

基于 Meanshift 的目标跟踪

杨安东

目录

1 选题背景	1
1.1 Mean Shift 目标跟踪	1
1.2 反向投影过程	2
1.3 Mean Shift	2
2 实验说明	2
2.1 实验目标	2
2.2 方案设计	2
2.3 实现细节	3
2.3.1 图像二值化与 mask 制作	3
2.3.2 Mean Shift 核函数选取	3
2.3.3 Mean Shift 核函数归一化	3
3 结果分析	3
参考文献	4

1 选题背景

Mean Shift 是一个历史很远的算法，但是目前也焕发着活力。在查阅了各类博客与论文后了解到，Mean Shift 方法最初由 Fukunaga 在 1975 年提出 [1]，在聚类、图像平滑、图像分割和跟踪方面都有着广泛的应用。本次实验的主要内容为实现 meanshift 算法，并测试目标跟踪的效果。

1.1 Mean Shift 目标跟踪

基于 Mean Shift 的目标跟踪方法最初于 2000 年由 Dorin Comaniciu 提出 [2]，是一个轻量级的方法，大致流程是首先在一帧图像中指定一个兴趣区域，之后在下一帧中寻找于兴趣区域特征相同的区域，这样就需要判断下一帧图像中哪个区域与目前帧兴趣区域中的目标最为相似，而判断两个图像的相似性，可以通过计算两个图像直方图分布来计算。因此通过比较图像的直方图便可以找到下一帧中对应目标位置的概率分布。在获得了概率分布之后，就可以使用 Mean Shift 方法来将兴趣区域的边框逐渐移动到下一帧中目标对应的位置。由此实现基于 Mean Shift 的目标跟踪方法，之后也有许多改进，例如使用不同核函数的方法 [3]，使用空间颜色直方图来计算匹配目标概率分布图的方法等 [4]。

1.2 反向投影过程

可以看出 Mean Shift 目标跟踪过程中有两个主要步骤，一个是计算目标匹配的概率图，也就是计算反向投影图，反向投影图获得的结果是一张概率密度图，与输入图像大小相同，每一个像素值代表了输入图像上对应点属于目标对象的概率，像素点越亮，代表这个点属于目标物体的概率越大。

以空间颜色直方图为例，我对这一过程的理解是：

1. 将图像准换位 HSV 格式。
2. 建立选定兴趣区域的 H-S 直方图。这个直方图就相当于兴趣区域的特征。
3. 在待检测的图像上不断移动检测区域，计算当前检测区域与之前计算出的目标直方图的相似度。计算的方法有相关系数，卡方系数，相交系数，巴氏系数等。
4. 得到一个概率图，概率越大则与给定目标直方图的相似度越大。

1.3 Mean Shift

Mean Shift 目标跟踪过程中另一个主要部分是使用 Mean Shift 调整兴趣区域的边框，Mean Shift 作为使用核密度估计的爬山算法，可以寻找数据空间中的最高点，对应于此就是反投影后的概率图中概率最大的部分。因此划定兴趣区域后即可使用 Mean Shift 将其移动到目标处，Mean Shift 方法的详细过程上课时董老师已经有详细的介绍，不再说明。

2 实验说明

2.1 实验目标

实验主要目标是测试 meanshift 算法的效果，因此不对视频进行处理，而是直接给定两帧由车辆前方摄像头拍摄的图像。如图 1所示。目标是在原始帧中给定圈住行人的兴趣区域，由算法在待检测帧中调整兴趣区域位置，也框出出行人对应位置。



(a) 原始帧



(b) 待检测帧

图1 目标跟踪输入图像

2.2 方案设计

按照之前的介绍，方案主要有两步骤，计算反投影图，使用 Mean Shift 在获得的反投影结果上寻找概率密度最大处。主要流程如下：

1. 读取原始帧，转换为 HSV 格式。
2. 根据给定兴趣区域，转为二值图像。
3. 计算兴趣区域的直方图，获得目标直方图。

4. 读取待检测帧，转为 HSV 格式。
5. 根据目标直方图反向投影，获得概率密度图。
6. 计算初始给定兴趣区域的重心。
7. 计算平移向量。
8. 根据平移向量移动兴趣区域。
9. 不断循环步骤 6 到 8 直到平移向量小于一定的值结束。
10. 获得在待检测帧上的正确兴趣区域。

2.3 实现细节

本实验中有几个实验细节需要说明：

2.3.1 图像二值化与 mask 制作

为了更明显的凸显物体轮廓，加速运算速度，对图像进行来二值化处理。同时根据二值化的结果来制作反投影时的蒙版，去除无用的噪声，加速运算。

2.3.2 Mean Shift 核函数选取

在最初实现时使用的是平顶的核函数，但是发现效果很不好，经常导致兴趣区域不能正好框住高概率的区域，判断可能是最后收敛时抖动太大，因此改用了二次的核函数 $1 - x^2$ ，并添加对应的归一化方法。

2.3.3 Mean Shift 核函数归一化

在使用二次的核函数 $1 - x^2$ 时发现更新的步长不再变化，在检查后没有发现什么问题，于是在查阅资料后发现，核函数不当的归一化会导致步长过小，无法有效移动兴趣区域。在对照后发现，我的归一化函数不是归一而是使用了未加权重的像素和，这会导致归一化不正确，从而影响了跌倒过程，在修改为使用加权后像素点权重加和来归一化后获得了尚可的效果。

3 结果分析

本文输入的兴趣区域如图 2(a) 所示，圈定了行人部分，可见人物的头部颜色与背景树叶和路灯柱交杂较为严重。



(a) 输入兴趣区域



(b) 目标跟踪结果

图2 算法实现输入与输出

计算出得反投影图如图 3所示,可见由于主要使用了颜色信息来就行反投影,而大楼与人物裤子颜色较为相似,因此大楼也有较高的匹配程度。此外目标人物的裤子较好的被识别了出来,而人物的上半身并没有被很好的匹配出来,在分析后可能的原因有:

1. 人物裤子颜色与周围环境色差较大,较为容易识别,而人物上半身与头部有较多的深色部分,与背景的树影尤其是路灯杆较为相似,在计算直方图中区别度不够大,可能导致匹配度低。
2. 图像除了人物变化,摄像头角度也出现了较大的变化,由于光线是实时计算的,可能会导致图像整体像素值出现了一定的变化,干扰了匹配过程。



图3 反投影获得的概率图

最终的结果如图 2(b) 所示,可见与反投影获得概率图中高亮部分较为匹配,可以很好的收敛到图像中的高概率部分,但是高概率部分没有很好的包括目标全身。也许是处理过程中对颜色阈值的设置不够合适导致的。

参考文献

- [1] K. Fukunaga and L. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, 21(1):32–40, 1975.
- [2] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, volume 2, pages 142–149 vol.2, 2000.
- [3] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5):564–577, 2003.
- [4] S. T. Birchfield and Sriram Rangarajan. Spatiograms versus histograms for region-based tracking. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 1158–1163 vol. 2, 2005.