

面向煤田勘探的多源地质数据AI融合处理与异常体识别研究

张建军

陕西省煤田地质集团有限公司 陕西 西安 710021

摘要: 煤田勘探正从传统单一方法向多技术协同、数据驱动模式转变,多源地质数据的综合解释与异常体智能识别成为提高勘探精度的关键。本文针对煤田勘探中地震、测井、电磁、钻探及遥感等多源数据存在的异构性、多尺度及不确定性问题,系统研究了基于人工智能的多源数据融合处理框架与异常体识别方法。构建了多级数据融合体系,探讨了深度学习在数据配准、特征提取及联合反演中的应用,提出了基于语义分割与目标检测的煤层、断层、陷落柱及采空区等异常体的智能识别模型。研究成果对于提升煤田勘探自动化水平、降低解释不确定性、保障煤矿安全高效开采具有重要的理论价值与工程意义。

关键词: 煤田勘探;多源地质数据;数据融合;人工智能;异常体识别

引言:煤炭作为我国主体能源,其安全高效开采依赖于精准的地质勘探成果。煤田勘探涉及地震、测井、电磁、钻探及遥感等多种技术手段,获取的数据具有多源、异构、多尺度的特点。传统人工解释依赖专家经验,效率低且主观性强,难以充分挖掘多源数据蕴含的丰富信息。人工智能技术特别是深度学习的快速发展,为多源地质数据融合处理与异常体识别提供了新的技术路径。开展面向煤田勘探的多源地质数据AI融合处理与异常体识别研究,对于提升勘探精度、降低解释不确定性、保障煤矿安全生产具有重要的理论价值与现实意义。

1 煤田勘探多源数据类型与特征分析

1.1 煤田地质勘探的主要数据源

煤田地质勘探形成了以地球物理勘探为核心、钻探验证为手段、遥感技术为辅助的多技术体系。地震勘探是主力方法,三维地震数据能够精细刻画煤层形态与构造展布,其振幅、频率、相干性等属性蕴含着丰富地质信息。测井数据包括自然伽马、电阻率、声波时差等常规曲线及成像测井资料,直接反映煤层及围岩物性特征,垂向分辨率高。电磁法勘探如瞬变电磁、可控源音频大地电磁等,对低阻体敏感,在采空区积水探测方面具有独特优势。钻探数据提供真实的煤层见煤点坐标、厚度及化验结果,是其他物探方法的验证标尺。遥感数据从宏观角度揭示区域构造格架与地表地质信息。各类数据源从不同侧面反映地质体特征,共同构成煤田勘探的多源数据基础。

1.2 多源数据的异构性与互补性

多源地质数据的异构性体现在数据格式、空间尺

度、物理属性等多个维度。地震数据以SEG-Y格式存储,测井数据为LAS格式,钻探数据为文本表格,格式差异增加了综合解释难度。空间尺度方面,地震勘探横向覆盖广但垂向分辨率有限,测井数据垂向分辨率高但仅反映井孔周边信息,遥感数据宏观性强但难以探测深部地质体,多尺度特征使数据融合面临挑战。然而,异构性背后隐藏着深刻的互补性。地震数据反映构造格架,测井数据标定岩性物性,电磁数据对流体敏感,钻探数据提供真实验证,遥感数据揭示区域背景^[1]。单一方法各有局限,多源数据联合解释能够相互印证、优势互补,有效降低多解性,提高地质解释可靠性。

1.3 煤田主要异常体类型及其地质特征

煤田勘探关注的主要异常体包括煤层、断层、陷落柱、采空区及岩浆岩侵入体等,各类异常体具有独特的地球物理响应特征。煤层在地震剖面上表现为强振幅、低频率、连续性反射同相轴,测井曲线上呈现“两高三低”特征。断层在地震数据中表现为同相轴错断、产状突变,相干体属性上呈现线性低相干带。陷落柱是煤田特有构造现象,地震反射紊乱、同相轴中断,电磁法呈现局部低阻异常,测井曲线出现畸变。采空区因煤层采出后顶板垮塌,地震反射杂乱,瞬变电磁对采空区积水尤为敏感,呈现明显低阻异常。岩浆岩侵入体使煤层结构破坏,地震反射出现不规则强振幅异常体。准确识别这些异常体及其空间分布,是煤田勘探的核心任务。

2 多源地质数据预处理与一致性处理

2.1 数据质量控制

多源地质数据原始质量影响融合与识别效果,要建

立严格质量控制流程。地震数据需评价原始单炮记录、识别干扰波、分析信噪比，严重噪声记录要重采或剔除。测井数据要深度匹配、环境校正、处理异常值，校正井径坍塌等影响，确保曲线反映地层信息。钻探数据要核对见煤点坐标等，剔除错误记录。电磁数据要检查原始曲线形态等，排除人文干扰。对缺失值，可用克里金插值等方法补全，保证数据完整。质量控制要详细记录，为后续融合处理提供追溯依据。

2.2 空间配准与坐标统一

多源数据空间参考框架不同，空间配准是融合基础。先统一至相同地理坐标系统，投影方式一致。地震数据以炮点等定位，构建三维网格，将时间域剖面转深度域。测井数据以井口坐标等定位，做井斜校正，将斜深转垂深。钻孔数据可直接作控制点。遥感数据几何精校正后可直接用。对多源数据空间偏差，用特征点匹配法，以钻孔见煤点为基准，校准地震解释层位，确保数据在同一空间框架严格对应，为融合打基础^[2]。

2.3 数据归一化与标准化

不同物探方法数据物理量纲和数值范围不同，直接融合会致量纲主导问题，要归一化与标准化。归一化将数据映射到[0,1]或[-1,1]区间，极差归一化适用于数值分布均匀的数据。标准化将数据转为均值为0、标准差为1的分布，适用于正态分布数据，能消除量纲影响且保留分布形态。地震多属性要分别计算统计特征后标准化。测井曲线可分层段标准化，保留岩性区分度。归一化参数基于全区统计确定并保存，归一化后数据转至统一度量空间，为特征融合创造条件。

2.4 多尺度数据融合的尺度匹配

多源地质数据具有显著的多尺度特征，地震勘探覆盖数十平方公里但垂向分辨率米，测井数据垂向分辨率厘米级但仅反映井孔周围，实现多尺度数据的有效融合必须解决尺度匹配问题。小波变换是多尺度分析的有效工具，通过小波分解将不同尺度的数据分解为不同频带的分量，实现尺度统一。对于测井数据的高分辨率信息，可向上采样至地震采样率，保留其高频细节作为约束条件。对于地震数据的横向连续性，可向下采样至与测井匹配的网格密度。金字塔融合策略将数据构建为多分辨率金字塔，由粗到精逐层融合，既保留宏观构造形态，又融入精细细节信息。多尺度特征提取则利用卷积神经网络不同层级的感受野，自动学习跨尺度的特征表达，实现从局部细节到整体结构的全面刻画。

3 基于 AI 的多源地质数据融合方法

3.1 数据融合层次划分

多源地质数据融合分三个层次。数据层融合最基础，在原始采样点级别直接组合多源数据，形成多通道输入数据体，适用于空间配准精度高、物理意义相近的数据，如地震多属性叠合等。特征层融合是深度学习主流，先从各数据源提取高维特征，再拼接或加权组合特征向量，输入网络学习，能挖掘非线性关联。决策层融合是对各数据源独立处理后的初步结果，用投票等方法综合决策，适用于异质性强的数据组合。实际应用常采用混合融合策略，在不同环节用不同层次方法，发挥数据互补优势。

3.2 基于传统机器学习的融合方法

传统机器学习方法在多源数据融合中仍有重要价值，适用于样本量有限、需可解释性的场景。主成分分析通过线性变换投影数据至主成分空间，实现降维融合，常用于地震多属性融合。独立成分分析假设数据由独立源信号混合而成，可分离不同物理来源信号成分。支持向量机用核函数映射数据，找最优分类超平面，能融合多源特征分类识别。贝叶斯融合方法将数据源不确定性建模，用贝叶斯公式计算后验概率，显式表达融合结果不确定性。这些方法计算效率高、可解释性强，可与深度学习方法结合。

3.3 基于深度学习的特征级融合

深度学习自动特征提取能力强，是多源地质数据特征级融合主流方法。卷积神经网络多通道输入设计，让地震属性等作为不同通道输入，卷积核自动学习跨通道联合特征。针对地震与测井尺度差异，可用双分支网络结构，深层特征空间融合^[3]。自编码器及其变体通过编码-解码结构学习低维流形表示，多源数据共同输入，强迫学习联合特征表达，重构误差用于异常检测。图神经网络将地质单元建模为图节点，通过消息传递实现井间信息传播融合，利用地质体拓扑关系。这些方法能自动发现复杂特征组合，提升融合效果。

3.4 联合反演中的AI融合

地球物理反演是从观测数据推测地下物性参数，联合反演同步拟合多种数据获一致地质模型。传统联合反演有计算复杂、正则化参数选择难等问题，AI方法提供新思路。深度学习代理模型用大量正演模拟数据训练神经网络，学习非线性映射，替代正演计算提升反演效率。基于物理信息神经网络的反演方法将物理方程作为损失函数一部分，满足物理规律约束，提高结果合理性。生成对抗网络可生成符合地质特征的地下模型作先验约束。AI驱动的反演能融合多类数据，获符合多种物理场观测的煤田地质模型，为异常体识别提供可靠

物性基础。

4 多源数据融合与异常识别一体化平台构建

4.1 平台总体架构

面向煤田勘探的多源数据融合与异常识别平台采用分层架构设计,自下而上分为数据管理层、预处理层、融合处理层、异常识别层与可视化输出层五个层次。数据管理层负责多源数据的接入、存储与管理,支持SEG-Y、LAS、SHP等主流格式,建立统一的数据索引与元数据库,实现高效检索与访问。预处理层集成质量控制、空间配准、归一化标准化、尺度匹配等功能模块,对原始数据进行一致性处理,生成可供融合的基础数据体。融合处理层封装传统机器学习与深度学习融合算法库,支持特征级融合、决策级融合及联合反演等多种融合模式。异常识别层部署基于目标检测与语义分割的深度学习模型,实现断层、陷落柱、采空区等异常体的自动识别与边界刻画。可视化输出层提供三维交互展示、成果图件自动生成、报告输出等功能。平台采用微服务架构,各模块独立部署、松耦合集成,便于功能扩展与算法更新。

4.2 关键技术模块设计

平台核心模块包括数据统一访问接口、融合算法集成框架、模型训练推理引擎与三维可视化引擎。数据统一访问接口屏蔽底层数据格式差异,提供统一的数据读取接口,支持按空间范围、数据类型、时间等多条件查询与数据抽取。融合算法集成框架采用插件化设计,支持PCA、ICA等传统算法与CNN、自编码器、图神经网络等深度学习算法的灵活插拔,提供算法对比评估工具。模型训练推理引擎基于TensorFlow或PyTorch开发,支持GPU加速分布式训练,内置样本增强、迁移学习、模型压缩等功能,降低模型应用门槛。三维可视化引擎基于OpenGL或VTK开发,支持地震数据体立体显示、测井曲线空间展布、异常体三维雕刻与半透明渲染,实现多源数据与识别结果的可视化融合,帮助地质人员直观理解地下地质结构^[4]。

4.3 人机交互与解释辅助

平台注重人机协同,充分发挥AI高效处理与专家经验判断的各自优势。交互式标注工具提供便捷的样本标注功能,支持在地震剖面、属性切片上快速勾画断层线、圈定异常体边界,构建高质量训练样本库。识别结果的交互修正允许地质人员对AI识别结果进行编辑调整,修改后的结果可作为新样本反馈至模型进行增量学习,实现模型的持续优化。多方案对比功能支持不同融合算法、不同参数设置下的识别结果并列显示,辅助解释人员综合判断。不确定性可视化采用透明度或颜色饱和度和度表达模型预测置信度,引导地质人员重点关注不确定区域。解释知识库记录典型异常体的特征模式与识别经验,在解释过程中提供相似案例推荐,辅助新手解释人员快速上手。平台以辅助而非替代为理念,提升解释效率的同时保证地质解释的可靠性与专业性。

结束语

本文围绕煤田勘探多源地质数据AI融合处理与异常体识别开展系统研究,分析了地震、测井、电磁、钻探等多源数据的异构性与互补性,建立了从质量控制、空间配准到尺度匹配的预处理流程,探讨了传统机器学习与深度学习方法在数据融合中的应用,构建了集数据管理、融合处理、异常识别与可视化于一体的平台架构。未来,随着多模态大模型、物理信息神经网络等技术的发展,多源地质数据融合与异常识别将朝着更精准、更智能、更可解释的方向迈进,为煤炭资源安全高效开发提供更强有力的技术支撑。结束语

参考文献

- [1]陈碧华.地质勘查技术在煤田勘探过程中的应用[J].内蒙古煤炭经济,2023(6):178-180.
- [2]李丽.多源数据融合技术及其在地质矿产勘查中的实践研究[J].中国建设信息化,2022(10):20-25.
- [3]伍超奇.人工智能在矿山地质中的应用现状及未来展望[J].世界有色金属,2024,(02):49-51.
- [4]刘恒,赵威,韩媛,等.人工智能与大数据分析在新一轮找矿突破战略行动中的应用与挑战[J].科技创新与应用,2024,14(02):20-23.