

元胞驻留视频群体状态预报方法

叶志鹏, 赵巍, 刘鹏, 唐降龙

(哈尔滨工业大学 计算机科学与技术学院, 150001 哈尔滨)

摘要: 视频场景人群状态预报是防止公共场所安全事故的一种重要措施. 为此提出了一种视频场景群体状态预报模型. 该模型使用格内驻留元胞自动机(inner-grid parking cellular automata, IPCA)模型预报视频中群体状态. 首先, 通过光流法进行运动跟踪作为 IPCA 的输入特征对场景进行建模, 并根据特征点的运动状态自适应地调整其在元胞格内的驻留时间以提高预报精度; 其次, 通过分析 IPCA 模型的预测结果判断群体微团间的相互作用得到视频帧的状态信息, 从而达到对场景中人群状态进行预报的目的; 最后, 利用当前场景的状态信息对该场景的预报输出进行反馈校正, 使模型能够及时地反映场景中群体状态的变化, 该模型不依赖于具体的视频帧, 以微团结构的碰撞作为异常检测依据. 实验结果表明, 与传统的检测方法相比, 该预报模型能够提前预报视频场景中的异常状态, 并具有较好的准确性.

关键词: 视频处理; 元胞自动机; 格内驻留; 状态检测和反馈; 状态预报

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 0367-6234(2014)09-0025-06

Inner-cellular automata method for video pedestrian state forecasting

YE Zhipeng, ZHAO Wei, LIU Peng, TANG Xianglong

(School of Computer Science and Technology, Harbin Institute of Technology, 150001 Harbin, China)

Abstract: Forecasting pedestrian states with the help of video analysis is a significant mean to prevent accidents in public places. An improved Inner-grid Parking Cellular Automata (IPCA) model is proposed to forecast pedestrian state. Motion tracking and scenario modeling are firstly achieved by optical flow method, and then motion states of pedestrian are used to adjust parking time in cellular adaptively to improve forecasting accuracy. Here, the state of an incoming video frame is acquired by analyzing the interaction between micelles, which is provided by the forecasting result of IPCA. After that, a feedback algorithm is employed to revise the forecasting result so that the model could reflect precisely the change of pedestrian state in the scenario. A criterion is also proposed to judge the abnormal state of a video frame, which is collision between two micelles. Compared to other traditional detection methods, IPCA based model has a good ability in predicting the abnormal state ahead of time and obtaining better accuracy.

Keywords: video processing; cellular automata; inner-grid parking; state detection and feedback; state forecasting

近年来随着社会的快速发展,大规模人员聚集和运动的情况和场所大量增加,公共场所的大规模安全事故屡有发生,造成了人员伤亡和财产损失,给活动的组织管理者和参与者带来明显的安全挑战.对人群状态预报可以提前获知人群运

动状况的变化,提高应急措施的效率和效果,将群体危机消除在萌芽状态,对防止安全事故的发生具有重要的意义.

近来,国内外学者通过元胞自动机 (cellular automata, CA) 模型,研究特定场景下大规模群体疏散问题. Burstedde 等^[1]首次提出了基于动态参数的地场 CA 模型; Yamamoto 等^[2]对 CA 进行了改进,提出了实数编码的 CA 模型,其最主要的改进是将速度的范围从整数扩展到了实数集合,从而增强了 CA 在描述群体运动中的灵活性和适应

收稿日期: 2013-10-21.

基金项目: 国家自然科学基金资助项目 (61171184, 61201309).

作者简介: 叶志鹏 (1989—), 男, 博士研究生;

唐降龙 (1960—), 男, 教授, 博士生导师.

通信作者: 唐降龙, tangxl@hit.edu.cn.

性.Hao 等^[3]提出了用于描述双向行人流的 CA 模型,该模型使用动态参数形式,将行人的转移收益分解为各参数之和,提出了行人视场的概念,并描述了实际情况下行人随大流和躲避拥挤区域的心理状态.Wang 等^[4]提出了一个用于描述室内行人行为的 CA 模型,将行人的运动分为 3 种形式:向目标移动、碰撞和躲避拥挤以及组团,并研究各种行为对应的代价函数对行人转移收益和路径选择的影响.这些描述群体状态或运动的 CA 模型侧重于描述群体在场景中的分布和速度间的关系,存在描述视频特征困难、不能结合视频特征提供 CA 模型参数修正方法,导致 CA 模型缺乏对群体状态的预报能力.

针对以上 CA 模型的局限性,本文将群体中的微团结构作为研究对象.微团结构是群体中具有相似运动特征的个体的集合.除了自主运动以外,微团之间的相互作用是影响群体状态的主要因素.本文研究认为微团中个体自主运动的的目的性和微团之间相互作用的随机性导致了群体成为确定性和随机性并存的系统.本文提出了格内驻留元胞自动机(inner-grid parking CA, IPCA)群体状态预报模型.首先通过光流法获取场景中群体的运动状态参数,然后将群体运动速度矢量作为模型的输入对 IPCA 模型进行初始化并根据特征点的速度设置其在元胞格内的驻留时间来完成场景的建模.运动特征点的驻留时间在特征点移出场景前会一直存在并根据不同的检测结果自适应地调整.本文提出了基于检测-校正的群体预报模型,检测视频中微团之间的相互作用,并将检测结果作为校正模型参数的依据,对群体状态进行预报.实验结果表明,本文提出的检测方法更具灵活性,检测-校正的参数调节使模型具有自适应能力,取得了良好的预报效果.

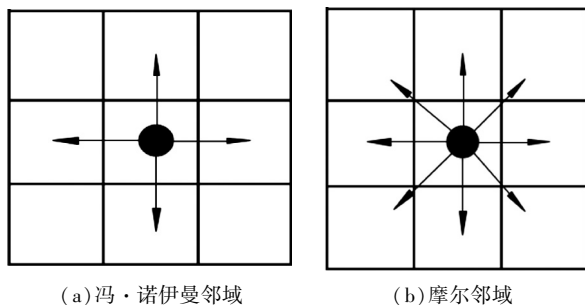
1 CA 模型

CA 是一种通用的建模方法,也是动力学系统中有关秩序、混沌、非对称等系统整体行为与复杂现象的一种有效研究手段.一个标准的元胞自动机 A 由元胞、元胞状态、邻域和状态更新规则构成.用数学表示为^[5]

$$A = (L, d, S, N, f).$$

式中: L 为元胞空间; d 为元胞自动机内元胞空间的维数; S 为元胞有限的、离散的状态集合; N 为某个邻域内所有元胞的集合; f 为局部映射或局部规则.在二维 CA 中常见的邻域有冯·诺伊曼邻域和摩尔邻域.图 1 中实心圆代表特征点当前位置,

在每个时间步特征点可以通过计算其周围 8 个元胞格的转移收益进行特定方向的移动,如图 1 所示.



(a) 冯·诺伊曼邻域

(b) 摩尔邻域

图 1 二维 CA 的邻域

对 CA 中元胞转移收益的计算采用动态参数模型^[3],如

$$P_{ij} = D_{ij} + E_{ij} + F_{ij} + C_{ij}.$$

式中 P_{ij} 为目标向其邻域内第 i 行第 j 列的元胞格移动的转移收益值.每个参数的意义及取值如下:

1) 方向参数 D_{ij} 反映了行人根据其当前运动方向选择下一个方向的代价为

$$D_{ij} = \begin{cases} 1, & \text{正上方元胞;} \\ a, & \text{左上、右上元胞;} \\ 0, & \text{当前行元胞;} \\ -a, & \text{左下、右下元胞;} \\ -1, & \text{正下方元胞.} \end{cases}$$

其中 $a \in (0, 1)$ 且 $a \in R$. 其意义为目标优先移向其正前方的元胞.

2) 空参数 E_{ij} 反映元胞格是否被占用为

$$E_{ij} = \begin{cases} 1, & \text{元胞为空;} \\ 0, & \text{行人所在当前元胞;} \\ -1, & \text{元胞被其他行人占据.} \end{cases}$$

3) 前向参数 F_{ij} 反映行人躲避拥挤的意愿为

$$F_{ij} = \frac{S_1 - S_2}{R}.$$

在现实中通常表现为行人不愿向人多可能造成拥挤的地区运动.其中: S_1 为市场中空元胞数目; S_2 为被占用的元胞数目; R 为行人视场大小,即特征点视场范围内元胞数目.

4) 内容参数 C_{ij} 反映行人希望随着和自己同方向行进的人一同行走的意愿,在现实中常表现为“随大流”为

$$C_{ij} = \frac{S_3 - S_4}{R}.$$

式中: S_3 为空元胞和场景中与当前行人运动方向相容的行人数之和; S_4 为场景中与当前行人运动方向不相容的行人数之和; R 为行人视场大小.本文将 CA 相容方向集定义为: {左上, 左下, 水平

向左}、{右上,右下,水平向右}、{左上,垂直向上,右上}和{左下,垂直向下,右下}.同时,各方向和自身也是相容的.

2 IPCA 的视频场景群体状态预报方法

本文将视频场景划分为多个正方形元胞格.场景的建模过程将光流法获取的每个运动特征点依据其在视频帧中的位置放置于相应的元胞格中,对每个特征点以元胞格的大小与速度幅值之比作为其在该元胞格内的驻留时间,表示特征点经过多少时间后可以移出该元胞格.本文利用 IPCA 的转移规则描述群体未来的运动情况.

2.1 IPCA 建模

视频场景中所有特征点组成的集和 P 定义为

$$P = \{P_i\}, P_i = \langle x_i, y_i, v_i \rangle$$

对所有 $P_i \in P$, 将其位置 (x_i, y_i) 和速度矢量 v_i 输入 IPCA, 构建初始状态 S_0 ; 依据预先设定的转移规则驱动模型预测 S_1 到 S_n 的状态.其过程如图 2 所示.

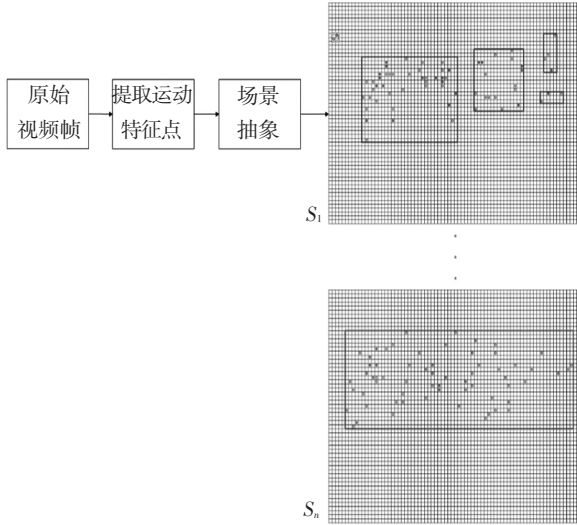


图2 IPCA 建模过程

2.1.1 IPCA 参数计算方法

针对以往 CA 的研究不能很好地描述视频场景的不足,本文通过对行人视场、消除建模误差、参数计算和获取方面做了相应的改进,得到 IPCA 模型.

1) 行人的视场形状由矩形改为三角形,这更加符合人类视野的扇形描述,如图 3 所示.

2) 引入特征点驻留时间 $t \propto \frac{1}{v}$, v 为其运动速度幅值.特征点是否可以移动取决于其已等待的时间 t_w 为

$$P_{ij} = \begin{cases} -\infty, & t_w < t; \\ \max P_{ij}, & t_w = t. \end{cases}$$

若 $t_w < t$ 则当前特征点对其他元胞的转移收益为 $-\infty$, 特征点将停留在当前元胞, $t_w + 1$, 否则依据动态参数模型计算并选择代价最大的元胞进行移动.驻留时间 t 的引入可以显著提高 CA 模型对真实场景描述的准确性.

3) 前向参数 F_{ij} 计算方法为

$$F_{ij} = \begin{cases} \frac{S_1 - S_2}{R}, & R_c = R; \\ \max P_{ij}, & R_c < R. \end{cases}$$

式中: R_c 为目标当前视场大小; R 为图 2 所示的视场大小.若一个区域内空元胞较多,则 F_{ij} 较大,使特征点倾向于向该区域移动;反之会使特征点趋于躲避该区域,因此 F_{ij} 描述了行人躲避拥挤意愿的强烈程度.根据视场大小计算 F_{ij} 可以避免因视场不足由不完整数据导致的计算误差,同时也简化了计算过程.

4) 通过对视频帧的光流分析和计算得到运动特征点的速度幅值和方向,具有更好的客观性.

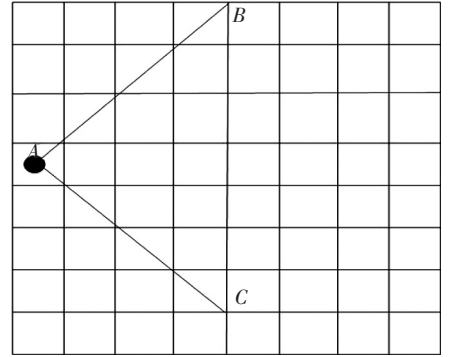


图3 三角形视场示意

2.1.2 IPCA 转移规则

当前元胞中特征点在下一时刻的转移状态依据其近邻元胞的当前状态,其合理程度直接影响了 CA 描述结果的准确性.本文提出的转移规则如下:

1) 特征点倾向于向转移收益最大的元胞格移动;

2) 当发生冲突时,如多个特征点试图同时移向一个元胞,则随机选择一个特征点进行移动,其他特征点保持原地不动;

3) 对每个特征点的驻留时间 t ,若当前特征点已驻留等待的时间小于 t ,则该特征点停留在所在元胞格不动,否则根据规则 1) 进行移动,移动后其 t_w 置零重新进入等待状态;

4) 日常生活中,人们总是倾向于最快地到达目的地,因此本文假设特征点不执行后退移动.

由转移规则可以看出,IPCA 在保持了原始模型简洁性的同时仍能够模拟复杂的现象.

本文使用 K -means 聚类算法将场景中不同位置的运动特征点划分为不同的微团. 因此对场景中行人的异常状态的检测转换为对每个微团的状态进行检测. 异常判别标准为微团间发生碰撞或重叠, 以表达行人间的相互撞击和推挤.

2.2 群体状态检测

微团是由空间内相邻运动特征点所组成的结构. 某一时刻场景中共存在两种状态的微团: 正常微团和异常微团. 本文认为处于正常状态的微团结构不与其他任何微团重叠; 反之, 则为异常微团. 此时对于高强度行人流来说, 人群将变得非常拥挤接近堵塞, 人员的聚集风险将大大增加^[6]. 本文采用一种基于微团碰撞的异常状态检测 (micell state detection, MSD) 方法. 将微团分为上、下、左、右 4 个子微团. 对每个子微团, 若与其他邻近微团发生重叠, 则该微团当前处于异常状态, 否则处于正常状态.

通常情况下光流法提取的特征点不可避免的受到来自环境中噪声的影响而产生错误的结果, 从而影响异常检测结果的准确性. 因此, 在最终的异常判断中, 通过设置异常帧在检测半径内所占百分比 s 对进行结果平滑, s 又称为敏感系数. 对于同一帧视频场景的预报结果, 若检测到的异常次数 t 与预报半径 m 满足

$$t > s \times m$$

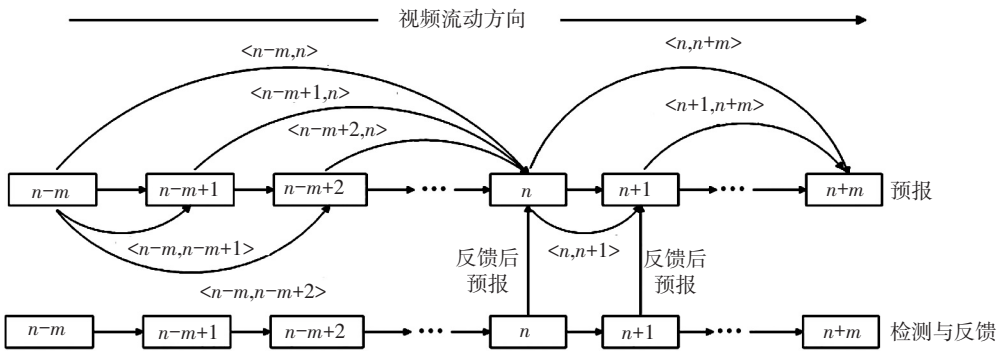


图 4 人群状态预报

对视频场景特征点集中任意两点 $X_1(x_1, y_1)$ 、 $X_2(x_2, y_2)$, 若其欧式距离 $\rho(x, y) \leq T$, 则认为特征点 X_1 与 X_2 是相邻的. 定义两个集合的和运算为

$$P_n \cup P_{n+1} = M \cup UM.$$

式中: M 为合并两帧中所有近邻点集组成的集合; UM 为非近邻点组成的集合. 特别地, 若 $P_n = \phi$, 则 $P_n \oplus P_{n+1} = P_{n+1}$, 初始状态 $\text{Result} = \phi$, P_0 为第 0 帧视频的特征点集合.

反馈 $\oplus$ 过程如下:

$\text{Result} = \{ \}$;

for $i = n - m$ to m

则判断当前帧为异常帧.

2.3 群体状态预报方法

本文提出的群体状态预报方法通过获取一定时间内的历史状态信息, 利用 IPCA 模型对处于预报半径 m 内的未来某一时刻的群体状态进行预测, 根据检测结果得到关于群体未来状态的预报值. 在预报时刻到来时, 利用对当前场景的特征检测结果反馈更新模型对当前时刻的预报输出. 其过程如

$$S(n+k) = F \left(\sum_{i=1}^m F(n-i, n) \oplus D(n), n+k \right).$$

式中: $S(n+k)$ 为第 $n+k$ 个时间步场景的状态, $k \in [1, m]$ 且 $k \in \mathbb{Z}$; $\sum_{i=1}^m F(n-i, n)$ 为模型从时间步 $[n-m, n-1]$ 对第 n 个时间步的状态预测的叠加; $D(n)$ 为模型对第 n 个时间步场景状态的检测结果; $\oplus$ 为反馈过程, 因此, $\sum_{i=1}^m F(n-i, n) \oplus D(n)$ 为模型对第 n 个时间步场景状态的输出结果; $F \left(\sum_{i=1}^m F(n-i, n) \oplus D(n), n+k \right)$ 为在第 n 个时间步输出结果的基础上对第 $n+k$ 个时间步场景状态的预测结果, 如图 4 所示. 其中 $\langle n-m, n-m+1 \rangle$ 为 $n-m$ 帧对 $n-m+1$ 帧状态的预测.

$$\text{Result} = \text{Result} \oplus P_i$$

对近邻点集 $M \subseteq \text{Result}$, 执行 1 次特征点集更新.

end for

return Result;

其中, 特征点集更新算法如下:

输入: $p_i = (x_i, y_i, \mathbf{v}_i) \in P_n$, $p_j = (x_j, y_j, \mathbf{v}_j) \in P_{n+1}$

If p_i 与 p_j 不相邻, $\text{Result} = \text{Result} \cup P_j$, 结束;

Else If $d_i = d_j$, $//d_i$ 为 $\mathbf{v}_i$ 的方向

$$x_i = \frac{x_i + x_j}{2}, y_i = \frac{y_i + y_j}{2},$$
$$v_i = \frac{v_i + v_j}{2}, // v_i \text{ 和为 } v_i \text{ 的幅值};$$

Else $v_i = v_i + v_j$, $P_{n+1} = P_{n+1} - (x_j, y_j, v_j)$. 若 $P_{n+1} = \phi$, 则

Result = Result $\cup P_i$ 结束, 否则 $j = j + 1$, 继续迭代

输出: Result

2.4 自适应预报参数修正

视频场景通常是不断变化的. CA 模型的预测输出值会随时间推移与实际情况有较大偏差; 同时, 光流法提取的特征点由于摄像机抖动等因素存在误差, 这两个因素会影响 IPCA 预报的准确度. 本文提出方法的回馈阶段主要解决以上问题.

利用当前状态的检测结果更新上一状态的预报值, 如

$$F(x_n) = D(x_n) + cF_{n-1}.$$

式中: c 为更新速率; F_{n-1} 为前 $n - 1$ 帧对当前帧的预报初值, 模型以此输出值作为预报起点, 向后预报一定时间步内场景的状态, 该时间长度作为预报半径 m . 因此, 对于变化较为剧烈的场景, 较小的 c 可以令模型的输出值更趋向于当前检测结果, 使模型增强模型容忍场景变化的能力; 对于变化较为平缓和连续的场景, 较大的 c 可以更多的保留场景中一定时间步内的历史信息, 使模型的更新过程更为准确.

2.5 IPCA 预报方法的特点

本文提出的 IPCA 视频场景群体状态预报方法具有两个特点:

1) 使用 IPCA 模型作为预报的基础描述真实视频场景. 与传统的 CA 模型相比, IPCA 模型能够

接受运动跟踪结果自适应地设置其初始状态; 通过引入元胞格内驻留时间, 使对特征点运动的描述更加准确, 降低了场景的建模误差从而能使模型更好地描述视频场景的状态;

2) 不断利用场景的最新信息对模型参数进行反馈调整以提高预报的准确度. 反馈过程使模型能够及时地反映场景中群体的状态变化.

传统的 CA 模型将所有特征点的速度均设置为一个固定值 1, 没有考虑不同人群行进速度不同的特点. 本文针对 CA 的局限性, 引入了与人群行进速度密切相关的元胞格内驻留时间, 进而改进了传统 CA 的转移规则, 提出了 IPCA 模型. IPCA 模型通过对场景人群状态的建模, 根据人群行进速度各不相同的特点, 将元胞内特征点移动速度的取值范围和方式进行建模, 并通过驻留时间消除行进中的位置误差, 从而达到预报群体状态的效果.

3 结果与分析

本文使用了 3 个数据集: PETS09 (帧尺寸 768 $\times$ 576, 正常情形)、WEB Dataset (帧尺寸 1 280 $\times$ 720, 异常情形) 以及作者在校园内拍摄的视频 Self Dataset (帧尺寸 1 280 $\times$ 720, 正常异常情形交替出现), 其中, UCF 的 WEB Dataset 中同方向奔跑的群体状态识别是具有挑战性的. 实验中取反馈更新速率 $c = 0.7$, 敏感系数 $s = 0.5$, 检测半径 $r = 10$, 预报半径 $m = 15$, 前向参数 D_{ij} 中 $a = 0.7$. 本文将所提出的基于 IPCA 的检测方法与近几年的主流检测方法 (文献 [7-12]) 进行了比较. 实验结果如表 1 所示. 表 2 为 CA 与 IPCA 作为预报模型, 采用 MSD 检测方法的结果. 表中: 数据左侧为漏报率; 右侧为误报率.

表 1 各方法比较结果 %

检测方法	PETS09 (406 帧)		SelfDataset (212 帧)		WebDataset (574 帧)	
文献 [7]	25.0	14.5	7.1	4.7	31.6	23.5
文献 [8]	18.9	13.8	22.6	11.3	25.4	30.9
文献 [9]	13.8	12.6	10.7	14.5	28.3	28.7
文献 [10]	17.5	13.2	16.4	12.9	22.7	30.6
文献 [11]	21.4	15.3	15.2	13.8	20.1	34.5
文献 [12]	12.7	10.4	8.6	9.4	18.6	22.9
IPCA+MSD	10.6	6.6	2.3	3.5	16.1	21.5

表 2 IPCA 与 CA 的预报精度比较 %

检测方法	PETS09 (406 帧)		SelfDataset (212 帧)		WebDataset (574 帧)	
CA+MSD	23.8	11.9	12.7	10.4	26.5	29.4
IPCA+MSD	10.6	6.6	2.3	3.5	16.1	21.5

从表 1 可以看出,与其他检测方法相比,IPCA 与 MSD 的预报方法具有最优性能.这是由于以往的方法局限于当前场景异常状态的检测,忽视了异常状态的发生是一个渐进过程,而本文模型不仅获取当前场景的状态,而且利用之前一段时间内场景的状态作为预报的知识参与模型的反馈修正过程后再进行状态检测,因此具有较好效果.

从表 2 结果可以看出,在采用相同检测方法的前提下,本文提出的 IPCA 与传统 CA 相比具有较高的预报精度.与传统 CA 相比,IPCA 的格内驻留特点不仅结合实际视频场景改进了 CA 的转移规则与元胞格参数的计算,同时根据人群行进速度不同的特点引入驻留时间使群体状态的建模更为准确,从而提高了整体模型针对群体状态预报的精度.

4 结 论

1)通过结合光流法获取的运动特征,自适应地设置每个特征点在元胞格内的驻留时间.

2)采用微团场景状态建模方法对场景中人群状态进行检测和预报.

3)利用当前场景的最新信息对模型综合历史信息给出的预报值进行反馈校正,以达到尽早发现并防止发生安全事故的目的.实验结果表明,本文方法具有较好的群体状态预报能力.

参考文献

[1] BURSTEDDE C, KLAUCK K, SCHADSCHNEIDER A et al. Simulation of pedestrian dynamics using a two-dimensional cellular automaton [J]. Physica A: Statistical Mechanics and its Applications, 2001, 295(3): 507-525.

[2] YAMAMOTO K, KOKUBO S, NISHINARI K. Simulation for pedestrian dynamics by real-coded cellular automata (RCA) [J]. Physica A: Statistical Mechanics and its Applications, 2007, 379(2): 654-660.

[3] HAO Y, GUAN Hongzhi, ZHANG Juan, et al. Study on bi-direction pedestrian flow using cellular automata simulation[J]. Physica A: Statistical Mechanics and its Applications, 2010, 389(3): 527-539.

[4] WANG Dalong, KWOK N M, JIA Xiuping, et al. A cellular automata based crowd behavior model[C]//IEEE Conference on Artificial Intelligence and Computational Intelligence. Sanya: IEEE, 2010:218-225.

[5] AMOROSO S, YALE N P. Decision procedures for surjectivity and injectivity of parallel maps for tessellation structures [J]. Journal of Computer and System Sciences, 1972, 6(5): 448-464.

[6] MA J, LO S M, SONG W G, et al. Modeling pedestrian space in complex building for efficient pedestrian traffic simulation[J]. Automation in Construction, 2013, 30: 25-36.

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

[8] MORRIS B, TRIVEDI M. Learning trajectory patterns by clustering: Experimental studies and comparative evaluation[C]//IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE, 2009:312-319.

[9] JUNSEOK K, LEE K M. A unified framework for event summarization and rare event detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012:1266-1273.

[10] YANG Cong, YUAN Junsong, LIU Ji. Sparse reconstruction cost for abnormal event detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. Colorado: IEEE, 2011:3449-3456.

[11] NATARAJAN P, WU Shuang, VITALADEVUNI S, et al. Multimodal feature fusion for robust event detection in web videos [C]//IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012:1298-1305.

[12] 朱海龙, 刘鹏, 刘家锋, 等. 人群异常状态检测的图分析方法[J].自动化学报, 2012, 38(5): 742-750.

(编辑 张 红)