

SHILIANG YAOSU XIETIAO YIZHIXING MOSHI
PIPEI FANGFA YANJIU

矢量要素协调一致性模式匹配方法研究

赵东保 著



中国水利水电出版社
www.waterpub.com.cn

要 索 内

本书主要研究... 本书主要研究... 本书主要研究...

矢量要素协调一致性模式匹配方法研究

赵东保 著



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书对矢量要素协调一致性模式匹配方法进行了系统研究。主要内容包括：绪论，矢量要素自动匹配问题剖析，同名点要素的自动匹配，同名线要素的自动匹配——以道路网为例，同名面要素的自动匹配，同名要素上同名特征点的自动匹配，空间信息内容集成原型系统的实现，总结和展望。

本书可供空间数据集成技术研究和应用人员使用。

图书在版编目（C I P）数据

矢量要素协调一致性模式匹配方法研究 / 赵东保著
— 北京：中国水利水电出版社，2015.4
ISBN 978-7-5170-3069-0

I. ①矢… II. ①赵… III. ①空间信息系统—矢量—
配方—方法研究 IV. ①P208

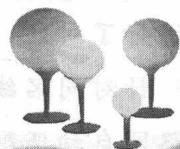
中国版本图书馆CIP数据核字(2015)第068286号

书 名	矢量要素协调一致性模式匹配方法研究
作 者	赵东保 著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座 100038) 网址: www. waterpub. com. cn E-mail: sales@waterpub. com. cn 电话: (010) 68367658 (发行部)
经 售	北京科水图书销售中心 (零售) 电话: (010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	中国水利水电出版社微机排版中心
印 刷	三河市鑫金马印装有限公司
规 格	170mm×240mm 16开本 8.25印张 158千字
版 次	2015年4月第1版 2015年4月第1次印刷
印 数	001—500册
定 价	28.00元

凡购买我社图书，如有缺页、倒页、脱页的，本社发行部负责调换

版权所有·侵权必究

前言



空间数据集成是地理信息科学领域的一个重要研究方向,全面解决该问题需要涉及多方面的研究工作,在这些研究工作中,对空间数据的内容进行集成则是其中至为关键的一环。空间数据的内容集成也就是对空间数据自身的位置、形状、语义、空间关系等信息进行集成和融合。实现空间数据内容集成的核心是能够对覆盖相同地区不同版本的矢量地图中的同名要素进行自动识别与匹配。以要素匹配为基础,就可以共享同名要素的语义信息,实现同名要素位置上的融合,对要素的变化部分进行自动检测与更新等。要素匹配技术是当前正在蓬勃发展的“人人参与空间数据采集计划”的关键技术之一。鉴于矢量要素自动匹配问题的重要理论与实践意义,本书对该问题的研究成果进行了归纳总结,系统地介绍了矢量要素协调一致性的模式匹配方法。

(1) 对矢量要素自动匹配问题进行了较为全面的分析和总结,对现有的各种矢量要素自动匹配方法进行了较为全面的分类与归纳。针对当前绝大多数匹配方法都是采用局部寻优策略来进行匹配的问题,分析了局部寻优策略的缺陷,并从全局寻优的角度出发,提出了矢量要素的全局最优匹配概念模型,为后续各种几何类型矢量要素的自动匹配奠定了统一基础。

(2) 针对同名点要素的自动匹配,根据点要素之间的匹配可信度指标,以矢量要素全局最优匹配概念模型为指导,构建了点要素的全局最优匹配模型,并将其转化为图论中的求最大带权二分图完备匹配问题进行求解,以点要素的自动匹配为例,指出了影响同名要素匹配

准确率的实际因素,通过大量随机仿真实验对全局寻优策略与局部寻优策略进行了对比,验证了全局寻优策略的优越性。

(3) 针对同名线要素的自动匹配,特地以具有典型结构特征的多尺度道路网自动匹配问题为例开展研究。以所有匹配道路结点与匹配道路弧段同时达至全局一致性为目标,构建了道路网的全局最优匹配模型,提出了基于概率松弛法的多尺度道路网匹配算法。与其他大多数道路网匹配算法相比,本书方法的特点在于在判断道路之间是否匹配时,不但考虑它们本身在各种特征指标上的相似性,还顾及领域内其他道路的匹配情况,所给出的实验结果表明该方法具有更高的匹配准确率,对数据的要求和限制更低,即便待匹配的两组道路网吻合得不够好,有较大的位置偏差,又存在 $1:N$ 匹配关系(也包括一定数量的 $M:N$ 匹配关系),该方法依然可以获得较为满意的匹配结果。

(4) 针对同名面要素的自动匹配,提出了基于几何矩和叠置分析的全局寻优的一对多面要素自动匹配方法。实验结果表明:与其他面要素自动匹配方法相比,本书方法即便在同名面要素之间存在较大甚至是非一致分布的位置偏差,而且还存在 $1:N$ 匹配关系时,仍能对同名面要素进行较为准确的匹配。

(5) 针对同名要素上同名特征点的自动匹配问题,对传统串匹配法进行了改进,一方面,提出了与旋转、缩放和偏移无关的基于形状相似度的矢量要素化简算法,以用于提取矢量要素形状边界的主要特征点;另一方面,提出了一种鲁棒的局部特征相似变换不变量。一系列实验结果表明:本书所提出的改进方法能够适应矢量要素复杂的形状特点,具有较强的抗噪能力,能够容忍矢量要素之间一定程度的尺度差异。

(6) 以本书的研究成果为基础,设计并开发了空间信息内容集成原型系统 MapMatcher,该系统能够实现基于形状特征的多个要素之间的自动配准,能够对同名点要素、线要素、面要素等进行自动匹配,具有一定处理复杂匹配情况的能力,在对同名要素完成匹配后,系统

还能进一步对同名要素上的同名特征点进行匹配,并能以多种方式相互传递同名要素之间的属性信息。

全书共分8章。

第1章介绍本书的研究背景和研究意义,并对国内外研究现状进行全面总结和分析,指出当前研究中存在的主要问题;介绍本书的研究内容、研究目标、拟解决的关键问题、研究方法和技术路线等内容。

第2章对本书所要研究的矢量要素自动匹配问题进行了全面介绍,分类归纳出在不同情况下需要对矢量要素进行自动匹配的各种类型划分。指出在对矢量要素进行自动匹配前,需要对各种数据进行处理。对现有的矢量要素自动匹配方法所采用的匹配策略的优缺点进行了分析和归纳,而后在给出全局一致性函数的前提下,构建本书所提出的矢量要素全局最优匹配的概念模型,并以该概念模型为基础,为后文的点、线、面等矢量要素的自动匹配提供理论指导。

第3章分析了同名点要素自动匹配的一般方法,对这些方法的特点进行归纳总结后,给出了点要素之间的匹配概率指标,从全局寻优的角度出发,建立点要素的全局最优匹配模型,随后给出基于二分图匹配的最优匹配模型的求解方法,最后分别以常见的地图接边以及随机仿真实验为实例,验证本书方法的有效性,对比全局寻优准则与局部寻优准则的匹配性能,指明全局寻优准则的优越性。

第4章对已有的线要素自动匹配方法进行了分类归纳,并指出了其中的主要问题所在。然后以典型结构特征的道路网为例,给出了道路结点和道路弧段的相似性度量方法,构建了矢量道路网的全局最优匹配模型,并利用概率松弛法对全局最优匹配模型进行求解,从而获得道路结点和道路弧段的匹配关系,随后通过实例对本书方法加以验证。

第5章首先对已有的面要素自动匹配方法进行了分类归纳,并指出了其中的主要问题所在。给出了利用全局形状特征以及低阶几何矩从多个潜在候选匹配面要素中快速确定匹配面要素组合的方法,并据

此提出了一种基于全局准则的一对多面要素自动匹配方法, 后续的实验结果表明该方法可以确保即便面要素之间存在较大的距离偏差, 且存在 1:N 匹配关系时, 也依然可对它们进行有效匹配。

第 6 章介绍了同名要素同名特征点自动匹配问题的来源, 对已有匹配方法进行了分类归纳, 分析归纳了矢量要素的形状特点, 对传统串匹配方法进行了改进, 将传统化简算法的绝对标量阈值改为形状相似度阈值, 构建了一种鲁棒的相似变换不变量, 并据此提出了同名要素上同名特征点的自动匹配方法, 后续的实验结果表明本书方法能够适应矢量要素的形状特点, 具有较强抗噪能力, 能否容忍一定程度尺度差异, 且能应对矢量要素之间的相似变换。

第 7 章以前 6 章研究成果为基础, 设计并开发了空间数据内容集成原型系统, 论述了该系统的总体设计思路及各个功能模块, 利用 C# 语言并结合 ArcEngine9.2 控件对系统进行了开发, 并简要给出了各个功能模块的运行实例。

第 8 章对全书的研究工作进行了综述, 总结了本书的研究成果与创新之处, 并展望了下一步研究方向。

本研究得到了国家自然科学基金青年基金项目 (41101373) 和河南省高校科技创新团队支持计划项目 (131RTSTHN023) 联合资助, 在此谨表谢意。

作者

2014 年 12 月

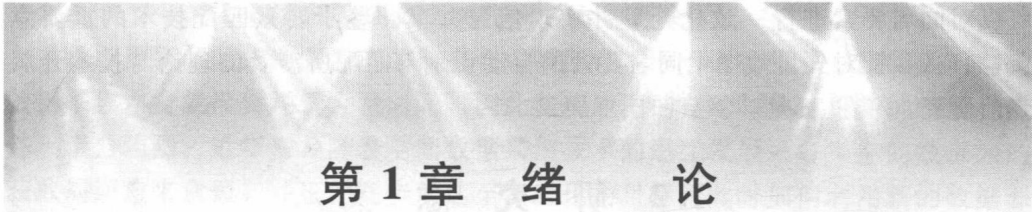
目 录

前言

第 1 章 绪论	1
1.1 研究背景	1
1.2 研究意义	2
1.3 国内外研究现状分析	4
1.3.1 国外研究现状	4
1.3.2 国内研究现状	9
1.3.3 当前研究中存在的问题	11
1.4 研究目标、研究内容及技术路线	12
1.4.1 研究目标	12
1.4.2 研究内容	12
1.4.3 技术路线	13
第 2 章 矢量要素自动匹配问题剖析	15
2.1 矢量要素自动匹配的类型划分	15
2.2 矢量要素自动匹配的数据预处理	18
2.3 矢量要素自动匹配的常用评判指标	19
2.4 矢量要素自动匹配的匹配策略	21
2.4.1 局部寻优策略	21
2.4.2 全局寻优策略	22
2.5 本章小结	24
第 3 章 同名点要素的自动匹配	25
3.1 同名点要素自动匹配的一般方法	25
3.1.1 距离邻近度	25
3.1.2 结构相似度	26
3.1.3 方位相似度	28
3.2 全局寻优的同名点要素自动匹配	28
3.2.1 同名点要素的全局最优匹配模型	29
3.2.2 点要素全局最优匹配模型的求解	29

3.2.3 全局寻优的点要素自动匹配算法步骤	30
3.3 实例验证与分析	31
3.3.1 地图接边实例及其结果分析	31
3.3.2 仿真实验及其结果分析	32
3.4 本章小结	38
第4章 同名线要素的自动匹配——以道路网为例	39
4.1 道路网自动匹配的一般方法	39
4.1.1 道路结点自动匹配的一般方法	40
4.1.2 道路弧段自动匹配的一般方法	41
4.1.3 道路网匹配一般方法的总结	46
4.2 全局寻优的矢量道路网自动匹配	46
4.2.1 道路网的全局最优匹配模型	46
4.2.2 基于概率松弛法的道路网全局最优匹配模型的求解	47
4.2.3 全局寻优的多尺度道路网自动匹配算法步骤	51
4.3 实例验证与分析	52
4.3.1 实例一及其结果分析	52
4.3.2 实例二及其结果分析	55
4.4 本章小结	57
第5章 同名面要素的自动匹配	58
5.1 面要素自动匹配的一般方法	58
5.1.1 基于重叠相似度的面要素匹配方法	59
5.1.2 基于形状相似度的面要素匹配方法	60
5.1.3 面要素匹配一般方法的总结	65
5.2 面要素的一对多自动匹配方法	66
5.2.1 候选匹配面要素组合的确定	66
5.2.2 基于几何矩提取面要素的质心	68
5.2.3 面要素全局最优匹配模型的建立与求解	71
5.2.4 面要素的一对多自动匹配算法步骤	72
5.3 实例验证与分析	72
5.4 本章小结	76
第6章 同名要素上同名特征点的自动匹配	78
6.1 同名要素上同名特征点自动匹配的一般方法	78
6.1.1 串匹配法的基本思路	79

6.1.2 串匹配法中局部特征量的构建	81
6.2 基于改进串匹配法的同名要素同名特征点的自动匹配	83
6.2.1 矢量要素的形状特点	83
6.2.2 对串匹配方法的改进	84
6.2.3 同名要素同名特征点自动匹配算法步骤	87
6.3 实例验证与分析	87
6.3.1 实例一及其结果分析	87
6.3.2 实例二及其结果分析	92
6.3.3 实例三及其结果分析	95
6.4 本章小结	97
第7章 空间信息内容集成原型系统的实现	98
7.1 系统的总体设计	98
7.2 系统的运行实例	100
7.3 本章小结	108
第8章 总结和展望	109
8.1 成果总结	109
8.2 推广应用前景及经济社会效益	111
8.3 展望	111
参考文献	113



第1章 绪 论

1.1 研究背景

随着空间数据采集手段的日新月异,人们已经积累了丰富的各种类型的空间数据,如何对这些数据进行有效融合和集成已逐渐成为地理信息科学需要迫切解决的问题之一。针对空间数据的集成,近些年来,人们在文件格式转换、空间数据模型的统一和空间数据互操作等方面已经开展了不少卓有成效的研究工作,然而对空间数据的内容本身,也即其位置、属性和空间关系等方面的集成则较少涉及。在实践中,不同组织和部门出于不同需要常常会反复采集覆盖相同地区的空间数据,尽管理论上讲,这些空间数据是对同一个客观实体的描述,应当完全相同,然而由于所使用仪器设备、测量手段不同以及采集过程中侧重点的不同,这些数据总是会存在着各种各样的差异,比较常见的差异如同名地要素位置上的差异、附着在同名要素上的属性信息的差异、由于制图综合原因所造成的各种类型要素详细程度上的差异,以及在采集数据时所处时段上的差异等。尽管存在着不少差异,但这些不同版本的空间数据又有着各自的优势与劣势,因此,如果能够对这些不同版本的空间数据进行有效融合和集成,就能够博采各个版本空间数据之所长,以获得一个集各种优势于一身的新的空间数据集,从而节省大量不必要的重复劳动和开支。出于这种对空间数据的集成、复用和共享的强烈需求,20世纪80年代就诞生了地图合并技术(Saalfeld, 1988)。地图合并(map conflation)技术一般是针对矢量地图而言的,它是地图编绘过程中的一个常见操作,在现阶段,关于地图合并,较完备的定义为:“在同名要素匹配的基础上,对相关要素的几何位置、空间关系和属性等进行调整,以实现同一地区不同来源地图数据库的集成和信息融合。”由以上定义可以看出,能否快速和准确地对同名要素进行匹配直接关系到地图合并质量的高低,是其首要和关键所在。

矢量要素的自动匹配技术尽管起源于地图合并领域,但其在其他领域也同样具有重要的基础性作用。例如,在地图变化检测领域,只有首先对同名要素进行准确匹配,才可以对地图的变化部分作出判断,又如,在最近蓬勃发展的“人人参与空间数据采集计划”(Goodchild, 2007)中,只有在识别出同名要素后,才



可以对同名要素进行一致化处理并实现信息共享。鉴于要素匹配技术的重要意义,有必要针对矢量地图中同名要素的自动识别与匹配所涉及的理论与技术开展专门研究。

1.2 研究意义

当前许多 GIS 应用项目所面临的一个主要问题就是如何将不同来源的数据集成在一起 (Devogele, 1998)。集成不同来源的信息是 GIS 的一个基本特征,但这并不意味着将不同来源的数据简单堆砌在一起就可以解决问题。要素匹配技术可以大大提高 GIS 系统集成不同来源空间数据的能力,可以说是空间数据集成领域的关键技术之一,其理论及实践意义归纳起来主要有三点。

1. 基于要素匹配实现空间数据的有效融合

数据的采集和更新是 GIS 工程中耗时耗力的一项工作,尽管人们在这个方面已经花费了大量的时间和金钱,但是当今的许多 GIS 应用系统一方面仍然存在数据不足和更新周期过长的问题,另一方面现有数据又得不到有效再利用,许多存储在各个数据库的现有数据在语义、存储格式、数据模型、编码方案等方面均可能不同,再加上这些数据有关文档也可能不全,致使这些数据的使用只能局限于原来的应用部门或原来的应用项目,形成一个个封闭的“数据孤岛”。如今,许多 GIS 工程项目所需要的数据仅来自于一个部门是难以满足所有应用要求的,例如同样是道路数据,测绘部门采集的数据位置精度很高,而交通部门采集的数据则会有丰富的道路属性信息。显然,这两者既各有所长,又各有所短,舍弃任何一个都非常可惜,而如果一项应用需要结合两方面的数据,就必须克服上述“数据孤岛”所带来的问题,找到一种合适的空间数据融合和集成手段。然而,目前大部分的商业 GIS 软件在这方面的功能都还较弱,只提供了数据格式转换、坐标系统变换等基本的数据集成功能,即便以专注于空间数据集成而著称的 FME 软件也同样缺乏基于内容集成的能力。当人们需要对空间数据的内容进行融合和集成时,主流的 GIS 平台软件,例如 ArcGIS,还仍然主要依赖于作业员的目视和手工编辑来消除不同来源空间数据的逻辑上的不一致性,效率低下。Goodchild (1996) 指出,随着空间数据的日益丰富,用户将经常会为了特定目的而需要使用多种来源的空间数据,在这种情况下,用户希望将这些数据进行某种形式的平均或匹配。然而,欲实现空间数据的融合,首先要解决的问题就是要确定对哪些数据融合,换言之,就是确定不同来源地图同名要素之间的匹配关系,只有确定了同名要素,才可以相互交换各种信息,实现所谓的融合。由此可见,要素匹配问题是空间数据融合问题的关键所在。



2. 基于要素匹配实现地图的变化检测和自动更新

地图更新问题是地图生产的核心与关键所在,由于以往很长一段时间内空间数据采集本身就需要耗费大量时间,因此实现地图的自动与快速更新的需求也就并不十分强烈。近年来,随着空间数据采集技术的迅猛发展,各类空间数据源的获取周期愈来愈短,并正在逐步地达至实时化的目标,此时如何利用新的数据源更新原有地图就日益引起人们的关注。目前, TM 遥感影像数据空间分辨率已经可达 0.6m 的精度,而与此同时,各种多波段、多时相的遥感影像卫星也层出不穷,蓬勃发展的数字摄影测量技术更是大大提高了传统测绘的生产效率,地图的生产周期日益缩短。在 GPS 领域, RTK 测量技术可以快速地获取点位坐标,其点位精度也早已可达 2cm,而精密单点定位技术也在蓬勃发展。如此众多的高精度和高现势性的数据正在呈几何级增长,如何将这些新数据融入到已有的数据库中,为用户提供更好的数据服务是数据管理和维护者面临着一个巨大课题。正如 Saalfeld (1996) 指出的那样,所有这些新的数据,无论它们精度有多高,只有将它们融入到已有的数据库中,才能显示出它们的价值。

在现实实践中,地图更新更多是针对 GIS 中的矢量地图数据库而言的,对矢量地图的更新可以分为利用遥感影像更新矢量地图和利用矢量地图更新矢量地图两种,而其中的后者又可以细分为四种情况。第一种情况是某部门对某区域进行了空间数据的最新采集,此时需要更新该区域的原有数据。第二种情况是某部门所使用的地图(从数据库)一直是购买自另一个部门(主数据库),假定从数据库已经在该地图上进行了各种编辑操作,现主数据库发生了更新,那么从数据库又该做何种更新。第三种情况是某部门已拥有某地区不同时段、不同尺度的地图,在过去并未建立这些不同地图之间的联系,而现在需要对地图的同名要素进行匹配,从而形成地图的版本化,这种情况也属于地图更新的范畴。第四种情况是由大比例尺地图的更新级联推理中小比例尺地图的更新,此时需要获知大比例尺地图出现了哪些更新,而大比例尺地图又和中小比例尺地图之间存在着怎样的要素匹配关系。无论是以上哪种情况,欲实现地图更新,都必须对地图变化的部分进行自动检测,可以说地图更新和变化检测是一个问题的两个方面,二者相辅相成,而变化检测是建立在要素匹配基础之上的。容易理解,当旧地图中的某个要素可以在新地图中找到与其匹配的要素,则可认为该要素没有变化,反之就是变化了的要素,这时就可用新的数据将其替换,从而实现地图的自动更新。

3. 基于要素匹配实现空间数据的质量评估和位置校正

空间数据质量是 GIS 应用系统的生命线,当数据质量无法得到保证时,所作出的各种分析推理都将缺乏可靠的依据。在实践中,可能会遇到这种情况:虽然手头并不缺乏空间数据,却不知道所拥有的空间数据的质量情况,或者仅仅是



了解一个大致信息。对于这种情况,如果同时还拥有该地区其他版本的空间数据,就可在对同名要素进行匹配后,通过对同名要素进行比对,可利用已知数据质量的空间数据对未知数据质量的空间数据进行质量评估。显然,不同版本的名要素可以视为不等精度下或者同等精度下的多次独立观测,对于多次独立观测,由测量误差理论可进行精度评定、消除残差、粗差定位等,从而对不同版本的空间数据作出相对质量评估。

除了可对空间数据质量进行相对评估外,更重要的是还可以利用高精度的空间数据对低精度的空间数据进行位置校正。在识别出同名要素后,考虑各种约束条件,诸如要素的形状、长度、角度等限制条件,可采用一定坐标转换模型对相关要素进行带限制条件的最小二乘平差,从而实现对矢量要素几何位置的精确纠正。另外,空间数据的质量不仅仅表现在空间位置准确性上,还表现在其他方面,如逻辑一致性上。例如,根据道路网网络分析的要求,加油站应该落在道路上,但是实测数据可能并不能满足这个要求,此时可按照一定准则,将加油站匹配到道路上的一个最合适的位置上。显然,所有这一切都必须以要素匹配技术作为基础。

1.3 国内外研究现状分析

矢量要素的自动匹配就其匹配方法而言,无外乎是从几何特征或者语义特征两个方面入手,几何特征主要包括位置、形状、拓扑结构等,语义特征主要包括属性信息等。下面将分别从国内外两个角度就基于几何特征的点要素匹配、线要素匹配、面要素匹配以及基于语义的要素匹配等四个方面对矢量要素自动匹配技术的研究现状进行归纳和总结。

1.3.1 国外研究现状

1. 点要素的自动匹配

这里所指的点要素一方面是指一些独立的点状地物,另一方面也是指具有重要标志意义的特征点,例如道路的交叉点、面状地物的质心等。点要素匹配的一种简单思路是直接根据先验知识进行判断,假定已知两幅地图各自的点位精度以及比例尺等信息,则由测量误差理论可推导出同名点的最大距离误差,将此距离误差作为阈值即可对点要素是否匹配作出判断。不过在实际情况中,由于点要素可能过于密集,再加上同名点的距离误差又较大,导致某个点要素存在多个潜在的候选匹配要素,这个时候就必须寻找一种可靠的匹配方法使匹配结果尽可能正确。Beerli 等人(2004)针对独立点状地物,依据模糊集理论,提出了一种通过计算可信度值来对点要素进行匹配的方法。随后,Beerli 等人(2005)又对当点



要素并非来自于两个图层,而是多个图层时的情况进行了讨论,进一步扩充了基于模糊集理论的点要素自动匹配方法。在地图合并中,点要素更多的是指那些具有重要标志意义的特征点,诸如道路交叉点,当对这些点要素进行匹配时,除了可以考虑距离因素外,还可以考虑拓扑结构和方位关系等特征信息,例如 Saalfeld (1988) 在对道路结点进行匹配时,就设计了一种可以称之为“蜘蛛编码”的八字节结点拓扑结构二进制编码方案,蜘蛛编码将道路结点所关联弧段的延伸方向分为八个连续不相交的角度区域,并据此对道路结点进行匹配。Samal (2004) 给出了一种基于上下文的要素匹配方法,该方法根据已知点要素的匹配结果,依据其他点要素与它们之间的相对方位关系,推导其他点要素可能的匹配关系。除此之外,在计算机视觉和模式识别领域,点的匹配问题被专门称作所谓“点模式匹配”问题 (point pattern matching), 其中不少经典算法均可以借鉴,诸如 Ballard (1981)、Fischler (1981)、Olson (1997)、Chui (2000) 等。

2. 线要素的自动匹配

线要素的自动匹配是地图合并中最为活跃的一个研究方向,甚至可以说是地图合并技术产生的原动力。早在 1980 年,美国人口调查局为了完成十年一次的人口普查,数字化了全国 278 个大中城市的纸质地图,共有 5500 多幅图,每幅的面积为 25 平方英里,这些地图主要是一些线要素,包含有丰富的属性,如名称、街区地址范围、街区类型等。1983 年,美国人口调查局又获得了美国地质测量局的高精度电子地图数据,其中仍然主要是线要素,但没有属性信息。美国人口调查局想将已有的数字化地图与新的美国地质测量局的地图进行地图合并,以便能检测并消除两幅地图中存在的不一致现象并将属性信息传递给美国地质测量局的地图数据库中,由于要匹配和合并的数据量很大,又必须保证匹配和合并的正确和快捷,因此有必要建立一个自动的地图合并系统。当时,美国地质测量局的地图数据并没有包含属性信息,故而要素匹配只能依靠一些几何测度来进行,为此 Saalfeld 等学者 (1988) 就开发了世界上第一个数字地图合并系统,他们提出了著名的橡皮筋技术 (rubber sheeting), 并利用该技术对地图中结点的位置不断进行调整和迭代,从而获取同名结点之间的匹配关系,并进而再获取线要素之间的匹配关系。橡皮筋技术后来还被引入到主流商业 GIS 软件中,如 ArcGIS 等。不过当时的数字地图合并系统还只能对线要素 1:1 的匹配关系进行处理。

由于道路网是一类非常重要的线状要素,和人们的生活息息相关,进入到 20 世纪 90 年代以后,对于线要素的自动匹配的研究大都集中到道路网的自动匹配上。这方面的代表性文献如 Walter 等学者 (1999) 在对道路网的匹配类型进行全面总结的基础上,直接依据道路网的弧段而非结点进行道路网匹配,其主要



思想是：首先利用所谓的缓冲区增长算法确定候选匹配弧段集合，然后给定一个互信息函数，互信息函数的取值大小不但与匹配弧段之间的形状、位置和拓扑等特征的差异有关，还与它们邻域范围内的其他弧段的匹配情况有关，然后采用 A^* 算法对互信息函数进行求解，并将最优解所对应的道路匹配关系作为最终解。该方法虽匹配效果较好，且可应对 $M:N$ 的匹配情况，但却非常耗时。Mantel (2004) 又进一步对 Walter 提出的缓冲区增长算法进行了改进，一定程度上加快了搜索候选匹配道路弧段的速度。Walter 直接利用弧段进行匹配的思想也得到了他人的引用，例如，Davis 等人 (2003) 在他们的开源产品 JCS (Java conflation suit) 中通过计算弧段之间的一种顶点 Hausdorff 距离来判断道路之间是否匹配，距离越小的则匹配的可能性越大。而 Deng (2007) 更是对传统 Hausdorff 距离进行了扩充，使其具有更强的鲁棒性和抗噪能力，能够容忍目标局部形状的变化，可用来判定非一对一的匹配模式。除了常见的 Hausdorff 距离外，更为常见的线状要素距离测度还有 Saalfeld (1993) 提出的一种 L2 距离测度，L2 距离实质上是对 Fréchet 距离的一种简化，两条曲线之间的 Fréchet 距离（也叫皮带距离）定义为所有单调参数化路径中的最小路径长度，这是最自然的距离测度，但由于需要处理所有可能的参数化方法使其计算非常复杂，其时间复杂度高达 $O(pq \log 2pq)$ ，其中 p 和 q 分别为两条曲线的线段数目。而 L2 距离则仅需线性时间，在计算效率上要快捷得多。不过，利用道路弧段之间的距离测度判断道路是否匹配只适用于待匹配的两组道路网在位置上极为接近的情况，当同名道路在位置上存在较大偏差时，很容易导致误匹配。

由于道路网中道路交叉点和道路弧段都蕴含了十分重要的特征信息，事实上人们更多的是将结点和弧段加以综合考虑。例如，Filin 等人 (2000) 所提出的方法首先根据道路网中结点的位置、拓扑等特征信息对结点进行初匹配，然后再根据结点所关联的弧段采用一种称之为“round trip walk”的方法对匹配结点做进一步调整。具体做法是从待匹配道路网的一点 A 沿其关联弧段走到其所有邻接点 B 、 C 等处，找到 B 、 C 等点在参考道路网中的匹配结点 B' 、 C' 等，如果从 B' 、 C' 等点出发存在关联弧段可返回 A' ，那么 A 与 A' 之间的匹配关系便可得到确认，反之则予以修改。如此这般，反复迭代，从而获得更为可靠的结点匹配关系，以此为基础再获得道路的匹配关系。不过该方法严重依赖于结点的初匹配结果，当存在较多结点初匹配不准确时，后续的迭代调整很难进行。Safra (2006) 针对移动环境下对道路网匹配的高效率要求，提出了一种简单、快速的方法，不过该方法显然只能针对待匹配的两组道路网在几何位置上甚为接近的情况。Volz 等人 (2006) 所提出的方法则是首先设置严格的阈值，而后找出满足严格阈值条件下的匹配结点，并将这些匹配结点作为种子点，从这些种子点出发根据弧段之



间的形状、距离等特征信息再找出满足严格阈值条件的匹配弧段, 然后根据已经获得的匹配结果放宽阈值的限制范围, 再次寻找匹配结点和匹配弧段。如此这般, 反复迭代, 直至没有新的匹配结点和匹配弧段出现为止。该方法的问题是: 它在各个迭代阶段所设置的阈值完全是针对特定数据进行统计调查后依据经验而得的, 并不具有通用性。相比较而言, Zhang (2007) 所提出的方法适用范围更广一些, 其方法在初始阈值确定上仍然是根据先验知识获得, 但随后在匹配过程中利用数理统计分析方法对初始阈值进行优化。但该方法只能适用于待匹配的两组道路网其同名道路在位置上存在大体一致的偏差的情况, 当同名道路位置上的偏差在某些地方大, 而在另一些地方小, 也即存在非一致偏差的情况, 该方法将很难奏效。Zhang (2006) 还专门针对大比例尺地图中的道路匹配问题进行了研究, 在大比例尺道路网中, 道路并非是道路中心线, 而是双线, 且在交叉口道路非常复杂, 加上还存在立交桥, 在这些情况下道路的匹配要更为困难一些。最近, Mustière (2008) 针对多尺度下道路网匹配问题进行了研究, 其方法也是首先对道路结点进行匹配, 在获得了道路结点的匹配关系后, 道路匹配问题就转化为从多个候选匹配道路弧段中寻找一个路径, 使这个路径所对应的道路与待匹配道路在某个指标上 (如形状、距离等) 最为相似, 而这个问题实质上对应于图论中求两点之间最短路径的问题。该方法不但可以处理道路弧段之间 $1:N$ 的匹配关系, 还可以处理道路结点之间存在的 $1:N$ 匹配关系。不过在该方法中, 道路结点的候选匹配结点通常数目很少, 也即待匹配的两组道路网数据在位置上也同样是甚为接近的, 如果道路结点的候选匹配结点过多, 他们的算法也将难以应对。此外, Kang (2004) 对多尺度道路网匹配问题也进行了研究, 其方法是首先根据大比例尺地图和中小比例尺的最小显示范围, 对大比例地图进行制图综合, 合并或者简化一些道路结点与弧段, 综合后再与中小比例尺的道路网进行匹配, 但其中又将面临制图综合这个国际难题。

除了道路网的自动匹配外, Blasby 等人 (2003) 在他们的开源产品 JCS 中也对河流网的匹配问题进行了研究, 他们的方法充分利用了河流网在拓扑结构上非常类似于一个树结构这一重要特征, 其匹配首先从河流网的主干河流开始, 然后逐步扩散到分支河流。不过, 河流网并不一直都是树结构, 相对而言, Kieler (2009) 所提出的方法更具通用性, 该方法首先对呈面状的河流构建 Delaunay 三角网, 然后据此提取河流中心线, 将河流组成一个网络, 再采用类似 Mustière (2008) 所述方法进行多尺度下河流网的自动匹配。

3. 面要素的自动匹配

面要素的自动匹配方法可归纳为以下三类: 第一类方法是将面要素的匹配转换为点的匹配, 通过对比面要素对应顶点或者质心之间的距离、点与点以及点与