

基于运动矢量的人群异常事件实时检测^①

张伟峰, 周 智, 赵 斌, 黄 露, 朱 明

(中国科学技术大学 信息科学技术学院, 合肥 230026)

摘 要: 近年来, 城市公共安全形势严峻, 给社会经济的可持续发展提出新的问题. 因而, 如何有效地监测突发人群异常状况, 已成为目前的一个研究热点问题. 由于人群运动目标众多且在不断变化, 所以很难通过运动目标的跟踪来研究人群异常. 研究表明: 当人群出现异常时, 最明显的变化就是人群的运动速度的大小和运动方向会突然发生变化, 比如, 由静止或慢速行走变为快速奔跑、突然改变前进的方向等, 相应的, 视频帧的运动矢量也会发生同样的变化, 由此, 提出基于运动矢量的人群异常快速检测算法. 实验结果表明, 本文提出的算法能够实时、有效的检测出人群运动的异常.

关键词: 人群异常; 运动矢量; 运动特征矩阵

引用格式: 张伟峰, 周智, 赵斌, 黄露, 朱明. 基于运动矢量的人群异常事件实时检测. 计算机系统应用, 2017, 26(8): 227-231. <http://www.c-s-a.org.cn/1003-3254/5922.html>

Real-Time Detection of Abnormal Event Based on Motion Vector

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

(University of Science and Technology of China, Hefei 230026, China)

Abstract: In recent years, the urban public security poses a new problem to the sustainable development of social economy. Therefore, how to effectively monitor the abnormal situation of the crowd has become a hot issue. Due to the large number of moving targets and the changing of the crowd, it is difficult to study the abnormalities of the crowd by tracking the moving objects. The study shows that when the crowd is abnormal, the most obvious change is the movement speed of the crowd and the direction of the movement will suddenly change. For example, from static state or slow walking to fast running, and the sudden change of the motion direction. The corresponding motion vector of the video frame will undergo the same change. Thus, we propose the fast detection algorithm for the crowd based on motion vector. The experimental results show that the algorithm proposed can detect the abnormal movement of human beings in real time and effectively.

Key words: crowd abnormal; motion vector; motion characteristic matrix

近年来, 随着社会经济的快速发展, 人民生活水平的提高, 人流高峰在大型广场、地铁站、火车汽车站等公共场所几乎时刻存在, 由于人群密度较大, 这给公共安全带来极大的隐患. 如果能够对人群进行有效的、实时的监控, 第一时间发现异常和采取相应的措施, 就能最大程度的减少意外发生和降低人身财产损失. 2014年12月31日23时35分, 正值跨年夜活动, 因很

多游客市民聚集在上海外滩迎接新年, 上海市黄浦区外滩陈毅广场东南角通往黄浦江观景平台的人行通道阶梯处底部有人失衡跌倒, 继而引发多人摔倒、叠压, 致使拥挤踩踏事件发生, 造成36人死亡, 49人受伤. 正是由于预防准备不足、现场管理不力、应对处置不当而引发的拥挤踩踏并造成重大伤亡和严重后果的公共安全责任事件. 因此, 为了更好的维护公共秩序, 保障

^① 基金项目: 中科院先导项目课题(XDA06011203)

收稿时间: 2016-12-13; 采用时间: 2017-01-09

公共场所的人群安全,需要投入越来越多的监控系统。

1 概述

牛津词典对异常的定义是:以不良或者有害的方式偏离普通类型,或者与正常的规则或制度相反。根据这个定义,异常行为可以被定义为区别于正常行为的不规则行为。因此,我们要做的就是通过正常的训练集训练出一个模型,并通过模型来识别异常行为。

异常事件的检测主要分为如下几个步骤:

① 预处理:减少视频图像中的干扰信息或冗余,提高图片质量。

② 运动区域检测:从图像序列中将人群运动目标相关的前景像素点和背景像素点分割开,提取与运动目标相关的有用信息。

③ 行为特征提取:提取视频场景中人群相关的图像特征并用来表征人群运动等信息(大多数学者基于光流向量来描述运动特征)

④ 行为建模:基于提取的特征建立描述人群和基本事件的模型。

⑤ 行为识别:设定特定的标准来判断是否发生了异常事件。一般采用的是机器学习的方法,常用的机器学习方法有 LDA^[1]、HMM^[2]、SVM^[3]、神经网络^[4]等。

目前,人群异常检测方法还有适用性和实时性两方面的不足。一方面,由于人群异常事件的多样性(踩踏、恐慌、打架斗殴等)和场景复杂性(路口、广场等),描述人群异常事件的数学模型都有一定的局限性。例如,流体动力学模型和混合动态纹理模型在某些场景下会失效,缺少广泛的适用性;另一方面,由于密集人群异常事件在短时间会造成较高的危害,这就要求检测方法要具有能够实时监测的能力。现有的各类方法,如:社会力模型^[5]、SVM、神经网络等算法复杂度高,实时性比较差,不能满足实际的需要。

针对目前检测方法在适用性和实时性方面的不足,本文提出了一种基于小区域运动矢量的实时检测方法。由于光流算法容易受到噪声、光照的影响,且计算图像的光流向量比较耗时,考虑到实时性的要求,我们采用基于运动矢量来表征各种场景下的运动特征。由于密集人群存在严重的遮挡等原因,因此采用行人跟踪的方法来检测人群异常是不现实的,运动矢量算法既不需要分离前景,也不需要背景建模,对各类场景都具有良好的适用性,而且在异常事件发生时,运动矢量会发生明显的变化,通过统计运动矢量的强度和方向的

分散程度可以很好的检测性能。

本文提出的算法有以下的优点:首先,算法对场景没有要求,具有广泛的适用性;其次,算法有很高的实时性,每帧的处理速度几十毫秒,最后,算法的准确率较高,可以达到几种经典算法水平。

2 算法设计与实现

本文的基本思路是:首先,将每一帧输入图像的运动矢量作为运动特征,然后把每一帧图像分为 $n \times n$ 的小区域块,接着分析每个区域块的运动矢量的运动强度和运动方向杂乱程度,进而确定小区域块是否存在异常,最后综合整帧图片的异常小区域块,如果异常小区域块超过可信区间范围,则判断该帧图像有异常事件发生。

2.1 H.264 视频编码标准简介

H.264 编码标准采用基于块的变换和预测的混合编码方式,H.264 协议里定义了三种帧:帧内编码帧(I 帧)、前向预测帧(P 帧)以及双向预测帧(B 帧)。I 帧利用离散余弦变换(DCT for Discrete Cosine Transform)独立编码,是视频序列的开始帧,不需要参考其他帧的信息;P 帧和 B 帧是预测编码方式,利用参考帧(I 帧或 P 帧)、运动矢量和残差 DCT 系数进行编码^[6]。

2.2 运动矢量的提取

I 帧:由于 I 帧不需要参考其他帧的信息,提取运动矢量时,直接提取零值。

P 帧:表示的是与前一帧(I 帧或 P 帧)的差别(也就是差别帧,P 帧没有完整数据,只有与前一帧差别的数据),编码时,在前一帧中找出 P 帧“某点”的预测值和运动矢量,取预测差值和运动矢量一起传送,解码时,根据运动矢量从前一帧中找出 P 帧“某点”的预测值并与差值相加,生成最终完整帧;根据压缩原理,可以直接提取“某点”的运动矢量。

B 帧:双向编码帧,参考前一帧和后一帧,提取运动矢量原理与 P 帧一样。

H.264 编码标准是基于块的,块模式包含 7 种,分别是 (16×16) 、 (16×8) 、 (8×16) 、 (8×8) 、 (8×4) 、 (4×8) 、 (4×4) ,兼顾计算量与准确性,统一将其归一化为 (4×4) 的块^[7]。

2.3 运动特征矩阵

上述可以提取整帧($M \times N$)运动矢量,运动矢量不能直接作为特征,因此我们使用运动矢量统计作为特征。具体过程如下。

图像 (i, j) 像素点处运动矢量 V_{ij} 的模长和方向分别为:

$$|V_{ij}| = \sqrt{(V_{ij}^x)^2 + (V_{ij}^y)^2} \quad (1)$$

$$\varphi_{ij} = \arctan\left(\frac{V_{ij}^y}{V_{ij}^x}\right) \quad (2)$$

根据计算的运动矢量角度,把运动矢量化成如表1所示的8个方向.

表1 运动矢量方向量化表

φ_{ij}°	0~45	45~90	315~360
φ'_{ij}	0	1	7

把每一帧视频图像分成 $n*n$ 的小区域块,则整帧图像被分成 $[M/n]*[N/n]$ 个小区域,统计每个小区域块全部运动矢量模长之和:

$$L_{i'j'} = \sum_0^k |V_{ij}| \quad (3)$$

其中 $i' = [i/n]$, 表示小区域在第 i' 行, $j' = [j/n]$, 表示小区域位于第 j' 列; k 表示该小区域运动矢量的个数, $L_{i'j'}$ 代表该小区域块的运动能量,如果运动能量大于阈值 δ_A , 则该小区域运动较剧烈,把该区域所有运动矢量的方向根据表1量化为方向值(0、1、2.....),并求出量化后方向值的方差:

$$\varphi_{i'j'} = \frac{1}{k} \sum_0^k \varphi'_{ij} \quad (4)$$

$$\sigma_{i'j'} = \frac{1}{k} \sqrt{\sum_0^k (\varphi'_{ij} - \varphi_{i'j'})^2} \quad (5)$$

如果方差 $\sigma_{i'j'}$ 大于阈值 δ_σ , 则标记为异常小区域,并把异常小区域的主方向量化成-1; 方差小于阈值 δ_σ , 则把 $\varphi_{i'j'}$ 量化成小区域块的主方向,并把小区域标记为异常存疑小区域; 如果运动能量小于阈值 δ_A , 则标记为非异常小区域,并把非异常小区域的主方向量化成-1.

根据上述原则,我们把标记为异常的小区域用2表示,异常存疑小区域用1表示,非异常的小区域用0表示,这样我们就对每帧视频图像建立一个只含有0、1和2三个值的矩阵,本文称之为运动特征矩阵(用 A 表示, $A \in R^{[M/n]*[N/n]}$).

本文就是基于运动矢量得到运动特征矩阵,得到运动特征矩阵以后,下节将根据运动特征矩阵提出判别规则,判断是否有异常事件的发生,并在几种场景下进行仿真实验进而验证算法的有效性.

2.4 异常事件判别

由于监控视频开始都是正常的,因此我们可以利用视频开始的帧,来初始化我们小区域块的运动能量阈值.根据上节所述,我们对每帧视频图像提取运动特征矩阵.根据运动特征矩阵,我们判定视频帧是否出现异常.

根据对人群行为的研究发现:当人群出现异常,如恐慌、打斗、追逐等,就会伴随着运动矢量的改变,具体表现为运动能量较大,运动方向较为杂乱.正如数据集中的人群一样,当未发生异常时,人群运动能量较小,且有一定的秩序性,当发生异常时,人群从发生异常的地方向周围各个方向快速散开,因此可以通过异常块的比例来判定帧是否出现异常.还有一种情形,当有两个人在人群中追逐,由于人群较为密集,因此也是一种人群中的异常行为.在这种情况下,两个相互追逐的人的运动矢量的运动能量较大,运动方向变化较快,在帧运动特征矩阵图中表现为连续(横向、竖向、斜向等)的几个小浅色块,因此可以通过帧运动特征矩阵中的连续小浅色块的数目,来判定帧是否发生局部异常.

综上所述,我们判定规则如下:

$$\begin{cases} \sum_{i'}^{[M/n]} \sum_{j'}^{[N/n]} (a_{i'j'} \% 2) > \varepsilon_1 * \left[\frac{M}{n}\right] * \left[\frac{N}{n}\right] \\ \frac{1}{k'} \sum_0^{k'} \sqrt{(\varphi_{i'j'} - E\varphi_{i'j'})^2} + \alpha T > \varepsilon_2 \\ Sequence(a_{i'j'} \% 2) > \varepsilon_3 \end{cases} \quad (6)$$

其中: $E\varphi_{i'j'} = \frac{1}{k'} \varphi_{i'j'}$, 表示异常存疑小区域主方向的均值, k' 表示异常存疑小区域的个数, $\varphi_{i'j'}$ 表示异常存疑小区域的主方向, $Sequence(a_{i'j'} \% 2)$ 表示连续的异常小区域数, T 即为判别条件1的左式.

当满足(6)式中的条件1时,说明整个视频帧运动较为剧烈,且运动方向较为杂乱,异常小区域数超过阈值,可能是密集人群出现了大规模的恐慌逃散等全局异常;当(6)式中的条件1不满足,条件2满足时,说明整个视频帧运动较为剧烈,异常存疑小区域数超过阈值,且异常存疑小区域的主方向较为杂乱,但人群不是很密集,可能是稍微稀疏一点的人群全局异常行为;当(6)式中的条件1和条件2都不满足,条件3满足时,说明不存在人群全局的异常的行为,可能是人群中局部的追逐、奔跑等局部异常行为,因此本文的判别规则可以判别出人群全局异常和局部异常行为,具有很好的适用性.

有些异常事件中,可能是一部分人进行聚集或者是多个人一起追逐,这也就是局部异常行为.这种异常情况也就说上面所说的不满足(6)式的条件1和条件2,

但满足条件3的情况,这种局部异常行为,必定会导致局部存在连续的异常小区域,条件3就是判断局部是否出现连续的异常小区域.通过条件3,我们可以判定存在局部的连续(横向连续、纵向连续、斜向连续)异常小区域,也就是判定了一部分人进行聚集或者多个人一起追逐等人群中局部异常行为.

2.5 算法流程

综上所述,本文算法主要分为以下三个步骤:

① 初始化:通过视频前50帧初始化上述提到的阈值,流程如图1所示.

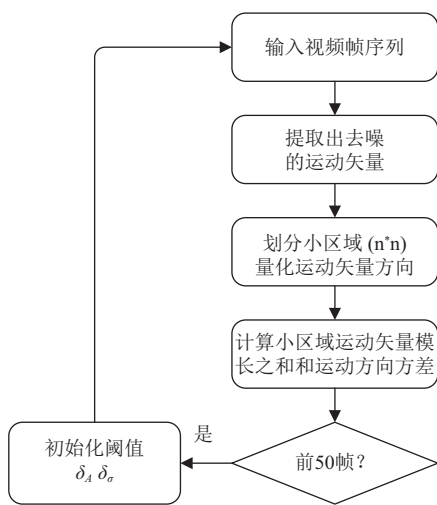


图1 初始化流程图

② 小区域异常判断:根据初始化以后的阈值来判断小区域是否发生异常,流程如图2所示.

③ 帧异常判断:根据异常事件判别条件来判断帧是否发生异常,流程如图3所示.

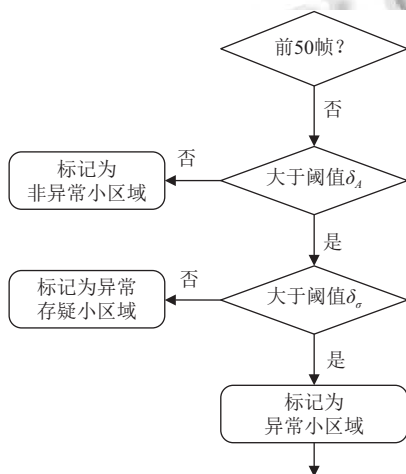


图2 小区域异常判断流程图

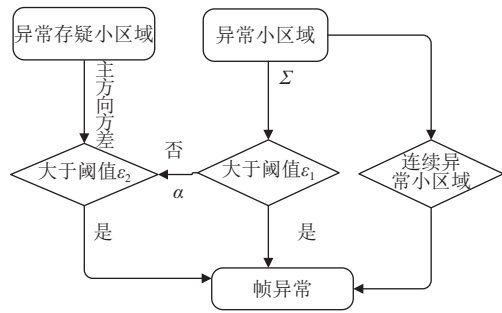


图3 帧异常判断流程图

3 实验与分析

本实验的硬件平台为 Intel i5 3.3 GHz+6 G 内存,软件平台为 VS2013+OpenCV+FFmpeg,选取了 UMN dataset^[8]中的2个场景以及自录的一个场景,总共3段视频作为我们的测试对象.

3.1 异常事件检测

根据我们提出的算法,分别对3段视频进行异常事件检测.经过大量的实验,发现当小区域块的大小选择为 20*20,运动能量的阈值为正常运动能量的 1.5~1.6 倍,连续异常小区域块为 3~4 个时,准确率最高.各个场景的实验如图4至图6(左边是原视频帧,中间是提取的运动矢量图,右边是对应的运动特征矩阵示意图,图中深色是异常存疑小区域块,浅色是异常小区域块).

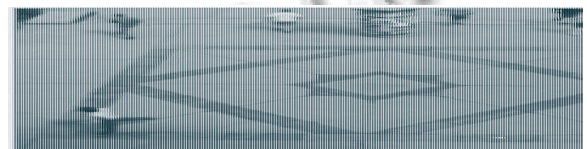


图4 场景1异常帧

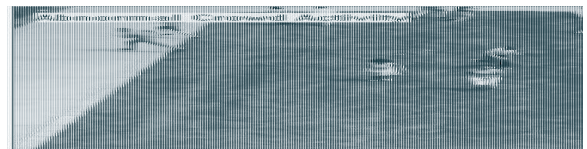


图5 场景2异常帧

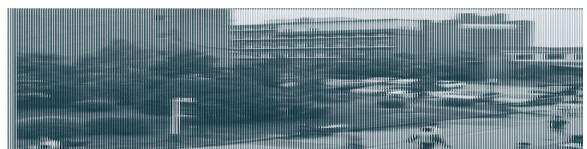


图6 场景3异常帧

3.2 实验结果与分析

在 UMN dataset 数据集上,平均每帧的处理时间为 25 ms 左右,在自录的数据集上,平均每帧的处理时间为 30 ms 左右,有的文章中处理每帧时间也是毫秒级,但是并没有把提取光流的时间算在内,这显然是不合理的,本文把提取运动矢量的时间也计算进来了,实现了真正意义上的实时性。

如表 2 所示,本文算法在几个场景下的准确率。

表 2 运动矢量方向量化表

视频源	UMN1	UMN2	自录
总帧数	2019	1450	2763
检测正确	2051	1380	2680
检测错误	78	70	83
准确率(%)	96.34	95.17	97.00

如表 3 所示,本文算法与几种经典算法相比较(社会力算法 SF^[5]、稀疏重构算法 SRC^[9]、混沌不变算法 CI^[10]、贝叶斯模型算法 BM^[11]),比较如下:

如表 4 所示,本文与几种经典算法(混合动态纹理算法 MDT^[12]、稀疏重构算法 SRC、贝叶斯模型算法 BM)在实时性方面的比较。

表 3 几种算法准确率(%)比较

视频源	SF	SRC	CI	BM	本文
UMN1	90.83	92.70	91.58	96.63	96.34
UMN2	84.41	90.52	90.62	99.09	95.17
自录	\	\	\	\	97.00

表 4 几种算法准确率比较

算法	秒/帧	CPU	视频源
SRC	110	2.6 GHz	USCD
MDT	25	3.0 GHz	USCD
BM	0.93	3.0 GHz	UMN
本文算法	0.035	3.3 GHz	UMN

4 结语

本文创新性地提出基于运动矢量的人群异常事件检测方法,相比与基于光流的算法真正意义上实现了人群异常的实时检测.实验结果表明,本文提出的算法不仅满足了实时检测(每帧毫秒级),在准确率上也达到了传统经典算法的水平.但是本文对运动量不是很大的异常事件,检测效果不明显。

下一步的工作,将侧重算法的改进,提出不同异常事件的模式,达到准确实现检测不同异常的事件。

参考文献

- Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 2003, 3: 993–1022.
- Eddy SR. Hidden Markov models. *Current Opinion in Structural Biology*, 1996, 6(3): 361–365. [doi: 10.1016/S0959-440X(96)80056-X]
- Cortes C, Vapnik V. Support-vector networks. *Machine Learning*, 1995, 20(3): 273–297.
- Haykin S. *Neural networks: A comprehensive foundation*. Englewood Cliffs: Prentice Hall PTR, 1994.
- Mehran R, Oyama A, Shah M. Abnormal crowd behavior detection using social force model. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*. Miami, FL, USA. 2009. 935–942.
- 王玉新, 刘冀伟, 李擎. H.264 视频码流中运动物体的提取. *微计算机信息*, 2006, 22(8S): 183–185.
- 邹铮, 闫玮, 谢剑斌, 等. 基于运动矢量的异常拥挤行为检测算法. *数学技术与应用*, 2016, (5): 120–122.
- University of Minnesota. Unusual crowd activity dataset of University of Minnesota. <http://mha.cs.umn.edu/movies/crowdactivity-all.avi>. [2015-07-11].
- Cong Y, Yuan JS, Liu J. Sparse reconstruction cost for abnormal event detection. *Proc. of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*. Providence, RI, USA. 2011. 3449–3456.
- Wu SD, Moore BE, Shah M. Chaotic invariants of Lagrangian particle trajectories for anomaly detection in crowded scenes. *Proc. of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA. 2010. 2054–2060.
- Wu S, Wong HS, Yu ZW. A Bayesian model for crowd escape behavior detection. *IEEE Trans. Circuits and Systems for Video Technology*, 2014, 24(1): 85–98. [doi: 10.1109/TCSVT.2013.2276151]
- Mahadevan V, Li WX, Bhalodia V, et al. Anomaly detection in crowded scenes. *Proc. of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA. 2010. 1975–1981.