

基于DBSCAN空间聚类的广州市区餐饮集群识别及空间特征分析

杨帆¹, 徐建刚^{※2}, 周亮^{3,4}

(1. 南京大学 地理与海洋科学学院, 中国江苏 南京 210023; 2. 南京大学 建筑与城市规划学院, 中国江苏 南京 210093; 3. 中国科学院 地理科学与资源研究所, 中国 北京 100101; 4. 兰州交通大学 测绘与地理信息学院, 中国甘肃 兰州 730070)

摘 要: 选取广州作为研究案例地, 通过百度地图 API 获取广州市区 27 037 个餐饮类 POI 点的空间数据, 在此基础上引入 DBSCAN 空间聚类算法, 将其识别为 397 个集群, 其在空间特征上呈现以天河南集群为主中心、以北京路及江南西两个集群为副中心的“一主两副”空间结构。根据集群的规模划分为 6 个等级, 发现不同等级的集群在数量上符合中心地理论模型, 并随宏观至微观呈现由基于 K=3 的市场原则向基于 K=4 的交通原则的转变。根据紧凑率、延伸度、密度及集中度等空间形态指标, 将集群划分为街道型、片区型、单体—片区型、单体型四类。本研究有助于更好地认识城市餐饮业集聚特征规律, 为深入认识城市实体空间提供支撑。

关键词: DBSCAN 聚类; 餐饮业; 空间特征; 广州

中图分类号 K901 文献标志码 A 文章编号 1000-8462(2016)10-0110-07

DOI: 10.15957/j.cnki.jjdl.2016.10.015

Cluster Identification and Spatial Characteristics of Catering in Guangzhou based on DBSCAN Spatial Clustering

YANG Fan¹, XU Jian-gang², ZHOU Liang^{3,4}

(1. School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, Jiangsu, China; 2. School of Architecture and Urban Planning, Nanjing University, Nanjing 210093, Jiangsu, China; 3. Institute of Geographical Science and Natural Resource Research, CAS, Beijing 100101, China; 4. School of Surveying and Mapping and Geographic Information, Lanzhou Jiaotong University, Lanzhou 730070, Gansu, China)

Abstract: The catering industry is an essential component of the urban service industry. Research on catering clusters is beneficial to grasp its spatial distribution pattern as well as to guide the development and overall arrangement of catering industry. However, little has been done in the quantitative researches related to catering clusters as well as the objective identification method for the cluster. Therefore, in this paper Guangzhou is selected as the location of case study and the spatial data of 27037 catering POIs is acquired by the Baidu Map API. On the basis of these, the DBSCAN spatial clustering algorithm is introduced to identify the data into 397 clusters. The “One main center and Two sub-centers” spatial structure is presented with the main center of Tianhenan cluster and two sub-centers of Beijing Road and Jiangnanxi clusters. According to the scale, the clusters are divided into six levels, and it is found that the clusters of different levels accord with the model of Central Place Theory in the quantitative terms. From the macroscopic view to microcosmic view, it is changed from the marketing principle based on K=3 to the traffic principle based on K=4. In accordance with spatial morphology indicators such as Compactness rate (COR), Elongatedness (ELG), Density (DENS) and Concentration degree (G), the clusters could be divided into four types, including street type, district type, monomer-district type and monomer type. This research is beneficial for the better understanding of cluster features in urban catering industry and provides support for the profound recognition on the urban entity space.

Key words: DBSCAN clustering; catering industry; spatial characteristics; Guangzhou

收稿时间 2016-06-05; 修回时间 2016-07-27

基金项目 国家自然科学基金项目(51278239)

作者简介 杨帆(1991—), 男, 广东清远人, 硕士研究生。主要研究方向为城市数据挖掘、城市与区域规划。E-mail: ficush@163.com。

※通讯作者 徐建刚(1960—), 男, 江苏涟水人, 教授, 博士生导师。主要研究方向为城市与区域规划、数字城市与规划。E-mail: xjg129@sina.com。

餐饮业作为城市服务业不可或缺的组成部分,是标示城市活力的重要指标,餐饮设施点是餐饮业的载体,对其的相关研究有助于掌握餐饮业布局规律、揭示城市发展的空间格局,为深入认识城市实体空间提供支撑,同时对提升城市服务业综合竞争力、辅助餐饮业发展规划与网点布局具有重大的意义。目前国内已有一定的城市餐饮空间格局研究^[1-6],有关研究主要围绕不同区域、不同类型的餐饮业分布情况及其空间分布驱动成因所展开,从方法来看以定性或半定性半定量方法为主,侧重于宏观层面的餐饮空间分布特征描述,鲜有对超大城市餐饮集群现象的定量探讨。餐饮集群是餐饮业空间分布格局的重要组成,对其进行识别与定量研究有助于理解城市的商业地理格局。然而,目前国内有关餐饮集群的定量研究并不多见,国外学者在社区级餐饮集群分布格局^[7]、都市区餐饮集群分布及空间关联^[8]、空间集聚性^[9]与快餐集群与学校的空间关系^[10-11]等方面展开了一定的定量研究,而国内有关餐饮集群的研究多集中在商业体系、品牌营销及微观层次的规划设计等领域,基本没有采用定量方法对餐饮集群乃至其他服务业集群进行识别与分析的相关研究。界定、识别与提取餐饮集群并分析其规模、等级、形态等特征,对餐饮业集群及空间格局研究至关重要,为此亟需在大量餐饮点数据中识别空间集群的定量方法。

餐饮集群是一类产业集群,实质是大量餐饮企业在地理空间上的集聚现象,相比起制造业等产业集群,餐饮集群往往依托于居民消费需求、大型商业中心或旅游景点等与地理区位相关的因素作为其形成的主导原因^[4],产业集群有两个显著特征:产业集聚和产业关联,鉴于餐饮点间的产业关联难以测度,地理集聚可作为识别餐饮业集群最主要的依据,即在地理空间上显著邻近的餐饮业点可被视为一个集群。空间聚类是定量识别产业集聚的高效方法,在地理学领域,传统的基于距离或层次划分的聚类方法等已经得到一定程度的应用,夏云^[12]及林冬云^[13]曾利用k-means聚类对产业集群或企业分布格局进行了空间数据挖掘研究。基于距离的聚类方法的缺点是只适合于发现球状簇。现实中的餐饮集群往往有着多种多样的形态,基于距离的聚类方法对于非球状簇显得无能为力。因此本文引入基于密度的DBSCAN(Density-Based Spatial Clustering of Applications with Noise)聚类算法,其基本思想是将空间中达到一定密度的区域划分为

簇。相比起基于距离的聚类算法,此算法更适合于在空间中发掘任意形状的簇,已在计算机领域相关研究中被用于处理获取自Twitter、Google Maps或基于GPS的空间数据^[14-16],然而在人文地理学上的应用^[17]还相对少见。因此,本文引入该空间聚类算法以识别餐饮点的空间集群,并为后续分析提供支撑。

在利用空间聚类有效识别餐饮集群的基础上,可以对集群的等级结构、空间形态等方面进行分析。餐饮集群的等级结构研究从某种意义上即是对中心地理论的验证和深化^[1],也是商业地理学一直以来的关注焦点,本文经空间聚类识别并提取餐饮集群,可以为其等级结构的定量研究提供充足支撑;同时,虽然国内针对城市餐饮业乃至服务业集群的形态已展开一定的研究^[3,18],但多从微观尺度的规划实践出发并以定性分析为主,本文通过对识别的集群进行几何形态分析,可进一步揭示餐饮业集群的空间特征。广州作为我国餐饮业高度发达的城市,人均餐饮消费在国内大城市排名居首^[19],本土饮食文化源远流长,从而素有“食在广州”的美誉,市区内餐饮设施点数量众多且分布密集。因此本文选择广州市区作为研究案例地,其范围即传统意义上的老八区,包括目前的越秀区、荔湾区、天河区、海珠区、白云区及黄埔区六区。基于获取自百度地图的广州市区餐饮点数据,通过DBSCAN空间聚类算法识别餐饮点的集群,以刻画其基本空间分布格局,在此基础上一方面根据规模对集群进行等级划分,依据中心地理论模型对餐饮集群的空间等级结构进行验证,另一方面,利用空间形态指标考察各餐饮集群并对其进行形态划分。

1 数据获取与研究方法

1.1 研究数据获取

从已有的餐饮点相关研究来看,其数据获取方式以问卷调查^[5]、实地调研^[3,6]、经济普查资料^[4]等为主,由于城市餐饮点数量巨大且分布广泛,前两种方式较耗费人力物力,且无法覆盖所有餐饮点,而经济普查资料因只统计规模以上餐饮企业亦无法实现对全体餐饮点的有效覆盖。进入大数据时代,网络数据成为城市实体空间研究的重要数据源,近年来如百度地图、高德地图、腾讯地图等在线地图服务商以及大众点评网、美团网等在线生活服务平台整合了大量的地理空间数据,相应的基于网络数据的餐饮空间格局研究亦逐渐兴起。考虑到

在线地图数据具有较高的覆盖性,本文选取中国最大的在线地图服务平台之一百度地图(<http://map.baidu.com/>)作为研究数据来源,通过其官方开放API在研究范围内爬取并筛选出餐饮类兴趣点(Point of Interest, POI)合计27 037个。获取数据包括餐饮点的名称、实际地址、空间坐标等信息。对上述获取数据经过坐标校正等预处理后,建立广州市区餐饮点空间数据库。

1.2 DBSCAN空间聚类

DBSCAN空间聚类算法是一种在数据挖掘、机器学习等领域已得到广泛应用的基于密度的聚类方法,能将具有足够高密度的区域划分为簇,该算法给出如下基本定义^[20]:①给定任意一点 p ,其半径 ε 内的邻域即点 p 的 ε 邻域;②若某点 p 的 ε 邻域至少包含最小数目 $MinPts$ 个点,则称点 p 为核心点;③给定点集 D ,若点 q 在点 p 的 ε 邻域内,而 p 是核心点,则称点 p 到点 q 关于 ε 和 $MinPts$ 直接密度可达;④若存在对象链 p_1, p_2, \dots, p_n ,其中 $p_1 = p, p_n = q$,则对于 $1 \leq i \leq n$,称点 p_i 到点 p_n 关于 ε 和 $MinPts$ 密度可达;⑤若存在点 $O \in D$,使点 p 和点 q 从 O 关于 ε 和 $MinPts$ 密度可达的,则称点 p 和点 q 关于 ε 和 $MinPts$ 密度相连。

本空间聚类算法的流程为:以点集 D 中某一个点 p 出发,若点 p 的 ε 邻域包含点多于 $MinPts$ 个,则表明点 p 是核心对象,创建以 p 为核心的簇,将其 ε 邻域中的直接密度可达的点加入该簇中。将与簇中所有核心对象直接密度可达的点加入簇中,迭代本过程直至所有与点 p 密度相连的点都加入簇中时,选定尚未被加入任意簇的另一个点出发,重复上述过程,直至没有新的点可加入任意簇中时,聚类算法结束,未被加入任何簇的点即为噪声点。算法需要设置两个参数:邻域半径 ε 及簇最小点数 $MinPts$,其中参数邻域半径 ε 对聚类结果的影响较大,只要设置恰当参数即可在空间中有效发现任意形状的簇。

为确定最优的算法参数,相关研究者提供了一种辅助参数的确定办法^[21]:计算每个点与其第 k 个最近邻点间的距离 $k-dist$,并对所有点根据 $k-dist$ 由大到小进行排序绘制排序 $k-dist$ 图,根据排序 $k-dist$ 图中曲线斜率出现明显变化的拐点时的 $k-dist$ 值,即可确定参数邻域半径 ε 的取值。在参数 ε 得到确定的基础上,统计点集中每个点的邻域 ε 内点数量的期望,即可得到参数簇最小点数 $MinPts$ 的最优取值:

$$MinPts = \frac{1}{n} \sum_{i=1}^n Count_i$$

其中 $Count_i$ 为点 i 的邻域 ε 内点的数量。通过本算法识别出的餐饮点簇,即为餐饮点的空间集群。

1.3 集群特征分析

根据相应的空间格局分析方法,基于ArcGIS平台分别生成通过空间聚类算法识别出的餐饮集群的最小边界矩形、最优拟合椭圆及最小凸包,用于代表集群的外形轮廓。在此基础上选取紧凑度(COR)、延伸度(ELG)、不对称度(ASM)、密度($DENSITY$)、集中度(G)、分形维数(D_c)等指标^[22](表1),以刻画识别出的餐饮集群的二维空间形态。

2 结果分析

2.1 餐饮集群识别结果及其空间格局

基于ArcGIS平台结合Python语言编程实现DBSCAN空间聚类算法,在运行聚类算法前,为确定聚类的最优参数,首先绘制餐饮点集的排序 $k-dist$ 图,将算法参数分别设置为邻域半径 $\varepsilon=80m$ 、簇最小点数 $MinPts=12$,基于上述参数根据各餐饮点的空间坐标进行聚类,在研究范围内的合计27 037个餐饮点中识别出397个集群(簇),并以各集群包含的餐饮点数量作为该集群的规模指标,将部分规模较大的集群以图表方式列出。同时,空间聚类算法识别出9 086个离群噪声点,此类餐饮点因没有与任何集群在指定阈值下密度相连而被认为是孤立的,表明其在空间上并没有与其他餐饮点形成有效集聚。在SPSS平台下对集群的规模 S 与规模位序 r 进行双对数曲线回归分析,结果呈现如下幂函数关系: $\ln S = 7.473 - 0.828 \ln r$ ($R^2 = 0.979, p < 0.001$),说明餐饮集群符合幂律分布规律。

在此基础上对餐饮点进行核密度分析,由分析结果(图1b)可见广州市区餐饮点主要在两片区域集中连绵分布:一是越秀区中西部、荔湾区东部及海珠区西北部,大致对应广州改革开放前的建成区范围(老四区),基本是市区内人口密度最高的地区,稠密的人口作为最直接的动力吸引餐饮业在此区域的集聚分布;二是天河区天河南商圈到石牌一带,是1990年代以来新兴的商业中心,是广州市区餐饮点核密度值最高的核心地区。

由空间聚类结果(表2、图1a)来看,聚类算法识别出的规模最大的前三个餐饮集群呈现“一个主中心、两个副中心”的空间结构,其中,主中心位于天河区的天河南商圈,拥有正佳广场、天河城等大型

表1 餐饮集群的空间形态指标

Tab.1 The spatial morphology indicators of catering cluster

指标	定义
紧凑度(COR)	$COR = \frac{4\pi A}{P}$, 其中A为最小凸包的面积、P为最小凸包的周长
延伸度(ELG)	$ELG = \frac{l}{w}$, 其中l为最小边界矩形的长轴长度、w为最小边界矩形的短轴长度
不对称度(ASM)	$ASM = 1 - \frac{b}{a}$, 其中a为最优拟合椭圆的长半轴、b为最优拟合椭圆的短半轴
密度(DENS)	$DENS = \frac{n}{A}$, 其中A为最小凸包的面积、n为集群内餐饮点个数
集中度 ^[23] (G)	$G = 100 \times \frac{\sqrt{\sum_{i=1}^n (N_i/n)^2}}{\sqrt{\sum_{i=1}^n (1/N)^2 - 1}}$, 其中n为集群内餐饮点个数, N为餐饮点的坐标个数, N_i 为第i个空间坐标对应的餐饮点个数
分形维数(D_c)	$D_c = \lim_{r \rightarrow 0} \frac{\ln C(r)}{\ln r}$; $C(r) = \frac{1}{n^2} \sum_{i,j=1}^n H(r-dij)$; $H(r-dij) = \begin{cases} 1 & r-dij \geq 0, i \neq j \\ 0 & r-dij < 0, i \neq j \end{cases}$ 。其中C(r)为空间关联函数、 $H(r-dij)$ 为Heavside阶跃函数

购物广场,是目前广州市体量最大的商圈,对应于核密度分析图中密度值最高的区域,其集群规模亦显著大于其他所有集群;两个副中心分别位于越秀区的北京路及海珠区的江南西,对应于核密度图中两个密度值次高区域,其中北京路一带是广州传统的商业中心区。餐饮业作为城市商业服务业的重要组成部分,与其布局存在着紧密的关联,天河南主中心与北京路副中心是广州商业服务业高度集聚的两个地区,亦恰好对应《广州市城市总体规划(2011—2020)》中划定的公共服务中心体系中的两个主中心。而另一个副中心江南西一带虽然并非市级核心商圈,但作为上世纪末广州市商业设施的重点建设中心之一,餐饮娱乐休闲功能相对突出^[24]。

通过上述对比可见经过DBSCAN聚类算法识别出的餐饮集群与核密度分析结果中的热点地区有很好的对应性,可用以研究餐饮点的空间分布格局,但传统的核密度分析方法相比,DBSCAN聚类算法能够判断哪些餐饮点在空间上形成显著集群,并定量识别集群的空间范围与数量规模,排除识别过程中的人为主观影响并为后续定量分析提供支撑。

2.2 餐饮集群的等级结构特征

在已得各集群规模的基础上,为进一步研究餐饮集群的等级结构特征,对其进行等级划分。考虑到各集群规模只是一维数据而非二维数据,因此利用k-means方法对其进行等级划分。为确定划分的最佳等级数,以轮廓系数作为评价指标,最终在 $n=2,3,\dots,8$ 中确定最佳划分等级数为 $n=6$,并基于此在SPSS平台上进行聚类,得到各等级集群的数量,并利用中心地理论模型对其进行验证,计算对

表2 广州市区规模前20的餐饮集群

Tab.2 The top 20 catering clusters in Guangzhou urban area

位序	集群	规模	地点	位序	集群	规模	地点
1	#245	1 093	天河南	11	#332	184	上社
2	#99	674	北京路	12	#268	178	京溪
3	#92	555	江南西	13	#262	171	广州东站
4	#42	390	西华路	14	#242	169	客村
5	#290	369	石牌	15	#107	168	白云万达广场
6	#333	341	龙洞	16	#60	166	宝业路
7	#155	284	中山三路	17	#252	163	梅花园
8	#156	260	建设街	18	#181	161	东晓南
9	#361	249	东圃	19	#347	155	棠下
10	#76	209	新市	20	#272	150	珠江新城东

应等级的中心地数量(低等级中心地在数量上包括所有更高等级的中心地)及等级内各集群的空间形态指数均值,在此基础上进行分析:

各等级集群在数量上符合中心地理论模型,但集群空间结构与理论模型存在差异。前三个等级的餐饮集群数量分别为1、2、6,对应等级的中心地数量分别为1、3、9,与 $K=3$ 的传统中心地理论格局在数量上恰好一致,第四等级与第三等级中心地数量比为 $K=3.22$,第五等级与第四等级中心地数量比为 $K=3.45$,而第六等级与第五等级中心地数量比为 $K=3.97$,低级中心地与高级中心地的数量比 K 值表现随等级的降低而逐渐增加的趋势。相对高等级的中心地呈现 $K=3$ 的基于市场原则的中心地规律,与吴郁文等对广州市商业区位的分析结果是类似的^[25],然而随着尺度由宏观到微观,广州市区餐饮的分布格局呈现由 $K=3$ 的基于市场原则的中心地理论向 $K=4$ 的基于交通原则的中心地理论的转变。其主要原因可能为在宏观尺度上,路网发达密布的市区大体符合市场原则的理想地表假设,可以

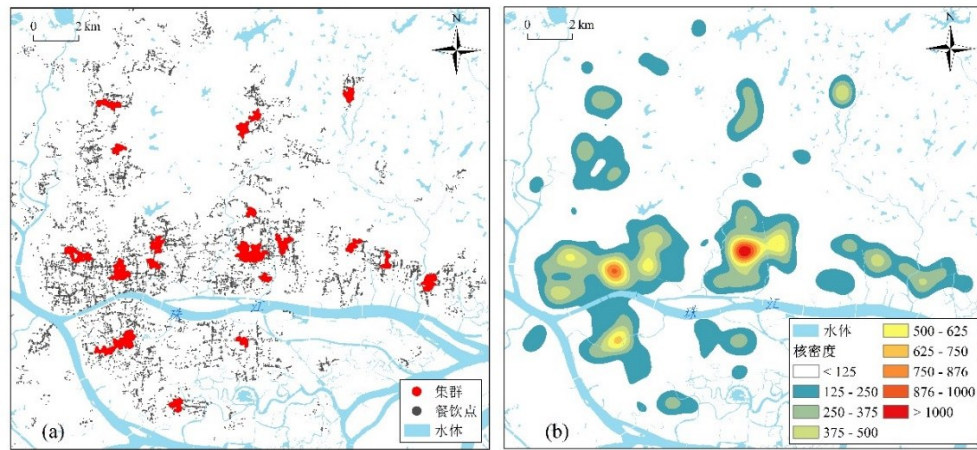


图1 广州市区规模前二十的餐饮集群及餐饮点核密度分析结果

Fig.1 The top 20 catering clusters in Guangzhou urban area and the Kernel Density analysis result of catering spots

表3 集群的等级结构及其形态指数均值

Tab.3 The grading structure of clusters and the mean value of its morphology indicators

等级	集群数量	中心地数量	COR	ELG	ASM	DENS	G	D _c
1	1	1	0.856	0.600	0.282	1 306.70	717.39	0.754
2	2	3	0.811	0.830	0.381	1 266.78	329.27	0.748
3	6	9	0.794	0.619	0.331	1 392.85	195.78	0.775
4	20	29	0.762	0.673	0.459	1 402.34	137.51	0.792
5	71	100	0.725	0.625	0.506	2 046.16	94.44	0.807
6	297	397	0.623	0.525	0.526	4 320.93	44.78	0.811

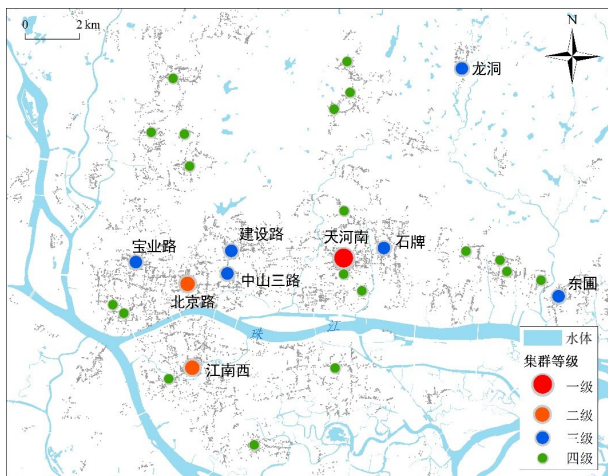


图2 前4个等级集群的空间结构

Fig.2 The spatial structure of clusters with top 4 levels

假设交通在大尺度上是大致均质的;然而随着尺度向微观的迈进,现实情况与均质地表交通的差异越发增大,交通因素对餐饮点格局的影响不容忽视,交通原则对低等级餐饮集群分布格局起着越来越强的支配作用,使低级与高级中心地的数量比趋近于 $K=4$ 。然而,由于城市实际情况与中心地理论的严格假设条件不尽相同,在人口分布、交通要素、地形条件等因素的不均匀性的作用下,各等级集群的服务范围、分布格局并未呈现与传统中心地理论模

型一致的空间结构形态,例如海珠区存在第三等级集群缺失现象,除副中心江南西外,只有客村、宝业路及东晓南三个第四等级集群,而白云区亦明显缺乏高等级餐饮集群。

根据等级划分结果,最高的两个等级餐饮集群分别对应上文的“一主两副”中心,是面向全市的市级餐饮业中心。第三等级、第四等级的餐饮集群基本可视为区级、次区级餐饮中心,结合所包含的餐饮点信息对等级较高的集群进行分析:①位于中心市区的北京路、西华路、宝业路、建设街等集群基本是自下而上逐步集聚而成的,本土广府餐饮特色显著,属于城市服务业集群中的城市—社区中心集群,中心市区密集的人口是其形成的主要驱动因素;②天河南(天河城广场、正佳广场)、中山三路(中华广场)、白云万达广场、广州东站(东方宝泰广场)等是伴随广州城市空间扩张而形成的餐饮中心,也基本对应城市总体规划及商业网点规划重点打造的公共服务次中心,作为典型的依托大型商业中心形成的餐饮集群,与商业中心关联紧密,业态档次相对较高,西餐、自助餐、日韩料理等时尚休闲类餐饮及连锁式快餐企业比例相对较大;③东圃、龙洞、石牌、上社、棠下等餐饮集群基本处于近郊的大型城中村,城中村作为非正规土地空间,其低廉

的空间成本与极高的居住人口密度吸引餐饮业的高度集聚,虽然其餐饮档次相对低下,以小吃快餐店为主导业态,但在空间上形成显著的规模集群,且由于城中村以外来居住人口为主,内部本土菜系(粤菜)比例显著偏少,然而在以往的研究中,此类基于城中村的餐饮集群相比起其他类型餐饮集群往往受到忽视。

2.3 餐饮集群的空间形态特征及其分类

利用表1中的空间形态指标,对识别出的各餐饮集群进行考察,发现各集群主要呈现如下空间形态特征:①从不同等级集群的形态指标均值(表3)来看,高等级的餐饮集群形态以连绵成片的团块状为主,因此有着相对较高的紧凑率(COR)和延伸度(ELG),随着餐饮集群等级的逐步降低,集群形态逐渐变为小规模沿街分布的条带状,使得其紧凑率、延伸度降低;赖静怡^[3]曾将城市餐饮集群按形态划分为建筑单体型、街道型、园区型三类,从本文的分析结果来看,小规模条带状集群对应于其划分的街道型集群,而大规模的团块状集群往往基于但不限于商业建筑单体,而是向外延伸连绵成片,从而呈现出“以点带面”的格局。②随着等级由高到低,餐饮集群反映变形程度的不对称度(ASM)指标趋于上升,表明低等级的餐饮集群不对称性有所增强、空间形态趋于多样化,使用基于距离的聚类方法将难以有效识别此类空间集群。③从集中度(G)指标来看,指标值最高的是主中心天河南($G=717.39$),其次依次分别为花城汇北区(#260, $G=666.75$)、柏德来商业城(#41, $G=666.75$)、富力商贸大厦(#29, $G=428.28$)、太阳新天地购物中心(#296, $G=418.52$)等,均为围绕某个商业中心而形成的餐饮集群,因为大多数餐饮点集中在某一栋或几栋建筑单体中,从而在空间上表现出高地理集中度;建设街、西华路等城市一社区中心型餐饮集群及棠下、石牌、上社等处于城中村的餐饮集群因餐饮点分布比较均匀,内部缺乏较大型的商业中心单体,因此地理集中度相对不高。④随着等级由高到低,集中度趋于下降,同时反映空间分散程度的指标分

形维数(D_c)趋于上升,说明等级越低的餐饮集群在空间上分布相对越均匀,而等级较高的集群中,餐饮点则倾向于集中在某些大型商业中心。⑤从密度(DENS)指标来看,随着等级由高到低密度平均值趋于上升,而在具有高集中度的集群中,密度指标可用于判断集群内餐饮点是只限于集聚的某一栋或几栋建筑单体内,或向外连绵成片式分布。基于上述分析,本文认为集群的空间形态指标对其形态类型划分具有一定的参考价值,其中紧凑率(COR)、延伸度(ELG)、集中度(G)和密度(DENS)四项形态指标的意义较大,故参照以上四项指标,将集群按形态特征划分为如表4所示的四种类型。

3 结论与讨论

3.1 结论

本文引入在数据挖掘领域被广泛运用的DBSCAN聚类算法,首次提出一种基于该空间聚类的餐饮集群识别方法,基于获取自百度地图的广州市区餐饮点数据,在利用聚类算法识别并提取餐饮集群的基础上,分析其空间格局及等级结构,并利用空间形态指标对餐饮集群进行研究和分类,主要得出如下结论:

第一,通过DBSCAN空间聚类算法,在研究范围内27 037个餐饮点中识别出397个集群及9 086个离群噪声点,以餐饮点的数量规模作为衡量集群大小的指标,广州市区餐饮点呈现以天河南集群为主中心、以北京路及江南西两个集群为副中心的“一主两副”空间结构,并研究其与城市服务业现状布局及相关规划的关系。此外,还发现集群的规模与按规模排序的位序呈现有显著的幂函数关系($R^2=0.979, p<0.001$),符合幂律分布的特征。

第二,根据规模将餐饮集群划分为六个等级,发现各等级集群在数量上基本符合中心地理论模型,并随着等级由高到低呈现由 $K=3$ 的市场原则向 $K=4$ 的交通原则的转变。根据紧凑率、延伸度及集中度等空间形态指标,可将集群划分为街道型、片区型、单体一片区型、单体型四类。随着等级由高

表4 按空间形态特征划分的四类餐饮集群
Tab.4 Four kinds of catering clusters divided by the spatial morphological characteristics

类型	空间形态指标特征	说明
街道型	低集中度、低紧凑率及延伸度	沿街道两侧条带状延伸分布,布局较为均匀,常见于低等级集群
片区型	低集中度、高紧凑率及延伸度	连绵成片团块状分布,片区内布局较为均匀,常见于高等级集群
单体一片区型	高集中度、中低密度	依托某一栋或几栋餐饮点高度集聚的建筑单体向外连绵成片,形成“以点带面”的格局,常见于高等级集群
单体型	高集中度、高密度	餐饮点在某一栋或几栋建筑单体集中,但未有向外连绵成片

到低,集群形态由连绵成片的团块状转变为沿街分布的条带状,形态不对称性有所增强,餐饮点在集群内的分布趋于分散。

3.2 讨论

基于本文分析结果,并对比相关规划,现状广州市区餐饮集群格局存在如下问题:高等级餐饮集群过度集中于中心城区,海珠区与白云区缺失区级餐饮集群。因此,在未来的城市餐饮产业空间布局规划中,需要继续加大力度促进中心城区餐饮点向外疏解。在城市总规与商业网点发展规划的布局方案指导下,建议依托城市中轴线南段、白云新城次中心分别培育海珠区、白云区餐饮业的区级中心,分担市级餐饮主中心的部分职能,以保证餐饮集群的结构合理与布局均衡。同时伴随着目前正在推进中的城中村改造工作,一批依托城中村形成的餐饮集群将逐渐受到冲击,需要借此机会对餐饮业空间布局进一步优化。针对现状餐饮集群,在相关规划中根据其特征及空间形态类型作出合理定位,提出具有针对性的发展策略,如发挥集中度指数较高的单体型集群对周边的带动效应,促进其向单体一片区型集群的转变。

随着大数据时代的到来,百度地图等在线网络平台提供了规模巨大、易于获取的POI数据,为更好地在城市研究与规划实践中发挥这些海量数据的价值,本文的研究方法同样可以运用于其他类型服务设施点集群的识别与研究中。另一方面,本文的研究还有不少有待完善之处,如在获取餐饮点的详细分类、人均消费、口碑评价等信息的基础上,可以对各集群的餐饮点风味、档次等进行更深入的分析,以全面挖掘城市餐饮集群的特征,深入探讨集群与商圈的空间关联等,上述问题有待在未来的研究中进一步完善。

参考文献:

- [1] 秦萧,甄峰,朱寿佳,等.基于网络口碑度的南京市餐饮空间分布格局研究——以大众点评网为例[J].地理科学,2014(7):810-817.
- [2] 张旭,徐逸伦.南京市餐饮设施空间分布及其影响因素研究[J].热带地理,2009,29(4):362-367.
- [3] 赖怡琳.广州市城市餐饮产业集群规划设计研究[D].广州:华南理工大学,2012.
- [4] 舒舍玉,王润,孙艳伟,等.城市餐饮业的空间格局及影响因素分析——以厦门市为例[J].热带地理,2012,32(2):134-140.
- [5] 李新阳.上海市中心市区餐饮业区位研究[D].上海:同济大学,2006.
- [6] 方嘉雯.北京五道口地区韩国餐饮业空间布局特征与集聚机理分析[J].经济地理,2014,34(12):106-113.
- [7] Newman L L. Commensality, Sustainability, and Restaurant Clustering in a Suburban Community[J]. Suburban Sustainability, 2014, 2(2): 2.
- [8] Leslie T F, Frankenfeld C L, Makara M A. The spatial food environment of the DC metropolitan area: Clustering, co-location, and categorical differentiation[J]. Applied Geography, 2012, 35(1): 300-307.
- [9] Prayag G, Landré M, Ryan C. Restaurant location in Hamilton, New Zealand: clustering patterns from 1996 to 2008[J]. International Journal of Contemporary Hospitality Management, 2012, 24(3): 430-450.
- [10] Austin S B, Melly S J, Sanchez B N, et al. Clustering of fast-food restaurants around schools: a novel application of spatial statistics to the study of food environments[J]. American Journal of Public Health, 2005, 95(9): 1575.
- [11] Nixon H, Doud L. Do fast food restaurants cluster around high schools? A geospatial analysis of proximity of fast food restaurants to high schools and the connection to childhood obesity rates[J]. Journal of Agriculture, Food Systems, and Community Development, 2011, 2(1): 181-194.
- [12] 夏云. 基于聚类方法的产业分布空间数据挖掘研究[D]. 南京: 东南大学, 2012.
- [13] 林冬云. 基于空间聚类法的北京市海淀区朝阳区企业分布研究[D]. 北京: 北京师范大学, 2005.
- [14] He Y, Tan H, Luo W, et al. MR-DBSCAN: a scalable MapReduce-based DBSCAN algorithm for heavily skewed data[J]. Frontiers of Computer Science, 2014, 8(1): 83-99.
- [15] Habibullayevich G K, Chen X, Shin H. Efficient filtering and clustering mechanism for google maps [J]. Journal of Advanced Management Science, 2013, 1(1): 107-111.
- [16] Yu X, Ding Y, Wan W, et al. Explore Hot Spots of City Based on DBSCAN Algorithm[C]//Audio, Language and Image Processing (ICALIP), 2014 International Conference on. IEEE, 2014: 588-591.
- [17] 丁娟,李俊峰. 基于Web地理图片的中国入境游客POI空间格局[J]. 经济地理, 2015, 35(6): 24-31.
- [18] 王兴中,等. 中国城市商娱乐场所微区位原理[M]. 北京: 科学出版社, 2009.
- [19] 李江涛,王旭东. 广州商贸业发展报告[M]. 北京: 社会科学文献出版社, 2014.
- [20] 李新延,李德仁. DBSCAN空间聚类算法及其在城市规划中的应用[J]. 测绘科学, 2005(3): 51-53.
- [21] 周红芳,王鹏. DBSCAN算法中参数自适应确定方法的研究[J]. 西安理工大学学报, 2012, 28(3): 289-292.
- [22] Yu B, Shu S, Liu H, et al. Object-based spatial cluster analysis of urban landscape pattern using nighttime light satellite images: a case study of China [J]. International Journal of Geographical Information Science, 2014, 28(11): 2328-2355.
- [23] 朱沁夫,李昭,杨樾. 用地理集中度衡量游客集中程度方法的一个改进[J]. 旅游学刊, 2011, 26(4): 26-29.
- [24] 林耿,阎小培. 广州市商业功能区空间结构研究[J]. 人文地理, 2004, 18(3): 37-41.
- [25] 吴郁文,谢彬,骆慈广,等. 广州市市区零售商业企业区位布局的探讨[J]. 地理科学, 1988(3): 208-217.