

DOI: 10.19650/j.cnki.cjsi.J2109059

基于行人组运动信息表达的人群异常检测^{*}

周生运, 张旭光, 方银锋

(杭州电子科技大学通信工程学院 杭州 310018)

摘要: 人群异常检测是智能化人群监控技术下的一个重要研究方向,在现有的方法中,异常检测的首要步骤就是获取运动信息,传统通过对视频帧均匀分块的方式并不能保证行人的完整性,提取的特征也不能准确反映行人的运动状态。本文提出了递进式人群分组的方法,先将人群运动场与帧差法结合分割图像获取人群前景,然后依据人群运动方向获取方向组,结合时空信息对方向组再聚类,得到更细致的行人组。对于每个行人组,利用人群能量特征去表征行人整体运动信息,并依据能量场构造了环块能量直方图特征来削弱行人四肢摆动的影响,最后与图像外观特征相结合用于人群异常检测。实验结果表明,本文方法在两个不同场景下帧级准确率达到83%和92%,像素级准确率达到64%和83%,与传统方法相比有较大提升。

关键词: 人群监控;人群异常检测;运动信息;行人组;环块能量直方图

中图分类号: TP391.41 TN948.6 TH701 **文献标识码:** A **国家标准学科分类代码:** 510.40

Crowd abnormal detection based on pedestrian group movement information expression

Zhou Shengyun, Zhang Xuguang, Fang Yinfeng

(School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: The crowd anomaly detection is an important research direction under intelligent crowd surveillance technology. In the existing methods, the first step of anomaly detection is to obtain motion information. The traditional way of uniform chunking of video frames does not guarantee the integrity of pedestrians, and the extracted features do not accurately reflect the motion status of pedestrians. In this article, an incremental crowd grouping method is proposed, which firstly combines crowd motion field with frame difference method to segment the image to obtain crowd foreground. Then, the direction group is achieved, which is based on crowd motion direction. Finally, Spatio-temporal information is combined to re-cluster the direction group to get a more detailed pedestrian group. For each pedestrian group, the crowd energy feature is used to characterize the overall pedestrian motion information, and the ring block energy histogram feature is constructed based on the energy field to weaken the effect of pedestrian limb swing, and finally combined with the image appearance features for crowd anomaly detection. Experimental results show that the proposed method achieves 83% and 92% accuracy at frame level and 64% and 83% accuracy at the pixel level in two different scenes, which is a significant improvement compared to the traditional method.

Keywords: crowd surveillance; crowd anomaly detection; motion information; pedestrian group; ring block energy histogram

0 引言

随着公共场所中的人群安全问题受到越来越多的重

视,人群异常检测研究在视频处理、计算机模式识别等^[1]领域已成为热点课题。从目前的研究来看,人群异常检测主要可以分为局部异常检测和全局异常检测^[2]。全局异常是将成场景中的人群看成一个整体并分析是否出现异常,无需关注场景中单个人的运动情况,对于行人之间

的遮挡也基本可以忽略。局部异常相对于全局异常更为常见,主要分析小部分个体与整体行为不一致的情况,在检测难度上,与全局异常相比,局部异常因为更容易受到其他因素干扰,检测起来要更加困难。

目前人群异常行为检测的方法可以分为两大类^[1],分别是基于视觉特征提取构建视觉模型的方法和基于物理特征构建物理模型的方法。第 1 种方法主要是利用计算机视觉领域相关技术提取特征构建模型,再利用分类器进行异常的检测^[3]。在基于视觉特征的方法中常用的两个模型就是隐马尔科夫模型和贝叶斯模型^[4],随着相关技术的发展,近年来提出的方法都对这两个模型进行了改进。如 Zhang 等^[5]在隐马尔科夫模型基础上提出了隐条件随机场模型对人群进行异常检测,相比于隐马尔科夫模型取得了更好的效果。Loy 等^[6]则提出了动态贝叶斯模型,通过两个级联的贝叶斯模型对人群行为进行建模,这种级联的结构具有允许大量不同模型的优势,每个模型对特定类型的异常都具有更高的敏感性。还有一些方法则是利用依据视觉技术构建出的光流用于人群异常检测,如 Sabih 等^[7]将 Lucas Kanade 光流法得到的光流特征用于表达帧间空间信息,并将其与表达时间信息的特征相结合,经分类器训练后用于人群异常检测,取得了较好的效果。Kotkar 等^[8]则是采用光流法和条纹流构建模型,对兴趣点进行检测和跟踪,最后采用高斯混合模型对兴趣点进行分类来识别人群中的异常事件。Lalit 等^[9]则是把光流和纹理特征结合用于人群异常检测。Roshtkhari 等^[10]则把视频序列作为时空立方体集合,运用概率统计学思想,以时空结构通过概率框架构建词包模型来进行异常检测。另一类方法是基于物理特征构建物理模型的方法。最具有代表的便是社会力模型(social force model, SFM),Mehran 等^[11]提出了一种基于社会力模型异常检测结构,通过社会力模型来估计像素粒子间的相互作用力,以此来确定异常位置。Chaker 等^[12]采用窗口的方法,从时间和空间上对视频场景进行划分构造出时空长方体,并使用局部社交网络进行建模,从检测结果上看取得了较好的效果。流场模型(flow field model, FFM)在异常检测方面也被广泛应用。如汪双承等^[13]采用的粒子流来描述人群的运动。钟志等^[14]则是把物理学中的能量概念引入到了人群异常检测中。还有些方法通过提取特征点,如角点、交叉点、边缘点等用于描述人群运动^[15]。

这两种方法的重点都在于特征提取,为了提取特征通常情况下都会先获取图像内的有效区域,大都是将连续图像帧平均的分成等大小、形状相同的区域块,然后提取每个区域块的特征如目标速度^[16]、颜色直方图^[17]或者深度特征^[18]等来建模进行异常检测。这种方式并没有考虑人群在场景中的分布情况,会把整体

的行人分割开来,不能保证行人个体的完整性,容易导致提取的特征不能准确反映行人的运动状态,影响检测结果的有效性。

为了解决上述这种问题,提出了基于行人组运动信息表达的人群异常检测。在实际场景中,对行人分组时要考虑到行人空间位置、速度以及运动方向等多种因素,而多种因素对运动状态的权衡参数又很难把控,在行人聚集的地方很容易造成产生错误分组,故本文提出了递进式的解决策略。先对人群的运动方向进行分组,然后对每个方向组人群根据其空间信息中的空间位置以及时间信息中的运动速度进行细化分组,通过这种方式可以很好的减少在行人之间距离过近导致的错误分组。对行人分组后,采用行人组能量特征来表征行人组整体的运动信息,并利用人群能量场构建环块能量直方图特征来削弱行人四肢摆动对整体运动的影响,最后将这两种运动特征与通过灰度共生矩阵提取出的外观特征结合用于异常检测。

1 算法描述

本文算法的详细流程如图 1 所示。先根据连续图像帧计算光流场,并引入分子扩散模型对其预处理,与帧差法结合分割图像获取运动人群,然后采用递进式策略对人群分组,提取出能量特征和环块能量直方图特征,最后将其与通过灰度共生矩阵提取的外观特征结合用于人群异常检测。

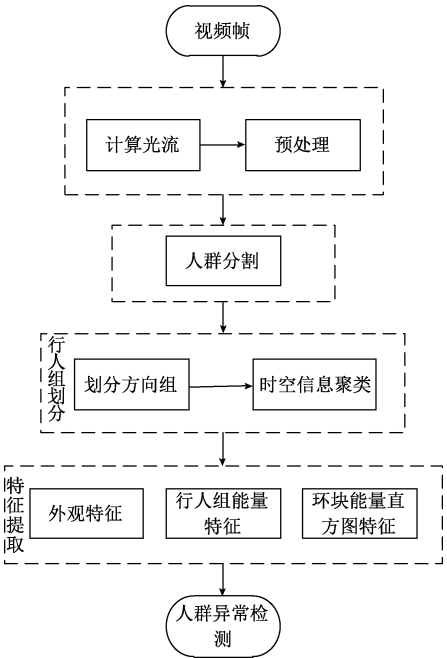


图 1 本文算法流程

Fig. 1 Algorithm flow of this article

2 运动场构建及人群分割

为了获取人群的运动信息,本文通过视频序列构建运动场来表征人群运动,以满足后续特征提取的需要。光流法是获取运动流场很实用的方法,如图2所示,本文采用的是 Horn-Schunck 光流法得到光流场,可以看出光流场在局部区域内并不平滑,光流方向和大小在局部内比较杂乱,会影响后续人群分组时的效果,所以本文引入分子扩散理论对原始光流场进行处理,然后再与帧差法相结合对图像分割,提取出运动目标。

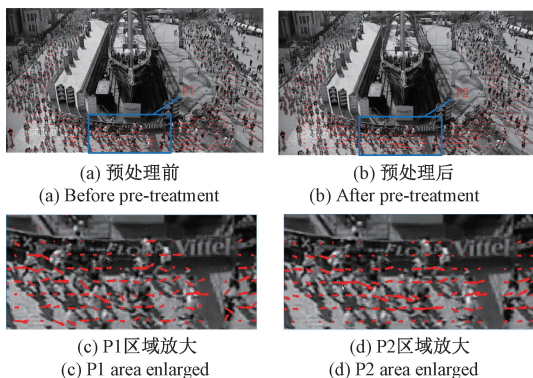


图2 马拉松视频帧光流场处理示例

Fig. 2 Marathon video frame optical flow field processing example

2.1 光流场预处理

受到流场分子扩散理论的启发,分子会从高浓度的地方向低浓度的地方进行扩散,扩散速度大小与浓度梯度成正比,是一种趋向于平衡态的熵驱动过程。将计算出的水平光流和垂直光流看做是浓度场,将每个像素点的光流速度看做是浓度。假设 $F_a = (u, v)$ 和 $F_b = (u, v)$ 是场中两点 a 和 b 的光流矢量,其中 u 和 v 分别表示矢量中的水平速度和垂直速度。则浓度场内两点间的扩散量可以表示为:

$$c_a(b) = \begin{cases} T_{a,b} \times e^{-k_1 \times \theta_{a,b} \times D_{a,b}}, & \cos(F_a, F_b) > 0 \\ 0, & \text{其他} \end{cases} \quad (1)$$

式中: $T_{a,b}$ 表示场内两点 a, b 的浓度梯度, $\theta_{a,b}$ 表示两点矢量的余弦值, $D_{a,b}$ 表示两点间的欧氏距离。这种处理的物理意义是,两点的距离越近,扩散量越大。两点光流矢量的夹角越小,两点在运动上更相似,扩散量越大。浓度场内扩散后一点 a 的浓度可以表示为:

$$C_a = c_a + \frac{1}{n_1} \sum_{b \in I, b \neq a} c_a(b) \quad (2)$$

其中, I 为场内 a 点的一块邻域; n_1 表示邻域内发生扩散的点数; c_a 表示扩散前 a 点浓度。将水平光流和垂直光流处理后重建的光流场 S 与处理前的光流场对比

如图2所示。图2(a)和(b)展示了预处理前和预处理后的光流场。图2(c)和(d)为光流场内 P1 区域和 P2 区域的放大图。从结果来看,处理后局部的光流变得更加有序,更符合行人整体的运动方向。

2.2 运动人群分割

为了判断场景中的异常行为,需要将运动人群与背景区域分割开。帧差法^[19]可以通过对比视频序列中的相邻帧来获取运动目标位置,当图像中不存在运动目标时,相邻两帧图像对应的像素点是基本相同的,当运动目标出现时,对应的像素点会发生较大变化。

如图3所示。图3(a)和(b)分别为视频帧图像和计算出的光流示意图 I_y 。图3(c)和(d)分别为帧差法计算出的差分图像 I_x 和人群分割的结果。差分图主要体现的是运动目标轮廓,容易缺失运动目标信息,而光流分布图 I_y 能很好的体现运动目标位置,对两个图像二值化后取其并集 $T_1 = I_x \cup I_y$, 经过形态学操作后,就可以提取到最终的人群分割图像。

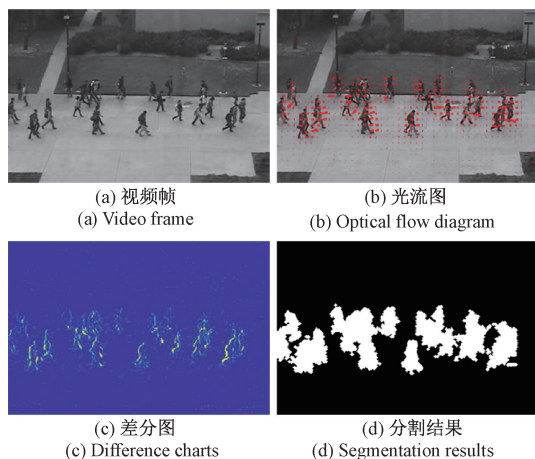


图3 运动人群分割

Fig. 3 Sports population segmentation

3 行人递进式分组

通过以上操作可以得到运动的人群对象,还需要对人群进行聚类分组,以提取特征进行异常检测。在人群场景中,行人的运动模式和其周围的行人表现出一致性,不同的一致性可以将人群划分为不同的小组。

行人的一致性主要反映在空间和时间上,空间信息主要表现在某一时刻行人的相互位置,时间信息主要表现在行人一段时间内的运动信息,如速度、方向。为了异常检测位置的准确性,希望能获得更精细的一致性组,采用了递进式解决策略,先根据行人的运动方向进行聚类,再对不同方向的人群结合其速度和空间位置进行聚类。通过这种方式,能很大程度上改善行人因空间位置过近导致的错误聚类问题。

3.1 方向组分割

在原始光流场处理后,个体行人内部光流趋于稳定,整体方向基本保持一致,可以通过对人群的主要走向对人群进行方向组的分割,计算方式如式(3)所示。

$$\cos\theta = \frac{\mathbf{F} \cdot \mathbf{R}_i}{|\mathbf{F}| \times |\mathbf{R}_i|}, i = 1, 2, 3, 4 \quad (3)$$

式中: $\mathbf{F}$ 为光流场 $\mathbf{S}$ 内一点的光流矢量, $\mathbf{R}$ 为场景道路延伸的方向向量, 根据与道路夹角的余弦值将人群分为 4 个主要运动方向。

方向组划分结果如图 4(b) 所示, 可以看出还存在小部分的噪声, 由于噪声的面积比较小, 对于与其他正常小组相连的噪声将其归入与正常组关联部分最长的组, 对于单独的噪声则直接舍弃掉, 则得到了最终的方向组划分结果, 如图 4(c) 所示。从处理后的结果来看, 对于不同方向的行人实现了较好的分割。

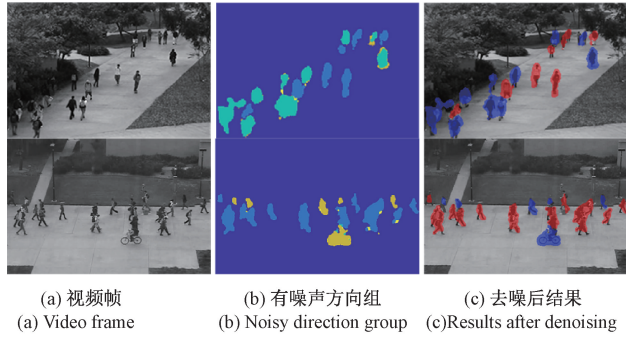


图 4 人群方向组划分

Fig. 4 People direction group division

3.2 时空聚类

通过以上的操作后,得到了人群的一个粗分组,再利用人群时间信息和空间信息可以对分组进行细化。行人的运动速度可以通过光流场 $\mathbf{S}$ 内光流大小反映,其空间位置反映了行人的实际分布。

本文同时依据光流场 $\mathbf{S}$ 内一点的光流速度与空间位置进行聚类,聚类每一步都基于均值漂移聚类算法。其中偏移均值的计算公式为:

$$m_{h,c}(X) = \frac{\sum_{i=1}^n X_i g\left(\left\|\frac{d(X, X_i)}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{d(X, X_i)}{h}\right\|^2\right)} - X \quad (4)$$

其中, $g(x) = -k(x)$, $k(x)$ 是核函数 $K(x)$ 的截面函数, h 是带宽。由于在实际场景中,行人在视觉上看起来更像上下宽左右窄的椭圆而非圆形,所以在偏移过程中,采用了椭圆的结构进行偏移。

对于场内一点的光流速度与空间位置,本文采用加权的欧几里得距离来评估粒子之间的相似度,计算公式如式(5)所示。

$$r(X, X_i) = \sqrt{\left(\frac{x - x_i}{a}\right)^2 - \left(\frac{y - y_i}{b}\right)^2} \quad (5)$$

$$d(X, X_i) = t_1 \times r_p(X, X_i) + t_2 \times r_q(X, X_i)$$

式中: a, b 分别为椭圆的短轴和长轴, r_p 表示两个点空间坐标的欧几里得距离, r_q 表示两点速度大小的欧几里得距离, t_1 和 t_2 为其权重。最终的聚类效果如图 5(b) 所示,可以看出对于场景中的行人实现了较为细致的分组。

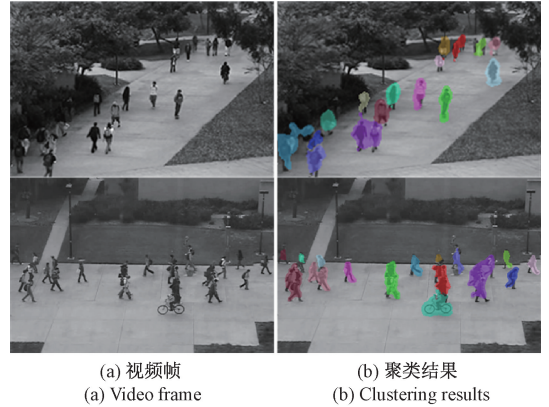


图 5 人群聚类结果

Fig. 5 Population clustering results

4 特征提取

4.1 外观特征提取

在场景中,异常区域并不是只有行人,也包含一些车辆、自行车等其他物体,这些物体和正常的行人在外观上有很大差异,纹理特征能很好地区分出图像中形状差别较大的物体。本文利用从灰度共生矩阵(gray-level co-occurrence matrix, GLCM)抽取出的几种纹理参数来描述人群小组的纹理特征。

对每个行人组构建灰度共生矩阵,矩阵内每个元素的计算公式如下:

$$P_{i,j} = \frac{M_{i,j}(d_1, \theta)}{n_2} \quad (6)$$

其中, $M_{i,j}(d, \theta)$ 表示像素点灰度分别为 i, j ; 距离为 d_1 ; 角度为 θ 的像素对个数; n_2 为像素对总数。通过概率 $P_{i,j}$ 可以提取出多种特征,考虑到实时性,本文选用了能量(energy)、熵(entropy)、对比度(contrast) 3 种特征值。

能量是用来度量图像分布均匀性和纹理粗细的特征,计算方式如下:

$$Energy = \sum_{i=1}^L \sum_{j=1}^L P_{i,j} \quad (7)$$

熵是基于信息理论的一种特征,反映了图像所包含的信息量。当图像的纹理越复杂,熵值越大,图像纹理越

简单,熵值越小。计算方式如下:

$$Entropy = - \sum_{i=1}^L \sum_{j=1}^L P_{i,j} \log_2 P_{i,j} \quad (8)$$

对比度可以反映图像的清晰度和纹理的沟纹深浅,纹理的沟纹越深,则其对比度越大,图像的视觉效果越清晰。计算过程如下:

$$Contrast = \sum_{i=1}^L \sum_{j=1}^L (i-j)^2 P_{i,j} \quad (9)$$

本文在构造共生矩阵中选用了4中不同的参数,其中距离 d_1 为1,角度 θ 分别选择 0° 、 45° 、 90° 、 135° ,并对求出的3种统计数据求平均后作为最终的纹理特征 $\bar{G}$ 。

4.2 运动特征提取

当异常事件发生时,行人的运动情况要比正常情况下强烈,通过能量特征可以很好的表征行人组整体运动情况,另外,由于行人四肢摆动很容易导致检测结果的误判,所以通过重心法依据能量场构建了环块能量直方图特征来削弱行人四肢摆动带来的影响。

1) 行人组能量特征提取

物理学中通常使用动能来表示运动物体的能量,结合动能的概念将光流场 S 内的点看成粒子,每个粒子的能量可以表示为:

$$E = \frac{1}{2}m(u^2 + v^2) \quad (10)$$

其中, u 表示水平光流, v 表示垂直光流,每个粒子的质量 m 都视为1。对场内所有点处理后得到能量场 S_1 ,如图6(b)所示。对于每个行人组,将组内所有点能量的均值 $\bar{E}$ 作为其能量特征。

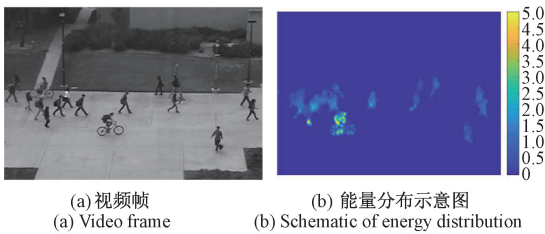


图6 能量场示意图

Fig. 6 Schematic diagram of energy field

2) 环块能量直方图特征提取

将每个小组的水平最小外接矩形提取出来,并以行人组重心向外以环形分块,取中值4作为分块的块数,并且为使环块之间面积差尽可能小,从内向外每个环块的宽度分别定为重心到矩形边长度的0.4、0.3、0.2、0.1,分块结果如图7(b)所示。并对每一块内的粒子能量做统计,依据训练集数据分析,将粒子能量大小分成8个范围并绘制直方图,并将直方图归一化。可以看出不同状态的行人组对应的直方图有很大差异。

由于粒子能量大的地方发生异常的概率比较大,故对8个不同范围的能量设置了不同的权重来获取每个环块的特征,由内到外第 b 个环块可以表示为:

$$H_b = \sum_{i=1}^8 \varepsilon_i \left(0.5 + \frac{N_i - O}{2 \times |N_i - O|} \right), b = 1, 2, 3, 4 \quad (11)$$

其中, ε_i 为直方图不同bin的权重; N_i 表示不同bin的频数; O 表示bin的阈值。则每个行人组的环块能量直方图特征 $\bar{H}$ 表示为:

$$\bar{H} = n_3 \times H_b, b = 1, 2, 3, 4 \quad (12)$$

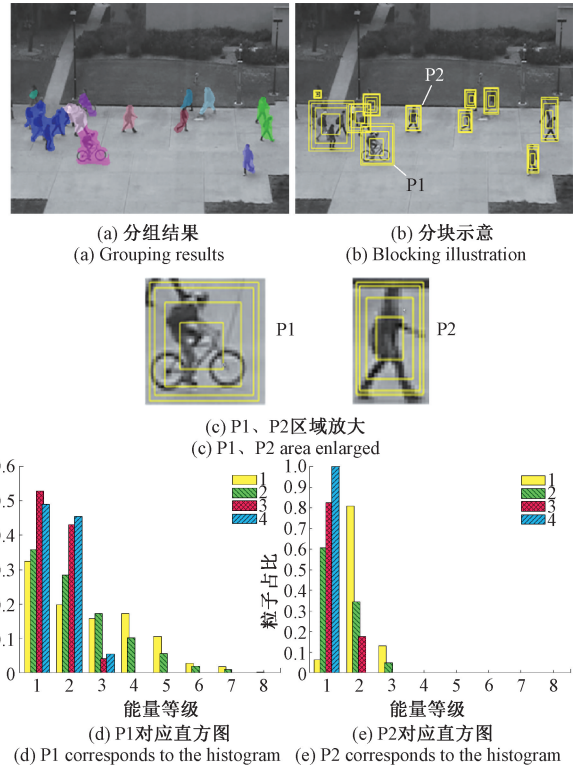


图7 环块能量直方图特征提取示意图

Fig. 7 Schematic diagram of ring block energy histogram feature extraction

对于分成的4个环块,越靠近重心位置的块应该有更的权重来减弱四肢摆动的影响,由重心由内向外的4个块的权重 n_3 分别设为0.45、0.25、0.2、0.1。

5 实验结果与分析

5.1 实验数据集及评估方式

本文采用了UCSD (University of California, San Diego) 数据集进行了实验,该数据集由美国加利福尼亚州统计学习视觉计算实验室提供并公开^[20],包含两个不同的场景Ped1和Ped2, Ped1包含34个训练视频序列和

36 个测试视频序列, Ped2 包含 16 个训练视频序列和 12 个测试视频序列, 分辨率分别为 238×158 , 360×240 。在人群异常检测过程中, 将外观特征和运动特征结合起来作为最终的特征描述符 $(\bar{G}, \bar{E}, \bar{H})$, 并采用一类 SVM (one-class support vector machine, OC-SVM) 检测人群中的异常运动, 并基于两种不同的评估标准将本文的方法与现有的局部人群异常检测方法进行了比较。

对于本文提出的方法, 使用受试者工作特征曲线 (receiver operating characteristic curve, ROC) 对其评估。ROC 曲线的横坐标表示假阳性率 (false positive rate, FPR), 纵坐标表示真阳性率 (true positive rate, TPR), 计算方式如下:

$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{FP + TN}$$

(13)

在得到 ROC 曲线后, 利用曲线下面积 (area under curve, AUC) 和等误差率 (equal-error rate, EER) 作为本方法整体的评价指标。

5.2 人群异常检测结果

在 UCSD 数据集中, 正常的事件是一些行人在缓步行走, 异常事件主要包括骑自行车、玩滑板、驾驶汽车以及使用轮椅者等, 这种事件在外观和运动信息上与正常事件都有较大差异。为了评价本文的方法, 实验采用了基于帧级别和像素级别两种不同的测量方式。帧级别是把数据集内每一帧作为检测单位, 只要该帧内存在异常像素点则该帧被视为异常帧, 无需精确检测出异常发生的具体位置, 通常用于评价全局异常。像素级别是把帧内的像素点作为检测单位, 只有该帧检测结果与数据集真值重合面积至少超过 40% 才认为该帧是异常帧, 通常用于评价局部异常。

图 8 和 9 为在 Ped1 和 Ped2 两个场景下的分组和异常检测结果, 其中 I 部分展示了 Ped1 和 Ped2 中运动人群的聚类分组结果, 可以看出人群分组结果比较细致。II 部分展示了 Ped1 和 Ped2 中的异常帧, 标记部分是检测出的异常位置, 从图里可以看出, 对于异常人群的定位比较准确, 即使帧内存在多个异常, 也能很好的检测出来。



图 8 UCSD Ped1 数据集中分组和异常检测结果
Fig. 8 Grouping and anomaly detection results in UCSD Ped1 dataset

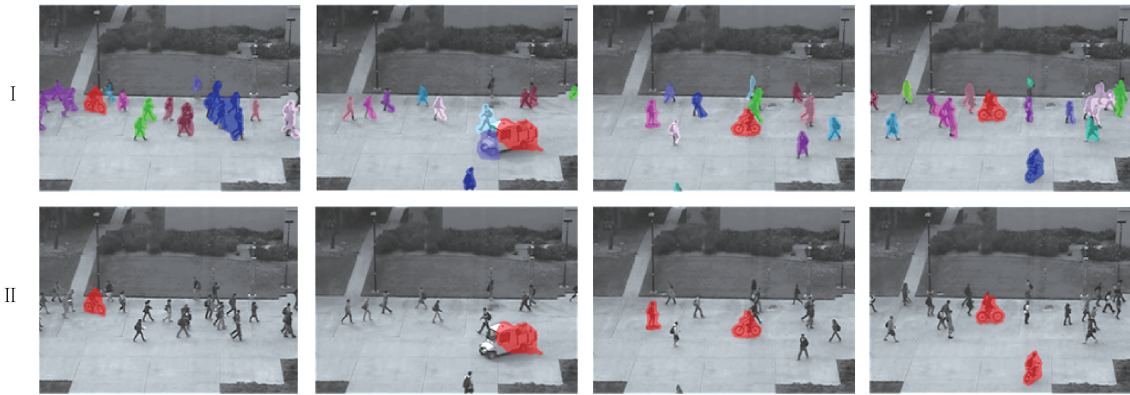


图 9 UCSD Ped2 数据集中分组和异常检测结果
Fig. 9 Grouping and anomaly detection results in UCSD Ped2 dataset

Ped1 和 Ped2 的帧级和像素级 ROC 曲线如图 10 和 11 所示,并将本文的方法与社会力模型 (social force, SF)^[11]、混合概率主成分分析 (mixture of probabilistic principal component analyzers, MPPCA)^[21]、动态纹理混合 (mixture of dynamic textures, MDT)^[22]、分层动态纹理混合 (hierarchy of MDT, H-MDT)^[23]、时空组合 (spatio-temporal composition, STC)^[24]、残差网络 (residual network, RESNET)^[25]、增量时空学习方法 (incremental spatio-temporal learner, ISTL)^[26]做了比较,如表 1 和 2 所示,本文方法与多个方法相比在两种测量级别上都有优势。从 AUC 值来看,在 Ped1 数据集中,相比于前 3 种方法在帧级上分别高出了 16%、24% 和 1%,在像素级上分别高出了 44%、44% 和 20%,并且与这 3 种方法相比对应

的 EER 值更低。相比于后两种方法,本文方法在帧级上也高出了 11% 和 8%。在与 H-MDT 和 STC 方法在 EER 值对比上略显不足,在 Ped2 数据集中,相比于前 3 种方法在帧级上分别高出了 37%、23% 和 9%,在像素级上更是都高出了 40% 以上。与 RESNT 方法相比在帧级上高出了 9%,与 STC 和 ISTL 方法相比在帧级上比较接近,但对应的 EER 值略逊一筹。从不同场景看,本文的方法在 Ped2 数据集的效果要优于 Ped1,这是因为在 Ped1 数据集中由于场景视角的影响,导致光流场的构建以及人群的分组效果变差,所以影响了后续的异常检测结果。总的来说,本文方法的优势在于分组后保证了行人的完整性,对于场景中存在的自行车、汽车等局部异常也能很好的检测出来,能够较为准确定位单处和多处的异常发生位置。

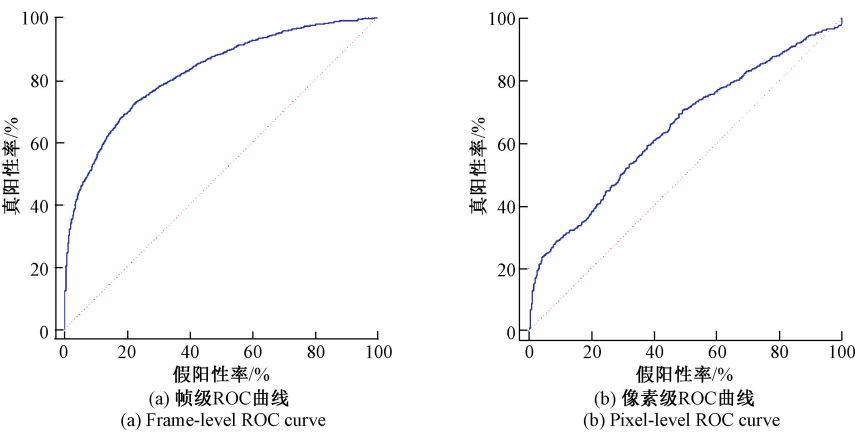


图 10 Ped1 数据集 ROC 曲线
Fig. 10 ROC curve of Ped1 dataset

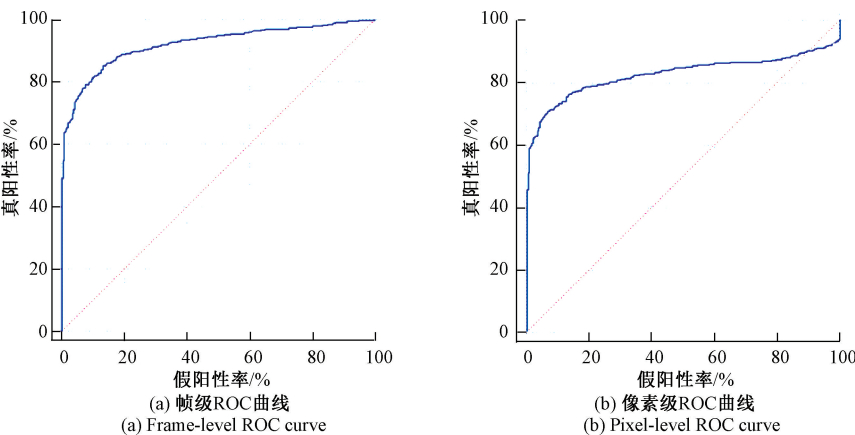


图 11 Ped2 数据集 ROC 曲线
Fig. 11 ROC curve of Ped2 dataset

表 1 Ped1 数据集帧级和像素级下的 EER 值和 AUC 值比较

Table 1 Comparison of EER and AUC values at frame level and pixel level for Ped1 dataset

方法	Ped1 帧级别		Ped1 像素级别	
	AUC	EER/%	AUC	EER/%
SF ^[10]	0. 67	31	0. 20	—
MPPCA ^[20]	0. 59	40	0. 20	83
MDT ^[21]	0. 82	25	0. 44	58
H-MDT ^[22]	—	17. 8	—	—
STC ^[23]	—	15	—	27
RESNET ^[24]	0. 72	—	—	—
ISTL ^[25]	0. 75	29. 8	—	—
Ours	0. 83	25	0. 64	40

表 2 Ped2 数据集帧级和像素级下的 EER 值和 AUC 值比较

Table 2 Comparison of EER and AUC values at frame level and pixel level for Ped2 dataset

方法	Ped2 帧级别		Ped2 像素级别	
	AUC	EER/%	AUC	EER/%
SF ^[10]	0. 55	42	0. 17	79
MPPCA ^[20]	0. 69	30	0. 13	82
MDT ^[21]	0. 83	25	0. 42	55
H-MDT ^[22]	—	18. 5	—	—
STC ^[23]	0. 92	13	—	26
RESNET ^[24]	0. 83	—	—	—
ISTL ^[25]	0. 91	8. 9	—	—
Ours	0. 92	14	0. 83	21

6 结 论

本文提出了一种基于行人组运动信息表达的人群异常检测方法,首先通过分子扩散理论来对光流场进行预处理,减少了行人内部光流的杂乱,使人群运动场更符合整体运动状态。然后采用了一种递进式的聚类策略,先对其方向分组,再对每个方向组依据时空信息聚类,避免了对图像均匀切块分割方式导致的行人完整性被破坏。在特征提取时考虑到行人在运动时姿态的变化,有效特征主要集中在行人重心位置,通过能量场构建了环状能量直方图特征来削弱行人四肢摆动及行人外围的噪声影响,可以更好地表征行人运动信息,最后将特征用于人群异常检测。实验结果表明,本文方法在帧级和像素级都有较高的准确率。在未来的工作中,将会着力于提高场景的适用性以及算法参数的优化方面。

参考文献

[1] XIE S, ZHANG X, CAI J. Video crowd detection and abnormal behavior model detection based on machine learning method [J]. Neural Computing and Applications, 2019, 31(1): 175-184.

[2] 吴新宇,郭会文,李楠楠,等. 基于视频的人群异常事件检测综述[J]. 电子测量与仪器学报, 2014, 28(6): 575-584.

WU X Y, GUO H W, LI N N, et al. A review of video-based crowd anomaly event detection [J]. Journal of Electronic Measurement and Instrumentation, 2014, 28(6): 575-584.

[3] YOGAMEENA B, NAGANANTHINI C. Computer vision based crowd disaster avoidance system: A survey [J]. International Journal of Disaster Risk Reduction, 2017, 22: 95-129.

[4] MA J, DAI Y, HIROTA K. A survey of video-based crowd anomaly detection in dense scenes [J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2017, 21(2): 235-246.

[5] ZHANG D, XU K, LU Y, et al. Abnormal crowd motion detection with hidden conditional random fields model [J]. International Journal of Multimedia and Ubiquitous Engineering, 2015, 10(10): 91-98.

[6] LOY C C, XIANG T, GONG S. Detecting and discriminating behavioural anomalies [J]. Pattern Recognition, 2011, 44(1): 117-132.

[7] SABIH M, VISHWAKARMA D K. Crowd anomaly detection with LSTMs using optical features and domain knowledge for improved inferring [J]. The Visual Computer, 2022, 38(5): 1719-1730.

[8] KOTKAR V A, SUCHARITA V. Praticle filtering based optical flow computing model for crowd anomaly detection using Gaussian mixture model [J]. International Journal of Recent Technology and Engineering, 2019, 6: 583-591.

[9] LALIT R, PURWAR R K. Crowd abnormality detection using optical flow and GLCM-based texture features [J]. Journal of Information Technology Research (JITR), 2022, 15(1): 1-15.

[10] ROSHTKHARI M J, LEVINE M D. An on-line, real-time learning method for detecting anomalies in videos using spatio-temporal compositions [J]. Computer Vision and Image Understanding, 2013, 117(10): 1436-1452.

[11] MEHRAN R, OYAMA A, SHAH M. Abnormal crowd behavior detection using social force model [C]. 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Miami, FL, USA, 2009: 935-942.

- [12] CHAKER R, AL AGHBARI Z, JUNEJO I N. Social network model for crowd anomaly detection and localization [J]. Pattern Recognition, 2017, 61: 266-281.
- [13] 汪双承, 管业鹏. 基于粒子流和能量模型的异常行为检测[J]. 电子器件, 2014, 37(6): 1081-1086.
WANG SH CH, GUAN Y P. Anomalous behavior detection based on particle flow and energy model [J]. Electronic Devices, 2014, 37(6): 1081-1086.
- [14] 钟志, 徐扬生, 石为人, 等. 群体异常检测 (英文)[J]. 仪器仪表学报, 2007, 28(4): 614-620.
ZHONG ZH, XU Y SH, SHI W R, et al. Population anomaly detection (in English) [J]. Chinese Journal of Scientific Instrument, 2007, 28(4): 614-620.
- [15] JIANG Y, WANG J, LIANG Y, et al. Combining static and dynamic features for real-time moving pedestrian detection[J]. Multimedia Tools and Applications, 2019, 78(3): 3781-3795.
- [16] 张伟峰, 周智, 赵斌, 等. 基于运动矢量的人群异常事件实时检测[J]. 计算机系统应用, 2017, 26(8): 227-231.
ZHANG W F, ZHOU ZH, ZHAO B, et al. Real-time detection of abnormal crowd events based on motion vectors [J]. Computer System Applications, 2017, 26(8): 227-231.
- [17] ABDELALI H A, ESSANNOUNI F, ESSANNOUNI L, et al. Fast and robust object tracking via accept-reject color histogram-based method [J]. Journal of Visual Communication and Image Representation, 2016, 34: 219-229.
- [18] ILYAS Z, AZIZ Z, QASIM T, et al. A hybrid deep network based approach for crowd anomaly detection[J]. Multimedia Tools and Applications, 2021, 80(16): 24053-24067.
- [19] 崔杰, 胡长青, 徐海东. 基于帧差法的多波束前视声呐运动目标检测[J]. 仪器仪表学报, 2018, 39(2): 169-176.
CUI J, HU CH Q, XU H D. Multi-beam forward-looking sonar motion target detection based on frame difference method [J]. Chinese Journal of Scientific Instrument, 2018, 39(2): 169-176.
- [20] LAB S V C. Anomaly detection and localization in crowded scenes [OL]. <http://www.svcl.ucsd.edu/projects/anomaly/>, 2014.
- [21] KIM J, GRAUMAN K. Observe locally, infer globally: A space-time MRF for detecting abnormal activities with incremental updates [C]. 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 2009(June): 20-25.
- [22] MAHADEVAN V, LI W, BHALODIA V, et al. Anomaly detection in crowded scenes [C]. 2010 IEEE Computer Society Conference On Computer Vision and Pattern Recognition, Bismarck, IEEE, 2010: 1975-1981.
- [23] LI W, MAHADEVAN V, VASCONCELOS N. Anomaly detection and localization in crowded scenes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 36(1): 18-32.
- [24] ROSHTKHARI M J, LEVINE M D. An on-line, real-time learning method for detecting anomalies in videos using spatio-temporal compositions [J]. Computer Vision and Image Understanding, 2013, 117(10): 1436-1452.
- [25] PANG G, YAN C, SHEN C, et al. Self-trained deep ordinal regression for end-to-end video anomaly detection [J]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020: 12170-12179.
- [26] NAWARATNE R, ALAHAKOON D, SILVA D, et al. Spatiotemporal anomaly detection using deep learning for real-time video surveillance [J]. IEEE Transactions on Industrial Informatics, 2020, 16(1): 393-402.

作者简介



周生运, 2019 年于山东建筑大学获得学士学位, 现为杭州电子科技大学硕士研究生, 主要研究方向为视频图像处理与人群行为分析。

E-mail: zsy2856234574@hdu.edu.cn

Zhou Shengyun received his B.Sc. degree from Shandong Jianzhu University in 2019. He is currently a master student at Hangzhou Dianzi University. His main research interests include video image processing and crowd behavior analysis.



张旭光 (通信作者), 2008 年于中科院长春光机所获得博士学位, 现为杭州电子科技大学教授, 主要研究方向为视频图像处理与人群行为分析。

E-mail: zhangxg@hdu.edu.cn

Zhang Xuguang (Corresponding author) received his Ph.D. degree from Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences in 2008. He is currently a professor at Hangzhou Dianzi University. His main research interests include video and image processing and crowd behavior analysis.