

# 行人检测与跟踪 中期报告

小组成员：无 24 梁栋 2012011092

无 24 李嘉龙 2012011096

无 24 王怡人 2012011099

## 一、 前期调研

按照前期调研的结果，我们的项目框架主要分为两大部分：可能存在行人区域（ROI）的识别、区域内是否为行人的判定。小组的任务也主要围绕这两个部分去完成。

### ROI 检测部分：

在 ROI 检测部分，主要使用的模型有帧差法、背景统计模型、编码本背景模型。其中在背景统计模型中，具体的算法有均值偏移与 Kmeans 聚类级联、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）。我们将根据实际情况，选取其中的几种实现 ROI 的识别，对不同提取算法进行测试，比较各种识别方法之间的效率和准确度，并选出在一定条件下最优的实验结果。

**ROI 识别预期结果：**背景提取较为准确，和原图像作差后能明显看到行人的轮廓，方便下一步行人检测的进行。注意要多测试几组视频，比较不同环境下的识别结果。

### 行人检测部分：

行人检测效果比较显著的有以下几种方法：小波变化、神经网络，HOG，shapelet 特征。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，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。

**预期实验结果：**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。

## 二、 已完成的工作

在之前的几周中，我们主要完成了行人检测与跟踪问题中的 ROI 识别，即背景提取部分。在本实验中，我们采取了混合高斯模型对背景进行建模，并借助 OpenCV 实现了运动前

景的提取，具体的算法及结果如下。

混合高斯模型

混合高斯模型使用  $K$ （通常为 3 到 5）个高斯模型来表征图像中各个像素点的特征，在新一帧图像获得后更新混合高斯模型，用当前图像中的每个像素点与混合高斯模型匹配，如果成功则判定该点为背景点，否则为前景点。




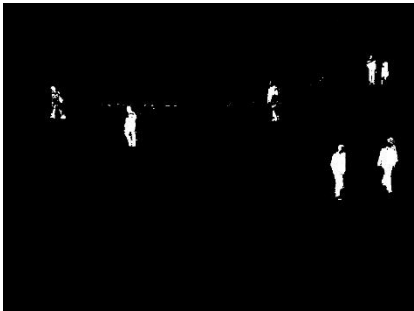
每个高斯模型，主要由方差和均值两个参数决定，对均值和方差的学习，采取不同的学习机制，将直接影响到模型的稳定性、精确性和收敛性。由于我们是对运动目标的背景提取建模，因此需要对高斯模型中方差和均值两个参数实时更新。为提高模型的学习能力,改进方法对均值和方差的更新采用不同的学习率；为提高在繁忙的场景下，大而慢的运动目标的检测效果，引入权值均值的概念，建立背景图像并实时更新，然后结合权值、权值均值和背景图像对像素点进行前景和背景的分类。


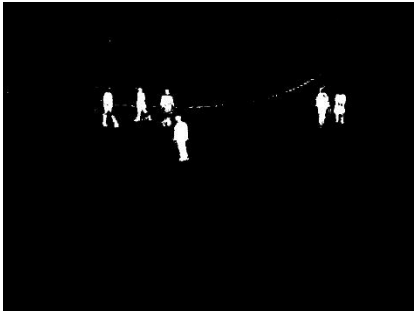




实现方式

本文借助 OpenCV 所提供的背景去除函数 BackgroundSubtractorMOG 来实现相关的背景检测与去除，采取  $K = 3$ ，学习率  $\rho = 0.001$ ，对助教所给的视频“PET\_View\_001.avi”进行背景检测，得到了较为满意的结果。

结果

具体的检测结果如下：

帧数	原始图像	检测结果
100		
200		

300		
400		
500		

### 结果分析

从上面的检测结果可以看出，混合高斯模型在背景检测上取得了较好的效果，能够准确地判断运动行人所在的区域，便于之后进行连通域分析以及行人检测。

## 三、 其余工作

### （一）行人检测

#### 1. 帧差法筛选

由于行人存在运动的特性，利用帧差法可以有效滤除静态背景噪声。不过帧差法筛选过程也会影响程序运行时间和效率等，因此，我们在这一部分还需要具体分析，根据需选择是否加入该环节。

#### 2. 连通域判别

我们已经得到了二值化的图像，下一步对连通域集合进行进一步分析，去除不可能包含行人的连通域。具体思路为：1）将面积过小的连通域判定为无效。2）比较连通域所

在的限定矩形框的长与高，比例关系超过一定阈值的判定为非行人。

### 3. HOG 特征分类器

- HOG 特征提取：提取正负样本的 HOG 特征
- 分类器训练：训练 SVM 分类器，得到训练好的模型，生成检测子
- 在线检测：用训练好的 SVM 分类器对测试文件的特征向量进行分类

#### （二）交互端口

在视频中标出感兴趣区域，对该区域内行走的行人进行检测。

#### （三）算法优化

- 完成全部功能后，对运行时间、正确率等性能进行优化
- 分析可能存在的错检和虚警现象，并进行程序和参数的调试
- 行人检测部分尝试 shapelet 特征（比 HOG+SVM 算法有更低的错误率），并进行两种算法的比较