

行人检测与跟踪 开题报告

小组成员：无 24 梁栋 2012011092

无 24 李嘉龙 2012011096

无 24 王怡人 2012011099

一、 相关研究综述

行人检测是近年来计算机视觉领域的研究热点，其目标为在输入的连续视频中能够快速识别各种状态下的行人，并能够对其进行跟踪、计数。

目前的行人检测技术的框架主要分为两个部分：可能存在行人的区域（ROI）的识别、区域内是否为行人的判定。其中每个部分执行的内容以及目前前沿的研究成果如下：

1. ROI 识别

由于本实验所需要处理的视频均为在同一地点摄像机相对静止时所拍摄得到的，因此可以采用背景减除以及分析帧间差异的方法来获得可能包含行人的待检测区域，从而提升行人检测的效率。

在这之中，最为重要的是背景提取。背景与前景是一组相对的概念，在本实验中，由于需要检测的对象是行人，因此除了行人之外的物体全部都应该被视作背景。背景提取的方式有很多，或复杂或简单，不过各种背景模型都有自己的适用场合，在此将做简要介绍：

A. 帧差法

在帧差法中，背景就是上一帧图像，每一帧与上一帧进行差分运算，提取的效果与运动前景对象的速度以及帧率有关。其优点是处理速度较快，稳定性较好，但缺点是可能出现物体的“空洞现象”，即两帧中物体重合的部分可能被当作背景。

B. 背景统计模型

对一段时间内的视频进行统计，然后计算其统计数据（例如平均值、平均差分、标准差、均值漂移值等），将统计数据作为背景。具体的算法有均值偏移与 Kmeans 聚类级联[6]、混合高斯模型、滑动高斯平均、Color（基于颜色信息的背景建模方法）[7]等。

C. 编码本背景模型

编码本的基本思路是：针对每个像素在时间轴上的变动，建立多个（或者一个）包容近期所有变化的 Box（变动范围）；在检测时，用当前像素与 Box 去比较，如果当前像素落在任何 Box 的范围内，则为背景。

在线检测时，对于输入的每一帧，将其与提取得到的背景做差，并进行二值化处理，即可得到可能包含行人的待检测区域。在此基础上，便可以进行下一步行人检测的部分。

2. 行人检测

特征提取与描述是目标检测中的一个关键问题，它强烈影响到分类器的设计及其性能。特征没有精确定义和普适法则，依研究问题和应用类型而定。常用的图像特征提取与描述方法包括颜色特征、形状特征、纹理特征和统计学习特征等。

行人检测 (Pedestrian Detection) 是目标检测 (Object Detection) 中的一类重要问题，是对图像及视频认知的关键技术。对于行人检测的分类判别，包括特征提取与分类器学习两个核心问题，效果比较显著的有以下几种方法：

A. 小波变换

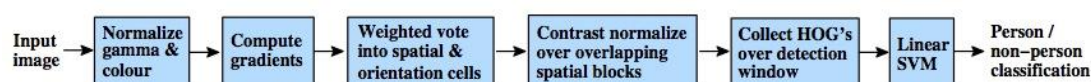
Oren 等人[3]利用小波表示法 (Wavelet Representation) 捕捉各个类别实例之间的结构相似性，提出小波模板 (Wavelet Template) 的方法在静态图像中检测目标。小波模板将目标定义为一组区域和它们之间的关系，该方法的关键在于使用一个小波基来表示模板，得到高效计算的算法和一种有效的学习策略。他们利用小波系数描述低层次的强度特性，根据相对强度和位置分布等信息选取合适的小波系数作为特征，并利用支持向量机 (Support Vector Machine, SVM) 进行分类。

B. 神经网络

Zhao等人提出基于图像分割和神经网络的识别方法[4]，该方法首先利用基于立体声视觉的前景/背景分割方法，提取出前景目标，每个目标通过训练好的神经网络被判定为行人和非行人。该算法包括三个核心步骤：首先将图像分割成不一致、不连续的子图作为候选目标，然后合并和拆分这些候选目标，得到符合行人大小、形状等约束条件的子图，最后将候选子图的强度梯度作为输入，利用训练好的神经网络进行行人识别。由于基于立体声检测和大数据训练出的神经网络，该方法具有良好的鲁棒性和实时性。

C. HOG

Dalal和Triggs提出方向梯度直方图 (Histogram of Oriented Gradient) 的方法[1]，HOG特征是一种在计算机视觉和图像处理中用来进行物体检测的特征描述子。它通过计算和统计图像局部区域的梯度方向直方图来构成特征。HOG特征结合SVM分类器已经被广泛应用于图像识别中，尤其在行人检测中获得了极大的成功。



D. Shapelet

Sabzmeydani等人定义了shapelet特征[5]，用于检测静态图像中的行人。每个shapelet特征覆盖了检测窗口的一个子窗口，shapelet是一系列底层特征利用AdaBoost算法组合

而成的中层特征，其中每个底层特征是从shapelet覆盖的子窗口提取出的位置、方向和强度信息。Shapelet特征可以有效对行人和非行人进行分类，并获得了比HOG+SVM算法更低的错误率。

二、 实验方案

正如之前的背景调研所述，本实验所需要完成的行人检测与跟踪任务主要可以分为两个部分，即 ROI 识别与行人检测。

在 ROI 检测部分，我们决定采取第二种即背景统计模型来进行我们的实验。除了[6]中的算法外，我们还计划采取 MOG（混合高斯模型）算法，同时借助[8]中提供的 BGSLibrary，对其他的背景提取算法进行测试，比较各种背景提取算法之间的效率与准确率。

在行人检测部分，首先我们将借助 OpenCV 实现 Dalal 在 2005 的 CVPR 发表的 HOG+SVM 的行人检测算法[1]，并使用 INRIA Person Dataset 进行训练与测试。在成功实现该算法的基础上，尝试采取 Zhu 等人提出的分类器级联结构[9]进一步提升效率。然后再通过图像分辨率变换等方法，尽可能实现实时性的行人检测。对于行人跟踪部分，在行人检测准确率较高的情况下比较容易解决。

实验所使用的语言为 C++ 及 Python，同时需要使用 OpenCV 以及 BGSLibrary[8]，最终的预期结果为：从电脑端输入待检测视频，经过短暂的背景提取后开始在线检测，实现接近实时的行人检测功能并保证一定的准确率。

参考文献：

- [1] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE, 2005.
- [2] Gavrilu, Dariu M. "The visual analysis of human movement: A survey." *Computer vision and image understanding* 73.1 (1999): 82-98.
- [3] Oren, Michael, et al. "Pedestrian detection using wavelet templates." *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. IEEE, 1997.
- [4] Zhao, Liang, and Charles E. Thorpe. "Stereo-and neural network-based pedestrian detection." *Intelligent Transportation Systems, IEEE Transactions on* 1.3 (2000): 148-154.
- [5] Sabzmeydani, Payam, and Greg Mori. "Detecting pedestrians by learning shapelet features." *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007.

- [6] “融合表观与运动信息的视频图像行人检测”
- [7] Horprasert, Thanarat, David Harwood, and Larry S. Davis. "A statistical approach for real-time robust background subtraction and shadow detection." *IEEE ICCV*. Vol. 99. 1999.
- [8] Sobral, Andrews. BGSLibrary: An OpenCV C++ Background Subtraction Library. IX Workshop de Visão Computacional (WVC'2013), Rio de Janeiro, Brazil, Jun. 2013.
- [9] Zhu, Qiang, et al. "Fast human detection using a cascade of histograms of oriented gradients." *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. Vol. 2. IEEE, 2006.
- [10] Walk, Stefan, et al. "New features and insights for pedestrian detection." *Computer vision and pattern recognition (CVPR), 2010 IEEE conference on*. IEEE, 2010.