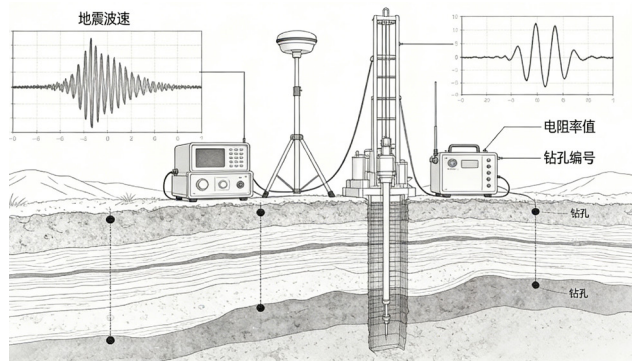


大数据驱动下岩土物探数据处理精度提升研究

文 | 蒋佳俊 张宗军

岩土物探数据处理精度直接影响地质勘查准确性，当前精度影响因素包括地质环境复杂性导致的信号干扰、噪声抑制能力不足、多源数据融合效果不佳等问题。大数据技术凭借强大的计算能力和智能化算法，为岩土物探精度提升开辟了新路径，能够有效打破传统方法的技术瓶颈。



(配图由 AI 生成)

岩土物探数据处理精度影响因素与大数据技术优势

岩土物探数据处理精度受岩土介质复杂物理特性强烈制约，成为影响地质勘查质量的核心因素，岩土体强非均质性导致物探信号传播过程中产生严重畸变，地层界面电阻率差异可达 2~3 个数量级，引起信号幅度衰减超过 40dB，深层地质信息提取面临巨大挑战。岩土工程中软弱夹层、破碎带等不良地质体厚度往往小于 0.5m，要求物探数据处理具备极高分辨率精度。传统处理方法在应对岩土介质频散效应、多重反射干扰时存在明显不足，当信噪比低于 5:1 时，有效信号识别率急剧下降，多种物探方法获取的数据在时空采样密度、物理响应机制方面存在显著差异，数据融合过程中信息损失严重。大数据技术通过海量数据挖掘能够识别岩土介质中微弱异常特征，机器学习算法可以建立复杂非线性映射关系，突破传统线性处理方法的精度瓶颈，分布式计算架构实现多维度数据协同处理，为岩土物探精度根本性提升奠定坚实基础。

大数据驱动的岩土物探数据处理精度提升关键技术

海量岩土物探数据的分布式存储与并行计算框架

岩土物探数据存储采用地质单元导向的分布式架构，按照土层与基岩及断层等地质类型建立独立存储集群，数据分片策略遵循地质空间连续性原则，单个分片容量控制在 500GB 以内，副本数量设置为 3，确保数据安全。存储介质分级配置，将活跃勘查区域数据部署于 NVMe SSD 获得微秒级访问延迟，历史数据迁移至 SATA HDD 实现成本优化，元数据索引采用 B+ 树结构记录地理坐标、地质年代、岩性编码等属性，索引深度控制在 4 层以内，数据压缩引入地质先验知识，利用岩土分层垂直相关性设计差分压缩算法，压缩

比达到 12:1，且保持地质界面信息完整性。计算任务调度基于地质剖面完整性原则，单个作业包含完整地质横截面，避免跨节点数据依赖，内存管理实施预测性缓存机制，根据勘查路径和地质走向预加载相关数据块，缓存容量配置为 128GB，支持大规模矩阵运算。负载均衡算法考虑不同物探方法计算密集度差异，CPU 密集型任务优先调度至高主频节点，内存密集型任务分配至大容量节点。

多源异构物探数据融合与智能预处理技术

多源数据统一采用岩土物理参数约束的配准算法，建立电阻率、密度、泊松比的三元关系模型：

$$\rho = a \cdot \sigma^b \cdot e^{c\nu} + d$$

式中， ρ 为电阻率 ($\Omega \cdot m$)； σ 为密度 (g/cm^3)； ν 为泊松比 (无量纲)； a 、 b 、 c 、 d 为岩土类型相关的拟合参数，其中 a 的量纲为 $\Omega \cdot m \cdot (g/cm^3)^{-b}$ ， d 的量纲为 $\Omega \cdot m$ 。空间配准精度控制在厘米级别，时间同步误差限制在毫秒范围内，数据质量评估建立包含信噪比、动态范围、线性度等技术指标的综合评价体系，异构数据格式转换采用 UK00A 标准，支持 SEG-Y、LAS、XYZ 等格式无损转换。预处理流水线集成去噪、校正及标准化 3 个核心模块，自适应中值滤波窗口动态调整范围为 3×3 至 15×15 像素，校正模块补偿温度漂移系数 $0.002/^\circ C$ 和湿度影响系数 $0.001/\%RH$ 。特征提取基于独立成分分析技术分离混合信号地质成分，成分数量根据累计贡献率 95% 原则确定，数据融合权重通过贝叶斯推理更新各数据源可信度分布，输出标准化多维地质特征向量。

自适应噪声抑制与信号重构优化算法

噪声模型构建涵盖工业干扰、自然电磁场、仪器热噪声三类主要噪声源，工频干扰频率锁定在 $50Hz \pm 0.5Hz$ ，自

然电磁场频率范围设定为 0.01 ~ 100Hz，滤波器设计采用级联结构，一级陷波器抑制工频及其谐波分量，品质因子 Q 值设置为 100，获得窄带抑制特性，二级带通滤波器提取有效信号频段。信号分解基于改进的经验模态分解算法，分解层数限制在 8 层以内，避免过度分解，停止准则设定为残余信号能量占比小于 5%，稀疏表示字典学习采用 K-SVD 算法训练岩土特征原子库，字典规模设置为 256×512，稀疏度约束参数 K 控制在 5 ~ 20。重构算法采用正交匹配追踪方

$$\min_{\alpha} \|y - D \cdot \alpha\|_2^2 + \lambda \|\alpha\|_1$$

法求解稀疏系数，其优化目标函数为：

式中，y 为观测信号向量；D 为字典矩阵 (256×512)； α 为稀疏系数向量； λ 为正则化参数； $\|\cdot\|_2$ 表示 L2 范数； $\|\cdot\|_1$ 表示 L1 范数。迭代收敛条件设定为残差能量下降至初始值 1% 以下，信号质量评估引入地质解释一致性指标，拟合误差容忍度设定为 ±10%，自适应参数调节基于粒子群优化算法，种群规模设置为 50，最大迭代次数限制为 200 代，确保收敛性。

基于深度学习的岩土介质特征智能识别方法

卷积神经网络架构针对岩土数据空间分辨率特征优化设计，网络深度设定为 18 层，包含 5 个残差块，每个残差块包含 3 个卷积层，输入层设计为多通道结构接收电阻率、自然电位、激发极化、磁化率等物探参数，输入尺寸标准化为 512×512×4 维张量。卷积核采用多尺度设计策略，浅层使用 3×3 小核提取局部细节，深层采用 7×7 大核捕获全局地质背景，步长统一设置为 1，损失函数设计融合分类损失和边界损失，分类损失采用加权交叉熵处理样本不平衡，边界损失通过 Sobel 算子强化地质界面识别，权重比例设定为 7 : 3。数据增强策略包含 ±30° 旋转变换模拟不同勘查方向，0.8 ~ 1.2 倍缩放变换适应地质尺度变化，±15% 幅值扰动模拟观测误差影响，网络正则化采用 Dropout 技术，丢弃概率从 0.1 线性增长至 0.5，批归一化层动量参数设置为 0.9。模型集成采用 5 折交叉验证训练多个子模型，预测结果通过加权投票融合，权重根据验证集性能动态分配，输出岩土介质类别概率分布和空间分布图。

精度提升效果评价与验证

典型地质条件下工程应用案例分析

选择华北平原软土地区作为典型验证场地，该区域软土层厚度 10 ~ 25m，含水量 40% ~ 80%，代表平原地区复杂地质条件，验证区域位于某高速铁路工程段，工程建设对地质勘查精度要求极高。建立传统物探方法与大数据驱动方法的对比验证体系，采用相同测量设备确保数据可比性，布设

高密度电阻率测线 12 条，总长度 3.6km，测点间距 2m，瞬变电磁测点 36 个，同步开展钻探验证，布设验证钻孔 18 个，深度 30 ~ 40m，获取完整地质剖面 and 土样。建立统一的数据质量控制标准，验证方案采用双盲原则，分别运用两种处理方法对相同原始数据进行处理，通过钻孔揭露的地质情况作为真值基准，评估两种方法的精度差异，为大数据技术应用效果提供客观验证依据。

精度提升效果定量评估与对比验证

基于华北平原软土地区验证数据，采用钻孔揭露的地质界面和室内试验获得的物性参数作为真值基准，对两种处理方法的精度进行定量评估。

表 1 验证结果表明大数据驱动方法显著提升了岩土物探数据处理精度，地质界面识别精度从 ±0.8m 改善至 ±0.3m，异常体定位精度提升了 60%，电阻率参数反演误差降低了 42.8%。处理效率方面，大数据方法将原本需要 48.5 小时的数据处理时间缩短至 12.3 小时，效率提升了 74.6%，满足工程快速决策需求。

表 1 华北平原软土地区精度提升效果对比

评价指标	传统方法	大数据方法	提升幅度
地质界面识别精度 (m)	±0.8	±0.3	62.5%
异常体定位精度 (m)	±1.5	±0.6	60.0%
电阻率反演误差 (%)	±15.2	±8.7	42.8%
数据处理时间 (h)	48.5	12.3	74.6%
信噪比改善 (dB)	3.2	8.9	178.1%

来源：济南市勘察测绘研究院

结束语

大数据驱动的岩土物探数据处理方法在精度提升方面取得了显著成效，通过构建完整的技术体系，有效解决了海量数据处理、多源信息融合、智能特征识别等关键技术难题。工程验证结果显示，该方法相比传统技术精度提升明显，为复杂地质环境下的岩土工程勘查提供了可靠技术保障。随着算法优化和应用推广，大数据技术将进一步推动岩土物探向智能化、精准化方向发展，为地质勘查和工程建设提供更高质量的技术支持。

作者简介：蒋佳俊 张宗军 济南市勘察测绘研究院

责任编辑：杨佳宇 投稿邮箱：zhouhl@staff.ccidnet.com