

DILI KONGJIAN CHANGJING XIANGSIXING DULIANG
LILUN FANGFA YU YINGYONG

地理空间场景相似性度量 理论、方法与应用

陈占龙 吴亮 周林 张丁文 叶文 编著



中国地质大学出版社
ZHONGGUO DIZHI DAXUE CHUBANSHE

责任编辑：彭琳

封面设计：魏少雄



DILI KONGJIAN CHANGJING XIANGSIXING DULIANG
LILUN FANGFA YU YINGYONG

地理空间场景相似性度量 理论、方法与应用



ISBN 978-7-5625-3883-7



9 787562 538837 >

定价:42.00元

图书在版编目(CIP)数据

地理空间场景相似性度量理论、方法与应用/陈占龙等编著. —武汉:中国地质大学出版社, 2016. 11

ISBN 978—7—5625—3883—7

I. ①地…

II. ①陈…

III. ①地理信息系统—研究

IV. ①P208

中国版本图书馆 CIP 数据核字(2016)第 287506 号

地理空间场景相似性度量理论、方法与应用

陈占龙 吴 亮 周 林 编著
张丁文 叶 文

责任编辑:彭 琳

责任校对:戴 莹

出版发行:中国地质大学出版社(武汉市洪山区鲁磨路 388 号)

邮政编码:430074

电 话:(027)67883511

传 真:67883580

E-mail:cbb@cug.edu.cn

经 销:全国新华书店

<http://www.cugp.cug.edu.cn>

开本:787 毫米×1092 毫米 1/16

字数:320 千字 印张:12.5

版次:2016 年 11 月第 1 版

印次:2016 年 11 月第 1 次印刷

印刷:湖北睿智印务有限公司

印数:1—500 册

ISBN 978—7—5625—3883—7

定价:42.00 元

如有印装质量问题请与印刷厂联系调换

《地理空间场景相似性度量理论、方法与应用》

著作委员会

主要编撰:陈占龙 吴 亮 周 林 张丁文 叶 文

参与编撰:安晓亚 江宝得 郭明强 龚 希

徐永洋 朱榕榕 冯齐奇 吕梦楼

胡宗哲 郑 阔 覃梦娇 白雪云

金 超

前 言

地理信息产业是以现代测绘和地理信息系统、遥感、卫星导航定位等技术为基础,以地理信息开发利用为核心,从事地理信息获取、处理、应用的高技术服务业。随着国民经济和社会信息化步伐的加快,地理空间信息资源作为国家重要的基础性、战略性资源,已经成为推动信息化发展的重要基础。地理空间信息事业是国民经济和社会发展的基础性工作,为国民经济和社会发展提供可靠、适用、及时的保障服务,是地理空间信息工作的出发点和落脚点。走信息化的发展道路,推进地理空间信息的发展,是信息社会对地理空间信息发展的基本要求。

我国地理信息产业形成于 20 世纪 90 年代末,经过短短 10 多年的快速发展,已初具规模,地理信息资源日渐丰富,技术创新取得了重大成果。与此同时,作为具有发展快、效益高、贡献大、需求广、潜力足、前景好、可持续、环境友好等特点的高技术服务业,地理信息产业的发展受到了高度的重视。

笔者收录整理了一系列关系地理空间场景的构成、描述、表达、度量、应用等方面的研究成果和学术见解,并从多个角度分析和讨论了地理空间场景的作用及用途,内容精细丰富,涵盖面广。例如书中研究了融合空间认知的空间场景相似性查询理论与方法,这对满足未来智能化空间信息检索需求具有重要的理论和应用价值。针对多尺度复杂场景约束下的空间信息查询,在查询方式上采用空间场景代替标准化的查询语言;引入特征矩阵对空间场景的独立特征场景进行形式化描述;利用独立成分分析方法对特征场景进行分类,提取特征场景的独立关系特征向量;从空间场景中进行空间语义理解和相似性度量,借助场景中稳定的空间参照建立尺度变化时空间场景相似性度量模型;对传统的空间索引进行扩展以支持空间场景相似性查询,通过用户参与查询结果的评价对相似性约束指标进行松弛化修正,逼近用户查询的主观意图,提高查询效率和智能化程度。笔者将空间场景相似性理论应用于草图地理信息检索,对拓展和丰富现有地理信息检索模式,满足未来智能化网络地理信息服务需求有极为重要的理论和应用价值。具体内容包括:利用特征向量法建立空间数据场景与草图场景间的混合相似性度量模型,先提取空间场景的各种特征值组成一个高维的特征向量,然后利用图谱和独立成分分析方法对提取的向量进行降维,这样可以降低场景相似性计算的复杂度;此外还采用扩展的 R 树和空间关系索引建立顾及空间关系的双重索引检索机制,以提高搜索效率;在场景的输出问题上,探究匹配关联图的生成以及约束条件的松弛化处理的问题,为检索与图的搜索之间建立相关的联系,从而使场景的输出更完备有效;在信息检索效率方面,把人的参与引入到信息检索的过程中,使检索结果更接近用户的主观意图,从而提高检索性能。通过本书的研究内容,可以为基于草图的地理信息检索提供理论和方法上的支撑。

该书共分为 8 章,分别从地理空间场景的概述、相关理论,场景的表达、度量,全局的一致性评价、智能反馈及应用等方面来深入地介绍和说明地理空间场景信息。其中第二章由安晓亚、叶文组织编写,第三章、第四章、第五章的主要内容分别由吴亮、周林、张丁文负责完成,陈占龙、叶文负责全书的整理、编辑等其他工作。书中内容丰富广泛,包含了地理空间场景中的

一系列主要内容和大量的细节问题,梳理了地理空间场景中几乎全部的知识点。该书的出版为 GIS(geographic information system)邻域和地图制图学领域的人士提供帮助及参考,也将为其他读者认识地理空间信息、地理空间场景等起到引领的作用。

陈占龙

2016 年 4 月 28 日

目 录

第一章 地理空间场景相似性概述	(1)
第一节 地理空间相似性	(1)
第二节 地理空间相似性的研究现状与分析	(2)
第三节 地理空间场景相似性的应用	(5)
第二章 空间场景相似性的基本问题	(7)
第一节 相似性的相关理论	(7)
第二节 空间相似性的基本问题	(8)
第三节 相似性度量	(10)
第三章 地理空间场景表达	(13)
第一节 空间关系分类	(13)
第二节 空间拓扑关系	(14)
第三节 空间方向关系	(32)
第四节 空间几何关系	(64)
第五节 空间语义关系	(71)
第四章 地理空间关系相似性度量	(78)
第一节 拓扑关系的相似性度量	(78)
第二节 空间方向关系的相似性度量	(95)
第三节 复合参考对象稳定性评估	(100)
第四节 空间几何形状相似性度量	(101)
第五节 空间距离关系的相似性度量	(109)
第六节 空间语义关系的相似性度量	(111)
第五章 全局一致性评价的地理空间场景层次化度量	(114)
第一节 空间场景相似性相关基础理论	(114)
第二节 地理空间场景的综合表达	(122)
第三节 地理空间场景层次化相似性度量模型	(125)
第四节 地理空间场景相似性度量的一致性综合评价	(129)
第六章 地理场景相似性指标自适应融合与智能反馈	(139)
第一节 地理场景相似性的约束性指标	(139)
第二节 地理场景相似性指标自适应融合方法	(139)
第三节 地理场景相似性指标松弛化策略	(146)

第四节	智能算法与地理场景相似性反馈·····	(149)
第七章	相似性度量在智能地理信息匹配检索中的应用 ·····	(160)
第一节	地理场景匹配检索的基本问题·····	(160)
第二节	地理场景中的特征场景划分·····	(161)
第三节	地理场景相似性查询索引树·····	(162)
第四节	地理场景相似性查询案例分析·····	(163)
第八章	相似性度量在制图综合评价中的应用 ·····	(182)
第一节	制图综合评价的基本问题·····	(182)
第二节	场景相似性度量在制图综合中的应用·····	(183)
主要参考文献	·····	(188)

第一章 地理空间场景相似性概述

地理信息是一种潜在的巨大的战略资源。如果把地理信息比作一座金矿,那么地理空间场景则是打开这座金矿的一把钥匙。下面对地理空间场景作具体的概述。

第一节 地理空间相似性

一、地理空间

地理空间是物质、能量、信息的数量及行为在地理范畴中的一种存在形式,特指形态、结构、过程、关系、功能的分布方式和分布格局在时间上的延续(抽象意义上的静止态),讨论所表达出的“断片图景”。地理空间的研究是地理学的基本核心内容之一。

二、地理空间认知

认知(cognition)是人类认识和了解我们生活的世界时所经历的各个过程的总称,包括感受、发现、识别、想象、判断、记忆、学习等。Neisser(1967)把认知定义为“感觉输入被转换、简化、加工、存储、发现和利用等诸过程”。认知是模拟人脑理智思维的信息处理过程,是一系列的感知、思考、学习、概念化、推理、问题求解、记忆和精神意象等精神活动所组成的复杂行为系统(章明,1989),因此,可以说认知就是“信息获取、存储转换、分析和利用的过程”,简言之,就是“信息的处理过程”。认知科学是当代新兴的研究门类,被看作 20 世纪世界科学的最重要的进展之一。认知科学是心的科学、智能的科学、思维的科学、知识及其应用的科学,它研究的目的就是说明和解释人在完成认知活动时如何进行信息加工的。认知科学也是一门新兴的独立学科,它研究智能系统,包括自然的智能系统和人工的智能系统,有专门的学科语言,并以此建构关于智能系统的认知过程的模型和理论(章士嵘,1992)。人类个体的认知能力有明显的差异,认知的方法也有明显的个性化特征,认知的结果不是必然的,认知的进程也充满了不确定性。认知与人的心理有关。

认知心理学(cognitive psychology)是以信息加工观点为核心的心理学,故又称为信息加工心理学,是研究人接受、编码、操作、提取和利用知识的过程,即感知觉、记忆、表象、语言、思维,以及儿童的认知发展和人工智能,强调人已有的知识结构对当前认知活动的决定作用,研究人进行信息加工的各个阶段的特点。核心是解释认知过程的内部心理机制,即如何获得、存储、加工和使用信息。认知心理学认为知觉是先感知对象的整体,后感知对象的部分,这也是人们对空间目标相似性判断时进行图形综合的原因。地理认知侧重于心理感知和分析,认知者既感知图上明显的信息和挖掘潜在的信息,又主动解译信息,形成对客观世界的整体认识。空间认知是从人类如何理解和表示空间的方式,以及空间信息的处理过程的角度去理解空间。

地理空间认知是指在日常生活中,人类如何逐步理解地理空间,进行地理分析和决策,包括地理信息的知觉、编码、存储、记忆和解码等一系列心理过程。地理空间认知的研究内容包括地理事物在地理空间中的位置(where)和地理事物本身性质(what)。地理空间认知作为认知科学与地理科学的交叉学科,需要在认知科学研究成果的基础上对地理空间的相关问题进行深入研究,因此,与认知科学研究相对应,地理认知研究主要包括地理知觉、地理表象、地理概念化、地理知识的心理表征和地理空间推理,涉及地理知识的获取、存储和使用。

三、地理空间对象

与地理空间位置或特征相关联的对象称为空间对象,它包括结点、弧段等几何元素。通过对地理实体或现象以及它们的相互关系、分布特征、空间特征进行分析和抽象,可以把空间对象定义为如下几种类型:

(1)点状地物。点是零维几何元素,它是三维数据模型中最基本的元素,曲线、曲面或其他形体均可用有序点集表示。点状地物是零维空间目标,包括独立地物点、结点以及在点元素的基础上加属性编码和属性表构成的点状地物。

(2)线状地物。线或弧段是一维几何元素,由起结点和终结点加一系列有序点集组成。弧段一般没有属性意义,它可能是某线状地物、面状地物或体状地物的一部分。线状地物是一维空间目标,由一条或若干条弧段组成,且有自己的属性编码和属性表。

(3)面状地物。面是二维几何元素,是物体上一个有限、非零区域,由一个外环和若干个内环界定范围。面有方向性,一般用外法矢方向作为该面的正向。面状地物是二维空间目标,由周边弧段组成,是有属性编码和属性表的曲面,它不仅可表示地表上的地物,还可以表示地表下的各种层面。

(4)体状地物。体是三维几何元素,是由封闭表面围成的空间,也是欧氏空间中非空、有界的封闭子集,边界是有限面的并集。体状地物是在体元素的基础上包含了自己的属性信息。

(5)复杂地物。复杂地物由上述4种中的若干个地物组成。

四、地理空间场景相似性的定义

空间相似关系具有两方面的含义:一是指空间目标几何形态上的相似,历史上著名的大陆漂移学说的提出就是源于大西洋两岸大陆轮廓线十分相似吻合;二是指空间物体(群)结构上的相似。形态相似分析本身就是分析的目的。形态相似在很多情况下是更深层次分析的基础,提供部分分析依据,这是因为形态是空间物体的特征之一,而属性特征是另一个重要方面。对形态的相似性分析有两种途径:一是在相似变换下图形吻合度的分析,二是基于形态参数的聚类分析或相关分析。

第二节 地理空间相似性的研究现状与分析

一、空间关系表达及其相似性建模

GIS 中的空间关系主要描述空间对象之间的各种关系,为 GIS 的分析提供基本的理论和

方法支持。空间关系主要包括距离关系、方向关系、拓扑关系及空间关系相似性等。从空间关系的内涵来看,空间关系模型应能够反映目标尺度、人类认知、目标层次、现象的不确定以及随时间变化等特性对空间关系的影响;从数学角度来看,空间关系模型必须是可形式化和可推理的,以方便操作和实现。空间关系从人们认知的角度对空间现象和目标间的关系进行建模,因而在空间数据查询、检索,空间数据挖掘、空间场景相似性评价以及图像理解等应用领域得到了广泛的应用。

1. 拓扑关系

拓扑关系主要关注地理实体的连接、包含、相邻和相离关系。空间拓扑关系表达有 3 种方法:基于点集拓扑的四元组、九元组和维扩展法。这些方法能够较好地表达两个空间目标的交集不为空的拓扑关系。基于 Voronoi 图的维扩展描述法可以表达空间目标间的交集为空的拓扑关系,特别是邻近关系。影响较大的空间拓扑模型是 Randell 等(1992)提出的区域连接演算(region connection calculus,简称 RCC)理论。RCC 理论基于空间区域连接关系 $C(x,y)$,由 $C(x,y)$ 可以推导出 RCC-8 模型,它区分了 8 种互不相交且联合完备(jointly exhaustive and pairwise disjoint,简称 JEPD)的 RCC 关系:(DC,EC,PO,TPP,TPPi,EQ,NTPP,NTPPi)。对于实际空间对象,该模型可以得到 8 种拓扑关系。

2. 方向关系

针对点对象,Andrew Frank(1991)提出了“基于圆锥”和“基于投影”两种方向划分方法。区域对象之间的方向关系比点对象之间的方向关系更复杂,可用一种 3×3 的“方向关系矩阵”表示。该矩阵对应参照对象最小边界矩形将平面划分为 9 个部分,通过考察每个部分与主对象的关系来确定参照对象和主对象之间的方向关系。

3. 距离关系

GIS 中可以定义多种地理空间中两点间的距离,使用比较普遍的两种地理距离概念是空间距离和定性距离。距离关系的表示分为绝对距离关系和相对距离关系。绝对距离关系直接表示两个空间对象之间的距离;相对距离关系通过与第 3 个对象的比较,间接表示两个对象间的距离。有时单纯的距离关系不足以用于推理,如“点 A 距离点 B 很远”且“点 B 距离点 C 很远”,要得到 A、C 两点之间的距离关系还要知道这 3 个点间的方向关系。方向与距离的结合称为位置信息。

二、基于空间相似性的实体匹配

实体匹配算法中常用的指标有空间信息指标和非空间属性信息指标。空间信息指标是实体的几何及拓扑信息等方面的指标,而非空间属性信息指标是使用实体的属性进行语义上的匹配。国内外很多文献根据面实体的特性,利用不同的信息指标对匹配方法进行了研究。如基于知识的非空间属性数据匹配策略,通过计算属性项的相似度值来确定匹配实体。该方法在很大程度上依赖于数据集的数据模型及属性数据类型等。从几何的角度来说,点连成线,线围成面,国外很多文献研究了点实体和线实体的匹配,如结合点的几何位置和蜘蛛编码的点匹配方法,也有利用点距及线段方向夹角进行线串的匹配。针对面实体的匹配,如利用面质心并结合多种匹配检验规则的几何匹配方法,大致步骤是先通过面实体栅格化后收缩来确定质心,然后将它矢量化,用点在面内的规则进行粗匹配,再结合多边形的面积和面密度进行匹配检

验,最终判断匹配情况。

三、基于空间关系相似性的图像检索

基于空间关系相似性的图像检索是通过对象间的方位关系或者拓扑关系的定性空间推理来完成的,它可以粗略地分为3类:精确匹配的检索、子图检索和相似性检索。精确匹配的检索需要两幅图像具有完全一致的对象集合,并且空间排列完全相同;子图检索需要查询图像的每个对象是否都出现在数据库图像中,并且两幅图中符号相同的对象要满足空间相似,即子图检索就是在数据库图像中找到“包含”查询图像的结果;相似性检索的条件更宽泛,只要两幅图的对象集合交集不为空即可,目标是通过查询符号化图像数据库,找到和查询图像具有相同符号的对象以及对象间具有一致的空间关系的图像。

四、空间相似性查询的反馈机制

交互性使得系统可以同时使用来自于计算机和人的能力,并能基于用户的反馈来动态地更新搜索权重。因此,首先用户不需要在搜索前指定一组精确的权值,相反,用户与系统交互,并指定它所认为的相关的反馈,动态地更新搜索权值;其次,计算机不再需要去理解高层次的概念,基于用户的反馈,嵌在搜索权重里的高层次概念会自动地被精确。目前用户相关性反馈主要应用于文本检索和基于内容的图像检索(CIBR)上,CIBR系统使用图像的虚拟内容,如颜色、纹理和形状要素作为检索图像数据库中图像的依据。

首先空间场景具有与图像相似的多义性,因为计算机基于搜索场景的多个特征要素在空间数据库中检索与它匹配的场景,同样地,基于内容的图像检索在检索前需要输入多个图像特征。其次场景匹配的检索结果可能返回多个候选场景,这与图像检索中返回多张相似图像的行为一致。同时,空间场景匹配和基于内容的图像检索都较侧重于用户的视觉感受。因此,把反馈机制引入到空间场景匹配中将会得到更好的匹配结果。

五、研究现状分析

关于空间相似性在GIS中应用的研究,国内外的文献较少。概括起来,国内外的研究主要有以下几点:

(1)郭仁忠(2001)在他的《空间分析》第二版中阐述了空间相似关系,吴立新、史文中在2003年所著的《地理信息系统原理与算法》中提到相似关系,他们都说明了这个问题的重要性,即空间相似关系在GIS中是个基础的问题。郭庆胜、丁虹(2005)开展了基于栅格数据的面状目标空间方向相似性研究。在其他学科中,如数学、心理学、生物学、地理学、认知科学、语言学、体育学和图像模式识别中也有一部分有关相似性应用方面的文章。

(2)国外学者对GIS也有大量的研究。代表性成果有Gudivada和Raghavan(1994)用对象质心之间的角度来定义基于方向的图像相似性;Sistla等(1997)提出了一个在三维空间中的方向和拓扑关系的集合以及基于这个集合的相似性测量;Lee和Hsu(2008)用二维串中的被编码的符号对象投影和串匹配算法来评价结构相似性;Nabil等(1998)结合投影的技术,用概念邻近来测量关系相似性;Rodriguez(1990)开发了一种计算模型——MD(Matching—Distance)模型去评价语义相似性;Goyal(2000)提出了一种能考虑空间目标在方向关系矩阵的方向片之间变化的空间方向相似性计算方法;Dimitris Papadias、Nikos Karacapilidis和Dinos

Arkoumanis(1999)在定义了拓扑、方向和距离的相似程度后提出了一种空间场景相似程度的计算方法。Bruns H T 和 Egenhofer M(1997)在空间拓扑关系概念邻域图和空间方向关系概念邻域图的基础上提出了一种空间场景相似性的描述方法。Dimitris Papadias、Nikos Karacilidis 和 Dinos Arkoumanis(1999)用模糊隶属函数定义了空间拓扑关系的相似性、空间方向关系的相似性和空间距离关系的相似性。Alex Hagen(2002)提出了模糊集方法,采用近似关系和分类来评估两幅地图的相似性。

第三节 地理空间场景相似性的应用

一、空间图形目标的识别

将分布在空间中复杂的地理目标简化并抽象成相对简单的图形,利于地理目标的表达,并可将这样的图形称之为“空间图形”。“图”的直观性、简洁性、信息量大等特点,使它成为表示地理目标的有效工具。无论地理目标多么复杂,都可以将它抽象成点、线、面 3 种基本空间图形,甚至可将它扩展为点群、线群、面群。

空间图形有自身的特征,包括位置、形态和分布。空间位置是借助空间坐标系来传递空间目标的位置信息;形态是图形的最基本几何特征,是可让人直接感受和判断的,还有另一些几何特征是必须用数值描述的,如线状目标的长度、弯曲程度、分维数等,面状目标的轴线、形心、轮廓周长、面积、凸壳等,曲面的坡向、坡度、结构线等;空间分布是反映同类空间目标群体的位置信息。

根据空间图形的特征,可利用拓扑、方向、距离等空间关系对空间实体的大小、形状、距离、连通性等进行判断,识别出相似的空间场景。

二、空间大数据分析挖掘

空间数据挖掘,也称基于空间数据库的数据挖掘和知识发现,它是从空间数据库中提取隐含的、用户感兴趣的知识,空间关系或其他有意义的一种模式。空间数据挖掘对于理解或重组空间数据,发现空间与非空间数据之间的内在联系,构造空间知识库、优化空间查询等方面起着重要的作用,因此广泛应用于智能 GIS、遥感影像处理、交通控制、环境研究和气象预报等领域。

空间数据挖掘所处理的对象是空间数据,具有非结构化、多尺度、高维数、模糊性、不确定性等特点。相比传统的数据挖掘,空间数据挖掘具有新的特点,它挖掘的数据具有复杂性,如点、线、面等对象;空间数据一般具有相关性,如自然资源、野生动植物和气候的变化是随空间渐变的,且包含空间维度信息。

空间数据挖掘主要是研究空间数据的分布模式、聚类、分类特征、属性间的依赖关系、时空相关和互相关特征等,从而发现大量的地理空间信息中所隐含的有用信息。如:空间关联规则,即空间对象间相邻、相连、共生、包含等关联关系;空间聚类规则,即特征相近的空间对象聚类的规则;空间特征规则,即一类或几类空间对象的普遍特征;空间区分规则,即多类空间对象间的不同特征;空间演变规则,即空间对象依时间的变化规则。

三、空间智能化查询

空间数据查询技术在地理信息系统中基于位置的服务、建筑规划与布局等方面得到了广泛的应用。典型的空間数据查询包括位置查询、范围查询、最近邻域与最近邻查询等。为了适应高级应用,出现了新的空间查询技术,包括位置选择与评价、路径选择、位置相关查询与服务等。现有典型的空間位置查询技术支持准确输入信息的查询,应用位置查询技术的系统需要用户提供准确的位置信息,这样难以满足支持模糊信息输入的智能化空間查询需求。空間位置信息查询的过程中经常会遇到不能提供准确输入信息的情况,例如在事故发生时,求救者不知所在位置的准确信息,只能提供周围环境信息,如标志性建筑等,因此要确定求救者的位置只能根据空間场景的相似来确定。

四、空間数据的一致性维护与匹配分析

多尺度数据匹配的困难主要在于3个方面:①同名地点位置误差;②同名地点形状差异较大;③存在大量的 $1:N$ 和 $M:N$ 匹配关系。同名地点位置误差的存在使得传统的基于距离邻近度和重叠相似度评价方法的匹配效果大大降低。由于尺度不同的原因,同名要素的形状往往变化较大,这使得基于形状相似性评价的适用性也在降低。另外,不同尺度条件下,相邻要素之间存在大量的合并操作,这造成同名要素之间的匹配关系大多以 $1:N$ 和 $M:N$ 为主,这类匹配关系的匹配问题是非常复杂的,特别是在普遍存在同名地点位置误差的情况下。对于多尺度要素匹配,在同名要素之间的位置和形状等特征都存在较大差异的情况下,要素之间的空間结构关系往往在匹配中发挥了突出的作用。传统基于单个要素的匹配搜索和相似性评价方法不便于进行空間结构特征的提取和全局一致性评价,所以基于空間场景的匹配候选对象的同步搜索和整体相似性评价的匹配策略可以对多尺度数据进行匹配及评价。

五、空間数据制图综合一致性评价

地图综合是一个知识简化的过程。因为通过综合得到的是关于事物主要本质的结论,而不是事物所有特性的罗列,在这个过程中就要区分主次,去伪存真,这是一个知识筛选、化简的过程。同时,地图综合又是一个重新组织知识、生成新知识的过程。经过重新组织得到的新知识往往对事物本质把握得更为准确,认识更为全面。

地图可以看作是在一定数据模型支持下的空間场景,从大比例尺、高分辨率状态到小比例尺、低分辨率状态的地图综合过程,可看作是两个空間场景的映射,包括空間目标的映射和空間关系的映射。

地图综合是一个充满了矛盾的过程,它的目的是在数据传送极小化的同时,使读者接收的信息达到极大化,所以地图综合是在众多的制图约束条件的竞争下找到一个折中的解决方法,这个竞争过程的规则就是综合的约束条件,因为有可能在满足了一个制图约束条件的同时就违反了另一个制图约束条件。例如在夸大一个符号或某个目标的某个重要特征的时候,就可能压盖了相邻的目标,引起了空間的冲突等,因此利用空間场景的相似对地图进行一致性评价可提高地图的质量。

第二章 空间场景相似性的基本问题

在客观事物的发展过程中,都存在着同和异。只有同,才能有所继承;只有异,事物才能往前发展。相似是客观事物存在的同与异矛盾的统一,变异是事物发展和运动过程中的差异。从系统论角度考虑,客观事物都可看作是一个系统,而每个系统都存在具体的属性和特征。在不同系统中,可能存在某些共有的属性或特征,这些共有的属性或特征可称之为相似特性,描述这些共有的属性或特征的量称之为相似特征值。当系统间存在相似特性时,则认为系统间存在相似性。相似性是反映特定事物间属性和特征的共同性和差异性。相似与相似性的联系与区别是:相似性的外部表现形式为相似现象,相似现象的内部本质属性即为相似性,即相似性与相似是本质与现象之间的关系。

第一节 相似性的相关理论

一、相似性的定义及其性质

相似性的本质是系统间客观特性存在相似。由于系统中属性和特征的客观性,因此相似性不依赖于人们的感性认知而存在。从哲学上来讲,事物之间的相同都是有条件和相对的,而事物之间的差异或相似是无条件和绝对的,即事物间普遍存在不同程度的相似。尽管相似是客观存在的,但客观世界中相似现象必然要反映到我们的感性认识和理性认识中来,因此,相似性与认知科学是有关联的。目前对这种关系有两种观点:①相似性是人类认知的基础,因为相似的事物往往具有类似的属性、行为或功能,这是人类认识事物、对事物进行分类或推理的基础之一;②相似性具有很强的不确定性,即相似性是人的一种主观认知,并不取决于对象的属性,它是人们根据自己对世界的认识、所处的环境和判别的目的所做出的一种主观的、整体的判断,换言之,相似性是因时、因地、因人而异的主观认知过程。

相似性具有以下性质:

(1)普遍性,事物之间普遍存在不同程度的相似。

(2)反身性,即自身与自身的特征相似。

(3)对称性,若 A 与 B 的特征相似,则 B 与 A 的特征相似。

(4)模糊性和不确定性,即 A 与 B 相似度的大小因度量方法的不同而具有不确定性。

(5)相似性是事物与事物之间的一种关系,而非事物自身的一种属性。

(6)相似性具有系统性与层次性。如果将事物看作系统,则事物之间的相似性便是系统之间的相似性,而系统又可以分为许多子系统,子系统间也具有相似性,所以相似性又具有层次性。

(7)相似性具有动态性,一切事物总是随着时间不断运动变化着,因此事物间的相似性及相似度会随着时间而变化。

(8)相似性度量具有目的性,相似判据必须因相似工程应用的不同而不同,必须紧密结合应用实际采用对应的度量方法。

二、相似性的分类

目前对相似性的分类尚未有明确且权威的观点,这里仅选其中最具代表性的分类。

(1)1965年,Torgerson将相似性分为两大类:客观相似性和主观相似性。客观相似性是指对象的多维特征之间具有某种函数对应关系,这种对应关系不以人的意志为转移而客观存在,一旦指定特征类别和函数形式,对象之间的相似度大小取决于对象特征的度量值;主观相似性是指人与研究对象之间的认知关系,是人主观认知的结果,相似度的大小因人、因环境而异。

(2)在心理学领域,Medin(1959)将相似性分为感知相似性和概念相似性。感知相似性是指两对象之间表面视觉特征的相似性程度,故而又指视觉相似性、表面相似性或物理相似性;概念相似性是指对象在语义特征上的相似性。

(3)周美立(2003)将相似性分为以下几类:①一般相似性和具体相似性。前者指广义上事物间普遍存在的相似性,包括自然、社会和心理学的各个领域及各系统的相似性;后者指系统间具体属性和特征的相似,如结构、功能、几何特征、物理特征、化学特征和生物特征相似等。②自然相似性与人工相似性。前者指客观存在的自然现象及系统间的相似性,是自然界演化过程中形成的;后者指依靠人的创造性活动,在人造系统或人造系统与自然系统间的相似性。③他相似性与自相似性。前者指不同系统间的相似性,而后者是指系统或者事物自身的局部与整体具有相似性,如分形就是研究自相似性的理论。④精确相似性、模糊相似性和可拓相似性。精确相似性指服从同一自然规律,可用相同的数学方程式描述,能精确识别和获取相似特征值,从而实现精确相似性度量;模糊相似性是指系统间相似特性不能用经典数学精确描述,相似性具有一定的模糊性;可拓相似性是指介乎于精确相似性与模糊相似性之间的相似性。⑤混合相似性。混合相似性不是单一性质的相似,而是一种综合性的系统相似。混合相似性可以对不同性质的相似特性分别用相应的方法去识别、分析和度量。

第二节 空间相似性的基本问题

一、空间相似性的概念及研究范畴

目前对于空间数据相似性及空间数据几何相似性的概念尚未有明确的定义,但已有许多概述,如Alec Holt(1997)对空间相似性的定义是:在一个特定的尺度和内容上被认为是相似的两个区域。广义上,空间相似性是根据特定内容和尺度(或比例尺)对空间的匹配或排序。多数学者将空间相似关系归类于空间关系的一种,与距离关系、拓扑关系、方向关系和相关关系并列,认为是空间信息学的理论基础之一。空间数据是指空间实体的属性、数量、位置及其相互关系的空间符号描述,可以是实体的地理坐标、点的高程、道路的长度、多边形的面积、像元的灰度,也可以是地名、注记等字符串,还可以是图形图像等多媒体成分。空间数据具有从宏观、中观到微观的层次,也具有空间性、时间性、多维性、海量性和空间关系复杂等特性。结合相似性的相关概念和空间数据的概念特点,空间数据相似性可认为是空间数据实体间的属

性、数量、位置及其相互关系等特征的共同性和差异性。空间数据几何相似性是空间数据实体间除属性特征之外的所有与几何特征相关的空间位置、形状、长度、面积、方向、分布及空间实体间相互关系的共同性与差异性。根据空间数据相似性度量的定义,相似性度量的对象可以是空间数据中的任意两个或两个以上的地理实体。空间数据几何相似性度量模型的研究内容包括以下几个方面:

(1)度量的数据集数量。参与度量的空间数据集必须至少是两个。

(2)度量的区域。参与度量的两个或两个以上数据集所在区域是同一区域或者部分属于同一区域。

(3)度量的数据类型。参与度量的数据包括矢量数据及其对应区域的遥感影像数据和栅格数据,矢量数据一般是指基础的地理信息数据。

(4)度量的对象或特征。必须是除属性信息之外的与几何特征相关的对象。矢量对象的几何相似性主要包括点集结构及分布特征的相似度量与匹配,线目标和面目标的几何相似性度量与匹配,以及包含实体间拓扑关系和方向关系相似性度量的混合相似性度量;影像的几何相似性主要侧重于线目标和面目标的轮廓形状;相似匹配的目标一般是属于不同数据集的同一实体。

(5)度量的尺度。参与度量的空间数据可以是不同尺度的矢量数据、不同分辨率的影像数据。

(6)度量的时相。参与度量的空间数据可以是不同时间产生或更新的空间数据。

二、空间数据相似性产生的根源

根据相似性的定义及相似性与差异性的关系可知相似性与差异性辩证统一的,是一个矛盾的两个方面,因此,要研究空间数据相似的产生、形成与演变的规律,就有必要研究同一区域处于不同数据集内的实体集合或同一实体因客观和主观原因而发生的变化。变化前后的空间数据既存在一定的差异性,同时也必然存在一定的相似性,即变化导致空间数据间产生差异性和相似性。因此,有必要对地理实体的变化进行分类和讨论。

地理实体的内容包括实体特征及其特征的变化。地理实体的特征包括空间特征、属性特征、空间关系特征和时间特征。空间特征描述了地理实体的空间位置、形状和分布等信息;属性特征描述了地理实体的特性,包括地理实体的专题属性及实体之间的非空间关系;空间关系特征描述实体之间的拓扑关系特征和方向关系特征;时间特征描述了地理实体的生存周期、随时间的变化状态,以及时间关系的相关性。时间特征可以在空间特征和属性特征中得到反映。特征的变化是指地理实体特征状态改变的过程,它的实质就是一个或者多个地理实体随着时间的推移在空间特征、属性特征、空间关系特征等方面各自单一的变化或者综合变化。在空间特征中,点实体的出现、消失、平移以及整个点集的分布特征发生变化,线实体的出现、消失、增加、减少、延伸、缩短以及产生位移、旋转、局部形状和整体形状的改变,面实体的面积增大、面积减小、出现、消失、新增、移位和旋转缩短、局部变形、整体变形等都能引起空间特征的变化。属性特征的变化包括地理实体属性项的变化,如属性项的增加、减少等,以及地理实体属性值的变化,如属性值的增加、减少和改变。空间关系特征包括空间拓扑关系和方向关系,地理实体空间和属性特征的变化也会引起空间拓扑关系和方向关系的变化。