



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

## 《武汉大学学报(信息科学版)》网络首发论文

题目: 测绘大数据时代数据处理理论面临的挑战与发展  
作者: 朱建军, 宋迎春, 胡俊, 邹滨, 吴立新  
DOI: 10.13203/j.whugis20210232  
收稿日期: 2021-05-11  
网络首发日期: 2021-06-09  
引用格式: 朱建军, 宋迎春, 胡俊, 邹滨, 吴立新. 测绘大数据时代数据处理理论面临的挑战与发展. 武汉大学学报(信息科学版).  
<https://doi.org/10.13203/j.whugis20210232>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI: 10.13203/j.whugis20210232

## 测绘大数据时代数据处理理论面临的挑战与发展

朱建军<sup>1</sup> 宋迎春<sup>1</sup> 胡俊<sup>1</sup> 邹滨<sup>1</sup> 吴立新<sup>1</sup>

1 中南大学地球科学与信息物理学院, 湖南 长沙, 410083

**摘要：**随着信息技术的发展、测绘大数据和人工智能的兴起，数据缺乏不再是一个问题。可是，现有的测绘数据处理技术一直追求数据的准确性（微观），而大数据研究则恰恰允许数据的混杂性、不确定性（宏观）。因此，尽管传统测绘数据处理理论在微观数据处理方面积累了大量的技术优势，而大数据的规模性和复杂性使得传统的计算模型和分析算法无法有效地支撑大数据的高效分析处理。作为开启智能时代“大门钥匙”的数据处理理论与方法，如何适应新技术的挑战与机遇是值得深入思考的问题。在大数据驱动下，大规模的数据挖掘、机器学习和深度学习等新思想和新方法正在蓬勃发展，极大地促进了场景内外多源异质大数据的融合，从而有效地从多种传感器数据中提取地表特征信息，不断提升测绘信息获取和分析能力。因此，测绘数据处理理论也需要同步跟进，现有的数据处理方法也需要进行智能化。结合智能测绘的前沿热点、发展趋势和存在的挑战，探索数据处理理论扩展的方向，一是希望能够推动测绘数据处理理论的进一步发展，二是希望为有兴趣研究测绘大数据领域的研究生提供学习参考。

**关键字：**智能测绘；大数据；人工智能；数据处理理论；挑战；发展

中图分类号：P208 文献标志码：A

智能测绘是人工智能与测绘科学技术相融合的新技术，它使测量数据的类型、规模发生了深刻变化，传统的测绘数据处理方法转向了智能数据处理方向，一些新的数据处理方法在测绘数据处理中不断呈现，如基于深度机器学习的方法、人工神经网络方法、智能目标识别算法等<sup>[1-4]</sup>。在实际测绘工程中，位置服务和测绘技术需要测量数据具有准确性，而这些新型的测量数据并不满足这一要求，它通常具有混杂性、不确定性<sup>[5]</sup>。对于小规模数据，虽然经典的数据处理理论有着丰富的算法，但对于大数据时代的新型数据也无能为力。在智能测绘中，只能近似或逼近地描述工程实际问题，无法建立精确的数学模型，这是因为模型参数和环境等多种因素的影响，总会出现一些不可忽略、难以估计或预测的不确定性。虽然现有的观测数据丰富，可很难满足独立性和平稳性等经典统计假设。如何实现某些非结构化和半结构化测绘数据的结构化，如何从大量非规范的数据中提取符合规范的数据，如何对位置大数据的复杂性、不确定性特征进行描述，如何降低位置大数据的异构性并从中获取有用的信息，是测绘大数据时代面临的挑战<sup>[5]</sup>。从数据处理研究的角度来说，这些问题可以描述为：一是对于带有不确定性的大数据如何保证参数估计或自适应学习算法的有效性，以及保证大数据中含有多少有效信息量；二是如何建立自适应学习或识别算法与控制算法的理论；三是如何计算由数据驱动的反馈控制机制应对不确定性的最大能力；四是如何调控复杂系统中具有“博弈”行为的对象等。上述4个问题是数据处理中基本的科学问题<sup>[6]</sup>，这需要我们研究宏观数据处理层面的新方法、新思维。作为开启智能时代“大门钥匙”的数据处理理论与方法，如何适应人工智能新技术的挑战与机遇，是值得深入思考的问题。

经典参数估计理论必须建立一个准确的数学模型（平差模型）。测绘大数据来源于多种不确定性数据，且混杂许多的错误信息和粗差，其表现为不准确、不精确、不完整以及过时和冗余特征，这些特征对于建

立准确的数学模型通常非常困难<sup>[6-7]</sup>。大数据研究中使用的方法主要是数据挖掘和机器学习，其强调的是数据之间的关联性，而对于数据的精确性和数据间的因果联系并不注重，表现为重数据、轻模型。随着大数据与超级计算方法等技术的飞速发展，基于深度神经网络的机器学习算法得到了迅速发展。来自各种反馈机制的复杂开放动态系统的学习数据通常不会满足经典统计分析所关注的独立性和平稳性等假设。因此，我们需要利用类似自适应估计的思想，研究和探讨测绘大数据场景下机器学习算法的预测或收敛情况。

## 1 测绘大数据的特征以及测量数据处理面临的挑战

测量数据处理理论是关于测绘数据处理方法的科学，它融合了数学、统计学、概率论、误差估计理论、测量平差基础等计算理论的现代数据科学，其核心任务就是对测绘数据进行分析处理，从中提取所需要的数据信息。李德仁<sup>[8]</sup>在《展望大数据时代的地理空间信息学》中，将测绘大数据的特点归纳为5个——体量大、速度快、模态多样、真伪难辨、价值巨大，并针对测绘大数据的数据海量、信息缺失、知识难觅特征，分析了时空数据挖掘理论和算法研究的重要性。从测量数据处理理论的角度，可以对测绘大数据概括为三大特征：

### 1) 大规模、分布式、泛在性

测绘大数据也称为时空大数据，它可以是无组织的专业测绘大数据，如多分辨率、多用途的卫星遥感、定位信息、摄影测量、地球物理、地形测绘等专业数据集合，也可以是空-天-地-海一体化的有组织的协同观测大数据，高精度实时测绘体系实现从静态到动态、地基到天基、区域到全局、室内到室外、被动到主动的快速智能测绘。这些大数据来源于多个学科，其特点是跨尺度、跨平台、跨参数，规模巨大、类型多样的大数据为地表变化的研究提供了海量信息支持。然而，测绘大数据都是从复杂的工程环境中产生的，数据之间、系统之间、层次之间、环境之间具有较强的耦合作用，数据质量和数据分析的结果无法得到有效的保证<sup>[9]</sup>。由探测数据、飞行数据、仪器状态数据组成的原始数据具有复杂多变性，准确解耦出测绘成果是非常困难的。大数据模态多样、真伪难辨的特性是复杂系统构建的难题，如何解释不同平台、不同尺度、不同参数之间的随机与非线性关联作用影响大数据的演化形态是现代测绘数据处理理论需要解决的问题。

### 2) 多源、异构、周期性

随着测绘对象范畴的扩大，多尺度、个性化、智能化、全天候的测绘服务型需求越来越多，复杂对象的数据感知面临大数据多源、异构等复杂的数据形态。不同学科和不同技术之间、不同观测工具和不同分辨率之间以及不同对象之间的观测将使得观测数据呈现多元化的特点。这些数据的类型、统计性质和数据结构完全不相同，形成了异构数据，它们随时间和空间交替变化表现出某种循环、重复或叠加，呈现周期性变化的特点。因此，测绘大数据通常表现出多源、异构、周期涌现等特征。传统的测绘数据处理理论已无法有效分析和预测大数据之间的变化规律和真实价值。大数据需要新型的不依赖于全体数据的算法理论支持，需要降低大数据中的不确定性影响，以提高非确定化和局部增量的机器学习能力，需要突破小规模数据处理中独立同分布的假设<sup>[7]</sup>。

### 3) 高速、更新快、时效短

空-天-地-海不同的观测技术以及各类观测设备的高速发展，使得数据获取的速度、更新周期越来越快，

信息量也越来越大,时效性越来越短。例如,当前在轨能够获取到高分辨率影像的卫星重访周期约为 1~5 d,而无人机、移动测量车、背包式的三维激光扫描仪等新型测绘技术则可以实时获取到相应的地理信息数据。再如,卫星对地观测数据也是一种高速、更新快、时效短的大数据,它不仅具有海量大数据的特征,同时也具有实时更新的数据流特征<sup>[10]</sup>。数据的快速处理与决策的及时性是大数据高速性的前提,如地图导航中路况的实时信息、共享出租车的行驶轨迹数据等,每隔 10 s 至 1 min 数据中心就需要进行更新。增强测绘数据的处理速度必须建立在测绘大数据的快速分析之上,如为了更好地预报地震发生的地点、时间和强度,地震观测中不仅需要快速获取数据,更需要数据的快速分析和处理。时空大数据处理的实时性需要利用各种方法进行有效挖掘,如何快速地处理海量时空数据是当前数据处理中需要重点研究的一个问题。

## 2 测绘大数据时代数据处理理论的发展方向

当前大数据的多源、异构以及跨平台、跨尺度、跨参数等形态特征导致了大数据处理的复杂性,现有的测绘数据处理方法是理论驱动和模型驱动,其优点是能够将不充足的数据与已知理论结合起来建立相关模型,但也有局限性,如不能处理大数据量,不能处理很多新类型数据、多媒体数据和高分辨率的遥感数据等。当前测绘数据量越来越大,并且还将更快地增长,而测绘技术对数据处理的有效性、精确性、实时性要求越来越高,传统的常规数据处理技术已无法适应测绘新技术的发展。

### 2.1 多源异构大数据的感知、融合与快速识别

地表观测大数据感知的核心是从大规模、分布式、泛在性的大数据中获取有价值的信息。如何从细粒度的个体层面更有效地感知数据并利用群体智能来提升大数据的感知能力,是数据处理理论需要解决的问题。

#### 1) 高质量数据的感知和获取技术

由于测绘对象范畴的扩大,数据感知将面临多源、异构以及跨平台、跨尺度、跨参数等复杂的数据特征,这使得数据采集容易出现缺失、冗余、粗差等问题。因而,数据的信噪比降低,利用大数据精准还原系统本征属性的智能学习方法变得困难<sup>[9]</sup>。人工智能算法在数据处理和理解等方面有较强的优势,如何对其进行改进以有效处理复杂数据特征的大数据,缩短信息提取的时间,是数据处理方法需要研究的新问题。例如,日均新增超过 100 TB 的出租车轨迹大数据集蕴藏着道路交通状态、城市居民出行规律、城市结构及其他高价值信息,可以结合神经网络、贝叶斯分类、支持向量机和遗传算法等对交通流量、市民出行目的、城市区域用车需求预测、电动出租车充电站选址等信息进行提取<sup>[11]</sup>。此外,德国宇航局 Koyama 等<sup>[12]</sup>首次利用星载合成孔径雷达(synthetic aperture radar, SAR)影像大数据集(300 TB)和降雨数据集联合深度学习算法,深入挖掘了降雨条件对 L 波段 SAR 数据后向散射性能的影响,并提出结合人工智能算法对受降雨影响的 SAR 影像自适应修正定标因子。现有的小规模数据或静态数据处理理论显然不能适应上述数据处理方法的变化,需要研究针对多源异构大数据的精确感知和高效获取新算法,分析算法精度、稳定性与数据规模之间的关系,研究兼顾精度和性能的高效获取模型。

#### 2) 大数据信息融合与质量控制

跨平台的多时序观测数据之间的信息融合可以弥补单一模式观测数据的不充分或数据缺失的问题。多



维、多源、跨平台的异质观测数据中既有可量化的数值型观测数据，也有其他非数值型数据，如图像、视频数据和描述性的文字信息等，因此异构信息的融合非常困难。大数据一般来源于复杂环境，与地质环境关联<sup>[7]</sup>。数据与数据之间、不同的尺度之间、不同的系统环境之间存在错综复杂的耦合作用，新旧数据的更新交换等问题严重影响着数据的质量和数据分析的结果。数据的复杂性、不确定性以及不同分辨率、不同物理意义、不同统计性质的数据，使得多传感器数据融合变得困难。此外，时序观测数据中还存在着各种噪声、不确定度、数据不完整和观测不同步等问题，因此大数据的质量问题更需要引起重视。在大数据数据挖掘方面，随着人工智能算法研究的不断推进，深度神经网络在大数据分析中表现出的独特优势越来越明显，从动力学角度对多源异构数据进行解耦与重建，利用数理统计规律，在时间尺度上降低各时间段间的相似性和大数据的样本数，从海量数据中智能学习提取关键信息，可以形成测绘大数据的一般性方法。例如，在地质灾害监测方面，通过对地震重灾区 62 个县市内的滑坡灾害与地质环境、降雨之间的关联分析，运用改进 BP(back propagation)神经网络、动态粒子群优化 BP 神经网络、深度置信网络等机器学习理论和方法，可建立滑坡灾害气象预警模型<sup>[13]</sup>。在全球地形测绘方面，可利用神经网络模型融合多源全球数字高程模型(digital elevation model, DEM)产品(SRTM1、ASTERGDEMv2 及激光雷达测高数据 ICESatGLAS 等)，获得相比传统 DEM 精度更高的融合 DEM 产品，其具有精度稳定、无空间空洞等优势<sup>[14]</sup>。现有的数据融合和质量控制技术是建立在数据规模小、数据结构化语义清楚的基础之上<sup>[7]</sup>。为了有效解决大数据信息获取的全面性和一致性问题，需要对大数据的多维、多源、跨平台以及相关性开展大数据融合和集成的新方法以及质量控制问题的研究。

### 3) 基于智能算法的目标快速识别技术

基于深度学习的目标识别算法通过模仿人脑神经网络结构对大数据进行学习、分析和解释，已成功应用于地下目标识别<sup>[15]</sup>和遥感地物目标识别等领域<sup>[16]</sup>。智能视觉融合了计算机和人工智能领域的最新成果，将它与深度学习方法相结合，智能目标快速识别算法将更加有效，只需要简单的加、乘运算，就能够快速生成目标分类结果。目前，除了输入实数域特征（高光谱参数、激光雷达数据等）进行模型训练及目标分类外<sup>[17-18]</sup>，学者已针对 SAR 数据复数特性提出复数域深度神经网络，进一步提高了极化 SAR 影像数据的目标分类精度<sup>[19-20]</sup>。然而，深度学习方法也有许多不足，如要求较高的训练数据，数据不充分或数据质量低时应用效果差。因此，我们需要在无监督学习、弱监督学习领域探索新型的智能算法，这样才能充分利用目标识别算法和智能算法为智能测绘服务。

## 2.2 大数据内容的理解与特征提取

人工智能算法能够对特征逐层理解并进行自动分析，完成几何观测、特征和要素解译等任务，实现从信息感知到结论判断的高级理解过程。在高精度几何观测和多维特征感知方面，机器有独特的优势，如何充分利用人机各自的优势，快速有效地研判地表形变的变化趋势，是机器人实现自主测绘的关键科学问题。

### 1) 基于智能算法的特征提取技术

随着卫星遥感和航空遥感测绘技术的发展，土地资源、森林覆盖、环境变化的调查和动态监测，城市扩张和土地利用变化的趋势研究获得了海量的观测数据。这些海量遥感数据具有多源、多时相、多波段、

多分辨率的特征,给数据处理带来了新挑战。基于深度学习模型对多源遥感数据进行时空特征的提取,可以获得高效的时空决策支持信息,能够动态感知和理解地球复杂系统之间的相互作用。例如,通过引入神经网络融合全波形 LiDAR (light detection and ranging) 和极化干涉 SAR 数据信息,可以对大范围森林区植被高度、冠层覆盖度、植被垂直结构等空间特征进行较为准确的预测,极大地简化了传统基于模型反演的森林区特征提取过程<sup>[21]</sup>。此外,基于遥感大数据的地表连续覆盖特征提取技术也逐渐成熟,不同传感器的优势和信息互补,使得深度学习成为目前能满足数据驱动下信息提取的最优智能模型之一<sup>[22]</sup>。

## 2) 基于智能算法的空间特征理解技术

基于智能算法的空间特征理解技术,一方面可以充分挖掘与利用信息,将目标状态信息转换为空间特征理解,并对时域空间进行预测;另一方面又可以通过表征学习,提取潜在的地理空间特征,对模型进行准确预测。例如,将基于迁移学习的卷积神经网络融入侧扫声呐沉船图像识别算法中,能够克服传统侧扫声呐图像人工判读存在效率低、耗时长、资源消耗大及主观不确定性强和过分依赖经验等问题<sup>[23]</sup>。从案例学习的角度出发,段佩祥等<sup>[24]</sup>利用朴素贝叶斯机器学习方法进行训练,得到主支流分类模型,进而对河段主支流关系进行识别,实现对树状河系的主支流自动识别、分类与分级。因此,空间特征学习或表征学习是基于智能算法的空间特征理解的关键技术。

## 3) 基于智能算法的时域空间预测技术

根据多维属性变量估计一个目标对象或地理变量在未知地点(时空)的数值是时域空间预测的基本原理。利用已知位置的属性数值推测未知点相同属性的数值称为空间插值法,它是 GIS 中常见的空间分析功能,包括反距离加权、不规则三角网络、和克里金法等。我们可以运用机器学习和深度学习探索时域空间预测和空间插值的新方法,并将其广泛应用在测绘领域。部分学者已在相关领域取得一定的研究成果,例如刘青豪等<sup>[25]</sup>提出一种基于深度学习的 InSAR (interferometric synthetic aperture radar) 地表沉降预测方法,其克服了传统预测方法因模型参数难以获取或相关数据缺乏而难以得到可靠预测结果的缺陷,对大范围时序形变的短期预测较为有效。在城市时空预测方面,基于极限学习机的城市扩展元胞自动机模型已被用于城市空间结构扩展模拟与预测方面,其相比于神经网络的模拟精度更高,训练时间更短,且在小样本条件下仍能保持较高的模拟预测精度<sup>[26]</sup>。在 GNSS(global navigation satellite system) 导航定位中,对一个或多个运动目标进行采样,可以建立运动学模型对目标的运动状况进行预测。为此,我们需要利用人工智能技术去理解目标的运动行为,学习目标运动的行为模式,挖掘目标行为与目标行动间的关联关系,从而对目标的运动行为进行预测。

# 3 结 语

为了适应大数据时代的新挑战,测绘学科需要根据测绘数据的特点建立数据科学体系,培养测绘领域的的数据科学人才<sup>[27]</sup>。当今测绘大数据和智能测绘正在兴起,数据处理智能化逐渐成为测绘学科的焦点。本文就传统的测量数据处理算法无法有效支撑大数据高效分析处理的问题展开讨论,结合智能测绘的前沿热点、发展趋势和存在的挑战,探索数据处理理论扩展的方向,一是希望能够推动测绘数据处理理论的进一步发展,二是希望为有兴趣研究测绘大数据领域的研究生提供学习参考。

## 参考文献

- [1] Liu Jingnan, Guo Wenfei, Guo Chi, et al. Rethinking Ubiquitous Mapping in the Intelligent Age[J]. *Acta Geodaetica et Cartographica Sinica*, 2020, 49(4): 403-414 (刘经南, 郭文飞, 郭迟, 等. 智能时代泛在测绘的再思考[J]. *测绘学报*, 2020, 49(4): 403-414)
- [2] Gong Jianya. Chances and Challenges for Development of Surveying and Remote Sensing in the Age of Artificial Intelligence[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(2): 1 788-1 796 (龚健雅. 人工智能时代测绘遥感技术的发展机遇与挑战[J]. *武汉大学学报 信息科学版*, 2018, 43(2): 1 788-1 796)
- [3] Li Deren, Yao Yuan, Shao Zhenfeng. Big Data in Smart City[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(6): 631-640 (李德仁, 姚远, 邵振峰. 智慧城市中的大数据[J]. *武汉大学学报 信息科学版*, 2014, 39(6): 631-640)
- [4] Ning Jinsheng. Research on the Development Strategy of Surveying and Mapping Science and Technology Transformation and Upgrading[J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(1): 1-9 (宁津生. 测绘科学与技术转型升级发展战略研究[J]. *武汉大学学报 信息科学版*, 2019, 44(1): 1-9)
- [5] Liu Jingnan. Big Data and Location Services[J]. *Science of Surveying and Mapping*, 2014, 39(3): 3-9 (刘经南. 大数据与位置服务[J]. *测绘科学*, 2014, 39(3): 3-9)
- [6] Guo Lei. Estimation, Control, and Games of Dynamical Systems with Uncertainty[J]. *Scientia Sinica: Informationis*, 2020, 50(9): 1 327-1 344 (郭雷. 不确定性动态系统的估计、控制与博弈[J]. *中国科学: 信息科学*, 2020, 50(9): 1 327-1 344)
- [7] Jin Xiaolong, Wang Yuanzhuo, Cheng Xueqi. Research System and Status of Big Data[J]. *Information and Communications Technologies*, 2013, 7(6): 35-43 (靳小龙, 王元卓, 程学旗. 大数据的研究体系与现状[J]. *信息通信技术*, 2013, 7(6): 35-43)
- [8] Li Deren. Towards Geo-Spatial Information Science in Big Data Era[J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(4): 379-384 (李德仁. 展望大数据时代的地球空间信息学[J]. *测绘学报*, 2016, 45(4): 379-384)
- [9] Palpanas T, Vlachos M, Keogh E, et al. Online Amnesic Approximation of Streaming Time Series[C]. *IEEE International Conference on Data Engineering*, Boston, MA, USA, 2004
- [10] Zheng Zhiming, Lü Jinhu, Wei Wei, et al. Refined Intelligence Theory: Artificial Intelligence Towards Complex Dynamic Objects[J]. *Scientia Sinica: Informationis*, 2021, 51(4): 678-690 (郑志明, 吕金虎, 韦卫, 等. 精准智能理论: 面向复杂动态对象的人工智能[J]. *中国科学: 信息科学*, 2021, 51(4): 678-690)
- [11] Wu Huayi, Huang Rui, You Lan, et al. Recent Progress in Taxi Trajectory Data Mining[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(11): 1 341-1 356 (吴华意, 黄蕊, 游兰, 等. 出租车轨迹数据挖掘进展[J]. *测绘学报*, 2019, 48(11): 1 341-1 356)
- [12] Koyama C N, Watanabe M. Assessing the Impact of Precipitation on L-band SAR Forest Observation: An ALOS-2 Big Data Case Study in the Tropics [C]. *The 13th European Conference on Synthetic Aperture Radar (EUSAR 2021)*, Online Conference, 2021
- [13] Huang Lu. Research on Meteorological Early-Warning Model of Landslides in Wenchuan Earthquake Area Based on Machine Learning[D]. *Wuhan: China University of Geosciences*, 2016 (黄露. 基于机器学习的汶川震区滑坡灾害气象预警模型研究[D]. 武汉: 中国地质大学, 2016)
- [14] Shen Huanfeng, Liu Lu, Yue Linwei, et al. A Multi-source DEM Fusion Method Based on Elevation Difference Fitting Neural Network[J]. *Acta Geodaetica et Cartographica Sinica*, 2018, 47(6): 854-863 (沈焕锋, 刘露, 岳林蔚, 等. 多源 DEM 融合的高差拟合神经网络方法[J]. *测绘学报*, 2018, 47(6): 854-863)
- [15] Yang Bisheng, Zong Zeliang, Chen Chi, et al. Real Time Approach For Under Ground Objects Detection from Vehicle-Borne Ground Penetrating Radar[J]. *Acta Geodaetica et Cartographica Sinica*, 2020, 49(7): 874-883 (杨

- 必胜,宗泽亮,陈驰,等. 车载探地雷达地下目标实时探测法[J]. 测绘学报, 2020, 49(7):874-883)
- [16] Qiu C, Tong X, Schmitt M, et al. Multi-level Feature Fusion-based CNN for Local Climate Zone Classification from Sentinel-2 Images: Benchmark Results on the So2Sat LCZ42 Dataset[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 2 793-2 806
- [17] Zhu X X, Tuia D, Mou L, et al. Deep Learning in Remote Sensing: A Review[J]. IEEE Geoscience and Remote Sensing Magazine, DOI:10.1109/MGRS.2017.2762307, 2017
- [18] Zhao Chuan, Guo Haitao, Lu Jun, et al. Airborne LiDAR Point Cloud Classification Based on Deep Residual Network[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(2):202-213 (赵传,郭海涛,卢俊,等. 基于深度残差网络的机载 LiDAR 点云分类[J]. 测绘学报, 2020, 49(2):202-213)
- [19] Zhu Xiaoxiang, Montazeri S, Ali M, et al. Deep Learning Meets SAR[J]. IEEE Geoscience and Remote Sensing Magazine, 2021
- [20] Zhang L, Dong H, Zou B. Efficiently Utilizing Complex-Valued PolSAR Image Data via a Multi-task Deep Learning Framework[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 157:59-72
- [21] Brigot G, Simard M, Colin-Koeniguer E, et al. Retrieval of Forest Vertical Structure from Polinsar Data by Machine Learning Using LiDAR-Derived Features[J]. Remote Sensing, 2019, 11(4): 381
- [22] Zhang Bing. Remotely Sensed Big Data Era and Intelligent Information Extraction[J]. Geomatics and Information Science of Wuhan University, 2018, 43(12):108-118 (张兵. 遥感大数据时代与智能信息提取[J]. 武汉大学学报·信息科学版, 2018, 43(12): 108-118)
- [23] Tang Yulin, Jin Shaohua, Bian Gang, et al. The Transfer Learning with Convolutional Neural Network Method of Side-Scan Sonar to Identify Wreck Images[J]. Acta Geodaetica et Cartographica Sinica, 2021, 50(2):260-269 (汤寓麟,金绍华,边刚,等. 侧扫声呐识别沉船影像的迁移学习卷积神经网络法[J]. 测绘学报, 2021, 50(2):260-269)
- [24] Duan Peixiang, Qian Haizhong, He Haiwei, et al. Naïve Bayes-based Automatic Classification Method of Tree-Like River Networks Upported by Cases[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(8):975-984 (段佩祥,钱海忠,何海威,等. 案例支撑下的朴素贝叶斯树状河系自动分级方法[J]. 测绘学报, 2019, 48(8):975-984)
- [25] Liu Qinghao, Zhang Yonghong, Deng Min, et al. Time Series Prediction Method of Large-Scale Surface Subsidence Based on Deep Learning[J]. Acta Geodaetica et Cartographica Sinica, 2021, 50(3):396-404 (刘青豪,张永红,邓敏,等. 大范围地表沉降时序深度学习预测法[J]. 测绘学报, 2021, 50(3):396-404)
- [26] Wang He, Zeng Yongnian. Urban Expansion Model Based on Extreme Learning Machine[J]. Acta Geodaetica et Cartographica Sinica, 2018, 47(12):1 680-1 690 (王鹤,曾永年. 城市扩展极限学习机模型[J]. 测绘学报, 2018, 47(12):1 680-1 690)
- [27] Zhao Pengda. Big Data Era Calls for Data Science in Various Fields of Science[J]. Chin Sci Tech Award, 2014(183): 29-30(赵鹏达. 大数据时代呼唤各科学领域的科学[J]. 中国科技奖励, 2014(183): 29-30)

## Challenges and Development of Data Processing Theory in the Era of Surveying and Mapping Big Data

ZHU Jianjun<sup>1</sup> SONG Yingchun<sup>1</sup> HU Jun<sup>1</sup> ZOU Bin<sup>1</sup> WU Lixin<sup>1</sup>

<sup>1</sup> School of Geoscience and Info-Physics, Central South University, Changsha 410083, China

**Abstract:** With the development of information technology, the rise of surveying and mapping big data and artificial intelligence, the lack of data is no longer a problem. However, the existing surveying and mapping data processing technology has been pursuing the accuracy of data (micro), and big data research just allows the data to be mixed and uncertain (macro). Therefore, although the traditional surveying and mapping data processing theory



has accumulated a large number of technical advantages in micro data processing, the large-scale and complexity of big data has become increasingly prominent, in which traditional calculation model and analysis algorithm cannot effectively support the efficient analysis and processing of big data. As the key to the intelligent era, data processing theory and method, how to adapt to the challenges and opportunities of new technology is worth our deep thinking. Driven by big data, new ideas and methods such as large-scale data mining, machine learning and deep learning are booming, which greatly promote the fusion of multi-source heterogeneous big data inside and outside the scene, effectively extract surface feature information from a variety of sensor data, and constantly improve the ability of surveying and mapping information acquisition and analysis. We think that the theory of surveying and mapping data also needs to be followed up, and the existing data processing methods need to be intelligent. Combined with the frontier hot spots, development trends and existing challenges of intelligent surveying and mapping, this paper explores the expansion direction of data processing theory. One is to promote the further development of surveying data processing theory, and the other is to provide reference for graduate students who are interested in entering the field of Surveying and mapping big data.

**Key words:** intelligent surveying and mapping; big data; artificial intelligence; data processing theory; challenges; development

First author: ZHU Jianjun, PhD, professor, specializes in error data processing and its application in InSAR. E-mail: zjj@csu.edu.cn

Corresponding author: SONG Yingchun, PhD, professor. E-mail: csusyc@mail.csu.edu.cn

Foundation support: The Key Research Project of Graduate Education and Teaching Reform in Hunan Province (2020JGZD014); the Research Project of Teaching Reform in Colleges and Universities of Hunan Province (HNJG-2020-0034); the Key Project of Graduate Education and Teaching Reform in Central South University (2020JGA011); the Key Project of Education and Teaching Reform in Central South University (2020JY001).

---

收稿日期: 2021-05-11

项目资助: 湖南省研究生教育教学改革研究重点项目 (2020JGZD014); 湖南省普通高等学校教学改革研究项目 (HNJG-2020-0034); 中南大学研究生教育教学改革研究重点项目 (2020JGA011); 中南大学教育教学改革研究重点项目 (2020JY001)。

第一作者: 朱建军, 博士, 教授, 主要从事误差数据处理及其在 InSAR 中的应用。zjj@csu.edu.cn

通讯作者: 宋迎春, 博士, 教授。csusyc@mail.csu.edu.cn