它们都是形态学中的重要运算。

开启运算可以把比结构元素小的突刺滤掉, 切断细长搭接而起到分离的作用。其运算符为“◦”, A 用 B 来开启记为 $A◦B$, 定义如下:

$$A◦B = (A \ominus B) \oplus B \quad (2.26)$$

闭合运算可以把比结构元素小的缺口或孔填充上, 搭接短的间断而起到连通的作用。其运算符为“•”, A 用 B 来闭合记为 $A•B$, 定义如下:

$$A•B = (A \oplus B) \ominus B \quad (2.27)$$

开启和闭合运算不受原点位置的影响, 无论原点是否包含在结构元素中, 开启和闭合运算的结果都是一定的。

2.5.4 数学形态学处理肤色区域

经颜色空间转换的二值化图像中不可避免地出现噪声, 本文采用数学形态学算子去除分割后的噪声, 并使得图像中重要的区域(肤色区域)仍能保持连贯的轮廓。

文中所使用的结构元素如(2.28)式所示:

$$B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (2.28)$$

数学形态学处理的具体步骤如下:

- (1) 使用(2.28)中定义的结构元素对二值化图像进行一次开运算和一次闭运算, 用于消除肤色与背景的连接区域以及一些小的噪声块;
- (2) 使用相同的结构元素进行 3 至 5 次腐蚀运算, 以尽可能消除肤色区域内较大的黑色杂块, 使之成为连通区域;

(3) 再使用相同的结构元素进行 3 至 5 次膨胀运算，使得腐蚀后过大的肤色区域恢复到腐蚀前的近似大小，同时平滑了各区域的边界。

经过形态学处理的效果如图 2.7 所示，从处理结果可以看到形态学处理可以过滤由于噪声或其他原因引起的一些较小的不可能为人脸的类肤色区域，减少候选区域，提高了检测速度。同时，处理中填充了肤色区域内较小的孔洞，从而为后续人脸器官的判定降低了误判的可能。



图 2.7 形态学处理前后的二值化图像

- (a) 形态学处理前的二值化图像
- (b) 形态学处理后的二值化图像

2.6 候选人脸区域的确定

形态学滤波后得到许多不连通的类肤色区域，由于背景复杂，还需要对这些类肤色区域进行一些必要的处理。如将不连通的区域进行区域归并，将裸露的手臂或腿部等非人脸区域予以删除等。

2.6.1 轮廓提取

二值图像的轮廓提取算法比较简单，就是掏空内部点，即如果原图中有一个点为黑，且它的八个相邻点都是黑色时，表明该点为内部点，则需将该点删除。联系到形态学处理的内容，实际上就相当于用九个点的结构元素对原图进行腐蚀，再用原图像减去腐蚀图像得到。

2.6.2 外接矩形的算法

通过形态学滤波和轮廓提取，各个类肤色区域的边界具有了连通性和闭合性，这样就

可以寻找到每一段边界的外接矩形，具体算法^[42]如下：

- (1) 给每一个边界点增加一个标志域：“边界跟踪的起始点”，“已检测过的边界点”和“未检测过的边界点”；
- (2) 在图像中的边界点中从上到下，从左到右寻找一个“未检测过的边界点”，将其标志域改为“边界跟踪的起始点”，为这一新的跟踪建立一个外接矩形，初始化其上下左右四个边界，进入(3)处理，若找不到转到(7)；
- (3) 从这个“边界跟踪的起始点”开始，按照四连通原则，在其上下左右的四个邻接点中寻找“未检测过的边界点”，因为不存在孤立的边界点，所以这个点一定可以找到，将其标志域改为“已检测过的边界点”，并将其作为当前操作点，进入(4)处理；
- (4) 从当前操作点开始，按照四连通原则，在其上下左右的四个邻接点中寻找“未检测过的边界点”或“边界跟踪的起始点”，如果找到这样的点，则进入(5)处理，否则转到(6)；
- (5) 如果找到的点是“未检测过的边界点”，则将其标志域改为“已检测过的边界点”，将其作为当前操作点，再转回(4)。如果找到的点是“边界跟踪的起始点”，要判断这个点与当前操作点是否“互相找到对方”，如果是，仍以原来那个当前操作点开始，转回(4)，如果不是，将“边界跟踪的起始点”的标志域改为“已检测过的边界点”，并进入(6)处理；
- (6) 将(2)中建立的外接矩形的四个边界确定下来，并转回(2)寻找新的候选人脸区域；
- (7) 整个二值化图像中的区域寻找完毕，得到一系列外接矩形。

在上述算法中，每找到一个新的“未检测过的边界点”，就判断其坐标值能否改变外接矩形四个边的位置，这样所有的边界点找完，外接矩形也就确定下来了。

2.6.3 矩形的归并

基于边界的外接矩形算法快速有效，但是它是从局部特性出发，没有考虑图像整体因素，这样会造成得到的初始矩形序列中的矩形相互覆盖、重叠等现象，因此有必要对这一初始矩形序列进行归并处理。具体的归并规则如下：

- (1) 当一个矩形完全包围另一个矩形时，抛弃内部的矩形；
- (2) 当两个矩形的边界接触或两个矩形出现重叠时，用这两个矩形公共的外接矩形来代替这两个矩形。

2.6.4 几何筛选

考虑到可能存在裸露的手臂或腿部等非人脸区域被检测出来的情况，需要由一些先验知识对这些候选区域进行初步的判别，剔除明显不是人脸的区域。在本文中，根据人脸的

几何特征采用了如下判定规则：

- (1) 如果候选区域面积太小，则删除该区域（人脸在图像中应占有一定的比例）。
- (2) 通过外接矩形的长宽比，删除一些非人脸区域。考虑到脖子区域对肤色区域合并的影响，此处定义长宽比的范围为[0.6, 2.8]。

除了以上方法外，还可以根据欧拉数、质心、椭圆近似的长短轴比例等方法剔除非人脸区域。实验中发现，只需使用上述方法当中的两种方法来进行筛选，就能够过滤掉大部分的非人脸区域。

通过形态学滤波、区域归并和几何筛选，从肤色分割的图像中得到一组候选人脸区域。图 2.8 是经区域归并和几何筛选后得到的候选人脸区域。

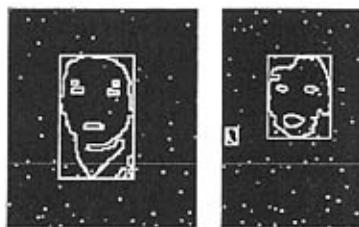


图 2.8 区域归并和几何筛选后得到的候选人脸区域

2.7 本章小结

本章主要从颜色空间的选择、肤色模型的建立、肤色分割算法的改进以及类肤色区域的预处理方面进行了详细的分析，并完成了以下工作：

- (1) 提出一种改进的 H_SI_I 颜色空间，实验证明，在该空间建立的肤色模型对光亮度的变化具有较强的鲁棒性；
- (2) 采用模糊聚类算法进行肤色的二次分割，并通过颜色空间中各点的密度概率统计结果设定聚类区域的判决法则。经过二次分割有效地提高了肤色分割的准确性；
- (3) 对类肤色区域进行了预处理，通过形态学滤波、区域归并和人脸几何特征规则，去掉了大部分背景噪声和非人脸区域，筛选出候选人脸区域。本章肤色分割算法的流程如图 2.9 所示。

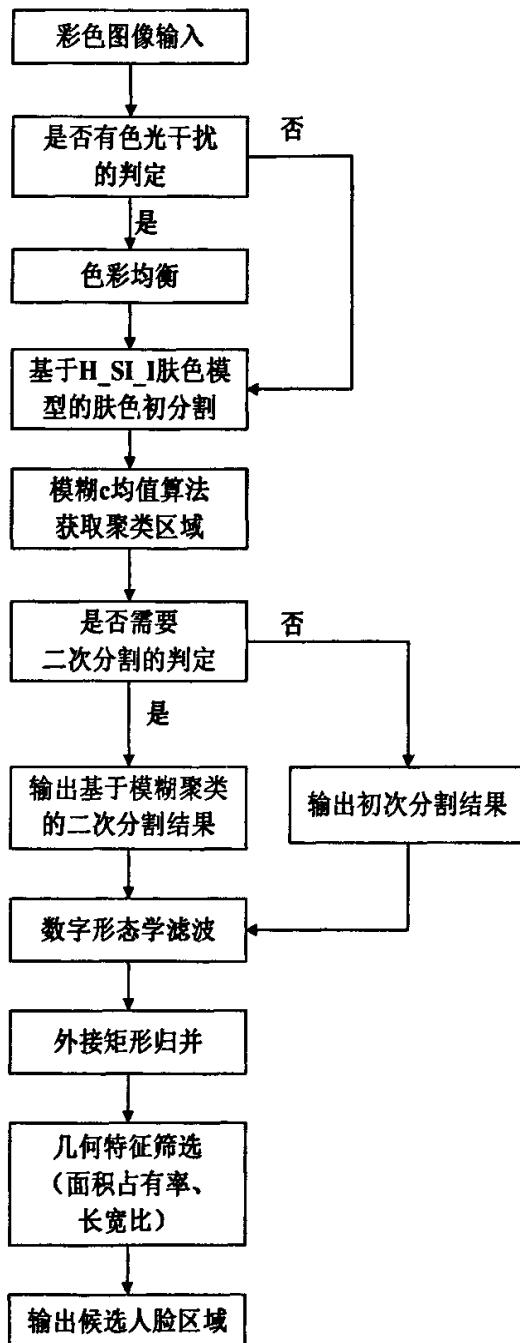


图 2.9 肤色分割算法流程图

第三章 特征提取和人脸定位

在简单背景下，仅利用肤色分割就可以将人脸直接检测出来，但在背景的类型和特征不受约束的复杂环境中，还必须利用其它的人脸特征才能够实现人脸检测的完整算法。

在诸如眉毛、眼睛、鼻子、嘴巴等人的脸部特征当中，双眼是最为稳定的特征，如果能准确地定位出双眼的位置，将会大大简化人脸检测的难度。因此，本章着重讨论人眼的提取算法，其流程如图 3.1 所示。

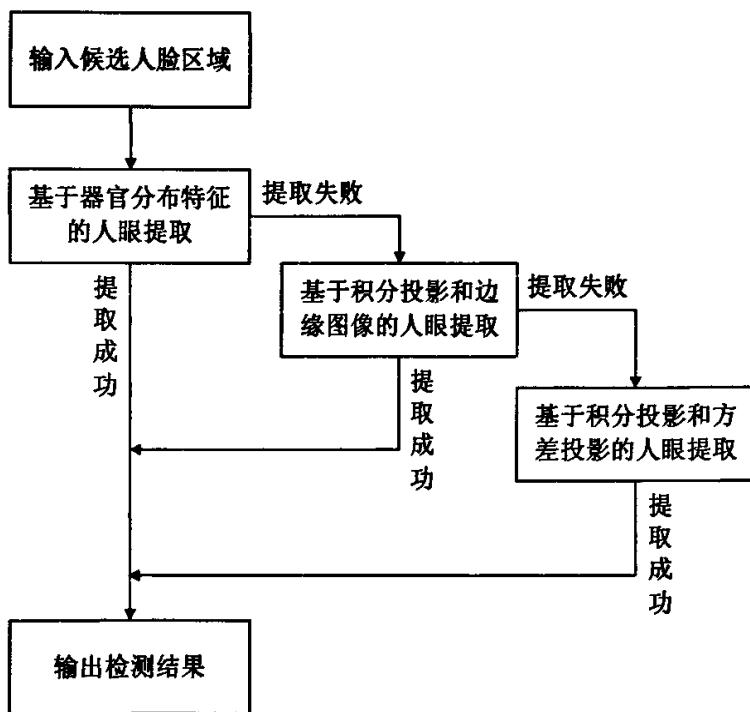


图 3.1 特征提取算法的流程图

算法分成三步来实现，它结合了器官分布特征、积分投影、边缘图像、方差投影等知识信息，能比较准确地提取出眼睛。在讨论时发现，算法中的三个步骤可独立成三种特征提取方法，因此本章中将特征提取算法的三个步骤用三种方法分别予以阐述。一旦双眼位置确定下来，即可根据“三停五眼”规则定位出人脸区域。

3.1 基于器官分布特征的人脸定位

候选人脸区域都包含一定数目的孔洞，若是真的人脸区域，这些孔洞中的一部分对应着眉毛、眼睛、鼻子、嘴巴等人脸器官的位置，根据人脸器官几何分布规律，可对孔洞进行检测，从而在获得双眼对的同时排除掉候选人脸中的非人脸区域。

3.1.1 眼睛提取

对每个候选人脸区上半区域进行搜索，将获得的孔洞两两匹配构成所有可能的眼睛对。然后根据人脸的几何特征对候选眼睛对进行判定，其判定规则如下：

- (1) 根据双眼的间距在人脸图像宽度中占一定比例，要求配对孔洞外接矩形的中心点之间的连线 D_b 大于外接矩形宽度的 $1/6$ ，其中 D_b 的表达式为：

$$D_b = \sqrt{(y_l - y_r)^2 + (x_l - x_r)^2} \quad (3.1)$$

- (2) 孔洞连线与水平轴的交角 θ_b 也应在一定的角度范围内，要求 $\theta_b < 20^\circ$ ，其中 θ_b 的表达式为：

$$\theta_b = \arctan \left| \frac{y_l - y_r}{x_l - x_r} \right| \quad (3.2)$$

经过上面两个规则，可以过滤掉很多不可能的候选眼睛对，有的已得到正确的结果。但是由于眉毛的影响，仍需做进一步确认。因此，在上面判定的基础上，增加一些判定规则：

- (3) 为了排除眉眼对，规定若在同一候选人脸区域中，有合乎(1)(2)判定的多个配对，则选择 θ_b 最小的两组配对。
- (4) 根据眉毛在眼睛上方的先验知识，在判定(3)的基础上，选择下方的一组配对，以排除眉毛对。

3.1.2 嘴唇提取

得到眼睛对后，可以进一步定位嘴唇。如图 3.2 所示，C、D 分别为 AE、BF 的中点，设 $AB=d$ ，则 $AE=1.4d$ 。若 A、B 为眼睛对，则嘴在 CDEF 内存在。通过搜索 CDEF 中是否存在孔洞来判定嘴的位置。

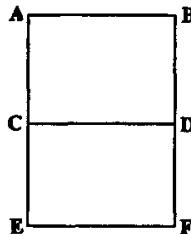


图 3.2 嘴区与双眼关系图

3.1.3 脸部定位

确定了双眼的位置，整个人脸的比例及位置都可以根据双眼的位置以及双眼的间距，按照“三停五眼”的规则输出人脸区域^[43]。即设眼睛对之间的距离为 d ，以两眼的中心为基准，取向上的 $0.5d$ 个象素高度，向下的 $1.5d$ 个象素长度，向左的 d 个象素长度，向右

的 d 个象素长度，输出一个正方形区域，正方形的上下边与两眼的连线平行。

综上所述，本节方法的流程图如图 3.3 所示。

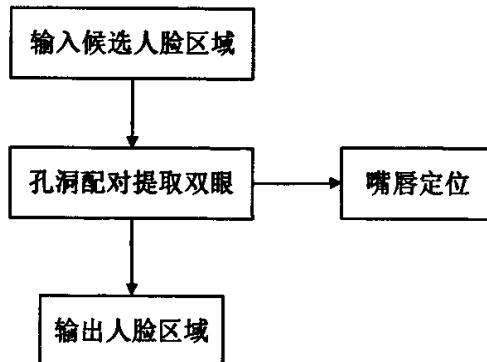


图 3.3 基于器官分布特征的人脸定位方法流程图

3.2 基于积分投影和边缘图像的人脸定位

若肤色分割后的二值图像没能在双眼区域形成相应的孔洞，则无法通过孔洞匹配来实现人眼检测。为此，本节主要从灰度信息入手，分析特征提取方法。

在 3.1 节方法未能检测出人脸的前提下，将彩色图像转换为灰度图像，可以发现灰度图中脸部特征处（如鼻子、眼睛、眉毛等）的灰度值要低于脸部其他区域。因此考虑通过对正面人脸进行灰度积分投影分析，粗略定位特征部位的位置。在本节方法中，首先对灰度图像使用水平积分投影估计人眼的水平位置，然后结合边缘图像和加窗处理，给出人眼的确定位置，并最终输出人脸定位结果。

3.2.1 主轴旋转

常见的图像中大部分人脸都是垂直的。但是，有部分人脸可能存在旋转倾斜。为了提高积分投影定位的准确度，必须对倾斜的人脸进行方向矫正。

在二维的几何形状中，主成分分析法（PCA）能用来检测形状空间的主轴方向^[41]，因此在本文中通过采用 PCA 方法获得候选人脸区域的主轴方向，然后计算水平主轴方向与水平方向的夹角得到倾斜的角度，再利用该角度实现方向矫正。

设 $f(x, y)$ 为候选人脸区域图像，由于该图像是二值化图像， $f(x, y)$ 只有 0 和 1 两种取值，因此很容易得到各点 (x, y) 之间的协方差矩阵 C_{cov} ， C_{cov} 按下面的公式计算：

$$C_{cov} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix} \quad (3.3)$$

C_{cov} 是个实对称矩阵，所以总能找到一组特征值 (λ_1, λ_2) 和与之对应的正交特征矢量 (V_1, V_2) ，其特征矢量的方向即为候选人脸图像的主轴方向。

设计算得到的特征矢量 V_1 的方向为水平主轴方向，那么就在 V_1 的方向上取两个点

(x_1, y_1) 和 (x_2, y_2) , 则图像的倾斜角度 θ 为:

$$\theta = \arctan \left| \frac{y_1 - y_2}{x_1 - x_2} \right| \quad (3.4)$$

根据(3.4)式, 将候选人脸图像旋转 $-\theta$, 即可实现人脸的方向矫正。

3.2.2 彩色图像灰度化

水平积分投影前, 首先要将彩色图像转换为灰度图像。一般情况下彩色图像每个象素用三个字节表示, 每个字节对应着 R、G、B 分量的亮度, 转换后的灰度图像的一个象素用一个字节表示该点的灰度值, 它的值在 0~255 之间, 数值越大, 该点越白, 即越亮, 越小则越黑。转换关系^[44]为:

$$I(x, y) = 0.11R(x, y) + 0.59G(x, y) + 0.3B(x, y) \quad (3.5)$$

其中 $I(x, y)$ 为转换后灰度图像在 (x, y) 点处的灰度值。观察(3.5)式, 发现其中绿色分量所占的比重最大, 因此本文在灰度化处理中直接使用 G 值作为转换后的灰度。

3.2.3 水平积分投影估计人眼的水平位置

设候选人脸区域大小为 $M \times N$, $I(x, y)$ 表示像素点 (x, y) 处的灰度值, 则图像的积分投影函数为:

$$h(y) = \frac{1}{M} \sum_{x=1}^M I(x, y) \quad (3.6)$$

$$v(x) = \frac{1}{N} \sum_{y=1}^N I(x, y) \quad (3.7)$$

其中 $h(y)$ 、 $v(x)$ 分别为水平投影积分函数和垂直投影积分函数。在本节中只利用到水平积分投影函数。

为了去除噪声影响, 本文使用一维 Gauss 模板对水平投影图像进行平滑预处理, 其具体步骤如下:

(1) 选择一维 Gauss 模板:

$$g(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left| -\frac{(y-y_0)^2}{2\sigma^2} \right| \quad (3.8)$$

其中 $y=1, 2, \dots, L_g$, L_g 为 Gauss 模板的长度, $y_0 = L_g/2$;

(2) 用 Gauss 模板对水平积分投影图像进行卷积运算:

$$h_0(y) = h(y) \otimes g(y) \quad (3.9)$$

其中 $h_0(y)$ 的长度为 $L_g + N$;

(3) 令 $h_s(y)$ 等于 $h_0(y)$ 的前 N 个数据构成的向量加上后 N 个数据构成的向量再除以 2, 则 $h_s(y)$ 即为平滑后的水平积分投影图, 即:

$$h_s(y) = \frac{1}{2} [h_0(1:N) + h_0(L_g : L_g + N)] \quad (3.10)$$

图 3.4(a)表示候选人脸区域灰度化后的水平积分投影图像, 图 3.4(b)表示经高斯函数平滑后的水平积分投影图像。观察投影图可发现, 在候选脸上半区通常有两个极小值点, 分别对应人的眉毛和眼睛部位, 而眉毛位于眼睛的上面, 因此可将投影图中上半区的第二个谷点对应的 y 值确定为人眼水平位置坐标。如图 3.4(c)所示。

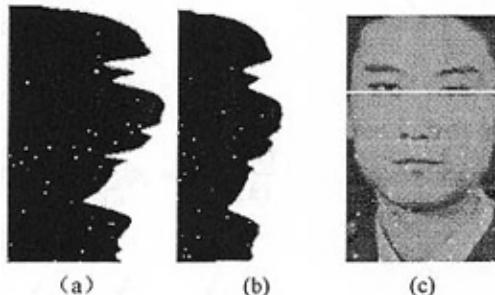


图 3.4 水平积分投影估计人眼水平位置的过程
 (a) 水平积分投影图
 (b) 高斯平滑后的水平积分投影图
 (c) 人眼水平位置的标定

3.2.4 灰度图像二值化

为了获得较好的边缘提取效果, 在边缘检测之前先对灰度图像进行二值化处理, 使人脸眼睛区域的对比度增强, 双眼更加突出。常用的二值化方法有直方图阈值法、最大类间方差阈值法等。对于待检测的图像, 其背景有时会出现各种复杂的情况, 如噪声干扰、背景灰度不均匀等, 并且存在光照变化的影响, 所以运用直方图阈值法将会使分割变得不可靠; 一维最大熵阈值分割对噪声很敏感, 二维最大熵阈值分割则使问题变得更复杂, 且计算量大为增加; 模糊阈值分割法的缺点是窗口大小和资格函数的选取对图像分割影响很大。对比起来, 由 Osta 提出的最大类间方差阈值法^{[45][46]}是建立在最小二乘原理基础上的聚类分割, 其方法严密、原理简单, 是一种公认有效的传统图像分割方法。它的基本思路是将直方图在某个阈值处分割成两组, 当被分成的两组的方差为最大时, 决定阈值, 因此, 它有动态选取阈值实现目标自动提取的功能。

最大类间方差自动阈值法的实现原理是: 设图像有 L 个灰度级, 其中灰度值为 i 的象素数为 n_i , 则总的象素数为:

$$N = \sum_{i=0}^L n_i \quad (3.11)$$

各灰度值出现的概率为:

$$p_i = \frac{n_i}{N} \quad (3.12)$$

设以灰度 k 为门限将图像分割成两个区域，灰度为 $1 \sim k$ 的象素和灰度为 $(k+1) \sim L$ 的象素分别构成区域 A 和 B。显然，区域 A 和 B 的概率分别为：

$$\omega_A = \sum_{i=1}^k p_i = \omega(k) \quad (3.13)$$

$$\omega_B = \sum_{i=k+1}^L p_i = 1 - \omega(k) \quad (3.14)$$

区域 A 和 B 的平均灰度分别为：

$$\mu_A = \frac{1}{\omega_A} \sum_{i=1}^k i p_i = \frac{\mu(k)}{\omega(k)} \quad (3.15)$$

$$\mu_B = \frac{1}{\omega_B} \sum_{i=k+1}^L i p_i = \frac{\mu - \mu(k)}{1 - \omega(k)} \quad (3.16)$$

式中 μ 是全图的平均灰度，

$$\mu = \sum_{i=1}^L i p_i = \sum_{i=1}^k i p_i + \sum_{i=k+1}^L i p_i = \omega_A \mu_A + \omega_B \mu_B \quad (3.17)$$

两个区域总的方差为：

$$\sigma^2(k) = \omega_A (\mu_A - \mu)^2 + \omega_B (\mu_B - \mu)^2 = \frac{[\mu \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (3.18)$$

从 $1 \sim L$ 改变 k ，使 $\sigma^2(k) = \max$ 的 k 便是区域分割阈值。

通过灰度图像二值化处理，增强了图像对比度，为边缘提取获得较好的效果提供了保证。

3.2.5 边缘检测

在灰度图像二值化后，就可以进行边缘检测了。常用的边缘检测算子^{[41][47]}有 Robert 算子、Sobel 算子、Prewitt 算子、LoG 算子、Canny 算子等。Robert 算子对具有陡峭的低噪声图像的响应度最好；Sobel 算子和 Prewitt 算子计算量小，对灰度渐变和噪声较多的图像处理较好，但它们只能探测到灰度垂直和水平方向变化的边缘，在眼睛的边缘提取中效果比较差；LoG 算子是一种二阶边缘检测方法，它将高斯平滑滤波和拉普拉斯锐化滤波器结合起来，先平滑掉噪声，再进行边缘检测，提取的边缘比其它算子完整，位置也比较准确，但是易出现伪边缘效应，且不能给出灰度变化的方向信息；Canny 算子是利用高斯函数的一阶微分提取边缘，且进行了非极大值抑制和形态学操作，因此能在噪声抑制和边缘检测之间取得较好的平衡，提取的边缘十分完整，连续性要优于其它算子。

为了有效抑制噪声，完美地提取眼睛边缘，本文选取 Canny 算子进行边缘检测。Canny 边缘检测的具体步骤如下：

(1) 用高斯滤波器平滑图像。高斯平滑函数为：

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3.19)$$

若用 $I'(x, y)$ 表示二值化后的灰度图像，则 $G(x, y)$ 对 $I'(x, y)$ 的平滑可表示为：

$$F(x, y) = G(x, y) \otimes I'(x, y) \quad (3.20)$$

在实际的程序中是用一维的高斯滤波器分别在 x, y 方向上滤波后得到平滑图像：

- (2) 用一阶偏导的有限差分来计算边缘的强度和方向。Canny 算子可由 $\nabla G(x, y) \otimes F(x, y)$ 计算得到边缘强度和方向，而实际处理时，是把 $\nabla G(x, y)$ 的二维卷积模板分解为一维滤波器：

$$\frac{\partial G(x, y)}{\partial x} = kxe^{-\frac{x^2}{2\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} = g_1(x)g_2(y) \quad (3.21)$$

$$\frac{\partial G(x, y)}{\partial y} = ky e^{-\frac{x^2}{2\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} = g_1(y)g_2(x) \quad (3.22)$$

其中

$$g_1(x) = \sqrt{k} xe^{-\frac{x^2}{2\sigma^2}}; \quad g_1(y) = \sqrt{k} ye^{-\frac{y^2}{2\sigma^2}} \quad (3.23)$$

$$g_1(x) = xg_2(x); \quad g_1(y) = yg_2(y) \quad (3.24)$$

将上述两个模板分别与 $F(x, y)$ 进行卷积，得到：

$$E_x = \frac{\partial G(x, y)}{\partial x} \otimes F(x, y) \quad (3.25)$$

$$E_y = \frac{\partial G(x, y)}{\partial y} \otimes F(x, y) \quad (3.26)$$

令

$$A(x, y) = \sqrt{E_x^2 + E_y^2} \quad (3.27)$$

$$\alpha(x, y) = \arctan \frac{E_y(x, y)}{E_x(x, y)} \quad (3.28)$$

则 A 为边缘强度， $\alpha(x, y)$ 为垂直于边缘的方向；

- (3) 对边缘强度进行非极大值抑制。若只得到全局的梯度不足以确定边缘，必须保留局部梯度最大的点，而抑制非极大值点。具体的判决方法就是：若象素点 (x, y) 的边缘强度大于沿梯度方向的两个相邻象素的边缘强度，则认为该点为边缘点，否则为非边缘点；
- (4) 用双阈值算法检测和连接边缘。为了减少假边缘数量，采用双阈值算法进行处理，即对非极大值抑制后的图像作用两个阈值。设两个阈值分别为 τ_1 和 τ_2 ，且 $2\tau_1 \approx \tau_2$ ，这样经双阈值处理后可以得到两个阈值边缘图像 $N_1(x, y)$ 和 $N_2(x, y)$ 。由于 $N_2(x, y)$ 使用高阈值得到，因而含有很少的假边缘，但有间断，不闭合。双

阈值法要在 $N_2(x, y)$ 中把边缘连接成轮廓，当到达轮廓的端点时，该算法就在 $N_1(x, y)$ 的八个邻点位置寻找可以连接到轮廓上的边缘，这样，算法不断地在 $N_1(x, y)$ 中收集边缘直到将 $N_2(x, y)$ 连接起来为止。

本文在检测边缘时将阈值 τ_2 取得较高，使得脸部的边缘点仅由眉、眼等特征位置的强边缘组成，而不会包括鼻的轮廓产生的弱边缘。图 3.5(a) 为提取边缘后的图像，边缘像素点用白色表示。利用 Canny 算子提取的眼睛边缘，虽然也不能完全避免连接线的断续，但并不影响提取边缘的完整性，提取结果还是令人满意的。

3.2.6 加窗处理定位眼睛

边缘检测后，为实现眼睛的快速定位，本文在边缘图像的候选人脸区域上加一个大小为 $K \times L$ 的窗， K 的取值应大于瞳孔的宽度，小于两眼的水平间距， L 的取值应大于眼睑的高度，小于眉眼的垂直距离。将此窗放在候选脸区边缘图的左端，高度为预估人眼水平位置对应的 y 值，然后在候选人脸区内沿人眼水平位置从左往右移动窗口，移动步长为 K 。每移动一次记录下窗口中白色像素点的数目，移动结束后找出窗口中白色像素点数目出现最大值和次最大值的两个位置，若这两个位置的间距大于候选人脸区域宽度的 $1/5$ 倍，即为两眼睛所在的位置。图 3.5 (b) 中灰色块表示窗口移动结束后找到的两眼位置。

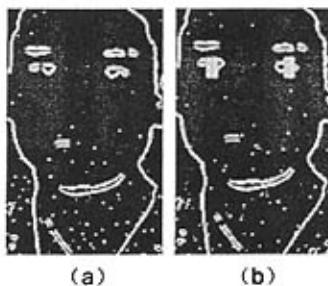


图 3.5 边缘检测和加窗处理定位人眼的过程

(a) Canny 算子检测后的边缘图像

(b) 加窗标定双眼的位置

双眼位置一旦确定，即可根据“三停五眼”规则输出检测到的人脸区域。

综上所述，本节方法的流程如图 3.6 所示。

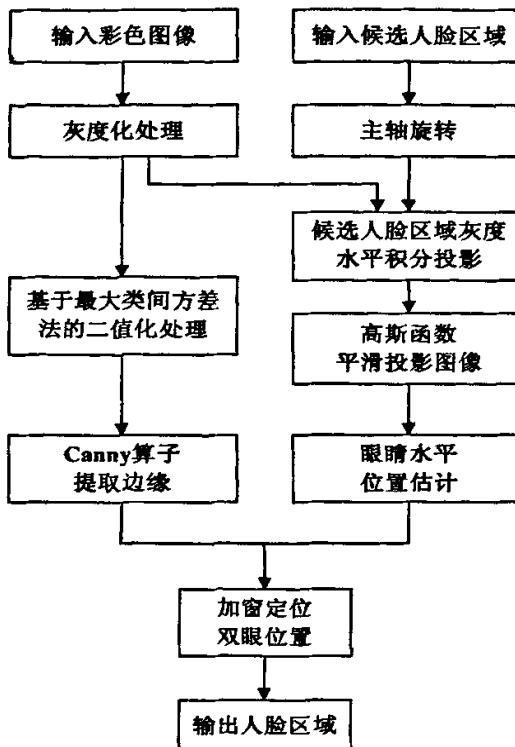


图 3.6 基于积分投影和边缘图像的人脸定位方法流程图

3.3 基于积分投影和差分投影的人脸定位

若 3.2 节算法还是不能正确地定位人眼，则在使用水平积分投影估计出人眼水平位置的基础上，再在粗估的水平位置附近进行垂直方差投影和水平方差投影，以此来得到两只眼睛的确定位置，并最终输出人脸定位结果。

3.3.1 方差投影

对正面人脸图像进行水平积分投影和垂直积分投影后，通过对投影曲线的分析可知，人脸的主要器官眼睛、鼻子、嘴巴的位置分别对应曲线的某个谷值或峰值区域，由此可以粗略地检测出人眼、鼻子和嘴巴的位置，如图 3.4 所示。但是在有些情况下，积分投影具有一定的局限性。如对一幅具有三种灰度值的图像，应用垂直积分投影如图 3.7 所示，垂直积分投影 $v(x)$ 是一条直线，体现不了图像内在的灰度值间的变化。为此，定义水平方差投影和垂直方差投影。设候选人脸区域大小为 $M \times N$ ， $I(x, y)$ 表示像素点 (x, y) 处的灰度值，则水平方差投影函数和垂直方差投影函数为^[48]：

$$\sigma_h^2(y) = \frac{1}{M} \sum_{x=1}^M [I(x, y) - h(y)]^2 \quad (3.29)$$

$$\sigma_v^2(x) = \frac{1}{N} \sum_{i=1}^N [I(x, y_i) - v(x)]^2 \quad (3.30)$$

计算图 3.7 的垂直方差投影曲线 $\sigma_v^2(x)$, 则可看到方差投影很好地体现了图像内在的灰度变化, 避免了积分投影的局限性。

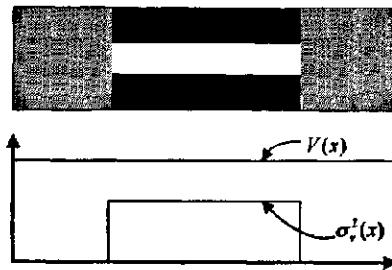


图 3.7 积分投影和方差投影的比较

灰度图像中, 人眼瞳孔的灰度值和眼白及上下眼睑的灰度值对比强烈, 因此本节采用方差投影来实现对人眼的准确定位。如图 3.8, 对左眼图像分别进行水平方差投影和垂直方差投影, 并对投影图像进行高斯平滑, 从投影结果可以看到, 由于瞳孔的灰度值和眼白及上下眼睑的灰度值对比强烈, 投影图在瞳孔边缘处形成明显的谷值区域, 所以瞳孔位置对应投影图像的峰值位置 P_1 和 P_2 。因此, 可以通过计算眼睛图像方差投影的极大值点实现对眼睛的准确定位。

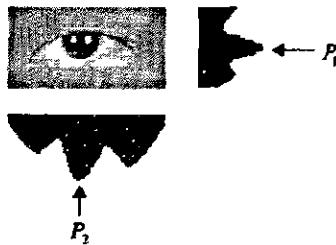


图 3.8 左眼方差投影图

本节算法开始部分和 3.2 节中的算法相同, 首先采用水平积分投影估计人眼的水平位置, 然后通过方差投影确定眼睛的位置。方差投影定位眼睛的具体步骤如下:

- (1) 设候选人脸区域大小为 $M \times N$, 在候选人脸区域取一个矩形, 矩形的高度是以粗估人眼的水平位置 y 为基准, 向上和向下都取 $\frac{N}{6}$ 个象素高度, 矩形宽度为 M 个象素;
- (2) 对(1)中所取矩形进行垂直方差投影, 并用高斯函数平滑投影图像。从投影结果可看到, 方差出现最大值和次最大值的位置即为两眼睛所在位置, 因此将这两个位置的横坐标记录下来, 即为两只眼睛对应的横坐标 x_l , x_r ;
- (3) 分别以 (x_l, y) , (x_r, y) 为中心取宽为 $\frac{M}{5}$, 高为 $\frac{N}{3}$ 的矩形;
- (4) 对(3)中的两个矩形各自都进行水平方差投影, 然后用高斯函数平滑投影图像。

从理论分析可知，投影中峰值对应的纵坐标应为眼睛的纵坐标，但若被投影的矩形内包括了眉毛，则会出现两个明显的峰值区域，此时得根据眉眼位置的几何关系，取纵坐标小的峰值点，将其纵坐标作为眼睛的纵坐标。

图 3.9(a)为垂直方差投影确定两只眼睛的横坐标的图像，图 3.9(b)为水平方差投影确定两只眼睛纵坐标的图像，在积分投影的基础上采用方差投影，对有一定倾斜的人脸图像仍能定位准确。

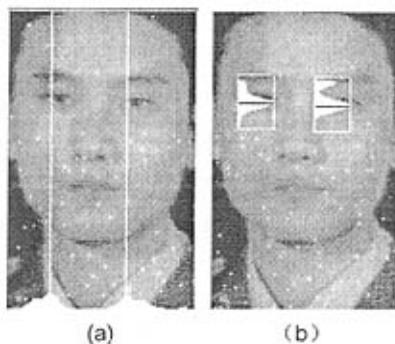


图 3.9 方差投影定位人眼的过程

- (a) 垂直方差投影确定眼睛横坐标
- (b) 水平方差投影确定眼睛纵坐标

综上所述，本节方法的流程图如 3.10 所示。

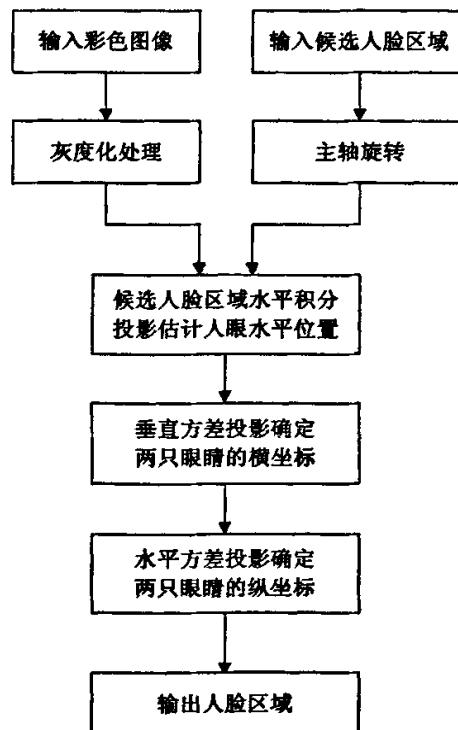


图 3.10 基于积分投影和方差投影的人脸定位方法流程图

第四章 实验结果及分析

4.1 实验测试结果

实验中的人脸图像主要来源于 Internet 上下载的图片和数码相机拍摄的照片，包括各种环境下、不同复杂背景、不同尺寸的人脸图像。测试图像共 129 幅，包括 149 个人脸。

整个人脸检测算法由两部分组成，第一部分是粗略定位人脸的位置，主要通过肤色分割算法得到候选人脸区域，第二部分为提取人脸特征并最终得到人脸检测结果。

第二部分的人脸特征提取算法，分成三步来实现，它结合了器官分布特征、积分投影、边缘图像、方差投影等知识信息，能比较准确地提取出眼睛。在讨论时，算法中的三个步骤独立成三种特征提取方法。用方法 1 代表基于器官分布特征的方法，用方法 2 代表基于积分投影和边缘图像的方法，用方法 3 代表基于积分投影和方差投影的方法，统计各种实验检测结果，得到表 4.1：

表 4.1 人脸检测结果统计（129 幅图像，149 个人脸）

	正确数	误检数	检测率	误检率
方法 1	112	0	75.2%	0
方法 2	111	2	74.5%	1.8%
方法 3	102	3	68.5%	2.9%
方法 1+方法 2	144	2	96.6%	1.4%
方法 1+方法 3	143	3	96.0%	2.1%
方法 2+方法 3	120	5	80.5%	4.0%
方法 1+方法 2+方法 3	146	5	98.0%	3.3%

图 4.1 为采用肤色分割算法和基于器官分布特征的方法，对一幅彩色图片进行检测时，得到的中间结果和最终输出结果。

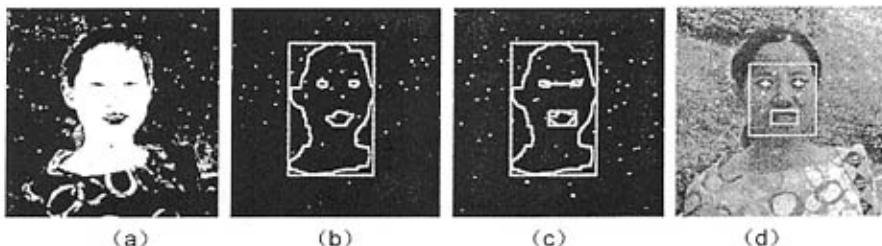


图 4.1 基于肤色分割算法和方法 1 检测得到的相关图像

- (a) 肤色分割结果
- (b) 得到的候选人脸区域
- (c) 基于器官分布特征的人眼和嘴巴定位
- (d) 人脸检测结果

图 4.2 为采用肤色分割算法和基于积分投影与边缘图像的方法，对一幅彩色图片进行检测时，得到的中间结果和最终输出结果。

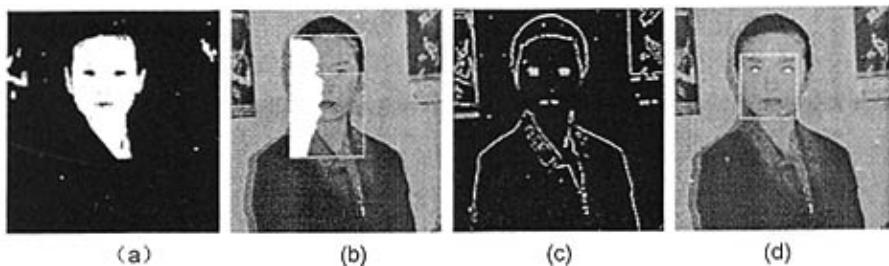


图 4.2 基于肤色分割算法和方法 2 检测得到的相关图像

- (a) 肤色分割结果
- (b) 水平灰度积分投影估计人眼水平位置
- (c) 加窗处理定位人眼位置
- (d) 人脸检测结果

图 4.3 为采用肤色分割算法和基于积分投影与方差投影的特征提取方法，对一幅彩色图片进行检测时，得到的中间结果和最终输出结果。

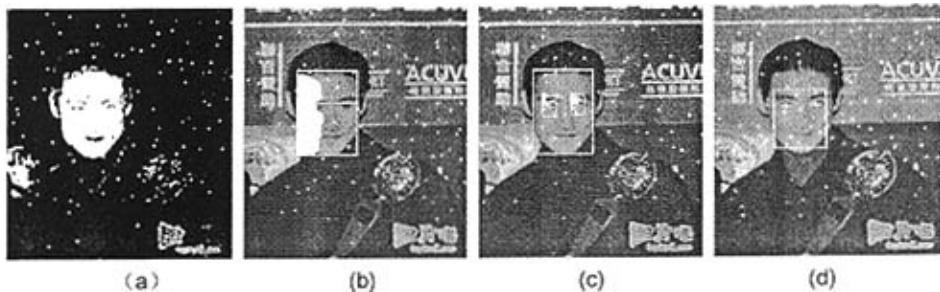


图 4.3 基于肤色分割算法和方法 3 检测得到的相关图像

- (a) 肤色分割结果
- (b) 水平灰度积分投影估计人眼水平位置
- (c) 方差投影定位人眼位置
- (d) 人脸检测结果

图 4.4 为部分实验测试结果。

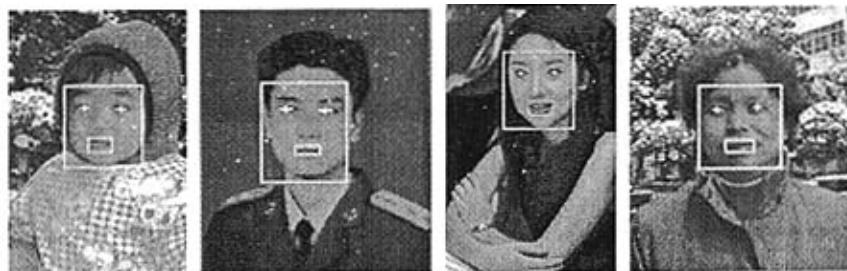




图 4.4 部分图像的人脸检测结果

4.2 实验结果分析

对实验结果进行具体分析，可得到一些结论。

在肤色分割中有如下特点：

- (1) 采用基于“灰度世界假设”的色彩均衡预处理技术，使得肤色分割算法更加稳健；
- (2) 改进的 H_SI_I 肤色模型在光亮度变化的条件下，比 HSI 肤色模型检测效果要好；
- (3) 基于模糊聚类方法的肤色再分割能极大降低背景干扰点的出现，提高肤色检测的准确性，但是模糊聚类算法的计算量大，降低了肤色分割速度；

基于器官分布特征的方法，有如下特点：

- (1) 检测方法简单，检测速度较快，能适用于小角度倾斜和偏转的人脸；
- (2) 图像漏检率比较高，主要是因为一些图像质量较差或眼睛被遮挡，使得肤色分割后没能在双眼区域形成相应的孔洞，无法通过孔洞匹配实现人眼检测，如图 4.5 所示，左边人脸的眼睛被头发部分遮挡，无法在眼睛处形成完整的孔洞，实现不了双眼的匹配，因而该人脸被漏检；

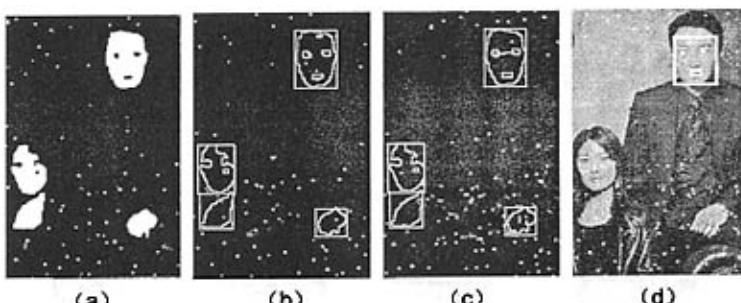


图 4.5 双眼匹配失败造成人脸漏检的实例

(a) 形态学处理后的肤色分割效果图

(b) 得到的候选人脸区域

(c) 双眼匹配的结果

(d) 人脸检测结果

- (3) 图像误检率较低，因为候选人脸区域和人脸相似而又不是人脸的情况，其发生的概率很小。

基于积分投影和边缘图像的方法，有下列不足：

- (1) 图像漏检率比较高，主要原因是由于该方法是种串行算法，在任何中间结果出错时，都会使得检测失败。方法中，通过水平灰度积分投影实现眼睛水平位置的估计是关键步骤，主轴旋转、高斯平滑的处理都是为了提高水平投影的准确性，但是对于一些图像仍然无法准确定位人眼水平位置，如图 4.6 所示，由于左边人脸的水平积分投影图像在眼部位置没有对应的谷点，因此估计不出人眼

水平位置，造成该人脸被漏检；

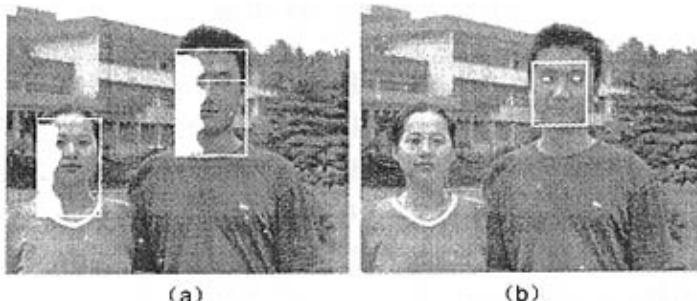


图 4.6 人眼水平位置估计失败造成人脸漏检的实例

- (a) 水平灰度积分投影估计人眼水平位置
- (b) 人脸检测结果

(2) 在遮挡及人眼完全闭合的情况下，加窗处理对于眼睛的提取不够理想。

基于积分投影和方差投影的方法，有如下特点：

- (1) 方差投影处理比加窗处理能更精确地定位人眼的位置，且对于小角度倾斜和偏转的人脸也同样适用；
- (2) 图像漏检主要是因为水平灰度积分投影对眼睛水平位置进行估计出现错误，如图 4.6 所示；
- (3) 图像误检主要出现在方差投影定位人眼的步骤，如图 4.7 所示，左边人脸在方差投影定位人眼时出现错误，造成该人脸定位上偏差较大。

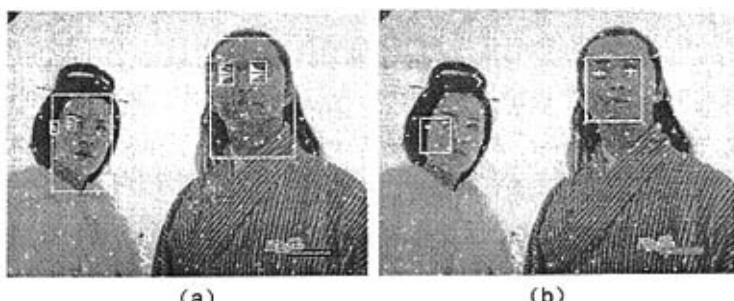


图 4.7 方差投影错误定位人眼造成人脸误检的实例

- (a) 方差投影定位人眼的图片
- (b) 人脸检测结果图

本文的特征提取算法将三种方法结合起来，应用在整体检测算法的不同阶段，有效地利用了它们的优点，克服了各自存在的缺陷，达到多种特征信息的优化组合。实验结果也证明，三种方法的结合，使得特征定位的精确性和人脸检测的准确率大幅提高。

第五章 总结与展望

本文针对复杂背景下的彩色图像，提出基于肤色分割和特征定位相结合的人脸检测算法，实现彩色图像中人眼的定位和人脸的检测。

本文首先探讨了肤色分割算法。根据光照理论，分析了 HSI 颜色空间对光亮度变化敏感的原因，建立起一个改进的颜色空间（H_SI_I 空间），在该空间的肤色模型对于光亮度变化具有强鲁棒性。同时，采用模糊 C 聚类算法进行肤色的再分割，提高了肤色分割的准确性。另外，采用色彩均衡预处理方法，使得肤色分割算法更为稳健。

特征提取是人脸检测中非常重要的步骤，针对候选人脸区域，本文分成三步来实现，正如文中所述，这三个步骤可独立成三种特征提取方法。通过实验发现将三种方法结合起来，应用在整体检测算法的不同阶段，能达到多种特征信息的优化组合。首先，通过器官分布特征可快速定位人脸，但是在图像质量较差或肤色分割不够好时，则须结合灰度信息来实现特征提取；其次，比较第二步和第三步的检测效果，积分投影和边缘图像结合，得到的人脸检测正确率要高些，但是积分投影和差分投影相结合时，特征的定位更精准。

本文的检测算法仍有许多不足，如在遮挡、人眼闭合、脸部倾斜角度较大时检测效果不理想。因此，下一步的研究工作主要包括：

- (1) 在形态学处理上做进一步改进，以提高检测算法的灵活性；
- (2) 特征提取上实现多元化提取，改进串行算法的不足，降低人脸检测的漏检率，且可以通过抽取特征的细节，如眼睑、下额轮廓线等，达到描述整个人脸的目的；
- (3) 在特征提取后，增加人脸验证环节，降低检测误检率；
- (4) 对算法进行扩展，考虑采用基于统计模型的方法实现对多角度、旋转人脸的特征定位和人脸检测；
- (5) 实现对视频序列图像中的人脸检测。利用运动信息可望提高运算速度和准确性。

致 谢

在论文即将完成之际，我的心情无法平静，从开始进入课题到论文的顺利完成，有多少可敬的师长、同学、朋友、亲人给了我无言的帮助，在这里请接受我诚挚的谢意！

首先我要衷心感谢我的导师杨卫平副教授。在攻读硕士学位期间，杨老师以他远见卓识的启发、教诲和指导，使我的专业水平和学习思维方法有了很大的提高。在论文的选题、协作和完成过程中，杨老师给予了我许多热情的帮助和鼓励，并在百忙之中抽出时间审阅了全稿，提出了许多宝贵的指导意见。杨老师渊博的学识、敏锐的思维、民主而严谨的作风使我受益匪浅，并终生难忘。从老师身上学到的不仅仅是知识和学问，更重要的是工作方法、治学态度、一丝不苟的工作作风和孜孜不倦的求索精神。我所取得的成绩都是他无私的帮助和精心培养的结果，在此谨向杨老师致以诚挚的谢意和崇高的敬意。

感谢周小清教授和赵鹤平教授。两位老师为我课题研究创造了良好的研究条件和工作环境。两位老师高深的学术造诣、严谨的治学作风、对教学工作的奉献精神都给我留下了深刻的印象，是我今后从事学术研究的楷模。

感谢我的同学和朋友，在学习上和生活上他们给予我许多支持和鼓励，和他们的融洽相处、热烈讨论、坦诚交往将会成为我人生中最美好的记忆。

最后我要感谢我的家人，他们总是鼓励、鞭策我要勤奋努力，不懈怠，没有他们在背后默默的支持，我也是不可能完成学业的，他们的爱是我永远前进的动力。

参考文献

- [1] M. H. Yang, D. J. Kriegman, Narendra Ahuja. Detecting Face in Image: A Survey, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 34-54, Jan. 2002
- [2] 梁路宏, 艾海舟等, 人脸检测研究综述. 计算机学报, 2002, 25(5): 449-458.
- [3] G. Yang, T. S. Huang. Human Face Detection in Complex Background. Pattern Recognition, Vol. 27, no. 1, pp. 53-63, 1994
- [4] Terrillon J C, Shirazi M N, Fukamachi iHetal. Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images. In: Proc Conference on Automatic Face and Gesture Recognition, Grenoble France, 2000. 54-61
- [5] 张欣, 徐彦君, 杜利民. 彩色图象中主要人脸特征位置的全自动标定. 中国图象图形学报, 2000, 5(A2): 138-143
- [6] 衣晓飞, 陈福接. 单人图象中人眼的快速定位. 计算机工程与科学, 2001, vol. 23, no. 3, pp. 15-24
- [7] Zabrodsky H, Peleg S, Avnir D. Symmetry as a Continuous Feature. IEEE Trans. PAMI, 1995, 17:1154-1166
- [8] Reisfeld D, Yeshurun Y. Robust detection of facial features by generalized symmetry. IAPR'92, 117-120
- [9] 卢春雨. 人脸自动识别若干问题研究与系统实现[博士学位论文]. 清华大学, 1998
- [10] M. F. Augusteijn, T. L. Skujića. Identification of Human Faces through Texture-Based Feature Recognition and Neural Network Technology. Proceedings of IEEE International Conference on Neural Network, pp. 392-398, 1993
- [11] Y. Dai, Y. Nakano. Face-Texture Model Based on SGLD and Its Application in Face Detection in a Color Scene. Pattern Recognition. 1996, Vol. 29, no. 6, pp. 1007-1017
- [12] Jie Yang, A. Weibel. A Real-time Face Tracker. Proceeding of WACV'96. (Sarasota Floreda, USA), 142-147
- [13] Lee C H, Kim J S, Park K H. Automatic Human Face Location in a Complex Background. Pattern Recognition, 1996, 29:1877-1889
- [14] E. Saber, A. M. Tekalp. Frontal-view Face Detection and Facial Feature Extraction Using Color, Shape and Symmetry Based Cost Functions. Pattern Recognition Lett. vol. 19:669-680, 1998
- [15] T. Sakai, M. Nagao, S. Fujibayashi. Line Extraction and Analysis Pattern

- Detection in a Photograph. Pattern Recognition, vol. 1, pp. 233-248, 1969
- [16] I.Craw, H.Ellis, J.Lishman. Automatic Extraction of Face Features. Pattern Recognition Lett, vol5, pp183-187, 1987
- [17] 梁路宏, 艾海舟, 何克忠. 基于多模板匹配的单人脸检测. 中国图象图形学报, 1999, Vol. 4(A), no. 10, pp. 825-830
- [18] A.Yuille, P.Hallinan, D.Cohen. Feature Extraction from Faces Using Deformable Templates. International Journal of Computer Vision, 1992, vol.8, no.2, pp. 99-111
- [19] M.Kass, A.Witkin, D.Terzopoulos. Snakes: Active Contour Models. International Journal of Computer Vision, 1988, vol. 1, no. 4, pp. 321-331
- [20] Moghaddam B, Pentland A. Probabilistic Visual Learning for Object Representation. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1997, vol. 19, no. 7, pp. 696-710, Jan.
- [21] M.H.Yang, N.Ahuja, D.Kriegman. Mixtures of Linear Subspaces for Face Detection. Proc. Fourth Int'l Conf. Automatic Face and Gesture Recognition, pp. 70-76, 2000
- [22] Burel G, Carel D. Detection and Location of Faces on Digital Images. Pattern Recognition Letters, 1994, 15:963-967.
- [23] Vincent J.M., Waite J.B., Myers D.J. Automatic Location of Visual Features by a System of Multilayered Perceptions. IEEE Proceedings F, 1992, 139(6):405-412.
- [24] Julle P, Marsh R. A Hierarchical Neural Network for Human Face Detection. Pattern Recognition. vol. 29, no. 5, pp. 781-787, 1996
- [25] Boser B E, Guyon I M, Vapnik V N. A Training Algorithm for Optimal Margin Classifier. In: Proc of the 5th ACM Workshop on Computational Learning Theory, Pittsburgh, 1992. 144-152
- [26] Gunn S R. Support Vector Machines for Classification and Regression. Image Speech and Intelligent Systems Research group, University of Southampton: ISIS Technical Report, ISIS1-98, 1998
- [27] Osuna E, Freund R, Girosi F. Training Support Vector Machines: An Application to Face Detection. In: Proc Computer Vision and Pattern Recognition, Puerto Rico, 1997, 130-136
- [28] Sung K, Poggio T. Example-based Learning for View Based Human Face Detection . IEEE Transaction on Pattern Analysis and Machine Intelligence, 1998, 20(1):39-51
- [29] Scholkopf B et al. Fast Approximation of Support Vector Kernel Expansions,

- and an Interpretation of Clustering as Approximation in Feature Spaces. In: Levi P, Schanz M, Ahlers R-J, May F eds. *Mustererkennung*. Berlin: Springer-Verlag, 1998. 124-132. <http://svm.first.gmd.de/publication.html>
- [30] Platt J C. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. MSR: Technical Report MSR-T R-98-14, 1998
- [31] Schneiderman H, Kanade T. Probabilistic modeling of Local Appearance and Spatial Relationships for Object Recognition. In: Proc. IEEE Conf. On Computer Vision and Pattern Recognition, Santa Barbara, California, 1998, 45-51
- [32] Schneiderman H. A Statistical Approach to 3D Object Detection Applied to Faces and Cars. PhD Thesis, CMU-RI-TR-00-06,
<http://www.cs.cmu.edu//afs/cs.cmu.edu/user/hws/www/>
- [33] Nefian A V, Hayers M H. Face Detection and Recognition Using Hidden Markov Models. In: Proc. IEEE Conf. on Image Proceedings, Chicago, 1998, 141-145
- [34] Nefian A V, Hayers M H. An Embedded HMM Approach for Face Detection and Recognition. In: Proc. IEEE Conf. on Acoustics, Speech, and Signal Processing, Phoenix, Arizona, 1999, Vol. 6:3533-3556
- [35] 张宏林. 数字图像模式识别技术及工程实践. 北京:人民邮电出版社, 2003
- [36] 郎锐. 数字图像处理学 VisualC ++实现. 北京:北京希望电子出版社, 2003
- [37] Rafael C. Gonzalez Richard E. Woods. 数字图像处理(第二版). 阮秋琦, 阮宇智. 北京:电子工业出版社, 2003
- [38] Aaron C Shumate, Hui Li. Color Balancing in Digital Cameras.
<http://ise.stanford.edu/class/phych.221/00/trek/>
- [39] Jie Yang, A. Weibel. A real-time face tracker. Proceeding of WACV'96 (Sarasota, Floreda, USA), 142-147
- [40] 边肇祺, 张学工等. 模式识别. 北京: 清华大学出版社, 2000
- [41] 章毓晋. 图象工程(上册)——图象处理和分析. 北京: 清华大学出版社, 1999
- [42] 张宏林, 蔡锐. VisualC++数字图像模式识别技术及工程实践. 北京:人民邮电出版社, 2003
- [43] S. H. Jeng, H. M. Liao, Y. T. Liu, M. Y. Chern. An Efficient Approach for Facial Feature Detection Using Geometrical Face Model. Proc. Of ICPR, pp. 426-430, 1996
- [44] 杨淑莹. VisualC++图像处理程序设计. 北京: 清华大学出版社, 北方交通大学出版社, 2003
- [45] 吴崇明, 王晓丹. 基于类别方差的多门限图像分割. 计算机工程与应用, 2001, (17):115-116
- [46] 孙即祥. 数字图像处理. 石家庄: 河北教育出版社, 1993
-

- [47] 荆仁杰, 叶秀清, 徐胜荣等. 计算机图像处理. 杭州: 浙江大学出版社, 1990
- [48] Feng G C, Yuen P C. Variance Projection Function and Its Application to Eye Detection for Human Face Recognition. Pattern Recognition, 1998, 19:899–906
- [49] 何斌, 马天予, 王运坚等. VisualC++数字图像处理(第二版). 北京:人民邮电出版社, 2002
- [50] 贾永红. 数字图像处理. 武汉: 武汉大学出版社, 2003

作者在学期间取得的学术成果

- [1] 张书真, 宋海龙, 杨卫平, 基于肤色模型和模糊聚类的肤色分割算法, 计算机应用研究 (已录用)
- [2] 张书真, 宋海龙, 杨卫平, 复杂彩色图像中的人脸检测, 湖南省研究生创新论坛 (已录用)
- [3] 张书真, 宋海龙, 杨卫平, 基于肤色模型和特征定位的人脸检测算法, 吉首大学学报 (已录用)