1. 绪论

1.1 课题研究背景及意义

随着计算机科学技术的不断发展，人工智能的概念被提了出来，使得计算机可以在很多场景上代替人类。机器视觉就是人工智能的一个主要应用领域，其主要目标便是通过对图像的处理和分析，模拟人类的视觉识别能力并做出相应的判断。

今年来，随着城市建设的不断发展，以及人们对高品质生活的不断追求，机器视觉领域得到了快速的发展，市场也有巨大的需求。行人检测作为机器视觉领域的一个重要研究方向，该技术在道路交通，智能安防以及娱乐产业等很多领域都有着及其重大的研究意义。其主要应用在以下几个领域:

1. 智能驾驶

我国的汽车保有量增长迅速，相应的交通事故也不断增长，交通事故的主要原因也是人为疏忽导致，行人检测装置检测到行人时就会提醒司机或者自动采取制动措施，从而避免交通事故的发生。行人检测在智能驾驶的作用主要体现在辅助驾驶和无人驾驶上。行人检测技术需要相关的图像采集设备或者雷达等，设备采集到数据后处理器应用相应的图像识别算法分析数据最终得出是否有行人的结论，无人驾驶汽车则要做出相应的措施去避让行人，辅助驾驶系统则通过语音提示的方式告诉司机，司机再做出合适的处理。

1. 安防监控

在很多时候我们需要对特定场合的行人进行监控，比如监狱外围区域，私人住宅的行人进行监控，发现可疑人物就会及时报警；军事基地还可以监控外来人员非法入境；或者监控中发现行人出现异常举动，就可以通知公安、消防等快速介入，保障生命财产不受损失。尤其在金融机构、人员流动大的公共场所更加需要智能化的安防监控系统。早期，监控主要通过人来对显示屏上的图像进行监视，不仅需要大量的人力和物力，而且监控效率并不能得到保证。因此，行人检测技术就有了用武之地，系统可以自动完成行人检测和人的行为分析，实现了对可疑人物、危险行为的准确预警。

1. 机器人导航

通过行人检测技术，机器人可以找到感兴趣的区域，然后导航到目的区域，典型的应用就是现在的无人机可以对行人进行跟随，可以在完全无人操作的情况，主动跟随目标，或者在行人的周围定点环绕。

本文主要研究的是行人检测技术在智能驾驶上的应用，ADAS(高级驾驶辅助系统)现在已经开始广泛的应用在我们的汽车上，为我们驾驶过程中提供了很多辅助功能。行人检测效果图如图1-1，行人识别就是辅助驾驶系统的一个主要功能，行人识别技术应用在辅助驾驶的难点是检测的精度和速度。所以人们对行人检测技术的算法性能提出了更高的要求，在保证检测的精度的情况下，还必须提高检测的速度，保证检测的实时性。在早期提出的算法中，鉴于汽车设备环境的限制，在精度和实时性上面都不适合应用在辅助驾驶上，在最近几年一些优秀的检测算法被提出了，其中基于CENTRIST特征的行人检测算法不管是检测精度和速度上都让行人检测技术应用在汽车辅助驾驶系统有了较好的效果。而且GPU的并行运算可以进一步优化检测的速度，从而保证了检测系统的实时性。

1.2 课题研究现状分析

行人检测算法在最近的十几年发展迅速，国内外的研究人员对行人检测算法的准确性和实时性进行了大量研究，接下来对这些研究工作进行分析。其中主要的研究来自于国内外的高校和科研机构，国内主要有清华大学、南京大学、中科院等，其中中科院计算机科学重点实验室及自动化研究所都取得了巨大成绩。除此之外，国内许多科技公司也在该领域投入了大量的人力物力，其中包括目前比较火热的百度，成立了百度大脑，并将行人检测技术应用在了无人驾驶上，还有类似商汤科技在人脸识别和目标识别领域都取得了巨大成就，并将技术转换为商业产品应用在了我们生活中，为我们的生活提供了便捷。国外的加州理工，麻省理工，谷歌，脸书等高校和企业都在进行行人检测课题的研究；每年都有很多优秀的论文发表在计算机视觉权威的机构上。



图1-1

1.2.1 行人检测算法的研究现状

行人检测目前最常用的方法是基于统计学习的方法，根据大量的样本构建行人检测分类器。提取特征主要有目标的灰度，边缘，纹理，颜色，梯度直方图等信息。分类器主要包括神经网络、SVM、adaboost以及当前十分火热的深度学习。统计学习目前的难点有：行人的姿态、服饰各不相同、复杂的背景、不同的行人尺度以及不同的关照环境；提取的特征在特征空间中的分布不够紧凑；分类器的性能受训练样本的影响较大；离线训练时的负样本无法涵盖所有真实应用场景的情况。

2005年法国研究人员Dalal在2005的CVPR发表的HOG+SVM的行人检测算法，基于HOG特征的行人检测提出，使得行人检测领域有了较大突破，HOG+SVM作为经典算法也别集成到opencv里面去了，可以直接调用实现行人检测。

基于统计学习方法的行人检测技术很重要的一点就是样本集，而当前关于行人检测主要的数据库有[MIT数据库](http://cbcl.mit.edu/software-datasets/PedestrianData.html)，INRIA数据库，[Caltech行人数据库](http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/)等。其中最常用的是INRIA和Caltech数据库，INRIA数据库是目前使用的最多的静态行人检测数据库，提供原始图片及相应的标注文件。训练集有正样本614张（包含2416个行人），负样本1218张；测试集有正样本288张（包含1126个行人），负样本453张。Caltech数据库是目前规模较大的数据库，采用车载摄像头拍摄，约10个小时左右，视频的分辨率为640x480，30帧/秒。标注了约250,000帧（约137分钟），350000个矩形框，2300个行人，另外还对矩形框之间的时间对应关系及其遮挡的情况进行标注。数据集分为set00~set10，其中set00~set05为训练集，set06~set10为测试集（标注信息尚未公开）。

HOG特征描述了一幅图像的方向梯度信息，提取该特征采用的滑动窗口扫描图像的方式，将窗口划分为大小相同的图像块（block），以块为单位提取图片的特征，这种方法可以保证提取到稳定的HOG特征，HOG+SVM的结合在INRIA数据库上的检测精度接近90%，但是HOG特征的提取计算量高，所以此方法的检测速度很慢，无法实现实时检测。

2009年，Dollar等提出了积分通道特征（Integral channel feature），该方法使用了梯度幅度特征，图像的LUV特征，及梯度直方图特征，为了简化计算使用了积分图像的方法实现特征的快速计算，该方法由于提取了多种特征来描述图像，所以检测精度可以达到86%，而且检测速度也可以得到提高。但是其复杂的特征提取导致检测还是无法做到实时性。

2011年吴建鑫等提出了基于轮廓信息的CENTRIST特征，作者通过大量的实验证明了获取行人边缘轮廓信息就能够准确描述一个行人的特征，而且CENTRIST特征的计算也比较简单，检测效果与积分通道特征相比也相差无几，而且作者在分类器上使用的是级联分类器，首先是一个线性SVM分类器快速找出可能的行人区域，再将结果交给第二层的HIK SVM分类器检测，因此保证了算法的检测精度，所以此方法是一个检测精度和速度都比较可行的一个算法。

1. 基于CENTRIST特征的行人检测算法原理

2.1引言

本章首先阐述了CENTRIST特征的特点，以及基于CENTRIST特征的行人检测算法的主要步骤，首先是如何提取图像的CENTRIST特征，以及采用级联分类器对图片进行滑动窗口检测，检测图片中是否有行人，最后比较了CENTRIST特征和HOG特征在检测准确率和速度方面的差距。

2.2 CENTRIST特征简介

2.2.1 CENTRIST特征的原理

CENTRIST（C4）特征能够很好的描述行人的特征，而且特征的提取相对传统方法计算更加简单，待特征提取完毕后就可以直接交给分类器进行判断的出结果。C4在目前最高精确度下可以达到20帧每秒的检测速度，并且是在只使用一个处理线程和不使用GPU等硬件的情况下达到的。能达到实时而精确的检测源于以下两点：第一，相邻像素差值的符号是描述轮廓的关键信息；第二，CENTRIST描述子非常适合做人体检测，因为它编码了符号信息并且可以隐式地表达全局轮廓。使用CENTRIST描述子和线性分类器，我们提出了一种不需要显式生成特征向量的计算方法，它不需要图像的预处理或特征向量的归一化，只需要O(1)时间去测试一个图片区域。

C4特征是针对行人的局部轮廓提出的，相邻像素的大小关系是编码的关键，具体像素的值对该特征的计算影响并不大。图片中的行人的衣服上的图像，绘画等纹理性特征对行人检测有较大影响。因此首先用Sobel边缘检测算子对这些纹理性的特征进行消除，同时又能较好的保留行人的整体轮廓。所以我们使用Sobel算子先对图像进行平滑处理，用Sobel梯度值代替像素的大小，这样就可以消除这些局部的高频纹理特征带来的不理影响。因此通过Sobel算子平滑后的原始图像*Ｉ*可以得到一个新的图像*Ｉ*’ ，*Ｉ*和*Ｉ*’的对应关系如下图2-1：



图 2-1

如图2-2所示，原始图像（a）经过Sobel进行平滑处理后，图像很好地过滤掉了一些不重要的纹理特征，（b）很好的描述了行人的边缘轮廓信息。



图 2-2

C4特征中一个重要的概念是CT（Census Transform）值，图像上某个像素的CT值由该像素的灰度值和其九邻域的像素的灰度值比较而得来，具体计算的方法如图2-1所示。若相邻的像素的灰度值比中心像素的灰度值大，则把这个位置的值置为0，否则置为1。然后将邻域内的八个数字按照从上到下从左往右的一个顺序排列起来得到一个八位的二进制数，最后将这个二进制数转为十进制数，这个十进制数就是中间那个像素的CT值。而且由于是一个八位的二进制数得到的，所以这个CT值的取值范围就是0到255。



图 2-3

我们通过计算CT值来代替图像的灰度值，从而得到新的图像，我们称为CT图像。为了更好的描述，在遍历一张图像时，使用一个256维度的CT值直方图来统计每个CT值的在图像中出现的频率。最终得到的这个直方图就称之为CENTRIST特征。

在实际应用中，把一个图像分为若干个小的图像块，我们取的是一个108\*36大小的检测窗口，然后把这个块分割为9\*4的图像小块（block），每2\*2的邻接小块组成一个超级块（super-block），所以通过计算得到一共有24个超级块，计算特征时就以一个超级块为单位，则整张图片的CENTRIST描述特征就是256\*24=6144维大小的直方图。（由于CT值的计算是与3\*3的模板进行匹配，因此超级块的边缘像素不会参与CT值的计算）。

2.2.2 CENTRIST特征的计算过程

上一节我们介绍了特征的计算原理，本节我们会介绍提取一副图像的CENTRIST特征的过程，包含以下步骤：

1. 读取原始图片到内存中；
2. 采用滑动国定大小的检测窗口对图片提取CENTRIST特征。但是由于真实图像中行人的大小各异，为了能够不漏检，因此我们必须要对图像进行缩放，使得各个尺度的行人能够被滑动窗口检测到，如图2-5所示，红色的矩形框就是检测窗口。
3. 使用Sobel算子对图像进行处理，构建Sobel图像；
4. 根据前面介绍的 计算CT值的方法计算CT图像，从而提取行人的轮廓信息；
5. 检测窗口在CT图像上滑动，并计算得到窗口的CENTRIST特征；

整个流程如图2-6所示，接下来我们会介绍如何利用我们提取到的CENTRIST特征对行人进行检测。



图 2-5



图 2-6

2.3 级联分类器

基于CENTRIST特征的行人检测算法使用了线性SVM分类器和HIK SVM分类器串联得到的级联分类器来实现快速、准确的检测。首先要对分类器进行训练。

2.3.1 SVM原理

2.3.1 训练级联分类器

(1)线性SVM的训练

训练分类器时正样本集P我们用的是108\*36的包含行人的图像，负样本集我们使用的是同样大小的不包含人的图像，把正负样本集输入训练线性分类器