

基于地理标记照片的北京市入境旅游流空间特征

秦 静¹, 李郎平², 唐鸣镝^{1,3}, 孙 岩¹, 宋昕芮¹

(1. 北京第二外国语学院旅游管理学院, 北京 100024; 2. 中国科学院地理科学与资源研究所, 资源与环境信息系统国家重点实验室, 北京 100101; 3. 北京林业大学园林学院, 北京 100083)

摘要: 城市作为现代旅游目的地的一种重要类型, 在集聚旅游发展要素、塑造城市休闲空间、提升城市品质等方面具有突出作用。城市旅游流是城市旅游要素关联、空间组织的路径, 对于城市交通组织方式、公共服务设施建设、休闲旅游产业布局等方面影响越来越大。本文立足城市内部微观尺度, 以入境旅游流空间特征为研究对象, 利用数据挖掘技术, 通过 Flickr API 接口获取了 2004-2015 年北京市范围内全部带有地理信息的照片及其属性信息, 通过构建的数据清洗方法控制初始数据质量, 建立基础数据库, 使用密度聚类(P-DBSCAN)、马尔科夫随机链等方法探索入境游客热点区域空间分布格局、流动空间特征。结果显示: ①北京市历史城区(2环之内区域)、中心城区(2环至5环之间区域)以及郊区(5环之外区域)的旅游热点区域(AOI)呈现出不同的空间特征, 其中历史城区 AOI 呈现出与景区、休闲娱乐、餐饮购物等相关产业空间重叠的趋势, 各热点区域边界模糊, 复合连绵成片。中心城区 AOI 集聚与城市功能区密切相关, 具有明确的城市功能取向。郊区 AOI 基本与旅游景点一致, 没有向景区外部延伸; ②传统历史文化类 AOI 访问热度占有绝对优势, 商业购物休闲 AOI 及现代城市功能 AOI 的影响力在逐渐增强; ③强入境旅游流主要集中在历史城区内部 AOI 之间, 以及中心城区的西北部与东部区域内部 AOI 之间, 郊区 AOI 之间没有较强的旅游流, 北京市入境旅游流整体呈现以历史城区为中心, 放射性分布的框架; ④各区域 AOI 之间均有较次一级旅游流联系, 北京市入境旅游流形成了初步的网络结构。

关键词: 地理标记照片; 旅游流; 旅游热点区域; 空间特征; 北京市

DOI: 10.11821/dlxb201808012

1 引言

旅游流的核心是旅游者的空间移动及其引发的各类经济社会发展效应, 各类旅游流因连接客源地与旅游地、旅游地与旅游地而成为旅游系统的神经中枢或纽带^[1]。入境旅游流是旅游流的重要类型, 研究入境游客在旅游热点城市的流动规律, 可反映入境游客市场的动态行为规律及趋势, 对于城市休闲游憩场点设置、公共服务设施配套、旅游业态开发等方面均具有重要意义^[2]。有关入境旅游流的研究依据空间尺度的不同又可以分为城市内部^[3-5]与不同城市之间^[6-11]游客流动规律。随着城市功能转型与旅游发展, 游客在城市内部流动的空间特征对城市交通格局、基础设施建设、旅游开发、城市环境等的影响日

收稿日期: 2017-07-27; 修订日期: 2018-04-11

基金项目: 北京市社会科学基金研究基地项目(17JDGLB004); 国家自然科学基金青年项目(71502005); 北京第二外国语学院种子计划孵化项目 [Foundation: Philosophy and Social Science Foundation of Beijing, No.17JDGLB004; National Natural Science Foundation of China, No.71502005; Seed Incubation Project of Beijing International Studies University]

作者简介: 秦静(1986-), 女, 山东济南人, 博士, 讲师, 研究方向为地理信息技术与大数据在城市与旅游中的应用。

E-mail: qjing1986@163.com

益凸显,城市内部游客路径轨迹空间的研究逐渐受到国内外学者的重视。

最初,城市内部旅游流空间特征的研究仅使用区域内的旅游者接待量,或直接采用旅游景区各入口处统计的人数,这是一种基于结果数据的旅游流分析^[12],此类数据揭示的是旅游者在一定区域内的空间分布规律,而非旅游者的空间移动特征,体现不出“流动性”和相关地域单元之间的旅游流联系。随后,旅游流空间特征的研究大部分是通过调查问卷的方法研究游客在城市内部各区域及景区(点)之间的流动与扩散路径^[13-14]。这种方式一定程度上可以反映游客在城市内部的动态分布情况,然而通过调查问卷采集数据往往需要花费很长时间并且只能获得有限的样本数量^[3,15],而使用有限随机的样本进行数据分析,难以准确全面的反应游客的行为特征。另外,从空间尺度上,通过统计数据及问卷调查的方式获取的数据基本是以行政区域或者城市内独立的景区(点)为研究范围,而游客真实的兴趣点往往并不和景区(点)范围完全重合,因此采用这两种方式无法真实反映游客在城市内部的空间分布特征。如何突破传统采集数据途径(问卷调查、统计年鉴等)和空间分析方法,科学、准确地获取旅游流数据,并运用地理信息技术进行空间表达,是城市内部游客空间行为研究的关键问题。

伴随着第3次因特网浪潮的出现、Web 2.0概念的提出以及相应的技术体系为各种信息应用带来的全新技术支撑,形成了新地理信息时代^[16]。通讯方式的改变、迅速发展的互联网社区促进了Flickr、Panoramio等新兴网络共享平台出现。这些社会网络平台中拥有数以百万计用户共享的地理标记照片,通过网络平台提供的接口,可以下载照片及包含在照片EXIF中的拍摄时间和经纬度等属性信息,利用这些信息进一步通过数据挖掘算法可以获取游客的移动轨迹,为旅游流的研究提供全面、海量、准确的一手数据^[17-19]。目前国内相关研究中,基于地理标记照片的旅游流空间特征的相关研究仍处于起步阶段,由于新的空间分析技术应用滞后,研究的尺度依然停留在区县等行政单元^[4-6,20]。以区县为研究尺度在一定程度上可以反映城市内部旅游流空间分布规律及游客流动规律,但不能精确的反映出游客真正的兴趣点及兴趣点的类型,进而无法精确的分析游客行为特征。国外相关研究中,基于地理标记照片的应用大部分集中在依据游客游览轨迹进行旅游线路推荐方面^[21-24],部分涉及旅游流空间特征的研究,仅以城市的某个区域或者单一景区内部旅游流为研究对象探讨游客转移规律^[3,25],而没有对城市市域范围内旅游流整体空间结构进行研究。因此,准确获取城市内部旅游热点区域(Area of Interest, AOI),以旅游AOI为研究单元进行游客流动性空间特征研究是旅游流研究需要进一步探索的问题。

综上,本文以城市内部旅游AOI为研究单元,以入境旅游流空间特征为研究对象,选取北京市为研究案例区,首先,利用数据挖掘技术,通过Flickr API接口获取2004-2015年北京市范围内全部带有地理信息的照片及其属性信息;其次,通过构建的数据清洗方法去除原始数据中错误信息,控制初始数据质量,建立基础数据库;然后,使用密度聚类(P-DBSCAN)方法获取AOI,分析北京市热点区域分布特征;最后,使用马尔科夫随机链,获取游客在不同热点间的转移概率,进而分析入境游客流动性空间特征,为城市旅游空间布局优化及目的地城市精准营销提供科学指导。

2 数据来源与研究方法

2.1 研究区域

北京市是一个具有3000年建城史和800年建都史的历史名城,拥有世界上最大的皇宫紫禁城、祭天神庙天坛、皇家花园北海、皇家园林颐和园,还有八达岭、慕田峪、司

马台长城等名胜古迹，是中国的政治、经济、文化、教育中心城市，同时也是重要的入境口岸城市。北京市的入境旅游接待量和外汇收入长期以来一直位居全国前列，对外交通网络非常发达，是国内、国际航空通达性最高的城市之一，已与全世界 70 多个国家的 190 多个城市开通了国际航线，是中国国际航空网络的核心节点。因此，无论是从入境旅游接待规模、外汇收入，还是航空交通条件来看，北京市入境旅游是中外互动交流的重要渠道和窗口，肩负着“世界认知中国，中国向世界展示自我”的双向功用，其发展对于北京市乃至中国都有着重要的意义。

2.2 数据来源及处理

本文通过 www.flickr.com/services/api 接口抓取 2004 年 1 月-2015 年 12 月北京市范围内 (115.7°E~117.4°E、39.4°N~41.6°N) 带地理标记的 flickr 图片及属性数据，共包括以下字段：用户 ID，照片 ID，用户所属区域，照片拍摄时间，照片上传时间，照片拍摄地点 (经纬，纬度)，照片标题，照片描述，照片评论。该数据集覆盖了从 2004-2015 年共 12 年用户共享数据，因研究对象为北京市入境旅游流，所以首先从数据集中去除用户所属区域属性为中国城市及空白的记录，然后合并各个年份数据，共 131787 条记录和 3515 个游客 (表 1)。将所有记录的地理坐标转换成通用墨卡托投影下的平面坐标，以便进行随后的空间计算。

表 1 各数据处理步骤后数据集的记录数、游客数和游览路径数
Tab. 1 Numbers of records, users and trajectories of dataset after each data processing procedure

数据处理步骤	数量		
	记录(条)	游客(人)	游览路径
原始数据集	131787	3515	N.A.
删除上传时间不晚于拍摄时间的记录	128835	3382	N.A.
删除具有相同游客、坐标和拍摄时间的“重复记录”	126166	3382	N.A.
初始质量控制	删除与那些具有相同游客和拍摄时间但不同坐标的“错误记录”	118180	3353
	具有相同游客和拍摄日期的记录		N.A.
	删除与那些具有相同游客和坐标但不同拍摄时间的“错误记录”	36308	2817
	具有相同游客和拍摄日期的记录		N.A.
	删除位于“热点区域”以外的记录	27473	2477
删除所有在“来回摆动路径”之上的记录	26334	2473	5836

注：“N.A.”代表“不适用 (not available)”。

2.3 初始质量控制

Flickr 共享平台上存在游客分享的数以百万计的大体量数据，这些数据由于设备精度等问题可能存在一些错误，对进行旅游流的分析会带来一定的干扰。为了保证研究结果的精确性，在进行数据分析前，采取了 4 步初始质量控制措施。① 将上传时间不晚于拍摄时间的记录删除。② 具有相同用户 ID、相同拍摄时间和相同坐标的记录被认为是“重复记录”。仅保留一条“重复记录”，其他的都将被删除。③ 对于同一游客，具有相同拍摄时间但不同坐标的记录被认为是“错误记录”。这是因为，同一游客不会在同一时间出现在不同位置。④ 分析发现，对于同一游客，相当多的记录具有相同坐标但却具有不同拍摄时间。这些记录主要是由“定位错误”引起的，而不是由“游客在同一地点拍摄了多张照片”或“游客在一段时间后又精确地回到原地点”等情况所引起的。因此，这些记录也被认为是“错误记录”。考虑到游客游览路径是按天计的，一天之内的一条错误记录即可给该游客在该天内的整条游览路径带来错误。因此，本文将那些与“错误记录”

具有相同游客和相同拍摄日期的记录都删除。数据集在每一步初始质量控制后的记录数与游客数如表1所示。在全部4步初始质量控制措施都实施后, 数据集共有36308条记录和2817个游客(表1)。

2.4 热点区域识别方法

研究游客时空分布特征, 首要的是获取城市内部AOI。国内很多研究在获取城市内部热点区域时采用核密度的方法^[4, 6, 26], 通过核密度方法可以探测游客在城市内部分布的冷热点区域, 但核密度分析结果通过分级设色来表示, 反应的是旅游热点分布的大致趋势, 并没有一个阈值把旅游热点与非热点区域明确区分开来, 因此, 无法进一步研究游客在不同旅游热点之间的转移路径。国外相关研究中, 常使用密度聚类的方法来获取城市内部旅游热点, 密度聚类方法又可依据选取阈值时采用照片密度(如DBSCAN算法)^[15]和游客密度^[3, 27](如P-DBSCAN算法)分为两种。采用照片密度为阈值进行密度聚类是以一定距离内照片的多少来划分热点区域的, 这种方法在反映游客热点时存在一定的缺陷, 一个区域照片的数量往往并不代表真实的游客数据量, 因为同一游客可能在某一区域拍摄多张照片。而以游客数量为阈值进行热点区域与非热点区域划分, 能够反映游客在城市内部的真实分布, 从而获取游客访问较多的热点区域。

本文使用P-DBSCAN算法^[28]来识别旅游AOI。P-DBSCAN算法利用一个邻域半径 θ 和一个游客数量阈值 δ 来识别所谓的“核心记录”。假设 D 是一个地理标记照片集合, 其中照片 p 的经纬度坐标为 (x_p, y_p) 。照片 p 与 q 之间的距离定义为 $Dist(p, q)$ 。则照片 p 的邻域照片集 $N_\theta(p)$ 定义为:

$$N_\theta(p) = \{q \in D, Owner(q) \neq Owner(p) | Dist(p, q) \leq \theta\} \quad (1)$$

式中: θ 为邻域搜索半径; $Owner(q)$ 为照片 q 的用户属性函数(用于表示照片 q 的所属游客), 也即如果照片 q 与照片 p 所属的用户不同, 并且它们之间的距离在 θ 之内, 就称照片 q 为照片 p 的邻域照片。定义 $NumOwner(p)$ 为邻域照片集 $N_\theta(p)$ 的所属用户总数量, δ 为用户数量阈值, 如果 $NumOwner(p)$ 大于 δ , 称照片 p 为“核心记录”。也就是说, 如果一条记录是核心记录, 那么至少有 δ 个除这条记录所对应的游客之外的其他游客在这条记录的半径为 θ 的邻域内拍摄过至少一张照片。直观地, 既然至少有 δ 个游客被吸引到核心记录的 θ 邻域内拍摄照片, 核心记录就应坐落在热点区域以内。换言之, 核心记录的空间分布即标明了热点区域的空间分布。为了使游客数量相对较少的远郊热点区域也能被有效地识别, 本文使用12作为游客数量阈值 δ 的值。领域半径 θ 的值则根据不同的区域的照片空间分布特点来决定。在所有热点区域被识别出来以后, 仅将在热点区域以内的记录用来生成和分析游客流动规律。

2.5 转移概率计算方法

游客游览路径按时间顺序连接了同一游客在同一天内所访问的不同AOI。在分析不同AOI之间的转移之前, 基于生成的初始游客游览路径采取了进一步质量控制处理, 利用专家经验识别出并剔除了那些不合理的“来回摆动路径”。“颐和园—前门/大栅栏—颐和园—前门/大栅栏”就是不合理的“来回摆动路径”的一个典型例子。在删除所有位于“来回摆动路径”之上的记录后, 本文得到一个最终的Flickr数据集(表1), 用来进行旅游流空间特征的分析。

基于至少含有两个不同AOI的游览路径, 可以统计得到任意两个不同AOI之间转移的全部游客的数量, 进一步可计算得出从每一个AOI出发的转移的个数以及到达每一个AOI的转移的个数。假设游客在不同AOI之间的游览路径是一个“稳定马尔科夫链”^[29], 定义 $A = \{a_1, a_2, a_3, \dots, a_n\}$ 为由P-DBSCAN密度聚类方法获取的AOI的集合, 并且游客对

于任意给定时间序列 $t_1 < t_2 < \dots < t_k$ 的转移路径为 $T = (a_i^{t_0}, a_i^{t_1}, \dots, a_i^{t_k})$ ，那么在稳定马尔科夫链的假设下，游客由位置 $a_i^{t_{k-1}}$ 转移到 $a_i^{t_k}$ 的概率与之前经历的路径无关，可以用恒等式 (2) 表示为：

$$P(a_i^{t_k} | a_i^{t_{k-1}}, a_i^{t_{k-2}}, \dots, a_i^{t_0}) = P(a_i^{t_k} | a_i^{t_{k-1}}) \tag{2}$$

那么游客从 AOI a_1 到 AOI a_2 的“转移概率”可以表述为从 AOI a_1 流入到 AOI a_2 的游客数量数比上从 AOI a_1 出发的游客总数量。

3 结果分析

3.1 AOI的空间识别

使用 ArcGIS 平台将全部地理标记图片按照经纬度显示在地图上，结果如图 1 所示。由图 1 可以看出，北京市 2 环之内（以下称历史城区）、2 环~5 环之间（以下称中心城区）与 5 环之外（以下称郊区）游客拍摄的地理标记照片的分布特征不相同，这也反映了各区域旅游发展现状与特征的不同。其中，历史城区照片分布最为密集，中心城区较为分散，郊区分布最为分散。通常情况下使用 P-DBSCAN 算法时，整个研究区域使用相同搜索半径^[3, 30]，然而研究案例区北京市不同区域地理标记照片分布特征的存在较大差异，因此，在提取 AOI 时，对不同区域设置不同的搜索半径，经过多次实验，最终确定的搜索半径分别为：历史城区 300 m，中心城区 600 m，郊区 1000 m。对于游客人数的设定，通过实验，选取游客总人数的 5%，即拍摄照片的游客人数达到 12 人就认为此处为 AOI。将上述两个参数带入 P-DBSCAN 中，获得的聚类结果如图 2 所示。

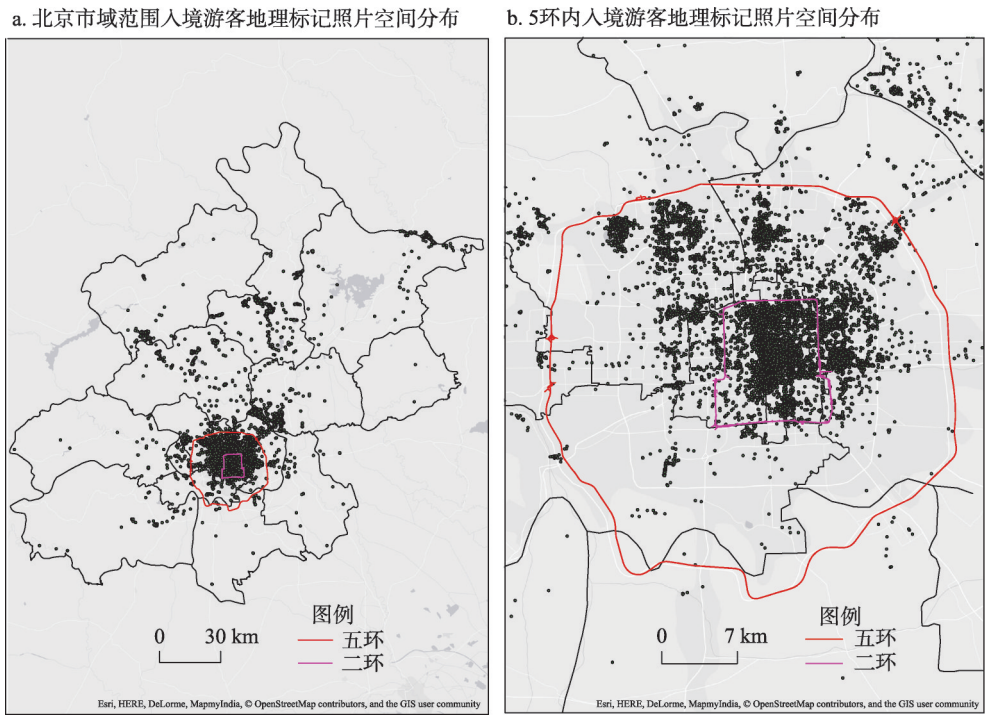


图 1 2004-2015 年北京市游客地理标记照片空间位置
Fig. 1 The locations of inbound geotagged photos in Beijing from 2004 to 2015

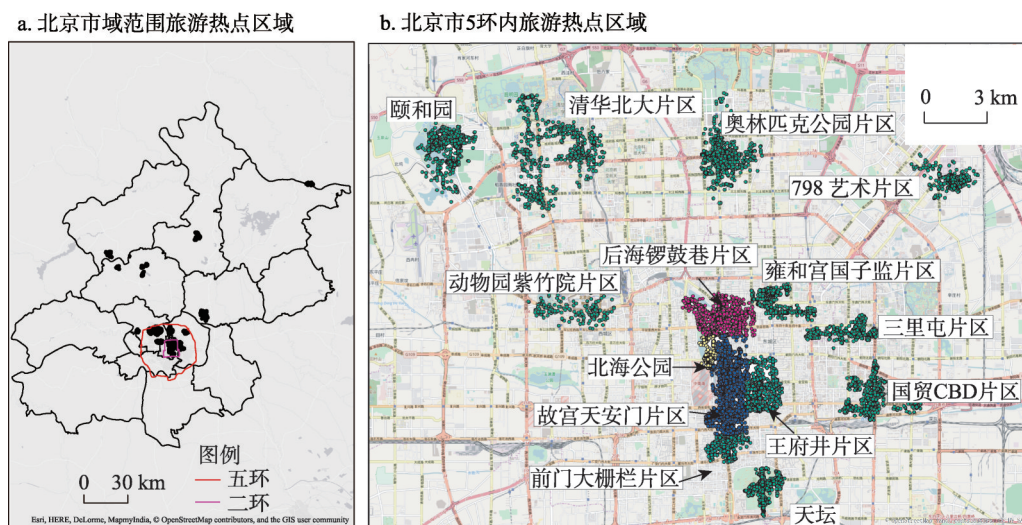


图2 通过P-DBSCAN算法聚类得到的北京市AOI

Fig. 2 Areas of interest (AOI) in Beijing by P-DBSCAN

3.1.1 AOI空间特征及属性特征 由图2所示, 历史城区、中心城区以及郊区的AOI呈现出不同的空间特征。历史城区AOI特点: 集聚成片, 主要集中分布在历史城区中轴部分及东部, 呈现出与景区、休闲娱乐、餐饮购物等相关产业空间重叠的趋势, 非单一功能构成, 边界模糊、复合连绵成片; 中心城区AOI特点: AOI分布与城市功能区密切相关, 聚集成片, 功能性强, 具有明确的城市功能取向; 郊区AOI特点: 基本都与旅游景区相吻合, 边界清晰, 没有向周围延展, 郊区景区(景点)主要以观光活动为主。基本以单一旅游景点为依托, 没有向景区外部延伸。

依据历史城区、中心城区与郊区AOI的上述特点, 对聚类结果进行了进一步的划分, 由图2热点区域聚类结果可以看出, 中心城区与郊区通过聚类分析均很好的划分出了热点区域, 对于历史城区连绵成片的AOI, 本文按景区/景点相互之间距离与类型进一步细分, 共得到21个热点区(表2)。

依据表2热点区域聚类结果, 根据不同热点区域主导功能与性质的不同, 上述AOI大致可分为传统历史文化、商业购物休闲、现代城市功能3种类型。传统历史文化类: 由传统历史景点与历史街区构成, 既包括标志性的历史文化古迹, 也包括传统的居住片区。前者是观光旅游流的主要空间载体, 如故宫天安门、天坛、北海; 后者则随着旅游活动的发展与集聚, 传统街巷的居住功能逐渐被商业休闲功能更迭, 形成具有地方空间形态与生活特色的新兴文化休闲区。比如后海锣鼓巷片区、雍和宫国子监片区。这类AOI数量最多, 共11个, 占热区总量的52.38%。商业购物休闲类: 由传统商业街区构成, 随着单一的商业功能与传统商业文化、休闲体验活动相结合, 逐渐衍生商业休闲街区, 形成与本地生活圈截然不同的地方商业特色, 承担着城市RBD (Recreational Business District) 的职能。如前门大栅栏片区、王府井片区, 因区域尺度与定位呈现出不同的业态与特色。这类AOI共3个, 占热区总量的14%。现代城市功能类: 由商务、教育、体育等主导产业功能构成, 是城市特色产业的集聚区, 与旅游协同共存的方式体现为产城融合, 比如北大清华高教区、国贸CBD商务区、奥林匹克公园、三里屯片区。这类AOI共4个, 占热区总量19%; 此外, 还有一些其他类型的热点区域, 如交通枢纽(首都机场), 城市公园(香山植物园, 动物园、紫竹院片区)也普遍受到入境游客的关注。

表2 北京市旅游热点区域(AOI)聚类结果
Tab. 2 Total areas of interest by P-DBSCAN in Beijing

热点区域名称	
郊区(7个)	八达岭、十三陵、居庸关、慕田峪、首都机场、香山植物园、金山岭 798艺术片区：含798艺术区、751设计园区等 国贸CBD片区：含国贸中心、中央电视台、建外SOHO、购物中心等 三里屯片区：含三里屯、工体、东四十条等
中心城区(7个)	清华北大片区：含清华、北大、中关村、圆明园、五道口等 动物园紫竹院片区：含动物园、紫竹院、国家图书馆等 奥林匹克公园片区：含奥运场馆、奥林匹克森林公园等 颐和园 雍和宫国子监片区：含雍和宫、国子监、簋街等 王府井片区：含王府井商业区 前门大栅栏片区：含前门大街、大栅栏历史街区、鲜鱼口历史街区等
历史城区(7个)	故宫天安门片区：含天安门、故宫、景山等 前海锣鼓巷片区：含什刹海、恭王府、钟鼓楼、锣鼓巷等 北海公园 天坛

综上，北京市历史城区、中心城区与郊区AOI分布呈现不同的空间层次：历史城区典型AOI主要以传统历史文化、商业购物休闲为主。由于酒店、景区等旅游产业要素的分布以及城市空间用地的影响，呈现城市中轴线与东部局部片区集聚的现象；中心城区典型AOI主要以城市功能区为主，也包括少量近郊的传统历史文化资源。其与城市功能结合，呈现商务、商业、教育、体育、观光为主的特色片区；郊区典型AOI主要以传统历史文化为主，以景区景点为主要空间载体，呈现观光游览为主的功能，形成特征单一的旅游区，以长城为代表。近年来随着休闲度假的发展，此类景点也呈现文化观光向深度体验转型的趋势，比如慕田峪国际文化村。

3.1.2 AOI的热度特征 为了进一步探讨入境游客对各热点区域的关注程度，对21个片区进行了热度排序（表3），结合热点区域空间分布，可以得出：传统历史文化类AOI占有绝对优势，与历史文化相关的AOI访问热度在前10中有7个；随着城市化的发展，AOI的类型、规模、数量都有增长，在原有历史文化AOI基础上，商业购物休闲AOI更加活跃，现代城市功能AOI的影响力也逐渐增强，如奥林匹克公园片区与国贸片区热度均在前10位；作为首都的城市中心性空间特征突出，体现为从传统轴线的故宫天安门到现代奥林匹克森林公园的城市轴线的延续。故宫天安门片区访问热度几乎是前5中其余热点访问人数的总和，显示出政治中心的特征；强烈的地方特色是受到游客青睐的主要原因，历史城区高热度AOI集聚体现了北京市独有的人文景观与文化积淀。

3.2 旅游流空间特征分析

不同旅游者所选择的不同旅游线路汇聚成旅游流，而旅游流的不同汇聚，塑造了旅游活动的总体形态，并最终形成了区域间访问的不同格局。在整体旅游流动格局中，流量大的旅游节点之间的旅游流联系，构成了城市主要的旅游流，并决定了旅游流的总体流动构架；小的旅游流接在大的旅游节点之上，决定了一个大的旅游节点向外辐射的能力及旅游流的拓展方向。因此，本文在研究旅游流的空间特征时，首先要获取不同AOI的流量；然后通过控制流量及转移概率，获取主要旅游流形成的总体流动空间构架；最

后获取接续在大旅游节点上的次旅游流的空间特征。

3.2.1 AOI 旅游流量统计 以一天为统计单元, AOI 的出度为该热点区域转向其他热点区域的游客总数量, 相应入度为由其他热点区域转入该热点区域的游客总数量, AOI 流量为该 AOI 出度与入度之和。通过计算不同 AOI 的流量统计如表 4 所示。由表 4 可以看出, 流量最高的 AOI 均在历史城区, 其次为中心城区, 郊区 AOI 则流量较小。作为城市旅游的典型活动空间, 北京市城区在标识吸引物、休闲场所、配套服务设施、公共基础设施等方面完备密集, 特别是历史城区, 更是因为历史与人文的积淀有着压倒性的优势, 成为城市旅游者的首选之地。而郊区的 AOI 由于交通距离、资源影响力、配套设施等原因, 吸引物之间转换需要较长的时间, 且缺少高品质的设施支撑, 因此游客往往当天返回主城区次日再发散式出游, 形成典型的基地型旅游目的地模式, 从而进一步加剧对主城区的依赖、加大主城区的流量。流量大的点, 一般也作为游客转移的中心节点, 承担了游客集散的作用。

3.2.2 北京市入境旅游流空间特征 把每一个旅游者的路径轨迹过程拆分成连续节点间的转移, 通过马尔科夫链将上一个节点向下一个特定节点的转移的可能性计算成转移概率, 从而考察不同节点的向外转移情况, 通过不同 AOI 之间转移概率的大小, 探讨不同强度旅游流空间特征。

(1) 不同热点之间强旅游流空间特征

由于存在 AOI 流量很小, 但转移概率大的情况, 因此本文在统计强旅游流时, 选取流量 (L) 与概率 (P) 两个指标为阈值, 排除流量非常小、不具备统计意义的点, 选取 $L \geq 50$ 、 $P \geq 0.2$ 的数据进行统计。与表 4 中 AOI 游客流量统计相呼应, 表 5 中旅游流转移概率进一步细化了旅游流节点的空间关系, 显示了距离较近的 AOI 之间转移概率比距离远的 AOI 高的一般规律。由表 4 与图 3 可得, 北京市较强的入境旅游流空间特征如下:

① 转移概率较高的区域集中在历史城区, AOI 距离较近、交通便捷, 比较容易安排在一天之内游览, 转移概率普遍较高, 形成相互关联的热点网络。

② 故宫天安门片区是概率最高的 AOI 转移节点, 由于其特殊地位, 集旅游观光、交通集散、城市中心等功能于一身, 使得该区域成为北京市旅游中重要的节点, 政府在此设置的旅游集散中心更是将郊区周边出游的人流吸纳其中, 凸显了中心节点的作用。

③ 郊区 AOI 转移概率普遍偏低, 游客往往会选择一天一个景点, 依托城市中心组织次日出游, 从而进一步加大主城区的游客流量。

表 3 北京市旅游热点区域(AOI)热度排名

Tab. 3 Popularity of areas of interest in Beijing

热度排名	热点区域(AOI)	访问人数
1	故宫天安门片区	1420
2	王府井片区	414
3	天坛	414
4	后海锣鼓巷	412
5	奥林匹克公园	336
6	颐和园	331
7	首都机场	244
8	雍和宫国子监片区	231
9	国贸片区	227
10	前门大栅栏片区	209
11	三里屯片区	188
12	慕田峪	145
13	清华北大圆明园	132
14	北海公园	115
15	798 片区	103
16	八达岭	99
17	紫竹院动物园片区	76
18	居庸关	52
19	十三陵	44
20	金山岭	38
21	香山植物园	13

④ 首都机场与国贸、王府井、清华北大等城市功能区转移概率高，在一定程度上说明游客的住宿选址、旅游偏好。

从整体上看，强入境旅游流主要集中在历史城区内部AOI之间，中心城区的西北部（颐和园、清华北大片区、奥林匹克公园片区）与东部（798片区、三里屯片区、国贸CBD片区）内部AOI之间，郊区景点之间没有较强的旅游流，主城区与其余各个分区之间有较强旅游流联系，但各分区之间没有强旅游流联系，北京市入境旅游流呈现以历史城区为中心，放射性分布的整体框架。

（2）不同热点之间次一级旅游流空间特征

为进一步探讨北京市入境旅游流空间特征，对不同热点之间转移概率与流量阈值进行下放，得到与上述旅游流相比较弱一级的旅游流特征。从图4中可以看出，无论是历史城区，中心城区，还是郊区，同一区域内部各AOI之间均有联系，历史城区各AOI之间联系最为强烈；历史城区，中心城区西北部与东部、郊区AOI之间存在相互联系，旅游流在各区域之间形成了初步的网络结构。

3.2.3 游客路径轨迹模式 根据上述实验得到的不同AOI之间转移的概率，进一步统计游客路径轨迹模式，同样以一天为时间单位，统计一天内游客游览路径中经过的AOI的数量结果如图5所示。由图5可得，随着AOI数量的增加，游客人数呈指数函数趋势降低。绝大部分游客选择一天游览1~3个AOI这种较为闲适的游览方式，只有较少数游客选择一天之内较为集中的游览4~6个AOI。由于4~6个景点的线路游览人数较少，没有统计规律，同时3.2.2节中已经对2个AOI之间转移概率做了详细分析。因

表4 北京市旅游热点区域(AOI)流量统计
Tab. 4 Travel flow of areas of interest in Beijing

编号	名称	入度	出度	流量(人)
19	故宫天安门片区	325	394	719
20	王府井片区	176	148	324
17	后海锣鼓巷	165	118	283
13	天坛	101	116	217
21	前门大栅栏片区	108	81	189
10	奥林匹克公园	87	63	150
16	雍和宫国子监片区	71	78	149
8	颐和园	61	62	123
14	国贸片区	70	49	119
15	三里屯片区	62	44	106
18	北海公园	47	42	89
9	清华北大圆明园	32	35	67
6	首都机场	26	30	56
2	慕田峪	17	36	53
11	798片区	21	30	51
3	八达岭	13	27	40
5	十三陵	14	19	33
12	紫竹院动物园片区	9	12	21
4	居庸关	2	17	19
1	金山岭	4	7	11
7	香山植物园	0	3	3

表5 热点之间转移概率($L \geq 50, P \geq 0.2$)

Tab. 5 Transition probability between different areas of interest ($L \geq 50, P \geq 0.2$)

起点	终点	概率P(%)
王府井片区	故宫天安门片区	0.587838
前门大栅栏片区	故宫天安门片区	0.580247
北海公园	故宫天安门片区	0.5
天坛	故宫天安门片区	0.362069
后海锣鼓巷	故宫天安门片区	0.347458
北海公园	后海锣鼓巷	0.285714
雍和宫国子监片区	后海锣鼓巷	0.282051
慕田峪	故宫天安门片区	0.277778
颐和园	故宫天安门片区	0.258065
798片区	国贸片区	0.233333
清华北大片区	颐和园	0.228571
故宫天安门片区	王府井片区	0.225888
国贸片区	三里屯片区	0.22449
颐和园	奥林匹克公园	0.209677
奥林匹克公园	故宫天安门片区	0.206349
三里屯片区	首都机场	0.204545
国贸片区	故宫天安门片区	0.204082
首都机场	国贸片区	0.2

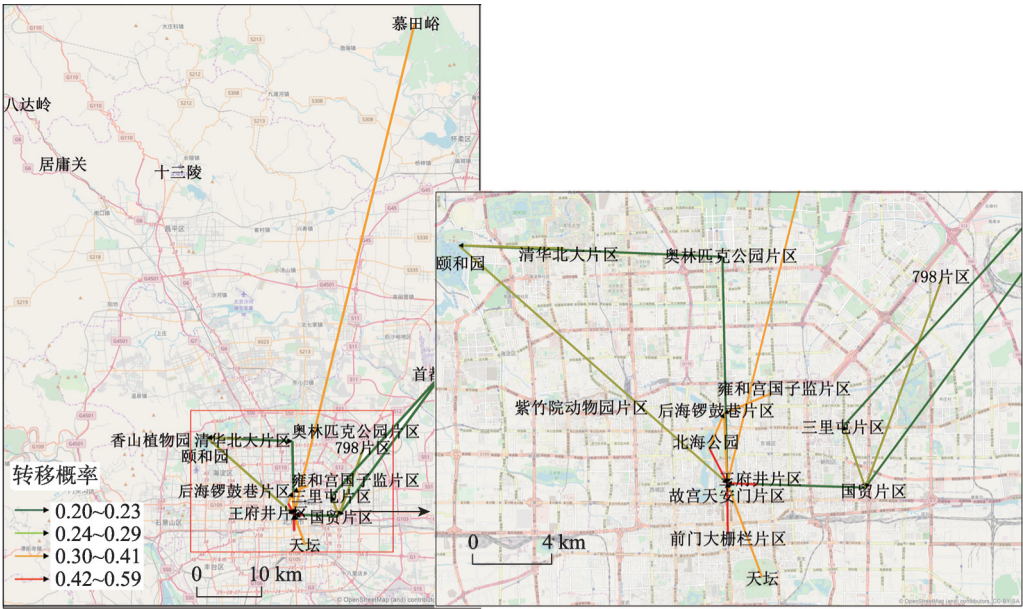


图3 北京市旅游热点之间主要流动路径($L \geq 50, P \geq 0.2$)

Fig. 3 Tourist flow among different values of AOI in Beijing ($L \geq 50, P \geq 0.2$)

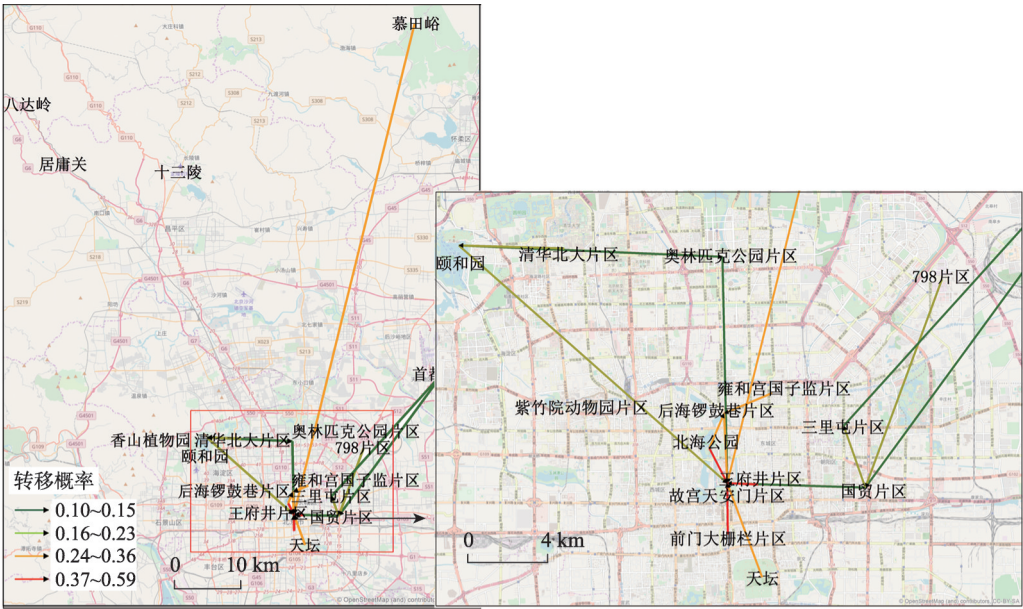


图4 北京市入境旅游流空间特征($L \geq 30, P \geq 0.1$)

Fig. 4 Tourist flow among different values of AOI in Beijing ($L \geq 30, P \geq 0.1$)

此，本节主要对选择单个及3个AOI作为目的地的一日游模式进行分析。

首先，选择单个AOI游览线路占有绝大多数的比例，对其进行统计结果如表6所示。由于AOI在历史城区与郊区的结构属性不同，历史城区内大多体现为组合的热点片区，不是单个景点，而郊区或中心城区边缘的AOI则表现为独立的景区或功能区，因此

将单个 AOI 线路对比之前表 3 的 AOI 游览热度时, 会发现略有不同。清华北大片区作为单一景点线路, 排名较靠前, 前门大栅栏片区及雍和宫国子监片区作为单一景点线路排名较热度靠后, 北海公园作为单一景点线路排名较热度有较大的后移。这种排名变化与 AOI 的区位、规模、性质都有相关。比如位于城区边缘的清华北大片区, 是相对独立、规模较大的现代城市功能区, 具有教育、文化、旅游等多种功能, 其旅游内容适合作为单一景点线路展开。位于城市中心的北海公园等, 由于规模很小、内容单一, 又与故宫、景山等中心景区相邻, 故很少被当做单一景点线路, 而是与周边 AOI 形成组合线路存在。

除了单个景点线路, 3 个以下串联景点的组合线路也占很大比例, 进行分析得出以下热门的游览线路 (表 7)。由表 7 可得, 除慕田峪、奥林匹克公园以外, 其余热门线路均在历史城区的 AOI 之间移动, 即便是慕田峪、奥林匹克公园, 也是与故宫天安门片区直接相连, 进一步凸显了历史城区中心性。

综上, 入境市场一日游的旅游流所形成的旅游空间, 绝大部分旅游活动集中在旧城之中, 特别是中轴线地带以及东城部分片区, 而中心城区与郊区由于景点之间距离的增加, 入境游客很少选择在一日之内同时游览多个 AOI (≥ 3)。

4 结论与讨论

以北京市入境旅游流为研究对象, 基于 Flickr 上游客分享的地理标记照片, 构建数据清洗方法, 去除错误数据, 保证了研究数据的精确性。通过密度聚类的方法, 探讨了入境 AOI 的空间分布特征及属性特征; 然后通过马尔科夫链计算热点区域之间游客的转移概率, 基于此分析了

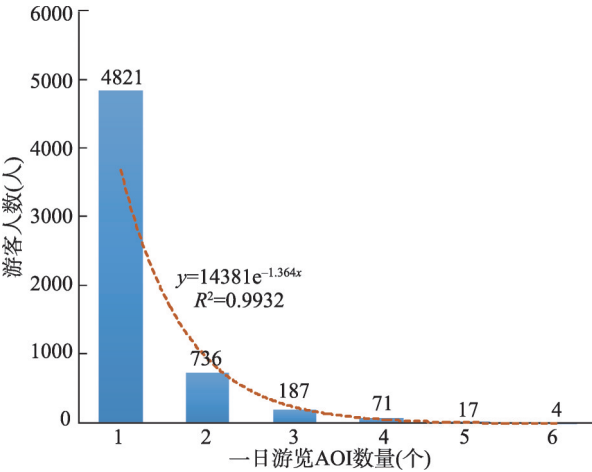


图5 北京市游客游览模式统计
Fig. 5 The statistics for inbound tourist flow patterns in Beijing

表6 北京市单个 AOI 游览模式游客数量统计
Tab. 6 The statistics for inbound tourists who choose one type of AOI in a day

AOI 名称	游客数量统计(人)
故宫天安门片区	1471
后海南锣鼓巷片区	400
王府井片区	312
清华北大圆明园片区	295
奥林匹克公园片区	278
天坛	277
国贸片区	254
首都机场	251
颐和园	251
三里屯片区	209
雍和宫国子监片区	191
前门大栅栏片区	121
慕田峪	109
798 片区	108
紫竹院动物园片区	82
八达岭	65
北海公园	53
金山岭	33
居庸关	33
十三陵	18
香山植物园	10

表7 北京市一日游3个AOI游览模式热门线路
Tab. 7 The hot tourist routes among three types of AOI in a day in Beijing

游客游览路线	经过区域
王府井片区—故宫天安门片区—前门大栅栏片区	历史城区
天坛—前门大栅栏片区—故宫天安门片区	历史城区
故宫天安门片区—天坛—王府井片区	历史城区
故宫天安门片区—后海锣鼓巷—雍和宫国子监	历史城区
天坛—雍和宫国子监—后海锣鼓巷	历史城区
天坛—故宫天安门片区—后海锣鼓巷	历史城区
故宫天安门片区—后海锣鼓巷—天坛	历史城区
故宫天安门片区—北海公园—后海锣鼓巷	历史城区
慕田峪—故宫天安门片区—王府井片区	郊区、历史城区
王府井片区—故宫天安门片区—奥林匹克公园片区	历史城区、中心城区

入境游客流动的空间特征及规律，得到如下结论：

(1) 北京市入境AOI在历史城区、中心城区以及郊区呈现出不同的空间特征。历史城区AOI集聚成片，主要集中分布在中轴部分及东部，呈现出与景区、休闲娱乐、餐饮购物等相关产业空间重叠的趋势，非单一功能构成，边界模糊、相互重叠，复合连绵成片；中心城区AOI分布与城市功能区密切相关，聚集成片，功能性强，具有明确的城市功能取向。郊区AOI基本都与旅游景区相吻合，边界清晰，没有向周围延展，郊区景区（景点）主要以观光活动为主。基本以单一旅游景点为依托，没有向景区外部延伸。

(2) 入境旅游流整体呈现以历史城区为中心，放射性分布的入境旅游流框架：强入境旅游流主要集中在历史城区内部AOI之间，中心城区的西北部（颐和园、清华北大片区、奥林匹克公园片区）与东部区域（798片区、三里屯片区、国贸CBD片区）内的AOI之间，郊区景点之间没有较强的旅游流，北京市各大片区之间强旅游流呈现以历史城区为中心，放射性分布。进一步对较弱一级的旅游流特征进行分析，历史城区、中心城区与郊区内部及之间相互均有联系，旅游流在各区域之间形成了初步的网络结构。

(3) 历史城区是北京市入境旅游空间的核心，历史城区既是旅游吸引物核心，也是旅游集散中心。通过AOI集聚特征分析可以得出由于酒店、景区等旅游产业要素的分布以及城市空间用地的影响，呈现城市中轴线与东部局部片区集聚的现象，尤其是天安门故宫片区的高游客访问量，使其交通集散、线路组织、公共服务、游客疏导等问题成为城市旅游管理的关键，必须引起政府重视。同时，政府也应加大培育历史城区其他城市功能区的吸引力，使其在突出自身特色基础上成为观光休闲的多元复合体，可以在一定程度上缓解点状城市中心性的问题。

(4) 中心城区的现代城市功能区是新兴的入境旅游空间热点，因核心性质不同而各具特色，且发展势头良好，具有充足的发展空间。现代城市功能区应在产城融合的基础上，加大区域公共服务设施、景观环境建设，在特定人流集聚的公共空间，提供更加多元的休闲消费供给，吸引更多的人前往，使其成为更具活力的城市功能单元。

(5) 郊区以历史古迹为主的观光景点区域是北京市入境旅游重要的空间拓展，目前由于功能过于单一，热点空间仅局限在景区内，且停留与转移概率都低。今后景区应大力推动与周边社区结合共建，改善郊区景区之间的交通联系，将人流疏导到更大的区域，同时培育休闲功能、深化游客体验，帮助景区从观光向休闲观光并重的方向发展。

道路交通、旅游线路服务、景点吸引力、城市公共配套服务等因素对北京市热点区域的分布及游客转移路径模式均有较大的影响,本文在实验结果分析及讨论时简单进行了说明,对于各类因素对旅游流影响的定量分析,将会在后续的研究中重点讨论。

参考文献(References)

- [1] Boniface B, Cooper C. The Geography of Travel and Tourism. Oxford: Butterworth-Heinemann, 1994: 1-6.
- [2] Wu Jinfeng, Pan Xuli. Characteristic of Beijing and Shanghai inbound tourists flow network. *Scientia Geographica Sinica*, 2010, 30(3): 370-376. [吴晋峰, 潘旭莉. 京沪入境旅游流网络结构特征分析. *地理科学*, 2010, 30(3): 370-376.]
- [3] Vu Huy Quan, Li Gang, Law Rob, et al. Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. *Tourism Management*, 2015, 46: 222-232.
- [4] Yang Xingzhu, Jiang Kai, Lu Lin. Urban tourist path trajectory spatial characteristics: An empirical analysis of geotagged photos. *Economic Geograph*, 2014, 34(1): 181-187. [杨兴柱, 蒋锴, 陆林. 南京市游客路径轨迹空间特征研究: 以地理标记照片为例. *经济地理*, 2014, 34(1): 181-187.]
- [5] Yang Xingzhu, Gu Chaolin, Wang Qun, et al. Spatial measure and simulation of urban tourism flow in Nanjing City China. *Scientia Geographica Sinica*, 2011, 31(7): 802-809. [杨兴柱, 顾朝林, 王群, 等. 南京市内部旅游客流空间测度与模拟. *地理科学*, 2011, 31(7): 802-809.]
- [6] Wu Jing, Yang Xingzhu, Sun Jingdong. The spatial characteristics of tourist flows in Nanjing based on the new geographic information technology. *Human Geography*, 2015, 30(2): 148-154. [吴静, 杨兴柱, 孙井东. 基于新地理信息技术的南京市游客流动性空间特征研究. *人文地理*, 2015, 30(2): 148-154.]
- [7] Wang Yiqi, Wu Jinfeng, Han Lining, et al. Geographical distribution and network structure characteristics of Beijing inbound tourist flows. *Journal of Arid Land Resources and Environment*, 2014, 28(6): 202-208. [王奕祺, 吴晋峰, 韩立宁, 等. 北京入境旅游流地理分布与网络特征研究. *干旱区资源与环境*, 2014, 28(6): 202-208.]
- [8] Wang Yongming, Ma Yaofeng, Wang Meixia. A study on spatial diffusion features and diffusion routes of inbound tourist flows of Beijing city. *Economic Geograph*, 2011, 31(6): 1019-1024. [王永明, 马耀峰, 王美霞. 北京入境旅游流空间扩散特征及路径. *经济地理*, 2011, 31(6): 1019-1024.]
- [9] Zhang Youyin, Ma Yaofeng, Ma Hongli, et al. Study on the dynamic shift of Beijing convergence-diffusion tourism flow equilibrium point. *Tourism Tribune*, 2009, 24(12): 31-35. [张佑印, 马耀峰, 马红丽, 等. 北京入境集聚扩散旅游流平衡点转移规律研究. *旅游学刊*, 2009, 24(12): 31-35.]
- [10] Liu Hongying, Ma Yaofeng. Coupling analysis of the linkages between inbound tourism flows and tourism economy: A case study of the westward diffusion of tourists through Shanghai. *Resources Science*, 2008, 30(8): 1162-1168. [刘宏盈, 马耀峰. 入境旅游流空间转移与省域旅游经济联系强度耦合分析: 以上海入境旅游流西向扩散为例. *资源科学*, 2008, 30(8): 1162-1168.]
- [11] Zhang Youyin, Ma Yaofeng, Gu Jing. Study on the potential energy tempo-spatial variation regulation of Beijing indirect inbound aggregating tourist flows. *Tourism Tribune*, 2011, 26(10): 31-35. [张佑印, 马耀峰, 顾静. 北京间接入境聚集旅游流流势时空演化规律研究. *旅游学刊*, 2011, 26(10): 31-35.]
- [12] Niu Yafei, Xie Libo, Liu Chunfeng. The variation tendency and proposed countermeasures for adjusting and controlling visitor flow in Beijing. *Geographical Research*, 2005, 24(2): 283-292. [牛亚菲, 谢丽波, 刘春风. 北京市旅游客流时空分布特征与调控对策. *地理研究*, 2005, 24(2): 283-292.]
- [13] Xuan Guofu, Lu Lin, Wang Degen, et al. Spatial characteristics of tourist flows in coast resorts: A case study of Sanya City. *Geographical Research*, 2004, 23(1): 115-124. [宣国富, 陆林, 汪德根, 等. 三亚市旅游客流空间特性研究. *地理研究*, 2004, 23(1): 115-124.]
- [14] Yang Xinzhu, Gu Chaolin, Wang Qun. Urban tourism flow network structure construction in Nanjing. *Acta Geographica Sinica*, 2007, 62(6): 609-620. [杨兴柱, 顾朝林, 王群. 南京市旅游流网络结构构建. *地理学报*, 2007, 62(6): 609-620.]
- [15] Hu Yingjie, Gao Song, Janowicz Krzysztof, et al. Extracting and understanding urban areas of interest using geotagged photos. *Computers, Environment and Urban Systems*, 2015, 54: 240-254.
- [16] Li Deren, Shao Zhenfeng. The new era of geographic information. *Science China Information Sciences*, 2009, 39(6): 579-587. [李德仁, 邵振峰. 论新地理信息时代. *中国科学 (F辑: 信息科学)*, 2009, 39(6): 579-587.]
- [17] Zheng Yantao, Zha Zhengjun, Chua Tatseng. Mining travel patterns from geotagged photos. *ACM Transactions on Intelligent Systems and Technology*, 2012, 3(3): 1-18.

- [18] 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. *Applied Geography*, 2015, 63: 408-417.
- [19] Paldino S, Bojic I, Sobolevsky S, et al. Urban magnetism through the lens of geo-tagged photography. *EPJ Data Science*, 2015, 4(1): 5.
- [20] Zhang Shaojie. Exploring the travel behaviors of the Xi'an tourism circle [D]. Xi'an: Shaanxi Normal University, 2015. [张少杰. 基于地理标记照片的大西安旅游圈游客时空行为研究[D]. 西安: 陕西师范大学, 2015.]
- [21] Quercia D, Schifanella R, Aiello L M. The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city//Proceedings of the 25th ACM conference on Hypertext and social media. ACM, 2014: 116-125.
- [22] El Ali A, Van Sas S N A, Nack F. Photographer paths: Sequence alignment of geotagged phhaotos for exploration-based route planning//Proceedings of the 2013 Conference on Computer Supported Cooperative Work. ACM, 2013: 985-994.
- [23] Yin H, Wang C, Yu N, et al. Trip mining and recommendation from geo-tagged photos//Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference on. IEEE, 2012: 540-545.
- [24] Chen Y Y, Cheng A J, Hsu W H. Travel recommendation by mining people attributes and travel group types from community-contributed photos. *IEEE Transactions on Multimedia*, 2013, 15(6): 1283-1295.
- [25] Orsi F, Geneletti D. Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. *Journal for nature conservation*, 2013, 21(5): 359-368.
- [26] Zhu Jinyue, Hu Tao. A study of the spatial distribution of tourists based on geotagged photos: Case of Hainan Province. *Tourism Forum*, 2016, 9(6): 17-22. [朱金悦, 胡涛. 基于地理标记照片的游客空间分布特征研究: 以海南省为例. 旅游论坛, 2016, 9(6): 17-22.]
- [27] Vu Huy Quan, Leung Rosanna, Rong Jia, et al. Exploring park visitors' activities in Hong Kong using geotagged photos//Information and Communication Technologies in Tourism 2016. Springer, Cham, 2016: 183-196.
- [28] Kisilevich S, Mansmann F, Keim D. P-DBSCAN: A density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos//Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application. ACM, 2010: 38.
- [29] Xia Jianhong, Zeepongsekul Panlop, Arrowsmith Colin. Modelling spatio-temporal movement of tourists using finite Markov chains. *Mathematics and Computers in Simulation*, 2009, 79(5): 1544-1553.
- [30] Xu Zhenxing, Chen Ling, Chen Gencai. Topic based context-aware travel recommendation method exploiting geotagged photos. *Neurocomputing*, 2015, 155(C): 99-107.

Exploring the spatial characteristics of Beijing inbound tourist flow based on geotagged photos

QIN Jing¹, LI Langping², TANG Mingdi^{1,3}, SUN Yan¹, SONG Xinrui¹

(1. The School of Tourism Management, Beijing International Studies University, Beijing 100024, China;
2. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences
and Natural Resources Research, CAS, Beijing 100101, China; 3. The School of Landscape Architecture,
Beijing Forestry University, Beijing 100083, China)

Abstract: As one of the most important modern tourism destinations, cities are playing a critical role in developing agglomerated tourism elements, shaping urban leisure space and promoting urban life quality. In particular, the city tourist flow provides a key path for connections between tourism elements and spatial organization, and its influence on urban traffic organization, public service facilities and leisure tourism industry layout has also been increasing. Motivated by these observations, this paper examines spatial characteristics of inbound tourist flows from a microscopic perspective within a city. Firstly, all photos with geographical and attribute information in Beijing from 2004 to 2015 are obtained from the Flickr API interface using data mining technology. Then, a series of data cleaning methods are used to enhance initial data quality and a database is established. Finally, spatial pattern of inbound tourist flows and spatial characteristics of areas of interest (AOI) are analyzed using various methods including P-DBSCAN and Markov chain. The main results are as follows: (1) Spatial characteristics of AOI in the historic center (the region inside the 2nd Ring Road), the central urban area (the region between the 2nd and 5th ring roads) and the suburbs (the region outside the 5th Ring Road) in Beijing are significantly different: the AOI in the historic center in Beijing cover spaces of scenic spots, leisure, entertainment, restaurants, shops and other related industries, and the boundaries between AOI are unclear, with one AOI connecting to another; AOI in the central urban area are closely related to urban functional areas, and has a definite orientation of urban functions; finally, AOI in the suburbs of Beijing are consistent with the tourist attractions, with the boundaries not going beyond scenic areas. (2) The traditional, historical and cultural AOI dominate other AOI. Meanwhile, the commercial, shopping and leisure AOI as well as modern city function related AOI are becoming more and more attractive. (3) Inbound tourist flows are mainly distributed in AOI of the historical districts and the northwest and east of the central urban areas. Moreover, there is no big tourist flow between the suburban regions. Overall, the inbound tourist flow in Beijing follows a radioactive distribution pattern centered at the historic center. (4) AOI of different regions are connected with secondary tourist flows, and the inbound tourism flow in Beijing has formed a preliminary network structure.

Keywords: geotagged photos; tourist flow; area of interest; spatial characteristics; Beijing