

文章编号:1671-8860(2010)08-0913-04

文献标志码:A

# 基于结构化空间关系信息的结点层次匹配方法

邓 敏<sup>1</sup> 徐 凯<sup>1</sup> 赵彬彬<sup>1,2</sup> 徐 锐<sup>1</sup>

(1 中南大学信息物理工程学院, 长沙市麓山南路 932 号, 410083)

(2 长沙理工大学交通运输工程学院, 长沙市赤岭路 45 号, 410076)

**摘要:** 提出了一种基于结构化空间关系信息的结点层次匹配方法。指出了在结点匹配中存在的不足, 在此基础上, 顾及结点及其连接弧段的结构特征, 结合距离、拓扑和方向三类约束, 对候选匹配点进行逐层筛选, 以得到最佳匹配结果。实验表明, 该匹配方法是一种行之有效的方法。

**关键词:** 结点; 匹配; 层次性; 结构化空间关系

**中图法分类号:**P208

在地理信息科学领域, 匹配技术已广泛应用于空间数据集成或融合<sup>[1, 2]</sup>、空间数据质量改善和评价<sup>[3, 4]</sup>、多尺度空间数据库的维护和更新<sup>[5, 6]</sup>、基于位置服务的导航<sup>[7]</sup>等方面。其中在多源空间数据集成或融合方面的应用较为广泛, 重点是解决不同来源的空间数据间的不一致性问题, 实现属性数据间的转换、综合空间分析等。1988 年, Saalfeld 首次利用匹配方法实现了地图数据的融合<sup>[8]</sup>, 该方法利用距离、连通度和蜘蛛编码来寻找可能的匹配点对, 并且只能处理 1:1 匹配关系。随后, 许多匹配算法相继被提出来<sup>[9-14]</sup>。分析已有的研究可以发现, 在空间数据集成、更新中, 目标匹配是其中的一个关键技术, 匹配指标(约束条件)的选取和运用的先后顺序对匹配结果

的优劣有着重要的影响。

本文主要探讨不同比例尺空间数据集成或更新中结点的匹配问题, 首先分析结点的结构特性, 并对结点拓扑分类, 指出现有描述方法存在的问题, 进而提出一种基于结构化空间关系的结点层次匹配方法。

## 1 结点拓扑分类、描述方法及存在的问题

### 1.1 结点拓扑分类

连通度是一个用来描述图形拓扑结构的基本拓扑不变量。文献[13]利用连通度将结点的拓扑关系分为 6 种类型, 如表 1 所示。

表 1 利用连通度对点的类型进行区分和描述

Tab. 1 Classification of Point Features in Terms of Connectivity Degree

| 示例  | 1 | 2 | 3 | 4 | 5 | 6 |
|-----|---|---|---|---|---|---|
| 连通度 | 1 | 2 | 3 | 4 | 5 | 6 |

但是在地图目标匹配中, 仅仅根据连通度来进行结点匹配, 容易发生误配的情况。如图 1 所示, 若仅根据一定距离范围内候选匹配点的连通度这一指标, 那么将误判为  $P_2$  与  $Q_1$  匹配。为了避免这种情形, 则需要对上述分类进行扩展, 从而更精确地匹配结点。

### 1.2 结点类型的细化和区分方法

针对图 1 的误配情况, 文献[15]提出了利用三条连接边之间的夹角来辅助判定候选匹配点。该方法将结点连通度为 3 的类型细分为三种(见表 2 第 2 列):① 三条连接边之间的夹角都是钝角, 或者说, 三个夹角都大于  $90^\circ$ ; ② 三条连接边

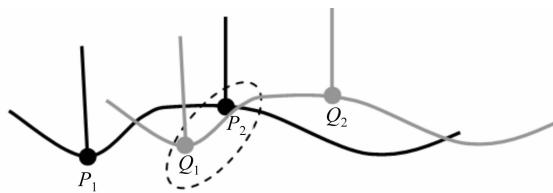


图 1 仅利用连通度指标导致误配的情形  
Fig. 1 A Case of Incorrect Matching Only by Considering the Degree of Connectivity

之间的三个夹角中有两个小于  $90^\circ$ ; ③ 三条连接边的夹角中存在一个夹角为平角, 即等于  $180^\circ$ 。但是, 在计算结点连接边之间的夹角时, 有互为周角的两个角度之分, 易产生歧义。为此, 本文提出利用邻域与连接边的交点连成的三角形类型来区分这三种不同的情形(见表 2 第三列), 即锐角三角形、钝角三角形和直角三角形。

表 2 结点连通度为 3 的三种情况

Tab. 2 Three Cases of Nodes with the Degree of Connectivity 3

| 图形 | 三条边之间的夹角             | 三角形类型 | 类 型 |
|----|----------------------|-------|-----|
|    | 至少有两个角大于 $90^\circ$  | 锐角三角形 | 1   |
|    | 三夹角中有两个小于 $90^\circ$ | 钝角三角形 | 2   |
|    | 存在一个角为 $180^\circ$   | 直角三角形 | 3   |

类似地, 对于结点连通度为 4 的情况, 可以根据文献[16]提出的交点局部顺序来进行区分, 分为两种类型(见表 3)。对于一个结点  $P_i$ , 其局部顺序  $L_o(P_i)$  定义为: 以  $P_i$  为圆心, 以无穷小正数  $\epsilon$  为半径画一个圆(即邻域), 则该圆与两个线目标的交点在圆上的排序即为点  $P_i$  的局部顺序。

表 3 结点连通度为 4 的两种情况

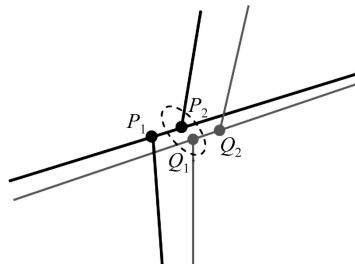
Tab. 3 Two Cases of Nodes with the Degree of Connectivity Equal to Four

| 图形 | 局部顺序 $L_o(P_i)$           | 类 型 |
|----|---------------------------|-----|
|    | $L_o(P_i) = A, B, A, B$   | 1   |
|    | $L_o(P_i) = (A, B, B, A)$ | 2   |

### 1.3 存在的问题及本文的研究策略

根据现有的结点区分及描述方法, 仍有可能出现误配的情况。如图 2 所示, 根据距离约束、连

通度约束以及角度(或三角形类型)约束, 点  $P_2$  可能与点  $Q_1$  匹配。事实上, 正确的匹配对应该是  $P_1 \rightarrow Q_1, P_2 \rightarrow Q_2$ 。通过比较  $P_2$  和  $Q_1$  可以发现, 它们连接的三条弧段中有两条弧段的方向基本是一致的, 而另一条是明显不一致的, 因此可判定  $P_2$  与  $Q_1$  非同名结点(即表达现实中同一地物的结点)。



Distance( $P_2, Q_1$ )  $\leq d_0$ ; Topo\_Node( $P_2$ ) = Topo\_Node( $Q_1$ ) = 3;  
Triangle\_Type( $P_2$ ) = Triangle\_Type( $Q_1$ ) =  $T_3$

图 2 基于距离、连通度和角度约束的错误匹配情形  
Fig. 2 A Case of Incorrect Matching Based on the Constraints of Distance, the Degree of Connectivity and Triangle Type

为了发展一种新的结点匹配方法, 本文提出一个新的匹配策略, 即层次匹配策略。首先, 定义三种类型空间关系约束, 分别为: ① 距离约束, 即待匹配点对的距离小于距离阈值; ② 结点连通度约束(亦称拓扑约束), 即待匹配结点对的连通度相同或小于给定的差异; ③ 结点连接的对应弧段方向小于给定的差异(亦称方向约束)。然后依次考虑这三类约束, 即以①  $\rightarrow$  ②  $\rightarrow$  ③ 的顺序来进行结点层次过滤, 达到最终正确匹配的目的。

## 2 基于结构化空间关系的结点层次匹配方法

### 2.1 利用距离约束来确定候选匹配点

匹配过程实质上是一个过滤的过程。对于参考数据集中的一个结点  $P_i$ , 首先给定一个距离阈值  $d_0$ , 它可以通过对一些已知的(或非常明显的)结点匹配对进行抽样统计分析得到, 具体数值与数据本身的精度、目标类型(即人工建造或自然形成)密切相关。然后在目标数据集中寻找距离阈值范围内的所有结点或顶点, 依次记为  $Q_{i1}, Q_{i2}, \dots, Q_{in}$ , 它们构成候选匹配点集(记为  $\Omega_{i1}$ )。这里使用的距离为欧氏距离, 即满足:

$$d(P_i, Q_{ij}) = ((x_{P_i} - X_{Q_{ij}})^2 + (y_{P_i} - y_{Q_{ij}})^2)^{1/2} \leq d_0 \quad (1)$$

## 2.2 利用结点连通度约束来确定候选匹配点

利用连通度约束条件对上述得到的候选匹配点集  $\Omega_{i1}$  进行筛选。由于参考数据集和目标数据集的比例尺、现势性可能不同, 导致结点的连通度具有一定的差异。本文定义结点拓扑约束为:

$$|\text{Topo\_Node}(P_i) - \text{Topo\_Node}(Q_j)| \leq 1 \quad (2)$$

于是, 连通度差异不大于 1 的结点为候选匹配点。根据式(1)和式(2)得到结点  $P_i$  的候选匹配点集为  $\Omega_{i2}$ 。

## 2.3 利用结点连接弧段方向来确定候选匹配点

引入结点连接弧段方向约束。首先, 对结点弧段的方向进行定义。方向描述可以分为定性和定量两种。定性描述有 4 方向(如东方向)、8 方向(如东南方向)或 16 方向(如东南南方向)。定量描述通常用方位角表示, 即以正北方向为零指向, 按顺时针方向依次递增, 取值为  $[0^\circ, 360^\circ]$ 。本文采用定量描述, 并以参考(或候选匹配)结点为起点, 来表达结点连接弧段的指向或方向。于是, 可以定义相应的连接弧段的方向满足约束:

$$|\text{dir}(P_i^k - Q_j^k)| \leq \delta_0 \quad (3)$$

式中,  $k$  为第  $k$  个相应的连接弧段;  $\delta_0$  为候选匹配连接弧段方向间的差异阈值。根据文献[13], 通常可取  $15^\circ$ 。实践中, 需要考虑数据来源、数据本身的精度以及比例尺跨度(或差异)等情况进行设定。如果  $P_i$  和  $Q_j$  所有的相应弧段的方向都满足式(3), 则认为  $P_i$  和  $Q_j$  是一对匹配点目标。

综上所述, 基于结构化空间关系的结点层次匹配过程可描述为: ①首先通过创建缓冲区(即距离约束)寻找到候选匹配对, 获得候选匹配点集  $\Omega_{i1}$ ; ②进行拓扑分析, 在寻找到的匹配点集中分析结点的连通度, 剔除拓扑不一致的候选匹配点, 获得候选匹配点集  $\Omega_{i2}$ ; ③定量方向判断, 比较结点连接弧段的方位角的差异是否在阈值范围内, 最终筛选得出精确的匹配结果, 具体流程见图 3。

## 3 算例分析

根据以上的层次匹配策略, 本文采用 ArcGIS Engine, 并在.NET 环境下实现了基于结构化空间关系的结点匹配。如图 4 所示, 将参考结点和目标结点叠加, 并根据所提出的结点匹配算法得到的匹配结果如图 5 所示。

以图 5(a)中的  $R_{11}$  点目标为例, 详细分析其匹配过程, 如表 4 所示。由距离约束获得  $R_{11}$  点

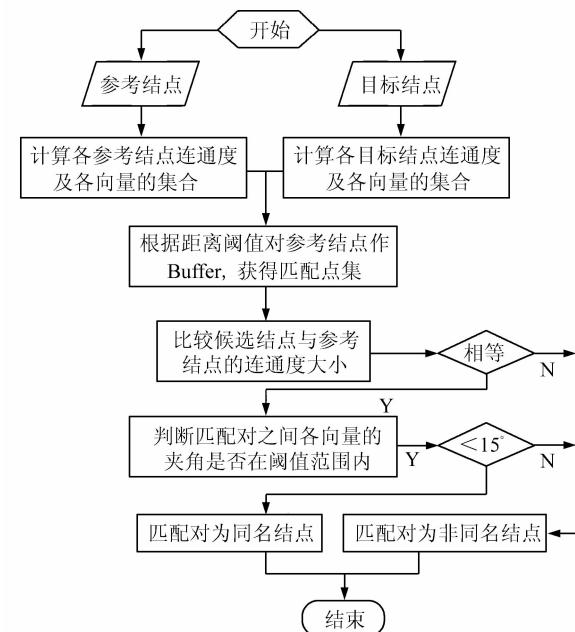


图 3 结点匹配流程图

Fig. 3 Process of the Matching Between Reference Node and Target Nodes

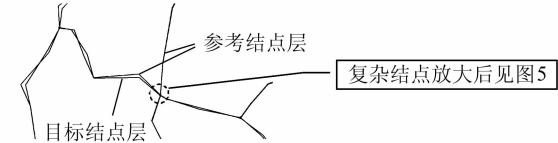


图 4 参考结点与目标结点叠加

Fig. 4 Overlapping of Reference Nodes and Target Nodes

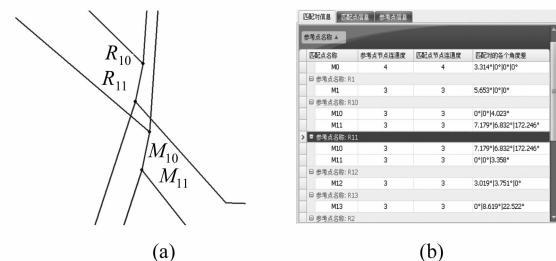


图 5 匹配结果

Fig. 5 Matching Result

目标的可能匹配对象为  $M_{10}$  和  $M_{11}$ , 进而分析参考点目标和候选匹配点目标的结点类型, 都是连通度为 3 的第 3 种类型。进一步分析结点的各向量方向, 由于  $R_{11}$  与  $M_{10}$  存在一个向量夹角为  $172.246^\circ$ (如图 5(b)所示), 大于阈值  $15^\circ$ , 而  $R_{11}$  与  $M_{11}$  的所有向量夹角都在阈值范围内, 故  $R_{11}$  与  $M_{11}$  匹配(程序显示中,  $7.179^\circ | 6.832^\circ | 172.246^\circ$  表示三个向量的夹角之差)。

表4 结点层次匹配方法示例

Tab. 4 Example of Hierarchical Method for Nodes Matching

| 结点类型   | 结点连通度 | 结点的各向量方向  | 匹配结果                    |
|--|-------|---|-------------------------|
| $R_{11}$ Topo_Node( $R_{11}$ )=3; Triangle_Type( $R_{11}$ )= $T_3$ | 3     |   |                         |
| $M_{10}$ Topo_Node( $M_{10}$ )=3; Triangle_Type( $M_{10}$ )= $T_3$ | 3     | $R_{11}$ 与 $M_{10}$ 存在一个向量夹角为 $172.246^\circ$ , 大于阈值 $15^\circ$ | $R_{11}$ 与 $M_{10}$ 不匹配 |
| $M_{11}$ Topo_Node( $M_{11}$ )=3; Triangle_Type( $M_{11}$ )= $T_3$ | 3     | $R_{11}$ 与 $M_{11}$ 所有向量夹角都在阈值范围内                               | $R_{11}$ 与 $M_{11}$ 匹配  |

## 4 结语

匹配问题是空间数据集成与更新中的一个非常关键的技术问题。本文针对多尺度地图数据更新中的匹配问题,提出了一种基于结构化空间关系的结点层次匹配方法。实验表明,该匹配方法是一种行之有效的方法。但是在实践中,有时亦需要考虑结点的其他几何特征,如曲率,尤其对于自然地理实体(如河流)。此外,该方法可望进一步扩展运用于其他类型空间目标(如线、面)的匹配,但需要考虑更多的匹配指标(如形状、尺寸等)以及这些指标的优先顺序。

## 参 考 文 献

- [1] 李德仁, 龚健雅, 张桥平. 论地图数据库合并技术[J]. 测绘科学, 2004, 29(1):1-4
- [2] Devogele T, Parent C, Spaccapietra S. On Spatial Database Integration[J]. International Journal of Geographical Information Science, 1998, 12: 335-352
- [3] Duckham M, Worboy F. An Algebraic Approach to Automated Information Fusion[J]. International Journal of Geographical Information Science, 2005, 19(5): 537-557
- [4] Anders K H, Bobrich J. MRDB Approach for Automatic Incremental Update[C]. ICA Workshop on Generalization and Multiple Representation, Leicester, England, 2004
- [5] Volz S. An Iterative Approach for Matching Multiple Representations of Street Data [C]. ISPRS Workshop Multiple Representation and Interoperability of Spatial Data , Hannover, Germany, 2006
- [6] Stigmar H. Matching Route Data and Topographic Data in a Real-time Environment [C]. The 10th Scandinavian Research Conference on Geographical Information Science , Stockholm, 2005
- [7] 邬伦, 张毅. 分布式多数据库系统集成技术[J]. 地理学与国土研究, 2002, 18(1): 6-10
- [8] Saalfeld A. Conflation: Automated Map Compilation[J]. International Journal of Geographical Information Systems, 1988, 2(3): 217-228
- [9] Gabay Y, Doytsher Y. Adjustment of Line Maps [C]. GIS/LIS Proceedings'94, Phoenix, Arizona, 1994
- [10] Jones C B, Kidner D B, Luo L Q, et al. Database Design for a Multi-scale Spatial Information System [J]. International Journal of Geographical Information Systems, 1996, 10(8): 901-920
- [11] Cobb A M, Chung M J, Foley H, et al. A Rule-based Approach for the Conflation of Attributed Vector Data[J]. GeoInformatica, 1998, 2(1): 7-35
- [12] Walter W, Fritsch D. Matching Spatial Data Sets: A Statistical Approach[J]. International Journal of Geographical Information Science, 1999, 13 (5): 445-473
- [13] Zhang M, Shi W, Meng L. A Generic Matching Algorithm for Line Networks of Different Resolutions[C]. The 8th ICA Workgroup on Generalization and Multiple Representation, Coruna, Spain, 2005
- [14] 童小华, 邓愫愫, 史文中. 基于概率的地图实体匹配方法[J]. 测绘学报, 2007, 36(2):210-217
- [15] 邓敏, 李志林, 祁华斌. GIS线目标间空间关系的集成表达方法[J]. 测绘学报, 2007, 36(4):421-427
- [16] Devogele T, Trevisan J, Raynal L. Building a Multi-scale Database with Scale-transition Relationships[C]. The 6th International Symposium on Spatial Data Handling. London: Taylor & Francis, 1996

第一作者简介:邓敏,教授,博士,博士生导师,主要从事GIS空间关系理论及应用、地理空间信息度量和时空数据挖掘的理论与方法研究,发表论文60余篇。

E-mail:dengmin208@tom.com

(下转第1000页)