

## 基于动态特征的突发暴力事件辨识算法

蒲静<sup>1</sup>,傅饶<sup>1</sup>,胡栋<sup>2</sup>,卢璇<sup>2</sup>

(1. 西华师范大学 计算机学院, 四川 南充 637002; 2. 电子科技大学 计算机科学与工程学院, 成都 611731)

(15451278@qq.com)

**摘要:**在智能监控系统中,公共场所发生的暴力事件应及时预警,否则可能造成严重的后果。针对该情况,对暴力事件发生时图像中反映出来的信息进行了全面的提取,提出一种基于动态特征的能量算法。该算法从运动的剧烈程度、运动的不规则程度和目标间的位置关系3个角度较为全面地描述了该事件发生时图像的数字特征,并通过引入动能和势能总结出能量函数,用于辨识是否发生突发暴力事件。最后通过对大量包含各类事件的视频进行实验,并得出结论,该方法在处理突发暴力事件时有较高的识别率。

**关键词:**暴力事件,光流特征,动态特征,能量函数,质量权重

**中图分类号:** TP391.41 **文献标志码:** A

## Sudden violence identification algorithm based on motion feature

PU Jing<sup>1</sup>, FU Rao<sup>1</sup>, HU Dong<sup>2</sup>, LU Xuan<sup>2</sup>

(1. College of Computer Science, China West Normal University, Nanchong Sichuan 637002, China;

2. College of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu Sichuan 611731, China)

**Abstract:** In the intelligent monitoring system, violence behavior in public place should be warned in time; otherwise, it can lead to serious consequences. In this paper, the information of image was abstracted comprehensively when violence behavior happened, and an energy algorithm based on motion feature was proposed. The digital features of image from moving intensity, moving irregularity and location between targets when violence behavior happened were described, and the energy function using kinetic energy and potential energy was summarized to identify violence behavior. Finally, lots of videos that include all kinds of behaviors were used to carry out experiments and draw the conclusion that this method has a higher identification rate when handles sudden violence incident.

**Key words:** sudden violence; optical-flow feature; motion feature; energy function; quality weight

### 0 引言

通常的行为识别主要针对单个目标,并通过提取特征—训练—归类<sup>[1]</sup>的模式辨别目标的具体行为(如走路、跑动、跳跃等)。针对多个目标的情况,文献[2]利用L-K算法<sup>[3]</sup>得到光流特征,对目标的运动有直观反映。文献[4]提出了一种基于图片架构的方法,它将人体框架划分为5~10个部分,通过研究关节角度和骨骼的走向,准确地提取出特征信息。文献[5]将此方法用于姿势估计,效果较好,但该算法计算量很大,不满足系统实时性的需求。

本文主要研究多个目标(人体)间的行为,对发生突发暴力行为(打斗)的情景予以预警处理。为此,本文在传统光流特征的基础上,提出了一种基于动态特征的暴力行为识别算法。该研究可以运用于小区、办公室以及监狱等公共场所。

### 1 传统的光流特征

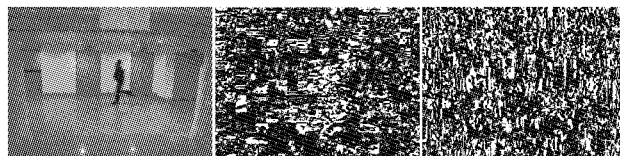
传统的光流法主要用于前景提取,它捕捉目标的运动速度,以此来区分前景和背景。该方法需要采集的参数列表如表1所示。

首先跟踪每个像素点的灰度信息,并观察连续帧之间该点的速度变化信息,分别利用水平方向和垂直方向来记录该信息,得到光流速度的复数形式  $x_{i,j}(n) + j \cdot y_{i,j}(n)$ , 其中

$x_{i,j}(n), y_{i,j}(n)$  分别是像素点  $(i,j)$  在第  $n$  帧的速度水平分量 and 垂直分量,如图1所示。

表1 传统光流特征参数表

| 参数      | 描述      |
|---------|---------|
| $X$     | 光流的水平分量 |
| $Y$     | 光流的垂直分量 |
| $V$     | 光流速度的幅值 |
| $Angle$ | 速度的方向角  |



(a) 原图 (b) 水平分量 (c) 垂直分量

图1 奔跑过程中某帧图像的速度水平分量和垂直分量

光流速度的幅值可由式(1)计算:

$$v_{i,j}(n) = \sqrt{x_{i,j}(n)^2 + y_{i,j}(n)^2} \quad (1)$$

速度的方向计算公式为:

$$\text{angle}_{i,j}(n) = \arctan \frac{y_{i,j}(n)}{x_{i,j}(n)} \quad (2)$$

图2描述了两人在普通见面和发生打斗时表现出的光流

收稿日期:2010-06-04;修回日期:2010-07-16。

**作者简介:**蒲静(1977-),女,四川南充人,讲师,硕士,主要研究方向:计算机网络、图形图像处理;傅饶(1982-),男,重庆人,讲师,硕士,主要研究方向:计算机网络、图形图像处理;胡栋(1984-),男,四川广安人,硕士研究生,主要研究方向:嵌入式实时软件、图形图像处理;卢璇(1985-),女,四川眉山人,硕士研究生,主要研究方向:嵌入式实时软件、图形图像处理。

特征,箭头描述了像素点的速度大小和方向角。

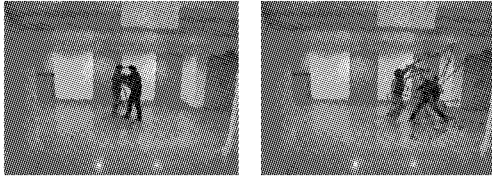


图2 会面与打斗时的光流特征在图像中的表现

由于表1中的特征仅仅描述了单个点的运动趋势,没有考虑到目标团块内部各点的关系以及团块间的位置关系,因此不能全面反应发生打斗行为时图像的特征。

## 2 光流特征的改进

本文在传统光流特征的基础上引进了几个新的动态特征,更全面地描述了多人间发生打斗行为时图像反映出来的数字信息,具体参数表如表2所示。

表2 动态特征参数表

| 参数                | 描述          |
|-------------------|-------------|
| $X$               | 动态特征的水平分量   |
| $Y$               | 动态特征的垂直分量   |
| $V$               | 速度的幅值       |
| $Angle$           | 速度的方向角      |
| $\Delta Angle$    | 相邻帧的方向角度差   |
| $\Delta MaxAngle$ | 与速度主方向的角度差  |
| $r$               | 目标团块之间的欧氏距离 |

### 2.1 相邻帧的方向角度差

根据式(2)计算得到的相邻帧速度方向角,即可得到相邻帧的方向角度差:

$$\Delta angle_{i,j} = angle_{i,j}(n) - angle_{i,j}(n-1) \quad (3)$$

### 2.2 速度主方向

遍历整个团块,寻找该团块的速度主方向。观察人体运动的特点发现,上半身主要表征了目标运动的主要趋势,而下半身呈周期性运动,因此选择上半身的速度方向为运动的主方向。

首先将跟踪得到的目标团块保存在下面的数据结构中:

```
struct CvBlob
{
    float x,y;
    float w,h;
    int ID;
}
```

结构中: $(x,y)$ 为团块的中心坐标; $w,h$ 分别是其宽度和高度; $ID$ 标识了团块号。

根据CvBlob结构的中心坐标确定上半身的区域,用四元组保存(left, right, up, down),且有 $left = x - w/2$ ,  $right = x + w/2$ ,  $up = y - h/2$ ,  $down = y$ 。

求该搜索框中速度幅值的均值 $V_{max}(n)$ ,令其为运动主方向,则有:

$$V_{max}(n) = \sqrt{\left(\sum_{i=up}^{down} \sum_{j=left}^{right} x_{i,j}(n)\right)^2 + \left(\sum_{i=up}^{down} \sum_{j=left}^{right} y_{i,j}(n)\right)^2} \quad (4)$$

主方向角 $angle_{max}(n)$ 为:

$$angle_{max}(n) = \arctan \frac{\sum_{i=up}^{down} \sum_{j=left}^{right} y_{i,j}(n)}{\sum_{i=up}^{down} \sum_{j=left}^{right} x_{i,j}(n)}$$

### 2.3 与速度主方向的角度差

根据前面的式(2)和式(4)可知:

$$\Delta angle M_{i,j}(n) = angle_{i,j}(n) - angle_{max}(n) \quad (5)$$

### 2.4 团块间的欧氏距离

在二维空间中,目标团块1( $x_1, y_1$ )和目标团块2( $x_2, y_2$ )之间的欧氏距离 $r$ 为:

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (6)$$

## 3 基于动态特征的能量法

第2章详细描述了表2中的各个特征,本文由此引入能量的概念。能量公式可分为动能部分和势能部分,动能部分描述了单个目标的运动剧烈程度和各部分运动方向的不规则程度,势能部分描述了目标间的位置与发生打斗行为间的联系,下面主要探讨其具体表达形式。首先研究动能共有的质量部分 $m_{i,j}(n)$ 。

### 3.1 质量权重 $m_{i,j}(n)$

$m_{i,j}(n)$ 描述了不同目标发生不同行为时所占的权重,文献[6]中提出了几种计算权重的方法,并进行了比较,其中一个较优的方法如下:

$$m_{i,j}(n) = \left( \frac{|\Delta angle_{i,j}(n)|}{\pi} \times 10 \right)^2 + \left( \frac{|\Delta angle M_{i,j}(n)|}{\pi} \times 10 \right)^2 \quad (7)$$

其中: $\Delta angle_{i,j}(n)$ 和 $\Delta angle M_{i,j}(n)$ 分别由式(3)和式(5)得出, $v_{i,j}(n)$ 为速度的幅值,由式(1)求得。

### 3.2 动能部分

将每个像素点看成一个运动的质点,根据大学物理中描述物体运动能量的动能公式 $E = \frac{1}{2}mv^2$ 可将图像中单个目标运动产生的动能描述为该团块外接矩形以内每个像素点的动能之和。因此,第 $k$ 个团块的动能为:

$$E_{mom}(n, k) = \sum_{i=up_k}^{down_k} \sum_{j=left_k}^{right_k} m_{i,j}(n) v_{i,j}(n)^2 \quad (8)$$

式(7)中的二元组 $(i,j)$ 表示每个像素点的位置信息, $n$ 为当前的帧数, $m_{i,j}(n)$ 为该点在当前帧动态特征的质量权重。

根据上面的结论,不妨设 $num$ 为图像中的目标团块总数,则可以根据整个图像的动能公式:

$$E_{mom}(n) = \sum_{k=1}^{num} E_{mom}(n, k)$$

并结合式(7)、(8)得到第 $n$ 帧图像的动能为:

$$E_{mom}(n) = \sum_{k=1}^{num} \sum_{i=up_k}^{down_k} \sum_{j=left_k}^{right_k} \left( \left( \frac{|\Delta angle_{i,j}(n)|}{\pi} \times 10 \right)^2 + \left( \frac{|\Delta angle M_{i,j}(n)|}{\pi} \times 10 \right)^2 \right) v_{i,j}(n)^2 \quad (9)$$

### 3.3 势能部分

由于本文研究的对象是人体,因此距离越近发生打斗等暴力事件的概率越大。联系到天体间的万有引力定律,可将图像看做是一个各目标相互牵扯的势能场,因此势能公式可以描述为:

$$E_{pot} = Km_1 m_2 / r^2$$

其中: $K$ 为平衡系数; $r$ 为两个目标团块间的欧氏距离,可由式(6)得到; $m_1$ 和 $m_2$ 分别是它们的质量权重。根据式(7),第 $n$ 帧时 $k$ 号团块的质量权重为:

$$m_k(n) = \sum_{i=up}^{down} \sum_{j=left}^{right} \left( \left( \frac{|\Delta angle_{i,j}(n)|}{\pi} \times 10 \right)^2 + \left( \frac{|\Delta angle M_{i,j}(n)|}{\pi} \times 10 \right)^2 \right) \quad (10)$$

因此,第  $n$  帧图像的势能为:

$$E_{pot}(n) = \begin{cases} 0, & num = 1 \\ \sum_{l=1}^{num} \sum_{k=l+1}^{num} \frac{Km_k m_l}{r^2}, & num = 2 \\ \frac{1}{2} \sum_{l=1}^{num} \sum_{k=l+1}^{num} \frac{Km_k m_l}{r^2}, & num > 2 \end{cases} \quad (11)$$

由式(9)和(11),第  $n$  帧图像产生的总能量为:

$$E(n) = E_{mom}(n) + E_{pot}(n) \quad (12)$$

#### 4 实验分析

本实验采用了行走、奔跑、正常会面、并排行走、超越及打斗视频若干段,分别利用式(12)计算每帧图像的能量总和。图3描绘了单人行走和奔跑场景的能量曲线。

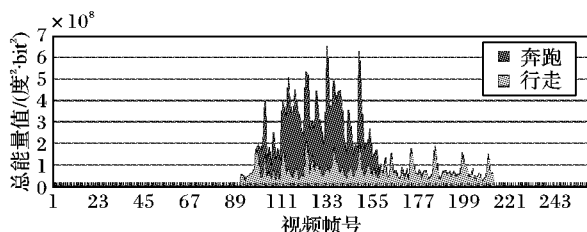


图3 行走和奔跑行为的能量曲线

由图3可知,行走时能量的局部峰值以大约11帧的周期呈周期性变化,但幅度变化不大;奔跑时周期性不明显,且能量值远大于行走。

图4描绘了两人正常会面和发生打斗行为场景的能量曲线。

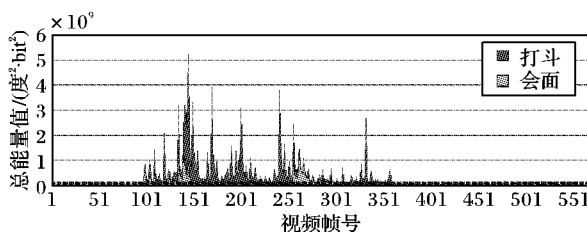
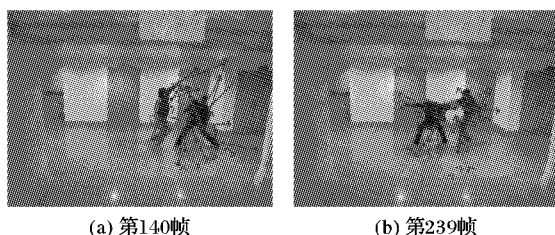


图4 会面和打斗行为的能量曲线

由图4可知,两人会面时(第148帧至第251帧)几乎处于静止状态,因此能量值趋近于0;发生打斗时,能量值很大,且变化极不规则。当动作变化非常剧烈且不规则时能量达到极大值,图5为打斗场景时第139和240帧的动态特征,此时能量值都处于局部峰值。



(a) 第140帧 (b) 第239帧

图5 打斗行为的动态特征信息表示

由于行为是一个连续的过程,不能单凭一帧的信息作判断,因此将上面的数据按当前帧之前的连续15帧进行积分。积分后每个时间点的值  $S(n)$  为前15帧能量值总和:

$$S(n) = \sum_{i=1}^{15} E(i)$$

$S(n)$  的曲线图如图6所示。

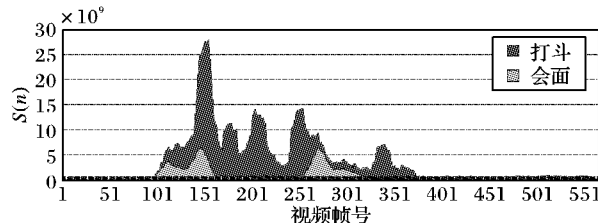
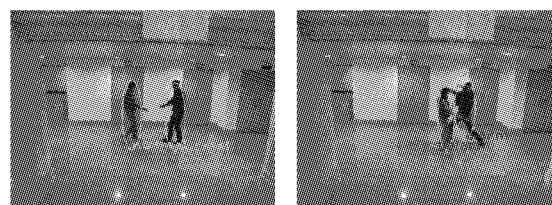


图6 会面和打斗行为下的积分能量累加值

若  $S(n)$  大于经验阈值  $\varepsilon$ ,则认为当前帧(第  $n$  帧)发生突发暴力行为,予以标记并预警,如图7所示。



(a) 会面 (b) 打斗

图7 正常会面和打斗场景下的预警图

本文将突发暴力行为(发生打斗)定义为异常行为,其他行为视为正常。若正常行为被判为异常,异常行为未被识别或未在对应帧识别,则记为误判。表3说明了各种行为场景下本文算法的识别率。

表3 MEF法对各类视频的处理结果

| 视频内容 | 视频总数 | 误判视频 | 准确率/%  |
|------|------|------|--------|
| 单人行走 | 25   | 0    | 100.00 |
| 单人跑动 | 25   | 2    | 92.00  |
| 正常会面 | 20   | 4    | 80.00  |
| 并排行走 | 20   | 3    | 85.00  |
| 超越   | 30   | 7    | 76.67  |
| 打斗   | 30   | 4    | 86.67  |

则总误判率 =  $20/150 \approx 13.33\%$ 。

#### 4 结语

由于本文提出的动态特征较全面地反应了暴力行为的特征,即速度快、速度方向变化不规则且目标距离近,因此该方法针对突发暴力事件的识别率较高。但当发生快速的超越行为时,识别率有一定的下降,针对该问题,将在接下来的研究中从整体运动趋势上作分析,提高识别率。另外,该算法中经验阈值  $\varepsilon$  的设定与具体环境有关,可以通过机器学习寻找一个自适应阈值,这也是接下来的工作。

#### 参考文献:

- [1] DONG JIANMIN, WANG RUOPENG. A novel smooth support vector machines for classification and regression[C]// ICCSE '09: 4th International Conference on Computer Science and Education. Washington, DC: IEEE Computer Society, 2009: 12-17.
- [2] ZHI ZHONG, YE WEIZHONG, XU YANGSHENG. Detecting human abnormal behaviors in crowd[J]. International Journal of Information Acquisition, 2007, 4(4): 281-290.
- [3] LUCAS B D, KANADE T. An iterative image registration technique with an application to stereo vision[C]// IJCAI81: Proceedings of the 7th International Joint Conference on Artificial Intelligence. Vancouver: Morgan Kaufman, 1981: 674-679.

(下转第425页)