

区域人群状态的实时感知监控

宋宏权, 刘学军, 闫国年, 张兴国

(南京师范大学 虚拟地理环境教育部重点实验室, 南京 210023)

摘要: 公众聚集场所人群高度聚集, 流动性大, 隐藏着巨大的安全隐患, 时常发生群死群伤的拥挤踩踏等突发公共事件。针对现有以视频监控的人群分析, 无法从空间视角掌握区域人群状态的时空格局, 本文提出了面向人群分析的视频 GIS 框架, 将视频数据映射至地理空间, 在 GIS 环境下提取人群密度、人群运动矢量场等人群特征。通过分析人群运动矢量场可得到人群运动模式及各方向人群主体运动速率。最后, 将视频监控系统与 GIS 进行有机集成, 设计并实现了以视频与 GIS 协同的区域人群状态实时感知监控系统。实验结果表明, 本系统可为大型集会活动的突发事件预防、人群疏导等提供决策依据。

关键词: 视频 GIS; 人群监测; 时空格局; 感知; 监控

DOI: 10.3724/SP.J.1047.2012.00686

1 引言

随着社会经济的快速发展, 大规模人群聚集活动频繁出现, 因人群密度过大而造成的人员伤亡事故屡见不鲜, 世界各国已将视频监控技术的研究上升到战略高度, 监控探头已遍及城市各个重要角落。实时监控人群聚集活动区域的人群状态, 可为人群管理、突发事件预警等提供重要依据。本文的区域是指在露天道路型公众聚集场所, 离散布设有大量监控摄像机的特定区域, 如南京市夫子庙步行街、北京市王府井大街等。人群状态包括人群密度、人群速度、人群运动方向等人群特征。然而, 目前, 对某区域的人群监控大都为特定出入口截面的人员总数统计^[1], 或统计单个监控摄像机的人群数量或密度^[2], 包括前景像元法^[3-5]、纹理分析法^[6-7]、目标识别法^[8]等, 而基于视频的人群运动分析局限于图像空间, 无法获得人群的真实运动速度与方向^[9-10]。

要完成大区域人群状态时空格局的实时监控, 需将视频分析系统与 GIS 进行有机集成, 实现视频分析与 GIS 之间的耦合与协同。从 20 世纪 70 年代开始^[11], 国内外学者就开始了将视频数据与 GIS

的集成研究, 但只局限于以下 3 个方面内容: (1) 将视频数据作为空间数据的属性, 采用超链接方式静态调用视频文件^[11-13]; (2) 将视频数据与 GPS 数据融合实现视频与空间数据的交互^[14-22]; (3) 将视频数据投影至二维平面, 实现视频图像与二维或三维地图中某个侧面的叠加^[23], 如 ArcGIS、Skyline 等软件, 或实现实景影像的简单量测^[24]。视频数据本身具有时空、表达直观、信息丰富、动态实时等特点, 而目前的研究并没有充分利用视频所蕴含的丰富信息, 因此, 将视频分析系统与 GIS 有机集成, 发展地理环境协同下的视频分析技术与方法, 实现区域人群状态实时感知监控系统, 是一个亟待解决并具有现实意义的研究课题。

2 面向人群分析的视频 GIS 框架

为实现对区域人群状态的实时感知监控, 本文将视频分析系统与 GIS 相结合, 提出了面向人群分析的视频 GIS 框架, 如图 1 所示。

本框架将视频数据映射至地理空间, 在 GIS 环境下提取人群密度、人群速度、人群运动方向等特征, 通过人群特征分析进一步得到人群运动模式及

收稿日期: 2012-11-01; 修回日期: 2012-12-10.

基金项目: 国家“十二五”科技支撑计划课题“视频 GIS 与突发公共事件的感知控制系统”(2012BAH35B02); 江苏省高校自然科学重大基础研究资助项目(10KJA420025)。

作者简介: 宋宏权(1986-), 男, 河南民权人, 博士研究生。研究方向: 视频 GIS, 地理信息系统设计、开发与应用。

E-mail: hongquansong@126.com

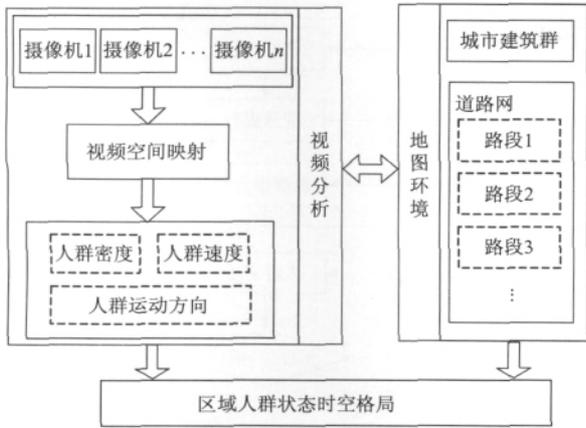


图1 面向人群分析的视频 GIS 框架

Fig. 1 Video-GIS framework for crowd analysis

人群主体运动方向,对人群高度聚集的监控场景进行报警处理,实现视频分析系统与GIS的协同,进而完成区域人群状态的实时感知监控。

3 人群特征提取与分析方法

3.1 视频数据空间映射

在地理空间提取人群特征,可得到具有统一地理参考量纲的真实值,且能够在GIS环境下对其进行表达与分析。为此,本文设计了视频数据空间映射方法,将视频影像中的人群活动区域映射至地理参考,使得图像数据具有统一的度量参考,进而实现地理空间下的人群特征提取与分析。具体实现流程如图2所示:①人群活动的区域大都为地平面,为避免其他动态目标对人群特征提取的干扰,在监控场景中选取人群活动区域;②由于视频数据大都为透视影像,需对选取的人群活动区域进行透视校正,以消除因目标距摄像机远近不同而产生的近大远小现象;③对透视校正后的人群活动区域图像,选取3个以上对应的地面控制点进行单应计算;④利用校正后人群活动区域图像与地理空间的单应关系,将其映射至地理参考,完成人群图像与地理空间的参考统一。

3.2 自适应人群密度估计

采用前景像元的人群密度估计方法在人群密度较小时具有良好的效果,随着人群密度的增大,因行人间相互遮挡使得此类方法的线性关系不再成立。由纹理分析的人群密度估计方法可知:高密

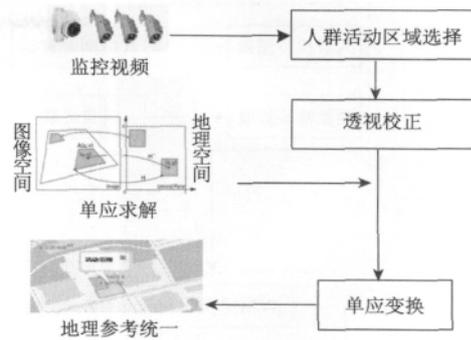


图2 视频数据空间映射流程

Fig. 2 Spatial video generation process

度人群在纹理上表现为细模式,低密度人群表现为粗模式^[25]。提取人群图像的纹理特征,训练人群密度等级分类器,并使用分类器实现对人群密度的等级分类。此类方法对高密度人群具有良好的分类效果。

面向低密度与高密度人群,本文设计了地理空间下的自适应人群密度估计方法,即当人群密度较低时采用以像元的方法,人群密度较高时采用纹理分析的方法。本方法分为模型训练与人群密度自适应估计两部分。具体如图3所示,模型训练的流程为:①选取人群图像样本数据,利用视频数据空间映射方法将人群图像统一至地理空间;②在地理空间进行人群运动前景的提取;③对人群前景图像进行特征提取,对于低密度人群图像,提取其边缘像元并生成二值图像,对于高密度人群图像,计算其灰度共生矩阵等纹理特征值;④根据标定的人数与前景边缘像元数的关系,拟合低密度人群数量估计模型,对于高密度人群图像,利用提取的纹理特征值与人群密度等级的对应关系,训练得到支持向量机(SVM)分类器,用于高密度人群的密度等级分类。

模型训练完成后,可实现对人群密度的自适应估计,具体步骤为:①获取视频流,并对要处理的人群图像进行空间映射处理;②在地理空间提取人群前景图像,进行形态学处理与二值化;③对二值化图像进行边缘检测得到前景边缘像元数目,利用以像元的低密度人群数量估计模型计算人群数量,并利用人群活动区域的面积计算人群密度值;④若人群密度值大于设定阈值,则提取前景图像的纹理特征值,并利用SVM人群密度分类器进行人群密度的分类,当分类结果为极高时,可报警通知人群密

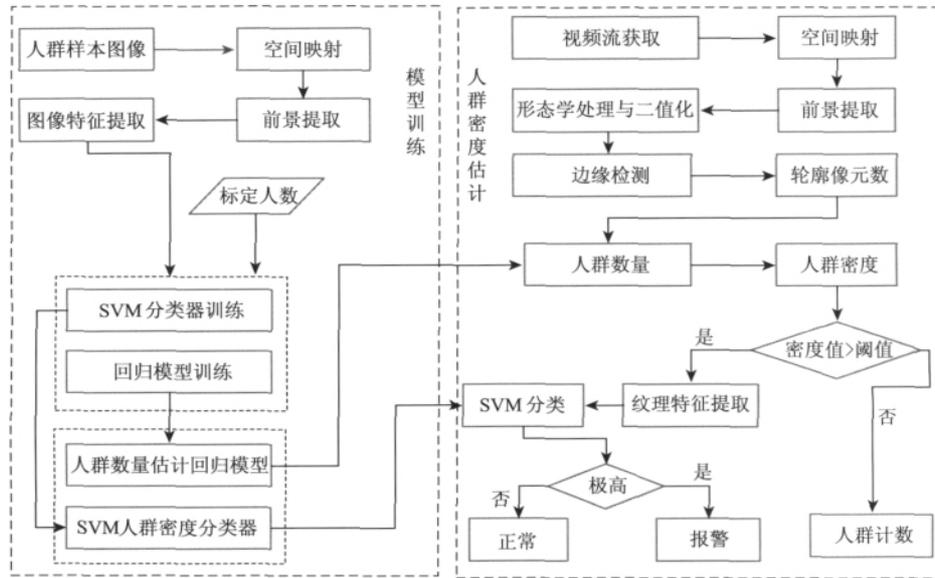


图 3 自适应人群密度估计与预警流程

Fig. 3 Adaptive crowd density estimation and the early warning process

度过高,易发生拥挤踩踏等突发事件。其中阈值的设定根据公共场所人群安全管理报告^[26]规定的人群密度建议值,结合我国人口的具体情况,对人群密度等级的设定如表 1 所示,本文设定的阈值为 1.53 人/m²。

表 1 人群密度等级标准

Tab. 1 Crowd density levels

等级	描述	密度(人/m ²)
很低	环境相对自由	<0.83
低	一般状况	0.83~1.53
中	中等密度	1.53~2.50
高	混乱状况下	2.50~3.30
很高	非常拥挤	3.30~5.50
极高	极易发生拥挤事故	>5.50

3.3 人群运动特征提取与分析

人群运动研究主要是通过人群运动速度定量的表征人群运动状态。人群运动速度一般是指人群整体表现出来的速度状态,不是由个人的速度决定,而是人群在行走过程中相互影响和制约表现出来的一种平均速度状态。人群运动速度包括大小和方向两个方面。人群运动速度大小是表征人群移动快慢的物理量。一般来讲,人群密度越大,人群的整体移动速度越慢。人群运动速度的方向性对人群流动有很大影响,不同方向的行人在行走

过程中会产生冲突与碰撞,当此局面进一步恶化易产生拥挤踩踏等突发事件。

本文在地理空间利用光流法提取人群运动矢量场,克服了现有以图像空间的方法不能获取其真实值的问题,并对矢量场进行分析,进而得到人群的运动模式与主体运动速度。具体流程如图 4 所示:①实时获取视频流,并对视频数据进行空间映射处理;②利用空间映射处理后的人群图像数据,在地理空间进行光流计算,求算地理参考下的人群运动光流场;③对人群运动光流场进行矢量分析,包括矢量场的主体方向分析、主体运动速度分析等,得到人群的主体运动趋势;④最后,根据人群运动趋势判断人群的运动模式,包括同向运动、双向运动、交叉运动、无序运动、聚拢运动与发散运动等。

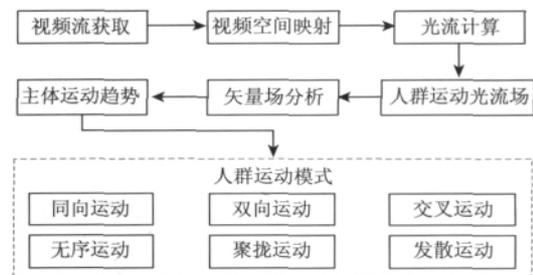


图 4 人群运动特征提取与分析流程

Fig. 4 Flow chart of the crowd movement character extraction and analysis

4 区域人群实时感知监控的集成分析

4.1 人群密度估计结果分析

利用视频数据空间映射方法将人群图像映射至地理空间,如图 5 所示,其中图 5a 为原始图像,图 5b 为选取的人群活动区域,图 5c 为透视校正后的人群活动区域图像,图 5d 为空间映射后人群图像与地图的叠加显示。以上操作,可得到人群活动区域的真实面积,在地理空间进行人群密度估计与人群运动分析,其结果能够在 GIS 环境下实时动态展示。与现有图像空间的处理方法相比,其无需进一步计算结果与真实值之间的转换。

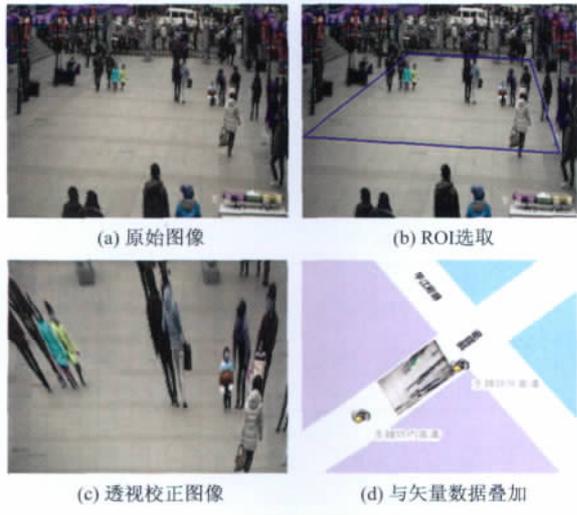


图 5 人群图像空间映射结果
Fig. 5 Spatial-video generated results

当人群密度等级为很低和低时采用像元的方法,密度等级为中、高、很高和极高时采用纹理的方法,故将阈值设置为 1.53 人/m²。图 6 为某摄像机低密度人群估计模型估计结果与人工标定结果的对比。为评价模型的精度,选取了 200 帧测试图像,采用式(1)来定量评价此模型的估计结果。

$$err = \frac{\sum_{i=1}^n \frac{|y(i) - p(i)|}{p(i)}}{n} \times 100\% \quad (1)$$

其中, $y(i)$ 为模型估计人数, $p(i)$ 为人工标定人数, n 为测试样本图像数量,根据此精度评价模型可得人群数量的估计结果精度为 89.53%。

对于高密度人群,本文使用纹理分析与 SVM 相结合的方法估计人群密度。综合考虑准确率与

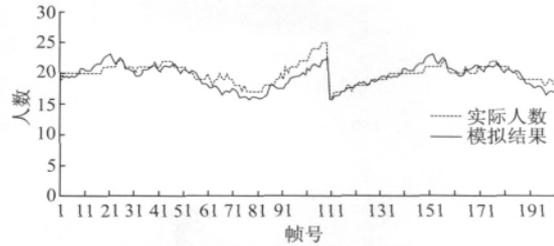


图 6 低密度人群数量估计结果
Fig. 6 Low density crowd counting results

算法效率,利用灰度共生矩阵描述图像纹理特征,经多次实验,将图像灰度级定为 16 级,距离设定为 8。实验得知在 0°、45°、135° 三个方向人群图像具有一致的纹理特征,与 90°方向的纹理特征差别较大,故选取 0°和 90°两个方向计算灰度共生矩阵。经反复测试,利用能量、相关性、对比度、熵、逆差矩等 5 个纹理特征的分类结果较好,利用本方法对中、高、很高、极高 4 类进行分类。为评价此方法的精度,对某摄像机的 4 类人群图像各选取 40 幅作为测试样本,其分类结果如表 2 所示,精度可达到 95.63%。

表 2 高密度人群等级分类结果
Tab. 2 High density crowd levels classification results

测试样本	分类结果				准确率 (%)
	中	高	很高	极高	
中(40)	38	0	0	0	95.63
高(40)	2	39	0	0	
很高(40)	0	1	38	2	
极高(40)	0	0	2	38	

4.2 人群运动实验结果分析

采用夫子庙步行街某监控摄像机数据用于人群运动分析实验。将人群图像数据映射至地理空间,用光流法在地理空间求算人群运动光流场。将运动矢量场数据投影至极坐标,如图 7(a)所示。可以看到,人群运动主要集中于西北-东南方向,进而可判断人群的运动模式为双向运动。利用此方法可有效地判断人群运动模式,如若人群运动矢量场以极坐标原点为圆心的圆状分布,则人群运动模式为聚拢或发散运动。对矢量场在东、南、西、北、西北、东北、东南、西南 8 个方向的累积矢量进行分析,其统计分析结果如图 7(b)所示。可以看出,在 8 个方向中西北和东南方向的累积矢量最大,也能够说明人群在地理空间中做西北-东南方向的双向运动,这与实际的观察情况相吻合。

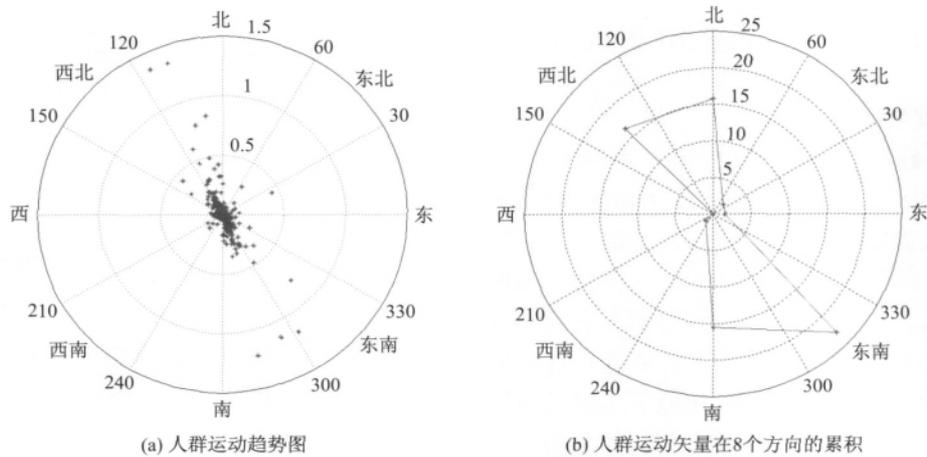


图7 人群运动分析结果

Fig. 7 Crowd movement analysis results

本实验每隔 100 毫秒捕获一帧图像,对捕获的相邻两帧图像进行光流计算,可进一步得到矢量场在地理空间各方向的运动速率,如表 3 所示,其中

位移的单位为 m,速率的单位为 m/s。从表中可以看出,在人群运动的主要方向西北和东南方向,人群的运动速率分别为 0.237m/s 和 0.223m/s。

表3 人群平均速度估计结果

Tab. 3 Crowd average velocity estimation results

	东	东北	北	西北	西	西南	南	东南
累积位移(m)	1.6427	1.8658	15.7674	16.5049	0.0000	1.2492	15.5104	22.9317
平均位移(m)	0.0083	0.0114	0.0279	0.0237	0.0000	0.0084	0.0210	0.0223
平均速率(m/s)	0.0830	0.1140	0.2790	0.2370	0.0000	0.0840	0.2100	0.2230

4.3 系统实现

本文将视频监控系统与 GIS 进行有机集成,结合视频数据空间映射、自适应人群密度估计、地理空间下人群运动特征提取与分析等算法,实现了区域人群状态实时感知监控系统。系统界面如图 8 所示,地图窗口实时显示各监控场景的人群密度等级和人群主体运动方向。系统左上角窗口实时显示各监控场景的人群状态,当人群密度大于 5.50 人/m²时系统将报警提示人群密度过高,易导致拥挤踩踏事故,用户可查看报警的视频场景,可根据具体情况对该监控区域进行人群疏导,以避免突发事件的发生。系统下侧显示的是重点监控摄像机人群数量实时统计图,用户可根据需要自行设定重点监控摄像机来实时监测人群数量变化。同时,本系统可对整个监控区域的历史数据进行时空分析,以便从时空的角度掌握该区域的人群流动规律,为安保人员的配置与布控提供可靠依据。

5 结论与讨论

实验表明,将视频数据映射至地理空间,使得人群图像数据具有与 GIS 统一的地理参考,解决了现有图像空间人群分析方法以像元为量纲的问题,可真实地在地理空间表达与分析人群特征数据。基于空间映射的视频数据进行自适应人群密度估计,低密度精度可达 90%,高密度分类结果精度在 95%以上,可满足人群密度估计的现实需求。在地理空间进行人群运动特征提取与分析,可判断监控场景的人群运动模式,同时可获得人群各方向的总体运动速率,可很好地掌握人群运动状态。基于面向人群分析的视频 GIS 框架,将视频监控系统与 GIS 进行有机集成开发的区域人群状态实时感知监控系统,可实时感知监控区域人群状态的时空格局,为大型集会活动的突发事件预防、人群疏导等提供决策依据,以避免群死群伤等恶性安全事故的发生。

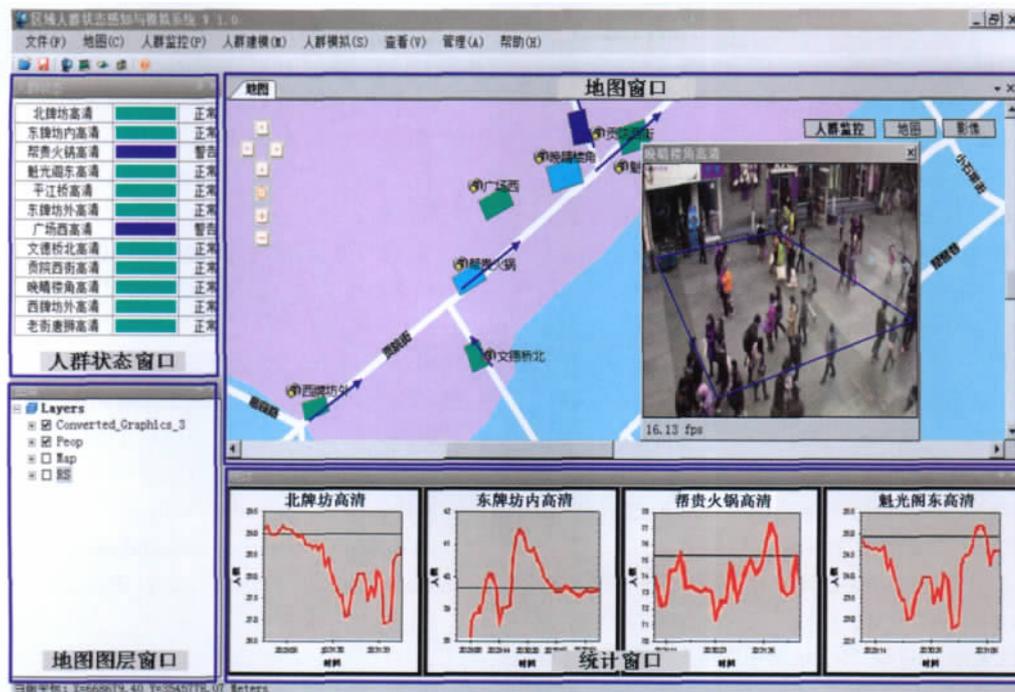


图 8 系统界面

Fig. 8 System interface

本文只实现了对区域人群密度、人群运动模式等人群状态的感知监控,如何根据人群运动速度、运动方向等特征检测人群的骤聚、骤散等异常状况,以及根据人群运动特征与人群密度之间的关系,推演空白监控区的人群状态;如何进行区域人群流动的情景模拟预测等,从大范围分布式视频监控数据中提炼与公共安全密切关联的行为、事件等语义信息,是需进一步要研究的内容。

参考文献:

- [1] 肖坦,杨棚,李黄煌.基于视频监控的铁路客运站人群密度分析算法[J].铁道通信信号,2010,46(8):80-82.
- [2] 杨琳,苗振江.一种人群异常行为检测系统的设计与实现[J].铁路计算机应用,2010,19(7):37-41.
- [3] 贾永华,呼志刚,浦世亮.基于视频前景边缘和特征检测的人群密度估计方法[J].中国公共安全(综合版),2011,(5):180-184.
- [4] 吴晟,葛万成.基于可变矩形框的人群密度估计算法[J].通信技术,2011,44(10):63-65.
- [5] Ma R, Li L, Huang W, *et al.* On pixel count based crowd density estimation for visual surveillance[C]. Proc. IEEE Conf. Cybernetics and Intelligent Systems, Singapore, 2004.
- [6] 刘福美,黎宁,张燕,等.一种基于图像处理的人群密度估计方法[J].计算机与数字工程,2011,39(5):118-122.
- [7] 麻文华,黄磊,刘昌平.基于置信度分析的人群密度等级分类模型[J].模式识别与人工智能,2011,24(1):30-39.
- [8] Rittsche J, Tu P H, Krahnstoeve N. Simultaneous estimation of segmentation and shape[C]. Proc. Computer Vision and Pattern Recognition, Washington, DC, 2005.
- [9] 王尔丹.人群运动与密度估计技术研究[J].安全,2005,(3):21-24.
- [10] Jacques Junior J C S, Muse S R, Jung C R. Crowd analysis using computer vision techniques[J]. IEEE Signal Processing Magazine, 2010, 27(5): 66-77.
- [11] Lippman A. Movie maps: An application of the optical videodisc to computer graphics[J]. SIGGRAPH'80, 1980, 14(3): 32-43.
- [12] Stefanakis E, Peterson M. Geographic hypermedia: Concepts and systems[M]. Lecture Notes in Geoinformation and Cartography, Springer, 2006, 1-21.
- [13] Klamma R, Spaniol M, Jarke M, *et al.* A hypermedia Afghan sites and monuments database[M]. Lecture Notes in Geoinformation and Cartography, Springer, 2006, Part III: 189-209.
- [14] Berry J K. Capture 'where' and 'when' on video-

- based GIS[J]. *GeoWorld*, 2000(9): 26-27.
- [15] Kim K H, Kim S S, Lee S H. The interactive geographic video[C]. *IEEE International Geoscience and Remote Sensing Symposium*, Japan, 2003.
- [16] Navarrete T, Blat J. VideoGIS: Segmenting and indexing video based on geographic information [C]. *The 5th AGILE Conference on Geographic Information Science*, Palma, Spain, 2002.
- [17] Hwang T H, Choi K H, Joo I H, *et al.* MPEG-7 metadata for video-based GIS applications [C]. *Japan*, 2003.
- [18] Lee S Y, Kim S B, Choi J H, Joo I H. Design and implementation of 4S-Van: A mobile mapping system [J]. *ETRI Journal*, 2006(3): 256-273.
- [19] 唐冰,周美玉. 基于视频图像的既有线路地理信息系统[J]. *铁路计算机应用*, 2001,10(11):31-33.
- [20] 孔云峰. 地理视频数据模型设计及网络视频 GIS 实现[J]. *武汉大学学报(信息科学版)*, 2010,35(2):133-137.
- [21] 宋宏权,孔云峰. Adobe Flex 框架中的视频 GIS 系统设计与开发[J]. *武汉大学学报(信息科学版)*, 2010, 35(6):743-746.
- [22] 丰江帆,张宏,沙月进. GPS 车载移动视频监控系统的的设计[J]. *测绘通报*, 2007,2:51-53.
- [23] 宋宏权,刘学军,闫国年,等. 基于视频的地理场景增强表达研究[J]. *地理与地理信息科学*, 2012,28(5):6-9.
- [24] 李德仁,郭晟,胡庆武. 基于 3S 集成技术的 LD2000 系列移动道路测量系统及其应用[J]. *测绘学报*, 2008,37(3):272-276.
- [25] Marana A N, Velastin S A, Costa L F, *et al.* Automatic estimation of crowd density using texture[J]. *Safety Science*, 1998, 28(3): 165-175.
- [26] Au S Y, Ryan M C, Carey M S, *et al.* Managing crowd safety in public venues: A study to generate guidance for venue owners and enforcing authority inspectors[R]. *HSE Contract Research Report*, 1993(53):993.

Real-time Monitoring for the Regional Crowds Status

SONG Hongquan, LIU Xuejun, LV Guonian and ZHANG Xingguo

(Key Laboratory of Virtual Geographical Environment, Ministry of Education, Nanjing Normal University, Nanjing 210023, China)

Abstract: With the rapid development of the social economy, the massive crowd gathering appears frequently. Personnel casualties often caused by higher crowd density. So, video surveillance technology has become a national policy in many countries. Surveillance cameras have been installed in various important places of the city. Real-time monitoring of the crowds status in crowd gathering area can provide important basis for crowd management and emergency warning. Existing video-based crowd analysis can only monitor crowd status for each camera separately. We cannot get the spatial-temporal patterns of regional crowd status from a spatial perspective. In this paper, we proposed a video-GIS framework for crowd analysis. Video frames can be mapped to geographic space based on the video-GIS framework. So we can process crowd images and extract crowd density, crowd movement vector field in GIS. Then the crowd movement pattern and the main direction of crowd movement can be acquired by the vector field analysis. Finally, we design and implement a real-time monitoring system for the regional crowd status using video surveillance system and GIS. Experimental results show that: (1) previous crowd analysis methods based on the image space can only measure results by the unit of pixels. It requires further conversion if we want to get the real value. But we can get the real value directly when we process crowd images in GIS using the method we proposed. (2) The accuracy of the pixel-based low-density crowd counting estimation results can be up to 90%. The classification accuracy of the high-density crowd levels support vector

(下转第 697 页)

channels. The process of watermark detecting is the inverse process of watermark embedding. Finally, the real-time copyright protection system for tile map is implemented and verified experimentally. By experiments, it has been confirmed that the proposed watermark algorithm is robust to cutting, splicing, and additive noise. Besides, the embedding has little effect on the size of original tile map.

Key words: tile data; indexing mechanism; watermark; copyright protection

(上接第 692 页)

machine classifier is more than 95%. So, they can fully meet the needs of crowd monitoring. (3) We can get the crowd movement pattern and the main movement direction by the analysis of crowd movement vector field in GIS. Also, we can obtain the speed of the crowd in different directions. These crowd characters all can be expressed in GIS. (4) The system we developed for the crowd monitoring can be applied to crowd management and emergency warning. It can provide decision making basis for emergencies prevention and crowd divert.

Key words: video-GIS; crowd status monitoring; spatial-temporal pattern; perception; real-time monitoring