

基于颜色布局描述子的改进 ViBe 算法

王彤¹, 王巍^{1,2,3*}, 崔益豪¹, 朱天宇¹

(1. 河北工程大学信息与电气工程学院, 河北邯郸 056038; 2. 物联网技术应用教育部工程研究中心(江南大学), 江苏无锡 214122;
3. 江南大学物联网工程学院, 江苏无锡 214122)
(* 通信作者电子邮箱 wangwei83@hebeu.edu.cn)

摘要:针对 ViBe 算法在检测运动目标时会有“鬼影”产生和在动态背景下会对目标检测过程中产生干扰造成运动目标误检的问题, 结合颜色布局描述子 (CLD) 提取关键帧进行三帧差分、形态学后处理的技术, 提出一种改进的 ViBe 算法。首先, 通过 CLD 提取视频关键帧图像; 然后, 将所选取关键帧图像进行三帧差分, 通过差分结果将含有运动目标的背景模型进行填充, 得到真实的背景图像, 再对运动目标进行检测, 以达到消除鬼影的目的; 最后, 在背景模型更新的阶段加入自适应阈值的形态学处理技术, 消除动态背景模型中的干扰信息。实验结果表明, 所提算法在运动目标检测时在避免鬼影、抗动态背景干扰等方面表现出优越性, 在相似度量阈值选取为 0.67 到 0.72 时, 所提算法的准确率最高可以达到 99.4%, 可以理想地检测出运动目标的位置信息。

关键词:运动目标检测; ViBe; 关键帧; 三帧差; 颜色布局描述子; 鬼影

中图分类号: TP391.41 **文献标志码:** A

Improved ViBe algorithm based on color layout descriptor

WANG Tong¹, WANG Wei^{1,2,3*}, CUI Yihao¹, ZHU Tianyu¹

(1. College of Information and Electrical Engineering, Hebei University of Engineering, Handan Hebei 056038, China;
2. Engineering Research Center of Internet of Things Technology Applications, Ministry of Education (Jiangnan University),
Wuxi Jiangsu 214122, China;
3. School of Internet of Things Engineering, Jiangnan University, Wuxi Jiangsu 214122, China)

Abstract: In view of the problem that ViBe (Visual Background extractor) algorithm sometimes produces “ghost” when detecting moving targets and the misdetection problem of moving targets caused by the interference produced by the algorithm in the target detection with dynamic background, an improved ViBe algorithm was proposed based on the techniques of three frame differencing and morphological post-processing carried on the key frames extracted by Color Layout Descriptor (CLD). Firstly, the video key frame images were extracted by CLD. Secondly, three frame differencing was performed on the selected key frame images. The background model containing the moving target was filled by the difference results, obtaining the real background image. Thirdly, the moving target was detected, so as to eliminate the “ghost”. Finally, the morphological processing technique with adaptive threshold was added in the updating stage of background model to eliminate the interference information in the dynamic background model. The experimental results show that the proposed algorithm has superiority in avoiding ghost and anti-dynamic background interference in moving target detection, and when the similar measurement threshold is selected from 0.67 to 0.72, the accuracy of the algorithm can be as high as 99.4%, indicating that the algorithm can ideally detect the position information of the moving target.

Key words: moving target detection; ViBe (Visual Background extractor); key frame; three frame differencing; Color Layout Descriptor (CLD); ghost

0 引言

近年来,随着大量监控摄像头的使用,视频监控数据爆炸性增长,对海量视频数据的存储、检索、浏览提出了严峻的挑战。为解决这一问题 Pritch 等^[1-2]和 Rav-Acha 等^[3]提出了视频浓缩技术,极大提高了视频浏览的速度,压缩了监控视频存储空间。其中运动目标检测作为视频浓缩技术中的第一步骤,

也是视频浓缩系统搭建的基础,目的是为了提取到完整的背景与运动目标,所以运动目标检测技术作为视频浓缩技术的关键步骤成为国内外研究学者关注的重点。

运动目标检测的基本思想是:从视频的帧图像中建立背景模型,将当前帧图像与背景模型比较,得到运动目标的信息。目前比较常用运动目标检测的方法主要有:帧差法^[4]、光流法^[5]和高斯混合建模 (Gaussian Mixture Model, GMM)^[6]。

收稿日期: 2019-07-11; **修回日期:** 2019-08-22; **录用日期:** 2019-09-02。 **基金项目:** 国家重点研发计划项目(2018YFF0301004); 国家自然科学基金资助项目(61802107); 教育部-中国移动科研基金资助项目(MCM20170204); 江苏省博士后科研资助计划项目(1601085C)。

作者简介: 王彤(1995—),女,河北邯郸人,硕士研究生,主要研究方向:人工智能、图像处理; 王巍(1983—),男,河北邯郸人,副教授,博士,CCF会员,主要研究方向:公共安全物联网、隐式人机交互; 崔益豪(1994—),男,河南洛阳人,硕士研究生,主要研究方向:群智能算法、人工智能、嵌入式系统; 朱天宇(1996—),男,江苏徐州人,硕士研究生,主要研究方向:无人机、无人驾驶。

帧差法对背影变化不敏感,当光照有细微变化影响图像像素有轻微变化,运行差分时可以相互抵消,稳定性较好,但是很难适应复杂的场景,所检测出来的目标有很大的“空洞”,并且帧间间隔的选择对目标检测结果影响较大。光流法计算量大,对光线太敏感,在光照变化大的场景下识别效果不佳。虽然GMM^[7]方法可以准确检测运动目标,且对光线的变化不敏感,但是建模过程复杂,计算量大,对慢速运动的目标检测效果不好。2009年, Barnich等^[8-9]提出了ViBe (Visual Background extractor)算法。ViBe算法的提出为运动目标检测提供了一个新的思路,使用ViBe完成运动目标检测相对于其他算法计算量小,处理速度快,样本衰减最优,具有较好的运动目标检测性能。

但是ViBe算法随着检测速度的提高,鬼影会被引进检测中造成前景误判,而且该算法在动态背景中容易受噪声和背景的干扰。基于高斯混合模型的ViBe算法^[10],虽然减少了鬼影出现,但是由于建立模型方法的复杂性,导致算法计算能力下降,使得算法缺少实时性;结合三帧差分的ViBe算法^[11],虽然在特定情况下可以避免“鬼影”的出现,但是在有些情况下会产生新的“鬼影”或空洞,并且选择的步长对其影响很大。而颜色布局描述子(Color Layout Descriptor, CLD)^[12-15]可以描述视频序列中帧图像与帧图像之间的颜色分布情况,获得图像的特征信息。所以本文提出了融合CLD选取关键帧对关键帧进行三帧差法^[16]提取真实的背景模型的ViBe算法来抑制鬼影的出现,在运动目标运动速度不同的场景下检测步长参数对结合三帧差分的ViBe算法的影响,同时探索本文算法是否能弥补结合三帧差分的ViBe算法在不用运动目标速度的条件下需要改变步长以适应环境变化的缺点。在背景更新阶段通过加入形态学处理和自适应阈值相结合的方法消除动态背景干扰。

1 基本算法描述

1.1 ViBe算法

ViBe算法是一种非参数化聚类背景建模方法,与其他算法相比ViBe算法仅用第一帧图像就可以完成背景模型初始化。它的基本思想是:为每一个像素点创建一个背景样本集,将每一帧画面的各个像素点与其对应的背景样本集作匹配,判断该像素点是否为背景,有时由于光线的变化、相机的抖动等原因导致背景发生了变化,所以当检测中满足一定条件时,会对背景样本集进行更新。该算法的核心部分包括背景模型初始化、背景像素点匹配、背景模型更新和异常处理模块。具体流程如下:

1) ViBe算法利用一个像素点与周围其他像素点的灰度值相似特点,随机选取 N 个该像素点的周围八邻域像素值,为每一个像素点建立一个背景样本集来初始化背景模型。

2) 以2-D欧氏空间的半径 R 作为判断阈值,对视频中的每一帧图像的每一个像素点与对应的背景模型中的 N 个样本值比较来判断该像素点是否为背景像素点。统计在背景样本集中匹配为背景点的个数 n ,与所设定的个数匹配阈值 $\#_{\min}$ 作比较,若匹配数大于该阈值则表示该点为背景点,如式(1)所示:

$$b(x, y, t) = \begin{cases} 0, & n \geq \#_{\min} \\ 1, & n < \#_{\min} \end{cases} \quad (1)$$

式(1)表示在图像中 (x, y) 点在 t 时刻被判定为背景像素点; $b(x, y, t)=1$,表示在图像中 (x, y) 点在 t 时刻被判定为前景像素点;最终得到一个二值化图像的检测结果。

3) ViBe算法采用随机选取像素点的方法来更新背景像素模型,经过运动目标检测后如果检测出像素点 x 是背景像素,那么从样本集中随机选择一个样本用该像素点的像素值 $v(x)$ 代替。这种更新方式避免视频中细微的抖动对背景建模的影响,避免产生重影和误差,使目标检测更加准确。

4) 在实际场景中,背景不是一成不变的,由于受到环境的影响,背景的变化会影响到算法的判断,将变化了的背景判断为前景。ViBe算法会采用保守更新和前景计数点相结合的方法,即:判断为前景的像素点不会被填充进背景模型中,同时如果该像素点连续被判定为前景点的话,则将其更新为背景点。

1.2 颜色布局描述子

颜色布局描述子(CLD)^[13]是用来描述颜色在图像中空间分布的颜色描述符,以一种紧凑的形式,有效地表达了图像的颜色空间分布,计算代价小,提供了图像与图像的匹配和高速的片段与片段的匹配,常用于图像检索与视频摘要技术。本文使用CLD检测视频中不相似的帧画面,选取本文算法需要的关键帧画面。

CLD的提取分为4个阶段,分别是:图像分割、获取区域代表颜色、离散余弦变换(Discrete Cosine Transform, DCT)和锯齿扫描系数的非线性量化。如图1所示,首先将图像从RGB色彩模式(RGB color mode, RGB)映射到YCrCb空间中,用 8×8 栅栏把图像分割成64个区域块,计算每一块的所有像素点各个颜色分量(Y, Cr, Cb)的平均值作为该区域块的代表颜色,对每个区域块的代表颜色进行二维DCT。

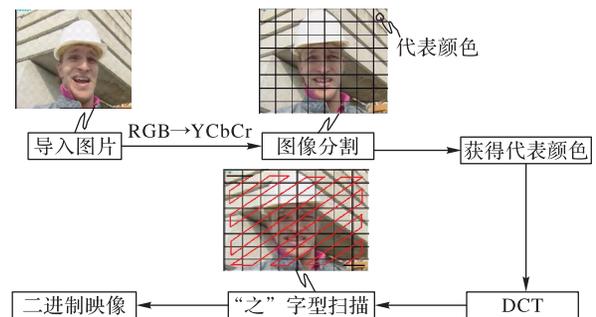


图1 颜色布局描述子原理图

Fig. 1 Principle diagram of color layout descriptor

然后对三组系数矩阵做之字扫描,由于经过了DCT,所以图像的主要信息全部集中在系数矩阵的左上角,矩阵的左上角是低频部分的集中区域,矩阵的右下角部分则是图像的冗余信息,找出低频系数进行量化,得到颜色布局特征参数。本文采用的是非线性量化,目的是减少系数矩阵的信息存储,使高频部分量化后趋于0。对于Y分量,取出6个低频系数分别为 $F(0,0)$ 、 $F(0,1)$ 、 $F(1,0)$ 、 $F(2,0)$ 、 $F(1,1)$ 、 $F(0,2)$,对于Cb, Cr颜色分量分别取 $F(0,0)$ 、 $F(0,1)$ 、 $F(1,0)$,3个颜色分量总共取得12个参数构成该图像的颜色局部描述符,图像间的相似程度用颜色布局描述符的特征向量的距离来判断。

$$D(img1, img2) = \sqrt{\sum_i w_Y (Y_{i, img1} - Y_{i, img2})^2 + \sum_i w_{Cb} (Cb_{i, img1} - Cb_{i, img2})^2 + \sum_i w_{Cr} (Cr_{i, img1} - Cr_{i, img2})^2} \quad (2)$$

其中: $D(img1, img2)$ 表示 $img1, img2$ 两个图像的相似性量度, $(Y_{i, img1}, Cb_{i, img1}, Cr_{i, img1})$ 和 $(Y_{i, img2}, Cb_{i, img2}, Cr_{i, img2})$ 分别是图像 $img1$ 与 $img2$ 的各颜色分量对应的第 i 个DCT系数。图像

越相似, $D(img1, img2)$ 的值越小。

2 基于 CLD 的 ViBe 算法改进

本文提出算法流程如图 2 所示。本文采用 CLD 描述子选取关键帧图像做差分运算, 得到监控视频中真实背景, 利用真实背景来建立背景采样样本集, 但是在监控的实际场景中, 背景并不是一成不变的, 所以使用自适应阈值判别检测误判背景, 使得动态背景下背景模型更加稳定。

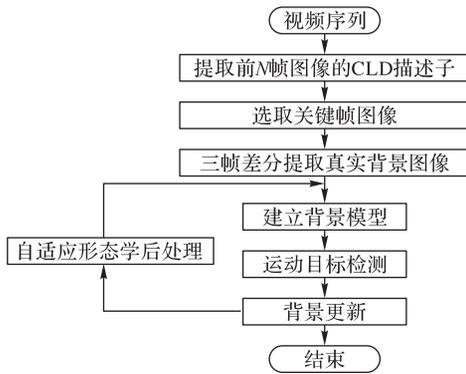


图 2 基于 CLD 的 ViBe 算法流程

Fig. 2 Flow chart of ViBe algorithm based on CLD

2.1 基于 CLD 描述子与三帧差分的背景初始化

由于 ViBe 算法选用的视频第一帧图像来建立背景模型, 所以当第一帧图像中有运动目标时, 会把运动目标像素点放入背景模型中当作背景样本值, 这样检测运动目标时, 会把运动目标误检为背景, 所以当该运动目标运动时, 会出现“鬼影”, 结合三帧差分的 ViBe 算法^[11]在一些场景下可以有效地解决这一问题; 但是这种方法得到的运动目标不连续, 会产生空洞, 而步长的选择又对其影响特别大, 步长过短时, 或者运动目标运动速度过慢, 帧差后的运动目标有可能发生重合, 因为像素点的像素值相差不多, 所以运动目标区域会有空白, 造成空洞, 甚至有些区域会产生新的“鬼影”。

本文提出采用 CLD 描述符来选取关键帧, 设定 CLD 处理帧数 N , 图像相似度阈值 a , 将提出的 3 个关键帧记作 $img1$ 、 $img2$ 、 $img3$, 作三帧差分, 得到背景图像。

1) 选取关键帧: 将视频图片前 N 帧图像的格式从 RGB 空间, 转换为 YCrCb 空间, 之后经过 DCT 和量化, 取得颜色局部描述符, 求得第 1 帧图像 $img1$ 的描述符与后面每一帧图像的相似性量度, 直到其值大于图像相似度阈值 a 小于 b , 获取该关键帧图像 $img2$, 再用相同的方法获取第三帧关键帧 $img3$ 图像。

2) 确定运动目标位置: 将获取的三帧关键帧图像作三帧差分运算, 即如式(3), 将得到的差分结果作“与”运算, 根据阈值 φ 来判断第 1 帧关键帧图像中的运动目标位置。

$$\begin{cases} Q1_{\text{minus}} = |img1 - img2| \\ Q2_{\text{minus}} = |img1 - img3| \end{cases} \quad (3)$$

$$Q1_{\text{minus}} \cap Q2_{\text{minus}} \quad (4)$$

3) 填充运动目标位置: 将 $img1$ 含有运动目标的位置用 $img3$ 中该位置填充, 得到的结果作为真实的背景图像。

4) 将得到的真实的背景作为检测运动目标的背景模型, 用其建立背景采样样本集, 从第一帧视频图像开始检测运动目标。

该算法具体流程如下所示:

算法 1 基于 CLD 的背景模型初始化。

输入: 视频序列; 阈值 a, φ ;

输出: 背景模型。

```

1) initialization (video,  $D(a, b)$ )
2) for  $i := 1:N$ ; //逐帧读取视频
3) 计算图像相似度量度:  $D(img1, img_i)$ 
4) for  $a < D(img1, img_i) < b$ 
5) print  $img1, img_i$  //输出满足条件的关键帧图像
6) end for
7) for  $i := i:N$ 
8) 再次计算图像相似度量度:  $D(img1, img_i)$ 
9) for  $a < D(img1, img_i) < b$ 
10) print  $img3 \leftarrow img_i$ 
11)  $Q1_{\text{minus}} \leftarrow img1 - img2$ 
12)  $Q2_{\text{minus}} \leftarrow img1 - img3$ 
13)  $Q(x, y) = Q1_{\text{minus}} \cap Q2_{\text{minus}}$ ;
//得到含有运动目标的所有像素点  $(x, y)$ 
14) if  $img2$  中位于  $img1$  中运动目标位置是背景)
15) then  $img2(x, y) = img1(x, y)$ 
//将  $img2$  中此位置的像素填充到  $img1$  中, 同时填充该像素的八邻域
End

```

2.2 动态背景下的背景更新策略

在实际应用的背景中, 有很多含有高频运动的背景即动态背景, 比如: 含有风吹动树叶的背景、阳光照射下水面波纹晃动的背景。在动态背景的场景下, 由于树叶或者水面像素的变化, 在检测计算中容易形成误判, 把这些动态背景的晃动误判为前景。在检测出来的背景处存在较多的白点现象, 这些白点包括噪声点以及受到动态背景影响的背景像素点。为了解决这一问题, 本文通过形态学处理, 通过开运算, 先腐蚀后膨胀来处理动态背景中小的噪声点, 将动态背景中小的连通域分离开, 平滑较大运动物体的边界轮廓且并不改变连通区域面积的大小, 然后通过计算前景连通域面积, 设立一个自适应的连通区域面积阈值, 即若连通区域面积小于当前帧中面积最大的前景连通区域面积的 $k\%$, 采用填补空洞的办法更新背景模型, 消除背景干扰点。该算法的具体实现流程如下所示。

算法 2 形态学后处理。

输入: 背景模型 $img2$, 检测后的二值化图像 p_i ;

输出: 更新的背景模型。

```

1) 定义函数 background update( $img2, p_i$ )
2) For  $i=1:N$ ;
3)  $p_i = \text{bw morph}(p_i, 'open')$ ;
//对检测后的二值化图像进行开运算处理
4)  $[L, num] = \text{bw label}(p_i, n)$ ; //对连通域进行标号
5)  $area = \text{region props}(num, 'Area')$ ;
//计算当前帧每个连通区域的面积
6)  $Ar = \text{cat}(1, area)$ 
7)  $ind \leftarrow \text{find}(Ar == \max(Ar))$  //读取到最大的连通区域标号
8) for  $m = 1:num$ 
9) If  $Ar(1, m) < Ar(1, ind) * k\%$ ;
//该连通区域的面积小于最大连通区域面积的  $k\%$ 
10) Then  $p_i[L(x, y)] = 1$ ;
//丢弃这个连通域, 将其像素值改为 1, 将第  $m$  区域更新为背景
11) Else 该连通域像素值不变
12) End

```

3 实验对比

本文实验是在系统类型为64位操作系统的Windows7联想台式电脑,内存为8GB的环境下实现的,并且本文算法是在编程运行环境为MatlabR2016a进行编写的。同时本章将本文算法与Van-Droogenbroeck和Olivier Barnich提出的ViBe算法和结合三帧差分法的ViBe算法进行实验对比。本实验中相关参数的选择如下:背景样本集中样本个数 $N=20$,样本集匹配点的个数阈值 $R=20$, $\#_{\text{min}}=2$,更新采样概率为16,真实背景检测阈值 $\varphi=30$,相似度阈值 $a=0.72$, $b=0.67$ 。本文分别在单运动目标场景、多运动目标场景、含有动态背景的场景对本文算法进行测试。将原始ViBe算法、结合三帧差分的ViBe算法与本文算法在鬼影消除方面进行比较,重点体现本文算法在运动目标运动速度不同的情况下有良好的鲁棒性,弥补结合三帧差分的ViBe算法在环境不同的情况下需要调整参数大小以适应其变化的缺点;将本文算法在动态背景下进行测试,并对其结果与单高斯背景建模、多高斯背景建模、原始ViBe算法进行对比分析。

3.1 鬼影消除策略实验研究

如图3所示,分别是ViBe算法、结合三帧差分的ViBe算法以及本文算法提取的背景模型以及检测结果。第一行所示的三张图片分别为ViBe算法、结合三帧差分的ViBe算法、本文算法所提取的背景模型。如图3(a)是原始视频第一帧图像,也是ViBe算法所选取的背景模型,该背景模型表现出ViBe算法第一帧建模的特点,保留了第一帧图像所包含的所有前景信息;图3(b)是结合三帧差分的ViBe算法提取的背景模型,通过算法的优化,该算法提取的背景模型虽然减少了前景信息量,但是仍然包含少量运动目标信息;图3(c)是通过本文算法提取出的背景模型,前两者算法所提取的背景模型都包含着运动目标信息,而本文算法提取到的背景模型完整且清晰,实验结果表明由于本文算法可以提取出完整、准确的背景模型,所以为以后的检测的准确性提供了保障。

图3(d)(e)(f)是ViBe算法、结合三帧差分的ViBe算法以及本文算法对原始视频第80帧图像进行运动目标检测的实验结果,实验表明本文算法可以完全避免“鬼影”的出现。图3(d)为ViBe算法对运动目标检测实验结果,图3(a)中ViBe算法提取的背景模型中的运动目标即黑色矩形框中的位置,被记录为背景信息,导致当运动目标运动时,运动目标所在位置的像素点被记录为前景信息,造成误判,形成图3(d)中白色

矩形框中的“鬼影”;图3(e)是结合三帧差分ViBe算法检测运动目标的实验结果,由于做帧差时,运动目标位置发生重叠,导致提取背景图像时,会把另一帧运动目标的位置替换到背景图相应位置,所以建立背景模型不准确,同样将部分运动目标信息读入背景模型中,白色矩形框中的导致“鬼影”出现;图3(d)是本文算法提取完整背景模型,成功消除了“鬼影”,得到真实运动目标,且具有良好的鲁棒性,不容易受到噪声干扰。

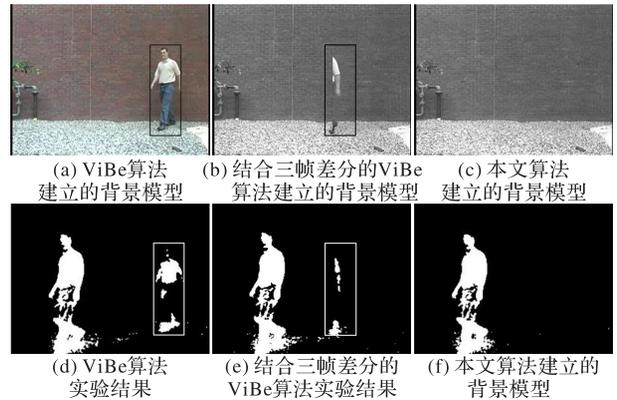


图3 三种算法在单目标场景下的实验结果

Fig. 3 Experimental results of three algorithms in single target scenarios

图4是在多目标运动的场景下对ViBe算法、结合三帧差分的ViBe算法以及本文算法进行测试,可以看到本文改进后的算法在运动目标多且混乱的情况下,仍然可以准确检测出运动目标位置。图4(a)为原视频原始图片,图4(b)为本文算法提取到的背景模型,可以看出本文算法提取到了完整的背景信息且不包含任何运动目标信息;图4(c)中右上角有许多误判的运动目标,由于运动目标离镜头远,运动目标面积小,且运动目标运动缓慢,造成有许多鬼影出现在检测画面中,所以在检测出来的运动目标中无法判断哪些是真正的运动目标,哪些是“鬼影”。所以该算法在多目标且环境复杂的场景下容易发生检测混乱;图4(d)和图4(e)效果相差不多,在检测运动速度较快的运动目标时,鬼影消除效果比较好,但是在检测图片的左上角处,存在运动目标运动缓慢的运动目标,所以仍然出现大量“鬼影”;图4(e)可以清楚地看出本文算法通过CLD描述子选取的关键帧做三帧差分还原出来的背景要比原ViBe算法和结合三帧差分的ViBe算法还原的背景模型要真实、可靠、准确,所以在检测过程中彻底消除了鬼影,提取出的运动目标清晰准确。

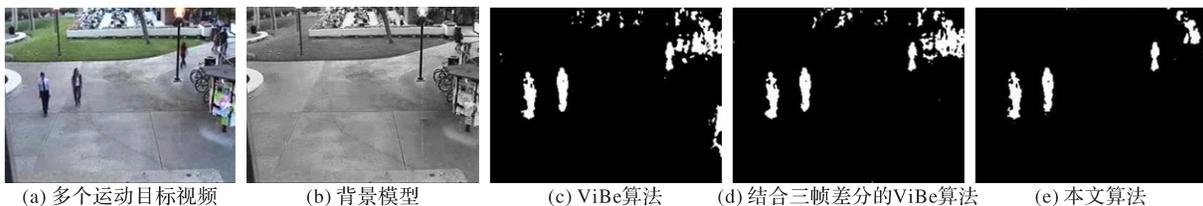


图4 三种算法在多目标场景下的实验结果

Fig. 4 Experimental results of three algorithms in multi-target scenarios

图5是在运动目标运动速度不同的情况下对算法的准确性进行测试。从图5的实验结果可以明显地看出,运动目标运动速率的不同,会影响算法的准确度。对于原ViBe算法,只要第一帧视频有运动目标,ViBe算法百分之百会检测出鬼

影,造成误判,所以运动目标运动快慢对该算法无显著影响。

图5测试结果显示,结合三帧差分的ViBe算法在运动目标速度慢的时候,即在行人的测试视频场景下,其检测结果与ViBe算法相同,当运动目标速度快时,即在城市汽车测试视

频场景下时,其“鬼影”的像素面积比 ViBe 算法产生的“鬼影”像素面积减少将近一半;在高速汽车的测试视频场景下,其检测结果与本文算法检测结果一致,且不产生“鬼影”。

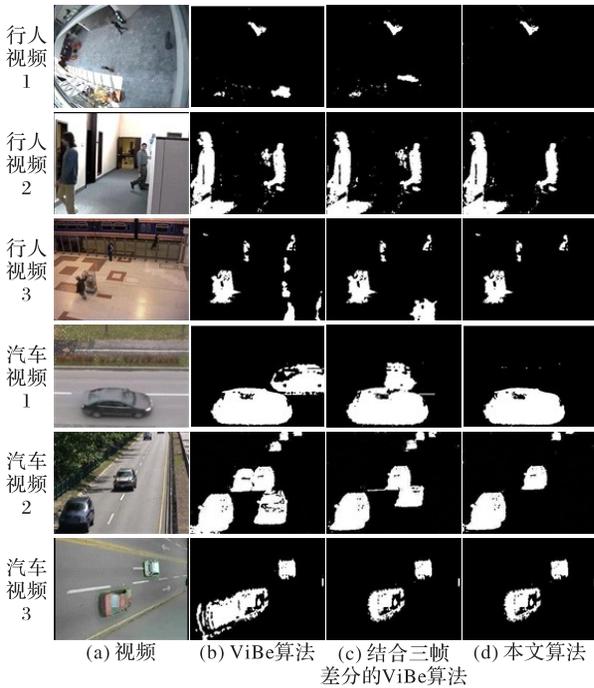


图5 三种算法在不同运动目标速度下的实验结果

Fig. 5 Experimental results of three algorithms at different moving target speed

从图6(a)中不难看出运动目标运动速度对结合三帧差的算法检测的准确率有很大的影响。

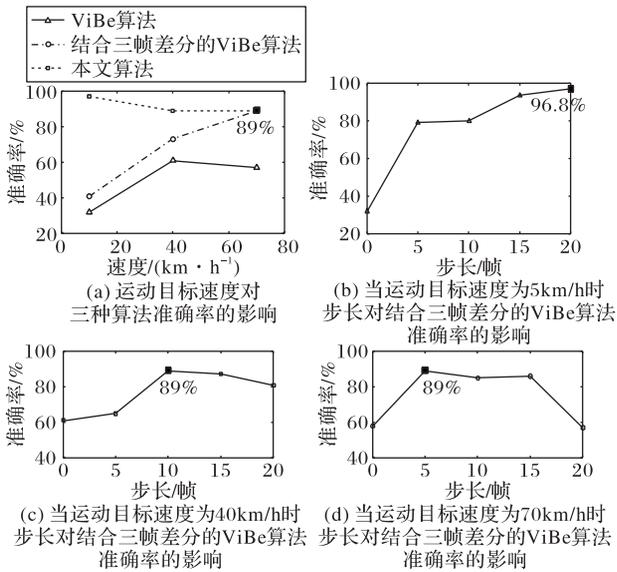


图6 运动目标速度与步长对结合三帧差分的ViBe算法的影响
Fig. 6 Effect of moving target speed and step size on ViBe algorithm combined with three frame difference

随着运动目标速度的变化,运动目标运动速度越快,那么该算法的准确率呈上升趋势,当运动目标速度达到70 km/h的时候,结合三帧差分的ViBe算法与本文算法的准确度达到一致,但是运动目标运动速度越慢,就越容易检测出鬼影,造成算法对前景的误判,影响ViBe算法准确度。而本文算法无论运动目标运动速度的快慢都有很好的鲁棒性,可以完整检测

出运动目标且完全避免鬼影的出现。图6(b)(c)(d)是结合三帧差分的ViBe算法在运动目标运动速度不同的情况下步长的选择对运动目标检测准确率的影响。实验表明,如图6(b)所示在运动目标在5 km/h的速度下运动,结合三帧差分的ViBe算法步长选取20帧时,其准确率可以与本文算法在运动目标速度为5 km/h时的准确率一致;图6(c)表明在运动目标速度为40 km/h时,结合三帧差分的ViBe算法步长选取10帧可以使准确率达到最大值与本文算法在运动目标速度为40 km/h时的准确率一致;图6(d)表明当运动目标速度为70 km/h时,步长选为5帧时,结合三帧差分的ViBe算法的准确率等于本文算法在运动目标为70 km/h时的准确率。经过两种算法的实验对比表明,结合三帧差分的ViBe算法只有在步长与运动目标速度达到一定条件时,其算法准确率才可以与本文算法的准确率一致,而本文算法不需要协调步长参数与运动目标速度就可以达到最大准确率。

3.2 针对运动目标的不同速度下算法参数的最优选择

表1、表2是在公共区域监控视频场景中运动目标运动速度不同的情况下,相似度量阈值的选择对准确率的影响。相似度量阈值决定了关键帧的选取,同时也决定了运动目标检测的准确率。经过6组行人测试视频与6组汽车测试视频的实验表明,本文算法对相似度量阈值的选取十分敏感,偏大或偏小都会影响运动目标检测的准确度。其中行人测试视频中的视频1和视频2是在公园里行走的行人的视频;视频3广场上来回行走的人群与骑着自行车的人的视频;视频4、5是校园教学楼走廊上下课的学生,视频6是实验楼大厅行走的研究人员。当本文算法的相似度量阈值选择为 $a=0.62, b=0.67$ 时,对6组基于行人的室外监控视频进行运动目标检测,得到的准确率最大值为94.4%,最小值为41.1%,虽然在检测视频3时,其准确率大于 $a=0.67, b=0.72$ 时对该视频检测的准确度,但是该阈值仅适用于视频3,其他视频的准确度并未有明显优势。当 $a=0.72, b=0.78$ 时,其检测后的准确率最高94.6%,最低48.4%,该阈值的设定对视频2、6有良好的适应力;但是对于视频5、6选择 $a=0.67, b=0.72$,可以使检测结果准确率更高。实验结果表明,本文算法选取相似量度阈值 $a=0.67, b=0.72$ 时,在室外行人运动的场景下的运动目标检测准确率更高、更稳定。

对基于汽车行驶的视频,同样用本文算法做相同的实验,在汽车测试视频中,视频1、2是汽车在城市主城区正常行驶的视频;视频3、4是在汽车在高架桥上行驶的测试视频;视频5、6是汽车在高速公路上行驶的测试视频。当相似度量阈值为 $a=0.62, b=0.67$ 时,检测准确率最大值为92.8%,最小值为40%,该阈值适用于汽车测试视频4,但对于其他测试视频效果并不理想;相似度量阈值为 $a=0.72, b=0.78$ 时,检测结果准确率最大值为93.3%,最小值为52.8%,在6段测试视频中,仅对视频4有不错的效果,不适用于测试视频中所有的场景。与上一组实验相同,当相似量度阈值为 $a=0.67, b=0.72$ 时,检测结果的准确率高且稳定。由两组实验可知在室外场景下,本文算法选择相似度量阈值为 $a=0.67, b=0.72$ 时,可是使本文算法的准确率性能达到最优。同时运动目标速度不同而选取的阈值相同进一步证明了本文算法的准确率只与相似量度阈值有关,运动目标的运动速度并不影响相似量度阈值的选择。经过两组实验测试表明本文所选取的阈值 $a=0.67, b=0.72$ 在不同的运动目标的运动速度下具有良好的稳定性。

表 1 本文算法在行人视频下选择不同相似度量阈值的准确率

Tab. 1 Accuracy of the proposed algorithm to select different similarity measurement thresholds under pedestrian video

相似度量阈值		不同行人视频的检测准确率/%					
<i>a</i>	<i>b</i>	视频 1	视频 2	视频 3	视频 4	视频 5	视频 6
0.62	0.67	62.3	78.2	94.4	56.3	82.2	41.1
0.67	0.72	92.0	97.1	89.2	87.2	93.7	95.6
0.72	0.78	62.7	94.6	59.0	48.4	57.8	81.0

表 2 本文算法在汽车视频下选择不同相似度量阈值的准确率

Tab. 2 Accuracy of the proposed algorithm to select different similarity measurement thresholds under car video

相似度量阈值		不同汽车视频的检测准确率/%					
<i>a</i>	<i>b</i>	视频 1	视频 2	视频 3	视频 4	视频 5	视频 6
0.62	0.67	62.3	56.1	68.5	92.8	78.4	40.0
0.67	0.72	91.7	74.6	90.1	99.4	93.8	95.0
0.72	0.78	62.7	69.2	52.8	93.3	69.4	69.1

3.3 动态背景下的背景更新策略实验研究

由于在动态背景下很难建立一个稳定的背景模型,所以在动态背景下进行运动目标检测的效果不好,许多算法都不

能很好地处理这个问题,本文所选取的视频是分别是具有水面波纹、喷泉以及晃动的树叶的动态背景视频。图 7 是在动态背景下对单高斯背景建模、多高斯背景建模、ViBe 算法与本文算法进行实验对比,检测算法在动态背景下的适应度。通过实验可以看出多高斯和单高斯在水面波纹晃动的环境下都会有误检的情况,单高斯背景建模虽然在动态背景下有很好的适应力,可是在检测运动目标时,有严重的空洞现象,而且容易受到噪声的影响;多高斯背景建模的背景建模初始化速度和背景更新的计算方法复杂,计算速度慢,虽然不会有空洞的现象,但是很容易受动态背景的影响,而且不具有抗噪性,检测质量很差;ViBe 算法由于它是第一帧图像背景建模所以很难适应动态背景的变化,检测结果所误检信息不仅包含因为动态背景变化所持有的背景信息还包括产生的“鬼影”;而本文算法在动态背景下仍然会稳定检测出前景信息,所以会产生较多且面积大的误检区域。图 7(f)的实验是针对动态背景情况下提出的本文算法,经过本文所提出的形态学后处理可以有效地去除动态背景对 ViBe 算法的干扰,鬼影和误检的虚警几乎为零,但是也有前景被误检为背景的地方,而且还有空洞产生,但是整体来看,本文算法大大提高了检测的准确性和清晰度,其抗噪性和鲁棒性皆优于其他三种算法,检测出的结果更接近真实的运动目标。

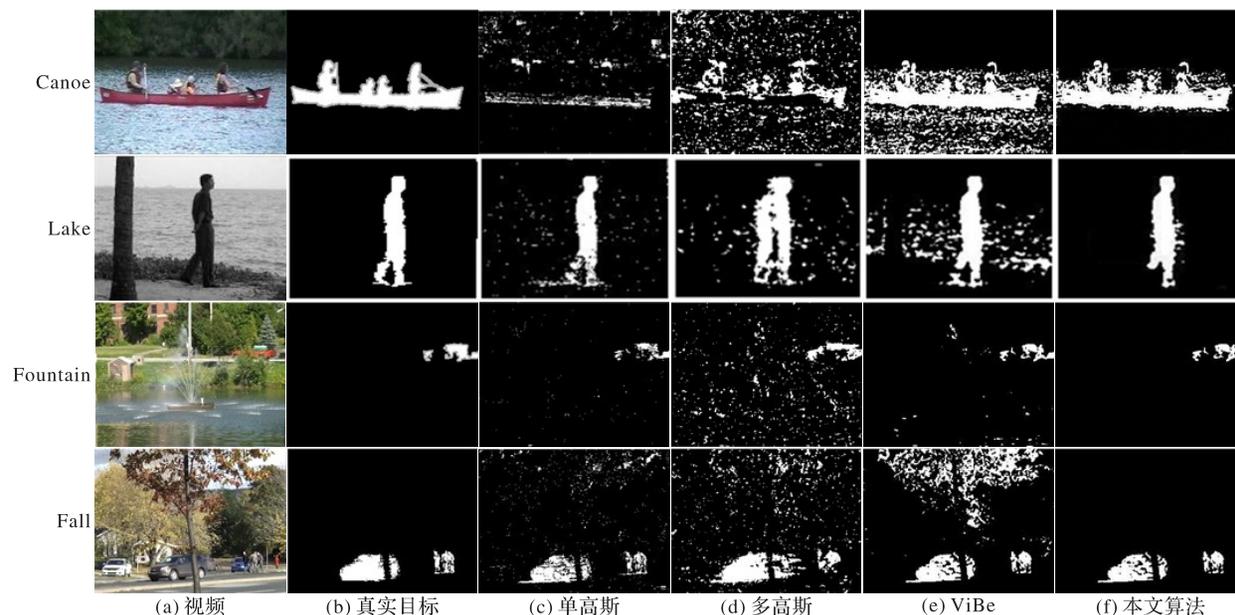


图 7 动态背景下不同算法的实验结果对比

Fig. 7 Experimental results comparison of different algorithms in dynamic background

3.4 时间复杂度对比实验

ViBe 算法是一种像素级的背景建模方法,它的初始帧建模的特点相对于其他算法就有初始化速度快的优点,并且 ViBe 算法的背景更新策略是以像素为单位进行更新,如果该算法在背景更新中将全部像素点更新,则更新部分的算法时间复杂度为 $O(\text{height} * \text{width})$,但是 ViBe 算法更新策略中,每一个点都有更新的概率,所以更新部分的时间复杂度就会降为 $O(1/\text{rate} * \text{height} * \text{width})$,一般 *rate* 的值取 16。经过对 ViBe 算法的分析,ViBe 算法初始帧建模的特点使其在该部分的代码时间复杂度为 $O(1)$,在 ViBe 算法中该复杂度可以忽略不计,估计出 ViBe 算法在目标检测时的时间复杂度为 $O[1 + (n * \text{height} * \text{width})^3 + n * 1/\text{rate} * \text{height} * \text{width}]$ 。根据时间复杂度的规则次数小的忽略,则 ViBe 算法最后的时间复杂度为 $O[(\text{height} * \text{width})^3]$ 。结合三帧差分的 ViBe 算法为了改善原始 ViBe 的“鬼影”的缺点,在背景建模部分做出了改进,将初

始帧建模,改为结合三帧差分建模,根据步长选择三帧帧图像做差分运算,得到的差分结果做填充,该算法在背景建模部分的改进并不涉及循环,故其时间复杂度仍然与原始 ViBe 算法保持一致,为 $O[(n * \text{height} * \text{width})^3]$ 。本文算法在结合三帧差分的 ViBe 的基础上加入了用 CLD 提取关键帧的方法,并且在 ViBe 算法的更新策略上加入了形态学处理以及设立自适应阈值的方法来让本文算法更好地适应动态背景的变化。由于在背景建模阶段加入的循环次数是根据图像相似程度来决定的,并且寻找关键帧图像时做出循环处理,该部分的循环次数不会超过视频总帧数 *n*,所以在本文算法中该部分的时间复杂度忽略不计,本文算法的时间复杂度即同样为 $O[(n * \text{height} * \text{width})^3]$ 。

通过对 ViBe 算法、结合三帧差分的 ViBe 算法以及本文算法的时间复杂度的估计与分析,可以发现三种算法时间复杂度是同一数量级的,并且都是由视频帧数以及像素值来决定的,只通

过估计时间复杂度无法比较三者的效率,但是算法的运行时间可以更直观地说明算法的效率与复杂度。表3是ViBe算法、结合三帧差分的ViBe算法以及本文算法在三种规格的视频下的运行时间比较,从中可以看出随着帧数的增多三种算法的运行时间都在增长,且结合三帧差分的ViBe算法以及本文算法的运行时间都高于原始ViBe算法,但是相差并不多。就此说明,随着本文算法抑制“鬼影”的性能以及适应动态背景的性能的提高,算法的运行时间也在增多。

表3 三种算法的运行时间对比

Tab. 3 Comparison of running time of three algorithms

视频信息		运行时间/s		
帧数 n	像素值	结合		
	($height^*$ $width$)	ViBe算法	三帧差分的 ViBe算法	本文算法
80	240×320	10.265 2	12.335 6	13.445 1
300	240×320	38.263 2	40.322 5	44.256 1
374	176×320	42.185 4	44.296 5	47.553 3

4 结语

本文提出了融合提取关键帧改进的ViBe算法,在原始的ViBe算法融合CLD描述子提取关键帧结合三帧差分法建立背景模型,以及加入自适应阈值的策略来控制动态背景下的干扰,完全避免了ViBe算法鬼影的出现,准确检测出运动目标的位置以及运动信息,大大降低了误检率,可以很好地适应动态背景。本文分别在单目标、多目标、运动目标运动速度不同和动态背景等多种场景下对此算法进行测试,与结合三帧差分的ViBe算法进行对比,实验表明在大多数情况下本文算法的准确率要优于结合三帧差分的ViBe算法。在运动目标运动速度快时,结合三帧差分法的ViBe算法的准确率会与本文算法准确率达到一致。经过在具有动态背景下的视频对单高斯背景建模、多高斯背景建模、ViBe算法和本文提出的算法进行测试,可以得出本文算法避免检测过程中虽然会出现空洞、残缺现象,但是大体来看,在动态背景下与其他三种算法相比更能克服动态背景的干扰,提高了原始ViBe算法的抗噪性和鲁棒性。

参考文献 (References)

- [1] PRITCH Y, RAV-ACHA A, PELEG S. Nonchronological video synopsis and indexing [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(11): 1971-1984.
- [2] PRITCH Y, RAV-ACHA A, GUTMAN A, et al. Webcam synopsis: peeking around the world [C]// *Proceedings of the IEEE 11th International Conference on Computer Vision*. Piscataway: IEEE, 2007: 1-8.
- [3] RAV-ACHA A, PRITCH Y, PELEG S. Making a long video short: dynamic video synopsis [C]// *Proceedings of the 2006 IEEE International Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 2006: 435-441.
- [4] 张玲,陈丽敏,何伟,等. 基于视频的改进帧差法在车流量检测中的应用 [J]. *重庆大学学报 (自然科学版)*, 2004, 27(5): 31-33, 73. (ZHANG L, CHEN L M, HE W, et al. Application of an improved frame-difference method based on video in traffic flow measurement [J]. *Journal of Chongqing University (Natural Science Edition)*, 2004, 27(5): 31-33, 73.)
- [5] GOYETTE N, JODOIN P M, PORIKLI F, et al. Changedetection.net: a new change detection benchmark dataset [C]// *Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Piscataway: IEEE: 2012: 1-8.
- [6] STAUFFER C, GRIMSON W E L. Adaptive background mixture models for real-time tracking [C]// *Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society: 1999: 246-252.
- [7] CHIU C C, KU M Y, LIANG L W. A robust object segmentation system using a probability-based background extraction algorithm [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(4): 518-528.
- [8] BARNICH O, VAN DROOGENBROECK M. ViBe: a powerful random technique to estimate the background in video sequences [C]// *Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. Piscataway: IEEE, 2009: 945-948.
- [9] BARNICH O, VAN DROOGENBROECK M. ViBe: a universal background subtraction algorithm for video sequences [J]. *IEEE Transactions on Image Processing*, 2011, 20(6): 1709-1724.
- [10] 张红瑞,马永杰. 融合混合高斯模型的改进的ViBe算法 [J]. *西北大学学报 (自然科学版)*, 2017, 47(3): 361-366. (ZHANG H R, MA Y J. An improved Vibe algorithm based on Gaussian mixture model [J]. *Journal of Northwestern University (Natural Science Edition)*, 2017, 47(3): 361-366.)
- [11] 胡小冉,孙涵. 一种新的基于ViBe的运动目标检测方法 [J]. *计算机科学*, 2014, 41(2): 149-152. (HU X R, SUN H. Novel moving object detection method based on ViBe [J]. *Computer Science*, 2014, 41(2): 149-152.)
- [12] FU S C, SIKORA T, PURL A. Overview of the MPEG-7 standard [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(6): 688-695.
- [13] KASUTANI E, YAMADA A. The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval [C]// *Proceedings of the 2001 International Conference on Image Processing* Piscataway: IEEE, 2001: 674-677.
- [14] MANJUNATH B S, OHM J R, VASUDEVAN V V, et al. Color and texture descriptors [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(6): 703-715.
- [15] YAMADA A. Results of core experiment on compact color layout descriptor [S]. [S. l.]: ISO/IEC JTC1/SC29/WG11 (MPEG), 1999: document No. M6446.
- [16] NIU L, JIANG N. A moving objects detection algorithm based on improved background subtraction [C]// *Proceedings of the 8th International Conference on Intelligent Systems Design and Applications*. Piscataway: IEEE, 2008: 604-607.

This work is partially supported by the National Key R&D Program of China (2018YFF0301004), the National Natural Science Foundation of China (61802107), the Ministry of Education-China Mobile Research Fund Project (MCM20170204), the Jiangsu Postdoctoral Research Support Program (1601085C).

WANG Tong, born in 1995, M. S. candidate. Her research interests include artificial intelligence, image processing.

WANG Wei, born in 1983, Ph. D., associate professor. His research interests include public safety Internet of things, implicit human computer interaction.

CUI Yihao, born in 1994, M. S. candidate. His research interests include swarm intelligence algorithm, artificial intelligence, embedded system.

ZHU Tianyu, born in 1996, M. S. candidate. His research interests include unmanned aerial vehicle, unmanned driving.