

# 基于邻近模式的多比例尺居民地松弛迭代匹配

张云菲<sup>1,2</sup> 黄金彩<sup>2</sup> 邓 敏<sup>2</sup> 房晓亮<sup>3</sup> 胡继萍<sup>2</sup>

1 长沙理工大学交通运输工程学院,湖南 长沙,410114

2 中南大学地球科学与信息物理学院,湖南 长沙,410083

3 中国电建集团中南勘测设计研究院有限公司,湖南 长沙,410014

**摘 要:**空间目标匹配是实现多源空间信息融合、空间对象变化检测与动态更新的重要前提。针对多比例尺居民地匹配问题,提出了一种基于邻近模式的松弛迭代匹配方法。该方法首先利用缓冲区分析与空间邻近关系检测候选匹配目标与邻近模式,同时计算候选匹配目标或邻近模式间的几何相似性得到初始匹配概率矩阵;然后对邻近候选匹配对进行上下文兼容性建模,利用松弛迭代方法求解多比例尺居民地的最优匹配模型,选取匹配概率最大并满足上下文一致的候选匹配目标或邻近模式为最终匹配结果。实验结果表明,所提出的多比例尺居民地匹配方法具有较高的匹配精度,能有效克服形状轮廓同质化与非均匀性偏差问题,并准确识别1:M、M:N的复杂匹配关系。

**关键词:**居民地匹配;松弛迭代;邻近模式;多比例尺;空间数据集成

**中图分类号:**P208      **文献标志码:**A

日益加快的城市化进程和飞速发展的数据采集手段促进了地理空间大数据的爆炸式增长<sup>[1]</sup>。随着众包数据的广泛普及,空间数据的获取和更新正经历从传统的自顶向下的主动更新到自底向上的众源更新的转变过程<sup>[2]</sup>。实际应用中,不同部门或个人在不同时间获取的不同精度、尺度、专题的多种空间数据集,彼此间相互独立,难以交互,形成了封闭的“信息孤岛”。如何有效地汇集和整合不同来源的空间数据已成为地理信息科学领域亟待解决的难题<sup>[3]</sup>。而空间目标匹配是实现多源空间信息融合、空间对象变化检测与动态更新的重要前提,其目的是建立不同来源空间数据中表达相同地物的同名目标的对应关系<sup>[4]</sup>。

作为地图表达的重要内容,面目标的匹配问题一直受到国内外学者的广泛关注。一些学者通过比较几何相似性进行面目标匹配<sup>[5-6]</sup>,常用的几何相似性包括距离、形状、面积、方向等<sup>[7]</sup>以及其他一些复杂形状描述算子<sup>[8-9]</sup>。基于几何相似性的匹配方法难以处理居民地形状轮廓同质化问

题,一些研究通过计算邻域结构相似性或拓扑、距离、方向等空间关系相似性来改进基于局部几何相似性的匹配方法<sup>[10-11]</sup>。文献[12]提出了一种居民地的松弛迭代匹配方法,通过上下文兼容性建模,迭代修正基于局部几何的初始匹配,从而得到全局一致的匹配结果。但是,不同比例尺居民地匹配存在较多1:M、M:N的对应关系,基于目标层次的松弛匹配模型可能存在多个目标匹配时各目标分配的匹配概率较低的问题<sup>[13]</sup>。其他学者采用增量式凸壳匹配、目标聚类、群对象探测、提取骨架线网眼的方法来识别不同比例尺的居民地匹配关系<sup>[14-17]</sup>,但匹配精度一定程度上受聚类方法、群对象探测或网眼提取算法的影响。

本文提出了一种改进的松弛迭代匹配方法,该方法扩展了原有匹配模型<sup>[12-13]</sup>,同时计算单个目标或邻近模式的总体匹配概率,并利用邻近上下文兼容关系迭代修正基于局部几何的初始匹配关系,最后基于收敛匹配矩阵选取上下文一致的匹配结果。

1 多比例尺居民地的松弛迭代匹配

松弛迭代匹配的主要思想与地理学第一定律类似,即某一候选匹配对的匹配概率除与自身相似性有关外,还与邻近候选匹配对的匹配概率和相对关系有关<sup>[18]</sup>。因此,松弛迭代匹配一般包括基于局部几何相似性的概率矩阵初始化和顾及邻近兼容关系的松弛迭代两个基本步骤。本文首先利用缓冲区和邻近关系检测候选匹配对(包括单个目标和多个目标构成的邻近模式),并计算候选匹配对间绝对几何相似性,得到初始匹配概率矩阵;然后根据邻近候选匹配对间的相对几何关系,计算邻近候选匹配对间的兼容系数和支持系数,迭代更新初始匹配概率矩阵,直到矩阵对应元素变化量小于给定阈值;最后基于收敛的匹配矩阵,结合冲突处理,选取最终匹配目标。

1.1 基于邻近模式的匹配矩阵初始化

传统的松弛迭代匹配方法通过逐个计算单个候选匹配目标间的几何相似性来估算初始匹配概率矩阵,忽略了 1 : M、M : N 的总体匹配概率<sup>[12]</sup>,本文同时计算单个候选匹配目标或多个目

标组成的邻近模式之间的匹配概率,建立跨比例尺的居民地面目标松弛迭代匹配模型。

1.1.1 候选匹配目标与邻近模式

假设待匹配的两份面状居民地数据为小比例尺数据  $S = \{s_i | i = 1 \cdots m\}$  和大比例尺数据  $T = \{t_j | j = 1 \cdots n\}$ ,其中, $s_i$ 和 $t_j$ 分别为数据  $S$ 、 $T$  中待匹配的居民地多边形, $m$ 和 $n$ 分别为对应的多边形数目。首先对  $S$  中所有的多边形质心构建 Delaunay 三角网,按文献[19]删除全局和局部长边后得到平均边长  $\tau$ ,然后以  $S$  中多边形质心为圆心, $\tau$  为圆半径,检测落在圆内的候选匹配目标集合。

如图 1(a)所示,检测得到  $s_0$  的候选匹配目标集合  $C_{s_0} = \{t_1, t_2, t_3, t_4\}$ ,则  $s_0$  与 4 个候选匹配目标之间存在  $C_1^4 + C_2^4 + C_3^4 + C_4^4 = 2^4 - 1$  种可能的匹配情况(图 1(b)~1(e)分别表示  $s_0$  与 1、2、3、4 个候选目标所有可能的匹配情况)。基于目标层次的松弛迭代方法仅计算  $s_0$  与单个候选目标的匹配概率,且分配至每个候选目标的匹配概率皆较低<sup>[12]</sup>。本文对候选匹配目标进行邻近模式组合,建立单个或多个目标与之匹配的概率矩阵。

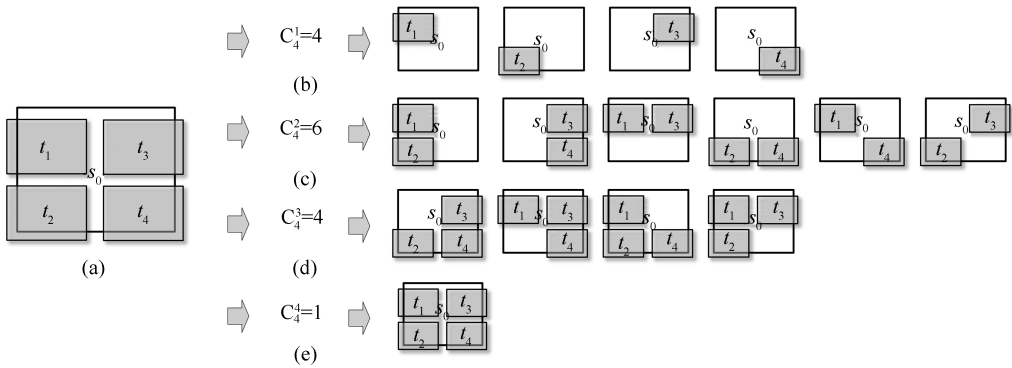


图 1 候选匹配目标的邻近模式  
Fig.1 Neighboring Patterns of Candidate Matching Objects

对于某一目标  $s_i \in S$ ,假设其候选匹配目标集合为  $C_{s_i}$ ,其中包含  $p$  个候选匹配目标,则从  $p$  个候选匹配目标中任意选取  $k$  个( $1 < k \leq p$ ),构成与之匹配的邻近模式集  $C_{p_i}^k$ 。该邻近模式集  $C_{p_i}^k$  中包含  $C_p^k$  种可能的邻近模式情况。如图 1(c)所示,从 4 个候选匹配目标中选取 2 个与  $s_0$  匹配的邻近模式集  $C_{p_i}^2$ ,该集合包含  $C_4^2 = 6$  种可能的组合情况。对于邻近模式集  $C_{p_i}^k$  ( $1 < k \leq p$ ) 中的任一元素,皆由任意  $k$  个候选匹配目标合并得到,本文采用文献[20]的方法来合并候选匹配目标多边形。

众源地理空间数据具有尺度不确定性,存在

匹配方向性问题。本文同样以  $\tau$  为缓冲半径,检测数据  $T$  中各目标  $t_j$  的候选匹配目标集合  $C_{t_j}$  (包含  $q$  个候选匹配目标)及邻近模式  $C_{q_j}^k$  ( $1 < k \leq q$ )。通过候选匹配目标检测与邻近模式分析,确定了初始候选匹配表  $\text{table}_{(m+m') \times (n+n')}$ ,其中  $m'$ 、 $n'$  分别为数据  $S$ 、 $T$  的候选匹配邻近模式数目。

1.1.2 几何相似性与初始匹配概率

对于  $(s_i, t_j) \in \text{table}_{(m+m') \times (n+n')}$ , $s_i$  与  $t_j$  可能为原始数据中单个居民地目标,亦可能是多个目标构成的邻近模式。本文分别计算  $s_i$  与  $t_j$  对应多边形间的几何相似性,相似性计算公式为:

$$\begin{cases} s_{\text{pos}}(s_i, t_j) = 1 - \frac{d(s_i, t_j)}{\tau} \\ s_{\text{dir}}(s_i, t_j) = 1 - \frac{w(s_i, t_j)}{\pi/2} \\ s_{\text{area}}(s_i, t_j) = \frac{\min(a(s_i), a(t_j))}{\max(a(s_i), a(t_j))} \\ s_{\text{shp}}(s_i, t_j) = \frac{\min(e(s_i), e(t_j))}{\max(e(s_i), e(t_j))} \end{cases} \quad (1)$$

式中,  $s_{\text{pos}}(s_i, t_j)$ 、 $s_{\text{dir}}(s_i, t_j)$ 、 $s_{\text{area}}(s_i, t_j)$  和  $s_{\text{shp}}(s_i, t_j)$  分别为候选匹配对  $(s_i, t_j)$  的位置、方向、面积和形状相似性;  $d(s_i, t_j)$  为  $s_i$  与  $t_j$  的质心距离;  $\tau$  为缓冲半径 (见 § 1.1.1);  $w(s_i, t_j)$  为  $s_i$  与  $t_j$  的主方向之差, 归一化为  $[0, \pi/2]$ , 主方向可根据文献[21]计算得到;  $a(\cdot)$  和  $e(\cdot)$  分别为对应多边形的面积和最小外接矩形长宽比。

若考虑所有的邻近模式可能计算量较大, 为剔除明显错误的候选匹配对, 本文首先通过设定相应的位置、方向、面积、形状阈值 (分别为  $T_{\text{pos}}$ 、 $T_{\text{dir}}$ 、 $T_{\text{area}}$ 、 $T_{\text{shp}}$ ), 对  $\text{table}_{(m+m') \times (n+n')}$  中包含邻近模式的初始候选匹配对进行粗筛, 删除相似性小于给定阈值的候选匹配对, 得到最终的候选匹配对表  $\text{table}_{M \times N}$ 。其中,  $M \leq m + m'$ ,  $N \leq n + n'$ 。然后, 计算  $\text{table}_{M \times N}$  中各候选匹配对  $(s_i, t_j)$  的初始匹配概率, 得到初始匹配概率矩阵  $\mathbf{P}^{(0)} = (p_{ij}^{(0)})_{M \times N}$ , 其中,

$$\begin{cases} p_{ij}^{(0)} = \frac{s(s_i, t_j)}{\sum_{t_k \in C_{s_i} \cup C_{p_i}} s(s_i, t_k)} \\ s(s_i, t_j) = s_{\text{pos}} \times s_{\text{dir}} \times s_{\text{area}} \times s_{\text{shp}} \end{cases} \quad (2)$$

式中,  $p_{ij}^{(0)}$  为  $(s_i, t_j)$  的初始匹配概率;  $s(s_i, t_j)$ 、 $s(s_i, t_k)$  分别为  $(s_i, t_j)$  和  $(s_i, t_k)$  的总体相似性;  $s_{\text{pos}}$ 、 $s_{\text{dir}}$ 、 $s_{\text{area}}$ 、 $s_{\text{shp}}$  分别为式 (1) 计算得到的位置、方向、面积和形状相似性;  $C_{s_i}$  和  $C_{p_i}$  分别为  $s_i$  的候选匹配目标集和候选匹配邻近模式集。

1.2 顾及邻近兼容关系的松弛迭代匹配

初始匹配概率仅反映了候选匹配对间的局部一致性, 未考虑邻域上下文结构特征, 而空间目标的匹配概率往往受到周围其他实体匹配程度和邻域结构特征的影响, 因此需要考虑邻近候选匹配对的相对几何关系和匹配程度, 启发式地修正基于局部相似性的初始匹配概率。

1.2.1 面目标邻域关系与相对兼容系数

由于面目标之间不存在显式的邻域关系, 本文通过如下方式确定数据 S 和数据 T 的目标邻域关系<sup>[22]</sup>: 数据 S 的目标邻域关系通过构建 Delaunay 三角网得到, 如图 2(a) 中,  $s_i$  的邻域定义为三角网中与  $s_i$  质心连接的其他面目标及其构

成的邻近模式, 即  $\mathcal{N}(s_i) = \{s_a, s_b, s_c, s_d, s_e, S_{x'}\}$ ; 而数据 T 的目标邻域关系则通过缓冲阈值  $\tau$  得到<sup>[12]</sup>, 如图 2(b) 中,  $t_j$  的邻域定义为与  $t_j$  质心距离小于  $\tau$  的其他面目标及其构成的邻近模式, 即  $\mathcal{N}(t_j) = \{t_f, t_g, t_h, t_p, t_q, T_{y'}\}$ 。

特别地, 邻近模式  $S_u = \{s_1 \cdots s_v\}$  的邻域定义为  $\mathcal{N}(S_u) = \mathcal{N}(s_1) \cup \mathcal{N}(s_2) \cup \cdots \cup \mathcal{N}(s_v) - S_u$ 。如图 2(a) 中,  $\mathcal{N}(s_a) = \{s_b, s_i, s_d \cdots\}$ ,  $\mathcal{N}(s_b) = \{s_a, s_i, s_c \cdots\}$ , 则邻近模式  $S_{x'} = \{s_a, s_b\}$  的邻域定义为  $\mathcal{N}(s_a) \cup \mathcal{N}(s_b)$  中剔除  $S_{x'}$  的目标集合, 即  $\mathcal{N}(S_{x'}) = \mathcal{N}(s_a) \cup \mathcal{N}(s_b) - S_{x'} = \{s_d, s_i, s_c \cdots\}$ 。

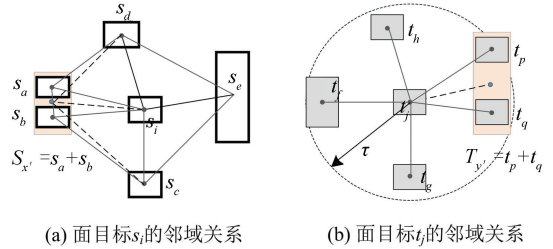


图 2 面目标邻域关系确定  
Fig.2 Determining the Neighboring Relation of Area Objects

邻域关系确定后, 计算邻近候选匹配对间的相对几何关系和支持程度, 启发式地更新基于局部几何的初始匹配矩阵。令  $s_h \in \mathcal{N}(s_i)$ ,  $t_k \in \mathcal{N}(t_j)$ , 则  $s_h$  和  $t_k$  分别表示  $s_i$  和  $t_j$  邻域内的面目标,  $(s_h, t_k)$  是  $(s_i, t_j)$  的邻近候选匹配对。分别计算  $(s_i, t_j)$  与  $(s_h, t_k)$  之间的相对位置、相对方向、相对面积和相对形状关系。相对位置关系计算公式如下:

$$\begin{cases} r_{\text{pos}}(s_i, t_j; s_h, t_k) = r_{\text{dis}} \times r_{\text{dir}} \times r_{\text{ratio}} \\ r_{\text{dis}} = 1 - \frac{|d(s_i, s_h) - d(t_j, t_k)|}{\max_{s_m \in \mathcal{N}(s_i), t_n \in \mathcal{N}(t_j)} (d(s_i, s_m), d(t_j, t_n))} \\ r_{\text{dir}} = 1 - \frac{\theta(s_i s_h, t_j t_k)}{\pi/2} \\ r_{\text{ratio}} = 1 - \frac{d(s_i, s_h)}{\max_{s_m \in \mathcal{N}(s_i)} d(s_i, s_m)} \end{cases} \quad (3)$$

式中,  $r_{\text{pos}}(s_i, t_j; s_h, t_k)$  表示  $(s_i, t_j)$  与  $(s_h, t_k)$  之间的相对位置关系, 由  $r_{\text{dis}}$ 、 $r_{\text{dir}}$ 、 $r_{\text{ratio}}$  相乘得到;  $r_{\text{dis}}$  表示  $(s_i, t_j)$  和  $(s_h, t_k)$  之间的相对几何距离关系, 当  $(s_i, t_j)$  匹配时, 其邻域内面目标  $(s_h, t_k)$  应具有较大的距离相似性;  $r_{\text{dir}}$  表示  $(s_i, t_j)$  和  $(s_h, t_k)$  之间的相对方向关系, 即  $(s_i, s_h)$  和  $(t_j, t_k)$  的角度相似性;  $r_{\text{ratio}}$  表示  $s_i$  邻域内面目标  $s_h$  的权重, 本文认为与  $s_i$  靠近的面目标具有较大的权重;

$d(a, b)$  表示两面目标  $a$  与  $b$  的质心距离;  $\mathcal{N}(\cdot)$  为对应目标的邻域;  $s_i s_h, t_j t_k$  表示两面目标的质心连线;  $\theta(s_i s_h, t_j t_k)$  表示质心连线之间的夹角, 归一化为  $[0, \pi/2]$ , 具体见图 3。相对方向关系计算公式为:

$$r_{\text{dir}}(s_i, t_j; s_h, t_k) = 1 - \frac{|\omega(s_i, s_h) - \omega(t_j, t_k)|}{\pi/2} \quad (4)$$

式中,  $r_{\text{dir}}(s_i, t_j; s_h, t_k)$  表示  $(s_i, t_j)$  与  $(s_h, t_k)$  之间面目标的相对主方向关系;  $\omega(a, b)$  表示两面目标  $a$  与  $b$  的主方向之差, 归一化为  $[0, \pi/2]$ , 具体见图 3。相对面积关系计算公式为:

$$r_{\text{area}}(s_i, t_j; s_h, t_k) = \frac{1}{1 + \left( \frac{a(s_i)}{a(s_h)} - \frac{a(t_j)}{a(t_k)} \right)^2} \quad (5)$$

式中,  $r_{\text{area}}(s_i, t_j; s_h, t_k)$  表示  $(s_i, t_j)$  与  $(s_h, t_k)$  之间的相对面积关系;  $a(\cdot)$  为对应目标的面积。相对形状关系计算公式为:

$$r_{\text{shp}}(s_i, t_j; s_h, t_k) = \frac{1}{1 + \left( \frac{e(s_i)}{e(s_h)} - \frac{e(t_j)}{e(t_k)} \right)^2} \quad (6)$$

式中,  $r_{\text{shp}}(s_i, t_j; s_h, t_k)$  表示  $(s_i, t_j)$  与  $(s_h, t_k)$  之间的相对形状关系;  $e(\cdot)$  为对应目标的最小外接矩形的长宽比。

综合以上 4 种相对几何关系, 得到  $(s_i, t_j)$  与  $(s_h, t_k)$  之间的兼容系数  $K(s_i, t_j; s_h, t_k)$ :

$$K(s_i, t_j; s_h, t_k) = s(s_i, t_j) \times \prod r_{\text{geo}} \quad (7)$$

式中,  $s(s_i, t_j)$  为  $(s_i, t_j)$  的总体相似性;  $\text{geo} = \{\text{pos}, \text{dir}, \text{area}, \text{shp}\}$ , 见式(3)~(6)。

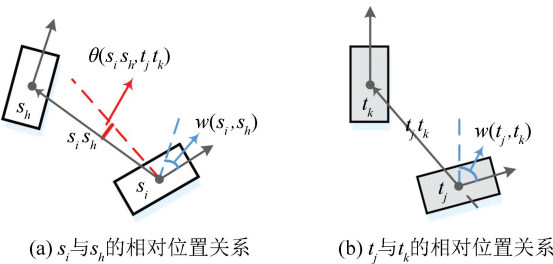


图 3 面目标的相对几何关系计算

Fig.3 Computing the Relative Geometry Relations of Area Objects

### 1.2.2 松弛迭代更新与匹配对选取

兼容系数反映了两两邻近候选匹配对间的一致性关系, 而通常邻近候选匹配对不止一个, 需要综合所有邻近候选匹配对间的兼容系数与匹配概率, 计算总体匹配支持度, 据此更新初始匹配矩

阵。由于邻域要素中可能包含单个目标和多个目标构成的邻近模式, 计算各邻域要素的支持系数时存在重复计算, 需要进行归组合并, 具体步骤为:

1) 计算  $s_i$  的各邻域要素  $s_h \in \mathcal{N}(s_i)$  的支持系数  $q_{ij,h}^{(r)}$ :

$$q_{ij,h}^{(r)} = \max_{t_k \in \mathcal{N}(t_j)} (K(s_i, t_j; s_h, t_k) \cdot P_{hk}^{(r)}) \quad (8)$$

并将各邻域要素的支持系数划分为两个集合  $Q_{ij}$  与  $Q_{ij}'$ , 其中,  $Q_{ij}$  存储单独邻域目标的支持系数,  $Q_{ij}'$  存储  $\mathcal{N}(s_i)$  中邻近模式的支持系数。式(8)中,  $P_{hk}^{(r)}$  为第  $r$  次迭代时  $(s_h, t_k)$  的匹配概率。

2) 按包含目标个数从小到大遍历  $Q_{ij}'$  中各邻近模式及支持系数, 不妨设遍历至  $S_{h'} = \{s_u \cdots s_v\}$  及其支持系数  $q_{ij,h'}^{(r)}$ , 若  $q_{ij,h'}^{(r)}$  满足:

$$q_{ij,h'}^{(r)} > \frac{1}{|S_{h'}|} \sum_{s_u \in S_{h'}} q_{ij,u}^{(r)} \quad (9)$$

则将  $q_{ij,h'}^{(r)}$  替换集合  $Q_{ij}$  中  $\{s_u \cdots s_v\}$  对应的支持系数, 否则继续遍历  $Q_{ij}'$  的下一个邻近模式及支持系数。

3) 当  $Q_{ij}'$  中所有的邻近模式及支持系数皆被判断是否替换到  $Q_{ij}$  后, 计算  $Q_{ij}$  中各支持系数的均值, 即为此次迭代候选匹配对  $(s_i, t_j)$  的总体支持系数  $q_{ij}^{(r)}$ 。

以图 2 为例, 候选匹配对  $(s_i, t_j)$  中  $s_i$  的邻域  $\mathcal{N}(s_i) = \{s_a, s_b, s_c, s_d, s_e, s_{x'}\}$ , 根据式(8)计算各邻域要素的支持系数并划分至集合  $Q_{ij}$  与  $Q_{ij}'$ 。如图 4 所示,  $Q_{ij}$  包含 5 个邻近目标  $s_a, s_b, s_c, s_d$  和  $s_e$ , 其支持系数分别为  $q_{ij,a}^{(r)}, q_{ij,b}^{(r)}, q_{ij,c}^{(r)}, q_{ij,d}^{(r)}$  和  $q_{ij,e}^{(r)}$ ;  $Q_{ij}'$  包含 1 个邻近模式  $S_{h'} = \{s_a, s_b\}$  及其支持系数  $q_{ij,h'}^{(r)}$ 。由于  $q_{ij,h'}^{(r)} > (q_{ij,a}^{(r)} + q_{ij,b}^{(r)})/2$ , 故  $q_{ij,h'}^{(r)}$  将替换  $Q_{ij}$  中  $s_a$  和  $s_b$  对应的支持系数  $q_{ij,a}^{(r)}$  和  $q_{ij,b}^{(r)}$ 。当  $Q_{ij}'$  中所有的邻近模式皆被判断后,  $Q_{ij}$  中支持系数的均值即为候选匹配对  $(s_i, t_j)$  的总体匹配支持程度。

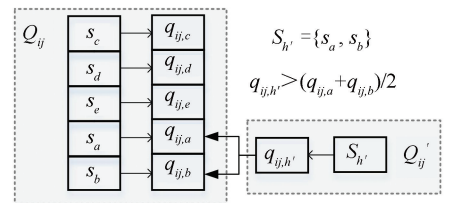


图 4 总体支持程度计算

Fig.4 Calculating the Total Support Degrees

根据以上方法计算各候选匹配对的总体匹配支持程度, 按照式(10)<sup>[23]</sup> 迭代更新匹配矩阵:

$$p_{ij}^{(r+1)} = \frac{p_{ij}^{(r)} + q_{ij}^{(r)}}{1 + \sum_{t_k \in C_{s_i} \cup C_{p_i}} q_{ik}^{(r)}} \quad (10)$$



式中,  $p_{ij}^{(r)}$  和  $p_{ij}^{(r+1)}$  分别为第  $r, r+1$  次迭代时( $s_i, t_j$ )的匹配概率;  $q_{ij}^{(r)}$  和  $q_{ik}^{(r)}$  分别为  $r$  次迭代时( $s_i, t_j$ )与( $s_i, t_k$ )的总体匹配支持度;  $C_{s_i}$  和  $C_{p_i}$  分别为  $s_i$  的候选匹配目标和邻近模式集合。当两次迭代期间各候选匹配对的匹配概率变化量小于给定的阈值  $\alpha$  时, 则迭代停止。

基于收敛的概率矩阵选取最终匹配结果。首先选择概率矩阵的每行和每列中匹配概率最大的候选匹配对, 加入队列 MQueue; 然后逐一判断 MQueue 的各匹配对, 当与其他匹配对存在公共的面目标时, 计算两组匹配合并前后的几何相似度, 若合并后的相似度大于合并前的, 则将两者合并为多对多匹配; 否则, 保留概率较高的匹配对, 概率较低的则选择概率次之的匹配目标或邻近模式。

表 1 实验数据统计与参数设置  
Tab.1 Experimental Data Statistics and Parameter Settings

实验数据	数据 S/	数据 T/	位置阈值	方向阈值	面积阈值	形状阈值	迭代阈值
	面目标数目	面目标数目	$T_{pos}$	$T_{dir}$	$T_{area}$	$T_{shp}$	$\alpha$
模拟数据	小比例尺/25	大比例尺/42	0.75	0.95	0.80	0.75	0.000 5
中国西安市数据	1 : 2.5 万/83	1 : 2 万/118	0.70	0.95	0.65	0.72	0.000 5
美国达拉斯市数据	OpenStreetMap/83	网络数据/80	0.87	0.93	0.90	0.87	0.000 5

注: 美国达拉斯市数据的下载地址分别为: OpenStreetMap 为 <http://download.geofabrik.de/north-america/us/texas.html>; 网络数据为 <https://gis.dallascityhall.com/shapezip.htm>。

图 5 显示了模拟居民地数据的匹配结果。图 5(a)为初始候选匹配对检测, 由于存在非均匀性偏差, 基于单一距离阈值检测得到多个明显错误的匹配对; 图 5(b)为利用文献[12]提出的松弛迭代匹配方法得到的结果, 说明通过上下文兼容性

## 2 实验结果与分析

为了验证本文方法的有效性和可靠性, 分别采用一份模拟数据、两份分别来自中国西安市和美国达拉斯市的居民地数据进行实验分析。如表 1 所示, 模拟居民地数据存在尺度差异, 但各自比例尺不确定; 中国西安市居民地数据具有不同比例尺, 分别是 1 : 2.5 万与 1 : 2 万; 美国达拉斯市居民地数据为非专业测绘数据, 存在尺度差异, 但比例尺不确定。表 1 中,  $T_{pos}$ 、 $T_{dir}$ 、 $T_{area}$ 、 $T_{shp}$  分别为邻近模式粗筛时使用的位置、方向、面积、形状阈值;  $\alpha$  为松弛迭代的收敛阈值, 统一设为 0.000 5。

建模能够有效克服建筑物形状轮廓同质化与非均匀偏差的匹配问题, 但基于单个目标相似性进行优化匹配, 存在一对多、多对多的漏匹配情况; 图 5(c)为本文匹配结果, 可以看到通过邻近模式组合能够有效识别 1 : M、M : N 的匹配关系。

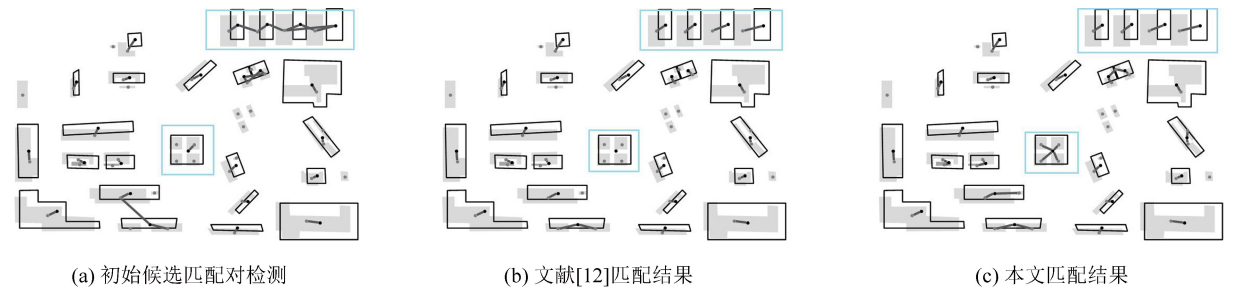


图 5 模拟居民地数据的匹配结果对比

Fig.5 Comparison of Matching Results of Simulated Residential Data

图 6(a)和图 6(b)分别显示了中国西安市和美国达拉斯市实验区数据的匹配结果。结果表明, 本文方法能够有效处理具有明显尺度差异的多源居民地数据匹配, 同时对尺度不确定的非专业测绘数据具有较好的匹配效果。

通过与人工匹配结果进行对比, 定量评价本文方法的匹配精度。3 种精度评价指标分别为准确率  $P$ 、召回率  $R$  和  $F$  值, 其计算公式为:

$$P = \frac{t_p}{t_p + f_p} \times 100\%, R = \frac{t_p}{t_p + f_n} \times 100\%, F = \frac{2PR}{P + R} \quad (11)$$

式中,  $t_p$  为正确匹配数;  $f_p$  为错误匹配数;  $f_n$  为漏匹配数。

从表 2 可以看出, 本文方法的准确率、召回率和  $F$  值皆高于 90%。而且使用模拟数据将本文

方法与文献[12]的方法进行对比,结果发现本文方法能够进一步改善基于目标层次的松弛迭代的匹配精度。

表 2 实验精度评价/%

Tab.2 Statistics of Precision, Recall and F Value/%

数据	方法	准确率 $P$	召回率 $R$	$F$ 值
模拟数据	文献[12]方法	100	68.57	81.35
	本文方法	100	100	100
中国西安市数据	本文方法	93.56	91.47	92.5
美国达拉斯市数据	本文方法	98.12	95.45	96.93

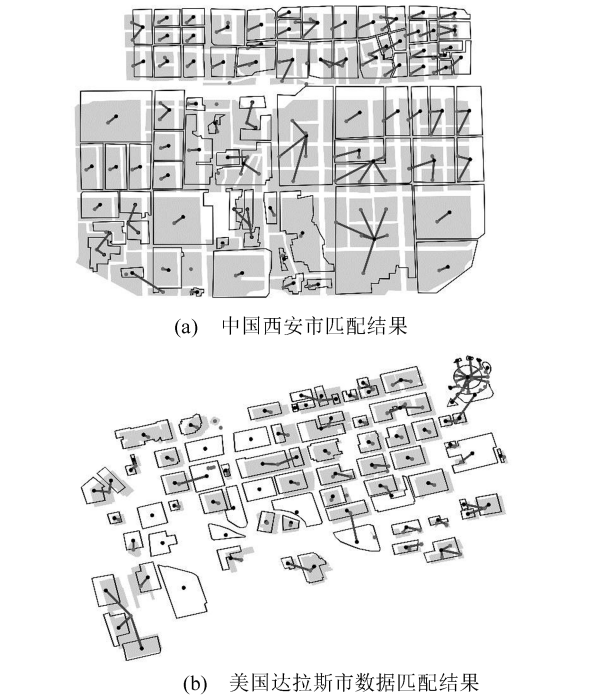


图 6 中国西安市和美国达拉斯市实验区数据匹配结果  
Fig.6 Matching Results of Test Datasets of Xi'an and Dallas

图 7 列出了居民地面目标  $C$  与候选匹配目标  $b$ 、 $c$ 、 $d$  在迭代过程中的概率变化情况。 $C$  与  $d$  几何相似性最大,初始匹配概率最高,而通过松弛迭代, $C$  与  $d$  的匹配概率不断下降, $C$  与  $c$  的匹配概率逐渐升高并达到最大,从而实现匹配结果上下文一致。说明顾及邻近兼容关系的松弛迭代匹配方法可以修正基于几何相似性的局部最优。

表 3 给出了两个邻近模式的匹配示例。对于示例 1,居民地  $A_1$  包含  $a_1$ 、 $a_2$ 、 $a_3$ 、 $a_4$  共 4 个候选匹配目标,通过邻近模式组合计算得到  $A_1$  与  $\{a_1, a_2, a_3, a_4\}$  的总体匹配概率为 0.218,明显高于其他匹配目标或组合。对于示例 2,  $B_1$  和  $B_2$  包含  $b_1$ 、 $b_2$ 、 $b_3$  共 3 个候选匹配目标,同样采用顾及邻近模式的匹配模型构建方法,识别了  $B_1 \rightarrow \{b_1, b_2\}$  和  $B_2 \rightarrow \{b_1, b_3\}$  两个  $1:M$  的匹配关系。而由于两个匹配关系存在公共面目标  $b_1$ ,因此通过

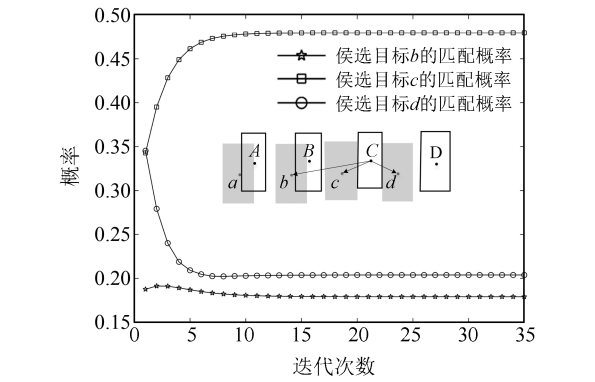


图 7 松弛迭代过程的概率变化  
Fig.7 Probability Change in Relaxation Labelling

几何相似性比较,将它们合并为一个  $M:N$  的匹配对,使最终匹配结果为  $\{B_1, B_2\} \rightarrow \{b_1, b_2, b_3\}$ 。说明本文提出的基于邻近模式的松弛迭代匹配方法能够准确识别  $1:M$ 、 $M:N$  的复杂匹配关系。

表 3 邻近模式的匹配示例

Tab.3 Matching Examples of Neighboring Pattern Combination

组合示例	数据 $T$	候选匹配组合	匹配概率
示例 1 	$A_1$	$\{a_1\}$	0.011
		$\{a_1, a_2\}$	0.046 9
		$\{a_1, a_2, a_3\}$	0.138
		$\{a_1, a_2, a_3, a_4\}$	<b>0.218</b>
		$\{b_1\}$	0.058 5
示例 2 	$B_1$	$\{b_2\}$	0.142 4
		$\{b_1, b_2\}$	<b>0.308</b>
		$\{b_1, b_2, b_3\}$	0.051
	$B_2$	$\{b_1\}$	0.070
		$\{b_3\}$	0.130
		$\{b_1, b_3\}$	<b>0.290</b>
		$\{b_1, b_2, b_3\}$	0.055

### 3 结 语

针对基于目标层次的松弛迭代匹配方法无法准确完整地识别  $1:M$ 、 $M:N$  的对应关系,本文提出了一种顾及邻近模式的多比例尺居民地匹配方法,该方法扩展了原有松弛迭代匹配模型,同时计算单个目标和多个目标组合的匹配概率和上下文关系,以达到多比例尺居民地数据全局一致性匹配的目的。实验表明,该方法能够准确识别多比例尺居民地中  $1:M$ 、 $M:N$  的匹配关系,同时对建筑物形状轮廓同质化、非均匀偏差较大的居民地数据保持了可靠的匹配精度,为居民地变化发现与动态更新提供重要的技术支撑。但是,当待匹配数据间比例尺差异极大时,单个目标可能

有多个候选匹配目标或邻近模式,从而大幅度增加了本文算法的计算复杂度。因此,下一步研究主要集中于邻近模式的优化筛选方法和空匹配建模方面,使之能够适应更大比例尺跨度。

### 参 考 文 献

- [1] Li Qingquan, Li Deren. Big Data GIS[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(6): 641-644(李清泉, 李德仁. 大数据 GIS [J]. 武汉大学学报·信息科学版, 2014, 39(6): 641-644)
- [2] Jiang B. Volunteered Geographic Information and Computational Geography: New Perspectives[M]// Sui D, Elwood S, Goodchild M. *Crowdsourcing Geographic Knowledge*. Dordrecht: Springer, 2013
- [3] Chen Jun, Li Zhilin, Jiang Jie, et al. Key Issues of Continuous Updating of Geo-spatial Databases[J]. *Geomatics World*, 2004, 2(5): 1-5(陈军, 李志林, 蒋捷, 等. 基础地理数据库的持续更新问题[J]. 地理信息世界, 2004, 2(5): 1-5)
- [4] Zhang Yunfei, Yang Bisheng, Luan Xuechen. Automated Matching Urban Road Networks Using Probabilistic Relaxation[J]. *Acta Geodaetica et Cartographica Sinica*, 2012, 41(6): 933-939(张云菲, 杨必胜, 栾学晨. 利用概率松弛法的城市路网自动匹配[J]. 测绘学报, 2012, 41(6): 933-939)
- [5] Zhang Qiaoping, Li Deren, Gong Jianya. Areal Feature Matching Among Urban Geographic Databases [J]. *Journal of Remote Sensing*, 2004, 8(2): 107-112(张桥平, 李德仁, 龚健雅. 城市地图数据库面实体匹配技术[J]. 遥感学报, 2004, 8(2): 107-112)
- [6] Zhao Dongbao, Sheng Yehua, Zhang Ka. An Algorithm for Multi-Scale One-to-many Areal Feature Matching Based on Geometry Moments and Overly Analysis[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(11): 1 371-1 375(赵东保, 盛业华, 张卡. 利用几何矩和叠置分析进行多尺度面要素自动匹配[J]. 武汉大学学报·信息科学版, 2011, 36(11): 1 371-1 375)
- [7] Hao Yanling, Tang Wenjing, Zhao Yuxin, et al. Areal Feature Matching Algorithm Based on Spatial Similarity[J]. *Acta Geodaetica et Cartographica Sinica*, 2008, 37(4): 501-506(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)
- [8] Fan H, Zipf A, Fu Q, et al. Quality Assessment for Building Footprints Data on OpenStreetMap[J]. *International Journal of Geographical Information Science*, 2014, 28(4): 700-719
- [9] Ai T, Cheng X, Liu P, et al. A Shape Analysis and Template Matching of Building Features by the Fourier Transform Method[J]. *Computers, Environment and Urban Systems*, 2013, 41(5): 219-233
- [10] Xu Junkui, Wu Fang, Qian Haizhong, et al. Settlement Matching Algorithm Using Spatial Similarity Relations as Constraints [J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(4): 484-488(许俊奎, 武芳, 钱海忠, 等. 一种空间关系相似性约束的居民地匹配算法[J]. 武汉大学学报·信息科学版, 2013, 38(4): 484-488)
- [11] Jiao Yangyang, Wang Hui, Zhai Renjian. Matching Method for Areal Feature Considering Context-Dependent Similarity [J]. *Journal of Liaoning Technical University (Natural Science)*, 2013, 32(7): 947-952(焦洋洋, 王卉, 翟仁健. 顾及邻域相似性的面要素匹配方法[J]. 辽宁工程技术大学学报(自然科学版), 2013, 32(7): 947-952)
- [12] Zhang X, Ai T, Stoter J, et al. Data Matching of Building Polygons at Multiple Map Scales Improved by Contextual Information and Relaxation [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 92(6): 147-163
- [13] Hu Jiping. Area Target Data Matching Method of Different Scales[J]. *Applied Technology*, 2016(4): 89-92(胡继萍. 不同比例尺面目标数据匹配方法[J]. 科技创新与生产力, 2016(4): 89-92)
- [14] Zhang Liping, Guo Qingsheng, Sun Yan. The Method of Matching Residential Features in Topographic Maps at Neighboring Scales [J]. *Geomatics and Information Science of Wuhan University*, 2008, 33(6): 604-607(章莉萍, 郭庆胜, 孙艳. 相邻比例尺地形图之间居民地要素匹配方法研究[J]. 武汉大学学报·信息科学版, 2008, 33(6): 604-607)
- [15] Huh Y, Kim J, Lee J, et al. Identification of Multi-scale Corresponding Object-Set Pairs Between Two Polygon Datasets with Hierarchical Co-clustering [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 88(2): 60-68
- [16] Xu Junkui, Wu Fang, Zhu Jiandong, et al. A Multi-to-Multi Matching Algorithm Between Neighborhood Scale Settlement Data[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(3): 340-345(许俊奎, 武芳, 朱建东, 等. 相邻比例尺居民地匹配[J]. 武汉大学学报·信息科学版, 2014, 39(3): 340-345)
- [17] Wang Xiao, Qian Haizhong, He Haiwei, et al. Matching Multi-source Areal Habitations with Skeleton Line Mesh of Blank Region[J]. *Acta Geodæt-*

ica et Cartographica Sinica, 2015, 44(8): 927-935  
(王骁, 钱海忠, 何海威, 等. 利用空白区域骨架线网眼匹配多源面状居民地[J]. 测绘学报, 2015, 44(8): 927-935)

[18] Yang B, Zhang Y, Luan X. A Probabilistic Relaxation Approach for Matching Road Networks[J]. *International Journal of Geographical Information Science*, 2013, 27(2):319-338

[19] Deng M, Liu Q, Cheng T, et al. An Adaptive Spatial Clustering Algorithm Based on Delaunay Triangulation[J]. *Computers, Environment and Urban Systems*, 2011, 35(4): 320-332

[20] Zhang Jing, Zhou Ye, Liu Yu. An Improved Algorithm for SDS Model Based Polygon Simplification and Aggregation [J]. *Journal of Image and Graphics*, 2006, 11(7):1 010-1 016(张晶,周烨,刘瑜. SDS 模型化简合并多边形的一个改进算法研究[J].中国图象图形学报, 2006, 11(7): 1 010-1 016)

[21] Duchène C, Bard S, Barillot X, et al. Quantitative and Qualitative Description of Building Orientation [C]. *The 5th Workshop on Progress in Automated Map Generalisation*, Barcelona, 2003

[22] Kim J O, Yu K, Heo J, et al. A New Method for Matching Objects in Two Different Geospatial Datasets Based on the Geographic Context[J].*Computers & Geosciences*, 2010, 36(9): 1 115-1 122

[23] Parent P, Zucker S W. Radial Projection: An Efficient Update Rule for Relaxation Labeling [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11(8): 886-889

Relaxation Labelling Matching for Multi-scale Residential Datasets  
Based on Neighboring Patterns

ZHANG Yunfei<sup>1,2</sup> HUANG Jincai<sup>2</sup> DENG Min<sup>2</sup> FANG Xiaoliang<sup>3</sup> HU Jiping<sup>2</sup>

1 School of Traffic and Transportation Engineering, Changsha University of Science & Technology, Changsha 410114, China

2 School of Geosciences and Info-physics, Central South University, Changsha 410083, China

3 Zhongnan Engineering Corporation Limited, Power China, Changsha 410014, China

**Abstract:** This paper proposes a relaxation labelling matching approach for multi-scale residential datasets based on neighboring patterns. Firstly, we detect the candidate matching objects and neighboring patterns by buffering analysis and spatial neighboring relations. Secondly, the geometric similarities of candidate matching objects or neighboring patterns are calculated to initialize the matching matrix that contains  $1 : 1$ ,  $1 : M$  and  $M : N$  relations. After that, the contextual information of neighborhood objects or patterns are explored to heuristically update the matching matrix to achieve a global consistency. The matching pairs with maximum probabilities are finally selected after context consistency detection. The experimental results and contrast analysis show that our method obtains high correct matching rates, efficiently overcomes the problems of shape homogeneity and uneven deviation, and can correctly identify complex  $1 : M$  and  $M : N$  matching objects in multi-scale residential datasets.

**Key words:** residential data matching; relaxation labelling; neighboring patterns; multiple scales; spatial data integration

**First author:** ZHANG Yunfei, PhD, specializes in multi-source and multi-scale spatio-temporal data conflation and updating. E-mail: zhangyunfei@csu.edu.cn

**Corresponding author:** DENG Min, professor. E-mail: dengmin@csu.edu.cn

**Foundation support:** The National Natural Science Foundation of China, Nos. 41601495, 41471385, 41501442; China Postdoctoral Science Foundation, No. 2015M582345; the Open Fund of State Key Laboratory of Resources and Environmental Information System; the Open Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, No. 17S01.