

# 基于电子地图兴趣点的城市建成区边界识别方法

许泽宁<sup>1,2</sup>, 高晓路<sup>1</sup>

(1. 中国科学院地理科学与资源研究所 中国科学院区域可持续发展分析与模拟重点实验室, 北京 100101;  
2. 中国科学院大学, 北京 100049)

**摘要:** 城市建成区边界是认识和研究城市的重要基础性信息,也是落实城市功能空间布局、实施界限管控的前提。但是,以往通过夜间灯光的强度、土地覆被或建筑覆盖率等信息获取城市空间范围的方法,由于受到数据精度和尺度限制,对城市社会经济活动的解释性不强,因而存在较大局限性。电子地图兴趣点(POI)作为城市空间分析的基础数据之一,直观且有效地反映了各类城市要素的集聚状况。本文基于POI与城市空间结构和城市要素空间分布的关联性,提出了一种新的通过POI密度分布来判别城市建成区边界的技术方法。为此,开发了Densi-Graph分析方法,用来分析POI密度等值线的变化趋势,在此基础上对城乡过渡地带的阈值识别的方法进行了理论分析,并讨论了单中心圆结构、双中心“鱼眼型”结构、双中心“子母型”结构等各类城市POI密度等值线的生长规律,证明了Densi-Graph分析方法的适用性。较之以往的城市建成区边界识别方法,这种方法的基础数据更加直观可信,分析结果也更加客观。运用这种方法,本文对全国地级以上城市的建成区边界进行了实证分析,探索了城市建成区边界的阈值及其与城市人口规模、城市所在区域之间的关系。

**关键词:** 城市边界;城市建成区;POI;Densi-Graph;核密度分析;城乡空间结构

DOI: 10.11821/dlxb201606003

## 1 引言

城市建成区边界是认识和研究城市的重要基础性信息,也是落实城市功能空间布局、实施界限管控的前提。在城市和区域规划中,城市建成区边界是确定未来城市增长边界的重要依据,也是通过土地利用政策和空间管控手段对城市增长进行管理所不可或缺的信息<sup>[1]</sup>,同时,城市建成区为评估城市发展和城镇化建设的质量提供了必不可少的数据支持。在快速城镇化背景下,中国城市发展超过规划范围的情况屡见不鲜。例如,Han等<sup>[2]</sup>通过卫星遥感影像和GIS技术发现,北京六环以内规划城市建设用地边界外部的城镇建设用地增长规模普遍高于边界内部的增长规模;田莉等<sup>[3]</sup>和徐毅松等<sup>[4]</sup>发现,广州和上海也有大量城市开发位于规划城镇建设用地边界之外。可见,只有及时有效地把握城市建成区扩展的实际情况,才能准确诊断城市发展中出现的问题并提出相应的空间管制措施,从而实现城市的紧凑集约发展和土地的高效利用。

在现有城市边界研究中,行政区划边界和城市规划建设用地的边界经常代替城市建

收稿日期: 2015-04-23; 修订日期: 2016-02-17

基金项目: 中国科学院重点项目(KZZD-EW-06-04); 国家自然科学基金项目(41171138) [Foundation: The Key Project of Chinese Academy of Sciences, No.KZZD-EW-06-04; National Natural Science Foundation of China, No.41171138]

作者简介: 许泽宁(1990-), 男, 博士, 主要从事城市地理、流动人口与城镇化研究。E-mail: xuzn.13b@igsrr.ac.cn

通讯作者: 高晓路(1969-), 女, 研究员, 博士生导师, 主要从事城市地理、城市和区域发展政策研究。

E-mail: gaoxl@igsrr.ac.cn

成区边界出现。尽管这些“边界”概念十分相近,但是用它们代替城市建成区边界不够精确,为避免混淆,有必要对这些概念进行界定。行政区划边界是城市的法定边界,是一种具有法律层面含义的界线,城市行政区与城市规划范围具有密切联系<sup>[5]</sup>,各类空间规划的范围界定通常以中心城区的行政区划边界作为参考。然而城市实际发展范围和中心城区行政区划无法完全重合,因此,在大部分情况下直接使用行政区划边界作为城市建成区边界显然并不合适。从狭义理解,城市规划建设用地边界是通过规划划定的城市增长边界(Urban Growth Boundary, UGB)<sup>[6]</sup>,即允许进行城市住宅、公共设施和基础设施建设的土地范围界限,是一种城市增长管理、控制无序蔓延的工具<sup>[7]</sup>。在实际情况中,既可能存在UGB内部的建设用地未被征用的情况,又可能出现建设用地超出规划范围的现象。近年来,为了推动城市建设,超出建设用地边界的土地过度开发成为各级城市的普遍现象<sup>[8]</sup>,出现了严重的建设用地增速过快<sup>[9]</sup>、整体利用效率不高<sup>[10]</sup>的问题。另外,即使土地被合理征用,可是如果依旧处于闲置状态,把它们视为城市建成区也有失偏颇。因此,本研究把城市建成区边界定义为实际已经建设并发展起来的城市空间范围,也就是说,建成区是展现城市居民日常生活和生产活动的空间范围。

为了获得对城市建成区边界更为客观的认识,学者们以遥感解译为基础,提出了基于形态学分析的方法——如通过夜间灯光的强度<sup>[11-13]</sup>、土地植被或建筑覆盖率<sup>[14-18]</sup>等信息来判别城市空间的实际范围,构建经济活动密度、与中心城经济联系的强度、土地利用和建筑物密度等指标作为划分城市建成区边界的参考依据。虽然形态学方法可以较为客观地描述城市建成区边界范围,分析过程也比较简单,但由于对城市经济活动及其联系方式的考虑不足,分析结果存在不确定性<sup>[19]</sup>。例如大城市边缘的一些城乡结合部地区,由于存在大量的人口活动,其经济、社会活动与城市联系十分紧密,应该被视为城市建成区的一部分。但这些地区在形态和结构上仍属于农村,很难通过遥感解译识别出来。另外,也有学者利用人口密度进行城市范围研究。人口作为劳动力或消费者,其密度代表着经济活动的强弱程度,人口密度因而成为研究社会经济现象的地理规律的基础<sup>[20]</sup>。周一星等<sup>[21]</sup>提出了城市实体地域的概念,并将城市统计区标准设定为2000人/km<sup>2</sup>。如日本的人口集中地区(densely inhabited district, DID)将4000人/km<sup>2</sup>作为高密度城镇化地区或高密度连绵区的识别依据<sup>[22]</sup>。但是,由于人口统计周期长、成本高,当城市之间人口流动性较强时,人口密度统计数据的实效性较弱,难以客观反映城市建成区边界的动态性变化。

值得注意的是,城乡边界阈值是判定城市建成区边界的关键要素,决定着最终边界的范围,如利用DMSP/OLS夜间灯光数据进行城市建成区的提取,重点就在于寻找最佳的阈值<sup>[10]</sup>。然而在已有研究中,阈值几乎完全依赖于人为判断,很难做到足够精确,相同阈值是否能够通用也缺乏相关的研究和结论<sup>[9, 23-24]</sup>,如Imhoff等<sup>[23]</sup>采用的灯光数据突变检测法通用性低,无法适用于区域差异较大的地区。同时,也有学者提出了城市边缘区(urban fringe)的概念<sup>[25-26]</sup>,以城乡之间的过渡性带状区域代替线状界线,也可以很好地反应城市边界,但不论是利用指标体系法亦或是遥感解译法,带状边界的确定仍然有一定的不足之处<sup>[27]</sup>。因此,亟需一种更为合理、客观和准确的方法来判别城市建成区边界。

大数据(Big Data)时代的到来给城市边界研究带来了新的思路与方法,体现出大量化、多样化和快速化特点。网络空间的数据体量不断扩大、数据类型丰富,常常以数据流的形式动态、快速地产生,具有很强的实效性<sup>[28]</sup>。电子地图兴趣点(point of interest, POI)数据是一类重要的地理空间大数据,在更新速度和获取成本方面,较遥感和人口密度数据有一定优越性。鉴于此,本文试图提出一种客观寻找城乡边界阈值的方法

法, 利用电子地图兴趣点 POI 数据, 以密度分析为基础来识别城市建成区的边界范围, 以期为认识和分析城市提供参考与帮助。

## 2 数据与方法

### 2.1 POI 与城市活动的关联

POI 是指具有地理标识的空间特征物, 包含名称、类别、经纬度等信息, 是空间大数据分析的基础性数据。电子地图中每一个地名、建筑、住宅小区、公园、学校、医院、公司、商场等都是一个兴趣点。电子地图中含有大量 POI 数据, 如 2013 年的百度电子地图中含有全国数千万条 POI 数据信息。POI 信息是数据供应商通过实地调查采集的, 信息丰富而准确, 更新及时, 而且贴近人们的生活。例如, 2008 年全国百度地图的 POI 数据共 600 多万条, 2011 年增加至 1100 多万条。随着电子地图的广泛使用, POI 数据的应用日益受到人们的重视。

POI 数据可以在空间上反映城市结构和与道路骨架相对应的分布特征。由于道路两侧的居民、商户较多, 大部分 POI 集中分布在道路两侧和街区区块当中, 当道路两侧的地标和商户较多时, POI 数据与道路骨架对应的分布特征更加明显, 这种空间分布模式直接体现出城市的肌理和各类城市活动的集聚状况。为了证明 POI 数据的稳定性和有效性, 对 2008 年以来不同年份百度电子地图的 POI 数据进行了对比分析。发现 POI 信息的增加主要发生在规模较小的城市和乡村, 也就是以前信息采集没有覆盖的地方。而在人口规模较大的城市、特别是地级以上城市的建成区, POI 数量和密度的变化很小, 沿道路两侧分布的布局方式并没有明显改变。

POI 密度的空间分布差异反映了不同的区域发展水平<sup>[29]</sup>, 为了验证兴趣点类型对密度分布的影响, 按照城市市辖区和县级市, 以及其他县乡的行政范围对各类兴趣点的比例分布进行了统计, 发现占 POI 总量 9% 的政府机构类 POI 在城乡之间的分布相对均匀, 而住宅、交通、企业、商业服务、公园广场、公共服务设施等其他种类的 POI 数量分布都存在显著的城乡分异 (表 1), 在市辖区和县级市范围内的密度明显较高。另外, 考虑到各城市功能定位不同 (如有的城市商业活动较为发达, 有的城市企业较多), 采用细类 POI 数据会产生较大误差。因此在全国尺度上宜采用全类别 POI 数据进行城市边界的分析。

### 2.2 POI 的核密度分析

城市作为一个社会经济综合体, 集聚效应和规模效益是其最基本的特征, 而这一特点恰好可以通过 POI 的密度动态反映出来。郊区和农村因地形地貌、土地利用等自然要素的差别和人口密度、聚落形式的不同, POI 密度与城市地区相比有很大的差异, 表现在 POI 密度分布图上, 城市建成区的 POI 相对密集, 而在建成区边界附近, POI 密度显著降低。城市地区的 POI 密度会显著高于城市外围的农村地区的这一规律对于判别城市建成区边界具有很大的参考价值, 由此利用密度的明显变化, 进行城市建成区边界的判定。

表 1 2008 年全国百度地图各类型 POI 的比例分配

Tab. 1 The proportion of different categories of POI data in cities and counties of China

	行政 机构	交通 站点	银行	商业 大厦	零售 行业	宾馆 酒店	餐饮 娱乐	医院	学校科 研院所	公司 企业	公园 广场	住宅 小区	综合 信息	总计
总体比例(%)	9.0	10.6	5.7	0.6	21.4	2.8	15.0	3.3	5.4	12.8	0.2	3.1	10.1	100
市辖区和县级市(%)	53.3	80.8	72.8	90.1	77.7	73.1	82.2	72.1	66.6	82.6	79.2	87.4	71.4	75.5
其他县乡(%)	46.7	19.2	27.2	9.9	22.3	26.9	17.8	27.9	33.4	17.4	20.8	12.6	28.6	24.5



为此, 首先建立 POI 概率密度的空间分布。非参数估计是概率密度估计的有效手段, 其中, 核密度估计方法 (Kernel Density Estimation) 可以对点数据进行高质量的密度估计, 且不会受到栅格大小和位置的影响<sup>[30]</sup>, 已经被广泛应用于地理学中的景观格局分析<sup>[31]</sup>、商业空间布局分析<sup>[32]</sup>、就业空间布局分析<sup>[33]</sup>、产业空间分布<sup>[34]</sup>等方面。

核密度分析将空间中任意一点周围的一定规则区域作为密度的计算范围, 通过计算该区域内观测数据的密度来分析观测对象的空间分布态势。距离中心点近的数据点被赋予较高的权重, 反之则赋予较低的权重, 每一个点的估计密度都是该区域所有点的加权平均密度。其中, 空间中任意一点  $i$  的核密度  $P_i$  被定义为中心处最高并向外不断降低的函数, 当距离达到一定值  $R$  以后密度为 0:

$$P_i = \frac{1}{n\pi R^2} \times \sum_{j=1}^n K_j \left(1 - \frac{D_{ij}^2}{R^2}\right)^2 \quad (1)$$

式中:  $K_j$  为研究对象  $j$  的权重;  $D_{ij}$  为空间点  $i$  与研究对象  $j$  的距离;  $R$  为选定规则区域的带宽 (bandwidth) ( $D_{ij} < R$ );  $n$  为带宽  $R$  的范围内研究对象  $j$  的数量。

研究表明, 带宽  $R$  的选择对核密度分析的结果有关键影响<sup>[35]</sup>。对于空间数据的核密度分析来说, 带宽需要根据点的实际空间分布以及研究问题合理确定, 较小的带宽适合反映密度分布的局部变化, 较大的带宽则能有效地反映全局尺度的空间变化。在进行城市建成区边界识别时, 即要考虑到城市尺度上 POI 数据的分布差异, 又要避免计算过程中将邻近城市或周边功能区纳入同一边界范围, 从而避免 POI 数据的空间分异特征被过于强化或弱化。Hinnerburg 等<sup>[36]</sup>对带宽和密度中心 (density-attractor) 数量的关系进行了研究, 发现存在一定数量的带宽区间使得密度中心保持稳定, 在这些区间内选择带宽是较为合理的。通过对 POI 密度分布的估算, 将带宽区间确定为 3000~3500 m, 此时分析结果有着良好的稳定性, 且能够满足城市空间结构分析的需要 (图 1)。

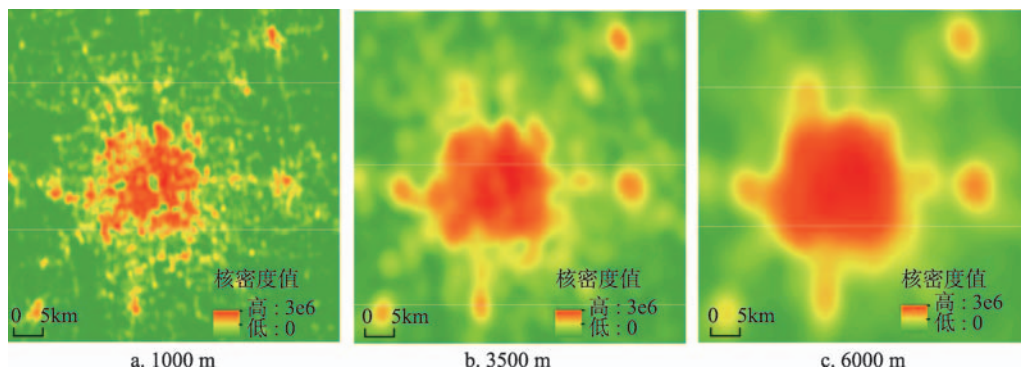


图 1 不同带宽情况下的核密度估计结果

Fig. 1 Kernel estimation of different bandwidths

### 3 利用 Densi-Graph 识别城市建成区边界的阈值

对 POI 数据进行核密度分析, 并根据核密度分析结果绘制核密度等值线 (图 2)。POI 的核密度等值线是以城市 POI 密度的最高点为中心 (有时是多个中心), 向四周不断扩展的不规则闭合曲线。其空间分布的特征表现为, 在密度中心附近, 等值线间的距离相对紧密, 在城市外围, 等值线的间距逐渐扩大, 到了城市外围的农村地区, 等值线的间距变得非常稀疏。那么, 如何确定核密度等值线由密至疏的临界值成为判别城市建成区边

界的关键问题。

为了解决这个问题,提出一种新的用于分析密度变化的方法,将之命名为“Densi-Graph”。Densi-Graph的基本思路是观察密度值 $d$ 与相对应闭合曲线围成面积 $S_d$ 的理论半径 $S_d^{(1/2)}$ 的变化情况,找到密度等值线的临界值,从而确定城市建成区边界。分别以密度值 $d$ 与理论半径增量 $\Delta S_d^{(1/2)}$ 为坐标轴做图,即Densi-Graph图。具体包括两个步骤。

第一步:刻画密度值 $d$ 与面积 $S_d$ 的关系。定义 $S_d^{(1/2)}$ 为 $S_d$ 的理论半径,并对理论半径的增量 $\Delta S_d^{(1/2)}$ 求导数。理论上说,当下式成立,即:

$$\lim \frac{d(\Delta S_d^{(1/2)})}{dd} = 0 \quad (2)$$

时,说明密度曲线是均匀的;当下式成立,即:

$$\lim \frac{d(\Delta S_d^{(1/2)})}{dd} > 0 \quad (3)$$

时,说明密度曲线向外发散。反之,如果导数值小于零,则说明密度曲线收缩。

第二步:临界值的判定。实际上,由于多个城市中心的存在或城市的非均匀扩张,POI密度分布在城市内部存在一定变化,因此,Densi-Graph曲线在城市建成区范围内表现为波动。尽管如此,从整个城市的尺度上看,在城市建成区的边界处应该存在一个具有全局意义的临界值点。 $r$ 定义为城市空间结构变化的容许值,那么当出现:

$$\lim \frac{d(\Delta S_d^{(1/2)})}{dd} > r \quad (4)$$

时,即波动趋势出现显著增长且该趋势不可逆,可以认为此处是一个全局拐点,此时的密度值即为城市建成区边界的临界值。由此即可确定城市建成区的边界。

对不同的城市结构对Densi-Graph的适用性进行分析:

假想1:单中心城市结构。假设城市仅有1个密度中心,且无限均匀扩张,其核密度等值线表现为单中心的同心圆结构(图3a)。由于POI密度从密度中心点向四周等值均匀扩张,Densi-Graph曲线表现为一条水平的直线(图3b)。

假想2:双中心城市结构。假设城市有2个密度中心,且无限均匀扩张。当两个密度中心密度值相等时,POI密度曲线从两个密度中心分别向四周等值均匀扩张,等值线在两个密度中心中点处合并,核密度等值线表现为双中心“鱼眼型”结构(图3c)。当两个中心独立生长且尚未融合时,Densi-Graph曲线表现为一条直线;当密度中心扩张到一定范围并导致相互融合时,Densi-Graph曲线在相交处陡降,之后随着等值线的不断扩展,曲线趋于水平(图3d)。

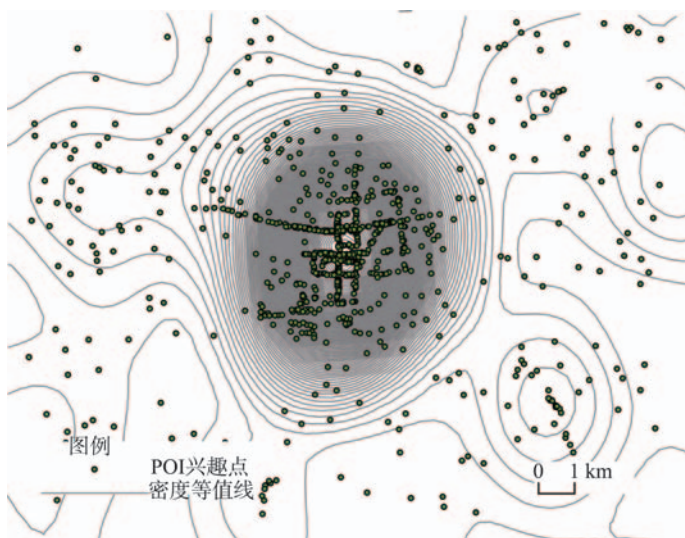


图2 POI核密度等值线实例

Fig. 2 An example of POI density contour

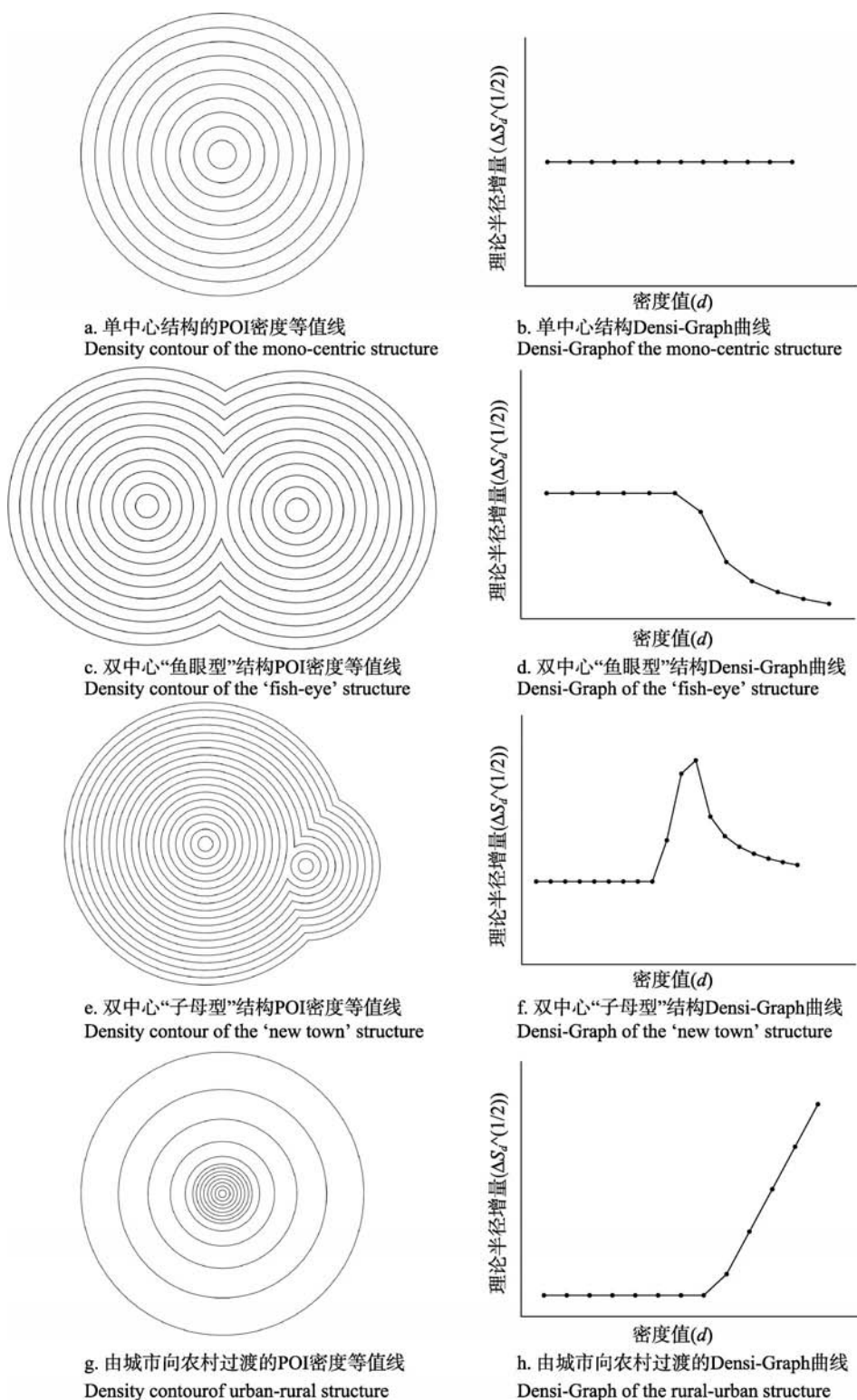


图3 不同假想城市结构的POI核密度等值线和Densi-Graph分析

Fig. 3 Densi-Graph analysis of the POI density contour for different spatial structures



假想3: 双中心城市结构。假设城市有2个密度中心, 且无限均匀扩张。当两个密度中心密度值不相等时, 核密度等值线先从一个密度中心向四周生长, 在扩张至一定程度时, 另一密度中心出现并与之同时生长, 核密度等值线表现为双中心“子母型”结构(图3e)。当两个中心独立生长且尚未融合时, Densi-Graph曲线表现为一条直线; 在另一个密度中心出现时, 曲线呈上升趋势; 在等密度曲线融合时, Densi-Graph曲线达到峰值, 之后随着等值线的不断扩展, 曲线趋于水平(图3f)。

假想4: 单中心城市向农村过渡。假设城市有1个密度中心, 在城市建成区的范围内均匀扩张, 在城市外围农村地区POI密度开始下降, 变得越来越稀疏。其核密度等值线表现为内密外疏的单中心圆结构(图3g)。在密度曲线均匀扩张的过程中, Densi-Graph曲线表现为一条水平的直线; 在由城市过渡到农村后, Densi-Graph曲线不断上升, 并在城市—农村临界处形成明显的拐点(图3h)。

通过以上分析可以看出, 在均匀扩张的城市空间范围内, 城市密度中心的数量和城市密度中心的融合都不会影响Densi-Graph曲线的整体走势, 曲线最终都会趋向水平。而在城市和农村的过渡地带, Densi-Graph曲线会出现上升且不可逆转的趋势。在四种假设条件下, 都可以较为理想地对计算结果做出合理的解释。对于实际城市空间的密度等值线, 都是由以上4种假设结构衍生出来的。随着城市的发展与扩张, 会出现不同的POI密度中心, Densi-Graph曲线整体趋势中表现为不同程度的细微波动。这类波动是可逆的, 会随着密度中心的融合而趋于水平。在城市边缘地带, POI密度的整体下降会使曲线产生不可逆的上升趋势。因此, 根据Densi-Graph曲线的这一特点, 可以有效对城市建成区的边界进行识别。

## 4 城市建成区边界识别的实证分析

基于2008年百度地图的POI数据, 对全国所有城市的建成区边界进行了识别分析。研究涉及的城市共有523个, 其中地级市226个, 县级市297个。分析结果表明, 这种方法在地级城市尺度上具有较好的适应性。这主要是因为这些城市的电子地图和POI数据比较完善。对于地级市以下的城市而言, 由于很多城市的POI数据尚不完善, 且POI分布较为分散, 计算过程中的相关参数有待进一步研究, 故不在文中讨论。对于一些紧邻的地级市(如广州和佛山、苏锡常等地区), 由于其城镇化水平较高、区域一体化高度发达, 城市建成区出现了融合的现象。在对这些地区进行分析时, 将这些地区作为一个整体进行边界的识别, 并利用城市行政区划边界将其分开。

对POI数据进行核密度分析, 并绘制出Densi-Graph曲线。接着对Densi-Graph曲线求导, 对其进行去量纲化处理。通过Densi-Graph分析发现, 特大城市局部地区POI密度分布的波动性较大, 而中小型城市POI密度分布的波动性则相对较小。波动较大主要是因为多中心的存在(如图3d、3f所示, 多中心的存在使Densi-Graph曲线出现较大的下落)以及存在一些大型绿地公园、交通和产业设施等较高等级的区域性服务设施(这些地方形成的谷地会使Densi-Graph曲线出现较大的上扬)所致; 波动较小则是由于要素集聚相对集中, 密度中心相对较少所致。但是, 根据不同城市分别提出容许值存在大量主观性, 也会产生极大误差。由于POI密度会在城乡交界地带产生明显变化, 因此把Densi-Graph曲线的增长率作为容许值的确定依据。在核密度曲线分布图上, Densi-Graph曲线的增长率代表相邻两条等值线之间距离的增长速度, 是一个不依赖于等值线间隔取值的无量纲数。根据全国地级市建成区边界的分析, 确定将5%的增长率作为容许值, 当

曲线的增长率超过5%, 且增长趋势不可逆, 则把该曲线作为城市建成区边界线。我们发现, 由于不同城市POI密度变化的差异, 特大城市(如北京、天津)在容许值处的曲线变化值比其他城市的2~3倍。

国内外有很多学者都将城市边界作为城乡间的带状区域进行研究, 本文仅对城市建成区的线状边界进行研究。在本文的基础上, 未来可以通过设定Densi-Graph曲线的取值范围, 对城市边缘区进行提取与识别。

不同空间结构的城市, 城市建成区边界的识别结果表现出很大差别。本文以北京(单中心结构)、包头(双中心结构)、兰州(带状结构)3个具有代表性的城市作为案例进行展示。以3500 m为带宽, 对3个城市进行核密度分析并生成密度等值线。计算每条密度等值线的理论半径并画出3个城市的Densi-Graph曲线, 利用Densi-Graph曲线计算不同城市的容许值, 从而计算出3个城市建成区边界的Densi-Graph值。图4是3个城市建成区边界识别的结果。可以看出, 基于POI数据的Densi-Graph分析方法对于单中心城市、双中心城市和串珠状城市的城市边界识别都具有良好的适用性。随着POI密度的降低, Densi-Graph曲线呈现出相同的走势, 而且均出现可以识别的拐点。由于城市内部空间结构的差异, 3个城市的Densi-Graph曲线都有不同程度的细小波动, 这种波动是由城市内部各密度中心的融合而引起的。最终城市建成区边界的阈值, 则选取在曲线开始显著发散的拐点处。

运用以上方法, 对全国226个地级市的城市建成区边界的Densi-Graph曲线阈值逐一进行了分析。结果表明, 不同人口规模和区域的地级市的城市建成区边界的阈值存在一定的规律性。据此, 对Densi-Graph曲线阈值与城市人口规模和城市所在地区的联系进行了总结和归纳(表2)。结果表明, 华东和华南比较相近, 华北、东北和华中地区比较相近, 西北、西南地区比较接近。可以看出, 随着人口规模的不断提升, 城市建成区边界的阈值逐级上升。另外, 华东、华南地区阈值最高, 华北、东北、华中地区其次, 西北、西南地区城市边界值最低。从一定意义上说, 这一结果印证了POI反映城市空间结构肌理和活动规律的假设。人口规模较大的城市, 通常具有较高的规模等级和中心性, 城市活动更加活跃, 其对郊区和周边乡村的带动性也比较强; 同理, 经济发展水平较高的华东和华南沿海城市也是这样。因此, 人口规模较大的城市 and 经济发展较好的地区表现出较高的城市建成区边界阈值, 而人口规模较小或经济发展水平偏低的地区的城市建成区边界阈值则较低。

## 5 结论

城市建成区边界是认识和研究城市的基础。针对现有研究和实践中对城市边界的认识比较含混、界定方法具有较大主观性等问题, 本文提出了一种新的利用POI核密度等值线的分布规律来区分城市建成区边界的分析方法, 并成功地将之用于地级市城市建成区边界的分析。与遥感、人口密度等建成区边界划定方法相比, 基于POI数据和Densi-Graph方法的建成区边界识别在实效性和客观性方面有较大提升。本文的主要价值有:

(1) 结合大数据分析思路, 利用电子地图的POI数据进行城市建成区边界的识别。通过对POI和城市空间结构与城市活动之间关联的分析, 证明了POI数据的应用价值, 特别是在全国地级市尺度城市边界识别中的可行性。这一实例对于拓展城市研究中的大数据应用具有一定价值。

(2) 开发了Densi-Graph分析方法, 用于城市建成区边界的识别, 并讨论了单中心结



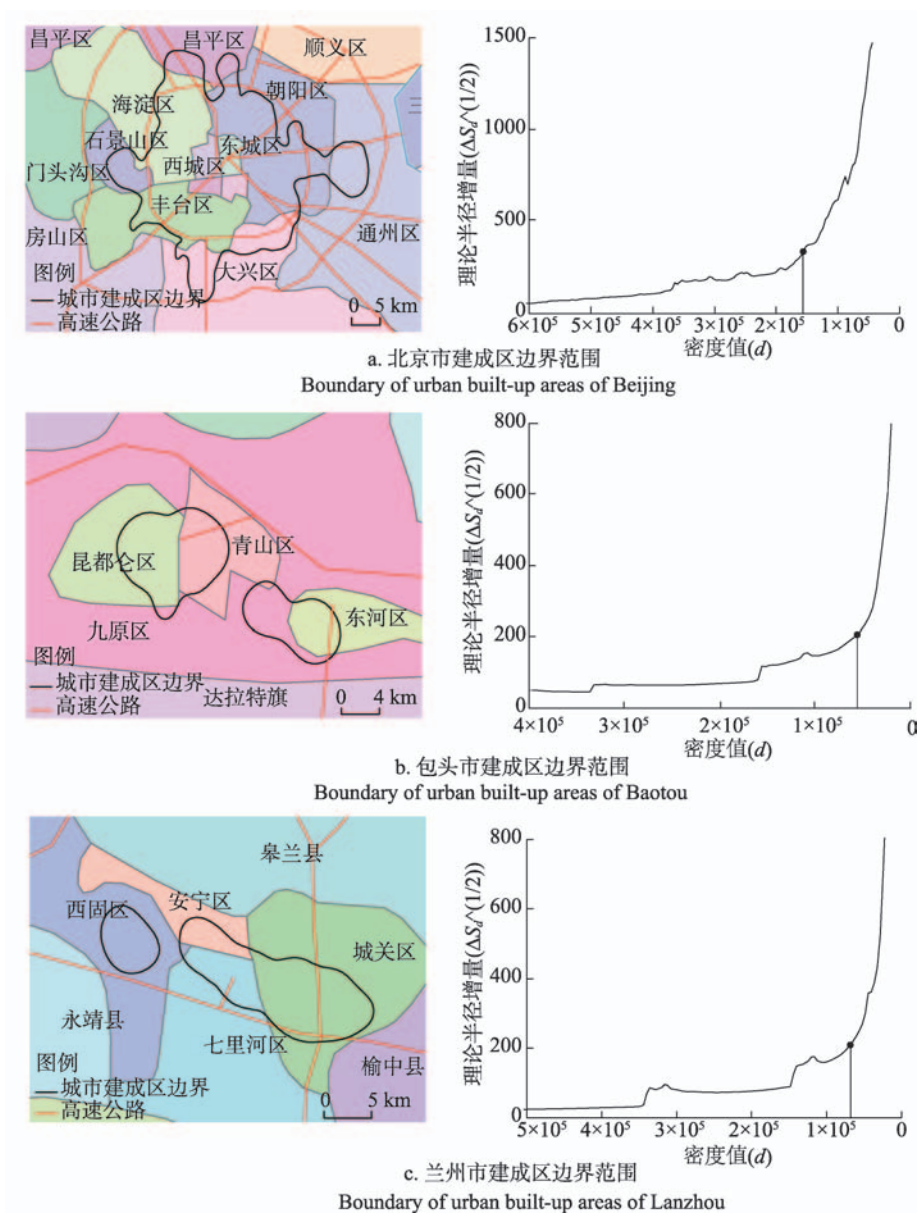


图4 典型城市的城市建成区边界范围

Fig. 4 Boundary of urban built-up areas of three typical cities

构、双中心“鱼眼型”结构、双中心“子母型”结构等不同空间结构城市的核密度等值线的分布情况，从而证明了Densi-Graph的理论合理性和实际可行性。Densi-Graph分析方法不仅可以用于城市建成区边界的识别，也可以为其他各种基于密度分布的地域空间分析提供一定的参考。

表2 不同规模和区域城市建成区边界的Densi-Graph曲线平均阈值的参数

Tab. 2 Average threshold values of the Densi-Graph contour of urban built-up areas in different cities

市辖区 人口规模	地区		
	华东、华南	华北、东北、华中	西北、西南
1000万以上	156000	156000	
500~1000万	124000	84000	72000
300~500万	76000	60000	48000
100~300万	60000	44000	44000

(3) 利用Densi-Graph分析方法, 对全国226个地级市的城市建成区边界进行识别, 并根据实证研究的结果, 提出了判别城市建成区的关键参数。例如, 核密度方程的带宽值(3000~3500 m之间), 地级市的Densi-Graph曲线容许值等, 这些参数为城市建成区边界的判定提供了有益的参考。研究还发现了全国地级市城市边界等值线阈值与城市人口规模和城市所在区域的相关规律。人口规模等级结构越高, 城市边界值越大; 华东、华南地区城市边界值最大, 华北、东北、华中地区其次, 西北、西南地区城市边界值最小。这些参数和新的知见为客观认识城市边界, 以及未来开展更加深入的城市研究奠定了科学基础。

(4) 这个方法也适用于更大范围的城市群的边界识别。本文的研究思路也适用于其他与空间边界有关的各类研究。例如, 通过重新确定核密度分析的带宽, 以“点—轴系统”发展理论为基础, 利用空间社会经济联系和城市体系的基础性数据, 可以对城市群的发育程度和城市群边界做出判断与识别。

致谢: 本研究得到了龙瀛和BCL研究团队在数据方面的大力支持及樊杰研究员、陈田研究员、吝涛、王芳的宝贵建议, 特此感谢。

## 参考文献(References)

- [1] Pendall R, Martin J, Fulton W. Holding the line: Urban containment in the United States. Washington D.C.: The Brookings Institution Center on Urban and Metropolitan Policy, 2002.
- [2] Han Haoying, Lai Shihkung, Dang Anrong, et al. Effectiveness of urban construction boundaries in Beijing: An assessment. Journal of Zhejiang University, 2009, 10(9): 1285-1295.
- [3] Tian Li, Lv Chuanting, Shen Tiyan. Theoretical and empirical research on implementation evaluation of city master plan: A case of Guangzhou City Master Plan (2001-2010). Urban Planning Forum, 2008(5): 90-96. [田莉, 吕传廷, 沈体雁. 城市总体规划实施评价的理论与实证研究: 以广州市总体规划(2001-2010年)为例. 城市规划学刊, 2008(5): 90-96.]
- [4] Xu Yisong, Shi Song, Fan Yu. Methodology of Shanghai urban master planning in new position. Urban Planning Forum, 2009(2): 10-15. [徐毅松, 石崧, 范宇. 新形势下上海市城市总体规划方法论探究. 城市规划学刊, 2009(2): 10-15.]
- [5] Wang Guoen, Zhang Yuanyuan. Urban growth boundary efficacy and its influence on administrative boundary adjustment. Planners, 2012, 28(3): 21-27. [王国恩, 张媛媛. 城市增长边界的效能及对行政区划调整的影响. 规划师, 2012, 28(3): 21-27.]
- [6] Long Ying, Han Haoying, Mao Qizhi. Establishing urban growth boundaries using constrained CA. Acta Geographica Sinica, 2009, 64(8): 999-1008. [龙瀛, 韩昊英, 毛其智. 利用约束性CA制定城市增长边界. 地理学报, 2009, 64(8): 999-1008.]
- [7] Xu Kang, Wu Shaohua, Chen Dongxiang, et al. The urban growth boundary determination based on hydrology effect: Taking Xinminzhou as an example. Scientia Geographica Sinica, 2013, 33(8): 979-985. [徐康, 吴绍华, 陈东湘, 等. 基于水文效应的城市增长边界的确定: 以镇江新民洲为例. 地理科学, 2013, 33(8): 979-985.]
- [8] Economic Information Daily. [http://jjckb.xinhuanet.com/gnyw/2010-01/11/content\\_201839.htm](http://jjckb.xinhuanet.com/gnyw/2010-01/11/content_201839.htm), 2010-01-11. [经济参考报. [http://jjckb.xinhuanet.com/gnyw/2010-01/11/content\\_201839.htm](http://jjckb.xinhuanet.com/gnyw/2010-01/11/content_201839.htm), 2010-01-11.]
- [9] Sun Pingjun, Zhao Feng, Xiu Chunliang. Study on the spatial differentiation of urban construction land input efficiency in China. Economic Geography, 2012, 32(6): 46-52. [孙平军, 赵峰, 修春亮. 中国城镇建设用地投入效率的空间分异研究. 经济地理, 2012, 32(6): 46-52.]
- [10] Sun Pingjun, Ding Sibao, Xiu Chunliang. Dynamic evolution research on non-conformity of dynamic of urban construction land input in China. Scientia Geographica Sinica, 2012, 32(9): 1047-1054. [孙平军, 丁四保, 修春亮. 中国城镇建设用地投入非协调性的动态演变研究. 地理科学, 2012, 32(9): 1047-1054.]
- [11] Shu Song, Yu Bolang, Wu Jianping, et al. Methods for deriving urban built-up area using night-light data: Assessment and application. Remote Sensing Technology and Application, 2011, 26(2): 169-176. [舒松, 余柏浪, 吴健平, 等. 基于夜间灯光数据的城市建成区提取方法评价与应用. 遥感技术与应用, 2011, 26(2): 169-176.]

- [12] Liu Qiping, Yang Yongchun, Fu Dongxia, et al. Urban spatial expansion based on DMSP\_OLS nighttime light data in China in light in 1992-2010. *Scientia Geographica Sinica*, 2014, 34(2): 129-136. [刘沁萍, 杨永春, 付冬霞, 等. 基于DMSP\_OLS灯光数据的1992-2010年中国城市空间扩张研究. *地理科学*, 2014, 34(2): 129-136.]
- [13] Chen Jin, Zhuo Li, Shi Peijun, et al. The study on urbanization process in China based on DMSP/OLS data: Development of a light index for urbanization level estimation. *Journal of Remote Sensing*, 2003, 7(3): 168-175. [陈晋, 卓莉, 史培军, 等. 基于DMSP/OLS数据的中国城市化过程研究: 反映区域城市化水平的灯光指数的构建. *遥感学报*, 2003, 7(3): 168-175.]
- [14] Welch R. Monitoring urban population and energy utilization patterns from satellite data. *Remote Sensing of Environment*, 1980, 9: 1-9.
- [15] Masek J G, Lindsay F E, Goward S N. Dynamics of urban growth in the Washington DC metropolitan area, 1973-1996, from Landsat observations. *International Journal of Remote Sensing*, 2000, 21: 3473-3486.
- [16] Mou Fengyun, Zhang Zengxiang, Chi Yaobin, et al. Dynamic monitoring of built-up area in Beijing during 1973-2005 based on multi-original remote sensed images. *Journal of Remote Sensing*, 2007, 11(2): 257-268. [牟凤云, 张增祥, 迟耀斌, 等. 基于多源遥感数据的北京市1973-2005年间城市建成区的动态监测与驱动力分析. *遥感学报*, 2007, 11(2): 257-268.]
- [17] Wu Hongan, Jiang Jianjun, Zhou Jie, et al. Dynamics of urban expansion in Xi'an City using Landsat TM/ETM+ data. *Acta Geographica Sinica*, 2005, 60(1): 143-150. [吴宏安, 蒋建军, 周杰, 等. 西安城市扩张及其驱动力分析. *地理学报*, 2005, 60(1): 143-150.]
- [18] Cai Dong, Li Manchun, Chen Zhenjie, et al. A method of division of urban fringe based on message entropy: A case study in Nanjing city. *Science of Surveying and Mapping*, 2010, 35(3): 106-109. [蔡栋, 李满春, 陈振杰, 等. 基于信息熵的城市边缘区的界定方法研究: 以南京市为例. *测绘科学*, 2010, 35(3): 106-109.]
- [19] Cheng Jicheng, Guo Donghua, Shi Wenzhong. *The Uncertainty of Remote Sensing Data*. Beijing: Science Press, 2004. [承继成, 郭华东, 史文中. *遥感数据的不确定性问题*. 北京: 科学出版社, 2004.]
- [20] Wang Fahui, Jin Fengjun, Zeng Guang. Analyzing regional density functions and growth patterns in China with a GIS-based method delineating influential regions of cities. *Geographical Research*, 2004, 23(1): 97-103. [王法辉, 金凤君, 曾光. 区域人口密度函数与增长模式: 兼论城市吸引范围划分的GIS方法. *地理研究*, 2004, 23(1): 97-103.]
- [21] Zhou Yixing, Shi Yulong. Toward establishing the concept of physical urban area in China. *Acta Geographica Sinica*, 1995, 50(4): 289-301. [周一星, 史育龙. 建立中国城市的实体地域概念. *地理学报*, 1995, 50(4): 289-301.]
- [22] Mao Qizhi, Long Ying, Wu Kang. Spatio-temporal changes of population density and exploration on urbanization pattern in China: 2000-2010. *City Planning Review*, 2015, 39(2): 38-43. [毛其智, 龙瀛, 吴康. 中国人口密度时空演变与城镇化空间格局初探: 从2000年到2010年. *城市规划*, 2015, 39(2): 38-43.]
- [23] Imhoff M L, Lawrence W T, Stutzer D C, et al. A technique for using composite DMSP/OLS "City Lights" satellite data to map urban areas. *Remote Sensing of Environment*, 1997, 61(3): 361-370.
- [24] Cheng F, Thiel K H. Delimiting the building heights in a city from the shadow in a panchromatic SPOT-image (Part 1). Test of forty-two buildings. *Remote Sensing*, 1995, 16(3): 409-415.
- [25] Cui Gonghao, Wu Jin. The spatial structure and development of Chinese urban fringe. *Acta Geographica Sinica*, 1990, 45(4): 399-411. [崔功豪, 武进. 中国城市边缘区空间结构特征及其发展. *地理学报*, 1990, 45(4): 399-411.]
- [26] GuChaolin, Chen Tian, Ding Jinhong, et al. The study of the urban fringes in Chinese megalopolises. *Acta Geographica Sinica*, 1993, 48(4): 317-328. [顾朝林, 陈田, 丁金宏, 等. 中国大城市边缘区特性研究. *地理学报*, 1993, 48(4): 317-328.]
- [27] Yu Wei, Song Jinping, Mao Xiaogang. Review on identification of connotation and spatial definition of urban fringe. *Areal Research and Development*, 2011, 39(5): 55-59. [于伟, 宋金平, 毛小岗. 城市边缘区内涵与范围界定述评. *地域研究与开发*, 2011, 30(5): 55-59.]
- [28] Wang Yuanzhuo, Jin Xiaolong, Cheng Xueqi. Network big data: Present and future. *Chinese Journal of Computers*, 2013, 36(6): 1125-1138. [王元卓, 靳小龙, 程学旗. 网络大数据: 现状与展望. *计算机学报*, 2013, 36(6): 1125-1138.]
- [29] Zhao W F, Li Q Q, Li B J. Extracting hierarchical landmarks from urban POI data. *Journal of Remote Sensing*, 2011, 15(5): 973-988.
- [30] Silverman B W. *Density Estimation for Statistics and Data Analysis*. London, UK: Chapman and Hall, 1986.
- [31] Wu Zhifeng, Li Yuehui, Bu Rencang, et al. Evaluation and simulation of historical range of variability of forest landscape pattern in Huzhong area. *Acta Ecologica Sinica*, 2013, 33(15): 4799-4807. [吴志丰, 李月辉, 布仁仓, 等. 呼中林区森林景观的历史变域模拟及评价. *生态学报*, 2013, 33(15): 4799-4807.]



- [32] Zhang Xun, Zhong Ershun, Zhang Xiaohu, et al. Spatial distribution and clustering of commercial network in Beijing during 2004-2008. *Progress in Geography*, 2013, 32(8): 1207-1215. [张珣, 钟耳顺, 张小虎, 等. 2004-2008年北京城区商业网点空间分布与集聚特征. *地理科学进展*, 2013, 32(8): 1207-1215.]
- [33] Wang Bo, Zhen Feng. An analysis of spatial distribution of urban employment in Nanjing. *Human Geography*, 2011, 26(4): 58-65. [王波, 甄峰. 南京市就业空间布局研究. *人文地理*, 2011, 26(4): 58-65.]
- [34] Xue Dongqian, Liu Hong, Ma Beibei. Characteristics of spatial distribution of cultural industries in urban area of Xi'an city, China. *Scientia Geographica Sinica*, 2011, 31(7): 775-780. [薛东前, 刘虹, 马蓓蓓. 西安市文化产业空间分布特征. *地理科学*, 2011, 31(7): 775-780.]
- [35] Heidenreich N B, Schindler A, Sperlich S. Bandwidth selection for kernel density estimation: A review of fully automatic selectors. *Advances in Statistical Analysis*, 2013, 97(4): 403-433.
- [36] Hinneburg A, Keim D A. An efficient approach to clustering in large multimedia databases with noise//Proc. of the 4th International Conference on Knowledge Discovery and Data Mining. AAAI Press, 1998: 58-65.

## A novel method for identifying the boundary of urban built-up areas with POI data

XU Zening<sup>1,2</sup>, GAO Xiaolu<sup>1</sup>

(1. Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** The boundary of urban built-up areas provides foundational information for urban studies and meets the requirements for urban pattern and urban spatial structure research. However, commonly used methods for identifying the boundary of urban built-up areas such as using remote sensing data of night-light and land use, cadastral data, and building coverage data, are limited in accuracy. To remedy this, this paper proposes to use POI (Point of Interest) data obtained from web maps, assuming that it well reflects the agglomeration of urban activities at higher precision. Based on the underlying connection between POI and the spatial distribution of urban activities, a new method called 'Densi-Graph' is proposed to identify the actual boundary of urban built-up areas with the contour map of the kernel density of POI, where the threshold value for the contour lines to make significant change from densely to loosely placed is picked, giving the boundary of urban and rural areas. Different contour structures for mono-centric, poly-centric and linear cities are discussed, whereby the Densi-Graph method using POI data is validated. The method is also used to study the boundaries of urban built-up areas in China's prefecture level cities. The relationships between the Density-Graph thresholds and the population and location of different cities are discussed. This study advances previous studies in presenting more reliable and objective data on the boundary of urban built-up areas.

**Keywords:** boundary; urban built-up areas; POI; Densi-Graph; kernel density analysis; urban-rural spatial structure