

大数据时代的城市计算

特邀编辑: 陈宝权 曾 琼
山东大学

关键词: 城市计算 大数据 智慧城市 智能决策

城市计算的兴起

随着城市化进程的不断推进,我国人口和资源迅速向城市集中,城市化比例已经超过了50%,成为世界上城市人口最多的国家。城市化虽然提高了人们的生活水平,但也使城市的社会环境变得日趋复杂:空间结构复杂、人口密度大、流动性大、社会关系庞杂、交通拥堵严重、事故和恶性事件频发。复杂的社会环境已成为城市发展的障碍,给城市管理和公共安全防护带来极大的挑战。除了管理意识和政策等方面的因素外,还在于管理决策者对城市复杂事件所涉及的对象、行为和状况、时间、地点、起因等关键要素缺乏及时、全面、精准的了解。随

着时间的推移,这些要素往往具有高度的动态性,现有的信息技术缺乏对这些要素进行及时捕获与感知的机制和手段,从而导致决策的不准确、不及时和服务的不到位、不便捷。这些问题已成为制约城市健康、可持续发展的瓶颈,如何有效解决迫在眉睫。

“城市计算”应运而生。随着大数据时代的来临,多样化、群体化的信息感知模式使得城市环境下的信息空间、物理世界和人类社会逐步融合、深度关联与互动,形成了一个具有闭环反馈回路的有机整体,即“信息-物理-社会”三元空间。城市环境下的三元空间融合系统是一个极其复杂、高度动态的系统,既涉及到信息的获取、处理、传输、分析、理解、反馈等环节,又涉及到人在回路中的分析推理和决策控制,如图1所示。

城市计算是一门交叉学科,是通过对城市异构海量数据的整合、分析和挖掘,实现城市环境下三元空间的智能融合,为现代城市复杂事件的综合感知、分析理解以及决策服务提供计算的模型。2009年,IBM公司提出了“智慧地球”的概念,城市数字化与智能化相关理论与技术的研究备受国际学术界的重视。《科学》(Science)杂志在2012年接连发表了题目为《你的家有多智能(How Smart is Your Home)?》^[1]与《你所在的城市有多智能(How Smart is Your City)?》^[2]的文章,指出了城市智能化的重要性,并指出,虽然通过各种信息采集设备可以实现对城市环境的数据获取,但要达到城市或家庭的智能化,关键是要建立一个统一的智能化支持环境,将各种信息无缝地整合在一起,形成一个有机的整



图1 “信息-物理-社会”三元空间示意图

体。近年来,城市计算在我国亦受到广泛关注,得到了政府的大力支持。国家在“十二五”科技规划中提出城镇化、工业化、信息化三化融合,着力发展新一代信息技术、现代服务业、智能城市等促进城市和城镇化可持续发展的科学与技术。《国家中长期科学和技术发展规划纲要(2006-2020年)》中特别提到,要重点研究城市基础数据获取与更新技术、城市多元数据整合与挖掘技术、城市多维建模与模拟技术、城市动态监测与应用关键技术、城市应急和联动服务关键技术等。

大数据时代的城市计算

现阶段我国城市计算仍然存在信息资源不完整、信息的横向整合和贡献严重不足、难以建立有效的智能分析与决策支持系统等问题。对于智能分析与决策支持系统的建立,很大程度上依赖于将数据的融合分析与信息的展示和现实场景无缝连接,通过对城市时空的增强及呈现为城市管理分析与决策提供更高质量的支持。

基于上述思想,笔者认为城市计算主要面临四个挑战:(1) **缺乏城市数据之间的时空精准耦合**。城市数据往往涉及不同领域(比如电信行业、电力行业、城市监管等),采集方式不同(比如监控视频、激光扫描、图像传输等),复杂多样、极度分散,内容混杂、不确定、不完整甚至相互矛盾。(2) **缺乏对城市数据智能性的分析**。城市数据以多模态的形式存在,具有时空跨度大、内容关联缺失、全局稀疏与局部冗余并存等特点。例如,A车辆可能在不同时刻被多个路口的监控拍摄,但由于监控仅覆盖局部区域,因此全局信息稀疏。(3) **缺乏城市数据信息处理的时效性**。城市活动具有高度动态性,演变过程快速而复杂,管理者难以做出及时的决策、响应和服务。(4) **缺乏城市情境关联式的决策支持**。城市环境下的事件对象具有不可预测性,欲实现精准的决策支持,要求计算模型具有较强的临场应变性。

计算设备日趋增强的处理能力、大数据时代计

算机各学科的交叉与融合等为应对上述挑战提供了可行的条件。对城市非结构化信息的时空耦合建立统一的表达方式(比如场景的三维时空模型),能够为城市计算提供数据基础,进一步分析城市环境下场景与时空对象的语义关联,建立信息空间与物理空间之间的联系,通过可视分析、增强现实等技术与现实场景无缝融合,能够基于城市场景呈现数据分析结果,辅助实现城市计算中的推理决策,从而实现人机智能的有机结合。

城市场景的三维时空表达与构建

城市非结构化信息的时空耦合依赖于城市环境下的三维空间信息,即城市对象与场景的三维时空模型。随着影像传感技术的发展,城市三维空间进入空、天、地综合一体化模式。特别是近年来激光测量技术快速发展,利用车载或机载激光雷达,结合传统视觉技术,能够快速获取城市空间信息,并进行大规模城市场景三维建模^[3]。美国政府资助了多个科研项目专注于城市级大场景三维建模,如加州大学伯克利分校的“大规模城市场景快速三维建模”和麻省理工学院的“城市扫描”项目。工业界的如谷歌和苹果等公司则推出了基于互联网的三维地图及应用平台,国内企业如腾讯、百度也纷纷开始研发类似的平台。随着移动设备的普及和传感器的小型化,业界已经开始考虑如何结合大数据与众包来构建智能城市,如采用来自互联网的300万张照片数据对罗马进行三维重建^[4];在德国海德堡大学的“开放式建筑模型”项目中,用户通过一个在线平台,在地图上标注位置并上传该位置上建筑的三维模型,进而渐进式地构建整个城市的建筑模型^[5]。

应用的不断深入对城市三维建模也提出了更高的要求,即场景规模越来越大、颗粒度越来越细、更新频率越来越高。因此,利用城市中的既有语义信息进行快速建模、按需实现城市对象及场景的多尺度建模、利用大量的多源异构数据进行渐进式建模等,成为城市计算构建三维时空模型的新思路。城市环境中的几何、影像、图表、文字等多源数据具有互补性,能够消除诸如激光扫描等单一数据源

的不完整性,从而获取场景更为全面的信息,生成高精度的三维几何模型。三维场景中的几何或者纹理特征(比如连续性、对称性、重复性、风格性、一致性等)往往蕴含着结构、功能和领域等语义信息。语义样本、先验规则等知识对于城市场景的渐进式建模起着非常重要的作用,例如,基于局部对象之间的对称性及规则性迭代优化建筑物层级结构,来实现建筑物的多样化建模^[6]。除此以外,利用机器人主动渐进式扫描场景,有利于消除物体自遮挡以及互遮挡等原因引起的三维扫描瓶颈,从而精细化构建场景三维模型^[7];传统车载扫描与以机器人为代表的精细化扫描方式能够实现闭环式采集处理,为城市场景粗、细粒度的自动化扫描提供了基础。

城市多模态数据的语义计算与融合

尽管城市三维场景为城市计算模型提供了融合非结构信息的基础,但信息空间的多模态数据间仍然存在语义融合问题。为刻画城市对象的完整语义和对象之间的互动关系,需要对城市对象的多模态信息进行有效的关联与语义融合。

城市数据的语义计算与融合涉及城市对象的三维空间定位、属性提取、属性关联分析等研究内容。已有的关于城市对象三维空间定位方法的研究主要基于单一模态或者低维度数据来对城市的公共设施、地标建筑、绿茵覆盖区域以及商业区等进行定位,比如利用全球卫星导航系统及图像处理技术的城市交通导航定位^[8]。对于城市对象属性信息的提取,目前也是主要集中在单一模态数据上,待提取的属性也相对简单,比如从航空或遥感图像中提取建筑物的形状、纹理或频谱特征^[9]。近年来兴起的细粒度物体识别^[10]为城市对象的精细标注提供了研究基础。在城市对象属性间的关联方面,跨数据属性的迁移学习(transfer learning)近年来逐渐引起人们的兴趣,最典型的例子是文本分析里的词、词袋、主题模型等概念在图像和视频分析里都有对应,如视觉词、视觉词袋、视觉主题模型等,这些大大推进了图像和视频中物体、场景的分类和识别的研

究^[11]。但是目前的迁移学习一般只考虑两个属性间的知识迁移,源空间和目标空间的数据属性较为单一,基于多种不同属性的协同学习理论与方法比较欠缺。

多模态数据的语义计算与融合的难点在于数据具有跨时空关联、海量异构、属性特征不易提取等特点。笔者认为,应充分利用人类视觉认知机理,将基于感知神经元的多尺度编码机理模型与多模态数据的分析理解相结合,并借用迁移学习的思想方法,把人类视听觉感知认知的原理推广到一般的多模态数据上。针对影像数据分散度高、关联性差、局部冗余的特点,可以利用多摄像头的影像数据的目标检测、跟踪方法,建立群体目标间的跨时空协同关联,探索跨时空影像对象之间互动的特点,特别是复杂环境中的群体效应、突发事件下的认知规律。针对城市大数据的海量冗余、异构多源、动态变化的特点,可利用多层次数据约简与非线性维数约减算法对其进行简化。除此以外,关键字提取、显著区域检测、几何测量、运动分析、物体识别等手段,也有利于提取城市对象的属性特征。

城市事件的临场分析与决策

复杂事件的临场感知分析与推理的主要目标是寻找事件的根源,发现三元空间中感兴趣事物的关联,利用数据挖掘技术预测事件的发展趋势,从而辅助分析与决策。

对三元空间中与复杂事件相关的视觉元素进行关联建模、理解和分析,是理解城市复杂事件的态势、分析复杂事件的缘由、制定应对措施等任务的核心和关键。美国南加州大学研究的“基于单视点的行为识别”^[12],就是通过协同分析多个摄像头信息来有效识别人体的行为。但是,现有方法往往局限于采用单个空间数据进行分析,形成难以突破的瓶颈。更重要的是,现有研究大多从局部数据出发分析少量对象的行为,如何基于三元空间的大量数据协同分析群体对象的行为模式也是亟待解决的问题。

数据挖掘是大数据时代理解和分析复杂事件的核心技术^[13],但其在部分场合和条件下存在严重的

局限性。数据挖掘技术的最佳应用场合是：待处理数据的属性和整合不随时间变化，数据完整干净，对分析目标有所认知。而智能城市复杂事件分析面临更多的难题，如数据来源迥异、不完整、不统一，且动态变化，规律无法提前预知。城市复杂事件分析关系到社会公共安全和生命财产安全，潜在风险大，对临场应变能力要求高。数据分析和挖掘技术只能提供决策的依据，机器智能无法完成的复杂事件的理解和分析必须借助人智能共同完成。

对于城市中的复杂对象与事件，可视分析能够帮助人类分析大规模数据，为决策者提供一个信息完备的可视化分析环境，有助于人类智能与机器智能的深度耦合。具有代表性的可视分析系统 Sand-box^[14] 提供了一个基于人机信息交互的、灵活的、表达力强的推理环境，支持视觉思维并提供分析过程模板，帮助用户实现可视分析推理。美国国家可视分析中心 (NVAC) 于 2005 年出版的《可视分析的研究和发展规划》报告^[15] 总结了可视分析方法和应用前景，指出可视分析的核心是采用可视化和用户交互方法，辅助用户从大尺度、复杂、矛盾甚至不完整的数据中快速挖掘有用的信息，以便做出有效决策。

因此，城市事件的临场分析与决策需要以视觉感知为基本通道，通过可视界面和交互，将人的智能特别是“只可意会，不可言传”的知识、个性化经验和直觉融入复杂数据的关联分析和推理决策过程中。简言之，临场复杂事件的分析与处理必须将人、数据和分析技术有机结合，在此过程中数据挖掘与交互可视分析方法各司其职，互助互补，迭代递进，从而高效推理决策。

专题文章介绍

本期专题邀请国家 973 计划项目“城市大数据的计算理论与方法”中的青年骨干科学家撰文，共同探讨大数据时代城市计算理论的进展。

城市环境下三维空间信息的构建是城市非结构化信息时空耦合的必要条件，城市三维场景数据为

城市计算整合其他多源信息提供了空间载体。山东大学教授陈宝权等人在《大规模城市场景建模与理解》一文中介绍了城市场景建模与理解领域的研究内容，包括国内外城市场景建筑和植物的三维重建、自动扫描与智能建模等方面的最新成果以及未来的发展方向。

城市计算依赖于数据处理的时效性和数据分析的智能性，异构时空大数据的管理、查询、挖掘等是城市计算中的技术基础。山东大学教授禹晓辉等人在《城市时空大数据的研究与应用现状》一文中介绍了城市时空大数据的管理、维护、查询以及挖掘等方面的研究进展，探讨了未来城市时空大数据的热点研究问题。

时空大数据具有数据量大、数据类型繁多、价值密度低以及更新速度快等特性，如何管理、查询以及挖掘有效信息是极具挑战的问题。时空数据度量空间为异构时空大数据处理提供了统一的表达方式，能够支持有效的时空大数据处理与分析。浙江大学教授高云君等人在《度量空间数据管理》一文中从度量空间的角度探讨了异构时空大数据的索引、查询技术以及未来的发展方向。

城市数据的可视分析是用户理解城市、分析城市、感知城市的强有力工具。北京大学研究员袁晓如等人在《城市移动数据知微探秘》一文中从深度与广度两个层面探索了城市数据中不同层次的分析对象以及不同类型的稀疏城市轨迹。

展望未来

大数据时代背景下的城市计算涉及计算机图形学、计算机视觉、数据可视分析、数据挖掘等领域知识，其研究成果为城市计算奠定了理论与实践基础。根据城市时空数据获取、多模态城市数据语义计算以及复杂事件临场分析等领域的现状及面临的挑战，笔者认为未来城市计算值得关注的动向主要有以下两方面。

- 社会空间中的活动主体——人类——将贯穿城市时空数据采集、分析与处理的全过程，众包策

略的机理分析与研究是未来值得关注的焦点。例如,利用众包机制对数据进行标注或语义分析,将有助于获取城市对象和场景的几何、材质和行为等属性信息,实现城市场景一致性的三维重构和动态更新。除此以外,利用深度学习从人类理解中探索城市场景理解机制,也将为城市非结构化数据时空耦合提供解决思路。

● 从城市计算中现有的对事实的了解分析逐步转换为对事件的溯源分析,使得现实数据的分析和模拟反演耦合更紧密(模拟反演包括物理的和事件的),有利于推断与决策。以城市计算中的污染物计算为例,城市三维时空模型作为数据基础,通过分析污染物扩散与迁移的数学模型,利用现有计算机技术对探测到的污染数据进行反演。城市计算不仅能够呈现当前的污染物浓度、影响范围等信息,还能够得到污染源的数目、位置、强度等关键信息,从而为政府管理人员、环境治理人员分析污染原因、指导应急隔离措施提供参考。■



陈宝权

CCF常务理事,长江学者特聘教授,山东大学计算机学院和软件学院院长。主要研究方向为基于车载移动激光扫描的大规模城市场景三维建模及海量数据可视化。baoquan@computer.org



曾 琼

CCF专业会员。山东大学博士。主要研究方向为计算机图形学、场景深度信息计算、语义相似性计算。qiong.zn@gmail.com

参考文献

- [1] Cook D J. How Smart Is Your Home? [J]. *Science*, 2012, 335(6076):1579-1581.
- [2] O'Grady M, O'Hare G. How Smart Is Your City? [J]. *Science*, 2012, 335(6076):1581-1582.
- [3] Li Y, Zheng Q, Sharf A, et al. 2D-3D fusion for layer decomposition of urban facades[C]// *Proceedings of*

International Conference on Computer Vision. IEEE Computer Society, 2011:882-889.

- [4] Frahm J M, Fitegeorgel P, Gallup D, et al. Building Rome on a Cloudless Day[C]// *Proceedings of European Conference on Computer Vision*. Springer-Verlag, 2010:368-381.
- [5] Uden M, Zipf A. Open Building Models: Towards a Platform for Crowdsourcing Virtual 3D Cities [J]. *Lecture Notes in Geoinformation & Cartography*, 2013:299-314.
- [6] Zhang H, Xu K, Jiang W, et al. Layered Analysis of Irregular Facades via Symmetry Maximization [J]. *ACM Transactions on Graphics*, 2013, 32(4):Article No.121.
- [7] Xu K, Huang H, Shi Y, et al. Autoscanning for coupled scene reconstruction and proactive object analysis [J]. *ACM Transactions on Graphics*, 2015, 34(6):1-14.
- [8] Marais J, Meurie C, Attia D, et al. Toward accurate localization in guided transport: Combining GNSS data and imaging information [J]. *Transportation Research Part C Emerging Technologies*, 2013, 43:188-197.
- [9] Mao F, Liu Z, Zhou W, et al. Extracting of urban features from high resolution remote sensing data based on multiscale segmentation[C]// *Proceedings of Urban Remote Sensing Event*. 2009:1-6.
- [10] Zhang N, Farrell R, Iandola F, et al. Deformable Part Descriptors for Fine-Grained Recognition and Attribute Prediction[C]// *Proceedings of International Conference on Computer Vision*. 2013:729-736.
- [11] Russell B C, Freeman W T, Efros A A, et al. Using Multiple Segmentations to Discover Objects and their Extent in Image Collections[C]// *Proceedings of IEEE Computer Society Conference on Computer Vision & Pattern Recognition*. IEEE Computer Society, 2006:1605-1614.
- [12] Nevatia R, Lv F. Automated single viewpoint human action recognition by matching linked sequences of key poses: US, US 8577154 B2 [P]. 2013.
- [13] Rajaraman A, Ullman J D. *Mining of massive datasets* [M]. Cambridge University Press, 2011.
- [14] Wright W, Schroh D, Proulx P, et al. The Sandbox for analysis: concepts and methods [J]. *Proceedings of ACM CHI*, 2006:801-810.
- [15] Munzner T, Johnson C, Moorhead R, et al. NIH-NSF Visualization Research Challenges report summary [J]. *IEEE Computer Graphics & Applications*, 2006, 26(26):20-24.