

【研究动态】

地理大数据在旅游领域的创新应用及学科影响和研究展望

蒋依依 高洁 郭佳明 徐海滨

【摘要】大数据正在改变捕捉和分析人类活动与行为的方式。多种新兴数据渠道不断涌现,为旅游领域研究提供了大量具有潜在应用价值的数据,克服了传统旅游研究中普遍存在的数据不足问题。本文以地理大数据的研究前沿为基础,明晰地理大数据在旅游领域人类旅游活动、旅游地理环境、旅游人地关系3个方面的理论支撑重点,梳理人类旅游活动/行为数据(UGC、设备数据、交易数据)和旅游地理环境数据(POI、客观环境数据)等地理大数据在旅游领域的创新应用,探讨地理大数据在理论研究更新、多源数据融合、分析方法突破3个方面的挑战和展望,为旅游领域时空认知、理论支撑、建模方法、分析框架等方面研究的深入开展提供一定参考。研究建议在理论上需要将科学研究范式标准化系统化,结合不同的事件和场景创造基于“过程-结构-机制”的中国旅游地理解释体系;数据上需要向大数据与其他数据的互补结合转变;分析技术需要向提高分析方法自适应性、涉入旅游现象特殊性变量的方向转变。

【关键词】地理大数据;旅游地理;用户生成数据;旅游行为;理论应用;地理信息科学;数据融合

【作者简介】蒋依依(1978-),女,贵州贵阳人,北京体育大学体育休闲与旅游学院,北京市冬奥文化与冰雪运动发展研究基地,教授,博士,博士生导师,研究方向体育旅游、奥运遗产、旅游地理,电子信箱:yiyijiang@bsu.edu.cn(北京100084);高洁(通讯作者)(1994-),女,河北邯郸人,北京体育大学体育休闲与旅游学院,博士生,研究方向为体育人文社会学、体育旅游,电子信箱:gj@bsu.edu.cn;郭佳明,徐海滨,北京体育大学体育休闲与旅游学院(北京100084)。

【原文出处】《地球信息科学学报》(京),2024.2.242~258

【基金项目】国家社会科学基金重点项目(21ATY001);中央高校基本科研业务费专项资金项目(2022YB016)。

1 引言

随着全球信息技术与工业化的发展与融合,物联网、云计算应运而生,传感器技术与移动定位技术飞速发展,人类对地表系统的刻画能力空前提升,位置感知技术等大数据正在改变我们捕捉和分析人类活动的方式^[1]。地理信息相关数据的多源性海量性逐步凸显,数据稀缺环境转变为数据丰富环境,数据驱动地理学的新模式已经出现^[2],地理大数据也为地理现象的分布格局、相互作用及动态演化提供了前所未有的社会感知手段^[3],使空间人文学与社会科学研究结合更加紧密^[4]。涉及地理大数据聚类分析方法、地理大数据异常探测、地理大数据预测建模等多方面的地理分析技术不断发展,在城市规划、智能交通、环境保护、公共安全

等领域研究应用成果不断涌现。尤其在旅游领域,海量地理信息与游客行为等数据的结合拓展了旅游地理的研究空间。

地理大数据与旅游统计数据、问卷数据等数据相比,涉及地理现象的时间、空间和属性维度,能够更加细致地刻画时空粒度、时空广度、时空密度、时空偏度和时空精度等“5度”特征^[5]。目前研究方法在大数据挖掘技术^[5]、时空分析方法^[6]、地理情感特征挖掘^[7]、地理大数据聚合^[8]、地理事件建模^[9]等方面已有较多探索,研究内容在旅游活动模式^[1]、旅游流动^[10]、社交媒体中的旅游偏好^[11]等方面已有涉猎,但地理大数据在旅游领域的分类和应用场景有待进一步梳理、使用挑战有待进一步明晰、研究内容也有待进一步拓展。本文收集了Web of Science

核心集数据库有关地理大数据在旅游应用的文献,在此基础上对地理大数据在旅游领域主题进行整理,筛选和剔除不相关文献,采用滚雪球查阅方式进行补充,最终获得 186 篇文献作为研究基础。从发表国家和地区看,中国发表论文数量最多,接近总量一半,体现了中国在旅游大数据领域研究的先进性,紧随其后的是美国、澳大利亚、西班牙等国。从研究领域分类来看,涉及社会科学、生态环境科学、计算机科学、商业经济学、自然地理等多个领域。从高被引文章看,Garcia - Palomares^[12]根据地理标签照片识别旅游热点景区的研究、Tenkanen^[13]评估社交媒体数据在游客监测方面的可用性研究、Miah 的大数据分析法^[14]、牟乃夏使用地理标记照片分析入境旅游流量时空分布和变化的框架^[15]、李渊通过时空分析对游客旅游行为的建模研究^[16]等均有较高的学术研究热度。通过文献分析,本文力图梳理地理大数据背景下旅游领域的理论支撑重点,明晰地理大数据在旅游领域应用的研究热点与应用场景,探讨地理大数据在旅游领域的现存挑战与研究趋势,以期对地理大数据在旅游领域应用中的理论支撑、创新应用、分析技术等方面提供借鉴。

2 地理大数据在旅游领域的理论支撑

地理大数据的应用为旅游研究提供了新数据和新技术,进而拓展了旅游相关理论的广度和深度。旅游作为多学科交叉融合的新兴学科,地理大数据融合能为包括旅游地理学在内的旅游学科底层理论构筑提供新的思路和范式。地理大数据数据和方法的多种应用凸显了该领域的新兴,但也反映了旅游现象的复杂性。相对于社会科学中成熟的理论体系,目前大数据的理论发展仍十分有限,其数据应用与理论支撑之间仍存在一定的学术鸿沟^[17]。但是地理大数据及其提供的社会感知手段,在“人-地”关系研究中发挥了重要作用,提供了一种“由人及地”的研究范式^[18]。本文以旅游地理学^[19-20]理论框架为基础,融合旅游消费者行为学^[21]等相关理论,从人类旅游活动、旅游地理环境以及旅游人地关系 3 个方面梳理了地理大数据背景下旅游领域在旅游行为、旅游流、旅游资源等相关理论中的研究进展(图 1)。

2.1 人类旅游活动相关理论研究

旅游者行为研究涉及经济学、心理学(普通心理学和社会心理学)、社会学、人类学、行为科学、地理学等多个学科^[21]以及旅游动机、旅游行为等多个理

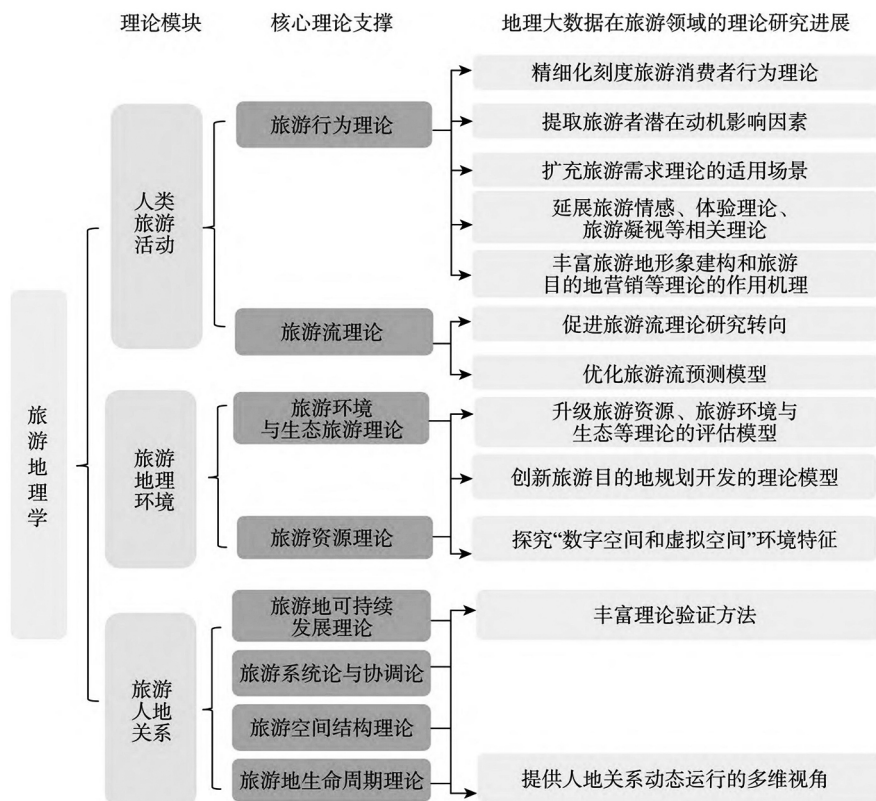


图 1 地理大数据在旅游领域的理论支撑和研究进展

论。从人类旅游活动层面来说,地理大数据的发展拓宽了旅游行为、旅游流等理论的研究广度,丰富了旅游需求、体验旅游等理论的应用场景。

地理大数据的引入为探究旅游动机、旅游需求、旅游体验等旅游行为理论提供了新的数据和方法,主要表现在精细化刻度旅游消费者行为理论、提取旅游者潜在动机影响因素、增加需求理论适用场景、延展旅游情感体验等方面。

(1) 精细化刻度旅游消费者行为理论。由于文本、图像和视频等地理大数据的大量引入,旅游空间格局认知要素分析、旅游者行为时空特征刻画和旅游目的地形象格局感知等实现了更精细化的刻度,进而使得旅游者的信息处理、经验主义、行为主义等研究范式的探索得以延展^[22]。例如,以往的研究大多使用政府统计和调查数据来检验旅行距离对游客行为的影响,现有学者开始利用新浪微博等数据分析中国苏州游客的旅行模式和消费模式,探究距离对酒店和餐厅支出等方面的影响差异^[23];使用带有地理标签的 Instagram 数据去获取游客在景点间的移动模式,描述不同景点相应的集群效应^[24]。

(2) 提取旅游者潜在动机影响因素。地理大数据的引入使旅游者动机中的旅游者人格分类、推拉因素、“资源-市场-产品”外部因素、“2W-2H”内部因素的挖掘更加细腻,能够更好协助研究者抽取隐性的潜在因素,具象化地分析旅游者动机。例如,已有研究大多基于空间可达性等地理距离视角去评析旅游者动机,而通过对地理大数据聚类及可视化分析可以发现游客的社会距离和地理距离均会对旅游者动机产生影响。

(3) 扩充旅游需求理论的适用场景。融入地理大数据后,游客需求理论在智慧化场景中的应用得到扩展,基于地理大数据的当地旅游活动推荐、旅游景点兴趣检测模型、智慧旅游个性化服务系统、智能决策的路径规划系统等都已经成为实践应用的重要方向,研究者也在逐步探究新媒体场景下的旅游者“需要-动机-行为”“刺激-反应”等相互关联、目的地规划的需求驱动结构^[26]等内容。

(4) 延展旅游情感、体验理论、旅游凝视等相关理论。问卷调查与深度访谈等方式,较多只能衡量单一层面和单一时间段的旅游者情绪情感,而基于地理大数据可以以颗粒化的形式评估不同旅游时间和不同旅游场景下游客的情感强度、持续时间、情感方向(正负向)等,从而能够更好地了解旅游者的态

度形成与决策改变的过程和影响因素,有助于学者提出不同的旅游态度改变行为假说、旅游者购买决策模型和旅游目的地选择模型。在体验经济背景下,感官体验维度的测量被放大和增强,学者们开始尝试使用文本挖掘和情感分析捕获和探索社交媒体评论中的五大感官体验^[27],也有学者将第六感内感受(interoception)纳入了新的探索维度^[28]。在游客满意度方面,相关研究从原有的问卷测量延展到从社交媒体文本中识别意向属性和旅游偏好,通过间接测量性能和重要性值来评估目的地的游客满意度,使得获取的样本量更大、成本更低,同时数据采集的灵活性更高^[29]。

(5) 丰富旅游地形象建构和旅游目的地营销等理论的作用机理。由于传播方式的变化,旅游目的地营销由注重传统媒体向新媒体的转变,尤其是与旅游者相关社交媒体口碑作用被放大。由于社交媒体分享习惯的常态化,研究者借助图像数据进一步整合了旅游目的地多个指标以丰富量化分析策略,搭建差异化的营销框架^[30],地理大数据也为旅游前后感知测度提供更充分的依据,弥合了目的地形象一致性研究中的理论和实证差距^[31],旅游者通过访问和利用谷歌街景地图和其他移动数字图像、制图工具以及社交媒体来创造城市变化的场所形象解读,为旅游地感知形象和认知形象路径的探索提供基础^[32],尤其是借助冬奥会等大型事件研究旅游体验的建构作用、投射形象的引致作用和品牌感知的晕轮作用对城市旅游地形象的影响^[33]。

信息和通讯技术的发展带来经济组织结构的变化,进而实现了空间形态由静态的位空间向流动的流空间转换^[34],地理大数据也为探究全球、国家、区域和地方不同尺度的旅游流规律机理等问题提供了研究基础,主要表现在促进旅游流理论的研究转向和优化旅游流的预测模型2个方面。①促进旅游流理论的研究转向。学术界有关于旅游流空间模式主要有圈层结构、核心边缘和空间扩散等理论^[35],信息技术促进了旅游流从中心地理论到中心流理论的研究转变,例如研究显示南京市餐馆密度的空间分布遵循传统的中心地理论,而南京市餐馆电子口碑的空间分布则更符合中心流理论^[36]。地理大数据可以更好地动态阐述空间扩散理论、细化旅游者个体行为和旅游流的社会群体流动行为以及有效解决旅游系统研究中旅游客源地、旅游目的地和联系媒介要素一体化整合的挑战^[33]。②优化旅游流的预测模

型。方法论方面,搜索引擎数据在旅游预测中应用最为广泛,时间序列计量预测模型仍占主导地位,人工智能方法仍在摸索和发展。模型建构方面,基于游客行为的空间规律性,地理大数据能够更持续、动态、精准地监测旅游者行为,识别影响游客忠诚度和满意度等相关因素,并借助优化仿真模型、神经网络模型、灰色预测方法、模糊理论和粗糙集理论等技术构建旅游空间需求预测模型。例如,过滤谷歌趋势数据以减少噪声干扰^[37]、利用通用嵌套时空(GNST)模型考虑内生和外生变量以及未观察到因素的空间和时间影响^[38]、开发图神经网络(GNN)框架注重区域实际联系^[39]、深度学习的时空融合图卷积网络(ST-FGCN)^[40]等方法,以改善旅游流量、旅游趋势、旅游消费等预测性能。

2.2 旅游地理环境相关理论研究

从旅游地理环境层面来说,地理大数据的发展有利于升级旅游资源、旅游环境与生态等理论的评估模型,创新旅游目的地规划开发的理论模型,同时也开始让学界思考从“实体空间”到“数字和虚拟空间”的环境特征。①升级旅游资源、旅游环境与生态等理论的评估模型。地理大数据的引入丰富了旅游服务设施空间分析、旅游资源评价、旅游地空间竞争、旅游环境容量测度等研究内容,也从多维度、多角度刻画了旅游资源适宜性、旅游地异质性和旅游场所活力等旅游地物理环境的特征和变化。例如通过整合手机定位数据和高分辨率遥感图像来评估城市绿地暴露水平,分析人口在不同时空尺度上接触城市绿地的游憩行为特征,探究促进城市可持续发展的相关因素^[41]。②创新旅游目的地规划开发的理论模型。William等^[42]丰富了智慧目的地(Innovation and Smart Destinations)等相关概念。地理大数据的研究综合利用了传统的地理数据分析方法和新兴的分布式云计算技术,在热门旅游资源识别^[43]、旅游目的地管理^[44]等多个方面构建出了多种能够处理大规模旅游数据的模型,为智慧旅游信息分析提供了有效的技术支持。有关目的地管理绩效和竞争力的评估正在向从价值链到价值网络、从物理空间到虚拟空间、从产品至上到品质聚焦、从静态固定到不断发展成形、从线性和时序到模型与实时动态的范式转变^[45]。③从探究“实体空间”到“数字空间和虚拟空间”的环境特征。不同虚拟空间环境可以游客带来差异化的硬件体验和心理体验,形成不同的感知价值,技术接受模型(TAM)、创新扩散理论、心流理

论、享乐理论等也在逐步应用于虚拟旅游研究中^[46]。

2.3 旅游人地关系相关理论研究

从旅游人地关系层面来说,地理大数据在旅游可持续发展等理论验证、旅游目的地动态运行的演化机制、目的地营销互动等方面发挥了一定作用。①丰富旅游可持续发展、旅游空间结构、旅游系统论与协调论、旅游地生命周期等理论的验证方法。“大数据+其他数据”的混合研究让人地互动的过程机制更加清晰。例如将传统游客统计数据和社会大数据结合,引入生态系统旅游压力测度,表明旅游压力可能是影响保护区生态系统可持续发展的决定性因素^[47];应用移动电话网络数据来评估日本海滩气候变化对人类活动社会影响和价值变化,以协助可持续发展决策^[48]。②提供人地关系动态运行的多维视角。地理大数据的应用有利于深化场景理论、空间生产理论、具身体验理论、演化地理学等相关理论,尤其是分析不同事件背景下(互联网传播、公共卫生事件、自然灾害等)旅游目的地的演化机制。例如,Yu等^[49]根据2019—2021年中国2.7715亿游客的手机数据,比较2次不同的COVID-19爆发浪潮导致全国游客流动性的变化,协助政府更好地理解公共卫生危机中的旅游管理以及大流行后恢复期间和未来爆发期间的决策。

3 地理大数据的种类及应用实践

3.1 地理大数据在旅游领域的分类

目前国际上比较认可的旅游大数据分类模式是以数据来源分为3类:①UGC数据(用户生成数据,User Generated Content),包括在线文字数据和在线照片数据;②设备数据(按设备),包括GPS数据、移动漫游数据、蓝牙数据等;③交易数据(按操作),包括网络搜索数据、网页访问数据、在线预订数据等^[50]。也有学者分为UGC数据、位置数据、事务数据3类^[51]或者UGC数据、操作数据(网络搜索)和物联网数据^[17]。此外,还可根据数据的存储形式分为本地数据和外部数据2类^[52];根据数据结构属性分为传统数据电子化、多源异构用户生产的UGC数据以及由旅游活动后台产生的时空行为数据^[53];根据数据应用类型划分社交生活数据、健康数据、商业数据、交通数据、科学研究数据5类^[54](表1)。

部分学者将地理大数据聚合定义为“多源地理大数据通过转换形成聚焦研究对象的多元数据集合的过程”,并根据关注对象的不同将地理大数据划分为对地观测大数据和人类行为大数据2类^[5]。结合

表 1 旅游大数据的类型划分			
分类标准	维度	举例	文献来源
数据来源	UGC 数据(用户生成)	在线文字数据和在线照片数据等	[50 – 52]
	设备数据(按设备)	GPS 数据、移动漫游数据/基站定位数据、蓝牙定位数据、Wifi 定位数据等	
	交易数据(按操作)	网络搜索数据、网页访问数据、在线预订数据等	
数据存储	本地数据	市政公共服务数据、目的地及景区自有数据、本地设备收集数据等	[52]
	外部数据	电信运营商数据、互联网在线数据、其他行业(非本地)数据等	
数据结构属性	传统数据电子化	企业内部管理系统,如员工考勤(打卡)记录;客户管理系统(CRM)等	[53]
	多源异构用户生产(UGC)数据	文字、图片、音频、视频等	
	由旅游活动后台产生的时空行为数据	GPS 轨迹数据、在线交易数据、网页访问数据等	
数据应用	社交生活数据	社交媒体服务、在线论坛、在线视频游戏和网络博客等	[54]
	健康数据	电子病历(EMR)、癌症登记数据、官方疾病暴发跟踪和流行病学数据、个人健康数据以及其他相关的生物医学信息等	
	商业数据	商业交易记录、在线商业评论、客户关系管理、超市会员记录、商店购物中心交易记录等	
	交通数据	GPS 轨迹、交通检查数据、社交媒体数据(签到数据、Waze 等平台数据)和手机数据(来自数据传输记录和蜂窝网络数据)等	
	科学研究数据	地震传感器、天气传感器、卫星图像、用于生物多样性研究的众包数据、自愿提供的地理信息和人口普查数据	

以上分类,将在旅游领域的地理大数据划分为人类旅游活动数据/行为数据、旅游地理环境数据两部分(表2)。人类旅游活动数据/行为数据大数据记录旅游者移动、社交、消费等各种行为的信息,涉及UGC数据、设备数据、交易数据等;旅游地理环境数据主要针对旅游客源地、旅游目的地、旅游景点/景区等不同的地表要素特征,涉及兴趣点(POI)数据、气象地貌客观环境数据等。与其他旅游大数据不同之处在于,地理大数据以地理信息为载体,结合其他社交媒体数据、网络交易数据等,综合了“旅游目的地”和“旅游者”的信息,更有利于可视化分析和模

拟游客行为特征、探究旅游系统中的人地关系等内容。

3.2 地理大数据在旅游领域的应用

3.2.1 人类旅游活动/行为数据

(1)UGC 数据

UGC 数据涉及数字视觉、文本和音频数据等,微博、抖音、Facebook、Instagram、Twitter、YouTube、Flickr 等多个社交媒体平台成为获取 UGC 数据的重要来源^[55]。UGC 数据驱动了地理空间语义、旅游目的地感知、情感体验^[52]等方面研究。其中地理文本数据和图像数据目前应用较为广泛。

表 2 地理大数据在旅游领域的应用举例			
渠道划分	数据类型	举例	
人类旅游活动/行为数据	UGC 数据	网络文本的在线签到数据、用户属地信息、自然语言文本的地理信息等;网络图像的地理标签信息等;音频数据的地理语义数据等	
	设备数据	与旅游者位移相关的 GPS 数据、移动漫游数据、蓝牙数据、RFID 数据和 WIFI 数据等	
	交易数据	交通消费数据、不同地区和时间在线预订数据和消费数据等	
旅游地理环境数据	POI 数据	工作、旅游、娱乐和寻路等目的地活动数据等	
	客观环境数据	气象数据、地质地貌数据、卫星遥感数据、无人机影像等	

地理文本数据驱动旅游地理空间语义研究。地理文本数据是指包含地理位置链接以及自然语言文本的数据集,这些链接既包括坐标明确地附加在文本中的地理标签,例如具有地理标记的推文或维基百科页面,也涉及新闻文章、旅游博客或历史档案等文本中明确提及的地点^[56]。地理文本数据通过提高地理信息特征的提取和匹配性能,促进了地理空间语义研究^[10],应用于旅游地情感分析^[57]、酒店景点评估^[58]等方面。例如,Wang^[58]对同一城市内的 TripAdvisor 酒店评论进行情感分析,发现顾客满意度存在空间依赖性;Wartmann^[59]通过收集游客的现场访谈、旅行博客以及 Filcker 标签,评估人们对不同景观特征的地方感差异。

地理图像数据助力多尺度评估游客的旅游目的地感知。多媒体和移动技术的快速发展让在线创建和共享大量旅行照片成为可能。与传统旅行调查或主动收集的 GPS 日志不同,地理图像数据的数据集样本量大,且可通过工具包轻松访问,能更快、更详细地提供信息和解决方案。一方面,借助地理图像数据,研究者可以更加立体形象地分析旅游者感知,包括旅游者的偏好运动轨迹^[60-61]、不同地点受欢迎程度^[62]等。另一方面,研究者可以根据地理图像数据模拟国际旅行行为和国家间旅行流动模式^[63],识别旅游景点发展潜力^[12],预测个性化旅游路线和偏好景点^[64],从而进行精准地目的地推荐、旅游规划和营销。旅游图像数据可以利用机器准确识别旅游场景、物体等环境信息,相较于地理文本数据可视化的程度更强,分析方法涵盖机器学习算法^[65]、混合集成学习方法^[66]等。随着流媒体发展进程加快,旅游视频图像的识别处理也正在成为技术研究的趋势之一。

(2) 设备数据

随着物联网的蓬勃发展,各种设备和传感器被开发和使用来追踪游客的行动,如 GPS 数据、移动漫游数据(手机信令数据)、蓝牙数据、RFID 数据和 WIFI 数据^[67]。以上结构化和非结构化类型的信息大数据不仅为旅游管理提供了海量的高质量的数据参考,作为官方统计数据的补充方法,已经广泛应用于旅游研究,显示出各自优势。设备数据能精准感知游客时空活动模式,为旅游管理者提供更高效的决策依据。目前研究较为常用的是 GPS 数据、手机信令数据。GPS 数据主要是通过轨迹记录仪和移动应用程序 2 种方式获取,尤其是公众参与地理信息

系统(PPGIS)以其游客行为刻画的高精度被较为广泛地应用在旅游目的地规划中。手机信令数据主要通过运营商无线电波的基站获取^[12],作为被动移动数据能够更客观的量化游客对时间和空间的使用。

公众参与地理信息系统提升旅游目的地规划有效性。公众参与地理信息系统(PPGIS)与土地利用规划的紧密耦合,成为旅游发展规划和监测的有效工具^[68]。目前 PPGIS 在旅游交通规划设计、识别和监测区域旅游发展偏好、公园景点空间决策、预测旅游相关发展、收集与旅游目的地的总体规划相关数据等方面发挥着重要作用。PPGIS 的研究案例地也趋向多元化,涵盖滑雪旅游目的地^[69]、休闲公园^[70]、国家公园^[71]、旅游景点^[72]等多个场景。PPGIS 方法也在根据应用场景不断创新,例如 Stewart 探讨了通过使用改进的 PPGIS 方法,即“社区行动地理信息系统”(CAGIS),分析居民对当前和未来当地旅游发展的反应^[73],让旅游规划者和管理者在项目实施前能充分考虑游客价值和偏好,为可持续的长期规划提供更深入的支持。

手机信令等数据推动旅游地理多元流研究的发展。物质、信息、能量等的移动或交换嵌入地理空间形成的地理多元流网络,为从地理和网络角度研究全球性问题提供了新的视角^[74]。手机信令数据微观层面上可以进行旅游行为机制解释分析,在流动性观测数据中分析比较游客对时间和空间的使用,量化游客对时间的使用、可视化游客的时空活动模式,系统性比较不同城市的旅游活动差异^[1],解析城市旅游地理特征与规律^[75],重建游客在逗留期间所走的路线,从而为优化旅游路线提供支撑;宏观层面可以对游客移动量进行解释,利用网络化挖掘搭建游客流动的迁徙网络。在数据提取方法上,Li 等^[76]提出了一种从手机数据中重建个人轨迹的方法(MDP-TR);在模型优化上,牟乃夏将轨迹数据模型分为轨迹点模型和轨迹段模型,提供了多种的聚类算法^[77]。移动漫游数据与 UGC 数据相比,属于被动移动数据,能够更客观地衡量游客时空行为。然而,目前此类数据的使用在隐私保护和道德方面还存在一定局限。

此外,GPS 数据和手机信令数据较广泛地应用于室外空间。室内定位数据的研究虽然较少,但是也在逐步进入研究视野。室内定位数据是一种同时具备地理空间位置与时间标签的轨迹大数据,蕴含丰富的个体行为信息^[78]。蓝牙、射频识别、无线局域网、红外线、超声波、紫蜂、超宽带、灯光和地磁感应

等数据均可用于室内定位^[79]。室内定位数据时空粒度细、定位精度高,在精度上可达到亚米级^[80],其所反映的主体与兴趣点间的空间关系更加明确。有关旅游领域室内定位研究,有学者目前已经尝试应用在博物馆空间中^[81],而酒店、景点、机场、旅游购物点等其他旅游场所的旅游者行为习惯和属性推断有待丰富,基于定位技术和穿戴设备的旅游者与室内环境交互作用等方面的研究也有待拓展。

(3) 交易数据

交易数据帮助旅游研究者洞察游客消费行为特征和旅游市场趋势。交易数据是直接从交易中获得的信息,记录了交易的时间、地点、所购买商品的价格、采用的支付方式、折扣等以及与交易相关的其他信息。在旅游领域,交易数据记录涉及旅游者进行旅游消费的相关操作(如网页搜索、网页访问、在线预订和购买等)和旅游市场供应商之间的交易、活动和事件等^[50]。目前研究侧重于交通交易数据和景区酒店等场景消费数据两方面。在与旅游密不可分的交通运输场景中,旅游交通交易数据呈现出迅速、海量发展趋势,高速公路 ETC、交通一卡通、移动支付等产生了海量的交易数据,有助于研究者探索旅游流量、流向、消费特征的研究。例如,Zhong 等^[82]通过新加坡的智能卡旅行记录中分析城市发展的空间结构;Domènec 等^[83]用由旅行卡生成的大数据监测城际公共交通的需求。通过对景区酒店等场景消费数据分析可以了解游客消费类别、游客消费偏好等,结合地理相关信息可以进行不同区域之间游客服务类型偏好、不同时间段旅游消费行为差异的对比研究,从而可以帮助优化旅游产品、改进旅游服务、实现旅游精准营销。将多种交易数据结合,例如将旅游交通数据与景区数据结合起来进行分析,可以探索景区外到景区内的旅游流量、客流布局,有助于深化景区的运营研究,改善管理和提升旅游体验。

3.2.2 旅游地理环境数据

(1) POI 数据

以兴趣点(POI)为主的数据为旅游研究提供了多元场景。POI 是一种数据集合的点数据,单体 POI 数据包含了实体名称、经纬度、地址等信息^[84]。作为一种主客共享的空间数据,POI 能够反映旅游要素的空间格局。研究尺度方面,POI 不仅包括在城市、区域大尺度的兴趣点,也产生了如室内监控等小尺度的兴趣点^[85];研究内容方面,POI 不仅应用在分析旅游业空间格局、热门景点布局 and 影响因素^[86]的研究

中,在旅游线路规划、个性旅游方案推荐、旅游客流引导等领域应用的有效性也得到验证^[87-88]。例如,Sarkar 等^[89]利用 PWP 算法根据游客的兴趣、行程受欢迎程度和行程费用等设计了一种 POI 多行程推荐引擎,实现推荐个性化;Gil 等^[86]通过分析来自搜索引擎的大数据来确定城市居民的兴趣点,分析 POI 与城市要素之间的关系,用以辅助城市规划和旅游政策的制定。

(2) 客观环境数据

客观环境数据主要通过遥感影像、街景图片、气候气象等数据对旅游地环境进行深度刻画,为旅游地理环境和旅游人地关系的研究提供了数据支撑。随着大数据时代的到来以及计算机处理技术的发展,遥感影像、街景图片、气候气象等多源数据大量介入,使得旅游地物理环境的刻画研究更加多维度和精细化,进而为人地交互视角下旅游者时空行为的机制探究提供了数据支撑。例如基于移动定位记录评估空气污染与出行行为之间的关系^[90]、开发沿海旅游气候指数评估沿海城市的旅游气候适宜性^[91]、测量气候变化对中国城市徒步旅行影响^[92]、模拟与预测景区扩容影响下的土地利用格局^[93]、利用街景图片分析游客感知和周边物理环境之间的关系^[94]等创新研究不断涌现。

3.2.3 地理大数据融合

地理大数据聚合的本质就是不同类型大数据的信息互补、增强与催生^[8]。地理大数据形式与结构多样的特征决定了数据融合存在多方面挑战,目前多源地理信息融合技术主要集中在矢量信息融合、地址信息一致化、同名实体识别与语义对齐、地理实体关系构建等方面^[95]。地理大数据的融合为旅游流、旅游数据集成平台的深入研究提供了一些新的思路。①利用多源地理数据丰富了旅游流的关联分析研究,基于游客移动流数据、航班流数据、旅游贸易数据等数据可以构建和识别旅游地理多元流网络,分析旅游者动机行为、决策选择行为、旅行行为和体验行为等方面的时序变化,结合国际关系流、国际贸易流、全球航班流等可以开展多元流网络的关联分析^[74],例如 Brahmabhatt^[96]使用基于网络度量的方法研究国际旅游与国际贸易流动之间的关系,Salas-Olmedo^[97]集合 Panoramio(观光)、Foursquare(消费)和 Twitter(住宿)3 种数据分析城市游客的数字足迹等;②基于多源地理异构数据平台的搭建让目

的地管理更加智慧化,有助于升级政府、旅游目的地、酒店的管理服务,例如 Fuchs 等^[44]设计了一款瑞典的目的地信息管理系统 DMIS - Are,用于获得有关游客在目的地旅游行程(旅游前、中、后)的相关信息,Zhu 等^[98]提出一种基于大数据融合的海洋旅游地理信息可视化系统设计方案。

4 地理大数据在旅游领域的研究展望

目前,地理大数据在旅游领域处于快速发展阶段。相对于传统数据,大数据规模更大、处理速度更快、实时性更好、数据价值和数据细节更丰富,尤其是地理多源大数据涉及地理现象的时间、空间和属性维度更丰富,时空粒度、时空广度、时空密度、时空偏度和时空精度的刻画更细致^[5],但是在数据质量的稳定性、数据的隐私性和安全性、技术操作的易用性、管理存储的便捷性等方面仍在存在一定的局限性。理论层面上,主要基于旅游地生命周期理论、旅游动机理论、旅游需求理论、社会网络等多学科理论开展研究;数据层面上,用户感知数据、位置数据等地理大数据的应用处于发展阶段;技术层面上,定量

分析及时空信息分析技术已经广泛应用,深度学习、自然语言处理等自适应性属性信息分析技术有待广泛引入。具体包括以下研究挑战及展望。

4.1 理论研究有待系统更新

旅游学、地理学、社会学、心理学等学科理论的引入和发展不仅丰富了地理大数据在旅游学研究的角度和思路,还为旅游现象以及旅游活动的现象解释提供了支撑。尽管与社会科学中公认的理论相比,地理大数据的理论发展非常有限^[17],在解决数据与理论之间的关系问题上仍有很大差距,但是在人类旅游活动、旅游地理环境、旅游人地关系等方面已经开始有了一定探索,初步形成了以旅游行为和旅游流、旅游资源和旅游环境与生态、旅游可持续发展和旅游目的地演化机制和营销等为代表的基于地理大数据的旅游研究理论支撑框架,但在研究范式的规范性和研究理论的创新上仍有不足(图2)。

研究范式上,目前大数据不仅仅是一种技术手段,也成为了一种自下而上的研究范式^[2]。在大数据视角下,世界由相关关系而非因果关系主导,这导致

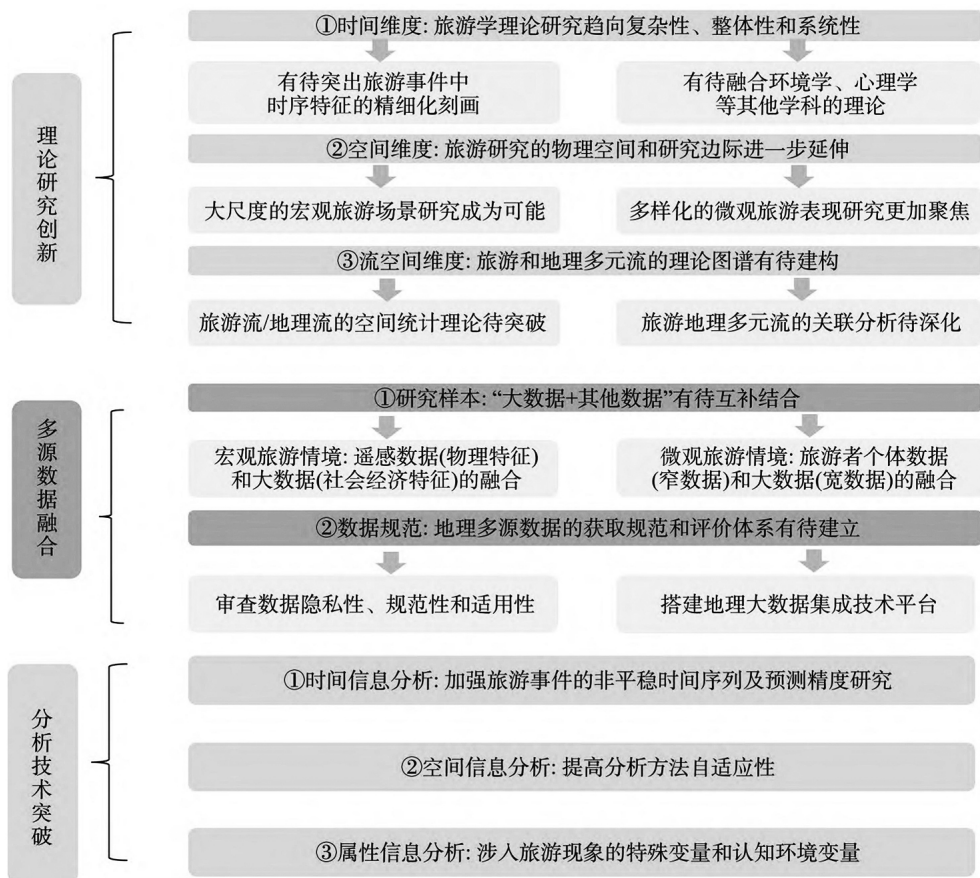


图2 地理大数据在旅游领域的研究展望

未来可以在旅游大数据研究中而无须依赖底层的结构化原理或机制,这种现象超出了旅游和地理研究中强调的传统科学范式^[99]。如何从精准描述现象到实现科学理论的惊险一跳仍然有更多空间值得探索,地理大数据的理论支撑和理论应用研究仍然有很大的发展空间。未来旅游大数据的理论研究重点可以放在科学的概念范式上,明确与地理大数据相关的智慧旅游目的地^[42]、旅游地理多元流空间^[74]等相关概念,建立问题和解决方案之间的联系,服务地理大数据研究科学范式的标准化系统化。

理论创新上,部分理论的使用仍然存在误区,既要明确理论的应用场景和适用范围,也要明晰研究问题和理论假设^[2],需要结合人类旅游活动、旅游地理环境、旅游人地关系等多方理论内容,从时间、空间、流空间等视角切入,创造基于“过程-结构-机制”的中国旅游地理解释体系^[100]。①在时间维度上,一方面非线性、不确定及多元的旅游事件时序特征精细化刻画以及旅游事件之间的反演、预测与延伸成为有待解决的议题之一,包括公共卫生事件^[49]、大型体育赛事^[101]、自然灾害等;另一方面旅游场景数据信息的密度和精细度进一步增强,迫使旅游场景研究需要从独立事件向整体性和系统性转向,也有待引入心理学、环境学等其他学科的理论,丰富大数据背景下旅游场景的研究范式及方法论。②在空间维度上,一方面地理大数据凭借着低成本、宽覆盖、高时效的数据特点,使得旅游事件的群体性、大场景研究成为可能,另一方面多样化的微观空间研究更加聚焦,未来需要综合多种开放数据在更精细的空间尺度上进行分析,如开展功能空间之间(例如酒店、景点、机场、旅游购物点等室内场景与地理空间信息相结合)的游客流动的网络结构和流动规律、以及超越“此时此地”空间性的游客行为模拟预测研究,以及旅游者在与旅游场景(消费、管理、营销、服务)交互过程中^[102]产生的行为差异、心理波动、生理表现等内容。③在流空间的维度上,目前地理流^[103]和旅游流^[35]的概念、模式分类和应用探讨已经十分深入。未来一方面有关构建和识别旅游流或地理流的空间统计理论仍有待突破^[103],另一方面旅游地理多元流的关联分

析有待进一步深化,例如从国际关系、国际金融、节事管理、人口社会学等多领域理论出发,探索国际关系、国际贸易、旅游人口移动、旅游交通网络等因素之间的相互关系与作用机制^[74],建构旅游地理多元流的知识图谱。

4.2 多源数据有待深度融合

基于前人的研究,人类旅游活动/行为数据和旅游地理环境数据的分类涵盖了绝大多数的数据类型,但仍有一些其他渠道的大数据因场景不同而具备多种属性,较难准确归类,地理大数据数据分类体系有待进一步优化。同时随着多源大数据表现出质度和量度的优化态势、多源数据的时空融合趋势和“大数据+其他数据”的组合互补等逐渐成为旅游学研究中重要的数据来源,这既可以发挥地理大数据的高精度优势,也可以结合其他数据在非空间维度的优势,让旅游机制揭示更微观、更具体。但是,在研究样本方面,旅游研究对象的准确性对数据收集和解释偏差的要求增加。目前旅游研究领域较为常见网络文本数据(微博、Twitter等)也仅能代表人群中的一个特殊子集^[104-105],例如基本位置检测方法可以在旅游监测中实现高精度,但存在高估游客数量、对用户的性别不敏感等问题^[106]。在数据规范性方面,随着对数据隐私、数据质量的要求增强,数据在网上数据收集和分析的道德规范方面更需要详细的规范与制约,不能仅仅因为数据表面公开而忽视对研究伦理的评估过程。同时,地理大数据的数量、类型与格式的不断丰富,也给存储、管理、处理、分析以及数据质量验证等方面带来了挑战^[107]。

基于以上的研究挑战,地理大数据的融合既要注重多源数据样本的获取,也要建立数据获取规范和评价体系。①在研究样本方面,未来可以基于大数据的时间信息、空间信息和属性信息,对旅游行为时间演变、旅游景区游客感知等社会人文特征进行机制揭示和规律挖掘,但是对于旅游景区物理环境及旅游者个体属性的信息挖掘及特征感知则需要引入其他数据,例如基于遥感影像的光谱特征、纹理特征、形状特征等数据信息,进行旅游景区地物类型、植被覆盖指数等环境变量的提取,进而弥补大数据在物理环境刻画上的不足^[108];而针对大数据在旅游

者年龄、性别、喜好等个体特征维度的信息缺失,可以融合大数据与用户画像数据,打破地理大数据这一“窄”数据维度信息不足的研究局限性,实现了从人地二元独立到人地交互结构性转变,进而为探究旅游研究中的人地交互机制提供更加丰富的研究数据。②在数据规范方面,需要建立针对旅游目的地、旅游者、旅游流等不同研究对象的空间尺度转换、时间序列存储、属性信息更新等地理大数据的获取规范和评价体系,并加强数据隐私性、规范性和适用性审查。另一方面,建立服务于旅游研究的地理大数据的计算框架与地理大数据集成技术平台也应受到广泛关注^[109-110]。

4.3 分析技术有待创新突破

随着自然语言处理、深度学习、空间分析、地理数学模型等数据分析技术的不断引入,基于地理大数据的大规模、高精度的旅游现象定量化感知逐渐成为旅游研究中重要的研究方法。目前在地理大数据的分析中,Lyu等^[17]已经较好地总结了现阶段有关UGC数据分析、设备数据分析和交易数据分析的主流方法和技术流程,数据分析基于信息科学技术基础逐渐规范。而由于旅游场景的复杂性与系统性,旅游现象除了简单的时空一阶特征,还应表现为旅游者之间的社交表达以及旅游者与旅游环境之间时空交互等地理大数据所体现的二阶特征,现有的分析技术也有待向提高分析方法自适应性^[111]、引入旅游现象特殊性变量涉入的方向转变^[112]。

未来地理大数据分析技术的创新在旅游领域可以从时间信息分析、空间信息分析、属性信息分析3个方面进行突破。①在对地理大数据的时间信息分析方面,连续时间序列下的旅游事件分析、客流预测、长期趋势变动已被广泛应用,但是仍需加强旅游事件的非平稳时间序列及预测精度的相关研究,如面向多元时间序列的预训练增强研究^[113]、突发事件对旅游影响的时间序列下的不均匀性特征等^[49]。②在对地理大数据的空间信息分析方面,核密度分析、最短路径分析、聚类分析、社会网络分析等空间分析技术已被广泛应用,而现有分析技术自适应性不强,难以实现更精准的、自发的、主动的、针对性的、个性化的旅游产品推荐和旅游导引。为此,在后

续研究中,需要针对旅游目的地的研究环境和旅游者个体的数据集特征,引入基础地理信息数据,实现技术研究参数的自动调整 and 适应。例如:在基于聚类分析技术开展旅游者旅游目的地选择偏好的研究中,可以根据旅游目的地的道路、河流、资源的空间分布关系,建立人地交互下的空间聚类分析模型^[114];在基于核密度分析方法开展旅游资源空间分布模式研究中,可以根据不同的资源类型和路网约束,自适应确定不同的搜索半径参数^[115]。多源地理数据与空间计量方法的结合成为研究的新趋势,采用新颖的空间估计方法产生了一个新的维度来调查跨景点的目的地内需求,来自邻近景点的游客流量溢出对景点需求有显著影响,客流溢出效应的强度和方向因景点位置而异^[116],这可以为决策者优化游客流量提供经验证据。③在对地理大数据的属性信息分析方面,现有的旅游环境情感表达和地理认知分析技术往往直接调用计算机或地理信息领域成熟的技术与方法,忽略了旅游现象研究的特殊性和认知环境变量的涉入。基于地理大数据的旅游领域研究需要紧跟自然语言处理的前沿技术,如加强面向旅游情景的旅游文本模型训练集建设^[117],提高冰雪旅游、海上冲浪等旅游场景的图片识别精度^[118]等。尤其是对于非结构化的社交媒体和多源数据,在将数据转化为时间序列进行预测之前,仍需在文本挖掘、情感分析和社交网络分析等方面取得方法论上的进步^[119]。

5 结论与讨论

首先,在理论方面,通过介绍人类旅游活动/行为数据和旅游地理环境数据的相关文献和案例,参考旅游地理学、旅游行为学等研究梳理了地理大数据在旅游领域(人类旅游活动、旅游地理环境、旅游人地关系)的理论支撑重点。其次,在应用层面概括阐述了UGC数据、设备数据、交易数据、POI数据、客观环境数据等地理大数据在旅游领域的创新应用。最后,从理论研究更新、多源数据融合、分析技术突破3个方面阐述了研究展望。在理论上需将自下而上的科学研究范式标准化系统化,结合不同的事件和场景探索丰富旅游学的理论边界,数据上有待向大数据与其他数据的互补结合转变,分析技术有待向提高分析方法自适应性、涉入旅游现象特殊性变

量的方向转变。由于文献收集等方面的局限性,本文在文献内容、数据分类、理论梳理、技术分析等方面仍有一些值得改进的地方,待后续研究进行补充和完善。

参考文献:

- [1] Xu Y, Li J, Xue J, et al. Tourism geography through the lens of time use: A computational framework using fine-grained mobile phone data[J]. *Annals of the American Association of Geographers*, 2021, 111(5): 1420-1444.
- [2] Zhang J. Big data and tourism geographies: An emerging paradigm for future study? [J]. *Tourism Geographies*, 2018, 20(5): 899-904.
- [3] 刘瑜, 詹朝晖, 朱递, 等. 集成多源地理大数据感知城市空间分异格局[J]. *武汉大学学报·信息科学版*, 2018, 43(3): 327-335.
- [4] 秦昆, 林琤, 胡迪, 等. 空间综合人文学与社会科学研究综述[J]. *地球信息科学学报*, 2020, 22(5): 912-928.
- [5] 裴韬, 刘亚溪, 郭思慧, 等. 地理大数据挖掘的本质[J]. *地理学报*, 2019, 74(3): 586-598.
- [6] 邓敏, 蔡建南, 杨文涛, 等. 多模态地理大数据时空分析方法[J]. *地球信息科学学报*, 2020, 22(1): 41-56.
- [7] 刘逸, 保继刚, 陈凯琪. 中国赴澳大利亚游客的情感特征研究——基于大数据的文本分析[J]. *旅游学刊*, 2017, 32(5): 46-58.
- [8] 裴韬, 黄强, 王席, 等. 地理大数据聚合的内涵、分类与框架[J]. *遥感学报*, 2021, 25(11): 2153-2162.
- [9] 杜云艳, 易嘉伟, 薛存金, 等. 多源地理大数据支撑下的地理事件建模与分析[J]. *地理学报*, 2021, 76(11): 2853-2866.
- [10] Raun J, Ahas R, Tiru M. Measuring tourism destinations using mobile tracking data[J]. *Tourism Management*, 2016, 57: 202-212.
- [11] Su S, Wan C, Hu Y, et al. Characterizing geographical preferences of international tourists and the local influential factors in China using geo-tagged photos on social media[J]. *Applied Geography*, 2016, 73: 26-37.
- [12] García-Palomares J C, Gutiérrez J, Mínguez C. Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS[J]. *Applied Geography*, 2015, 63: 408-417.
- [13] Tenkanen H, Di Minin E, Heikinheimo V, et al. Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas[J]. *Scientific reports*, 2017, 7(1): 1-11.
- [14] Miah S J, Vu H Q, Gammack J, et al. A big data analytics method for tourist behaviour analysis[J]. *Information & Management*, 2017, 54(6): 771-785.
- [15] Mou N, Yuan R, Yang T, et al. Exploring spatio-temporal changes of city inbound tourism flow: The case of Shanghai, China[J]. *Tourism management*, 2020, 76: 103955.
- [16] Li Y, Yang L, Shen H, et al. Modeling intra-destination travel behavior of tourists through spatio-temporal analysis[J]. *Journal of Destination Marketing & Management*, 2019, 11: 260-269.
- [17] Lyu J, Khan A, Bibi S, et al. Big data in action: An overview of big data studies in tourism and hospitality literature[J]. *Journal of Hospitality and Tourism Management*, 2022, 51: 346-360.
- [18] 杨学习, 邓敏, 刘瑜. 社会感知与地理空间智能的研究动态与展望——“社会感知与地理空间智能”专栏导读[J]. *地理与地理信息科学*, 2022, 38(1): 1-4.
- [19] 黄震方, 黄睿. 基于人地关系的旅游地理学理论透视与学术创新[J]. *地理研究*, 2015, 34(1): 15-26.
- [20] 保继刚, 楚义芳. 旅游地理学. 第3版[M]. 高等教育出版社, 2012: 3-7.
- [21] 孙九霞, 陈钢华. 旅游消费者行为学[M]. 大连: 东北财经大学出版社, 2015: 10-20.
- [22] Zhang K, Chen Y, Li C. Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing[J]. *Tourism Management*, 2019, 75: 595-608.
- [23] Xue L, Zhang Y. The effect of distance on tourist behavior: A study based on social media data[J]. *Annals of Tourism Research*, 2020, 82: 102916.
- [24] Zhou X, Chen Z. Destination attraction clustering: Segmenting tourist movement patterns with geotagged information[J]. *Tourism Geographies*, 2021: 1-21.
- [25] Kirilenko A P, Stepchenkova S O, Hernandez J M. Comparative clustering of destination attractions for different origin markets with network and spatial analyses of online reviews[J]. *Tourism Management*, 2019, 72: 400-410.
- [26] Park S, Zu J, Xu Y, et al. Analyzing travel mobility patterns in city destinations: Implications for destination design[J]. *Tourism Management*, 2023, 96: 104718.
- [27] Mehraliyev F, Kirilenko A P, Choi Y. From measurement scale to sentiment scale: Examining the effect of sensory experiences on online review rating behavior[J]. *Tourism Management*,

2020,79:104096.

[28] Li H, Li M, Zou H, et al. Urban sensory map: How do tourists "sense" a destination spatially? [J]. Tourism Management, 2023, 97: 104723.

[29] Chen J, Becken S, Stantic B. Assessing destination satisfaction by social media: An innovative approach using Importance - Performance Analysis [J]. Annals of Tourism Research, 2022, 93: 103371.

[30] Xiao X, Fang C, Lin H, et al. A framework for quantitative analysis and differentiated marketing of tourism destination image based on visual content of photos [J]. Tourism Management, 2022, 93: 104585.

[31] Li Y, He Z, Li Y, et al. Keep it real: Assessing destination image congruence and its impact on tourist experience evaluations [J]. Tourism Management, 2023, 97: 104736.

[32] Speake J, Kennedy V, Love R. Visual and aesthetic markers of gentrification: Agency of mapping and tourist destinations [J]. Tourism Geographies, 2021, 1: 1 - 22.

[33] 徐琳琳, 周彬, 虞虎, 等. 2022 年冬奥会对张家口城市旅游地形象的影响研究——基于 UGC 文本分析 [J]. 地理研究, 2023, 42(2): 422 - 439.

[34] 孙中伟, 路紫. 流空间基本性质的地理学透视 [J]. 地理与地理信息科学, 2005(1): 109 - 112.

[35] 钟士恩, 张捷, 韩国圣, 等. 旅游流空间模式基本理论: 问题分析及其展望 [J]. 人文地理, 2010, 25(2): 31 - 36.

[36] Xu F, Zhen F, Qin X, et al. From central place to central flow theory: An exploration of urban catering [J]. Tourism Geographies, 2019, 21(1): 121 - 142.

[37] Bokelmann B, Lessmann S. Spurious patterns in Google Trends data: An analysis of the effects on tourism demand forecasting in Germany [J]. Tourism management, 2019, 75: 1 - 12.

[38] Jiao X, Chen J L, Li G. Forecasting tourism demand: Developing a general nesting spatiotemporal model [J]. Annals of Tourism Research, 2021, 90: 103277.

[39] Sáenz F T, Arcas - Tunez F, Muñoz A. Nation - wide touristic flow prediction with Graph Neural Networks and heterogeneous open data [J]. Information Fusion, 2023, 91: 582 - 597.

[40] Li C, Zheng W, Ge P. Tourism demand forecasting with spatiotemporal features [J]. Annals of Tourism Research, 2022, 94: 103384.

[41] Song Y, Huang B, Cai J, et al. Dynamic assessments of population exposure to urban greenspace using multi - source big data [J]. Science of the Total Environment, 2018, 634: 1315 - 1325.

[42] Williams A M, Rodriguez I, Makkonen T. Innovation and

smart destinations: Critical insights [J]. Annals of Tourism Research, 2020, 83: 102930.

[43] Peng X, Huang Z. A novel popular tourist attraction discovering approach based on geo - tagged social media big data [J]. ISPRS International Journal of Geo - information, 2017, 6(7): 216.

[44] Fuchs M, Höpken W, Lexhagen M. Big data analytics for knowledge generation in tourism destinations: A case from Sweden [J]. Journal of destination marketing & management, 2014, 3(4): 198 - 209.

[45] Stienmetz J L, Fesenmaier D R. Traveling the network: A proposal for destination performance metrics [J]. International Journal of Tourism Sciences, 2013, 13(2): 57 - 75.

[46] 赖勤, 钱莉莉, 应天煜, 等. 虚拟旅游研究综述——基于 Scopus 数据库的文献计量与内容分析 [J]. 旅游科学, 2022, 36(1): 16 - 35.

[47] Chun J, Kim C K, Kim G S, et al. Social big data informs spatially explicit management options for national parks with high tourism pressures [J]. Tourism Management, 2020, 81: 104136.

[48] Kubo T, Uryu S, Yamano H, et al. Mobile phone network data reveal nationwide economic value of coastal tourism under climate change [J]. Tourism Management, 2020, 77: 104010.

[49] Yu L, Zhao P, Tang J, et al. Changes in tourist mobility after COVID - 19 outbreaks [J]. Annals of Tourism Research, 2023, 98: 103522.

[50] Li J, Xu L, Tang L, et al. Big data in tourism research: A literature review [J]. Tourism Management, 2018, 68: 301 - 323.

[51] 陆保一, 韦俊峰, 明庆忠, 等. 基于知识图谱的中国旅游大数据应用研究进展 [J]. 经济地理, 2022, 42(1): 230 - 240.

[52] 邓宁, 牛宇. 旅游大数据: 理论与应用 [M]. 北京: 旅游教育出版社, 2019: 14 - 15.

[53] 李君轶, 刘逸, 肖文杰, 等. “旅游大数据研究与应用前沿”系列笔谈 [J]. 旅游论坛, 2022, 15(1): 1 - 24.

[54] Tsou M H. Research challenges and opportunities in mapping social media and Big Data [J]. Cartography and Geographic Information Science, 2015, 42(sup1): 70 - 74.

[55] Bauder M. Using Social Media as a Big Data Source for Research: The Example of Ambient Geospatial Information (AGI) in Tourism Geography [M] // Geographies of Digital Culture. Routledge, 2017: 39 - 51.

[56] Hu Y. Geo - text data and data - driven geospatial semantics [J]. Geography Compass, 2018, 12(11): e12404.

[57] Cataldi M, Ballatore A, Tiddi I, et al. Good location, terrible food: Detecting feature sentiment in user - generated reviews

[J]. Social Network Analysis and Mining, 2013, 3 (4): 1149 – 1163.

[58] Wang M, Zhou X. Geography matters in online hotel reviews[J]. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2016, 41: 573 – 576.

[59] Wartmann F M, Acheson E, Purves R S. Describing and comparing landscapes using tags, texts, and free lists: An interdisciplinary approach[J]. International Journal of Geographical Information Science, 2018, 32(8): 1572 – 1592.

[60] Zhang Z, Chen R J, Han L D. Strategic sustainable development through the lenses of tourists' preferences: A geotagged photo approach[J]. Journal of Sustainable Development, 2019, 12 (3): 12 – 21.

[61] Vu H Q, Li G, Law R, et al. Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged. photos[J]. Tourism Management, 2015, 46: 222 – 232.

[62] Da Mota V T, Pickering C, Chauvenet A. Popularity of Australian beaches: Insights from social media images for coastal management[J]. Ocean & Coastal Management, 2022, 217: 106018.

[63] Yuan Y, Medel M. Characterizing international travel behavior from geotagged photos: A case study of flickr[J]. PloS one, 2016, 11(5): e0154885.

[64] Sun X, Huang Z, Peng X, et al. Building a model – based personalised recommendation approach for tourist attractions from geotagged social media data[J]. International Journal of Digital Earth, 2019, 12(6): 661 – 678.

[65] Barchiesi D, Preis T, Bishop S, et al. Modelling human mobility patterns using photographic data shared online[J]. Royal Society Open Science, 2015, 2(8): 150046.

[66] Wan L, Hong Y, Huang Z, et al. A hybrid ensemble learning method for tourist route recommendations based on geo – tagged social networks[J]. International Journal of Geographical Information Science, 2018, 32(11): 2225 – 2246.

[67] Shoval N, Ahas R. The use of tracking technologies in tourism research: The first decade[J]. Tourism Geographies, 2016, 18(5): 587 – 606.

[68] Brown G, Weber D. Using public participation GIS (PPGIS) on the Geoweb to monitor tourism development preferences[J]. Journal of Sustainable Tourism, 2013, 21(2): 192 – 211.

[69] Waleghwa B, Heldt T. Exploring the use of public participation GIS in transportation planning for tourism at a Nordic destination[J]. Scandinavian Journal of Hospitality and Tourism, 2022: 1 – 25.

[70] Brown G, Schebella M F, Weber D. Using participatory GIS to measure physical activity and urban park benefits[J].

Landscape and Urban Planning, 2014, 121: 34 – 44.

[71] Wolf I D, Wohlfart T, Brown G, et al. The use of public participation GIS (PPGIS) for park visitor management: A case study of mountain biking[J]. Tourism Management, 2015, 51: 112 – 130.

[72] Kantola S, Uusitalo M, Nivala V, et al. Tourism resort users' participation in planning: Testing the public participation geographic information system method in Levi, Finnish Lapland[J]. Tourism Management Perspectives, 2018, 27: 22 – 32.

[73] Stewart E J, Jacobson D, Draper D. Public Participation Geographic Information Systems (PPGIS): Challenges of implementation in Churchill, Manitoba [J]. The Canadian Geographer/Le Géographe Canadien, 2008, 52(3): 351 – 366.

[74] 秦昆, 喻雪松, 周扬, 等. 全球尺度地理多元流的网络化挖掘及关联分析研究[J]. 地球信息科学学报, 2022, 24 (10): 1911 – 1924.

[75] Reif J, Schmücker D. Exploring new ways of visitor tracking using big data sources: Opportunities and limits of passive mobile data for tourism[J]. Journal of Destination Marketing & Management, 2020, 18: 100481.

[76] Li M, Gao S, Lu F, et al. Reconstruction of human movement trajectories from large – scale low – frequency mobile phone data [J]. Computers, Environment and Urban Systems, 2019, 77: 101346.

[77] 牟乃夏, 徐玉静, 张恒才, 等. 移动轨迹聚类方法研究综述[J]. 测绘通报, 2018(1): 1 – 7.

[78] 承达瑜, 秦坤, 裴韬, 等. 基于室内定位数据的群体时空行为可视化分析[J]. 地球信息科学学报, 2019, 21(1): 36 – 45.

[79] 舒华, 宋辞, 裴韬. 室内定位数据分析与应用研究进展[J]. 地理科学进展, 2016, 35(5): 580 – 588.

[80] 邓中亮, 张森杰, 焦继超, 等. 基于高精度室内位置感知的大数据研究与应用[J]. 计算机应用, 2016, 36(2): 295 – 300.

[81] Kanda T, Shiomi M, Perrin L, et al. Analysis of people trajectories with ubiquitous sensors in a science museum [C]// Proceedings 2007 IEEE International Conference on Robotics and Automation. IEEE, 2007: 4846 – 4853.

[82] Zhong C, Arisona S M, Huang X, et al. Detecting the dynamics of urban structure through spatial network analysis[J]. International Journal of Geographical Information Science, 2014, 28 (11): 2178 – 2199.

[83] Domènech A, Miravet D, Gutiérrez A. Mining bus travel card data for analysing mobilities in tourist regions[J]. Journal of Maps, 2020, 16(1): 40 – 49.

[84] 窦旺胜,王成新,薛明月,等. 基于 POI 数据的城市用地功能识别与评价研究——以济南市内五区为例[J]. 世界地理研究,2020,29(4):804–813.

[85] 郑姗姗. 移动通信中多系统合路平台(POI)及监控的研究[D]. 南京理工大学,2014.

[86] Gil E, Ahn Y, Kwon Y. Tourist attraction and Points of Interest (POIs) using search engine data: Case of Seoul[J]. Sustainability,2020,12(17):7060.

[87] Santos F, Almeida A, Martins C, et al. Using POI functionality and accessibility levels for delivering personalized tourism recommendations[J]. Computers, Environment and Urban Systems, 2019,77:101173.

[88] Bin C, Gu T, Sun Y, et al. A personalized POI route recommendation system based on heterogeneous tourism data and sequential pattern mining[J]. Multimedia Tools and Applications, 2019,78(24):35135–35156.

[89] Sarkar J L, Majumder A. A new point-of-interest approach based on multi-itinerary recommendation engine[J]. Expert Systems with Applications,2021,181:115026.

[90] Xu Y, Liu Y, Chang X, et al. How does air pollution affect travel behavior? A big data field study[J]. Transportation Research Part D: Transport and Environment,2021,99:103007.

[91] Gao C, Liu J, Zhang S, et al. The coastal tourism climate index(CTCI): Development, validation, and application for Chinese coastal cities[J]. Sustainability,2022,14(3):1425.

[92] Liu J, Yang L, Zhou H, et al. Impact of climate change on hiking: Quantitative evidence through big data mining[J]. Current Issues in Tourism,2021,24(21):3040–3056.

[93] 吴振华,俞钦平,王亚蓓,等. 景区扩容影响下的桂林市土地利用情景模拟与预测[J]. 水土保持通报,2022,42(5):131–139.

[94] Zhang F, Zhou B, Liu L, et al. Measuring human perceptions of a large-scale urban region using machine learning[J]. Landscape and Urban Planning,2018,180:148–160.

[95] 刘纪平,王勇,胡燕祝,等. 互联网泛在地理信息感知融合技术综述[J]. 测绘学报,2022,51(7):1618–1628.

[96] Brahmabhatt J, Menezes R. On the relation between tourism and trade: A network experiment[C]//2013 IEEE 2nd Network Science Workshop. IEEE,2013:74–81.

[97] Salas-Olmedo M H, Moya-Gómez B, García-Palomes J C, et al. Tourists' digital footprint in cities: Comparing Big Data sources[J]. Tourism Management,2018,66:13–25.

[98] Zhu W, Hou Y, Wang E, et al. Design of geographic information visualization system for marine tourism based on data mining[J]. Journal of Coastal Research,2020,103(SI):1034–

1037.

[99] 吴必虎,黄潇婷,刘培学,等. 中国旅游地理研究:成果应用转化和研究技术革新[J]. 中国生态旅游,2021,11(1):52–65.

[100] Lu L, Bao J, Huang J, et al. Recent research progress and prospects in tourism geography of China[J]. Journal of Geographical Sciences,2016,26(8):1197–1222.

[101] Choe Y, Baek J, Kim H. Heterogeneity in consumer preference toward mega-sport event travel packages: Implications for smart tourism marketing strategy[J]. Information Processing & Management,2023,60(3):103302.

[102] 李云鹏. 旅游场景驱动的大数据应用[J]. 旅游学刊,2017,32(9):4–6.

[103] 裴韬,舒华,郭思慧,等. 地理流的空间模式:概念与分类[J]. 地球信息科学学报,2020,22(1):30–40.

[104] Boyd D, Crawford K. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon[J]. Information, Communication & Society,2012,15(5):662–679.

[105] DeLyser D, Sui D. Crossing the qualitative-quantitative divide II: Inventive approaches to big data, mobile methods, and rhythm analysis[J]. Progress in Human Geography,2013,37(2):293–305.

[106] Liu Z, Zhang A, Yao Y, et al. Analysis of the performance and robustness of methods to detect base locations of individuals with geo-tagged social media data[J]. International Journal of Geographical Information Science,2021,35(3):609–627.

[107] Li S, Dragicevic S, Castro F A, et al. Geospatial big data handling theory and methods: A review and research challenges[J]. ISPRS Journal of Photogrammetry and Remote Sensing,2016,115:119–133.

[108] Ramaano A I. Geographical information systems in sustainable rural tourism and local community empowerment: A natural resources management appraisal for Musina Municipality Society[J]. Local Development & Society,2022:1–32.

[109] Wang S W, Liu Y, PADMANABHAN A. Open cyber-GIS software for geospatial research and education in the big data era[J]. SoftwareX,2016,5:1–5.

[110] Xiong W, Chen L. HiGIS: An open framework for high performance geographic information system[J]. Advances in Electrical and Computer Engineering,2015,15(3):123–132.

[111] Sanyal S, Xu Y, Wang S, et al. APOLLO: A simple approach for adaptive pretraining of language models for logical reasoning[J]. arXiv Preprint arXiv:2212.09282,2022.

[112] Wu D C, Cao C, Liu W, et al. Impact of domestic

tourism on economy under COVID - 19: The perspective of tourism satellite accounts[J]. Annals of Tourism Research Empirical Insights, 2022, 3(2): 100055.

[113] Shao Z, Zhang Z, Wang F, et al. Pre - training enhanced spatial - temporal graph neural network for multivariate time series forecasting[C]//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2022: 1567 - 1577.

[114] Zhan J, Mao J, Liu Y, et al. Learning discrete representations via constrained clustering for effective and efficient dense retrieval[C]//Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. 2022: 1328 - 1336.

[115] Song J, Xing H, Zhang H, et al. An Adaptive Network - Constrained Clustering (ANCC) model for fine - scale urban functional zones[J]. IEEE Access, 2021, 9: 53013 - 53029.

[116] Kim Y R, Liu A, Stienmetz J, et al. Visitor flow spillover effects on attraction demand: A spatial econometric model with multisource data[J]. Tourism Management, 2022, 88: 104432.

[117] Yang S, Duan X, Xiao Z, et al. Sentiment classification of Chinese tourism reviews based on ERNIE - Gram + GCN[J]. International Journal of Environmental Research and Public Health, 2022, 19(20): 13520.

[118] Qin X, Yin D, Dong X, et al. Passenger flow prediction of scenic spots in Jilin province based on convolutional neural network and improved quantile regression long short - term memory network[J]. ISPRS International Journal of Geo - information, 2022, 11(10): 509.

[119] Li X, Law R, Xie G, et al. Review of tourism forecasting research with internet data[J]. Tourism Management, 2021, 83: 104245.

Big Geodata in Tourism Research: Innovative Application, Disciplinary Influence, and Research Prospect

Jiang Yiyi Gao Jie Guo Jiaming Xu Haibin

Abstract: The way we capture and analyze human activity and behavior is changing because of big data. A variety of new data sources have emerged to supplement the official data, offering a significant amount of data with potential application value for the research of tourism and leisure while overcoming the common problem of insufficient data in traditional tourism research. Based on the research frontier of big geodata, this paper explains the theoretical foundation of tourism under the background of geographic multi - source big data at three levels: human tourism activities, tourism geographical environments and destinations, and the relationship between tourists and tourist destinations. This paper summarizes the application of big geodata, such as human tourism activity data (e. g. , UGC data, device data, transaction data) and tourism geographical environment data (e. g. , POI, environmental data). This paper discusses the challenges and prospects of big geodata in three aspects: research paradigm and theory, multi - source data fusion, and analysis methods. For the research paradigm and theory, there is the requirement for standardize and systematize the scientific research paradigm by combining different events and scenarios to create an interpretation system of Chinese tourism geography based on "process - structure - mechanism". In terms of multi - source data fusion, the combination of big data and other data is necessary. In terms of analysis methods, efforts are still needed to improve the adaptability of analysis methods and incorporate the specific variables of tourism phenomena.

Key words: big geodata; tourism geography; UGC; travel behavior; theoretical application; geographic information science; data fusion