

车辆辅助驾驶系统中基于计算机视觉的行人检测研究综述

贾慧星¹ 章毓晋¹

摘要 基于计算机视觉的行人检测由于其在车辆辅助驾驶系统中的重要应用价值成为当前计算机视觉和智能车辆领域最为活跃的研究课题之一. 其核心是利用安装在运动车辆上的摄像机检测行人, 从而估计出潜在的危险以便采取策略保护行人. 本文在对这一问题存在的困难进行分析的基础上, 对相关文献进行综述. 基于视觉的行人检测系统一般包括两个模块: 感兴趣区分割和目标识别, 本文介绍了这两个模块所采用的一些典型方法, 分析了每种方法的原理和优缺点. 最后对性能评估和未来的研究方向等一系列关键问题给予了介绍.

关键词 行人检测, 车辆辅助驾驶系统, 感兴趣区分割, 目标识别

中图分类号 TP391.41

A Survey of Computer Vision Based Pedestrian Detection for Driver Assistance Systems

JIA Hui-Xing ZHANG Yu-Jin

Abstract Computer vision based pedestrian detection has become one of the hottest topics in the domain of computer vision and intelligent vehicle because of its potential applications in driver assistance systems. It aims at detecting pedestrians appearing ahead of the vehicle using a vehicle-mounted camera, so as to assess the danger and take actions to protect pedestrians in case of danger. In this paper, we give detailed analysis of the difficulties lying in the problem and review most of the literature. A typical pedestrian detection system includes two modules: regions of interest (ROIs) segmentation and object recognition. This paper introduces the principle of typical methods of the two modules and analyzes their respective pros and cons. Finally, we give detailed analysis of performance evaluation and propose some research directions.

Key words Pedestrian detection, driver assistance system, ROIs segmentation, object recognition

1 引言

车辆辅助驾驶系统中基于计算机视觉的行人检测是指利用安装在运动车辆上的摄像机获取车辆前面的视频信息, 然后从视频序列中检测出行人的位置. 由于它在行人安全方面的巨大应用前景, 成为智能车辆、计算机视觉和模式识别领域的前沿研究课题. 欧盟从 2000 年到 2005 年连续资助了 PROTECTOR^[1] 和 SAVE-U^[2] 项目, 开发了两个以计算机视觉为核心的行人检测系统; 意大利 Parma^[3] 大学开发的 ARGO 智能车也包括一个行人检测模块; 以色列的 MobilEye^[4] 公司开发了芯

片级的行人检测系统; 日本本田汽车公司^[5] 开发了基于红外摄像机的行人检测系统; 国外的大学如 CMU^[6]、MIT^[7, 8] 和国内的西安交通大学^[9]、清华大学^[10] 也在该领域做了许多研究工作.

车辆辅助驾驶系统中基于计算机视觉的行人检测属于计算机视觉中人体运动分析的研究范畴, 其主要任务是在运动摄像机下快速准确地检测行人. 本文主要针对这一特定领域对相关的文献进行综述, 重点分析常用方法的原理和优缺点, 以期对相关的科技人员起到指导作用. 对监控系统和体育运动分析领域中人体检测感兴趣的读者可以参考综述文献^[11~14].

行人检测除了具有一般人体检测具有的服饰变化、姿态变化等难点外, 由于其特定的应用领域还具有以下难点: 摄像机是运动的, 这样广泛应用于智能监控领域中检测动态目标的方法便不能直接使用; 行人检测面临的是一个开放的环境, 要考虑不同的路况、天气和光线变化, 对算法的鲁棒性提出了很高的要求; 实时性是系统必须满足的要求, 这

收稿日期 2006-3-14 收修改稿日期 2006-6-17

Received March 14, 2006; in revised form June 17, 2006

国家自然科学基金 (60573148), 教育部高等学校博士学科点专项科研基金 (20060003102) 资助

Supported by National Natural Science Foundation of P. R. China (60573148), Specialized Research Fund for the Doctoral Program of Higher Education (20060003102)

1. 清华大学电子工程系 北京 100084

1. Department of Electronic Engineering, Tsinghua University, Beijing 100084

DOI: 10.1360/aas-007-0084

表 1 感兴趣区分割方法
Table 1 Methods for ROIs segmentation

	原理	优点	缺点	典型文献
基于运动的方法	通过检测运动区域进行分割	受行人姿态的影响较小	检测不到静止行人	[15]
基于距离的方法	通过测量目标到汽车的距离进行分割	受行人姿态、颜色和光照的影响较小	需要额外的测距设备,增加了系统的造价和复杂性	雷达[2], 立体视觉[1, 6, 16, 17]
基于图像特征的方法	通过分析图像上的一些典型的行人特征进行分割	可以对得到的图像数据直接操作	很难定义出比较鲁棒且易于检测的特征	边缘 [3]、熵 [18]、纹理 [4]、“热点”[5, 19]
基于摄像机参数的方法	通过摄像机的几何关系确定搜索区域	可以只对感兴趣的区域处理	需要对系统参数进行标定,受车辆振动的影响	[3, 4, 19]

就要求采用的图像处理算法不能太复杂。

由上面的论述可以知道,系统的鲁棒性和实时性构成了一个矛盾,为了解决这种矛盾,现有的行人检测系统一般包括两个模块:感兴趣区 (Regions of Interest, ROIs) 分割和目标识别。ROIs 分割的目的是从图像中提取可能包含行人的窗口区域作进一步验证,以避免穷尽搜索,提高系统的速度。目标识别是行人检测系统的核心,它对得到的 ROIs 进行验证,以判断其中是否包含行人,它的性能决定了整个系统可以达到的精度和鲁棒性。本文首先讨论这两个模块可以采用的方法和各种方法的优缺点,最后介绍一些现有的行人检测系统的性能并对未来的发展趋势做出详细的阐述。

2 ROIs 分割

根据分割所用的信息,可将 ROIs 分割的方法分为基于运动、基于距离、基于图像特征和基于摄像机参数四种方法。各种方法的基本原理和优缺点见表 1。在实际系统中,各种方法通常结合在一起使用,彼此互补。

2.1 基于运动的方法

基于运动的方法通过检测场景中的运动区域来得到 ROIs。由于摄像机是运动的,所以要想应用该方法,必须首先补偿车辆自身的运动。例如 Stein^[15]等首先利用光流法估计车的运动,补偿车辆的运动后利用时间差分从而得到运动物体。基于运动的方法的优点是不受姿态的影响,比较鲁棒。缺点是只能检测运动的行人。所以目前主要应用在智能监控领域,在车辆辅助驾驶系统中往往只作为一个辅助的手段^[4]。

2.2 基于距离的方法

基于距离的方法通过测量目标到汽车的距离来得到 ROIs。可以用来测距的传感器主要包括雷达^[2]和立体视觉^[1, 6, 16, 17]。当雷达与摄像机结合时,必须要考虑各个传感器间的时间同步和空间同步问题。

而立体视觉本身就是图像,便于后期的图像处理算法直接进行处理,所以目前大多数基于视觉的检测系统都采用立体视觉来分割出 ROIs。基于距离的方法的优点是比较鲁棒,受光照和行人外貌的影响较小。缺点是需要额外的测距设备,从而增加了系统的造价和复杂度。

2.3 基于图像特征的方法

基于图像特征的方法指通过检测与行人相关的图像特征从而得到 ROIs。对于可见光图像来说,常用的特征包括垂直边缘^[3]、局部区域的熵^[18]和纹理^[4]等。对于红外图像来说,主要根据人体尤其是人脸的温度比周围环境温度较高这一特征,通过检测一些“热点”^[5, 19] (Hot spot) 来得到 ROIs。基于图像特征的方法的优点是直接利用了图像信息,不需要额外的传感器。缺点是由于环境的复杂多变,很难定义出对每种场景都适用的特征。对这种方法的要求是特征检测算法不能太复杂,整个系统的速度不应该低于对所有窗口进行穷尽识别的速度。目前采用这种方法进行分割的算法通常和后面的目标识别算法结合在一起,从而形成了一个单步检测算法。例如 Viola^[20, 21]等设计的级联分类器的前几级就可以视为一个 ROIs 分割算法,后面的几个分类器被视为识别算法。

2.4 基于摄像机参数的方法

摄像机的安装位置和摄像机参数也是一个很重要的考虑因素。它对行人在图像上出现的位置和每个位置上目标的大小给出了很多限制,合理利用这些限制可以大大地缩小搜索空间^[3, 4, 19]。该方法的优点是只需要处理可能发生危险的区域。缺点是需要对摄像机参数进行标定,并且受车辆振动的影响。

3 目标识别

根据利用的信息的不同,目标识别可以分为基于运动的识别和基于形状的识别两种方法^[22],各种方法的优缺点和相关文献可以参考表 2。在一个实

表 2 行人目标识别方法
Table 2 Methods for pedestrian recognition

	基于运动的方法	基于形状的方法		
		基于明确人体模型的方法	基于模板匹配的方法	基于统计分类的方法
原理	通过分析行人步态 (Gait) 的周期性来识别行人	构造明确的人体参数模型来表示行人	通过模板表示行人	通过分类器对行人进行识别
优点	受颜色、光照的影响较小, 比较鲁棒	具有明确的模型, 方便处理姿态和遮挡问题	计算方法简单	不需要人工设置大量参数、比较鲁棒
缺点	只能识别运动行人, 需要多帧, 影响实时性	建模和求解比较复杂	需要很多模板对付姿态问题, 匹配比较耗时间	需要大量的训练数据
典型文献	[18, 23~25]	[10, 19, 26~28]	[1, 3, 29~31]	NN [6, 32~34]、SVM [7~9, 36~39]、Adaboost [4, 40,42]

际应用系统中, 这两种方法通常结合在一起使用^[4].

3.1 基于运动的方法

基于运动的识别方法指通过分析人运动时的步态 (Gait) 特征来识别行人. 人体的步态具有特定的周期性, 通过分析图像序列的周期性, 然后与行人步态的周期性的模式相比较, 就可以识别出行人. 基于运动的识别的好处是避免了人的纹理和光线变化的影响, 缺点是只能识别运动的行人, 并且要分析很多帧才能给出判决结果, 影响系统的实时性. 周期性的检测方法包括傅立叶变换和神经网络两种方法. Cutler 和 Davis^[23] 首先计算不同时间间隔图像之间的相关性, 然后对得到的相关信号利用短时傅立叶变换进行分析, 进而识别出行人. Wöhler^[24] 等通过一个自适应时间延迟神经网络 (Adaptive time delay neural network) 对图像序列进行分析, 从而判断出该序列是否是人体的运动序列. 以上两篇文献都是对行人的全身进行分析, 由于行走的时候双腿的周期性更加明显, 所以还可以通过分析双腿的周期性来识别行人^[18, 25].

3.2 基于形状的方法

基于形状的识别方法指通过分析目标的灰度、边缘和纹理信息来对目标进行识别. 基于形状的方法的优点是可以检测出静止的行人, 缺点是容易产生大量的“虚警”(False positives). 如表 2 所示, 可将基于形状的行人识别方法分为基于明确人体模型 (Explicit human model) 的方法、基于模板匹配 (Template matching) 的方法和基于统计分类 (Statistical classification) 的方法.

3.2.1 基于明确人体模型的方法

基于明确人体模型的方法指根据人体结构的知识, 构造一个明确的 2D 或 3D 参数模型, 通过提取图像的底层特征来求解模型, 从而识别行人. 这种方法的优点是具有明确的模型, 可以处理遮挡问题, 并且可以推断出人体的姿态. 缺点是模型比较

难构建, 模型求解也比较复杂. 该方法能否成功依赖于 ROIs 分割的准确度. Zhao^[26] 针对正面和侧面人体分别建立了一个 2D 的平移-旋转-尺度不变的人体概率模型, 该模型不但包括人体的整体形状和各个部件的形状, 而且包括各个部件的大小和空间位置等信息, 这些信息为算法的求解提供了限制条件, 可以根据这个模型推断出被遮挡或者漏掉的部分. 采用 2D 模型的还有 Sun^[10] 和 Yasuno^[27] 等人, Sun 等将人体分成头部、躯干和左右腿四个部分, 而 Yasuno 等将人体分为头部和躯干两个部分. 相对于 2D 模型, 3D 模型在处理姿态、视角变化和遮挡问题上有比较大的优势. 在 Bertozzi^[19, 28] 等人开发的利用红外图像检测行人的系统中, 针对不同的姿态和行人的衣着, 建立了一系列 3D 人体模型来对红外图像中的行人进行识别.

3.2.2 基于模板匹配的方法

基于模板匹配的方法通过存储一些灰度或者轮廓模板来表示行人, 识别的时候只需要度量模板与输入窗口的距离就可以识别行人. 基于模板匹配的算法的优点是计算简单, 缺点是由于行人姿态的复杂性, 很难构造出足够的模板以处理不同的姿态. 基于模板匹配的最典型的算法是由 Gavrilu^[29] 提出的基于轮廓的分层匹配算法, 该算法已经应用于 PROTECTOR^[1] 项目. 为了解决行人姿态的问题, Gavrilu 构造了将近 2500 个轮廓模板对行人进行匹配, 从而识别出行人. 为了解决众多模板引起的速度下降问题, Gavrilu 采用了由粗到细的分层搜索策略以加快搜索速度. 匹配的时候通过计算模板与待检测窗口的距离变换 (Distance transform, DT) 图像之间的 Chamfer^[30] 距离来度量两者之间的距离. 采用 DT 图像而不是原始图像计算距离的好处是, 得到的距离测度是模板变换参数的平滑函数, 方便了快速搜索. 基于轮廓的 Chamfer 距离的识别方法还被成功地应用于 [31]. 除了对人的全身通过模板匹

配进行识别外, 对于人体的局部部件同样也可以采用模板匹配的方法进行识别. 例如, Broggi^[3] 等人利用不同大小的二值图像模板来对人头和肩部进行建模, 通过将输入图像的边缘图像与该二值模板进行比较从而识别行人, 该方法被用到意大利 Parma 大学开发的 ARGO 智能车中.

3.2.3 基于统计分类的方法

基于统计分类的方法通过机器学习从一系列训练数据中学习得到一个分类器, 用该分类器来表示行人, 然后利用该分类器对输入窗口进行识别. 基于统计分类的方法的优点是比较鲁棒, 缺点是需要很多训练数据, 并且很难解决姿态和遮挡的问题. 基于统计模型的方法主要包括两个步骤: 特征提取和分类器设计. 特征提取的目的是降低数据的维数, 得到能反映模式本质属性的特征, 方便后面的分类; 分类器设计属于机器学习领域的范畴, 其目的是得到一个计算复杂度较低, 并且推广性较好的分类器. 针对行人识别问题, 可根据分类器的设计方法将现有的基于统计分类的方法分为基于神经网络 (NN) 的方法, 基于支持向量机 (SVM) 的方法和基于 Adaboost 的方法.

神经网络通过学习的方法将模式的特征隐含在一系列的网路参数之中, 可以描述极为复杂的模式, 已经成功地应用在模式识别领域中的光栅字符识别和人脸检测上, 它在行人的检测上也有广泛的应用^[6, 32~34]. Zhao^[6] 采用图像梯度的幅值作为特征, 采用三层的前馈神经网络作为分类器来识别行人, 采用梯度而不是原始灰度图像或者二值图像的原因是为了消除光照的影响, 同时避免二值化时阈值的选择问题. Szarvas^[33] 利用卷积神经网络 (Convolutional neural network, CNN) 来识别行人, CNN 直接利用原始图像灰度值作为输入, 将特征的提取过程作为隐含节点包含在网络的结构中, 通过学习算法自动选取最佳的分类特征. Franke^[34] 等人直接利用原始图像灰度作为神经网络的输入特征.

SVM 是 Vapnik^[35] 提出的基于结构风险最小化原理 (Structural risk minimization principle, SRM) 的统计学习理论, 比基于经验风险最小化 (Empirical risk minimization principle, ERM) 的神经网络方法具有更好的泛化能力. 最早提出利用 SVM 进行行人检测的是 Oren^[7] 等人, Oren 利用 Haar 小波特征, 结合 SVM 算法构造了一个静态图片上的行人检测系统. 该检测算法首先将原始图像灰度空间转换到过完备 (Over-complete) 的 Haar 小波系数空间, 相对于原始灰度特征空间, Haar 小波系数空间可以更有效地表征行人, 从而为利用 SVM 分类器进行分类打下良好的基础. 基于 Haar 小波

特征和 SVM 分类的行人识别方法引起了广大学者的兴趣, 出现了许多改进的算法^[8, 36, 37]. Mohan^[8] 等人提出了基于部件 (Component-based) 的识别策略来解决人体的姿态问题. Grubb^[36] 设计了两个 SVM 分类器, 一个针对正面和背面图像, 另外一个针对侧面图像, 然后将这两个分类器的判决结果进行融合. Oren 的方法同样可以直接应用到红外图像里行人的识别上^[37]. 对于红外图像, 由于人体和背景的灰度差别比较明显, 图像灰度值也可以直接作为 SVM 算法的输入^[5, 16]. 除此之外, 形状描述符特征^[38]、Gabor^[9] 特征、梯度方向的直方图特征^[39]与 SVM 的结合也广泛地应用于现有的行人检测系统中.

Adaboost 是一种分类器组合的策略, 它的目的是将一些弱分类器组合成一个强分类器. Adaboost 得到的分类器具有较好的推广性能, 现在广泛地应用于模式识别和计算机视觉领域. 最早将 Adaboost 应用到计算机视觉领域的是 Viola 等人, 利用矩形特征、Adaboost 算法和级联分类器成功地实现了第一个实时人脸检测系统^[20]. 近几年的人脸检测算法几乎都是这一检测方法的改进. Viola 同样将该方法应用到了监控系统中行人的检测上^[21], 该方法同时利用两帧的信息, 利用一系列矩形模板提取外貌和运动信息, 从而实现了监控系统中的低分辨率的行人的检测. Abramson^[40] 将 Viola 的方法直接应用到了汽车辅助系统中的行人识别上. 除了方便快速计算的矩形特征外, SIFT^[41] (Scale invariant feature transform) 特征由于其尺度不变的良好特性, 它与 Adaboost 的结合也越来越受到学者的重视. Shashua^[4] 等给出了车辆辅助驾驶系统中的一个系统的行人检测方案, 它将识别阶段分成两个部分, 在单帧识别阶段利用 Adaboost 训练得到的分类器对行人进行识别, 在多帧识别阶段通过分析行人的运动信息来进行综合的判决. 在使用 Adaboost 训练分类器时, 利用了类似 SIFT 特征的梯度方向的直方图特征, 为了解决行人的姿态问题, 该方法根据姿态的不同, 手工地将训练集分成了不同的子集, 从而大大降低了类内的变化. 基于 Adaboost 的检测方法同样可以结合分块检测策略以降低类内变化, Mikolajczyk^[42] 等人将人体分成七个部分, 针对每一个部件利用类 SIFT 特征和 Adaboost 建立起一个检测器, 然后将检测结果利用概率图模型进行融合从而识别站立的行人.

4 现有的系统和性能

4.1 评价方法

基于计算机视觉的行人检测的最难的问题是不

表 3 典型行人检测系统的实验结果
Table 3 Experiments results of typical existing pedestrian detection systems

作者	ROIs 分割	目标识别	实验结果
Broggi ^[3]	利用摄像机参数和竖直边缘的对称性	利用人头的模板匹配	仅仅给出了几张图片作例子，无详细测试结果
Gavrila ^[29]	无	分层模板匹配和 RBF 分类器	基于检测的评价方法：用 900 张彼此无遮挡的图片做测试，正确检测率为 45%~75%，虚警率不详
Zhao ^[6]	立体视觉	利用梯度图像的神经网络分类器	基于分类的评价方法：用 8400 个窗口作评价，正确率为 85.2%，虚警率为 3.1%
Shashua ^[4]	利用纹理特征	基于 SIFT 特征的 Adaboost 分类器	对于向马路内运动的行人的检测率为 96%，在 5 个小时的驾驶过程中只出现一个虚警。
Oren ^[7]	无	基于 Haar 特征的 SVM	基于分类的评价方法：检测率为 69.6%，虚警率为 1:15000
Mohan ^[8]	无	基于 Haar 特征的 SVM，将行人分块	基于分类的评价方法：检测率为 90%，虚警率为 1:10000

同的算法的性能评价，由于评价数据和评价方法的不同，很难比较出不同算法的优劣。目前文献中的评价方法主要分为两类，基于分类器的评价和基于检测的评价。

基于分类器的评价针对“目标识别”这一步骤中采用的分类器的性能进行评价，正面测试样本是一系列和训练样本同样大小的包含行人的窗口，正确率一般用正确分类的窗口数量和总的窗口数量的比值来表示；负面测试样本是一系列不包含行人的窗口，虚警率用错误分类的窗口数量和总的窗口数量的比值来表示。基于分类器的评价方法的优点是简单、直接。缺点是这不是针对行人检测问题本身的评价，很难评价出一个行人检测系统的整体性能。

基于检测的评价是对行人检测问题本身的评价。测试数据一般是一个图片集或者视频序列，人工对行人在图像或视频中的位置进行标注。评价时将检测到的行人的位置和标注的位置进行比较，如果误差小于一个阈值，则判决为正确的检测结果，否则视为一个虚警。正确率被定义为正确检测到的目标个数和手工标注的目标个数的比值；虚警率被定义为一个数据集上出现的总的虚警的个数或者平均出现一个虚警个数的帧数。该评价方法的优点是是对检测问题本身的评价，直接评价了 ROIs 分割和目标识别两者结合在一起的性能；缺点是由于手工标注的主观性和判决正确检测结果阈值设置的不同，使得很难比较各个系统的优劣。

以上评价的方法仅仅是从模式识别的角度进行的评价，如果从行人检测的目的即防撞报警的角度出发，性能的评价更加复杂。例如，针对不同的行人对安全的影响，Shashua^[4] 等人将行人分为横穿马路的行人、静止的行人和沿着马路走的行人，针对这三种情况分别给出检测结果。Gavrila^[1] 在 PROTECTOR 项目的最终测试中为建立一个系统级的测试评价方法作出了尝试。

除了评价方法的不同外，各种算法用来做测试的数据集也不完全相同，使得算法的评价十分困难，建立一个公共的训练和测试数据库是该领域必须解决的一个问题。

4.2 典型系统的性能

虽然行人检测的文献很多，但给出详细实验结果的并不太多，并且彼此采用的数据集和评价方法的差别很大，使得很难进行公平的比较。但为了对目前的系统的性能有一个直观的认识，表 3 给出了一些典型系统的实验结果和所采用的方法。其中的大部分结果仅仅还停留在 PC 上的仿真阶段，已经在汽车上做过测试的主要包括 Broggi^[3] 等人的基于竖直边缘和简单人头模板验证的系统，Gavrila^[29] 等人的基于分层模板匹配的系统，Zhao^[6] 等人的基于神经网络的系统和 Shashua^[4] 等人的基于 Adaboost 的系统。

5 总结与展望

本文介绍了车辆辅助驾驶系统中基于计算机视觉的行人检测研究的最新进展。在分析基于计算机视觉的行人检测难点的基础上，介绍了行人检测系统的组成和常用方法。基于计算机视觉的行人检测系统一般包括 ROIs 分割和目标识别两个模块。ROIs 分割的目的是快速确定行人可能出现的区域，缩小搜索空间，目前常用的方法是采用立体摄像机或雷达的基于距离的方法，其优点在于速度比较快、比较鲁棒。目标识别的目的是在 ROIs 中精确检测行人的位置，目前常用的方法是基于统计分类的形状识别方法，其优点在于比较鲁棒。目前这一领域的最大问题是缺乏标准的测试数据库和测试方法。随着技术的进展，我们认为以下几个方面将有望成为未来的研究热点：

- 1) 数据库和测试方法的标准化
- 目前，行人检测的最大的问题是没有一个统一

的测试数据和测试标准, 研究还处于探索阶段, 建立一个公共的数据库和标准的测试方法是该领域的一个首要任务.

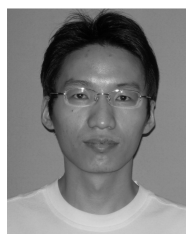
2) 多传感器融合

我们应当看到基于视觉的方法有其难以克服的弱点, 首先是受光照和天气的影响较大, 很难开发出对各种环境都鲁棒的算法, 目前最成功的系统也仅仅是针对天气良好情况的结果^[4]. 其次是立体视觉的测距的精度很差, 很难估算出行人的精确位置, 从而很难估计对行人造成的危险. 第三, 目前的行人检测算法仅仅处理站立的行人, 而对于其它姿态的行人和自行车、摩托车等情况缺乏相应的处理. 问题的解决依赖于多种传感器的融合. 超声传感器可以测量近距离内的目标, 可以应用于停车和倒车等情况; 23GHz 的微波雷达可以用来检测 20 米内的目标; 77GHz 的激光雷达可以检测 120 米内的目标; 红外图像可以用来在夜间检测目标. 将这些传感器有机地融合起来就可以构成一个覆盖所有空间和时间的检测系统, 保证系统在任何条件下正常地工作, 所以基于多传感器融合的行人检测系统是目前的研究热点, 也是将来一段时间的研究方向.

References

- Gavrila D M, Giebel J, Munder S. Vision-based pedestrian detection: the protector system. In: Proceedings of IEEE Intelligent Vehicles Symposium. Parma, Italy. IEEE, 2004. 13~18
- Tons M, Doerfler R, Meinecke M M, Obojski M A. Radar sensors and sensor platform used for pedestrian protection in the EC-funded project SAVE-U. In: Proceedings of IEEE Intelligent Vehicles Symposium. Parma, Italy. IEEE, 2004. 813~818
- Broggi A, Bertozzi M, Fascioli A, Sechi M. Shape-based pedestrian detection. In: Proceedings of IEEE Intelligent Vehicles Symposium. Dearborn, USA. IEEE, 2000. 215~220
- Shashua A, Gdalyahu Y, Hayun G. Pedestrian detection for driving assistance systems: single-frame classification and system level performance. In: Proceedings of IEEE Intelligent Vehicles Symposium. Parma, Italy. IEEE, 2004. 1~6
- Xu Feng-Liang, Liu Xia, Fujimura K. Pedestrian detection and tracking with night vision. *IEEE Transactions on Intelligent Transportation Systems*, 2005, **6**(1): 63~71
- Zhao Liang, Thorpe C. Stereo and neural network-based pedestrian detection. *IEEE Transactions on Intelligent Transportation Systems*, 2000, **1**(3): 148~154
- Oren M, Papageorgiou C, Sinha P, Osuna E, Poggio T. Pedestrian detection using wavelet templates. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Juan, Puerto Rico. IEEE, 1997. 193~199
- Mohan A, Papageorgiou C, Poggio T. Example-based object detection in images by components. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(4): 349~361
- Cheng Hong, Zheng Nan-Ning, Qin Jun-Jie. Pedestrian detection using sparse Gabor filter and support vector machine. In: Proceedings of IEEE Intelligent Vehicles Symposium. Vienna, Austria. IEEE, 2005. 583~587
- Sun Hui, Hua Cheng-Ying, Luo Yu-Pin. A multi-stage classifier based algorithm of pedestrian detection in night with a near infrared camera in a moving car. In: Proceedings of 3rd IEEE International Conference on Image and Graphics. Hong Kong, China. IEEE, 2004. 120~123
- Aggarwal J K, Cai Q. Human motion analysis: a review. *Computer Vision and Image Understanding*, 1999, **73**(3): 428~440
- Gavrila D M. The visual analysis of human movement: a survey. *Computer Vision and Image Understanding*, 1999, **73**(1): 82~98
- Pentland A. Looking at people: sensing for ubiquitous and wearable computing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(1): 107~119
- Wang Liang, Hu Wei-Ming, Tan Tie-Niu. Recent developments in human motion analysis. *Pattern Recognition*, 2003, **36**(3): 585~601
- Stein G P, Mano O, Shshua A. A robust method for computing vehicle ego-motion. In: Proceedings of IEEE Intelligent Vehicles Symposium. Detroit, USA. IEEE, 2000. 362~368
- Liu Xia, Fujimura K. Pedestrian detection using stereo night vision. *IEEE Transactions on Vehicular Technology*, 2004, **53**(6): 1657~1665
- Bertozzi M, Binelli E, Broggi A, Rose M D. Stereo vision-based approaches for pedestrian detection. In: Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Universita di Parma, Italy. San Diego, CA, USA: IEEE, 2005. **3**: 16
- Curio C, Edelbrunner J, Kalinke T, Tzomakas C, Werner von Seelen. Walking pedestrian recognition. *IEEE Transactions on Intelligent Transportation Systems*, 2000, **1**(3): 155~163
- Bertozzi M, Broggi A, Fascioli A, Graf T, Meinecke M M. Pedestrian detection for driver assistance using multiresolution infrared vision. *IEEE Transactions on Vehicular Technology*, 2004, **53**(6): 1666~1678
- Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Marriott, Hawaii. IEEE, 2001. **1**: 511~518
- Viola P, Jones M, Snow D. Detecting pedestrians using patterns of motion and appearance. In: Proceedings of 9th IEEE International Conference on Computer Vision. Paris, France. IEEE, 2003. **2**: 734~741
- Bertozzi M, Broggi A, Cellario M, Fascioli A, Lombardi P, Porta M. Artificial vision in road vehicles. *Proceedings of the IEEE*, 2002, **90**(7): 1258~1271
- Cutler R, Davis L S. Robust real-time periodic motion detection, analysis and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(8): 781~796
- Wöhler C, Kressler U, Anlauf J K. Pedestrian recognition by classification of image sequences: global approaches vs. local spatio-temporal processing. In: Proceedings of 15th International Conference on Pattern Recognition. Barcelona, Spain. IEEE, 2000. **2**: 540~544
- Ran Yang, Zheng Qin-Fen, Weiss I, Davis L S, Abd-Almageed W, Zhao Liang. Pedestrian classification from moving platforms using cyclic motion pattern. In: Proceedings of International Conference on Image Processing. Genoa, Italy. IEEE, 2005. **2**: 854~857
- Zhao Liang. Dressed Human Modeling, Detection, and Parts Localization [Ph.D. dissertation], CMU-RI-TR-01-19, Robotics Institute, Carnegie Mellon University, 2001

- 27 Yasuno M, Yasuda N, Aoki M. Pedestrian detection and tracking in far infrared images. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshop. Washington, DC, USA: IEEE Press, 2004. 125~125
- 28 Broggi A, Fascioli A, Grisleri P, Graf T, Meinecke M. Model-based validation approaches and matching techniques for automotive vision based pedestrian detection. In: Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Universita di Parma, Italy. San Diego, CA, USA: IEEE, 2005. 3: 1
- 29 Gavrila D M. Pedestrian detection from a moving vehicle. In: Proceedings of the 6th European Conference on Computer Vision-Part II. London, UK: Springer-Verlag, 2000. 37~49
- 30 Barrow H G, Tenenbaum J M, Bolles R C, Wolf H C. Parametric correspondence and Chamfer matching: two new techniques for image matching. In: Proceedings of 5th International Joint Conference on Artificial Intelligence. MIT, Cambridge, Massachusetts: MIT Press, 1977. 659~663
- 31 Leibe B, Seemann E, Schiele B. Pedestrian detection in crowded scenes. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, CA. IEEE, 2005. 1: 878~885
- 32 Tate S, Takefuji Y. Video-based human shape detection by deformable templates and neural network. In: Proceedings of 6th International Conference on Knowledge-Based Intelligent Information & Engineering Systems. Podere d'Ombriano, Crema, Italy: IOS Press, 2002. 280~285
- 33 Szaras M, Yoshizawa A, Yamamoto M, Ogata J. Pedestrian detection with convolutional neural networks. In: Proceedings of IEEE Intelligent Vehicles Symposium. Las Vegas, Nevada. IEEE, 2005. 224~229
- 34 Franke U, Joos A. Real-time stereo vision for urban traffic scene understanding. In: Proceedings of IEEE Intelligent Vehicles Symposium. Dearborn, MI. IEEE, 2000. 273~278
- 35 Vapnik V. *The Nature of Statistical Learning Theory*. Berlin: Springer-Verlag, 1999
- 36 Grubb G, Zelinsky A, Nilsson L, Rilebe M. 3D vision sensing for improved pedestrian safety. In: Proceedings of IEEE Intelligent Vehicles Symposium. Parma. IEEE, 2004. 19~24
- 37 Andreone A, Bellotti F, Gloria A D, Lauletta R. SVM-based pedestrian recognition on near-infrared images. In: Proceedings of 4th International Symposium on Image and Signal Processing and Analysis. Zagreb. IEEE, 2005. 274~278
- 38 Dai Cong-Xia, Zheng Yun-Fei, Li Xin. Layered representation for pedestrian detection and tracking in infrared imagery. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, CA. IEEE, 2005. 3: 13~13
- 39 Dalai N, Triggs B. Histograms of oriented gradients for human detection. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego. IEEE, 2005. 1: 886~893
- 40 Abramson Y, Steux B. Hardware-friendly pedestrian detection and impact prediction. In: Proceedings of IEEE Intelligent Vehicles Symposium. Parma, Italy. IEEE, 2004. 590~595
- 41 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, 60(2): 91~110
- 42 Mikolajczyk K, Schmid C, Zisserman A. Human detection based on a probabilistic assembly of robust part detectors. In: Proceedings of European Conference on Computer Vision. Prague, Czech Republic. Springer, 2004. 69~82



贾慧星 清华大学电子工程系博士生。2003 年获得北京交通大学电子信息工程学院学士学位。研究领域为模式识别, 计算机视觉, 智能车辆。本文通信作者。Email: jiahx03@mails.tsinghua.edu.cn
(**JIA Hui-Xing** Ph.D. candidate in Department of Electronic Engineering at Tsinghua University. He received his

B.S. degree from Beijing Jiaotong University in 2003. His research interests include pattern recognition, computer vision, and intelligent vehicle. Corresponding author of this paper.)



章毓晋 清华大学电子工程系教授。主要科学研究领域为其积极倡导的图象工程(图象处理、图象分析、图象理解及其技术应用)和相关学科。已在国内外发表了 250 多篇图象工程研究论文, 出版了专著《图象分割》和《基于内容的视觉信息检索》(科学出版社)。主

页: <http://www.ee.tsinghua.edu.cn/~zhangyujin/>.

(**ZHANG Yu-Jin** Professor of Department of Electronic Engineering, Tsinghua University. His research interests include image processing, image analysis, and image understanding, as well as their applications. He has published extensively on image engineering and related areas, with more than 250 research papers, and two monographs *Image Segmentation* and *Content-based Visual Information Retrieval* (Science Press). Homepage: <http://www.ee.tsinghua.edu.cn/~zhangyujin/>.)