

# 基于改进视觉背景提取的运动目标检测算法

莫邵文<sup>1</sup> 邓新蒲<sup>1</sup> 王 帅<sup>1</sup> 江 丹<sup>1</sup> 祝周鹏<sup>2</sup>

<sup>1</sup>国防科学技术大学电子科学与工程学院, 湖南 长沙 410073

<sup>2</sup>中国人民解放军 61541 部队, 北京 100094

**摘要** 针对视觉背景提取算法(ViBe)中出现的鬼影问题、不能很好适应背景高频扰动和摄像机抖动问题以及由于采用空间邻域扩散机制引起背景更新错误问题,提出一种改进的视觉背景提取算法。该算法结合视觉显著性判断背景模型中存在的鬼影目标,通过判断背景模型中每个像素点的鬼影程度,结合模糊准则自适应改变时间子采样因子,加快消除鬼影的速度;通过建立一个闪烁程度矩阵,判断背景高频扰动程度来设置自适应匹配阈值,加入小目标丢弃和空洞填充策略;统计前景像素 24 邻域区域的像素点个数,判断前景像素点是否为摄像机抖动或者背景更新错误引起的噪点,提高算法的稳健性。结果表明,改进后的算法可以很好地弥补经典 ViBe 算法的不足,准确率与识别率等指标均大大提升。

**关键词** 机器视觉; 运动目标检测; 视觉背景提取; 鬼影消除; 视觉显著性; 动态背景; 自适应阈值

**中图分类号** TP391.4 **文献标识码** A

**doi:** 10.3788/AOS201636.0615001

## Moving Object Detection Algorithm Based on Improved Visual Background Extractor

Mo Shaowen<sup>1</sup> Deng Xinpu<sup>1</sup> Wang Shuai<sup>1</sup> Jiang Dan<sup>1</sup> Zhu Zhoupeng<sup>2</sup>

<sup>1</sup> College of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan 410073, China

<sup>2</sup> Unit 61541 of the Chinese People's Liberation Army, Beijing 100094, China

**Abstract** Aiming at the problems of ghost, background turbulence in high frequency, camera jitter and error of background update caused by spatial propagation technique in classic visual background extractor(ViBe) algorithm, an improved ViBe algorithm is proposed. Combined with visual saliency, the new method determines whether the ghost target exists in the background model or not, and adaptively changes the time subsampling factor through the level of ghost for each pixel in the background model, which can improve the rate of ghost elimination. Self-adaptive threshold is adopted in the process of model matching by establishing a blinking degree matrix to judge the high-frequency disturbance level of background, so that the background model is better suitable for the dynamic background. Small object discard and hole filling strategies are added to the new method. It can judge if a foreground pixel is a noise point caused by camera jitter or an error of background update by counting pixel numbers in 24-connected neighboring region of foreground pixels. Therefore, it can improve the robustness of the algorithm. Experiments demonstrate that the improved algorithm is a good way to make up for the deficiency of the original ViBe algorithm. The accuracy and recognition rate are improved greatly.

**Key words** machine vision; moving object detection; visual background extractor; ghost elimination; visual saliency; dynamic background; adaptive threshold

**OCIS codes** 150.1135; 040.1880; 040.1520

收稿日期: 2016-01-21; 收到修改稿日期: 2016-02-01

基金项目: 国家自然科学基金(61401474)

作者简介: 莫邵文(1992—),男,硕士研究生,主要从事空间信息获取与处理、图像处理方面的研究。

E-mail: msw1992@163.com

导师简介: 邓新蒲(1966—),男,博士,副教授,硕士生导师,主要从事空间信息获取与处理方面的研究。

E-mail: drden@sina.com

## 1 引 言

运动目标检测技术是计算机视觉和数字图像处理领域的一个重要研究内容,在智能视频监控等领域有着广泛的应用。运动目标检测的目的就是从图像序列中将运动区域从背景图像中提取出来。运动区域的有效分割对于目标跟踪、分类和行为理解等后续处理至关重要,因此目标检测的处理效果和速度直接影响到整个智能视频监控系统的准确性和稳定性。但在复杂的现场环境中,诸多干扰因素,如动态背景、摄像机抖动、鬼影、阴影等,给该技术带来了很大的挑战。

目前常用的运动目标检测算法有背景建模法<sup>[1]</sup>、光流法<sup>[2]</sup>、帧间差分法<sup>[3]</sup>,其中光流法和帧间差分法都基于各自的核心算法思想,已经形成完整的算法体系。背景建模法的核心思想是从当前帧图像中减去已经提前建立好的背景图像,而背景图像通常是指不包含任何运动目标的视频序列的背景。基于 Piccardi<sup>[4]</sup>的研究,一个实用的背景建模方法必须能够适应快速的光照变化、背景抖动(摄像机抖动)、复杂背景(背景包含树叶抖动、水纹波动)和背景场景变化(停车)等现象。现有的背景建模算法大致分为两类<sup>[5]</sup>:1)参数化的背景模型,该类方法为每个像素点建立一个参数模型用以估计背景,目前为止最受欢迎的是混合高斯模型(GMM)<sup>[6]</sup>,混合高斯背景建模是基于像素样本统计信息的背景表示方法,可以对复杂动态背景进行建模,但是该算法复杂度高,且在无规律变化的场景中有较高的虚警率;2)基于非参数化、像素样本的背景模型,如核密度估计方法<sup>[7-9]</sup>,此类方法由于依赖过去观察到的像素值,所以可以巧妙地避免参数估计这一步骤,然而由于这些方法采用先进先出的方式更新它们的像素模型,所以它们不能处理速度变化的目标。码本(CodeBook)<sup>[11-12]</sup>算法是一种压缩样本的背景提取算法,能够处理一些重复性的变化,但它仍然容易受到像刮风引起的树木摇晃这样无规则性运动的影响。

Barnich 等<sup>[13-14]</sup>提出了基于像素的非参数化随机样本模型(ViBe),它采用像素样本值建立背景模型,先将检测帧的像素值与对应模型进行匹配,再通过与固定阈值的比较来判断是属于前景还是属于背景,最后将被判为背景的像素点及其邻域点的背景模型以一定概率进行随机更新。该算法具有计算量小、内存占用少、处理速度快、检测效果好和应对噪声稳定可靠等特点,非常适合嵌入照相机等小型设备中。但是传统的 ViBe 算法仍然存在初始化时容易引入鬼影、阴影,无法适应动态背景和摄像机抖动以及在背景更新时采用空间邻域扩散机制容易引起更新错误的问题。针对这些缺点,提出一种改进的视觉背景提取算法。

本文从以下 4 个方面对 ViBe 算法进行了改进:1)结合背景模型的视觉显著性,判断前景是否为鬼影,再根据背景模型中每个像素点鬼影程度的高低自适应地改变时间子采样因子,加快消除鬼影的速度;2)通过建立一个闪烁程度矩阵判断背景高频扰动程度,将匹配阈值自适应赋值,最大程度地消除动态背景的影响;3)统计前景像素 24 邻域区域的像素点个数,判断前景像素点是否为摄像机抖动导致背景更新错误引起的噪点,将噪点消除的同时最大限度地保留检测到的正确前景形状,而且可以进一步消除动态背景带来的影响;4)采用小目标丢弃与空洞修补策略,使得检测结果更加准确。但是这里不对阴影做特别的处理。虽然现在有许多算法研究了阴影造成的问题<sup>[15-16]</sup>,但是对如何合适地处理阴影的问题存在争议。由于阴影具有物理、形态和时间特征,好的算法应该同时考虑阴影的这 3 个特征,忽略阴影带来的影响同时把阴影看作前景。通过仿真证明,改进后的算法可以有效地消除鬼影,消除动态背景、摄像机抖动和背景更新错误带来的负面影响,提高检测准确率和降低虚警率。

## 2 ViBe 算法描述

ViBe 建模算法是一种通过收集背景样本来建立背景模型的技术,主要由三部分组成:背景像素模型建立及分类过程、像素模型初始化方法和背景模型的更新策略。

### 2.1 背景像素模型建立及分类过程

ViBe 算法把前景检测问题看成一个分类问题,背景模型里存储一组由图像采样得到的样本值,将检测帧的像素值和背景模型里的样本值进行比较来判断此像素是前景或背景。假设  $v(x)$  为图像中位于位置  $x$  的像素值, $v_i$  代表在给定的欧式颜色空间中索引为  $i$  的背景样本值。每个背景像素  $x$  通过  $N$  个背景样本值的集合进行建模,该集合定义为

$$M(x) = \{v_1, v_2, \dots, v_{N-1}, v_N\}. \quad (1)$$

若以  $v(x)$  为球心、 $r$  为半径的球  $S_r[v(x)]$  与  $M(x)$  的交集满足一定的基数阈值  $R_\#$ , 即

$$\{S_r[v(x)] \cap \{v_1, v_2, \dots, v_N\}\} \in R_\#, \quad (2)$$

则认为  $v(x)$  是背景像素。

图 1 描述在 2 维(2D)欧式彩色空间  $(C_1, C_2)$  [13] 内一个像素的背景样本集与一个新的像素  $v(x)$  的对比过程。

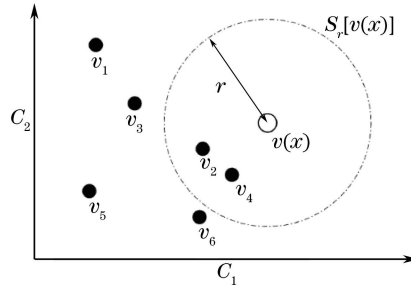


图 1 二维欧式空间匹配示意图

Fig. 1 Diagram of model matching system in two-dimensional European space

## 2.2 背景模型初始化方法

ViBe 算法的背景模型初始化可以看作像素模型样本值的填充过程。由于单帧图像中缺乏时域信息, 所以假设邻域像素具有相似的时空分布[17]。基于这个假设, 算法随机地选择像素点的邻居点的像素值作为它的初始模型样本值。背景模型初始化的数学模型为

$$M^0(x) = \{v^0[y \mid y \in N_G(x)]\}, \quad (3)$$

式中  $M^0(x)$  为第 1 帧初始化的背景模型集合,  $N_G(x)$  为像素点  $x$  的具有相似时空分布的邻居集合,  $y$  为像素  $x$  的领域中随机选择的一个邻居像素,  $v^0(y)$  为第 1 帧所选取的像素点的像素值。

初始化方法的优点是对噪声和光照灵敏可以即时地对目标进行检测, 但是容易引入鬼影区域。

## 2.3 背景模型的更新策略

ViBe 算法的更新策略采用保守的更新策略加空间邻域扩散机制的方法, 应该包含以下 3 个重要的部分: 1) 无记忆的更新策略; 2) 通过随机的时间重采样来扩展被前景像素遮挡的背景像素; 3) 通过扩展背景像素空间的机制来保证空间一致性。该方法是当一个像素点被确定为背景点时, 使匹配上的像素值以  $1/\phi$  的概率替代背景样本中的任一样本值( $\phi$  称为时间子抽样因子), 同时采用空间信息邻域扩散机制将背景信息扩散到相邻像素, 在更新匹配像素的同时对其邻域像素的背景模型进行随机更新。

这种样本邻域传播机制有效保证了空间一致性, 但对于前景、背景交界处的像素会造成信息的错误传播, 产生误检。

# 3 改进的 ViBe 算法

## 3.1 鬼影消除

“鬼影”是一系列连通的、被检测为运动目标但实际上不是任何运动目标的像素点[18]。在这种特别的情况下, “鬼影”的产生是由于把运动目标作为背景抽样得到的不正确的初始化结果。ViBe 算法是利用空间邻域扩散机制消除鬼影, 但是此机制消除鬼影的速度过慢, 尤其是在时间子采样因子设为较大默认值的时候。

结合背景样本中存在的视觉显著性来判断鬼影的存在, 进而由模糊准则进一步设置自适应时间子采样因子, 加快背景更新的速度, 消除鬼影。显著性的概念是在人们研究人类视觉感知机制的过程中提出来的[19-20], 可以简单描述为场景中的元素吸引视觉注意力的能力。这种能力通常是指目标因为具有某些属性而拥有能够比较容易与周边的环境区分出来的特性, 视觉显著性计算模型是对这种吸引注意力的能力进行建模和量化评估。Dick 等[21] 通过实验表明在背景是固定的情况下, 运动的目标是具有显著性特征的, 是在视觉注意力分配前的视觉处理阶段就能将背景和运动目标区分开。如果用来初始化背景模型的第 1 帧图片

本身存在遮挡住背景的前景目标,即存在鬼影,就可以结合显著性目标检测判断鬼影的存在。考虑到运动目标检测的实时性需求,采用基于频域信息的显著性目标检测方法<sup>[22]</sup>。改进的总体思想是训练出的背景模型里有显著性区域,同时发现这块区域一直被检测为前景,满足这 2 个条件判断为鬼影区域,如果确定像素点是鬼影区域,则此像素背景样本模型中的样本值用这个像素点的像素值进行更新,同时设置自适应时间子采样因子。

具体算法步骤如下:1)判断鬼影像素。为每一个像素  $x$  设置一个计数器  $f_{\text{count}}(x)$ ,初始化为 0。当检测到待检测帧的像素为前景像素时,判断背景模型中此像素是否具有显著性,如果具有显著性则像素的  $f_{\text{count}}(x)$  加 1,则为

$$f_{\text{count}}(x) = \begin{cases} f_{\text{count}}(x) + 1, & f(x) = 1 \cap S(x) = 1 \\ 0, & \text{else} \end{cases}, \quad (4)$$

式中  $f(x)$  标志像素  $x$  是否为前景像素,其值为 1 表示此像素为前景,为 0 表示此像素为背景。 $S(x)$  表示检测帧前 1 帧保存的背景模型对应像素  $x$  的视觉显著性判断结果,当像素  $x$  具有视觉显著性时其值为 1,否则其值为 0。当  $f_{\text{count}}(x)$  大于预设阈值  $R_{\text{count}}$  时,则认为像素  $x$  为鬼影像素,将其背景模型进行更新,意味着对于在较长时间内前景像素值没有发生变化的像素,该方法认为它是鬼影像素,把它更新到背景模型中去。2)自适应更新时间子采样因子。在背景更新过程中,一般遵循这样的准则:如果满足条件 A,那么执行操作 B,但不能完全满足条件 A 时,是否执行操作 B 就要视条件 A 的满足程度而定,其越趋向于满足条件 A,那么越趋向于执行操作 B。而 ViBe 算法在检测前景目标时,背景是否应该更新取决于背景模型中像素点鬼影程度的高低,所以在背景更新的过程中引入模糊集合论原理。模糊集合为处理不严密信息提供了一种形式,令  $Z$  为元素(对象)集, $z$  表示  $Z$  的一类元素,即  $Z = \{z\}$ ,该集合称为论域。 $Z$  中的模糊集合  $A$  由隶属度函数  $\mu_A(z)$  表征,它是在区间  $[0, 1]$  内与  $Z$  的元素相关的一组实数。因此模糊集合是一个由  $z$  值和(赋予  $z$  成员等级的)相应隶属度函数组成的序对,即

$$A = \{z, \mu_A(z) \mid z \in Z\}. \quad (5)$$

借鉴模糊理论,背景模型中像素点的鬼影程度等价于  $\mu_A(z)$ ,  $z$  等价于待检测像素。当背景模型中像素点鬼影程度较高时,背景更新的概率应该增大,即时间子采样因子应该减小,因此不仅可以加快鬼影消除的速度,而且可以一定程度消除动态背景的影响。具体实施的方法是:在检测鬼影的同时,设置一个鬼影因子  $f_{\text{num}}(x)$ ,当像素点  $x$  由改进算法第一步判断为鬼影像素时,对应像素点的  $f_{\text{num}}(x)$  进行加  $N_0$  操作;相反,当像素点  $x$  判断为不是鬼影像素时,对应像素点的  $f_{\text{num}}(x)$  进行减  $N_1$  操作,则为

$$f_{\text{num}}(x) = \begin{cases} f_{\text{num}}(x) + N_0, & g(x) = 1 \\ f_{\text{num}}(x) - N_1, & g(x) = 0 \end{cases}, \quad (6)$$

式中  $g(x)$  的值代表像素  $x$  是否为鬼影像素,值为 1 代表鬼影像素,为 0 代表不是鬼影像素。 $N_0$  和  $N_1$  为固定常数,一般  $N_0$  取的较大而  $N_1$  取的较小,这与待检测目标的运动速度和大小有关。将  $f_{\text{num}}(x)$  的范围限定在  $[0, \phi]$ ,其中  $\phi$  为时间子采样因子。因此二次时间子采样因子为

$$\varphi = \phi * F(x), \quad (7)$$

式中  $F(x)$  为模糊函数,定义为

$$F(x) = 1/f_{\text{num}}(x), f_{\text{num}}(x) \in [0, \phi]. \quad (8)$$

### 3.2 背景高频扰动消除

背景高频扰动实际上就是由动态环境多背景造成的,它通常会在检测的前景结果中引入闪烁噪声,增加目标检测的虚警率。在 ViBe 算法中,采用全局的匹配阈值  $R$ ,但这通常无法适应动态背景的目标检测。一般在实际应用场景中,背景总会存在高频扰动的情况,对于高动态区域(例如树叶的抖动,湖面的水纹等)应该增大阈值  $R$ ,以使背景不被检测为前景;而对低动态区域,应该减小阈值  $R$ ,以便检测到更多的目标像素。因此,需要定义一个衡量背景动态变化的指标。文献[23]提出删掉通过检测一个像素点是否经常在前景与背景之间切换来判断该像素是否为闪烁像素,如果是闪烁像素则从更新掩码中消除,但是这种方法对于摄像机高频抖动产生的目标误检是无效的。借鉴文献[20]的思想,同样采取类似于  $\Sigma - \Delta$ <sup>[24]</sup>的方法,定义一个闪烁程度矩阵,通过这个闪烁程度矩阵来判断背景的动态变化程度,设置自适应阈值。具体做法是:建立一个

存储有每个像素点闪烁值的闪烁程度矩阵,其表示范围为 $[0, 150]$ 。存储之前,如果当前帧得到的前景二值图像像素点和上一帧得到的前景二值图像对应像素点结果不同,表明该像素前景背景发生切换,则该像素的闪烁值增加 15,否则像素闪烁值减少 1。如果像素点的闪烁值大于 30,则判断为闪烁像素。通过统计闪烁程度矩阵的闪烁像素数目来修改匹配阈值  $R'$ ,可表示为

$$R' = \begin{cases} R(1 + \lambda_{inc}), m \geq R_{blinking} \\ R(1 - \lambda_{dec}), m < R_{blinking} \end{cases}, \quad (9)$$

式中  $m$  为闪烁像素数目,  $R_{blinking}$  为匹配阈值  $R'$  自适应更新的边界阈值,  $\lambda_{inc}$  和  $\lambda_{dec}$  为固定的参数。匹配阈值  $R'$  不能无限增大,因为当  $R'$  太大时,会检测到形状不完整的目标,出现漏检的情况;同时  $R'$  也不能无限减小,因为当  $R'$  太小时,会检测到过多的虚假目标,增加检测虚警率。因此经过反复实验  $R'$  的取值范围应该限定为 $[20, 40]$ 。当出现背景动态扰动时,闪烁像素会不断增多,当超过设定的阈值时,匹配阈值  $R'$  会逐渐增大,可以适应动态背景的影响;而当背景相对稳定时,  $R'$  会逐渐减小,可以检测到更多的细微变化的前景。由于阈值选择不恰当会导致目标空洞的出现,同时为了在自适应阈值的基础上进一步消除动态背景对目标检测带来的影响,所以在检测目标时加入丢弃小目标和填充目标空洞的策略,具体做法是求取前景图像中的联通区域集合,如果集合中的元素(即联通区域)只有 1 个像素则将其更新为背景,接着判断集合中的元素是否有漏洞,如果有则对其进行填充。

### 3.3 摄像机抖动和背景更新错误导致噪点的消除

原 ViBe 算法采用了随机选取邻域像素的信息来构建背景模型,这在很大程度上可以保证背景模型中的信息量足够使用,但另一方面如果某一个背景像素点与其邻域像素点像素值相差很大,如摄像机的抖动容易造成背景同一区域邻域的快速变化或者出现斑马线的边缘地带,即背景某一像素点为白色,其邻域像素点一边为白色,一边为灰色。由于背景模型在构建和更新时,模型在邻域随机选取,故很容易掺杂到错误的样本点,导致前景检测结果出现噪点。对于这种问题,通常采用形态学处理方法,例如均值滤波或者图像开闭操作去除噪点,但是此方法存在一个很大缺点,会导致原来检测到的前景目标形状不完整,产生负面影响。

通过实验观察可以看出,原 ViBe 算法的检测结果表明,噪点单独出现,周围无其他噪点,因此提出一种通过遍历前景目标邻域的策略来判断该像素点是否为噪点的方法。具体做法是采用遍历检测前景像素的 24 邻域区域的像素点个数的方法,如果 24 邻域中的前景像素点个数超过给定阈值  $N_{th}$ ,则判定该前景像素点为真实前景目标像素点,反之,则该像素点为噪点。将此像素点更新为背景像素,并将此像素点的背景模型进行更新,因为选取的是 24 邻域,所以阈值  $N_{th}$  不宜取值过大,根据反复实验可得  $N_{th}$  的经验取值为 3。需注意的是,采取统计前景像素点 8 邻域区域的像素点个数的方法容易破坏原来检测目标的形状。

## 4 实验结果与分析

测试的硬件平台的 CPU 为 Intel Core i7,内存为 4 G,操作系统为 Win7 旗舰版。实验数据来源于 www.changedetection.net 网站提供的 5 个视频序列:Badminton、Traffic、Overpass、Canoe 和 Boats。所有实验中,样本个数  $N = 20$ ;最小匹配个数  $R_{min} = 2$ ;初始距离阈值  $R = 20$ ;阈值  $R$  自适应更新的边界阈值  $R_{blinking}$  为图像大小的 4%;时间子采样因子  $\phi = 12$ ;24 邻域前景像素阈值  $N_{th} = 3$ ;自增适应参数  $\lambda_{inc} = 0.05$ ;自减适应参数  $\lambda_{dec} = 0.5$ 。这里,在动态背景区域出现真实运动前景时,为了使距离阈值  $R$  能够快速减小以减少漏警的概率,设置自减适应参数  $\lambda_{dec}$  为自增适应参数  $\lambda_{inc}$  的 10 倍。实验过程中,所有结果均没有经过形态学后处理。

### 4.1 鬼影消除验证分析

实验的输入是一组连续的 Pedestrian 序列,该图像序列是行人在路上行走的监控图像,在第 607 帧时,静止的行人开始行走,在原来静止的地方产生了鬼影。图 2 为分别使用原算法与改进后算法进行鬼影消除的效果。其中,图 2(a)~(d)分别对应第 607,667,717,767 帧图像,为输入图像;图 2(e)~(h)是原算法空间邻域扩散机制的鬼影消除效果;图 2(i)~(l)是结合视觉显著性和自适应时间子采样因子的鬼影消除效果。从图中可以看出,改进后的 Vibe 算法消除鬼影的速度大大加快。原算法和改进算法的检测结果中的鬼影像素的数量随帧数的变化如图 3 所示,其中纵坐标表示鬼影像素的数量,可以看出,改进后的 ViBe 算法的检测

结果中的鬼影像素的数量减少的速度大大加快,改进算法在 40 帧之后鬼影已经完全消除,从而具有更为准确的检测结果。

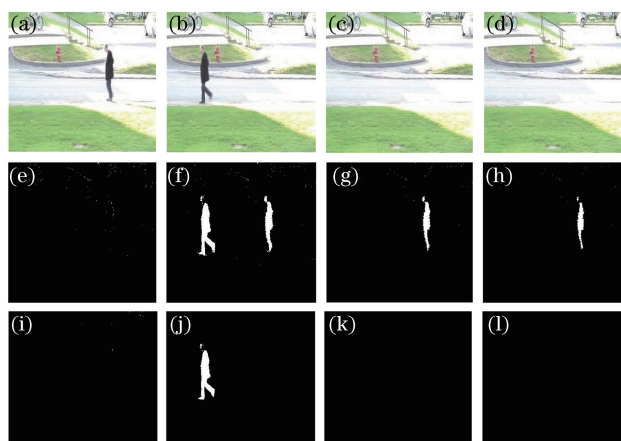


图 2 (a)~(d) 第 607, 667, 717, 767 帧原始图像; (e)~(h) 应用 ViBe 算法的第 607, 667, 717, 767 帧图像检测结果; (i)~(l) 改进型 ViBe 算法的第 607, 667, 717, 767 帧图像检测结果

Fig. 2 (a)~(d) Original images of 607th, 667th, 717th, 767th frames; (e)~(h) outcomes of 607th, 667th, 717th, 767th frames by the original ViBe algorithm; (i)~(l) outcomes of 607th, 667th, 717th, 767th frames by the improved ViBe algorithm

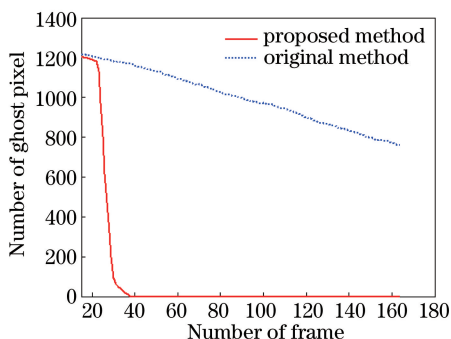


图 3 鬼影像素的数量随帧数的变化

Fig. 3 Number of ghost pixel versus number of frame

#### 4.2 背景高频扰动消除验证分析

图 4 为原始 ViBe 算法和改进后应用自适应阈值的 ViBe 算法的对比效果图。其中图 4(a)为输入的 Overpass 序列的其中一帧图像,图 4(b)为原始 ViBe 算法的检测效果图,图 4(c)为加入自适应阈值  $R$  后的 ViBe 算法的检测效果图。由图 4(b)可知,背景中树叶的抖动明显,树叶被误检测为运动目标,极大地影响

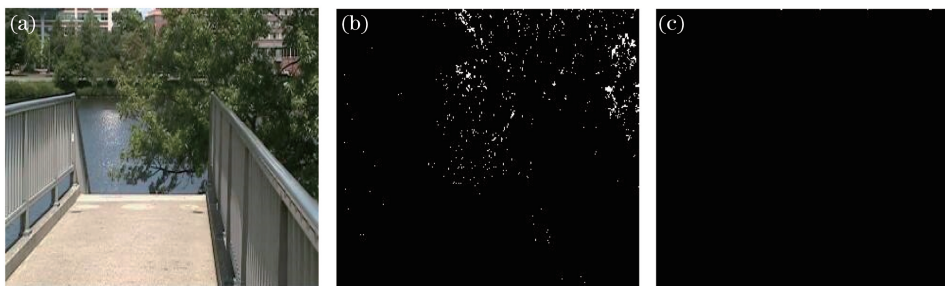


图 4 (a) Overpass 序列的其中一帧原始图像; (b) 使用 ViBe 算法后的 Overpass 序列的检测二值图; (c) 使用加入自适应  $R$  后的 Overpass 序列的检测二值图

Fig. 4 (a) An original image in Overpass sequence; (b) detection binary outcome by the original ViBe algorithm; (c) detection binary outcome by the improved ViBe algorithm

了前景目标检测的准确性。通过和图 4(c)比较可以看出加入自适应阈值  $R$  后的算法可以很好地适应动态背景变化(树叶扰动),大大地降低了目标检测的虚警率。

#### 4.3 摄像机抖动和背景更新错误导致噪点消除验证分析

图 5 和图 6 分别为第 14 帧和第 44 帧 Badminton 视频序列的检测结果,其中(a)~(d)分别是当前帧、真实值、ViBe 算法和本文算法的检测结果。由图 5 可以看出,在检测目标的过程中摄像机抖动带来背景强烈的晃动,原始 ViBe 算法将其当作虚假目标(图中的羽毛球球和羽毛球网),造成大量的虚警,无法消除摄像机抖动带来的负面影响,而用于邻域检测的改进型 ViBe 算法很好地消除了摄像机抖动带来的负面影响,同时最大程度地保留了正确的目标形状。而且由图 5(c)和 5(d)可以看出在第 14 帧时目标检测都出现了鬼影,这是由于用来初始化背景模型的第 1 帧图像存在运动目标造成的。但是图 6(c)和(d)中在第 44 帧时原始 ViBe 算法并没有消除鬼影,而改进型 ViBe 算法检测到的结果中鬼影已经消除了,进一步说明了改进型 ViBe 算法可以加快鬼影消除的速度。

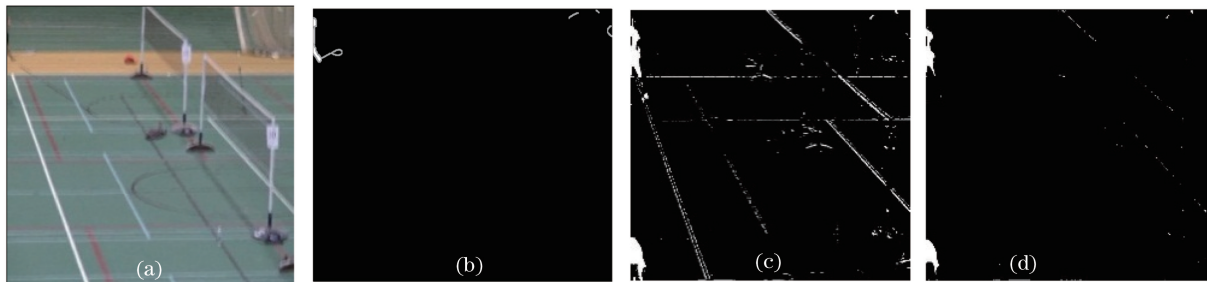


图 5 第 14 帧 Badminton 视频序列的测试结果比较。(a) 当前帧;(b) 真实值;(c) ViBe 算法;(d) 本文算法

Fig. 5 Comparison of test results of ViBe and proposed method in Badminton sequence of 14th frame.

(a) Input frames; (b) ground truth; (c) ViBe method; (d) proposed method

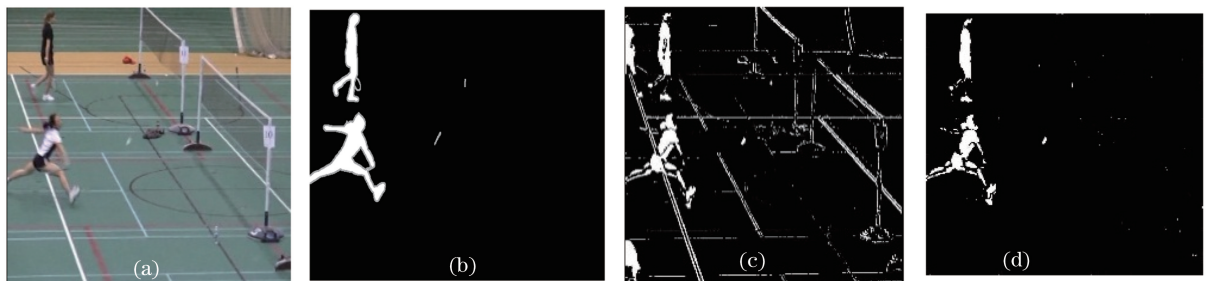


图 6 第 44 帧 Badminton 视频序列的测试结果比较。(a) 当前帧;(b) 真实值;(c) ViBe 算法;(d) 本文算法

Fig. 6 Comparison of test results of ViBe and proposed method in Badminton sequence of 44th frame.

(a) Input frame; (b) ground truth; (c) ViBe method; (d) proposed method

#### 4.4 算法比较验证分析

图 7 给出了 GMM 算法、ViBe 算法、三帧差分法和本文算法在其余 4 组场景(Overpass、Canoe、Traffic、Fall)下的检测效果。从测试结果可以看出,在 Overpass 场景下,GMM 算法和 ViBe 算法都将晃动的树枝和树叶误检测为运动目标,有较多的错误前景点,三帧差分法由于有逻辑与的步骤,所以很大程度地消除树叶树枝晃动的影响。但是如果不进行形态学处理,只能检测到运动目标的轮廓。而本文算法则可以很好地消除树枝树叶晃动带来的负面影响,只将运动的人检测为运动目标,且形状提取也比其他算法好;在 Canoe 场景和 Fall 场景下,GMM 算法和 ViBe 算法都将湖面水纹的运动和剧烈的树叶晃动误检测为目标前景,出现了大量的虚假目标,严重影响了目标检测的精度。由于 Canoe 场景和 Fall 场景的背景扰动更加剧烈,所以三帧差分法也不能消除湖面水纹运动带来的负面影响,且同样只能提取出运动目标的轮廓。而本文算法则很好地抑制了湖面水纹和树叶剧烈晃动的干扰;在 Traffic 场景中,受摄像机抖动的影响,GMM 算法和 ViBe 算法都将马路边缘和斑马线误检测为运动目标,而且 ViBe 算法还将树叶的影子误检测为运动目标。而本文算法则可以适应多模态的背景,所以不受影响。三帧差分法由于有逻辑与的步骤,可以很大程度地消

除摄像机抖动带来的负面影响,但是无法得到完整的目标前景,甚至只能得到目标轮廓的一部分。而本文算法很好地抑制了摄像机抖动带来的影响,同时完整地提取出了前景的形状。

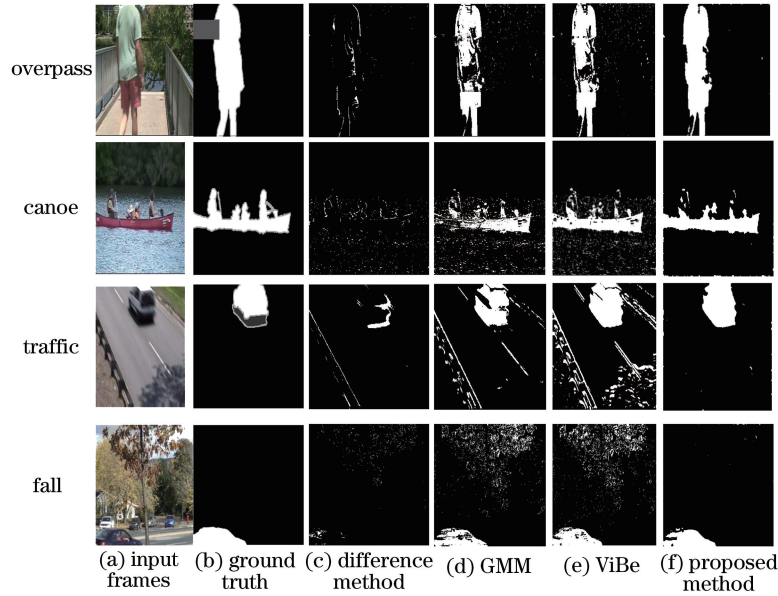


图 7 Change detection 数据集测试结果

Fig. 7 Test results of the change detection dataset

通过以上实验测试可以得出如下结论:1)与原始 ViBe 算法相比,改进后的 ViBe 算法可以极大地加快消除鬼影的速度;2)GMM 算法和 ViBe 算法不适用于动态背景场景和摄像机抖动场景的目标检测,而改进后的 ViBe 的算法克服了上述两种算法的不足,在上述 5 组测试序列中展现了良好的性能,扩大了算法的应用范围;3)三帧差分法虽然速度很快,但是只适合作为辅助性的目标检测算法。

为了定量比较几种算法的性能,采取准确率和召回率作为量化指标,定义为

$$f_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}, \quad (10)$$

$$f_{\text{recall}} = \frac{N_{\text{TN}}}{N_{\text{TN}} + N_{\text{FN}}}, \quad (11)$$

式中  $N_{\text{TP}}$  表示正确检测的前景数;  $N_{\text{FP}}$  表示错误检测的前景数;  $N_{\text{TN}}$  表示正确检测的背景数;  $N_{\text{FN}}$  表示错误检测的背景数。  $f_{\text{recall}}$  反映了正确检测的像素占真值图(ground truth)中正确前景数的百分比;  $f_{\text{precision}}$  反映了正确检测的像素占检测出的前景总像素的百分比。

图 8 给出了 5 组场景(Badminton、Traffic、Overpass、Canoe、Fall)下 4 种算法的  $f_{\text{precision}}$  与  $f_{\text{recall}}$  的直方图,总共计算了 6409 帧序列图片,采取了分别求取每帧图片中对应算法的 2 个指标,然后取平均值的方法。由图 8 可以看出,由于本文算法大大降低了鬼影、动态背景和摄像机抖动引起的错误前景检测数,因此在准

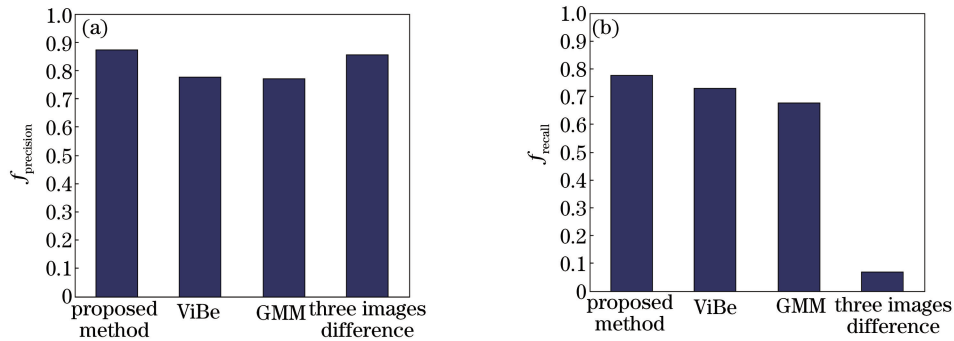


图 8 5 组场景下 4 种算法的  $f_{\text{precision}}$  与  $f_{\text{recall}}$  的直方图

Fig. 8 Histogram of  $f_{\text{precision}}$  and  $f_{\text{recall}}$  based on four algorithms in five sequences



确率指标上明显高于 ViBe 算法和 GMM 算法,比原始 ViBe 算法提高了 10 个百分点,而三帧差分法在这个指标上高于 ViBe 算法和 GMM 算法是因为它只检测出目标轮廓,正确检测的前景数小,而它又能抑制动态背景等因素的影响,所以错误前景检测数也很小,导致准确率偏高,不具有参考价值。在召回率指标上,本文算法相比 ViBe 算法和 GMM 算法也有了一定程度的提高,而三帧差分法在这个指标上很低是因为它只检测出目标轮廓,错误检测的前景数很大,正确检测的前景数小,所以也可以看出三帧差分法并不适用于实际场景的目标检测。

## 5 结 论

针对 ViBe 算法无法消除鬼影、无法适应动态背景和摄像机抖动、背景容易更新错误的不足,提出了一种改进型 ViBe 算法。将原算法、GMM 算法、三帧差分法和改进算法进行对比实验,并使用评价指标对实验结果进行评价。为了加快鬼影消除的速度,结合视觉显著性理论判断背景模型中存在的鬼影目标,通过判断背景模型中每个像素点的鬼影程度结合模糊准则,自适应改变时间子采样因子;为了适应动态背景,消除背景的高频扰动,设置一个闪烁程度矩阵来判断背景的高频扰动程度,进而改变匹配阈值;为了使检测结果更加准确,加入了小目标丢弃与空洞修复方案进行优化;为了消除摄像机抖动引起的噪点和背景邻域更新错误带来的负面影响,通过统计前景像素 24 邻域区域的像素点个数来判断前景像素点是否为噪点,在噪点消除的同时最大限度地保留了检测到的正确前景形状,并进一步消除动态背景带来的影响。结果表明,改进型算法误检率明显下降,准确率和召回率两个指标均有所提高,稳健性大大提高,适用于更多的应用场景。

## 参 考 文 献

- 1 Benezeth Y, Jodoin P, Emile B, *et al.*. Review and evaluation of commonly-implemented background subtraction algorithms[C]. 19th International Conference on Pattern Recognition, 2008: 1-4.
- 2 Ding Qi, Gu Guohua, Xu Fuyuan, *et al.*. Moving target detection on moving camera with the presence of strong parallax [J]. Laser & Optoelectronics Progress, 2015, 52(9): 091501.  
丁 祺, 顾国华, 徐富元, 等. 强视差下的移动相机运动目标检测[J]. 激光与光电子学进展, 2015, 52(9): 091501.
- 3 Bernd Jahne. Digital image processing[M]. Berlin: Springer-Verlag, 2000: 7-13.
- 4 Piccardi M. Background subtraction techniques: A review[C]. 2004 IEEE International Conference on Systems, Man and Cybernetics, 2004, 4: 3099-3104.
- 5 Zhuang Zhemin, Zhang Congyou, Yang Jinyao, *et al.*. Investigation on visual background extractor based on gray feature and adaptive threshold[J]. Journal of Electronics & Information Technology, 2015, 37(2): 347-352.  
庄哲民, 章聪友, 杨金耀, 等. 基于灰度特征和自适应阈值的虚拟背景提取研究[J]. 电子与信息学报, 2015, 37(2): 347-352.
- 6 Chen Yin, Ren Kan, Gu Guohua, *et al.*. Moving object detection based on improved single Gaussian background model [J]. Chinese J Lasers, 2014, 41(11): 1109002.  
陈 银, 任 侃, 顾国华, 等. 基于改进的单高斯背景模型运动目标检测算法[J]. 中国激光, 2014, 41(11): 1109002.
- 7 Elgammal A, Harwood D, Davis L. Non-parametric model for background subtraction [M]. Berlin: Springer-Verlag, 2000: 751-767.
- 8 Elgammal A, Duraiswami R, Harwood D, *et al.*. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance[J]. Proceedings of the IEEE, 2002, 90(7): 1151-1163.
- 9 Tavakkoli A, Nicolescu M, Bebis G, *et al.*. Non-parametric statistical background modeling for efficient foreground region detection[J]. Machine Vision and Applications, 2008, 20(6): 395-409.
- 10 Monari E, Pasqual C. Fusion of background estimation approaches for motion detection in non-static backgrounds [C]. 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, 2007: 347-352.
- 11 Kim K, Chalidabhongse T, Harwood D, *et al.*. Background modeling and subtraction by codebook construction[C]. 2004 IEEE International Conference, 2004, 5: 3061-3064.
- 12 Kim K, Chalidabhongse T, Harwood D, *et al.*. Real-time foreground-background segmentation using codebook model [J]. Special Issue on Video Object Processing, 2005, 11(3): 172-185.
- 13 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.

- 14 Barnich O, van Droogenbroeck M. ViBe: A powerful random technique to estimate the background in video sequences [C]. 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, 2009: 945-948.
- 15 Sanin A, Sanderson C, Lovell B. Shadow detection: A survey and comparative evaluation of recent methods[J]. Pattern Recognition, 2012, 45(4): 1684-1695.
- 16 Prati A, Mikic I, Trive di M M, *et al.*. Detecting moving shadows: Algorithms and evaluation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(7): 918-923.
- 17 Jodoin P M, Mignotte M, Konrad J. Statistical background subtraction using spatial cues [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(12): 1758-1763.
- 18 Shoushtarian B, Bez B. A practical adaptive approach for dynamic background subtraction using an invariant colour model and object tracking[J]. Pattern Recognition Letters, 2005, 26(1): 5-26.
- 19 Borji A, Itti L. State-of-the-art in visual attention modeling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 185-207.
- 20 Li Qingzhong, Zang Fengni, Zhang Yang. Ship target detection for moving video maritime surveillance[J]. Chinese J Lasers, 2014, 41(8): 0814001.  
李庆忠, 臧凤妮, 张 洋. 动态视频监控中海上舰船目标检测[J]. 中国激光, 2014, 41(8): 0814001.
- 21 Dick M, Ullman S, Sagi D. Parallel and serial process in motion detection[J]. Science, 1987, 237(4813): 400-402.
- 22 Hou X D, Zhang L Q. Saliency detection: A spectral residual approach [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2007: 1-8.
- 23 van Droogenbroeck M, Paquot O. Background subtraction: Experiments and improvements for ViBe [C]. 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2012: 32-37.
- 24 Manzanera A, Richefeu J C. A new motion detection algorithm based on  $\Sigma$ - $\Delta$  background estimation [J]. Pattern Recognition Letters, 2007, 28(3): 320-328.