**行人识别模块和限界设定程序的融合实现阶段总结**

1. **阶段目标及工作安排**

1、 行人识别模块和限界设定程序的融合，9月15日前完成；

2、 利用前期采集数据进行训练，并改进程序，9月30日前完成；

1. **工作完成方案或思路**

行人识别的总体框架如下：



图1 行人识别的总体框架

收集行人数据集。监督学习需要大量标定数据的支持。为了得到一个适应性强的分类器，需要采集各种气象条件、各种场景以及各种姿势的行人图像。从图像中抠出行人并规范化尺寸，得到正样本集。从无人的图像中随机采样形成负样本集。



图2 行人上）正样本和下）负样本示例

学习支持向量机分类器。给定已标注的行人训练样本集，训练支持向量机得到一个模型，模型将每一个样本分配到其中一类，不同的类的样本被尽可能大的间隔分开。学习得到的就是这个分割超平面参数。通过拉格朗日乘子及拉格朗日对偶性将学习过程转换为凸二次规划问题，可采用序列最小最优化（SMO）等方法求解。支持向量机的具体内容参见*Support-Vector Networks*和*Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines*等论文。

计算行人检测区域。中国成人中男人的肩宽大致在330mm到489mm之间，女人的肩宽大致在302mm到468mm之间。（参见《GB10000-88 中国成年人人体尺寸》）。图3到图6给出典型场景下，不同前视距离处一个“平均人”的大小，以矩形框标示，此处平均人肩宽取近似均值400mm。图中最外侧的两条曲线是根据4.88m警戒区域宽度计算出来的行人警戒线，中间的两条曲线是铁轨线。这个“平均人”在图中水平灰线以后的尺寸小于HOG行人检测器窗口尺寸（16×32），所以有效的行人检测区域，应定为人的像素宽度刚好为行人检测器窗宽时，所在扫描线以下的区域。

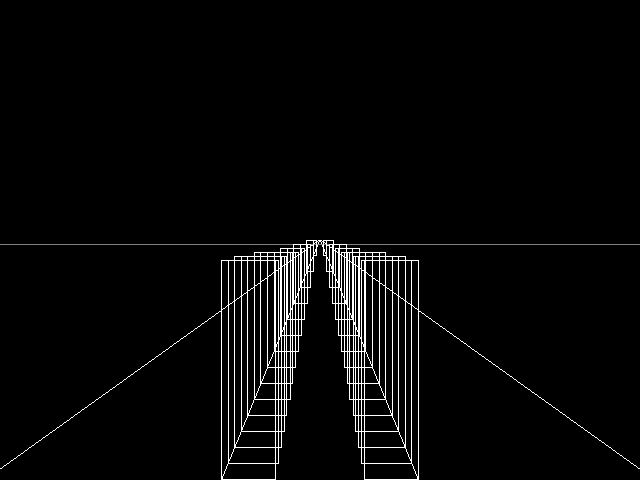


图3直轨时不同距离平均人的尺寸估计

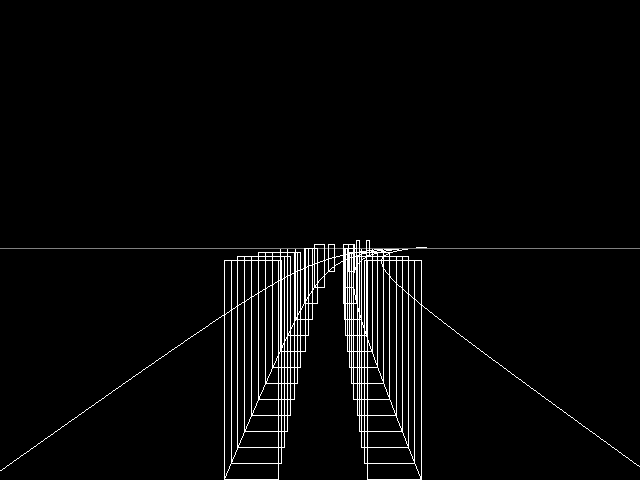


图4 铁轨水平曲率大于零时不同距离平均人的尺寸估计

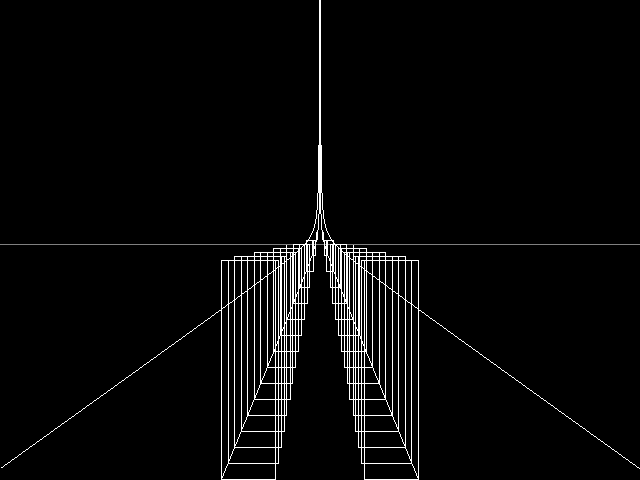


图5 铁轨垂直曲率大于零时不同距离平均人的尺寸估计

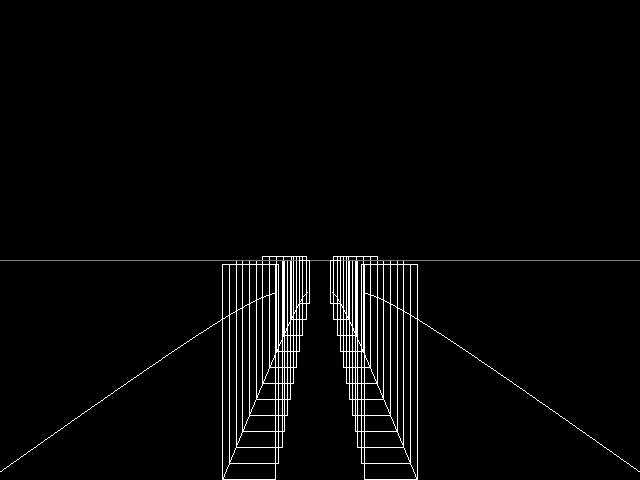


图6 铁轨垂直曲率小于零时不同距离平均人的尺寸估计

传输至GPU。将主机上的图像数据拷贝到设备端。

下采样。为了实现多尺度行人识别，需要利用原图构建图像金字塔。将图像数据拷贝到纹理空间实现快速的双线性插值。

计算梯度图像。利用行sobel算子和列sobel算子对图像进行行卷积和列卷积，得到梯度的幅值和方向。

计算梯度方向直方图。梯度方向直方图提供对物体形状的描述。设定梯度方向角取值范围为[0，180]，将其离散化为9个区间，设定4×4的胞，利用胞内每个像素的梯度幅值对方向直方图进行权重投影，那么每个胞就生成了一个9维特征向量。

正规化直方图。由于许多因素（如气象条件、物体材质等）会影响图像梯度，为了得到与这些因素无关的直方图，需要对其进行正规化处理。设定一个块包含2×2的胞，在8×8的块内进行直方图正规化，得到一个36维的归一化特征向量。设定检测区域大小为16×32，连接区域内所有块的特征向量，形成最终的756维特征向量。

线性SVM分类。将检测窗口的特征向量输入最大间隔分类器，计算特征向量与分隔超平面的距离。距离为负数，表示检测样本不是行人，负数结果的绝对值越大，结果的置信度越高。距离为正数，表示检测样本是行人，正数结果越大，结果的置信度越高。

传输至主机。将设备端的分类结果拷贝到主机。

格式化分类结果。统计正的检测结果的数量，以及正的结果位置、缩放系数和得分。

非最大值抑制。分类器处理不同位置、不同尺度的检测区域，对同一目标得到相近的检测结果，即相近的位置和缩放系数，所以需要融合检测结果，对同一目标只输出一个检测结果。可以采用表示位置、尺度的3D空间中的均值偏移算法实现检测结果的非最大值抑制。

行人检测得分的规范化处理。分类器输出的结果∈(-∞,+∞)，不太利于使用者直观地判断结果的可信度。所以需要进行规范化处理，将其值限制在[0,1]范围内，这直接对应了人们常用的“百分之几的可能性”的感性判断。采用sigmoid函数规范化分类器结果，这个函数可以将R∈(-∞,+∞)平滑地映射至(0,1)。



其中score和normscore分别是分类器输出值和其归一化的值，参数*T*∈[0,+∞)控制判定的严格程度。



图7 sigmoid函数，T=1

行人限界。结合之前计算的行人警戒线，将足迹处于左、右行人警戒线之外侧的行人检测结果去除。足迹以行人检测器给出的矩形框的下底边的中点近似。

行人检测器超参优选。

1. **工作完成情况**

截止到9月15日已完成行人识别模块和限界程序的融合。下图是利用从火车上采集到的一段进站的原始数据，运行的结果。图8中行人警戒线的宽度设置为4.88m，即行人需要和铁轨保持1.7225m的距离，才不会触发报警。图9是另一场景的测试结果，此处行人警戒线的宽度设置为6.88m。

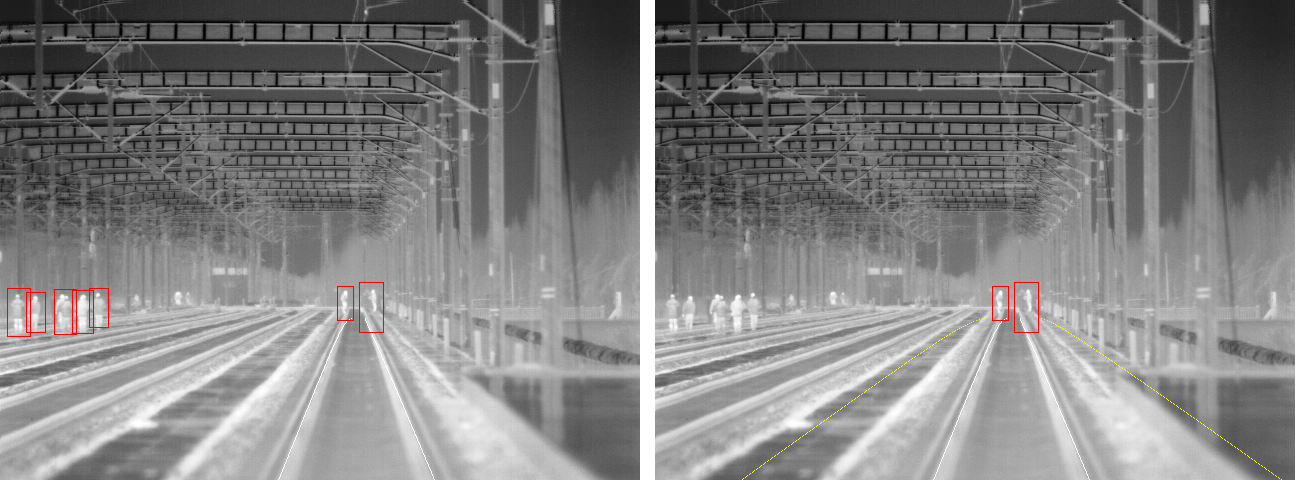


图8 行人识别及不限界和限界的对比。左右图中间两根线是检测到的铁轨。

左图)不对行人限界。右图)对行人限界。两侧的线是4.88m行人警戒线。



图9 行人识别及其限界。中间两根线是检测到的铁轨。

两侧的线是6.88m行人警戒线。进入警戒区的行人用矩形框标注。

矩形框左上角的数字表示其置信度。

设置了四组超参，并在测试集上比较各自配置下分类器的性能，如下表。在设置A的配置下，分类器的准确度、精度和召回率均最优。最终选择设置A的参数配置行人检测器。

表1 HOG算子超参设置

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Cell | Block | Window | Feature Length | Accuracy | Precision | recall |
| A | 4×4 | 2×2 | 16×32 | 756 | 99.02% | 98.73% | 96.65% |
| B | 5×5 | 2×2 | 15×30 | 360 | 98.22% | 97.22% | 94.3% |
| C | 6×6 | 2×2 | 18×36 | 360 | 97.44% | 96.48% | 91.29% |
| D | 8×8 | 2×2 | 16×32 | 108 | 94.31% | 92.7% | 79.46% |

1. **未完成工作原因分析**