

学校代号 10532

学 号 S12093012

分 类 号 TP391

密 级 公 开



湖南大学  
HUNAN UNIVERSITY

## 硕士学位论文

# 基于面部视觉多特征融合的驾驶员 疲劳检测方法研究

学位申请人姓名	周云鹏
培 养 单 位	电气与信息工程学院
导师姓名及职称	朱青 副教授
学 科 专 业	电子科学与技术
研 究 方 向	图像处理技术
论 文 提 交 日 期	2015 年 4 月 22 日

学校代号：10532

学    号：S12093012

密    级：公开

## 湖南大学硕士学位论文

# 基于面部视觉多特征融合的驾驶员 疲劳检测方法研究

学位申请人姓名：周云鹏

导师姓名及职称：朱青 副教授

培 养 单 位：电气与信息工程学院

专 业 名 称：电子科学与技术

论文提交日期：2015 年 4 月 22 日

论文答辩日期：2015 年 5 月 14 日

答辩委员会主席：黎福海 教授

Research on method of driver drowsiness detection based on fusion  
of multi-face clues

by

ZHOU Yunpeng

B.E.( Hunan Institute of Technology)2012

A thesis submitted in partial satisfaction of the

Requirements for the degree of

Master of Engineering

in

Electronics Science and Technology

in the

Graduate School

of

Hunan University

Supervisor

Associate Professor ZHU Qing

April, 2015



## 摘 要

近年来机动车数量的急剧增加，在给人们的生活和工作带来了很大便利的同时，随之而来的交通问题也愈发严重。据相关部门统计，由疲劳驾驶引发的交通事故占了很大的比例，所以开发一个实时监测驾驶员疲劳状态的系统具有很重要的实际价值，而疲劳检测算法的研究是系统实现的基础，因此本文研究了一种基于驾驶员面部图像信息的疲劳检测算法。

本文提出了一种基于面部视觉多特征融合的驾驶员疲劳状态检测算法，通过分析驾驶员眼睛、嘴巴以及整体面部表情的状态，并采用模糊系统融合推理判断驾驶员的疲劳状态。首先，采用基于 Haar 特征的级联 Adaboost 检测算法确定驾驶员的脸部位置，然后利用 Gabor 滤波器增强面部局部特征的边缘信息以准确定位眼睛和嘴部，采用一种旋转不变的 LBP 金字塔特征对眼睛进行特征描述，训练线性 SVM 分类器判别眼睛的开闭状态；并根据嘴部二值图像的张开面积及宽高比判断嘴部的开闭状态；同时采用多模态特征的距离测度学习算法判别驾驶员整体面部的疲劳表情。最后根据眼睛和嘴部的状态以及疲劳表情，计算出四个能够描述驾驶员状态的疲劳参数，利用模糊系统融合推理得出驾驶员的疲劳状态，并根据推理结果予以不同的预警措施。

实验结果表明面部局部特征眼睛和嘴部状态判别的算法有较高的准确率，能够满足实际检测系统的要求，而整体面部表情状态判别由于情况比较复杂，外界干扰比较多，但是本文算法的准确率也能够满足实际应用的需要；用于信息融合的模糊系统的合理性也在仿真实验中得以验证。最终，本文的第 5 章验证了整个算法流程的完整性和合理性。

**关键词：**疲劳检测，Gabor 滤波，LBP，距离测度学习，模糊系统

## Abstract

Recent years, the number of vehicles increased rapidly, which brings a large amount of convenience to people's daily life and work, but the traffic problems are getting worse. According to the statistics of relevant department, the traffic problems caused by fatigue driving take a great proportion, so developing a system to real-time detect the drowsiness state of the driver has great practical value, whose foundation is drowsiness detecting algorithm, therefore, the thesis proposes a drowsiness detecting algorithm based on the driver's face image.

A detecting algorithm, which is based on the fusion of facial multi-parameters of drowsiness, is proposed, analyzing the state of the driver's eyes and mouth, and the facial expression, and using the fuzzy system to reason the drowsiness state of the driver. Firstly, the driver's face is localized by the detecting algorithm of cascaded Adaboost based on Haar features, and then the marginal information of the face is enhanced using Gabor filter to localize the eyes and the mouth precisely, and the state of the eyes is classified as being open and closed by a linear SVM classifier trained by an rotation-invariant LBP features of the eye images; and the state of the mouth is determined according to the opening-area and the ratio of width-height of the binary image of the mouth; meanwhile the drowsy expressions of the driver are discriminated using multi-view distance metric learning algorithm. Finally, the state of the driver can be reasoned by the fuzzy system using four drowsiness coefficients calculated by the state of the eyes and the mouth, and the facial expression, and there are different forewarning measures to the reasoned results.

The results of the experiments show the accuracy of the state discrimination algorithms of the eyes and the mouth, of which the average accuracy of the recognition meet requirements of practical detecting system, however, for the discrimination of facial expressions, although the circumstance is complex and there are a lot of disturbances outside, the correction rate of the algorithm proposed in this thesis can still satisfy requirements of the practical application; the reasonability of the fuzzy system used to information fusion is verified in the simulation experiment as well. Finally, the reasonability and integrity of the whole algorithm are proved in the fifth chapter of this thesis.

**Key Words:** Drowsiness Detection; Gabor Filter; LBP; Distance Metric Learning; Fuzzy System

# 目 录

学位论文原创性声明和学位论文版权使用授权书 .....	I
摘 要 .....	II
Abstract .....	III
目 录 .....	V
第 1 章 绪论 .....	1
1.1 研究背景及意义 .....	1
1.2 驾驶员疲劳检测技术及国内外研究现状 .....	2
1.2.1 国外的研究现状 .....	2
1.2.2 国内的研究现状 .....	4
1.2.3 疲劳检测系统存在的关键问题 .....	6
1.3 课题的来源 .....	6
1.4 本文的研究思路及主要研究内容 .....	6
1.4.1 本文的研究思路 .....	6
1.4.2 本文的主要研究内容 .....	7
第 2 章 面部局部特征定位及疲劳状态判别方法 .....	8
2.1 人脸检测 .....	8
2.2 眼睛的定位及状态判别 .....	9
2.2.1 眼睛的定位 .....	9
2.2.2 眼睛的状态判别 .....	14
2.3 嘴部的定位及状态判别 .....	20
2.3.1 嘴部的定位 .....	20
2.3.2 嘴部状态判别及实验结果 .....	21
2.4 本章小结 .....	21
第 3 章 整体面部特征的疲劳状态判别方法 .....	22
3.1 整体面部表情的特征提取 .....	22
3.2 基于距离测度学习的整体面部疲劳表情判别 .....	25
3.2.1 距离测度学习的基本概念 .....	25
3.2.2 多模态测度学习的疲劳表情判别 .....	28
3.2.3 疲劳表情判别的实验结果 .....	32
3.3 本章小结 .....	35
第 4 章 基于模糊推理的疲劳状态判别方法 .....	36



4.1 模糊系统的结构 .....	36
4.1.1 模糊化接口 .....	36
4.1.2 规则库 .....	36
4.1.3 模糊推理 .....	37
4.1.4 去模糊化 .....	37
4.2 疲劳状态判别的模糊系统设计 .....	38
4.2.1 定义输入输出变量及其隶属函数 .....	38
4.2.2 规则库的定义 .....	39
4.2.3 模糊推理与去模糊化 .....	42
4.3 疲劳状态判别的仿真实验结果 .....	44
4.4 本章小结 .....	48
第 5 章 驾驶员疲劳状态判别的实验结果 .....	49
5.1 测试样本的采集 .....	49
5.2 测试的实验结果 .....	50
5.3 本章小结 .....	54
总结与展望 .....	55
1. 工作总结 .....	55
2. 工作展望 .....	56
参考文献 .....	57
致谢 .....	61
附录 A 攻读学位期间所发表的学术论文目录 .....	62
附录 B 攻读硕士学位期间参与的科研项目 .....	63

# 第1章 绪论

## 1.1 研究背景及意义

随着国民经济的迅速发展，交通行业已经成为经济建设最重要的基础，汽车对人们工作和生活的影响力也越来越大。据中国公安部统计，截止 2014 年 11 月底，我国机动车驾驶人数量突破 3 亿大关，其中汽车驾驶人 2.44 亿人；全国民用机动车保有量达 2.64 亿辆，其中汽车 1.54 亿，驾驶人数量居世界第一，汽车数量居世界第二位。而与此同时，随着机动车的数量迅速增加，随之而来的交通安全问题也越来越严重，交通事故引起的经济损失和人员伤亡也逐年增加。据统计，2012 年全国接报涉及人员伤亡的道路交通事故 4.6 万起，造成 1.1 万人死亡、5 万人受伤，分别上升 17.7%、16.5%和 12.3%。因此，汽车安全驾驶一直是研究的热门方向，其中安全气囊等一些能够保障驾驶员安全的装置已经被普及，但是这些设计均是被动的保护措施，而不能主动降低交通事故的发生率。但是如果在事故发生前能够对驾驶员提前预警，就能够避免一些交通事故的发生，这项技术称之为汽车主动安全技术。被动的保护措施与主动安全技术相结合的话，就很大程度能够降低交通事故的发生率，并且能减少交通事故中的人员伤亡。本文研究的驾驶员疲劳检测技术就是汽车主动安全技术领域中的一个重要研究方向。

交通事故分析表明，很多的交通事故是由于驾驶员疲劳驾驶引起的，因为如果驾驶员处于疲劳状态，其注意力就不能集中，驾驶能力也会随之下降，反应速度也会明显变慢，对于突发情况的处理能力就远不如正常状态。2010 年美国汽车协会也发布了一份基于美国高速公路安全管理局(NHTSA)的交通事故统计数据的分析报告，表明该年美国至少有 17%造成伤亡的交通事故是由于疲劳驾驶引起的；法国也有调查数据显示由驾驶员疲劳驾驶所引起的交通事故占总交通事故的 21%；而数据表明，在德国，5 起交通事故中平均就有 1 起是由疲劳驾驶引起的<sup>[1]</sup>。

驾驶人员长时间连续行车后，会产生生理机能和心理机能的失调，使驾驶员反应变得迟钝、注意力不集中、动作不灵活、视觉模糊，从而客观上导致驾驶机能下降，引起交通事故。而在较长时间的驾驶过程中，驾驶员的疲劳程度是逐渐积累的，如果能够根据驾驶员眼部、嘴部及头部特征及时有效的监测，并且必要时预警提醒驾驶员，这样就能很大程度上避免发生交通事故。据统计，如果驾驶员比交通事故时的反应快 0.5s，就有 60%的交通事故能够避免<sup>[2]</sup>。因

此研究实时性的疲劳驾驶检测系统具有非常实际的意义。

## 1.2 驾驶员疲劳检测技术及国内外研究现状

汽车驾驶属于感知型劳动，是一项综合性很强的技术工作<sup>[3]</sup>。虽然疲劳是一种正常的生理现象，但是当驾驶人在疲劳状态下驾驶车辆，其对周围环境的感知能力、形势判断能力以及对车辆的操控能力都有不同程度的下降，因此就很容易发生交通事故。驾驶员疲劳检测就是针对这一问题，希望能够监测驾驶员的精神状态，在驾驶员出现疲劳迹象的时候及时提醒，有效地防止交通事故的发生。

目前比较常见的疲劳检测方法概括起来可分为三类：基于驾驶员生理信号的检测方法、基于车辆行驶参数的检测方法以及基于驾驶员行为特征的检测方法<sup>[4]</sup>。

### 1. 基于驾驶员生理信号的检测方法

该方法是一种非常可观、准确率非常高的方法，主要是根据脑电图信号<sup>[5]</sup>、心电图信号<sup>[6]</sup>、肌电图信号<sup>[7]</sup>等生理信号判断驾驶员的状态。通过在驾驶员的身上安装相关的传感器，实时监控其生理信号的变化。这种方法由于驾驶员需要安装很多的检测仪器，本身会对驾驶员有一定的干扰，而且检测的设备非常贵，所以很难推广普及。

### 2. 基于车辆行驶参数的检测方法

该方法主要根据车道偏离<sup>[8]</sup>、方向盘转动的频率及幅度，对方向盘的平均握力变化<sup>[9]</sup>等判断驾驶员的状态。驾驶员在正常状态和疲劳状态下对整个车辆的控制是存在区别的，所以通过不同的传感器采集车辆的行驶参数也可以判断驾驶员的状态。

### 3. 基于驾驶员行为特征的检测方法

此类方法是以驾驶员为检测主体，主要根据驾驶员眼睛的状态<sup>[10,11]</sup>、头部的位置<sup>[12]</sup>、嘴部的状态<sup>[13]</sup>、面部表情等参数来判断驾驶员的状态。通过摄像头采集驾驶员的驾驶状态图像<sup>[14]</sup>，然后分析其面部特征状态，以及一些行为特征来判断驾驶员的状态。这种方法由于是一种非侵入式的检测方式，并且其检测设备非常的简单，所以近些年来，一直是大家研究的主要方向。

#### 1.2.1 国外的研究现状

1996年美国科学家 Knippling 和 Kanianthra 等通过测量眼睛的开闭状态及其运动和眼睛的一些生理学表现形态来判断驾驶员的疲劳状态，同时也认为通过眼睛的状态来判断疲劳状态是非常有效的<sup>[15]</sup>。但是这种方法行之有效的根本前提条件是所选择的疲劳参数能够描述驾驶员的疲劳状态信息，并且该疲劳信息

能够被测量。

美国 Attention Technologies 公司设计了一款根据驾驶员的眼部特征监测其疲劳状况并预警的产品（DD850 Driver Fatigue Monitor）<sup>[16]</sup>，装置如图 1.1 所示。它通过使用红外的传感器采集驾驶员眼部信息，并根据驾驶员眼部的状态计算 PERCLOS 参数作为疲劳报警依据，公司生产的整个设备可以安装在车子的仪表盘上，并且其预警的方式和敏感度可以根据需要调整。



图 1.1 DD850 装置示意图

日本一家公司利用多种传感器发明了一个能够判断驾驶员疲劳状态的系统<sup>[17]</sup>。系统根据驾驶员头部与传感器之间的距离确定驾驶员头部在三维场景中的位置，并统计一段内头部位置的变化判断驾驶员是否疲劳；在方向盘上也安装了检测驾驶员心跳速度的感应器，它可以定时间隔检测驾驶员的心率，并可以根据驾驶员心率的变化判断驾驶员状态，因为人在疲劳状态下，心跳速度较于正常状态会明显变慢；另外，该系统还安装了检测行车速度的传感器，根据驾驶员的行车速度变化判断驾驶员的疲劳状态。最终，系统融合以上的多种参数，综合判断驾驶员的疲劳状态，并且若检测到驾驶员出现异常状态，系统则可以自动预警，提醒驾驶员。

美国 Digital Installations 开发的 S.A.M(Steering Attention Monitor)系统<sup>[18]</sup>，是一种监测方向盘非正常运动的疲劳报警装置，如图 1.2 所示。方向盘正常运动时传感器装置不报警，但是若方向盘 4s 不运动，S.A.M.就能发出警报提醒驾驶员，直到方向盘重新正常运动警报就会自动停止。另外，S.A.M.被固定在仪表盘旁边，并且方向盘下边安装了磁性带，可以监测方向盘的运动情况。



图 1.2 S.A.M.系统的车载装置图

英国的一家公司综合分析驾驶员的睡眠情况、已行驶的时长，驾驶员方向盘的操作情况等各种驾驶状况判断驾驶员的疲劳状态<sup>[19]</sup>。装置在运行前需要输入驾驶员前 24 小时的睡眠情况。当检测到驾驶员属于疲劳驾驶时，就触发预警警报，提醒驾驶员停车休息。系统还内置闹钟，会在合理时长之后叫醒驾驶员，并会自动重置驾驶时间<sup>[20]</sup>。

2001 年，美国 Rensselaer Polytechnic 学院的 Ji Qiang 教授通过统计驾驶员眼睑的运动，头部运动，面部表情等一些参数，利用贝叶斯网络融合决策出驾驶员的状态<sup>[21]</sup>。但是他在眼睛定位的方法上做了改进，通过一个外部的硬件电路来控制两种不同波长的红外光源，根据采集到的亮瞳孔图像和暗瞳孔图像的差分图像来确定眼睛的位置。

2004 年，欧盟启动了一项被称为“AWAKE”的驾驶行为综合监控系统，该系统通过检测驾驶员的眼睛状态、方向盘的操作以及车道线偏离的情况，并研究各种特征与疲劳之间的联系，最后再融合判断驾驶员的疲劳状态，并根据疲劳等级采用不同的方式提醒驾驶员<sup>[22]</sup>。

Zhang 等基于 RBF 神经网络进行眼睛状态识别，并根据眼睛的状态计算 PERCLOS 参数判断驾驶员的疲劳状态，但此方法的计算量较大<sup>[23]</sup>。Sun 等提取眼睛的边缘信息，再采用 Hough 变换来检测眼球，并以此来计算驾驶员眼睛的开闭程度来判断驾驶员的疲劳状态，此方法预处理较复杂，计算较大，不能满足实时性<sup>[24]</sup>。Ralph 根据眼睛的状态和头部的位置计算三个能够代表驾驶员疲劳状态的疲劳参数，再利用 SVM 训练一个分类器，判断驾驶员是否需要疲劳预警<sup>[25]</sup>。Mandalapu 利用 SVM 训练一个嘴部状态判别的分类器，根据嘴部信息判断驾驶员的疲劳状态，此方法由于嘴部状态个体差异性较大，因此判别误差较大且鲁棒性不够<sup>[26]</sup>。Luis M 采用有限状态机对眼睛的各种状态之间的联系建模，计算 PERCLOS、眨眼频率等参数融合决策出驾驶员的疲劳状态<sup>[27]</sup>。

### 1.2.2 国内的研究现状

深圳长途客运公司的研究员周鹏通过对电子工程、人体生理学、现代神经学等领域的研究，深入分析了驾驶员疲劳状态的起因，并提出应该消除驾驶员在长时间行车过程中产生的大脑麻痹和异常疲劳等问题，才能从根本上降低交通事故的发生率<sup>[28]</sup>。基于此思想，他研究开发了一套能够帮助驾驶员消除疲劳的佩戴式设备，这也是中国最早的佩戴式疲劳防止器，它能够促进驾驶员血液和体液的循环，帮助其大脑保持清醒的状态，从而缓解驾驶员出现疲劳等异常情况。

中国单片机公共实验室南京研发中心联合南京远驱科技有限公司研究出一套疲劳检测设备 gogo850，装置如图 1.3 所示，其主要检测驾驶员眼睛的开闭

情况，尤其还增加了瞳孔识别技术，因为疲劳状态的瞳孔很暗，因此当出现驾驶员睁眼瞌睡时也可以被识别出，并且系统处理是基于红外图像，所以该系统在白天和晚上均可以运行，另外系统还可以对驾驶员戴眼镜的情况进行判别，整个系统的实用性很强。



图 1.3 gogo850 装置示意图

吉林大学智能车辆课题组开发的 JLUVA-DFWS 系统基于肤色特征定位人脸区域，然后分别检测出眼睛和嘴唇的特征状态信息，最后依据眼睛和嘴唇张度推断驾驶员的精神状态，确定是否对驾驶员进行预警<sup>[29]</sup>。

比亚迪“疲劳驾驶预警系统 (BAWS)”<sup>[1]</sup>基于驾驶员生理图像反应，由 ECU 和摄像头两大模块组成，利用驾驶员的面部特征、眼部信号、头部运动性等推断驾驶员的疲劳状态，并进行报警提示和采取相应措施的装置。对驾乘者给予主动智能的安全保障<sup>[30]</sup>。

2002 年，浙江大学研制了一个防驾驶员瞌睡的装置，根据灰度积分投影的方法快速确定人眼的位置。首先利用图像在垂直方向上的灰度投影确定人脸的边界，然后根据水平方向的灰度投影曲线定位人眼，最后利用阈值分割及边缘检测判断眼睛的开闭状态，通过统计一段时间内驾驶员眼睛的闭合时间、闭合的频率等参数，来确定驾驶员当前的疲劳状态<sup>[31]</sup>。

重庆大学导航与制导实验室开发了“高速公路智能型汽车行驶主动安全预警系统”，它主要应用了模式识别、信息融合等技术，为驾驶员提供一种辅助驾驶的功能，并且该系统通过了国家重型质量监督检测中心的测试与鉴定，已经达到实际应用的阶段，并且效果处于国际优先水平。

北京航空航天大学研究的眼动测量仪，根据图像传感器采集眼睛运动的模拟图像，再传给图像处理系统，经滤波、增强、以及数模转换产生包括眼睛位置信息的数字信号。其中数字检测系统主要负责两个瞳孔中心的坐标检测和数模转换，并需要和计算机实时通讯，将眼睛的位置坐标传输给计算机以完成标定等工作。计算机对传感器传来的图像信息和标定后的位置信息进行数据处理，再根据结果判断驾驶员的眼动状态，并以此结果来判断驾驶员的状态。

### 1.2.3 疲劳检测系统存在的关键问题

近年来，视频监控已经成为了一个很热门的方向，基于图像的疲劳检测技术研究已经取得了很大的进展，但是由于环境和疲劳成因的复杂性，我们仍然有很多的难题需要解决。

- 1) 由于检测系统需要适应白天和夜间两种环境，天气会导致光照环境非常复杂多变，并且车辆是在不停的运动，不同的外界环境都会影响车内的环境，这样对图像的预处理与算法的鲁棒性就有非常高的要求。然而适用于实验室稳定单一环境下的算法通常很难适用于室外多变的环境，因此系统的鲁棒性也是导致检测系统不能大范围推广的一个重要的原因。
- 2) 产生疲劳的原因非常的复杂，并且由于人与人之间的个体差异非常的明显，生活习惯也不尽相同，所以根据单一的图像信息很难来判断驾驶员是否疲劳。
- 3) 缺少公共测试集，因此很难衡量一个系统的性能。至今对于疲劳检测的研究领域依然没有一个权威的公共测试的样本集，所以所有的算法都是基于研究人员自己的测试样本，这样算法的普适性就很难评判。
- 4) 由于开发的检测系统最终是要安装在车上的，所以设备的大小与成本都限制了算法的复杂度不能太高，并且如果太复杂的算法很难满足实时性的要求。

## 1.3 课题的来源

本课题源于三个项目：国家 863 重点项目“机器视觉感知与智能车载控制项目(2012AA112312)”、国家自然科学基金项目“自主车辆的高质量三维场景认知与导航避障控制方法研究项目(61175075)”、交通运输部科技项目“基于物联网的车路协同主动安全技术研究及工程应用项目(201231849A70)”。

## 1.4 本文的研究思路及主要研究内容

### 1.4.1 本文的研究思路

本文通过融合面部多种疲劳参数来判断驾驶员的疲劳状态，其总体框图如图 1.1 所示。首先，采用基于 Haar 特征的级联 Adaboost 分类器检测人脸，再利用 Gabor 滤波器增强面部边缘信息并结合面部几何特征定位眼睛和嘴部的位置，并且分别判断眼睛和嘴部的开闭状态；同时根据多模态特征距离测度学习的算法对人脸表情进行判断其是否疲劳；最后，通过统计一段时间内眼睛和嘴部的状态以及面部表情状态计算出几个能代表驾驶员疲劳状态的疲劳参数，采用模糊系统推理得出驾驶员最终的疲劳状态。

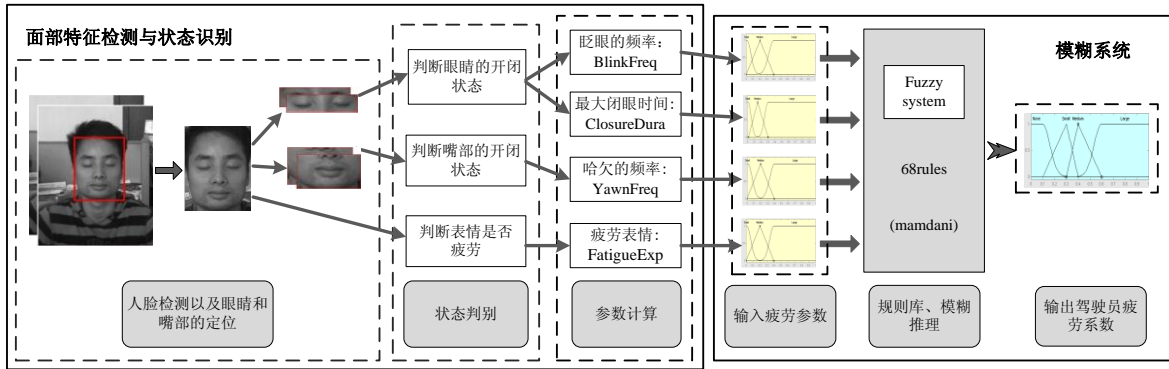


图 1.1 疲劳检测系统的总体框图

### 1.4.2 本文的主要研究内容

第 1 章：绪论。首先简单介绍了本课题的研究背景与意义，说明研究疲劳驾驶技术的必要性，然后概述了在驾驶员疲劳检测方面国内外的研究现状以及所存在的问题。最后还简单介绍了文章的主要研究思路和研究内容。

第 2 章：面部局部特征检测及状态判别。首先简单介绍了人脸检测，在此基础上分别详述了眼睛和嘴部的定位，以及它们的状态的判别。

第 3 章：整体面部表情的疲劳状态识别。首先介绍了用于面部特征描述的三种算子，再引入了一种基于二元组核函数的距离测度学习框架，并介绍了测度学习问题的一种求解方法，基于此还介绍了根据多模态特征的测度学习算法判断驾驶员疲劳表情。

第 4 章：基于模糊推理的信息融合。首先简单介绍了模糊系统的组成部分，并根据本文的需要设计了一个用于驾驶员疲劳信息融合推理的模糊系统，并且最后对系统做了验证性的仿真实验，证明了这个模糊系统推理的合理性。

第 5 章：实验结果分析。综合第 2、3、4 章的内容，组合了一个驾驶员疲劳检测的系统。检验了来自五个不同人的模拟驾驶视频样本，验证了系统的有效性。

总结与展望。首先总结了本文的研究工作，并阐述了文中需要进一步研究的内容。



## 第2章 面部局部特征定位及疲劳状态判别方法

研究表明，疲劳状态有很多特征，包括打瞌睡，打哈欠等一些很直观的表现形式，而本文要求根据驾驶员的面部图像信息判断驾驶员的疲劳状态，所以本章将主要介绍驾驶员眼睛和嘴部的定位及疲劳状态判别<sup>[32]</sup>，其状态判定结果是判断驾驶员疲劳状态的重要依据。本章的算法流程图如图 2.1 所示。

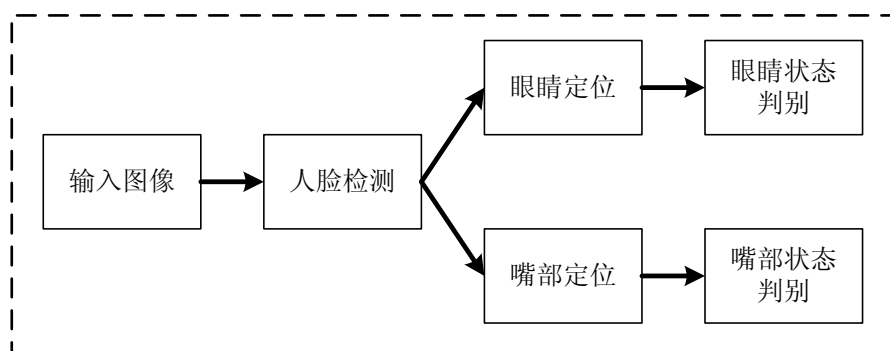


图 2.1 面部局部特征定位及状态判别流程图

### 2.1 人脸检测

人脸检测是整个系统的第一步，也是非常重要的一个环节，因为它的准确率直接影响了整个系统的性能。近年来，人脸检测的算法<sup>[33]</sup>越来越多，并且很多也已经逐步趋于成熟的阶段，无论从检测的速度和准确性都有很大的提高，可以满足工程应用的要求。本文采用基于 Haar 特征<sup>[34]</sup>的级联 Adaboost 人脸检测算法，此算法不仅具有很高的准确性并且有很高的鲁棒性，其中 Haar 特征由于计算过程中引入了积分图像，所以提取的速度非常快，并且此特征对人脸有很强的描述性，所以是一种非常适用于人脸检测的特征。Adaboost 级联算法<sup>[35]</sup>是通过结合 Haar 特征训练多个弱的人脸分类器，再将它们级联组合成一个强的人脸分类器。

由于人脸检测的算法已经比较成熟，文中对于算法的细节就不再详述，只对算法流程简单描述。由于样本之间的采集环境存在差异，因此首先需要对样本进行简单的预处理，包括图像增强，将所有的样本转换到同一个亮度区间内。再对人脸样本和非人脸样本提取 Haar 特征，再结合 Adaboost 算法训练一个人脸分类器。对于系统输入的图像，首先需要经过简单的预处理，然后就加载之前训练的好人脸分类器，再基于滑动窗口法在图片上搜索人脸，其算法流程如图 2.2 所示。

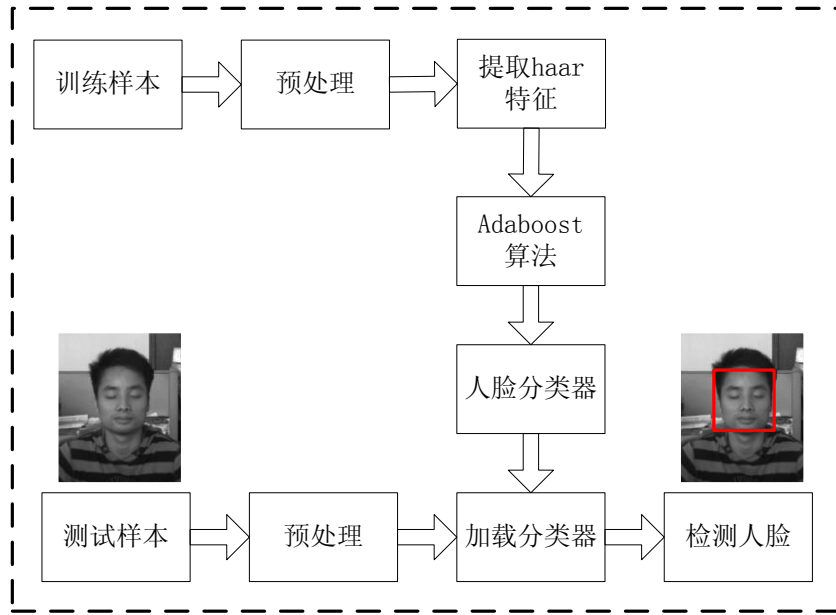


图 2.2 人脸检测流程框图

## 2.2 眼睛的定位及状态判别

### 2.2.1 眼睛的定位

人的面部有一些边缘非常清晰的特征，例如眉毛、眼睛和嘴巴，所以如果能够增强面部图像的边缘信息，就非常有利于这些局部特征的定位。因此，本文提出了一种基于增强面部边缘信息，并且结合面部的几何特征规律定位眼睛和嘴部的算法。目前边缘增强的方法有很多，其中 Gabor 滤波器<sup>[36]</sup>对于图像的边缘信息非常敏感，并且滤波器的方向和尺度可以多元选择，对于边缘非常复杂的图像依然有很好的增强效果，所以本文选择通过 Gabor 滤波器来增强眼睛和嘴部的边缘信息。

#### 2.2.1.1 Gabor 滤波器简介

在图像处理、模式识别等领域中，Gabor 滤波器是一个常用于边缘检测的线性滤波器，并且由于其对频率和方向的表示接近人类视觉系统对于频率和方向的表示，所以也常被用于描述图像的纹理信息和细节信息。在空间域中，一个二维的 Gabor 滤波器是一个正弦平面和高斯核函数的乘积。二维 Gabor 滤波器的定义如式 (2.1) 所示。

$$g(x, y, \lambda, \theta, \varphi, \sigma_x, \sigma_y) = \exp\left(-\frac{1}{2}\left(\frac{x'^2}{\sigma_x^2} + \frac{y'^2}{\sigma_y^2}\right)\right) \times \exp(i * (2\pi \frac{x'}{\lambda} + \varphi)) \quad (2.1)$$

其中  $x' = x \cos \theta + y \sin \theta$ ,  $y' = -x \sin \theta + y \cos \theta$ ;  $\lambda$  表示正弦函数的波长,  $\theta$  表示滤波器核函数的方向,  $\varphi$  表示相位偏移,  $\sigma_x, \sigma_y$  表示高斯函数的标准差。由 (2.1)

式可以看出 Gabor 滤波器的表达式是一个复函数，其实部和虚部的表达式分别为式（2.2）和式（2.3）。

$$g(x, y, \lambda, \theta, \varphi, \sigma_x, \sigma_y) = \exp\left(-\frac{1}{2}\left(\frac{x'^2}{\sigma_x^2} + \frac{y'^2}{\sigma_y^2}\right)\right) \times \cos\left(2\pi \frac{x'}{\lambda} + \varphi\right) \quad (2.2)$$

$$g(x, y, \lambda, \theta, \varphi, \sigma_x, \sigma_y) = \exp\left(-\frac{1}{2}\left(\frac{x'^2}{\sigma_x^2} + \frac{y'^2}{\sigma_y^2}\right)\right) \times \sin\left(2\pi \frac{x'}{\lambda} + \varphi\right) \quad (2.3)$$

在空域中，二维 Gabor 滤波器中主要参数有  $\theta$ 、 $\lambda$ 、 $\sigma_x, \sigma_y$ ，分别表示滤波器的核函数方向、纹理的波长和高斯窗口的大小。下面将具体分析一下这几个参数的意义<sup>[37]</sup>如下：

$\theta$ ：随着参数  $\theta$  的变化，二维 Gabor 滤波器能够响应图像不同方向的纹理特征，当  $\theta = 0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{\pi}{2}$  时，二维 Gabor 滤波器函数的实部变化如图 2.3 所示。

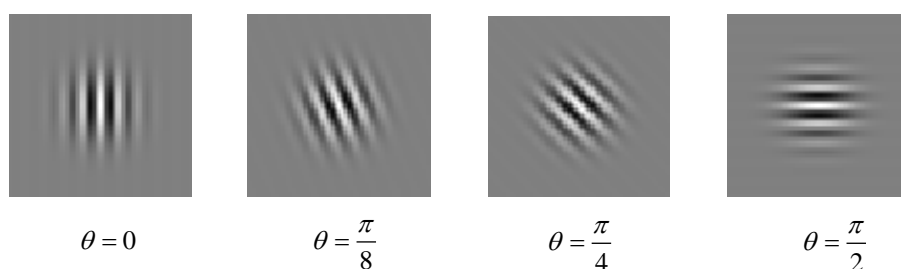


图 2.3 二维 Gabor 滤波器函数实部的纹理方向随参数  $\theta$  的变化

$\lambda$ ：随着波长  $\lambda$  的变化，滤波器能够响应不同频带的图像特征。当  $\lambda = 4, 4\sqrt{2}, 8, 8\sqrt{2}$  时，滤波器函数的实部如图 2.4 所示。

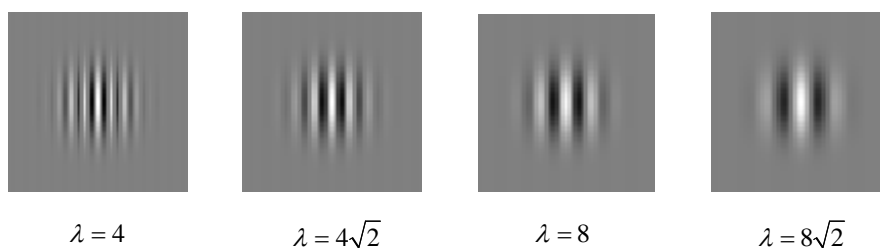
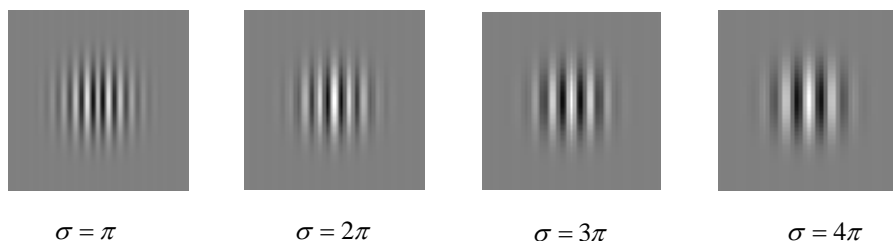


图 2.4 二维 Gabor 滤波器函数实部随参数  $\lambda$  的变化

$\sigma$ ：随着参数  $\sigma$  的增大，滤波器的窗口大小会随之增大，并且当  $\sigma_x$  与  $\sigma_y$  的大小相等时，滤波器的窗口为圆形，否则，滤波器的窗口为椭圆形。当  $\sigma_x = \sigma_y = \pi, 2\pi, 3\pi, 4\pi$  时，滤波器的窗口大小如图 2.5 所示。


 图 2.5 二维 Gabor 滤波器函数实部随参数  $\lambda$  的变化

### 2.2.1.2 图像形态学处理的基本概念

#### 1. 腐蚀

腐蚀主要用于消除边界点，使边界向内部收缩，同样也可以用于消除小且无意义的物体。对结构元素  $B$  平移  $a$  后得到  $B'$ ，若  $B'$  包含于  $X$ ，我们记下这个  $a$  点，所有满足上述条件的  $a$  点组成的集合称为  $X$  被  $B$  腐蚀(Erosion)的结果,数学公式表示如式(2.4)。

$$E(X) = \{a | B' \subset X\} = X \ominus B \quad (2.4)$$

示意图如图 2.7 所示。

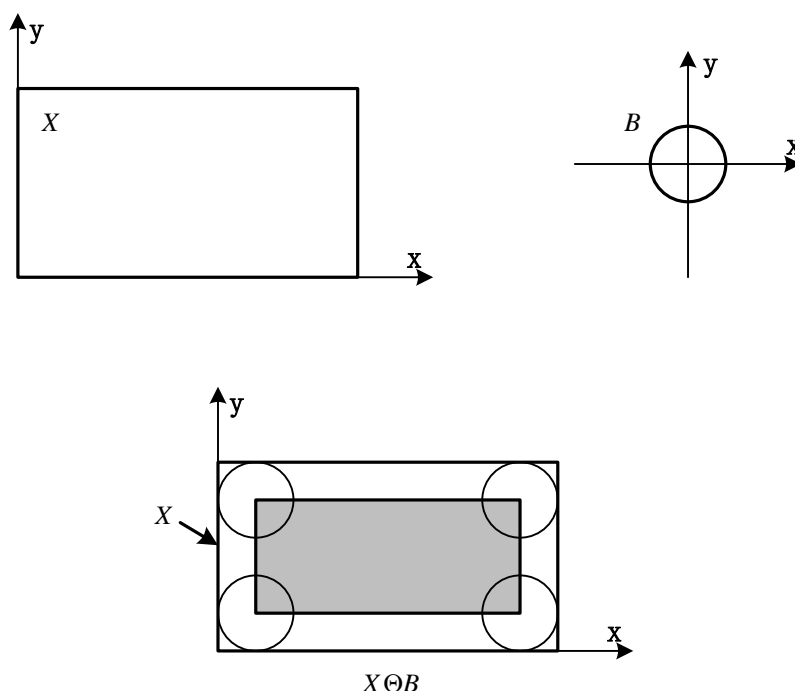


图 2.7 腐蚀的原理图(阴影部分表示腐蚀后的结果)

实际应用中，形态学处理对象一般都是二值图像，那么接下来我们模拟一下在图像处理的过程中，腐蚀对每个像素点的作用，如图 2.8 所示。 $X$  表示被处理的图像，其中黑色的点表示需要处理的像素点， $B$  表示结构元素。腐蚀的过程就是拿  $B$  的中心点分别与  $X$  上的点对比，如果  $B$  上的所有的点都能够在  $X$

的范围内，则保留改点，否则去掉这个点。

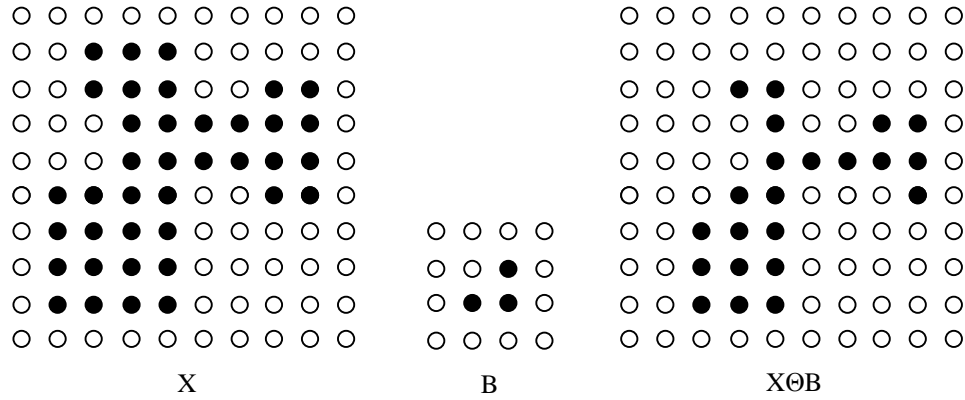


图 2.8 图像处理中腐蚀过程示意图

(左图表示被处理的图像，中间表示结构元素，右图为腐蚀后的结果)

## 2.膨胀

膨胀是将与物体接触的所有背景点合并到该物体中，使边界向外部扩张的过程。可以用来填补物体中的空洞。把结构元素  $B$  平移  $a$  后得到  $B'$ ，若  $B'$  击中  $X$ ，则记下这个  $a$  点。所有符合上述条件的  $a$  点组成的集合称为  $X$  被  $B$  膨胀的结果。用公式表示如式(2.5)。

$$D(X) = \{a | B' \uparrow X\} = X \oplus B \quad (2.5)$$

示意图如图 2.9 所示。

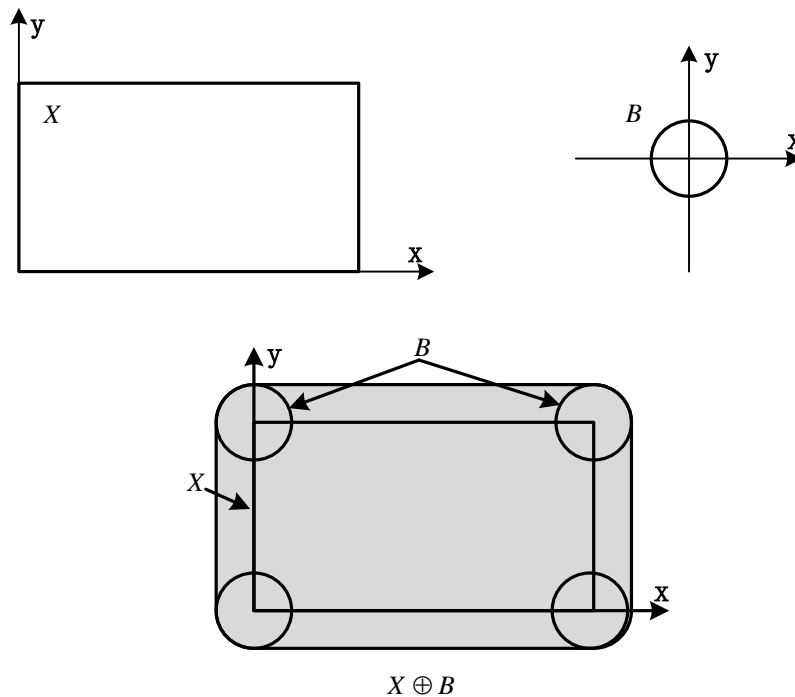


图 2.9 膨胀的原理图(阴影部分表示膨胀后的结果)

图像处理过程中，膨胀对各个像素的作用如图 2.10 所示。其中  $X$  表示被处理的图像。类似于腐蚀过程，拿  $B$  的中心点和  $X$  上的点及  $X$  周围的点一个一

个地对比,如果  $B$  上有一个点落在  $X$  的范围内,则该点就为黑色有效的像素点。

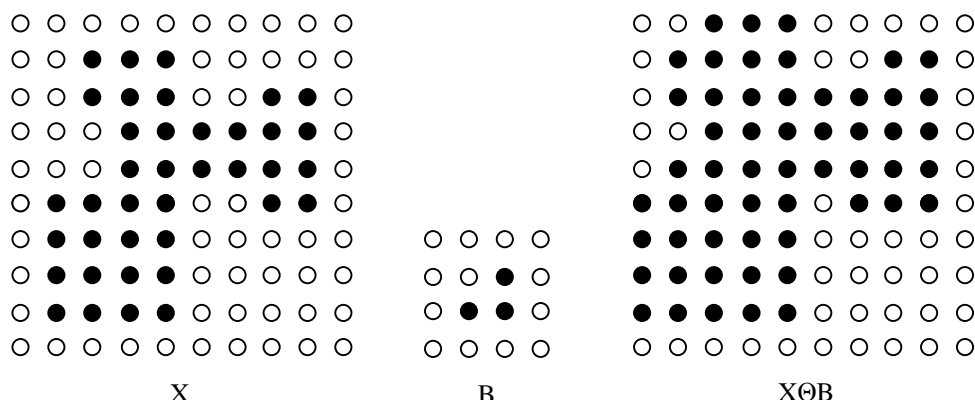


图 2.10 图像处理中膨胀过程示意图

(左图表示被处理的图像, 中间表示结构元素, 右图为膨胀后的结果)

### 2.2.1.3 眼睛定位

为了准确定位眼睛和嘴部的位置, 本文先使用 Gabor 滤波器增强眼睛和嘴部的边缘信息。根据图 2.3, 显然当图像的边缘特征的方向与 Gabor 滤波器核函数的方向垂直时, 滤波器有最强的响应, 图 2.6 表示滤波器核函数的方向角  $\theta$  取不同的值所对应的人脸响应结果。从图中可以看出, 在单组滤波器中, 当  $\theta = \frac{\pi}{2}$ , 得到了最理想的边缘信息, 这是因为眼睛和嘴部边缘信息主要集中在水平方向上, 而组合多组滤波器虽然能够得到更多面部边缘的细节, 但同样也带来了很多的干扰, 所以本文取  $\theta = \frac{\pi}{2}$ ; 根据人脸图像的尺寸, 需要选择合适大小的滤波窗口, 取  $\sigma_x = \sigma_y = 1.44$ ; 同时基于面部器官的边缘长度确定滤波器的波长  $\lambda = 8$ 。

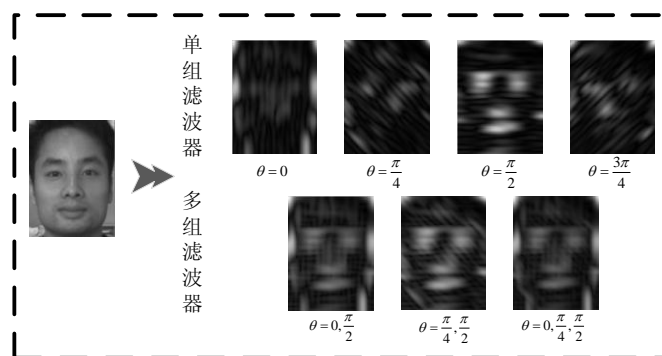


图 2.6 人脸的响应图像随核函数方向角  $\theta$  的变化

本小节主要研究对眼睛的定位, 首先采用 Gabor 滤波器增强面部图像的边缘信息, 并且为了验证算法的普适性, 实验部分考虑了睁眼和闭眼两种情况, 其算法流程如图 2.11 所示。实验结果表明无论眼睛是睁开还是闭合状态, 其边缘增强之后的图像均能将面部主要的特征区域(眉毛、眼睛、鼻头、嘴巴)与

脸颊、额头等区域区分开来。另外，人的面部器官分布一般也遵循一定的几何规律（三庭五眼），所以根据几何特征可以在脸部区域粗略地确定眼睛的位置，对分割出来的图像二值化处理，所得的二值图像通常包含眼睛，部分眉毛和部分鼻子。在此基础上，经过一些简单的形态学处理<sup>[38]</sup>就可以准确的确定眼睛的位置。

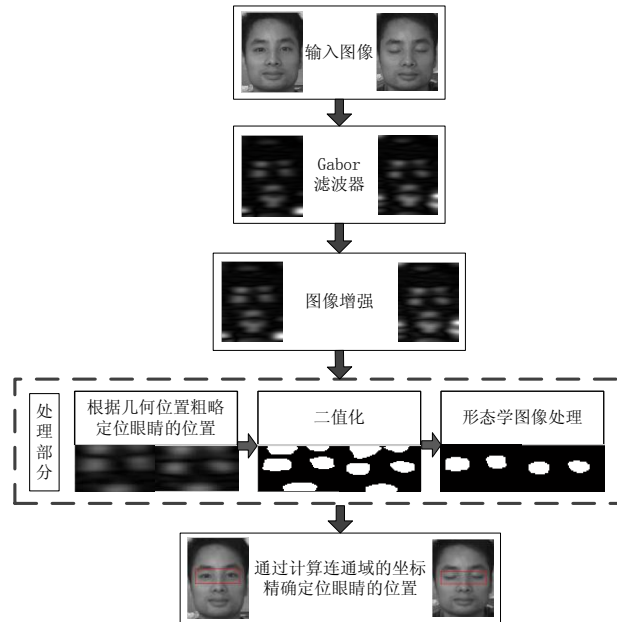


图 2.11 眼睛定位的流程图

### 2.2.2 眼睛的状态判别

人在疲劳的状态下，眨眼的频率和眼睛闭合的时间均会发生较大的变化，所以判断眼睛的状态是驾驶员疲劳状态判定最为直观和重要的依据，因此准确判断驾驶员眼睛的状态对于整个系统非常重要。目前对于眼睛状态判定最简单快速的方法是投影法，根据眼睛在水平方向上的投影判断眼睛的开闭程度，但是该方法对于光照变化非常的不稳定，并且如果驾驶员与传感器之间的距离发生变化时其判定的阈值就不再准确，所以该方法不太适合运用在实际应用中。本文将采用鲁棒性较高的特征分类的方法判断眼睛的状态：通过使用一种旋转不变的 LBP 特征，构造一个高斯金字塔<sup>[39]</sup>描述眼睛，利用 SVM<sup>[1]</sup>训练出一个线性分类器判断眼睛的开闭状态。

#### 2.2.2.1 LBP 算子的简介

LBP (Local Binary Pattern) 是一种用来描述图像局部纹理特征的算子，它是由 Oulu 大学的 T.Ojala 等人<sup>[40]</sup>于 1996 年提出的，其基本的思想是把中心像素点的灰度值当作阈值，并与它的邻域内的像素点比较得到的编码值用来表述其局部纹理特征。LBP 算子作为目前最常用的纹理特征描述算子之一，其主要的优点包括：(1) LBP 算子具有灰度不变性，也就是对于光照的变化不敏感。

(2) LBP 算子的提取过程没有任何其它预设的参数变量，因此人为干涉的因素就比较少。(3) 计算速度非常快。只需要比较中心像素点与相邻像素点的灰度值，然后再编码即可。由于它以上的优点，所以已经被广泛地应用于纹理分类、人脸识别等领域。

### 1. 基本的 LBP

基本的 LBP<sup>[41]</sup>算子是将 3\*3 矩阵中的中心灰度值与其邻域中的 8 个灰度值进行比较，大于中心灰度值的记为 1，否则为 0，由此可以得到一个 8 位的二值编码，其值记为该中心点的编码值，基本原理如图 2.12 所示，计算过程如下所示：

中心像素点的灰度值为 123，其相邻的八个像素点的灰度值分别为 145、132、124、114、134、34、45、31。因此其 LBP 算子的编码为 00001111，特征值为： $128*0+64*0+32*0+16*0+8*1+4*1+2*1+1*1=15$ 。

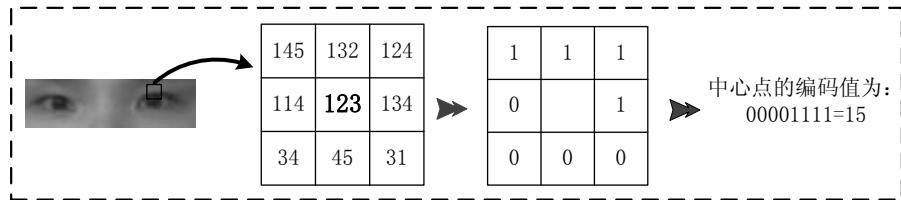


图 2.12 LBP 算子的原理

### 2. LBP 算子的扩展

#### (1) 圆形 LBP 算子

基本的 LBP<sup>[41]</sup>算子由于只能覆盖相邻的 8 个像素点，所以不能满足不同尺寸和频率的纹理需要。为了适应不同尺度的纹理特征，Ojala 对 LBP 算子又进行了发展和改进，用圆形邻域代替了方形邻域，并且允许在邻域半径内有任意多的采样点，为了表示本文用  $(P, R)$  来表示一个圆形的邻域，其中  $P$  表示邻域内采样点的个数， $R$  表示邻域的半径，如图 2.13 所示。以图像中的某一个像素点  $(x, y)$  为中心的 LBP 特征值计算公式如 (2) 所示。

$$LBP_{P,R}(x, y) = \sum_{p=0}^{P-1} S(f(x, y) - f(x_p, y_p)) 2^p \quad (2.6)$$

其中  $S(z) = \begin{cases} 1, z \geq 0 \\ 0, z < 0 \end{cases}$ ， $f(x, y)$  表示该像素点的灰度值， $(x_p, y_p)$  表示其邻域的采样点。



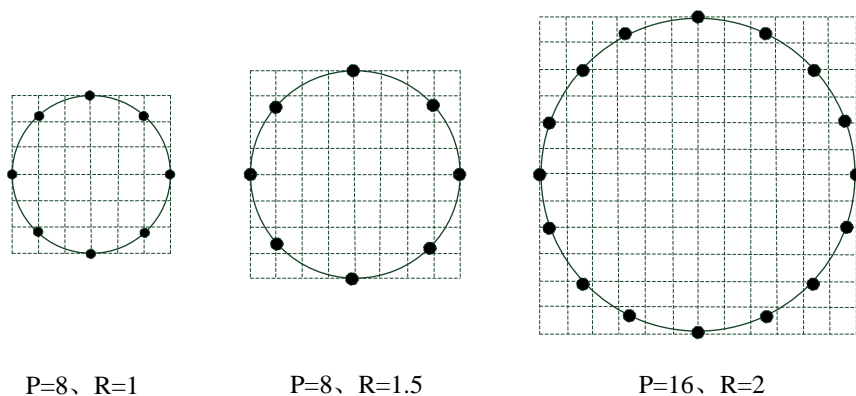


图 2.13 圆形 LBP 算子

圆形 LBP 算子与基本的 LBP 算子在原理上是一致的，但是由于其半径取值的灵活性，圆形 LBP 能够适应更多尺度大小的纹理特征，需要注意的是，如果其采样点不在图像坐标上，则采用它们的内插点代替。另外，圆形 LBP 算子对于 LBP 算子的发展有着决定性的作用，后来研究的很多关于 LBP 算子的特性均是在圆形 LBP 基础上发展的，例如旋转不变性等。

## (2) LBP 的等价模式

LBP 算子能够产生多种不同的二进制模式，对于半径为  $R$  的圆形邻域内含有  $P$  个采样点的 LBP 算子可以产生  $2^P$  种模式，因此随着邻域集内采样点数的增加，二进制模式的种类会迅速增加，并且其中很多模式在图像中出现的概率很低，这样不仅不利于纹理特征的提取、分类，而且会浪费很多的资源。

为了解决二进制模式过多的问题，Ojala 又提出了 LBP 的“等价模式”（Uniform Pattern）。在实际图像中，绝大多数 LBP 模式最多只包含两次从 1 到 0 或从 0 到 1 的跳变，甚至有时这样的模式能够包括整幅图像 90% 以上的模式类别，因此，Ojala 定义当某个 LBP 模式所对应的二进制编码中 0 到 1 或 1 到 0 的跳变最多只有两次时，则认为该模式为 LBP 等价模式类。如 00000000（0 次跳变），11110000（一次跳变），00110000（两次跳变）都是等价模式类。另外其他非等价模式类归为一类。例如 11011110（三次跳变），11010011（四次跳变）。

等价模式大大减少了 LBP 算子的模式种类，对于含有  $P$  个采样点的 LBP 算子由原来的  $2^P$  减少为  $P(P-2)+2$  种。对于  $3 \times 3$  邻域内 8 个采样点来说，二进制模式由原始的 256 种减少为 58 种，如图 2.14 所示，这样对 LBP 特征作统计直方图的时候大大减少了特征的维数，从而可以减少高频噪声带来的影响。

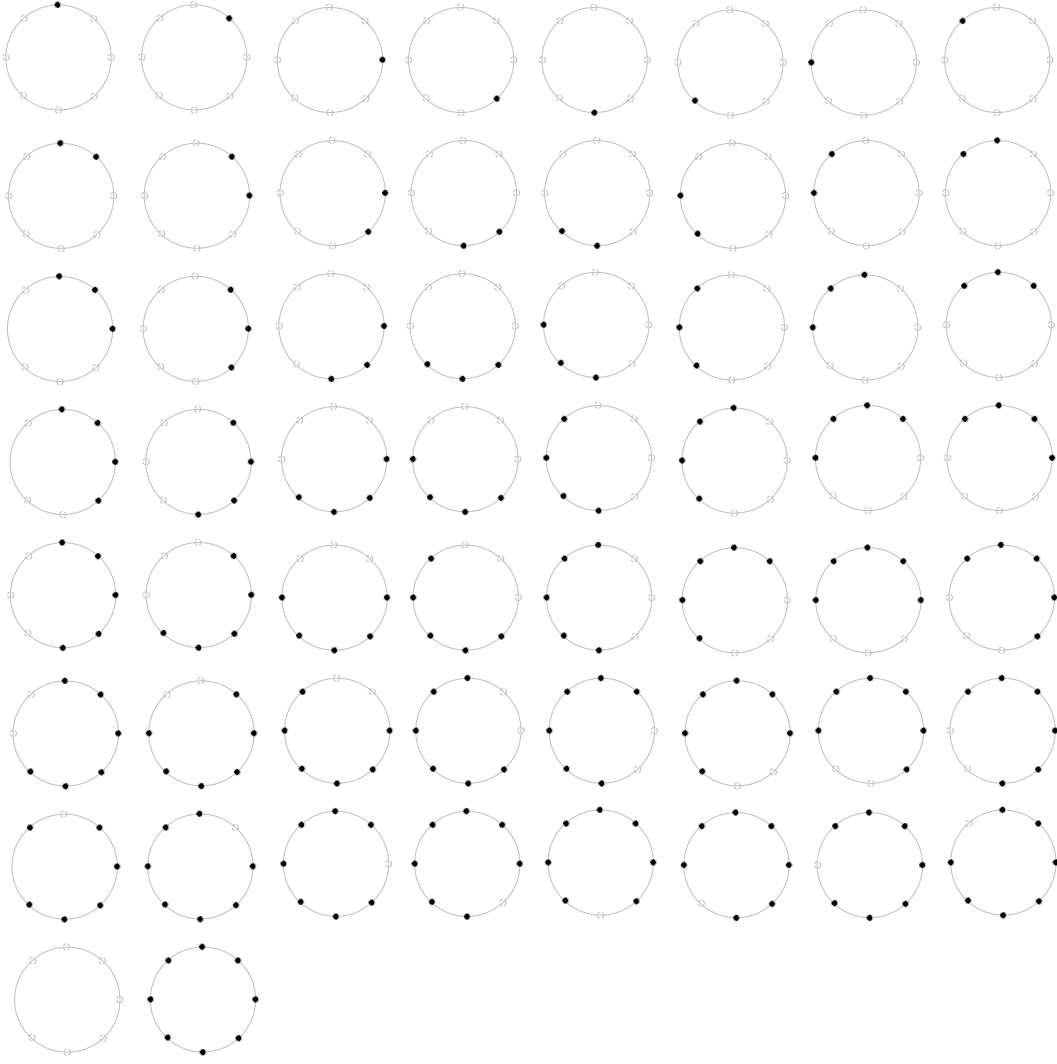


图 2.14 LBP 算子的均匀模式

### (3) 旋转不变的 LBP 算子

如前文所描述，LBP 算子是灰度不变的，但却不是旋转不变的。图像的旋转就会得到不同的 LBP 值。而在实际情况中图像的旋转是经常发生的事情，所以旋转不变的 LBP 算子对于实际应用有更好的普适性。

Zhao<sup>[42]</sup>等人提出了一种旋转不变的 LBP 算子，即对等价模式  $U_p(n, r)$  的统计直方图  $h_l(U_p(n, r))$  作离散傅里叶变换，记作  $H(n, u)$ ，即  $H(n, u) = \sum_{r=0}^{P-1} h_l(U_p(n, r)) e^{-i2\pi ur/P}$ 。当输入变量发生一个旋转  $a$  之后，统计结果记为  $h_l(U_p(n, r-a))$ ，则  $H'(n, u) = H(n, u) e^{-i2\pi ua/P}$ 。由此不难发现  $H'(n, u) \bullet \overline{H'(n, u)} = H(n, u) e^{-i2\pi ua/P} \overline{H(n, u) e^{-i2\pi ua/P}} = H(n, u) \bullet \overline{H(n, u)}$ 。（其中  $\overline{H(n, u)}$  为  $H(n, u)$  的共轭复数）此结果中明确表明  $H(n, u) \bullet \overline{H(n, u)}$  的值与样本旋转的角度无关，所以 LBP 特征  $|H(n, u)| = \sqrt{H(n, u) \bullet \overline{H(n, u)}}$  是旋转不变的。

### 2.2.2.2 基于 3 层 LBP 特征金字塔的眼睛状态判别

本文采用特征分类的方法对眼睛状态分类。首先对所有的训练样本提取上述旋转不变的 LBP 特征，但是如果只提取单一尺度下的 LBP 特征，会导致丢失很多细节信息，所以本文构造一个三层的 LBP 特征金字塔来描述眼睛图像，这样可以增强特征的描述能力。其具体的提取过程描述如下：本实验采用的训练样本大小均为  $80 \times 20$ ，首先在原图像 ( $80 \times 20$ ) 上提取一次 LBP 特征作为金字塔的第一层，然后采用高斯滤波器对图像平滑滤波，再对图像降采样缩放到  $40 \times 10$ ，再对图像提取一次 LBP 特征作为金字塔的第二层，类似地，对上一层的图像作类似处理，再提取 LBP 特征作为金字塔的第三层。最后，组合三个部分的特征，作为最后的描述特征，然后再通过 SVM 训练一个能够判断眼睛的状态的线性分类器。具体过程如图 2.15 所示。

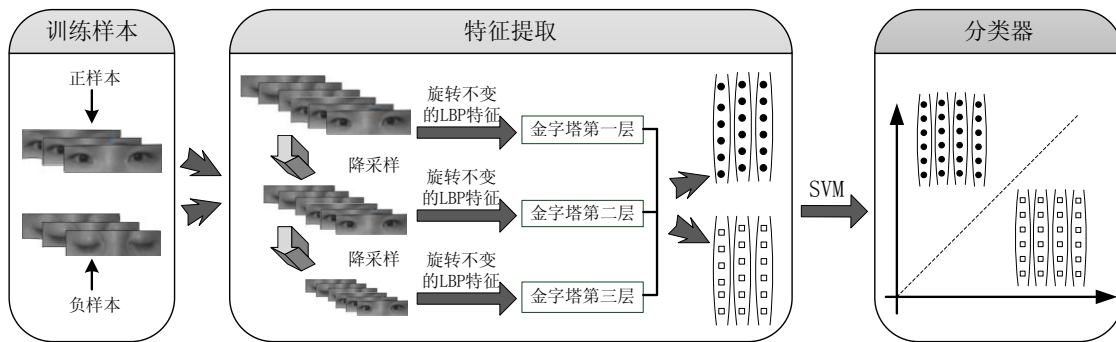


图 2.15 眼睛状态判别分类器的训练框图

### 2.2.2.3 眼睛状态判别实验结果

本实验是为了验证上述眼睛状态判别算法的准确性和鲁棒性。实验中，为了说明文中所提出的特征描述方法的有效性，分别用一层、二层、三层高斯金字塔进行特征描述，对包括来自不同的 10 个人的 943 个正样本和 851 个负样本（其中每个样本的大小为  $80 \times 20$ ）进行四种方式的特征提取，各种特征的维数如表 2.1 所示。随机从中抽取 700 个正样本和 700 个负样本组成训练样本，分别根据每种特征训练 SVM 线性分类器，再把剩余的 243 个正样本和 151 个负样本作为测试样本做验证性实验。以此重复 15 次实验，实验结果统计如图 2.16 所示。

表 2.1. 眼睛状态判别的特征维数

特征类型	特征维数
一层金字塔: (8,1)	76
二层金字塔: (8,1)	152
三层金字塔: (8,1)	228
(8,1)+(16,2)	352

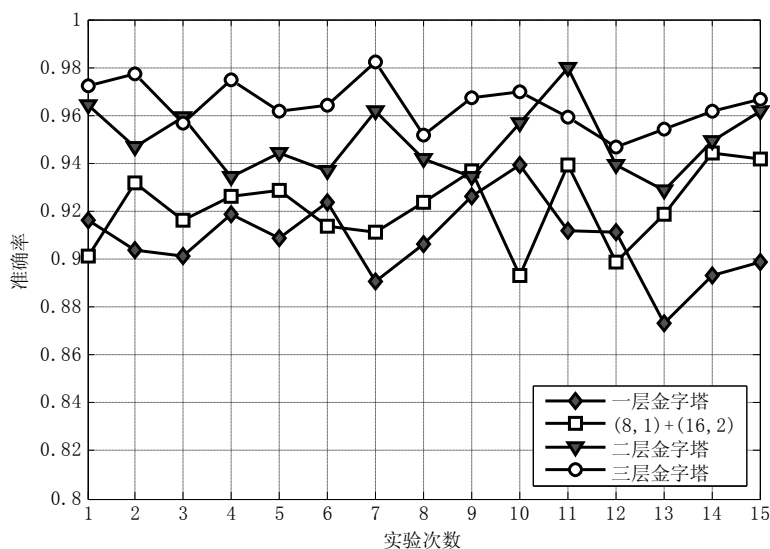


图 2.16 眼睛状态分类实验结果

从图 2.13 中可以看出,单层的(8,1)特征虽然特征维数比较低,但是只包含了图像大部分的纹理信息,所以准确相对较低。改进的方法可以分别从尺度上和空间上来增加,比较红色和深绿色曲线,容易看出在尺度上增加的效果更好,这是因为图像在缩放之后(8,1)特征比原图上取(16,2)特征包含更多能够描述图像的纹理信息,同时在特征的维数上,尺度模型也明显的优于空间模型。比较蓝色和红色曲线,说明三层金字塔通常具有更高的准确率,而且三层金字塔的特征维数也比较低,计算量小,判别速度很快。其中第 2 组实验的 ROC 曲线如图 2.17 所示,根据每种特征的 AUC 值可以证明此类 LBP 特征具有很强的描述性,而且三层金字塔相较之下性能最佳。

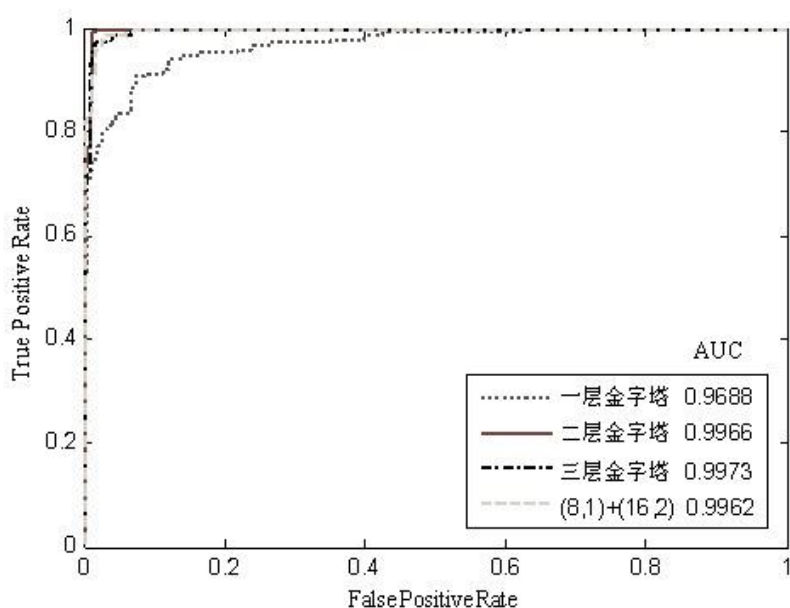


图 2.17 第 2 组实验的 ROC 曲线

## 2.3 嘴部的定位及状态判别

### 2.3.1 嘴部的定位

疲劳的状况下，人会不停的打哈欠，所以嘴部的状态也是判断驾驶员是否疲劳的重要指标。算法首先需要确定嘴部的位置，类似于对眼睛的定位，先采用 Gabor 滤波器和梯度图像<sup>[43]</sup>增强嘴部的边缘信息，再根据人脸的几何特征对嘴部粗略定位，再将其分割出来的图像二值化，并且经过形态学处理消除其它部分的干扰，最终就能精确分割出嘴部的二值图像。算法流程图如图 2.18 所示。其中二值图像的形态学处理过程如图 2.19 所示，首先对分割出来的嘴部二值图像作腐蚀处理，消除一些干扰区域与嘴部的粘连，再去除一些粘连在边界上和面积小的连通域，然后对图片作膨胀处理，填补嘴部中出现空洞的连通域，最后为了尽量保证嘴部状态的真实性和完整性，还需要对膨胀之后的图像作腐蚀处理，其中两次腐蚀操作的结构元素的半径设置为膨胀操作结构元素的一半。

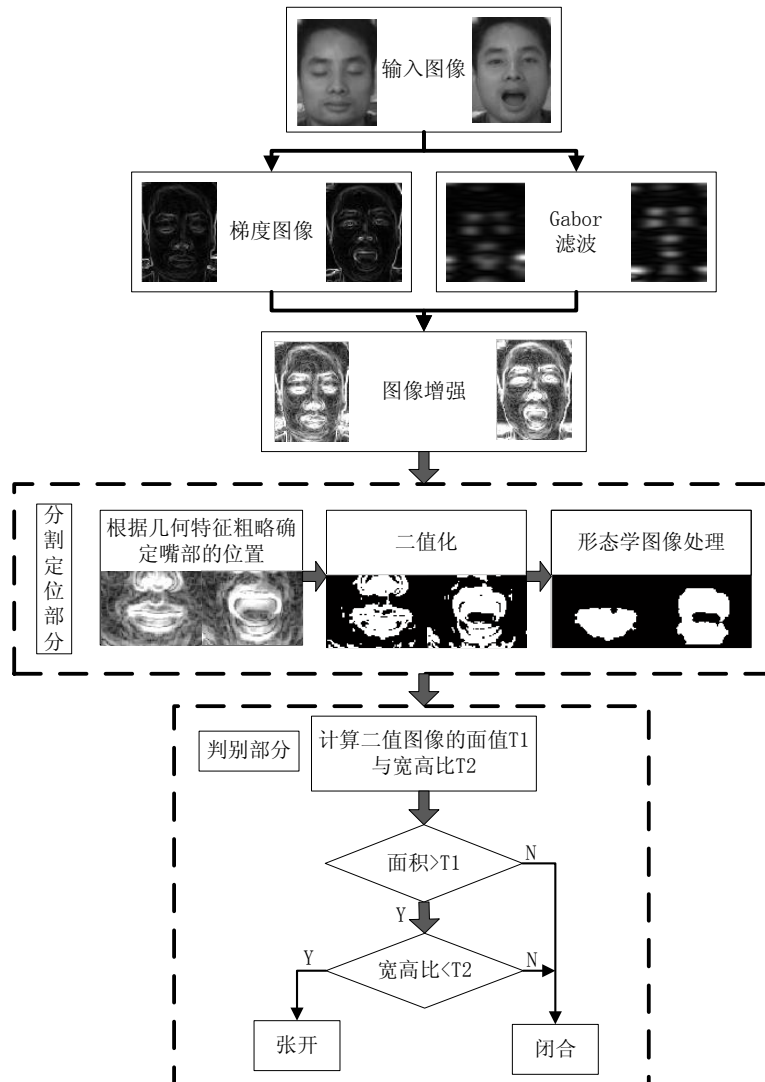


图 2.18 嘴部定位与状态判别的流程图



图 2.19 嘴部图像的形态学处理过程

### 2.3.2 嘴部状态判别及实验结果

本文根据两个参数来判别嘴部开闭状态：嘴部的面积以及宽高比。当嘴部的二值图像的像素个数小于  $T_1$  时，将其判断为闭合状态；对于大于此阈值的样本，再根据其嘴部二值图像的长宽比来进一步判断，如果长宽比大于  $T_2$  时，依然判断它是闭合状态，反之被判断是张开的。根据上述分割方法，统计来自 10 个人的 2000 个样本（张开、闭合样本各 1000 个，大小为  $160 \times 120$ ）的嘴部面积以及宽高比的均值，最终确定两个判断阈值  $T_1=130$ ， $T_2=1.6$ 。验证实验包括 720 帧张开状态和 540 帧闭合状态的样本，实验结果如表 2.2 所示。

表 2.2 嘴部状态判别的实验结果

嘴巴的状态	测试样本数	识别正确数	准确率
张开	720	666	92.5%
闭合	540	496	91.85%

从表 2.2 可以看出本文采用的方法对嘴部开闭状态判别有较高的准确率，但是当光线较暗时，会导致不能准确分割出嘴部图像而导致最终的判断错误，另外所取的判断阈值也只是对于大部分人是合适的，所以个体差异也会对判断结果造成一定的影响。

## 2.4 本章小结

本章主要介绍了面部局部特征（眼睛和嘴部）的定位及状态判别。首先，根据基于 haar 特征的 Adaboost 算法检测人脸之后，采用 Gabor 滤波器增强整体面部的边缘信息，并结合面部的几何特征准确定位眼睛和嘴部。采用三层的 LBP 特征金字塔描述眼睛，并且结合 SVM 训练一个线性分类器判断眼睛的开闭状态，同时根据分割出来的嘴部二值图像的面积和宽高比来判断嘴部的开闭状态。实验结果表明，无论眼睛和嘴部状态判别的准确率均在 90% 以上，因此说明本文的方法具有很高的准确率和鲁棒性。本章中对于眼睛和嘴部状态判别的结果都将作为驾驶员疲劳状态判定的依据，具体计算过程在第四章中将会详细描述。

## 第3章 整体面部特征的疲劳状态判别方法

上一章主要研究了面部局部特征状态与驾驶员的疲劳状态的联系，而本章将根据整体面部图像判断驾驶员的疲劳状态。当人处于疲劳状态时，面部表情会明显的异于正常状态，因此面部表情状态可以作为疲劳状态判定的一个依据。本文中，根据需要将驾驶员的表情分为两类：疲劳和非疲劳，并采用多模态特征的距离测度学习算法对疲劳表情分类。在图像分类算法中，根据距离度量两幅图像之间的相似度是非常普遍的，而距离测度学习<sup>[44]</sup>是通过有限的训练样本学习一个更适合样本分类的距离测度空间，从而提高分类效果。本文采用多模态特征指的是根据多种特征下的平均距离对样本分类，这样可以提高系统的准确率和鲁棒性。本章的算法流程图如图 3.1 所示。

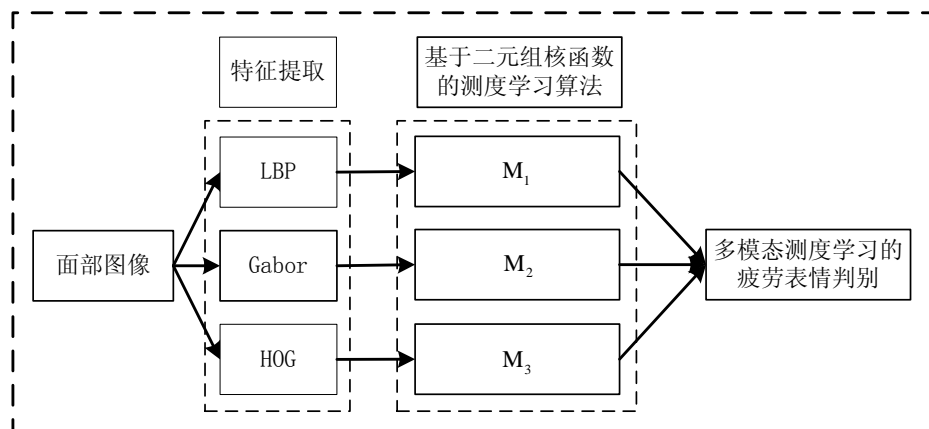


图 3.1 整体面部特征的疲劳状态判别流程图

### 3.1 整体面部表情的特征提取

#### 1. 局部二值模式（LBP）

LBP 算子是一种用于局部纹理特征提取的算子，很适合用来提取图像中局部的细节，如前文所描述，LBP 特征也已经被广泛的应用在特征分类等领域，所以本文对整体面部图像提取 LBP 特征描述驾驶员的面部表情。其中 LBP 算子的介绍以及提取方法已在前文中详细介绍，在本章节中不再赘述。

#### 2. Gabor 特征

Gabor 滤波器对于图像的边缘非常敏感，能够提供良好的方向和尺度的选择特性，因此对于面部主要特征和表情的变化具有很强的描述性，能够捕捉对应于空间和频率的局部信息。而且 Gabor 特征对于光照的变化不敏感，而驾驶环境的光照情况会随外界环境的变化不断变化所以光照不变性是非常有必要的。所以本文也选择采用 Gabor 特征来描述驾驶员的面部表情。

### 3.方向梯度直方图（HOG:Histogram of Oriented Gridients）

HOG<sup>[45]</sup>是一种常用于描述图像局部纹理的特征，如它的名字所言，特征需要先对整幅图像分区，然后分别计算每一区域中不同方向上的梯度值，然后作统计直方图，最后再将各个部分组合在一起作为 HOG 特征。具体的原理和计算方法如下所示：

#### （1）分割图像

HOG 是一种局部特征，因此如果对整幅图片提取 HOG 特征并不能得到很好的描述效果。从信息论的角度，一张 480\*320 的图片，总共有十几万的像素点，也就是说原始的特征维数有十几万，但是用 HOG 特征描述整张图像时，最多也就 360 个 bin，这样就会损失很多的细节。因此，一般来说，只有图像区域比较小的情况，基于统计原理的直方图对于该区域才有表达能力。如果图像区域比较大，那么两个完全不同的图像的 HOG 特征，也可能很相似，但是如果区域较小，这种可能性就很小。因此需要对样本图片分成若干区块，然后对每个区块计算 HOG 特征，最后再组合所有小区块的 HOG 特征成最终图像特征，这样的组合也包含了图像的几何（位置）特性。例如，人脸图像的上部分区块一般都是眼睛而下部分的特征则是嘴部的 HOG 特征。

其中分区块的方法有交叠和非交叠两种，如图 3.2 所示。交叠，顾名思义指的是分割出的区块互相交叠，有重合的区域，图 3.2（1）图中的阴影部分则表示相邻两个 cell 的交叠区域。非交叠则指的是所分的区块不交叠，没有重合的区域如图 3.2（2）所示。这两种分块的方法各有利弊，其中交叠的方法能够提取更完整的 HOG 特征，这样可以增强特征的描述性，因此提高分类效果。例如对一张人脸分块，非交叠的方式很大可能会把一只眼睛分在两个 cell 中，这样的描述破坏了人脸的特征，不利于分类，但是如果采用交叠的分块方式，就可以将一只完整的眼睛分在下一个 cell 中。但是很明显，交叠的分块方式需要计算很多的重复区域，因此会增加整个特征提取的计算量，有些时候当样本比较大时甚至会爆发维数灾难。因此采用哪种分块方式需要根据实际应用的情况来做选择。

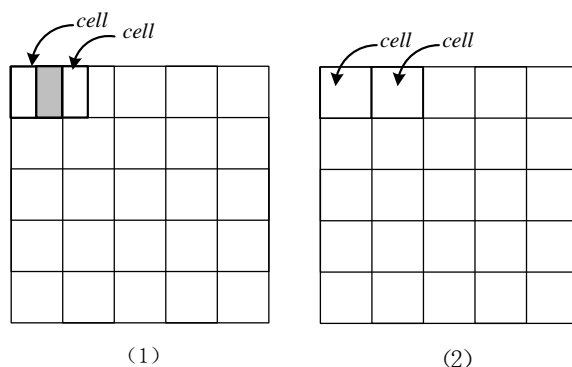


图 3.2 HOG 特征的分块方法



## (2) 计算每个区块的方向梯度直方图

将图像分割后，接下来就要计算每个 cell 的方向梯度直方图。步骤如下：利用任意一种梯度算子，例如：sobel, laplacian 等，对该 cell 进行卷积，计算得到每个像素点处的梯度方向和幅值。具体公式如下：图像任意像素点  $(x, y)$  的梯度<sup>[46]</sup>分别为式(3.1)和(3.2)。

$$I_x(x, y) = H(x+1, y) - H(x-1, y) \quad (3.1)$$

$$I_y(x, y) = H(x, y+1) - H(x, y-1) \quad (3.2)$$

其中  $I_x(x, y)$  和  $I_y(x, y)$  分别表示像素点  $(x, y)$  处水平方向的梯度和垂直方向的梯度。而该像素点处的梯度幅值与梯度方向分别如式(3.3)和(3.4)所示。

$$M(x, y) = \sqrt{I_x^2(x, y) + I_y^2(x, y)} \quad (3.3)$$

$$\theta(x, y) = \tan^{-1} \frac{I_y(x, y)}{I_x(x, y)}, \theta \in [0, 360^\circ) \text{ or } [0, 180^\circ) \quad (3.4)$$

求解图像中每个 cell 的梯度方向，将  $360^\circ$  根据需要分割成若干个 bin。例如，如果分割成 12 个 bin，则每个 bin 包含  $30^\circ$ ，根据每个像素点的梯度方向，利用双线性内插法将其幅值累加到直方图中。也就是说，如果这个像素的梯度方向是  $30^\circ \sim 60^\circ$ ，则在直方图第 2 个 bin 的计数加 1，统计完该 cell 中所有的像素点就可以得到该 cell 的梯度方向直方图，统计过程如图 3.3 所示。直方图包括 12 个 bin 也就是说该 cell 的特征向量为 12 维。

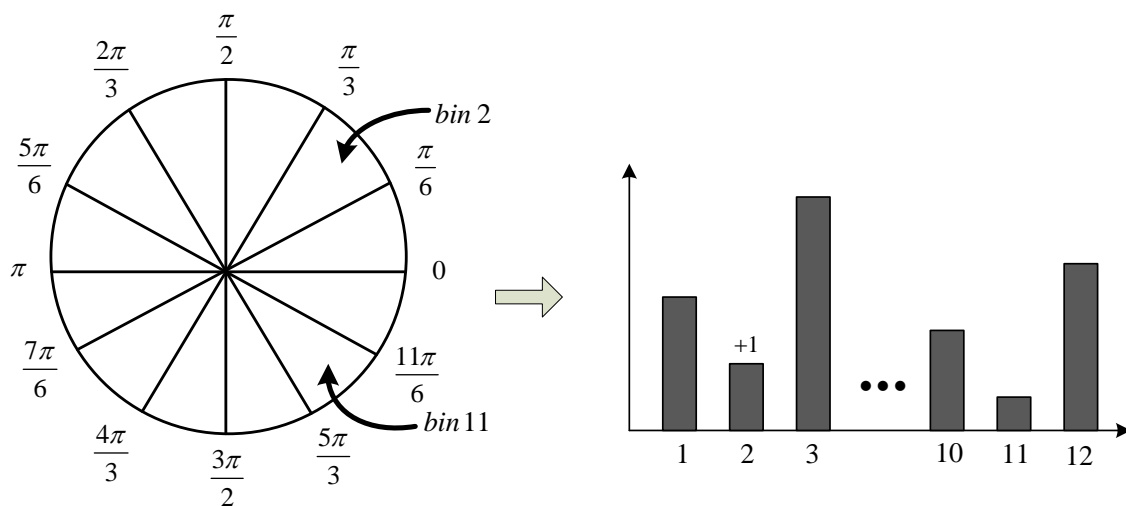


图 3.3 像素点梯度方向的统计规则

## (3) 组合细胞单元(cell)成块(block)，再归一化块内的梯度直方图

将相邻的几个 cell 组合成一个 block，如图 3.4 所示。组合该 block 中包括的所有 cell 的 HOG 特征，再对 block 的特征进行颜色、亮度的归一化。这一步主要是用来去掉光照、阴影等影响，对于光照影响不剧烈的图像，可以省略这个步骤。

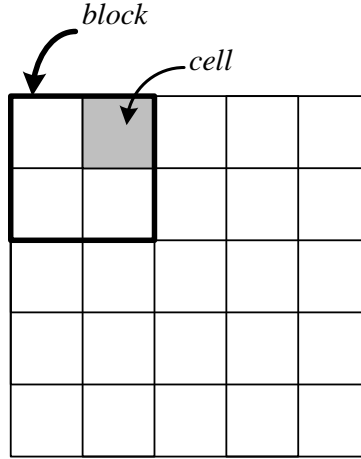


图 3.4 图像分块中 cell 与 block 的联系

#### (4) 组合整幅图像的 HOG 特征

将从每个 block 中提取出的 HOG 特征首尾相连, 组合成一个大的一维向量, 这就是该图像的 HOG 特征。该特征的维数计算如下: 如果把一副  $64 \times 128$  样本图像分割为 32 个 cell (每个 cell 包括  $16 \times 16$  个像素点), 然后相邻的  $2 \times 2$  个 cell 组合成一个 block, 并且每个 cell 包括 12 个 bin, 则每个块内的特征包括  $4 \times 12 = 48$  维。以 non-overlap 为例, 样本图像的特征维数就是:  $48 \times 8 = 384$  维。

### 3.2 基于距离测度学习的整体面部疲劳表情判别

#### 3.2.1 距离测度学习的基本概念

##### 3.2.1.1 测度空间

如果存在一个映射  $D: \Omega \times \Omega \rightarrow \mathbb{R}_0^+$  ( $\Omega$  表示一个向量空间), 对于  $\forall \vec{x}_i, \vec{x}_j, \vec{x}_k \in \Omega$ , 如果它满足下面四个性质:

1. 三角不等性:  $D(\vec{x}_i, \vec{x}_j) + D(\vec{x}_j, \vec{x}_k) \geq D(\vec{x}_i, \vec{x}_k)$
2. 非负性:  $D(\vec{x}_i, \vec{x}_j) \geq 0$
3. 对称性:  $D(\vec{x}_i, \vec{x}_j) = D(\vec{x}_j, \vec{x}_i)$
4. 可区别性:  $D(\vec{x}_i, \vec{x}_j) = 0 \Leftrightarrow \vec{x}_i = \vec{x}_j$

则这个映射关系为一个测度, 严格来讲, 如果一个映射只满足前面三个性质, 则称之为伪测度。一般情况下, 为了简化复杂度, 通常把伪测度也认为是一个测度。

### 3.2.1.2 空间距离与距离测度学习

在欧氏空间内，两个向量之间的距离计算公式如式(3.5)所示。

$$D(\vec{x}_i, \vec{x}_j) = (\vec{x}_i - \vec{x}_j)^T (\vec{x}_i - \vec{x}_j) \quad (3.5)$$

如果经过一个线性变换  $\vec{x}' = \mathbf{L}\vec{x}$  之后，在新的空间，这两个向量之间的距离计算如式(3.6)所示。

$$D_L(\vec{x}_i, \vec{x}_j) = \|\mathbf{L}(\vec{x}_i - \vec{x}_j)\|_2^2 = (\vec{x}_i - \vec{x}_j)^T \mathbf{L}^T \mathbf{L} (\vec{x}_i - \vec{x}_j) \quad (3.6)$$

可以证明，如果  $\mathbf{L}$  是一个满秩的矩阵，则上式定义了一个有效的测度，不然，称之为有效的伪测度。通常情况下，更常定义一个方阵  $\mathbf{M}$  来替代  $\mathbf{L}$ ，如式(3.7)所示。

$$\mathbf{M} = \mathbf{L}^T \mathbf{L} \quad (3.7)$$

因此，两个向量的距离就可以表示为式(3.8)。

$$D_M(\vec{x}_i, \vec{x}_j) = (\vec{x}_i - \vec{x}_j)^T \mathbf{M} (\vec{x}_i - \vec{x}_j) \quad (3.8)$$

由此可见，距离测度学习包括两种方法，一种是学习测度矩阵  $\mathbf{M}$  或者学习一个线性变换的矩阵  $\mathbf{L}$ ，这两种情况的最终的效果是一样的，但是求解的过程却是有区别的。如果求解  $\mathbf{L}$ ，对于  $\mathbf{L}$  本身是没有任何约束条件的，但是如果直接求解  $\mathbf{M}$ ，则要求是一个半正定的方阵。而且若已知变换矩阵  $\mathbf{L}$  则可以确定  $\mathbf{M}$ ，但已知测度矩阵  $\mathbf{M}$  却不能确定  $\mathbf{L}$ 。因为只要  $\mathbf{P}\mathbf{P}^T = \mathbf{I}$ ，就都可以满足  $\mathbf{P}\mathbf{M}\mathbf{P}^T = \mathbf{L}\mathbf{L}^T$ 。本文采用的方法则是直接求解测度矩阵  $\mathbf{M}$ 。

### 3.2.1.3 几种常用的距离测度学习方法

在分类算法中，人们经常用图像之间的欧氏距离来衡量这两幅图像的相似度，而很多时候欧氏距离并不能作为一个很好的判定标准，所以距离测度学习的方法就是通过利用已有的样本学习一个测度，将样本映射到一个更容易区分的空间。

#### 1. PCA 主成分分析

PCA<sup>[47]</sup>是将高维空间的样本通过一个线性变换  $\vec{x}_i \rightarrow \mathbf{L}\vec{x}_i$  映射到低维空间，使得变换空间的样本的方差尽可能的大。若原空间的样本均值为  $\vec{\mu} = \frac{1}{n} \sum_{i=1}^n \vec{x}_i$ ，则变换空间的样本均值为  $\mathbf{L}\vec{\mu}$ ，此时的样本方差为式(3.9)。

$$\frac{1}{n} \sum_{i=1}^n (\mathbf{L}^T \vec{x}_i - \mathbf{L}^T \vec{\mu})^2 = \mathbf{L}^T \mathbf{C} \mathbf{L} \quad (3.9)$$

其中  $\mathbf{C} = \frac{1}{n} \sum_{i=1}^n (\vec{x}_i - \vec{\mu})^T (\vec{x}_i - \vec{\mu})$ ，恰好是样本的协方差矩阵。上文已经提到，

需要求解的变换矩阵  $L$  需要满足使变换空间的样本方差最大化，如式 (3.10) 所示。

$$\max_L L^T C L, \text{ s.t. } L^T L = I \quad (3.10)$$

## 2. LDA 线性判别分析

LDA<sup>[48]</sup>就是通过  $\vec{x}_i \rightarrow L\vec{x}_i$  变换之后是的是类间的方差与类内的方差比最大。类间的方差和类内的方差分别由类间和类内的协方差矩阵  $C_b$  和  $C_w$  计算而来，分别如式 (3.11) 和 (3.12) 所示。

$$C_b = \frac{1}{C} \sum_{c=1}^C \vec{\mu}_c \vec{\mu}_c^T \quad (3.11)$$

$$C_w = \frac{1}{n} \sum_{c=1}^C \sum_{i \in \Omega_c} (\vec{x}_i - \vec{\mu}_c)(\vec{x}_i - \vec{\mu}_c)^T \quad (3.12)$$

其中  $\vec{\mu}_c$  表示第  $C$  类的平均值， $\Omega_c$  表示第  $C$  类样本。

经过线性变换之后，类间和类内的协方差矩阵分别是  $L^T C_b L$  和  $L^T C_w L$ ，我们需要做的就是寻找一个最优的  $L$  使得类间和类内的方差比最大。因此我们需要优化的目标函数如式 (3.13)。

$$\max Tr\left(\frac{L^T C_b L}{L^T C_w L}\right), \text{ s.t. } L L^T = I \quad (3.13)$$

## 3. LMNN

LMNN<sup>[49]</sup>通过学习一个线性变换  $L$  ( $M=L^T L$ ) 将样本转换到一个新的空间，从而提升 kNN 分类器的效果。在目标样本  $\vec{x}_i$  的  $k$  邻域内，使得同类样本之间的距离能够尽可能的小，而异类样本  $\vec{x}_j$  与  $\vec{x}_i$  之间的距离则需要满足条件：

$1 + D_M(\vec{x}_i, \vec{x}_j) \leq D_M(\vec{x}_i, \vec{x}_l)$  (其中  $\vec{x}_j$  为  $\vec{x}_i$  的  $k$  邻域中的同类样本)。其损失函数如下式 (3.14) 所示。

$$\begin{aligned} \xi(M) = (1 - \mu)\xi(\text{pull}) + \mu\xi(\text{push}) = (1 - \mu) \sum_{i,j \rightarrow i} D_M(\vec{x}_i, \vec{x}_j) \\ + \mu \sum_{i,j \rightarrow i} \sum_l (1 - y_{ij}) \left[ 1 + D_M(\vec{x}_i, \vec{x}_j) - D_M(\vec{x}_i, \vec{x}_l) \right]_+ \end{aligned} \quad (3.14)$$

其中  $[z]_+ = \max(0, z)$ 。

目标函数包括两项，第一项： $\xi(\text{pull}) = \sum_{i,j \rightarrow i} D_M(\vec{x}_i, \vec{x}_j)$ ，主要是用来惩罚目标样本  $k$  邻域中距离目标样本  $\vec{x}_i$  太远的同类样本。第二项：

$\xi(\text{push}) = \sum_{i,j \rightarrow i} \sum_l (1 - y_{ij}) \left[ 1 + D_M(\vec{x}_i, \vec{x}_j) - D_M(\vec{x}_i, \vec{x}_l) \right]_+$ ，是用来惩罚目标样本  $k$  领域中

距离目标样本  $\vec{x}_i$  太近的异类样本。当异类样本  $\vec{x}_l$  与  $\vec{x}_i$  之间的距离不满足条件时， $\xi(\text{push})$  就会产生一个损失，使得  $\vec{x}_l$  与  $\vec{x}_i$  的距离增大。

实际情况中，通常会引入一个松弛变量  $\xi$  来模拟第二项的效果，因此需要优化的目标函数就可以表示为式(3.15)。

$$\begin{aligned} \min \quad & (1 - \mu) \sum_{i,j \rightarrow i} (\vec{x}_i - \vec{x}_j)^T \mathbf{M} (\vec{x}_i - \vec{x}_j) + \mu \sum_{i,j \rightarrow i, l} (1 - y_{il}) \xi_{ijl} \\ \text{s.t.} \quad & (\vec{x}_i - \vec{x}_l)^T \mathbf{M} (\vec{x}_i - \vec{x}_l) - (\vec{x}_i - \vec{x}_j)^T \mathbf{M} (\vec{x}_i - \vec{x}_j) \geq 1 - \xi_{ijl} \\ & \xi_{ijl} \geq 0 \\ & \mathbf{M} \geq 0 \end{aligned} \quad (3.15)$$

目标函数的实现效果如图 3.5 所示。

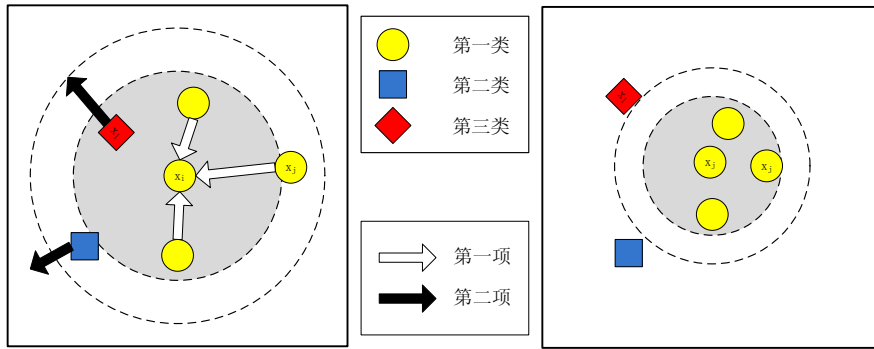


图 3.5 LMNN 的原理示意图

### 3.2.2 多模态测度学习的疲劳表情判别

距离测度学习的关键在根据已有的样本训练求得一个合适的测度矩阵  $\mathbf{M}$ ，而该问题的模型及求解有很多种方法，而本文将引入一种基于二元组核函数的测度学习框架，它能够统一很多目前很优秀的测度学习的算法，其不仅能够缩短训练时间，还能够根据现有的工具包求解。本文也基于此框架学习了三种特征的测度矩阵，并根据三种测度下的平均距离判别驾驶员的疲劳表情。

#### 3.2.2.1 基于二元组核函数的测度学习

目前大部分测度学习的模型都是根据凸或者非凸的优化方法，这样很多方法都不适用于大规模的样本，并且学习一个测度矩阵需要很长的时间。WANG<sup>[50]</sup>提出了一种基于二元组核函数的测度学习框架，很大程度上能够解决了一般模型求解困难、训练时间长的问题，其建模过程如下所示：

##### 1. 二元组样本定义

首先明确一点，测度学习跟一般传统的监督学习的方法不一样，测度学习总是需要讨论同类或者异类之间的关系，因此第一步，需要将训练样本分成一

系列的二元素样本。假设训练样本为  $D=(x_i, y_i)$ ，其中  $x_i$  表示样本， $y_i$  表示该样本的标签。对于每一个训练样本  $x_i$ ，需要找它与它最近的  $m$  个同类样本  $\{x_{i,1}^s, \dots, x_{i,m}^s\}$  和  $n$  个异类样本  $\{x_{i,1}^d, \dots, x_{i,n}^d\}$ ，这样就组成了  $(m+n)$  组二元素样本  $\{(x_i, x_{i,1}^s), \dots, (x_i, x_{i,m}^s), (x_i, x_{i,1}^d), \dots, (x_i, x_{i,n}^d)\}$ 。这样就由这些二元素样本组成了新的训练样本集，表示为  $z_k=(x_{k,1}, x_{k,2}), k=1, 2, \dots, N$ ，样本的标签用  $h_k$  表示，当  $y_i = y_j$  时，这组样本的标签被定义为  $h=1$ ，否则  $h=-1$ 。

## 2. 二元素核函数定义

在文中，WANG 对于二元素的样本定义了一个核函数如式(3.16)。

$$K(x_i, x_j) = \text{tr}(x_i x_i^T x_j x_j^T) = [x_i^T x_j]^2 \quad (3.16)$$

其中  $\text{tr}(\bullet)$  表示矩阵的迹。

如果将这个核函数延伸到二元素样本中，假设两组二元素为  $z_i=(x_{i,1}, x_{i,2})$  和  $z_j=(x_{j,1}, x_{j,2})$ ，则定义相应的二次多项式核为式(3.17)。

$$\begin{aligned} K_D(z_i, z_j) &= \text{tr}\left((x_{i,1} - x_{i,2})(x_{i,1} - x_{i,2})^T (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,2})^T\right) \\ &= \left[(x_{i,1} - x_{i,2})^T (x_{j,1} - x_{j,2})\right]^2 \end{aligned} \quad (3.17)$$

## 3. 基于二元素核函数的测度学习的模型

目标函数的定义包括正则项和损失项：正则项为测度矩阵的 F-范数，如式(3.18)所示，损失项如式(3.19)所示。

$$\gamma_{\text{svm}}(\mathbf{M}) = \frac{1}{2} \|\mathbf{M}\|_F^2 \quad (3.18)$$

$$\rho_{\text{svm}}(\xi) = C \sum_l \xi_l \quad (3.19)$$

其中还包括了两个约束条件，第一个表明同类样本之间的距离需要小于 1，而异类样本之间的距离应该大于 1；第二个表明松弛变量均大于 0，因为很难让所有的样本都满足第一个约束条件，松弛变量就是为了使目标函数的求解更加的灵活，所以目标函数为式(3.20)。

$$\begin{aligned} \min_{\mathbf{M}, b, \xi} \quad & \frac{1}{2} \|\mathbf{M}\|_F^2 + C \sum_l \xi_l \\ \text{s.t.} \quad & h_l \left( (x_{l,1} - x_{l,2})^T \mathbf{M} (x_{l,1} - x_{l,2}) + b \right) \geq 1 - \xi_l \\ & \xi_l \geq 0, \forall l \end{aligned} \quad (3.20)$$

### 3.2.2.2 基于二元素核函数测度学习的模型求解

本小节详述了基于二元素核函数测度学习模型的求解过程，根据推导发现根据定义的二元素核函数可以将目标函数的形式转换成 SVM<sup>[51]</sup>求解模型的形式

式，因此可以采用现有的 **LibSVM** 工具包<sup>[52]</sup>求解，大大降低了求解的难度，比一般的测度学习求解过程要简单很多。

具体的推导过程如下所示，第一步需要求原问题的对偶问题，步骤如下：

(1)原问题的拉格朗日函数：

$$L(\mathbf{M}, b, \xi, \alpha, \beta) = \frac{1}{2} \|\mathbf{M}\|_F^2 + C \sum_l \xi_l - \sum_l \alpha_l \left[ h_l \left( (x_{l,1} - x_{l,2})^T \mathbf{M} (x_{l,1} - x_{l,2}) + b \right) - 1 + \xi_l \right] - \sum_l \beta_l \xi_l \quad (3.21)$$

(2)  $L(\mathbf{M}, b, \xi, \alpha, \beta)$  分别对  $\mathbf{M}, b, \xi$  求导得式(3.22)，(3.23)，(3.24)：

$$\begin{aligned} \frac{\partial L(\mathbf{M}, b, \xi, \alpha, \beta)}{\partial \mathbf{M}} = 0 &\Rightarrow \mathbf{M} - \sum_l \alpha_l h_l (x_{l,1} - x_{l,2})(x_{l,1} - x_{l,2})^T = 0 \\ &\Rightarrow \mathbf{M} = \sum_l \alpha_l h_l (x_{l,1} - x_{l,2})(x_{l,1} - x_{l,2})^T \end{aligned} \quad (3.22)$$

$$\frac{\partial L(\mathbf{M}, b, \xi, \alpha, \beta)}{\partial b} = 0 \Rightarrow \sum_l \alpha_l h_l = 0 \quad (3.23)$$

$$\frac{\partial L(\mathbf{M}, b, \xi, \alpha, \beta)}{\partial \xi_l} = 0 \Rightarrow C - \alpha_l - \beta_l = 0 \Rightarrow 0 < \alpha_l < C, \forall l \quad (3.24)$$

(3)将(3.22)、(3.23)和(3.24)三个式子代入式(3.21)中，可以得到原问题的对偶问题，如式(3.25)所示。

$$\begin{aligned} \max_{\alpha} & -\frac{1}{2} \sum_{i,j} \alpha_i \alpha_j h_i h_j \left[ (x_{i,1} - x_{i,2})^T (x_{j,1} - x_{j,2}) \right]^2 + \sum_i \alpha_i \\ & = \max_{\alpha} -\frac{1}{2} \sum_{i,j} \alpha_i \alpha_j h_i h_j K_D(z_i, z_j) + \sum_i \alpha_i \\ & \text{s.t. } 0 \leq \alpha_l \leq C, \forall l \\ & \quad \sum_l \alpha_l h_l = 0 \end{aligned} \quad (3.25)$$

式(3.25)表明该模型的求解与 **SVM** 的求解非常相似，所以可以使用现有的工具包求解。可是实际情况下，对于很多种的测度学习方法（例如 **LMNN**<sup>[53]</sup> 和 **ITML**<sup>[54]</sup>）的求解，要求  $\mathbf{M}$  是一个半正定的矩阵<sup>[55]</sup>，这样很多现有的工具箱就不能使用。而对于问题(3.25)，作者采用两步贪婪策略求解，首先忽略半正定这个约束条件，直接采用 **SVM** 工具箱学习一个初步的测度矩阵  $\mathbf{M}$ ，然后再采用次梯度映射的方法将  $\mathbf{M}$  映射到一个半正定的空间内，所以很多现有的 **SVM** 包都可以求解对这个问题。

### 3.2.2.3 基于多模态距离测度学习算法的疲劳表情判别

根据前两小节训练得到测度矩阵  $\mathbf{M}$  之后，在该测度下结合 **KNN** 分类算法<sup>[56]</sup>对样本进行分类。根据样本描述特征的数量，距离测度学习分类算法可以分为单模态和多模态两类。

单模态的测度学习就是根据训练样本的一种描述特征，训练求得该特征的

测度矩阵  $\mathbf{M}$ ，然后在所求的测度下估计测试样本的类别。首先计算测试样本在该测度下与所有训练样本的距离，计算公式如(3.26)所示，如果 KNN 分类器的近邻元素变量  $n=5$ ，则根据与之最近的五个样本的类别投票判断测试样本的类别，以本文中判断驾驶员的疲劳表情为例，如果最近的五个样本有三个或三个以上的样本为疲劳，则判断该测试样本为疲劳状态，反之为正常状态。

$$D_{\mathbf{M}_1 \dots \mathbf{M}_K}(z_i, z) = (z_i - z)^T \mathbf{M}(z_i - z) \quad (3.26)$$

其中  $z_i$  ( $i=1, 2, N$  表示训练样本， $z$  表示测试样本。

单一的特征描述会丢失很多的细节信息，但如果结合多种描述特征，则可以尽可能的避免这种情况，所以本文采用多模态特征测度学习来判断驾驶员的疲劳表情。顾名思义，多模态特征测度学习就是对训练样本提取多种特征，再分别对每种特征求测度矩阵  $\mathbf{M}$ ，然后计算测试样本在每种测度下与所有训练样本的距离，再求其所有测度下的平均距离作为 KNN 分类器分类的依据，其距离的计算公式如式(3.27)。

$$D_{\mathbf{M}_1 \dots \mathbf{M}_K}(z_i - z) = \frac{1}{K} \sum_{k=1}^K (z_i - z)^T \mathbf{M}_k (z_i - z) \quad (3.27)$$

其中  $z_i$  ( $i=1, 2, N$  表示训练样本， $z$  表示测试样本， $k$  表示提取的特征的种类， $\mathbf{M}_k$  表示该种特征对应的测度矩阵。

在本章第一节已对面部图像提取了 LBP、Gabor、HOG 三种特征，根据多模态特征测度学习的方法，需要分别计算三种特征的测度矩阵  $\mathbf{M}_1$ 、 $\mathbf{M}_2$ 、 $\mathbf{M}_3$ ，再通过式(3.27)计算测试样本在三种测度下与每个训练样本的平均距离，并以此来判断该测试样本属于疲劳还是非疲劳状态，其流程图如图 3.6 所示。



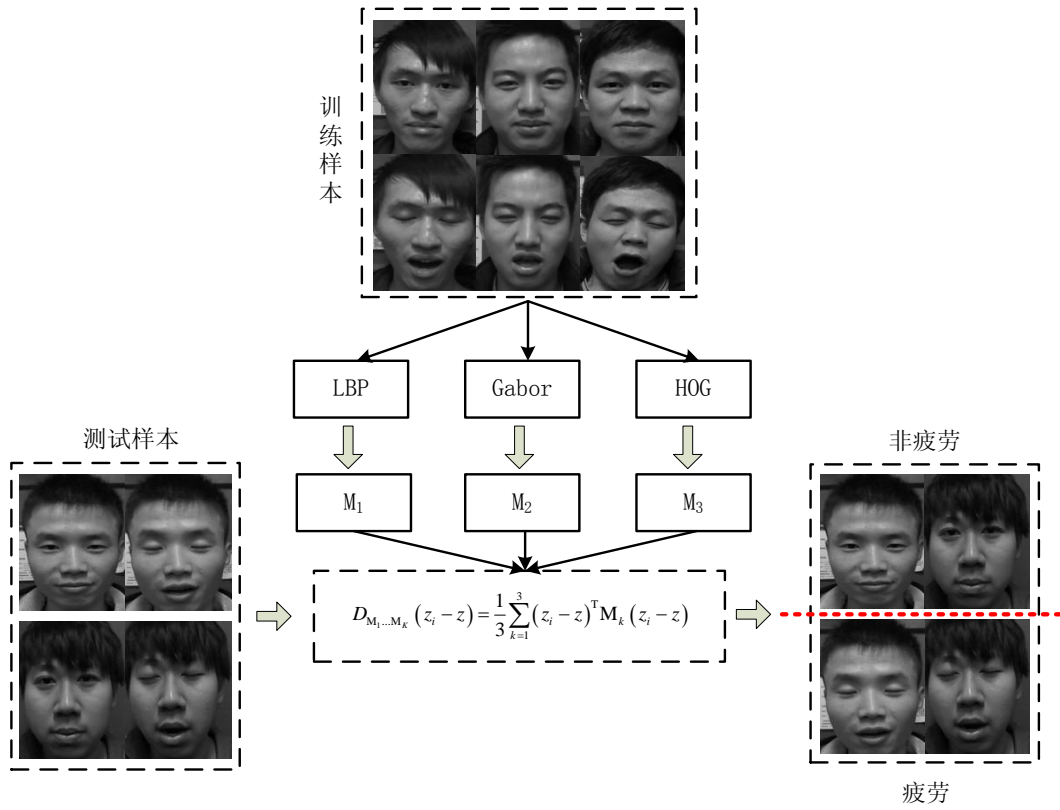


图 3.6 多模态测度学习的疲劳表情判别的原理图

### 3.2.3 疲劳表情判别的实验结果

驾驶员在疲劳的状态下，会有很多的疲劳表现，其中从图像信息中比较容易辨别的有打瞌睡，打哈欠等特点，所以实验中疲劳状态的样本主要包括了闭眼、张嘴以及伴随打哈欠过程出现的一些表情变化。本文对 20 个人在正常状态和模拟疲劳的状态下采集了 252 个正样本（正常）和 626 个负样本（疲劳），然后从中随机挑选了 3 个人作为测试样本（包括 42 个正样本，94 个负样本），剩余的 17 个人则作为训练样本。其中部分正样本如图(3.7)所示，部分负样本如图(3.8)所示。

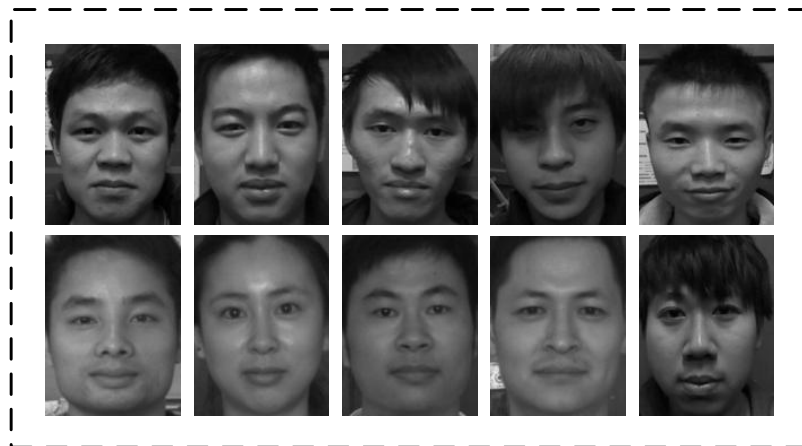


图 3.7 部分正样本

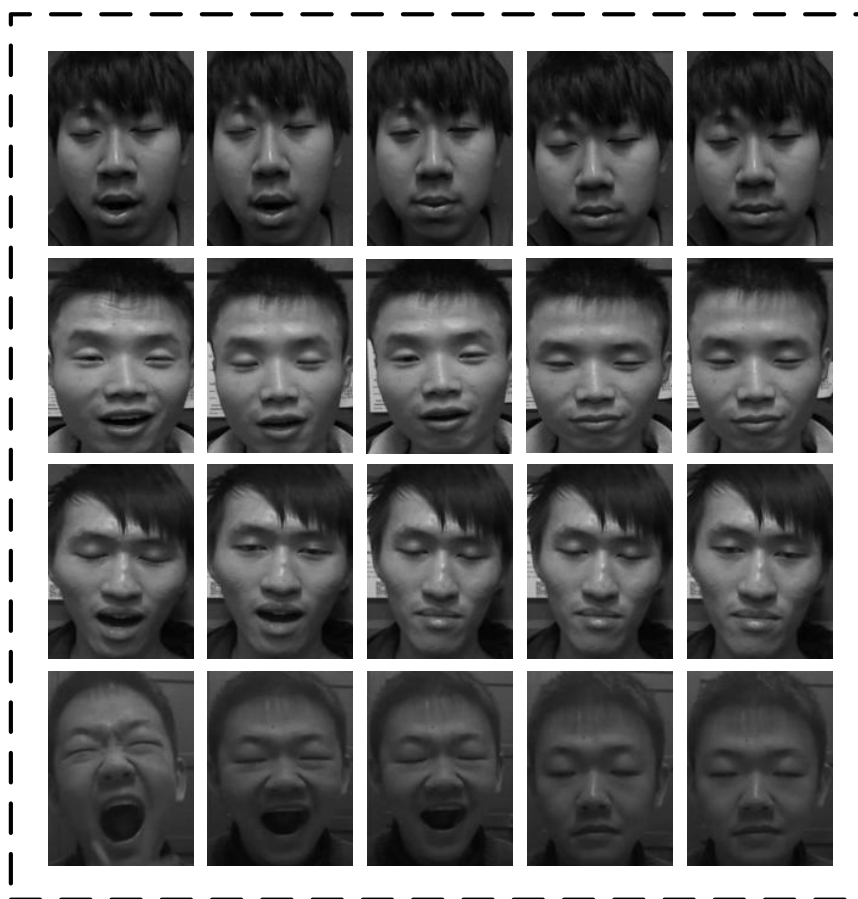


图 3.8 部分负样本

分别对训练样本和测试样本提取 LBP, Gabor, HOG 三种特征, 基于二元组核函数的测度学习框架, 根据训练样本求得每种特征的测度矩阵  $M$ , 再根据多模态特征测度学习算法对测试样本作验证性的实验。并且为了验证多模态特征测度学习的有效性, 本文也分别对 LBP, Gabor, HOG 三种特征做了单模态特征的验证实验, 实验结果如表 3.1 所示。

表 3.1 疲劳表情判断的实验结果

	准确率
LBP	72.3%
Gabor	75.2%
HOG	69.3%
多模态	81.4%

从表 3.1 中可以看出多模态特征相较于单模态特征有更高的准确率, 表明多种特征对于一张图片的描述比单一特征包含了更多图像信息, 这与理论也是一致的, 当描述特征的完整性越高时, 其分类效果也相应有所提高。

分析几组实验的结果表明: 导致面部表情判别错误的原因有很多, 首先, 整体面部表情种类非常复杂, 训练的样本并不能包含所有的状态; 其次, 由于有些疲劳表情与正常表情的差异非常小, 例如当某些疲劳样本与正常样本只有

眼睛状态发生细微变化时，如图 3.9 所示，此时尽管采用多种特征描述，但是依然非常难判断其状态；另外，由于样本之间存在个体差异，面部某些主要的特征会有较大的差别，比如有些测试者的眼睛比较小，其某些疲劳状态的表情与正常状态的表情就更加难以区分，如图 3.10 所示；最后，有些测试者由于头发比较长，会遮挡面部比较明显的特征，如图 3.11 所示，这种情况也会增加状态判断的难度。

综上所述，整体面部表情的判别时一个非常复杂的过程，包括测试者本身的差别与外界带来的干扰，以及疲劳状态判别的复杂性，所以整体面部表情判别的难度非常大。

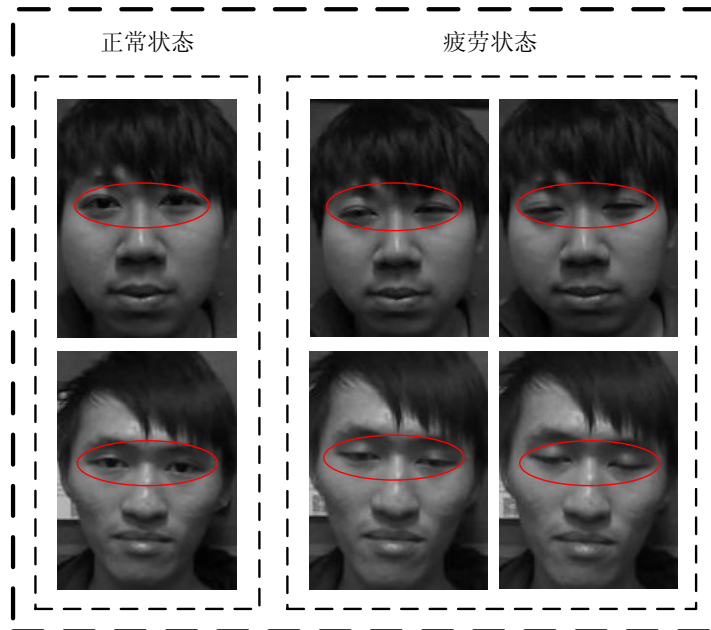


图 3.9 面部疲劳表情差异较小的样本

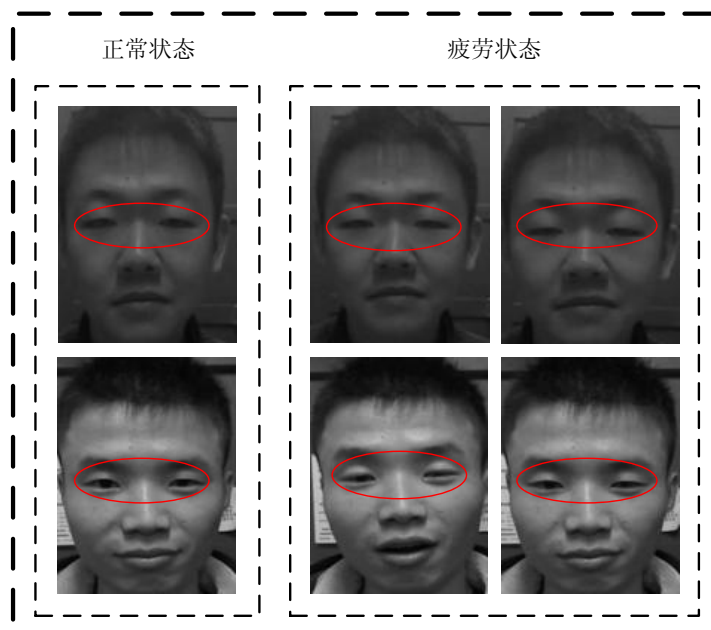


图 3.10 面部主要特征差异较小的样本

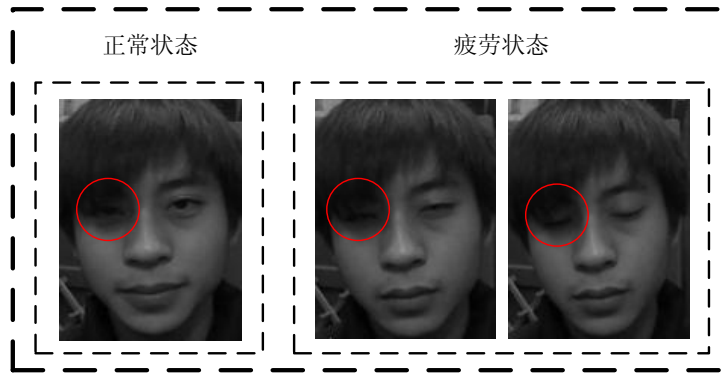


图 3.11 面部表情被遮挡的样本

### 3.3 本章小结

本章主要介绍了根据多模态特征测度学习判断驾驶员的疲劳表情。首先介绍了三种用于面部表情特征描述的算子（LBP、Gabor、HOG），然后介绍了一些关于测度学习的基本概念和三种常用的测度学习算法，并且引入了一种基于二元组核函数的测度学习框架，并以此学习了 LBP、Gabor、HOG 三种特征的测度矩阵，采用多模态测度学习算法综合判别驾驶员的疲劳表情。本章最后的实验部分也证明了算法的有效性，并且说明多模态测度学习在准确率上明显优于单模态。

## 第4章 基于模糊推理的疲劳状态判别方法

本文前两章分别判断了面部局部特征状态和整体面部特征状态，主要包括眼睛和嘴部的状态，以及疲劳表情的判别，这些都将是判断驾驶员疲劳状态的重要依据。而本章就是通过设计一个模糊系统<sup>[57]</sup>，融合上述的疲劳信息推理判断驾驶员的疲劳状态。本章的算法流程图如图 4.1 所示。

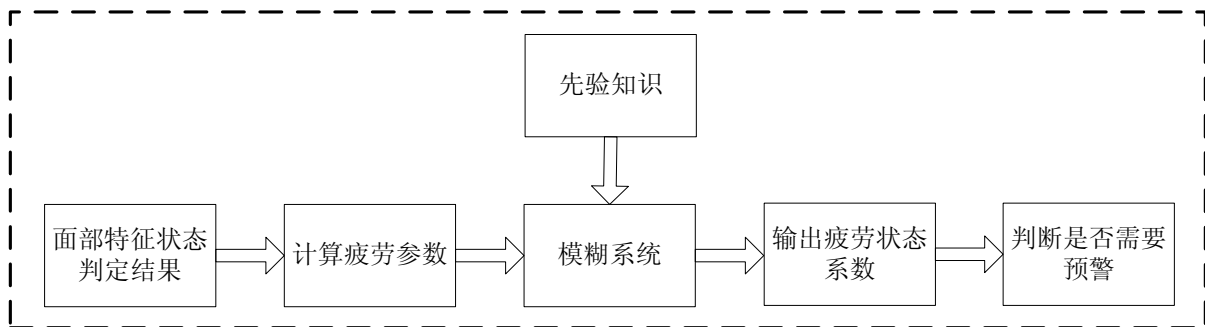


图 4.1 基于模糊推理的疲劳状态判别流程图

### 4.1 模糊系统的结构

#### 4.1.1 模糊化接口

模糊化就是在系统的输入输出论域上定义语言变量，将精确的输入输出值转换为模糊的语言值。通常在语言变量的论域上将其划分为有限的几个等级，例如本文输出变量疲劳参数有四个等级：None, Small, Medium, Large。并且需要根据经验定义每个变量的隶属函数，以确定变量属于某一个等级的隶属度。显然，变量的等级越多也就是意味着相应的推理规则越多，推理过程越复杂，但是如果等级过少推理系统又很难达到预期的效果，所以在定义变量状态时需要同时考虑系统的复杂性和准确性。

#### 4.1.2 规则库

规则库由若干条推理规则组成，这些规则是根据专家的经验总结得出，实际上它就是用语言的表现形式来描述输入输出变量之间的关系。目前，规则库有很多种表达形式，如很多 IF...is...and...is...THEN...is...语句的组合，或者一个矩阵表等，但是最常用的表示形式还是 IF...THEN 语句形式，其中 IF 部分表示输入变量的状态，THEN 部分则表示输出变量的模糊值。所有的控制规则组合起来就是模糊系统的规则库，为模糊推理过程提供推理的依据。如上段中描述，规则的数量与输入变量的个数有直接的关系，并且规则数量越多也并不代

表模糊系统的准确性越高，因为推理的准确性与规则的准确性有直接的联系。

### 4.1.3 模糊推理

模糊推理是整个系统中非常关键的一个环节，就是根据规则库中蕴涵的输入与输出之间的联系推理得出输出变量的模糊状态。目前推理方法有很多种，但是复杂的方法意味着非常大的计算量，这样就在实际应用的过程中很难满足实时性的要求，所以通常使用的有两种相对简单的算法 Zadeh 和 Mamdani。大致模糊推理的过程包括如下几个步骤：

- 1.根据输入变量的值确定触发的规则。由于规则库中包含很多条规则，但是每组输入参数并不能触发所有的规则，而只是选择性的触发其中几条相关的规则。
- 2.计算触发每条触发的规则的可信度。由第一步已经确定了会被触发的规则，接下来要分别计算每条规则会被触发的可信度，从而来推理最有可能发生的结果。
- 3.确定模糊系统总的输出。综合所有被触发的规则的可信度，计算得出最终的模糊输出的结果。

### 4.1.4 去模糊化

经过模糊推理之后，系统输出的只是输出论域上的模糊子集，但是我们实际上需要的是一个准确的控制量，所以需要解模糊的过程，常用的去模糊化方法有两种。

#### 1.最大隶属度方法

最大隶属度去模糊化就是直接从输出的模糊子集中选取隶属度最大的元素作为系统的输出值，即  $w' = \max_{w \in W} \mu(w)$ ，但是当输出的模糊集合中最大隶属度的元素不止一个时，则取所有这些元素的平均值并采用四舍五入取整之后作为精确的输出值，如式(4.1)所示。

$$w' = \frac{1}{N} \sum_{j=1}^N w_j, w_j = \max(\mu(w)) \quad (4.1)$$

其中  $N$  表示最大隶属度的元素个数

最大隶属度输出计算非常简单，但是它只考虑了最大隶属度的输出值，因此会丢失很多的信息，所以这种方法只能应用在要求不高的情况下。

#### 2.加权平均法

加权平均法就是将模糊输出集合中的各个元素及其对应的乘子加权求平均值，并采用四舍五入的方法取整得到精确的输出值，这种方法广泛的应用在工业控制中，其计算表达式如式(4.2)所示。

$$w' = \frac{\sum_{i=1}^n w_i k_i}{\sum_{i=1}^n k_i} \quad (4.2)$$

其中  $k_i$  表示提前设置好的加权系数，不同系统具有不同的加权系数。如果当  $k_i$  表示各个元素的隶属度时，该方法也称为重心法。

## 4.2 疲劳状态判别的模糊系统设计

模糊系统在本文中主要是用来融合各个疲劳参数，推理出驾驶员最终的疲劳参数<sup>[58]</sup>。根据本文的第二章和第三章，作者通过统计 30 帧内眼睛、嘴部和表情的状态计算出几个能够代表驾驶疲劳状态的疲劳参数，再经过自己设计的模糊系统推理得出驾驶员的疲劳状态<sup>[59]</sup>。本章的下一节，作者也有做仿真实验来验证所设计的系统的准确性。具体的设计步骤如下：

### 4.2.1 定义输入输出变量及其隶属函数

统计连续 30 帧内眼睛、嘴部和面部表情的状态，确定模糊系统的四个输入变量：眨眼频率（*BlinkFreq*）、闭眼持续最长时间（*ClosureDura*）、哈欠频率（*YawnFreq*）、疲劳表情（*FatigueExp*）；一个输出变量：疲劳参数（*DC*）。每个变量的计算公式如下所示，其中  $n$  表示统计的帧数，本文取  $n=30$ 。

(1)眨眼的频率:  $BlinkFreq = \frac{n_i}{n}$ ， $n_i$  表示统计的 30 帧内闭眼的帧数；

(2)闭眼持续最长时间:  $ClosureDura = 0.1n_i$ ， $n_i$  表示统计的 30 帧内连续闭眼的最大帧数；

(3)哈欠的频率:  $YawnFreq = \frac{n_i}{n}$ ， $n_i$  表示统计的 30 帧内嘴部张开的帧数；

(4)疲劳表情:  $FatigueExp = \frac{n_i}{n}$ ， $n_i$  表示统计的 30 帧内被判断为疲劳表情的帧数。

参考相关文献<sup>[59]</sup>，根据先验知识定义输入变量的隶属函数分别如式(4.3)~(4.6)所示。

$$\mu_{BlinkFreq} = \begin{cases} \mu_S(x) = -5x + 1, 0 \leq x \leq 0.2 \\ \mu_M(x) = \begin{cases} 5x, 0 \leq x \leq 0.2 \\ -5x + 2, 0.2 < x \leq 0.4 \end{cases} \\ \mu_L(x) = \begin{cases} 5x - 1, 0.2 \leq x \leq 0.4 \\ 1, x > 0.4 \end{cases} \end{cases} \quad (4.3)$$

$$\mu_{CloseDura} = \begin{cases} \mu_S(x) = -4x + 1, 0 \leq x \leq 0.25 \\ \mu_M(x) = \begin{cases} 4x, 0 \leq x \leq 0.25 \\ -4x + 2, 0.25 < x \leq 0.5 \end{cases} \\ \mu_L(x) = \begin{cases} 4x - 1, 0.25 \leq x \leq 0.5 \\ 1, x > 0.5 \end{cases} \end{cases} \quad (4.4)$$

$$\mu_{YawnFreq} = \begin{cases} \mu_S(x) = -5x + 1, 0 \leq x \leq 0.2 \\ \mu_M(x) = \begin{cases} 5x, 0 \leq x \leq 0.2 \\ -5x + 2, 0.2 < x \leq 0.4 \end{cases} \\ \mu_L(x) = \begin{cases} 5x - 1, 0.2 \leq x \leq 0.4 \\ 1, x > 0.4 \end{cases} \end{cases} \quad (4.5)$$

$$\mu_{FatigueExp} = \begin{cases} \mu_S(x) = -5x + 1, 0 \leq x \leq 0.2 \\ \mu_M(x) = \begin{cases} 5x, 0 \leq x \leq 0.2 \\ -5x + 2, 0.2 < x \leq 0.4 \end{cases} \\ \mu_L(x) = \begin{cases} 5x - 1, 0.2 \leq x \leq 0.4 \\ 1, x > 0.4 \end{cases} \end{cases} \quad (4.6)$$

输出变量的隶属函数如式(4.7)所示。

$$\mu_{DC} = \begin{cases} \mu_N(x) = \begin{cases} 1, 0 \leq x \leq 0.1 \\ -5x + 1.5, 0.1 < x \leq 0.3 \end{cases} \\ \mu_S(x) = \begin{cases} 5x - 0.5, 0.1 \leq x \leq 0.3 \\ -10x + 4, 0.3 < x \leq 0.4 \end{cases} \\ \mu_M(x) = \begin{cases} 10x - 3, 0.3 \leq x \leq 0.4 \\ -5x + 3, 0.4 < x \leq 0.6 \end{cases} \\ \mu_L(x) = \begin{cases} 5x - 2, 0.2 \leq x \leq 0.6 \\ 1, x > 0.6 \end{cases} \end{cases} \quad (4.7)$$

输入变量的隶属度函数在 Matlab 中的仿真图如图 4.2 所示。输出变量的隶属度函数在 Matlab 中的仿真图如图 4.3 所示。

#### 4.2.2 规则库的定义

人在疲劳的状态下，身体机能都会有所下降。此时，眨眼的频率会明显的高于正常状态，并且每次眨眼的时间会变长，同时可能会出现打哈欠的情况，面部表情也会发生变化。因此，根据本文定义四个输入变量可以模拟人从清



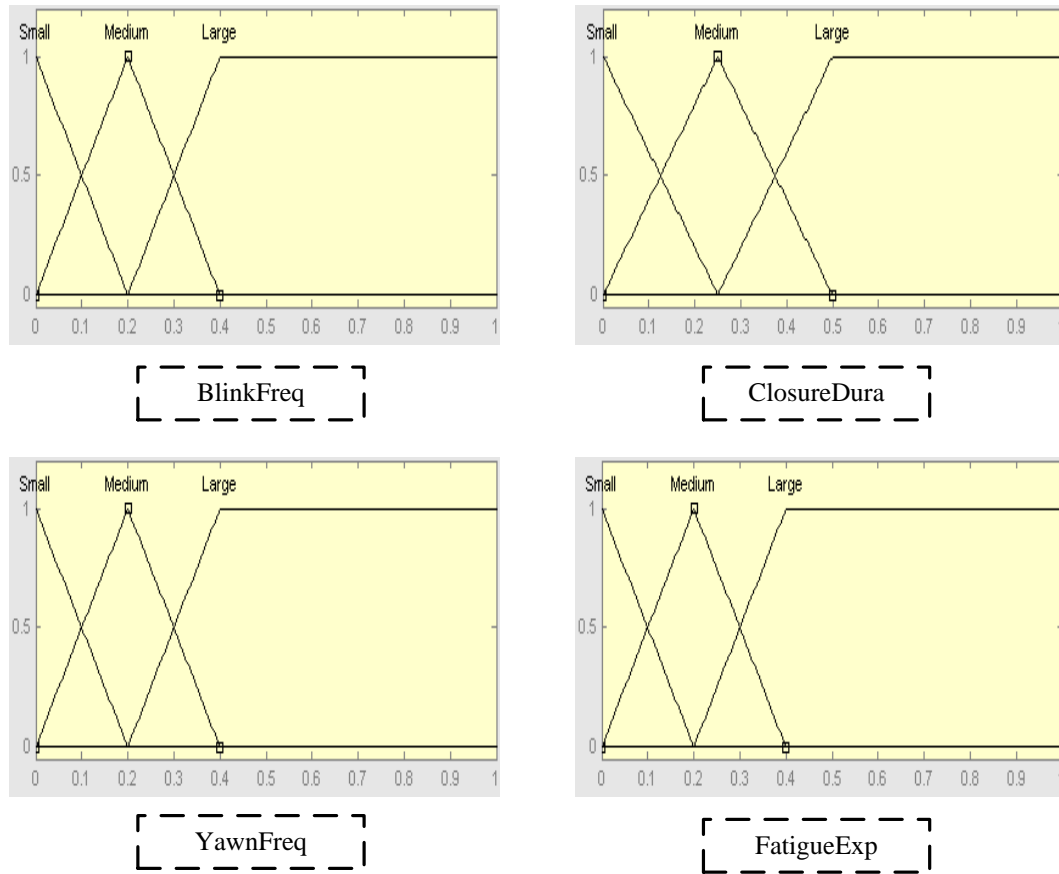


图 4.2 输入变量的隶属度函数

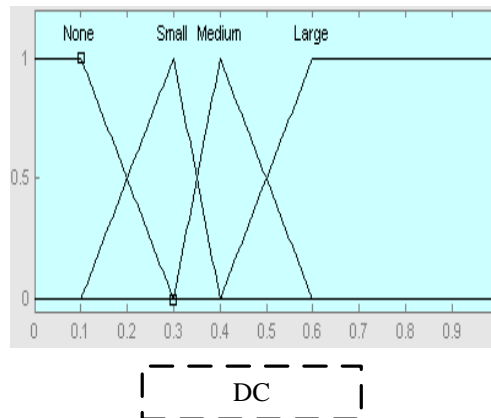


图 4.3 输出变量的隶属度函数

醒到疲劳的变化过程，随着各个参数的增大，可以表明驾驶员的疲劳状态是在逐步加重的。因此作者基于实际情况设置了一组推理规则，能够根据各个输入参数的变化，模拟人疲劳状态变化的过程。

设计的模糊系统包括四个输入变量，并且每个变量有三个等级，所以理论上推理的规则库应该有  $3^4$  条规则。但是根据实际情况，有一些规则是不存在的，所以我们剔除了其中的一些不合常理的规则，例如：

(1) If (BlinkFreq is Small) and (ClosureDur is Small) and (YawnFreq is Large) and (FatigueExp is Large) then (DC is Medium)。

(2) If (BlinkFreq is Small) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Large) then (DC is Medium)。

(3) If (BlinkFreq is Small) and (ClosureDur is Small) and (YawnFreq is Large) and (FatigueExp is Small) then (DC is Small)。

(4) If (BlinkFreq is Small) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Small) then (DC is Small)。

(5) If (BlinkFreq is Small) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Medium) then (DC is Medium)。

除去其中的 17 条规则，最后规则库只保留了 64 条，部分规则如下所示。

(1) If (BlinkFreq is Small) and (ClosureDur is Small) and (YawnFreq is Small) and (FatigueExp is Small) then (DC is None)

(2) If (BlinkFreq is Small) and (ClosureDur is Medium) and (YawnFreq is Small) and (FatigueExp is Small) then (DC is None)

(3) If (BlinkFreq is Small) and (ClosureDur is Small) and (YawnFreq is Small) and (FatigueExp is Medium) then (DC is None)

(4) If (BlinkFreq is Small) and (ClosureDur is Medium) and (YawnFreq is Medium) and (FatigueExp is Small) then (DC is Small)

(5) If (BlinkFreq is Small) and (ClosureDur is Small) and (YawnFreq is Medium) and (FatigueExp is Medium) then (DC is Small)

(6) If (BlinkFreq is Medium) and (ClosureDur is Medium) and (YawnFreq is Small) and (FatigueExp is Small) then (DC is Small)

(7) If (BlinkFreq is Medium) and (ClosureDur is Medium) and (YawnFreq is Medium) and (FatigueExp is Small) then (DC is Small)

(8) If (BlinkFreq is Medium) and (ClosureDur is Large) and (YawnFreq is Medium) and (FatigueExp is Medium) then (DC is Medium)

(9) If (BlinkFreq is Large) and (ClosureDur is Medium) and (YawnFreq is Medium) and (FatigueExp is Medium) then (DC is Medium)

(10) If (BlinkFreq is Large) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Medium) then (DC is Medium)

(11) If (BlinkFreq is Large) and (ClosureDur is Large) and (YawnFreq is Medium) and (FatigueExp is Medium) then (DC is Medium)

(12) If (BlinkFreq is Medium) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Large) then (DC is Medium)

(13) If (BlinkFreq is Large) and (ClosureDur is Large) and (YawnFreq is Large) and (FatigueExp is Medium) then (DC is Large)

- (14) If (BlinkFreq is Large) and (ClosureDur is Large) and (YawnFreq is Medium) and (FatigueExp is Large) then (DC is Large)
- (15) If (BlinkFreq is Medium) and (ClosureDur is Large) and (YawnFreq is Large) and (FatigueExp is Large) then (DC is Large)
- (16) If (BlinkFreq is Large) and (ClosureDur is Large) and (YawnFreq is Large) and (FatigueExp is Large) then (DC is Large)

#### 4.2.3 模糊推理与去模糊化

在此我们假设一组输入参数：BlinkFreq=0.25；ClosureDura=0.3；YawnFreq=0.3；FatigueExp =0.15。

##### 1.确定触发的规则

$$\mu_{BlinkFreq} : \mu_M(x)=0.75, \mu_L(x)=0.25;$$

$$\mu_{ClosureDura} : \mu_M(x)=0.8, \mu_L(x)=0.2;$$

$$\mu_{YawnFreq} : \mu_M(x)=0.5, \mu_L(x)=0.5;$$

$$\mu_{FatigueExp} : \mu_S(x)=0.25, \mu_M(x)=0.75;$$

根据上述八种隶属度，会触发规则库中的 16 条规则，如下所示。

- (1)If BlinkFreq is Medium) and (ClosureDur is Medium) and (YawnFreq is Medium) and (FatigueExp is Small) then (DC is Small)
- (2)If BlinkFreq is Medium) and (ClosureDur is Large) and (YawnFreq is Medium) and (FatigueExp is Small) then (DC is Medium)
- (3)If BlinkFreq is Medium) and (ClosureDur is Large) and (YawnFreq is Large) and (FatigueExp is Small) then (DC is Medium)
- (4)If BlinkFreq is Medium) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Small) then (DC is Medium)
- (5)If BlinkFreq is Large) and (ClosureDur is Medium) and (YawnFreq is Medium) and (FatigueExp is Small) then (DC is Medium)
- (6)If BlinkFreq is Large) and (ClosureDur is Large) and (YawnFreq is Medium) and (FatigueExp is Small) then (DC is Medium)
- (7)If BlinkFreq is Large) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Small) then (DC is Medium)
- (8)If BlinkFreq is Large) and (ClosureDur is Large) and (YawnFreq is Large) and (FatigueExp is Small) then (DC is Medium)
- (9)If BlinkFreq is Medium) and (ClosureDur is Medium) and (YawnFreq is Medium) and (FatigueExp is Medium) then (DC is Medium)
- (10)If BlinkFreq is Medium) and (ClosureDur is Large) and (YawnFreq is Medium)

- and (FatigueExp is Medium) then (DC is Medium)
- (11)If BlinkFreq is Medium) and (ClosureDur is Large) and (YawnFreq is Large) and (FatigueExp is Medium) then (DC is Large)
- (12)If BlinkFreq is Medium) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Medium) then (DC is Medium)
- (13)If BlinkFreq is Large) and (ClosureDur is Medium) and (YawnFreq is Medium) and (FatigueExp is Medium) then (DC is Medium)
- (14)If BlinkFreq is Large) and (ClosureDur is Large) and (YawnFreq is Medium) and (FatigueExp is Medium) then (DC is Large)
- (15)If BlinkFreq is Large) and (ClosureDur is Medium) and (YawnFreq is Large) and (FatigueExp is Medium) then (DC is Large)
- (16)If BlinkFreq is Large) and (ClosureDur is Large) and (YawnFreq is Large) and (FatigueExp is Medium) then (DC is Large)

2.确定 16 条触发的规则之后，需要计算每条规则的可信度，并且根据计算确定模糊系统总的输出，最后还需要对输出去模糊化。这两部分的计算部分很简单，主要是取小和‘与’操作，但是由于本文选取了四个输入变量，其计算步骤非常的繁琐，不便于描述，所以具体的推理过程在文中不再详述。而本文的仿真实验是直接采用了 matlab 中的模糊推理工具箱，模糊系统总界面如图 4.4 所示，本组实验推理去模糊化之后的结果如图 4.5 所示。

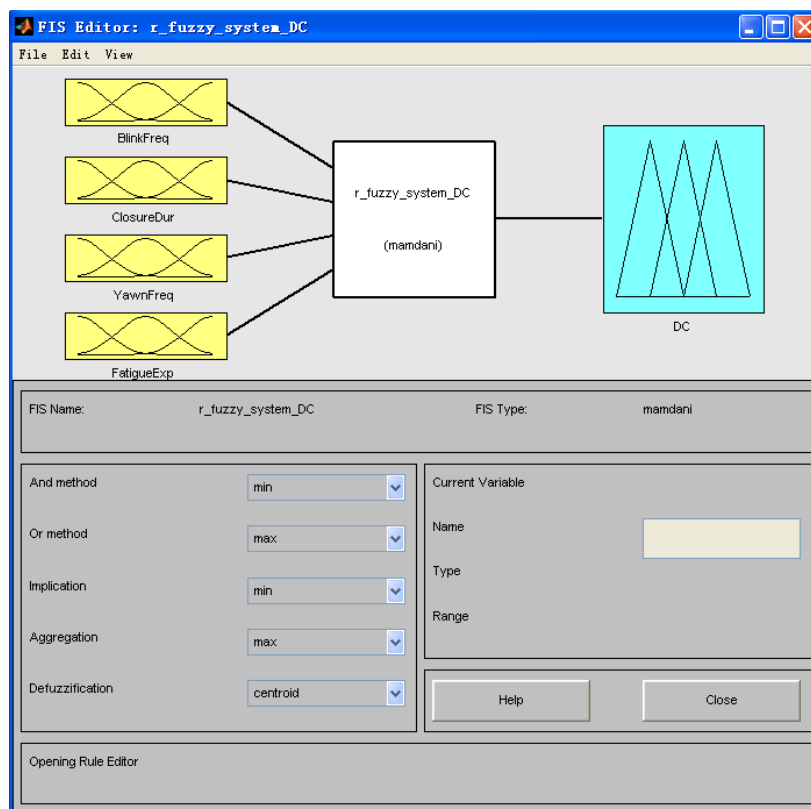


图 4.4 matlab 模糊推理工具箱

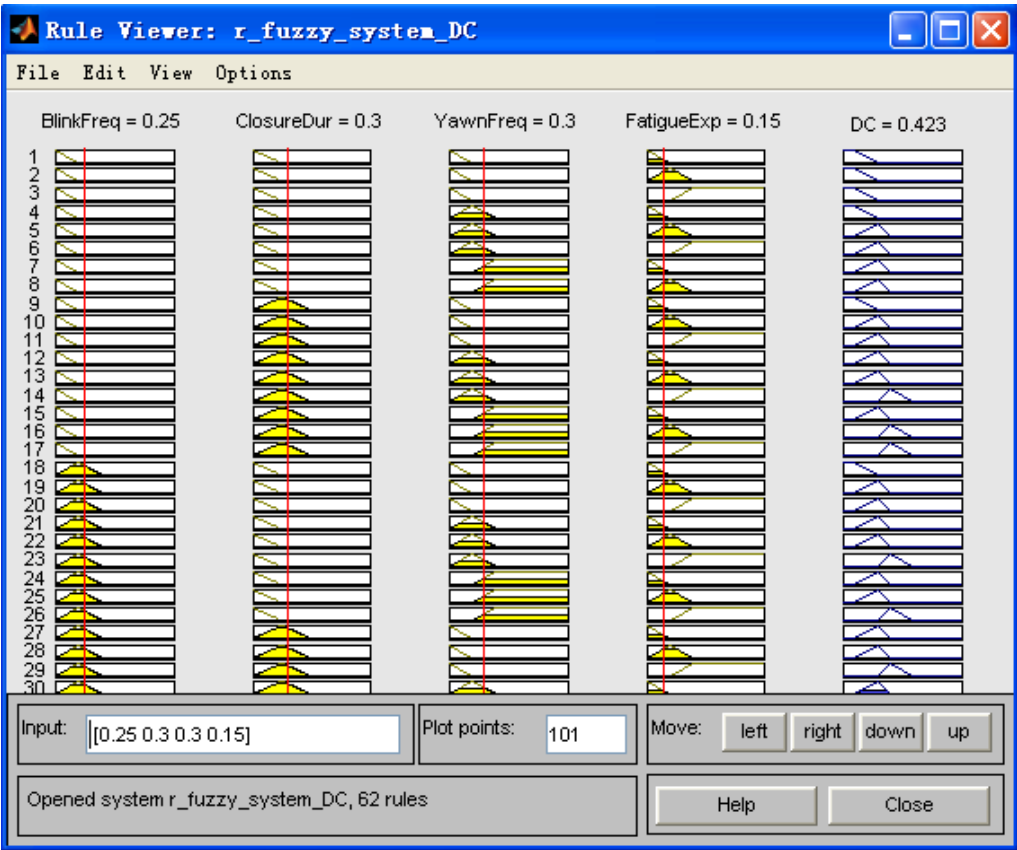


图 4.5 matlab 模糊系统工具箱的推理结果

4.3 疲劳状态判别的仿真实验结果

为了验证所设计的模糊系统的合理性，作者根据实际情况设置了几组数据作为系统的输入，通过各个疲劳参数的组合变化模拟一个人从清醒到疲劳的过程，然后再观察系统的输出是否符合实际情况。系统的验证实验的 matlab 仿真图如图 4.6~4.11 所示，统计实验结果如表 4.1 所示。

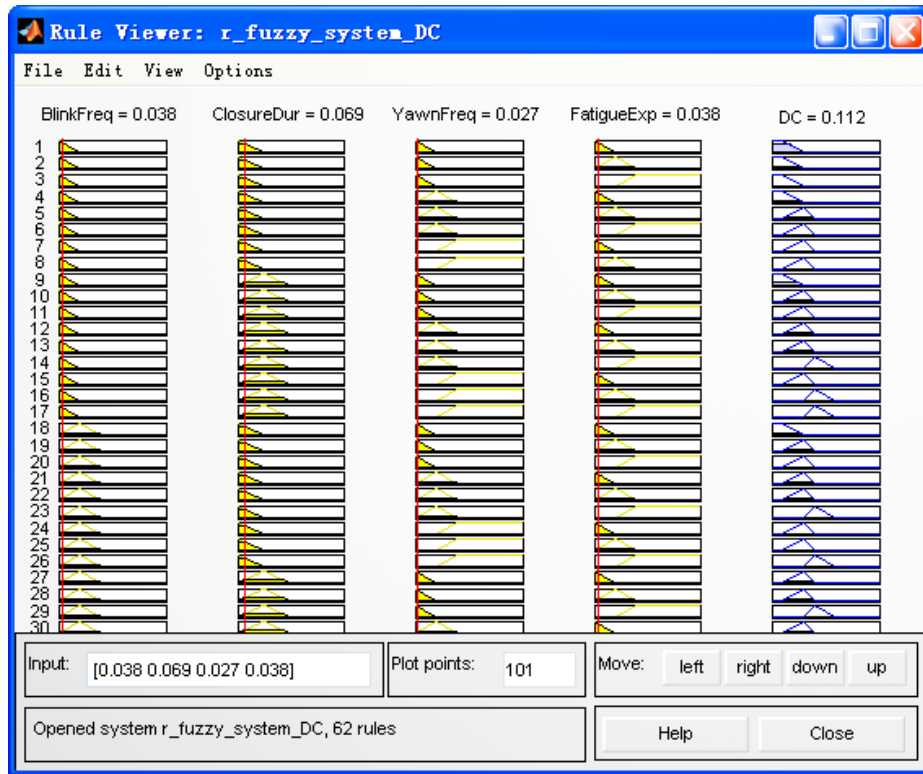


图 4.6 第一组实验的仿真图

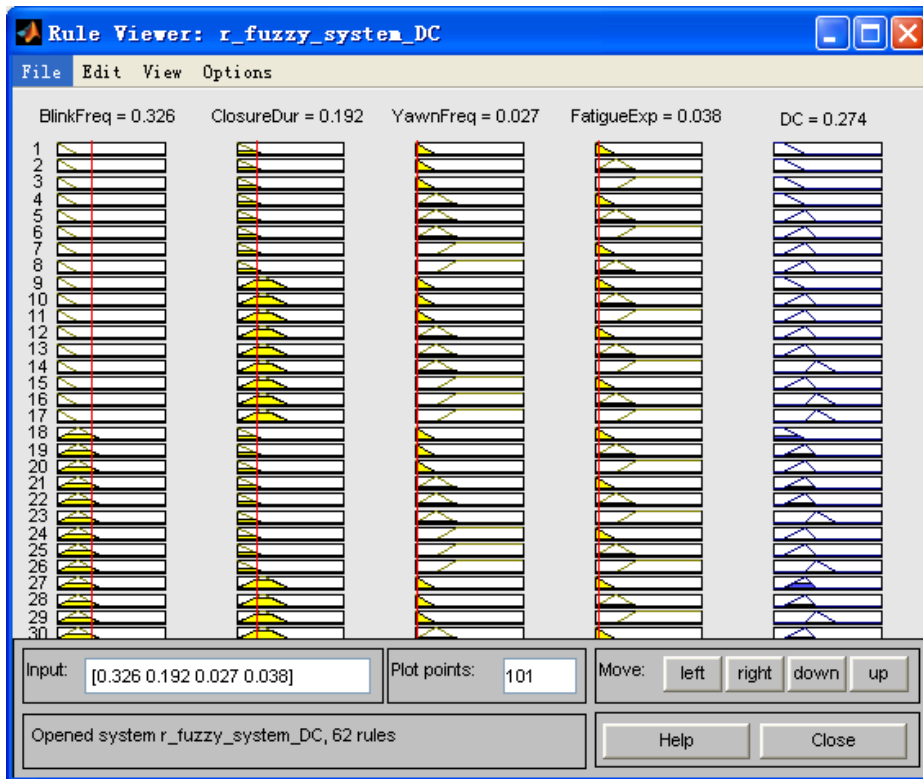


图 4.7 第二组实验的仿真图

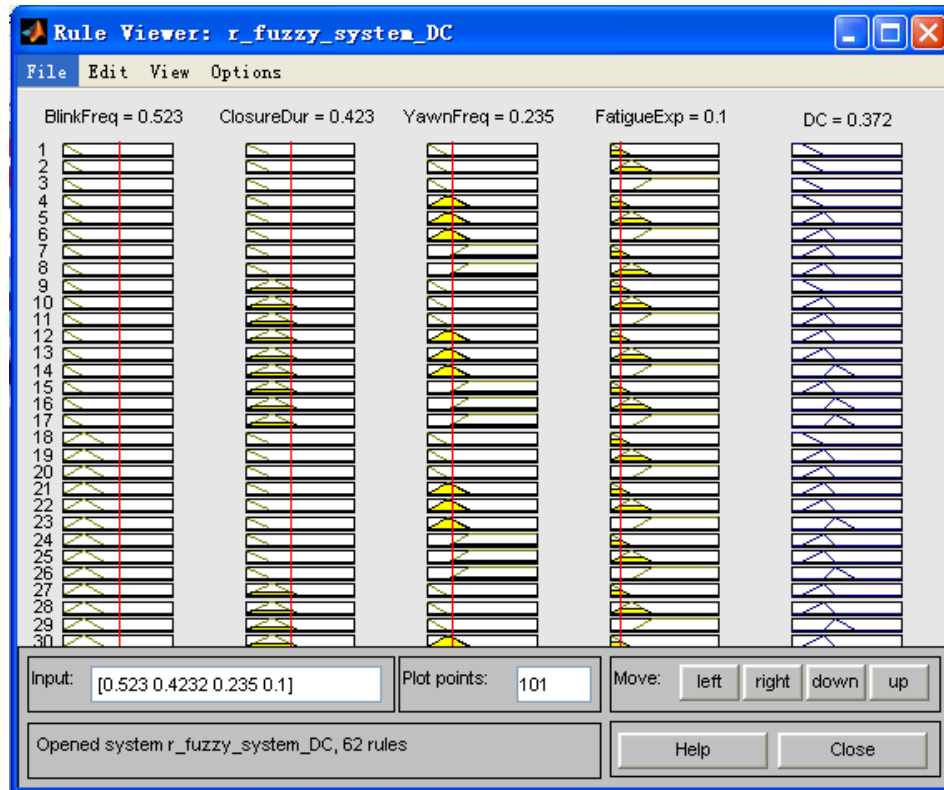


图 4.8 第三组实验的仿真图

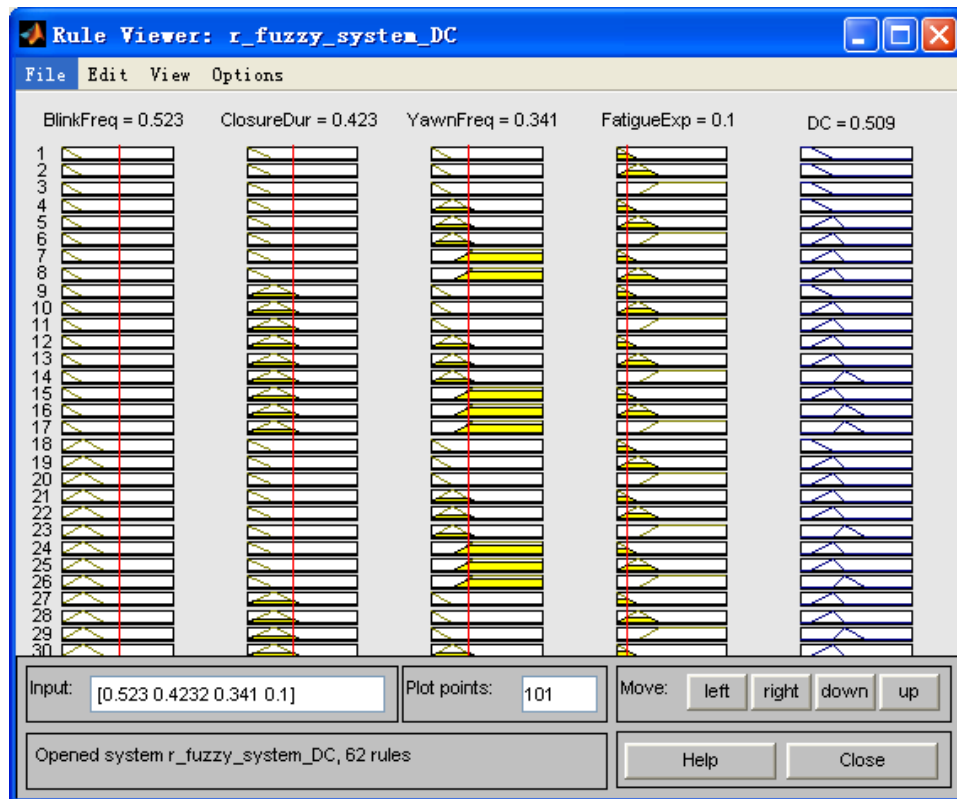


图 4.9 第四组实验的仿真图

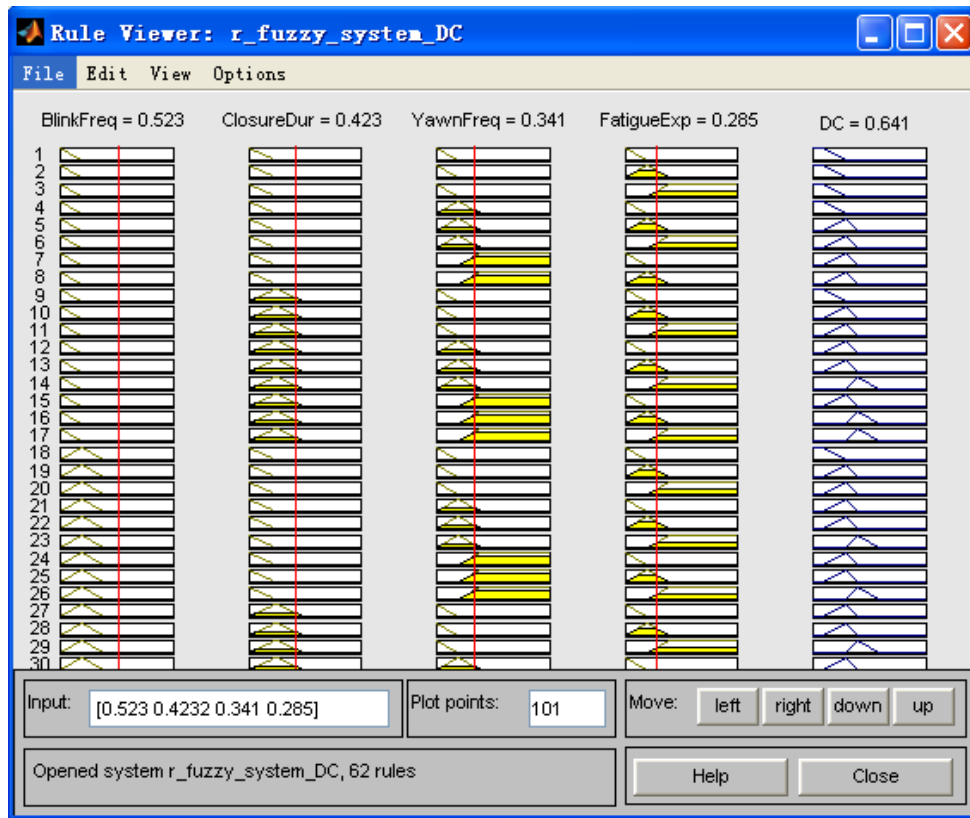


图 4.10 第五组实验的仿真图

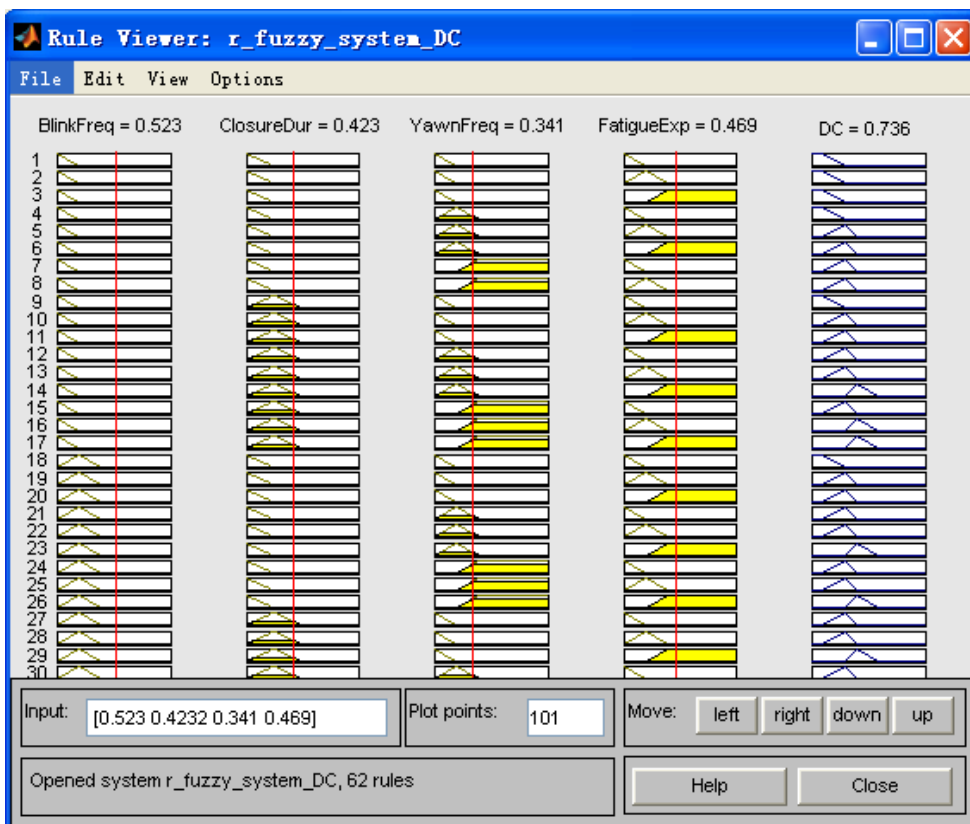


图 4.11 第六组实验的仿真图



表 4.1 疲劳信息融合推理的仿真实验结果

组别	BlinkFreq	ClosureDura	YawnFreq	FatigueExp	DC
1	0.038	0.069	0.027	0.038	0.112
2	0.326	0.192	0.027	0.038	0.274
3	0.523	0.423	0.235	0.1	0.372
4	0.523	0.423	0.341	0.1	0.509
5	0.523	0.423	0.341	0.285	0.641
6	0.523	0.423	0.341	0.469	0.736

表 4.1 中，第 1 组数据表明当四个输入变量的值很小时，输出变量 DC 的值也非常小，表明驾驶员的状态很清醒。一般情况下，人们在出现疲劳状况初期眨眼的频率会增加，所以第 2 组实验数据中眼睛部分的两个疲劳参数有所增大，此时疲劳系数 DC 的输出值也相应增大。第 3 组实验数据表示当人眼睛的疲劳状态越来越严重并出现了打哈欠的现象，此时系统输出变量 DC 则表明驾驶员已经进入了轻度疲劳的状况。第 4 组数据相较第 3 组表明驾驶员打哈欠的状况变严重，而此时的疲劳系数则表明驾驶员进入了中度疲劳的状态。第 5、6 组实验数据表示驾驶员的面部表情出现了疲劳迹象，倦容越来越明显，系统对应的输出系数也说明驾驶员此时已经进入了很严重的疲劳状态。这个仿真实验是模拟一个驾驶员状态从清醒到疲劳的全过程，实验结果也与实际的情况基本相符，因此说明本文中设计的模糊系统有良好的推理性能，能满足实际需求。

#### 4.4 本章小结

本章首先介绍了模糊系统的构成，详述了各个部分的组成，然后在此基础上根据本文的需要构造了一个模糊系统，包括输入输出变量的选取和隶属度函数的定义、规则库的定义、推理过程以及去模糊化的过程，本章最后也利用 matlab 的模糊推理工具箱对设计的系统做了验证性实验，结果表明其推理的过程基本符合实际情况，因此证明该系统是合理有效的。

## 第5章 驾驶员疲劳状态判别的实验结果

本文的第 2、3、4 章组成了一个完整的驾驶员疲劳检测系统，包括系统的输入，面部特征检测及状态判别，信息融合，驾驶员疲劳状态输出。本章将根据五组模拟驾驶的视频图片验证整个系统的准确性。

### 5.1 测试样本的采集

本实验所用的样本均是模拟驾驶环境在实验室中采集的，相机安装在人脸的正上方以保证能采集到完整的人脸图像并不阻碍驾驶员前方的视线，所用的传感器是 PointGrey FL3-GE-13S2M 黑白相机，图片的大小为 640\*480，相机的参数以及照片如表 5.1 和图 5.1 所示。实验的测试样本包括分别从五个同学的模拟驾驶视频中随机挑选的连续 30 帧图像，部分测试样本的面部图像如图 5.2 所示。

表 5.1 Point Grey 相机参数设置

相机型号	FL3-GE-13S2C-C (Color)
镜头型号	Computar(2/3'',D:67.1°, H:56.3°, V:43.7°) 焦距：8mm
图像分辨率	640*480(最高可达 1288*964)
采集帧率	25 帧/秒(最高可达 31 帧/秒)
像素深度	24bit
数据接口	Gigabit Ethernet
传输速率	1000M bit/s
曝光时间	0.03ms~32s(自动曝光)
电源供应	12-24V DC, <2.5W



(1) PointGrey 相机

(2) 镜头

图 5.1 采集样本的传感器

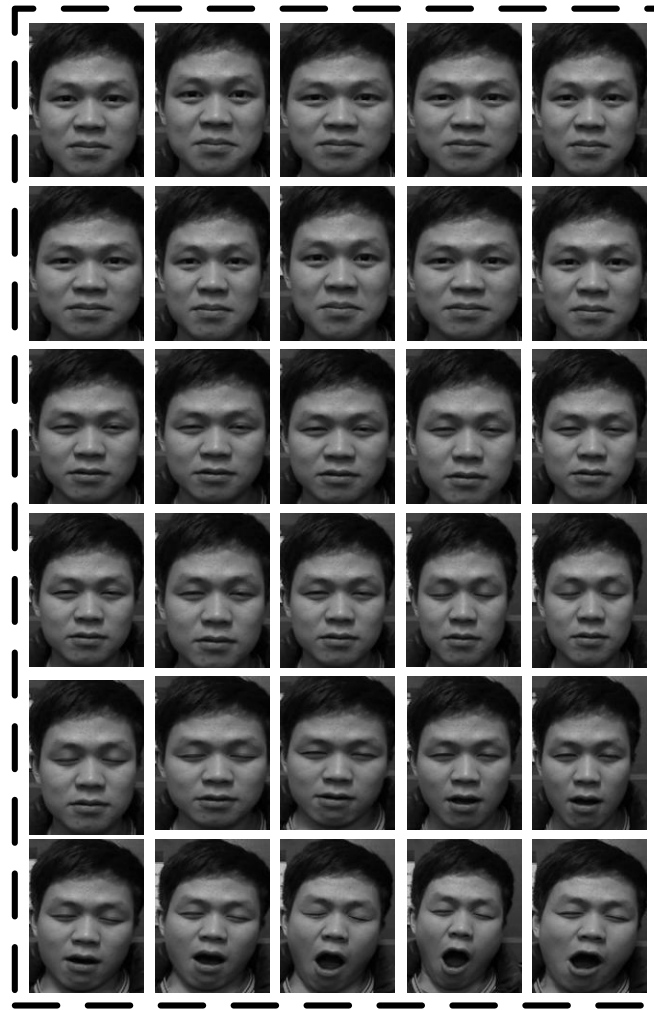


图 5.2 部分测试样本的面部图像

## 5.2 测试的实验结果

本实验是为了验证由第 2 章、第 3 章以及第 4 章组成的疲劳检测系统的准确性，所以实验主要包括两个部分：首先需要统计 30 帧内闭眼、张嘴和疲劳表情的帧数，然后再根据各部分状态判别的结果计算四个疲劳参数，并且利用模糊系统推理得出驾驶员的疲劳状态系数，并以此来判断驾驶员是否需要预警。根据模糊系统输出的疲劳状态系数，对驾驶员状态的判定规则如表 5.2 所示。

表 5.2 驾驶员状态判定标准及预警措施

疲劳状态系数 DC	驾驶员 状态	预警措施
( 0 , 0.2]	清醒	无
(0.2 , 0.4]	轻微	预警器 1s 响一次
(0.4 , 0.6]	中度	预警器 1s 响两次
>0.6	严重	预警器连续响铃

第一部分：对面部局部特征及整体面部的状态判别结果如表 5.3 所示，其中 ‘( )’ 表示标定的真实值。

表 5.3 驾驶员面部特征状态的统计结果

实验 组别	眼睛闭合 (帧)	眼睛连续闭合 (帧)	嘴部张开 (帧)	疲劳表情 (帧)
1	6 (5)	3 (3)	4 (3)	0 (3)
2	18 (17)	12 (15)	7 (7)	14 (12)
3	4 (4)	2 (2)	5 (5)	4 (6)
4	8 (9)	5 (4)	6 (7)	6 (7)
5	2 (2)	1 (1)	3 (2)	3 (3)

第二部分：根据第一部分的实验结果计算疲劳参数，并根据第四章设计的模糊系统融合推理出驾驶员的疲劳状态，实验结果如表 5.3 所示，其中 ‘( )’ 表示由标定的真实值计算得出的疲劳参数，其中第一组和第四组的仿真实验图如图 5.3~5.6 所示。

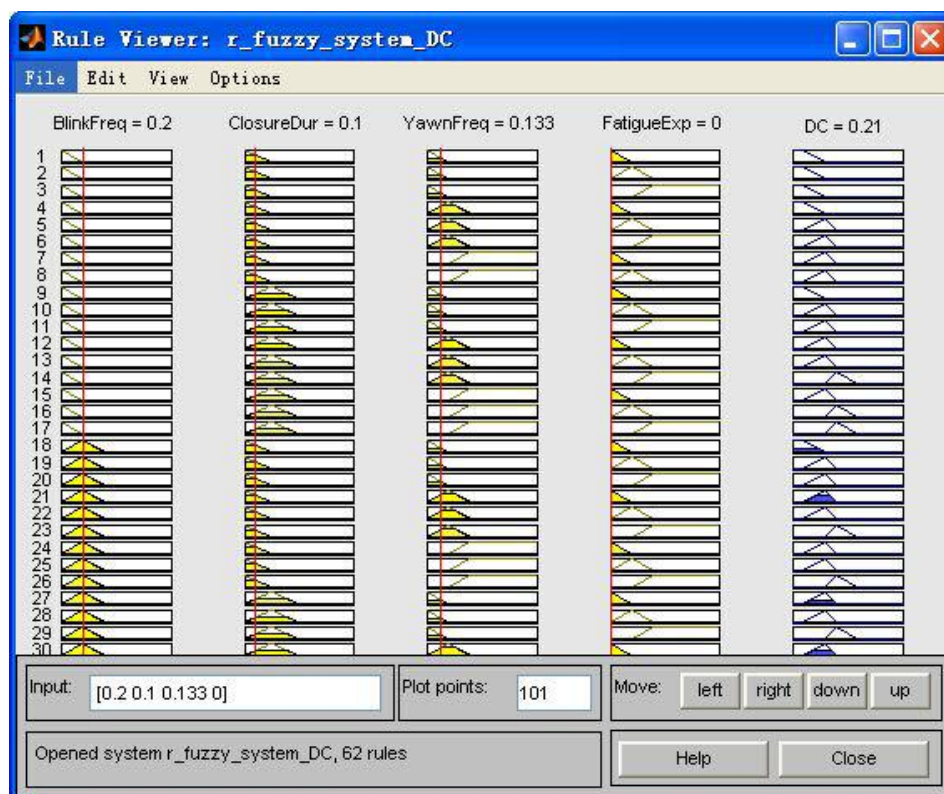


图 5.3 第一组实验结果的推理结果

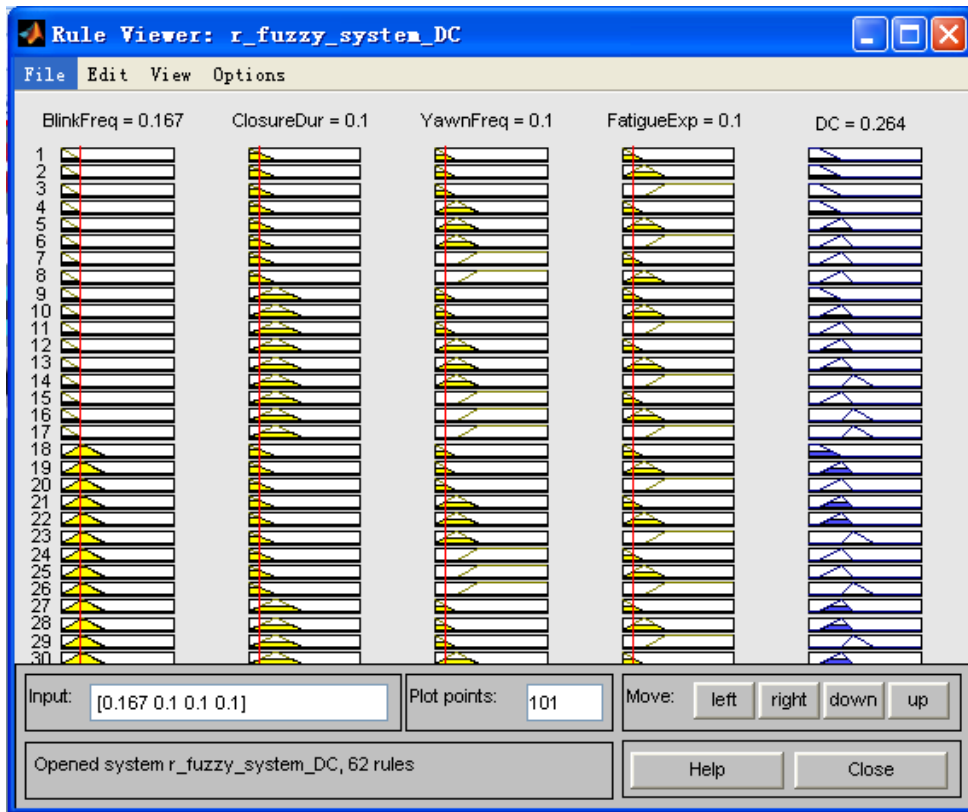


图 5.4 第一组实验标定值的推理结果

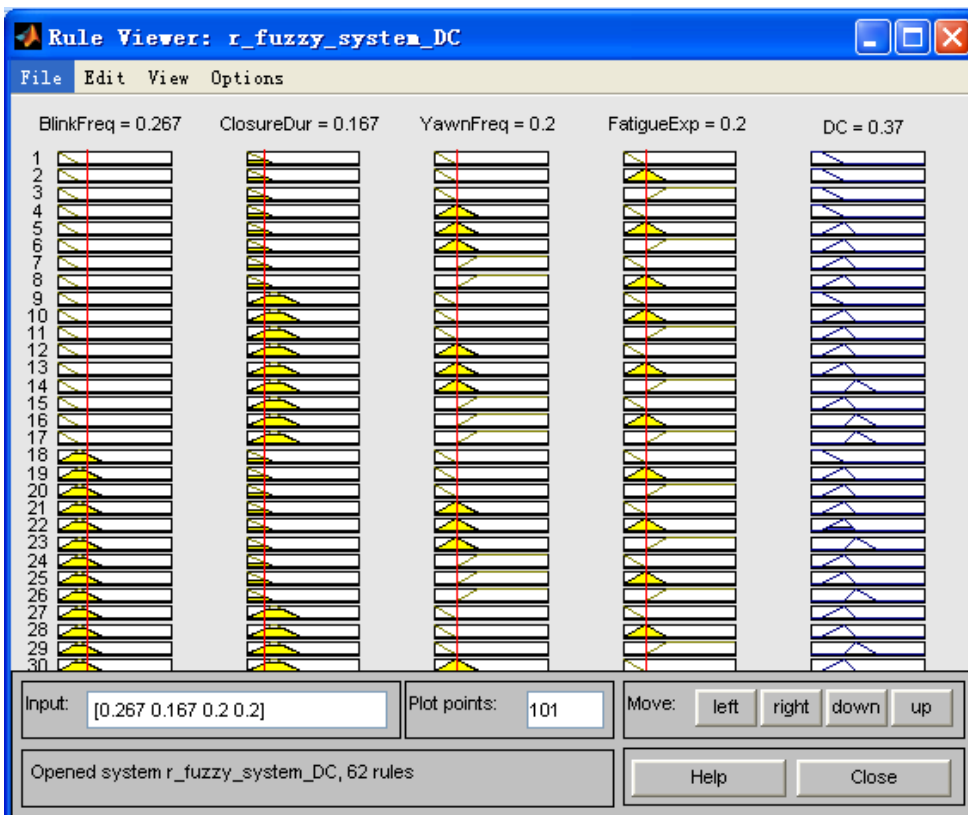


图 5.5 第四组实验结果的推理结果

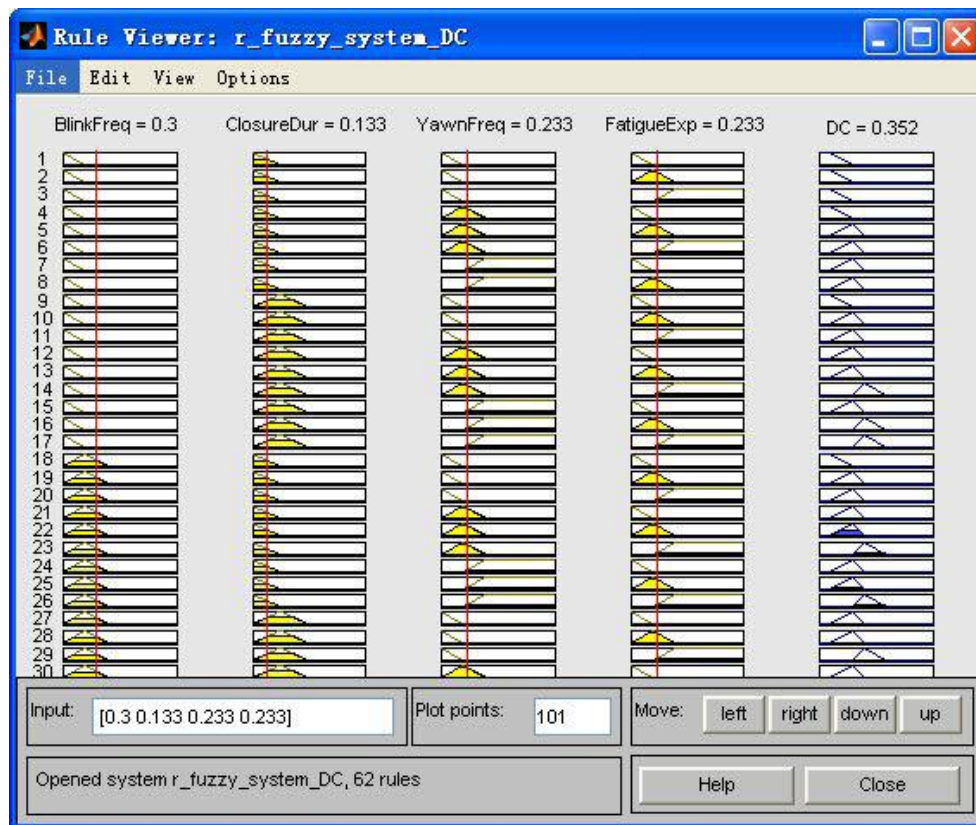


图 5.6 第四组实验标定值的推理结果

表 5.4 驾驶员疲劳状态判别的实验结果

实验组别	BlinkFreq	ClosureDura	YawnFreq	FatigueExp	DC	是否预警
1	0.200	0.100	0.133	0.000	0.21	是
	(0.167)	(0.100)	(0.100)	(0.100)	(0.264)	(是)
2	0.600	0.400	0.233	0.467	0.749	是
	(0.567)	(0.500)	(0.233)	(0.400)	(0.742)	(是)
3	0.133	0.067	0.167	0.133	0.218	是
	(0.133)	(0.067)	(0.167)	(0.200)	(0.223)	(是)
4	0.267	0.167	0.200	0.200	0.37	是
	(0.300)	(0.133)	(0.233)	(0.233)	(0.352)	(是)
5	0.067	0.033	0.100	0.100	0.172	否
	(0.067)	(0.033)	(0.067)	(0.100)	(0.162)	(否)

从表 5.3 和表 5.4 中可以看出，系统无论是状态判别还是推理结果都有很高的准确率，与标定的真实值的误差在可控范围内，能够满足驾驶员疲劳预警的需求。分析表明判别误差主要是因为被检测者头部出现了较大幅度的运动导致眼睛被遮挡，或者由于光线问题使嘴部边缘模糊，导致分割出现问题。

## 5.3 本章小结

本章中，从模拟驾驶的视频中采集了五组测试样本来验证整个系统的准确性，首先判断样本眼睛、嘴部以及表情的状态，再根据判断结果计算疲劳参数，并利用上章设计的模糊系统推理得出驾驶员的疲劳状态系数，并且根据该系数判断驾驶员是否需要提醒，实验结果也表明该系统有较高的准确率。



## 总结与展望

### 1. 工作总结

疲劳驾驶问题日趋严重，所以驾驶员疲劳检测技术的研究已经成为了安全辅助驾驶领域的主要研究方向。本文提出了一种基于面部视觉多特征融合的驾驶员疲劳检测算法，通过统计分析驾驶员眼睛、嘴部和整体面部表情的疲劳状态，根据模糊系统融合推理出驾驶员的疲劳状态，其无论从系统的整体框架还是算法实现的细节上均在现有算法的基础上有所改进，实验部分也验证了算法的有效性。论文研究的主要工作如下：

（1）面部局部特征的检测与疲劳状态判别。由于系统主要是根据驾驶员的面部信息判断其疲劳状态，所以系统第一步是采用基于 Haar 特征的级联 Adaboost 检测算法确定驾驶员的面部位置。在检测到人脸之后，采用 Gabor 滤波器增强面部主要特征的边缘信息，并结合面部的几何特征定位眼睛和嘴部。采用一种旋转不变的 LBP 金字塔特征对眼睛进行特征描述，训练线性 SVM 分类器判别眼睛的开闭状态，由于旋转不变的 LBP 特征具有旋转不变性，因此能够克服驾驶员在驾驶过程中发生的倾斜等情况带来的干扰；根据嘴部的张开面积及宽高比判断嘴部的开闭状态。眼睛和嘴部的状态都是系统判断驾驶员疲劳状态的重要依据。

（2）整体面部特征的疲劳状态判别。采用多模态特征的测度学习算法判断驾驶员的疲劳表情，首先对整体面部图像提取了 LBP, Gabor, HOG 三种特征，再分别计算每种特征下的测度矩阵  $M$ ，并计算测试样本在三种测度下的平均距离，再结合 KNN 分类器对驾驶员的疲劳表情分类。实验表明多模态特征的判别算法在鲁棒性和准确性方面都明显的优于单模态特征。疲劳表情的判别结果也是驾驶员疲劳状态判定的重要依据。

（3）基于模糊推理的信息融合。首先根据先验知识设计一个能够满足驾驶员疲劳状态推理的模糊系统。统计一段时间内驾驶员眼睛、嘴部以及表情的状态判别结果，并计算出四个能够代表驾驶员疲劳状态的疲劳参数（眨眼频率（BlinkFreq）、闭眼持续最长时间（ClosureDura）、哈欠频率（YawnFreq）、疲劳表情（FatigueExp）），再根据之前设计的模糊系统融合推理得出驾驶员的疲劳状态系数，并以此来判断驾驶员的疲劳状态。



## 2. 工作展望

由于作者水平和时间有限，本文对于驾驶员疲劳检测技术只作了初步的研究，其算法的普适性和系统的实时性、完整性还需要提高。如果要将其开发成实际应用的系统还需要进行大量的研究工作。

本文的实验样本均是在实验室中模拟驾驶环境采集的，而实验室内的光照变化相较于真实的驾驶环境较为单一，并且疲劳表现非常复杂，模拟疲劳状态采集到的图像不能覆盖所有的实际情况，比如说疲劳表情的样本远不止实验中涉及到的几种非常明显的疲劳特征。因此如果将文中算法运用到实际应用中，本文的很多算法的普适性和鲁棒性都需要提高。另外，实验采用均是正脸样本，而实际过程中驾驶员的脸部发生遮挡、旋转、低头等情况是不可避免的，所以算法上还需要进一步改进。

实际中人疲劳的特征非常多，远比本文中所采用的四个参数要复杂很多，因此如果为了保证系统的稳定性，需要融合更多的疲劳特征，但是要减少信息融合的计算量，以使系统能够达到实时性的要求。

模糊系统融合多个疲劳参数，其定义规则库需要很多的先验知识，包括疲劳参数与疲劳状态的度量，以及各个疲劳参数之间的联系等，但是由于研究条件有限，本文定义规则时仅是参考了一些相关的论文，所以可能在推理的细节上还需要完善。

## 参考文献

- [1] 吕健健.基于贝叶斯网络的驾驶员疲劳评估方法研究[大连理工大学硕士学位论文].辽宁:大连理工大学,2013,1-15
- [2] Nobe S A,Wang F Y.An overview of recent developments in automated lateral and longitudinal vehicle controls.In:2001 IEEE Int Conf on Systems,Man and Cybernetics,Tucson,AZ,United States,2001,5,3447-3452
- [3] 程文冬,付锐,袁伟等.驾驶人疲劳监测预警技术研究与应用综述.中国安全科学学报,2013,23(1):155-160
- [4] 刘刚.基于眼睛状态判断的疲劳检测:[大连理工大学硕士学位论文].辽宁:大连理工大学通信与信息系统系,2013,1-2
- [5] Lal s,Craig A,Boord P.Development of an algorithm for an EEG-based driver fatigue countermeasure.Journal of Safety Research,2003,34(3),321-328
- [6] Jeong I C,Lee D H,Park S W.Automobile driver's stress index provision system that utilizes electrocardiogram.In:2007 IEEE Intelligent Vehicles Symposium,Istanbul,Turkey,2007,652-656
- [7] Katsis C D,Ntouvas N E,Bafas C G.Assessment of muscle fatigue during driving using surface EMG.In:Proc of the IASTED Int Conf on Biomedical Engineering,Innsbruck,Austria,2004,259-262
- [8] Adachi T,Butsuen T,Kamiyama S.Application of real-time simulation to the lane-keeping system.In:Proc of VNIS'94-1994 Vehicle Navigation and Information Systems,Yokohama,Japan,1994,415-418
- [9] Baronti F,Lenzi F,Roncella R.Distributed sensor for steering wheel grip force measurement in driver fatigue detection.In:2009 Design,Automation and Test in Europe Conf and Exhibition,Nice,France,2009,894-897
- [10] 孟勇.人眼状态及疲劳检测技术研究:[西安电子科技大学硕士学位论文].西安:西安电子科技大学通信与信息系统系,2012,3-14
- [11] 吴从中,贺伟,汤润森等.光照自适应的疲劳驾驶检测系统.电子测量与仪器学报,2012,26(1):60-66
- [12] Lee D,Oh S,Heo S.Drowsy driving based on the driver's head movement using infrared sensors.In:2nd Int Symposium on Universal Communication,Osaka, Japan,2008,231-236
- [13] 邬敏杰,穆平安,张彩艳.基于眼睛和嘴巴状态的驾驶员疲劳检测算法.计算

- 机应用与软件,2013,30(3):25-54
- [14] 郭永彩, 苏渝维, 高潮.基于 FPGA 的红外图像实时采集系统设计与实现. 仪器仪表学报,2011,32(3):514-519
- [15] Knipling R R,Wang J s,Kanianthra J N.Current NHTSA drowsy driver R&D.In:Proceeding of the Int Technical Conf on the Enhanced Safety of Vehicles.Melbourne,Australia,1996
- [16] 刘志强.驾驶员疲劳检测方法.农业机械学报,2006,37(4):210-217
- [17] 录庆.日本发明防止驾驶员打瞌睡系统.南方农机,2004,(5):46
- [18] STEERING ATTENTION MONITOR[EB/OL]. <http://www.carkits.com.au/index.htm>
- [19] 赵李坤.基于人脸特征识别的驾驶员疲劳检测方法研究:[中南大学硕士学位论文].湖南:中南大学,2013,2-18
- [20] 赵李坤.基于人脸特征识别的驾驶员疲劳检测方法研究:[中南大学硕士学位论文].湖南:中南大学,2013,5-23
- [21] Ji Q, Zhu Z W. Real-time nonintrusive monitoring and prediction of driver fatigue. IEEE Transactions on Vehicular Technology,2004,53(4),1052-1068
- [22] Boveris S.Driver fatigue monitoring technologies and future ideas.In:Proc AWAKE Road Safety Workshop,2004
- [23] Zhang Y,Zhao X Y, et al. A time delay neural etwork model for simulating eye gaze data. Journal of Experimental and Theoretical Artificial Intelligence 20-11,23(1): 111-126
- [24] Sun W,Tang I Q. Driver fatigue driving detection base- d on eye state. International Journal of Digital Content Technology and its Applications,2011, 10(5): 307-314
- [25] Ralph O M, Seong G K. Visual Analysis of Eye State and Head Pose for Driver Alertness Monitoring, IEEE Transaction on Intelligent Transportation System,2013,14(3): 1462-1469
- [26] Mandalapu S,Preeti B. Driver fatigue detection using mouth and yawning analysis. International Journal of Computer Science and Security,2008, 8(6):183-188
- [27] Luis M B,Miguel A S. Real time system for monitoring driver vigilance. IEEE Transactions on intelligent Transportation Systems,2006,7(1):63-77
- [28] 周鹏.疲劳事故隐患消除技术与方法.汽车电器,1998,8:27-30
- [29] 邸巍.基于视觉的全天候驾驶员疲劳与精神分散状态监测方法研究[吉林大学博士学位论文].吉林:吉林大学,2010,2-10

- [30] 关玉萍.基于视频的驾驶员疲劳检测技术的研究[哈尔滨理工大学硕士学位论文].黑龙江:哈尔滨理工大学,2010,3-12
- [31] 李峰,曾超,徐向东.驾驶防瞌睡装置中人眼快速定位方法研究.光学仪器, 2002,24(4):70-72
- [32] Song F Y, Tan X Y, Liu X, et al. Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients. Pattern Recognition,2014
- [33] Huang C,Ai H Z,Li Y,et al.High-Performance Rotation Invariant Multi-view Face Detection. IEEE Transaction on Pattern Analysis and Machine Learning,2007,29(4):671-686
- [34] Paul V,Michael J.Rapid object detection using a boosted cascade of simple features.IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2001,1:511-518
- [35] Yoav F,Robert E,Schapire.A Decision-Theoretic Generalization of On-Line Learning and an Applicaition to Boosting.Journal of Computer and System Sciences,1997,55:119-139
- [36] Shen L L,Bai L. A review on gabor wavelets for face recognition. Pattern analysis and applications,2006,9:273-292
- [37] 李云峰.基于 Gabor 小波变换的人脸识别:[大连理工大学博士学位论文].辽宁:大连理工大学,2005,44-55
- [38] 文华.基于数学形态学的图像处理算法的研究[哈尔滨工程大学硕士学位论文].黑龙江:哈尔滨工程大学,2007,8-17
- [39] 丘亚伟.基于弱监督空间金字塔模型的图像分类研究:[山东大学硕士学位论文].山东:山东大学计算机科学与技术学院,2013,17-24
- [40] Timo O,Matti P,Topi M. Multiresolution gray-scale and rotation invariant texture classification with local bin- ary patterns.IEEE Transactions on Pattern Analysis and Machine Intelligence,2006,24(7):971-987
- [41] 王玲.基于 LBP 的特征提取研究:[北京交通大学硕士学位论文].北京: 北京交通大学计算机与信息技术学院,2009,23-45
- [42] Zhao G Y, Timo A. Rotation-invariant image and vid- eo description with local binary pattern features. IEEE Transactions on Image Processing, 2012, 21(4):1465 -1477
- [43] 高丽,杨树元,李海强.一种基于标记的分水岭图像分割新算法.中国图像图形学报,2007,12(6):1025-1033
- [44] 张杰.基于距离测度学习的图像分类方法研究:[复旦大学硕士学位论文].上

- 海: 复旦大学计算机科学技术学院,2010,15-31
- [45] 王天江.基于 HOG 特征的图像人体检测技术的研究与实现[华中科技大学硕士学位论文].湖北:华中科技大学,2008,15-29
- [46] 李凤英.基于静态图像的人体检测[天津大学硕士学位论文].天津:天津大学,2013,18-33
- [47] Jolliffe T. Principal Component Analysis. Springer-Verlag, New York, 1986.
- [48] 曹娟,张勇东,李锦涛.一种基于密度的自适应最优 LDA 模型选择方法.计算机学报,2008,10:234-240
- [49] Kilian Q.W, Lawrence K. S.Distance Metric Learning for Large Margin Nearest Neighbor Classification. Journal of Machine Learning Research 10, 2009:207-244
- [50] Wang F Q, Zuo W M, Zhang L, et al. A Kernel Classification Framework for Metric Learning. IEEE Transactions on Neural Networks and Learning Systems, 2014
- [51] Chang C C, Lin C J. Libsvm: A library for support vector machines. ACM Trans Intelligent System Technology, 2011, 2:1-27
- [52] Fan R E, Chen P H, Lin C J. Working set selection using the second order information for training SVM. Journal of Machine Learning Research, 2005, 6, 1889-1918
- [53] Domeniconi C, Gunopulos D, Peng J. Large margin nearest neighbor classifiers. IEEE Transaction on Neural Networks, 2005, 16(4):899-909
- [54] Davis J, Kulis B, Jain P. Information on-theoretic metric learning. In: Proc of 24th Int Conf on Machine Learning, 2007, 209-216
- [55] Lanckriet R G, Cristianini N, Bartlett P, et al. Learning the kernel matrix with semidefinite programming. Journal of Machine Learning Research, 2004, 5:27-72
- [56] Kumar P, P. H. S. Torr, and Zisserman A. An invariant large margin nearest neighbour classifier. In: Proc of the Eleventh IEEE Int Conf on Computer Vision. Brazil, 2007, 1-8
- [57] 张乃尧.典型模糊控制器的结构分析.模糊系统与数学.1997,11(2):10-21
- [58] 吴康华.基于 PERCLOS 的驾驶员疲劳检测系统设计:[浙江大学硕士学位论文].浙江:浙江大学,2008,2-35
- [59] 王秀.驾驶员疲劳检测系统的研究:[华南理工大学硕士学位论文].广州:华南理工大学,2010,43-63

## 致谢

论文完成之际，也表示我三年的硕士研究生学习即将就要结束了。这三年生活和学习都过得很充实，感觉收获了很多，其中也肯定少不了老师和同学们的帮助，没有他们的帮助就不会有我今天的成绩，在此对他们表示衷心的感谢。

首先要衷心的感谢我的导师朱青老师。这三年的学习期间，朱老师给了我很多的帮助和指导。学习上，朱老师给予我极大的空间，围绕课题的研究方向只要是我感兴趣的她都非常支持。遇到困难时，朱老师不仅尽力帮助我们解决，也总是鼓励我们要积极乐观的面对，也正是这种乐观的心态，让我在这三年的学习中感受到了很多的乐趣。另外，还要非常感谢王耀南老师，刚进入研究生阶段，我陷入了一阵迷茫，不知道如何选择学习的方向，最终是在朱老师的帮助下，让我加入了王老师的 863 项目。王老师虽然平时工作特别忙，但也总是抽时间给我们开会，帮我们指引方向和分享他多年来学术研究的经验。也是在两位老师严格要求和悉心指导下，才使我顺利的完成了学位论文的工作。

其次，非常感谢实验室的卢笑博士。在整个三年的学习中，无论是项目还是论文中遇到的问题，师姐都很热心的帮忙解答，总能给我一些很实际的建议，让我少走了很多的弯路。平时还会推荐我看一些的专业相关的论文资料，帮助我了解很多经典算法以及发展的方向。还需要特别感谢实验室的其他同学澎湃、张楚金、高小龙、张斌、旷小红等，由于实验都需要大量的样本，所以每次都麻烦他们帮忙拍照片、录视频，这样才使得论文的实验部分能够顺利完成。衷心的感谢大家无私的帮助。

最后要感谢我的父母，无论是在经济上还是精神上他们都给予了我极大的帮助。每次遇到困难时，他们也总是开导我，帮我分担，也从来不会因为家里的事而打扰我的学习，使我能够专心的学习。他们也一直是我努力奋斗的动力，非常感谢他们这些年来对我的无私奉献和支持。

感谢这三年中所有帮助和支持过我的人，向你们表示最崇高的敬意！

感谢参与学位论文审阅和答辩的专家老师们，对于论文中存在的纰漏，敬请各位老师指正，不胜感激。

## 附录 A 攻读学位期间所发表的学术论文目录

- [1]周云鹏,朱青,王耀南,卢笑,凌志刚.面部多特征融合的驾驶员疲劳检测方法.电子测量与仪器学报,2014,10(28):1140-1148

## 附录 B 攻读硕士学位期间参与的科研项目

[1]参与机器视觉感知与智能车载控制项目（国家 863 重点项目），该项目已经与中国公路咨询工程咨询有限公司联合申报了中国公路学会科学技术奖的特等奖。

[2]参与自主车辆的高质量三维场景认知与导航避障控制方法研究项目（国家自然科学基金项目）。

[3]参与基于物联网的车路协同主动安全技术研究及工程应用项目（交通运输部科技项目）。

三个项目主要是面向智能交通的视觉感知领域，以机器视觉控制技术为核心，主要研究复杂环境下智能交通信息快速检测和识别等技术，本人在三个项目中主要负责驾驶员疲劳检测部分，通过融合驾驶员面部多种疲劳特征来判断驾驶员的疲劳状态。