

大连理工大学

硕士学位论文

基于FPGA的驾驶员疲劳检测系统的设计与实现

姓名：李栋

申请学位级别：硕士

专业：计算机系统结构

指导教师：马洪连

20091201

摘 要

机动车辆作为社会发展、人类文明和科学技术进步的产物，为人类文明的进步和社会经济的发展做出了巨大贡献。然而，随着机动车辆的增加和交通运输业的发展，交通事故的发生也日渐频繁，给人们的生产生活带来了十分严重的影响。无论是国内还是国外的交通事故统计中，疲劳驾驶造成的生命财产损失的数字都是触目惊心的。因此，如何能够有效地检测并预防驾驶员疲劳驾驶，对于减少交通事故的发生有着非常重要的意义。

本课题的主要目标是开发一套基于 FPGA 的自然光下的驾驶员疲劳检测系统。本文分析对比了当前国内外较为流行的各种检测方法，并研究了这些方法的关键和难点，提出了基于机器视觉的、无接触、无生理负荷的驾驶员疲劳检测方法，建立了基于 FPGA 的实时驾驶员疲劳检测系统，实时采集、处理并显示图像，首先通过肤色建模来定位人脸，然后通过几何特征、灰度水平积分投影、区域生长定位人眼区域，最后通过眼睛状态结合 PERCLOS 疲劳评测标准及眨眼频率判断驾驶员是否疲劳。

系统采用 Xilinx 的 Virtex-II Pro FPGA，该芯片内部集成 IBM PowerPC405 硬核处理器。课题采用软硬件协同设计理念，针对图像处理计算量大的特点，进行合理的软硬件划分，将 FPGA 灵活的硬件设计和实现与处理器的强大软件功能有机结合，通过硬件对视频图像进行部分处理，大大减轻了嵌入式处理器的负担，将不易于硬件实现的算法处理部分交给 PowerPC 处理，提高系统检测的实时性。

关键词：FPGA；疲劳检测；机器视觉；软硬件协同设计

Design and Implementation of Drivers' Fatigue Detection System Based on FPGA

Abstract

Motor vehicle which as the product of the progress of human civilization and society has made great contributions to the economic development .However, with the increase of motor vehicles and transportation development, traffic accidents occurs more frequently,which bring people's production and life serious impact.Both domestic and foreign, in the traffic accident statistics due to driving fatigue losses of life and property of the Numbers are astounding.Therefore, how to effectively monitor and prevent driving fatigue, to reduce traffic accidents has the extremely vital significance.

The main objective of this subject is to develop a set of FPGA-based natural light, driver fatigue detection system.Comparing the current domestic and international popular detection methods' key technology and research result, the proposed driver fatigue detecting method is based on computer vision, non-contact, non-physiological load. Using FPGA logic resource design image data acquisition module, vga display module and image processing module .Firstly ,use face-color model to detect and locate face region , secondly,aiming at the eyes location,the paper bring the mothed of geometric features ,integral projection ,region growth to locate the eyes.finally through the eyes of the state of fatigue with PERCLOS metrics and frequency of wink to determine whether the driver fatigue.

In this paper the HW-SW codesign flow is used,the system is based on Xilinx's Virtex-II Pro FPGA, in the chip there are two IBM PowerPC405 processor.assign proper work to hardware and software,make good combination of FPGA's flexible logic resource design and powerful processor software functions, assign part of the video image processing to hardware greatly reducing the burden of embedded processors,assign other procedure that not easy to process by hardware module to PowerPC by write C code.Experiments result shows the hardware and software efficient cooperation. make system achieve high speed and performance, this system has high real-time performance and accuracy.

Key Words: FPGA;Fatigue detection;machine vision;HW-SW Codesign

大连理工大学学位论文独创性声明

作者郑重声明：所呈交的学位论文，是本人在导师的指导下进行研究工作所取得的成果。尽我所知，除文中已经注明引用内容和致谢的地方外，本论文不包含其他个人或集体已经发表的研究成果，也不包含其他已申请学位或其他用途使用过的成果。与我一同工作的同志对本研究所做的贡献均已论文中做了明确的说明并表示了谢意。

若有不实之处，本人愿意承担相关法律责任。

学位论文题目：基于PCA的驾驶员疲劳检测系统的设计与实现

作者签名：李栋 日期：2009年12月20日

大连理工大学学位论文授权使用授权书

本人完全了解学校有关学位论文知识产权的规定，在校攻读学位期间论文工作的知识产权属于大连理工大学，允许论文被查阅和借阅。学校有权保留论文并向国家有关部门或机构送交论文的复印件和电子版，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。

学位论文题目：基于IP卡的驾驶员疲劳检测系统的设计与实现
作者签名：李栋 日期：2009年12月20日
导师签名：马洪才 日期：2009年12月20日

1 绪论

1.1 课题背景与实际意义

随着社会的进步和人类文明的发展,机动车辆日益增多,机动车辆为经济的发展做出了巨大贡献。然而,机动车辆的增加使交通事故发生量呈现不断攀升的趋势。近年来,我国国民经济快速发展,作为支柱产业的交通运输业发展迅速,交通事故的发生日趋频繁,为国家和人民群众造成了沉重的负担和严重的影响。表 1.1 是 2001-2006 年我国道路交通事故发生的统计数据^[1]。

表 1.1 2001-2006 年我国道路交通事故发生的统计数据
Tab. 1.1 2001-2006 road traffic accident statistics of china

年份	事故次数	死亡人数	受伤人数	损失(亿元)
2001	755142	106231	546580	30.9
2002	773137	109381	562074	33.2
2003	567507	104372	494174	33.6
2004	567753	99217	451810	27.7
2005	450254	98738	469911	18.84
2006	378781	89455	431139	14.9

中华医学会联合相关部门于 2003 年在北京和沪杭高速公路对 516 名驾驶员进行了警觉度的测试及相关的问卷调查;这项调查显示:24%的驾驶员有疲劳驾驶的感觉,50%的被调查者曾在驾车中打过瞌睡。仪器检测结果显示 20%的被调查者存在反应时间慢的现象,50%的被调查者睡眠存在问题。

国外的疲劳驾驶引起的交通事故数据同样触目惊心。美国国家公路交通安全管理局 NHTSA (National Highway Traffic Safety Administration) 近年来的调查^[2]显示:平均每年与疲劳驾驶有关的车祸达 56000 起,造成 76000 人受伤,1544 人死亡该数字占交通事故死亡总人数的 3.7%,调查还发现有开车打瞌睡的经历的受访者占总受访者数的 27%。

国内外交通事故的数据显示:道路交通事故中因驾驶员疲劳驾驶造成的约占总数的 10%至 25%,由此可知,驾驶疲劳已成为造成交通事故的最主要原因之一,调查还显示驾驶员职业病的诱因之一也是疲劳驾驶。因此,开展驾驶员疲劳检测的研究工作意义重大,具很高学术价值及实用价值。

1.2 国内外发展现状分析

常用的疲劳检测方法有主观和客观两种，其中主观研究方法主要包括使用如下问卷调查形式，如斯坦福睡眠尺度表、主观调查表、睡眠习惯调查表和驾驶员自我记录表等。使用主观调查表的优点是使用简单，但是由于填写人对表格条目的理解会存在一定的偏差，并且由于填表人的主观原因，其准确性不能很好的保证，使得结果往往不能令人满意，因此主观方法不能作为疲劳驾驶评测的标准尺度^[3]。

客观的方法有 EMG（肌电图）、EOG（眼电图）、EEG（脑电图）、EKG（心电图）、通过鼻+声传感器测得的呼吸气流、胸腔部传感器测得的呼吸效果等测量方法。这些方法固然准确，但是这些方法所得结果都不是驾驶时得到的，而是超前或滞后的，不具备非实时性。在驾驶室有限的空间内安装复杂的检测仪器也是这些方法主要缺点。并且驾驶员在未进入驾驶室和进入驾驶室的精神状态是不尽相同，即使仪器再精确其检测结果也会受到很大影响。因此，开发一种车载的、实时的、非接触式的、无生理负担的疲劳检测系统已成为国内外专家学者和技术人员的共同目标^[4]。

1.2.1 国外研究状况

(1) 日本尼桑公司研制出的疲劳报警系统能自动提醒驾驶员驾车时不要打瞌睡，将该装置安装到车内驾驶员的前方，装置中的电眼能实时的采集司机的眼部图像，然后经过中央处理器的分析做出判断，进行相应的报警处理。该公司的研究人员将司机的眨眼次数、频率作为驾驶员是否疲劳的依据，如果驾驶员单位时间内眨眼过多，系统就会进行报警，并同时喷放出含有柠檬味的醒脑香气来消除司机的困意。

(2) 美国 Electronic Safety Products 公司开发了一款疲劳检测装置通过检测方向盘的运动来进行判断，该装置可用于各种型号车辆。司机正常驾驶时方向盘处于正常运动，此时检测系统不报警，当方向盘持续 4s 无运动时，系统就会产生报警，报警会一直持续到方向盘继续正常运动为止。该装置通过在方向盘杆上装有的传感器来监测方向盘的运动^[5]。

(3) 英国 Loughborough 大学睡眠研究中心研制的疲劳检测设备可以避免处于疲劳状态的驾驶员开车过程中打瞌睡。通过检测驾驶员在驾驶期间表现出的疲劳迹象进行判断，其特点是引入了参考系数进行综合评测，其中包括车辆行驶状况、驾驶员持续驾驶的时间等。当系统检测出驾驶员的疲劳迹象时就会通过声音和图像发出报警。

(4) 位于澳大利亚墨尔本市的奥斯汀医院也研制了一款疲劳报警器，该疲劳报警系统可以向表现出疲劳状态的驾驶员发出疲劳警告。研究小组的大量实验显示，驾驶员

在疲劳的早期阶段会对红灯及对面来车等刺激反应减慢，该系统内安装有红外线传感器，能检测驾驶员的眼睑活动情况及眨眼的速度，系统可以在驾驶员还未意识到自己疲劳的时候发出警告。使用这款警示器的驾驶员必须习惯佩戴一个大框、粗黑的眼镜，这对于爱美的女士以及平时不佩戴眼镜的人来讲不大美观和方便。

(5) 德国研究人员同样也发明了眼镜状的疲劳检测装置，如果司机在驾驶过程中打磕睡，装置会使车内产生强光，使声音发生器发出声音，从而引起司机警觉。如果未能成功，声音会持续增强。如司机仍未清醒，装置会接通制动器，强制刹车。

1.2.2 国内研究状况

我国从 20 世纪 60 年代就开始对驾驶疲劳检测方法进行实验性的研究。但是由于种种原因，我国在这方面的研究还处于起步阶段，截至目前还没有很成熟的技术问世。

(1) 深圳长途汽车公司的技术人员结合现代神经学、人体生理学、电子工程学等多学科分析疲劳驾驶事故的诱因，研究表明消除司机驾驶过程中的异常疲劳和大脑麻痹是消除疲劳事故隐患的主要途径。研制了佩戴于司机脚踝和手腕的“驾驶员疲劳事故预防器”。

(2) 上海交通大学的学者研究的疲劳检测装置，使用高灵敏度的传感器检测驾驶员驾驶过程中的方向盘、踏板运动状态来判别驾驶员的精神状态，若方向盘和踏板较长时间不动，说明驾驶员已经出现困意正在打磕睡。

(3) 浙江大学的学者研究的防疲劳驾驶装置通过实时监测司机的眼睛闭合时间及频率，来判断当前驾驶人员的注意力集中程度，从而识别驾驶员是否疲劳，给出相应的报警处理。

(4) 吉林大学和中国农业大学的学者，利用图像处理的方法对驾驶员的眼睛特征进行实时跟踪从而判断驾驶员的精神状态^[6]。

通过对比国内外的疲劳检测方法及技术可以看出，我国对于车辆驾驶中疲劳驾驶测评方法的研究同发达国家相比，还处于起步阶段、还有较大的差距。因此，研究如何利用图像处理技术、脸部及眼部检测与定位技术、PERCLOS 疲劳检测方法、眨眼频率疲劳检测方法相结合，开发出一种车载的、非接触式的、实时的、无生理负担的驾驶员疲劳检测系统是当前的一个研究热点，也是本课题研究的初衷。

1.3 本文的研究内容与组织结构

1.3.1 论文的研究内容

本文在深入研究了国内外各种驾驶疲劳检测方法的原理、关键技术和成果后,进行了以下方面的研究与设计:

(1) 深入分析比较了各种检测方法的关键技术及原理,采用基于眼部状态疲劳判定方法,设计并实现了一种非接触性、自然光照、实时的、无生理负担的疲劳检测系统。

(2) 基于软硬件协同设计理念,在 XILINX 公司的 Virtex II Pro 芯片上搭建系统,针对图像处理计算量大的特点,进行合理的软硬件划分,通过 VHDL 语言实现视频采集、处理、显示模块以及部分易于硬化的图像算法,将不易于硬件实现的图像处理部分交给 PowerPC,编写 C 程序实现人脸定位,人眼定位,状态识别等算法。这样就将 FPGA 灵活的硬件设计和实现与 PowerPC 处理器的强大软件功能有机结合,通过硬件对视频图像进行部分处理,大大减轻了嵌入式处理器的负担,通过软硬件配合提高系统检测的实时性。

(3) 利用肤色特征对人脸区域进行了初步定位,利用肤色在 YCbCr 颜色空间中的聚类特性消除光照对检测结果的影响。

(4) 通过脸部几何特征结合灰度积分投影粗定位眉眼区域,通过自适应阈值分割及区域生长精确定位眼球位置。

(5) 通过眼睛状态结合 PERCLOS 评测算法及眨眼频率判断驾驶员是否疲劳。

1.3.2 论文的组织结构

本文的内容和结构如下:

第一章 绪论。简要介绍了本课题的应用背景及实际意义,并在分析了国内外发展现状的基础上提出了自己的设计理念。

第二章 系统总体方案设计。介绍系统的整体结构,软硬件划分、软硬件协同工作机理以及本文采用的疲劳驾驶检测识别方法。

第三章 视频采集显示部分设计与实现。详细介绍系统视频采集、显示、存储模块的设计与实现。

第四章 人脸检测定位的研究与设计实现,研究目前人脸检测、人脸建模的主要方法,详细介绍人脸检测功能的设计与实现过程。

第五章 眼睛检测定位的研究与设计实现,分析人眼检测与定位的常用方法,提出根据脸部几何特征及水平积分投影粗定位与区域生长精确定位的方法。

第六章 对眼部状态进行分析，并将眼部状态数值化，采用 PERCLOS 与眨眼频率相结合的方法对疲劳进行检测。

最后是对本次开发设计的疲劳检测系统的总结以及对该系统的进一步展望。

2 疲劳检测系统总体方案设计

疲劳检测系统关乎人的生命安全，合理的设计方案是保证疲劳检测系统准确性、实时性、鲁棒性的基础，本章将对驾驶员疲劳检测系统的总体方案进行分析与设计。

2.1 疲劳检测方案

本文采用基于机器视觉的疲劳检测方法。在 FPGA 上完成系统设计，通过高速 CCD 摄像机实时采集驾驶员头部图像，正常驾驶中驾驶员的头部在一定范围内，左右移动不会超过 15 厘米，前后移动不超过 5 厘米，头部会有轻微的晃动，偶尔会有扭头但很快会恢复原位，眼睛正视前方。对 CCD 摄像机采集到的头部图像本系统通过图像处理快速定位并进行眼睛状态的识别与分析。本文中的疲劳检测系统要求系统能适应驾驶员头部在一定程度上倾斜、适应车内的光线变化、适应驾驶员的表情变化等情况。

对于这样一个关乎人的生命的安全系统来说以下性能指标是必要也是必须的。首先，要满足高实时性。疲劳驾驶检测系统关乎人的生命安全，高实时性能保证驾驶员疲劳检测系统能够做出及时的预警，使驾驶员及时做出的调整，消除不安全因素、减少危险的发生。其次，鲁棒性即健壮性要高。系统的鲁棒性至关重要，其具体体现是系统对异常情况也能做出相应的处理和判断，鲁棒性已成为考察安全检测系统的优劣的重要指标。最后，准确率高。作为如此重要的安全系统必须保证系统的高准确率和低误报率，准确率同样是评测疲劳系统好坏的重要指标。

基于机器视觉的疲劳驾驶检测系统的工作流程如图 2.1 所示。处理一帧时，首先，对捕获的图像进行光照补偿；然后，进行人脸检测以缩小眼睛定位的区域，降低搜寻人眼的系统开销；接着，进行人眼定位；最后，进行眼睛状态识别和疲劳分析。

2.2 疲劳检测系统

传统的基于 PC 的疲劳检测系统体积大，不能满足便携的要求；而采用通用的 DSP 芯片组成的系统，外围电路复杂，设计与调试都需要较长的时间，并且系统的可扩展性和可移植性不好。基于视频的实时目标检测系统中，要求处理时间尽可能短，实时性尽可能高。而图像数据量巨大、处理速度慢已成为制约计算机视觉进入实用的瓶颈问题。随着 FPGA 芯片集成度、运算速度、数据吞吐率的不断提高，基于高速 FPGA 的实时图像处理系统的研究得到广泛的关注，这种基于 FPGA 的系统也代表了实时图像处理系统的发展趋势。FPGA 是从 ASIC 的基础上发展起来的，较 ASIC 开发更加灵活。FPGA 芯片除了具有 ASIC 芯片低功耗、小型化等特点外，还具有一系列 ASIC 没有的优点诸如

高可靠性、易修改性、设计开发周期短、研制成本低廉、开发工具先进。更重要的是，使用 FPGA 进行系统开发还可根据后期现场的实际情况，对其进行现场编程，可方便的擦写原来的逻辑，非常方便的完成现场调试和现场功能改进、其灵活的设计让工程师兴奋不已，并且随着集成电路工艺的飞速发展，FPGA 芯片的价格也日益降低。可编程逻辑器件 FPGA 可以使用内部逻辑资源通过编写硬件描述语言实现个性化、可定制的数字功能模块，诸如根据系统实际需要设计并实现的视频采集、储存、显示模块、图像并行处理的算法模块等，使用 FPGA 可以把所有数字模块集成到单个器件中，真正实现 SOPC（可编程片上系统）达到预期的结果。更令人兴奋的是现在的中高端的 FPGA 内部还配有嵌入式处理器如 PowerPC、ARM 等，有些还配有数字信号处理 DSP 等。随着 SOPC 及软硬件协同概念的提出，使得在 FPGA 上进行图像处理更加灵活，本文就是在高端 FPGA 上通过软硬件协同设计的方式完成图像处理，高效的实现 SOPC 系统。

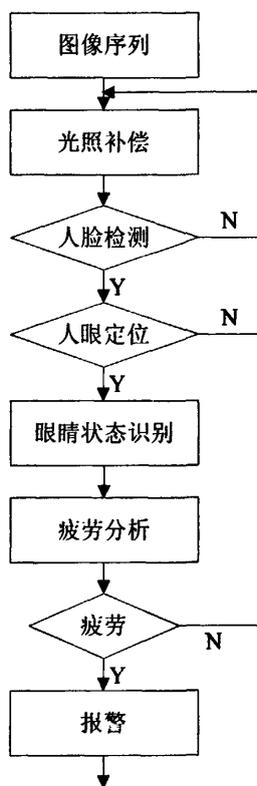


图 2.1 基于视觉的驾驶疲劳检测流程图

Fig. 2.1 Flow chart of vision-based fatigue detection system

本文在 Xilinx 的 Virtex-II Pro XC2VP30 FPGA 芯片上实现疲劳检测算法,该芯片内部集成 IBM PowerPC405。这样本文就可以利用 FPGA 中的可编程逻辑资源构成该嵌入式系统的图像采集模块、图像处理模块、图像显示模块及部分算法模块等。针对图像处理计算量大的特点,进行合理的软硬件划分,通过硬件对视频图像进行部分算法处理,大大减轻了嵌入式处理器的负担,将不易于硬件实现的算法处理部分交给 PowerPC 处理,这样就能使得 FPGA 灵活的硬件设计和实现与处理器的强大软件功能有机地相结合,通过软硬件配合提高系统检测的实时性,高效地实现 SOPC 系统。并且可以在外围电路不变的情况下通过 FPGA 内部的电路设计,使系统的功能得到升级和增强。

相比 Xilinx 公司以前推出的 FPGA 产品 Virtex-II Pro 系列产品结构和性能都有了较大的改进,其主要特色如下: Virtex-II Pro 内部的每个 CLB 模块都包含 4 个 Slice,芯片内部嵌入更大容量的 BlockRAM、BlockRAM 可配置为 RAM、双口 RAM、FIFO 等随机存储器用于 FPGA 内部的随机存取,内部还配有数字时钟管理模块(Digital Clock Manager, DCM)较之前的 DLL 提供更加灵活的时钟管理功能,支持更多的 I/O 接口标准等用于连接不同设备。内嵌的 18×18 专用硬件乘法器多达 556 个,使其更加适用于处理数字信号^[7]。

ISE 和 EDK 是 Xilinx 公司提供的可供开发 Virtex-II Pro 系列 FPGA 的开发工具。其中 VHDL 和 Verilog 等硬件描述语言编写的文件在 ISE 上进行编译。本系统的硬件程序都是在 ISE 下开发完成的。EDK 是 Xilinx 公司推出的用于在 FPGA 片内 32 位嵌入式处理器开发的工具套件,可用于处理器程序开发,EDK 为工程师提供了标准外设及多种设计工具,使工程师可以进行 PowerPC 上的开发。EDK 中包含硬件编译器、硬件仿真器、软件编译器如 Eclips, HW-SW debugger 等等。工程师可以通过集成在 EDK 中的 XPS (Xilinx Platform Studio) 在 Windows 或 Linux 的图形界面下,方便地调用各种内部集成工具,完成整个 SOPC 系统的开发,更重要的是软硬件的开发可以同时进行。设计需要 EDK 与 ISE 配合使用进行 SOPC 的开发,其中嵌入式系统的硬件部分需要通过 ISE 编译完成、在 EDK 中可以调用 ISE 的功能、在 EDK 下可以完成系统硬件生成和软件的编译。

2.3 系统软硬件协同的研究

软硬件协同是本文的一个创新,因此这里有必要先研究一下软硬件协同的方案,所谓软硬件协同设计是指对整个系统的软件部分、硬件部分使用统一的描述方式和工具进

行集成开发,进而完成整个系统的设计验证,并最终跨越软硬件界面进行系统优化。软硬件协同设计的设计流程如下:

首先,通过硬件描述语言(HDL)和软件编程语言(如C语言)进行系统描述,并进行模拟仿真和验证系统功能;

然后,对系统进行功能划分,从软硬件角度分别实现,分别用硬件描述语言和软件编程语言进行设计,并对综合起来的系统进行功能验证和性能预测等仿真确认来进一步分析;

再次,待无问题时即可进行软件和硬件详细设计;

最后,最后进行系统功能及性能测试。

软件的设计与实现在嵌入式处理器上进行,对于本文的实时系统选择一款性能卓越的处理器是十分必要的也是必需的,在 Virtex-II Pro 芯片中可以例化若干个软核处理器 MicroBlaze 及两块 PowerPC405 硬核处理器。此种由设计人员在设计的过程中下载的处理器核称为“软核”,MicroBlaze 软核具有较高的可配置性,它允许开发者根据自己的设计需求进行适当的选择和自定义;硬核是指在 FPGA 芯片出厂之前,处理器核通过授权的形式,已经由厂商嵌入于 FPGA 中,PowerPC405 处理器硬核是 IBM 公司专门为 Xilinx 的 FPGA 芯片开发的处理器产品。

本系统是一个疲劳检测的实时系统,疲劳检测关乎生命安全,对实时性要求很高,处理器必须有较强的处理能力。一般来说,硬核相对于软核具有更快的速度,更高的性能。本文选取 PowerPC405 作为系统的核心处理器,实现人脸检测算法,人眼定位算法,人眼状态识别算法等。嵌入 Virtex-II Pro 芯片中的 PowerPC405 处理器硬核是一种 32 位哈佛结构的 RISC 核处理器。它可以工作在 300Mhz 频率下,具有 420 Dhrystone MIPS 的运行速度。Virtex-II Pro 芯片中的 PowerPC405 处理器支持低功耗设计,支持全部 PowerPC 用户指令集 (UIISA),具有 32 个 32 位通用寄存器、16Kbit 指令缓存 (Cache) 以及 16Kbit 数据缓存。支持专用片上存储接口 (On Chip Memory, OCM) 及专用的 IBM CoreConnect 总线标准结构。其中,处理器局部总线 (Processor Local Bus, PLB)、片上外设总线 (On Chip Peripheral Bus, OPB) 和设备控制寄存器总线 (Device Control Register, DCR) 构成了 IBM CoreConnect 总线结构。处理器还支持多种调试方式和 2 级中断。

图 2.2 是 PowerPC 的体系结构示意图,详细了解体系结构有利于充分利用 PowerPC 资源,编写出高效健壮的代码。

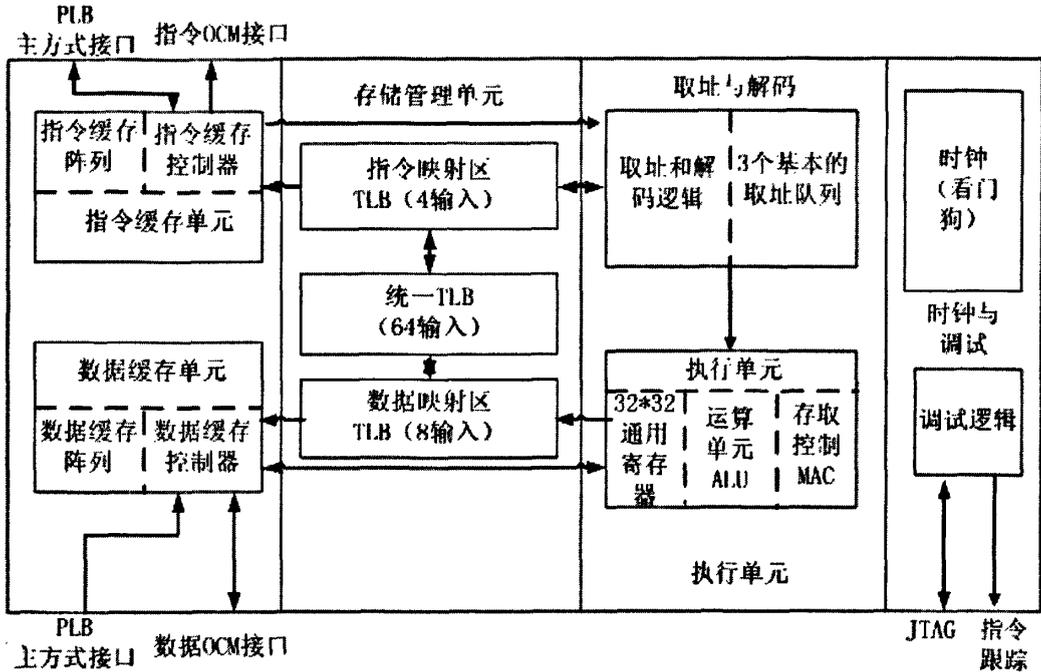


图 2.2 PowerPC405 处理器体系结构示意图
 Fig. 2.2 Processor Architecture of PowerPC405

PowerPC 中的 CoreConnect 总线架构是由 IBM 公司开发的总线连接解决方案。在 Virtex II Pro 系列 FPGA 中不同的 IP 核就是通过 CoreConnect 总线连接起来构成完整的系统。得益于 CoreConnect 技术的运用，使得在一个标准平台内的系统核、处理器核和外设核的集成和复用都变得非常容易实现，这也极大的提高了系统的整体性能。

本设计在详细研究 PowerPC 总线结构后采用如下方案，将视频采集、存储、处理、显示模块通过 VHDL 编写成 IP 核，并遵循 PowerPC 的 CoreConnect 总线标准将这些模块连接到 PLB 总线上，作为 PowerPC 外设的一部分，从而完成整个系统的软硬件协同工作，CoreConnect 总线标准架构如图 2.3 所示，下一节将对软硬件协同的实现进行具体的阐述。

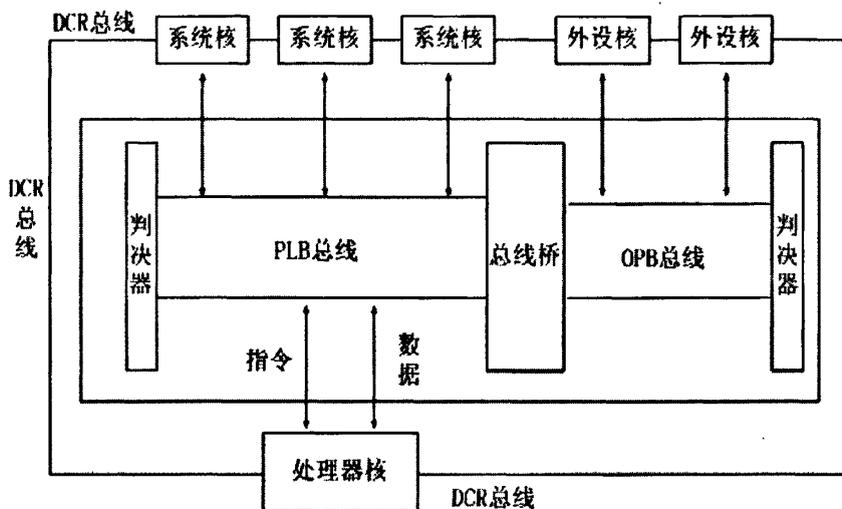


图 2.3 CoreConnect 总线架构

Fig. 2.3 Architecture of CoreConnect

2.4 系统总体方案设计

本系统分为软硬件两个部分，硬件部分通过 VHDL 语言编写功能模块使用 FPGA 内部可编程资源实现，如图 2.4 所示灰色部分，软件部分通过在 PowerPC 上编写 C 语言程序实现，如图 2.4 所示 PowerPC 白色部分；其中硬件部分包括视频采集模块，帧存储模块，视频输出模块，光线补偿模块，中值滤波模块，阈值分割模块；软件部分包括肤色建模程序，膨胀与腐蚀程序，第一次分割程序，第二次分割程序，精确定位程序，状态识别程序，疲劳分析程序。

视频采集模块：通过 I2C 总线控制逻辑控制视频采集芯片产生需要的视频流，并将视频流解码，采集到的图像数据存储到帧存中。

帧存储模块：存储连续三帧，通过地址操作图像数据。

视频输出模块：将帧存中的视频图像输出到 VGA 显示器。

光线补偿模块：实现对采集到的人脸视频图像的光照补偿，提高识别的准确性。

中值滤波模块：对眼部区域的图像进行去噪处理，同时保持图像的细节，提高识别的准确性。

阈值分割模块：通过动态自适应阈值方法对眼部图像进行阈值分割。

肤色建模程序：通过人脸肤色模型阈值化色调图像，得到粗略的包含人脸候选区（灰度值是 255）和背景（灰度值是 0）的二值分割图。

膨胀与腐蚀程序：对分割出来的皮肤区域进行膨胀与腐蚀操作，除去噪声对肤色分割的影响。

第一次定位程序：通过人脸几何特征，将人脸区域进行第一次分割。

第二次定位程序：通过对脸部区域的水平灰度积分投影粗定位眼部区域。

精确定位程序：通过区域生长方式，对二值化的图像进行区域生长精确定位人眼。

状态识别程序：提取眼睛状态，为疲劳分析提供依据。

疲劳分析程序：使用 PERCLOS 及眨眼频率方法进行疲劳评测。

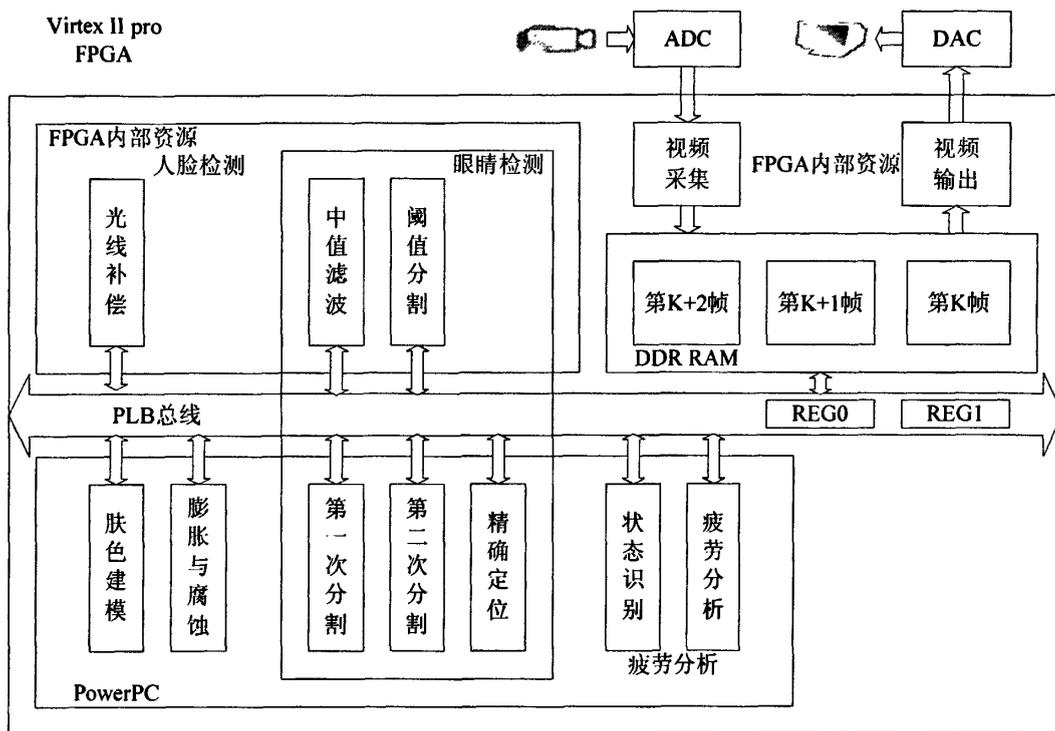


图 2.4 系统框图

Fig. 2.4 System architecture diagram

如何使软硬件协同工作是系统实现的关键问题。Xilinx 提供的片上可编程系统 (SOPC) 解决方案是以硬核 (PowerPC) 或软核 (MicroBlaze) 为核心处理器，以硬件描述语言编写的功能模块为外设，通过符合 IBM CoreConnect 标准的 OPB 或 PLB 总线将处理器与外设连接起来，完成软硬件协同。

Xilinx Platform Studio 提供了以硬件语言描述的外设到总线的标准接口 (IPIF) 及其 BSP, IPIF 内部提供的功能模块包括 S/W 复位, 模块鉴别寄存器 (MIR), 中断机制, 寄存器, 读写 FIFO, 地址译码器, 直接存储器存取 (DMA) 等。用户可以根据实际要求选择相应功能模块, 定制符合设计需求的 IPIF。IPIF 可以形象的比喻成一个符合 IBM CoreConnect 总线标准的外壳 (wrapper), 手动改写 IPIF 中 user_logic 文件可将用户逻辑与 IPIF 中的功能模块连接起来, 应用软件通过使用 BSP 即可访问控制用户逻辑^[8]。

本文根据实际需要定制了具有 S/W 复位, 寄存器, 地址译码器, DMA。寄存器 0 对应硬件模块的 Spdone, 寄存器 1 对应硬件模块的 Hpdone。设计中使用 DMA 方式将图像数据储存在 DDR RAM 中, 如图 2.4。

系统上电后, FPGA 通过 I2C 总线对视频采集 AD 芯片进行配置, 产生遵从 ITU BT.656 标准的 NTSC 格式的 YCbCr 数据以及图像数据时钟, FPGA 在接收到 NTSC 格式数据流后, 从数据流中解析出有效的灰度值 Y, Cr, Cb 并将数据存储到帧存中, 图像帧存放在 DDR RAM 中, 本设计存储连续三帧图像, 第 K+2 帧为刚刚采集到的, 第 K+1 帧用于处理, 进行人脸检测, 人眼定位, 疲劳分析等处理, 第 K 帧是已经处理完的帧用于显示。帧存储的帧存切换的机制见 3.2 节。每次采集到的图像经过人脸检测, 人眼定位, 疲劳分析等一系列处理后, 产生帧存切换信号, 将刚刚处理完的帧显示出来。其中人脸检测, 及人眼定位与检测部分是由软硬件协同完成的, 软硬件切换通过定制的两个寄存器协助完成, 这两个寄存器软硬件可以同时访问的, 当硬件处理结束时产生 Hpdone, 此时软件开始进行后续工作, 当软件处理结束时产生 Spdone, 此时硬件开始工作, 当所有处理结束时需要切换帧存也是通过 Spdone 完成的。这里还根据 IBM 的 Core Connect 标准将硬件部分作为一个外设连接到 PLB 总线上。这样就把软硬件作为一个有机的整体进行协同, 提高了系统工作效率及实时性。

3 图像采集显示模块的设计与实现

图像的采集与显示是本系统的基础, 如何根据本系统的实际情况采集到相应的图像数据供给疲劳检测算法处理, 并将处理后的图像显示出来是本章研究的主要问题。

3.1 视频采集模块的设计与实现

使用 FPGA 内部逻辑资源通过编写 VHDL 语言生成芯片控制模块对 ADV7183B 芯片的配置完成以后, 解码芯片就可以正常工作^[9]。视频解码芯片将 NTSC 制式的模拟视频信号转换成符合 ITU-R BT.601 标准的数字视频流。本文通过硬件描述语言 VHDL 编写视频解码模块用于从 ITU-R BT.601 标准视频流中解码出视频中的有效数据。对于 NTSC 制式的视频数据, 每行为 858 像素, 一帧 525 行。其中, 有效的数据为 720 行, 第 20~263 行是奇数场视频数据, 283~525 行是偶数场视频数据, 其余为行场消隐数据。NTSC 制式的图像每秒扫描 30 帧。NTSC 制式的视频信号分辨率为 858×525, 其中有效视频数据为 720×487 像素。如表 3.1 所示。

表 3.1 NTSC 垂直时序参考
Tab. 3.1 NTSC vertical timing reference

Line	F	V	H(EAV)	H(SAV)	Notes
1-3	1	1	1	0	Blanking
4-19	0	1	1	0	Blanking
20-263	0	0	1	0	Active
264-265	0	1	1	0	Blanking
266-282	1	1	1	0	Blanking
283-525	1	0	1	0	Active

每行数据包含 YCbCr 视频数据信号和行控制信号。视频数据是以 27M 字/秒的速率传送的, 其顺序为: Cb, Y, Cr, Y, ……其中, Cb, Y, Cr, 这三个数据指的是相同地址的亮度和色差信号的取样, 后面的 Y 数据对应于下一个亮度取样。每一行的开始 276 字节为行控制信号, 其中最前面的 4 字节为 EAV 信号 (有效视频结束), 紧接着是 268 个固定填充数据, 最后的 4 字节是 SAV 信号 (有效视频起始)。ITU-R BT.656 标准视频流数据的每行数据的结构如图 3.1。

SAV 和 EAV 信号包含 3 字节的前导，其内容为：FF、FF、00；最后 1 字节 XY 包含该行位于整个数据帧的位置以及如何区分 SAV、EAV 的重要信息。XY 字节各比特的数据中最高位 bit7 为固定数据 1；F=0 表示偶数场，F=1 表示奇数场；V=0 表示该行为有效视频数据，V=1 表示该行没有有效视频数据；H=0 表示为 SAV 信号，H=1 表示为 EAV 信号；P3-P0 为保留信号，由 F、V、H 信号计算生成； $P3=V \text{ XOR } H$ ； $P2=F \text{ XOR } H$ ； $P1=F \text{ XOR } V$ ； $P0=F \text{ XOR } V \text{ XOR } H$ 。

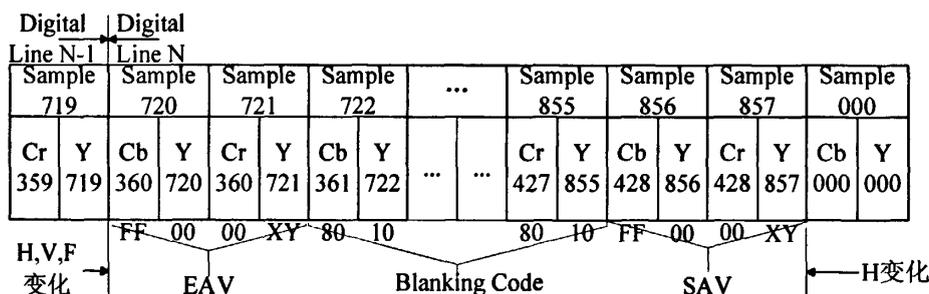


图 3.1 NTSC 制式视频每行数据结构
Fig. 3.1 Data structure of NTSC video data

视频采集模块在每个时钟的上升沿读取从解码芯片传来的 8 位数据，视频解码模块检测一行数据的开始标志是否为 FF0000XY；当检测到 SAV 信号或 EAV 信号时，提取 H、F、V 信号。然后开始解码，与此同时开启行列计数器，开始对接下来的图像数据进行解码。通过视频解码模块视频流被解码成如图 3.2 所示。其中每个 Cb, Cr 各复用一次，通过复用使 YCbCr 变成 4: 4: 4，这样就使每个像素的 YCrCb 值独立出来。通过视频解码将 8bit 数据变成了 24bit 数据。

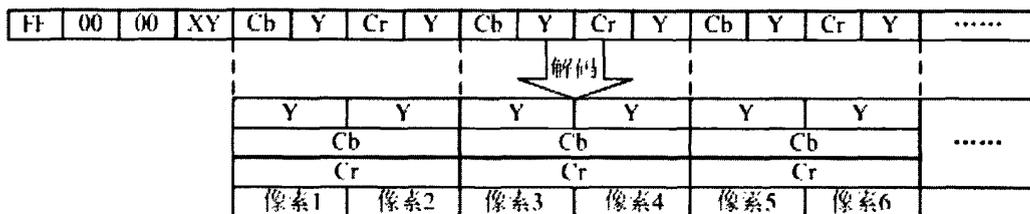


图 3.2 视频解码示意图
Fig. 3.2 Video decoder diagram

视频采集模块根据行计数及指示信号可以提取需要的行，根据列计数及指示信号可以从每行视频数据中分解出需要显示视频图像的 Y, Cr, Cb 分量数据，并同时随图像数据产生视频数据在帧存中存放的地址，如图 3.5 所示。

本文采集视频图像的有效像素中的 256×256 存入帧存，由于视频流来自于隔行扫描的 NTSC 摄像头，因此需要取有效的奇数场数据中的 256 行，以及对应的偶数场数据中的 256 行。写数据的时候需要跳地址写，根据行同步信号或者 SAV 来区分行，换行的时候地址要加额外的 256，用来存放夹杂在奇场信号中的偶场信号，直到出现偶场信号地址切换为初始基地址加 256，接下来的处理方法的同奇数场，具体的地址分配见图 3.3。这样就使隔行的视频数据变成逐行数据并存入了帧存模块。

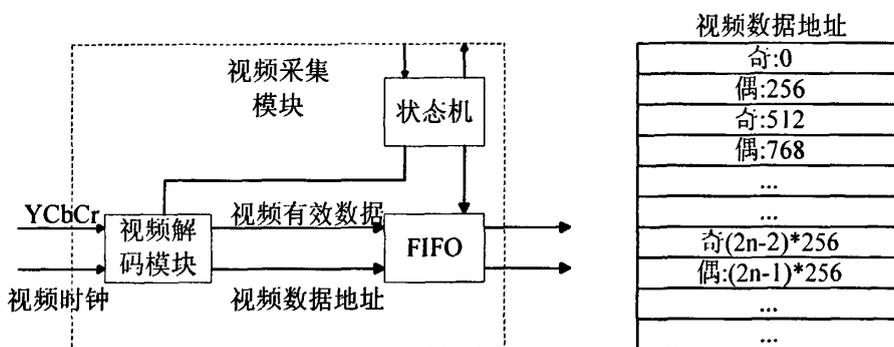


图 3.3 采集模块及视频数据存储地址

Fig. 3.3 Video capture module and video store address

3.2 帧存储的设计与实现

帧存储是整个系统的关键，视频流数据需要到帧存中暂存等待处理，实现疲劳检测算法的软硬件都需要对帧存中的图像数据进行处理，视频显示模块需要将处理后的图像显示出来。因此本系统中，存储器即帧存储模块可以作为本系统的核心，这也正是冯·诺依曼计算机的基本思想。设计中视频采集模块首先对视频数据进行解码，然后将一行数据暂存在 FIFO 中，最后通过 DMA 方式将图像数据存入 DDR RAM 中。

帧存中存放连续三帧图像，大小为 $256 \times 256 \times 24$ Bit。存储连续三帧。工作方式如图 3.4。某一时刻 t 帧存中的数据为 $FK, FK+1, FK+2$ ，其中 FK 内存中驻留时间最长的帧， $FK+2$ 是内存中驻留时间最短的帧，从图中可以看出此时 FK 作为显示帧， $FK+1$ 做为正在处理的帧， $FK+2$ 作为最新采集的数据帧。当对 $FK+1$ 的应用疲劳检测算法检测结束

后，开始采集新的一帧 FK+3。此时 FK 已经没有使用价值可以用 FK+3 覆盖 FK，此时，驻留时间最长的帧为 FK+1，驻留时间最短的帧为 FK+3，驻留时间最长的帧为已经处理过的帧，用于显示。驻留时间最短的帧的为刚刚采集的未经过任何处理的帧，等待处理。当 FK+2 被处理完时，再次采集新的一帧 FK+4，覆盖之前的显示帧 FK+1。每次处理完一帧图像，都将新采集的帧覆盖刚刚显示的帧，即最新帧覆盖最老帧的策略。这样只需要三帧的存储空间就可以完成相应的操作。

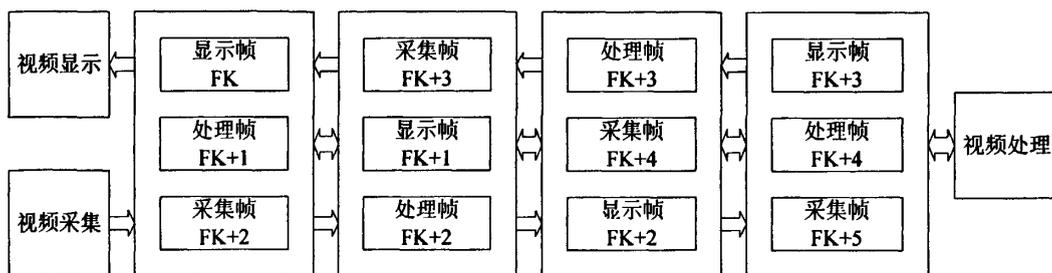


图 3.4 帧存切换方式
Fig. 3.4 Frame switch method

3.3 视频显示模块的设计与实现

视频显示模块用来将帧存中图像数据显示到 VGA 显示器上。VGA 显示器所采用的色彩是 RGB 色彩空间本文中的帧存中的数据是 Y, Cr, Cb 数据，这里需要将 YCbCr 空间转化成 RGB 空间，通过如图 3.5 的 YCrCb 转换 RGB 模块，RGB 数据经过 FIFO 转至 VGA 显示器上，视频显示模块根据 VGA 的工业标准产生 VGA 控制信号 HS（水平同步信号），VS（垂直同步信号），使得 VGA 显示器将帧存中的信号显示出来。

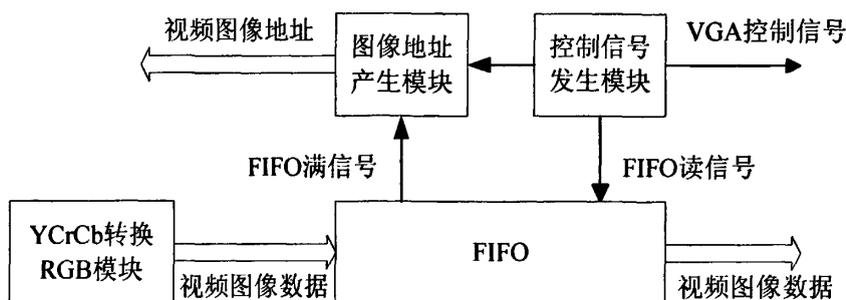


图 3.5 视频显示模块
Fig. 3.5 Video display module

4 人脸检测与定位

基于机器视觉的疲劳检测系统必须准确地检测眼睛，若直接在整幅图像中直接寻找眼睛，由于背景的不确定性，在背景中可能出现若干符合眼睛条件的图像，这样很容易造成误判。如果能在脸部范围内寻找眼睛就会相对准确很多，为了快速准确检测眼睛本文首先确定人脸区域^[10]，然后再在定位出的人脸区域内进一步检测定位眼睛，这样可以提高系统的准确性及实时性。

4.1 人脸检测方法的研究与设计

人脸检测问题最初在人脸识别研究中提出，是人脸识别的关键环节。早在上世纪六七十年代国内外学者就开始了对面脸识别的研究，近些年技术逐步成熟。早期的人脸识别主要针对诸如无背景的图像等具有较强约束条件的人脸图像进行研究，并假设脸部位置为可获得的或已知的。随着信息化时代的到来，电子商务，数码产品，身份识别等应用的发展，人脸识别已跃升为最有潜力的生物身份验证方法，上述应用都要求人脸识别系统能够自动对大多数环境图像具有一定的自适应性，而并非早期的脸部位置已知或易获得，这使得人脸检测开始作为一门独立的研究课题受到国内外越来越多的研究者的重视。如今，人脸检测已广泛应用于数字图像视频处理、内容检索、视觉检测等领域。

常用的人脸检测方法有基于知识和基于统计两大类。

(1) 基于知识的人脸检测

基于知识的人脸检测方法是利用人脸特征知识建立规则，常采用模式识别理论，通过假设与验证进行基于进行模式识别的人脸检测^[11]。

基于运动的方法：该方法的思想是通过分析运动信息将人脸从复杂背景中分割出来，人相对于背景不变的图像来说通常是运动的，这些运动包括眨眼、张嘴等，常见的运动方法是帧间相减法其中包括背景差分法，帧间差分法，光流法可用于图像的动态分析^[12]。

基于纹理及脸部颜色的方法：肤色中的亮度受光照引起的亮度变化影响较大，色度受光照的影响较小并且在颜色空间中的分布有一定的聚类性，可以使用颜色信息将脸部从背景中分割出来。基于肤色特征的优点是处理速度快，实时性高，非常适合实时视频图像的人脸定位处理^[13]。缺点是检测出的人脸区域不够准确。

基于对称性的方法：脸部及脸部各器官都存在一定的对称性。因此国内外有学者通过连续对称性检测方法检测圆形区域的对称性，从而判定是否为人脸。

基于轮廓提取的方法：人脸检测可通过脸部轮廓近似椭圆这一特征来进行。Hough 变换是检测椭圆的常用方法，Hough 变换方法具有较好的鲁棒性和容错性，能够处理被噪声干扰的图像以及边界区域间断的图像，缺点是计算量大，识别速度慢，有文献提出通过检测半圆来减少计算量^[14]。

基于器官分布方法：脸部器官的空间位置分布规则同样可以作为检测的依据。基于器官分布的方法首先检测几个器官分布可能位置，然后将这些位置的坐标进行分别组合，通过各器官分布的几何关系属性准则作为特征通过分类器进行分类筛选，最终找到可能存在的人脸区域。

(2) 基于统计的人脸检测

基于统计的人脸检测方法的基本思想是将人的脸部图像视为一个多维向量，这样就可以通过多维空间下的分布信号的检测问题来解决人脸检测问题。

事例学习的方法：通过对非人脸样本集和人脸样本集的学习训练产生分类器。人工神经网络是通常使用的方法，神经网络方法能够避免了复杂的特征提取工作，能根据样本自我训练学习有很强的自适应和鲁棒性。

基于子空间的方法：该方法利用特征脸的方法，根据脸部模式在特征脸的子空间中的投影系数基本相似这一特性，首先将子图在特征脸空间投影后重建，然后比较原图及重建后的图判定原图是否为人脸^[15]。

基于空间匹配滤波器的方法：首先为每类模式都建立一个模板，识别时将待识别模式与模板匹配，固定的相关性模板就相当于一个匹配滤波器，在检测目标过程中，通过候选位置的图像与参考模板间的匹配度判定是否检测到了待测目标^[16]。

本设计要求人脸检测算法能适应驾驶室内光照变化、驾驶员头部旋转及表情变化等各种情况，疲劳检测系统关乎驾驶员的生命，该实时系统对运算量和运行速度的要求较高，本文选择肤色信息定位人脸，基于肤色的人脸检测的运算量较低，执行效率较高，能够满足实时性的要求。而且肤色能够与大部分的背景颜色相区别，并且不受面部表情、头部角度等细节的影响，具有很高的实用性和鲁棒性。

人脸检测流程如图 4.1：首先对采集到的图像进行光线补偿，实验室或车内的光线不是很充足，并且忽明忽暗，本文采用的是硬件方法的直方图修正法，接着进行人脸肤色建模，得到二值图像，对建模后的人脸区域进行数学形态学处理，腐蚀膨胀两次，能够补偿由于各种原因引起的脸部区域提取的图像点的离散化问题，使之更好的连通。通过肤色建模提取出的图像区域可能包括一些其他非脸部小连通区域因此需要去除非脸区域的操作，去除的方法是考察各目标区域连通的像素数目，如果大于某一个阈值则认

为该区域是人脸区域保留之，否则去除该区域。再对图像进行膨胀、腐蚀操作从而使人脸区域更好地连为一体。最后可以得到人脸区域的上下左右边界位置。

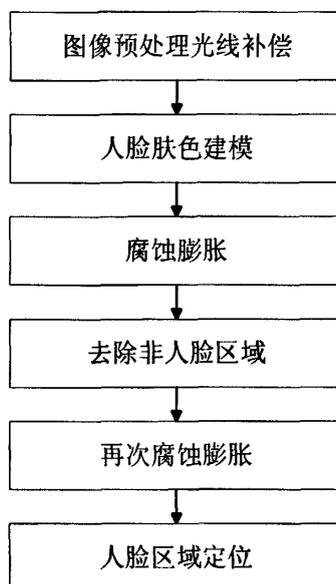


图 4.1 人脸检测流程

Fig. 4.1 Face detection flow

4.2 光线补偿模块的设计与实现

车内光照不均，这样很容易引起人脸各处颜色的不一致，使得相同的人脸在不同光线的环境下得到差异较大的图像。这种影响将大大地降低图像检测的准确率。通过光照补偿处理即可将不均匀脸部图像处理并还原成亮度均匀易于特征提取的脸部图像。常见光照补偿方法有：

(1) 局部图像增强方法

该方法将图像分为若干图像小块，将每块中的最小灰度值作为前景图像对背景照度的粗略估计。接着将粗略估计扩展到整个图像，采用插值函数方法对图像进行平滑扩展处理，进而得到某一个照度的背景图。为了纠正光照的不均匀影响用原始图像中减去背景照度图像。

(2) 同态滤波增强方法

该方法将图像的照明反射模型作为频域处理基础，通过增强对比度和压缩灰度范围改善图形。同态滤波就是对由照明分量和反射分量两个分量组合而成的图像进行取对数运算，通过运算将两个分量的乘性关系转化成加性关系，进而进行光线补偿。

(3) 高斯滤波方法

该方法通过高斯滤波器对图像进行模糊化操作，以达到对图像照度的一个最优估计。使用高斯滤波法可以通过设计高斯滤波器并取滤波器模板的尺寸和标准背离率，得到对图像的照度估计的较好效果以进行较好的光照补偿。

(4) 最优估计方法

最优估计法通过线性函数近似待处理图像的整体亮度。通过将这个函数从图像中减走来补偿光照条件的变化。最后使用直方图均衡化方法，对源图像灰度分布范围进行相应的扩展，从而增加图像的对比度。

(5) 直方图修正方法

直方图修正是图像灰度级变换的最常用方法之一，其目的是消除成像设备感应曲线的不同和光照条件变化给图像处理带来的影响，通过这种方法直方图变成均匀分布的形式。直方图修正处理可以使图像对比度增加，图像质量提高。该方法在医学影像方面有很广泛的应用：医学方面为了改善 X 射线操作人员的工作条件往往使用较低强度的 X 射线曝光，但使用低强度的 X 射线曝光获得的 X 光片灰度集中在暗区，图像中的许多细节无法看清，通过直方图修正法可以使灰度级别分布在合适人眼的亮度区域，就可以使 X 片中关节、筋骨等细节变得清晰。此外直方图修正法还广泛应用于公安军事等照片的处理、工业中非可见光下成像的零件无损检查方面等。

本设计利用 FPGA 内部可编程资源通过 VHDL 语言编写硬件模块实现直方图修整，大大提高了光线补偿的速度。直方图修正法首先需将直方图归一化，即将图像灰度范围从 $[Z_l, Z_k]$ 归一化为 $[0,1]$ 。设变换前的图像的任一灰度级 Z 归一化为 r ，变换之后的图像的任一个灰度级 Z 归一化为 s ，则 r, s 应当满足： $0 \leq r \leq 1$ ， $0 \leq s \leq 1$ 。所以该直方图修整方法所对应公式如下：

$$s = T(r) \text{ 或 } r = T^{-1}(s) \quad (4.1)$$

式中变换函数为 $T(r)$ ，该函数必须满足下列条件：

- (1) $T(r)$ 是在 $0 \leq r \leq 1$ 区域内的单调函数，并且单调递增。
- (2) $T(r)$ 在 $0 \leq r \leq 1$ 区域内并且必须满足 $0 \leq T(r) \leq 1$ 。

其中条件 (1) 保证了图像的灰度级从黑到白的变化次序, 条件 (2) 保证了变换后图像的像素灰度级仍然可以在允许的灰度级范围中。图像逆变换函数 $T^{-1}(s)$, 也同样需要满足上述条件^[17]。

若用 $p_r(r)$ 表示原图像的灰度级概率密度函数, $p_s(s)$ 表示变换后图像的灰度级概率密度函数, 若已知 $p_r(r)$ 和 $T(r)$, 并且 $T^{-1}(s)$ 也是单调递增的, 根据概率论中所学的知识, 则 $p_s(s)$ 可以通过如下公式求出:

$$p_s(s) = p_r(r) \frac{dr}{ds} = T^{-1}(s) \quad (4.2)$$

通过上面的推导不难看出, 灰度直方图修整法对图像进行增强的实质就是选用合适的变换函数 $T(r)$ 来变换修整图像灰度级概率密度函数 $p_k(r)$, 从而得到灰度级的概率密度为 $p_s(s)$ 分布的新图像。

直方图的均衡化是直方图修整法的常用方法。采用均衡化的方法的主要原因是使图像熵尽可能大, 即为了能从图像中获得尽量多的信息量, 这样就常常要求 $p_s(s) = \text{常数}$, 这样就变成了直方图均匀化, 也就是说对于那些图像中所有灰度出现频率相等的图像的信息量最大。因此, 可以得出结论对直方图进行均衡化意味着令 $p_s(s) = 1, s \in [0, 1]$, 由公式 4.2 可得

$$ds = \frac{p_r(r)dr}{p_s(s)} = \frac{p_r(r)}{1} = p_r(r)dr \quad (4.3)$$

从而

$$s = T(r) = \int_0^r p_r(w)dw \quad (4.4)$$

从上式可知 $T(r)$ 为分布累计函数。对数字图像处理需引出离散形式的公式:

$$p_r(r_k) = \frac{n_k}{n} \begin{cases} 0 \leq T_k \leq 1 \\ k = 0, 1, \dots, L-1 \end{cases} \quad (4.5)$$

式中 L 是待变化的图像灰度级数; 其中 n_k 为待变化图像中第 k 个灰度级的像素数; $p_r(r_k)$ 为第 k 个灰度级数出现的概率; n 是图像中包含的总像素数。

根据式 4.3 和式 4.5 可变得变换函数进行直方图均匀化处理如下式:

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n} \begin{cases} 0 \leq T_k \leq 1 \\ k = 0, 1, \dots, L-1 \end{cases} \quad (4.6)$$

本文充分挖掘 FPGA 内部资源，用 FPGA 内部可编程资源通过 VHDL 语言编写光线补偿模块如图 4.2，较传统的软件方法提高上百倍的速度。

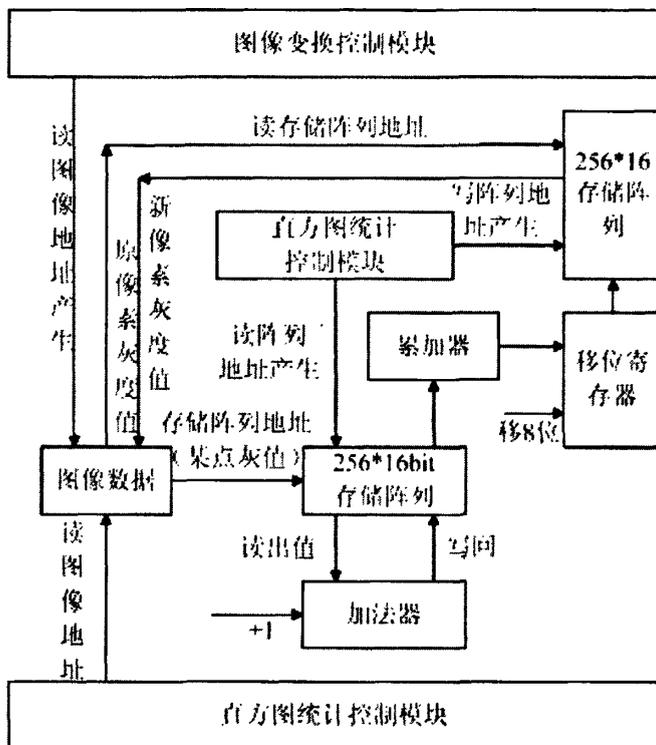


图 4.2 光线补偿模块原理图
Fig. 4.2 Light compensation Module diagram

对于待处理的图像，如果考虑所有象素灰度值集中在一个灰度级上的极端情况，那么存储单元大小应为是 256×256 即 2^{16} ；图像中有 256 个灰度级，因此本文需构造一个 256×16 位的存储阵列用于存储统计结果。这里将输入图像象素的灰度值作为地址，存储单元的内容值为图像中该灰度值象素的总个数。对每一个象素处理时，都需要将该象素点的当前统计个数从存储阵列中读出，送到加法器进行加 1 运算后再被送入至寄存器，然后将寄存器中的数据回写至该存储单元。本文采集的图像为 256×256 大小，变化后的灰度值为：

$$Y_k = 255 \sum_{j=0}^k \frac{n_j}{256 \times 256}, \text{ 即 } Y_k \sim \frac{1}{256} \sum_{j=0}^k n_j \quad (4.7)$$

从上式看到变换过程包括累加及除法过程。当一帧图像的数据统计结束后，直方图统计结果被存入存储阵列中，对得到的存储阵列进一步运算得到累加结果，接着地址发生器产生 0 至 Y 的地址（ Y 为当前灰度，对应灰度级 0 至 Y ）。地址值从 0 开始一直变换到 Y ，通过 16 位并行加法器对从每个地址读出数据进行累加，累加的结果被送入数据寄存器。除以 256 在硬件上可通过右移 8 位来实现。最后的工作是将变换后的灰度值存入新的存储空间，将输入图像灰度值作为地址，从查找表读出数据作为新图像的灰度，处理后的图像即为均衡化后的图像。下图为经直方图修正法进行光照补偿前后的两幅图片。



图 4.3 光照补偿效果图

Fig. 4.3 Illumination compensation effect diagram

4.3 肤色检测粗定位人脸区域

4.3.1 肤色模型的选择

近年来色彩信息在计算机视觉研究领域得到越来越广泛的应用，国内外很多学者都对基于肤色的人脸检测作出了深入的研究^[18]。基于肤色的人脸检测的优点是它不依赖于面部的细节特征，并且能较好的适应头部旋转、表情变化、姿态变化等，具有较高的鲁棒性^[19]。此外对于具有复杂背景的彩色图像特别是在人脸个数未知的情况下，根据肤色和大多数背景物体的颜色具有一定的差异这一特性，可以利用肤色信息可以将脸部区域从大量复杂背景中快速地分割出来。

在计算机系统中，常用的颜色模型有以下几种。

(1) RGB 格式

RGB 格式对应人类视觉的三基色，即红、绿、蓝三个波段，该格式是对人眼的光谱量化性质的近似。利用 RGB 三基色能自然表征多种颜色，并且多数的数字图像采集

设备都是以 CCD 技术为核心的，都是直接感知色彩的 RGB 三个分量，因此三基色模型已发展成为图像成像、显示、打印等设备的基础，具有非常重要的作用。

(2) HIS 格式

HIS 格式是人类观察色彩的方式的反映，这种格式也有很有利于图像的处理。其优点是该格式将亮度与反映色彩本质特性的色度及饱和度分开。在一些图像处理中若能将亮度分量从色彩中提取出去，而只通过反映色彩本质特性的色度和饱和度进行聚类化分析，就会获得比较好的处理效果，HIS 格式的这一特性使其在色彩图像处理和计算机视觉的研究中得到广泛的应用。

(3) YCbCr 格式

YCbCr 格式的色彩空间是以演播室质量指标为目标的 CCIR601 编码方案中采用的颜色模型，该格式被广泛应用于电视的色彩显示等诸多领域。其中 Y 代表亮度分量，Cb 代表蓝色色度分量，而 Cr 代表红色色度分量。YCbCr 格式还具有 HIS 格式中亮度与色度分量分离的优点。

目前常采用的 YCbCr 作为肤色统计空间，YCbCr 颜色空间的三个分量可以直接由 RGB 经过线性变得到，这样就避免了非线性空间带来的奇异性。YCbCr 色彩空间与 RGB 色彩空间的变换关系如下面公式所示^[20]：

$$\begin{bmatrix} Y \\ Cb \\ Cr \\ 1 \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1440 & 0 \\ -0.1687 & -0.3313 & 0.5000 & 128 \\ 0.5000 & -0.4187 & -0.813 & 128 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix} \quad (4.8)$$

YCbCr 色彩空间较其他颜色空间具有以下优点：首先，该空间的构成原理与人类视觉感知过程相类似；其次，YCbCr 色彩空间具有将色彩中的亮度分量分离出来的优点；再次，YCbCr 与 RGB 之间的换算关系简单，避免了非线性空间的奇异性。再则，其色彩空间是离散的，采用该空间易于实现聚类算法。最后，大量统计表明肤色在 YCbCr 色彩空间的聚类性较好。

本文在 YCbCr 色彩空间通过肤色模型分割脸部的，脸部区域的定位由模型建立、模型运用两部分组成。模型的建立主要是通过对大量肤色像素集进行统计分析进而确定模型中的参数。模型运用就是使用建立好的模型进行脸部区域提取。国内外大量的研究表明，尽管不同种族、不同性别、不同年龄的人脸肤色看上去有所不同，但这种不同主要是集中在亮度上的，在去掉亮度的色度空间中，不同人脸肤色分布都具有明显的聚类性，因此可以根据这个原理分割脸部区域^[21]。肤色模型大体上都分为三种：高斯模型、神经网络模型和色彩区域模型。

(1) 高斯模型

肤色分布近似于高斯分布是高斯模型的理论基础。根据这个特点,利用二维高斯概率公式即可计算出图像中每一个像素是否属于肤色的概率,并可以将这些概率值构成一个肤色概率图,概率图中具有较高肤色概率值的像素所构成的区域就是初定位出的可能人脸区域。

(2) 神经网络模型

神经网络模型是通过大量的肤色和非肤色样本训练学习得到一个分类器,并通过分类器来判断像素是否为肤色,完成脸部区域的初步定位。

(3) 色彩区域模型

色彩区域模型主要是利用了人脸肤色在色彩空间的聚类性,将满足条件的区域标记为肤色区域。首先通过统计确定模型,然后利用该模型判别新的像素是否为肤色。对于任意一幅的图像,如果某个像素满足给定的条件即为肤色,否则就是非肤色。国外学者 Anil K.Jain 从 Heinrich-Hertz-Institute(HHI)图像库中的 137 幅图像中手工选取了 853571 个肤色像素点,并将这些像素点绘制在 YCbCr 空间和它的二维投影子空间 CbCr 空间中^[22]。该方法中的阈值选择不当会使肤色检测率下降,因此使用色彩区域模型的难点在于如何确定阈值。

综合比较上述三种模型,色彩区域模型方法计算量小、算法简单,速度快,适于本文中的肤色检测。高斯模型,肤色的检测率较高,误检率小,但是模型参数的难确定,检测速度慢,因此不适合快速的肤色检测。神经网络模型同样存在计算复杂和参数难确定的缺点,并且当脸部区域与类肤色背景相连的情况发生时,极易被定位为一个整体。

4.3.2 肤色建模标定人脸区域

对于彩色视频图像,在确定肤色模型后就可以进行肤色信息的检测。通过建立好的肤色模型,将一幅彩色图像转变为二值图像,其中 255 表示皮肤区域,0 表示非皮肤区域。Anil K.Jain 经过对大量肤色像素的分析得出结论,肤色在 CbCr 色度空间有类似椭圆的聚类特性。

通过肤色建模的方式定位人脸的步骤如下:

(1) 非线性转换

研究表明亮度 Y 在 YCrCb 色彩中并不是完全独立于色度信息而存在的,因此肤色的聚类区域会随 Y 的不同而呈非线性变化。研究还表明色度值对亮度值总是存在着一定的非线性依赖关系,这个特性会给图像的检测造成很大的影响^[22],所以本设计首先将 YCrCb 空间进行一次非线性的转换,通过这次转换来消除色度对亮度的依赖关系。

非线性转换公式如下所示:

$$C_i(Y) = \begin{cases} (C_i(Y) - \bar{C}_i(Y)) \frac{W_{C_i}}{W_{C_i}(Y)} + \bar{C}_i(Y) & Y < K_l \text{ 或 } Y > K_h \\ C_i(Y) & Y \in [K_l, K_h] \end{cases} \quad (4.9)$$

其中 $i = b, r$ 。

$W_{C_i}(Y)$, $\bar{C}_i(Y)$ 的公式如下所示:

$$W_{C_i}(Y) = \begin{cases} WL_{C_i} + \frac{(Y - Y_{\min})(W_{C_i} - WL_{C_i})}{K_l - Y_{\min}} & Y < K_l \\ WH_{C_i} + \frac{(Y_{\max} - Y)(W_{C_i} - WH_{C_i})}{Y_{\max} - K_h} & K_h > Y \end{cases} \quad (4.10)$$

其中 W_{C_i} , WL_{C_i} , WH_{C_i} , K_l , K_h 为常数, $Y_{\min}=16$, $Y_{\max}=235$, $i=b, r$ 。

$$C_b(Y) = \begin{cases} 108 + \frac{(K_l - Y)(118 - 108)}{K_l - Y_{\min}} & Y < K_l \\ 108 + \frac{(Y - K_h)(118 - 108)}{Y_{\max} - K_h} & K_h < Y \end{cases} \quad (4.11)$$

$$C_r(Y) = \begin{cases} 154 - \frac{(K_l - Y)(154 - 144)}{K_l - Y_{\min}} & Y < K_l \\ 154 - \frac{(Y - K_h)(154 - 144)}{Y_{\max} - K_h} & K_h < Y \end{cases} \quad (4.12)$$

(2) 人脸肤色颜色建模

所谓建模就是指根据已知的对象特征来为对象建立一个模型, 通过这个模型来对对象进行判断、检测、识别等。人脸建模通常可分为色彩建模和几何建模两种方式, 其中色彩建模的模型比较简单并且执行速度较快, 经非线性变换后的肤色区域在 $CrCb$ 空间近似一个椭圆, 实现人脸色彩建模的公式如 4.13 式中的 Cr , Cb 是经上述非线性转换的 $Cr(Y)$, $Cb(Y)$ 。

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} Cb - c_x \\ Cr - c_y \end{bmatrix} \quad (4.13)$$

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1 \quad (4.14)$$

其中 $a, b, ec_x, ec_y, c_x, c_y, \theta$ 为系数常量, x, y 为中间变量, $a = 25.39$, $b = 14.03$, $ec_x = 1.60$, $ec_y = 2.41$, $c_x = 109.38$, $c_y = 152.02$, $\theta = 2.53$ 。

这种几何模型相对于色彩模型来说具有实现简单，匹配速度快，精度高等优点。模型通过计算离心率来查看相似度，通过相似度标记处肤色像素点，分割出皮肤区域，本设计使用 255、0 分别代表皮肤区域和非皮肤区域，得到经肤色建模后的二值图像。本文在 PowerPC 上编写 C 语言程序完成肤色建模。

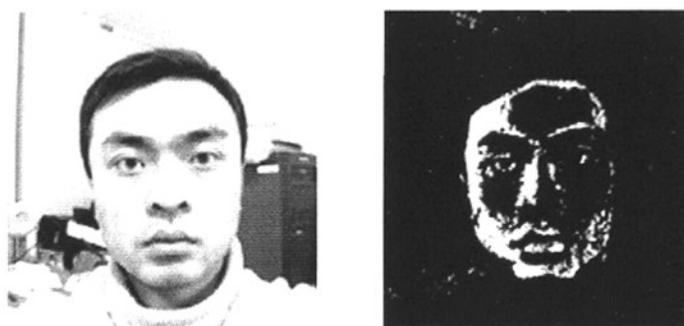


图 4.3 肤色建模示意图
Fig. 4.3 Face color modeling diagram

4.4 数学形态学定位脸部区域

上一节已经对人脸进行了肤色建模，得到了人脸的二值图像，为了定位人脸需要使人脸二值区域更好的连通，这里应用到了数学形态学的方法。

数学形态学 (Mathematical Morphology) 是一门诞生于 1964 年的新兴学科，它是法国的科学家在研究岩石结构时建立的一门学科。数学形态学的基本思想是使用一定形态的结构元素去提取或量度图像中的对应形状从而达到对图像识别和分析的目的。通过使用数学形态学的应用可以简化图像数据，保持图像的基本形状特性，并且除去不相干的结构 [23]。

(1) 图像的腐蚀运算

集合是二值形态学中的运算对象，在实际运算中涉及到的两个集合认为是不对等的。若假设 A 为图像集合，假设 B 为结构元素，数学形态学运算就是使用 B 对 A 进行的操作。实际上结构元素本身也是 1 个图像集合。运算时需要为每个结构元素指定 1 个原点，该原点即为结构元素参与数学形态学运算的参考点。

腐蚀的运算符为 θ ，使用 B 对 A 进行腐蚀可写作 $A\theta B$ ，定义为： $A\theta B = \{a | B_a \subset A\}$ 该式的含义是把结构元素 B 平移 a 后得到 B_a ，若 A 中包含 B_a ，则记下该 a 点，所有满足上式的 a 点组成的集合即被称为 A 被 B 腐蚀的结果，其过程如图 4.4 所示。图中 A 是

待处理的对象, B 是所选结构元素, 原点为 B 的圆心。则对于任意一个在阴影部分的点 a , B_a 都包含于 A , 所以 A 被 B 腐蚀的结果即为阴影部分。阴影部分在 A 的范围之内, 就像 A 被剥掉了一层似的, 这就是该操作被称为腐蚀的原因。

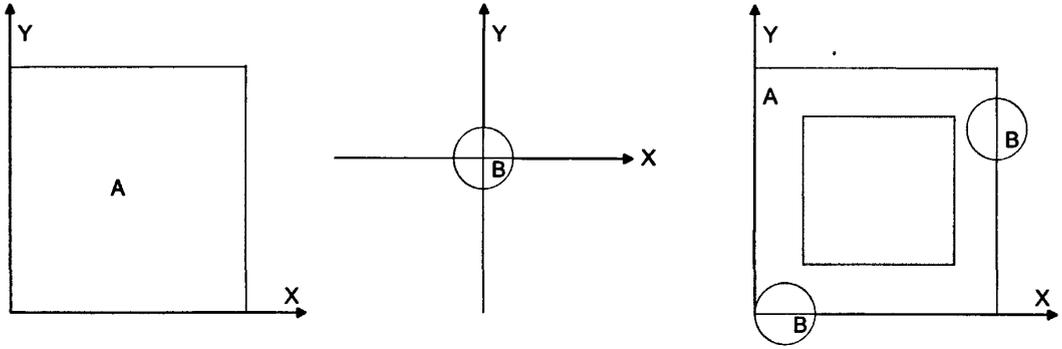


图 4.4 腐蚀运算示意图
Fig. 4.4 Corrosion computing diagram

二值图像的腐蚀的过程如图 4.5 所示: 左边是待处理的二值图像 A (其中要处理的是黑色图像即黑点), 结构元素为 B , 拿 B 的原点和图像 A 上的每个黑点进行对比, 如果 B 上的所有黑点都在 A 的黑点范围内, 则保留该点, 否则去掉; 右边是经腐蚀运算后的结果。结构元素 B 以及其原点的选取决定了二值图像的腐蚀效果。实际应用中往往采用线性的结构元素从水平和垂直方向对图像进行腐蚀处理。

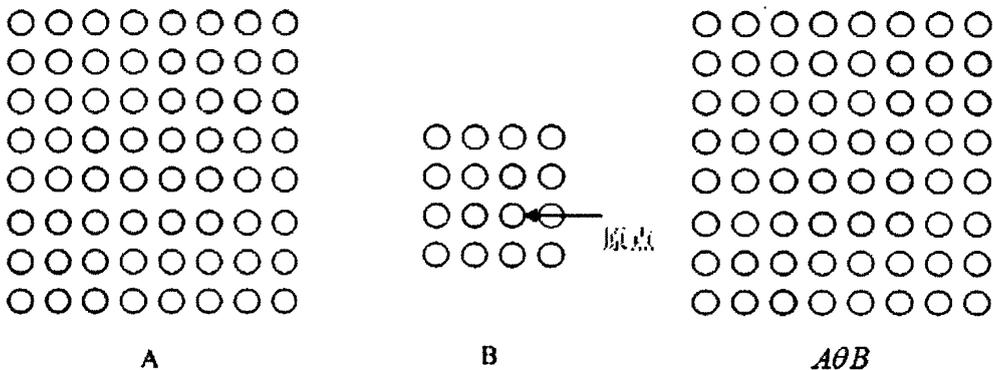


图 4.5 二值图像腐蚀的过程
Fig. 4.5 The process of corrosion of binary image

(2) 图像的膨胀运算

膨胀是腐蚀的对偶运算，设 A 为待处理的图像， B 为所选的结构元素，通过结构元素 B 对图像 A 进行膨胀处理的示意图如图 4.6 所示。

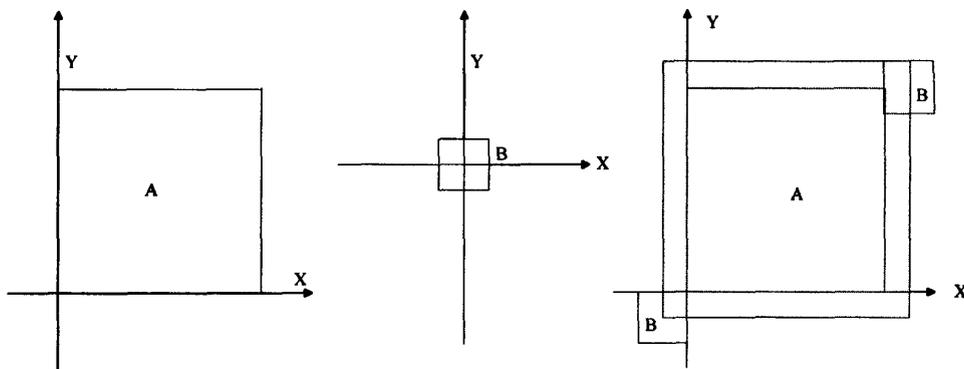


图 4.6 膨胀运算示意图
Fig. 4.6 Dilation computing diagram

膨胀的运算符为 \oplus ，使用 B 来膨胀 A 可以写作 $A \oplus B$ ，其定义为： $A \oplus B = \{a | B_a \uparrow A\}$ ，该式表示将结构元素 B 平移 a 后得到 B_a ，若 B_a 击中 A ，则记下该 a 点。满足上述定义的 a 点组成的集合即为 A 被 B 膨胀的结果。图 4.6 中 A 表示的大正方形为被处理的对象， B 所示的小正方形为结构元素，原点为 B 的中心。对于任意一个在阴影部分的点 a ， B_a 击中 A ，所以阴影部分即为 A 被 B 膨胀的结果。

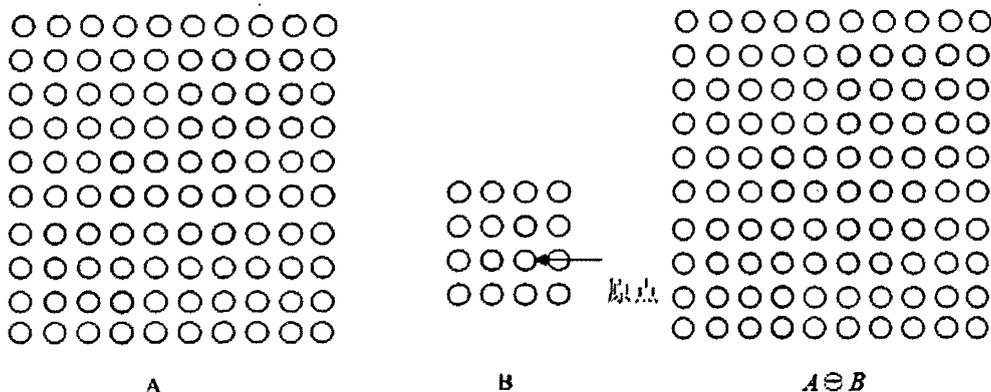


图 4.7 二值图像膨胀的过程
Fig. 4.7 The process of Dilation of binary image

二值图像的膨胀过程如图 4.7 所示：被处理的图像为 A ，结构元素为 B ，其原点已经标出。膨胀的具体步骤是，拿 B 的原点和 A 上的每一个点及 A 周围的点进行比对，只要 B 中有一个点落在 A 的范围内，则该点就为黑。

(3) 图像开、闭运算

膨胀和腐蚀可以级连使用。可先对图像进行腐蚀然后膨胀其结果称为开运算，或先对图像进行膨胀然后腐蚀其结果称为闭运算。无论开运算还是闭运算，其膨胀和腐蚀使用的是同一个结构元素。

开运算具有消除细小物体，分离纤细处的物体以及平滑较大物体边界的作用。开运算的运算符为 \circ ， A 用 B 来作开运算写作 $A \circ B$ ，定义为 $A \circ B = (A \ominus B) \oplus B$ 。

闭运算具有填充物体内部细小空洞，连接邻近物体和平滑边界的作用。闭运算的运算符为 \bullet ， A 用 B 来作闭运算写作 $A \bullet B$ ，其定义为 $A \bullet B = (A \oplus B) \ominus B$ 。

本设计对经肤色建模后的图像进行膨胀腐蚀，用于补偿由于各种原因引起的人脸提取图像点的离散化问题，使之更好的连通，如图 4.8 所示。



图 4.8 原图像、膨胀、腐蚀

Fig. 4.8 Original image, dilated image, corroded image

为了使人脸区域更好的连成一体，再次对图像的再进行膨胀、腐蚀操作，去除较小的白色区域，人脸部即是白色方框所示的区域。如图 4.9 所示。



图 4.9 再次膨胀, 再次腐蚀, 定位脸部

Fig. 4.9 dilated again, corroded again, face location

5 眼睛检测与定位

眼睛是脸部最复杂的器官,眼睛的这种复杂性给基于图像处理方法的人眼定位和特征提取带来一定的难度。本文在第四节已经粗定位出了脸部的区域,这一节本文将在这个区域中精确定位眼球,为使用 PERCLOS 评测方法和计算眨眼频率提供数据支持。

5.1 眼睛检测方法的研究与设计

对于本文的疲劳检测系统选择一种可靠的眼睛检测、定位的算法至关重要,算法的选择关系到系统的实时性,安全性,准确性,目前常用的眼睛检测定位算法包括 Hough 变换法,变形模板法,对称变换法。

(1) Hough 变换法: Hough 变换法对随机噪声不敏感,常用于检测定位直线或解析曲线,具有较强的抗干扰能力,并且允许待检测的边界曲线不连续。因此当图像中存在部分重叠或遮挡导致的物体表示有限时,都可以用 Hough 变换进行图像分割^[24]。

Hough 变换公式如下所示:

$$H(\Omega) = \sum_{i=1}^N p(X_i, \Omega) \quad (5.1)$$

其中,

$$p(X, \Omega) = \begin{cases} 1, \forall (X, \Omega): \{\Lambda: f(X, \Lambda) = 0 \cap C_{\Omega} \neq \emptyset\} \\ 0, \text{otherwise} \end{cases} \quad (5.2)$$

$H(\Omega)$ 是 Hough 变换在 Ω 处的值, C_{Ω} 为参数空间中以 Ω 为中心且大小受限的累加器单元, $f(X, \Lambda) = 0$ 为待测物体的参数化约束方程, $\{X_1, X_2, \dots, X_H\}$ 是图像空间中的特征点。

Hough 变换法利用眼球的圆形特征检测眼睛,在边缘检测后用 Hough 变换寻找到眼球。由于眼球的上半部分较多地被眼睑覆盖,有学者使用检测下半圆的方法,该方法更加可靠省时。若图像空间为 (i, j) , i, j 分别表示眼球圆心的行和列。三维变换空间为 (i_o, j_o, R) , 其中 i_o, j_o 分别代表眼球圆心的行和列, r 为半径。下半圆的数学公式为:

$$i = i_o + \sqrt{r^2 - (j - j_o)^2} \quad (5.3)$$

Hough 变换抗干扰能力强,缺点是需要搜索整个参数空间 (i, j, R) 来寻找未知半径的圆,搜索空间非常巨大。国内外的研究人员针对该问题一直在做相关方面的研究,利用边缘方向的先验知识信息,降低计算负担^[24]。

(2) 变形模板法: 根据所选模板的维数可分为二维可变形模板和三维可变形模板。眼睛定位时, 不断改变眼睛模板的眼球中心位置参数 (x,y) 和尺度参数 S , 当模板与人脸图像取得最佳匹配时, 其中心位置 (x,y) 就是眼睛在人脸图像中的位置, 其对应的尺度 S 就是眼睛的大小。对于正面人脸图像基于二维可变形模板的方法提取效果很好, 但二维模板不适用于具有垂直旋转角度的人脸图像。三维可变形眼睛模板匹配技术可以解决旋转问题。眼睛模板源于人脸的一般三维模型。三维可变形眼睛模板用数学公式来表示为:

$$Eye3D = EyeModel(X, S, \theta, X_{eye}, Y_{eye}, Z_{eye}) \quad (5.4)$$

通过变换公式中的参数可以产生各种形状的三维模板。模板匹配时需将三维模板映射成二维眼睛模板, 再针对二维眼睛模板采用最小值搜索算法求得最佳匹配。

(3) 对称变换法

对称性分析可以应用于目标识别、视觉监视和形状表示等问题的研究中。点对称即中心对称和轴对称是目前研究较多的对称变换法, 人脸的嘴巴、鼻子、眼睛都具有很强的轴对称性。有学者根据双眼具有的高对称性, 利用广义对称变换来求得具有高对称性的点即为双眼的位置^[25]。广义对称变换法将图像中所有像素点同等对待, 为每个像素点都计算其对称值^[26]。

以上这些常用的眼睛定位方法各具特色, 适用的场合各不相同。霍夫变换法通过眼睑形成的椭圆特征或检测瞳孔的圆形特征进行定位, 但计算量大, 速度慢。变形模板法能够准确地确定人眼形状及大小, 但收敛速度过慢; 对称变换同样存在运算量大的缺点, 无法满足本文实时性的要求。

对于本文的疲劳检测系统眼睛定位算法需要满足实时性的要求, 并且只要求定位眼睛, 不需要涉及特定的表情, 上一章的人脸检测定位操作已经初步标定除了人脸区域, 首先结合脸部三庭五眼的几何特征及人眼的灰度信息初步划定眼部区域并通过积分投影方法粗定位人眼, 然后通过区域生长方法精确定位眼球, 眼睛检测定位流程如图 5.1 所示。



图 5.1 人眼定位流程
Fig. 5.1 Eye location flow

5.2 眉眼区域的定位

人脸的眼睛、眉毛、嘴巴、下颚等区域较周围区域暗，因此对人脸区域作灰度的水平积分投影在这些位置处会出现灰度局部极小值。脸部区域的这种特征如图 5.2 所示。可以利用灰度水平积分投影曲线各极小值点之间存在的位置关系，粗定位眉眼区域。

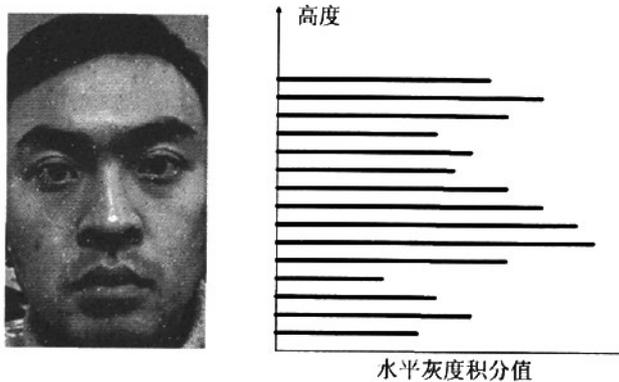


图 5.2 人脸图像及其水平灰度积分投影示意图
Fig.5.2 face image and its gray integral projection diagram

基于水平灰度积分投影的人脸特征提取是一种常用的方法,这种方法计算简单,适于粗定位,国内外很多学者已将投影函数成功运用于定位面部特征^{[27][28]}。后来又有学者提出改进的方法,利用灰度方差投影图寻找特征点,但这种改进方法的缺点是计算复杂度较大。另外有学者提出利用边缘检测结合积分投影来定位眼睛,这种方法在有些方面会提高定位准确性,但定位极易受遮挡物边缘的影响,并且边缘提取减弱了原图的灰度信息,定位效果有时也会受到一定的影响。

对于本设计中的疲劳检测系统为减少计算量提高实时性,这里结合人脸部的几何特性及图像水平灰度积分投影粗定位眼睛区域。首先,根据人脸“三庭五眼”^[29]的面部几何比例特征(眉心以上部分脸部占整个脸部的 1/3,眉心至鼻头占整个脸部的 1/3,鼻头至下额占整个脸部的 1/3)初步裁剪图像,只去除顶部的 1/5 和底部的 1/3,见图 5.3。这种根据面部几何特征的切割大大减少计算量。

在进行完第一次切割后,需要对这幅图像进行水平灰度积分投影,若第一次分割提取的眉眼区域图像的大小为 $M \times N$, 则水平灰度积分投影函数为:

$$p_x(y) = \sum_{x=1}^M f(x,y) \quad (5.5)$$

其中, x 分别代表某一个像素点的横坐标和 y 代表该点的纵坐标, $f(x,y)$ 该点的像素值。



图 5.3 脸部区域第一次分割

Fig.5.3 Face region's first division

人的脸部特征是在眉毛与眼睛处的会存在两个区间极小值,这里为了减小计算量将第一次脸部分割后的图像进行灰度积分投影,得到区间极小值与区间次极小值,这两个

区间极小值不存在绝对的大小关系,用两个极小值的中点纵坐标来作为眉眼区域中心坐标,该位置向上向下各取脸部区域高度的 1/9,得到如图 5.4 所示眉眼区域。



图 5.4 脸部区域第二次分割
Fig.5.4 Face region's second division

5.3 眉眼区域的中值滤波

一般的,视频系统获取的图像在形成、传输、接收和处理的过程中,通常会受到来自外部和内部干扰。如脉冲噪声干扰,这类噪声主要来源于图像传感器、信道传输、A/D 转换等、以及数字化过程中的量化噪声,另外还有来自的电磁波干扰等。这些噪声在很大程度上影响了图像质量,对图像分割、目标检测、特征提取、图像识别等处理造成了严重的影响,因而消除噪声是图像处理中非常重要的一个环节。滤波是去除噪声的主要方法,滤波操作需要的既能去除噪声,又能保持图像的细节。中值滤波是一种较好的滤波方法,非常适合作为本设计的预处理操作。中值滤波算法比较固定,没有特别复杂的计算,FPGA 处理速度快,因此中值滤波很适合在 FPGA 上实现。本设计中使用了一种快速的在 FPGA 上实现的中值滤波算法,提高了滤波的速度。

中值滤波首先需要通过比较找到中值,比较通过排序完成,其中冒泡排序最为常用,比较次数是影响速度的重要因素。假设滤波窗口为 $n \times n$,如果通过传统的排序运算来求数据的中值时,当求出第 $(n \times n + 1)/2$ 个值时比较即可以停止。比较次数为:

$$(n^2 - 1) + (n^2 - 2) + \dots + (n^2 - 1 - \frac{n^2 - 1}{2}) = \frac{3}{8}(n^4 - 1) \quad (5.6)$$

随着滤波窗口尺寸的增大,运算次数急剧增加这就使得处理速度变得很慢。有学者提出的一种用于二维中值滤波的改进算法能很好地解决了这个问题^[30]。本文将该方法在 FPGA 上进行实现,并行化的比较进一步提高了中值滤波的速度。

FPGA 硬件资源丰富,用户可根据实际需要定制特定数字功能模块,本文通过 VHDL 语言编写中值滤波模块,以 3×3 的窗口为例,文中使用一种快速并行中值滤波方法如图 5.5。其设计思想是将二维中值滤波阵列降维至一维,步骤如下:首先对每一列进行升序排序,显然比较次数为 $3 \times 3 = 9$ 次,然后取第一行中的最大值、第二行中的中间值、第

三行中的最小值，比较总次数为 $3 \times 3 = 9$ 次，最后取这三个值的中值，所需要的比较次数为 3 次，共需要 21 次比较的运算，次数减少了不少。

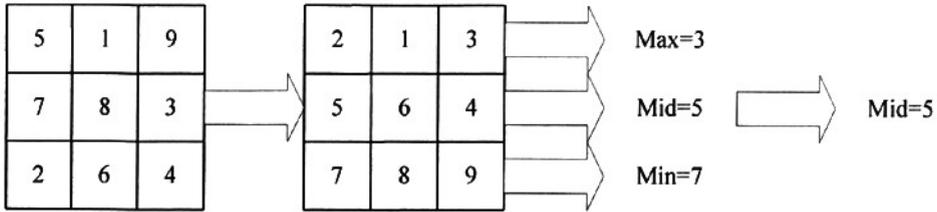


图 5.5 快速中值滤波法
Fig.5.5 Fast median filtering method

对于 $n \times n$ 的滤波窗口，含有 $n \times n$ 个像素的二维阵列降维为 $2n+1$ 个仅含 n 个像素的一维阵列取中值。将算法比较次数如下公式：

$$\frac{n \times (n-1)}{2} \times n \times 2 + \frac{n \times (n-1)}{2} = \frac{n \times (n-1) \times (2n+1)}{2} \quad (5.7)$$

快速中值滤波只需 $n(n-1) \times (2n+1)/2$ 次运算，所需的运算次数大大减少，并行化的处理使速度明显提高。对于 5×5 窗口，比较的次数为 110 次，比原来的次数减少了一半。对于 9×9 窗口，比较次数是 684，只为传统方法的 1/4 多一点。如表 5.1。

表 5.1 算法比较
Tab. 5.1 Comparison of two algorithms

	3*3	5*5	7*7	9*9	11*11
传统方法	30	234	900	2460	5490
快速方法	21	110	315	684	1265

经快速中值滤波方法滤波后的眉眼区域图像如图 5.6 所示。



图 5.6 眉眼区域中值滤波效果图
Fig.5.6 Eyebrow eye area after median filtering

5.4 眼球的精确定位

5.4.1 阈值分割模块的设计与实现

阈值分割法是一种基于区域分割的图像处理技术，该方法的基本原理是：通过设定的不同的特征阈值，将图像像素点分成若干类。常用的特征通常来自原始图像的灰度值或彩色特征以及由原始灰度或彩色值变换得到的特征。设原始图像为 $f(x,y)$ ，根据阈值 T 对图像进行二值化变换即确定当前像素是背景点还是前景点。在经粗定位后的眼睛窗口，分析眉眼窗口直方图，本设计为了消除光照影响采取自适应阈值分割法，设计中通过 VHDL 语言编写硬件模块完成眼部区域阈值分割，算法的基本步骤是：首先，计算眼睛窗口灰度直方图，通过硬件实现；然后，求取峰值 T_{max} ，取阈值为 $T=T_{max}-30$ ；最后，把像素值与阈值比较，二值化眉眼窗口。

眉眼区域图像尺寸小于 64×256 ，考虑到所有象素灰度值集中在一个灰度级上的极端情况，本文将每个存储单元设计为 14 位，图像共有 256 个灰度级，故需构造一个 256×14 位的存储阵列用来存储统计结果。将图像象素的灰度值作为输入地址，该灰度值的象素个数作为存储单元的内容。当读取到一个象素值后，存储阵列的相应地址的数据被读出，送到加法器加 1 后再写入该存储单元。一帧结束后，存储阵列中的数据即直方图统计结果，如图 5.7 所示。

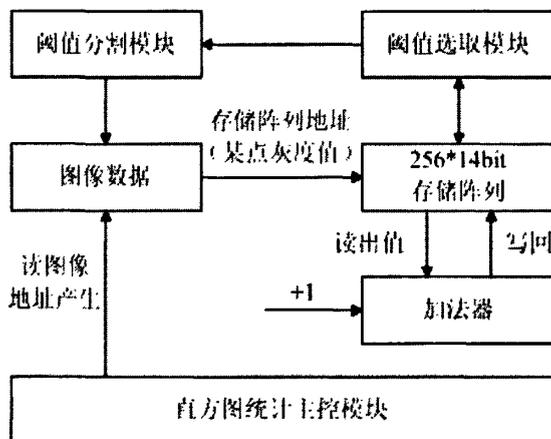


图 5.7 自适应阈值分割模块

Fig.5.7 Self-adaptive threshold segmentation module

阈值选取模块从直方图统计结果选出阈值，阈值为 $T=T_{max}-30$ ，阈值分割模块根据选取的阈值将大于阈值的点置为 255，将小于阈值的点置为 0，形成一幅二值图像。

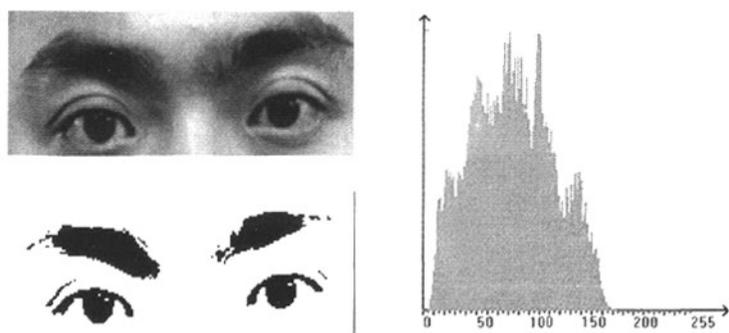


图 5.8 自适应阈值分割眉眼区域

Fig.5.8 Binarize eyebrow and eye area image with adaptive threshold method

5.4.2 区域生长定位眼球中心

经自适应阈值分割后的图像在眉毛、眼球处呈黑色，周围肤色区域呈白色。接下来要想得到眼球的坐标位置、面积大小就需要对二值化的图像进行处理，可以通过连通区域生长来确定眼部区域的大小及中心坐标。本设计中通过区域生长精确定位眼球。

将具有相似性质的像素集合起来构成区域是区域生长的基本思想。首先需要对待处理的区域找一个种子像素点作为区域生长的起点，然后将种子像素周围邻域中与种子像素有相同或相近性质的像素合并到种子像素所在的区域中。每当新的点被合并后都再用新的区域重复这一过程，直到没有可接受的邻近点时区域生成过程结束，这样一个区域就生长完成。灰度级、组织、彩色、梯度等都可以作为相似度准则，相似性通过所确定的阈值来判定。

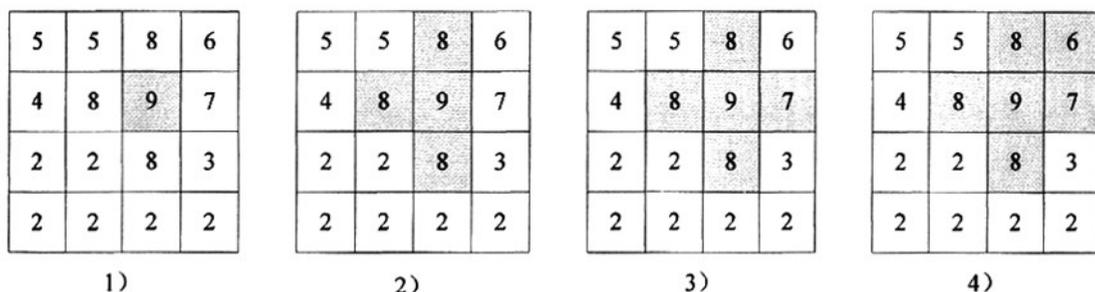


图 5.9 区域生长示例

Fig.5.9 Regional growth examples

图 5.9 的相似性准则为邻近点的灰度级与物体的平均灰度级之差小于 2。图 5.9 中被接受的点和起始点均为灰色，其中 1) 图是输入图像；2) 图是第一次接受的邻近点；3) 图是第二次接受的邻近点；4) 图是从 9 开始生长的结果^[31]。

本设计中针对的是阈值分割后的二值图像进行区域生长，区域生长研究中三个非常重要的问题是种子点的选取，生长准则及终止条件。

初始点的选择通常是通过人工干预加入的交互信息即告诉计算机初始点的位置。初始点是该算法的关键所在。本设计使用螺旋队列的方法来寻找种子点。在算法的具体实现中将眼部区域分为左右两部分处理，找到每部分的中心 (x, y) ，在图 5.10 中 (x, y) 点即为方格图中心 1 的位置。按图 5.10 中螺旋队列中数字 1, 2, 3, 4, ……的顺序搜索区域生长的种子点。当搜索到的某点 P 对应点的 8 邻域灰度值均为 0 即可停止螺旋队列的搜索，然后将 P 作为区域生长的初始点种子点。

21	22	23	24	25
20	7	8	9	10
19	6	1	2	11
18	5	4	3	12
17	16	15	14	13

图 5.10 螺旋队列示意图

Fig.5.10 Spiral queue diagram

在本设计的二值化眼睛窗口中，只有黑白两种颜色，所以在确定生长准则时，是黑色则可以生长，在 PowerPC 上编写 C 语言程序实现区域生长算法，将眼部区域分为左右两个部分计算处理，每部分计算方法相同，精确定位双眼眼球如图 5.11 所示。

区域生长的具体步骤如下：

- (1) 通过螺旋队列方式寻找初始点 (x_seed, y_seed) ，将初始点加入空队列。
- (2) 判断队列是否为空。如果队列非空，以队头 (x_head, y_head) 为中心，考虑 $(currentx, currenty)$ 的 4 邻域像素点 (x, y) ，如果满足生长条件（黑色点为 255）就将 (x, y) 合并到这个区域内，同时将 (x, y) 加入队列尾部。如果为空跳至 (4)。

(3) 从队列中取出队头即当前处理点($currentx, currenty$), 并将该像素的横坐标值 $currentx$ 乘以像素值 $f(x, y)$ 累加, 将该像素的横坐标值 $currenty$ 乘以像素值 $f(x, y)$ 累加, 将该像素 $f(x, y)$ 累加, 回到步骤 (2)。

(4) 求的质心的坐标, 并记录这个坐标, 坐标如公式 5.8 所示。

$$\bar{x} = \frac{\sum_{x=0}^{j-1} \sum_{y=0}^{k-1} xf(x, y)}{\sum_{x=0}^{j-1} \sum_{y=0}^{k-1} f(x, y)}, \quad \bar{y} = \frac{\sum_{x=0}^{j-1} \sum_{y=0}^{k-1} yf(x, y)}{\sum_{x=0}^{j-1} \sum_{y=0}^{k-1} f(x, y)} \quad (5.8)$$

(5) 重复 (1) 至 (4), 直到图像中的每一个点(x, y)都有一个归属, 此时生长结束, 进入第 6 步。

(6) 根据连通区域的面积, 取出面积最大的两块区域。

(7) 根据所求得的重心坐标, 取重心在下的连通区域即为一侧眼球的位置, 重心在上的连通区域即为一侧眉毛。



图 5.11 通过区域生长精确定位眼球

Fig 5.11 precise positioning eye through regional growth method

6 疲劳检测的实现

6.1 眼部状态数值化

所谓眼部状态数值化就是采用数值的方式来量化眼睛的状态, 本文将上下眼睑高度和眼部区域面积作为疲劳评测的指标。

在上一章本文已经定位了眼球的中心位置即眼部区域重心坐标, 眼球的由上下眼睑覆盖, 因此可以通过上下眼睑的高度 H 量化眼睛睁开的程度。疲劳状态的数值化是以 H 作为根据的。系统对眼部图像进行眼睛状态检测, 图 6.1 是各种状态下的原始眼部图像、眉眼二值图像和眼球二值图像, H 为上眼睑最高点与下眼睑最低点之间的垂直距离, 单位为像素。由于人的头部可能有略微倾斜, 本文采用的计算 H 的方法如下: 首先, 分别计算每只眼睛的上下眼睑的高度 H , 计算方法是, 找到眼球重心, 重心坐标向左向右各 3 个像素区域内共有 7 条竖线, 计算每条竖线所包括的像素数, 取最大值即为眼睛上下眼睑高度。然后, 取两只眼睛眼睑高度的均值作为此时的上下眼睑高度 H , 作为后续疲劳分析的一个量化指标。

眼部区域面积为眼部二值图像中的黑色像素点的个数, 统计眼部二值图像的黑色点的像素总数记为 SUM , 这一步操作在 5.5.2 节区域生长精确定位眼球中计算出。图 6.1 为若干种眼部状态, 从左到右依次为睁眼、半睁、微睁、闭眼。

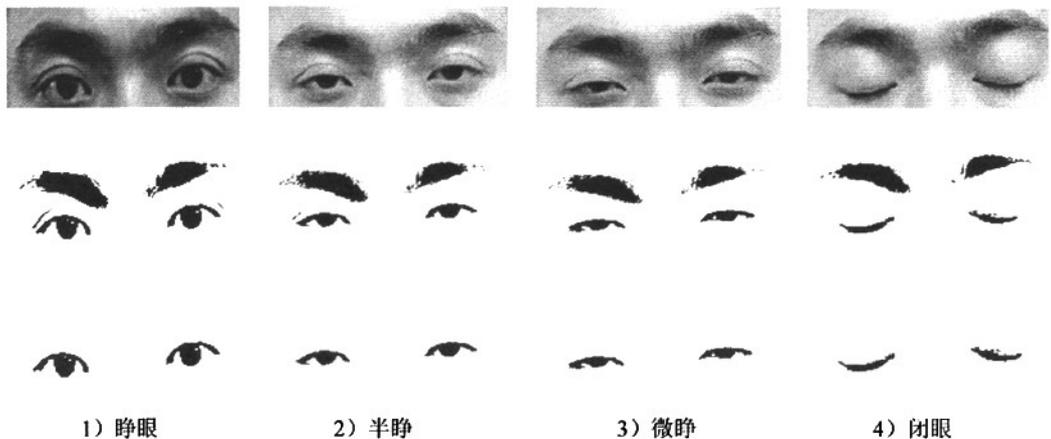


图 6.1 眼睛若干状态分析示意图

Fig.6.1 Some state diagram of eyes

6.2 疲劳检测方法的研究

目前基于视觉的驾驶疲劳判定方法主要有以下四种：

(1) 眨眼频率

所谓眨眼即眼睛从睁开到闭合再到睁开的过程。通常人眼每分钟眨动 10 至 15 次，每次眨眼间隔时间 4 至 5 秒，每次眨眼持续时间约为 0.2 秒。通过统计驾驶员在驾驶时的单位时间内眨眼的次数可以作为判定是否出现疲劳驾驶的方法。

(2) 平均眨眼持续时间

在疲劳时，人的眨眼持续时间会显著提高，眼睛中飞入异物等异常情况也会提高，这些都会影响驾驶员正常的驾驶。因此一些疲劳检测的研究中把眨眼持续时间作为判断疲劳的指标^[32]。

(3) 平均眼睛闭合的速率

通过计算一段时间内眼睛从最大程度睁眼到闭眼平均所需要的时间，如果这个时间大于某个阈值，则认为驾驶员处于疲劳驾驶的状态。

(4) PERCLOS 原理

PERCLOS (Percentage of Eyelid Closure Over the Pupil Time) 是指单位时间内眼睛闭合的时间。大量实验表明驾驶员眼睛闭合的时间越长，驾驶员疲劳感越强，因此 PERCLOS 已成为很多疲劳检测系统所采用的评测标准^{[33][34]}。

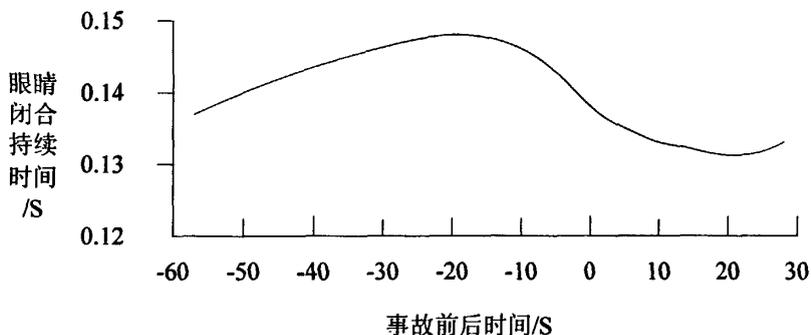


图 6.2 眼睛闭合时间与事故发生前后的时间关系

Fig.6.2 Relation between the time of eye closure and accidents

图 6.2 为眼睛闭合与事故发生的时间关系，从图中可以看出，眼睛闭合时间与疲劳程度有着极其密切的关系。美国联邦公路管理局 FHWA (Federal Highway Administration)

的一项研究表明 PERCLOS 值可以很好的作为车载的、实时的、非接触式的疲劳测评方法。如图 6.3 所示，通过测量出 t_1 - t_4 就能计算出 PERCLOS:

$$PERCLOS = \frac{t_3 - t_2}{t_4 - t_1} \times 100\% \quad (6.1)$$

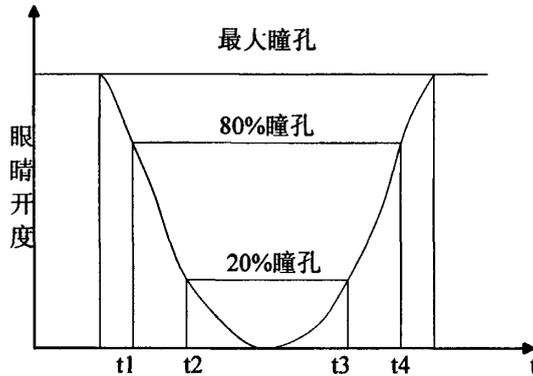


图 6.3 PERCLOS 测量原理

Fig.6.3 PERCLOS measuring principle

通常使用 PERCLOS 方法判断眼睛闭合有以下三种标准：P70 标准：该标准认为眼睑遮住瞳孔的面积超过 70%就计为眼睛闭合。P80 标准：该标准认为眼睑遮住瞳孔的面积超过 80%就计为眼睛闭合。EM 标准：该标准认为眼睑遮住瞳孔的面积超过一半就计为眼睛闭合。美国国家公路交通安全局（NHTSA）的实验数据表明 PERCLOS 中的 P80 与疲劳发展程度的相关性较好。P80 大于 0.4 时^[35]则认为驾驶员处于疲劳状态，需要对驾驶员提出相应警告^[36]。

6.3 基于 PERCLOS 与眨眼频率的疲劳检测方法

本文选择与疲劳程度相关性最好的 PERCLOS 原理作为疲劳分析方法；单纯用 PERCLOS 值来判断也不够准确，当驾驶员走神时眼睛睁开目光呆滞此时 PERCLOS 值也很低，此时使用 PERCLOS 就不能准确的判断了，但此时的疲劳状态可以通过眨眼频率得到及时的判断。本文采用的方法相对于通过测试头部的位置、打哈欠等方法更加准确、稳定。疲劳分析流程如图 6.4 所示。通常情况人眼每分钟内眨动 10 至 15 次，平均每次眨眼间隔时间 4 至 5 秒，每次眨眼持续时间约为 0.2 秒。驾驶员困倦或因走神引起的目光呆滞都能使眼睛眨眼频率过低。本文的系统设定眨眼频率的正常区间为 5 次/分到 20 次/分。

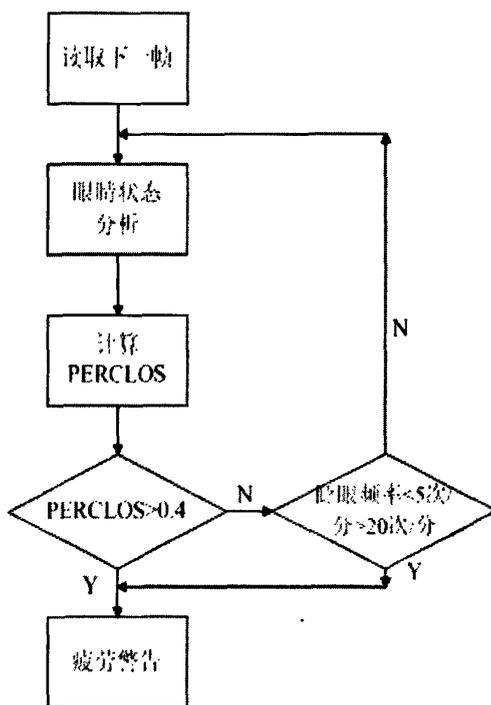


图 6.4 疲劳分析流程图
Fig.6.4 Fatigue Analysis Flow

PERCLOS 是指单位时间内眼睛闭合的时间。6.1 节已经对人眼若干种状态进行了分析，并将眼部状态数值化，将眼睛状态与 H 和 SUM 相关联，假设眼睛正常睁开时眼睛高度 H 的值为 H_MAX ，眼睛面积为 SUM_MAX ，闭眼时眼睛面积为 SUM_MIN ，单位均为像素。 H_MAX ， SUM_MAX ， SUM_MIN 都是通过使用者的具体情况统计出来的，由于不同人的个体差异，每个人眼睛大小不同，驾驶员初次使用时需要将系统置位，使系统重新统计 H_MAX ， SUM_MAX ， SUM_MAX 。

对于上下眼睑高度，用 $H_MAX/3$ 为眼睛正常状态与非正常状态的门限值。令 $isOpenJudgeByH(i)$ 为判断眼睛睁闭的变量，该变量表示第 i 帧中的眼睛是否闭合，则变量 $isOpenJudgeByH(i)$ 可以通过公式 6.2 计算。

$$isOpenJudgeByH(i) = \begin{cases} 1 & H(i) > H_MAX \times 2/3 \\ 0 & H(i) \leq H_MAX \times 2/3 \end{cases} \quad (6.2)$$

$isOpenJudgeByH(i)$ 为第 i 帧眼睛的状态，其中 1 代表眼睛睁开，0 代表眼睛闭合， $H(i)$ 表示第 i 帧的上下眼睑的高度。

对于眼部区域面积， SUM_MIN 为闭眼时无眼球即只有眼睫毛的面积，对于第 i 帧眼球露出的面积为：

$$AREA(i) = SUM(i) - SUM_MIN \quad (6.3)$$

眼球面积为：

$$AREA_EYE = SUM_MAX - SUM_MIN \quad (6.4)$$

用 $AREA_EYE \times 1/5$ 作为眼睛正常与非正常状态的阈值。设 $isNormalJudgeByArea(i)$ 表示第 i 帧的眼睛是否正常睁开，则 $isNormalJudgeBySUM(i)$ 可以用公式(6.5)计算。

$$isOpenJudgeByArea(i) = \begin{cases} 1 & SUM(i) > AREA_EYE \times 1/5 \\ 0 & SUM(i) \leq AREA_EYE \times 1/5 \end{cases} \quad (6.5)$$

$$isOpen(i) = isOpenJudgeByH(i) \ \&\& \ isOpenJudgeByArea(i) \quad (6.6)$$

每 60 秒计算一次 PERCLOS 值来判断驾驶员的疲劳程度，PERCLOS 的计算公式为：

$$PERCLOS = \text{眼睛闭合帧数} / \text{总帧数} \times 100\% \quad (6.7)$$

其中，总帧数为本系统中在 60 秒内处理的视频帧数： $15 \times 60 = 900$ 帧。

眨眼频率为一分钟内眼睛由睁到闭或由闭到睁得次数。对于每一帧图像都有眼睛睁、闭两种状态被判别出来。若用“1”代表睁眼，用“0”代表闭眼，一段时间内的所有图像帧会对应一个 1, 0 组成的二进制序列。

6.4 实验结果

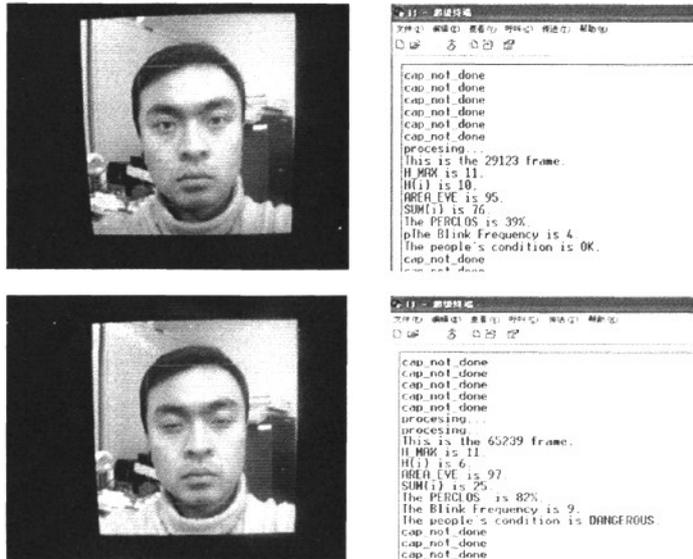


图 6.5 实验结果

Fig.6.5 Experiment result

实验是在室内模拟驾驶条件下进行的，被测者头部放松左右旋转不超过 15 度，无大幅度的左右前后移动；系统采集图像大小为 256×256，采样频率为 15 帧/秒。本文记录 3 分钟内，采集到的 15×60×3 共 2700 帧图像。记录 3 分钟内被测者的眨眼次数，进而得到驾驶员的眨眼频率。面部边界及眼部边界用红色边框标出，调试过程中通过超级终端打印的结果，实际应用中，给出声光等报警，结果如图 6.5。对 20 名测试者进行 3 分钟的试验验证，计算 PERCLOS 及眨眼频率。实验结果如表 6.1 所示。实验结果显示 PERCLOS 大于 80% 时或眨眼频率低于 5 次/分钟及高于 20 次/分钟，即可判断属于疲劳。

表 6.1 实验结果
Tab. 6.1 Experiment result

测试人员	眼睛闭合总 帧数	眼睛睁开总 帧数	PERCLOS	眨眼频率 (次/分钟)	是否疲劳
1	1701	999	63	7	否
2	2268	432	84	3	是
3	1944	756	72	30	是
4	2349	351	87	3	是
5	1080	1620	40	8	否
6	2214	486	82	4	是
7	1766	934	65	4	否
8	1508	1192	56	7	否
9	2058	642	76	5	否
10	2243	457	53	3	否
11	1890	810	70	4	否
12	1620	1080	60	8	否
13	1512	1188	56	9	否
14	1971	729	73	5	否
15	2295	405	85	3	否
16	2255	445	83	3	是
17	2192	508	81	4	是
18	1971	729	73	30	是
19	1175	1525	44	9	否
20	2133	567	79	4	否

结 论

本文在比较国内外对疲劳生理特征的研究结果和现有的疲劳检测报警装置后提出用 PERCLOS 结合眨眼频率作为疲劳程度的判定依据,开展了基于 FPGA 的驾驶员疲劳实时检测系统研究,该系统基于计算机视觉、无接触、无生理负荷的驾驶员疲劳检测方法,取得了一些有价值的成果。

本文主要是在实验室自然光照条件下,通过 CCD 摄像机实时拍摄被测试者的面部,提取最能反应疲劳的眼睛状态特征,判断被测试者是否疲劳。所作的工作具体如下:

(1) 构建基于 FPGA 的疲劳检测系统,采用 Xilinx 的 Virtex-II Pro FPGA。采用软硬件协同设计理念,进行合理的软硬件划分,将 FPGA 灵活的硬件设计和实现与处理器的强大软件功能有机结合,通过软硬件配合提高系统检测的实时性。

(2) 通过硬件描述语言利用 FPGA 内部可编程资源根据具体需要编写视频采集、显示、帧存模块等,满足了系统的特定要求,并且保证了系统的可扩展性和可移植性。其中直接采集 YCrCb 数据减少了传统方法 RGB 到 YCrCb 颜色空间转换的开销。

(3) 利用人的皮肤颜色在 YCrCb 颜色空间的聚类分布特点,提出了基于人脸肤色模型的驾驶员面部定位方法。该方法可靠性高,具有良好的动态定位能力,而且对驾驶员坐姿的变化有较好的适应能力。人脸定位过程中通过硬件描述语言利用 FPGA 内部可编程资源用纯硬件方式完成光线补偿,提高系统的实时性。

(4) 提出了脸部几何特征、水平灰度积分、区域生长法相结合的人眼定位算法。提出了自适应阈值分割结合区域生长法提取眼睛上下眼睑高度及眼部区域面积,有效提取到了驾驶员的疲劳参数。其中阈值分割及中值滤波部分都是用纯硬件方式实现大大提高系统的实时性。

(5) 提取最能反应疲劳的眼睛状态特征,PERCLOS 结合眨眼频率作为疲劳程度的判定依据,大大提高系统的准确性。

本文中的检测主要针对自然光的情况,今后需考虑非自然光全天候的情况,可以通过红外摄像头进行视频采集;检测目前处理速度是 15 帧/秒,疲劳检测的准确率为 75%,速度和准确率还有待于进一步提高,可以通过 FPGA 内部 DSP 及软核模块提高系统的处理能力进而提高系统的处理速度,采用更加精确的人眼检测算法提高疲劳状态的识别率。

参 考 文 献

- [1] 公安部国家安全生产监督管理局. 关于 2002 年全国道路交通事故及预防工作情况的通报 [J]. 道路交通管理, 2005:42-43.
- [2] 2004 年全国道路交通事故情况 [J]. 道路交通管理, 2005:5-7.
- [3] 周玉彬, 俞梦孙. 疲劳驾驶检测方法的研究 [J]. 医疗卫生装备, 2003, 10(6):25-28.
- [4] 毛品, 初秀民, 严新平, 等. 汽车驾驶员驾驶疲劳监测技术研究进展 [J]. 中国安全科学学报, 2005, 15(3):108-112.
- [5] 张灵聪, 王正国, 朱佩芳, 等. 汽车驾驶疲劳研究综述 [J]. 人类工效学, 2003, 9(1):39-42.
- [6] 王荣本, 郭烈, 顾柏园, 等. 基于机器视觉的行车安全综合保障系统研究 [J]. 山东交通学院学报, 2006, 14(2):1-6.
- [7] 王诚, 薛小刚, 钟信潮. FPGA/CPLD 设计工具 Xilinx ISE 5. x 使用详解 [M]. 北京:人民邮电出版社, 2003.
- [8] 赵峰, 马迪铭, 孙炜, 等. FPGA 上的嵌入式系统设计实例 [M]. 西安:西安电子科技大学出版社, 2008.
- [9] 薛小刚, 葛毅敏. Xilinx ISE 9. x FPGA/CPLD 设计指南 [M]. 北京:人民邮电出版社, 2007.
- [10] Hsieh Ing-Sheen, Fan Kuo-Chin. A statistic approach to the detection of human faces in color nature scene [J]. Pattern Recognition, 2002, 25(2):1583-1596.
- [11] 艾海舟, 肖习攀, 等. 人脸检测与检索 [J]. 计算机学报, 2003, 26(7):874-881.
- [12] Karlekar J, Desai U B. Finding faces in color images using wavelet transform [J]. Image Analysis and Processing, 1999, 13(36):1085-1088.
- [13] Craw I, Ellis H, Lishman J. Automatic extraction of face features [J]. Pattern Recognition, 1987, 5(2):183-187.
- [14] Kim H, Kang W, Shin J, et al. Face Detection Using Template Matching and Ellipse Fitting [J]. IEEE Trans, 2000(11):2008-2011.
- [15] Albert P, Jordi Vitrib. Topological Principal Component Analysis for Face Encoding and Recognition [J]. Pattern Recognition letters, 2001, 22(627):7692-7761.
- [16] Govindaraju V, Srihari. S. N, Sher DB. A computational model for face location [J]. Computer Vision, 1990, 23(5):718-721.
- [17] 高守传, 姚领田. Visual C++ 实践与提高—数字图像处理与工程应用篇 [M]. 北京:中国铁道出版社, 2006.
- [18] Jones M. J, Rehg J. M. Statistical color models with application to skin detection [C]. IEEE Computer Vision and Pattern Recognition, Fort Collins, Colorado, 1999:274-280.
- [19] 孙晓玲. 人脸识别系统中眼睛定位方法的研究 [D]. 济南:山东师范大学, 2008.

- [20] Rafael C. Gonzalez, Richard E. Woods. 数字图像处理(第二版)[J]. 阮秋琦, 阮宇智译. 北京: 电子工业出版社, 2003.
- [21] 陶霖密, 彭振云, 徐光祐. 人体的肤色特征[J]. 软件学报, 2001, 5.
- [22] Hsu R L, Abdel-Mottaleb M, Jain A K. Face detection in color images[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2002, 24(5):696-706.
- [23] Linda G. Shapiro, George C. Stockman. Computer Vision[M]. 北京:机械工业出版社, 2005.
- [24] Milan sonka, Vaclav Hlavac, Roger Boyle. 图像处理分析与机器视觉(第二版)[M]. 北京:人民邮电出版社, 2003.
- [25] 杜平, 张燕昆, 刘重庆. 基于广义对称变换的人脸检测和面部特征提取[J]. 计算机仿真, 2003, 20(2):117-121.
- [26] 林维训, 潘钢, 吴朝晖, 等. 脸部特征定位方法[J]. 中国图象图形学报, 2003, 8(8):849-859.
- [27] Dieckmann U, Wagner T, Wagner T. SESAM: A biometric person identification system using sensor fusion[J]. Pattern Recognition Letters, 1997, 18(9):827-833.
- [28] Karin Sobottka, Ioannis Pita S. A fully automatic approach to facial feature detection and tracking[C]. In:AVBPA' 97 Department of Informatics, University of Thessaloniki, Greece, 1997, 77-84.
- [29] 陈启泉, 邱文字, 陈维斌. 标准正面人脸图像的特征提取[J]. 华侨大学学报(自然科学版), 2000, 21(4):413-418.
- [30] Kolte Priyadarshan, Smith Roger, Su Wen. A Fast Median Filter using Altivec[C]. IEEE International Conference on Computer Design (ICCD' 99). 1999: 384-391.
- [31] 谷口庆治. 数字图像处理[M]. 北京:科学出版社, 2002.
- [32] 王磊. 基于计算机视觉的驾驶员疲劳瞌睡检测方法的研究[D]. 济南:山东大学, 2005.
- [33] 刘刚. 基于视频的驾驶员疲劳状态检测研究 [D]. 西安:西安理工大学, 2006.
- [34] 郑培, 宋正河, 周一鸣. 基于 PERCLOS 的机动车驾驶员驾驶疲劳的识别算法[J]. 中国农业大学学报, 2002, 7(2):104-109.
- [35] 杨杰, 卢业力. 驾驶员疲劳度的实时检测[J]. 武汉理工大学学报·信息与管理工程版, 2005, 27(5):21-24.
- [36] 郑培. 机动车驾驶员驾驶疲劳测评方法的研究[D]. 北京:中国农业大学, 2001.

攻读硕士学位期间发表学术论文情况

- 1 马洪连, 李栋, 丁男, 谢海潮, 魏战宇, 齐伟. 基于 FPGA 的多运动目标实时检测系统的研究. 中国智能自动化会议 (CIAC'2009): 1407-1412, 2009 年 9 月. (本硕士学位论文论文第三、五章)
- 2 李栋, 马洪连, 丁男. 基于 FPGA 及机器视觉的运动目标检测系统. 大连理工大学研究生院网络学刊, 2009 年 10 月. (本硕士学位论文论文第三、五章)

致 谢

值此论文完成之际，回顾两年多的读研历程，感觉在学习和生活上都有很大的收获，这其中有自己的拼搏与奋斗，但更要归功于导师、同学、家人和朋友的帮助和鼓励。

首先我要衷心感谢我的导师马洪连教授在我读研两年多时间里对我的精心培养和悉心照顾。两年半来，马老师渊博的学识、严谨的科研态度、忘我的工作精神给我留下了深刻的印象，也使我学到了很多东西，同时也为我今后的学习、工作和生活树立了良好的榜样。

感谢实验室的丁男老师、张益嘉老师，在读研期间中，他们给予了我很大支持和帮助，并在论文完成期间为我提出了很多建设性意见。同时向本实验室的谢海潮、王佳英、齐伟、徐军亮、宫学生、韩秀峰、魏战宇等表示感谢，他们都在论文的完成过程中给了我很大的帮助和支持，同时我也从他们身上学到了不少东西。

最后，我要特别感谢我的家人和朋友，因为有了他们多年来一如既往的关爱、支持和鼓励，才能使我不断的成长、不断的进步，使我能在前进的道路上始终微笑面对困难，战胜自我。

向所有给予过我帮助和支持的人们表示衷心的感谢！