

空间数据融合的研究进展:从经典方法到扩展方法

张景雄^{1,2} 刘凤珠¹ 梅莹莹¹ 唐韵玮³

1 武汉大学遥感信息工程学院,湖北 武汉,430079
2 武汉大学测绘学院,湖北 武汉,430079
3 中国科学院遥感与数字地球研究所数字地球重点实验室,北京,100094

摘 要:随着数据获取方式和技术的多样化,多源空间数据持续累积,迫切需要研究数据集成方法,以更好地为地学研究与应用提供信息和技术支持。将各种用于集成多源空间数据的处理过程统称为数据融合,并从经典方法和扩展方法的角度,分别综述相关研究进展。为综述经典方法的研究进展,依据空间对象数据模型和场数据模型,理清了数据融合及其相关数据处理较成熟的方法;对于扩展方法,阐述了多点地统计方法、统计一机理型方法、多尺度分析与重构方法、信息论方法等。为更有效地支持多源异构数据环境下的数据融合,讨论了尺度不匹配、语义不一致性、时间维等问题。

关键词:数据融合;地统计学;机理型;信息论;数据支撑

中图法分类号:P208 **文献标志码:**A

随着测绘、遥感、计算机、通信等技术的发展,获取地球空间信息的方式越来越多样化,使得空间数据量持续增长,并呈现出鲜明的异构性(即不同来源、分辨率、精度和格式等特点),尤其是步入大数据时代后,数据来源纷呈,数量激增^[1-3]。空间数据不断积累和丰富,有必要进行这些数据的整合,从而最大限度地发掘和利用这些数据资源,实现数据融合与信息增值,更好地服务于空间分析与应用。

本文将所有对多源地理空间数据进行集成和整合的过程称为数据融合,是影像融合和地图合并的一种概括^[1, 4-6],是从不同的数据源合并为一个新的、更优的数据集的过程。“新”和“优”大多表现在融合后的数据集的分辨率的提高和/或精度的改善^[7],虽然有时候融合只是为了相互印证。另外,本文的“新”和“优”也表现在融合后的数据的格式更符合特定的应用环境和要求,如受传输信道容量限制或简化数据处理并提高效率的数据压缩处理。

数据融合首先出现在 1980 年代^[8],主要考虑如何消除或缩小地图数据的几何不一致性,提高地图的几何精度^[9-11]。其应用颇多,例如,将旧图上的属性数据转到位置精度更高的新图上来;或

通过比较不同时期的地图数据进行变化检测,用来整合多源土地利用数据,实现土地变更信息快速精准获取;在地理空间数据库更新方面,可利用大比例尺地图对小比例尺地图进行级联更新^[12]。数据精化^[13]可认为是数据融合的一个应用,它侧重于精化数据的空间尺度(分辨率),提高数据的精度。

一般地,空间数据(非空间数据亦然)依其测量尺度水平分为名目、次序、等距、等比数据,前两者统称为类别型或定性数据,后两者合称为连续型或定量数据。地理空间概念化的模型(即地理数据模型)有两种:一是实体或对象模型,认为现实世界由离散实体(表示为 $ID(x, attri)$)构成,其中 x 代表地理空间位置, $attri$ 为属性(可为定性或定量的关于对象特性的描述);另一模型为场模型,由定义在问题域中点位 x 的若干单值函数构成,包括连续型变量 $Z(x)$ (如地形高程)或类别型变量 $C(x)$ (如土地覆盖类型)。空间对象可分为 bona fide 型和 fiat 型; bona fide 型对象是指真实存在的空间实体或能诚实反映地理分布的特征;另一类对象是 fiat 型,即人工或抽象的^[7]。

值得说明的是,相对于本文的范畴,空间对象的重点是其位置,而其属性被认为是相关应用领

域的专业人员所应关注的,因而讨论对象数据融合时,前者侧重于点位数据的配准和合并,后者一般只涉及属性数据的转换或合并。另有网络模型,但因其表达的方法是基于场或对象而相应地将其归并为场模型或对象模型,不作为第三种模型讨论^[7]。但是,网络数据的融合问题本身很重要,虽然本文不予深入讨论。

本文将综述地图合并和影像融合等较成熟的方法和应用,探讨数据融合的新方法。本文中,成熟的方法被称为经典方法,而相对新颖的方法被称为扩展方法。经典方法的综述将不会是“大”而“全”的,着眼点是方法,尤其是扩展方法的提炼。当然,本文在探讨扩展方法时的侧重点是场数据,但这不排除空间对象的点位数据在融合处理时可运用基于场模型的方法。

1 经典方法

根据数据源类型不同,可以将空间数据融合分为:(1)栅格数据与栅格数据之间的融合,如影像融合;(2)栅格数据与矢量数据融合,如利用遥感影像进行地图更新(以矢量数据为主)、基于辅助信息(如矢量数据)改善遥感影像分类(以栅格数据为主);(3)矢量数据与矢量数据之间的融合,如数字地图合并问题^[9,14-15]。上述的数据类型的区分依据是数据结构,而数据模型(即空间对象模型和场模型)更有助于讨论不同特性的数据,因为它是地理空间概念化的框架(同一数据模型可以使用不同数据结构,同一数据结构可以用于不同数据模型)。

本文将回顾空间对象数据和场分布数据融合方面的研究进展。分开讨论不表示融合时不同时处理对象数据与场数据。相反地,它们经常作为输入,服务于数据融合和空间问题求解。

1.1 空间对象数据

1.1.1 地图合并的概述

作为空间数据采集和集成的核心技术,地图合并除广泛用于空间数据库的集成和信息共享外,还用于同一地区不同来源地图数据库之间的图形配准、差异分析和质量评价,以及地图数据库的局部变化检测和自动更新^[14]。

地图数据库的合并一般包括 4 个过程:(1)数据预处理;(2)同名实体的匹配;(3)图形数据的合并;(4)属性数据转换或合并。根据使用的匹配准则,地图合并分为几何合并、语义合并和拓扑合并。

几何合并就是将一幅地图的特征转换到另一幅地图(目标地图)上,使两个地理空间数据库的几何差异性最小。距离作为主要的绝对几何参数用于衡量两个地理空间数据库之间的差异性,可通过待匹配的元素计算得到(一般地,欧氏距离用于点匹配,平均距离、Hausdorff 距离和离散 Frechet 距离用于线匹配)。语义异质性会产生地理空间数据库同质元素的匹配误差。语义合并是通过比较候选同名点实体的语义信息进行匹配,使待合并的数据库之间的语义差异性最小。语义合并评价方法主要基于元素之间的语义关系,利用语义过滤器剔除与合并过程无关的元素,从而建立语义联系。它可分为基于本体论的评价方法和使用人工智能的评价方法两类。拓扑合并是几何合并和语义合并共同作用的结果,其评价方法涉及全局信息调整^[14]。地图合并相关研究见文献^[16-23]。

1.1.2 实体匹配

空间实体匹配是通过对空间对象的几何、拓扑和语义进行相似性度量,识别出不同来源空间数据图层中的同一地物,从而建立不同空间数据集之间同名实体间的联系,以及探测不同数据集之间的差异或变化。根据判别依据,匹配可分为几何匹配、拓扑匹配和语义匹配。

几何匹配属于强条件匹配,即通过计算参照对象与源对象之间的几何相似度进行目标匹配。常用的几何相似性度量指标包括距离、形状相似性、角度或方向相似性等。拓扑匹配属于弱条件匹配,指通过计算候选同名实体的拓扑关系度量作为匹配的依据。此匹配方法的优点在于克服了几何匹配方法的不足,缺点是同一地物在两幅图中的拓扑关系的微小差异都将导致匹配失败,故常与其他匹配方法结合使用。语义匹配通过比较候选匹配目标的语义信息作为匹配依据。

根据目标类型,匹配分为点目标匹配(包括点与点匹配、点与面匹配)、线目标匹配(包括线与线匹配、线与面匹配)和面目标匹配(即面与面匹配)。点目标匹配一般根据几何位置差异、结构相似度以及语义相似度来匹配。线目标匹配的依据包括距离量度、几何形状、拓扑关系、图形结构、属性等。面目标匹配也可以根据距离(位置)、形状、大小、拓扑关系、属性数据等来实现。

匹配算法分为同类要素间的匹配和不同类要素间的匹配。不同类要素间的匹配问题将在 1.1.3 节简述。同类要素对象(如道路网、河流、居民地等)间的匹配及一致性维护,主要通过几何特

征、属性特征及拓扑关系的相似性建立匹配关系,利用 Delaunay 三角网、共轭点控制变换等几何方法实现一致性改正。

日新月异的城市建设使得人们出行越来越依赖于新的导航数据,而导航地图产品的现势性远不能满足日益增长的要求。因此,导航数据更新一直是近年来研究的热点技术,路网数据匹配技术则是其中的关键技术之一。道路数据的匹配,可归为线要素匹配。针对道路网线要素特点,可根据线段斜率对线段进行投影、分割,提取线要素特征点,将复杂的道路网匹配转化为特征点匹配,从而提高整体的匹配效率^[22]。考虑到道路网交叉口的结构形态特点,可使用局部网络结构描述道路的形态特征与拓扑关系,并通过比较待匹配道路间的结构相似性来确定最优的道路匹配^[19]。对于同名道路存在较大位置偏差甚至是非均匀偏差等问题,可考虑使用基于全局寻优策略的道路网自动匹配方法^[20]。

1.1.3 数据更新

空间数据库建立及更新需集成多种来源数据,存在匹配中的不一致问题。这种不一致可能发生在不同主题、不同比例尺、不同时代的数据集成中,其表现形式包括要素表达不一致、邻近关系冲突、拓扑结构不一致、地理匹配关系矛盾等。

本文以地形图上的水网和等高线为例,介绍不同类要素对象间的匹配以及一致化处理。等高线本来是为了表达地形高程这一场分布的(见1.2节),但本文为讨论方便,将其与线状对象——河流联系起来。等高线作为地势起伏与地形凹凸特征的表达,与水网表达有着密切的约束关系。在空间数据库集成与匹配中,由于来源不同,这两种数据往往出现不一致,导致出现水网“爬坡”现象。利用 Delaunay 三角网模型在等高线基础上通过弯曲分析提取地形特征,基于河流分布与谷地走势的相关性,可建立水网与等高线间的匹配关系,进而识别两者间的一致性。在此基础上利用仿射变换等几何方法完成一致化操作,实现二者的匹配^[21]。

变化检测是空间数据库更新的研究重点。国内外学者提出了对象匹配、拓扑联动及空间叠加等变化信息检测方法。为了实现 GIS 数据更新中变化信息的自动提取,结合模式识别的理论与方法,可使用神经网络决策树方法进行变化信息识别^[22]。该方法能准确快速地识别出变化信息,有助于提高地理信息系统(geographic information system, GIS)数据库动态更新的自动化与智

能化水平^[23]。

在电子地图的生产实践中,经常需要发现不同道路网信息几何上的差异,来实现道路数据的质量分析、道路数据的压缩和地图更新等。为此,文献^[24]提出了一种基于扫描线的计算道路网信息几何差异的算法,该算法可以有效地检测出道路网中的变化信息,具有很强的实用性。

本节回顾的是空间对象(如道路等线状对象)变化检测方面的若干研究进展。直接与场分布(如遥感影像及其所反映的土地覆盖和地形等信息)有关的变化检测的问题将在1.2.4节中予以阐述。

1.2 场分布数据

前文已述及,场分布数据包括连续型和离散型两类。场模型有 6 种变体:规则点、非规则的点、等值线、规则格网(栅格)、多边形、非规则三角网^[7]。例如,遥感数字影像是常见的栅格数据,土地覆盖通常使用多边形或规则格网模型,而地形高程数据经常使用上述 6 种模型。

数据可以根据是否经过处理而分为原始数据和派生数据。原始影像经一般的处理,如格式转换、影像增强等,得到的数据不是本文特指的派生数据。派生数据是指由原始数据派生的特征或信息,如从立体像对派生数字高程模型数据。对派生数据而言,数据融合所包涵的数理意义需加以领会,因为派生数据其实是关于特定空间分布的信息,如航测技术派生的地形数据、遥感反演的地表温度等信息产品。当然,这些信息产品包含了成像过程和数据处理的信息损失和精度降低的影响。

为简便起见,下文将分别叙述原始影像数据(原始影像灰度值及其简单处理后的影像数据)和派生数据的融合方法及其典型应用。原始影像数据简称为影像数据,其他数据简称为非影像数据(指遥感反演信息产品、GIS 图层、一般非规则格网表示的场分布数据),以示区别。

1.2.1 影像数据融合

显示地表光谱反射或散射特性的遥感影像(光学和微波)是场分布数据的常见例子。其融合的研究积累较深厚,旨在集成单一传感器的多波段数据或不同传感器的多波段数据所含的空间信息和光谱信息,获得比单波段数据更为丰富的信息,以增强影像的可视化与解译能力^[4,25-27]。根据融合处理所处的不同阶段,影像融合可分像素级、特征级和决策级 3 种层次。基于像素级的影像融合是直接对采集获得的数据进行融合处理,

是最低级的影像融合。显然,像素级的影像融合应理解为基于像素量测值的融合,而不能简单地理解为基于像素(作为数据结构和数据支撑)本身,因为下文将叙述的特征级和决策级的融合也可以使用像素的栅格数据结构。特征级的影像融合需首先从多源数据中提取特征,如利用影像边缘检测、影像分割提取线状和面特征,利用特征变换从影像量测值提取判别特征、构建特征空间,然后根据某种算法进行特征融合。从原始影像中提取的特征因子由环境因子决定,例如其位置、形状、周围环境等。特征级融合即是从多源数据中提取出相似的特征,进行有效的集成。决策级的融合是利用含有信息附加值的“影像”(意指由原始影像派生的结果,称为决策,如影像分类结果)进行处理的一种方法。其即是对多源数据分别进行处理以提取信息,然后将获取的信息按照一定的决策规则进行融合,提高数据的容错性,以便对所研究对象有更好的理解。不难看出,像素级融合是特征级和决策级融合的基础,特征级融合本身也是决策级融合的基础。

与空间对象数据相同,影像数据融合前(尤其是像素级的),至少要经过配准、地理编码等预处理。地理编码有至关重要的作用,因为若没有经过配准,融合后的影像会存在伪彩色或伪特征;而配准又包括影像数据经重采样后获得一致的分辨率和具有相同的地图投影(重投影需要地理编码)。遥感数据的几何校正、辐射校正是其有效融合的前提和基础。影像配准的不可避免的不精确性导致偶发差异甚至异常变化。文献[28]利用对影像配准的局部调整,得到一个对误配准不敏感的异常性变化的指标。

像素级影像融合的方法研究积累深厚,所涉及领域不限于遥感和 GIS^[25,29-30]。常用方法包括基于色彩模型的 IHS 变换的影像融合^[31]、KL (Karhunen-Loeve)变换、Brovey 变换(通过归一化后的多光谱波段与高分辨率影像的乘积来增强影像信息)、基于多分辨率分析的融合方法(注重影像的多分辨率分解与重构)、地统计学中的协同克里格^[32]。前述的多尺度影像融合方法可以基于小波变换。小波变换使用一组不同尺度的带通滤波器对原始信号进行滤波,以便将其分解到一系列频带上,并在分频基础上进行影像融合。

上述有些方法(如多分辨率分析和克里格)可能泛化用于特征级和决策级融合,但经常需经必要调整和修改。特征级融合方法的独特性表现在其与面向对象融合处理的无缝对接、与具体应用

的信息有效性和针对性等方面;而决策级融合方法的优势是可以汇集众多信息融合方法,如 Bayesian 条件概率方法、证据推理方法、模糊集合和模糊逻辑方法等。

1.2.2 全色锐化与影像超分辨率

全色锐化是一种用于提高多光谱影像空间分辨率同时保留其光谱信息的像素级融合技术。根据各类方法的主要技术,可进行分类:(1)分量替换方法,包括 IHS、主成分分析(principal component analysis, PCA)和正交变换,因为这些方法都利用了许多分量的线性变换和分量代替运算;(2)相对光谱贡献方法,包括 Brovey 变换、全色加多光谱方法,应用光谱波段的线性组合代替影像分量进行融合;(3)高频(信息)注入方法,其中高频信息是通过原始影像减去经过低通滤波的全色影像后获得的;(4)基于影像的统计方法,包括 Price 法、空间自适应法和贝叶斯法;(5)多分辨率方法,包括拉普拉斯金字塔、小波变换法和轮廓分析法,以及多分辨率分析法与其他类别方法的任意组合^[33]。

与全色锐化不同但有联系的是影像超分辨率。研究者发现,在同一场景下获取多帧影像时,所获取的低分辨率序列丢失的细节不尽相同,对其进行数学算法的处理可以提高影像的分辨率。本文将对所获取的同一场景的多帧连续影像序列进行处理,重建一帧高分辨率影像的过程称为分辨率重构。

文献[34]提出了基于频域的方法,将多帧降采样不含噪声的影像变换到频率域进行解混叠处理,利用傅里叶变换的位移特性将不同帧的信息连接起来,在频率域求得最终解,继而重建得到一帧空间域的高分辨率影像。文献[35]提出了空间域插值的方法,将降质过程分为模糊和降采样两部分,并只对降采样进行处理而不进行反卷积处理来恢复模糊的部分,具有一定局限性。文献[36]研究了基于小波的超分辨率重构算法,指出第二代小波更适合于超分辨复原与重建。基于概率论的方法是超分辨率重构中被广泛研究的方法,其中研究最多的是基于最大后验概率(maximum a posteriori probability, MAP)的算法。这类算法比较灵活,尤其在引入正则项部分^[37]。文献[38]综述了影像的超分辨率重构方面的进展,包括对上述成果的总结。

近几年来,有关超分辨率重建的研究越来越多,研究者针对不同的应用提出了很多新方法。文献[39]提出用较高分辨率和较低分辨率影像的

稀疏表示实现影像超分辨率重构。文献[40]通过学习主字典和副字典,并用稀疏表示恢复主高频和副高频的信息,得到更好的重构结果。文献[41]通过两个阶段稀疏性进行超分辨率重构,第一个阶段产生一个较低分辨率的影像,再根据重构字典产生一个高分辨率的影像。这些成果大多是基于单幅影像的,将有助于基于多帧影像的超分辨率的研究。

显然,与全色锐化不同,影像超分辨率使用同一场景的多帧连续影像序列,根据的物理机制不同,用途也不同。这是因为全色锐化使用的是多波段影像,利用全色波段较高的空间分辨率信息,融合多光谱波段的信息,以提高多光谱波段影像的空间分辨率。经典的超分辨率重构中输入的是影像序列,不涉及全色锐化中的全色波段数据与多光谱波段数据的融合问题。

1.2.3 非影像数据的融合

上述的影像融合属于场数据融合比较特殊的类型,因为其规则的数据支撑和较为深厚的研究积淀。经典影像融合的诸多方法并不能直接或简单地用于一般的非影像数据的融合。这些非影像数据具有信息特质,也常常是非规则格网采样或表达方式。它们的融合处理应遵从随机场(random fields, RF)或区域化变量的理论和方法。RF的概念使得这些非影像数据的融合方法的基本出发点是地统计学。当然,影像数据的融合也可以应用地统计学。在此基础上,可以发展宽泛的方法体系。下文讨论的侧重点是影像融合方法(像素级)的延伸,即考虑了数据支撑的非规则性、数据空间覆盖的不完整性和不一致性等复杂性的若干方法,并以高程数据为例。

场分布数据可以是规则或非规则采样的,一般是有限数据支撑的,而非点支撑的。前文已述及,数据精化可视作为各种为提高空间分辨率(可能包括精度)为目的^[42],但不限于特定数据类型的数据融合方法的另称。某一数据源参与融合的主要方式可以是提供有用的信息,例如关于某场分布的精细尺度的空间结构信息。

例如,通过地貌变化检测,可用内插的方法实现数字高程模型(digital elevation model, DEM)的更新与精化^[13];文献[43]提出了3种DEM数据的精化方法,以修补断裂线,保证地形影像的连续性。文献[44]采用克里格法精化机载激光高程数据和雷达卫星高程数据。文献[45]研究了涵盖无缝合并、缺失数据填充、含噪数据平滑等技术环节的DEM数据融合方法,并基于ASTER

GDEM2和CGIAR-CSI v4.1,创建了一个新的DEM数据产品(90 m空间分辨率,覆盖了全球约91%面积),名曰EarthEnv-DEM90,具有数据质量增强等优点。相关研究为快速获取大范围、高分辨率、精确的DEM等信息产品提供技术支撑。

上文中提及的克里格法是地统计方法的一种,用于空间预测,即给定空间邻域的某RF(或与之相关的RF)的若干点位样本数据,估计未知点位的RF值。由于空间数据可认为是均值(确定性)与残差(随机性)之和,故可利用地统计学方法进行数据融合(前文已经提及了利用该方法的案例)。较之于传统的统计方法,地统计学方法的优越性主要是所预测的变量值是最优无偏估计,并提供该预测值的克里格方差(用于预测值的不确定性评估)^[46]。

地统计学还有一个方法是随机模拟。由于随机模拟可以生成大量等概的关于某RF的实现,便于按需要统计派生反映某空间查询(如某坡度值范围内的山体面积、各作物面积统计、某污染物浓度超标的概率风险值等)结果的不确定性的指标(如方差、中误差、概率等),除了统计计算均值外。虽然克里格法和随机模拟均提供不确定性的估值,但前者只能评价单点位(或块段)的某RF误差统计(或指示变换后的概率),而后者可以估计任何空间邻域和空间查询的不确定性指标^[47-48]。上述的克里格和随机模拟方法均应冠以“协同”二字,因为融合的数据环境是多元(变量)的。文献[49]描述了一个利用非规则分布的少量高程点数据与粗分辨率DEM格网数据的条件模拟方法生成较高分辨率DEM的实例。

1.2.4 变化检测

与空间对象数据相对应,变化检测也是场分布数据(如遥感影像)融合的应用之一。例如,遥感影像是获取空间变化信息的重要数据源,可使用包括影像差值、分类后比较等变化检测算法,进行变化信息提取。

文献[50]综述了基于多时相、多光谱、多传感数据,面向生态系统监测的变化检测方法,主要有:(1)分类后比较;(2)复合分析;(3)影像差值法;(4)影像比值法;(5)变换法,主要是针对多光谱和高光谱遥感影像,通过多变量数据变换减少或消除影像的波段相关性,包括PCA、多元变换探测(multivariate alteration detection, MAD)和独立成分分析(independent component analysis, ICA);(6)变化向量分析(change vector analysis, CVA);(7)多维时相特征空间分析方法,其中,影

像叠置作为一种数字增强技术,用于在显示屏上勾勒变化,该影像由 3 个波段的最大值组合而成。由它们组成的复合影像的各波段由阴极射线管的蓝、绿、红色彩枪显示时,变化将以独特的颜色呈现。这种方法不但可以基于光谱波段数据,还可以基于归一化植被指数(normalized difference vegetation index, NDVI)和归一化湿度指数(normalized difference water index, NDMI)等。文献[51]也综述了常见的遥感影像检测算法。

国内外学者还提出了许多改进算法,如基于光谱梯度差分的土地覆盖变化检测方法^[52]、K-均值聚类用于合成孔径雷达(synthetic aperture radar, SAR)影像变化检测^[53]、面向对象的变化检测算法^[54]。除了双时相影像,多时相、多尺度数据的融合处理也是变化检测的重要方向,如基于不同分辨率影像的子像素级土地覆盖变化检测^[55]、结合景观学度量指标和多时相遥感影像的土地覆盖变化检测^[56]。

针对时间序列影像,文献[50]讨论了基于时相轨迹的分析方法。该方法比较有关指标在生长季节或年份的时相进程曲线(也称时间轨迹)。高时间频率的数据获取不仅提高了变化检测的时间分辨率,而且极大地便利了物候特征的描述。例如,某变化参数可计算为某一月份的植被指数(vegetation index, VI)值与平均值的差,除以标准偏差。如果该参数异常地高于或低于某一阈值,一季节性或年际变化将被标记。上述 PCA 方法也可用于基于时间序列的变化检测,其优势包括可以区分不同时间频率的变化,如 10 a 尺度的生产力的变化、季节性变化、传感器相关指数的飘移以及与厄尔尼诺南半球震荡现象有关的 VI 变化等。时相轨迹分析方法适用于不同传感器,还可以使用不同指标,如 VI、地表温度、空间结构数据等。

对于非影像数据,文献[57]描述了一个用于调适立体测图生成的 DEM(连续空间覆盖但较低精度)与激光测高点云数据(高时间分辨率、稀疏空间覆盖、较高精度)之间差异的新方法。其中的高程改正值是从反映高程变化信息的详细的高程变化时空历史数据中确定的。

此外,一些其他学科领域的方法也用于变换检测,如基因算法^[58]、信息测度方法^[59]。对于地统计方法,它们的优势亦可在变化检测中得以发挥:这可视为将融合运算的加法改为变化检测时的减法,协克里格的方程求解可以便利地推导出,而空间自相关和变量间协相关函数同前^[7]。

2 扩展方法

尽管数据融合研究和应用有长足的进步,但是空间数据的多源、多尺度特征,还有语义、精度、格式等方面的不一致性问题,致使数据融合研究难度较大,相关的技术未有令人满意的解决方案,其应用的水平和自动化程度不够高。因此,经典方法需作必要和有益的补充和完善,形成数据融合的扩展方法。本节将总结先进的地统计方法、统计-机理方法、多尺度分析方法、信息论方法等,作为扩展方法的发展起点和基础。

本文的扩展方法与经典方法的划分不是唯一或绝对的,而是相对意义上的。另外,相对于它们的真实点位,空间对象的位置数据可视为特殊的场分布数据,虽然拓扑关系等空间数据的特性需加以考虑。可以探索场数据(尤其是非影像数据)融合方法库中适用于点位数据的子集或补集。

2.1 地统计方法的扩展

除了普通的地统计方法(如协同克里格),多点地统计方法亦可用于数据融合。多点地统计是相对于两点地质统计学提出的,利用训练图像代替变差函数来表征特定类型下的空间差异性 or 纹理特性。传统的地质统计学以变差函数为工具,定量研究地质变量的空间相关性。但它只能描述空间上两点之间的相关性,即在二阶平稳或者本征假设的前提下空间上任意两点之间的相关性,难以表征复杂的空间结构和再现复杂目标的几何形态。多点地统计学正是在这样的背景下得到研究和发展的。

该方法可以重现问题域的空间格局^[60-61],其最大优势在于利用训练影像量化空间结构:在训练影像上每次可搜寻多个点来表达空间的连续性和可变性,克服了变异函数的限制。

多点地统计模拟基于传统地统计的序贯指示模拟算法,分为基于像素点的模拟和基于模式的模拟。对于连续型场数据,多点地统计模拟的一个常用功能就是空间外推,即用一个区域的已知数据来推算另一区域缺失数据的值。对于 DEM 数据,用基于模式的模拟方法能够更精确地捕捉到地形特征。基于模式的多点地统计模拟包含了一种局部图案相似性匹配的思想,有利于地形数据局部自相似特征的再现。这反映在模拟输出的 RF 实现里。多点地统计方法应顾及数据的多元特征,如同协同克里格(或协同随机模拟)之于克里格(或随机模拟)。

文献[62]提出了基于多点地统计模拟(multiple-point geostatistical simulations, MPS)的DEM数据融合方法。MPS可以根据训练图案的空间相关性获取相似图案。基于MPS的FILTERSIM算法,可用于数据融合。然而,传统的FILTERSIM算法并没有顾及空间结构的非平稳性。文献[62]提出了一种改进的FILTERSIM算法,并通过实验验证了传统地统计差值、传统FILTERSIM算法和改进的FILTERSIM算法3种融合方法的效果。实验结果表明,改进的FILTERSIM算法能得到更高的精度,并保存更多的空间结构。

多点地统计主要应用于地质、石油行业,近年来涉及遥感领域。多点地统计方法能存储及再现地物的几何结构,如利用多点地统计在粗尺度遥感图像的基础完成了超分辨率重建^[63]。多点地统计模拟能有效地融合尺度差别较大的数据,并保证影像的连续性,对多源、多尺度数据的整合具有很大的发展潜力,值得进一步的探索。

地统计学领域地统计反演模型(geostatistical inverse modeling, GIM)方法体现了贝叶斯融合策略。其中,先验信息以反映空间结构的空间协变差模型形式表达,量测数据用于获得验后分布信息^[64]。它的一大优势是提供了一种空间结构建模的数据驱动方法。在地统计学和应用中,空间结构建模的重要性不言而喻。因此,除数据精化的功效外,GIM将有利于空间结构建模。

2.2 统计-机理型方法

地统计方法隐含假设RF场分布S是均值m与残差r之和,即 $S=m+r$ 。若考虑某问题域离散化为n个格网点,可得向量表示的模型 $S=m+r$ (其中, $S=[S_1 \cdots S_n]^T$, m、r类似定义)。对此区域化变量模型,考虑数据采样的数理模型,可建立相应的模型-数据关系,进而导出统计-机理型数据融合方法。

设有采样矩阵(雅可比矩阵,描述场变量一采样数据间的卷积过程) H (N行n列),量测误差e(N分量列向量,反映模型-数据不匹配)。于是,得到类似GIM的线性模型,也就是待估场分布S与采样数据值Z(N分量列向量)之间的线性关系: $Z=HS+e$ 。这种机理型模型在数据融合中大有作为,并将拓宽GIS和空间分析的地理信息服务范围和专题类别,如水文、生物量、人文信息等,不再单单是传统的地图要素、高程、地类界等经典测绘信息范畴。

文献[65]描述了一个时空数据融合方法

(spatio-temporal data-fusion, STDF),使用了 $Z=HS+e$ 模型。它基于降维的Kalman平滑处理,为处理海量遥感数据而设计,以有效应对传感器的“脚印”、量测误差和空间覆盖等特性。研究者利用该方法集成特性互补的日本温室气体观察卫星(greenhouse gases observing satellite, GO-SAT)数据和美国国家航空航天局(National Aeronautics and Space, NASA)大气红外测深仪(atmospheric infrared sounder, AIRS)数据,以最优地估计全球低大气层摩尔分数。

常规的数据融合和地图合并大多未考虑源数据和集成后的数据的不确定性传递问题。虽然融合的性能评估时有些考虑了输出数据的精度指标,但大多只是相对意义上的^[66]。作为不确定性传播链条的源头之一,数据获取时的量测误差须纳入融合处理的相关算法中,一来更客观地确定融合权重,二来更准确地量化误差传播。上述的统计-机理模型有助于融合过程的不确定性分析,因为量测误差e直接纳入了模型。

前述的融合处理所涉及的场变量是连续型的。类别型变量的融合处理可以使用多数投票、概率加权、证据推理等传统方法。判别空间^[67]的概念和模型强化了类别信息的提取、尺度转换、不确定性分析的科学可重复性。利用判别空间模型,可以首先进行连续型源数据(作为判别向量的部分或全部)的融合操作,然后在融合后的判别向量的基础上进行类别信息提取:将判别向量输入所用分类器,输出各制图单元的最或然类别代码或隶属于各备选类别的隶属度。显然,判别空间属于机理型模型。

统计-机理型可用于多尺度数据融合,并有助于多尺度Kalman滤波等方法^[68]的利用。可以说,小波变换提供了一个多分辨率信号分解的架构,而多尺度Kalman滤波则是与之对应的一个多尺度分析与估计(重构)方法,以支持多尺度信号处理。多尺度Kalman滤波是高效的递归滤波器,能从一系列的不完全及包含噪声的测量中估计动态系统的状态。它实现了一种基于分层树的尺度递归算子。其中,树的每一层对应不同的空间尺度,且树的底层对应着最精尺度,即最高分辨率。树结构有利于创建尺度-递归线性动态模型,它定义于树的局域,代表粗略-精细的进化过程。另外,线性量测与过程变量存在一个形如 $Z=HS+e$ 的线性关系。

原始最精细尺度的过程变量的估计可由两步计算过程完成。第一步是精细-粗略尺度的计算

过程(是时域 Kalman 滤波的推广),第二步是粗略-精细尺度的计算过程。多尺度空间树模型可用于多分辨率影像分析、信息融合和数据同化^[7,68-70]。多分辨率信号分析与重构是数据融合的有力方法,因为其提供了机理型方法的基础,可以纳入数据生成机制或不可直接量测的过程参量的动态机制。

2.3 信息论方法

数据融合的处理方法可以基于信息论^[71-72],这至少反映在数据配准和融合运算本身。基于互信息(一种相似性测度)的(医学)影像配准具有普适性和高精度等优点,其在遥感领域的应用较新。文献[71]提出了一个广义偏体积估计方法(generalized partial volume estimation,GPVE),用于计算互信息。实验结果表明,高阶 GPVE 可显著地减少通常的联合直方图插值导致的人工效应,并较基于流行相似测度(如均方差值和标准协相关)的配准结果更具一致性^[71]。

不同数据(图层)的融合旨在合成新的图层,使之提供更多的信息与对场景的解译^[72]。为此,可采用信息论方法,使得其能从两个输入的影像(源数据)自动选择最富含信息的像素值。融合的准则是基于两输入数据图层间创建的信息通道,使人们能量化每一像元灰度值相关的信息量。该特指信息由分解该信道互信息的 3 种方式获得。评价原则是基于融合图层的信息量,它可用于分析和修改初始的像素值(通过对每一输入数据对结果贡献加权的方法)。对每一图层,特指信息是与每一像素每一灰度值相连的,并可以计算获得。融合的数据由每对匹配像素信息含量最高的像素值获得。

然而,当源数据不是相同分辨率的影像数据时,或者融合后的数据层不是输入数据的线性函数时,上述信息量方法并不能直接适用,需发展适用于一般数据融合处理的信息加权算法。

3 讨 论

3.1 数据支撑与尺度不匹配

在本文提及的融合处理中,数据支撑是所涉及的核心概念之一,处理不当将成为技术软肋、应用的障碍。数据支撑指数据所隐含的几何特性,如形状、大小、方向等^[73]。作为规则或非规则格状型的场分布数据,格数据是基于有限数据支撑的,即是底层场变量经卷积而生成的数据^[7]。影像数据是具有规则支撑的格状型数据,而地学研

究和应用中涉及许多具有非规则支撑的格状型数据,如资源普查、环境监测、生态建模和社会经济统计数据^[74-75]。点支撑数据可视为特例,以简化问题。

多源数据往往是多(空间)尺度的。尺度的主要表现形式是数据支撑。相对于非规则的数据支撑,规则的格网或者像素可能使融合处理更方便,尤其是在它们彼此大小成整数倍比例时,如地球观测系统(systeme probatoire d'observation de laterre, SPOT)的全色波段与多光谱波段的像素大小之间。一般的不同数据支撑的存在,使得数据融合时面临尺度不匹配问题。为解决此瓶颈问题,应在理清基本概念和逻辑框架的基础上,以区域化变量理论和信息论为指导,重点研究不规则数据支撑、尺度不匹配的场分布数据(包括通过或转换为场模型表示的空间数据,如点位偏差数据)的处理分析机制,探索地统计学用于数据融合的新途径,集成多尺度信号分析与重构方法和信息论方法等,提升数据融合的系统性和科学性^[7]。

3.2 语义不一致性

除了尺度不匹配,数据融合中的另一障碍是语义不一致性。数据的语义是指某一数据源的数据取值的含义或所指。例如,高程值本来是指“纯”地面的海拔高程。如果地表被树木或建筑物遮挡,航测或其他技术所量测和推算的高程值就不是真正的高程值。因此,来源不同的高程数据,其语义不一致性将造成融合处理的复杂性,影响融合的精度。

上述的讨论是面向场分布数据的。空间对象和位置数据是 GIS 的范畴。地图合并的关键是同名对象的匹配。传统的匹配是基于几何的,近年来越来越多的讨论聚焦于语义匹配。语义不一致性将造成对象数据合并的困难,需妥善处理。文献[2]提出了一个 4 组分-元组的空间特征定义方法和基于几何和语义相似度的线性组合的融合过程,以及相应的误差测度方法。进一步的研究将是语义畸变的形式化以及语义畸变与几何畸变现有模型的关联性。另外,需要研究通用的模型,用以估计误差,以便进行几何和语义匹配^[3]。

语义不一致性常常与尺度不匹配纠缠在一起,使得融合处理的难度系数增大,影响数据融合的处理效率和精度。另外,新地理学的出现,如基于含有社区产生的内容的网络地图等众源数据的使用^[76],地图特征的解释越来越多样化。于是,数据融合的架构需足够普适,灵活处理语义不一致性和尺度不匹配,以适应复杂数据环境下的融

合要求,满足应用需求。

3.3 时间维度

空间数据的重要特性之一是其空间和时间交织。上述的诸多处理隐含了静态数据,即仅考虑了空间特性。

变化监测是数据的时间维显现作用的应用场合。变化检测可以基于双时相数据,也可以基于时间序列。前者相对简单,后者复杂些,但可以提供量化地表过程物候特征的源数据。其他基于非影像数据的时间序列的时空分析方法也可考虑。

基于多帧影像的超分辨率技术也利用了时间维信息。具体说,同一场景下获取的降采样(低分辨率)影像序列包含了该场景高分辨率影像的信息。每帧影像所包含的信息不尽相同,对其进行数学处理,如正则化或者概率方法,可以恢复较高分辨率影像,即提高影像分辨率。

除了上述多时相影像或数据的融合处理,现代的信息获取方式日益实时和连续。传感网代表着时代特征。传感网是一个用于环境监测,由空间分布的且能够进行内部通信的传感器组成的系统。地理空间传感网是一种互联网环境下包含数据采集、查询和处理的新型观测信息服务系统。研究地理空间数据和信息服务资源的整合是提高传感网资源有效利用的重要途径。文献[77]设计并开发了地理空间传感网信息服务系统——GeoSensor,实现了对遥感、移动原位等类型传感器及其观测数据等信息资源的在线搜索、即时获取、网络控制和实时制图。

很明显,时间维将使尺度不匹配和语义不一致性的问题愈发复杂化。本节讨论的3方面的问题需加以全盘考虑,并进行适当处理,以实现时空数据的配准与融合。例如,时序影像和海量传感网的数据处理将涉及时空配准^[78]等技术处理。时空分析也将是未来研究的重点之一。

4 结 语

随着空间信息获取技术的发展和方式的多样化,多源空间数据与日俱增,数据融合的研究应运而生。地理信息科学领域有关空间数据融合的研究经历了近三十年的发展,取得了可喜的进步。经典的数据融合方法在地图合并、影像和非影像型场数据融合等方面已经并将继续发挥积极作用。扩展型的方法,如地统计方法(包括多点地统计方法和GIM)、统计-机理型方法、基于树模型和Kalman滤波的多尺度信号分析与重构方法、

信息论方法均可用于数据融合,尤其是场分布数据,并将可以在空间点位数据的合并中加以灵活运用。

多源异构数据的几何、属性、语义、尺度、管治、误差等方面的差异或不一致性使得非常规的融合方法的研究很有必要,需将其与常规方法有机整合,用以建立数据融合系统化研究的架构。另外,在整合数据融合的有关概念和原理、建立数据融合架构的基础上,应发展相关技术,并以典型应用为牵引,完善理论研究。

空间数据融合与信息融合的研究意义重大,也任重道远。本文的视角和聚焦点为:空间数据对象模型和场模型的辩证统一观点,场分布数据的子类(原始影像数据和非原始影像数据)划分,数据本身(点位、属性、连续数值、离散类别等)及其重要维度(包括尺度、语义、时空特性等)的探究,经典方法和扩展方法的历史与未来考量。可以预见,系统化的研究将有利于完善地理信息科学的理论和方法体系,并将对地理信息产业和应用产生积极影响。

参 考 文 献

- [1] Li Deren, Gong Jianya, Zhang Qiaoping. On the Conflation of Geographic Databases[J]. *Science of Surveying & Mapping*, 2004, 29(1):1-4(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] Adams B, Li L, Raubal M, et al. A General Framework for Conflation[J]. *GIScience*, 2010(9):14-17
- [3] Grant M, Krzysztof J, Benjamin A. A Weighted Multi-attribute Method for Matching User-Generated Points of Interest[J]. *Cartography and Geographic Information Science*, 2014, 41(2): 125-137
- [4] Zhang Jingxiong. Scale, Uncertainty and Fusion in Spatial Information[M]. Wuhan: Wuhan University Press, 2008(张景雄. 空间信息的尺度、不确定性与融合[M]. 武汉: 武汉大学出版社, 2008)
- [5] Chen Huanxin, Liu Dongyong, Xu Mingshi, et al. Research on the Flowline and Status of Spatial Data Fusion[J]. *Geomatics World*, 2013, 20(5): 26-31(陈焕新, 刘栋永, 徐明世, 等. 空间数据融合的框架流程及发展现状研究[J]. 地理信息世界, 2013, 20(5): 26-31)
- [6] Tong X, Liang D, Jin Y. A Linear Road Object Matching Method for Conflation Based on Optimization and Logistic Regression[J]. *International Journal of Geographical Information Science*, 2014,

- 28(4): 824-846
- [7] Zhang J X, Atkinson P M, Goodchild M F. Scale in Spatial Information and Analysis[M]. Boca Raton: CRC Press, 2014
- [8] Saalfeld A. Conflation Automated Map Compilation [J]. *International Journal of Geographical Information System*, 1988, 2(3): 217-228
- [9] Tong X, Shi W, Deng S. A Probability-Based Multi-measure Feature Matching Method in Map Conflation [J]. *International Journal of Remote Sensing*, 2009, 30(20): 5 453-5 472
- [10] Podobnikar T, Vrečko A. Digital Elevation Model from the Best Results of Different Filtering of a Li-DAR Point Cloud[J]. *Transactions in GIS*, 2012, 16(5): 603-617
- [11] Fu D, Chen B, Wang J, et al. An Improved Image Fusion Approach Based on Enhanced Spatial and Temporal the Adaptive Reflectance Fusion Model [J]. *Remote Sensing*, 2013, 5(12): 6 346-6 360
- [12] Chen Jun, Liu Wanzeng, Zhang Jianqing. Research Progress of the Model and Method for GIS Database Updating[J]. *Geomatics World*, 2008, 6(3): 1-6 (陈军, 刘万增, 张剑清, 等. GIS 数据库更新模型与方法研究进展[J]. 地理信息世界, 2008, 6(3): 1-6)
- [13] Liu Jianjun, Wang Donghua, Shang Yaoling, et al Updating and Refinement of National 1 : 50 000 DEM Database[J]. *Geomatics World*, 2012(1): 18-20(刘建军, 王东华, 商瑶玲, 等. 国家 1 : 50 000 数字高程模型更新与精化[J]. 地理信息世界, 2012 (1): 18-20)
- [14] Ruiz J J, Ariza F J, Ureña M A, et al. Digital Map Conflation: A Review of the Process and A Proposal for Classification[J]. *International Journal of Geographical Information Science*, 2011, 25(9): 1 439-1 466
- [15] Song W, Keller J M, Haithcoat T L, et al. An Automated Approach for the Conflation of Vector Parcel Map with Imagery[J]. *Photogrammetric Engineering and Remote Sensing*, 2013, 79(6): 535-543
- [16] 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): 204-209 (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): 204-209)
- [17] Xu Feng, Deng Min, Zhao Binbin, et al. A Detailed Investigation on the Methods of Object Matching [J]. *Journal of Geo-information Science*, 2009, 11(5): 657-663 (徐枫, 邓敏, 赵彬彬, 等. 空间目标匹配方法的应用分析[J]. 地球信息科学学报, 2009, 11(5): 657-663)
- [18] Zhang Feng, Liu Nan, Liu Renyi, et al. Research of Cadastral Data Modelling and Database Updating Based on Spatio-Temporal Process[J]. *Acta Geodaetica et Cartographica Sinica*, 2010, 39(3): 303-309 (张丰, 刘南, 刘仁义, 等. 面向对象的地籍时空过程表达与数据更新模型研究[J]. 测绘学报, 2010, 39(3): 303-309)
- [19] Luan Xuechen, Yang Bisheng, Li Qiuping. Pattern-Based Node Matching Approach for Road Networks [J]. *Acta Geodaetica et Cartographica Sinica*, 2013, 42(4): 608-614 (栾学晨, 杨必胜, 李秋萍. 基于结构模式的道路网节点匹配方法[J]. 测绘学报, 2013, 42(4): 608-614)
- [20] Zhao Donghua, Sheng Yehua. Research on Automatic Matching of Vector Road Networks Based on Global Optimization[J]. *Acta Geodaetica et Cartographica Sinica*, 2010, 39(4): 416-421 (赵东保, 盛业华. 全局寻优的矢量道路网自动匹配方法研究 [J]. 测绘学报, 2010, 39(4): 416-421)
- [21] Yang Min, Ai Tinghua, Liu Pengcheng. The Matching and Consistency Correcting in the Integration of Contour and River Network[J]. *Acta Geodaetica et Cartographica Sinica*, 2012, 41(1): 152-158 (杨敏, 艾廷华, 刘鹏程, 等. 等高线与水网数据集成中的匹配及一致性改正[J]. 测绘学报, 2012, 41(1): 152-158)
- [22] Guo Li, Li Hongwei, Zhang Zejian, et al. Geometry Matching Method for Transportation Road Network Data Based on Projection[J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(9): 1 113-1 117 (郭黎, 李宏伟, 张泽建, 等. 道路网信息投影匹配方法研究[J]. 武汉大学学报·信息科学版, 2013, 38(9): 1 113-1 117)
- [23] Guo Taisheng, Zhang Xinchang, Liang Zhiyu. Research on Change Information Recognition Method of Vector Data Based on Neural Network Decision Tree[J]. *Acta Geodaetica et Cartographica Sinica*, 2013, 42(6): 937-944 (郭泰圣, 张新长, 梁志宇. 神经网络决策树的矢量数据变化信息快速识别方法 [J]. 测绘学报, 2013, 42(6): 937-944)
- [24] Zhang Yun, Li Qingquan, Cao Xiaohang, et al. An Algorithm for Detecting the Geometric Difference between the Road Networks[J]. *Acta Geodaetica et Cartographica Sinica*, 2009, 37(4): 521-525 (张韵, 李清泉, 曹晓航, 等. 一种道路网信息几何差异检测算法[J]. 测绘学报, 2009, 37(4): 521-525)
- [25] Pohl C, van Genderen J L. Review Article Multi-sensor Image Fusion in Remote Sensing: Concepts, Methods and Applications[J]. *International Journal of Remote Sensing*, 1998, 19(5): 823-854

- [26] Simone G, Farina A, Morabito F C, et al. Image Fusion Techniques for Remote Sensing Applications [J]. *Information Fusion*, 2002, 3(1): 3-15
- [27] Zhang Jixian. The Trend of Development of Multi-Source Remote Sensing Data Fusion[J]. *Geomatics World*, 2011, 9(2): 18-20 (张继贤.多源遥感数据融合的发展趋势[J]. 地理信息世界, 2011, 9(2): 18-20)
- [28] Theiler J, Wohlberg B. Local Coregistration Adjustment for Anomalous Change Detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(8): 3 107-3 116
- [29] Hall D L, Llinas J. An Introduction to Multisensor Data Fusion[J]. *Proceedings of the IEEE*, 1997, 85(1): 6-23
- [30] James A P, Dasarathy B V. Medical Image Fusion: A Survey of the State of the Art[J]. *Information Fusion*, 2014, 19: 4-19
- [31] Carper W J. The Use of Intensity-Hue-Saturation Transformations for Merging SPOT Panchromatic and Multispectral Image Data[J]. *Photogrammetric Engineering and Remote Sensing*, 1990, 56(4): 459-467
- [32] Memarsadeghi N, Le Moigne J, Mount D M, et al. A New Approach to Image Fusion Based on Cokriging[C]. The Eighth International Conference on Information Fusion, Philadelphia, PA, 2005
- [33] Amro I, Mateos J, Vega M, et al. A Survey of Classical Methods and New Trends in Pansharpening of Multispectral Images [J]. *Eurasip J. Adv. Sig. Process*, 2011, 79: 1-8
- [34] Tsai R Y, Huang T S. Multiframe Image Restoration and Registration [J]. *Advances in Computer Vision and Image Processing*, 1984, 1(2): 317-339
- [35] Ur H, Gross D. Improved Resolution from Subpixel Shifted Pictures[J]. *Cvqip Graphical Models & Image Processing*, 1992, 54(2): 181-186
- [36] Bose N K, Chappalli M B. A Second-Generation Wavelet Framework for Super-Resolution with Noise Filtering [J]. *International Journal of Imaging Systems & Technology*, 2004, 14(2): 84-89
- [37] Schultz R R, Stevenson R L. Improved Definition Image Expansion[C]. IEEE International Conference on Acoustics, Speech, and Signal Processing, San Francisco, CA, 1992
- [38] Su Heng, Zhou Jie, Zhang Zhihao. Survey of Super-Resolution Image Reconstruction Methods[J]. *Acta Automatica Sinica*, 2013, 39(8): 1 202-1 213 (苏衡, 周杰, 张志浩. 超分辨率图像重建方法综述[J]. 自动化学报, 2013, 39(8): 1 202-1 213)
- [39] Yang J, Wright J, Huang T, et al. Image Super-resolution as Sparse Representation of Raw Image Patches[C]. IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, 2008
- [40] Zhang J, Zhao C, Xiong R, et al. Image Super-Resolution via Dual-Dictionary Learning and Sparse Representation[C]. IEEE International Symposium on Circuits and Systems, Seoul, South Korea, 2012
- [41] Nath A G, Nair M S, Rajan J. Single Image Super Resolution from Compressive Samples Using Two Level Sparsity Based Reconstruction[J]. *Procedia Computer Science*, 2015, 46: 1 643-1 652
- [42] Atkinson P M, Pardo-Iguzquiza E, Chica-Olmo M. Downscaling Cokriging for Super-Resolution Mapping of Continua in Remotely Sensed Images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(2): 573-580
- [43] Elaksher A F, Bethel J. Refinement of Digital Elevation Models in Urban Areas Using Breaklines Via a Multi-photo Least Squares Matching Algorithm [J]. *Journal of Terrestrial Observation*, 2010, 2(2): 67-80
- [44] Hosford S, Baghdadi N, Bourguine B, et al. Fusion of Airborne Laser Altimeter and Radarsat data for DEM Generation[C]. Geoscience and Remote Sensing Symposium, Toulouse, France, 2003
- [45] Robinson N, Regetz J, Guralnick R P. Earth ENVDEM90: A Nearly-Global, Void-Free, Multi-scale Smoothed, 90 m Digital Elevation Model from Fused ASTER and SRTM Data[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 87: 57-67
- [46] Yoo E H, Kyriakidis P C. Area-to-Point Kriging in Spatial Hedonic Pricing Models[J]. *Journal of Geographical Systems*, 2009, 11(4): 381-406
- [47] Zhang J, Goodchild M F. Uncertainty in Geographical Information[M]. Boca Raton: CRC Press, 2002
- [48] Li D, Zhang J, Wu H. Spatial Data Quality and Beyond[J]. *International Journal of Geographical Information Science*, 2012, 26(12): 2 277-2 290
- [49] Kyriakidis P C, Shortridge A M, Goodchild M F. Geostatistics for Conflation and Accuracy Assessment of Digital Elevation Models[J]. *International Journal of Geographical Information Science*, 1999, 13(7): 677-707
- [50] Coppin P, Jonckheere I, Nackaerts K, et al. Digital Change Detection Methods in Ecosystem Monitoring: A Review[J]. *International Journal of Remote Sensing*, 2004, 25(9): 1 565-1 596
- [51] Lu D, Mausel P, Brondizio E, et al. Change Detection Techniques[J]. *International Journal of Remote Sensing*, 2004, 25(12): 2 365-2 407

- [52] Chen Jun, Lu Miao, Chen Xuehong, et al. A Spectral Gradient Difference Based Approach for Land Cover Change Detection[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2013, 85: 1-12
- [53] Zheng Y, Zhang X. Using Combined Difference Image and k-Means Clustering for SAR Image Change Detection[J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(3): 691-695
- [54] Dronova I, Gong P, Wang L, et al. Object-Based Analysis and Change Detection of Major Wetland Cover Types and Their Classification Uncertainty During the Low Water Period at Poyang Lake, China[J]. *Remote Sensing of Environment*, 2011, 115(12): 3 220-3 236
- [55] Ling F, Li W, Du Y, et al. Land Cover Change Mapping at the Subpixel Scale With Different Spatial-Resolution Remotely Sensed Imagery[J]. *IEEE Geoscience and Remote Sensing Letters*, 2011, 8(1): 182-186
- [56] Fichera C R, Modica G, Pollino M. Land Cover Classification and Change-Detection Analysis Using Multi-temporal Remote Sensed Imagery and Landscape Metrics[J]. *European Journal of Remote Sensing*, 2012, 45(1): 1-18
- [57] Schenk T, Csatho B, Veen C V D, et al. Fusion of Multi-sensor Surface Elevation Data for Improved Characterization of Rapidly Changing Outlet Glaciers in Greenland[J]. *Remote Sensing of Environment*, 2014, 149: 239-251
- [58] Celik T. Change Detection in Satellite Images Using a Genetic Algorithm Approach[J]. *IEEE Geoscience and Remote Sensing Letters*, 2010, 2(7): 386-390
- [59] Gueguen L, Soille P, Pesaresi M. Change Detection Based on Information Measure[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(11): 4 503-4 515
- [60] Zhang Chi, Ge Yong, Bai Hexiang, et al. Multiple-point Simulation and its Application and Development in RS Image Classification[J]. *Remote Sensing Technology & Application*, 2010, 25(2): 296-302 (张弛, 葛咏, 白鹤翔, 等. 多点模拟及其在遥感影像信息提取中的应用及进展[J]. *遥感技术与应用*, 2010, 25(2): 296-302)
- [61] Tahmasebi P, Hezarkhani A, Sahimi M. Multiple-point Geostatistical Modeling Based on the Cross-correlation Functions[J]. *Computational Geosciences*, 2012, 16(3): 779-797
- [62] Tang Y, Zang J, Jing L, et al. Digital Elevation Data Fusion Using Multiple-Point Geostatistical Simulation[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2015(1): 1-13
- [63] Michalak A M, Bruhwiler L, Tans P P. A Geostatistical Approach to Surface Flux Estimation of Atmospheric Trace Gases[J]. *Journal of Geophysical Research*, 2004, 109(D14109): 1-8
- [64] Tang Y, Atkinson P M, Zhang J. Downscaling Remotely Sensed Imagery Using Area-to-Point Cokriging and Multiple-point Geostatistical Simulation[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, 101: 174-185
- [65] Nguyen H, Katzfuss M, Cressie N, et al. Spatio-Temporal Data Fusion for Very Large Remote Sensing Datasets[J]. *Technometrics*, 2014, 56(2): 174-185
- [66] Li S, Li Z, Gong J. Multivariate Statistical Analysis of Measures for Assessing the Quality of Image Fusion[J]. *International Journal of Image and Data Fusion*, 2010, 1(1): 47-66
- [67] Goodchild M, Zhang J, Kyriakidis P. Discriminant Models of Uncertainty in Nominal Fields[J]. *Transactions in GIS*, 2009, 13(1): 7-23
- [68] Willsky A S. Multiresolution Markov Models for Signal and Image Processing[J]. *Proceedings of the IEEE*, 2002, 90(8): 1 396-1 458
- [69] Wackernagel H, Bertino L. Geostatistics and Sequential Data Assimilation[M]. Netherlands: Springer, 2005
- [70] Bai Yulong, Li Xin, Han Xujun. A Review of Error Problems for Land Data Assimilation Systems[J]. *Advances in Earth Science*, 2011, 26(8): 795-804 (摆玉龙, 李新, 韩旭军. 陆面数据同化系统误差问题研究综述[J]. *地球科学进展*, 2011, 26(8): 795-804)
- [71] Chen H M, Varshney P K, Arora M K. Performance of Mutual Information Similarity Measure for Registration of Multitemporal Remote Sensing Images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2003, 41(11): 2 445-2 454
- [72] Bramon R, Boada I, Bardera A, et al. Multimodal Data Fusion Based on Mutual Information[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(9): 1 574-1 587
- [73] Gotway C A, Young L J. A Geostatistical Approach to Linking Geographically Aggregated Data from Different Sources[J]. *Journal of Computational and Graphical Statistics*, 2007, 16(1): 1-21
- [74] Calvert K, Luciani P, Mabee W. Thematic Land-Cover Map Assimilation and Synthesis: The Case of Locating Potential Bioenergy Feedstock in Eastern

Ontario[J]. *International Journal of Geographical Information Science*, 2014, 28(2): 274-295

[75] Villarreal M L, Lii CVR, Petrakis R E. Conflation and Aggregation of Spatial Data Improve Predictive Models for Species with Limited Habitats: A Case of the Threatened Yellow-Billed Cuckoo in Arizona, USA[J]. *Applied Geography*, 2014, 47: 57-69

[76] Goodchild M F. Citizens as Sensors: The World of Volunteered Geography[J]. *Geo Journal*, 2007, 69(4): 211-221

[77] Chen Nengcheng, Yang Xunliang, Wang Xiaolei, Design and Implementation of Geospatial Sensor Web Information Public Service Platform[J]. *Journal of Geo-information Science*, 2013, 15(6):887-894(陈能成, 杨训亮, 王晓蕾. 地理空间传感网信息公共服务平台的设计与实现[J]. 地球信息科学学报, 2013, 15(6): 887-894)

[78] Ledesma-Carbayo M J, Kybic J, Desco M, et al. Spatio-Temporal Nonrigid Registration for Ultrasound Cardiac motion Estimation[J]. *IEEE Transactions on Medical Imaging*, 2005, 24(9): 1 113-1 126

Progress in Spatial Data Fusion: From Classic Approaches to Extended Methods

ZHANG Jingxiong^{1,2} LIU Fengzhu¹ MEI Yingying¹ TANG Yunwei³

1 School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

2 School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China

3 Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academic of Science, Beijing 100039, China

Abstract: With diversification of data acquisitions and associated technologies, there is steady accumulation of multi-source spatial data, prompting more research on effective and fast data integration to provide information and technical support for geospatial research and applications. In this paper, procedures for integrating multi-source data are referred to as data fusion. Research developments are reviewed by considering them as classic and extended types of methods, respectively. To review the progress in classic methods for data fusion and related data processing, this paper clarifies some of the relatively well-established methods using a typology of object- and field-based models of spatial data. For extended methods, the paper discusses multi-point geostatistics, statistical-mechanistic methods, multi-scale signal analysis and reconstruction, and information-theoretic strategies. Some of the issues, such as scale mismatch, semantic inconsistency, and the temporal dimension, are discussed in the hope of better supporting the fusion of multi-sourced and heterogeneous data. Further research will enhance theoretical foundations of geographic information to enrich methodologies for spatial data and their analyses, and help add to the applicability and value of spatial information.

Key words: data fusion; geostatistics; mechanistic; information theory; data support

First author: ZHANG Jingxiong, PhD, professor, specializes in remote sensing, spatial statistics, and GIScience. E-mail: jxzhang@whu.edu.cn

Foundation support: The National Natural Science Foundation of China, Nos. 41471375, 41501489.