

散列式面状注记自动配置技术研究

张志军¹ 李霖¹ 于忠海¹ 应申¹

(1 武汉大学资源与环境科学学院,武汉珞喻路 129 号,430079)

摘要:参照国家标准图式规范,结合格式塔心理学,总结了散列式面状注记的配置规则,提出了一种新的自动配置方法。首先,用凸包格网法计算注记的候选位置;然后,用影响注记位置的格式塔因子对候选位置进行质量评价;最后,依据冲突规则得到全局最优解。该方法有效地利用了散列式面状要素周边的区域,扩充了注记的可调节性,并被成功地应用于 1:5 万地形图的数字制图中。

关键词:散列式面状注记;凸包格网法;格式塔;质量评价

中图法分类号:P283.7

我国大中比例尺地形图中需要注记的面状要素很多在几何形态上是散列分布的,这些散列分布的地理要素具有相同注记名称,且相互之间的距离较小。散列式面状注记的本质和目的就是用一个注记指代多个面状地物,其配置过程需要解决两个主要问题:注记在什么位置才是最合理的指代,以及如何计算该位置。目前,面状注记配置算法大多数都是基于定位线和域的^[1-4]。对于散列式面状注记,文献[5]中提出了一种散列居民地注记方案,并不适用于散列水系的注记。

本文参照《国家基本比例尺地图图式》(简称 GB/T20257),为散列式面状注记定义了 3 种配置模式,分析了影响注记位置的格式塔因子,借此引入以格式塔原则为理论依据且符合行业标准的散列式面状注记配置规则,并提出了一种新的注记配置方法——凸包格网散列注记算法予以形式化表达。

1 散列式面状注记配置规则

我国大中比例尺地形图中,需要注记的散列式面状要素主要有居民地和水系。按 GB/T20257 的要求,居民地名称注记一般采用水平行列注出;海、海湾、海港、江、河、湖、沟渠、水库等名

称按自然形状排列注出,名称一般注在河流、湖泊的内部;对于面积较小的散列式面状要素可将名称注在多边形外部。因此,本文为散列式面状注记定义 3 种配置模式:居民地配置模式、水系配置模式和小面积配置模式。本文将共享同一个注记的要素集称为一个注记体。水系配置模式旨在在注记体内部寻找合适的配置点;居民地配置模式旨在在注记体的凸包内寻找合适的配置点;小面积配置模式旨在在注记体的凸包外寻找合适的配置点。

在众多格式塔因子中,影响散列式面状注记位置的主要有:注记与要素相似性因子;注记的拓扑一致性因子;注记共同性因子;注记的视觉方向。因此,本文认为高质量的散列式面状注记应满足几个基本原则:① 注记易读性;② 注记整体性;③ 位置优先性;④ 形状相似性;⑤ 地物关联性。

根据上述 5 项基本原则,定义了与 3 种配置模式相对应的配置规则(见表 1)。

2 凸包格网散列注记算法

2.1 算法思想

目前,国内外学者普遍认可将地图注记自动配置过程分为候选注记位置的产生、候选注记位置的评价、注记位置的选择 3 个阶段^[6,7]。凸包格

收稿日期:2011-04-05。

项目来源:国家自然科学基金资助项目(40871178,41001307);国家 863 计划资助项目(2008AA121601,2007AA12Z241);国家测绘局测绘基金资助项目;地理空间信息工程国家测绘局重点实验室经费资助项目(200907);数字制图与国土应用工程国家测绘局重点实验室开放基金资助项目(GCWD200902)。

表 1 散列式面状注记配置规则表

Tab. 1 Configure Rules for Hash Area Feature

配置模式	优先配置原则	配置规则
居民地配置模式	易读性、整体性、关联性、位置优先性	注记不能与注记体压盖 尽量配置在注记体的凸包内部 注记与各要素平均距离最小
水系配置模式	形状相似性、位置优先性、整体性	注记应沿注记体的凸包的主骨架线排列 注记应包含在注记体的内部
小面积配置模式	易读性、关联性	注记不能与注记体压盖 优先注记于注记体凸包右边

网散列式面状注记算法也遵循这个过程：① 候选位置的确定。从散列式面状要素的形态和属性出发，根据配置规则确定居民地、水系和小面积要素的候选位置。② 候选位置的评价。用注记质量评价函数对候选位置进行评价，评价函数中既包含了注记与要素之间的空间关系，也考虑到了注记之间的冲突。③ 位置的选取。根据冲突规则进行冲突检测，对有冲突的注记进行组合优化，从而得到全局最优解。

2.2 候选位置的确定

居民地配置模式和水系配置模式用固定格网大小(经验值取半个字体大小)将注记体的凸包格网化。居民地配置模式以与注记体不压盖的格网为候选位置，水系配置模式则以注记体所压盖的格网为候选位置。小面积配置模式下需先将凸包向外扩展一定距离(经验值取一个注记长度)的缓冲区再进行格网化，以与注记体不压盖的格网为候选位置。格网越小，候选位置越多，算法精度越高，注记效果越好，但效率会受到一定的影响。

2.3 候选位置的评价

候选位置的评价是多个评价指标的综合。为每个评价指标定义“理想”和“可接受”两种状态。“理想”状态得分为 1，“可接受”状态得分为 0，单个评价指标的得分一般都在 0.0~1.0 之间。如某个单项指标得分为负，则表示该候选位置不可接受，可直接剔除。得分高于 1.0 时按 1.0 参与计算。最终得分根据配置规则中各评价指标的权重进行综合计算。

1) 要素压盖因子。用注记和要素相交部分的面积来评价注记与要素之间的压盖情况，其量化公式为：

$$FLOverlap(l_i, F) = \begin{cases} Area(l_i \cap F)/Area(l_i), & \text{水系模式} \\ 1 - Area(l_i \cap F)/Area(l_i), & \text{其他模式} \end{cases}$$

2) 注记冲突因子。用注记矩形和其他注记矩形相交的面积来评价注记与其他注记之间的冲

突情况，其量化公式为 $LLOverlap(l_i, l_j) = 1 - Area(l_i \cap l_j)/Area(l_i)$ 。所有配置模型中都认为注记与注记之间的冲突是不合理的。

3) 关联性因子。用注记与要素之间的距离来评价注记与要素之间的关联程度，其量化公式为：

$$Association(l_i, F) = \begin{cases} \sum Dist(l_i, f_i)/[M \times Length(l_i)], & \text{居民地模式} \\ 1 - Dist(l_i, F)/Length(l_i), & \text{小面积模式} \end{cases}$$

在居民地配置模式中，取与每个要素的距离的平均值；在小面积配置模式中，取与注记体凸包的距离；水系配置模式中注记应包含在要素内，故不需要此因子。

4) 相似性因子。用注记与要素主骨架线之间的距离来评价注记与要素自然分布的相似程度，其量化公式为 $Similarity(l_i, F) = 1 - Dist(l_i, Skeleton(F))/Length(l_i)$ 。此因子只用于水系配置模式中。

5) 视觉方向因子。用注记配置点和要素重心的水平夹角($\theta, \theta \in (0^\circ, 360^\circ)$)与主方向的夹角来评价注记与要素方位一致性，其量化公式为 $Direct(l_i, F) = 1 - Abs(90^\circ \times Int(\theta \% 90^\circ)/45^\circ - \theta \% 90^\circ)/45^\circ$ 。

6) 位置优先性因子。用注记方位来评价注记方位与视觉习惯的一致性。根据我国的阅读习惯，方位优先级从大到小为右、下、上、左，其量化公式为：

$$DirectPriority(l_i, F) = \begin{cases} 0.75, & \theta \in [0, 45^\circ) \cup [315^\circ, 360^\circ) \\ 0.25, & \theta \in [45^\circ, 135^\circ,) \\ 0, & \theta \in [135^\circ, 225^\circ) \\ 0.5, & \theta \in [225^\circ, 315^\circ) \end{cases}$$

式中， M 为组成注记体的要素总数； N 为注记总数； l_i 为第 i 个注记， $i \in (1, N)$ ； L 为注记的集合； f_i 为第 i 个要素， $i \in (1, M)$ ； F 为组成注记体的各要素的集合； $Skeleton(F)$ 为求取注记体 F 的主骨架线函数； $Area()$ 为求面积的函数； $Dist(a, b)$ 为求 a 和 b 的距离； $Length(l)$ 为求注记 l 长度的函数。

2.3.1 居民地配置模式

居民地配置模式以到各要素平均距离最小、与注记体不压盖的配置点为最优位置，其质量评价指标包括要素压盖因子、注记冲突因子和关联性因子。评价函数为：

$$ResidentScore =$$

$$\begin{cases} -1, & FLOverlap(l_i, F) < 1 \text{ 或 } LLOverlap(l_i, F) < 1 \\ Association(l_i, F), & \text{其他} \end{cases}$$

2.3.2 水系配置模式

水系配置模式将注记打散为多个注记单元,并沿注记体凸包的主骨架线注记于要素内部。其质量评价指标包括要素压盖因子和相似性因子,评价函数为:

DrainageScore =
$$\begin{cases} -1, \text{FLOverlap}(l_i, F) < 0.8 \\ 0.5 \times \text{Similarity}(l_i, F) + 0.5 \times \\ \text{FLOverlap}(l_i, F), \text{其他} \end{cases}$$

2.3.3 小面积注记

小面积配置模式以靠近注记体、靠近优先级最高的主方向线为最优位置。其质量评价指标包括要素压盖因子、注记冲突因子、关联性因子、视觉方向因子、位置优先级因子。评价函数为:

LitterScore =
$$\begin{cases} -1, \text{FLOverlap}(l_i, F) < 1 \text{ 或} \\ \text{LLOverlap}(l_i, F) < 1 \\ 0.25 \times \text{Association}(l_i, F) + 0.25 \times \text{Direct}(l_i, F) \\ + 0.5 \times \text{DirectPriority}(l_i, F), \text{其他} \end{cases}$$

2.4 注记位置的选择

将注记候选位置按其质量评价得分降序排列,第一个作为初始位置参与检测,其他位置作为

该注记的候选位置存储。根据制图知识定义注记检测规则,如面状水系注记需包含在水系要素中、注记不应压盖道路、注记不能压盖注记等。参照上述规则将注记与其他图层要素和注记进行冲突检测,检测结果有 3 种情况:没有冲突、与要素冲突和与注记冲突。对于没有冲突的注记,直接选择其初始位置;对于与注记冲突的注记,在候选位置中选择没有冲突的次优位置,如果所有的候选位置都与要素冲突,则选择初始位置并作为无解情况标注,需要手工调整;对于与注记冲突的注记,将所有可能冲突的注记作为一个组,运用组合优化算法得到全组最优解。

3 实验与分析

以 1:5 万比例尺地形数据为实例进行实验,与 ArcGIS 中 Maplex 注记引擎的效果进行比较。

3.1 实验结果

实验数据中包含散列面状居民地 68 个,散列面状水系 3 个,注记自动配置耗时 45 s,手工调整耗时 10 min。实验结果如图 1 所示。图 1 中注记结果清晰、无歧义,注记与要素的指代关系明确。

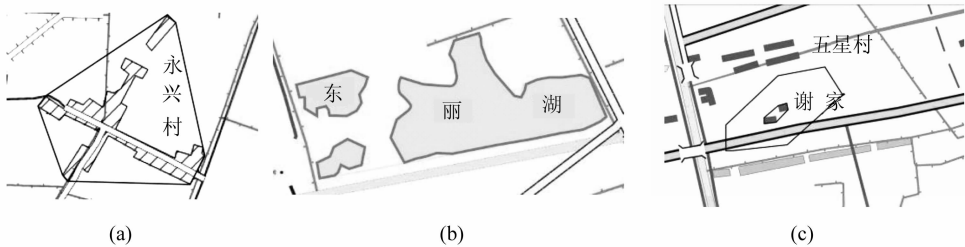


图 1 散列面注记自动配置效果示意图

Fig. 1 Results of Auto-Placement for Hash Area Feature

图 1(a)为居民地配置模式下的注记结果,从图中可以看出,注记左边的要素为“永兴村”(质量评分为 0.84)的“主要”要素,其他为“次要”要素,同时垂直排列的注记与注记体的走势一致,符合格式塔心理学的相似性原则和简化性原则。图 1(b)为水系配置模式下的注记结果,“东”、“丽”、“湖”(质量评分为 0.96)沿要素的凸包的主骨架线排列且都注记于要素内部,符合格式塔心理学的连续性原则和拓扑一致性原则。图 1(c)为小面积配置模式下的注记结果,根据配置规则,要素正右方为最优位置,如图中“五星村”(质量评分为 0.86),当注记与其他要素冲突时,将自动移至次优位置,如图中“谢家”(质量评分为 0.78),符合格式塔心理学的视觉方向一致性。

3.2 对比实验

一般的散列式注记的配置方法(如 ArcGIS 中 Maplex 注记引擎)没有考虑到水系和居民地的配置差异,而是对散列要素中每个要素都进行标注,然后由制图人员选择如何处理副本。如图 2(a)为保留副本的情况,图 2(b)为在页面 2 cm 缓冲区内隐去副本的情况,其注记结果不符合本文所提出的散列式面状注记的基本原则。如图 2(b)中“双梅村”的阅读严重受到了要素的干扰;右边中间的要素由“双梅村”还是“双梅桥”来指代存在着歧义。

图 2(c)为凸包格网法注记效果。通过对比实验可以看出,凸包格网算法比一般散列注记算法效果理想;同时算法思路清晰,容易实现,具有

很高的运行效率。

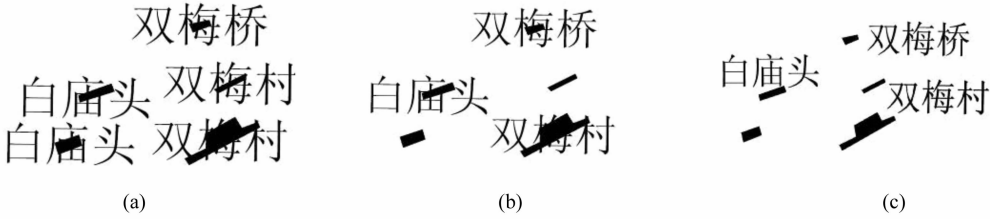


图 2 对比实验效果图

Fig. 2 Results of Different Labeling for Hash Area Feature

4 结 语

本文讨论了散列式面状注记的配置规则和自动配置方法,并详细阐述了 3 种配置模式。以 1 : 5 万地形图数据为实例进行实验,验证了本文算法的可行性和高效性。由于注记配置过程中尚未顾及其他图层要素,加上注记配置自身的复杂性,有时候还需要人工交互。下一步的工作应考虑多图层要素同时注记的问题,并提供更多的处理方案,如图外指向注记等。

参 考 文 献

[1] Ahn J, Freeman H. AUTONAP:an Expert System for Automatic Map Name Placement[C]. International Symposium on Spoitial Data Handing, Zurnich, Switzerland, 1984

[2] 马飞. 用数学形态学自动快速寻找地图注记位置[J]. 武汉测绘科技大学学报, 1996,21(2):150-153

[3] 姜永发,张书亮,兰小机,等. 长对角线法实现 GIS 中矢量地图面状地物汉字注记的自动配置[J]. 武汉大学学报·信息科学版, 2005,30(6):544-548

[4] 赵静,罗兴国,张汝云. 一种新的电子地图注记算法:格网法[J]. 计算机工程, 2008(7):278-279

[5] 张晓通,李霖,舒亚东,等. 面状要素注记智能化配置方法研究[J]. 武汉大学学报·信息科学版, 2008,33(7):762-765

[6] Chirie F. Automated Name Placement with High Cartographic Quality: City Street Maps[J]. Cartography and Geographic Information Science, 2000, 27(2): 101-110

[7] Edmondson S, Christensen J, Marks J, et al. A General Cartographic Labelling Algorithm[J]. Cartographica: the International Journal for Geographic Information and Geovisualization. 1996, 33(4): 13-24

第一作者简介:张志军,博士生,研究方向为空间数据管理及可视化表达。
E-mail:zhangzhijun@whu. edu. cn

Auto-Labeling of Hash Anea Features

ZHANG Zhijun¹ LI Lin¹ YU Zhonghai¹ YING Shen¹

(1 School of Resources and Enoinonmental Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: We focus on how to automatically place annotation for hash area features and proposes a new method with reference to the compilation specifications for “cartographic symbols for national fundamental scale maps” and Gestalt psychology theory. At first, the candidate positions are generated with convex-grid Method. Then, they are evacuated with the Gestalt factors on the annotation. Finally, the global optimum location of annotation is determined by “conflict” rule. Utilizing the peripheral zone of features is a notable merit of this method. The method has been successfully used in producing topographic maps at 1 : 50 000 scale.

Key words: discrete area labeling; convex-grid; Gestalt; quality evaluation