

赵玄,严家斌,皇祥宇,等. 基于K中心点聚类分析的大地电磁阻抗识别[J]. 地球科学与环境学报, 2018, 40(6): 779-786.

ZHAO Xuan, YAN Jia-bin, HUANG Xiang-yu, et al. Magnetotelluric Impedance Recognition Based on K-medoids Clustering Analysis [J]. Journal of Earth Sciences and Environment, 2018, 40(6): 779-786.

基于K中心点聚类分析的大地电磁阻抗识别

赵玄¹, 严家斌^{1,2*}, 皇祥宇^{1,2}, 胡涛¹

(1. 中南大学地球科学与信息物理学院, 湖南长沙 410083; 2. 中南大学有色金属资源与地质灾害探查
湖南省重点实验室, 湖南长沙 410083)

摘要:在大地电磁场复杂的观测环境中, 信号会受到不同类型或程度的噪声干扰。传统阻抗估计或提取方法(如功率谱法、Robust法等)难以获得稳定的阻抗值。利用大地电磁阻抗的实虚分量特性, 受噪声干扰小的信号阻抗分布集中, 受噪声干扰严重的信号阻抗分布散乱。引入K中心点聚类分析对阻抗进行提取与识别, 并利用阻抗欧氏距离来描述阻抗间的相似性, 依据相似性把受干扰小的信号阻抗划分到一类, 受干扰大的信号阻抗划分到不同的类。依据相干度准则和紧凑性准则等类的选取准则, 甄别出干扰环境中阻抗所在的最佳类。通过仿真实验和实例分析, 验证了K中心点聚类分析能在噪声环境中识别出高质量的信号, 恢复出真实阻抗值。

关键词:阻抗估计; 实虚分量; 聚类分析; K中心点; 阻抗欧氏距离; 相干度; 紧凑性; 数值模拟

中图分类号: P631.3⁺25

文献标志码: A

文章编号: 1672-6561(2018)06-0779-08

Magnetotelluric Impedance Recognition Based on K-medoids Clustering Analysis

ZHAO Xuan¹, YAN Jia-bin^{1,2*}, HUANG Xiang-yu^{1,2}, HU Tao¹

(1. School of Geosciences and Info-physics, Central South University, Changsha 410083, Hunan, China;

2. Key Laboratory of Non-ferrous Resources and Geological Hazard Detection of Hunan Province,
Central South University, Changsha 410083, Hunan, China)

Abstract: In the complicated observation environment, the magnetotelluric signal is disturbed by different types or degrees of noise. Traditional estimation and extraction methods are hard to obtain stable impedance, such as the power-spectrum method and Robust method. Based on the physical properties of the real and imaginary components of magnetotelluric impedance, the impedance of signal disturbed slightly by noise distributes intensively, and the impedance of signal disturbed severely by noise distributes dispersedly. So, the impedance extraction and identification based on K-medoids clustering analysis was introduced, and then the euclidean distance for impedance was proposed to describe the similarity of impedance. The K-medoids clustering analysis can divide the impedance of signal disturbed slightly by noise into the same cluster and other impedance of signal disturbed severely by noise into different clusters based on the similarity. According to the selection criteria of cluster including coherence and compactedness criteria, the optimal cluster of impedance in the disturbed environment was

收稿日期: 2018-03-16; 修回日期: 2018-07-11 投稿网址: <http://jese.chd.edu.cn/>

基金项目: 国家自然科学基金项目(40874055); 湖南省自然科学基金项目(14JJ2012)

作者简介: 赵玄(1992-), 男, 安徽阜阳人, 工学硕士研究生, E-mail: 2625476614@qq.com。

* 通讯作者: 严家斌(1969-), 男, 湖南常德人, 教授, 工学博士, E-mail: cspyy@csu.edu.cn。

chosen. Based on the simulation experiment and instance analysis, the K-medoids clustering analysis can recognize high quality signal in noisy environment and recovery the real impedance.

Key words: impedance estimation; real and imaginary components; clustering analysis; K-medoids; euclidean distance for impedance; coherence; compactedness; numerical simulation

0 引言

大地电磁法(MT)诞生于 20 世纪 50 年代初,该方法具有工作效率高、成本低廉、勘探深度大、高阻层的屏蔽作用小等特点^[1-2],广泛应用于矿产勘查、地下水与地热勘探、油气普查等方面,产生了巨大的经济效益^[3-5]。随着城市的扩展和人类工业活动的加剧,大地电磁场信号的观测越来越困难,干扰越来越严重,从受干扰的观测数据中估计稳定的阻抗或视电阻率是当前地球物理电磁信号处理的难点之一。Sims 等较早提出依据最小二乘法的张量阻抗估计,该方法要求各数据道之间的噪声相互独立,但当电磁资料存在相关噪声以及观测误差不服从高斯分布规律时,由最小二乘法估算的阻抗会产生严重偏差^[6]。为了克服上述问题,Egbert 等将稳健性估计(Robust Estimation)引入到地磁转换函数中,并详细介绍了加权最小二乘法和改进算法^[7]。为了进一步进行稳健性估计,Chave 等详细介绍了地球物理中几个重要的物理量:功率谱、相干度和转换函数,并综述了 Robust 估计理论,从统计学原理出发梳理了 M 估计和极大似然估计法^[8]。针对海洋环境下的大地电磁法,柳建新等提出了基于相关归一 Robust 法,根据相关系数的变化改进 Robust 法的权系数^[9]。张刚等将基于重复中位数估计的 Robust 法应用于长周期大地电磁阻抗张量估算中^[10]。Chave 等提出了基于 Stable 分布的最大似然阻抗估计法,根据数据自身的噪声分布进行加权,能更合理地给出阻抗估计值^[11-12]。常规的阻抗估计法对噪声辨识度很低,未利用大地电磁阻抗的物理特性分辨高质量数据,导致噪声严重干扰信号数据,使得阻抗结果发生严重偏差。由于大地电磁数据中高质量信号数据不能有效利用,所以聚类分析算法被引入电磁数据处理中。李晋等介绍了基于递归分析和聚类分析的大地电磁信噪辨识及分离方法^[13],该方法虽然保留了大地电磁信号低频段的缓慢变化信息,改善了低频段大地电磁数据质量,但对电场和磁场的处理是独立进行的,对电磁场的分量间相关性考虑不足。

本文从大地电磁阻抗的实虚分量特性出发,通过定义阻抗的欧式距离构建阻抗相似性度量,利用其相

似性对受干扰阻抗、轻微受干扰阻抗、未受干扰阻抗进行 K 中心点聚类分析,识别出高质量的信号,并基于仿真实验和实例分析验证该方法的有效性。

1 原理

聚类分析(Clustering Analysis)是数据挖掘的主要方法之一^[14-17],主要研究数据之间的相关性,这种相关性大小是通过一定的相似性准则来判断的,然后依此将对象分簇或聚类。在频率域内针对某一地质构造,高质量的大地电磁信号数据的阻抗实虚分量必然在真实阻抗附近浮动,其阻抗间存在相似性,而与噪声数据阻抗间存在差异性^[18]。这一特性十分适合采用聚类分析进行阻抗数据的优选,筛选出高质量的数据。采用聚类分析的结果使得信号数据尽可能归为一类,信号数据和噪声归为不同类。

1.1 相似性度量

一个聚类分析过程的质量取决于对度量标准的选择。为了度量数据对象之间的相似或者接近程度,需要定义一些相似性度量标准。聚类分析中,为了表示两个数据对象之间的相似度,一般采用特征空间中的距离作为度量标准来计算两个样本间的相异度^[16-17]。为了描述阻抗数据之间的相似或者接近程度,提出了阻抗(Z)的欧式距离。其表达式为

$$d_{ij}^z(f) = \sqrt{|\operatorname{Re} Z_i(f) - \operatorname{Re} Z_j(f)|^2 + |\operatorname{Im} Z_i(f) - \operatorname{Im} Z_j(f)|^2} \quad (1)$$

式中: $d_{ij}^z(f)$ 为同一频率 f 下第 i 段数据的阻抗 Z_i 和第 j 段数据的阻抗 Z_j 之间的欧式距离; $\operatorname{Re} Z_i(f)$ 、 $\operatorname{Im} Z_i(f)$ 分别表示阻抗 Z_i 的实部、虚部; $\operatorname{Re} Z_j(f)$ 、 $\operatorname{Im} Z_j(f)$ 分别表示阻抗 Z_j 的实部、虚部。

由式(1)可知,欧式距离越小表示高质量阻抗之间的相似或接近程度越大。相对于受噪声干扰的数据,受干扰较小的信号数据的阻抗欧式距离小。

1.2 K 中心点聚类分析

K 中心点(K-medoids)聚类分析是为了降低 K-means 算法^[19]对噪声孤立点的敏感度而提出的新算法。K 中心点聚类分析选择类或簇中最靠近均值中心点的一个对象来代表该类或簇的中心点,可以有效消除孤立点对聚类效果的影响^[20]。在聚类分析中只需要计算类或簇的中心点,并选择该类或簇内靠近中心点的对象作为新中心点,依次循环迭代,

直到得到稳定的类中心^[21]。

以频率域阻抗为例,对于包含 N 个阻抗数据对象的集合 X ,随机选取 K 个数据对象作为初始聚类中心点,并把 K 个初始聚类中心点表示成 K 个初始类。计算集合内所有数据对象与各个初始聚类中心点的阻抗欧式距离,依据就近分配原则,将数据对象划分到最靠近的类内,重新分配每个数据对象直至所有数据对象分配完毕,计算新的聚类中心点。 K 中心点聚类分析步骤为:①输入 $K, X=[\text{Re } Z, \text{Im } Z]$; ②选择 K 个初始聚类中心点, $C[1]=X[1], C[2]=X[2], \dots, C[K]=X[K]$; ③计算第 i 个类 $X[i]$ 与第 ②步中 K 个初始聚类中心点的阻抗欧式距离,找到离它最近中心点 $C[m]$,将其分配到 $C[m]$ 所在的类内,并对该点作标记;④计算出 m 个类内的所有数据对象均值,选择靠近均值的数据对象作为 m 个类的新中心点;⑤重复第 ③、④步,直至所有数据分配到类内以及达到设定迭代次数。

1.3 类的选取准则

经过 K 中心点聚类分析后,每个频点下的阻抗数据被划分为 K 个类。把每个类的中心点作为该类阻抗的代表值,每个频点下将会得到 K 个阻抗值和对应的视电阻率。从 K 个结果中挑选出最佳阻抗值,依据电磁法理论定义类的选取准则。

(1)相干度准则。相干度可以衡量输入信号和输出信号的相关性,相干度越大,数据相关性越好,信号质量越好。因此,设定相干度阈值,筛选出满足阈值对应的类。

(2)紧凑性准则。为了描述数据聚集程度,提出了紧凑性概念。以某一频点下的阻抗为例,对于包含 N 个阻抗数据对象的集合 $X=\{X_1, X_2, \dots, X_N\}$, $X=[\text{Re } Z, \text{Im } Z]$,假设数据集合 X 被划分为 K 个类,每个类中数据对象个数为 $\{n_1, n_2, \dots, n_k\}$,且 $n_1+n_2+\dots+n_k=N$ 。以第 j 类为例,该类的紧凑性 $C_j(f)$ 表达式为

$$D_j^z(f) = \sum_{i=1}^{n_j} \sum_{j=1}^{n_j} d_{ij}^z(f) \quad (2)$$

$$C_j(f) = \frac{D_j^z(f)}{|n_j|^2} \quad (3)$$

式中: $D_j^z(f)$ 为第 j 类的类内距离; n_j 为第 j 类中数据对象的个数, $j=1, 2, 3, \dots, k$; k 为类个数。

紧凑性 $C_j(f)$ 值越小,说明类内数据对象聚集的越紧凑,数据之间相似程度越高,更符合高质量信号数据的阻抗高相似性特征,紧凑性好的类被认为是高质量信号数据所在的类。满足相干度准则的条

件下,选择紧凑性好的类^[22]。

1.4 算法步骤

算法步骤为:①分别对 M 段时间域数据采用最小二乘法求取初始阻抗值,得到每个频点下 M 个阻抗值;②输入 $K, X=[\text{Re } Z, \text{Im } Z]$,对第 ①步中阻抗组成的集合 X 采用 K 中心点聚类分析,每个频点下的阻抗被划分到 K 个类内;③计算每个频点下各类的相干度和紧凑性;④依据相干度准则和紧凑性准则,筛选出符合要求的类以及相对应的阻抗值;⑤根据第 ④步中的阻抗值计算出相对应的视电阻率。

2 仿真实验

大地电磁资料中存在着多种噪声,使得数据处理异常困难,常规方法估算的结果会发生偏移^[23-24]。通过在大地电磁仿真数据中加入噪声,模拟实际情况下噪声对信号的影响,并对比基于 K 中心点聚类分析和 Robust 法的估算结果。首先利用蒙特卡罗(Monte Carlo)法产生互不相关的随机时间序列,作为电道或者磁道数据,通过傅里叶变换把时间域序列转化到频率域,根据层状介质的视电阻率理论公式,加入设定的电性模型参数,计算出磁道或者电道频率域数据,再通过傅里叶反变换转换到时间域。本文设定的电性模型参数为:地下结构为两层均匀介质,第一层介质电阻率为 $500 \Omega \cdot \text{m}$,厚度为 250 m ,第二层介质电阻率为 $100 \Omega \cdot \text{m}$ 。上述仿真实验产生了 x 方向电道分量 E_x 和 y 方向电道分量 E_y 各 50 段, x 方向磁道分量 H_x 和 y 方向磁道分量 H_y 各 50 段,各段数据采样长度为 2 048,采样频率为 10 000 Hz。在大地电磁仿真实验的电道或者磁道中加入类三角波噪声和脉冲噪声。

(1)随机选出 15 段磁道仿真数据,在选择的数据段不同位置加入类三角波噪声、脉冲噪声。信号与类三角波噪声幅值之比为 1:4,噪声数据长度占每段数据总长度的 10%;信号与脉冲噪声幅值之比为 1:50。对 50 段数据分别采用 Robust 法和 K 中心点聚类分析估计视电阻率和相位。由图 1(a)和图 2(a)可知,不含噪声的磁场频谱能量是逐渐增大的。在磁场中加入类三角波噪声和脉冲噪声[图 1(b)],这两种噪声影响所有频点的磁场频谱[图 2(b)],使得频谱能量明显高于原始信号能量。对上述 50 段数据先进行 K 中心点聚类分析,对分类后的阻抗按照相干度准则和紧凑性准则筛选出最佳阻抗值所在的类,计算最佳阻抗值所对应的视电阻率和相位。以 24.41 Hz 频点为例,阻抗数据被划分到

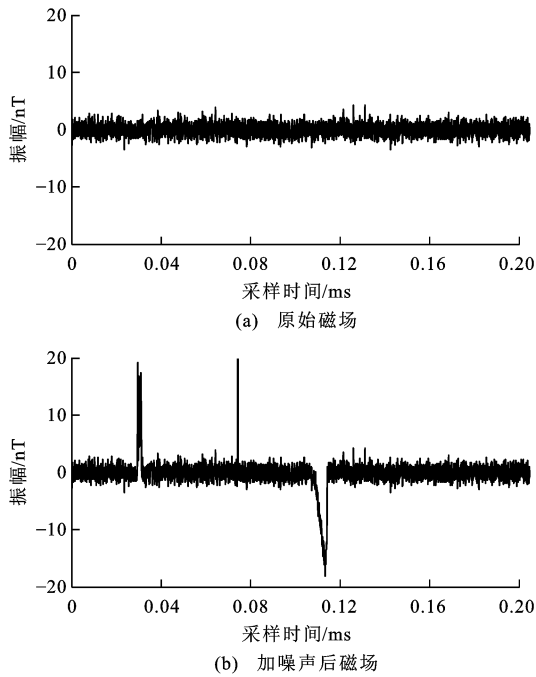


图 1 原始磁场和加噪声后磁场的时序序列
Fig. 1 Time Series of Magnetic Field with and Without Noise

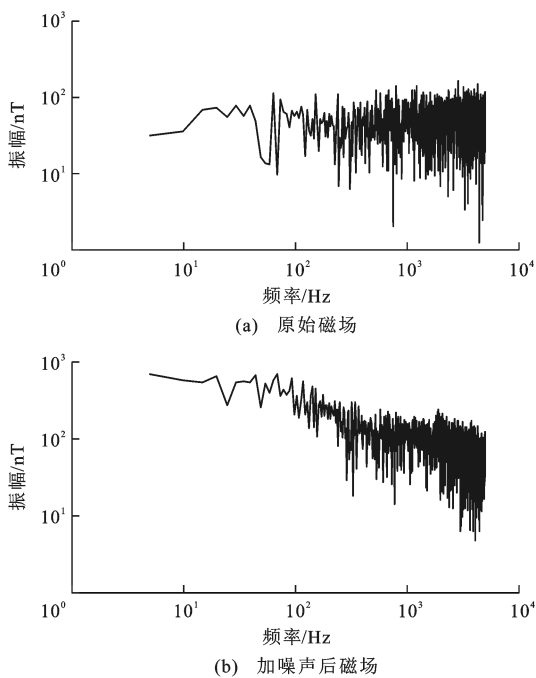


图 2 原始磁场和加噪声后磁场的频谱
Fig. 2 Frequency Spectra of Magnetic Field with and Without Noise

3 个区域内(图 3)。表 1 是该频点下划分成 3 个类对应的参数,由于第 3 类的相干度和紧凑性很好,所以选择第 3 类为最佳类,并计算最佳类所对应的视电阻率及相位。对其他频点采用相同计算方法,得到视电阻率曲线(图 4)和相位曲线(图 5)。从图 4、5 可以看到,基于 K 中心点聚类分析估算的结果更接近

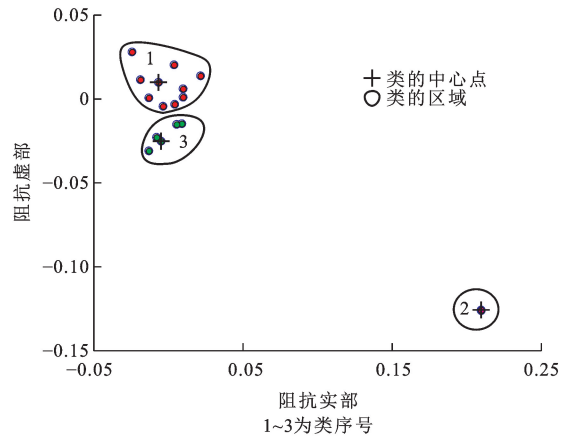


图 3 磁场噪声阻抗聚类图 ($f=24.41\text{ Hz}$)
Fig. 3 Clustering Diagram of Impedance of Magnetic Field with Noise ($f=24.41\text{ Hz}$)

理论曲线,当频率高于 1 000 Hz 时,基于 Robust 法估算的视电阻率小于理论值。相对于 Robust 法,基于 K 中心点聚类分析可以分离和识别出高质量数据对应的阻抗,削弱了噪声的干扰,使结果更为可靠。

表 1 基于 K 中心点聚类分析的磁场阻抗识别评价参数 ($f=24.41\text{ Hz}$)

Tab. 1 Evaluation Parameters of Impedance Recognition for Magnetic Field Based on K-medoids Clustering Analysis ($f=24.41\text{ Hz}$)

类序号	相干度	紧凑性	视电阻率/ $(\Omega \cdot \text{m})$	相位/ $(^\circ)$	点数
1	0.30	1.6×10^{-3}	0.5	54.2	10
2	0.08	1.8×10^{-3}	2.2	-78.6	5
3	0.99	8.8×10^{-7}	194.3	31.0	35

注:理论上,视电阻率为 $194.3\ \Omega \cdot \text{m}$,相位为 31.0° 。

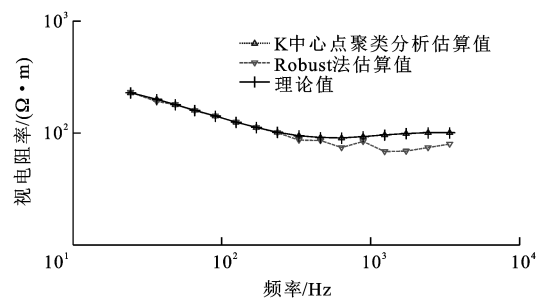


图 4 磁场噪声视电阻率曲线
Fig. 4 Apparent Resistivity Curves of Magnetic Field with Noise

(2)随机选出 10 段电道仿真数据,在选择数据段的不同位置加入类三角波噪声、脉冲噪声。信号与类三角波噪声幅值之比为 1 : 4,噪声数据长度占每段数据总长度的 10%;信号与脉冲噪声幅值之比为 1 : 50。对 50 段数据分别采用 Robust 法和 K 中心点聚类分析估计视电阻率和相位。由图 6(a)和

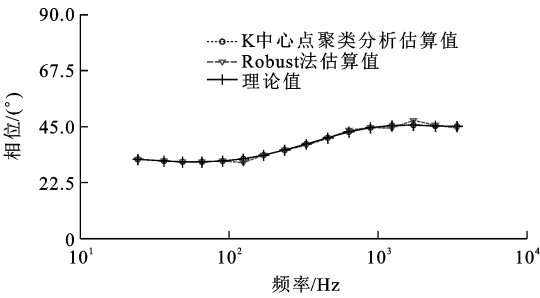


图 5 磁场噪声相位曲线

Fig. 5 Phase Curves of Magnetic Field with Noise

图 7(a)可知,不含噪声的电场频谱能量是逐渐增大的。在电场中加入类三角波噪声和脉冲噪声[图 6 (b)],这两种噪声影响所有频点的电场频谱[图 7 (b)],使得频谱能量明显高于原始信号能量。对上述 50 段数据进行 K 中心点聚类分析,对分类后的阻抗按照相干度准则和紧凑性准则,筛选出最佳阻抗值所在的类,并计算其对应的视电阻率和相位。以 43.95 Hz 频点为例,阻抗数据被划分到 4 个区域内(图 8)。表 2 是该频点下划分成 4 个类分别对应的参数,第 2 类的相干度和紧凑性很好,选择第 2 类作为最佳类,并计算其对应视电阻率和相位。对其他频点采用相同计算方法,得到视电阻率曲线(图 9)和相位曲线(图 10),基于 K 中心点聚类分析的估算结果更接近理论值,而基于 Robust 法估算的视电阻率曲线和相位曲线都出现了波动,结果偏离了理论值,由此说明基于 K 中心点聚类分析可以

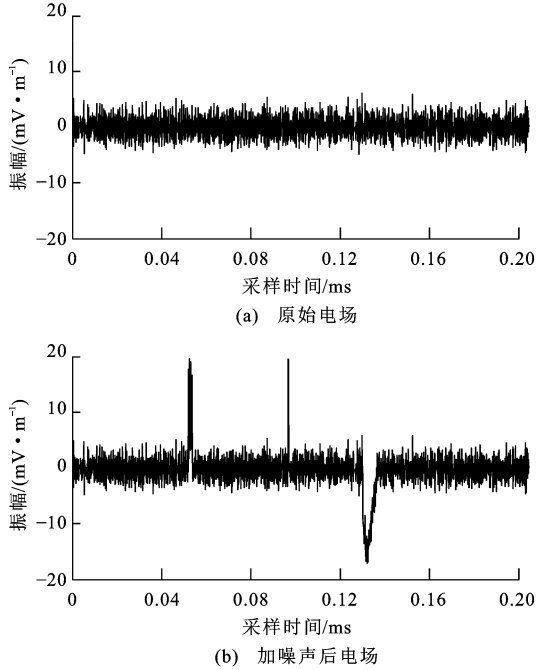


图 6 原始电场和加噪声后电场的时间序列

Fig. 6 Time Series of Electric Field with and Without Noise

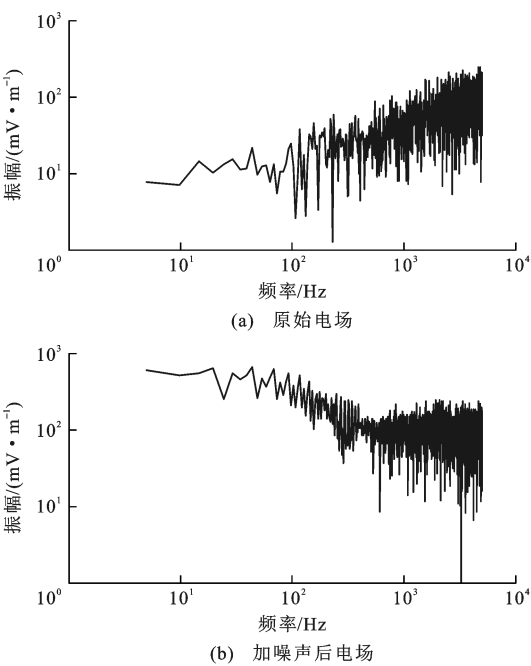


图 7 原始电场和加噪声后电场的频谱

Fig. 7 Frequency Spectra of Electric Field with and Without Noise

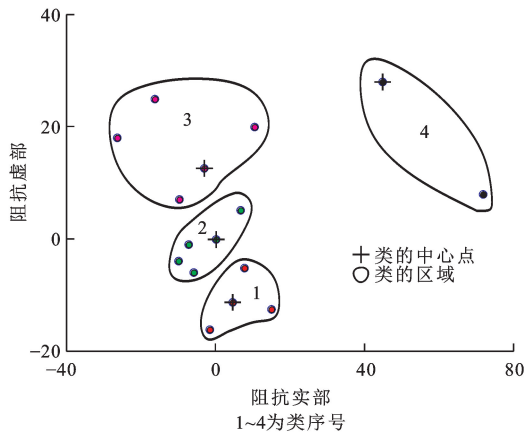


图 8 电场噪声阻抗聚类图($f=43.95\text{ Hz}$)

Fig. 8 Clustering Diagram of Impedance of Electric Field with Noise ($f=43.95\text{ Hz}$)

表 2 基于 K 中心点聚类分析的电场阻抗识别评价参数($f=43.95\text{ Hz}$)

Tab. 2 Evaluation Parameters of Impedance Recognition for Electric Field Based on K-medoids Clustering Analysis ($f=43.95\text{ Hz}$)

类序号	相干度	紧凑性	视电阻率/ $(\Omega\cdot\text{m})$	相位/ $(^\circ)$	点数
1	0.08	1.60	440 000.0	67.9	4
2	0.99	0.02	186.1	30.9	39
3	0.09	2.70	480 000.0	76.5	5
4	0.16	8.40	8 100 000.0	-32.0	2

注:理论上,视电阻率为 $186.1\ \Omega\cdot\text{m}$,相位为 30.7° 。

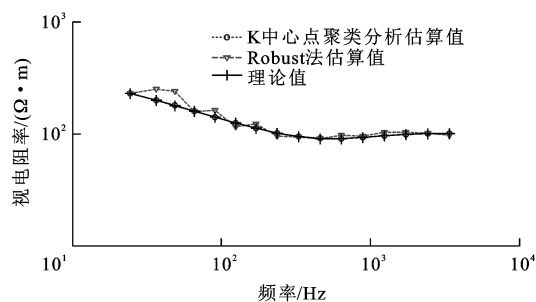


图 9 电场噪声视电阻率曲线

Fig. 9 Apparent Resistivity Curves of Electric Field with Noise

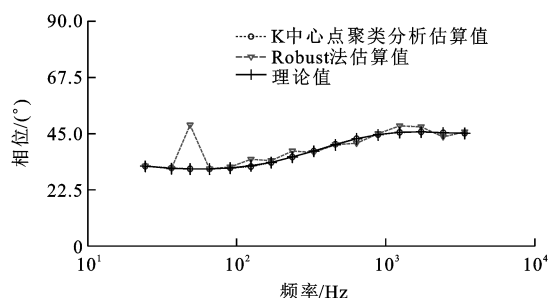


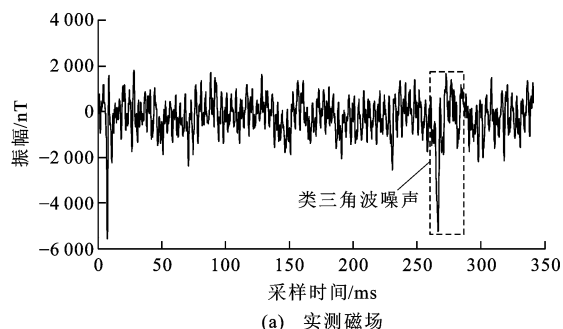
图 10 电场噪声相位曲线

Fig. 10 Phase Curves of Electric Field with Noise

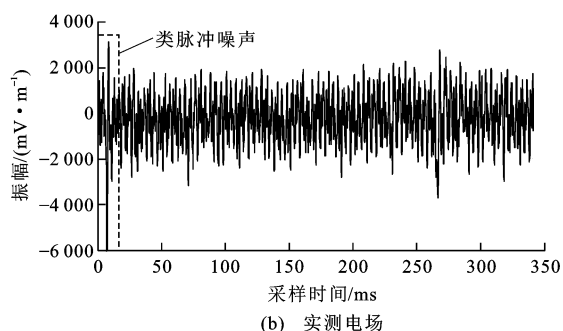
识别和筛选出高质量信号数据的阻抗,进而估算出更为可靠的结果。

3 实例分析

为验证基于 K 中心点聚类分析的实际应用效果,选取云南省昭通市牛栏江天花板水电站田坝村堆积体处电磁成像系统观测数据进行验证。观测区周围无高压线,但居民较多。图 11 是采集的原始时间序列数据,磁场在水平轴 260~280 ms 之间存在似类三角波噪声,电场在水平轴 0~20 ms 之间存在明显的似类三角波噪声和脉冲噪声。该测点电道和磁道信号中存在明显的类三角波噪声和脉冲噪声(图 11),同时还存在幅度较小的类方波等噪声。由仿真实验可知,不含噪声的电磁数据频谱能量逐渐增大,而测点的频谱能量逐渐减小,说明噪声影响了所有频点的频谱(图 12)。对测点数据采用和仿真实验一样的 K 中心点聚类分析,对阻抗数据进行聚类分析。以频点 82.5 Hz 为例,该频点下阻抗经过 K 中心点聚类分析,划分成 6 个区域(图 13),每个区域代表一个类。由图 13 可知,第 2、3、6 类内数据点比较分散,说明类内的阻抗受噪声影响很大,而第 1、4、5 类内的数据点分布集中,说明高质量数据比较多。由表 3 可知,第 1、4、5 类所对应的相干度和紧凑性很好,而第 5 类所对应的紧凑性最好,因此,



(a) 实测磁场



(b) 实测电场

图 11 实测磁场和实测电场时间序列

Fig. 11 Time Series of Measured Magnetic Field and Electric Field

选择第 5 类为最佳类,把其对应的视电阻率作为最终的视电阻率。对其他频点采用相同计算步骤,得到视电阻率曲线(图 14)。由图 14 可知,相对于基于 Robust 法的估算结果,基于 K 中心点聚类分析的估算结果更为光滑连续,不会出现异常跳动,更贴近真实地下结构,说明该方法可以识别出高质量数据的阻抗,降低了噪声干扰,使结果更为可信。

表 3 基于 K 中心点聚类分析的实测阻抗识别评价参数 ($f=82.5$ Hz)

Tab. 3 Evaluation Parameters of Measured Impedance Recognition Based on K-medoids Clustering Analysis ($f=82.5$ Hz)

类序号	相干度	紧凑性	视电阻率/($\Omega \cdot m$)
1	0.97	0.191 7	251.9
2	0.55	2.628 6	343.8
3	0.51	3.010 9	71.3
4	0.94	0.182 0	265.5
5	0.95	0.132 2	353.5
6	0.86	0.331 4	449.7

4 结 语

(1)从大地电磁阻抗的实虚分量特性出发,定义了阻抗欧氏距离,描述阻抗数据之间的相似性,提出了基于 K 中心点聚类分析的电磁信号识别及阻抗提取方法,并依据类的选取准则,识别出可靠的阻抗

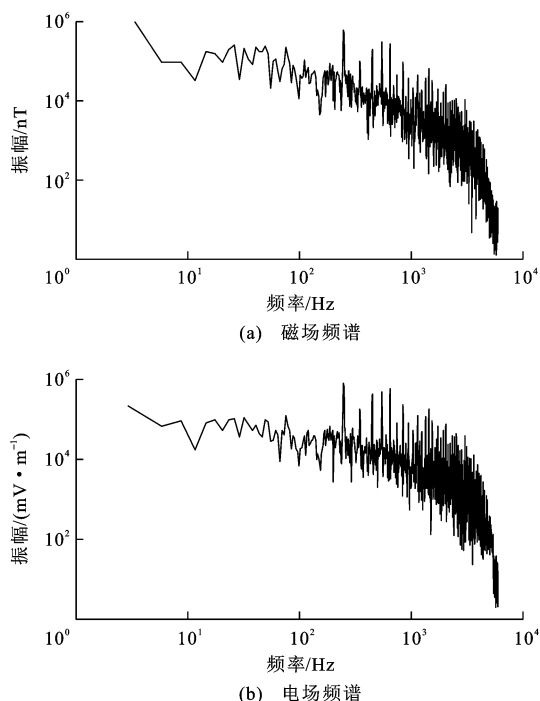


图 12 实测磁场和电场频谱

Fig. 12 Frequency Spectra of Measured Magnetic Field and Electric Field

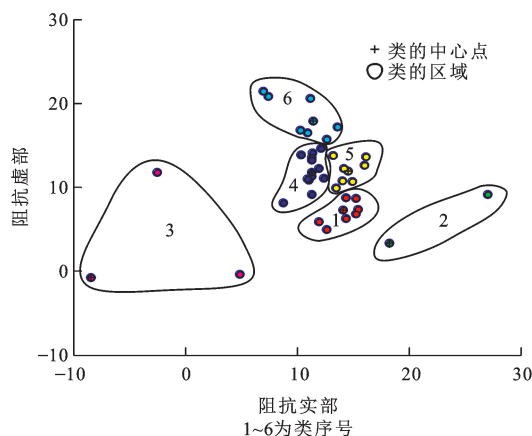
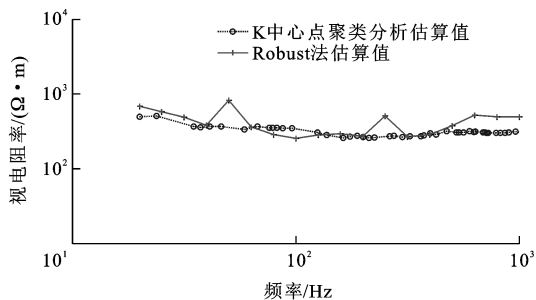
图 13 实测数据阻抗聚类图 ($f=82.5$ Hz)Fig. 13 Clustering Diagram of Impedance of Measured Data ($f=82.5$ Hz)

图 14 实测数据视电阻率曲线

Fig. 14 Apparent Resistivity Curves of Measured Data

值。通过仿真实验和实例分析,基于 K 中心点聚类分析可以识别并提取出高质量数据的阻抗,得到稳定阻抗值和视电阻率,提高估算的可靠性和稳定性。

(2)Robust 法更适用于磁场受干扰小时的情况,当电场或者磁场含有异常幅度大的噪声时,Robust 法估算的结果不再稳定,会使结果发生偏差;而 K 中心点聚类分析是利用阻抗的物理特性,高质量数据的阻抗会集中分布且具有高度的相似性,因此,受影响较小,得到的阻抗更合理有效。由于个别异常点的存在和初始聚类数目的选择都会影响聚类算法效率,所以本文的方法还需要进一步的优化。

参考文献:

References:

- [1] CAGNIARD L. Basic Theory of the Magneto-telluric Method of Geophysical Prospecting[J]. Geophysics, 1953,18(3):605-635.
- [2] 何继善. 可控源音频大地电磁法[M]. 长沙:中南工业大学出版社,1990.
HE Ji-shan. Controlled Source Audio-frequency Magneto-telluric[M]. Changsha: Central South University of Technology Press,1990.
- [3] 杨长福,徐世浙. 国外大地电磁研究现状[J]. 物探与化探,2005,29(3):243-247.
YANG Chang-fu, XU Shi-zhe. The Present Situation of Magnetotelluric Researches Abroad[J]. Geophysical and Geochemical Exploration, 2005,29(3):243-247.
- [4] KAUFMAN A A, KELER G V. 大地电磁测深法[M]. 刘国栋,译. 北京:地震出版社,1987.
KAUFMAN A A, KELER G V. Magnetotelluric Sounding Method[M]. Translated by LIU Guo-dong. Beijing: Seismological Press,1987.
- [5] 陈乐寿,王光愕. 大地电磁测深法[M]. 北京:地质出版社,1990.
CHEN Le-shou, WANG Guang-e. Magnetotelluric Sounding Method[M]. Beijing: Geological Publishing House, 1990.
- [6] SIMS W E, BOSTICK F X, SMITH H W. The Estimation of Magnetotelluric Impedance Tensor Elements from Measured Data[J]. Geophysics, 1971, 36(5):938-942.
- [7] EGBERT G D, BOOKER J R. Robust Estimation of Geomagnetic Transfer Functions[J]. Geophysical Journal International, 1986, 87(1):173-194.
- [8] CHAVE A D, THOMSON D J, ANDER M E. On the Robust Estimation of Power Spectra, Coherences, and Transfer Functions[J]. Journal of Geophysical Re-

- search; Solid Earth, 1987, 92(B1): 633-648.
- [9] 柳建新, 严家斌, 何继善, 等. 基于相关系数的海底大地电磁阻抗 Robust 估算方法[J]. 地球物理学报, 2003, 46(2): 241-245.
LIU Jian-xin, YAN Jia-bin, HE Ji-shan, et al. Robust Estimation Method of Sea Magnetotelluric Impedance Based on Correlative Coefficient[J]. Chinese Journal of Geophysics, 2003, 46(2): 241-245.
- [10] 张刚, 王绪本, 张伟, 等. 重复中位数估计 Robust 算法在长周期大地电磁中的应用[J]. 物探与化探, 2011, 35(6): 813-816, 828.
ZHANG Gang, WANG Xu-ben, ZHANG Wei, et al. The Application of Robust Algorithm Based on Repeated Median Estimator in the Long-period Magnetotelluric Survey[J]. Geophysical and Geochemical Exploration, 2011, 35(6): 813-816, 828.
- [11] CHAVE A D. Magnetotelluric Data, Stable Distributions and Improprity: An Existential Combination[J]. Geophysical Journal International, 2014, 198(1): 622-636.
- [12] 郭俊昌, 郭荣文, 柳建新. 大地电磁阻抗估计方法研究进展[J]. 工程地球物理学报, 2017, 14(4): 409-414.
GUO Jun-chang, GUO Rong-wen, LIU Jian-xin. The Process of Research in Magnetotelluric Impedance Estimation Method[J]. Chinese Journal of Engineering Geophysics, 2017, 14(4): 409-414.
- [13] 李晋, 汤井田, 燕欢, 等. 基于递归分析和聚类的大地电磁信噪辨识及分离[J]. 地球物理学报, 2017, 60(5): 1918-1936.
LI Jin, TANG Jing-tian, YAN Huan, et al. Identification and Separation of Magnetotelluric Signal and Noise Based on Recurrence Analysis and Clustering[J]. Chinese Journal of Geophysics, 2017, 60(5): 1918-1936.
- [14] 王立伟. 数据挖掘研究现状综述[J]. 图书与情报, 2008(5): 41-46.
WANG Li-wei. The Summarization of Present Situation of Data Mining Research[J]. Library and Information, 2008(5): 41-46.
- [15] 王娟. 一种基于遗传算法的 K-means 聚类算法[J]. 微型机与应用, 2011, 30(20): 71-73, 76.
WANG Juan. A K-means Clustering Algorithm Based on Genetic Algorithm[J]. Microcomputer and Its Applications, 2011, 30(20): 71-73, 76.
- [16] 苏新宁, 杨建林, 邓三鸿, 等. 数据挖掘理论与技术[M]. 北京: 科学技术文献出版社, 2003.
SU Xin-ning, YANG Jian-lin, DENG San-hong, et al. Data Mining Theory and Technology[M]. Beijing: Scientific and Technical Documentation Press, 2003.
- [17] 孙即祥. 现代模式识别[M]. 北京: 高等教育出版社, 2008.
SUN Ji-xiang. Modern Pattern Recognition[M]. Beijing: Higher Education Press, 2008.
- [18] 蒋玲霞. 基于能量聚集的 Robust 阻抗估计方法[D]. 长沙: 中南大学, 2009.
JIANG Ling-xia. Robust Estimation of Impedance Based on Energy Assembled[D]. Changsha: Central South University, 2009.
- [19] LAI J Z C, LIAW Y C. Improvement of the K-means Clustering Filtering Algorithm[J]. Pattern Recognition, 2008, 41(12): 3677-3681.
- [20] 毛国君, 段立娟. 数据挖掘原理与算法[M]. 2 版. 北京: 清华大学出版社, 2007.
MAO Guo-jun, DUAN Li-juan. The Principle and Algorithm of Data Mining[M]. 2nd ed. Beijing: Tsinghua University Press, 2007.
- [21] 李川, 姚行艳, 蔡乐才. 智能聚类分析方法及其应用[M]. 北京: 科学出版社, 2016.
LI Chuan, YAO Xing-yan, CAI Le-cai. Intelligent Clustering Analysis Method and Application[M]. Beijing: Science Press, 2016.
- [22] 周世兵, 徐振源, 唐旭清. K-means 算法最佳聚类数确定方法[J]. 计算机应用, 2010, 30(8): 1995-1998.
ZHOU Shi-bing, XU Zhen-yuan, TANG Xu-qing. Method for Determing Optimal Number of Clusters in K-means Clustering Algorithm[J]. Journal of Computer Applications, 2010, 30(8): 1995-1998.
- [23] 朱威, 范翠松, 姚大为, 等. 矿集区大地电磁噪声场源分析及噪声特点[J]. 物探与化探, 2011, 35(5): 658-662.
ZHU Wei, FAN Cui-song, YAO Da-wei, et al. Noise Source Analysis and Noise Characteristics Study of MT in an Ore Concentration Area[J]. Geophysical and Geochemical Exploration, 2011, 35(5): 658-662.
- [24] 汤井田, 徐志敏, 肖晓, 等. 庐枞矿集区大地电磁测深强噪声的影响规律[J]. 地球物理学报, 2012, 55(12): 4147-4159.
TANG Jing-tian, XU Zhi-min, XIAO Xiao, et al. Effect Rules of Strong Noise on Magnetotelluric(MT) Sounding in the Luzong Ore Cluster Ares[J]. Chinese Journal of Geophysics, 2012, 55(12): 4147-4159.