

# 基于光流的 ATM 机异常行为实时检测<sup>①</sup>

周 智, 张伟峰, 赵 斌, 黄 露, 朱 明

(中国科学技术大学 信息科学技术学院, 合肥 230022)

**摘 要:** 异常行为检测在自助银行智能监控系统领域中有广泛的应用前景. 本文针对此应用领域, 提出了基于区域光流特征的异常行为检测方法. 首先利用混合高斯模型来表示背景像素的变化并自适应更新背景模型, 用背景差法从视频序列中提取运动前景; 采用 lucas-kanade 光流法计算出运动区域内的光流信息. 采用基于幅值的加权方向直方图描述行为, 计算区域内直方图的运动熵发现候选异常区域, 再利用支持向量机进行分类. 从实验结果可以看出, 能够较好的识别出异常事件, 并且实时性较好, 能够满足实际应用需求.

**关键词:** 异常行为检测; 动作识别; 背景建模; 区域光流特征; 支持向量机

引用格式: 周智, 张伟峰, 赵斌, 黄露, 朱明. 基于光流的 ATM 机异常行为实时检测. 计算机系统应用, 2017, 26(9): 232-237. <http://www.c-s-a.org.cn/1003-3254/5929.html>

## Real-Time Detection of ATM Abnormal Events Based on Optical Flow

ZHOU Zhi, ZHANG Wei-Feng, ZHAO Bin, HUANG Lu, ZHU Ming

(School of Information Science and Technology, University of Science and Technology of China, Hefei 230022, china)

**Abstract:** Abnormal behavior detection has a wide application prospect in the field of self-service banking intelligent monitoring system. In this paper, an anomaly detection method based on regional optical flow feature is proposed. Firstly, the mixed Gaussian model is used to represent the change of the background pixels and the background model is updated. The motion foreground is extracted from the video sequence with the background difference method. The optical flow information in the moving region is calculated with the lucas-kanade optical flow method. The weight-oriented histogram is used to describe the behavior, and the motion anomaly region of the histogram is calculated by using the motion entropy of the histogram. Then the SVM is used to classify the anomaly regions. From the experimental results, it can be seen that the abnormal events can be identified better and the real-time performance is better, which can meet the practical application requirements.

**Key words:** anomaly detection; action recognition; background modeling; regional optical flow feature; support vector machine

随着 ATM 机日益广泛的使用, 其使用安全性也受到了广泛的关注, 近些年, 有许多 ATM 机抢劫, 打架等危害人身安全的事件发生. 传统的视频监控主要是人工视频监控, 随着视频数据的迅速增加, 这种方式存在着效率低, 耗费大量人力物力, 且易发生漏报和误报的情况. 为了更好的解决这类问题, 智能视频监控系统的研究变得尤为重要, 其中针对于人的异常行为的检测

识别是其中一项重要的研究方向.

近年来, 针对监控视频中的人体异常行为检测和识别领域, 已有一些方法识别某些特定场景下人体的行为, 主要有 2 类: ① 基于运动轨迹的分析<sup>[1-3]</sup>. 首先是跟踪运动目标, 得到运动目标的运动轨迹; 对目标的正常轨迹进行分析和建模; 最后对新轨迹的模型与原来的正常轨迹模型进行对比, 从而判断是否存在异常行

① 基金项目: 中科院先导项目课题(XDA06011203)

收稿时间: 2016-12-14; 采用时间: 2017-01-12

为. 此类方法要求能够准确的对运动目标进行提取和跟踪, 一旦出现遮挡或丢失情况, 则对检测结果有影响, 有较大的局限性. ② 基于人体特征的分析<sup>[4-7]</sup>. 此类方法关键在于提取人体的特征如人体轮廓结构、运动方向等. 这些特征能够代表人体的不同行为, 但是比较容易受噪声的影响, 特征不易提取.

针对于此, 本文提出了一种基于区域光流特征的异常行为检测方法. 该方法首先用混合高斯模型从视频中提取有效的运动目标, 并且进行运动区域标记; 然后运动LK光流法计算运动区域的光流, 采用基于幅值大小的光流方向直方图描述光流特征, 利用运动熵发现候选异常区域, 最后利用SVM进行分类.

## 1 运动目标检测

运动目标检测是智能视频分析技术的一个重要内容. 其目的是从监控视频中的每一帧中提取运动目标区域, 然后根据需要进行目标分类, 提取出所需要的运动人体目标. 这几乎是所有进行人体目标检测所需要的处理步骤, 其提取结果的好坏对后续处理有着极大的影响, 如目标跟踪、特征提取、行为检测等. 目前, 在运动目标检测的过程中, 主要有三种方法: 帧间差分法、光流法、背景减除法. 本文采取了背景减除法进行运动目标检测, 下面主要简述一下其基本原理.

背景减除法<sup>[8]</sup>是一种有效的运动对象检测算法, 基本思想是利用背景的参数模型来近似背景图像的像素值, 将当前帧与背景图像进行差分比较实现运动区域的检测. 其中区别较大的像素区域被认为是运动区域, 而区别较小的像素区域被认为是背景区域. 背景减除法必须要有背景图像, 并且背景图像必须是随着光照或者外部环境的变化而实时更新的, 因此背景减除法的关键是背景建模及其更新. 在背景建模方法中, 最著名的就是高斯背景建模方法, 使用高斯概率分布进行像素概率分布的描述, 是最常见的概率分布模型. 本文就是采用这种方法进行背景建模.

混合高斯背景建模法建立在单高斯模型方法的基础上, 也就是采用  $K$  个高斯模型进行混合, 通常  $K$  取 3~5. 在某时刻某一个点的像素值为  $X_t$  的概率密度为:

$$p(X_t) = \sum_{i=1}^K \omega_{i,t} p(x_{i,t}, \mu_{i,t-1}, \sigma_{i,t-1})$$

其中,  $\mu_{i,t-1}$  为  $t-1$  时刻第  $i$  个高斯模型的期望值,  $\sigma_{i,t-1}$  为  $t-1$  时刻第  $i$  个高斯模型的标准差,  $\omega_{i,t}$  是  $t$  时刻

第  $i$  个高斯模型的权重.  $p(x_{i,t}, \mu_{i,t-1}, \sigma_{i,t-1})$  为  $t-1$  时刻像素  $X_t$  的高斯概率密度函数, 具体表示如下:

$$p(x_{i,t}, \mu_{i,t-1}, \sigma_{i,t-1}) = \frac{1}{2\pi^{n/2} |\sigma_{i,t-1}|} \exp\left\{-\frac{1}{2} (x_{i,t} - \mu_{i,t-1})^T \Sigma_{i,t}^{-1} (x_{i,t} - \mu_{i,t-1})\right\}$$

混合高斯背景建模的具体实现过程分为两步: 第一步, 背景训练: 首先对视频序列进行一段时间的训练, 训练时间过短会影响  $K$  个高斯模型的建模效果, 但是训练时间过长会有对整个系统的启动时间造成拖延, 通常取 30~70 帧的时间进行训练. 初始背景通常都是相对静止的, 故该高斯模型的初始权值较大, 而其他高斯模型的初始权值较小, 方差取值较大. 第二步, 模型匹配: 采用第一步训练得到的  $K$  个高斯模型, 通常使用模板匹配的方法来对前景和背景做出区分. 对于图像的几个像素点, 利用公式:

$$|I(x, y) - \mu(x, y)| < W * \sigma(x, y)$$

进行匹配, 如果小于阈值则判断为背景点, 否则为前景运动目标. 如果某像素点始终是静态的背景, 该点就会与某一高斯模型匹配, 其方差的波动范围就会比较小; 但是如果某像素点是运动目标, 其与高斯模型匹配不能完全匹配, 那么其方差会有较大的波动.

现实的监控场景中, 外部环境时刻变化着, 所以要想保持背景与实际所处的环境保持一致, 必须要对背景模型进行自适应实时更新. 更新率  $\alpha$  是用来决定背景模型更新速率的, 如果  $\alpha > 1/2$ , 那么当前帧对新背景决策较大, 背景模型更新较快、变化较大; 如果  $\alpha < 1/2$ , 则是前一时刻的背景对新背景的贡献较大, 背景模型更新较慢、变化不大.  $\alpha$  不能去太大, 否则会增大将当前帧的运动目标判定为背景的模型, 进而影响运动目标提取的准确性, 所以对于更新率取值的选择要非常谨慎. 本文经过实验比对, 选择  $\alpha=0.001$ , 保证了前景目标提取的准确性.

在提取出前景后, 如图 1 所示, 使用形态学处理, 可以得到比较完整的运动前景区域, 从而可以得到感兴趣的人体运动区域.

## 2 特征选择及分类器

### 2.1 区域光流的计算

光流是很好的运动特征, 主要有 3 个部分构成: 水平偏移量  $x$ , 垂直偏移量  $y$  和像素的速率  $val$ , 分别表示了像素的方向和速度. 异常行为的特征可以用光流

特性较好地表示. 人的异常行为只发生在运动目标区域, 所以只需要计算对前景提取的区域计算光流即可.

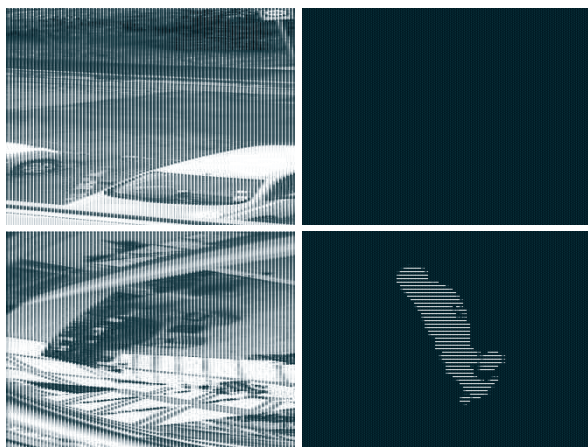


图1 高斯建模法前景提取

采用 Lucas-Kanade 算法计算光流, 即令  $m \times m$  特征窗口的光流为  $(\mu, \nu)$ , 满足光流约束条件  $I_x\mu + I_y\nu + I_t = 0$ , 可得:

$$\begin{bmatrix} I_{x1} & I_{y1} \\ I_{x2} & I_{y2} \\ \vdots & \vdots \\ I_{xn} & I_{yn} \end{bmatrix} \begin{bmatrix} \mu \\ \nu \end{bmatrix} = - \begin{bmatrix} I_{t1} \\ I_{t2} \\ \vdots \\ I_{tn} \end{bmatrix}$$

式中,  $n$  为特征窗口内的像素个数 ( $n = m^2$ ),  $I_x$  和  $I_y$  为图像的空间梯度,  $I_t$  为图像的时间梯度. 通过求解上述方程可得:

$$\begin{bmatrix} \mu \\ \nu \end{bmatrix} = \begin{bmatrix} \sum I_{x_i}^2 & \sum I_{x_i}I_{y_i} \\ \sum I_{x_i}I_{y_i} & \sum I_{y_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum I_{x_i}I_{t_i} \\ -\sum I_{y_i}I_{t_i} \end{bmatrix}$$

区域光流计算的计算复杂度为  $O(\eta N/m^2)$ , 其中  $\eta$  为前景占整个图像的比率,  $N$  为图像像素点的总个数,  $m^2$  为特征窗口大小. 在一般场景下, 前景区域不会超过当前视频面积的 40%, 从而会大大减少运算量; 同时, 基于运动区域的光流计算可以有效的去除由环境产生的扰动, 可以采用阈值筛选法, 提取那些有用的光流特征.

## 2.2 光流特征直方图

运动区域中的每个像素的速度为:

$$V_{i,j}(n) = \sqrt{\Delta X_{i,j}^2 + \Delta Y_{i,j}^2}$$

式中,  $\Delta X_{i,j}$ 、 $\Delta Y_{i,j}$  分别为像素  $pix(i, j)$  在水平和垂直方向的偏移量, 像素点的运动方向与水平方向的夹角为:

$$Angle_{i,j}(n) = \arctan \frac{\Delta Y_{i,j}}{\Delta X_{i,j}}$$

在每一帧图片中, 运动区域的光流特征都可以表示为:  $E = \{F_1, F_2, \dots, F_n\}$ ,  $n$  为图片中前景区域的个数. 每个前景区域的特征  $F_i = (f_{i1}, f_{i2}, \dots, f_{ij}, \dots, f_{ij})$ ,  $f_{ij} = \{X_{ij}, Y_{ij}, \theta_{ij}, V_{ij}\}$ , 其中  $(X_{ij}, Y_{ij})$  代表光流特征像素点的位置,  $J$  表示运动前景区域特征点的个数,  $\theta_{ij}$ 、 $V_{ij}$  分别为对应的像素点的光流特征的方向和速度. 从截取的部分视频序列中感兴趣区域的光流特征的速度及方向的分布图如下所示, 可以看出在序列中正常的行为的光流特征相对来说分布较为均匀, 光流的方向相似; 而异常的行为的光流特征相对来说分布较为杂乱, 光流的方向不一致.

直方图就是对数据进行统计, 将统计值组织到一系列事先定义好的 bin 中. bin 中的数值是从数据中计算出的特征的统计量, 这些数据可以是速度、方向、色彩或其他任何特征.

在本文中, 我们对方向直方图做一些改进, 一般的方向直方图是将方向范围划分为不同的角度区间, 然后将计算的方向特征归类到不同的区间上, 虽然方向直方图能比较好地体现运动区域内光流矢量的方向变化, 但是只考虑方向的变化, 却忽视了幅值大小所带来的影响, 幅值大小在客观上体现了人体动作行为的剧烈程度. 因此, 我们考虑采用基于幅值权值的方向直方图, 幅值大的光流矢量的权重也增大. 首先对光流矢量的幅值进行归一化处理, 将其作为光流矢量的方向的权值; 接着累计求和做方向直方图的大小, 假设运动区域  $R$  的方向直方图为  $OH(R) = \{oh_j(R)\}_{j=1,2,\dots,n}$ ,  $n$  是方向直方图区间的个数, 在本文中, 区间按照  $30^\circ$  划分,  $n=12$ .

$$oh_j(R) = \lambda_{uni} \sum_{i=1}^S \omega_{F_i} \delta(q(F_i) - j)$$

上式中,  $S$  是运动区域  $R$  内的光流特征的个数,  $\omega_{F_i}$  为第  $i$  个光流特征方向的权值,  $q(F_i)$  表示光流特征所对应的直方图的区间,  $j$  是方向区间,  $\delta$  为克罗内克函数,  $\lambda_{uni}$  为归一化参数. 从上式中看出,  $\omega_{F_i}$  反映了光流特征在运动方向上的幅值大小, 与一般的方向直方图相比, 加入幅值的影响因素, 更能真实地体现运动的强度和分布情况, 降低了噪声对特征提取的影响, 在图 2 中我们可以看出正常行为和异常行为的光流直方图分布很不一样, 正常行为的光流直方图分布较为集中, 而异常行为的直方图分布比较分散. 通过基于幅值大小



的光流方向直方图统计, 我们针对运动区域的直方图进行熵的计算:

$$E_H = - \sum_{j=1}^m oh_j(R) \log_2(oh_j(R))$$

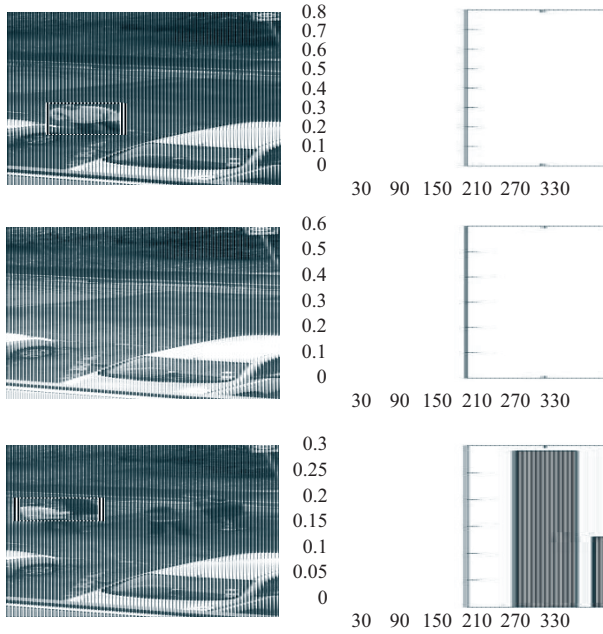


图2 正常行为和异常行为的光流直方图比较

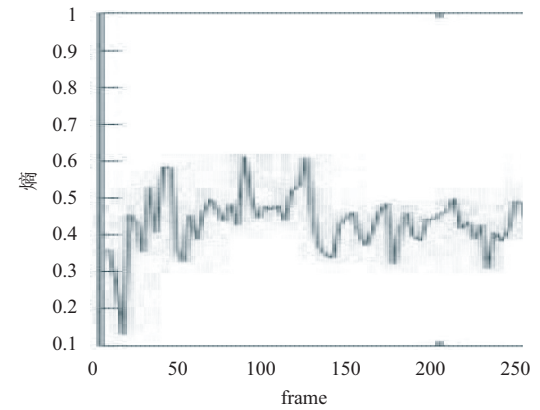
$E_H$  的值越大, 则运动区域内的光流向量方向分布比较混乱, 方向不一致, 可以反映出区域内的人体运动行为杂乱无章, 有很大可能是发生了异常行为. 将  $E_H$  归一化后, 取值范围为 (0, 1). 假设图片有  $K$  个运动区域, 则当前图片最大的熵  $E_{MAX} = \max(E_H^1, E_H^2, \dots, E_H^K)$ , 当  $E_{MAX} > T$  的时候, 我们判定有可能发生了异常行为, 标定为候选异常行为区域, 如图3所示, 阈值  $T$  本文中取 0.55.

### 2.3 SVM 分类器

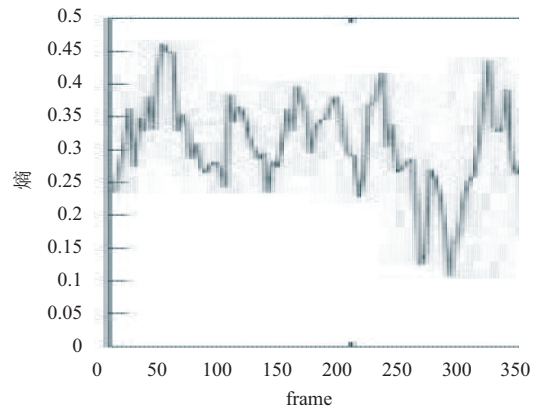
对于异常行为来说, 高维的特征很难通过简单的模板匹配或者欧式距离来进行分类. 为了解决这个问题, 这里我们采用 SVM<sup>[9]</sup> 做为分类器. 通常, 在样本数量不多的情况下, SVM 只需要较少的支持向量个数就能决定最终结果, 是一种很好的分类器, 具有较好的鲁棒性. 但是标准的 SVM 对于小样本数据容易造成过拟合的现象, 在数据样本多的时候会出现运行速度很慢, 泛化能力低等缺点. 因此出现了许多双目标函数的支持向量机, 如 TWSVM<sup>[10]</sup>, 基本思想是找出两对非平行的分类超平面, 本文采用了 TWSVM 分类方法.

对于非线性分类器, 定义如下:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n y_i a_i^* K(x_i \times x) + b^* \right\}$$



(a) 异常行为熵值变化



(b) 正常行为熵值变化

图3 正常行为和异常行为的熵变化曲线图

这里,  $x \in R^N$  是一个特征向量,  $n$  表示支持向量的个数,  $x_i$  是支持向量,  $y_i \in \{-1, 1\}$  表示分类标签 (-1 表示负样本, 1 表示正样本).  $a_i^*$  和  $b^*$  通过 SVC 学习算法得到.  $K(\cdot)$  表示核函数, 核函数的作用是将高维特征转换为低维特征, 从而降低计算复杂度. 本文中, 采取高斯径向基函数, 它适用于非线性分类问题.

$$K(x_i \cdot x) = \exp \left\{ - \frac{|x_i - x|^2}{2\sigma^2} \right\}$$

在本文中, 我们采用 2.2 节提出的区域光流直方图作为特征向量, 运动区域的直方图进行熵的计算进行异常行为的初筛之后, 选定了候选的运动区域, 为了考虑上下文信息, 异常行为的过程是连续的, 单帧图片可能不太适用于描述特征, 所以我们采用连续的 8 帧做为特征描述, 因此, 总共的特征有  $12 \times 8$  维特征.

### 3 实验结果与分析

本文方法的框架如图4所示,在我们的实验中,采用了三种数据集做为测试, BEHAVE 数据集、CAVIAR 数据集以及我们自己模拟的 ATM 数据集. 实验配置为 win10 系统, 4G 内存, Intel core i5 处理器.

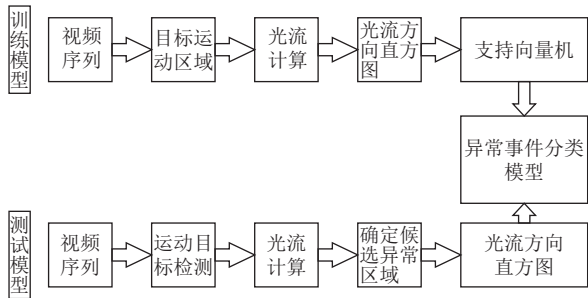


图4 本文算法流程图

BEHAVE 数据集: 我们采用 20 个片段做为训练样本. 同时, 我们与 HOG, HOF, HNF<sup>[11]</sup>, MoSIFT<sup>[12]</sup>和 ViF<sup>[13]</sup>算法进行了对比. 对于这些时空特征, 字典大小选择为 500. 表1可以看出在 60 个视频片段上的实验效果, 我们的方法在检测打斗事件中, 表现很好, 如图5所示. 在出现人体严重遮挡的情况下, 很难检测, 光流信息因为遮挡, 无法有效提取. 另外, 我们将所有人群行走、站立、讨论以及其他正常动作都视为一类, 即正常行为. 总的来说, 我们的方法能够很好的区分正常和异常事件, 在遮挡和人群之间做一些夸张动作的情况下, 会出现误报.

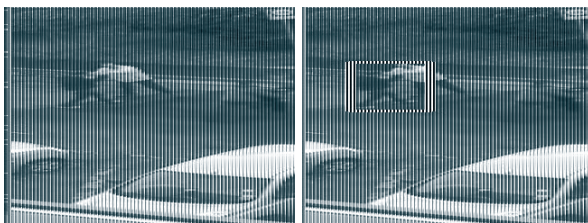


图5 BEHAVE 数据集异常检测结果

表1 BEHAVE 数据集的分类结果

Algorithms	Accuracy(±SD)(%)	AUC(%)
Ours	84.29±0.16	87.88
HOG	58.69±0.35	63.22
HOF	59.91±0.28	58.93
HNF	57.97±0.31	60.89
ViF	81.02±0.19	84.92
MoSIFT	62.02±0.23	65.78

我们可以看到, 在比较的六种方法中, 我们的方法

有最好的表现, 其次为 ViF、MoSIFT、HOG 和 HNF, HOF 的检测效果最差. 因为 BEHAVE 数据集包括许多行走, 奔跑及骑车等行为, 基于 MoSIFT、HOG、HNF 和 HOF 特征的分类方法很难区分.

CAVIAR 数据集: 我们选择了 4 个打斗片段和 19 个正常片段当做测试集, 我们的分类模型由 BEHAVE 数据集训练. 同样的, 我们与 HOG, HOF, HNF, MoSIFT 和 ViF 算法进行了对比. 表2可以看出在 23 个视频片段上的实验结果, 我们的方法能够较好的检测出异常事件, 如图6所示, 漏检的情况多发生在距离摄像头较远处, 光流信息提取偏少; 有些误检的情况是由于摄像头的抖动导致的光流计算错误. 另外, 尽管我们的训练集使用的是 BEHAVE 数据集, 但是可以看出, 训练出的模型同样能够有效的检测 CAVIAR 数据集中的行为.

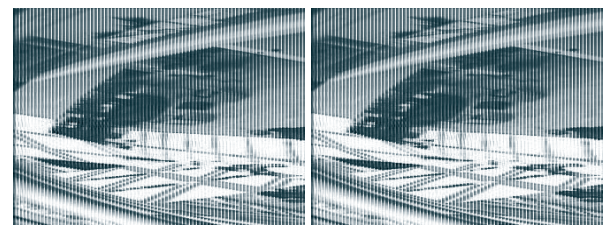


图6 CAVIAR 数据集异常检测结果

表2 CAVIAR 数据集的分类结果

Algorithms	Accuracy(±SD)(%)	AUC(%)
Ours	86.75±0.15	89.68
HOG	59.37±0.33	64.12
HOF	60.93±0.27	59.63
HNF	59.05±0.30	62.19
ViF	83.92±0.17	86.42
MoSIFT	64.03±0.22	67.39

模拟的 ATM 数据集: 模拟的 ATM 机取款监控视频, 摄像头高度为 3m 左右. 视频内包括: 单人取款, 单人徘徊; 双人排队取款, 双人靠拢交谈, 多人取款等正常行为; 双人打架, 抢劫; 多人打架, 抢劫等异常行为. 视频大小为 640\*480, 帧率为 30fps. 每段视频长度大约 100s. 为了验证在实际场景下算法的有效性, 实验选择了六种行为类型, 如图7所示, 本文的算法能够很好的检测出抢劫以及打架等异常行为, 漏检发生在人体严重遮挡的情况下; 另外, 在正常的单人及双人交谈情况下误报率较低, 出现误报的情况是因为我们模拟了动作幅度较大的正常行为. 在此, 我们选取了 ViF 作为

对比算法,从表3可以看出,我们的算法在误检率和检测率方面,表现的都要比ViF算法要好。

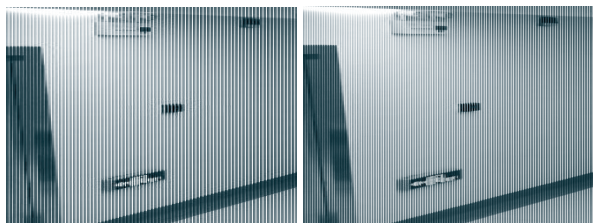


图7 模拟 ATM 数据集异常检测结果

表3 ATM 数据集的对比检测结果

视频类型	视频个数	本文算法		对比算法(ViF)	
		报警数	检测率(%)	报警数	检测率(%)
单人取款	15	1	93.3	2	86.7
双人交谈	10	1	90	2	80
抢劫	15	13	86.7	12	80
打架	15	14	93.3	13	86.7

另外,我们选取ViF算法,来进行实时性的对比,文中提及的其他时空特征算法无法保证实时性。表4可以看出对比的结果,我们的算法在实时性方面也要好于ViF算法。

表4 与ViF算法的实时性比较

Algorithms	Per frame(ms)
Ours	35
ViF	44

## 4 结语

本文提出了一种基于区域光流的异常事件实时检测方法,相对于基于时空特征的检测方法,本文提出的算法有着较高的准确率,同时有较好的实时性。不仅如此,我们还针对于具体的ATM机场景进行模拟实验,也有较好的表现。本文提出的方法目前也存在一些问题,对于人体遮挡及光线昏暗处,检测效果不是很好。

本文下一步将优化检测算法,提高人体遮挡及光线昏暗处的检测准确率,能更好地检测出异常事件。

## 参考文献

1 Bouttefroy PLM, Bouzerdoum A, Phung SL, *et al.* Abnormal behavior detection using a multi-modal stochastic learning

approach. International Conference on Intelligent Sensors, Sensor Networks and Information Processing. Sydney, Australia. 2008. 121-126.

2 Tehrani MA, Kleihorst R, Meijer P, *et al.* Abnormal motion detection in a real-time smart camera system. 3rd ACM/IEEE International Conference on Distributed Smart Cameras. Como, Italy. 2009. 1-7.

3 王晓龙. 基于轨迹分析的暴力行为识别算法研究[硕士学位论文]. 上海: 上海交通大学, 2015.

4 Xiang T, Gong SG. Incremental and adaptive abnormal behaviour detection. Computer Vision and Image Understanding, 2008, 111(1): 59-73. [doi: 10.1016/j.cviu.2007.06.004]

5 Fathi A, Mori G. Action recognition by learning mid-level motion features. IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, USA. 2008. 1-8.

6 Adam A, Rivlin E, Shimshoni I, *et al.* Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2008, 30(3): 555-560. [doi: 10.1109/TPAMI.2007.70825]

7 谢锦生, 郭立, 陈运必, 等. 基于时空惊奇计算的视频异常检测方法. 中国科学院研究生院学报, 2013, 30(1): 83-89.

8 Piccardi M. Background subtraction techniques: A review. 2014 IEEE International Conference on Systems, Man and Cybernetics. The Hague, Netherlands. 2004. 3099-3104.

9 Cortes C, Vapnik V. Support-vector networks. Machine Learning, 1995, 20(3): 273-297.

10 Jayadeva, Khemchandani R, Chandra S. Twin support vector machines for pattern classification. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(5): 905-910. [doi: 10.1109/TPAMI.2007.1068]

11 de Souza FDM, Chavez GC, do Valle Jr EA, *et al.* Violence detection in video using spatio-temporal features. 2010 23rd SIBGRAPI Conference on Graphics, Patterns and Images. Gramado, Brazil. 2010. 224-230.

12 Bertini M, Del Bimbo A, Seidenari L. Multi-scale and real-time non-parametric approach for anomaly detection and localization. Computer Vision and Image Understanding, 2012, 116(3): 320-329. [doi: 10.1016/j.cviu.2011.09.009]

13 Hassner T, Itcher Y, Kliper-Gross O. Violent flows: Real-time detection of violent crowd behavior. 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, RI, USA. 2012. 1-6.