

DOI:10.19651/j.cnki.emt.2005499

# 一种新的过套管电阻率测井数据预处理方法

马 营<sup>1</sup> 汪小军<sup>1</sup> 张光洲<sup>2</sup> 孙业杰<sup>1</sup>

(1.中国石油集团测井有限公司辽河分公司 盘锦 124000; 2.中国石油集团测井有限公司大庆分公司 大庆 163000)

**摘 要:** 为解决传统的过套管电阻率测井数据预处理方法存在的主观性强,适用性较差以及不能反映地层中的薄层信息,连续曲线分辨率低等问题,设计了一种新的数据预处理方法。首先采用数据清洗原理剔除同一深度上的异常测量值,再根据信号稀疏表示原理,构建观测矩阵和稀疏变换矩阵,通过最优化算法及约束条件求解信号的稀疏表示系数,得到等深度间隔的连续过套管电阻率测井曲线。在SL油田5口井数据上的实验结果为数据中的异常测量值得到了有效剔除,且形成的连续测井曲线与裸眼井微电位曲线具有良好的薄层对应性。结果表明,所提方法较传统的预处理方法,适用性好,具备较强的薄层分辨能力和更大的推广应用价值,且形成的软件通过在SL油田动态监测中的应用,为油田开发和持续稳产提供了有力的技术支撑。

**关键词:** 过套管电阻率测井;信号稀疏表示;数据预处理

**中图分类号:** TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.4030

## New data pre-processing method of through casing resistivity logging

Ma Ying<sup>1</sup> Wang Xiaojun<sup>1</sup> Zhang Guangzhou<sup>2</sup> Sun Yejie<sup>1</sup>

(1.Liaohu Branch of CNPC Logging Co.,Ltd., Panjin 124000, China; 2.Daqing Branch of CNPC Logging Co.,Ltd., Daqing 163000, China)

**Abstract:** To solve the problem of worse feasibility, lower resolution of continuous curve without thin beds information in stratum of traditional data pre-processing methods in through casing resistivity logging, a new data pre-processing method is designed in this paper. Abnormal values on same depth are eliminated by data cleaning theory and the sparse representing coefficients of signal are solved by optimizing algorithm and constraints through constructing observation matrix and sparse transforming matrix based on signal sparse representing theory to get the continuous through casing resistivity logging curve with equal space of depth interval. Experimental result is that abnormal values are eliminated effectively and formed logging curves by this method has good correspondence of thin beds with open-hole micronormal curves on logging data of five wells in SL oilfield. Results show that compared with traditional pre-processing methods, the method has good usability, stronger thin beds dividing ability and more greater application value. The implemented software provides powerful technical support for development and sustained and stable production in the oilfield during the application of dynamic supervising of SL oilfield.

**Keywords:** through casing resistivity logging; signal sparse representation; data pre-processing

## 0 引 言

油田开发中后期,剩余油监测变得越来越重要。过套管电阻率测井是在套管内多个不等间隔的深度点上测量地层电阻率的一种生产测井方法,属于电法测井技术,不受孔隙度的影响,且探测径向范围较大,在剩余油监测方面较放射性测井技术有着较为独特的优势,并在辽河、江苏等油田得到了广泛应用<sup>[1-3]</sup>。

过套管电阻率测井首先测量由套管和地层引起的电压

降,得到套管和地层的电阻率之和,再测量套管电阻率。两步测量结果的差值即为地层电阻率测量值。在测量过程中,电流发射电极和测量电极的电极尖通过推靠刺入套管,减少电极与套管的接触电阻。为防止套管壁沾污或套管腐蚀严重影响测量结果,保证数据可靠性,仪器固定在目的层段的某个深度点,进行重复测量;同时由于电极深入套管,仪器只能选择性的在目的层段的某些深度点进行测量,测量数据表现为不等深度间隔的点测、多值数据。因此在过套管电阻率测井数据进行处理解释之前,必须进行预处理,

收稿日期:2020-12-07

剔除单个深度上的异常数据,并形成等深度间隔的连续曲线。

文献[4]采用手工选择的方式,由解释人员根据主观经验判断异常值,但这样造成不同的解释人员得到不同的结果,且缺乏统一的判别标准;文献[5]根据相同深度点上所有测量值的平均值,计算每个测量值残差,选择残差最小的测量值作为初始智能选点,但这种方法只选择一个测量值,造成预处理误差较大,用户经常发现测量值与裸眼井原状地层电阻率曲线对比时误差较大,不合需求;文献[6-8]采用 Grubbs 准则剔除异常测量值,需要构建与危险率和测量次数相关的临界值表,完全依赖地区经验,造成该方法适应性和推广性较差。以往的研究成果<sup>[4-8]</sup>采用不同的插值方法(包括线性插值、拉格朗日插值、样条插值、抛物线插值等),对不等深度间隔的数据进行插值,形成等深度间隔的连续曲线。但有些插值方法会造成奇异点,在实际井处理中,需要用户凭借经验进行选择插值方法;同时,插值方法形成的连续光滑曲线,漏失了地层中重要的薄层信息,造成曲线分辨率较差。

经过实际测量数据的研究比对发现,异常测量值一般表现为离群的孤立点,因此本文根据数据清洗原理,采用基于距离的孤立点检测算法,检测出异常测量值,并进行剔除,其优点是算法简单、高效,不需要任何先验统计模型。同时,根据信号稀疏表示原理,将过套管电阻率在不等间隔深度上的测量点,看作是不等采样间隔的时序采样信号,通过构建观测矩阵和信号的稀疏变换矩阵,求解最优的信号稀疏表示系数,由稀疏变换矩阵,反变换得到原始信号。通过在 SL 油田实际井的预处理结果,以及与裸眼井测量的高分辨率微电位曲线的比对,验证异常测量值剔除方法的有效性以及本文方法形成的等深度间隔连续测井曲线的地层分辨率,以及薄地层的分辨能力。本文方法的研究及形成的软件可应用于 SL 油田的过套管电阻率测井资料预处理,为各采油厂的油藏动态监测提供了优质、高效服务。

## 1 基于孤立点检测的异常测量值剔除

孤立点检测作为数据清洗中实例检测的一个重要任务<sup>[9]</sup>,通常分为基于样本距离的检测算法<sup>[10]</sup>、基于样本密度的检测算法<sup>[11]</sup>和基于模型的检测算法<sup>[12]</sup>。基于样本密度算法将周围近邻样本密度小的数据点看作是孤立点,通过样本  $k$  近邻的距离信息计算局部样本密度,但在数据稀疏的情况下,无法将正常点和孤立点有效分开。基于模型的检测算法需要使用测量正常值拟合检测使用的模型,而在过套管电阻率测井环境下,每个深度点的正常测量值是未知的,因此该算法不适用。

本文采用基于样本距离的孤立点检测算法,基于距离的孤立点定义为与  $k$  个近邻的平均距离在前  $n$  大的  $n$  个数据点<sup>[9]</sup>。因此其基本算法是遍历数据集中的每个数据点,计算其在整个数据集中的  $k$  近邻,再根据每个数据点到  $k$

个近邻的平均距离对所有数据点进行排序,选出距离最大的前  $n$  个数据点作为孤立点。但算法对于每个数据点需要遍历整个数据集,算法时间复杂度大。因此,本文采用了基于简单剪枝策略和随机数据顺序的检测算法,具体算法描述如下:

- 1) 对所有数据点随机排序,设定近邻数  $k$  和输出的孤立点数目  $n$ ;
- 2) 将用于剪枝的剪切阈值 cutoff 置为 0,清空孤立点集合;
- 3) 从整个样本集  $D$  中顺序取出一部份样本,形成集合  $B$ ;并将  $B$  中每个样本的邻域集合清空;
- 4) 对于样本集  $D$  中的每个样本  $d$ ,在集合  $B$  中寻找互不相同的样本  $b$ ,如果  $b$  的邻域数目  $< k$ ,或者  $b$  与  $d$  的距离  $< b$  与其邻域集合的最大距离,则将  $b$  的邻域集合与样本  $d$  的合并集中与  $b$  的距离最近的  $k$  个样本放入  $b$  的邻域集合;同时,计算  $b$  与其邻域集合内  $k$  个样本的平均距离,如果小于阈值 cutoff,则将  $b$  从集合  $B$  中移除;
- 5) 计算  $B$  与孤立点集合的合并集中每个样本点到  $k$  个近邻的平均距离,根据从大到小的距离排序,将前  $n$  个样本点放入孤立点集合;
- 6) 计算孤立点集合中每个样本到  $k$  个近邻的平均距离,将 cutoff 更新为最小距离值;
- 7) 返回步骤 3),直至样本集  $B$  为空;输出孤立点集合,作为最终结果。

可以看到,cutoff 在每次循环时进行了更新,随着处理样本的增多,cutoff 随之增加,提高了剪枝效率。但实际使用时,孤立点数目不能确定,因此本文在该算法的基础上,采用  $3\sigma$  准则,在步骤 5)中,将不符合  $3\sigma$  准则的样本点放入孤立点集合,而不是固定数目  $n$ ,使算法能够自适应的选择  $n$  值。

## 2 基于信号稀疏表示的连续测井曲线生成

在将异常测量值剔除后,计算非异常值的平均值,作为该深度点的实际测量值。在此基础上,根据信号稀疏表示原理生成等深度间隔的连续测井曲线。

在信号稀疏表示原理中,如果原始信号能够进行稀疏变换,则采用远低于奈奎斯特采样定理要求的采样频率,采集得到的观测信号,可以通过最优化问题求解,恢复原始信号,这一点已经得到理论证明<sup>[13]</sup>。

对原始信号  $f$  进行稀疏变换,如式(1)所示。

$$\alpha = Df \quad (1)$$

其中,  $f \in \mathbf{R}^{N \times 1}$  为  $N$  个采样点的等间隔采样原始信号;  $D \in \mathbf{R}^{M \times N}$  为的稀疏变换矩阵;  $\alpha \in \mathbf{R}^{N \times 1}$  为  $f$  进行稀疏变换的稀疏表示系数,则有  $f = D^H \alpha$ ,  $D^H$  为  $D$  的共轭转置阵,满足  $D^H D = 1$ 。原始信号  $f$  通过如下观测模型:

$$y = \Phi f \quad (2)$$

得到  $M$  个点的观测信号  $y \in \mathbf{R}^{M \times 1}$ ;通过由  $(0,1)$  组成的观测矩阵  $\Phi \in \mathbf{R}^{M \times N}$ ,可以由原始信号  $f$  得到不等采样

间隔的低密度观测信号  $y$ , 其中  $N$  远大于  $M$ ; 在本文中, 将深度信号看作时序采样信号, 则原始信号  $f$  等深度间隔序列为  $[d_1, d_1 + \Delta t, d_1 + 2\Delta t, \dots, d_1 + (N - 1)\Delta t]$ , 其中  $\Delta t$  为深度采样间隔; 信号  $y$  的观测深度序列为  $[d_1, d_2, \dots, d_M]$ , 则根据深度点  $d_i, i = 1, \dots, M$  与信号  $f$  的深度  $d_1 + j\Delta t, j = 0, \dots, N - 1$  最小差值的序号对应  $(i, j)$ , 将预先置零的观测矩阵  $\Phi$  位置  $(i, j + 1)$  处置 1。由式 (1) 和 (2) 可知:

$$y = A\alpha, A = \Phi D^H \quad (3)$$

由于  $\alpha$  是稀疏的, 可通过求解最优化问题, 得到最稀疏的表示系数  $\tilde{\alpha}$ , 通过  $f = D^H \tilde{\alpha}$  反变换, 得到原始信号  $f$  的估计, 作为生成的等深度间隔测井曲线。其中, 最优化问题描述如下:

$$\tilde{\alpha} = \arg \min_{\alpha} \|\alpha\|_1, \text{ s.t. } y = \Phi D^H \alpha \quad (4)$$

通过系数  $\alpha$  的  $l_1$  范数的最小化, 不断提高其稀疏性, 而约束条件  $y = \Phi D^H \alpha$  使  $\alpha$  的估计值向真实值收敛。本文采用正交匹配追踪算法<sup>[14-16]</sup>, 求解上述优化问题。根据稀疏表示原理, 将  $y$  看作要稀疏表示的信号,  $\Phi D^H$  为过完备原子库, 为  $M \times N$  大小的矩阵, 其中每列为一个原子, 表示为  $a_i, i = 1, \dots, N$ ; 初始化残差  $r_0 = y$ , 迭代次数  $t = 1$ ; 计算每个原子  $a_i$  与残差  $r_{t-1}$  的内积, 将内积最大的原子以列向量形式加入初始为空的临时原子库  $A_t$ , 再通过求解  $\hat{\theta} = \arg \min_{\theta} \|y - A_t \theta\|_2$  这一最小二乘问题得到已选原子的稀疏系数估计  $\hat{\theta}$ , 更新残差  $r_t = y - A_t \hat{\theta}$ , 迭代次数  $t$  加 1, 重复上述过程, 直至迭代次数  $t$  达到预设值。算法结束后, 稀疏系数估计  $\hat{\theta}$  即为稀疏表示系数  $\alpha$  的估计值。本文经过实验结果验证, 稀疏变换矩阵  $D$  采用  $M \times N$  维单位矩阵  $I$  的傅里叶变换域。

### 3 实验结果及分析

实验在 SL 油田 Z12-X3013、Y22、X855-X23 等 5 口井的实测过套管电阻率数据上, 验证本文预处理方法的有效

性, 如图 1~4 所示。图 1(a) 和图 3(a) 分别给出了 Z12-X3013 井 1 238~1 251 m 和 Y22 井 1 218~1 231 m 的过套管电阻率实测数据的可视化显示; 其中, 大实心圆点表示原始测量值, 可以看到, 在横向上, 大圆点有多个值, 说明仪器在某个深度上进行了重复测量; 而在纵向上, 可以看到原始测量点的分布是不等深度间隔的。同时, 软件提供了深感应曲线 HRID, 代表原状地层电阻率, 方便用户比对。这里, 对 HRID 曲线值和过套管电阻率测量值, 使用 0.1~100  $\Omega \cdot m$  的对数坐标进行了刻度。图 1(b) 和图 3(b) 给出了使用基于孤立点检测的异常测量值剔除结果, 其中的三角形表示剔除的异常值, 可以看到负值和明显的离群孤立点都被检测出来, 说明了方法的有效性。本文中对单个深度上只有两个测量值的情况, 算法不做检测, 只是将其中可能的负值剔除。在过套管电阻率测井连续曲线生成过程中, 本文将在对应深度点上剔除异常值后的平均值, 作为当前深度点的实际测量值, 分别采用三次样条插值方法和本文方法, 在 Z12-X3013 井 1 238~1 251 m 井段和 Y22 井 1 218~1 231 m 井段生成深度采样间隔为 0.1 m 的连续曲线。图 2(a) 和图 4(a) 分别给出了 3 次样条插值方法生成的连续曲线, 其中小实心圆点代表被插值点; 可以看到, 插值曲线极为平滑, 与 0.1 m 深度采样间隔的高分辨微电位实测曲线比对发现, 插值曲线分辨率较低, 不能识别微电位曲线反应的薄层, 特别是图 4(a) 中 1 228~1 230 m 段由于插值方法的局限性, 生成的曲线出现了急剧陡峭的拐点, 电阻率值下降到 0.3  $\Omega \cdot m$ , 明显与同深度的深感应和微电位实测数据不符。图 2(b) 和图 4(b) 分别给出了两口井基于信号稀疏表示方法生成的预处理曲线, 深度上的轮廓与插值曲线具备较高的相似性, 保持了插值曲线的大致走向, 但包含了更丰富的地层细节信息, 通过与微电位曲线对比发现, 两者具备良好的薄层对应性, 说明生成的预处理曲线的地层分辨率更高, 能够有效识别地层中的薄地层。图 4(b) 中 1 228~1 230 m 段的处理结果比图 4(a) 的插值结果更符合实际的地层电性特征。

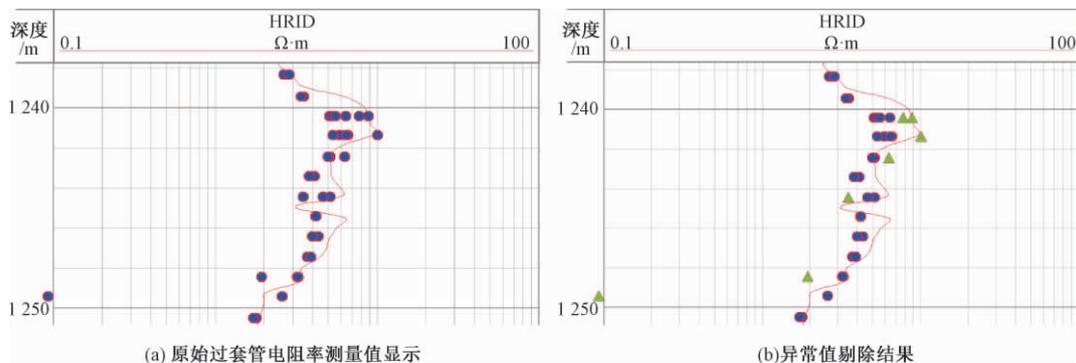


图 1 Z12-X3013 井 1 238~1 251 m 实际测量值及异常值剔除结果

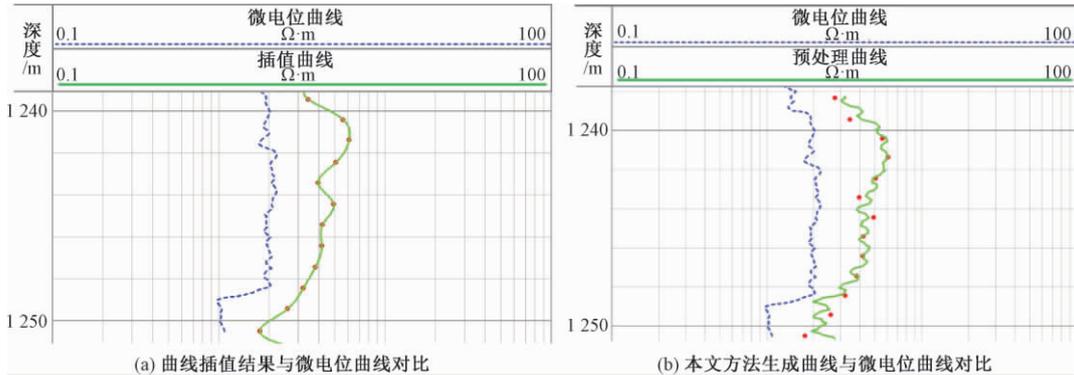


图 2 Z12-X3013 井 1 238~1 251 m 连续曲线生成及对比

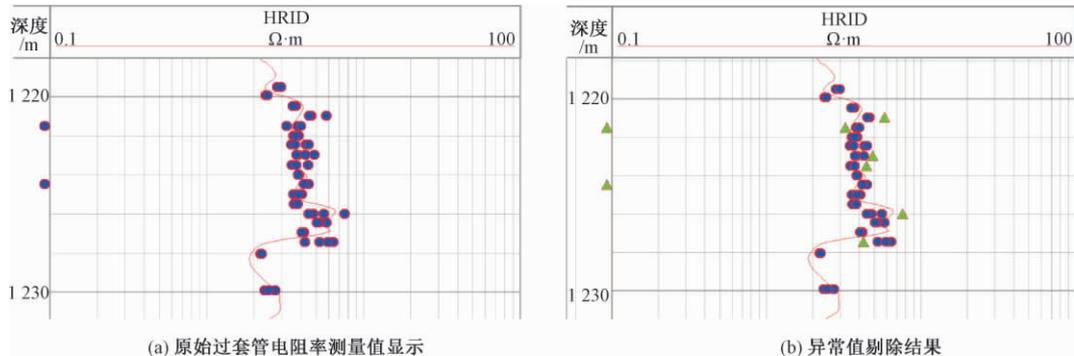


图 3 Y22 井 1 218~1 231 m 实际测量值及异常值剔除结果

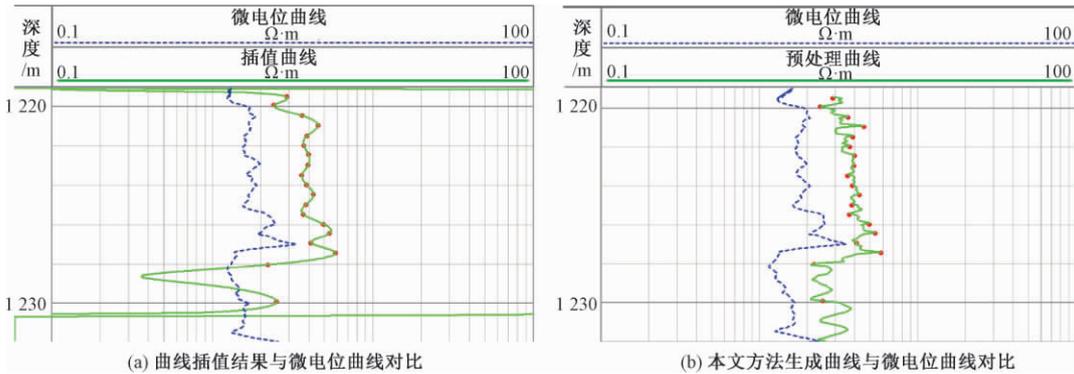


图 4 Y22 井 1 218~1 231 m 连续曲线生成及对比

### 4 应用实例

本文将上述异常值剔除方法和等深度间隔连续测井曲线生成方法,通过 Visual C++ 在 Visual Studio 软件开发平台上进行了编程实现,除了算法实现外,还开发了图形界面,为用户提供直观和可视化的预处理结果呈现方式。

本文方法形成的软件已经应用于 SL 油田的过套管电阻率测井资料处理过程中。图 5 所示为 Z12-X3013 井的套后解释成果图,其中根据裸眼井测井资料,在 1 219. 875~1 227. 25 m 解释为 19 号水层,可以看到产液剖面道中该

层孔隙度较大,含水饱和度较高,说明该层储集物性较好,故在动态监测中,重点对该层进行了过套管电阻率测井。电阻率曲线道中,给出了根据本文预处理方法生成的过套管电阻率连续曲线,可以看到 1 220. 8~1 224. 3 m 处过套管电阻率曲线明显大于裸眼的深中感应曲线,电阻率值平均提升了  $5 \Omega \cdot m$ ,说明该处在油井生产一段时间后出现了油气聚集,故套后解释为油水同层;而 21 号层由于套后电阻率比裸眼电阻率增加数值较小,说明该层油气聚集量较少,故根据套后解释标准,定为含水水层。22 和 23 号层由于套后电阻率与裸眼电阻率相比明显下降,说明该井生产后,在这两个层位注入水侵入地层孔隙,造成水淹,故分别

解释为弱水淹层和中水淹层。由本文方法生成的高分辨率过套管电阻率曲线,对裸眼解释的 19 号层进行了更加精细的套后评价,实现了产层的精细划分。对套后解释的

20 和 21 号层进行试油,日产油 0.5 t,水 9.3 m<sup>3</sup>,符合套后解释结论,验证了本文软件在使用过套管电阻率测井进行油田剩余油监测这一实际应用中的有效性。

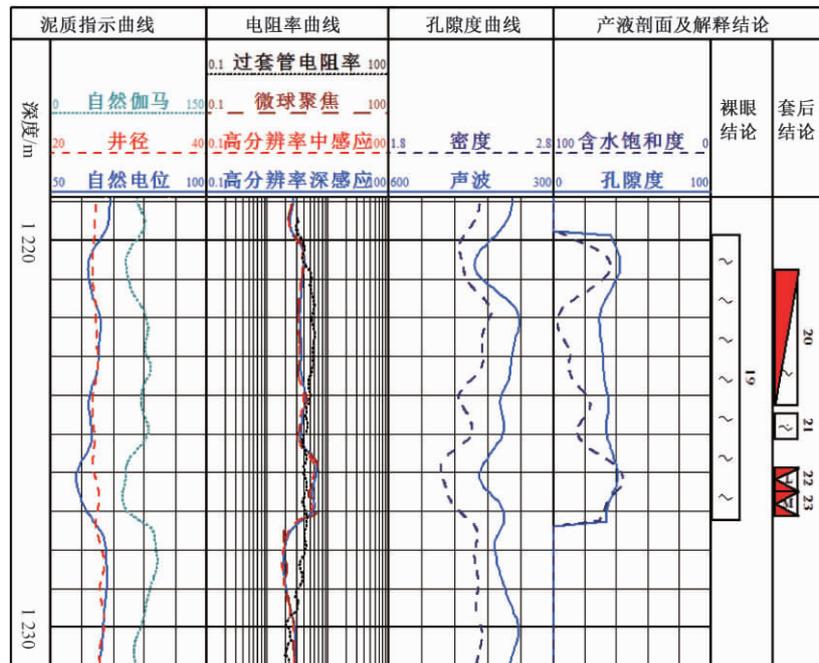


图 5 Z12-X3013 井 1 219~1 231 m 套后解释成果

## 5 结 论

本文针对过套管电阻率测井数据特征,采用基于孤立点检测的异常值剔除方法和基于信号稀疏表示的连续曲线生成方法,进行测井资料的预处理,在 SL 油田的储层动态监测和生产管理中发挥了重要作用。本文结合实验分析,比较插值方法和本文预处理方法的实际处理成果,表明了本文方法生成的高分辨率曲线,在薄地层信息识别方面的有效性。

### 参考文献

- [1] 万吉庆. 俄罗斯过套管电阻率测井技术应用效果评价[J]. 物探与化探, 2010, 34(3):362-366.
- [2] 王正国, 李丽, 杨贵凯, 等. 俄罗斯过套管电阻率测井技术在江苏油田的应用[J]. 国外测井技术, 2010(3): 40-42.
- [3] 孔庆东. 过套管电阻率测井在海外油田的适用性分析[J]. 测井技术, 2020, 44(2):141-145.
- [4] 岳喜洲. 过套管电阻率测井解释方法研究[D]. 青岛: 中国石油大学(华东), 2010.
- [5] 张守伟, 孙建孟, 伍东, 等. 俄罗斯过套管电阻率测井资料预处理软件的设计与开发[J]. 计算机测量与控制, 2010, 18(5):1189-1192.
- [6] 孙中春, 罗兴平, 周继宏, 等. 过套管电阻率测井资料预处理方法[J]. 测井技术, 2011, 35(2):151-154.
- [7] 谢关宝, 李三国, 李永杰. 过套管电阻率测井资料预

处理方法研究[J]. 科技导报, 2012,30(3):57-60.

- [8] 袁瑞. 过套管电阻率测井资料预处理方法研究[D]. 荆州: 长江大学, 2013.
- [9] 苗润华. 基于聚类和孤立点检测的数据预处理方法的研究[D]. 北京: 北京交通大学, 2012.
- [10] PAMULA R, DEKA J K, NANDI S. An outlier detection method based on clustering [C]. Second International Conference on Emerging Applications of Information Technology, IEEE Computer Society, 2011:253-256.
- [11] 吕奔, 高茂庭. 基于密度的孤立点检测算法改进研究[J]. 现代计算机, 2015 (6):62-67.
- [12] 夏娟. 孤立点检测算法研究[J]. 信息系统工程, 2014(2):142-142.
- [13] 申伟, 石平. 单演信号随机加权融合的 SAR 图像目标识别方法[J]. 电子测量与仪器学报, 2020, 34(9): 186-192.
- [14] 马云飞, 贾希胜, 胡起伟, 等. 基于 EMD 和 BCS 的振动信号数据修复方法[J]. 仪器仪表学报, 2019, 40(3):154-162.
- [15] 孔舒亚, 白新有, 王亮, 等. 基于压缩感知的 SAR 自适应窄带干扰抑制方法[J]. 国外电子测量技术, 2019, 38(7):65-69.
- [16] 元沐南, 李晓风, 李哲茹, 等. 基于压缩感知的三维足型重建平台[J]. 电子测量技术, 2020,43(9):99-103.

### 作者简介

马营, 大专, 高级技师, 主要研究方向为生产测井资料采集。

E-mail:1959720094@qq.com