

文章编号: 2095-2163(2023)10-0056-04

中图分类号: X924.3

文献标志码: A

# 基于运动历史图像的异常行为识别算法研究

赵琴<sup>1</sup>, 赵团结<sup>1</sup>, 郑新桥<sup>1</sup>, 秦琴<sup>1</sup>, 龙念<sup>1</sup>, 邓超<sup>2,3</sup>

(1 武昌工学院 国际教育学院 武汉 430065 ; 2 武汉科技大学 汽车与交通工程学院 武汉 430065 ;

3 武汉科技大学 智能汽车工程研究院 武汉 430065)

**摘要:** 针对日益突出的安全监护问题,本文建立了一个能够较好地识别和预测人体摔倒等异常行为的算法模型。首先,通过人体动作实验分析人体基本的动作特征,运用背景减和混合高斯背景提取算法分别求解人体静态和动态前景,获得人体运动历史图像,运用运动历史图像进行运动分割,并通过 Sobel 算子求梯度判定人体运动方向,结合运动角度将得到的特征与训练出的动作库里的动作进行特征匹配,进而判断人体行为是否异常。通过这种方法构建的视频监控系统兼具实时性和智能性,可以为视频监控系统的开发提供理论依据。

**关键词:** 安全监护; 异常行为; 前景; 运动历史图像; 特征匹配

## Research on abnormal behavior recognition algorithm based on motion history images

ZHAO Qin<sup>1</sup>, ZHAO Tuanjie<sup>1</sup>, ZHENG Xinqiao<sup>1</sup>, QIN Qin<sup>1</sup>, LONG Nian<sup>1</sup>, DENG Chao<sup>2,3</sup>

(1 College of International Education, Wuchang Institute of Technology, Wuhan 430065, China;

2 College of Automobile and Traffic Engineering, Wuhan University of Science and Technology, Wuhan 430065, China;

3 Institute of Intelligent Automobile Engineering Research, Wuhan University of Science and Technology, Wuhan 430065, China)

**【Abstract】** To solve the increasingly prominent safety monitoring problem, an algorithm model which can effectively identify and predict abnormal behaviors such as human falls is proposed in this study. Firstly, through human motion experiments, the basic motion features of the human body are analyzed. Background subtraction and mixed Gaussian background extraction algorithms are used to separate the static and dynamic foreground of the human body respectively to obtain human motion history images. Motion segmentation is performed with motion history images, and the direction of human motion is determined through Sobel algorithm gradient. The obtained features are matched with the trained actions in the action library based on the motion angle. Then abnormal human behavior is recognized. The video surveillance system constructed using this method possesses real-time and intelligent capabilities, providing a theoretical basis for the development of video surveillance systems.

**【Key words】** safety monitoring; abnormal behavior; foreground; motion history image; feature matching

## 0 引言

随着计算机视觉技术的不断发展、计算机硬件的处理速度以及存储能力不断提升,监控设备成本的逐渐下降,摄像头等监控装置已经被广泛的应用到各种场合,视频监控系统也逐步在监护领域扮演着越来越重要的角色。目前人体运动异常行为分析是计算视觉领域最活跃的研究主题之一,其核心是利用图像处理、图像分析、计算机视觉分析等技术从

图像序列中检测、识别、跟踪行人并对其行为进行理解和描述,进而找到异常行为<sup>[1]</sup>。

有学者提出了许多异常行为识别算法,李一波<sup>[2]</sup>等使用运动能量图像(MEI)的傅里叶描述子来区分人的摔倒行为及正常的躺下行为,这种方简单有效,但由于只是对 MEI 图像进一步的处理,丢失了大量的信息,不能用于提取人单独某个部位的动作,如当人的两只手都在挥动时,用这种方法得不到手的挥动的方向;也有些学者使用人体关节模型图

**基金项目:** 武昌工学院科学研究项目(2022KY24); 国家自然科学基金(52002298); 四川省无人系统智能感知控制技术工程实验室开放课题(WRXT2022-001); 教育部产学研合作协同育人项目(202102580026)。

**作者简介:** 赵琴(1986-),女,学士,助理实验员,主要研究方向:智能化设备; 邓超(1986-),男,博士,副教授,主要研究方向:自动驾驶。

**通讯作者:** 邓超 Email: woec@wust.edu.cn

**收稿日期:** 2022-10-13

的方法,如张军<sup>[3]</sup>等提出一种简化的关节模型图方法,由四肢和躯干组成,定义了5个对人体运动产生影响的角度参数确定人的行为,从而实现对人体异常行为识别,有益于识别人体不同部位的运动,如判别手在运动还是脚在运动,从而识别出人体的动作,但由于只是针对某一帧的图像提取动作,并不能用来表述行为特征;也有学者使用运动方向来提取人的动作,胡芝兰<sup>[4]</sup>等根据不同行为具有不同运动方向这一特点,使用块运动方向描述不同行为,并用支持向量机(SVM)分类器对异常行为进行识别这种方法可以得到人的运动方向,进而提取出人的行为,但由于不同的动作有可能运动方向也是相同的,因此也会出现问题;吴艳平<sup>[5]</sup>等使用运动历史图像(MHI),提取人体运动方向和运动区域,最后由运动方向来识别动作,但这种仅依靠动作方向的方法只对特定情况下的异常行为检测有效。在动作分割领域,Kahol<sup>[6]</sup>等提出了层次活动分割(Hierarchical Activity Segmentation, HAS)的概念;也有些学者使用隐马尔科夫链(HMM)来分割连续动作,如Lee和Kim<sup>[7]</sup>提出复杂动作的非典型动作阈值模型,但是这种方法随着动作复杂性的增加,状态数也会大大的增加,计算困难。

由于人的动作很多,本文选择其中一部分基本动作进行特征描述,包括行走、摔倒、蹲下、站起、躺着不动。为了识别人体的异常行为,本文改进MHI算法,使得异常行为检测具有普适性。首先,结合人体运动基本特征处理人体对应的前景,得到运动历史图像,运用运动历史图像进行运动分割,并判定运动方向;其次,结合运动角度将得到的特征与训练出的动作库里的动作进行特征匹配,并组合这些动作,进而判断人体行为是否异常。

## 1 前景提取

为了提取出人体行为特征,必须识别出视频中的人的位置。首先需要得到视频的前景,前景提取算法有很多。本文重点考虑背景减和混合高斯前景提取算法。

### 1.1 背景减

对于一个视频,首先储存一幅认为是背景的图像 $B(b(x,y))$ ,当视频中的每一帧传入时,记作 $A(a(x,y))$ ,将当前帧减去背景图像。再设定一个阈值,若减去后的值小于这个阈值,则认为是背景,否则,认为是前景。用公式(1)表示, $f(x,y)$ 为1时表示此图像此处为前景,为0表示为背景。

$$f(x,y) = \begin{cases} 0 & |a(x,y) - b(x,y)| < T \\ 1 & \text{else} \end{cases} \quad (1)$$

其中, $T$ 表示阈值。

背景减算法适用于背景几乎不变的情形,算法简单。然而,如果背景本身出现了小规模晃动或抖动,求解的前景效果并不好;另外,对于前景与背景相差不大的情况效果也不好。为了在背景有小幅抖动的情况下也能够得到准确的前景,采用另一种前景提取算法混合高斯背景模型进行处理。

### 1.2 混合高斯背景提取算法(GMM)

混合高斯背景提取算法认为任何一个点的像素在时间轴上符合高斯分布,可以用多个高斯分布构建背景模型。通过不断更新背景模型来适应背景的变化,当背景变化时也能够得到较好的效果。对于图像中的点, $t$ 时刻该点属于前景的概率,如式(2)所示:

$$p(x_t) = \sum_{i=1}^k w_{i,t} \eta(x_t, \mu_{i,t}, \sigma_{i,t}) \quad (2)$$

其中, $w_{i,t}$ 为 $t$ 时刻第 $i$ 个高斯分布的权值, $\eta(x_t, \mu_{i,t}, \sigma_{i,t})$ 为以 $\mu_{i,t}$ 为均值、以 $\sigma_{i,t}$ 为方差的高斯分布(正态分布)在 $x_t$ 处的概率的值。

取最大的高斯通道数 $K$ 为3~7个,每个像素高斯通道总是按照优先级 $w_{i,t}/\sigma_{i,t}$ 从大到小排列。一般对于 $k$ 有 $1 \leq k \leq K$ 。背景模型初始化时,取当前像素的值为均值 $\mu_0$ 。设定一个较大的值 $\sigma_0$ ,设定一个较小的值 $w_0$ 。

开始训练背景模型时,当 $t$ 时刻一个新的像素 $x_t$ 到来时,按照优先级 $w_{i,t}/\sigma_{i,t}$ 的高低,开始将其与已有的 $k$ 个高斯通道匹配,匹配条件如式(3)所示:

$$|x_t - \mu_{k,t}| < 2.5\sigma_{k,t} \quad (3)$$

匹配的同时进行参数更新,若匹配成功,参数更新如式(4)所示:

$$\begin{cases} w_{i,t} = (1 - \alpha) w_{i,t-1} + \alpha M_{i,t} \\ \mu_{i,t} = (1 - \beta) \mu_{i,t-1} + \beta x_t \\ \sigma_{i,t} = (1 - \beta) \sigma_{i,t-1}^2 + \beta (x_t - \mu_{i,t})^T (x_t - \mu_{i,t}) \\ \beta = \alpha \eta(x_t, \mu_{i,t}, \sigma_{i,t}) \end{cases} \quad (4)$$

其中, $\alpha$ 、 $\beta$ 为学习率。

对于参数 $M_{i,t}$ ,若匹配则为1,不匹配时为0。若不匹配,则权值按式(4)更新,均值、方差保持不变。如果 $k$ 个高斯通道都不匹配,就新增加一个高斯分布。如果此时 $k = K$ ,则去掉一个优先级最低的高斯通道。新增的高斯通道的均值为 $x_t$ ,方差可以设置为一个较大的值,权重可以设置为一个较小的值。

每一次更新后,都按照优先级  $w_{i,t}/\sigma_{i,t}$  由高到低对高斯分布重新排序。取前  $B$  个高斯分布组合描述背景,如式(5)所示:

$$B = \arg \min_b \left( \sum_{i=1}^b w_{i,t} > T \right) \quad (5)$$

其中,  $T$  为阈值。

只要前  $B$  个高斯组合分布中有一个与  $x_t$  匹配,就认为为背景,否则认为前景。

按照上述方法,可以根据视频流不断的更新背景的模式,达到获取前景的目的。使用 GMM 方法可以抑制背景规律运动,如小草的晃动以及电风扇的转动等。但这个方法也有缺点,即不能用来识别非运动的目标。而且人一开始就在视频中时,刚开始得到的前景会有问题。

由于在所建立的人体动作模型中动作是已知的,而用来表示这些动作的视频也是已知的。因此,在提取前景时,当提取静止的状态(如躺在地上不动)时,本文使用背景减的方法;当提取运动的状态(如跑步)时,本文使用 GMM 方法。

## 2 MHI 特征的提取

得到前景的掩码后对前景进行处理,为后续的动作识别做准备。本文选用 MHI (Motion History Image) 对前景进行处理。MHI 可以保存一段时间的前景图像,并且算法复杂度低,适合提取动作。

MHI 图像主要是保存最近一段时间的前景图像,其中像素值赋值为对应前景的时间戳。若有重复交叉的前景部分,以最新的前景的时间戳为准。非前景的部分以及一段时间之前的前景赋值为 0。首先进行初始化,将 MHI 图像全部赋值为 0。当接收到新图像时,更新方式如式(6)所示:

$$tMHI_{\delta}(x,y) = \begin{cases} 0 & \text{if } MHI_{\delta}(x,y) < (\tau - \delta) \\ \tau & \text{else} \end{cases} \quad (6)$$

其中,  $\tau$  为当前的时间戳,  $\delta$  为前景的保存时间。

由于时间戳是浮点数类型,因而 MHI 图像也是一个浮点数类型的图像。如果直接显示,当  $\tau \geq 1$  时,将无法分辨不同时刻之间的区别。因此,在显示时对 MHI 图像做一个映射,将其映射到 0~255 的灰度区间,并用 8 位无符号整形数表示。视频连续 8 帧原图及使用映射后的 MHI 图像如图 1 所示,可见当前的前景最亮,而较早的前景较暗。

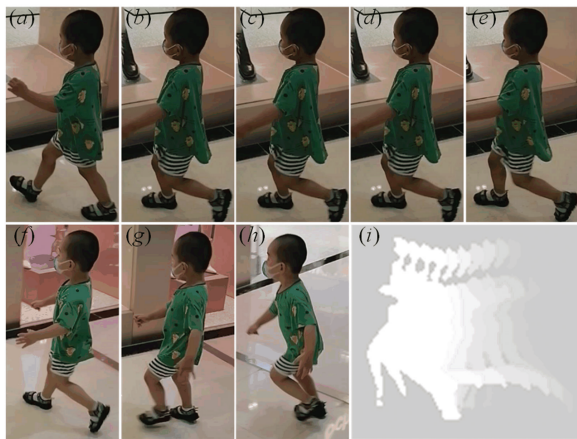


图1 原图及映射后的 MHI 图像

Fig. 1 Original and MHI Image

## 3 MHI 图像的处理及动作特征的表达

为了处理 MHI 图像,首先可以对 MHI 图像做动作分割并提取运动方向,然后判断人的动作。

### 3.1 动作分割

当提取动作特征时,需要对动作进行分割。算法流程如下:

(1) 遍历 MHI 图像,寻找属于当前前景(即  $I(x,y) = \tau_t$ ) 的点,该点是在当前前景边界上的点;

(2) 判断该点周围的点中是否有与该点邻近时间戳的点(即  $I(x,y) = \tau_{t-1}$ ),若没有则继续逆时针遍历当前前景边界上的点,直到有这种点,则进行第 3 步,或回到初始边界点为止;

(3) 用一个新值标记这个相邻点,并用新值标记这个相邻点的连通域,若这个连通域周围存在此连通域前一个时间戳的连通域,则此点周围的连通域也使用该连通域的标记,再以这个连通域为基准,依此不断寻找周围前一个时间戳的连通域并标记,若不能标记,则返回第 2 步;

(4) 标记完后,将标记的区域提取出来,就得到了分割后的运动区域。

### 3.2 运动方向提取

依据得到的 MHI 图像可以求得属于前景的每一点的梯度方向。使用 Sobel 算子求梯度,Sobel 算子模板如式(7)所示:

$$x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \quad y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad (7)$$

求解  $x$  和  $y$  方向的导数  $F_x(x,y)$  和  $F_y(x,y)$ ,并由此求得梯度方向,如式(8)所示:

$$\phi(x, y) = \arctan \frac{F_y(x, y)}{F_x(x, y)} \quad (8)$$

梯度方向即运动方向, 需要说明的是并非所有在 MHI 图像中表示前景的点都可以表示运动方向。如梯度为 0 的点是前景内部的点, 不能用来表示运动方向; 边界上的点才能表示运动方向, 有些边界上的点旁边是背景, 梯度很大, 梯度大的点不能用来求运动方向, 即用来求运动方向的点的梯度值不能太大也不能太小。

为了求得运动方向, 就要求全局梯度方向。先求符合要求的点的梯度方向, 取其中次数最多的梯度方向  $\phi_{ref}$  作为全局梯度方向, 采用公式(9)来求全局梯度方向:

$$\bar{\phi} = \phi_{ref} + \frac{\sum_{x,y} \text{angDiff}(\phi(x, y), \phi_{ref}) \times \text{norm}(\tau, \delta, t, MHI_{\delta}(x, y))}{\sum_{x,y} \text{norm}(\tau, \delta, t, MHI_{\delta}(x, y))} \quad (9)$$

其中,  $\text{angDiff}(\phi(x, y), \phi_{ref})$  为当前点的梯度方向和  $\phi_{ref}$  之间的差值,  $\text{norm}(\tau, \delta, t, MHI_{\delta}(x, y))$  为将 MHI 图像的值从  $\tau$  和  $\tau - \delta$  之间归一化到 0~1 后得到的值, 用来作为差值的权值, 即时间上离现在越近的点, 求得的梯度方向的权值越大。同时, 也可以对之前动作分割中得到的区域求运动方向, 效果如图 2 所示。图 2 方框即是识别出来的运动区域, 而圆中指针的方向就是识别出来的运动方向。



图 2 运动区域及方向的提取

Fig. 2 Extraction of motion area and direction

## 4 结束语

视频监控系统在监护领域扮演着越来越重要的角色。本文建立了一个能够较好地识别和预测人体摔倒等异常行为的算法模型。结合人体运动基本特征处理人体对应的前景, 得到运动历史图像, 运用运动历史图像进行运动分割, 并判定运动方向, 结合运动角度将得到的特征与训练出的动作库里的动作进行特征匹配, 并组合这些动作, 进而判断人体行为是否异常。通过这种方法构建的视频监控系统兼具实时性和智能性, 可以为视频监控系统的开发提供理论依据。未来将开展更多的实验, 对具体的应用场景进行图像标定, 对该算法模型的识别精度和实时性进行验证。

## 参考文献

- [1] 李瑞峰, 王亮亮, 王珂. 人体动作行为识别研究综述[J]. 模式识别与人工智能, 2014, 27(1): 35-48.
- [2] 李一波, 丁裕林. 基于智能监控的独居老人室内异常行为检测[J]. 计算机应用与软件, 2014, 31(2): 188-190.
- [3] 张军, 刘志镜. 基于模糊理论的行人异常动作检测[J]. 模式识别与人工智能, 2010, 23(3): 421-427.
- [4] 胡芝兰, 江帆, 王贵锦, 等. 基于运动方向的异常行为检测[J]. 自动化学报, 2008, 34(11): 1348-1357.
- [5] 吴艳平, 崔宇, 胡士强. 基于运动图像序列的异常行为检测[J]. 计算机应用研究, 2010, 27(7): 2741-2744.
- [6] YOON H S, SOH J, BAE Y J, et al. Hand gesture recognition using combined features of location, angle and velocity [J]. Pattern recognition, 2001, 34(7): 1491-1501.
- [7] LEE H K, KIM J H. An HMM-based threshold model approach for gesture recognition [J]. IEEE Transactions on pattern analysis and machine intelligence, 1999, 21(10): 961-973.