

引文格式: 孔令彦,关丽,丁燕杰,等:空间大数据支持下的城市区域评估及其关键技术[J].测绘通报,2017(8):100-105.DOI: 10.13474/j.cnki.11-2246.2017.0263.

空间大数据支持下的城市区域评估及其关键技术

孔令彦,关丽,丁燕杰,陈思
(北京市测绘设计研究院,北京 100038)

摘要:针对人口拥挤、交通拥堵、环境污染、住房困难等一系列“城市病”问题,为了客观、定量评估城市现状,找到“城市病”根源,本文面向城市区域评估,利用空间大数据全面、客观、准确的特点,对城市进行全方位的评估,从人口、资源、环境、住房、公共设施、道路交通、经济发展及城市安全等多方面为治理“城市病”提供基础数据。本文从城市区域评估的内涵入手,设计了城市区域评估的总体流程,建立了全面的评估体系,并对实现城市区域评估的关键技术进行了阐述。通过城市区域评估的方式,对城市现状客观扫描,将使“城市病”的病重程度得到客观度量,是根治“城市病”的首要条件。

关键词:城市病;城市评估;城市现状评估;空间大数据

中图分类号: P208

文献标识码: A

文章编号: 0494-0911(2017)08-0100-06

Urban Examination Evaluation System and Its Key Technologies Based on Spatial Big Data

KONG Lingyan, GUAN Li, DING Yanjie, CHEN Si
(Beijing Institute of Surveying and Mapping, Beijing 100038, China)

Abstract: For “urban disease” problem, such as overcrowding, traffic congestion, environmental pollution, housing difficulties and so on, in order to assess urban situation objectively and quantitatively and find the causes of “urban disease”, this paper presents concept of “urban examination” based on spatial big data in spatial information domain. Hoping to play the role of comprehensive, objective and accusation of spatial big data, to evaluate urban in full range is important, such as population, resources, environment, housing, public facilities, road transportation, economic development and urban safety. Its results will provide basic data for governance of “urban disease”. In this paper, concept and connotation of “urban examination” is presented firstly, and overall process of “urban examination” is designed, and comprehensive assessment system is established. Finally key technologies to retaliate “urban examination” are described. By the way of “urban examination”, to measure the illness degree of urban objectively is the first condition of position of “urban disease”.

Key words: urban disease; urban examination; urban situation assessment; spatial big data

城市是一个巨系统,涉及社会、经济、生活等各方面。按照经济学解释,城市是一个坐落在有限空间地区内的各种经济市场(包括住房、劳动力、土地、运输等)相互交织在一起的网络系统。近些年,随着城市规模的日益扩大,现代大中城市发展过程中暴露出诸多的问题,统称为“城市病”(urban disease)。所谓“城市病”,是常常与人口在大城市过度聚集相伴的若干社会管理和公共服务问题的统称。如果以城市发展愿景作为参照,按聚集于发达、宜居与文明作为衡量标准,可以将城市病的具体表现形式总结为人口拥挤、交通拥堵、环境污染、住房困难、资源供应紧张、安全风险增加、城市秩序恶化、贫富差距过大、存在低收入人群成片聚集区及城

市缺少文化特征等。现阶段我国城市病较为突出地表现为人口拥挤、交通拥堵、环境污染、住房困难等。根治“城市病”是一个复杂而漫长的过程,需要社会多部门、多领域发挥专长协调治理。但目前亟须解决的问题是:城市哪里生病了,病的有多重,这个量化的指标需要定期给出来,这样专家才能针对性地给出治理方案。

正是基于此,本文立足城市区域评估,利用空间大数据全面、客观、准确的特点,对城市进行全方位的评估,找到“城市病”根源,为专家治理提供基础数据。城市生态系统作为有机体,城市区域评估是对城市进行的全方位或专项评估。针对城市现状评估方法方面国内外鲜有专门文献进行论述。与此类

收稿日期: 2016-12-12

基金项目: 北京市科学计划课题(Z151100002115057)

作者简介: 孔令彦(1979—),男,硕士,高级工程师,主要从事地理信息服务相关工作。E-mail: kongly@163.com

似的工作,如规划实施评估、宜居城市、健康城市,由于侧重点不同,评估结果往往具有片面性;同时,由于数据条件和技术条件的限制,使相关工作只进行了定性评估,没有实现真正意义的定量评估。如规划实施评估主要是将规划项目实施后的现状与规划成果进行对比,评估规划实施效果^[1-5]。宜居城市、健康城市方面主要是借助多层次指标对城市的宜居方面进行评价^[6]。随着大数据分析技术的发展,对城市进行客观评估的数据源及分析手段已趋成熟,具备对城市进行客观评估的条件,评估结果也将为城市多部门多学科提供基础服务。本文从城市区域评估的内涵入手,设计了城市区域评估的总体流程,建立了评估体系,并对实现城市区域评估的关键技术进行了阐述,真正实现客观、定量地评估城市现状。

1 城市区域评估的特点

1.1 概念

城市区域评估是指针对城市发展过程中一系列社会管理和公共服务问题而进行的分析和评价,是城市现状评估的表现形式。

1.2 特点

城市是一个有机体,其健康发展对于城市中人的影响很大,只有城市健康了,城市中的人才能健康。城市区域评估具有空间位置、定期性和动态性等特征。

(1) 城市区域评估具有空间位置特征。由于城市区域评估是基于空间大数据进行的城市现状评估研究,因此研究成果具备空间位置特征,即解决了发生地点的问题,这也是城市区域评估区别于其他评估的根本特征,能够清晰表达出城市空间的概念。

(2) 城市区域评估具有定期性。定期对城市进行评估,实现预防为主的目标,从城市中人的行为方式和致病原因上发现影响城市健康的因素;定期对城市进行评估,实现对城市的连续监测,可以了解环境、经济、社会等有关因素对城市有机体的影响;定期对城市进行评估,了解城市的健康状态,采取最佳强度,提高城市抗病能力,及早发现危险因素。

(3) 城市区域评估具有动态性。评估区域位置、范围大小、评估时间、评估间隔的不同,都会对评估体系内指标内容的选择有影响,这些动态的变化主要是实现城市区域评估的有效性,同时保证城市区域评估结果的客观描述和安全隐患的提示。从评估成本上讲,从长远考虑,城市区域评估的支出远远小于后期的治理费用,就像目前北京的雾霾污染治

理,如果早预警早治理,治理成本估计会小很多。

2 城市区域评估总体流程设计

城市区域评估的总体流程分为以下 4 个步骤,即输入区域数据、确定评估指标、制定健康指数、输出评估报告。

(1) 输入区域数据。区域数据是城市区域评估的血液,因此数据的准确性和全面性很重要。数据种类主要包括业务运营数据、普查数据、监控数据、社会网络数据、主动感知数据和空间基础数据等几类。其中,业务运营数据包括公交 IC 刷卡数据、水电煤气数据、业务审批数据、出租车 GPS 轨迹数据等;普查数据包括人口普查、经济普查、地理国情普查数据等;监控数据包括视频监控、交通监控、环境监控等;社会网络数据包括微博、论坛数据等;主动感知数据包括关于温度、湿度、PM_{2.5} 等环境的感知数据、手机定位数据等;空间基础数据包括空间电子地图、地形图、三维数据等。这些数据数量庞大、结构多样,利用技术难度较大,而空间化技术可以在一定程度上方便它们的利用。城市大数据空间化后,信息变得直观,并拓展了分析和利用途径,可以更加有效地支持信息发掘和知识发现。

(2) 确定评估指标。根据城市发展阶段与实际情况,在有效对接城市发展目标的情况下,建立科学、动态的评价城市的评估指标体系。

(3) 制定健康指数。参考宜居城市、健康城市等体系,制定符合城市区域的健康指数。同时,在世界范围内,一些与所研究城市地理纬度和气候相当、位置相似的发达国家的城市,在城市健康发展方面,为我们提供了有效的经验,如东京、伦敦;而一些城市,如新加坡堪为世界的范本。以这些城市为样板,根据评估指标制定出合理的可参照的健康指数。

(4) 测出评估报告。针对历史数据和现状数据,开发出针对性的数学模型,进行指标测定,将城市分析的数值与评价指标所参照的指标数值进行对比,找出差异,形成评估报告,初步分析城市发展的问题和弊端。

3 城市区域评估体系的构建

城市区域评估体系的构建分为评估项目的选取、分类指标的构建两个组成部分。

3.1 城市区域评估项目的选取

本研究根据世界卫生组织(WHO)提出的健康城市的具体标准和宜居城市的研究成果,研究北京市城市区域评估的项目选取问题^[7-8]。WHO 整理公

布了健康城市的 10 项具体标准及其内容。健康城市主要涉及 7 大领域(即政治领域(领导参与、政策制定)、经济领域(就业、收入、住房)、社会领域(文化、教育、福利、保障)、生态环境(生态平衡、污染控制和资源保护)、生物、化学和物理因素(医疗卫生技术及其服务和营养供给及其安全卫生等)、社区生活(健康的社区邻里关系、文明的风尚等)、个人行为(心理卫生、行为矫正和健康生活方式的鼓励等))。

结合城市实际情况和矛盾比较突出的方面,最终确定了城市区域评估项目,包括 8 个专项,分别是人口、资源、环境、住房、交通、公共设施、经济发展和城市安全,见表 1。这 8 个专项中,其中 4 项属于空间信息领域关联度大的项目,即住房、交通、公共设施和城市安全,也就是说这 4 个项目通过空间数据方式可以实现更好的表达。另外 4 个项目数据关联度小的项目,即人口、资源、环境、经济发展,但考虑到城市区域评估的全面性,也需要加入其中。

3.2 城市区域评估分类指标的构建

在城市区域评估体系中,每个专项具体细分为

若干评估指标,以人口为例,说明城市区域评估的指标体系构成,见表 2。在分类指标的设计上,考虑指标的可操作性,选取的指标多数能从政府公开数据中计算获得,另外一部分则需要依靠评估区域的社会感知数据计算获得,如手机信令数据或交通刷卡数据等。城市区域评估分类指标的构建还要考虑其动态性,因为评估的区域特点不同,需要指标能够做到动态调整,这无形中增加了指标设计的难度。

表 1 城市区域评估的项目设置

序号	项目	内容
1	人口	人口、就业等
2	资源	水资源、能源资源、土地资源等
3	环境	生态、大气、水等
4	住房	住房短缺、保障住房等
5	交通	交通拥堵、通勤时间等
6	公共设施	教育、医疗、福利等
7	经济发展	产业、产出等
8	城市安全	预防灾害、城市安全、食品安全等

表 2 人口专项的分类指标设计

专项	指标	单位	模型方法	指标类型
人口	人口总数	万人	/	定量
	人口密度	%	万人/平方公里	定量
	年龄金字塔结构	%	不同年龄段人口数/总人口数	定量
	性别结构	%	男(女)/总人口数	定性
	职业结构	%	不同教育程度人口数/总人口数	定性
	教育结构	%	不同行业人口数/总人口数	定性
	高中阶段毛入学率	%	[某学年本市生源高中在学人数÷某年本省籍(15-18)岁人口数]×100%	定量
	高等教育毛入学率	%	[高等教育在学人数÷(18-22)岁人口数]×100%	定量
	户籍人口小学入学数	人	/	定量
	户籍人口高中入学数	人	/	定量
	外地人口小学入学数	人	/	定量
	外地人口中学入学数	人	/	定量
	外地人口高中入学数	人	/	定量
	就业密度	人	(居民+就业)/总面积	定量
	登记就业率	%	指城镇登记从业人数与城镇从业人数和城镇登记失业人数之和的百分比	定量
	登记失业率	%	/	定量

3.3 城市区域评估指标的成果

为了验证城市区域评估指标的可行性,研究组对 3 个街道共 64 km² 进行了“评估”评估,得到诸多评估成果,这里只选取其中部分成果说明验证过程。在人口专项中,首先将人口数量与空间分布相结合,形成研究区域 64 km² 的人口热力图,如图 1 所示。该图能够反映研究区内的人口分布均匀度情况,进一步分析分布原因,可以指导区域内产业或人口布局。

在此基础上,研究流动人口对于研究区域人口分布的贡献率,计算了区域内 3 个街道的流动人口 2010—2014 年度数量变化,如图 2 所示。从流动人口的数量和 5 年的变化趋势分析,街道 1 在 2014 年度流动人口大幅度增加,需要进一步了解背后的原因;街道 2 流动人口 5 年来变化不大,有下降趋势;街道 3 近 5 年流动人口数量先降低后升高,需要进一步了解区域情况。

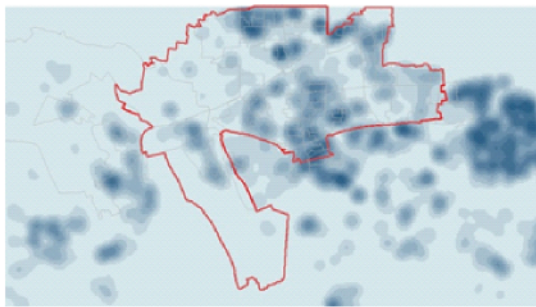


图 1 研究区域人口分布热力图

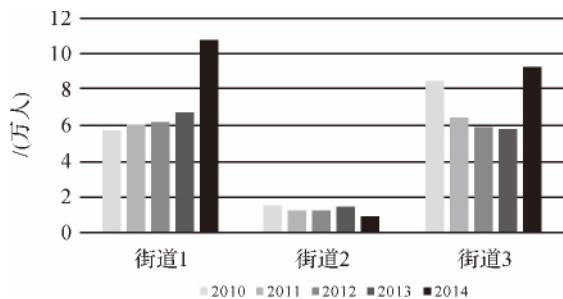


图 2 研究区 3 个街道的流动人口 2010—2014 年度数量变化

如在公共设施专题中,基于教育科研设施的分布情况(如图 3 所示)分析教育科研设施实际服务范围有效性,如图 4 所示。本次针对服务半径有效性分析是运用 GIS 基于丰台区现有道路分布情况对教育科研设施的实际服务半径和服务范围与居住用地相叠加并与相关规划标准所给予的服务半径进行对比得到的结果。评价过程中,服务半径大于 1000 m 为较好,800—1000 m 之间为一般,低于 600—800 m 为较差,低于 600 m 为差。从图 4 可知,有约 40% 区域距离教育科研设施过远,另外该区域存在教育科研设施分布不均现象。

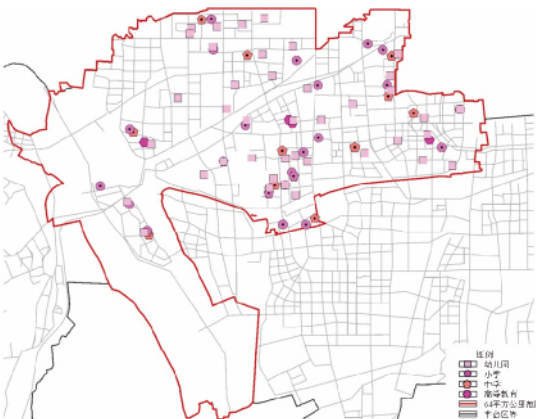


图 3 教育科研设施分布

4 城市区域评估的关键技术

空间大数据种类多样,为了城市区域评估的顺

利实施,需要多类技术进行支撑,其中包括大数据分析技术、并行处理技术和数据可视化表达技术等。

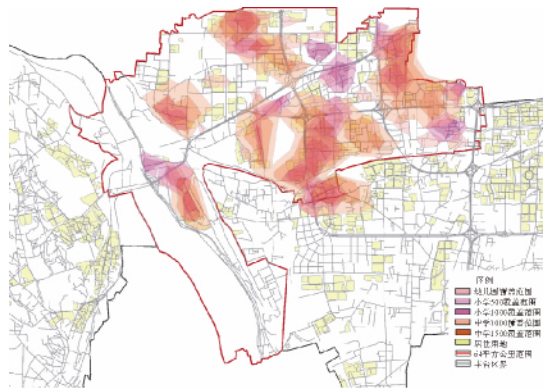


图 4 教育科研设施实际服务范围分析

4.1 大数据分析技术

大数据时代对于数据分析、管理都提出了不同程度的新要求,许多传统的数据分析技术和数据库技术已经不能满足现代数据应用的需求。为了给大数据处理分析提供一个性能更高、可靠性更好的平台,Doug Cutting 模仿 GFS,开发了一个云计算开源平台 Hadoop^[9-10]。目前 Hadoop 已经发展为一个包括分布式文件系统(HDFS)、分布式数据库(HBase、Cassandra)及数据分析处理(MapReduce)等功能模块在内的完整生态系统(Ecosystem),现已成为目前最流行的大数据处理平台^[11-13]。在这个系统中,以 MapReduce 算法为计算框架,HDFS 是一种类似于 GFS 的分布式文件系统,可以为大规模的服务器集群提供高速度的文件读写访问。HBase 是一种与 BigTable 类似的分布式并行数据库系统,可以提供海量数据的存储和读写,而且兼容各种结构化或非结构化的数据。如图 5 所示。

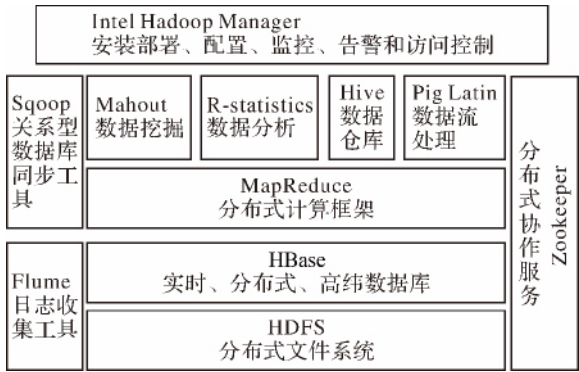


图 5 Hadoop 组件结构

4.2 并行处理技术

MapReduce 技术由 Google 公司于 2004 年提出,

作为一种典型的数据批处理技术被广泛应用于数据挖掘、数据分析、机器学习等领域,已经成为大数据处理的关键技术^[14]。MapReduce 的数据分析流程如图 6 所示。由图 6 可以看出,MapReduce 系统主要由两个部分组成: Map 和 Reduce。MapReduce 的核心思想在于“分而治之”,也就是说,首先将数据源分为若干部分,每个部分对应一个初始的键值 (Key/Value) 对,并分别交给不同的 Map 任务区处理,Map 对初始的键值对进行处理,产生一系列中间

结果 Key/Value 对,MapReduce 的中间过程 Shuffle 将所有具有相同 Key 值的 Value 值组成一个集合传递给 Reduce 环节; Reduce 接收这些中间结果,并将相同的 Value 值合并,形成最终的较小 Value 值的集合。MapReduce 系统简化了数据的计算过程,避免了数据传输过程中大量的通信开销,使得 MapReduce 可以应用于多种实际问题的解决方案中,在各个领域均有广泛的应用。

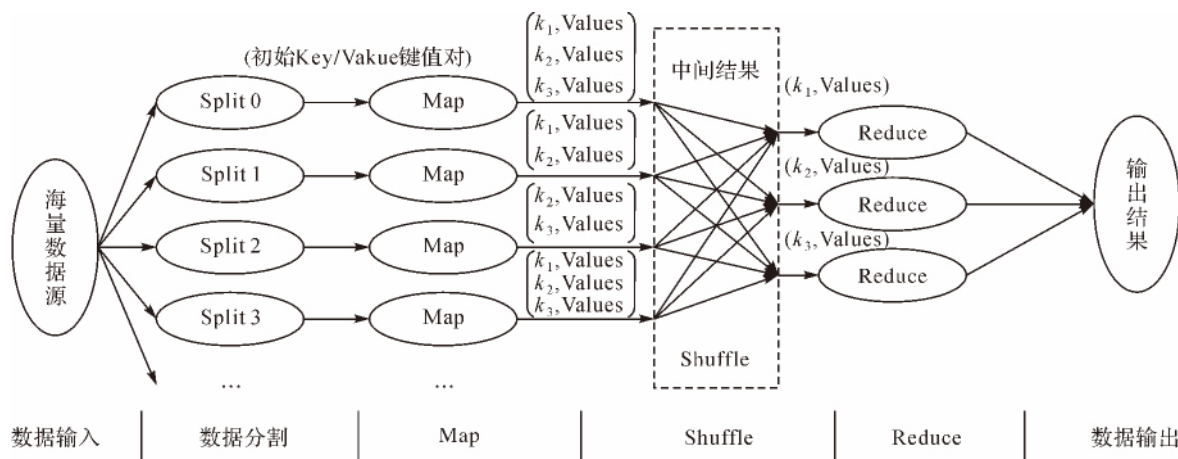


图 6 MapReduce 的数据分析流程

4.3 数据可视化表达技术

城市大数据由于数据量极其庞大,要直观反映其中的内容是非常困难的。而空间化后,用二、三维地图反映其空间分布,用颜色、线型反映信息的内容(如 GDP、人口数量、交通量等),繁杂的信息就能变得有序易懂。数据可视化(data visualization)技术是指运用计算机图形学和图像处理技术,将数据转换为图形或图像在屏幕上显示出来,并进行交互处理的理论、方法和技术。由于图形化的方式比文字更容易被用户理解和接受,数据可视化就是借助人脑的视觉思维能力,将抽象的数据表现成为可见的图形或图像,从而帮助人们发现数据中隐藏的内在规律。可视化技术能够迅速和有效地简化与提炼数据流,帮助用户交互筛选大量的数据,有助于用户更快地从复杂数据中得到新的发现。用形象的图形方式向用户展示结果,已作为最佳结果展示方式之一率先被科学与工程计算领域采用。常见的可视化技术有原位分析(insitu analysis)、标签云(tag cloud)、历史流(history flow)、空间信息流(spatial information flow)、不确定性分析等。可以根据具体应用需要选择合适的可视化技术^[15]。可视化技术作为解释大数据最有效的手段之一,最初是被科学与

计算领域运用,它对分析结果的形象化处理和显示,在很多领域得到了迅速而广泛的应用。如百度迁徙应用数据可视化表达技术所形成的北京国贸地区通勤情况就是很好的例证,如图 7 所示。



图 7 北京国贸地区通勤动态情况

5 结 语

本文立足空间信息领域,基于空间大数据提出了城市区域评估的概念,利用空间大数据全面、客观、准确的特点,对城市进行全方位的评估,从人口、资源、环境、住房、公共设施、道路交通、经济发展及城市安全等多方面为领域专家治理“城市病”提供基础数据。在城市区域评估指标体系中,分类指标

的动态性和可操作性是需要充分考虑的,有些指标的设计往往由于数据无法获取而被迫放弃。同时,城市区域评估是基于空间大数据,在数据表达与数据分析方法上,也区别与一般的社会经济数据评价,“以图代数”的方式大大提高了评估结果的直观性,也直观地体现了城市区域评估的空间位置特征。接下来,笔者将继续扩大研究区范围和类型,一方面完善评估指标体系,另一方面,推进城市区域评估的产业化进程,为“城市病”提供全面、可靠的数据评估结果。

参考文献:

- [1] 平宗莉,闫整. 国外规划实施评价理论与方法及其在国内应用的研究进展[J]. 山东国土资源, 2010, 26(2): 33-36.
- [2] 石莹怡. 城市规划量化监测与评估的体系框架研究[D]. 广州: 华南理工大学, 2011.
- [3] 何灵聪. 基于动态维护的城市总体规划实施评估方法和机制研究[J]. 规划师, 2013, 29(6): 18-23.
- [4] 郑德高,闫岩. 实效性和前瞻性: 关于总体规划评估的若干思考[J]. 城市规划, 2013, 37(4): 37-42.
- [5] 邢谷锐,蔡克光. 城市总体规划实施效果评估框架研究[J]. 城市问题, 2013, 6(6): 23-227.
- [6] 吴良镛. 人居环境科学导论[M]. 北京: 中国建筑工业出版社, 2001.
- [7] TIMMER V, SEYMOAR N K. The World Urban Forum 2006: Vancouver Working Group Discussion Paper[C]// The Livable City. [S.l.]: [s.n.], 2006.
- [8] 李业锦,张文忠,田山川,等. 宜居城市的理论基础和评价研究进展[J]. 地理科学进展, 2008, 27(3): 101-109.
- [9] 刘智慧,张泉灵. 大数据技术研究综述[J]. 浙江大学学报(工学版), 2014, 48(6): 957-972.
- [10] 宫夏屹,李伯虎,柴旭东,等. 大数据平台技术综述[J]. 系统仿真学报, 2014, 26(3): 489-496.
- [11] 曾元武,陈泽鹏,方晓乐,等. 大数据时代下地理信息公共平台建设展望[J]. 测绘通报, 2015(11): 84-87.
- [12] 肖建华,王厚之,彭清山,等. 地理时空大数据管理与应用云平台建设[J]. 测绘通报, 2016(4): 38-42.
- [13] 王凯,曹建成,王乃生,等. Hadoop 支持下的地理信息大数据处理技术初探[J]. 测绘通报, 2015(10): 114-117.
- [14] 潘巍,李战怀. 大数据环境下并行计算模型的研究进展[J]. 华东师范大学学报(自然科学版), 2014, 5(9): 43-53.
- [15] 吉根林,赵斌. 面向大数据的时空数据挖掘综述[J]. 南京师大学报(自然科学版), 2014, 37(1): 1-7.

(上接第 70 页)

- [3] ZOU Q, CAO Y, LI Q, et al. CrackTree: Automatic Crack Detection from Pavement Images [J]. Pattern Recognition Letters, 2012, 33(3): 227-238.
- [4] XU B. Automatic Inspection of Pavement Cracking Distress [J]. Journal of Electronic Imaging, 2006, 15(1): 185-188.
- [5] LI Q, ZOU Q, ZHANG D, et al. FoSA: F* Seed-growing Approach for Crack-line Detection from Pavement Images [J]. Image and Vision Computing, 2011, 29(12): 861-872.
- [6] KOCH C, BRILAKIS I. Pothole Detection in Asphalt Pavement Images [J]. Advanced Engineering Informatics, 2011, 25(3): 507-515.
- [7] SUN L, ZHANG Y M. Weighted Neighborhood Pixels Segmentation Method for Automated Detection of Cracks on Pavement Surface Images [J]. Journal of Computing in Civil Engineering, 2016, 30(2): 11.
- [8] 钱彬,唐振民,徐威,等. 子块鉴别分析的路面裂缝检测[J]. 中国图象图形学报, 2015, 20(12): 1652-1663.
- [9] KOCH C, GEORGIEVA K, KASIREDDY V, et al. A Review on Computer Vision Based Defect Detection and Condition Assessment of Concrete and Asphalt Civil Infrastructure [J]. Advanced Engineering Informatics, 2015, 29(2): 196-210.
- [10] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object Detection with Discriminatively Trained Part-based Models [J]. IEEE Transactions on Software Engineering, 2010, 32(9): 1627-1645.

(上接第 79 页)

- [10] 郭海涛,刘智,张保民. 基于遗传算法的快速影像匹配技术的研究[J]. 测绘学院学报, 2001, 18(51): 20-22.
- [11] 熊兴华. 遗传算法及其在影像处理与分析中应用[J]. 测绘学报, 2004, 33(2): 74-81.
- [12] 罗代建. 遥感影像配准与融合技术研究[D]. 成都: 电子科技大学, 2009.
- [13] 常青,张斌,邵金玲. 基于 SIFT 和 RANSAC 的特征图像匹配方法[J]. 华东理工大学学报(自然科学版), 2012, 38(6): 97-101.
- [14] 陈华,邓喀中,张以文,等. 结合 SIFT 和 RANSAC 算法的 InSAR 影像配准[J]. 测绘通报, 2015(12): 30-33.
- [15] KE Y, SUKTHANKAR R, HUSTON L. Efficient Near Duplicate Detection and Sub-image Retrieval [C]// Proceedings of ACM Multimedia Conference. [S.l.]: ACM, 2004.