

引用格式:褚鼎立,陈红,王旭光.基于WOA和PCA的盲信号提取算法[J].电光与控制,2019,26(7):36-39. CHU D L, CHEN H, WANG X G. A blind signal extraction algorithm based on WOA and PCA[J]. Electronics Optics & Control, 2019, 26(7):36-39.

## 基于 WOA 和 PCA 的盲信号提取算法

褚鼎立, 陈红, 王旭光  
(国防科技大学电子对抗学院, 合肥 230037)

**摘要:** 针对线性混叠信号中独立源的盲提取问题,提出一种基于鲸鱼优化和主成分分析的有序盲提取算法。该算法以信号峰度绝对值大小作为目标函数,利用鲸鱼优化算法计算抽取向量,同时利用主成分分析对混合信号降维,去除混合信号中存在的关于已提取信号的成分,从而按照源信号峰度顺序进行盲提取。仿真实验表明,提出的算法对盲信号提取是有效的,而且提取精度是较高的,有一定的实际应用价值。

**关键词:** 信号处理; 盲信号提取; 鲸鱼优化算法; 主成分分析

中图分类号: TN911.72 文献标志码: A doi:10.3969/j.issn.1671-637X.2019.07.007

## A Blind Signal Extraction Algorithm Based on WOA and PCA

CHU Ding-li, CHEN Hong, WANG Xu-guang

(Institute of Electronic Countermeasure, National University of Defense Technology, Hefei 230037, China)

**Abstract:** Aiming at the blind extraction of independent sources in linear aliasing signals, this paper proposes an ordered blind extraction algorithm based on Whale Optimization Algorithm (WOA) and Principal Component Analysis (PCA). The algorithm takes the absolute value of the signal kurtosis as the objective function, uses the WOA to calculate the decimation vector, and PCA to reduce the dimension of the mixed signal for removing the components of the extracted signal existing in the mixed signal. Thus the blind extraction is implemented in the order of the source signal kurtosis. Simulation experiments show that the proposed algorithm is effective for extracting blind signals, and the extraction accuracy is high, which has certain practical application value.

**Key words:** signal processing; blind signal extraction; Whale Optimization Algorithm (WOA); Principal Component Analysis (PCA)

### 0 引言

盲信号处理是当前信号处理领域研究的重点。仅从观测到的混沌信号中恢复出所有源信号,这一过程称为盲源分离(Blind Source Separation, BSS)。目前,针对盲源分离问题已经有许多成熟的算法,如基于非负矩阵分解的算法<sup>[1]</sup>、基于信号序列结构预测的算法<sup>[2]</sup>、基于差分进化的算法<sup>[3]</sup>、非线性代价函数<sup>[4]</sup>以及基于智能优化算法的盲源分离<sup>[5-7]</sup>。盲源提取(Blind Source Extraction, BSE)是盲源分离技术的一种扩展<sup>[8]</sup>,目的是从观测到的混沌信号中恢复出一个或多个源信号,使用更加灵

活。对于盲源提取,各国的学者也做了大量的工作,如CICHOCKI<sup>[9]</sup>, LI<sup>[10]</sup>等早期对源信号的盲提取进行了研究工作,提出了一些盲提取算法。近年来,也有诸多学者在混沌信号的盲提取上进行研究,如陈寿齐<sup>[11]</sup>等提出基于峭度的稳健特定信号盲提取算法,计算量相对于内点惩罚函数法、约束尺度法等要小一些;李勇<sup>[12]</sup>等提出一种基于二阶统计量的去除干扰波的盲提取算法,并应用于地震数据的处理;任婕<sup>[13]</sup>等利用线性预测模型描述信号的时序结构,提出一种基于高阶统计量的自回归参数估计的盲信号提取算法,在不含噪和含高斯白噪声的情况下都有良好的提取效果。随着智能优化算法的发展,越来越多的学者将智能优化运用到盲提取中,如陈雷<sup>[14]</sup>等提出了基于细菌觅食优化的盲信号提取算法,通过多次提取和去相关消源过程实现对混沌信号的盲提取,但未对算法的收敛速度进行分析;李广明<sup>[15]</sup>等将人工蜂群算法应用于混沌信号的盲提取,提出了一个低复杂度的目标函数,利用马尔可夫模型分析了算法的有效

收稿日期:2018-08-16

修回日期:2018-08-31

基金项目:国家自然科学基金(61571446);安徽省自然科学基金(KY13C152)

作者简介:褚鼎立(1993—),男,山东烟台人,硕士生,研究方向为盲源分离。

性,并对算法进行理论分析。

针对混叠信号中微弱信号的盲提取问题,本文提出一种基于鲸鱼优化算法(WOA)和主成分分析(PCA)的盲提取算法。鲸鱼优化算法(WOA)全局寻优能力更强,本文将其用于解决盲提取问题,算法使用峰度作为目标函数,结合PCA降维,可将混叠信号按源信号的峰度顺序逐个提取,在保证分离精度的同时,计算复杂度较低,鲁棒性更好。

## 1 问题描述

假设  $n$  个零均值且彼此相互独立的源信号  $s(t) = (s_1(t) \ s_2(t) \ \dots \ s_n(t))^T$ , 线性瞬时混叠后,得到  $m$  个观测信号  $x(t) = (x_1(t) \ x_2(t) \ \dots \ x_m(t))^T$ , 即

$$x(t) = As(t) \quad (1)$$

式中:  $A \in \mathbf{R}^{m \times n}$  为混合矩阵,  $A$  为列满秩矩阵;  $t$  为时域采样点。盲源提取的思想是仅根据  $s(t)$  的统计独立假设条件,从观测信号  $x(t)$  中逐次分离出源信号  $s(t)$ , 分离过程表示为

$$y(t) = w^T x(t) = w^T As(t) \quad (2)$$

式中:  $y(t)$  为从观测信号  $x(t)$  中分离出的一个源信号; 在  $m \geq n$  时, 抽取向量  $w$  为  $m$  维的列向量, 令  $g = w^T A$ , 此时  $g$  为  $n$  维的行向量, 当只有一个元素为 1, 其余元素为 0 时,  $y(t)$  为  $s(t)$  中对应  $g$  中 1 的位置的源信号的估计, 通过  $n$  次抽取, 可得所有源信号的估计。

传统方法通常会进行  $n$  次抽取, 按照一定准则逐次抽取对源信号的估计, 本文利用鲸鱼优化算法(WOA)对分离准则的目标函数进行寻优, 同时利用主成分分析对抽取一个源信号后的观测信号进行降维, 最终只需进行  $n-1$  次抽取, 简化了计算复杂度。

## 2 基于鲸鱼优化和 PCA 的微弱信号盲提取

### 2.1 鲸鱼优化算法

鲸鱼优化算法(WOA)是 MIRJALILI<sup>[16]</sup> 等于 2016 年提出的新型智能优化算法, 该算法模仿鲸鱼捕食行为: 鲸鱼通过包围猎物、发泡网攻击和搜索捕食 3 种方式进行位置更新。

#### 2.1.1 包围猎物

WOA 算法假设当前最优解是猎物位置或者最接近猎物的鲸鱼位置, 在定义最佳鲸鱼个体后, 其他鲸鱼向当前最优位置靠近, 位置更新的方式为

$$X(j+1) = X^*(j) - A_1 \cdot D \quad (3)$$

$$D = |C \cdot X^*(j) - X(j)| \quad (4)$$

式中:  $A_1$  和  $C$  为系数变量;  $D$  是距离向量;  $X^*(j)$  为记录的最佳鲸鱼位置(最优个体);  $X(j)$  表示当前鲸鱼位置;  $j$  表示迭代次数。

系数变量的算式如下

$$A_1 = 2a \times r - a \quad (5)$$

$$C = 2r \quad (6)$$

式中:  $a$  取值范围在 0 到 2 之间, 以  $a = 2 - \frac{2j}{M}$  的方式更新,  $M$  为最大迭代次数;  $r$  为  $[0, 1]$  内的随机数。

#### 2.1.2 发泡网攻击

本文是模拟座头鲸攻击猎物的行为, 数学模型如下。

1) 收缩包围: 即减少式(5)中的  $a$  值。在这个过程中,  $A_1$  的取值范围是  $[-a, a]$ , 一旦这个范围发生变动,  $a$  也会相应地改变, 若  $A_1$  在  $[-1, 1]$  范围内, 鲸鱼的目标位置将确定在当前位置和猎物位置之间, 这表现出座头鲸的局部搜索能力。更新方式按式(3)。

2) 螺旋式位置更新: 鲸鱼以螺旋式的游走方式向当前最优个体靠近, 将螺旋式数学模型引入, 建模如下

$$X(j+1) = D' e^{bl} \cos(2\pi l) + X^*(j) \quad (7)$$

$$D' = |X_j^* - X_j| \quad (8)$$

式中:  $D'$  表示第  $i$  头鲸当前位置和最优个体之间的距离;  $b$  为对数螺旋形状常数;  $l$  是  $[-1, 1]$  内的随机数。

座头鲸在围捕猎物时不仅以螺旋形状游向猎物, 还要收缩包围圈, 以 50% 的概率作为阈值来更新鲸鱼位置。算式为

$$X(j+1) = \begin{cases} X^*(j) - A_1 \cdot D & p < 0.5 \\ D' e^{bl} \cos(2\pi l) + X^*(j) & p \geq 0.5 \end{cases} \quad (9)$$

式中:  $p$  为  $[0, 1]$  内的随机数。

#### 2.1.3 搜索捕食

鲸鱼在进行发泡网攻击时, 还有一定几率进行搜索捕食。搜索捕食是以系数  $A_1$  变化为基础的, 若  $A_1$  超出  $[-1, 1]$  范围, 距离数据  $D$  随机更新, 此时, 鲸鱼为了寻找猎物, 会偏离原来的目标猎物, 使算法具有一定的全局搜索能力。算式为

$$D = |C \cdot X_{\text{rand}} - X(j)| \quad (10)$$

$$X(j+1) = X_{\text{rand}} - A_1 \cdot D \quad (11)$$

需要注意的是, 这里的  $X_{\text{rand}}$  不是随机的位置, 而是当前种群中随机一个鲸鱼所在的位置。

鲸鱼优化算法(WOA)通过上述 3 种位置更新方式进行寻优, 其中式(3)和式(7)进行局部寻优操作, 式(11)扩展搜索范围进行全局寻优, 在收敛性和寻优精度上较以往算法有了一定提升。

## 2.2 主成分分析降维

主成分分析(PCA)是由卡尔·皮尔逊于 1901 年提出的, 用于分析数据及建立数理模型<sup>[17]</sup>。PCA 在尽可能减少信息损失的情况下对数据进行降维, 同时把数据中对方差贡献最大的特征保留下来。该方法对数据集的协方差矩阵进行处理, 特征值分解后, 去除小特

征值后,进行投影空间重构,达到降维的目的。

对于给定的数据集  $X_{m \times n}$ ,用 PCA 对  $X_{m \times n}$  进行降维处理,具体步骤如下。

1) 计算样本均值。

$$m_1 = \frac{1}{n} \sum_{k=1}^n x_k \quad (12)$$

式中:  $x_k$  为  $X_{m \times n}$  中第  $k$  列的向量;  $m_1$  为  $X_{m \times n}$  每行的均值向量。

2) 计算  $X_{m \times n}$  的协方差矩阵。

$$S = \sum_{k=1}^n (x_k - m_1)(x_k - m_1)^T \quad (13)$$

3) 计算协方差矩阵  $S$  的特征值  $\lambda$  和特征向量  $e$ 。

$\lambda = (\lambda_1 \ \lambda_2 \ \dots \ \lambda_m)$ , 其中  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ ;  $e = (e_1 \ e_2 \ \dots \ e_m)$ , 特征向量  $e$  与特征值  $\lambda$  一一对应。

4) 新的投影空间。

取前  $p$  ( $p < m$ ) 个最大特征值对应的特征向量构成新的投影矩阵  $W_{m \times p}$ , 将原数据集投影到新的空间中可得  $X_{\text{new}} = W^T \cdot X_{m \times n}$ , 此时  $X_{\text{new}}$  为降维后的数据集, 维度为  $p$  ( $p < m$ )。

### 2.3 盲提取原理及步骤

根据式(2), 确定了抽取向量  $w$  便能从观测信号中提取源信号。这里, 利用信号峰度作为抽取的度量。  $y(t)$  的峰度为  $K(y) \triangleq \text{cum}(y, y, y, y)$ , 当  $K(y)$  为正时,  $y(t)$  服从超高斯分布; 当  $K(y)$  为负时,  $y(t)$  服从亚高斯分布; 当  $K(y)$  为零时,  $y(t)$  服从高斯分布。一般设  $E(s(t)s(t)^T) = I$  和  $z(t) = Bx(t)$ ,  $B$  为白化矩阵, 白化后式(2)可改为  $y(t) = w^T z(t)$ , 即  $E(z(t)z(t)^T) = I$ , 此时  $E(wz(t)z(t)^T w^T) = wE(z(t)z(t)^T)w^T = 1$ , 而  $E(z(t)z(t)^T) = 1$ , 可知  $\|w\|^2 = 1$ , 所以可得抽取向量  $w$  的代价函数为  $w = \arg \max_w |K(y)|$ 。

抽取向量  $w$  的参数作为鲸鱼优化算法(WOA)的待优化参数, 在约束项  $\|w\|^2 = 1$  的限制下, 对  $w$  进行霍夫变换可得各元素, 即

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{m-1} \\ w_m \end{pmatrix} = \begin{pmatrix} \cos \alpha_1 \\ \sin \alpha_1 \cos \alpha_1 \\ \vdots \\ \sin \alpha_1 \sin \alpha_2 \cdots \cos \alpha_{m-1} \\ \sin \alpha_1 \sin \alpha_2 \cdots \sin \alpha_{m-1} \end{pmatrix} \quad (14)$$

式中:  $\alpha_i \in [0, 2\pi]$ 。霍夫变换后, 待求解的参数减少一个, 参数的搜索空间减少从而减少了寻优时间。

鲸鱼优化算法(WOA)根据峰度绝对值最大准则抽取源信号, 每提取一个源信号, 需要对原始观测信号进行退化和降维, 退化过程如下

$$x_{\text{new}}(t) = x_{\text{old}}(t) - hy(t) \quad (15)$$

式中:  $h = E(x_{\text{old}}(t)y(t))/E(y^2(t))$ , 向量  $h$  可以使

$E(x_{\text{old}}(t) - hy(t))^2$  最小<sup>[18]</sup>;  $x_{\text{new}}(t)$  为退化后的信号,  $x_{\text{old}}(t)$  为退化前的信号。退化处理后需要利用 PCA 对退化后的信号进行降维, 降维的步骤见 2.2 节。

基于鲸鱼优化和 PCA 的盲提取步骤如下。

1) 对观测信号中心化、预白化。

2) WOA 参数初始化。抽取向量  $w$  中各元素作为鲸鱼的位置信息。

3) 利用 WOA 进行寻优, 根据峰度绝对值最大(或者最小)准则得到最佳抽取向量  $w$ 。

4) 利用式(2)抽取一个源信号。

5) 利用式(13)对当前信号  $x_{\text{old}}(t)$  进行退化处理, 得到不包含已提取信号成分的退化信号  $x_{\text{new}}(t)$ 。

6) 利用 PCA 对退化后的信号  $x_{\text{new}}(t)$  做降维处理, 得到降维后的信号  $x'_{\text{new}}(t)$ 。

7) 若  $x'_{\text{new}}(t)$  维度大于 1, 跳至步骤 1), 否则结束。

以上步骤是将观测信号中的源信号按照峰度绝对值的大小顺序逐个提取出, 当实际情况只需要提取峰度最大或峰度最小的单个源信号时, 将步骤进行到 4) 就可得到。

### 3 仿真验证

本节基于 Matlab2014a 平台仿真验证并分析本文算法的有效性。

实验中所采用的仿真信号如下:  $s_1 = \sin[2\pi \cdot 2t/(fs)] \cdot \cos[2\pi \cdot 50t/(fs)]$ ;  $s_2 = 0.5 \cdot \sin[2\pi \cdot 9t/(fs)] \cdot \sin[2\pi \cdot 100t/(fs)]$ ;  $s_3 = \sin[2\pi \cdot 50t/(fs)]$ 。

采样点数为 5000, 采样频率为 5000 Hz, 三路信号的峰度分别为  $-0.0469$ ,  $-0.0029$ ,  $-0.3750$ 。混叠矩阵为

$$A = \begin{pmatrix} 1.6 & 1 & 1 \\ 1 & 1.8 & 1 \\ 1 & 1 & 2 \end{pmatrix}。$$

仿真实验中, 鲸鱼优化算法(WOA)的参数设置: 鲸鱼种群数为 30, 寻优次数为 100。实验仿真结果见图 1 ~ 图 3。

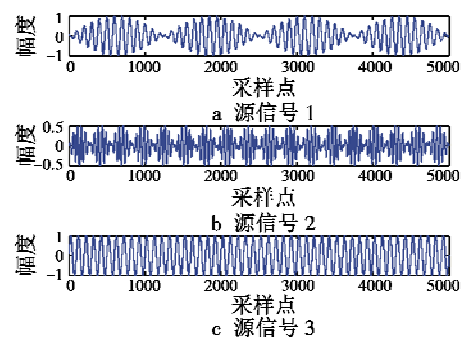


图1 源信号

Fig. 1 Source signals

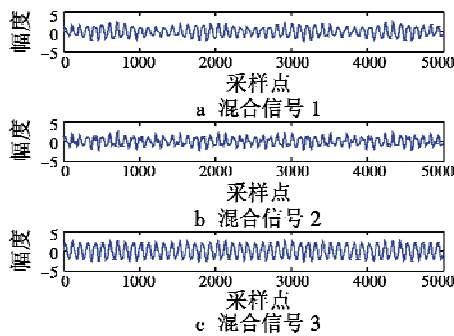


图2 混合信号

Fig. 2 Mixed signals

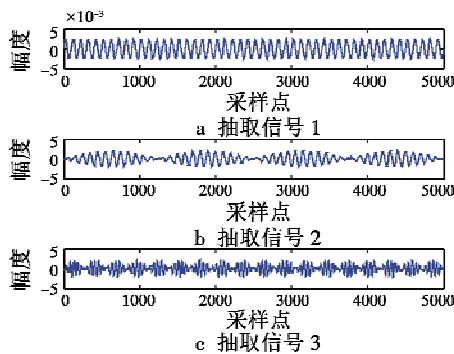


图3 抽取信号

Fig. 3 Extracted signals

图1为仿真生成的源信号波形,图2为经混叠矩阵后的混合信号,图3为按照峰度大小顺序抽取出的源信号。通过比较图1和图3的波形可以看出,本文算法能够有效抽取所需的源信号,为了更客观地评价本文算法的抽取效果,引入抽取信号与源信号的相似系数 $\zeta_{ij}$ 作为评价指标,相似系数 $\zeta_{ij}$ 定义为

$$\zeta_{ij} = \zeta(s_i, y_j) = \frac{\left| \sum_{t=1}^n s_i(t)y_j(t) \right|}{\sqrt{\sum_{t=1}^n s_i^2(t) \sum_{t=1}^n y_j^2(t)}} \quad (16)$$

式中: $n$ 表示采样点数目。当 $y_j$ 与 $s_i$ 的相关性越强时, $\zeta_{ij}$ 越接近1,即若由相似系数 $\zeta_{ij}$ 构成的相似系数矩阵 $\zeta$ 每行每列有且仅有一个元素接近1,其他都接近0时,抽取效果(分离效果)最佳。经仿真计算,实验分离信号与源信号

的相似系数矩阵为 $\zeta = \begin{pmatrix} 0.0525 & 0.9934 & 0.1017 \\ 0.1268 & 0.1077 & 0.9861 \\ 0.9905 & 0.0389 & 0.1316 \end{pmatrix}$ ,可

见已经将源信号很好地抽取出来,抽取精度也是非常高的,说明了算法的有效性。

#### 4 结论

在复杂的电磁环境中准确提取出有用信号具有重要的现实意义和研究前景。本文针对盲信号的抽取问题,提出了一种基于鲸鱼优化和PCA的盲提取算法,

以信号峰度绝对值大小为序,利用鲸鱼优化算法(WOA)从混叠信号中抽取源信号,并通过PCA降维去除混合信号中存在的已抽取的源信号成分,最终逐个抽取服从任意分布的源信号。

#### 参考文献

- [1] 魏乐. 基于非负矩阵分解算法进行盲信号分离[J]. 电光与控制, 2004, 11(2): 38-41.
- [2] 陈雷, 张立毅, 郭艳菊, 等. 基于时间可预测性的差分搜索盲信号分离算法[J]. 通信学报, 2014, 35(6): 117-125.
- [3] 张朝柱, 张健沛, 孙晓东. 基于自适应粒子群优化的盲源分离[J]. 系统工程与电子技术, 2009, 31(6): 1275-1278.
- [4] CRUCES-ALVAREZ S A, CICHOCKI A, AMARI S. From blind signal extraction to blind instantaneous signal separation: criteria, algorithms, and stability [J]. IEEE Transactions on Neural Networks, 2004, 15(4): 859-873.
- [5] 刘俊豪. 基于粒子群算法和鱼群算法的盲源分离的研究[D]. 太原: 太原理工大学, 2006.
- [6] 张银雪, 田学民, 邓晓刚. 基于改进人工蜂群算法的盲源分离方法[J]. 电子学报, 2012, 40(10): 2026-2030.
- [7] 栾先冬. 基于萤火虫算法的语音信号盲源分离研究[D]. 兰州: 兰州交通大学, 2016.
- [8] SHI Z, ZHANG H, JIANG Z. Hybrid linear and nonlinear complexity pursuit for blind source separation [J]. Journal of Computational & Applied Mathematics, 2012, 236(14): 3434-3444.
- [9] CICHOCKI A, THAWONMAS R, AMARI S. Sequential blind signal extraction in order specified by stochastic properties [J]. Electronics Letters, 1997, 33(1): 64-65.
- [10] LI Y, WANG J, ZURADA J M. Blind extraction of singularly mixed source signals [J]. IEEE Transactions on Neural Networks, 2000, 11(6): 1413-1422.
- [11] 陈寿齐, 沈越泓, 马明, 等. 基于峭度的稳健特定信号盲提取[J]. 系统仿真学报, 2008, 20(22): 6097-6099.
- [12] 李勇, 陈明, 张建跃. 基于二阶统计量的盲提取算法[J]. 石油地球物理勘探, 2011, 46(2): 272-274.
- [13] 任婕, 朱立东. 基于自回归高阶估计的盲信号提取算法[J]. 空间电子技术, 2012, 9(3): 5-8.
- [14] 陈雷, 张立毅, 郭艳菊, 等. 基于细菌觅食优化的盲信号提取算法[J]. 计算机应用研究, 2012, 29(2): 451-454.
- [15] 李广明, 胡志辉. 基于人工蜂群算法的混沌信号盲提取[J]. 物理学报, 2016, 65(23): 44-51.
- [16] MIRJALILI S, LEWIS A. The whale optimization algorithm [J]. Advances in Engineering Software, 2016, 95: 51-67.
- [17] 李远博, 曹蕊. 基于PCA降维的协同过滤推荐算法[J]. 计算机技术与发展, 2016, 26(2): 26-30.
- [18] 王荣杰, 周海峰, 詹宜巨. 一种基于人工蜂群优化的有序盲源抽取方法[J]. 电信科学, 2012, 28(5): 43-48.