

基于特定听觉特征的盲信号提取方法

王心怡

(昆明船舶设备研究试验中心, 云南昆明 650051)

摘要: 针对参考信号波形难以构造等情况, 提出一种利用期望源信号听觉特征进行信号盲提取的方法, 根据源信号的听觉特征先验知识, 将提取信号与期望信号听觉特征向量的接近程度作为度量, 使用非线性最小二乘法进行迭代计算分离向量, 提取出与期望信号在听觉特征上最为接近的源信号。对这种方法进行了理论推导, 并使用不同类型的信号进行了仿真。仿真结果表明, 该方法具有良好的效果, 且由于仅利用了信号的听觉特征, 而不需要参考信号的具体波形, 对于不同类型的信号, 具有广泛的适应性, 具有良好的应用价值。

关键词: 盲源分离; 盲源提取; 听觉特征; 非线性最小二乘

中图分类号: TB566

文献标识码: A

文章编号: 1000-3630(2018)-05-0496-05

DOI 编码: 10.16300/j.cnki.1000-3630.2018.05.015

Blind signal extraction based on specific auditory feature

WANG Xin-yi

(Kunming Shipbuilding Equipment Research and Test Center, Kunming 650051, Yunnan, China)

Abstract: In the case that the reference signal waveform is difficult to provide, a blind extraction method based on auditory feature is proposed in this paper. According to the known auditory feature, the accessibility of auditory features between expected signal and extracted signal is treated as a measure, the nonlinear least square algorithm is used to calculate the separation vector in order to extract the source signal that is closest to the expected signal in auditory feature. The theory of the method is described in this paper, and the simulation is also performed with different kinds of signals. The simulation result shows the effectiveness of the method, which needs no waveform but only the auditory feature of reference signal. And, the method is adaptive to different kinds of signals and has a good performance in practical applications.

Key words: blind source separation; blind source extraction; auditory feature; nonlinear least square

0 引言

近年来, 盲源分离广泛应用于生物医学、无线通信、图像处理、阵列信号处理等领域^[1-3]。在实际应用中, 感兴趣的信号往往只是某一个或几个源信号, 将特定的源信号提取出来, 就称为盲提取。文献[4]提出了一种将特定频带信号提取出来的盲提取方法。文献[5]提出了首先将具有最大峭度的信号提取出来的盲提取算法。文献[6-7]提出了一种针对特定峭度信号的盲提取方法, 它可以将峭度在某一范围内的源信号提取出来。另一类盲提取方法是参考独立分量分析(Independent Component Analysis-Reference, ICA-R)方法, 将信号的先验知识以参考信号的形式加以利用, 只提取出期望的源信号。在 ICA-R 中, 通过引入一个接近性度量函数, 将传

统独立分量分析(Independent Component Analysis, ICA)问题转化为一个新的约束最优化问题, 接近性度量函数常用的有分离输出信号与参考信号的均方误差, 或分离输出信号与参考信号的瞬时相关性, 使用这些接近性度量, 通常需要根据源信号的波形构建参考信号, 如源信号的符号函数或同频正弦信号等。

在很多情况下, 从混合信号中提取特定的期望源信号存在很大的困难, 主要原因有两个方面: 一是难以找到一种合适的度量准则。例如, 多个源信号的频谱有很大重叠, 峭度也很接近, 在提取期望的源信号时难以确定一个合适的峭度范围; 二是难以根据期望的源信号构建参考信号, 限制了 ICA-R 等方法的使用, 例如很多时候只能得到源信号的某些统计特性, 而没有源信号的波形样本, 不具备源信号波形有关的先验知识, 无法构建参考信号; 并且有些源信号具有较强的时变性, 已有源信号的波形先验知识与混合信号中的该源信号接近性很差, 也无法准确估计出该源信号。

收稿日期: 2017-06-15; 修回日期: 2017-08-20

作者简介: 王心怡(1983-), 男, 内蒙古锡林浩特人, 硕士, 高级工程师, 研究方向为声学信号处理。

通讯作者: 王心怡, E-mail: wxy345135@163.com

本文提出一种基于源信号听觉特征的盲提取方法，将提取信号与期望源信号听觉特征的接近程度作为度量准则，提取出与期望信号在听觉特征上最接近的源信号，这种方法不需要构建参考信号，也不需要源信号波形的先验知识，对源信号的频段、峭度等也没有限制，只要源信号的听觉特征是稳定的，就能够以较好的效果提取出期望的源信号。

1 信号的听觉特征

信号的听觉特征是保证正确提取源信号的重要因素。听觉特征有很多种，用于信号盲提取的听觉特征应具有以下特点：(1) 具有很高的区分能力，能够充分体现不同的源信号之间的差异，而对同一个源信号在状态变化时保持相对稳定；(2) 特征参数在受到环境噪声影响和信道变化时能够保持稳健；(3) 易于计算等。

对于声学信号，可以使用的特征包括线谱、线性预测系数(Linear Prediction Coefficient, LPC)、Mel 频率倒谱系数(Mel-Frequency Cepstrum Coefficient, MFCC)等。根据研究，使用 MFCC 作为特征参数，在不同环境下可以保持较好的稳定性^[8-10]，对于不同的源信号具有较好的区分能力，并且对于同一个源信号，MFCC 在信号的基频、相位等参数发生较大变化时仍保持稳定，因此，本文采用 MFCC 作为信号的听觉特征。

MFCC 模拟人耳的听觉机理，人耳在嘈杂的环境中，以及各种变异情况下仍能正常地分辨出各种语音，具有很高的识别正确率，其关键因素之一就是耳蜗的作用。耳蜗实质上相当于一个滤波器组，其滤波作用是在对数频率尺度上进行的，在 1 000 Hz 以下为线性尺度，而 1 000 Hz 以上为对数尺度，这就使得人耳对低频信号比对高频信号更敏感，因此可以根据这个特性，构造一组类似耳蜗的滤波器来模仿人耳对声音的特征提取方式。

MFCC 参数的计算过程是将信号通过临界频带滤波器组，然后转换到倒谱域，流程如图 1 所示。

首先对信号进行预处理，包括分帧和加窗，分帧是将信号分成短的时段，每一帧的信号可以看作是平稳的，并对每一帧信号加窗，常用的是 Hamming 窗。然后对每一帧加窗信号作傅里叶变换(Fast Fourier Transformation, FFT)，并求平方，得到的能量谱经过 Mel 滤波器组，得到一组系数 X_j ，再对 X_j 求常用对数，得到相应频带的对数能量谱，再进行离散余弦变换(Discrete Cosine Transform, DCT)，得到倒谱域的 p 个 MFCC 系数 m_i ， $i=1, 2,$

$$\dots, p, \text{ 即}^{[8]}$$

$$m_i = \sum_{j=1}^M \lg(X_j) \cos\left[\frac{(j-0.5)\pi i}{M}\right] \quad (1)$$

其中， M 为 Mel 滤波器组的滤波器个数。

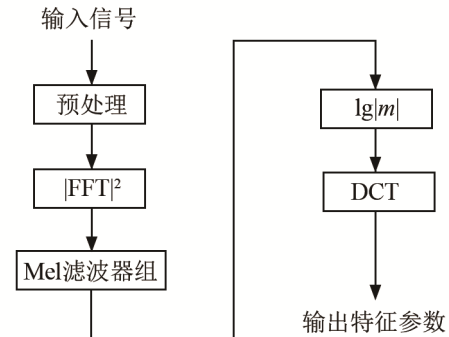


图 1 MFCC 参数提取流程
Fig.1 Flow chart of MFCC parameter extraction

2 基于听觉特征的盲提取

设有 n 个源信号，观测到的 k 维混合信号为源信号的瞬时线性混合，即

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

其中， \mathbf{s} 为 n 维源信号， \mathbf{A} 为 $k \times n$ 维混合矩阵， \mathbf{x} 为 k 维观测信号，本文讨论 $k=n$ 的情况。

提取的信号为 $\mathbf{y} = \mathbf{w}^T \mathbf{x}$ ，提取出与期望信号在听觉特征上最接近的源信号，等价于最小化代价函数 J ：

$$J(\mathbf{w}) = E\{(\mathbf{M}(\mathbf{w}^T \mathbf{x}) - \mathbf{M}_r)^2\} \quad (3)$$

其中， $\mathbf{M}(\mathbf{y})$ 为输出信号 \mathbf{y} 的 MFCC 参数向量， \mathbf{M}_r 为期望的源信号的 MFCC 参数向量先验值，是通过已有的期望源信号样本获得的。

$E\{\mathbf{h}^2\}$ 作为向量 \mathbf{h} 的度量，通常采用 $E\{\mathbf{h}^2\} = \|\mathbf{h}\|_2^2$ 。求解式(3)的极小值，是一个非线性最优化问题，现有的多种迭代方法，如 Gauss-Newton 法、Levenberg-Marquardt 法^[11]等，都可以用来求解，本文使用 Gauss-Newton 法，迭代求解使代价函数 J 最小的分离向量 \mathbf{w} ，从而得到听觉特征与先验 MFCC 最接近的提取信号 $\mathbf{y} = \mathbf{w}^T \mathbf{x}$ 。

3 仿真实验验证

仿真使用 3 个源信号，分别是频率为 500 Hz 的方波信号、谐波成分为 150、350、550、750、950、1 150、1 350 Hz 的多频正弦信号和粉红噪声信号，3 个源信号的波形如图 2 所示，仿真信号采样率为 10 kHz，横坐标为采样点数，纵坐标为波形幅度。将 3 个源信号：方波、多频正弦信号和粉红噪声信

号按照混合矩阵:

$$A = \begin{bmatrix} 1.3 & 0.52 & 0.75 \\ 0.85 & 0.9 & -0.6 \\ -0.83 & -1.05 & -0.32 \end{bmatrix}$$

进行加权得到的 3 通道混合信号, 其波形图如图 3 所示。

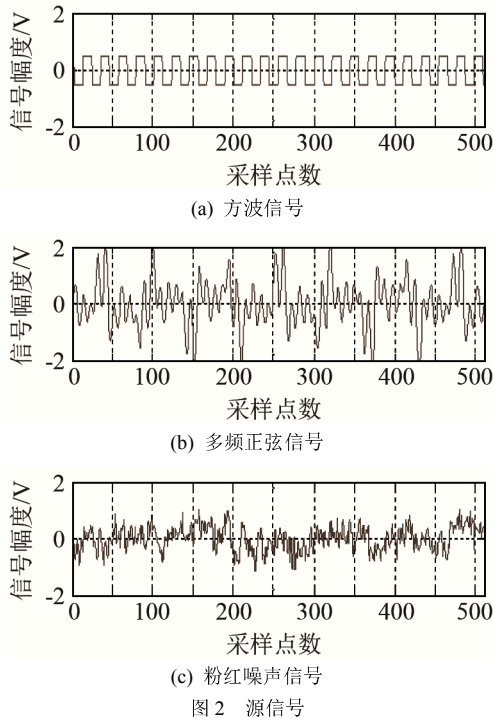


Fig.2 Source signals: 1. square wave signal, 2. multi-frequency sinusoidal signal, 3. pink noise signal

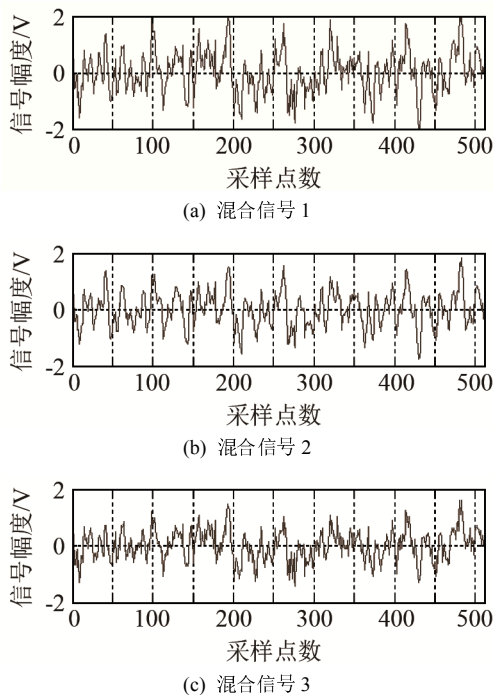


Fig.3 Mixed signals corresponding to Fig.2

使用本文方法对源信号 2 进行提取, 听觉特征

MFCC 先验值是采用与源信号 2 相同的谐波能量比构造的多频正弦信号, 取多个采样时段计算, 再求平均值。这个统计的先验 MFCC 向量与任意一段符合该谐波能量比的信号都具有较强的相似性。

为了评价算法性能, 采用提取出的信号 y 与真实的期望源信号 s_i 的相关系数作为度量

$$\rho = \frac{|E\{y^T s_i\}|^2}{\sqrt{E\{|y|^2\}E\{|s_i|^2\}}} \quad (4)$$

式(4)中, ρ 的值越接近 1, 表示提取出的信号与真实的源信号越相似。

真实源信号和使用本文方法提取出的信号如图 4 所示, 提取的源信号与真实源信号的相关系数为 0.979 9。

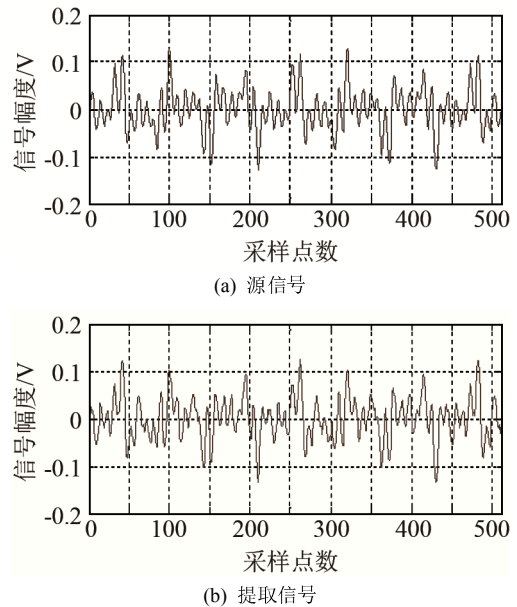


Fig.4 Extraction of the expected source signal

图 5 是提取信号的听觉特征 MFCC 向量与先验特征的比较。由图 5 可以看出, 提取信号在听觉特征上与先验值很相似, 计算出的相关系数为 0.923 6。

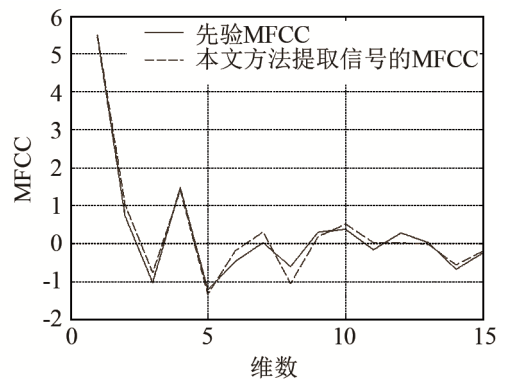


Fig.5 Auditory feature of the extracted source signal

对峭度相近的信号进行仿真，3 个源信号分别为两种不同乐器的独奏音乐信号及汽车引擎声音信号，如图 6 所示，源信号的归一化峭度值分别为 -0.343、-0.146、-0.346。将源信号按照混合矩阵

$$A = \begin{bmatrix} 0.85 & -0.3 & 1.22 \\ -0.7 & -1.15 & 0.86 \\ -0.55 & 0.78 & 1.56 \end{bmatrix}$$

进行加权得到混合信号，混合信号波形如图 7 所示。

使用本文方法提取其中一个音乐信号。使用音色相同的乐器的其他音乐片段信号进行听觉特征 MFCC 的计算，对多段信号样本分别计算 MFCC，并取平均值。

提取出的幅度归一化音乐信号如图 8 所示，该乐器的先验 MFCC、提取信号的 MFCC 如图 9 所示，由图 9 可以看出，本文方法提取出的源信号与该乐器在听觉特征上更为接近。

根据仿真结果可以看出，提取信号与真实源信号的相似程度较高，听觉特征接近期望的源信号，

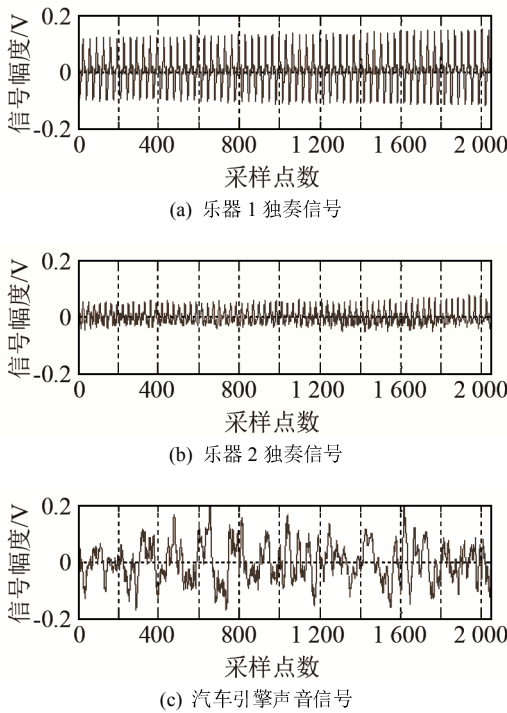
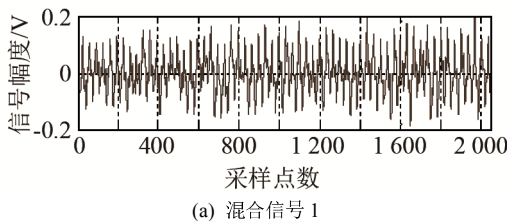
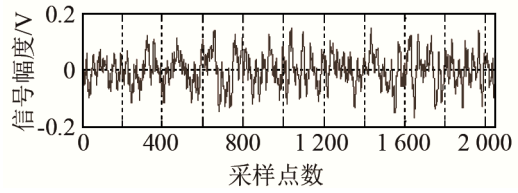


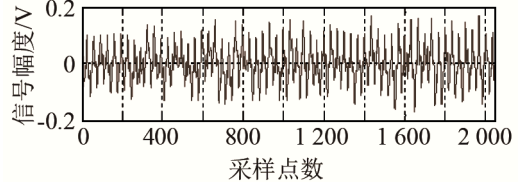
Fig.6 Source signals: 1 & 2. solo music signals of two different instruments, 3. sound of car engine



(a) 混合信号 1



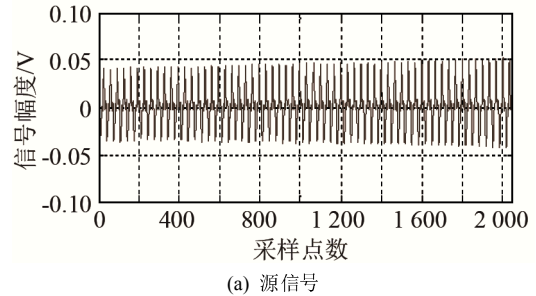
(b) 混合信号 2



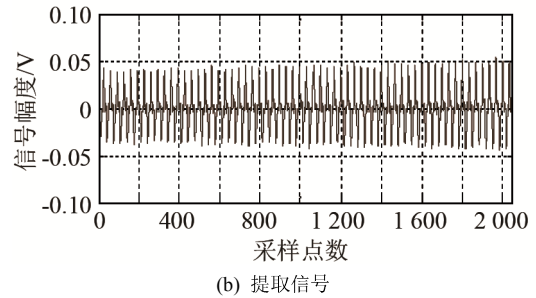
(c) 混合信号 3

图 7 图 6 中源信号的混合信号

Fig.7 Mixed signals corresponding to Fig.6



(a) 源信号



(b) 提取信号

图 8 提取出的音乐信号

Fig.8 Extracted music signal closest source signal

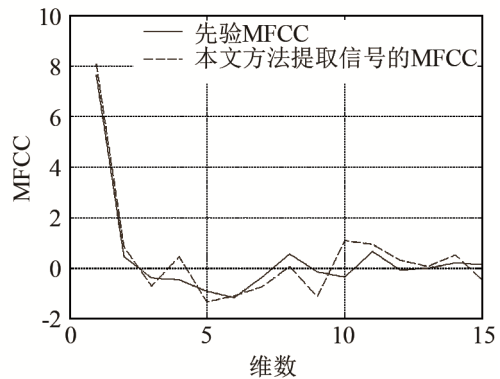


图 9 提取出的音乐信号的听觉特征

Fig.9 Auditory feature of the extracted music signal

当混合的源信号数目较多、无法获取当前混合的源信号波形参考信息时，本文方法能够较好地提取期望的源信号。

4 结 论

本文提出一种以期望信号听觉特征作为参考的盲提取方法,使用非线性最小二乘法求解分离向量,提取出特定的源信号。在无法得到参考信号具体波形的情况下,能够有效地实现信号提取。仿真结果表明,本文方法对不同类型的信号都有较好的稳健性,不受峭度等特性的影响。本文方法具有较为广泛的适应性和良好的应用价值。

参 考 文 献

- [1] COMON P. Independent component analysis, a new concept[J]. *Signal Processing*, 1994(36): 287-314.
- [2] CICHOCKI A, AMARI S. Adaptive blind signal and image processing[M]. New York, USA: Wiley, 2002.
- [3] 陈晋央, 吴琰. 基于独立分量分析的波束空间波达方向估计算法[J]. *声学技术*, 2012, 31(2): 227-231.
CHEN Jinyang, WU Ying. Beam-space direction of arrival (DOA) algorithm based on independent component analysis[J]. *Technical Acoustics*, 2012, 31(2): 227-231.
- [4] CICHOCKI A, THAWONMAS R, AMARI S. Sequential blind signal extraction in order specified by stochastics properties[J]. *Electronics Letters (S0013-5194)*, 1997, 33(1): 64-65.
- [5] CICHOCKI A, RUTKOWSKI T, SIWEK K. Blind signal extraction of signals with specified frequency band [C]//Proceedings of the 2002 12th IEEE Workshop on Neural Networks for Signal Processing, Switzerland. USA: IEEE, 2002: 515-524.
- [6] 季策, 王艳茹, 王晓宇. 一种基于标准峭度的新型复数盲分离算法[J]. *东北大学学报(自然科学版)*, 2015, 36(5): 614-617.
JI Ce, WANG Yanru, WANG Xiaoyu. A new complex blind source separation algorithm based on standard kurtosis[J]. *Journal of Northeastern University(Natural Science)*, 2015, 36(5): 614-617.
- [7] 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, 11(22): 6093-6096.
CHEN Shouqi, SHEN Yuehong, MA Ming, et al. Robust extraction of specific signal based on kurtosis[J]. *Journal of System Simulation*, 2008, 11(22): 6093-6096.
- [8] 刘若伦, 张家琦. 乘法模型下西洋乐器音色特征[J]. *声学技术*, 2009, 28(3): 269-275.
LIU Ruolun, ZHANG Jiaqi. Timbre features of western instrument under multiplication model[J]. *Technical Acoustics*, 2009, 28(3): 269-275.
- [9] 石超雄, 李钢虎, 何会会, 等. 基于提升小波变换的MFCC在目标识别中的应用[J]. *声学技术*, 2014, 33(4): 372-375.
SHI Chaoxiong, LI Ganghu, HE Huihui, et al. Application of the lifting wavelet transform based MFCC in target identification[J]. *Technical Acoustics*, 2014, 33(4): 372-375.
- [10] 朱知萌, 郭育, 王冠, 等. 水下蛙人呼吸声 Mel 倒谱特征的实验研究[J]. *声学技术*, 2017, 36(3): 286-292.
ZHU Zhimeng, GUO Yu, WANG Guan, et al. Experimental research on Mel cepstrum feature of the underwater diver breathing-signal[J]. *Technical Acoustics*, 2017, 36(3): 286-292.