



汽车技术创新与研发系列丛书

驾驶员视觉分散 特征识别及检测方法

路玉峰 朱淑亮 著



机械工业出版社
CHINA MACHINE PRESS

汽车技术创新与研发系列丛书

驾驶员视觉分散特征 识别及检测方法

路玉峰 朱淑亮 著



机械工业出版社

《驾驶员视觉分散特征识别及检测方法》一书从视觉分散对驾驶员驾驶能力的影响研究入手,开展驾驶员视觉分散检测技术的研究,并重点研究基于视频图像分析的驾驶员视觉分散特征的提取方法。具体内容包括:研究视觉分散对驾驶性能的影响;研究多姿势下驾驶员面部、面部特征点精确定位的方法;研究驾驶员面部姿势的提取方法,提出利用核主元分析估计驾驶员面部姿势的方法;研究驾驶员视线方向提取方法,提出基于 Multi-PCA (多主元分析)的视线方向估计方法;研究 PCA 实现原理,分析常用 PCA 应用于识别时存在的问题;研究驾驶员转向行为识别,提出根据手部位置标准差来识别驾驶员转向行为的方法。

本书适合从事汽车设计、汽车理论研究等工作的技术人员阅读参考,也适合车辆工程等专业的研究生学习。

图书在版编目 (CIP) 数据

驾驶员视觉分散特征识别及检测方法/路玉峰,朱淑亮著. —北京:机械工业出版社, 2015. 7

ISBN 978-7-111-50151-0

I. ①驾… II. ①路…②朱… III. ①驾驶员 - 视觉识别 - 研究
IV. ①B842. 2

中国版本图书馆 CIP 数据核字 (2015) 第 094052 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑:杜凡如 责任编辑:刘 焯

版式设计:赵颖喆 责任校对:张晓蓉

封面设计:张 静 责任印制:刘 岚

北京圣夫亚美印刷有限公司印刷

2015 年 7 月第 1 版第 1 次印刷

169mm × 239mm · 11.75 印张 · 225 千字

0001—1500册

标准书号:ISBN 978-7-111-50151-0

定价:39.80 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

服务咨询热线:010-88361066

机工官网:www.cmpbook.com

读者购书热线:010-68326294

机工官博:weibo.com/cmp1952

010-88379203

金书网:www.golden-book.com

封面无防伪标均为盗版

教育服务网:www.cmpedu.com



前言 Preface

据不完全统计,全世界每年约有 60 万人死于交通事故,约 1000 万人受伤,其中驾驶员注意分散是引起汽车交通事故的重要原因,占撞车事故的 25% ~ 37%。相关研究表明,驾驶员注意分散的频率与发生事故的概率具有很高的相关性(相关系数为 0.72),相比较于正常情况下,驾驶员分心时(如拨打电话)发生事故的概率将提高 3 倍。

驾驶员视觉分散研究是驾驶员注意分散研究的一个重要组成部分,在最初的研究中没有详细分类注意分散的类型,而是把所有的类型统称为注意分散。驾驶员视觉分散是众多交通事故的诱因,并且随着车载信息系统的增加,将会引起驾驶员越来越多的视觉分散行为,从而引发更多的交通事故。检测驾驶员视觉分散并警告驾驶员,可减少类似原因造成的交通事故。

本书从视觉分散对驾驶员驾驶能力的影响研究入手,开展驾驶员视觉分散检测技术研究,并重点研究基于视频图像分析的驾驶员视觉分散特征的提取方法。主要包括:

(1) 设计实验让驾驶员阅读 4 处位置上的 2 类文本信息,使其产生 8 种不同的视觉分散。分析驾驶员观察不同位置处视觉分散时车辆的 SDLP (偏离道路中心距离的标准差),以研究视觉分散对驾驶性能的影响。根据驾驶过程中驾驶员视线变化的特点,建立基于驾驶员面部姿势与视线方向识别,并包含转向行为识别的视觉分散检测模型。

(2) 研究多姿势下驾驶员面部特征点精确定位的方法。研究利用肤色混合高斯模型预定位人脸区域,然后根据眉毛、嘴唇位置精确定位驾驶员面部的方法。针对眉毛区域灰度值低、变化剧烈的特点,研究基于联合投影函数定位驾驶员眉毛上边缘的方法。研究背景滤除的方法,克服面部横摆角度较大时眉毛定位不准的缺点。研究利用唇色多项式模型及嘴唇比人脸肤色更红的特点定位驾驶员嘴唇区域下边缘。



(3) 研究驾驶员面部姿势的提取方法, 提出利用核主元分析估计驾驶员面部姿势的方法。研究获取标准样本图像的方法, 设计样本图像采集系统。利用核主元分析把高维面部图像存在的流形结构嵌入到二维空间, 建立估计面部姿势的标准曲线。根据姿势曲线拟合圆, 提出利用拟合圆心及姿势曲线上与新投影点最近的 2 个点, 来估计新投影图像对应角度的方法, 估计精度可满足一定实际需要。

(4) 研究驾驶员视线方向提取方法, 提出基于 Multi-PCA (多主元分析) 的视线方向估计方法。针对驾驶环境中精确提取视线方向的困难, 把视线方向分为 5 类 (上、下、左、右、前), 为每种视线方向的图像建立特征空间, 利用主元分析提取每类视线的共有的统计特征, 再根据测试样本在每类特征空间下的重构误差来进行分类。该方法充分运用了 PCA 变换的最佳逼近性能, 并提取每类视线图像的独有特征, 实验结果表明该方法可以获得更高的识别准确率。

(5) 研究驾驶员转向行为识别, 提出根据手部位置标准差来识别驾驶员转向行为的方法。驾驶员在十字路口处的转向过程中, 视线方向偏离车辆前方的时间将超过 2s, 检测系统会误认为是视觉分散。研究驾驶员驾驶过程中双手位置变化的特点, 根据转向过程中驾驶员手部位置变化剧烈的特性, 提出利用双手位置标准差识别驾驶员转向行为的方法, 以防止将驾驶员转弯行为误判为视觉分散。

(6) 研究结构化道路的检测方法。提出基于领域主成分分析法的道路图像平滑方法。为增强道路图像对比度, 作者提出基于 Sin 函数的图像拉伸曲线; 提出图像分区垂直积分投影的车道线检测方法; 针对投影曲线平滑后曲线轮廓易改变的缺点, 提出新平滑方法, 该方法能在保持原始投影曲线形状基本不变的前提下进行平滑。作者研究了车道线边缘点的调整方法, 以便能更加精确地检测出车道线边缘。为提取车道标志线的特征参数, 采用最小二乘法对车道线进行边缘拟合。

(7) 驾驶疲劳作为视觉分散的特殊情况, 本书采用多信息融合技术对驾驶员疲劳进行检测。根据采集信息特点, 采用分布式信息融合结构, 在决策级根据粗糙集理论进行驾驶员疲劳信息融合判断。针对粗糙集理论的要求, 研究条件属性的离散归一化问题, 建立单个检测特征量与驾驶员疲劳程度之间的量化关系, 研究数据的约简方法, 得出最小决策算法, 通过决策算法判断驾驶员疲劳程度。

由于著者水平有限, 书中不足之处, 敬请读者批评指正。

著 者



目 录

Contents

前言

第 1 章 绪论	1
1.1 研究背景及意义	1
1.1.1 研究背景	1
1.1.2 研究意义	2
1.2 驾驶员视觉分散研究历程	3
1.3 视觉分散对驾驶能力影响研究现状	4
1.3.1 评测方法	5
1.3.2 实验方法	6
1.4 驾驶员视觉分散检测研究现状	7
1.4.1 特定作业检测	8
1.4.2 认知分散检测	9
1.4.3 视觉分散检测	10
1.5 本书研究方法 with 主要内容	11
参考文献	12
第 2 章 视觉分散对驾驶能力影响分析及检测模型	17
2.1 概述	17
2.2 视觉分散影响机理	17
2.3 基于实车实验的视觉分散影响分析	19
2.3.1 实验用车及场地	19
2.3.2 实验内容	19
2.3.3 实验过程	21
2.3.4 结果分析	21
2.4 驾驶员视觉分散检测模型	24



2.4.1 驾驶员注意特征评价	24
2.4.2 驾驶员视线特点	25
2.4.3 视觉分散检测模型	26
2.5 本章小结	27
参考文献	27
第3章 多姿势下驾驶员面部精确定位	31
3.1 概述	31
3.2 基于运动信息的头部区域检测	32
3.2.1 运动目标检测方法	32
3.2.2 基于帧间差分法的驾驶员头部检测	33
3.3 建立肤色混合高斯模型	37
3.3.1 彩色空间选择	37
3.3.2 YCbCr 空间下的肤色混合高斯模型	38
3.4 驾驶员面部预定位	40
3.4.1 基于图像分解和 MSR 算法的光照补偿	41
3.4.2 肤色区域检测	50
3.4.3 驾驶员面部预定位	52
3.5 驾驶员面部精定位	54
3.5.1 联合投影函数	54
3.5.2 眉毛上边缘定位	55
3.5.3 嘴唇检测	57
3.5.4 驾驶员脸部精定位	63
3.6 驾驶员面部图像归一化	64
3.7 本章小结	66
参考文献	66
第4章 基于核主元分析的驾驶员面部姿势估计	70
4.1 概述	70
4.2 核主元分析	71
4.2.1 基本原理	71
4.2.2 实现方法	72
4.2.3 核函数及其参数选择	74
4.3 基于核主元分析的驾驶员面部姿势估计	75
4.3.1 驾驶员面部姿势模型	75
4.3.2 图像样本采集	75
4.3.3 建立姿势曲线	76



4.3.4 面部姿势估计	80
4.4 实例分析	82
4.4.1 面部横摆角估计	82
4.4.2 核函数对估计精度的影响	83
4.4.3 面部俯仰角估计	84
4.5 本章小结	85
参考文献	85
第5章 基于 Multi-PCA 的驾驶员眼睛视线方向识别	87
5.1 概述	87
5.2 Multi-PCA	88
5.2.1 K-L 变换	88
5.2.2 PCA 实现方法	90
5.2.3 Multi-PCA 算法	91
5.3 基于 Multi-PCA 的眼睛视线方向识别	92
5.3.1 驾驶员眼睛区域定位	92
5.3.2 建立特征空间	93
5.3.3 测试重构误差及特征空间相似度	94
5.3.4 Multi-PCA 与 PCA 比较与分析	96
5.4 本章小结	96
参考文献	97
第6章 驾驶员转向行为识别	99
6.1 概述	99
6.2 转向类型	100
6.3 驾驶员手部检测及位置分析	100
6.3.1 手部检测	100
6.3.2 手部位置分析	101
6.4 粒子滤波跟踪	104
6.4.1 粒子滤波跟踪原理	104
6.4.2 序贯重要采样	105
6.4.3 重采样	106
6.4.4 基于直方图的粒子滤波算法	106
6.5 驾驶员手部跟踪	107
6.5.1 状态转移模型	107
6.5.2 颜色特征及其观测模型	107
6.5.3 手部跟踪	107



6.6 驾驶员转向行为识别	108
6.7 本章小结	112
参考文献	112
第7章 车道检测研究	114
7.1 概述	114
7.2 结构化道路检测	115
7.2.1 图像预处理	115
7.2.2 车道标志线分区投影	120
7.2.3 标志线特征点检测	123
7.2.4 车道标志线拟合	125
7.2.5 车道标志线跟踪	126
7.2.6 实验分析	130
7.3 非结构化道路检测	131
7.3.1 图像预处理	132
7.3.2 彩色图像分割	133
7.3.3 纹理特征分析	138
7.3.4 图像恢复	139
7.3.5 边缘拟合	139
7.3.6 实验分析	139
7.4 车辆偏离分析	140
7.4.1 车辆轨迹模型	140
7.4.2 车道偏离率	141
7.4.3 车道偏离量变化率	143
7.4.4 实验分析	144
7.5 本章小结	145
参考文献	146
第8章 基于信息融合的驾驶员疲劳检测	148
8.1 概述	148
8.2 信息融合	148
8.3 基于粗糙集理论的驾驶员疲劳检测	150
8.3.1 粗糙集理论	150
8.3.2 驾驶疲劳相关条件属性分析	152
8.3.3 疲劳程度判断	157
8.4 本章小结	160
参考文献	160



第 9 章 检测系统设计与检测实验	163
9.1 概述	163
9.2 视觉分散检测实验	163
9.3 驾驶疲劳检测实验	169
9.4 本章小结	172
第 10 章 结论与展望	173
后记	177

1.1 研究背景及意义

1.1.1 研究背景

据不完全统计，全世界每年约有 60 万人死于交通事故，约 1000 万人在交通事故中受伤^[1]。近几年，我国每年约有 10 万人死于车祸，而车祸造成的直接经济损失每年约有 20 亿元^[2]，统计结果如表 1-1 所示。

表 1-1 我国交通事故统计

年 份	事故次数/(万次)	死亡人数/(万人)	直接经济损失/(亿元)
2006	37.88	8.95	14.9
2007	32.72	8.16	12.0
2008	26.52	7.35	10.1
2009	23.83	6.78	9.1
2010	39.06	6.52	9.3
2011	21.08	6.24	10.8
2012	20.42	6.00	11.7

根据欧洲及美国的统计数据发现，交通事故中的 80%~90% 是人为因素造成的，并且约有 85% 的事故是由于驾驶员的原因造成的，车辆本身的因素引起的事故约为 10%，而环境因素造成的事故也仅占 5% 左右^[3]。

美国国家公路交通安全管理局（NHTSA）与美国弗吉尼亚州科技运输协会（VTTI）于 2006 年 4 月 20 日发布一个调查报告，称驾驶员注意分散是大部分事故和近距离碰撞的主要因素，其中 80% 的碰撞和 65% 的几乎碰撞（Near Crash）都与驾驶员注意分散有关^[4]。高速公路的交通事故中有 50% 以上是由于驾驶疲劳，或由所见目标单调而造成注意分散、甚至打瞌睡等原因造成的^[5]。日本学者对众多交通事故进行了统计研究，发现驾驶员注意分散是众多交通事故的诱因，占到诱因总数的 36.0%^[6]。来自法国的统计表明高速公路上 1/3 的交通事



故是由驾驶员注意分散引起的^[7]。NHTSA^[8]公布一项包括 723 次碰撞事故原因的调查研究，1284 名驾驶员与这些事故有牵连，其中 22.7% 的碰撞事故是由驾驶员注意分散引起的，结果如表 1-2 所示。

表 1-2 引发交通事故诱因的百分比

碰撞事故原因	占总数百分比（%）
驾驶员注意分散	22.7
车辆速度	18.7
驾驶员饮酒	18.2
驾驶员感知错误	15.1
驾驶员决策错误	10.1
驾驶员能力限制	6.4

Stutts^[9]等分析了由警察所记录的 1995-1999 年之间发生在美国的交通事故，并对引起注意分散的部分原因进行总结，发现驾驶员因观察车外人、物、事件等而引起的注意分散占总数的 29.4%，分析结果如表 1-3 所示。

表 1-3 引起驾驶员注意分散原因的百分比

引起注意分散原因	占总数百分比（%）
车外人、物、事件	29.4
调节收音机、CD	11.4
车内其他人	10.9
移动物体	4.3
车内其他设备、物体	2.9
调节空调、车辆设备	2.8
吃食物、喝饮料	1.7
使用移动电话	1.5
吸烟	0.9
其他注意分散因素	25.6
未知注意分散	8.6

通过以上分析发现，虽然各研究机构统计结果不尽相同，但是都表明驾驶员注意分散是交通事故一个重要诱因。其中驾驶员最容易发生的注意分散类型是视觉分散^[10]，并且随着车载信息系统的增加以及交通环境的复杂化，引起驾驶员视觉分散的因素会越来越多。因此需要对驾驶员视觉分散进行重点研究，以找出相应解决办法。

1.1.2 研究意义

研究驾驶员视觉分散，有助于我们寻找减少其影响的措施，使行车更加安全；或者研究出一套检测驾驶员视觉分散的系统，在必要时对驾驶员进行警告，



预防此类交通事故,这具有重大社会意义。检测视觉分散的系统能够应用在多个方面,如汽车、飞机的驾驶等,只要操作人员视觉分散将引起严重后果的场合,都可以应用该技术产品。可以预见,如果这样的系统能够切实有效地工作,不但能减少人员伤亡,而且能避免因此产生的经济损失。

研究并检测驾驶员视觉分散还具有理论意义。研究驾驶员视觉分散需要运用到认知行为方面的理论,对驾驶员认知行为的研究必将促进认知科学的发展^[11]。检测驾驶员视觉分散需要涉及图像处理中多种相关算法,因此该方面研究可以促进图像处理相关领域的发展。另外,检测驾驶员视觉分散实际上是一种状态识别,许多问题是模式识别中长期以来都没有彻底解决的问题,并且已有的成果也不一定适合检测驾驶员的要求,因此该研究还可以促进模式识别技术的发展^[12]。

1.2 驾驶员视觉分散研究历程

驾驶员视觉分散研究是驾驶员注意分散研究的一个重要组成部分,在最初的研究中没有详细分类注意分散的类型,而是把所有的类型统称为注意分散。注意分散与驾驶疲劳一样,也会影响驾驶员的驾驶性能。驾驶员疲劳时发生交通事故的概率要比正常时高4~6倍^[4],因此驾驶疲劳问题更容易引起人们的关注,国内外学者也对该问题进行了较深入的研究。但随着驾驶员注意分散引起交通事故的增加,人们逐渐认识到驾驶员注意分散研究的重要性。

早在1915年人们就认为刮水器的单调重复动作会分散驾驶员的注意力。到20世纪30年代随着收音机设备进入驾驶室,人们也逐渐认识到该设备会引起驾驶员注意分散,因此美国许多州政府要禁止其应用^[13]。

而真正的研究则开始于20世纪60年代,如研究车载收音机对汽车驾驶员心理的影响等,结果表明听音乐有助于驾驶员更平稳地操纵汽车^[14]。随后Brown又研究了打电话对驾驶能力的影响,发现打电话时驾驶员的感知判断能力会受到不利影响,出现错误的概率将增加1.7%~22.2%^[15]。

随着信息技术在汽车上的应用越来越广泛,例如手机^[16]、导航^[17]、互联网^[18]、收音机^[19]、CD机^[20]等对驾驶能力的影响也受到重视。另外,与新技术无关的事物或者行为,如乘客、路边广告牌与指示牌、驾驶员自身行为(吃食物、喝饮料、吸烟)等也会分散驾驶员的注意力^[21,22]。资料表明,由于驾驶员在行驶过程中饮食时引起的交通事故数量比打电话时发生的交通事故量高0.2%,因为驾驶员吃东西、喝饮料发生的概率要远大于驾驶员打电话的概率^[9]。

到1990年前后开始有人提出减少驾驶员注意分散的措施,如出台相关法律



等^[16]。1993 年 Leiser^[23] 研究在车内使用语音交互的信息系统,以降低驾驶员的注意力分散。但后来有的研究认为使用语音交互的电子邮件会使驾驶员反应时间延长 30%, 出事故的概率增加 3.5% ~ 38.5%, 并且使碰撞时的速度增加 27.3% ~ 80.7%^[24]。因此,语音交互系统对减少驾驶员注意分散的作用不大^[25]。

进入 21 世纪,则有越来越多的研究者提出对驾驶员注意分散进行检测,以在危险情况下警告驾驶员。驾驶员出现注意分散时主要体现在眼睛特征、生理指标(脑电波、皮肤电导率)或者驾驶能力参数(对车辆速度、侧向位移、转向、加减速等进行控制)的变化上,根据这些参数的变化,研究者开发出众多注意分散检测办法。如 2001 年 Waard 等把行车参数作为判断指标,并利用神经网络综合各个指标来识别驾驶员是否在使用移动电话,正确识别率可以达到 89%^[26]。

2005 年 Pettitt 等为了研究驾驶员注意分散的统一性,对驾驶员注意分散进行了详细定义,并根据驾驶员受影响的信息指标,把驾驶员注意力分散的类型定为:视觉、认知、身体、听觉等四种分散形式^[27],为全面开展驾驶员注意力分散研究奠定了基础。

国内对驾驶员注意分散的研究起步较晚,并且针对驾驶员视觉分散的研究也较少。早期研究主要是感性的,主要分析了交通事故的诱因,真正研究开始于 21 世纪。2004 年前后吉林大学的王荣本、施树明、董兵亮等提出检测驾驶员嘴部状态来识别驾驶员注意分散的方法^[28,29]。2004 年 5 月,我国出台了新的交通安全法,对开车打手机的行为做了严格的交通法律、法规限制,罚款金额达到 200 元,并罚扣 3 分,以减少打手机引起的注意分散。2005 年 Chen^[30] 等通过研究使用车内信息系统对驾驶员视觉、听觉、精神等的影响,发现使用这些设备会降低驾驶员的反应速度,更重要的是如果驾驶员出现听觉分散,那么再采用一个声音信号警告驾驶员的话会使他的反应速度变得更慢。Liu^[31,32] 等通过对 12 位驾驶员实验研究发现,驾驶时使用手机会使驾驶员的反应时间增加 11.9%,行驶速度降低 5.8%。Liu Ning^[33] 等通过在驾驶模拟器上的实验研究认为驾驶员对前方车辆的反应时间、驾驶员的心率等可以作为评价驾驶注意分散的指标。

1.3 视觉分散对驾驶能力影响研究现状

表征驾驶能力的工作指标主要包括:车辆偏离道路中心位置的标准差(SDLP-Standard Difference of Lane Position)、速度保持和控制能力、对外部事件的反应时间、车距保持、劳动强度、对安全驾驶参数的警惕性等。研究者已经开发出许多不同的方法来研究视觉分散对这些参数的影响。



1.3.1 评测方法

1.3.1.1 多作业分析

人类只有有限的视觉资源用来处理各种作业。同时处理两种作业时,如果两种作业要求的视觉资源超出了驾驶员总的视觉资源,那么分配到某个作业的视觉资源将会减少^[34]。对驾驶员视觉分散来说,这些研究主要用于测试某个车内设备对驾驶能力的影响,或者其他事物对驾驶能力的影响,如驾驶员与乘客交谈等对驾驶车辆的影响,不过它的有效性还需进一步验证。

1.3.1.2 眼睛扫视研究法

驾驶时的视觉扫视行为早在20世纪60年代就得到广泛研究。测试视觉行为的方法有眼睛扫视记录和视线遮蔽技术。眼睛扫视技术通过记录眼睛对视野内特定物体扫视的频率和持续时间来测量视觉行为。在驾驶车辆时,驾驶员通过一系列简短(1~2s)的扫视来完成第二作业。眼睛扫视技术可以记录眼睛偏离道路的总时间,也就是完成某作业需要的视觉要求^[19]。视线偏离道路的总时间是一个被广泛接受的有效测量第二作业视觉需求的方法。可以应用视频记录仪来记录驾驶员的眼睛和头部动作。通过分析记录的视频来计算眼睛扫视数据^[17]。当今成熟的跟踪设备已经能实时计算眼睛的扫视路径、眼睛闭合时间、头部转动等。

1.3.1.3 视觉遮蔽技术

尽管眼睛跟踪系统很先进,但是这些系统也很昂贵、费时、难安装、难校准。Green^[18]等于2001年提出采用视觉遮蔽技术作为眼睛跟踪系统的一个变通。该技术利用一个眼罩,以不同的时间间隔来部分或者完全遮住驾驶员眼睛,模拟驾驶时驾驶员与一个车内设备交互时视线在道路与设备间的转换。由于视觉遮蔽技术的成本小且容易实施,被很多研究者认为是很有潜力的测试方法。

1.3.1.4 15s 规则

美国汽车工业协会(SAE)根据车内导航系统,发布一个注意分散最高允许水平的标准^[17]。标准认为如果一个作业在静止车辆上能在15s内完成,那么这种作业引起的驾驶员注意分散不严重,可以在驾驶时完成。与其他方法相比,此标准比较简单易行。但是实验发现,静止车辆上完成作业的时间与运动车辆上完成的时间没有相关性,因此该标准的效果不是很好^[35]。并且该规则没有考虑到速度控制、目标检测等作业。

1.3.1.5 建立驾驶员认知模型的方法

2001年 Salvucci^[11,36]提出建立驾驶员注意分散认知模型,并在ACT-R理论框架下仿真注意分散对驾驶能力的影响。通过在驾驶模拟器上验证仿真的结果,



发现可以用建立模型的方法来评测注意分散对驾驶能力的影响。该方法的优点是成本低，不需要特殊实验设备，但是建立驾驶员模型比较困难，建模软件——ACT-R 不容易掌握。

这些评测方法都有各自的优缺点。到底应用哪种方法合适，主要取决于要评测何种人机交互界面，特别是该种行为主要影响驾驶员的哪个方面。一般研究中不使用 15s 规则，因为它的有效性还有待于进一步的验证。

1.3.2 实验方法

1.3.2.1 实车实验法

研究视觉分散对驾驶能力影响的最直接方法，是在道路上采集实际车辆行驶数据，该方法要求受试人员在道路上驾驶车辆，并完成各种视觉分散的作业，车辆上安装自动收集表明驾驶能力数据的仪器。然后，将收集的数据与基准数据进行比较，基准数据通常是正常驾驶时驾驶员的表现。该方法采集的数据来自真实环境，因此比较可靠，但是也比较耗时、昂贵，因此很少用来测量驾驶员视觉分散。短距离测试轨迹的方法也能获得真实的数据，也经常用来测试视觉分散对驾驶能力的影响^[37,38]。该方法让受试者在一条测试道路或者封闭道路上驾驶车辆。通过自动仪器或者人工观察收集数据，然后与基准数据比较。此法接近真实环境，并且在一个封闭路上驾驶降低了危险性^[16]。但是，操作者使用仪器的熟练性和人工记录的准确性将影响测量准确性。

1.3.2.2 驾驶模拟器实验法

驾驶模拟器能让受试者在一个相对真实并且安全的环境下进行测试，因此很多研究者应用驾驶模拟器来测试视觉分散。1999 年 Reed^[39,40]等应用驾驶模拟器对驾驶员视觉分散进行研究，并与实车实验的结果进行比较，认为驾驶员在实车与驾驶模拟器上的驾驶能力表现相同。模拟器的特性不同会影响测量的真实性和有效性。高质量的模拟器能提供相对真实的驾驶环境，具有布局合理的彩色树木、建筑物、路标等，并能使车辆振动、摇摆，以给驾驶员一个更加真实的驾驶感觉。

相比在道路和试车跑道上的实车实验，驾驶模拟器具有很多优势。例如，驾驶模拟器更安全、驾驶环境更容易控制、改造成本低，并且容易收集大量代表驾驶能力的参数，如速度控制、侧向偏移，还有驾驶员的眼动和扫视行为。当然，驾驶模拟器也有不利的因素。模拟器与车辆的不同会影响测试结果，并且驾驶员在了解模拟器不能对他们造成伤害的前提下，容易做些对安全驾驶不利的事情，如双手离开转向盘、视线偏离道路场景等也会影响实验结果。



1.4 驾驶员视觉分散检测研究现状

驾驶员视觉分散是影响驾驶能力的最主要因素,目前减少驾驶员视觉分散的方法大体有:

1) 开发不占用视觉信息的人机交互界面,如研究基于语音的信息交互系统。但基于语音的电子邮件系统仍会使驾驶员出事故的概率增加 $3.5\% \sim 38.5\%$ ^[24],因此语音交互系统对减少驾驶员视觉分散的作用不大^[25]。

2) 建立防止视觉分散的法规制度,比如禁止安装车载收音机、禁止驾车时使用手机。执行这样的规定,虽然可以减少驾驶员视觉分散现象,但也损害了其有益用处。如驾驶车辆时听音乐可以缓解驾驶员的疲劳,在紧急情况下可以使用手机报警。并且,有人研究认为在交通不算繁忙的高速公路上打电话并不影响驾驶员的驾驶能力^[41]。

3) 对驾驶员的注意力状态进行检测,当发现驾驶员视觉不集中时对其进行适当预警,将其注意力转移到驾驶过程中。21世纪开始,越来越多的人研究驾驶员状态检测的问题。

通过检测驾驶员的生理指标、行为、驾驶能力,可以判断驾驶员是否发生视觉分散行为。首先,驾驶员发生视觉分散与驾驶能力指标(对速度、侧向位移、转向、加减速等的控制)的变化有很强相关性^[34]。例如,在观察驾驶室内设备时会使平均车速降低 $5\% \sim 20\%$,使车辆侧偏距离的标准差增大 10% 以上,并且会使驾驶员的小范围内变换转向的频率增加 10% 左右。另外,根据驾驶员出现视觉分散时眼睛特征、生理指标(脑电波、皮肤电导率),或者驾驶能力的变化规律,可以识别出驾驶员是否出现视觉分散行为^[42]。

驾驶员视觉分散的检测可以分为3个步骤,参数获取、状态判断、发出警告。

1) 应用到的参数主要有驾驶员特征、行车参数。常用驾驶员特征主要包括视线方向、面部姿势、心率、皮肤电导率等,提取驾驶员的心率或者皮肤电导率等需要与驾驶员身体接触的设备,会造成驾驶员的不适。因此人们一般采用摄像机来获取驾驶员面部图像,再采用图像处理等方法来提取驾驶员的视线方向或者面部姿势等^[43]。应用的行车参数主要有车速、转向角度、侧向位移、加减速等,这些参数需要特定的传感器来测量。

2) 状态判断主要是区分驾驶员是否发生视觉分散,可以根据设定的阈值直接判断或者利用数据融合法判断,采用的融合方法有HMM^[44](隐马尔科夫模型)、ANN^[45](人工神经网络)、SVM^[46](支持向量机)等。

3) 发出警告是在检测出驾驶员视觉分散以后,采用声、光或者振动信号使



驾驶员的视觉注意力恢复到其应该注意的地方,但这个提醒信号不能加深驾驶员的视觉分散^[47,48]。

1.4.1 特定作业检测

对驾驶员视觉分散的检测开始于对具体行为的检测,如对驾驶员使用手机的检测,并在研究初始阶段统称为注意分散检测,图 1-1 列出部分比较常见的注意分散作业。



图 1-1 部分注意分散作业

2001 年 Waard^[26]把驾驶员 10s 内的行车参数(转向盘转角、加速踏板位置、车辆的侧偏距离等 14 项)作为判断依据,并利用神经网络综合这些指标来识别驾驶员是否在使用手机,正确识别率为 89.3%,但误识率较高(7.1%)。Edenborough^[49]根据检测驾驶员注意分散的需求,提出汽车级、实时性的图像检测系统所应满足的具体条件,并提取了驾驶员的眼睛特征与面部姿势。2004 年前后,吉林大学的王荣本、施树明、董兵亮等应用图像处理获取驾驶员嘴部特征,然后利用 BP 神经网络判断驾驶员是否在讲话^[28,50]。Yamakita^[51,52]采用图像处理的方法获得驾驶员视线方向在 2 维平面上分布变化、眨眼频率、嘴巴状态,并根据 3 项参数的变化来判断驾驶员是否在打电话,可以识别出 3 种影响程度不同的打电话行为,在 3 种情况下的正确识别率可达 90%,但在个别情况下会低于 50%。为提高图像处理的速度,并且便于检测系统在驾驶室内的安装,Moreno^[53]构造了基于 DSP 的驾驶员状态检测系统框架,可以实时处理驾驶员面部图像,但是没有进行具体的注意分散检测实验。图像处理方法是提取驾驶员参数最好的方法之一,然而所获取图像的质量是一个关键性因素。因此 Michimasa^[54]等开发了一个可以嵌入到后视镜里面的图像传感器,并在驾驶模拟器中进行实验,证明采集的图像可满足检测驾驶员眼睛状态的要求。

在交叉路口左转向时发生的碰撞事故主要是由于驾驶员没有注意红色交通信号引起的。White^[47]提出利用车辆在交叉路口处的速度与加速特性对驾驶员是否注意红灯进行检测。在检测到驾驶员未注意红灯信号时,提出利用闪光作为警告使驾驶员的注意力集中到信号灯上。警告信号不但要能使驾驶员意识到自



已发生注意分散,还要能把注意力吸引到驾驶员应该注意的地方。Cristy^[48]研究发现利用空间声音信号或者振动触觉的警告方法,在80%的情况下可以改变驾驶员视觉注意方向到指定的方位。

引起驾驶员注意分散的事物各式各样,因此不可能按驾驶员的具体行为来检测所有注意分散,把驾驶员注意分散按一定规律分为几种类型进行检测可以解决这个问题。根据注意分散对驾驶员不同方面的影响,可分为4类:认知分散、视觉分散、听觉分散、身体分散,每类注意分散的影响都有各自特点。认知分散会使驾驶员的视线过多地集中于道路中心,使驾驶员在整个视野范围内的事件检测能力降低,但是对行驶路线的控制有很小的影响。而视觉分散会使驾驶员的视线过多偏离前方道路,使驾驶员不能很好地保持行驶路线,并且反映不同形式注意分散的参数也不相同。2007年Pohl^[55]发现人脸朝向可以反映视觉分散,而认知分散主要可由眼睛视线方向特征反映出来。因此,可以对驾驶员不同形式的注意分散进行分类检测,可以实现大范围内的驾驶员注意分散检测。影响驾驶员最深的是认知分散、视觉分散,现在研究也主要围绕这两个方面进行检测。

1.4.2 认知分散检测

驾驶员在出现认知分散时,视线方向会发生特定变化,眼睛在5s内扫视环境的次数将会由7.53降低到6.72,并且减少对周围环境与仪表的观测,将更多地注视道路中心15°的范围之内^[56]。作为对认知作业影响的补偿,在损害周围视觉的代价上,驾驶员会增加观察道路中心的时间,并在空间上集中视觉焦点于道路中心。驾驶员的视线变化虽然与驾驶环境的要求、是否饮酒、疲劳程度有关,但视线集中于道路中心的特点与认知分散具有很强的相关性^[57]。例如,视线在水平与垂直方向上的变化范围会显著降低,更长时间盯视道路,观察仪表、后视镜等的频率都有所降低。

Victor^[57]等研究了驾驶员出现认知分散时视线、生理指标、侧向性能与径向参数的特点。听觉作业产生的认知负担会使驾驶员的道路保持能力有所改进,因为这时驾驶员更多地采取微调转向盘的方法。另外,还发现听觉作业产生的认知负担多于视觉负担,也会导致驾驶员更多地观察道路中心。Victor采用在一个时间段内视线集中于道路中心的时间来识别认知分散,并在集中时间大于92%时给出警告(风窗玻璃出两道闪光),并且认为视线角度的标准差对检测认知分散更敏感些。

在分类方法上,大量研究发现SVM分类方法要好于其他的分类方法。如SAVE-IT(Safety Vehicle Using Adaptive Interface Technology)项目研究发现SVM分类效果要好于HMM^[46]。AIDE(Adaptive Integrated Driver-vehicle Interface)项



目与 Kutila 都采用了 SVM 分类的方法^[44]，并且 Kutila 把驾驶员的视线、面部姿势与行车参数（道路偏移）融合到一起对驾驶员认知分散进行识别，识别率可达 65%~80%。

Liang^[45,46]等利用驾驶员驾驶过程中一段时间内的眼睛运动特征（扫视、注视等）、驾驶特征（注视时间、反应时间、车辆偏离道路的距离、转向盘的变化等）的 19 个参数建立 Bayesian 神经网络检测模型与 SVM 检测模型，发现 SVM（检测率为 96.08%）的检测效果要好于 Bayesian 神经网络（检测率为 80.1%），并且 SVM 的训练速度也比 Bayesian 神经网络快。两种方法都力图实现实时检测驾驶员认知分散，但是由于传感器、计算模型等原因未能实现实时检测。

1.4.3 视觉分散检测

驾驶员视觉分散主要指驾驶员视线偏离车辆前方的行为，在这里驾驶员观测后视镜也被认为是一种视觉分散行为。Kutila^[7]对驾驶员正常状态与视觉分散状态下的视线进行深入研究，发现驾驶不同型号车辆时驾驶员的视线范围有所不同，驾驶轻型车时驾驶员的视线范围：俯仰角的变化范围是 $-5^{\circ} \sim 12^{\circ}$ ，横摆角的变化范围是 $-10^{\circ} \sim 8^{\circ}$ 。Hughes^[58]研究发现驾驶员视线俯仰角与横摆角的变化范围都在 $\pm 12^{\circ}$ 内，如果驾驶员的视线方向偏离这个角度范围，则认为驾驶员发生了视觉分散。

Zhang^[59]通过研究在视觉分散情况下，驾驶员在驾驶模拟器上反应时间、行驶路线等表现，发现驾驶员视线偏离的持续时间、偏离的角度与反应时间、路线偏离的标准差（SDLP-Standard Deviation of Lane Position）之间有很强相关性。视线偏离的持续时间每增加 25%，驾驶员对前车制动信号的反应时间延长 0.39s，SDLP 增大 6cm。因此根据驾驶员的视线偏离与持续时间、发生频率等可以判断驾驶员是否发生视觉分散，一般不需要采用 SVM 等复杂的分类方法。

Batista^[60]提出一个根据视频分析来检测驾驶员注意水平的框架，但该系统要依赖于瞳孔的检测，需要用到复杂的照明设备，因此不容易在驾驶室内实现。其检测瞳孔的主要目的是跟踪驾驶员的面部姿势，并且实验中仅检测了驾驶员的疲劳程度。

为便于检测系统在驾驶室内实现，Smith^[61]等提出利用单个摄像机检测驾驶员的视觉注意。首先，获取驾驶员的头面部图像，然后根据眼睛与人脸边缘的距离等特征提取驾驶员的面部姿势以及视线方向，得到的姿势与视线角度不是很精确；并且随着驾驶员眼睛状态的变化，其眼睛的特征点会更加不稳定，因而面部姿势的估计偏差越来越大。

驾驶员在变道时要转移视线来提前观察对应道路，因此容易把驾驶员此时的情况误判为危险状态。Lee^[62]等提出采用两个摄像机分别获取驾驶员视线方向



与道路方向、车辆行驶方向。然后根据视线方向与道路方向的关系、道路方向与车辆行驶方向的关系判断当时驾驶环境的危险程度,可以减少变道时的误判。但这种方法在车辆经过十字路口、T 字路口而转向时并不能有效降低误判,因为此时驾驶员需要较多地观测车辆两侧,而在这样的路口并没有合适的分道线来判断车辆是否在转向。因此,如果能增加一个驾驶员转向行为识别系统,可以进一步提高视觉分散检测率。

1.5 本书研究方法与主要内容

综合前面驾驶员视觉分散研究的现状以及检索到的文献,可以看出当前对驾驶员视觉分散的研究多集中在国外,我国在这方面的研究还较少。研究驾驶员视觉分散的目的是减少视觉分散对驾驶员的影响,目前比较可行的方法是检测驾驶员视觉分散,以给出适当警告。检测驾驶员视觉分散的核心是提取驾驶员视觉分散特征,如视线方向等的提取,目前主要是利用眼动仪来获得驾驶员的视线方向,而眼动仪高昂的价格限制了视觉分散检测系统的实际应用潜力,因此需要开发相对廉价的检测设备。另外,目前的视觉分散检测系统中,大多没有考虑驾驶员转向行为对视觉分散检测的影响,容易造成虚警现象的发生。因此,本书提出采用两个摄像机获取驾驶员面部、手部图像,利用图像处理、模式识别理论提取视线特征,以检测驾驶员视觉分散的方法,并根据手部运动特征判断驾驶员的转向行为,以减少虚警现象。本书还对驾驶员被动视觉分散(驾驶疲劳)检测方法进行了初步研究。

本书主要研究内容如下。

1. 研究视觉分散对驾驶性能的影响

作者提出采用实车实验的方法,研究同一水平面上不同位置视觉分散对驾驶能力的影响;研究视觉分散影响驾驶员的机理;研究不同车速下视线偏离角度与路线保持能力之间的相关性;分析基于脑电、皮肤电、行为评测驾驶员注意水平的方法;确定检测驾驶员视觉分散的视线变化范围,根据应用需要给出驾驶员视觉分散检测模型。

2. 研究多姿势下精确定位驾驶员面部的方法

研究基于人脸肤色混合高斯模型检测驾驶员面部图像的方法,研究驾驶员面部图像的光照补偿方法。研究驾驶员眉毛定位方法,并设计联合投影函数;研究利用多项式模型及唇色特点的方法定位驾驶员嘴唇区域;研究驾驶员面部图像归一化的方法。

3. 研究驾驶员面部姿势估计的方法

作者提出基于核主元分析估计驾驶员面部姿势的方法;分析核主元分析实



现原理,研究利用核主元分析估计面部姿势步骤;研究获取标准样本图像的方法,设计样本图像采集系统,确定估计面部姿势的标准曲线;研究姿势曲线的特点,根据姿势曲线拟合圆,提出利用拟合圆心及姿势曲线上与新投影点最近的2个点,来估计新投影图像对应角度的方法;研究面部俯仰角的估计方法、研究面部横摆角的估计方法;研究不同的核函数、核函数参数的变化对估计面部姿势的影响。

4. 研究驾驶员眼睛视线方向的识别方法

研究 PCA 实现原理,分析常用 PCA 应用于识别时存在的问题;针对驾驶环境中精确提取视线方向的困难,提出应用 Multi-PCA (多主元分析) 来识别视线方向的方法;研究适合视线方向提取的眼睛区域定位方法;研究不同视线特征空间的相似度,以证明方法的有效性。

5. 转向行为识别方法

研究不同驾驶行为中驾驶员双手位置变化的特点;研究基于粒子滤波算法的驾驶员手部跟踪方法;研究不同的驾驶行为的特征。作者提出了根据双手位置的标准差来识别驾驶员的转向行为,提出了驾驶员转向行为识别算法。

6. 研究运用信息融合技术检测驾驶员疲劳程度的方法

采用分布式信息融合结构,在决策级根据粗糙集理论判断驾驶员疲劳程度;建立单个检测特征量与驾驶员疲劳程度之间的量化关系;针对粗糙集理论要求,研究检测特征量的离散归一化问题,并对数据进行约简,得出最小决策算法;根据最小决策算法判断驾驶员疲劳程度。

参考文献

- [1] D. Dawson, N. Lamond, K. Donkin, et al. Quantitative Similarity between the Cognitive Psychomotor Performance Decrement Associated with Sustained Wakefulness and Alcohol Intoxication [C]. Proceedings of the Third International Conference on Fatigue and Transportation. Western Australia, 1998: 1-3.
- [2] 司银霞. 限速标志对驾驶行为影响研究 [D]. 长春: 吉林大学. 2006: 2.
- [3] D. Parker, J. T. Reason. Driving Errors, Driving Violations and Accident Prevention [J]. Ergonomics. 1995, 38 (1): 1036-1048.
- [4] S. G. Klauer, T. A. Dingus, V. L. Neale. The Impact of Driver Inattention on Near-Crash/Crash [C]. Virginia: National Highway Traffic Safety Administration, 2006: 1-10.
- [5] 郑培, 宋正河, 周一鸣. 机动车驾驶员驾驶疲劳测评方法的研究状况及发展趋势 [J]. 中国农业大学学报. 2001, 6 (6): 101-105.
- [6] 晓阳. 驾车分心是交通事故的主因 [J]. 农业机械化与电气化. 2001, (2): 28.
- [7] M. Kutila. Methods for Machine Vision Based Driver Monitoring Applications [D]. Tampere



- University Doctor Thesis. 2006: 16-17.
- [8] D. L. Hendricks, J. C. Fell, M. Freedman. The Relative Frequency of Unsafe Driving Acts in Serious Traffic Crashes [C]. Washington, D. C.: National Highway Traffic Safety Administration Office of Research and Traffic Records Research and Evaluation Division, 1999: 1-10.
- [9] J. C. Stutts, D. W. Reinfurt, L. Staplin. The Role of Driver Distraction in Traffic Crashes [C]. Washington, D. C.: AAA Foundation for Traffic Safety, 2001: 5-20.
- [10] M. Basset, C. Cudel, V. Georges, et al. Visual Characterization of the Road Driver's Behaviour [C]. IEEE International Workshop on Intelligent Signal Processing. Faro, Portugal. 2005: 288-291.
- [11] 刘雁飞, 吴朝晖. 驾驶 ACT-R 认知行为建模 [J]. 浙江大学学报 (工学版). 2006, 40 (10): 1657-1662.
- [12] 李衡峰. 基于综合集成的驾驶疲劳识别 [D]. 长沙: 中南大学. 2005: 1-8.
- [13] R. Karlsson. Evaluating Driver Distraction Countermeasures [D]. Linköping University. 2004: 1.
- [14] I. D. Brown. Effect of a Car Radio on Driving in Traffic [J]. Ergonomics, 1965, 8 (4): 475-479.
- [15] I. D. Brown, A. H. Tickner, D. C. Simmonds. Interference between Concurrent Tasks of Driving and Telephoning [J]. J. Appl. Psychol, 1969, 53 (5): 419-424.
- [16] M. J. Goodman, F. D. Bents, Tijerina. An Investigation of the Safety Implications of Wireless Communication in Vehicles [C]. Washington, D. C.: Department of Transportation NHTSA, 1997: 1-10.
- [17] E. Farber, J. Foley, S. Scott. Visual Attention Design Limits for ITS In-Vehicle Systems: The Society of Automotive Engineers Standard for Limiting Visual Distraction while Driving [C]. Transportation Research Board Annual General Meeting, Washington D. C. USA, 2000: 2-3.
- [18] K. Young, M. Regan, M. Hammer. Driver distraction: A Review of the Literature [D]. Victoria Australia: Holden Ltd, 2003: 1-7.
- [19] D. Haigney. Mobile (Cellular) Phone Use and Driving: A Critical Review of Research Methodology [J]. Ergonomics. 2001, 44 (2): 132-143.
- [20] J. W. Jenness, R. J. Lattanzio, M. Toole. Voice-Activated Dialing or Eating a Cheeseburger: Which is More Distracting during Simulated Driving [C]. Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, Pittsburgh, PA., 2002: 12.
- [21] B. Simons-Morton, N. Lerner, J. Singer. The Observed Effects of Teenage Passengers on the Risky Driving Behavior of Teenage Drivers [J]. Accident Analysis & Prevention. 2005, 37 (6): 973-982.
- [22] B. Wallace. Driver Distraction by Advertising: Genuine Risk or Urbanmyth? [J] Municipal Eng. 2003, 156 (3): 185-190.



- [23] R. Leiser. Driving Future Vehicles [M]. UK: Routledge, 1993.
- [24] J. D. Lee, B. Caven, S. Haake, et al. Speech-based Interaction with In-Vehicle Computers: The Effects of Speech-Based E-Mail on Drivers' Attention to the Roadway [J]. Human Factors. 2001, 43 (4): 631-639.
- [25] T. A. Ranney, J. L. Harbluk, Y. I. Noy. Effects of Voice Technology on Test Track Driving Performance: Implications for Driver Distraction [J]. Human Factors. 2005, 47 (2): 439-454.
- [26] D. D. Waard, K. A. Brookhuis, N. Hernandez- Gress. The Feasibility of Detecting Phone-Use Related Driver Distraction. International Journal of Vehicle Design [J]. 2001, 26 (1): 85-95.
- [27] M. Pettitt, G. Burnett, A. Stevens. Defining Driver Distraction [C]. World Congress on Intelligent Transport Systems, San Francisco, 2005: 1-12.
- [28] Wang Rongben, Guo Lie, Tong Bingliang, et al. Monitoring Mouth Movement for Driver Fatigue or Distraction With one Camera [C]. The 7th International IEEE Conference on Intelligent Transportation Systems, 2004: 314-319.
- [29] 童兵亮. 基于嘴部状态的疲劳驾驶和精神分散状态监测方法研究 [D]. 长春: 吉林大学. 2004: 29-41.
- [30] W. H. Chen, C. Y. Lin, J. L. Doong. Effects of Interface Workload of In-Vehicle Information Systems on Driving Safety [J]. Transportation Research Record. 2005, 19 (37): 73-78.
- [31] B. S. Liu, Y. H. Lee. Effects of Car-Phone Use and Aggressive Disposition During Critical Driving Maneuvers [J]. Transportation Research Part F: Traffic Psychology and Behaviour. 2005, 8 (4): 369-482.
- [32] Liu Bor-Shong, Lee Yung-Hui. In-Vehicle Workload Assessment: Effects of Traffic Situations and Cellular Telephone Use [J]. Journal of Safety Research. 2006, 37 (1): 99-105.
- [33] Liu Ning, Zhang Kan, Sun Xianghong. The Measurement of Driver's Mental Workload: A Simulation-Based Study [C]. International Conference on Transportation Engineering, Chengdu, China, 2007: 1187-1193.
- [34] J. Engström, E. Johansson, J. Östlund. Effects of Visual and Cognitive Load in Real and Simulated Motorway Driving [J]. Transportation Research Part F: Traffic Psychology and Behaviour. 2005, 8 (2): 97-120.
- [35] L. Tijerina, E. Parmer, M. J. Goodman. Driver Workload Assessment of Route Guidance System Destination Entry While Driving: A Test Track Study [C]. Proceedings of the 5th ITS World Congress, Seoul, Korea, 1998: 5.
- [36] D. D. Salvucci. Predicting the Effects of In-Car Interface Use on Driver Performance: An Integrated Model Approach [J]. International Journal of Human-Computer Studies. 2001, 55 (1): 85-107.
- [37] P. Green, E. Hoekstra. M. Williams. Further on-the-road Tests of Driver Interfaces: Ex-



- amination of a Route Guidance System and Car Phone [C]. Washington, D. C. : Federal Highway Administration, 1993: 1-5.
- [38] P. J. Cooper, Y. Zheng. Turning Gap Acceptance Decision-Making: Impact of Driver Distraction [J]. Journal of Safety Research. 2002, 33 (3): 321-335.
- [39] M. P. Reed, P. A. Green. Comparison of Driving Performance on-Road and in Low-Cost Simulator Using a Concurrent Telephone Dialing Task [J]. Ergonomics. 1999, 42 (8): 1015-1037.
- [40] S. T. Godley Triggs, T. J. Fildes, B. N. Fildes. Driving Simulator Validation for Speed Research [J]. Accident Analysis and Prevention. 2002, 4 (5): 589-600.
- [41] K. A. Brookhuis, G. Vries, D. Waard. The effects of mobile telephoning on driving performance [J]. Accident Analysis & Prevention. 1991, 23 (4): 309-316.
- [42] P. Green. Visual and Task Demands of Driver Information Systems [C]. Society of Automotive Engineers, Warrendale, Michigan, USA, 1999: 1-20.
- [43] J. C. McCall, M. M. Trivedi. Visual Context Capture and Analysis for Driver Attention Monitoring [C]. IEEE Conference on Intelligent Transportation Systems, Washington, D. C., United States, 2004: 332-337.
- [44] M. H. Kutila, M. Jokela, T. Mäkinen, et al. Driver Cognitive Distraction Detection: Feature Estimation and Implementation [J]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering. 2007, 221 (9): 1027-1040.
- [45] Y. Liang, J. D. Lee, M. L. Reyes. Nonintrusive Detection of Driver Cognitive Distraction in Real Time Using Bayesian Networks [J]. Transportation Research Record, Human Performance, User Information, Simulation, and Visualization. 2007, 20 (18): 1-8.
- [46] Y. Liang, M. L. Reyes, J. D. Lee. Real-time Detection of Driver Cognitive Distraction Using Support Vector Machines [J]. IEEE Transactions on Intelligent Transportation Systems. 2007, 8 (2): 340-350.
- [47] B. White, R. Ferlis. Algorithm for Predicting Inattentive Signal Violators in an Infrastructure-Based Intelligent System [C]. Transportation Research Record. 2004, 1886: 85-91.
- [48] H. Cristy. Using Spatial Warning Signals to Capture a Driver's Visual Attention [C]. Sixth International Conference on Multimodal Interfaces, State College, PA, USA, 2004: 350.
- [49] N. Edenborough, R. Hammoud, A. Harbach. Driver State Monitor from DELPHI [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, United States, 2005: 1206-1207.
- [50] 施树明, 金立生, 王荣本, 等. 基于机器视觉的驾驶员嘴部状态检测方法 [J]. 吉林大学学报 (工学版). 2004, 34 (2): 232-236.
- [51] Takahashi Kenichi, Yamada Keiichi, Nakano Tomoaki, et al. Method of Detecting Concentration on Cellular Phone Call from Facial Expression Change by Image Processing [C]. IEEE International Conference on Systems, Man and Cybernetics, Waikoloa, HI, United States, 2005: 3444-3448.



- [52] M. Yamakita, K. Takahashi, K. Yamada, et al. Measurement of Driver's Consciousness by Image Processing 2 -Detection of Concentration on Cellular Phone Call from Facial Expression Change Coping with Individual Differences [C]. IEEE International Conference on Systems, Man and Cybernetics, Taipei, 2006: 1688-1692.
- [53] F. Moreno, F. Aparicio, W. Hernandez. A Low-cost Real-Time FPGA Solution for Driver Drowsiness Detection [C]. The 29th Annual Conference of the IEEE Industrial Electronics Society, Roanoke, VA, United States, 2003: 1396-1401.
- [54] Michimasa ITOH, Yoshiyuki MIZUNO, Shin YAMAMOTO. Driver's Status Monitor [C]. Proceedings of the 21st International Conference on Data Engineering, Tokyo, Japan, 2005: 1-8.
- [55] J. Pohl, W. Birk, L. Westervall. A Driver-distraction-based Lane-keeping Assistance System [J]. Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering. 2007, 221 (4): 541-552.
- [56] J. L. Harbluk, N. Y. Ian, P. L. Trbovich, et al. An on-road Assessment of Cognitive Distraction: Impacts on Drivers' Visual Behavior and Braking Performance [J]. Accident Analysis and Prevention. 2007, 3 (2): 372-379.
- [57] T. W. Victor, J. L. Harbluk, J. A. Engström. Sensitivity of Eye-Movement Measures to In-Vehicle Task Difficulty [J]. Transportation Research Part F: Psychology Behaviour. 2005, 8 (2): 167-190.
- [58] P. K. Hughes, B. L. Cole. The Effect of Attentional Demand on Eye Movement Behavior when Driving [C]. Vision in Vehicles II. Amsterdam, Netherlands, 1988: 221-230.
- [59] H. Zhang, M. R. Smith, G. J. Witt. Identification of Real-Time Diagnostic Measures of Visual Distraction with an Automatic Eye Tracking System [J]. Human Factors. 2006, 48 (4): 805-821.
- [60] J. P. Batista. A Real-time Driver Visual Attention Monitoring System [J]. Lecture Notes in Computer Science. 2005, 3522 (1): 200-208.
- [61] P. Smith, M. Shah, N. V. Lobo. Determining Driver Visual Attention with One Camera [J]. IEEE Transactions on Intelligent Transportation Systems. 2003, 4 (4): 205-218.
- [62] J. D. Lee, J. D. Li, L. C. Liu. A Novel Driving Pattern Recognition and Status Monitoring System [J]. Lecture Notes in Computer Science. 2006, 43 (19): 504-512.

2.1 概述

视觉分散指驾驶员视线偏离车辆前方的行为，主要包括3种类型：

- 1) 视线被物体遮挡。
- 2) 驾驶员忽视道路紧急情况。例如，由于观看风景等，没有观察道路。
- 3) 选择性视觉分散，驾驶员在观察道路，但没有识别道路上的潜在风险^[1]。

本书研究的对象是第2种类型，并简称为视线偏离。如果视线偏离的持续时间较短，那么视觉分散对驾驶能力影响不大，但当持续时间超过2s时，那么驾车出事故的风险将急剧升高^[2]。

驾驶员在正常驾驶和视觉分散时会表现出不同的驾驶能力^[3]。表征驾驶员驾驶能力的工作指标主要包括：保持车辆在道路上侧向偏移的稳定性、车速保持和控制能力、对外部事件的反应时间、车距控制、工作强度、对安全驾驶参数的警惕性等^[4]。视线偏离持续时间的长短对驾驶员的影响已经得到深入研究，在偏离角度方面，主要研究了视线在垂直方向变化时对驾驶能力的影响^[5]，特别是对驾驶员反应时间的影响上^[6,7]。而侧向偏离对碰撞与安全的度量来说是更好的标准，因为即使几米的偏离也经常导致交通事故的发生^[8]。

本章中作者对同一水平面上、不同偏离角度下视觉分散的影响进行了研究，重点研究视觉分散对车辆侧向偏移的影响，并给出视觉分散检测模型。

2.2 视觉分散影响机理

如果驾驶车辆时驾驶员进行的第二作业需要视觉功能参与，将会导致视觉资源在两个任务中的竞争分配，而视觉资源是有限的，这必然导致分配于驾驶任务的视觉资源减少，形成视觉分散。根据眼睛运动特征（眼睛扫视频率、盯视平均持续时间）可以直接测量第二作业需要的视觉资源^[9]。



当驾驶员视觉注意从道路上转移开时（因为第二作业、视觉遮蔽等），驾驶员不能获得控制车辆所需的信号，在一个短的时间段内只能保持转向盘于固定的转角。在这样的时间段内，驾驶员出错的概率将会增大，导致行驶路线的不稳定甚至偏出道路。另外，视觉需求与行驶路线的保持具有很强的相关性^[10]。驾驶员保持的行驶路线是由“安全边界”指导的，驾驶员要努力保持行驶在这个边界内的区域中。

由于时间分配的不合理性导致行驶路线保持的错误，这就要求驾驶员进行手动补偿，这个动作将大于在正常行驶时的补偿动作，并更严重地分散驾驶员视觉注意。应用转向盘转向频率来研究第二作业对转向的影响，发现转向频率不是一个简单的第二作业能力需求的函数，还包括基本作业与第二作业间的相对复杂关系，以及驾驶员特征，不同作业需要的能力。并且，视觉显示的复杂性显著影响转向频率，老年驾驶员在这方面表现得更为突出。视觉负担还会导致驾驶员降低车速。这被解释为一个补偿效应，驾驶员要减少基本作业负担来保持驾驶能力在一个相对安全的水平。进行第二视觉作业时阻止驾驶员对信号、事件的识别能力。例如，Greenberg^[10]发现视觉分散将导致驾驶员对紧急情况检测率的降低。

当前，许多研究已经证明眼睛扫视与驾驶能力间具有很强的相关性^[11]。随着扫视次数与总的扫视持续时间的增加，车辆偏离行驶路线的次数也随之增加^[12]。Dingus^[13]等的研究表明扫视持续的平均时间与路线偏离距离成正相关性。虽然 Tijerina^[14]研究后给出的结论不相同，但这种相背离的结论可以通过视线平均偏离时间没有超过 2s 来解释^[15]，因为持续 2s 以下的视线偏离对驾驶员保持路线的能力没有多大影响。Zwahlen 与 DeBald^[16]的研究也表明当驾驶员闭上眼睛、视线偏离道路时，车辆偏离道路中心的标准差将比正常行驶时增大。另外，视觉分散还会增加驾驶员对前车制动信号的反应时间。驾驶员在视觉分散时释放加速踏板的时间要比正常驾驶时延长 0.4s^[17]。

除了持续时间，视线偏离的角度是描述视线偏离的另一个指标。不同偏离角度的视觉分散对驾驶员产生的影响也不同。Summala^[5]等在实验中要求驾驶员持续注视一个偏离道路的地方。要求参加实验者大声读出这三个偏离处（7°，23°，38°）的数字。虽然驾驶员还能够用眼睛的余光观察道路来控制车辆，但是随着偏离角度的增加，驾驶员路线保持的能力也跟着降低。实验应用的偏离位置主要是在垂直方向上发生变化，然而视线在垂直方向偏离与水平方向偏离时产生的影响不尽相同，因此 Lamble^[6]等研究了视线在两个方向上变化时对驾驶能力的影响，发现随着偏离角度的增加，驾驶员对前方车辆制动信号的反应时间也增加了。Horrey 与 Wickens^[7]发现视线偏离到中控台附近时，会减缓驾驶员对紧急情况反应速度。也有研究人员进行了实验研究，让参加者注视道路



中心,并且观察放置在不同位置的红色信号,并在信号发出时按下按键。实验发现红色信号放置在驾驶员正常视线范围的中心时,驾驶员的反应时间最短,如果信号放置在车辆外部,那么驾驶员对信号的反应时间将会增加14%~25%。

2.3 基于实车实验的视觉分散影响分析

研究驾驶员注意分散可以在驾驶模拟器上进行,也可以在真实车辆上进行。但模拟器与真实车辆的不同会影响测试结果,并且驾驶员在了解模拟器不能对他们造成伤害的前提下,容易做些对安全驾驶不利的事情,如双手离开转向盘、视线偏离道路场景等,也会影响实验结果。因为场景显示在屏幕上,视觉分散目标在驾驶模拟器上,两者的相互位置关系不能很好地模拟融为一体,所以会对研究不同位置视觉分散产生不利影响。

测量视觉分散对驾驶能力影响的最直接方法是在道路上采集数据,该方法要求受试人员在道路上驾驶车辆,并完成各种视觉分散的作业。该方法采集的数据来自真实环境,因此比较可靠,但是也比较耗时、昂贵。短距离测试轨迹的方法也能获得真实的数据,也经常用来测试视觉分散对驾驶能力的影响。该方法让受试者在一条测试道路或者封闭道路上驾驶车辆,通过自动仪器或者人工观察收集数据,然后与基准数据比较。此法接近真实环境,并且在封闭路上驾驶降低了危险性。

2.3.1 实验用车及场地

作者的实验选择了山东大学车辆研究所的多功能教练用车,教练车在前排乘客位置有一辅助制动装置,在出现紧急情况时前排乘客可以帮助驾驶员制动车辆以保证行车安全。我们选择一条封闭的道路作为实验场地,以避免其他车辆、行人等对实验的影响。实验车道共350m,分为两部分:实验准备部分——150m、实验数据采集部分——200m。实验场地及实验用车如图2-1a所示。

2.3.2 实验内容

要求驾驶员在实验道路中心(驾驶员正对分道线)以20km/h、10km/h的恒定速度驾驶教练用车,并在驾驶车辆期间阅读粘贴在风窗玻璃不同位置处的词语、数字。这四个位置分别位于驾驶员右侧23°、39.2°、48.36°、55.7°处,如图2-1b所示。

实验中采用的词语信息如:“地球 劳动节”、“火山 北京市”“中国世界杯”等简单词语;数字信息如:“1 2 3”、“3 5 2”、“7 8 9”等连贯或者不连贯的3位数字。采用一号黑体字将词语、数字信息打印在白色打印纸上,并裁剪至适





a) 实验用车及场地



b) 文本信息位置

图 2-1 实验用场地、车辆、文本位置

当大小，粘贴在以上四个位置。选择阅读比较简单的文本信息的原因是，阅读这样的信息需要很少认知能力，因此能较少地引起驾驶员的认知分散^[18]。根据放置位置的不同、信息类型的不同，可以把驾驶员驾驶车辆时的视觉分散分为9类，其中第9类的视线偏离角度为 0° ，代表基本驾驶（无视觉分散）。9类内容



如表 2-1 所示。

表 2-1 视觉分散类型

信息偏离角度/(°)	信息类型	
	词语	数字
55.7	1	2
48.6	3	4
39.2	5	6
23	7	8
0	9 (基本驾驶)	

2.3.3 实验过程

选择 10 位男性驾驶员,并且他们中间没有人参加过类似实验,驾龄均在 4 年左右,并且经常驾驶车辆,行驶里程均在 5000km 以上,年龄分布在 24 ~ 37 岁之间,平均年龄为 29.6 岁。

首先,让驾驶员驾驶车辆 1000m 来熟悉车辆,并告知其需要观察文本信息的位置,以及各种作业任务。在熟悉实验作业后,驾驶员进入驾驶室直到感觉正常。数据记录人员坐在前排乘客座位上,负责告知驾驶员需要阅读文本的位置,并记录数据,在有紧急情况时制动车辆。每位驾驶员需要完成 20 次实验模块,其中两次为在静止车辆上阅读不同位置的文本(词语、数字)。其他 18 次分别为在 10km/h、20km/h 的车速下完成表 2-1 中的 9 种视觉分散作业。要求驾驶员以恒定速度驾驶车辆,并建议他们保持好行驶路线。

驾驶员在实验车道始端发动车辆,在实验车道准备部分加速车辆至 10km/h 或者 20km/h,保持车速恒定并调整车辆行驶在道路中心。当车辆行驶至实验数据采集部分,提醒驾驶员进行阅读文本信息的作业。

2.3.4 结果分析

通过让驾驶员阅读数量不等的文本来模拟视觉分散,变化文本字数与内容以产生两个级别的视觉分散。另外,这些文本所处的位置不同,让驾驶员产生不同的视线偏离角度。这些阅读作业可以代表典型驾驶中能遇到的第二作业(如观察车外路标等)。

需要采集的数据主要有车辆偏离道路中心线距离与驾驶员阅读文本信息需要的时间。在车辆尾部中间悬挂一滴水装置,流速控制在 2.0L/min,以水迹中心与分道线最近边缘间的距离作为车辆偏离道路中心的距离。根据车速的不同、





文本信息的不同,测量不同距离内的车辆偏离距离,主要保证能测量到驾驶员视觉分散时间段内经过的所有路程,以及大约 1s 的驾驶员方向调整期。记录员坐在前排乘客座位,应用秒表记下驾驶员阅读文本的时间,以此作为驾驶员视线偏离道路的时间。其中时间记录起始点为记录员提醒驾驶员阅读文字的时间,终点为驾驶员阅读文本完毕的时间。因为车速对驾驶员阅读文本速度的影响不大,我们仅统计 20km/h 车速下视线偏离的平均时间,如图 2-2 所示。

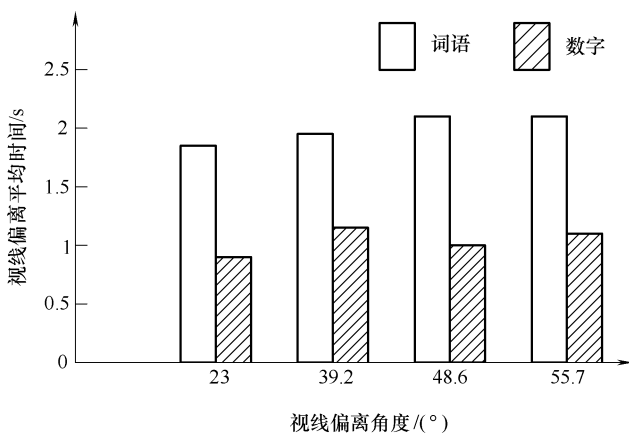


图 2-2 视线偏离道路平均时间

根据图 2-2 可以看出,阅读不同文本需要的时间是不同的,阅读词语要比阅读数字多用接近 1s 的时间,这其中的原因主要是词语的单字数(5 位)要比数字(3 位)多 60%。随着文本位置偏离角度的增加,视线偏离持续的时间也有所增加,从 23°到 55.7°,视线偏离平均时间大约增加 20%。时间增加的主要原因是驾驶员头部转动较大角度时需要的时间比较长,另外,驾驶员搜索较远处的目标也比较耗时。

计算驾驶员在不同视觉分散情况下路线偏离的标准差(SDLP- Stand Difference of Lane Position)如图 2-3 所示。

从图 2-3 中可以看出,20km/h 车速下的 SDLP 要比 10km/h 下的 SDLP 大 1.5 ~ 3cm。SDLP 的变化主要是由于驾驶员保持方向不稳引起的,因此车速高时 SDLP 变化要剧烈一些。SDLP 在基本驾驶时大约是 1cm,而在视觉分散时最多超过 4cm,这说明层次较深的视觉分散对驾驶员的影响还是比较大的。SDLP 随着视线偏离角度的增大而增大,并且阅读词语要比数字时大,这在车速高、视线偏离角度大时更明显。另外,我们还可以看出,虽然 SDLP 随着驾驶员视线偏离角度增加时变化明显,但是 SDLP 都比较小(最大不超过 5cm)。这可能有两方面的原因造成这个现象:

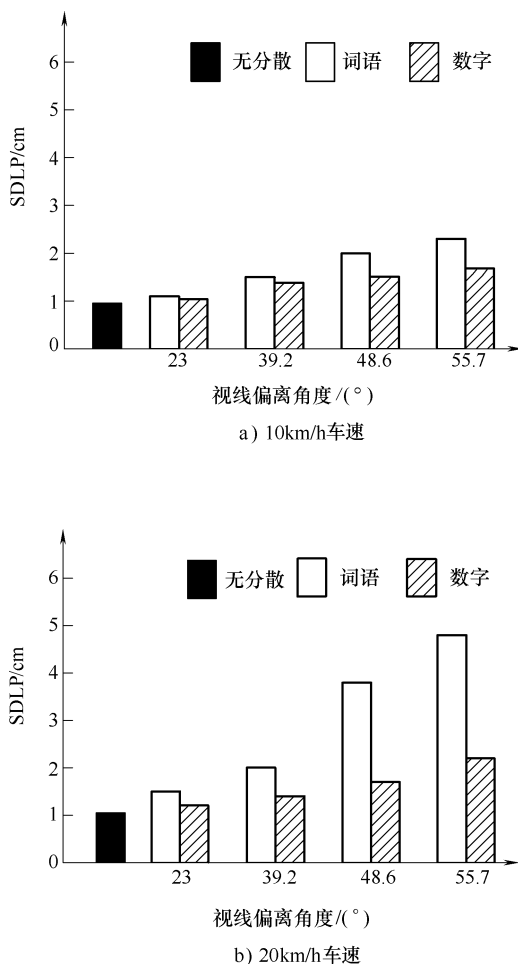


图 2-3 车辆偏离车道中心距离标准差 (SDLP)

1) 驾驶员被告知的任务就是保持行驶路线的稳定, 因此驾驶过程中驾驶员刻意保持了行驶路线的稳定性。

2) 为保证安全, 实验中采取了较低的车速, 车辆在横向上的速度分量也较小, 因此车辆不易偏离道路中心。

在这里应用的变量主要有视线偏离时间、视线偏离角度、SDLP。表征驾驶能力最直接的变量是 SDLP, 如果其他变量也能揭示驾驶能力变化, 那么它应该是与 SDLP 是共变的。在两种车速、两种视觉分散类型下, 视线偏离角度与 SDLP 的皮尔逊相关系数在 0.785 ~ 0.814 之间 ($p < 0.05$), 预示着二者之间具有很强相关性。

随着文本数目的增加, 或者文本偏离角度的增加, SDLP 计算出来的值是基



准驾驶时的 3~4 倍。这说明驾驶员在文本信息变得复杂,或者偏离率增加时会受到更大影响。因为这些驾驶能力变量与交通安全在所有驾驶情况下都有关,这就意味着减少视觉分散将会减少交通事故是合情理的。因为及时地对道路事件做出反应,对高速公路的交通来说至关重要,当前研究表明应该对视觉分散的驾驶员进行警告,使驾驶员避免车辆前方的威胁因素。

2.4 驾驶员视觉分散检测模型

2.4.1 驾驶员注意特征评价

注意是一切认知过程的基础,是心理活动对一定事物的指向和集中^[19]。

首先,人们的注意行为与大脑是密不可分的。大量研究表明,额叶执行注意控制的储存和执行功能^[20]。人们的注意水平制约于大脑组织,通过检测脑电波可以判断人们的注意水平^[21]。Huang^[22]等进行的实验中根据 α 波平均值把脑电波分为 5 个等级,通过对照这 5 个等级与当事人的实际情况,发现 α 脑电波平均值能很好地反映注意水平高低。因此脑电分析是一种直接的注意相关分析方法。但是脑电信号很微弱,需要特殊的采集设备,限制了这种分析方法在许多场合的应用。另外,脑电波的影响因素还很多,例如情绪、疲劳等,这会给基于脑电的注意水平分析带来一定不利因素。

再者,皮肤电导率也是一个与人的注意水平直接相关的生理指标^[23]。皮肤电导率,也就是皮肤电反应,是指当人受到内部或外部刺激而觉醒水平升高时,皮肤变为电的良导体的现象。强烈的情绪、惊讶的事情、艰难的作业、深呼吸、训练、疼痛等都会导致皮肤导电率的提升,因此只有在限定很多条件的情况下,才能区分出到底是什么导致了皮肤电导率的升高。由此看来,虽然注意水平变化是引起皮肤电导率反应的一个直接因素,但在很多场景下不能应用皮肤电导率反应来分析人的注意水平。例如,在驾驶员注意水平检测上就有些困难,因为这时驾驶员不但要注意集中,还要有一定的加速、制动、转向等动作,而这些动作也会引起皮肤电导率的提升。

最后,注意与人们的行为之间也存在一个深层次的关系。人们普遍认为,注意转移和眼动是密切相关的。实验表明眼动和注意转移通常是趋于同一位置和方向的^[24]。眼动是给出人类行为的重要线索。一些人类行为因素的专家认为,对驾驶作业来说周围视觉与中央凹视觉一样重要。Recarte 与 Nunes^[25]研究了眼睛盯视的焦点、持续期、序列间的关系,并着眼于精神作业、视觉场景、口头作业如何影响驾驶员。其重点放在第二作业和正常驾驶时视线的不同上。结果显示,处理视觉场景时要长序列和短时间的盯视,在处理口头作业时需要短时



间的盯视。

当人们注视的物体在人脸前方偏转 20° 以上时,为跟踪该目标人们不仅要转动眼睛,还需要转动头部:头—眼系统的这种组合运动看起来是要保证眼睛在眼眶的中心位置^[26]。因此,考虑头部运动来判断注意方向也是合理的。Pelz^[27]的实验展示出头部方向是揭示注意视野的最基本线索,这个视野包括注意的实际焦点。

这样判断一个人的注意方向不仅需要眼睛信息,还需要头部转动角度信息。早在 1824 年, Wollaston 就发现判断一个人的注视方向不单单需要虹膜、瞳孔与眼白的相对位置,还需要头部姿势的信息^[28]。有研究者的实验也发现,决定一个人的注意方向同样还需要一些不同的线索,但是他们认为头部姿势与眼睛对判断注视方向一样重要^[29]。

综合起来看,根据驾驶员的视线方向检测驾驶员是否发生视觉分散是比较合适的方法。另外, Zhang^[30]等的研究也表明驾驶员的视线特点与驾驶员的驾驶能力、反应时间之间有很强的相关性,可以作为检测驾驶员视觉分散的指标。因此,我们根据驾驶员的视线特征来检测驾驶员是否发生视觉分散。

2.4.2 驾驶员视线特点

大量实验表明,驾驶员在正常行驶时,其视线方向的角度有 $80\% \sim 90\%$ 时间是在俯仰角与横摆角的 $\pm 12^\circ$ 范围内变化^[30],如图 2-4 所示。另外,驾驶员发生视觉分散时,驾驶员视线只有不到 50% 的时间在这个范围内。因此,如果驾驶员的视线方向超出这个范围,则可以认为驾驶员视线偏离前方,出现视觉分散^[31]。

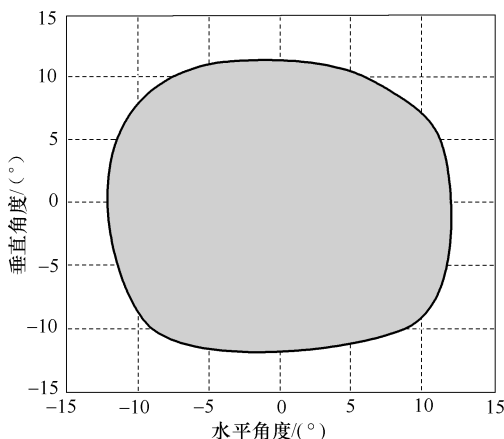


图 2-4 驾驶员视线角度变化范围



并不是所有情况下驾驶员的视线偏离都会造成危险。如果偏离时间小于 1s, 那么对驾驶能力影响不大。另外, 低频率的视线偏离也不会给驾驶能力造成严重影响。只是持续时间较长、频率高的视线偏离才会影响驾驶员的驾驶能力。为判断驾驶员是否视觉分散, 就需要确定持续时间及频率的阈值, 持续时间或者频率高于相应阈值的情况被认为是不可接受的。因为确定此阈值的复杂性较高, 不同研究者建立的阈值也不尽相同。

1987 年 Zwahlen^[8]等认为: 视线偏离 4 次, 每次不超过 2s 是可行的, 但是连续 4 次 2s 的视线偏离是不可接受的。较短时间 (例如 4s) 内的 4 次连续扫视也是危险的, 因为在大脑中存储的道路信息的总量较低。Blanco^[32]定义了两种阈值: “红线”标准, 视线偏离严重影响驾驶能力; “黄线”标准, 驾驶能力发生明显变化。Blanco 提出的红线标准是视线偏离的次数大于 9 次, 或者持续时间大于 2.5s 的视线偏离。基于 85 个百分点的准则, 假设手动调节收音机被认为是可接受的视觉分散, 汽车制造协会制定的标准是持续时间 2s 或者偏离次数 10 次。

驾驶员视线偏离的持续时间与车辆发生事故的风险存在一定的正相关性, 如果视线偏离前方的持续时间不超过 2s, 那么危险系数升高很小; 但当视线偏离前方道路的持续时间大于 2s, 那么危险系数将升高很快^[33], 不管是出于什么目的造成了视线偏离。

总体来看, 检测视觉分散的阈值一般有两个方面:

- 1) 单次视线偏离持续的时间。
- 2) 一段时间 (例如 20s) 内视线偏离的次数。

若以视线偏离的次数为阈值, 则需要较长的时间窗口, 而交通事故发生在很短的时间内, 因此通过检测驾驶员视线偏离次数来判断视觉分散, 对预防交通事故的作用不大。本书仅涉及持续时间上的阈值。根据其他研究者确定的阈值, 以及我们自己的实验结果, 确定的阈值为 2s, 即如果视线偏离道路的持续时间超过 2s, 则认为是不可接受的视觉分散, 在本书中简称为视觉分散。

2.4.3 视觉分散检测模型

视线角度包括面部姿势与眼睛视线角度 (本书中不加说明的视线均指二者的合成)。在驾驶环境中, 面部姿势对预测驾驶员的视线方向是很有益的, 如驾驶员在看后视镜或窗外风景时都会伴有特定的面部姿势^[34], 根据驾驶员的面部姿势能判断驾驶员有没有处于视觉分散的状态^[35]。当驾驶员的面部朝向的角度偏离正常视线范围时, 就可以判断驾驶员视线角度偏离正常行驶方向, 因为不会有驾驶员在偏转面部时还会注视车辆前方。但是, 面部姿势在正常视线范围之内时, 驾驶员的眼睛视线也可能偏离车辆前方, 这时就需要判断眼睛视线



方向。

驾驶员在变换车道、转向时，需要转移视线来观察对应道路，特别是在十字路口转向时，驾驶员视线偏离车辆前方的持续时间会超过 2s，这也被认为是一种视觉分散行为。但这并非是一个危险情况，所以需要在视觉分散检测模型中增加一个驾驶员转向行为识别的模块，以降低视觉分散检测的虚警率。

因此，我们确定一个基于驾驶员面部姿势与驾驶员眼睛视线的检测，并包括转向行为识别的视觉分散检测模型，如图 2-5 所示。

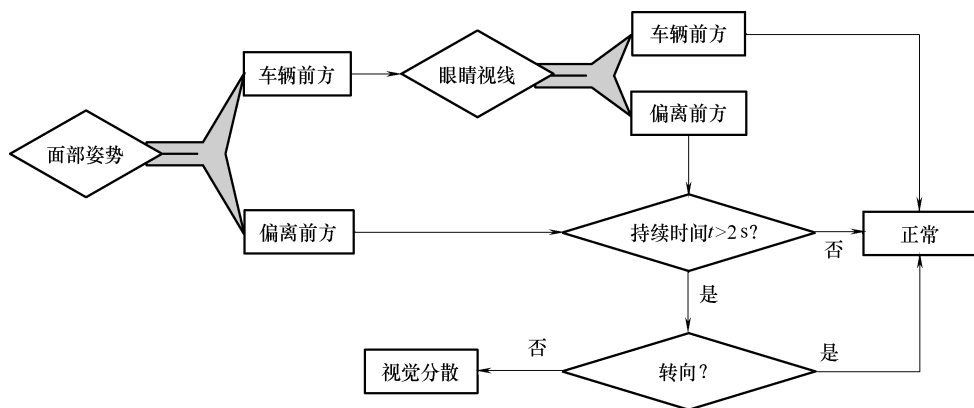


图 2-5 驾驶员视觉分散检测模型

2.5 本章小结

本章研究了视觉分散对驾驶员的影响机理，并设计实验研究视觉分散对驾驶能力的影响。让驾驶员阅读数量不等的文本来模拟视觉分散。实验结果表明，在同一水平面上，随着文本位置偏离角度的增加，视觉分散对驾驶能力的影响也增大。通过分析各种注意水平评价方法，发现根据视线特征可以实现无侵入性检测驾驶员视觉分散。作者根据驾驶过程中驾驶员视线变化的特点，建立了基于驾驶员面部姿势与视线方向识别，并包含转向行为识别的视觉分散检测模型。

参考文献

- [1] M. Pettitt, G. Burnett, A. Stevens. Defining Driver Distraction [C]. World Congress on Intelligent Transport Systems, San Francisco, 2005: 1-12.
- [2] S. G. Klauer, T. A. Dingus, V. L. Neale. The Impact of Driver Inattention on Near-



- Crash/Crash [C]. Virginia: National Highway Traffic Safety Administration, 2006: 1-10.
- [3] P. Green. Visual and Task Demands of Driver Information Systems [C]. Society of Automotive Engineers, Warrendale, Michigan, USA, 1999: 1-20.
- [4] J. L. Harbluk, N. Y. Ian, P. L. Trbovich, et al. An on-road Assessment of Cognitive Distraction: Impacts on Drivers' Visual Behavior and Braking Performance [J]. Accident Analysis and Prevention. 2007, 3 (2): 372-379.
- [5] H. Summala, T. Nieminen, M. Puntio. Maintaining Lane Position with Peripheral Vision during In-Vehicle Tasks [J]. Human Factors. 1996, 38 (3): 442-451.
- [6] D. Lamble, M. Laakso, H. Summala. Detection Thresholds in Car Following Situations and Peripheral Vision: Implications for Positioning of Visually Demanding In-Car Displays [J]. Ergonomics. 1999, 42 (6): 807-815.
- [7] W. J. Horrey, C. D. Wickens. Driving and Side Task Performance: The Effects of Display Clutter, Separation, and Modality [J]. Human Factors. 2004, 46 (4): 611-624.
- [8] H. T. Zwahlen, D. P. DeBald. Safety Aspects of CRT Touch Panel Controls in Automobiles [C]. Proceedings of the 16th International Symposium on Automotive Technology and Automation, Florence, Italy, 1987: 193-212.
- [9] J. Engström, E. Johansson, J. Östlund. Effects of Visual and Cognitive Load in Real and Simulated Motorway Driving [J]. Transportation Research Part F: Traffic Psychology and Behaviour. 2005, 8 (2): 97-120.
- [10] J. Greenberg, L. Tijerina, R. Curry. Evaluation of Driver Distraction Using an Event Detection Paradigm [J]. Journal of the Transportation Research Board. 2003, 1843 (1): 1-9.
- [11] H. Zhang, M. R. Smith, G. J. Witt. Identification of Real-Time Diagnostic Measures of Visual Distraction with an Automatic Eye Tracking System [J]. Human Factors. 2006, 48 (4): 805-821.
- [12] L. Tijerina, S. Kiger, T. Rockwell, et al. Heavy Vehicle Driver Workload Assessment Task 7 A: In-Cab Test Message System and Cellular Phone Use by Heavy Vehicle Drivers on the Road [C]. Washington, D. C.: Office of Crash Avoidance Research National Highway Traffic Safety Administration, 1996: 1-5.
- [13] T. A. Dingus, M. C. Hulse, D. V. McGehee, Etc. Driver Performance Results From the Travtek IVHS Camera Car Evaluation Study [C]. Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting, Santa Monica, CA, 1994: 1118-1122.
- [14] L. Tijerina, E. B. Parmer, M. J. Goodman. Individual Differences and In-Vehicle Distraction While Driving: A Test Track Study and Psychometric Evaluation [C]. Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, Santa Monica, CA, 1999: 982-986.
- [15] T. H. Rockwell. Spare Visual Capacity in Driving Revisited: New Empirical Results for an Old Idea [C]. Vision in Vehicles. Netherlands, Amsterdam, 1988: 317-324.
- [16] H. T. Zwahlen, D. P. DeBald. Safety Aspects of Sophisticated In-Vehicle Information Dis-



- plays and Controls [C]. Proceedings of the Human Factors Society 30th Annual Meeting, Santa Monica, CA, 1986: 256-260.
- [17] J. D. Lee, D. V. McGehee, T. L. Brown, et al. Collision Warning Timing, Driver Distraction, and Driver Response to Imminent Rear-End Collisions in a High-Fidelity Driving Simulator [J]. Human Factors. 2002, 44 (1): 314-334.
- [18] C. M. MacLeod. Half a Century of Research on the Stroop Effect: An Integrative Review [J]. Psychological Bulletin. 1991, 109 (1): 163-203.
- [19] 罗跃嘉, 魏景汉. 注意的认知神经科学研究 [M]. 北京: 高等教育出版社, 2004: 1-34.
- [20] 朱镛连. 认知功能的药物康复 [J]. 中国康复理论与实践. 2007, 13 (1): 10-15.
- [21] M. Pilu. On the Use of Attention Clues for an Autonomous Wearable Camera [C]. Bristol: Hardcopy Technology Laboratory HP Laboratories, 2003.
- [22] Haung Wei Ng, Yasuhito Sawahata, Kiyoharu Aizawa. Summarization of Wearable Videos Using Support Vector Machine [C]. IEEE International Conference on Multimedia and Expo, Lausanne, Switzerland, 2002: 325-328.
- [23] A. Lockerd, F. Mueller. LAFCam: Leveraging Affective Feedback Camcorder [C]. Conference on Human Factors in Computing Systems, Cambridge USA, 2002: 574-575.
- [24] 马勇. 基于眼动分析的汽车驾驶员视觉搜索模式研究 [D]. 西安: 长安大学, 2006: 10-20.
- [25] M. A. Recarte, L. M. Nunes. Mental Workload While Driving: Effects on Visual Search, Discrimination, and Decision Making [J]. Journal of Experimental Psychology: Applied. 2003, 9 (2): 119-137.
- [26] M. Argyle, M. Cook. Gaze and Mutual Gaze [M]. Cambridge: Cambridge University Press, 1996.
- [27] J. B. Pelz. Visual Representations in a Natural Visuo-motor [D]. New York University. 1995: 15-37.
- [28] W. H. Wollaston. On the Apparent Direction of Eyes in a Portrait [J]. Philos. Trans. R. Soc. London Ser. B. 1824, 11 (4): 247-256.
- [29] S. R. H. Langton. The Mutual Influence of Gaze and Head Orientation in the Analysis of Social Attention Direction [J]. The Quarterly Journal of Experimental Psychology A. 2000, 53 (3): 825-845.
- [30] P. K. Hughes, B. L. Cole. The Effect of Attentional Demand on Eye Movement Behavior when Driving [C]. Vision in Vehicles II. Amsterdam, Netherlands, 1988: 221-230.
- [31] H. Zhang, M. R. Smith, G. J. Witt. Identification of Real-Time Diagnostic Measures of Visual Distraction with an Automatic Eye Tracking System [J]. Human Factors. 2006, 48 (4): 805-821.
- [32] M. Blanco. Effects of In-Vehicle Information Systems (IVIS) Tasks on the Information Processing Demands of a Commercial Vehicle Operations (CVO) Driver [D]. Virginia Polytech-





- nic Institute and State University. 1999: 41-80.
- [33] S. G. Klauer, T. A. Dingus, V. L. Neale. The Impact of Driver Inattention on Near-Crash/Crash [C]. Virginia: National Highway Traffic Safety Administration, 2006: 1-10.
- [34] P. A. Beardsley. A Qualitative Approach to Classifying Head and Eye Pose [C]. the Fourth IEEE Workshop on Applications of Computer Vision, Princeton, 1998: 208-213.
- [35] Toshiki Matsui, Naoki Suganuma, Naofumi Fujiwara. Driver's Head Pose Measurement and Corner Center Detection [C]. SICE- ICASE International Joint Conference, Busan, Korea, 2006: 2834-2839.

3.1 概述

定位人脸区域是估计驾驶员面部姿势的前提。另外，在人脸定位的基础上还可以在一个相对小的范围内，更容易地定位眼睛等人脸特征点。因此，本章首先对驾驶员面部定位问题进行研究。基于人脸定位在各种应用中的重要性，已有众多研究者开发了各种各样的算法，主要有：基于知识的方法、基于不变特征的方法、基于外部特征的方法、基于模板匹配的方法 4 种^[1]。虽然人们对该问题进行了深入研究，但是人脸姿势、表情等还是困扰人脸定位效果的难题。为解决人脸姿势对人脸定位的影响，许多学者提出各种解决办法，如 Lee 提出基于人脸特征点定位的方法，这种方法可以定位小角度范围内变化的人脸，当人脸姿势变化范围较大时，由于特征点的丢失使该方法失效^[2]。Chen 等提出基于 Harr-like 特征（局部）与 Gabor 特征（全局）相结合的方法定位多姿势人脸^[3]。因为人脸肤色在多角度下都具有相似的聚类性^[4]，可以先根据人脸肤色的特征进行人脸区域大致定位，然后根据人脸轮廓采用神经网络、模板匹配、三角结构等方法进行人脸精确定位^[5]。

分析各种人脸定位方法可以看出，利用肤色特征可以在多种姿势下检测出驾驶员的面部。基于肤色的人脸定位的难点是肤色检测效果容易受光照变化的影响，为提高不同光照下人脸检测的鲁棒性，需要对采集的图像进行一定的光照补偿，并根据经验公式对变换后的 YCbCr 值进行非线性校正。

基于肤色的人脸定位算法一般会把人的额头、颈部同时提取出来，而这两部分并不利于检测驾驶员面部姿势。因为头发的遮挡，会使相同姿势下的人脸额头具有不同的外观特征，而不同面部姿势下的颈部区域并没有多大变化。而嘴唇、眼睛等脸部特征是判断人脸姿势的显著特征^[4]。因此，分析认为利用精确定位眉毛、嘴唇区域的方法可以限定人脸的上下边界，以分割额头、颈部区域。由于人脸检测应用场合的不同，在精确定位人脸时大多没有包含人耳。而人的耳朵在估计人脸姿势时是一个很有用的线索，因为根据平面图像中耳朵的数量、



位置、变形等情况，可以判断人脸姿势的大体角度（正面、偏左、偏右等）。

另外，根据本研究的实际情况，可归纳出驾驶员视频图像具有以下特点：

- 1) 图像中最多只有一个人脸。
- 2) 图像易受侧光影响，导致光照分布不均匀。
- 3) 眼睛区域不会受强光照射的影响。
- 4) 驾驶员头部会轻微运动。

综合以上各方面考虑，本章提出基于头部运动信息与颜色信息相结合的方法来检测驾驶员面部位置。首先，利用运动信息及肤色模型进行人脸预定位，再根据耳朵、眉毛、嘴唇的特征进行人脸精确定位的方法，并分割面部图像作为估计驾驶员面部姿势的样本，最后根据需把分割的面部图像进行大小归一化。

3.2 基于运动信息的头部区域检测

3.2.1 运动目标检测方法

由于驾驶员在驾驶车辆时，头部是不断运动的，而在驾驶员视频中的其他物体是相对不动的，因此可以利用头部运动信息来检测头部位置。运动目标检测按背景变化与否可分为摄像机运动、静止两种模式，按技术划分可分为光流法和图像差分法^[6]。

光流法根据运动目标随时间变化的光流特性，可以有效地提取和跟踪运动目标^[7]，但计算相当复杂，对噪声敏感，且对硬件要求也比较高，因此并不适合在实际中应用。

图像差分法是检测复杂背景中运动目标的一种常用方法，主要包括：背景差分法和帧间差分法。

背景差分法是将两幅图像相减结果中的像素值与阈值进行比较，如果像素值大于阈值，则认为是前景点，否则为背景点。该方法计算速度快，可以做到实时检测，不足之处是受环境光线变化影响较大，背景更新较为复杂。算法流程图如图 3-1 所示。

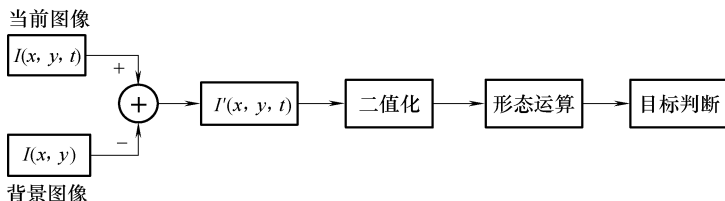


图 3-1 背景差分法算法流程图



帧间差分法是将视频图像序列中的相邻两帧或三帧图像相减，目的是获取图像中运动像素点的时间梯度信息，以此来提取图像中的运动区域。算法流程图如图 3-2 所示：

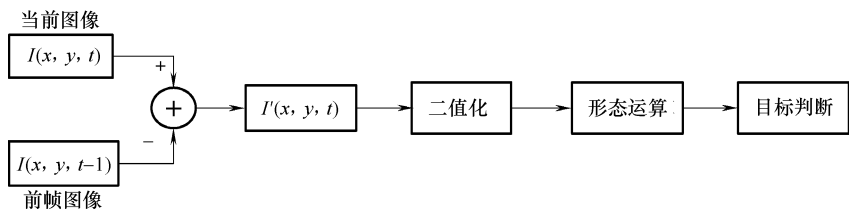


图 3-2 帧间差分法算法流程图

简单图像差分算法主要存在以下缺陷^[8]：

- 1) 当两帧图像对应坐标稍有偏差时，即使两帧图像非常相似，其差分后的图像灰度值也不会趋于零。
- 2) 原始灰度图像可能受到噪声干扰，这会影响差分图像的后续处理，可以将连续多帧差分图像进行与运算，从而有效地抑制噪声；也可以对差分图像进行中值滤波或形态学滤波，图像差分的处理过程一般如图 3-3 所示。

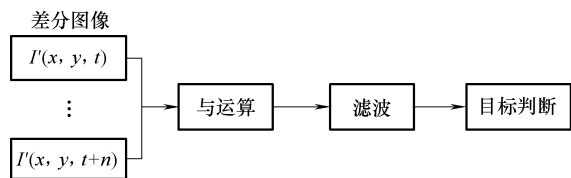


图 3-3 图像差分处理流程图

3.2.2 基于帧间差分法的驾驶员头部检测

驾驶员在驾驶过程中，由于眼睛不仅需要注视车道前方，还需观察车道左右侧及后视镜等，这就需要头部时刻保持轻微的转动或摇动，基于此特点可以采用多帧图像差分的方法进行驾驶员头部区域检测。具体算法如下。

- 1) 将彩色视频转换为灰度视频后，取长度为 n 的视频段，本书中 $n=6$ ，视频序列见图 3-4，将图 3-4 进行相邻帧间差分，得到 $5(n-1)$ 帧差分图像，并将差分图像进行归一化处理，使其像素在 $0 \sim 255$ 范围内，如图 3-5 所示，可以看出由于驾驶员头部的微动作，差分图像中出现了驾驶员的头部轮廓。
- 2) 对差分后图像按照式 (3-1) 进行滤波处理，滤波目的为去掉像素值变化较小的像素点，滤波图像如图 3-6 所示。





图 3-4 原始视频序列



图 3-5 差分图像



图 3-6 滤波图像



$$I'_c(x,y) = \begin{cases} I_c(x,y) & \text{if } I_c(x,y) \geq 50 \\ 0 & \text{else} \end{cases} \quad (3-1)$$

式中 I'_c ——滤波后差分图像；

I_c ——滤波前差分图像。

3) 将滤波后的图像进行或运算，使其像素值在 0 ~ 255 范围内分布，如图 3-7 所示。或运算的目的是尽可能得到较为全面的目标轮廓（面部轮廓）图像。

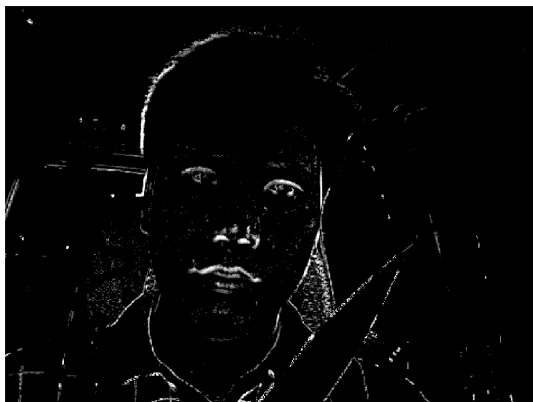


图 3-7 或运算图像

4) 对或运算图像进行数学形态运算，以去掉孤立点或填充空洞；然后分别在水平、垂直方向上进行积分投影，投影公式见式 (3-2)、式 (3-3) 所示。投影曲线如图 3-8、图 3-9 所示，通过分析投影曲线即可定位出头部位置，如图 3-10 所示。

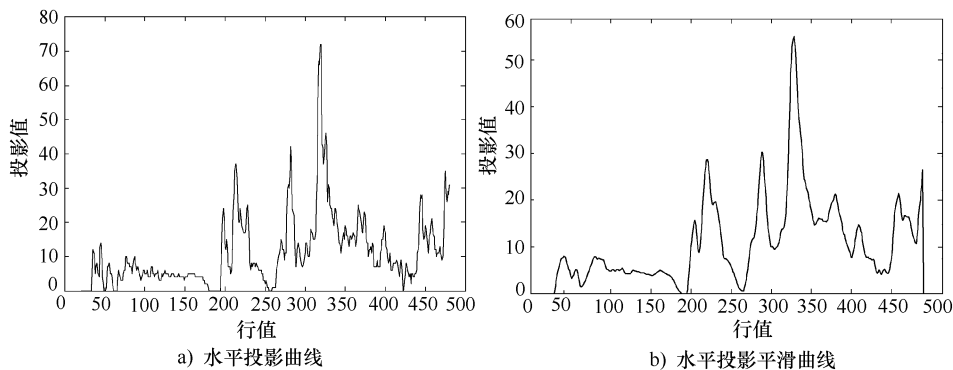


图 3-8 水平投影曲线

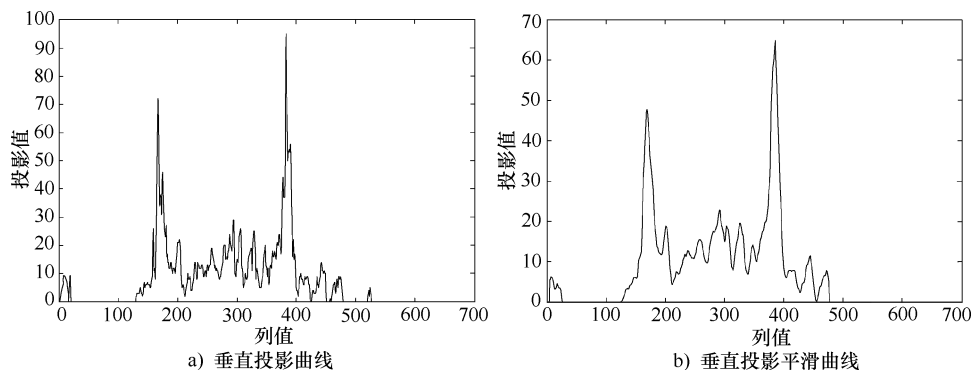


图 3-9 垂直投影曲线

水平积分投影

$$H(i) = \sum_{j=1}^n I(i, j) \quad (3-2)$$

式中 H ——水平积分投影函数；

I ——投影图像；

i, j ——图像像素点坐标, $1 \leq i \leq m, 1 \leq j \leq n$;

m ——图像行数；

n ——图像列数。

垂直积分投影

$$V(j) = \sum_{i=1}^m I(i, j) \quad (3-3)$$

式中 V ——垂直积分投影函数；

I ——投影图像；

i, j ——图像像素点坐标, $1 \leq i \leq m, 1 \leq j \leq n$;

m ——图像行数；

n ——图像列数。



图 3-10 头部定位效果

投影曲线存在众多毛刺, 不利于曲线处理, 需进行曲线平滑处理, 平滑公式如下列公式所示

$$H'(i) = \frac{\sum_{i_1=i-3}^{i_1=i+3} H(i_1)}{7} \quad (3-4)$$

$$V'(j) = \frac{\sum_{j_1=j-3}^{j_1=j+3} V(j_1)}{7} \quad (3-5)$$



式中 $4 \leq i \leq m-3$, $4 \leq j \leq n-3$ 。

基于差分算法的头部检测定位流程图如图 3-11 所示。

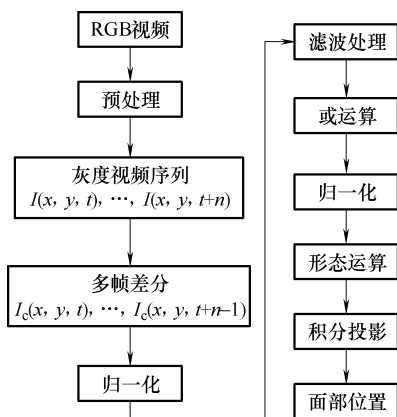


图 3-11 基于差分的头部定位流程图

基于帧间差分的头部检测算法计算量小，可以达到实时运算的要求。然而，由于在驾驶室内可能存在其他运动物体的干扰，使得该方法的检测精度不是很高，经常会出现检测失败的情况。为此，该方法需与其他的检测方法相结合使用，若通过差分法不能检测到人脸，则接着启用第二种检测手段。本书采用肤色检测法检测驾驶员的面部。

3.3 建立肤色混合高斯模型

3.3.1 彩色空间选择

在计算机中应用的彩色空间有 RGB、CMYK、YCrCb、YUV、YIQ、HSI 等，其中 RGB 是最基本的彩色系统，对应于人类视觉的三基色即红、绿、蓝三种颜色，该空间由 R（红），G（绿）和蓝（B）三个分量组成，分别与三维空间中的三个轴对应。原点对应于黑色，离原点最远的顶点对应于白色，其他颜色落于由红、绿和蓝三基色组成的彩色立方体中。

不同彩色空间下的人脸检测有所不同，检测效果也不尽相同。Benjamin 等^[9]发现在 HIS 和 YCrCb 空间上的实验效果要相对理想。而 YCrCb 空间由于转换方法简单，并且能够实现亮度和色度分量比较彻底的分离^[10]，因此我们可以选择在 YCrCb 彩色空间建立肤色模型。其中 Y 指亮度，通常把 Cb，Cr 叫做色度，因而实现了亮度和色度的分离，并且在 YCrCb 彩色空间中肤色的聚类特性



比较好。

RGB 和 YCrCb 空间的转换如式 (3-6) 所示

$$\begin{pmatrix} Y \\ C_b \\ C_r \\ 1 \end{pmatrix} = \begin{pmatrix} 0.2990 & 0.5870 & 0.1140 & 0 \\ -0.1687 & -0.3313 & 0.5000 & 128 \\ 0.5000 & -0.4187 & -0.0813 & 128 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \\ 1 \end{pmatrix} \quad (3-6)$$

3.3.2 YCbCr 空间下的肤色混合高斯模型

肤色模型，用一种代数（解析）或查表的形式来表达哪些像素的色彩属于肤色，或者表征出某一像素的色彩与肤色的相似程度。肤色模型是与彩色空间相联系的，也就是说肤色模型的建立需要结合不同的彩色空间，下面在 YCrCb 彩色空间中建立肤色模型。人脸肤色建模的方法很多，如高斯模型、神经网络模型、几何模型等，并各自具有优缺点^[11]。几何模型与神经网络模型本质上属于二值化模型。对于高斯模型，它是通过计算像素的概率值构成连续数据信息，并由此得到一个肤色概率图，然后根据肤色概率图上的不同数值来完成肤色像素的定位，是一个适合于本书所需人脸定位的肤色模型。

不同人的皮肤颜色在 YCrCb 彩色空间中色度信息 C_b 、 C_r 具有相似的分布特性，并且人脸皮肤的颜色信息 C_r 、 C_b 组成矢量 $X = (C_r, C_b)^T$ 的分布基本符合高斯分布的规律^[12]，如图 3-12 a 所示。

该肤色分布模型的概率密度函数可以表示为

$$f(X) = \frac{1}{\sqrt{2\pi} |C|^{1/2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T C^{-1} (X - \mu) \right\} \quad (3-7)$$

式中 $\mu = (\bar{C}_r, \bar{C}_b)^T$ 。

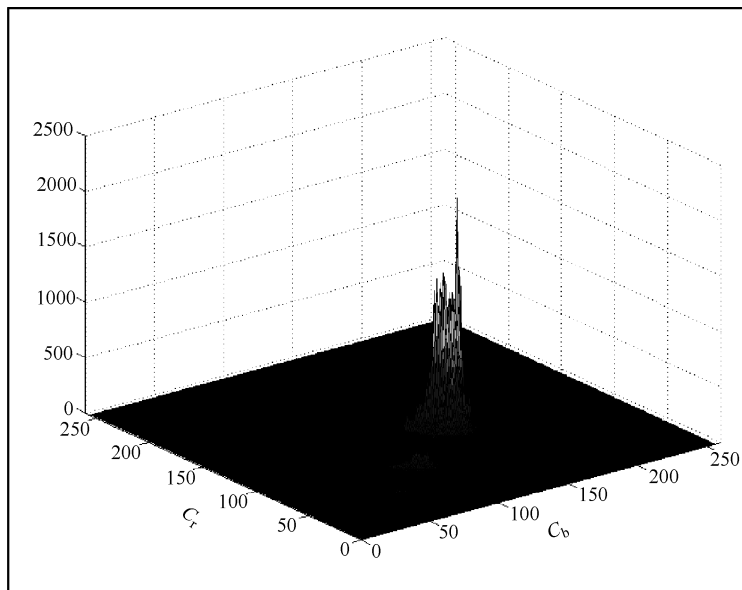
该向量中两个值分别指的是颜色分量 C_r 、 C_b 的均值， C 是 C_b 、 C_r 的协方差矩阵

$$C = \begin{bmatrix} \sigma_{rr}^2 & \sigma_{rb} \\ \sigma_{br} & \sigma_{bb}^2 \end{bmatrix} \quad (3-8)$$

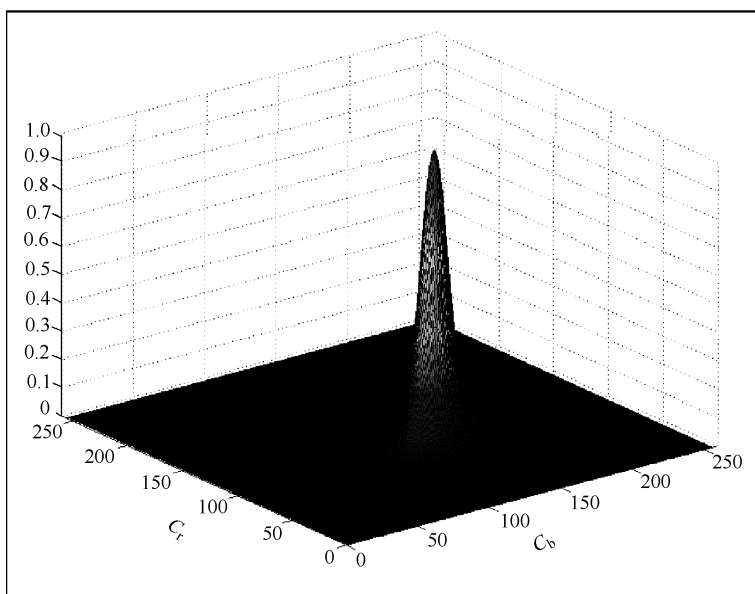
式中 σ_{rr}^2 、 σ_{bb}^2 ——分别是 C_r 、 C_b 的方差；

σ_{rb} 、 σ_{br} ——分别是 C_r 、 C_b 的协方差。

然而，人脸部位在高光、阴影的影响下，人脸肤色模型不再是单一的高斯模型，肤色模型可以表示为 4 个高斯函数组成的简化混合高斯模型^[13,14]



a) 人脸肤色 C_b 、 C_r 分布图



b) 人脸肤色高斯模型

图 3-12 YCrCb 空间中人脸肤色分布



$$\begin{cases} f_1(x) = 0.0053 \exp\left(-\frac{1}{2}\left(X - \begin{bmatrix} 90.3 \\ 169.2 \end{bmatrix}\right)^T \begin{bmatrix} 25.8 & 0 \\ 0 & 36.5 \end{bmatrix}^{-1} \left(X - \begin{bmatrix} 90.3 \\ 169.2 \end{bmatrix}\right)\right) \\ f_2(x) = 0.0037 \exp\left(-\frac{1}{2}\left(X - \begin{bmatrix} 100.2 \\ 160.5 \end{bmatrix}\right)^T \begin{bmatrix} 49.8 & 0 \\ 0 & 37.3 \end{bmatrix}^{-1} \left(X - \begin{bmatrix} 100.2 \\ 160.5 \end{bmatrix}\right)\right) \\ f_3(x) = 0.0025 \exp\left(-\frac{1}{2}\left(X - \begin{bmatrix} 110.9 \\ 145.1 \end{bmatrix}\right)^T \begin{bmatrix} 79.2 & 0 \\ 0 & 81.6 \end{bmatrix}^{-1} \left(X - \begin{bmatrix} 110.9 \\ 145.1 \end{bmatrix}\right)\right) \\ f_4(x) = 0.0028 \exp\left(-\frac{1}{2}\left(X - \begin{bmatrix} 125.1 \\ 135.3 \end{bmatrix}\right)^T \begin{bmatrix} 64.0 & 0 \\ 0 & 50.6 \end{bmatrix}^{-1} \left(X - \begin{bmatrix} 125.1 \\ 135.3 \end{bmatrix}\right)\right) \end{cases} \quad (3-9)$$

式中 $X = [C_b \ C_r]$;

x ——像素点。

则某像素属于该混合高斯模型的概率为

$$f(x) = 0.16f_1(x) + 0.21f_2(x) + 0.40f_3(x) + 0.23f_4(x) \quad (3-10)$$

根据 $C_b C_r$ 组成的向量即可计算该像素属于皮肤的概率。

3.4 驾驶员面部预定位

基于肤色定位人脸的缺点是定位效果容易受光照变化的影响,为提高不同光照下人脸定位的鲁棒性,可以对采集的图像进行一定光照补偿,并根据经验公式对变换后的 YCrCb 值进行非线性校正。

光照对图像的影响主要分为图像过亮和图像过暗。光照太强,可导致 CCD 的感光元件超载,失去分辨能力,导致图像曝光过度并且完全丧失细节信息,无法进行修补;光照太弱使得 CCD 感光较少,导致曝光不足,采集的图像偏暗,隐藏较多细节信息,同样不利于后续处理。人脸检测系统常对光照条件做出一定的限制,采集图像的光照基本均匀,仅允许小范围内的光照强度变化。然而,实际的图像采集环境中往往会有偏光、侧光、高光的存在,导致采集的图像亮度不均、过亮或过暗,这些都会使人脸检测率大幅度下降。Adini 等人^[15]的研究显示,光照变化对人脸检测系统性能影响巨大,同一个人在不同光照条件下,图像差距会非常大。Jacobs 等人^[16]对人脸光照图像进行详细分析后得出结论:人脸图像的像素值会随着光照条件的改变而变化,很多时候会大于人的不同所引起的差异。在文献[17]中指出光照条件变化是导致人脸识别率下降的主要原因,光照问题是图像处理中的关键问题之一。图 3-13 为 Yale Faces 库中人在不同光照条件下的图像,可以看出像素值随光照强度变化而出现剧烈变化。

在人脸检测系统中解决光照问题的方法有:

1) 建立光照模型或人脸模型,例如光照锥方法^[18]和基于光照样本的方法^[19],此类方法理论性强,但其假设过多,在实际场景中难以满足条件。

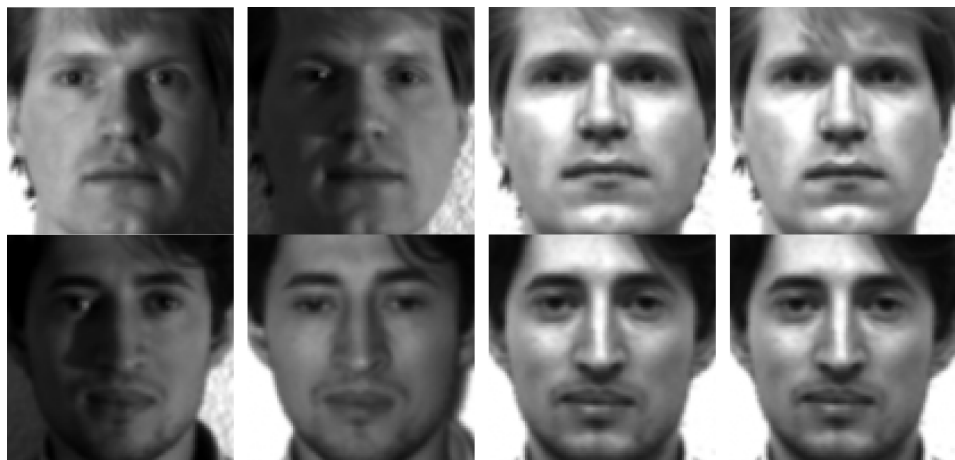


图 3-13 不同光照条件下的人脸图像

2) 提取光照不变特征, 如基于边缘的图像特征有: 图像边缘、轮廓、Gabor 特征和角点等^[20]; 基于图像的方法有: Eigenface、Fisher-face 方法, 以及熵图像等^[21,22]。

3) 光照补偿法。对于均匀光照变化常用的方法有: 直方图均衡法、Gamma 校正法、对数变换法等^[23]; 对于非均匀光照变化常用的方法有: 自适应直方图均衡法和分块直方图均衡法、非线性自适用法、方差滤波器法^[24,25]。此类算法较简单, 但这些方法对图像的校正并不理想, 在性能上往往难以达到理想效果。

本章提出一种新的基于图像分解和 MSR 算法相结合的光照补偿方法, 该方法共分四个部分: 图像分解、光照补偿、图像合成和光照均衡。基于空间梯度方法的图像分解首先将原始图像分解为高频、低频图像, 光照信息完全包含在低频图像中; 然后对低频图像应用 MSR 算法进行光照补偿, 将较暗区域增亮; 将补偿后的低频图像和高频图像合成, 并进行光照均衡得到新图像。

3.4.1 基于图像分解和 MSR 算法的光照补偿

环境中的光照和物体材料自身的一些属性决定了物体表现出不同的颜色。物理光学研究了光照射物体时所表现的特性, 当光照射物体时, 由于物体属性不同, 而导致吸收不同波长的光, 例如: 当用白光照射物体时, 如果物体只吸收黄色光, 则按补色原理, 该物体显示的是蓝色, 物体表面颜色取决于该物体反射出光的波长。

光照模型一般分为局部光照模型和全局光照模型两种。局部光照模型只研究直接光源对物体的照射, 不考虑周围环境中其他反射光的影响, 并假定物体表面光滑, 所用材料是理想材料。该模型虽然与实际的光照效果相差较大, 但





由于计算简单,并能模拟部分不透明物体的光照效果,使物体的显示具有真实感,因此,该模型具有一定的实用价值。“朗伯反射模型”属于局部光照模型。

根据朗伯反射模型理论,图像可以表示为

$$I(x, y) = \rho(x, y) n(x, y)^T s(x, y) \quad (3-11)$$

式中 $\rho(x, y)$ ——图像坐标点 (x, y) 对应的物体表面反射系数,仅由物体材质决定,而与光照无关, $0 \leq \rho(x, y) \leq 1$;

$n(x, y)$ ——图像坐标 (x, y) 对应物体的表面法向量;

$s(x, y)$ ——点光源在 (x, y) 点上的入射方向,其幅值为入射光强度。

全局光照模型除考虑光源照射物体的效果外,还考虑周围物体反射光对它的影响,因此物体表面颜色是所有照射到该物体表面上的光综合后的效果,其强度为所有光强度之和,这种模型虽较真实地模拟了客观世界中的三维物体,但计算较复杂^[26]。

若定义 $r(x, y) = \rho(x, y) n(x, y)^T$, $l(x, y) = s(x, y)$ 则可以得到全局的辐照度模型

$$I(x, y) = r(x, y) \times l(x, y) \quad (3-12)$$

式中 $r(x, y)$ ——物体表面 (x, y) 点的反射系数, $0 \leq r(x, y) \leq 1$;

$l(x, y)$ ——光照度。

Retinex 理论是 Land 等^[27]于 1963 年提出的,是解释人类视觉如何调节感知到物体颜色和亮度的模型,Retinex 是由 Retina (视网膜) 和 Cortex (大脑皮层) 两个单词合成而来的。该模型的基本思想是:人感知到某点的光照并不只取决于该点的绝对光照值,还和其周围的光照值有关,Retinex 诠释了同一物体在不同光源或光线下颜色恒定 (Color Constancy) 的机理^[28-30]。

用 Retinex 理论去除光照影响时,从模型上来讲就是要由图像 $I(x, y)$ 中求解反射图像 $R(x, y)$ 和光照图像 $L(x, y)$ 。为求解该不定方程,由 Retinex 理论衍生出的 Retinex 去光照算法做了如下约束:光照 $L(x, y)$ 在图像空间上是平滑的,而反射系数 $R(x, y)$ 为常数, $(x, y) \in \Omega$, Ω 为一小邻域;同时要求 $L(x, y) > R(x, y)$,即光照的值接近图像的亮度值。

根据 Retinex 理论,原始图像 $I(x, y)$ 是光照图像 $L(x, y)$ 和反射图像 $R(x, y)$ 的乘积,即

$$I(x, y) = L(x, y) \times R(x, y) \quad (3-13)$$

光照图像 $L(x, y)$ 决定了一幅图像中像素所能达到的动态范围,而反射图像 $R(x, y)$ 决定了一幅图像的内在性质。式 (3-12) 与式 (3-11) 类似,可见 Retinex 理论符合辐照度模型。Retinex 理论的实质就是从原始图像 $I(x, y)$ 中获得物体的反射图像 $R(x, y)$,以消除光照不均的影响,保持图像颜色恒常性,压缩图像动态范围,提高对比度,有效地显示淹没在阴影区域中的细节,以改善图像



的视觉效果,因而该理论在图像增强处理中有着广泛应用。

Retinex 算法首先将图像像素转化到对数域^[27], 对式 (3-12) 两侧同时进行对数运算, 得到式 (3-13)

$$\log I(x, y) = \log L(x, y) + \log R(x, y) \quad (3-14)$$

经对数运算后式 (3-12) 由乘法运算简化为加法运算, 提高了系统的计算能力。Land^[31] 提出了输出函数为

$$R'(x, y) = \log I(x, y) - \log(F(x, y) * I(x, y)) \quad (3-15)$$

式中 $R'(x, y) = \log R(x, y)$ 。

Jobson 等人^[28] 将 $F(x, y) = e^{-|r|^2/c^2}$ 命名为中心/包围函数, 并且表示为

$$F(x, y) = F(x, y, \sigma) = Ke^{-(x^2+y^2)/\sigma^2} \quad (3-16)$$

式中 K ——由归一化条件 $\iint F(x, y, \sigma) dx dy = 1$ 来确定;

σ ——高斯函数的标准偏差。

标准差 σ 的选择会直接影响图像增强的效果, 当 σ 取较小值时, 能够较好地进行动态范围的压缩; 反之, σ 取较大值时, 图像色感一致性较好。在实际处理中, 通常是在动态范围压缩和图像色感一致性之间寻找平衡点, 如果要在图像的动态范围压缩及图像的色感一致性这两方面都取得良好的效果, 需采用多尺度 Retinex 算法。

多尺度 Retinex (Multi-Scale Retinex, MSR) 算法不仅可以很好地将图像进行动态压缩, 还可更好地保持图像的细节信息^[32], 计算公式如下

$$R(x, y) = \sum_{i=1}^n \omega_i [\log(I(x, y)) - \log(F_i(x, y) * I(x, y))] \quad (3-17)$$

式中 ω_i ——对应尺度权重系数;

n ——尺度数目;

$$\sum_{i=1}^n \omega_i = 1$$

$F_i(x, y)$ ——对应尺度中心函数。

试验结果显示^[33]: σ 一般取大、中、小三个尺度, 每个尺度的权重根据侧重点的不同需要来选取。

Retinex 算法的物理意义是: 在对数空间中将原图像减去高斯函数与原图像的卷积, 如式 (3-14) 所示, 卷积实际上是图像的平滑部分 (光照图像), 剩下的是图像的高频部分 (反射图像), 从而突出了原图像中的细节。因此, 对于图像中存在的阴影区域, 通过 Retinex 算法增强后, 可以突出暗区域中的细节。

将图像分解为高频和低频两部分, 高频图像主要包含图像的细节成分, 低频图像包含图像的光照成分, 这样在对图像进行光照补偿时, 只需补偿低频图





像即可。因此,该方法可以最大限度地保留图像的边缘细节,以便图像的后续处理。

图像梯度是图像处理中的一个常用的局部不连续度测量方法,不连续度反映了图像中每个像素点与邻域像素点的不连续水平,将当前像素值与邻域像素值进行差分运算得到梯度值,具体计算如下

$$\nabla I = |I(x, y) - I(i, j)| \quad (3-18)$$

式中 $I(i, j)$ ——当前像素 $I(x, y)$ 邻域的像素值。

对于彩色图像,分别计算每个通道的差分。当差分值满足 $\nabla I > k$ 时,认为图像颜色变化明显,通过实验得 $k = 15$ 。

定义局部空间有效平均梯度

$$\overline{\nabla I(x, y)} = \frac{1}{N_{(x, y)}} \sqrt{\sum_{(i, j) \in \Omega} k_{(i, j)} [I(x, y) - I(i, j)]^2} \quad (3-19)$$

式中 Ω ——像素点 (x, y) 的邻域;

(i, j) ——像素在邻域中的位置,选用 3×3 邻域;

$k_{(i, j)}$ ——决定差分有效性的参数,定义如下

$$k_{(i, j)} = \begin{cases} 1 & \text{if } \nabla I_{(i, j)} > 15 \\ 0 & \text{else} \end{cases} \quad (3-20)$$

$N_{(x, y)}$ ——归一化系数,计算公式如下

$$N_{(x, y)} = \begin{cases} \sum_{(i, j) \in \Omega} k_{(i, j)} & \text{if } \sum_{(i, j) \in \Omega} k_{(i, j)} > 0 \\ 1 & \text{if } \sum_{(i, j) \in \Omega} k_{(i, j)} = 0 \end{cases} \quad (3-21)$$

如果 $k_{(i, j)} > 15$,认为是一次有效差分。有效差分的平均值称为局部空间有效平均梯度,这种梯度处理接近各向异性扩散,可以有效地保持图像细节,梯度图像如图 3-14b 所示。



图 3-14 人脸图像



可以看出梯度图像只包含图像的细节，并不包含变化缓慢的光照信息，梯度图像又称为高频图像，因此，对原图像进行光照补偿可以先滤掉该部分与光照无关的图像。

图像分解是在空间域上进行的像素级分解，把原始输入图像分解为梯度图像和光照图像两部分，梯度图像包含细节信息，光照图像包含光照信息，分解定义如下

$$I'(x, y) = I(x, y) - \nabla I(x, y) \quad (3-22)$$

式中 $I'(x, y)$ ——光照图像；

$I(x, y)$ ——原始输入图像。

对原始图像图 3-14a 进行图像分解，分解后图像如图 3-15 所示。



图 3-15 光照图像

从图 3-15 中可以看出，由于去掉了图像的大部分细节信息，使得分解后的光照图像变得模糊不清，但是该图像包含了大量光照信息，光照图像又称为低频图像。对原始图像的光照补偿就可以转化为对光照图像的光照补偿，补偿后再还原为原始图像即可。

采用 MSR 方法补偿彩色图像，对彩色图像的处理方法有两种：

一是在 RGB 每个通道上分别采用 MSR 方法进行光照补偿。

二是从 RGB 三个通道中计算光照信息——亮度，只对亮度运用 MSR 方法进行补偿。

本书采用第二种方法进行光照补偿，取 RGB 三个通道上的最大值，作为亮度 $H(x, y)$ 的值，即

$$H(x, y) = \max(I'_R(x, y), I'_G(x, y), I'_B(x, y)) \quad (3-23)$$

式中 $I'_R(x, y)$ 、 $I'_G(x, y)$ 、 $I'_B(x, y)$ ——分别为光照图像 I' 在点 (x, y) 处的 R、G、B 三分量像素值。





通过计算三个通道的最大值求出光照信息，只对光照信息进行 MSR 运算，这样可节省大量计算时间。对 MSR 补偿算法采用三个尺度，即大、中、小三个尺度取平均，实验测试得出三尺度分别为 20、100、260 时效果最好。通过式 (3-16)，求得 $H(x,y)$ 经 MSR 处理后的亮度 $H'(x,y)$ 。

通过用 MSR 方法对光照成分进行光照补偿后，获得了增强后的亮度图像。利用光照图像中的色彩信息，通过简单的线性变换即可恢复出增强后图像的色彩信息，由于恢复后的图像各像素点的 R、G、B 三分量的比例关系保持不变，因而能够较好地保留原始图像中的颜色信息。增强后彩色图像的恢复方法如下^[34]

$$I_z''(x,y) = r(x,y) \times I_z'(x,y) \quad (3-24)$$

式中 I_z'' ——恢复后的光照图像；

z ——对应颜色通道，分别为 R、G、B；

$r(x,y)$ ——对应点 (x,y) 处的比例因子；

$I_z'(x,y)$ ——原始光照图像。

比例因子 r 计算如下

$$r(x,y) = \frac{H'(x,y)}{H(x,y)} \quad (3-25)$$

式中 H' ——经 MSR 处理后图像亮度分量；

H ——原始图像亮度分量。

图 3-16 为恢复后的光照图像，与图 3-15 对比可以看出，图像左侧较暗区域亮度增大，该方法有效地对图像进行了光照补偿。

用 MSR 方法对低频图像（光照图像）进行光照补偿后，去除了光照信息的影响，较好地解决了偏光问题，但缺少图像细节，即高频图像，需将高、低频图像重构，具体方法如下

$$I_h(x,y) = I''(x,y) + \overline{\nabla I(x,y)} \quad (3-26)$$

式中 $I_h(x,y)$ ——重构后的彩色图像；

$I''(x,y)$ ——通过式 (3-23) 得到的低频图像；

$\overline{\nabla I(x,y)}$ ——式 (3-18) 得到的高频图像。

重构后的图像超出了显示设备的动态范围，利用下式进行归一化

$$I_h' = 255 \times \frac{I_h - I_{h \min}}{I_{h \max} - I_{h \min}} \quad (3-27)$$

式中 $I_{h \min}$ ——重构图像 I_h 的最小像素值；

$I_{h \max}$ ——重构图像 I_h 的最大像素值。

重构后的图像经过归一化后，如图 3-17 所示。



图 3-16 恢复后的光照图像



图 3-17 重构图像

基于图像统计特性的 S 曲线函数，可以增强图像中的较暗区域或降低图像中的较亮区域。 S 曲线定义如下

$$I_s(x, y) = 255 \left(\frac{k_1}{k_2 + e^{-k_3 I(x, y)/k_4}} - k_5 \right) \quad (3-28)$$

式中 I_s ——均衡后图像；

I ——原始彩色 RGB 图像， $0 \leq I(x, y) \leq 255$ ；

k_1 、 k_2 、 k_5 ——归一化系数；

k_3 ——决定曲线倾斜度的参数；

k_4 ——像素特性参数，决定曲线水平方向位移的参数，在此 $k_1 = 2$ 、 $k_2 = 1$ 、 $k_3 = 1$ 、 $k_5 = 1$ 。

从图 3-18 中可以看出： k_4 值越小，曲线越陡峭，对应输出像素值变大； k_4 值越大，曲线越平缓，对应输出像素值变小。反方向调节像素值，当处理暗区域时， k_4 应该取较小值，使图像变亮；处理亮区域时， k_4 应该取较大值，使图像变暗。 k_4 取局部图像的平均值，具体计算公式如下

$$k_4(x, y) = ce^{-(x^2 + y^2)/\sigma^2} \times I(x, y) \quad (3-29)$$

式中 σ ——高斯分布的标准偏差；

c ——由归一化条件 $\iint k_4(x, y) dx dy = 1$ 确定。

图 3-19 为利用式 (2-17) 对重构图像 (图 3-17) 进行均衡后的结果，可以看出均衡后的图像亮度分布更加均匀，人脸肤色特点更加明显，为后续检测奠定了基础。

本书所用算法流程如图 3-20 所示。

为验证算法的有效性进行实验研究，主要与 McCann's Retinex、DCT、小波及直方图算法进行比较。效果如图 3-21、图 3-22 所示。

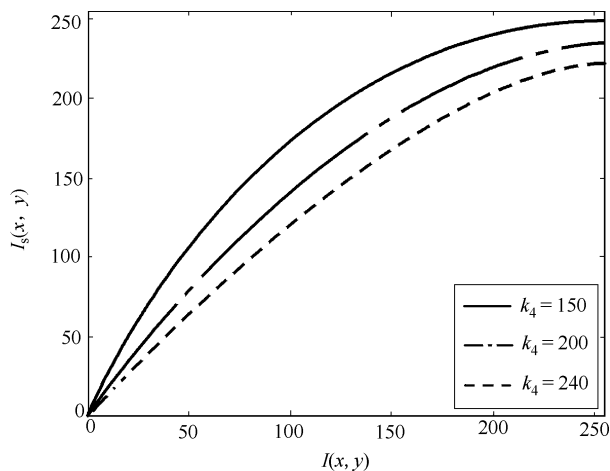


图 3-18 S 曲线



图 3-19 均衡图像

图 3-21 为原始图像（图 3-14a）采用直方图补偿法、离散余弦补偿法、小波补偿法及 Retinex 补偿法处理后的效果图。图 3-22 为另一组实验图像。图 3-21a 和图 3-22b 为采用直方图算法的结果，发现图像阴暗处的亮度依然较低，不能很好地显示出细节；图 3-21b 和图 3-22c 为离散余弦变换结果，结果显示图像整体依然较暗，细节不清楚，甚至不如直方图处理的结果；图 3-21c 和图 3-22d 为小波算法，图像整体变亮，但丢失部分细节，图像对比度变小；图 3-21d 和图 3-22e 为 McCann's Retinex 算法处理结果，图像整体变亮，但丢失较多细节信息，特别是肤色信息丢失严重。图 3-19 和图 3-22f 为本书算法处理

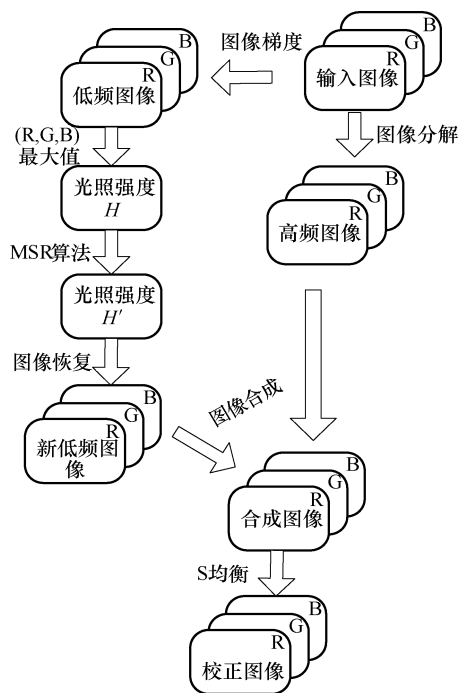


图 3-20 光照补偿流程



图 3-21 各处理方法比较



图 3-22 各处理方法比较

结果，与前面四种算法相比，图像亮度 and 对比度都得到了很好的增强，表现出了更多的图像细节信息。

3.4.2 肤色区域检测

根据式 (3-5) 计算转换后图像中像素属于模型的概率值，利用归一化的方法将概率值转化为 0 ~ 255 之间的整数值，得到被检测彩色图像的灰度图，灰度图可直观反映出各像素点的颜色与肤色的相似程度，灰度值越高的地方越接近肤色，如图 3-23a 所示，并计算灰度图像对应的灰度直方图，如图 3-23b 所示。

选取合适的阈值，将灰度图像二值化，可以分割出图像的肤色区域。二值



化方法主要有双峰直方图阈值分割法、模糊聚类分析法、最大类间方差阈值分割法等^[35]。为自适应获取二值化图像的阈值，本书采用二值化效果较好的最大类间方差阈值分割法^[12]。

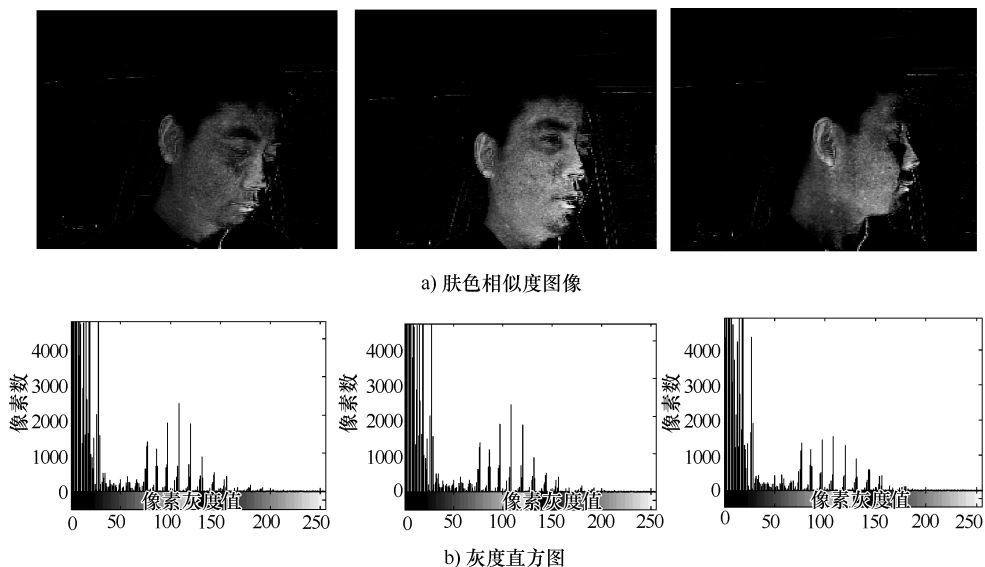


图 3-23 肤色相似度图像及灰度直方图

最大类间方差阈值分割法将灰度直方图在某一阈值处分割成两组，当两组之间的方差为最大时决定阈值。设一幅图像的灰度值为 $1 \sim M$ 级，灰度值为 i 的像素数为 n_i ，可以得到以下参数。

总像素数

$$N = \sum_{i=1}^M n_i \quad (3-30)$$

各灰度值像素数出现的频率为

$$P_i = n_i / N \quad (3-31)$$

然后将所有的像素用 K 将其分成两组 $S_0 = \{1 \sim K\}$ 和 $S_1 = \{K+1 \sim M\}$ ，则 S_0 发生的概率为

$$\eta_0 = \sum_{i=1}^K p_i = \eta(K) \quad (3-32)$$

S_1 发生的概率

$$\eta_1 = \sum_{i=K+1}^M p_i = 1 - \eta(K) \quad (3-33)$$

S_0 的平均值



$$\lambda_0 = \sum_{l=1}^K \frac{ip_l}{\eta_0} = \frac{\lambda(K)}{\eta(K)} \quad (3-34)$$

S_1 的平均值

$$\lambda_k = \sum_{l=1}^K \frac{ip_l}{\eta_l} = \frac{\lambda - \lambda(K)}{1 - \eta(K)} \quad (3-35)$$

式中 $\lambda = \sum_{i=1}^M ip_i$ ——图像所有像素灰度的平均值;

$\lambda_k = \sum_{i=1}^M ip_i$ ——灰度值低于 K 的像素的灰度平均值。

定义 S_0 、 S_1 两组间的方差为

$$\begin{aligned} \sigma^2(K) &= \eta_0(\lambda_0 - \lambda)^2 + \eta_1(\lambda_1 - \lambda)^2 \\ &= \eta_0\eta_1(\lambda_1 - \lambda_0)^2 \end{aligned} \quad (3-36)$$

寻找 $1 \sim M$ 间使 $\sigma^2(K)$ 取得最大值的 K , 即可得到阈值 K 。利用该阈值对上述相似度图像进行二值化, 二值化计算公式为

$$I(x, y) = \begin{cases} 1 & I(x, y) \geq K \\ 0 & I(x, y) < K \end{cases} \quad (3-37)$$

得到二值化图像如图 3-24 所示, 白色区域为人脸皮肤。



图 3-24 相似度图像二值化结果

3.4.3 驾驶员面部预定位

由二值化图像可以看出, 驾驶员面部图像周围存在部分白色区域, 这对定位驾驶员的面部区域存在不利影响, 可以利用数学形态学方法消除这些区域。

数学形态学以几何学为基础, 着重研究了图像的几何结构, 看是否能将这个结构元素很好地填放在该图像内部, 腐蚀和膨胀是最基本的数学形态学运算^[36, 37]。

1. 膨胀

已知二值图像 A , 如果 $A_{b1}, A_{b2}, \dots, A_{bn}$ 是由二值图像 $B = \{b_1, b_2, \dots, b_n\}$ 中像素为 1 的点平移得到, 则 A 由 B 平移的并称为 A 被 B 膨胀, 即

$$A \oplus B = \cup A_{b_i} \quad (3-38)$$

2. 腐蚀

腐蚀是膨胀的相反过程。二值图像 A 经二值图像 B 腐蚀后在 p 点仍为 1 的



充分必要条件是： B 平移到 p 后， B 中的 1 像素也是 A 中的 1 像素。 A 被 B 腐蚀可用式 (3-38) 表示

$$A \ominus B = \{p | B_p \subseteq A\} \quad (3-39)$$

两种运算的不同组合可以得到不同的处理效果。对二值化后的图像先进行开运算，消除面部周围的细小干扰区域，然后再进行闭运算，填充嘴部区域的细小空洞，连接临近黑色区域，并进行平滑边界。再对白色连通区域内的黑色孔洞进行填充，得到图像如图 3-25（最左边一列），可看出驾驶员面部周围的杂散点已经被滤除，并且形成一个完整的脸部轮廓。

由于二值图像投影包含了图像的很多信息，所以投影是二值图像的一种简洁表示方式。在某些应用中，投影可以作为物体识别的一个特征。

3. 二值图像投影

给定一条直线，用垂直该直线的一簇等间距直线将该二值图像分割成若干条，每一条内像素值为 1 的像素量为该条图像在给定直线上的投影值。计算二值化图像的水平、垂直方向投影，如图 3-25 所示，其中中间列为水平投影，最右边列为垂直投影。

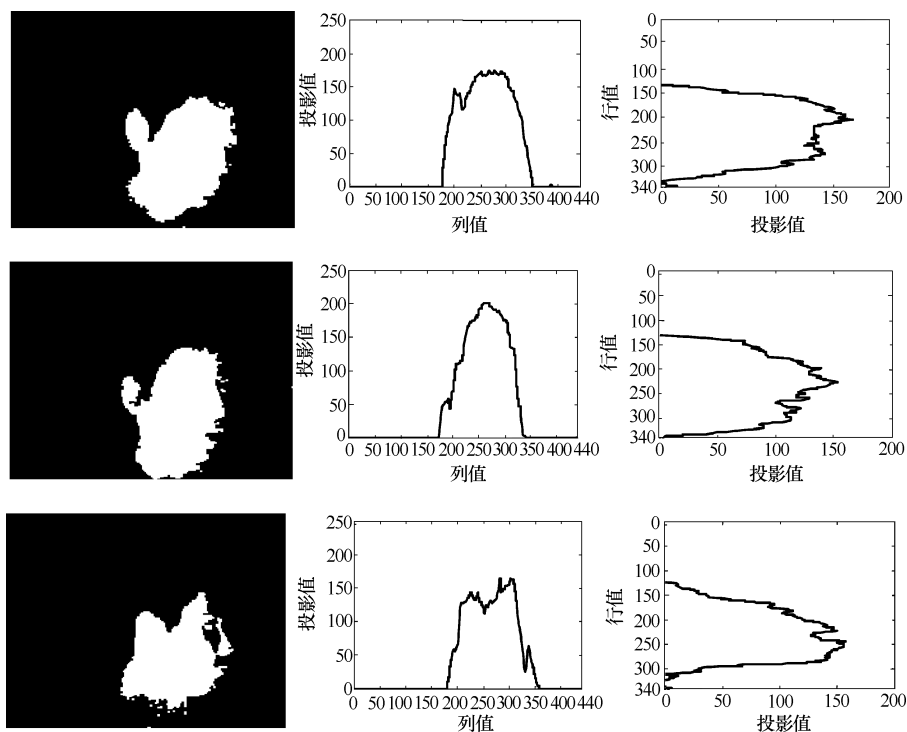


图 3-25 形态滤波后的图像及投影曲线



因为人脸对应区域经二值化转换后形成像素值为 1 的密集区，所以对应区域的投影值较大。可以利用这一投影特征初步定位人脸区域并分割该区域，如图 3-26 所示。



图 3-26 面部图像分割

由于驾驶员颈部颜色与人脸颜色相近，定位面部时没能把该区域排除在外，这对估计驾驶员的面部姿势有不利影响，因此需要进一步精确地定位人脸。

3.5 驾驶员面部精定位

人脸内显著的特征点有眉毛、眼睛、鼻、嘴唇等，这对估计驾驶员的面部姿势很重要。分析这些特征点的相对位置关系，发现可以通过定位眉毛上边缘与嘴唇下边缘，来实现驾驶员面部上、下边界的精确定位。

3.5.1 联合投影函数

投影是一种重要的图像分析方法，二维图像可以由正交的两个一维投影函数来分析，维数的降低在减小计算量的同时也使其更便于分析。积分投影函数不但被成功地应用于人脸识别，还可以应用于边界分析，从而确定出人脸各个特征的位置^[38]。

1. 积分投影算法

假设 $I[x, y]$ 表示点 (x, y) 处的像素灰度值，在区间 $[x_1, x_2]$, $[y_1, y_2]$ 内的垂直积分投影函数和水平积分投影函数分别为 $I_v(x)$ 、 $I_h(y)$ ，则定义为

$$I_v(x) = \sum_{y=y_1}^{y_2} I(x, y) \quad (3-40)$$

$$I_h(y) = \sum_{x=x_1}^{x_2} I(x, y) \quad (3-41)$$

2. 微分投影算法

式 (3-39)、式 (3-40) 为积分投影函数，现在还经常应用到微分投影函



数,其概念最早由 Feng 和 Yuen^[39]提出,并同时提出了一种利用微分投影函数定位眼睛的简单方法。

假设 $I[x, y]$ 表示点 (x, y) 处的像素灰度值,在区间 $[x_1, x_2]$, $[y_1, y_2]$ 内的垂直微分投影函数和水平微分投影函数分别表示为 $D_v(x)$ 和 $D_h(y)$,则定义相邻像素间差值绝对值的和为

$$D_v(x) = \sum_{y=y_1}^{y_2} |I(x, y) - I(x-1, y)| \quad (3-42)$$

$$D_h(y) = \sum_{x=x_1}^{x_2} |I(x, y) - I(x, y-1)| \quad (3-43)$$

根据式 (3-41) 和式 (3-42) 可以看出,微分投影算法反映的是像素某一行或某一列中相邻像素的变化程度,即像素灰度值变化越剧烈投影值 $D_v(x)$ 和 $D_h(y)$ 就会越大。

3. 微积分联合投影算法

积分投影函数只考虑了各行各列均值的变化,却不能反映均值相同情况下同一行列内灰度值的变化,而微分函数则主要反映了同一行列内像素灰度值变化的剧烈程度。两种方法具有互补性,故将两者结合,其中垂直方向投影函数的结合公式为^[12]

$$M_v = (1-k)D_v - k \times I_v \quad (3-44)$$

式中 k ——权值系数,该系数确定方法为

$$k = \frac{\sum_{i=1}^m D_v(i)/I_v(i)}{m} \quad (3-45)$$

式中 m ——图像矩阵行数。

由于存在噪声干扰等问题的影响,投影曲线存在较多毛刺,需对投影曲线进行平滑处理。平滑公式为

$$D_v^*(x) = \frac{1}{L+1} \sum_{i=x-L/2}^{x+L/2} D_v(x) \quad (3-46)$$

式中 L ——取值跟人脸在图像中的大小有关,实验中取 $L=6$ 。

联合投影公式是采用微分投影值与权值 (k) 相乘减积分投影与权值 ($1-k$) 相乘的方法,如果计算结果小于 0 则归一化为 0。从联合投影公式可以看出,如果某个区域像素灰度值的平均值较小,而且变化比较剧烈的话,可以提高该区域投影值与其他区域投影值的对比度。

3.5.2 眉毛上边缘定位

分析眉毛区域可以看出,眉毛所在的水平区域不但灰度值的平均值低,而





且对于整个水平区域来说,其灰度值变化也很剧烈,所以可以应用垂直积分、微分联合投影的方法定位驾驶员眉毛区域。一方面,在人脸区域进行垂直方向积分投影,利用眉毛区域灰度值较小的特点,积分投影 I_v 越小越可能是眉毛。另一方面,在眉毛区域进行垂直方向微分投影,利用人脸图像中眉毛区域的灰度值变化大的特点,差分后绝对值累加得到的 D_v^* 值较大。

把预定位的人脸区域灰度化,然后采用微积分联合投影算法计算垂直投影,实验结果如图 3-27 所示。图像左侧白色区域对应该行像素的投影值。



图 3-27 直接对预定位人脸进行联合投影变换

图 3-27 中可以看出,投影曲线有多个峰值区域,定义从上向下的峰值为第 1、2、3……峰值区。因为眉毛上部为额头,该区域灰度变化比较均匀,而且灰度值较高,因此第 1 峰值区一般为眉毛区域。图 3-27 中也可以看出第 1、2 幅图像的第 1 峰值区正好对应应在眉毛区域,与我们的推理相符,但是第 3 幅图像由于背景的影响,没有使第 1 峰值区对应应在眉毛区域,因此,不能直接在预定位的人脸区利用联合投影的方法定位眉毛。

经过大量实验发现,背景影响投影的情况多发生在面部姿势旋转比较大时。因为这时鼻尖突出人脸的大致范围,在垂直方向上没有与人鼻连接的肤色区域,人脸预定位就包含了部分背景。这时预定位图像的一侧为头发区域,另一侧即为背景区域,需要消除背景区域的影响。这就需要先检测头发区域,那么另一侧就是背景。

采用亮度信息和正则化后的 R—G 颜色信息对头发颜色进行描述^[40]。采用高斯模型来对样本中的发色信息的进行拟合,应用如下的判别函数 $H(x,y)$ 来识别是否是头发

$$H(x,y) = \begin{cases} 1 & p(x,y) > Th \\ 0 & p(x,y) < Th \end{cases} \quad (3-47)$$

这里 (x,y) 为像素点所处的位置, $p(x,y)$ 为像素点在发色概率模型中的概率密度,而 Th 则为经验选定的判断阈值。

因为头发、眉毛在人脸上部,因此取人脸预定位的上 1/3 区域进行头发检



测，如图 3-28a 所示。对预定位人脸区域进行肤色变换，在图像上 1/3 区域寻找非肤色的最高外接矩形，认为该区域为背景图像并滤除。对精简后的图像进行灰度化，并进行垂直投影变换得到投影图像，如图 3-28b 所示。可以看出在消除背景影响后，第一峰值区能准确对应眉毛区域。

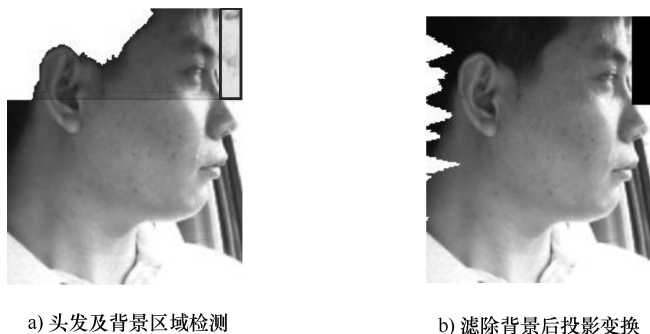


图 3-28 侧面人脸图像投影变换

从上往下以投影值第一次突变的位置即为眉毛上边缘，作为驾驶员面部定位的精确上边缘，定位结果如图 3-29。



图 3-29 定位眉毛上边缘

图中可以看出利用联合投影算法，可以在人脸呈现不同姿势的情况下准确定位眉毛区域。

3.5.3 嘴唇检测

由于人的个体差异及光照变化等原因的影响，使得嘴唇检测成为一个复杂的问题，嘴唇特征提取对周围环境的变化（如胡须、牙齿及舌头等干扰）异常敏感。由于唇色与肤色的不同使得颜色检测成为一个较为有效的方法，文献 [41] 研究显示，在 HSI 颜色空间，色调（ H 分量）对于嘴唇颜色和皮肤颜色具有很高的差异性，通过 H 的差异性来检测嘴唇，该方法只适合于彩色图像；Wang^[42] 使用 Fisher 算法最大化唇色与肤色的分离度以区分嘴唇与人脸的颜色；





Eveno 等^[43]使用彩色滤波器增大唇与皮肤的差异性；Nefian^[44]提出应用线性判别式区分唇与皮肤区域；文献 [45] 采用 ASM 检测嘴唇的轮廓，该方法运算量巨大；文献 [46] 利用监督法区分嘴巴形状，采用 ASM 曲线提取轮廓，然后利用 SVM 进行嘴巴状态分类；Hennecke 等^[47]使用变形模板构建四次多项式来检测跟踪嘴唇；Coianiz 等^[48]通过两条抛物线构建模板检测嘴唇；Benoit 等^[49]采用运动信息，通过人类视觉系统模型检测嘴巴状态；Zhang 等^[50,51]提出基于 Markov 随机场，利用色调及边缘信息进行嘴唇检测及分割，由于没有严格的模型约束，只能得到粗糙的分割结果；文献 [52] 利用基于区域分割的方法检测嘴唇，利用 Markov 随机场结合运动和颜色信息检测嘴巴，然后采用 ASM 检测嘴巴的外部轮廓和内部轮廓，本方法可较为准确地检测出嘴巴位置，但 Markov 随机场需为颜色分布的松弛处理进行初始化，另外，由于嘴唇的形状没有限制可能导致检测失败；文献 [53] 用统计的方法从人脸中提取嘴唇，该模型包含了一定数量的特征点，特征点可最优内外部的轮廓曲线，但由于该模型是非线性的，所以必须进行一个优质的初始化。

综合以上信息，本书作者提出采用基于人脸几何分布特征与颜色检测相结合的方法进行嘴唇的检测。首先，根据面部特征的几何分布缩小嘴唇检测的范围，并划定标准肤色区域。然后，在 HSI 空间中将候选区域内的 H 分量与标准肤色内的 H 分量进行对比，映射出嘴唇区域。

首先，将根据人脸特征分布情况，划分出标准肤色区域和嘴唇检测区域。标准肤色区域内只包含面部皮肤，作为进行嘴唇检测的对比区域；嘴唇区域内包括嘴巴，可能有牙齿、胡须等。

在划分区域前，需对人脸姿势进行纠正，使两眼处于同一水平线上，以便准确分割出两个区域。根据两眼中心位置校正人脸姿态，模型如图 3-30 所示。

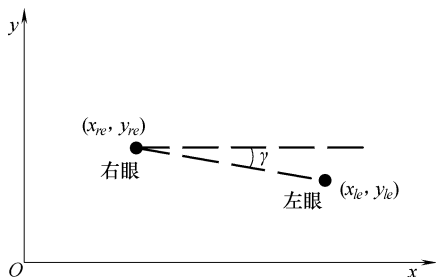


图 3-30 人脸姿态纠正示意图

假设两眼连线与水平方向的夹角为 γ ，其计算公式如下

$$\gamma = \begin{cases} 2\pi - \gamma' & \text{if } \gamma' > 0 \\ \gamma' & \text{else} \end{cases} \quad (3-48)$$

式中 $\gamma' = \arctan\left(\frac{y_{le} - y_{re}}{x_{le} - x_{re}}\right)$;

(x_{le}, y_{le}) ——驾驶员左眼中心坐标；

(x_{re}, y_{re}) ——驾驶员右眼中心坐标。

原始图像以点 (x_c, y_c) 为中心，逆时针旋转 γ 后，新坐标计算公式如下所示



$$\begin{cases} x_1 = (x - x_c) \cos \gamma - (y - y_c) \sin \gamma + x_c \\ y_1 = (x - x_c) \sin \gamma + (y - y_c) \cos \gamma + y_c \end{cases} \quad (3-49)$$

本书中图像旋转中心取驾驶员右眼中心, 即 $x_c = x_{re}$, $y_c = y_{re}$ 。姿态纠正效果如图 3-31 所示。

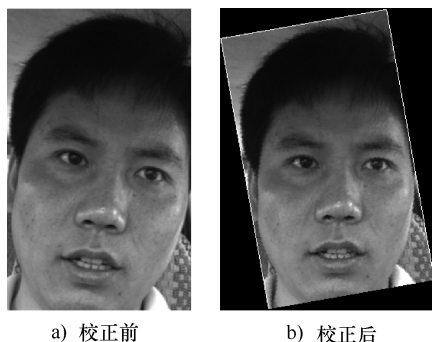


图 3-31 姿态纠正效果图

由面部特征分布可知, 定位出眼睛的位置后, 可以根据双眼坐标划定嘴唇区域。在嘴唇区域内进行检测, 由于缩小了检测范围, 因此可以降低搜索时间, 提高检测效率和精度。根据人脸几何特征——“三庭五眼”关系, 嘴唇位于两眼中间下方约脸长 $1/3$ 的区域内, 如图 3-32 所示。

根据眼睛的已知位置, 划分一个嘴唇的检测区域。区域划分规则如下:

- 1) 横坐标以左、右眼的外眼角为界。
- 2) 纵坐标以人脸下方 $1/3$ 处及人脸下边缘为界。

标准肤色区域 A 的位置在眼睛与嘴巴区域之间, 如图 3-33 所示。

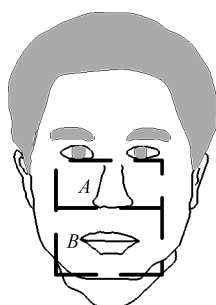
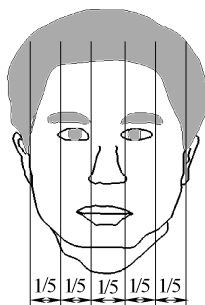
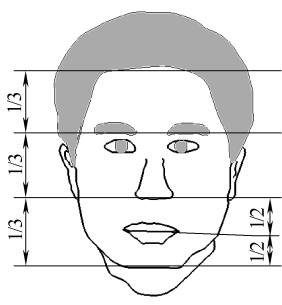


图 3-32 “三庭五眼”示意图

图 3-33 区域示意图

图 3-34 中 A 区域为标准肤色区域, 是检测嘴巴的对比区域; B 区域为嘴巴待检测区域。图 3-33 的分割效果如图 3-34 所示:

在图 3-34b 区域内, 由于鼻孔、牙齿、胡须的影响, 嘴唇检测并不是一个二

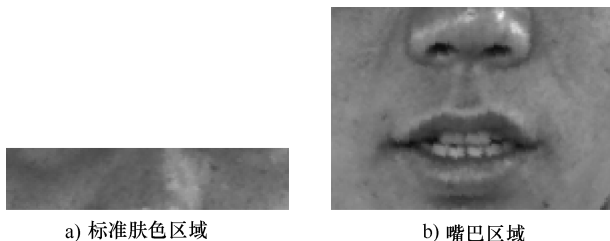


图 3-34 面部区域划分

分类问题（非皮肤即嘴唇）。彩色图像比灰度图像能提供更丰富的信息，而颜色是区分嘴唇与人脸皮肤最为直观的特征，因此可充分利用嘴唇的颜色比皮肤颜色红这一特点进行检测。本节通过对 RGB 彩色图像进行色彩空间变换，转换到 HSI 空间以分离亮度信息，选取区分肤色和唇色比较明显的 H 分量进行嘴唇检测和定位。图 3-34 的 H 分量值空间分布如图 3-35 所示，可以看出，标准肤色区域中 H 值分布比较集中；而嘴巴区域中， H 值非常明显的分为两个部分，在嘴唇位置处 H 值比其他位置的值大，两处的值相差较大，且各自数值聚类性好，容易区分。

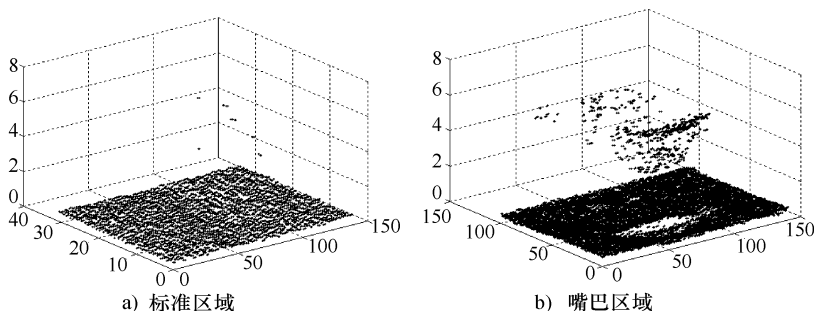


图 3-35 H 分量空间分布

将 H 分量的数值按照下式归一化到 0 ~ 255 范围内

$$H^* = 255 \frac{H - H_{\min}}{H_{\max} - H_{\min}} \quad (3-50)$$

式中 H_{\max} 、 H_{\min} ——分别为 H 的最大和最小值，归一化后灰度图像如图 3-36 所示。

从图 3-36 中可以看出，标准区域的 H 值分布均匀，嘴巴区域在嘴唇及鼻孔处 H 值与其他位置的差别较大，验证了上述结论的正确性。

在 HSI 空间进行嘴唇映射的具体步骤如下所述。

1) 计算 A 区域的 H 分量均值

$$\overline{H}_A = \frac{5}{4n_A} \left(\sum_{(i,j) \in \Omega_A} [H(i,j)] - S_{H_{\max}} - S_{H_{\min}} \right) \quad (3-51)$$



式中 n_A ——A 区域的总像素数;

$H(i,j)$ ——A 区域中 (i,j) 处 H 分量值;

Ω_A ——A 区域;

S_{Hmax} ——A 区域内 H 分量值前 10% 最大的加和;

S_{Hmin} ——A 区域内 H 分量值前 10% 最小的加和。

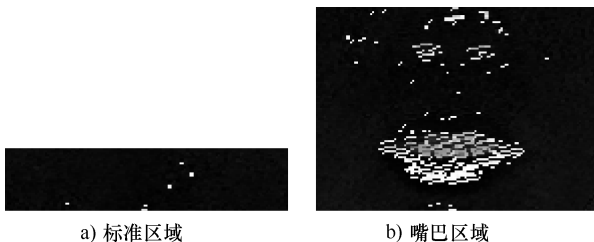


图 3-36 H 分量映射灰度图

2) 计算 B 区域内 H 的相对变化率

$$K(i,j) = \frac{H(i,j)}{H_A} \quad (3-52)$$

式中 $K(i,j)$ —— B 区域内点 (i,j) 的相对变化率;

$H(i,j)$ —— B 区域中 (i,j) 处 H 分量值。

如果某点的颜色较红, H 分量值就会较大, 即相对变化率 K 就会变大, 基于该特性, 只要 K 满足以下关系就可映射出嘴唇

$$K > \Delta K$$

通过试验得出 $\Delta K = 1.7$ 时, 可以精确地检测出嘴唇, 然后根据式 (3-52) 映射出嘴唇轮廓, 映射图如图 3-37 所示。

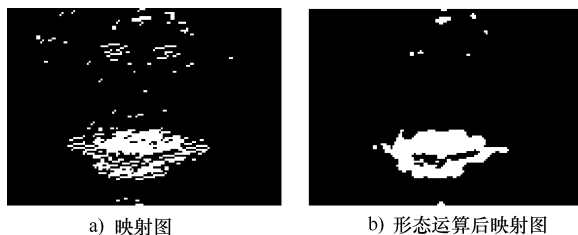


图 3-37 嘴巴区域灰度图

$$I(i,j) = \begin{cases} 255 & K(i,j) > \Delta K \\ 0 & K(i,j) \leq \Delta K \end{cases} \quad (3-53)$$

式中 $I(i,j)$ ——嘴唇区域对应的灰度图。

对同样三段视频, 进行嘴唇检测, 嘴唇检测效果如图 3-38 所示, 可以看出驾驶员在很多姿势下, 都能较准确地检测出嘴唇的外轮廓。



视频一



视频二



视频三

图 3-38 嘴唇检测



3.5.4 驾驶员脸部精定位

定位眉毛与嘴唇区域的边界后，以眉毛上边缘作驾驶员面部上边界，以嘴唇下边缘作面部下边界，以面部预定位左右边缘为左右边界，然后在该区域内对面部预定位图像进行肤色检测。对肤色检测后的图像进行二值化水平投影。以投影值由0突变的位置作为右边界，定位左、右边界如图3-39所示。



图3-39 精确定位驾驶员面部左右边界

综合精确定位的驾驶员面部上、下、左、右边界，可以得到驾驶员面部精确定位结果，如图3-40所示。其中较小方框为精确定位的驾驶员面部图像，并以“o”标记4个顶点，较大方框为驾驶员面部预定位结果。



图3-40 驾驶员面部精确定位及分割结果

从图3-40中可以看出，本书提出的方法在驾驶员多个面部姿势下都能精确





地定位面部图像，并包括了我们需要的面部特征点。图 3-40 中还可以看出驾驶员精确定位的左右边界与预定位的左右边界相差不大，精定位与预定位的差别主要在于上下边界的不同。

3.6 驾驶员面部图像归一化

采用上述由粗到精的逐步定位方法获得的面部图像，无论其高度还是宽度基本上都是不同的。而大小不一的图像不利于估计驾驶员面部姿势，需要进行归一化处理。文中采用了基于插值变换的归一化方法^[53]。

根据原图像中的位置来决定像素点的灰度值，设 $f(x,y)$ 是原图像， $g(x_0,y_0)$ 是归一化后的图像，设 (x_0,y_0) 是图像 $g(x_0,y_0)$ 中的任意一点，对应于 $f(x,y)$ 中的点 (a,b) ，根据 (a,b) 的具体情况来表示 $g(x_0,y_0)$ 中各像素点的值，公式为

$$\begin{cases} a = \frac{wid'}{wid}x_0 \\ b = \frac{heig'}{heig}y_0 \end{cases} \quad (3-54)$$

式中 wid ——原图像的宽度；

$heig$ ——原图像的高度；

wid' ——归一化后图像的宽度；

$heig'$ ——归一化后图像的高度。

当 a, b 为整数的时，有 $g(x_0,y_0) = f(a,b)$ ；当 a, b 不是整数时，通过插值变换得到整数坐标值。

1. 最近邻插值

最近邻插值算法中，归一化图像 $g(x_0,y_0)$ 的灰度值用 (a,b) 邻接的四个网络点 (i,j) 、 $(i+1,j)$ 、 $(i,j+1)$ 、 $(i+1,j+1)$ 中最接近的点的灰度值来近似表示。即

$$g(x_0,y_0) = f(a',b') \quad (3-55)$$

式中 (a',b') —— (a,b) 距其最近的一个点。

例如，图 3-41 中 $(i,j+1)$ 是 (a,b) 的最近邻点 (a',b') 。

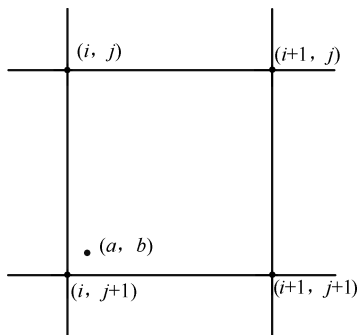


图 3-41 最近邻插值

2. 双线性插值

双线性插值中 $g(x_0,y_0)$ 的灰度值采用 (a,b) 邻近的四个网络点 (i,j) 、 $(i+1,j)$ 、 $(i,j+1)$ 、 $(i+1,j+1)$ 的灰度值按



照下面的公式决定，即

$$g(x_0, y_0) = f(i, j)(1 - \eta)(1 - \beta) + f(i, j + 1)\beta(1 - \eta) + f(i + 1, j)\eta(1 - \beta) + f(i + 1, j + 1)\eta\beta \quad (3-56)$$

式中 i, j ——非负整数， $\eta = a - i$ ， $\beta = b - j$ 。

双线性插值计算量较大，但图像质量高，不会出现像素不连续的情况。由于双线性插值具有低滤波器的性质，所以可能会使图像轮廓在一定程度上变得模糊。

3. 三次线性插值

三次线性插值中 (x_0, y_0) 的灰度值是用与 (a, b) 临域的 16 个网络点的灰度值来近似表示，需要增加计算量，并且计算方法复杂。

分别采用 3 种方法对图 3-13 的面部图像进行归一化实验，结果如图 3-42 所示。



图 3-42 各种插值方法的归一化效果





图 3-42 中从左至右分别为最近邻插值、双线性插值和三次线性插值结果。三种插值方法得到的效果基本相同,但是三次线性插值归一化图像需要的计算时间比较长,另外,应用最近邻插值方法归一化灰度图像时得到的边缘效果不如另外两种方法得到的效果好。因此,文中采用双线性插值方法对分割的驾驶员面部图像进行归一化。

3.7 本章小结

精确定位驾驶员面部图像是估计驾驶员面部姿势、视线方向的前提,因此本章提出由粗到精的驾驶员面部定位方法。首先,在 YCbCr 彩色空间下检测肤色区域并根据二值图像投影进行人脸预定位;然后,在人脸预定位的区域内,提出利用联合投影的方法定位眉毛区域,并根据唇色二次多项式模型及其比人脸肤色更红的特点定位嘴唇区域;最后,以眉毛上边缘、嘴唇下边缘为人脸精定位的上、下边界,并在该区域内再利用肤色模型定位左、右边界。根据需要对精定位的人脸区域利用双线性插值法进行大小归一化。从定位效果上看出,我们提出的方法可以精确定位不同姿势下的驾驶员面部图像。

参考文献

- [1] M. H. Yang, D. J. Kriegman, N. Ahuja. Detecting Faces in Images: A Survey [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002, 24 (1): 34-58.
- [2] Kyoung-Mi Lee. Component-based Face Detection and Verification [J]. Pattern Recognition Letters. 2008, 29 (3): 200-214.
- [3] Hsiuao-Ying Chen, Chung-Lin Huang, Chih-Ming Fu. Hybrid-boost Learning for Multi-Pose Face Detection and Facial Expression Recognition [J]. Pattern Recognition. 2008, 41 (3): 1173-1185.
- [4] Rein-Lien Hsu, Mohamed Abdel-Mottaleb, Anil K. Jain. Face Detection in Color Images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002, 24 (5): 696-706.
- [5] C. Lin. Face Detection in Complicated Backgrounds and Different Illumination Conditions by Using Yeber Color Space and Neural Network [J]. Pattern Recognition Letters. 2007, 28 (16): 2190-2200.
- [6] 王亮, 胡卫明, 谭铁牛. 人运动的视觉分析综述 [J]. 计算机学报. 2002, 25 (3): 225-237.
- [7] Barron J, Fleet D, Beauchemin S. Performance of Optical Flow Techniques [J]. International Journal of Computer Vision. 1994, 12 (1): 42-77.
- [8] 许宏. 基于图像差分的关键帧检测技术 [J]. 计算机工程与设计, 2010, 31 (12):



2849-2852.

- [9] D. Benjamin, J. S. Boaz, K. H. Francis. Comparison of Five Color Models in Skin Pixel Classification [C]. International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, RATFG-RTSV, 1999: 58-63.
- [10] S. L. Phung, A. Bouzerdoun, D. Chai. A Novel Skin Color Model in YCrCb Color Space and Its Application to Human Face Detection [C]. IEEE International Conference on Image Processing, Rochester, USA, 2002: 289-292.
- [11] 盛敬. 驾驶员疲劳监控系统中人脸检测与识别研究 [D]. 辽宁: 东北大学. 2006.
- [12] 朱淑亮. 基于视频图像分析的驾驶员疲劳检测方法的研究 [D]. 济南: 山东大学, 2008: 26.
- [13] H. Greenspan, J. Goldberger, I. Esher. Mixture Model for Face-Color Modeling and Segmentation [J]. Pattern Recognition Letters. 2001, 22 (14): 1525-1536.
- [14] 王志强. 基于彩色图像多线索信息的人脸检测综合研究 [D]. 武汉: 华中科技大学. 2004: 22-25
- [15] Adini, Y, Moses, Y, Ullman, S. Face Recognition; The Problem of Compensating for Changes in Illumination Direction [J]. IEEE Trans on Pattern Analysis and Machine Intelligence. 1997, 19 (7): 721-732.
- [16] Jacobs D W, Belhumeur P N. Basri R. Comparing Images Under Variable Illumination [C]. In Proceedings Conference on Computer Vision and Pattern Recognition, Santa Barbaric. 1998: 610-617.
- [17] Phillips P J, Grother P, Micheals R J et al. Face Recognition Vendor Test 2002: Overview and Summary [R]. National Institute of Standards and Technology, 2003.
- [18] Ishiyama R, Sakamoto S. Geodesic Illumination Basis: Compensating for Illumination Variations in any Pose for Face Recognition [C]. Proceeding of 16th Internet. Conf. Pattern recognition. 2002: 297-301.
- [19] Lee S W, Moon S H. Face Recognition under Arbitrary Illumination Using Illuminated Exemplars [J]. Pattern Recognition. 2007, 40 (5): 1605-1620.
- [20] 刘应乾, 曹茂永, 孙农亮. 基于活动轮廓和 Gabor 滤波器的细胞分割 [J]. 光电子·激光, 2007, 18 (10): 1255-1259.
- [21] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1997, 19 (7): 711-720.
- [22] Wang H T, Wang Y C. Generalized Quotient Image [R]. Proceeding of CVPR, 2004: 498-505.
- [23] Savvides M, Kumar V. Illumination Normalization Using Logarithm Transform for Face Recognition [C]. Proceeding of IAPR AVBPA. 2003: 549-556.
- [24] Xie X D, Lam K M. Face Recognition under Varying Illumination Based on a 2D Face Shape Model [J]. Pattern Recognition. 2005, 38 (2): 221-230.





- [25] 张钰, 姚素英, 徐江涛. 一种非线性自适应图像光照影响去除方法 [J]. 光电子·激光, 2009, 20 (10): 1390-1393.
- [26] 郑范. 人脸图像光照预处理算法研究 [D]. 广州: 华南理工大学. 2010: 8-20.
- [27] Land E, McCann J J. Lightness and Retinex theory [J]. Journal of the Optical Society of America. 1971, 61 (1): 1-11.
- [28] Jobson D J, Rahman Z, Woodell G A. Properties and Performance of a Center/Surround Retinex [J]. IEEE Transactions on Image Processing. 1997, 6 (3): 451-462.
- [29] Funt B, Cardei V, Barnard K. Learning Color Constancy [C]. Proc Fourth IS & T/ SID Color Image Conference. Arizona: Scottsdale. 1996: 58-60.
- [30] Munteanu C, Rosa A. Color Image Enhancement using Evolutionary Principles and the Retinex Theory of Color Constancy [C]. Neural Networks for Signal Processing XI, 2001. Proceedings of the 2001 IEEE Signal Processing Society Workshop, North Falmouth, MA, USA. 2001: 393-402.
- [31] Land E H. An Alternative Technique for the Computation of the Designator in the Retinex Theory of Color vision [C]. Proceedings of the National Academy of Sciences of the United States of America. 1986, 83 (10): 3078-3080.
- [32] Rahman Z, Jobson D J, Woodell G A. Retinex Processing for Automatic Image Enhancement [J]. Journal of Electronic Imaging. 2004, 12 (3): 32-45.
- [33] Li X M. Image Enhancement algorithm based on Retinex theory [J]. Application and Research of Computer. 2004, 2 (9): 235-237.
- [34] 王守觉, 丁兴号, 廖英豪, 等. 一种新的仿生彩色图像增强方法 [J]. 电子学报, 2008, 36 (10): 1970-1973.
- [35] 王飞, 李定主. 在车牌识别项目中对阈值选定法的一点改进 [J]. 机械工程与自动化, 2007, (4): 49-50.
- [36] 王飞. 在车牌自动识别系统中对字符的分割与识别的研究 [D]. 太原: 中北大学, 2007.
- [37] 闫敬文. 数字图像处理技术与图像图形学基本教程 [M]. 北京: 科学出版社, 2002.
- [38] 耿新, 周志华, 陈世福. 基于混合投影函数的眼睛定位 [J]. 软件学报. 2003, 14 (8): 1394-1400.
- [39] G. C. Feng, P. C. Yuen. Variance Projection Function and Its Application to Eye Detection for Human Face Recognition [J]. Pattern Recognition Letters. 1998, 19 (9): 899-906.
- [40] 王江涛. 基于视频的目标检测、跟踪及其行为识别研究 [D]. 南京: 南京理工大学. 2008.
- [41] Zhang X, Broun C C, Mersereau R M, et al. Automatic Speech Reading with Applications to Human-computer Interfaces [C]. EURASIP J. Appl. Signal Process, Special Issue on Audio-Visual Speech Processing. 2002, 1: 1228-1247.
- [42] Wang R, Gao W, Ma J. A Novel Approach to Robust and Fast Locating Lip Motion [C]. Proceedings of 3rd International Conference on Multi modal Interface. 2000: 582-589.



- [43] Eveno N, Caplier A, Coulon P Y. New Color Transformation for Lips Segmentation [C]. Proceedings of International Conference on Image Processing, Thessaloniki, Greece. 2001: 3-8.
- [44] Nefian A, Liang L, Pi X, et al. A Couple HMM for Audio-visual Speech Recognition [C]. Proceedings of International Conference on Acoustics Speech and Signal Processing, Orlando, Florida. 2002: 2013-2016.
- [45] Mathews I, Cootes T F, Bangham J A. Extraction of Visual Features for Lip Reading [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2002, 24 (2): 198-213.
- [46] Yuen P C, Lai J H, Huang Q Y. Mouth State Estimation in Mobile Computing Environment [C]. In: Proc. Sixth International Conference on Automatic Face and Gesture Recognition, Seoul, Korea. 2004: 17-19.
- [47] Hennecke M E, Prasad K V, Stork D G. Using Deformable Templates to Infer Visual Speech Dynamics [C]. Proceedings of 28th Annual Asilomar Conference on Signals, Systems, and Computers. 1994: 578-582.
- [48] Coianiz T, Torresan I L, Caprile B. 2D Deformable Model for Visual Speech Analysis [C]. Speech Reading by Humans and Machines, Stork D, Hennecke MEds. Berlin, Germany: Springer-Verlag. 1996, Volume150, Series F: Computer and Systems Sciences : 391-398.
- [49] Benoit A, Caplier A. Hypovigilance Analysis: Open or Closed Eye or Mouth? Blinking or Yawning Frequency? [C]. IEEE Conference on Advanced Video and Signal Based Surveillance, 2005. AVSS Italy. 2005: 207-212.
- [50] Zhang X, Mersereau R M. Lip Feature Extraction Toward an Automatic Speech Reading System [C]. Proceedings of IEEE International Conference on Image Processing, Vancouver, BC, Canada. 2000, 3: 226-229.
- [51] Zhang X, Mersereau R M, Clements M A, et al. Visual Speech Feature Extraction for Improved Speech Recognition [C]. Proceedings of International Conference on Acoustics Speech and Signal Processing, Orlando, Florida. 2002: 1993-1996.
- [52] Liévin M. , Luthon F. Nonlinear Color Space and Spatiotemporal MRF for Hierarchical Segmentation of Face Features in Video [J]. IEEE Transactions on Image Processing. 2004, 13 (1): 63-71.
- [53] 李学勇. 金属标牌压印凹凸字符的特征提取和识别方法研究 [D]. 济南: 山东大学. 2008.

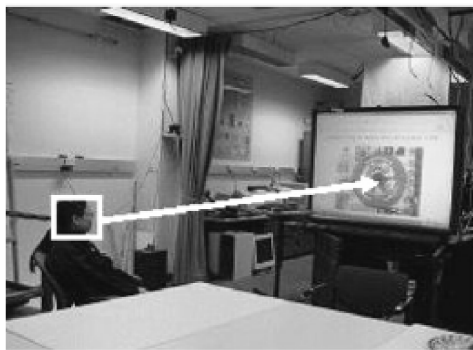
基于核主元分析的驾驶员面部姿势估计

4.1 概述

在许多智能系统中，一个实质的目标是通过图像识别人员的身份和活动，面部姿势是其中一个很重要的线索^[1]，面部姿势还可以揭示一个人的注意方向，因此许多研究者致力于研究如何根据图像来估计一个人的面部姿势^[2,3]。面部姿势估计已经在许多场合中得到应用，如会议中判断参会者的注意方向，估计驾驶员的注意方向，人机交互中利用面部姿势控制程序等，如图4-1所示。根据面部姿势可以检测出85%的驾驶员视觉分散行为^[2]，另外，通过估计操作者的注意方向实现人机交互，也是一个很热门的人机界面课题。面部姿势估计能否成功应用于实际，取决于姿势估计模型的精确度和鲁棒性。除了能分析人们的注意方向，估计面部姿势还有利于处理人脸检测和识别问题，可以提高算法的检测、识别效果^[4]。



a) 判断驾驶员注意方向



b) 人机交互中控制程序

图4-1 面部姿势估计典型应用

基于面部姿势估计在诸多方面的重要应用，众多研究者采用了图像处理、



模式识别的方法对该问题进行了深入研究,并提出许多估计算法。现有的方法大体上可以分为两类:几何分析法^[3-6]、降低数据维数法^[7-9]。

1. 几何分析法

首先,应用人脸特征点定义通用人脸模型或者人脸参考结构,然后,根据人脸相对运动时图像尺寸的缩放、平移、旋转来复原 3D 面部姿势。通常根据 5 个特征点的相对位置来计算姿势角度,5 个特征点分别是 4 个眼角点与鼻尖。然后,根据不同姿势图像中人脸特征点具有的不同几何关系来分析面部姿势。另外, Braathen 等提出利用多粒子滤波器来跟踪这 5 个特征点,再利用它们间的几何关系来估计面部姿势。也有根据图像差分与人脸大体呈椭圆形的特点来估计面部姿势^[1]。实际上,建立侧面人脸几何模型还比较困难,虽然 3D 建模方法可以获得很高精度,但是很难处理单个平面图像。此类方法的特点是需要精确定位人脸的特征点,人脸特征点的定位精度对估计结果影响很大,然而,在不同姿势的人脸图像中定位特征点本身就是个很困难的问题,并且人脸表情发生变化时会影响人脸特征点的相对位置,也会影响估计精度。

2. 降低数据维数法

主要应用主元分析 (Principal Component Analysis, PCA)、最佳 Gabor 过滤器特征空间^[7]、核辨别分析 (Kernel Discriminant Analysis, KDA)、流形学习的方法、支持向量分类法 (SVC)^[10]、多视角特征空间^[11],也有人提出应用小波神经网络的方法来估计面部姿势。此类方法实际上是把面部姿势估计看作一个模式分类问题,这种方法的缺陷是不能通用化,对于不同的人需要重新进行训练。

核主元分析 (Kernel Principal Component Analysis, KPCA) 作为主元分析的非线性推广形式,被认为是特征提取的一种有效手段,已经越来越多地被应用到文字识别、人脸检测与识别等问题中^[12-14]。根据流形学习理论,不同姿势的高维人脸图像,具有一个潜在的低维流形结构^[8]。应用 KPCA 的方法可以把这种潜在流形结构嵌入到低维空间,实现高维数据的可视化。因此,本书提出利用核主元分析估计驾驶员面部姿势的方法。

4.2 核主元分析

4.2.1 基本原理

核主元分析的基本思想是利用一个非线性函数 Φ 将原始输入空间的数据映射到一个高维的特征空间 F 中,使在输入空间无法线性分类的数据变换到线性可分的高维空间,然后在高维空间 F 中利用主元分析方法进行特征提取^[13, 15],



其原理如图 4-2 所示。

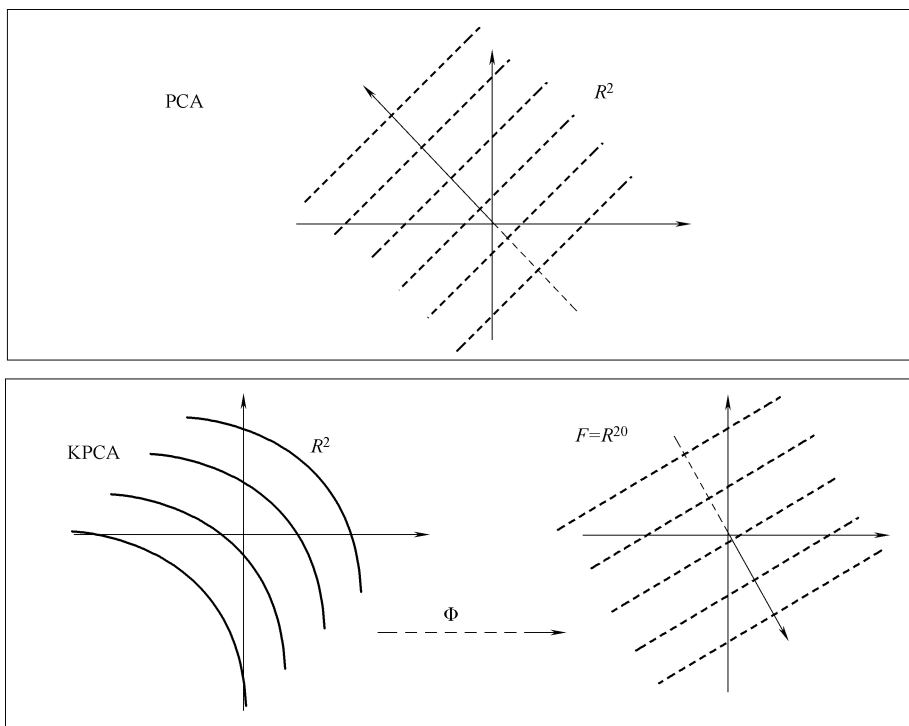


图 4-2 核主元分析基本原理示意图

4.2.2 实现方法

我们可以进行如下算法推导^[14]。设训练样本为 x_1, x_2, \dots, x_M ，定义非线性映射

$$\Phi: R^d \rightarrow F, x \rightarrow \Phi(x) \quad (4-1)$$

将样本从原始输入空间映射到高维特征空间 F 。假设样本映射到高维空间中后均值为零，即

$$\frac{1}{M} \sum_{i=1}^M F(x_i) = 0$$

这样在特征空间 F 中相应的协方差矩阵是

$$C^\Phi = \frac{1}{M} \sum_{k=1}^M \Phi(x_k) \Phi(x_k)^T \quad (4-2)$$

对应的特征方程为

$$\lambda^\Phi u = C^\Phi u \quad (4-3)$$



容易知道, 对于所有满足 $\lambda^\Phi \neq 0$ 的特征向量 u , 它位于 $\Phi(x_1), \Phi(x_2), \dots, \Phi(x_M)$ 所张成的空间中。

因此存在系数 $\alpha_i (i=1, 2, \dots, M)$, 满足

$$u = \sum_{i=1}^M \alpha_i \Phi(x_i) \quad (4-4)$$

把式 (4-4) 代入式 (4-3) 中, 我们可以得到

$$\lambda^\Phi \left(\sum_{i=1}^M \alpha_i \Phi(x_i) \right) = \left(\frac{1}{M} \sum_{k=1}^M \Phi(x_k) \Phi(x_k)^T \right) \left(\sum_{i=1}^M \alpha_i \Phi(x_i) \right) \quad (4-5)$$

在式 (4-5) 的两边同时乘上 $\Phi(x_j), j=1, 2, \dots, M$, 则

$$\lambda^\Phi \left(\sum_{i=1}^M \alpha_i \Phi(x_i) \cdot \Phi(x_j) \right) = \left(\frac{1}{M} \sum_{k=1}^M \Phi(x_k) \Phi(x_k)^T \right) \left(\sum_{i=1}^M \alpha_i \Phi(x_i) \cdot \Phi(x_j) \right), \quad j=1, 2, \dots, M \quad (4-6)$$

定义 $M \times M$ 的矩阵 K

$$K_{ij} = K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \quad (4-7)$$

这里 $K(x_i, x_j)$ 即是核函数。

这样方程 (4-6) 就可以写为

$$M \lambda^\Phi K \alpha = K^2 \alpha \quad (4-8)$$

这里 α 为由 α_i 组成的列向量。由于 K 是对称矩阵, 因此下列方程的解, 即是方程 (4-8) 的解

$$M \lambda^\Phi \alpha = K \alpha \quad (4-9)$$

令 $\lambda_1^\Phi \leq \lambda_2^\Phi \leq \dots \leq \lambda_m^\Phi$ 为方程 (4-9) 的特征值, u^1, u^2, \dots, u^m 为相对应的特征向量。我们需要将 $u^k, k=1, 2, \dots, m$ 在特征空间进行归一化, 即

$$u^k \cdot u^k = 1, k=1, 2, \dots, m \quad (4-10)$$

利用式 (4-4) 和式 (4-9), 我们可得

$$\begin{aligned} u^k \cdot u^k &= \left(\sum_{i=1}^M \alpha_i^k \Phi(x_i) \right) \cdot \left(\sum_{j=1}^M \alpha_j^k \Phi(x_j) \right) \\ &= \sum_{i,j=1}^M \alpha_i^k \alpha_j^k K(x_i, x_j) \\ &= \alpha^k K \alpha^k \\ &= \lambda_k^\Phi (\alpha^k \cdot \alpha^k) \\ &= 1 \end{aligned} \quad (4-11)$$

我们将归一化后的向量 u^1, u^2, \dots, u^m 记为

$$U = [u^1, u^2, \dots, u^m]$$

对于任何一个测试样本 x , 它在 F 中的映射为 $\Phi(x)$, 则 $\Phi(x)$ 在 U 上的投影为 $U^T \cdot \Phi(x) = [u^1, u^2, \dots, u^m]^T \cdot \Phi(x)$



$$\begin{aligned}
 &= [u^1 \cdot \Phi(x), u^2 \cdot \Phi(x), \dots, u^m \cdot \Phi(x)]^T \\
 &= \left[\sum_{i=1}^M \alpha_i^1 (\Phi(x_i) \cdot \Phi(x)), \sum_{i=1}^M \alpha_i^2 (\Phi(x_i) \cdot \Phi(x)), \dots, \right. \\
 &\quad \left. \sum_{i=1}^M \alpha_i^m (\Phi(x_i) \cdot \Phi(x)) \right]^T \\
 &= \left[\sum_{i=1}^M \alpha_i^1 k(x_i, x), \sum_{i=1}^M \alpha_i^2 k(x_i, x), \dots, \sum_{i=1}^M \alpha_i^m k(x_i, x) \right]^T \\
 &= [y_1, y_2, \dots, y_m]^T \\
 &= y
 \end{aligned} \tag{4-12}$$

y 就是测试样本 x 对应的非线性主元。这样我们只需要通过简单的计算核函数就能得到非线性主元，而不必要知道非线性映射 Φ 的具体形式。

在特征空间中变量的均值不为零时，可以通过如下方法进行转化，即

$$\sum_{k=1}^M \Phi(x_k) = 0 \tag{4-13}$$

如果令

$$\bar{\Phi}(x_i) = \Phi(x_i) - \frac{1}{M} \sum_{i=1}^M \Phi(x_i) \tag{4-14}$$

则训练样本 x_1, x_2, \dots, x_M 在高维特征空间中对应的映射 $\Phi(x_1), \Phi(x_2), \dots, \Phi(x_M)$ 的均值就转化为零了。此时式 (4-7) 对应的核函数矩阵相应地变为

$$\begin{aligned}
 \bar{K}_{i,j} &= \bar{k}(x_i, x_j) \\
 &= \bar{\Phi}(x_i) \cdot \bar{\Phi}(x_j) \\
 &= \left(\Phi(x_i) - \frac{1}{M} \sum_{p=1}^M \Phi(x_p) \right) \cdot \left(\Phi(x_j) - \frac{1}{M} \sum_{q=1}^M \Phi(x_q) \right) \\
 &= \Phi(x_i) \cdot \Phi(x_j) - \frac{1}{M} \sum_{p=1}^M \Phi(x_p) \cdot \Phi(x_j) - \frac{1}{M} \sum_{q=1}^M \Phi(x_q) \cdot \Phi(x_i) \\
 &\quad + \frac{1}{M^2} \sum_{p=1}^M \sum_{q=1}^M \Phi(x_p) \Phi(x_q) \\
 &= K_{i,j} - \frac{1}{M} \sum_{p=1}^M K_{p,j} - \frac{1}{M} \sum_{q=1}^M K_{i,q} + \frac{1}{M^2} \sum_{p=1}^M \sum_{q=1}^M K_{p,q}
 \end{aligned} \tag{4-15}$$

这样我们就可以从 $K_{i,j}$ 计算出 $\bar{K}_{i,j}$ 。

这样与式 (4-9) 相对应的特征方程为

$$M \bar{\lambda}^{\Phi} \bar{\alpha} = \bar{K} \bar{\alpha} \tag{4-16}$$

4.2.3 核函数及其参数选择

常用的核函数有以下几种^[14,16]。



1) 多项式核函数

$$k(x, z) = (b * \langle x, z \rangle + c)^d \quad (d \in N; b, c \geq 0) \quad (4-17)$$

2) 高斯 (Gaussian) 核函数

$$k(x, z) = \exp\left(-\frac{\|x - z\|^2}{\sigma^2}\right) \quad (4-18)$$

式中 $\sigma > 0$, 用来控制核函数的宽度。

3) Sigmoid 内积核函数

$$k(x, z) = \tanh(\alpha \langle x, z \rangle + \beta) \quad (4-19)$$

式中 $\alpha > 0, \beta < 0$, 分别用来控制核函数的斜率和偏移量。

4) 一般径向基 (RBF) 核函数

$$k(x, z) = \exp(-\rho d(x, z)) \quad (4-20)$$

式中 $\rho > 0$;

$d(x, z)$ ——可以是任意一种距离度量。

另外, 可以利用已知的核函数来构造新的核函数^[1,2,4]。

核函数选择的好坏将直接影响到核主元分析方法的性能。参数选择的原则主要有^[17]以下几条。

1) 如果关于训练集无任何先验知识, 则只能分别试用不同的核函数, 归纳误差最小的核函数就是最好的核函数。

2) 如果对于训练样本有一定的先验知识, 则所选用的核函数应该与被估计实值函数的功率谱的形状相匹配。

综合以上的讨论, KPCA 的算法可以归结如下^[13]:

- 1) 选定核函数 $K(x_i, x_j)$, 根据式 (4-7) 计算矩阵 K 。
- 2) 求解特征方程 (4-9), 并利用式 (4-11) 将特征向量归一化。
- 3) 对于测试样本, 利用式 (4-12) 求得其非线性主元。

4.3 基于核主元分析的驾驶员面部姿势估计

4.3.1 驾驶员面部姿势模型

人的面部姿势可以由其在 X 、 Y 、 Z 三个方上的旋转角度来合成, 驾驶员出现视觉分散时面部姿势的变化主要体现在绕 Y 轴与绕 Z 轴的旋转上^[1], 如图 4-3 所示。因此, 我们主要对这两个方向上的姿势变化, 即 α (面部俯仰角)、 β (面部横摆角) 进行估计, 并合成最后的驾驶员实际面部姿势。

4.3.2 图像样本采集

利用 KPCA 算法估计面部姿势首先需要建立标准姿势估计曲线, 因此需要

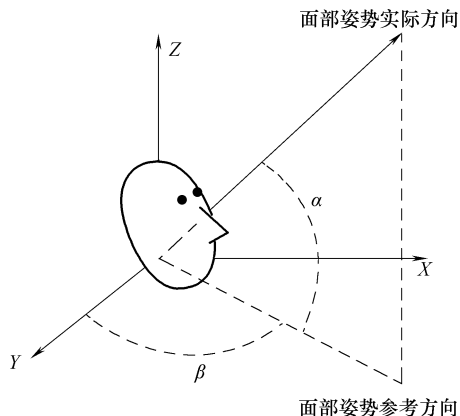


图 4-3 驾驶员面部姿势模型

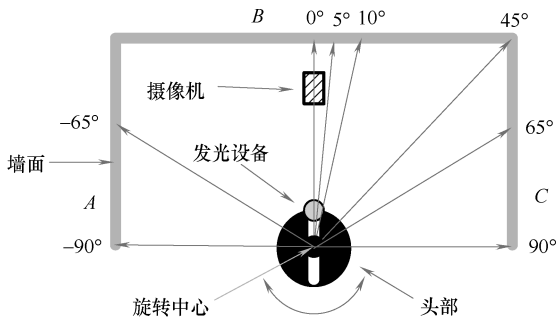
一组姿势角度已知的面部图像，将这些图像作为核主元分析的训练样本。为保证采集的面部图像对应一个比较精确的角度，我们自行设计了一套图像采集系统，如图 4-4 所示。

该采集系统需要在一个具有三面光滑墙面 A、B、C 的房间内进行，确定旋转中心，保证旋转中心到三面墙壁的垂直距离相等，该距离在我们系统中为 180 像素。采集图像时让被采集图像的人坐在椅子上，保证人的头部在旋转中心。在人面部固定一个教学用光笔，标定其发光方向与人脸朝向一致。以旋转中心向 B 墙面做垂线得到交点即为 0°对应点，然后根据几何关系每隔 5°在墙面上标记一点，标记点距地面的高度为 130cm。5°、10°等对应点与 0°对应点的水平距离为 15.75 像素、31.74 像素等。摄像机固定在人体前方一定距离处，轴线对准人面部中心。采集图像时让实验人员坐在椅子上保持身体基本不动，然后旋转头部，每次转动头部都使光源发出的光照射在墙上的标记点，这样每转动一次转过的角度为 5°。采集图像时让实验人员从 65°处逆时针转动头部到 -65°处，每转动一次采集一幅图像，共得到 27 幅图像，其中 0°时摄取的图像为正面人脸图像。

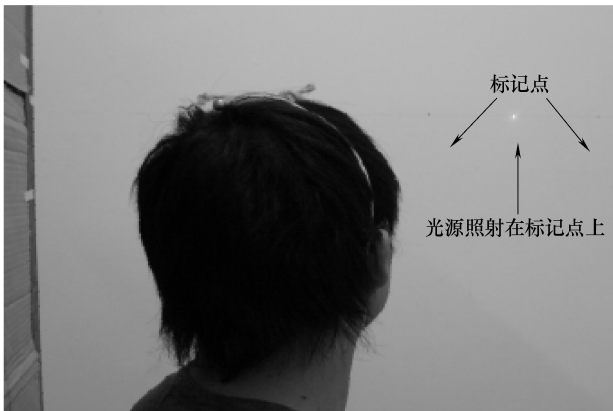
采用图 4-4 的方法采集 5 位实验人员不同姿势下的头肩图像，并利用第 3 章中的人脸定位算法，分割人脸，并归一化大小为 95 × 145 像素的面部图像。采集到的一位实验人员某姿势下的图像与归一化后的图像如图 4-5 所示。

4.3.3 建立姿势曲线

根据流形学习理论，不同姿势的高维人脸图像，具有一个潜在的低维流形结构^[15,16]。应用 KPCA 的方法可以把这种潜在流形结构嵌入到低维空间，实现高维数据的可视化。应用 KPCA 估计面部姿势主要分两步：



a) 图像采集示意图



b) 图像采集现场

图 4-4 图像样本采集系统

- 1) 建立姿势曲线。
- 2) 估计新面部图像的姿势角度。

由于人们不能直接感知高于 3 维的数据集，需要把高维人脸图像存在的流形嵌入到低于 3 维的空间。对于面部单纯的俯仰角、横摆角变化，它们实际上是一个平面（人脸）旋转变化的，因此把高维人脸图像存在的流形嵌入到一个 2 维空间可以揭示其变化的本质^[18]。

将图 4-5b 中获得已知面部姿势的图像作为核主元分析的训练样本，建立姿势曲线。首先，将获得的彩色图像转化为灰度图像，因为在计算机中灰度图像都由一个矩阵表示，则可按图 4-6 所示的方法把图像矩阵转化为列向量。如果每幅图像的大小是 100×100 像素，则转化后的向量是 $100 \times 100 = 10000$ 维的列向量。

选定核函数 $K(x_i, x_j)$ ，根据式 (4-7) 计算矩阵 K ，然后求解特征方程

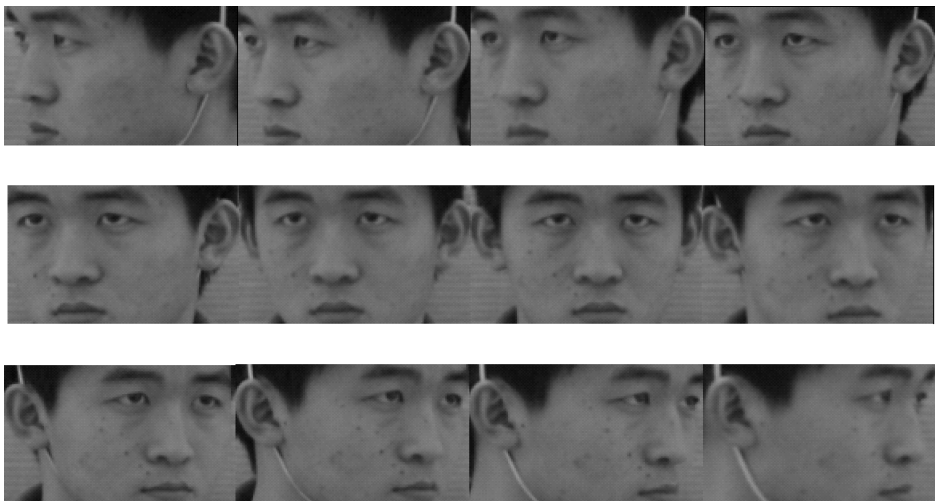


图 4-5 归一化后部分样本图像

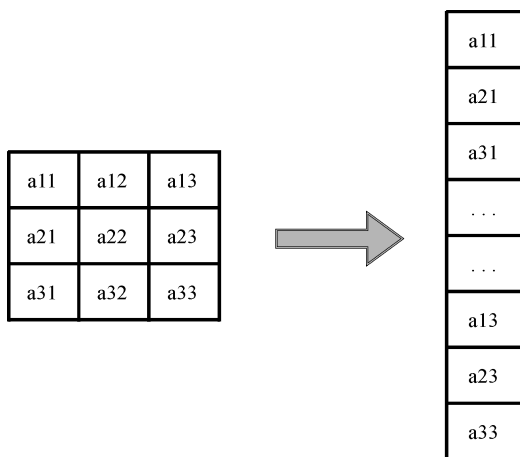


图 4-6 图像矩阵转化为列向量

(4-9)，并利用式 (4-11) 将特征向量归一化。假设得到特征值为 $A_1, A_2, A_3, \dots, A_n$ (从大到小)，它们对应的特征向量分别为 $X_1, X_2, X_3, \dots, X_n$ 。计算每个样本在 $X_1, X_2, X_3, \dots, X_n$ 上的投影，可以得到 M (训练样本数) 组 $(X_1^1, X_2^1, X_3^1, \dots, X_n^1)$ 、 $(X_1^2, X_2^2, X_3^2, \dots, X_n^2)$ 、 $(X_1^3, X_2^3, X_3^3, \dots, X_n^3)$ 、 $(X_1^M, X_2^M, X_3^M, \dots, X_n^M)$ 投影值。根据样本在任意两个特征向量上的投影值，就可以建立姿势曲线。应用最多的是前几个最大特征值对应的特征向量，其中最大特征值对应的特征向量称为主元，相比较其他特征向量它包含了更多的图像信息。但据主元建立的面部姿势曲线，



并不一定最适合估计面部姿势。以样本在 1、2、3、4 特征向量（前 4 个最大特征值对应的特征向量）上的投影值为例进行说明，建立姿势曲线如图 4-7 所示。

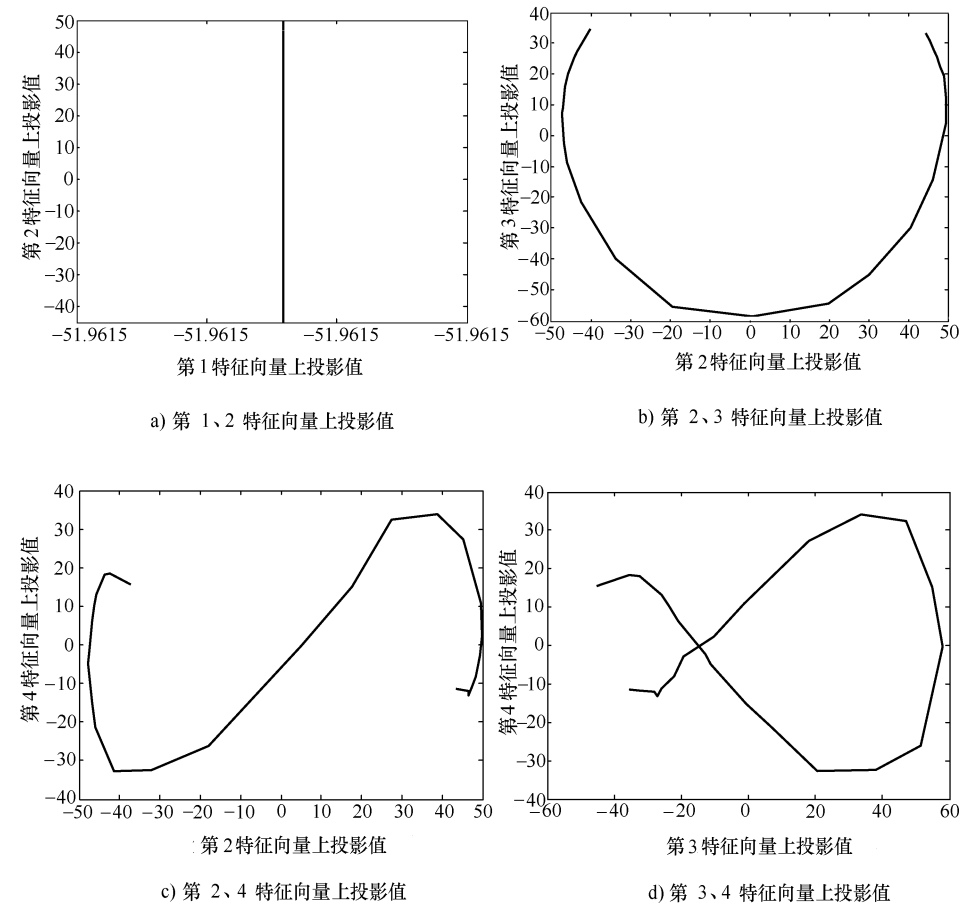


图 4-7 利用不同特征向量上的投影值建立姿势曲线

根据图 4-7a 曲线可以看出由于样本在第 1 特征向量上的投影为常数，因此第 1 特征向量上的投影值不能反映面部姿势变化的本质，不适合于估计面部姿势。图 4-7d 曲线存在交叉点，在估计姿势角度时在交叉点处容易引起歧义。图 4-7c 曲线虽然没有交叉，但若把曲线分为三段单调曲线的话，这三段曲线的横向距离太近。这些特征都不利于应用临近点估计新投影点的角度。综合起来看，只有图 4-7b 曲线最适合用来估计面部姿势角度，因为该曲线不仅光滑，而且左右对称，反映了高维面部图像存在的流形结构。因此本书应用样本在第 2、3 特征向量上的投影值来建立姿势曲线，并根据新样本在 2、3 特征向量上的投影值来估计其对应的角度。



在同一空间嵌入不同人的姿势曲线，如图4-8所示。图中可以看出不同人的姿势曲线虽然不能完全重合，但是姿势曲线的形状大体是相同的，左右基本上对称。并且，可以认为对称轴线通过姿势曲线上 0° 角（正面人脸）对应的点。每个人面部旋转的角度是沿某条姿势曲线逐渐变化的。根据不同个体姿势曲线大体相同的特点，可以应用同一个人的姿势曲线来估计不同人的面部姿势，而不至于引起较大误差。

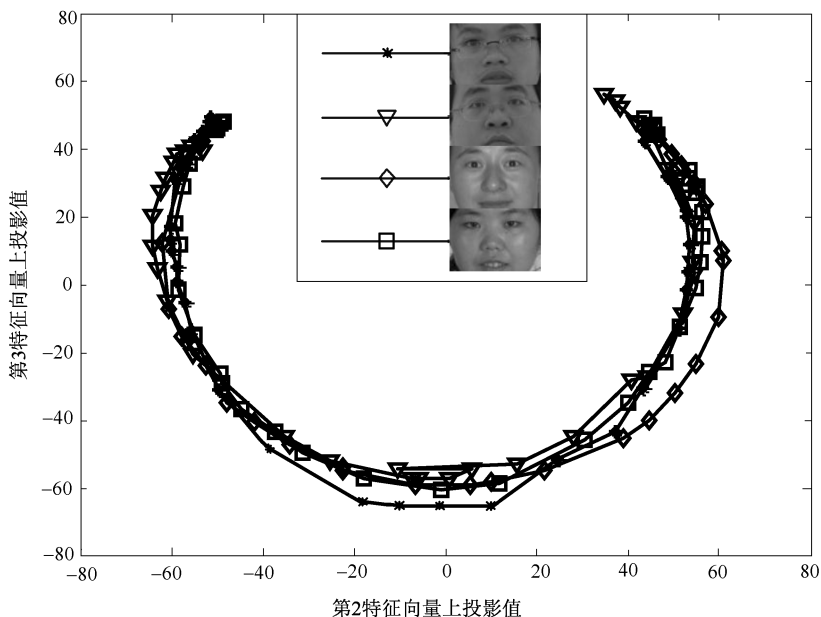


图 4-8 同一空间嵌入不同人的姿势曲线

4.3.4 面部姿势估计

根据图4-8可以看出，不同面部姿势图像在曲线上对应点的分布规律基本相同，呈现出中间稀疏，两端相对密集的特点。不同人的姿势曲线都大体呈圆形，根据姿势曲线利用最小二乘法拟合圆，发现它们的圆心相距比较近，并且半径的长度也比较接近。因此我们提出根据拟合圆心、新投影点及与新投影点最近两点之间的位置关系，估计新面部图像对应角度的方法。

建立如图4-9的坐标系，其中横轴代表图像在第2特征向量上的投影值，纵轴代表图像在第3特征向量上的投影值， O 为坐标原点。根据训练样本对应的点拟合圆，得到圆心为 $P(x_0, y_0)$ ，半径为 R ，则圆方程为

$$(x - x_0)^2 + (y - y_0)^2 = R^2 \quad (4-21)$$



假设有一测试样本图像 I ，计算其在 X_2 、 X_3 上的投影值，得到测试样本对应的坐标点 T ，然后根据 T 的邻近点 T_1 、 T_2 ，以及圆心 P 的坐标估计新坐标点 T 对应的角度。

求解的具体算法如下。

根据核主元分析算法计算测试样本图像在嵌入空间对应点的坐标 $T(x_3, y_3)$ 。

连接点 P 、 T 得到直线 PT

$$\frac{x - x_0}{y - y_0} = \frac{x - x_3}{y - y_3} \tag{4-22}$$

计算 PT 与拟合圆的交点，假设与 T 较近的交点为 J 。

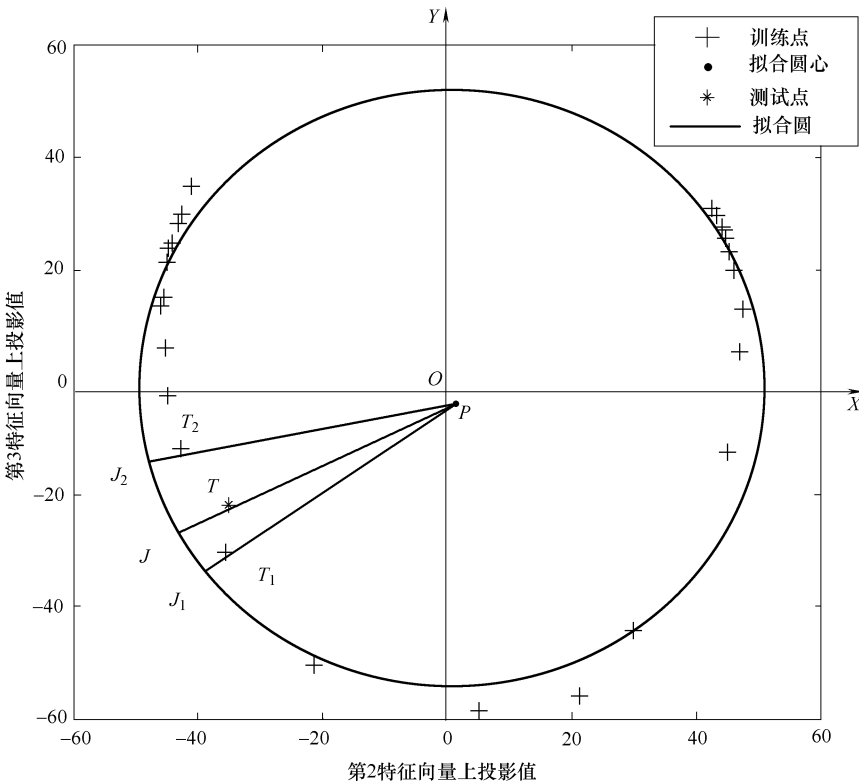


图 4-9 新图像角度估计

寻找距离 T 最近的训练点，假设为 $T_1(x_1, y_1)$ 、 $T_2(x_2, y_2)$ 。连接 P 与 T_1 、 T_2 点，得到直线 PT_1 、 PT_2

$$\frac{x - x_0}{y - y_0} = \frac{x - x_1}{y - y_1} \tag{4-23}$$

$$\frac{x - x_0}{y - y_0} = \frac{x - x_2}{y - y_2} \tag{4-24}$$



计算 PT_1 、 PT_2 与拟合圆的交点，找到与 J 最近的交点 J_1 、 J_2 ，则 T 点对应的角度应该为

$$\beta = (-5^\circ) \times \frac{\widehat{JJ_1}}{J_1J_2} + \beta_1 \quad (4-25)$$

式中 β_1 ——训练点 T_1 对应的角度；

$\widehat{JJ_1}$ 、 $\widehat{J_1J_2}$ ——拟合圆上 JJ_1 、 J_1J_2 的弧长。

因为图 4-9 所示的是新投影点落在左半平面时的情况，对应的姿势角度为负值，所以在 5° 前面加个负号；如果新投影点落在右半平面则对应角度为

$$\beta = 5^\circ \times \frac{\widehat{JJ_1}}{J_1J_2} + \beta_1 \quad (4-26)$$

在式 (4-25)、式 (4-26) 中 β_1 都是对应与 0° 角最近的已知点对应的角度，因为与 $\widehat{JJ_1}$ 、 $\widehat{J_1J_2}$ 相对的圆心角较小，在实际计算中可以用线段 JJ_1 、 J_1J_2 的长度代替弧长。

4.4 实例分析

4.4.1 面部横摆角估计

选择 1 位测试人员的样本图像作为训练样本，选择高斯核函数 ($\sigma = 10000$) 建立姿势估计曲线，并对其他 4 位测试人员的样本图像进行角度估计实验。4 位测试人员样本图像的估计精度（偏差绝对值）如图 4-10 所示。

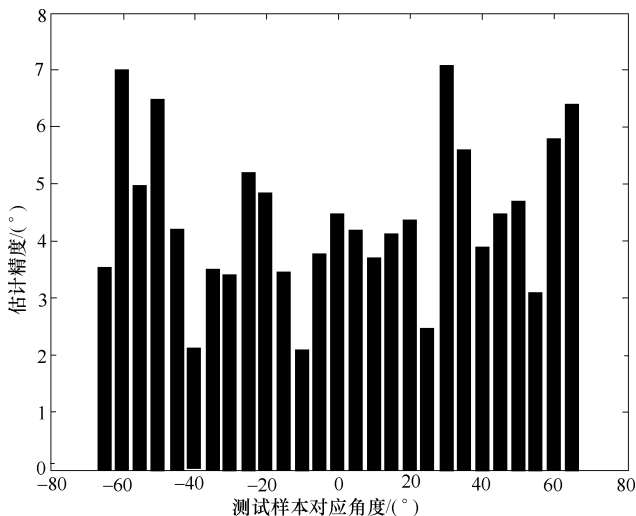


图 4-10 样本图像的估计精度



根据图 4-10 来看,虽然角度估计精度没有明显变化规律,但是在 $\pm 20^\circ$ 的范围内估计精度相对来说比较稳定,估计精度大体在 4° 左右,当面部旋转角度偏大时,估计精度稳定性有所降低,最大偏差达到 7° ,最小偏差只有 2° 左右。与其他现有方法相比,我们的估计方法具有一定优势。Fisher Manifold Learning (FML)^[19] 估计精度较高,但是估计范围较小,在 $[-10, 10]$ 度内估计精度在 3° 左右; KPVA + SVC^[20] 对间隔为 10° 图像的识别率为 97%; Isomap 与 LPP^[21,22] 的估计精度在 11° 与 15° 左右; LEA^[8] 方法的估计精度较高,但计算过于复杂,不适合于在检测驾驶员视觉分散时实际应用;几何分析方法的精度也较高,但是需要定位人脸的多个关键特征点,在多姿势人脸中定位特征点本身就比较困难,故不适合本书需要。

4.4.2 核函数对估计精度的影响

常用的核函数主要有 4 种,前面用到的核函数主要是高斯核函数。我们应用高斯核函数下不同参数建立的姿势曲线如图 4-11a 所示,应用不同的核函数建立的姿势曲线如图 4-11b 所示。

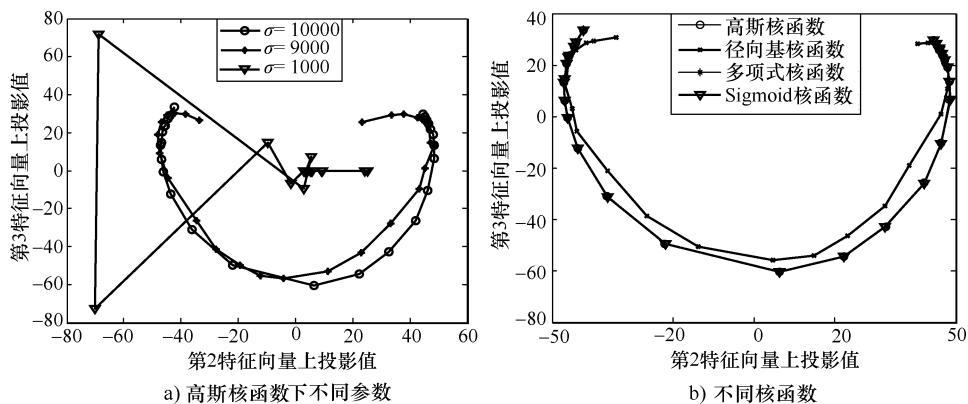


图 4-11 应用不同核函数、参数建立姿势曲线

从图 4-11a 中可以看出,对于同一个核函数,当选择的核函数参数不一致时建立的姿势曲线会有很大差别,有时甚至无法进行估计计算。如图 4-11a 中 3 条曲线是高斯核函数分别在 $\sigma = 10000$ 、 $\sigma = 9000$ 与 $\sigma = 1000$ 时作出的姿势曲线,在 $\sigma = 1000$ 时姿势曲线已经严重变形,多个点重合到一起以致无法进行角度估计。因为还没有一个选择核函数参数的依据,这就需要通过实验来选择核函数的参数,对于相同的训练样本,如果应用某核函数及参数能得到比较平滑姿势曲线,则认为该核函数及参数则被认为适合于估计面部姿势。

从图 4-11b 中可以看出,应用不同核函数建立的姿势曲线大体相同,曲线



大部分位置重合，只是应用径向基核函数建立姿势曲线时有较大差别，因此选择不同核函数对估计精度的影响不会太大。表 4-1 列出应用相同训练样本、相同测试样本而不同的核函数时平均估计精度。

表 4-1 不同核函数对估计精度影响

核 函 数	平均估计精度/(°)
高斯核函数 ($\sigma = 10000$)	4.58
多项式核函数 ($b = c = d = 1$)	4.66
径向基核函数 ($\rho = 10000$)	4.48
Sigmoid 核函数 ($\alpha = 2 \times 10^{-10}$, $\beta = -5$)	4.71

4.4.3 面部俯仰角估计

估计俯仰角的方法类似于估计横摆角，首先建立俯仰角姿势曲线，然后将测试图像投影到姿势曲线所在空间，再估计新图像对应的角度。

采用类似于图 4-4 的图像采集方法采集样本图像，只不过把水平方向上的距离变化改变为垂直方向上的距离变化。受人体生理结构影响，人脸面部俯仰角的自由变化范围较小，实验采集样本图像的变化范围是 $[-30^\circ, 30^\circ]$ ，相邻样本间的角度间隔为 5° 。

采用高斯核函数 ($\sigma = 10000$)，根据图像在第 2、3 特征向量上的投影建立姿势估计曲线如图 4-12a 所示。根据图 4-12a 建立的姿势曲线估计其他测试人员样本图像的角度，样本图像估计的平均精度（偏差绝对值）如图 4-12b 所示。根据图 4-12b 我们可以看出估计精度在 5° 左右，估计精度低于横摆角的估计精度，这可能与训练的数量有关，同样 0° 左右样本的估计精度较高。

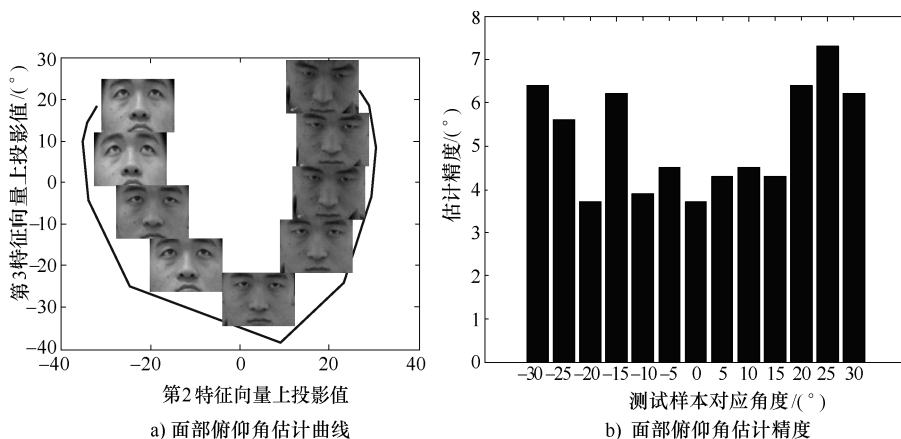


图 4-12 面部俯仰角估计曲线及估计精度



4.5 本章小结

根据面部姿势可以检测出大部分的驾驶员视觉分散行为。本书作者提出利用核主元分析把高维面部图像存在的流形结构嵌入到二维姿势空间,利用训练样本建立姿势曲线并拟合圆,然后把新图像投影到相同空间,利用拟合圆心及姿势曲线上与新投影点最近的2个点,估计新投影图像对应的角度。利用核主元分析算法可以把不同人的姿势曲线嵌入到同一空间,并能保持各个姿势曲线原来本质。因此,利用同一条姿势曲线可以估计不同人的面部姿势而不致引起较大误差。实验表明我们提出的估计算法效果良好,可以满足一定实际需求。

参考文献

- [1] J. Wu, M. M. Trivedi. A two-Stage Head Pose Estimation Framework and Evaluation [J]. Pattern Recognition. 2008, 41 (3): 1138-1158.
- [2] N. J. Emery. The Eyes Have it: The Neuroethology, Function and Evolution of Social Gaze [C]. Neuroscience and Biobehavioral Reviews. 2000, 24: 581-604.
- [3] K. Huang, M. M. Trivedi. Driver Head Pose and View Estimation with Single Omni directional Video Stream [C]. Proceedings of the first International Workshop on In-Vehicle Cognitive Computer Vision Systems, in Conjunction with the Third International Conference on Computer Vision Systems, Seattle, USA, 2003: 709-714.
- [4] B. Braathen, M. S. Bartlett, J. R. Movellan. 3-D Head Pose Estimation from Video by Stochastic Particle Filtering [C]. the Eighth Annual Joint Symposium on Neural Computation, 2001: 1-7.
- [5] 王珂,尹宝才,王雁来. 人脸特征跟踪和头部姿势估计(英文)[J]. 北京工业大学学报. 2005, 31 (2): 220-224.
- [6] 毋立芳,张斯聪,赵晓晴,等. 一种人脸姿势估计新方法[J]. 信号处理, 2006, 22 (1): 61-64.
- [7] Y. Wei, L. Fradet, T. Tan. Head Pose Estimation Using Gabor Eigenspace Modeling [C]. IEEE International Conference on Image Processing (ICIP2002), Rochester, New York, USA, 2002: 281-284.
- [8] Y. Fu, T. S. Huang. Graph Embedded Analysis for Head Pose Estimation [C]. Proceedings of the Seventh International Conference on Automatic Face and Gesture Recognition, Southampton, UK, 2006: CD-Room.
- [9] 陈家大,赖剑煌,冯国灿. 一种人脸姿势判别与正脸合成的新方法[J]. 计算机研究与发展. 2006, 43 (8): 1477-1484.
- [10] Y. Li, S. Gong, H. Liddell. Support Vector Regression and Classification Based Multi-View





- Face Detection and Recognition [C]. IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France, 2000: 300-305.
- [11] S. Srinivasan, K. L. Boyer. Head Pose Estimation Using View Based Eigenspaces [C]. the 16th International Conference on Pattern Recognition, Quebec, Canada, 2002: 302-305.
- [12] 史明光. 蛋白质相互作用预测方法的研究 [D]. 合肥: 中国科学技术大学. 2009.
- [13] 张燕昆. 基于核方法的人脸识别技术的研究 [D]. 上海: 上海交通大学. 2003.
- [14] 李学勇, 路长厚, 李建美. 基于重构误差的单类子空间压印凹凸字符识别 [J]. 机械工程学报, 2008, 44 (5): 181-185.
- [15] 黄国宏, 邵惠鹤. 核主元分析及其在人脸识别中的应用 [J]. 计算机工程, 2004, 30 (13): 13-14.
- [16] S. T. John, C. Nello. Kernel Methods for Pattern Analysis [M]. Cambridge: Cambridge University Press, 2004.
- [17] 孔锐. 基于核的学习方法及其在人脸识别中的应用研究 [D]. 合肥: 中国科学技术大学, 2004.
- [18] N. Hu, W. Huang, S. Ranganath. Head Pose Estimation by Non-Linear Embedding and Mapping [C]. IEEE International Conference on Image Processing, Singapore, 2005: CD-ROM.
- [19] L. Chen, L. Zhang, Y. Hu, et al. Head Pose Estimation Using Fisher Manifold Learning [C]. in Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, in Conjunction with ICCV2003, 2003: 203-207.
- [20] S. Z. Li, Q. D. Fu, L. Gu, et al. Kernel Machine Based Learning for Multi-View Face Detection and Pose Estimation [C]. in Proceedings of Eighth IEEE International Conference on Computer Vision, 2001: 674-679.
- [21] B. Raytchev, I. Yoda, K. Sakaue. Head Pose Estimation by Nonlinear Manifold Learning [C]. IEEE Conf. on ICPR04, 2004: 462-466.
- [22] 王笑. 基于支持向量机分类的步态识别方法研究 [D]. 西安: 西安电子科技大学. 2007.

基于Multi-PCA的驾驶员 眼睛视线方向识别

5.1 概述

根据驾驶员面部姿势，可以识别出大部分的注意分散行为^[1]。但在驾驶员面部姿势朝前时，驾驶员的眼睛也可能会观察其他方向，根据第3章的分析可知，这也是一种注意分散行为。因此，还需要对驾驶员的眼睛视线方向进行识别。眼睛视线方向识别的应用可以分为诊断应用与交互应用^[2]。尽管视线已经成功地作为一种计算机输入方式，但一般需要在严格控制的外部环境下应用，成功应用也只限于武器操纵以及残障人士服务上。

传统的眼睛视线识别方法大多是侵入性的，现在研究方法大多是非侵入性的^[3,4]。非侵入性的方法一般是基于机器视觉的，应用摄像机捕捉眼部图像，再根据图像分析视线特性。瞳孔与角膜的反光点及位置信息常被用来估计视线方向。但是要设定很多假设条件，并且这些假设一般都不是真实的^[5]。这种方法需要精确定位瞳孔与反光点。为避免精确定位这些特征点，Baluja^[6]等提出应用神经网络的方法估计视线方向。Perrone^[7]应用虹膜或者瞳孔在图像中的形状特征来估计视线方向，因为虹膜与瞳孔都是圆形的，圆状物体在以不同角度向平面投影时，会形成不同的椭圆形投影，根据椭圆的特征估计视线方向。这种方法的缺点是需要较高分辨率的眼睛区域图像。

可靠、准确、无接触地识别视线方向是比较困难的，特别是在驾驶环境下，驾驶员不能佩戴任何额外设备。因此，为满足检测驾驶员视线方向的需要，Bebis^[8]把视线方向分为5种（前、上、下、左、右），并提出采用PCA（Principal Component Analysis——主元分析）的方法来识别视线方向。主元分析是把所有不同类型的样本放在一起进行特征提取，以所有类型样本的最优重建为目的，因此对于描述不同类型样本之间的差异而言它不一定是最优的描述。从这个意义上说用它来描述人眼视线识别的特征是不充分的，因而限制了PCA识别率的进一步提高^[9]。

为此，本书作者提出了一种改进的眼睛视线方向识别方法，为不同视线情况下的人眼图像各建立一特征空间，在每个特征空间下重建待识别人眼图像，并计



算重建误差, 根据重建误差来设计分类器。特征空间的构造是以某类人眼图像的最优重建为目的, 用来重建同类图像时可以获得最小的重建误差, 而重建其他类图像的误差将会较大, 将测试图像分类到能获得最小重建误差的特征空间。这样在识别样本图像时, 需要进行多次 PCA 运算, 可以称之为 Multi-PCA^[10]。

5.2 Multi-PCA

PCA (主分量分析), 也称 K-L 变换^[11], 它是一种基于统计特征的最佳正交变换。

5.2.1 K-L 变换

设有一个 n 维的向量 \mathbf{x} , 欲将其转换为 l 维 ($l < n$) 的向量 \mathbf{y} 以简化计算, 即

$$\mathbf{x} = (x_1, x_2, \dots, x_l, \dots, x_n) \rightarrow \mathbf{y} = (x_1, x_2, \dots, x_l) \quad (5-1)$$

最简单的方法就是将向量 \mathbf{x} 直接截断, 即

$$\mathbf{x} = (x_1, x_2, \dots, x_l, \dots, x_n) \rightarrow \mathbf{y} = (x_1, x_2, \dots, x_l, 0, 0 \dots, 0) \quad (5-2)$$

此时的均方误差为

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{x_l^2 + x_{l+1}^2 + \dots + x_n^2} \quad (5-3)$$

从式 (5-3) 中可以看出, 转换误差是非常大的。而 PCA 是以转换前后的均方根误差最小为出发点的。

设向量 \mathbf{x} 的均值向量为 $\bar{\mathbf{x}} = E(\mathbf{x})$, 其相关矩阵为 $\mathbf{R}_x = E[\mathbf{x}\mathbf{x}^T]$, 协方差矩阵为 $\mathbf{C}_x = E((\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})')$, \mathbf{x} 经正交变换后得向量 $\mathbf{y} = (y_1, y_2, \dots, y_l, \dots, y_n)'$, 即

$$\mathbf{y} = T\mathbf{x} = (t_1 t_2 \dots t_n)' \mathbf{x} = \begin{pmatrix} t'_1 \\ t'_2 \\ \dots \\ t'_n \end{pmatrix} \mathbf{x} \quad (5-4)$$

从而 \mathbf{y} 的各个分量表示为

$$y_i = \mathbf{t}'_i \mathbf{x}, i = 1, 2, \dots, n \quad (5-5)$$

而

$$\mathbf{x} = (T')^{-1} \mathbf{y} = T\mathbf{y} = (t_1 t_2 \dots t_n) \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \sum_{i=1}^n y_i \mathbf{t}_i \quad (5-6)$$



在式 (5-6) 中选择 m 项, 则在最小均方差准则下的线性估计 $\hat{\mathbf{x}}$ 为

$$\hat{\mathbf{x}} = \sum_{i=1}^m y_i \mathbf{t}_i \quad (5-7)$$

式中 $1 \leq m < n$ 。

此时的均方误差为

$$\begin{aligned} \varepsilon^2(m) &= E[(\mathbf{x} - \hat{\mathbf{x}})'(\mathbf{x} - \hat{\mathbf{x}})] = \sum_{i=m+1}^n E[y_i^2] = \sum_{i=m+1}^n E[y_i y_i'] \\ &= \sum_{i=m+1}^n \mathbf{t}_i' E[\mathbf{x} \mathbf{x}'] \mathbf{t}_i = \sum_{i=m+1}^n \mathbf{t}_i' \mathbf{R}_x \mathbf{t}_i \end{aligned} \quad (5-8)$$

向量 \mathbf{x} 通过正交变换得到向量 \mathbf{y} , 我们希望变换前后的均方误差 $\varepsilon^2(m)$ 最小。可以在 \mathbf{T} 为标准正交阵的约束下 (即 $\mathbf{t}_i' \mathbf{t}_i = \delta_{ij}$) 构造准则函数

$$J = \sum_{i=m+1}^n \mathbf{t}_i' \mathbf{R}_x \mathbf{t}_i - \sum_{i=m+1}^n \lambda_i (\mathbf{t}_i' \mathbf{t}_i - 1) \quad (5-9)$$

使 $\frac{\partial J}{\partial \mathbf{t}_i} = \mathbf{0}$, 可得 $(\mathbf{R}_x - \lambda_i \mathbf{I}) \mathbf{t}_i = \mathbf{0}$, $i = m+1, \dots, n$, 即

$$\mathbf{R}_x \mathbf{t}_i = \lambda_i \mathbf{t}_i, i = m+1, \dots, n \quad (5-10)$$

式 (5-10) 表明, λ_i 是 \mathbf{x} 相关矩阵 \mathbf{R}_x 的特征值, \mathbf{t}_i 是 \mathbf{R}_x 对应的特征向量。

将式 (5-10) 代入 (5-8) 式中, 可得

$$\varepsilon^2(m) = \sum_{i=m+1}^n \mathbf{t}_i' \lambda_i \mathbf{t}_i = \sum_{i=m+1}^n \lambda_i \quad (5-11)$$

在 \mathbf{x} 的估计式中, 如果保留前 m 个 y_i , 而余下的 $(n-m)$ 个分量 y_i 分别由预设的 $(n-m)$ 个常数 b_i 代替, 此时的估计式为

$$\hat{\mathbf{x}} = \sum_{i=1}^m y_i \mathbf{t}_i + \sum_{i=m+1}^n b_i \mathbf{t}_i \quad (5-12)$$

估计的均方误差为

$$\varepsilon^2(m) = E[(\mathbf{x} - \hat{\mathbf{x}})'(\mathbf{x} - \hat{\mathbf{x}})] = \sum_{i=m+1}^n E[(y_i - b_i)^2] \quad (5-13)$$

最佳的 b_i 可以通过计算 $\partial \varepsilon / \partial b_i$ 求得, 此时

$$b_i = \mathbf{t}_i' \bar{\mathbf{x}} \quad (5-14)$$

此时估计的均方误差为

$$\varepsilon^2(m) = \sum_{i=m+1}^n \mathbf{t}_i' \mathbf{C}_x \mathbf{t}_i \quad (5-15)$$

下面求在 \mathbf{T} 为标准正交阵约束下 (即 $\mathbf{t}_i' \mathbf{t}_i = \delta_{ij}$) 使式 (5-15) 取得最小值的转换向量 \mathbf{t}_i , 仿照前面的做法, 构造准则函数

$$J = \sum_{i=m+1}^n [\mathbf{t}_i' \mathbf{C}_x \mathbf{t}_i - \lambda_i (\mathbf{t}_i' \mathbf{t}_i - 1)] \quad (5-16)$$



使 $\frac{\partial J}{\partial \mathbf{t}_i} = \mathbf{0}$, 可得

$$\mathbf{C}_x \mathbf{t}_i = \lambda_i \mathbf{t}_i, \quad i = m+1, \dots, n \quad (5-17)$$

如果将式 (5-17) 代入式 (5-15), 则

$$\varepsilon^2(m) = \sum_{i=m+1}^n \lambda_i \quad (5-18)$$

由线性代数的有关知识知, 因为 $\mathbf{R}_x - \mathbf{C}_x$ 为非负定阵, 故有

$$\lambda_i(\mathbf{R}_x) \geq \lambda_i(\mathbf{C}_x), \quad i = 1, 2, \dots, n \quad (5-19)$$

式中 $\lambda_i(\mathbf{R}_x)$ 和 $\lambda_i(\mathbf{C}_x)$ ——分别表示 \mathbf{R}_x 和 \mathbf{C}_x 的大的特征值。

因此, 可以得出结论, 第一种估计的均方误差比第二种要大。

5.2.2 PCA 实现方法

假设有 l 类训练样本, 每类训练样本的个数分别为 c_1, c_2, \dots, c_l , 则训练样本总数为 $c_1 + c_2 + \dots + c_l = M$ 。

若每个样本的大小都为 $a \times b$, 则这 M 个样本可以表示成 $a \times b$ 维列向量 $\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \dots, \mathbf{\Gamma}_M$, 并且可以求出 M 个列向量的均值

$$\mathbf{\Psi} = \frac{1}{M} \sum_{i=1}^M \mathbf{\Gamma}_i \quad (5-20)$$

每个向量 $\mathbf{\Gamma}_i$ ($i=1 \dots M$) 与均值向量的差为

$$\mathbf{\Phi}_i = \mathbf{\Gamma}_i - \mathbf{\Psi} \quad (5-21)$$

那么训练样本的协方差矩阵

$$\mathbf{C} = \frac{1}{M} \sum_{i=1}^M \mathbf{\Phi}_i \mathbf{\Phi}_i^T = \mathbf{A} \mathbf{A}^T \quad (5-22)$$

式中 $\mathbf{A} = [\mathbf{\Phi}_1, \mathbf{\Phi}_2 \dots \mathbf{\Phi}_M]$ 。

利用 \mathbf{C} 的正交特征向量建立样本特征空间。对于 N 维样本向量, 协方差矩阵 \mathbf{C} 的大小为 $N \times N$ 维, 因为 N 一般比较大, 对 \mathbf{C} 求特征值和特征向量很困难。我们需要找到一个简便的计算方法。根据奇异值分解原理, 我们可以求较小的 $M \times M$ 矩阵 \mathbf{L} 的特征值 λ 与特征向量 \mathbf{v} , 将得到的特征向量线性变化后可得 \mathbf{C} 的特征向量^[12]。

首先计算 $M \times M$ 矩阵 \mathbf{L}

$$\mathbf{L} = \mathbf{A}^T \mathbf{A} \quad (5-23)$$

的特征向量 \mathbf{v}_l ($l=1 \dots M$)。

矩阵 \mathbf{C} 的特征向量 \mathbf{u}_l ($l=1 \dots M$) 可由差值向量 $\mathbf{\Phi}_i$ ($i=1 \dots M$) 与 \mathbf{v}_l ($l=1 \dots M$) 的线性组合得到

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2 \dots \mathbf{u}_M] = [\mathbf{\Phi}_1, \mathbf{\Phi}_2 \dots \mathbf{\Phi}_M] [\mathbf{v}_1, \mathbf{v}_2 \dots \mathbf{v}_M] = \mathbf{A} \mathbf{V} \quad (5-24)$$



式中 U —— C 特征向量的矩阵;

L 的特征值 λ 即为 C 的特征值 λ 。

样本向量的特征主要体现在前面最大的几个特征值对应的特征向量上, 再根据重构需要达到的精度与计算复杂度, 选择最大的 p 个特征值对应的特征向量, 正交归一化, 以此为基础构成特征空间。然后, 把样本向量投影到特征空间中, 都会得到一个投影坐标。而根据这个投影坐标, 可以实现样本的重构。

5.2.3 Multi-PCA 算法

Multi-PCA 算法就是对待识别样本进行多次 PCA 运算, 具体算法概括如下。

1. 构建特征空间

1) 分别为每类选择训练样本, 并转换为列向量, 即

$$L_k = [\Gamma_1, \Gamma_2, \dots, \Gamma_{M_k}] \quad (5-25)$$

式中 k ——第 k 类训练样本, $k=1, 2, \dots, l$;

M_k ——第 k 类样本的训练样本个数。

2) 利用下式实现样本标准化

$$\Phi = \Gamma - \Psi^k \quad (5-26)$$

式中 Ψ^k ——第 k 类样本的均值。

3) 求各类特征空间的协方差矩阵 C^k ; 并对 C^k 进行特征值分解, 分别求出各类子空间的主分量 $u_1^k, u_2^k, \dots, u_{m_k}^k$, 其中 m_k 是第 k 个子空间所采用的主分量个数; 则由主分量构建的各类样本的特征空间为

$$U^k = [u_1^k, u_2^k, \dots, u_{m_k}^k] \quad (5-27)$$

2. 样本识别

1) 通过式 (5-26) 对测试样本标准化。

2) 将标准化后样本 Φ 向各类特征空间投影

$$\Omega^k = (U^k)^T \Phi \quad (5-28)$$

则 $\Omega^k = [w_1, w_2, \dots, w_{m_k}]^T$ 是测试样本在各个特征空间下的投影坐标。

3) 求测试样本在各类特征空间中的重构样本

$$\hat{\Phi}^k = \sum_{i=1}^{m_k} w_i u_i \quad (5-29)$$

4) 计算样本在每个特征空间中的重构误差

$$e_d^k = \|\Phi - \hat{\Phi}^k\| \quad (5-30)$$

使误差值最小的特征空间 k 就是测试样本所应属的类别空间。



5.3 基于 Multi-PCA 的眼睛视线方向识别

5.3.1 驾驶员眼睛区域定位

在本书第3章中所述驾驶员面部图像精确定位的基础上,本书作者提出采用联合投影函数与角点检测相结合的方法定位眼睛区域。

对精定位人脸进行水平方向的微积分联合投影,如图5-1a所示。图中可以清楚看到,眉毛、眼睛都对应一投影的峰值区,这样以眉毛区域下边缘作为眼睛区域上边界。在眼睛区域的下边缘,因为灰度值的变化也形成一个小的峰值,以该峰值为眼睛区域下边界,定位眼睛区域上下边界,如图5-1a所示。对该区域进行肤色检测,消除左右两侧头发对后续边缘、角点检测工作的影响,精简结果如图5-1b所示。

眼角是眼睛区域中比较稳定的特征点,随年龄、表情、胖瘦等改变发生的变化很小,应用 Susan 算子可以准确检测出眼睛的内外角点^[13]

$$c(\vec{r}, \vec{r}_0) = \begin{cases} 1 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| \leq t \\ 0 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| > t \end{cases} \quad (5-31)$$

$$n(\vec{r}_0) = \sum_{\vec{r}} c(\vec{r}, \vec{r}_0) \quad (5-32)$$

$$R(\vec{r}_0) = \begin{cases} g - n(\vec{r}_0) & \text{if } n(\vec{r}_0) < g \\ 0 & \text{others} \end{cases} \quad (5-33)$$

式中 \vec{r}_0 ——当前像素点;

\vec{r} ——区域范围内的其他点;

$I(\vec{r})$ ——该像素灰度值;

$c(\vec{r}, \vec{r}_0)$ ——两个像素灰度值的相关性度量;

t ——像素间差异阈值,一般设为27;

$n(\vec{r}_0)$ ——USAN 面积,该面积越小,则边缘强度越大;

g ——几何阈值,可以理解为 USAN 面积的最大值,超过这个值就认为

该点不是边缘,通常设定为 $\frac{3}{4}n(\vec{r}_0)_{\max}$ 。

用以上算法获得的眼部区域边缘图像,如图5-1c所示。提取图中黑色边缘曲线的角点,即可获得准确的两个外眼角点。以眼睛的两个外眼角点为眼睛区域左右边界,并结合眼睛区域的上下边界,定位眼睛区域如图5-1d所示,并分割该区域。

利用双线性插值算法把分割的眼睛区域归一化为 45×150 像素的图像,如



图 5-2 所示。

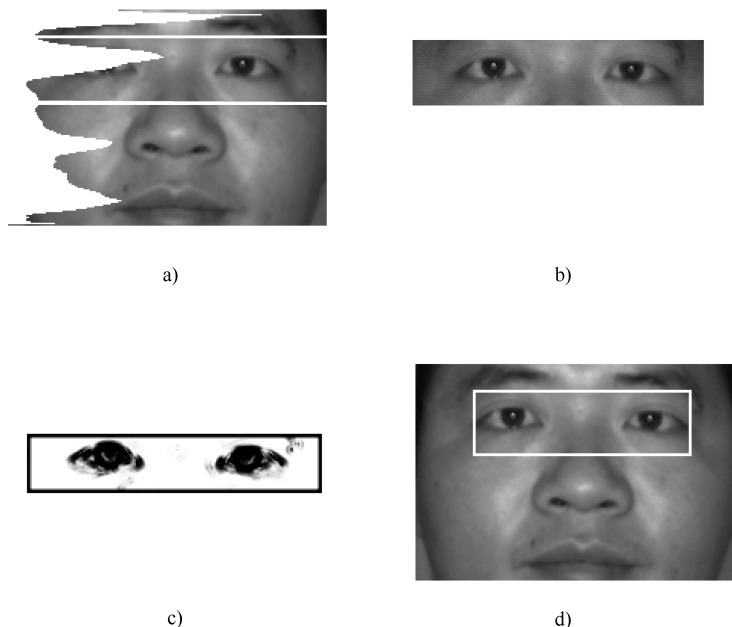


图 5-1 眼睛区域定位

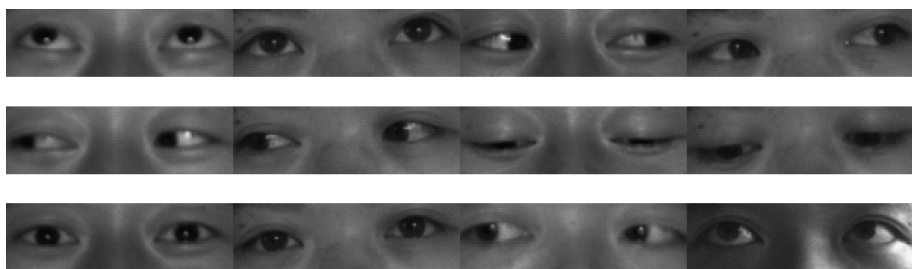


图 5-2 眼睛区域图像归一化效果

5.3.2 建立特征空间

选择 25 位实验者，摄取他们在观察 5 个不同方向时的面部图像，并分割眼睛区域图像，每种视线方向上共获得 25 幅图像。根据视线方向分类（上—U，下—D，左—L，右—R，前—S），需要建立 5 个特征空间。每类选择 10 幅眼睛图像，按照图 4-2 所示方法转换成列向量。利用训练样本构造协方差矩阵，根据 5.2.2 小节给出的方法计算矩阵的特征向量和特征值，选择最大的 p 个特征值





对应的特征向量构成 p 维特征空间，即为该视线方向人眼图像的特征空间^[16]。

确定降维空间（特征空间）的维数（ p ）是一个很重要的问题，采用下述的方法来确定特征空间的维数^[14,15]：

- 1) 将特征值按照从小到大顺序排列 $\lambda_1 > \lambda_2 > \dots > \lambda_M$ 。
- 2) 通过下面两式来确定特征空间的维数 p

$$\sum_{i=1}^M \lambda_i / \sum_{i=1}^p \lambda_i > \varepsilon_1 \quad (5-34)$$

$$\lambda_{p+1} / \lambda_p < \varepsilon_2 \quad (5-35)$$

式中 M ——原始空间的维数；

ε_1 、 ε_2 ——位于 0 - 1 之间的系数。

式 (5-34) 确定了降维后特征空间能量占原始空间能量的比例；而式 (5-35) 保证在逼近误差差别不大的情况下尽量降低特征空间的维数。通常情况下， ε_1 通过指定数值或实验来确定，本书取 $\varepsilon_1 = 0.99$ ；而 ε_2 通过观察特征值的分布情况或实验来确定，本书取 $\varepsilon_2 = 0.6$ 。

5.3.3 测试重构误差及特征空间相似度

既然 Multi-PCA 以最小重建误差来分类样本，我们需要测试样本图像在不同特征空间下的重构误差。图 5-3 是图 5-2 中 5 种不同视线方向样本图像在 5 个不同特征空间下的重构误差。横坐标表示 5 个特征空间，纵坐标分别表示重构误差。每条曲线上的标记点代表一个样本图像在不同特征空间下的重构误差，可

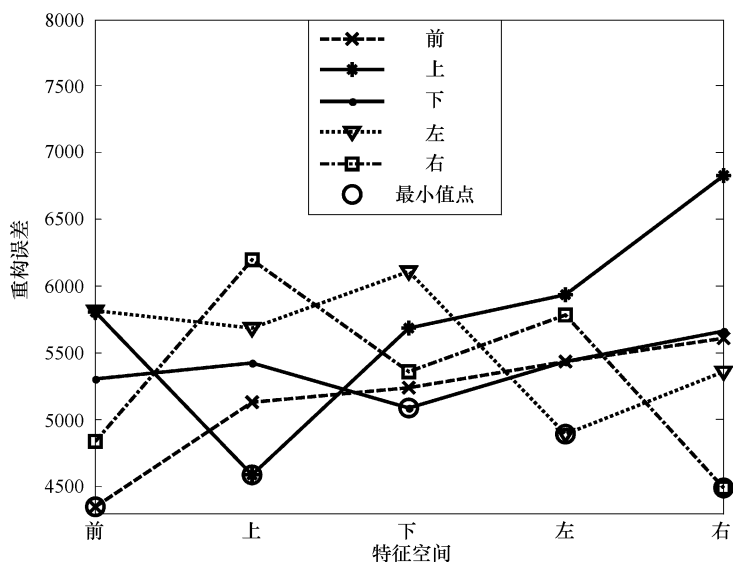


图 5-3 不同样本图像在不同特征空间下的重构误差



以看出每类样本只在其自身对应的特征空间中具有最小重构误差。图 5-3 中每一列标记点可以当成是同一特征空间重构同一样本时的重构误差，因此我们还可以看出，每一特征空间对应的重构误差也只有在重构本类样本时具有最小值。这两个特征可以说明 Multi-PCA 可以有效地识别视线方向。

另一个可以说明算法有效性的指标是特征空间相似度，它是衡量特征空间相似性的重要指标，常以特征空间距离来衡量特征空间的相似度。Sun^[17]等在分析以往空间距离方法的基础上提出了一种新的空间相似度定义方法，并在基于 Bayes 的人脸识别上得到成功应用。文中采用 Sun^[17]的子空间相似度定义，对不同眼睛视线方向的特征空间进行测试。

设 U 和 V 分别为两个特征空间，其维数分别为 m 和 n ； u_1, u_2, \dots, u_m 和 v_1, v_2, \dots, v_n 分别是特征空间 U 和 V 的正交基。则两个特征空间的距离定义为

$$d(U, V) = \sqrt{\max(m, n) - \sum_{i=1}^m \sum_{j=1}^n (u_i^T v_j)^2} \quad (5-36)$$

两个特征空间的相似度定义为

$$\omega_1 = d(U, V) / \sqrt{\max(m, n)} \quad (5-37)$$

当且仅当 $\omega_1 \leq 0.5$ 时，称两个特征空间相似。 ω_1 越小表示两个特征空间的相似度越高，为表达起来更直观我们将特征空间相似度定义为

$$\omega = 1 - d(U, V) / \sqrt{\max(m, n)} \quad (5-38)$$

这样在 ω 越大时表明两个特征空间的相似度越高，当且仅当 $\omega > 0.5$ 时，称两个特征空间相似。

根据子空间相似度的定义，对 5 类视线特征空间彼此间的相似度进行了测试，相似度计算结果如表 5-1。从表 5-1 中可以看出，各类视线的特征空间除和其自身空间相似度高外，同其他空间的相似度都很低，说明文中方法适合识别视线方向。

表 5-1 特征空间之间的相似度

特征空间	相似度 (%)				
	前	上	下	左	右
前	100				
上	30	100			
下	31	23	100		
左	29	29	28	100	
右	34	25	31	25	100



5.3.4 Multi-PCA 与 PCA 比较与分析

1. 识别效果比较

对于 Multi-PCA 利用 5.3.2 小节构造的特征空间,对剩余样本进行识别实验。而 PCA 法用 5 类视线方向的各 10 个样本,共 50 个训练样本构建一个特征空间,其余样本当成测试样本。随着特征向量个数 p 选择的不同,识别率也不相同,在此都选择最大识别率进行比较。除了识别率,对识别系统来说,另一个重要的指标是实时性,因此也对识别时间进行比较。利用配置为 Intel Pentium Dual E2180 2.00GHz CPU,2G 内存的计算机对两种方法进行实验对比,实验结果如表 5-2:

表 5-2 识别效果对比

	训练时间/s	单样本识别时间/s	最高识别率 (%)
PCA	3.375	0.062	84
Multi-PCA	1.53	0.078	89.33

从表 5-2 中可以看出,Multi-PCA 的训练时间还不到 PCA 训练时间的 50%,这是因为在建立特征空间时 Multi-PCA 需要进行特征值分解的矩阵 $[(45 \times 150) \times (10 \times 1)]$ 要比 PCA 特征值分解的矩阵 $[(45 \times 150) \times (10 \times 5)]$ 小得多,所以尽管进行特征值分解的矩阵数量多 4 个,但是训练样本的总时间还是要比 PCA 方法少;而在识别阶段,由于 Multi-PCA 要将待识别的样本向多个特征空间中进行投影,所以识别时间要比 PCA 法多。从最高识别率来看,PCA 法最大识别率为 84%,而 Multi-PCA 法最大识别率可以达到 89.33%,识别率明显提高。

2. 分析

由于 Multi-PCA 算法的中心思想是以分类为目的,因此有助于提高识别率。因为在建立特征空间时至少需要 2 个样本,所以 Multi-PCA 算法对单样本识别问题无效,而 PCA 可以把所有类别的样本一起训练建立特征空间,因此 PCA 可以用来识别单样本问题。

5.4 本章小结

鉴于驾驶环境应用的需要,本章提出应用 Multi-PCA 识别驾驶员视线方向的方法。首先,分析 PCA 方法用作识别方法时的不足,然后介绍了 Multi-PCA 算法。在第 3 章精确定位驾驶员面部技术的基础上,提出采用联合投影函数与



眼角检测相结合的方法定位驾驶员眼睛区域。建立 5 类特征空间,并测试不同样本在不同特征空间下的重构误差,发现只有在样本类别与特征空间类别相同时才能获得最小重建误差。测试特征空间相似度,进一步证明 Multi-PCA 算法的有效性。最后,对不同视线时的眼睛区域图像进行识别实验,并对 Multi-PCA 与 PCA 的识别效果进行对比,发现 Multi-PCA 算法可以获得较高识别率。

参考文献

- [1] K. Huang, M. M. Trivedi. Driver Head Pose and View Estimation with Single Omni directional Video Stream [C]. Proceedings of the first International Workshop on In-Vehicle Cognitive Computer Vision Systems, in Conjunction with the Third International Conference on Computer Vision Systems, Seattle, USA, 2003: 709-714.
- [2] C. H. Morimoto, M. R. M. Mimica. Eye Gaze Tracking Techniques for Interactive Applications [J]. Computer Vision and Image Understanding. 2005, 98 (1): 4-24.
- [3] C. Colombo, A. D. Bimbo. Interacting Through Eyes [J]. Robotics and Autonomous Systems. 1997, 19 (3): 359-368.
- [4] R. Stiefelhagen, J. Yang, A. Waibel. A Model-Based Gaze Tracking System [J]. International Journal of Artificial Intelligence Tools. 1997, 6 (2): 193-209.
- [5] T. Hutchinson, K. White, K. Reichert, et al. Human-Computer Interaction Using Eye-Gaze Input [J]. IEEE Trans. on Systems Man Cybernetics. 1989, 19 (6): 1527-1534.
- [6] S. Baluja, D. Pomerleau. Non-Intrusive Gaze Tracking Using Artificial Neural Networks [J]. Advances in Neural Information Processing Systems. 1994, (6): 753-760.
- [7] M. Perrone. Estimating Human Gaze Direction [C]. Computational Vision and Active Perception Laboratory, Royal Institute of Technology, Sweden, 1995: 2-14.
- [8] G. Bebis, K. Fujimura. An Eigenspace Approach to Eye-Gaze Estimation [C]. International Conference Parallel and Distributed Computing Systems, Las Vegas, USA, 2000: 1-6.
- [9] 刘青山, 卢汉青, 马颂德. 综述人脸识别中的子空间方法 [J]. 自动化学报, 2003, 29 (6): 900-911.
- [10] 路玉峰, 王增才, 刘学忠. 提高 PCA 识别率的新算法 [J]. 光学技术. 2008, 34 (1): 10-13.
- [11] 孙即祥. 现代模式识别 [M]. 长沙: 国防科技大学出版社, 2002.
- [12] 边肇祺, 张学工. 模式识别 [M]. 北京: 清华大学出版社, 1999.
- [13] 顾华, 苏光大, 杜成. 人脸的眼角自动定位 [J]. 红外与激光工程. 2004, 33 (4): 375-379.
- [14] 李学勇. 金属标牌压印凹凸字符的特征提取和识别方法研究 [D]. 济南: 山东大学, 2008.





- [15] 李学勇, 路长厚, 李建美. 基于重构误差的单类子空间压印凹凸字符识别 [J]. 机械工程学报, 2008, 44 (5): 181-185.
- [16] 宋怀波. 低质量可见光图像的处理技术和识别方法研究 [D]. 济南: 山东大学. 2009.
- [17] L. Wang, X. Wang, J. Feng. Subspace Distance Analysis with Application to Adaptive Bayesian Face Recognition [J]. Pattern Recognition. 2006, 39 (3): 456-464.

6.1 概述

驾驶员在十字路口转向时，其视线偏离车辆前方的持续时间会超过 2s（大约为 3s^[1]），这也被认为是一种视觉分散行为。但这并非是一个危险情况，所以在视觉分散检测模型中增加了一个驾驶员转向行为识别的模块。本章对驾驶员的转向行为进行识别研究。

识别驾驶员转向行为最简单的方法似乎是检测转向灯的使用。然而，这个看起来简单明了的二值信号也可能是含糊不清的。例如，驾驶员使用该信号时，即可能是暗示将要转向还有可能是变道。并且，驾驶员在使用转向信号时并不是一致的，例如在实车与模拟实验中发现，驾驶员只在一半（甚至不到一半）的转向、变道过程中使用转向灯^[2]。

Lee^[3]采用两个摄像机分别获取驾驶员视线方向与道路方向、车辆行驶方向，然后根据视线方向与道路方向的关系、道路方向与车辆行驶方向的关系判断当时驾驶环境的危险程度，可以减少将变道误判为视觉分散的次数。但这种方法在车辆经过十字路口、T 形路口而转向时，并不能有效降低误判。因为，此时驾驶员需要较多地观测车辆两侧，而在这样的路口并没有合适的分道线来判断车辆是否在转向。

驾驶行为前的固定动作，如制动前的姿势转换^[4]、加速前的眼睛运动^[5]，可以是识别驾驶行为的充分依据。Liu^[6]等利用转向角、转向速率、车辆速度、加速度来建立驾驶行为的隐马尔科夫模型，来识别驾驶员的转向、停车等行为，识别率基本在 50% 左右。Kuge^[7]应用转向角、转向速率、转向力等建立驾驶员转向行为的隐马尔科夫模型，实验证明模型可以识别驾驶员不同的变道行为。但是，应用到的车辆参数需要特殊的传感器来测量。

除了应用车辆的行驶参数，还可以利用驾驶员的动作来识别其行为，如驾驶员的面部姿势、眼睛视线方向等^[8]。Cheng 等根据驾驶员转向时面部、手部的姿势来建立了转向模型以识别驾驶员转向行为^[1]。但这并不适合于本书的应用，



因为本书识别驾驶员转向的目的，是防止把驾驶员转向时的视线偏离情况误认为是视觉分散行为，所以不能再应用视线（面部姿势）来识别驾驶员转向行为。因此，我们提出采用视频分析检测、跟踪驾驶员的手部位置，根据手部位置标准差识别驾驶员转向行为的方法。

6.2 转向类型

车辆行驶过程中驾驶员需要完成各种各样的转向行为，如在高速路斜道、曲线道路、避免路障、U形弯处等。在交叉路口处的转向比较典型，需要车辆的前进方向变化90°。人们会在T形路口或者十字路口，或者是在由支道到干道的转向处，遇到类似转向过程。这些路口具有数目不等的车道，连接单向或者双向的道路。在十字路口处有两类转向方式：一是不间断行驶，车辆没有停止行为的转向（慢转）；二是先停止再等待至少2s 然后开始进入转向^[1]。在这里我们对十字路口处的第一类转向方式进行识别研究。

6.3 驾驶员手部检测及位置分析

6.3.1 手部检测

在转向盘右上方安装摄像机，来获取驾驶员手部运动视频，其中镜头对准转向盘中心，与转向盘平面呈19.6°夹角，获取手部图像如图6-1a所示。

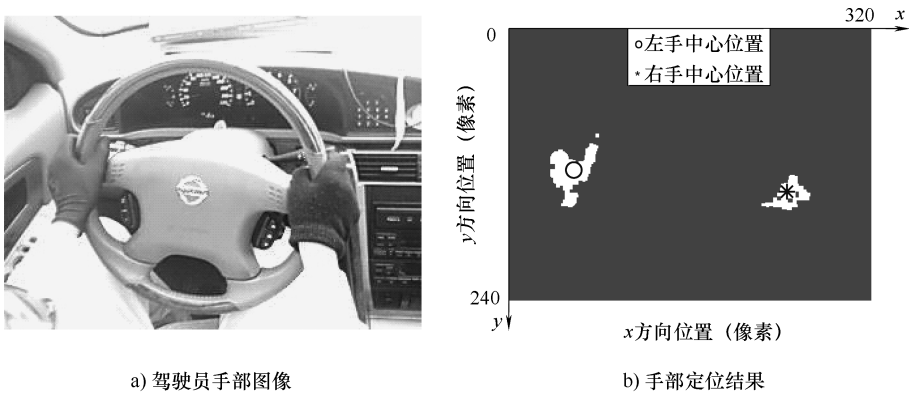


图 6-1 驾驶员手部图像及定位结果

最直接的手部检测方法，是根据第3章建立的混合肤色模型进行双手检测。但在转向过程中，驾驶员的双手有可能交叉，或者相互重合，以至于在图像中



不能区分左、右手。为在图像中容易区分驾驶员的左、右手，我们提出让驾驶员双手佩戴不同颜色手套（左手红色、右手蓝色，都与背景颜色差别较大）的方法来区分其左、右手。戴手套的方法对驾驶员来说侵入性较小，例如，很多驾驶员就习惯于戴手套驾驶车辆。

所摄取图像的色彩空间为 RGB 空间，在 RGB 空间下根据红颜色的 R 分量较高而 G、B 分量较低；蓝颜色的 B 分量较高而 R、G 分量较低。根据这些特征选择合适阈值检测驾驶员的左右手，并对检测后的图像进行腐蚀、膨胀^[9,10]操作去除杂散小区域。在检测出双手后根据投影值确定左右手的中心位置，定位效果如图 6-1b 所示。

6.3.2 手部位置分析

让驾驶员驾驶车辆沿着典型的道路行驶，并在驾驶员左转向、右转向、变道行驶、直线行驶时采集驾驶员的手部运动视频。每种驾驶行为下各采集 50 段视频，每段大约持续 20s。根据 6.3.1 小节中的方法确定驾驶员左右手中心的 x 、 y 位置，作为驾驶员手部位置，其变化规律如图 6-2 所示。

从图中我们可以清楚地看出，转向行为与其他行为的不同。驾驶员左右手的位置在转向时变化比较剧烈，特别是在 y 方向上的变化更加明显，并且主要集中在驾驶员转向开始或者即将结束时。既然驾驶员手部位置在转向时变化剧烈，那么我们可以根据手部位置的标准差来判断驾驶员是否在转向。选择每种驾驶行为中视频比较完整、图像质量较好的 10 段，计算手部位置标准差的平均值，结果如表 6-1 所示。

表 6-1 手部位置标准差 (单位：像素)

	左手位置标准差		右手位置标准差	
	x 方向	y 方向	x 方向	y 方向
左转向	7.1	29.2	10.9	24.4
右转向	6.5	13.9	7.3	20.6
变道行驶	2.3	6.4	2.4	4.0
直线行驶	1.7	4.2	1.9	4.1

因为转向等行为所耗费的时间不一致，因此表 6-1 中统计标准差时使用的窗口宽度的宽度也不一致。其中，转向的起始点为驾驶员手部开始向相应方向移动的时间，终点为驾驶员手部停止向原来位置恢复运动的时间，其中转向的平均时间是 10.7s，如表 6-2 所示。



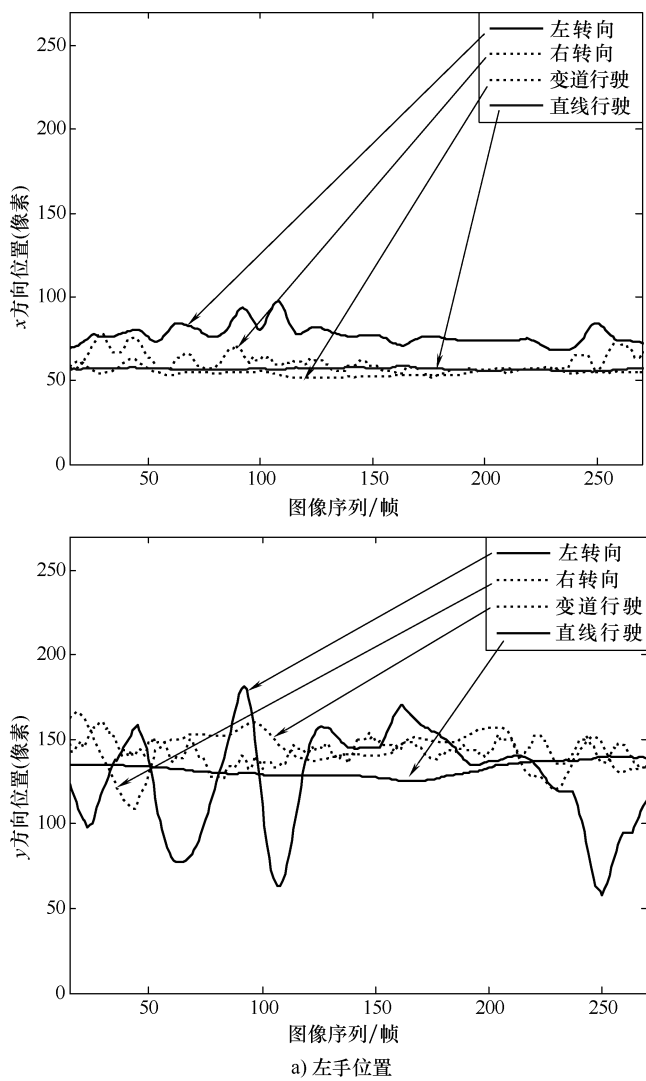


图 6-2 不同驾驶行为中左右手位置变化规律

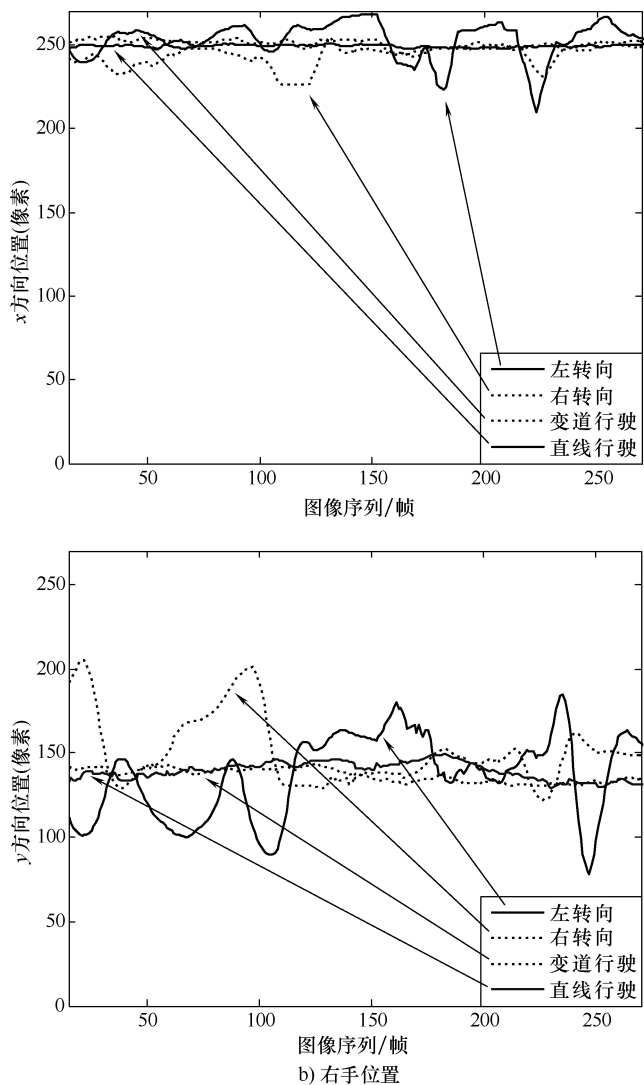


图 6-2 不同驾驶行为中左右手位置变化规律（续）

表 6-2 驾驶员转向时间统计（单位：s）

	最 长	最 短	平 均	标 准 差
时间	14.4	7.3	10.7	2.5

表 6-1、表 6-2 中统计的数据是在没有信号灯及其他车辆、行人干扰情况下进行的转向，因为信号灯、干扰物的影响会使驾驶员的驾驶行为变得非常复杂。



6.4 粒子滤波跟踪

在视频图像的每一帧中定位驾驶员的双手位置显然会降低系统的实时性。根据视频图像的特点，利用相邻两帧图像间的相关性，采用目标跟踪方法可以进一步提高驾驶员双手的定位速度。

粒子滤波，又称为序列蒙特卡洛方法，其基本思想是通过一系列加权粒子来逼近目标的后验概率密度。

6.4.1 粒子滤波跟踪原理

假设目标的状态转移模型为

$$x_t = f(x_{t-1}) + Aw_t \quad (6-1)$$

式中 f ——关于状态 x_{t-1} 的非线性函数， $f: \mathbb{R}^{n_x} \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}^{n_x}$ ；

w_t ——一组独立同分布的离散过程噪声序列， $t \in N$ ；

n_x ——状态向量的维数；

n_w ——过程噪声向量的维数；

N ——自然数集合。

式 (6-1) 中假定了运动系统是一个一阶马尔可夫过程，本节以下的公式推导也都服从该假设。令目标的观测模型为

$$y_t = g(x_t) + Bv_t \quad (6-2)$$

式中 g ——关于状态 x_t 的非线性函数， $g: \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \rightarrow \mathbb{R}^{n_y}$ ；

v_t ——一个独立同分布的离散观测噪声序列， $t \in N$ ；

n_y ——分别为观测向量的维数；

n_v ——观测噪声向量的维数。

如图 6-3 所示，贝叶斯滤波原理的实质是试图用所有已知信息来构造系统状态变量的后验概率密度，即用系统状态转移模型预测状态的先验概率密度，再使用最近的观测值对其进行修正，得到后验概率密度^[11]。

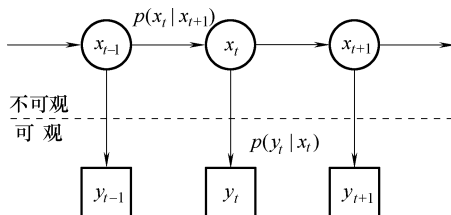


图 6-3 目标跟踪的状态-空间模型



1. 预测

假设已经得到了 $t-1$ 时刻的后验 pdf $p(x_{t-1} | Y_{t-1})$ ，那么利用系统的状态转移模型 [如式 (6-1) 所示]，可以获得 t 时刻状态变量的先验 pdf。由 Chapman-Kolmogorov 方程得

$$p(x_t | Y_{t-1}) = \int_{x_{t-1}} p(x_t | x_{t-1}) p(x_{t-1} | Y_{t-1}) dx_{t-1} \quad (6-3)$$

上述公式的推导利用了系统状态转移模型中所描述的一阶马尔科夫过程，即 $p(x_t | x_{t-1}, Y_{t-1}) = p(x_t | x_{t-1})$ ，而通过系统的运动模型以及统计量 w_{t-1} ，可以确定状态预测概率模型 $p(x_t | x_{t-1})$ 。

2. 更新

在 t 时刻有新的观测值 y_t 引入，那么可以用它来更新先验 pdf $p(x_t | Y_{t-1})$ 来获得当前状态的后验 pdf $p(x_t | Y_t)$ 。根据贝叶斯公式和马尔科夫过程的性质有

$$p(x_t | Y_t) = \frac{p(y_t | x_t) p(x_t | Y_{t-1})}{p(y_t | Y_{t-1})}$$

其中，标准化常量 $p(y_t | Y_{t-1}) = \int p(y_t | x_t) p(x_t | Y_{t-1}) dx_t$ 由观测模型式 (6-2) 和统计量 v_t 所定义的似然函数 $p(y_t | x_t)$ 所决定。对于非线性、非高斯系统想采用上述的卡尔曼滤波方法来得到贝叶斯解是很困难的，而粒子滤波则提供了一个有效手段，得到这种情况下的贝叶斯解。

6.4.2 序贯重要采样

序贯重要采样 (Sequential Importance Sampling, SIS) 是粒子滤波的核心技术。如果系统的整个递归过程符合马尔科夫假设，那么 t 时刻，系统状态的后验 pdf 可用一组带有权值的粒子 $\{x_t^{(i)}, w_t^{(i)}\}_{i=1}^N$ 表示为

$$p(x_t | Y_t) \approx \sum_{i=1}^N w_t^{(i)} \delta(x_t - x_t^{(i)}) \quad (6-4)$$

式中 $w_t^{(i)}$ ——粒子对应的权值，且 $\sum_{i=1}^N w_t^{(i)} = 1$ ；

$\delta(\cdot)$ ——狄拉克函数。

一般令 $x_t^{(i)} \sim q(X_t | Y_t)$ ，并且满足

$$q(X_t | Y_t) = q(x_t | X_{t-1}, Y_t) q(X_{t-1} | Y_{t-1}) \quad (6-5)$$

这样，便可以通过前一时刻的已知采样 $x_{t-1}^{(i)} \sim q(X_{t-1} | Y_{t-1})$ ，以及当前的先验状态 $x_t^{(i)} \propto q(x_t | X_{t-1}, Y_t)$ 来获取 t 时刻的采样 $x_t^{(i)} \propto q(X_t | Y_t)$ ，这里 $q(X_t | Y_t)$ 称为重要密度函数。这样式 (6-4) 的权值定义为

$$w_t^{(i)} \propto \frac{p(x_t^{(i)} | Y_t)}{q(x_t^{(i)} | Y_t)} \quad (6-6)$$

事实上，每一时刻粒子的权值都是通过层层迭代演化而来的，由贝叶斯公





式有

$$p(x_t | Y_t) \propto p(y_t | x_t) p(x_t | x_{t-1}) p(x_{t-1} | Y_{t-1}) \quad (6-7)$$

代入式 (6-6)，得到权值的迭代公式

$$w_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(y_t | x_t) p(x_t | x_{t-1})}{q(x_t | X_{t-1}, y_t)} \quad (6-8)$$

在给定重要性分布函数 $q(X_t | Y_t)$ 的条件下，式 (6-8) 给出递推计算重要性权值的方法。

6.4.3 重采样

为了避免序贯重要性采样方法的粒子退化现象，Gordon^[12,13]等人提出了重采样方法。

Tan^[14,15]等提出了一个有效采样尺度的概念，定义有效采用尺度为

$$N_{eff} = \frac{N_s}{1 + \text{var}(\omega_t^{(i)})} \quad (6-9)$$

式中 $\omega_t^{(i)}$ ——粒子的真实权值。

但式 (6-9) 很难进行精确估计，因此一般利用 \hat{N}_{eff} 代替 N_{eff} 来近似计算

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N_s} (\omega_t^{(i)})^2} \quad (6-10)$$

式中 $\omega_t^{(i)}$ ——归一化后的粒子权值。

6.4.4 基于直方图的粒子滤波算法

Nummiaro^[16]等提出了一种基于直方图的粒子滤波方法。

Color Particle Filter 跟踪算法

在时间步 $k-1$ ，已知 $S_{k-1}^i = \{x_{k-1}^i, b_{k-1}^i | i=1, 2, \dots, M\}$

步骤 1 对每个粒子计算其累计权重

$$c_{k-1}^0 = 0$$

$$c_{k-1}^i = c_{k-1}^{i-1} b_{k-1}^i | i=1, 2, \dots, M,$$

从而有 $S_{k-1}^i = \{x_{k-1}^i, b_{k-1}^i, c_{k-1}^i | i=1, 2, \dots, M\}$

步骤 2 重采样，从 S_{k-1}^i 中选取 M 个状态（可以重复）。

步骤 3 应用模型 $x_k^i = f(k, x_{k-1}^i) + w_k$ 预测 k 时间步的状态。

步骤 4 对新状态 x_k^i ，根据 $p(y | x)$ 求与之对应的权重 b 。

步骤 5 将权重归一化为 $\sum_{i=1}^M b_k^i = 1$ ，同时构建状态组合。

$$S_k^i = \{x_k^i, b_k^i | i=1, 2, \dots, M\}$$



步骤6 由 $S_k^i = \{x_k^i, b_k^i \mid i=1, 2, \dots, M\}$ 估计目标在时间步 k 的状态

$$E[S_k] = \sum_{i=1}^M b_k^i x_k^i$$

6.5 驾驶员手部跟踪

6.5.1 状态转移模型

选择两个矩形区域来表示驾驶员两只手，以一只手为例进行说明，其状态向量可以定义为

$$X = \{x, y, x', y', s_x, s_y, s'_x, s'_y\} \quad (6-11)$$

式中 x, y ——矩形的中心坐标；

x', y' ——矩形中心的移动速度；

s_x, s_y ——矩形的宽和高；

s'_x, s'_y ——矩形宽和高的变化速度。

采用一阶自回归模型对这些参数建模，则有

$$X_t = X_{t-1} + V_t \quad (6-12)$$

式中 V_t ——多变量标准正态分布的噪声。

6.5.2 颜色特征及其观测模型

定义目标矩形区域内的颜色分布为 $p(x) = \{p_u(x)\}$, $u=1, 2, \dots, m$, m 为直方图的级数。采用核函数对直方图进行加权，使靠近目标矩形中心区域的权值增大^[11]

$$p_u(x) = B \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (6-13)$$

式中 δ ——Kronecker delta 函数；

k ——核函数，例如 Epanechnikov 核函数。

以 q 表示目标模型， $p(x)$ 为候选目标 x 的特征表达，两者的相似性可表示为

$$d_c[p(y), q] = \sqrt{1 - \sum_{u=1}^m \sqrt{p_u(y) q_u}} \quad (6-14)$$

相应的观测似然值（即权重）可以定义为

$$b_c = \frac{1}{\sqrt{2\pi}\sigma_c} \exp - \frac{d_c^2}{2\sigma_c^2} \quad (6-15)$$

式中 σ_c ——观测高斯噪声的标准方差。

6.5.3 手部跟踪

为了验证跟踪算法的定位性能，进行驾驶员双手跟踪实验。跟踪中粒子数



均取 $N=200$ 。所有的实验均是在 Intel Pentium Dual E2180 2.00Ghz CPU, 2G 内存的 PC 上实现的, 在 Matlab2006 的软件环境下可以达到大约 7 帧/s 的处理速度。为了避免跟踪的累积误差, 跟踪进行一定时间后, 重新检测手部位置。如果检测到的手部位置与跟踪结果相差较大, 则以检测结果为准, 初始化跟踪模块, 以提高跟踪系统的鲁棒性。双手跟踪结果如图 6-4 所示。

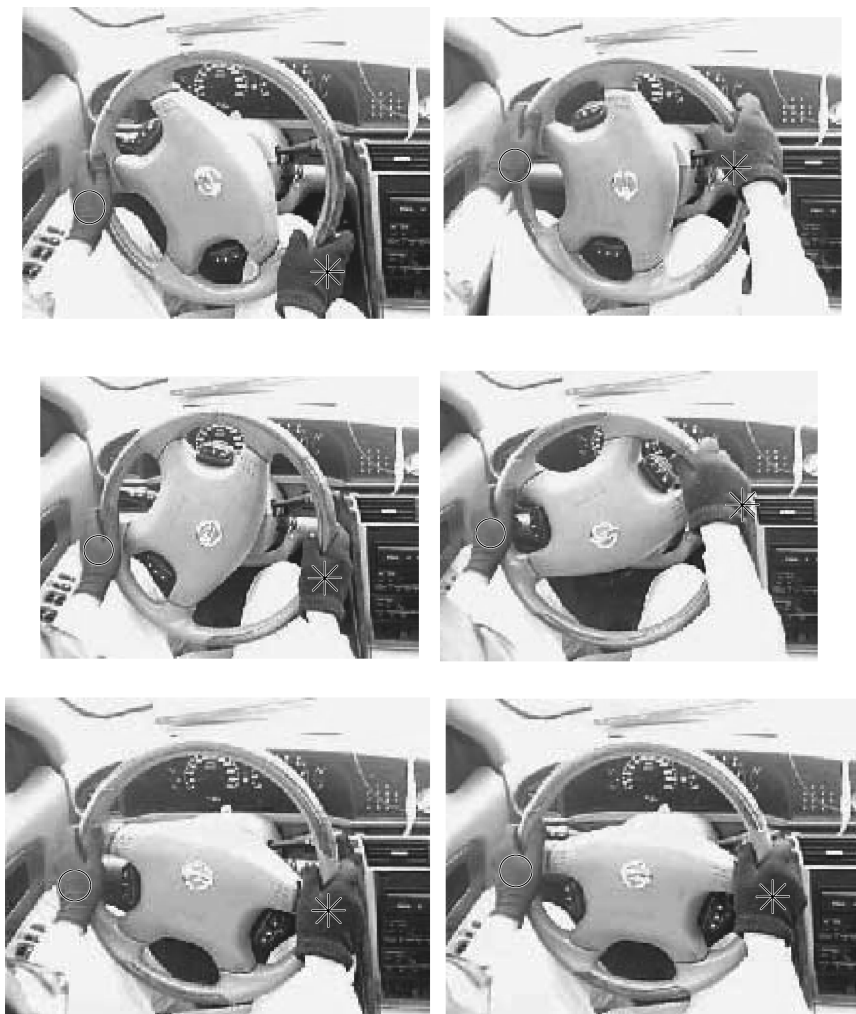


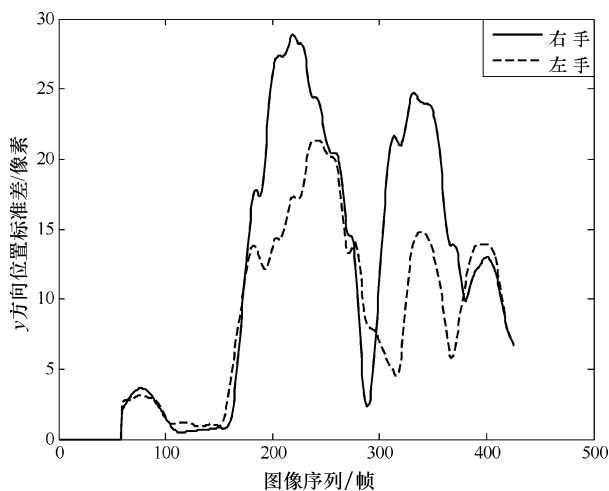
图 6-4 驾驶员双手跟踪结果

6.6 驾驶员转向行为识别

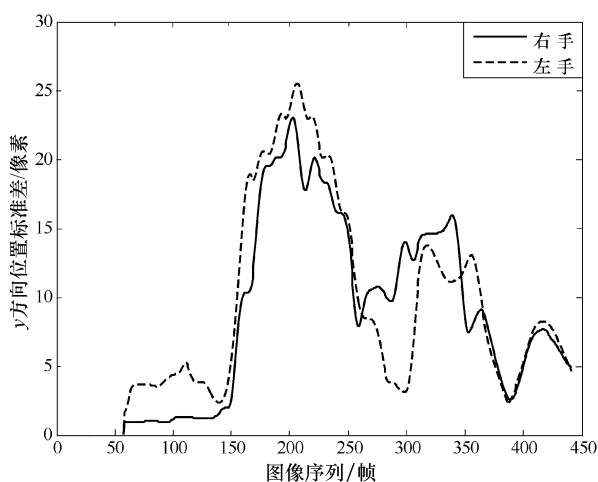
根据图 6-3 及表 6-1, 可以看出驾驶员在转向时手部的运动要比其他驾驶情



况时复杂得多。设置合适的阈值,利用手部位置的标准差可以识别出驾驶员的转向行为。本书识别转向的目的是防止检测系统把驾驶员此时的视线偏离误认为是视觉分散行为,这就要求统计手部位置标准差的时间窗口不能过长,因此本书选择50帧图像中的手部位置来计算标准差(帧率=30帧/s)。驾驶员左右手的位置在 y 方向上的变化更剧烈一些,因此以驾驶员左右手在 y 方向上位置的标准差为判断依据。这样从视频图像的第50帧开始计算标准差,计算结果如图6-5所示。



a) 右转向



b) 左转向

图 6-5 驾驶员双手 y 方向位置标准差

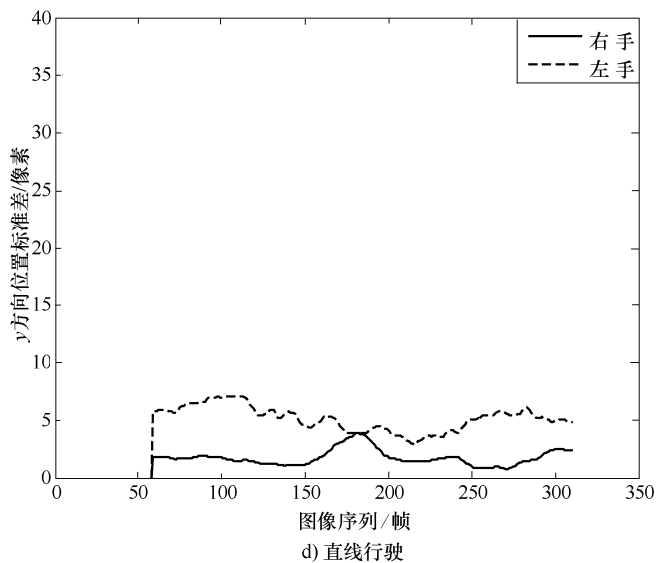
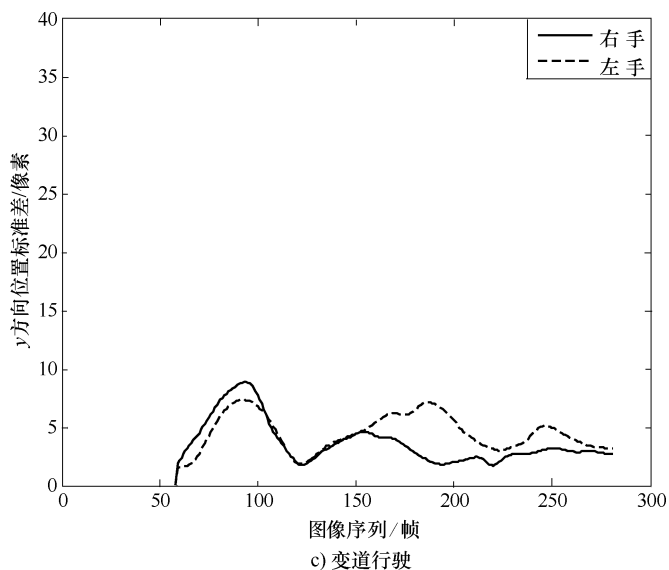


图 6-5 驾驶员双手 y 方向位置标准差 (续)

从图 6-5 可以看出驾驶员左右手在 y 方向上的位置标准差，在转向时要明显大于其他驾驶行为时的标准差。特别是在转向行为的始末，双手位置标准差明显偏大。因为在转向行为的中间，驾驶员需要在一段时间内保持转向角度的稳定，所以手部位置的标准差有所降低。驾驶员在行车过程会发生一只手离开转



向盘的行为，如在换档、调节收音机等的时候。但是，驾驶员不会双手同时离开转向盘，因此选择驾驶员两只手位置的标准差为判断依据，只有两只手位置的标准差同时超过某个阈值，才认为驾驶员发生了转向行为。考虑到两种现象对转向行为识别的影响，并假设视频开始部分驾驶员没有转向行为，我们提出转向行为识别算法如图 6-6 所示。

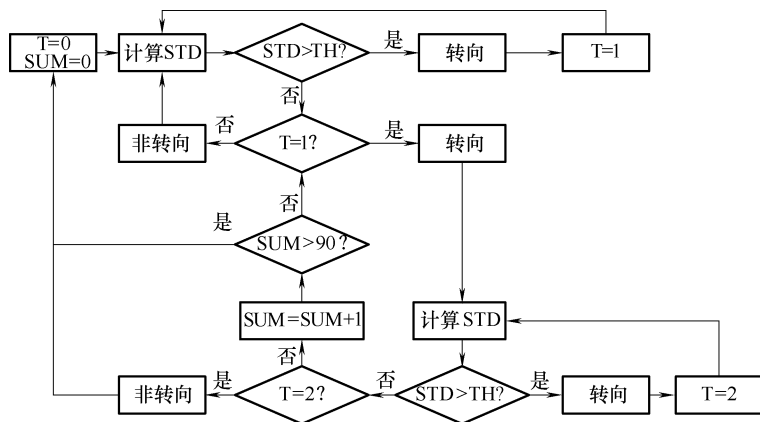


图 6-6 转向行为识别算法

(图中 STD—双手位置的标准差、TH—阈值、T—标志位、SUM—非转向帧数和)

设置标志位的目的是区分转向开始与结束部分，计算非转向帧数目的是为了区分驾驶员是否在进行转向，因此在识别时不再区分左右转向，并把直线行驶、变道行驶定义为其类型。设定阈值为 10 像素，对表 6-1 中涉及的驾驶行为进行识别实验，识别结果如表 6-3 所示。

表 6-3 转向识别结果

(%)

	左 转 向	右 转 向	直 线 行 驶	变 道 行 驶
识别为转向	93.5	94.6	4.7	5.2

整体来看识别率较高，分析错误识别的例子，发现把转向误识为非转向的图像主要发生在转向开始的 15 帧图像内。而把非转向识别为转向的情况主要发生在转向结束后，这时识别错误的图像较少，一般在 5 帧以内。在变道、直线行驶的过程中也有少量图像被识别为转向行为，主要原因是阈值设置较小的问题，增大阈值可以防止这部分的识别错误，但是会增加其他情况下的错误识别率。



6.7 本章小结

本章对驾驶员的转向行为进行识别研究,提出利用驾驶员手部位置的标准差识别转向行为的方法,并研究了视频图像中驾驶员双手的定位及跟踪。本书作者首先研究了驾驶员双手的定位方法,并分析了不同驾驶行为时驾驶员双手位置的变化特点,发现驾驶员在转向时双手位置变化比较剧烈,可以根据这个特征识别驾驶员的转向行为。为提高驾驶员双手定位的实时性,研究了粒子滤波算法跟踪驾驶员双手的方法。最后提出转向行为识别算法,实验证明此算法能有效识别驾驶员的转向行为。

参考文献

- [1] Shinko Yuanhsien Cheng, Mohan Manubhai Trivedi. Turn-Intent Analysis Using Body Pose for Intelligent Driver Assistance [J]. IEEE Pervasive Computing. 2006, 5 (4): 28-37.
- [2] D. D. Salvucci, A. Liu. The Time Course of a Lane Change: Driver Control and Eye-Movement Behavior [J]. Transportation Research Part F. 2002 (5): 123-132.
- [3] J. D. Lee, J. D. Li, L. C. Liu. A Novel Driving Pattern Recognition and Status Monitoring System [C]. Lecture Notes in Computer Science. 2006, 4319: 504-512.
- [4] M. Helander. Applicability of Drivers' Electrodermal Response to the Design of the Traffic Environment [J]. J. Applied Psychology. 1978, 63 (4): 481-488.
- [5] M. Land, D. N. Lee. Where We Look When We Steer [J]. Nature. 1994, 369 (6483): 742-744.
- [6] A. Liu, A. Pentland. Towards Real-Time Recognition of Driver Intentions [C]. IEEE Conference on Intelligent Transportation System, Boston, MA, USA, 1997: 236-241.
- [7] N. Kuge, T. Yamamura, O. Shimoyama. A Driver Behavior Recognition Method Based on a Driver Model Framework [J]. Soc. Automotive Engineers Publication. 2000, 109 (6): 469-476.
- [8] D. D. Salvucci, H. M. Mandalia, Kuge Nobuyuki, et al. Lane-Change Detection Using a Computational Driver Model [J]. Human Factors. 2007, 49 (3): 532-542.
- [9] 章毓晋. 图象分割 [M]. 北京: 科学出版社, 2001.
- [10] 闫敬文. 数字图像处理技术与图像图形学基本教程 [M]. 北京: 科学出版社, 2002.
- [11] 王江涛. 基于视频的目标检测、跟踪及其行为识别研究 [D]. 南京: 南京理工大学. 2008: 27-30.
- [12] N. Gordon, D. Salmond, A. Smith. Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation [J]. IEE Proceedings on Radar and Signal Processing. 1993, 140 (2): 107-113.



- [13] 刘阳. 基于粒子滤波的机器人视觉跟踪研究与实现 [D]. 沈阳: 大连理工大学. 2007.
- [14] T. N. Tan, G. D. Sullivan, K D Baker. Model-based Localization and Recognition of Road Vehicles [J]. International Journal of Computer Vision. 1998, 27 (1): 5-25.
- [15] 邢睿智. 彩色图象序列中基于贝叶斯滤波的目标跟踪 [D]. 大连: 大连理工大学. 2005.
- [16] K. Nummiaro, E. Koller-Meier, L. V. Gool. An Adaptive Color-based Particle Filter [J]. Image and Vision Computing. 2003, 21 (1): 99-110.

车道检测研究

7.1 概述

在智能车辆、机器人等导航过程中经常应用到道路检测。道路可分为两类：结构化道路和非结构化道路。结构化道路一般是指高速公路和部分结构化较好的公路，针对这类道路检测问题就可简化为车道线或道路边界的检测。非结构化道路一般指结构化程度较低的道路，道路可能凹凸不平，道路上可能有阴影、水迹、泥沙等，改变了道路的表面特征，非结构化道路的检测主要为检测道路的边界。

目前，各种基于视觉的结构化道路检测技术基本上采用两种检测途径^[1-3]：一种是整体路面法，利用区域检测方法实现，主要是针对整个道路路面设计算法；另一种是道路边缘或车道线检测法，通过对道路边缘或车道线的检测来实现道路的检测，这种检测方法简单实用，能适应现有高速公路的结构化环境，与基于整体法的道路检测技术相比，这种道路检测方式具有更快的图像处理速度和更高的检测精度，也具有更好的应用前景。典型的检测方法主要有：AU-RORA 系统^[4]利用可变模板检测车道标志线；Klgue^[5]利用可调整的车道结构模板检测车道边界，该方法不需设定亮度变化的阈值；GOLD 系统^[6]利用模式匹配技术检测车道标志线；LANA 系统^[7]通过在频域内捕捉边缘信息的大小和方向进行边缘检测；Wang 等^[8,9]在边缘潜在区域，采用 B-样条 Snake 模型和 Catmull-Rom 样条模型检测车道曲线；TFALDA 系统^[10]利用车道边缘的起始位置、方向性及图像灰度值等特征检测车道标志线；VioLET 系统^[11]通过可变滤波器检测车道标志线，并用可变模板估计车道曲线。

非结构化道路检测算法可以分为三大类：基于道路特征的方法^[11-15]、基于道路模型的方法^[16]和基于神经网络的方法^[17]等。基于道路特征的检测方法，根据道路区域与非道路区域在边缘、纹理、颜色等方面的不同，判断道路区域；基于模型的检测方法首先对道路模型做出假设，如直线、抛物线、样条曲线等模型，然后根据道路图像匹配道路模型，检测出的道路区域比较完整，但在道路结构比较复杂的情况下，很难建立准确的模型与道路匹配；基于神经网络的



方法,具有神经网络的学习特性,但需要大量的训练样本。

本书对于结构化道路的检测,采用分区投影的方法检测车道线边缘点,然后用边缘点拟合出车道线的边缘方程;对于非结构化道路的检测,采用模糊 C 均值方法将彩色道路图像聚类分块,然后通过分析各图像块的特征来检测道路区域,最后进行边缘拟合。

7.2 结构化道路检测

对于结构化道路的检测,本书简化为检测车道线,国内高速公路主要分为直道和弯道两种类型。按国家标准^[18],高速公路转弯半径需大于 650m,即使在 40m 以内弯道也可以近似地认为是直道,且误差小于 0.01m^[19]。因此,车辆前方近距离的车道标志线可以用直线进行拟合,鉴于此情况,本书将只检测直线车道标志线,建立直线模型。

现有大部分方法只是基于结构特征进行分析,没有考虑到车道标志线在图像中的空间位置信息,这样处理的缺点是运算量大,鲁棒性差,对外界条件要求严格,路边环境对检测结果干扰大。本书的车道检测算法考虑了空间位置信息,算法首先是基于以下几点假设:

- 1) 针对结构化道路的检测,车道标志线较为明显。
- 2) 车辆行驶基本(或大多数时间)位于两车道标志线中间。
- 3) 车辆偏离检测只需检测前方较近一段距离,即检测的车道标志线假设为直线段。
- 4) 初始检测时,两车道线中间无其他标志线或文字等。

7.2.1 图像预处理

在检测车道标志线之前,为提高检测精度,减小检测误差,特对图像进行预处理,本书主要进行灰度图像平滑和灰度图像拉伸处理。

7.2.1.1 灰度图像平滑

图像平滑是一种实用的图像处理技术,能减弱或消除图像中的高频分量,而不影响低频分量。高频分量主要对应图像中的区域边缘等像素值具有较大、较快变化的部分,平滑滤波将高频分量滤去以减小局部灰度起伏,均匀像素值。常用图像方法有:线性平滑、非线性平滑和自适应平滑。

自适应平滑的思想为:中心点像素值的大小取决于其邻域中占主要成分的像素值,滤波器设计如下。

首先,定义中心点与其邻域的像素值偏差,公式如下

$$d_{\Omega}(i,j) = |I_{\Omega}(i,j) - I_{\Omega}(x,y)| \quad (7-1)$$





式中 $I_{\Omega}(x,y)$ ——灰度图像中心点 (x,y) 像素值；
 $I_{\Omega}(i,j)$ ——灰度图像中心点 (x,y) 邻域像素值；
 Ω ——中心点 (x,y) 的邻域，本书取 5×5 矩阵；
 d_{Ω} ——邻域内的偏差矩阵。

如果中心像素点 (x,y) 邻域的像素值分布均匀，则偏差矩阵 d_{Ω} 的值变化不大，通过计算方差 D_d 判断邻域像素值是否存在较大变化。若满足 $D_d < 5$ ，说明中心像素点邻域的像素值变化较小，图像颜色单一，则中心像素点的平滑，直接计算其邻域的均值即可得到；如果 $D_d \geq 5$ ，则需按照下面的方法计算邻域像素的主要成分。

利用 Otsu 算法将 d_{Ω} 分为两类 $\{d_1, d_2\}$ ，所对应个数为： $\{n_1, n_2\}$ ， $n_1 + n_2 = 25$ 。根据邻域内占主要成分的像素值平滑中心点的像素值，平滑公式如下

$$I(x,y) = \frac{\sum_{k=1}^N I_k(i,j)}{N} \quad (7-2)$$

式中 $N = \max(n_1, n_2)$ ；
 $I_k(i,j)$ ——取 N 相对应的像素值。

图 7-1 即为道路图像的平滑结果，采用该平滑算法可以更好地均匀图像的灰度值，同时还可以尽量保持图像的原始特征。

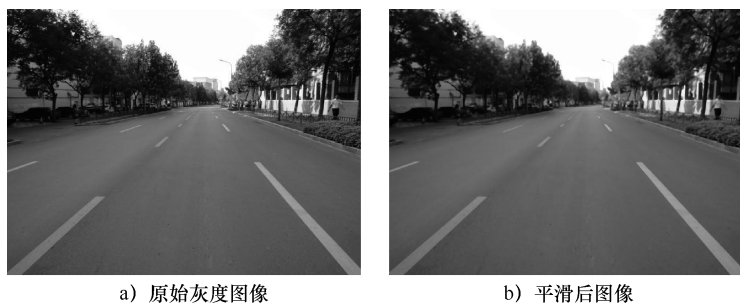


图 7-1 图像平滑

7.2.1.2 灰度图像拉伸

灰度拉伸是指根据灰度直方图的分布，有选择地拉伸某一段内像素的灰度值以改善输出图像的质量，灰度拉伸是一种最简单、最有效的图像对比度增强方法^[20]。灰度拉伸可以使灰度动态范围加大，图像对比度得到扩展，图像更清晰，特点更明显，灰度拉伸是图像增强的重要手段。

图像灰度拉伸可分为线性变换和非线性变换两大类，其中线性变换又分为简单线性变换和分段线性变换两种，它们都是在某一像素范围内成固定比例的增大（或减小）像素值，由于比例不变，并不能最大化拉伸像素范围；而非线性变换可在整个像素范围内随着像素值的变化，其拉伸比例也跟随变化，能较

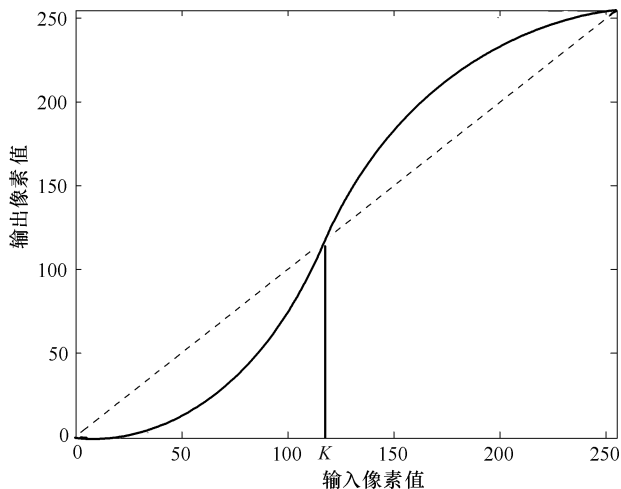


好改善图像的像素分布，主要有 Sigmoid 函数和双曲正切函数（Th）拉伸等。本书采用非线性函数拉伸灰度图像。

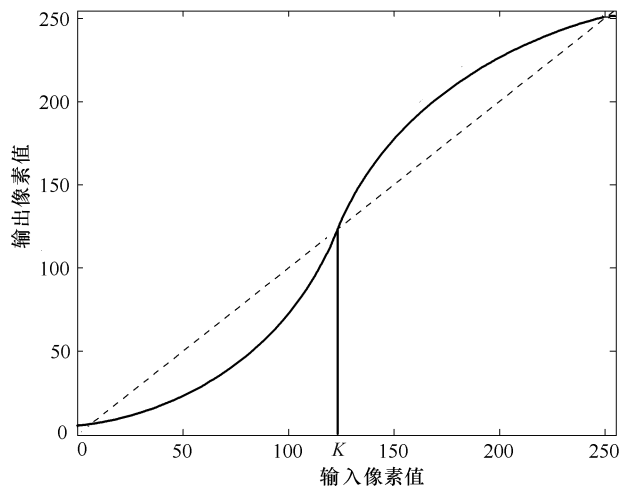
(1) Sigmoid 函数定义如下

$$S(x) = \frac{255}{1 + e^{-k_1 \frac{(x-k)}{k}}} \tag{7-3}$$

式中 k ——灰度分割参数；
 k_1 ——曲线陡度控制系数；
 x ——像素值， $x \in [0, 255]$ 。
Sigmoid 函数曲线如图 7-2a 所示。

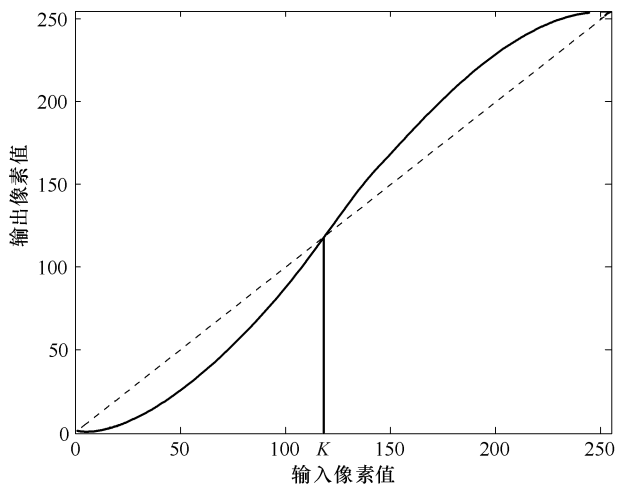


a) Sigmoid

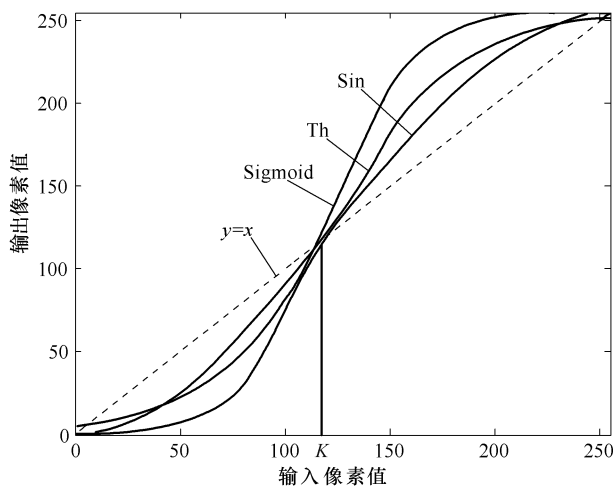


b) Th

图 7-2 拉伸函数曲线



c) Sin



d) Sigmoid、Th、Sin

图 7-2 拉伸函数曲线 (续)

(2) 双曲正切函数定义如下

$$T(x) = \frac{e^{k_1 \frac{(x-k)}{k}} - e^{-k_1 \frac{(x-k)}{k}}}{e^{k_1 \frac{(x-k)}{k}} + e^{-k_1 \frac{(x-k)}{k}}} \quad (7-4)$$

式中 k ——灰度分割参数;

k_1 ——曲线陡度控制系数;

x ——像素值, $x \in [0, 255]$ 。



双曲正切函数曲线如图 7-2b 所示。

从图 7-2a 和图 7-2b 中可以看出, Sigmoid 函数、双曲正切函数的曲线陡峭(陡度控制系数已达极限), 斜率较大会使图像像素值变化剧烈, 不适合本书的图像处理。

本书采用基于 \sin 函数的调整曲线, 函数定义如下所示

$$\sin(x) = \begin{cases} k[\sin(\frac{\pi}{2}(\frac{x}{k}-1)) + 1] & 0 \leq x < k \\ (255-k)\sin(\frac{\pi(x-k)}{2(255-k)}) + k & k \leq x \leq 255 \end{cases} \quad (7-5)$$

式中 k ——灰度分割参数;

x ——像素值, $x \in [0, 255]$ 。

拉伸曲线如图 7-2c 所示, 当 $x < k$ 时, 曲线位于直线 $y = x$ 的下方, 图像像素值缩小; 当 $x \geq k$ 时, 曲线位于直线 $y = x$ 的上方, 图像像素值增大。物理意义为: 图像像素点越暗, 拉伸后会更暗, 图像像素点越亮, 拉伸后更亮, 加大了图像的对比度。根据实际情况本书中 $k = 120$ 。

从图 7-2d 中可以看出, 三种曲线都能实现灰度拉伸功能, 但是 Sigmoid、Th 函数曲线由于函数自身性质, 其偏离直线 $y = x$ 的距离较远, 即曲线斜率波动较大, 这样会使拉伸后的灰度图像像素值波动较大, 改变原始图像的纹理特征, 丧失很多信息, 而 \sin 函数曲线变化较为平缓, 不会大尺度改变像素值, 从而可以在保留更多信息的前提下进行拉伸灰度。各种方法的图像拉伸结果如图 7-3 所示。

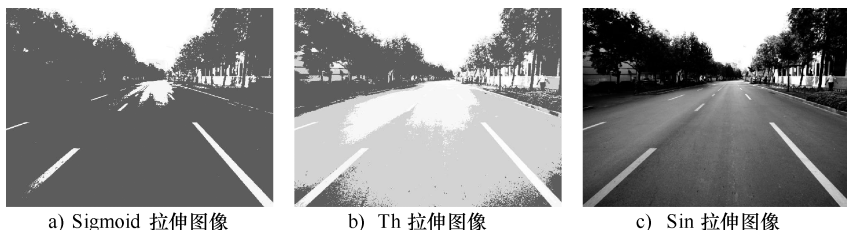


图 7-3 灰度拉伸效果图

图 7-3 为图 7-1c 的灰度拉伸图, 可以看出 Sigmoid 拉伸后图像对比度增强, 车道线变得更白, 车道变得更黑, 这样有利于图像的后续处理。然而, 道路中某些区域由于光照的影响, 其亮度较大, 由于拉伸曲线斜率较大, 使得拉伸后其亮度更大, 变成白色, 改变了原图像; 双曲正切拉伸效果同样存在此缺点; 而 \sin 函数在不改变原始图像的条件下, 增大了图像的对比度, 符合图像处理要求。





7.2.2 车道标志线分区投影

假设图像中间偏下位置（车辆正前方较近区域）为结构化道路及车道标志线部分，此处图像纹理简单，干扰较少，路面与车道标志线的灰度值差别较大，因此采用投影方法即可区别出路面和车道标志线。由于图像透视关系，车道标志线在图像中呈现倾斜状态，如图 7-4 所示。直接在全局进行灰度投影不能定位出车道标志线，因此，需将图像进行分割，然后分区域投影，投影曲线的谷点即为车道标志线边缘。此外，分区域投影还可以最大限度去除干扰的影响。

将道路图像分割成条状区域，如图 7-4 所示。图 7-5 为车道标志线模型示意图，现只取其中两个区域分析，每个区域内有两小段车道标志线，车道标志线为平行四边形，顶点分别标注为 $\{A_1 \sim A_8, B_1 \sim B_8\}$ ，显然由于位置关系 B_1 与 A_2 、 B_3 与 A_4 、 B_6 与 A_5 、 B_8 与 A_7 的坐标点重合，分别将这两个区域按照式 (3-3) 进行垂直积分投影。

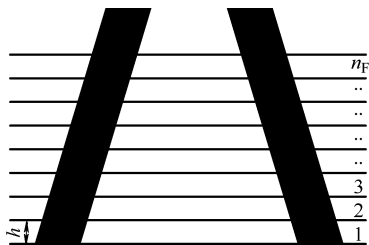


图 7-4 图像分割示意图

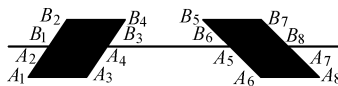


图 7-5 车道标志线模型

图 7-5 的垂直积分投影如图 7-6 所示， B_1 与 A_2 、 B_3 与 A_4 、 B_6 与 A_5 、 B_8 与 A_7 的横坐标相同，点 A_1 、 B_1 、 A_4 、 B_4 、 A_5 、 B_5 、 A_8 、 B_8 的横坐标即为车道标志线边缘的横坐标，而纵坐标在图像分割时已经确定，由此可以检测出车道边缘点。

由于道路灰度像素值的波动，投影曲线也会出现起伏、毛刺等现象，如果不进行平滑处理，在求极值时很容易陷入局部最优，因此必须对投影曲线进行平滑。曲线平滑最常用的方法就是邻域数值平均。然而直接进行邻域数值平均。由于在拐点处邻域中的数值相差很大，平均后真实值就会改变巨大，曲线形状也会跟随改变，因此需对原始平滑算法进行改进。原始邻域数值平均平滑公式如式 (7-6) 所示

$$\overline{V(x)} = \frac{1}{L} \sum_{x_1=x-L/2}^{x_1=x+L/2} V(x_1) \quad (7-6)$$

式中 x ——横坐标位置；

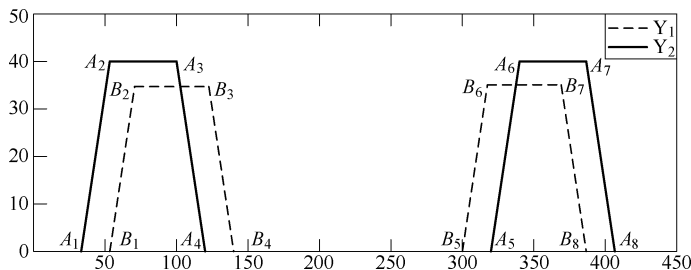


图 7-6 车道标志线投影示意图

L ——邻域长度。

改进策略：由于本书中采用积分投影，因此曲线不会出现因干扰而造成的剧烈波动。首先，判断邻域值与中心值的差值 d ，如果差值大于设定阈值 D ，则剔除该像素值，该值不参加平滑运算，即只在波动幅值较小时进行平滑运算，这样可尽最大程度上保留曲线形状。改进后的平滑公式如下

$$\overline{V(x)} = \frac{\sum_{x_1=x-L/2}^{x_1=x+L/2} k(x_1) V(x_1)}{\sum_{x_1=x-L/2}^{x_1=x+L/2} k(x_1)} \quad (7-7)$$

式中 $\overline{V(x)}$ ——在 x 处的平滑投影曲线值；

x ——曲线横坐标， $3 \leq x \leq N-2$ ；

N ——横坐标最大值；

$V(x_1)$ ——原始投影曲线；

L ——邻域长度，本书中 $L=5$ ；

$k(x_1)$ ——有效平滑系数。

计算公式如下

$$k(x_1) = \begin{cases} 1 & \text{if } d(x_1) \leq D \\ 0 & \text{else} \end{cases} \quad (7-8)$$

d 差值计算公式如下

$$d(x_1) = |V(x_1) - V(x)| \quad (7-9)$$

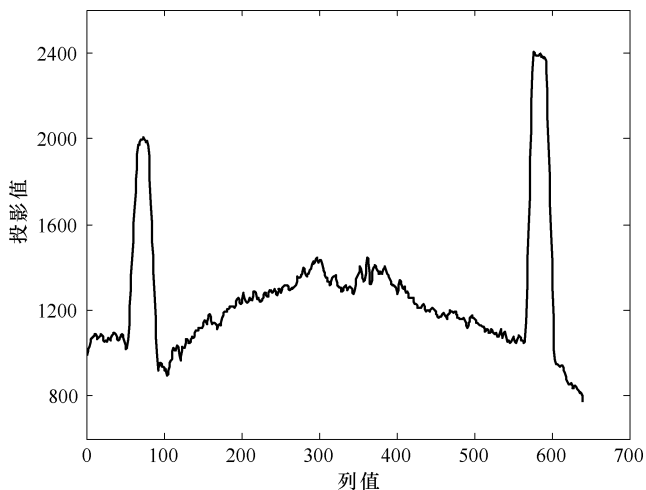
式中 $x-L/2 \leq x_1 \leq x+L/2$ ，本书中阈值取 $D=100$ 。

图 7-7 为图 7-3c 灰度拉伸图像中部分区域的垂直积分投影，图中两波峰分别为左右车道线，标记为 B_L 、 B_R ，这两个波峰各自波谷对应的横坐标为 x_{L1} 、 x_{L2} 、 x_{R1} 、 x_{R2} ，每条车道线对应两谷点距离为 $L_L = x_{L2} - x_{L1}$ ， $L_R = x_{R2} - x_{R1}$ 。

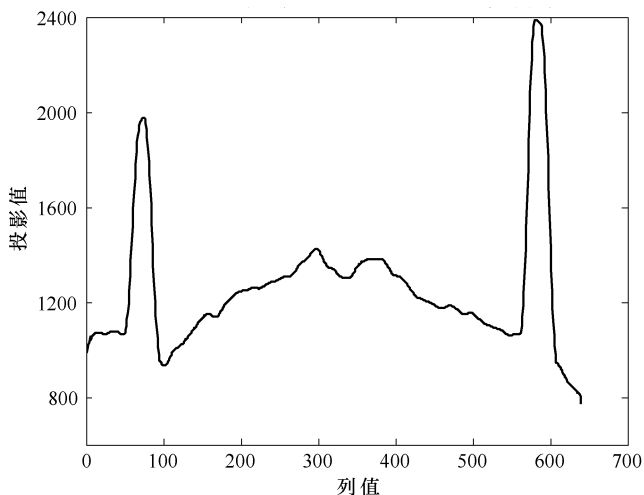
表 7-1 为图 7-7 所对应三条曲线的谷点位置值，从中可以看出，曲线平滑后谷点的位置都发生了变化，谷点分别向两侧移动，这是由于在平滑过程中曲线



的较大值和较小值共同参与了平滑计算，曲线值相差很大，从而引起曲线的波动。然而，图 7-7c 曲线的数值比图 7-7b 曲线的值更接近图 7-7a 曲线的值，说明改进后的平滑算法对数值的改变较小，曲线移动距离小，比传统平滑算法更接近真实曲线形状。

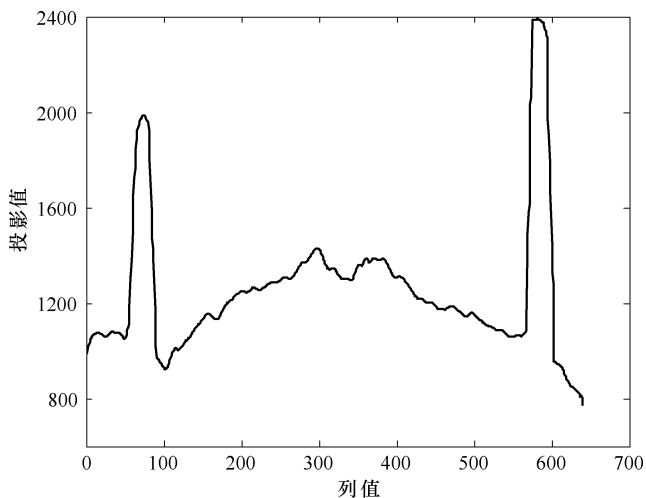


a) 未平滑曲线



b) 原始平滑曲线

图 7-7 垂直积分投影



c) 本书平滑曲线

图 7-7 垂直积分投影 (续)

表 7-1 投影曲线坐标比较

分 组	x_{L1}	x_{L2}	x_{R1}	x_{R2}	L_L	L_R
未平滑曲线	53	92	566	605	39	39
原始平滑曲线	50	95	562	609	45	47
改进平滑曲线	52	93	565	607	41	42

7.2.3 标志线特征点检测

7.2.3.1 车道标志线投影特性

车道标志线投影特性如下:

1) 由于车道标志线宽度相同且处在同一水平线上, 因此, 图像中的投影宽度应该相近, 即 $|A_4(x) - A_1(x)| \approx |A_8(x) - A_5(x)|$ 。

2) 由于标志线与路面对比度大, 两峰值的投影像素值会远大于两侧谷点投影像素值。

3) 根据平行四边形的对称关系, 峰值两侧谷点的投影像素值应该接近, 即投影曲线纵坐标相同。

7.2.3.2 算法具体实现

1) 将图像的下半部分水平分割, 分割高度 $h = 10$ 。

2) 在每个分割区域内分别进行垂直积分投影, 并平滑投影曲线。

3) 在每一个投影曲线中, 从中间坐标 x_M 开始分别向两侧寻找谷点。第一



个谷点假设为 S 点, 继续向前寻找, 当找到新的谷点后, 将新谷点定义为 S 点, 上一谷点定义为 E 点, 即 S 、 E 点依次沿着寻找方向前进。 S 、 E 点分别对应横坐标为 x_S 、 x_E , 当满足如下条件时, 停止搜索, 定位出车道标志线边缘, 其对应谷点坐标即为图 7-6 中的 A_1 、 A_4 、 A_5 、 A_8 (B_1 、 B_4 、 B_5 、 B_8)。

① 车道标志线宽度满足 $w_1 < |x_S - x_E| < w_2$, 即两谷点水平距离合适, 本书中 $w_1 = 20$, $w_2 = 100$ 。

② 由于两谷点之间为车道线而之前扫描过的区域为车道, 因此两谷点之间像素平均值大于之前扫描区域的平均值, 即 $Mean(V(x_S : x_E)) > Mean(V(x_E : x_M))$ 。

③ 两谷点之间最大像素值 $V_{\max} = Max(V(x | x \in [x_S, x_E]))$ 远大于谷点处像素值, $V_{\max} - Max(V(x_S), V(x_E)) > \Delta V$, 本书 $\Delta V = 500$ 。

④ 由于图像透视关系, 谷点横坐标 x_S 、 x_E 会在上一分割区域谷点横坐标的内侧, 如不符合条件, 此分割区域定为无效区域, 此约束可保证车道标志线为直线。

4) 返回步骤 3) 在下一个分割区域内计算, 直至全部区域检测完成。

经过以上 4 个步骤即可取得一系列车道标志线边缘处的分散点, 定义为车道标志线特征点, 记为 L_i , $i = 1, 2, 3, 4$, 分别表示从左到右四组特征点集合。分别对每组点进行直线拟合, 就可以检测出整个车道标志线边缘。

7.2.3.3 特征点调整

通过寻找垂直积分投影曲线的谷点定位出车道标志线的特征点, 因为积分关系等可能引入干扰, 使得特征点会偏离真实位置。因此, 需要对特征点 L_i 进行调整。

调整规则: 在特征点水平方向上的较小区域内计算各相邻像素灰度值的梯度, 梯度最大为车道标志线与车道的交界, 该点为车道线的边缘点即为新特征点。

定义梯度:

$$\Delta d(x) = |I(i, x) - I(i, x - 1)| \quad (7-10)$$

式中 I ——灰度图像像素值;

i ——特征点行坐标;

x ——特征点邻域坐标, $x \in (x_L - l/2, x_L + l/2]$;

x_L ——任一特征点横坐标;

l ——特征点邻域大小, 本书中取 $l = 5$ 。

特征点调整后需满足如下关系

$$x'_L = \{x | I(x) > I(x - 1), \max(\Delta d) = \Delta d(x)\} \quad (7-11)$$

式中 x'_L ——调整后特征点。



7.2.4 车道标志线拟合

经过上述处理后,已经最大限度去除了干扰点,得到大量的车道标志线特征点。常见的车道标志线参数提取方法有最小二乘法拟合和 Hough 变换方法。前者较后者计算量小,本书中特征点已经精确定位出,因此不存在最小二乘法抗干扰能力差的问题,故采用最小二乘法拟合车道标志线。

假设 $f(x)$ 为原函数, $g(x)$ 为拟合函数, $(x_i, f(x_i)) (i=0, 1, \dots, N)$ 为数据点,寻找 $g(x)$ 满足下式误差值最小

$$E_{\min} = \sum_{i=0}^N [f(x_i) - g(x_i)]^2 \quad (7-12)$$

设 x 和 y 之间的函数关系由下面的直线方程给出

$$y = ax + b \quad (7-13)$$

式 (7-13) 中有两个待定参数: a 表示直线的斜率; b 表示截距。 (x_i, y_i) 为实际测量的 N 组数据, $i=1, 2, \dots, N$ 。下面利用最小二乘法把观测数据拟合为直线。

按照式 (7-12) 要求,使式 (7-14) 达到最小

$$E = \sum_{i=1}^N [y_i - (ax_i + b)]^2 \quad (7-14)$$

当 E 达到最小时即为对参数 a, b 的最佳估计,根据式 (7-14) 应有

$$\frac{\partial}{\partial a} \sum_{i=1}^N [y_i - (ax + b)]^2 = -2 \sum_{i=1}^N (y_i - ax - b)x_i = 0 \quad (7-15)$$

$$\frac{\partial}{\partial b} \sum_{i=1}^N [y_i - (ax + b)]^2 = -2 \sum_{i=1}^N (y_i - ax - b) = 0 \quad (7-16)$$

整理后得到方程组

$$\begin{cases} bN + a \sum x_i = \sum y_i \\ b \sum x_i + a \sum x_i^2 = \sum x_i y_i \end{cases} \quad (7-17)$$

解方程组便可求得直线参数 a 和 b 的最佳估计值,即

$$a = \frac{N(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{N(\sum x_i^2) - (\sum x_i)^2} \quad (7-18)$$

$$b = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_i y_i)}{N(\sum x_i^2) - (\sum x_i)^2} \quad (7-19)$$

本书中的 y_i 是由图像所在分区决定的

$$y_i(n_f) = (n_f - 1) \times h$$

式中 n_f ——图像分区序号, $n_f=1, 2, \dots, n_F$;

n_F ——分区数;



本书中 x_i 为 y_i 所对应的 x'_i 。

7.2.5 车道标志线跟踪

为节省检测时间,提高抗干扰能力,在检测出车道标志线后,需对其进行跟踪。由于视频图像刷新很快,可采用建立感兴趣区域 (Region Of Interest, ROI) 的方法缩小检测范围,在 ROI 内利用方向滤波器进行滤波检测。建立感兴趣区域时,采用与前一帧中识别出的车道标识线平行的方法,如图 7-8 所示,实线表示前一帧图像中识别出的车道标识线,虚线在图像中所包含区域为感兴趣区域, R_L 和 R_R 分别为左、右区域。

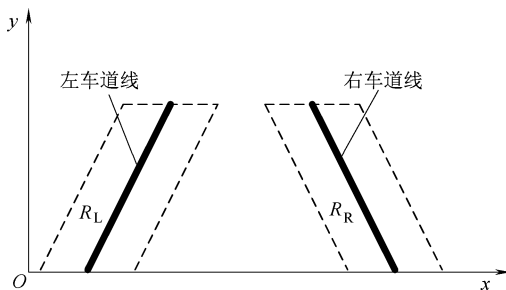


图 7-8 ROI 示意图

7.2.5.1 方向可调滤波器

1. 方向可调滤波器原理

方向可调滤波器^[21]是由 Freeman 和 Adelson 开发的一种方向特征探测器。如果函数 $f(x,y)$ 可以表示为其自身旋转后的函数的线性组合,则称 $f(x,y)$ 是可调的。方向可调滤波器是用一组不同方向上的“基本滤波器”的线性组合来实现其在任意方向上的响应。假设 $f(x,y)$ 为滤波器的脉冲响应,方向可调滤波器可表示为

$$f^\theta(x,y) = \sum_{j=1}^M k_j(\theta) f^{\theta_j}(x,y) \quad (7-20)$$

式中 $f^\theta(x,y)$ —— $f(x,y)$ 旋转 θ 的脉冲响应;

$f^{\theta_j}(x,y)$ ——基滤波器;

M ——基本滤波器的数目;

$k_j(\theta)$ ——第 j 个差值函数;

θ_j ——第 j 个基角。

任何方向上的滤波器均可用有限个“基本滤波器”的线性组合表示。方向滤波器结构示意图如图 7-9 所示。

在极坐标系中应用方向可调滤波器,设 $r = \sqrt{x^2 + y^2}$, $\varphi = \arg(x,y)$, 则

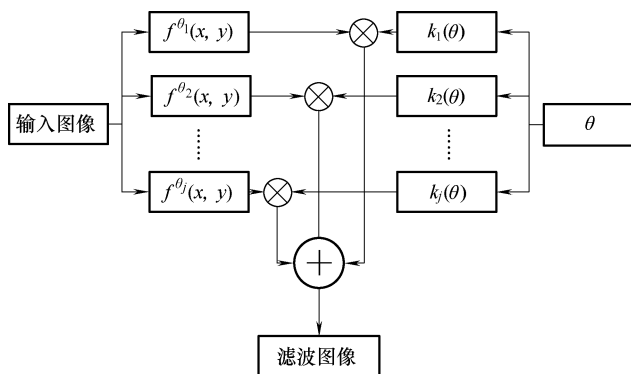


图 7-9 方向滤波器结构示意图

$f(r, \varphi)$ 的傅里叶级数表达式为

$$f(r, \varphi) = \sum_{n=-N}^{n=N} a_n(r) e^{jn\varphi} \quad (7-21)$$

式 (7-20) 应满足以下理论^[21]。

理论 1 当且仅当插值函数 $k_j(\theta)$ 满足

$$\begin{pmatrix} 1 \\ e^{j\theta} \\ \vdots \\ e^{jN\theta} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ e^{j\theta_1} & e^{j\theta_2} & \cdots & e^{j\theta_M} \\ \vdots & \vdots & \vdots & \vdots \\ e^{jN\theta_1} & e^{jN\theta_2} & \cdots & e^{jN\theta_M} \end{pmatrix} \begin{pmatrix} k_1(\theta) \\ k_2(\theta) \\ \vdots \\ k_M(\theta) \end{pmatrix} \quad (7-22)$$

时式 (7-20) 可表达成式 (7-21) 的形式。如果 $a_n(r) = 0$, 则应该去掉式 (7-22) 左边的第 n 行和右边矩阵的第 n 行。

理论 2 设 T 为式 (7-21) 中函数 $f(x, \varphi)$ 的非零系数 $a_n(r)$ 的个数, 则式 (7-20) 中基滤波器的最小数目为 T , 即 $M \geq T$ 。

理论 3 设 $f(x, y) = W(r)P_N(x, y)$, 其中 $W(r)$ 是任意窗口函数, $P_N(x, y)$ 是 x, y 的 N 次多项式, 只需 $(2N+1)$ 个基函数的线性组合就能够使 $f(x, y) = W(r)P_N(x, y)$ 旋转到任意角度。如果 $P_N(x, y)$ 只包含偶 (奇) 次项, 则只需要 $(N+1)$ 个基函数就能使 $f(x, y)$ 旋转到任意角度。

2. 方向可调滤波器设计

由于高斯函数的任意阶导数都是方向可调的, 所以本书选用基于二维高斯函数的二阶偏导数作为方向可调滤波器。二维高斯函数如式 (7-23) 所示

$$G(x, y) = e^{-\frac{(x^2+y^2)}{2}} \quad (7-23)$$

G_2^{0*} 为本书设计的一个基函数

$$G_2^{0*}(x, y) = (x^2 - 1)e^{-\frac{(x^2+y^2)}{2}} \quad (7-24)$$



G_2^{0*} 只包含偶次项的二次多项式, 依照理论 3 可知需要 3 个基函数。根据式 (7-22) 对于插值函数 $k_j(\theta)$ 有

$$\begin{pmatrix} 1 \\ e^{j2\theta} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ e^{j2\theta_1} & e^{j2\theta_2} & e^{j2\theta_3} \end{pmatrix} \begin{pmatrix} k_1(\theta) \\ k_2(\theta) \\ k_3(\theta) \end{pmatrix} \quad (7-25)$$

由于对称和对噪声的抗干扰性等原因, 通常在 $0^\circ \sim 180^\circ$ 之间选择基函数, 本书选用 $\theta_1 = 0^\circ$, $\theta_2 = 60^\circ$, $\theta_3 = 120^\circ$ 三个方向上的基函数

$$k_j(\theta) = \frac{1}{3} [1 + 2\cos(2(\theta - \theta_j))] \quad (7-26)$$

则根据式 (7-13), $G_2^{0*}(x, y)$ 滤波器在任意方向上的响应为

$$G_2^\theta = k_1(\theta) G_2^{0*} + k_2(\theta) G_2^{60*} + k_3(\theta) G_2^{120*} \quad (7-27)$$

通常基滤波器的 $x-y$ 是不可分离的, 导致该算法非常耗时, 为加快运算速度, 设计 $x-y$ 可分离的基滤波器, 该滤波器只包含最少的数目, 本书 G_2^θ 采用如下形式

$$G_2^\theta = k_a(\theta) G_{2a} + k_b(\theta) G_{2b} + k_c(\theta) G_{2c} \quad (7-28)$$

其中

$$G_{2a} = 0.9213(x^2 - 1)e^{-\frac{(x^2+y^2)}{2}}$$

$$G_{2b} = 1.843xye^{-\frac{(x^2+y^2)}{2}}$$

$$G_{2c} = 0.9213(y^2 - 1)e^{-\frac{(x^2+y^2)}{2}}$$

$$k_a(\theta) = \cos^2(\theta)$$

$$k_b(\theta) = -2\sin(\theta)\cos(\theta)$$

$$k_c(\theta) = \sin^2(\theta)$$

图 7-10 显示了构成滤波器的三个基函数的三维图; 图 7-11 为滤波器在不同角度下的三维图。

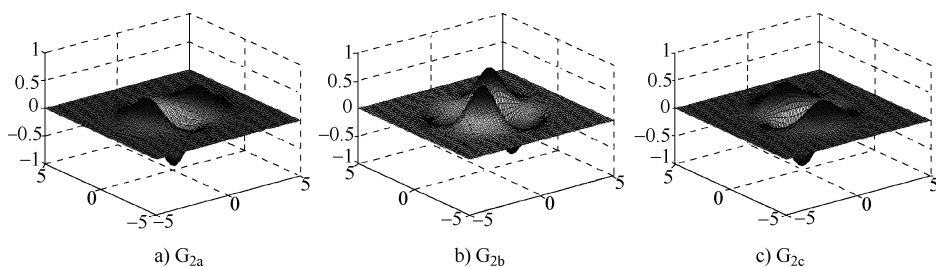


图 7-10 基函数三维图

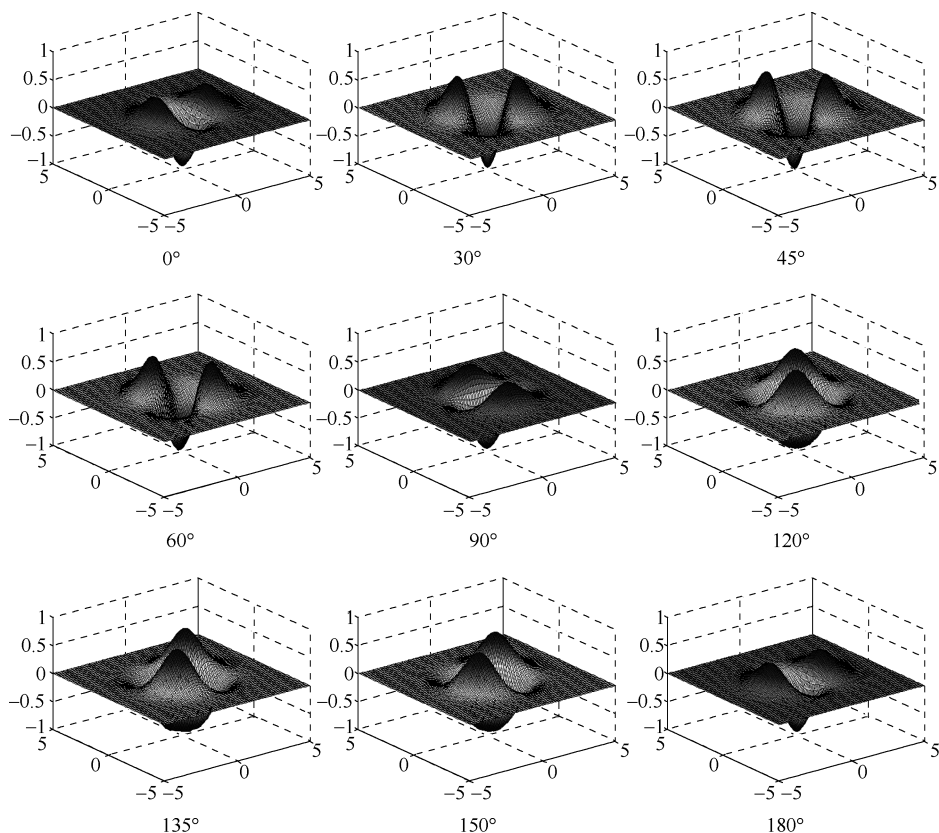


图 7-11 滤波器在不同角度下的三维图

7.2.5.2 基于方向滤波器的车道标志线检测

对感兴趣区域进行方向滤波就是把 ROI 图像 I_R 与 G_2^θ 进行卷积, 得到 f_R

$$\begin{aligned} f_R(x, y) &= I_R(x, y) * G_2^\theta(x, y) \\ &= (I_R * G_{2a})k_a(\theta) + (I_R * G_{2b})k_b(\theta) + (I_R * G_{2c})k_c(\theta) \end{aligned} \quad (7-29)$$

其中 θ 为基角, 取值如下

$$\theta = \arctan(a_i) \quad (7-30)$$

式中 a_i ——上一帧图像中车道标志线拟合直线 $y = a_i x + b_i$ 的斜率;

$i = 1, 2, 3, 4$, 表示从左到右两条车道标志线的轮廓线序号。

根据式 (7-28) 构造 G_{2a} 、 G_{2b} 、 G_{2c} 的 3×3 运算模板如下

$$G_{2a} = \begin{bmatrix} 0.0449 & 0.1221 & 0.0449 \\ 0.1221 & 0.3319 & 0.1221 \\ 0.0449 & 0.1221 & 0.0449 \end{bmatrix}$$



$$G_{2b} = \begin{bmatrix} 0.25 & 0 & 0.25 \\ 0 & 0 & 0 \\ 0.25 & 0 & 0.25 \end{bmatrix}$$

$$G_{2c} = \begin{bmatrix} 0.0449 & 0.1221 & 0.0449 \\ 0.1221 & 0.3319 & 0.1221 \\ 0.0449 & 0.1221 & 0.0449 \end{bmatrix}$$

在感兴趣区域内进行方向滤波，效果如图 7-12 所示，可以看出在标志线感兴趣区域内，标志线被检测出来。

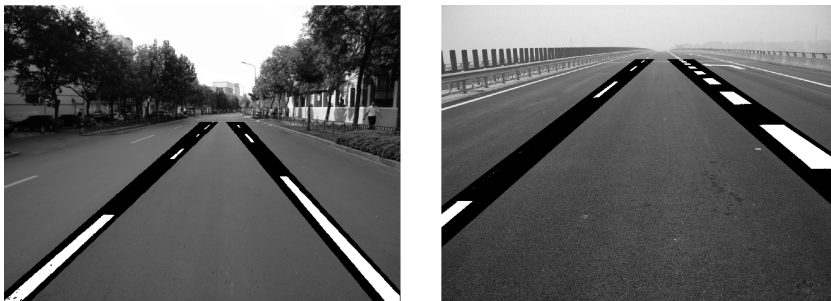


图 7-12 ROI 内方向滤波效果

7.2.5.3 车道标志线特征点确定

检测出车道标志线后，在每个 ROI 内，计算水平方向的梯度 d_H ，定义如下

$$d_H(i, j) = |f(i, j) - f(i, j+1)| \quad (7-31)$$

式中 f ——方向滤波后图像；

i, j ——分别为 ROI 内行坐标和列坐标。

从车道检测图 7-12 中可以看出，水平梯度 d_H 值最大且像素为白色的两点为车道边缘线，该点坐标为车道标志线特征点。为加快运算速度，降低运算时间，本书每隔 5 行计算一次水平梯度来寻找特征点。

7.2.5.4 线性拟合

对于求得特征点，通过最小二乘法进行线性拟合，重建车道线，拟合方法如前所述。

7.2.6 实验分析

部分结构化道路检测结果如图 7-13 所示，其中图 7-13a、图 7-13b 为市区内结构化较好车道检测结果，图 7-13c、图 7-13d、图 7-13e、图 7-13f 为高速公路车道检测效果。可以看出该检测方法能很好地检测出每条车道线的外边缘。

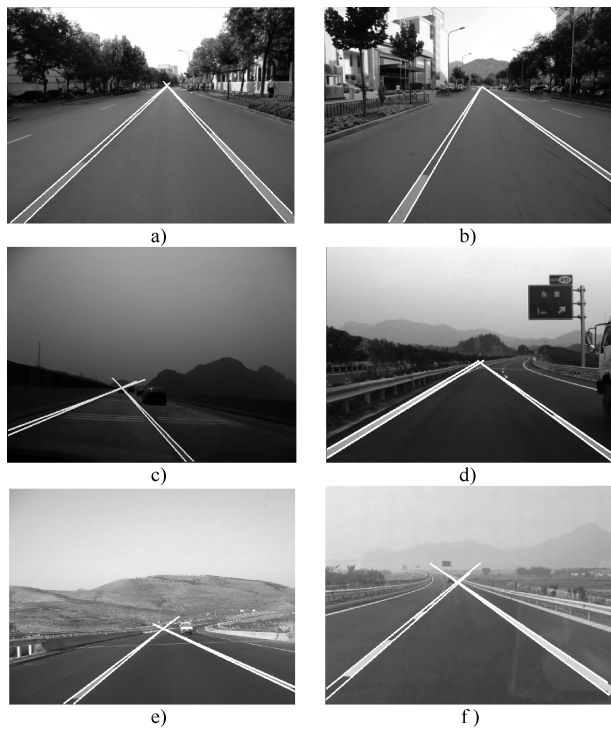


图 7-13 结构化道路检测图

表 7-2 所示为结构化道路的检测结果，可以看出，高速公路的检测率高于城市道路的检测率，主要因为高速公路的车道线清晰、车距较大；而市内道路标志线清晰度不够及车距较小，道路两侧存在建筑物和树木的影响。总体上看，由于结构化道路路面干扰少，只检测路面上的标志线，检测效果良好。造成检测失败的原因主要可归纳为以下几种情况：①道路标志线遭到破坏。②车道中间还有其他的标志线或符号。③光照影响，容易在路面形成镜面反射。④其他车辆影响。⑤道路存在积水或杂物，遮住车道线，无法检测。

表 7-2 结构化道路检测结果

路 型	总 数	正 确 数	错 误 数	正 确 率
高速公路	6748	6421	327	95.2%
城市道路	5890	5448	442	92.5%

7.3 非结构化道路检测

在检测非结构化道路之前，首先根据实际情况，做出如下假设：





- 1) 路面颜色相近, 像素值相差不大。
- 2) 路面颜色与周围颜色不同。
- 3) 路面图像位于整幅图像的下半部分。
- 4) 路面图像面积在整幅图像中所占比例较大。

假设 1)、2) 符合常识; 假设 3)、4) 与摄像机拍摄角度有关, 通过调整可以达到要求, 因此, 这些假设是合理正确的。

本节采用的是基于道路特征的检测方法。首先, 将图像进行重采样, 减小图像尺寸, 降低计算数据量; 然后将图像进行大尺度平滑, 过滤掉大量的纹理信息, 去除纹理对检测算法的干扰; 接着在 HSI 空间中对 H 分量做直方图, 通过爬山法寻找直方图的峰点, 并进行初始图像分割, 得到进行模糊 C-均值 (FCM) 聚类算法所需的输入量, 并进行精确图像分割聚类; 然后将分割后的图像进行数学形态运算, 进一步消除干扰; 最后根据道路特征识别出路面图像, 并进行图像恢复以消除由于图像重采样造成的定位误差。

7.3.1 图像预处理

7.3.1.1 图像重采样

由于道路在图像中占较大比例且道路颜色较为均匀, 并且在道路检测过程中不涉及图像纹理方面的计算, 因此, 为简化计算量, 将图像进行等间隔重新采样, 采样规则如下

$$I(x', y', z) = I\left(\text{Round}\left(\frac{x'}{\eta}\right), \text{Round}\left(\frac{y'}{\eta}\right), z\right) \quad (7-32)$$

式中 η ——图像缩放因子, $0 < \eta < 1$ 图像缩小, $\eta > 1$ 图像放大, $\eta = 1$ 图像尺寸不变;

(x', y', z) ——重采样后彩色图像像素坐标值;

Round ——取整函数;

$z = 1、2、3$, 彩色图像分量通道。

在图 7-14 中表示了 $\eta = 0.7$, $\eta = 0.5$ 时图像与原始图像 (图 7-15a) $\eta = 1$ 的 H 通道的直方图 (另外两通道类似), 该直方图已通过高斯滤波, 可以看出, 图像缩小后, 直方图的形状及像素强度分布没有发生变化, 这说明图像的性质没有变化, 只是像素点个数减少了, 因此, 图像缩小后不会影响后续道路的检测。图 7-15b 为 $\eta = 0.5$ 时, 图 7-15a 的重采样图像。

7.3.1.2 图像大尺度平滑

由于图像中包含丰富的细节信息 (纹理), 这对道路的检测会造成相当大的干扰, 甚至会完全检测不出道路。非结构化道路检测不需要图像的细节信息, 因此可以去掉图像细节, 平滑待测图像, 将其转换为计算机容易识别理解的形

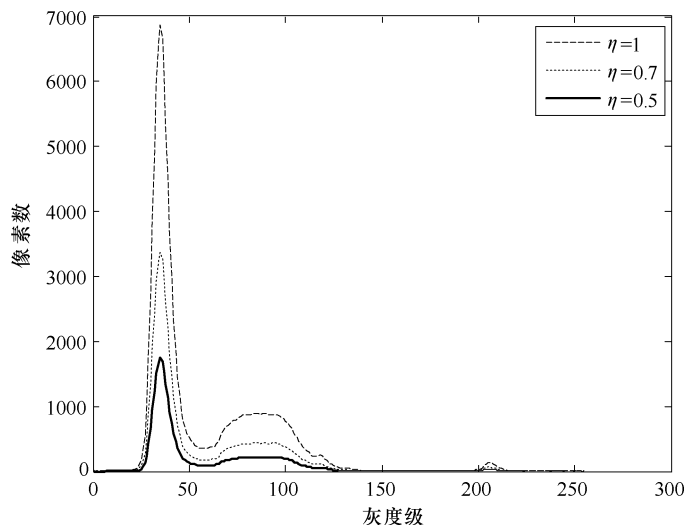


图 7-14 不同图像尺寸下直方图

式（图像在感观上会变得模糊，不细腻，纹理不清楚）。

非结构道路图像的平滑与结构化道路的平滑方法一致，唯一的差别是中心像素点的邻域范围变大了，本节中邻域 Ω 取 9×9 矩阵，具体算法在此不再赘述。

图 7-15c 为对图 7-15b 采用大尺度平滑后的结果，可以看出图像像素变得更加均匀，消除了很多纹理信息。

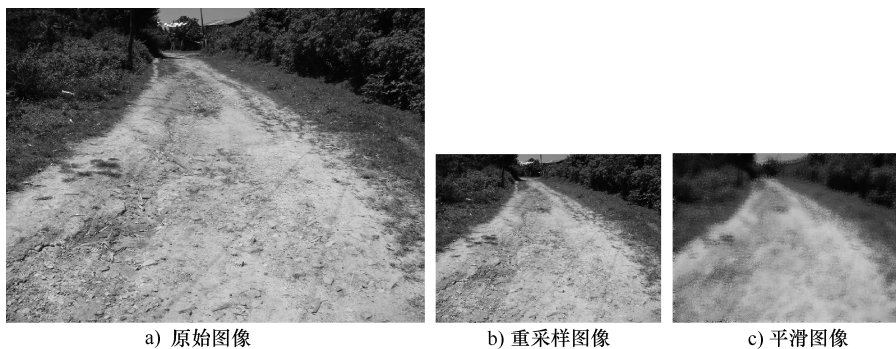


图 7-15 道路图像

7.3.2 彩色图像分割

7.3.2.1 图像空间

因为用 RGB 描述的图像很容易受光照变化的影响，并且 RGB 三个分量之间有很高的相关性，因此直接对 RGB 图像进行图像分割往往得不到所要求的结果。



HSI 彩色模型一方面接近人对颜色的感知, 并受环境光照变化影响较小; 另一方面, 彩色图像的每个均匀彩色区域都对应一相对一致的色调 (H), 仅用色调信息即可进行彩色图像分割。

7.3.2.2 图像分类

1. 直方图极值点

直方图统计 256 级灰度图像的灰度在某个灰度级上的像素个数, 是一种非常重要的常用图像处理方法, 显示了整幅图像像素值的分布情况。通过寻找直方图的峰点和谷点作为聚类初始值进行图像分割是一种经典的算法。对于色调 (H) 分量图像直方图, 如图 7-14 所示。

为能准确找出真实的极值点, 去除曲线中的干扰, 用高斯平滑滤波器对直方图曲线平滑, 平滑公式如下

$$\begin{aligned}\overline{H(x)} &= H(x) * g(x, \sigma_g) \\ &= \int_{-\infty}^{\infty} H(u) g(x - u, \sigma_g) du \\ &= \int_{-\infty}^{\infty} H(u) \frac{1}{\sqrt{2\pi}\sigma_g} e^{-\frac{(x-u)^2}{2\sigma_g^2}} du\end{aligned}\quad (7-33)$$

式中 $H(x)$ ——原始直方图;

$\overline{H(x)}$ ——平滑后直方图;

σ_g ——标准协方差, σ_g 越大直方图越平滑。

通过爬山法寻找曲线上所有的峰点和谷点。爬山法是一种属于人工智能的局部择优算法, 是对深度优先搜索算法的改进, 它利用反馈信息生成解的决策寻找极点。爬山法的计算步骤为: 从当前点开始, 与邻域点的值进行比较, 若当前点的值最大, 那么返回当前点, 作为最大值 (即山峰最高点); 反之就用最高的邻域点来替换当前点, 实现向山峰最高处攀爬的目的, 如此循环直到达到最高点。对于寻找谷底算法正好相反。

峰点满足如下条件

$$P_s = (i, (\overline{H(i)} | \overline{H(i)} > \overline{H(i-1)} \& \overline{H(i)} > \overline{H(i+1)})) \quad (7-34)$$

谷点满足如下条件

$$V_s = (i, (\overline{H(i)} | \overline{H(i)} < \overline{H(i-1)} \& \overline{H(i)} < \overline{H(i+1)})) \quad (7-35)$$

式中 P_s ——峰点;

V_s ——谷点

$i=0, 1, \dots, 255$ 为像素值范围。

爬山算法容易陷于局部最优, 根据非结构化道路的实际情况, 做如下约束。

1) 寻找到的各个极值点间横坐标 (像素值) 之差应大于一定阈值, 即两个极值点的颜色不同。



2) 极值点对应像素数量需超过一定值, 即该点颜色对应的像素数较多。

在道路检测中, 假定道路在图像中占较大比例, 因此只需取前 N 个峰点即可, 其中必定包含道路像素值。在此取 $N=5$, 若实际峰值数小于 N , 则以实际数目为准。

2. 区域初始化

通过爬山法得到图像直方图的主要峰点和谷点。定义 P 为主要峰点的像素强度值

$$P = (i_1, i_2, \dots, i_c) \quad (7-36)$$

式中 i ——峰点像素强度值;

c ——峰点个数, $c \leq N$ 。

经爬山算法将直方图分割分成多个区域, 但是仅利用颜色信息进行分割, 并不能保证相同颜色的像素都属于同一个区域, 因此经过颜色分类的像素还需要结合空间信息才能够形成有意义的区域。具体算法如下:

- 1) 形成所有可能的簇质心 V 。
- 2) 将图像点归到最近的簇质心, 并在每个簇质心中形成像素序列 X_i 。
- 3) 去掉像素个数小于 U 的所有簇质心 (U 取图像总像素数的 20%, $U = 20\% m \times n$)。
- 4) 重新分配像素到最近的簇质心类别中。
- 5) 通过像素序列 X_i 更新对应的簇质心点 v_i 。

3. 区域合并

经过区域初始化后, 得到了一个简单的区域分割, 然而某些区域颜色非常接近, 可以合并到一起, 以便简化簇质心。通过基于颜色的相似性进行合并, 最简单的颜色相似性算法为计算欧几里得距离, 假设 $V = (v_1, v_2, \dots, v_c)$ 为簇质心向量, C 为簇质心个数。合并算法如下:

- 1) 设定最大欧几里得距离阈值 D_c 为正整数值, $D_c = 10$ 。
- 2) 通过下式, 计算欧几里得距离 D

$$D(c_j, c_k) = \sqrt{(R_j - R_k)^2 + (G_j - G_k)^2 + (B_j - B_k)^2} \quad (7-37)$$

式中 $1 \leq j$;

$k \leq C$;

$j \neq k$;

R 、 G 、 B ——分别对应三分量的簇质心像素值。

- 3) 将满足 $D(c_j, c_k) < D_c$ 的质心合并。
- 4) 更新簇质心, 返回步骤 2), 直至条件 3) 不满足为止, 完成合并。

7.3.2.3 基于模糊 C 均值聚类的道路图像分割

Bezdek^[22] 在 20 世纪 70 年代提出了模糊 C 均值聚类 (FCM) 算法, 该算法





用值在 $[0, 1]$ 区间的隶属度判别每个样本点属于某个聚类的程度, 同一样本属于各类别的隶属度之和为 1。把 n 个样本向量 V_i 分为 C 个模糊组, 并求每组的聚类中心, 使得非相似性指标中的价值函数达到最小。

Dunn^[23] 依据 Ruspini 定义的集合模糊划分概念, 给出 FCM 算法如下

$$\min J_2(U, V; X) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^2 d_{ik}^2 \quad (7-38)$$

式中 u_{ik} ——样本 X_i 属于第 k 个聚类的程度;

$d_{ik} = \|x_k - v_i\|_A$, 是第 i 个聚类中心与第 k 个样本间的欧几里得距离。

式 (7-38) 中参数满足

$$\begin{aligned} \sum_{i=1}^c u_{ik} &= 1, 1 \leq k \leq n \\ u_{ik} &\in [0, 1], 1 \leq i \leq c, 1 \leq k \leq n \\ 0 &< \sum_{k=1}^n u_{ik} < n, 1 \leq i \leq c \end{aligned}$$

Bezdek^[24] 将式 (7-38) 推广到更一般情形, 给出了如下表达式

$$\min J_m(U, V; X) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m d_{ik}^2 \quad (7-39)$$

式 (7-39) 中参数满足

$$\begin{aligned} \sum_{i=1}^c u_{ik} &= 1, 1 \leq k \leq n \\ u_{ik} &\in [0, 1], 1 \leq i \leq c, 1 \leq k \leq n \\ 0 &< \sum_{k=1}^n u_{ik} < n, 1 \leq i \leq c \end{aligned}$$

式中 $m \in [1, \infty]$ ——加权系数;

J_m ——类内误差的加权平方和目标函数。

FCM 算法通过迭代运算达到目标函数 J_m 的最小化, 此时得到 X 的一个最优模糊 C 划分 $U^* = [u_{ik}^*]$

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)}} \quad (7-40)$$

其中

$$\begin{aligned} d_{ik} &= \|x_k - v_i\|_A, 1 \leq i \leq c, 1 \leq k \leq n \\ v_i &= \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m}, 1 \leq i \leq c \end{aligned}$$



FCM 算法的具体计算过程如下:

1) 初始化。给定分类数 C , 迭代截止误差 $\varepsilon > 0$, 算法的最大迭代次数 T_{\max} , 初始聚类中心 $V = (v_{10}, v_{20}, \dots, v_{c0})$ 。

2) $t = 1, 2, \dots, T_{\max}$ 。

① 计算 $U_t = [u_{ik,t}]_{c \times n}$, 令 $d_{ik,t} = \|x_k - v_t\|_A$ 。如果 $d_{ik,t} = 0$, 则 $u_{ik,t} = 1$, 且对 $j \neq i$, $u_{jk,t} = 0$, $1 \leq i \leq c$, $1 \leq k \leq n$;

如果 $d_{ik,t} > 0$, 则

$$u_{ik,t} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)}}, 1 \leq i \leq c, 1 \leq k \leq n;$$

② 计算 $V_{t+1} = (v_{1,t+1}, v_{2,t+1}, \dots, v_{c,t+1})$

其中

$$v_{i,t+1} = \frac{\sum_{k=1}^n (u_{ik,t}) x^k}{\sum_{k=1}^n (u_{ik,t})^m}, 1 \leq i \leq c$$

③ 判断, 如果 $\|V_{t+1} - V_t\| = \sum_{i=1}^c \|v_{i,t+1} - v_{i,t}\| \leq \varepsilon$, 则终止迭代。

④ 否则进行下一个 t 。

图 7-16 为部分非结构化道路彩色图像分块结果, 可以看出, FCM 方法能按照图像颜色不同将图像分块, 道路区域被完整地分离出来。

7.3.2.4 图像形态运算

1. 开启、闭合运算

开启运算通常可以平滑图像轮廓, 去掉轮廓上的毛刺, 截断狭窄的山谷; 而闭合运算可以去除区域中的小孔, 填平狭窄的断裂, 细长的沟壑以及轮廓的缺口。通过 FCM 分类出的各图像块, 由于外界条件变化及道路的不均匀性导致检测结果中夹杂着较多非道路点, 边缘毛刺较多, 内部存在小孔 (如图 7-16a'、图 7-16c'、图 7-16d'), 因此, 为得到准确的检测结果, 本书采用 3×3 方形结构元先对图像进行开启运算, 然后再进行闭合运算去掉干扰点, 平滑图像。

2. 像素连通性计算

开启、闭合运算结束后, 道路区域能较好的检测出来, 然而除此之外还会有些较大的干扰点, 开启、闭合运算不能消除。通过计算像素的连通性, 利用填充算法可去掉干扰点。本书选用八连通类型进行检测, 与标准区域不在一个连通区域的点记为干扰点。



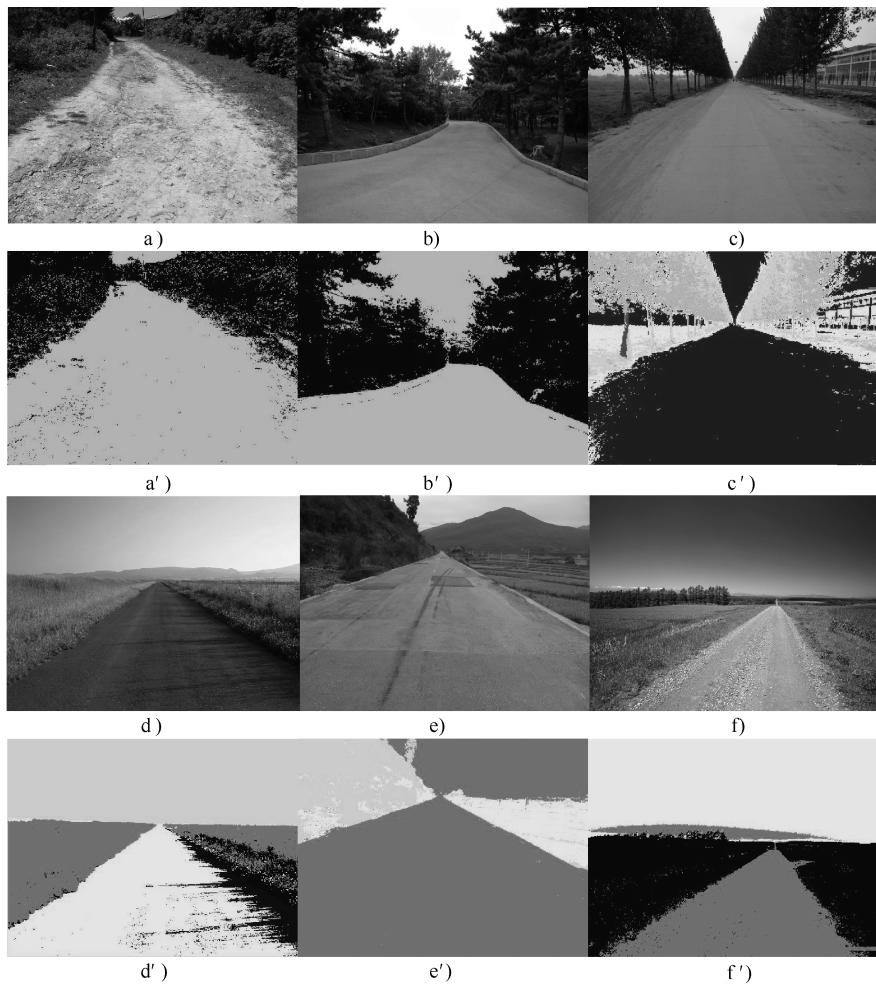


图 7-16 颜色分割效果图

7.3.3 纹理特征分析

将彩色道路图像进行分割后，可以准确地分割出路面图像，然而并不能判定哪块分割区域为路面，在此根据路面图像的假设条件及各分割块的特性进行如下判断。

- 1) 计算分割块标准差 σ_x ，标准差值较小的是道路的可能性较大。
- 2) 计算分割块的质心 M ，质心处在整幅图像的中间偏下位置是路面的可能性较大。
- 3) 计算分割块的面积 S ，面积较大的是路面图像的可能性较大。



综合以上3条规则,分析各图像块,找出路面图像。

图7-16a'中图像块特性如下表7-3所示:

表7-3 图像块特性

序号	标准差 $\sigma_x(R, G, B)$	质心 M	面积 S
1	(0.2941, 0.3001, 0.4432)	(60, 197)	38067
2	(0.1200, 0.1103, 0.1168)	(197, 201)	79577
3	(0.8404, 0.8833, 1.1289)	(11, 189)	1938

序号与图像块对应关系:1号对应中间黑色区域;2号对应下方道路区域;3号对应中间上方白色区域。从表7-3中可以看出道路区域的标准差小于其他两块区域的标准差;道路图像块质心位置符合实际情况,处于中间偏下;道路面积较大,符合假设。因此,完全可以利用这三个特性,检测出道路区域,为提高准确性,需综合考虑。

7.3.4 图像恢复

7.3.4.1 图像重采样误差分析

原始图像大小为 $m \times n$,重采样后图像大小为 $\eta m \times \eta n$ 。行方向上用 ηm 个像素点代替原来的 m 个像素点;列方向上用 ηn 个像素点代替原来的 n 个像素点。采样后丢失大量信息,每个像素代替原来的 $m/\eta m$ ($n/\eta n$) 个像素点,也就是在图像恢复后,会有 $1/\eta$ 个像素点的误差 ($0 < \eta < 1$)。

7.3.4.2 图像恢复

由于图像重采样,将图像缩小后图像会丢失部分信息,这样就给道路最终的检测、定位带来一定的误差,为消除误差,将图像恢复。具体规则如下:

- 1) 取原采样率的倒数为新采样率 $\eta' = 1/\eta$,将图像恢复。
- 2) 以恢复后的道路轮廓点为中心像素点,在 $2/\eta \times 2/\eta$ 邻域内,计算原图像中像素值的欧几里得距离,距离最大且像素点与新轮廓点满足八邻域规则的点为下一个新的轮廓点,遍历所有轮廓点,得到新轮廓。

7.3.5 边缘拟合

道路边缘拟合方法与结构化道路的拟合方法完全相同,在此不再赘述。图7-17为部分非结构化道路拟合直线实例。

7.3.6 实验分析

本书对沙土路、水泥路、柏油路等路型进行了实际测试,部分道路的检测效果如图7-17所示,实验结果如表7-4所示。





图 7-17 道路拟合曲线

从表 7-4 中可以看出，沙土路的检测率最低，水泥路次之，柏油路最高。沙土路检测失败的原因有：①道路与周围环境颜色相近，不易区分；②道路中存在水迹等影响；③道路不均匀，色彩不统一；④光照影响，强光照射，图像超出显示范围，无法提供有效色彩信息；⑤道路轮廓极不规则，无法拟合出有效直线；⑥树荫影响，路面出现斑驳的阴影；⑦天气原因。

表 7-4 非结构化道路检测结果

路型	总数	正确数	错误数	正确率
沙土路	3800	3269	531	86.0%
水泥路	4266	3839	427	89.9%
柏油路	3690	3468	222	94.0%

水泥路和柏油路由于比沙土路具有更好的结构特征，路面比沙土路均匀，因此检测率比沙土路高，之所以柏油路的检测率比水泥路的高，是因为柏油路的色彩与环境的对比更加明显。其检测失败的原因主要有以下几点：①道路中存在水迹等影响；②道路破损严重，路面整体色彩极不均匀；③光照影响，强光照射，图像超出显示范围，无法提供有效色彩信息；④天气原因。

7.4 车辆偏离分析

7.4.1 车辆轨迹模型

由于前面已经假设车辆行驶前方较近区域为直线车道，因此车辆行驶轨迹



有以下几种情况,如图 7-18 所示。

图 7-18a 为车辆在车道线中间直线行驶,这是一种理想情况,现实生活中由于驾驶员驾驶习惯决定车辆在中间线附近直线行驶,如图 7-18b、图 7-18c 所示,但由于安全等原因,不会特别靠近左右车道线;而图 7-18d、图 7-18e 为车辆斜线行驶即车辆行驶方向与车道线方向不同,之间存在夹角,这种情况除停车外大多是由于驾驶员对车辆失去有效控制而引起的,极易发生交通事故。

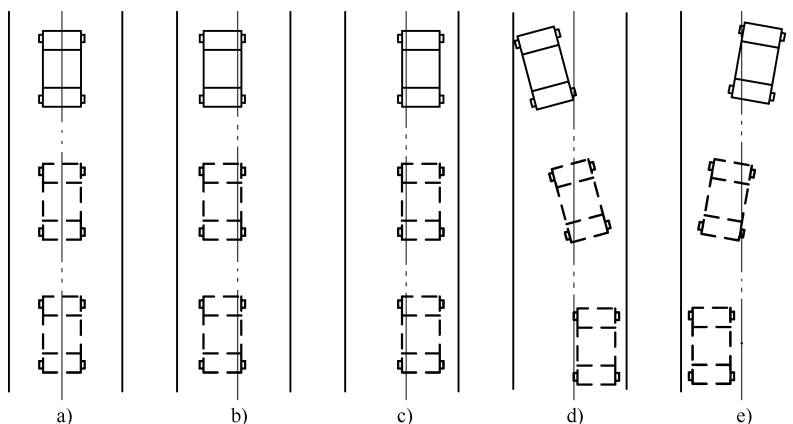


图 7-18 车辆行驶轨迹

可以得出车辆行驶轨迹不安全的情况:

- 1) 车辆虽沿车道线方向直线行驶,但车辆严重偏离理想轨迹,离车道线太近甚至于压线行驶。
- 2) 车辆沿车道线斜线方向行驶,按此方向行驶即将驶出车道。

本书将对以上这两种情况进行检测,对于第一种情况,采用车道偏离率来表示车辆当前偏离车道的程度;对于第二种情况,采用车道偏离量变化率来反映车辆在垂直车道线方向上的速度。

7.4.2 车道偏离率

车道偏离率 ξ ^[25]反映了车辆偏离车道的程度,其计算公式如下

$$\begin{aligned}\xi &= d_l/d_r \\ &= \left(\frac{\pi}{2} - \theta_L\right) / \left(\theta_R - \frac{\pi}{2}\right)\end{aligned}\quad (7-41)$$

式中 d_l 、 d_r ——摄像机距左右车道线的垂直距离, $d_l + d_r = D$;

D ——车道宽度;

θ_L 、 θ_R ——分别表示左、右车道线的倾角(如图 7-19 所示)。



θ_L 、 θ_R 按式 (7-42) 求解

$$\theta_i = \begin{cases} \theta'_i & \text{if } \theta'_i \geq 0 \\ \theta'_i + \pi & \text{if } \theta'_i < 0 \end{cases} \quad (7-42)$$

式中 $\theta'_i = \arctan((\overline{a_i})/2)$;

i ——代表 L 、 R ，表示左右车道；

$\overline{a_i}$ ——对应车道线的平均斜率， $\overline{a_L} = (a_1 + a_2)/2$ ， $\overline{a_R} = (a_3 + a_4)/2$ ；

a_1 、 a_2 、 a_3 、 a_4 ——分别为左、右车道线边缘拟合直线的斜率。

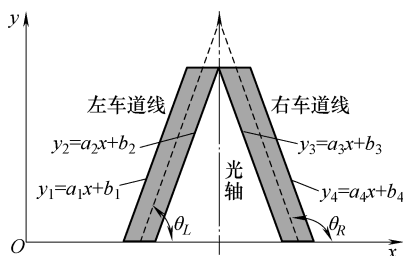


图 7-19 车道示意图

式 (7-41) 需满足的条件为车辆沿车道线方向直线行驶，但当驾驶员失去了对车辆的有效控制，车辆在车道中沿斜线方向行驶时，根据文献 [26] 的证明，车辆的车道偏离率同样可以用该公式近似。

车辆行驶时光轴与车道线倾角变化情况如图 7-20 所示，车道偏离率 ξ 及车道线倾角具有以下几种变化规律：

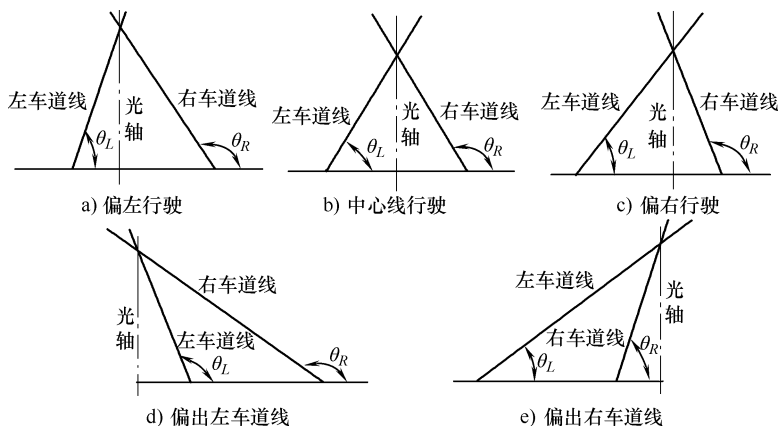


图 7-20 车辆偏离示意图

1) 当车辆靠近左车道线行驶时 (图 7-20a 所示): θ_L 变大且 $\theta_L \in (0, \pi/2)$, θ_R 变大且 $\theta_R \in (\pi/2, \pi)$, 此时, ξ 变小 $\xi \in (0, 1)$ 。



- 2) 当车辆靠近车道中心行驶时 (图 7-20b 所示): $\theta_L + \theta_R = \pi$, $\xi \approx 1$ 。
- 3) 当车辆靠近右车道时 (图 7-20c 所示): θ_L 变小且 $\theta_L \in (0, \pi/2)$, θ_R 变小且 $\theta_R \in (\pi/2, \pi)$, 此时, ξ 变大 $\xi \in (1, \infty)$ 。
- 4) 当车辆偏出左车道线时 (图 7-20d 所示): $\theta_L \in (\pi/2, \pi)$, $\theta_R \in (\pi/2, \pi)$, $\theta_L < \theta_R$, $\xi \in (-1, 0)$ 。
- 5) 当车辆偏出右车道线时 (图 7-20e 所示): $\theta_L \in (0, \pi/2)$, $\theta_R \in (0, \pi/2)$, $\theta_L < \theta_R$, $\xi \in (-\infty, -1)$ 。

车道偏离率 ξ 与车道偏离位置之间的关系如图 7-21 所示。

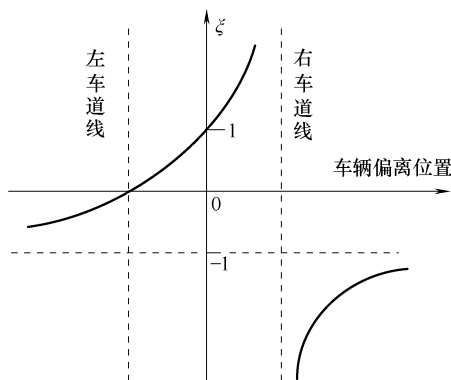


图 7-21 ξ 取值范围示意图

从图 7-21 中看出, ξ 值与车辆位置关系一一对应, ξ 的大小可以反映出车辆的偏离情况, 因此, 可将 ξ 作为衡量车道偏离程度的一个重要依据。

7.4.3 车道偏离量变化率

对于车辆沿斜线方向行驶的情况, 可以通过车道偏离量变化率 ζ 来判断, 当 ζ 值较大则说明车辆正在迅速的靠近 (远离) 某车道线, 发生了车辆沿斜线行驶的情况。 ζ 定义如下

$$\begin{aligned}\zeta &= \frac{|d_p(t) - d_p(t-1)|}{D} \\ &= \left| \frac{1}{2} \left[\frac{\xi(t) - 1}{\xi(t) + 1} - \frac{\xi(t-1) - 1}{\xi(t-1) + 1} \right] \right|\end{aligned}\quad (7-43)$$

式中 $d_p(t)$ ——车辆在当前时刻的偏离距离;

$d_p(t-1)$ ——车辆在上一个时刻的偏离距离;

D ——车道宽度;

$\xi(t)$ ——当前时刻的车道偏离率;

$\xi(t-1)$ ——上一个时刻的车道偏离率;



t ——当前时间变量。

7.4.4 实验分析

图 7-22 为车道偏离率 ξ 随视频序列 F 的变化曲线图，图 7-22a 为车辆沿车道中心正常驾驶时的曲线，可以看出 ξ 值在 1 附近波动，波动量不大；图 7-22b 为车辆按“S”形轨迹前进，并在行进过程中，偏离出左右车道；图 7-22c 为车辆从左车道线外进入车道后又偏离出右车道线的偏离率变化曲线；图 7-22d 为车辆从右车道线外进入车道后又偏离出左车道线的偏离率变化曲线。可以看出 ξ 值随车辆位置的不同出现较大的变化， ξ 值完全可以作为一个评价车辆当前位置的指标。

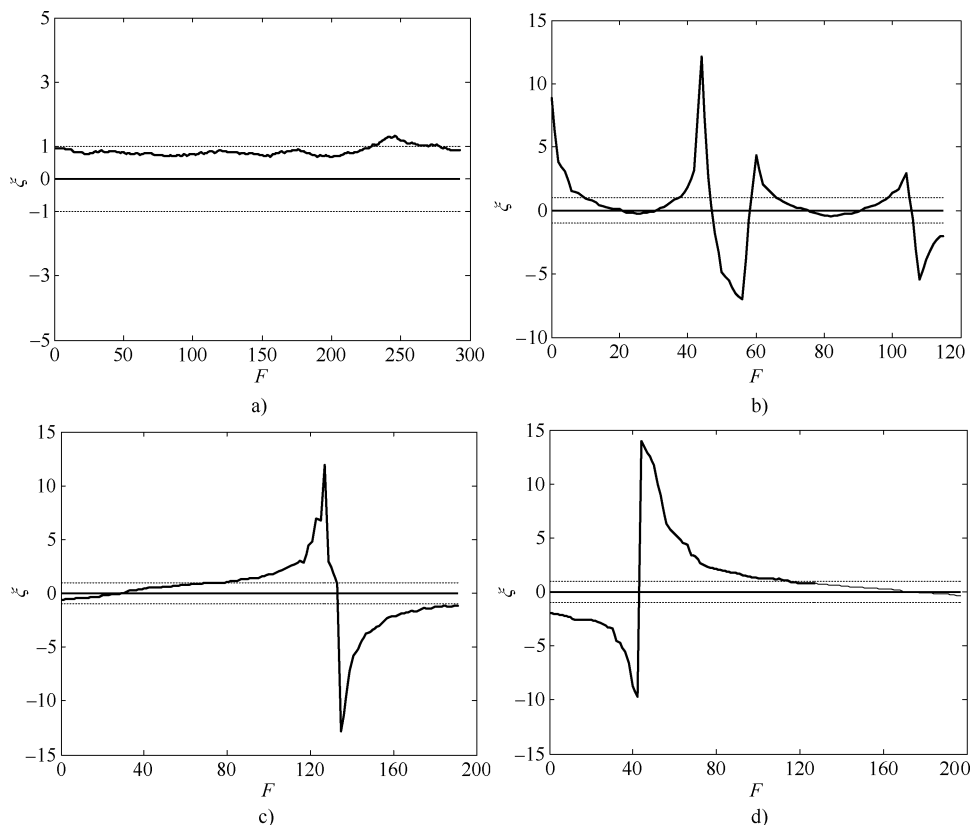
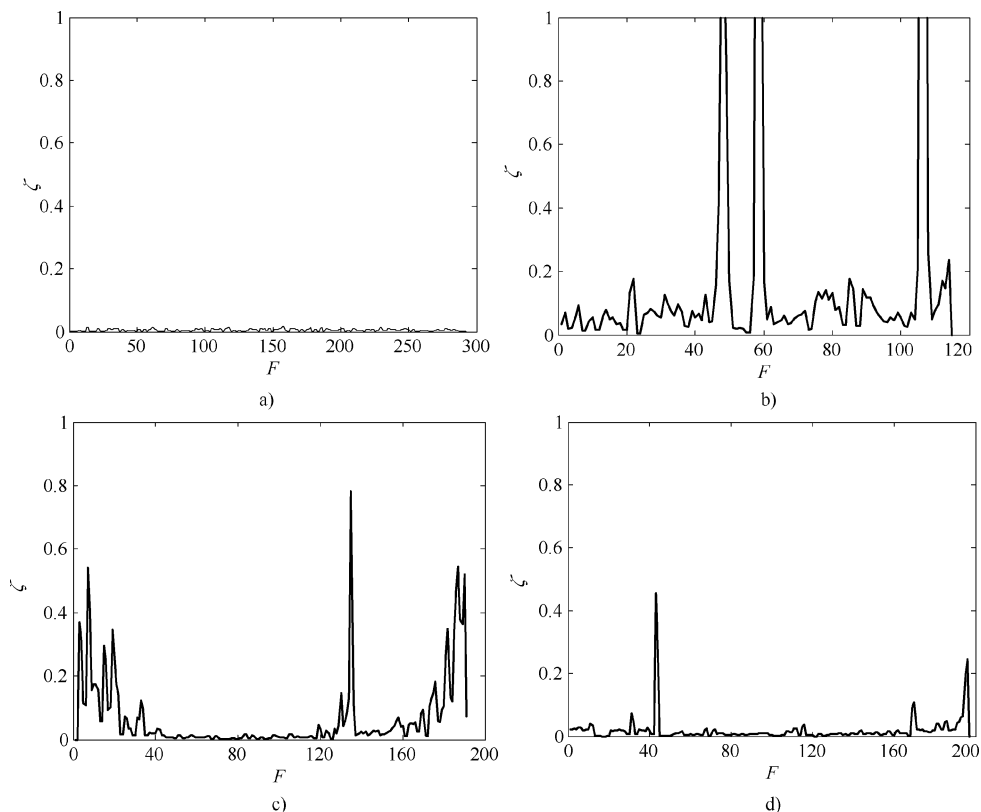


图 7-22 ξ 变化曲线

图 7-23 为图 7-22 对应车辆行驶轨迹的车道偏离量变化率 ζ 的曲线图，从中可以看出在偏离率 ξ 变化剧烈的位置， ζ 对应值就大，说明该处车辆行驶方向与

图 7-23 ζ 变化曲线

车道线方向并不平行, ζ 表征了车辆的行驶状态。

7.5 本章小结

本章主要研究了结构化道路及非结构化道路的检测和车辆行驶状态判断等问题, 具体总结如下:

1) 结构化道路检测: 在检测初期, 对道路图像实施平滑及灰度拉伸预处理, 以均匀局部区域像素值, 并加大车道线与道路的对比度。由于结构化道路车道标志线明显, 根据道路视频图像的特点, 采用图像分块投影的方法检测车道线边缘; 对检测出的边缘点再次进行微调, 提高检测精度; 然后, 采用最小二乘法进行线性拟合, 得到车道边缘线。该方法的优点是抗干扰能力强, 不易陷入局部最优。

2) 非结构化道路检测: 非结构化道路的复杂性决定了图像中的像素值波动





较大,同时对道路的检测不需要丰富的纹理信息,因此,对图像进行了大尺度平滑,尽量减少纹理对检测的干扰;然后,采用FCM方法将彩色道路图像分块,道路区域被完整、独立地分为一个图像块;根据道路图像特点分析各图像块判断出道路区域;最后对道路边缘进行线性拟合。本节还研究了图像尺寸对检测结果的影响。

3) 车道偏离计算分析:本书采用车道偏离率和车道偏离量变化率来表示车辆的行驶轨迹。车道偏离率反映了车辆当前的行驶位置,车道偏离量变化率反映了车辆偏离量的变化情况,根据此值可以预测出车辆的行驶方向。

参考文献

- [1] Aufrere R, Chapuis R, Chausse F. A Model Driven Approach for Real-time Road Recognition [J]. Machine Vision and Application. 2001, 13 (1): 95-107.
- [2] Chapuis R, Aufrere R, Chausse F. Accurate Road Following and Reconstruction by Computer Vision [J]. IEEE Transactions on Intelligent Transportation Systems. 2002, 3 (4): 261-270.
- [3] Kuan D, Phipps G. Autonomous Robotic Vehicle Road Following [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1988, 10 (5): 648-658.
- [4] Chen M, Jochem T, Pomerlean D. Aurora: a Vision-based Roadway Departure System [C]. Proc. IEEE Conf. Intell. Robots and Systems. 1995: 243-248.
- [5] Kluge K, Lakshmanan S. A Deformable-template Approach to Lane Detection [C]. Proc. IEEE Intell. Vehicle Symp.. 1995: 54-59.
- [6] Bertozzi M, Broggi A. Gold: a Parallel Real-time Stereo Vision System for Generics Obstacle and Lane Detection [J]. IEEE Trans. Image Process. 1998, 7 (1): 62-81.
- [7] Kreucher C, Lakshmanan S. Lana: a Lane Extraction Algorithm that Uses Frequency Domain Features [J]. IEEE Trans. Robot. Autom.. 1999, 15 (2): 343-350.
- [8] Wang Y, Teoh E K, Shen D. Lane Detection Using B-snake [C]. Int. Conf. Information Intelligent and Systems, Bethesda, MD, USA. 1999: 438-443.
- [9] Wang Y, Shen D, Teoh E K. Lane Detection Using Spline Model [J]. Pattern Recognit. Lett.. 2000, 21 (8): 677-689.
- [10] Yim Y U, OH S Y. Three-feature Based Automatic Lane Detection Algorithm (TFALDA) for Autonomous Driving [J]. IEEE Trans. Intell. Transp. Syst.. 2003, 4 (4): 219-225.
- [11] Mccall J C, Trivedi M M. Video-based Lane Estimation and Tracking for Driver Assistance: Survey, System, and Evaluation [J]. IEEE Trans. Intell. Transp. Syst.. 2006, 7 (1): 20-37.
- [12] 管琰平, 贺跃, 刘培志等. 基于彩色图像的非结构化道路检测 [J]. 计算机应用, 2005, 25 (12): 2931-2934.



- [13] Crisman J D, Thorpe C E. Unscarf. A Color Vision System for the Detection of Unstructured Roads [C]. Proceedings of the IEEE International Conference on Robotics and Automation. 1991, 04: 2496-2501.
- [14] Zhang J Y, Nagel H H. Texture-Based Segmentation of Road Images [C]. Proceedings of the Intelligent Vehicles' 94 Symposium. 1994. 10: 260-265.
- [15] Manduchi R. Bayesian Fusion of Color and Texture Segmentations [C]. The Proceedings of the Seventh IEEE International Conference on Computer Vision. 1999. 09: 956-962.
- [16] 苏开娜, 任文君, 易小琳等. 基于运动模型的道路识别与跟踪算法的研究 [J]. 中国图象图形学报, 2000, 5 (3): 226-230.
- [17] 李青, 郑南宁, 马琳等. 基于主元神经网络的非结构化道路跟踪 [J]. 机器人, 2005, 27 (5): 247-251.
- [18] 交通部第一公路勘察设计院. JTJ 011—1994 公路路线设计规范 [S]. 北京: 人民交通出版社, 1994.
- [19] 沈岷, 毛建国, 李舜酩. 单目视觉车道标志线识别算法及其 ARM 实现 [J]. 南京航空航天大学学报, 2008, 40 (20): 209-213.
- [20] 贾云得. 机器视觉 [M]. 北京: 科学出版社, 2000: 65-67, 70.
- [21] Freeman W T, Adelson E H. The Design and Use of Steerable Filters [J]. IEEE Trans. Pattern Analysis and Machine Intelligent. 1991, 13 (9): 891-906.
- [22] Bezdek J C. Aphysical Interpretation of Fuzzy ISODATA [J]. IEEE Transaction on SMC. 1976, 6: 387-390.
- [23] Dunn J C. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well Separated Clusters [J]. Journal of Cybernetics. 1974, 3 (3): 32-57.
- [24] Bezdek J C. Pattern Recognition with Fuzzy Objective Function Algorithms [M]. New York : Plenum Press. 1981.
- [25] Joon Woong Lee, Chang Doo Kee, Un Kun Yi. A New Approach for Lane Departure Identification [C]. Intelligent Vehicles Symposium, 2003. Proceedings. IEEE. 2003: 100-105.
- [26] 余厚云, 张为公. 基于摄像机模型的运动车辆车道偏离检测 [J]. 东南大学学报, 2009, 39 (5): 933-936.

基于信息融合的驾驶员疲劳检测

8.1 概述

驾驶员视觉分散的一种极端情况是驾驶员出现疲劳（瞌睡）状态。此时，驾驶员的眼睛出现闭合现象，不能很好地检测到视线方向。针对此问题我们提出对驾驶员疲劳检测进行研究。只用一种方法或者一个标准来判断驾驶员的疲劳状态，存在着众多不足。首先，单一传感器获取的信息量不足，由于信息量太少而不能全面表达出驾驶员的疲劳程度；其次，单一的疲劳检测手段极易受外界环境的干扰，检测结果也会受到影响，系统鲁棒性差。为克服以上缺点，本章将用多传感器测量多个特征，采用不同的判断标准进行疲劳检测，最后将检测信息进行融合得出判断结果。

8.2 信息融合

信息融合是人类或其他逻辑系统中常见的功能，人类能把来自人体各个传感器（如眼、耳、鼻、四肢等）的信息（如视觉、听觉、味觉、触觉等）组合起来，并使用先验知识去估计、理解周围环境，最终做出相应判断。由于社会的不断进步，使得信息具有表现形式的多样性，信息关系的复杂性，信息数量的巨大性，以及要求信息处理的实时性等特点，这些都已超出人脑的综合处理能力。因此，从 20 世纪 70 年代起，多传感器信息融合技术（Multi-Sensor Data Fusion, MSDF）作为一门新兴的学科便迅速地发展起来。多信息融合的定义可概括为：将来自多个传感器或多源的信息进行综合处理分析，以便得到更为准确、可靠的结论^[1-3]。

信息融合技术主要具有以下优点：

- 1) 信息融合可改善检测性能，增加系统可信度。通过对多种测量的有效融合，提高了检测的有效性，降低了事件的不确定性。
- 2) 信息融合可加强系统的生存能力。多传感器的测量信息之间存在一定的



冗余度,在有若干传感器损坏或受到干扰而不能正常工作时,总会存在一部分传感器可以提供信息,使系统能够继续正常运行。

3) 信息融合扩大了时间和空间的覆盖范围。多传感器可以在不同时间从多方面对目标进行检测,扩大了时空覆盖范围,进而增加了系统的检测能力。

信息融合按照不同标准具有不同的分类方法,例如按照融合技术、融合算法、融合结构分类等^[4]。

信息融合由于是对多信息的综合处理而使其具有本质上的复杂性。根据传统的算法理论,近年来开发出一些新的基于统计推断、人工智能以及信息论的新方法^[5],其主要内容介绍如下。

1. 信号处理与估计理论方法

该方法主要包括小波变换、加权平均、卡尔曼滤波、最小二乘法等线性估计理论,以及扩展卡尔曼滤波、高斯滤波等非线性估计理论等^[6-9]。近年来,粒子滤波^[10]和马马尔科夫链^[11]等非线性估计技术已受到越来越多学者的重视。期望极大化算法^[12]为求解在观测数据不完整情况下的参数估计与融合问题,提供了新的思路。通过建立一系列的参数优化指标,应用最优化方法获得参数的最优估计,代表算法主要有极小化风险法^[13]和极小化能量法^[14]。

2. 信息论方法

信息论方法通过运用优化信息度量的方法融合多源数据,代表算法主要有熵方法和最小描述长度法^[15]。

3. 人工智能方法

人工智能方法主要包括遗传算法、模糊逻辑、神经网络、逻辑模板法、品质因数、专家系统以及基于规则的推理等,这些方法已在信息融合领域中取得了一定的成果^[16]。

4. 决策论方法

该方法常在高级别的决策融合中应用。文献 [17] 借助决策论方法融合可见光、红外以及雷达数据用于报警分析。

5. 统计推断方法

统计推断方法主要包括随机集理论、贝叶斯推理、证据推理、经典推理以及支持向量机等^[18]。

6. 几何方法

几何方法是根据环境以及传感器模型的几何属性来实现多传感器的信息融合。文献 [19] 通过对不确定椭圆球体积进行极小化的几何方法实现多信息融合。





8.3 基于粗糙集理论的驾驶员疲劳检测

8.3.1 粗糙集理论

粗糙集 (Rough Set, RS) 理论是由波兰数学家 Z. Pawlak 在 20 世纪 80 年代初提出的, 是处理不确定和模糊知识的数学工具^[20]。

8.3.1.1 知识与决策系统

知识是对对象的一个分类, 为了处理知识而采用决策系统来标识知识。决策系统是由对象组成的, 对象的属性描述了其具有的信息。决策系统 S 可表示为

$$S = (U, C, D, V, f) \quad (8-1)$$

式中 U ——论域, 是对象的有限集合, $U = \{x_1, x_2, \dots, x_n\}$;

C ——条件属性值;

D ——决策属性值;

$R = C \cup D$, 是属性集合, $C \cap D = \emptyset$;

$V = \bigcup_{r \in R} V_r$, 是属性值的集合, V_r 表示属性 $r \in R$ 的值域;

$f: U \times R \rightarrow V$, 是信息函数, 为 U 中各对象 x 的属性指定唯一值。

可用表格法来表示知识, 行表示对象, 列表示属性, 一个表可作为一个知识库, 知识库和知识表达系统间一一对应。决策表定义如下:

$S = (U, A)$ 表示形式化的知识表达系统, 且 C 为条件属性 $C \subset A$, D 为决策属性 $D \subset A$, 决策表 $T = (U, A, C, D)$ 简称为 CD 决策表。设 X 是 U 中根据条件属性 C 定义的分类, Y 是 U 中根据决策属性 D 定义的分类, 对每个 x_i, y_i 定义如下函数

$$d_x: Des_C(x_i) \rightarrow Des_D(y_i): x_i \cap y_i \neq \emptyset \quad (8-2)$$

式中 d_x ——决策表 T 中的决策规则;

如果满足当 $x \neq y$, $d_x/C = d_y/C$ 时, 有 $d_x/D = d_y/D$, 则称决策规则是协调的, 否则为不协调的。

8.3.1.2 粗糙集

令 $X \subseteq U$, R 为 U 上一等价 (不可分辨) 关系, 当 X 能用 R 描述时, 称为 X 是 R 可定义的, 否则 X 是 R 不可定义的。 R 粗糙集可用上近似集和下近似集这两个精确集来描述, 具体如下

$$R_*(X) = \{X \in U: R(X)X\} \quad (8-3)$$

$$R^*(X) = \{X \in U: R(X) \cap X \neq \emptyset\} \quad (8-4)$$

式中 $R_*(X)$ ——下近似集;



$R^*(X)$ ——上近似集；

U ——全域；

R —— U 的等价关系；

$R(X)$ ——包含 X 的等价， $X \subseteq U$ 。

集合 $Bn_R(X) = R^*(X) - R_*(X)$ 称为 X 的 R 边界域， $Pos_R(X) = R_*(X)$ 称为 X 的 R 正域， $Neg_R(X) = U - R^*(X)$ 称为 X 的 R 负域。 $Pos_R(X)$ 是由根据知识 R 判断肯定属于 X 的 U 中元素组成的集合， $R^*(X)$ 是根据知识 R 判断可能属于 X 的元素组成的集合， $Bn_R(X)$ 是根据知识 R 无法判断是否属于或不属于 X 的元素组成的集合， $Neg_R(X)$ 是根据知识 R 判断肯定不属于 X 的元素组成的集合。

8.3.1.3 属性的依赖性和重要性

在决策系统 S 中，条件属性 C 与决策属性 D 之间的关系非常重要，决策属性 D 对条件属性 C 的依赖度定义为

$$k = \gamma_C(D) = \frac{Pos_C(D)}{|U|} \quad (8-5)$$

式中 $Pos_C(D)$ ——决策属性 D 在条件属性 C 下的正域。

依赖度 k 表示了条件属性 C 下能确切划入决策类 U/D 的对象比率，表示了决策属性对条件属性的依赖程度。

8.3.1.4 决策表约简与求核

决策表约简的目的是化简条件属性，约简后的决策表在决策功能不变的前提下具有最少条件属性。约简主要分为以下两类。

1. 基于决策表的约简

(1) 对于协调决策表

设 (C, D) 为协调决策表， $a \in C$ ，如果 $(C - \{a\}, D)$ 协调，则属性 a 可以去掉，否则不可。如果所有属性 $a \in P$ 是 (C, D) 中不可省略的，则称 (C, D) 是独立的；如果属性子集 $R \subseteq C$ ， (R, D) 是独立协调的，则称 (R, D) 为 (C, D) 的约简，所有约简的集合称为核 $Core(C, D)$ 。

(2) 对于不协调决策表

设 (C, D) 为不协调决策表， $a \in C$ ，当满足 $Pos_C(D) = Pos_{C - \{a\}}(D)$ 时，属性 a 可以省略，否则不可省略。

2. 基于决策规则的约简

决策规则的约简是通过去掉每个决策规则的不必要条件来实现的。假设 $\theta \rightarrow \psi$ 为 CD 规则， $a \in C$ ，当且仅当： $\theta \rightarrow \psi$ 为真蕴含 $\theta | (C - \{a\}) \rightarrow \psi$ 为真时，属性 a 是可以省略的，否则不可省略。





8.3.2 驾驶疲劳相关条件属性分析

8.3.2.1 驾驶疲劳分析

1. 疲劳含义

疲劳又称疲惫，是一种主观不适的感觉，客观上会失去其完成原来所从事的正常活动或工作的能力^[21]。疲劳可分为以下几类。

(1) 体力疲劳

体力疲劳主要是因为人体持续了较长时间和较大强度的体力劳动而引起乳酸及其他代谢产物的堆积，造成肌肉张力下降及运动耐久性降低，最终使人产生疲惫无力的感觉，削弱了体力，局部过度紧张的肌肉可能会出现酸痛现象。

(2) 脑力疲劳

脑力疲劳主要是由于大脑持续了较长时间的与中枢神经活动有关的脑力劳动引起的，由于二氧化碳的堆积，刺激呼吸中枢，还可能会导致打哈欠，是一种不愿意再做任何活动的懒惰感觉，意味着身体急需休息。

(3) 心理疲劳

心理疲劳是一种常见的带有主观体验性质的长期性疲劳，又称心因性疲劳或精神性疲劳。与体力疲劳和脑力疲劳不同，心理疲劳大多是由情绪低落引起的，它不是发生在学习或工作进行之中，往往在刚刚开始甚至还没开始时就已出现。人在心理疲劳时对学习、工作的热情和兴趣会明显降低，甚至产生厌倦情绪。

(4) 混合性疲劳

混合性疲劳又称综合性疲劳，由以上两种或多种疲劳共同组成。最常见的为脑力疲劳与体力疲劳共存、脑力疲劳与心理疲劳共存。

驾驶疲劳主要是指驾驶员在驾驶车辆时，由于驾驶动作反复、枯燥、注意力高度集中、睡眠不足所引起的心理机能和生理机能下降的现象。驾驶疲劳一般为混合性疲劳，通常为体力疲劳和脑力疲劳共存。

2. 疲劳的危害

驾驶员疲劳会使视力下降，注意力不集中，驾驶动作迟缓，对驾驶员的知觉、感觉、判断产生影响，对速度、距离、环境等判断失误，有时甚至会产生幻觉或短暂的睡眠，这些都可能引发交通事故。

文献[22]从人体生理学的角度提出了工作——休息周期概念，把驾驶过程分为疲劳积累期和疲劳抑制期。理想情况下，一个完整的周期应该是驾驶工作达到疲劳稳定期，然后休息直至疲劳完全消除。而在实际情况中，驾驶员并不能将每个工作——休息周期全部进行到底，往往是还没有休息到疲劳完全消



除又开始了一个新的周期,这样疲劳会在体内不断堆积,人体抗疲劳能力不断下降,疲劳稳定期逐渐缩短,最终直接进入疲劳恢复期,恢复期很长,不仅危害人体健康而且影响行车安全。图 8-1 为长时间驾驶时的疲劳变化示意图。

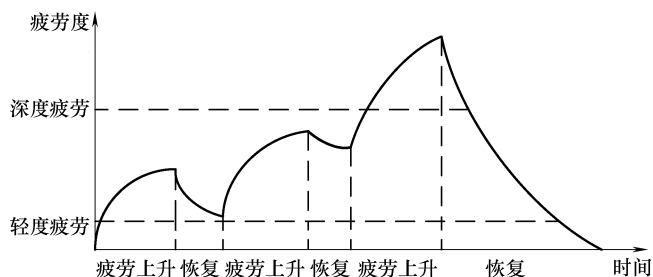


图 8-1 疲劳变化示意图

驾驶员产生疲劳时,除有循环机能的变化、呼吸机能的变化、血液的变化、神经机能的变化、尿的变化、体温的变化等以外,在心理状态方面也会引起多种变化,这些变化主要通过驾驶员的自觉症状和他觉症状表现出来。主要有以下几个方面:

1) 感觉器官功能和感知能力降低。主要表现为:不自觉的频频点头(打瞌睡),难以保持抬头姿势;视线模糊,眼睛发红、发干;打哈欠,肌肉放松,脸发木,面无表情,眼睑下垂,甚至闭眼;眨眼频率变大,眨眼持续时间增长。

2) 记忆、思维、判断能力下降。主要表现为:反应迟钝、反应时间变长,判断决策迟缓,经常漏掉重要的行车信息或忘记驾车技术规范和交通法规,走错行车路线。

3) 驾驶车辆动作不协调、不灵活。主要表现为:操作动作的连贯性降低,力量减弱,操作不到位;操作动作僵硬,节奏缓慢,失去方向感等。

4) 在车辆行驶状态上主要表现为:车速不稳定,速度忽快忽慢,行车方向控制不好,车辆左右摇摆呈“S”形路线前进等。

5) 主观上不愿意继续驾驶,信心、耐心和自我控制力减退,容易激动、急躁和开快车,无法按驾驶操作要求继续驾车。

驾驶员疲劳时较为明显的外在特征主要有以下几点:

- 1) 脸发呆,打哈欠。
- 2) 不自觉的频频点头(打瞌睡),难以保持抬头姿态。
- 3) 肌肉放松,面无表情,眼睑下垂,甚至闭眼。
- 4) 眨眼频率变大,眨眼持续时间变长。
- 5) 失去方向感,车辆偏出车道,或行驶轨迹呈“S”形曲线。





6) 随意变换车速, 行驶速度不定。

驾驶员由于在性格、年龄、文化、习惯、驾龄等存在个体差异, 在疲劳时对于某些特征的反应是不同的。因此, 通过某一症状判断驾驶员是否疲劳的可靠性不高, 而将多个症状进行分析判断, 综合考虑, 可以大大提高驾驶员疲劳检测的准确率。

8.3.2.2 PERCLOS 特征

Walt Wierwille^[23]认为 PERCLOS 是最具潜力的疲劳测定方法之一。美国国家公路交通安全局 (NHTSA)^[24]讨论了 PERCLOS 与其他眼睛活动测量方法的有效性对比, 研究结果显示, 该方法非常适合用于检测驾驶员的精神状态, 利用 PERCLOS 可以实现实时检测。美国联邦公路管理局 (FHWA)^[25]也推荐测量驾驶员的 PERCLOS 作为疲劳检测方法。通过大量实验发现 PERCLOS 中的 P80 参数与疲劳程度有较好的线性相关性^[26,27]。

P80: 认为眼睑遮住瞳孔面积的 80%, 就认为眼睛闭合, 以此计算在一段时间内眼睛闭合时所占的时间比例。

图 8-2 为眼睛睁开程度 (百分比) 与时间之间的关系, t_1 表示眼睛在完全睁开后开始闭合的时刻; t_2 表示眼睛闭合 80% 对应时刻; t_3 表示眼睛睁开 20% 对应时刻; t_4 表示眼睛完全睁开对应时刻。

根据定义, P80 判断标准的 PERCLOS 值为

$$P_{80} = \frac{t_3 - t_2}{t_4 - t_1} \times 100\% \quad (8-6)$$

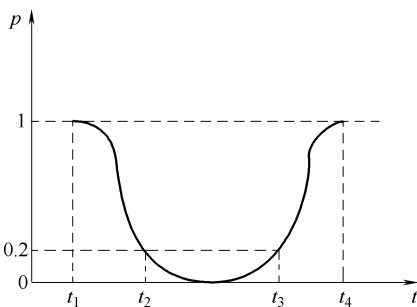


图 8-2 PERCLOS 示意图

PERCLOS 测量方法: 由于时间 t 与视频中图像序列是相互对应的, 所以可采用帧数计算 P_{80}

$$P_{80} = \frac{N_c}{N_T} \quad (8-7)$$

式中 N_c ——眨眼过程中眼睛闭合所对应的帧数;

N_T ——眨眼一次对应帧数。

研究表明^[28], 当 $P_{80} < 0.2$ 时, 驾驶员精神好, 处于清醒状态; 当 $0.2 \leq P_{80} < 0.4$ 时, 驾驶员为轻度疲劳状态; 当 $P_{80} \geq 0.4$ 时, 驾驶员为深度疲劳状态, 该标准可以作为量化驾驶员疲劳程度的参考值。

8.3.2.3 眨眼参数

眨眼是一种快速的睁闭眼运动, 常见的眨眼有两种: 一种是保护性的, 是



对外界刺激的快速反应，医学上称为“角膜反射”；第二种眨眼的起因至今没有一个合理的解释，是每个人平时常做的动作，并无外界刺激存在，由于是在不知不觉中完成的，所以叫做“不自主运动”。

眼睛眨眼时间 T_{Blink} 定义为眼睛从开始闭合到再次完全睁开持续的时间。如图 8-2 所示， $T_{Blink} = t_4 - t_1$ 。驾驶员在清醒和疲劳时眨眼时间的长短有明显不同，人在疲劳时，眼睑的眨动一般较频繁，眼睛闭合时间也较长。在驾驶过程中，驾驶员的眼睛闭合时间若超过一定时间，则表明他出现疲劳。若眼睛闭合时间达到 0.5s 表示人已经疲劳，此时驾驶很容易发生交通事故^[29]。

眨眼时间 T_{Blink} 测量方法

$$T_{Blink} = t_4 - t_1 = \frac{N_{TB}}{FPS} \quad (8-8)$$

式中 t_1, t_4 ——如图 8-2 所示；

N_{TB} ——眨眼一次对应帧数；

FPS ——图像的刷新频率（帧/s）。

眼睛眨眼频率 f_{Blink} 定义为在单位时间内出现眨眼的次数。通常情况下，人的眨眼频率平均为 15 ~ 20 次/min，每次持续 200 ~ 300ms，眨眼频率会受到疲劳程度和疾病的影响^[30]。人疲劳时会表现出眨眼速度变慢，频率变大的现象，因此，统计眨眼频率也可作为衡量疲劳程度的一项指标。

眨眼频率 f_{Blink} 计算方法

$$f_{Blink} = \frac{N_{Blink}}{\Delta T_{Blink}} \quad (8-9)$$

式中 N_{Blink} ——时间段 ΔT_{Blink} 内眨眼次数；

ΔT_{Blink} ——检测时间段。

8.3.2.4 点头参数

驾驶员疲劳时，会出现点头动作，不能保证正常的视线方向，同时为了清醒头脑，又将头抬起，驾驶员不断地重复点头——抬头动作，该动作的发生频率，如果超过一个预定的值，则可以认为已经发生了疲劳。同时，如果点头时间过长，则说明驾驶员已临近睡眠状态，因此该现象也可作为判断驾驶员是否疲劳的一个标准。Philip^[31] 于 1998 年设计了一个测量点头动作的仪器，研究表明，微小的点头动作和疲劳有很强的相关性。

根据驾驶员低头角度 β 判断是否点头， β 随时间变化如图 8-3 所示。当 $\beta > \beta_r$ 时，定义为一次有效的点头动作。通过试验得出，当 $\beta_r = \pi/18$ 时，驾驶员的点头会出现较大幅度动作，因此设阈值 $\beta_r = \pi/18$ 。图 8-3 中在 $t_2 t_3$ 时间段内为头部的点头姿态。

点头频率 f_{Nod} 计算方法



$$f_{Nod} = \frac{N_{Nod}}{\Delta T_{Nod}} \quad (8-10)$$

式中 N_{Nod} ——在 ΔT_{Nod} 时间段内点头的次数；

ΔT_{Nod} ——检测时间段。

点头时间 T_{Nod} 定义为：驾驶员发生点头动作到头部姿势正常时所用时间，其值为

$$T_{Nod} = t_3 - t_2 = \frac{N_{TN}}{FPS} \quad (8-11)$$

式中 N_{TN} ——发生一次点头动作所对应的图像帧数；

FPS ——图像的刷新频率（帧/s）。

8.3.2.5 哈欠特征

人疲劳时往往会打哈欠，以提醒人体大脑已疲劳，需要睡眠休息。一般认为打哈欠是脑缺氧的表现，通过打哈欠的深呼吸运动来增加血液中的氧含量，排出尽可能多的二氧化碳，从而使人精力更加充沛。人在一生之中都会打哈欠，完成一次打哈欠的时间大约需要 6s，在这期间人体全身的神经、肌肉得到完全放松，使人在生理和心理上得到较好的休息，对人体具有重要的生理保护作用。因此，可以通过检测驾驶员是否打哈欠来判断疲劳程度。基于嘴巴状态^[32,33]的判断方法是近几年发展较快的一种疲劳检测手段，驾驶员驾驶车辆时，嘴巴基本上处于闭合状态；而说话时，嘴巴以较快的速度进行张闭动作；当驾驶员打哈欠时，嘴巴张开到最大，并且持续较长一段时间，因此，嘴巴的张合状态也是判断驾驶员是否疲劳的一个依据。

通过大量试验发现，大多数人在闭嘴或者轻声交谈时，嘴巴宽高比 $K_M > 1.2$ ；当打哈欠时 $K_M < 0.7$ ，通过 K_M 值即可判断驾驶员是否在打哈欠。当 $K_M < 0.7$ 且持续时间大于 4s 时，就认为驾驶员在打哈欠。

打哈欠频率 f_{Yawn} 计算方法

$$f_{Yawn} = \frac{N_{Yawn}}{\Delta T_{Yawn}} \quad (8-12)$$

式中 N_{Yawn} ——在 ΔT_{Yawn} 时间段内打哈欠的次数；

ΔT_{Yawn} ——检测时间段。

8.3.2.6 车道偏离特征

疲劳驾驶时，由于驾驶员反应迟钝、注意力分散，车辆有可能偏离车道而不按正常轨迹行驶。本书主要检测两个技术指标：一是车道偏离率 ξ ；二是车道偏离量变化率 ζ 。根据试验得出车辆偏离距离 d 在 $d < 0.3$ 之内属于正常偏离范

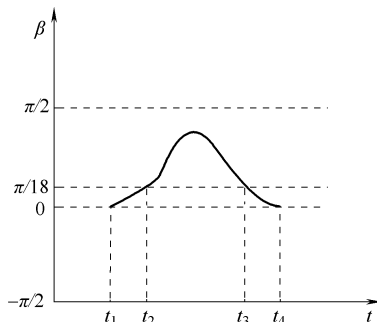


图 8-3 点头程度示意图



围；偏离距离 $0.3 \leq d < 0.7$ 时车辆出现了偏离； $d \geq 0.7$ 时车辆严重偏离车道线。偏离距离 d 与车道偏离率之间的关系如图 8-4 所示。对于车道偏离量变化率，当 $\zeta < 0.1$ 时车辆行驶轨迹属于正常范围； $0.1 \leq \zeta < 0.15$ 时车辆偏离量变化较大； $\zeta \geq 0.15$ 时车辆偏离量变化很大。

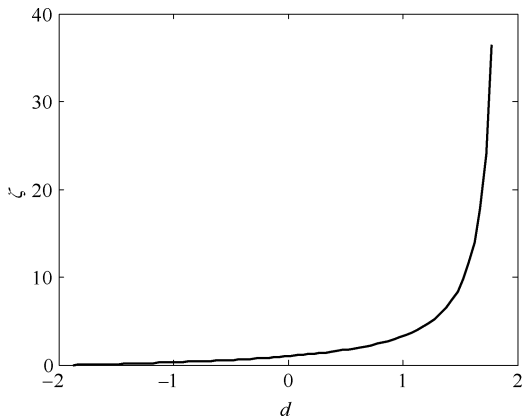


图 8-4 偏离率与偏离距离关系

8.3.3 疲劳程度判断

考虑到在驾驶员疲劳检测过程中，各检测量的数据类型不同，且数据量较大的特点，本书采用分布式信息融合结构，首先对各传感器的原始观测数据进行初步分析处理，做出本地判决结论；然后将结论信息传递到融合中心，在决策级按照粗糙集理论将多方面的数据做进一步的合成处理，得出最终判断。融合结构如图 8-5 所示。

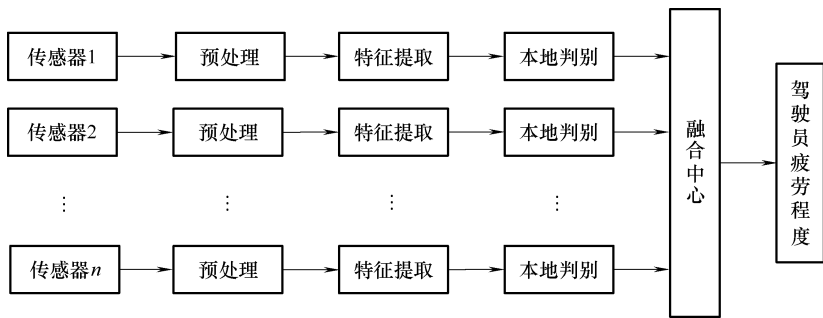


图 8-5 决策层融合结构

8.3.3.1 属性离散归一化

本系统中将下列参数作为粗糙集中的条件属性： $\{a = P80, b = \text{眨眼时间},$



c = 眨眼频率, d = 点头时间, e = 点头频率, f = 打哈欠频率, g = 车道偏离率, h = 车道偏离量变化率}。由于这些参数是连续的数值型数据, 因此需先将数据进行离散化处理, 否则粗糙集理论将得不到正确的结论。

本书根据驾驶疲劳相关条件属性的分析, 采用数据分割的方法进行离散。由于不同属性表示的意义不同, 因此数据的分割区域亦不相同。本书将属性离散为三个等级, 分别以 $\{0,1,2\}$ 表示, 具体如下:

- 0: 表示在正常范围内;
 - 1: 表示超出正常范围程度较小;
 - 2: 表示超出正常范围程度较大。
- 具体离散规则如表 8-1 所示。

表 8-1 属性离散化

条 件 属 性	0	1	2
P_{80}	$P_{80} < 20\%$	$20\% \leq P_{80} < 40\%$	$P_{80} \geq 40\%$
眨眼时间/s	$T_{Blink} < 0.3$	$0.3 \leq T_{Blink} < 0.5$	$T_{Blink} \geq 0.5$
眨眼频率	$f_{Blink} < 20$	$20 \leq f_{Blink} < 25$	$f_{Blink} \geq 25$
点头时间/s	$T_{Nod} < 0.3$	$0.3 \leq T_{Nod} < 0.5$	$T_{Nod} \geq 0.5$
点头频率	$f_{Nod} < 2$	$2 \leq f_{Nod} < 5$	$f_{Nod} \geq 5$
打哈欠频率	$f_{Yawn} < 2$	$2 \leq f_{Yawn} < 5$	$f_{Yawn} \geq 5$
车道偏离率	$0.7 < \xi < 1.4$	$0.4 < \xi \leq 0.7$ or $1.4 \leq \xi < 2.3$	$\xi \leq 0.4$ or $\xi \geq 2.3$
车道偏离量变化率	$\zeta < 0.1$	$0.1 \leq \zeta < 0.15$	$\zeta \geq 0.15$

决策属性 O 代表驾驶员疲劳程度, $O = \{0,1,2\}$, 表示驾驶员的三种疲劳状态:

- 0: 驾驶员处于清醒状态, 能够胜任驾驶任务;
- 1: 驾驶员处于轻度疲劳状态, 此时驾驶员虽感到疲劳但仍可驾驶车辆, 但是驾驶员需要休息, 如果继续驾驶车辆将达到深度疲劳状态;
- 2: 驾驶员处于深度疲劳状态, 此时继续驾驶将存在重大安全隐患, 易发生交通事故。

人体的疲劳感觉具有渐进性, 如图 8-1 所示。随着驾驶时间的延长, 疲劳度呈现缓慢增长。然而, 由于人体的抑制作用, 在一段时间内疲劳程度也会呈现缓慢的递减现象。无论哪种情况, 疲劳度都是缓慢变化的, 特别是疲劳度的增长, 因此, 为增加疲劳检测的准确性, 需取一段时间内的条件属性进行分析判断。本书取当前时刻前 60s 内的条件属性数据进行统计, 当前条件属性取值取决于统计中数量最多的属性值。然而, 对于车道偏离率和车道偏离量变化率这两个条件属性是个例外, 进行信息融合时, 属性值取当前较小时间段 (0.5s) 内的测量值, 这是由于人体疲劳具有渐进性, 然而车辆的行驶由于车速较快,



属性值存在突变，因此，不能在较长范围内进行统计处理。信息融合过程如图 8-6所示。

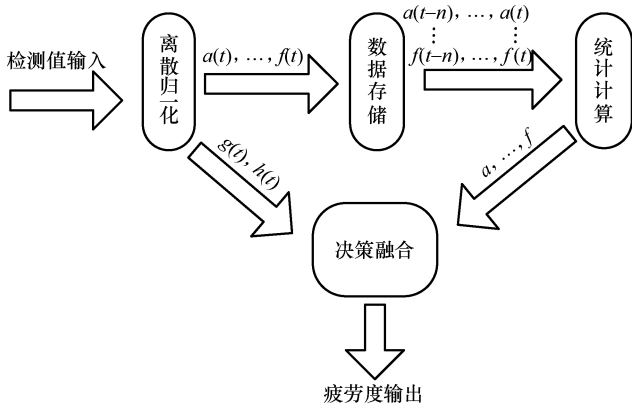


图 8-6 信息融合过程

部分离散化后的数据如表 8-2 所示。

表 8-2 决策表

<i>U</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>O</i>
1	2	1	1	2	1	0	1	2	2
2	1	0	0	0	0	0	0	0	0
3	1	2	2	1	1	1	0	1	2
4	2	2	2	0	0	0	1	0	2
5	1	0	0	2	2	1	1	1	2
6	1	1	2	2	2	0	1	1	2
7	1	1	1	1	2	2	2	0	2
8	1	0	0	1	1	1	2	1	2
9	1	1	1	1	1	1	1	1	2
10	1	1	1	1	1	1	0	2	2
11	0	0	0	1	1	0	0	1	1
12	0	0	1	1	0	0	0	1	1
13	1	0	1	0	0	0	0	1	1
14	0	0	1	1	0	0	1	1	1
15	0	0	0	1	1	0	1	1	1
16	1	0	0	1	0	0	0	0	1
17	0	1	1	0	0	1	1	0	1
18	1	0	1	0	0	0	0	1	1
19	0	0	1	1	0	0	1	0	1
20	0	0	0	1	0	1	0	1	1





8.3.3.2 数据约简

由于以上任何一个条件属性只要其值为 2, 就可判断驾驶员已经出现了深度疲劳状态, 因此, 任何一个条件属性都是必须的, 不能约简, 决策表是协调的, 其核值为 $\{a, b, c, d, e, f, g, h\}$, 共有 264 个最小决策算法。部分决策如下所示

$$\begin{aligned} & a_2 \rightarrow 2, b_2 \rightarrow 2, c_2 \rightarrow 2, d_2 \rightarrow 2, e_2 \rightarrow 2, f_2 \rightarrow 2, g_2 \rightarrow 2, h_2 \rightarrow 2, \\ & a_1 b_1 c_1 d_1 e_1 f_1 g_1 h_1 \rightarrow 2, a_1 b_1 c_1 d_1 e_1 f_1 g_1 h_0 \rightarrow 2, a_0 b_0 c_0 d_0 e_0 f_0 g_0 h_0 \rightarrow 0, \\ & a_1 b_0 c_0 d_0 e_0 f_1 g_1 h_0 \rightarrow 2, a_0 b_0 c_0 d_0 e_0 f_0 g_1 h_0 \rightarrow 2, a_0 b_0 c_0 d_1 e_1 f_0 g_0 h_1 \rightarrow 1, \\ & a_0 b_0 c_1 d_1 e_0 f_0 g_0 h_1 \rightarrow 1, a_1 b_0 c_1 d_0 e_0 f_0 g_0 h_1 \rightarrow 1, a_0 b_0 c_1 d_1 e_0 f_0 g_1 h_1 \rightarrow 1, \\ & a_0 b_0 c_0 d_1 e_1 f_0 g_1 h_1 \rightarrow 1, a_1 b_0 c_0 d_1 e_0 f_0 g_0 h_0 \rightarrow 1, a_0 b_1 c_1 d_0 e_0 f_1 g_1 h_0 \rightarrow 1, \\ & a_1 b_0 c_1 d_0 e_0 f_0 g_0 h_1 \rightarrow 1, a_0 b_0 c_1 d_1 e_0 f_0 g_1 h_0 \rightarrow 1, a_0 b_0 c_0 d_1 e_0 f_1 g_0 h_1 \rightarrow 1. \end{aligned}$$

根据最小决策算法即可判断驾驶员是否疲劳, 以及疲劳程度。

8.4 本章小结

本章主要研究了运用多信息融合技术进行驾驶员疲劳检测的问题。考虑到检测数据类型不同、数据量大的特点, 采用分布式融合结构, 在决策级用粗糙集理论进行融合判断。研究了粗糙集中条件属性的离散归一化问题, 将检测值根据评判标准归一化到疲劳程度上。研究了数据约简方法, 寻找出判断驾驶员是否疲劳的 264 个最小决策方法。

参考文献

- [1] Llinas J, Waltz E. Multisensor Data Fusion [M]. Artech House, Norwood, Massachusetts. 1990.
- [2] Hall D L. Mathematical Techniques in Multisensor Data Fusion [M]. Artech House, Boston, London. 1992.
- [3] 关键, 何友, 彭应宁. 多传感器分布式检测综述 [C]. 第七届全国雷达年会论文, 南京, 1999, 11: 355-359.
- [4] 杨露菁, 余华. 多源信息融合理论与应用 [M]. 北京: 北京邮电大学出版社, 2006.
- [5] 韩崇昭, 朱洪艳等. 多源信息融合 [M]. 北京: 清华大学出版社, 2005.
- [6] Nnez J, Otazu X. Multi-resolution-based Image Fusion with Additive Wavelet Decomposition [J]. IEEE Transactions on Geoscience and Remote Sensing. 1999, 37 (3): 1204-1211.
- [7] Petrovic V S, Xydeas C S. Gradient-based Multiresolution Image Fusion [J]. IEEE Transactions on Image Processing (SI057-7149), 2004, 13 (2): 228-237.
- [8] Alspach D L, Sorenson H W. Nonlinear Bayesian Estimation Using Gaussian Sum Approximations [J]. IEEE Transactions on Automatic Control. 1972, 17 (4): 439-448.



- [9] Rudolph van der Merwe, Arnaud Doucet, Nando de Freitas. The Unscented Particle Filter [R]. Technical Report: Cambridge University Engineering Department. 2000.
- [10] Carine Hue, Jean-pierre Le Cadre, et al. Sequential Monte Carlo Methods for Multiple Target Tracking and Data Fusion [J]. IEEE on signal Processing. 2002, 50 (2): 309-325.
- [11] Radford M. Neal. Probabilistic Inference Using Markov Chain Monte Carlo Methods. Technical Report CRG-TR-93-1 [R]. Department of computer science University of Toronto. 1993.
- [12] Logothetis A, Krishnamurthy V, et al. Expectation Maximization Algorithms for MAP Estimation of Jump Markov Linear Systems [J]. IEEE Transactions on Signal Processing. 1999, 47 (8): 2139-2156.
- [13] Richardson J M, Marsh K A. Fusion of Multi-sensor Data [J]. The International Journal of Robotics Research. 1988, 7 (6): 78-96.
- [14] James Joseph Clark. Data Fusion for Sensory Information Processing Systems [M]. Springer-Verlag New York, LLC. 2001.
- [15] Joshi R, Sanderson A C, et al. Minimal Representation Multisensor Fusion Using Differential Evolution [C]. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans. 1999, 29 (1): 63-76.
- [16] Nelson C I, Fitzgerald D S. Sensor Fusion for Intelligent Alarm Analysis [J]. IEEE Transactions on Aerospace and Electronic Systems. 1997, 12 (9): 18-24.
- [17] S. M. C. Peers. A Blackboard System Approach to Electromagnetic Sensor Data Interpretation [J]. Expert Systems. 1998, 15 (3): 197-215.
- [18] Bloch I. Information Combination Operators for Data Fusion: A Comparative Review with Classification [C]. IEEE Trans. Syst., Man, Cybern. A. 1996, 26 (1): 52-67.
- [19] Mongi A, Abidi, Rafael C, et al. Data Fusion in Robotics and Machine Intelligence [M]. Orlando, FL: Academic Press. 1992.
- [20] Pawlak Z. Rough Sets [J]. International Journal of Computer and Information Science. 1982, 11 (5): 341-356.
- [21] 同一鸣, 毛恩荣. 车辆人机工程学 [M]. 北京: 北京理工大学出版社, 1999.
- [22] 金键. 驾驶疲劳机理及馈选模式研究 [D]. 成都: 西南交通大学, 2002: 05, 29-30.
- [23] Wierwille, W W, Ellsworth, et al. Research on Vehicle Based Driver Status Performance Monitoring: Development, Validation, and Algorithms for Detection of Driver Drowsiness [R]. National Highway Traffic Safety Administration Final Report. 1994.
- [24] 夏芹, 义伟, 朱学峰. 基于 PERCLOS 的驾驶疲劳监控方法进展 [J]. 自动化技术与应用, 2008, 11: 81-83.
- [25] Bishop R. Survey of Intelligent Vehicle Applications Worldwide [C]. Proceedings of the IEEE intelligent Vehicles Symposium. 2000, 25-30.
- [26] 于兴玲, 王民, 张立材. 基于 PERCLOS 的驾驶员眼睛状态检测方法 [J]. 汽车电子, 2007, 23 (5): 251-253.
- [27] Dinges D F, Grace R. PERCLOS: a Valid Psychophysiological Measure of Alertness as As-





- essed by Psychomotor Vigilance [R]. Federal Highway Administration, Office of Motor Carriers. 1998; 26-29.
- [28] 郑培, 宋正河, 周一鸣. 基于 PERCLOS 的机动车驾驶员驾驶疲劳的识别算法 [J]. 中国农业大学学报, 2002, 7 (2): 104-109.
- [29] Barr L, Popkin S, Howarth H. An Evaluation of Emerging Driver Fatigue Detection Measures and Technologies [R]. Volpe Nat. Transp. Syst. Center, Cambridge, MA, Tech. Rep. 2009.
- [30] 淘海, 刘爽. 频繁眨眼不正常 [J]. 大众健康, 2007, 4: 19.
- [31] Kithil Philip W, Jones Roger D, John MacCuish. Development of Driver Alertness Detection System Using Overhead Capacitive Sensor Array [C]. SAE Technical Paper Series 982292. SAE International. 1998.
- [32] Liao W, Zhang W, Zhu Z, et al. Toward a Decision-theoretic Framework for Affect Recognition and User Assistance [J]. International Journal of Human Computer Studies. 2006, 64 (9): 847-873.
- [33] Ji Q, Zhu Z W, Lan P L. Real Time Non-intrusive Monitoring and Prediction of Driver Fatigue [J]. IEEE Transactions on Vehicular Technology. 2004, 53 (4): 1052-1068.

9.1 概述

前面章节分别研究了驾驶员面部姿势、视线方向、驾驶行为、车道线的识别及相关检测理论，本章根据视觉分散检测模型建立驾驶员视觉分散检测系统及驾驶疲劳检测系统，并进行一定的实验研究，以验证视觉检测系统的可行性及驾驶疲劳检测的可行性。

9.2 视觉分散检测实验

检测系统的硬件构成是由多方面的因素共同决定的，所要达到的性能指标是决定系统硬件配置的最关键因素，同时系统硬件的选择也是很广泛的。由于本书所需要实现的系统是一个实验性质的小型系统，对性能的要求不是很高，所以选择了一般的硬件配置。硬件系统主要完成驾驶员面部、手部视频采集，以及图像处理、模式识别等任务。因此，该系统的硬件组成主要包括两台摄像机、一台微型计算机，还包括部分板卡、数据线、电源线及逆变器等辅助器件。

其中主要硬件的型号：

摄像机 1 主要用来采集驾驶员的面部视频，因为需要较清晰的面部图像，所以采用成像质量较好的数字工业摄像机。型号为 AVT Guppy F-003C，逐行 SONY CCD，最大分辨率 $656 \times 494\text{dpi}$ ，最大帧率为 58 帧/s，镜头为 COMPUTAR 5mm 镜头。具有自动增益控制、自动白平衡、自动曝光控制等功能。采用 IEEE 1394a 接口方式与计算机进行数据传递。

摄像机 2 主要用来采集驾驶员的手部视频，不需要太高质量的视频图像，因此选择成本较低的极速辉煌 05 摄像机。其性能指标如下：301plh 芯片，CMOS 传感器，最大分辨率 $800 \times 600\text{dpi}$ ，最大帧率为 30 帧/s，具有完善的自动亮度、自动白平衡控制，可提供色彩饱和度、对比度等高级数码影像控制功能。采用 USB 2.0 接口方式与计算机进行数据传递。



计算机 计算机是执行处理算法的核心部分，本次使用的是普通 PC 机，基本配置为：Intel Pentium Dual E2180 2.00Ghz CPU，2G 内存。

检测系统软件在 Windows XP 操作系统下 Matlab2006a 的软件环境下实现。根据视觉分散检测模型，检测系统需要包括以下五个部分：

- 1) 采集驾驶员面部、手部彩色视频图像，并进行图像预处理；
- 2) 利用肤色特征，进行人脸区域定位，以估计驾驶员面部姿势；
- 3) 在人脸区域内进行眼睛定位，以识别驾驶员眼睛视线方向；
- 4) 定位并跟踪驾驶员双手位置，判断驾驶员是否在转向；
- 5) 判断驾驶员是否发生视觉分散，若是则给出警告。

系统软件框架如图 9-1 所示。

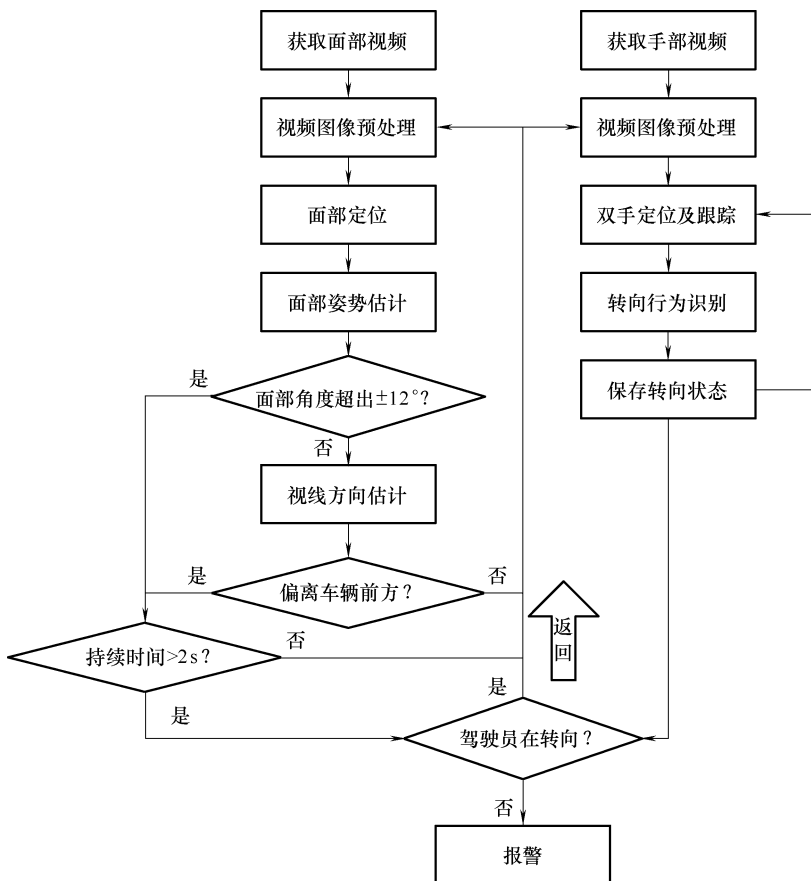


图 9-1 视觉分散检测系统软件框架

因为在行驶的车辆上进行视觉分散作业毕竟很危险，所以让驾驶员在车流量较少的道路上低速驾驶并完成多种作业：①直线行驶时观察仪表板；②直线



行驶时调节收音机；③完成十字路口处转向的驾驶行为。完成作业的目的是验证检测系检测视觉分散的效果，以及能否避免转向过程中的虚警现象。

选择车内空间较大的中兴 SUV 车辆，并在场地内安装摄像机，摄像机 1 正对驾驶员面部，摄像机 2 朝向转向盘中心，安装位置如图 9-2a 所示。在道路行驶中把摄像机的帧率均设定在 30 帧/s，获取的视频图像如图 9-2b 所示。因为使用的车辆不同，获取的驾驶员手部图像与前面章节不尽相同。但不影响检测驾驶员的转向行为，因为我们应用的参数是驾驶员手部位置标准差。在分析手部位置时，建立坐标系与图 6-1b 相同。



a) 摄像机在车辆内的安装位置



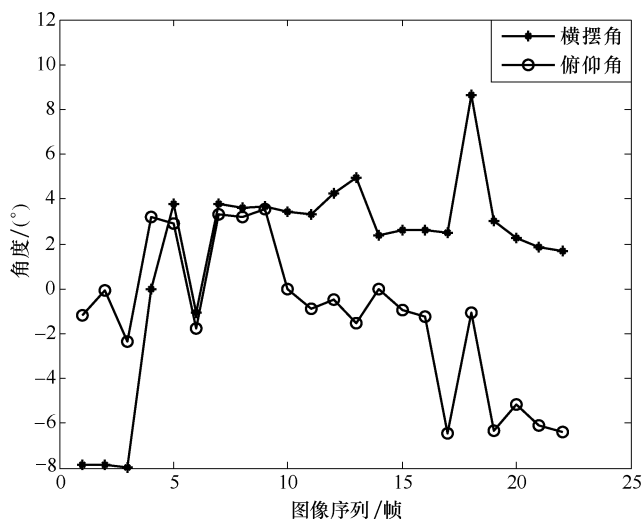
b) 摄取的手、面部图像

图 9-2 摄像机安装及获取图像

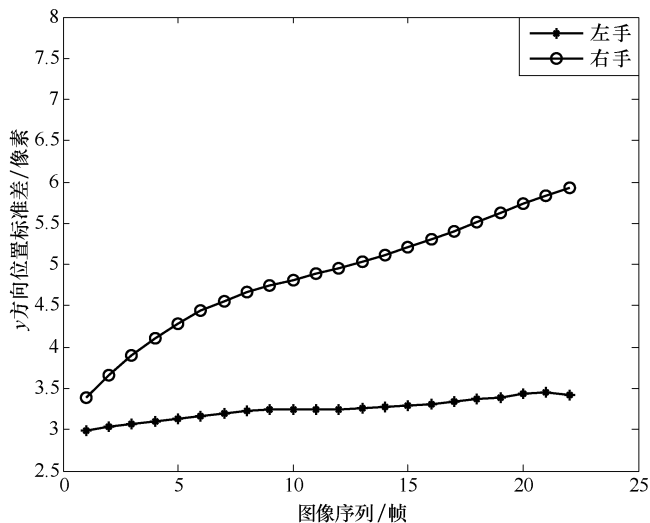
采用前述章节的相关理论对本节提出的三种作业进行识别与检测实验，得到的实验结果及分析如下。

1. 观察仪表盘

图 9-3 中可以看出驾驶员观察仪表盘时面部姿势变化较小，面部俯仰角与



a) 驾驶员面部姿势



b) 驾驶员手部位置标准差

图 9-3 驾驶员观察仪表盘

横摆角都在 $\pm 12^\circ$ 范围以内，因此需要判断驾驶员的眼睛视线方向。根据第 5 章的方法对眼睛视线方向进行识别，其中连续 20 帧（共 22 帧）识别为视线方向偏离，持续时间不到 1s 对驾驶能力的影响较小。

2. 调节收音机

图 9-4 中可以看出驾驶员调节收音机时面部姿势变化较大，面部横摆角偏

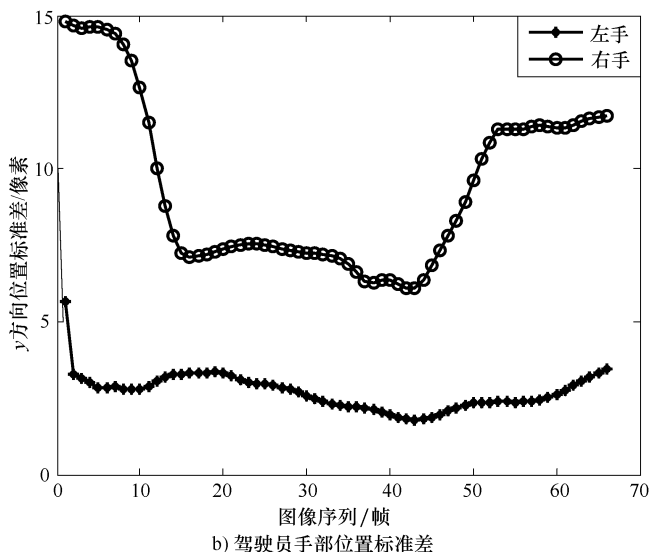
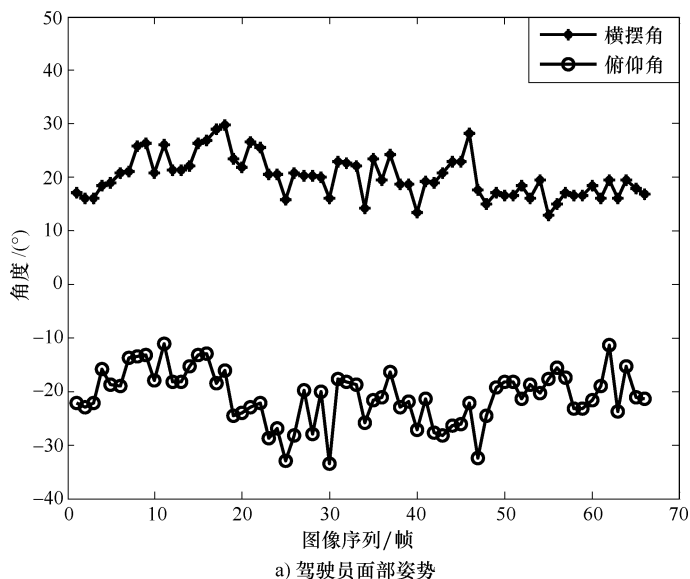


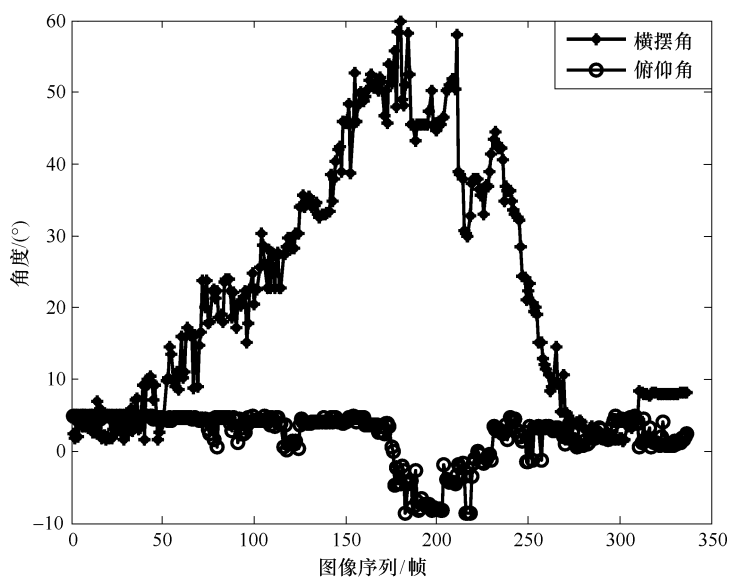
图9-4 驾驶员调节收音机

转达到 20° 左右,俯仰角在 -20° 左右变化,并且持续时间超过2s,需要判断驾驶员是否在转向。因为只有驾驶员右手的位置标准差变化较大,所以可以判定驾驶员并没有进行转向行驶,这时需要对驾驶员进行警告。

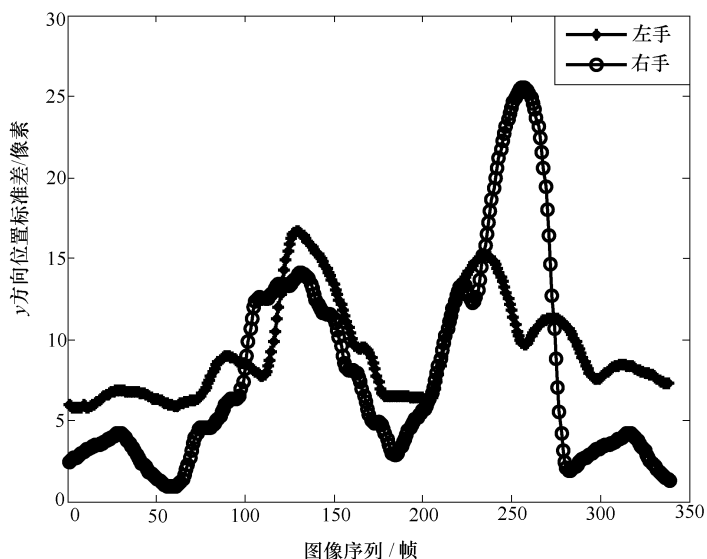
3. 十字路口转向

从图9-5a中我们可以看出,在十字路口的转向过程中驾驶员面部姿势变化比较明显,特别是横摆角明显超出 $\pm 12^{\circ}$ 的正常视线范围,并且持续时间超过





a) 驾驶员面部姿势



b) 驾驶员手部位置标准差

图 9-5 十字路口转向

2s。根据转向行为识别算法，可以识别出驾驶员是在进行转向，因此不需要警告驾驶员，避免虚警的发生。



9.3 驾驶疲劳检测实验

为研究信息融合方法对驾驶员疲劳检测的有效性,本书作者设计了一个检测性实验,考虑到驾驶安全因素,本书作者采用了模拟各种状态的实验方法。图9-6、图9-7、图9-8、图9-9为四段实际检测数据。其中,图9-6为驾驶员清醒时采集的数据;图9-7为驾驶员轻度疲劳时采集的数据;图9-8、图9-9为驾驶员深度疲劳时采集的数据。

从图9-6a~图9-6h中可以看出,离散后的8个条件属性值在0与1之间波动,其中,P80、眨眼时间周期、眨眼频率、点头时间分别对应的值 a 、 b 、 c 、 d 大部分时间在正常范围内,偶尔会稍微超出正常范围,单靠这些指标中任何一个都不能判断驾驶员的疲劳程度;点头频率 e 在大部分时间内值都为1,也不能

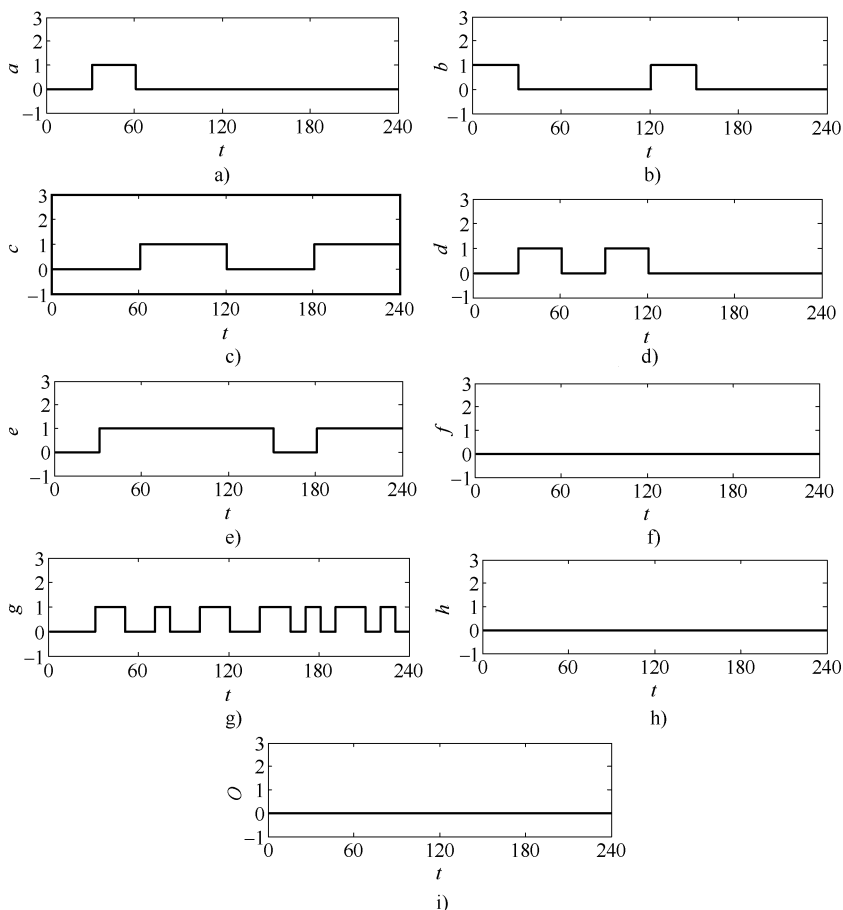


图9-6 数据融合曲线 (一)



说明驾驶员的疲劳程度； f 值在此段时间内都为 0，说明驾驶员没有打哈欠； g 值在 0、1 之间变化较为频繁，同时 h 值一直为 0，说明此段时间驾驶员的车辆在车道中心边缘行驶，并未出现失去控制的车辆偏离。综合这些指标，根据最小决策算法判断驾驶员处在清醒状态。综合判断疲劳程度输出如图 9-6i 所示。

图 9-7a ~ 图 9-7h 为驾驶员轻度疲劳时的数据，曲线在开始阶段， a 、 b 值为 1 偏大，其他值为 0，都在正常范围内；然而 c 、 d 、 e 、 g 值为 1，根据融合规则，判断驾驶员处于轻度疲劳阶段；随着各属性值恢复正常，判断驾驶员恢复到清醒状态。综合判断疲劳程度输出如图 9-7i 所示。

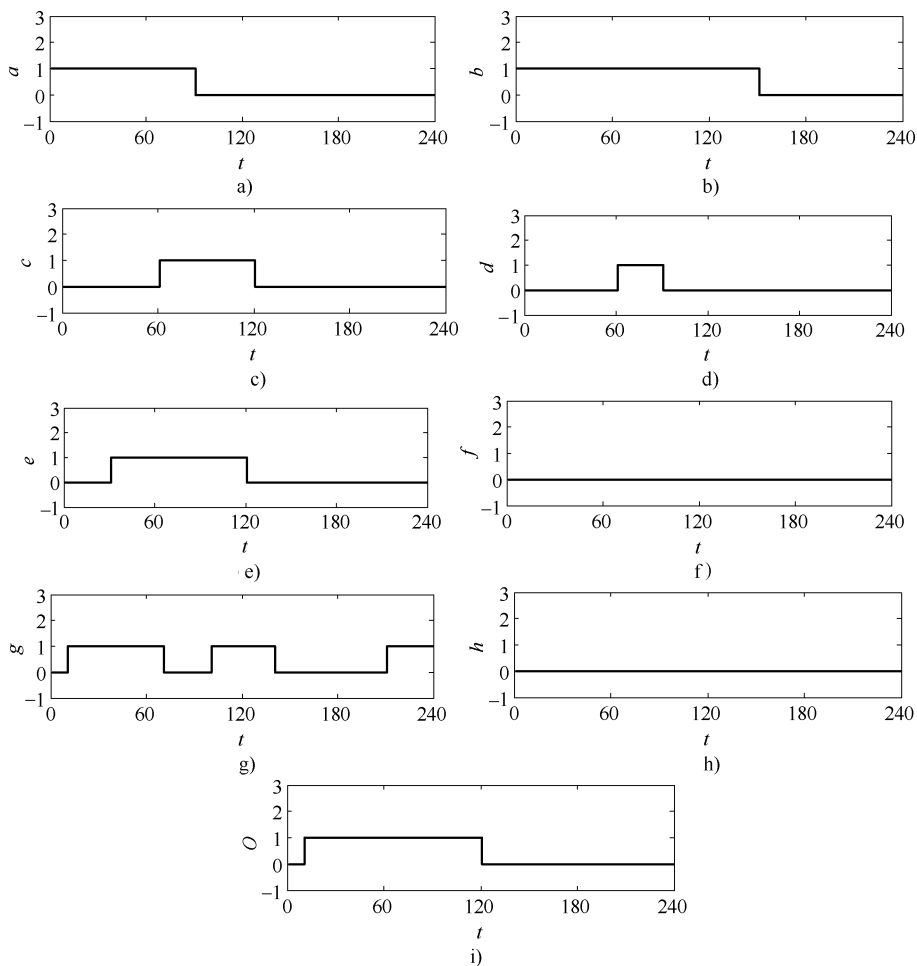


图 9-7 数据融合曲线 (二)



图9-8a~图9-8h为驾驶员驾驶过程中出现深度疲劳的曲线,可以看出 a 、 b 、 c 、 e 值都能达到最大值, h 经常取1,说明车辆在车道中变向行驶,并未沿车道线方向行进。此时,车辆处于高度危险状态,必须马上发出警报;后半阶段随着驾驶员逐渐清醒,各检测值恢复到正常水平。综合判断疲劳程度输出如图9-8i所示。

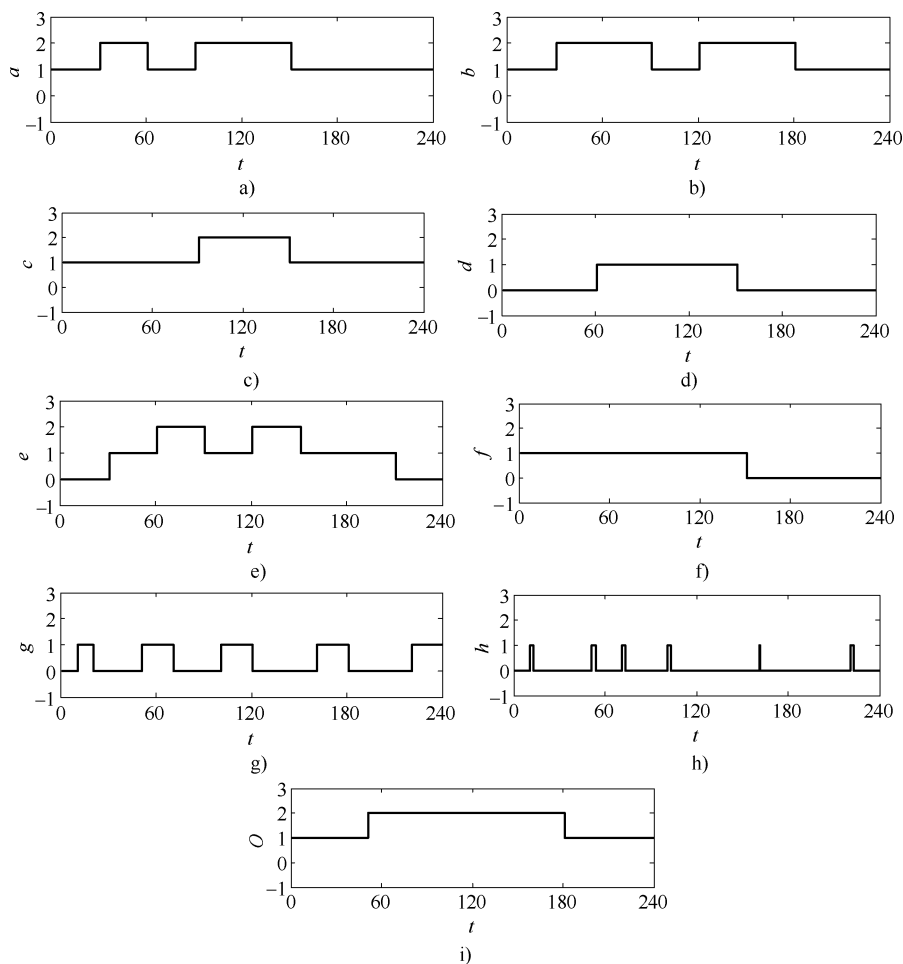


图9-8 数据融合曲线 (三)

图9-9a~图9-9h同样是驾驶员处在疲劳时的检测曲线,根据融合规则,检测疲劳程度输出如图9-9i所示。

可以看出,采用信息融合规则后,通过检测多个特征量,综合考虑各特征状态,根据融合决策判断驾驶员疲劳程度,比单一检测方法更准确、更全面。

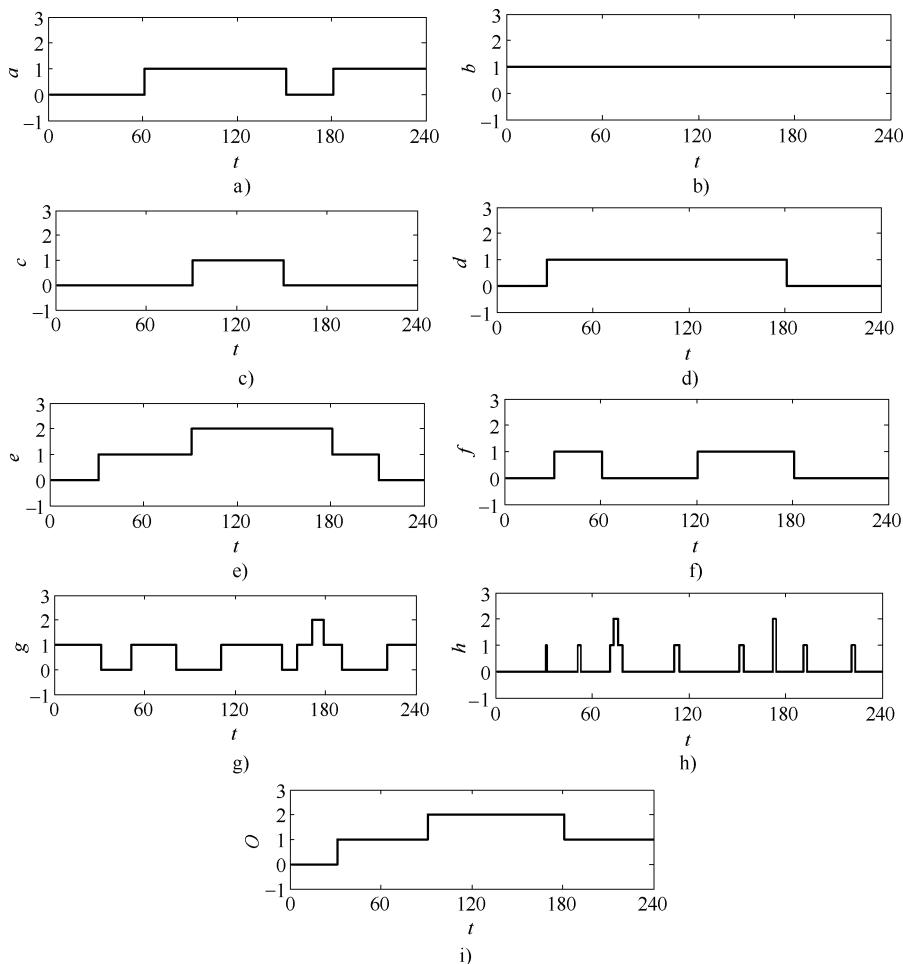


图 9-9 数据融合曲线（四）

9.4 本章小结

本章中作者根据视觉分散检测模型，建立驾驶员视觉分散检测系统的软硬件系统，并进行实验研究。摄取驾驶员在驾驶过程中进行各种视觉分散作业，以及转向时的手、面部视频图像，利用前面章节的研究成果进行检测实验，结果表明检测系统可以检测出驾驶员的各种视觉分散行为，并能有效防止转向时的虚警现象。在模拟状态下获取驾驶员疲劳状态各种视频进行疲劳检测实验，结果表明采用信息融合算法能够减少误判，检测结果更具可靠性。

本书研究了基于视频图像分析的驾驶员视觉分散检测，并重点研究了驾驶员面部姿势、眼睛视线方向特征的提取方法，以及驾驶员转向行为的识别方法，还对驾驶员被动视觉分散（驾驶疲劳）检测方法进行了初步研究。本书作者建立了驾驶员视觉分散检测模型、面部图像定位、面部姿势提取、眼睛视线方向识别、转向行为识别、车道线识别、驾驶疲劳检测、设计驾驶员视觉分散检测系统等模型和系统，并取得如下阶段性研究成果。

1. 建立了驾驶员视觉分散检测模型

作者研究了视觉分散影响驾驶员的机理，驾驶员发生视觉分散时不能及时获得控制车辆所需的信号，只能保持转向盘以固定的转角，导致行驶路线的偏离。通过分析驾驶员观察四处文本信息时车辆的 SDLP（偏离道路中心距离的标准差），发现 SDLP 随着视线偏离车辆前方角度的增大而增大。作者研究了视觉分散的特点，认为通过视频分析检测行为的方法适合于检测驾驶员视觉分散。本书讨论了视觉分散检测的阈值，根据驾驶过程中驾驶员视线变化的特点，建立了基于驾驶员面部姿势与视线方向识别，并包含转向行为识别的视觉分散检测模型。

2. 研究了适合面部姿势提取的驾驶员面部精确定位方法

针对驾驶员面部姿势多变的特点，作者提出先采用肤色混合高斯模型预定位人脸区域，再根据眉毛、嘴唇位置精确定位驾驶员面部的的方法。针对眉毛区域灰度值低、变化剧烈的特点，采用联合投影函数定位眉毛区域上边缘。针对驾驶员面部姿势偏转较大时，少量背景图像会影响眉毛定位的不足，作者研究了一种在头发检测基础上，滤除背景图像的方法，实验证明效果良好；研究了利用唇色多项式模型及唇色比人脸肤色更红的特点，定位驾驶员嘴唇区域下边缘的方法；研究了驾驶员面部图像归一化的方法。实验证明，本书提出的由粗到精的定位方法可以精确定位多种姿势下的驾驶员面部，为面部姿势估计奠定基础。

3. 提出了利用核主元分析估计驾驶员面部姿势的方法

研究核主元分析实现原理，作者研究了利用核主元分析估计面部姿势步骤；



研究了获取标准样本图像的方法,设计了样本图像采集系统。利用核主元分析把高维面部图像存在的流形结构嵌入到二维空间,建立提取面部姿势的标准曲线,发现可以在同一空间下嵌入多人的姿势曲线,而不至于使姿势曲线形状发生较大改变。利用训练样本建立姿势曲线并拟合圆,然后把测试样本图像投影到相同空间,根据拟合圆心及姿势曲线上与新投影点最近的2个点,估计测试样本图像对应的角度,估计精度可满足一定实际需求。作者研究了不同核函数、核函数参数对估计面部姿势的影响,发现核函数参数的变化对提取精度的影响较大,在核函数参数选择合理的情况下,不同核函数对估计精度影响不大。

4. 提出了应用 Multi-PCA 识别驾驶员眼睛视线方向的方法

本书研究了 PCA 实现原理,分析了常用 PCA 应用于识别时存在的问题。作者针对驾驶环境中精确提取眼睛视线方向的困难,提出了应用 Multi-PCA 来识别眼睛视线方向的方法。研究了应用联合投影函数与眼角定位相结合定位眼睛区域的方法;研究不同视线特征空间的相似度,发现只有同类空间具有较高的相似度;研究不同眼睛视线方向图像在不同特征空间的重构误差,发现只有在样本图像与特征空间的类别相同时才能获得最小重构误差,从而也证明了方法的可行性。实验证明 Multi-PCA 方法的识别率高于 PCA 方法的识别率。

5. 提出了驾驶员转向行为的识别方法

驾驶员在十字路口处的转向过程中,视线方向偏离车辆前方的持续时间会超过 2s。为防止检测系统将此误识为驾驶员视觉分散,作者提出对驾驶员转向行为进行识别研究。作者研究了基于视频分析的驾驶员双手定位方法,为提高双手定位的实时性,研究了基于粒子滤波算法的驾驶员双手跟踪方法;研究驾驶过程中驾驶员手部位置的特点,发现转向过程中驾驶员手部位置变化相对剧烈,因此提出利用双手位置标准差识别驾驶员转向行为的方法。作者提出设置标志位方法,解决转向过程中部分的标准差较小而不利于转向行为识别的缺点,实验证明所提算法可以识别驾驶员的大部分转向行为。

6. 研究了道路检测跟踪方法

针对结构化道路特点,本书作者提出了道路图像分区垂直积分投影的车道线检测方法。为提高检测率,首先对道路图像进行预处理,提出了基于邻域主成分分析法的图像平滑方法。为增强图像对比度,作者提出基于 Sin 函数的图像拉伸曲线。作者研究了分区图像中车道标志线投影关系问题;针对投影曲线平滑后曲线轮廓易改变的缺点,提出了新平滑方法,该方法能在保持原始投影曲线形状基本不变的前提下进行平滑。为准确检测出车道线边缘点,作者提出了车道线边缘点调整方法;为得到车道线参数,采用最小二乘法对车道线进行线性拟合。作者研究了车道标志线的跟踪方法,基于视频图像的连续性,确定车道线感兴趣区域,在该区域中采用方向可调滤波器对车道线进行精确定位跟踪,



研究了跟踪过程中方向角的选取问题。实验结果表明,该检测方法能够准确检测出道路中的车道线。

对于非结构化道路检测,本书作者假设道路图像与周围环境颜色之间具有一定的差异性,提出了基于 FCM 分割的方法检测道路区域,并对 FCM 中的初始化问题进行了研究。针对非结构化道路的特点,作者提出了道路图像重采样及大尺度平滑的思想,并分析了重采样对检测结果的影响;讨论了直方图在图像区域划分中的应用问题;分析了道路区域图像块的特性,并提出了判断标准。实验结果显示,采用大尺度平滑算法能够提高检测率,图像重采样可以降低计算时间,FCM 方法能够较好的检测出道路区域。对于车辆行驶轨迹的检测,本书选用两个参数进行衡量,采用车道偏离率衡量车辆当前位置,采用车道偏离量变化率判断车辆行驶方向。书中分析了车道偏离率取值范围与车辆位置之间的量化关系,给出了车道偏离量变化率的计算方法。

7. 研究了多信息融合技术在疲劳检测中的应用

根据检测信息特点,本书采用了分布式信息融合结构进行疲劳程度判断。作者研究了粗糙集理论在决策级信息融合中的应用,根据粗糙集理论要求,分析了单个检测特征量与驾驶员疲劳程度之间的量化关系,并将条件属性进行离散归一化。对于表征疲劳程度的决策属性问题,作者将驾驶疲劳程度分为清醒、轻度疲劳、深度疲劳三种状态;研究了决策表的约简问题,根据实际情况,建立判断驾驶疲劳程度的 264 个最小决策方法。实验结果证明,采用信息融合技术可以更加全面地检测驾驶疲劳程度。

8. 建立了视觉分散检测系统的软硬件框架,并对驾驶员视觉分散进行实验研究

根据视觉分散检测模型,本书建立了驾驶员视觉分散检测系统的软、硬件框架,并进行实验研究;摄取驾驶员在驾驶过程中进行各种视觉分散作业以及转向时的手、面部视频图像,利用已有研究成果进行检测实验,结果表明检测系统可以检测出驾驶员的各种视觉分散行为,并能有效防止转向时的虚警现象。

本书的主要创新点如下:

1) 提出利用核主元分析估计驾驶员面部姿势的方法。设计精确角度下面部图像采集系统以获得标准样本。利用核主元分析,把高维面部图像存在的流形结构嵌入到二维空间,建立估计面部姿势的标准曲线,并拟合圆。将测试样本投影到相同空间,根据拟合圆心及姿势曲线上与新投影点最近的 2 个点,估计测试样本图像对应的角度。实验表明,该方法可以用同一个人的姿势曲线估计不同人的面部姿势,克服传统模式分类方法需要为不同人建立不同姿势曲线的缺点。

2) 提出基于 Multi-PCA 的眼睛视线方向识别方法。针对驾驶环境中精确提





取人眼视线方向的困难,本书提出采用 Multi-PCA 识别驾驶员 5 类眼睛视线方向的方法。首先,为每类人眼视线方向分别建立特征空间。然后,以测试样本在每类特征空间下的重构误差来实现分类。实验证明,该方法充分应用了 PCA 的最佳逼近性和每类视线方向的独有特征,获得了比常用 PCA 方法更高的识别率。

3) 提出根据手部位置标准差来识别驾驶员转向行为的方法。针对检测系统会把驾驶员转向行为误识为视觉分散的缺点,提出对驾驶员转向行为进行识别。根据转向时驾驶员双手位置变化剧烈的特点,提出利用双手位置标准差来识别驾驶员转向行为的方法。

4) 提出了基于图像分区垂直积分投影的车道线检测方法。首先,对道路图像进行预处理,提出了基于邻域主成分分析法的图像平滑方法。为增强图像对比度,作者提出了基于 Sin 函数的图像拉伸曲线。根据分区图像中车道标志线投影关系寻找车道线边缘点。针对投影曲线平滑后曲线轮廓易改变的缺点,提出了新平滑方法,该方法能在保持原始投影曲线形状基本不变的前提下进行平滑。

尽管对基于视频图像分析的驾驶员视觉分散检测的研究取得了一些成果,但由于时间与条件的限制,作者认为还需要在以下几个方面作进一步的研究和完善:

1) 文中主要在均匀光照的情况研究了驾驶员视觉分散特征的提取方法,为增强检测系统的鲁棒性,还需要研究光照变化对特征提取的影响。如偏光、阴影情况时的特征提取。

2) 本书重点对视觉分散的检测算法进行研究,为在实际中应用,需要进一步研究检测系统的实时性。如研究驾驶员面部、眼睛区域跟踪,考虑采用基于 DSP 等芯片的检测系统。

3) 本书主要应用面部姿势作为驾驶员视觉分散的判断依据,还需要进一步研究眼睛视线方向与面部姿势间的关系,并提高眼睛视线方向的提取精度,如采用高分辨摄像机以获取高质量图像,再开发驾驶环境中可行的提取方法。

4) 由于人员个体差异及驾驶习惯不同,导致人体疲劳时各检测指标可能不同,需进一步完善检测指标与驾驶员疲劳程度之间的量化关系,以提高疲劳检测系统的准确率。

后 记

自 2005 年读博士以来，我就从事驾驶员注意分散检测研究的相关工作。10 年来对驾驶员注意分散进行了大量研究，也积累了很多成果。本书内容主要来自于我的博士论文与朱淑亮博士论文的部分章节，以及两位博士的最新研究内容。

在过去的日子里，从学习科研到日常生活，得到了众多老师、同学、朋友和家人的关怀与帮助，这是最为值得珍惜的财富，我将永远铭记于心。在此，我要向所有帮助、关心过我的人致以诚挚的谢意！

感谢齐鲁工业大学的陈兆强老师及原山东轻工业学院的高立营、陈向东、王涛、张青等同学，他们为实验样本的采集付出了艰辛劳动。

衷心感谢我的家人和亲友，特别是我的爱人李春玲女士，感谢他们多年来在生活和精神上给予我的无私关爱和支持，对他们的感激之情无法用语言表达，我只有努力工作，用更大的成就来回报他们的殷切期望，献上我最真心的感谢！

本著作是国家自然科学基金（项目编号：51375275）、山东省高等学校科技计划（编号：J14LB01）项目的阶段性成果，并得到齐鲁工业大学特色名校建设项目的支持，在此一并表示感谢。著作由齐鲁工业大学的路玉峰博士与烟台大学的朱淑亮博士合作完成。

路玉峰

2015 年 2 月 3 日于泉城

内容提要

本书从视觉分散对驾驶员驾驶性能的影响研究入手,开展驾驶员视觉分散检测技术研究,并重点研究基于视频图像分析的驾驶员视觉分散特征的提取方法。主要研究内容:

(1) 研究视觉分散对驾驶性能的影响。(2) 研究多姿势下驾驶员面部、面部特征点精确定位的方法。(3) 研究驾驶员面部姿势的提取方法,提出利用核主元分析估计驾驶员面部姿势的方法。(4) 研究驾驶员视线方向提取方法,提出基于Multi-PCA(多主元分析)的视线方向估计方法。(5) 研究驾驶员转向行为识别,提出根据手部位置标准差来识别驾驶员转向行为的方法。(6) 根据视觉分散检测模型,建立驾驶员视觉分散检测系统的软、硬件框架,并进行视觉分散检测的实验研究。

地址:北京市百万庄大街22号
邮政编码:100037

电话服务

服务咨询热线:010-88361066

读者购书热线:010-68326294

010-88379203

网络服务

机工官网:www.cmpbook.com

机工微博:weibo.com/cmp1952

金书网:www.golden-book.com

教育服务网:www.cmpedu.com

封面无防伪标均为盗版



机械工业出版社微信公众号



扫一扫,
更多精彩汽车
生活图书任你选

上架指导 交通运输 / 汽车安全

ISBN 978-7-111-50151-0



9 787111 501510 >

ISBN 978-7-111-50151-0

策划编辑◎杜凡如 连景岩

编辑邮箱◎605285955@qq.com

定价:39.80元