

DOI:10.3969/j.issn.1003-0972.2018.01.011

视频和 GIS 协同的人群状态感知模型

邓仕虎^{1,3}, 张兴国^{2*}, 王小勇³, 朱俊丰³, 王 秀²

(1.重庆市地理信息中心, 重庆 401121; 2.信阳师范学院 地理科学学院, 河南 信阳 464000;
3.重庆知行地理信息服务有限公司, 重庆 401121)

摘要:针对监控视频在图像空间难以统一、宏观及真实量化人群状态的问题,提出了一种基于地理信息技术进行人群状态地图展示、量化分析及预警的方法。首先,通过摄像机厂商或标准视频流接口接入监控视频;然后,将视频图像调整为较小尺寸,采用光流法进行计算,获取图像空间下的光流场;最后,将各个监控视频相应的光流场映射至地理空间,即可在地图中观察人群运动状态,并可通过散点内插、等值线等进行分析,通过相关阈值的设置实现人群异常检测与预警。本文基于 MATLAB、ArcGIS Engine、C#等技术,采集多个实验视频,并研发了基于 GIS 的人群状态感知原型系统,对相关算法进行了验证。结果表明,该方法相对于传统图像空间下的人群状态检测方法,具有人群运动可定位、可量测、可宏观观察和预警等优势。

关键词:GIS; 视频; 人群; 光流; 异常行为

中图分类号:P208 **文献标志码:**A **文章编号:**1003-0972(2018)01-0059-05

Crowd Status Analysis Based on Surveillance Videos and GIS

DENG Shihu^{1,3}, ZHANG Xingguo², WANG Xiaoyong³, ZHU Junfeng³, WANG Xiu²

(1. Chongqing Geomatics Center, Chongqing 401121, China;

2. College of Geography Science, Xinyang Normal University, Xinyang 464000, China;

3. Chongqing Zhixing Geographic Information Consulting Services Ltd., Chongqing 401121, China)

Abstract: In order to solve the problem that surveillance videos were difficult to be unified in the image space and the crowd status was difficult to be perceived macroscopically and quantitatively, a new method was proposed based on GIS, which can be used to display, analyze and early warn crowd status through 2D map. Firstly, the surveillance video was accessed through the API of camera manufacturers or the standard video stream interface. Secondly, the video images were adjusted to a smaller size and the optical flow field in image space was calculated through some optical calculation algorithm. Finally, the optical flow field of each surveillance video was mapped to the corresponding geographic space. Then through GIS, the movement status of the crowd can be observed in the map. Two analysis methods, including scatter point interpolation and contour map, were put forward. The crowd alarm can be realized through relevant thresholds setting. Based on MATLAB, ArcGIS Engine and Microsoft Visual Studio C#, a prototype system was developed to validate the algorithm using several crowd videos. The results showed that the new method had the advantages of geographic location, measurement, macroscopic observation and early warning, etc. when compared with the previous methods.

Key words:GIS; video surveillance; crowd; optical flow; abnormal behavior

0 引言

随着社会经济的发展,人群聚集活动日趋增多,致使群体性事件时有发生。如何及时发现人群状态的异常并采取有效措施进行疏导已成为国内

外学者和政府管理部门关注的热点。目前,在大型安保活动中,除了加强警力部署外,视频监控系统的已成为必不可少的构成要素,发挥着越来越重要的作用。

收稿日期:2017-03-25;修订日期:2017-08-16; *通信联系人, E-mail: zhangxingguo2012@163.com

基金项目:国家自然科学基金项目(41401436);对地观测技术国家测绘地理信息局重点实验室开放课题(K201510);信阳师范学院博士科研启动基金项目(0201402)

作者简介:邓仕虎(1981—),男,重庆人,高级工程师,博士生,主要从事 GIS 理论与方法研究。

当前,基于视频获取人群运动状态的研究已取得了较好的研究成果.在视频图像空间,通过图像处理、光流计算等获取运动特征,通过机器学习的方法实现人群异常行为的检测^[1].为了扩大人群视频监控的视域范围,胡学敏等基于鱼眼相机对人群进行估计,并采用透视加权的方法来消除透视畸变^[2].在视频流的实时计算方面,姬丽娜等基于 GPU 多流处理器进行加速,有效地提高了人群实时计数效率^[3].近年来,许多新技术也在视频人群信息提取中得到应用,效果较好,如基于序的空间金字塔池化网络、压缩感知、社会计算和深度学习等^[4-6].图像采用像素坐标,缺乏真实的度量坐标,人群可量测的运动状态参数获取困难,一些 GIS 学者尝试将视频与 GIS 结合起来,也取得了一定的效果^[7-9].

针对监控视频在图像空间难以统一、宏观及真实量化人群状态的问题,本文将视频与 GIS 相结合,将图像空间的光流场映射至地理空间,在地图中进行人群状态的展示、分析与预警.融合地理信息及实时视频分析数据的人群感知方法,相对于传统的方法更为便捷,将会有效降低安保人员的疲劳度,提高突发事件实时监控与预警能力.

1 人群状态感知模型

1.1 技术路线

人群状态感知流程主要包括四大过程,即视频采集及处理、光流计算、矢量光流场映射至地理空间、人群状态检测与分析,如图 1 所示.



图 1 人群状态感知技术路线

Fig. 1 Crowd status perception flow chart

第一,通过手机或相机在商场、医院、车站、银行等室内外人群密集、易发生突发公共事件的场所采集实验数据,包括正常的人群行为视频和异常行为视频,如打架、追捕、聚集、骤散、逆行等,为人群异常行为检测提供测试数据.基于高清遥感影像或 GPS,采集摄像机中心点的位置.基于相机标定或直接设定得到摄像机的高度、倾角、方位角、主距等内外参信息.第二,计算各个监控视频的光流场.以视频中前后两帧图像为基础,采用相关的光流计算方法,得到当前视频的实时光流场.第三,将区域内各个视频的光流场映射至地理空间.视频图像空间所计算的光流场,其坐标为像素坐标,难以在地理空间中进行定位、量测或分析,需要通过图

像空间与地理空间的映射方法将其投射至地图中.第四,在地理空间中,任何一条矢量光流都具有明确的空间参考,很容易了解人群的位置、方向、速度、加速度等运动参数,是人群状态检测及预警的重要基础.

1.2 关键技术

基于视频提取动态目标的运动矢量场,主要依赖于图像空间,光流场以像素为单位,其方向以图像空间为基础,无法获取动态目标在地理场景中真实的度量值(包括位置、方向和速度等).为了使人群状态可定位、可量测、可分析,并具有一览性的优势,需要重点解决光流场的提取及光流场至地理空间的映射问题.

1.2.1 Lucas-Kanade 光流计算方法

光流场计算是人群状态感知最为重要的关键技术之一.光流场的计算最初是由美国学者 HORN 和 SCHUNCK 提出,并在研究中不断推进光流场相关算法的改进和优化.光流场应用很广范,可以用于行人检测、行人的步态分析、运动分割等,也有学者对光流法执行效率进行评价^[10].

本文采用 Lucas-Kanade 光流法来获取人群的运动状态.该方法首先将视频转换为序列图像,并转换为序列灰度图,约束在较短时间内,运动目标相同点的灰度值不会发生变化,依此构建方程,计算光流矢量.人群状态的变化可以通过场模型来表达.序列灰度图像的变化是人群运动的表现,灰度图变化越快,人群的运动越快.在临近时间内,相同特征点灰度值不变的约束,可以形成一个矢量场.矢量场中各个矢量的大小和方向即人群运动的快慢和方向.需要注意的是,该矢量场并不是人群真实的运动速度,仅仅是人群运动的透视视图.

Lucas-Kanade 方法根据临近帧灰度梯度的变化,寻找各个像素的运动矢量.以 HORN 与 SCHUNCK 的研究为基础,产生了 Lucas-Kanade 光流算法,效果较好.该方法是一种基于梯度的局部参数化光流估计方法.该方法假设在图像空间中一个小的窗口内,光流矢量保持不变,则可以采用最小二乘法计算其光流矢量.

假设图像平面一个像素点 (x, y) , 在 t 时刻的亮度为 $E(x, y, t)$, $u(x_0, y_0)$ 为像素点的水平变化分量, $v(x, y)$ 为像素点的垂直变化分量.

$$u = dx/dt, \quad (1)$$

$$v = dy/dt. \quad (2)$$

在经过时间间隔 Δt 后该点对应点亮度为 E

$(x + \Delta x, y + \Delta y, t + \Delta t)$, 当 Δt 很小趋近于 0 时, 假设该点亮度不变, 如式(3).

$$E(x, y, t) = E(x + \Delta x, y + \Delta y, t + \Delta t). \quad (3)$$

当该点的亮度有变化时, 将移动后点的亮度由泰勒公式展开, 如式(4).

$$E(x + \Delta x, y + \Delta y, t + \Delta t) = E(x, y, t) + \frac{\partial E}{\partial x} \Delta x + \frac{\partial E}{\partial y} \Delta y + \frac{\partial E}{\partial t} \Delta t + \epsilon, \quad (4)$$

若二阶导数为无穷小, Δt 为 0, 如式(5).

$$-\frac{\partial E}{\partial t} = \frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} = \nabla E g w, \quad (5)$$

式中: $w = (u, v)$, 该式为基本的光流约束方程.

令 $E_x = \partial E / \partial x$, $E_y = \partial E / \partial y$, $E_t = \partial E / \partial t$ 表示图像中像素点灰度沿 x, y, t 方向的梯度, 可将上式改写成: $E_x u + E_y v + E_t = 0$, 即灰度一致约束条件. 输入的视频尺寸为 $X \times Y$; 检测当前图像的所有 Harris 角点作为特征点; 然后运用金字塔 LK 光流方法, 取 $w \times w$ 的特征窗口, 根据光流约束条件求解特征点的光流 (u, v) , 如式(6).

$$\begin{bmatrix} E_{x1} & E_{y1} \\ E_{x2} & E_{y2} \\ \dots & \dots \\ E_{xn} & E_{yn} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -E_{t1} \\ -E_{t2} \\ \dots \\ -E_{tn} \end{bmatrix}, \quad (6)$$

即 $A v = -b$.

然后, 采用最小二乘法, 如式(7)和(8).

$$A^T A v = A^T (-b) \text{ 或 } v = (A^T A)^{-1} A^T (-b), \quad (7)$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum E_{xi}^2 & \sum E_{xi} \sum E_{yi} \\ \sum E_{xi} \sum E_{yi} & \sum E_{yi}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum E_{xi} \sum E_{ti} \\ -\sum E_{yi} \sum E_{ti} \end{bmatrix}. \quad (8)$$

实验视频共 n_{Frame} 帧图片, 时间为 t_{Sum} 秒, 则每帧图像所需时间可通过式(9)来计算.

$$t_{\text{Each}} = t_{\text{Sum}} / n_{\text{Frame}}. \quad (9)$$

解算中, 求和从 1 到 n , 通过图像导数计算, 并各自累加得出. 同时还需权重函数 $W(i, j, k)$, $i, j, k \in [1, m]$ 来突出窗口中心点的坐标, 采用高斯函数进行计算效果明显, 可较好地得到当前视频的实时光流场. 该算法对特定应用场景(如车站人群)监测效果较好, 具有广阔的应用前景.

1.2.2 光流场映射至 2D 地理空间

由视频得到的光流场, 其坐标为像素坐标, 难以在地理空间中进行定位、量测或分析等, 需要通过图像空间与地理空间的映射方法将其投射至地图中. 地图明确的空间参考、宏观观测视角, 有助于整体把握大区域人群的运动态势. 光流场映射至 2D 地理空间是本文另一关键技术.

传统的方法需要计算 2D 单应矩阵. 首先, 在图像中选择 3 个及以上控制点, 得到控制点的图像坐标; 然后, 在高清遥感影像或地图中选择相应点, 得到控制点的空间直角坐标; 最后, 计算 2D 单应矩阵 H . 该方法不仅需要交互选择, 而且要求具备高清遥感影像或地形图, 适合于小型的视频监控场景, 难以在大型的视频监控系统中使用. 特别是, 当获取目标的三维几何信息时, 2D 单应方法无能为力. 在此背景下, 张兴国等^[11]提出了监控视频与 2D 地理空间的互映射模型, 可以较好地实现图像空间与二维地理空间的一一对应.

为了得到人群真实的三维信息, 顾及 PTZ 摄像机, 本文引入了监控视频映射至 2D 地理空间的互映射模型. 该模型要求摄像机内外参已知, 便于计算. 相机模型以齐次方程来表示, 如式(10).

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & u_0 & 0 \\ & f_y & v_0 & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} R & -R \bar{C} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (10)$$

式中: (X, Y, Z) 为 GIS 中的空间直角坐标, \bar{C} 为空间直角坐标系下摄像机中心的位置, (x, y) 为图像空间坐标, (u_0, v_0) 为像主点坐标, $f_x = f/p_x$, $f_y = f/p_y$ 表示等效焦距, s 为坐标轴的非正交性, R 为旋转变换.

式(10)中, 摄像机的内外参数隐藏在矩阵当中. 为了便于使用, 通过对比分析摄影测量学和计算机视觉的相机模型, 得到了便于使用的相机模型, 如式(11).

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & u_0 \\ & f v_0 \\ & & 0 \end{bmatrix} \begin{bmatrix} a_1 & b_1 & c_1 \\ -a_2 & -b_2 & -c_2 \\ -a_3 & -b_3 & -c_3 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (11)$$

式中: $u_0 = x_{\text{oph}}$, $v_0 = H_{\text{pic}} - y_{\text{oph}}$ 为像主点的坐标, H_{pic} 为图像高度, 与摄影测量学中的内外参数含义保持一致.

对于平面直角坐标系下某点, 若高程为 Z_h , 可实现图像坐标到平面直角坐标的转换, 如式(12).

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f a_1 - u_0 a_3 & f b_1 - u_0 b_3 & (f c_1 - u_0 c_3) Z_h + s_1 \\ f a_2 - v_0 a_3 & -f b_2 - v_0 b_3 & (-f c_2 - v_0 c_3) Z_h + s_2 \\ -a_3 & -b_3 & (-c_3) Z_h + s_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (12)$$

式中: $s_1 = (f a_1 - u_0 a_3)(-X) + (f b_1 - u_0 b_3) \cdot (-Y) +$

$$(fc_1 - u_0c_3)(-Z), s_2 = (fa_2 - v_0a_3)(-X) + (-fb_2 - v_0b_3) \cdot (-Y) + (-fc_2 - v_0c_3)(-Z), s_3 = (-a_3)(-X) + (-b_3)(-Y) + (-c_3)(-Z).$$

2 实验与分析

2.1 实验平台与数据

本文基于 MATLAB、OpenCV、VS C# 2012、ArcEngine 作为实验平台,并将各环境下实现的算法进行集成,开发了一个人群状态感知的原型系统. MATLAB 和 OpenCV 下的计算机视觉基础库较多,已经实现了众多基础功能,如视频实时接入、视频文件播放、视频帧的获取、图像尺寸调整等,本文以此为基础,实现了所述算法,并封装为 DLL. 在 VS C# 2012 下调用相关 DLL,并结合 GIS 二次开发组件 ArcEngine,开发完成了原型系统.

本文以某高校的餐厅及教学楼前作为实验区域,采集了多个包含人群的视频数据,并搜集了该区域的 QuickBird 遥感影像及高精度矢量地图. 为了获取摄像机的内外参,采用灭点、灭线的方法来计算主距,结合高清遥感影像,获取摄像机的外参(约束旋转角为 0). 从而,较好地实现了视频数据到地理空间的映射.

2.2 实验结果及分析

实验主要包括三方面内容,即光流提取、光流至地理空间映射、光流的空间分析. 基于 Lucas-Kanade 算法,实现了餐厅、教学楼前视频数据光流信息提取,光流真实反映了人群运动的趋势. 如图 2,箭头的指向为光流的方向,其长度代表光流的值. 当人群中某人运动正常时,箭头呈现方向趋同,长度相似,否则箭头会出现杂乱,长度会长一些.



图 2 人群异常行为检测

Fig. 2 Abnormal behavior detection

基于视频图像空间与地理空间的互映射模型,可以将视频中提取的光流实时映射到 GIS 中. GIS

中,以箭头的形式来表达,具有明确的空间参考和度量信息,可以从全局把握区域人群的运动状态. 如图 3,原型系统中展现了某时刻的运动光流场.

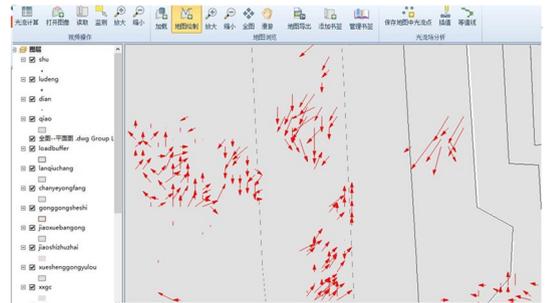
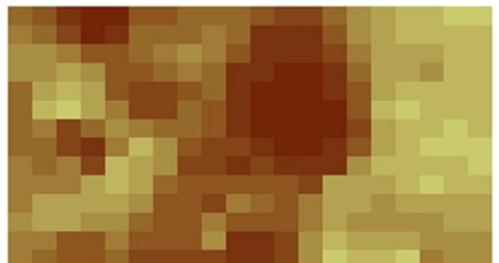


图 3 地图中的光流场

Fig. 3 The optical flow field in the map

基于 GIS 可以对光流场进行空间分析,本文进行了散点内插和等值线生成实验. 散点内插选择了 IDW 方法. 如图 4(a),可以看出颜色越深速度越快,可根据颜色的深浅了解人群运动的速度场. 基于内插的 DSM,可以自动生成等值线,等值线的分布也反映了人群的运动状态,如图 4(b). 值的变化量越小,线的间距就越大. 值上升或下降得越快,线的间距就越小.



(a) 光流值内插图



(b) 光流场等值线图

图 4 光流场空间分析

Fig. 4 Spatial analysis of the optical field

3 结论

基于 GIS 的人群感知方法,不仅解决了人群的空间定位、空间量算及一览性问题,而且也提供了面向人群的空间分析方法,将会有效增强传统的视频监控系统. 同时,实时的人群运动信息对于地理信息科学也是至关重要的,该方法可促进视频在 GIS 中的深入应用. 本文所提出的模型通过实验

证已取得了较好的实验效果,但尚需在光流的尺度效应、精度评价、运行效率、多摄像机下光流的融合等方面深入研究. 视频图像的尺寸对光流场有重要影响,针对人群状态感知而言,视频图像的最佳尺寸是一个值得研究的问题. 摄像机距离人群的远近也影响光流信息的提取,距离或者深度的有效

区间是光流计算的前提. 另外,多个摄像机距离人群远近不同、摄像机的内外参各异,光流的融合模型是大场景下人群光流场可视化的关键. 如何进一步提高运算速度、实时性、抗噪性及系统的鲁棒性是后续研究的重点.

参考文献:

- [1] 赵永涛,陈庆奎. 采用光流矢量共生矩阵分析的实时人群异常检测[J]. 小型微型计算机系统,2017,38(6):1353-1357.
ZHAO Yongtao, CHEN Qingkui. Real-time abnormal crowd detection utilizing optical flow co-occurrence matrix analysis[J]. Journal of Chinese Computer Systems, 2017,38(6):1353-1357.
- [2] 胡学敏,郑宏,郭琳,等. 利用鱼眼相机对人群进行运动估计[J]. 武汉大学学报(信息科学版),2017,42(4):537-542.
HU Xuemin, ZHENG Hong, GUO Lin, et al. Crowd monitor estimation using a fisheye camera[J]. Geomatics and Information Science of Wuhan University, 2017,42(4):537-542.
- [3] 姬丽娜,陈庆奎,陈圆金,等. 基于 GPU 的视频流人群实时计数[J]. 计算机应用,2017,37(1):145-152.
JI Lina, CHEN Qingkui, CHEN Yuanjin, et al. Real-time crowd counting method from video stream based on GPU[J]. Journal of Computer Applications, 2017,37(1):145-152.
- [4] 时增林,叶阳东,吴云鹏,等. 基于序的空间金字塔池化网络的人群计数方法[J]. 自动化学报,2016,42(6):866-874.
SHI Zengli, YE Yangdong, WU Yunpeng, et al. Crowd counting using rank-based spatial pyramid pooling network[J], ACTA AUTOMATICA SINICA, 2016,42(6):866-874.
- [5] 李冲,王蓉,陈鹏. 基于压缩感知的人群密度估计方法研究[J]. 中国人民公安大学学报(自然科学版),2016,22(1):68-72.
LI Chong, WANG Rong, CHEN Peng. Crowd density estimation based compressive sensing[J]. Journal of People's Public Security University of China (Science and Technology), 2016,22(1):68-72.
- [6] 操晓春,荆丽桦,王蕊,等. 基于社会计算和深度学习的社交网络特定内容监控[J]. 计算机科学,2016,43(10):1-8.
CAO Xiaochun, JING Lihua, WANG Rui, et al. Specific content monitoring on social network based on social computing and deep learning[J]. Computer Science, 2016,43(10):1-8.
- [7] SONG Hongquan, LIU Xuejun, LYU Guonian, et al. Video scene invariant crowd density estimation using geographic information systems[J]. China Communications, 2014,11(11):80-89.
- [8] 侯景伟,孔云峰,张迪. 基于 ArcEngine 的公安视频监控网络算法设计与实现[J]. 信阳师范学院学报(自然科学版),2013,26(3):415-418.
HOU Jingwei, KONG Yunfeng, ZHANG Di. Algorithm design of the video surveillance network for police and its realization based on ArcEngine [J]. Journal of Xinyang Normal University(Natural Science Edition), 2013,26(3):415-418.
- [9] 宋宏权,王丰,刘学军,等. 地理环境下的群体运动分析与异常行为检测[J]. 地理与地理信息科学,2015,31(4):1-5,11,133.
SONG Hongquan, WANG Feng, LIU Xuejun, et al. Crowd movement analysis and abnormal behavior detection under graphical environment[J]. Geography and Geo-Information Science, 2015,31(4):1-5,11,133.
- [10] 李亚楠,赵耀,林春雨,等. 基于图像分割的金字塔 Lucas-Kanade 光流法提取深度信息[J]. 铁道学报,2015,37(01):63-68.
LI Yanan, ZHAO Yao, LIN Chunyu, et al. Depth information extraction with pyramid Lucas-Kanade optical flow method based on image segmentation [J]. Journal of the China Railway Society, 2015,37(1):63-68.
- [11] 张兴国,刘学军,王思宁,等. 监控视频与 2D 地理空间数据互映射[J]. 武汉大学学报(信息科学版),2015,40(8):1130-1136.
ZHANG Xingguo, LIU Xuejun, WANG Saining, et al. Mutual mapping between surveillance video and 2D geospatial data[J]. Geomatics and Information Science of Wuhan University, 2015, 40(8): 1130-1136.

责任编辑:张建合