



最优信息熵约束的居民地点状要素选取方法

刘万增¹ 陆辰妮² 霍亮³ 吴晨琛¹ 赵婷婷¹ 朱秀丽¹

1 国家基础地理信息中心,北京,100830

2 国家测绘产品质量检验测试中心,北京,100830

3 北京建筑大学测绘与城市空间信息学院,北京,102616

摘要:实现多种约束下的地图信息的负载均衡是制图综合的难点之一。在中小比例尺地图中,对于乡镇及村庄居民点进行尺度转换,需要综合考虑其行政级别、拓扑和度量关系,以使地图信息负载量在一定尺度下达到合理。提出一种基于最优信息熵约束的居民地点状要素选取方法,在最优信息熵约束下,调整度量关系约束,优先考虑语义关系,保留行政级别高的居民点,对行政级别低的居民点,如果不是道路端点,且不满足度量关系约束,则删除该点,不断迭代,直到满足最优信息熵约束。采用 1:250 000 居民地点数据进行实验,实现了维护拓扑一致性、级别优先性、度量合理性的居民地点状要素选取,在有效地保持地图的负载均衡和可读性的同时,实现了地图有效信息量的最大化。采用最优信息熵约束进行居民点选取,在整体上可以保留居民点群空间分布的疏密特征,效果上能够达到图幅信息量的负载均衡。

关键词:点状要素选取;地图制图;最优信息熵约束;Delaunay 三角网;Voronoi 图

中图分类号:P283

文献标志码:A

高质量、高效率地图制图是测绘公益性服务保障的重要任务,其基于不同比例尺的基础地理信息数据,按照制图规范,通过尺度转换、符号映射等,实现地理要素数据的快速符号化、自动注记及图廓整饰等。在数据尺度转换的过程中,尤其当比例尺缩小时,会造成地图负载量过大,引起符号冲突、注记压盖等问题,影响地图的可视化表达效果,导致读图困难^[1-2]。因此,在数据尺度转换的过程中,为了实现地图负载均衡,需要进行要素选取。居民地是地图上重要的地理要素,在中小比例尺地图中,对地图负载量影响较大的是不依比例的乡镇及村庄居民点。本文以不依比例居民点选取为例,研究基于最优信息熵约束的居民地点状要素选取方法。

针对居民点选取,现有的大多数算法如居民地空间比率算法、分布系数法、重力模型算法和圆增长算法等^[3-7],属于基于点的增量式选取算法,优先保存地图上相对重要的点,但未考虑居民点的拓扑和空间分布特征。沈婕等^[8]基于路划网眼划分兴趣点来进行两者之间的拓扑关系维护。Zhou 等^[9]提出了基于云平台点要素综合约

束构建方法,在云计算环境下提高了点选择的效率。艾廷华等^[10]提出了基于 Voronoi 图的点群综合算法,很好地保持点群空间分布特征,并使点群的度量得到有效传递,但其未考虑点的属性意义。遗传算法和神经网络等智能算法^[11-12]将点群综合作为一个优化问题来处理,虽然能保持综合前后的密度对比以及排列规律,但是点群的属性特征并未考虑。因此,需要研究一种新的居民点选取方法,在保证行政级别优先性、地图拓扑一致性和度量合理性的同时,使地图的负载量在一定尺度下达到合理。

信息熵是一种信息量测度方法,本文使用信息熵对居民点选取结果的信息量进行度量,将信息熵与信息模糊度进行融合,基于最优信息熵约束对居民地点状要素进行选取,实现了维护拓扑一致性、级别优先性、度量合理性的居民地点状要素选取。总体思路是在居民点要素的选取中综合考虑居民点要素的行政级别属性、居民点与道路的相接关系及点群的空间分布特征。在语义关系约束下,优先保留行政级别高的居民点;在拓扑关系约束下,尽可能保留道路首端或末端

收稿日期:2019-10-24

项目资助:国家重点研发计划(2018YFC0807005)。

第一作者:刘万增,博士,高级工程师,主要从事 GIS 空间关系、数据库更新及应急制图技术方法研究工作。luwnzg@163.com

通讯作者:陆辰妮,硕士。libra688@126.com

的居民点,避免出现断头路;在度量关系约束下,整体上保留居民点群空间分布的疏密特征,在效果上通过最优信息熵约束,达到图幅信息量的负载均衡等。

1 约束定义

在居民点选取过程中必须维护其与邻近要素空间完整性约束,如空间关系的合理性、语义关系合理性、点群空间分布的合理性等。本文将居民点之间的空间完整性约束归结为最优信息熵约束、度量关系约束、语义关系约束与拓扑关系约束。

1.1 最优信息熵约束

Shannon^[13]提出了信息熵的概念,基于概率论建立了语法层次的信息量度量,将排除冗余后的平均信息量定义为信息熵;Sukhov^[14]采用统计采样的方法完成了地图信息量化测量工作,提出了符号信息熵的概念;Neumann^[15]提出了一种通过对偶图计算符号类型统计分布的拓扑信息熵计算方法;Li等^[16]提出拓扑信息熵、几何信息熵和专题信息熵的定量计算方法;Knopfli^[17]将地图符号感知的不确定性称为模糊度,解决了符号之间的感知相似性问题;Björke^[18]进一步定义了两个相邻地理要素值域为 $[0, 1]$ 的感知相似度函数,成功地将信息论应用在地图综合领域,提出了一种基于熵的地图评估框架。

在居民点选取中,由于尺度的变化,从大比例尺到小比例尺,经过要素选取后,信息量必然减少,为确保地图尺度转换后的地图信息负载均衡,本文提出将最优信息熵作为一种约束条件来度量选取后的居民地点状要素信息量,利用居民地点要素的 Voronoi 图,计算每一个居民地点要素的空间分布概率及其与邻近点的相似度函数,进而计算选取后居民地点群的信息熵和信息模糊度,将居民地点群的信息熵减去信息模糊度得到该点群的有效信息量。有效信息量越大,说明信息模糊度越小。当有效信息量最大时,信息熵和信息模糊度的值达到最优匹配,称此信息熵为最优信息熵。

1.2 度量关系约束

在居民地点状要素选取中,点群的空间分布特征是点要素选取后需要保留的重要信息,一般利用点群密度来表示。点群密度可以定义为特定数目的居民点所占用的分布空间,为计算方便,本文将之转化为相邻点间的欧氏距离约束。

在计算几何中, Delaunay 三角网中的三角形的边是具有邻近关系的点对连接,表达了点的影响范围^[19-20]。本文基于 Delaunay 三角网计算居民点的一阶邻近点,居民点 M_i 与其一阶邻近 M_{i+1} 之间的欧氏距离计算公式为:

$$t_i = \sqrt{(X_{M_i} - X_{M_{i+1}})^2 + (Y_{M_i} - Y_{M_{i+1}})^2} \quad (1)$$

设 d_0 为两个一阶邻近居民点无法分辨出的最大距离, 即当 $t_i < d_0$ 时, 表示两个相邻点在视觉上无法分辨, 可以认为在该比例尺下两个点在视觉上完全重合; 设 d_1 为可以分辨两个相邻点的最小视距, 当 $t_i > d_1$ 时, 两个相邻点在视觉上没有任何重叠, 可以完全区分开。本文假定居民点选取后的目标比例尺为 1:500 000, 制图规范规定居民点的符号大小为 1.2 mm, 因此为保证选取后的视觉效果, 取 $d_0 = 2$ mm, $d_1 = 10$ mm。

本文将 Delaunay 三角网确定的一阶邻近居民点间的距离 D 作为度量关系约束值, 以 $D_0 = d_0 \times \frac{500\,000}{1\,000\,000} = 1$ km 为初始值, 以 $D_1 = d_1 \times \frac{500\,000}{1\,000\,000} = 5$ km 为计算的上限值, 不断迭代, 取得最优信息熵所对应的 D 值即为最终的度量关系约束值。

1.3 语义关系约束

在居民地点状要素的选取中, 居民点的取舍主要取决于居民点要素的行政级别属性。其等级用分类代码表示: AA 代表国名, AB 代表省(直辖市、自治区、特别行政区)行政地名, AC 代表自治州、盟、地区行政地名, AD 代表地级市行政地名, AE 代表县级市行政地名等, 依次类推。

在居民点的取舍中, 可根据语义关系约束建立以下规则: 当设相邻居民点 M_i 的等级大于点 M_{i+1} 的等级时, 删除点 M_{i+1} ; 反之, 则删除点 M_i 。

1.4 拓扑关系约束

GIS 中拓扑关系是指用结点、弧段和多边形所表示的实体之间的邻接、关联、包含和连通关系。居民点群的拓扑关系约束主要考虑居民点与邻近点及居民点与邻近道路线要素的拓扑关系。本文主要考虑以下两种情况: (1) 居民点是否重合, 若两个不同等级的居民点重合, 将会导致注记压盖, 因此, 需要视情况删除等级低的居民点; (2) 居民点是否在邻近道路线要素的端点上, 若删除与道路端点相接的居民点, 将会导致出现断头路的情况, 因此, 与道路两端相接的居民点为必须保留的点。

2 算法设计

最优信息熵约束的居民地点状要素选取方法的总体算法流程如图1所示。

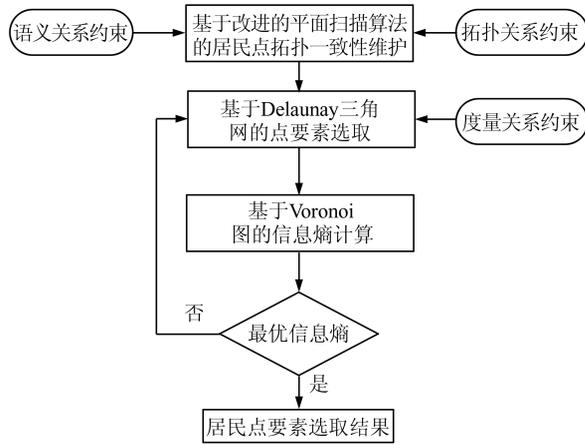


图1 总体算法流程图

Fig.1 Flowchart of Overall Algorithm

基于以上约束实现居民地点状要素选取,需要3个步骤:(1)基于改进的平面扫描算法的要素拓扑一致性维护,目的是去除居民点群中的重合点并检测居民点与道路的邻接关系;(2)基于Delaunay三角网的点要素选取,旨在保证居民点优先级的基础上,利用度量约束实现居民点的选取;(3)利用Voronoi图计算信息熵,即对选取结果进行评价,如果不符合预期值,则不断迭代调整度量关系约束值(即 D 值),并返回第(2)步重新执行,直到取得最优信息熵,算法结束。

2.1 基于改进的平面扫描算法的居民点拓扑一致性维护

利用语义关系约束与拓扑关系约束判断重合点与断头路,若直接进行点点、点线两两判断,其计算量大且效率低,时间复杂性为 $O(n^2)$ 。为了提高效率,本文采用改进的平面扫描算法^[21]。

选择需要选取的居民点数据和道路数据,设点集 P 包含 n 个居民点数据,点集 L 包含 m 条道路的 $2m$ 个端点数据。将点集 P 和点集 L 合成一个点集 Q ,将点集 Q 中每一个点 q_i 的 x 坐标值按照从小到大的顺序排序 $(x_1, x_2 \dots x_{n+2m})$;选择直线 $x=x_1$ 沿 x 方向进行扫描,当扫描到第 i 个点与第 $i+1$ 个点的 x, y 坐标都相等时($i=1, 2 \dots n+2m$),进行点类型判断。若两个点的类型均为居民点,则表明 q_i 与 q_{i+1} 为重合点,进一步判断两个点的语义关系,若第 i 个点的等级高于第 $i+1$ 个点的等级,则保留 q_i ;反之,保留 q_{i+1} 。若第 i 个点类型为居民点,第 $i+1$ 个点类型为道路端点,或第 i 个

点类型为道路端点,第 $i+1$ 个点类型为居民点,则表明 q_i 或 q_{i+1} 是在道路两端的点,为必须保留的重要点。扫描结束后,将点集 Q 中类型为居民点且需要保留的点选择出来,得到点集 Q_1 。

定义点集 Q 的数据结构为 $(x, y, \text{type}, \text{class}, \text{isMarked}, \text{remain})$, x, y 分别为点的横纵坐标; type 代表点的类型; class 是居民点的等级; isMarked 的值为0或-1,该值初始化为0, $\text{isMarked}=-1$ 代表该点为重合点; remain 的值为0或1,0代表删除该点,1代表保留该点。

2.2 基于Delaunay三角网的居民点要素选取

通过预处理过程得到了包含 n_1 个离散点的点集 Q_1 。

如图2(a)所示,运用逐点插入法来构建点集 Q_1 的Delaunay三角网,可得到包含 m 个三角形的三角形列表 T ;初始度量约束值取 $D=D_0$;为了优先考虑等级高的点,按照等级字段排序后遍历点集 Q_1 。

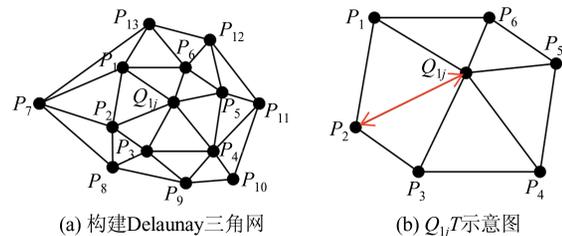


图2 利用Delaunay三角网选取点要素的基本原理

Fig.2 Principle of Point Feature Selection Based on Delaunay Triangulation

如图2(b)所示,假设点 Q_{1j} 为等级最高的点,则由以点 Q_{1j} 为顶点的所有三角形组成的三角形集为 $Q_{1j}T$,逆时针搜索点 Q_{1j} 的一阶邻近点 $P_1, P_2, P_3, P_4, P_5, P_6$,计算点 Q_{1j} 与其一阶邻近点的距离,若其距离大于 D ,说明该区域居民点密度较小,需要保留该一阶邻近点;反之,进入下一步判断。当点 Q_{1j} 与其一阶邻近点的距离小于 D 时,则需判断该一阶邻近点否为重要点,若是,则保留该点;否则,删除该点。点集 Q_1 遍历结束后,得到点集 Q_2 。

2.3 基于Voronoi图的信息熵计算

将选取后居民点集合 X 的信息熵表示为 $H(X)$ ^[13],则:

$$H(X) = - \sum_{i=1}^n P(x_i) \ln P(x_i) \quad (2)$$

式中, $P(x_i)$ 为居民点 x_i 的分布概率; n 为地图中的居民点总数。

居民点 x_i 的分布概率 $P(x_i)$ 的计算是信息熵

计算的基础。由于点群的 Voronoi 图是对空间目标影响区域的剖分,定义了点群目标的空间分布特征^[21-24],某点剖分区域面积越大,则认为该点占有或影响的空间越大,其分布概率也就越大。因此,本文采用基于 Voronoi 图剖分的信息熵度量方法^[16],将居民点的 Voronoi 剖分单元面积与制图区域总面积之比作为该点的分布概率,不仅反映了实体对空间占有的分布,同时,也体现出地物之间的邻接关系^[2,16]。

建立的居民点群的 Voronoi 图如图 3 所示。对整个制图区域进行剖分,假设 S 是整个区域的面积,且被分割为 n 个面积为 S_i 的子区域,其中, $i=1, 2 \dots n$, n 为居民点群个数,则每个居民点的分布概率为:

$$P(x_i) = \frac{S_i}{S} \quad (3)$$

则式(2)可以改写为:

$$H(X) = - \sum_{i=1}^n \frac{S_i}{S} (\ln S_i - \ln S) \quad (4)$$

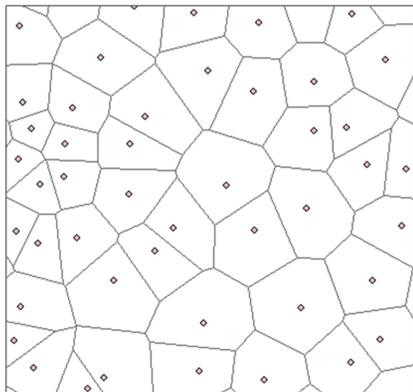


图 3 居民点群的 Voronoi 图

Fig.3 Voronoi Diagram of Residential Point Group

设信息模糊度为 $H(X|Y)$ ^[17],其定义为:

$$H(X|Y) = - \sum_{j=1}^k P(y_j) \sum_{i=1}^n P(x_i|y) \ln P(x_i|y) \quad (5)$$

式中, x_i 为第 i 个居民点; y_j 为 x_i 给定范围内的第 j 个相邻点; $P(y_j)$ 为居民点 y_j 的分布概率; $P(x_i|y)$ 通过归一化对居民点 x_i 与 y_j 的相似度函数的集合进行条件概率计算,其定义为:

$$P(x_i|y) = \frac{\mu(x_i|y)}{\sum_{i=1}^n \mu(x_i|y)} \quad (6)$$

式中, $\sum_{i=1}^n P(x_i|y) = 1$; $\mu(x_i|y)$ 为居民点 x_i 与 y_j 的相似度函数^[18]。

在本文提出的选取方法中,居民点 x_i 和 y_j 之

间的欧氏距离 t 造成它们视觉分离,按照 §1.2 的度量关系约束定义,定义相似度函数 $\mu(t)$ 为:

$$\mu(t) = \begin{cases} 1, & t < d_0 \\ 1 - \frac{t - d_0}{d_1 - d_0}, & d_1 \geq t \geq d_0 \\ 0, & t > d_1 \end{cases} \quad (7)$$

则式(5)可改写为:

$$H(X|Y) = - \sum_{j=1}^n \frac{S_j}{S} \sum_{i=1}^n \frac{\mu(t_i)}{\sum_{i=1}^n \mu(t_i)} \ln \frac{\mu(t_i)}{\sum_{i=1}^n \mu(t_i)} \quad (8)$$

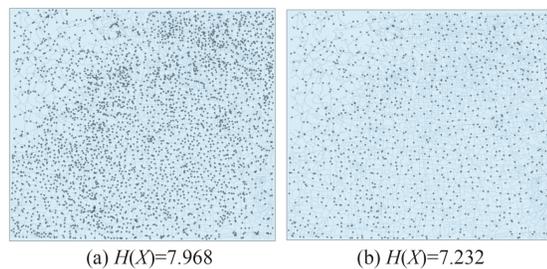
则有效信息量 $R(X)$ 可通过式(4)~式(8)得到,即:

$$R(X) = H(X) - H(X|Y) \quad (9)$$

通过调整度量关系约束值 D , 当 $R(X) = \max[H(X) - H(X|Y)]$ 时,地图有效信息量取得最大值,其对应的 $H(X)$ 为最优信息熵,表示以值 D 为度量关系约束选取的结果最合理。

3 应用与分析

为了验证本文方法的有效性,以某区域 1:250 000 居民点数据和城市道路数据为例,假设综合后目标比例尺为 1:500 000,图 4(a) 为选取前居民点的分布情况,有 3 421 个居民点,其 $H(X) = 7.968$ 。



(a) $H(X)=7.968$ (b) $H(X)=7.232$

图 4 选取效果对比图

Fig.4 Comparison Charts of Selection Results

表 1 为利用上述算法计算出的居民点数据在不同的度量约束下的 $H(X)$ 和 $R(X)$ 的值。由表 1 可以看出, $H(X)$ 随着度量约束值的增加而减小,而 $R(X)$ 随着度量约束值增加呈现先增后减的规律,在度量约束为 3 km 时得到 $R(X)$ 的最大值为 3.869。

图 5 进一步把表 1 的数据可视化,可以明显看出, $H(X)$ 、 $H(X|Y)$ 、 $R(X)$ 随度量约束值的增加而变化的趋势。

图 4(b) 为选取后在 $R(X)$ 最大时的效果,选取后的群点个数为 1 516,总体点数减为原来的

0.443倍,对应最优信息熵 $H(X)=7.232$,最优度量约束值为3 km。与图4(a)比较,信息熵虽然减小了,但其有效信息量达到最大,因此,从视觉效果上看,选取后居民点符号间的关系更趋合理,并较好地保持了选取后的居民地点群的总体结构特征。

表1 实验结果

Tab.1 Experimental Results

度量约束值/km	信息量/个	$H(X)$	$H(X Y)$	$R(X)$
1.0	3 024	7.883	4.119	3.764
1.5	2 687	7.782	4.109	3.673
2.0	2 242	7.617	3.947	3.670
2.5	1 849	7.430	3.670	3.760
3.0	1 516	7.232	3.363	3.869
3.5	1 283	7.066	3.316	3.750
4.0	1 162	6.963	3.311	3.652
4.5	1 102	6.903	3.236	3.667
5.0	1 082	6.880	3.213	3.667

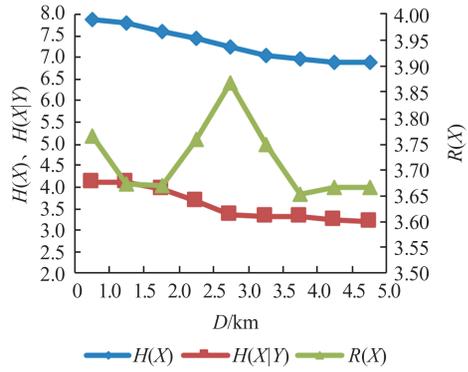
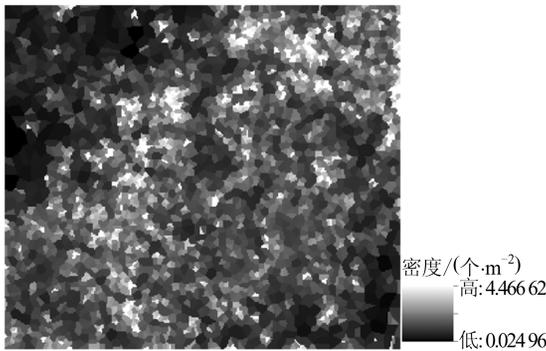


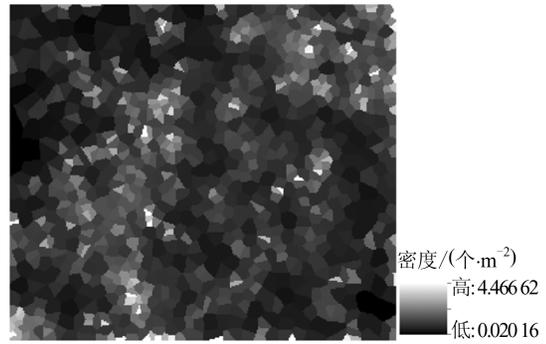
图5 实验结果统计分析图

Fig.5 Flowchart of Statistical Analysis

1) 点群空间分布的合理性。图6(a)、图6(b)分别为选取前、选取后的居民点的密度图,灰度值越高代表密度越高,从视觉效果上可以看出,采用本文方法对居民点选取后,较好地保持了原图的密度分布特征。



(a) 实验前原居民点群密度分布图



(b) 实验后居民点群密度分布图

图6 点群空间分布的合理性验证展示

Fig.6 Rationality Verification of Spatial Distribution of Point Group

2) 语义关系合理性。如图7(a)所示,廊坊市为地级市行政地名,级别为AD,周围部分级别低

于AD的居民点没有被选取上。图7(b)中的结果保留的点的等级均高于周围其他点的等级。



(a) 实验前廊坊市相邻点分布



(b) 实验后廊坊市相邻点分布

图7 属性主次关系合理性验证结果展示

Fig.7 Rationality Verification of Attribute Primary and Secondary Relationship

3)空间关系的合理性。选取前,重合点情况如图 8(a)所示,两个居民点名称分别为娄村与娄村满族,坐标相同,造成了注记冲突;选取后,重合点保留结果如图 8(b)所示,保留级别更高的娄

村满族点。选取前,重要点如图 8(c)所示,名为下场的点为自然村地名(BB);虽然等级较低,但其与道路邻接而保留下来,重要点保留结果如图 8(d)所示。

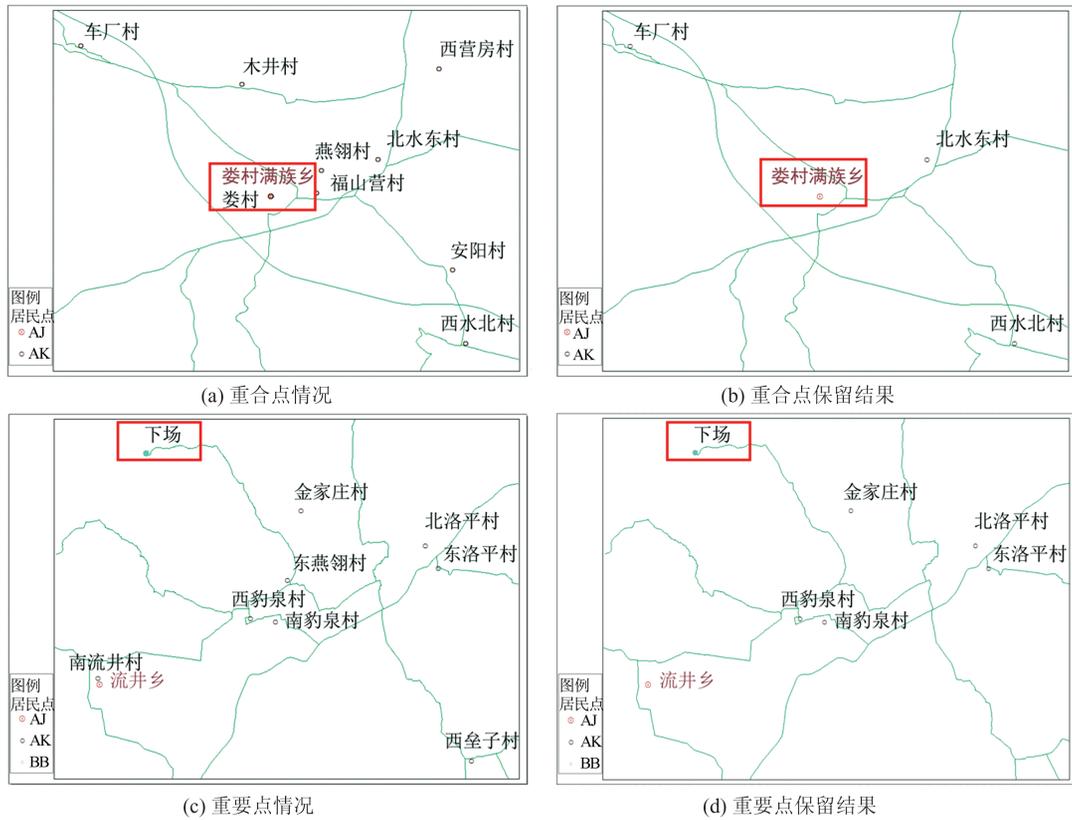


图 8 空间关系的合理性验证结果展示

Fig.8 Rationality Verification Results of Spatial Relationships

4 结 语

本文基于最优信息熵约束提出了居民地点状要素选取的迭代优化方法。该方法具有以下 4 个优点:(1)保持了居民点群的空间分布特征;(2)顾及了居民地点的行政级别差异;(3)保留了其与道路的空间关系的合理性;(4)实现了地图的有效信息量最大化。基于以上方法,开发了 ArcMap 插件,用于应急地图及政府工作用图编制,取得了较好的应用效果。

本文为了突出最优信息熵约束在居民点选取中的控制效果,将问题进行了简化,信息熵的计算没有考虑图上居民点注记所占用的空间,在一定程度上影响了居民点选取后地图表达的整体效果,今后的研究将对该方法进行拓展和优化。

参 考 文 献

[1] Wang Jiayao, Qian Haizhong. Cartographic-Genera-
lization-Knowledge and Its Application[J]. *Geomatics
and Information Science of Wuhan University*,
2006,31(5):382-386 (王家耀,钱海忠.制图综合知
识及其应用[J].武汉大学学报·信息科学版,2006,
31(5):382-386)

[2] Ai Tinghua, He Yakun, Du Xin. Information Entro-
py Change in GIS Data Scale Transformation [J].
Geography and Geo-Information Science, 2015, 31
(2):7-11 (艾廷华,何亚坤,杜欣.GIS数据尺度变
换中的信息熵变化[J].地理与地理信息科学,
2015,31(2):7-11)

[3] Yan Xiongfeng, Ai Tinghua, Yang Min. A Simplifi-
cation of Residential Feature by the Shape Cognition
and Template Matching Method[J]. *Acta Geodaeti-
ca et Cartographica Sinica*, 2016, 45(7):874-882
(晏雄锋,艾廷华,杨敏.居民地要素化简的形状识

- 别与模板匹配方法[J]. 测绘学报, 2016, 45(7): 874-882
- [4] Gao Kai, Yang Min, Zhang Yuepeng. A Method of Automatic Selection of Hash-Style Habitation with Spatial Distribution Characteristics Preserved [J]. *Journal of Geomatics Science and Technology*, 2015, 32(6): 626-630 (高凯, 杨敏, 张跃鹏. 保持空间分布特征的散列式居民地综合选取方法[J]. 测绘科学技术学报, 2015, 32(6): 626-630)
- [5] Yu Yanping, Shen Jie, Shang Zaiying. A Study on the Time Complexity of Point Cluster Selection and Simplification Algorithms [J]. *Journal of Nanjing Normal University (Natural Science Edition)*, 2012, 35(1): 111-116 (于艳平, 沈婕, 尚在颖. 点群选取与化简算法时间复杂度研究[J]. 南京师范大学学报(自然科学版), 2012, 35(1): 111-116)
- [6] Qian Haizhong, Wu Fang, Deng Hongyan. A Model of Point Cluster Selection with Circle Characters [J]. *Science of Surveying and Mapping*, 2005, 30(3): 83-85 (钱海忠, 武芳, 邓红艳. 基于CIRCLE特征变换的点群选取算法[J]. 测绘科学, 2005, 30(3): 83-85)
- [7] Yan Haowen, Wang Jiayao. A Generic Algorithm for Point Cluster Generalization Based on Voronoi Diagrams [J]. *Journal of Image and Graphics*, 2005, 10(5): 633-636 (闫浩文, 王家耀. 基于Voronoi图的点群目标普适综合算法[J]. 中国图象图形学报, 2005, 10(5): 633-636)
- [8] Shen Jie, Zhu Yueqin, Wu Peng, et al. Parallel Computing Method for POIs Selection Based on Stroke Mesh Decomposition [J]. *Acta Geodaetica et Cartographica Sinica*, 2015, 44(S): 54-61 (沈婕, 朱月琴, 吴鹏, 等. 兴趣点选取的路网分割并行计算法[J]. 测绘学报, 2015, 44(S): 54-61)
- [9] Zhou Jiemin, Shen Jie, Yang Shuai, et al. Method of Constructing Point Generalization Constraints Based on the Cloud Platform [J]. *ISPRS International Journal of Geo-Information*, 2018, 7(7), DOI: 10.3390/ijgi7070235.
- [10] Ai Tinghua, Liu Yaolin. A Method of Point Cluster Simplification with Spatial Distribution Properties Preserved [J]. *Acta Geodaetica et Cartographica Sinica*, 2002, 31(2): 175-181 (艾廷华, 刘耀林. 保持空间分布特征的群点化简方法[J]. 测绘学报, 2002, 31(2): 175-181)
- [11] Cai Yongxiang, Guo Qingsheng. Points Group Generalization Based on Kohonen Net [J]. *Geomatics and Information Science of Wuhan University*, 2007, 32(7): 626-629 (蔡永香, 郭庆胜. 基于Kohonen网络的点群综合研究[J]. 武汉大学学报·信息科学版, 2007, 32(7): 626-629)
- [12] Deng Hongyan, Wu Fang, Qian Haizhong, et al. A Model of Point Cluster Selection Based on Genetic Algorithms [J]. *Journal of Image and Graphics*, 2003, 8(8): 970-976 (邓红艳, 武芳, 钱海忠, 等. 基于遗传算法的点群目标选取模型[J]. 中国图象图形学报, 2003, 8(8): 970-976)
- [13] Shannon C E A. Mathematical Theory of Communication [J]. *The Bell System Technical Journal*, 1948, 27: 379-423
- [14] Sukhov V I. Information Capacity of a Map: Entropy [J]. *Geodesy and Aerophotography*, 1967, 10(4): 212-215
- [15] Neumann J. The Topological Information Content of a Map: An Attempt at a Rehabilitation of Information Theory in Cartography [J]. *Cartographica*, 1994, 31(1): 26-34
- [16] Li Zhilin, Huang Peizhi. Quantitative Measures for Spatial Information of Maps [J]. *International Journal of Geographical Information Science*, 2002, 16(7): 699-709
- [17] Knopfli R. Communication Theory and Generalization [C] // Graphic Communication and Design Contemporary Cartography, New York, USA, 1983
- [18] Bjørke J T. Framework for Entropy - Based Map Evaluation [J]. *Cartography and Geographic Information Systems*, 1996, 23(2): 78-95
- [19] Wu Fang, Gong Xianrong, Du Jiawei. Overview of the Research Progress in Automated Map Generalization [J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(10): 1 645-1 664 (武芳, 巩现勇, 杜佳威. 地图制图综合回顾与展望[J]. 测绘学报, 2017, 46(10): 1 645-1 664)
- [20] Shen Jin, Liu Jiping, Lin Xiangguo, et al. A Method for Delaunay Triangulation by Integration of Distance Transformation and Region Adjacency Graphics [J]. *Geomatics and Information Science of Wuhan University*, 2012, 37(8): 1 000-1 003 (沈晶, 刘纪平, 林祥国, 等. 集成距离变换和区域邻接图生成Delaunay三角网的方法研究[J]. 武汉大学学报·信息科学版, 2012, 37(8): 1 000-1 003)
- [21] Chen Jun, Liu Wanzeng, Li Zhilin, et al. The Refined Calculation Method of Topological Relationships Between Line Objects [J]. *Acta Geodaetica et Cartographica Sinica*, 2006, 35(3): 255-260 (陈军, 刘万增, 李志林, 等. 线目标间拓扑关系的细化计算方法[J]. 测绘学报, 2006, 35(3): 255-260)
- [22] Liu Wanzeng, Chen Jun, Yan Chaode, et al. Model for Calculating Topology Stable Area of the Plan Discrete Points [J]. *Acta Geodaetica et Cartographi-*

- ca Sinica*, 2012, 41(1):127-132 (刘万增, 陈军, 闫超德, 等. 平面离散点集拓扑邻近稳定区域计算模型[J]. 测绘学报, 2012, 41(1):127-132)
- [23] Zou Yafeng, Liu Yaolin, Kong Xuesong, et al. Optimization of Rural Residential Land Based on Weighted-Voronoi Diagram [J]. *Geomatics and Information Science of Wuhan University*, 2012, 37(5): 560-563 (邹亚锋, 刘耀林, 孔雪松, 等. 加权 Voronoi 图在农村居民点布局优化中的应用研究[J]. 武汉大学学报·信息科学版, 2012, 37(5): 560-563)
- [24] Yan Haowen, Wang Bangsong. A MWVD-Based Algorithm for Point Cluster Generalization[J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(9): 1 088-1 091 (闫浩文, 王邦松. 地图点群综合的加权 Voronoi 算法[J]. 武汉大学学报·信息科学版, 2013, 38(9): 1 088-1 091)

Selection Method of Residential Point Features Constrained by Optimal Information Entropy

LIU Wanzeng¹ LU Chenni² HUO Liang³ WU Chenchen¹ ZHAO Tingting¹ ZHU Xiuli¹

¹ National Geomatics Center of China, Beijing 100830, China

² National Quality Inspection and Testing Center for Surveying and Mapping Products, Beijing 100830, China

³ School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 102616, China

Abstract: Objectives: The load balancing of map information under multiple constraints is one of the difficulties in cartographic generalization. In small and medium-scale maps, it is necessary to comprehensively consider their administrative levels, topologies and metric relationships for the scale conversion of townships and village residential point features to make the map information load reasonable at a certain scale. **Methods:** This paper proposes a method for selecting residential point features based on optimal information entropy constraints. Under the constraints of optimal information entropy, the metric relationship constraints are adjusted, the semantic relationships are prioritized, the residential point features with higher administrative levels are reserved, and for the residential point features with lower administrative levels, if they are not the endpoints of the road and do not satisfy the metric relationship constraint, then the points are deleted, and the process is iterated until the optimal information entropy constraint is satisfied. **Results:** Experiments with 1 : 250 000 residential point data have realized the selection of residential location elements that maintain topological consistency, level priority, and metric rationality. The load balancing and readability of the map are effectively maintained, meanwhile, the amount of effective information of the map is maximized based on the algorithm. **Conclusions:** The optimal information entropy constraint is adopted for the selection of residential points, which can retain the density characteristics of the spatial distribution of the residential point group as a whole, and achieve the load balancing of map information in effect.

Key words: point feature selection; cartography; optimal information entropy constraint; Delaunay triangulation; Voronoi diagram

First author: LIU Wanzeng, PhD, senior engineer, mainly engaged in the research of GIS spatial relationship, database updating and emergency mapping technology. E-mail: luwnzg@163.com

Corresponding author: LU Chenni, master. E-mail: libra688@126.com

Foundation support: The National Key Research and Development Program of China (2018YFC0807005).

引文格式: LIU Wanzeng, LU Chenni, HUO Liang, et al. Selection Method of Residential Point Features Constrained by Optimal Information Entropy[J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(8):1178-1185. DOI:10.13203/j.whugis20190305 (刘万增, 陆辰妮, 霍亮, 等. 最优信息熵约束的居民地点状要素选取方法[J]. 武汉大学学报·信息科学版, 2021, 46(8):1178-1185. DOI:10.13203/j.whugis20190305)