

密级: \_\_\_\_\_



**中国科学院大学**  
University of Chinese Academy of Sciences

## 硕士学位论文

基于激光和视觉信息融合的行人跟踪研究

作者姓名 \_\_\_\_\_ 高山 \_\_\_\_\_

指导教师: \_\_\_\_\_ 焦建彬 教授 \_\_\_\_\_

学位类别: \_\_\_\_\_ 工程硕士 \_\_\_\_\_

学科专业: \_\_\_\_\_ 物流工程 \_\_\_\_\_

研究所: \_\_\_\_\_ 中国科学院大学工程管理与信息技术学院 \_\_\_\_\_

二零一三年 四月



**Pedestrian tracking based on laser and image data fusion**

**By**

**Shan Gao**

**A Paper Submitted to**

**University of Chinese Academy of Sciences**

**In partial fulfillment of the requirement**

**For the degree of**

**Master of Logistics Engineering**

**College of Engineering & Information Technology**

**April, 2013**



**中国科学院大学直属院系**  
**研究生学位论文原创性声明**

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

**中国科学院大学直属院系**  
**学位论文授权使用声明**

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密的学位论文在解密后适用本声明。

作者签名：

日 期：

导师签名：

日 期：



## 摘 要

在物流系统中，车辆起着举足轻重的作用。车辆辅助驾驶系统致力于通过检测跟踪车辆周围的行人目标，做出有效预警，为驾驶员和行车安全提供有效保障。本文以辅助安全驾驶原型系统为平台，基于激光传感器和视觉传感器的融合信息方法进行行人跟踪，用于安全辅助驾驶，在物流工程领域中有着重要的意义与应用价值。

本文提出一种基于激光深度信息和图像信息融合的实时行人跟踪方法，通过激光点滤波，聚类以及激光点模式分析完成对行人的初步定位，然后经过激光测距仪和摄像机之间的数据配准，将行人的潜在位置投影到图像层；进而利用快速行人检测技术完成行人位置确定并进行特征提取，最终在连续多帧之间进行数据关联，达到对行人实时检测跟踪的目标。实验结果及对比分析验证了本文所提算法的快速性和有效性。本文主要研究成果：

1、搭建了安全辅助驾驶原型系统平台。本文利用多传感器和移动机器人平台架设了原型平台，并在真实交通场景中进行了数据采集。激光测距仪和 CCD 摄像头固定安装后，本文对两种传感器进行了统一的坐标参数配准，并进行了实验验证。该系统对大雾，雨雪，光照差等环境同样适用；

2、利用激光分类模式和视觉方法进行行人检测。本文利用新的激光分类模式，在激光坐标系下对目标进行滤波聚类，模式分析，得到候选行人区；经过投影转化到图像层，利用 HOG+SVM 方法对候选区域进行小尺度变化的确认性检测识别，从而实现对行人跟踪的初始化；

3、提出基于信息融合的多目标跟踪模型。本文根据信息融合中数据关联基本理论，充分融合了激光的深度信息和图像信息，推导出多行人跟踪模型，并通过深度坐标的动态参数对多行人进行权值的调整，进行模型的在线更新，从而使得求解过程与客观视觉观测原理一致，合理解决遮挡条件下多目标跟踪丢失，错误标号等问题。并利用改进二部图匹配原理提出了模型的求解方法。

**关键词：行人跟踪，信息融合，深度信息，数据关联，遮挡模型**

## Abstract

- **Shan Gao** (Logistics Engineering)

Directed by: **Jianbin Jiao** (Professor)

The safe transport of the vehicle plays important role in the logistics. One of the most important branches in the intelligent transport system (ITS) technology is advanced driving assistant system (ADAS), which aims at detecting and tracking the obstacles around vehicles and providing warning information for drivers. Based on information fusion from multi-sensor, we can fully take advantage of high-quality information. Therefore, this subject of this thesis has important theoretical significance and wide practical value in the field of Logistics Engineering.

This thesis proposes a robust pedestrian tracking framework based on laser and image data fusion. The multi-information should run along: filtering, coarse clustering, projection, detecting and association to finish the multi-pedestrian real-time tracking job. The main content can be listed as following:

1. Setting up an ADAS prototype platform. Fixed with the laser rangefinder and CCD camera, the platform could finish the data collection in real transportation scenario, which can also be applied in smog, rain, and bad illumination condition.

2. Detecting pedestrian through information fusion. A new laser information clustering method has been adopted, after that the segments are projected to the image layer to form candidate regions, then the valid vision-detection following by.

3. Proposing a new association-based multi-pedestrian tracking model. Based on the depth and vision information, this thesis formulates an online learning tracking model and giving explicit occlusion reasoning to solve the complex occlusion problem.

**KEY WORDS: Pedestrian Tracking, Information Fusion, Laser Scanner, Data Association, Occlusion Model**

## 目 录

摘 要 .....	I
目 录 .....	III
图目录 .....	V
表目录 .....	VI
<b>第一章 绪论</b> .....	1
<b>1.1 课题背景和研究意义</b> .....	1
<b>1.2 国内外研究现状</b> .....	2
<b>1.2.1 车辆辅助安全驾驶系统</b> .....	2
<b>1.2.2 信息融合的基本理论</b> .....	4
<b>1.2.3 多目标跟踪</b> .....	6
<b>1.3 基于信息融合的目标跟踪的关键问题</b> .....	7
<b>1.4 本文的研究内容</b> .....	8
<b>1.5 本文的组织结构</b> .....	9
<b>第二章 车辆辅助安全驾驶原型平台及多传感器配准</b> .....	11
<b>2.1 传感器对比</b> .....	11
<b>2.2 实验平台搭建</b> .....	12
<b>2.2.1 激光测距仪 LMS291 介绍</b> .....	13
<b>2.2.2 单目摄像头介绍</b> .....	15
<b>2.3 基于激光测距仪与单目摄像头的配准</b> .....	16
<b>2.4 本章小结</b> .....	24
<b>第三章 基于激光信息的行人快速检测研究</b> .....	25
<b>3.1 激光数据预处理</b> .....	25
<b>3.2 基于激光信息的粗分类</b> .....	27
<b>3.2.1 激光信息</b> .....	27
<b>3.2.2 宽高比特征</b> .....	28
<b>3.2.3 候选区域投影</b> .....	29
<b>3.3 基于视觉信息的行人目标的精检测</b> .....	30
<b>3.4 本章小结</b> .....	32
<b>第四章 基于激光和视觉信息融合的多目标鲁棒跟踪</b> .....	35
<b>4.1 基于数据关联的多目标跟踪介绍</b> .....	35
<b>4.2 基于激光与视觉信息融合的多目标跟踪</b> .....	37
<b>4.2.1 跟踪模型的构建</b> .....	37

4.2.2 跟踪模型的求解 .....	39
4.2.3 跟踪模型的在线更新 .....	41
4.2.4 跟踪模型在遮挡问题中的求解 .....	43
4.3 多目标跟踪模型的实验验证 .....	44
4.3.1 实验数据集介绍 .....	44
4.3.2 实验的评测标准 .....	45
4.3.3 实验结果及分析 .....	46
4.4 本章小结 .....	48
总结与展望 .....	51
参考文献 .....	53
个人简介及发表文章目录 .....	58
致 谢 .....	59

## 图目录

图 1-1 像素级融合 .....	5
图 1-2 特征级融合 .....	5
图 1-3 决策级融合 .....	5
图 2-1 系统平台搭建 .....	12
图 2-2 激光测距仪和扫描示意图 .....	13
图 2-3 激光测距原理图 .....	14
图 2-4 传感器相对位置和摄像头 .....	15
图 2-5 图像像素坐标系和物理坐标系 .....	17
图 2-6 线性摄像机成像模型 .....	18
图 2-7 成像过程 .....	19
图 2-8 激光测距仪与相机安装示意图 .....	22
图 2-9 标定实验设备 .....	23
图 2-10 不同位姿的棋盘标定板 .....	23
图 2-11 激光点投影变换 .....	24
图 3-1 激光测距仪返回的原始数据可视化 .....	25
图 3-2 从激光点中产生候选目标 .....	26
图 3-3 人体在图像上的显示比例 .....	29
图 3-4 二维标定板 .....	29
图 3-5 候选行人区域投影结果 .....	30
图 3-6 HOG 特征和人体检测示意图 .....	31
图 3-7 检测结果示意图 .....	32
图 4-1 集群点与单一点示意图 .....	38
图 4-2 改进二部图的匹配 .....	40
图 4-3 方向模型 .....	42
图 4-4 运动模型 .....	42
图 4-5 遮挡中的目标关联 .....	43
图 4-6 遮挡模型求解示意图 .....	44
图 4-7 SDL-1 的跟踪结果 .....	47
图 4-8 SDL-2 的跟踪结果 .....	47
图 4-9 SYNC 的跟踪结果 .....	48

## 表目录

表 2-1 多种传感器比较 .....	11
表 3-1 激光模式特征 .....	27
表 3-2 基于激光信息分类算法 .....	28
表 4-1 跟踪模型求解算法 .....	41
表 4-2 数据集介绍 .....	44
表 4-3 评测标准 .....	45
表 4-4 对比实验结果 .....	46

## 第一章 绪论

### 1.1 课题背景和研究意义

随着我国社会和经济的现代化发展，近年来物联网，物流行业得到了极大地繁荣。汽车作为货物的载体和物流系统中重要的组成部分。而近年来汽车数量的猛增，不仅使城市道路不堪重负，也给城市交通安全带来严重的挑战，成为日益严重的社会问题。根据公安部交通管理局在网站上公布的统计数字，2010年，全国接报的道路交通事故总计 3906164 起，同比增加高达 35.9%。其中，涉及人员伤亡的道路交通事故有 219521 起，共造成 65225 人死亡、254075 人受伤，直接财产损失 9.3 亿元。计算下来，平均我国每天有 179 人因为交通事故死亡，696 人在事故中受伤。世界卫生组织的研究结果同时也表明，道路交通事故比空难、海滩事故严重得多，造成的死亡人数高于战争，带来巨大的经济损失。据统计，道路交通事故带来的经济损失在低收入国家约占国民生产总值的 1%，在中等收入国家为 1.5%，而在高收入国家为 2%；全球每年由道路交通事故造成的损失估计高达 5180 亿美元<sup>[1][2]</sup>。而在这些受害者中，绝大多数是缺乏保护的行人，行人是道路交通的主要参与者，同时也极易受到伤害。在近年来的国际智能车辆和智能交通会议中，交通安全保障技术已经成为研究的新热点，而其中的车载汽车安全辅助驾驶系统的开发更是受到了广泛的关注。

汽车安全辅助驾驶是智能交通系统(Intelligent Transport System, 简称 ITS)中的一项主要研究内容，该系统主要利用传感器技术、计算机技术、通信技术等先进技术，掌握驾驶员、道路以及车辆周围行人其他车辆的状态信息，并做出综合的分析和判断，以便在存在安全隐患的情况下给司机发出预警信息，提醒司机注意规避危险。目前，该研究主要从以下几个方面<sup>[3]</sup>展开：

1、行人检测跟踪：行人检测跟踪是通过机器视觉、红外线以及激光测距仪等传感器感知车辆前方的行人，并在必要情况下向驾驶员发出警告以避免与邻近车辆的行人发生碰撞。这一研究对于保障行人的安全具有及其重要的意义。

2、车辆检测跟踪：车辆的检测跟踪是利用多种传感器探测车辆前方、侧方、

后方车辆的信息，包括它们的速度、位置、和当前车辆的车距等。这一技术主要是为了避免车辆因车距不足导致的追尾、碰撞等交通事故。

3、车道检测：车道检测主要是利用机器视觉、激光等传感器或埋设于路面下的磁钉，使车辆始终在车道线内运行，防止车辆因为偏离当前的行驶车道而导致交通事故的发生，从而提高行车的安全性。

4、交通标志检测：道路交通标志是重要的道路交通安全辅助设施，可向驾驶员提供各种引导和约束信息。实现道路交通标志信息的实时准确获取，为行车安全提供了至关重要的保障。在汽车安全辅助驾驶系统中，交通标志信息的检测通常是通过图像识别系统实现的。

5、视觉增强：人眼的作用能力是十分有限的，当出现暴风雨、暴风雪、大雾、沙尘、光线不足等情况时，仅凭人眼通常很难清楚辨识前方路面的情况。视觉增强的研究目的即为解决以上问题，以便为处于恶劣环境中的驾驶员提供良好的视野，保障其驾驶安全。视觉增强的实现有两种方法，一种采用 CCD、红外传感器、激光波雷达及 GPS 等传感器进行道路信息获取和融合，提取低能见度、低照度下交通环境的有用信息并为驾驶员提供实时的图像显示；另一种方法则通过除去挡风玻璃上的雨水和霜、提高汽车前照灯的智能化管理等手段，改善驾驶员的视觉环境，为其提供良好的视野。

本课题的研究受以下项目资助：

- 1、“基于多源数据的飞行器进近威胁目标检测跟踪及行为预测”，国家自然科学基金重点项目（课题编号：61039003），2011.01-2014.12，在研。
- 2、“飞行器威胁目标识别与图像鲁棒匹配理论与方法”，国家 973 计划子课题（课题编号：2010CB731804-2），2010.01-2014.12，在研。

本文所研究的激光深度与图像信息融合的目标跟踪与基于多源数据的飞行器威胁目标跟踪具有方法上的共同之处。

## 1.2 国内外研究现状

### 1.2.1 车辆辅助安全驾驶系统

鉴于多传感器信息融合在汽车安全辅助驾驶中的重要性，国内外众多汽车制造商、大学和科研院都将多种传感器应用于其开发的汽车安全辅助驾驶系统

上, 并取得了积极的成果。进入二十一世纪, 许多汽车生产厂商、大学以及研究机构陆续开展了汽车安全辅助驾驶中行人检测跟踪技术的研究。针对这一研究普遍使用的传感器包括: 激光测距仪、微波雷达、红外摄像头、光学摄像头等。这些传感器可对车辆周围环境进行非接触式探测, 以获取行人等障碍物及其轮廓、速度、和车辆的距离等信息。本文根据不同传感器的使用状况, 对国内外现有的行人检测跟踪系统做了简要的介绍。

从 20 世纪 70 年代开始, 激光测距传感器就被应用到智能车辆和汽车安全辅助驾驶领域, 主要用来进行障碍物检测。使用激光测距仪的典型装置有德国乌尔姆大学的 K.Ch.Fuerstenberg 与 IBEO 汽车传感器公司<sup>[4]</sup>研究开发的 IBEO 激光测距仪。该扫描器安装在车辆前方的正中央, 具有 180 度的视角, 以便能够扫描到所有相关的目标物体。目标的深度可以根据每次扫描的位置信息获得。使用激光测距仪, 可以准确、快速地获得目标物体的距离, 并根据宽度信息进一步地将目标物分成行人、自行车、轿车、卡车等类别<sup>[5]</sup>。美国军方研究的 DEMO-III 智能车辆以及德国 Braunschweig 科技大学开发研制的智能车辆均采用了雷达与机器视觉融合技术用于障碍物探测和跟踪<sup>[6][7]</sup>; 丰田公司使用毫米波雷达和视觉传感器共同探测前方障碍物<sup>[8]</sup>; 谷歌公司在 2010 年开发的自动驾驶汽车采用摄像机、雷达传感器和激光测距仪, 并通过详细的地图指引汽车在路上行驶<sup>[9]</sup>。

在国内, 从 20 世纪 80 年代末开始介入智能车辆的研究, 研究的领域主要集中于路径识别、自动转向等自主导航技术, 对智能车辆安全辅助驾驶技术还没有进行过深入系统的研究开发。目前国内从事智能车辆的研究单位主要有清华大学、北京大学、吉林大学、国防科技大学、沈阳自动化研究所、长安大学、北京理工大学等。清华大学的汽车安全与节能国家重点实验室研究的汽车主动碰撞系统采用了激光雷达和毫米波雷达, 实现了对车辆前方车辆的检测跟踪<sup>[10]</sup>; 吉林大学研制的 JUT 系列试验车利用视觉信息和激光雷达信息融合技术实现路径识别、跟踪和保持本车与前方车辆安全车距等功能<sup>[11]</sup>; 国防科技大学以轻型越野汽车为平台, 集成安装了激光雷达、视觉、组合导航系统等多种传感器, 研制成功了我国第一辆紧凑结构轻型越野自动驾驶汽车<sup>[12]</sup>。在对城市道路交通环境下的行人运动检测与跟踪方面, 中科院自动化研究所在跟踪人体

运动中使用了运动模型和人体关节模型对行人进行建模分析：浙江大学人工智能研究所也对行人运动分析进行了一定的研究，采用视频图像序列对人体被遮挡部位的动作进行跟踪，首帧图像采用手工标注特征点，但是由于该方法基于单目视觉技术，无法准确估计被遮挡部位的位置，并且人工干预较多。而清华大学的智能车辆 THMR-V 和吉林大学的智能车辆 JUTIV- III，虽然也在校园道路或城市道路中进行过试车，但它们仅考虑了道路上的静止障碍物，对运动的障碍如车辆和行人考虑较少。

总的来说，我国在智能车辆以及汽车安全辅助驾驶的开发研究尤其是行人的检测跟踪方面，近几年内虽然取得了一些可喜进展，但由于受经济条件投入等各方面因素的限制，和国外相比而言还存在一定的差距。

### 1.2.2 信息融合的基本理论

信息融合（又称为数据融合）是一种对多种信息的获取、表示及对其内在联系进行综合处理和优化的技术，它充分利用多种数据资源，通过对各种数据信息的合理支配与使用，将各种数据在空间上和时间上的互补与冗余信息依据某种优化准则组合起来，产生对观测环境的一致性解释和描述<sup>[13]</sup>。信息融合方法于 20 世纪 70 年代被提出，首先应用于军事领域。在早期的研究中，融合的对象主要集中于传感器数据，因此又被称为传感器信息融合。单一传感器只能获得环境或被测对象的部分信息段，而多传感器信息经过融合能够完善准确的反映环境特征。多传感器信息融合技术通过对这些传感器在时间和空间上的冗余或互补信息依据某种准则进行组合，以获得被观测对象的一致性解释或描述。随着融合技术的发展，传感器的意义成为了广义上的，数据来源也更加广泛，物理传感器数据，软件传感器数据，历史数据，先验信息等数据逐渐进入到融合系统中<sup>[14]</sup>。

按照融合系统中数据抽象的层次，融合可划为三个级别：像素级融合、特征级融合和决策级融合<sup>[15]</sup>。像素级融合是指在融合算法中，要求进行融合的传感器数据间具有精确到一个像素的匹配精度的任何抽象层次的融合。特征级融合是指从各个传感器提供的原始数据中进行特征提取，然后融合这些特征，一般来说，提取的特征信息应该是数据信息的充分表示量或充分统计量。决策级融合是指在融合之前，各传感器数据源都经过变换并获得独立的身份估计，信

息根据一定准则和决策的可信度对各自传感器的属性决策结果进行融合，最终得到整体一致的决策<sup>[15]</sup>。决策级融合是一种高层次的融合，直接针对具体决策目标，充分利用特征级融合所得出的各类特征信息，并给出简明而直观的结果。

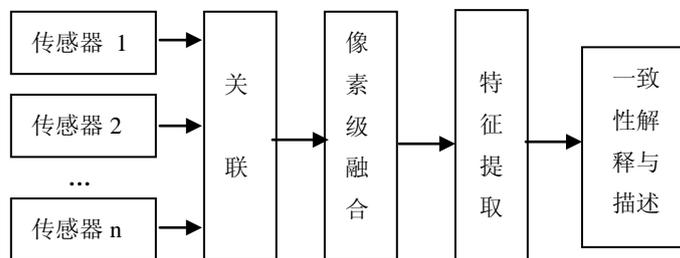


图 1-1 像素级融合

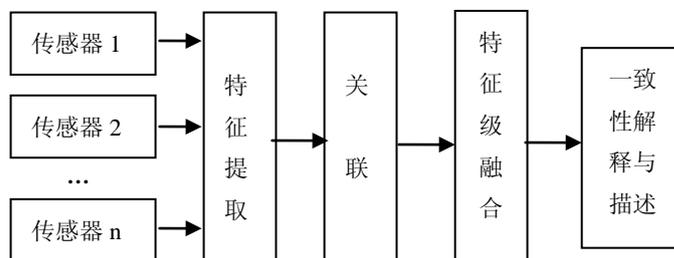


图 1-2 特征级融合

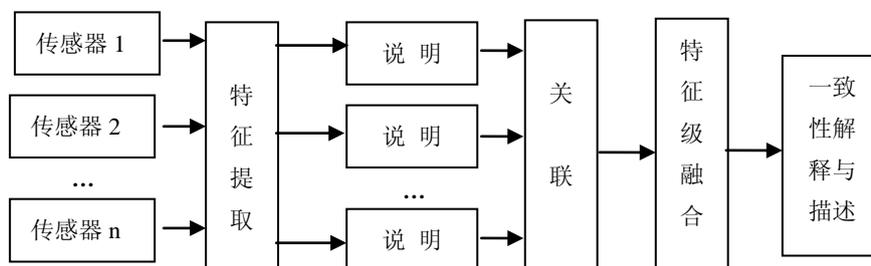


图 1-3 决策级融合

这三种融合层次各有优缺点：像素级融合尽可能多的保持了原始信息，能够提供其他两个层次融合所不具有的微小信息，但面临的是大量信息，处理代价高且实时性较差，又因为融合在信息最低层进行的，传感器的原始数据具有不确定性、不完全性和不稳定性，要求在融合过程中有较高的纠错能力；特征级融合实现了客观的数据压缩，保留原始信息主要特征，有利于实时处理，但由于损失了有用细节信息，使得融合性能有所下降；决策级融合则具有容错性强、通信中信息量小、抗干扰能力强、融合过程处理数据代价低等优点，其缺点是

预处理开销大，数据损失量最大，精度低。

### 1.2.3 多目标跟踪

多目标跟踪是利用传感器所获得的测量对多个目标的状态进行持续精确（或优化）的估计和预测。它通过综合利用统计估计、随机决策、控制以及智能计算等现代科学理论，将经过初级处理的传感器数据依据一定规则进行判决和关联，分解为对应虚警、新目标和已确认目标等不同信源产生的观测集合，从而判定被跟踪的目标数目并估计出各目标的状态参数。

单传感器多目标跟踪的功能要素包括轨迹起始、数据关联、状态估计、轨迹终止。轨迹起始是一种建立新的目标轨迹文件的决策方法，包括假定轨迹形成、轨迹初始化和轨迹确认。数据关联是根据一定原则判定测量来源是属于虚警或新轨迹还是已确认的目标轨迹。状态估计则是利用关联结果估计出每条轨迹的状态。

避免资源浪费，当被跟踪的目标离开所跟踪范围将被摧毁后，系统必须能够自动终止轨迹并删除这些已终止轨迹的文件。上述要素中，核心是数据关联和状态估计。如果不考虑其它因素，多目标跟踪就是单纯的估计问题。然而在实际跟踪过程中往往存在多种干扰因素，例如目标的随机机动、存在杂波的测量环境、非分辨测量、编队目标跟踪、非均匀杂波环境<sup>[16]</sup>、非高斯的噪声污染<sup>[17,18,19]</sup>、被动目标跟踪<sup>[20]</sup>、非顺序测量数据处理<sup>[21]</sup>等，使得这一问题不再是简单的估计问题。目标运动状态的不确定性和测量来源的不确定性是多目标跟踪系统所必须克服和解决的两个难点，也就是通常所说的机动目标跟踪问题和数据关联问题。

现在绝大部分的算法是基于动态背景下单目标<sup>[22,23,24,25]</sup>或者多目标同时有些又利用了红外，激光等信息进行跟踪。这些跟踪算法中，部分算法是针对缩小目标搜索范围提出的，这类算法主要是通过滤波方法预测目标下一个时刻可能出现的位置，从而通过缩小目标的搜索范围来缩短目标搜索时间以满足系统实时性的要求；部分是针对提高跟踪匹配的准确性而提出的，这类算法主要是通过选取稳定的、高区分性的特征，并可以与红外、激光等信息相融合，从而提高对目标的搜索速度和匹配的正确度。

### 1.3 基于信息融合的目标跟踪的关键问题

信息融合作为一种信息综合处理技术，实际上是许多传统学科和新技术的集成。为了进行信息融合，需要将多维数据在不同融合层次上采取不同的数学方法对数据进行综合处理，最终达到融合的目的。而目标跟踪系统需要对多元信息进行筛选，进而利用融合方法建立跟踪模型完成目标跟踪任务。其中要解决的几个关键问题：

1、数据对准。在多传感信息融合系统中，每个传感器提供的观测数据都在各自的参考框架之内。在对这些信息进行组合之前，必须首先将它们变换到同一个参考框架中去。但要注意的是，由于多传感时空配准的舍入误差必须得到补偿。

2、同类或异类数据。多传感器提供的数据在属性上可以是同类也可以是异类的，而且异类多传感器较之同类传感器，其提供的信息具有更强的多样性和互补性；但同时由于异类数据在时间上的不同步，数据率不一致以及测量维数不匹配等特点，使得对这些信息的在跟踪模型中处理更困难。

3、传感器观测数据的不确定性。由于传感器工作环境的不确定性，导致观测数据包含有噪声成分。在融合处理中需要对多源观测数据进行分析验证，并补充综合，在最大限度上降低数据的不确定性。在跟踪系统中对于不确定数据的取舍要有一定标准

4、不完整、不一致及虚假数据。在多传感信息融合系统中，对传感器接收到的量测数据有时会存在多种解释，称为数据的不完整性；多传感数据往往也会对观测环境做出不一致甚至相互矛盾的反馈；另外，由于噪声及干扰因素的存在，往往存在一些虚假的测量数据。跟踪系统需要能够对这些不完整数据、不一致数据以及虚假数据进行有效的融合处理。

5、数据关联。数据关联问题广泛存在，需要解决单传感时间域上的关联问题，以及多传感空间域上的关联问题，从而能确定来源于同一目标源的数据。跟踪模型中，尤其在完成多目标跟踪时，对于多个数据返回值，对同一目标的对应关系，成为跟踪系统的主要任务。

6、粒度。多传感器提供的数据可能是在不同的粒度级别上。这些数据可以

是稀疏的，也可以是稠密的；它们可能分别处于数据级、特征级，或是符号级等各种不同的抽象级别上，所以一个可行的跟踪系统应该可以工作在各种不同的粒度级别上，将这些多粒度数据在一个框架下进行融合。

#### 1.4 本文的研究内容

基于多传感器信息融合的多目标跟踪可以有效的解决单一传感器普遍存在的有效探测范围小、数据可靠性低以及易受外部环境影响等问题，对于提高多目标跟踪的准确性，可靠性具有及其重要意义。本文利用多元信息融合技术进行多目标关联，汲取了激光测距仪的深度信息，弥补二维图像缺失的像素深度值，从根本上解决了单一视觉技术处理图像的难题。而在驾驶环境中，利用多传感器的数据测量对车辆前方行人进行检测与跟踪，并对潜在的危险进行预警，为实现辅助安全驾驶是本文的主要目的。

本文研究的主要内容：

1、搭建了安全辅助驾驶系统原型平台。本文利用多传感器和移动机器人平台架设了模拟平台，并在真实交通场景中进行了数据采集。激光测距仪和 CCD 摄像头固定安装后，本文对两种传感器进行了统一的坐标参数配准，并进行了实验验证。该系统对大雾，雨雪，光照差等环境同样适用。

2、利用激光分类模式和视觉方法进行行人检测。本文利用新的激光分类模式，在激光坐标系下对目标进行滤波，聚类，模式分析，以及粗分类，得到候选行人区；经过投影转化到图像层后，利用 HOG+SVM 方法对候选区域进行小尺度变化的确认性检测识别。

3、提出了基于信息融合的多目标跟踪模型。本文根据信息融合中数据关联基本理论，推导出多行人跟踪模型，充分融合了激光的深度信息和图像信息，进行模型的在线更新。并利用改进二部图匹配原理提出了多项式时间的求解方法。

4、解决复杂遮挡条件下的目标跟踪问题。本文利用激光测距仪的深度信息，在跟踪模型上通过深度坐标的动态参数对多行人进行权值的调整，使得求解过程与客观视觉观测条件一致，合理解决遮挡条件下多目标跟踪丢失，错误标号等问题。

本文用激光和视觉图像融合进行前方多行人跟踪，通过上述问题的求解。不仅可以克服单独应用视觉图像传感器进行目标跟踪时易受天气状况和光照条件变化影响且无法得到检测跟踪对象的深度信息的缺点，也可以克服激光测距无法判断障碍物类别，无法可视处理，冗余报警的缺点。系统同时使用两种不同种类的传感器，利用各自的优点完成对路面行人的快速检测跟踪，实现优势互补，解决单一传感器无法克服的困难。

## 1.5 本文的组织结构

第一章，绪论。论述基于多传感器信息融合的行人跟踪的研究背景和意义，分析国内外基于多传感器信息融合行人跟踪的研究现状，同时描述智能交通中汽车辅助安全驾驶系统的现状及发展趋势，列出本文的主要研究目的和研究内容。

第二章，实验平台原型搭建和多传感器配准。在列举了不同传感器之间参数性能比较后，介绍了本文选择的激光测距仪和视觉摄像头，以及他们的具体参数配置。详细介绍了本文用来模拟真实驾驶环境下的实验平台。

第三章，基于激光深度信息的快速行人检测算法。主要论述了激光传感器和视觉传感器联合在用于行人目标检测方面的优势、具体特征选取方法以及融合方法。详细描述了激光分类模式和候选行人区投影以及视觉检测方法。

第四章，基于激光和视觉信息融合的多目标鲁棒跟踪。主要描述了本文提出的多目标实时跟踪算法模型，并详细介绍了模型的求解过程，在线更新，以及在复杂遮挡情况下，深度信息对遮挡问题的模型具体求解的作用。最后在三个数据集分别对本文的跟踪模型以及另外三种方法进行实验，实验结果显示了本文提出的跟踪模型具有良好的稳定性，实时性。

最后，总结了本文的主要工作，对未来工作的方向的展望，以及对如何提高行人，汽车等多种目标检测跟踪算法的鲁棒性、准确性等研究热点和难点进行了探讨。



## 第二章 车辆辅助安全驾驶原型平台及多传感器配准

目前行人检测跟踪领域中常用的传感器有光学摄像头、红外摄像头、立体摄像机，激光测距仪等，从使用传感器的角度而言，现有的行人检测跟踪工作可以分为两大类：第一类是产业界的一些汽车厂商为代表，除了基本的光学摄像头，还尝试综合运用红外摄像头，立体摄像机以及激光雷达等较为昂贵的传感设备，期望通过多传感器的信息融合来获得更好的综合检测性能；另一类是研究者，更专注于简单，廉价且容易普及的光学摄像头，试图通过算法软件层的优化，在保证实时检测跟踪的基础上争取达到较好的实时性能。

### 2.1 传感器对比

表 2-1 多种传感器比较

传感器类型	优点	缺点
红外摄像头	1、适合夜间工作，也可用于雨雪等特殊天气； 2、可获得目标的轮廓信息。	1、测量的有效距离和视角范围有限； 2、容易受到热源干扰； 3、精度，速度，稳定度不如激光； 4、成本比光学相机高，但低于立体摄像机，激光测距仪。
光学摄像头	1、可以获得目标物的轮廓，纹理，颜色等信息； 2、检测有效距离比较大，视觉较广； 3、价格低廉，维护方便，易于普及。	1、视频数据处理复杂，处理速度偏慢。 2、光照，天气，遮挡影响显著。 3、双目视觉得到距离信息在稳定度，精度，速度方面不如激光。
激光测距仪	1、测定目标距离和速度优势明显； 2、适用于夜晚，雾霾，雨雪等天气。	1、无法获得目标的纹理，颜色，轮廓信息。 2、成本相对较高。

表 2-1 列举了目前行人检测跟踪研究领域中常用的多种传感器，并详细对比了他们的优点和缺点。从表中可以看到，几种主要的传感器在性能上都有各自的优势和不足。光学摄像头价格低廉，容易普及，在信息处理角度而言也可以通过获取目标的轮廓、纹理、颜色等信息有效的提高检测效果，视角宽度和

测量的有效距离也是一大优势；但同时，视频数据处理过程复杂，光照、天气、地形、人流等路况因素也给检测带来很大困难，实际系统中这些问题必须考虑。相比之下，红外摄像头的系统测量的有效距离较低，成本较高，且容易收到热源的干扰，但是红外摄像头优势在于夜间环境以及雨雪等特殊天气下的测量，在这些特殊情况下，红外信息可以更好的描述当前的道路交通状况，在处理技术方面，红外摄像头与光学摄像头也有相似之处。激光测距仪，其在测定目标距离和行进速度方面有着无可比拟的优势，但无法获得目标的纹理和全部的轮廓等信息，不利于目标的进一步分类。综合而言，激光测距仪与其他传感器在性能上有所互补，更适合与其他传感器搭配使用。

本文搭建的实验平台充分利用光学摄像头与激光测距仪的优势互补进行行人检测跟踪系统的研究，在吸取现有处理方法优势的同时，着重用于交通复杂场景下的多行人遮挡条件下的跟踪。

## 2.2 实验平台搭建

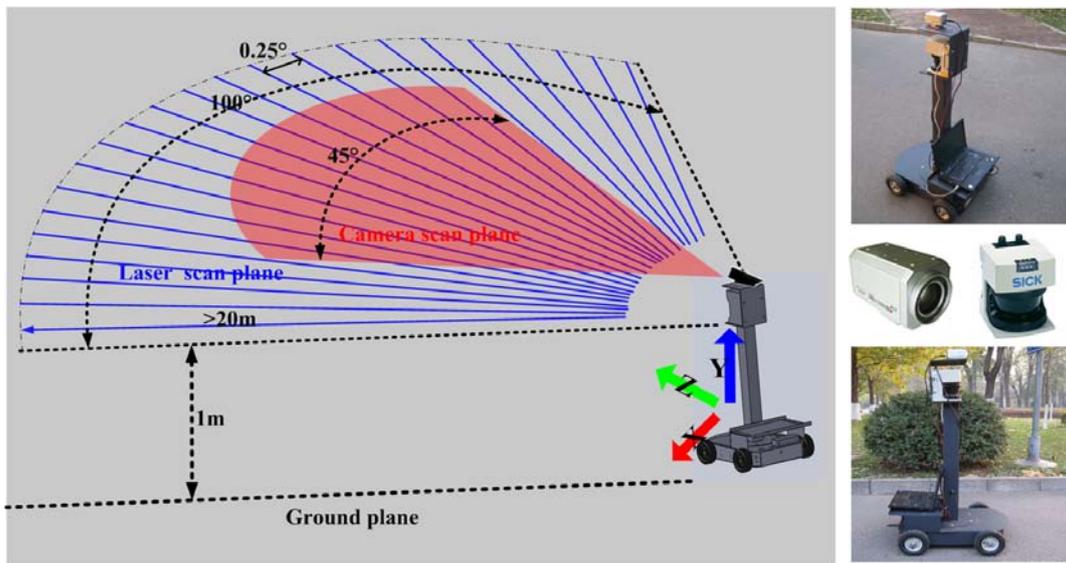


图 2-1 系统平台搭建

在选择了合适的激光测距仪和 CCD 摄像头之后，本文搭建了用于模拟驾驶员在真实驾驶场景下的观测环境障碍物应大致位于图像中间部分，经大量试验验证，本文激光测距仪安装高度为地面以上 1m 处，摄像头的高度在激光测距仪正上方 20cm 处，如图 2-1 所示。模拟移动平台在校园道路进行移动，完

成对行人和车辆等其他障碍物的扫描。

### 2.2.1 激光测距仪 LMS291 介绍

激光传感器是在 20 世纪 60 年代发展起来的一种新型测距雷达，现已进入实用化阶段。激光雷达的波束窄、波长短等独特优点使激光具有极高的角分辨能力、距离分辨能力和速度分辨能力。此外，激光雷达可以获取目标的反射特性、距离信息、速度信息等信息。作为主动式测距成像系统的激光测距仪，照射条件、光线对其基本没有影响。激光测距仪通过二维扫描激光束或者三维光面，获取大量、准确的空间距离信息。因此目前激光测距仪在机器视觉、智能车辆及机器人等领域都有广泛的应用<sup>[26]</sup>。

本文使用的激光测距仪是德国 SICK 电子公司生产的脉冲式激光测距系统 LMS291。该激光的扫描系统不需要额外的反射镜和位置标记。内部安装的旋转反射镜决定了发射脉冲激光束探测的是周围环境的扇形区域，接收到的反射光的次序决定目标形状的轮廓。其自动滤波修正功能对户外使用具有鲁棒性，可以使返回结果不受雨滴和雪花等影响。图 2-2 (a)和图 2-2 (b)分别为 LMS291 的外观图和激光扫描的示意图，图中激光光束用蓝色部分表示，障碍物用中间的灰色矩形方块表示。

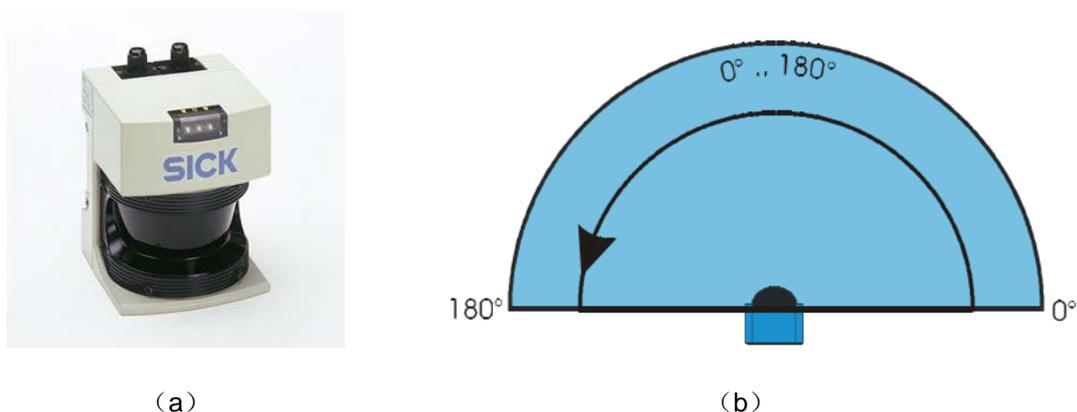


图 2-2 激光测距仪和扫描示意图

#### (4) 激光的视场角选择

本文中选择的视场角的参数是：视场角为  $100^\circ$ ，精度为  $0.25^\circ$ ，401 个激光数据点。视角选  $100^\circ$  时，能够获取车辆前方可能会对车辆造成威胁的目标。另外，线扫描面的高度为距离地面 0.9m。

(5) 激光深度原始数据获取

激光发射器发出激光脉冲波，当激光波碰到物体后，部分能量返回，当激光接收器收到返回激光波，且返回波的能量触发阈值，激光测距仪即可计算它到物体的距离值；激光测距仪连续不断地发射激光脉冲波，这些激光脉冲波打在高速旋转镜面上，将激光脉冲波发射向各个方向从而形成一个二维区域的扫描面。在激光测距仪的扫描范围内，输出每个测量点的距离，根据这些距离信息，则可以计算物体的外形轮廓、坐标定位等<sup>[27]</sup>。将光速为  $c$  的激光射向被测目标，测量它返回的时间，由此求得激光器与被测目标间的距离  $s$ 。如图 2-3 所示为激光测距原理图。

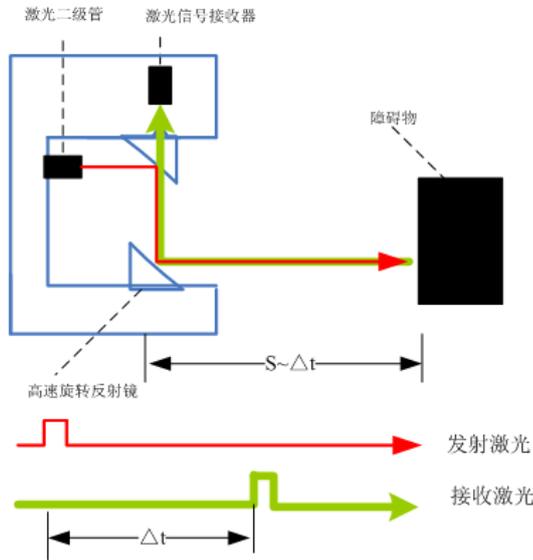


图 2-3 激光测距原理图

激光测距仪所返回的数据是二维扫描平面内一组长度有限的离散数据点，其中每个数据表示在该元素对应角度方向上最近目标的距离值。具体的离散数据点与测距仪的分辨率有很大关系。扫描数据由极坐标形式给出 ( $d$  表示距离， $\phi$  表示扫描角度)：

$$l_n = (d_n, \phi_n)^T, n = 1, \dots, N. \quad (3-1)$$

将其转换为笛卡尔坐标系表示为：

$$u_n = (x_n, y_n)^T, n = 1, \dots, N. \quad (3-2)$$

其中  $x_n = d_n \cos \phi_n$ ,  $y_n = d_n \sin \phi_n$ ,  $N$  为扫描数据点个数，与激光测距仪的角度分

分辨率有关。

激光测距仪的测距性能受到物体表面光学特性以及混合像素现象影响<sup>[2]</sup>，激光扫描数据中存在大量的噪声点和无效点，要通过激光测距仪进行目标检测，需要考虑如何精简数据，从大量数据中提取有用信息。通过聚类方法对大量激光数据点进行处理并从中提取有用信息，是一种行之有效的方法。

### 2.2.2 单目摄像头介绍

摄像机镜头的选择主要包括以下两个方面：

#### (1) 视场角（焦距）参数的选择

镜头的焦距和视场角是一组对应且矛盾的参数。焦距变大，则视场角会变小，即视野范围小，而远距离处的物体分辨率较高；反之，焦距变小时，则视场角大，即视野范围变大，此时图像的畸变很严重导致远距离处的物体分辨率较低，CCD 摄像机成像会变得不清楚。本文选用维视公司 VS-870HC 工业数字相机如图 2-4 (b)所示。

本文根据真实驾驶环境的实际情况，为了保证激光测距仪与 CCD 摄像机的视场范围可以达到相互配合的效果，考虑到本文针对车道正前方向上的障碍物给予更大的关注，这样要求 CCD 视野区应覆盖整个前方道路区域，即视场范围应大一些。最终选择的水平视场角  $45^\circ$ ，焦距为 16mm。



(a)



(b)

图 2-4 传感器相对位置和摄像头

#### (2) 摄像机安装高度（俯仰角）的选择

当 CCD 镜头俯仰一定的角度时，其主光轴与前方路面相交于一定的距离处。俯仰角及其安装高度的选择标准是保证摄像头采集前方 2~20m 内的障碍物

在视野中，本文选择摄像头水平于地面，高度模拟驾驶员坐在驾驶位置时的高度，如图 2-4 (a)所示。

### 2.3 基于激光测距仪与单目摄像头的配准

客观世界的三维场景投影到普通相机的二维平面是损失深度信息的透视投影过程，因此通过 CCD 相机拍摄的图片可以捕获周围环境而获取二维视觉信息，而激光测距仪可以提供准确的深度信息，合理的融合二者所提供的信息，可以用于对周围环境中特定障碍目标的检测与识别。

激光测距仪和视觉传感器的数据表示方法不同，这导致了异类数据对准的困难。在融合摄像机和激光测距仪数据时，首要考虑的是二者之间数据的统一标定，目的是为了确定激光空间扫描平面在世界坐标系中的位置，从而解决激光测距仪和摄像机之间的数据对准问题。空间物体表面某点的三维几何位置与其在图像中对应点之间的相互关系是由摄像机成像的几何模型决定的，这些几何模型参数即为摄像机参数，本文通过实验与计算来求解这些参数，即为摄像机标定。下面给出摄像机标定理论方法。

#### ■ 摄像机标定所涉及的坐标系

在计算机视觉中，图像是空间物体通过成像系统在像平面上的反映，即空间物体在像平面上的投影。图像上每一个像素点的灰度反映了空间物体表面某点的反射光的强度，而该点在图像上的位置与空间物体表面对应点的几何位置有关。成像变换涉及到不同坐标系之间的转换，为了定量的描述相机成像过程，定义以下四个坐标系：图像物理坐标系、图像像素坐标系、摄像机坐标系与世界坐标系<sup>[28] [29]</sup>。

##### (1) 像素坐标系

像素坐标系代表着计算机内部的数字图像所用的坐标系。摄像机采集的图像以标准电视信号的形式输入计算机，经计算机数模转换板变换成数字图像，该数字图像在计算机内为  $M \times N$  数组， $M$  行  $N$  列图像的每一个元素数值为图像像素点的亮度。如图 2-5 所示，图像坐标系定义为  $u_0v_0$ ，坐标系原点  $o_0$  位于图像左上角，其中每一像素坐标  $(u, v)$  为该像素在数组中的行和列数。

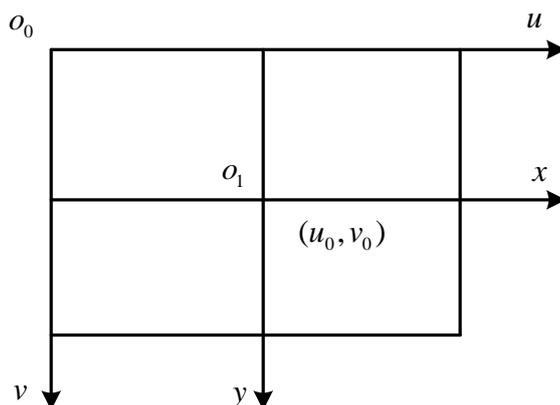


图 2-5 图像像素坐标系和物理坐标系

### (2) 物理坐标系

物理坐标系的原点为透镜光轴与成像平面的焦点，是在摄像机内形成的平面坐标系。在图像像素坐标系中，由于点  $(u, v)$  并未用物理单位表示出该像素在图像中的位置，因此建立以物理单位（如毫米）表示的物理坐标系  $xy$ ，该坐标系以图像内某一点  $o_1$  为原点， $x$ 、 $y$  轴分别与  $u$ 、 $v$  轴平行。如图 2-6 所示， $xo_1y$  坐标系中，原点  $o_1$  定义为摄像机光轴与图像平面的交点， $(x, y)$  表示以毫米为单位的像素坐标系的坐标。在像素坐标系中，若  $o_1$  在  $uv$  坐标系下的坐标为  $(u_0, v_0)$ ，每个像素在  $x$  轴与  $y$  轴方向上的物理尺寸为  $dx$  和  $dy$ ，则图像坐标系与像素坐标系点的关系为：

$$u = \frac{x}{dx} + u_0, \quad (2-3)$$

$$v = \frac{y}{dy} + v_0. \quad (2-4)$$

矩阵形式为：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (2-5)$$

### (3) 摄像机坐标系

摄像机坐标系以小孔相机模型的焦距中心为原点，以相机光轴为  $z_c$  轴建立相机坐标系。如图 2-6 所示， $o$  点为摄像机光心作为坐标系原点， $x_c$  轴  $y_c$  轴与图像坐标系的  $x$ 、 $y$  轴平行， $z_c$  轴为相机光轴，与图像平面垂直且其与图像平面的交点为像平面坐标系原点  $o_1$ 。

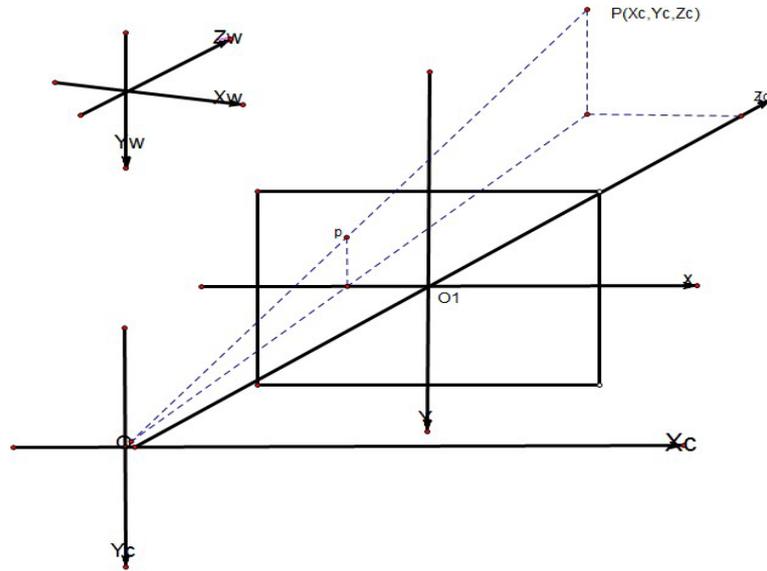


图 2-6 线性摄像机成像模型

#### (4) 世界坐标系

世界坐标系也称真实世界坐标系，全局坐标系，是客观世界的绝对坐标，可以是任意定义的三维空间坐标系。需要在环境中选择一个基准坐标系来描述相机和其他物体的位置，它是客观世界的绝对坐标，由  $x_w, y_w, z_w$  轴组成，如图 2-6 所示。摄像机坐标系与世界坐标系之间的关系用旋转矩阵  $R$  与平移向量  $t$  来描述。空间中的点  $P$  在两个坐标系下的齐次坐标关系为：

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (2-6)$$

#### ■ 线性相机成像模型

在计算机视觉研究中，三维空间中的物体到像平面的投影关系即为成像模型，理想的投影成像模型是光学中的中心投影，也称针孔模型。

针孔模型假设物体表面的反射光都经过一个针孔而投影到像平面上，即满

足光的直线传播条件。针孔模型由光心（投影中心）、成像面和光轴组成。本文把针孔模型整理成数学形式上简单表示的等价形式。交换针孔和图像平面，针孔中的点为投影中心。这样，每一条光线从远处的某点出发，到达投影平面的中心。如图 2-6 所示， $oo_1 = f$  为摄像机焦距， $P = (X_C, Y_C, Z_C)$  由通过投影中心的光线投影到图像平面上，相应的图像点为  $p = (x, y)$ 。

由三角形的相似关系可得：

$$\frac{x}{f} = \frac{X_C}{Z_C}, \quad (2-7)$$

$$\frac{y}{f} = \frac{Y_C}{Z_C}. \quad (2-8)$$

矩阵表示为：

$$Z_C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix}. \quad (2-9)$$

从世界坐标系到图像坐标系的成像变换可以通过以下 3 个步骤来实现（本文未考虑摄像机畸变过程），如图 2-7 所示。

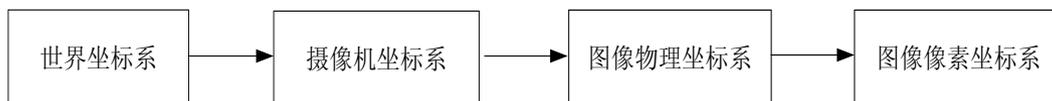


图 2-7 成像过程

矩阵表示为：

$$Z_C \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \quad (2-10)$$

整理得到：

$$Z_C \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_x & 0 & u_0 & 0 \\ 0 & a_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \quad (2-11)$$

其中,  $a_x = \frac{f}{dx}, a_y = \frac{f}{dy}, u_0, v_0$  只与摄像机内部结构有关, 这 6 个参数称为摄像机内部参数。

平移向量  $t = [t_x, t_y, t_z]^T$  表示世界坐标系原点在摄像机坐标系下的坐标。

矩阵  $R$  是正交旋转矩阵, 表示为:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}. \quad (2-12)$$

正交旋转矩阵含 3 个独立变量 (也就是 3 个方向的旋转角度), 加上平移向量中的  $t_x, t_y, t_z$ , 这 6 个参数为相机外部参数, 完全由相机相对于世界坐标系的方位决定。

#### ■ 摄像机外部标定

根据需要标定参照物与否, 摄像机标定分为传统的摄像机标定方法和摄像机自标定方法。传统的摄像机标定是在一定的相机模型下, 基于特定的实验条件, 如形状、尺寸已知的标定物, 经过对其进行图像处理, 利用一系列数学变换, 求取相机模型的内部参数和外部参数; 摄像机自标定方法不依靠标定参照物, 仅利用摄像机在运动过程中周围环境的图像与图像之间的对应关系对相机进行标定。因其中的未知参数过多而难以得到稳定结果, 故而自标定方法即使非常地灵活, 但却并不成熟。所以, 在应用场合所要求的精度很高且摄像机的参数不经常变化时, 传统标定方法为首选。且自标定方法主要应用于对精度要求是很高的环境。

传统的摄像机标定方法中, 基于 2D 平面标定板的摄像机标定 (又称为张正友标定法) 是一种适合应用的灵活方法。该方法需要摄像机在两个以上不同的方位拍摄同一平面标定板, 摄像机和 2D 平面标定板都可以自由移动, 且内

部参数保持不变。假设 2D 平面标定板在世界坐标系中的  $Z=0$ ，那么，通过线性模型分析可以计算出摄像机参数的优化解，进而用基于最大似然法进行非线性求精。在这个过程中得出考虑镜头畸变的目标函数后就可以求出所需的摄像机内、外部参数。这种标定方法具有较好的鲁棒性，且无需制作精度极高的标定板，具有极强可操作性和实用性。

张正友标定法<sup>[29]</sup>首先将式 (2-11) 表示如下：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = sA \begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{bmatrix}. \quad (2-13)$$

其中，假定二维平面标定板在世界坐标系中的  $Z_w=0$ ，因此  $r_3$  向量可以忽略。定义投影矩阵  $H$ ，将式 (2-13) 改写为：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix}, \quad H = sA \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}. \quad (2-14)$$

通过式 (2-13) 由定义的空间已知点及其图像点坐标，可解线性方程得到投影矩阵  $H$ 。然后利用投影矩阵  $H$  求解摄像机内部和外部参数，令  $H$  为：

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = sA \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}. \quad (2-15)$$

由式 (2-15) 中  $r_1, r_2$  相互正交得到内参数  $A$  的两个约束条件：

$$h_1^T A^{-T} A^{-1} h_2 = 0, \quad (2-16)$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2. \quad (2-17)$$

文献<sup>[29]</sup>中通过将  $A^{-T} A^{-1}$  改写为对称的  $B$  矩阵，由约束条件 (2-16) (2-17)，在  $B$  矩阵的封闭解中求得摄像机的内部参数。同时，每幅图像相对于标定板的外部参数可由内部矩阵和投影矩阵  $H$  求出。

#### ■ 摄像机与激光测距仪联合标定

激光测距仪与 CCD 相机标定技术<sup>[30]</sup>是在张正友提出的移动平面模板的标定方法的基础上，加入激光测距仪所在的激光坐标系（二维平面）。方法核心在于通过同一标定板同时被激光测距仪和摄像机捕获的方法来求得激光坐标系和

摄像机坐标系之间的旋转和平移矩阵。图 2-8 描述了两种传感器的安装位置：

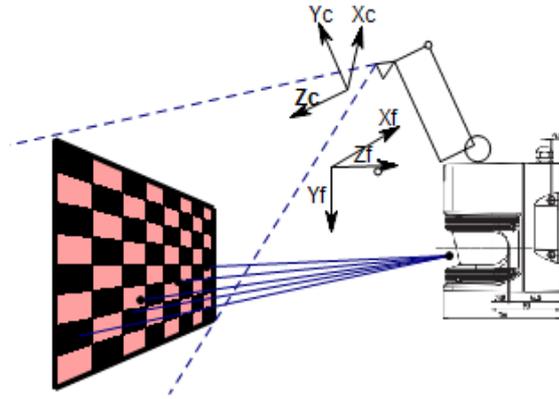


图 2-8 激光测距仪与相机安装示意图

从摄像机坐标系  $p$  到激光坐标系  $p^f$  的变换关系如 (2-18) 式，其中  $\Phi$  为旋转矩阵，表示两个传感器之间的相对方向； $\Delta$  为平移向量，表示两个传感器之间的相对位置。

$$p^f = \Phi p + \Delta \quad (2-18)$$

在相机坐标系下，将标定板参数化为一个 3 维向量  $N$ ，此向量的方向平行于标定板的法向方向， $\|N\|$  等于摄像机到标定板的距离。在摄像机坐标系下取标定板上任一点  $p$ ，由于  $p$  位于被参数化为  $N$  的标定板上，则有：

$$N \cdot p = \|N\|^2 \quad (2-19)$$

其中，由 (2-19) 可得到 (2-20)。N 可以通过相机的外参数矩阵  $[R, t]$  求出

$$N = -r_3(r_3^T \cdot t) \quad (2-20)$$

由 (2-18)、(2-19) 可推导出 (2-21)。通过变换标定板的位置，可以得到一组不同的  $N$  以及相应的激光点  $p^f$ ，即一组约束条件。

$$N \cdot \Phi^{-1}(p^f - \Delta) = \|N\|^2 \quad (2-21)$$

通过变换标定板的位置，可以得到一组不同的  $N$  以及相应的激光点  $p^f$ ，即一组约束条件。求解方程 2-21 可求得相机坐标系与激光坐标系之间的旋转和平移关系  $\Phi$  和  $\Delta$ ，即完成激光测距仪与摄像机的联合参数标定。

本文完成了联合参数标定实验。首先将两个传感器的相互位置固定，采集 10 种以上不同位姿的棋盘格图片（标定板大小  $10 \times 10$  格，每格 76mm），得到同步的激光距离值。如图 2-9，利用 CCD 摄像头与激光测距仪构成了基本的

测量系统。使激光测距仪的中心与标定板齐平进行数据采集，图 2-10 显示了本文采集的不同位姿黑白标定板的图像，并且得到了同步的激光距离数据。



图 2-9 标定实验设备



图 2-10 不同位姿的棋盘标定板

本文接下来利用采集的数据进行标定实验<sup>[57]</sup>。通过棋盘黑白标定板图像，利用张正友标定法的 matlab 工具包求得摄像机内参数矩阵以及每张图像对应的旋转矩阵和平移向量  $[R, t]$ 。并利用同步采集的激光距离值，利用公式  $N \cdot \phi^{-1}(p^f - \Delta) = \|N\|^2$  解方程求得  $\Phi$  和  $\Delta$ （其中， $N$  通过相机的外参数矩阵  $[R, t]$  求出，对应的激光点  $p^f$  从数据中读取）。最后，可得实验所固定位置下两个传感器的投影变换参数：即旋转矩阵  $\Phi$  和平移向量  $\Delta$ 。

$$\Phi = \begin{Bmatrix} -0.020806332201310 \\ -0.189339386091943 \\ -0.014229928182967 \end{Bmatrix}$$

$$\Delta = \begin{Bmatrix} 0.999600970302630 & 0.007238477362317 & 0.027303930404171 \\ -0.01157123350284 & 0.986716574692589 & 0.162038599667324 \\ -0.02576832794854 & -0.162289881608188 & 0.986406603588164 \end{Bmatrix}$$

利用求得的变换参数，利用  $p^f = \phi p + \Delta$  将激光点投影到图片上以验证结果是否正确，如图 2-11 所示：



图 2-11 激光点投影变换

图 2-11 中红色点为配准工作完成后，按照提出的坐标参数变换，图像上模拟标出的激光点位置，这与实验所用标定板的位置是基本重合的，证明上述坐标变换矩阵投影理论是正确的。

## 2.4 本章小结

本章完成了实验平台搭建和多传感器配准工作，为后续的检测跟踪搭好实际的操作平台。首先，本章列举了不同传感器之间参数性能比较，随后介绍了本文选择的激光测距仪和视觉摄像头，以及他们的具体参数配置。并详细介绍了本文用来模拟真实驾驶环境下的实验移动平台。最后推导了多传感器之间的转移矩阵变换方法，并用具体的实验验证了该配准方法是正确有效的。

### 第三章 基于激光信息的行人快速检测研究

在车辆辅助安全驾驶系统中，利用多传感器对行人进行跟踪之前首先需要在危险区域内实现障碍物的识别，即通过前视激光测距仪和视觉摄像头在前方对汽车可能造成威胁的障碍物的识别，从而实现了跟踪目标的自动初始化。通过对激光传感器所成前视距离图像的分析计算，可以识别出障碍物的位置和距离信息。本文采用的基于行人目标检测算法分为三个步骤：激光数据预处理，激光模式粗分类和视觉检测确认。

#### 3.1 激光数据预处理

为实现后续数据处理、行人的检测跟踪，本文中本文自行开发 Visual C++ 软件实现数据采集，并提取一帧扫描图像的有用数据点信息，图 3-1 为通过编程实现的直角坐标系下的激光数据图像。

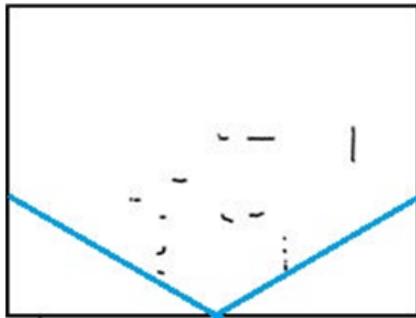


图 3-1 激光测距仪返回的原始数据可视化

激光测距仪以固定的频率扫描周围环境，所获得的距离信息反映了周围物体的深度，几何位置，形状。激光测距仪所返回的数据是其二维扫描平面内一组长度有限的离散数据点，其中每个数据点表示在该元素对应角度方向上最近目标的距离值，具体的离散数据点与激光测距仪的角度分辨率有关，本文利用每次返回的 401 个离散点距离数据，扫描数据由极坐标形式给出，即：

$$l_z = (d_z, \phi_z)^T, z = 1 \dots 401. \quad (3-1)$$

在笛卡尔坐标系表示为：

$$u_z = (x_z, y_z)^T, z = 1 \dots 401. \quad (3-2)$$

其中  $x_z = d_z \cos \phi_z, y_z = d_z \sin \phi_z$ 。激光测距仪对不同材质、不同颜色的物体存在测量误差，在无效数据中，很大一部分表现为最大测距范围的值，另外还有因阳光辐射、车辆颠簸及反射物体材质造成的噪声点，可通过滤波将这部分数据去掉。

令  $L = \{l_z\}$ ，其中  $z = 1, \dots, Z$  表示一组激光点距离数据，将集合  $L$  与核模板  $[-1, 1]$  卷积，点间距在一定范围内的点保留，即得到粗分类结果集合：

$$C = \{c_n\}, n = 1, \dots, N。$$

在激光测距仪提供的深度数据中，障碍物表现为相互靠近且长度在一定阈值范围内的一组点，因此首先需对离散的激光深度数据进行聚类分析。由于不能事先确定深度数据中包含的类别数目，因此使用层次聚类方法，自下而上划分。类与类之间的距离采用最短距离标准，即在每一步，两类之间的距离是它们两个最近点间的距离。

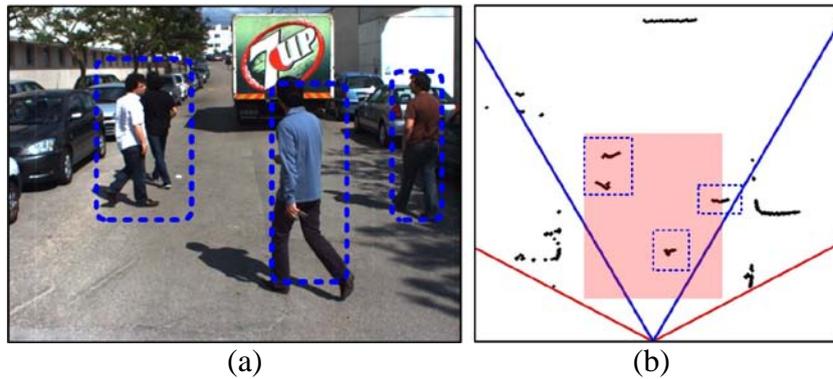


图 3-2 从激光点中产生候选目标

从粗分类结果  $C = \{c_n\}_{n=1}^N$  中聚类：聚类初始，将每一点作为一类，计算新类和单个样本间的距离。若相邻两类  $c_{n-1}$  和  $c_n$  的间距属于某一试算阈值范围，则认为它们在同一片段内，否则就将其分开，并以当前的单个样本数据点作为新增类的起始点。每个聚类片的特征值  $\{f_m, d_m\}$  分别表示为聚类中心位置和聚类片的长度，本文保留片段长度在  $T([t_1, t_2])$  范围内的聚类片段作为候选目标，其中，阈值  $T([t_1, t_2])$  通过对目标障碍物的扫描训练得到其长度平均值来设定。

在此过程中，大量的无效激光点被排除。本文得到的是表示类似障碍物的

一组候选集合  $S = \{s_m\}$ ，其中  $m = 1, \dots, M$ ，表示候选目标的数量。这些候选目标在横向长度上与所检测目标类似，其特征包括了深度、长度及位置信息。表 3-1 表示了候选目标生成步骤。图 3-2 表示了从激光点中产生的候选目标结果。

### 3.2 基于激光信息的粗分类

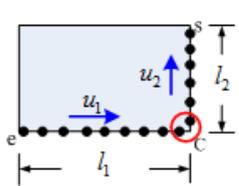
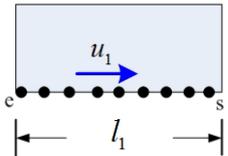
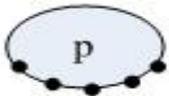
#### 3.2.1 激光信息

利用激光测距仪实现障碍物的识别，就是通过前视激光传感器对试验车前方能探测到的对汽车可能造成威胁的障碍物的判别。通过对激光测距仪返回的距离信息的分析计算，可以识别出障碍物的位置和速度信息。本文只讨论当障碍物为行人的情况，对采集到的激光距离数据的特点进行分析来识别行人，同时可获得行人的具体位置坐标、行人宽度和行人离汽车的距离。

在经过了上一节对激光数据的坐标系变换，聚类分析后，可以得到直角坐标系下如图 3-2 的片段激光点。从图中，当完成激光距离数据经过坐标变换和的聚类分析后，从图中的结果得到如下信息：

- 1，行人目标在距离图像中表现为小的团簇。
- 2，车辆的边缘具有标准的几何形状：线段或者弧线
- 3，目标（车辆，行人）与背景的背景边界非常明显。

表 3-1 激光模式特征

类型	结构	特征参数
双轴		C: 两轴交点 u: 一个轴的向量 l: 单个轴的长度
单轴		u: 单轴的向量 L: 单轴的长度
无轴		P: 中心点 r: 半径

对于车辆行驶过程中，路面可能出现的目标，本文在激光层根据聚类结果的激光轴数对目标做了如表 3-1 的分类<sup>[31]</sup>：

- 1， 两轴结构的目标，具有两条轴线的物体，且两个轴呈现一定的角度，一般两个轴上的激光点分布均匀且在直线上。这样的目标是车辆和墙角。
- 2， 单轴结构的目标，只有一条轴线，激光点分布均匀，且在一条直线上。具有这种结构的目标是正前方的车辆，或者是前方的一面墙。
- 3， 零轴结构的目标，没有明显的轴线，激光分布为一个团簇，可以计算这个团簇的聚类中心，和结构半径。具有单轴结构的目标很可能是道路上的行人，树木和垃圾筒等。

具体的激光模式分类算法在表 3-2 有详细描述。

表 3-2 基于激光信息分类算法

分类算法
输入： $L = \{l_z\}_{z=1}^Z$ 激光测距仪扫描一次输出激光数据
输出： $\{f_m, d_m\}_{m=1}^M \subset S, T$
<ol style="list-style-type: none"> <li>1 <math>T([t_1, t_2]) \leftarrow</math> 目标障碍物长度平均值；</li> <li>2 <math>C = \{c_n\} \leftarrow</math> 粗分割滤波 (<math>\{l_z\}</math>)；</li> <li>3 <math>S = \{s_m\}, \{f_m, d_m\} \leftarrow</math> 层次聚类 (<math>\{c_n\}</math>)。</li> </ol>

### 3.2.2 宽高比特征

激光测距仪只能得到目标的宽度，但是高度不能获得，行人在图像上的特征与真实的行人有所不同，行人在图像中的大小不仅受实际大小的影响，且反比于行人到镜头的距离。同一行人随着距离摄像机的距离远近，所成的图像大小不相同，距离越近，行人在图像中大小越大，反之越小。此外，图像中面积大小相同的行人在实际中可能并不相同。因此，仅仅利用行人的面积等特征识别行人会失效。但是激光测距仪能够返回目标的深度信息，利用这一点，经过小孔成像的原理，可以求得在特定深度上的行人在图像上的大小如图 3-3(a)所示。假设行人的深度为  $d$ ，摄像机的焦距为  $f$ ，假设行人的高度 170cm，则在图

像上的高度理论上应该是  $h \approx Hf/d \approx 170f/d$ 。但是图像上存在像素的畸变，矫正后的行人在图像上的投影高度应该满足如图 3-3(b)所示的曲线<sup>[32]</sup>。

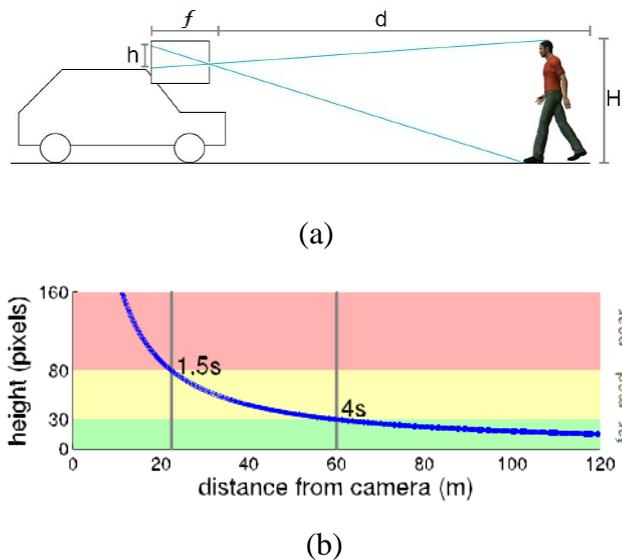


图 3-3 人体在图像上的显示比例

### 3.2.3 候选区域投影

为了在图像上找到激光定位出的候选行人区域，需要通过标定的方法解决激光测距仪和摄像机之间的数据对准问题。本文利用 2.2 节中提到的激光测距仪与摄像机联合标定方法，通过标定板同时被激光测距仪和摄像机捕获的方法来求得激光坐标系和摄像机坐标系之间的旋转和平移矩阵。

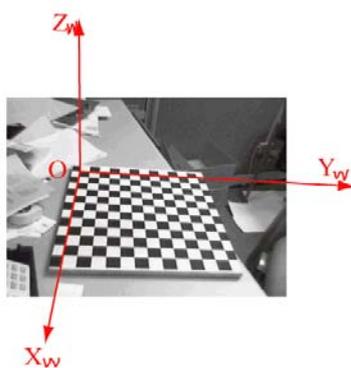


图 3-4 二维标定板

首先使用张正友提出的基于移动平面模板的标定方法<sup>[29]</sup>，用同一物体（二维棋盘）的多个图像来计算摄像机内部参数以及每个视场的旋转和平移。其中二维标定板平面位于世界坐标系中  $Z_w=0$  的平面上，标定二维棋盘上的每一个特征点，可得其在图像上的投影位置，而其相对于世界坐标系的位置在制作时

即可精确测定。由定义的空间已知点及其图像点坐标，可解线性方程得到投影矩阵  $\mathbf{H}$ 。得到投影矩阵  $\mathbf{H}$  后，可按照张的方法<sup>[29]</sup>求解摄像机内外参数。

接下来利用文献<sup>[30]</sup>中提出的激光测距仪与相机标定技术进行两种传感器的联合标定，目的是利用联合标定得到的配准参数，将激光空间得到的候选目标投影映射到与激光数据同步的图片上，获得图片上的行人候选区域。利用式 (3-22)，其中  $\mathbf{N}$  为标定板参数化后的 3 维向量，方向平行于标定板法向， $\mathbf{N}$  可以通过相应的摄像机外部参数矩阵 $[\mathbf{R}, \mathbf{t}]$ 求出； $p^f$  为激光坐标系中位于标定板上的某一点； $\Phi$  为旋转矩阵，表示两个传感器之间的相对方向； $\Delta$  为平移向量，表示两个传感器之间的相对位置。通过变换标定板的位置，可以得到一组约束条件

$$\mathbf{N} \cdot \phi^{-1}(p^f - \Delta) = \|\mathbf{N}\|^2. \quad (3-22)$$

求解方程可求得摄像机坐标系与激光坐标系之间的旋转和平移关系  $\Phi$  和  $\Delta$ ，由此即完成了激光测距仪与摄像机的联合标定。获得了  $\Phi$  和  $\Delta$ ，即可将激光空间的候选目标位置投影变换到图像空间，也就获得了图像上的行人候选区域。投影结果如图 3-5 所示。

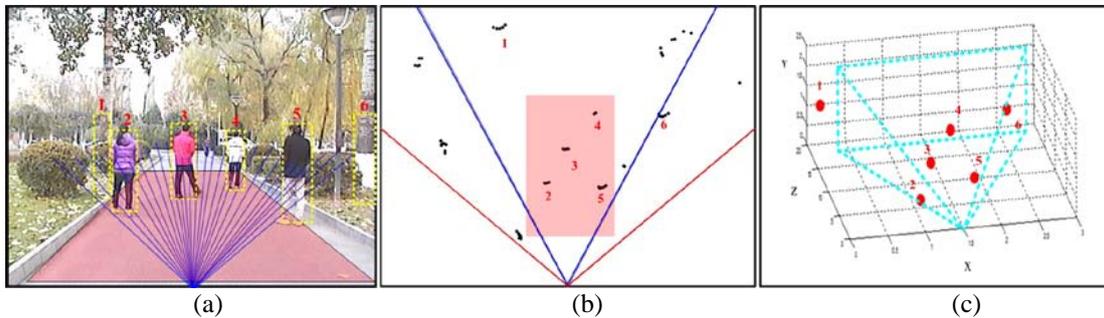


图 3-5 候选行人区域投影结果

从图中的投影结果，(a)中可以看到六个候选区域，但是只有中间四个候选区域为行人；(b)中描述的是六个候选区域在激光坐标系下的位置，可以获得这些区域的位置，深度和宽度信息；(c)中描述的是在三维空间下六个目标候选区域的位置。

### 3.3 基于视觉信息的行人目标的精检测

在获得候选行人目标的粗分类后，需要利用在图像行人检测方法对这些候

选行人区域进行再次检测确认。如果不用视觉检测方法进行再确认，误检率很高，比如图 3-5 所示，有六个候选行人目标具有单轴结构的激光模式，但是一号和六号候选行人目标都是虚警目标，他们都不是行人，但在激光模式分类中具有相同的结构，所以用视觉行人检测方法在候选行人区域内进行再检测确认这一步骤是必要的。此外，利用激光信息进行粗分类的为跟踪初始化带来巨大好处。比如，经过粗分类后，仅需要在上述六个候选区域进行行人检测，而不需要对整幅图像进行尺度放缩的全搜索检测。

在检测之前，需要提取目标特征并训练分类模型。本文所采用的方法为经典的 HOG+SVM 人体目标检测方法。在视觉检测的部分需要在检测前针对不同目标训练不同的分类器。

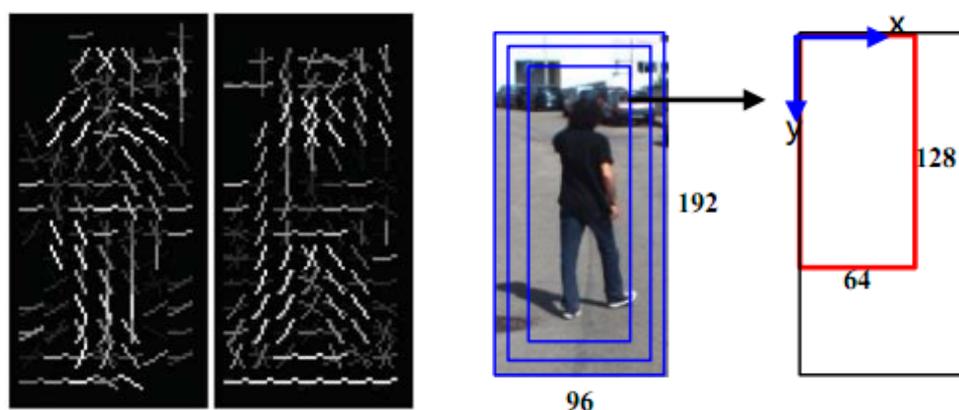


图 3-6 HOG 特征和人体检测示意图

**HOG 特征提取：**将样本图像（人体样本为  $64 \times 128$  像素的训练样本）按  $8 \times 8$  个像素分割为若干个单元（cell）；将相邻的 4 个 cell（田字结构）划分为一个 block（利用每个 cell 进行窗口滑动而生成）。将  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  的梯度方向平均划分，其中规定每 20 度分成一个 bin，即 9 个方向 bins。对每个 cell，将所有像素的所有梯度方向投影以建立各自的梯度方向直方图。将每个 block 中含有的 4 个 cell 的梯度方向直方图连接起来，形成 36 维的向量。再将所有的 block 内的 36 维向量归一化后在向量中连起来，由此得到每个训练样本的 HOG 特征向量。图 3-6 为 HOG 特征示意图，反映了人体的轮廓。HOG 特征提取之后，训练一个 SVM 分类模型用于目标检测。对于候选行人区域，通过逐像素扫描的图像块来判断该图像块内是否存在行人目标，如图 3-7。

**SVM 分类器：**SVM 是一种基于结构风险最小化准则的学习方法，其核心

思想是在样本空间或特征空间建立一个超平面作为决策曲面，构造最优平面，使得正例样本和反例样本之间的分隔边界最大化，从而达到最大的泛化能力。SVM 的主要优点是对目标模式变化的鲁棒性，已经被证明是一种更系统的学习线性和非线性的分类边界的方法<sup>[33]</sup>，并且在目标检测等应用中表现出了很好的性能。

采用激光模式分类确定行人候选区域并映射在图像中进行检测验证的方法，速度比单纯用 HOG+SVM 进行全局检索有了明显的提高，此过程中 95% 以上的区域被激光排除，在行人候选区域内只用小尺度下可以完成检测，根据人体宽高比例模型的投影区进行检测，这样可以大大减少循环的运算量。检测结果如图 3-7 所示，其中(a)为原始图像，(b)为经过聚类分析后的激光片段，(c)为利用激光特征选取的候选目标在图像层上的投影，(d)为利用图像检测方法进行确认行人检测结果，可见图中在激光层聚类结果中与行人类似的车体片段在图像检测确认中被排除。

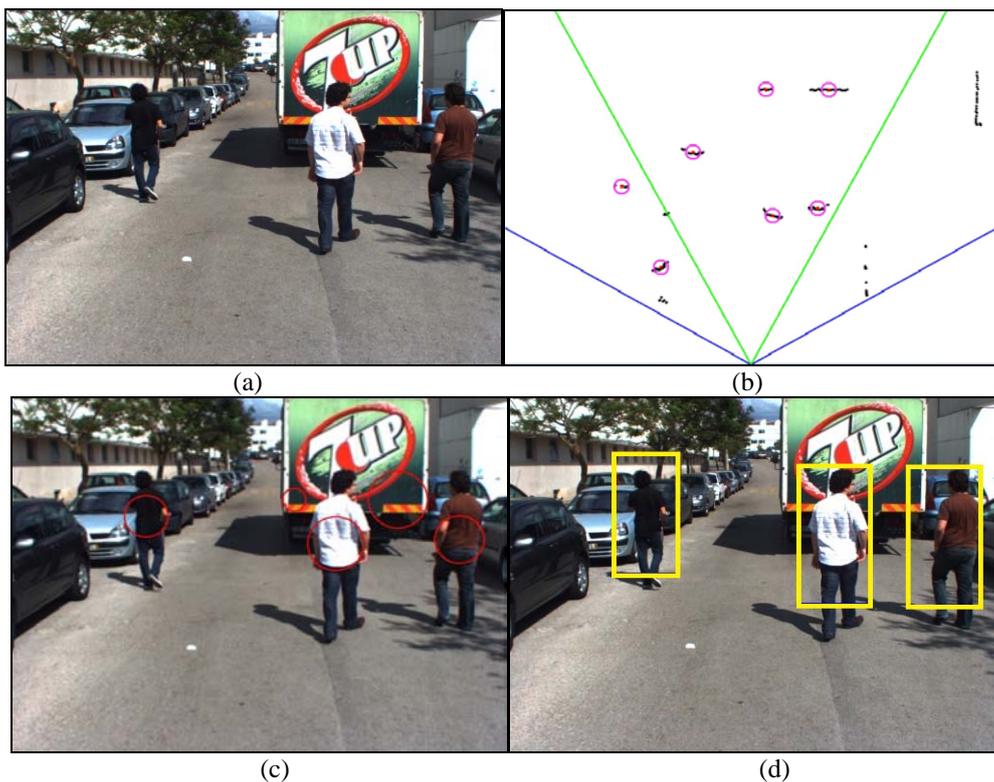


图 3-7 检测结果示意图

### 3.4 本章小结

本章介绍了如何利用激光模式和视觉特征对车辆前方的目标进行分类和确

认。在实时车辆驾驶条件下，对路面道路的精确判断建立在对目标障碍物准确快速的分类识别基础上。本章利用激光分类模式首先完成对危险区域内障碍物的粗分类，进而利用多传感器联合标定配准后的坐标变换矩阵，将候选行人区域投影到图像层，最后利用 HOG+SVM 分类方法在投影候选区内对行人确认性检测。该方法最大程度上缩小了视觉检测时对整个图像区域进行多尺度检测搜索所用的时间，对实时安全辅助驾驶具有决定性意义。



## 第四章 基于激光和视觉信息融合的多目标鲁棒跟踪

数据关联的理论建立于上世纪六七十年代，期间研究人员提出了众多数据关联算法。依据关联逻辑的不同，这些算法大致可以分为三类：一是面向目标的方法，假设跟踪过程中的目标数目是固定的，测量来自于已知的目标或杂波，典型算法包括概率数据关联算法<sup>[34]</sup>（PDA）和联合概率数据关联算法<sup>[35]</sup>（JPDA）；二是面向测量的方法，假设测量来自已知的目标、新目标或杂波，典型算法为多假设跟踪算法<sup>[36]</sup>（MHT）；三是面向轨迹的方法，该类方法假设轨迹未被检测、已经终结、与测量相关。就所采用的手段而言，数据关联也可分为两类，一类是基于贝叶斯估计的方法，包括 PDA、JPDA、MHT 等；另一类是非贝叶斯估计方法，如最近邻算法<sup>[37]</sup>（NN）、广义多维分配算法<sup>[38]</sup>（S-D Assignment）及 Viterbi 算法<sup>[39]</sup>。此外也有不少学者研究利用期望最大化算法<sup>[40,41,42]</sup>（EM）、神经网络<sup>[43]</sup>（ANN）和模糊逻辑<sup>[44]</sup>等方法来解决数据关联问题。

### 4.1 基于数据关联的多目标跟踪介绍

在多目标跟踪问题中，目标的数目是未知的并且随时间无规律变化，例如在驾驶过程中，车辆前方出现的行人的数目和时间都是不确定的。另外，无论多么先进的行人检测技术，都有虚警和误报的产生，传感器每时每刻都接收到随机数目的目标测量。在任何时刻，不能确定哪个测量应该用来更新哪个目标的状态，因此不能直接运用标准的贝叶斯滤波技术。此外，由于目标的不确定性，传统的单目标跟踪技术也不适用于多目标跟踪技术中。下面简单介绍几种传统的多目标跟踪方法<sup>[45]</sup>。

#### ■ 全局最近邻

全局最近邻（GNN）滤波器是将 NN 滤波器推广到了多目标跟踪情形中。在已知每个目标上一时刻状态估计的均值和协方差的情况下，GNN 滤波器首先运用卡尔曼滤波的预测部分得到每个目标测量的预测以及相对应的协方差。在数据关联部分，通过最小或最大化整体损失函数得到目标和测量唯一的联合

关联，损失函数可以是距离也可是似然概率和，限制条件是每个测量量最多与一个目标轨迹进行关联。在更新部分，该方法简单地假设得到的联合关联是正确的，然后直接运用卡尔曼滤波，用相关联的测量去更新每个目标的状态。自然，这种方法会遇到和 NN 滤波器一样的问题和限制。

#### ■ 联合概率数据关联

联合概率数据关联(JPDA)滤波器是 PDA 滤波器在多目标跟踪中的推广，它能够处理多个目标数目已知的情况。在该方法中，除了互联概率的计算以外，单个目标状态的迭代传播过程与 PDA 滤波器基本一致。JPDA 滤波器使用互联事件和互联概率来避免将多个测量分配给多个目标时引起冲突。然而，互联概率的计算复杂度伴随目标数目和测量数目呈指数型增长。基本的 JPDA 滤波器在计算上的不可行性导致了許多近似算法的产生，比如次优选择策略<sup>[46,47]</sup>以及基于马尔科夫链蒙特卡洛(MCMC)的策略<sup>[48,49]</sup>。此外，由于 JPDA 滤波器只能处理目标数目已知并且固定的情况，所以对于目标数目未知并且不断变化的场景，出现了许多新颖、实用的近似算法，比如说联合整体概率数据关联(JIPDA)滤波器<sup>[50]</sup>。此外，Vermaak<sup>[51]</sup>采用序贯蒙特卡洛(SMC)方法，他将 JPDA 滤波器推广到了非线性、非高斯的数据模型中。

#### ■ 多假设跟踪

多假设跟踪(MHT)滤波器<sup>[52,53]</sup>是一种基于非实时数据关联策略的多目标跟踪方法，MHT 滤波器通过寻找先前所有时刻的目标与观测的所有可能关联组合，构成目标所有可能的轨迹，来降低关联的不确定性。所有的观测与真实目标或者虚假目标的所有关联集合称作假设，在任意时刻，MHT 滤波器会保留假设中后验概率较大的子集部分。当新的观测数据集到达后，会在原假设的基础上产生新的假设，同时应用贝叶斯准则更新相应的后验概率值。需要注意的是，产生新假设的时候，有三种选择：一个测量可以与已存在的轨迹进行关联，也可以被当作噪声，还可以初始化一条新的轨迹。从这个方面来讲，MHT 滤波器自身具有初始化轨迹和终止轨迹的功能，所以它能够处理目标数目未知且随时间变化的情况。因此，MHT 滤波器的基本思想是向后传递后验概率较高的多个假设，在每个时刻寻找后验概率最高的假设作为当前的最优关联。得到最优的关联之后，运用卡尔曼滤波来更新各个目标的状态。R.Singer<sup>[54]</sup>第一次

提出传递关联假设的基本思想，主要是处理单目标跟踪的情况，Reid<sup>[55]</sup>将这种思想完整且系统地运用到了多目标跟踪中。MHT 滤波器最大的缺陷是假设的数目会随时间呈指数型增长，此点与 JPDA 相似。在实际应用中，MHT 滤波器通常需要设定关联门限或者使用各种启发式规则来限制假设数目的过快增长，从而降低计算复杂度。

## 4.2 基于激光与视觉信息融合的多目标跟踪

第三章完成了多传感器对行人的精确快速检测后，对具有标号的多目标进行跟踪是系统达到实时性的另一个重要课题。本节将从模型的构建，求解，更新，以及对复杂遮挡模型的验证等多方面进行详细阐述。

### 4.2.1 跟踪模型的构建

本方法提出的模型从最大后验概率 MAP (Maximum-A-Posteriori) 问题入手。假设，场景中各个目标的观测向量为： $s = (T_1, T_2, \dots, T_n)$ ，每个目标的观测向量表示为： $T_i = \{p_i, o_i, v_i, f_i\}$ 。其中  $p_i$  表示目标的位置坐标； $o_i$  表示目标的运动方向； $v_i$  表示目标的运动速度； $f_i$  表示目标的表象特征(如颜色直方图，方向梯度等)。对于一段独立的跟踪轨迹片段估计可以用一连串目标的观测向量来表示，例如  $s = (T_1, T_2, \dots, T_n)$ 。而对于同一个目标的跟踪轨迹估计可以使用一连串的轨迹片段来表示，例如  $S = \{s_k\}$ 。

对于给定的观测向量  $T$ ，数据关联的目标就是最大化后验概率  $S$ ，即：

$$S^* = \arg \max_S P(S | T) = \arg \max_S P(T | S)P(S). \quad (4-1)$$

假设行人目标的运动是独立的，则可以将式 4-1 分解成

$$S^* = \arg \max_S \prod_i P(T_i | S) \prod_{s_k \in S} P(S_k). \quad (4-2)$$

然后进行  $S = -\log S^*$  操作，则式 4-2 可以重新写成

$$\begin{aligned}
 S &= \arg \min_S \sum_{s_k \in S} -\log P(S_k) + \sum_i -\log P(T_i | S) \\
 &= \arg \min_S \sum_{s_k \in S} -\log P(S_k) - \sum_i (1-t_i) \log P(T_i = 0 | S) + t_i \log P(T_i = 1 | S) \\
 &= \arg \min_S \sum_{s_k \in S} -\log P(S_k) - \sum_i t_i \log \frac{P(T_i = 1 | S)}{P(T_i = 0 | S)} \\
 &= \arg \min_S \sum_i \left( -\log P_{sim}(T_{i_{k-1}} | T_{i_k}) \right) - \sum_i t_i \log \frac{P(T_i = 1 | S)}{P(T_i = 0 | S)}.
 \end{aligned} \tag{4-3}$$

式 4-3 中的条件概率  $P(T_i | S)$  可以用一个二项分布来描述：

$$P(T_i | S) = \begin{cases} p_i & \exists s_k \in S, T_i \in s_k \\ 1 - p_i & otherwise \end{cases}. \tag{4-4}$$

对于  $p_i$ ，本文将它定义成深度因子，将在 4.2.4 节中做详细介绍。式 4-4 中，先验概率可以用相似函数来表示：

$$P_{sim}(T_{i_{k-1}} | T_{i_k}) = P_{app}(T_{i_{k-1}} | T_{i_k}) P_{pos\&vel}(T_{i_{k-1}} | T_{i_k}) P_{ori}(T_{i_{k-1}} | T_{i_k}). \tag{4-5}$$

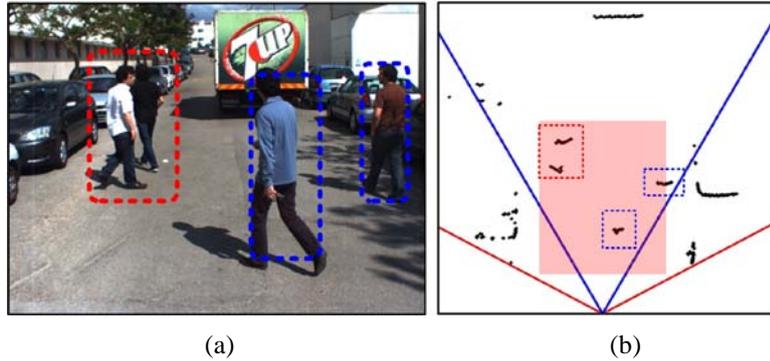


图 4-1 集群点与单一点示意图

在式 4-5 中，条件概率  $P(T|S)$  是独立分布的，因为它主要描述的是每个行人的运动特点。但是，与其他方法不同的是，本文用一个新的观测空间分类来描述行人位置上的“亲近关系”。根据被观测目标的关系，本文将观测空间  $S$  划分为两部分  $S = \{S^\alpha + S^\beta\}$ ，如图 4-1 所示。红色矩形框中的行人位置更加靠近，轨迹也更容易混淆；而蓝色矩形框中行人相对独立，轨迹更容易区分。基于此，本文将式 4-1 重新写成

$$S^* = \arg \max_S P(S^{\alpha+\beta}) P(T | S^{\alpha+\beta}), \tag{4-6}$$

进一步展开

$$\begin{aligned}
 S^* &= \arg \max_S P(S^{\alpha+\beta})P(T | S^{\alpha+\beta}) \\
 &= \arg \max_S \left( P(S^\alpha) + P(S^\beta) \right) \left( P(T | S^\beta) + P(T | S^\alpha) \right) \\
 &= \arg \max_S \prod P(S^\alpha) \prod P(T | S^\alpha) + \prod P(S^\beta) \prod P(T | S^\beta).
 \end{aligned} \tag{4-7}$$

在式 4-7 中, 等式右边应该有四项, 但是  $P(S^\alpha)P(T | S^\beta)$  和  $P(S^\beta)P(T | S^\alpha)$  的值很小, 在等式中可以近似地省略。为了满足一个轨迹片段只属于一个行人的约束, 引入指示变量表示轨迹片段属于某一个行人的轨迹点。

$$\begin{aligned}
 t_{j,i} &= \begin{cases} 1 & \text{if } T_j \text{ is right after } T_i, \\ 0 & \text{otherwise} \end{cases}, \\
 t_i &= \begin{cases} 1 & \text{if } T_i \in S_k \\ 0 & \text{otherwise} \end{cases}.
 \end{aligned} \tag{4-8}$$

对于观测空间  $S$  来说, 轨迹是非重叠的, 则指示变量满足  $t_i = \sum_j t_{j,i} \leq 1$ 。

此时将式 4-3 带入式 4-8 中, 则目标方程可以重写为:

$$\begin{aligned}
 S &= \arg \min_{S^\alpha} \sum_{i,j} C_{j,i} t_{j,i} + \sum_i C_i t_i + \arg \min_{S^\beta} \sum_{i,j} C_{j,i} t_{j,i} + \sum_i C_i t_i \\
 &= \arg \min_{S^\alpha} \sum_{i,j} C_{j,i} t_{j,i} + \sum_i C_i t_i + \arg \min_{S^\beta} \sum_{i,j} C_{j,i} t_{j,i}
 \end{aligned} \tag{4-9}$$

其中引入的参数可以表示为:

$$C_{j,i} = -\log P_{sim}(T_{j-1} | T_i), \tag{4-9}$$

$$C_i = -\log P(T_i | S) - \log \frac{p_i}{1-p_i} = \log \frac{1-p_i}{p_i}. \tag{4-10}$$

从新的目标方程式 4-9 可以得到两类行人目标, 一类是在  $S^\alpha$  空间内和周围行人关系较为密切的行人, 对应于目标方程中的前半部分, 称作集群点; 另一类是处在  $S^\beta$  空间单独的行人对应于目标方程中的后半部分, 称作单一点。

#### 4.2.2 跟踪模型的求解

数据关联问题可以转化成图算法进行求解。二分图又称作二部图, 它是这样定义的: 设  $G=(V,E)$  是一个无向图。如顶点集  $V$  可分割为两个互不相交的子集, 并且图中每条边依附的两个顶点都分属两个不同的子集。则称图  $G$  为二分图。给定一个二分图  $G$ , 在  $G$  的一个子图  $M$  中,  $M$  的边集中的任意两条边都

不依附于同一个顶点，则称  $M$  是一个匹配。二分图的最佳匹配，它是指在二分图中，边带权值，使得匹配所在权值和最大的就为最佳匹配。

本文对多目标进行跟踪时所用的数据关联方法采用了二部图的最佳匹配方法，如图 4-2。本文把每一帧数据中多行人抽象成一个点，而相邻帧之间多行人轨迹的关联可以看成是在二部图中寻找最佳匹配问题，相邻帧之间前一帧每个行人到下一帧的所有行人存在转移代价，构成了二部图中间的边的权值。本文中的多目标的关联问题，则是在二部图最佳匹配的基础上，根据目标方程式 4-9 对二部图进行了改变，在原来的二部图基础上加了一条边，而这条边用于多行人的在线学习的自我更新权值。目标方程式 4-9 中  $C_i$ 、 $C_{j,i}$  分别构成本帧与两帧之间的转移代价， $t_i$ 、 $t_{j,i}$  是 0-1 指示变量。

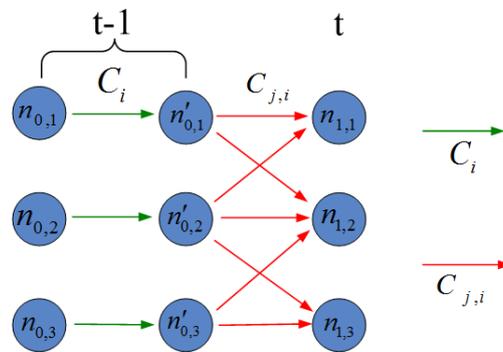


图 4-2 改进二部图的匹配

对于目标方程式 4-9 的求解最常用到的是 KM 算法，该算法的时间复杂度是多项式时间；此外，子模块问题用图割法也容易达到全局最优。但是，根据式 4-8 的约束，本文的目标方程不能算成是子模块问题，因此想找到一个多项式时间复杂的算法，并且达到全局最优很难。所以本文采用改进迭代算法进行搜索，使问题达到局部的最优，而时间复杂度控制在多项式时间内。

对于搜索的基准本文使用贪婪算法在没有出现复杂遮挡问题下，找到多帧之间的最小费用路径作为每次搜索的基准。而费用最小的边在本文中多是以很低的转移代价或者是负代价形式出现。对于已经选择的二部图中的节点，给这些节点进行顺序编号  $l = \{0,1,\dots,n\}$ ，接着在临近点周围（也就是具有“亲近位置”的节点范围内），进行标号互换，重新计算代价，如果代价比原来的基准小，保留这次的标号变换，完成循环，找到具有最小代价的匹配。具体的算法细节

在表 4-1 中有详细介绍。

由于实际的目标数目是有限的，并且本文在前面章节已经将行人按照位置关系进行了集群点和单一点的划分，划分后按集群的循环次数要比穷举法小很多。在行人目标比较少且比较分散时，往往初始的贪心算法求出的二部图匹配解就是最终的全局最优解。只有在遮挡情况比较严重时，行人密集出现的极端情况，算法的循环数目达到最高。所以整个算法的平均时间复杂度是在实时跟踪系统可以接受的范围以内。

表 4-1 跟踪模型求解算法

求解算法	
输入：	改进二部图，每条边的权值 $\{C_i\}$ ；观测向量 $S$
初始化：	利用贪心算法找到一组相对小的初始匹配解 $L'$ 。并计算总代价 $S^*$
1	For $i < n$ do
2	设置最大代价是 $S' = +\infty$
3	For $j=i, \dots, n$ 并且 $C_{i,j} \in G$ do
4	交换 $N'_{t-1,i}$ 和 $N_{t,j}$ 的标号，并计算新的代价 $S''$
5	if $S'' < S'$ , $S' = S''$
6	end
7	If $S' < S^*$ , $S^* = S'$ , 保留此次匹配解 $L$
8	end
输出：	新的匹配解 $L$

### 4.2.3 跟踪模型的在线更新

本节将介绍本文提出的多目标跟踪模型如何进行融合激光测距仪和图像信息进行信息融合并在线更新的方法。在线更新后的转移代价转化成为上节所提到的改进二部图的权值边。

为了充分利用激光测距仪返回的深度信息，本文不光利用了原始的距离信息以及坐标变换后的深度信息，还提出了具有方向区分的方向模型，用来区分行人在行进中的方向信息。如图 4-3 所示，本文将实际运动方向划分为 0-8 九

个方向区（0 代表静止）。而针对这九种方向，他们之间的关联度是不一样的，相连方向的关联度要大于处在对角方向的关联度，如图中红色虚线框内的“8”方向，和它关系紧密的是“1”，“7”，和“0”。“0”与 1-8 中的任何方向都有比较紧密的关系。

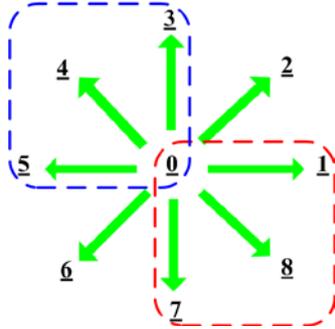


图 4-3 方向模型

利用上述方向关系，在进行目标运动信息在线更新时，有了更多关联标准，比如图 4-4 中，在  $t, t+1$  连续两个时刻，根据原来的前面的运动方向信息，将不会混淆两个轨迹。如  $T_a$  在  $t-1$  时刻的运动方向是“8”，而  $T_b$  在  $t-1$  时刻的运动方向是“2”。当两个目标出现交叉换位时，即使他们的颜色轮廓形状信息相同，也可以通过他们具有方向的运动信息加以区分。

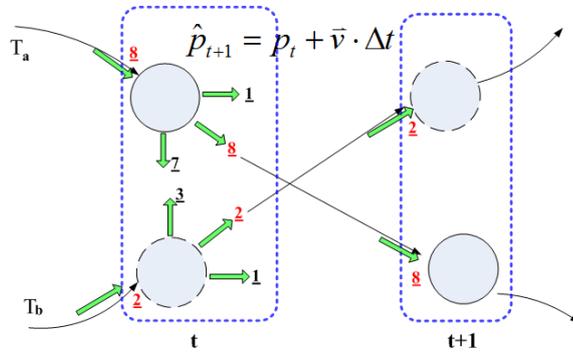


图 4-4 运动模型

在第 4.2.1 节，跟踪模型的构建中，方程式 4-5 中的条件概率用到了相似函数

$$P_{sim}(T_{i_{k-1}}^L | T_{i_k}^L) = P_{app}(T_{i_{k-1}} | T_{i_k}) P_{pos\&vel}(T_{i_{k-1}} | T_{i_k}) P_{ori}(T_{i_{k-1}} | T_{i_k}). \quad (4-11)$$

对于目标的运动方向，运动信息和视觉特征，本文用高斯分布模型对他们进行统一的建模，将他们的乘积作为改进二部图的边的权值进行在线更新。

$$P_{ori}(T_m^L | T_n^L) = G(|ND_m - ND_n|, \sum ND), \quad (4-12)$$

$$P_{mot}(T_m^L | T_n^L) = G(|P_n - P_m|, \sum P) = G((P_n + \bar{v}_n \Delta t - P_m), \sum P), \quad (4-13)$$

$$P_{app}(T_m^L | T_n^L) = G(sim(h_m, h_n), \delta_c). \quad (4-14)$$

#### 4.2.4 跟踪模型在遮挡问题中的求解

多目标跟踪模型的好坏，很大一部分程度上取决于目标被遮挡时，模型是否依然能够不丢失目标。本节主要内容就是利用本文的跟踪模型，解决多目标长时间复杂遮挡问题。

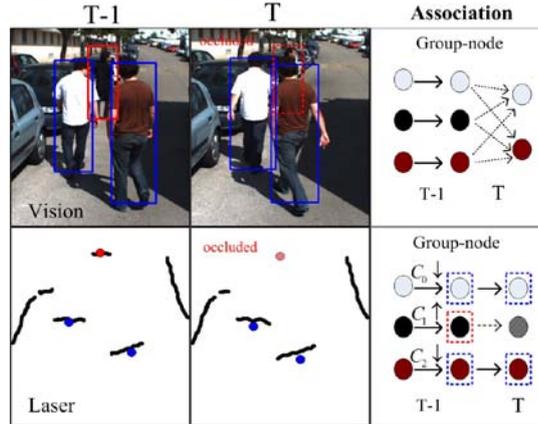


图 4-5 遮挡中的目标关联

在方程 4-4 中，条件概率  $P(T|S)$  满足二项分布，在方程中  $p_i$  将作为区分目标在深度顺序上的参数，并将它作为改进二部图的转移代价，进行搜索求解。本文将  $p_i$  定义成 sigmoid 函数：

$$p_i = [1 + \exp(\bar{z} - z_i)]^{-1}. \quad (4-15)$$

这个方程建立在深度方向，也就是距离观测点的远近，本文坐标系中的  $Z$  轴方向。如图 4-6 (a) 中所示，三个行人组成集群点，他们之间很容易出现遮挡情况。此时本文在深度方向  $Z$  轴上完成对深度参数的确定，首先确定集群点的平均深度  $\bar{z}$ （如图中所示的绿线），进而求出每个行人所处的深度到平均深度  $\bar{z}$  的相对深度  $\bar{z} - z_i$ ，最后带入到方程 4-4 中求出  $p_i$  深度参数。 $p_i$  可能为正数负数，也可能为零，而且根据 sigmoid 函数特点， $p_i$  在的取值范围在 -1 到 1 之间。此时本文再将方程式 4-15 带入到方程式 4-10 可以得到一个可能为正数或负数，也可能为零的转移代价  $C_i$ 。比如在图 4-6 (a) 中，红色矩形内的行人处在容易被遮挡的位置，按照上面求取深度参数的方法，他会得到一个正的，且比价大的代价  $C_i$ ；而相对应的蓝色矩形框内的行人处在靠近观测者的位置，他们就能得

到一个负的转移代价  $C_i$ 。

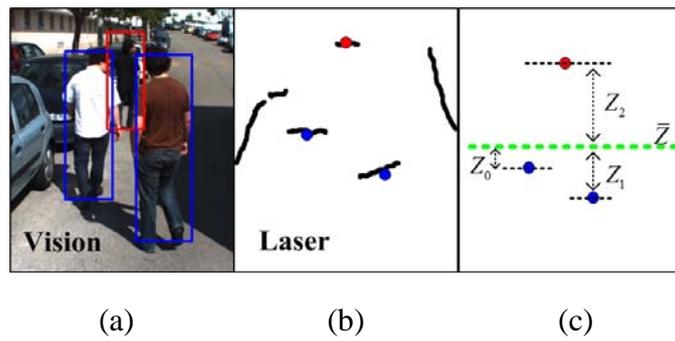


图 4-6 遮挡模型求解示意图

对于集群点的情况本文可以按照上述的求取与平均深度的深度差来计算深度参数，而对于单独点，也就是单一人时，与平均深度的深度差为零，此时的  $P_i$  等于 1，而转移代价通过带入方程 4-10 可以求出为 0。所以本为的目标方程刚好分为两部分，一部分是包含相似函数的转移代价  $C_{j,i}$  和包含深度参数  $C_i$  的集群点；第二部分为只包含相似函数的转移代价  $C_{j,i}$  的单独点。

在搜索过程中，拥有负值转移代价的边优先被选择，而被赋予正值且转移代价比较高的边处在末位被选择，而这两者正好对应与遮挡者和被遮挡者的位置，符合客观规律。

### 4.3 多目标跟踪模型的实验验证

在完成了平台搭建，多传感器的参数配准以及跟踪模型的推导求解之后，本节利用实验对本文提出的算法进行验证。本节包含对数据集的介绍，跟踪结果评价标准，以及和多种算法比较后的跟踪结果分析。

#### 4.3.1 实验数据集介绍

表 4-2 数据集介绍

数据集	图像分辨率	行人数目	激光分辨率
SYNC	1024 × 768	2-4	100°, 0.25°
SDL-1	640 × 480	4	100°, 0.25°
SDL-2	640 × 480	5-10	100°, 0.25°

本文在实验验证阶段采用三段数据集,如表 4-2 所示,ISR-UC-implidar-sync 数据集<sup>[56]</sup>,和 SDL 数据集。其中 SDL 数据集为本文第二章中介绍的实验搭建的平台在真实交通场景中采集的驾驶场景激光深度信息和视觉图像数据,包括两组数据 SDL-1,2。三个数据集的激光测距仪精度均为  $0.25^\circ$ ,水平扫描范围  $100^\circ$  (.txt 格式,每个文件中包括 401 个十进制数据,单位 mm)。其中 SDL-1,2 的图像分辨率均为  $640 \times 480$ , SYNC 数据集的图像分辨率为  $1024 \times 768$ 。在验证算法鲁棒性方面,实验所用的数据集中行人的数目也做了区分,从 2-4 个,到 10 个左右的行人数目。

### 4.3.2 实验的评测标准

跟踪结果的评测标准有很多,在单目标跟踪的评测标准中,经常用到跟踪正确结果帧数与总帧数的比值等类似的评价指标。但是对于多目标跟踪的评测标准,单一评测标准是无法对算法做出全面详细的评价。既要考虑多行人中间一个体的跟踪准确率还要考虑算法对所有行人的跟踪性能。

本文采用如表 4-3 中所描述的评测标准。前三个指标(GTP, LF, PF)描述的是单个行人在每一帧中的跟踪结果,分为位置和标号。中间的四个指标(GTI, MT, ML, PL)描述的是行人在整个视频中轨迹的跟踪结果。最后一项 Sp 是算法处理不同数据集的平均速度。对于括号中含  $\uparrow$  的指标,得分越高性能越好;括号中含  $\downarrow$  的指标,得分越低性能越好。

表 4-3 评测标准

指标	定义
GTP	真实数据中行人的个数。
LF( $\downarrow$ )	标号错误,与真实数据不同。
PF( $\downarrow$ )	位置错误,与真实数据不同。
GTI	真实数据中行人的ID标号。
MT( $\uparrow$ )	大部分轨迹跟踪正确,轨迹片段80%以上与真实数据相同。
ML( $\downarrow$ )	大部分轨迹跟踪错误,轨迹片段20%以下与真实数据相同。
PL( $\downarrow$ )	部分跟踪错误,跟踪正确的轨迹占到20-80%之间。
Sp( $\uparrow$ )	处理速度,帧/每秒。

### 4.3.3 实验结果及分析

本文在实验中采用了另外三种方法作为对比试验，如表 4-4 所示。第一种方法“Laser (NN)”只应用激光的深度信息，利用“最近邻 (NN)”方法对行人的轨迹点进行关联。第二种方法“Laser+Fea”利用激光测距仪的深度信息和图像中新人特征 (HOG+颜色) 在本文所提出的跟踪模型框架进行实验。第三种方法“Laser+Mot”是利用激光测距仪和目标的运动信息在本文所提出的跟踪模型进行实验。最后一种方法“Laser+ANM”是利用本文的在线学习的跟踪模型。四种方法在三段数据集上的表现如表 4-4 所示。

表 4-4 对比实验结果

数据集	方法	GTP	PF(↓)	LF(↓)	GTI	MT(↑)	ML(↓)	PL(↓)	Sp(fps)
SYNC	Laser (NN)	2360	612	566	10	4	2	4	<b>60-80</b>
	Laser+Fea	2360	229	312	10	7	1	2	25-30
	Laser+Mot	2360	268	347	10	6	2	2	25-35
	Laser+ANM	2360	<b>0</b>	<b>0</b>	10	<b>10</b>	<b>0</b>	<b>0</b>	20-25
SDL-1	Laser (NN)	384	189	189	4	1	1	2	<b>60-80</b>
	Laser+Fea	384	68	68	4	2	1	1	25-30
	Laser+Mot	384	72	72	4	2	1	1	25-35
	Laser+ANM	384	<b>6</b>	<b>6</b>	4	<b>4</b>	<b>0</b>	<b>0</b>	20-25
SDL-2	Laser (NN)	1954	512	495	31	13	8	10	<b>55-70</b>
	Laser+Fea	1954	274	294	31	18	6	7	25-30
	Laser+Mot	1954	240	286	31	17	7	7	25-30
	Laser+ANM	1954	<b>43</b>	<b>46</b>	31	<b>25</b>	<b>3</b>	<b>3</b>	18-25

从表中可以看出，在三段数据集上，利用本文提出的方法在除了 Sp 以外的其他所有指标中都做到了最好。在速度方面，虽然没有达到前面三种方法相同的速度，但是仍然能够满足实时跟踪的标准。

图 4-7, 4-8, 4-9 是实验方法在三个数据集上的跟踪结果，三个图中第一行均为激光测距仪在交通场景测量数据通过坐标变化后在直角坐标系下的二维成像。其中两条蓝线之间代表摄像机的视场范围，两条红线之间是激光测距仪的扫描范围。

SDL-1 数据集：如图 4-7 所示，视频中有四个行人，整个视频中四个行人出现了反复的遮挡关系，如果只利用单纯的视觉信息进行跟踪，穿红色衣服的行人与紫色衣服的行人由于在很长一段时间内是处于遮挡和被遮挡关系，容易被混淆，在本文提出的跟踪模型中，加入深度动态权重因子后，紫色衣服行人

的转移代价为负值，在进行模型求解中优先会被选择。而处在远处的红色行人和白色行人则具有比较大的转移代价，处在末位被匹配选择的位置。

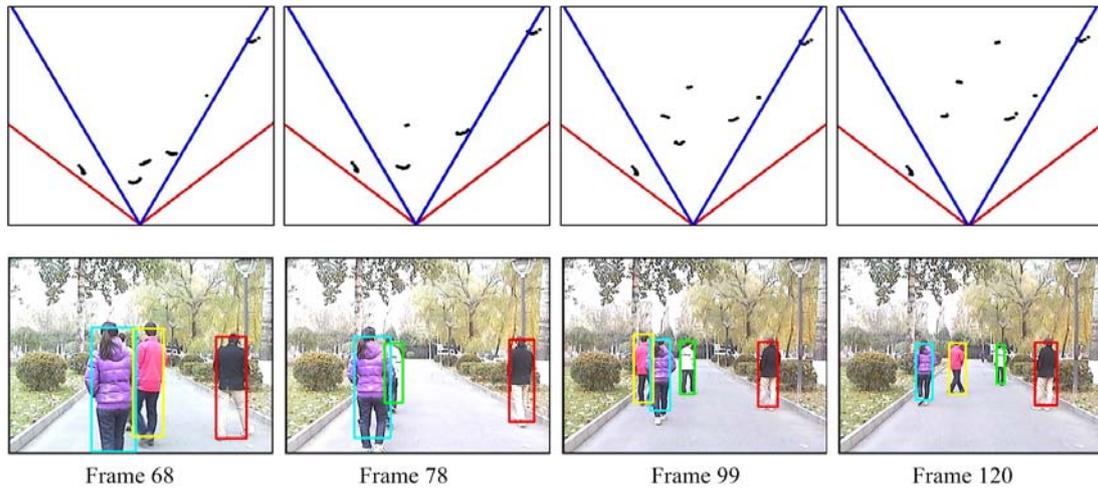


图 4-7 SDL-1 的跟踪结果

**SDL-2 数据集：**该数据集记录了长时间内，位于丁字路口出现在车辆前方的行人运动状况，实验对每个行人进行了标号，本文提出的跟踪模型，面对十个左右数目的且具有不同行进方向行人，进行了正确的标号保持。尤其处在车辆正前方的危险区域内，我们的检测跟踪方法没有失误，只是在危险区域外围的潜在危险区域内出现了行人的标号前后不一致的标号错误。

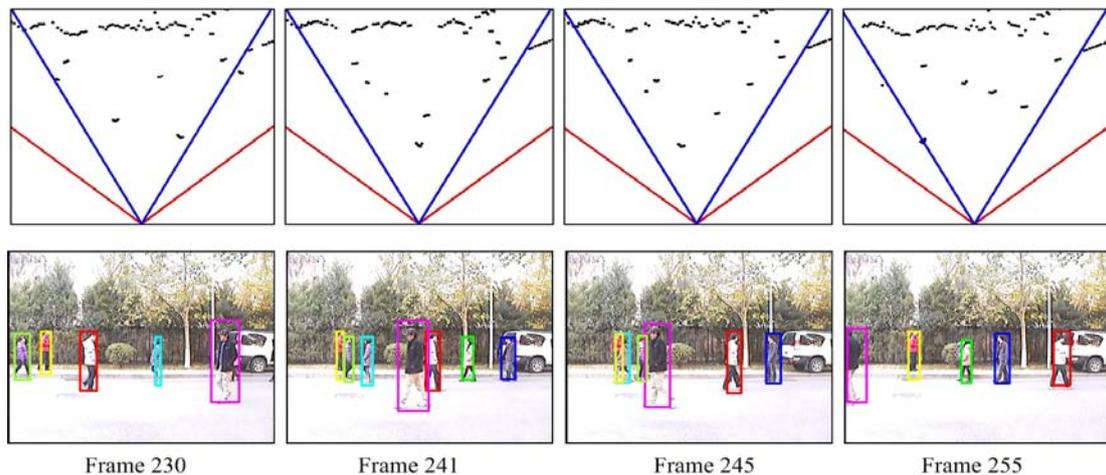


图 4-8 SDL-2 的跟踪结果

**SYNC 数据集：**在这个数据集中，四个行人一直处在周围背景比较复杂的道路上进行交错前进，(b)，(c)，(d)，(e)分别代表了试验中的四种方法，从(c)，

(d)的实验结果,可以看出利用运动信息和视觉信息单一为模型解法中二部图的权重边赋值,结果要明显好于只用最近邻方法进行关联的跟踪结果,而且加入到动态深度方向的权重因子后,遮挡情况出现后,仍然可以不丢失目标,完成全局的多目标跟踪。而在加入了多种颜色,运动,方向等信息后(e),并且在集群点内,模型对于不同深度的行人基于区分性的深度动态权重因子后,即使出现多次反复的近距离遮挡,算法仍然可以做到丢失目标。

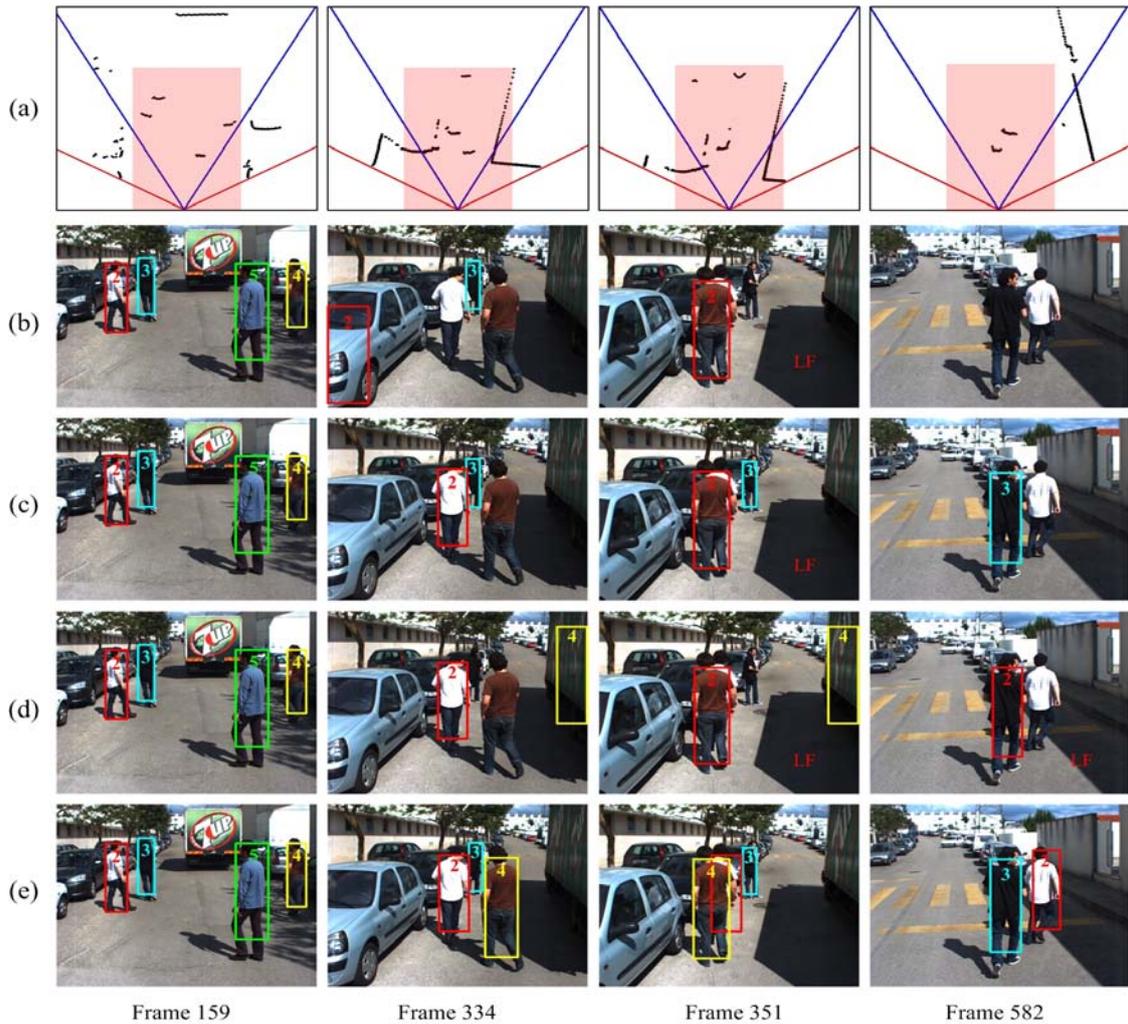


图 4-9 SYNC 的跟踪结果

#### 4.4 本章小结

本章首先提出了基于多元信息融合的复杂多目标跟踪模型,然后利用改进二部图对该模型进行了多项式复杂度的求解。并实现了模型如何应用多元信息

进行在线更新。此外本节介绍了该模型在复杂遮挡情况下，在深度方向利用激光距离信息进行深度权重的调节，进行遮挡问题的求解。最后本节利用三个数据集对提出的算法进行了验证，结果显示本节提出的多目标跟踪模型是鲁棒高效的。



## 总结与展望

本文通过融合激光测距仪返回的深度数据和从 CCD 摄像机采集的图像信息来实现对多行人目标进行快速检测和准确跟踪。研究内容对于利用安全辅助驾驶有着及其重要的意义，在物流系统研究领域有着重要实际应用价值。

本文首先论述了信息融合中行人跟踪系统的研究背景意义以及国内外研究现状，综述了多传感器信息融合中跟踪系统的研究方法，提出了本文主要的研究内容和理论基础；本文在第二部分，论述了包含激光测距仪和视觉传感器的实验平台，并用实验验证了本文采用的联合参数配准方法；在第三章，从激光测距仪和视觉检测多方向提出了行人检测三个步骤：激光模式分类，候选行人区域生成，候选区域视觉确认检测；在本文的最后一章，提出了基于数据关联的多目标跟踪模型，融合了激光深度信息和视觉特征的在线更新方法，并用改进二部图匹配的方法给出了模型的解法。最后提出了在深度方向上的动态权重调整方法，并在三个具有复杂遮挡情况的数据集上进行对比试验，实验结果及对比分析验证了本文所提出算法的准确性和实效性。

本文还需要从以下几个方面进行后续研究，完善基于信息融合的多目标跟踪系统：

- 1、通过大量的实验发现，在激光测距仪在扫描到玻璃等具有透明穿透性物质是，激光无法返回到机关激光测距仪，在处理这类假象“无穷远”数据时，将在后续的模式分类中基于更全面的分析。

- 2、在汽车行驶在路面情况不平整的道路时，激光测距仪的抖动比较大，数据的配准不会出现偏差，但是利用宽高比投影到图像层时，会出现和原有高度不相符的区域，造成检测中的漏检。



## 参考文献

- [1] 牛学军, 朱茵, 楼涛. 道路交通安全风险预控管理对策. 综合运输[J], 2008.(10).
- [2] 郑文贵, 李向云. 1996~2003 年全国交通事故伤害的时间序列分析. 中国卫生事业管理 [J]. 2006:105-107.
- [3] 徐翠, 基于计算机视觉的汽车安全辅助驾驶若干关键问题研究 [D]. 中国科学技术大学博士学位论文, 2009.
- [4] Fuerstenberg, K.C., Scholz, J.. Reliable Pedestrian Protection using Laserscanners [C]. In Proc. of the IEEE Intelligent Vehicles Symposium. 2005:142-146.
- [5] Ewald, A., Willhoeft, V. Laserscanners for obstacle detection in automotive applications [C]. In Proc. of the IEEE Intelligent Vehicles Symposium. 2000:682-687.
- [6] P. Bellutta, R. Manduch. Terrain perception for DEMO III [J]. Proceedings of IEEE Intelligent Vehicles Symposium, 2000(10): 326-331.
- [7] Jan C Becker, Andreas Simon. Sensor and navigation data fusion for an autonomous vehicle [J]. Proceedings of IEEE Intelligent Vehicles Symposium, 2000(10): 156-161.
- [8] 刘晓君. 丰田先进安全车 [J]. 世界汽车, 1998(2).
- [9] 谷歌测试无人驾驶汽,行程 14 万英里 [J/OL]. 新华网(2010-10-12). [http://news.xinhuanet.com/world/2010-10/12/c\\_12648406.htm](http://news.xinhuanet.com/world/2010-10/12/c_12648406.htm).
- [10] 徐有春, 王荣本,李兵. 世界智能车辆近况综述 [J]. 汽车工程, 2001, 23(5):289-295.
- [11] 高德芝, 段建民, 郑榜贵. 智能车辆环境感知传感器的应用现状[J]. 现代电子技术, 2008(19).
- [12] 刘大学. 用于越野自主导航车的激光雷达与视觉融合方法研究 [D]. 国防科技大学博士学位论文, 2009.
- [13] Y. Wang, Q. Zheng, J. P. Zhang. Real-time detection of small target in IR grey

- image based on mathematical morphology [J]. *Infrared & Laser Eng.*, 2003, 2(1):28-31.
- [14] Dave L, Hall and James Linas. Introduction to multi-sensor data fusion [J]. *Proceedings of the IEEE*, 85 (1), 6-23
- [15] 韩崇昭, 朱洪艳, 段战胜等, 多源信息融合 (第二版) [M], 清华大学出版社, 2009.
- [16] S. B. Colgrove, S. J. Davey. The Probabilistic Data Association Filter With Multiple Nonuniform Clutter Regions [C]. In *Proc. IEEE Radar Conf.*, 2000: 65~70.
- [17] E. Daeipour, Y. Bar-Shalom. An IMM Approach for Target Tracking with Glint Noise [J]. *IEEE Trans on AES*, 1995, 31(2): 706~715.
- [18] E. Daeipour, Y. Bar-Shalom. Layered IMM for Tracking Maneuvering Targets in the Presence of Glint [C]. In *Proc. Conf. Info. Sciences and Systems*, Princeton, NJ, March 1996.
- [19] E. Daeipour, Y. Bar-Shalom. IMM Tracking of Maneuvering Targets in the Presence of Glint [J]. *IEEE Trans on AES*, 1998, 34(3): 996~1003.
- [20] A. Farina. Target tracking with bearings-only measurements [J]. *Signal Processing*, 1999, 78 (1): 81~78.
- [21] Y. Bar-Shalom. Update with Out-Of-Sequence Measurements in Tracking: Exact Solution [J]. In *Proc. of SPIE Conf. on Signal and Data Processing of Small Targets*, Vol. 4048, 2000:541~556.
- [22] S. Avidan. Ensemble tracking [C]. In *CVPR*, 2005, volume 2, pages 494–501.
- [23] R. T. Collins, Y. Liu, and M. Leordeanu. Online selection of discriminative tracking features [J]. *PAMI*, 2005, 27(10):1631–1643.
- [24] C. Stauffer and E. Grimson, Learning Patterns of Activity Using Real-Time Tracking [J], *IEEE Trans. Pattern Analysis and Machine Intelligence*, Aug. 2000, vol. 22, no. 8, pp. 747-757.
- [25] Z. j. Han, J. b. Jiao, B. c. Zhang, Q. x. Ye, and J. z. Liu, Visual Object Tracking via Sample-Based Adaptive Sparse Representation (AdaSR) [J], *Pattern*

- Recognition, no.44, pp.2170 - 2183, 2011.
- [26] Sick\_LMS-200\_激光测距仪文档 [Z].
- [27] 刘志花. 面向立体视觉的匹配算法 [D]. 中国科学院研究生院博士学位论文, 2010.
- [28] G. Bradski, A. Kaehler. Learning OpenCV (于仕琪译) [M]. 北京: 清华大学出版社, 2009.
- [29] Z. y. Zhang. Flexible camera calibration by viewing a plane from unknown orientations [C]. Proceedings of the IEEE International Conference on Computer Vision, 1999(3):666-673.
- [30] Q. Zhang, R. Pless. Extrinsic calibration of a camera and laser rangefinder (improves camera calibration) [C]. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004(3):2306–2310.
- [31] H.Zhao, J.Sha, Y. Zhao, Detection and Tracking of Moving Objects at Intersections Using a Network of Laser Scanners [J], IEEE Trans. Intell. Transp. Syst., JUNE. 2012, vol. 13, no. 2, pp. 655-670.
- [32] P. Dollar, C. Wojek, B. Schiele, and P. Perona, Pedestrian Detection: An Evaluation of the State of the Art [J], IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, vol. 34, no. 4, pp. 743-761.
- [33] V. Vapnik. The nature of statistical learning theory [J]. Springer Verlag, 1995.
- [34] Y. Bar-Shalom, E. Tse. Tracking in a cluttered environment with probabilistic data association [J]. Automatic, 1975, 11(9): 451~460。
- [35] Y. Bar-Shalom, T. E. Fortmann. Tracking and Data Association [M]. Academic Press, 1988.
- [36] D. B. Reid. An Algorithm for Tracking Multiple Targets [J]. IEEE Trans on Auto. Control, 1979, AC-24:843~854.
- [37] H. Leung, Z.Hu and M. Blanchette. Evaluation of Multiple Radar Target Trackers in Stressful Environments [J]. IEEE Trans on AES,1999, 35(2): 663~673.
- [38] K. R. Pattipati, T. Kirubarajan and R. L. Popp. Survey of Assignment Techniques for Multitarget Tracking [C]. In Proc. Of The Workshop on

- Estimation, Tracking and Fusion: A Tribute to Yaakov Bar-Shalom, Monterey, CA, May,2001:51~69.
- [39] A. Gad, F. Majdi and M. Farooq. A Comparison of Data Association Techniques for Target Tracking in Clutter [C]. In Proc. of 5<sup>th</sup> International Conf. On Information Fusion, Annapolis, MD, 2002: 1126~1133.
- [40] H. Gauvrit, J.P Le Cadre and C. Jauffret. A Formulation of Multitarget Tracking as an Incomplete Data Problem [J]. IEEE Trans on AES, 1997, 33(4): 1242~1257.
- [41] D. Avizatour. A Maximum Likelihood Approach to Data Association [J]. IEEE Trans on AES, 1992,28(2): 560~565.
- [42] A. Logothetis, V. Krishnamurthy and J. Holst. A Bayesian EM Algorithm for Optimal Tracking of a Maneuvering Target in Clutter [J]. Signal Processing, 2002, 82 (2002): 473~490.
- [43] 敬忠良.神经网络跟踪理论及其应用 [M]. 北京: 国防工业出版社,1997.
- [44] 刘源. 基于模糊信息处理的数据融合方法研究 [D]. 博士论文, 西安电子科技大学, 1999 年,1 月.
- [45] 乔向东,信息融合系统中目标跟踪技术研究 [D]. 博士论文, 西安电子科技大学,2003 年,4 月.
- [46] Roecker J. Suboptimal joint probabilistic data association [J]. in IEEE Trans. Aero-space & Electronic Systems, vol. 29, no. 2, pp. 510–517, Apr 1993.
- [47] J. A. Roecker. A class of near optimal JPDA algorithms [J]. in IEEE Trans. Aerospace & Electronic Systems, vol. 30, no. 2, pp. 504–510, Apr 1994.
- [48] S. Oh, S. Russell, and S. Sastry. Markov chain monte carlo data association for general multiple target tracking problems [C]. In Proc. IEEE Conf. Decision & Control, Dec 2004.
- [49] S. Oh and S. Sastry. A polynomial-time approximation algorithm for joint probabilistic data association [C]. In Proc. American Control Conference, 2005.
- [50] D. Musicki. Joint integrated probabilistic data association [J]. IEEE Trans. Automatic Control, vol.AC-40, no. 3, pp. 1093–1099, 2004.
- [51] J. Vermaak, S. Godsill, and P. Perez. Monte Carlo filtering for multi-target

- tracking and data association [J]. in IEEE Trans. Aerospace & Electronic Systems, vol. 41, no. 1, pp. 390–332, Jan 2005.
- [52] Blackman S. Multiple Target Tracking with Radar Applications [M]. Norwood: Artech House, 1986.
- [53] Blackman S. Design and Analysis of Modern Tracking Systems [M]. Norwood: Artech House, 1999.
- [54] R. Singer and J. Stein. An optimal tracking filter for processing sensor data of imprecisely determined origin in surveillance systems [C]. In Proc. IEEE Conf. Decision & Control, Florida, USA, pp. 171–175, Dec 1971.
- [55] D. Reid. An algorithm for tracking multiple targets [J]. IEEE Trans. Automatic Control, vol. AC-24, no. 6, pp. 843–854, 1979.
- [56] L. Oliveira, U. Nunes. Semantic fusion of laser and vision in pedestrian detection [J]. Pattern Recognition, 2010:3648-3659.
- [57] 武博. 融合激光深度和图像特征的快速行人检测研究 [D]. 中国科学院研究生院博士学位论文, 2012.

## 个人简介及发表文章目录

### 个人简介

高山 男 汉族 中共党员

- 2006年9月至2010年7月 南开大学 学士
- 2010年9月至2013年7月 中国科学院大学 硕士

### 已发表文章目录

- **Shan Gao**, Zhenjun Han, Yang Xu, Qixiang Ye, and Jianbin Jiao, "Real-time pedestrian learning-tracking with information fusion," Proc. of International Conference On Internet Multimedia Computing and Service, 2012. (EI)

## 致 谢

在中国科学院研究生院（中国科学院大学）攻读硕士学位三年的学习生活中，没有什么成就，没有什么奖励，也没有什么精彩。一直认为自己是个普通平凡并且默默无闻的学生，不会有多大的成绩，是三位指导我研究学习的老师，焦建彬教授，叶齐祥副教授，韩振军老师给我一如既往的支持与鼓励，才让我不至于在研究道路上迷失，停滞，掉队。写毕业论文到了最后的总结和致谢环节，才发现自己心里满满的全是感谢。

本论文的研究工作是在焦建彬教授，叶齐祥副教授，韩振军老师的悉心指导下完成的。首先，我要感谢导师焦建彬教授在我攻读硕士学位期间从理论和实验方面给予的大量的、极其有益的指导，对我的每一个实验和每一篇论文都给予极大的支持。感谢叶齐祥副教授对我实验中的每一步，论文中的每一句都精益求精。感谢韩振军老师兼师兄，对我学习生活中的无论是小灵感，小思路，还是小困难，小抱怨都关怀备至。三位恩师在科研上精益求精，在学术上认真严谨，生活中的认真负责都令我敬佩。

其次，感谢中科院大学模式识别与智能系统实验室的所有已经毕业和还没毕业的师兄师姐，师弟师妹，我们一起科研，一起奋斗，一起成长，一起锻炼，一起欢笑，一起流泪。没有和他们的并肩作战，很难想象怎样在实验室里一周度过 60 个小时。

感谢我的父母，女朋友，亲戚好友，他们给我无私的爱和无条件的支持，永远是我前进中最坚强的后盾和依靠，愿他们永远安康幸福。

感谢参加开题及中期评阅的各位指导老师和专家，他们丰富的经验和细致的指导对论文方向和研究进度的指点给整个研究工作带来了巨大的帮助。

最后，感谢参加论文评审和答辩的所有老师。

高 山

2013 年 4 月