

密级: _____

中国科学院研究生院

硕士学位论文

融合激光深度与图像特征的快速行人检测研究

作者姓名: 武博

指导教师: 叶齐祥 副教授 中国科学院研究生院

学位类别: 工程硕士

学科专业: 物流工程

培养单位: 中国科学院研究生院工程教育学院

2012年4月

Fast pedestrian detection based on laser and image data
fusion

By
Bo Wu

A Paper Submitted to
Graduate University of Chinese Academy of Sciences
In partial fulfillment of the requirement
For the degree of
Master of Logisticis Engineering

College of Engineering

April, 2012

中国科学院研究生院直属院系 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：

中国科学院研究生院直属院系 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

现代社会的物流是物品从供应地到接收地的实体流动过程，在此过程中车辆的安全运输非常重要。车辆行驶智能安全保障技术中的行人防撞安全预警致力于通过检测车辆周边的行人目标，为驾驶员和行人提供安全保障。多传感器信息融合技术通过对多类同构或异构传感器数据进行综合，形成比单一信源更可靠、更完全的融合信息。本文以基于激光传感器和视觉传感器的融合信息进行行人检测用于防撞安全预警为主要的研究内容，在物流工程领域中有着重要的意义与应用价值。

本文提出了一种基于激光深度信息和图像轮廓信息融合的行人检测算法，该算法很好的综合了激光传感器和视觉传感器的优点，有效的提高了基于单一传感器的行人检测的精度和效率。本文的主要工作如下：

1、不同空间的特征提取与表示。在激光空间，对获得的激光深度信息进行处理，通过滤波和聚类分析将噪声点去掉；在图像空间，提取了人体样本的 HOG 特征并训练二级 SVMs 分类模型用于行人检测。

2、激光深度与视觉信息的融合。为了在图像上找到激光定位出的 ROI 区域以便合理融合两种信息，首先需要通过标定的方法解决激光扫描仪和摄像机之间的数据配准问题；其次，根据聚类片的特征值来判断候选目标；最后利用前期进行传感器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，获得了图片上的 ROI 区域。本算法可以实现快速的行人检测，并在一定程度上提高检测的精度。

3、针对传感器信息串行检测在互有遮挡的情况下的行人检测中的失效情况，对基于激光深度信息和图像轮廓信息融合的行人检测算法进行改进，提出了一种基于动态规划与信息融合的解决方案，在一定程度上有效的解决了复杂前景中的快速人体检测。

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。

关键词：行人检测，信息融合，深度信息，HOG 特征，支持向量机，动态规划

Abstract

- **Bo Wu** (Logistics Engineering)

Directed by: **Qi-xiang Ye** (Associate Professor)

The safe transport of the vehicle is very important in the logistics. One of the most important branches in the intelligent vehicle security technology is crash safety warning, which aims to detect the obstacles around vehicles for early warning the driver. Based on information from multi-sensor, we can get high quality information. Therefore, this subject of this paper has important theoretical significance and wide practical value in the field of Logistics Engineering.

In this paper, we proposed a pedestrian detection framework based on laser and image data fusion. The main content can be listed as following:

1) Feature extraction and representation in different space. In laser space, data from laser scanner is clustered into segments, each of which implies a candidate. The SVM classifier on Histogram of Oriented Gradient (HOG) features is used in image space.

2) Information fusion of Multi-sensor. The segments are projected to the image domain to form regions of interest (ROI), given camera calibration parameters.

3) In order to solve the problem of the failure in obstructed pedestrian detection of complex environment, we propose a dynamic programming-based solution.

Experiments report over 30 times higher speed than the state-of-the-art method and a comparable detection rate.

KEY WORDS: Pedestrian Detection, Information Fusion, Laser Scanner, HOG Features, SVM Classifier, Dynamic Programming

目 录

摘 要	I
目 录	III
图目录	V
表目录	VI
第一章 绪论	1
1.1 课题背景和研究意义	1
1.2 国内外研究现状	2
1.3 本文的研究内容	3
1.4 本文的组织结构	5
第二章 多传感器信息融合的研究基础	7
2.1 多传感器信息融合的基本理论	7
2.1.1 信息融合的概念	7
2.1.2 信息融合的层次	8
2.1.3 信息融合的方法	9
2.2 基于多传感器信息融合的目标检测简介	11
2.3 本章小结	12
第三章 基于激光深度信息与图像特征信息融合的行人检测研究	13
3.1 基于激光深度特征与图像轮廓信息融合的基本理论	13
3.1.1 激光深度特征提取	14
3.1.2 图像轮廓特征提取和特征分类	16
3.1.3 摄像机与激光扫描仪配准	17
3.2 基于激光深度特征与图像轮廓融合的行人检测算法	24
3.2.1 候选目标生成	24
3.2.2 ROI区域的确定	26
3.2.3 候选目标的确认	27
3.3 实验结果与对比分析	28
3.3.1 实验数据	28
3.3.2 标定板数据及配准参数	28
3.3.3 训练数据集	30
3.3.4 测试数据集	31
3.3.5 检测效果及对比分析	32
3.4 本章小结	33

第四章 基于动态规划与信息融合的行人检测算法	35
4.1 基于多传感器信息串行检测的局限性	35
4.2 基于动态规划与信息融合的行人检测	36
4.2.1 动态规划的基本原理	37
4.2.2 基于动态规划与信息融合的行人检测算法	39
4.3 检测效果	41
4.4 本章小结	41
总结与展望	43
参考文献	45
个人简介及发表文章目录	51
致 谢	53

图目录

图 1-1 基于单一传感器的人体目标检测图	4
图 1-2 基于传感器融合的检测框架	5
图 3-1 激光测距原理图	14
图 3-2 图像像素坐标系和物理坐标系	18
图 3-3 线性摄像机成像模型	19
图 3-4 成像过程	20
图 3-5 激光扫描仪与相机安装示意图	23
图 3-6 扫描仪返回的原始数据可视化	24
图 3-7 从激光点中产生候选目标	25
图 3-8 二维标定板	26
图 3-9 ROI区域生成结果	27
图 3-10 HOG特征示意图	28
图 3-11 人体检测示意图	28
图 3-12 标定实验设备	29
图 3-13 不同位姿的棋盘标定板	29
图 3-14 激光点投影变换	30
图 3-15 人体检测训练样本示例	30
图 3-16 同步采集的图片及其激光距离数据	31
图 3-17 检测结果示意图	32
图 3-18 检测速度比较	32
图 3-19 检测率比较	33
图 4-1 典型的激光点群	35
图 4-2 串行检测针对三类点群的有效性	36
图 4-3 串行检测失效情况	36
图 4-4 基于动态规划与信息融合的行人检测方法示意图	40
图 4-5 基于动态规划与信息融合的行人检测结果与原串行方法对比图	41

表目录

表 3-1 激光空间候选目标生成算法	26
表 4-1 基于动态规划与信息融合的检测算法	40

第一章 绪论

1.1 课题背景和研究意义

现代社会的物流是物品从供应地到接收地的实体流动过程，在此过程中利用车辆进行的运输过程非常重要。随着公路交通和汽车工业的发展，交通事故呈上升趋势，避免交通事故或减轻事故造成的伤害成为各国政府和社会亟待解决的问题，因此世界各国在研究智能车辆自主导航技术的同时，投入大量人力物力，致力于提高汽车行驶安全性。智能车辆技术按功能主要分为自主导航和安全保障两个方面^[1]。自主导航技术的应用依赖于整个智能交通系统（ITS）的建立和完善，短期内难以达到实用化，而安全保障技术却可以独立应用于辅助驾驶系统，通过对周围的行驶环境做出检测从而判断可能对驾驶员造成的威胁，因此对于解决因驾驶员主观因素产生的交通事故可以提供技术支持。智能车辆安全保障技术分为安全监测与预警和主动安全保障^[1]，安全监测与预警主要指借助传感器和报警系统来监测车辆驾驶者状况、车辆隐患、特殊环境等，从而帮助驾驶员增进安全驾驶状态，而其中针对车辆周围的人体目标检测致力于通过传感器对周边环境进行非接触探测来提高道路交通安全。

对于传感器来说，没有任何一种可以保证在任意时刻都提供完全可靠的信息。而多传感器信息融合技术通过多类同构或异构传感器数据进行综合以获得比单一传感器更多的信息，形成比单一信源更可靠、更完全的融合信息，从而得到有利于决策的对被感知对象的更加精确、更加完备的描述^[2]。利用多传感器信息融合来解决单一传感器普遍存在的有效探测范围小、数据可靠性低和易受外部环境影响等困难问题，从而快速、准确的对路面人体目标进行检测，对于提高路面行人防撞安全预警的速度和精度、进而提高车辆运输行驶过程中的安全性有着重要的意义，在物流工程领域中有着重要的应用价值。

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。

本文所研究的融合激光深度与图像特征的目标检测与基于多源信息的飞行器威胁目标检测具有方法上的共同之处。

1.2 国内外研究现状

鉴于多传感器信息融合在智能交通发展中的重要性，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辅助安全驾驶系统上，并取得了积极的成果。

在国外，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^{[3][4]}；丰田公司使用毫米波雷达和机器视觉共同探测前方障碍物^[5]；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^[6]。在国内，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，实现了对本车前方车辆的探测^[7]；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与跟踪及保持本车与前方车辆安全车距等功能^[8]；国防科技大学以轻型越野汽车为底盘，集成了激光雷达、可见光视觉、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^[9]。

在各种辅助安全驾驶系统中，基于多传感器信息融合的行人检测受到各方的广泛重视。根据其中利用传感器信息的不同时序，该研究可概括为两类：基于传感器信息串行的行人检测和基于传感器信息并行的行人检测。

基于传感器信息串行的行人检测，利用不同传感器得到的信息分别进行特征提取或计算，在上一个传感器信息分类结果的基础上利用下一个传感器信息进行分类判断。日本DENSO公司IT实验室的Szarvas等人，通过激光传感器搜索行人可能出现的ROI区域，通过卷积神经网络CNN对投影到图像上的ROI进行分类^[10]；意大利Parma大学的A.Broggi等人提出了一种环境主导的搜索，针对城市交通中路边停着的车对行人造成遮挡的问题，通过激光点定位出一个可能有

行人出没的危险区域（如被路边车遮挡而只出现了部分人体），避免了大范围的无用搜索^{[11][12][13]}。

基于传感器信息并行的行人检测，利用不同传感器同时收集并处理同一场景的不同数据，同步的分类检测数据并将结果进行融合。瑞士的ETH Surich大学的L. Spinello等人对传感器独立检测后的结果进行了融合，在激光空间，用adaboost方法提取特征，将其提供给条件随机场（CRF）来对行人和车辆进行分类，在图像空间，应用隐式形状模型（ISM）来对图像内的物体进行分类，传感器信息的融合利用了kalman滤波在跟踪过程中得到实现^[14]；葡萄牙Coimbra大学的L.Oliveira等人提出了一种将激光和视觉信息通过语义信息进行融合的方法，通过马尔可夫逻辑网（Markov logic network）对基于空间关系上的各个人体部分进行融合，该方法对于解决行人检测中的遮挡问题有较好的效果^{[15][16]}。

对于串行和并行的传感器信息融合检测方法来说，前者的判定分类速度快，但对各部分传感器单独应用的分类判决有效性有较高要求；后者不会使后续进行处理的传感器丢失信息，但处理速度比串行慢且误判较多。如何合理地确定融合思路，以便快速、全面的检测在车辆行驶环境中出现的行人，还存在着很多的问题和挑战。因此，我们需要研究快速有效的融合方法来进行路面人体的检测。

1.3 本文的研究内容

基于多传感器信息融合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的影响等难题，对于提高目标检测的可靠性具有重要意义。

本文主要研究基于轮廓和深度信息融合的人体目标检测。基于视觉信号的目标特征表示主要依赖于形状和轮廓等特征，特别是对于车辆、行人等目标的检测。目前对于特征的描述主要分为对色彩的处理和对轮廓的提取，这两方面最具有代表性的研究成果是Viola提出的Haar-Like算法^[18]和Dalal^[19]采用的HOG（及改进特征v-HOG，LBP-HOG等^[20]）算法。这些特征在很多检测上取得了成功。然而，这两种特征都存在着各自的不足，前者对于颜色和光照过于

敏感,后者由于是一种区域内的方向统计,容易受到纹理的影响。如图 1-1(a)(b)所示,其中(a)受光照影响发生误检,(b)受纹理影响将不在同一深度层次的物体误检为人体,且基于视觉信息的目标检测运算速度较低,无法满足辅助安全驾驶对实时性较高的要求。利用激光扫描仪可以比较清晰的区分具有不同大小以及深度的物体(如人体和墙面),但对具有相似长度和深度的物体(如人体和车体片段)则不能识别,如图 1-1(c)(d)所示,其中的部分车体返回的激光点群无法依赖激光信息进行区分。通过融合激光扫描特征与图像特征,利用视觉的方法对激光扫描仪所不能区分的具有相似特征长度的物体区域进行检测,既可很好的识别和区分不同类别的物体,又由于减少了图像搜索区域而提高检测速度。所以,本文利用激光测距得到的深度信息和CCD摄像机得到的图像信息进行融合,得到车辆行驶前方障碍物的深度信息并对障碍物区域进行判断,从而避免因单一传感器不稳定造成的障碍物目标多报和误报。



图 1-1 基于单一传感器的人体目标检测图,其中(a)和(b)是基于视觉信息的人体目标检测结果图,(c)和(d)是基于激光信息的人体目标检测结果图。

多传感器提供的往往是异类平台、不同时相的异质信息,这给融合处理带来了困难。造成信息不一致的主要原因有:

1. 信息描述空间不同。在多传感器系统中,各个传感器提供的信息均为对目标环境在自身传感器空间的描述,由于各传感器的异类属性导致

2. 信息维数不匹配。多传感器获取的信息维数不一定相同，且图像质量水平也不完全一致。

如何融合深度和轮廓信息，利用两种信息的优势互补达到快速、准确检测危险区域的障碍物，成为本文的研究内容。利用不同传感器进行融合检测，首要问题是传感器的异步传输问题，可通过对激光扫描仪的端口读取数据进行与摄像机同频率的数据截取而实现。融合的另一问题是激光扫描仪与摄像机的信息描述空间不同，可以通过对摄像机和激光扫描仪进行外部标定来表示激光坐标系到图像坐标系的变换关系。

本文以行人检测作为主要内容，所提出的方法对于检测车辆等其他障碍物同样有效。论文的检测框架如图 1-2 所示，我们通过激光空间获得的候选目标位置信息，投影映射到图像相关位置上得到图像空间的感兴趣区域（ROI）。在 ROI 内，利用基于外观的方法，提取特定目标（人体）的特征并训练分类模型用于对 ROI 内目标的检测。

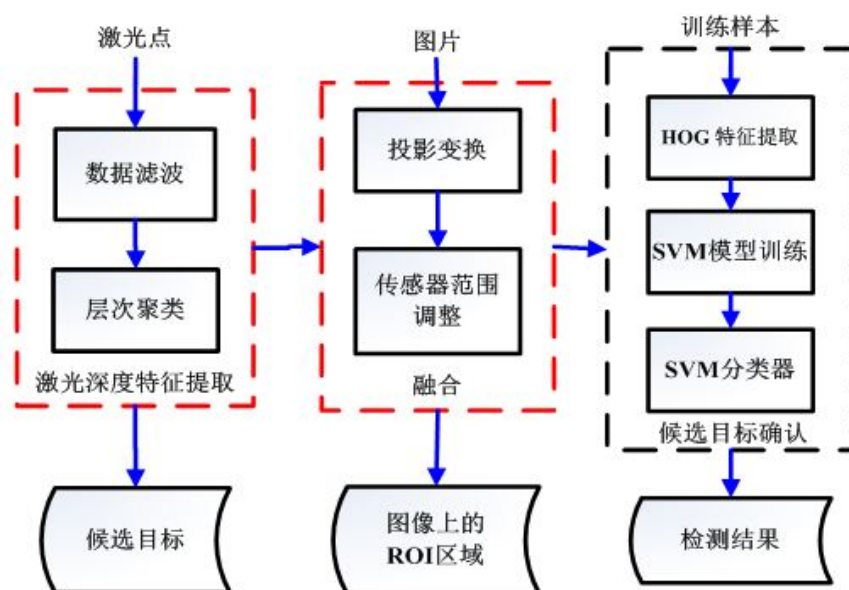


图 1-2 基于传感器融合的检测框架

1.4 本文的组织结构

第一章，绪论。主要论述了基于多传感器信息融合的行人检测的研究背景和意义，分析了国内外多传感器信息融合用于人体检测的研究现状，同时分析

了其在辅助驾驶系统的应用现状及发展趋势，列出了本文的主要研究目的和研究内容。

第二章，多传感器信息融合的研究基础。主要综述了用于目标检测的多传感器信息融合的主要研究内容和方法，同时也对基于激光深度特征与图像特征融合的行人检测的基本思想进行了简单介绍。

第三章，基于激光深度与图像特征融合的行人检测。主要论述了激光传感器和视觉传感器在用于目标检测方面的优势、具体特征选取方法以及融合方法，详述了基于激光深度与图像特征融合的行人检测算法的三个步骤。实验结果与对比分析论述了行人检测实验过程所使用的数据获得和实验步骤，将本文提出的方法与其他方法做了比较并列举了部分实验结果。

第四章，基于动态规划与信息融合的行人检测。针对复杂前景中的行人检测中串行方法失效的问题，提出一种基于动态规划以及信息融合的有效检测方案。

最后总结了本文的主要工作，展望了未来工作的方向，以及对如何进一步提高检测算法的鲁棒性、准确性等研究热点和难点的探讨。

第二章 多传感器信息融合的研究基础

2.1 多传感器信息融合的基本理论

2.1.1 信息融合的概念

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^[21]。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^[28]。

由于应用广泛，因此各个行业对于信息融合的定义有所不同。美国三军组织实验室理事联合会 JDL（Joint Directors of Laboratories）给出的信息融合定义为大多数研究者所接受，定义从军事应用出发，将融合定义为一个过程^[28]：信息融合是一种多层次、多方面的处理过程，包括对来自单一和多源的数据进行关联、相关和综合处理，从而提高状态和身份估计的精度，以及对战场态势和威胁的重要程度进行适时完整的评价。定义包括三个要点，数据融合是多信源、多层次的处理过程，每个层次代表信息不同抽象程度；信息融合过程包括数据的检测、关联、估计与合并；信息融合的输出包括低层次上的状态身份估计和高层次上的总战术态势的评估。信息融合的实质是针对多维信息进行关联或

综合分析, 选取适当的融合模式和算法用以提高信息的质量, 为高效信息的获取提供支持。

2.1.2 信息融合的层次

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^[2]。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^[2]。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。

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细微信息, 但面临的是大量需要处理的信息, 处理代价高且实时性差, 又由于融合是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, 要求在融合时有较高的纠错能力; 特征级融合实现了客观的数据压缩, 有利于实时处理, 但由于损失了一部分有用信息, 使得融合性能有所降低; 决策级融合具有容错性强、通信量小、抗干扰能力强、融合中心处理代价低等优点, 其缺点是预处理花销大, 数据损失量最大, 精度低。总之, 数据级融合的信息处理量大、处理时间长、实时性较差; 特征级融合能够保留足够数量的重要信息和实现信息压缩, 从而有利于实时处理; 决策级融合在融合之前, 每种传感器的信号处理装置已完成决策或分类任务。信息融合根据一定的准则和决策的可信度做最优决策, 以便具有更好的实时性和容错性, 使得在一种或几种传感器失效时也能工作。

综上, 信息融合层次越高则用于融合的信息抽象程度越高, 对传感器的同质要求就越低; 反之, 信息融合层次越低则用于融合的信息抽象程度越低, 但融合后保持的细节信息越多则要求信息间的配准性高。

2.1.3 信息融合的方法

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。经过这些年的发展，信息融合技术的主要方法主要分为五类^[27]：信号处理与估计理论方法、统计推断方法、信息论方法、决策论方法以及人工智能方法。

信号处理与估计方法包括最小二乘、加权平均、Kalman滤波等线性估计技术，还有扩展Kalman滤波、高斯滤波^[34]等非线性估计技术。期望极大化（EM）算法^[35]为求解在具有不完全观测数据的情况下的参数估计与融合问题提供了新的思路。

经典的统计推断方法包括Bayes推理、证据推理^[36]以及Random set^[37]理论、支持向量机（SVM）理论等。

信息论是运用概率论与数理统计的方法研究信息、信息熵、通信系统、数据传输、密码学、数据压缩等问题的应用数学学科。信息论将信息的传递作为一种统计现象来考虑，给出了估算通信信道容量的方法。信息传输和信息压缩是信息论研究中的两大领域。这两个方面又由信息传输定理、信源—信道隔离定理相互联系。信息论采用优化信息度量的方法融合多源数据。典型算法有熵方法^[38]、最小描述长度原理MDL^[39]等。

决策论是根据信息和评价准则，用数量方法寻找或选取最优决策方案的科学，是运筹学的一个分支和决策分析的理论基础。在实际生活与生产中对同一个问题所面临的几种自然情况或状态，又有几种可选方案，就构成一个决策，而决策者为对付这些情况所取的对策方案就组成决策方案或策略。在多源信息融合问题中，决策论方法被用于高级别的决策融合。如Fitzgerald在论文^[40]中采用决策论方法融合可见光、红外及毫米波雷达数据用于报警分析。

人工智能方法包括神经网络、遗传算法、模糊逻辑、自然语言处理、基于规则的推理以及专家系统。

在以上提及的大量的算法中，常用的^[41]有Bayes方法、证据理论推理、模糊理论神经网络方法和SVM理论。

■ Bayes 方法

Bayes方法提供了一种计算假设概率的方法，这种方法是基于假设的先验概率、给定假设下观察到不同数据的概率以及观察到的数据本身而得出的。其方法为，将关于未知参数的先验信息与样本信息综合，再根据贝叶斯公式，得出后验信息，然后根据后验信息去推断未知参数的方法。Bayes方法的难点在于寻找合适的概率分布^[42]，当数据来源于低档传感器时更加困难；在实际中，先验概率难以确定，造成了坏的推理结果。

■ 证据理论

证据理论是由Dempster于1967年首先提出，并由他的学生Shafer于1976年进一步发展起来的一种不精确推理理论，也称为Dempster/Shafer证据理论(D-S证据理论)，属于人工智能范畴，最早应用于专家系统中，具有处理不确定信息的能力。作为一种不确定推理方法，证据理论的主要特点是：满足比Bayes概率论更弱的条件；具有直接表达“不确定”和“不知道”的能力^[43]。证据理论在表示和处理不确定问题时具有明显的优势，然而，随着信源数的增长，其运算量成指数增长^[44]。其判决规则常具有很大的主观性。

■ 模糊理论

模糊理论是在美国加州大学伯克利分校电气工程系的L.A.zadeh教授于1965年创立的模糊集合理论的数学基础上发展起来的，主要包括模糊集合理论、模糊逻辑、模糊推理和模糊控制等方面的内容。模糊理论以模糊集合为基础，其基本精神是接受模糊性现象存在的事实，而以处理概念模糊不确定的事物为其研究目标，并积极的将其严密的量化成计算机可以处理的讯息，不主张用繁杂的数学分析即模型来解决模型。模糊理论进一步放宽了概率论定义中的制约条件，从而可以对数字化信息进行宽松建模^[45]。模糊理论可以与其他信息融合算法结合，用于提高信息融合处理的精度^[46]。

■ 神经网络

人工神经网络是由大量处理单元互联组成的非线性、自适应信息处理系统，试图通过模拟大脑神经网络处理、记忆信息的方式进行信息处理。人工神经网络具有自学习和自适应的能力，可以通过预先提供的一批相互对应的输入-输出数据，分析掌握两者之间潜在的规律，最终根据这些规律，用新的输入数据来推算输出结果，这种学习分析的过程被称为“训练”^[47]。神经网络会训练一种模式，然后预测给出的数据属于哪一类，属于有指导的分类方法。信息融合

过程接近人类的思维过程，与人脑神经系统有较大的相似性，因此利用神经网络的结构优势和其高速并行处理能力进行信息融合是一种有效方法。但进行融合处理时，存在学习过程中运算量大，寻找全局最优解也比较困难的问题。

■ SVM 理论

Vapnik等人在多年研究统计学习理论基础上对线性分类器提出了另一种设计最佳准则。其原理从线性可分说起，然后扩展到线性不可分的情况，甚至扩展到使用非线性函数中去，这种分类器被称为支持向量机(Support Vector Machine, 简称SVM)。

2.2 基于多传感器信息融合的目标检测简介

基于多传感器融合的目标检测研究分为串行和并行两类，其中基于传感器信息串行的目标检测利用不同传感器得到的信息分别进行特征提取或计算，在上一个传感器信息分类结果的基础上利用下一个传感器信息进行分类判断。而基于传感器信息并行的目标检测利用不同传感器同时收集并处理同一场景的不同数据，同步的分类检测数据并将结果进行融合。两种检测思路各有优缺，需针对具体应用进行合理地取舍。串行检测分类速度快，但对各传感器单独应用的有效性有较高要求；并行检测不会丢失信息，但处理速度比串行慢且误检率高。

在本文中，我们使用了激光扫描仪和 CCD 摄像机两种传感器。对于前景较为简单的、不存在相互遮挡的行人检测来说，单独应用激光传感器进行初步选择时可靠性较高，因此我们利用串行的方法进行检测，通过融合从激光扫描仪得到的深度数据和从 CCD 摄像机得到的图片信息，来实现对行人目标的准确和快速检测。通过在激光空间获得的障碍物位置信息，投影映射到图像相关位置上得到图像空间的感兴趣区域 (ROI)。在 ROI 内，利用基于外观的方法，提取特定目标 (人体) 的特征并训练分类模型用于对 ROI 内目标的检测。整体算法分三个步骤进行：

- 1、候选目标生成阶段，对获得的激光深度信息进行处理，通过滤波和聚类分析将噪声点去掉，并且根据聚类片段的特征值来判断候选目标。

- 2、ROI 区域生成阶段，利用前期进行传感器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 ROI 区域输入第三步进行确认性检测。

3、候选目标的确认阶段，在获得 ROI 后，利用视觉检测的方法对此区域进行候选目标的确认性检测。HOG 特征提取之后，我们训练一个 SVM 分类模型用于目标检测。

上面提出的算法考虑了通常的情况，即无遮挡的行人检测。在这种情况下，检测范围内的各个物体在激光扫描条件下是可分的。但在复杂前景中有遮挡的行人检测状况下，会出现如下问题：由于距离过近，不同物体的激光点密集而平均，聚类后也无法区分。在这种情况下利用串行思路进行检测时，激光扫描仪无法为图像检测提供正确的信息，从而使得图像检测无法继续进行，最后导致无法检测到行人目标。针对这种有遮挡的行人检测情况下利用串行方法会导致目标丢失的弊端，本文提出了一种基于动态规划与信息融合的行人检测方案。

检测算法中涉及到的具体方法有：特征提取，包括深度特征提取（数据滤波和层次聚类）；轮廓特征提取（HOG 特征提取和 SVM 分类）。特征对准，包括摄像机外部标定；摄像机与激光扫描仪联合标定。以上具体方法的理论依据及实现将在第三章和第四章中详细阐述。

2.3 本章小结

本章是多传感器信息融合的一些基本理论。主要综述了多传感器信息融合的基本理论以及用于目标检测的多传感器信息融合的主要研究内容和方法，总结了现有的多传感器信息融合检测的串行和并行两种思路，最后对基于激光深度特征与图像特征融合的行人检测基本思想进行简单介绍，引出下一章中对算法具体理论和实现细节的详细阐述。

第三章 基于激光深度信息与图像特征信息融合的行人检测研究

CCD 摄像机价格便宜且能够提供丰富的视频资料信息，通过采用计算机视觉技术对视频资料进行分析，丰富的视频信息可以用来进行相关的应用，如道路检测、交通标识识别以及行人、障碍物的检测与识别等^[2]，但其存在着存储量小、计算速度慢和模型泛化能力的局限性；激光扫描仪的算法实现简单，性能较好，但其分辨率较低，提供的信息不如图像信息丰富^[49]。相对于单一的视觉传感器来说，多传感器信息融合技术通过多类同构或异构传感器数据进行综合，可以获得比单一传感器更多的信息，形成比单一信源更可靠、更完全的融合信息。利用多源信息融合来解决单一传感器普遍存在的有效探测范围小、数据可靠性低和易受外部环境的影响等困难问题，对于提高目标检测的可靠性有着重要意义。

在本文中，我们选择激光扫描仪和 CCD 摄像机这两种传感器，通过融合激光深度特征与图像特征进行行人检测，不仅可以克服基于视觉信息进行行人检测时易受天气状况和光照条件变化影响、无法有效获取检测对象的距离信息且计算复杂度非常大等缺点，也可以克服激光扫描测距精度较低、噪声大、无法直观判断障碍物类别等缺点。本章中基于激光深度特征与图像特征融合的行人检测可以快速可靠的检测行人目标，其中涉及到的基本理论和算法步骤将在本章中进行详细介绍。

3.1 基于激光深度特征与图像轮廓信息融合的基本理论

利用激光扫描仪的距离信息与 CCD 摄像机的图像信息融合进行目标检测，距离信息与图像信息是在不同的传感器平台上采集的，其信息描述空间存在差异性，需要解决的问题有：不同空间内各自信息特征的提取与描述；图像空间内分类模型的训练；两种空间表示坐标的配准。

涉及的具体方法与理论有：

- 特征提取：深度特征提取和轮廓特征提取。

- 特征对准：摄像机外部标定和摄像机与激光扫描仪配准。

3.1.1 激光深度特征提取

■ 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，且返回波的能量足以触发阈值，激光扫描仪计算它到物体的距离值；激光扫描仪连续不停地发射激光脉冲波，激光脉冲波打在高速旋转的镜面上，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。在激光扫描仪的扫描范围内，扫描仪输出每个测量点的距离，根据此距离信息，可以计算物体的外形轮廓、坐标定位等^[50]。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 3-1 所示为激光测距原理图。

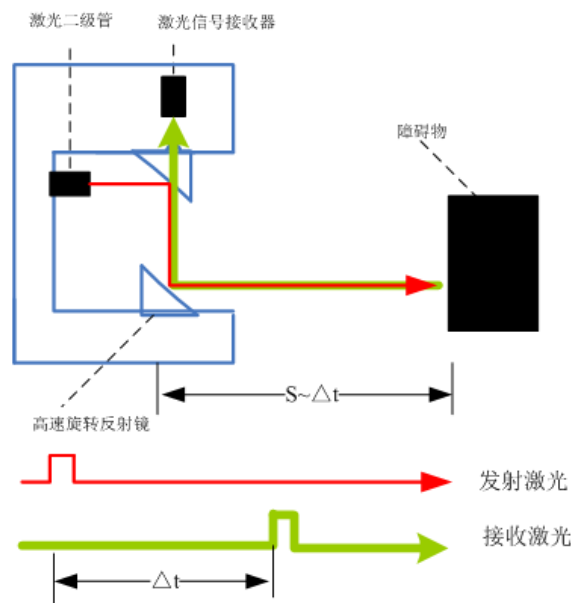


图 3-1 激光测距原理图

激光扫描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，其中每个数据表示在该元素对应角度方向上最近目标的距离。具体的离散数据点与扫描仪的角度分辨率有关。

扫描数据由极坐标形式给出（ d 表示距离， φ 表示扫描角度）：

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

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, 与激光扫描仪的角度分辨率有关。

激光扫描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^[2], 激光扫描数据中存在大量的噪声点和无效点, 要通过激光扫描仪进行目标检测, 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。

■ 基于深度信息的层次聚类分析

作为一种常用的模式识别方法, 聚类分析在处理大规模数据集中发挥着重要作用。在目前的应用中, **K-means**法和系统聚类法是聚类分析中最为常用的两种方法。其中的**K-means**法计算速度快, 但需要事先根据样本空间分布指定分类的数目。当样本的变量数超过 3 个时, 该方法的可行性较差。而层次聚类法(**Hierarchical clustering methods**) 则由于类与类之间的距离计算方法多种多样, 且不需要事先指定分类数目, 而使其适应我们的激光点聚类要求。该方法的基本思想^[51]是: 先将 n 个样本各自看成一类, 并规定样本间的距离和类间的距离。开始时, 每个样本自成一类, 类与类之间的距离与样本之间的距离是相同的, 在所有的类中, 选择距离最小的两个类合并成一个新类, 并计算所得新类和其它各类的距离; 接着再将距离最近的两类合并, 这样每次合并两类, 直至将所有的样本都合并成一类为止。

层次聚类法在进行聚类的过程中, 需要计算类与类之间的距离。根据类与类之间的距离计算方法的不同, 层次聚类法分为单连接法、完全连接法、平均连接法、组平均连接法与离差平方和法等^{[51][52]}。其中的单连接法首先将距离最近的样本归入一类, 即合并的前两个样本是它们之间有最小距离和最大相似性; 然后计算新类和单个样本间的距离作为单个样本和类中的样本间的最小距离, 尚未合并的样本间的距离并未改变。其中认为两类之间的距离是它们两个最近点间的距离, 因此单连接法又称最短距离法。

用 d_{ij} 表示样本 i 和 j 之间的距离, G_1, G_2, \dots 表示类。单连接法定义类间的距离为两个类中最近样本间的距离, 如类 G_x 和 G_y 间的距离 G_{xy} 表示为:

$$G_{xy} = \min_{i \in G_x, j \in G_y} d_{ij} \quad (3-3)$$

由于单连接法每次合并新类，都是将该类与其它类中距离最近的两个样本之间的距离作为该类与其它类的距离，因此这种聚类方法的逐次并类距离之间的差距可能会越来越小。因此，该方法具有距离收缩的性质^[51]。但是单连接方法认为，只要单个样本之间的相异度小，就认为两个组是紧邻的，而不管组间其他样本间的差异大小。这种方法倾向于合并由一系列本身位置（原始数据集中样本的排列）靠近的样本。

3.1.2 图像轮廓特征提取和特征分类

在图像特征提取阶段，我们使用基于外观的方法，该方法从大量的训练数据中学习一类目标的特性。首先，每一个训练样本都用一组局部或全局特征来表示，然后，通过训练一个分类器来学习目标和非目标的决策边界，最后，利用该决策边界来对新输入的图像进行判断。

对于训练数据，通常需要大量不同外观的人体数据作为正例样本和许多不同的非人体数据作为反例样本。特征提取方法和分类器的选择是基于外观的检测方法的核心问题^{[56][57]}。局部方向编码，Haar小波特征，梯度方向直方图特征（Histograms of Oriented Gradients, HOG）都已经被用于特征表述，其中，HOG被广泛地用于表述、分析和识别图像，因为它计算简便有效，并且对噪声和图像局部变化具有一定的鲁棒性。在分类方法方面，基于统计学习的分类器目前占据了主导地位。近邻分类器、人工神经网络、支持向量机（Support Vector Machine, SVM）、Adaboost等机器学习理论广泛用于多种信号源图像中的目标检测。

在本文中，我们提取目标的HOG特征，并结合SVM进行目标检测。

■ HOG 特征提取

Dalal 和 Triggs^[19]在2005年提出了HOG特征用于人体检测，该特征提出的思想是基于一个基本观点：物体的局部外观和形状能够用局部像素的梯度值的分布或边缘信息来描述，即使不知道对应的梯度值或边缘位置的精确信息。HOG特征跟边缘方向直方图（edge orientation histograms）、尺度不变特征变换（scale-invariant feature transform descriptors）以及形状上下文方法（shape contexts）有很多相似之处，不同之处在于它是在一个均匀分布的单元内进行计算的，并且为了提高性能，又将这些单元组成一些相互重叠的块。

HOG特征描述为：首先将图像分成小的连通区域，然后采集区域单元中各像素点的梯度方向直方图，最后把这些直方图组合起来构成HOG特征。为了提高性能，还可以将这些局部直方图在图像的更大的范围内（文献^[19]中称为区间或block）进行对比度归一化。与其他的特征描述方法相比，HOG特征对图像形变和光照有很好的适应性。

■ 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，构造出最优平面，使得正例和反例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^[48]，并且在目标检测应用中表现出了很好的性能。

3.1.3 摄像机与激光扫描仪配准

客观世界的三维场景投影到相机的二维平面是一个损失深度信息的透视投影过程，因此通过 CCD 摄像机拍摄的图片可以捕获周围环境而获取二维视觉信息，而激光扫描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扫描仪数据时，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，从而解决激光扫描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，我们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则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变换，为了定量的描述摄像机成像过程，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^{[53][54]}。

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 3-2 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。

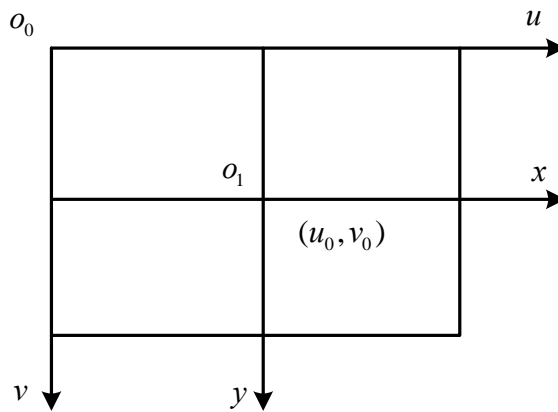


图 3-2 图像像素坐标系和物理坐标系

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 3-2 所示， x_0y_0 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：

$$u = \frac{x}{dx} + u_0 \quad (3-4)$$

$$v = \frac{y}{dy} + v_0 \quad (3-5)$$

矩阵形式为:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3-6)$$

3) 摄像机坐标系

摄像机坐标系以小孔摄像机模型的焦距中心为原点, 以摄像机光轴为 z_c 轴建立摄像机坐标系。如图 3-3 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 轴为摄像机光轴, 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 。

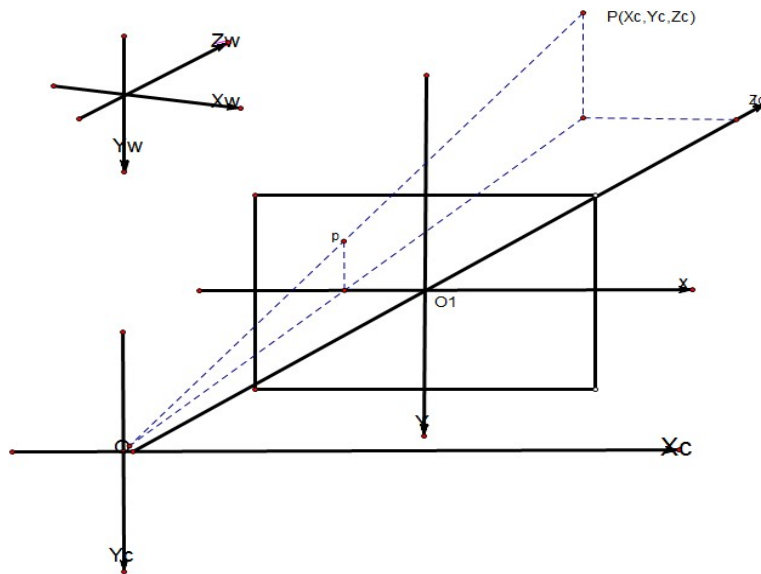


图 3-3 线性摄像机成像模型

4) 世界坐标系

世界坐标系也称为真实或现实世界坐标系, 或全局坐标系, 是客观世界的绝对坐标, 由用户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摄像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 3-3 所示。摄像机坐标系与世界坐标系之间的关系用旋转矩阵 \mathbf{R} 与平移向量 \mathbf{t} 来描述。空间中的点 \mathbf{P} 在两个坐标系下的齐次坐标关系为:

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (3-7)$$

■ 线性摄像机成像模型

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,也称为针孔模型。

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。针孔模型主要有光心(投影中心)、成像面和光轴组成。我们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,针孔中的点被理解为投影中心。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 3-3 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:

$$\frac{x}{f} = \frac{X_C}{Z_C} \quad (3-8)$$

$$\frac{y}{f} = \frac{Y_C}{Z_C} \quad (3-9)$$

矩阵表示为:

$$Z_C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} \quad (3-10)$$

从世界坐标系(三维场景)到数字图像(图像坐标系)的成像变换可以通过以下 3 个步骤来实现(本文不考虑摄像机畸变),如图 3-4 所示。

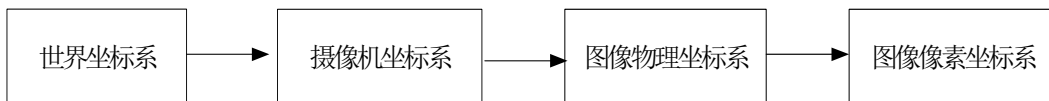


图 3-4 成像过程

矩阵表示为:

$$Z_C \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (3-11)$$

整理后得到:

$$Z_C \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_x & 0 & u_0 & 0 \\ 0 & a_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (3-12)$$

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:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (3-13)$$

正交旋转矩阵含有 3 个独立变量 (表示 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 个参数为摄像机外部参数, 完全由摄像机相对于世界坐标系的方位决定。

■ 摄像机外部标定

根据是否需要标定参照物, 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摄像机模型下, 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和计算方法, 求取摄像机模型的内部参数和外部参数; 摄像机自标定方法不依赖于标定参照物, 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摄像机进行标定。因其中的未知参数太多而难以得到稳定结果,

故而自标定方法虽然非常地灵活，但却并不是很成熟。一般来说，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。而自标定方法主要应用于精度要求不高的场合。

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）是一种适合应用的新型灵活方法。该方法要求摄像机在两个以上不同的方位拍摄同一个平面标定板，摄像机和 2D 平面标定板都可以自由移动，且内部参数始终不变。假定 2D 平面标定板在世界坐标系中的 $Z=0$ ，那么，通过线性模型分析就可计算出摄像机参数的优化解，然后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，可操作性和实用性强。

张正友标定法^[55]首先将式（3-12）表示如下：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = sA \begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{bmatrix} \quad (3-14)$$

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 向量可忽略。定义投影矩阵 H ，将式（3-14）改写为：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix}, \quad H = sA \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \quad (3-15)$$

通过式（3-14）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：

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = sA \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \quad (3-16)$$

由式（3-16）中 r_1, r_2 相互正交得到内参数 A 的两个约束条件：

$$h_1^T A^{-T} A^{-1} h_2 = 0 \quad (3-17)$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2 \quad (3-18)$$

文献^[55]中通过将 $A^{-T} A^{-1}$ 改写为对称的 B 矩阵，由约束条件（3-17）（3-18），在 B 矩阵的封闭解中求得摄像机的内部参数。同时，每副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。

■ 摄像机与激光扫描仪联合标定

激光扫描仪与CCD标定技术^[60]是在张正友提出的基于移动平面模板的标定方法的基础上，加入激光测距仪本身所在的激光坐标系（二维扫描面）。方法核心在于通过标定板同时被激光扫描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 3-5 表示了标定时传感器的安装位置：

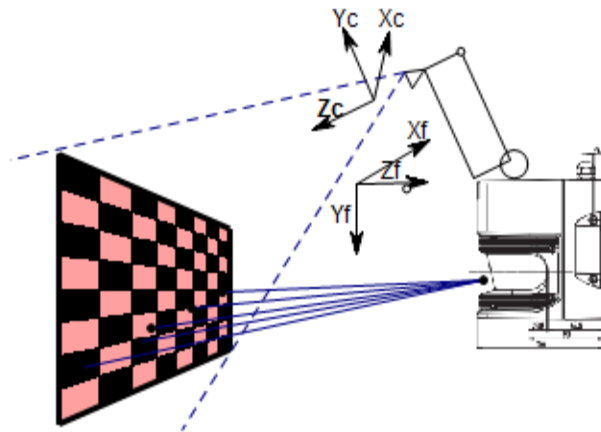


图 3-5 激光扫描仪与相机安装示意图

从摄像机坐标系 p 到激光坐标系 p^f 的变换关系如 (3-19) 式，其中 Φ 为旋转矩阵，表示两个传感器之间的相对方向； Δ 为平移向量，表示两个传感器之间的相对位置。

$$p^f = \phi p + \Delta \quad (3-19)$$

在摄像机坐标系下，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， $\|N\|$ 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：

$$N \cdot p = \|N\|^2 \quad (3-20)$$

其中，由 (3-20) 可得到 (3-21)。 N 可以通过相机的外参数矩阵 $[R, t]$ 求出。

$$N = -r_3(r_3^T \cdot t) \quad (3-21)$$

由 (3-19)、(3-20) 可推导出 (3-22)。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。

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

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 ϕ 和 Δ ，由此即完成了激光扫描仪与摄像机的联合标定。

3.2 基于激光深度特征与图像轮廓融合的行人检测算法

基于激光深度特征如图像轮廓融合的人体目标检测算法分为三个步骤，候选目标生成，ROI 区域生成和候选目标的确认。

3.2.1 候选目标生成

激光扫描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几何位置和形状。实验选择的 2 维激光扫描仪为 SICK LMS291，其扫描范围是 100° ，角度分辨率为 0.25° ，距离范围在 80 米之内。激光扫描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，其中每个数据表示在该元素对应角度方向上最近目标的距离，具体的离散数据点与扫描仪的角度分辨率有关，在本实验中每次返回 400 个离散点距离数据，扫描数据由极坐标形式给出，即：

$$l_z = (d_z, \phi_z)^T, z = 1 \dots 400 \quad (3-23)$$

在笛卡尔坐标系表示为：

$$u_z = (x_z, y_z)^T, z = 1 \dots 400 \quad (3-24)$$

其中 $x_z = d_z \cos \phi_z, y_z = d_z \sin \phi_z$ 。如图 3-6 所示为激光扫描仪返回的离散数据可视化。

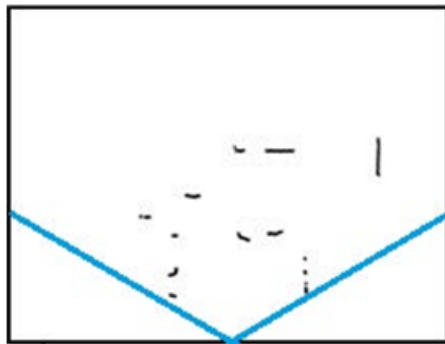


图 3-6 扫描仪返回的原始数据可视化

激光扫描仪对不同材质、不同颜色的物体存在测量误差，在无效数据中，很大一部分表现为零值和超出最大测距范围的值，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得以保留，即得到粗分类结果集合： $C = \{c_n\}, n = 1, \dots, N$ 。

在激光扫描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，因此需对离散的数据进行聚类分析。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。

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，我们保留片段长度在 $T([t_1, t_2])$ 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。我们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 3-7 表示了从激光点中产生的候选目标结果。

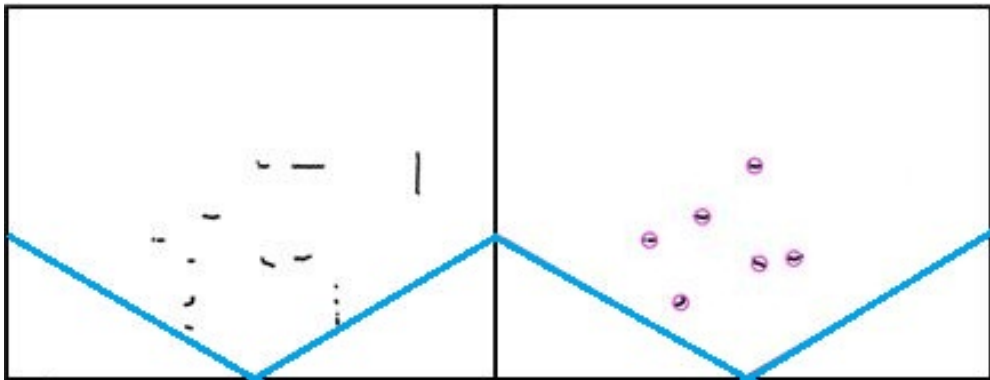


图 3-7 从激光点中产生候选目标

表 3-1 激光空间候选目标生成算法

输入: $L = \{l_z\}_{z=1}^Z$ 扫描一次输出激光数据点

输出: $\{f_m, d_m\}_{m=1}^M \subset S, T$

- 1 $T([t_1, t_2]) \leftarrow$ 目标障碍物长度平均值;
- 2 $C = \{c_n\} \leftarrow$ 粗分割滤波 ($\{l_z\}$);
- 3 $S = \{s_m\}, \{f_m, d_m\} \leftarrow$ 层次聚类 ($\{c_n\}$)。

3.2.2 ROI区域的确定

为了在图像上找到激光定位出的 ROI 区域, 需要通过标定的方法解决激光扫描仪和摄像机之间的数据对准问题。我们利用 3.1.3 中提到的激光扫描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。

首先使用张正友提出的基于移动平面模板的标定方法^[55], 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 H 。得到投影矩阵 H 后, 可按照张的方法^[55]求解摄像机内外参数。

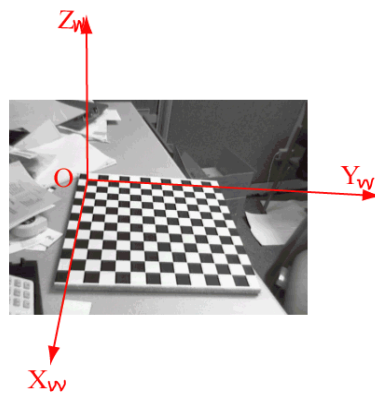


图 3-8 二维标定板

接下来利用文献^[60]中提出的激光扫描仪与 CCD 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ROI区域输入第三步进行确认性检测。利用式(3-22)，其中 N 为标定板参数化后的3维向量，方向平行于标定板法向， N 可以通过相应的摄像机外部参数矩阵 $[R,t]$ 求出； p^f 为激光坐标系中位于标定板上的某一点； Φ 为旋转矩阵，表示两个传感器之间的相对方向； Δ 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。

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

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 Φ 和 Δ ，由此即完成了激光扫描仪与摄像机的联合标定。获得了 Φ 和 Δ ，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ROI区域。投影结果如图3-9所示。

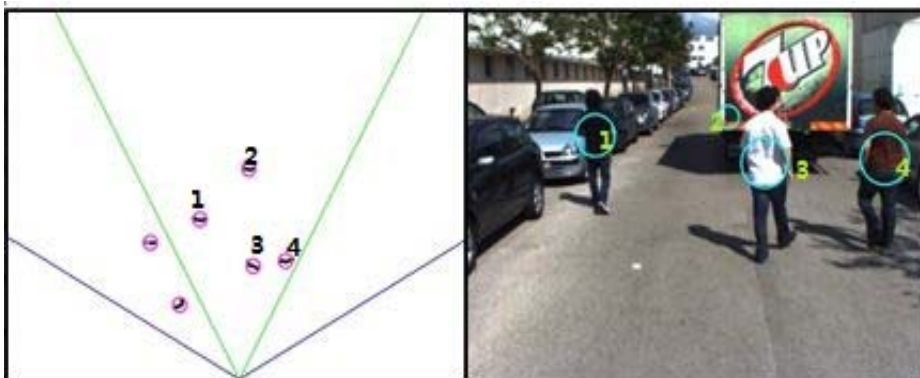


图 3-9 ROI 区域生成结果

3.2.3 候选目标的确认

在获得 ROI 后，需要利用视觉检测的方法对此区域进行检测，以对候选目标进行确认性的检测。

在检测之前，需要提取目标特征并训练分类模型。本文所提出的方法应用于人体目标检测，方法对车辆检测同样适用。在视觉检测的部分需要针对不同目标训练不同的分类器（如人体、车辆等）。

HOG 特征提取：将样本图像（人体为 64×128 像素的训练样本）按 8×8 个像素分割为若干个单元（cell）；将每相邻的 4 个 cell（田字结构）划分为一个 block（利用每个 cell 进行滑动而生成）。将 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 的梯度方向平均划分，其中规定每 20 度为一个 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 维向量归一化后串连起来，由此得到每个训练样本的 HOG 特征向量。图 3-10 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。对于该 ROI 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目标，如图 3-11。



图 3-10 HOG 特征示意图

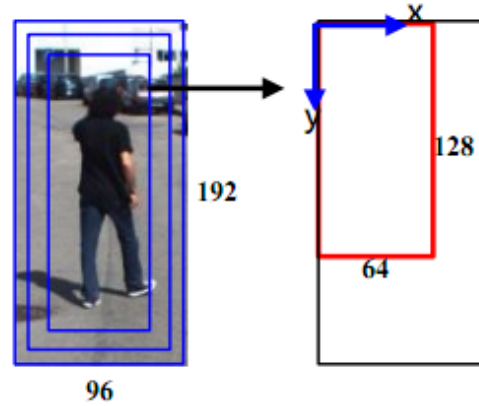


图 3-11 人体检测示意图

3.3 实验结果与对比分析

本文完成了基于激光深度与图像轮廓融合的人体目标检测算法的初步工作。本节中我们利用实验室自有的激光和摄像机设备，进行了激光传感器与摄像机的联合标定实验；利用ISR-UC-implidar-sync^[15]中的实验数据，实现了本文的算法并与其他检测方法进行了比较。

3.3.1 实验数据

前期实验 激光传感器与摄像机的联合标定实验中用到的不同位姿标定板图像及其对应激光距离数据，经过计算获得传感器配准参数；候选目标确认阶段进行目标检测需要的训练数据集（本文列举了行人检测需用的训练数据集）。

测试阶段 实验阶段所用的测试数据集包括一组视频中的连续图片及其匹配的不同步激光距离数据。

3.3.2 标定板数据及配准参数

我们进行了联合标定实验。首先将两传感器的相互位置固定，采集 10 张不

同位姿的棋盘格图片（标定板为 10×10 格，每格 76mm），得到其同步的激光距离数据。如图 3-12，利用单目摄像头 vcap2800 与 SICK LMS-291 激光扫描仪构成了基本的测量系统。使激光扫描仪的中心与标定板齐平进行数据采集，图 3-13 显示了我们采集的不同位姿棋盘标定板的图像，同时得到了同步的激光距离数据。



图 3-12 标定实验设备



图 3-13 不同位姿的棋盘标定板

接下来利用采集的数据进行标定实验。通过棋盘标定板图像，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 $[\mathbf{R}, \mathbf{t}]$ 。利用同步采集到的激光距离数据，利用公式 $N \cdot \phi^{-1}(p^f - \Delta) = \|N\|^2$ 解方程求得 Φ 和 Δ （其中， N 通过相机的外参数矩阵 $[\mathbf{R}, \mathbf{t}]$ 求出，对应的激光点 p^f 从数据中读取）。最后，可得实验位置下两传感器的投影变换参数：即旋转矩

阵 Φ 和平移向量 Δ 。

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

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

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的正确性，如图 3-14 所示：



图 3-14 激光点投影变换

3.3.3 训练数据集



图 3-15 人体检测训练样本示例

本文研究的是基于激光深度与图像轮廓融合的人体目标检测。在训练阶段，主要使用正视及侧视的人体作为正例样本。采用 6000 副正面及侧面人体作为正例，5000 副非人体图像作为反例，样本均被归一化为 64×128 像素。图 3-15

表示了本文使用的一些正例和反例样本，具有多样性和差异性。

3.3.4 测试数据集

本文采用ISR-UC-imglidar-sync数据集^[15]，其中包括 2157 视频帧数据及其相应的 2157 组激光距离数据 (.txt格式，每个文件中包括 401 个十进制数据，单位mm)。图 3-16 为部分测试数据。



图 3-16 同步采集的图片及其激光距离数据

3.3.5 检测效果及对比分析

采用激光测距仪搜索 ROI 区域并在图像检测中进行验证的方法，速度比单纯用 HOG+SVM 进行全局检索有了明显的提高，且后续的视觉验证可以将上一部分激光搜索的错误区域排除掉。检测结果如图 3-17 所示，其中(a)为原始图像，(b)为经过聚类分析后的激光点群，(c)为利用激光特征选取的候选目标在图像上的投影，(d)为利用图像检测方法进行确认检测后的结果，可见图中的长度类似的车体片段在图像检测中被排除了。

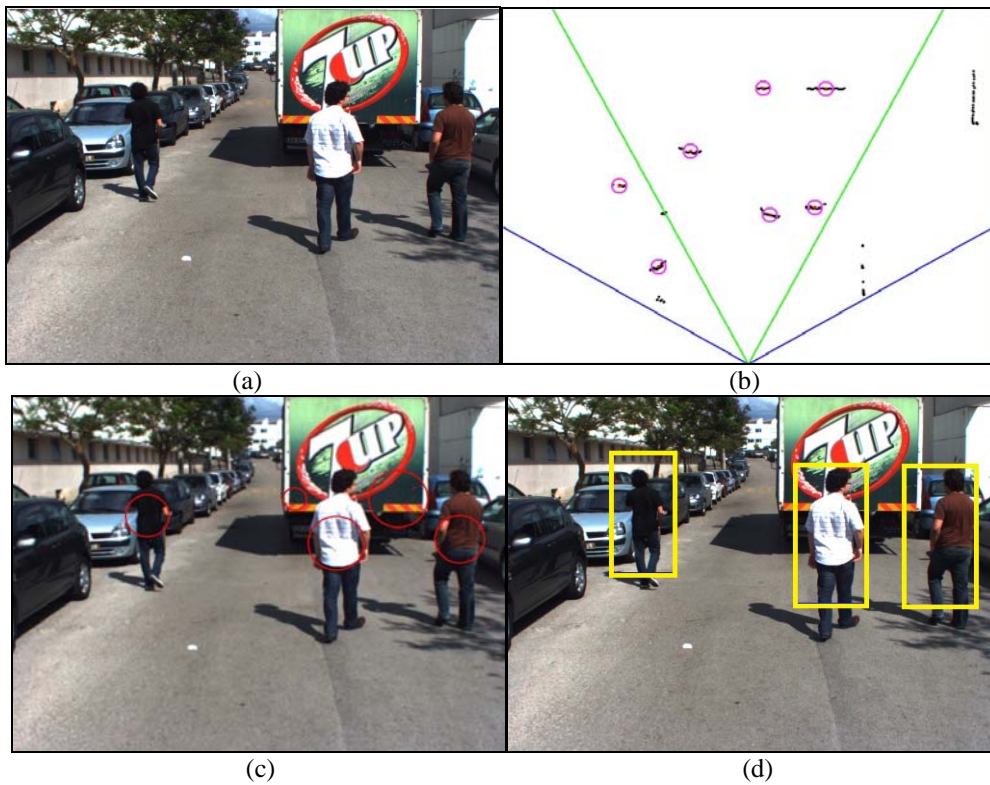


图 3-17 检测结果示意图

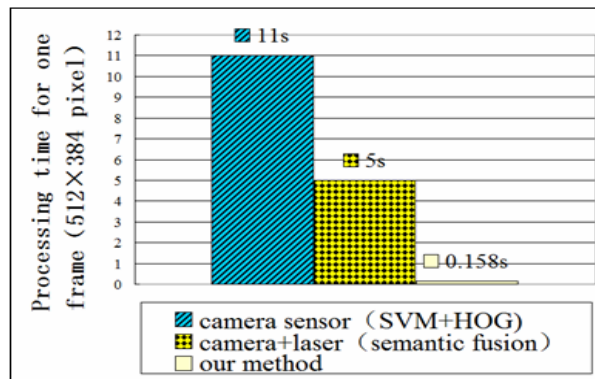


图 3-18 检测速度比较

我们利用激光扫描仪与摄像机信息融合进行行人目标检测，比单独应用视觉方法有了明显的速度提高。图 3-18 显示了本方法与HOG+SVM方法以及文献^[15]中提出的传感器语义融合方法在速度上进行的比较，可见速度的明显提高。

在检测率方面，图 3-19 显示了本方法与文献^[15]中提出的传感器语义融合方法比较的结果。可见我们的方法检测率较高且性能稳定。

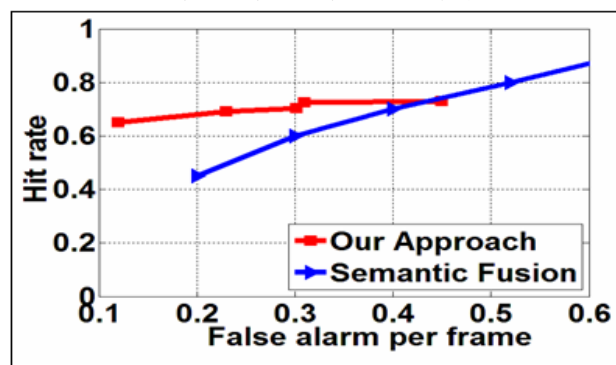


图 3-19 检测率比较

3.4 本章小结

本章对基于激光深度特征与图像轮廓融合的人体目标检测进行了论述和研究。首先论述了激光传感器和视觉传感器的联合在用于目标检测方面的优势，然后阐述了具体特征选取方法以及融合方法。最后，本章对实验过程中需要的数据进行介绍，并且将本文的算法与目前的其他检测方法进行了比较。最后的实验结果及对比分析验证了本文所提出算法的快速性和有效性。

第四章 基于动态规划与信息融合的行人检测算法

4.1 基于多传感器信息串行检测的局限性

在针对行驶环境下的行人防撞检测实验中，对激光扫描仪获得的激光点进行聚类划分后，激光点群可分为如下几类：单个行人或车辆片段、较长的墙面或其他平面物体、零星噪声点以及距离较近的人群组合。如图 4-1 所示为典型的激光点群示意图及其对应图像的投影结果，其中标注序号为 1 号的点群为人体或类似片段物体，2 号点群为连续且平整的车体或墙体片段，3 号为零星噪声点，4 号为距离较近的人-车组合片段。

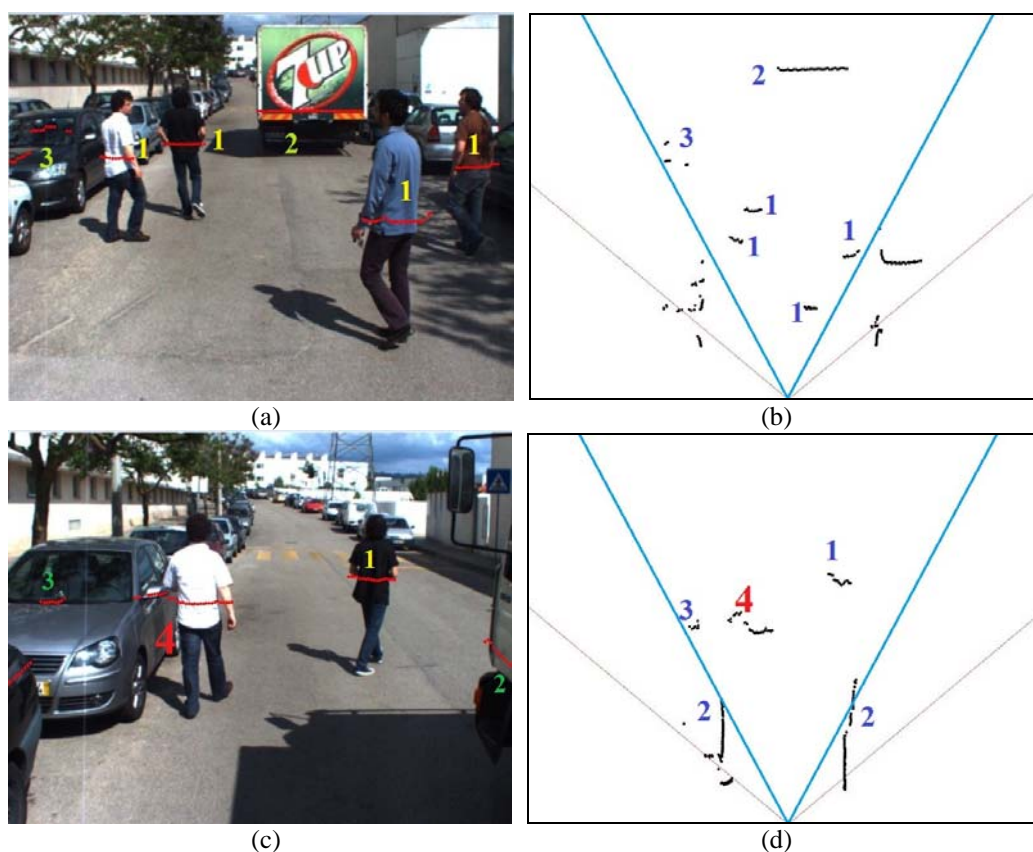


图 4-1 典型的激光点群，其中(b)是(a)图对应的激光扫描平面图，(d)是(c)图对应的激光扫描平面图。

在这几类点群中，通过层次聚类、阈值限制和方向判断，单个行人或车辆片段、较长的墙面或其他平面物体以及零星噪声点均可被准确的划分，便于下

一步投影后的判断。如图 4-2 所示为 4-1(a)经过处理后的结果，经过聚类分析后，较长且平整的车体片段和零星噪声点被排除，如 4-2(a)所示；在此基础上将候选区域交由图像检测进行处理，长度与人体类似的非人体被排除，如 4-2(b)所示。

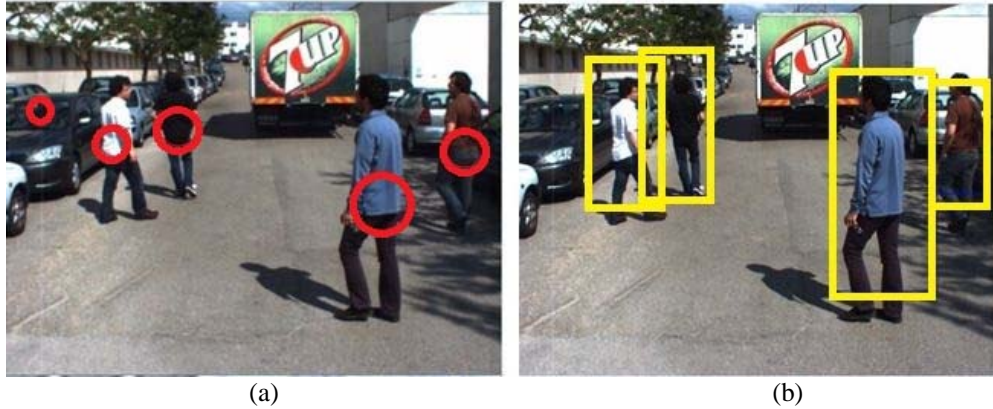


图 4-2 串行检测针对三类点群的有效性

但在处理距离较近的人-人或人-车组合等情况时，则会出现由于聚类错误导致的“候选”漏选或聚类中心的选取发生较大偏差的情况，如图 4-3 所示，使得投影映射后交给图像处理的 ROI 区域丢失或不准确，从而直接影响了检测效果。



图 4-3 串行检测失效情况

4.2 基于动态规划与信息融合的行人检测

上节中基于激光特征与图像轮廓融合的人体检测算法考虑的是通常的情况，即目标检测范围内的各个物体在激光扫描条件下是可分的，从而可以利用层次聚类法找到不同的激光点片段，并根据片段的长度来选择“候选”。但是，

无论是用于聚类时进行分类判别的阈值，还是用于选择“候选”时所用的长度阈值，都是根据经验和试算确定的，并不一定普适。在针对有遮挡下的行人检测中存在着这样的情况：由于距离过近，不同物体间的激光点密集而平均，聚类后也无法区分。这种情况下，激光扫描仪无法为图像检测提供正确的信息使得图像检测无法进行下去，最后导致了由于串行检测而导致的目标丢失。

针对有遮挡情况下的人体检测中多传感器串行检测方法失效的问题，本文提出一种基于动态规划与信息融合的检测方案。

4.2.1 动态规划的基本原理

动态规划（dynamic programming）^[61]是运筹学的一个分支，是一种在数学和计算机科学中使用的解决多阶段决策过程最优化的数学方法。上世纪 40 年代，Richard Bellman 最早使用动态规划这一概念表述，通过遍历寻找最优决策求解问题的求解过程。1953 年，Richard Bellman 将动态规划赋予现代意义，该领域被 IEEE 纳入系统分析和工程中。在“动态规划”（dynamic programming）一词中，programming 是指对生成活动的优化策略，在此意义上，规划意味着找到一个可行的活动计划。动态规划把多阶段决策问题作为研究对象。

■ 多阶段决策问题

根据问题本身的特点，可以将其求解的全过程划分为若干个相互联系的阶段（即把问题划分为多个相互联系的子问题），在它的每一阶段都需要作出决策，并且在一个阶段的决策确定以后再转移到下一阶段。

■ 多阶段决策过程

前一阶段的决策会影响后一个阶段的决策，从而影响整个过程。各个阶段所确定的决策构成一个决策序列。一般来说，由于每一阶段可供选择的决策往往不止一个，因此对于整个过程就会有許多可供选择的策略。

■ 最优策略

若对应于一个策略，可以由一个量化的指标来确定这个策略所对应的活动过程的效果，则不同策略就有各自的效果。在所有可供选择的策略中，对应效果最好的策略成为最优策略。把一个问题划分为若干个相互联系的阶段选取其最优策略，这类问题称为多阶段决策问题。

多阶段决策过程最优化的目标是要达到整个活动过程的总体效果最优。由

于各段决策间有机制的联系着，本段决策的执行将影响到下一段的决策，以致会影响总体效果，因此决策者在每段决策时不应仅考虑本阶段最优，还应考虑对最终目标的影响，从而做出对全局效果最优的决策。解决多阶段决策问题的思路有三种：穷举法、局部最优路径方法和动态规划方法。其中，穷举法通过列举出所有可能发生的方案和结果，再对它们逐一进行比较来求出最优解，此法虽可求出结果，但过于繁琐，费时费力；局部最优路径方法并不顾及全局最优而是选择当前最优路径，此法的结果通常是错误的；动态规划方法寻求解的过程为，首先将问题划分为若干个阶段，每次的选择总是综合后继过程的一并最优进行考虑，在各段所有可能状态的最优后继过程都已求得的情况下，全局的最优路线即可求得。相比穷举法和局部最优法，动态规划属于比较科学有效的方法。

■ 动态规划基本思想

把一个比较复杂的问题分解为一系列同类型的更易求解的子问题，便于计算机应用。整个求解过程分为两个阶段，先按整体最优的思想逆序的求出各个子问题中所有可能状态的最优决策与最优路线值，然后再顺序的求出整个问题的最优策略和最优路线。计算过程中，系统的删减了所有中间非最优的方案组合，从而使计算工作量比穷举法大为减少。

■ 动态规划的基本步骤

1. 将实际问题分为 N 个子问题；
2. 定义状态变量 s_k ，使其既能正确描述过程的状态，又能满足无后效性；
3. 定义决策变量以及各阶段的允许决策集合；
4. 建立状态转移方程；
5. 构造动态规划函数基本方程并求解。

■ 动态规划的模型

$$f_k(s_k) = \text{opt}\{R_k(s_k, p_k(s_k)) + f_{k+1}(s_{k+1})\}, k = 1, 2, \dots, N \quad (4-1)$$

其中， N 表示将系统分为 N 个分段， s_k 为第 k 阶段的状态变量， $p_k(s_k)$ 表示在第 k 阶段的状态为 s_k 时的决策变量， $R_k(s_k, p_k(s_k))$ 表示第 k 子过程的指标函数（也称为其阶段效益）， $f_k(s_k)$ 表示从状态 s_k 触发到达终点的最优值。

4.2.2 基于动态规划与信息融合的行人检测算法

在具有复杂前景的、有遮挡的人体检测情况中，利用激光点群扫描遇到的瓶颈问题是：由于激光点分布比较平均且集中，经过聚类处理后无法区分，导致串行算法的失效。考虑到这种类型点群的分布特点，我们认为可以将问题看成一类多阶段决策的问题：通过将点群划分为不同区段的组合，最后根据各区段的具体得分情况来判断最终的全局最优组合。

本文提出一个基于动态规划与信息融合的检测方法。如图 4-4 所示，首先对密集而平均的激光点群进行分割，将激光点群按照激光点个数进行划分（图中的方框为不同激光区段划分后的投影结果），不同的分割组合形成一个动态规划网络。然后利用 HOG+SVM 对每个分割段进行分析，以获得每个分割段的得分响应，相应的，为了排除由于相邻两个分割段具有物理长度上的明显差异而引起的误差，加入误差惩罚因子进行平衡，如此即形成一种激光信息与图像信息互相反馈求解的机制。每个分割区间是动态规划的子问题，每一种点群划分的组合即为动态规划网络中的一条路线，利用动态规划去求得的最优解，即为点群划分组合的 SVM 得分最高且各个划分区段的物理长度差异不大的点群分割方式。求得最优解的同时，也就得到了最优化的目标检测结果。

如式(4-2)，其中 N 表示每种划分中包含分段的数目， s_i 为第 i 阶段的状态变量 ($i=1,2,\dots,N$)， $\langle w, \phi(x_i) \rangle$ 为每个分段区域内的 SVM 检测得分，定义 μ 为针对物理差异的惩罚因子， $L(s_i) - L(s_{i-1})$ 为相邻的划分之间在物理观测上的差异，变量 F 为动态规划中的任一可行解，使得 F 值最大的划分方法被认为是最优解。在式(4-3)中，利用动态规划方法求解最优划分。

$$F = \frac{1}{N} \left(\sum_{i=1}^N \langle w, \phi(x_i) \rangle - \mu \sum_{i=1}^N (L(s_i) - L(s_{i-1})) \right) \quad (4-2)$$

$$\max_{s_i} (F) \quad (4-3)$$

表 4-1 表示了基于动态规划与信息融合的检测算法。

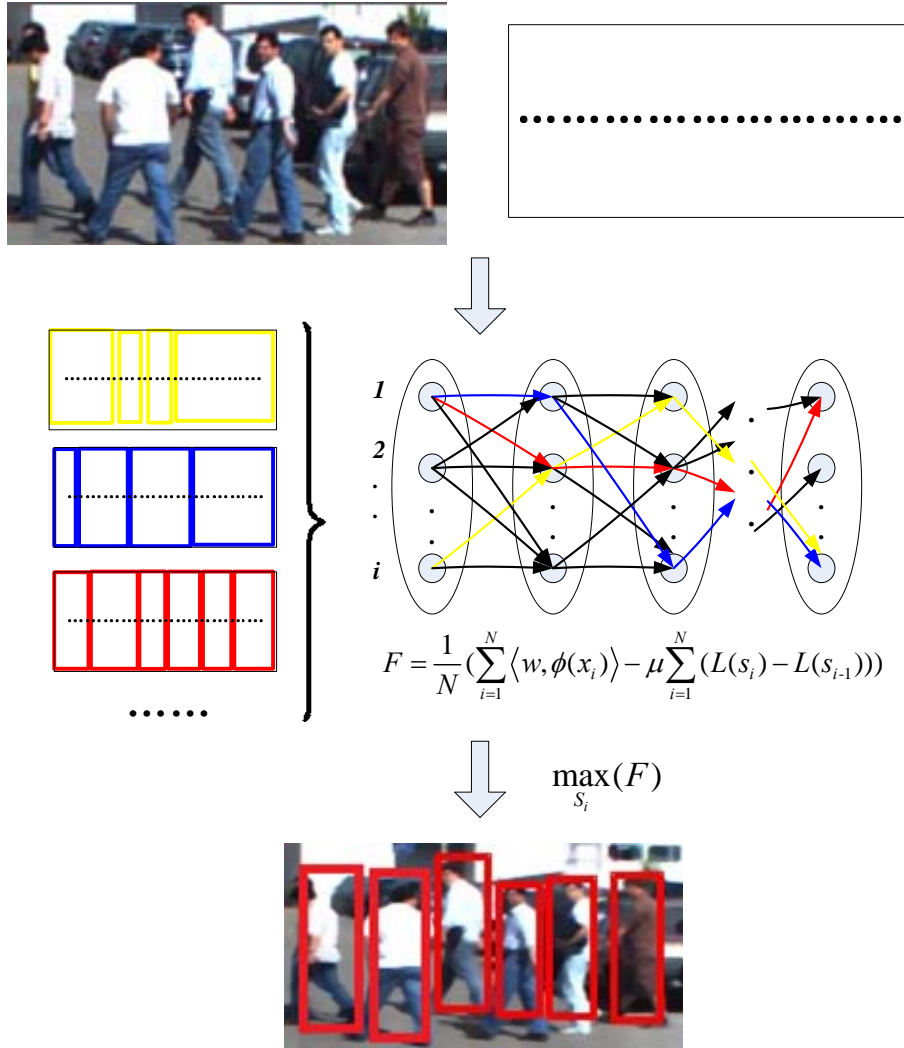


图 4-4 基于动态规划与信息融合的行人检测方法示意图

表 4-1 基于动态规划与信息融合的检测算法

输入: $\{l_z\}_{z=1}^Z$ 扫描一次输出激光数据点

输出: 最优解 $\max_{s_i}(F)$

- 1 激光点聚类, 得到距离相近的激光点群;
- 2 建立激光点群划分, 动态规划网络;
- 3 **for each** i **do**

$$F \leftarrow (\langle w, \phi(x_i) \rangle, L(s_i) - L(s_{i-1}));$$

end

- 4 $\max_{s_i}(F)$

4.3 检测效果

我们提出的改进方法是针对使串行检测失效的有遮挡情况下的行人检测的。图 4-5 是我们利用第三章提出的串行方法和改进方法进行有遮挡情况下行人检测的对比结果图，左列为串行方法检测结果，右列为改进方法检测结果。可见改进方法针对难以区分的激光点群有较好的效果。

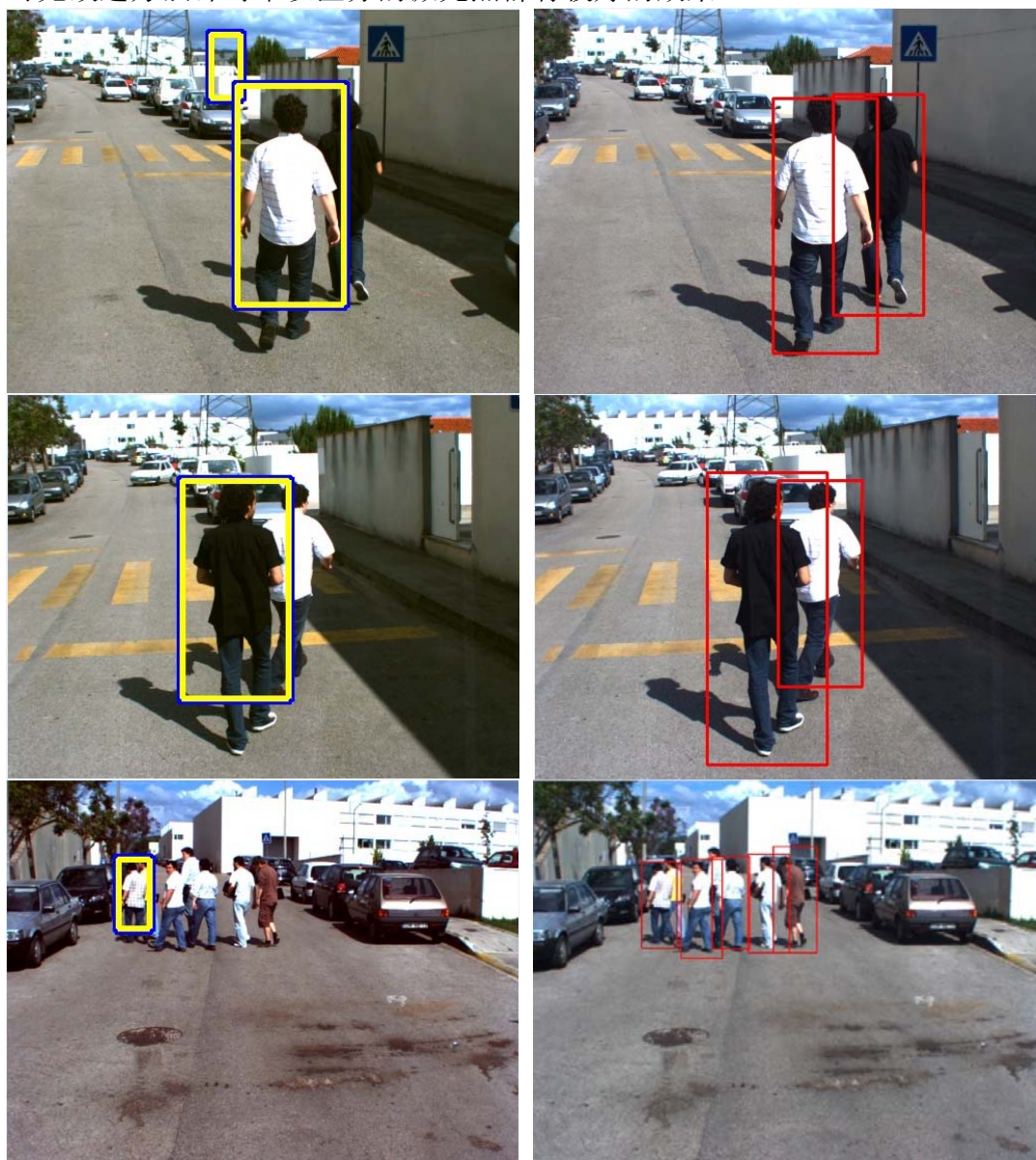


图 4-5 基于动态规划与信息融合的行人检测结果与原串行方法对比图，左列为串行方法，右列为改进方法

4.4 本章小结

本章首先分析了用于防撞检测中可能存在的几种典型的激光点群，分别对

于传感器串行融合检测的有效性进行了分析，接下来针对利用串行方法在有遮挡情况下进行行人检测时存在的缺点，提出了一种基于动态规划与信息融合的复杂前景下行人检测的解决方案。

总结与展望

本文通过融合从激光扫描仪得到的深度数据和从 CCD 摄像机得到的图片信息来实现对行人目标的准确和快速检测。研究内容对于利用行人防撞安全预警来提高车辆运输行驶过程中的安全性有着重要意义，在物流工程领域有着重要应用价值。

本文首先论述了多传感器融合用于行人检测的研究背景和意义及国内外研究现状，综述了多传感器信息融合的研究基础，提出了本文最主要的研究框架；接下来，本文对于多传感器信息融合的基本理论进行综述，为我们提出的方法奠定了理论基础；在第三部分，本文论述了激光传感器和视觉传感器的联合在用于目标检测方面的优势及可行性，详细的论述了应用两种传感器进行信息融合过程中所涉及的具体理论和方法，在此基础上提出了具体的三个检测步骤：候选目标生成，ROI 区域生成和候选目标确认检测。同时，针对利用传感器信息串行方法进行复杂前景的、互有遮挡的行人检测时存在的缺点，本文提出了一种利用动态规划与信息融合方法进行行人检测的解决方案。最后的实验结果及对比分析验证了本文所提出算法的快速性和有效性。

对于无遮挡的行人检测，本文先利用激光扫描仪迅速排除大量无关区域，找到可能区域，再利用图像检测方法对此区域进行确认性的检测，该方法迅速且有效；对于互有遮挡的行人检测，串行方法无法解决问题，本文针对此情况提出了一种有效的解决方案。本文所提出的检测方法虽然在一定程度上满足了需求，但仍存在一些问题，需要进一步的深入研究：

- 1、串行检测方法在目标确认阶段受到具体的分类模型限制，只能单独对一类物体（行人或车辆）进行检测。若模型具有良好的适应性，则该方法可扩展到更多障碍物的检测中。

- 2、针对复杂前景下的行人检测方案还停留在理论阶段，速度也比不上基于传感器信息串行的检测方法，其具体实现还需要进行改进。

在将来的工作中，需要对图像检测阶段的具体模型和算法进行改进，使其具有良好的适应性。同时需对基于动态规划的检测方法进行进一步的研究。

参考文献

- [1] 王荣本, 李斌, 储江伟, 郭克友. 世界智能车辆行驶安全保障技术的研究进展 [J]. 公路交通科技, 2002(19).
- [2] 周芳, 韩立岩. 多传感器信息融合技术综述 [J]. 遥测遥控, 2006(27).
- [3] P Bellutta, R Manduch. Terrain perception for DEMO III [J]. Proceedings of IEEE Intelligent Vehicles Symposium, 2000(10): 326-331.
- [4] 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.
- [5] 刘晓君. 丰田先进安全车 [J]. 世界汽车, 1998(2).
- [6] 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.
- [7] 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.
- [8] 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).
- [9] 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.
- [10] M. Szarvas, U. Sakai, J. Ogata. Real-time pedestrian detection using LIDAR and convolutional neural networks [J]. Proceedings of the IEEE International Symposium on Intelligent Vehicles, 2006:213-218.
- [11] A.Broggi, P. Cerri, S. Ghidoni, P. Grisleri, H. Gi. A new approach to urban pedestrian detection for automatic braking [J]. IEEE Transactions on Intell. Transp. Syst. 2009(10): 594-605.
- [12] A.Broggi, P. Cerri, S. Ghidoni, P. Grisleri, and H. G. Jung. Localization and analysis of critical areas in urban scenarios [J]. IEEE Intell. Vehicles Symp., Eindhoven, 2008:1074-1079.
- [13] G. Alessandretti, A. Broggi, P. Cerri. Vehicle and guard rail detection using

- radar and vision data fusion [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2007: 95-105.
- [14] L. Spinello, R. Triebel, R. Siegwart. A trained system for multimodal perception in urban environments [C]. *Proceedings of the Workshop on Safe Navigation in Open and Dynamic Environment: Applications to Autonomous Vehicles (ICRA)*, 2009.
- [15] L. Oliveira, U. Nunes. Semantic fusion of laser and vision in pedestrian detection [J]. *Pattern Recognition*, 2010:3648-3659.
- [16] L. Oliveira, U. Nunes. Context-aware pedestrian detection using lidar [J]. *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2010, pp. 773–778.
- [17] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2001.
- [18] C. Papageorgiou, T. Evgeniou, and T. Poggio. A trainable pedestrian detection system [C]. *IEEE Transactions on Intelligent Vehicles*, 1998(10): 15-33.
- [19] N. Dalal, Bill Triggs. Histograms of oriented gradients for human detection [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
- [20] F.j Kong, Q.x. Ye, and J.b Jiao. On-road vehicle detection using histograms of multi-scale orientations [C]. *Proceedings of the IEEE Youth Conference on Information, Computing and Telecommunications*, 2009.
- [21] 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.
- [22] 吴艳, 杨万海, 李明. 多光谱与高分辨率图像融合算法研究 [J]. *光子学报*, 2003, 32(2):174-178.
- [23] A. Goshtasby, S. Nikolov. Image fusion: Advances in the state of the art [J]. *Information Fusion*, 2007, 8(2):114-118.
- [24] A. Gholipour, N. Kehtarnavaz, R. Briggs, M. Devous, and K. Gopinath. Brain functional localization: A survey of image registration techniques [J]. *IEEE*

- Transactions on Medical Imaging*, 2007, 26(4):427-451.
- [25] H. Lester, S. R. Arridge. A survey of hierarchical non-linear medical image registration [J]. *Pattern Recognition*. 1999, 32:129-149.
- [26] J. P. Heather, M. I. Smith. Multimodal image registration with applications to image fusion[C]. *The 8th International Conference of Information Fusion*, 2005,1: 372-379.
- [27] 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
- [28] Dave L, Hall and James Llinas. Introduction to multi-sensor data fusion [J]. *Proceedings of the IEEE*, 85 (1), 6-23
- [29] Glorol L, Calhoun, and Mark H. Synthetic vision system for improving unmanned aerial vehicle operator situation awareness [J]. *Proceedings of SPIE Enhanced and Synthetic Vision*, 2005: 219-230.
- [30] M. S. Choi, and W. Y. Kim. A novel two stage template matching for rotation and illumination invariance [J]. *Pattern Recognition*, 2002, 35(1): 119-129.
- [31] E. Fishler. Spatial diversity in radars-models and detection performance [J]. *Proceedings of IEEE Trans. on Signal Processing*. 2006, 54(3): 823-838.
- [32] L. J. Kramer, S. P. Williams, and R.E. Bailey. Simulation evaluation of synthetic vision as an enabling technology for equivalent visual operations [J]. *Proceedings of SPIE Defense and Security Symposium*, 2008.
- [33] M. Janic. Modeling the capacity of closely-spaced parallel runways using innovative approach procedures [J]. *Transportation Research Part C-Emerging Technologies*, 2008.
- [34] Alspach D L, and Sorenson H W. Nonlinear Bayesian estimation using Gaussian sum approximations [J]. *IEEE Transactions on Automatic Control*., 1972,17(4):439-448.
- [35] Molnar K J, and Modestino J W. Application of the EM algorithm for the multi-target/multi-sensor tracking problem [J]. *IEEE Transactions on Signal Processing* ,1998,46(1):115-129.
- [36] Murphy, R.R. Dempster-Shafer theory for sensor fusion in autonomous

- mobile robots [J]. IEEE Transactions on Robotics and Automation, 1998,14(2):197-206.
- [37] Goutsias John, and Mahler Ronald P. S. Random sets : theory and applications [M]. New York: SpringerVerlqg, 1997.
- [38] J Manyika, H DurrantWhyte. Data fusion and sensor management: a decentralized information-theoretic approach [M]. New York: Ellis Horwood, 1994.
- [39] Barron A, Rissanen J, Bin Yu . The minimum description length principle in coding and modeling [J]. IEEE Transactions on Information Theory, 1998:44(6):2743-2760.
- [40] Nelson, C.L, Fitzgerald, D.S. Sensor fusion for intelligent alarm analysis [J]. IEEE Transactions on Aerospace and Electronic Systems, 1997, 12(9):18-24.
- [41] Valet,L ,Mauris G and et al. A statistical overview of recent L iterature in information fusion [G]. ISIF2000:MoC3-22-Moc3-29.
- [42] 王璇 ,李春升 ,周荫清. 多传感器信息融合技术 [J]. 北京航空航天大学学报 ,1994.
- [43] 段新生. 证据理论与决策. 人工智能 [M]. 北京,中国人民大学出版社 , 1993.
- [44] Barnett J A. Computational methods for a mathematical theory of evidence [C]. Proceedings of the 7th International Joint Conference on Artificial Intelligence. Vancouver, B. C. , Canada. 1981:868-875
- [45] Nauck D, and Kruse R. Obtaining interpretable fuzzy classification rules from medical data [J], Artificial Intelligent in Medicine, 1999, 16 (2): 149-169.
- [46] Doyle R S. and Harris, C. J. Multi-sensor data fusion for helicopter guidance using neuro-fuzzy estimation algorithms[C]. IGARSS'1995: 1392-1397.
- [47] Zeidenberg and Matthew. Neural networks in artificial intelligence [M]. Ellis Horwood Limited, 1990.
- [48] V. Vapnik. The nature of statistical learning theory. Springer Verlag, 1995.
- [49] 王水清. 高速公路车辆检测系统 [J]. 测控技术, 2001(20).

- [50] Sick_LMS-200_激光测距仪文档 [Z].
- [51] 胡雷芳. 五种常用系统聚类分析方法及其比较 [J]. 浙江统计, 2007(4).
- [52] Rui Xu. Survey of clustering algorithms [J]. IEEE Transactions on Neural Networks. 2005, 16(3):645-678.
- [53] 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.
- [54] G. Bradski, A. Kaehler. Learning OpenCV(于仕琪译) [M]. 北京: 清华大学出版社, 2009.
- [55] 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.
- [56] Z.h. Sun, G. Bebis, R. Miler. On-Road vehicle detection: A Review [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(5).
- [57] 金立生, 王荣本, Bart Van Arem, 郭烈. 先进驾驶员辅助系统中的车辆探测研究综述 [J]. 汽车工程, 2007, 29(2).
- [58] 张学工. 关于统计学习理论与支持向量机 [J]. 自动化学报, 2000(26):32-42.
- [59] 马颂德, 张正友. 计算机视觉—计算理论与算法基础 [M]. 北京: 科学出版社, 1998:52-59.
- [60] Q. Zhang, R. Pless. Extrinsic calibration of a camera and laser range finder (improves camera calibration) [C]. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004(3):2306–2310.
- [61] 吴祈宗. 运筹学 [M]. 北京: 机械工业出版社, 2010.

个人简历及发表文章目录

个人简历

武博 女 汉族 党员

- 2005年9月至2009年7月 河北工业大学 学士
- 2009年9月至2012年7月 中国科学院研究生院 硕士

已发表文章目录

- **B. Wu, J. Liang, Q. Ye, Z. Han and J. Jiao**[✉]. Fast pedestrian detection with laser and image data fusion [C]. *Proceedings of the 6th International Conference on Image and Graphics (ICIG)*, 2011. (EI)

致 谢

在中国科学院研究生院攻读硕士学位三年的学习生活中，我经历了诸多坎坷，也付出了艰辛的努力，同时也得到了很大的收获。在毕业论文完成之际，由衷地感谢这三年来曾经给予我无数帮助的老师、同学、朋友和家人。

本课题的研究工作是在焦建彬教授和叶齐祥副教授的悉心指导下完成的。首先，我要感谢导师叶齐祥老师在我攻读硕士学位期间从理论和实验方面给予的大量的、极其有益的指导和他在我论文的撰写和修改中倾注的心血，感谢焦建彬教授在我的学习中对我每一点进步的指导和鼓励以及在生活和思想上对我的关心和帮助。两位恩师在科研上精益求精，在学术上认真严谨，他们的科学精神令人敬佩。

其次，感谢韩振军师兄在理论学习和项目开发过程中给我提供的耐心的引导和帮助，感谢他在我论文中倾注的心血。感谢师弟梁吉祥和高山，他们在实验阶段给予了我很大的帮助和支持，使实验进度能够顺利进行。

感谢我的好友李策、纪颖夏、高文、小汪以及实验室的刘一飞等同学，三年的科研生活中大家在忐忑时互相鼓励，在瓶颈时献计献策，在茫然时互相提点。这些一起度过的日子在我记忆中永远不会褪色。

感谢我的父母，他们给了我巨大的、无私的爱和永远无条件支持，永远是最坚强的后盾和避风的港湾，愿他们永远平安喜乐。

感谢参加开题及中期评阅的各位老师和专家们，他们丰富的经验和无私的工作对论文方向和研究进度的把握和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各位老师。

武博

2012年4月