

前言：

实验之初，准备使用深度学习的方法，但由于本课程是图像处理，而了解到深度学习的方法几乎不需要对图像进行处理，因此经过思考后决定使用传统方法HOG+SVM对人头进行识别。

实验过程：

在进行实验之前,首先要规划与思考整个实验的流程。大致如下:预处理,提取样本,训练,识别,返回调整参数与预处理方式,循环。

1.预处理：

对于传统方法而言，预处理是最重要的一步，做好了预处理，后面的调试就可以省下许多时间。

预处理中必要且最常见的就是灰度化处理，可以降低数据维度同时几乎不会影响原图的质量。但是我使用自己定义的灰度化函数时,效率却非常低,几乎1帧就需要1秒的时间,而调用了cv2中的灰度化函数后,灰度化处理几乎没有花费时间,令人费解。



原图



灰度化后的图像

灰度化只是最简单的第一步处理，接下来尝试进行滤波以消除摄像机导致的画质问题以及光照原因导致的明暗度不同的问题。滤波可以用于消除或抑制噪声干扰，在实验了中值滤波、高斯滤波、均值滤波等方式后，发现高斯滤波效果最好，可以部分解决视频的模糊等问题。但光照的去除并没有找到很好的去除方法。



滤波后的图像

滤波完成后感觉人头的边缘不够明显，因此使用边缘锐化再次对之前的图像进行处理。简单的对图像进行了一下锐化，其中没有遇到什么困难。



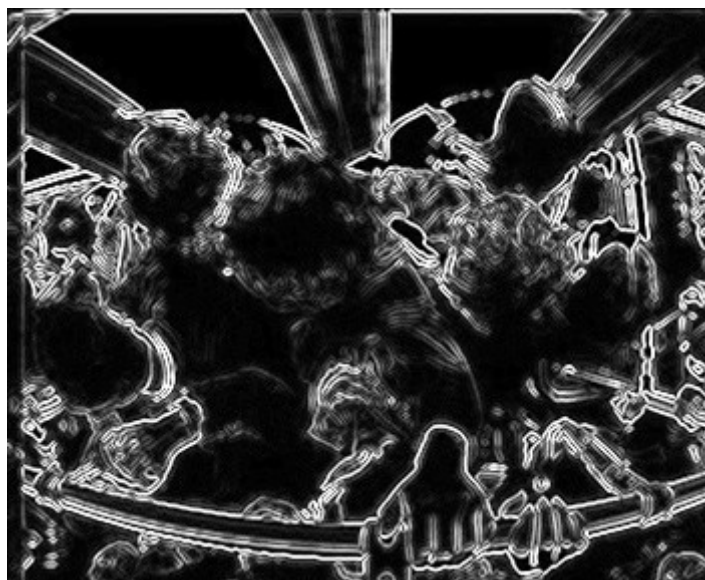
锐化后的图像

简单的滤波不能解决光照的问题，但在打开经过之前的预处理视频后，发现车辆行驶时，所造成的影响是在人头部分的大圆中产生一个个的小圆，这种情况顿时令我想起了形态学的处理方法。

因此对经过滤波的视频尝试使用形态学处理，然而腐蚀膨胀、开闭运算都试过了，感觉效果都不算很好，其中开运算的效果算是最好的，但是依然不能屏蔽光照的影响。同时这四种形态学处理方式所用的核矩阵，在本例中椭圆核最适合，其他核都会使人的头部模糊失真，本实验中使用的椭圆核矩阵为 $\{[[0, -1, 0], [-1, 5, -1], [0, -1, 0]]\}$ 。

无奈之下，准备使用开运算暂时处理一下，准备再寻找是否还有别的预处理方式进行优化。准备放弃的时候，遇到了惊喜，在一篇文章中发现形态学处理不只有腐蚀膨胀和开闭运算，还有形态学梯度、顶帽和黑帽三种方法。

尝试使用了形态学梯度算法后，发现效果惊人，基本可以以肉眼将人头识别出来，分割效果很好，同时还将背景提取了出去，后续不需要再对背景进行差值处理了。



形态学梯度处理结果

接下来进行均衡化，二值化考虑到可能会损失太多信息以及车内在灰度化后颜色深的地方太多，二值化可能会导致很多误识别，因此没有采用二值化处理。只是在对于光照的去除上，还是没有很好的处理方式。

最后两天看到可能 gamma 变换对此有比较好的效果，可惜时间不怎么来的及了，准备完成作业后自己再尝试一下。

2.提取样本：

完成预处理后，根据经过预处理的视频，截取正样本。

这个过程比较简单，在网上都有别人写好的代码，稍作修改即可完成。

截取后的正样本尺寸并不一致，因此需要批量设置尺寸。考虑到人头多是圆的，因此新的尺寸设置为 64×64 。同时，考虑到人头是可以随意旋转的，因此将最初的正样本进行了 5 次随机旋转后得到的 5 张新图片也加入正样本中。

负样本最初在网上找了几张内部无人的公交车图片，进行同样的预处理后，从中批量随机截取 64×64 的图片，这个过程也比较简单，自己随便写了个脚本完成。

3.训练：

第一次的训练困难的地方在于从网上获取的代码不适应我的环境和本次实验的条件，因此在调整代码上需要花费一定的精力。

SVM 的核类型选择了线性核，因为在浏览相关博客时看到一位作者介绍在样本量较多时核的选取影响不是很大，同时考虑到我的电脑性能很差，又是在虚拟机上运行，为了节省时间，所以选择线性核进行训练。

训练后将会进行识别过程，在识别过程中，会发现识别效果不尽人意的地方，如误识别和未识别等。这些未识别和误识别的样本将加入相应样本集中再次训练新的分类器，用于识别。

在训练时，遇到了很多问题，大部分都经过一段时间的研究后解决掉了，但是仍然有一个问题无法解决：在一次训练过程中，训练了一段时间之后，遇到了内存不足的问题，尝试了一些方法无法解决后，只好降低样本数量，这可能会导致训练结果变差，但目前无能为力。

4.识别：

识别过程就是将之前训练好的分类器加载，然后对目标视频进行检测。

在识别过程中发现，很多误识别情况多位于车门以及车中的扶手上，而且这些地方不可能存在人头，因此决定在进行识别时，将获取到的每一帧图片进行裁剪，裁去上下两端的车门和扶手两部分，只针对会存在人头的车中部分进行识别。

同时，考虑到检测不需要时时进行，因此可以间隔 3 帧或 5 帧进行一次检测，这样由于人的视觉暂留效果，既不会影响识别效果，同时又降低了计算消耗。完成识别后，检查识别效果，如果不满意，首先尝试增加正负样本重新训练分类器，若是结果距离预期差距极大，则需要考虑整个流程是否有问题。

实验结果：

由于样本数量不足，同时没有将误识别样本加入负样本再次训练，因此训练效果不是很好，出现部分人头未识别以及将地板识别为人头的情况。令人意外的是算法的效率惊人的高，几乎可以达到实时识别的效果。

改进方向：

- 1.使用深度学习方法进行识别。
- 2.寻找去除光照干扰的方法，获取更好的样本与处理视频。
- 3.将未识别的人头加入正样本，误识别的加入负样本，多次重复训练。
- 4.调整训练参数。