

基于微博语义分析的重庆主城区风貌感知评价

易 峥^{1,2}, 李继珍¹, 冷炳荣¹, 陈 敏¹

(1. 重庆市规划设计研究院, 重庆 401147; 2. 四川美术学院公共艺术学院, 重庆 400053)

摘 要:采用社交网络数据对城市进行感知与评价是一种新的城市感知定量研究方法。本文将新浪微博签到数据作为城市感知的数据来源,通过文本挖掘和语义分析方法,探索用户对重庆主城区城市风貌的感知与评价。针对重庆主城区特色风貌片区,研究形成了签到地图、情绪地图和对象地图。签到地图客观反映重庆主城区新浪微博用户活跃程度的空间分布特征;情绪地图挖掘活跃用户对空间的情绪表达和正负面态度;对象地图解析出现这种签到热力分布和情绪分布的原因,发现引发正负面态度的情绪对象。通过分析用户对规划师所设定的风貌要素载体的典型意见,将要素载体归纳为有感知积极、有感知消极、无感知和有感知未设定4类。今后应进一步从要素识别和价值判断两个方面为规划师塑造和管控城市风貌提供参考意见。

关键词:城市风貌感知与评价;新浪微博;语义分析;重庆;签到地图;情绪地图;对象地图

1 引言

城市风貌是由城市的自然环境、建成环境等物质环境与历史文化等非物质环境组成,既包括体现城市历史文化和人文气质的“风”,也包括展现城市物质空间特色的“貌”(马武定, 2009; 段德罡等, 2010; 戴慎志等, 2013)。城市风貌以人们的主观意愿为基础,是人们所倡导的文化在城市物质空间上的反应(段德罡等, 2010),也是城市在人的感知层面上区别于其他城市的形态表征(王建国, 2007)。因此,可将城市风貌视为一种审美对象,需要审美主客体的互动方能产生的一种审美意象,并应该强调审美主体对客体的感受与体验。

城市风貌研究的历史经历了从功能主义的客体化研究向人文主义倾向的主体化研究,再向后现代主义倾向的主客体统一的发展脉络(杨昌新, 2014)。凯文·林奇(2001)开创了城市风貌主体化研究领域,后来的中西方城市设计和风貌规划开始不断地实践他的理论,增加主客体互动研究。城市风貌规划一直是城市特色塑造和城市形象展示的重

要手段,研究主要针对城市风貌评价、风貌特色规划、风貌管控等展开。由于城市风貌特色与人们的建设发展理念和主导价值观念直接相关,城市风貌的难以描述特性使其不易为规划所控制,在城市建设中易被随意诠释(王建国, 2007),因此感知和评判城市风貌在风貌规划中显得尤为重要。对于城市风貌的感知,主要采取传统的行为偏好调查研究和凯文·林奇的城市意象调查方法,即通过问卷调查和规划设计师的现场调查,对研究区域进行要素识别、价值评估、主题提炼、意象建构,推演出城市风貌特色构成、规划塑造方法及管控手段。这种研究方法依赖的是专业人士的感知能力和认知水平,而受制于主观判断、客体规模、意愿表达真实性等的影响,审美客体的参与深度和广度有限,因而对公众感知的测度存在偏差,需要探索新的方式来实现主客体的互动。

近年来出现了对城市感知的定量研究。心理学研究中的语义分析法是定量而简单获取人们对城市空间感知的一种方法。语义分析采用SD量表法(Semantic Differential),即通过设定形容词对进

收稿日期:2017-06;修订日期:2017-08。

基金项目:中华人民共和国住房和城乡建设部科技计划项目(2017-R2-022)[Foundation: Science and Technology Project of Ministry of Housing and Urban-Rural Development of the People's Republic of China, No.2017-R2-022]。

作者简介:易峥(1973-),女,重庆人,博士,教授级高级工程师,主要从事城市与区域规划等方面研究,E-mail: loisayi@126.com。

引用格式:易峥,李继珍,冷炳荣,等. 2017. 基于微博语义分析的重庆主城区风貌感知评价[J]. 地理科学进展, 36(9): 1058-1066. [Yi Z, Li J Z, Leng B R, et al. 2017. Perception and evaluation of cityscape characteristics using semantic analysis on microblog in the main urban area of Chongqing Municipality[J]. Progress in Geography, 36(9): 1058-1066.]. DOI: 10.18306/dlkxjz.2017.09.002

行问卷调查,获取对事物的差异评价,从而达到量化分析的目的(王德等, 2009; 赵渺希等, 2013; 张励杰, 2015)。SD量表法的优势在于研究对象指向明确,研究过程可控,语义归纳相对容易(纪娇娇等, 2015),但SD量表的形容词为研究者所设计,带有先入为主的特征,会影响到受访者的评价。另外,设定的形容词对可能存在概括不全的问题,难以准确全面表达客体的感受与体验(林江豪等, 2012)。互联网的发展为语义研究提供了新的可能,一方面是海量文本数据的可获得性空前增加,twitter、微博、微信、贴吧、论坛/BBS等社交网络的飞速发展使得广泛快速获取公众态度变为可能,自媒体数据成为公众感知研究的新数据源;另一方面文本挖掘和文本分析工具的发展促进了对海量文本数据的运用,通过对非结构化文本信息的挖掘、处理、提取特征、分类、聚类等,获取客体行为数据、情感数据与经验数据(秦萧等, 2013; 杨振山等, 2015; 甄峰等, 2015),为解决客体参与样本数量有限、调查精力有限、时空尺度有限等问题提供了可能。近年来基于自然语言处理的语义研究方法逐步成熟(林江豪等, 2012),研究者使用twitter、微博用户数据,构建情感分析模型和算法,分析用户情绪(Go et al, 2009; Kouloumpis et al, 2011; Saif et al, 2012; Fan et al, 2014),但是很少对产生情绪的原因进行研究。

本文探索运用互联网公共开放数据,基于自然语言的语义分析方法,开展大尺度的城市风貌感知与评价。相比于传统研究方法,基于互联网数据和语义分析法的感知研究可带来3个转变:一是受访者范围更大,不局限于问卷受访人群,使用网络资源的人群均可成为潜在的研究对象;二是研究视角发生了改变,不再基于调查者的主观视角,而是基于体验者亲身经历或感悟的视角;三是分析内容更为全面,不仅包括有关城市风貌的形容词,还可对情感色彩词组等关键词进行大规模交叉分析。

微博是基于用户关系的信息分享、传播和获取平台,是快捷自主的社交网络。研究使用的微博签到数据用于城市感知评价具有以下特征:一是空间性,签到地带有经纬度,通过空间位置帮助识别感知地点;二是时间性,有签到的具体时间;三是地点属性,签到地POI属性有地点分类,包括20个大类193个中类692小类;四是高数据量,2016微博用户发展报告显示:新浪微博月活跃用户2.97亿人,是微博元年2009年的10倍,其中月均登录频次在15

天以上的占51.4%。微博文本的主要形式是文字+配图,反映博主实时记录,自主抒发情感,选择转发等行为,旅游和娱乐是重要的兴趣标签,旅游目的地提及次数和提及人数大幅增长,人均提及次数为12次/年。用微博签到数据和文本来测度公众对城市的感知,可部分克服缺乏数据来源和数据分析方法的问题,开拓了认知城市的方法。但是,微博用户组成缺乏广泛性,用户以青年白领群体为主。2016年微博用户发展报告显示,微博用户中18~30岁的占70%,78%为大学及以上学历(微博数据中心, 2017);另外微博文本的短文字、碎片化和泛主题特征,使其被应用于某一特定主题研究时,文本数据处理难度大,准确提取并定义特征难度大,需要人工辅助决策。本文先结合重庆主城区特色风貌片区,人工提取筛选锁定一部分特征点,再抓取特征点上的签到信息,人工逐一梳理提取有用信息等办法进行相关的分析和处理。基于微博签到数据,挖掘用户的签到分布、情绪分布以及引起积极或消极情绪的对象,进一步梳理风貌要素载体的典型意见,进行风貌要素感知分类,为塑造和管控城市风貌提供参考意见。

2 研究区域与方法

2.1 研究范围

本文研究范围为重庆市主城区,重庆主城区以大山大水著称,拥有3000多年建城史,山水与城市交融形成了独特的城市风貌。重庆主城区面积5473 km²,结合山水资源(美丽山水)、历史文脉(巴渝名城)、山地建筑特色(立体都市)进行空间判读,重庆主城区风貌规划中具有重点风貌特色的区域约180 km²,包括一岛两江城市风貌核心、风貌展示轴、山水展示带、传统风貌区、城市中心、城市门户等(图1),重点对这些区域的微博数据进行挖掘和分析。

2.2 研究方法

2.2.1 数据获取与挖掘

本文以新浪微博签到数据作为数据源,共分析文本约80万条,截止时间为2016年10月。结合重庆主城区特色风貌片区,人工提取筛选出代表城市风貌的150个特征点,通过调用新浪微博API接口,首先抓取这些特征点所对应的新浪微博兴趣点(POI),再抓取每个POI点的签到数量及微博文本。

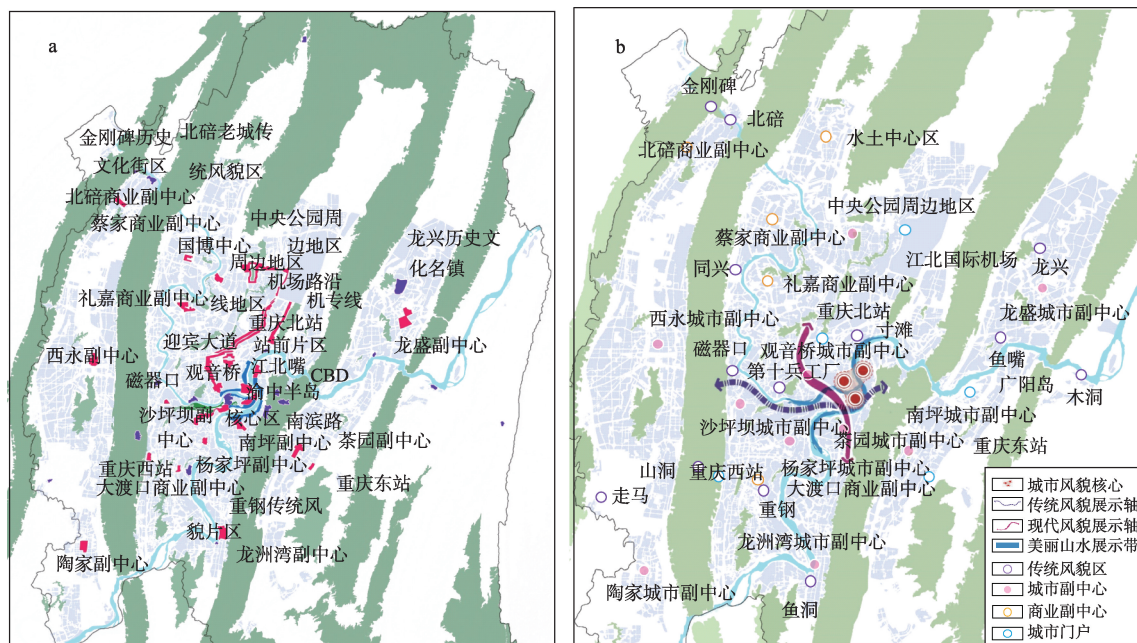


图1 重庆主城区特色风貌要素(a)和风貌区(b)

(资料来源:《重庆市主城区城市风貌总体设计》)

Fig.1 Cityscape characteristic factors (a) and spatial divisions (b) in the main urban area of Chongqing Municipality

2.2.2 数据处理

对微博文本进行预处理,包括分词和去停用词。分词是将连续的字序列按照一定的规范重新组合成词序列,国内已经开发出针对中文的分词软件,本文采用编程语言Python调用中文分词组件“结巴(Jieba)”进行中文分词和词性标注。停用词是指去除对文本意思表达贡献度很小的连接词、代词、介词、语气词等虚词和标点符号,同样采用Python编程去掉停用词,以及“@用户”、“分享图片”、“#标签#”和链接“http://...”等与文本意思表达不相关的信息。

3 结果与分析

3.1 签到地图:基于用户签到的点评活跃空间

签到地图反映了用户在重庆主城区点评活跃程度的空间分布,用于识别被关注的区域和关注的热点。重庆市主城区新浪微博POI点签到热点区域主要分布在:地标建筑集中的渝中半岛、观音桥等消费休闲娱乐功能突出的城市副中心区域,南山、大礼堂等旅游风景区,南滨路、磁器口等滨江地带,江北机场和重庆北站等交通枢纽地区,以及大学城地区。这些都是具有重庆特色风貌的区域,人们更愿意去这些地方游玩并进行签到活动和发表

自己的意见。结合这些区域,重点分析人工筛选出与城市风貌相关的具有代表性的150个特征点的签到情况,据此可得出这些POI点的签到次数热力图(图2)。

将150个POI点的签到次数进行排序(图3),发现签到次数最多的是“江北机场”,达34万多次,远远超过其他POI点上签到次数;签到次数排名前10位的POI点有5个在渝中区,其次是沙坪坝区有3个,渝北区仅有2个交通枢纽。通过网站(<http://weibo.cn/>)可获得每个POI点的提及次数,以此来辅助研究和分析签到次数,将提及次数进行排序,得到提及次数排名前10的POI点。风貌片区150个POI点提及次数排名前2名“沙坪坝”和“解放碑”的提及次数明显超过其他特征点,分别达到36万和31万次;提及次数排名前10的POI点中有4个是城市中心和副中心,有1所教育机构重庆大学,有洪崖洞等3处景点,1处商场,1处地名“石桥铺”。

3.2 情绪地图:基于表情和文本语义的情感信息挖掘

研究对用户在其签到所在地发表的微博信息进行情感分析,反映大众在虚拟环境对当前空间的正负面态度。微博正文用表情和文本表达情感。本文首先基于表情情感词典来研究风貌片区内签到用户的情绪地图,然后基于文本情感来进一步探索风貌片区签到用户的情绪空间分布。

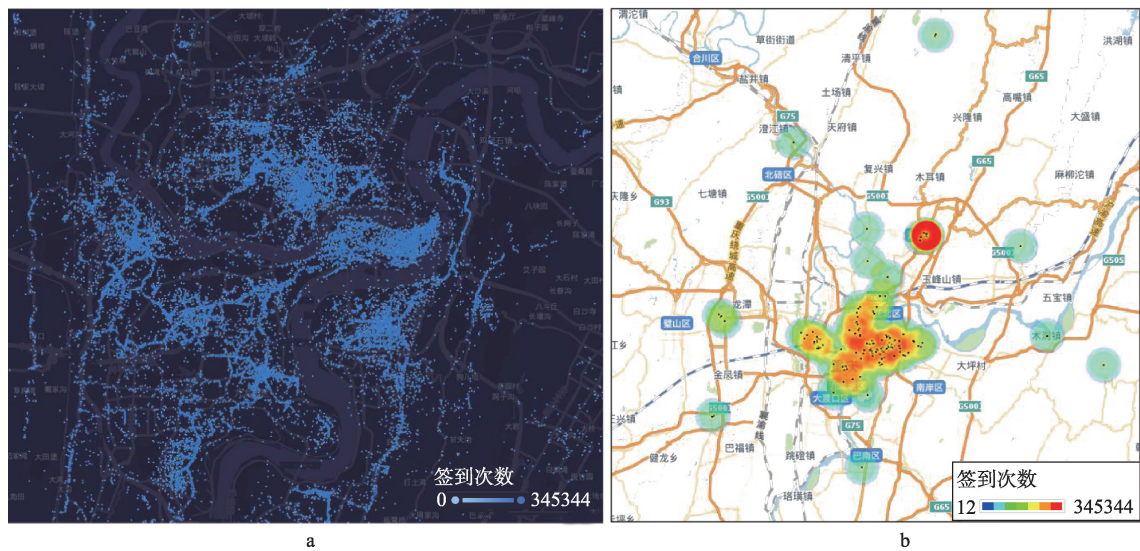


图2 重庆市主城区微博POI点签到地图(a)和签到次数热力图(b)

Fig.2 Sign-in map (a) and heat chart (b) of microblog point of interest (POI) in the main urban area of Chongqing Municipality

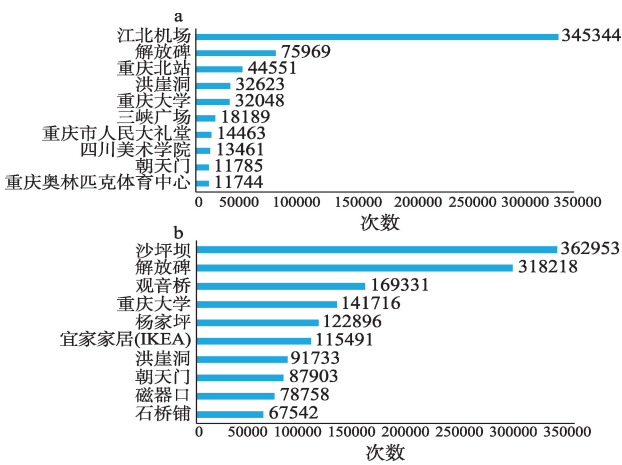


图3 150个POI点中签到次数(a)和提及次数(b)前10位
Fig.3 Top 10 sign-in (a) and mentions (b) of the 150 POI

3.2.1 基于表情情感词典的情绪地图

一般通过表情情感判断用户情绪的准确度是较高的,当微博用户想很明确表达自己的情感时,会专门发出表情符号,所以本文先基于表情研究签到用户的情绪分布。首先需要构建微博表情情感词典,通过整合新浪微博普通表情、动漫表情和魔法表情,最后得到一共包括2008个表情的表情情感词典。通过微博主题中的表情倾向性分数进行情感分析。倾向性分数为正负表情得分之和,其中正表情根据其表达的积极情绪程度取值为1~10,负表情根据其表达的消极情绪程度取值为-1~-10,中立表情为0。通过专家打分法对表情情感词典中2008个表情进行打分,获得各表情的倾向性分数(表1)。

表1 微博表情情感词典(部分)

Tab.1 Microblog emotion dictionary (partial)

表情极性	表情	倾向性分数
正表情	[太开心]	10
	[亲亲]	6
	[馋嘴]	2
中立表情	[话筒]	0
	[思考]	0
	[思考]	0
	[围观]	0
负表情	[黑线]	-1
	[晕]	-5
	[怒]	-9

基于上述表情情感词典,采用Python对每一条带有表情的签到文本数据进行情感分数计算,情感分数正的为积极情感,情感分数负的为消极情感,情感分数零的为中立情感。这样就得到每一条带有表情微博签到文本的情感分数,然后进一步计算每一个POI点上积极情感、消极情感和中立情感的微博数量占比。另外,通过ArcGIS软件采用核密度分析方法(输出像元大小为50 m,搜索半径为1000 m)制作基于表情情感词典的情绪地图(图4a)。

从图4a的情绪地图发现,风貌片区内几乎所有POI点上积极情感微博数量占比都在55%以上,可

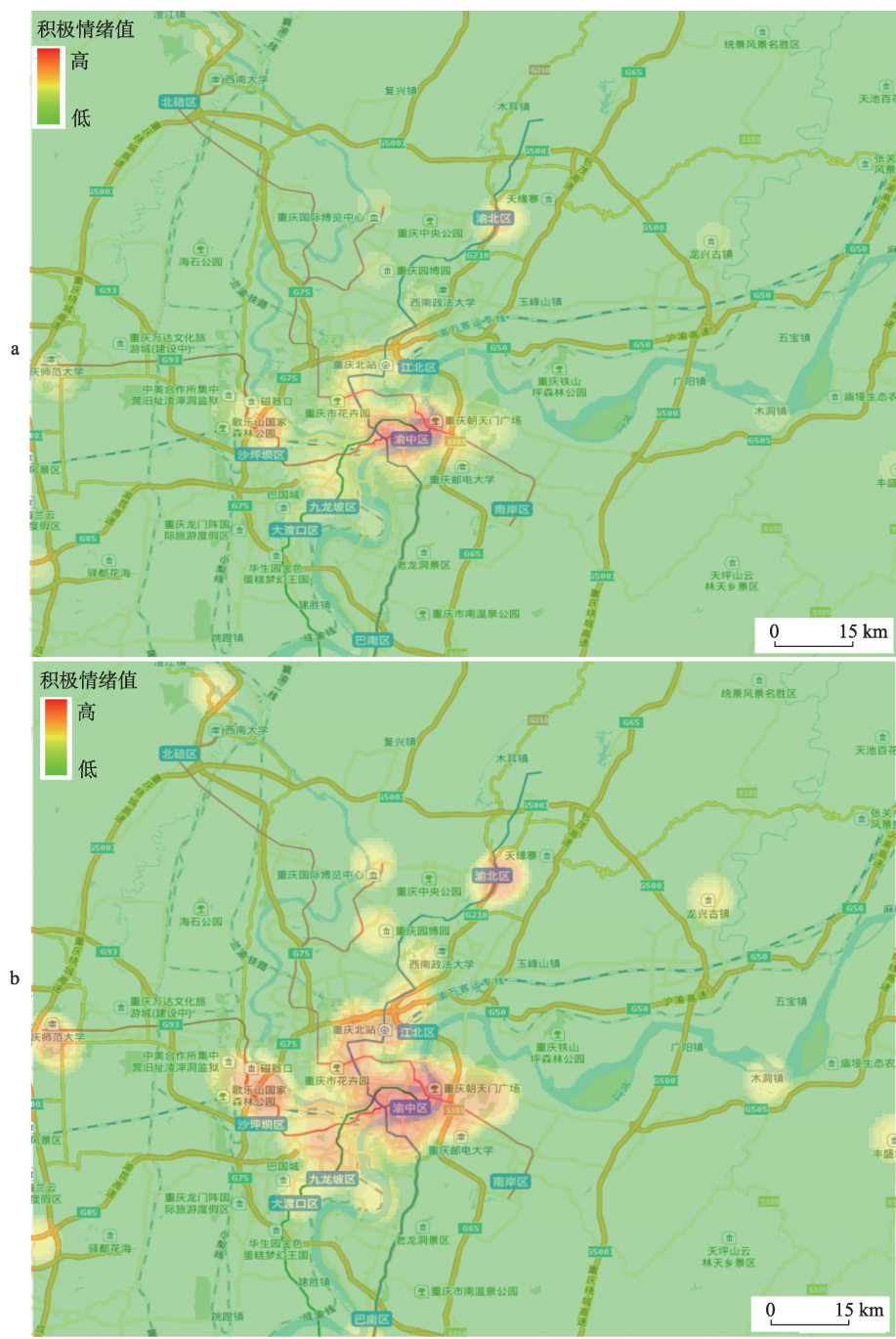


图4 基于表情情感词典的情绪地图(a)和基于文本情感的情绪地图(b)

Fig.4 Emotion map based on the emotion dictionary (a) and emotion map based on the text sentiment (b)

见大家对重庆各个地方的印象还是比较积极的。其中,积极情感微博数量占比排名前几位的为磁器口、一棵树、烈士墓和白公馆等旅游景点;积极情感微博数量占比最低但仍在50%以上的POI点主要为重庆北站北广场、重庆北(龙头寺)火车站、重庆大剧院和重庆喜来登酒店等地。

3.2.2 基于文本情感的情绪地图

进一步将微博表情和文本结合进行情感分析,

采用Python调用处理中文文本的Python库SnowNLP实现。通过对微博文本数据(包含微博表情符号)逐条进行情感分析,预测输入文本数据属于积极和消极的概率,返回的值在0~1之间,将<0.5的划为消极,0.5的划为中立,>0.5的划为积极。这样就可得到每个POI点上积极情感、消极情感和中立情感微博数量占比。同样通过ArcGIS软件采用核密度分析方法(输出像元大小为50 m,搜索半径为

1000 m)制作基于文本情感的情绪地图(图4b)。

基于文本情感的情绪地图显示,风貌片区内大部分POI点上积极情感微博数量占比都在60%以上,均以积极情绪为主。其中,积极情感微博数量占比相对较高的为重庆游乐园、一棵树、重庆大剧院、三峡博物馆和十八梯等地点;积极情感微博数量占比最低但仍在50%以上的POI点为菜园坝火车站广场、重庆江北机场、菜园坝汽车站和沙坪坝火车站等交通枢纽。

由于微博文本数据包含了微博表情数据,在基于文本情感的情绪地图中,积极情绪集中分布的区域范围,比基于表情情感词典的情绪地图更广泛。总体上,用户对特色风貌区的评价以积极情绪为主,微博签到用户情绪较积极的地方主要为旅游景区和城市中心,而情绪相对较低的地方主要为交通枢纽地区,其分析结果可进一步用于后续的对象地图研究。

3.3 对象地图:识别产生积极或消极情绪的主要原因

签到热力分布和情绪分布只是状态,为了寻找背后的原因,需要再次对自然语义进行处理,将其称作对象地图。研究使用玻森(Boson)中文语义开放平台 BosonNLP,识别让用户产生积极和消极情绪的对象。BosonNLP 提供词性分析、情感分析、命名实体识别、典型意见和文本聚类等功能接口,本

文主要用到“典型意见”接口来对每一个POI所发表的微博文本进行文本内容归类,获得多文本的签到数据中用户主要的典型意见,然后再从用户典型意见中人为归类,统计让用户产生积极和消极情绪的对象。通过对文本的阅读,以及大家对重庆主城区的关注内容,将典型意见归类为夜景、道路、城市建筑、城市环境品质、娱乐、饮食和交通等方面进行总结和分析。结合情绪地图的分析结果,通过 ArcGIS 软件对基于文本情感的分析结果进行重分类、栅格转矢量等处理,对区域进行简单的分块,再依次统计每块区域签到用户的典型意见分类。

如图5所示,在渝中半岛核心区域及周围区域(区域1),令用户产生绝对积极情绪的对象为夜景;而该区域的交通、道路让用户产生绝对的消极情绪;对于城市建筑、城市环境品质和娱乐、饮食出现了积极和消极情绪分化的情况。在南山周围区域(区域2),除交通让用户产生绝对消极情绪,夜景、城市建筑等让用户产生绝对的积极情绪外,对于道路、城市环境品质、娱乐和饮食同样出现了积极和消极情绪分化的情况。在观音桥步行街及其周围区域(区域3),饮食、城市环境品质、娱乐和夜景是该区域让用户产生积极情绪的主要对象,而建筑和道路情况让用户产生绝对消极情绪;在沙坪坝磁器口及周边区域(区域4),只有道路和交通产生绝对消极



图5 主要风貌片区对象地图

Fig.5 Object map of cityscape characteristic spatial divisions

情绪,其余娱乐、饮食、城市建筑、环境品质均让用户产生了绝对积极情绪。在奥林匹克体育中心周围区域(区域5),娱乐、夜景和城市建筑令用户产生绝对积极情绪,而城市环境品质和饮食则主要令用户产生消极情绪;在大学城及周围区域(区域6),城市环境品质、建筑、娱乐和饮食等是用户产生积极情绪的对象,而交通和道路是引起用户绝对不满的对象;在江北机场周边区域(区域7),娱乐让用户产生绝对积极情绪,而对于城市建筑则基本为消极情绪,对于城市环境品质产生情绪分化,且主要是消极情绪。

城市风貌所关注的建成环境,对于城市建筑和城市环境品质的评价以积极为主,特别是大学城地区、南滨路地区,解放碑、南滨路等集中的夜景展示区都获得了高度的关注和积极正面的评价。美食之都作为重庆城市文化的载体,受到关注和积极正面的评价。娱乐作为在特色风貌区中人的活动,受到积极的评价。对道路和交通的评价以的消极为主,其中在解放碑城市风貌核心区中,出现了大量关于交通的签到和评价,且基本为负面情感。

3.4 感知评价:风貌要素感知分类

重庆主城区风貌总体设计中提出了4大风貌要

素,分别为山城形、两江势、都市格和巴渝风,研究针对4大要素的典型载体,抽取部分典型意见做进一步分析(表2),可将风貌感知分为4类。多数风貌载体属有感知积极类,如山上城、滨水岸线、四山风光、两江交汇、城市中心、城市夜景、传统风貌区,受到用户的关注和认知,得到积极的评价。部分载体属有感知消极类,如在门户地区,对交通的消极情绪代替了对其形象的关注,最为典型的就是对重庆北站交通组织的不满。部分载体属无感知类,如桥、景观大道。此外,还发现一些有感知未设定类,如长江索道、李子坝穿楼轻轨、四川美术学院、皇冠大扶梯、重庆大学校园等作为受到高度感知的设施或地点,但没有被列入风貌要素。

下文继续尝试回答微博用户视角下的城市感知能为规划师塑造和管控城市风貌提供哪些参考。从要素识别来看,微博用户关注的城市风貌与规划师提炼的多数吻合,对少数受到高度感知但未被列入风貌要素的设施或地点应予以增补,对于现状感知度不高的风貌要素和地点应加强塑造与宣传;从价值评估来看,多数意见反映的正是规划师试图塑造和管控的主题与意象,体现了城市风貌的应有价值,用户在风貌片区的活动偏爱提示规划师应重

表2 四大风貌要素典型意见
Tab.2 Typical opinions on four cityscape characteristic factors

风貌要素	典型载体	典型意见示例	
		积极	消极
山城形	四山	黄葛古道空气好、悠然南山、天然氧吧、欢乐垂钓、踏青、风景无限好、南山看夜景好安逸哦、春意盎然、花的海洋、舒适惬意	无聊、人太多、交通拥堵
	城中山	登高望远美不胜收、鸟瞰重庆、赏山城夜景、眺望重庆森林、无与伦比的美丽	
	山上城	十一楼又是一楼又是四楼的奇妙城市、烟雨朦胧的山城、悬崖上的吊脚楼、洪崖洞现实版千与千寻、洪崖洞古香古色、立体重庆、车水马龙、穿梭在楼里的重庆	
两江势	水岸线	夜跑、春暖花开、夕阳下的渝中半岛、有逼格、很不错、吹吹风喝茶、夜色阑珊、环境很好、喷泉好美、山城江城不夜城、爱情人行道南滨路、江边的风好舒服	3D城市迷路、除了上坡就是下坡、导航都迷路
	桥	无	
巴渝风	两江交汇	两江夜游美景尽收眼底、烟雾蒙蒙美丽江景、码头上真惬意	堵车
	古镇	走马观花、古镇走马还不错值得去、偏岩镇水里吃饭打牌、悠哉悠哉、好大一棵树	
	传统风貌区	重庆人民大礼堂雄伟壮丽、最美街道中山四路、磁器口好山好水好风景、繁华中的废墟十八梯、市井风情、雄伟壮丽的大礼堂、十八梯果然名不虚传已醉已腿抽	
都市格	城市中心	朝天门码头与解放碑CBD组图好看、魅力无限的城市、华丽街灯的街道、大气的奥体中心、博物馆看起来很高大上/到处都是摩天大楼、洋气的重庆解放碑	解放碑显得好渺小
	城市门户	无	坑爹的重庆北站的南北广场、最乱火车站没有之一、无力吐槽此车站只觉得晕
	夜景	夜景好迷人、夜景不错、真好看、很美、美丽、好美、好漂亮、太酷、超美、震撼美、超级棒、霸气、美翻了、美爆、美美哒、美呆、太绚了、大赞、令人心醉、美不胜收、小香港	夜景没有想象中美好
	景观大道	无	无

视公共空间功能策划与活动植入,而消极意见提示规划师应重视的规划问题,如交通的引导与疏解、便利上下坡的交通设施的补充等。对于规划师划定的线性风貌区,如景观大道,要实现整体感知不太可能,应着重于分段塑造和对重点地区的营造。

4 结论

(1) 本文基于新浪微博签到数据,采用语义分析方法,分析挖掘与城市风貌相关词汇的词频、语义关联、情绪表达等内容,对重庆主城区城市风貌进行感知和评价,形成“签到地图”、“情绪地图”和“对象地图”3张图。“签到地图”客观地反映了新浪微博用户在空间上的活跃程度,发现市民或游客在主城区签到活动的地理空间分布和集聚规模主要集中于景区、城市中心、交通枢纽、重要公共设施等;“情绪地图”挖掘用户对空间的情绪表达,反映市民或游客对风貌片区的正负面态度,发现用户对风貌片区的积极情绪多于消极情绪;“对象地图”解析出现这种签到热力分布和情绪分布的原因,通过归类情绪对象,发现用户对于夜景、饮食和娱乐是绝对积极情绪,对于城市建筑和城市环境品质是比较积极的情绪,而对于交通和道路则基本为消极情绪。通过分析用户对规划师所设定的风貌要素载体的典型意见,研究归纳出有感知积极类、有感知消极类、无感知类和感知未设定类共4类风貌载体。研究对规划设定的风貌要素载体进行典型意见梳理后,分别从要素识别和价值判断2个方面为规划师塑造和管控城市风貌提出了增补要素点、增强公共空间功能策划与活动植入、改善交通消极影响、以及线性空间的塑造等参考意见。

(2) 采用社交网络数据对城市风貌特色进行感知与评价是一种新的城市风貌研究方法,还处于摸索阶段,已有研究多以签到分布规律和对文本的情绪分析为主,未涉及原因分析。本文在微博签到和情绪分析的基础上,以对象地图的形式分析了签到热力与情绪分布的原因,并进行了风貌要素感知分类评价。本文尝试解决传统风貌研究中客观性不足和难于量化等问题,研究结论可提升规划师对风貌要素的识别与价值判断,提升规划方案编制的科学性与说服力。城市规划强调以人为中心的价值导向,人本主义的设计必须建立对人类活动和人类感知的认知,而目前仍欠缺好的测定手段与方法。

社交媒体数据由个体自主发布,是个体意见、态度、情绪的巨大数据源,挖掘分析社交媒体数据,将其与城市规划研究相结合具有极大的应用前景。但社交媒体也有局限性,如微博用户的非广泛性导致结论只反映了部分人群的观点和态度,文本的泛主题性会增加数据处理的难度,例如本文中大量在机场的签到和文本都是关于到达与离开,与城市门户形象无关联。未来需要设法获取更广泛的数据源,不断优化文本挖掘和分析工具,提高数据的针对性,以便更多地在城市规划设计中运用。

参考文献(References)

- 戴慎志,刘婷婷. 2013. 面向实施的城市风貌规划编制体系与编制方法探索[J]. 城市规划学刊, (4): 101-108. [Dai S Z, Liu T T. 2013. A study on cityscape planning system and planning methods from the view of administration and implementation[J]. Urban Planning Forum, (4): 101-108.]
- 段德罡,孙曦. 2010. 城市特色、城市风貌概念辨析及实现途径[J]. 建筑与文化, (12): 79-81. [Duan D G, Sun X. 2010. City characteristics, city style concept and approach[J]. Architecture & Culture, (12): 79-81.]
- 纪娇娇,申帆,黄晟鹏,等. 2015. 基于语义网络分析的微信公众平台转基因议题研究[J]. 科普研究, 10(2): 21-29. [Ji J J, Shen F, Huang S P, et al. 2015. A semantic network analysis of issues about genetically modified food on wechat official account admin platform[J]. Science Popularization, 10(2): 21-29.]
- 凯文·林奇. 2001. 城市意象[M]. 方益萍,何晓军,译. 北京: 华夏出版社. [Kevin L. 2001. Chengshi yixiang[M]. Fang Y P, He X J, Trans.. Beijing, China: Huaxia Press.]
- 林江豪,阳爱民,周咏梅,等. 2012. 一种基于朴素贝叶斯的微博情感分类[J]. 计算机工程与科学, 34(9): 160-165. [Lin J H, Yang A M, Zhou Y M, et al. 2012. Classification of microblog sentiment based on naïve bayesian[J]. Computer Engineering & Science, 34(9): 160-165.]
- 马武定. 2009. 风貌特色: 城市价值的一种显现[J]. 规划师, 25(12): 12-16. [Ma W D. 2009. Cityscape: An indication of city value[J]. Planners, 25(12): 12-16.]
- 秦萧,甄峰,熊丽芳,等. 2013. 大数据时代城市时空行为研究方法[J]. 地理科学进展, 32(9): 1352-1361. [Qin X, Zhen F, Xiong L F, et al. 2013. Methods in urban temporal and spatial behavior research in the Big Data Era[J]. Progress in Geography, 32(9): 1352-1361.]
- 王德,张昀,崔昆仑. 2009. 基于SD法的城市感知研究: 以浙江台州地区为例[J]. 地理研究, 28(6): 1528-1536. [Wang D, Zhang Y, Cui K L. 2009. Study of city perception based on SD method: A case study of Taizhou, Zhejiang[J]. Geographical Research, 28(6): 1528-1536.]

- 王建国. 2007. 城市风貌特色的维护、弘扬、完善和塑造[J]. 规划师, 23(8): 5-9. [Wang J G. 2007. Conservation, improvement, integrity, and construction of urban feature[J]. Planners, 23(8): 5-9.]
- 微博数据中心. 2017. 2016 微博用户发展报告[EB/OL]. 2017-01-11[2017-08-13]. <http://data.weibo.com/report/reportdetail?id=346>. [Weibo Shuju Zhongxin. 2017. 2016 weibo yonghu fazhan baogao[EB/OL]. 2017-01-11[2017-08-13]. <http://data.weibo.com/report/reportdetail?id=346>.]
- 杨昌新. 2014. 城市风貌研究的历史视野[J]. 城市问题, (4): 30-36, 73. [Yang C X. 2014. The historical horizon of cityscape research[J]. Urban Problems, (4): 30-36, 73.]
- 杨振山, 龙瀛, Douay N. 2015. 大数据对人文—经济地理学研究的促进与局限[J]. 地理科学进展, 34(4): 410-417. [Yang Z S, Long Y, Douay N. 2015. Opportunities and limitations of big data applications to human and economic geography: The state of the art[J]. Progress in Geography, 34 (4): 410-417.]
- 张劲杰. 2015. 基于博客视角的日照游客行为研究[D]. 曲阜: 曲阜师范大学. [Zhang M J. 2015. Jiyu boke shijiao de Rishao youke xingwei yanjiu[D]. Qufu, China: Qufu Normal University.]
- 赵渺希, 邵琳, 林韵莹. 2013. 外地游客与本地居民的城镇空间感知差异: 以广东省佛山市南海区西部城镇为例[J]. 旅游科学, 27(2): 46-58. [Zhao M X, Shao L, Lin Y Y. 2013. Perceptual differences of urban space between visitors and local residents: A case study of west towns in Nanhai, Guangdong Province[J]. Tourism Science, 27(2): 46-58.]
- 甄峰, 王波, 秦萧, 等. 2015. 基于大数据的城市研究与规划方法创新[M]. 北京: 中国建筑工业出版社. [Zhen F, Wang B, Qin X, et al. 2015. Urban studies and innovation in urban planning methods based on Big Data[M]. Beijing, China: China Architecture & Building Press.]
- Fan R, Zhao J C, Chen Y, et al. 2014. Anger is more influential than joy: Sentiment correlation in Weibo[J]. PLoS One, 9 (10): e110184.
- Go A, Bhayani R, Huang L. 2009. Twitter sentiment classification using distant supervision[R]. Project Report CS224N, Palo Alto, CA: Stanford University: 12.
- Kouloumpis E, Wilson T, Moore J. 2011. Twitter sentiment analysis: The good the bad and the omg[C]//Proceedings of the 5th international AAAI conference on weblogs and social media. Barcelona, Spain: AAAI: 164.
- Saif H, He Y L, Alani H. 2012. Semantic sentiment analysis of twitter[C]//Proceedings of the 11th international semantic web conference on the semantic web: ISWC 2012. Berlin Heidelberg, Germany: Springer: 508-524.

Perception and evaluation of cityscape characteristics using semantic analysis on microblog in the main urban area of Chongqing Municipality

YI Zheng^{1,2}, LI Jizhen¹, LENG Bingrong¹, CHEN Min¹

(1. Chongqing Planning & Design Institute, Chongqing 401147, China;

2. Sichuan Fine Arts Institute, Chongqing 400053, China)

Abstract: Using social network data to apperceive and evaluate cityscape characteristics is a new method for the quantitative study of urban features. Taking the sign-in data of Sina microblog as data sources of cityscape characteristics perception, this study used text mining and semantic analysis to explore the perception and evaluation of cityscape characteristics in the main urban area of Chongqing Municipality. The study resulted in three maps about the main urban area of Chongqing Municipality, including the sign-in map, the emotion map, and the object map. The sign-in map objectively reflects the distribution of level of activity of the users; the emotion map mines users' emotional expression of positive and negative attitudes on their sign-in space; the object map analyzes the reasons behind the activity distribution and emotional expression to find out urban features that cause positive and negative attitudes. By analyzing the users' typical opinions on urban features designed by planners, the factors are classified into four categories: perceived and positive, perceived and negative, unconscious, and perceived but unspecified. The study provides further references on factor identification and quality evaluation to planners for shaping and controlling urban features.

Key words: perception and evaluation of cityscape characteristics; Sino microblog; semantic analysis; Chongqing Municipality; the sign-in map; the emotion map; the object map