

引文格式:吕峥,孙群,赵国成,等.顾及方向关系的农村居民地聚类方法[J].武汉大学学报(信息科学版),2023,48(4):631-638.

DOI:10.13203/j.whugis20200546

Citation:LÜ Zheng, SUN Qun, ZHAO Guocheng, et al. A Clustering Method of Rural Settlement Considering Direction Relation [J]. Geomatics and Information Science of Wuhan University, 2023, 48(4):631-638. DOI:10.13203/j.whugis20200546



顾及方向关系的农村居民地聚类方法

吕 峥¹ 孙 群¹ 赵国成¹ 陆川伟² 胡健健³

1 信息工程大学地理空间信息学院,河南 郑州,450000

2 信息工程大学数据与目标工程学院,河南 郑州,450000

3 31016部队,北京,100088

摘要:农村居民地空间分布具有独特的规律性和复杂性,Voronoi图在表达居民地分布特征方面有显著优势。针对当前空间聚类较少考虑实体方向关系的问题,基于Voronoi图提出一种顾及方向关系的农村居民地聚类方法。首先,构建距离约束的Voronoi图,并构建居民地实体间的Voronoi邻近图;然后,利用无向特征与有向特征来综合评价居民地实体间的聚集强度;最后,消除聚集强度小于阈值的实体对的邻近关系,得到聚类结果。采用浙江省宁波地区部分农村居民地数据进行实验,结果表明,所提方法能够有效聚类不同分布模式的居民地,聚类结果符合人的认知习惯。

关键词:空间聚类;方向关系;农村居民地;距离约束的Voronoi图

中图分类号:P208

文献标识码:A

收稿日期:2020-10-15

DOI:10.13203/j.whugis20200546

文章编号:1671-8860(2023)04-0631-08

A Clustering Method of Rural Settlement Considering Direction Relation

LÜ Zheng¹ SUN Qun¹ ZHAO Guocheng¹ LU Chuanwei² HU Jianjian³

1 Institute of Geospatial Information, University of Information Engineering, Zhengzhou 450000, China

2 Institute of Data and Target Engineering, University of Information Engineering, Zhengzhou 450000, China

3 Troops 31016, Beijing 100088, China

Abstract: Objectives: The spatial distribution of rural settlements has unique regularity and complexity. In order to reduce the difficulty of cartographic generalization, we can cluster rural settlements first. Voronoi diagram has significant advantages in expressing the distribution characteristics of settlements, but now spatial clustering seldom considers direction relation between entities. Direction relation is an important part of spatial relation. In theory, the introduction of direction relation in spatial clustering can help to improve the clustering effect. Therefore, based on Voronoi diagram, this paper proposes a clustering method of rural settlement considering direction relation. **Methods:** First, Voronoi diagrams with distance constraint(DC-Voronoi) are constructed, and Voronoi proximity relations among settlement entities are determined. Second, undirected features are calculated based on the area of entities and the area of DC-Voronoi polygons. Directed features are calculated based on offset direction and offset distance of entities in DC-Voronoi polygons. Aggregation strength values of all entity pairs are calculated by combining undirected features and directed features. Finally, clustering result is obtained by eliminating the proximity of entity pairs whose clustering strength value is less than the threshold. Taken the data of rural settlements in Ningbo as an example, this paper sets silhouette coefficient as result evaluation index. **Results and Conclusions:** Compared with the clustering results of density-based spatial clustering of applications with noise method and clustering by fast search and find of density peaks method, the results show that the proposed method can effectively cluster rural settlements with different distribution patterns, and can accurately identify the

基金项目:国家自然科学基金(41571399,41771487,41901397);河南省中原学者项目(202101510001)。

第一作者:吕峥,博士生,主要从事多源矢量数据融合及制图综合研究。lvzheng_xd@163.com

通讯作者:孙群,博士,教授。13503712102@163.com

boundary of rural settlements. The clustering granularity is moderate, and the clustering results accord with people's cognitive habits.

Key words: spatial clustering; direction relation; rural settlement; Voronoi diagram with distance constraint

居民地是地图上基本地理要素之一,其分布具有复杂性和特有的规律性。部分大比例尺地图数据中,居民地实体缺少详细的行政信息,难以将其按照行政归属进行划分,但居民地综合往往需要以居民地聚落为单位,因此需要对居民地在属性不完善的情况下进行空间聚类,挖掘其空间分布规律,将目标区域内所有居民地实体聚合成一个个聚落单元,以辅助后续的制图综合过程。空间聚类旨在依据相似性度量准则将空间实体集划分成若干簇,使得簇内部相似且外部相异^[1]。现在的空间聚类方法大致可分为:(1)基于划分的聚类方法,以 *K*-Means^[2] 为代表方法;(2)基于层次的聚类方法,如利用等边长正交格网进行层次聚合聚类^[3];(3)基于密度的聚类方法,如具有噪声的基于密度的聚类方法^[4](density-based spatial clustering of applications with noise, DBSCAN)、基于快速搜索和发现密度峰值的聚类方法^[5](clustering by fast search and find of density peaks, CFSFDP)等;(4)基于网格的聚类方法,如统计信息网格聚类方法^[6](statistical information grid, STING)等;(5)基于图论的聚类方法,如多级图划分聚类分析方法^[7]等;(6)混合聚类方法,如融合图论与密度思想的混合空间聚类方法^[8]。当前的聚类研究成果很丰富,但不同聚类方法各有优缺点,适用场景不同。

大比例尺地图中,居民地以面的形式进行表达,不同居民地实体及居民地聚落之间差异较大,不同类型的居民地聚类时所用方法有所不同。在城市区域,居民地实体分布密集,形状较规则,对城市居民地进行空间聚类或空间分布模式识别多采用基于图论的方法^[9-10];在农村区域,居民地实体形状不规则,居民地聚落内部密度不均匀,聚落间密度差异大,进行聚类时要自适应密度。Voronoi图对密度、邻近关系等能有效表达,在聚类中应用广泛^[11-12],其中文献^[13]将其用于面目标位置聚类,对密度不均、任意形状分布的面目标集群聚类效果良好。当前的空间聚类方法对实体的几何表现、度量关系及拓扑关系等进行了充分利用,但较少考虑方向关系,因此本文结合空间认知准则和农村居民地空间分布特

点,对 Voronoi图在农村居民地聚类中的应用方式进行拓展,提出了一种顾及方向关系的农村居民地聚类方法。

1 农村居民地聚类中的方向关系

1.1 农村居民地分布情况分析

农村聚落的形成、分布、规模和发展受自然条件、人文因素的综合影响,具有不同的分布规律、集聚特征^[14-15]。农村居民地是地图上重要的地理要素,在大比例尺地图上通常划分为街区式居民地、散列式居民地、分散式居民地、窑洞式居民地、蒙古包及棚房^[16]。

街区式居民地多为房屋建筑密集分布的区域,聚集中心和外围轮廓较清晰;散列式居民地多由普通房屋分布所组成,也有明显的村庄范围,部分村庄有较集中的中心;分散式居民地多由普通房屋组成,但房屋分布依靠天然地势,无分布规律;窑洞式居民地大多呈排、多层分布,与坡壁方向一致;蒙古包、棚房等数量较少,且多为季节性。

在上述农村居民地类型中,分散式居民地通常不进行聚类,而是直接根据分布范围、分布特征及密度对比进行选取;成排、多层分布的窑洞式居民地呈线状分布,适合采用基于图论的聚类方法;蒙古包、棚房数量较少,通常无需聚类;街区式居民地、散列式居民地在地图上占比大,具有相对清晰的村庄范围,对空间聚类有一定的需求。因此,本文的聚类对象主要为街区式居民地和散列式居民地。农村的平面形态大致可分为团块型、条带型和点簇型^[17],其中街区式居民地以团块型为主,散列式居民地以条带型和点簇型为主。

1.2 方向关系研究的必要性

几何和关系是空间信息的两种表现方式,前者侧重地理实体的位置、形态和分布的表达,需严格定义空间坐标;后者侧重关系概念的认知和表达^[18]。关系表现不能精确表达地理信息,但人类能对其进行推理、理解和交流,高层次的空间分析和理解离不开空间关系。空间关系主要包括度量关系、拓扑关系和方向关系等3种基本类

型,其中方向关系用于描述目标对象与参照对象的相对位置关系。几何表现、度量关系与拓扑关系在空间聚类方法中应用广泛,常用的面积、长度、密度及分布模式等特征是对几何表现的应用,度量关系和拓扑关系多用于构建地理实体间的邻近关系,方向关系在空间聚类中的应用则较少。

Voronoi图对居民地实体的几何表现和关系表现均能有效表达,因此在空间聚类中经常使用。当前的聚类方法主要利用Voronoi图的几何表现、拓扑邻近评价实体的聚集强度,几何表现主要指Voronoi多边形的面积及其衍生参数,拓扑邻近主要用于确立实体间的邻近关系,但对某些实体的分布情况仍不能准确描述,例如密集分布的聚落的外围实体可能因其Voronoi多边形的面积较大而被判定为离群分布。在几何表现和

拓扑邻近的基础上引入方向关系,能够更准确识别实体的分布状态。

街区式居民地与散列式居民地的Voronoi图具有显著的方向特征,如图1所示。由图1可知,街区式居民地与散列式居民地聚落内部实体的Voronoi多边形相对面积小,实体在Voronoi多边形内近似居中分布,被其他实体紧密环绕,因此无明显的方向性;聚落边缘实体的Voronoi多边形相对面积较大,实体在Voronoi多边形内的分布有明显的偏向和背离,方向性显著;离群分布的居民地实体的Voronoi多边形相对面积大,实体在Voronoi多边形内无显著偏向和背离。由此可知,实体的分布状态与实体的方向特征、Voronoi图的几何特征关联紧密。仅从形态上而言,聚落边缘的居民地实体向内聚集、向外排斥,可以由此区分不同的聚落。

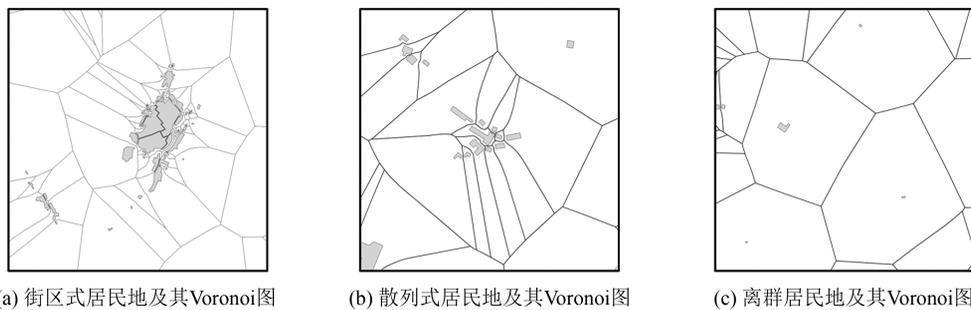


图1 居民地聚落的方向特征

Fig.1 Directional Characteristics of Rural Settlements

为了充分利用构建的Voronoi图,本文采用基于Voronoi的空间方向关系描述模型,所涉及的方向关系包括两部分:(1)居民地实体间的方向关系;(2)居民地实体与对应Voronoi多边形的方向关系。

1.3 居民地实体间的方向关系

居民地实体间的方向关系指目标实体对象与其一阶Voronoi邻近对象间的方向关系。实体间的空间邻近关系的定义是空间聚类的核心问题。实体间的邻近关系包括欧氏距离邻近与Voronoi(Delaunay)邻近两种,Voronoi邻近在度量邻近目标空间分布时更具优势^[19]。描述方向关系及后续聚类过程皆以Voronoi邻近为基础。设实体集 $P = \{p_1, p_2, \dots, p_n\} \subset R^2$,对于 $\forall p_i, p_j \in P, i \neq j$,若满足如下关系,则实体 p_i 与 p_j 为Voronoi邻近:

$$v_i \cap v_j \neq \emptyset \quad (1)$$

式中, v_i, v_j 分别为 p_i, p_j 的Voronoi多边形。

方向关系模型分为定量描述和定性描述两类,定性描述模型在方向关系表达和推理使用更

普遍,但缺少量化信息,不适用于数值计算。以方位角的形式量化描述居民地实体间的方向关系,便于后续计算,采用文献[20]中方法进行实体间方位角的计算: $\forall p_i, p_j \in P, i \neq j, p_j$ 为 p_i 的一阶Voronoi邻近对象,两者Voronoi多边形的公共边 $e_{ij} = v_i \cap v_j$,连接 e_{ij} 的端点,其垂线与指北方向线的夹角即为 p_j 相对于 p_i 的方位角,记作 θ_{ij} 。

1.4 居民地实体与对应Voronoi多边形的方向关系

居民地实体与对应Voronoi多边形的方向关系,即居民地实体偏离其Voronoi多边形几何中心的方向,是方向关系用于聚类的关键,包括偏离方向和偏离度。Voronoi图的势力范围是实体间相互妥协而得到的独属范围,分析面积可得到实体的分布密度,分析边界可得到邻近关系,分析实体在其Voronoi多边形内相对位置得到的偏离方向和偏离度在一定程度上能够反映实体的聚散情况。

通过实体和Voronoi多边形的几何中心定量计算偏离方向。 $\forall p_i \in P$,其偏离方向为Voronoi

多边形几何中心 c_{vi} 指向实体几何中心 c_{pi} 的方向, 可以方位角的形式表达为:

$$a_i = \text{atan2}(dy, dx) \in [-\pi, \pi] \quad (2)$$

2 顾及方向关系的居民地聚类

不同类型的居民地聚落具有不同分布模式, 不同分布模式的居民地聚落其 Voronoi 图表现也不同, 综合考虑居民地实体和 Voronoi 图的几何表现及关系表现, 能够更准确判断居民地的分布模式, 从而指导居民地聚类。

2.1 距离约束的 Voronoi 图

普通 Voronoi 图是一种空间剖分的数据模型, 表现为一组生长元同时向四周扩张直至相遇所形成的各生长元空间势力范围集合^[21]。普通 Voronoi 图在解决空间关系、空间分析等问题研究中发挥着重要作用, 具有空间邻近、空间密度与势力范围等优良空间性质, 因此在空间聚类中应用广泛。但在空间聚类中使用普通 Voronoi 图评价实体间的邻近关系存在一些问题。由地理学第一定律可知, 地理事物或属性在空间中

相互关联, 相关性大小与彼此间的距离负相关^[22]。地图上的居民地实体是对现实世界中地理实体的抽象, 空间影响有限, 当居民地实体间的距离增大到一定程度时, 两者将不可能属于一个居民地聚落, 在一定尺度下其相关性可以忽略不计。普通 Voronoi 图在剖分空间时, 稀疏分布或聚落边缘的居民地实体的势力范围远超其真实的势力范围, 从而导致原本距离远、相关性小的居民地实体间产生不必要甚至错误的邻近关系。为解决上述问题, 在居民地聚类中采用距离约束的 Voronoi 图。如图 2 所示, 设实体集 $P = \{p_1, p_2, \dots, p_n\} \subset R^2$, 对于 $p \in R^2$, 存在 $p_i, p_j \in P, p_i$ 的 DC-Voronoi 多边形可表示为:

$$v_{DCi} = \{d(p_i, p) \leq \min[d(p_j, p), r_i], i \neq j \text{ 且 } i, j \in N\} \quad (3)$$

式中, $d(\cdot)$ 为欧氏距离算子; r_i 为 a_i 的最大影响距离。实体集 P 的距离约束的 Voronoi 图可表示为 (见图 2(c)):

$$V_{DC} = \{v_{DC1}, v_{DC2}, \dots, v_{DCn}\} \quad (4)$$

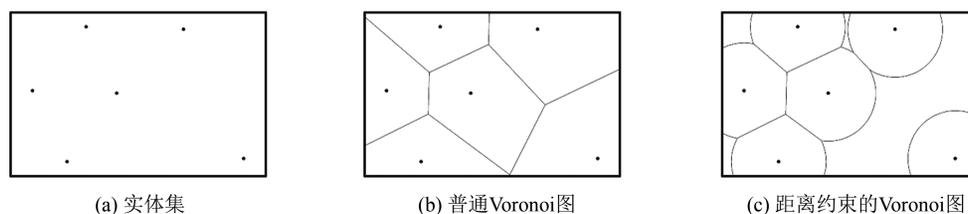


图2 Voronoi图对比

Fig.2 Comparison of Voronoi Diagrams

与普通 Voronoi 图相比, DC-Voronoi 图能更合理地控制居民地实体的影响范围, 消除大量不合理的邻近关系。如图 3(a) 所示, 使用普通 Voronoi 图判断邻近关系时, 居民地实体 I 与居民地聚落 II、居民地聚落 III 的外侧实体普遍存在邻近关系, 但从现实角度出发, 其欧氏距离过大, 空间相关性小, 不应存在邻近关系。如图 3(b) 所示, 在 DC-Voronoi 图剖分下的居民地实体 I 与居民地聚落 II、居民地聚落 III 空间隔离, 相互不影响, 更符合人们的空间认知。

在参数设置方面, 最大影响距离可以人工设置, 也可以根据实体特征自适应计算。文献[23]动态计算各实体的影响范围阈值, 但在聚类中使用 DC-Voronoi 图旨在消除不合理邻近关系, 同时使 Voronoi 多边形的形状更加规则, 居民地实体的地理特征差异在最大影响距离上的表现差距并不大, 因此本文选择人工设置固定阈值的方

法, 阈值需要尽可能小, 但要保证当实体间距大于阈值时两者不会直接聚集。

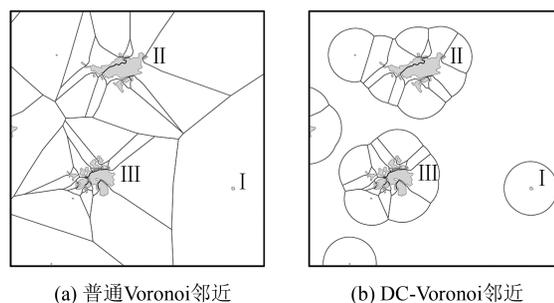


图3 Voronoi邻近关系对比

Fig.3 Comparison of Voronoi Proximity Relations

2.2 聚集强度评价

居民地实体的分布具有方向性, 目标对象对不同方向的邻近参照对象的聚集倾向不同, 即居民地实体的聚集倾向具有各向异性。无向特征和有向特征均能对实体空间分布进行不完全描

述,需要在无向特征的基础上引入有向特征来综合评价居民地实体间的聚集强度。

2.2.1 无向特征

实体分布密集的区域,Voronoi 势力范围小;反之,分布稀疏的区域,Voronoi 势力范围大。势力范围没有方向性,无向特征的确定需要围绕势力范围展开。在基于 Voronoi 图的空间聚类方法中,常用的无向特征有 Voronoi 面积、面积差(Voronoi 面积-实体面积)、面积比(实体面积/Voronoi 多边形面积)及比率(面积差/面目标周长)等。居民地实体自身具备面积属性,在描述居民地实体的聚散特征时需要考虑,而 Voronoi 面积未考虑实体面积;面积差与面积比是同一性质的两种表达方式,但面积比更直观且无需进行归一化;比率的极值相差过大,不够直观。故使用面积比作为评价聚集强度的无向特征,计算如下:

$$\rho_i = s(p_i) / s(v_{DCi}) \quad (5)$$

式中, $s()$ 为面积算子。

在计算面积比时,与普通 Voronoi 图相比,使用 DC-Voronoi 图更有优势。普通 Voronoi 图在居民地分布稀疏的区域势力范围受约束较少,稀疏区域与密集区域的 Voronoi 面积差异巨大,这使得密集区域的面积比正常而稀疏区域的面积比较小,稳定性较差,不利于准确定量描述无向特征。DC-Voronoi 图则因为距离约束的存在使得 Voronoi 面积存在上限,故彼此差距不大,面积比分布在一个受控制的区间内,稳定性较高。

2.2.2 有向特征

偏离度为有向特征,表示实体偏离其 Voronoi 多边形几何中心的程度,是方向性的数值表现。方向性越显著,则偏离度越大,反之偏离度越小。偏离度并不能孤立地判断实体的分布状态,聚落中心的实体和离群分布的实体均无显著方向性,但分布状态截然不同,需要结合无向特征综合判断。

为了与偏离方向保持一致,通过实体几何中心和 Voronoi 多边形间的相对位置计算偏离度,计算时需要顾及 Voronoi 多边形的大小对偏离度的影响,在几何中心间距相等且 Voronoi 多边形形状相同的情况下,偏离度与 Voronoi 面积负相关,计算如下:

$$\zeta_i = \frac{d(c_{vi}, c_{pi})}{\sqrt{s(v_{DCi}) - s(p_i)}} \quad (6)$$

狭义的偏离度仅指实体在偏离方向的偏离程度,所能体现的方向性单一,而实体的一阶 Voronoi 邻近对象分布在各个方向,为评价实体在这

些方向的偏离程度,需要对偏离度进行扩展。方位角与偏离方向角相差越小,则实体在该方向的偏离程度越高,其数值随方位夹角变化的规律符合余弦函数,对偏离度进一步推导可得各向偏离度:

$$\zeta_{ij} = \frac{d(c_{vi}, c_{pi}) \times \cos(\alpha_i - \theta_{ij})}{\sqrt{s(v_{DCi}) - s(p_i)}} \quad (7)$$

式中, ζ_{ij} 表示 p_i 对 p_j 的偏离度; α_i 为偏离方向角; θ_{ij} 为 p_j 相对于 p_i 的方位角。

2.2.3 聚集强度

面积比是无向特征,在偏离度相同的情况下,面积比越大则分布越聚集,是评价聚集强度的基础。偏离度是有向特征,具有各向异性,有向特征和无向特征可相互补充。综合面积比和偏离度, $\forall p_i, p_j \in P, i \neq j, p_j$ 为 p_i 的一阶 Voronoi 邻近对象, p_i 对 p_j 的单向聚集强度为:

$$\delta_{ij} = \rho_i + \zeta_{ij} \quad (8)$$

单向聚集度评价的是 p_i 对 p_j 的单向态度,两个实体能否直接聚集还需要考虑 p_j 对 p_i 的态度。聚集强度与实体间欧氏距离成反比,考虑到同一聚落内邻近居民地实体间距离的最大数量级为 10^2 ,则实体间的聚集强度为:

$$\varphi_{ij} = \frac{(\delta_{ij} + \delta_{ji}) \times 100}{d(p_i, p_j)} \quad (9)$$

一般而言,聚集强度值小于0,表示两个实体相互拒绝,大于0表示相互接受,但实际聚类时,若聚集强度为较小正值仍应被拒绝,因此需要结合实际情况给定阈值 σ ,聚集强度值小于 σ ,则认为实体间不能直接聚集。

2.3 聚类结果评价

空间聚类是无监督分类,聚类质量的评价是聚类分析的重要内容。理想的聚类结果应具有良好的凝聚度和分离度,即类簇内部尽可能紧密,类簇之间尽可能疏远,同时要保持居民地聚落的自然空间形态。轮廓系数^[24]是一种基于相对评价法的聚类结果评价指标,能够兼顾凝聚度和分离度,定义如下:

$$\tau_i = \frac{b(i) - a(i)}{\max[a(i), b(i)]} \quad (10)$$

式中, $a(i)$ 表示第 i 个实体与类簇内其他实体的平均距离,是类簇内部凝聚度的体现; $b(i)$ 表示第 i 个实体与最近簇内实体的平均距离,是类簇间分离度的体现。 $\tau_i \in [-1, 1]$,数值越大,聚类结果越好。

所有实体的轮廓系数的均值即为当前空间

聚类结果的整体轮廓系数(silhouette coefficient, SC),可作为聚类结果的评价指标。

3 实验与分析

采用宁波地区部分1:1万农村居民地作为实验数据,实验区域包括3 607个居民地实体,如图4所示。对实验数据进行初步分析,在居民地密集分布的区域,人工识别的类簇中邻近居民地间距为0~100 m;居民地稀疏分布的区域,人工识别的类簇中邻近居民地间距为100~250 m。为确保距离约束不影响合理的邻近关系,对最大影响范围 r 适当扩大,取值300 m。

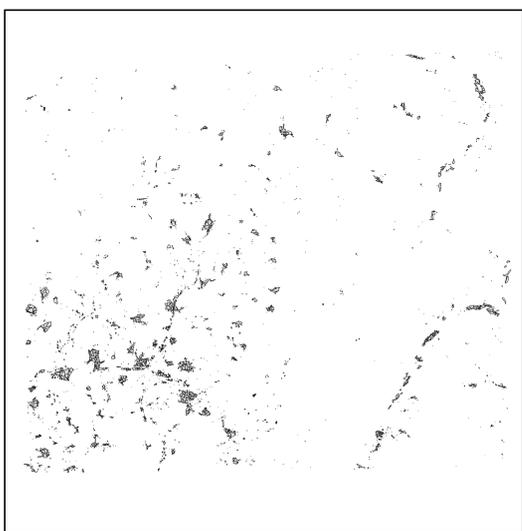


图4 实验区域

Fig.4 Experimental Area

首先根据实验数据及最大影响范围 r 构建DC-Voronoi图,并由此构建居民地实体间的Voronoi邻近图。然后计算所有居民地实体的面积比和偏离度,进一步根据偏离方向和实体间方位角计算实体间的聚集强度,部分计算结果如表1所示。

若实体间的强度小于阈值 σ ,则断开居民地间的邻接关系。阈值 σ 需要人工确定, $\sigma > 0$ 表示实体间有聚集的可能性,在区间 $[0, 0.4]$ 对 σ 按步长0.05依次取值进行空间聚类,并计算聚类结果的整体轮廓系数,结果如图5所示。由图5可知,当 $\sigma = 0$ 时,本文方法表现最佳。

在 $\sigma = 0$ 的情况下对聚类结果进行分析,以凸包的形式显示聚类结果,如图6所示。由图6可知,本文方法能有效识别不同类型的居民地聚落,在不同密度下均有良好的适应性。在高密度区域能识别出聚集分布的团块状居民地

表1 居民地实体间的聚集强度值

Tab.1 Aggregation Intensity Values Between Entities

居民地ID	面积比	偏离方向/(°)	偏离度	方位角/(°)	聚集强度
2052	0.897	209.75	0.018	240.50/60.50	35.883
2054	0.898	212.29	0.014		
1484	0.003	91.45	0.732	119.69/299.69	5.946
1678	0.159	220.75	0.133		
673	0.001	261.56	0.023	261.82/81.82	0.014
1347	0.003	76.33	0.019		
3168	0.019	317.99	0.809	116.03/296.03	-0.535
3187	0.032	102.39	0.887		

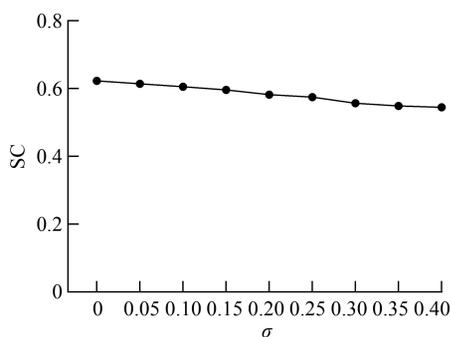
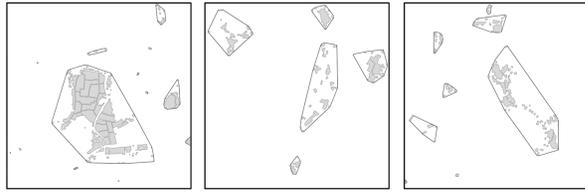


图5 不同阈值下整体轮廓系数

Fig.5 SC Under Different Thresholds

聚落并和周边的低密度居民地实体明显区分;在低密度区域对不同分布模式的点簇型居民地具备良好的划分能力;在分布情况较为复杂的情况下也能对线状分布的条带型居民地进行有效聚类。

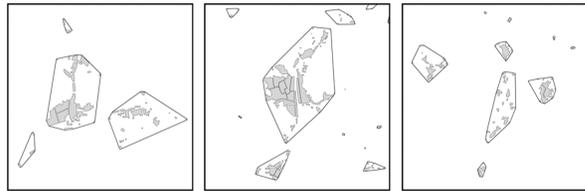
在居民地聚类中,基于密度的聚类方法较为常用,DBSCAN方法及CFSFDP方法是典型代表。分别利用本文方法、DBSCAN方法及CFSFDP方法对局部区域进行聚类实验,结果如图7所示。由图7可知,DBSCAN方法对“颈”结构的处理效果较差,会导致两个高密度区域被“颈”连接而成为一个类簇;与DBSCAN方法相比,CFSFDP方法的聚类结果更准确,对聚集现象明显的居民地能有效聚类,凝聚度和分离度都得到了保证,但聚落周边的离群居民地对聚类结果有一定影响,且部分区域聚类粒度过小;本文方法充分顾及居民地的空间分布特征,对不同分布模式的居民地均能有效聚类,聚类粒度适中,居民地聚落边界识别准确,能有效处理“颈”结构,聚类结果更符合人们的认知习惯。



(a) 高密度区域 (b) 低密度区域 (c) 分布较为复杂

图 6 各类型居民地聚类结果

Fig.6 Clustering Results of Different Types of Rural Settlements



(a) 本文方法



(b) DBSCAN



(c) CFSDP

图 7 实验区域局部聚类结果对比

Fig.7 Comparison of Local Clustering Results in Experimental Area Using Different Clustering Methods

4 结 语

本文对基于 Voronoi 的聚类方法进行扩展,在充分考虑农村居民地空间分布特征的基础上,采用 DC-Voronoi 图提取居民地实体间的邻近关系,综合无向特征和有向特征以评价居民地实体间的聚集强度,再以此为基础进行空间聚类。采用浙江省宁波地区部分农村居民地数据进行实验分析,结果表明,本文方法能够有效应对不同分布类型的农村居民地,聚类结果符合人们的认知习惯。但将方向关系应用于聚类的方式仍有待进一步研究,如何更准确地描述居民地实体的偏离方向及方向关系的不同应用方式,并考虑实体形状对方向关系的影响,也是后续研究的方向。

参 考 文 献

[1] Li Zhilin, Liu Qiliang, Tang Jianbo. Towards a

Scale-Driven Theory for Spatial Clustering[J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(10): 1534-1548. (李志林,刘启亮,唐建波.尺度驱动的空间聚类理论[J].测绘学报,2017,46(10):1534-1548.)

[2] Jain A K. Data Clustering: 50 Years Beyond K-Means[J]. *Pattern Recognition Letters*, 2010, 31(8): 651-666.

[3] Lin Heng, Gong Wei, Shi Shuo. Hierarchical Agglomerative Clustering Algorithm with Equilateral Orthogonal Grid[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(5): 786-791. (林恒,龚威,史硕.利用等边长正交格网进行层次聚合聚类[J].武汉大学学报(信息科学版),2018,43(5):786-791.)

[4] Ester M, Kriegel H P, Sander J, et al. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise[C]//The 2nd International Conference on Knowledge Discovery and Data Mining, Portland, Oregon, USA, 1996.

[5] Rodriguez A, Laio A. Clustering by Fast Search and Find of Density Peaks[J]. *Science*, 2014, 344(6191): 1492-1496.

[6] Wang W, Yang J, Muntz R R. STING: A Statistical Information Grid Approach to Spatial Data Mining [C]//The 23rd International Conference on Very Large Data Bases, Athens, Greece, 1997.

[7] Jin Cheng, An Xiaoya, Chen Zhanlong, et al. A Multi-level Graph Partition Clustering Method of Vector Residential Area Polygon[J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(1): 19-29. (金澄,安晓亚,陈占龙,等.矢量居民地多边形多级图划分聚类方法[J].武汉大学学报(信息科学版),2021,46(1):19-29.)

[8] Shi Yan, Liu Qiliang, Deng Min, et al. A Hybrid Spatial Clustering Method Based on Graph Theory and Spatial Density[J]. *Geomatics and Information Science of Wuhan University*, 2012, 37(11): 1276-1280. (石岩,刘启亮,邓敏,等.融合图论与密度思想的混合空间聚类方法[J].武汉大学学报(信息科学版),2012,37(11):1276-1280.)

[9] Liu Huimin, Hu Wenke, Tang Jianbo, et al. A Method for Recognizing Building Clusters by Considering Functional Features of Buildings[J]. *Acta Geodaetica et Cartographica Sinica*, 2020, 49(5): 622-631. (刘慧敏,胡文柯,唐建波,等.顾及功能语义特征的建筑物空间分布模式识别方法[J].测绘学报,2020,49(5):622-631.)

[10] Xing Ruixing, Wu Fang, Gong Xianyong, et al. A Recognition Method of Homogeneous Two-Tuple

- Collinear Pattern in Building Groups[J]. *Journal of Geomatics Science and Technology*, 2021, 38(1): 91-96. (行瑞星, 武芳, 巩现勇, 等. 一种建筑群同质二元组直线模式的识别方法[J]. 测绘科学技术学报, 2021, 38(1): 91-96.)
- [11] Kang Shun, Li Jiatian, Wu Hao. An Extraction Method for Point Pattern Convergence Under Voronoi Adjacency Relation[J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(5): 649-657. (康顺, 李佳田, 武昊. Voronoi邻近关系支持下的点模式趋同提取方法[J]. 测绘学报, 2017, 46(5): 649-657.)
- [12] Yu Li, Gan Shu, Yuan Xiping, et al. Multi-scale Clustering of Points Synthetically Considering Lines and Polygons Distribution [J]. *Acta Geodaetica et Cartographica Sinica*, 2015, 44(10): 1152-1159. (余莉, 甘淑, 袁希平, 等. 综合线面特征分布的点目标多尺度聚类方法[J]. 测绘学报, 2015, 44(10): 1152-1159.)
- [13] Yu Li, Gan Shu, Yuan Xiping, et al. Position Clustering for Polygon Object Under Dual-constraints [J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(10): 1250-1259. (余莉, 甘淑, 袁希平, 等. 克服双重约束的面目标位置聚类方法[J]. 测绘学报, 2016, 45(10): 1250-1259.)
- [14] Liu Yansui. Research on the Urban-Rural Integration and Rural Revitalization in the New Era in China [J]. *Acta Geographica Sinica*, 2018, 73(4): 637-650. (刘彦随. 中国新时代城乡融合与乡村振兴[J]. 地理学报, 2018, 73(4): 637-650.)
- [15] Zhou Yang, Huang Han, Liu Yansui. The Spatial Distribution Characteristics and Influencing Factors of Chinese Villages [J]. *Acta Geographica Sinica*, 2020, 75(10): 2206-2223. (周扬, 黄晗, 刘彦随. 中国村庄空间分布规律及其影响因素[J]. 地理学报, 2020, 75(10): 2206-2223.)
- [16] State Administration of Quality Supervision, Inspection and Quarantine, People's Republic of China, China National Standardization Management Committee. GB/T 12343. 1—2008 Compilation Specifications for National Fundamental Scale Maps-Part 1: Compilation Specifications for 1:25 000 1:50 000 1:100 000 Topographic Maps [S]. Beijing: Chinese Standard Publishing House, 2008. (中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. GB/T 12343. 1—2008国家基本比例尺地图编绘规范第1部分: 1:25 000 1:50 000 1:100 000地形图编绘规范[S]. 北京: 中国标准出版社, 2008.)
- [17] He Yanhua, Zeng Shanshan, Tang Chengli, et al. Differentiation Characteristics and Forming Mechanism of Rural Settlements in Central China [J]. *Acta Geographica Sinica*, 2013, 68(12): 1643-1656. (贺艳华, 曾山山, 唐承丽, 等. 中国中部地区农村聚居分异特征及形成机制[J]. 地理学报, 2013, 68(12): 1643-1656.)
- [18] Du Shihong, Luo Liqun, Zhao Wenzhi, et al. Research Progress in Multi-scale Spatial Relations [J]. *Journal of Geo-Information Science*, 2015, 17(2): 135-146. (杜世宏, 雒立群, 赵文智, 等. 多尺度空间关系研究进展[J]. 地球信息科学学报, 2015, 17(2): 135-146.)
- [19] Li Jiatian, Wu Huajing, Gao Peng, et al. Landmark Extraction via Composite Features of Voronoi Diagram [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(9): 1448-1454. (李佳田, 吴华静, 高鹏, 等. 一种复合Voronoi几何特征的地标提取新方法[J]. 武汉大学学报(信息科学版), 2020, 45(9): 1448-1454.)
- [20] Yan Haowen, Guo Renzhong. A Formal Description Model of Directional Relationships Based on Voronoi Diagram [J]. *Geomatics and Information Science of Wuhan University*, 2003, 28(4): 468-471. (闫浩文, 郭仁忠. 基于Voronoi图的空间方向关系形式化描述模型[J]. 武汉大学学报(信息科学版), 2003, 28(4): 468-471.)
- [21] Li Jiatian, Luo Fuli, Yu Li, et al. The Gradient Voronoi Diagram and Construction Algorithm [J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(2): 163-170. (李佳田, 罗富丽, 余莉, 等. 梯度Voronoi图及其构建算法[J]. 武汉大学学报(信息科学版), 2016, 41(2): 163-170.)
- [22] Miller H J. Tobler's First Law and Spatial Analysis [J]. *Annals of the Association of American Geographers*, 2004, 94(2): 284-289.
- [23] Gong Xianyong. Research on Settlement Generalization Methods Considering Spatial Pattern and Road Networks [D]. Zhengzhou: Information Engineering University, 2017. (巩现勇. 顾及分布特征和道路网约束的居民地综合方法研究[D]. 郑州: 信息工程大学, 2017.)
- [24] Dinh D T, Fujinami T, Huynh V N. Estimating the Optimal Number of Clusters in Categorical Data Clustering by Silhouette Coefficient [M]//Singapore: Springer, 2019.