

引用格式:曹元晖,刘纪平,王勇,等.基于POI数据的城市建筑功能分类方法研究[J].地球信息科学学报,2020,22(6):1339-1348. [Cao Y H, Liu J P, Wang Y, et al. A study on the method for functional classification of urban buildings by using POI data[J]. Journal of Geo-information Science, 2020,22(6):1339-1348. ] DOI:10.12082/dqxxkx.2020.190608

## 基于POI数据的城市建筑功能分类方法研究

曹元晖<sup>1</sup>,刘纪平<sup>1,2\*</sup>,王勇<sup>1</sup>,王良杰<sup>3</sup>,吴文周<sup>4</sup>,苏奋振<sup>4</sup>

1. 中国测绘科学研究院,北京 100830; 2. 河南省科学院地理研究所,郑州 450052; 3. 清华大学环境学院,北京 100084;  
4. 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

### A Study on the Method for Functional Classification of Urban Buildings by Using POI Data

CAO Yuanhui<sup>1</sup>, LIU Jiping<sup>1,2\*</sup>, WANG Yong<sup>1</sup>, WANG Liangjie<sup>3</sup>, WU Wenzhou<sup>4</sup>, SU Fenzhen<sup>4</sup>

1. Chinese Academy of Surveying and Mapping, Beijing 100830, China; 2. Institute of Geographical Sciences, Henan Academy of Sciences, Zhengzhou 450052, China; 3. School of Environment, Tsinghua University, Beijing 100084, China; 4. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

**Abstract:** As the carrier of human activity and social development, buildings are the most important geographical entities that constitute the spatial structure of a city. It is one of the urgent tasks in the construction of smart cities in China to build elaborate digital models of urban buildings. Classifying a large amount of buildings by their functions facilitates urban functional area division and urban spatial cognition, thus assisting the government in population estimation, land management, urban planning, and smart city construction. In this paper, POI (Point of Interest) with rich semantic information including name, address, and types was used as the main data source, because it was more accessible and updated more frequently than the traditional geographic information data. The process of finding out the functional type of a building was similar with identifying urban functional areas by using POI data, but there existed the problem of low classification rate due to the sparsity of POI. Therefore, to improve the traditional quantitative identification of urban functional areas, this study attempted to calculate the weighted frequency density ratio of each type of POIs inside and within a certain range around a building. Experimenting on more than 5000 buildings near South Shawa Bridge in the west of Beijing, the study found that 93.04 percent of the buildings were effectively classified into different functional types: residential, commercial, public service, and other three mixed types. The classification rate has been greatly improved compared with that of the traditional method. These classified buildings showed the spatial distribution of functional areas more clearly and precisely than blocks used in identifying urban functional areas, since too many multi-functional blocks with very limited practical meaning were identified by using the traditional method. In order to calculate the classification accuracy, more than 2000 randomly selected buildings were manually divided into functional classes with the assistance of POI and AOI data. The overall classification

收稿日期:2019-10-17;修回日期:2019-12-09.

基金项目:国家重点研发计划项目(2017YBF0503601、2017YFB0503502);中国测绘科学研究院基本科研业务费项目(AR1904)。[ **Foundation items:** National Key Research and Development Program of China, No.2017YBF0503601, 2017YFB0503502; The Basal Research Fund of Chinese Academy of Surveying and Mapping, No.AR1904. ]

作者简介:曹元晖(1995—),女,安徽合肥人,硕士生,从事空间数据分析与挖掘研究。E-mail: caoyuanhui17@mails.ucas.ac.cn

\*通讯作者:刘纪平(1967—),男,陕西高陵人,博士,研究员,主要从事政府地理信息系统研究和应用。E-mail: liujp@casm.ac.cn

accuracy reached 91.18 percent compared with the manually classified result. The classification error was mainly caused by the shortage of POI and the poor data quality, which could be avoided by merging multi-source POI to improve the data quality or applying various Internet location information, such as the social media data and the real estate transaction data. However, by using easily accessible web POI data, the proposed method, which can replace manual classification in an automated way, has greatly improved the effectiveness of classifying large number of buildings into different functional types, and shown higher accuracy than existing researches.

**Key words:** the functional classification of building; Point of Interest; frequency density; inverse distance weight; urban buildings; urban planning; urban functional districts; spatial distribution

**\*Corresponding author:** LIU Jiping, E-mail: liujp@casm.ac.cn

**摘要:**对建筑物进行建模与分析是智慧城市建设的重要任务之一。将城市中数量庞大的建筑物按功能分类,辅助认知城市内部空间结构,对政府部门开展人口估计,土地管理,城市规划等工作具有重要意义。本文以蕴含丰富语义信息的兴趣点(POI, Point of Interest)作为主要信息源,针对POI分布稀疏导致大量建筑物无法识别出功能的问题,改进了传统的城市功能区定量识别方法。该方法为建筑物内部及周边一定区域范围内的POI赋予反距离权重,通过计算不同类型POI的加权频数密度比例来识别建筑物功能类型。文中以北京市西四环中路附近5000多栋建筑物为例进行实验验证,实现了将目标区域内的建筑物按功能类型划分为居住、商业、公服和3种混合类型,识别率达93.04%,与人工判别的结果对比得出总体分类精度达91.18%。该方法采用易于获取的互联网POI数据,可以实现大范围建筑物功能类型的快速自动化识别,丰富了城市建筑模型语义属性,扩展了POI数据的应用范围。

**关键词:**建筑功能分类;兴趣点;指标频数密度;反距离加权;城市建筑物;城市规划;城市功能区;空间分布

## 1 引言

随着城镇化与智慧城市建设的不断推进,精细化的城市规划和管理需要及时、全面地监测与分析城市内部空间结构的分布和变化。目前城市的规划与管理通常以区县、街道等行政区划为单位,也有一些研究以规则格网、路网对区域进一步细化分割,但随着城市内部空间结构的日益复杂化,大比例尺下的应用场景需要更精细的空间单元。楼宇建筑是人类活动的主要场所,经济发展的重要载体,作为城市空间认知的基本对象,筑成了城市的形态结构<sup>[1-2]</sup>。建筑信息的采集、管理、建模和共享,在城市规划<sup>[3]</sup>、土地利用<sup>[4]</sup>、人口估计<sup>[5]</sup>和建筑能耗计算<sup>[6]</sup>等实际应用中发挥着重要作用,已成为近年来的研究热点<sup>[7]</sup>。随着测绘技术的进步,通过高分影像提取<sup>[8]</sup>、三维激光扫描<sup>[9]</sup>等方法可以高效获取大范围建筑物的几何与位置信息,批量生产建筑物的平面和立体数据,但上述方法均无法获取建筑物的语义属性,在实际生产过程中,楼宇建筑的名称、类型等属性信息主要靠人工进行采集和管理,存在成本高、效率低、更新慢等问题。

目前,已有不少方法用于解决建筑物分类问题,如用支持向量机<sup>[10]</sup>、随机森林<sup>[11]</sup>等机器学习方法,根据建筑的形状、光谱、空间、纹理信息识别别

墅、公寓、排屋或城中村等住宅类型。这些研究大多采用LiDAR点云、高光谱遥感影像等数据,利用遥感数据通常可以实现根据地物的物理特征区分城市中水体、绿地、道路、建筑等土地利用类型或按建筑材料、形状区分建筑物类别,但从中很难进一步获取建筑物的社会经济属性,因此这些研究均没有对建筑的功能分类进行讨论。聚焦建筑物功能分类的研究目前较少,Smith等<sup>[12]</sup>的研究中将伦敦部分地区的建筑物划分为居住、零售、办公、教育等12种类型,依据的是土地利用、不动产交易、住宅类型等社会经济数据,这类数据在国内难以直接获取。相比较而言,曲畅等<sup>[13]</sup>采用的POI数据数据获取成本低,更新快,是目前城市研究中被广泛使用的信息来源,通过对不同类型的POI做核密度分析,根据建筑物内各类型的核密度值判断建筑物功能也取得了较准确的分类结果。

POI(Point of Interest)也称兴趣点,可用于指代各种类型的地理实体,因其具有语义丰富,获取周期短,成本低,更新速度快,客观性和现势性强的优势<sup>[14]</sup>,目前已有许多研究以POI为数据源进行城市中心<sup>[15]</sup>、地标<sup>[16]</sup>和功能区<sup>[17]</sup>的识别。其中,利用POI识别城市功能区的方法可以为建筑物功能分类提供思路,因为建筑物的功能类别与城市功能区的分类体系相似,对建筑功能进行划分可以反映出城市

中功能区的分布<sup>[13]</sup>,功能类型相似的建筑聚集分布则形成了城市功能区。但两者的不同之处在于,现有的城市功能区划分方法通常先将城市分割成规则格网、街区构成的基本单元<sup>[18-19]</sup>,再对空间单元内各种地理实体的功能属性进行概括,识别出的混合功能单元居多,但对混合功能类型的定义又缺乏规范性;建筑物功能分类以具有实际意义的建筑范围作为空间单元,描述区域功能分布的空间粒度更细,除了高层建筑可能具有多种功能类型外,多数建筑物通常具有明确的功能语义。

本文对现有的城市功能区识别方法进行移植和改进,提出利用建筑范围内部及周边不同类型POI的加权频数密度比例来划分建筑物功能类型,解决了POI数据分布不均导致的部分建筑物内没有POI作为分类依据的问题,实现了大规模建筑物快速自动分类。

## 2 建筑物功能分类方法

### 2.1 频数密度比例法

POI可以代表各种地理实体,一栋建筑应有其对应的POI点,建筑中具有特定功能的实体,如医院里的科室、商场中的店铺、写字楼中的公司等也应具有对应的POI点。对于每一栋建筑,依据空间范围筛选出落入该建筑范围内的POI,就可以根据这些POI的类型构建指标频数密度(Frequency Density, FD)和类型比例(Category Ratio, CR)来判断建筑物的功能。

$$F_i = \frac{n_i}{N_i} \quad (i=1, 2, 3, \dots, n) \quad (1)$$

$$C_i = \frac{F_i}{\sum_{i=1}^n F_i} \quad (i=1, 2, 3, \dots, n) \quad (2)$$

式中: $i$ 表示类型; $n$ 表示类型的数量; $n_i$ 表示建筑物内第 $i$ 类POI的数目; $N_i$ 表示POI数据集中第 $i$ 类POI的总数; $F_i$ 表示第 $i$ 类POI占该类POI的频数密度; $C_i$ 表示第 $i$ 类POI的频数密度占建筑物内所有类型POI频数密度的比例。

计算出各类型的频数密度比例后,现有研究多以0.5为阈值<sup>[19-20]</sup>,将大于0.5的类型作为建筑物类型,若所有类型的CR值均小于0.5则判为混合类型。这种方法虽然最终呈现的划分结果较为准确,但在特定情况下会产生错误的划分结果,如当类型为 $C_1=0$ 、 $C_2=0.49$ 、 $C_3=0.51$ 时,若以0.5为划分标准,

则该建筑会被判断为第3类,但从CR指标来看,判断该建筑为第2、3类组成的混合类型更为合理。因此本文提出将各类型CR指标的均值和标准差作为类型划分的依据,判断的过程如下:

(1)计算各类指标的均值 $E$ 与方差 $S$ ;

(2)若 $E=0$ ,则该建筑物的类型为空,否则按从小到大的顺序排列各指标为 $v_1, v_2, v_3, \dots, v_n$ ;

(3)用格拉布斯法判断最大值是否偏离均值 $E$ 很远,计算最大值和最小值的 $G_i$ 值,公式为 $G_i=|v_i-E|/S$ ,然后与查表获得的临界值 $GP(i)$ 比较;

(4)若 $G_i \geq GP(i)$ ,则认为建筑物类型为最大值 $v_i$ 对应的类型;

(5)若 $G_i < GP(i)$ ,则认为建筑物类型为混合类型,找出不小于均值 $E$ 的指标值 $v_m-v_i$ ,认为建筑物类型为由 $v_m-v_i$ 对应类型组成的混合类型。

格拉布斯法<sup>[21]</sup>是一种检验异常值的方法,通过一组数中某一个值的残差和这组数的标准差的比值来判断该值是否与其他值存在较大差异。应用到本文的类型划分问题中,对于单一功能的建筑物,其对应类型的频数密度比例应明显高于其他类型的比例,否则该建筑物为混合功能类型。

### 2.2 反距离加权的频数密度

POI数据大多为众包模式生产或由网络爬虫采集,数据质量通常难以保证,具体表现为数据丢失,位置偏移、属性不规范等;此外,POI数据主题性强,在空间中并非均匀分布,如在邻近道路和大型商业综合体的地方,POI的分布通常较为密集,而在占地面积较大的工业园区、校园、小区等地,POI的分布则较为稀疏。以上2个方面原因往往导致许多建筑物范围多边形内没有POI的分布,此时传统的频数密度比例方法无法进行建筑功能类型的划分。

文献[13]采用核密度的方法,先计算每类POI的核密度分布,再根据建筑范围内每类核密度的值划分建筑类型。这种方法存在一定的局限性,①核密度分析时的带宽选择和划分类型时核密度阈值的选择都需要经过多次实验才能取得最优结果;②建筑物的类型是离散型而非数值型的属性,虽然存在相同类型的建筑聚集分布的情况,但并非所有类型的建筑附近出现同类型建筑的概率就一定很大,例如一条道路两侧的建筑类型可能完全不同,而上述核密度方法难以顾及这样的分异性。

核密度方法本质上利用了建筑物周围邻近的POI来推测该建筑物的类型,依据这一思想,本文将



参与计算指标频数密度的POI范围扩展到建筑物所在的兴趣区(Area of Interest, AOI)和街区(Block)<sup>[12]</sup>。为密度频数比例方法中的密度频数 $F_i$ 赋予反距离权重(式(3)),即距离建筑物越近的POI属于该建筑的概率越大。

$$F_i = \frac{\sum_{j=0}^m w_j}{N_i} \quad (i=1, 2, 3, \dots, n; j=0, 1, 2, \dots, m) \quad (3)$$

式中: $m$ 为范围内第 $i$ 类的POI的数量,权重 $w_j$ 的计算采用以POI至建筑物距离为自变量的一维高斯函数(式(4))。反距离加权中常用倒数形式的反函数,这种方法会为近邻分配很大的权重,随距离的增加权重迅速降低,这使算法对噪声数据十分敏感,不适用于不确定性强的POI数据,因此本文选用曲线更为平滑的高斯函数。

$$f(x) = ae^{-\frac{(x-b)^2}{2c^2}} \quad (4)$$

高斯函数的一维图表现为钟状曲线(Bell Curve), $a$ 为峰值, $b$ 为峰值处的自变量值,同时 $x=b$ 也是钟型的对称轴, $c$ 为标准方差,表征钟型的宽度。本文中自变量 $x$ 的实际意义为POI点到建筑物平面范围多边形几何中心的距离,当POI接近建筑物时权重应接近1,POI距建筑物太远时权重接近0但不为负,因此令 $a=1, b=0$ ;对于范围内没有POI的建筑物,本文对其100 m缓冲区内的POI的类别进行统计和计算,以判断建筑功能类型,根据图1中不同 $c$ 取值的函数图像,本研究取 $c=30$ 的曲线作为距离权重函数。

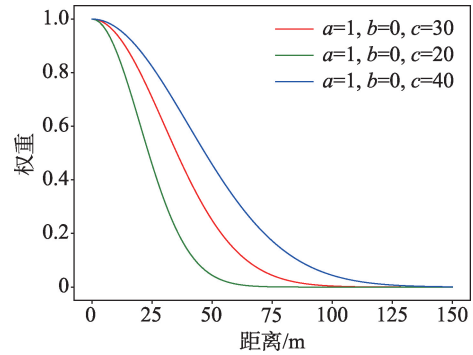


图1 距离权重函数

Fig. 1 Inverse distance weight function graph

### 3 研究区概况与数据处理

#### 3.1 研究区概况及数据源

本文研究区域为北京市西四环中路附近区域,如图2所示。采用的数据包括2019年5月通过深网爬虫<sup>[22]</sup>由百度地图API(<http://lbsyun.baidu.com/index.php?title=webapi>)获取的5552条建筑物平面范围多边形数据、255条AOI数据、12 189条POI数据,由腾讯地图API([https://lbs.qq.com/webservice\\_v1/index.html](https://lbs.qq.com/webservice_v1/index.html))获取的443条POI数据,以及从OpenStreetMap下载的705条道路数据<sup>[23]</sup>。道路数据中多数道路具有名称、类型和最大速度等属性;建筑物多边形数据不具有任何属性信息,仅包含建筑物的平面位置和形状;AOI数据多数为门址(xx路xx号院)范围多边形,以及个别公园与购物中心;

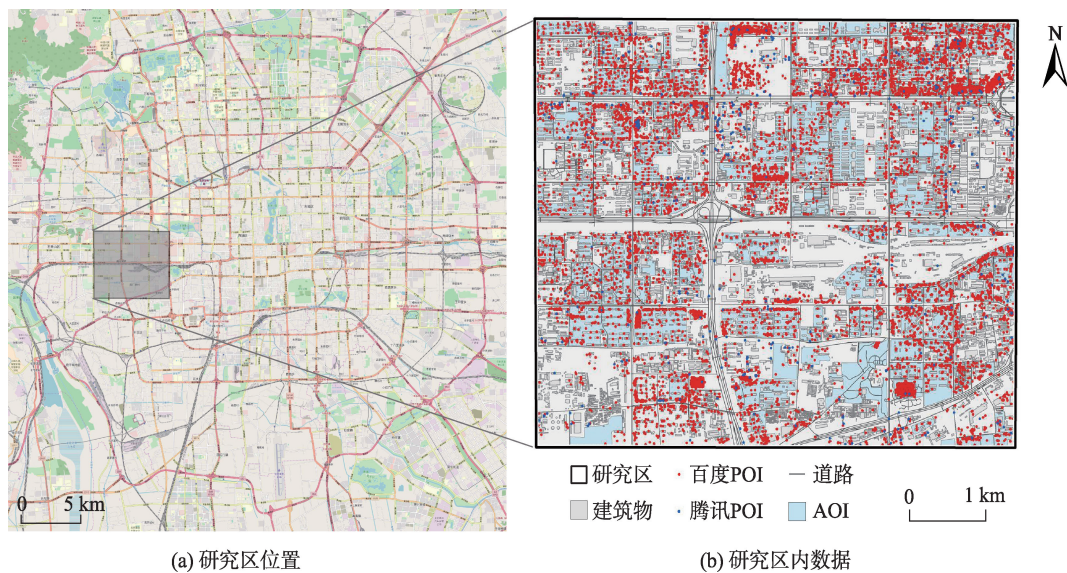


图2 2019年北京市南沙窝桥附近部分区域的POI、AOI、道路与建筑物数据

Fig. 2 POI, AOI, road and building data of the study area near the South Shawo bridge in Beijing in 2019



POI则包含名称、地址、位置和类型信息。从图2可看出,POI大多沿道路两侧分布,且在一些区域分布极为密集,而在另一些区域分布特别稀疏。

### 3.2 数据预处理

虽然POI数据蕴含了丰富的空间和语义信息,但数据质量较差,存在位置偏移、语义不规范、不确定性等问题,难以直接利用。本文主要利用POI的类别属性,由于POI的类别过多,需要对其进行剔除和归并处理。从图3的POI类型中可以看出,POI可以指代一切具有位置的实体,这些实体空间范围有大有小,认知度有高有低,并非所有POI都能为建筑功能推测提供有效信息,甚至造成干扰,需先将空间粒度小,公众认知度低的点,如公厕、公交站、报刊亭等,从原始数据中剔除。接着对剩余的POI点按建筑功能类型重分类,参考《土地利用现状分类标准(GB/T 21010-2017)》<sup>[24]</sup>,本文将建筑物按功能分为公共、居住和商业3大类。将各个类型的POI归并入建筑功能的3大类中,重分类后的POI数据如图4所示。

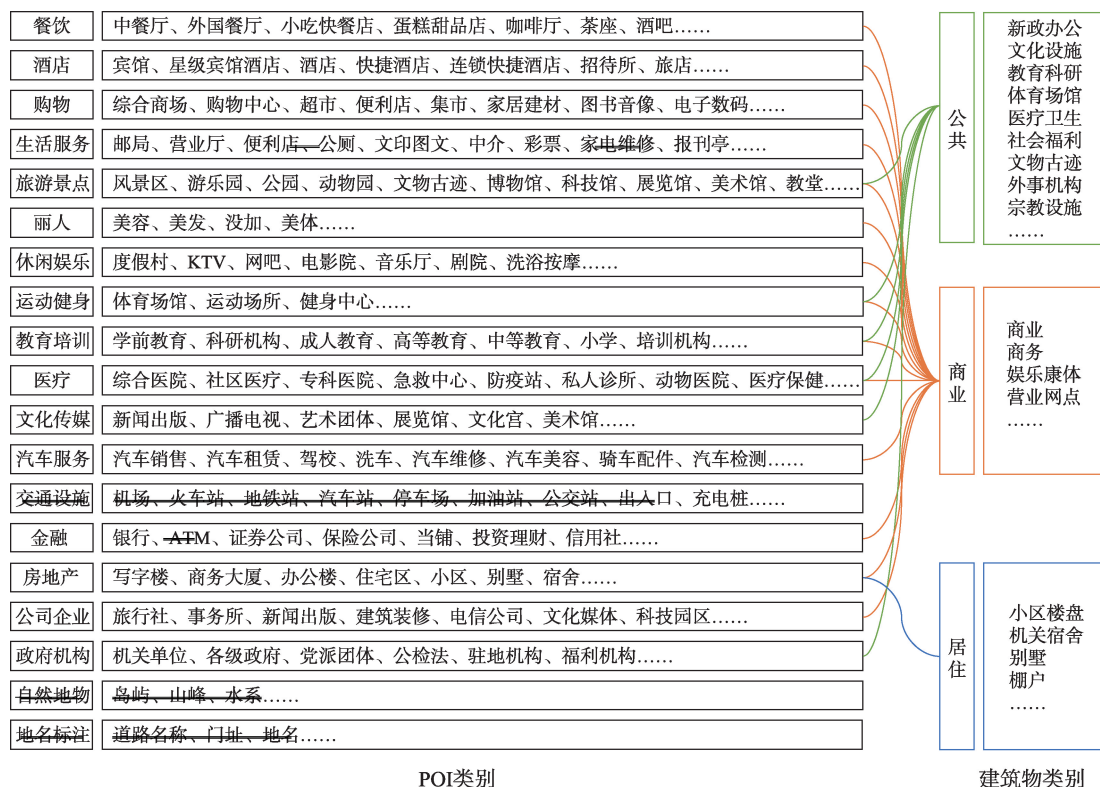
从图4可看出,许多建筑物多边形内没有POI

分布,根据本文提出的方法,参与分析建筑物类型的POI分布范围扩展到了建筑物所在AOI或街区内,因此,还需准备研究区域内的区块数据。选取研究区域内的OpenStreetMap道路数据,从中删除等级较低的道路,如living\_street, residential, foot-way等,然后在ArcMap中用线状道路分割研究区域,形成街区块,再用AOI数据对街区块进行分割,结果如图5所示。

## 4 结果与分析

### 4.1 建筑分类结果

根据本文提出的建筑物功能分类方法,研究区域内建筑物的功能分类结果如图6(a)所示。由于本文方法将建筑物周围邻近的POI点纳入建筑物类型指标的计算,有效避免了POI分布稀少地区建筑物功能难以识别的问题,从而大大提高了识别率。若依照传统频数密度比例法,研究区域的建筑物功能识别结果如图7(a)所示,识别率仅为34.18%,大量分布在POI低密度地区的建筑物因内部没有样点而无法识别类型。对传统方法进行改



注:双删除线表示在数据预处理时将这类别的POI数据从原始数据集中剔除。

图3 POI与建筑物类别映射

Fig. 3 Category map between POIs and buildings

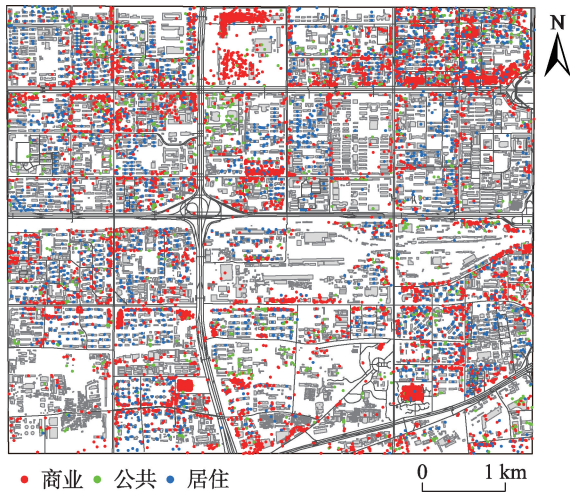


图4 研究区内3类POI的空间分布  
Fig. 4 Spatial distribution of three types of POIs in the study area

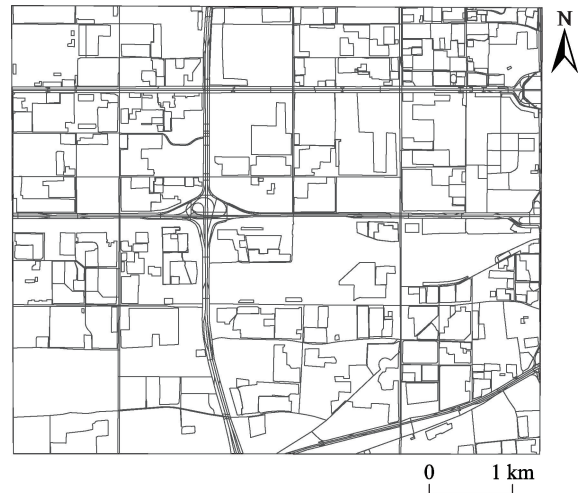
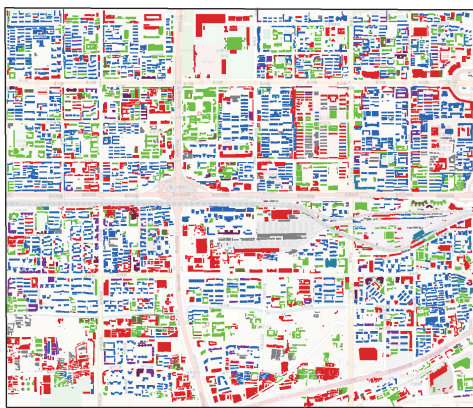
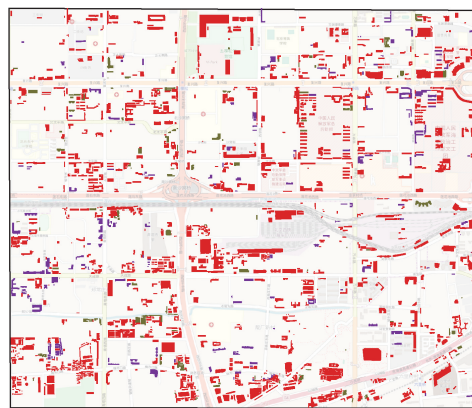


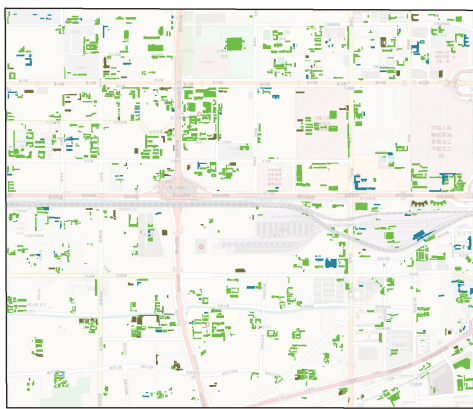
图5 用道路与AOI划分研究区形成的空间单元  
Fig. 5 Space units formed by dividing the study area with roads and AOIs



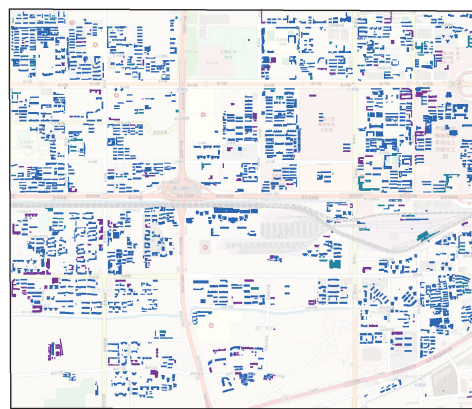
(a) 建筑物功能分类结果



(b) 商业类建筑空间分布



(c) 公共类建筑空间分布



(d) 居住类建筑空间分布

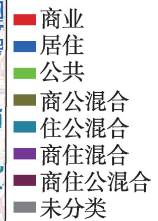


图6 2019年北京市南沙窝桥附近区域5552栋建筑物的功能分类结果与3大类建筑物的空间分布  
Fig. 6 Functional classification results of 5552 buildings in the area near the South Shawo Bridge in Beijing in 2019 and the spatial distribution of three major types of buildings

进后,研究区域内的5222栋建筑共有5166栋被识别出功能类型,识别率达93.04%,与传统方法相比

提高了58.86%,弥补了传统方法在POI分布稀少的区域,如城郊、居住区等地识别率低的问题。实验



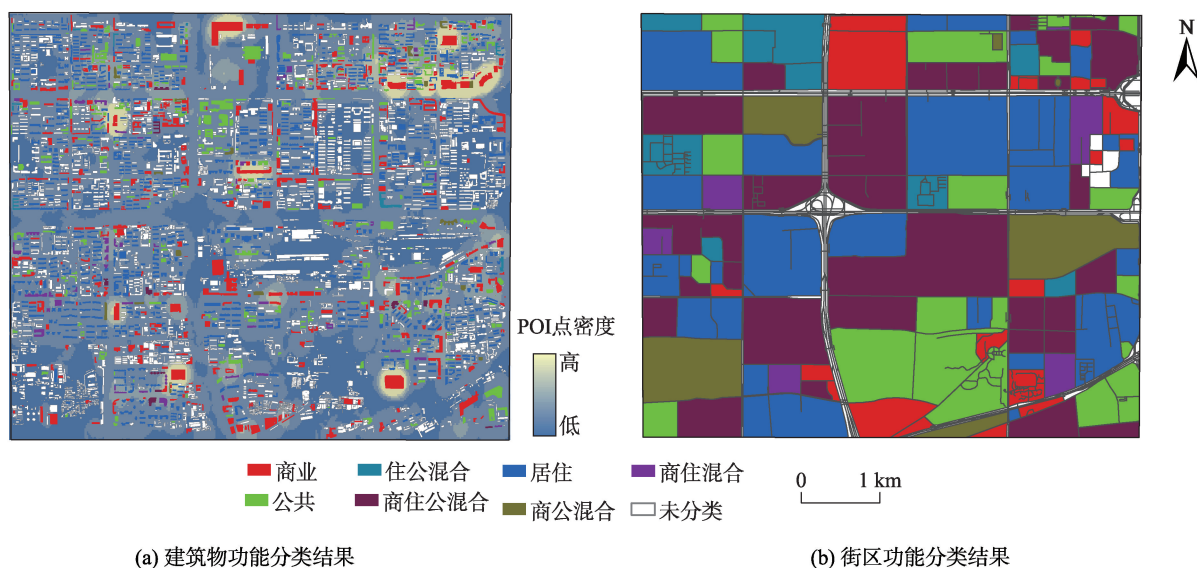


图7 采用传统频数密度比例法的研究区建筑物与街区功能分类结果

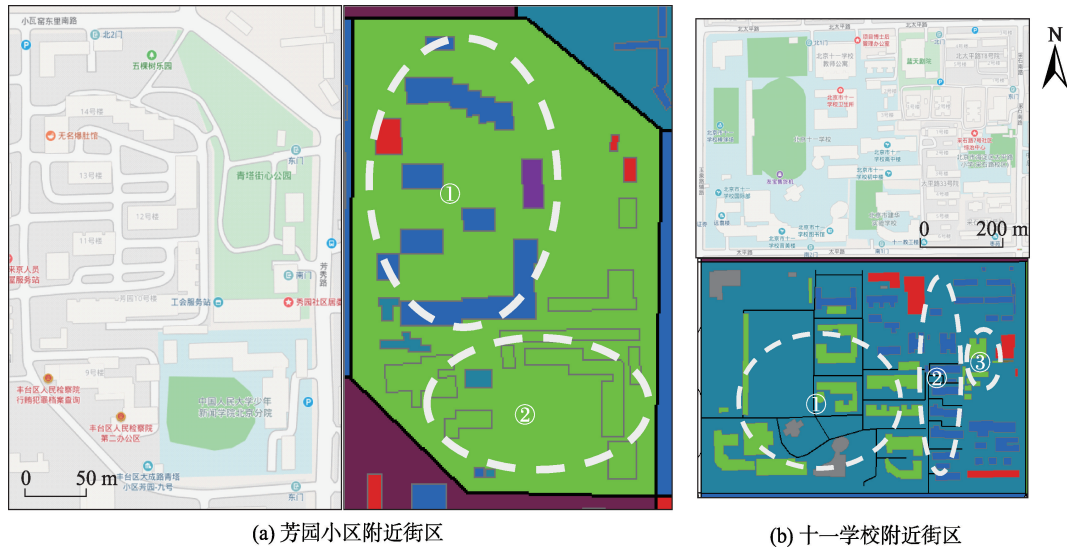
Fig. 7 Functional classification results of buildings and blocks in the study area using the traditional frequency density ratio method

中识别出的建筑物按功能分为居住、商业、公共和混合类型,分别占建筑物总数的26.40%,36.35%,20.23%和10.06%。从分类中可以看出,商业类建筑物(图6(b))大多分布在道路附近,对道路做50 m和100 m的缓冲区分析,并分别与商业类建筑物叠加,发现约50%商业类建筑分布在道路的50 m缓冲区内,约70%的建筑分布在道路的100 m缓冲区内,且呈现出沿道路的条带状分布规律,这与城市中商业设施通常分布在交通可达性较好的地区的规律相符<sup>[25]</sup>;居住类的建筑(图6(d))多数在街区内部呈大片的组团聚集分布,符合我国大部分城市居住小区形式的居住空间聚集分布特点<sup>[26]</sup>;公共类的建筑(图6(c))除了医院、学校等较大规模的公共服务机构呈现组团模式的聚集分布之外,更多地呈现较为零散的分布特点,主要因为多数行政办公场所和文体设施通常分布在独栋建筑中,且占地范围较小。

将图6所示本文方法的建筑物分类结果和图7(b)所示的采用传统方法识别的街区块功能类型划分结果对比分析,建筑物分类结果在表现区域功能时的空间粒度更细,混合功能的空间单元更少,并且体现出了街区块内部的功能分异,在大比例尺下反映了更丰富的空间和语义信息。图8中以2个街区为例,比较百度地图、按街区块划分的城市功能区和建筑物功能分类结果对同一区域的不同描述,展示街区内部由建筑物功能体现的街区内部的功能空间分异。图8(a)是芳园小区内的一个街区,

从百度地图中可看,出该街区主要由北部的居住区、公园和南部的政府机构、学校组成,传统的城市功能区识别方法将该街区识别为公共功能区,而建筑物功能分类的结果可以区分街区北部的居住功能(①)和南部的公共功能(②)。图8(b)是北京十一学校所在街区,该街区被分为住-公混合功能区,而建筑物分类结果可以体现该街区由西向东按照公共①—居住②—公共③的功能分异。从图6可看出,城市中不同的功能实体在空间上是相互耦合的,例如居住小区内部通常有配套的活动中心、便民商店等,而当空间单元的范围较大时,各个单元的功能趋于混合,难以体现差异性。因此,从精细化空间管理的角度看,以建筑物为描述区域功能的基本空间单元更能准确反映区域的功能特征。以建筑物功能分类的结果为基础,将功能类型相似的建筑物进行聚类,即可在小比例尺下获得比传统城市功能区划分与识别更加精准的区域功能分布与概括结果。这一过程还可以应用于制图综合中,为制图综合提供语义信息支撑。建筑物功能类型的划分对城市规划具有重要意义,城市通常要经历规划—发展—调整—再规划的循环过程,在这一过程中,建筑物功能分类信息反映了区域的发展现状,可用于评估区域是否按规划发展,为调整区域的功能结构提供参考。另外,建筑物的功能与建筑物内部的人类活动息息相关。在一天中的不同时间段内,不同功能类型建筑物内的人口聚集情况





注:图(a)中①所示区域对应百度地图中芳园小区的建筑;②所示区域对应百度地图中丰台区人民检察院与中国人民大学少年新闻学院北京分院的建筑。图(b)中①所示区域对应百度地图中北京十一学校的建筑;②所示区域对应百度地图中太平路33号院的建筑;③所示区域对应百度地图中太平路小学的建筑。

图8 研究区内部分街区的建筑物功能分类结果与百度地图所示实际情况的对比  
Fig. 8 Comparison of building functional classification results of some blocks in the study area with Baidu map

也会发生规律性的变化,因此建筑物功能分类结果可作为推测人口密度和建筑物能耗的依据,服务于配套设施的规划和紧急情况的应对。

#### 4.2 分类精度评价

从研究区内随机选取2200个建筑进行人工判定类型,利用本文采用的POI和AOI数据辅助判别,并将人工判定的结果与自动分类的结果进行比较,形成自动分类结果的误差矩阵,如表1所示。由误差矩阵计算出总体分类精度为91.18%,较基于核密度的方法<sup>[13]</sup>有一定提高。居住类的预测准确率为90.61%,公共类为92.97%,商业类为94.73%,混合类为87.43%。

将未识别出类型和与人工判别结果不一致的建筑叠加在POI点密度图上,如图9所示,发现未识

别出类型的建筑主要位于POI分布极少的区域,如军事管理区、铁路设施用地等敏感区域;与人工判别结果不一致的建筑约有85%位于POI分布的低密度区域,少数分布在高密度区域的建筑物外部邻近的POI较多而内部POI极少。总体来看POI密度对分类结果影响较大,在POI分布少的区域,建筑物周边没有与之相关的POI,信息支撑弱,所以建筑

表1 与人工判读结果对比的建筑物功能自动分类误差矩阵  
Tab. 1 Automatic classification error matrix of building functions compared with manual classification results

自动分类结果/个	人工判读/个			
	居住	公共	商业	混合
居住	1283	8	2	8
公共	46	304	6	7
商业	50	12	252	9
混合	37	3	6	167

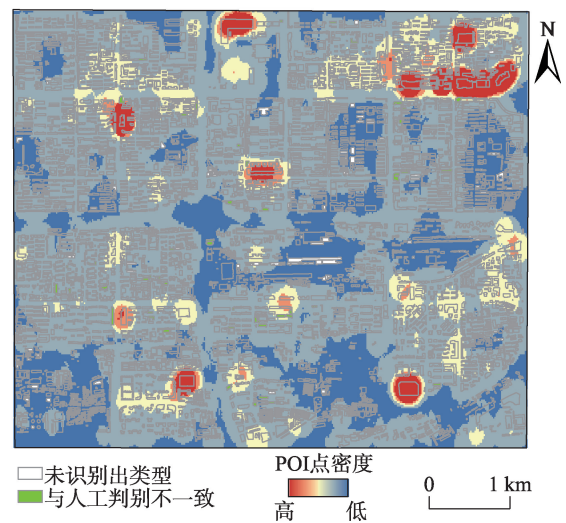


图9 POI密度与建筑物功能识别率和分类精度的关系  
Fig. 9 The relationship between POI density and building function recognition rate and classification accuracy

物的功能类型更容易被错分或无法识别。

此外,在检查与错判建筑物相关的POI点时,发现部分POI数据存在偏离真实位置较远,类别属性错误等问题,从而导致建筑物类型错判,因此,POI数据质量也是影响建筑类型识别准确性的重要因素之一。POI数据的质量可以从位置精度、语义精度、完整性、现势性几个方面来衡量。位置偏移、语义错误为识别建筑功能类型提供错误的线索,若建筑物本身就位于POI稀疏的区域,此类错误会对分类精度产生较大的影响,可通过拓扑检查或多源POI匹配的方法发现并剔除此类POI。POI完整性差是指在数据采集过程中由于web API接口限制等因素导致原本应有的POI数据丢失,这会使建筑物功能类型无法识别的概率提高,同时很容易将混合类建筑误判为单一功能建筑,因此应在数据采集时尽量保证其完整性,或通过融合多源数据来弥补数据的丢失。POI现势性差则会导致无法及时发现功能类型发生变化的建筑。

## 5 结论及展望

本文利用POI数据进行建筑功能类型的划分,对城市功能区识别中常用的指标频数密度比例法做出改进,并通过扩大POI搜索范围和对指标频数密度比例进行反距离加权解决了POI稀疏区域建筑物功能难以识别的问题,以北京市西四环中路附近的5000多栋建筑验证了该方法的可行性和准确性,识别出93.04%的建筑物的功能类型,将建筑物分为居住、商业、公共和3种混合类型,分类精度可达91.18%,识别率和分类精度较传统方法均有所提高。

由于POI数据成本低、易于获取、更新速度快等优势,改进后的指标频数密度比例法能够实现大范围建筑物功能的快速识别,为人工建模、遥感影像提取、倾斜摄影测量等手段获取的具有平面范围的建筑物单体数据赋予功能类型属性,以更精细的空间粒度描述区域的功能布局,从而为政府部门开展城市精细化管理、区域发展现状分析与规划调整、人口估计与应急管理等工作提供信息支撑。

然而POI数据也具有局限性,其数据密度和数据质量对分类的准确性影响较大。因此在今后的研究中,可考虑从以下3个方面提高分类的准确性:

- ① 采用融合多源POI数据的方法提高数据质量;
- ② 引入互联网中其他类型的泛在位置数据综合分析,如社交媒体签到数据、房产交易数据、工商企业数据等;
- ③ 在三维空间内对空间粒度进行进一步细化,如大中城市内的高层建筑中不同楼层往往有不同功能,可在高度上将其分割为多个具有单一功能类型的实体。

### 参考文献(References):

- [1] Hecht R, Meinel G, Buchroithner M. Automatic identification of building types based on topographic databases: A comparison of different data sources[J]. *International Journal of Cartography*, 2015,1(1):18-31.
- [2] 叶娟娟,杨昕,熊礼阳,等.南京市老城区城市建筑点格局研究[J].*地球信息科学学报*,2015,17(11):1404-1411. [Ye J J, Yang X, Xiong L Y, et al. Point pattern analysis of urban buildings in Nanjing old town[J]. *Journal of Geo-information Science*, 2015,17(11):1404-1411. ]
- [3] Yamamura S, Fan L, Suzuki Y. Assessment of urban energy performance through integration of BIM and GIS for smart city planning[J]. *Procedia Engineering*, 2017,180: 1462-1472.
- [4] Yokoya N, Ghamisi P, Xia J, et al. Open data for global multimodal land use classification: Outcome of the 2017 IEEE GRSS Data Fusion Contest[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018,11(5):1363-1377.
- [5] Kunze C, Hecht R. Semantic enrichment of building data with volunteered geographic information to improve mappings of dwelling units and population[J]. *Computers, Environment and Urban Systems*, 2015,53:4-18.
- [6] Evans S, Liddiard R, Steadman P. Modelling a whole building stock: domestic, non-domestic and mixed use[J]. *Building Research & Information*, 2019,47(2):156-172.
- [7] Gröger G, Plümer L. City GML- Interoperable semantic 3D city models[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2012,71:12-33.
- [8] Li M, Stein A, Bijker W, et al. Urban land use extraction from very high resolution remote sensing imagery using a Bayesian network[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016,122:192-205.
- [9] Gilani S, Awrangjeb M, Lu G. An automatic building extraction and regularisation technique using lidar point cloud data and orthoimage[J]. *Remote Sensing*, 2016,8(3):258.
- [10] Henn A, Römer C, Gröger G, et al. Automatic classification of building types in 3D city models[J]. *GeoInformati-*

- ca, 2012,16(2):281-306.
- [11] Huang Y, Zhuo L, Tao H, et al. A novel building type classification scheme based on integrated LiDAR and high-resolution images[J]. *Remote Sensing*, 2017,9(7):679.
- [12] Smith D, Crooks A. From buildings to cities: techniques for the multi-scale analysis of urban form and function [EB/OL]. <https://discovery.ucl.ac.uk/id/eprint/20360/>,2010.
- [13] 曲畅,任玉环,刘亚岚,等.POI辅助下的高分辨率遥感影像城市建筑物功能分类研究[J].*地球信息科学学报*, 2017,19(6):831-837. [ Qu C, Ren Y H, Liu Y L, et al. Functional classification of urban buildings in high resolution remote sensing images through POI-assisted analysis[J]. *Journal of Geo-information Science*, 2017,19(6): 831-837. ]
- [14] 许泽宁,高晓路.基于电子地图兴趣点的城市建成区边界识别方法[J].*地理学报*,2016,71(6):928-939. [ Xu Z N, Gao X L. A novel method for identifying the boundary of urban built-up areas with POI data[J]. *Acta Geographica Sinica*, 2016,71(6):928-939. ]
- [15] 雷英哲,田晶,林繆鹏,等.道路网与兴趣点相结合的城市中心提取方法[J].*测绘学报*,2015,44(S1):42-48. [ Lei Y Z, Tian J, Lin L P, et al. A method for automatic delineation of city centers using POI and road networks[J]. *Acta Geodaetica et Cartographica Sinica*, 2015,44(S1):42-48. ]
- [16] Lin Y, Cai Y, Gong Y, et al. Extracting urban landmarks from geographical datasets using a random forests classifier[J]. *International Journal of Geographical Information Science*, 2019,33(12):2406-2423.
- [17] Yuan J, Zheng Y, Xie X. Discovering regions of different functions in a city using human mobility and POIs[A]. *ACM. Proceedings of the 18<sup>th</sup> ACM SIGKDD international conference on Knowledge discovery and data mining* [C]. 2012.186-194.
- [18] 谷岩岩,焦利民,董婷,等.基于多源数据的城市功能区识别及相互作用分析[J].*武汉大学学报·信息科学版*, 2018,43(7):1113-1121. [ Gu Y Y, Jiao L M, Dong T, et al. Spatial distribution and interaction analysis of urban functional areas based on multi-source data[J]. *Geomatics and Information Science of Wuhan University*, 2018,43(7): 1113-1121. ]
- [19] 邓悦.基于多源兴趣点数据的城市功能区划分方法研究 [D].北京:中国测绘科学研究院,2018. [ Deng Y. Research on the method of dividing urban functional areas based on multi-source points of interest data[D]. Beijing: Chinese Academy of Surveying and Mapping, 2018. ]
- [20] 池娇,焦利民,董婷,等.基于POI数据的城市功能区定量识别及其可视化[J].*测绘地理信息*,2016,41(2):68-73. [ Chi J, Jiao L M, Dong T, et al. Quantitative identification and visualization of urban functional area based on POI data[J]. *Journal of Geomatics*, 2016,41(2):68-73. ]
- [21] 熊艳艳,吴先球.粗大误差四种判别准则的比较和应用 [J].*大学物理实验*,2010,23(1):66-68. [ Xiong Y Y, Wu X Q. The generalizing application of four judging criterions for gross errors[J]. *Physical Experiment of College*, 2010, 23(1):66-68. ]
- [22] 王勇.深网POI信息获取与一致性处理方法研究[J].*测绘学报*,2017,46(3):399. [ Wang Y. Research on crawling and consistency processing of POIs from deep web[J]. *Acta Geodaetica et Cartographica Sinica*, 2017,46(3):399. ]
- [23] OpenStreetMap Data Extracts[DB/OL]. <http://download.geofabrik.de/asia/china.html>, 2019.
- [24] GB/T 21010-2017.土地利用现状分类[S].北京:中国标准出版社,2017. [ GB/T 21010-2017. Current land use classification[S]. Beijing: Standards Press of China, 2017. ]
- [25] 陈晨,王法辉,修春亮.长春市商业网点空间分布与交通网络中心性关系研究[J].*经济地理*,2013,33(10):40-47. [ Chen C, Wang F H, Xiu C L, et al. The relationship between the spatial distribution of commercial networks and street centrality in Changchun[J]. *Economic Geography*, 2013,33(10):40-47. ]
- [26] 邹颖,卞洪滨.对中国城市居住小区模式的思考[J].*世界建筑*,2000(5):21-23. [ Zou Y, Bian H B. Comments on urban residential district pattern in China[J]. *World Architecture*, 2000(5):21-23. ]