

基于速度场简化的人群行为分析^①

李 楠^{②**} 张志敏^{③*}

(* 中国科学院计算技术研究所集成应用中心 北京 100190)

(** 中国科学院研究生院 北京 100049)

摘 要 针对视频监控中人群异常行为难于被检测的问题,提出了一种基于速度场简化的人群行为分析方法。速度场简化方法利用拓扑简化和聚类算法对人群运动的速度场进行分析,提取出人群运动的宏观行为信息,之后利用这些信息检测人群中发生的各种异常事件。与其它方法相比,速度场简化方法具有不需要训练的优点,并且具有较好的鲁棒性。在实验中,利用速度场简化方法对人群整体行为异常、人群速度异常、人群间行为异常这三类事件进行了检测。实验结果表明,该方法的检测结果与真实情况十分接近,能够较好地检测出与人群相关的各类异常事件。

关键词 视频监控, 人群行为分析, 速度场简化, 拓扑分类, 聚类算法

0 引 言

计算机视觉研究中的密集人群监控是近年来兴起的一个方向。密集人群场景往往十分复杂,通常包含数十个甚至上百个相互作用的独立个体,这些个体的外观十分相似,互相之间也存在着较为复杂的遮挡关系,因而传统的个体目标检测与分割、个体目标跟踪、个体目标行为分析等方法不适用于对密集人群的监控。人群行为分析是密集人群监控研究中的一项十分具有挑战性的课题。人群的行为分析主要指通过提取人群运动信息,如速度、方向、轨迹等,检测与定位人群中发生的异常事件。与传统的基于个体目标的行为分析不同,密集人群的行为分析更加侧重于人群整体运动行为的分析。

现有人群行为分析方法描述人群运动的数学模型往往并不具有分类能力,因此需要训练一个额外的分类器用于检测人群中的异常行为。Ihaddadene 等人^[1]用速度矢量的方向、幅度等特征的统计变化来检测异常事件的发生。Mehran 等人^[2]用社会力模型来描述目标之间的相互作用并利用词包模型来区分正常事件和异常事件。但是,这种方法并不能准确地提取出异常事件发生的位置。Wu 等人^[3]提出使用混沌变量来描述目标运动,并利用李雅普诺

夫指数和相关维数两种混沌不变量作为特征,训练高斯混合模型对异常事件进行分类和定位。Mahadevan 等人^[4]使用混合动态纹理对正常行为建模和训练,利用高斯混合模型和判别显著方法分别对时域和空间域上的异常事件进行检测和定位。Kratz 等人^[5]使用图像序列时空域的梯度信息训练隐式马尔科夫链模型,用于异常事件的检测。本文提出了一种基于速度场简化的人群行为分析方法,该方法不需要通过任何训练即可直接检测异常事件,并且具有较好的鲁棒特性。它首先利用拓扑简化算法提取出人群运动的主要动力学特征,检测人群整体异常行为和人群整体异常速度事件,然后利用聚类分析算法估算出图像中的人群的数目,并用其检测人群间异常行为。实验表明,此方法能有效检测出人群整体异常行为、人群异常速度、人群间异常行为这三类异常事件。本文的主要贡献在于首次将速度场简化的相关理论引入到人群行为分析中,并针对人群行为特点对其进行了改进。

1 速度矢量场的拓扑简化

根据微积分的基本理论,人群在图像上的所有运动状态可以用一个随时间变化的速度矢量场函数

① 国家自然科学基金(60703019)资助项目。

② 男,1984 年生,博士生;研究方向:智能视频监控,嵌入式系统;E-mail: linan@ict.ac.cn

③ 通讯作者,E-mail: zzm@ict.ac.cn

(收稿日期:2010-10-21)

$v(x, t)$ 描述。其中, x 表示图像的空间坐标, t 表示时间。通常情况下, 图像中包含的像素点数目较多, 导致速度矢量场的维数较高, 直接对速度场进行分析较为困难, 所需要的处理时间也相对较长。因此, 需要对速度场进行简化, 提取出其中的主要结构信息, 以减少行为分析算法的运算量。在现有的速度场简化方法中, 拓扑简化是一种非常优秀的方法。

速度矢量场的拓扑简化方法最早由 Helman 和 Hesselink^[6] 提出。拓扑简化方法的核心思想是将速度场的坐标点分成两类: 速度矢量非 0 的常点 (regular point) 和速度矢量为 0 的奇点 (singular point), 并通过奇点和连接奇点的流线来描述速度场的形状、结构以及演变过程。因为奇点的数量远远少于原始数据点的数目, 所以拓扑简化方法能够大大减少描述速度场所需要的数据量。在计算流体力学中, 这一方法被广泛应用于流体速度场数据的简化和可视化。下面简要介绍拓扑简化理论。

根据常微分方程理论^[7], 对于平滑的二维速度场, 速度矢量在邻域范围内可近似地用线性流方程

$$v(x, t) = \frac{dx}{dt} \approx A(t)x - b(t) \quad (1)$$

描述, 其中, $A(t)$ 为 2×2 的矩阵, $b(t)$ 为二维矢量, 描述了速度场中奇点的位置。如果 $A(t)$ 可逆, 在奇点附近, $v(x, t)$ 的局部形状由矩阵 $A(t)$ 的两个特征值决定。根据特征值的不同, 奇点按拓扑等价关系可以分为以下 6 类:

(1) 鞍点: A 的两个特征值为符号相反的实数, 见图 1(a)。

(2) 汇聚节点: A 的两个特征值均为正实数, 见图 1(b)。

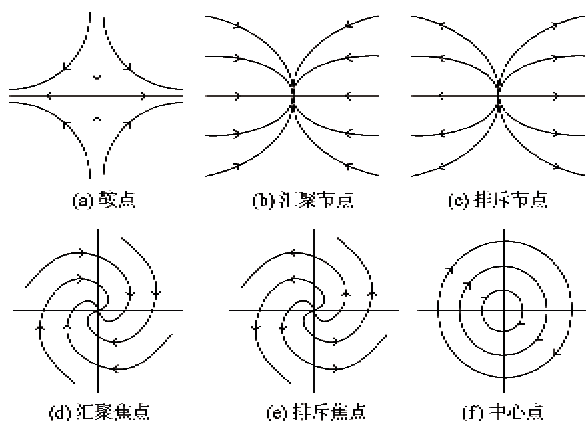


图1 奇点的拓扑分类

(3) 排斥节点: A 的两个特征值均为负实数, 见图 1(c)。

(4) 汇聚焦点: A 的两个特征值为实数部分为正的共轭复数, 见图 1(d)。

(5) 排斥焦点: A 的两个特征值为实数部分为负的共轭复数, 见图 1(e)。

(6) 中心点: A 的两个特征值为共轭虚数, 见图 1(f)。

拓扑等价性保证了当一个同胚映射 (具有连续逆映射的连续一一映射称为同胚映射) 作用在 A 所描述的线性流上得到的线性流, 其奇点的类型与 A 所描述的线性流奇点类型一致。换言之, 属于同一类型奇点的速度场之间, 可能在形态上存在较大差异, 在线性流参数的具体数值上也可能相差较大。但是奇点的分类对于这些差异并不敏感, 因此这种分类方法具有较好的鲁棒特性。图 2 显示了属于同一类型奇点的三种线性流。

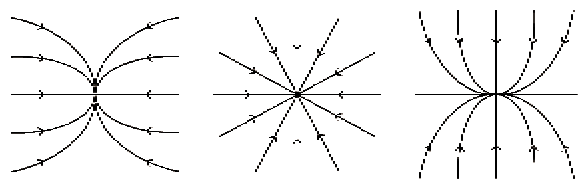


图2 具有相同类型奇点的三种线性流

2 基于速度场简化的人群行为分析

在视频监控领域中, 人群的行为分析主要指通过分析人群的运动变化, 对人群的当前状态分类, 并利用人群状态的变化进而检测出人群运动过程中发生的异常事件。在人群的行为分析中, 主要关注以下三大类行为事件:

(1) 人群速度变化事件

这类事件主要描述人群整体运动速度的变化过程。通常, 在发生异常事件时, 人群会快速疏散开来, 因此可以将其状态分为两类:

静止或行走: 人群以正常速度行走, 即 $v_{\text{平均}} \leq v_{\text{正常}}$;

奔跑: 人群以高速运动, 即 $v_{\text{平均}} > v_{\text{正常}}$ 。

(2) 人群宏观行为事件

人群宏观行为事件主要描述在一段时间内, 人群整体形态的变化, 其状态可以分为以下三类:

聚集: 人从四周向某一点聚拢, 形成密集人群;

散开: 人群从某一点向四周发散开来;

正常: 人群维持原有形状 (或密度)。

(3) 人群间行为事件

人群间行为事件主要用于检测在一段时间内, 人群数量的变化, 其状态可以分为以下三类:

分流:原本是一个人群分流成两个或多个;

合流:两个或多个个人群聚合成一个;

正常:人群维持原有状态, 数量上没有出现增减。

人群的异常行为检测往往通过上述状态的变化来定义。具体的异常行为定义根据监控场景的不同而不同。在人群的异常检测中, 通常要求不能漏报异常事件, 同时尽可能不要误报。因为人群的运动比较缓慢, 异常事件发生的具体时间较难确定, 所以人群的异常行为检测并不要求非常精确地预报事件发生的具体时间点, 可以存在一定的误差。

利用速度场简化的方法, 可以将上述三大类事件的检测与分类统一到一个框架之内。基于速度场简化的人群行为分析方法主要分为两个部分。第一部分改进了拓扑简化算法, 用线性流模型来描述人群的宏观运动, 并使用随机采样一致性 (random sample consensus, RANSAC) 算法^[8]提取出线性流模型中的参数, 用于人群整体异常行为和人群整体异常速度事件的检测; 第二部分使用速度场的聚类算法, 近似地估算出当前场景中的人群数目, 用于人群间异常行为的检测。

2.1 人群整体行为分析与人群平均运动速度提取

在人群整体的行为分析中, 人群的聚集和散开行为事件是两种经常出现的异常事件。这两种人群的运动行为可以通过上文介绍的拓扑简化方法, 用汇聚或排斥类型的奇点来描述。如图 3 所示的图像, 当人群散开时, 可以用一个位于发散中心的汇聚类型的奇点来描述人群的这一宏观行为。与其它方法相比, 基于奇点的拓扑分类具有更加坚实的数学理论基础, 并且这种分类蕴含在线性流模型之中, 因而不需要训练额外的分类器。值得一提的是, 在人群的行为分析中, 并不需要区分奇点分类中的节点与焦点, 因为二者描述的人群行为较为接近, 可以合并成一类。另外, 在人群运动的速度场中出现鞍点的情况较为罕见, 因此本文不考虑这种情况。

在传统的拓扑简化方法中, 奇点可以直接通过寻找速度为 0 的点定位, 而奇点临域的线性流模型参数也可以直接通过数值差分得到。但是, 这种方法对于人群运动的速度场不再适用, 因为: (1) 在人群图像中的速度场中, 存在着大量的速度为 0 的背景点, 传统的方法不能过滤这种背景点, 因而会影响

奇点的提取; (2) 人群运动的速度场存在一定的噪声, 比如人体手臂和腿部的前后摆动等; (3) 拓扑简化方法要求速度场是连续的, 但是人群在图像上运动的速度场, 是不连续的或是分段连续的。

为了解决上述问题, 首先选定一个速度阈值, 挑选出速度大于这一阈值的所有像素点, 将背景点过滤掉。然后, 使用 RANSAC 算法^[8]来定位奇点和估计线性流模型的参数, 并对当前图像速度场进行拓扑分类。RANSAC 算法是一种在强干扰环境中估计模型参数的有效方法。

由 RANSAC 算法得到的线性流参数 A 特征值, 可以判断当前速度场是否发生全局异常。如果两个特征值的实数部分都大于 0 并且二者之比小于一个阈值, 则人群当前运动为散开异常; 如果两个特征值的实数部分均小于 0 并且两者之比小于一个阈值, 则人群当前运动为聚集异常。

最后, 如果当前图像速度场分类结果属于某种异常, 还需要计算图像中出现这一异常事件的概率, 以降低误报率。如果当前图像中发生某一种全局异常, 那么图像中多数点的运动都会反映出这种异常。因此可以近似地用所有按照某种全局异常运动的点的数目 N_{abnormal} 占有有效运动点 N_{total} 的比例来表示当前图像是否发生异常的概率, 即

$$P(\text{abnormal}) = \frac{N_{\text{abnormal}}}{N_{\text{total}}} \quad (1)$$

而某一特定的有效运动点的运动是否属于异常可以由其光流速度矢量与线性流估计结果的差异决定。如果这一差异较小, 则认为这一点的运动属于异常运动; 如果这一差异较大, 则认为这一点的运动属于正常运动。

具体的密集人群宏观行为检测算法步骤如下:

步骤 1: 利用光流提取图像中像素点的速度场;

步骤 2: 滤除速度场中速度幅值小于某一阈值的点;

步骤 3: 根据线性流参数 A 的特征值进行判断当前速度场是否发生异常;

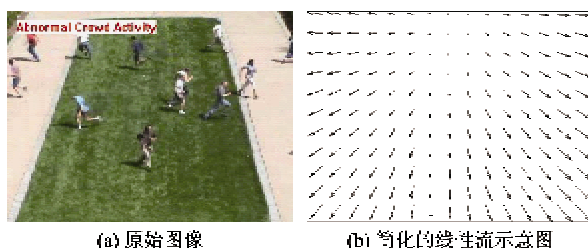


图 3 人群原始图像与对应的线性流

步骤 4:如果步骤 3 检测结果发生异常,计算异常事件的概率;否则发生异常事件的概率为 0。

步骤 5:对步骤 4 的检测结果进行平滑滤波,如果平滑后的结果大于一个事先设定的阈值,则报告发生相应事件。

线性流模型除了能够描述人群的宏观行为外,还可以描述人群运动的平均速度。人群运动的平均速度可以近似地认为与线性流模型中的矩阵 A 的欧式范数存在正比关系,即

$$v_{\text{人群}} \propto \|A\|_2 = \sqrt{\lambda_{\max}(A^T A)} \quad (2)$$

其中, $\lambda_{\max}$ 表示矩阵 A 的最大特征值。这样,通过比较这一范数与预先设定的阈值之间的大小关系,即可确定图像中是否出现异常速度事件。

2.2 人群间行为分析

基于拓扑简化的行为分析方法虽然能够较好地检测人群整体运动的行为,但是在其处理过程中并没有考虑人群的边界,因而不能用于检测人群之间发生的各种异常事件,所以需要额外的算法来检测人群间出现的异常。

考虑到一般情况下,同一人群的运动模式较为相似,而不同人群的运动差异较大,因而可以将作相似运动且距离较近的像素点划分为一个聚类,近似地估计出当前场景中的人群数目。然后,利用这一信息检测人群间异常事件的发生。

与传统的聚类算法不同,描述人群运动的速度场函数聚类涉及的数值有两组:速度场的平面坐标 x 和速度矢量 v 。并且,这两组数值在描述不同点之间相似程度时的侧重点并不相同:平面坐标更关心平面距离的远近,因而通常用欧式距离描述^[9];而速度矢量更关心速度方向之间的差异,通常用速度矢量的归一化内积描述^[10]。对归一化内积稍作处理,可以得到速度场数据之间的相似度函数

$$d(x_1, v_1, x_2, v_2) = \|x_1 - x_2\| \left(1 - \frac{v_1 \cdot v_2}{\|v_1\| \|v_2\|}\right) \quad (3)$$

不同点的速度和坐标越相近,其相似度函数越接近于 0。

利用这一相似度函数,并参照传统 k 平均聚类算法,可以构建出速度场 k 平均聚类算法。具体的速度场聚类算法如下:

步骤 1:随机指定 k 个聚类中心的坐标位置;

步骤 2:对每个数据点,指定其聚类编号为距其最近的聚类中心的编号;

步骤 3:初始化 k 个聚类的平均速度;

步骤 4:对每个数据点,指定其编号为与其速度数据相似度最高的聚类编号;

步骤 5:对每个聚类,更新聚类的中心坐标和平均速度矢量;

步骤 6:如果存在相似度较高的聚类,将其合并并更新当前聚类数目;

步骤 7:如果聚类编号发生变化,跳转回步骤 4;如果没有发生变化,返回聚类编号。

图 4 显示了在人工合成的速度场上使用速度场聚类算法的效果,图像平面按照速度方向的不同被分为 11 个不同的区域。

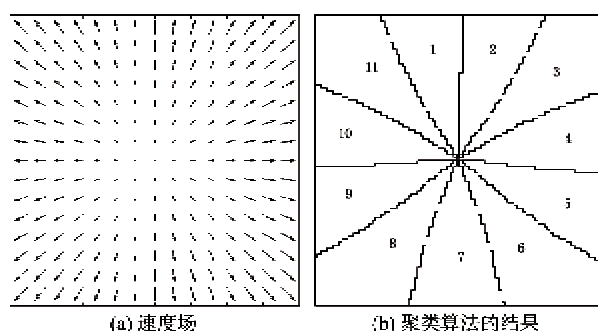


图 4 聚类算法效果图

3 实验与结果

3.1 人群整体异常行为检测实验

人群整体异常行为检测实验使用明尼苏达大学的密集人群行为测试集^[11]。此测试集包含 3 个场景共 11 组测试序列,每一组序列开始时人群处于正常状态,然后发生异常事件,人群以不同的方式散开。在本实验中,异常事件被定义为人群突然散开。图 5 显示了在正常情况和发生异常情况时的图像。



图 5 人群整体异常行为检测序列图像

在实验中,图像被划分成 8×8 的方格区域,并估计在每一个方格区域运动的速度矢量。人群速度场的提取使用文献[12]中的光流算法。为了比较本文算法的性能,使用社会力算法^[2]和混沌不变量算法^[3]作为对比。

图6显示在发生异常情况时的原始光流速度场和由线性流方程重构的速度场。可以看出,重构的速度场简化了原始光流速度场中的细节,同时较好地反映了原始光流速度场中的全局特征。

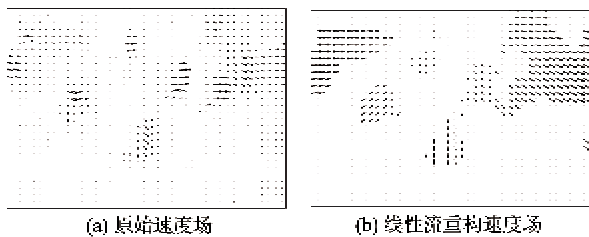
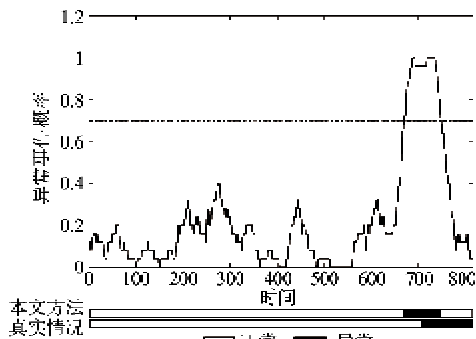
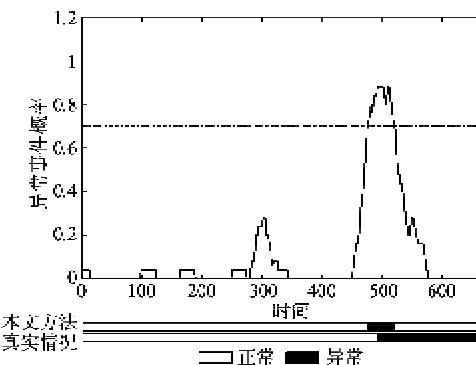


图6 速度场简化效果图

图7为人群宏观异常事件检测的检测结果。图中的横坐标为时间轴,纵坐标为人群发生发散事件这种异常情况的可能性,通过对代表当前人群整体状态的数值进行窗口长度为25的平均值滤波得到,



(a) 序列2检测结果



(b) 序列8检测结果

图7 人群宏观行为检测结果

其数值越大越有可能发生异常事件。图中的虚线表示事先设定的阈值,如果发生异常事件的概率大于这一阈值,则认为发生异常事件。图7下方的带状图显示了利用速度场简化的事件检测结果与数据集自带的真实情况之间的比较,其中白色区域代表正常情况,黑色区域代表发生异常事件。可以看出,线性流的拓扑分类能够很好地将人群的正常活动与异常活动区分开来,能够较好地检测出异常事件。图8为社会力算法、混沌不变量算法和本文算法的性能比较。在三种方法中,社会力算法检测性能最差,而混沌不变量算法最好。本文算法性能虽略差于混沌不变量算法,但是相差不大。表1为这三种算法的各方面比较。在三种方法中,社会力算法和混沌不变量算法都需要借用机器学习理论学习人群运动的正常行为。这两种方法将人群运动的速度场变化看作一个随机过程,并使用多帧图像共同建模人群行为,所以训练和检测时间均较长,异常检测的时间精度也较低;而本文方法利用拓扑分类理论对速度场分类,所以不需要训练,且时间精度较高。

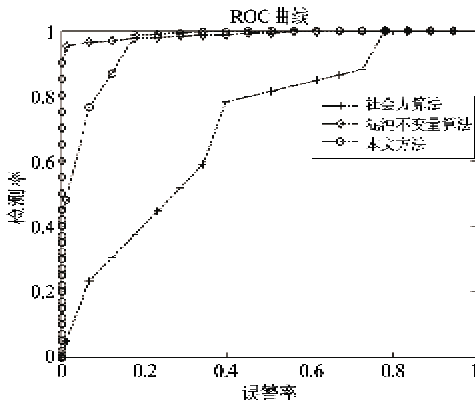


图8 不同方法性能比较图

表1 宏观行为检测中不同算法的比较

	训练时间 (min)	检测时间 (s)	时间精度 (帧)	异常检测 准确度
社会力算法	32	5.3	10	一般
混沌不变量 算法	47	30.4	10	最好
本文方法	不需要	4.4	1	较好

3.2 人群速度异常事件检测实验

人群速度异常行为检测测试集使用 PETS 2009 数据集^[13]中的一组人群序列。在这组序列中,人群首先从右边走入,然后开始奔跑出画面。之后,人群从左边走入,从某一时刻开始,奔跑出画面。在这一

实验中,异常事件定义为人群平均运动速度较快。图9显示了部分测试序列的截图。为了比较算法的性能,使用HOG算法^[14]作为对比。

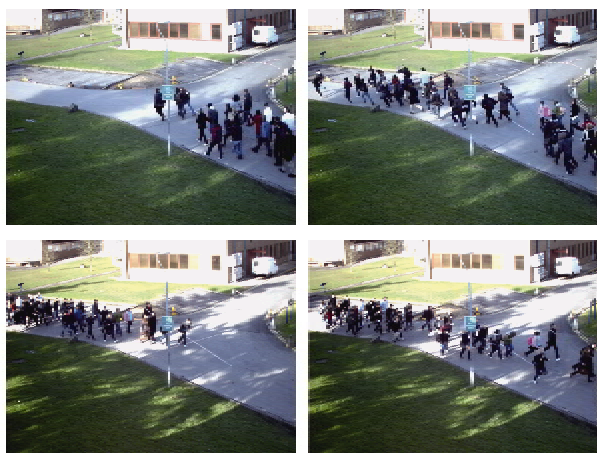


图9 速度异常检测序列图像

图10为人群异常速度检测的结果。其中图中实线代表线性流方程中矩阵A的欧式范数,经过窗口长度为25的平均值滤波处理;虚线代表预先设定的阈值。如果矩阵的范数大于事先设定的阈值,则认为出现了异常速度事件。图10下方的带状图显示了本文方法检测结果与手工标定的真实情况的比较。可以看出,线性流模型的欧式范数基本反映了人群的平均运动速度,利用这一方法可以可靠的检测人群运动中的速度异常。

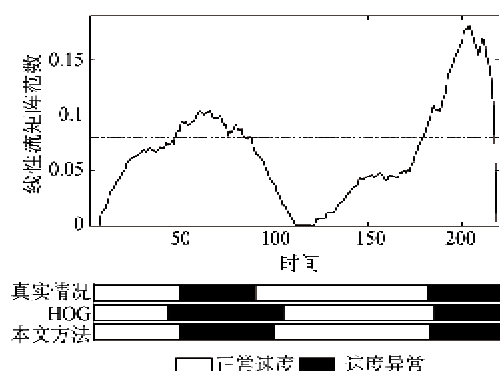


图10 速度异常检测结果

表2显示了HOG算法与本文方法在速度异常检测中的性能比较。其中,表格中 TP 指算法检测

表2 速度异常与人群间异常性能比较

	速度异常			人群间异常		
	TP	FP	敏感度	TP	FP	敏感度
HOG	202	18	0.92	100	30	0.77
本文方法	205	15	0.93	106	24	0.82

结果中与真实结果一致的帧数, FP 指与真实结果不一样的帧数,敏感度定义为 $TP/(TP+FP)$ 。由表可见,与HOG算法相比,使用速度场简化的方法更加精确。

3.3 人群间异常行为检测实验

人群间行为检测测试集使用PETS 2009数据集^[12]中的一组人群序列。这组序列描述的场景是开始时人群从左方进入画面,到达岔路口,分叉成两股人流,继续前进。在这一实验中,异常事件被定义为人群出现分流。与速度异常检测实验类似,使用HOG算法作为对照。

图11为测试序列第48帧和第127帧图像和聚类分析的结果。图12为人群间异常行为检测结果。图中实线代表平均聚类个数,由每帧图像的聚类个数经过窗口长度为25的平均值滤波处理得到。图像下方带状图代表手工标定的真实情况。可以看出,当人群的分流开始后,平均聚类个数显著提升,使用速度场聚类方法检测结果与真实结果十分接近。表2中显示了HOG算法与本文方法在人群间

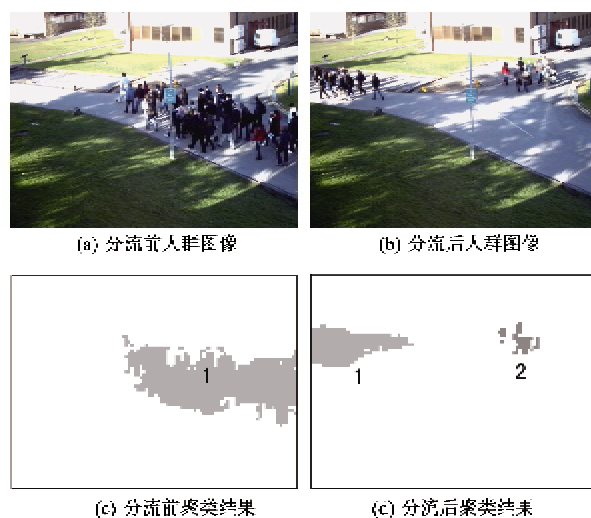


图11 人群间异常行为图像和分析结果

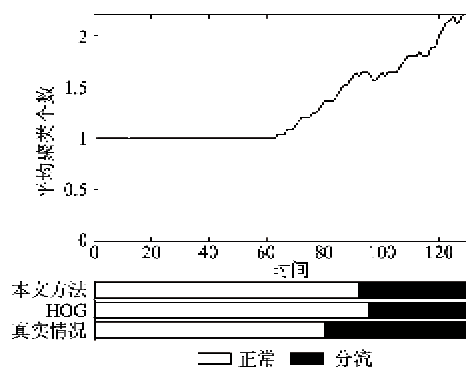


图12 人群间异常行为检测结果

异常检测中的性能比较。本文方法的检测精度比 HOG 算法高出 5%。

以上所有算法均使用 MATLAB 实现,实验环境为一台具有 Xeon 2.66G CPU 的 Linux 服务器。光流速度场提取平均耗时 4.3s,速度场的拓扑简化平均耗时 0.11s,速度场的聚类分析平均耗时 0.026s。

4 结 论

本文提出一种基于速度场简化的人群行为分析方法。速度场简化的人群行为分析方法分为两个部分,第一部分利用拓扑简化提取出人群运动的主要动力学特征,用于人群整体异常行为和人群整体异常速度事件的检测;第二部分利用聚类分析技术,计算出人群数目的变化,用于人群间异常行为的检测。由于速度场简化方法能够从大量的局部数据中提取出全局信息,因而具有较好的鲁棒特性,降低了异常事件检测的误报率。在实验中,以两组公开的人群行为分析序列作为测试对象,比较了三类异常事件的检测结果与真实情况的差异。实验结果表明,使用速度场简化方法能够有效地检测出各种人群整体运动过程中发生的异常事件。

参考文献

- [1] Ihaddadene N, Djeraba C. Real-time crowd motion analysis. In: Proceedings of 19th International Conference on Pattern Recognition, Tampa, USA, 2008. 1-4
- [2] Mehran R, Oyama A, Shah M. Abnormal crowd behavior detection using social force model. In: Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, USA, 2009. 935-942
- [3] Wu S, Moore B E, Shah M. Chaotic invariants of Lagrangian particle trajectories for anomaly detection in crowded scenes. In: Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010. 2054-2060
- [4] Mahadevan V, Li W, Bhalodi V, et al. Anomaly detection in crowded scenes. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010. 1975-1981
- [5] Kratz L, Nishino K. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In: Proceedings of 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Miami, USA, 2009. 1446-1453
- [6] Helman J, Hesselink L. Visualizing vector field topology in fluid flows. *IEEE Computer Graphics and Applications*, 1991, 11(3):36-46
- [7] Arnold L. Ordinary Differential Equations. New York: Springer, 1992
- [8] Fischler M A, Bolles R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of ACM*, 1981, 24(6): 381-395
- [9] Duda R O, Hart P E, Stork D G. Pattern Classification. Wiley-Interscience, 2001
- [10] Dhillon I S, Modha D S. Concept decompositions for large sparse text data using clustering. *Machine Learning*, 2001, 42(1): 143-175
- [11] Unusual crowd activity dataset of the University of Minnesota. Available at: <http://mha.cs.umn.edu/movies/crowdactivity-all.avi>
- [12] Black M J, Anandan P. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 1996, 63(1):75-104
- [13] PETS 2009 Dataset. Available at: <http://www.cvg.rdg.ac.uk/PETS2009/>
- [14] Garate C, Bilinsky P, Bremond F. Crowd event recognition using HOG tracker. In: Proceedings of the 12th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS-Winter), Snowbird, USA, 2009. 1-6

Crowd behavior analysis based on velocity field simplification

Li Nan^{***}, Zhang Zhimin^{*}

(^{*} Integration Application Center, Institute of Computing Technology, Chinese Academic of Sciences, Beijing 100190)

(^{**} Graduate University of Chinese Academic of Sciences, Beijing 100049)

Abstract

To solve the problem that crowd abnormal behavior detection is difficult to be achieved during video surveillance, this paper proposes a crowd behavior analysis method based on velocity field simplification. The method of velocity field simplification uses topological simplification and clustering analysis to extract the information of a crowd, and then uses the information to detect the anomaly in the crowd. The biggest advantage of velocity field simplification method over others is that it does not need the time-consuming training. The proposed method was tested on the detection of three categories of abnormality: crowd formation/dispersal, changes in crowd speed and crowd splitting/merging. The experiments indicate that the proposed method can efficiently extract the main behaviors of crowd motion and the detection results are similar to the ground truth.

Key words: video surveillance, crowd behavior analysis, velocity field simplification, topological classification, clustering algorithm