

学校代号 10532

分 类 号 U461.91

学 号 S160200283

密 级 公开



湖南大学
HUNAN UNIVERSITY

硕士学位论文

基于机器视觉的 多特征疲劳检测研究

学位申请人姓名	张婷
培 养 单 位	机械与运载工程学院
导师姓名及职称	曹立波教授
学 科 专 业	机械工程
研 究 方 向	汽车安全
论文提交日期	2019 年 4 月 19 日

学校代号：10532

学 号：S160200283

密 级：公开

湖南大学硕士学位论文

基于机器视觉的多特征疲劳检测 研究

学位申请人姓名： 张婷

导师姓名及职称： 曹立波教授

培 养 单 位： 机械与运载工程学院

专 业 名 称： 机械工程

论文提交日期： 2019 年 4 月 19 日

论文答辩日期： 2019 年 5 月 12 日

答辩委员会主席： 刘敬平教授

Research on Multi-feature Fatigue Detection Based on Machine Vision

By

ZHANG Ting

B.E. (Southwest Jiaotong University) 2016

A thesis submitted in partial satisfaction of the

requirements for the degree of

Master of Science

in

Mechanical Engineering

in the

Graduate School

of

Hunan University

Supervisor

Professor CAO Libo

April, 2019

摘 要

作为日渐走入每家每户的代步工具，汽车带来便利的同时也造成了交通事故的频发，智能驾驶相关技术随之赢来了空前的热潮，驾驶员疲劳检测技术是其重要组成部分。

本文在深入研究相关方法原理的基础上，考虑到应用推广性，并为更准确稳定地检测驾驶员疲劳程度，提出一种非接触式疲劳检测算法，开发了基于机器视觉的多特征融合的驾驶员疲劳检测系统。该方法首先利用相机拍摄驾驶视频，经一系列图像预处理后，通过基于 MB-LBP 特征的 AdaBoost 算法初步定位人脸，然后利用堆叠主动形状模型 Stasm (Stacked Active Shape Model) 和相机位姿测量提取并计算疲劳特征，最后通过极限学习机 ELM(Extream Learning Machine)分类器判断驾驶状态，移植到开发板 Jetson TX2 上，实现车载式疲劳检测。本文主要研究内容在于：

1.针对经典 AdaBoost 算法使用的 haar 检测特征受光照影响较大的问题，改用具有灰度不变性的 MB-LBP 特征，在提高人脸检测速度的同时减少光照对人脸检测的影响。

2.研究了一种特征点定位及特征提取的方法和一种基于视觉的旋转角度测量的方法：利用在 ASM(Active Shape Model)基础上改进的 Stasm 进行关键特征点定位，准确并统一提取面部状态参数；在 Stasm 的基础上结合相机位姿测量求取头部相对于相机的实时角度，较准确地提取头部状态参数。

3.为提高算法鲁棒准确性，使用多个具体疲劳特征共同判断疲劳状态。将各种基础状态参数结合特定驾驶员的初始化，计算 PERCLOS (Percentage of Eyelid Closure)，眨眼频率，哈欠频率，点头频率等 4 个具体疲劳特征。

4.为提高状态分类的速度和可靠度，使用 ELM 进行分类。搭建实验平台并采集志愿者模拟驾驶的驾驶视频，以由上述算法计算得到的 4 个具体疲劳特征为样本集，对分类器进行训练和测试，完成对驾驶状态的判断。

程序移植到开发板 Jetson TX2 上，实现车载式疲劳检测。通过测试验证，在不同光照条件下，本文疲劳检测方法平均检测率为 86.19%，为驾驶疲劳预警提供信息依据，避免驾驶员进入驾驶疲劳状态，从而实现更安全的驾驶。

关键词：疲劳检测；图像预处理；AdaBoost；MB-LBP；堆叠主动形状模型；相机位姿测量；极限学习机

Abstract

As a means of transportation for every household, automobiles bring convenience and also cause frequent traffic accidents. Intelligent driving technology has won an unprecedented upsurge. Driver fatigue detection technology is an important part of it.

On the basis of intensive study of relevant methods and principles, considering the generalization of application, and in order to detect driver fatigue more accurately and steadily, this paper proposes a non-contact fatigue detection algorithm, and develops a driver fatigue detection system with multi-feature fusion based on machine vision. Firstly, the driving video captured by camera goes through a series of image preprocessing. And the face is initially located by AdaBoost algorithm based on MB-LBP features. And then the fatigue features are extracted and calculated by using Stasm (Stacked Active Shape Model) and camera pose measurement. Finally, the driving status is judged by ELM (Extream Learning Machine) classifier and it is transplanted to development board Jetson TX2, realizing on-board fatigue detection. The main research contents of this paper lies in:

1. Aiming at the problem that the haar detection features used in the classical AdaBoost algorithm are greatly affected by illumination, MB-LBP features with gray invariance are used to improve the speed of face detection and reduce the influence of illumination on face detection.

2. Study a method of location for feature and feature extraction, and a method of rotation angle measurement based on vision. Key feature points are located by Stasm which is improved on the basis of ASM (Active Shape Model), extracting the facial state parameters accurately and uniformly. On the basis of Stasm, the real-time head angle relative to camera is obtained by combining camera pose measurement, extracting the head state parameters accurately.

3. In order to improve the robustness and accuracy of the algorithm, several specific fatigue features are used to judge the fatigue state together. Four specific fatigue features, including PERCLOS (Percentage of Eyelid Closure), blink frequency, yawn frequency and nod frequency, were calculated by combining various basic state parameters with the initialization of the specific driver.

4. In order to improve the speed and reliability of state classification, ELM is used for classification. Set up the experimental platform and collect simulated driving videos of volunteers. Take the four specific fatigue features calculated by the above algorithm as the sample set, which train and test the classifier to complete the judgment

of driving status.

The program is transplanted to the development board Jetson TX2 to realize on-board fatigue detection. Through testing and verification, under different illumination conditions, the overall detection rate of fatigue detection method in this paper reaches 86.19%. It provides information basis for driving fatigue early warning, avoiding driving fatigue to achieve safer driving.

Key words: Fatigue Detection; Image Preprocessing; AdaBoost; MB-LBP; Stasm; Camera Pose Measurement; ELM

目 录

学位论文原创性声明和学位论文版权使用授权书.....	1
摘 要	11
Abstract	111
第 1 章 绪论	1
1.1 课题的研究背景及意义	1
1.2 疲劳驾驶检测技术分类	3
1.2.1 基于生理信号的疲劳检测方法	3
1.2.2 基于车辆行为的疲劳检测方法	4
1.2.3 基于驾驶员行为的疲劳检测方法	5
1.3 驾驶员疲劳检测国内外研究现状	5
1.3.1 国外疲劳检测研究现状	5
1.3.2 国内疲劳检测研究现状	6
1.4 本文算法流程和结构安排	7
第 2 章 图像预处理	9
2.1 光照条件对人脸成像的影响	9
2.2 图像灰度化	11
2.3 图像光照补偿	12
2.4 图像降噪	13
2.5 图像预处理流程步骤及效果展示	15
2.6 本章小结	15
第 3 章 基于 AdaBoost 的人脸检测	17
3.1 人脸检测方法概述	17
3.2 经典 AdaBoost	18
3.2.1 haar 矩形特征	18
3.2.2 积分图	20
3.2.3 迭代 AdaBoost	20
3.3 改进的 AdaBoost	22
3.3.1 LBP 特征	22
3.3.2 AdaBoost 的不同	25
3.4 效果对比与实验分析	26
3.4.1 效果上的提升	26
3.4.2 速度上的提升	27
3.4.3 实验分析	28

3.5 本章小结	30
第 4 章 特征提取与计算	31
4.1 面部基础特征	31
4.1.1 特征提取方法概述	31
4.1.2 主动形状模型	32
4.1.3 堆叠主动性状模型	36
4.1.4 面部基础特征提取	36
4.2 头部基础特征	37
4.2.1 相机标定	37
4.2.2 相机位姿测量	42
4.2.3 头部基础特征提取	43
4.3 特征状态参数初始化及阈值	44
4.4 具体疲劳特征计算	45
4.5 本章小结	48
第 5 章 实验分类与移植测试	49
5.1 疲劳驾驶实验	49
5.1.1 实验平台搭建	49
5.1.2 实验设计与流程	51
5.1.3 数据处理	51
5.2 疲劳判断分类	52
5.2.1 极限学习机算法	52
5.2.2 极限学习机分类应用	53
5.3 算法移植	54
5.3.1 软硬件环境	54
5.3.2 移植算法	56
5.4 测试与验证	57
5.5 小结	58
总结与展望	60
参考文献	62
附录 A 攻读学位期间所发表的学术论文	67
致谢	68

第 1 章 绪论

本章详细介绍了驾驶员疲劳检测技术的研究背景、研究意义，介绍了驾驶员疲劳检测技术的三种检测方法并分析了其优势与劣势，为后期的选择提供重要的参考，并概要叙述了驾驶员疲劳检测技术国内外的研究应用现状，最后介绍了本文基于视觉的多特征疲劳检测研究各章节的算法流程和结构安排。

1.1 课题的研究背景及意义

2018 年我国汽车产量完成 2780.9 万辆，销量完成 2808.1 万辆，年产销量连续 10 年蝉联全球第一，但车市出现 28 年来首次年度下降，我国汽车销量 2001-2010 年年均增长速度高约 24.0%，2010-2018 年年均增长速度也达约 5.7%^[1]。因购置税优惠政策的全面退出和中美贸易战宏观经济方面的原因，2018 年出现了难得一见的回落，但是据中国汽车工业协会，我国汽车产业目前仍然处于普及期，具有较大的增长空间，并且向高质量方向的逐步发展与转型，汽车行业前景依旧可观，二胎政策也进一步刺激着国内本来就存在的汽车刚需。我国汽车保有量持续增加，截至 2018 年底已高达 2.4 亿辆，较 2017 年多出 2285 万辆，增长 10.51%^[1]。

汽车作为一种主要的出行选择带来了巨大舒适便捷的同时，所发生的交通事故也造成了很多的财产损失与身体伤害。全球范围内，2015 年交通事故作为第九大导致死亡原因，其死亡人数高于 120 万^[2]。根据公安部官方公布数据，2014 年我国共发生交通事故 196812 起，其中 58523 人死亡，211882 人受伤，约 10.7 亿元人民币造成了损失^[3]。相比于 2013 年的 31604 人，死亡人数增加了 8.5%；相比于 2012 年的 30222 人，死亡人数更是增加了 13.46%。各种各样的交通事故，带来了不可逆的财产损失和人身伤害，并造成了非常严重的社会影响。所以，如何减少驾驶汽车的碰撞，做到安全行车赢得了世界范围内的关注与研究。

美国国家公路安全管理局调查得知，驾驶员疲劳驾驶、分心驾驶与不熟练的驾驶行为成为导致交通事故的主要因素^[4]。据统计，疲劳驾驶导致的交通事故已经达到交通事故总数的 20% 左右，甚至达到特大交通事故的 40% 以上，高达交通事故死亡率的 83%^[5]。疲劳驾驶作为汽车安全中一重要课题，检测驾驶过程中的疲劳程度并给予驾驶员提醒具有很大的必要。

疲劳是指脑力或体力劳动后出现的心理上和生理上的不适状态，主观感受上缺乏动力和警惕，造成全身各项机能下降，一般会出现希望放松和休息的想法^[6]。疲劳虽然是费神后出现的正常的生理反应，但在驾驶过程中驾驶员若发生驾驶疲劳的话，可能会导致无法弥补的人身财产损失。驾驶疲劳是指在颠簸路况或者过于单一或过于紧张的行车条件下，长时间连续驾驶所造成的心理和生理上的不适，

频繁眨眼、持续闭眼，并伴随着一定的打哈欠，点头，伸懒腰等动作。驾驶疲劳会让驾驶员注意力不够集中、视野不够宽、反应不够灵敏、操作迟钝、自我控制力减退，更有甚者目光呆滞、停止操控，非常容易造成交通事故^[7]。

究其原因，分析疲劳驾驶机理发现，造成疲劳驾驶的主要原因包括^[8]：

1. 驾车前缺少睡眠是最主要的原因。这种情况最不易察觉故而最危险，驾驶员本以为在车外环境中清醒下来并且才开始驾驶，造成一种清醒的错觉，因为其实身体本身就进入了疲劳状态，此刻已经思维力不够灵敏、反应力不够快捷了，坐到舒适的车内环境中开始驾驶后极易进入瞌睡状态。

2. 长时间连续驾驶。经济压力之大和生活节奏之快，驾驶员放弃休息，坚持无间断驾驶，在长时间的精神紧张状态下进行着频繁重复的驾驶控车动作，驾驶员精神上和体力上容易进入疲劳状态。

3. 驾驶环境。车内环境与车外环境都会影响驾驶员的精神状态。车内的温度、湿度、光线以及声音等因素都会影响大脑皮质，一般车内环境会被我们调至驾驶员舒适的程度，在舒适的环境中难免容易被催眠。城市道路上复杂的行车条件会造成驾驶员情绪紧张，所带来的较大的精神压力，让驾驶员驾驶过程会比较费劲而容易进入疲劳状态；乡村路况中太过单一的行车条件，过于单调的驾驶动作也容易催眠驾驶员。

4. 驾驶速度。高速行驶时，视野会变窄，驾驶员会高度紧张而容易疲劳；速度过低时，会产生舒适驾驶的错觉，放松下来的神经也容易被催眠。

不同的疲劳状态也会产生不同的驾驶行为，分别从车速控制、方向盘调整 and 身体动作进行阐述，列表如下^[9]。

表 1.1 不同疲劳状态下的驾驶行为

驾驶行为	清醒状态	疲劳状态	极度疲劳状态
车速控制	正常加速减速	减速变慢	速度变换缓慢 甚至不变化
方向盘控制	可以做出正确及时的调整和判断	调整不及时，甚至判断错误	停止控制方向盘
身体动作	正常	频繁眨眼、持续闭眼， 打哈欠，点头，伸懒腰	长时间闭眼， 摇晃身体

开发一个驾驶员疲劳检测与预警系统，在驾驶员开始疲劳时提醒驾驶员，甚至辅之以听觉上的刺激帮助驾驶员抵抗疲劳，或者直接减少给油以降低行车速度，有效率地让行车更安全。交通事故统计结果表明^[10]，若驾驶员反应速度增加 0.5s，那么就可以减少将近 60% 交通事故的发生。由此可见，疲劳检测与预警在很大程度上可以避免不少的交通事故，提高路上行车安全性，不仅具有重要的社会意义和经济研究价值，也符合目前汽车智能安全化的发展潮流。生活在日新月异的世

界和不断发展的社会中，人们的脚步也随之越来越紧促，越来越大的物质上和精神上压力也接踵而至，疲劳问题自然也会越来越普遍。日前，疲劳检测与预警技术逐渐受到各个国家的重视，并开始列入衡量车辆智能与质量指标，也逐渐成为消费者购车的重要考虑因素。

总的来说，在这样一个汽车成为刚需而交通事故频发的时代，疲劳驾驶作为一个极大的安全隐患，难以被驾驶员察觉，难免产生不清醒的驾驶行为，疲劳驾驶成为交通事故很重要而不容忽视的原因。研发疲劳检测技术，识别驾驶员的疲劳状态，并作出相应的预警给驾驶员以提醒，避免驾驶员疲劳反应迟缓而延迟做出正确的操作以导致不必要的碰撞，可以大大的提高行车安全性，让驾驶员少了这种正常生理反应的后顾之忧，更加安心的驾驶。

1.2 疲劳驾驶检测技术分类

目前疲劳检测方法包括：第一种方法是利用脑电图、心电图、肌电图等检测生理行为，此需驾驶员佩戴相关复杂设备，不太方便并且高成本，而难以推广^[11]；第二种方法是利用车速传感器、方向盘角度传感器、踏板位移传感器获取车辆行驶状态，因易受到路况、个人驾驶习惯、车型等外界因素影响，很难形成判断标准，检测率不高，故通常用来辅助判断驾驶员的疲劳状态^[12]；第三种方法是基于视觉的疲劳检测，凭借简单易行并相对可靠的优点，得以广泛应用。

1.2.1 基于生理信号的疲劳检测方法

用生理信号来反映人的生理状态最可靠、最及时。该检测方法通常需要相关医疗设备辅助来采集有关的生理信号以进行疲劳检测，生理信号包括脑电图、心电图、肌电图、眼电图、脉搏信号等。在驾驶过程中，借助动态脑电扫描仪、动态心电仪、肌电仪等仪器采集对应的生理信号，分析生理参数的变化并参照人体疲劳参数指标，判断驾驶员是否疲劳。

1.通过脑电图进行疲劳检测

脑电图是将电极经过头皮的脑细胞群，放大其自发性与节律性生物电位并记录下来的图形，最能准确描述大脑活动，可以用来直接判断驾驶员的精神状态^[13]。Kecklund 等人早在 1993 年就提出，脑电与人的疲劳程度有很大的相关性，尤其是 theta 脑电波随着疲劳程度的提高上升明显，因此可以通过分析 theta 脑电波的不同来判断驾驶员的疲劳程度^[14]。Lal 等人于 2003 年证明了脑电波和疲劳程度的相关性^[15]。故经研究，脑电波完全能够作为判断疲劳程度的重要依据。

2.通过心电图进行疲劳检测

心电图是经动态心电仪在体表记录心脏的每一心跳周期所发生的生物电位放大后详细的图形。心电图通过人体的心脏跳动情况计算得到心率和心率变异性，借此来进行疲劳判断。Lal 等人研究得出心电图随着疲劳程度的增加呈现出节律性下降趋势的结论^[16]。Jeong

等人将电极安装于方向盘与驾驶座椅上，通过心率变异性分析驾驶员的疲劳状态^[17]。吴群建立心电图 R 波频段极值非线性特征与驾驶员疲劳程度的相关模型，模拟出驾驶员不同疲劳程度下 R 波频段的心电信号，作为驾驶员驾驶过程中判断是否疲劳的参考^[18]。

3. 通过其他生理信号进行疲劳检测

眼电图是记录人眼光感受器复合体因光适应改变而产生的眼睛静电位变化轨迹。通常利用待测者眼部上的一对电极，获得左右方向或上下方向的数据，分析得到其中极大值幅值、上升与下降时间 3 个特征值进行疲劳程度判断的依据^[19]。

肌电图记录了人体肌肉运动情况，有肌肉肌电图和表面肌电图两类，驾驶疲劳的检测一般使用表面肌电图中的幅值与平均频率。Katsis 等人研究发现表面肌电图幅值随驾驶员疲劳程度的增加而呈上升趋势，肌电平均频率却相反^[19]。

除研究单一生理信号对疲劳的反应外，Craig Ashley 等人融合了多个生理信号进行多特征综合分析，更有效的进行了驾驶员疲劳检测。

虽然生理信号作为精神状态最为直观的反应，具备较高的准确性、可靠性和及时性，但是其昂贵的硬件支持限制了他的商业价值，其接触性的佩戴要求降低了驾驶过程中的舒适度，也极大的限制了他的推广与应用。

1.2.2 基于车辆行为的疲劳检测方法

当驾驶员进入疲劳状态时，对于车辆的驾驶操控力明显不同于清醒状态，从而出现不正常的车辆行为数据，据此开展基于车辆行为的疲劳检测方法的研究。具体来说，该方法是利用各种传感器采集方向盘转角、加速板踩踏情况、刹车板踩踏情况、车辆加速情况及结合打转向灯情况的车辆轨迹等一系列车辆行为，间接识别判定驾驶员的疲劳状态。基于车辆行为的疲劳检测方法不需要驾驶员佩戴设备，非接触式的优点提高了驾驶舒适度，安装于车辆内部的传感器更是为一系列车辆行为参数的获取提供了很大的方便，数据的精确采集性与容易处理性为该方法的应用赢得了更大的可能性。

该方法凭借其独特的优势，赢得了各界广泛的深入研究与实际应用。Hu 等人研究表明，方向盘转角与驾驶员疲劳程度相关，并用实验数据提出了其中确切的关系^[20]。Wang 等进一步分析了车辆横向加速度、纵向加速度和对应的方向盘转角等数据与驾驶员疲劳程度的相关性，发现其中横向加速度最能反应驾驶员程度^[21]。严新平等探究了驾车操作数据与车辆状态行驶数据等因素和驾驶员疲劳程度的关系，建立了一套车载式驾驶员行为情况^[22]。

基于车辆行为的疲劳检测方法虽然有它独特的优势，但是其关键的弊端在于可靠性。该方法受到驾驶员驾驶习惯、驾驶水平、路面颠簸情况、车辆自身特性等多种因素影响，多变和不可控的影响因素大大的降低了基于车辆行为的疲劳检测方法的准确性，限制了该方法的实际应用。

1.2.3 基于驾驶员行为的疲劳检测方法

基于驾驶员行为的疲劳检测方法利用摄像头采集驾驶员驾驶过程中面部特征与头部姿态等相关参数,然后采用模式识别技术对驾驶员驾驶状态进行分类解析,进而判断驾驶员的疲劳状态。驾驶员行为指眼部状态、嘴部状态和头部状态,这些参数和疲劳状态关系密切,在不同程度上反映了驾驶员的驾驶疲劳情况。

驾驶员行为中眼部状态是反映疲劳状态最主要的特征。该方法利用车载摄像头对驾驶员实时的拍摄画面,识别驾驶员眼睛后提取其眼睛的特征,进而对驾驶员的疲劳状态进行判断。日本 Nissan 公司发现上眼睑与眉毛两曲率的大小关系可以反映眼睛的开闭状态^[23]。卡内基梅隆提出 PERCLOS 作为疲劳检测的指标,根据眼睛闭合程度与所占比例来判断疲劳程度。Acioglu 等人^[24]和 Nguyen 等人^[25]用 PERCLOS 来进行了有效的疲劳检测。

人体进入疲劳状态时会伴随一定的哈欠活动,在驾驶员疲劳时与驾驶员说话或正常驾驶时嘴部状态有所不同,所以嘴部状态在一定程度上也可以反映驾驶员的疲劳状态。该方法是基于随车实时采集的驾驶员驾车面部视频,识别驾驶员嘴巴后提取其嘴部的特征计算得到哈欠频率后,辅助驾驶员疲劳状态的判断。童兵亮等人^[26]通过嘴部特征来进行疲劳判断。Chu 等人分类嘴部几何特征输入到三级 BP 网络中进行人体精神状态的分类。

类似地,疲劳时也会伴随点头动作,头部姿态在不同精神状态下也有所不同,头部运动也可以用来辅助判断疲劳。ASCI 通过所设计的头部位置传感器得到驾驶员的头部位置变化,进而判定精神状态。Philip 等人研究表明点头动作与驾驶员疲劳状态具有不错的联系,并开发了一套点头检测疲劳系统。

基于驾驶员行为的疲劳检测方法建立在车载视频的基础上,作为一种非接触式检测方法,不会引起驾驶员的反感,并且驾驶员面部特征与头部姿态也能够可靠直接地反映驾驶员的精神状态,具有较好实用性和准确性,因此是检测驾驶员疲劳状态较合适的方法。

1.3 驾驶员疲劳检测国内外研究现状

疲劳驾驶的频发高危性,国内外或早或晚地对其进行了各式各样的研究,都取得了丰富有效的成果。

1.3.1 国外疲劳检测研究现状

国外早在 20 世纪初就开始关注驾驶员疲劳检测问题,随着科技水平的提高,在近 20 年疲劳检测研究取得了显著进展。

Grace 等人差分图像得到瞳孔位置与大小以计算 PERCLOS^[27]。Cui Xu 等人基于局部二值特征(Local Binary Pattern, LBP)训练得到正闭眼分类器^[28]。Yufeng

Lu 等人通过灰度投影得到眼部特征,判断眼睛睁闭^[29]。Electronic Safety Products 公司推出的 Steering Attention Monitor 产品,利用传感器测量方向盘角度,以持续 4 秒钟未变化为阈值判定疲劳^[30]。大众的疲劳检测系统通过比较实时的驾驶员操纵行为与随车记录的驾驶员操纵行为,并综合车距、打转向灯情况等参数,判断驾驶员的疲劳程度^[31]。美国约翰霍普金斯大学(Johns Hopkins University)开发了 The Drowsy Driver Detection System 系统,利用多普勒雷达经信号处理获得驾驶员眼部状态,进行疲劳判断^[31]。美国莱斯大学开发了一种红外线动态眼球扫描仪,通过扫描眼白部分得到的眼白占比进行眼睛开闭的判断。Attention Technologies 公司推出的 Driver Fatigue Monitor (DD850) 产品被美国交通运输部在全美国推广使用,基于亮瞳效应利用红外 CCD 摄像头得到 PERCLOS 以实现驾驶员疲劳检测^[32]。

1.3.2 国内疲劳检测研究现状

我国对驾驶疲劳开始关注和研究较晚,与国外有一定的差距,早期研究方向存在局限,未拓展到基于驾驶员行为的疲劳检测方法上来。目前,疲劳检测研究越来越受到科研工作者和相关机构的重视,收效明显,其中值得一提的如下:

深圳一公司推出一款起到驾驶疲劳检测预警作用的手表,根据检测得到的脉搏情况、红外光谱情况、生物电情况等特征进行疲劳检测,疲劳时发出预警^[33]。浙江省金华市高速交警支队则开发了一款挂耳式设备,基于电子平衡原理得到头部运动角度,超出阈值便发出报警^[34]。吉林大学智能车辆课题组^[35]研发了 JLUVA-DFWS 系统,通过眼睛、嘴巴的张合度判断驾驶员的疲劳程度,进一步通过经基于眼嘴在脸廓位置变化训练 BP 神经网络所得到的面部朝向决定是否预警,当夜间驾车时利用红外光谱检测驾驶员瞳孔与普尔钦光斑的位置关系得到面部朝向。清研微视的 ADAS“护驾卫士”已经实现小批量量产,应用于长途客运、货运公司^[36]。上海径卫视觉科技有限公司推出的 RDT-200 通过 PERCLOS 来判定疲劳,效果较好^[37]。刘志强等人基于 DSP 处理器开发了驾驶员疲劳检测系统,通过闭眼持续时间进行疲劳判断,具有 86.89%的准确率^[38]。

如今,市场上已经出现了一些疲劳检测产品,但由于检测效果和成本等因素,其实际应用并没有形成规模。目前国内市场上的疲劳检测产品实时性与准确性不高,其较高的误报警率也引起了众多驾驶员使用者的反感,似乎达不到实际应用要求。纵观该技术的现状,目前的热潮是基于机器视觉、图像处理技术、模式识别技术和包括 PERCLOS 在内的多因素判定因素综合检测驾驶员的疲劳状态,研制开发车载式舒适性、准确性及实时性都高的驾驶员疲劳检测预警系统,也成为本文研究开展的最初原因。

1.4 本文算法流程和结构安排

本文针对驾驶员疲劳检测方法进行相关研究，提出一套完整的基于视觉的多特征疲劳检测算法，算法流程图如图 1.1 所示。首先设计合适的流程步骤，对相机采集到的驾驶图像进行图像预处理，然后选择基于 MB-LBP 特征的 AdaBoost 人脸检测算法完成人脸的初步定位，其次通过 Stasm 算法提取面部基础特征，并结合相机位姿测量提取头部基础特征，在特定驾驶员初始化和特定阈值化后，计算具体疲劳特征，最后通过 ELM 分类器实现疲劳检测。

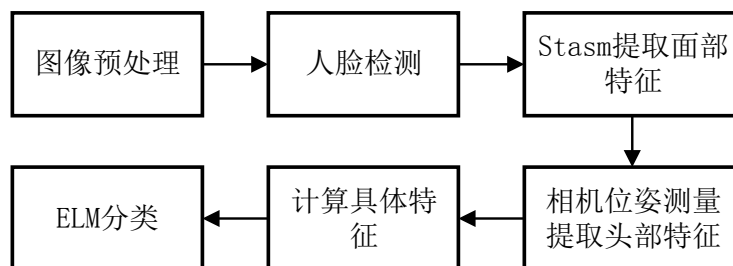


图 1.1 疲劳检测算法流程图

本文共有五章，各章节内容的结构安排如下：

第 1 章，阅读大量相关研究课题的文献资料，介绍驾驶疲劳检测技术的研究背景及课题意义，介绍并比较驾驶员疲劳驾驶检测技术分类方法，总结介绍驾驶员疲劳检测国内外研究现状，并简要介绍本文基于视觉的多特征疲劳检测算法的流程和论文章节内容的结构安排。

第 2 章，介绍图像预处理各步骤的不同方法，并选择适合本文算法的方法，完成本文图像预处理流程设计，即经加权平均法进行图像灰度化，经直方图均衡进行图像光补偿，经中值滤波进行图像降噪。

第 3 章，简要概述人脸检测的不同方法，具体介绍并对比分析基于不同特征的 AdaBoost 算法，通过加州理工学院人脸图像库的检测试验，选取较 haar 特征而言检测既更准确又更快速的 MB-LBP 特征，通过不同光照条件下人脸视频的检测试验，验证基于 MB-LBP 特征 AdaBoost 人脸检测算法的光照影响程度。

第 4 章，进行疲劳特征的提取与计算。介绍不同的面部基础特征提取方法，并选取 Stasm 这一整体化方法进行面部基础特征的提取。再结合相机位姿测量，经相机标定获得相机参数，进行头部基础特征的提取。对于特定驾驶员需进行特征状态参数的初始化，并计算其阈值。在此基础上，设计具体疲劳特征的计算流程，并通过包括驾驶员清醒、疲劳和非常疲劳状态的视频的具体疲劳特征的输出，定性验证该具体疲劳特征作为驾驶状态指标的可分类性。

第 5 章，进行实验分类与移植测试。搭建模拟驾驶实验平台，进行驾驶视频的采集，经视频数据分类处理，利用前面章节的算法流程为 ELM 分类器的训练

和测试提供不同驾驶状态下的具体疲劳特征数据集，经程序封装整体化，移植至 Jetson TX2 嵌入式平台中验证其疲劳检测效果。最后通过试验验证不同光照条件下本文疲劳检测算法可靠性。

第 2 章 图像预处理

在光的刺激下，经物体表面反射，反射光信息被感光体收集并处理而形成图像。数字图像的形成过程会受到环境因素与摄像头因素的影响，因光照条件和随机噪声而与失去自然图像的某些信息^[19]，对图像的识别与特征的提取造成不便，所以在对图像进行检测之前，完成图像预处理十分必要。驾驶场景存在各种各样的光照条件，采集到的驾驶员图像有明显的光线差异与多种噪声，若不进行图像预处理，那些成像质量差的图像很难进行下一步的图像识别和疲劳检测。图像预处理分为空间域的图像增强和频域的图像增强，可从减少光照与噪声的影响两方面着手。本节通过对比不同光照条件下人脸图像直方图分析光照因素的影响，然后比较分析了各种图像灰度化方法和图像滤波去噪方法，选择合适驾驶员疲劳检测场景的方法并设计了一套适合的流程，同时通过实验验证了效果。

2.1 光照条件对人脸成像的影响

摄像头是类似于人眼结构的仿生机器，通过物体的反射光和透射光由镜头聚焦至机器感光物质上而形成图像。其中图像质量的高低与光照强度、角度及被拍摄物材质的反射性质有很大的关系，不同的被拍摄物或者在不同的光照条件下会有不同的呈现效果，自然而然，不同光照强度、角度下的人脸成像会有所不同^[39]。Adini 等人早在上个世纪就研究了人脸识别对光照的敏感度，并表明光照条件比起不同人脸对人脸成像造成的影响差异更大^[40]。从如图 2.1 所示的 Yale 大学人脸数据库中挑选一组有各种光照条件下同一个人的人脸，如图 2.2 所示，其很好地表现了不同光照强度、光照角度对人脸投影成像的影响，光照强度会影响图像整体性灰度的不同，光照角度会引起人脸图像局部阴影状况甚至是阴阳脸。

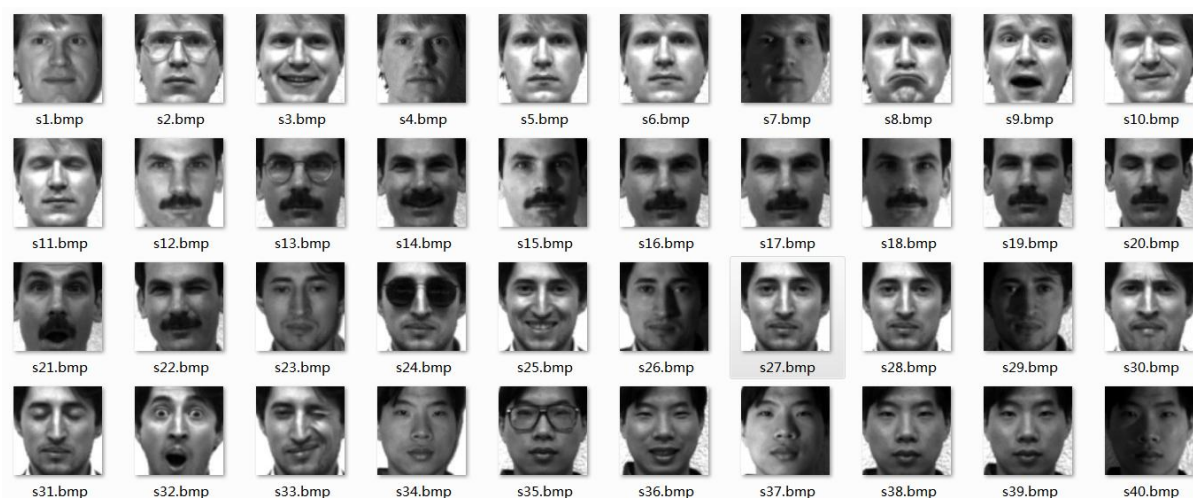
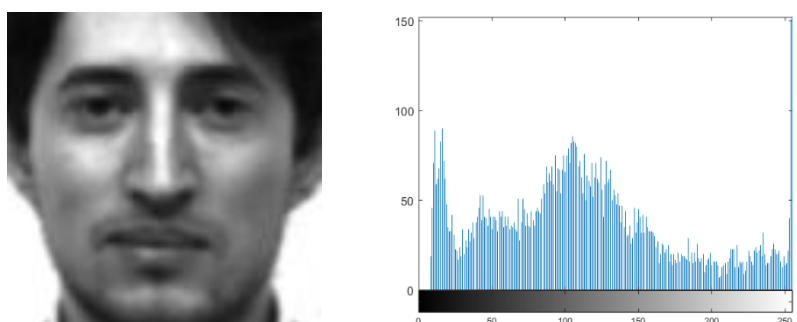


图 2.1 Yale 大学人脸数据库

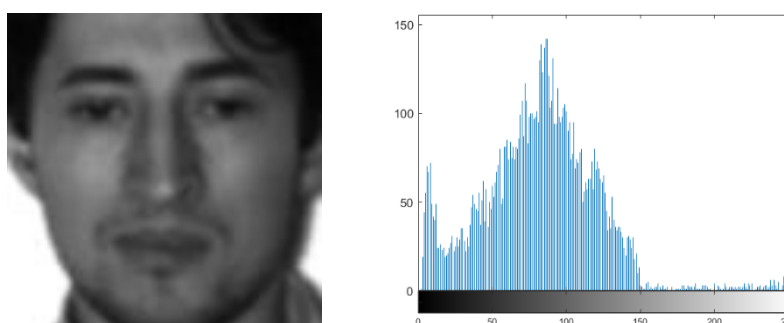


图 2.2 不同光照下的人脸图像

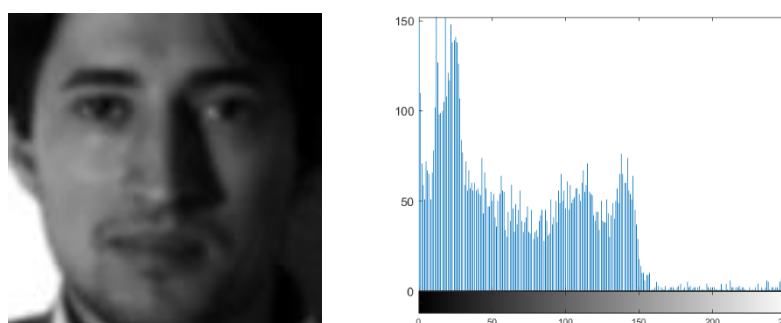
进一步地，对图 2.2 中不同光照下的四副人脸图像绘制了各个灰度级像素点个数的灰度直方图，完成进行直方图分析。如图 2.3(a)所示，正常光照人脸图像的灰度直方图覆盖了灰度级较宽的范围，在各个灰度级上分布较均匀，图像具有多变的灰度色调和较高的对比度；如图 2.3(b)所示，正面弱光人脸图像灰度直方图的整体灰度级变小并且集中压缩至 0-150 级了；如图 2.3(c)所示，左侧光使得人脸左亮右黑，使得整体灰度级进一步变小，并向 0-160 级集中；如图 2.3(d)所示，右侧光灰度分布与左侧光类似。



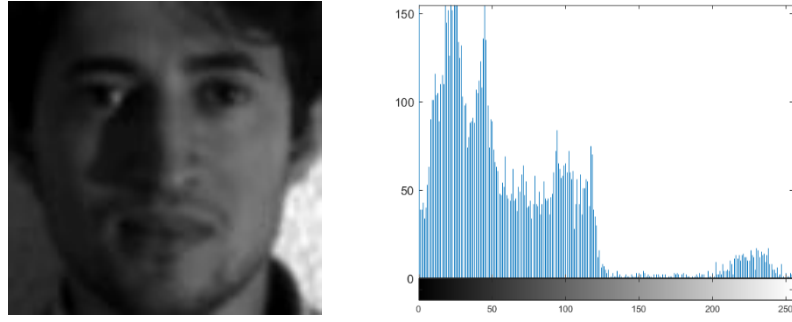
a) 正面正常光照人脸图像及灰度直方图



b) 正面弱光人脸图像及灰度直方图



c) 左侧光人脸图像及灰度直方图



d) 右侧光人脸图像及灰度直方图

图 2.3 不同光照下人脸图像及灰度直方图

而在实际驾驶过程中，白天夜晚光照强度有明显不同，光照角度的不同使得驾驶位的光线不均匀，使得光照条件已成为影响驾驶员人脸成像的一个重要因素。对从 Yale 大学人脸数据库中提取的不同光照下的人脸图像所进行的灰度直方图分析，灰度分布的不均匀很大程度上影响了人脸成像，故对其进行光补偿有很大的必要。目前光照补偿技术研究深入并且应用广泛，基于图像处理的光照补偿技术主要是通过改变图像的灰度分布使其更均匀，提高图像的整体可见性与局部细节可测性，为后面的图像识别与检测提供一个好的基础。

2.2 图像灰度化

彩色图像灰度化作为彩色图像处理的基础步骤，能够将多通道的数据综合成灰度级别这一个通道，降维操作在不影响后期特征提取的前提下，大大减少了数据运算量。灰度将纯黑色作为基准，描述不同饱和度的黑色，从 0 级开始到 255 级结束表示从纯黑色到纯白色灰度值的变化。人眼可观察到的彩色图像所构成的 RGB 颜色空间由红、绿、蓝三种颜色分量组成，每种分量根据其黑色饱和度有其对应的灰度值^[41]。

彩色图像灰度化是用每一个图像像素的一个灰度值来反映该像素的多通道分量，包括分量法、最大值法、平均值法、加权平均法等等。分量法是指根据应用情况选取 R、G、B 三分量中某一分量作为灰度化图像的灰度值；平均值法是求其三分量的平均值来作为最终值；加权平均法根据重要性将三分量给以合适权重再求和平均，现依据人眼对颜色的敏感程度，有式如：

$$f_k(i, j) = 0.299 \times R_k(i, j) + 0.587 \times G_k(i, j) + 0.114 \times B_k(i, j), k = 1, 2, 3, \dots \quad (2.1)$$

各个图像灰度化方法效果如图 2.4 所示。我们可以看到，分量法等单一通道的处理方法片面的处理某一通道的信息，丢低了图像其他通道的灰度信息，效果并不是很好；平均值法和加权平均法等综合多通道的信息，更全面，并且加权平均法较平均值法更符合人眼的观察，边缘信息的特征对比度更加突出，更适合人脸检测。因此，本文对人脸图像的灰度化选择简单有效的加权平均法。



图 2.4 各种图像灰度化方法效果

2.3 图像光照补偿

光线变化与天气变化等都是驾驶过程中常遇到的情况，阴天或夜晚光线暗环境和驾驶座位位置的偏置会降低后期识别检测的精确度。高光造成图像的过度曝光与缺少光线所带来的图像阴影都会导致细节信息的盲区，对后期人脸识别尤其是检测特征的提取造成不可弥补的损失。所以，图像的光照补偿可以保留一定的图像特征信息，成为基于图像处理的驾驶员疲劳检测技术的关键步骤之一。

基于图像处理的图像光照补偿将灰度值重新分布以实现补光，常见的方法包括直方图均衡、参考白、Gammer 变化、GrayWorld 色彩均衡等。

直方图均衡化凭借其简单性和自动性的特点应用广泛。直方图均衡化的原理是，自动确定变化函数以生成有均匀直方图的输出图像，输出图像灰度值覆盖范围得以扩大，提高了整体对比度^[42]。

一幅数字灰度图像灰度级 r_k 出现的概率近似为

$$P_r(r_k) = \frac{n_k}{n}, k = 0, 1, 2, \dots, L - 1 \quad (2.2)$$

式 2.2 中， n 是图像像素的总数， n_k 是隶属灰度级 r_k 的像素数量， L 是图像灰度级的总数。数字图像的离散变换过程如式 2.3 所示，各对应像素的灰度级从输入图像中的 r_k 映射为输出图像中的 s_k 。

$$s_k = \sum_{j=0}^k P_r(r_j) \quad (2.3)$$

驾驶过程中，由于光线强度和光线角度的原因，摄像头下的人脸成像会存在整体偏暗或局部偏暗的情况，此时通常经直方图均衡化进行光线补偿。如图 2.5 所示，直方图均衡化得到了不错的效果，弱光下的人脸图像经均衡提高了整体亮度，方便后期检测；偏光下的原始图出现了阴阳脸现象，使阴影一侧失去了特征提取的可检测性，而直方图均衡化补充了原始图丢失的信息。图 2.5(c)对比图 2.3(b)和图 2.5(f)对比图 2.3(d)可以发现，不管是在弱光环境下还是在偏光环境下，灰度级呈向右平铺趋势。因此，直方图均衡化是一种十分有效的光线补偿方法，改善人脸灰度的分布情况以提高亮度和对比度，完善突出人脸边缘特征信息。

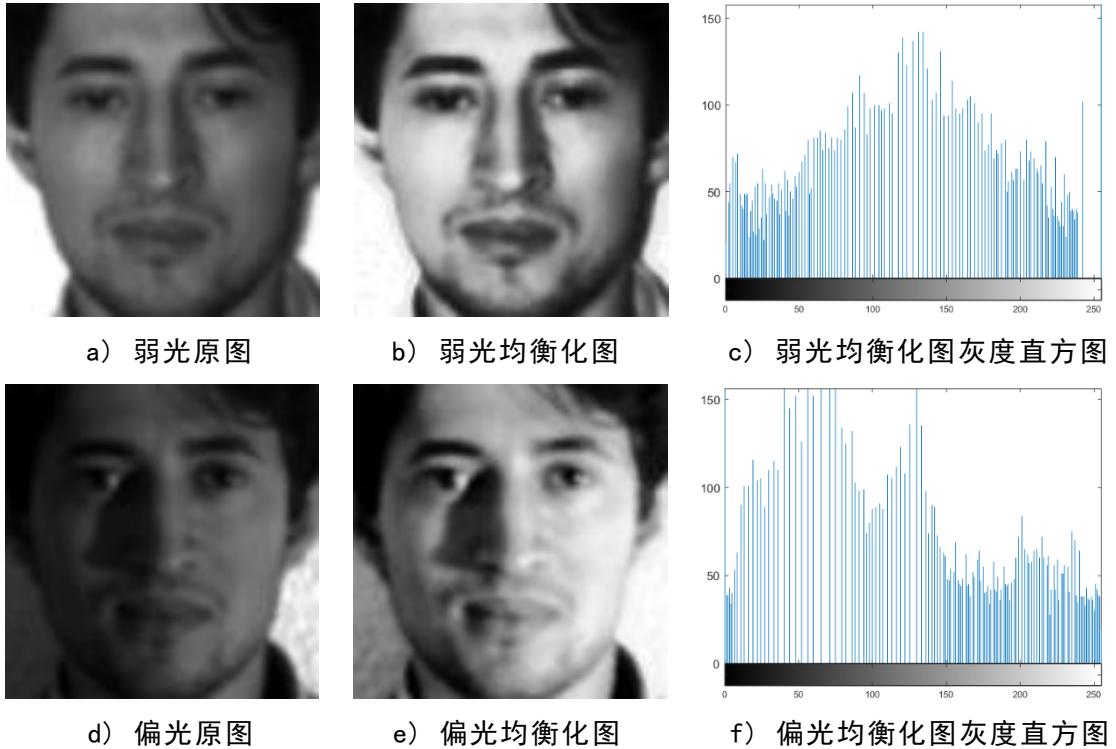


图 2.5 弱光与偏光人脸图像直方图均衡化效果

2.4 图像降噪

驾驶员图像会受到相机自身的设备因素和环境等外界因素的影响而存在或多或少的噪声，如高斯噪声、椒盐噪声等。原始图像中各种不可避免的噪声的存在降低了图像的品质，甚至模糊图像，干扰图像局部特征信息的检测与提取^[43]。因此为了避免噪声影响视觉效应，更为了避免噪声为后期更高层次的图像处理带来麻烦，驾驶员图像必须降噪，改善人脸图像质量。

图像降噪有空间域的滤波方法和变换域的滤波方法，有线性滤波和非线性方法，常用的包括均值滤波、高斯滤波、中值滤波等等^[44]。

均值滤波是一种邻域平均法，计算图像中每一个像素周围像素的平均值并作为该像素滤波后的值，此方法作为一种经典的线性滤波方法，平均值的计算相对

简单，能够有效地抑制高斯噪声，得到广泛应用。但是均值滤波会模糊图像，丢失图像边缘的特征信息，故不适用于疲劳特征的提取。

高斯滤波是在均值滤波的基础上改进的一种邻域加权平均滤波方法，将呈高斯分布的滤波窗口作为被扫描像素的加权值，用邻域加权平均值来更新该像素的灰度值，并滑动该滤波窗口扫描整个图像，得到高斯滤波后的图像。高斯滤波是与正态分布做卷积，理论来说更符合原始图像的分布规律，所造成的模糊程度相对于均值滤波稍低，但是加权的操作也增加了它计算的复杂性，使得它的使用效率并不是最高的。

无论是均值滤波还是高斯滤波都属于线性滤波的一种方法，会考虑滤波窗口内包括噪声在内的所有像素灰度值，所以线性滤波的方法不会去除噪声，只是让它成为更柔和却依然可见的散粒而已，而线性滤波在柔和噪声的同时也不同程度的模糊了图像。而非线性滤波更新后的像素值并没有包含滤波窗口内所有像素值，成为多数情况下更好的选择，尤其是在处理椒盐噪声时，非线性滤波的效果更好。

基于排序统计理论，中值滤波是一常用的非线性滤波方法，它将滤波窗口像素的灰度值进行排序并取中值作为该像素更新的灰度值，其滤波窗口的滑动扫描与线性滤波类似。中值滤波不依赖于滤波窗口内差异很大的像素，让滤波后的像素更接近真实值，进而消除孤立的噪点，尤其能够非常好地滤除掉椒盐噪声。中值滤波不仅能够避免线性滤波器难免产生的细节模糊，保留疲劳检测中重要的边缘特征，还可以很好地过滤脉冲干扰所带来的噪声，并且在实际运算时不用统计，为计算带来了不少方便。因此，疲劳检测中人脸图像的滤波方法选取简单高效的中值滤波，在降噪的同时还保留了检测所需的边缘信息。



图 2.6 经典滤波方法效果对比

如图 2.6 所示，将拍摄所得的具很多噪点的人脸图像，经各种滤波方法进行图像降噪，得到滤波窗口同为 3×3 的均值滤波和中值滤波后的人脸图像。首先从除噪的滤波效果进行对比，可以很明显地发现均值滤波只是让噪点变得柔和但仍

然存在，而中值滤波很好地去除了噪点，尤其对于椒盐噪声滤波反响超好；然后从滤波后图像的模糊程度进行对比，可以看出中值滤波较均值滤波等非线性滤波更接近原图，清晰程度更高，方便后续疲劳检测中脸部特征的提取。所以，本文选择中值滤波来进行图像降噪。

2.5 图像预处理流程步骤及效果展示

本文人脸图像预处理的流程步骤如图 2.7 所示，首先对输入的原始图像通过加权平均的图像灰度化，在保留了多通道信息的同时又有更好视觉效应，进行图像降维；然后对人脸灰度图像通过直方图均衡，自动简单地拓展了图像像素灰度的覆盖范围，在适当调节整体亮度的同时又补充完善了阴影部分的图像信息，进行图像光线补偿；最后采取中值滤波，在去除噪点的同时又最大保留了图像的边缘信息，进行图像降噪，完成人脸图像的图像预处理并输出待检测图像。

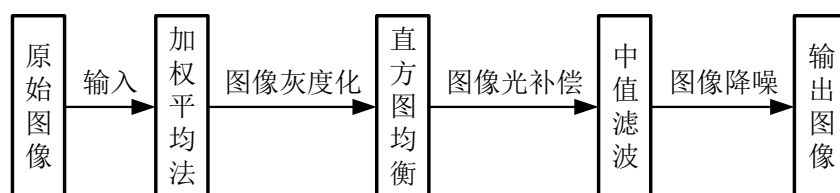


图 2.7 本文人脸图像预处理流程图

人脸原始图像经上述流程步骤得到输出图像，本文图像预处理方法的图像效果如图 2.8 所示。比较图像预处理前后图像，可以发现图像预处理改善了弱光条件下图像质量，图像整体亮度得到提高，并且在降噪的同时保留了图像特征的边缘细节。



a) 原图 b) 图像预处理效果图

图 2.8 人脸图像预处理效果

2.6 本章小结

本章从人脸图像对光照条件的敏感度出发，说明了尽可能地减少光照对驾驶员疲劳检测所造成的干扰的重要。然后分别介绍了图像预处理的常规步骤的不同方法并依次通过效果展示与方法分析选择了应用于驾驶员疲劳检测场景最适合的

方法：首先介绍了图像灰度化四种方法并选择加权平均值法，并且通过直方图均衡完成光补偿，接下来介绍了图像降噪的三种方法并选择中值滤波方法。最后根据每个图像预处理步骤所选定的方法，设计了本文驾驶员疲劳检测人脸的图像预处理步骤，并完成试验验证。

第 3 章 基于 AdaBoost 的人脸检测

基于视频的驾驶员疲劳检测是在摄像头采集到的驾驶员人脸图像的基础上提取人脸特征相关信息，进而计算驾驶员疲劳相关特征，实现疲劳检测。由于图像背景容易受到复杂环境的干扰，直接检测较小且特征不特别明显的面部特征难度较大，效果并不好。因此我们换用由大至小、依次检测的思路，先识别人脸再在这个基础上提取各个人脸特征，能够在大大提升算法准确性的同时也增加了算法的处理速率。

3.1 人脸检测方法概述

人脸检测是一种基础又被普遍使用的机器视觉技术，被实际应用于身份验证、真人识别、情绪推测等领域。人脸检测是确定被检测图像中有没有人脸并框出来以从图像中分离出人脸区域。本疲劳检测方法提取特征过程中，首先进行人脸检测，能够将人脸从复杂背景中分离出来，在这个基础上再进行面部、头部特征的检测，提高了算法的准确度及处理效率。由于人脸信息多和非接触式的便捷获取等优势，人脸检测相关技术广泛应用于生物信息检测领域。另一方面，计算机视觉处理技术的日渐成熟和计算设备的更新迭代，也为人脸检测技术提供了越来越好的研究基础。日前，人脸检测方法主要有基于经验知识、模板匹配和样本统计的人脸检测^[19]。

基于经验知识的人脸检测是根据皮肤颜色、面部轮廓等人脸特征信息进行人脸检测。其中肤色相对稳定，与图像背景差别大，容易分割出来，尤其赢得研究者的青睐。何俊^[45]等人提出肤色模型进行人脸的一个初始定位，肤色模型是将颜色空间转换到 HSV，选取合适的色度阈值分离出人脸并进行二值化，肤色模型检测方法直观，速度较快，但是可靠性不是那么好。

基于模板匹配的人脸检测方法利用预先建立的标准模板与图像中像素或区域去依次匹配。Sakai 等人早在上个世纪 60 年代就通过眼鼻、面部轮廓等局部特征模板进行正面面部模型搭建^[46]。Kass 等人更是搭建一个主动人脸轮廓模型^[47]。基于模板匹配的检测方法实现起来相对简单，但是标准模板决定了检测效果，而且不适用于人脸旋转等情况。

基于样本统计的人脸检测方法通过带有标签的人脸正负样本进行各种人脸分类器的训练，再用训练好的分类器来检测人脸，如主量分析法、支持向量机、AdaBoost。其中 AdaBoost 方法可靠、稳定、高效，广泛应用于人脸检测，因此本文选择这种方法实现人脸的初步定位，为人脸特征的检测剔除复杂的背景，且留下相对简单稳定的人脸对象，在缩小范围的同时大大增加了算法的可靠性。

3.2 经典 AdaBoost

AdaBoost 复杂的检测过程也是一点一滴积累下来的，从 1984 年强弱学习概念的提出^[48]，到 1990 年需提供训练正确分类下限的 Boosting 算法的提出^[49]，终于在 1995 年 Freund 和 Schapire 在这些基础上最早提出了迭代算法 AdaBoost(Adaptive Boosting)^[50]，紧接着 Viola 和 Jones 把积分图应用到 AdaBoost 中，给人脸检测提速，最终达到算法检测的实时性^[51]。AdaBoost 不需要依赖主观的经验知识和经不起变换匹配的模板，只需要利用客观的脸部特征提早训练好人脸分类器，从而实现高效人脸检测。

AdaBoost 通过分别以逐渐增大的各种基础特征依次扫描整个待测图像，经该图事先计算好的积分图查表得到基础特征的值，选取阈值便得到一个弱分类器，进一步线性加权构建强分类器，最后简单排序串联搭建级联分类器。其中，弱分类器和强分类器的权重需要随着训练轮数迭代更新。弱分类器权重根据上一轮样本分类的错误情况进行调整，增大对分错样本的权重，集中对分错样本的训练力度，提高弱分类器的可靠性。强分类器权重根据上一轮弱分类器分类的错误率进行调整，减小错误率高的弱分类器的权重，提高强分类器整体的准确度。

3.2.1 haar 矩形特征

像素的信息丰富却分布零散，不利于图像局部特征的表达，而基础特征可以编码特定区域的状态。在相同的数据量下，相对于基于像素的检测而言，基于特征的检测提速不少^[52]。

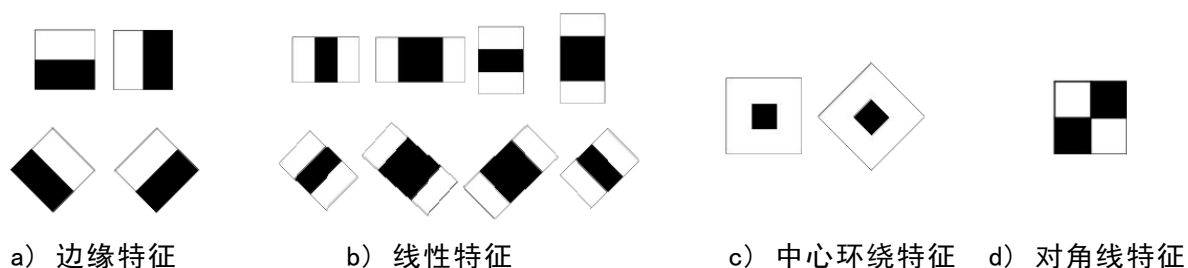


图 3.1 haar 特征类型

Haar 矩形特征可以表达该矩形框内的图像灰度值。Haar 特征一开始由 Papa 等提出，也被 Viola 和 Jones^[51]用来进行人脸检测，达到了非常不错的效果。Haar 矩形特征是由简单的黑白矩形框组成，如图 3.1 所示，人脸检测根据人脸的灰度分布特点用到的基础特征主要是前三种。Haar 矩形特征可以表达图像灰度信息，因种类、大小、位置的不同而非常丰富。根据人脸眼睛颜色较周围深，鼻梁较两侧颜色浅，嘴巴的颜色较周围深，可以用简单的 haar 矩形特征表示出来，如图 3.2 所示，(b) 中用中心环绕特征和边缘特征来表达眼睛灰度分布，(c) 中用线性特征来表达鼻子的灰度分布，(d) 中用边缘特征来表达嘴巴的灰度分布。

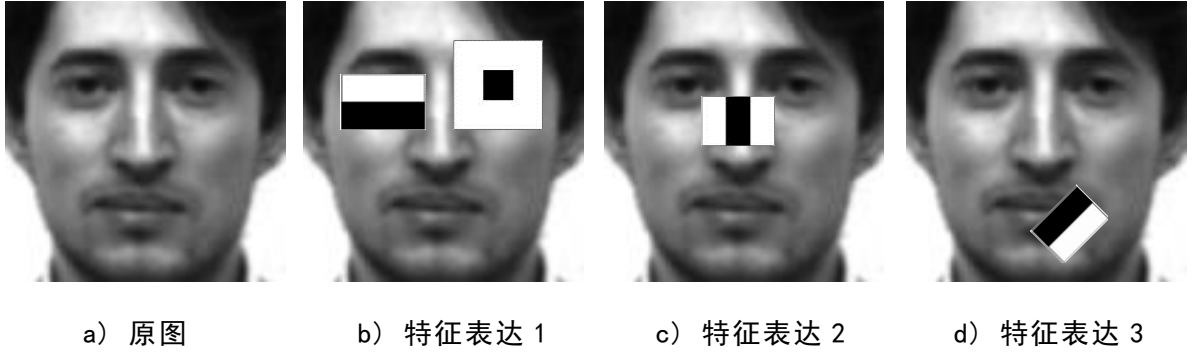


图 3.2 haar 特征对人脸特征的表达

各种各样的 haar 矩形特征最后会统一到数据上来，数字图像中用一个数值来表达，haar 矩形特征的值是由白色矩形框内所有像素的和减去黑色矩形框内所有像素的和所得。多个 haar 矩形特征进一步构成一个图像局部特征的描述，计算其中 haar 矩形特征个数的方法由 Lienhart 等人提出^[53]。在 $W \times H$ 的图像局部特征中，有如图 3.3 所示类型和尺寸的 haar 矩形特征，其在水平方向上的最大缩放比例取 $X = [W/w]$ ，在竖直方向上的最大缩放比例取 $Y = [H/h]$ ，然后可以计算得到这两种 haar 矩形特征在图像局部特征中的个数：

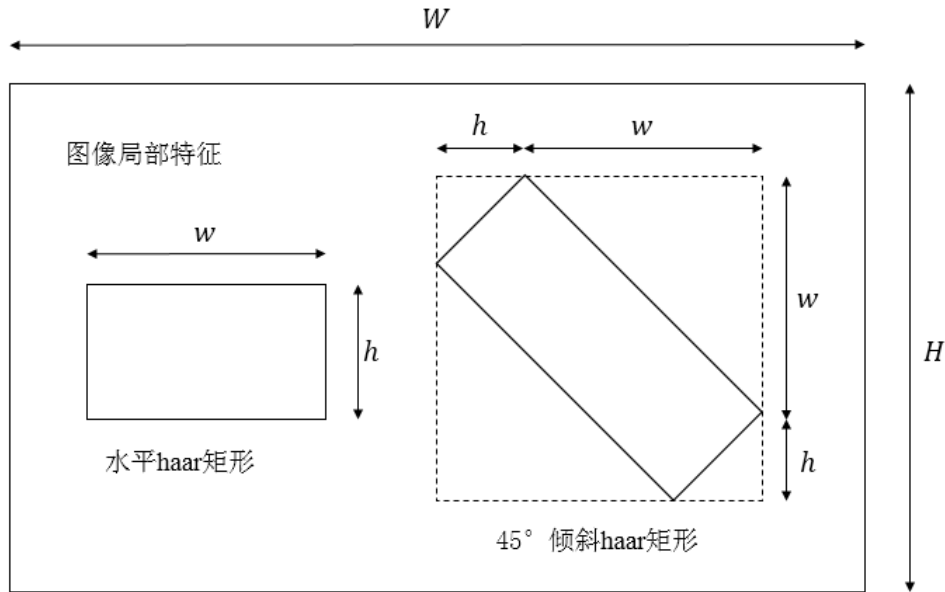


图 3.3 图像局部特征中 haar 特征示意

该图像局部特征中水平 haar 矩形的数量为：

$$XY \left(W + 1 - w \frac{X+1}{2} \right) \left(H + 1 - h \frac{Y+1}{2} \right) \quad (3.1)$$

该图像局部特征中 45° 倾斜 haar 矩形的数量为：

$$XY \left[W + 1 - (w + h) \frac{X+1}{2} \right] \left[H + 1 - (w + h) \frac{Y+1}{2} \right] \quad (3.2)$$

3.2.2 积分图

图像的基础特征数量庞大，为样本集带来了巨大的计算压力，检测的实时性急切地要求寻找一种方法解放庞大的计算量。积分图的提出大大简化了基础特征值的计算，为实时监测提供了可能。积分图构造了带有位置信息的中间表格，某图像区域的像素灰度值总和可以由该区域的位置查表得到，使得图像中所有区域的像素和只需要遍历一次图像进行计算即可，减少运算次数的同时免去了很多枯燥重复的运算。

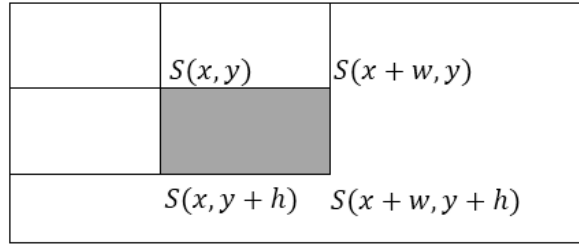


图 3.4 图像区域求和

图像的积分图是带有位置索引的所有像素点积分图的集合。图像中某一像素点的积分图是指以整个图像左上角为起点，以该像素位置为终点，这两点为对角所构成的矩形框内所有像素灰度值之和，如式 3.4 所示。像素点积分图的计算不需要按照定义求和，可以用与上一像素点增量的形式求得。如图 3.4 所示，图像阴影区域所有像素点之和可以通过积分图简单的加减计算得到，效率很高：

$$Sum(Rec) = S(x, y) + S(x + w, y + h) - S(x + w, y) - S(x, y + h) \quad (3.3)$$

其中， $S(x, y)$ 指像素点 (i, j) 的积分图：

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j) \quad (3.4)$$

由式 3.3 可知，不管尺度如何转变，任意水平矩形区域的像素和可以通过查找四次积分图进行常量计算得到，进而计算 haar 特征值。

3.2.3 迭代 AdaBoost

特征的计算问题已经由积分图解决，而 AdaBoost 迭代方法是用来从不同种类、位置和大小基础特征中选取最有效的，haar 特征相当于 AdaBoost 中的弱分类器。如表 3.1 所示，不同窗口大小的 haar 特征数量不同且多^[54]。

表 3.1 不同大小窗口的 haar 特征数量

窗口	特征数
16×16	32384
20×20	78460
24×24	162336
30×30	394725

AdaBoost 迭代级联算法首先用不同的 haar 特征扫描训练样本集，得到不同的弱分类器，每一次训练选取错误率最小的弱分类器并根据该最小的错误率加权形成该次训练的强分类器，经过若干次训练后形成若干个强分类器，通过识别速率进行重新排序形成单通道多层级级联分类器。对同一含 N 个样本的训练集每次训练遍历 S 次，总共训练 T 次，AdaBoost 迭代级联算法整个流程图如图 3.5 所示。

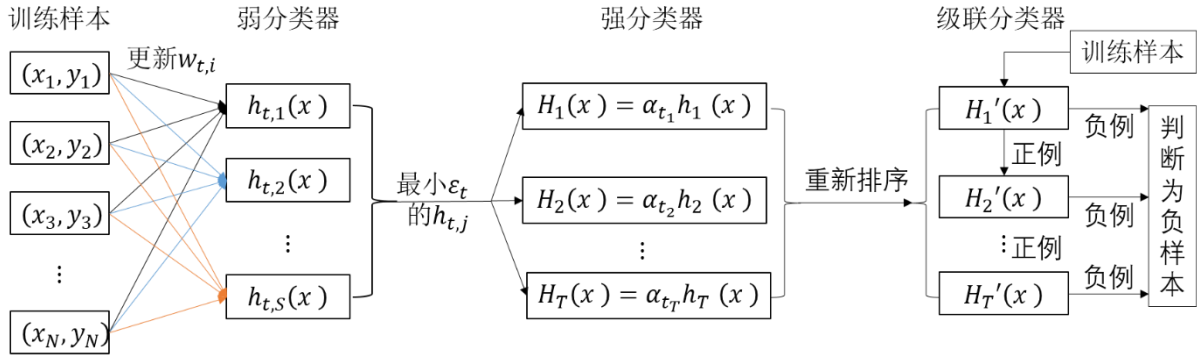


图 3.5 AdaBoost 流程图

弱分类器的生成在于在 haar 特征值基础上选择一个合适的阈值，其生成的具体过程^[19]如表 3.2 所示。

表 3.2 弱分类器生成流程

正样本权重总和 P ，正样本权重总和 N ，特征值排序后 f_i 前正样本权重总和 P' ， f_i 前负样本权重总和 N' ，偏置符 p_i

- 1 依次计算训练集中样本特征值 $f_i(x)$
- 2 排序所有样本特征值，并获得范围 $[f_{\min}, f_{\max}]$
- 3 **for** $f_i(x) \leftarrow f_{\min}$ **to** f_{\max}
- 4 **do** 计算误差 $e = \min\{P' + (N - N'), N' + (P - P')\}$
- 5 选择最小 e 所对应的阈值，作为该弱分类器合适的阈值 θ_i
- 6 得到该特征 $f_i(x)$ 对应的弱分类器 $h_i(x) = \begin{cases} 1, & p_i f_i(x) < p_i \theta_i \\ 0, & \text{其他} \end{cases}$

一次训练中每遍历一次训练样本集便生成一个弱分类器，每次训练遍历 S 次共生成 S 个弱分类器。弱分类器的迭代更新主要在于样本权重的变化，首先平均初始化样本权重并归一化，然后根据此次训练弱分类器的最小误差并结合样本分类正确与否，更新下一次的样本权重。这种样本权重的更新方法能够让下一次的训练中错误样本得到更高的关注度，从而让弱分类器越来越有效。但是弱分类器的阈值设得较低，所以其效果仅仅比随机分类好一点，虽然容易处理但是可靠性不高。

每次训练产生一个强分类器，训练 T 次共产生 T 个强分类器。强分类器只与该次训练的弱分类器有关，而与上次训练的强分类器无关。强分类器选取本次训练中误差最小的弱分类器，并且其权重由此最小误差计算得到。现实生活中很难一

次性找到错误率很低的分类器，但如果是找大量分类勉强的分类器，然后对其选择迭代便很容易，整体分类器的效率也可以达到不错。AdaBoost 的强分类器就是采取这样的思路，迂回却效果好。

强弱分类器的更新迭代的具体流程^[19]如表 3.3 所示。

表 3.3 强弱分类器迭代流程

训练集中有 N 个样本，其中有 p 个正样本， n 个负样本	
1	平均初始化样本权重 $w_{1,i} = \begin{cases} 1/(2p), & y_i = 1 \\ 1/(2n), & y_i = 0 \end{cases}$
2	for $t \leftarrow 1$ to T
3	do (a)归一化样本权重 $w_{t,i}$
4	(b)计算遍历整个样本的弱分类器的误差 $\varepsilon_j = \sum_{i=1}^N w_t h_j(x_i) - y_i $
5	(c)选择有最小误差 ε_t 的弱分类器 $h_t(x)$ ，并计算 $\beta_t = \varepsilon_t / (1 - \varepsilon_t)$
6	(d)更新样本权重 $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$ ，其中样本正确分类时 $e_i = 0$ ，否则 $e_i = 1$
7	(e)计算强分类器权重 $\alpha_t = \log(1/\beta_t)$ ，
8	得到强分类器 $H(x) = \begin{cases} 1, & \sum_{i=1}^T \alpha_t h_t(x) \geq 0.5 \sum_{i=1}^T \alpha_t \\ 0, & \text{其他} \end{cases}$

一个强分类器实现的实际效果往往达不到实际检测要求，我们需要将多个强分类器串联搭建级联分类器以提高可靠性。如图 3.5 所示，先将所有强分类器由简单至复杂依次排序，然后把图像子窗口输入到级联分类器中，若被其中一个强分类器判断为背景区域则被弃，只有所有的强分类器全部判断为目标才会最终输出为目标区域。强分类器如此排序与串联方式可以保证非目标区域很快的过滤掉，从而把精力和时间集中到目标区域上，最后输出检测目标。

3.3 改进的 AdaBoost

考虑到经典 AdaBoost 中使用的 haar 特征灰度计算量挺大，并且绝对灰度容易受到光照的影响，所以可以寻找其他的基础特征，以避免这些缺陷，其中 LBP(Local Binary Pattern)特征就是一种很好的替代选择^[55]。LBP 特征表达的是相对灰度，可以很好地适应不同的光照，另外一方面数值比较减少了大量的计算，可以大幅提高整个算法的效率。

3.3.1 LBP 特征

LBP 特征表达了图像纹理信息，由 Ojala 等人早在 1996 年提出^[56]，然后经逐渐改善，又在 2002 年提出多尺度的、具有旋转不变性的改进 LBP 特征。LBP 特征关键思想在于，将全部周围像素都与中心像素灰度值比较的结果转化为二进制码，描述图像局部纹理信息。LBP 特征计算非常简单，具有较强的灰度不变性，

分类效果好，作为一种无参数方法，无需提前设定其分布。凭借其众多的优势，LBP 特征得到广泛应用。

1. 最初 LBP 特征

如图 3.6 所示，最初的 LBP 特征就是取图像中 3×3 像素作为矩形子窗口，将中央像素灰度值作为该窗口阈值，分别与周围 8 个像素作比较，像素灰度值超过该阈值的像素值更新为 1，未超过该阈值的像素值更新为 0，将得到的二进制码从左上角顺时针依次排布组成一个 8 位二进制数，转至十进制数作为该子窗口的 LBP 码，以反映此局部区域的纹理特征。

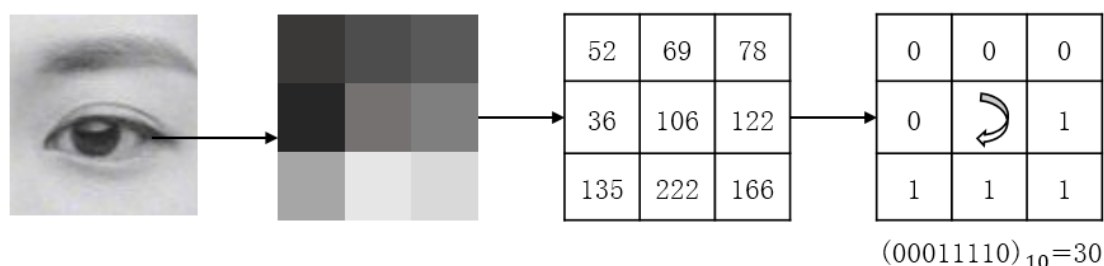


图 3.6 最初 LBP 特征示意图

比起 haar 特征，最初 LBP 特征检测速度更快，还有一个显著的优势在于 LBP 特征的灰度不变性。如图 3.7 所示，在包括白光和暗光在内的不同光照条件下，LBP 特征的灰度分布基本可以保持不变^[55]。

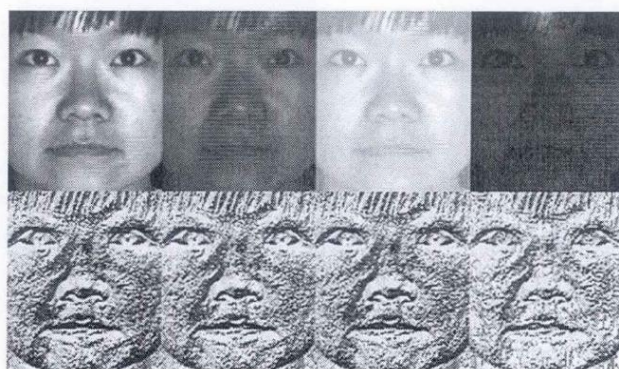


图 3.7 LBP 特征的灰度不变性

2. 圆形 LBP 特征

由于最初的 LBP 特征只表达了 3×3 邻域范围内的灰度变化信息，噪点给它带来了很大的误差，很难描述大尺度结构的图像灰度变化。从最初 LBP 特征尺度方面的局限性出发，Ojala 等人完善了最初 LBP 特征，提出有 P 个采样点、以 R 为圆形半径的圆形 $LBP_{(P,R)}$ 特征，可以表达各种大小邻域的纹理信息，实现不同尺度的表达。几种大小的圆形 $LBP_{(P,R)}$ 特征如图 3.8 所示。

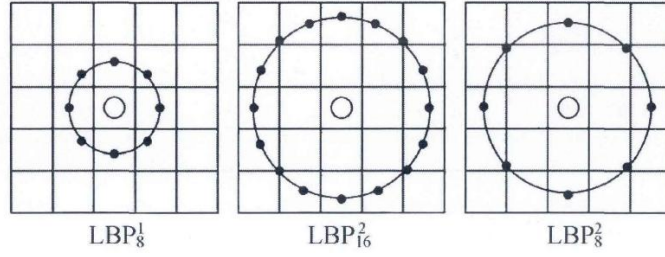


图 3.8 圆形 $LBP_{(P,R)}$ 特征示例

容易看出，旋转图像得到的 LBP 码因其起点的不同可能会发生变化，所以 LBP 特征没有旋转不变性。针对这个问题，Maenpaa 等人继续改进 LBP 特征，提出其旋转不变模式，先持续旋转 LBP 特征中的圆形邻域获得一组 LBP 码，再取当中最小的作为该圆形邻域的 LBP 码。如图 3.9 所示，很好地展示了在旋转不变模式下的 LBP 特征求取唯一确定的 LBP 码的过程，值为 30 的 LBP_8^1 特征经旋转可以得到 8 个其他值的 LBP 特征，但旋转不变模式下只有一个编码为 00001111 的旋转不变模式 LBP。 LBP_8^1 特征有 256 种形式，但是其旋转不变模式可能性只有 32 种。

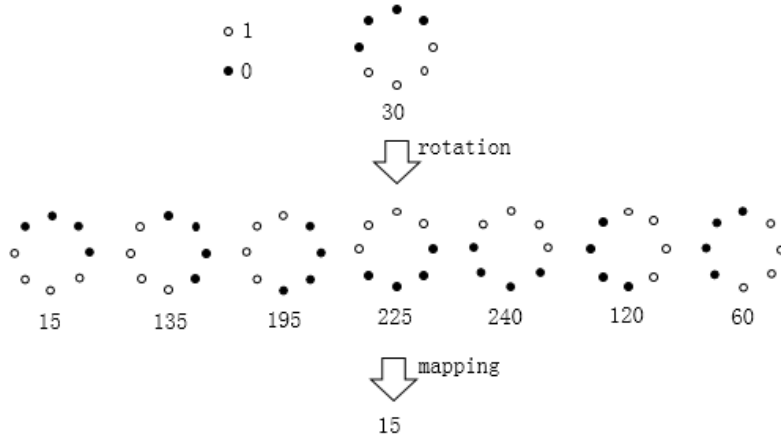


图 3.9 旋转不变模式 LBP 特征示例

3.MB-LBP 特征

上文为解决最初 LBP 特征在尺度描述上的限制问题，圆形 LBP 特征扩大尺度的同时却忽略了影响最大的最邻近像素，MB-LBP 特征因为是将多个相邻区域内的大小来作对比，所以能够用来表达图像大尺度结构的灰度变化信息，又因为是将区域像素的灰度平均值来作对比，所以可以对噪声进行过滤而减少噪声干扰。因此，本文利用 MB-LBP 特征^[57]来作为检测人脸的基础特征。

MB-LBP 特征的计算如图 3.10 所示，首先将待检测的矩形区域划分成 3×3 个相同大小的矩形块区，并计算每个矩形块区内所有像素的平均值，然后将边上的 8 个矩形块的平均值与中心矩形块作比较，若比中心矩形块的大则二值化为 1，反之为 0，从二值化窗口的左上角顺时针依次取值得到 MB-LBP 特征的二进制编码，最后转换成十进制数就得到了 MB-LBP 特征值。MB-LBP 特征值的具体计算如式

3.5 所示，其中 g_c 是中心矩形块内所有像素灰度的平均值， g_j 是第 j 个相邻矩形块内所有像素灰度的平均值， j 是从取值为 0 的位于左上角的相邻矩形块开始顺时针依次递增。

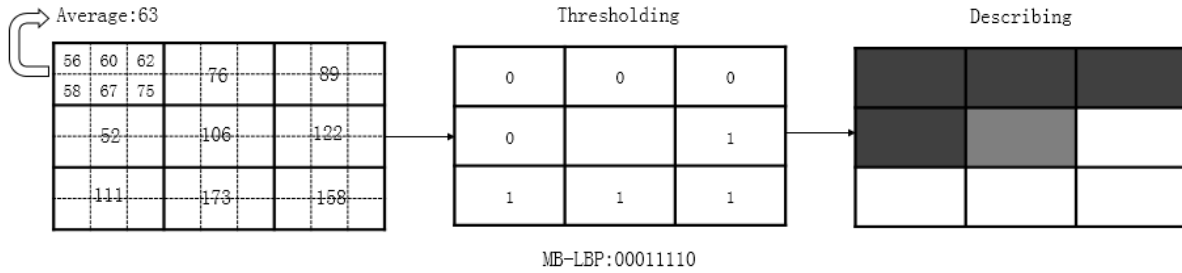


图 3.10 MB-LBP 特征计算示意

$$LBP = \sum_{j=0}^7 s(g_j - g_c) 2^j \quad (3.5)$$

$$s(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases} \quad (3.6)$$

总的来说，与 haar 特征和其他几种 LBP 特征，MB-LBP 特征有一些显著的优势，如表 3.4 所示。因此，本文的 AdaBoost 方法选用 MB-LBP 特征作为检测基础特征。

表 3.4 MB-LBP 特征的优势

序号	参照特征	MB-LBP 特征的优势
1	haar 特征	灰度不变性，在很大程度上减少了光照的影响
2	haar 特征	数量少，非度量，不需排序特征值经阈值分类，增快速度
3	圆形 LBP	尺度更大，大尺度下也考虑到了邻近像素，可靠性更好
4	其他 LBP	用平均灰度值替代整个矩形块，减少了噪点干扰
5	其他 LBP	特征大故图像所需的特征数目较少，在很大程度上提了速

3.3.2 AdaBoost 的不同

基于 MB-LBP 特征的 AdaBoost 与基于 haar 特征的分类器生成和迭代更新大同小异，只有弱分类器的衡量值和分类器中权重归一化处理不同，其他基本相同，故对相同的不加以赘述，重点阐述这两个不同点。

1. MB-LBP 特征的衡量值

与 haar 特征本质不同的是，MB-LBP 特征值仅仅只是个二元模式(共 256 种)的符号，不可度量，无大小之分。因为它的这种性质，基于 MB-LBP 特征的 AdaBoost 方法中弱分类器的生成没有类似 haar 特征值的阈值，所以我们需要选择一种合适的方法来找到在基于 MB-LBP 特征中类似于 haar 特征值的一个衡量值。

本文构造基于 MB-LBP 特征的 AdaBoost 中的弱分类器采用的是多分支树结

构的方法^[55], 分支总数为 256, 图像待测区域内第 j 个特征所对应的值 $f_j(x)$ 如式 3.7 所示:

$$f_j(x) = \begin{cases} a_0, & x^k = 0 \\ \dots & \dots \\ a_m, & x^k = m \\ \dots & \dots \\ a_{255}, & x^k = 255 \end{cases} \quad (3.7)$$

其中另外补充说明的参数: 离散值 m 从 0 取值到 255, 分别对应上文所提到的 MB-LBP 特征值 256 种二元模式符号; a_m 取值范围是 $[-1, 1]$, 需要通过训练样本获得, 如式 3.8 所示:

$$a_m = \sum_{i=1}^N w_i y_i \delta(x_i^k = m) / \sum_{i=1}^N w_i \delta(x_i^k = m) \quad (3.8)$$

通过遍历共含 N 个样本训练集, 得到 a_m 。其中, $\delta(x_i^k = m)$ 是一个判断函数, 表示若第 i 个样本的第 k 个分量等于 m , δ 函数值为 1, 反之为 0。 $a_m > 0$ 时的样本更有可能是正样本, 因此也可以找到一个基于 MB-LBP 特征的弱分类器的阈值 θ , 而基于 MB-LBP 特征的样本 $f(x)$ 函数值相当于基于 haar 特征的特征值, 当样本 $f(x) \geq \theta$ 时属正样本, 并且样本 $f(x)$ 函数值越大越有可能属于正样本。

2. MB-LBP 特征样本权重的归一化

为了让分类器的样本权重更稳定, 不像基于 haar 特征的分类器正负样本权重使作为整体进行统一归一化的, 基于 MB-LBP 特征的正负样本权重是分别进行归一化的:

$$w_{t,i} = w_{t,i} / 2 \sum_{i=1}^{m_1} w_{t,i} \quad (3.9)$$

$$w_{t,j} = w_{t,j} / 2 \sum_{j=1}^{m_2} w_{t,j} \quad (3.10)$$

式 3.9 和式 3.10 表示, 对含 m_1 个正样本、含 m_2 个负样本的训练集进行的第 t 次训练时, 正样本权重 $w_{t,i}$ 在正样本训练集中进行归一化处理, 而负样本权重 $w_{t,j}$ 在负样本训练集中进行归一化处理。

3.4 效果对比与实验分析

Viola 和 Jones 早在 2001 年就用基于 haar 特征的 AdaBoost 方法完成人脸检测, 效果显著, haar 特征和 AdaBoost 方法开始普遍应用到人脸检测领域。但是相比于 haar 特征或者最初 LBP 特征, MB-LBP 特征却有一些优势, 在人脸检测方面能够达到更好的效果, 本小节便从效果、速度进行详细的对比, 并通过实验进行验证。

3.4.1 效果上的提升

从特征结构上来看, MB-LBP 特征具有更加细致的结构, haar 特征通常只有 2 个框最多也只有 4 个框, 但是 MB-LBP 特征却始终都有 9 个矩形框, 所以 MB-

LBP 特征可以描述人脸更加精细的灰度特征。如图 3.11 所示，中心 haar 特征对眼睛的描述只是周围整体与眼球框的简单的差值，但是 MB-LBP 特征却将周围整体分成了 8 个小矩形，分别描绘了与中心矩形的一个更细致的变化。

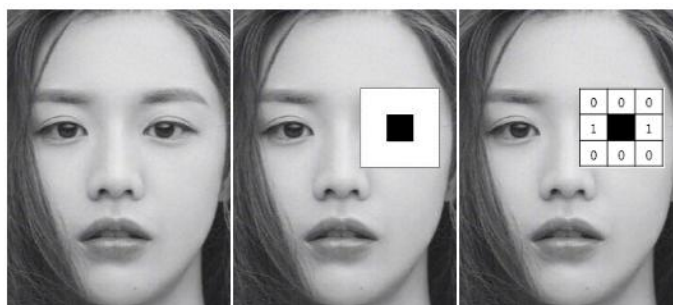


图 3.11 特征结构的对比

效果上更主要的提升是，MB-LBP 特征大幅度地减少了光照强度的影响。LBP 特征的灰度不变性，使得检测抗光照强度干扰能力很强。在此基础上，MB-LBP 特征更是降低了噪点的干扰，能够在更恶劣的光照环境中，达到不错的人脸检测效果。如图 3.12 所示，在过暗或者高光的条件下，基于 MB-LBP 特征的分类器能够识别多种姿态下的人脸。但是 haar 特征较 MB-LBP 特征对光照更敏感，在不太好的光照条件下往往无法有效的检测到人脸。因此，MB-LBP 特征可以有效地应用于亮度范围更广的图像，在光线多变复杂的驾驶环境中拥有更好的检测效果。

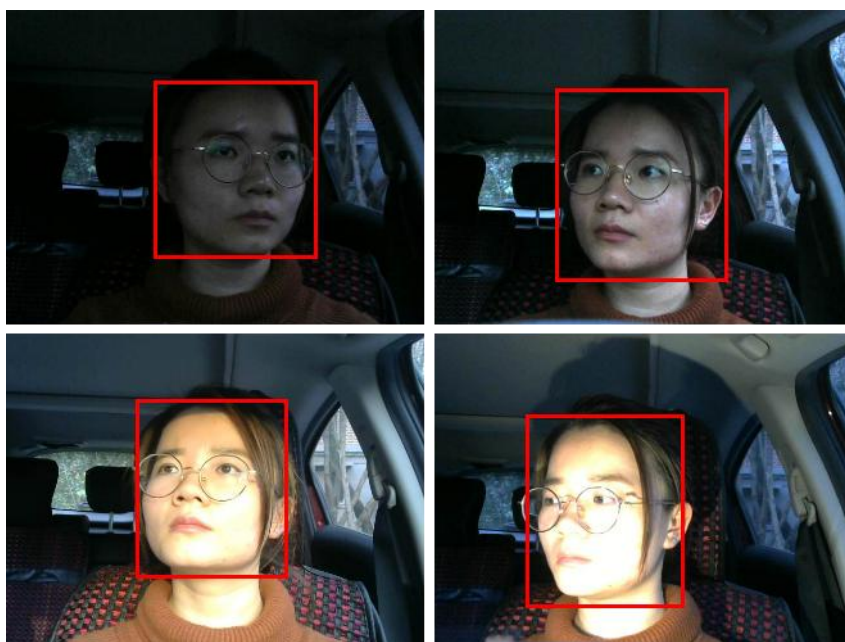


图 3.12 MB-LBP 特征分类器的检测效果

3.4.2 速度上的提升

相对于 haar 特征，一方面 MB-LBP 特征的数量少很多，haar 特征的数量几乎几十倍于 MB-LBP 特征；另外一方面，MB-LBP 特征无大小之分，迭代训练时无需排序特征值^[55]。所以，MB-LBP 特征的训练和检测速度比 haar 特征要快很多。

Zhang Lun^[57]等人分别利用基于 haar 特征、最初 LBP 特征、MB-LBP 特征的 AdaBoost 算法训练了各自对应的 50 个特征,并在同一个测试集上进行人脸检测,得到的差异十分明显,如图 3.13 所示。由图可得,训练 haar-like 特征、最初 LBP 特征、MB-LBP 特征获得的分类器在特征个数相同时其误检率依次递减,在误检率相同时检测率依次递增。切换另一个角度来看,当分类器误检率相同时,所需的 MB-LBP 特征的特征个数最少,那么 MB-LBP 特征分类器的训练和检测速率也最快。

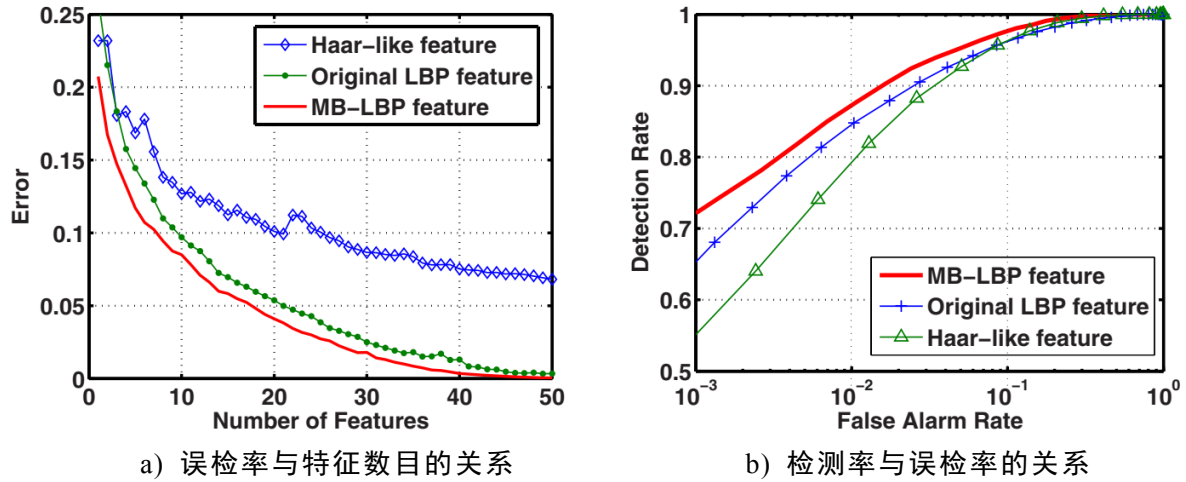


图 3.13 各种特征的检测结果对比

为具体地验证 MB-LBP 特征在速度方面的优越性,将数量同为 100 的 haar 特征和 MB-LBP 特征各自来训练分类器,依次记录训练、检测所需时间,如表 3.4 所示。由表可知,基于 MB-LBP 特征的分类器训练时间和平均检测时间是 haar 特征的一半,故 MB-LBP 特征的分类器速度较 haar 特征快一倍,体现了其拥有更好的速度性能。

表 3.4 haar 与 MB-LBP 特征的分类器时间对比

分类器特征	训练时间(s)	平均检测时间(ms)
haar 特征	5815	66
MB-LBP 特征	2592	32

3.4.3 实验分析

考虑到要贴合现实生活里摄像头所拍到的驾驶员图片,本文实验样本选取加州理工学院人脸图像库,该实验样本共有 450 张 896×592 的上半身人体图像,光线和背景各不相同。从中随机选择了 30 张图像,如图 3.14 所示。

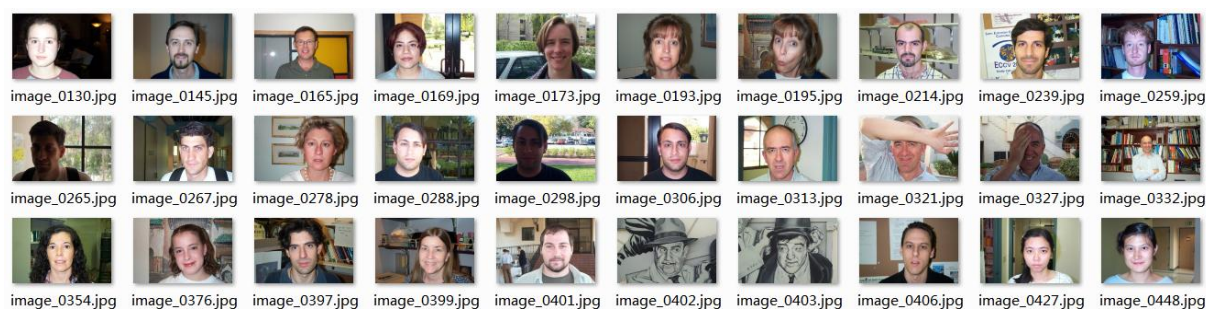


图 3.14 部分实验样本展示

将该样本作为同一个测试集，分别输入到基于 haar 特征的 AdaBoost 和基于 MB-LBP 特征的 AdaBoost 得到的分类器当中去，进行实验验证，两种方法的图片检测效果如图 3.15 所示，数值检测结果如表 3.5 所示。



a) 基于 haar 特征的 AdaBoost 检测 b) 基于 MB-LBP 特征的 AdaBoost 检测

图 3.15 两种方法的检测效果

表 3.5 两种方法的测试结果

人脸检测方法	检出数量 (张)	误检数量 (张)	检测率	误检率	检测时间 (ms)
haar 特征	425	42	94.44%	8.65%	59
MB-LBP 特征	438	30	97.33%	6.67%	32

由于测试集样本的主要区别在于光线和背景，测试样本分辨率高，人脸数几乎只有一个，两种方法的检测率都比较高，误检率也不高，但是两者还是存在一定差异。由表 3.5 可知，从人脸图像的检出数量、误检数量、检测率、误检率和检测时间这几个方面的对比可以看出 haar 特征、MB-LBP 特征对于人脸检测的性能差异，后者不仅效果更好，而且速度更快，所以本文人脸检测方法选取基于 MB-LBP 特征的 AdaBoost 方法。

为验证基于 MB-LBP 特征的 AdaBoost 方法在实际驾驶环境下人脸检测的具体情况，本文分别于白天和夜晚自行拍摄了两段驾驶员视频，每段视频帧速为 30 帧/s，持续 15 分钟。将两段视频序列依次输入基于 MB-LBP 特征的 AdaBoost 方法得到的分类器中，检测人脸，检测效果如图 3.16 所示，统计结果如表 3.6 所示。

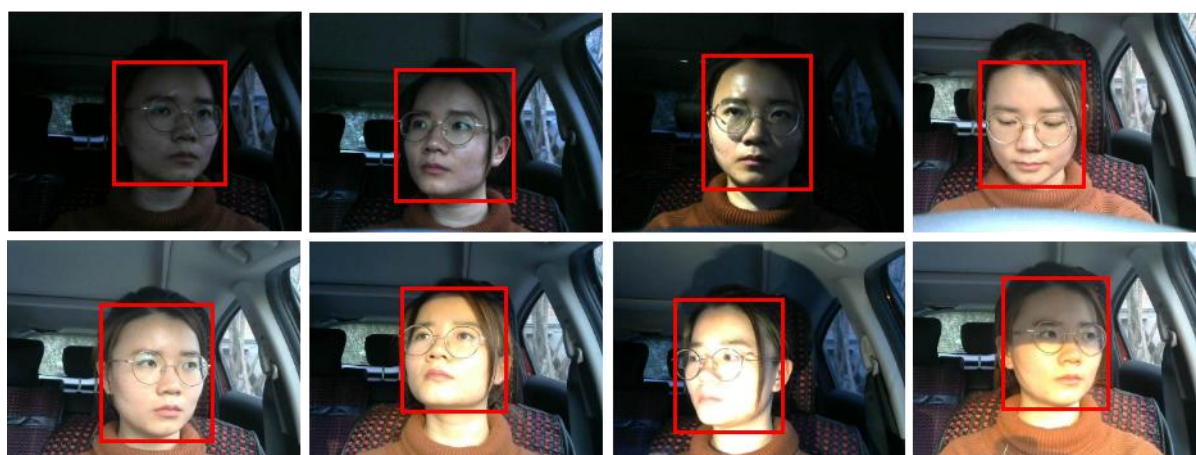


图 3.16 人脸检测效果图

从图 3.16 可以看出，尽管是在夜晚这种光线差的环境下，给与一定的光补偿后，光照强度也可满足本文检测方法的要求；还有不管在阳光直射存在曝光还是半边遮阳时，依然能准确检测人脸。因此本文方法在各种光照条件下都准确检测到人脸。

表 3.6 不同光照下视频序列测试统计

视频环境	检测率	误检率	平均检测时间(ms)
白天	98.13%	6.24%	30
夜晚	96.73%	7.14%	32

夜晚在一定补偿的情况下，人脸检测统计效果只是略差于白天。由此可见，基于 MB-LBP 特征的 AdaBoost 方法受光照的干扰很小，两段视频序列人脸检测率平均值达到 97.43%，误检率平均值只有 6.69%，检测每帧图片平均花费 31ms 的时间。

3.5 本章小结

本章首先简单介绍了人脸检测三种分类方法，并选择 AdaBoost 方法，对经典的 AdaBoost 方法的流程从 haar 特征、积分图到弱分类器、强分类器和级联分类器的搭建进行详细的介绍，接着详细地介绍了改进的 AdaBoost 方法，其中对 LBP 特征各种形式作对比，选择 MB-LBP 特征，并介绍了 MB-LBP 特征构成的 AdaBoost 与经典方法的不同地方。最后经实验分析对比，MB-LBP 特征训练的分类器在效果和速度上都有了很大的提升，并且经白天和夜晚两组拍摄视频序列验证，本文所选择的人脸检测方法，即基于 MB-LBP 特征的 AdaBoost 算法，有 97.43% 的平均检测率，仅 6.69% 的平均误检率，和 31ms/帧的检测速率。

第 4 章 特征提取与计算

前两章图像预处理和人脸初始化定位都是为了得到更准确、更有效的疲劳特征，本章所讲的特征提取与计算成为驾驶员疲劳检测过程中最为关键的一步。眼部状态、嘴部状态以及头部姿态是其重要的检测指标。本章通过 Stasm 得到面部基础特征，结合相机位姿测量得到点头角度，经初始化并求得特定驾驶员各参数阈值，计算具体的疲劳特征，为后续疲劳检测做好数据准备。

4.1 面部基础特征

4.1.1 特征提取方法概述

对于人脸各特征的提取与计算，分为基于知识和基于机器学习的方法。

基于知识的方法检测验证器官位置、肤色特征、纹理、灰度、形状规则等特征，是不是与人脸的经验知识相符。WeiSun 等人^[58]结合人脸的结构特征“三庭五眼”，如图 4.1 所示，再在一定的区域范围内分别对眼睛、鼻子等进行针对性的检测。

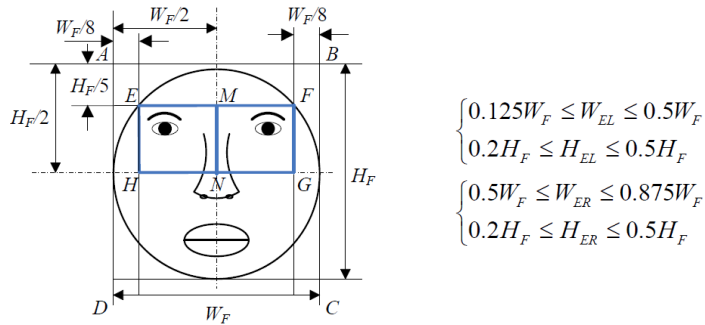


图 4.1 “三庭五眼”结构图

对眼睛局部而言，姚胜等人^[59]提出对分割出的眼部矩形区域进行水平积分，通过睁眼与闭眼时水平积分的明显不同，选取合适的阈值可以判断眼睛开闭，该方法思路清晰，不过投影图像中干扰噪声太大，曲线不规则，阈值的选取比较复杂，而且当脸部稍微倾斜一点后，会出现明显的不可靠。另外，WeiSun^[58]等人提出通过区域连通分析进一步去噪，然后通过水平与垂直的积分投影得到眼睛的宽高比，进而判断眼睛开闭。

对嘴部局部而言，也可以类似地考虑上述两种方法，用的比较普遍的是唇色聚类的方法，但对于实际的背景复杂的图像而言，唇色相对于肤色的聚类分布存在变动。对此黄永慧等人^[60]提出一种自动调试分割唇色的常量参数的方法，首先手动提取 200 幅图像唇部，得到与对应 AdaBoost 识别得到的 ROI 区域的平均比值，然后通过逐渐逼近的方法求出调整的常量参数，实验结果也表明该方法能够

更好地适应各类图像，完成唇动检测。

基于机器学习的方法指利用一种特征或者形状通过学习训练得到分类器或者模型，再直接应用于实际的检测与匹配。其中，典型的代表是眼睛、嘴巴的 haar 分类器和主动形状模型 ASM。主动形状模型是将整个局部特征作为一个整体来进行检测，并且同时实现了检测与校准，特征提取效果好速度快。堆叠主动形状模型 Stasm 是在主动形状模型基础上进一步完善得到的，拥有更好的性能，因此本文的特征提取方法选择 Stasm 算法。

4.1.2 主动形状模型

ASM 算法提取统计参数化后的特征。它的思想是：手动标记每一个训练样本的关键特征点，得到一组形状向量集，并在每一层图像高斯金字塔上建立初始形状模型和局部轮廓模型，据此从图像高斯金字塔的低分辨率至高分辨逐渐搜索测试目标，即在形状初始化后，寻找距当前初始点马氏距离接近的点，作为合适的匹配特征点^[61]。ASM 算法流程图如图 4.2 所示。

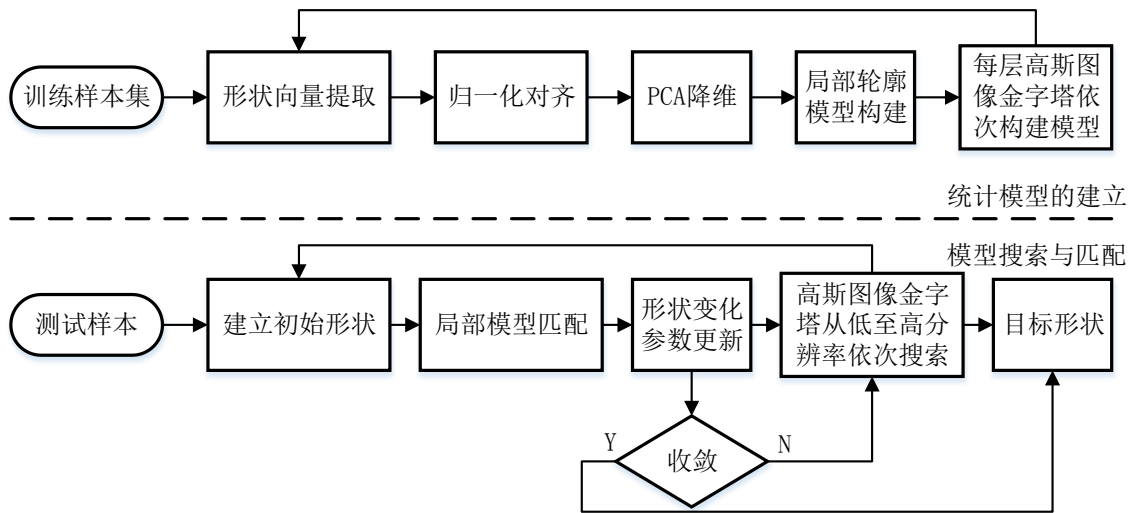


图 4.2 ASM 算法流程图

ASM 算法主要包括高斯图像金字塔中统计模型的建立和模型搜索与匹配这两大部分：

1.ASM 统计模型的建立

ASM 通过一系列关键特征点构成形状向量来描述某一形状，被称为点分布模型。一般来说，关键特征点首选那些能充分表达物体形状和生理特征的点，如边缘，拐角，T 连接等处，并在此基础上等间隔地取点以连接上述轮廓特征点。

如图 4.3 所示，按序人工标记每一个人脸样本，得到 68 点形状模型。在含 n 个图像样本的训练集中，第 i 幅图像样本的形状向量可表示为 $s_i = [x_{i0}, y_{i0}, \dots, x_{ij}, y_{ij}, \dots, x_{i67}, y_{i67}]$ ，式中 (x_{ij}, y_{ij}) 指第 i 幅图像样本的第 j 个关键特征点的坐标， $i = 1, 2, \dots, n$ ， $j = 0, 1, \dots, 67$ 。

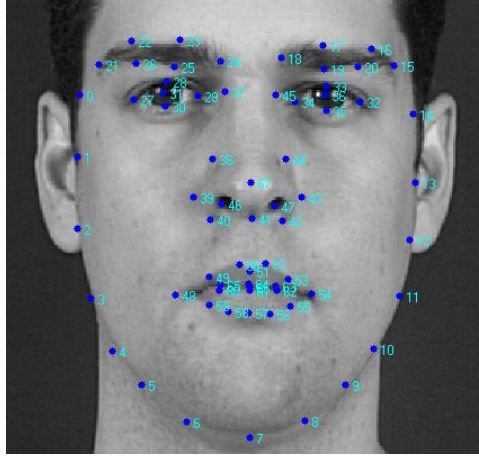


图 4.3 人脸 68 点形状模型

在每一层高斯图像金字塔上建立统计模型，包括初始形状模型和局部轮廓模型。

首先，对于初始形状模型，可通过归一化对齐和主成分分析提取。归一化对齐实质上指的通过平移、旋转和缩放进行最小化对应点坐标之和 $D = \sum |x_i - \bar{x}|$ 的对齐操作。对齐过程^[61]如下：

- (1) 移动每一个人脸样本，使其重心重合于坐标原点；
 - (2) 任选其中一个样本作为平均形状的初始估计，并将其缩放至 $|\bar{x}| = 1$ ，记录其方向 \bar{x}_0 作为参考框架的默认方向；
 - (3) 将训练集中所有样本对齐当前平均形状；
 - (4) 重新计算对齐后的样本训练集的平均形状；
 - (5) 将当前平均形状方向与 \bar{x}_0 对齐，并将其缩放至 $|\bar{x}| = 1$ ；
 - (6) 判断迭代后平均形状是否发生比较大的变化，若没有收敛，返回 (3)。
- 主成分分析实质上是将数据降维，以提高数据运算速度与分析精度过程^[62]：
- (1) 计算平均形状

$$\bar{s} = \frac{1}{n} \sum_{i=1}^n s_i \quad (4.1)$$

- (2) 计算协方差矩阵

$$S = \frac{1}{n-1} \sum_{i=1}^n (s_i - \bar{s})(s_i - \bar{s})^T \quad (4.2)$$

- (3) 计算 S 的特征值和特征向量，并将特征值由大到小排列，取前 t 个组成集合 $\{\lambda_1, \lambda_2, \dots, \lambda_t\}$ ， t 需要满足：

$$\sum_{i=1}^t \lambda_i / \sum \lambda_i \geq \eta \quad (4.3)$$

其中，通常情况下 η 取 95%~98%。

- (4) 取 S 前 t 个特征值所对应的特征向量，记为主成分特征向量 $P = [P_1, P_2, \dots, P_t]$ 。据此建立一个线性可变初始形状模型

$$s = \bar{s} + Pb \quad (4.4)$$

其中，形状变化参数 b 代表每一形状分量在整个模型中的占比，更新 b 可得到

不同的形状，并且为防止模型的过拟合和减少噪声等其他因素的影响，需限制 b 在 $[-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i}]$ 范围内。

然后，关于局部轮廓模型^[62]，它反映了每个关键特征点附近的局部纹理特征。如图 4.4 所示，对于第 i 幅图像样本中的第 j 个关键特征点，在与相邻关键特征点连线的法线方向上，其两边的 k 个像素灰度值按序组成长度为 $2k + 1$ 的灰度向量 g_{ij} ，为减少光照等的影响，对 g_{ij} 求导得到 g_{ij}' ，并进行归一化处理：

$$g_{ij} = g_{ij}' / \sum_{j=1}^{2k+1} g_{ij}' \quad (4.5)$$

类似地，可以得到样本训练集中每一幅图像样本中的每一个关键特征点的局部轮廓模型，通过平均灰度向量 \bar{g}_j 和其协方差矩阵 S_g 表达训练集中样本的第 j 个关键特征点的局部轮廓模型。

$$\bar{g}_j = \frac{1}{n} \sum_{i=1}^n g_{ij} \quad (4.6)$$

$$S_g = \frac{1}{n-1} \sum_{i=1}^n (g_{ij} - \bar{g}_j)(g_{ij} - \bar{g}_j)^T \quad (4.7)$$

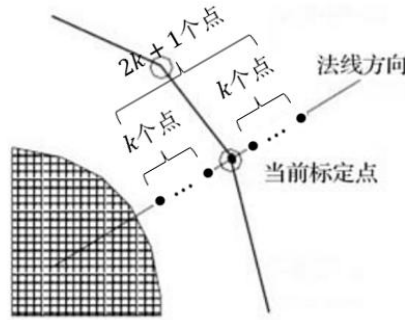


图 4.4 局部灰度模型

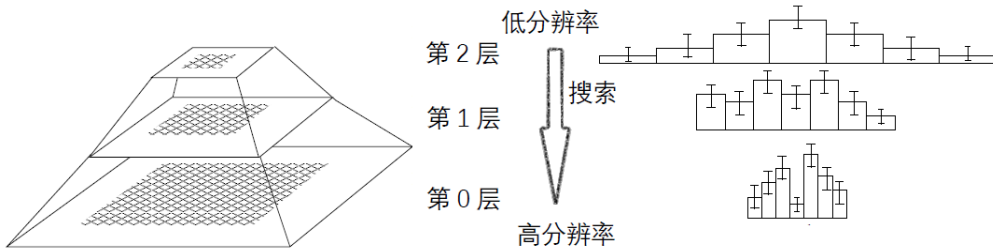


图 4.5 高斯图像金字塔的示意图

在此基础上，在高斯图像金字塔的每一层中建立初始形状模型和局部轮廓模型。它的引入是为了进一步提高算法的准确性和鲁棒性，从低分辨率到高分辨率逐步地开展搜索，在提速的同时也能减少在错误图像结构中的滞留，如图 4.5 所示。高斯图像金字塔构建^[61]的具体步骤是，首先对图像第 L 层进行高斯内核卷积实现平滑，然后将所有的偶数行和偶数列去除进行亚采样，得到第 $L + 1$ 层。某一关键特征点局部轮廓模型的灰度向量中像素数不随图像金字塔层数而改变，但在第 L 层像素尺寸是原始图像（第 0 层）的 2^L 倍，由此看出在分辨率低的图像层中包含有更多地图像信息。

2. 模型搜索与匹配

构建待检测图像的高斯图像金字塔，在之前样本集训练的基础上，在待检测图像的高斯图像金字塔中的每一层中建立初始形状，然后利用局部轮廓模型寻找当前关键特征点的灰度的适合匹配点，接着通过更新形状变化参数 b 修正初始形状，实现在高斯图像金字塔中该层中对应特征点的匹配，搜索是由高斯图像金字塔的低分辨率至高分辨率逐次进行的。

对于关键特征点的局部轮廓模型的搜索与匹配^[61]，可以在当前关键特征点每边取样 $m(m > k)$ 个像素，然后对比在取样过程中每个 $2(m - k) + 1$ 可能位置上对应局部轮廓模型的匹配效果，取样本灰度向量和平均灰度向量的马氏距离 $f(g_s)$ 最小时的关键特征点作为最佳匹配点。

$$f(g_s) = (g_s - \bar{g}_j)^T S_g^{-1} (g_s - \bar{g}_j) \quad (4.8)$$

对于将模型 s 匹配于新的图像特征点集 Y ，在于更新形状变化参数 b ，使 $|Y - sT_{X_t, Y_t, s, \theta}|$ 最小，具体迭代过程如下：

- (1) 初始化 $b = 0$;
- (2) 生成模型 $s = \bar{s} + Pb$;
- (3) 寻找合适的旋转矩阵 $T_{X_t, Y_t, s, \theta}$ ，使 Y 对齐于 s ;
- (4) 求旋转矩阵的逆矩阵，将 Y 投影到模型参考系中 $y = T_{X_t, Y_t, s, \theta}^{-1}(Y)$;
- (5) 通过缩放 $1/(y \cdot \bar{s})$ 倍，将 y 投影到 \bar{s} 的切向平面;
- (6) 为匹配 y ，在约束的前提下，更新形状变化参数 $b = P^T(y - \bar{s})$;
- (7) 迭代后若旋转矩阵与形状变化参数发生了明显变化，则继续转至(2)。

而在高斯图像金字塔的低分辨率层会允许比较大的移动匹配，从而汇聚到高分辨率层，算法的具体步骤如下：

- (1) 设层数 $L = L_{max}$;
- (2) 当 $L \geq 0$ 时
 - (a) 计算图像第 L 层模型中点的位置;
 - (b) 利用局部轮廓模型搜索当前关键特征点每边的 n_s 个像素点;
 - (c) 更新形状变化参数 b ，实现待检测图像中与模型训练集中的对应关键特征点的匹配;
 - (d) 判断寻找到的当前特征点的局部轮廓模型中的合适匹配点数目是否少于 $n_s/2$ 个，或者该层的迭代数已达到最大值，若出现上述情况，继续转至(2a)步;
 - (e) 若 $L > 0$ ，使 $L = L - 1$;

最后的结果由第0层收敛后的参数得到。

4.1.3 堆叠主动性状模型

Stasm 只是在 ASM 的基础上改进了其中的局部轮廓模型。Stasm 在 ASM 的基础上对于局部轮廓模型的具体改进在于^[63]:使用 2 维局部纹理特征而不是 1 维;为更好地实现初始形状的定位,进行两次 ASM 串联堆叠搜索,以第一次搜索的结果作为第二次搜索的初始形状;为节省计算时间简化协方差矩阵。

本文训练了 Stasm 模型,并进行匹配测试。如图 4.6 所示,本文得到的 Stasm 模型可以完美地匹配人脸的各个关键点。

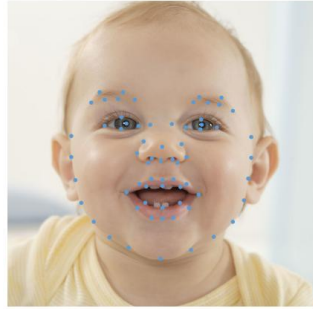


图 4.6 Stasm 匹配结果

4.1.4 面部基础特征提取

在堆叠主动形状模型的基础上,通过该关键点的横纵坐标可以得到面部基础特征。为避免头部晃动使驾驶员距摄像头远近不同造成对面部特征产生干扰,采取眼睛高宽比与嘴巴高宽比作为面部基础特征参数。

根据眼部、嘴部几个特征点的坐标便容易得到眼睛高宽比和嘴部高宽比,如图 4.7 所示,眼部高宽比为 $(y_{28} - y_{30}) / (x_{29} - x_{27})$,因后期嘴部哈欠阈值取正常状态的 1.6 倍,故嘴部高宽比选用外嘴唇的,具体为 $(y_{51} - y_{57}) / (x_{54} - x_{48})$ 。

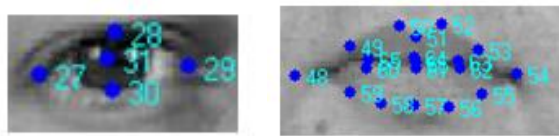


图 4.7 眼嘴高宽比

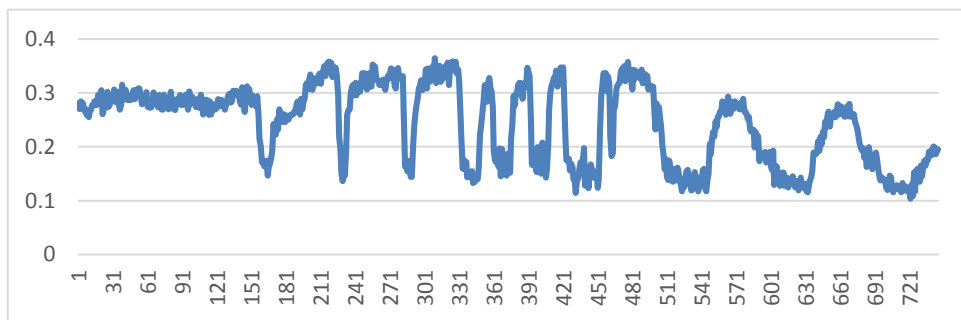


图 4.8 眼睛高宽比

一段眨眼视频的眼睛高宽比数据输出如图 4.8 所示,可以看到三次正常眨眼到三次快速疲劳眨眼到一次抗拒性挤眼再到三次缓慢疲劳眨眼的过程中,眼睛高

宽比有很规律的明显变化。视频中眼睛不同状态截图如图 4.9 所示, 其中 ear 指此时眼睛高宽比, blink 指该视频中此时为止眨眼总数, eclose 指此时连续闭眼帧数。

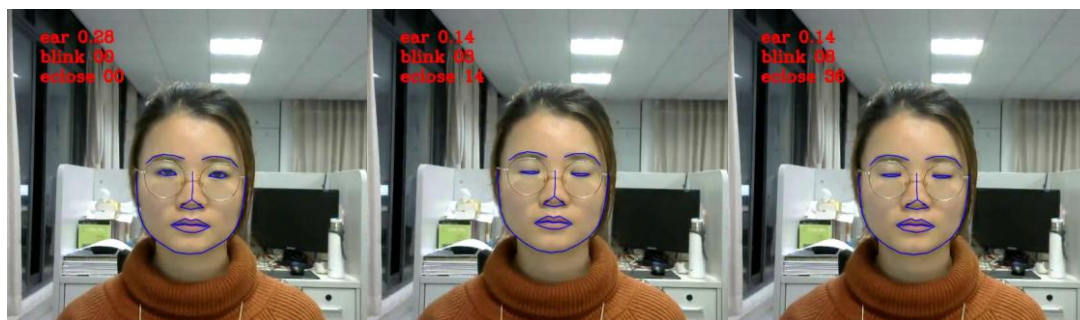


图 4.9 眼睛不同状态

一段嘴部动作视频的嘴部高宽比数据输出如图 4.10 所示, 可以看出这段视频中开始嘴部无动作到开始讲话再到打了三个哈欠的嘴部三种状态下的高宽比有明显的不同。视频中嘴部不同状态截图如图 4.11 所示, 其中 mar 指此时嘴部高宽比, yawn 指该视频中此时为止打的哈欠总数, mopen 指此时连续张嘴帧数。

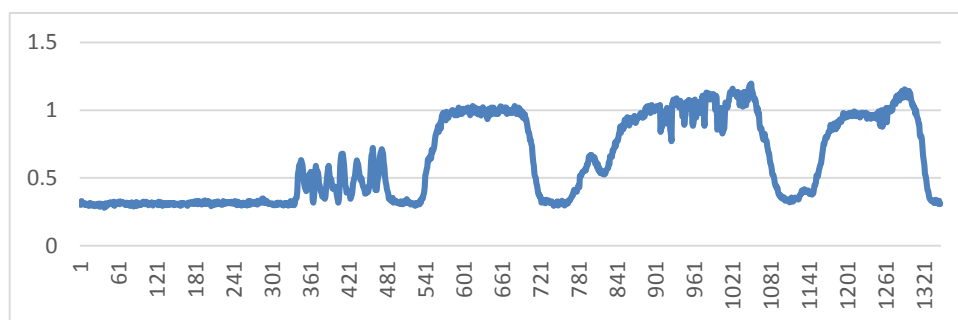


图 4.10 嘴部高宽比

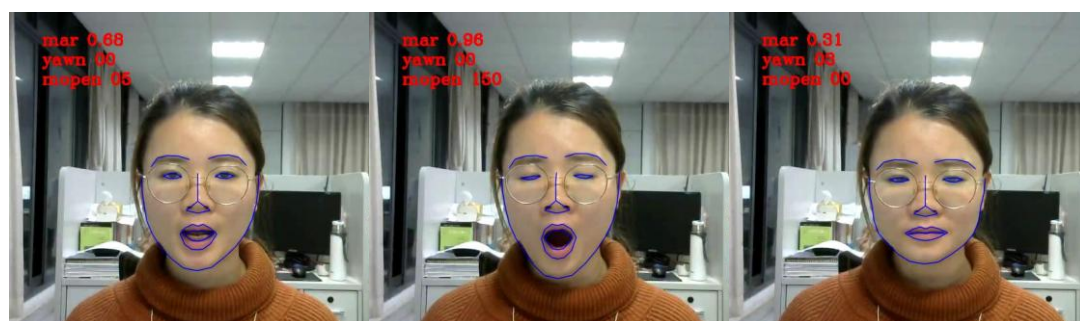


图 4.11 嘴部不同状态

4.2 头部基础特征

除了面部特征外, 疲劳时还会伴有点头的动作, 故可以加入头部状态。在 Stasm 特征形状匹配的基础上, 再结合相机位姿测量方法可以得到点头角度。相机位姿测量方法需先通过相机标定得到相机参数。

4.2.1 相机标定

物体经相机成像, 是通过线性相机模型建立真实物体和成像位置的联系, 其

中由于镜头的生产与安装方面的误差还会存在相机畸变，所以在相机标定之前，我们需要了解一下这两个概念^[64]。

1. 线性相机模型

通过小孔成像的原理，线性相机模型完成世界坐标系 $O_w-X_wY_wZ_w$ 、相机坐标系 $O_c-X_cY_cZ_c$ 、图像坐标系 xO_1y 和像素坐标系 uO_0v 彼此之间的转换，如图 4.12 所示。

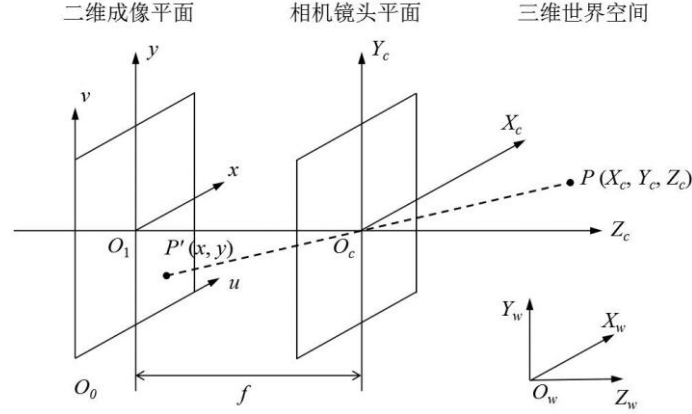


图 4.12 线性相机模型

世界坐标系是用来表达相机安装位置和其他物体位置的一个标准坐标系。相机坐标系以相机光学镜头所在中心为原点，所在平面为 $X_cO_cY_c$ ，光轴为 Z_c 轴，描述了被拍摄物相对于相机的具体位置。图像坐标系是被拍摄物经相机投影通过感光芯片成像的二维平面，与相机镜头平面平行并相距焦距 f ，以投影中心为原点，理想状态下与成像中心重合。像素坐标系表达的是经感光芯片数字化存储像素的二维序列，像素数量代表的是分辨率，像素坐标系与图像坐标系位于同一平面并平行，以图像左上角为原点，与其他坐标系不同在于单位是像素而不是长度。

$$\begin{aligned}
 & \text{像素坐标系2D} \\
 & \quad \text{图像坐标系2D} \\
 & \quad \quad \text{相机坐标系3D} \\
 & \quad \quad \quad \text{世界坐标系3D} \\
 & Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \\
 & = \begin{bmatrix} f/dx & 0 & u_0 & 0 \\ 0 & f/dy & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_2 M_1 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}
 \end{aligned}$$

图 4.13 线性相机模型公式

如图 4.13 所示，对于某一具体位置，三维的世界坐标系经旋转平移得到三维的相机坐标系，经投影变换得到二维的图像坐标系，最后经平移和单位换算得到二维的像素坐标系。值得注意的是从三维转换至二维的过程中，是多对一的转换，

并不是一一对应的关系。

相机标定得到的相机内部参数指公式中矩阵 M_2 提到的参数，包括主点坐标 (u_0, v_0) 、焦距 f 、像素 x 轴单位长度 dx 、像素 y 轴单位长度 dy ；相机外部参数指公式中的矩阵 M_1 ，包含旋转矩阵和平移向量。相机内部参数由相机生产制造过程中的内部结构决定，相机外部参数表达了相机坐标系相对与世界坐标系的具体位置。

2. 相机畸变

实际应用中，由于摄像头加工与安装误差，真实的成像会与理想的相机线性模型产生偏差，这种相对于实际物体成像的变形偏差就是图像畸变，分为径向畸变和切向畸变。

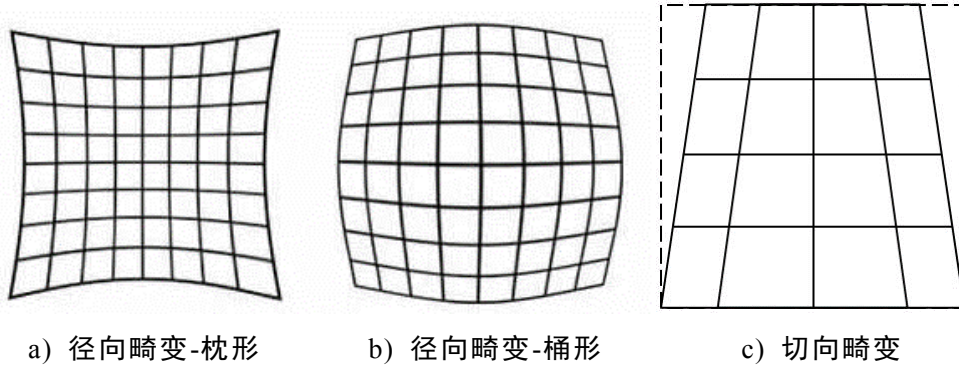


图 4.14 图像畸变示意图

径向畸变是由于镜头加工时曲率不均或者选取的广角镜头曲率太大所造成的，使得成像以投影中心为中心沿径向产生枕形或桶形的变形，如图 4.14(a)(b)所示。切向畸变是由于镜头安装时与感光芯片不平行所造成的，侧投影使得成像产生切向变形，如图 4.14(c)所示。畸变校正公式如式 4.9 所示：

$$\begin{cases} x' = x + \boxed{\text{径向校正}} + \boxed{\text{切向校正}} \\ y' = y + \boxed{\text{径向校正}} + \boxed{\text{切向校正}} \end{cases} \quad (4.9)$$

其中， (x, y) 为校正前的图像坐标， (x', y') 为校正后的图像坐标， $r^2 = x^2 + y^2$ ， k_1 、 k_2 、 k_3 为径向畸变的校正参数， p_1 、 p_2 为切向畸变的校正参数。

3. 张正友相机标定

相机参数和畸变参数只与相机自身内部结构有关，而不受外界使用环境的影响，所以可以由离线标定获得。在各种各样的标定方法中，张正友相机标定法既有传统标定方法的高精度，又有自标定方法对设备的低要求，而且还操作简单，适用于本文相机标定。张正友相机标定首先通过被标定相机以不一样角度拍摄多张同一棋盘格图案，然后自动检测棋盘格图案上的特征点，最后利用单应性矩阵和闭式解算法得到相机内外参，并利用最大似然估计非线性优化包括畸变校正参数在内的全部参数。

作为可视化相机标定 APP，MATLAB R2018a 中的 Camera Calibration 使用张

正友相机标定方法，实用高效，其可视化操作台如图 4.15 所示。具体的标定步骤如图 4.15 所示：

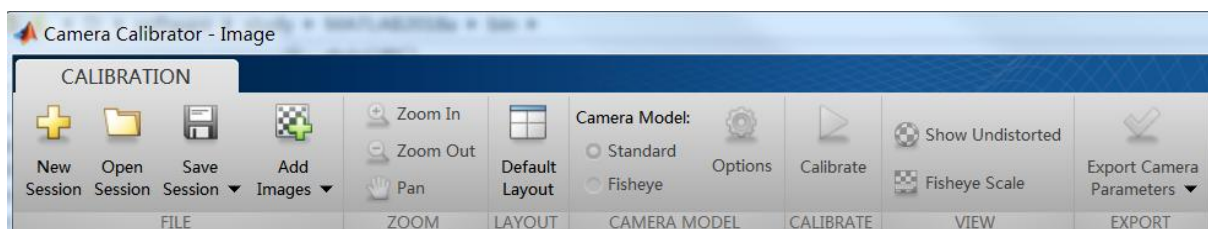


图 4.15 MATLAB 标定 APP 操作台

首先准备好棋盘格图案，采用 7×10 个尺寸为 $59\text{mm} \times 59\text{mm}$ 小方格，如图 4.16 所示，为保证拍摄过程中的平整性，用 A2 纸打印后平贴于绘图木板上，完成标定板的制作：

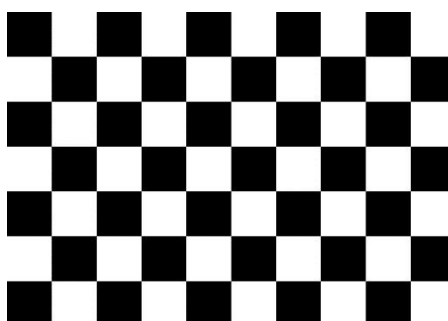


图 4.16 标定棋盘格

然后固定相机，拍摄并保存旋转移动过程中每个位置上的标定板图案。拍摄过程中在保持图案姿势多样化的同时，标定板与相机平面所在平面夹角应小于 45° 。本文标定图像选择如图 4.17 所示的 18 张图片，利用这这图片进行相机标定。



图 4.17 标定图像

最后，通过 MATLAB R2018a 中的 Camera Calibration App 进行相机标定。添加图像后自动检测特征点，如图 4.18(a)所示，软件会筛选掉不合适的图像，然后

再次进行图像的添加，如此反复至 18 张如图 4.17 所示的图像；点击 Calibrate 进行标定，得到标定后重投影的特征点，如图 4.18(b)所示；同时得到每幅图像特征点重投影的误差，可重新对标定图像进行选择后更新标定结果，如图 4.18(c)所示；并且还会计算相机与每次拍摄时标定板的相对位置并分别以相机的视角和标定板的视角显示出来，如图 4.19。

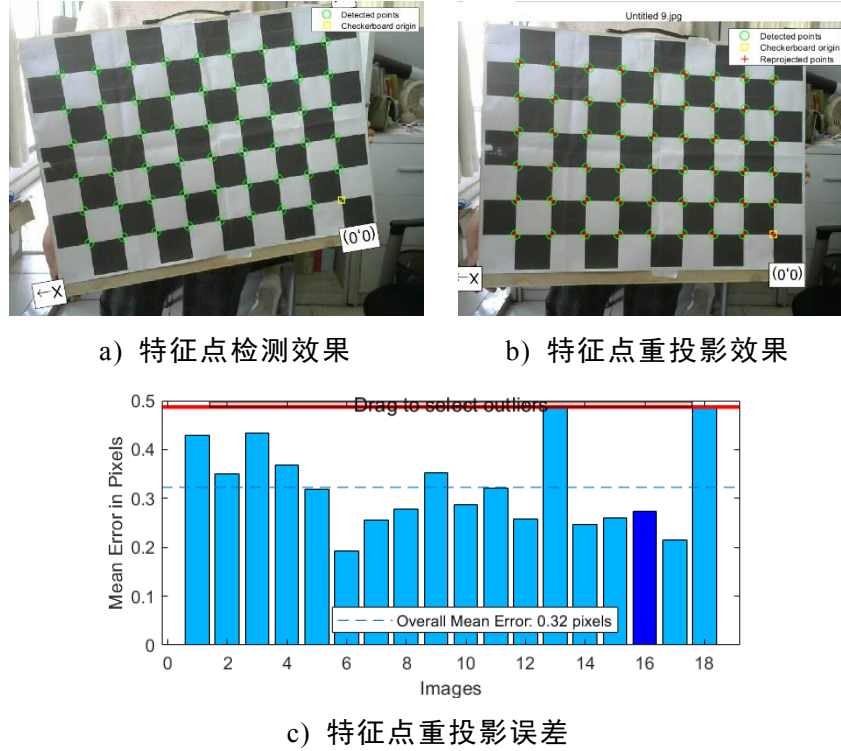


图 4.18 特征点检测与重投影

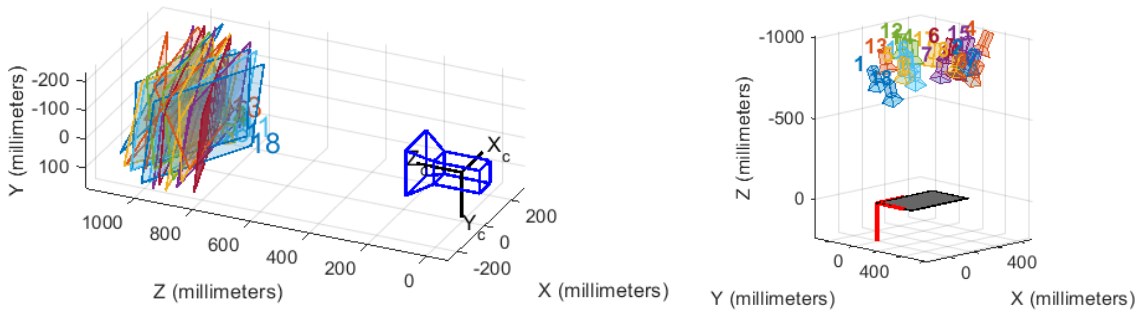


图 4.19 相机与标定图像的相对位置

最终标定得到的相机内参矩阵 M_2 为：

$$M_2 = \begin{bmatrix} 814.4112 & 0 & 316.1399 & 0 \\ 0 & 813.9187 & 248.1261 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.10)$$

其中，相机焦距为(814.4112, 813.9187)，主点坐标为(316.1399, 248.1261)，另外畸变参数 $k_1 = 0.0195$ ， $k_2 = 0.2873$ ， $k_3 = 0$ ， $p_1 = 0$ ， $p_2 = 0$ 。

相机外部参数与世界坐标系的选择有关，描述了相对于相机坐标系的位置关系，包括旋转矩阵和平移向量。MATLAB R2018a 中的相机标定对每一副标定图像都输出了对应的旋转矩阵和平移向量，都是选取右下方的第一个标记角点为原

点，垂直图像的方向为 Z 轴。故选取其中如图 4.20 所示的一张图像，输出此世界坐标系下的相机外部参数：

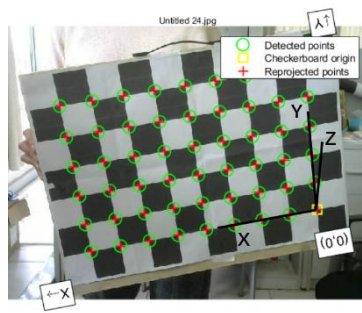


图 4.20 某棋盘格的世界坐标系

$$R = \begin{bmatrix} -0.9815 & 0.1727 & -0.0826 \\ -0.1485 & -0.9591 & -0.2410 \\ -0.1208 & -0.2243 & 0.9670 \end{bmatrix} \quad (4.11)$$

$$T = [276.3144 \quad 88.1869 \quad 970.5299] \quad (4.12)$$

4.2.2 相机位姿测量

相机位姿测量是由有确定相对位置的几个点，并在某坐标系中找到他们的具体位置，与其在相机照片中的成像，并结合相机参数，求解出相机位于该坐标系内的旋转角度与平移量，其关键就在于对 PNP 问题的求解。而关于 PNP 问题就是指通过世界中的 N 个特征点与图像成像中的 N 个像素点，计算出其投影关系，从而获得相机或物体位姿的问题。其实相机位姿测量的实质就是求取人脸这个坐标系下相机的外部参数，通过人脸坐标系与相机坐标系的旋转矩阵经欧拉角转换计算得到头部姿态。

本文用欧拉角来表达头部姿态，如图 4.21 所示，3 个欧拉角分别沿空间 3 条轴旋转，其中俯仰角 *pitch* 描述了头部俯仰角度，偏航角 *yaw* 描述了头部左右偏转角度，滚动角 *roll* 描述了头部摆动幅度。

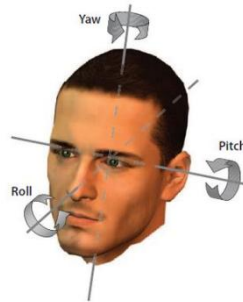


图 4.21 头部姿态欧拉角

相机位姿测量流程如下：

首先准备好函数 `solvePnP` 的原型中所需的相机内部参数与畸变参数，上文通者 Matlab 的相机标定工具箱得到了相关相机标定参数：相机焦距 (814.4112, 813.9187)，主点坐标 (316.1399, 248.1261)，畸变参数 $k_1 = 0.0195$, $k_2 =$

0.2873, $k_3 = 0$, $p_1 = 0$, $p_2 = 0$, 将其依次输入模型中。

然后确定鼻尖、两眼角与两嘴角等五个特征点的相对位置。如图 4.22 所示, 以鼻尖为原点, 面部朝向为 Z 轴, 以面部左右对称往上为 Y 轴, 建立世界坐标系。以本人头部模型为三维标准模型, 量得三维标准模型各点的相对位置, 其中鼻尖坐标为(0,0,0), 左眼角的坐标为(-475,350,-370), 右眼角的坐标为(475,350,-370), 左嘴角的坐标为(-240,-310,-240), 右嘴角的坐标为(240,-310,-240), 下巴的坐标为(0,-660,-210)。

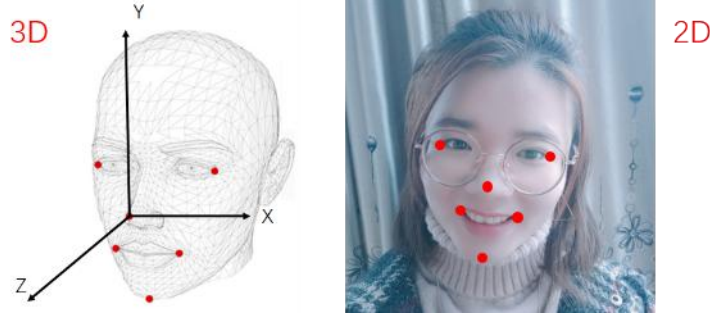


图 4.22 头部三维模型与二维图像

接着输入该 6 个特征点对应的二维图像坐标, 这是由上节中 Stams 匹配所得到的关键点获得, 鼻尖、左右眼角、左右嘴角和下巴分别对应 Stams 形状中的点 67、点 27、点 32、点 48、点 54、点 7。

最后调用 OpenCV 中的 solvePnP 函数, 输出旋转矩阵并由此计算欧拉角, 获得头部姿态。旋转矩阵 R 与 3 个欧拉角的关系^[65]就是:

$$R = R_x R_y R_z \quad (4.13)$$

$$\text{其中 } R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}, \quad R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(yaw) & -\sin(yaw) \\ 0 & \sin(yaw) & \cos(yaw) \end{bmatrix},$$

$$R_y = \begin{bmatrix} \cos(pitch) & 0 & \sin(pitch) \\ 0 & 1 & 0 \\ -\sin(pitch) & 0 & \cos(pitch) \end{bmatrix}, \quad R_z = \begin{bmatrix} \cos(roll) & -\sin(roll) & 0 \\ \sin(roll) & \cos(roll) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

所以, 3 个欧拉角可通过以下计算得到:

$$yaw = \text{atan2}(R_{32}, R_{33}) \quad (4.14)$$

$$pitch = \text{atan2}(-R_{31}, \sqrt{R_{32}^2 + R_{33}^2}) \quad (4.15)$$

$$roll = \text{atan2}(R_{21}, R_{11}) \quad (4.16)$$

其中, $pitch$ 便是我们想要的头部角度, 另外 yaw 可用于分心检测。

4.2.3 头部基础特征提取

本文中, 头部角度则可以通过计算随相机固定而不动的相机坐标系与随头部

转动而转动的世界坐标系的角度而得到。根据相机位姿测量方法，一段头部姿态视频点头角度输出数据如图 4.23 所示，为清楚地看到点头的幅度，此处输出的点头角度是初始化后的实际测得的头部角度与初始化头部角度的差值，可以看到两次仰头和三次疲劳点头的点头角度有明显的不同。视频中头部不同姿态截图如图 4.24 所示，其中 *pitch* 指此时测得的头部角度，*pitchstd* 指初始化窗口的头部角度均值，并以此为标准与 *pitch* 相减取绝对值作为点头角度 *pitchdiff*，*pitchtimes* 指该视频中此时为止点头总数，*pitchcon* 指此时连续点头帧数。

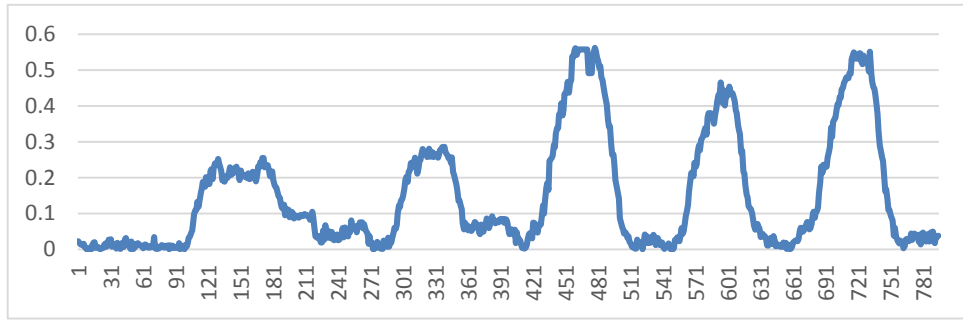


图 4.23 点头角度



图 4.24 头部不同姿态

4.3 特征状态参数初始化及阈值

因清醒状态下各特征的个体性差异，本文考虑了针对待检测者进行面部特征与头部角度的初始化。假定驾驶员在驾驶初始阶段属于清醒状态，取前 100 帧的面部特征与头部高度的中位数作为该驾驶员清醒状态下的特征，再结合事先在经验知识的基础上修正的每个人清醒与疲劳状态下各特征的关系，得到该驾驶员疲劳状态下特征的阈值，从而实时获取检测阶段的特征进而判断该驾驶员是否进入疲劳状态。

眼睛各种状态类别的不同状态参数采用 P80 的标准，如图 4.25 所示，取眼睛闭合超过 80% 认为闭合，小于 20% 认为完全睁开。打哈欠阈值是嘴部开度达到平时开度的 1.6 倍，且持续时间超过 3s^[66]。疲劳驾驶的点头阈值取在正常头部活动范围的 30%，即 0.34^[65]。

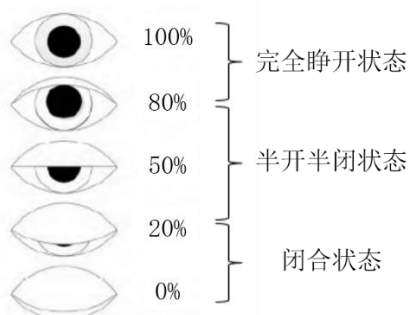


图 4.25 眼睛睁开程度示意图

4.4 具体疲劳特征计算

现实生活中，在人处于疲劳状态时，除了疲劳闭眼这一主要特征外，往往还伴随着打哈欠、点头。故以眼部状态为主，以嘴部与头部状态为辅的疲劳特征能大大增加算法的准确性与鲁棒性。本文通过 PERCLOS、眨眼频率、哈欠频率、点头频率等疲劳特征确定驾驶员的疲劳程度。

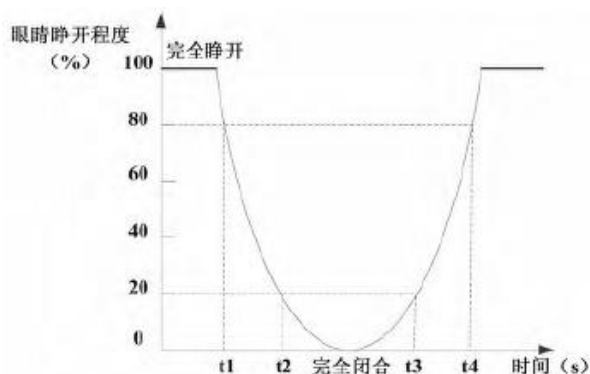


图 4.26 PERCLOS 计算原理示意

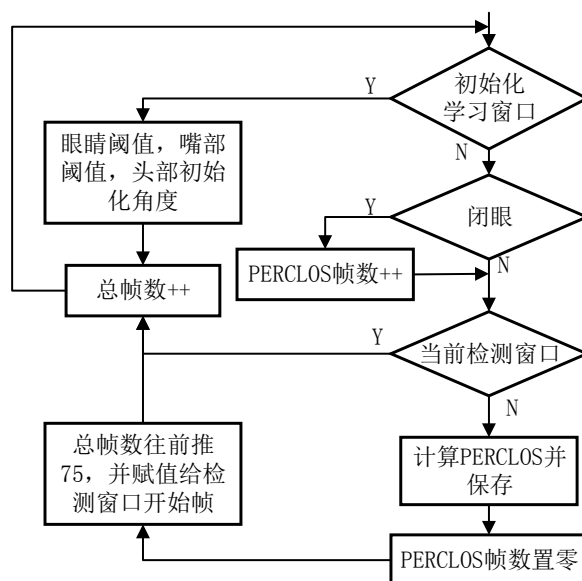


图 4.27 提取 PERCLOS 框图

美国公路交通安全局 (NHTSA) 指出在九种疲劳检测指标中, PERCLOS 最能反映驾驶员的疲劳程度^[67]。如图 4.26 所示, PERCLOS 指眼睛闭到某个程度的

时间占比,采用 P80 标准,故可根据定义,计算 PERCLOS 的 f_e 值 $f_e = [(t_3 - t_2)/(t_4 - t_1)] \times 100\%$ 。本文取 300 帧作为一个时间窗口^[68],其中有 K_e 帧的眼睛闭合程度大于 80%,其 PERCLOS 的 f_e 值 $f_e = (K_e/300) \times 100\%$,提取框图如图 4.27 所示。

眨眼频率也能反映疲劳程度。清醒状态下的眨眼频率为 3-4s/次,而轻度疲劳状态下伴随着较快的眨眼频率,极度疲劳状态下伴随着较慢的眨眼频率。本文定义当眼睛闭合且闭合时间在 3-60 帧之间为眨眼。本文取一个时间窗口 300 帧内的眨眼次数,进而计算眨眼频率,其提取框图如图 4.28 所示。

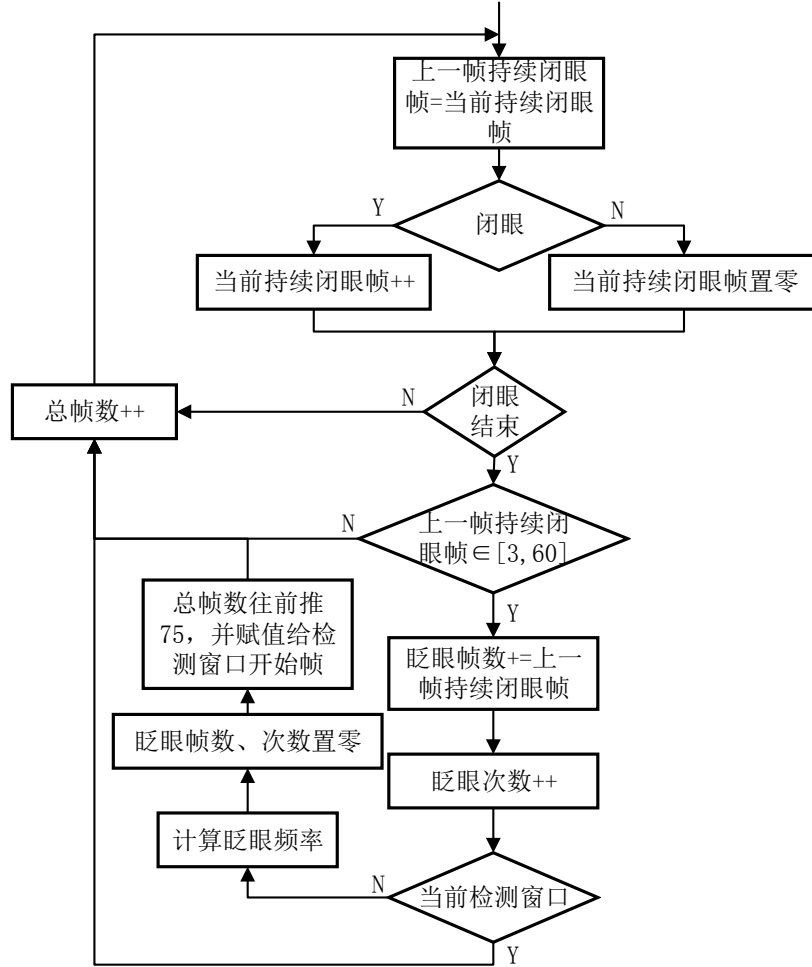


图 4.28 提取眨眼频率框图

哈欠是人体疲劳时表现出来的一个重要特征,当打哈欠时,嘴部明显持续张大。采用类似 PERCLOS 的思想,设计 PMRCLOS 变量,根据 $f_m = (K_m/N_m) \times 100\%$ 计算 PMRCLOS 的 f_m 值,其中一个时间窗口内总帧数 N_m 中嘴部开度超过阈值的总帧数为 K_m 。统计 300 帧内驾驶员被检测为打哈欠的帧数与总帧数之比为哈欠频率,其提取框图如图 4.29 所示。

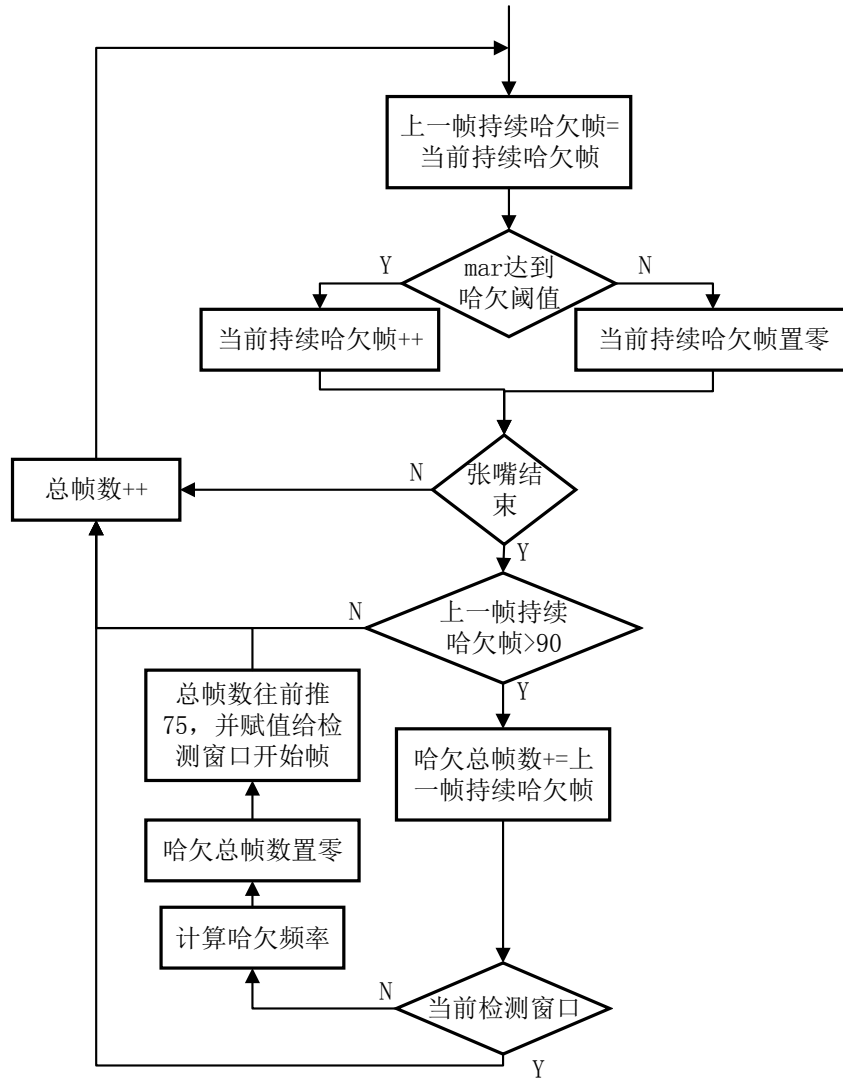


图 4.29 提取哈欠频率

疲劳状态下也会伴随着频繁点头的情况，统计 300 帧内驾驶员点头帧数与总帧数之比为点头频率，其提取框图如图 4.30 所示。

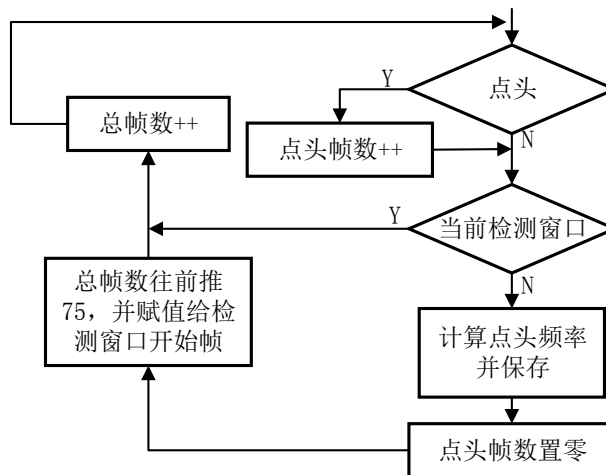


图 4.30 提取点头帧数框图

一段驾驶员视频具体特征数据输出如图 4.31 所示，驾驶员状态由包含说话的清醒状态（0-4）到伴随哈欠的疲劳状态（5-10）再到伴随点头的非常疲劳状态（11-13）。可定性观察到驾驶员 3 种疲劳状态下的 4 个具体疲劳特征有较明显的可分类性。

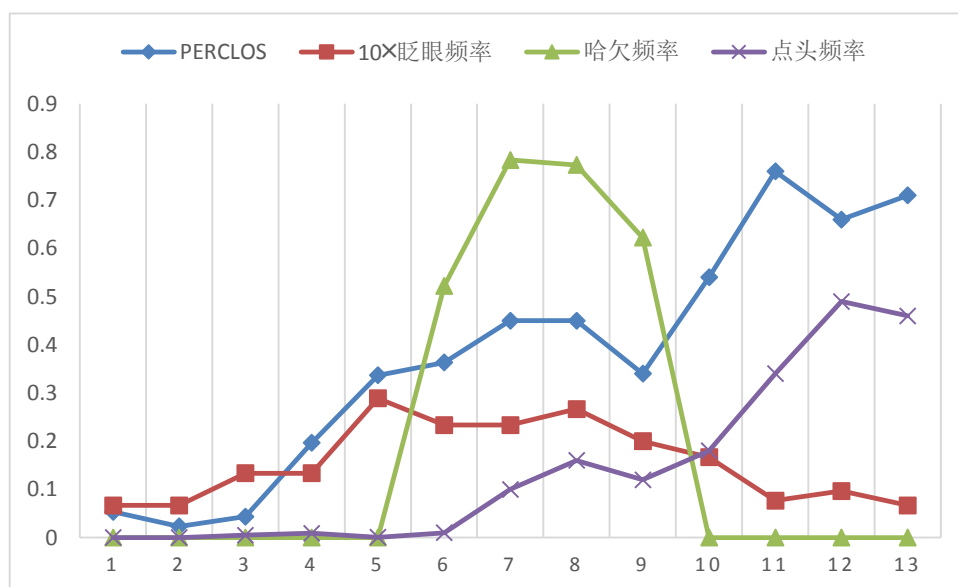


图 4.31 驾驶员视频具体特征输出

4.5 本章小结

本章主要内容是提取基础特征并计算具体疲劳特征。首先简要概述了各种特征提取方法，详细介绍了主动形状模型，并选择在 ASM 基础上改善的 Stasm，实现人脸关键点的匹配，达到了吻合良好的效果；然后利用 Stasm 获得眼睛高宽比、嘴部高宽比等面部基础特征，并结合相机位姿测量方法获得点头角度等头部基础特征；再对某位特定驾驶员进行基础特征的初始化与与阈值计算；最后在基础特征的基础上计算得到 PERCLOS、眨眼频率、哈欠频率、点头频率等具体疲劳特征。

第 5 章 实验分类与移植测试

前述章节完成了视频的具体疲劳特征的提取计算，本章介绍要实现整个疲劳检测剩下的其他步骤，即实验分类和移植测试。首先搭建模拟驾驶平台，采集驾驶实验视频并对其进行后期处理，通过上述章节的算法计算具体疲劳特征，构成带有标签的样本集；然后输入到搭建的极限学习机中进行训练及模型测试，将训练好的分类模型封装好，与上述章节算法程序构成统一的整体，一起移植到嵌入式开发板中，实现车载驾驶员疲劳检测；最后进行测试以验证系统效果及准确率。

5.1 疲劳驾驶实验

考虑到疲劳驾驶的危害性，我们通过模拟驾驶进行疲劳驾驶实验。为采集实验数据，设计了基于模拟驾驶仪和摄像头实验平台的模拟驾驶实验，并对实验视频进行整合与处理，得到 7 组带有标签的具体疲劳特征数据，为后续分类器的训练做准备。

5.1.1 实验平台搭建

首先，我们通过型号为 PR-L-3-42-AW-AT-MT-200 的模拟驾驶仪和 Logitech®C270 型网络摄像头搭建实验平台。

如图 5.1 所示，该模拟驾驶仪由富朗巴软件科技有限公司生产，具备汽车所有驾驶操作，有手动挡和自动挡。为减少驾驶难度和让驾驶员更快地进入疲劳状态，本实验选自动挡，故志愿者实验时只用简单操控方向盘、油门及刹车。为模拟真实的驾驶场景，采用相邻成 160° 的三块屏幕作为模拟驾驶的驾驶员视野。



图 5.1 模拟驾驶仪



图 5.2 UC-winRoad 鸟瞰地图

模拟驾驶仪中的虚拟环境用 UC-winRoad10.1.1 软件进行建模搭建，采用普通城市道路，包含直道，弯道，单行道、双向 4 车道(中间设置隔离栅栏)；有随机红绿灯，需遵守红绿灯，为让驾驶员更容易进入驾驶疲劳状态未设定车流。其鸟瞰地图如图 5.2 所示，其驾驶员视野如图 5.3 所示。



图 5.3 UC-winRoad 驾驶员视野



图 5.4 Logitech®C270 型网络摄像头



图 5.5 实验平台搭建示意图

采集驾驶员人脸信息的相机选择 Logitech®C270 型网络摄像头，如图 5.4 所示，其具体参数如表 5.1 所示。从表 5.1 可以看到，该摄像头拥有 360p、480p、720p 等 3 种拍摄分辨率，拍摄分辨率过小的人脸图像会存在模糊，造成细节信息的丢失，而拍摄分辨率过大的人脸图像数据大，让计算过程更加复杂，降低算法的它运行速度。考虑到这两方面，本文选取 480p 的拍摄分辨率，也就是 640×480 像素，在满足一定的算法实时性要求下最多地捕捉驾驶员人脸图像信息，其拍摄帧率设定为 30fps。

表 5.1 Logitech®C270 型网络摄像头参数

参数	规格
感光芯片类型	CCD
USB 类型	高速 USB2.0
聚焦类型	定焦
焦距范围	4.0 毫米
视场（FOV）	60°
光学分辨率（真实）	1280×960 1.2MP
图像拍摄（16:9 W）	360p、480p、720p
视频拍摄（16:9 W）	360p、480p、720p
帧速率（最大值）	30fps@640×480

如图 5.5 所示，为寻找拍摄驾驶员人脸区域最好的位置，将摄像头垫高固定于模拟驾驶仪仪表盘正中央。

5.1.2 实验设计与流程

每个志愿者进行两阶段实验，第一阶段处于清醒状态，第二阶段处于疲劳状态。清醒阶段实验要求：志愿者在实验前晚必须保证睡眠充分，前一天不能饮酒，实验在上午 10:00 进行。疲劳阶段实验要求：志愿者在实验前晚的睡眠限制在 4 小时以内，建议 03:00-07:00，实验当天上午进行正常的工作学习，实验在午饭后约 13:00 进行。为防止连续实验给驾驶者带来身体与精神上的伤害，同时也为避免两阶段实验相互影响，同一志愿者的两组实验相隔 5 天进行。招募 8 名志愿者，实验安排如表 5.2 所示：

表 5.2 志愿者实验的时间安排

时间	第 1 天	第 2 天	第 3 天	第 4 天	第 5 天	第 6 天	第 7 天	第 8 天
上午（清醒）	1 号	3 号	5 号	7 号	2 号	4 号	6 号	8 号
下午（疲劳）	2 号	4 号	6 号	8 号	1 号	3 号	5 号	7 号

志愿者到达实验现场后，先填写个人信息，包括年龄，性别，驾龄等。在实验开始前，每位驾驶志愿者对自己的精神状态进行自评，清醒状态时可以进行清醒阶段实验，疲劳状态时可以进行疲劳阶段实验。学习使用并适应模拟驾驶仪 10 分钟，正式清醒状态实验 30 分钟，正式疲劳状态实验 1 小时。

5.1.3 数据处理

采集 7 名志愿者的驾驶视频，利用文献[69]上的标准，如表 5.3 所示，通过 Corel Video Studio X9 进行视频分割与合并，如图 5.6 所示，将其分为清醒视频，疲劳视频和非常疲劳视频，为后续算法共提供 7 组共 21 段视频，如图 5.7 所示。

表 5.3 疲劳判断标准

精神状态	分数	状态描述
清醒	0	眼睛正常睁开，眨眼果断，眼球活动踊跃，注意力集中，头部端正
疲劳	1	眼睛想要闭合，眼球活跃程度下降，目光呆滞，有哈欠，下意识点头，为抵制疲劳而挤眼摇头
非常疲劳	2	眼睛闭合趋势严重，甚至闭眼持续长，出现头部歪斜、点头



图 5.6 整合视频示意图

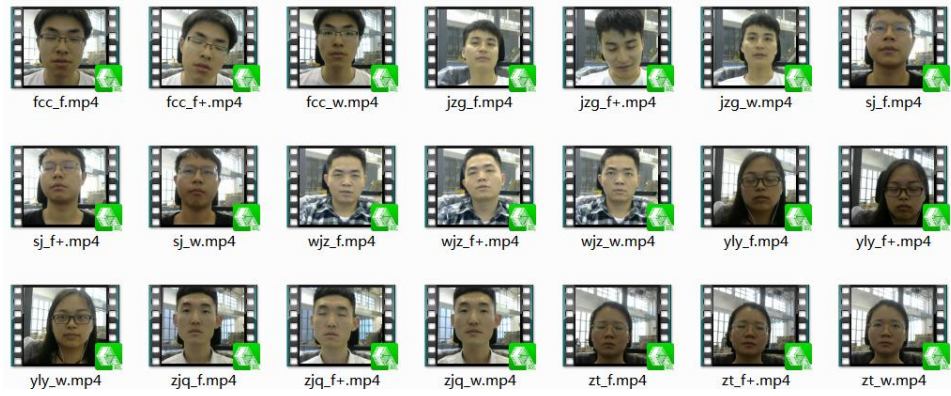


图 5.7 所有整合的视频

5.2 疲劳判断分类

以处理好的志愿者驾驶视频输出得到的具体疲劳特征结合其对应的标签作为样本集，采用 k 折交叉验证法，对疲劳判断分类器进行训练及测试。本文疲劳判断分类器选用快速准确的极限学习机 ELM。

5.2.1 极限学习机算法

神经网络大多通过梯度向下的方法，虽然基本能达到不错的效果，但是其迭代运算量大而耗时长；并且缺少全局意识，容易陷入一种局部最优。针对此问题，南洋理工大学的黄广斌教授提出了 ELM。

相对于 BP 算法和 SVM，ELM 有其自己的优势：需手工调整的参数只有隐层神经元数，用起来更方便；处理分类速度更快；拥有更好的模型泛化能力；激活函数更丰富，更复杂，可包括非线性函数、不连续函数、微分函数、非微分函数等。

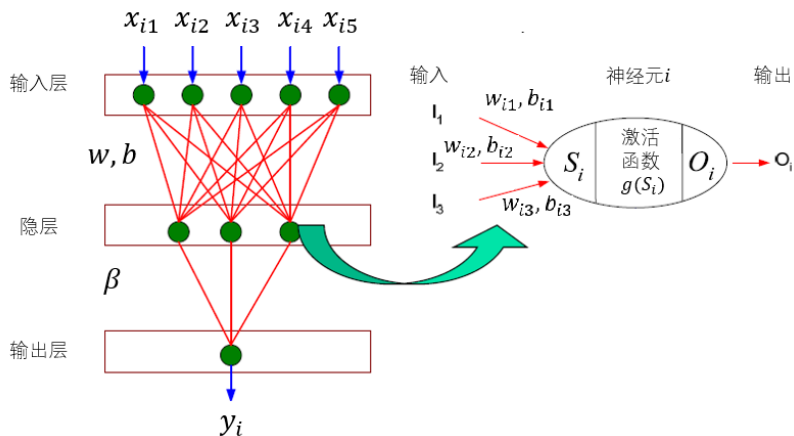


图 5.8 ELM 原理

ELM 是基于单隐层前馈神经网络 SLFNs (Single Hidden Layer Feed Forward Networks) 的新型神经网络学习算法。不同于传统神经网络算法，SLFNs 输入层

到隐层与隐层到输出层的权值无需迭代求解，SLFNs 随机给定输入层到隐层的权值 w 与偏移量 b ，利用最小二乘法求解隐层到输出层的权值 β ，一次性完成整个学习过程，如图 5.8 所示。

设有 N 个不同的带有标签的训练样本 (x_i, r_i) ，其中 $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$ ， x_{in} 指第 i 个样本输入的第 n 个判断特征； $r_i = [r_{i1}, r_{i2}, \dots, r_{im}]^T$ ， r_{im} 指第 i 个样本输出的第 m 类结果。

带 L 个隐层节点的 SLFNs 的理论输出：

$$y_i = \sum_{j=1}^L \beta_j G(\omega_j \cdot x_i + b_j) \quad (5.1)$$

式中训练样本序列数 $i = 1, 2, \dots, N$ ，隐层激活函数 G 形式有 Sigmoid、tanh、ReLU、Leaky ReLU、ELU 等可供选择^[71]，隐层第 j 个节点的输入权值 ω_j 和偏移量 b_j 随机给定，隐层第 j 个节点的输出权值 β_j 为待求值。

训练过程中实际输出应尽可能逼近理论输出，使 $\sum_{i=1}^N \|y_i - r_i\| = 0$ ，并改写成矩阵形式：

$$Y = H\beta = R \quad (5.2)$$

式中 $Y = [y_1, y_2, \dots, y_N]$ ， $R = [r_1, r_2, \dots, r_N]$ ， $\beta = [\beta_1, \beta_2, \dots, \beta_L]$ ，隐层输出矩阵：

$$H = \begin{bmatrix} G(\omega_1 \cdot x_1 + b_1) & \cdots & G(\omega_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ G(\omega_1 \cdot x_N + b_1) & \cdots & G(\omega_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \quad (5.3)$$

故 SLFNs 的训练转化成通过最小二乘法求解 Moore-Penrose 广义逆 H^+ 进而得到隐层的输出权值矩阵 β ：

$$\beta = H^+ R = (H^T H)^{-1} H^T R \quad (5.4)$$

5.2.2 极限学习机分类应用

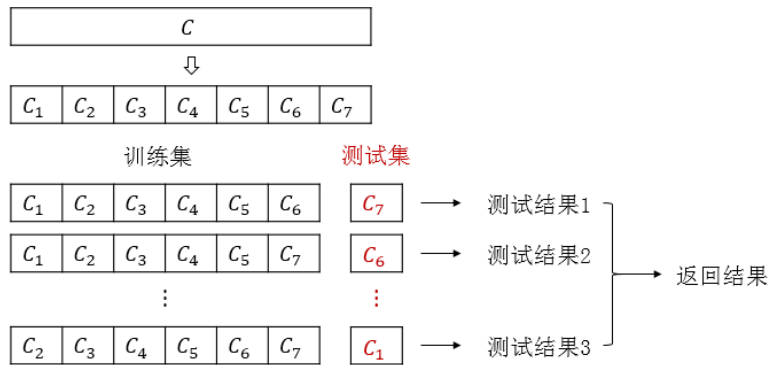


图 5.9 7 折交叉验证示意图

每名志愿者共 180 组带标签样本，包含清醒驾驶、疲劳驾驶、非常疲劳驾驶各 60 组，故 7 名志愿者总共 1260 组样本。考虑到样本数量不是太多，对这 7 名志愿者样本采取交叉验证法^[72]进行训练，每次取 6 名志愿者样本作为训练集，剩

下 1 名志愿者样本作为测试集，这样就得到了 7 套训练/测试集，从而进行 7 次训练和测试，最终用这 7 套所有的训练结果来共同评价该训练模型的性能。其 7 折交叉验证示意图如图 5.9 所示。最后将在 7 次训练得到的 7 个训练模型当中选取效果最好的作为应用的最终分类模型。

具体的训练与测试是利用 MATLAB 的 `elmtrain` 函数与 `elpredict` 函数，其中隐层神经元取 800 个，激活函数取 Sigmoid 函数，应用类型取代表分类的 1。

表 5.4 ELM 分类准确率

实际驾驶 状态	ELM 分类状态			测试集样 本数	分类准确率
	清醒	疲劳	非常疲劳		
清醒	375	36	9	420	89.29%
疲劳	23	366	31	420	87.14%
非常疲劳	0	52	368	420	87.62%
总计	准确分类: 1109			1260	88.02%

如表 5.4 所示，疲劳判断 ELM 分类模型整体性能的分类准确率可达到 88.02%，值得注意的是，在该模型中驾驶员非常疲劳状态不会误判为清醒状态。

5.3 算法移植

首先在 PC 端搭建好本文疲劳检测算法，然后移植到嵌入式板子上，在嵌入式平台中实现疲劳检测。

5.3.1 软硬件环境

首先 PC 端实现本文疲劳算法的搭建，包括疲劳特征的提取、分类器的训练等复杂流程。PC 端硬件配置如表 5.5 所示。软件环境选择 Visual Studio 2012 编译器下的 OpenCV 库和 Matlab 共同编程，编写语言选择 C++，便于移植到嵌入式平台。OpenCV 库是跨平台开源视觉库，广泛应用于图像处理。在 Matlab 里搭建好超限学习机，并封装成函数以被 OpenCV 调用，实现算法在不同软件中的切换，形成一个整体，实现疲劳检测。PC 端测试平台如图 5.10 所示。

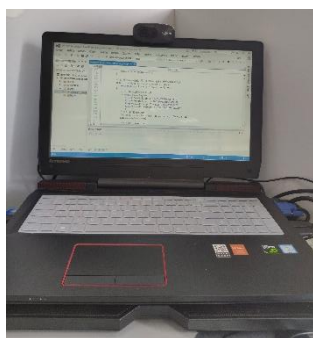


图 5.10 PC 端平台



图 5.11 嵌入式平台 Jetson TX2

表 5.5 PC 端配置参数

配置	参数
CPU	
处理器型号	Intel 酷睿 i5 6300HQ
处理器主频	2.3GHz
核心数	四核
内存	海力士 DDR4 2133MHz 8G
GPU	
显卡芯片	NVIDIA GeForce GTX 960M
显存	4GB

车载智能辅助驾驶系统的应用推广除了依赖于各种人工智能技术的发展外，还必须依赖于嵌入式处理器强大的图像处理和识别等运算能力，技术都要依托于嵌入式得以施展。现在具备生产智能驾驶功能处理器能力的公司只有一小部分，国内较有名气的有地平线、寒武纪等，国外则有 NVIDIA、mobileye 等。国内地平线公司主推征程处理器。国外恩智浦推出的 S32v234 处理器功能十分强大：CPU 有 4 颗 ARM Cortex A53，集成有 ISP 图像处理模块和 APEX 图像加速模块。但是 S32v234 处理器局限在于无法使用通用的摄像头，其所专用的 OV 摄像头限制了它的推广应用。

作为 NVIDIA 的 TX 系列的第二代处理器，嵌入式平台 Jetson TX2^[73]的功能强大，相当于一台人工智能超级计算机，如图 5.11 所示。Jetson TX2 的硬件配置参数如表 5.6 所示：

表 5.6 Jetson TX2 的硬件配置参数

配置	参数
CPU	双核 64 位 NVIDIA Denver 2、四核 ARM A57
GPU	NVIDIA Pascal 架构的 256 核 GPU
Camera	1.4Gpix/s
Memory	8G 128 bit LPDDR4

Jetson TX2 作为速度超级快、超节能高效的嵌入式 AI 计算机设备，单模块功耗仅 7.5 瓦，基于 NVIDIA Pascal™ GPU 架构，搭载 8GB 内存，内存带宽达 59.7Gb/s。借助 4K×2K 60Hz 视频编解码性能和 1400MPix/s 的摄像头接口性能，其成为深度学习、计算机视觉和 GPU 计算的优秀系统。TX 系列具有 NVIDIA 4-Plus-1™和 4 个 ARM® Cortex™-A57 的 CPU，在此基础上，Jetson TX2 采用基于 NVIDIA Pascal™架构的 256 核 GPU，性能强大；Jetson TX2 拥有 Jetson TX1 所有模块的功能，功能完善；Jetson TX2 体积小的同时效率高，十分适合智能机器人、无人机、便携医疗设备和智能汽车等终端设备。此外，Jetson TX2 使

用的是 Ubuntu 系统，方便开发人员直接进行嵌入式产品的开发。因此，为方便后期 ADAS 集成系统也能在同一块嵌入式板子上运行，本文的疲劳检测移植的硬件平台选择 NVIDIA Jetson TX2。

该嵌入式平台的软件环境以 Jetson Development Pack 3.0 SDK 工具包为主要软件包，包含最全面的人工智能计算方法。JetPack 作为一个一体化的软件包，按需求来提供相关解决方案，对该平台全部开发应用软件工具进行打包安装，其中的软件包括 TensorRT, cuDNN, CUDA 工具包, VisionWorks, Streamer 和 OpenCV。

5.3.2 移植算法

驾驶员疲劳检测系统作为汽车高级辅助驾驶系统重要的一部分，故其需要实现车载。我们的 PC 端过于厚重，过于昂贵，适用不了车载。而嵌入式硬件平台小巧的同时性能也有很大的提升，具备驱动应用编程接口 API，拥有良好的可移植性，嵌入式平台比 PC 机更能满足汽车 ADAS 的整体要求。因此，本文把前期在 PC 端研究开发的算法移植到 Jetson TX2 嵌入式平台中，实现车载的疲劳检测。

整个算法的移植包括 OpenCV 视觉库的移植和程序的移植。

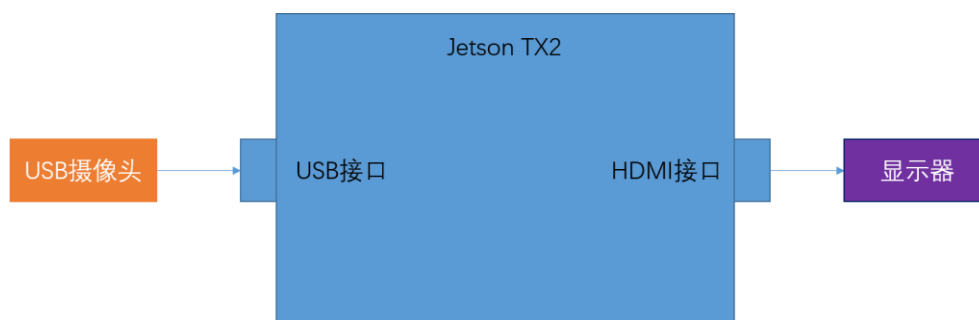


图 5.12 嵌入式平台数据流



图 5.13 嵌入式平台检测效果

OpenCV 视觉库主要用于图像处理和视频流的输入输出。JetPack 打包了 OpenCV 的嵌入式平台库。因为 OpenCV 开源，并且编写语言有一部分是高效的 C 语言，适当处理的话不用额外的外部支持也可自行完整地编译链接生成执行程序，所以 OpenCV 具有高效的移植性，广泛应用于嵌入式平台中。而 Jetson TX2 具备功能强大的硬件资源，能够直接编译 OpenCV，详细流程是：为保证算法的一致性，先在 OpenCV 官网上找到与 PC 端相同版本的 OpenCV 源码及相关模块并下载下来，对其依赖库进行安装；接着修改文件

CMakeList.txt 后通过 Cmake 配置 OpenCV；最后执行 make 编译后，执行 make install 对编译完成的 OpenCV 动态库进行安装。

在 Jetson TX2 配置好 OpenCV 库后，可直接经移动 U 盘将包含有 ELM 训练得到的权重参数的算法程序拷至已搭建好软件开发环境的 Jetson TX2 中，增加视频帧的缩放率以提快在嵌入式平台的图片处理速率，修改应用程序以直接调用车载摄像头。车载疲劳检测系统通过车载摄像头进行视频采集，输入到 Jetson TX2 嵌入式平台上实现疲劳检测，其数据流如图 5.12 所示，其达到的嵌入式平台移植结果如图 5.13 所示。

5.4 测试与验证

在正常光、曝光、弱光等不同光照条件下进行测试，验证算法的可靠性。

每种光照下，测试样本选取 3 段已知的不同驾驶状态的视频，经本文算法程序运行，输出具体特征值及状态判定结果，运行过程中 3 种状态检测效果如图 5.14、图 5.15、图 5.16 所示，经统计数据组数得到算法检测准确率如表 5.7、表 5.8、表 5.9 所示。



图 5.14 正常光下驾驶状态检测效果

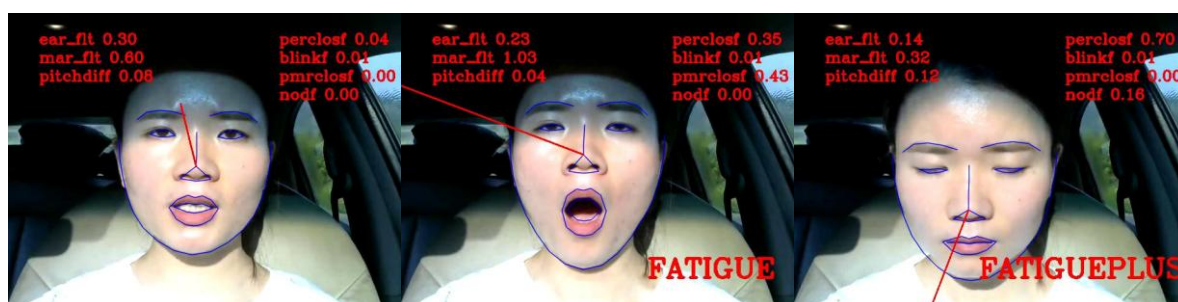


图 5.15 曝光下驾驶状态检测效果



图 5.16 弱光下驾驶状态检测效果

图 5.14、图 5.15、图 5.16 中左侧从上至下表示该帧实时的眼睛高宽比、嘴部高宽比、点头角度，右侧从上至下表示上一个检测窗口的 PERCLOS、眨眼频率、哈欠频率、点头频率。从图中可以看出，在各种光照条件下，Stasm 形状大体上能够实现正确匹配。但在曝光条件下，闭眼时由于光线原因会存在形状过匹配的情况；在弱光条件下，由于缺少光线会存在匹配形状缩小化的情况，故降低了算法检测准确率。

表 5.7 正常光照下算法检测准确率

实际驾驶	ELM 分类状态			测试集样	分类准确率
状态	清醒	疲劳	非常疲劳	本数	
清醒	62	6	2	70	88.57%
疲劳	6	60	4	70	85.71%
非常疲劳	0	9	61	70	87.14%
总计	准确分类：183			210	87.14%

表 5.8 曝光条件下算法检测准确率

实际驾驶	ELM 分类状态			测试集样	分类准确率
状态	清醒	疲劳	非常疲劳	本数	
清醒	62	5	3	70	88.57%
疲劳	9	60	1	70	85.71%
非常疲劳	1	10	59	70	84.29%
总计	准确分类：181			210	86.19%

表 5.9 弱光条件下算法检测准确率

实际驾驶	ELM 分类状态			测试集样	分类准确率
状态	清醒	疲劳	非常疲劳	本数	
清醒	59	9	2	70	84.29%
疲劳	4	60	6	70	85.71%
非常疲劳	1	9	60	70	85.71%
总计	准确分类：179			210	85.24%

由表 5.7、表 5.8、表 5.9 可知，正常光照条件下，非常疲劳驾驶状态下误判为清醒结果组数为零，算法整体分类准确率为 87.14%；曝光条件下，算法整体分类准确率为 86.19%；弱光条件下，算法整体分类准确率下降为 85.24%。故本文疲劳检测算法可适用于各种光照条件，其平均分类准确率为 86.19%。

5.5 小结

基于模拟驾驶仪和 USB 摄像头采集了 7 名志愿者的驾驶视频，通过前面几章

的算法计算其具体疲劳特征，结合对应标签构成样本集，训练及测试疲劳判断 ELM 分类模型，得到分类准确率达 88.02% 的 ELM，将训练好的分类模型封装好与提取具体疲劳特征的算法程序整体化，移植至 Jetson TX2 嵌入式平台中，实现车载式疲劳检测。最后对整个算法进行测试验证，平均分类准确率为 86.19%。

总结与展望

驾驶员疲劳驾驶为社会带来了严重的人员伤亡和财产损失，驾驶员疲劳检测技术作为智能驾驶的关键技术之一，对提高行车安全性有很大的必要性。对此，本文对基于视觉的多特征疲劳检测系统进行了完整的研究。本文的研究内容及完成的工作主要包括为以下几点：

1.通过大量相关课题的文献阅读和研究，介绍了驾驶员疲劳检测的研究背景及课题意义，介绍和比较了疲劳驾驶检测技术 3 种分类方法，总结介绍了驾驶员疲劳检测国内外研究现状，并由此提出本文的研究方法与内容，即基于视觉的多特征疲劳检测研究。

2.首先设计了图像预处理流程，通过加权平均法实现了图像灰度化，通过直方图均衡实现了图像光补偿，通过中值滤波实现了图像降噪，降维处理提高了算法处理速率，图像质量的改善也为后续检测算法提高了其鲁棒性。

3.然后通过人脸检测对人脸进行初步定位，为接下来的特征点匹配缩小范围。改进了 AdaBoost 人脸检测算法，改用 MB-LBP 特征，人脸平均检测时间加快了近一倍，人脸检测更加细致了，并且大大减少了光照强度的影响。通过基于加州理工学院人脸图像库的试验对比了基于 haar 特征的 AdaBoost 和基于 MB-LBP 特征的 AdaBoost 人脸检测，前者检测率为 94.44%，检测时间为 59ms/张，后者检测率增加至 97.33%，检测时间缩短至 32ms/张。又通过白天和夜晚的视频序列的试验检测，验证了本文基于 MB-LBP 特征的 AdaBoost 人脸检测算法，白天检测率达 98.13%，晚上检测率达 96.73%，受光照影响不大。

4.接着提取并计算疲劳特征。先经 Stasm 算法提取了面部基础特征，再结合相机位姿测量提取了头部基础特征。特定驾驶员状态初始化并得到其基础特征的具体阈值，通过求取 PERCLOS、EBF、哈欠频率，点头频率的流程设计，进一步计算得到了这 4 个具体疲劳特征，并且通过比较包含清醒、疲劳和非常疲劳状态的驾驶视频的 4 个具体疲劳特征输出，定性观察到明显的可分类性，说明这 4 个指标可以很好地反映驾驶员驾驶状态。

5.最后完成了实验分类和移植测试。搭建了基于模拟驾驶仪和相机的疲劳驾驶实验平台，采集了 7 个志愿者的驾驶视频，根据标准分割拼接视频完成了驾驶视频的分类。利用前面章节算法得到的 4 个具体疲劳特征构成样本集，经交叉验证法训练和检测了 ELM 分类器，其整体性能的分类准确率可达到 88.02%，并且驾驶员非常疲劳状态不会误判为清醒状态。封装统一了算法程序并移植到了嵌入式平台 Jetson TX2 开发板上，实现了疲劳检测。不同光照条件下进行了疲劳检测，对测试结果进行分析，整个算法平均分类准确率为 86.19%，验证了疲劳检测算法的可靠性。

综上所述, 本文提出了一套完整的基于机器视觉的多特征疲劳检测算法。首先提出图像预处理流程并测试效果良好, 然后对比分析了基于 haar 特征和基于 MB-LBP 特征的 AdaBoost 人脸检测算法, 通过试验测试选择更高效的后者, 并通过试验验证了基于 MB-LBP 特征的 AdaBoost 人脸检测算法在白天和夜晚的检测效果, 接着提出基于 Stasm 整体化提取面部基础特征, 再结合相机位姿测量提取头部基础特征, 搭建模拟驾驶试验平台进行驾驶视频的采集, 计算具体疲劳特征, 经 ELM 分类器进行疲劳判断, 将整体化程序移植至 Jetson TX2 开发板上, 最后通过试验验证本文检测算法的准确性。

然而, 由于个人能力以及时间的限制, 本疲劳检测算法虽完整但不够完善, 为弥补其不足, 可以从以下几个方面来进行后续的研究工作:

1. 本文虽从各个方面尽量地减少了光照的影响, 但本文疲劳检测的数据来源是视觉, 故对于曝光、黑暗和佩戴墨镜等极端条件下会造成检测来源数据的黑洞, 故后期可以考虑使用红外夜视摄像头采集信号, 或者结合心理信息或车辆行为进行联合检测。

2. 基于相机位姿测量提取头部姿态受关键点位置影响大, 关键点的不对称会导致头部姿态的抖动, 后期可以考虑更可靠的方法提取头部姿态, 提高算法稳定性。

3. 驾驶员疲劳是一个持续深入的过程, 而本文为顾及算法实时性采用了一个 10s 的检测窗口进行疲劳判断, 未考虑检测窗口时间上的累积性, 故以后可以记录每个检测窗口的判断并作为之后检测窗口判断的指标之一。

4. 考虑整个 ADAS 处理能力的需要, 选择了性能强大的 Jetson TX2 作为移植平台, 但是其价格昂贵, 若仅仅进行驾驶员疲劳检测存在很大的资源浪费, 后期可以结合具体的应用需求, 选择性价比最高的嵌入式开发板。

5. 本文的判断结果仅仅通过屏幕上的状态显示进行输出, 而实际应用过程中需加入预警的效果, 该疲劳检测系统才会存在意义, 故后期可以利用蜂鸣器搭建频率不同的分级报警机制。本设计还可以集成到高级驾驶辅助系统 ADAS 中, 针对当前 ADAS 系统频繁报警, 干扰驾驶员的正常驾驶, 影响驾乘体验的问题, 提出基于驾驶员疲劳状态的 ADAS 分级报警机制, 当驾驶员处于清醒的状态下减少报警频次, 而当驾驶员出现疲劳状况, 则调高报警频次; 在保证驾驶安全的情况下, 提高驾乘体验。

总而言之, 本文完成了基于机器视觉的多特征疲劳检测研究。经图像预处理后, 通过基于 MB-LBP 特征的 AdaBoost 进行人脸的初步定位, 提取并计算疲劳特征, 通过 ELM 分类器进行驾驶状态分类, 移植至 Jetson TX2 嵌入式开发板中, 并通过试验以完成验证。同时, 本文还存在很多不足, 有待后续进一步改进和完善。

参考文献

- [1] 赵黎. 2018 年中国汽车市场盘点. 汽车纵横, 2019, (2): 35-37.
- [2] World Health Organization. Global Status Report on Road Safety 2015. Geneva: World Health Organization, 2015, 63-64.
- [3] 公安部交通管理科学研究所. 中华人民共和国道路交通事故统计年报(2015 年度). 无锡: 公安部交通管理局, 2016.
- [4] Liberty E, Mazzae E, GarrottR, et al. NHTSA Driver Distraction Research: Past Present and Future, 2001, 33-34.
- [5] 陈勇, 黄琦, 刘霞, 等. 一种全天候驾驶员疲劳检测方法研究. 仪器仪表学报, 2009, 30(3): 636-640.
- [6] 黄丽娜. 乌饭树叶提取物抗大鼠精神疲劳作用及其机制研究: [第二军医大学硕士学位论文]. 上海: 第二军医大学, 2008, 11-13.
- [7] 夏国华, 郑安钦. 浅谈疲劳驾驶的预防. 中国科技博览, 2010(4): 186-186.
- [8] 谷成利. 谈疲劳驾驶的成因与预防措施. 农机使用与维修. 2018(9): 73.
- [9] 王军. 驾驶员疲劳检测算法研究: [北京交通大学硕士学位论文]. 北京: 北京交通大学, 2017, 2.
- [10] National Transportation Safety Board. Special Investigation Report-Highway Vehicle and Infrastructure Based Technology. For the Prevention of Rear-end collisions. NTSB number SIR-01/01. MAY 2001.
- [11] 白金蓬, 黄英, 江宜舟, 等. 驾驶状态实时监测系统. 电子测量与仪器学报, 2014, 28(9): 965-973.
- [12] 石坚, 吴远鹏, 卓斌, 等. 汽车驾驶员主动安全性因素的辨识与分析. 上海交通大学学报, 2000, 34(4): 441-444.
- [13] Aurelio Piazzzi, Corrado Guarino Lo Bianco, Massimo Bertozzi. Quintic G2 – Splines for the Iterative Steering of Vision-based Autonomous Vehicles. IEEE Transactionson Intelligent Transportation Systems, 2002, 3(1): 27-36.
- [14] Kecklund G. Sleepiness in Long Distance Truck Driving. Ergonomics, 1993, 36.
- [15] Lal S K, Craig A, Boord P, et al. Development of Analgorithm for an EEG-based Driver Fatigue Counter Measure. Journal of Safety Research, 2003, 34(3): 321-8.
- [16] Lal S K, Craig A. Driver fatigue: Electroencephalography and Psychological Assessment. Psychophysiology, 2002, 39(3): 313-321.
- [17] Jeong I C, Dong H L, Park S W, et al. Automobile Driver's Stress Index Provision System that Utilizes Electrocardiogram. Intelligent Vehicles Symposium. IEEE, 2007: 652-656.

- [18] 吴群. 基于心电信号的驾驶疲劳检测方法研究: [浙江大学硕士学位论文]. 杭州: 浙江大学, 2008, 4-5.
- [19] 张智腾. 基于卷积神经网络的驾驶员疲劳检测: [湖南大学硕士学位论文]. 长沙: 湖南大学, 2018, 2-3.
- [20] Hu X, Eberhart R, Foresman B. Modeling drowsy driving behaviors. IEEE International Conference on Vehicular Electronics & Safety. 2010, 13-17.
- [21] Wang M S, Jeong N T, Kim K S, et al. Drowsy Behavior Detection Based on Driving Information. International Journal of Automotive Technology, 2016, 17(1): 165-173.
- [22] 初秀民, 严新平, 吴超仲, 等. 基于计算机视觉的驾驶员转向操作实时监测研究. 汽车工程, 2005, 27(5): 522-524.
- [23] Lin C T, Chang C J, Lin B S, et al. A Real-time Wireless Brain-computer Interface System for Drowsiness Detection. IEEE Transactions on Biomedical Circuits & Systems, 2010, 4(4): 214-222.
- [24] Acioglu A, Ercelebi E. Real Time Eye Detection Algorithm for PERCLOS Calculation. 2016 24th Signal Processing and Communication Application Conference (SIU). IEEE, 2016.
- [25] Nguyen T P, Chew M T, Demidenko S N. Eye Tracking System to Detect Driver Drowsiness. International Conference on Automation. IEEE, 2015.
- [26] 童兵亮. 基于嘴部状态的疲劳驾驶和精神分散状态监测方法研究: [吉林大学硕士学位论文]. 长春: 吉林大学, 2004, 2-60.
- [27] Grace R, Byrne V E, Bierman D M, et al. A Drowsy Driver Detection System for Havy Vehicles. Digital Avionics Systems Conference. 2002.
- [28] Cui X, Ying Z, Wang Z. Efficient Eye States Detection in Real-time for Drowsy Driving Monitoring System. International Conference on Information & Automation. 2008.
- [29] Lu Y, Li C. Recognition of Driver Eyes' States Based on Variance Projections Function. International Congress on Image & Signal Processing. IEEE, 2010.
- [30] Kan M, Kan M Shan S, et al. Funnel-structured Cascade for Multi-view Face Detection with Alignment-awareness. Neurocomputing, 2017, 221: 138-145.
- [31] Kim Y, Kim Y, Hahn M. Detecting Driver Fatigue Based on the Driver's Response Pattern and the Front View Environment of an Automobile. International Symposium on Universal Communication, 2008 : 237-240.
- [32] 苟群森. 基于安卓的多特征疲劳实时检测系统设计与实现: [电子科技大学硕士学位论文]. 成都: 电子科技大学, 2016, 3-4.

- [33] 唐杰. 基于多信息融合的实时疲劳检测与预警系统研究: [南京航空航天大学硕士学位论文]. 南京: 南京航空航天大学, 2018, 3-4.
- [34] 王安. 喷淋——预防疲劳驾驶新点子. 道路交通管理, 2013, (10): 51.
- [35] 陈艺丹. 基于面部子空间特征的疲劳驾驶检测软件研究: [长春理工大学硕士学位论文]. 长春: 长春理工大学, 2012, 3-4.
- [36] Saroj K L L, Ashley Craig. Reproducibility of the Spectral Components of the Electroencephalogram during Driver Fatigue. International Journal of Psychophysiology, 2005, 55(2): 137-143.
- [37] Seeing Machines. FaceLAB V5 系统. <http://www.seeingmachines.com>, 2014-3-13.
- [38] 刘志强, 秦洪懋, 汪旻等. 驾驶疲劳监测系统 DDDS 设计方法. 江苏大学学报(自然科学版), 2008, 29(1): 25-28.
- [39] 庄连生. 复杂光照条件下人脸识别关键算法研究: [中国科学技术大学博士学位论文]. 合肥: 中国科学技术大学, 2006, 12-13.
- [40] Adini Y, Moses Y, Ullman S. Face Recognition: the Problem of Compensating for Changes in Illumination Direction. Pattern Analysis & Machine Intelligence IEEE Transactions, 1994, 19(7): 721-732.
- [41] 肖剑, 雄峰. 局部对比度增强的彩色图像灰度化参数化算法研究: [温州大学硕士学位论文]. 温州: 温州大学, 2016, 7-8.
- [42] Yin H , Fu P , Meng S . An Efficient Face Detection Method in Color Images. Knowledge-based Intelligent Information & Engineering Systems, International Conference, Kes, Melbourne, Australia, September. DBLP, 2005.
- [43] 王耀南. 计算机图像处理与识别技术(第 1 版). 北京: 高等教育出版社, 2001, 72-73.
- [44] 毛星云, 冷雪飞, 等. OpenCV3 编程入门(第 1 版). 北京: 电子工业出版社, 2015, 154-159.
- [45] 何俊, 房灵芝, 蔡建峰, 何忠文. 基于 ASM 和肤色模型的疲劳驾驶检测. 计算机工程与科学, 2016, 38(7).
- [46] Sakai T, Nagao M, Fujibayashi S. Line Extraction and Pattern Detection in a Photograph. Pattern Recognition, 1969, 1(3): 233-248.
- [47] Kass M, Witkin A, Terzopoulos D. Snakes: Active Contour Models. International Journal of Computer Vision, 1988, 1(4): 321-331.
- [48] Valiant L G. A Theory of the Learnable. Communications of the ACM, 1984, 27(11): 1134-1142.

- [49] Schapire R E. The Strength of Weak Learnability. Machine Learning, 1990, 5(2): 197-227.
- [50] Freund Y, Schapire R E. A Decision-theoretic Generalization of On-line Learning and an Application to Boosting. Journal of Computer and System Sciences, 1995, 90(4): 23-37.
- [51] Paul Viola, Michael Jones. Rapid Object Detection Using a Boosted Cascade of Simple Features. In Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition. 2001, 1: 1511-1517.
- [52] Kearns M. The Computational Complexity of Machine Learning. Cambridge: MIT Press, 1990. 109-116
- [53] 刘艺, 龚卫国, 李伟红. 双层结构 Adaboost 健壮分类器用于人眼睛精确定位. 计算机应用, 2008, 28(3): 207-308.
- [54] Wei Y. Face Alignment by Explicit Shape Regression. IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2012: 2887-2894.
- [55] 彭发超. 基于视觉的驾驶员疲劳检测算法研究: [湖南大学硕士学位论文]. 长沙: 湖南大学, 2016, 22-39.
- [56] Ojala T, Pietikinen M, Harwood D. A Comparative Study of Texture Measures with Classification Based on Featured Distributions. Pattern Recognition: The Journal of the Pattern Recognition Society, 1996, 29(1): 51-59.
- [57] Zhang L, Chu R, Xiang S, et al. Face Detection Based on Multi-Block LBP Representation. Advances in Biometrics, International Conference, Icb, Seoul, Korea, August. 2007.
- [58] Wei Sun, Xiaorui Zhang, Wei Zhuang, Huiqiang Tang. Driver Fatigue Driving Detection Based on Eye State. International Journal of Digital Content Technology and its Applications, 2011, 5(10), 307-314.
- [59] 姚胜, 李晓华, 张卫华, 等. 基于 LBP 的眼睛开闭检测方法. 计算机应用研究, 2015, 32(6): 1897-1901.
- [60] 黄永慧, 潘保昌, 梁坚, 等. 一种自适应唇区检测及定位方法. 计算机工程与应用, 2010, 46(21): 17-20.
- [61] Cootes T F, Taylor C J. Statistical Models of Appearance for Computer Vision. <http://www.isbe.man.ac.uk>. 2004, 12-28, 37-41.
- [62] 戈新良, 杨杰, 张田昊, 等. 改进的主动形状模型方法在人脸特征点定位中的应用. 上海交通大学学报, 2007, 41(8).
- [63] Milborrow S, Nicolls F. Locating Facial Features with an Extended Active Shape

- Model. European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2008.
- [64] 谭力凡. 机器视觉与毫米波雷达融合的前方车辆检测方法研究: [湖南大学硕士学位论文]. 长沙: 湖南大学, 2018, 37-46.
- [65] 文芳. 基于 SDM 的疲劳驾驶状态检测方法研究: [长安大学硕士学位论文]. 西安: 长安大学, 2018, 36.
- [66] 白中浩, 刘浏, 焦英豪, 曹松. 基于 ASM+的多特征融合驾驶员疲劳检测方法. 电子测量与仪器学报, 2016(12).
- [67] DINGES D F, GRACE R. PERCLOS: A Valid Psychophysiological Measure of Alertness as Assessed by Psychomotor Vigilance. US Department of Transportation, Federal Highway Administration, Publication Number FHWA-MCRT-98-006, 1998.
- [68] 刘志强, 宋雪松, 汪彭, 等. 基于眼部特征的疲劳驾驶辨识方法研究. 重庆理工大学学报 (自然科学版), 2016, 30(10).
- [69] 张希波, 成波, 冯睿嘉. 基于方向盘操作的驾驶人疲劳状态实时检测方法. 清华大学学报: 自然科学版, 2010(7): 1072-1076.
- [70] Huang G B, Bai Z, Kasun L L C, et al. Local Receptive Fields Based Extreme Learning Machine. Computational Intelligence Magazine, 2015: 10(2): 18-29.
- [71] 于兹文. 驾驶员疲劳状态分析及其预警算法研究: [吉林大学硕士学位论文]. 长春: 吉林大学, 2018, 61-63.
- [72] 周志华. 机器学习 (第 1 版). 北京: 清华大学出版社, 2017, 26-27.
- [73] 宋灵杰. 基于卷积神经网络的交通标志分类与识别研究: [湖南大学硕士学位论文]. 长沙: 湖南大学, 2018, 48-49.

附录 A 攻读学位期间所发表的学术论文

- [1]湖南大学. 基于ASM和ELM的驾驶员疲劳检测报警系统 V1.0, 软件著作权, 2019SR0166183, 2019-02-21.

致谢

在湖南大学的三年研究生学习及生活时光即将过去，回想三年，自己得到了成长也收获了不少，这些都要感谢我的导师、同门、家人及朋友的教导鼓励与帮助支持。

首先，感谢我的导师曹立波教授。曹老师所带领的曹门团队人才济济，资源丰富，每周一次的学术报告活动开阔了我的视野，彼此间的讨论交流让我了解到同门研究课题的同时，让曹老师更好地提供指导和建议，另外工作汇报和打卡习惯共同营造了整个团队浓厚的学术氛围。曹老师知识渊博，思路开阔，为我的研究课题提供了宝贵的意见和学习平台及资源。曹老师鼓励创新，给予了我足够的学术自由度，让我能顺利地完成我的课题研究。曹老师严谨治学的态度，大胆创新的进取精神和谦逊平和的处事风格深深地影响了我，再次感谢曹老师的教导与帮助！

同时，感谢颜凌波老师和吴俊老师。颜老师追求学术创新的精神和吴老师踏实严谨的学术态度令我敬佩并值得学习。

此外，感谢冯谢星博士。冯师兄给我的研究课题提供了十分珍贵的意见，保证了课题前进方向的正确性，在课题遭遇瓶颈的时候为我指出关键点，并不断鼓励我，让我的研究课题得以顺利进行。冯师兄对项目方向的把控度，钻研学术时认真勤勉的态度和迎难而上的不懈坚持与积极乐观深深撼动我，再次感谢冯师兄的帮助与鼓励！

同时感谢我的同门廖家才和李伟等博士，感谢我的同门谭力凡、李俊义、张颖麟、陈龙、樊冰冰、隆旭、石海兵等师兄师姐，感谢我的同门陶武康、张韵仙、喻志强、李伟、张晓、张乐祺、张建强、房建伟、江志国、吴金正、梁煜，感谢我的同门夏家豪、方晨晨、孙杰、乐路遥、付梦佳、常启瑜、向国梁、韦添元、曹文轩、罗荣华、张超、吴恒基等师弟师妹，同门的陪伴与关心让我在研究生期间充满了温馨与快乐，能够遇见优秀的各位同门让我感到莫大的幸运。另外，感谢参与我研究课题志愿者实验的同门，感谢你们的帮助与支持！

感谢我的室友邹思、段娇娇、李金燕等博士学姐对我生活上的照顾与陪伴，很荣幸能够遇见优秀的你们，你们的认真勤勉、乐观大方、对生活睿智的思考触动了我很多。特别感谢家人以及朋友一直以来温暖的陪伴，谢谢你们的包容、鼓励与支持。谢谢段娇娇学姐借车给我，也谢谢家人假期间陪我一起做测试实验。

最后，感谢湖南大学，给了我三年研究的学习平台，并结识了很多优秀的人，感谢一路上给我鼓励使我成长的所有人，谢谢你们！