

基于游记语义的游客情绪空间特征研究

——以安徽省黄山景区为例

蔡伟娜¹, 刘超^{1,2}, 许乐天¹

(1. 中国地质大学(武汉)地理与信息工程学院, 湖北 武汉 430078;

2. 中国地质大学(武汉)国家公园实验室, 湖北 武汉 430074)

摘要: 游记是游客表达感知情绪的一种重要方式。对现有游记进行情绪空间特征分析, 可以为游客旅游路线规划和景区形象定位提供指导。基于安徽省黄山景区游记文本, 通过情绪词汇提取和语义分析, 构建情绪表达度和情绪积极度两个指数, 并借助反距离权重插值法和空间自相关莫兰指数进行情绪空间特征分析。研究表明: 黄山景区游客情绪表达度总体较低, 高表达度节点呈线状集中分布, 情绪积极度总体较高, 且山上明显高于爬山途中; 具有文化背景的自然景观节点, 如飞来石、猴子观海、始信峰等, 更易表现出高表达度—高积极度的情绪特征; 表达度和积极度的莫兰指数均接近于0, 表明节点情绪的空间自相关性低, 即节点上的游客情绪不会受到周围节点的影响。

关键词: 游客情绪; 情绪表达度; 情绪积极度; 空间特征; 黄山

中图分类号: F59 **文献标识码:** A **文章编号:** 1674-5841(2022)01-0042-19

情绪在心理学上被认为是一种人类对客观事物感觉和知觉的外部性反应, 可以通过表情、肢体语言以及文字语言等方式表达, 通常伴随着特定的动机, 会对个体本身及其周围人群的行为产生一定影响^{[1]2}。游客情绪即游客在旅游活动中产生的感觉和知觉的表达, 感觉来源于旅游地客观事物的直接刺激, 知觉则受到游客在旅游活动发生前的经历记忆和意识的影响。情绪是游客一种较为强烈的感知表达, 而旅游游记中体现的游客情绪是经过从感知形象到记忆意象再到表达语象的逐步衰减^{[2]133}而后表达出的语义文字。因

作者简介: 蔡伟娜(2000—), 女, 安徽滁州人, 2018级本科生, 研究方向为旅游地理学; 许乐天(1994—), 男, 湖北黄冈人, 2019级硕士研究生, 研究方向为旅游地理学、旅游地多方参与管理。

通讯作者: 刘超(1980—), 男, 湖北鄂州人, 副教授, 博士, 硕士生导师, 研究方向为人文地理、游客时空行为。

此, 在游记文本中提取的情绪语义可以代表游客在整个旅行过程中印象最深刻的情绪体验, 直接对游客本人、游记读者及景区管理人员产生真实有效影响, 可以作为游客情绪分析的数据源^{[3]1}。

目前, 学者对游客情绪的研究大致可以归为三类。第一类是探究游客情绪的影响因素, 包括媒体宣传^{[4]610}、大众口碑^{[4]610}、景区服务^[5-7]、旅游资源^[8-9]、旅游互动^[10]等; 第二类为游客情绪的效应研究^[11-14], 包括对游客认知^[15]、游客满意度^{[16]143}等方面的影响; 第三类是旅游情绪的测度方法研究。这类研究的普遍做法是采用问卷、访谈的形式对主体情绪进行量化^{[16]144, [17]}, 但问卷质量会对评价结果产生较大影响。Shanshi Li 等^[18]则通过皮肤电分析、面部肌肉活动等生理测量法对游客情绪进行测度。近年来, 少数学者^{[3]11, [19]154, [20-21]}将文本挖掘和语义分析引入到该领域中, 取得了良好效果。

游客情绪本身的空间特征对游客路线设计规划和景区管理的参考作用, 目前还处于被忽视的状态。从游客情绪特征出发, 沿原始数据处理流程向上溯源, 可以找到游客情绪产生最的原始原因, 保留最真实的分析结果。本研究通过挖掘游记文本中的情绪词并进行分类统计, 构建情绪表达度和积极度两个指标, 从情绪表达程度和内容两个维度刻画黄山景区的游客情绪特征, 以期帮助游客根据个人偏好选择合适的游览景点, 从而提高整个旅游行程的满意度。此外, 旅游景区也可以根据真实的游客情绪反馈, 通过调整规划设计来打造或提升景区形象。

一、数据来源与研究方法

(一) 数据来源

根据“认知—情感—意动”理论, 游客在旅游过程中受景观刺激产生感知, 进而产生情感上的波动, 从而进行撰写游记等意向行动^[22]。游记作为旅游活动之后游客自发输出的语义文本, 不会受到研究者设定的引导性情境的污染, 可以客观表达不同景点位置上游客真实的心理情绪状态^[23-24]。目前, 已有相关学者把游记作为数据源应用到旅游研究^[25-27], 利用爬虫软件在旅游出行网站上进行不设限游记爬取, 确保样本数据涵盖不同出游方式、不同游伴数量、不同出游时间、不同人均日消费, 从而获得具有良好代表性的样本^[28]。保继刚等曾以旅游游记和评论为数据源研究游客情感特征和评价旅游目的地情感, 并取得了满意的效果^[29-30]。本研究将从网络游记中提取游客的情绪词汇, 进而刻画游客情绪的空间特征。

自2000年起, 众多游客自发地在携程网站的游记板块分享旅游经历, 创作出大量网络游记, 为学者研究旅游提供了一个可靠的数据源^{[2]130}。为了规避旅游季节性波动的影响, 游记文本的选取需覆盖一整年的全部旅游季节

和重要时间节点^[31]。利用八爪鱼采集器在携程网站爬取 2015 ~ 2018 年的游记文本, 获取 2017 年的游记 183 篇, 在数量上与这 4 年的平均水平 193 篇相近, 同时没有受到 2019 年末暴发的新冠肺炎疫情的影响, 具有普遍性和代表性。另外, 根据 2011 ~ 2020 年百度指数数据, 黄山的日均值达到 6919, 在同类型景区中名列前茅。在知网期刊库中以“旅游”分别并含“黄山”“华山”“泰山”“庐山”“丹霞山”五座国内著名山岳景观地为关键词进行检索, 发现 2010 ~ 2019 年这 10 年来, 有关黄山的年发文量遥遥领先, 每年的年发文量占比均超过 40%, 部分年份甚至超过 50%, 如图 1 所示。由此可见, 黄山的游客关注度和学术关注度极高, 以其作为案例地来探索游客情绪具有较强的说服力。

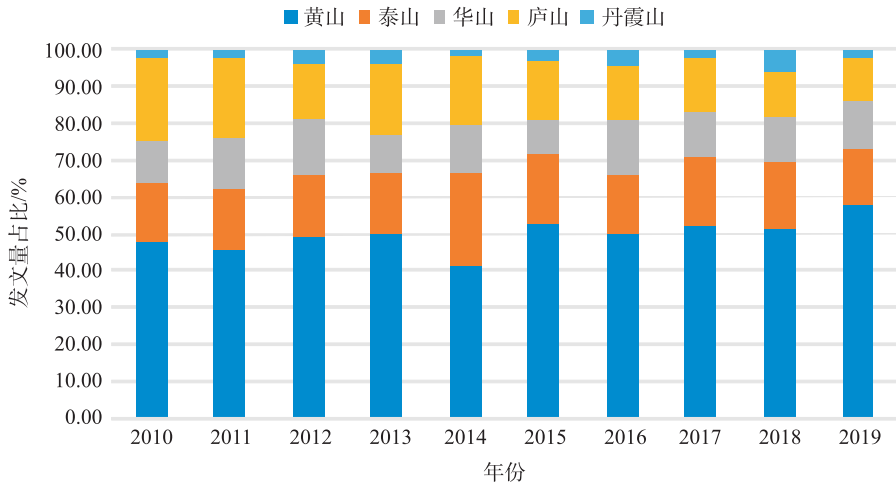


图1 国内著名山岳型景区年发文量统计

因此, 本文以在携程网站无差别爬取得到的 2017 年全年以黄山作为目的地的 183 篇游记作为数据来源。在经过仔细阅读后, 删除重复、广告类假游记和无关游记, 最终选取其中的 164 篇作为样本数据。

根据舍费尔^[32]等人的抽样公式计算样本规模:

$$n = \frac{N}{(N-1) \cdot g^2} + 1 \quad \text{公式 (1)}$$

式 (1) 中: n 为样本规模; g 为抽样误差; N 为母体大小。以 2017 年黄山景区游客量 337 万人次作为母体大小, 10% 的抽样误差, 计算的样本规模为 101, 即 164 篇游记样本在 90% 的置信水平下符合抽样要求。

(二) 研究方法

1. 情绪词编码

扎根理论是一种自下而上的理论构建方法^[33], 在对原始资料进行概念化、范畴化和理论抽象化等步骤后, 最终归纳出一套理论体系, 用以刻画研

究对象的某种属性。本研究依托扎根理论对游记文本中表达游客情绪的词汇进行编码归类, 具体步骤如下所述。首先, 参考知网 HOWNET 情感词典, 从积极情绪和消极情绪两个维度构建了自定义情绪词汇表。其次, 利用 ROST CM6 (ROST ContentMining, 后文缩写为 ROST) 词频分析软件对能够表现游客情绪的词汇进行了筛选, 并将其补充进情绪词汇表。筛选标准如下: (1) 直接表达情绪的词汇, 如“高兴”“悲伤”等; (2) 对旅游地项目和服务带有感情色彩的评价词, 如“秀美”“无聊”等; (3) 反映游客本身心理状态的词汇, 如“舒适”“平静”等。根据情绪词分词和筛选结果, 结合上下文语境核对样本游记中的每一个情绪词语义, 将意思相近的情绪词归并为同一类情绪主题建立归并分类词表 (见表1), 以规避游客用词差异对研究结果产生的负面影响。最后, 按照情绪主题分类, 结合自定义词表和归并词表利用 ROST 软件辅助主题情绪词频统计。

表1 游客情绪分类词表

情绪极性	情绪主题	具体情绪词汇 (归并词汇)
积极情绪	开心	幸福 开心 惬意 喜悦 高兴 快乐 愉悦 笑着 欢乐 欢快 欢笑 快活 兴冲冲 兴奋 激动 欢聚 喜出望外 心情大好 好开森 亢奋 得意
	美感	美景 雄伟 开阔 气势 漂亮 秀丽 好看 风采 秀美 优美 叹为观止 着迷 江山如画 豪迈 拍案叫绝 最美 很美 多美
	舒适	方便 舒服 舒适 自在 舒坦 随心 自得 宜人
	奇特	神奇 神秘 奇特 奇妙 奇迹 好奇 新奇
	震撼	震撼 折服 惊诧 惊讶 惊喜 震惊
	赞叹	赞叹 赞美 欣赏 观赏
	满意	满足 圆满 顺利 满意 不舍
	安静	安静 宁静 安详
	期待	向往 愿景 期待
	有趣	有趣 收获
消极情绪	遗憾	遗憾 可惜 失望 后悔 无奈
	劳累	疲惫 累趴 劳累
	惊险	危险 惊险 紧张 害怕 恐怖 发憷 胆颤 胆战心惊 恐惧 挑战
	其他	无望 扫兴 气愤 惆怅 凄迷 痛苦 彷徨 无聊 徘徊 崩溃 忧伤郁闷

资料来源: 作者利用 ROST 软件对黄山游记进行分析归纳。

2. 节点情绪具象化

由于游记是游客基于回忆意象创作的, 因此在创作时很容易受到主体感知情况的影响。具有标志性旅游吸引物的景点, 更容易在游客脑海中留下相

对强烈的情绪印象,从而被游客利用文字表达出来。基于此,本文选取了66个黄山景观,并将这些具有点、线、面三种不同形态特征的景观在地图上具化为具有明确坐标的节点进行分析。对猴子观海、飞来石等具有明确的坐标位置的点状景观,将坐标点具化为空间节点;对云谷索道、玉屏索道等线状景观,则取中点位置作为其空间节点;而对西海大峡谷等面状景观或景区,则直接将百度地图提供的坐标点具化为空间节点。

在此基础上,本研究对164篇游记中出现的情绪词频数进行了统计。由于汉语的特殊属性,具体的情绪词可能会因为语境的不同而产生含义上的变化,从而对最终的情绪识别产生影响。因此,本研究没有使用ROST等情感分析软件,而是采用人工判读的方式对情绪词在游记文本中出现的位置进行逐一定位,然后将其与节点进行匹配,并按照如表1所示的情绪词分类结果进行统计,得到每一个节点各类情绪词频矩阵,从而将节点代表的情绪特征进行具象化表达。

3. 情绪指数构建

在后旅游活动中,游客经由游记表达自己对于旅游活动的印象体验,并试图通过这种表达影响人们对于特定景观的先验感知,这个过程在心理学上被称为“印象管理”^[34-35]。这一表达的过程,往往都包含着主体的需要和意愿^[13],是游客对于自身情绪的一种再加工过程。从这个意义上说,单纯地以词频高低来判断节点的情绪极性并不合适。基于此,本研究引入情绪表达度、情绪积极度两个概念,分别从游客对于节点的情绪印象、节点的情绪表达状况两个方面对于节点情绪进行解读。

(1) 情绪表达度

通常情况下,游客会在游记中针对特定景点的视觉形象、记忆意象进行情绪表达^{[2]133}。因此,情绪词出现的频数在一定程度上可以反映游客对于该景点的情绪印象的强弱。本研究通过计算节点情绪词的篇频数来测度游客对于节点的情绪印象,具体公式如下:

$$E_i = \frac{c_i}{n} \quad \text{公式(2)}$$

式(2)中: E_i 为第*i*个节点的情绪表达度; c_i 为所有样本数据中第*i*个节点的情绪词总频数; n 为样本量。 E_i 的值越高,表明游客针对该节点的情绪表达越频繁,即该景点能够给游客留下更多的情绪印象。

(2) 情绪积极度

在对节点间的情绪进行比较前,首先需要确定游客对各个节点的情绪表达状况。本研究通过计算积极情绪词在节点情绪词总数中所占的比重对情绪表达状况进行判读,具体公式如下:

$$L_i = \frac{c_{ij}}{c_{ij} + c_{ix}} \quad \text{公式 (3)}$$

式 (3) 中: L_i 为第 i 个节点的积极度; c_{ij} 、 c_{ix} 分别为所有样本数据中第 i 个节点对应的积极情绪词总频数和消极情绪词总频数。其中, $0 \leq L_i \leq 1$, 其值越接近于 1, 说明该节点上游客表达出的情绪越积极。

4. 空间特征分析

首先, 利用地理信息系统软件 ArcGIS 中反距离权重插值工具^[36-37], 将分级后的游客情绪指标进行可视化分析, 分别得到节点情绪表达度散点拟合图 (见图 2) 和节点情绪积极度散点拟合图 (见图 4)。其次, 利用 GeoDa 的空间自相关分析工具包, 分别计算情绪表达度和情绪积极度的全局 Moran's I 指数^[38], 并将节点情绪的空间自相关性进行数值化表达, 以判断游客情绪的空间分布是否具有集聚特征以及该特征是否具有显著性。

二、结果分析

(一) 情绪表达度分析

1. 情绪表达度分级

利用研究方法中所构建的情绪表达度公式计算每个节点的情绪表达度, 并对结果进行升序排列, 得到节点情绪表达度汇总表 (见表 2)。将节点情绪表达度按照表 2 中的顺序依次绘制到图上, 并以指数型函数做拟合, 即可得到节点情绪表达度散点拟合图 (见图 2)。基于与函数的拟合效果, 本研究分别选取 $i=48$ 和 $i=60$ 两个节点对应的情绪表达度 0.09、0.2 作为阈值, 将游客情绪表达度划分为三个等级 (见表 3)。

表 2 节点情绪表达度汇总表

序号	节点	表达度	序号	节点	表达度	序号	节点	表达度
1	云谷寺	0.006	23	鳌鱼洞	0.024	45	光明顶山庄	0.085
2	通天塔	0.006	24	翡翠谷	0.024	46	北海宾馆	0.085
3	仙人翻桌	0.006	25	云谷景区	0.024	47	白鹅岭	0.085
4	廊桥	0.006	26	步仙桥	0.030	48	百步云梯	0.091
5	仙人指路	0.006	27	排云楼宾馆	0.030	49	玉屏索道	0.091
6	龙爪松	0.006	28	探海松	0.030	50	西海饭店	0.098
7	竖琴松	0.006	29	玉屏峰	0.030	51	猴子观海	0.098
8	团结松	0.006	30	猴谷	0.030	52	清凉台	0.098
9	太平索道	0.012	31	北大门	0.037	53	白云宾馆	0.110
10	梦幻景区	0.012	32	鲫鱼背	0.043	54	丹霞峰	0.116
11	飞龙瀑	0.012	33	揽海亭	0.043	55	玉屏景区	0.116
12	群峰顶	0.012	34	一线天	0.049	56	北海景区	0.122
13	玉屏楼	0.012	35	玉屏楼宾馆	0.049	57	莲花峰	0.146

表2 (续)

序号	节点	表达度	序号	节点	表达度	序号	节点	表达度
14	仙人晒靴	0.012	36	狮子峰	0.049	58	黑虎松	0.159
15	南大门	0.012	37	西海山庄	0.049	59	始信峰	0.171
16	梦笔生花	0.012	38	白云景区	0.055	60	排云亭	0.171
17	回音壁	0.012	39	鳌鱼峰	0.061	61	天都峰	0.232
18	慈光阁	0.012	40	石笋矽	0.061	62	飞来石	0.256
19	连理松	0.012	41	白鹅山庄	0.073	63	云谷索道	0.311
20	雨伞松	0.012	42	九龙瀑	0.073	64	迎客松	0.311
21	天海景区	0.024	43	狮林大酒店	0.073	65	光明顶	0.530
22	石笋峰	0.024	44	西海景区	0.079	66	西海大峡谷	0.726

资料来源：作者利用 ROST 统计及 Excel 计算输出。

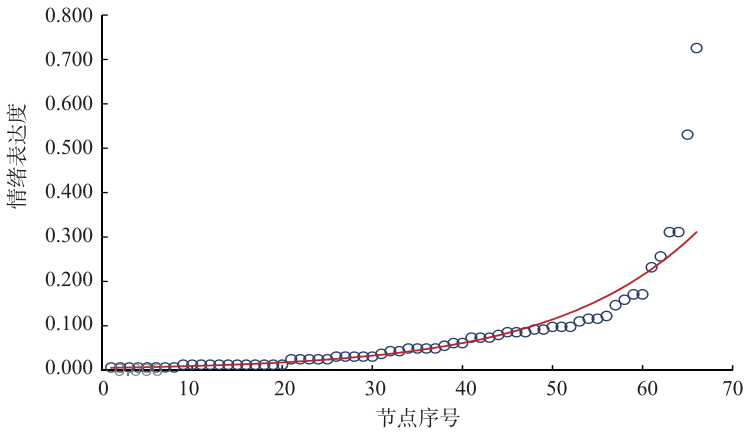


图2 节点情绪表达度散点拟合图

表3 情绪表达度分级表

表达度范围	表达等级	节点
$0.2 \leq E$	高	天都峰 飞来石 云谷索道 迎客松 光明顶 西海大峡谷
$0.09 < E < 0.2$	较高	百步云梯 玉屏索道 西海饭店 猴子观海 清凉台
		白云宾馆 丹霞峰 玉屏景区 北海景区 莲花峰
$E \leq 0.09$	低	黑虎松 始信峰 排云亭
		云谷寺 通天塔 仙人翻桌 廊桥 仙人指路 龙爪松
		竖琴松 团结松 太平索道 梦幻景区 飞龙瀑
		群峰顶 玉屏楼 仙人晒靴 南大门 梦笔生花
		回音壁 慈光阁 连理松 雨伞松 天海景区 石笋峰
		鳌鱼洞 翡翠谷 云谷景区 步仙桥 排云楼宾馆
		探海松 玉屏峰 猴谷 北大门 鲫鱼背 揽海亭
		一线天 玉屏楼宾馆 狮子峰 西海山庄 白云景区
		鳌鱼峰 石笋矽 白鹅山庄 九龙瀑 狮林大酒店
		西海景区 光明顶山庄 北海宾馆 白鹅岭

资料来源：作者利用 ROST 统计及 Excel 计算输出。

从结果上看, 不同节点间的情绪表达度存在较大差异, 出现明显的分级现象, 结合相关学者对于黄山景观关注度研究^{[2]129-135}, 可以看出分级与节点的知名度存在一定关联。西海大峡谷、光明顶、迎客松等知名节点的情绪表达度高居前列, 而云谷寺、通天塔、仙人翻桌等知名度相对较低的节点的表达度则整体偏低。整体而言, 高情绪表达度的节点数量偏少, 大部分节点的情绪表达度偏低, 未能给游客留下较多的情绪印象。

2. 表达度空间特征分析

将情绪表达度数据导入 ArcGIS 中, 利用反距离权重插值法进行景区全范围内插值, 并以不同颜色和形状对表达度等级进行呈现, 实现了情绪表达度在空间上的可视化(见图3)。从空间特征上来看, 低表达度节点数量较多且沿旅游路线连片分布, 高表达度节点数量较少且位置分散。这是因为游客对具体景观的情绪表达会经历“景象—视象—意象—语象”的逐层衰减的传递过程, 进而导致能给游客留下深刻印象且在游记中表达的节点数量并不多^{[2]133}。从线路上来看, “鲫鱼背—排云亭”沿线的高表达度节点最多, 涵盖了“飞来石”“光明顶”“迎客松”“天都峰”四个高表达度节点。在该线路上, 游客的情绪表达欲望明显增强。因此, 可以将其作为代表线路进行进一步的宣传和推广。

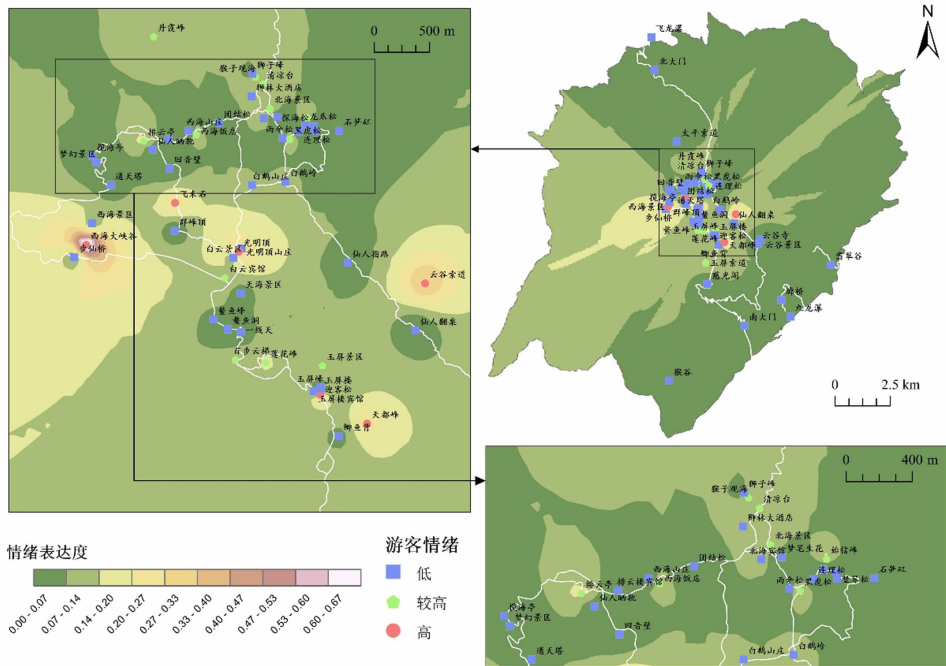


图3 情绪表达度空间分布

为了进一步了解节点在空间上的分布状况, 以及节点之间的位置关系是否会

对表达度结果造成影响,本研究引入了全局莫兰指数进行节点间的空间自相关分析。利用 GeoDa 软件进行计算,所得结果为 -0.078 。从数值上看,不同节点对应的情绪表达度之间没有明显的空间自相关性,其分布在空间上具有一定的随机性,即游客在某一节点上的情绪表达状况不会受到其他节点的影响。

(二) 情绪积极度分析

1. 情绪积极度分级

利用公式(3)计算每个节点的情绪积极度,并对结果进行升序排列,得到节点情绪积极度汇总表(见表4)。将节点情绪积极度按照表2的顺序依次绘制到图上,并以对数型函数做拟合,得到节点情绪积极度散点拟合图(见图4)。基于其数值聚类特征,本研究分别选取拟合图中的两个自然断点处的情绪积极度值0.6和1作为阈值,对情绪积极度进行聚类划分(见表5)。

从结果上看,黄山景区内游客情绪积极度总体较高,除了云谷寺、太平索道等少数节点外,其余节点的情绪积极度都较高,部分节点甚至出现情绪积极度值为1的情况。整体而言,节点数量与情绪等级之间呈现出漏斗式结构,即情绪等级越高,对应的节点数量也越多。这表明游客对黄山风景区的印象较好,整体的情绪感知较为积极。

表4 节点情绪积极度汇总表

序号	节点	积极度	序号	节点	积极度	序号	节点	积极度
1	云谷寺	0.000	14	白鹅山庄	0.750	27	飞来石	0.857
2	太平索道	0.000	15	天海景区	0.750	28	揽海亭	0.857
3	鲫鱼背	0.286	16	西海大峡谷	0.765	29	玉屏索道	0.867
4	百步云梯	0.467	17	北海宾馆	0.786	30	莲花峰	0.875
5	梦幻景区	0.500	18	鳌鱼峰	0.800	31	玉屏楼宾馆	0.875
6	步仙桥	0.600	19	排云楼宾馆	0.800	32	白云景区	0.889
7	西海饭店	0.688	20	探海松	0.800	33	玉屏景区	0.895
8	光明顶山庄	0.714	21	天都峰	0.816	34	白鹅岭	0.929
9	白云宾馆	0.722	22	排云亭	0.821	35	清凉台	0.938
10	丹霞峰	0.737	23	云谷索道	0.824	36	迎客松	0.941
11	始信峰	0.750	24	北大门	0.833			
12	猴子观海	0.750	25	西海景区	0.846	37~66	飞龙瀑等 30个景区 ^①	1.000
13	一线天	0.750	26	光明顶	0.851			

资料来源:作者利用 ROST 统计及 Excel 计算输出。

^①30个景区依次为:飞龙瀑、九龙瀑、狮子峰、石笋峰、玉屏峰、群峰顶、通天塔、仙人翻桌、廊桥、猴谷、石笋缸、玉屏楼、仙人指路、仙人晒靴、南大门、梦笔生花、鳌鱼洞、回音壁、慈光阁、翡翠谷、西海山庄、狮林大酒店、龙爪松、竖琴松、团结松、连理松、雨伞松、黑虎松、云谷景区、北海景区。

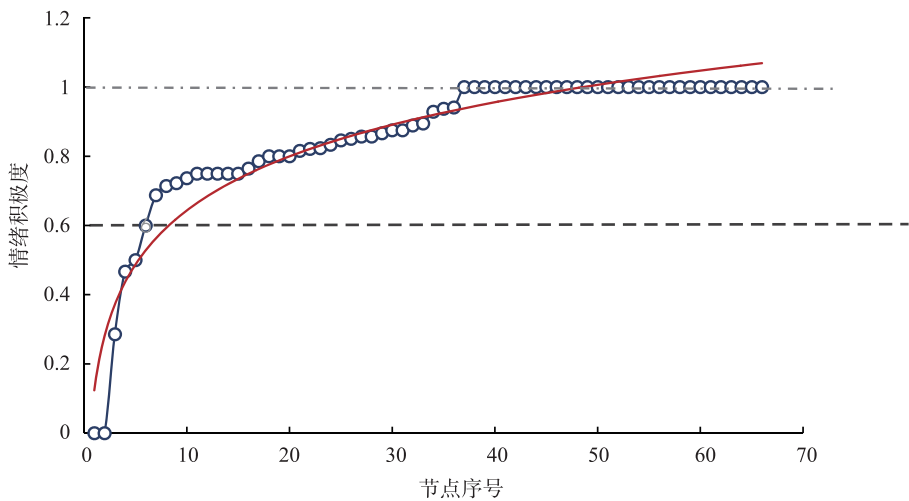


图4 节点情绪积极度散点拟合图

表5 情绪积极度分级表

积极度范围	节点情绪等级	节点
$L = 1$	积极	飞龙瀑 九龙瀑 狮子峰 石笋峰 玉屏峰 群峰顶 通天塔 仙人翻桌 廊桥 猴谷 石笋缸 玉屏楼 仙人指路 仙人晒靴 南大门 梦笔生花 鳌鱼洞 回音壁 慈光阁 翡翠谷 西海山庄 狮林大酒店 龙爪松 竖琴松 团结松 连理松 雨伞松 黑虎松 云谷景区 北海景区
$0.6 < L < 1$	比较积极	西海饭店 光明顶 山庄 白云宾馆 丹霞峰 始信峰 猴子观海 一线天 白鹅山庄 天海景区 西海大峡谷 北海宾馆 鳌鱼峰 排云楼宾馆 探海松 天都峰 排云亭 云谷索道 北大门 西海景区 光明顶 飞来石 揽海亭 玉屏索道 莲花峰 玉屏楼宾馆 白云景区 玉屏景区 白鹅岭 清凉台 迎客松
$L \leq 0.6$	消极	云谷寺 太平索道 鲫鱼背 百步云梯 梦幻景区 步仙桥

资料来源: 作者利用 ROST 统计及 Excel 计算输出。

2. 积极度空间特征分析

将情绪积极度数据导入 ArcGIS 中, 利用反距离权重插值法进行景区全范围内插值, 并以不同颜色和形状对不同等级进行呈现, 实现了情绪积极度在空间上的可视化 (见图 5)。总体来看, 情绪节点呈明显的线状分布, 与游览道路相契合; 积极节点在通天塔至石笋缸的东西向条带状空间内呈现集聚特性, 此区域内黄山松景观较多, 有黑虎松、雨伞松、团结松等近 10 个奇松景观, 说明黄山松景观给游客带来高积极度的情绪体验; 比较积极节点相对均

匀地分布在游览路线周围，从光明顶到天都峰的西北至东南向路段几乎全为比较积极节点；消极情绪节点总体数量少，主要产生原因可以归纳为两类。

第一类是因为地势陡峭而产生的惊恐之感。如：“一共两条路上天都峰，新路和老路，很多90度的大斜坡，并且好多地方没有扶手和绳子感觉有点危险。我是从新路上，老路下，下山的时候经过鲫鱼背真是被吓尿了，不要问我鲫鱼背长什么样，因为我过的时候脑子一片空白”“我翻出临出发前酒店给的地图看了看，离光明顶也不远了，海拔高差也不大了。走起，继续前行，不休息。光明顶的海拔高度1860m，这一段路又是爬山——发憊啊！”“接着就是百步云梯，这个是我觉得最恐怖的，比天都峰还恐怖，因为那个阶梯是倾斜的，并且头一直是望着天走的，时不时还遇到下阶梯的旅客”。

第二类主要集中在光明顶周围，因为天气等原因没有能观赏到预期的云海、日出、日落等景观而产生的遗憾之感。如：“第二天凌晨四点半左右，我们到了猴子观海看日出，很遗憾，日出还是没看到，不过看看早上的山峰也是不错的”“去了离酒店很近的猴子观海台，本是可以望见美丽的群山与云海，却被大雾遮得严严实实，雨后溪水哗哗声像正在下雨一样，很遗憾由于大雾没有看到日出”“4点钟出发到光明顶看日出，可惜当天云层太厚，遮住了太阳”。消极节点总体数量较少，零散地分布在节点集中区地外围，大多因地势险峻给游客带来恐惧、惊险等感受。

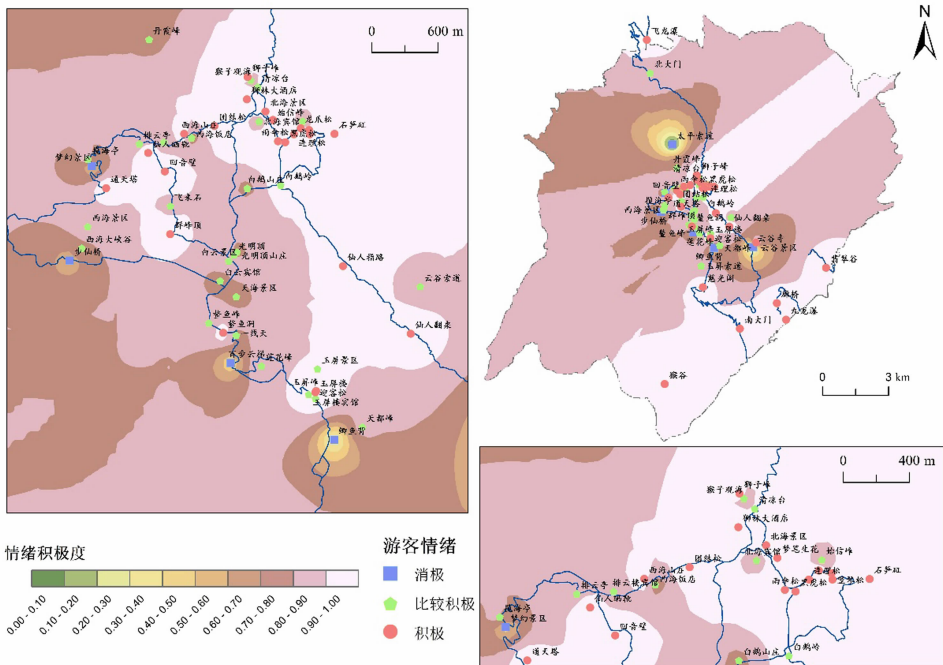


图5 情绪积极度空间分布

从插值所得面状数据来看, 黄山景区内大部分地区的情绪积极度都在 0.7 之上, 景区核心区大于边缘区, 东南地区大于西北地区, 太平索道和云谷寺周围是消极情绪相对集中的地方, 主要是由于部分游客恐高和云谷寺收费存在地域差别造成的游客消极情绪产生。“不公平啊, 我们是从山西过来的, 与河北相邻, 为什么他们能享受这个优惠, 而我们没有呢——气愤!(云谷寺)”“到了站也是排队等缆车, 太平索道一共两部大容量缆车, 往返各一部, 每次要排一个多小时的队伍才能坐上缆车。北坡下山的索道非常陡峭, 每过一个基站就是一段急速下降, 引起车厢里一阵阵大呼小叫”。

为了进一步了解节点情绪的空间分布特征及节点之间的位置关系是否会对情绪结果产生影响, 本研究参考梁嘉祺^{[19]154}等的做法, 通过 GeoDa 构建了 Queen 邻接空间权重矩阵, 并将积极情绪词频和消极情绪词频导入软件中分别做单变量空间自相关分析, 所得的全局莫兰指数分别为 -0.087、-0.036。两者都为负值, 且均接近于 0, 说明两种极性情绪的空间分布总体上接近随机状态, 即游客两种极性情绪的产生均不会受到其他节点位置的影响。从具体数值上看, 积极情绪的莫兰指数要小于消极情绪, 表明积极情绪在空间上的异质性分布更为明显, 即相比较而言, 游客积极情绪的产生更不易受到节点相互间位置的影响。

(三) 情绪积极度和表达度四象限分析

1. 节点情绪分类

在剔除西海大峡谷和光明顶两个“异值”(单独作为一类, 在下文进行分析)后, 分别以节点情绪表达度、节点情绪积极度作为横纵坐标, 以 (0.09, 0.6) 为原点绘制情绪表达度和积极度四象限散点图(见图 6)。其中, 落在第二象限上的节点数量最多, 共计 43 个, 占总数的 65.15%, 其对

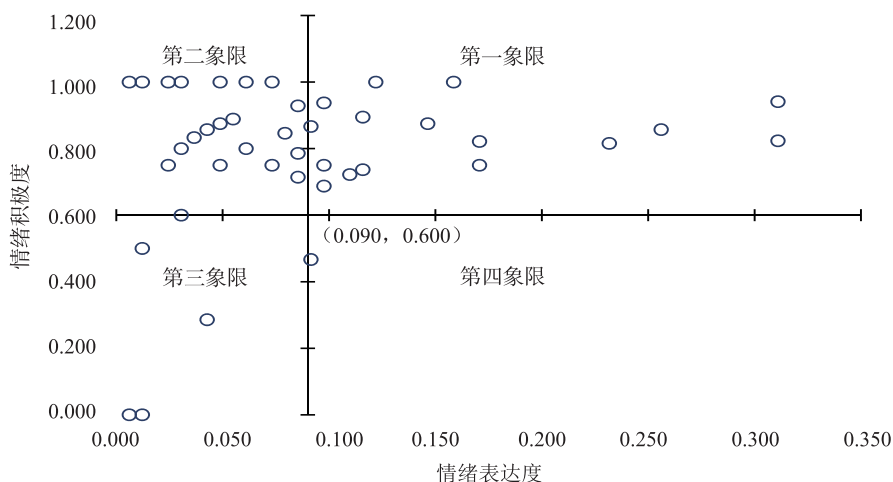


图 6 情绪表达度和积极度四象限散点图

应的情绪属性为低表达度—高积极度（见表6），这说明大部分节点虽然可以给游客带来积极的情绪体验，但这种情感体验的强度不够，很容易在“景象—视象—意象—语象”的逐级转化的过程中被过滤，不足以形成游客的情绪“爆点”。

表6 交叉分类节点统计

象限	属性	数量	节点名称
I	高情绪表达度 高情绪积极度	18	玉屏索道 西海饭店 猴子观海 清凉台 白云宾馆 丹霞峰 玉屏景区 北海景区 莲花峰 黑虎松 始 信峰 排云亭 天都峰 飞来石 云谷索道 迎客松 光明顶 西海大峡谷
II	低情绪表达度 高情绪积极度	43	通天塔 仙人翻桌 廊桥 仙人指路 龙爪松 竖琴松 团结松 飞龙瀑 群峰顶 玉屏楼 仙人晒靴 南大门 梦笔生花 回音壁 慈光阁 连理松 雨伞松 天海景区 石笋峰 鳌鱼洞 翡翠谷 云谷景区 排云楼宾馆 探海松 玉屏峰 猴谷 北大门 揽海亭 一线天 玉屏楼宾馆 狮子峰 西海山庄 白云景区 鳌鱼峰 石笋砦 白鹅山庄 九龙瀑狮 林大酒店 西海景区 光明顶 山庄 北海宾馆 白鹅岭
III	低情绪表达度 低情绪积极度	4	云谷寺 太平索道 梦幻景区 步仙桥 鲫鱼背
IV	高情绪表达度 低情绪积极度	1	百步云梯

资料来源：作者利用 ROST 统计及 Excel 计算输出。

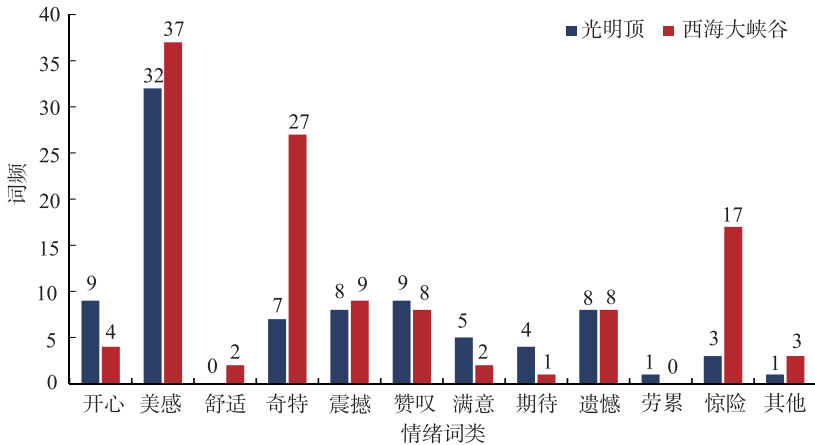


图7 光明顶、西海大峡谷情绪内容

情绪表达度越高说明该节点能够给游客产生的印象体验和情绪刺激就越多。这里选取表达度位居前二的光明顶和西海大峡谷两个节点, 结合研究方法中节点情绪具象化情绪词编码结果, 对节点的具体情绪内容(见图7)进行分析。在这两个节点上, 游客的主要情绪表达均是“美感”, 说明景色是游客情绪评价中的重要内容。相比光明顶, 西海大峡谷刺激游客产生了更多“奇特”“惊险”的情绪, 在一定程度上说明节点上的某些“异质化”的特点, 可能会为游客带来更多的情绪体验。

2. 可视化描述性分析

利用 ArcGIS 将不同情绪属类的节点进行空间可视化表达(见图8)。其中, 高情绪表达度—高情绪积极度的节点, 即能够给游客留下深刻印象并带来积极情绪体验的节点相对集中分布在天都峰—揽海亭—始信峰沿线的游步道两侧。这一类节点主要具备如下特征。

(1) 基于文学影视作品、神话传说等文化属性的加持, 节点本身已经具有较高的知名度, 对游客来说形成了更多的记忆点

这类节点如光明顶、飞来石、猴子观海等。“‘飞来石’, 作为黄山第一奇石, 赢得了古今中外游人的极高赞誉。1987年入央视36集电视剧《红楼梦》片头主题曲镜头, 名扬海内外。我也是从那时起就知道了这一‘不知名’的奇石, 犹如仙境。如今近在眼前, 一晃三十年, 已到中年, 生命之短暂, 大自然的伟大”“光明顶乍一听, 顿然有点金庸似武侠的风采, 这时不知不觉海拔已经1800米, 这里也是黄山观日出的绝好地点之一。”“猴子为何要望太平呢? 这里面有个故事……”

(2) 节点本身是景区内其他特色景色的载体或绝佳观赏点

基于这种与旅游地吸引物的联系, 节点能够获得更多的知名度。如丹霞峰、排云亭、清凉台等作为云海、日出、晚霞等景观的绝佳观赏地点, 在游客之间形成了一种口碑传播。“从西海大峡谷乘地轨索道上, 爬山到光明顶还不到4点钟, 在光明顶附近等到快6点钟, 拍下这组落日的照片。由于前面的云海, 落日好像海上升起的朝阳”“如果住排云亭附近, 那么可以选择丹霞峰、排云亭观看日出和日落。如果是住在天海则可以选择在光明顶或鳌鱼峰观看日落和日出。当然这几个地方也是观看云海、晚霞和奇峰幽谷的最佳地点。”“看云海最佳地点: 玉屏楼观前海、清凉台观后海、白鹅岭观东海、排云亭看西海、光明顶看天海。”

对于低—高类节点来说, 虽然其能够给游客留下较高的积极情绪体验, 但总体印象不深刻。这一类节点在景区内的数量最多, 主要分布路线与高一—高类相似。低—低节点和高—低节点数量较少, 只零散地分布在景区核心区以外的边缘地区。

分别利用 Excel、GeoDa 计算情绪表达度和情绪积极度的数值相关系数以及空间自相关莫兰指数，所得结果为 -0.06 、 -0.012 。从相关系数上看，情绪表达度与情绪积极度之间是相互独立的，游客的情绪表达度与情绪积极度之间不会相互影响。从莫兰指数上看，游客的情绪表达度与情绪积极度的空间自相关性较弱，即节点上的游客情绪体验特征不会受到周围节点情绪体验的影响。

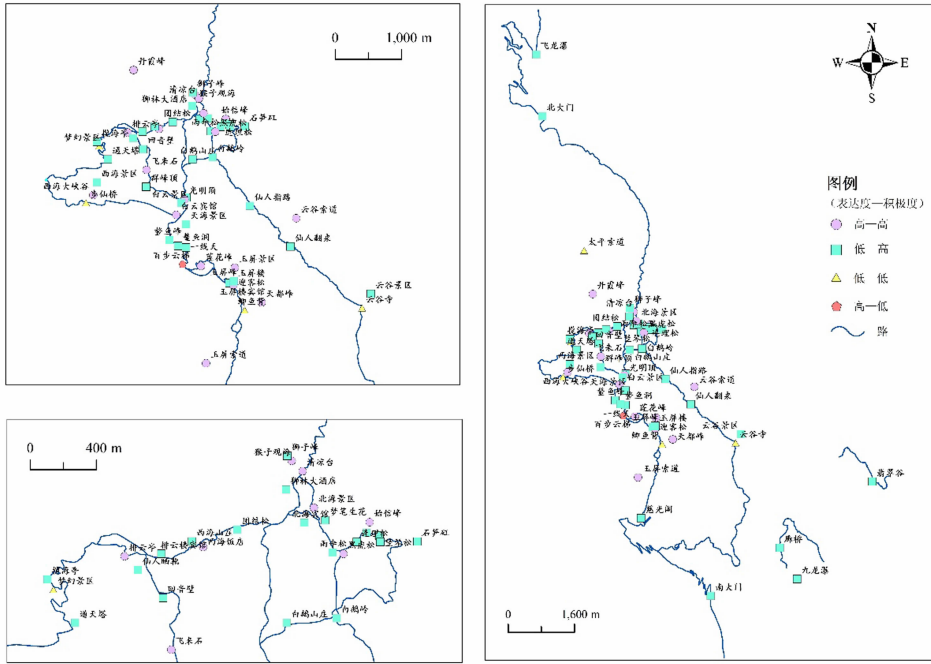


图8 情绪表达度和积极度交叉属性空间分布

三、结论与讨论

(一) 结论

本文基于游记文本，通过情绪词频统计和构建积极度和表达度情绪指数，分析黄山景区游客情绪的空间特征，主要结论如下。

1. 情绪表达度分析

整体而言，黄山景区内高情绪表达度的节点数量较少，66个研究节点中仅有6个为高表达度，占比不到10%。高情绪表达度的节点能够给游客带来更强的情绪印象，留下更深刻的旅游体验感。因此，黄山风景区在节点形象的定位和宣传方面还需调整，即需要基于节点找到能够引发游客共情的异质面做更深层次的艺术加工。

2. 情绪积极度分析

黄山景区内游客总体呈现积极情绪。从线路上来看,“天都峰—揽海亭—始信峰”一线的高情绪积极度节点数量分布最多。具体而言,“奇松”是最为典型的产生高情绪积极度的景观类型。消极情绪的产生,大多与地势险要造成的身体疲劳和心理恐惧以及因天气原因未能如愿欣赏到预期景观有关。

3. 情绪积极度表达度综合分析

总体上,在具备文化特质的自然景观节点上,游客情绪更易呈现出高表达度—高积极度的特征。其原因是带有影视文化等人文要素属性的节点,知名度相对较高,在旅游活动之前已经在游客脑海中生成了预设形象,在具体的旅游情境中更易产生人景交互的情绪活动。基于此,在进行景区的开发和宣传时,应当重视对于节点文化要素的挖掘;游客情绪的空间自相关性较弱,单个节点的游客情绪表达度和积极度大小受其周围节点的影响的程度较小。整体而言,游客情绪在空间上的分布呈现出一定的随机性,不具有明显的集聚特征。

(二) 讨论

利用游记文本进行游客情绪研究过程中,游记文本数据的选择和处理中所遇问题和解决方案如下。

1. 游记数据优势和抽样方法

游记文本出于游记作者自觉、自由撰写,不会受到景区管理者、研究者等人的影响,最大限度地保留了游客真实的心理状态。其研究价值(代表性和典型性)已经被保继刚等多位领域内学者认可,并作为研究数据源使用;游记样本(内容)的选择,采取无差别爬取。游记作者涵盖了多种出游方式(自驾游、跟团游)、多种游伴数量(个人游、家庭游、朋友同事游)、多种出游时间(一日游、二日游、三日游)、多种人均日消费(350~3500元)等实际旅游情境。通过详细阅读,发现所选旅游情境分类已经饱和——可以代表各种类型游客的情绪状态;样本规模(数量),我们利用舍费尔抽样公式进行了抽样合理性的验证。以2017年游记数量为样本量(抽样规模),2017年黄山景区游客量(统计数据)为游客总量(母本大小),验证结果表明本文的抽样在90%的置信水平下符合抽样要求,可以代表游客群体的总特征,对于景区规划和游客路线选择具有指导意义。

2. 游客情绪测定方法

通过爬虫软件爬取的原始游记文本存在重复、广告类假游记和无关游记等现象,会影响游记词频对游客真实情绪的代表性。本研究采用人工精读的方式对游记文本进行筛选,以保证研究结果的合理性和准确性;同时采用人

工方式进行情绪词的搜索、定位以及识别,在一定程度上规避了因语境造成的语义误差,提高了情绪识别的准确性,但也影响了情绪识别的效率。未来可以考虑将人工与软件结合起来,在保证准确度的同时,提高识别的效率。考虑到界定及识别的难度,本研究在情感属类划分时只选取了积极和消极两个维度,但在现实情境中,游客的情感可能会有更多的层次。因此,未来需要对情感属类做进一步的细分,以便更为真实的反映游客情绪状态。本研究在对情绪表达度、积极度以及最后的四象限分类过程中,都依据的是样本本身呈现的数据聚集特征选取阈值,从而在最大程度上保留了数据样本的原始状态,避免了人为确定而造成的主观偏差^{[19]156},具有一定的推广性。相比直接根据每个维度情绪词词频数量来确定情绪极性,本研究提出的情绪积极度和情绪表达度两个指标保留了节点间的细微差别特征,在情绪判别上更为客观。

[参考文献]

- [1] 王福顺. 情绪心理学 [M]. 北京: 人民卫生出版社, 2018.
- [2] 廖启鹏, 刘超, 李维. 游客记忆视角的景观关注度研究: 以黄山风景区为例 [J]. 人文地理, 2019, 34 (6).
- [3] 易婷婷, 梁明珠. 旅游博客情绪展示对旅游意向的影响研究: 基于情绪感染理论的实证分析 [J]. 旅游研究, 2019, 11 (1).
- [4] 李晓琴. 旅游体验影响因素与动态模型的建立 [J]. 桂林旅游高等专科学校学报, 2006 (5).
- [5] 何琼峰. 基于扎根理论的文化遗产景区游客满意度影响因素研究: 以大众点评网北京 5A 景区的游客评论为例 [J]. 经济地理, 2014, 34 (1): 139, 168 - 173.
- [6] 张丽. 旅游黄金周景区拥挤度与游客体验影响因素研究 [D]. 南宁: 广西师范大学, 2008.
- [7] 李恒, 崔婷婷. 安徽省旅游景区游客情绪体验影响因素研究 [J]. 中国商论, 2018 (33): 64 - 66.
- [8] 崔婷婷. 旅游资源独特性和服务质量对游客公民行为的影响 [D]. 马鞍山: 安徽工业大学, 2019.
- [9] YÜKSEL A, YÜKSEL F. Shopping risk perceptions: effects on tourists emotions, satisfaction and expressed loyalty intentions [J]. *Tourism Management*, 2007, 28 (3): 703 - 713.
- [10] 张筱璇, 高晓琪. 旅游幸福感影响因素的元分析 [J]. 区域治理, 2019 (49): 232 - 235.
- [11] XU Z, ZHANG H, ZHANG C, et al. Exploring the role of emotion in the relationship between museum image and tourists' behavioral intention: the case of three museums in Xi'an [J]. *Sustainability*, 2019, 11 (3): 1 - 20.
- [12] YAN Q, ZHOU S, WU S. The influences of tourists emotions on the selection of electronic word of mouth platforms [J]. *Tourism Management*, 2018, 66: 348 - 363.
- [13] KYUNG L H, SEOB K D, YONG K K, et al. Effects of planned versus unplanned purchase on chinese tourists' emotions and attitudes toward fashion products-focused on the mediating role of shopping emotion [J]. *Journal of the Korean Society of Clothing and Textiles*, 2015, 39 (4):

625-639.

- [14] 徐曼. 基于游客情绪的旅游目的地形象对行为意图的影响 [D]. 西安: 陕西师范大学, 2018.
- [15] 刘欢, 岳楠, 白长虹. 红色旅游情境下情绪唤起对游客认知的影响 [J]. 社会科学家, 2018 (3): 84-90.
- [16] 王芳, 谢朝武, 吴必虎. 景观空间尺度对游客满意度的影响及情绪体验的中介作用研究 [J]. 人文地理, 2018, 33 (6).
- [17] 余升翔, 费勇安, 田云章, 等. 旅游消费者情绪量表的维度及检验 [J]. 统计与决策, 2019, 35 (5): 36-40.
- [18] LI S, SCOTT N, WALTERS G. Current and potential methods for measuring emotion in tourism experiences: a review [J]. Current Issues in Tourism, 2015, 18 (9): 805-827.
- [19] 梁嘉祺, 姜珊, 陶犁. 基于网络游记语义分析和 GIS 可视化的游客时空行为与情绪关系实证研究: 以北京市为例 [J]. 人文地理, 2020, 35 (2).
- [20] 于静. 基于微博大数据的游客情感及时空变化研究 [D]. 西安: 陕西师范大学, 2015.
- [21] 方梦静, 郑钰旦, 夏兆焯, 等. 基于微博大数据的游客情感时空变化特征: 以杭州西溪国家湿地公园为例 [J]. 西南大学学报 (自然科学版), 2020, 42 (3): 156-164.
- [22] 原欣伟, 窦天苗, 李延, 等. 在线用户社区成员持续参与意愿的影响因素研究: 基于“认知—情感—意动”理论视角 [J]. 现代情报, 2018, 38 (5): 45-52.
- [23] 缪章伟, 陆净岚. 酒店顾客满意度评价体系研究: 基于 tripadvisor.cn 的杭州高星级酒店评论数据 [J]. 旅游研究, 2021, 13 (1): 73-84.
- [24] 陈航, 王跃伟. 基于旅游者情感的目的地品牌评价研究: 以互联网旅游日记为例 [J]. 人文地理, 2018, 33 (2): 154-160.
- [25] 胡传东, 李露苗, 罗尚焜. 基于网络游记内容分析的风景区骑行体验研究: 以 318 国道川藏线为例 [J]. 旅游学刊, 2015, 30 (11): 99-110.
- [26] 刘超, 胡梦晴, 林文敏, 等. 山岳型景区旅游形象感知研究: 基于 2014~2016 年黄山网络游记分析 [J]. 山地学报, 2017, 35 (4): 566-571.
- [27] 王硕, 曾克峰, 童洁, 等. 黄金周风景名胜区旅游客流量与网络关注度相关性分析: 以庐山、华山、八达岭长城风景名胜区为例 [J]. 经济地理, 2013, 33 (11): 182-186.
- [28] 王君怡, 吴晋峰, 王阿敏. 旅游目的地形象认知过程: 基于扎根理论的探索性研究 [J]. 人文地理, 2018, 33 (6): 152-160.
- [29] 刘逸, 保继刚, 朱毅玲. 基于大数据的旅游目的地情感评价方法探究 [J]. 地理研究, 2017, 36 (6): 1091-1105.
- [30] 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.
- [31] 余向洋, 张圆刚, 朱国兴, 等. 基于熵测度方法的旅游季节性研究 [J]. 地理科学进展, 2019, 38 (8): 1240-1247.
- [32] 刘超. 调查评价法在非景区型森林景观价值评估中的应用 [J]. 生态经济, 2012 (1): 163-165.
- [33] 刘琦, 韩思雨. 基于扎根理论的征地管理体系构建 [J]. 中国房地产, 2020 (33): 39-46.
- [34] 余玉婷, 王萍, 黄镛玉. 新媒体下的自我呈现: 基于亲密度的朋友圈印象管理模式 [J]. 教育现代化, 2019, 6 (64): 236-237.

- [35] 史雅楠. 社交时代自拍叙事中的理性追寻 [D]. 石家庄: 河北经贸大学, 2020.
- [36] 李海涛, 邵泽东. 空间插值分析算法综述 [J]. 计算机系统应用, 2019, 28 (7): 1-8.
- [37] 李小文, 曹春香, 常超一. 地理学第一定律与时空邻近度的提出 [J]. 自然杂志, 2007 (2): 69-71.
- [38] 孙一可, 宫辉力, 陈蓓蓓, 等. 综合莫兰指数和交叉小波的不均匀沉降量化分析 [J]. 国土资源遥感, 2020, 32 (2): 186-195.

Spatial Characteristics of tourists' Emotion in Huangshan Scenic Spot Based on Travel Notes Semantics

CAI Weina¹, LIU Chao^{1,2}, XU Letian¹

(1. School of Geography and Information Engineering, China University of Geosciences, Wuhan 430078, Hubei, China;

2. National Park Laboratory, China University of Geosciences, Wuhan 430074, Hubei, China)

Abstract: Nowadays, travel notes have been an important media for tourists to express the emotion they perceived from their travel experiences. Based on the analysis of emotional spatial characteristics of existing travel notes, a guideline can be provided both in the planning of travel routes for tourists and image positioning for scenic spots. After the mining of emotional vocabularies and semantic analysis of travel notes for Huangshan scenic spots, two indexes are utilized to evaluate emotional expression level and emotional positiveness respectively. Also, the inverse distance weight interpolation and space self-related Moran index are used for the analysis of emotional spatial characteristics. The result shows that the emotional expression level of Huangshan scenic spot is generally low, and most nodes with relative high-level are centralized distributed in a linear concentration. While the emotional positivity is generally high, and the positiveness level on the mountain is significantly higher than that on the way. Natural landscape nodes with cultural background such as Flying-Over Rock, Monkey Watching the Sea, Shixin Peak and so on are more likely to show high level in emotional expression and emotional positiveness. The Moran's I of emotional expression and positiveness are both close to 0, which shows a low spatial autocorrelation. Namely, the tourists' emotion of the nodes will not be affected by the surrounding nodes.

Keywords: tourist emotion, emotional expression, emotional positivity, spatial characteristics, Huangshan