

学校代号 10532

学 号 S1610W0599

分 类 号 TP393

密 级 公开



湖南大学
HUNAN UNIVERSITY

工程硕士学位论文

基于多特征的疲劳驾驶识别算法研究 与实现

学位申请人姓名 徐青云

培 养 单 位 信息科学与工程学院

导师姓名及职称 胡峰松副教授 胡斌晖高级工程师

学 科 专 业 计算机技术

研 究 方 向 图像识别

论 文 提 交 日 期 2019 年 04 月 15 日

学校代号：10532

学 号：S1610W0599

密 级：公开

湖南大学工程硕士学位论文

基于多特征的疲劳驾驶识别算法研究 与实现

学位申请人姓名：徐青云

导师姓名及职称：胡峰松副教授 胡斌晖高级工程师

培 养 单 位：信息科学与工程学院

专 业 名 称：计算机技术

论文提交日期：2019年04月15日

论文答辩日期：2019年05月04日

答辩委员会主席：秦拯 教授

The Research and Implementation of Fatigue Driving Recognition Algorithm Based on Multiple Features

by

Xu Qingyun

B.E.(Wuhan Textile University)2016

A thesis submitted in partial satisfaction of the

Requirement for the degree of

Master of Engineering

In

Computer Technology

in the

Graduate School

of

Hunan University

Supervisor

Associate Professor Hu Fengsong

May, 2019

摘 要

疲劳驾驶引发的道路交通事故已使人们的生命财产安全受到了严重的威胁。如何提前检测驾驶员的疲劳程度并预警，从根本上避免交通事故的发生已成为目前疲劳驾驶检测技术研究领域中的一大热点。本文对现有的驾驶员疲劳检测方法进行分析和总结后，提出一种基于驾驶员行为多特征加权求和的疲劳识别方法。主要研究工作有：

1. 图像预处理及人脸检测与跟踪。由于采集的视频图像在存储过程中可能会受到各类噪声不同程度的污染。因此预先对图像进行平滑去噪及光照补偿处理，以确保人脸检测的准确性。然后采用高精度的基于 Harr-like 的 AdaBoost 人脸检测算法检测人脸。在检测到图像中的人脸后，采用判别尺度空间的跟踪算法（DDST）对面部区域实时跟踪。

2. 针对于疲劳时人眼状态变化，本文提出了一种基于 SVM 多特征融合的睁闭眼状态识别方法。首先采用基于级联回归树算法对检测到或跟踪到的人脸进行特征点定位，根据人眼特征点的位置定位人眼区域。由人眼的 12 个特征点计算人眼纵横比 EAR，根据 EAR 来识别眼部睁闭状态；使用自适应域方法计算人眼二值图像的黑色像素累积差值来识别人眼睁闭状态。最后将两种方法得到的特征值作为 SVM 分类器的输入参数进行模型训练，使用训练好的模型对睁闭眼状态进行识别。

3. 对于嘴部疲劳状态的识别，主要根据嘴部 10 个特征点计算嘴部高宽比 MAR，并根据设定的 MAR 阈值来识别打哈欠状态。头部疲劳状态则针对头部运动计算二维垂直方向的点头频率。计算定位的左、右眼区域的两个中心点位置，取两中心点的中点作为头部位置运动分析点，以该检测点在垂直方向上坐标 y 随时间的变化情况，计算时间周期内的点头频率。

4. 根据眼、嘴、头部的疲劳特征提出了一种基于多特征加权求和的疲劳识别方法，并建立疲劳识别模型。根据人眼睁闭状态提取眼部疲劳参数持续闭眼时间、闭眼帧数比；通过眨眼检测得到眨眼频率；通过打哈欠检测得到打哈欠次数及持续时间；通过分析头部运动状态得到点头频率。将这些指标进行加权求和来评估驾驶员的疲劳等级并进行相应的预警。经实验验证，表明本文提出的方法识别准确率较高，实时性好。

关键词：AdaBoost 算法；特征点定位；睁闭眼状态识别；多特征疲劳识别；

Abstract

Road traffic accidents caused by fatigue driving have seriously threatened people's lives and property. How to detect the driver's fatigue level in advance and give early warning to avoid traffic accidents fundamentally has become a hot spot in the research field of fatigue driving detection technology. In this paper, after analyzing and summarizing the existing driver fatigue detection methods, a fatigue identification method based on facial multi-feature weighted sum is proposed. The main research work is as follows:

1. Image preprocessing and face detection and tracking. Since the captured video images may be contaminated with various types of noise in different degrees during the storage process. In order to ensure the accuracy of face detection, image denoising and illumination compensation are processed in advance. Then AdaBoost face detection algorithm based on Harr-like is used to detect faces. After the face in the image is detected, the facial region is tracked in real time using the tracking algorithm of discriminating scale space.

2. In view of the state change of human eyes in fatigue, we propose an open and closed eye state recognition method based on SVM multi-feature fusion. Firstly, a cascade regression tree algorithm is used to locate the feature points of the detected or tracked face, and the human eye region is located according to the position of the human eye feature points. The aspect ratio EAR of human eyes was calculated from 12 feature points of human eyes, and the state of eye opening and closing was identified according to the EAR. An adaptive domain method is used to calculate the cumulative difference of black pixels in the binary image of human eyes. Finally, the eigenvalues obtained by the two methods are used as input parameters of the SVM classifier for model training, and the trained model is used for the recognition of the open and closed state of human eyes.

3. For the identification of the fatigue state of the mouth, the aspect ratio MAR of the mouth was mainly calculated based on 10 characteristic points of the mouth, and the yawning state was identified according to setting the MAR threshold. In the head fatigue state, the head frequency in the two-dimensional vertical direction is calculated according to the head movement. The position of the two center points in the positioned left and right eye area was calculated, and the midpoint of the two

center points was taken as the head position motion analysis point. The change of the vertical coordinate y of the detection point with time was used to calculate the nod frequency within the time period.

4. The fatigue identification model was established according to the fatigue characteristics of eyes, mouth and head, and a fatigue identification method based on multi-feature weighted sum is proposed. The duration and closed eye frame ratio of eye fatigue parameters were extracted according to the condition of eye opening and closing. Blink frequency can be obtained by blink detection. The number and duration of yawning were measured by yawning test. The head frequency is obtained by analyzing the head motion state. These indicators are weighted and summed to evaluate the fatigue level of drivers and give corresponding early warning. Experimental results show that the method proposed in this paper has high accuracy and good real-time performance.

Key words: AdaBoost algorithm; feature points location; Eyes open and close state recognition; Multi-feature fatigue identification;

目 录

学位论文原创性声明.....	I
学位论文版权使用授权书.....	I
摘 要.....	II
Abstract.....	III
目 录.....	V
插图索引.....	VII
附表索引.....	IX
第 1 章 绪 论.....	1
1.1 研究背景与意义.....	1
1.2 疲劳驾驶检测方法简述.....	2
1.2.1 基于车辆的检测方法.....	2
1.2.2 基于驾驶员的检测方法.....	3
1.2.3 基于信息融合的检测方法.....	4
1.3 国内外研究现状.....	5
1.3.1 国外研究现状.....	5
1.3.2 国内研究现状.....	6
1.4 本文组织结构与研究内容.....	7
第 2 章 人脸检测及面部目标跟踪.....	9
2.1 图像预处理.....	9
2.1.1 图像平滑去噪.....	9
2.1.2 基于动态阈值的光照均衡处理.....	11
2.1.3 实验结果分析.....	12
2.2 基于 AdaBoost 算法的人脸检测.....	14
2.2.1 人脸检测方法分类.....	14
2.2.2 基于 Haar-like 特征的 AdaBoost 人脸检测算法.....	15
2.2.3 实验结果分析.....	18
2.3 基于高置信度的面部目标跟踪算法.....	19
2.3.1 跟踪算法分类.....	19
2.3.2 DSST 算法介绍.....	21
2.3.3 高置信度更新策略.....	24
2.3.4 本文算法步骤.....	24

2.3.5 实验结果分析.....	25
2.4 本章小结.....	27
第 3 章 基于 SVM 的睁闭眼状态识别及打哈欠状态识别.....	28
3.1 基于级联回归树的人脸特征点定位.....	28
3.1.1 人脸关键点定位算法分类.....	28
3.1.2 基于级联回归树算法介绍.....	29
3.1.3 实验结果分析.....	31
3.2 基于 SVM 的睁闭眼状态识别.....	31
3.2.1 人眼检测定位.....	32
3.2.2 基于人眼特征点计算人眼纵横比.....	33
3.2.3 基于自适应阈值计算人眼黑色像素累积差值.....	35
3.2.4 实验结果分析.....	38
3.3 基于嘴部纵横比的打哈欠状态识别.....	41
3.3.1 嘴部状态识别方法.....	41
3.3.2 基于纵横比的嘴部状态识别.....	41
3.3.3 实验结果分析.....	42
3.4 本章小结.....	44
第 4 章 基于多特征加权求和的疲劳状态识别.....	45
4.1 疲劳信息提取.....	45
4.1.1 眼部疲劳信息提取.....	45
4.1.2 嘴部疲劳信息提取.....	48
4.1.3 头部疲劳信息提取.....	48
4.2 基于多特征加权求和的疲劳状态识别.....	49
4.2.1 多特征加权求和疲劳状态识别方法.....	49
4.2.2 建立疲劳识别模型.....	51
4.2.3 整体实现流程.....	52
4.3 实验结果分析.....	53
4.4 本章小结.....	56
结 论.....	57
参考文献.....	59
附录 A 攻读学位期间的学术成果.....	64
附录 B 攻读学位期间参与的项目.....	65
致 谢.....	66

插图索引

图 2.1	去噪结果对比图.....	13
图 2.2	光照处理前后对比图.....	14
图 2.3	Haar-like 特征分类.....	16
图 2.4	矩形 D 区域像素和计算示意图.....	17
图 2.5	基于 AdaBoost 算法不同角度人脸检测结果.....	18
图 2.6	DSST 算法原理图.....	21
图 2.7	位置滤波器样例.....	21
图 2.8	目标位置估计过程.....	22
图 2.9	尺寸滤波器样例.....	23
图 2.10	目标尺寸估计过程.....	24
图 2.11	面部目标跟踪流程图.....	25
图 2.12	中心位置误差曲线图.....	26
图 2.13	重叠率曲线图.....	26
图 3.1	人脸特征点模型.....	29
图 3.2	不同角度人脸特征点检测结果.....	31
图 3.3	基于特征点的人眼定位示意图.....	33
图 3.4	人眼六个关键点示意图.....	33
图 3.5	左、右眼 EAR 结果图.....	34
图 3.6	ERA 均值结果图.....	35
图 3.7	人眼睁闭眼过程黑色像素数量.....	36
图 3.8	人眼连续两帧黑色像素数量差值.....	36
图 3.9	人眼黑色像素数量累积差值.....	36
图 3.10	自适应阈值黑色像素累积差值算法流程.....	37
图 3.11	自适应阈值人眼黑色像素数量累积差值.....	37
图 3.12	嘴部 10 个关键点示意图.....	41
图 3.13	嘴部 MAR 检测结果.....	42
图 3.14	YawDD 数据集部分样本数据.....	42
图 3.15	打哈欠检测过程.....	43
图 3.16	阈值寻优结果图.....	43
图 4.1	睁闭眼过程示意图.....	46
图 4.2	部分数据眨眼检测结果图.....	47

图 4.3	EAR 阈值、帧数 K 值寻优结果图.....	47
图 4.4	嘴部状态示意图.....	48
图 4.5	头部运动分析图.....	49
图 4.6	疲劳状态识别模型.....	51
图 4.7	疲劳识别流程图.....	52
图 4.8	疲劳状态识别测试数样例.....	53

附表索引

表 1.1	近六年中国机动车交通事故数据	1
表 1.2	疲劳检测算法分析	5
表 2.1	椒盐噪声各算法指标对比	13
表 2.2	高斯噪声各算法指标对比	14
表 2.3	AdaBoost 算法与阈值肤色模型人脸检测准确率对比	19
表 2.4	DSST 跟踪算法评估	26
表 3.1	人眼检测算法对比	32
表 3.2	不同情况下的人眼二值图像	35
表 3.3	睁闭眼状态检测结果	40
表 3.4	不同算法识别结果对比	40
表 4.1	眼部疲劳状态取值条件	48
表 4.2	各疲劳特征参数的权重取值	50
表 4.3	疲劳值与疲劳等级对应关系表	50
表 4.4	疲劳指标权值寻优	54
表 4.5	不同环境下疲劳识别结果	54
表 4.6	白天戴眼镜疲劳识别结果	55
表 4.7	疲劳检测准确率对比	56
表 4.8	各模块平均运行时间	56

第 1 章 绪 论

1.1 研究背景与意义

近几十年来，随着人们生活质量的不断提高，汽车数量也在逐步增加，它已经成为了人们出行的主要的交通工具。根据《2018 年国民经济和社会发展统计公报》的统计数据显示，到 2018 年末，全国民用汽车的数量为 24028 万辆，其中私人汽车数量为 20730 万辆，比 2017 年末的数量上长了 10.9%。而民用轿车有 13451 万辆，比上一年数量增长了 10.4%^[1]。随着汽车数量的速猛增长，其引发的道路交通事故也在直线上升，汽车在给人们生活带来快捷与便利的同时，频繁的道路交通事故也带来了惨重的经济损失，使人民的生命受到了巨大的威胁，疲劳驾驶已成为全球一个严重而亟待解决的问题。

由 2018 年全球道路安全状况报告可以知，每年死于道路交通事故的人次已达到 135 万，而主要受害人群的年龄范围为 5-29 岁^[2]，道路交通事故伤害现已成为全球人员伤亡的一项最大原因。由下表国家统计局统计的交通事故数据可知中国的道路交通事故问题极其严重，中国每年因道路交通事故死亡的人数约达 6 万人次以上，直接经济损失约 10 亿元。在过去六年中，机动车导致的交通事故达 1099539 起，致 33978 人死亡，造成的直接经济损失约为 64.7 亿元^[3]。

表 1.1 近六年中国机动车交通事故数据

年份	发生起数/起	死亡人数/人	受伤人数/人	直接损失/万元
2017	139412	59166	188585	115556.2
2016	192585	58803	205355	114586.4
2015	170130	54279	181528	98928.6
2014	180321	54944	194887	103386.0
2013	183404	55316	198317	100034.1
2012	190756	57277	210554	114199.5

从表 1.1 近六年来我国机动车交通事故的统计数据可以看出，虽然国家已经陆续采取了很多相关措施来控制道路交通事故的发生，但是我国每年由机动车造成的道路交通事故起数并没有明显的下降。大部分道路交通事故的发生主要是由于

驾驶员在行车过程中的警觉性、注意力、反应速度、判断力和决策能力有所下降，而超速、酒驾、疲劳驾驶等为主要原因。在这些危险因素中，驾驶员疲劳驾驶占比重是仅仅次于酒驾和超速行驶的，每年因疲劳驾驶而导致的道路交通事故占 20%^[4]，疲劳驾驶已逐渐成为人员死亡的重要因素。据美国国家公路交通安全管理局统计，2015 年有 90000 起车祸涉及疲劳驾驶，这些事故导致约 41000 人受伤，800 多人死亡。在 2016 年，几乎每十起车祸就有一起是由于疲劳驾驶造成的^[5]。由德国保险公司协会对高速交通事故的统计，约 25% 的交通事故是由疲劳驾驶导致的^[6]。

在实际行车中，我们很难遏制疲劳驾驶的现象，因为影响疲劳的因素太多，比如睡眠质量、驾驶时间、服用的药物有催眠作用、喝的饮料含酒精成分、心情的好坏等，这将致使我们很难判断驾驶员是否疲劳驾驶。这些因素会对驾驶员在行车过程中的思维、感知能力、警惕性、反应能力、判断能力和操控行为造成不同影响。当驾驶员进入疲劳状态时会变得动作迟缓、反应慢，对周围的行车状况没有判断力，甚至忘记操控行驶中的车辆，在这种状态下很容易发生交通事故，造成不可估量后果。

如何提前检测驾驶员的疲劳程度并及时预警，从根源上避免引发交通事故已成为目前疲劳驾驶研究领域中的一大难点，也是一大热点。为此，为了能够快速有效的判断驾驶者的疲劳状态，在其处于疲劳状态时及时提醒其停车休息，极大减少由于疲劳驾驶而引起的交通事故起数，提出一种实时而高效的疲劳驾驶识别方法来保障人们的生命财产安全具有重大意义。

1.2 疲劳驾驶检测方法简述

由于在全球因疲劳驾驶引发的危害较严重，而疲劳监控检测技术对预防及降低道路安全事故具有优良效果。随着相关研究领域技术的发展，疲劳驾驶检测技术已受到国内外众多研究学者的关注。经过相关研究数据调查，目前对于如何检测驾驶员疲劳状态的方法是多种多样的，但大体主要还是从基于车辆、基于驾驶员和基于信息融合三个方面来进行检测监控^[7]。

1.2.1 基于车辆的检测方法

主要通过采集车辆驾驶参数，分析参数的异常波动来判断疲劳状态。该类检测方法包括方向盘转角程度检测、方向盘转向握力检测，车速检测、车辆偏移检测、制动踏板力检测和加速踏板力检测等^[8-12]，目前的车辆大多均配备有不同类型的传感器，用于收集行车车速、方向盘角度、燃油消耗和发动机转速等实时参数，通过单独或综合分析这些数据可以间接地检测驾驶人员的疲劳状态。

方向盘转向角度检测：当驾驶员处于疲劳状态时，注意力会分散，此时控制

方向盘的振幅将增加，方向盘的角度值在一段时间内不会有显著的变化，方向盘转向频率会减小。通过采集和分析方向盘传感器在时间、振幅和频率上的方向盘角度参数来判断疲劳驾驶是一种较常用的方法。

方向盘握力检测：驾驶员困倦时，在行车过程中手握方向盘的握力要较正常驾驶时有所减弱。

车速检测：主要使用 CCD 摄像头和车载传感器来检测车辆的速度、横向加速度及横向位移等参数。

该类方法主要根据传感器实时采集的数据来进行疲劳驾驶检测，该项技术相对比较成熟且算法简单，故而此方法在该疲劳检测研究领域也受到了广泛的关注。但是其分析结果容易受到个人驾驶习惯、天气、车辆特性及道路状况等外界环境因素的影响，健壮性不强^[8]。而且只有在驾驶员即将发生交通事故时才能检测到异常，不能提前预警。因此，该类方法的分析结果最好用作辅助检测指标而不是主要的检测指标^[9]。

1.2.2 基于驾驶员的检测方法

基于驾驶员的方法可以分为基于驾驶员生理参数和基于驾驶员行为特征这两种检测方法：

(1) 基于生理参数检测：相关研究表明，当驾驶员处于疲劳状态时，生理反应会变慢，身体对外界的刺激反应会有所延迟，生理指标会偏离正常值。因此，通过生理传感器采集的驾驶员生理参数可用来判断驾驶员是否处于疲劳状态。其主要是基于脑电信号（EEG）、心电信号（ECG）及肌电信号（EMG）等^[13, 14]生理参数来进行检测的。

EEG 检测：记录大脑电波的变化的大脑电波，脑电图的变化反映了大脑神经细胞的电生理活动。EEG 信号已被证明是疲劳驾驶检测最具预测性和最可靠性的信号之一，因为它是大脑活动的直接量度^[15]。

ECG 检测：心率（HR）指数和心率变异性（HRV）指数是判断疲劳程度的重要指标，当驾驶员疲劳时，ECG 信号将出现明显规律性下降。

EMG 检测：当驾驶员从清醒过度到疲劳状态时，EMG 信号的频率将会持续下降，其振幅会逐渐增大^[16]。因此，对 EMG 信号进行分析可以判断驾驶员疲劳程度。

该类方法可靠性较强，很少有误检情况，能客观准确地反映驾驶员疲劳程度。但需要驾驶员穿戴数据采集设备采集相应数据，用户的体验感较差不说可能还会因影响驾驶员的正常操作而引发交通安全事故。因此，该方法在实际疲劳检测应用中受到一定限制。而且生理参数个体差异性较大，它会受驾驶员的性别、年龄、体型等因素影响，不利于采用统一标准进行疲劳判断。

(2)基于行为特征检测：当驾驶员比较困倦时，其面部特征将与清醒状态的面部特征不同。因此，利用计算机视觉技术分析驾驶员的面部特征数据，是一种有效的实时疲劳驾驶检测方法。该方法提取的特征参数主要有眼动特性(眨眼频率，PERCLOS，眼睛睁闭程度，注视方向等)、嘴部状态(打哈欠频率等)及头部位置^[17]。由于头部和面部特征的变化比较明显，易于被检测到，因此，目前该领域的研究焦点也聚集在这一块。

眼动特性检测：卡内基梅隆大学驾驶研究中心开发了 PERCLOS 系统^[17]。经实验证明疲劳程度与闭眼时间的长短有密切关系，眼睛闭合的时间越长则疲劳程度越严重。John Stern 通过研究发现，人在正常情况下眨眼一次眼睛闭合持续的时间约为 0.2s，每分钟眨眼 10-25 次，若驾驶过程中驾驶员眼睛持续闭合时间达到 0.5s 以上或者每分钟眨眼次数低于 10 次则容易发生安全事故^[19, 20]。

打哈欠检测：正常情况下，人嘴是闭合的，在说话时嘴部的张开程度也较小，而在打哈欠时嘴部持续张开时间比较长且张开程度明显比前面两种状态要大。因此在精确定位嘴部的情况下，可根据嘴部的宽高比或者似圆度来判断嘴部的张开程度。

头部运动检测：在行车过程中，驾驶员在困倦时头部位置会较清醒状态时的位置有所偏离，头部会出现上下点头或歪头等情况，这些可以作为检测疲劳的特征。因此，可以通过头部位置来评估驾驶员是否处于疲劳状态。

该类疲劳检测方法所需设备简单，且有公认的基于眼部状态判断疲劳驾驶的国际标准，它可以更快速效地确定驾驶员的疲劳程度。此外，它实现了非接触式的特征提取，不会影响驾驶员的正常驾驶。其缺点是识别算法较为复杂，特征提取及检测结果易受到遮挡和光照等因素的影响。但是随着机器学习等相关技术的发展进步，现已可以从图像中快速准确提取有用的特征用于分析疲劳状态。因此，基于驾驶员行为特征的疲劳监控检测方法将是疲劳驾驶检测领域未来发展的主流研究方向。

1.2.3 基于信息融合的检测方法

由于受驾驶员个体差异、驾驶环境、车辆特性及道路交通状况等多因素影响，如果只以单一信息标准来判断驾驶员的疲劳驾驶状态，很难保证疲劳检测的可靠性。通过将前面两种方法提取的多特征参数进行信息进行融合来建立疲劳检测模型，在很大程度上可以提高疲劳识别预警系统的准确性和可靠性，目前众多研究者都将精力集中在多特征信息融合共同检测疲劳驾驶状态的研究上。

该类检测方法融合多种疲劳特征，相比于基于单一特征信息疲劳检测方法的检测精度和可靠性都有所提高，但是使用现有技术提取各种特征并建立基于信息融合检测方法的模型存在较大挑战，且建立的疲劳检测模型对复杂环境的适用性

较差。

通过综合分析对比现有的疲劳驾驶检测方法，得到的目前各类检测方法的优缺点分析结果如表 1.2 所示。目前大多数的疲劳驾驶检测预警产品都是基于多特征信息融合来判断驾驶员的疲劳程度，这保证了检测结果的准确性。但是这类检测设备的生产成本较高，不利于疲劳驾驶检测系统的推广应用，故而现在急需研究成本低、精度高和可靠性强的产品。基于驾驶员行为特征的疲劳检测方法实现了非接触式实时采集疲劳特征，成本低、可靠性强，且随着计算机技术的发展，疲劳分析的准确性也将得到提高。

表 1.2 疲劳检测算法分析

方法	特点	优点	缺点
基于车辆参数检测	通过传感器实时采集数据来判断疲劳程度	获取特征信息方式简单，对人和车的干扰性小	易受到外界环境因素影响，准确性不够高
基于驾驶员生理参数检测	通过穿戴设备直接采集相应生理数据进行检测	测量技术成熟，方法可靠性较强，能客观准确地反映驾驶员疲劳程度	接触式检测，体验感差，会影响驾驶员的正常操作
基于驾驶员行为特征检测	通过监控驾驶员的面部和头部行为来判定疲劳状态	非接触式特征提取，干扰性低，可以准确地判断驾驶员的疲劳程度	算法复杂，特征提取易受光照和遮挡等因素影响
基于信息融合检测	通过建立多种疲劳特征信息融合模型来预测疲劳程度	较基于单一特征信息疲劳检测方法的精度和可靠性更高	信息融合模型建立困难，且检测模型对复杂环境的适用性较差

1.3 国内外研究现状

经不断研究实验及应用发现疲劳驾驶状态识别及预警技术可以有效的预防道路交通事故的发生，因而目前疲劳驾驶状态检测及预警技术已受到国内外研究人员的高度重视，疲劳驾驶检测技术已逐步从研究领域转移到工业应用领域，现有的一些研究成果也形成产品并开始投入使用。

1.3.1 国外研究现状

文献[21]通过 54 名驾驶员分别驾驶 10 辆车进行驾驶实验，提取车道位置相关的 11 个变量，使用 KNN、LDA、贝叶斯分类器、高斯混合模型和人工神经网络五种分类器建立检测模型，实验证明疲劳判别效果较为理想。

文献[22]在真实驾驶环境下对 20 名睡眠不足的受试者收集的多通道 EEG 数据进行研究，发现驾驶员处于疲劳状态时， δ 波相对频带比增加而 γ 波相对频带比

和 KL 熵明显减少，故可以该现象来判断驾驶人的状态。

Sangeetha 和 Kalpanadevi 开发了一种用 ECG 电极和 PIC 微控制器实现疲劳检测的硬件系统。将电极固定在驾驶员的手上采集驾驶员的 ECG 信号，并将采集的信号与已存储的疲劳驾驶 ECG 信号进行比较，当两个信号之间的相似性达到阈值时，说明驾驶员处于困倦状态^[23]。

美国 Iteris 开发的 AutoVue 系统^[24]，通过在车内前置 CCD 摄像头来采集驾驶数据，根据数据检测驾驶员的行车轨迹，若驾驶车道与原车道有偏离时系统将进行预警提示。

ASCI (Advanced Safety Concepts Inc) 公司开发了一种车载头部位置测量传感器设备，将传感器安装在驾驶员的座位上，通过实时获取驾驶人头部到传感器的位置，可以计算出驾驶员的头部在三维立体空间中的位置。通过实时监测跟踪头部位置，分析其运动的动态变化规律来推断驾驶人是否处于疲劳状态^[25]。

欧盟等汽车公司资助的 AWAKE 项目是一个基于驾驶行为特征的综合监测系统^[26]。该系统使用了多特征参数，它融合驾驶员眼动特性和使用不同类型传感器采集的方向盘转向握力、方向盘转角、车道轨迹、加速度、踏板制动力等车辆信息来实现驾驶员疲劳状态等级评估。对于评估结果采用声学光学刺激及安全带抖动等方式对驾驶员进行预警提示。

美国 Electronic Safety Products 公司开发的汽车方向盘监控装置 S.A.M 通过在方向盘下方安装光电传感器，来检测方向盘的转动幅度、方向和速度等运动情况，若方向盘持续四秒没有产生相对运动，系统则发出警报提示驾驶员，直至方向盘回归正常运动状态为止^[27]。

Mercedes Benz 配备的注意力辅助系统，使用高灵敏度传感器监控车辆加速踏板力、制动踏板力、方向盘角度与转向速度和各种外部环境参数，这些参数用于对驾驶员的驾驶行为进行科学分析。每次在行车的前几分钟内系统会分析这些驾驶行为参数并存档，在后面的行车过程中不断将当前传感器获取数据与存档数据进行比较分析，当检测到驾驶员处于疲劳状态时，系统会发出指令进行预警^[28]。

日本丰田汽车公司在 2008 年发布了一款疲劳驾驶检测预警系统^[29]。该系统通过在驾驶员座位前方安装摄像头来监控其眼部状态，计算眼睛上下眼睑的开合度是否达到系统设定的阈值，一旦达到阈值则检测为驾驶员在打盹，此时系统则会发出警报提醒，自动进行刹车减速以减少交通事故的发生。

1.3.2 国内研究现状

数十年来，由于我国道路交通事故起数较多，其中由疲劳驾驶造成的事故占比较高，因而疲劳驾驶检测技术近年来备受关注，但因国内相关技术研究发展时期较短，其研究相较于国外还在初期研究阶段，以下是我国的一些研究现状。

(1)基于车辆检测。文献[30]使用 AR 模型分析时间序列并从固定阶模型中提取参数作为输入向量,通过斯坦福嗜睡量表将驾驶疲劳状态分为三级作为输出向量,然后使用 SVM 分类器模型,利用 CV 方法优化 SVM 参数。文献[31]使用样本熵的方法来分析方向盘旋转特性,发现样本熵在正常驾驶状态比较大,表明样本熵对于区分疲劳驾驶和正常驾驶是可行的。文献[32]通过收集驾驶员双手对方向盘的握力,引入小波变换以从时域和小波系数中提取疲劳相关特征。经过比较了 SVM、LDA 和 KNN 这三种分类器的性能后,发现方向盘握力检测可以有效地区分疲劳和清醒状态。文献[33]通过模拟驾驶获得车辆驾驶的加速器状态、车速等参数。在去噪之后,使用小波变换在频域中计算数据,通过进行小波分解提取特征向量,然后构造最小距离分类器来识别驾驶员的疲劳状态。

(2)基于驾驶员检测。文献[34]对 15 名 20-30 岁的健康男性采集 EEG 信号,实验发现在疲劳状态下从 EEG 信号中提取的能量特征明显有所上升,而提取的样本熵在疲劳状态下明显下降,根据这一特性可以判断驾驶员的状态。文献[35]计算瞳孔的高宽比来判断眼睛的睁闭状态,清醒状态下,瞳孔的水平长度与垂直长度相同。实验表明当瞳孔长宽比大于 20%时眼睛是睁开的,否则眼睛是闭合状态。戴诗琪等在定位到嘴部后,统计驾驶员单位时间内打哈欠的次数,一次哈欠记为嘴巴宽高比值大于阈值 0.6 的连续帧大于 15 帧,然后使用卷积神经网络识别人眼状态,提取驾驶员的眼部疲劳特征进行疲劳预警,平均准确率达 92%^[36]。

由清华大学和东南大学几位博士共同组建的南京研发中心联合南京远驱科技有限公司研发的疲劳预警系统 gogo850 平安行^[37],通过 PERCLOS 算法检测眼睛的开合程度来分析驾驶员的精神状态,若检测到驾驶员处于困倦状态则发出疲劳预警。系统采用非接触方式采集人脸红外图像,这使得产品在强光和弱光下都能对眼部进行识别,且无论是否戴眼镜都能保持其检测的准确性。另装置也有瞳孔识别功能,可以对有睁眼睡觉习惯的人进行瞌睡检测,系统很具实用性。

(3)基于信息融合检测。文献[38]选择 PERCLOS 的 P80 算法和眨眼频率作为眼部状态特征参数、车辆越车道线作为车辆特征参数。利用支持 SVM 数据融合算法,整合基于人眼状态和基于车道轨迹偏离的疲劳判别数据建立疲劳检测模型。实验结果表明基于数据融合模型较单一数据信息的疲劳检测方法更为精准。

1.4 本文组织结构与研究内容

疲劳驾驶识别研究领域现还存在很多难点,在不同的场景下如戴太阳镜、光照明暗等都会影响识别的准确率,其研究方法包括根据单一疲劳信息识别和多特征疲劳信息识别,本文主要根据多特征疲劳信息来进行疲劳驾驶识别。

本文的组织结构及各个章节的主要研究内容如下:

第 1 章主要阐述了关于疲劳驾驶算法研究的背景与意义,同时对国内外在疲

劳检测方面现的研究现状进行相关介绍，并对目前较为常用的疲劳驾驶检测方法的优缺点进行分析后明确本文的研究方法。

第 2 章主要介绍了人脸检测及跟踪算法。为了提高人脸检测的准确率，先对视频图像进行了预处理，我们比较分析了中值滤波、均值滤波、高斯滤波及自适应中值滤波后，使用了自适应中值滤波对图像进行平滑去噪，并使用基于动态阈值的光照均衡算法对光照明暗不一的图像进行光照补偿处理。然后简要概括了现有的人检测方法，并将基于 Harr-like 特征的 AdaBoost 算法与基于阈值肤色模型算法进行比较，实验发现 AdaBoost 算法对人脸检测的效果更好。最后介绍了判别尺度空间的的面部目标跟踪算法（DDST），实验表明该算法对面部跟踪效果较好。

第 3 章主要实现眼部与嘴部状态的识别。首先我们实现了基于级联回归树的人脸特征点定位，根据人脸特征点的位置，我们可以准确的定位到人眼及嘴部区域。①人眼睁闭状态识别。在定位到人眼后，我们提出了计算人眼纵横比和计算人眼二值图像黑色像素累积差值的方法识别人眼睁闭状态，为了提升识别精度，将计算的这两个特征值作为 SVM 的输入特征进行模型训练，使用训练的 SVM 分类器对人眼状态进行识别，实验表明该方法可以很好的识别人眼睁闭状态。②打哈欠检测。根据人脸定位的嘴部特征点位置，我们发现计算嘴部高宽比 MAR 可以判断嘴部闭合、讲话、打哈欠等状态，为了验证该方法的准确性及健壮性我们在 YawDD 打哈欠部分数据集上实验来寻找最优 MAR 阈值来识别打哈欠状态。

第 4 章主要是疲劳状态识别的实现。为了准确得到眨眼频率，首先在 ZJU 眨眼数据集上进行了眨眼检测，然后根据眼部特征睁眼频率、持续闭眼时间、眼睛闭合占比，嘴部特征打哈欠次数、打哈欠持续时间和点头频率建立基于多特征加权求和的疲劳等级识别模型，根据识别的疲劳值与对应的疲劳等级来判断驾驶者是否处于疲劳状态，若疲劳则及时进行相应的预警。为了验证系统性能，最后在 NTHU 驾驶员疲劳检测视频部分数据集上实验测试。

最后，对本文研究内容的相关工作进行总结，并对下一步的研究工作进行展望。

第 2 章 人脸检测及面部目标跟踪

为了提高后期眼部、嘴部检测定位及特征信息提取的准确率及减少其计算时间，利用图像处理和图像识别技术进行人脸检测，去除无关背景信息的干扰，缩小人眼及人嘴检测定位的范围是有必要的。准确的人脸检测是人眼、嘴部区域定位及疲劳状态等级识别的基础和关键。本文经过比较分析现有人脸检测方法后采用基于 Haar-like 特征的 Adaboost 人脸分类器算法对面部区域进行检测，为了不用对每一帧视频图像都进行人脸检测定位，在人脸检测的基础上采用 DSST 跟踪算法进行人脸实时跟踪。

2.1 图像预处理

因为图像在采集和存储过程中可能会受到各种噪声污染和不同程度的破坏，导致图像降质^[39]。而且在实际行车中可能由于各种因素的影响使采集的图像出现光照分布不均匀等现象，进而对后面的人脸检测、特征信息提取和疲劳状态识别产生不良影响。故在对视频图像进行检测识别和特征参数提取之前，需要先对视频序列图像进行图像预处理，以消除图像中对检测无用的信息，保持原始信息的完整性和增强有用信息的可检测性，保证后续对感兴趣区域检测识别的准确性。本节介绍图像滤波去噪和光照均衡技术，主要为提高人脸检测定位的准确性。

2.1.1 图像平滑去噪

视频图像去噪的最终目的是改善采集图像的质量，保留原始图像中携带的有用信息。通过滤波去噪技术可以有效的解决实际图像由于噪声干扰而导致图像质量下降的问题，增大信噪比，更好的让图像适合特定场景的应用。

图像滤波技术大致可分为两类：空间域滤波和频域滤波^[40]。空域滤波处理是在空间域直接对图像各像素的邻域进行运算处理，包括点处理和模板处理。其运算过程为在待处理的图像矩阵中逐点地移动模板，滤波器在该中心像素点的响应输出通过定义的滤波器系数与滤波模板扫过区域内的相应像素值的关系来计算；而频域滤波处理则是先对图像进行傅里叶变换转换到频率域，然后在频域进行滤波处理，最后将计算后的结果图像逆转换到空间域。这里主要介绍空间域滤波处理方法中的中值滤波、均值滤波、高斯滤波算法和提出的自适应中值滤波算法，并作相关比较分析。

(1)中值滤波。中值滤波是一种适用于去除椒盐噪声和斑点噪声的基于排序统计的非线性平滑去噪方法，它用中心点邻域内的中值来替换中心点的像素值，所

以那种像素值为 0 或 255 的孤立斑点很容易被消除掉。其原理如式(2.1)所示，其中 $S(x,y)$ 为点 (x,y) 处的 $N \times N$ 模板 (N 一般为奇数) 邻域内的像素点集， $Med()$ 表示中值函数，将以像素点 (x,y) 为中心的邻域中的各个像素点的灰度值由大到小进行排列，输出处于中间位置的像素灰度值， $f(x,y)$ 表示 (x,y) 处中值滤波处理后的新像素值。

$$f(x,y) = Med(S(x,y)) \quad (2.1)$$

(2)均值滤波。均值滤波算法利用事先设定好的模板遍历图像中的各个像素点，用中心像素点邻域内像素点的平均灰度值作为滤波处理后中心像素点的输出值。它是一种线性滤波，对去除高斯噪声很有效。其算法如式(2.2)所示，其中 S 为 $N \times N$ 阵列中点 (x,y) 处的所有邻域的像素点集， $f(i,j)$ 为点 (i,j) 处的灰度值， H 为模板运算系数矩阵， $f(x,y)$ 为图像平滑去噪后像素点 (x,y) 的值。

$$f(x,y) = H * \sum_{(i,j) \in S} f(i,j) \quad (2.2)$$

其中模板系数矩阵 H 代表着点 (x,y) 像素值所占的权值，权值越大则对中心点像素的影响就越大，常用的模板矩阵 H 如式(2.3)所示。

$$H = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (2.3)$$

(3)高斯滤波。高斯滤波器是一种线性滤波器，为了能够在有效的抑制噪声平滑图像的同时更好的保留图像细节，高斯滤波器的模板系数权值则随着像素点与模板中心点距离的增大而减小，这使得远离中心点的像素点权值较小。所以，高斯滤波器对图像的模糊程度较小，不会导致图像严重失真。高斯滤波器的模板是对二维高斯函数进行离散化而得到的高斯函数值作为模板的系数。二维高斯函数表达式为：

$$h(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.4)$$

其中 (x,y) 为点坐标， σ 为高斯分布的标准差。对于一个 $(2k+1) \times (2k+1)$ 阶矩阵模板，经过对式(2.4)离散化后，模板系数矩阵中各元素值的计算公式表示为：

$$h(i,j) = \frac{1}{2\pi\sigma^2} e^{-\frac{(i-k-1)^2+(j-k-1)^2}{2\sigma^2}} \quad (2.5)$$

标准差 σ 的取值对模板系数的生成很重要， σ 代表着数据的离散程度。如果 σ 取值较小，那么使得模板的中心系数远大于周围其他的系数，滤波操作变为中心值的点运算，这样对图像的平滑效果就不是很明显；反之，若 σ 取值较大，则比较类似均值模板，对图像的平滑效果比较明显。高斯滤波一般 σ 取值 0.8，采用 3×3

高斯模板系数矩阵，其系数模板表示为：

$$w = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (2.6)$$

(4)自适应中值滤波。自适应中值滤波与中值滤波不同的是，它在噪声密度较大时也可以同时兼顾去噪和保留图像的细节信息。自适应中值滤波可以根据预先设定好的模板动态改变滤波模板尺寸的大小，同时可以判断当前像素是不是噪声，如果是则用邻域中值替换当前像素值。其算法处理过程分两步：

步骤 A：令 $\begin{cases} A_1 = Z_{med} - Z_{min} \\ A_2 = Z_{med} - Z_{max} \end{cases}$ ，如果 $A_1 > 0$ 且 $A_2 < 0$ 则转到步骤 B，否则增大滤波模板尺寸，记增大后的模板尺寸为 S 。如果 $S \leq S_{max}$ 则重复步骤 A，否则令 $Z_{xy} = Z_{med}$ ，输出 Z_{xy} 。

步骤 B：令 $\begin{cases} B_1 = Z_{xy} - Z_{min} \\ B_2 = Z_{xy} - Z_{max} \end{cases}$ ，如果 $B_1 > 0$ 且 $B_2 < 0$ 则输出 Z_{xy} ，否则输出 Z_{med} 。

其中 S 为滤波模板矩阵尺寸，点 (x, y) 为滤波矩阵的中心点， S_{xy} 表示以点 (x, y) 为中心的滤波区域， S_{max} 为滤波窗口允许的最大窗口尺寸。 Z_{min} 为滤波窗口中最小像素值， Z_{max} 为滤波窗口中最大像素值， Z_{med} 为滤波窗口中像素值的中值， Z_{xy} 为点 (x, y) 处的像素值。

2.1.2 基于动态阈值的光照均衡处理

由于系统采集到的驾驶员图像色彩信息易受到光照亮度、光源颜色与位置等因素的影响，导致驾驶员视频图像受光分布不均匀。为了能从不能强度光照的人脸图像中准确检测到人脸并进行特征提取，我们需要先对图像进行光照均衡处理。比较常见的光照补偿算法有 GrayWorld 色彩均衡算法，直方图均衡化，参考白算法^[41]，基于对数、指数、幂次等非线性变换法。本文主要采用基于动态阈值的光照补偿算法根据(2.7)式将图像从 RGB 颜色空间转为 YCbCr 颜色空间对图像进行光照均衡处理^[42]。

$$\begin{cases} Y = 0.299 * R + 0.587 * G + 0.144 * B \\ Cb = 128 - 0.1687 * R - 0.3313 * G + 0.5 * B \\ Cr = 128 + 0.5 * R - 0.4187 * G - 0.0813 * B \end{cases} \quad (2.7)$$

处理过程分为基于动态阈值检测参考白点和图像像素调整两步。对于参考白点的选取：首先将图像根据适当的高宽比（块大小）分为 M 块区域，对每个块计算 Cb、Cr 的平均值 M_b 、 M_r ，然后由式(2.8)分别计算其的平均绝对差值 D_b 、 D_r 。

$$\begin{cases} D_b = \frac{1}{N} \sum_{i,j} |C_b(i,j) - M_b| \\ D_r = \frac{1}{N} \sum_{i,j} |C_r(i,j) - M_r| \end{cases} \quad (2.8)$$

其中 N 为图像块的像素点总数， $C_b(i,j)$ 和 $C_r(i,j)$ 为像素点 (i,j) 的 Cb 、 Cr 值（色度值）。对于每个区域块，如果 D_b 和 D_r 都过小则说明该块色彩分布较均匀，不需要再作处理。然后对每个需要处理的区域块的 M_b 、 M_r 、 D_b 、 D_r 求和取平均值后作为整幅图像的 M_b 、 M_r 、 D_b 、 D_r 值，满足关系式(2.9)的则为图像近白区域的像素点集。

$$\begin{cases} |C_b(i,j) - (M_b + D_b * \text{sign}(M_b))| < 1.5 * D_b \\ |C_r(i,j) - (1.5 * M_r + D_r * \text{sign}(M_r))| < 1.5 * D_r \end{cases} \quad (2.9)$$

基于亮度值（ Y 值），选择近白区域亮度值前 10% 的像素点作为参考白点。

对于图像的调整：为了使整个图像的亮度保持相同，可以通过参考白点的在 RGB 各通道的平均值和整个图像的最大亮度值（ Y 值的最大值）来得到每个通道的增益。计算公式如下：

$$\begin{cases} R_{gain} = \frac{Y_{max}}{R_{avg}} \\ G_{gain} = \frac{Y_{max}}{G_{avg}} \\ B_{gain} = \frac{Y_{max}}{B_{avg}} \end{cases} \quad (2.10)$$

其中 R_{avg} 、 G_{avg} 、 B_{avg} 为参考白点在 RGB 各通道的平均值， Y_{max} 为图像中像素点的最大亮度值。

对于图像中每个像素点的像素值通过以下方式进行调整：

$$\begin{cases} R' = R * R_{gain} \\ G' = G * G_{gain} \\ B' = B * B_{gain} \end{cases} \quad (2.11)$$

其中 R 、 G 、 B 为图像的原始像素值， R' 、 G' 、 B' 为图像调整后的像素值。

2.1.3 实验结果分析

(1) 图像平滑去噪

为了比较四种滤波方法对图像平滑去噪的效果，对测试图像分别加了强度为 0.1 的椒盐噪声和均值为 0.1、方差为 20 的高斯噪声，然后使用四种滤波方法分别对加椒盐噪声和高斯噪声图像进行平滑去噪处理，并对四种方法的处理结果进

行比较分析，对比图如图 2.1 所示。

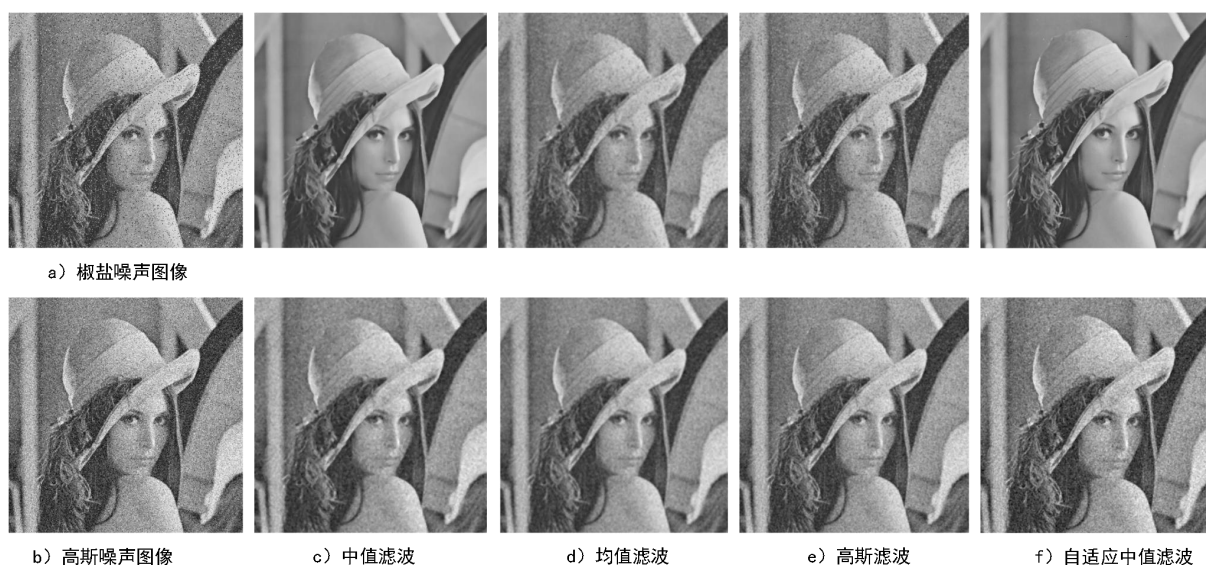


图 2.1 去噪结果对比图

经过对比发现，中值滤波和自适应中值滤波对椒盐噪声的去噪效果明显比其他两种方法好，而均值滤波对高斯噪声有较好的去除效果。

为了更客观的验证每种方法的去噪效果，分别计算每种滤波方法对图像处理前后的均方误差 MSE 和峰值信噪比 PSNR 及算法运行时间 T。MSE 和 PSNR 计算公式如式(2.12)、(2.13)所示，结果分析如表 2.1 和 2.2。

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [f(x,y) - f^*(x,y)]^2 \quad (2.12)$$

其中 $f(x,y)$ 表示大小为 $M \times N$ 的加噪图像， $f^*(x,y)$ 表示滤波去噪后的图像。

$$PSNR = 20 * \log_{10} \left(\frac{MAX}{\sqrt{MSE}} \right) \quad (2.13)$$

其中 MAX 为图像中可以采用的最大像素值。

表 2.1 椒盐噪声各算法指标对比

算法	评价指标		
	MSE	PSNR	T(s)
中值滤波	0.01702	19.4298	0.0139
均值滤波	0.01835	17.6062	0.0243
高斯滤波	0.01793	17.6546	0.0297
自适应中值滤波	0.00874	20.0532	0.7453

表 2.2 高斯噪声各算法指标对比

算法	评价指标		
	MSE	PSNR	T(s)
中值滤波	0.05790	22.5023	0.0135
均值滤波	0.05714	23.7135	0.0257
高斯滤波	0.04628	24.6644	0.0176
自适应中值滤波	0.07643	22.1808	0.6889

经过综合分析，前面三种滤波方法都能达到一定程度的去噪效果，但都需要提前设定好滤波模板，而且在对图像进行滤波去噪时，往往会使图像中的边缘细节和轮廓变得模糊，需要后期再对图像进行锐化处理来突出图像的边缘信息。因此为了能在平滑去噪的同时最大程度的保留图像细节信息，本文使用自适应滤波算法来提高图像去噪能力。

(2)光照均衡处理

根据 2.12 节介绍的光照处理算法进行实验，光照处理前后对比如图 2.2 所示。可以看出不论是光照过亮还是光照不足的图像经过动态阈值的光照均衡处理后，图像的亮度都得到了较好的改善。



图 2.2 光照处理前后对比图

2.2 基于 AdaBoost 算法的人脸检测

2.2.1 人脸检测方法分类

人脸检测技术的研究最早始于 1970 年代，不过彼时采用的方法比较简单，经过几十年的发展，人脸检测也逐渐成为人们的一个研究方向，且技术已日渐成熟。根据统计及分析大致可将人脸检测算法分为两大类：基于先验知识的检测方法和基于统计学习的检测方法。

(1) 基于先验知识的方法,主要是根据人脸器官的特征分布及其几何位置关系,利用轮廓特征、对称性、肤色、纹理特征等特征信息来检测图像中的人脸^[43]。基于先验知识的方法包括:基于肤色特征、基于模板匹配、基于人脸几何特征等人脸检测方法。

基于肤色特征建立模型来检测人脸是目前的主流方法。肤色模型的建立需要先将图像转换到 YCbCr 或 HSV 颜色空间,因为肤色信息在 Cb、Cr 及 H 分量上具有很好的类聚性,常用的模型主要有:阈值模型和高斯模型。该方法由于肤色与大多背景信息的颜色区别较大,所以其检测方法简单、直观、快速,但是其在复杂背景下易受类肤色背景干扰。基于模板匹配的方法通过预先定义一个标准的人脸模板,计算被测图像与模板的相似度来判断图像中的人脸。该类方法的难点在于怎么建立一个可以适应各种复杂环境的人脸模板,模板的建立存在很大的局限性,且模板匹配算法时间复杂度较大,准确性也不是很理想。基于面部几何特征的方法利用人脸面部的几何特征位置关系来定位人脸,该方法对人脸位姿的变化具有较强的适应能力。

(2) 基于统计学习的方法,采用统计学习的方法获取人脸模型的特征参数,对大量的人脸正样本与非人脸负样本进行训练并构造分类器,通过训练的分类来检测人脸^[44],它将人脸检测问题转化为二分类问题。它可以通过增加学习数据的数量来扩充检测模式的范围,提高检测系统的鲁棒性。该方法包括基于神经网络、基于支持向量机、基于 AdaBoost 算法、基于马尔可夫模型、基于线性子空间等算法。该方法需要进行大量样本训练,能在各种复杂背景环境下实时检测图像中的人脸,虽然算法计算量较大,但具有良好的鲁棒性和准确性,是目前的主流研究方法,其中 AdaBoost 算法是最为经典常用的算法。

经过综合分析上述检测方法后,本文使用基于 Haar-like 特征的 AdaBoost 分类器来进行人脸检测,下面主要就 AdaBoost 算法进行介绍与分析。

2.2.2 基于 Haar-like 特征的 AdaBoost 人脸检测算法

AdaBoost 是 Freund 和 Schapire 提出一个实际可用的自适应 Boosting 算法,是对传统 Boosting 算法的改进提升。后来,Viola 和 Jones^[45]在 AdaBoost 算法的基础上,对 AdaBoost 使用 Haar-like 特征和积分图法训练出的强分类器进行级联,然后使用级联分类器进行人脸检测。目前基于 Haar-like 特征的 AdaBoost 分类器算法是人脸检测的主流算法,其优点是可以稳定、快速、高效地检测人脸。

(1) Haar-like 特征

Haar-like 特征最早是由 Papageorgiou 等提出的原始矩形特征应用于人脸表示,它定义了三种类型四种基本特征结构。之后 Lienhart 和 Maydt^[46]在其提出的 Haar 特征基础上加入了旋转 45° 角对角矩形特征对 Haar-like 特征库进行了扩

展。目前最常使用的 Haar-like 特征是扩展后的 Haar-like 特征库，它可以分为三类（如图 2.3 所示）：线性特征、边缘特征、点特征（中心特征）、对角线特征。Harr-like 特征值是指白色矩形内所有像素灰度值之和与黑色矩形内所有像素灰度值之和的差值，它反映了图像的灰度变化情况。Haar-like 特征可以有效的提取图像的纹理特征，通过平移、缩放模板提取不同位置和尺度的特征值。

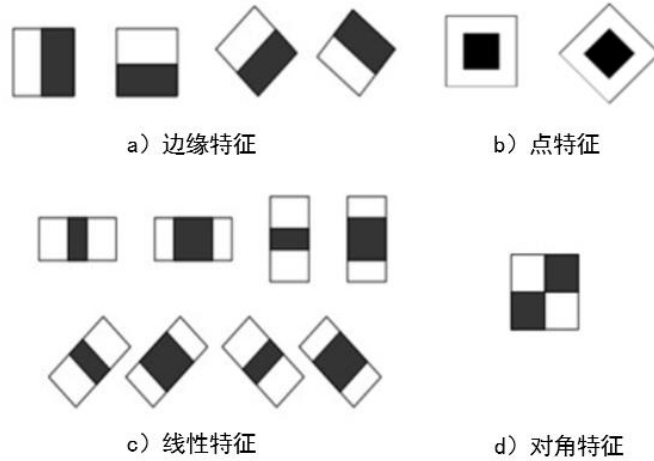


图 2.3 Haar-like 特征分类

(2) 积分图

由于矩形模板的类别、大小及位置的变化，使得即使检测窗口的尺寸很小也含有非常多的矩形特征值。如在确定了特征的形式之后，在大小为 24×24 的检测窗口内矩形特征数量也可以达到数十万个。由于特征数量巨大，所以解决特征的快速计算显得尤为重要。

积分图算法只需要遍历一次图像就能求出图像中任意矩形区域的像素和，很大程度的提高了图像特征值的计算效率。其主要思想：计算图像每个矩形区域从起点到各个点的像素之和，并将计算每个区域的值都作为一个元素保存数组中，当要后续需要计算某个区域的像素和时，可以直接使用数组索引得到目标区域的值，不用重新计算从而加快了计算。

对于积分图上任意一点 (i, j) 的值是指从灰度图像的左上角与当前点所围成的矩形区域内所有像素点灰度值之和。其积分图计算公式见式(2.14)：

$$I'(i, j) = \sum_{x=0}^{x \leq i} \sum_{y=0}^{y \leq j} I(x, y) \quad (2.14)$$

其中 $I(x, y)$ 为点 (x, y) 处的灰度值。积分图通过迭代运算还可以简化表示为下式：

$$I'(i, j) = I'(i, j-1) + I'(i-1, j) - I'(i-1, j-1) + I(i, j) \quad (2.15)$$

其中边界点 $I'(i, -1) = 0, \forall i$; $I'(-1, j) = 0, \forall j$; $I'(-1, -1) = 0$ 。

得到积分图后，矩形区域的特征值计算，只与此特征矩形的端点的积分图有关，所以不论特征矩形的尺度变换如何，计算特征值所消耗的时间都是固定的。计算两个矩阵区域像素和之差只需要计算特征区域端点的积分图来进行简单加减运行就可以了，这样可以实现快速计算任意矩形区域的特征值。

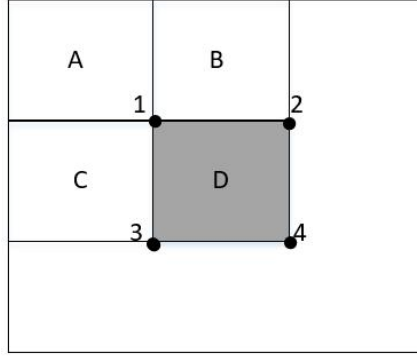


图 2.4 矩形 D 区域像素和计算示意图

以图 2.4 中的 D 区域来对积分图算法进行说明：

端点 1 的积分 $I'_1 = \text{Sum}(A)$ ；

端点 2 的积分 $I'_2 = \text{Sum}(A) + \text{Sum}(B)$ ；

端点 3 的积分 $I'_3 = \text{Sum}(A) + \text{Sum}(C)$ ；

端点 4 的积分 $I'_4 = \text{Sum}(A) + \text{Sum}(B) + \text{Sum}(C) + \text{Sum}(D)$ ；

其中 $\text{Sum}(N)$ 表示区域 N 的所有像素之和，那么区域 D 的所有像素之和为：

$$\text{Sum}(D) = I'_1 + I'_4 - I'_2 - I'_3 \quad (2.16)$$

(3) AdaBoost 算法原理

AdaBoost 算法是一种分类器算法，其算法原理为：利用积分图快速计算图像的 Haar-like 特征，通过训练迭代选取最优弱分类器，按照加权投票的方式将弱分类器构造为一个强分类器，再将训练得到的若干强分类器串联成一个级联结构的层叠分类器，从而提高分类器的检测速度和准确率。算法通过正负样本集的概率分布，来训练若干弱分类器，每循环一次则更新一次样本权重，经 T 次循环后，得到 T 个弱分类器，通过权重叠加，最终得到的强分类器。

给定一个训练数据集：

$$T = \{(x_i, y_i)\}, \quad i = 1, 2, \dots, N \quad (2.17)$$

其中 x_i 为训练的图像， y_i 属于 x_i 正确分类标签集 $\{-1, +1\}$ ，若 $y_i = 1$ 表示图像为正样本，即为人脸图像，若 $y_i = -1$ 则表示图像为负样本，即图像不包含人脸。对于样本的训练算法流程如下：

① 首先初始化训练数据的权值分布，每一个样本最开始的权值相同，令

$$D_1 = \{w_{11}, w_{12}, \dots, w_{1i}, \dots, w_{1N}\}, \quad w_{1i} = \frac{1}{N}, \quad i = 1, 2, \dots, N \quad (2.18)$$

其中 D_1 表示第一次迭代， w_{1i} 表示第一次迭代第 i 个样本的权值。

② 进行多轮迭代，令 $m=1,2,\dots,N$ ， m 为迭代的轮数。在具有权值分布的 D_m 数据上学习得到误差最低的弱分类器 $H_m(x):x \rightarrow \{-1,+1\}$ ，其分类误差率为：

$$\varepsilon_m = P(H_m(x_i) \neq y_i) = \sum_{i=1}^N w_{mi} I(H_m(x_i) \neq y_i) \quad (2.19)$$

③ 弱分类器每轮迭代的权重系数为：

$$\alpha_m = \frac{1}{2} \log \left(\frac{1 - \varepsilon_m}{\varepsilon_m} \right) \quad (2.20)$$

④ 更新训练集的权值分布：

$$D_{m+1} = \{w_{m+1,1}, w_{m+1,2}, \dots, w_{m+1,i}, \dots, w_{m+1,N}\} \quad (2.21)$$

$$w_{m+1,i} = \frac{w_{mi}}{Z_m} \times \begin{cases} e^{-\alpha_m}, & H_m(x_i) = y_i \\ e^{\alpha_m}, & H_m(x_i) \neq y_i \end{cases}, i=1,2,\dots,N \quad (2.22)$$

其中 Z_m 为归一化因子， $Z_m = \sum_{i=1}^N w_{mi} e^{(-\alpha_m y_i H_m(x_i))}$

⑤ 经过不断迭代，组合各个弱分类器最终得到强分类器：

$$G(x) = \text{sign} \left(\sum_{m=1}^N (\alpha_m H_m(x)) \right) \quad (2.23)$$

2.2.3 实验结果分析

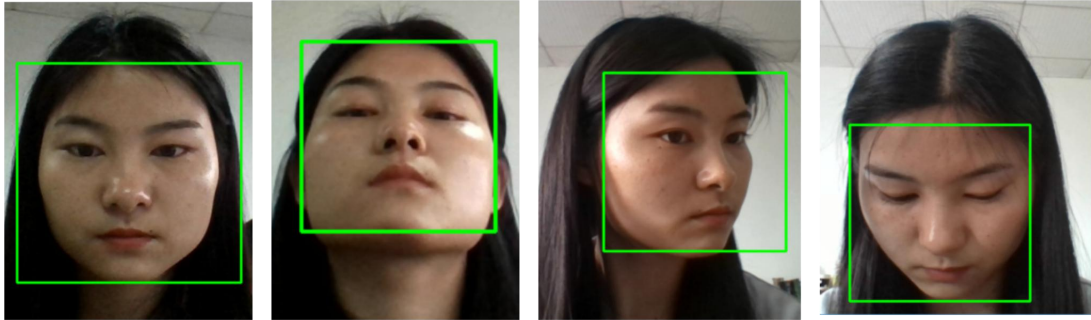


图 2.5 基于 AdaBoost 算法不同角度人脸检测结果

由于开源库 OpenCV 中封装好了基于 haar 特征的 AdaBoost 人脸检测算法，本文利用 OpenCV 中自带训练好的 haarcascade_frontalface_default.xml 分类器文件进行人脸检测，CascadeClassifier 为 OpenCV 定义的级联分类器类，其中封装了多尺度检测的方法，该方法输入待检测图像，通过加载检测人脸的 xml 分类器文件对待测图像进行人脸检测，输出可能的人脸区域矩形框。图 2.5 为对不同角度人脸检测的结果，表 2.3 为在有类肤色背景干扰和无干扰的情况下，使用 AdaBoost 算法与基于阈值肤色模型进行人脸检测的准确率对比。

表 2.3 AdaBoost 算法与阈值肤色模型人脸检测准确率对比

算法名 干扰情况	阈值肤色模型准确率	AdaBoost 算法准确率
类肤色干扰	93.4%	98.7%
无干扰	96.3%	99.2%

经对上图表的分析，发现当视频图像存在类肤色背景或者人体其他部位存在类肤色区域，这些类肤色的干扰使得基于阈值肤色模型的人脸检测算法检测范围不够精准，可能导致出现误检的情况。而 AdaBoost 主要根据 haar 特征进行人脸分类检测，能排除类肤色的干扰，计算效率和准确率都很高，不用做特征筛选可以快速检测出人脸，因此本文使用 AdaBoost 算法进行人脸检测。

2.3 基于高置信度的面部目标跟踪算法

考虑到在实际行车过程中，驾驶员的脸部位置变化范围较小，如果对于视频图像的每一帧都进行人脸检测定位，不仅会增加时间复杂度，而且不能充分利用连续帧之间的相互关系。因此，为了能在后续视频图像中更好地定位人脸，提高检测的准确性和鲁棒性，当在第一次检测到人脸后，便采用人脸跟踪算法实时跟踪检测到的人脸。

2.3.1 跟踪算法分类

目标跟踪（object tracking）是根据一组给定的图像序列，利用相邻两帧的区域匹配从图像序列中建立目标链，从而对图像中物体的运动形态进行分析。目前可大致将跟踪算法分为三大类：传统跟踪算法、基于深度学习的跟踪算法，基于相关滤波的跟踪算法^[47]。

(1)传统的跟踪算法。现今的目标跟踪算法中比较常见的传统跟踪算法包括：Kalman 滤波、粒子滤波、模板匹配算法、Meanshift 算法等。

Kalman 滤波是一种高效的迭代递归滤波器，利用反馈控制对物体的过去、现在及将来的状态进行分析与估计，通过目标上一帧的位置来预测当前帧的大致位置，利用最小均方误差来更新目标的当前状态^[48]。该算法复杂度小，对目标的位置估计较准确，但是易受遮挡、目标旋转、背景运动等因素的影响，且实时性较差，当目标运动较快速时可能出现延迟现象。

粒子滤波算法是一种基于贝叶斯的统计滤波器，其中心思想是提取目标的特征，随机获取一个粒子集合分布在搜索目标周围，计算其所处位置的颜色特征与目标的颜色特征的相似度，根据相似值对粒子做加权平均，然后由权重重新采样分配粒子数，权重大的粒子保留，通过不断迭代粒子的位置和权值来跟踪目标。

该算法很好的解决了由于遮挡而丢失跟踪目标的情况。但是由于该方法需要大量的样本数量才能很好地近似系统的后验概率密度，故算法的复杂度较高

Mean Shift 算法是基于鲁棒的核密度梯度估计迭代算法^[49]。其采用颜色直方图作为特征，通过不断迭代找搜索窗口中概率密度函数的最大值，从而在搜索窗口内找到质心的位置，然后以该质心作为目标新的位置。该算法实时性较好，计算量小，目标被挡或变形时，也能准确跟踪。但该算法没有对模板进行更新，当目标尺寸改变时，不能很好的跟踪。

模板匹配算法利用先验知识提取特征建立模板模板，然后将参考模板在待检测图像不断搜索进行匹配，计算它们之间的相似度，相似度最高的即为目标的位置。该算法简单且跟踪准确率高，计算速度快，抗噪性好。但是对建立的目标目标要高，不适用于旋转过大的目标跟踪。

(2)基于深度学习的跟踪算法。近年来深度学习方法在目标跟踪领域有不少成功应用，并逐渐在性能上超越传统方法，取得巨大的突破。深度学习在目标跟踪领域的其主要问题在于训练数据的缺失，深度模型需要对大量标注训练数据进行有效学习，而目标跟踪仅提供第一帧的边界框作为训练数据。

DLT 算法是第一个将深度网络运用于单目标跟踪的^[50]，该算法提出了“离线预训练加在线微调”的思路，很大程度的解决了基于深度学习跟踪中训练样本不足的问题。现有目标跟踪算法利用现有大规模分类数据集预训练的 CNN 分类网络来提取特征，如直接使用 VGG-Net 来获得目标的特征，之后再用观测模型进行分类获得跟踪结果。这种做法既避开了跟踪时直接训练大规模 CNN 样本不足的困境，也充分利用了深度特征强大的表征能力。基于深度学习的跟踪算法在跟踪精度和运行速度上有很大的提升，但对目标跟踪的实时性不够好。

(3)基于相关滤波的跟踪算法。该类跟踪方法由于其跟踪效果好、计算速度快且能实时跟踪而受到了众多研究者的关注。相关滤波器用相关性表示两个信号的相似度，其使用快速傅里叶变换来大幅度提升运算速度。该方法的基本思想为将输入的特征回归为高斯分布来训练滤波器，在后续连续帧中寻找预测分布中的响应最大值来得到跟踪目标的位置。Bolme 等^[51]是首个将相关滤波应用于跟踪领域的，其提出的最小输出平方误差和 (Minimum Output Sum of Squared Error, MOSSE) 跟踪算法引入了傅里叶变换将复杂的相似关系计算转换为点乘运算，大大提高了跟踪速度。随后在 MOSSE 算法基础上拓展了很多改进的跟踪算法，如引入核方法的化相关滤波器 (Kernel Correlation Filter, KCF) 利用循环矩阵计算使得跟踪速度惊人；加入尺度估计的相关滤波器 (Discriminative Scale Space Tracker, DSST) 可以处理运动目标尺度的变化^[52]。

由于传统的跟踪方法易受很多因素的影响不能很好的对目标跟踪，基于深度学习的跟踪方法需要大量的样本来进行学习，且对目标的跟踪实时性较差。而基

于相关滤波的方法在大幅度提升运算速度的同时既能实时跟踪还能保证跟踪精度。考虑到驾驶员在驾驶时基本不会有外物的遮挡，且要保证实时跟踪，因此本文使用 DSST 算法对检测到的人脸进行跟踪。

2.3.2 DSST 算法介绍

DSST 算法是在 MOSSE 算法的基础上改进的，虽然 MOSEE 算法在提升跟踪精度的同时降低了计算的复杂度，极大地改善了相关滤波器跟踪算法的性能，但在求解滤波器时，其输入是图像的灰度特征，模型所使用的特征维数太低，不能很好的反映目标的纹理、边缘等特性。且只估算了目标区域中心点在帧间的平移运动，而没有考虑目标在运动过程中的尺度变化，当目标尺度发生改变时不能很好的跟踪。

基于 MOSSE 算法的不足，文献[53]提出了一种三维尺度空间相关滤波器 translation-scale 联合跟踪方法。DSST 将原来的灰度特征替换为 HOG 特征，使得能够更好的对目标特征进行描述。另外为了更好的适应跟踪目标的尺度变化，增加了一个尺度相关滤波器，通过两个滤波器分别跟踪位置变化和尺度变化。二维位置滤波器（Translation Filter）用于评估目标位置变化，一维尺度滤波器（Scale Filter）用来进行目标尺度估计，三维联合位置和尺度滤波器 translation-scale 用于目标定位。这两个滤波器相对独立，因此可以使用不同的特征和特征计算方式进行训练及测试，其训练和测试原理如图 2.6 所示。

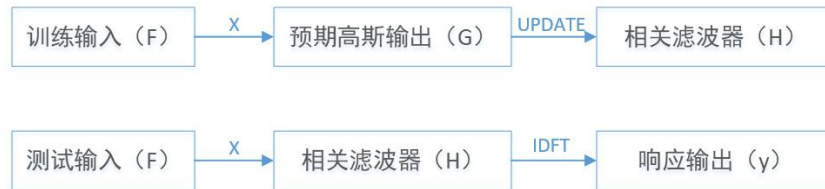


图 2.6 DSST 算法原理图

(1) 位置滤波器

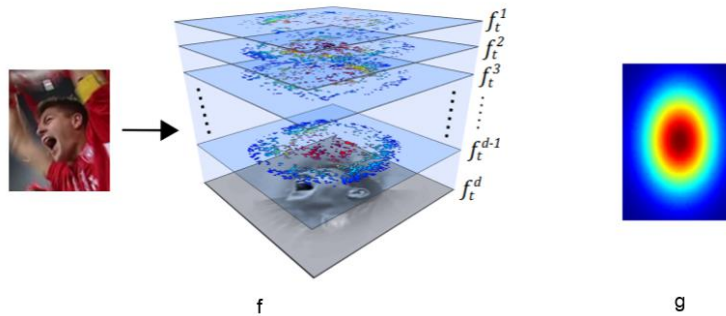


图 2.7 位置滤波器样例

① 滤波器训练

如图 2.7 所示, 采集一个尺寸为目标 2 倍大小的样本, 对其每个像素提取 d 维特征, 特征图可记作 f^l , $l=1,2,\dots,d$ 。为了构造最佳相关滤波器 h , 在不同特征维度 l 上最小化下面目标函数:

$$\varepsilon = \left\| g - \sum_{l=1}^d h^l \star f^l \right\|^2 + \lambda \sum_{l=1}^d \|h^l\|^2 \quad (2.24)$$

其中 \star 表示循环相关, l 表示特征的某一维度, λ 为正则项的系数, 设值为 0.01。 λ 项是为了避免滤波器频域参数求解过程中出现分母为零的情况, 同时也可以控制滤波器参数变化范围, λ 越小滤波器参数变化范围越大。预期相关输出 g 为具有参数化标准偏差的高斯函数, f^l , h^l , g 都具有相同的维度和大小。

对式(2.24)做傅里叶变换, 通过求偏导并令导数为 0, 得到滤波器

$$H^l = \frac{\overline{GF^l}}{\sum_{k=1}^d \overline{F^k F^k} + \lambda}, l=1,2,\dots,d \quad (2.25)$$

其中大写字母表示离散傅里叶变换 (DFT) 后的相应值, 即对 f 的每一个维度的特征做二维的 DFT 得到 F^l , 对 g 做二维的 DFT 得到 G 。

对于所有训练样本 $\{f_1, f_2, \dots, f_t\}$, 为了简化(2.25)式的计算, 分别更新滤波器 H_t^l 的分子 A_t^l 和分母 B_t^l , 其计算式如下:

$$\begin{aligned} A_t^l &= (1-\eta)A_{t-1}^l + \eta \overline{GF_t^l}, l=1,2,\dots,d \\ B_t &= (1-\eta)B_{t-1} + \eta \sum_{k=1}^d \overline{F_t^k F_t^k} \end{aligned} \quad (2.26)$$

其中 η 表示应学习率 ($\eta=0.025$), t 表示样本数。将 G 和 F 都代入上述等式, 可以得到滤波器模板 H 的值。式(2.25)简化后的计算式为:

$$H_t^l = \frac{A_t^l}{B_t} \quad (2.27)$$

② 目标位置估计

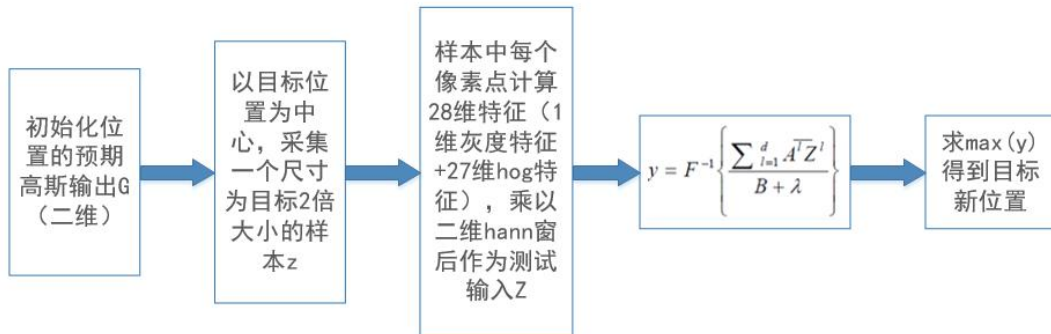


图 2.8 目标位置估计过程

对于第 t 帧图像的特征图 z_t ，同样求取 z 每一维度的二维的 DFT 得到 Z_t^l ，目标位置通过求逆向 DFT 得到最大相关滤波响应值 y_t 来确定：

$$y_t = \mathbf{F}^{-1} \left\{ \frac{\sum_{l=1}^d \overline{A_{t-1}^l} Z_t^l}{B_{t-1} + \lambda} \right\} \quad (2.28)$$

其中 A_{t-1}^l 和 B_{t-1}^l 是前一帧中更新的滤波器的分子和分母。

(2) 尺度滤波器

尺度滤波器训练过程中的模型更新及滤波响应求解尺度的过程与位置滤波器是一致的。

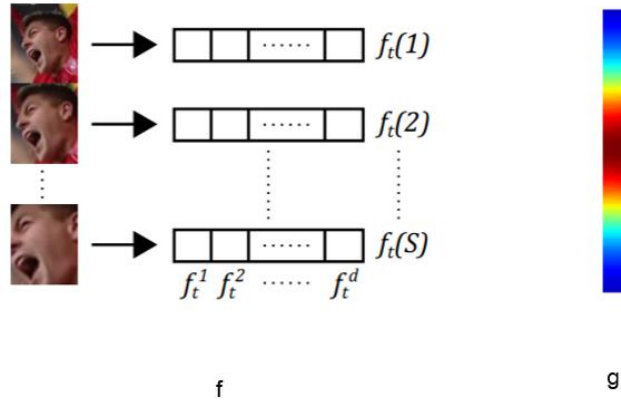


图 2.9 尺寸滤波器样例

① 滤波器训练

对应于图 2.9，以目标位置为中心进行尺度缩放采样，尺度选择原则如下：

$$a^n P \times a^n R, n \in \left\{ \left\lceil -\frac{S-1}{2} \right\rceil, \dots, \left\lfloor \frac{S-1}{2} \right\rfloor \right\} \quad (2.29)$$

其中 $P \times R$ 表示当前帧中的目标尺度， a 为放缩因子（ $a=1.02$ ）， S 为尺度滤波器的大小（ $s=33$ ）。

根据式(2.29)对目标图像进行放缩，选取 S 种不同尺度的样本，对每个样本提取 d 维 hog 特征，构成一个层数为 S 的金字塔。以该特征作为训练样本，每一个维度的特征 f^l 为一个 $1 \times S$ 的向量，对 f 的每一个维度的特征做一维的 DFT，得到 F^l ，对 g 做一维的 DFT 得到 G ， g 是高斯函数构造的输出响应大小为 $1 \times S$ ，根据式(2.27)得到相关滤波器 H ，用来预测输出尺度。

② 尺度估计

在新的一帧中，先利用二维的位置相关滤波器来确定目标新的候选位置，再利用一维的尺度相关滤波器以当前中心位置为中心点，获取 S 个不同尺度的候选块，分别提取 d 维特征组成新的特征图 z ，求取每一维度的 DFT 得到 Z^l 。然后根据式(2.28)求 y 的值， y 为 $1 \times S$ 维的向量， y 中最大值所对应的尺度为最终目标的

尺度。

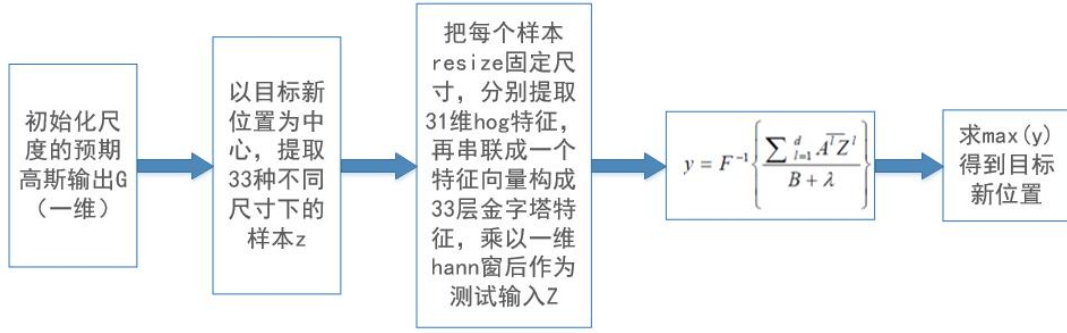


图 2.10 目标尺寸估计过程

2.3.3 高置信度更新策略

由于 DSST 算法需要人工标记初始帧的位置，且在目标被外物遮挡或目标丢失时不能很好的进行跟踪，因此需要利用目标检测期间跟踪结果的反馈来确定模型更新。响应图的峰值和波动可以在一定程度上揭示跟踪结果的置信度。因此引入了两个置信度指标，最大响应值 F_{\max} 和平均峰值相关能量(average peak-to correlation energy, APCE)。一般来说 F_{\max} 越大跟踪效果越好，APCE 反映响应图的波动程度和检测到的目标的置信水平^[54]。

$$APCE = \frac{|F_{\max} - F_{\min}|^2}{\text{mean}(\sum_{w,h} (F_{w,h} - F_{\min}))} \quad (2.30)$$

其中 F_{\max} 、 F_{\min} 表示响应最大值和最小值， $F_{w,h}$ 表示响应图 (w,h) 位置的值。当检测到的目标与正确的目标极其匹配时，响应图应该只有一个尖峰并且在所有其他区域中是平滑的，此时 APCE 将变得更大，相关峰越尖锐，定位精度越高。如果对象被遮挡或者丢失，APCE 将显著减小。若当前帧的 F_{\max} 和 APCE 都大于比率 β_1 ， β_2 ($\beta_1 = 0.7$ ， $\beta_2 = 0.45$) 时，当前帧中的跟踪结果被认为是高置信度的，此时才进行模型更新，否则则需要重新对当前帧进行人脸检测。

2.3.4 本文算法步骤

本文在 DSST 算法的基础上引入了高置信度更新策略，面部跟踪流程图如图 2.11 所示，具体算法步骤如下：

1) 将基于 AdaBoost 算法已检测到的人脸区域及大小作为目标的初始位置 P_1 和 S_1 。并对人脸区域进行位置滤波器及尺度滤波器训练得到位置模型 A_1^{trans} 、 B_1^{trans} 和尺度模型 A_1^{scale} 、 B_1^{scale} 。

2) 根据上一帧目标位置 P_{t-1} 和尺寸 S_{t-1} ，在当前帧 I_t 中采集一个尺寸为前一帧

目标 2 倍大小的样本 Z_t^{trans} ，利用 Z_t^{trans} 和上一帧位置模型 A_{t-1}^{trans} 、 B_{t-1}^{trans} ，根据式(2.28)计算 y_t^{trans} 的最大值得到目标的新位置 P_t 。

3) 根据确定的目标新位置，以当前新位置位置为中心点利用一维尺度相关滤波器，根据放缩规则获取 33 个不同尺度的候选样本，分别提取特得到 Z_t^{scale} 。然后利用 Z_t^{scale} 和尺度模型 A_{t-1}^{scale} 、 B_{t-1}^{scale} ，根据式(2.28)计算 $1 \times S$ 维向量 y 的值， y 中最大值所对应的尺度为最终目标的尺度。

4) 若最大响应值 F_{max} 和平均峰值相关能量 APCE 均满足更新策略条件，则在帧 I_t 中根据位置 P_t 和 S_t 提取特征 f_t^{trans} ， f_t^{scale} 并根据式(2.26)更新位置模型 A_t^{trans} 、 B_t^{trans} 和尺度模型 A_t^{scale} 、 B_t^{scale} 。否则则在帧 I_t 中重新进行人脸检测。

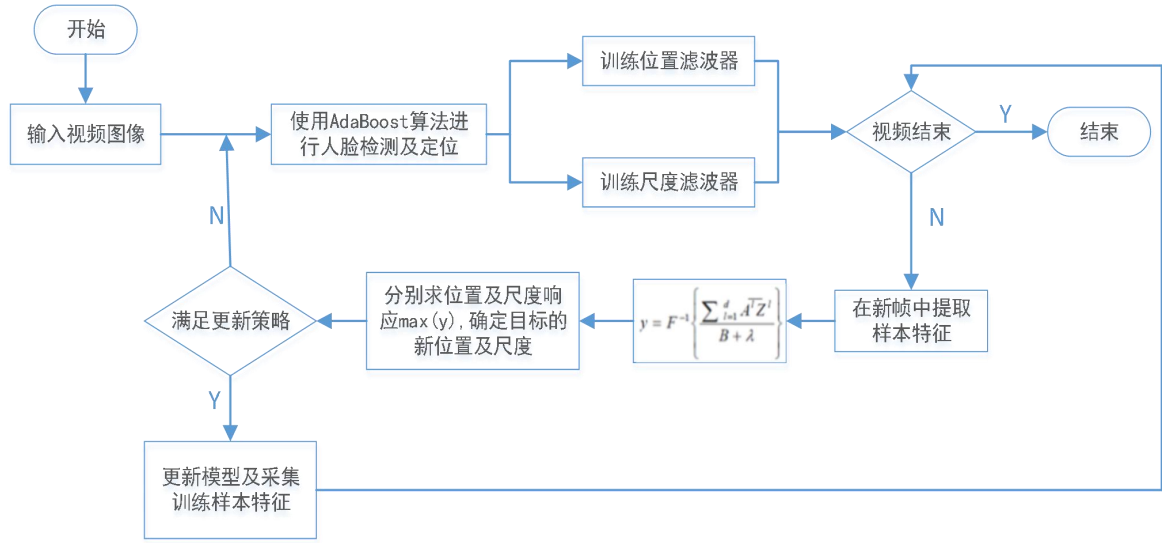


图 2.11 面部目标跟踪流程图

2.3.5 实验结果分析

DSST 算法在进行滤波器训练时需要提前指定目标区域，因此将基于 AdaBoost 算法已检测到的人脸作为跟踪的目标区域。使用 matlab 进行仿真，测试数据为 4 段视频，视频中包含面部与摄像头距离不断变化、面部左右移动、头部上下移动和头部微旋转等动作。将 AdaBoost 算法检测到的人脸位置存为文本数据与跟踪到的人脸位置进行比较。

实验评估以中心位置误差 (centre location error, CLE)、重叠率 (overlap precision, OP) 及跟踪速度每秒帧数 (frames per second, FPS) 这三个值为标准。中心位置误差为 DDST 算法预估的目标中心位置与真实目标中心位置的欧式距离，阈值设为 20 像素；重叠率是指预测的目标区域面积与真实目标区域面积的重叠程度，阈值设为 0.5。平均 FPS 值越大则跟踪速度越快，实时性越好；平均 CLE 值越小则跟踪精度越高，平均 OP 值越高说明跟踪成功率越高。

表 2.4 DSST 跟踪算法评估

视频序列	CLE	OP	FPS
视频 1	2.47	0.896	69.5
视频 2	5.93	0.783	56.1
视频 3	5.09	0.748	62.3
视频 4	4.34	0.771	65.6

本文对每个视频以三个值的平均值来进行评估，结果如上表，并在每个视频中选择连续 200 帧来绘制重叠率和中心位置误差曲线图，结果如图 2.12 和 2.13，可以看出该算法跟踪面部的跟踪精度和成功率很高。

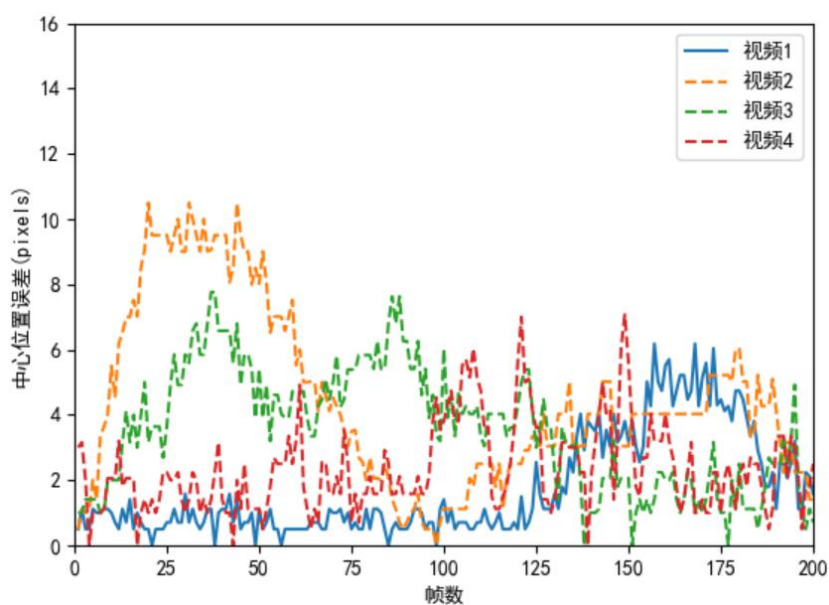


图 2.12 中心位置误差曲线图

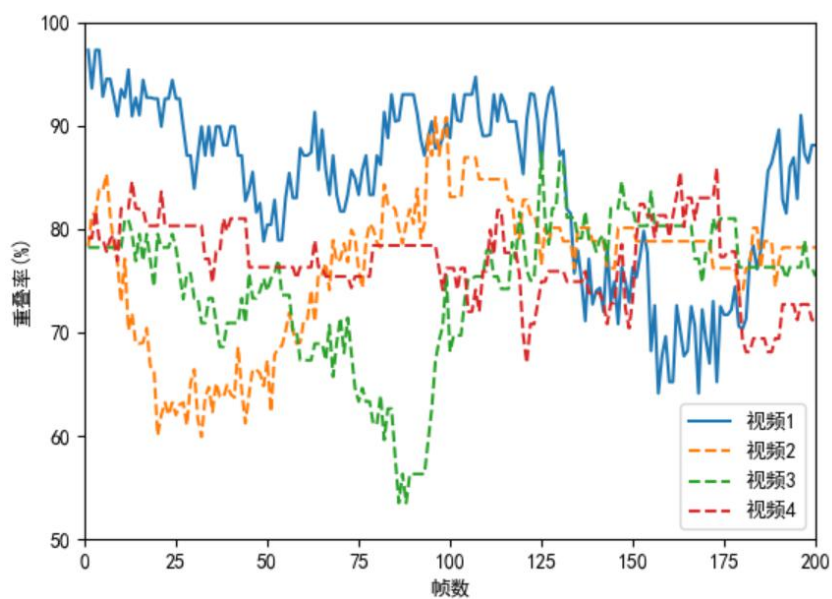


图 2.13 重叠率曲线图

2.4 本章小结

本章主要对人脸检测算法和面部目标跟踪算法进行了研究。为保证人脸检测的准确性率，本文选择自适应中值滤波算法对图像进行平滑去噪，很好的在去噪的同时又保留了图像的细节信息。然后采用基于动态阈值的方法进行光照补偿，解决了图像在不同亮度情况下光照分布不均问题。在图像预处理之后，利用基于 harr-like 特征的 AdaBoost 算法检测人脸，然后使用 DSST 目标跟踪算法来跟踪检测到的面部，并根据实际的跟踪情况根据高置信度更新策略及时的更新面部位置及尺度，充分利用连续帧之间的相互关系，提高了检测的精度。

第3章 基于SVM的睁闭眼状态识别及打哈欠状态识别

为了减少外界环境因素对人脸特征部位的检测及定位,本文在人眼及人嘴检测定位之前先使用了基于及级联回归树算法对人脸的特征点进行定位,然后在人脸特征点定位的基础上,根据特征点的位置快速准确的定位人眼及嘴部。最后提出了基于SVM的人眼睁闭状态识别的方法,以人眼纵横比和人眼二值图像黑色像素累积差作为SVM的输入参数进行模型训练,使用训练好的模型对睁闭眼状态进行识别,实验发现该方法识别的精度较高。

3.1 基于级联回归树的人脸特征点定位

人脸关键点定位的目的是在人脸检测的基础上,进一步确定脸部特征点的位置(如眼睛、眉毛、鼻子、嘴巴、脸部外轮廓)。考虑到在实际行车过程,驾驶者头部会有各种运动使得面部发生不同角度的旋转,致使对面部特征的检测不准确,为了提高后续对眼部、嘴部等特征区域定位的准确性,本文使用基于级联回归树面部特征点标定算法检测人脸特征。

3.1.1 人脸关键点定位算法分类

根据不同场景的应用,定义的人脸关键点的数量也各不相同,少则几个,多则上百个。该类算法的核心是如何精确计算出人脸中关键点的位置,涉及到的算法大致可以分为三种^[55]。

ASM和AAM经典算法。ASM算法由英国学者Tim Cootes在1995年提出,ASM采用参数化的采样形状来构成对象形状模型,并利用PCA方法建立描述形状的控制点的运动模型,最后利用一组参数组来控制形状控制点的位置变化从而逼近当前对象的形状,该方法只单纯利用对象的形状,因此准确率不高。Cootes等人在ASM算法的基础上提出AAM算法,基本思路是将人脸的纹理特征和各个特征点之间的位置进行约束结合。与ASM的不同之处是该算法不仅利用了对象的形状信息而且还利用了对象的纹理信息。

基于回归的算法。基于级联回归树(Ensemble of Regression Trees, ERT)算法^[56]的人脸关键点检测方法学习每个关键点的局部特征,然后将特征组合起来,使用线性回归检测关键点,该方法检测速度快,可精确估计面部特征点的位置。基于CPR(Cascaded Pose Regression)算法通过一系列回归器将一个指定的初始预测值逐步细化,每一个回归器依靠前一个回归器的输出来执行简单的图像操作,整个系统可自动的从训练样本中学习。

基于深度学习的算法。香港中文大学的 Tang Xiaoou 教授所在研究组提出了 CNN 的级联结构，首次将 CNN 应用到人脸关键点定位上。总体思想是由粗到细，实现 5 个人脸关键点的精确定位。网络结构分为 3 层：level 1、level 2、level 3。每层都包含多个独立的 CNN 模型，负责预测部分或全部关键点位置，在此基础上平均来得到该层最终的预测结果，该方法在关键点定位上取得了不错的效果。

3.1.2 基于级联回归树算法介绍

ERT 算法是 Kazemi 和 Sullivan 提出的基于级联回归树的人脸关键点定位算法。其选取了标记人脸 68 个关键特征点模型，如图 3.1 所示，并提出了一个基于梯度增强算法的通用框架，用于学习级联回归树，并使用级联回归树来直接从像素强度的稀疏子集估计面部的地标位置。该算法包括两个过程：训练建立模型和模型拟合。



图 3.1 人脸特征点模型

① 建立模型

该算法使用了两层回归来建立数学模型。第一层回归迭代式为：

$$\hat{S}^{(t+1)} = \hat{S}^{(t)} + r_t(I, \hat{S}^{(t)}) \quad (3.1)$$

其中 S 是形状向量， $S = (X_1^T, X_2^T, \dots, X_p^T)^T \in R^{2p}$ 表示图像 I 中所有 p 个面部标志的坐标， $X_i \in R^2$ 是图像 I 中第 i 个面部标志的坐标 (x, y) 。 $\hat{S}^{(t)}$ 为第 t 次迭代预测的特征点坐标集合形状矢量， $\hat{S}^{(t+1)}$ 为第 $t+1$ 次迭代预测的结果，每个回归器 $r_t(I, \hat{S}^{(t)})$ 在级联中预测来自图像的更新向量，其输入为当前的训练图片和形状向量，其输出则是对所有的关键点的位置更新量。在该层的级联回归器中，每经过一级级联回归器，就会对所有关键点位置进行一次更新得到更准确的位置。

第二层回归则是回归器 r_t 的内部迭代。假设有训练数据集 $\{(I_1, S_1), \dots, (I_n, S_n)\}$ ， n 为样本数， I_i 为人脸图像， S_i 为图像 I_i 对应的人脸关键点位置形状向量。为了学习级联中的回归函数 r_t ，根据训练数据创建人脸图像的三元组 $(I_{\pi_i}, \hat{S}_i^{(t)}, \Delta S_i^{(t)})$ ，其中 I_{π_i} 为数据集中人脸图像， $\hat{S}_i^{(t)}$ 为第一层级联回归第 t 次迭代预测关键点形状向

量, $\Delta S_i^{(t)}$ 是真实值和预测值差值。

$$\pi_i \in \{1, 2, \dots, n\} \quad (3.2)$$

$$\hat{S}_i^{(0)} \in \{S_1, \dots, S_n\} \setminus S_{\pi_i} \quad (3.3)$$

$$\Delta S_i^{(0)} = S_{\pi_i} - \hat{S}_i^{(0)} \quad (3.4)$$

$$\hat{S}_i^{(t+1)} = \hat{S}_i^{(t)} + r_t(I_{\pi_i}, \hat{S}_i^{(t)}) \quad (3.5)$$

$$\Delta S_i^{(t+1)} = S_{\pi_i} - \hat{S}_i^{(t+1)} \quad (3.6)$$

通过上式不断迭代该过程, 直到学习了 T 级回归 r_0, r_1, \dots, r_{t-1} 的级联。

对于训练数据 $\{(I_{\pi_i}, \hat{S}_i^{(t)}, \Delta S_i^{(t)})\}_{i=1}^N$, 学习率 $0 < \nu < 1$, 回归函数 r_t 通过使用具有平方误差损失总和的梯度树增强算法来进行学习, 具体算法如下:

1) 初始化函数 $f_k(I, \hat{S}^{(t)})$, 其中 $k = 1, \dots, K$:

$$f_0(I, \hat{S}^{(t)}) = \arg \min_{\gamma \in R^{2p}} \sum_{i=1}^N \|\Delta S_i^{(t)} - \gamma\| \quad (3.7)$$

2) 通过 N 次迭代拟合回归树 r_{ik} , 得到弱回归函数 $g_k(I, \hat{S}^{(t)})$, 其中 $i = 1, \dots, N$, r_{ik} 表达式如下:

$$r_{ik} = \Delta S_i^{(t)} - f_{k-1}(I_{\pi_i}, \hat{S}_i^{(t)}) \quad (3.8)$$

3) 根据得到的弱回归函数更新 $f_k(I, \hat{S}^{(t)})$:

$$f_k(I, \hat{S}^{(t)}) = f_{k-1}(I, \hat{S}^{(t)}) + \nu * g_k(I, \hat{S}^{(t)}) \quad (3.9)$$

4) 重复 2、3 步骤, 直至迭代 K 次得到 $f_K(I, \hat{S}^{(t)})$

5) 得到回归函数 $r_t(I, \hat{S}^{(t)}) = f_K(I, \hat{S}^{(t)})$

② 模型拟合

通过 K 次迭代得到回归模型, 模型拟合的具体步骤如下:

1) 初始化每幅人脸图像的特征点形状向量, 所有图像初始形状相同。

2) 建立特征池, 在特征池中随机选取两个点, 根据图像的特征点形状, 计算每幅图像在这两个点处的像素差。

3) 构造回归树。随机产生一个分裂阈值, 若图像的像素差值小于阈值则往左分裂, 否则往右分裂, 将所有图片按此方法分裂则图像分为左右两部分。重复该过程若干次, 通过最小化平方误差得到最优节点 θ^* , 目标函数如下:

$$E(Q, \theta) = \sum_{s \in \{l, r\}} \sum_{i \in Q_{\theta, s}} \|r_i - \mu_{\theta, s}\| \quad (3.10)$$

其中待选节点为 θ , l, s 分别代表左右子树, $\mu_{\theta, s}$ 代表按照当前划分产生的结果。

得到最优节点后保存下这两个特征点的坐标值和分裂阈值。然后每一个节点的分裂都重复该步骤，直到分裂到叶子节点。

4) 计算每个叶子节点的残差。计算每一个图像的当前形状和真实形状的差值，将在同一个叶子节点中的所有图像的差值和作平均，将该残差保存到叶子节点中。

5) 更新每张图像的形状。将当前形状 S 更新为当前形状加残差即 $(S, \Delta S)$ 。

6) 重复 2、3、4 过程，直至最后得到的特征点形状向量表示真实形状。

3.1.3 实验结果分析

Dlib 是一个跨平台的开源库，其中提供了许多关于机器学习、深度学习、图像处理等算法的实现。由于 Dlib 开源库实现了 ERT 算法，并在 iBUG 300-W 数据集上训练好了人脸关键点检测器，该检测器能够在任意人脸找到这 68 个特征点，因此本文使用 Dlib 开源库实现的算法来检测人脸关键点。实验结果如图 3.2 所示，从实验结果可以看到，ERT 算法对不同的面部表情和头部方向都有较好的稳健性，可以很好的实现不同角度的人脸特征点定位。

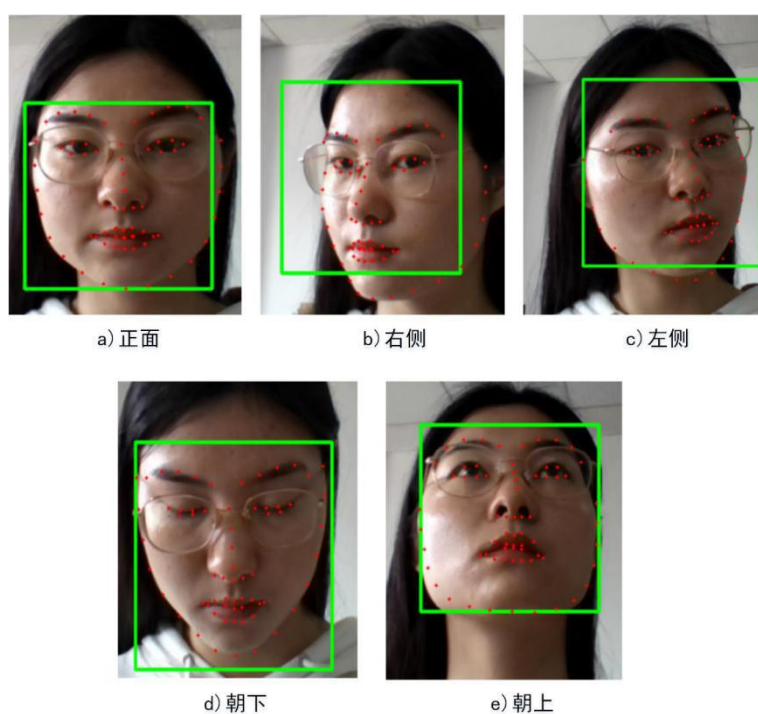


图 3.2 不同角度人脸特征点检测结果

3.2 基于 SVM 的睁闭眼状态识别

由 3.1 节中可以知道，人脸特征点定位可以较好的处理脸部的旋转，因此本文在人脸特征点定位的基础上使用特征点的位置来定位人眼。在检测到人眼后，

提出了一种基于 SVM 的人眼状态识别方法，该方法主要综合了人眼的两个特征：人眼纵横比和人眼二值图像区域黑色像素累积差值。

3.2.1 人眼检测定位

相比于人脸检测，想要更为精确的对人眼检测显得更为困难，虽然人眼也具有轮廓特征，但其检测易受眉毛的影响，有时还伴随着眨眼等动作。目前常用的人眼检测算法比较多，如：红眼效应的算法、Hough 变换检测算法、模板匹配算法、灰度投影算法、基于统计学习的算法等。各算法的优缺点分析如表 3.1 所示。

基于红眼效应的算法：利用了人眼瞳孔对两种红外光源反射程度的差异特性，计算同一时刻两种反射图像的差，反射相同的区域会被滤掉，留下反射不同的区域即为人眼区域。

基于 Hough 变换检测算法：Hough 变换将图像从图像空间转换到参数空间，以参数方程的形式描述人眼的边界轮廓，常用的有圆和椭圆两种轮廓形状。

基于模板匹配的算法：首先建立双眼的灰度模板，在检测的人脸图像中遍历搜索，与建立的模板匹配度最高的区域即为眼睛的位置。

基于灰度投影的算法：利用人眼区域比周围肤色区域灰度值低且灰度变化大的特性，通过计算水平和垂直方向的积分投影曲线的极值点来确定人眼的位置。

基于统计学习的算法：通过对大量的人眼图像和非人眼图像提取特征进行训练得到人眼分类器，最后通过分类器对待检测的图像进行检测来判断是否为人眼。

表 3.1 人眼检测算法对比

算法名称	优点	缺点
红眼效应法	简单快速，精度高	依赖红外设备，只有睁眼时才能检测到
Hough 变换法	有效减少因噪声的影响，鲁棒性高	计算量大，闭眼时无法检测到
模板匹配法	准确率高	计算量大，实时性差
灰度投影法	背景简单时，实时性好，准确率高	易受光照和眉毛的影响，不稳定
统计学习法	准确率高，鲁棒性好	计算量大，需要提前训练分类器，实时性较差

通过上表对各种人眼检测算法的分析，都有一定的局限，为了能简单快速的定位人眼，故本文采用在面部关键点检测的基础上，根据人眼特征点的位置进行眼部区域定位。

如图 3.3 中 a) 人脸特征点模型所示，根据图中关键点序号我们可以知道每个特征点的位置，如左眼的序号为 36-41，右眼的序号为 42-47。根据眼部特征点的序号，我们提取的左、右眼部区域如图 3.3 中 b) 所示的蓝色区域。其定位计算

规则如下：

$$\begin{cases} W = 1.6 * W_e \\ H = 3 * H_e \end{cases} \quad (3.11)$$

其中 W_e 为人眼特征点 36 和 39 水平距离， H_e 为特征点 37、41 和 38、40 垂直距离的平均值，而 W 和 H 为定位的眼部区域的宽和高。

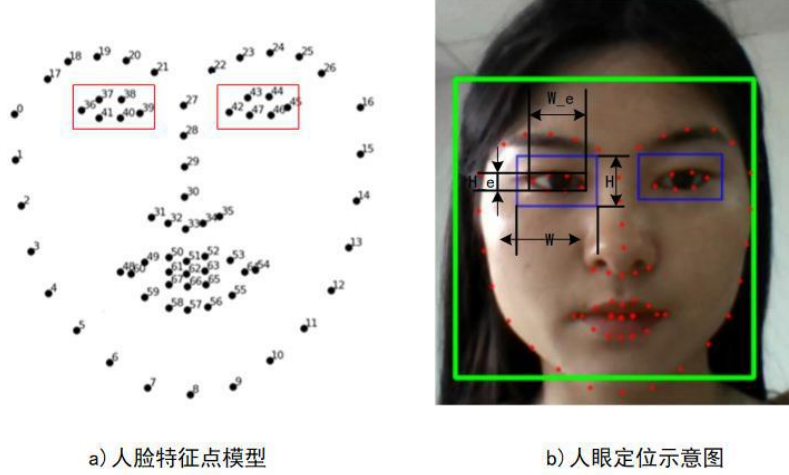


图 3.3 基于特征点的人眼定位示意图

3.2.2 基于人眼特征点计算人眼纵横比

为了能准确快速识别眼睛的睁闭状态，我们计算眼睛的纵横比（eye aspect ratio，EAR），睁眼时的纵横比在个体之间的差异基本很小，并且对于图像的均匀缩放和面部的旋转是完全不变的^[57]。如图 3.4 为左眼在睁开与闭合状态下检测到的 6 个关键点，眼睛纵横比计算式为：

$$EAR = \frac{\|P_2 - P_6\| + \|P_3 - P_5\|}{2\|P_1 - P_4\|} \quad (3.12)$$

其中分子表示眼睛垂直特征点之间的欧氏距离，分母为眼睛水平特征点之间的欧氏距离。

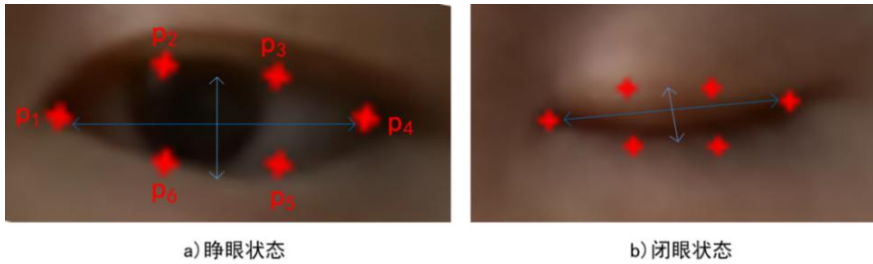


图 3.4 人眼六个关键点示意图

以左眼为例，根据这六个特征点，我们可以计算垂直关键点之间及水平关键

点之间的欧氏距离，两点的欧氏距离计算式为：

$$Dis(a,b)=\sqrt{(P_a \cdot x - P_b \cdot x)^2 + (P_a \cdot y - P_b \cdot y)^2} \quad (3.13)$$

其中 $P_a \cdot x$ 、 $P_a \cdot y$ 分别为点 a 的坐标 x 和 y 。则眼睛的水平欧氏距离距离可表示为：

$$Eye_h = Dis(P_1, P_4) \quad (3.14)$$

$$Eye_v = Mean(Dis(P_2, P_6), Dis(P_3, P_5)) \quad (3.15)$$

其中 $Mea(A, B)$ 表示取 A 和 B 的平均值。此时眼睛的纵横比可表示为：

$$EAR_{left} = \frac{Eye_v}{Eye_h} \quad (3.16)$$

根据式 (3.16)，我们计算视频图像连续 200 帧左、右眼的纵横比，得到的结果如图 3.5 所示：

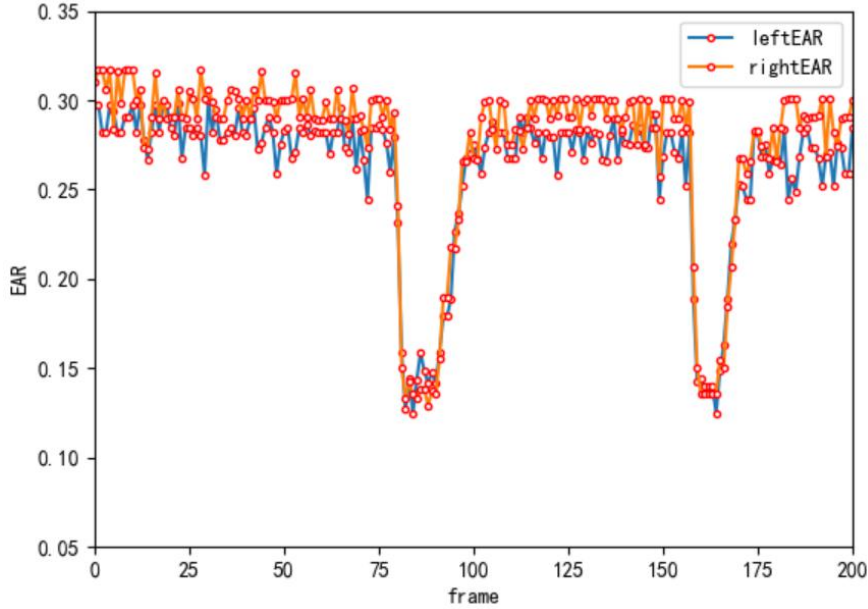


图 3.5 左、右眼 EAR 结果图

由上图可以看出，当眼睛睁开时 EAR 的变化很小，基本上是恒定的，但是在眼睛闭合时 EAR 值变小约接近零。根据结果图可知双眼基本同步闭合或者睁开，为更为精确的识别眼部状态，故取双眼 EAR 的平均值作为眼睛睁闭识别的特征：

$$EAR = Mean(EAR_{left}, EAR_{right}) \quad (3.17)$$

根据上式，进行眼部状态识别，计算双眼的 EAR 均值，结果如下图。当发生眨眼时，EAR 值迅速减小接近于 0，然后慢慢增加接近于正常情况下睁眼时的 EAR 值。根据这一现象，我们可以用 EAR 值作为识别睁闭眼状态的特征值，同时也可以根据 EAR 值进行眨眼检测。

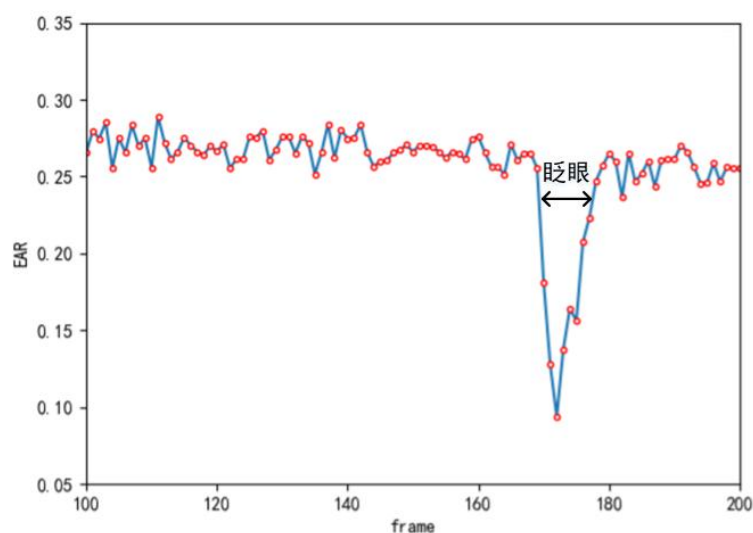



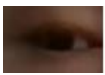


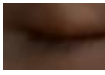




图 3.6 ERA 均值结果图

3.2.3 基于自适应阈值计算人眼黑色像素累积差值

根据 3.2.1 节中的方法定位人眼区域，然后选取局部自适应阈值算法对人眼图像进行二值化，经过形态学开操作和中值滤波处理后，能较好呈现眼睛轮廓及细节，下表为右眼不同情况下的二值图像。

表 3.2 不同情况下的人眼二值图像

人眼状态	人眼图像	局部自适应图像	人眼二值图像
右睁眼			
右半闭眼			
右闭眼			

当人眼闭合时，尽管可能会受睫毛和眼睑等暗区域的影响，但是最大暗部瞳孔区域不会出现。所以与睁眼相比，当眼睛闭合时，二值图像中的黑色像素的数量会急剧减少。但是，由于黑色像素的数量会随着人眼与摄像头之间距离的改变而改变。当距离变大时，在图像中眼睛区域被缩小，因此黑色像素的数量减少。图 3.7 为右眼睁闭过程中眼睛区域黑色像素数量，从图中可以看出，到第 57 帧我们可以设定阈值来区分睁闭眼，但是从第 109 帧开始当人眼远离摄像头时，不论睁眼还是闭眼状态，人眼的黑色像素数量都减少，此时根据阈值已经不能判断人眼睁闭状态了。

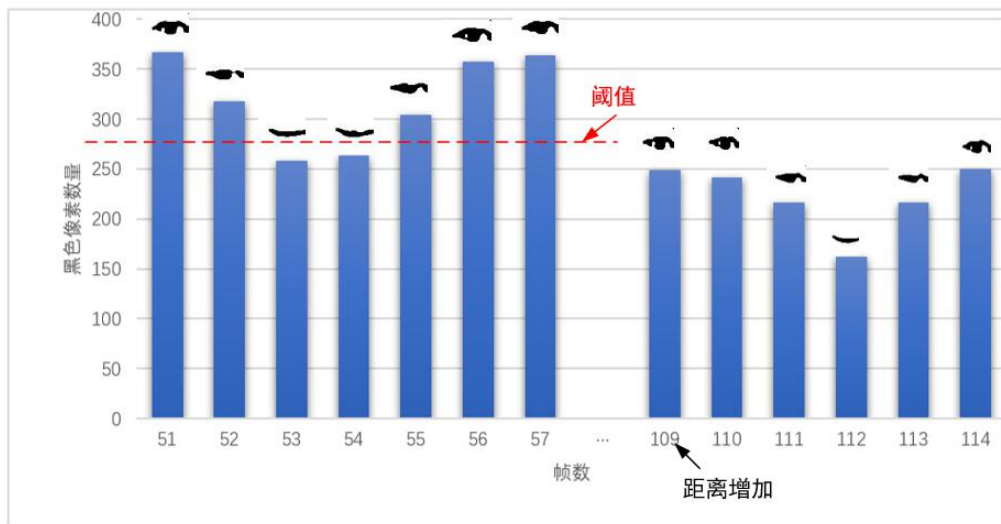


图 3.7 人眼睁闭眼过程黑色像素数量

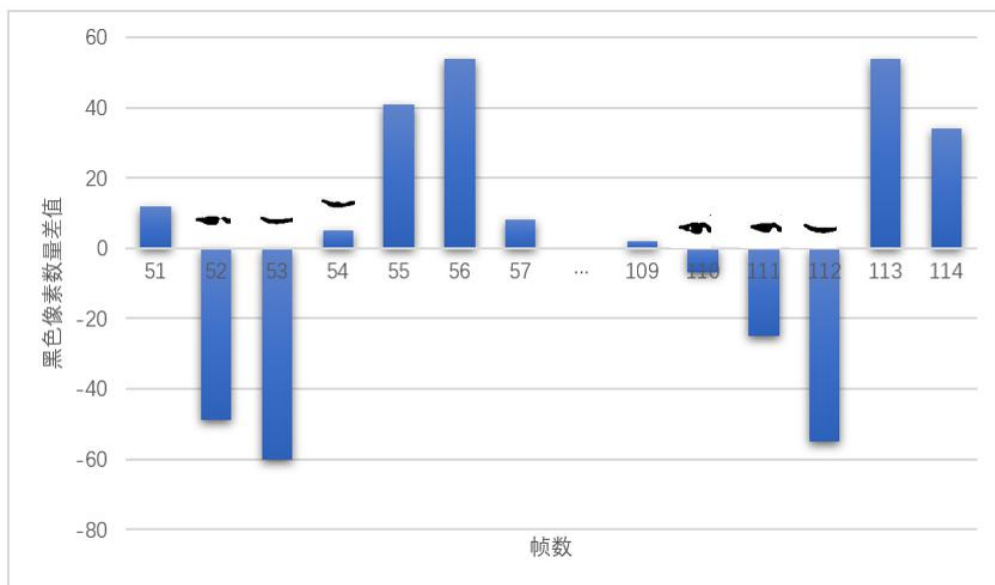


图 3.8 人眼连续两帧黑色像素数量差值

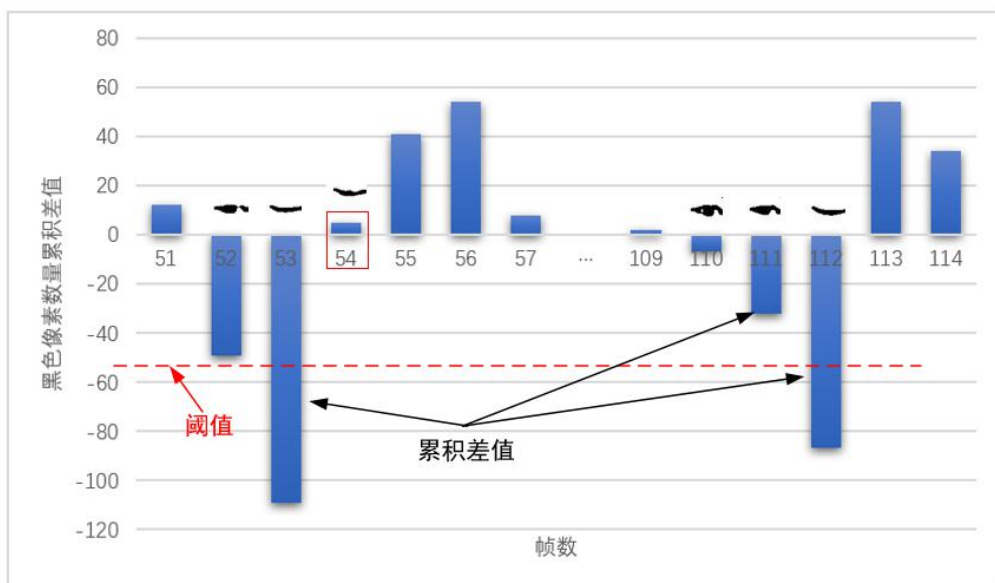


图 3.9 人眼黑色像素数量累积差值

为了减少人眼与摄像头距离因素的影响，我们将人眼图像归一为同一尺寸，计算连续两帧之间的黑色像素数量差，一般在两个以上的连续帧中可以观察到闭眼的动作，因此当差值满足多于两帧小于 0 时，累积连续的差值，设定累积差值阈值来判断睁闭状态。但是此种方法由图 3.8 和 3.9 可以看在第 54 帧时由于差值大于 0 并没有累积差值，从而导致其被错误的识别为睁眼状态。

因此为了解决这一问题，本文使用自适应阈值的方法累积差值^[58]。规定“状态 0”和“状态 1”两种状态，当人眼区域二值化图像的黑色像素的差值小于 0 时，从“状态 0”变为“状态 1”。在“状态 1”时，若差值小于阈值 $T(t)$ ，累积差值并保持状态不变；若差值大于阈值 $T(t)$ ，不累积差值且状态变为“状态 0”。

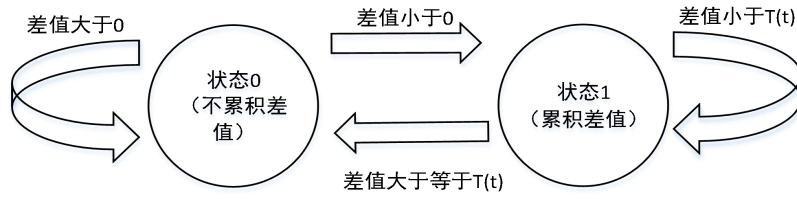


图 3.10 自适应阈值黑色像素累积差值算法流程

基于自适应阈值黑色像素累积差值计算方法如下：

$$F = \begin{cases} \Delta N(t), & (\text{状态}0) \\ \sum_{\Delta N(t) \leq T(t)} \Delta N(t), & (\text{状态}1) \end{cases} \quad (3.18)$$

$$T(t) = \alpha * |D(t-1)|, \alpha \in [0,1] \quad (3.19)$$

其中 $N(t)$ 是第 t 帧的黑色像素数量， $\Delta N(t)$ 为第 t 帧与第 $t-1$ 帧之间的差值， $D(t-1)$ 为“状态 1”中在第 $t-1$ 帧的累积差值， α 为 0 到 1 之间的恒定值，最优 α 值通过检测睁闭眼的准确性确定。

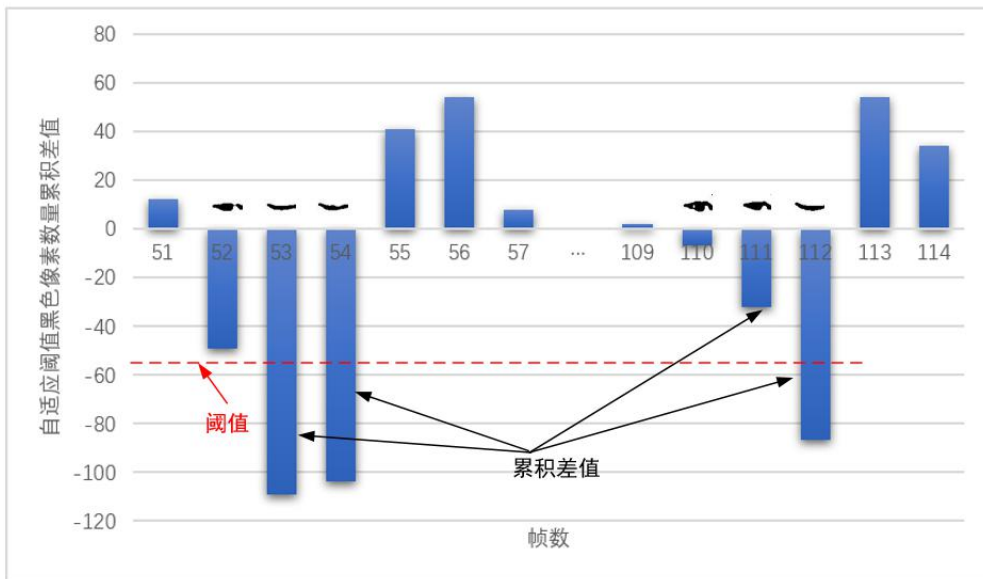


图 3.11 自适应阈值人眼黑色像素数量累积差值

通过根据在第 $t-1$ 帧的累积差值而改变自适应阈值 $T(t)$ 可以正确的将第 54 帧识别为闭眼。图 3.11 为使用自适应阈值计算人眼二值图像的黑色像素累积差值的结果图，可以看出该方法能较好的识别闭眼状态。

3.2.4 实验结果分析

为了能更为精准的识别人眼睁开闭状态，我们使用人眼纵横比和人眼黑色像素累积差值作为支持向量机 SVM 分类器的输入参数，使用训练好的分类器来识别图像中人眼的状态。

SVM 是一种解决二分类问题可以监督学习的机器学习算法，其本质就是找到距分类样本点间隔最大的分类超平面，使得训练的正负样本间隔最大。该算法可以用于数据的分类及回归分析，解决小样本、分线性和高维数学等问题^[59]。

本文使用 SVM 分类器来进行二分类，主要包括数据选取、数据处理、特征参数归一化、模型训练及测试五部分。

(1) 数据选取

我们从 ZJU 眨眼视频数据集^[60]的 80 个视频中选取睁眼样本 2000 张，闭眼样本 1000 张；从 NTHU 驾驶员疲劳检测视频数据集^[61]中选取睁眼样本 2000 张，闭眼样本 1000 张；同时自己采集睁眼样本 2000 张、闭眼样本 4000 张。总共采集睁、闭眼样本图像各 6000 张，其中有戴眼镜和不带眼镜的情况，每个样本中都包含人脸。

(2) 数据处理

先对每个样本进行人脸关键点定位，然后根据 3.2.2 的方法计算人眼纵横比，根据 3.2.3 的方法计算人眼黑色像素累积差值，即对每个样本提取两个特征值。

① 计算特征值 EAR

由于眼睛的纵横比 EAR 对于图像的均匀缩放和旋转是完全不变的，因此，对于每个样本，在定位到双眼关键点位置后，根据(3.17)式直接计算双眼纵横比的均值作为该样本的第一个特征值 F_1 。

② 计算人眼黑色像素累积差值

由于人眼区域黑色像素的数量会随着人眼与摄像头之间距离的改变而改变，因此对每个样本，根据(3.11)式定位到右眼区域后，将人眼区域放缩为同一尺寸，再计算其黑色像素值。对于不同实验者的样本数据，都以该实验者半睁眼状态的右眼黑色像素值作为第一帧的比较值，则该实验者第一帧的黑色像素累积差值为第一帧的黑色像素值减去半睁眼状态的黑色像素值，该实验者余下样本数据的黑色像素累积差值则根据(3.18)式及(3.19)式进行累积。将黑色像素累积差值作为每个样本数据的第二个特征值 F_2 。

我们将睁眼样本和闭眼样本分开处理。对于每个闭眼样本图像，根据上述方

法得到两个特征值后保存到相应的文本文件中，每行为一个样本数据，每列为一个特征值；对每个睁眼样本图像进行与闭眼样本相同的处理。

(3) 特征参数归一化

由于每个样本提取的两类特征参数的数值之间量纲有所差异，导致数值较小的特征参数在模型训练过程贡献较小，因此为平衡各特征参数在模型训练过程中的权重，需要对两类特征参数数据进行归一化处理：

$$y_i = \frac{2x_i - x_{\max} - x_{\min}}{x_{\max} - x_{\min}}, i = 1, 2, \dots, N \quad (3.20)$$

其中 y_i 为归一化后的结果值，归一化后的值在区间 $[-1, 1]$ 内； x_i 为原始特征值， x_{\max} 和 x_{\min} 分别为 x_i 中的最大值和最小值，训练样本数为 N 。

根据步骤(2)得到样本特征值文本数据后，读取两个文件中的特征值并存储在二维数组中，数组的每行为一个样本，每列为一个特征值，每个样本对应的类别标签保存到标签数组中。计算二维数组中每列的最大值 x_{\max} 及最小值 x_{\min} ，对数组中的每列，由式(3.20)计算该列每个特征值 x_i 归一化后的结果值 y_i ，当数组处理完毕后得到的二维数组值即为所有样本特征值归一化后的值。

(4) 模型训练及参数寻优

SVM 分类器可表示为：

$$f(x) = \text{sign}(\sum_{i=1}^N \alpha_i y_i K(x, x_i) + b) \quad (3.21)$$

N 为训练样本的数量； $y_i \in \{-1, 1\}$ 是训练样本的类别标签，1 表示闭眼，-1 表示睁眼； $K(x, x_i)$ 表示核函数；常数 b 是偏差项； α_i 通过求解具有线性约束的二次规划问题得到。

SVM 有四种核函数：线性核函数（LINEAR）、多项式核函数（POLY）、径向基核函（RBF）、SIGMOD 核函数。在进行分类器训练前，需要选择合适的核函数，由于当特征与分类标签之间的关系为非线性时，RBF 核函数能够处理此种情况，故本文采用 RBF 核函数进行模型训练。RBF 核函数有两个待定变量，用来控制损失函数的惩罚系数 C 和控制非线性问题变换到高维空间后的线形可分性核参数 γ ，这两个变量的选取对预测的精度有决定性作用。

为了寻找最优惩罚系数 C 和核变量 γ 提高模型预测的准确率，我们采用 K-CV 交叉验证法对参数 C 和 γ 进行寻优。从收集到的 12000 组特征值中选取 8000 组均分为 10 组，每次选 9 组作为训练集，剩余 1 组作为验证集。训练集与验证集的特征值则根据步骤(3)进行归一化处理，其类别标签保存在对应的类别标签数组中。经寻优发现，当参数 $C=2.04$ ， $\gamma=0.9$ 时，模型预测分类的效果较好。

(5) 实验检测

① 实验评估参数

为了对训练模型预测人眼睁闭状态性能进行评估，选取准确率(Accuracy)、精确率(Precision)与召回率(Recall)作为评估参数。对于测试集的每个样本，识别的结果可能会出现如下四种情况：

TP(True Positive): 表示测试样本预测为闭眼状态，实际也是闭眼状态。

FP(False Positive): 表示测试样本预测闭眼状态，实际为睁眼状态。

TN(True Negative): 表示测试样本预测为睁眼状态，实际为睁眼状态。

FN(False Negative): 表示测试样本预测为睁眼状态，实际为闭眼状态。

三个评估参数计算式分别如下：

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (3.22)$$

$$precision = \frac{TP}{TP + FP} \quad (3.23)$$

$$recall = \frac{TP}{TP + FN} \quad (3.24)$$

② 实验结果与分析

从样本数据中选择剩余的 4000 组数据对睁闭眼状态进行测试，其测试结果如下表。

表 3.3 睁闭眼状态检测结果

眼睛状态	识别状态		评估参数		
	睁眼	闭眼	准确率%	精确率%	召回率%
睁眼	1907	93	96.78%	95.48%	98.20%
闭眼	36	1964			

根据上表可以看出，提出的方法对睁闭眼状态识别的准确率较高，下表为使用不同算法的识别结果对比，文献[62]基于动态阈值计算肤色像素与非肤色像素的比例，若比例大于设定值则判断为闭眼。实验表明提出的特征融合训练分类器的方法对人眼睁闭状态的识别准确率高于单一特征的人眼状态识别方法，也比文献[62]提出的方法效果更好。

表 3.4 不同算法识别结果对比

算法	闭眼识别准确率%	睁眼识别准确率%	平均准确率%
基于纵横比方法	92.16%	91.33%	91.75%
基于黑色像素累积差方法	93.74%	92.61%	93.18%
基于 SVM 特征融合方法	98.20%	95.35%	96.78%
文献[62]	-	-	96.55%

3.3 基于嘴部纵横比的打哈欠状态识别

3.3.1 嘴部状态识别方法

人在困倦时会出现打哈欠、瞌睡等现象，当打哈欠时嘴部大幅度张开，可以明显区别于讲话状态，根据这一特点，可以使用下面几种方法来识别嘴部状态。

基于纵横比的方法：人在疲劳状态时会出现打哈欠的现象，此时嘴部由闭合状态到张开状态再到闭合状态。此过程中，嘴部出现大幅度张开，此时上下嘴唇之间的距离也大大增加，而嘴巴闭合时，两唇之间的距离几乎接近于零。由于当头部远离或者接近摄像头时，计算的两唇之间的距离会变小或变大，因此为了更准确的判别其状态，可以使用嘴部纵横比来判断嘴部的状态。

基于面积的方法：嘴部张开时如唱歌、讲话、打哈欠，嘴部的二值图像中的黑色像素数量会增加，通过计算嘴部二值图中的黑色像素所占面积与整个嘴部区域所占面积的比值可以判断嘴部张开或闭合状态。

基于近似张角的方法：通过嘴部定位方法得到嘴部最小外接矩形区域及该矩形的长、宽和中心坐标。通过计算该矩形四个顶点的坐标，可以得到嘴角坐标、上下嘴唇的中点坐标，最后通过这三个点的连线计算嘴部的张角来判断嘴部的状态。

3.3.2 基于纵横比的嘴部状态识别

嘴部检测定位的方法很多，由于在 3.1 节中通过人脸关键点定位我们已经知道嘴部特征点的位置序号为 48-67，因此我们可以根据特征点的序号来定位嘴部及识别其状态。

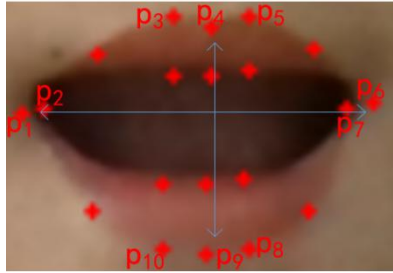


图 3.12 嘴部 10 个关键点示意图

本文主要通过计算嘴部纵横比（MAR）来判断嘴部状态，为了使 MAR 值更为精准，如图 3.12 所示，标记的 P_1-P_{10} 为用来计算 MAR 的 10 个特征点，欧氏距离的计算公式可参照(3.13)式。

$$MAR = \frac{Mean(Dis(P_3, P_{10}), Dis(P_4, P_9), Dis(P_5, P_8))}{Mean(Dis(P_1, P_6), Dis(P_2, P_7))} \quad (3.25)$$

我们知道，正常驾驶情况下，嘴部时处于闭合状态的；当与他人说话时，嘴

唇处于开合不断变化状态，且张开幅度不大；而当处于疲劳打哈欠状态时，嘴巴张开幅度很大且持续时间较长^[63]。为了判断嘴部状态如讲话、打哈欠等，使用基于高宽比的方法进行状态模拟，检测结果如图 3.13，由图可知当 $MAR \leq 0.4$ 时，此时嘴巴是闭合的；当 $0.4 < MAR \leq 0.8$ 时，为正常讲话状态；当 $MAR > 0.8$ 时处于打哈欠状态。根据上述分析，可以使用 MAR 作为特征来识别嘴部状态。

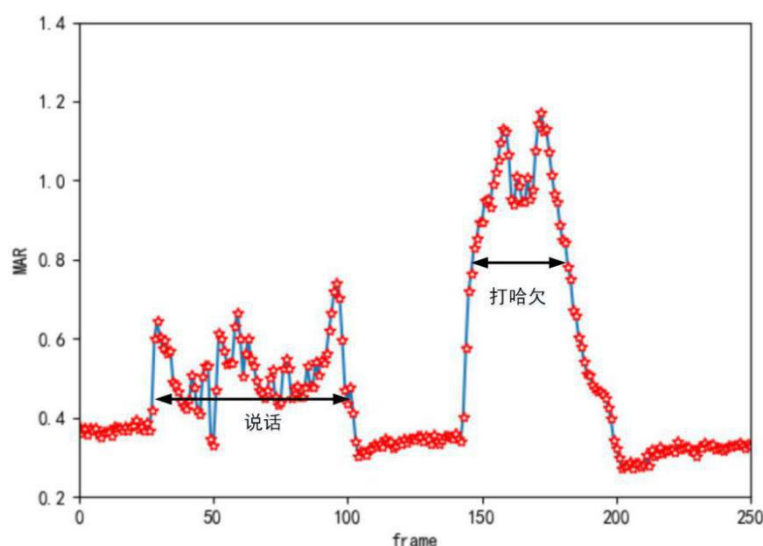


图 3.13 嘴部 MAR 检测结果

3.3.3 实验结果分析



图 3.14 YawDD 数据集部分样本数据

由 3.3.2 节可知，根据嘴部高宽比 MAR 可以判断嘴部闭合、讲话、打哈欠等状态，为了验证该算法的准确性及健壮性，我们在 YawDD 数据集^[64]上进一步进行实验。该数据集包含两个在不同照明条件下拍摄的具有各种面部特征的不同性别和种族的驾驶员视频集，可以用于打哈欠检测。

第一个数据集中提供了 322 个视频，包括不同种族的男性和女性驾驶员，每个驾驶员的视频分为 3 种不同的情况：驾驶时不说话、驾驶时说话或唱歌、驾驶时打呵欠。第二个数据集中提供了 13 个女驾驶员和 16 个男驾驶员视频，每个视频组合了驾驶员不说话、说话或唱歌、打呵欠等状态。数据集中部分样本展示如图 3.14。



图 3.15 打哈欠检测过程

为了正确识别打哈欠状态，我们在第二个数据集的 29 个视频上进行实验，不断改变 MAR 的阈值和连续帧数值，寻找最优高宽比 MAR 的阈值 M 及连续多少帧 MAR 值高于阈值 M 则判断为打哈欠的阈值 K 。找到最优阈值 $M = 0.7$ 和 $K = 15$ 后，在第一数据集上检测打哈欠，准确率约为 95.7%，因此在后续疲劳检测中，可以检测打哈欠状态作为疲劳特征。图 3.15 为一个打哈欠过程检测示意图，下图为阈值 M 及连续帧数 K 寻优的结果。

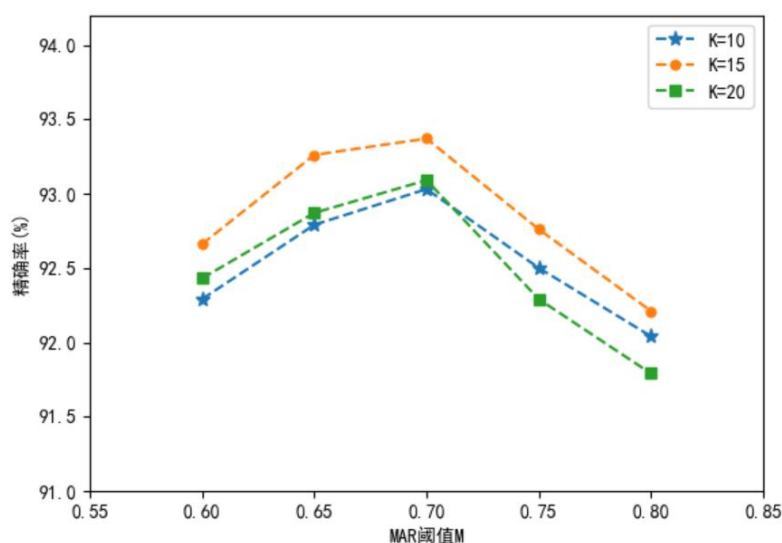


图 3.16 阈值寻优结果图

3.4 本章小结

本章主要介绍了关于人脸特征点定位、人眼状态识别及嘴部状态识别的相关方法。首先我们介绍了基于级联回归树的人脸特征点定位算法，实验证明使用该算法可以准确的对人脸的特征点进行定位。为了准确识别人眼的睁闭状态，提出了以人眼纵横比和人眼二值图像黑色像素累积差为特征的基于 SVM 人眼状态识别的方法，实验发现该方法识别的精度较高。同时我们在 YawDD 数据集上进行了打哈欠测试来验证基于纵横比的嘴部状态识别方法的准确性。

第 4 章 基于多特征加权值的疲劳状态识别

在上一章中主要介绍了眼部状态的识别，而在疲劳驾驶识别中，眼部特征是最能反映驾驶员是否疲劳的。根据眼部信息如眨眼频率、持续闭眼时、瞳孔运动状态、眼睛开闭程度等都能识别疲劳状态，但是根据单一的眼部特征存在一定的局限性，因此本文使用多个特征来判断疲劳状态。由于人在困倦时面部表情会有明显的变化，如会伴有眼睛闭合、不断打哈欠及连续点头等现象，本文主要根据眼部、嘴部及头部的状态来提取疲劳信息，然后综合这些疲劳信息建立疲劳状态识别模型，采用多特征加权和值来判断驾驶员的疲劳状态。主要提取的信息有闭眼帧数所占比（ECR）、眨眼频率（BF）、最长持续闭眼时间（MECT）、打哈欠频率（YF）及点头频率（NF）等。

4.1 疲劳信息提取

4.1.1 眼部疲劳信息提取

当人处于疲劳状态时，会出现眨眼频率增加，闭眼时间增长，打哈欠等现象，严重时甚至会出现打瞌睡的现象。根据研究发现，人在正常情况下每分钟眨眼 10 次到 25 次不等，眨眼一次眼睛闭合持续的时间约为 0.2s 左右。根据这一现象，本文选取最能表现疲劳状态的三个眼部指标基于 PERCLOS 准则的 ECR、MECT 和 BF 作为眼部疲劳特征参数。

(1) 基于 PERCLOS 准则的 ECR

PERCLOS 准则是公认用于疲劳驾驶检测最有效可靠的准则，它计算时间周期内人眼闭合时间占总时间的百分比^[65]。根据对人眼闭合的定义不同，该准则包含 3 个判定标准： EM 、 P_{70} 和 P_{80} 。其中 P_{80} 是最适合识别疲劳驾驶的，它表示眼睑遮住瞳孔面积超过 80% 的时间占时间的比例。由于实际检测时很难精确计算眼睑遮住瞳孔的面积，而在上章已经很好的实现了闭眼状态判断，因此本文通过时间周期内闭眼帧数占总帧数的百分比（Eye Close Ratio, ECR）作为眼部特征参数：

$$ecr = \frac{n}{N} \times 100\% \quad (4.1)$$

其中 n 为时间周期内闭眼帧数， N 为时间周期内总帧数。

(2) 最长持续闭眼时间

最长持续闭眼时间（Max Eye Close Time, MECT）：眼睛从完全闭合再到完全睁开持续的时间，即如图 4.1 中 t_2 到 t_4 所经历的时间。人在疲劳状态下，人眼闭

合时长往往会超过 1.5s。若视频速度为每秒 f 帧，时间周期内闭眼持续帧数为 K_c ，即则一个时间周期内持续闭眼时间为：

$$mect = K_c * \frac{1}{f} \quad (4.2)$$

若时间周期内持续闭眼时间超过阈值，则该特征参数视为疲劳状态。

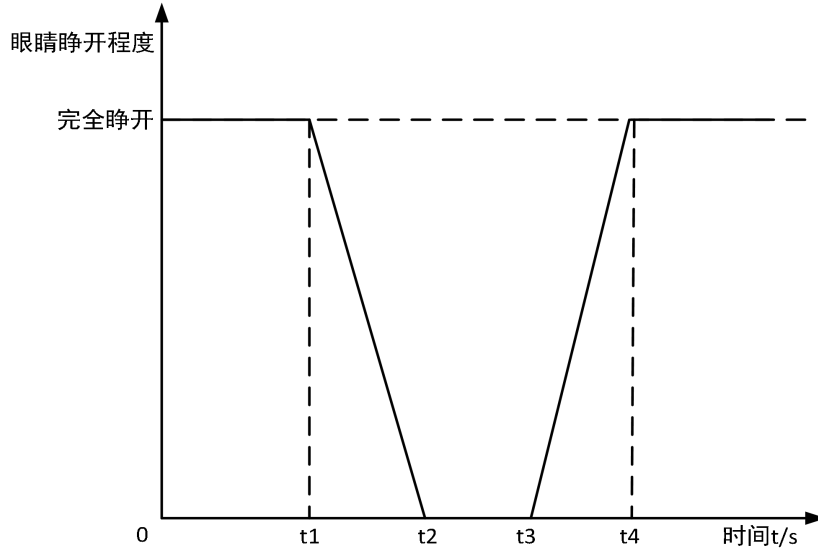


图 4.1 睁闭眼过程示意图

(3) 眨眼频率

眨眼频率（Blink Frequency, BF）：单位时间内的眨眼次数。一次眨眼时间为上图从 t_1 到 t_4 所经历的时间，人在清醒状态下，平均每分钟眨眼约为 10 - 25 次，疲劳时眨眼次数会增加，但走神分心或严重疲劳时，眨眼次数会减少。因此，可以统计时间周期内眨眼次数，若超出正常范围则该特征参数视为疲劳状态。

由 3.2.2 节可知根据 EAR 值可以进行眨眼检测。分析图 3.6 的 EAR 值计算结果，发现经历一次眨眼 EAR 值先减小直至接近于零，然后逐渐增大至正常睁眼状态值。我们以 E 作为 EAR 的阈值， K 为当 EAR 小于 E 时连续多少帧数记一次眨眼的阈值。当 EAR 小于阈值 E 时，眼睛开始闭合，当其值接近于正常睁眼状态值即大于 E 时，眼睛完全睁开，我们统计该过程中 $EAR < E$ 的连续帧数 F ，当 $EAR \geq E$ 时若 F 大于设定的连续帧数阈值 K 则记眨眼一次。

为了寻找最优阈值 E 和 K ，我们在 ZJU 眨眼数据集^[60]上进行实验，ZJU 中 80 个视频包含四种主题：未戴眼镜的正面视频、戴薄边框眼镜的正面视频、戴黑框眼镜的正面视图及未戴眼镜向上仰角的视频，每个主题 20 组视频，每个视频中眨眼次数一到六次不等，数据集中总共包含 255 次眨眼。图 4.2 为部分数据视频检测情况，图 4.3 为不同 EAR 阈值 E 值与连续视频帧数阈值 K 值对眨眼检测结果的精确率情况。

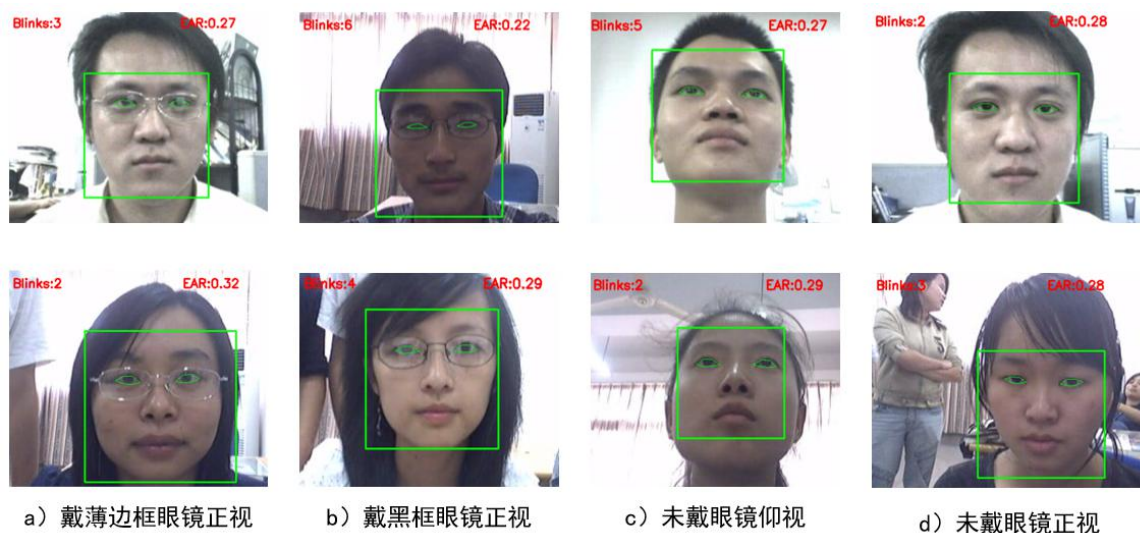


图 4.2 部分数据眨眼检测结果图

根据下图的结果，我们在提取眼睛疲劳参数眨眼频率时，选择计算 EAR 小于阈值 $E = 0.24$ 时的连续帧数，当 EAR 大于该阈值时若连续帧数也大于阈值 $K = 3$ ，则记一次眨眼，计算时间周期内眨眼次数即为眨眼频率。

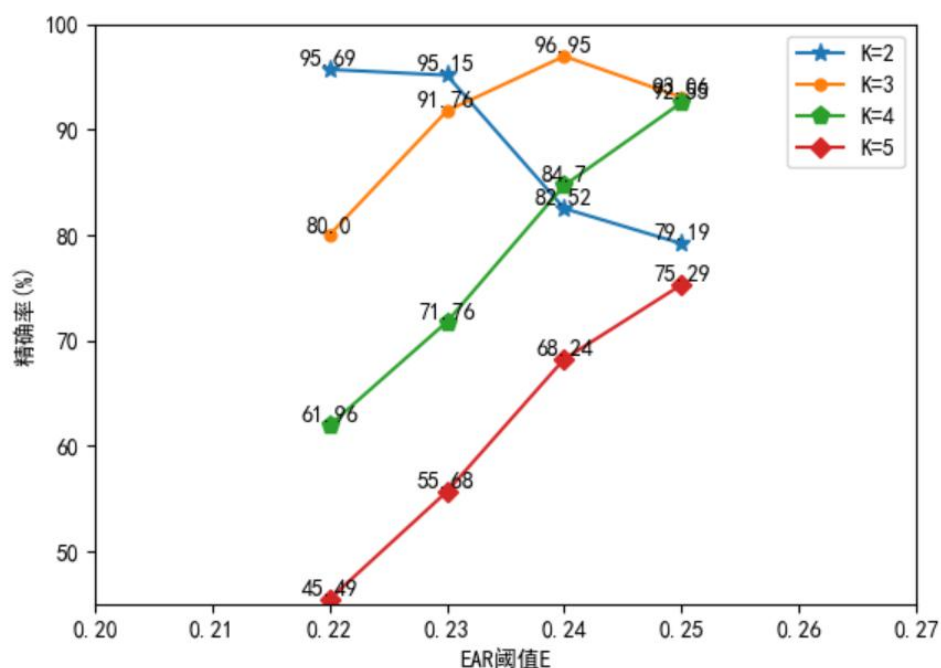


图 4.3 EAR 阈值、帧数 K 值寻优结果图

我们以 60s 为一个时间周期，对周期内人眼状态进行统计分析，得到眼部疲劳特征统计值。以 0 表示清醒状态，1 表示疲劳状态，记最长闭眼时间为 $mect$ ，闭眼帧数所占比为 ecr ，眨眼次数为 bf ，经实验及参考相关文献，得出三个眼部疲劳特征值之间的疲劳阈值如下表所示：

表 4.1 眼部疲劳状态取值条件

特征参数	MECT		ECR		BF	
状态取值	0	1	0	1	0	1
取值条件	$mect < 1.5$	$mect \geq 1.5$	$ecr < 0.20$	$ecr \geq 0.20$	$10 \leq bf \leq 25$	$bf < 10 \parallel bf > 25$

4.1.2 嘴部疲劳信息提取

当驾驶者处于困倦状态时，会接连不断的打哈欠，每次哈欠持续嘴部张开时间约 6 秒，此时则需要停车休息，不宜继续驾驶。根据该现象，我们可以检测时间周期内司机打哈欠的次数来评估其是否疲劳。由 3.3.2 节可以知道当嘴部纵横比 MAR 连续 15 帧大于 0.7 时，我们则记一次打哈欠。下图 t_1 到 t_4 的时间差即为一次哈欠时间，当嘴部张开程度超过阈值时，我们检测是否打哈欠。以 0 表示正常状态，1 表示疲劳状态，嘴部疲劳状态取值条件如下式：

$$YF = \begin{cases} 0, & yf < N \text{ and } yt < t \\ 1, & yf \geq N \text{ or } yt \geq t \end{cases} \quad (4.3)$$

其中 yf 表示打哈欠的次数， yt 为打一次哈欠持续的时间，取 $N = 3$ ， $t = 4s$ 。

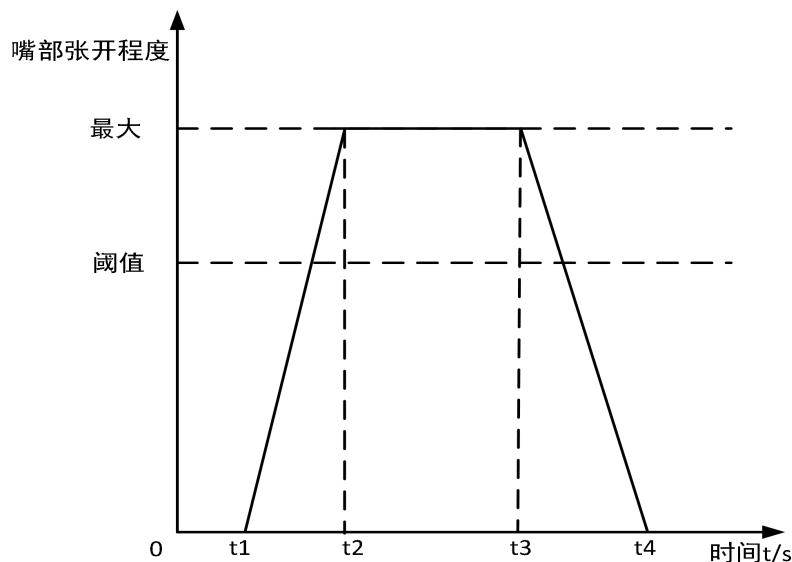


图 4.4 嘴部状态示意图

4.1.3 头部疲劳信息提取

人处于困倦状态时，反应会变迟缓，对头部的控制能力会有所下降，从而产生头部下垂现象。为了保持清醒，又不断抬头，因此会出现低头与仰头上下往复运动的现象。当驾驶员频繁出现该现象时，说明其已处于较疲劳状态，随时可能

出现交通事故，可见驾驶员驾驶过程中的点头频率的检测，是对头部运动分析的关键，也是疲劳驾驶检测的重要因素。当时间周期内点头频率超过某个阈值时，可以认为驾驶员处于疲劳状态。

依据定位到的眼部特征点位置信息，从实时性和准确性出发，本文提出了一种基于二维垂直方向的点头频率特征分析方法。取定位的双眼的中心点连线的中点作为头部位置检测点，根据该检测点在垂直方向上坐标 y 随时间的变化情况，计算时间周期内的点头频率。图 4.5 为驾驶员打瞌睡时 y 值与帧数之间的关系图：

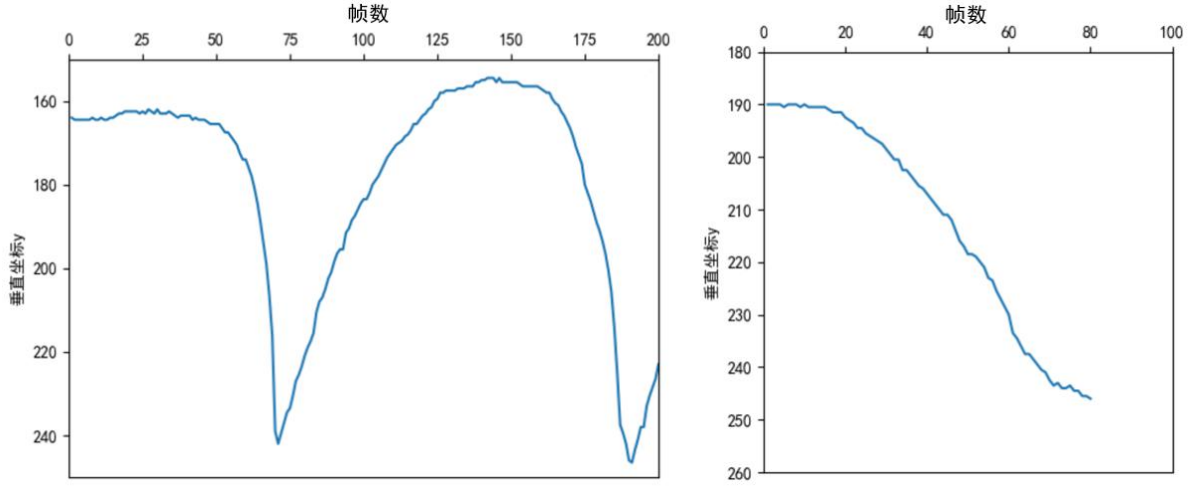


图 4.5 头部运动分析图

算法过程如下：当视频帧数较多时，图像可近似拟合为曲线，计算曲线极值点，极值点可将曲线分成许多单调的曲线。经实验得出，统计时间周期内单调递减段极小值点 y 值大于初始位置 50 像素的极值点个数，即为点头次数 nf ；若曲线没有极小值点，则判断曲线是否单调递减，若为单调递减，则点头次数 nf 为 1，否则为 0。NF 取值如式 4.4 所示：

$$NF = \begin{cases} 0, & nf < N \\ 1, & nf \geq N \end{cases} \quad (4.4)$$

若时间周期内点头次数 nf 大于某个阈值，则 NF 疲劳特征参数值为 1，否则为 0，经实验取 $N=8$ 疲劳状态检测准确率最高。

4.2 基于多特征加权和方法的疲劳状态识别

4.2.1 多特征加权和疲劳状态识别方法

只依据单一的疲劳特征来评估疲劳状态比将疲劳特征融合进行识别的准确率要低^[36]，因此，根据上述眼部、嘴部及头部的疲劳特征参数，本文提出基于多特征加权和的方法来识别疲劳等级，具体步骤如下：

(1)根据眼部、嘴部及头部的疲劳特征指标对疲劳判别的准确率各自取权重值，特征参数加权和计算下式：

$$F = \sum V_i * W_i, \sum W_i = 1, i = (ECR, MECT, BF, NF, YF) \quad (4.5)$$

其中 V_i 为特征参数的特征值， W_i 为其对应得权重值，五个特征参数值分别乘以其对应的权重值，求和后得到特征参数加权疲劳值 F 。

(2)通过模拟疲劳进行实验寻优，确定了眼部、嘴部及头部的五个疲劳特征指标各自的权重值，其对应特征的权重取值如下表所示：

表 4.2 各疲劳特征参数的权重取值

特征参数	眼 部						嘴 部		头 部	
	MECT		ECR		BF		YF		NF	
特征值 V_i	0	1	0	1	0	1	0	1	0	1
权重 W_i	0.2		0.1		0.2		0.2		0.3	

(3)根据疲劳参数加权后值的不同，将状态分为三个等级：清醒、疲劳、重度疲劳。综合特征参数的权重值和疲劳等级，将特征参数加权值与疲劳等级相对应，根据对应关系便能判断出驾驶员的驾驶状态。对应关系如表 4.3 所示：

表 4.3 疲劳值与疲劳等级对应关系表

疲劳等级	清醒状态	疲劳状态	
		疲劳	重度疲劳
特征参数加权值 F	$F < 0.3$	$0.3 \leq F < 0.7$	$F \geq 0.7$

算法伪代码如下表：

算法 1 多特征加权和的疲劳识别算法描述

Input: 疲劳判断标准 value

Output: 疲劳值 F

```

1: function judgeFatigueState(value)
2:    $V_{NF}$  = getHeadState(headLocation)
3:   if  $V_{NF} == value$  then
4:      $F += V_{NF} * W_{NF}$ 
5:   end if
6:    $V_{BF}$  = eyeBlinkDetection(eyeThreshold, eyeConsecFrames)
7:   if  $V_{BF} == value$  then
8:      $F += V_{BF} * W_{BF}$ 
9:   end if

```

(续上表)

算法 1 多特征加权求和的疲劳识别算法描述

```

10:    $V_{ECR}, V_{MECT} = \text{eyeStateRecognition}(Eigenvalues)$ 
11:   if  $V_{ECR} == value$  then
12:        $F += V_{ECR} * W_{ECR}$ 
13:   end if
14:   if  $V_{MECT} == value$  then
15:        $F += V_{MECT} * W_{MECT}$ 
16:   end if
17:    $V_{YF} = \text{MouthState}(mouthThreshold, emouthConsecFrames)$ 
18:   if  $V_{YF} == value$  then
19:        $F += V_{YF} * W_{YF}$ 
20:   end if
21:   return  $F$ 
22: end function

```

4.2.2 建立疲劳识别模型

根据上述提出的多特征加权求和疲劳识别方法，建立的疲劳识别模型如下图所示：

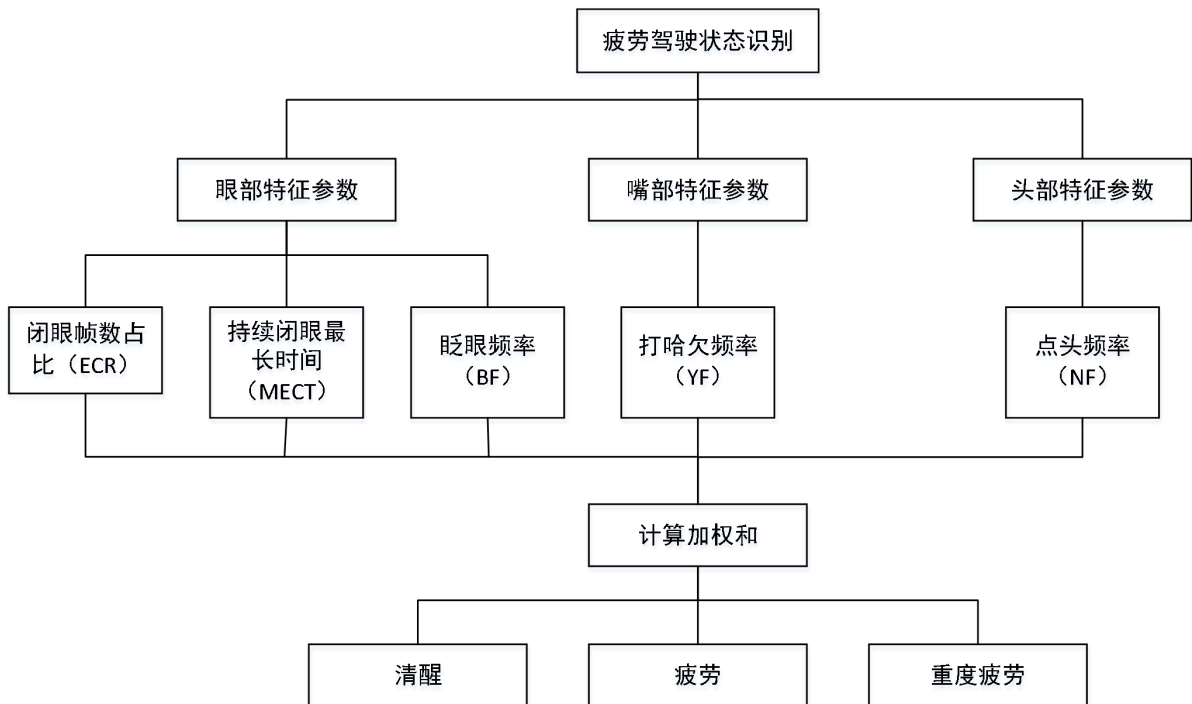


图 4.6 疲劳状态识别模型

4.2.3 整体实现流程

(1)基于多特征加权求和的疲劳状态实时识别实现的流程图如下图所示。

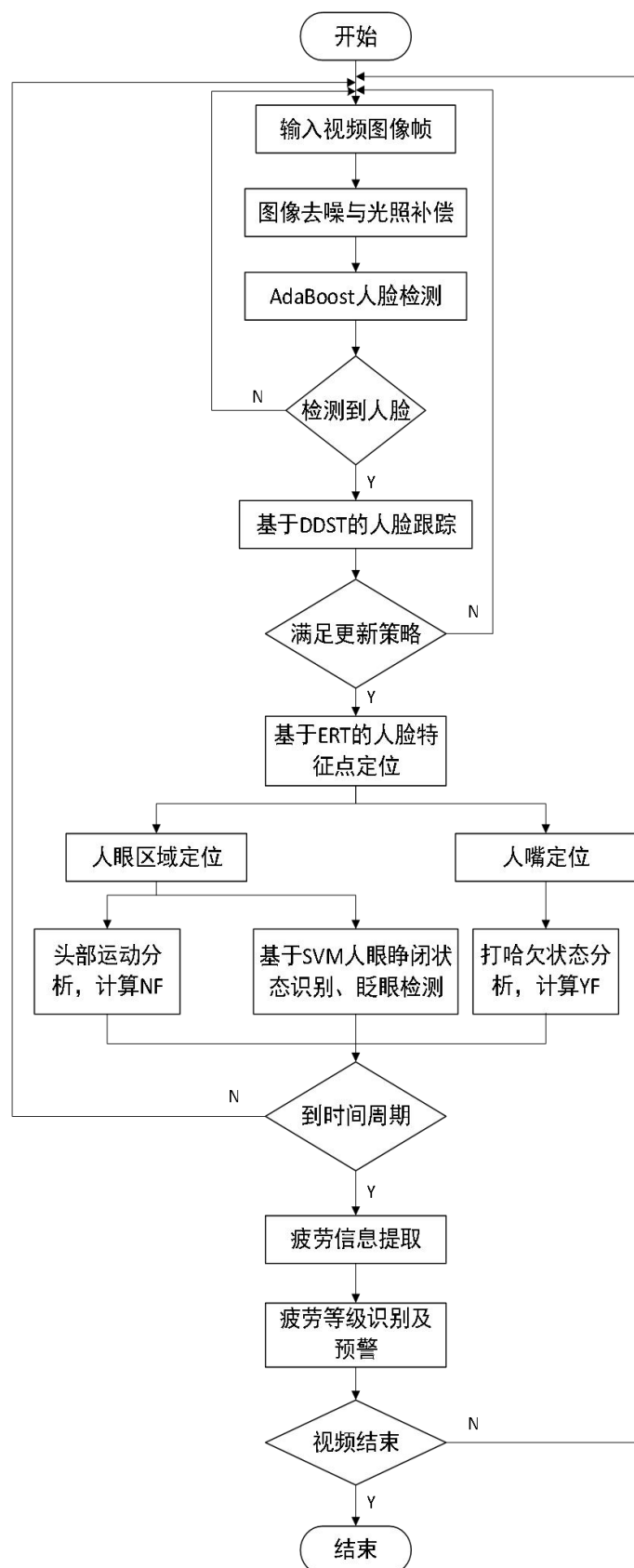


图 4.7 疲劳识别流程图

(2)疲劳驾驶识别的主要步骤为：

- 1) 实时获取视频单帧图像，并对图像平滑去噪及光照均衡处理；
- 2) 人脸检测定位及跟踪。若当前已检测到人脸则进行人脸跟踪，否则继续检测人脸，若跟踪不满足更新策略，则下一帧继续人脸检测；
- 3) 人脸特征点定位及人眼定位。若检测且跟踪到人脸，则对人脸特征点进行定位，并根据各个特征点的位置序号对人眼、人嘴区域进行定位；
- 4) 人眼睁开状态识别。若准确定位了人眼，则计算人眼纵横比 EAR 和人眼二值图像的累计黑色像素差，并融合这两个特征基于 SVM 分类器进行人眼睁开状态判断。
- 5) 疲劳信息提取。根据人眼睁开状态计算时间周期内闭眼帧数所占比 ECR 及最长持续闭眼时 MECT；根据人眼纵横比 EAR 进行眨眼检测，计算时间周期内的眨眼频率 BF；根据嘴部状态计算周期内打哈欠频率 YF；根据定位的眼部位置，分析头部运动状态计算时间周期内的点头频率 NF。
- 6) 时间周期结束时，计算这五个特征参数的加权和，根据加权和值进行疲劳状态判定并进行相应语音预警；若视频未结束则开始下一个周期循环；若时间周期未结束，则进入下一帧图像的特征参数提取，由此循环，直到视频结束时退出。

4.3 实验结果分析

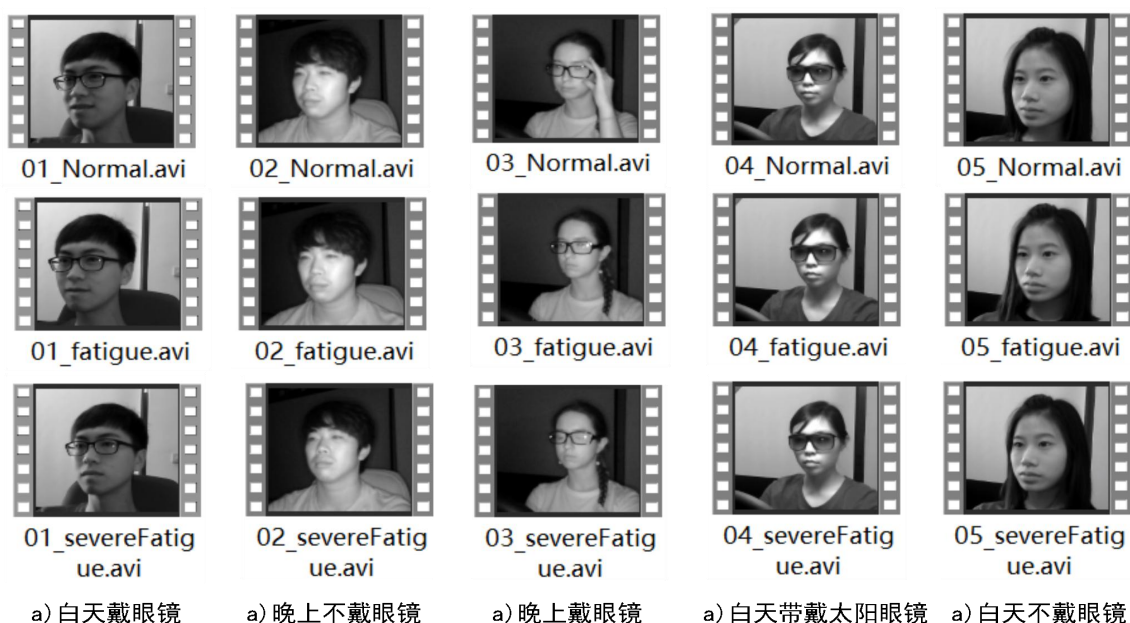


图 4.8 疲劳状态识别测试数据样例

为了验证研究方法的性能，本文的实验在 64 位操作系统的 PC 上进行，采用 python 编程语言，并结合 Opencv 2.4.13 和 Dlib18.17 函数库进行实验分析。实验测试数据来自于 NTHU 驾驶员疲劳检测视频数据集^[61]，该测试数据中有 5 种不同

场景：白天戴眼镜、戴太阳镜和不戴眼镜，晚上戴眼镜和不戴眼镜。每个场景中包含 16 组数据，每组数据包含清醒、疲劳和重度疲劳状态，从测试数据的每个场景中选择一组数据展示如图 4.8。

以 60s 为一个周期检测驾驶员疲劳状态，从 5 种不同场景的每个场景选取 11 组数据共 165 个视频用于寻找各疲劳指标的最优权值，每个权值在 0.1-0.6 之间变化，表 4.4 为各疲劳指标部分不同权值的选取对疲劳状态识别准确率的影响，由表中数据可知当各疲劳指标权重值为表 4.2 中的值时，疲劳识别率最高。疲劳等级识别准确率的计算如下式：

$$\text{疲劳等级准确率} = \frac{\text{正确识别疲劳等级的次数}}{\text{总识别次数}} \quad (4.6)$$

表 4.4 疲劳指标权值寻优

特征参数 权重 W_i 取值组数	MECT	ECR	BF	YF	NF	疲劳等级准确率
1	0.1	0.1	0.1	0.4	0.3	75.15%
2	0.1	0.1	0.1	0.3	0.4	75.76%
3	0.2	0.1	0.1	0.3	0.3	77.56%
4	0.1	0.2	0.1	0.3	0.3	74.55%
5	0.1	0.1	0.2	0.3	0.3	78.18%
6	0.2	0.2	0.1	0.2	0.3	86.67%
7	0.2	0.2	0.1	0.3	0.2	76.97%
8	0.2	0.1	0.2	0.2	0.3	89.70%
9	0.2	0.1	0.2	0.3	0.2	89.09%
10	0.1	0.2	0.2	0.2	0.3	86.06%
11	0.1	0.2	0.2	0.3	0.2	83.03%
12	0.2	0.2	0.2	0.2	0.2	84.85%

选取各疲劳指标最优权重值，对每个场景剩余 5 组数据共 75 个视频进行疲劳状态识别，表 4.5 为疲劳识别的结果；表 4.6 为在白天戴眼镜情况下 15 个视频的各特征参数的具体计算结果、疲劳值与对应的疲劳识别结果。

表 4.5 不同环境下疲劳识别结果

场景	真实状态	清醒(次)	疲劳(次)	重度疲劳(次)	疲劳等级准确率
白天戴眼镜	清醒	5	0	0	100%
	疲劳	0	5	0	
	重度疲劳	0	0	5	

(续上表)

场景	真实状态	清醒(次)	疲劳(次)	重度疲劳(次)	疲劳等级准确率
白天不戴眼镜	清醒	5	0	0	100%
	疲劳	0	5	0	
	重度疲劳	0	0	5	
白天戴太阳镜	清醒	5	0	0	53.3%
	疲劳	3	2	0	
	重度疲劳	2	2	1	
晚上不戴眼镜	清醒	5	0	0	93.3%
	疲劳	0	5	0	
	重度疲劳	0	1	4	
晚上戴眼镜	清醒	5	0	0	86.7%
	疲劳	0	5	0	
	重度疲劳	0	2	3	

表 4.6 白天戴眼镜疲劳识别结果

组数	状态	眼部		嘴部		头部		疲劳值 F	识别结果
		mect(s)	ecr	bf(次)	yf(次)	yt(s)	nf(次)		
1	正常	0.089	0.016	27	2	2.500	0	0.2	正常
	疲劳	0.122	0.027	41	0	0	10	0.5	疲劳
	重度疲劳	0.878	0.082	44	2	4.633	9	0.7	重度疲劳
2	正常	0.103	0.022	24	1	3.000	3	0	正常
	疲劳	0.367	0.045	38	4	9.300	1	0.4	疲劳
	重度疲劳	2.637	0.381	9	1	4.863	2	0.7	重度疲劳
3	正常	0.256	0.015	25	1	1.133	1	0	正常
	疲劳	6.344	0.444	19	1	6.267	0	0.5	疲劳
	重度疲劳	1.878	0.416	34	2	4.536	0	0.7	重度疲劳
4	正常	0.133	0.012	17	3	3.800	0	0.2	正常
	疲劳	0.211	0.025	23	3	3.033	13	0.5	疲劳
	重度疲劳	2.889	0.391	26	1	8.727	2	0.7	重度疲劳
5	正常	0.112	0.031	18	1	4.467	0	0.2	正常
	疲劳	1.422	0.208	27	2	1.894	11	0.6	疲劳
	重度疲劳	2.778	0.293	23	3	5.637	8	0.8	重度疲劳

由上表可以看出，提出的疲劳识别方法在白天的识别准确率比夜晚要好，在戴太阳镜时识别的精度较低，但是就整体而言，识别效果较好。

最后，将本文方法的实验结果与文献[20]及文献[36]的实验结果进行一个对比，表 4.7 为对比结果，可以看出本文提出的方法比文献[20]及文献[36]的方法略好。文献[20]采用基于 CNN 的方法对眼部和嘴部状态的进行识别，并据此状态提取眨眼频率、PERCLOS 和打哈欠等疲劳信息进行疲劳状态识别；文献[36]使用深度卷积神经网络实现睁闭眼状态的识别，最后融合 PERCLOS 值和打哈欠次数实现疲劳驾驶检测及预警。

表 4.7 疲劳检测准确率对比

场景	本文的准确率	文献[20]的准确率	文献[36]的准确率
有眼镜	93.35%	93.29%	89.60%
无眼镜	96.65%	96.47%	93.15%

下表为本文方法每帧各模块的平均运行时间，由表可知整体运行时间为：159.5903ms，检测到人脸后运行时间约 17.1003ms。当出现人脸错检或跟踪目标丢失时，立即进入下一帧的检测，即使一个时间周期内有 3-5 秒的错检，也可以满足 30 帧/秒以上的处理速度，可见该疲劳识别方法具有良好的实时性。

表 4.8 各模块平均运行时间

功能	图像预处理	人脸检测	人脸跟踪	特征点定位	疲劳信息提取	疲劳识别
时间/ms	119.32	23.17	8	3.04	6.06	0.0003

4.4 本章小结

本章主要介绍了一种融合三类疲劳参数指标加权和的疲劳识别方法，其主要评估疲劳的指标有：①眼部疲劳信息眨眼频率、最长眼睛持续闭合时间及基于 PERCLOS 的闭眼帧数所占比；②嘴部指标打哈欠次数及打哈欠持续时间；③头部疲劳信息的点头频率。根据上述疲劳特征参数，建立了疲劳识别模型，根据相应的比重进行加权计算疲劳值来判断驾驶者的疲劳程度，实验表明本文的疲劳识别方法准确率较高且实时性较好。

结 论

随着人们生活的质量的不断提高,汽车成为了人们出行的主要交通工具,汽车在给人们带来方便和快捷的同时其频繁造成的交通事故也带来了大量的人员伤亡和经济损失。虽然国内目前已对超速和酒驾行为采取了有力的惩罚措施,使得交通事故带来的后果有所缓解,但在疲劳驾驶监控方面,仍处于研究阶段。由于每年由疲劳驾驶而导致的交通事故给人们的生命财产安全造成重大威胁,因此提出一种有效的提前进行疲劳检测及实时预警驾驶员的方法具有重大意义。

1. 本文主要工作

本文提出了一种基于多特征加权和的疲劳驾驶识别方法对驾驶员的疲劳程度进行检测并给与相应的预警,具体工作如下:

(1) 图像预处理。主要对图像进行平滑去噪及光照补偿处理,根据实验结果分析了各类算法的优缺点,最后选择自适应中值滤波和基于动态阈值的光照补偿法对图像进行平滑滤波去噪及均衡明暗度。

(2) 人脸检测与人脸跟踪。简要介绍了现有人脸检测的各类方法的原理,基于阈值肤色模型提出了一种改进的积分投影算法,与 AdaBoost 算法进行比较分析后,发现改进的算法对有类肤色干扰的图像检测效果不是很好,因此本文选择基于 haar-like 特征的 AdaBoost 算法进行人脸检测,该算法检测的效率和准确率都很高。在检测到人脸后,介绍了 DDST 目标跟踪算法的原理及其跟踪流程,经实验发现该方法能很好的对人脸进行跟踪。

(3) 人脸特征点及人眼定位。为了减少外界环境因素的影响,快速的定位人眼和嘴部区域,本文在特征区域定位之前使用了基于及级联回归树算法对人脸的特征点进行定位,然后根据特征点的位置快速准确的定位人眼及嘴部区域。

(4) 人眼睁闭状态识别。提出了一种基于 SVM 的睁闭眼判断方法。①根据定位的人眼特征点,我们提出计算人眼纵横比 EAR 来判断人眼睁眼、眨眼、闭眼状态,经实验分析发现根据 EAR 的阈值可以区别这几种状态。②根据定位的人眼区域,使用局部自适应阈值算法对人眼图像进行二值化,经过形态学处理后,根据自适应阈值方法计算人眼二值图像黑色像素的累积差判断人眼睁闭状态。在这两种方法的基础上,本文提出将这两种特征融合作为 SVM 的模型输入进行分类器训练,然后使用训练好的分类器进行人眼状态识别,实验发现基于 SVM 的人眼睁闭状态识别比使用前面两种单一方法的效果要好。

(5) 嘴部状态识别。首先介绍了嘴部状态识别的方法,然后根据定位的嘴部 10 各特征点,计算嘴部高宽比 MAR,经实验分析发现可以设定阈值来区分闭嘴、

讲话、打哈欠等状态，最后我们在 YawDD 打哈欠数据集上进行了打哈欠检测寻找最优阈值。

(6) 点头频率计算。提出了一种基于二维垂直方向的点头频率特征分析算法。计算定位的左右眼区域的两个中心点位置，取两中心点的中点作为头部位置运动分析点，以该检测点在垂直方向上坐标 y 随时间的变化情况，计算时间周期内的点头频率。

(7) 疲劳状态识别。根据人眼睁开状态判断结果，提取眼部特征参数最长持续闭眼时间和闭眼帧数占比，并进行了眨眼检测实验。将眼部疲劳指标眨眼频率、最长持续闭眼时间、闭眼帧数占比，嘴部疲劳指标打哈欠次数、打哈欠持续时间，头部疲劳指标点头频率进行信息融合，加权得到疲劳值，根据疲劳值得到对应的疲劳等级并进行相应的语音预警提示，经实验验证本文提出的方法准确率高实时性较好。

2. 下一步的工作与展望

由于驾驶员行车过程中存在很多不确定因素，国内疲劳驾驶行为识别技术仍处于研究阶段，且该项技术研究是多个学科的综合，设计并实现一个方便快捷、具有较高准确性、实时性和鲁棒性的系统仍具有较大的挑战。本文研究的算法虽然在一定的程度上能进行疲劳识别，但仍存在一些缺陷，需做如下几个方面的改进：

(1) 提高准确性和鲁棒性。在实际行车过程中，需要对驾驶员实时检测，本文疲劳识别方法在太阳光很强和夜晚时的识别率还不理想，若能针对不同光照环境，进行不同的光照处理方法，识别方法将适用于更多的驾驶环境。

(2) 融合车辆特征。本文只是基于驾驶员的行为特征进行疲劳驾驶识别，若是能融合车辆特征如车辆偏移、车速、方向盘转向等特征参数将使得识别结果更为准确。

(3) 扩展开发平台。本文只在 Windows 系统上进行实验，不利于推广与使用，若能够搭建一个可移植的实时检测系统，更有利于实现商业化使用。

若在今后的研究中能对以上几点不足之处进行改进，并取得较好的成果，则疲劳驾驶检测识别技术将到达一个新的阶段，同时，这对后面的学术研究以及实现商用都具有很大的意义与价值。

参考文献

- [1] 《2018 年国民经济和社会发展统计公报》
http://www.stats.gov.cn/tjsj/zxfb/201902/t20190228_1651265.html, 2019-02-28
- [2] World Health Organization. Global Status Report On Road Safety 2018.
http://www.who.int/violence_injury_prevention/road_safety_status/2018/en/,
2018, 15(4): 286
- [3] National Bureau of statistics of China. China Statistical Yearbook.
<http://www.stats.gov.cn/english/>, 2018
- [4] 刘援朝, 孙忠友, 魏玉桂. 机动车驾驶员注意及相关因素的调查研究[J]. 社会心理学, 2007, Vol(z1): 89-100
- [5] NHTSA. Drowsy Driving. <https://www.nhtsa.gov/risky-driving/drowsy-driving>,
2017
- [6] 邹昕彤. 基于表情与头部状态识别的疲劳驾驶检测算法的研究:[吉林大学硕士论文]. 吉林: 吉林大学, 2017, 1-14
- [7] 晏朝. 疲劳对驾驶适宜性的影响及识别方法研究:[长安大学硕士论文]. 西安: 长安大学, 2017, 1-7
- [8] Sheng Yang Shi, Wen Zhong Tang, Yan Yang Wang. A Review on Fatigue Driving Detection. In: ITM Web of Conferences 12. 2017, 14-21
- [9] Zhang J Y, Qiu W W, Fu H J, et al. Review of Techniques for Driver Fatigue Detection[J]. Applied Mechanics and Materials, 2013, Vols(433-4350) : 928-931
- [10] A Mittal, K Kumar, S Dhamija, et al. Head movement-based driver drowsiness detection: A review of state-of-art techniques[C]. In: IEEE International Conference on Engineering and Technology. Coimbatore: IEEE, 2016, 903-908
- [11] A Mashko. Review of approaches to the problem of driver fatigue and drowsiness[C]. In: Smart Cities Symposium Prague. Prague: IEEE, 2015, 1-5
- [12] B. M. K. Kumari, P. R. Kumar. A survey on drowsy driver detection system[C]. In: 2017 International Conference on Big Data Analytics and Computational Intelligence. Chirala: IEEE, 2017, 272-279
- [13] 胡鸿, 易灿南, 廖远志等. 疲劳驾驶研究进展综述[J]. 价值工程, 2015, Vol(18): 254-256
- [14] Stork M, Skala J, Weissar P, et al. Various approaches to driver fatigue detection: A review[C]. In: International Conference on Applied Electronics. Pilsen: IEEE,

- 2015, 239-244
- [15] Jap B T, Lal S, Fischer P. Comparing combinations of EEG activity in train drivers during monotonous driving[J]. Expert Systems with Applications, 2011, 38(1):996-1003
- [16] Balasubramanian V, Adalarasu K. EMG-based analysis of change in muscle activity during simulated driving[J]. Journal of Bodywork and Movement Therapies. 2007, Vol 11(2):151-158
- [17] Vural E, Bartlett M, Littlewort G, et al. Discrimination of Moderate and Acute Drowsiness Based on Spontaneous Facial Expressions[C]. In: 20th International Conference on Pattern Recognition. Istanbul: IEEE, 2010, 3874-3877
- [18] R. Grace, V. E. Byrne, D. M. Bierman, et al. A drowsy driver detection system for heavy vehicles[C]. In: 17th Digital Avionics Systems Conference. Bellevue, WA, USA: IEEE, 1998, vol.2(I36): 1-8
- [19] 付川云. 疲劳状态下驾驶人生理及眼动特征研究:[哈尔滨工业大学硕士论文]. 黑龙江: 哈尔滨工业大学, 2011, 2-41
- [20] 耿磊, 袁菲, 肖志涛等. 基于面部行为分析的驾驶员疲劳检测方法[J]. 计算机工程, 2018, 44(1): 274-279
- [21] Friedrichs F, Yang B. Drowsiness monitoring by steering and lane data based features under real driving conditions[C]. In: Signal Processing Conference. European: IEEE, 2010, 209-213
- [22] C. Papadelis, C. Kourtidou Papadeli, P. D. Bamidis et al. Indicators of Sleepiness in an ambulatory EEG study of night driving[C]. In: International Conference of the IEEE Engineering in Medicine & Biology Society. New York: IEEE, 2006, 6201-6204
- [23] M. Sangeetha, S. Kalpanadevi. Driver Fatigue Management System using Embedded ECG Sensor[J]. International Journal for Scientific Research and Development. 2015, Vol 3(4): 1220-1224
- [24] M. I. Chacon-Murguia, C. Prieto-Resendiz. Detecting Driver Drowsiness: A survey of system designs and technology[J]. IEEE Consumer Electronics Magazine. 2015, vol(4): 107-119
- [25] 幸坚炬. 基于卷积神经网络的人脸识别在疲劳驾驶检测中的应用:[广东技术师范学院硕士论文]. 广东: 广东技术师范学院, 2017, 2-5
- [26] T. Brandt, R. Stemmer, A. Rakotonirainy. Affordable visual driver monitoring system for fatigue and monotony[C]. In: 2004 IEEE International Conference on Systems, Man and Cybernetics. The Hague: IEEE, 2004, vol(7): 6451-6456

- [27] 蔡伽. 基于人眼状态检测的疲劳驾驶识别研究:[河北科技大学硕士论文]. 河北: 河北科技大学, 2017, 2-41
- [28] Abdullah M H, Raman K J, Azman A, et al. Driver Fatigue Detection[C]. In: Information Science and Applications 2016. Lecture Notes in Electrical Engineering: pringer Singapore, 2016, vol 376: 269-278
- [29] 李延枫. 基于眼部识别的疲劳驾驶检测系统设计:[成都理工大学硕士论文]. 四川: 成都理工大学, 2017, 2-35
- [30] Zhang L, Yang D, Ni H, Driver, et al. Fatigue Detection Based on SVM and Steering Wheel Angle Characteristics[C]. In: Proceedings of the 19th Asia Pacific Automotive Engineering Conference and SAE-China Congress 2017: Selected Papers. Singapore: Springer, 2017, vol 486:729-738
- [31] Z. L. Li, X Jin, B J Wang, et al. Study of Steering Wheel Movement under Fatigue Driving and Drunk Driving Based on Sample Entropy[J]. Applied Mechanics and Materials, 2015, Vols.744-746:5, 2001-2005
- [32] Li F, Wang XW, Lu BL. Detection of Driving Fatigue Based on Grip Force on Steering Wheel with Wavelet Transformation and Support Vector Machine[J]. Neural Information Processing. 2013, vol 8228: 141-148
- [33] M. Mao, L. Du. Research on drive fatigue detection using wavelet transform[C]. In: 2007 IEEE International Conference on Vehicular Electronics and Safety. Beijing: IEEE, 2007, 1-4
- [34] A. Zhang and Y. Chen, EEG feature extraction and analysis under drowsy state based on energy and sample entropy[C]. In: International Conference on Biomedical Engineering & Informatics. Chongqing: IEEE, 2013, 501-505
- [35] P. Chen. Research on driver fatigue detection strategy based on human eye state[C]. In: 2017 Chinese Automation Congress. Jinan: IEEE, 2017, 619-623
- [36] 戴诗琪, 曾智勇. 基于深度学习的疲劳驾驶检测算法[J]. 计算机系统应用, 2018, 27(7): 115-122
- [37] 黎亚平, 周杰, 黄磊等. 国内外驾驶疲劳状态检测技术的现状与发展[J]. 上海学报, 2010(05)
- [38] 李娟, 王富, 王维锋等. 基于数据融合的疲劳驾驶检测算法[J]. 武汉工程大学学报, 2016, 38(5): 505-510
- [39] 唐娅琴. 几种图像平滑去噪方法的比较[J]. 西南大学学报(自然科学版), 2009, vol 31(11), 125-128
- [40] 张铮, 倪红霞, 苑春红等. 精通 Matlab 数字图像处理与识别[M]. 北京: 人民邮电出版社, 2013, 85-143

- [41] Rein-Lien Hsu, M. Abdel-Mottaleb, A. K. Jain. Face Detection in Color Images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002, 24(5): 696-706
- [42] Weng Ching-Chih, Chen Homer, Fuh Chiou-Shann. A novel automatic white balance methd for digital still cameras[J]. Proc IEEE Int Symp Circuits Syst, 2005, vol 4: 3801-3804
- [43] 董立新. 基于先验知识的人脸检测算法研究:[大连理工大学硕士论文]. 辽宁: 大连理工大学, 2010, 2-32
- [44] 张明慧, 张明超, 张尧禹. 基于统计的人脸检测方法研究[J]. 电脑编程技巧与维护, 2012(18): 86-86
- [45] Viola P, Jones M J. Robust Real-Time Face Detection[J]. International Journal of Computer Vision, 2004, 57(2): 137-154
- [46] Lienhart R, Maydt J. An Extended Set of Haar-like Features for Rapid Object Detection[C]. In: International Conference on Image Processing. Rochester, NY, USA: IEEE, 2002, 900-903
- [47] 钟国崇. 基于相关滤波的单目标跟踪算法研究:[南昌航空大学硕士论文]. 江西: 南昌航空大学, 2018, 4-16
- [48] 张娟, 毛晓波, 陈铁军. 运动目标跟踪算法研究综述[J]. 计算机应用研究, 2009, 26(12): 4407-4410
- [49] Feng X, Jiang Z. An improved mean shift algorithm for object tracking[C]. In: International Conference on Multimedia Technology. Hangzhou: IEEE, 2011, 4833-4836
- [50] Wang N, Yeung D Y. Learning a Deep Compact Image Representation for Visual Tracking[C]. In: NIPS'13 Proceedings of the 26th International Conference on Neural Information Processing Systems. USA: ACM, 2013, vol(1) : 809-817
- [51] D. S. Bolme, J. R. Beveridge, B. A. Drape, et al. Visual object tracking using adaptive correlation filters[C]. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, CA, USA: IEEE, 2010, 2544-2550
- [52] M Danelljan, G Häger, F Khan, el at. Accurate Scale Estimation for Robust Visual Tracking[C]. In: In Proceedings of the British Machine Vision Conference. BMVA Press, 2014, vol(65):1-11
- [53] M Danelljan, G Häger, F Khan, el at. Discriminative Scale Space Tracking[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence. 2017, vol 39(8): 1561-1575
- [54] Wang M, Liu Y, Huang Z. Large Margin Object Tracking with Circulant Feature

- Maps[C]. In: IEEE Conference on Computer Vision and Pattern Recognition, 2017, vol 1: 4021-4029
- [55] Yang H, Jia X, Loy C C, et al. An Empirical Study of Recent Face Alignment Methods[J]. Computer Science, 2015, 1-12
- [56] Kazemi V, Sullivan J. One Millisecond Face Alignment with an Ensemble of Regression Trees[C]. In: IEEE Conference on Computer Vision and Pattern Recognition. USA: IEEE, 2014, 1867-1874
- [57] Tereza Soukupová, Jan Čech. Real-Time Eye Blink Detection using Facial Landmarks[C]. In: 21st Computer Vision Winter Workshop. 2016, 1-8
- [58] Lee W O, Lee E C, Park K R. Blink detection robust to various facial poses[J]. Journal of Neuroscience Methods, 2010, 193(2):356-372
- [59] C Cortes, V Vapnik. Support-Vector Networks.[J]. Machine Learning. 1995, 20(3): 273-297
- [60] G Pan, L Sun, Z Wu , et al. Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcam[C]. In: 2007 IEEE 11th International Conference on Computer Vision. Rio de Janeiro: IEEE, 2007, 1-8
- [61] Weng C H , Lai Y H , Lai S H . Driver Drowsiness Detection via a Hierarchical Temporal Deep Belief Network[C]. In: Asian Conference on Computer Vision. Springer, 2016, 117-133
- [62] 杨露, 杨万坤. 面向疲劳驾驶的阈值法人眼睁开识别研究[J]. 知识经济, 2013(14), 105-106
- [63] S Abtahi, S Shirmohammadi, B Hariri, et al. A yawning measurement method using embedded smart cameras[C]. In: IEEE International Instrumentation and Measurement Technology Conference. Minneapolis: IEEE, 2013, 1605-1608
- [64] S Abtahi, M Omidyeganeh, S Shirmohammadi, et al. YawDD: A Yawning Detection Dataset[C]. In: ACM Multimedia Systems. Singapore: ACM, 2014, 24-28
- [65] Qing W, Bingxi S, Bin X, et al. A PERCLOS-Based Driver Fatigue Recognition Application for Smart Vehicle Space[C]. In: Third International Symposium on Information Processing. Qingdao: IEEE, 2010, 437-441

附录 A 攻读学位期间的学术成果

1. 申请的软件著作权

[1] 胡峰松, 徐青云, 徐蓉, 彭清舟. 一种基于 IPV6 的车载疲劳驾驶预警系统 V1.0: 中国, 2018SR1087845

2. 申请的发明专利

[1] 胡峰松, 张永, 王冕, 李苍, 全夏杰, 徐青云. 一种基于音频指纹的分片音频检索方法: 中国, 发明专利, 201710044606. 9

附录 B 攻读学位期间参与的项目

1.参与的项目

[1] 赛尔网络下一代互联网技术创新项目：基于 Ipv6 的疲劳驾驶预警系统研究，2018-2019

[2] 湖南掌上医疗信息技术有限公司：共享大夫诊所平台，2017-2018

致 谢

岁月如流水般从生命中淌过，仿佛什么也没留下，然而我知道它必定为我们带来了些什么。三年时光一晃便从指尖流逝，我们总感慨时间不够用，总有留恋不已的风景，就如现在的我。在湖大的三年，将是我毕生难忘的三年，因为在这里我遇见了亦师亦友的人生导师，结识了一群可爱、热情、幽默风趣、乐于助人、热爱生活、有信念、有理想的小伙伴，他们为我的人生画卷增添了丰富的色彩。临近毕业，我知道即使有再多的不舍，我们仍需各自满怀憧憬继续前行，在此，向来过我生命中的你们表达我最真挚的感谢与祝愿！

由衷地感谢我的导师胡峰松教授，在第一次见面交谈时，胡老师给我的印象就是慈眉善目、很有亲和力。在这三年的学习生活中，胡老师给予了我很多帮助和关怀，在学习上，他总以其渊博的专业学识与严谨的学术态度耐心地给与我建议与指导。毕业论文从选题到具体实现过程中，胡老师从大方向到细节处理为我严格把关，他总能给我一些有启发性的建议，本论文能顺利的完成，离不开胡老师的悉心指导。在学习之余，胡老师也会和我聊天谈心，讲讲所见所闻，教导我为人处事的道理，让我受益匪浅。胡老师在学习和生活上都给予了我很大的帮助，是我的良师益友，其与人为乐的处世态度值得我学习，在此，向胡老师表达我崇高的敬意。

逝水流年，感恩这地球是“缘”的，因为心中有梦，我们在湖大这所千年学府美丽的遇见。感谢与我相伴七年的熊和饼饼，从大学到研究生，你们在我失意的时候给予我安慰与帮助，在我自负的时候给我警醒，在我快乐的时候与我共享，感恩我们一起度过了人生最灿烂的年华。感谢我的室友兼好友小悦悦，缘分奇妙的让我们分离又遇见，感谢这三年中你形影不离的陪伴，我们相互交流相互学习，互帮互助，从你身上我学到了什么叫不惹事也不怕事，自己的正当权益坚决捍卫。感谢身边的小伙伴慧灵、阿蓉、莹妹、敏妹、小芳、婷宝宝及我的师兄师姐、师弟师妹们，感谢你们一直以来的陪伴、帮助、鼓励与监督，是你们让我这三年的学习生活不枯燥，使我的生活充满感动与爱。

同时，我必须感谢在成长路上一直陪伴、理解与支持我的父母，感谢你们的无私付出，让我可以无忧快乐的成长；感谢你们孜孜不倦的教导，母爱如海，父爱如山，是你们让我领略了人世间美好的风景；你们是我前进的动力，是我精神的支助，爱尽在不言中！愿你们今后喜乐安康，让我能回报你们的养育之恩。

最后，感谢评审专家们及答辩组的各位老师花费宝贵的时间对本论文评阅，万分感谢！