

中图法分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2011)02-0236-08

论文索引信息: 张水发, 张文生, 丁欢, 杨柳. 融合光流速度与背景建模的目标检测方法 [J]. 中国图象图形学报, 2011, 16(2): 236-243

融合光流速度与背景建模的目标检测方法

张水发, 张文生, 丁欢, 杨柳

(中国科学院自动化研究所, 北京 100190)

摘要: 为了克服传统基于像素的背景建模方法不能很好地描述背景运动的问题, 提出了一种融合光流速度与背景建模的目标检测方法。结合像素的灰度信息、空间信息和时间信息计算出每个像素的光流速度, 利用光流速度在时间域上的统计信息为背景建立光流速度场模型。利用建立的背景模型快速、准确地实现运动目标的检测。实验结果表明, 融合光流速度的背景建模方法能有效地描述背景的运动, 显著降低运动背景产生的噪音, 鲁棒地实现运动目标检测。

关键词: 背景建模; 光流法; 目标检测; 混合高斯

Background modeling and object detecting based on optical flow velocity field

Zhang Shuifa, Zhang Wensheng, Ding Huan, Yang Liu

(Institute of Automation, Chinese Academy of Sciences, Beijing 100190 China)

Abstract: The traditional pixel-based background model cannot represent the background motion efficiently. In this paper, a novel strategy is proposed to model background and track moving objects based on optical flow velocity field. Statistics on intensity, spatial and temporal information of pixels are extracted to generate the optical flow field, which is used to formulate a novel background model for tracking moving objects efficiently and exactly. This optical-flow-field-based strategy can reduce noise generated by background motion significantly and track moving objects robustly, as illustrated in our experiments.

Keywords: background modeling; optical flow; object detecting; mixture of Gauss

0 引言

运动目标检测是从图像序列中提取运动目标, 有效分割运动目标对目标的分类、跟踪及行为理解等后期处理具有非常重要的意义。由于实际场景中, 背景受光照、天气或其他因素影响, 并不能保证完全静止, 导致背景物体以及前景物体的运动同时

存在于待检测图片中, 使运动目标检测成为一项相当困难的工作, 成为目前计算机视觉和图像理解研究的难点和热点问题之一。

背景减除法^[1-6]是目前最常用的运动目标检测方法, 通过对背景的学习得到背景模型, 通过当前帧与背景模型比较, 将不符合背景模型分布的部分判定为运动目标。理想情况下, 背景完全静止, 可表示为一帧背景图像; 但大多数情况下, 由于边缘像素抖

收稿日期: 2009-09-08; 修回日期: 2009-12-11

基金项目: 国家基础研究发展计划项目(973)(2004CB 318103); 国家高技术研究发展计划项目(863)(2007AA012138578); 国家自然科学基金项目(60033020); 中国科学院海外杰出人才研究计划项目(06S3011S01); 软件工程国家重点实验室开放基金支持项目(SKLSSE2008-07-19)。

第一作者简介: 张水发(1985—), 男。现为中国科学院自动化研究所复杂系统与智能科学重点实验室模式识别与智能系统专业硕士研究生, 主要研究方向为图像处理、行为识别。E-mail: zhangs85@gmail.com。

动、光照变化、设备噪音等因素,背景中的物体自身存在或多或少的运动。针对这种情况,国内外学者采用各种模型表示背景,其中典型的有:高斯模型^[2]和混合高斯模型^[3]用一个或多个高斯分布表示背景,可用于复杂、运动的背景建模,但当背景快速运动时,一个或者多个高斯模型并不能准确表示背景;非参数方法^[4]利用像素点最近几帧的灰度值信息,使用窗函数密度估计得到该像素点的灰度值分布,但该方法计算每帧图像需遍历全部训练集,计算速度较慢;本征方法^[5]将视频帧投影到本征背景空间,通过计算视频帧与背景帧在本征空间的差值实现前景检测,该方法利用主成分分析方法实现降维,检测速度快,但背景模型是通过初始化几十帧图片获得,因此资源消耗大,自适应能力差。上述几种建模方法均属于基于像素的背景建模方法,能够获得较细致的运动目标形状,但由于对像素点的时间信息和空间信息利用不充分,当背景物体不规则运动时,不能准确地描述背景运动。

本文试图跳出传统的基于像素背景建模方法,提出融合光流速度的背景建模算法。利用灰度值的统计信息描述背景,综合利用像素的灰度信息、空间信息、时间信息,以及光流速度在时间域上的统计信息,从像素运动速度入手,对背景进行像素级建模。用基于梯度的方法计算光流速度,然后利用其在时间域上的统计信息建立背景的光流速度场^[7]模型,最后用于运动目标检测;同时在检测过程中实时更新背景模型,适应背景变化。本文算法利用光流速度准确地描述了背景运动,克服了树叶晃动等背景不规则运动对建模造成的不利影响,同时具有较好的自适应性,当背景中光照变化、物体被移走或遗留时,算法均能自适应背景变化。

1 光流法

光流(optical flow)是空间运动物体被观测表面上像素点运动的瞬时速度,光流按照空间位置排列组成光流场。Barron等人^[8]按理论基础与数学方法不同,将光流计算方法分成5种:基于梯度的方法^[9-11]、基于匹配的方法^[12]、基于能量的方法^[13]、基于相位的方法^[14]和神经动力学方法^[15]。

基于梯度的方法又称为微分法^[9-11],利用时变图像像素灰度的时空微分计算其速度矢量,由于基于梯度的方法计算简单且结果较好,该方法得到了

广泛的应用和研究;基于匹配的光流计算方法^[12]包括基于特征和区域两种,对快速运动和亮度变化具有鲁棒性,但其计算的光流通常很稀疏,而且特征提取和精确匹配也十分困难;基于能量的方法^[13]对输入图像序列进行时空滤波处理,但是会降低光流估计的空间和时间分辨率,并且存在高计算负荷的问题;基于相位的方法^[14]是根据带通滤波器输出的相位特性确定速度,速度估计结果比较精确,但是计算复杂度较高,边缘处运动估计的时间分辨率较差,并对输入图像序列的时间混叠比较敏感;近几年有人提出用神经网络建立视觉运动感知的神经动力学模型来测量光流^[15],但是这一技术尚不成熟。综上所述,光流场的计算技术经过多年的发展,已经形成一些比较经典的算法,Barron等人^[11]研究表明,基于梯度的方法计算复杂度较低,并且能得到比较精确的光流,因此本文选用基于梯度的方法^[9-11]计算光流速度。

在文献[11]中Horn-Schunck提出基于梯度的方法,假设 $f(x, y, t)$ 是关于变量 x, y, t 的连续函数。设 t 时刻,点 (x, y) 处的灰度值为 $f(x, y, t)$;在 $(t + \Delta t)$ 时刻,运动到点 $(x + \Delta x, y + \Delta y)$,灰度值为 $f(x + \Delta x, y + \Delta y, t + \Delta t)$,由于 Δt 是一个极小的时间段,假设

$$f(x, y, t) = f(x + \Delta x, y + \Delta y, t + \Delta t) \quad (1)$$

将式(1)右边用泰勒公式展开,略去二次无穷小项,得

$$\frac{\partial f}{\partial x}u + \frac{\partial f}{\partial y}v + \frac{\partial f}{\partial t} = 0 \quad (2)$$

式(2)是基本的光流约束方程,它反映了运动图像时间梯度与空间梯度间的时空微分关系。

在整个视频图像范围内,运动速度矢量场平滑变化,即相邻的图像像素点之间具有相似的速度矢量。速度矢量逐像素的变化可以通过速度矢量分量的空间梯度的模平方和(用 E_c^2 表示)来定量给出,可以证明图像的速度场越平滑,梯度值就越小。因此附加的光流平滑约束条件就是要求 E_c^2 最小。

定义像素点光流误差为

$$E_b = \frac{\partial f(x, y, t)}{\partial x}u(x, y, t) + \frac{\partial f(x, y, t)}{\partial y}v(x, y, t) + \frac{\partial f(x, y, t)}{\partial t} \quad (3)$$

通过求像素点光流误差与速度场逐像素变化数值加权后的最小值来估计每一个像素点处的速度矢量。

$$E^2 = \min_{V(x,y,t)} \iint_A [E_b^2(V) + a^2 E_c^2(V)] dx dy \quad (4)$$

式中 A 代表连续图像的区域范围。 α^2 用来控制平滑约束条件的强度。通过最小化 E^2 计算光流速度矢量,得到 $u(x,y,t)$ 和 $v(x,y,t)$ 的迭代方程

$$\begin{aligned} u^{n+1} &= u^n - \frac{f_x [f_x u^n + f_y v^n + f_t]}{[a^2 + f_x^2 + f_y^2]} \\ v^{n+1} &= v^n - \frac{f_y [f_x u^n + f_y v^n + f_t]}{[a^2 + f_x^2 + f_y^2]} \end{aligned} \quad (5)$$

式中 f_x 表示 $\frac{\partial f(x,y,t)}{\partial x}$, f_y 表示 $\frac{\partial f(x,y,t)}{\partial y}$, f_t 表示 $\frac{\partial f(x,y,t)}{\partial t}$, n 是迭代次数。

通过 Gauss-Seidel 迭代得到了每个像素 x 方向和 y 方向的光流速度,通过判断该速度大小可以把每个像素划分为前景点或背景点,实现运动目标检测。如图 1 所示,图 1(a)是从 Vssn06 video4 中截取的一帧,图 1(b)是用 Horn-Schunck 基于梯度的方法运动目标检测得到的二值图像,白色部分表示运动目标。从检测结果看出,在检测目标物体的同时,把运动背景也检测为运动目标,产生了很大噪音。因此,为了消除运动背景对运动目标检测带来的影响,我们提出了融合光流速度的背景建模方法。



(a) 原始图像

(b) 二值图像

图 1 光流法运动目标检测

Fig. 1 Object detecting based on optical flow

2 基于光流速度场背景建模

光流法可充分利用像素的时空信息计算图像中每一个像素点的速度矢量,形成图像运动场,且能在不需要知道场景任何先验知识的条件下能得到完整的运动信息,适合检测相对复杂的运动,可用于动态背景。因此利用光流法计算出的光流速度场构造背景模型。

利用基于梯度的光流计算方法,提出了一种融合光流速度场的背景建模方法检测运动目标:首先利用融合梯度的光流法综合每个像素的时间信息、

空间信息和灰度信息计算每个像素的光流速度;然后统计每个像素的光流速度在时间域上的分布信息,构造背景模型。

传统的基于像素背景建模方法没有充分利用像素的时间信息和空间信息,当背景物体不规则运动时,很难描述出背景运动;本文算法的优点在于综合利用了像素的灰度信息、时间信息、空间信息以及光流速度在时间域上的统计信息,通过光流速度更好的描述了背景运动。

传统的光流法通过判断像素运动幅度大小实现运动目标检测,当背景运动幅度大时,经常把背景运动误检为前景目标;本文算法在运动目标检测之前利用光流速度在时间域上的统计信息,建立光流速度场模型,描述背景运动,在检测过程中,利用建立的背景模型,能有效消除背景运动的影响。

2.1 建立光流速度场背景模型

令 $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 表示一个像素上用来训练的 n 个样本, $\mathbf{x}_t (t \in [1, n])$ 包括 t 时刻像素上的灰度信息、像素邻域内的灰度信息以及 $t-1$ 时刻,对应像素上的灰度信息与其邻域内的像素灰度信息;用 $\mathbf{c} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ 表示每个像素。 $\mathbf{u}, \mathbf{v}, \mathbf{w}$ 具体表示为

$$\mathbf{u} = \{u_{\min}, u_{\max}, u_diff_{\min}, u_diff_{\max}\} \quad (6)$$

$$\mathbf{v} = \{v_{\min}, v_{\max}, v_diff_{\min}, v_diff_{\max}\} \quad (7)$$

$$\mathbf{w} = \{w_{\min}, w_{\max}, w_diff_{\min}, w_diff_{\max}\} \quad (8)$$

$u_{\min}, u_{\max}, v_{\min}, v_{\max}$:所有训练样本在 x, y 方向速度的最小值和最大值; $u_diff_{\min}, u_diff_{\max}, v_diff_{\min}, v_diff_{\max}$:相邻两帧在 x, y 方向速度差最小值和最大值。 w_{\min}, w_{\max} :所有训练样本速度的最小值和最大值 ($w = \sqrt{u^2 + v^2}$); $w_diff_{\min}, w_diff_{\max}$:相邻两帧训练样本速度差的最小值和最大值。

背景建模时,利用 $\mathbf{x}_t (t \in [1, n])$ 中包含的灰度信息、空间信息和时间信息计算像素 x, y 方向速度 u_t, v_t 和总速度 w_t ,分别与 u_t, v_t, w_t 做差得到 $u_diff_t, v_diff_t, w_diff_t$,对于背景中的每一个像素点,分别得到 x 方向和 y 方向速度的最大值和最小值以及相邻两帧训练样本 x 方向和 y 方向速度差的最大值和最小值,并在此基础上得到总速度的最大值和最小值以及相邻两帧训练样本间总速度差的最大值和最小值,建立的背景模型为

$$\mathbf{c} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\} \quad (9)$$

式中 $\mathbf{u} = \{\min_t \{u_t\}, \max_t \{u_t\}, \min_t \{u_diff_t\}, \max_t \{u_diff_t\}\}$, $\mathbf{v} = \{\min_t \{v_t\}, \max_t \{v_t\}, \min_t \{v_diff_t\}, \max_t \{v_diff_t\}\}$,

$\max_i \{v_diff_i\}$, $w = \{\min_i \{w_i\}, \max_i \{w_i\}, \min_i \{w_diff_i\}, \max_i \{w_diff_i\}\}$, $t \in [1, n]$, 是 n 帧训练样本中的第 t 帧。 $\min_i \{x_i\}$, $\max_i \{x_i\}$ 分别表示 N 帧训练样本中的最小值和最大值。

具体实现算法如下:

1) 初始化

$u_{\min}, u_diff_{\min}, v_{\min}, v_diff_{\min}, w_{\min}, w_diff_{\min}$ 初始化为最大整数;

$u_{\max}, u_diff_{\max}, v_{\max}, v_diff_{\max}, w_{\max}, w_diff_{\max}$ 初始化为最小整数;

预读一帧图像, 为计算光流速度提供参考帧。

2) 对所有的 $t \in [1, n]$

(1) 读入一帧图像, 结合上一帧图像信息计算光流速度 u_i, v_i, w_i ($w_i = \sqrt{u_i^2 + v_i^2}$)

(2) 与上一帧相应像素的速度 $u_{i-1}, v_{i-1}, w_{i-1}$ 做差得 u_diff, v_diff, w_diff

(3) 把 $u_i, v_i, w_i, u_diff, v_diff, w_diff$ 分别与 $u_{\min}, v_{\min}, w_{\min}, u_diff_{\min}, v_diff_{\min}, w_diff_{\min}$ 比较, if ($u_i < u_{\min}$) then $u_{\min} = u_i$, 依次更新 $v_{\min}, w_{\min}, u_diff_{\min}, v_diff_{\min}, w_diff_{\min}$, 再把 $u_i, v_i, w_i, u_diff_i, v_diff_i, w_diff_i$ 分别与 $u_{\max}, v_{\max}, w_{\max}, u_diff_{\max}, v_diff_{\max}, w_diff_{\max}$ 比较, if ($u_i > u_{\max}$) then $u_{\max} = u_i$, 同理更新 $v_i, w_i, u_diff_i, v_diff_i, w_diff_i$ 。

经过背景学习, 建立了光流速度场模型, 描述背景运动。当背景物体运动状态改变时, 有可能把背景误检为运动目标。因此, 在检测过程中还需要更新背景模型。

2.2 更新光流速度场背景模型

光照变化、天气变化、背景物体运动状态改变都可能使原来的模型失效。由于本文为背景建立光流速度场模型, 在计算光流速度时结合了像素灰度信息、空间信息以及相邻帧信息, 当光照变化或天气变化时, 会产生一帧或短时间误报, 只要背景物体运动状态没变化, 就可以不用考虑背景模型的更新, 所以本文背景模型更新只针对背景物体运动状态的改变。

针对背景物体运动改变的3种可能(运动→不运动、不运动→运动、运动速度变化), 使用两种不同的策略对背景模型进行像素级更新。运动目标检测期间, 为每个像素点创建两个队列, 分别记录当前像素点最近 N 帧被判定为前景点的次数和速度信息, 以此来决定是否更新背景模型。两个队列的说明如下:

$nZ(t)$ 表示 (u, v, w) 在 0 邻域范围的次数:

$$nZ(t) = \begin{cases} nZ(t-1) + 1 & (u, v, w) \text{ 在 } 0 \text{ 的邻域范围内} \\ nZ(t-1) & (u, v, w) \text{ 不在 } 0 \text{ 的邻域范围内} \end{cases} \quad (10)$$

$FG(t)$ 表示像素点被判断为前景点的次数:

$$FG(t) = \begin{cases} FG(t-1) + 1 & \text{当前像素点被判断为前景点} \\ FG(t-1) & \text{当前像素点不被判断为前景点} \end{cases} \quad (11)$$

背景模型更新策略:

$$(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \begin{cases} (\mathbf{0}, \mathbf{0}, \mathbf{0}) & (\mathbf{u}, \mathbf{v}, \mathbf{w}) \text{ 不在 } 0 \text{ 的邻域内且} \\ & nZ(t) - nZ(t-N) > \kappa N \\ (\mathbf{u}, \mathbf{v}, \mathbf{w})_{\text{new}} & \text{最近 } N \text{ 个像素点有 } \gamma N \text{ 个被检测为前景点} \end{cases} \quad (12)$$

κ, γ 表示背景更新的阈值, 这里 κ 表示像素点在最近 N 帧中有 κ 的概率光流速度为 0 或者近似为 0, κ 理想情况下为 1 与背景运动物体宽度所占像素数与运动范围的像素数的比例之差, γ 表示像素点在最近 N 帧中有 γ 的概率光流速度为 $(u, v, w)_{\text{new}}$, γ 理想情况下为 $(1 - \kappa)$, 但是实际情况下, 并不能确定背景物体宽度与其运动范围的像素数, 所以通过先验知识对 κ, γ 取值, 本文针对背景中树叶的运动进行更新的 κ, γ 为 0.9, 0.1。

2.3 融合光流速度的背景模型的运动目标检测

运动目标检测可以分为 4 步: 二值化、去噪、数字形态学滤波和目标检测。将像素划分为背景点或者前景点, 符合背景模型分布的为背景点, 不符合背景模型分布的为前景点, 用 $B(x)$ 表示第 t 帧中像素 x 为背景像素还是前景像素。

首先利用像素的灰度信息和空间信息、时间信息计算出光流速度 u_i, v_i, w_i , 然后判断光流速度是否符合背景模型的分布, 符合则认为是背景点, 否则认为是前景点:

$$B(x) = \begin{cases} 0(\text{背景}) & u_i, v_i, w_i \text{ 符合背景分布} \\ 1(\text{前景}) & u_i, v_i, w_i \text{ 不符合背景分布} \end{cases} \quad (13)$$

$$u_diff_{\min} < u_i - u_{\min} < u_diff_{\max} \parallel u_diff_{\min} < u_i - u_{\max} < u_diff_{\max} \quad (14)$$

$$v_diff_{\min} < v_i - v_{\min} < v_diff_{\max} \parallel v_diff_{\min} < v_i - v_{\max} < v_diff_{\max} \quad (15)$$

$$w_diff_{\min} < w_i - w_{\min} < w_diff_{\max} \parallel w_diff_{\min} < w_i - w_{\max} < w_diff_{\max} \quad (16)$$

当 u_i, v_i 对式 (14) (15) 同时成立时, 则认为 u_i, v_i, w_i 符合背景分布, 属于背景点, 当 u_i, v_i 对式 (14) (15) 都不成立时, 则认为 u_i, v_i, w_i 不符合背景分布, 属于前景点, 当 u_i, v_i 其中之一对式 (14) (15) 成立时, 如果 w_i 对式 (16) 成立, 则认为 u_i, v_i, w_i 符合背景分布, 属于背景点, 否则认为 u_i, v_i, w_i 不符合背景分布, 属于前景点。

3 实验结果及比较分析

为了验证本文提出的方法, 我们主要在 Vssn06 (4th ACM International Workshop on Video Surveillance & Sensor Network) 视频库上进行测试。Vssn06 包含多个室内、室外场景, 提供的测试数据包括: 背景运动、光照渐变、光照突变、阴影、物体遗留等, 并提供了与运动目标对应的前景掩膜视频。在本文实验中主要在包含背景运动的 video4、光照突变的 video8、车辆停止在背景中的 tunnelvideo 视频中进行测试。同时将实验结果与相同视频经混合高斯^[3]、均值滤波和光流法处理后的结果进行比较。实验在 VC++6.0 平台下完成, 实验得到的结果都是进行了形态学滤波后的处理结果。

3.1 背景存在运动条件下的运动目标检测

图 2(c) — (f) 是在背景存在显著运动时的检测结果, 实验对比了均值滤波、混合高斯、光流法和本文融合光流速度场背景建模算法的检测效果。图 2(a) 是 Vssn06 video4 视频中截取的一帧, 分辨率为 384×240 , 背景存在复杂运动, 尤其是左下角小黄花运动幅度大, 运动速度快; 而且当人运动到如图 2(a) 所示位置时, 背景像素与人身体的像素灰度值接近, 也给检测带来了很大的影响。实验效果如图 2 所示, 图 2(b) 为 Vssn 提供的理想二值图像, 图 2(c) — (f) 分别是用均值滤波、混合高斯、融合光流速度场背景建模后检测得到的运动目标, 图 2(e) 是光流法检测得到的运动目标。

从实验结果可以看出均值滤波背景建模检测到的前景存在大量的噪音; 混合高斯背景建模比均值滤波更好的适应背景像素的运动, 但是对于背景快速变化的边缘, 仍然不能准确描述, 并且当前景像素与背景像素亮度及颜色特征比较接近时会出现漏检现象; 直接用光流法检测到的前景包含了图像中背景像素的运动; 因为建模时为运动背景建立了光流速度模型, 所以本文融合光流速度场背景建模的方

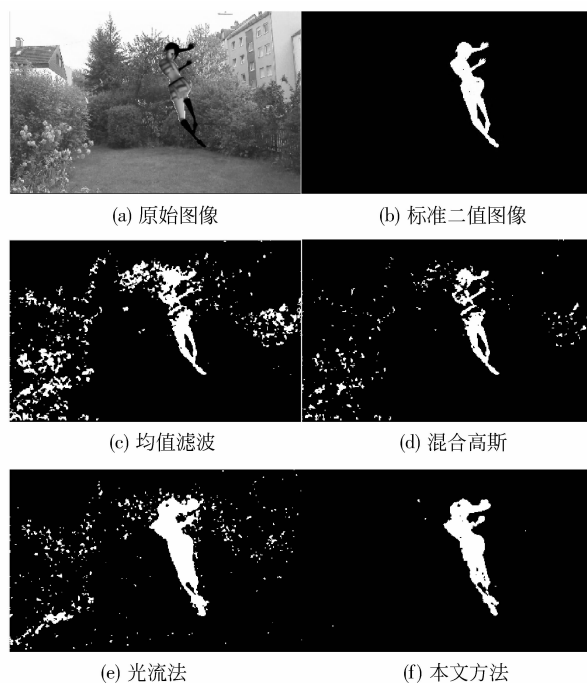


图 2 背景存在运动条件下的运动目标检测
Fig. 2 Object detecting based on moving background

法能很好地去除运动背景, 检测出完整的目标区域, 并很好的去除了由于背景物体运动产生的噪音。

利用上述实验得到的结果对均值滤波、混合高斯、光流法以及融合光流速度场的背景减除法进行像素级的分析。表 1 给出了分析对比结果, TP 与 TN 分别是指将目标样本判断正确与错误的比例, FP 与 FN 分别是指将非目标样本判断错误与正确的个数。对于 video4 这种复杂的户外场景, 背景存在快速明显的运动, 均值滤波存在大量的由背景运动产生的噪音, 并且检测到的目标区域也不完整; 相比之下, 混合高斯建模受背景运动的影响较小, 但是依然无法适应快速的背景运动, 而且因为图像帧目标所在的区域存在背景运动, 检测时目标像素的亮度及灰度值符合混合高斯模型分布, 所以, 存在漏检现象, 在表 1 中表现为 TP 为 76.00%, 略小于均值滤波, 而 FP 为 1.73%, 比较均值滤波为好; 用光流

表 1 定量分析结果/%

Tab. 1 The results of quantitative analysis/%

检测方法	TP	TN	FP	FN
均值滤波	83.99	16.01	6.19	93.81
混合高斯	76.00	24.00	1.73	98.27
光流法	96.92	3.08	4.52	95.48
本文方法	94.50	5.50	1.25	98.75

法直接检测效果也不理想,受背景运动影响较大,FP 为 4.52%,仅比均值滤波略好;融合光流速度场背景建模效果很好,不仅克服了背景运动的影响,还能较完整地描述目标区域,目标区域的错误检测仅为 5.50%,比均值滤波、混合高斯建模的错误检测率 16.01%、24.00% 好了很多,同时,FP 也比混合高斯少了 30%,而且从图中可以清楚地看出,FP 主要集中在目标物体的边缘,这是因为当目标物体运动时,其周围存在光流场,对后续运动目标检测影响较小。

3.2 光照改变条件下的运动目标检测

图 3 左边一列是在光照强度没有发生改变时的检测结果,右边一列是在光照发生突然变化时的检测结果,实验对比了当场景中光照改变时均值滤波、混合高斯、光流法和本文融合光流速度场背景建模的检测效果。video8 的特点是背景没有运动物体,但是在检测过程中右上角区域会出现光照变化。图 3(a)(b)(c)分别是视频中截取的第 1、899、1 088 帧,第 1 帧和第 1 088 帧右上角光照发生了明显变化,给检测和更新带来了很大的困难。

从图 3 检测结果中可以看出,利用均值滤波和混合高斯背景建模检测到的前景并没有消除光照条件变化带来的影响,并且在该区域出现的运动物体没被检测到;video8 视频中背景没有运动物体,因此融合光流速度场背景建模建立的背景模型中光流速度几乎处处为 0。此时融合光流速度场背景建模检测到的前景与光流法检测到的前景几乎一样。当光照没有变化时,融合光流速度场背景建模检测到的前景与均值滤波、混合高斯背景建模得到的效果一样,当光照发生变化时,融合光流速度场背景建模不仅能很好地检测到目标区域,而且能很好地消除光照变化的影响。

video8 视频中,从 138 帧光照开始变化到 597 帧,光照恢复为背景学习时的样子,从 970 帧开始光照又发生了一次变化,知道第 1 189 帧才恢复成原来的样子,共 680 帧。从表 2 的结果可以看出均值滤波和混合高斯容易受光照变化影响,当背景更新率为 0 时,只要背景光照发生变化,均值滤波和混合高斯建模就会把因为光照变化的部分误检为运动目标。当加入更新率后,误检的情况有所改善,但是受背景更新影响,检测到的目标前景出现了拖影和漏检。相比之下,本文方法因为结合了图像的前后帧信息计算背景模型,所以只有在光照改变的短时间内出现少量误报。

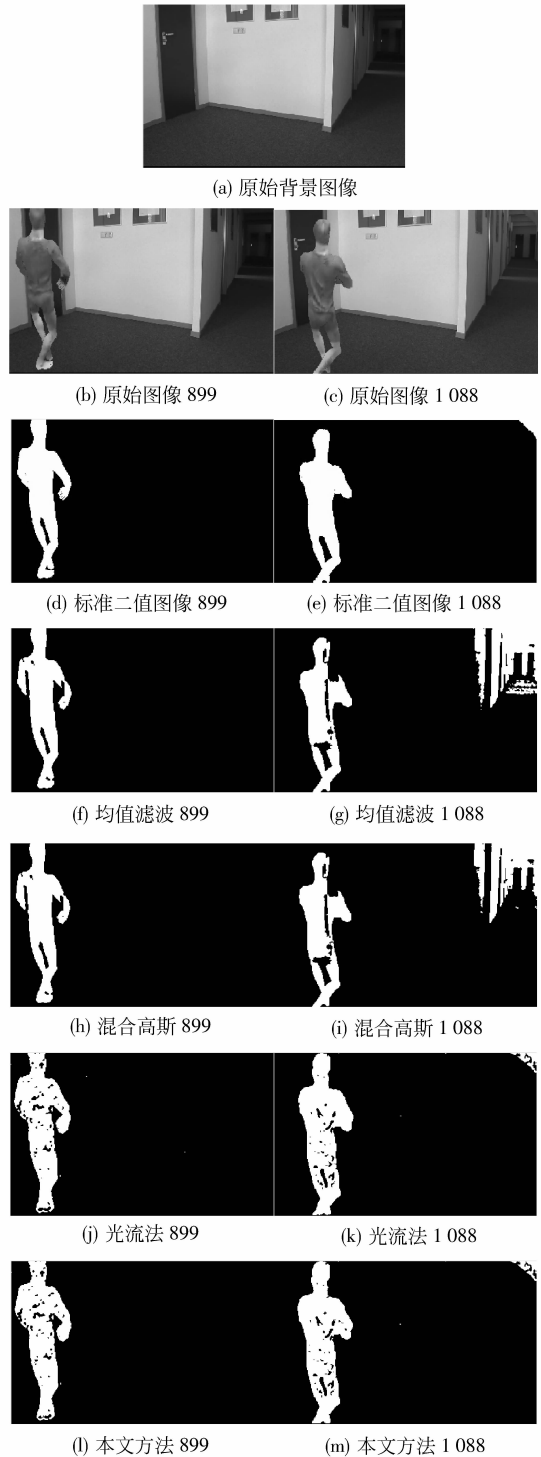


图 3 背景光照变化条件下的运动目标检测

Fig. 3 Object detecting based on illumination changes

3.3 有遗留物体条件下的运动目标检测

图 4 第 1 列是汽车在运动时的检测结果,第 2 列是同一辆汽车静止在场景中的检测结果。实验对比了均值滤波、混合高斯和本文融合光流速度场背景建模算法在场景中有遗留物体的检测效果。

表 2 光照变化错误检测结果/%

Tab. 2 The results of wrong detect based on illumination changes/%

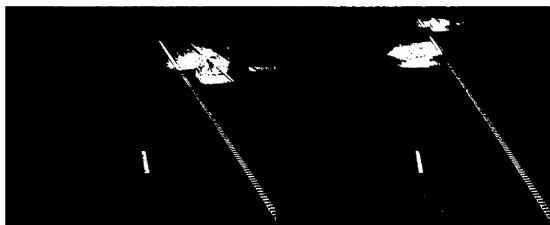
学习率	$\alpha = 0$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
均值滤波	100	65.88	18.24	11.03
混合高斯	100	45.59	15.44	9.20
本文方法	4.12	4.12	4.12	4.12



第 2 958 帧

第 4 505 帧

(a) 原始图像



(b) 均值滤波



(c) 混合高斯



(d) 本文算法

图 4 有遗留物体条件下的运动目标检测

Fig. 4 Object detecting based on left object

如图 4 所示,第 1 行第 1 列和第 2 列分别是 Vssn05 tunnelvideo 中的第 2 958、4 505 帧。在第 2 958 帧,当汽车进入场景时,均值滤波、混合高斯和本文融合光流速度场背景建模都检测到汽车的运动,但是在第 4 505 帧,当汽车停止不动后,均值滤波和混合高斯依然把静止的汽车检测为运动目标,而本文算法很好的把静止的汽车检测为背景。

tunnelvideo 视频中共有 8 005 帧,其中 3 个场景大约 5 142 帧存在遗留车辆。由表 3 的结果可以看出均值滤波和混合高斯容易受遗留车辆的影响,把遗留物体误检为运动目标。当加入更新率后,经过一段时间的背景更新,可以把遗留的车辆学习到背景中,但是受背景学习的影响,非遗留区域容易出现拖影和漏检的现象,尤其当学习率过大时,这种现象严重影响了正常运动目标检测。相比之下,本文方法结合了图像前后帧信息计算背景模型,在物体遗留的短时间内出现少量误报,大部分情况把遗留物体检测为背景。

表 3 遗留物体错误检测结果/%

Tab. 3 The results of wrong detect based on left objects/%

学习率	$\alpha = 0$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
均值滤波	100	62.26	20.11	13.56
混合高斯	100	47.59	14.63	10.22
本文方法	3.65	3.65	3.65	3.65

4 结 论

提出了一种融合光流速度的背景建模方法,把光流速度引入背景模型,充分利用了像素的空间信息、时间信息和光流速度信息,更好地描述了背景运动。算法既能从整体上描述背景运动,又能很好地描述每个像素的运动,从而减少了运动背景在运动目标检测中的影响,并能适应背景光照变化、物体丢失和遗留等因素的影响,能得到更加鲁棒的运动目标检测结果。本文算法在多个视频场景上与经典的均值滤波、混合高斯背景建模方法、光流法进行比较,实验结果表明本文提出的融合光流速度场背景建模方法能很好地适应背景光照变化、物体丢失和遗留等因素的影响,在保持较低漏检数的情况下,能显著地去掉运动背景的影响。

融合光流速度的背景建模方法也存在一些不足之处,即当目标区域的颜色特征与具有显著运动的背景区域重叠时,融合光流速度的背景建模方法仍然可能漏检,这是今后将要解决的问题。

参考文献 (References)

- [1] Li L, Luo R. Context-controlled adaptive background subtraction [EB/OL]. [2009-12-3]. http://www.cvg.rdg.ac.uk/PETS2006/PETS2006_PROCEEDINGS.pdf.

- [2] Wren C, Azarbayejani A. Pfinder: Real-time tracking of the human body [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 780-785.
- [3] Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction [J]. Pattern Recognition, 2004, 2(17): 28-31.
- [4] Elgammal A, Duraiswami R. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance [EB/OL]. [2009-12-3]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.20.215&rep=rep1&type=pdf>.
- [5] Shlens J. A tutorial on principal component analysis [C] // Systems Neurobiology Laboratory. San Diego, American: University of California, 2005: 1-26.
- [6] Wang liang, Hu Weiming. A survey of visual analysis of human motion [J]. Journal of Computers, 2002, 25(3): 225-237. [王亮, 胡卫明. 人运动的视觉分析综述 [J]. 计算机学报, 2002, 25(3): 225-237.]
- [7] Papenberg N, Bruhn A. Highly accurate optic flow computation with theoretically justified warping [J]. International Journal of Computer Vision, 2006, 67(2): 141-158.
- [8] Barron J, Fleet D. Performance of optical flow techniques [J]. International Journal of Computer Vision, 1994, 12(1): 43-77.
- [9] Bruhn A, Weickert J. Variational optical flow computation in real time [J]. Image Processing, IEEE Transactions on, 2005, 14(5): 608-615.
- [10] Bruhn A, Weickert J. Lucas/Kanade Meets Horn/Schunck: Combining local and global optic flow methods [J]. International Journal of Computer Vision, 2005, 61(3): 211-231.
- [11] Horn B, Schunck B. Determining optical flow [J]. Artificial Intelligence, 1981, 17(1-3): 185-203.
- [12] Yang Changqing, Wang Xiaotong. Fast point matching method based on feature optical flow [J]. Opto-Electronic Engineering, 2006, 33(004): 85-88. [杨常清, 王孝通. 基于特征光流的角点匹配快速算法 [J]. 光电工程, 2006, 33(004): 85-88.]
- [13] Feghali R, Mitiche A. Spatiotemporal motion boundary detection and motion boundary velocity estimation for tracking moving objects with a moving camera: a level sets PDEs approach with concurrent camera motion compensation [J]. IEEE Transactions on Image Processing, 2004, 13(11): 1473-1490.
- [14] Gautama T, Van Hulle M. A phase-based approach to the estimation of the optical flow field using spatial filtering [J]. IEEE Transactions on Neural Networks, 2002, 13(5): 1127-1136.
- [15] Grossberg S, Kelly F. Neural dynamics of binocular brightness perception [J]. Vision Research, 1999, 39(22): 3796-3816.