

doi: 10.6046/zrzyg.2022006

引用格式: 石善球. 基于几何与属性匹配的地理信息数据融合更新方法[J]. 自然资源遥感, 2023, 35(1): 251–257. (Shi S Q. A geographic data fusion and update method based on geometric and attribute matching[J]. Remote Sensing for Natural Resources, 2023, 35(1): 251–257.)

基于几何与属性匹配的地理信息数据融合更新方法

石善球

(江苏省基础地理信息中心, 南京 210013)

摘要: 针对多源地理信息数据在尺度、几何位置和属性等方面存在不一致性而造成难以融合更新这一现状, 探讨了一种基于几何与属性匹配技术的地理信息数据融合更新方法。该方法首先通过广义 Voronoi 图获取候选集, 有效地提高候选集的获取效率并减少不相关目标对候选集的影响; 接着, 基于几何和属性匹配等关键技术实现点、线、面 3 种不同几何类型数据的匹配分析; 最后, 基于匹配结果从参考地理信息数据中提取增量数据, 并完成目标数据的融合更新。实验结果表明, 采用几何与属性匹配的地理信息数据融合更新方法, 能够高效识别并提取增量数据, 对监测数据更新模式的创新探索有一定的借鉴意义。

关键词: 几何匹配; 属性匹配; 融合更新; 增量数据

中图法分类号: P 208 **文献标志码:** A **文章编号:** 2097-035X(2023)01-0251-07

0 引言

自然资源信息化需要把握的重点之一就是突出数据的整合与融合^[1]。随着自然资源统筹管理的逐步落实^[2-3], 各种来源的地理信息数据开始汇集, 问题也随之而来——数据在尺度、几何位置和属性等方面存在不一致性, 这给自然资源调查监测历史数据建库带来诸多不便。自然资源部颁布的《新型基础测绘试点建设技术大纲》中也特别强调, 要将“研究多源(元)数据融合和智能化处理技术”作为一项重要的核心技术研究。因此, 要想从多个维度和层次实现对数据的整合与融合, 开展多源地理信息数据整合与融合技术的研究显得尤为重要。

目前, 数据整合与融合更新的主要方法有: 基于日志的方法、基于变化检测表(影子表)的方法、基于时间戳字段的方法和基于自动匹配的方法^[4]。考虑到各种来源的地理信息数据互相之间无权访问更改日志, 无权添加影子表或时间戳字段, 将比例尺相近的多源地理信息数据进行自动匹配是整合与融合的最彻底、最精确的方法。现有匹配算法可分为几何匹配、拓扑匹配和属性匹配 3 类^[5]。几何匹配

方法的优点是充分利用数据的几何信息和特征, 不足之处是多源数据的几何位置往往存在差异^[6]; 拓扑匹配克服了几何匹配方法的不足, 但是该方法通常需要其他相似特征一起使用^[7]; 属性匹配的效率高, 但是属性匹配方法过于依赖数据模型以及属性数据类型^[8]。

本文基于几何与属性匹配理论探讨了点、线、面 3 种不同类型多源地理信息数据的融合更新方法, 并通过省级基础测绘和地理国情数据验证该方法的可行性。

1 研究方法

多源地理信息数据融合更新方法步骤为: 首先, 对 2 套不同来源的地理信息数据进行坐标、格式统一等预处理; 然后, 根据空间数据类型对质量检查合格的数据进行图层关联; 再以待更新地理信息数据为源实体构建广义 Voronoi 图, 在此基础上从参考地理信息数据中获取待匹配候选集; 接着, 待更新地理信息数据与待匹配候选集进行几何匹配和属性匹配; 最后, 基于匹配结果从参考地理信息数据中提取增量数据并完成待更新地理信息数据的融合更新。技术路线如图 1 所示。

收稿日期: 2022-01-05; 修订日期: 2022-07-11

基金项目: 江苏省自然资源科技项目“面向国土空间决策支持的智慧江苏时空大数据情势分析”(编号: 2020035)资助。

作者简介: 石善球(1980-), 男, 教授级高级工程师, 主要从事测绘地理信息数据处理研究。Email: 6919368@qq.com。

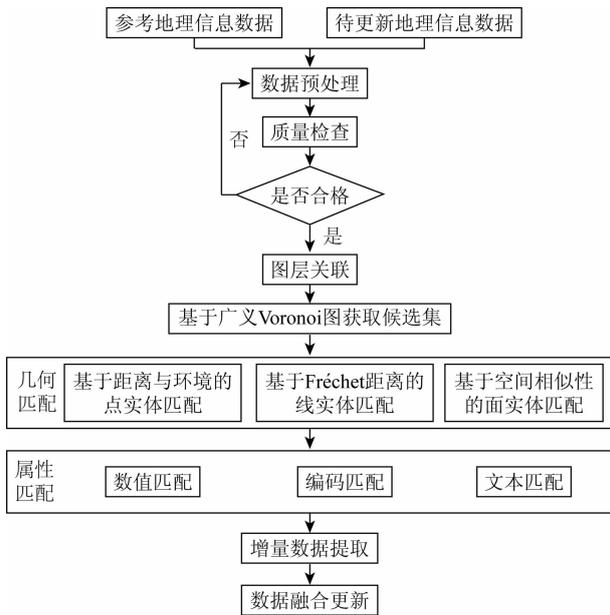


图1 多源地理信息数据融合更新技术路线

Fig. 1 Technology roadmap of geographic information data fusion and update

2 关键技术

2.1 候选集获取

通常实体匹配一般需要进行数据预处理、候选

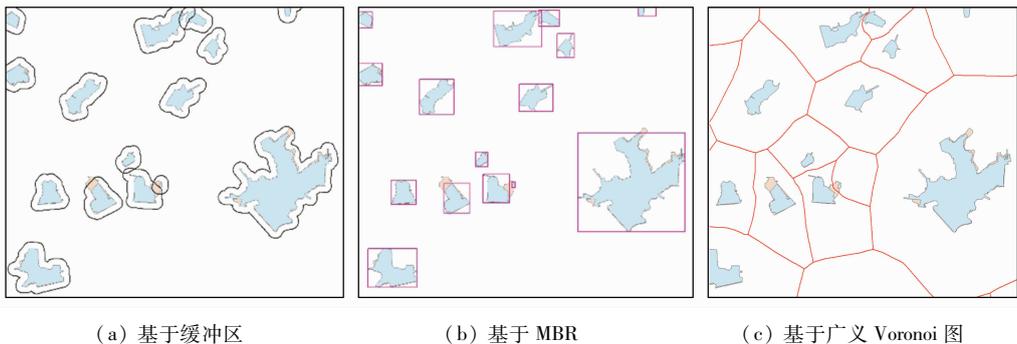


图2 候选集获取方法对比

Fig. 2 Comparison of candidate set acquisition methods

2.2 几何匹配

几何匹配原理是先度量实体的一个或几个几何特征的相似度,然后通过阈值来判断是否属于同名实体。常见的几何特征包括实体之间的距离、形状描述和方向趋势等,不同匹配算法对这些几何特征在数学描述上略有差异,如点实体匹配多采用距离指标来衡量相似度^[12];线实体匹配最常用的指标有线的方向、Hausdorff距离和Fréchet距离^[13];面实体则多利用面的形状特征,主要包括面积、相似度、不变距与质心度等^[14]。鉴于省级基础测绘和地理国情数据的特点,本文针对点、线、面3种实体分别采用3种不同的匹配方法。

集获取、相似度计算、匹配与结果分析等步骤,其中候选集获取是实体匹配需要解决的问题之一。传统的候选集获取方法有缓冲区方法^[9]和最小外包矩形(minimum bounding rectangle, MBR)方法^[10],但前者存在阈值依赖性过大的问题,后者误匹配或漏匹配较多。考虑到Voronoi图不受目标位置、距离和比例尺等因素影响的优势,采用基于Voronoi图的候选集获取方法。实际生产中为了同时满足点、线、面的应用,本文采用文献[11]改进的广义Voronoi图获取候选集。

基于广义Voronoi图的候选集获取方法思路如下:首先,将比例尺相对较小的面实体作为源实体(蓝色),另一个数据集作为目标实体(棕色),对源实体构建广义Voronoi图。在此基础上,遍历每个Voronoi多边形,搜索相应的目标实体中的要素作为匹配候选集。3种不同候选集获取方法对比结果如图2所示。与基于缓冲区方法(图2(a))和基于MBR方法(图2(b))相比,利用基于广义Voronoi图方法(图2(c))可快速地得到与源实体对应的候选集搜索区域,搜索范围不重叠,有效地提高了候选集的获取效率并减少不相关要素对候选集的影响。

2.2.1 基于距离与环境的点实体匹配

多源地理信息数据之间经常会出现一方局部区域点实体比较密集,而另一方则比较稀疏,甚至出现一方多点合一,另一方只有一个实体的现象。为提高匹配质量,本文采用基于距离与环境的点实体匹配技术。

首先,设 m_1 和 m_2 分别为源实体和匹配候选集的点位精度, m 为2个同名点的距离精度,则距离阈值 r 的计算公式为:

$$r = 3m = 3 \sqrt{m_1^2 + m_2^2} \quad (1)$$

接着,计算源实体 a 和候选匹配实体 b 之间的距离相似度 $D_{a,b}$,计算公式为:

$$D_{a,b} = 1 - \frac{d}{r}, \quad (2)$$

式中 d 为源实体和候选匹配实体之间的欧式距离。

然后,采用文献[15]的方法计算源实体 a 和候选匹配实体 b 之间的环境相似度 $E_{a,b}$,具体过程如下:以当前点实体 a 为中心在其周围构建一个 2×2 的格网区域,实体 a 周围左上、右上、右下、左下 4 个方位的方格区域分别用 LT, RT, RB, LB 表示,格网的边长等于距离阈值 r ,则四元组 $E_a = (p_{LT}, p_{RT}, p_{RB}, p_{LB})$ 表示点实体 a 的环境特征, p 表示各方格区域的点密度,用某个区域包含的点的个数来计算,例如 LT 区域包含 2 个点,则 $p_{LT} = 2$ 。针对 2 个待匹配点实体 f_a 和 f_b 的环境特征,依次比较它们环境特征四元组中对应项的大小,如果相等则计数变量加 1,设最后计数变量的大小为 N' ,则环境相似度计算公式为:

$$E_{a,b} = \frac{N'}{4}. \quad (3)$$

最后,利用各个特征的相似度计算源实体和候选匹配实体的综合图形相似度 $P_{a,b}^p$,计算公式为:

$$P_{a,b}^p = \alpha_1 D_{a,b} + \beta_1 E_{a,b}, \quad (4)$$

式中 α_1 和 β_1 分别为距离相似度和环境相似度的权重,可根据具体情况设定,但必须满足 $\alpha_1 + \beta_1 = 1$ 。

2.2.2 基于 Fréchet 距离和线段模型的线实体匹配

由于多源地理信息数据之间存在采集标准的差异,所以同名对象之间大多是 $1:N$, $M:1$ 或者 $M:N$ 匹配类型,为提高匹配质量,通过建立线段模型,抽象为 $1:1$ 匹配类型来处理。

具体思路如下:首先,将源实体和匹配候选集实体分别在结点(交叉点)处打断;然后,依据唯一

$$P_{a,b}^A = \alpha_2 sim_d(a,b) + \beta_2 sim_{sh}(a,b) + \gamma_2 sim_{si}(a,b), \quad (7)$$

式中 α_2 , β_2 和 γ_2 分别为距离相似度、形状相似度和面积相似度的权重,可根据具体情况设定,但必须满足 $\alpha_2 + \beta_2 + \gamma_2 = 1$ 。

2.3 属性匹配

属性匹配原理是利用不同数据源对相同地理现象的基本性质描述相同或相近的特点。本文鉴于多源地理信息数据属性数据类型和属性表达的内容层面不同,研究了数值、编码和文本 3 种不同的属性匹配方法。在此基础上利用各种属性项的相似度计算要素的综合相似度。

2.3.1 数值匹配

主要用于计算 integer 型、float 型和 double 型的数值之间的相似度,如道路宽度和水库面积等。其

的实体编码将源实体和匹配候选集实体分别进行合并连接;接着,再将源实体和匹配候选集实体分别在结点(交叉点)处打断;最后采用文献[13]的方法,基于 Fréchet 距离计算源实体和候选匹配实体之间相似度 $P_{a,b}^L$,计算公式为:

$$P_{a,b}^L = 1 - \frac{F_d}{r}, \quad (5)$$

式中 F_d 为源实体和候选匹配实体之间的 Fréchet 距离。Fréchet 距离是计算 2 条曲线距离的算法,用来判断 2 条曲线的相似度。设有 2 条连续曲线,即 $f: [0,1] \rightarrow R^2$, $g: [0,1] \rightarrow R^2$, θ 和 ρ 分别为连续非减的实函数,且 $\theta[0] = \rho[0] = 0$, $\theta[1] = \rho[1] = 1$,则 Fréchet 距离 $F_d(f,g)$ 计算公式为:

$$F_d(f,g) = \inf \max_{s \in [0,1]} \{ \text{dist} \{ f[\theta(s)], g[\rho(s)] \} \}. \quad (6)$$

2.2.3 基于空间相似性的面实体匹配

面实体匹配同样包含 $1:N$, $M:1$ 或者 $M:N$ 匹配等类型,其中 $M:N$ 匹配类型是面实体匹配重点解决的问题,因此,面实体的匹配需要采用计算多个几何特征的方法。本文采用文献[16]的方法,分别计算源实体 a 和候选匹配实体 b 之间的距离相似度 $sim_d(a,b)$,形状相似度 $sim_{sh}(a,b)$ 和面积相似度 $sim_{si}(a,b)$ 。如果实体间是一对一的匹配关系,则直接根据相似度计算公式计算;如果是一对多或者是多对多的匹配关系,则先获得源实体集和候选匹配实体集,然后再根据相似度公式计算实体集之间的匹配度。

在此基础上,利用各个特征的相似度计算源实体和候选匹配实体的综合图形相似度 $P_{a,b}^A$,计算公式为:

相似度 $p_{a,b}^N$ 计算公式为:

$$p_{a,b}^N = \frac{\min [V_a, V_b]}{\max [V_a, V_b]}, \quad (8)$$

式中 V_a 和 V_b 分别为 a 和 b 2 个待匹配实体的一个属性项值。

2.3.2 编码匹配

编码匹配是通过比较它们前 n 位的编码是否相同来计算其相似程度。设 $sNum()$ 表示编码的个数; $sCode()$ 表示某要素的编码,若 $m = sNum[sCode(a)]$, $n = sNum[sCode(b)]$,且 $m \leq n$,则将 $sCode(a)$ 与 $sCode(b)$ 中前 m 个数字对应相减,若字符串中从第一位开始存在 $k \leq m$ ($k = 0, 1, 2, \dots, m$) 个连续的 0 数字,则编码相似度 $p_{a,b}^C$ 计

算公式为:

$$p_{a,b}^c = \frac{k}{n} \quad (9)$$

2.3.3 文本匹配

文本匹配以文本的方式来描述要素的名称和位置等信息。本文考虑属性值漏输或错输的情况从 2 个方面进行匹配。一是 2 个字符串中相同字符的数目,设为 $iCount$; 二是 2 个字符串最长公共子串的字符数目,设为 $iCommon$, 则文本字符串的相似度指标计算公式为:

$$p_{a,b}^T = \alpha_3 \frac{iCount}{m_{a,b}} + \beta_3 \frac{iCommon}{m_{a,b}}, \quad (10)$$

式中: $m_{a,b}$ 为 2 个字符串中个数最少的个数; α_3 和 β_3 为可调参数,但必须满足 $\alpha_3 + \beta_3 = 1$ 。

2.3.4 综合属性相似度

利用各种属性项的相似度指标计算源实体 a 和候选匹配实体 b 的综合属性相似度 $p_{a,b}^S$, 计算公式为:

$$p_{a,b}^S = \left\{ \sum_{k=1}^{N_s} [sim(F_{a,k}, F_{b,k}) ESW_{aK}] \right\} / N_s, \quad (11)$$

式中: $F_{a,k}$ 为实体 a 的第 K 个属性; $F_{b,k}$ 为实体 b 的第 K 个属性; $sim(F_{a,k}, F_{b,k})$ 为实体 a 和实体 b 第

K 个属性计算出来的匹配相似度值; $0 \leq K \leq N_s$; N_s 为实体总的属性项个数; ESW_{aK} 为 a_K 对应的权重,一般取值都为 1。

2.4 增量提取及融合

从多源矢量参考数据中提取图形改变、属性改变、图形和属性改变的 3 种数据,同时获取对应自然资源监测数据的唯一实体编码,在此基础上完成数据的融合更新。假设 T_1 为几何匹配阈值; M_1 为属性匹配阈值。则识别规则为:

$$\begin{cases} P_{a,b} \leq T_1, p_{a,b}^S > M_1 & \text{识别为图形改变} \\ P_{a,b} > T_1, p_{a,b}^S \leq M_1 & \text{识别为属性改变} \\ P_{a,b} \leq T_1, p_{a,b}^S \leq M_1 & \text{识别为图形和属性改变} \end{cases} \quad (12)$$

3 应用与评价

3.1 融合应用

基于江苏省全域 2020 年度国情监测数据更新 2019 年度江苏省全域 1:10 000 基础测绘数据(重要要素),其中 2020 年度国情监测数据涉及 14 个图层,共计 191 667 个要素,基础测绘数据涉及 10 个图层,共计 441 544 个要素。采用本文方法,分类开展了基于几何与属性匹配的多源地理信息数据融合更新实验,具体融合结果如表 1 所示。

表 1 基础测绘数据融合更新结果

Tab. 1 Results of geographic information data fusion and update

要素名称	融合更新内容	分类名称	基础测绘数据要素数量/个	国情监测数据要素数量/个	几何更新要素数量/个	属性更新要素数量/个	几何属性更新要素数量/个	
水系	重要水系	等级河流	5 068	5 048	3 380	56	3 324	
		水库	1 718	1 581	465	44	421	
	重要水系附属设施	水闸	45 836	15 500	1 346	8 881	457	
		船闸	417	12	2	3	4	
		泵站、干堤等	85 683	67 983	634	12 716	1 034	
居民地及设施	重要设施或标志性建筑	干堤	6 134	5 675	3 679	854	2 918	
		自来水厂	916	1 435	887	69	801	
		污水处理厂	727	888	384	51	335	
		医院	15 941	5 717	1 745	2 829	399	
	学校	9 794	14 135	5 773	1 094	1 509		
飞机场	机场及附属设施	机场	30	12	0	0	1	
铁路	高铁、动车相关铁路	铁路	2 668	453	373	71	302	
交通	高速公路及国省道	等级公路	57 784	61 368	39 393	25 421	13 972	
		火车站	181	135	30	5	24	
	重要交通附属设施	大型交通枢纽,火车站,县级及以上的长途汽车站,高速、国省道相关的大型桥梁、隧道,大中型渡口、大中型码头等	汽车站	722	719	222	46	172
		桥梁	199 020	6 505	17 466	17 427	314	
		隧道	736	188	125	17	119	
		渡口	57	70	53	9	34	
		码头	6 080	2 595	479	363	326	

(续表)

要素名称	融合更新内容	分类名称	基础测绘数据要素数量/个	国情监测数据要素数量/个	几何更新要素数量/个	属性更新要素数量/个	几何属性更新要素数量/个	
县级及以上行政境界	主要是行政区划的调整	省界	64	46	35	0	35	
		市界	36	36	3	0	3	
		区县界	177	175	27	0	27	
境界与行政区	经国家或省级政府颁布的自然保护区、国家森林公园、AAAA 级及以上风景名胜旅游区、世界自然或文化遗产、高新技术开发区、经济开发区、农业开发区、保税区等	自然文化保护区	115	157	139	1	138	
		国有农林、牧场	57	107	83	16	80	
		开发区、保税区	150	197	70	9	61	
地名	乡镇级及以上行政地名	主要是行政区划调整或名称变更后的名称、驻地的更新	乡镇级以上	1 433	930	104	63	62

分析表 1 可知,其中居民地及设施要素中的学校、污水处理厂和自来水厂,水系要素中的泵站、水库,境界与行政区要素中的市界、区县界、开发区、保税区、国有农林、牧场和交通要素中的等级公路、汽车站和火车站,地名要素中的乡镇级以上行政地名等 13 个分类的几何匹配度都在 60% 以上,所以这些要素可以通过几何与属性匹配的方法完成利用 2020 年度国情监测数据更新 2019 年度基础测绘数据(重要要素),而等级河流、干堤和隧道等其他要素由于整体匹配度较低,不建议采用本文融合更新的方法。

实验结果表明,针对几何位置采集标准不一致的点实体匹配,以属性相似度为主,同时辅以距离相似度和环境相似度等条件可提高融合质量;针对属性高度不吻合的线实体,以其中一方数据的属性为准,通过 Fréchet 距离等几何匹配方法可实现数据的准确、快速融合;用空间相似性匹配基于广义 Voronoi 图方法获取的候选数据集可提高面实体的融合精度。

3.2 匹配质量评价

通常根据匹配精度和匹配效率衡量多源矢量数据匹配算法质量的优劣。目前匹配精度应用最为广泛的评价准则是查全率 R 、查准率 P 和综合评价指

标 F 。查全率是指正确匹配的实体数目与数据集全部实体总数的百分比;查准率是指正确匹配的实体数目与实际检索到的匹配实体数目的百分比,综合评价指标兼顾查全率和查准率;匹配效率通过评价得到匹配结果所耗费的时间反映^[17]。 F 的计算公式为:

$$F = \frac{2PR}{P + R} \quad (13)$$

从国情监测和基础测绘数据中选取部分点、线、面数据进行匹配质量评价。候选集获取的评价以面实体匹配为例,其评价结果如表 2 所示;在候选集获取的基础上,点实体、线实体和面实体的匹配评价结果如表 3—5 所示。分析表 2 可知,在不考虑创建搜索区耗时的前提下,基于 Voronoi 图的候选集获取与基于缓冲区或基于 MBR 的候选集获取相比,不管是在匹配精度(综合评价指标提升约 17%)还是在匹配效率(提高近 1 倍)上都有所提高。分析表 3—5 可知,采用本文提出的匹配算法虽然匹配效率有所降低,但匹配精度都得到不同程度的提高。特别是线实体的匹配精度,其综合评价指标提高了近 50%,点实体和面实体的综合评价指标也提高了约 1% 和 4%。

表 2 候选集获取评价

Tab. 2 Evaluation of candidate set acquisition

匹配算法	数据		匹配精度			搜索区域创建时间/s	匹配效率/s
	名称	数量/个	R	P	F		
基于缓冲区和空间相似性的匹配	国情监测数据	1 483	0.503 5	0.732 3	0.596 7	5.7	12.6
	基础测绘数据	1 698					
基于 MBR 和空间相似性的匹配	国情监测数据	1 483	0.493 5	0.794 4	0.608 8	6.9	15.0
	基础测绘数据	1 698					
基于 Voronoi 图和空间相似性的匹配	国情监测数据	1 483	0.612 6	0.813 3	0.698 8	183.0	8.2
	基础测绘数据	1 698					

表 3 点实体匹配质量评价

Tab. 3 Quality evaluation of point entity matching

匹配算法	数据		匹配精度			匹配效率/s
	名称	数量/个	<i>R</i>	<i>P</i>	<i>F</i>	
基于距离相似度的点实体匹配	国情监测数据	21 770	0.535 4	0.730 0	0.617 7	13.8
	基础测绘数据	20 770				
基于距离与环境的点实体匹配	国情监测数据	21 770	0.518 4	0.776 0	0.621 6	22.7
	基础测绘数据	20 770				

表 4 线实体匹配质量评价

Tab. 4 Quality evaluation of line entity matching

匹配算法	数据			匹配精度			匹配效率/s
	名称	原始弧段数/个	构建线段模型	<i>R</i>	<i>P</i>	<i>F</i>	
基于 Fréchet 距离的线实体匹配	国情监测数据	225 060	—	0.501 9	0.580 8	0.538 5	209.3
	基础测绘数据	64 663	—				
基于 Fréchet 距离和线段模型的线实体匹配	国情监测数据	225 060	59 165	0.760 2	0.863 3	0.808 5	613.1
	基础测绘数据	64 663	56 170				

表 5 面实体匹配质量评价

Tab. 5 Quality evaluation of polygon entity matching

匹配算法	数据		匹配精度			匹配效率/s
	名称	数量/个	<i>R</i>	<i>P</i>	<i>F</i>	
基于空间相似性的面实体匹配	国情监测数据	1 483	0.612 6	0.813 3	0.698 8	8.2
	基础测绘数据	1 698				
基于属性和空间相似性的面实体匹配	国情监测数据	1 483	0.603 9	0.903 7	0.724 0	14.1
	基础测绘数据	1 698				

4 结论

针对多源地理信息数据在尺度、几何位置和属性等方面存在不一致性而造成难以融合更新这一现状,提出了一种基于几何与属性匹配的地理信息数据融合更新方法。

1)通过广义 Voronoi 图获取候选集,有效地提高候选集的获取效率并减少不相关目标对候选集的影响。

2)点实体匹配中引入环境相似度,有效解决了点实体匹配中数据集整体呈强覆盖,而部分密集的区域呈现弱覆盖的匹配问题。

3)线实体匹配前通过建立线段模型,抽象为 1:1 匹配类型来处理,可有效提高匹配质量。

4)考虑多个几何特征并结合实体属性进行面实体匹配,可提高 $M:N$ 匹配类型的质量。

通过利用 2020 年度国情监测数据更新 2019 年度基础测绘数据(重要要素)验证了该方法的可行性。研究成果对监测数据更新模式的探索有一定的借鉴意义。本文匹配方法中用到的权重和阈值多数依赖于人工经验,有待进一步研究。

参考文献 (References):

- [1] 公海燕. 地理信息软件的技术进阶与应用创新[J]. 中国测绘, 2020(10):61-63.
- Gong H Y. Technology advancement and application innovation of geographic information software [J]. China Surveying and Mapping, 2020(10):61-63.
- [2] 胡 茜,张 亮,石婷婷. 多源地理要素变化识别研究[J]. 地理空间信息, 2020, 18(6):53-55,7.
- Hu H, Zhang L, Shi T T. Research on the change of multi-source geographical elements [J]. Geospatial Information, 2020, 18(6):53-55,7.
- [3] 李 倩. 我国将构建自然资源统一调查监测体系[J]. 国土资源, 2018(8):14-15.
- Li Q. China will build a unified survey and monitoring system for natural resources [J]. Land and Resources, 2018(8):14-15.
- [4] 王 峰,安晓亚,朱 璇. 地理空间数据增量更新版本化管理方法研究[J]. 地理空间信息, 2021, 19(2):26-29,6.
- Wang F, An X Y, Zhu X. Research on versioning management method for incremental update of geospatial data [J]. Geospatial Information, 2021, 19(2):26-29,6.
- [5] 杨 灿,汪齐松. 多源地理矢量空间数据融合研究[J]. 测绘通报, 2019(12):112-115.
- Yang C, Wang Q S. Research on multi-source geospatial spatial data fusion [J]. Bulletin of Surveying and Mapping, 2019(12):112-115.
- [6] 孙 群. 空间数据相似性研究的若干基本问题[J]. 测绘科学技术学报, 2013, 30(5):439-442.
- Sun Q. Research on some fundamental issues of spatial data similarity [J]. Journal of Geomatics Science and Technology, 2013, 30(5):439-442.
- [7] 安晓亚,孙 群,肖 强,等. 一种形状多级描述方法及在多尺度空间数据几何相似性度量中的应用[J]. 测绘学报, 2011, 40

- (4):495-501,508.
- An X Y, Sun Q, Xiao Q, et al. A shape multilevel description method and application in measuring geometry similarity of multi-scale spatial data [J]. *Acta Geodaetica et Cartographica Sinica*, 2011, 40(4):495-501,508.
- [8] 程钢, 卢小平. 顾及通名语义的汉语地名相似度匹配算法[J]. *测绘学报*, 2014, 43(4):404-410,418.
- Cheng G, Lu X P. Matching algorithm for chinese place names by similarity in consideration of semantics of general names for places [J]. *Acta Geodaetica et Cartographica Sinica*, 2014, 43(4):404-410,418.
- [9] 刘海龙, 钱海忠, 黄智深, 等. 采用 Stroke 层次结构模型的道路网匹配方法[J]. *测绘科学技术学报*, 2013, 30(6):647-651,657.
- Liu H L, Qian H Z, Huang Z S, et al. Road network matching method with stroke-hierarchical model [J]. *Journal of Geomatics Science and Technology*, 2013, 30(6):647-651,657.
- [10] 孙群. 多源矢量空间数据融合处理技术研究进展[J]. *测绘学报*, 2017, 46(10):1627-1636.
- Sun Q. Research on the progress of multi-sources geospatial vector data fusion [J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(10):1627-1636.
- [11] 万洋洋. 基于 Voronoi 图的多尺度面状居民地匹配方法研究[D]. 南昌: 江西师范大学, 2016:23-26.
- Wan Y Y. The matching method based on Voronoi diagram for multi-scale polygonal residential areas [D]. Nanchang: Jiangxi Normal University, 2016:23-26.
- [12] 张桥平, 李德仁, 龚健雅. 地图合并技术[J]. *测绘通报*, 2001(7):6-8.
- Zhang Q P, Li D R, Gong J Y. Map conflation [J]. *Bulletin of Surveying and Mapping*, 2001(7):6-8.
- [13] 程绵绵, 孙群, 季晓林, 等. 改进平均 Fréchet 距离法及在化简评价中的应用[J]. *测绘科学*, 2020, 45(3):170-177.
- Cheng M M, Sun Q, Ji X L, et al. An improved average Fréchet distance method and application in simplification evaluation [J]. *Science of Surveying and Mapping*, 2020, 45(3):170-177.
- [14] 邹静, 陈永刚, 龚金琪, 等. 一种利用多维目标分割比的矢量图形匹配算法[J]. *武汉大学学报(信息科学版)*, 2020, 45(10):1626-1632.
- Zou J, Chen Y G, Gong J Q, et al. An efficient matching algorithm based on vector graphics using multi-dimensional object segmentation ratio [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(10):1626-1632.
- [15] 吴建华. 矢量空间数据实体匹配方法与应用研究[D]. 武汉: 武汉大学, 2008.
- Wu J H. Research on methods of entity matching and its application in vector spatial data [D]. Wuhan: Wuhan University, 2008.
- [16] 郝燕玲, 唐文静, 赵玉新, 等. 基于空间相似性的面实体匹配算法研究[J]. *测绘学报*, 2008, 37(4):501-506.
- Hao Y L, Tang W J, Zhao Y X, et al. Areal feature matching algorithm based on spatial similarity [J]. *Acta Geodaetica et Cartographica Sinica*, 2008, 37(4):501-506.
- [17] 刘善磊, 潘九宝, 李梦梦, 等. 多源地理信息数据匹配质量评价研究[J]. *地理空间信息*, 2022, 20(1):118-121,9.
- Liu S L, Pan J B, Li M M, et al. Research on quality evaluation of multi-source geographic information data matching [J]. *Geospatial Information*, 2022, 20(1):118-121,9.

A geographic data fusion and update method based on geometric and attribute matching

SHI Shanqiu

(Provincial Geomatics Center of Jiangsu, Nanjing 210013, China)

Abstract: The inconsistency of multi-source geographic data in scale, geometric position, and attribute cause difficult data fusion and update. This study proposed a fusion and update method for geographic data based on geometric and attribute matching. First, the candidate set was acquired using the generalized Voronoi diagram, thus effectively improving the acquisition efficiency and reducing the impact of unrelated targets on the candidate set. Then, the matching analysis of point, line, and plane data was made using key techniques such as geometric and attribute matching. Finally, based on the matching results, the incremental data were extracted from the reference geographic information data, followed by fusion and update of target data. The experimental results show that the method proposed in this study can efficiently identify and extract incremental data and serves as a reference for the innovative exploration into the update mode of monitoring data.

Keywords: geometric matching; attribute matching; fusion and update; incremental data

(责任编辑: 陈理)