
目 录

一.实验名称.....	1
二.实验目的.....	1
三.实验原理.....	1
3.1 前言	1
3.2 高斯混合模型	2
四.实验结果.....	6
4.1 三通道运用	6
4.2 形态学操作.....	6
五.实验总结.....	9

一.实验名称

基于混合高斯模型的动态背景分割

二.实验目的

探索如何对 Wavingtrees 等具有动态背景的数据集进行有效的建模并分割，检测前景物体。

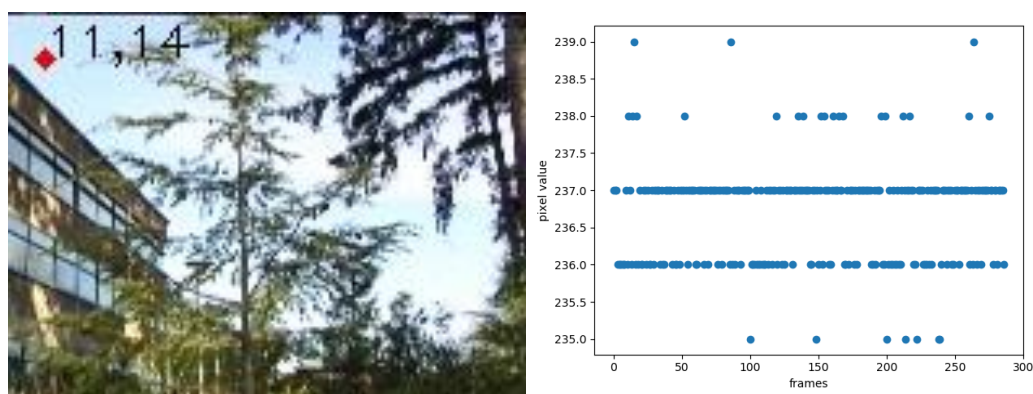
三.实验原理

3.1 前言

时域中的同一个点的像素值看做是一个像素过程，由一组像素组成，如果这里只考虑灰度图的话，对于点 x_0, y_0 处于时间 t 的像素历史值为

$$\{X_1, \dots, X_t\} = \{I(x_0, y_0, i): 1 \leq i \leq t\} \quad (1)$$

其中 I 代表着图像序列，由于各种因素，环境中的背景往往不是完全不变的，下图展示在 wavingtrees 中两个像素点随着帧数其像素值的变化。



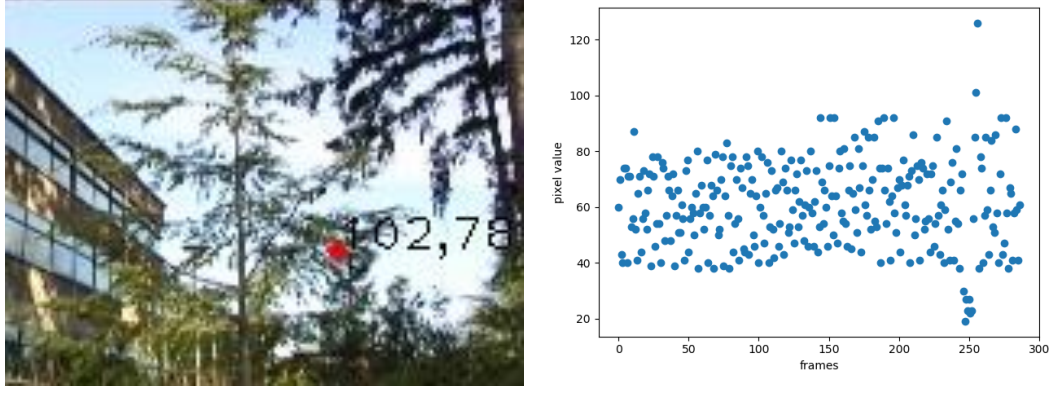


图 3.1 像素值变化图

从上图可以看出，我们在第一帧中选取了两个点，一个点坐标值为(11,14).位于较为远离树的位置，另一个点坐标值为(102,78)。位于树与人交叉的位置。从右上角与右下角的图表示像素值随着帧数的变化。横坐标为帧数，纵坐标为像素值的大小。可以明显的看出，随着时间的推移，即随着帧数的增加，像素点的值的分布较为分散，特别是前景与背景交叉的区域。所以单个模型，比如单个高斯模型难以适应这种场景。采用高斯混合模型就可以较好的适应这种变换。

3.2 高斯混合模型

3.2.1 背景建模

(1) 高斯模型定义

鉴于由于背景环境的复杂变化，单个高斯模型难以适应，考虑对每个像素值序列 $\{X_1, \dots, X_1\}$ 建立 K 个高斯分布综合描述。当前像素值得到的概率密度函数由以下公式进行描述：

$$P(X_t) = \sum_{i=1}^K (w_{i,t} * \eta(X_t, u_{i,t}, \Sigma_i, t)) \quad (2)$$

K 是该点对应的高斯模型的最大数目， $w_{i,t}$ 是时间 t 时的第 i 个高斯模型的权重，即当前高斯模型对于该像素点对应的所有模型占得权重， $u_{i,t}$ 是时间 t 时的第 i 个高斯模型的期望， Σ_i, t 是时间 t 时的第 i 个高斯模型的协方差矩阵， η 代表高斯概率密度方程：

$$\eta(X_t, u_{i,t}, \Sigma) = \frac{1}{(2\pi)^2 |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu_t)^T \Sigma^{-1} (X_t - \mu_t)} \quad (3)$$

考虑到程序的空间与时间复杂度， K 的取值一般为 3-5.而且协方差矩阵被简化为

$$\Sigma_{k,t} = \sigma_k^2 I \quad (4)$$

因为在这里我们假设 RGB 三通道的数据具有互相独立而且相同的方差。虽然有些不符合现实，但是在损失了一些精度的代价下减少了程序的时间复杂度。

(2) 训练过程

这里先采用若干张图片进行训练，即更新每个响度点对应的高斯混合模型的个数，高斯模型的参数等。背景建模的算法流程图如下

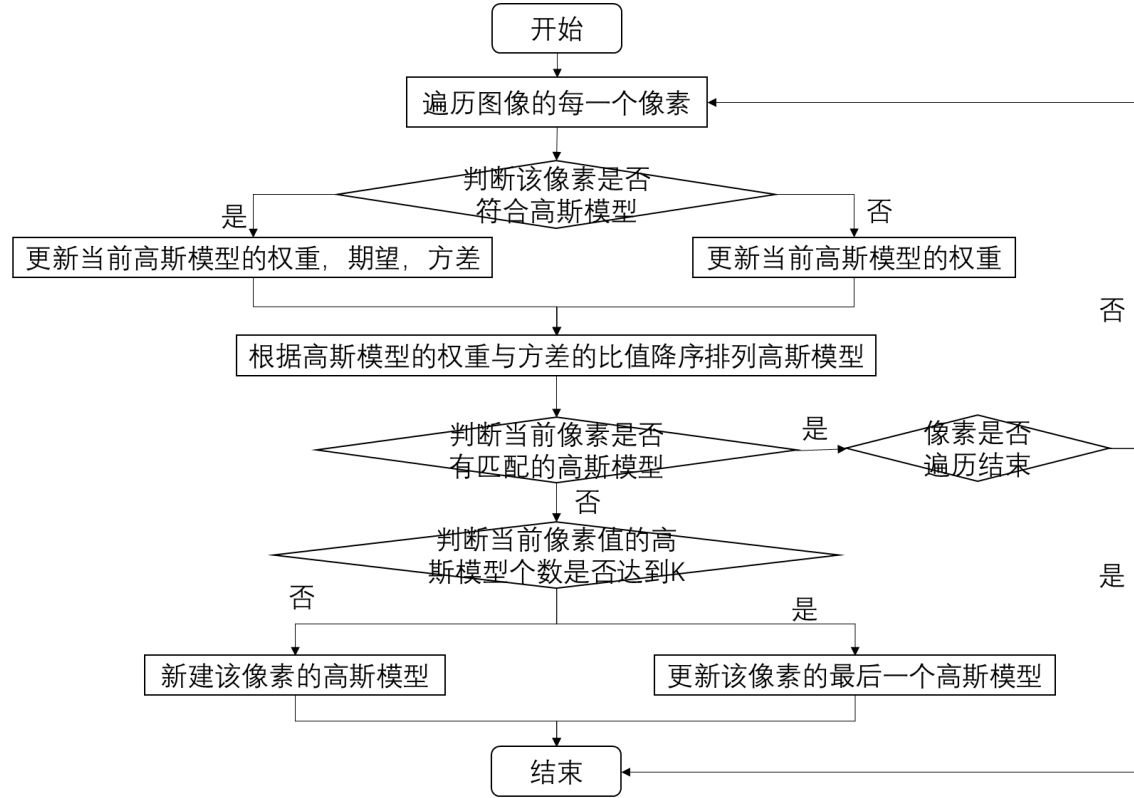


图 3.2 算法流程图

在上图中，判断一个像素点是否服从高斯分布：像素值是否位于高斯分布的 2.5 倍标准差内。对于每个像素点对应的高斯模型，如果该像素值匹配到了某个高斯模型，则更新公式如下：

$$\mu_t = (1 - \alpha)\mu_{t-1} + \alpha X_t \quad (5)$$

$$\sigma_t^2 = (1 - \alpha)\sigma_{t-1}^2 + \alpha(X_t - \mu_t)^T(X_t - \mu_t) \quad (6)$$

$$\omega_{k,t} = (1 - \alpha)\omega_{k,t-1} + \alpha(M_{k,t}) \quad (7)$$

对于没有匹配到的高斯模型，其方差与期望不变，权重的更新按照上述公式。其中当像素点匹配了某个高斯模型的时候 $\alpha(M_{k,t})$ 值为 1，否则值为 0。

如果该像素值与其对应的高斯模型都不匹配，则判断当前像素值的高斯模型个数是否已到达上限阈值 K。

1.如果达到的话，侧取权重与方差比值最低的那个高斯分布重新进行初始化。权重赋予一个较低的值，这个值在程序中取得 0.05.方差赋值为一个较高的值，程序中取值为 30.高斯分布的期望赋值为该像素点的值。

2.如果没有达到的话。则新建一个高斯分布。初始化步骤同上。权重赋予一个较低的值，这个值在程序中取得 0.05.方差赋值为一个较高的值，程序中取值为 30.高斯分布的期望赋值为该像素点的值。

当前像素值与其对应的高斯模型匹配结束之后，把该像素值对应的高斯模型按照权重与方差的比例降序排列，接着对所有高斯模型对应的权重进行归一化。

3.2.2 前景检测

每个像素值对应的各个高斯分布的参数会一直发生改变，当一个物体趋于静止的时候，对应的像素高斯分布的方差会趋于更小，其对应的权重会趋于变得更大，相对于一个运动的物体，他一般不会匹配到像素对应的高斯模型，运动的物体会让高斯模型的方差变得更大，权重变得更低。

我们按照高斯模型的权重与方差的比值进行降序排列，由阈值 T （程序中取值为 0.7）选出 B 个高斯分布作为当前像素判断是否属于背景的模式：

$$B = \operatorname{argmin}_b (\sum_{k=1}^b w_k > T) \quad (8)$$

如果一个新的像素值服从该 B 个高斯模型中的某一个，我们就认为该像素属于背景。但是如果该像素服从某个高斯分布，但是该高斯分布位于该 B 个高斯模型之后，该像素仍然会被判定为前景。 T 的取值不宜过大，否则会产生一个多峰分布模型，背景模型如果是单峰值的话，会选出一个可能性最高的高斯分布。

3.2.3 形态学处理

背景往往不是一直不变的，有些背景可能会一直处于一种细微变化的状态，比如本实验给出的数据集 wavingtrees，背景中的树一直处于在摇晃的状态，所以在检测前景的时候总是会把一些树的枝丫之类识别为前景。

为了去除一些被误识别为前景的树枝，考虑采用以下三种形态学操作。

（1）中值滤波：中值滤波常被用来处理椒盐噪声，这里的树枝可以看成是前景中的椒盐噪声。中值滤波的原理就是将当前像素值替换为周围领域像素值的中值。

（2）腐蚀：在周围都是背景的情况下，被误识别为前景的树枝可以经过腐蚀操作进行消除，但是这个操作会对前景中的人轮廓有一定的影响。

(3) 膨胀：由于中值滤波或者腐蚀操作会对前景中的人造成一定的影响，即会让人的轮廓出现空洞，所以可以采用膨胀操作在一定程度上减弱这个影响。

3.2.4 三通道的考虑

这里运用的三通道其实就是在三个通道上进行背景建模与前景检测。

背景建模：每个通道的像素点都会建立数目最多为 K 个的高斯模型。每个通道的建模过程跟 3.2.1 小节与 3.3.2 小节描述的一样。

前景检测：原来是灰度图的单通道进行判断，其接受为前景就认为是前景，而现在对于三通道就需要三个通道都认为是前景才认为当前像素是前景。

四.实验结果

本实验采用在线训练，在线测试的方式，为了对比 3.2 小节描述的形态学操作与三通道的有效性。实验中做了多组对比实验。

4.1 三通道运用

不采用任何的形态学操作，运用 RGB 三通道与只利用单通道的效果如下图 4.1 所示。



(a)单通道

(b)多通道

图 4.1 rgb 效果对比图

在图 4.1 中，左边的图是仅仅考虑单通道的时候第 247 帧的效果，右边的图是考虑三通道之后第 247 帧的效果。可以明显看出考虑了三通道之后，前景中的人轮廓显得更加丰满了，而对于单通道的情况，人的轮廓中就会存在许多的空洞区域。原因就在于三通道的時候是每个通道都接受为前景才认为是前景，从而是前景的判断更加准确合理。

4.2 形态学操作

实验中考虑了两种形态学组成操作对图像进行处理。

在实验中我们对比了在考虑三通道的情况下，采用开运算，即先腐蚀后膨胀；与先进行中值滤波，在进行膨胀的效果。如下图 4.2 所示。



(a)开运算

(b)中值滤波与膨胀

图 4.2 形态学操作对比图

在上图 4.2 中，左边的图是第 247 帧采用了开运算的效果，其中核为十字架形，大小为 3；右边的图为先进行中值滤波，再采用膨胀操作之后的效果，其中中值滤波的核大小为 7，膨胀操作的核大小为 3。从实验结果可以看出。直接利用开运算没有先进行中值滤波，再进行膨胀的效果好，前者会使人的轮廓产生一些空洞。这证明了中值滤波的效果比腐蚀的效果要好。

当形态学操作采用先进行中值滤波，后进行膨胀操作之后，对比了膨胀不同的核大小对于结果的影响。仍然采用第 247 帧的效果对比。

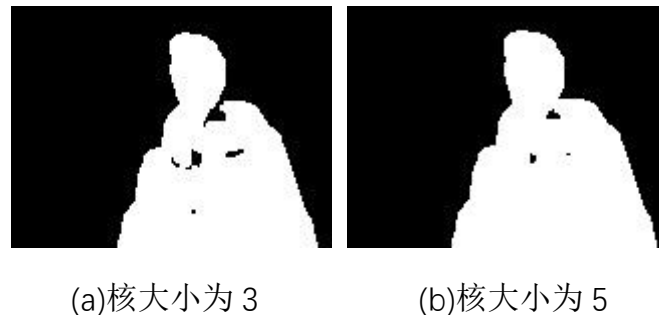


图 4.3

在图 4.3 中，左图表示膨胀的核大小为 3，右边的图表示核的大小为 5。可以看出当采用核大小为 5 的时候，人的轮廓显得更加丰满了。证明了核大小为 5 比 3 要好。

该实验的最后结果就是采用先进行中值滤波，后进行膨胀操作之后的效果。其中中值滤波的核大小为 7，膨胀操作的核大小为 5。

最后我们对比了 opencv 自带的 createBackgroundSubtractorMOG2 函数处理的效果与本实验最后得到的效果图。如下图 4.4 所示。

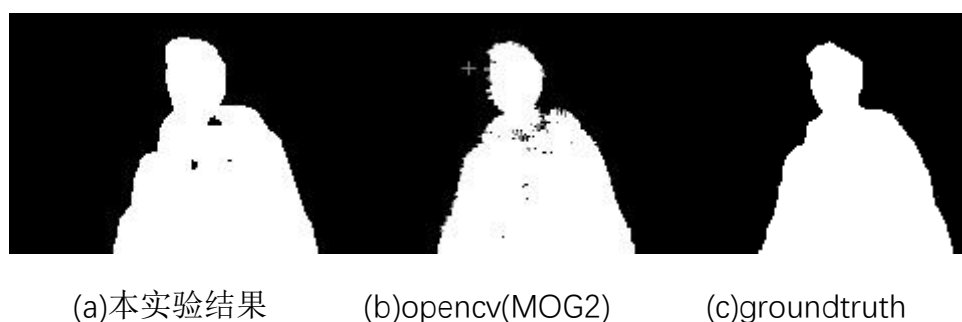


图 4.4

在图 4.4 中，仍然是展示第 247 帧的效果。左边的图是本实验最后得到的最好结果图，中间的图是 opencv 自带的 MOG2 函数得到的效果图。右边的图是 groundtruth。可以看出本实验得到的效果图已经和后面两张图非常类似，这就证

明了本实验的算法的有效性。不过其实也可以看出，由本实验得出的图由于采用了膨胀操作，导致人的轮廓有些略大。

五.实验总结

建立高斯混合模型可以较为有效的分割出背景与前景。得到了像素值在时间域上的拟合分布情况。运用高斯混合模型在 wavingtrees 上取得了较好的结果。

实验中运用的形态学操作一定程度上消除了被检测为前景的树，三通道在一定程度上消除了被误检测为背景的人的轮廓，使得人的轮廓更加的清晰。这都证明了实验中形态学操作与三通道运用的有效性。

但是从实验中也可以看出高斯混合模型对于这种动态背景还是不鲁棒的，仅仅是 wavingtrees 中背景中的树的轻微晃动，就会出现把树误检测为前景的情况，将前景中的人误检测为背景。这些情况在没有添加形态学操作与三通道的时候尤为明显。所以可以推断当背景发现较为明显的变化时，比如物体突然移动，光照突然变化。高斯混合模型都是无法适应这种变换的。