

# 运动目标精确检测算法<sup>①</sup>

## An Accurate Detection Algorithm of Moving Objects

白雪 黄廷磊 (桂林电子科技大学 计算机与控制学院 广西 桂林 541004)

**摘要:** 针对运动目标检测中背景多变和干扰多样的问题,提出了一种运动目标精确检测方法。首先在 HSV 空间,利用无约束的学习方式迅速建立三向量背景模型,各像素点根据其像素本身的混乱程度采取不同个数的三向量模型进行描述,然后在运动目标检测过程中及时的根据背景变化对背景模型进行更新,同时根据各种干扰的特点,分别实现了光照变化、物体的移入移出和运动目标短暂停留的干扰检测和排除,使背景模型抗干扰性和收敛速度得到有效提高。实验结果表明,该算法准确性高、抗干扰性强。

**关键词:** 无约束学习 运动目标检测 背景建模 干扰消除

从现实场景中准确的检测出运动目标,在工业控制、机器人视觉、生物组织运动分析等诸多方面具有重要意义。学者们已经提出了很多运动目标检测的算法,目前应用较多的运动目标检测方法可归纳为以下三种:光流法<sup>[1,2]</sup>、帧间差分法<sup>[3,4]</sup>、背景消减法<sup>[5-7]</sup>。光流法检测准确,但是计算复杂,很难应用的实时视频处理中。帧间差分法运算速度快,对场景中光线的动态变化不太敏感,但目标提取效果差,经常出现检测的运动目标比真实物体大得多的情况。背景消减法实现简单,计算速度很快,可以精确提取运动区域,完全的提供目标数据。

然而在实际应用中,很难得到一副“纯净”的背景图像,场景往往存在许多干扰因素使背景发生变化,因此要实现目标的精确检测,其算法必须具备较强的抗干扰性,主要有以下几个方面的干扰:(1)场景光线亮度变化的干扰;(2)场景中存在非静态的背景像素的轻微扰动;(3)场景中的背景物体移动。但是现存的运动目标检测算法,当背景存在上述干扰时,精度不高或实时性差。因此,本文提出了一种基于无约束学习的运动目标检测算法,在 HSV 空间建立背景模型,根据背景的混乱程度和各种干扰情况对模型进行更新,在检测速度和效果上都得到了满意的结果。

## 1 背景建模

### 1.1 RGB 到 HSV 的彩色转换

将 RGB 颜色空间的图像转化到 HSV 颜色空间,因为 HSV 颜色空间较 RGB 彩色模型更有利于干扰消除。在 HSV 彩色模型<sup>[8]</sup>中,表示色度;表示饱和度;表示纯度。

RGB 颜色空间到 HSV 颜色空间的转化公式为:

$$H = \begin{cases} \cos^{-1} \theta / 2\pi & B \leq G \\ (2\pi - \cos^{-1} \theta) / 2\pi & B > G \end{cases} \quad (1)$$

其中,

$$\theta = \frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} \quad (2)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}$$

$$V = \max(R, G, B) \quad (3)$$

其中 R, G, B 分别为 RGB 颜色空间三个通道的归一化的值。

### 1.2 建立背景模型

对于从 0 到 T 时刻的视频序列,可表示为时间序列  $\{X_1, X_2, \dots, X_T\} = \{I(i, j, t), 0 \leq i \leq M, 0 \leq j \leq N, 0 \leq t \leq T\}$ ,

<sup>①</sup> 基金项目:2008 年广西研究生科研创新项目(200810595081 1M421)

收稿时间:2009-03-11

每一时刻的视频图像的大小为  $M \times N$ ,  $I(i, j, t)$  表示像素点  $(i, j)$  在  $t$  时刻的颜色值, 则本算法采用  $L$  个三分量模型  $(h, s, v)$  的线性组合来描述  $I(i, j, t)$  的特征为:

$$p(I(i, j, t)) = \sum_{l=1}^L \omega_{i,j,t}^l (h_{i,j,t}^l, s_{i,j,t}^l, v_{i,j,t}^l) \quad (4)$$

其中  $\omega_{i,j,t}^l$  表示像素点  $(i, j)$  在  $t$  时刻第  $l$  个三分量模型的权重  $h_{i,j,t}^l, s_{i,j,t}^l, v_{i,j,t}^l$  分别表示像素点  $(i, j)$  在  $t$  时刻第  $l$  个三分量模型的  $h, s, v$  彩色分量。三分量模型的个数由该像素本身的复杂程度来决定。当然三分量模型的个数也不能任意增加, 一部分在整个的视频序中出现极少的特征应被替换, 算法设定描述某个像素的最大三分量模型的个数为  $L_{\max} = 5$ 。这样既提高了算法的计算效率, 又保证了算法的精确性。

### 1.3 背景模型的学习和更新

在长时间的视频监控中, 背景模型不可能是一成不变的, 因此, 需要对背景模型进行更新, 以适应背景的变化。

#### ① 背景模型的学习更新

在背景模型的学习阶段本文采用无约束学习的方式。将  $t$  时刻的图像输入, 像素点  $I(i, j, t)$ , 按照  $\omega_{i,j,t}^l$  从大到小将  $I(i, j, t)$  与背景模型中的三分量模型进行匹配, 设定经验阈值常量  $\lambda$ , 即若满足  $\|h_{i,j,t} - h_{i,j,t}^l\| + \|s_{i,j,t} - s_{i,j,t}^l\| + \|v_{i,j,t} - v_{i,j,t}^l\| < \lambda$ , 则与  $l$  个模型匹配, 若与几个模型的距离都不大于  $\lambda$ , 则根据距离最小准则。否则, 若满足  $l < L_{\max}$ , 则增加一个新的三分量模型到背景模型中。但若新的三分量模型已经达到最大值, 虽然建立这个特征的三分量模型, 但并不马上替换掉原有的  $l$  个三分量模型中权重最小的一个, 而是设定一个变量  $p$ 。若匹配, 则  $p = p + 1$ , 否则  $p$  值不变, 计算  $p$  与总的比较次数的比, 当大于  $n\%$  时, 用这个三分量模型代替原有的权重最小的一个。在替换之前的这段时间里, 这个新的三分量模型的  $h, s$  和  $v$ , 分量也和其它的三分量模型一样更新。

#### ② 背景模型的参数更新

背景学习与更新采用相同的迭代等式更新参数。具体方法是, 当输入的视频图像像素与已存在的第  $l$  个三分量模型匹配时, 则按下式更新第  $l$  个三分量模型:

$$\begin{cases} h_{i,j,t+1}^l = (1 - \varepsilon)h_{i,j,t}^l + \varepsilon h(i, j, t) \\ s_{i,j,t+1}^l = (1 - \varepsilon)s_{i,j,t}^l + \varepsilon s(i, j, t) \\ v_{i,j,t+1}^l = (1 - \varepsilon)v_{i,j,t}^l + \varepsilon v(i, j, t) \\ \omega_{i,j,t+1}^l = (1 - \gamma)\omega_{i,j,t}^l + \gamma \zeta_{i,j,t}^l \end{cases} \quad (5)$$

其中  $\varepsilon$  为模型的更新速率, 当不匹配  $\varepsilon = 0$  时,  $\gamma$  为权值的更新速率, 对于匹配三向量模型  $\zeta = 1$ , 不匹配三向量模型当匹配  $\zeta = 0$ 。从公式可以看出, 背景模型的更新速度随  $\gamma$  的增大而加快, 背景模型的可靠性随  $\varepsilon$  减小而增强, 所以在本文中  $h, s$  和  $v$  分量更新因子取较小值, 而权重的更新因子取较大值, 这样在保证背景模型稳定性的同时收敛速度也提高了。

## 2 环境干扰分析及解决方法

### ① 光照变化

场景中的光照变化一般可分为两种情况, 一种是光照亮度缓变, 如日光随时间缓慢变化; 另一种是光照亮度突变, 如开关灯等。三分量模型在模型更新过程中能适应光照亮度缓变, 而对于光照亮度突变, 背景模型不能快速适应, 而出现误检。为了避免此种情况, 我们检测当前帧中亮度分量发生较大变化的像素个数, 若变化像素个数大于整帧像素个数的  $q\%$ ,  $q$  是阈值常量, 则背景模型用当前的背景值替代:

$$\begin{cases} h_{i,j,t+1} = h(i, j, t) \\ s_{i,j,t+1} = s(i, j, t) \\ v_{i,j,t+1} = v(i, j, t) \end{cases} \quad (6)$$

### ② 场景本身的变化

在运动目标检测中, 场景本身也是经常变化的。背景像素的轻微扰动是常见的, 如树叶的轻微晃动等, 这种扰动使背景图像中的某些像素点在不同的时刻表现出不同的特征, 三分量模型本身能够通过背景模型的学习和更新排除这种扰动。但是, 当场景中出现物体的移入移出, 这种背景大范围的变化时, 则会出现误检, 我们采用相邻帧差和隔帧差相结合的检测法。帧间差法对物体运动的敏感性高, 实时性好, 能有效的消除物体移入移出产生的误检。其基本算法是:

对于视频序列的连续三帧为  $X_{t-1}, X_t$  和  $X_{t+1}$ , 计算相邻两帧源图像的绝对差图像  $D_1 = |X_t - X_{t-1}|$ ,  $D_2 = |X_{t+1} - X_t|$ , 对  $D_1, D_2$  取阈值进行二值化, 得到

二值图像  $B_1$  和  $B_2$ 。在计算  $D_3 = |X_{t+1} - X_{t-1}|$  对  $D_3$  取阈值进行二值化, 得到二值图像  $B_3$ 。比较  $B_1, B_2, B_3$ , 若  $B_1$  与  $B_3$  相同且  $B_2$  中该目标不被判断为运动目标, 则物体移入移出, 否则为运动目标。

### ③ 运动目标的不规则运动

在运动目标检测中, 运动目标的运动速度是不规则的, 有时可能会短暂停留后继续运动, 这时可能把运动目标误检为背景。为了解决这个问题本文设定了一个阈值  $a$ , 当像素大于  $a$  次被判定为背景时, 才认为为背景像素。

## 3 运动目标的检测

根据第二部分的分析, 本文提出的基于无约束学习的运动目标检测精确算法的步骤如下:

### 3.1 建立背景模型的步骤

① 将视频序列图像按第 1.1 节所述方法进行 RGB 到 HSV 的彩色转换, 由第 1.2 节所述方法初始化基于无约束学习的三向量背景模型。

② 输入后续帧序列, 由第 1.3 节所述的匹配方法与对应背景模型匹配。如没有匹配的子模型则执行第三步, 若有匹配的子模型, 则执行第四步。

③ 按照第 1.3 节所述的三向量模型更新方法创建新三向量模型。执行第四步。

④ 按照第 1.3 节所述的背景模型的参数更新方法更新背景模型。重新执行第二步。

### 3.2 运动目标检测步骤

① 如第 3.1 节所述建立基于无约束学习的三向量背景模型后, 执行第二步。

② 按照第 2 节所述的方法, 检查场景中是否有干扰因素, 并排除干扰。转第三步。

③ 将输入视频序列的每个像素点和对应的各三向量模型进行比较, 若有匹配, 则该点属于背景模型, 否则是前景。提取出所有的前景像素点。执行第四步。

④ 由于噪声的影响等因素, 可能使运动目标内的区域出现断点, 所以用四邻域的区域标号法和数学形态学的方法对该图像进行进一步的处理, 去除噪声和断点。最后得到准确的运动区域。

## 4 实验结果

在室内室外针对多种场景条件进行了实验, 本文均得到了较好的效果。如图 1(a)~(c)分别为在昏暗的

走廊里采集的一段视频序列的图像帧、其背景重建结果和检测结果, 可以看到本文的算法将运动的人物被正确的检测了出来。图 2(a)~(c)分别为在含有很多噪声的雨天采集的一段视频序列的图像帧、其背景重建结果和检测结果, 可以看到应用本文的算法取得了较好的检测结果。图 3(a)~(c)分别为在实验室场景下采集的视频序列中的两帧和由本文的方法的检测结果, 可以看出在采集过程中光照发生了突变, 本文的方法仍然能够正确的检测出运动目标。图 4(a)~(c)分别为在教学楼门口采集的一段视频序列中的两帧和检测结果, 可以看出原来场景中的垃圾筐被移走, 应用本文的算法仍然准确的检测出了运动目标。图 5(a)~(c)分别为在校园路上采集的视频序列的一帧、其背景重建结果和检测结果, 在该视频序列中, 运动目标做了运动的短暂停留, 本文的算法仍然有效的将运动目标检测了出来。图 6(a)~(c)分别为实验室中采集的视频序列的一帧、传统的混合高斯模型运动目标检测算法的检测结果和本文算法的检测结果, 由图可以看出, 在背景有风扇转动干扰的情况下, 传统的混合高斯背景模型运动目标检测算法出现了误检测, 把运动的风扇当成了运动目标, 而本文的算法取得了较好的检测效果。图 7(a)~(c)分别为在雨天采集的视频序列的一帧, 传统的混合高斯模型运动目标检测算法的检测结果和本文算法的检测结果, 由图可以看出在背景有很多雨点噪声和树枝晃动的情况下, 传统的混合高斯背景模型运动目标检测算法出现了误检测, 把一些树木和静态的物体都当成了运动目标, 而本文的算法将运动目标正确的检测了出来。对上述七种视频序列测试时, 本文方法的视频处理速度达到了平均 10.91 帧/秒, 满足实时性要求。结果表明本文所提出的方法对昏暗的灯光、含有大量噪声的雨天、树叶晃动, 光照突变、物体的移入移出、运动目标短暂停留这几类扰动都有较强的适应能力, 取得了较好的检测效果。

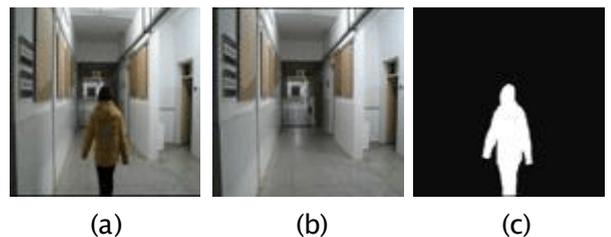


图 1 走廊中运动目标检测实验

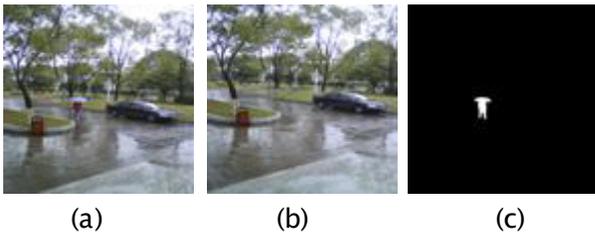


图 2 雨天运动目标检测实验

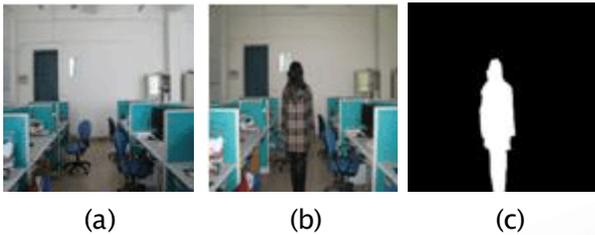


图 3 光照突变运动目标检测实验

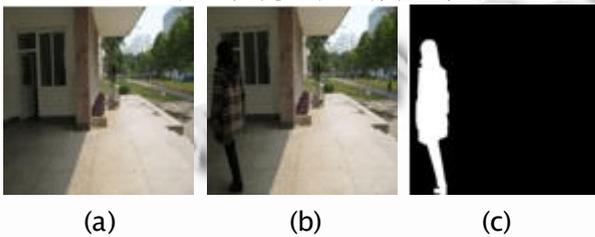


图 4 物体移入移出运动目标检测实验

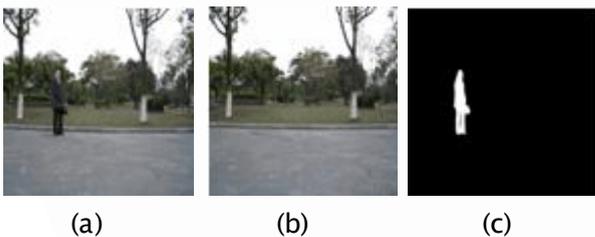


图 5 运动目标短暂停留运动目标检测实验

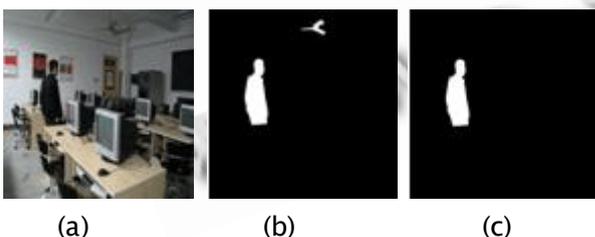


图 6 室内场景运动目标检测对比实验

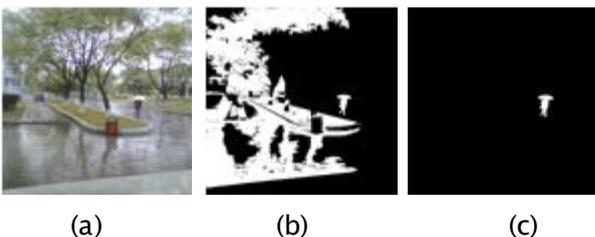


图 7 室外场景运动目标检测对比实验

## 5 结论

本文分析了运动目标检测面临的一些问题，并针对这些问题提出了一种运动目标精确检测算法，对视频场景进行无约束学习和快速更新，建立不等数目基于 HSV 颜色空间的三向量模型描述的稳定的可靠的背景模型，加快了建模速度；采用经验阈值和距离最小准则相结合的匹配准则，使模型匹配的准确性很高。考虑了干扰因素的影响，同时提出了有效的干扰检测和处理方法，实验分析表明，该算法对各种场景具有较好的检测效果和检测速度。

## 参考文献

- 1 Smith SM, Brady JM. ASSET-2: Real-Time motion segmentation and shape tracking. *IEEE Trans. on PAMI*, 1995,17(8):814 – 820.
- 2 Hiraiwa A, Fuse K, Komatsu N, et al. Accurate estimation of optical flow for fully automated tracking of moving-objects within video streams. *IEEE International Symposium on Circuits and Systems*, 1999,4:515 – 519.
- 3 Meier T, Ngan KN. Automatic segmentation of moving objects for video object plane generation. *IEEE Trans. on Circuits and Systems for Video Technology*, 1998, 8(5):525 – 538.
- 4 Lipton A, Fujiyoshi H, Patil R. Moving target classification and tracking from real-time video. *Applications of Computer Vision*, 1998. WACV'98. Proceedings, Fourth IEEE Workshop on. Princeton, NJ, USA: IEEE, 1998.8 – 14.
- 5 Xu J, Hu F, Huo H. Vehicle detection and separation based on 2D spatial temporal image analysis. *Journal of Shanghai Jiaotong University*, 2002,36(6):887 – 890.
- 6 Zhao T, Nevatia R. Tracking multiple humans in complex situations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2004,26(9):1208 – 1221.
- 7 Elgammal A, Duraiswami R, Harwood D, Davis L. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. *Proc. of the IEEE*, 2002,90(7):1151 – 1163.
- 8 Lee CC, Chen CL. A Mach-Zehnder joint transform correlator using encoding technique with reflective spatial light modulators based on the HSV color space. *IEEE 8th International Conference on Signal Processing (ICSP'06)*. Guilin, China, 2006.