

一种可跨摄像机的人群密度估计模型^{*}

宋宏权^{1,2} 讲师 刘学军^{2**} 教授 闫国年² 教授 张兴国² 讲师

(1 河南大学 生命科学学院 河南 开封 475004

2 南京师范大学 虚拟地理环境教育部重点实验室 江苏 南京 210023)

学科分类与代码: 6208030(公共安全预测预警)

中图分类号: X924.3

文献标志码: A

基金项目: 国家自然科学基金资助(41201402, 41201417)。

资助项目: 国家“十二五”科技支撑计划项目(2012BAH35B02)。

【摘要】 为实现人群密度估计模型的跨摄像机应用,提出一种模型构建方法。利用地理信息系统(GIS)技术,将视频数据统一至GIS空间坐标系,在GIS中实现人群图像的尺度统一,在GIS空间构建模型以实现其可跨摄像机应用,并将视频监控系统与GIS有机集成,研发人群密度实时感知系统。研究结果表明,基于提出的模型构建方法,利用某摄像机数据构建的模型能被应用于其他摄像机监控场景,无需为各摄像机构建模型。所研发人群密度实时感知系统,已成功应用于“2012年江苏·秦淮灯会”。

【关键词】 地理信息系统(GIS); 视频GIS; 跨摄像机; 人群密度估计; 纹理分析; 支持向量机(SVM)

A Cross-camera Adaptive Crowd Density Estimation Model

SONG Hong-quan^{1,2} LIU Xue-jun LÜ Guo-nian² ZHANG Xing-guo²

(1 College of Life Sciences, Henan University, Kaifeng Henan 475004, China

2 Key Lab of Virtual Geographic Environment, Ministry of Education, Nanjing Normal University, Nanjing Jiangsu 210023, China)

Abstract: In previous crowd density estimation methods, an estimation model obtained by one camera cannot be applied to other cameras. In order to solve this problem, a method for cross-camera crowd density estimation was worked out. The proposed model mapped video data to a GIS coordinate reference system. So authors can process crowd images under geo-referenced coordinate system using GIS at same spatial scale. Test results show that one crowd estimation model obtained by a particular camera in GIS environment can be applied to other cameras. So cross-camera estimation model overcomes the dependence on surveillance scene and greatly improves the efficiency of crowd density estimation modeling. Finally, a system was designed and implemented for real-time monitoring regional crowds spatial-temporal distribution pattern based on cross-camera model, video surveillance system and GIS. This system has been successfully used in ‘Jiangsu Qinhuai lantern festival, 2012’. It can provide early warning information and a sound basis for safety and security decision making.

Key words: geographic information systems(GIS); video-GIS; cross-camera; crowd density estimation; texture analysis; support vector machine(SVM)

* 文章编号: 1003-3033(2013)12-0139-07; 收稿日期: 2013-10-16; 修稿日期: 2013-11-30

** 通信作者为刘学军教授。

0 引言

随着社会经济快速发展,车站、广场、步行街等人群密集场所越来越多,因人群密度过高而造成的人员伤亡事故屡见不鲜^[1]。人群密度是对公众聚集场所进行风险评估与安全管理的重要指标,实时监测人群密度具有重要现实意义。近年来,基于视频的智能人群感知技术已成为研究热点^[1]。利用视频分析估计人群密度的方法可分为3类:像素法、个体分割法和纹理分析法。像素法最早由Davies等^[3]提出,它是通过提取人群前景及其边缘像素数来构建人群密度估计模型的。个体分割法是在视频序列中识别人群个体,对分割的人群对象计数的一种方法^[4]。其缺点是当人群密度较大时,因遮挡导致较大误差。1998年,Marana等^[5]提出了基于纹理分析的人群密度估计方法。高密度人群在纹理上表现为细模式,低密度人群表现为粗模式,据此纹理特征实现高密度人群的密度估计。此后的研究^[6-13]在以上方法的基础上,进一步改进了算法的效率和准确率,但大都针对特定监控场景,其场景依赖性强,对设备的稳定性要求较高,监控探头姿态稍有变化便会影响模型精度。若监控区域有大量探头,则模型的训练样本较多,计算起来势必浪费时间和人力。

因监控探头安装高度、姿态及内部参数不同,场景中相同目标具有不同的图像表现特征,导致特定场景人群分析模型不具普适性。另外,当前基于视频分析的人群监控只针对特定监控探头^[13],无法实时感知公众聚集场所的人群时空分布格局。地理信息系统(GIS, Geographic Information Systems)是一种在统一空间坐标系下采集、处理、管理、分析,并集可视化地理空间数据为一身的工具,笔者将开发基于GIS的可跨摄像机人群密度估计模型的构建方法,以克服模型的场景依赖性,提高模型的构建效率。

1 可跨摄像机人群密度估计模型构建

1.1 视频数据的GIS空间映射

将视频数据统一至GIS空间坐标系,解决不同监控设备图像的尺度多样化问题,是构建可跨摄像机人群密度估计模型首要解决的问题。公众聚集场所大都为平面场景,如车站广场、步行街等。图1描述了视频数据与GIS空间坐标系的几何关系。点C为摄像机位置,摄像机拍摄的图像投影至像平面I,平面T为透视校正图像,G为GIS空间坐标系参考

平面。G平面中任一点 $P(x_g, y_g)$ 在I中的位置为 $p(u, v)$,在T中的位置为 $P_t(x_t, y_t)$,视频数据GIS空间映射即建立点p与点P的变换关系,实现图像空间I到GIS空间G的映射。

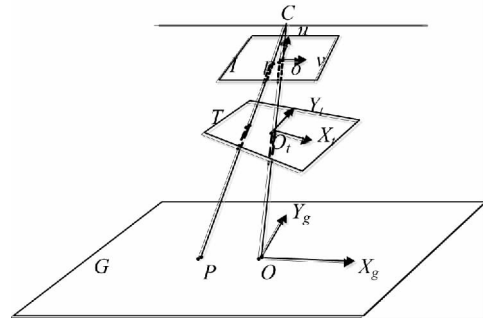


图1 视频数据地理空间映射模型

Fig.1 Spatial video generation model

透视校正后,令 (x_g, y_g) 为GIS空间某点坐标, (x_t, y_t) 为该点在校正后图像中的坐标,则二者存在式(1)所示变换关系:

$$\begin{bmatrix} x_g \\ y_g \\ 1 \end{bmatrix} = \begin{bmatrix} k_1 & k_2 & t_x \\ k_3 & k_4 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ 1 \end{bmatrix} \quad (1)$$

式中: t_x 与 t_y 为平移向量参数; k_1, k_2, k_3 与 k_4 为仿射变换参数。可通过3组以上对应点求算变换矩阵,实现监控视频数据的GIS空间映射。

1.2 低密度人群估计模型

采用基于像素的方法,利用映射后的人群图像构建低密度人群估计模型。具体步骤如图2所示:①选取低密度人群图像的活动区域(ROI, Region of Interest)并进行GIS空间映射处理;②在GIS空间提取人群活动前景图像块;③对前景图像进行边缘检测、形态学处理与二值化等操作;④标定人群数量,根据像素数与人群数量间关系,利用最小二乘法拟合可跨摄像机低密度人群数量估计模型;⑤根据人群数量与ROI面积,求算人群密度。

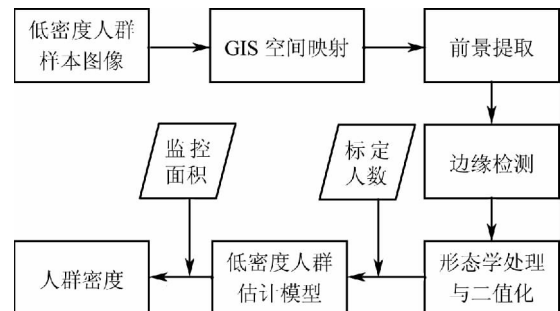


图2 低密度人群估计模型构建方法

Fig.2 Modeling process for low-density crowd estimation

1.3 高密度人群估计模型

将纹理分析与支持向量机(SVM, Support Vector Machines)结合,在GIS空间构建高密度人群密度估计模型。灰度共生矩阵是一种分析图像纹理特征的重要方法^[15-16],SVM是一种广义线性分类器,对基于小样本的高维非线性系统具有较好拟合效果^[17-18]。模型构建流程如图3所示:①选取不同密度等级的高密度人群图像,对其进行GIS空间映射处理;②在GIS空间提取人群活动范围前景图像块;③计算前景图像的灰度共生矩阵,提取纹理特征值;④根据人群密度等级与纹理特征值,训练可跨摄像机SVM分类器;⑤在GIS中提取人群图像的纹理特征,利用SVM分类器确定人群密度等级。

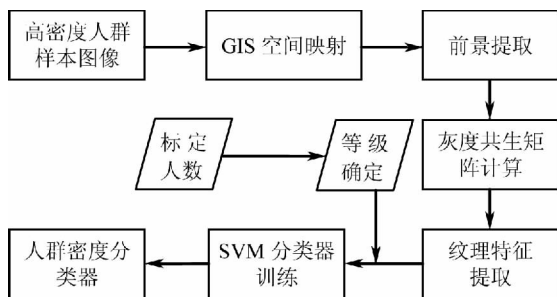


图3 高密度人群估计模型构建方法
Fig.3 Modeling process for high-density crowd estimation

1.4 自适应人群密度估计

基于低密度与高密度跨摄像机人群密度估计模型,在人群密度估计过程中设定阈值,自适应选取计算模型。具体方法为:①输入监控视频序列,对人群图像进行GIS空间映射处理;②在GIS中提取人群活动前景图像块;③对前景图像进行边缘检测、形态学处理等操作;④若前景边缘像素数目小于设定阈值,选取低密度人群估计模型计算人群密度;⑤若前景边缘像素数目大于设定阈值,提取前景图像的纹理特征,选取可跨摄像机SVM人群密度分类器估计人群密度等级。

2 试验结果与应用案例

2.1 视频数据的GIS空间映射结果

在图4a中选取ROI,如图4b中边框所示。对该区域进行透视校正,结果如图4c所示。将透视校

正图像统一至GIS空间坐标系(图4d),实现与遥感影像、矢量数据等空间数据的叠加。

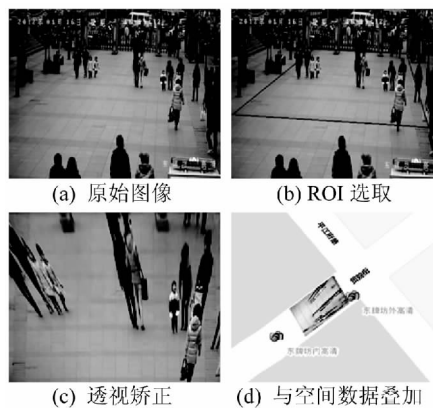


图4 视频数据地理空间映射结果

Fig.4 Spatial-video generated results

转换模型对视频数据GIS空间映射结果具有决定性影响,采用式(2)分析转换模型精度。

$$\sigma_m = \sqrt{\frac{\sum_{i=1}^n (\Delta_{xi}^2 + \Delta_{yi}^2)}{n}} \quad (2)$$

式中: Δ_{xi} 、 Δ_{yi} 分别为第*i*个待定位点转换前后在*X*、*Y*方向的差值;*n*为待定位点个数; σ_m 为模型转换精度。 σ_m 越大,转换精度越低,反之,精度越高。试验采用南京夫子庙步行街某监控探头的监控图像,根据4组待定位点与控制点对应关系,求得转换模型参数如下:

$$\begin{bmatrix} k_1 & k_2 & t_x \\ k_3 & k_4 & t_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.003\ 900 & 0.010\ 086 & 668\ 758.227 \\ -0.006\ 937 & 0.015\ 622 & 3\ 545\ 945.270 \\ 0 & 0 & 1 \end{bmatrix}$$

选取8个待定位点,利用该转换模型计算其对应的GIS坐标,采用式(2)精度评价模型,可得转换模型中误差为1.57cm(表1)。

2.2 人群密度估计

根据公共场所人群安全管理报告,结合我国人口具体情况,设定人群密度等级(表2)。利用夫子庙地区某监控探头(摄像头1)的数据作为训练样本,构建可跨摄像机人群密度估计模型。

表1 转换模型精度表
Table 1 Transformation model accuracy

点号	待定位点坐标		实测坐标		模型计算坐标		差值	
	x_i	y_i	x_g	y_g	x	y	Δx	Δy
1	120.712	-58.728	668 758.480	3 545 943.527	668 758.468	3 545 943.515	0.012	0.012
2	378.099	-86.668	668 759.351	3 545 941.312	668 759.363	3 545 941.293	-0.012	0.019
3	346.772	-136.622	668 759.062	3 545 940.718	668 759.046	3 545 940.730	0.016	-0.012
4	161.352	-185.728	668 758.143	3 545 941.263	668 758.132	3 545 941.249	0.011	0.014
5	102.932	-239.915	668 757.716	3 545 940.820	668 757.693	3 545 940.808	0.023	0.012
6	349.312	-288.175	668 758.443	3 545 938.332	668 758.465	3 545 938.345	-0.022	-0.013
7	41.971	-293.255	668 757.236	3 545 940.381	668 757.247	3 545 940.398	-0.011	-0.017
8	163.045	-184.882	668 758.161	3 545 941.232	668 758.142	354 5941.251	0.019	-0.019

中误差 $\sigma_m = 1.57 \text{ cm}$

表2 人群密度等级
Table 2 Crowd density levels

等级	描述	密度/(人·m ⁻²)
很低	自由环境	<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	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



图5 试验场景示例图像

Fig. 5 Test cameras' sample images

在试验区,选取3个场景进行试验分析。图5为试验场景示例图像,其分辨率为1 280 × 720。利

用传统基于像素的人群密度估计方法分别构建摄像头1和2和3的低密度估计模型,分别为: $y = 0.0017x + 34.1951$, $y = 0.0029x + 36.2948$, $y = 0.0042x + 18.2352$,其中, x 为前景像素数目, y 为估计人数。利用式(3)评价模型的精度。

$$err = \frac{\sum_{i=1}^n |y(i) - p(i)|}{n} \times 100\% \quad (3)$$

式中: $y(i)$ 为估计人数; $p(i)$ 为人工标定人数; n 为样本数量。对各摄像头选取100帧低密度人群图像,得出各模型的精度分别为91.43%,90.67%,90.78%。试验将摄像头1的估计模型用于摄像头2和3,其精度分别为74.34%和53.21%,与实际人数差别较大。

基于前述方法,拟合的低密度人群数量估计模型为: $y = 0.01973x - 5.3126$ 。选择摄像头1及另外2个监控探头数据各100帧,根据估计与标定结果(图6),得此模型的精度分别为89.77%,89.73%,89.47%。

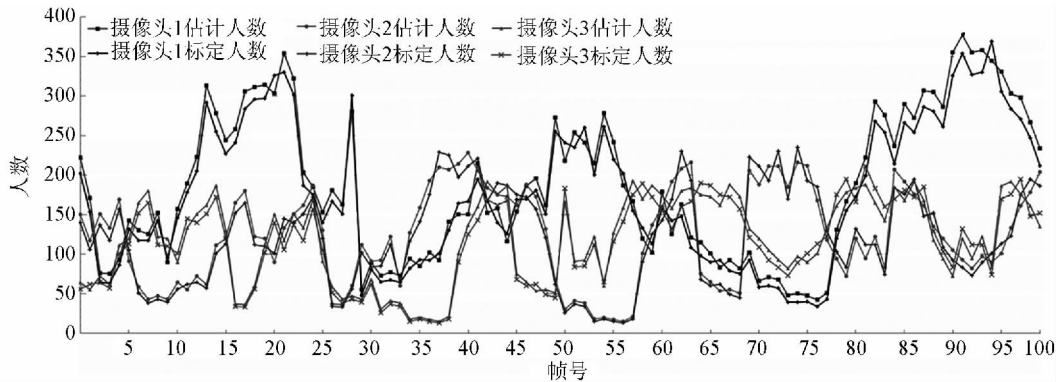


图6 低密度人群模型测试结果

Fig. 6 Test results of low crowd density estimation model

综合考虑准确率与算法效率, 根据试验结果, 将图像灰度级定为 16 级, 距离设定为 8, 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。3 个监控场景面积分别为 271.78, 150.62 和 135.53 m², 当人群密度大于 1.53 人/m² 时为中密度人群, 模型选取阈值分别为 416 人、230 人和 207 人, 则像素数阈值分别为 21 354, 11 926 和 10 760。选取中、高、很高、极高 4 类人群图像各 40 幅作为训练样本, 同时选取这些

作为测试样本。选择其他 2 个监控场景 4 类人群图像各 40 幅, 用来验证 SVM 分类器的可靠性(表 3)。可知, 传统基于纹理特征的 SVM 分类(T-SVM, Traditional SVM) 精度略高于笔者提出的可跨摄像机 SVM 方法(C-SVM, Cross-camera SVM), 但需对各场景进行训练, 而 C-SVM 只需对某监控场景训练即可将模型应用于其他场景。若将 T-SVM 训练的某场景分类器应用于其他场景, 其分类精度不足 50%, 不能满足应用需求。

表 3 试验分类结果

Table 3 Experimental classification results

测试样本	中(40)		高(40)		很高(40)		极高(40)		准确率/%		
	C-SVM	T-SVM	C-SVM	T-SVM	C-SVM	T-SVM	C-SVM	T-SVM	C-SVM	T-SVM	
摄像头 1	中	38	40	1	2	0	0	0	0	95.63	97.50
	高	2	0	39	38	2	1	0	0		
	很高	0	0	0	0	38	39	2	1		
	极高	0	0	0	0	0	0	38	39		
摄像头 2	中	39	40	3	1	0	0	0	0	94.37	96.88
	高	1	0	37	39	1	2	0	0		
	很高	0	0	0	0	38	38	3	2		
	极高	0	0	0	0	1	0	37	38		
摄像头 3	中	38	39	1	1	0	0	0	0	95.00	96.25
	高	2	1	39	39	2	2	0	0		
	很高	0	0	0	0	37	38	2	2		
	极高	0	0	0	0	1	0	38	38		

2.3 应用案例

基于可跨摄像机的人群密度估计模型, 研发了区域人群密度时空格局实时感知系统, 并成功将其应用于“2012 年江苏·秦淮灯会”的安保工作。系统界面如图 7 所示, 左上角区域为各监控探头的人群状态, 当人群密度为极高时显示为黑色, 表明该场景人群密度过高, 系统将自动报警提示该探头监控

区域易发生拥挤踩踏等突发事件。地图窗口显示监控探头及人群密度的空间分布状态, 可实时监测该区域人群的空间分布格局。用户可实时监测各监控场景人群数量的动态变化, 实时查看视频场景, 根据具体情况疏导人群, 以避免突发公共事件的发生。同时, 能够对监控区域的历史数据进行时空分析, 图中右下窗口为相应监控探头在某时段人群数量的动态变化过程分析, 可直观掌握区域人流的时空特征, 为安保人员配置与布控提供可靠依据。

4 结论

笔者认为监控摄像机内外参数差异造成的人群图像尺度多样性, 是导致人群密度估计模型具有场景依赖性的内在原因, 通过研究得出以下结论:

1) 利用 GIS 技术将视频监控图像数据映射在 GIS 空间坐标系, 可统一各监控场景人群图像的参考系, 解决不同监控场景间人群图像的空间尺度多样化问题。

2) 在 GIS 空间构建的低密度和高密度人群估计模型, 可实现跨摄像机应用, 无需为各监控场景构

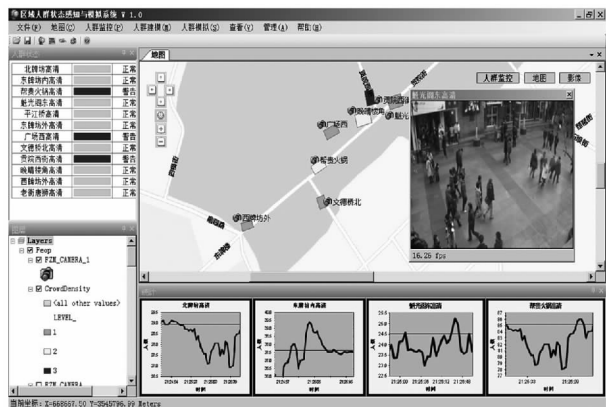


图 7 系统界面

Fig. 7 System interface

建模型。

3) 笔者基于提出的模型构建方法,将视频监控系统与GIS有机集成,研发了人群密度实时感知系

统,只需选取某监控场景进行模型构建,即能实现对监控区域内各监控场景的人群密度估计。

参考文献

- [1] 徐方,魏东,魏星. 公众聚集场所人群疏散基础数据的分析[J]. 中国安全科学学报,2008,18(4): 137-145.
XU Fang, WEI Dong, WEI Xing. Analysis on the original data of crowd evacuation in public gathering places[J]. China Safety Science Journal, 2008, 18(4): 137-145.
- [2] 宋宏权,刘学军,闫国年,等. 区域人群状态的实时感知监控[J]. 地球信息科学学报,2012,14(6): 686-692.
SONG Hong-quan, LIU Xue-jun, LÜ Guo-nian, et al. Real-time monitoring for the regional crowds status[J]. Journal of Geo-Information Science, 2012, 14(6): 686-692.
- [3] Davies A, Yin J, Velastin S. Crowd monitoring using image processing[J]. Electronics & Communication Engineering Journal, 1995, 7(1): 37-47.
- [4] Lin Sheng-fuu, Chen Jaw-yeh, Chao Hung-xin. Estimation of number of people in crowded scenes using perspective transformation[J]. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE, 2001, 31(6): 645-654.
- [5] Marana A, Velastin S, Costa L, et al. Automatic estimation of crowd density using texture[J]. Safety Science, 1998, 28(3): 165-175.
- [6] 肖坦,杨栩,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.
XIAO Tan, YANG Xu, LI Huang-huang. Railway passenger station crowd density estimation algorithm based on video surveillance[J]. Railway Signalling & Communication, 2010, 46(8): 80-82.
- [7] Hussain N, Yatim H, Hussain N L, et al. CDES: A pixel-based crowd density estimation system for Masjid al-Haram[J]. Safety Science, 2011, 49(6): 824-833.
- [8] 吴晟,葛万成. 基于可变形框的人群密度估计算法[J]. 通信技术,2011,44(10): 63-65.
WU Sheng, GE Wan-cheng. Crowd density estimation algorithm based on changeable rectangle[J]. Communications Technology, 2011, 44(10): 63-65.
- [9] Rittsche J, TU P, Krahnstoeve N. Simultaneous estimation of segmentation and shape[C]. Proceeding of Computer Vision and Pattern Recognition 2005: 486-493.
- [10] 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.
MA Wen-hua, HUANG Lei, LIU Chang-ping. Crowd density classification based on confidence analysis[J]. Pattern Recognition and Artificial Intelligence, 2011, 24(1): 30-39.
- [11] Cho S, Chow T. A fast neural learning vision system for crowd estimation at underground stations platform[J]. Neural Processing Letters, 2004, 10(2): 111-120.
- [12] Kilambi P, Ribnick E, Joshi A, et al. Estimating pedestrian counts in groups[J]. Computer Vision and Image Understanding, 2008, 10(1): 43-59.
- [13] Zhao Tao, Nevatia R, Wu Bo. Segmentation and tracking of multiple humans in crowded environments[J]. IEEE Trans. on Pattern Analysis Machine Intelligence, 2008, 30(7): 1198-1211.
- [14] 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.
YANG Lin, MIAO Zhen-jiang. Design and implemented of crowd abnormal behavior detection system[J]. Railway Computer Application, 2010, 19(7): 37-41.
- [15] 王波,姚宏宇,李弼程. 一种有效的基于灰度共生矩阵的图像检索方法[J]. 武汉大学学报·信息科学版,2006,31(9): 761-764.
WANG Bo, YAO Hong-yu, LI Bi-cheng. An efficient image retrieval approach based on generalized image co-occurrence matrix[J]. Geomatics and Information Science of Wuhan University, 2006, 31(9): 761-764.
- [16] 常利利,马俊,邓中民,等. 基于灰度共生矩阵的织物组织结构差异分析[J]. 纺织学报,2008,29(10): 44-46.
CHANG Li-li, MA Jun, DENG Zhong-min, et al. Study on identification of fabric texture based on gray-level co-occurrence matrix[J]. Journal of Textile Research, 2008, 29(10): 44-46.
- [17] 占文凤,陈云浩,周纪,等. 基于支持向量机技术的背景城市热岛模拟:热岛强度空间格局曲面模拟及其应用[J]. 测绘学报,2011,40(1): 96-103.
ZHAN Wen-feng, CHEN Yun-hao, ZHOU Ji, et al. Spatial simulation of urban heat island intensity based on the support vector machine technique: a case study in Beijing[J]. Acta Geodaetica et Cartographica Sinica, 2011, 40(1): 96-103.

[18] 陈杰, 邓敏, 肖鹏峰, 等. 结合支持向量机与粒度计算的高分辨率遥感影像面向对象分类 [J]. 测绘学报, 2011, 40(2): 135 - 141.
 CHEN Jie , DENG Min , XIAO Peng-feng , et al. Object-oriented classification of high resolution imagery combining support vector machine with granular computing [J]. Acta Geodaetica et Cartographica Sinica , 2011 , 40(2) : 135 - 141.

[19] Au S , Ryan M , Carey M , 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: 1 - 993.



作者简介: 宋宏权 (1986 -) 男, 河南民权人, 博士, 讲师, 主要从事遥感与 GIS 应用、空间分析等方面的研究。E-mail: hqsong@henu.edu.cn.



作者简介: 刘学军 (1965 -) 男, 陕西合阳人, 博士, 教授, 博士生导师, 主要从事 DEM 及其地形和 GIS 空间分析、视频 GIS、空间数据不确定性等方面的研究。E-mail: liuxuejun@njnu.edu.cn.

《中国安全科学学报》再次被收录为 中国科技核心期刊



《2013 版中国科技期刊引证报告(核心版)》以“中国科技论文与引文数据库”(CSTPCD)为基础,在各个学科领域中选择学术质量较高、影响较大的重要期刊作为来源期刊,即“中国科技核心期刊”。报告汇集了 20 余项文献计量学指标,并根据科学计量学原理计算所有来源期刊的综合评价总分和排名。

《中国安全科学学报》再次被收录为中国科技核心期刊,同时,《中国科技期刊引证报告(核心版)》显示 2012 年《学报》的主要指标如下:核心总被引频次为 1 865,在 5 种安全科学技术类期刊中排名第一;核心影响因子为 0.604,在 5 种安全科学技术类期刊中排名第三;综合评价总分为 69.4,在 5 种安全科学技术类期刊中排名第一;学报扩散指标为 74.00,学科影响指标为 1.00。