

Acoustic Gas Influx Detection based on Blind Source Separation

Gao Ying, Ou Shifeng

*Institute of Science and Technology for opto-electronic Information, Yantai University, Yantai, China
claragaoying@gmail.com*

Abstract: This thesis presents a new approach for acoustic gas influx detection based on blind source separation with some complicated additive noises. Firstly, it treats the model of acoustic gas influx detection system as an instantaneous mixing model of blind source separation. Then, the additive noise of this model is figured out and subtracted from the system. Finally, the least-square approach is employed to estimate the parameters of the system without noise. This proposed approach does not rely on any statistic characteristic of the additive noise and can work well under low SNR conditions. Synthetic data are applied to validate not only the independence upon the statistics of the noise, but also the effectiveness of the proposed method at low SNR conditions.

Keywords: Acoustic Gas Influx Detection; Blind Source Separation; least-square; additive noise

基于盲源分离的井口声发射气侵检测方法

高颖, 欧世峰

烟台大学光电信息科学技术学院, 烟台, 中国, 264005
claragaoying@gmail.com

【摘要】在复杂加性噪声背景下, 基于盲源分离技术实现井口声发射气侵检测。首先基于系统辨识理论, 将井口声发射气侵检测的系统模型纳入到盲源分离的瞬时混合模型中; 接着估计出该系统的加性噪声, 将其减去后得到不含噪声的线性系统; 最后利用最小二乘法实现对该系统的参数估计。加性噪声背景下的仿真实验表明, 本文方法不仅有效地消除了一般方法对于噪声特性的限制, 而且在低信噪比下仍能取得较为理想的辨识结果。

【关键词】声发射气侵检测; 盲源分离; 最小二乘; 加性噪声

1 引言

在钻井过程中, 地层里的气体侵入井筒会破坏泥浆性能, 降低泥浆比重, 此时容易引起井内压力不稳定、发生井涌、导致停钻除气, 甚至会造成井喷事故。不仅影响正常生产, 更污染环境、严重浪费自然资源, 并可能造成巨大的经济损失及危害工程操作人员的人身安全。因此, 气侵检测是石油钻井工业长期面临且亟待解决的理论和工程课题。

基于声波在气液两相流中传播的特性, Bang等人^[1]提出了在井口以高压水枪做声源的实时检测方法。但是迄今为止, 此方法仅处于实验油气侵入检测的阶段。因为在多次反射的复杂环境下, 要使方法投入商业运作, 实现上述功能, 信号多时间延迟等参数估计还是尚未完

全解决的难题。尤其是, 在井口声发射井筒气侵检测中, 噪声往往是非高斯的, 难以直接应用这些方法。为避免声发射传播时差检测方法对信号多时间延迟估计的不足, 文献[2]将钻井液视为一个线性系统, 利用系统辨识理论提出了一种能抑制高斯噪声的井口声发射气侵检测方法, 然而该方法只对高斯噪声有抑制作用。据此, 本文为避免一般系统辨识方法对于噪声特性的限制, 基于盲源分离技术, 提出一种新的井口声发射气侵检测方法, 简称为BL(Blind source separation - Least square)法。

2 基本原理

2.1 盲源分离基本概念

盲源分离^[3]是指从观测的多道混合信号中分离并恢复出各个源信号的过程。这里所谓的“盲”具有两个方面的含义: ①源信号是未知的; ②信号的混合方

资助信息: 烟台大学博士启动基金(WL08B19)资助

式是未知的。根据源信号混合方式的不同，盲源分离可分为线性混合（系统）盲源分离和非线性混合（系统）盲源分离两大类。线性混合盲源分离又是由瞬时混合盲源分离和卷积混合盲源分离所构成的。在盲源分离理论中，具有普遍意义的是线性混合盲源分离，当今人们对于盲信号分离的研究也绝大部分是针对线性混合过程的。

瞬时混和盲源分离是线性混和盲源分离中最为基础的形式，其可表述为源信号 $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$ 经过传输系统（混合矩阵 $\mathbf{A}$ ），得到系统输出的观测信号 $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ ，即 $x_i(t)$ 在 t 时刻只依赖于原始信号在时间 t 上的数值，数学表达式为

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (1)$$

这里矩阵

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

可逆，它表示系统的混合矩阵。这里需要说明的是：由于“盲”，故与传统数学处理方法的不同，在式(1)中源信号和混合矩阵都是未知的，只有观测信号是已知的。

那么，盲源分离目的就是要找到一个线性变换 $\mathbf{W}$ （即分离矩阵），将它施加到观测信号 $\mathbf{x}(t)$ 上，从而得到源信号 $\mathbf{s}(t)$ 的估计 $\hat{\mathbf{s}}(t)$ 。这里的 $\mathbf{W}$ 从信号处理的角度可称之为空间滤波器，当 $\mathbf{W}$ 通过某种优化算法得到后，便有

$$\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{A}\mathbf{s}(t) \quad (2)$$

这里，输出信号 $\hat{\mathbf{s}}(t)$ 的每个分量都是源信号 $\mathbf{s}(t)$ 经矩阵 $\mathbf{W}\mathbf{A}$ 传递后的线性组合，而其组合系数即是 $\mathbf{W}\mathbf{A}$ 中对应行向量的各个元素。

2.2 线性系统辨识原理

线性时不变系统（见图 1），其中 $x(t)$ 是输出信号； $\boldsymbol{\theta} = [\theta_0, \theta_1, \dots, \theta_{p-1}]$ 表示未知系统的参数向量， p 为其向量长度（假设其已知）；输入信号向量 $\mathbf{u}_t = [u(t), u(t-1), \dots, u(t-p+1)]^T$ 为可观测数据； $v(t)$ 是系统输出端的加性随机噪声。

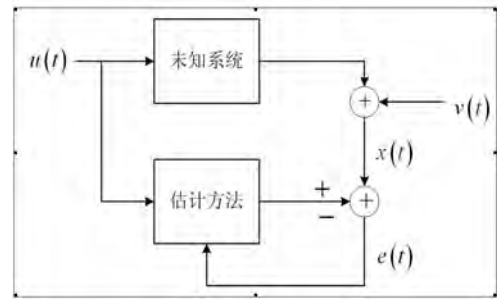


Figure 1. Curve: system result of standard experiment

图 1. 线性系统辨识的模型结构

则上述线性系统的数学模型可表示为

$$x(t) = \boldsymbol{\theta} \mathbf{u}_t + v(t) \quad (3)$$

那么，系统辨识^[4]就是在给定有限观测数据 $\{x(t), u(t)\}$ 的条件下，设计某一优化算法来获取未知系统参数 $\boldsymbol{\theta}$ 的估计^[4]。

3 基于盲源分离的声波气侵检测方法

3.1 声波气侵检测系统

在钻杆转动和泥浆循环过程中，可以通过气侵检测系统参数的变化来判断是否发生气侵^[2]。

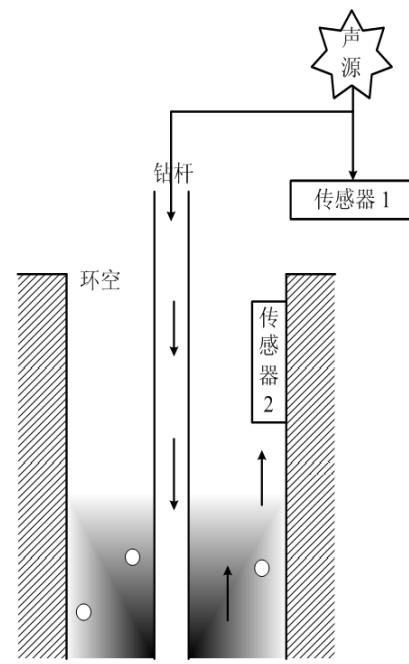


Figure 2. Acoustic Gas Influx Detection System

图 2. 井口声发射气侵检测系统的组成框图

图 2 为声波气侵检测系统的组成框图。声源旁的传感器 1 用于接收声源发射信号，视为系统输入。靠

近井口环空导管内壁的传感器 2 用于接收声源信号经钻杆和环空内泥浆的传播信号，声波信号通过钻井液反射被传感器 2 接收到的视为系统输出。系统中的其它信号则视为噪声。

根据图 2 所示的声波气侵检测系统框图，传感器 1 和传感器 2 接收到的信号可分别表示为

$$r_1(k) = s(k) + v_1(k) \quad (4)$$

$$r_2(k) = h(k) * s(k) + v_2(k) \quad (5)$$

其中， k 表示离散采样时间； $s(k)$ 为声源脉冲发射信号； $h(k)$ 为钻井液系统的脉冲响应； $v_1(k)$ 和 $v_2(k)$ 分别为传感器 1 和传感器 2 接收到的噪声信号。这里我们不对噪声信号的特性作任何要求，即不但允许噪声信号为文献[2]中设定的高斯信号，还可以允许其为非高斯信号，本文中不再对声波发生技术的原理和选择进行详细探讨。

3.2 检测模型设计

声源脉冲发射信号 $s(k)$ 通常选为伪二进制信号，并认为其与加性噪声 $v_2(k)$ 相互独立，现可将式(4)和式(5)所表示的声波气侵检测系统模型转化为矩阵的形式，如下

$$\mathbf{r}(k) = \mathbf{H} \cdot \mathbf{s}(k) + \mathbf{n}(k) \quad (6)$$

其中，传感器 1 和传感器 2 接收到的信号合并到一起作为观测信号 $\mathbf{r}(k) = [r_2(k), r_1(k), \dots, r_1(k-p+1)]^T$ ，传感器 2 接收到的噪声信号和声源信号合并到一起作为源信号 $\mathbf{s}(k) = [v_2(k), s(k), \dots, s(k-p+1)]^T$ ，而传感器 1 接收到的噪声信号 $\mathbf{n}(k) = [0, v_1(k), \dots, v_1(k-p+1)]^T$ 可看作为观测信号 $\mathbf{r}(k) = [r_2(k), r_1(k), \dots, r_1(k-p+1)]^T$ 的观测噪声，式(6)中混合矩阵 $\mathbf{H}$ 的表达式为

$$\mathbf{H} = \begin{bmatrix} 1 & h_0 & \cdots & h_{p-1} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (7)$$

则得到声波气侵检测系统的盲源分离模型，即式(6)，其中 p 表示系统阶数，一般可采用自相关函数等方法确定，实现较为简单，这里视其为已知条件。可以看出，式(6)给出的模型为标准的盲源分离的含噪瞬时混合模型。

3.3 算法流程

传统的系统辨识方法多对噪声环境存在一定的要求，如最小二乘 (LS, least square) 法需要假设加性噪声为白噪声，而最大似然 (ML, maximum likelihood) 法则需要在噪声统计特性已知的情况下，才能给出正确的辨识结果，这样的限制条件大大局限了系统辨识方法的适用范围。而由于盲源分离技术无需对信号源做过多的假设，因此可以考虑将背景噪声作为一个信号源（即可以是高斯分布的，也可以是非高斯分布的信号），利用盲源分离技术将其估计出来，然后从系统输出中减去已经估计出的噪声，进而得到无加性噪声干扰的观测数据，最后再利用常规的最小二乘法估计未知系统的参数。这样则无需对背景噪声特性作限定，放宽了文献[2]中处理对象的范围，又可以获得较为准确的估计结果。

由于声源脉冲发射信号 $s(k)$ 通常采用伪随机二进制信号，它是一种典型的独立同分布非高斯信号，因此信号 $[s(k), \dots, s(k-p+1)]^T$ 间满足相互独立且非高斯的条件。而加性噪声 $v_2(k)$ 则与信号向量 $\mathbf{s}(k)$ 中的各个分量间相互独立，并且即使其为高斯分布信号时，式(6)中的源信号 $\mathbf{s}(k) = [v_2(k), s(k), \dots, s(k-p+1)]^T$ 仍能满足盲源分离理论中最多只有一个信号为高斯分布的必要条件。另外，式(6)的 $\mathbf{H}$ 为满秩矩阵，混合矩阵不存在欠定的情况。因此，式(6)满足盲源分离技术中所有的基本假设条件，采用适当的具有抗噪特性的盲源分离算法^[5]，就很容易得到该瞬时混合模型的分离矩阵。

综上所述，本文方法可通过以下步骤来实现：第一步，根据瞬时混合盲源分离中的非高斯最大化算法，计算式(6)的分离矩阵 $\mathbf{W}$ ；第二步，搜寻分离矩阵 $\mathbf{W}$ 第一列中绝对值最大的值 $w_{i1}(k)$ ，确定 i 值；并计算噪声的真值估计 $\hat{v}_2(k) = y_i(k)/w_{i1}$ ；第三步，将估计的噪声从系统观测信号中减去，重新构造系统的输出信号 $\hat{r}_2(k) = r_2(k) - \hat{v}_2(k)$ ；第四步，利用数据 $\{\hat{r}_2(k), s(k)\}$ ，根据最小二乘法获取未知无噪线性系统的参数估计。

4 仿真实验

4.1 实验设计

仿真实验选择在两种不同类型的噪声（白噪声和有色噪声）背景下，将本文提出的方法（BL 法）与常用的最小二乘法（LS 法）进行对比。系统输入数据 $\{s(k)\}$ 采用 Bernoulli 序列，即 $s(k) = \pm 1$ ；未知系统参数 $\mathbf{H}$ 的真值为 $\mathbf{H} = [0.5 \quad -1.0 \quad 0.6]^T$ 。

本文仿真实验中的系统加性噪声 $v_2(k)$ 分别采用高斯白噪声和有色噪声信号。设定好实验系统的先验条件后，则要选择一性能指标参数作为检验方法有效性的依据。实验中采用均值误差 (ME) 作为定义算法全局性能的性能指标参数，其计算公式如下

$$ME = Mean \left\{ \frac{\|\hat{\theta} - \theta\|}{\|\theta\|} \right\} \quad (8)$$

其中 $\|\cdot\|$ 表示欧几里德距离 (Euclidean norm) ; $Mean(\cdot)$ 表示取均值, 仿真实验中对 50 次 Monte Carlo 实验的参数估计结果取平均值。

4.2 结果分析

首先给出在高斯白噪声背景下, 本文提出的 BL 法和 LS 法的性能对比结果。表 1 给出的是当信噪比 $SNR=10dB$ 时, 两种方法参数估计的结果, 该结果由估计参数的平均值和相对真值的标准偏移共同构成。从表 1 中的参数估计结果以及均值误差 (ME) 可以看出, BL 法和 LS 法受白噪声的影响均较小, 皆能给出准确的辨识结果且估计性能相当; 并且随着数据采样长度的增加, 两种方法的辨识精度都有所提高。

表 2 给出的是在色噪声背景下, 当信噪比分别为 $SNR = -5dB$ 时, BL 法和 LS 法参数辨识的结果。表中当加性噪声为较复杂的有色噪声时, LS 法的性能恶化较为严重, 其在给定有限观测数据的条件下给出的辨识结果与真值偏差较大。而本文方法的辨识性能则明显优于 LS 法, 因此通过 BL 法来对未知系统进行参数估计能更有效地抑制背景噪声的干扰, 且在较低信噪比情况下能够给出较为理想的辨识结果。

图 3 是 BL 法和 LS 法在色噪声背景下, 两种方法估计性能随 SNR 变化的曲线。从图中可以看出: 在色噪声背景下, BL 法的辨识误差要明显低于 LS 法, 即前者具有较强的辨识能力。

5 结论

针对加性噪声背景下线性系统的辨识问题, 本文将盲源分离技术与气侵检测的系统辨识理论相结合, 提出了一种声波气侵检测方法。该方法利用瞬时混合盲源分离无需对信源信号进行过多限制的技术优势, 将声波气侵检测系统模型纳入到盲源分离的瞬时混合模型中, 进而无需对系统加性噪声类型进行任何先验假定, 且在较低信噪比下仍具有较好的估计性能。

References (参考文献)

- [1] Bang J, Mjaaland S, Jensen L K, Hendriks P. Acoustic gas kick detection with wellhead sonar. Journal of petroleum technology, 1995, 47(2): 111-112.
- [2] 唐斌, 熊英. 井口声发射钻井液油气侵入检测——三阶统计的系统辨识方法. 仪器仪表学报, 2001, 22(3): 261-264.
- [3] Cichocki A, Amari S. Adaptive blind signal and image processing: learning algorithm and applications. New York: John Wiley & Sons, 2002, 1st edition.
- [4] 李鹏波, 胡德文等. 系统辨识基础. 北京: 中国水利水电出版社, 2006, 第 1 版.
- [5] Hyvärinen A, Oja E. A fast fixed-point algorithm for independent component analysis. Neural Computation, 1997, 19(7): 1483-1492.

Table 1. Performance comparison of BL and LS method

表 1. BL 法与 LS 法的性能比较
(高斯白噪声背景下, $SNR=10dB$)

采样值	$N = 500$		$N = 1000$	
算法	BL 法	LS 法	BL 法	LS 法
$\theta_1 = 0.5$	0.5034 ± 0.0145	0.5033 ± 0.0146	0.4995 ± 0.0102	0.4995 ± 0.0103
$\theta_2 = -1.0$	0.9967 ± 0.0124	-0.9969 ± 0.0123	-0.9996 ± 0.0101	-1.0090 ± 0.0102
$\theta_3 = 0.6$	0.6003 ± 0.0185	0.6002 ± 0.0187	0.6006 ± 0.0113	0.6005 ± 0.0118
ME	0.0247	0.0246	0.0164	0.0169

Table 2. Performance comparison of BL and LS method

表 2 BL 法与 LS 法的性能比较
(色噪声背景下, $SNR=-5dB$)

采样值	$N = 200$		$N = 400$	
算法	BL 法	LS 法	BL 法	LS 法
$\theta_1 = 0.5$	0.5095 ± 0.0310	0.4538 ± 0.1424	0.5013 ± 0.0169	0.5356 ± 0.1048
$\theta_2 = -1.0$	-0.9976 ± 0.0355	-0.0490 ± 0.0913	-1.0021 ± 0.0253	-0.9621 ± 0.0844
$\theta_3 = 0.6$	0.6029 ± 0.0217	0.5549 ± 0.1413	0.5987 ± 0.0215	0.6354 ± 0.1020
ME	0.0498	0.1948	0.0323	0.1424

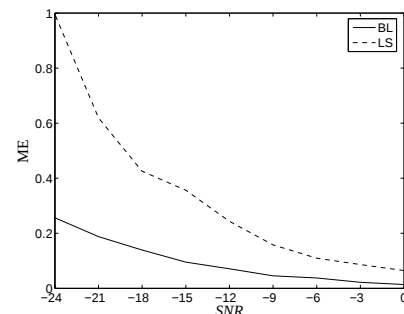


Figure 3. Performance comparison of BL and LS by SNR

图 3. 在色噪声背景下, BL 法与 LS 法性能随 SNR 变化的曲线