

室内单目监控视频动态前景 提取方法研究

刘冰清

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2023年3月21日; 录用日期: 2023年5月16日; 发布日期: 2023年5月23日

摘要

针对室内单目监控视频光照不均等原因产生的阴影干扰和设备、环境因素产生的抖动问题, 本文提出了一种特征定位与改进帧差法融合的动态前景目标提取方法。首先, 基于像素级消减动态阴影特征并干扰使用高斯滤波进行降噪处理以减少监控视频冗余特征信息。其次, 使用特征匹配算法获取连续两帧图像的差分, 初步捕获动态前景目标。然后, 使用自适应阈值和形态学处理方法改进帧差法精确提取前景动态目标。最后, 实验验证方法的有效性和精准性, 本文室内单目监控视频动态前景目标提取算法准确度达到92.2%, 有效消除室内单目监控视频中的阴影和抖动干扰现象。

关键词

目标检测, 室内监控视频, 动态前景提取, 改进帧差法

Research on Dynamic Foreground Extraction Method for Indoor Monocular Surveillance Video

Bingqing Liu

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Mar. 21st, 2023; accepted: May 16th, 2023; published: May 23rd, 2023

Abstract

Aiming at the shadow interference caused by uneven illumination of indoor monocular surveil-

lance video and the jitter problem caused by equipment and environmental factors, this paper proposes a dynamic foreground target extraction method based on feature location and improved frame difference method. Firstly, dynamic shadow features are reduced at the pixel level and Gaussian filtering is used for noise reduction to reduce redundant feature information of surveillance video. Secondly, the feature matching algorithm is used to obtain the difference between two consecutive frames of images to initially capture the foreground dynamic target. Then, the adaptive threshold and morphological processing method are used to improve the frame difference method to accurately extract the foreground dynamic target. Finally, the effectiveness and accuracy of the method are verified by experiments. The accuracy of the dynamic foreground object extraction algorithm in indoor monocular surveillance video reaches 92.2%, which effectively eliminates the shadow and jitter interference in indoor monocular surveillance video.

Keywords

Target Detection, Indoor Surveillance Video, Dynamic Foreground Extraction, Improved Frame Difference Method

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

视频监控设备是保障社会公共安全和公民人身安全的重要工具，识别提取复杂监控视频有效信息成为计算机视觉学课中一个重要研究内容之一。提取室内监控视频人体动态目标有利于优化智能监测系统，前景提取步骤是目标识别、视频摘要和视频浓缩技术的第一步，对后续模块的构建和操作以及视频摘要算法的性能上产生极大影响[1]。复杂视频动态前景提取算法主要有三个方向：方向一是背景差值建模，陈树[2]等基于背景差值建模原理提出一种融合局部二值模式(LBSP)背景建模和 ViBe 算法的 LOBSTER 算法来提高背景模型对动态背景的适应能力。方向二是基于特征信息提取算法，Zhao [3]等通过基于 YOLOv4 目标框选算法初步识别视频动态目标，再使用自顶而下的 AlphaPose 姿态评估框架提取关节点信息作为人体目标行为分析的动态特征变化数据。方向三是深度学习算法，Chang [4]等使用 YOLOv3 识别动态目标后构建时空三维卷积神经网络(3D-CNN)提取视频动态前景目标。Hsu [5]等基于生成对抗网络(GAN)结构模型的动态目标提取方法，根据帧序列捕捉运动中时空变化属性来识别提取视频运动目标信息特征实现动态目标识别。

根据上述分析，本文提出一种融合特征匹配和自适应阈值改进帧差算法(Feature Location and Improved Frame Difference Fusion Algorithm, HFID)减少背景静态区域冗余特征和干扰源对提取视频前景目标特征的影响。主要贡献有：1) 融合像素级降噪方法和消除阴影方法进行预处理，解决背景差值建模方法识别结果出现鬼影问题、抖动噪声干扰问题和数据预处理耗时问题。2) 引入 SIFT 算法特征点匹配和视频帧间图像特征点定位，从时序特征点动态变化角度生成差分图像，解决基于骨骼点姿态识别框架因目标遮挡提取准确性低问题。3) 使用自适应阈值原理和图像形态学变化改进帧差法，解决基于深度学习的动态目标前景识别因训练数据量少导致过拟合在测试数据中识别运动缓慢目标运动误判为背景问题，并解决视频动态前景目标识别完整度低问题。本文研究对象范围仅聚焦在室内单目摄像机监控视频下动态前景目标提取方法研究，暂不考虑针对室外监控视频的动态场景。

2. 视频数据预处理

室内监控视频主要存在视频抖动模糊和光影变化两个主要的干扰源，本文使用 HSV 颜色空间消除阴影，再对视频图像进行灰度化和滤波降噪处理，以下是视频预处理的详细阐述。

2.1. 基于 HSV 颜色空间阴影消除方法

室内单目监控视频的阴影主要分为本影和运动阴影。本影是运动目标自身的阴影区域，而运动阴影[6]是前景目标运动时受室内光影影响在背景区域产生的投影。运动阴影在动态前景目标提取过程中被误判为动态目标，影响视频动态目标提取的精准度。运动阴影区域色调和饱和度与背景其他区域基本一致但亮度较低，本文使用 HSV 色彩空间分离像素亮度来区分视频动态目标本影和运动阴影干扰。室内监控视频中光源不稳定和光影变化导致 HSV 颜色空间判断鬼影方法在前背景分离阶段产生较大误差。从视频图像像素级进行分析，降低阴影消除算法对光源变化的敏感度，减小阴影消减操作对后续前景提取算法精确性的影响。

2.2. 降噪处理

从二维信号角度分析，抖动噪声[7]属于高频段信息，分析数据噪声直方图确定抖动噪声属于高斯噪声。本文使用高斯滤波算法处理抖动噪声，高斯滤波器根据像素点领域的像素加权均值代替该像素点的像素值。高斯函数[8]的傅里叶变化频谱是单瓣频谱，高斯滤波降噪是以目标像素点为中心建立窗口，计算其内部像素点均值并替换目标像素点的信号值，处理后的图像保留大部分图像信息且降噪过程不受高频信号的干扰。为提升数据处理速率将阴影消减和降噪处理进行同步处理，当 $RG(x,y)$ 结果为 1 时，像素点是阴影部分和抖动噪声干扰信号，使用加权平均的方法用邻域像素信息替代该像素值；当 $RG(x,y)$ 为 0 时，该中心像素点是前景目标自身阴影即本影部分。如公式(1)所示。

$$RG(x,y) = \begin{cases} 1, C_{\min}^n \leq t \leq P_{\max}^f \\ C_{\min}^n = \sum_i^n \frac{R_{HSV}(x,y) + G(x,y)}{R_{HSV}(x,y)G(x,y)E_i^n \bar{R}^n} \begin{bmatrix} x^n \\ y^n \end{bmatrix} \\ P_{\max}^f = \sum_i^f \frac{R_{HSV}(x,y) + G(x,y)E_i^f}{R_{HSV}(x,y)G(x,y)E_i^f \bar{R}^f} \begin{bmatrix} x^n \\ y^n \end{bmatrix} \\ G(x,y) = \sqrt{2\pi}\sigma e^{-2\pi^2\sigma^2 D^2(x,y)} \\ 0, \text{other} \end{cases} \quad (1)$$

其中 $R_{HSV}(x,y)$ 是单一像素点亮度分量值， $G(x,y)$ 是高斯滤波函数目标像素点信号值， E_i^n 是单帧图像像素值， \bar{R}^n 是在该单帧图像下像素点 (x,y) 的领域亮度集合， \bar{R}^f 是单帧图像全部亮度值集合。为进一步减少前景提取算法的运行时间和冗余特征信息，本文使用加权平均法灰度化处理每张单帧图像，灰度图像可以有效反映图像整体和局部区域的亮度等级及特征区别、增加目标区域对比度、提高特征提取和目标分割的准确性。

3. 融合特征匹配和自适应阈值改进帧差算法

3.1. 算法整体设计

本文基于时间维度监控视频动态目标的运动差异性提出 HFID 算法，图 1 是整体结构。

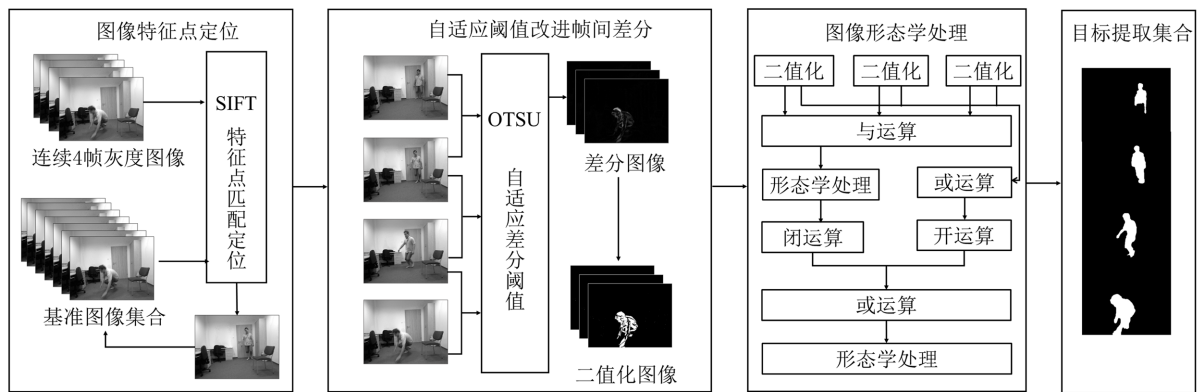


Figure 1. Overall structure diagram of HFID

图 1. HFID 整体结构图

3.2. 基于 SIFT 算法的目标特征定位

特征信息定位是动态目标提取算法的重要步骤之一。姿态估计算法提取关节数据成为提取动态目标特征信息的主流方法[9]，但室内监控视频易出现障碍物遮挡导致姿态估计准确性下降，因此基于姿态估计关节特征定位算法适用于实验场景。

轻量化的尺度不变特征转化(Scale Invariant Feature Transform, SIFT)特征匹配算法对抖动干扰不敏感，对图像旋转、尺度变化和光照变化有良好的抗干扰性[10]。本文引入 SIFT 算法对视频连续单帧图像进行特征匹配。特征匹配差异在目标运动区域明显，在背景区域特征信息匹配差值小，基于时间一致性分析全域特征信息匹配对齐结果后消除背景区域像素信息可以减少视频数据大面积静态背景区域，精简目标特征点信息能降低冗余特征信息对动态前景目标提取的干扰。依据 SIFT 特征点匹配结果对视频时间序列连续单帧图像动态目标特征点进行校正对齐实现精准差分。图 2 是基于 SIFT 算法连续单帧视频图像特征点匹配定位原理流程图：

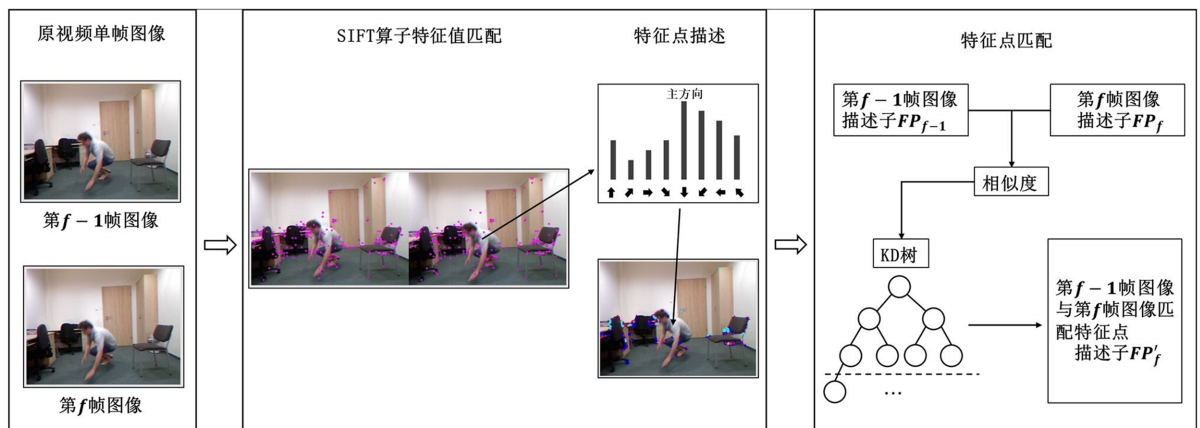


Figure 2. Flow chart of SIFT features location

图 2. SIFT 特征定位流程图

通过构建高斯尺度空间生成空间序列来确定视频单帧图像特征主导方向和描述符，可变参数的高斯函数和原始输入函数卷积得到该空间序列如公式(2)所示：

$$\begin{aligned}
L(x, y) &= G(x, y, \sigma) \times I(x, y) \\
\theta(x, y, \sigma) &= \tan^{-1} \frac{L(x, y+k, \sigma) - L(x, y-k, \sigma)}{L(x+k, y, \sigma) - L(x-k, y, \sigma)} \\
\psi_{k+1} &= \psi_k \left[\theta(x, y, \sigma); x^{(0)}, x^{(1)}, \dots, x^{(k)} \right], k=8 \\
m(x, y, \sigma) &= \sqrt{[L(x, y+k, \sigma) - L(x, y-k, \sigma)]^2 + [L(x+k, y, \sigma) - L(x-k, y, \sigma)]^2}
\end{aligned} \tag{2}$$

$I(x, y)$ 是像素点信息, σ 为标准差表示偏离距离均方差。迭代统计特征单元梯度方向 θ 作为主方向并计算其主峰值。特征点描述子之间的欧氏距离相似度作为特征点匹配的指标。在时间维度上图像背景存在较为细微的位移变化, 需要不断更新基准图像集合提高特征点匹配度以提升算法检测动态目标的实时性。假设基准图像样本 S_i^f 中有 P_1 特征点, 当前单帧图像中有 P_2 特征点, 二者的匹配点有 P 对, 则第 f 帧和第 $f+1$ 帧的特征点匹配度公式(3)如下:

$$R = \frac{\sqrt{\sum_{i=1}^n (x_i + y_i)^2} TN(x_i, y_i) P_1}{\max(P_1, P_2) FN(x, y)} \tag{3}$$

根据基准图像集合的不断更新进行自适应变化, 时序变化不断更新基准图像特征点匹配度使得静态背景区域在筛选离合后占总体特征值信息降低, 最终提升特征点匹配的灵活度和算法实用性的目标。根据特征匹配对采用仿射变换模型描述相邻单帧图像之间的几何变换如公式(4):

$$\begin{bmatrix} x^{f+1} \\ y^{f+1} \end{bmatrix} = \begin{bmatrix} a_1 S_x^f + b_1 S_y^f + c_1 \\ a_2 S_x^f + b_2 S_y^f + c_2 \end{bmatrix}, f \in q \tag{4}$$

将变换后的当前单帧图像更新到基准图像集合作为下一帧图像的基准图像样本 S_i^f , 根据基准图像得到当前单帧图像的变换矩阵后, 对当前单帧图像进行几何变换与基准图像对齐, 其结果作为改进帧差法的连续帧间输入图像。

3.3. 自适应阈值改进帧差法的动态目标提取

经典帧差法是一种轻量级视频动态目标提取算法, 通过比较图像间灰度值差值和固定阈值判别前景背景的方法[11]。帧差法对光线变化不敏感, 但单一固定阈值差分判断动态目标易导致提取结果出现“空洞”。本文使用动态变化信息自适应阈值算法 OTSU [12]代替经典帧差法固定阈值生成差分图像, 并引入图像形态学处理解决动态目标提取结果不完整问题。

假设图像存在灰度阈值 T 将像素集合 N 分为两部分且像素均值不同, 即灰度级在 $1-T$ 灰度区间内的图像域为像素区 A 类, 灰度级在 $T+1 \sim G-1$ 灰度区间的图像域是像素区 B 类, 分别如公式(5)计算灰度均值占总图像均值占比再对结果进行类间差分求得单帧图像区域分割差分阈值。公式(7)再对实验输入图像 A 、 B 区域进行类间差分。公式(8)基于差分均值拟合更新以实现视频帧目标动态特征像素变化自适应更新帧间差分阈值, 增加时序维度求类间差分最大阈值方法以实现随视频时间戳变化动态更新最大差分阈值更新帧间差分阈值得到最大差分阈值 T^* 。HFID 算法关键步骤伪代码如表 1 所示。

$$\begin{aligned}
m_0 &= \frac{1}{g_S} \sum_{i=0}^{g_S} iP_i, m'_0 = \frac{1}{g_E} \sum_{i=0}^{g_E} iP_i, m'_B = \frac{1}{P_{B'}} \sum_{i=T+1}^{g_E-1} iP_i \\
m_A &= \frac{1}{P_A} \sum_{i=0}^T iP_i, m_B = \frac{1}{P_B} \sum_{i=T+1}^{g_S-1} iP_i, m'_A = \frac{1}{P_{A'}} \sum_{i=0}^T iP_i
\end{aligned} \tag{5}$$

$$\begin{aligned}\sigma_{B'}^2(T) &= P_{A'}(m'_A - m'_0)^2 + P_{B'}(m'_B - m'_0)^2 \\ \sigma_{ESA'}^2(T) &= P_{A'}(m'_A - m'_0)^2 + P_B(m_B - m_0)^2\end{aligned}\quad (6)$$

$$\begin{aligned}\sigma_{ESB'}^2(T) &= P_A(m_A - m_0)^2 + P_{B'}(m'_B - m'_0)^2 \\ \{T_i, i \in f\} &= \arg \frac{\sigma_{B'}^2(T) \cdot \sigma_{ESA'}^2(T) \cdot \sigma_{ESB'}^2(T)}{\sigma_{B'}^2(T) + \sigma_{ESA'}^2(T) + \sigma_{ESB'}^2(T)}\end{aligned}\quad (7)$$

$$\sigma_{OTSU}^2(T^*) = \max\{\sigma_{B'}^2(T), \sigma_{ESA'}^2(T), \sigma_{ESB'}^2(T)\} \quad (8)$$

Table 1. Key steps of HFID algorithm pseudo code**表 1.** HFID 算法关键步骤伪代码

$HFID(S_i^f, E_i^n, R, F)$

输入: S_i^f 是基准图像, E_i^n 是视频连续四帧灰度图像, R 是基于 SIFT 算法特征匹配像素点占比, F 是监控视频图像总帧数

输出: 监控视频动态目标区域 D_f

```

1 for  $i=1$  to  $F$  do
2    $A' \leftarrow E_i^n R$ 
3    $B' \leftarrow E_i^n (1-R)$ 
4    $T_{\max}^i \leftarrow \max(otsu(B', T_i), otsu(E_i^n S_i^f A', T_i), otsu(E_i^n S_i^f B', T_i))$ 
5    $d_i \leftarrow E_i^n - RE_{i-1}^n$ 
6   if  $d_i \geq T_{\max}^i$ 
7     then  $d'_i \leftarrow 1$ 
8    $Tre_i^n \leftarrow threshold(d'_i), c_i \leftarrow orFun(Tre_i^n)$ 
9    $D_i \leftarrow morphologyImage(c_i)$ 
10  else  $d'_i \leftarrow 0$ 
11   $D_f \leftarrow d'_i$ 
12 end for
13 return  $D_f$ 

```

4. 实验结果与分析

本文使用分辨率为 640 像素 \times 480 像素的 UR Fall Detection 数据集[13]作为研究数据集, 该数据集包含 70 个视频序列, 其中 30 组跌倒视频序列和 40 组日常生活活动视频序列。此外, 为验证本文提出的基于室内监控视频动态前景目标提取方法的鲁棒性, 本文视频 320 像素 \times 240 像素的 CDnet 2014 数据集[14]作为实验对比数据集, 该数据集包含走动、跑步等更多行为活动。实验的基本硬件配置: CPU 是 Intel(R) Core(TM) i5-1135G7 2.42 GHz; 内存是 16 G、Windows64 位的操作系统; 实验使用 Python 库。

4.1. 算法提取时间分析

本文选取复杂视频动态目标提取算法中基于背景差值统计学的 LOBSTER 算法[2]、基于姿态关节特征提取的 AlphaPose 算法[3]和基于深度学习的 3D-CNN 算法[4]与 HFID 算法比较在室内监控视频场景下动态目标提取总耗时(单位 ms), 结果如表 2 所示。

Table 2. Comparison table of overall running time
表 2. 总运行时间比较表

分辨率(px)	LOBSTER	AlphaPose	3D-CNN	HFID
640 × 480	0.97	1.83	1.92	0.87
320 × 240	0.65	1.56	1.85	0.79

从实验结果可以看出图像分辨率越大算法耗时越长。HFID 算法与背景差值建模 LOBSTER 算法总体耗时均低于 1 ms。动态目标姿态特征提取框架 AlphaPose 框架和 CNN 算法需要对输入数据进行预标注和正常非正常数据分类,提取动态前景目标运行总耗时均高于 1 ms。因此, HFID 提取室内监控视频动态目标的耗时和资源消耗更小。

4.2. 算法有效性分析

本文引入如公式(9)准确率、召回率和综合评价指标三个指标[15]对本文前景提取方法进行定量算法性能评估。准确率表示算法提取结果中真正前景像素占监控视频单帧图像总像素的比值,召回率又称查全率表示提取结果中前景提取区域像素占真正前景像素的比值,综合评价指标是准确性和召回率加权调和综合指标,是反映算法提取效果重要指标之一。

$$\begin{aligned}
 Precision &= \frac{TP}{TP + FP} \\
 Recall &= \frac{TP}{TP + FN} \\
 F_1 &= \frac{2 \times (Precision \times Recall)}{(Precision + Recall)}
 \end{aligned} \tag{9}$$

本文选取基于背景差值算法的动态前景提取目标算法经典帧差法[11]、基于背景建模的 LOBSTER 算法[2]、基于 AlphaPose 姿态特征提取框架的动态目标提取算法[3]、基于深度学习的 3D-CNN 动态前景目标算法[4]和 GAN 无监督学习方法[5]作为对比算法,与 HFID 在 UR Fall Detection [13]室内监控视频跌倒行为场景下进行性能对比实验,结果如表 3 所示。

Table 3. Performance comparison of dynamic foreground extraction algorithms
表 3. 动态前景提取算法性能对比

方法	准确率(%)	召回率(%)	综合评价指标(%)
经典帧差法	60.6	32.1	44.5
LOBSTER	78.4	66.2	71.8
AlphaPose	80.5	75.4	77.9
3D-CNN	90.9	85.7	88.2
GAN	91.3	88.5	89.9
HFID	92.2	89.7	90.9

从表 3 中看出,经典帧差法和 LOBSTER 算法受室内光影干扰和抖动噪声干扰导致算法准确率和召回率均低于 80%,HFID 的综合评价指标比经典帧差法和 LOBSTER 分别提高 46.5%和 19.1%,弥补经典帧差法在复杂场景下动态前景目标提取鬼影问题。AlphaPose 算法的动态目标提取算法准确率是 80.5%,但综合评价指标比 HFID 算法低 13.1%,HFID 算法比避免姿态评估算法过拟合问题。基于深度学习算法的

视频动态目标提取准确性和完整性较高但依旧存在过拟合问题, HFID 比 CNN 和 GAN 动态目标提取算法分别增加 2.7%和 1.1%。从上述分析结果验证 HFID 算法在室内单目视频场景下动态目标提取准确率达 92.2%, 优于类似原理的经典帧差法和 LOBSTER 算法, 算法耗时和综合性指标优于基于深度学习的动态前景目标提取算法, 且比 AlphaPose 框架更轻量化。因此, HFID 算法对室内单目监控视频动态提取具有有效性, 准确率和召回率相较于其他算法有所提升。

本文使用 UR Fall Detection 数据集[13]中 fall-01-cam0 组场景下算法在目标不同运动行为识别结果对比实验。实验选取第 3 帧、第 52 帧、第 90 帧和第 103 帧的动态前景目标, 第 3 帧视频图像动态目标进入走入单目摄像机视野画面, 第 52 帧视频动态目标在视野内走动, 第 90 帧视频图像动态目标出现躯干倾斜, 存在较多动态阴影干扰, 第 103 帧视频图像动态在视野内跌倒, 存在设备剧烈抖动干扰。本文选择 LOBSTER [2]、GMM 背景建模算法[16]、AlphaPose 姿态特征识别框架[3]、3D-CNN [4]和 GAN 深度学习算法[5]与 HFID 算法在以上数据场景下进行对比实验, 结果如图 3 所示。LOBSTER 算法的提取结果连通性差无法进行目标区域分析研究。Liu [16]等提出一种基于高斯混合模型(Gaussian Mixture Model, GMM)前景提取算法构建背景更新模型来降低部分动态背景区域被误判为前景运动目标的概率, GMM 算法仅对运动较快目标提取效果较好并且能有效划分视频动态区域和本体阴影区域, 但对动态阴影干扰敏感。AlphaPose 框架的动态前景目标提取算法抗干扰性强, 但受运动速度影响提取结果不完整。基于深度学习算法的前景提取算法对视频运动目标提取完整性较强, 受光影和抖动噪声影响导致提取结果准确性下降。实验验证 HFID 算法能有效避免动态阴影干扰, 增加图像形态学处理能有效提升 HFID 算法提取结果的连通性, 有利于对视频动态前景目标区域的分析研究。HFID 算法在正常人体运动场景和跌倒异常场景识别结果完整性和精确性高对目标速度变化不敏感, 有效提取室内单目监控视频场景下的动态前景目标。

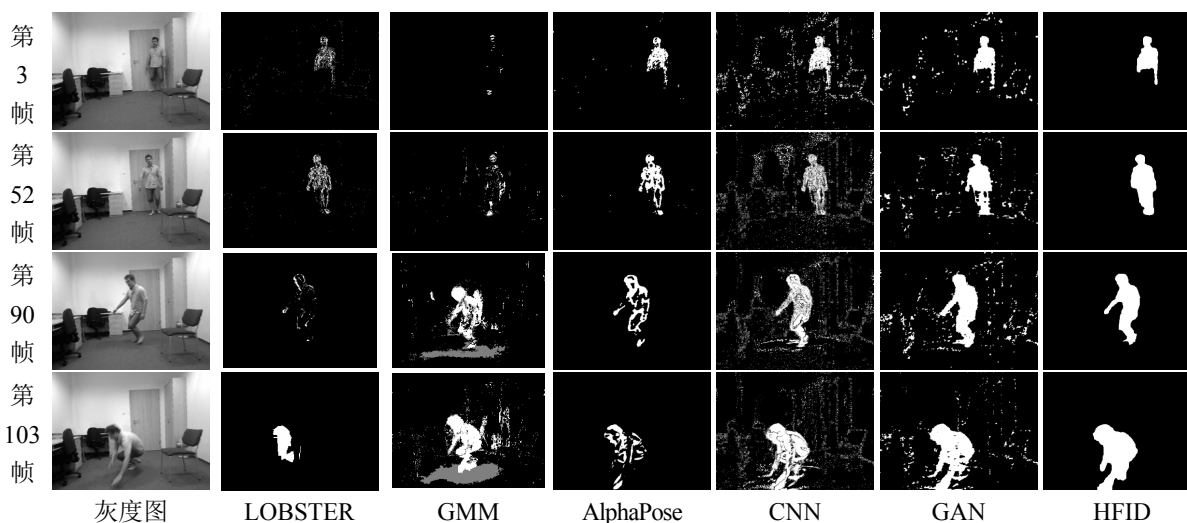


Figure 3. Comparison of target extraction results based on UR Fall Detection

图 3. 基于 UR Fall Detection 目标提取结果对比

4.3. 不同场景目标提取分析

本文使用不同室内单目监控视频场景作对比实验, 场景一是 CDnet 2014 通用数据集[14]的室内打印机监控视频(CopyMachine)场景, 该场景前景动态目标运动速率接近人体静止状态且存在大量光影灯光干扰, 视频分辨率为 720×480 。场景二是 CDnet 2014 数据集[14]的办公室监控视频(Office)场景, 该场景记录动态目标拿书动作, 是一组动静态组合场景数据即部分肢体运动迅速, 分辨率为 360×240 。场景三是 UR Fall

Detection [13]数据集中人体正常坐姿(Set)场景, 该场景干扰源较少且运动速率较缓慢但目标特征点较集中, 分辨率为 640×480 。场景四是 UR Fall Detection 数据集[13]中坐姿跌倒(Fall)场景, 该场景运动速率较快且存在大量抖动噪声和动态阴影干扰, 分辨率为 640×480 。实验选择 LOBSTER [2]、GMM 背景建模算法[16]、AlphaPose 姿态特征识别框架[3]、3D-CNN [4]和 GAN 深度学习算法[5]与 HFID 算法在以上数据场景下进行对比实验, 结果如图 4 所示。基于背景差值建模的 LOBSTER 算法和背景建模 GMM 算法易受抖动噪声干扰且对运动缓慢目标提取结果空洞。AlphaPose 姿态特征识别框架对抖动噪声和光影变化抗干扰力强, 但障碍物遮挡导致特征提取不足导致目标提取结果不完整。基于深度学习算法的 CNN 和 GAN 对光影变化敏感。HFID 算法能精准识别局部肢体运动目标, 对异常跌倒人体目标识别完整性高且对光影变化和抖动噪声抗干扰能力强, 相较于其他算法在不同场景下识别提取动态前景目标效果好。

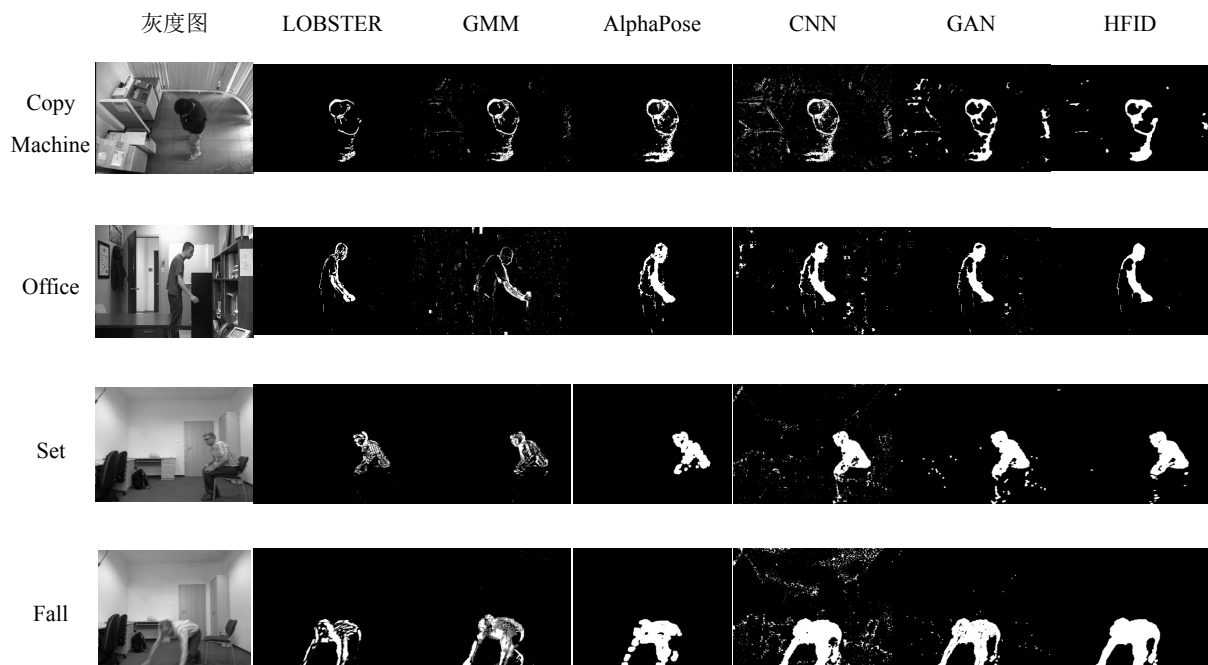


Figure 4. Comparison of target extraction results in different scenes
图 4. 不同场景下目标提取结果对比

5. 总结

本文提出一种融合特征匹配和自适应阈值改进帧差算法的动态前景目标提取方法, 相较于传统帧差法在一定程度上提升了室内单目监控视频中动态前景目标提取的准确性和抗干扰能力, 并解决因目标运动幅度较小前景被误判为背景区域问题。本文根据室内监控视频特点结合 HSV 和高斯滤波器实现视频数据预处理, 混合 SIFT 算法特征匹配定位和 OTSU 算法改进帧间差分法实现准确提取视频动态目标。实验表明 HFID 算法在室内单目监控视频下, 动态前景目标识别的准确度达 92.2%且算法运行时间较短具有实用性, 能够完整提取室内单目监控视频中的动态前景目标。

参考文献

- [1] Ghatak, S., Rup, S., Didwania, H., *et al.* (2021) GAN Based Efficient Foreground Extraction and HGWOSA Based Optimization for Video Synopsis Generation. *Digital Signal Processing*, **111**, Article ID: 102988. <https://doi.org/10.1016/j.dsp.2021.102988>
- [2] 陈树, 丁保阔. 动态背景下自适应 LOBSTER 算法的前景检测[J]. 中国图象图形学报, 2017, 22(2): 161-169.

-
- [3] Zhao, X.D., Hou, F.X., Su, J.F., *et al.* (2022) An Alphapose-Based Pedestrian Fall Detection Algorithm. *8th International Conference, ICAIS 2022*, Qinghai, 15-20 July 2022, 650-660. https://doi.org/10.1007/978-3-031-06794-5_52
- [4] Chang, C.W., Chang, C.Y. and Liu, Y.Y. (2022) A Hybrid CNN and LSTM-Based Deep Learning Model for Abnormal Behavior Detection. *Multimedia Tools and Applications*, **81**, 11825-11843. <https://doi.org/10.1007/s11042-021-11887-9>
- [5] Hsu, W.W., Guo, J.M., Chen, C.Y., *et al.* (2022) Fall Detection with the Spatial-Temporal Correlation Encoded by a Sequence-to-Sequence Denoised GAN. *Sensors*, **22**, Article No. 4194. <https://doi.org/10.3390/s22114194>
- [6] Sheng, M.X., Hou, W.W. and Jiang, J.C. (2020) Implementation of Accelerating Video Preprocessing Based on ZYNQ Platform Resource Management. *Journal of Physics Conference Series*, **1544**, Article ID: 012112. <https://doi.org/10.1088/1742-6596/1544/1/012112>
- [7] Han, P., Huang, T.Q., Weng, B., *et al.* (2021) Overcome the Brightness and Jitter Noises in Video Inter-Frame Tampering Detection. *Sensors*, **21**, Article No. 3953. <https://doi.org/10.3390/s21123953>
- [8] Arnal, J. and Sucar, L. (2020) Hybrid Filter Based on Fuzzy Techniques for Mixed Noise Reduction in Color Images. *Applied Sciences*, **10**, Article No. 243. <https://doi.org/10.3390/app10010243>
- [9] Pei, S.W., Wang, X.R., Qin, W., *et al.* (2021) STARS: Spatial Temporal Graph Convolution Network for Action Recognition System on FPGAs. *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, Madrid, 12-16 July 2021, 1469-1474. <https://doi.org/10.1109/COMPSAC51774.2021.00218>
- [10] Zheng, L., Yang, Y. and Tian, Q. (2018) SIFT Meets CNN: A Decade Survey of Instance Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 1224-1244. <https://doi.org/10.1109/TPAMI.2017.2709749>
- [11] 陈权, 黄俊, 徐访. 改进视觉背景提取算法在室内的研究[J]. 小型微型计算机系统, 2021, 42(6): 1250-1255.
- [12] Xing, J.W., Yang, P. and Qingge, L.T. (2020) Automatic Thresholding Using a Modified Valley Emphasis. *IET Image Processing*, **14**, 536-544. <https://doi.org/10.1049/iet-ipr.2019.0176>
- [13] Kwolek, B. and Kepski, M. (2014) Human Fall Detection on Embedded Platform Using Depth Maps and Wireless Accelerometer. *Computer Methods and Programs in Biomedicine*, **117**, 489-501. <https://doi.org/10.1016/j.cmpb.2014.09.005>
- [14] Wang, Y., Jodoin, P., Porikli, F., *et al.* (2014) CDnet 2014: An Expanded Change Detection Benchmark Dataset. *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2014)*, Columbus, 23-28 June 2014, 393-400. <https://doi.org/10.1109/CVPRW.2014.126>
- [15] Yang, H.H. and Qu, S.R. (2018) Real-Time Vehicle Detection and Counting in Complex Traffic Scenes Using Background Subtraction Model with Low-Rank Decomposition. *IET Intelligent Transport Systems*, **12**, 75-85. <https://doi.org/10.1049/iet-its.2017.0047>
- [16] Liu, J., Gu, Q.Y., Cheng, D.P., *et al.* (2022) VSLAM Method Based on Object Detection in Dynamic Environments. *Frontiers in Neurorobotics*, **16**, Article ID: 990453. <https://doi.org/10.3389/fnbot.2022.990453>