

一种视频监控中基于航迹的运动小目标检测算法

孙怡峰 吴 疆* 黄严严 汤光明

(中国人民解放军战略支援部队信息工程大学 郑州 450000)

摘 要: 针对视频监控中运动小目标难以检测的问题, 该文提出一种基于航迹的检测算法。首先, 为了降低检测漏警率, 提出区域纹理特征与差值概率融合的自适应前景提取方法; 其次, 为了降低检测虚警率, 设计航迹关联的概率计算模型以建立疑似目标在视频帧间的关联, 并设置双门限以区分疑似目标中的真实目标与虚假目标。实验结果表明, 与多种经典算法相比, 该算法能对定量范围内的运动小目标以更低的漏警率和虚警率实施准确检测。

关键词: 运动目标检测; 小目标检测; 航迹关联

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2019)11-2744-08

DOI: 10.11999/JEIT181110

A Small Moving Object Detection Algorithm Based on Track in Video Surveillance

SUN Yifeng WU Jiang HUANG Yanyan TANG Guangming

(Information Engineering University, PLA Strategic Support Force, Zhengzhou 450000, China)

Abstract: To solve the problem that small moving object is difficult to be detected in video surveillance, a track-based detection algorithm is proposed. Firstly, in order to reduce missing alarm, an adaptive foreground extraction method combining regional texture features and difference probability is presented. Then, for reducing false alarm, the probability computing model of track correlation is designed to establish the correlation of suspected objects between frames, and double-threshold are set to distinguish between true and false positive. Experimental results show that compared with many classical algorithms, this algorithm can accurately detect small moving object within the quantitative range with lower missing and false alarm.

Key words: Moving object detection; Small object detection; Track correlation

1 引言

目标检测是计算机视觉领域的基本任务之一, 现有目标检测的研究对象大都是占图像稍大部分的目标, 缺少对小目标的关注^[1]。国际光学工程学会(Society of Photo-optical Instrumentation Engineers, SPIE)把小目标定义为成像小于图像尺寸0.12%的目标^[2]。一般来说, 小目标是因自身体积小或成像距离远形成的, 如远处无人机、车辆等, 有限的像素和信息, 使得对其检测仍然充满挑战^[3]。及时地发现小目标, 有利于提高预警距离, 在军用、民用领域都将发挥巨大作用, 如可用于战场侦测敌方目标、自动驾驶避让障碍等。

本文主要研究视频监控中的运动小目标, 且假定摄像机固定。此外, 本文研究的小目标是占一定图像尺寸且大致满足SPIE定义的面目标, 而非成像极小的点目标。在该前提下, 可采用基于深度学

习的检测方法, 区域卷积神经网络(Regions with CNN, RCNN)是这类方法的开山之作, 文献[4]提出的Faster RCNN算法以区域推荐网络(Region Proposal Network, RPN)替代RCNN中的选择性搜索, 较RCNN在检测精度、速度上大幅提升, 后续衍生出部分关注小目标检测的算法: 文献[1]改进RPN获得推荐区域, 提出结合上下文信息的CNN模型对推荐区域进行分类和回归; 文献[5]提出特征金字塔网络(Feature Pyramid Network, FPN)以融合位置信息强的浅层特征和语义信息强的深层特征。这类方法尽管因采用特征融合的策略提高了检测表现, 但由于缺少公开的小目标数据集, 实验中的小目标通常并未达到SPIE定义的要求, 当目标小到一定程度时, 检测漏警率很高。还可采用基于目标运动特性的检测方法: 帧差法通过视频相邻帧的差分运算判断对应像素值的改变, 从而提取运动目标; 光流法建立运动向量场根据像素点的速度向量特征区分运动目标; 背景建模法将视频各帧图像与构建的背景模型做差分, 将区别较大的像素区域

收稿日期: 2018-11-29; 改回日期: 2019-03-14; 网络出版: 2019-05-17

*通信作者: 吴疆 liam181113@163.com

视为运动目标,该方法简单,适用范围更广^[6]。混合高斯模型(Mixtures Of Gaussians, MOG), MOG2^[7],视觉背景提取(Visual Background extractor, ViBe)和ViBe+^[8]均是经典的背景建模法,MOG算法通过多个高斯分布的混合对各个背景像素建模;MOG2算法为每个背景像素选择适当数量的高斯分布;ViBe算法建立像素点的样本集,通过将新像素点值与样本集进行匹配来判断是否属于背景;通过对距离函数、更新因子等改进,得到ViBe+算法^[9]。这类方法的主要问题是检测虚警率高,难以正确区分复杂动态背景、噪声与真实运动小目标。

针对以上两类检测方法存在的问题,本文提出一种视频监控中基于航迹的运动小目标检测算法。为了降低检测漏警率,提出区域纹理特征与差值概率融合的自适应前景提取方法,使介于预设面积范围内的疑似运动小目标(下称疑似目标)能被检出;为了降低检测虚警率,提出航迹关联的双门限虚假目标过滤方法,通过建立上一帧航迹与当前帧疑似目标的关联,计算各航迹的维持时间与撤销判断次数从而基于双门限过滤虚假目标。实验结果表明,本文算法能够有效地检测视频监控中的运动小目标。

2 问题分析

基于深度学习的检测方法对单帧图像进行检测时,能以外接矩形框形式给出目标的位置信息和识别目标所属的类别,但由于小目标特征不明显,此类方法对小目标的检测漏警率高。基于目标运动特性的检测方法依据目标在视频帧间因位置移动产生的运动信息对存在的运动目标进行检测,对于视频各帧均能给出相应的运动前景二值图,图1(c)是利用单高斯背景建模^[10]得到的运动前景二值图,图中白色像素点被认为属于运动目标,但真实目标如图1(b)所示。评估算法的检测性能常采用如文献^[6]中的定量法,其规则为:定义TP(True Positive)为被正确检测为运动前景像素点的个数, TN(True Negative)为被正确检测为背景像素点的个数, FP(False

Positive)为被错误检测为运动前景像素点实际属于背景的个数, FN(False Negative)为被错误检测为背景像素点实际属于运动前景的个数,通过统计各帧检测结果中TP, TN, FP和FN值可计算实验比较量,如漏警率、虚警率等。由图1(c)可见,基于目标运动特性的检测方法虚警率高,其原因是受到动态背景、噪声和不连通前景的干扰,由此产生大量虚假目标。同时,由于小目标距离远,成像时色彩和纹理信息趋于弱化,存在其与背景的颜色亮度基本相同的情况,也易造成小目标被漏检。

3 本文算法

实施运动小目标检测,不仅可以利用视频单帧图像中的目标特征信息,还可以利用目标在帧间的关联信息。因此,本文首先利用所提区域纹理特征与差值概率融合的自适应前景提取方法实现对视频各帧中疑似目标的初步检测,然后建立疑似目标在视频帧间的关联关系,并通过所提航迹关联的双门限虚假目标过滤方法区分虚假目标和运动小目标,最后输出各帧检测结果。

3.1 疑似目标的初步检测——区域纹理特征与差值概率融合的自适应前景提取

3.1.1 区域纹理特征与差值概率融合的自适应二值化阈值设置

在摄像机固定的条件下,背景建模法常被用于运动目标检测,其计算视频各帧图像与背景模型的差值,根据预设固定阈值区分前景和背景,差值大于阈值的像素点被认为是前景。固定阈值设置过高易使小目标被漏检,而过低会导致全局性的虚检。为此,本文提出首先对视频各帧进行直方图灰度拉伸^[11]以提高图像对比度,然后采用单高斯背景建模法构建背景模型,最后将由各帧图像与背景模型得到的差值图像划分为若干个区域,基于各个区域的纹理特征和差值概率自适应地确定二值化阈值。

基于单高斯背景建模法得到的像素差值可与预设阈值比较以区分前景和背景,一般来说,差值在0附近的像素点应属于背景,较大差值的像素点应属于前景,阈值可由差值概率密度分布确定。以本

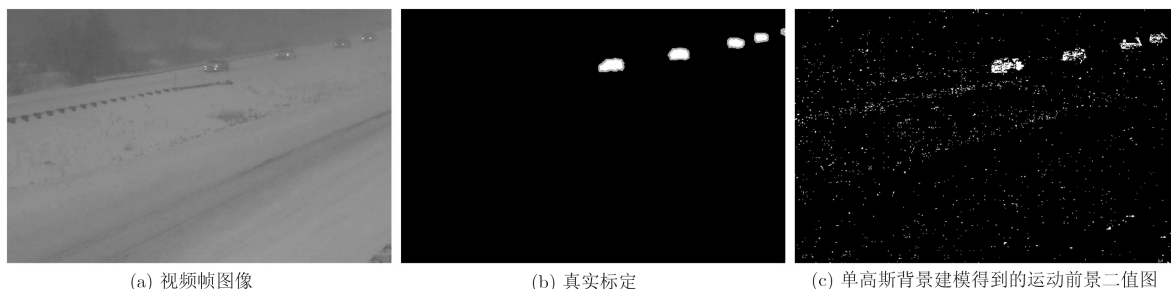


图1 视频帧图像与单高斯背景建模得到的运动前景二值图

文自建视频库中Camera 01视频为例,其差值图像直方图如图2所示,可见该差值分布具有递减、长尾的特点,在其它视频中也表现出该特点。本文选取具有类似递减、长尾特性的分布(Gamma, 指数, Weibull, Pareto分布)对上述差值分布进行拟合,以决定系数^[12]度量拟合优度。实验表明Weibull分布拟合效果最好,因此,本文认为差值服从Weibull分布。该分布的累积分布函数如式(1)所示,其中 η , m 分别表示比例和形状参数。由此可依据式(2)确定阈值,其中 thr 表示阈值, $G^{-1}(\cdot)$ 表示Weibull分布的逆累积分布函数, p_G 为预设概率,式(2)表示差值概率为 p_G 时选取的阈值,但该阈值是全局的,未充分考虑图像的纹理特征。

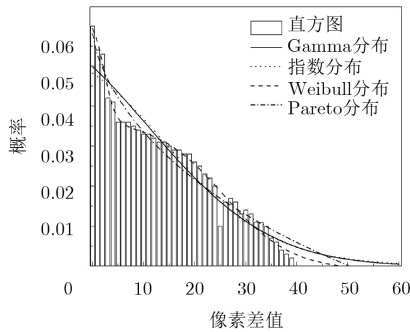


图2 差值分布拟合曲线

$$G(\partial; \eta, m) = \begin{cases} 1 - e^{-(\partial/\eta)^m}, & \partial > 0 \\ 0, & \partial \leq 0 \end{cases} \quad (1)$$

$$\text{thr} = G^{-1}(p_G; \eta, m) \quad (2)$$

对减少漏检和虚检而言,在图像不同区域设置不同的二值化阈值将比图像全局使用统一的阈值更为有效。划分区域的粒度越细,效果应越明显。因此,本文将差值图像划分为固定数目 N^2 个区域,记为 $A_q (q = 1, 2, \dots, N^2)$,经验性地设置 $N = 10$,各区域的尺寸由式(3)决定,其中 w, h 分别表示划分区域的宽和高, W, H 分别表示差值图像的宽和高, $\lfloor \cdot \rfloor$ 代表向下取整。若存在边缘未被划分的区域,将其纳入相邻区域。确定各区域的阈值应考虑其纹理特征,利用灰度共生矩阵得到的纹理特征之一“倒数差分矩IDF^[13]”反映了图像纹理局部变化的大小。若图像纹理变化缓慢, IDF较大,反之较小。IDF的计算如式(4)所示,其中 a, b 是距离为 d , 角度为 ϕ 的某两像素的灰度级, ϕ 常取 $0^\circ, 45^\circ, 90^\circ$ 或 135° , $P_{\phi,d}(a, b)$ 表示共生矩阵的元素,即 (a, b) 灰度级对出现的次数。将差值图像按上述划分方式得到的各个区域中,存在前景运动目标或复杂背景(如树木)的区域纹理变化快, IDF小;仅存在纯净背景(如天空)的区域纹理变化慢, IDF大。本文

按式(5)计算上述区域 A_q 的倒数差分矩IDF $_q$ 以度量区域纹理变化快慢,其中 $\phi \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ 。

$$\left. \begin{aligned} w &= \lfloor W/N \rfloor \\ h &= \lfloor H/N \rfloor \end{aligned} \right\} \quad (3)$$

$$\text{IDF} = \sum_{a,b} \frac{P_{\phi,d}(a, b)}{1 + (a - b)^2} \quad (4)$$

$$\text{IDF}_q = \frac{1}{4} \sum_{\phi} \sum_{a_q, b_q} \frac{P_{\phi,1}(a_q, b_q)}{1 + (a_q - b_q)^2} \quad (5)$$

为了减少漏检,对存在前景运动目标或复杂背景的区域可设置低阈值,由此导致的虚检通过后续方法过滤,而对仅存在纯净背景的区域应设置高阈值以减少虚检。因此,本文按式(6)确定各区域的自适应阈值,其根据各区域的倒数差分矩修正差值概率,其中 $\text{IDF}_{\max} = \max\{\text{IDF}_1, \text{IDF}_2, \dots, \text{IDF}_{N^2}\}$, p_G 经验性地设置为0.9。

$$\text{thr}_q = G^{-1} \left(\frac{\text{IDF}_q}{\text{IDF}_{\max}} p_G; \eta, m \right) \quad (6)$$

3.1.2 采用质心标记的疑似目标初步检测

图3可视化了疑似目标初步检测的过程,图3(a)是用矩形框标记了小目标、尺寸为 $720 \text{ pix} \times 480 \text{ pix}$ 的blizzard视频帧图像。图3(b)是经过直方图灰度拉伸得到的图像,与图3(a)相比,对比度明显提高。图3(c)是利用单高斯背景建模得到的运动前景二值图,其采用3.1.1节中的二值化阈值设置方法,图中白色像素代表包含了真实运动目标的前景,由于噪声等干扰,导致前景中存在虚假目标。采用中值滤波^[14]的非线性平滑处理后的运动前景二值图如图3(d)所示,相比于图3(c),噪声减少,但前景连通性欠佳。如图3(e)所示,通过形态学闭运算^[14]先膨胀后腐蚀的处理,一定程度上缓解了前景不连通的现象。在面积过滤处理中,计算前景中各连通区的面积,设置需要检测目标的尺寸占比,如选择占该图像尺寸0.10%~0.12%的目标进行检测,一方面滤除了部分噪声干扰,另一方面也实现了可控目标大小地检测运动小目标。相比于图3(e),图3(f)中前景连通区数量减少,图中有多个疑似目标,为方便描述,以各自质心坐标标记,记为 $(\text{centroidx}^i, \text{centroidy}^i)$,其中 $\text{centroidx}^i, \text{centroidy}^i (i = 1, 2, 3, \dots)$ 分别表示第 i 个疑似目标的质心在图像坐标下的 x 坐标、 y 坐标,而原图实际上只有一个满足SPIE定义的运动小目标,有效过滤其它虚假目标是本文运动小目标检测的关键。

3.2 航迹关联的双门限虚假目标过滤

3.2.1 虚假目标过滤的总体思想

用前述方法只能得到视频各帧以质心坐标表示

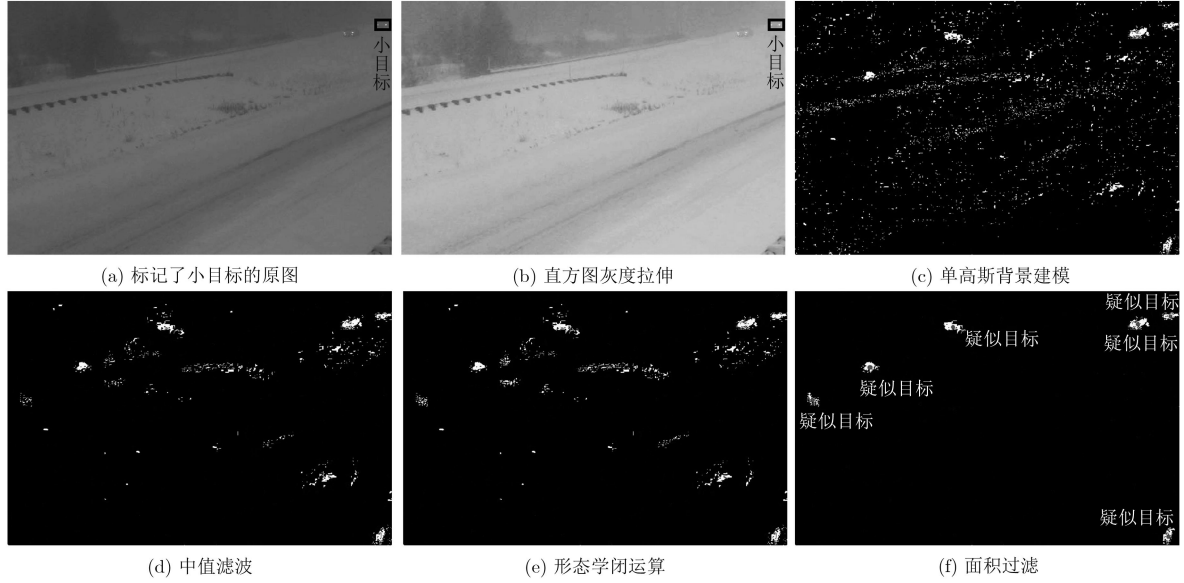


图3 疑似目标初步检测的可视化过程

的疑似目标，需要判断相邻帧中哪些疑似目标属于同一个目标，从而得到同一个目标在视频帧间的一系列位置以建立航迹(航迹)。对航迹的处理包括航迹起始、关联、撤销3个部分^[15]。其中，航迹关联是判断当前帧疑似目标是否属于已有航迹的方法。对于同一个目标，若航迹关联1次，意味着航迹维持了1帧，由此计算航迹维持帧数，本文将其称为航迹维持时间。

目标状态是由目标特征(如速度、位置)构成的向量，Kalman滤波、粒子滤波等常被用于估计目标状态。当目标运动模型为线性时，Kalman滤波是最小均方误差准则下的最佳线性滤波。由于距离摄像机远，小目标非线性的运动在成像中表现不明显，可以认为其运动趋于线性、匀速。因此，本文选取Kalman滤波估计目标状态。噪声无法形成稳定航迹，而树叶晃动等复杂背景的运动是非线性的，基于Kalman滤波进行航迹关联，不能得到较长的航迹维持时间。而由不连通前景产生的虚假目标，尽管其航迹维持时间会比因树叶等产生的虚假目标航迹维持时间更长，但仍然是短时的，可以通过设置航迹维持时间的门限进行过滤。

3.2.2 目标状态的Kalman滤波

本文选取目标质心在图像坐标下的位置和速度作为目标特征，记为 $\mathbf{X}_k = [x_k \ x_{v_k} \ y_k \ y_{v_k}]$ ，其中， \mathbf{X}_k 为某目标在第 k 帧的目标状态， $x_k, x_{v_k}, y_k, y_{v_k}$ 分别代表第 k 帧该目标质心的 x 坐标、 x 方向速度、 y 坐标和 y 方向速度。由3.1.2节方法获取相邻帧中同一个疑似目标的质心位置(centroidx, centroidy)计算得到的目标特征称为目标量测，将某目标在第 k 帧的目标量测记为 $\mathbf{Z}_k = [\dot{x}_k \ \dot{x}_{v_k} \ \dot{y}_k \ \dot{y}_{v_k}]$ ，其计算如式(7)

所示。其中，centroidx_k, centroidy_k, centroidx_{k-1}, centroidy_{k-1}依次代表第 k 帧疑似目标质心的 x 坐标、 y 坐标观测值和第 $k-1$ 帧中同一个疑似目标质心的 x 坐标、 y 坐标观测值， t 为视频帧周期。

$$\left. \begin{aligned} \dot{x}_k &= \text{centroidx}_k \\ \dot{x}_{v_k} &= (\text{centroidx}_k - \text{centroidx}_{k-1})/t \\ \dot{y}_k &= \text{centroidy}_k \\ \dot{y}_{v_k} &= (\text{centroidy}_k - \text{centroidy}_{k-1})/t \end{aligned} \right\} \quad (7)$$

Kalman滤波依据上一帧目标状态的估计值和当前帧目标量测来估计当前帧目标状态，使其尽可能地拟合目标真实的运动，核心公式如式(8)~式(12)所示。

$$\hat{\mathbf{X}}_{k|k-1} = \mathbf{F} \hat{\mathbf{X}}_{k-1|k-1} \quad (8)$$

$$\mathbf{P}_{k|k-1} = \mathbf{F} \mathbf{P}_{k-1|k-1} \mathbf{F}^T + \mathbf{Q} \quad (9)$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}^T \mathbf{S}_k^{-1} \quad (10)$$

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{K}_k \mathbf{Y}_k \quad (11)$$

$$\mathbf{P}_{k|k} = (\mathbf{E} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_{k|k-1} \quad (12)$$

式(8)中 $\hat{\mathbf{X}}_{k|k-1}$ 是在未获得 \mathbf{Z}_k 时对第 k 帧目标状态的预测值， \mathbf{F} 为状态转移矩阵，由于小目标的运动趋于线性、匀速，本文 \mathbf{F} 如式(13)所示， $\hat{\mathbf{X}}_{k-1|k-1}$ 是通过Kalman滤波对第 $k-1$ 帧目标状态的估计值，当 $k-1=1$ ，需要赋初值 $\hat{\mathbf{X}}_{1|1} = [\text{centroidx}_1 \ 0 \ \text{centroidy}_1 \ 0]$ 。式(9)中 $\mathbf{P}_{k|k-1}$ 是指在未获得 \mathbf{Z}_k 时对第 k 帧协方差的预测值， $\mathbf{P}_{k-1|k-1}$ 是通过Kalman滤波对第 $k-1$ 帧协方差的估计值，依据文献[16]对 $\mathbf{P}_{1|1}$ 赋值， \mathbf{Q} 表示过程噪声协方差矩阵。式(10)中 \mathbf{K}_k

表示Kalman增益, \mathbf{H} 表示观测矩阵, $\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}$ 是新息的协方差, \mathbf{R} 表示观测噪声协方差矩阵。式(11)中 $\hat{\mathbf{X}}_{k|k}$ 是指获得 \mathbf{Z}_k 后对第 k 帧目标状态的估计值, $\mathbf{Y}_k = \mathbf{Z}_k - \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1}$ 为新息。式(12)中 $\mathbf{P}_{k|k}$ 是指获得 \mathbf{Z}_k 后对第 k 帧协方差的估计值, \mathbf{E} 为单位阵。

$$\mathbf{F} = \begin{pmatrix} 1 & t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (13)$$

3.2.3 航迹关联的概率计算模型

由3.1.2节初步检测得到的各个疑似目标对应了一条初始航迹, 其目标状态和协方差矩阵的初值由 $\hat{\mathbf{X}}_{1|1}$, $\mathbf{P}_{1|1}$ 给定。虚假目标过滤的过程如图4所示, 为了排除不连通前景的干扰, 首先遍历当前帧各个疑似目标, 若其它疑似目标的质心落在以该疑似目标的质心为圆心、半径为 r 的圆中, 则以平均质心位置表示这些疑似目标。为了避免来源于不同目标的运动前景被视为同一个目标产生的多个不连通前景, 经验性设置了较小值 $r = 10$ 。为了对视频帧间同一个疑似目标进行关联, 提出航迹关联的概率计算模型。在第 $k-1$ 帧航迹与第 k 帧疑似目标的航迹关联中, 对于 $k-1$ 帧的各条航迹, 计算以该航迹在第 $k-1$ 帧中关联的疑似目标的质心为圆心, $D = 2t\sqrt{(\hat{\mathbf{X}}_{k-1|k-1}[0])^2 + (\hat{\mathbf{X}}_{k-1|k-1}[2])^2}$ 为半径的圆内的第 k 帧疑似目标的关联概率。对于不在圆内的第 k 帧疑似目标, 将关联概率 p 记为0。反之, 若航迹是第 $k-1$ 帧新起始的, 则按式(14)计算航迹关联概

率, 式(14)是在目标量测 \mathbf{Z}_k 与 $\mathbf{Z}_{1:k-1}$ 相关联的概率服从高斯分布前提下构建的, 其中 $\mathbf{Z}_{1:k-1}$ 表示某条航迹 $\mathbf{Z}_{1:k-1} = \{\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_{k-1}\}$, $p(\mathbf{Z}_k|\mathbf{Z}_{1:k-1})$ 代表 \mathbf{Z}_k 与 $\mathbf{Z}_{1:k-1}$ 的关联概率, c 为 \mathbf{Y}_k 的维数, 否则按式(15)计算。

$$p = p(\mathbf{Z}_k|\mathbf{Z}_{1:k-1}) = \frac{1}{(2\pi)^{c/2} \sqrt{\mathbf{S}_k}} \exp\left(-\frac{1}{2} \mathbf{Y}_k^T \mathbf{S}_k^{-1} \mathbf{Y}_k\right) \quad (14)$$

$$p = p(\theta_k) = (1 - \gamma)p(\mathbf{Z}_k|\mathbf{Z}_{1:k-1}) + \gamma p(\theta_{k-1}) \quad (15)$$

式(15)中 γ 为权重系数, 经验性地设置为0.3, $\theta_k = \{\mathbf{Z}_{1:k} \text{ 表示一条真实航迹}\}$, 即 $\mathbf{Z}_{1:k}$ 来源于同一个目标。至此, 对于视频各帧, 按式(14)或式(15)可计算上一帧航迹与当前帧疑似目标的关联概率。设上一帧航迹有 U 条, 当前帧疑似目标有 V 个, 则可建立关联概率表 $\text{APT}_{U \times V}$ 。按表1所示的关联规则进行关联, 对各条航迹, 用变量 $\text{Rel}(\mathbf{Z}_{1:k})$ 来表示航迹维持时间, 其初值为0, 若航迹关联1次, 对应航迹维持时间 $\text{Rel}(\mathbf{Z}_{1:k})$ 加1。

3.2.4 基于双门限的虚假目标过滤

上一帧航迹与当前帧疑似目标的关联关系存在3种情况: 一是航迹存在关联的疑似目标; 二是航迹不存在关联的疑似目标; 三是疑似目标不存在关联的航迹。对于关联上的航迹和疑似目标, 不是简单地认为对应疑似目标就是运动小目标, 而是设置航迹维持时间门限予以确定。对于上一帧任意一条航迹, 若关联上了当前帧的某个疑似目标且航迹维持时间超过预设门限 rel_num , 即 $\text{Rel}(\mathbf{Z}_{1:k}) >$

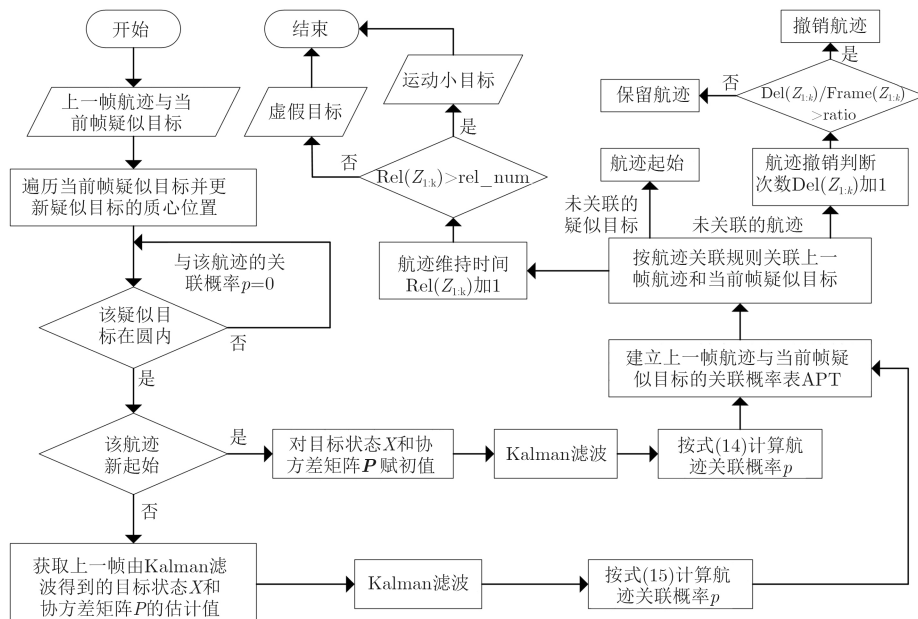


图4 航迹关联的双门限虚假目标过滤

表1 航迹关联规则

(1) do;
(2) for $u = 1, 2, \dots, U$;
(3) 寻找 $\text{APT}_{U \times V}$ 中第 u 行中的最大值 a_{uv} ，记录其列号 v ;
(4) if第 v 列的最大值等于 a_{uv} ;
(5) break;
(6) end if;
(7) end for;
(8) 关联 $\{\mathbf{Z}_k^u\}$ 与 $\{\theta_{k-1}^u\}$ ，删除 $\text{APT}_{U \times V}$ 中的第 u 行和第 v 列元素， $U = U - 1, V = V - 1$;
(9) while $\text{APT}_{U \times V}$ 中存在元素大于0.

rel_num，则认为该疑似目标是运动小目标，否则是虚假目标，实验表明rel_num为5时检测效果最好。对于未关联的上一帧航迹，也并非简单地撤销航迹，而是用变量 $\text{Del}(\mathbf{Z}_{1:k})$ 记录该航迹进入撤销判断的次数，其初值为0，若该航迹无疑似目标关联，则将航迹撤销判断次数 $\text{Del}(\mathbf{Z}_{1:k})$ 加1，当航迹撤销判断次数与航迹总帧数的比值大于另一预设门限ratio时，即 $\text{Del}(\mathbf{Z}_{1:k})/\text{Frame}(\mathbf{Z}_{1:k}) > \text{ratio}$ ，其中 $\text{Frame}(\mathbf{Z}_{1:k}) = \text{Del}(\mathbf{Z}_{1:k}) + \text{Rel}(\mathbf{Z}_{1:k})$ ，认为该航迹不稳定，才予以撤销，否则继续保留该航迹，经验性地设置 $\text{ratio} = 0.7$ 。对于未关联的当前帧疑似目标，将其作为新的航迹进行起始。需要指出的是，视频的前几帧由于还未建立稳定的航迹，使得本文检测结果类似于传统方法，但随着时间推移，通过航迹关联，将能显著改善虚检问题。

4 实验结果与分析

由于缺少公开的小目标数据集，本文同时选取运动目标检测公开数据集changeDetection dataset2014中包含了较小目标的视频和自建视频库，分别是：blizzard视频，存在雪花等动态背景；highway视频，包含的目标比SPIE定义稍大；Camera库(Camera 01, Camera 02)，相比于其它两个视频，包含更多的小目标，且存在树叶晃动、光照变化等多种干扰。需要明确的是，本文提出的运动小目标检测是对一段连续时间内运动小目标的持续检测，而不关注单帧图像的检测性能。为了验证算法的有效性，本文选取MOG2, ViBe+, Faster RCNN和文献[5]算法进行对比，下面从定性和定量两个方面对实验结果加以分析说明。

4.1 实验结果定性分析

定性实验结果如图5所示，第1行是标记了小目标的视频典型帧图像，第2~5行分别是MOG2, ViBe+, Faster RCNN和文献[5]算法的检测结果；第6行为本文算法检测结果。由图5可知，本文算法

总体上优于其余4种算法。在blizzard视频帧中，MOG2, ViBe+算法检出小目标的局部，MOG2同时产生较多虚检，Faster RCNN和文献[5]算法均未检出小目标，而本文算法较完整地检出小目标，且不存在虚假目标。在highway视频帧中，除Faster RCNN算法外，其余算法均检出小目标，MOG2同时检出大量虚假目标。在Camera 01视频帧中，本文算法正确检出5个小目标，受动态背景干扰，MOG2检出大片虚假目标，ViBe+算法检出小目标的局部，Faster RCNN未检出小目标，而文献[5]算法检出1个小目标(汽车)。这主要是因为基于深度学习的算法受目标尺寸约束大，通过多层卷积、池化操作，小目标特征已不明显，因此表现为Faster RCNN算法难以检出小目标，而在其基础上增加特征融合策略的文献[5]算法能够检出稍大的小目标。Camera 02视频帧检测结果与Camera 01类似，MOG2和ViBe+算法产生大量虚检，Faster RCNN未检出小目标，文献[5]算法检出1个小目标(稍大电动车)，而本文算法正确检出3个小目标。

4.2 实验结果定量分析

文献[6]中评估算法性能的定量法关注各个像素点的检测情况，但当同一个运动目标被检测为多个不连通前景时，会使计算机误认为出现了多个小目标，本文认为运动小目标检测问题中更应关注检出运动小目标的个数情况。由此本文提出如下算法性能评估方法：定义TP为运动小目标被检出的个数，FN为运动小目标未被检出的个数，FP为非运动小目标(各种干扰引起的虚假目标)被检测为运动小目标的个数。以漏警率MA(Missing Alarm)和虚警率FA(False Alarm)作为实验比较量，其中， $\text{MA} = \text{FN}/(\text{TP} + \text{FN})$ ，代表被漏检的运动小目标在运动小目标中的占比，该值越小越好； $\text{FA} = \text{FP}/(\text{TP} + \text{FP})$ ，代表被检测为运动小目标中虚假目标的占比，该值越小越好。对于上述各个算法的检测结果，若在原图存在运动小目标的位置上检出运动前景，则认为检出了该运动小目标，若出现多个检出的运动前景，则计算各个运动前景的连通区面积，取面积最大的连通区作为运动小目标的检测结果，而将其余前景视为虚假目标；若在原图存在运动小目标的位置上未检测出运动前景，则认为该运动小目标未被检出；若在原图不存在运动小目标的位置上检测出运动前景，认为该运动前景为虚假目标。由此，计算各帧的TP, FN和FP值，然后统计连续100帧TP, FN和FP值从而计算MA和FA值。

表2、表3分别给出了5种算法在不同视频中的MA和FA值。从定量实验结果看，本文算法在MA, FA两指标上的综合表现最佳。在blizzard视频中，

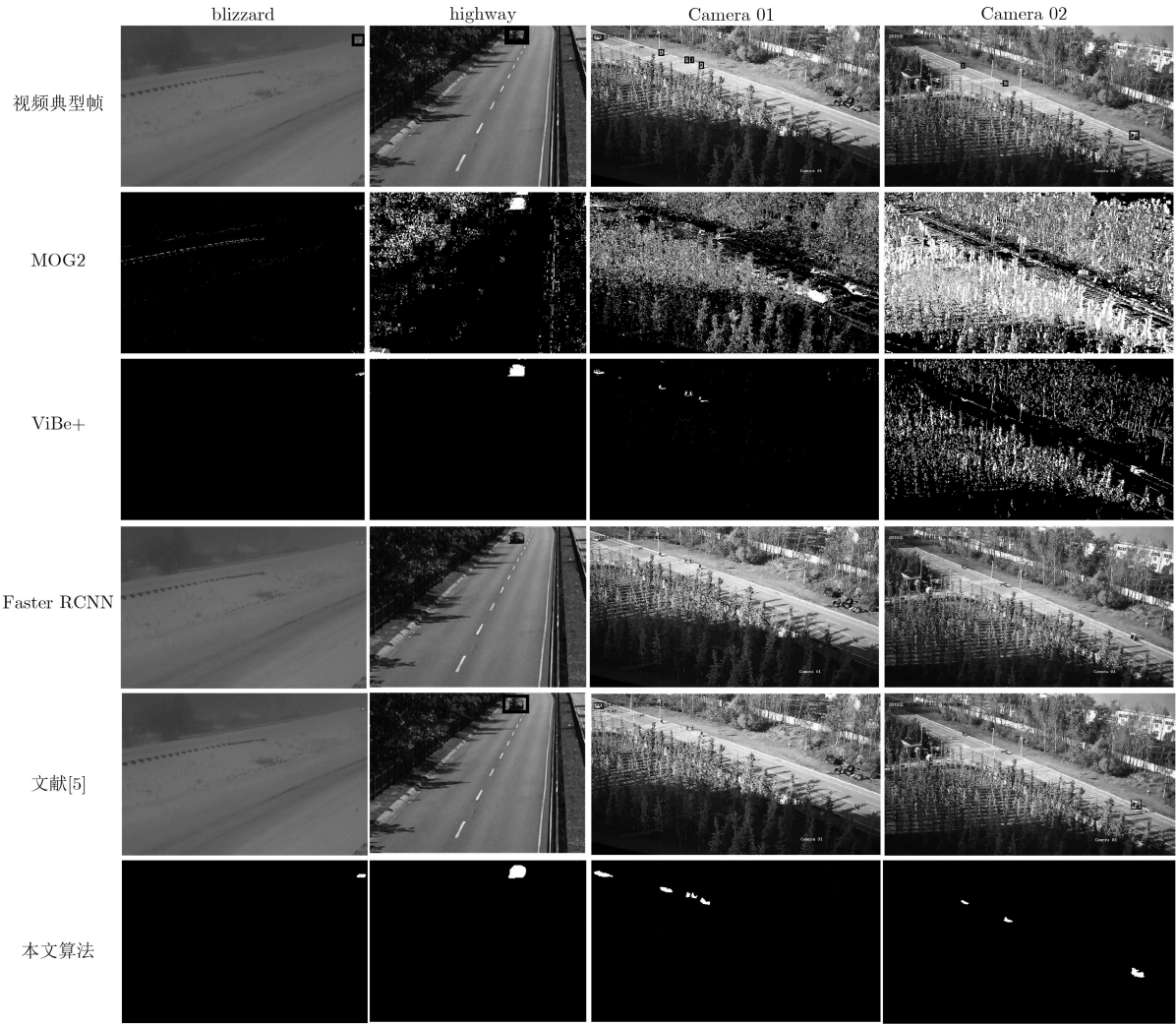


图 5 定性实验结果

表 2 5种算法在不同视频中的MA值比较

视频	图像尺寸(pix)	检测范围(%)	像素数	MOG2	ViBe+	Faster RCNN	文献[5]	本文算法
blizzard	720 × 480	0.10~0.12	345~414	0.15	0.78	1.00	1.00	0.28
highway	320 × 240	0.12~0.30	92~230	1.00	0.16	1.00	0.50	0.05
Camera 01	1920 × 1080	0.01~0.12	207~2488	0.38	0.29	1.00	0.86	0.11
Camera 02	1920 × 1080	0.01~0.12	207~2488	0.39	0.28	1.00	0.77	0.13

表 3 5种算法在不同视频中的FA值比较

视频	图像尺寸(pix)	检测范围(%)	像素数	MOG2	ViBe+	Faster RCNN	文献[5]	本文算法
blizzard	720 × 480	0.10~0.12	345~414	0.71	0.21	0.00	0.00	0.18
highway	320 × 240	0.12~0.30	92~230	0.37	0.61	0.00	0.00	0.29
Camera 01	1920 × 1080	0.01~0.12	207~2488	0.51	0.25	0.00	0.00	0.13
Camera 02	1920 × 1080	0.01~0.12	207~2488	0.52	0.17	0.00	0.00	0.14

Faster RCNN和文献[5]算法MA值最高为1, FA值最低为0, 而本文算法MA值为0.31, FA值为0.18; 在highway视频中, 本文算法MA值最低为0.10, FA值为0.23; 在Camera 01和Camera 02视频中, 除基于深度学习的算法(Faster RCNN和文献[5]算

法, 两种算法尽管FA值最低为0, 但MA值很高)外, 在小目标仅占图像尺寸0.01%~0.12%的条件下, 本文算法MA和FA值均最低, 表明在现实场景视频监控中, 本文算法能够实现对运动小目标的有效检测。

5 结束语

本文研究视频监控中的运动小目标检测问题，最突出的特点是将对该问题的处理分解为疑似目标的初步检测和虚假目标的过滤两个阶段，并以降低漏警率和虚警率为目的，对应提出区域纹理特征与差值概率融合的自适应前景提取方法和航迹关联的双门限虚假目标过滤方法。实验结果表明，本文算法能够有效地降低检测漏警率和虚警率，较好地解决了视频监控中运动小目标难以检测的问题。

参考文献

- [1] CHEN Chenyi, LIU Mingyu, TUZEL O, *et al.* R-CNN for small object detection[C]. Proceedings of the 13th Asian Conference on Computer Vision, Taipei, China, 2016: 214–230.
- [2] 邓鹤, 魏艳涛, 童名文, 等. 基于改进的局部反熵算子的小目标检测[J]. 通信学报, 2013, 34(4): 60–69. doi: [10.3969/j.issn.1000-436x.2013.04.007](https://doi.org/10.3969/j.issn.1000-436x.2013.04.007).
DENG He, WEI Yantao, TONG Mingwen, *et al.* Small target detection based on modified local reverse entropy operator[J]. *Journal on Communications*, 2013, 34(4): 60–69. doi: [10.3969/j.issn.1000-436x.2013.04.007](https://doi.org/10.3969/j.issn.1000-436x.2013.04.007).
- [3] CAO Guimei, XIE Xuemei, YANG Wenzhe, *et al.* Feature-fused SSD: Fast detection for small objects[C]. Proceedings of SPIE 10615, Ninth International Conference on Graphic and Image Processing, Qingdao, China, 2017: 106151E.
- [4] REN Shaoqing, He Kaiming, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
- [5] LIN Tsungyi, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 936–944.
- [6] 杨依忠, 汪鹏飞, 胡雄楼, 等. 基于鲁棒主成分分析的运动目标检测优化算法[J]. 电子与信息学报, 2018, 40(6): 1309–1315. doi: [10.11999/JEIT170789](https://doi.org/10.11999/JEIT170789).
YANG Yizhong, WANG Pengfei, HU Xionglou, *et al.* Moving object detection optimization algorithm based on robust principal component analysis[J]. *Journal of Electronics & Information Technology*, 2018, 40(6): 1309–1315. doi: [10.11999/JEIT170789](https://doi.org/10.11999/JEIT170789).
- [7] ZIVKOVIC Z and VAN DER HEIJDEN F. Efficient adaptive density estimation per image pixel for the task of background subtraction[J]. *Pattern Recognition Letters*, 2006, 27(7): 773–780. doi: [10.1016/j.patrec.2005.11.005](https://doi.org/10.1016/j.patrec.2005.11.005).
- [8] VAN DROOGENBROECK M and PAQUOT O. Background subtraction: Experiments and improvements for ViBe[C]. Proceedings of 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Providence, USA, 2012: 32–37.
- [9] 薛阳, 张亚飞, 杨天宇, 等. 一种针对抖动视频序列的运动目标检测算法[J]. 激光与光电子学进展, 2018, 55(9): 091506.
XUE Yang, ZHANG Yafei, YANG Tianyu, *et al.* A moving object detection algorithm aiming at jitter video sequence[J]. *Laser & Optoelectronics Progress*, 2018, 55(9): 091506.
- [10] WREN C R, AZARBAYEJANI A, DARRELL T, *et al.* Pfinder: Real-time tracking of the human body[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 780–785. doi: [10.1109/34.598236](https://doi.org/10.1109/34.598236).
- [11] 赵春江. C#数字图像处理算法典型实例[M]. 北京: 人民邮电出版社, 2009: 29–31.
ZHAO Chunjiang. C# Digital Image Processing Algorithm for Typical Examples[M]. Beijing: Posts & Telecom Press, 2009: 29–31.
- [12] 李长春, 王艳杰, 马春艳, 等. 多目标无人机微型凝视高光谱成像仪辐射校正[J]. 测绘通报, 2019(1): 60–64. doi: [10.13474/j.cnki.11-2246.2019.0012](https://doi.org/10.13474/j.cnki.11-2246.2019.0012).
LI Changchun, WANG Yanjie, MA Chunyan, *et al.* Research on radiance correction of mini snapshot high spectrometer load on UAV using multi-target[J]. *Bulletin of Surveying and Mapping*, 2019(1): 60–64. doi: [10.13474/j.cnki.11-2246.2019.0012](https://doi.org/10.13474/j.cnki.11-2246.2019.0012).
- [13] THEODORIDIS S, 李晶皎, 王爱侠, 王骄, 等译. 模式识别[M]. 4版. 北京: 电子工业出版社, 2016: 282–288.
THEODORIDIS S, LI Jingjiao, WANG Aixia, WANG Jiao, *et al.* translation. Pattern Recognition[M]. 4th ed. Beijing: Publishing House of Electronics Industry, 2016: 282–288.
- [14] GONZALEZ R C, WOODS R E, 阮秋琦, 阮宇智, 等译. 数字图像处理[M]. 3版. 北京: 电子工业出版社, 2017: 423–430.
GONZALEZ R C, WOODS R E, RUAN Qiuqi, RUAN Yuzhi, *et al.* translation. Digital Image Processing[M]. 3rd ed. Beijing: Publishing House of Electronics Industry, 2017: 423–430.
- [15] 陈钱, 钱惟贤, 张闻文. 红外目标探测[M]. 北京: 电子工业出版社, 2016: 298–300.
CHEN Qian, QIAN Weixian, and ZHANG Wenwen. Infrared Target Detection[M]. Beijing: Publishing House of Electronics Industry, 2016: 298–300.
- [16] 姚天任, 孙洪. 现代数字信号处理[M]. 2版. 武汉: 华中科技大学出版社, 2018: 30–40.
YAO Tianren and SUN Hong. Advanced Digital Signal Processing[M]. 2nd ed. Wuhan: Huazhong University of Science & Technology Press, 2018: 30–40.

孙怡峰: 男, 1976年生, 副教授, 硕士生导师, 研究方向为图像处理、信息安全。

吴疆: 男, 1995年生, 硕士生, 研究方向为目标检测、图像处理。

黄严严: 女, 1978年生, 高级工程师, 研究方向为人工智能、信息安全。

汤光明: 1963年生, 教授, 博士生导师, 研究方向为网络安全、信息安全。