

基于小波变换和模式识别随钻测量系统通信信号处理方法研究

董夏妮 王成林

(北京航空航天大学 电子信息工程学院 北京 100191)

摘要: 基于泥浆脉冲信道的通信信号处理是石油随钻测量系统的关键技术之一。针对恶劣工况环境下的泥浆脉冲信号数据处理,分析了泥浆脉冲信道衰减特点及其噪声特性,建立了泥浆脉冲信号的滤波、识别模型,给出结合小波变换、相关滤波、模式识别、聚类相关的组合处理方法,避免了以 FIR 滤波、门限法识别的组合方法,其实时性差、门限值对解码结果影响大的局限性。现场测试结果表明,该方法能够有效解算井下测量数据,改善了信号识别率,提高了实时性。

关键词: 随钻系统;泥浆脉冲;小波变换;信号识别

中图分类号: TN911.72 **文献标识码:** A **国家标准学科分类代码:** 510.4030

MWD system based on wavelet transformation and pattern recognition communication signal processing method and research

Dong Xiani Wang Chenglin

(Electronic information engineering college, Beihang University, Beijing 100191, China)

Abstract: signal processing based on mud pulse is one key of technologies of MWD system. In this paper, the characteristics of mud pulse channel and its noise is analyzed, signal filtering model and identification model is established, one signal processing method consists of wavelet transform, correlation filtering, pattern, recognition and clustering is given for mud pulse signal data processing under bad working environment, which avoid to limitations of bad real-time performance and bad influence from the threshold Field test results show that the method can decode down hole measurement data effectively and improve the signal recognition rate and real-time performance.

Keywords: while drilling system; mud pulse; wavelet transform; signal recognition

1 引言

随钻测量通信地面监测系统配合井下数据传输系统,对泥浆脉冲采样信号进行信号滤波识别解码,从而得到井下测量数据。有效的泥浆脉冲信号处理手段能对井下作业进行实时监测并为钻井提供有价值的参考数据。

在高温高压深井,噪声情况复杂,信号几乎完全淹没于噪声,如何提取井下数据是测井成败的关键一步。传统泥浆脉冲信号处理^[1]通常以 FIR 滤波为主,难以完成恶劣环境下信号还原。而信号识别门限值的设定对时隙要求很高的 M-ary 码影响很大,需要有经验的测井人员进行设定^[2]。基于上述原因,首先对泥浆脉冲信道,信道内噪声以及编码方式进行分析;在此基础上研究以多分辨率小波变换为主的滤波识别算法,运用数据统计学原理对信号进行识别解

码;最后对解码结果进行分析和验证,并具体运用到系统工程中,系统框图如图 1 所示。针对实际测井工况,使用小波变换、模式识别、聚类的组合数据处理方法,有效提高了信号识别率、测量实时性。

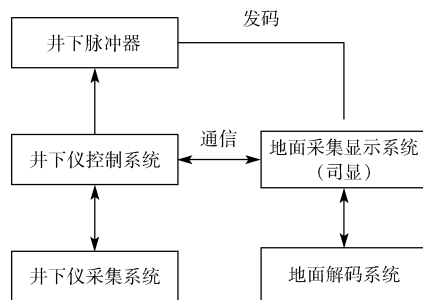


图 1 本系统随钻泥浆系统框图

2 泥浆脉冲信道分析及信号编码方式

2.1 信道传输特点

泥浆脉冲传输无线随钻测量主用通信方式,传输距离远,但传输速率低。

信号在泥浆中传播,对低频系统的衰减幅度可以简化为如下^[3-4]:

$$P = P_0(1 - e^{-\alpha d}) \quad (1)$$

式中: P_0 指井口采集到的最大压力值; P 指井口采集压力值; d 指井深; f_s 指信号频率; a 指套管的几何特性、弹性特性及泥浆特性的相关系数; u 指脉冲信号传输速度。

选择合适的信号传输频率,能够减少能源损耗以达到在特定环境有最佳的信号传输结果。

2.2 信道噪声分析

泥浆脉冲传输受现场工况的限制,信号在传输过程中不仅大幅衰弱^[5-6],而且存在多种类型干扰,主要包括:机械部件周期性活动引起的压力脉冲产生的干扰(泵噪声)、泥浆中的气泡产生的干扰、活动钻具等机械结构压力突变产生的信号干扰和杂波、传感器、信号传输过程中产生的电磁噪声。这些噪声产生的原因复杂多样,与泵噪声相比随机性很大,幅度比较小,频带比较宽,可近似认为是白噪声,与泵噪声和有用信号均不相干。

总的来说,该泥浆脉冲信号通过部分响应基带传输进行的,基带带宽为0.2~2 Hz。噪声存在时频域周期性及非周期性噪声,噪声带宽大于信号带宽,且频带发生交叠。

2.3 M-ary 编码定义

目前泥浆通信系统普遍采用 M-ary 编码方法进行信源编码,有两种编码方式。

1)四分码,占8个槽宽,第一个槽为一个空槽,每个脉冲占据两个槽宽,最后空两个槽;编码为3、2、1、0;

2)八分码,占12个槽宽,编码方式相同,编码为7、6、5、4、3、2、1、0。

3 泥浆脉冲信号去噪方法分析

本 MWD 系统,针对不同测井环境设计了多种算法,以提高滤波结果的可靠性,实时性,简化运算。

对于信道理想的环境,使用滑动平均滤波、中值滤波就能提取被测信号。但对于存在复杂噪声的泥浆脉冲信号,通常采用 FIR 带通滤波、小波变换,甚至混合若干种不同滤波方法消除噪声,同时结合相关滤波算法以提高可识别准确率。

3.1 小波消噪

泥浆脉冲信号并不是简单的平稳线性时不变信号,而小波变换对非平稳信号或者时变信号都有较好的应用。小波的多分辨率特性^[7]能使信号按照不同频率分解成各个子信号。选择一种小波基 $\Psi(t)$ 通过伸缩平移等运算对信号进行多尺度细化分析。

小波分解的过程可以认为是一个低通滤波过程以及一个高通滤波过程,而重构信号的过程即为分解算法的逆过程。

本系统算法根据 Mallat 改进,定义 $h(n)$ 为低通权系数, $g(n)$ 为高通权系数, $h(n)$ 和 $g(n)$ 有限长度,设初始已知值如下:

$$\begin{cases} h_1, h_2, \dots, h_N, N \text{ 为偶数} \\ g_1, g_2, \dots, g_N, N \text{ 为偶数} \\ c_n^0 = x_n, n = 1, 2, \dots, M \end{cases} \quad (2)$$

$$c_k^{j+1} = \sum_{n \in Z} h_{n-2k} c_n^j, j = 0, 1, \dots, J-1 \quad (3)$$

$$d_k^{j+1} = \sum_{n \in Z} g_{n-2k} c_n^j, j = 0, 1, \dots, J-1 \quad (4)$$

式中: J 为实际应用过程中分解的步数,最多不超过 $\log_2 M$ 。

3.1.1 小波尺度参数及小波基选取

选取合适的尺度参数是平衡滤波特性以及算法特性的关键点。

计算重构信号和实际信号相关系数表明,对大部分小波基而言,尺度参数为7是信号可以近似认为是滤波后的有用信号且运算量的平衡点,而其细节信号可以认为是时频域的噪声。

小波基的选取^[8-9],通常根据正交性,紧支撑性,对称性等基本特征进行分析的。

消失距决定小波变换后能量的分布以及平滑函数的收敛特性。消失距越大,小波滤波器的带宽越长^[10]。

然而测试结果可能并不完全指向某一种小波,因此定义误差模型如下。

设待测信号 V_s 的原始信号为 V ,经过小波消噪后信号为 V_f ,则:

$$e = \frac{r_1 \sqrt{\|V - V_f\|^2} + r_2 \max(|V - V_f|)}{r_1 \sqrt{\|V - V_s\|^2}} + r_3 \frac{VV_f}{\sqrt{(VV)(V_f V_f)}} \quad (5)$$

式中: r_1 为整体数值偏差因子; r_2 为极值偏差因子; r_3 为角度偏差因子; $r_1 + r_2 + r_3 = 1$;本文选取 $r_1 = 2r_2 = 2r_3 = 0.5$ 。

均衡考虑整体偏差以及局部偏差, e 越小则偏差越小。在系统测试模式对标准正弦信号小波分析,根据计算结果动态选取误差最小的小波基进行工程运用。

3.2 相关滤波

由于泥浆脉冲信号非常微弱,不易识别。且除了泵噪声外,大部分其他噪声的频率和相位都是随机的,和有用信号及周期性泵噪声均不相关,因此可以通过相关器的处理对其进行滤波^[11],同时也能提高幅值,增强其抗干扰的能力。

设待测信号: $f_1(t) = V_s(t) + n_1(t)$

参考信号: $f_2(t) = V_r(t)$

式中: $V_s(t)$ 为有效信号, $n_1(t)$ 为噪声信号, V_r 为相

关器。

互相关函数：

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [V_s(t) + n_1(t)] V_r(t - \tau) dt =$$

$$R_{sr}(\tau) + R_{nr}(\tau) = R_{sr}(\tau)$$

离散化，

$$R(k) = R_{sr}(k) = \frac{1}{N} \sum_{n=0}^{N-1} s(n)r(n-k)$$

$$(k = 0, 1, 2, \dots, M-1) \quad (7)$$

本文选取与脉冲信号相匹配的相关函数与经过预处理的泥浆脉冲信号进行相关。

根据泥浆脉冲编码方式，可以选用方波，正弦波，三角波作为相关函数，如图 2 所示。

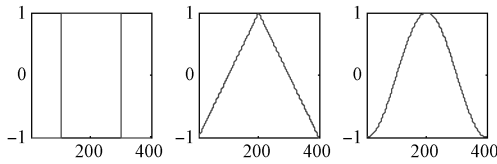


图 2 三种数字相关器

本文选取脉冲宽度为 2 s 的正弦波进行相关运算，能对脉冲周围的信号进行有效陷波，提高波形识别率。

3.3 泵压消除

根据噪声分析，泥浆脉冲存在低频泵压干扰，且接近有用信号频段，是主要的干扰源。

本文利用小波更多层次的分解，从未去除泵压干扰的信号中获得泵压曲线。

4 信号识别方法

4.1 相似性测度方法

M-ary 编码是一种时分编码方式，对时隙要求很高，传统的码元识别一般是通过设置门限，来获得发送的脉冲码元，但这种方法解码质量很大程度取决于门限值，不同的门限值对波形的识别乃至解码都存在很大影响。

采用模式识别的方法对各个 M-ary 码元的进行求解，避免了通过门限获得脉冲的缺点，一定程度能提高波形识别正确率以及解码正确率。根据实际工况本文采用若干种相似性测度^[12-13]，加以不同的权重比，对波形进行识别，从而更大程度适应复杂的工况条件，以下为几种使用的测度方法。

1) 欧氏距离 D ：

$$D = \|X_i - X_j\|^2 \quad (8)$$

2) 夹角余弦 S (范围为 0~1)：

$$S(X_i, X_j) = \frac{X_i X_j}{\sqrt{(X_i X_i)(X_j X_j)}} \quad (9)$$

3) Tanimoto 测度 T (范围为 0~1)：

$$T(X_i, X_j) = \frac{X_i X_j}{X_i X_i + X_j X_j - X_i X_j} \quad (10)$$

4) Soergel 距离(范围为 0~1)：

$$D_s = \frac{\sum_{k=0}^n |x_{ik} - x_{jk}|}{\sum_{k=0}^n \max(x_{ik}, x_{jk})} \quad (11)$$

除了相似性测度方法，相关系数一定程度上也能用来验证信号间的相关性，相关系数(范围为 0~1)如下：

$$r(X, Y) = \frac{\sum_{i=0}^n (x_i - \bar{x}) \sum_{i=0}^n (y_i - \bar{y})}{\sqrt{\sum_{i=0}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=0}^n (y_i - \bar{y})^2}} \quad (12)$$

相关系数 $r(X, Y)$ 越大，说明两个信号越相似。当 $r(X, Y) = 0$ ，则两个信号完全不相干；当 $r(X, Y) = 1$ ，则两个信号完全相关。

4.2 同步头查找

Mary 编码是根据不同时间槽上存在的脉冲来进行编码，因此同步头查找的准确度对解码准确性起到至关重要的一步。

建立三槽宽的标准同步头，计算标准同步头和信号的相关系数获得同步头位置，相关性高达 0.978。

4.3 模式识别重建 M-ary 编码信号

除同步头之外，建立四分码标准波形，八分码标准波形。设置标准波形为两个槽宽的正弦信号。

最终对波形的识别，通过表决机制来判定数据位。若这几种相似性测度的结果不一致，则发出警告，表示可能出现误码。分析测度结果可知该段码元 M-ary 编码结果。

本文尝试采用聚类算法进行模式识别的训练，标准波形库具有多个初始化实例，采用基于距离的聚类算法^[14-15]，根据不同测度的距离来衡量数据间的相似度，具体方法如下。

- 1) 选择标准码元库的码元作为初始的簇中心。
- 2) 将标准化后的信号段和标准码元库进行模式识别计算，得到对应簇，即 M-ary 码元。
- 3) 更新簇值，即将该信号段作为标准码元库一段能够代表该码元的信号段，并重新计算簇中心。
- 4) 重复以上步骤。

聚类识别的优势在于有效适应当前工况特有环境并排除异常数据干扰，对几乎被噪声淹没的信号依然具有甄别能力。

5 算法结果分析

设噪声为高斯白噪声，给出识别概率模型满足正态分布，其概率密度函数近似：

$$\rho = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (13)$$

式中： σ^2 为方差； μ 为数学期望； μ 为标准码元库中线； x 为实测码元中线。

本文基于某油田 S72-25x，桩古 18-1ST 井的泥浆脉冲数据，使用本 MWD 系统上位机如图 3 进行长时间作业，表明本文的方法可以有效获得井下传输数据。对概率密度大

于60%的码元,综合统计解码正确率约为98.9%,取一段新疆深井信号的解码结果,综合比较不同上位机方案对泥浆脉冲信号的解码结果,如表1所示。

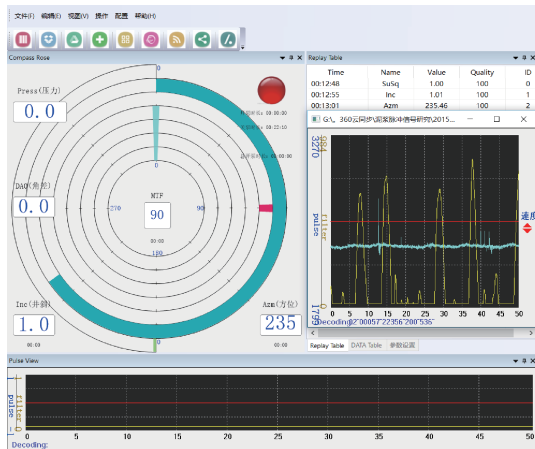


图3 运行中本MWD系统界面

表1 解码质量分析

组别	某研究所仪器	国外某仪器	本课题仪器
Inc	1.01	1.010 989 011	1.01
Azm	235.46	235.458 984 4	235.46
aTFA	9 999.00	-49.5	-49.5
Temp	257.00	257	257
Magf	71.45	0.714 529 915	71.45
DipA	80.53	80.549 450 55	80.55
Grav	0.998 5	0.998 778 999	0.998 7

排除数据精确位数,可以看出本MWD上位机系统解码结果和国外某仪器一致。

6 结 论

本文分析泥浆脉冲信道传输特性和噪声,确定以小波变换消除噪声为主的滤波方法,针对M-ary编码方式,建立了标准码元库,通过模式识别的方法,获得M-ary码元信号,并聚类方法以不断调整优化M-ary码元标准库。

井下实际数据解码结果表明,模式识别和聚类分析是一种可以在石油测井行业深入推广的运算效率高,信号识别率高的算法。该方法能够有效、快速地解算井下测量数据。

参考文献

- [1] 远方,巩宪锋. 随钻泥浆脉冲器信号处理技术研究[J]. 煤矿机械,2011,32(10):72-74.
- [2] PETROVIC J, PETROVIC V, WHITE M R, et al. System and method for downhole telemetry: U. S. Patent 8,154,420[P]. 2012-4-10.
- [3] 刘文庭. 随钻泥浆信号处理及其DSP技术实现[D]. 青岛:中国石油大学,2012.
- [4] 李闪. 随钻泥浆脉冲信号检测与处理技术研究[D]. 青岛:中国石油大学,2010.
- [5] 杨威. 随钻泥浆脉冲信号去噪算法和解码研究[D]. 成都:西南石油大学,2015.
- [6] 杨小龙. 连通井随钻系统信号处理模块的设计与实现[D]. 广州:广东工业大学,2014.
- [7] 张磊邦,唐荣斌,蒋建波,等. 基于小波变换的心音特征值提取算法研究[J]. 电子测量技术,2014,37(6):43-46.
- [8] 陶珂,朱建军. 多指标融合的小波去噪最佳分解尺度选择方法[J]. 测绘学报,2012,41(5):749-755,748.
- [9] 陈文渊. 随钻测量系统信号测量的关键技术研究[D]. 重庆:重庆大学,2011.
- [10] 涂兵,李德胜,贺建,等. 泥浆脉冲信号互相关滤波算法的研究[J]. 仪器仪表学报,2010,31(8):43-46.
- [11] 季少石. 模式识别中两种相似性测度算法比较[J]. 甘肃科技,2011,27(9):18-19.
- [12] 丁世飞,朱红,许新征,等. 基于熵的模糊信息测度研究[J]. 计算机学报,2012,35(4):154-159.
- [13] 涂涂兵,李德胜,林恩怀. 基于聚类算法的MWD泥浆脉冲信号识别研究[J]. 传感技术学报,2012,25(8):1172-1176.
- [14] 盛九朝. 动态聚类算法在ESM信号处理中的应用[J]. 舰船电子对抗,2013,36(2):55-57,61.
- [15] 龙玲,陈庆,刘飞. 电磁波随钻测量干扰信号消除算法研究[J]. 仪器仪表学报,2014,35(9):2144-2152.

作者简介

董夏妮,1991年出生,北京航空航天大学研究生,电子信息工程学院。

Email:shinnydong@126.com