

一种语音信号欠定盲分离的新方法^①

白琳 陈豪

(中国空间技术研究院西安分院, 西安 710000)

摘 要: 文章提出了一种语音信号欠定盲分离的新方法, 引入了一种伪提取矢量的概念, 并通过在信号采样点处选取对应的伪提取矢量, 以恢复取值占优的源信号采样点的值, 来实现欠定盲源分离。通过将该方法与传统的基于线性规划的欠定盲源分离方法分别仿真表明: 文章所提出方法由于在信号的各采样点处无需优化, 因此大大提高了信号分离的速度, 语音信号的分离速度比基于线性规划的方法快数倍。

关键词: 语音信号; 欠定盲源分离; 稀疏表征; 伪提取矢量

0 引言

独立信号源的盲分离是指在不知道源信号和传输信道任何先验知识的情况下, 仅通过观察信号恢复出各个独立源信号的过程^[1]。考虑最常见的线性形式为: 设 $s(k) = (s_1(k), s_2(k), \dots, s_n(k))^T$ 是 n 个零均值未知独立的源信号矢量, $x(k) = (x_1(k), x_2(k), \dots, x_m(k))^T$ 是经过信道传输混合后 m 个观测信号的矢量。其数学模型可以表示为:

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

其中 A 是一个 $m \times n$ 阶的未知矩阵, 称为混合矩阵。

在大部分关于盲信号分离的研究中, 一般假设 $m > n$ (观测信号的数目大于源信号的数目)。而 $m < n$ (欠定情形) 时的各种盲分离算法, 包括 JADE 算法、ICA 算法以及各种盲提取方法等, 都无法实现所有源信号被分离。考虑到语音信号在频域中具备一定的稀疏特性, 因此可以采用基于稀疏表示 (sparse representation) 的两步法 (two-step approach) 在频域中实现语音信号的欠定盲分离^[2-4]。即首先估计混合矩阵 A , 在已知 A 的基础上再求解源信号 $s(t)$ ^[5-7]。

文章在假定 A 已经估计出来的前提下, 集中研究如何估计源信号 $s(t)$ 。假设在观测信号的每个采样点存在 1 个提取矢量, 该提取矢量不同于适应模型 (源信号数目等于观测信号数目) 下的盲提取矢量, 文中称其为伪提取矢量, 通过其提取取值占优的源信号的采样值来恢复源信号。文中方法相比于基于线性规划的两步法分离源信号, 分离速度大大提高。

1 基于线性规划的欠定盲分离方法

信号具备一定的稀疏性是进行欠定盲源分离的前提。因此, 如果信号不够稀疏, 首先应对观测信号进行 Fourier 变换或者小波变换转化为稀疏信号。记稀疏化了的观测信号为 $x(k)$, $k = 1, \dots, T$, 对于式 (1) 的混合模型, 由文献 [6] 可知, 在不考虑噪声的情况下, 稀疏信号盲分离可归结为求解如下线性规划问题:

$$\begin{cases} \min_{s(t)} \sum_i |s_i(k)| \\ s. t. : As(k) = x(k) \quad k = 1, \dots, T \end{cases} \quad (2)$$

对于每个采样点 $k = 1, \dots, T$, 都可确定一个优化问题, 从而欠定盲源分离问题就转化为求解 T 个优化问题。从上述基于线性规划的欠定盲源分离方法可以看出, 如果信号的采样点较多, 依次进行多个优化需要很大的计算量, 因此, 这种算法的计算时间很长, 分离速度很慢。

2 基于伪提取矢量的欠定盲分离新方法

文章给出了一种新的源信号恢复的方法: 在欠定盲源分离中, 存在一个提取矢量组 $\vec{w} = (w_1, w_2, \dots, w_n)$, 文中把这些矢量 w_1, w_2, \dots, w_n 称为伪提取矢量, 而伪提取矢量 w_i ($i = 1, 2, \dots, n$) 能恢复出第 i 个源信号取值占优的各个采样点的值。下面对这一方法进行简单证明。

证明: 假设在 T_1, T_2 ($T_1, T_2 = 1, 2, \dots, T$) 采样点

收稿日期: 2011-04-28

处源信号 s_i 取值占优。即 $s_j(T_1) \neq 0, s_j(T_2) \neq 0$, 当 $j=i$ 而 $s_j(T_1)=0, s_j(T_2)=0$, 当 $j \neq i (i, j \in 1, 2, \dots, n)$ 。设提取矢量 $w_i (i \in 1, 2, \dots, n)$ 与 $r_i (i \in 1, 2, \dots, n)$ 分别能恢复出 $s_i(T_1)$ 与 $s_i(T_2)$ 。即:

$$\begin{aligned} w_{i1}x_{1T_1} + \dots + w_{iT_1}x_{T_1T_1} + \dots + w_{im}x_{mT_1} &= \\ s_{iT_1} \quad r_{i1}x_{1T_2} + \dots + r_{iT_2}x_{T_2T_2} + \dots + r_{im}x_{mT_2} &= s_{iT_2} \end{aligned} \quad (3)$$

由于 w_i 和 r_i 分别都是 m 维矢量,故除了等式(3)以外,要求解 w_i 和 r_i 分别还需要 $m-1$ 个等式。不失一般性,假设 s_i 的第 $1, 2, \dots, m-1$ 个采样点为非取值占优时刻,即 $s_i(k) \approx 0 (k=1, 2, \dots, m-1)$ 。于是,可得 m 元线性方程组:

$$\begin{cases} w_{i1}x_{1T_1} + \dots + w_{im}x_{mT_1} = s_{iT_1} \\ w_{i1}x_{11} + \dots + w_{im}x_{m1} = 0 \\ \dots \\ w_{i1}x_{1m-1} + \dots + w_{im}x_{mm-1} = 0 \end{cases}$$

$$\text{与} \begin{cases} r_{i1}x_{1T_2} + \dots + r_{im}x_{mT_2} = s_{iT_2} \\ r_{i1}x_{11} + \dots + r_{im}x_{m1} = 0 \\ \dots \\ r_{i1}x_{1m-1} + \dots + r_{im}x_{mm-1} = 0 \end{cases} \quad (4)$$

由于在 $T_1, T_2 (T_1, T_2 \in 1, 2, \dots, T)$ 采样点处,源信号 s_i 取值占优,故式(1)的混合模型可近似写为:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \dots & 0 & \dots & 0 & \dots \\ \cdot & \vdots & & \vdots & \cdot \\ \cdot & s_{iT_1} & \ddots & s_{iT_2} & \cdot \\ \cdot & \dots & & \dots & \cdot \\ \dots & 0 & \dots & 0 & \dots \end{bmatrix} =$$

$$\begin{bmatrix} \dots & x_{1T_1} & \dots & x_{1T_2} & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \dots & x_{mT_1} & \dots & x_{mT_2} & \dots \end{bmatrix} \quad (5)$$

于是,

$$\begin{aligned} x_{1T_1} &= a_{1i}s_{iT_1}, x_{2T_1} = a_{2i}s_{iT_1}, \dots, x_{mT_1} = a_{mi}s_{iT_1}, \\ x_{1T_2} &= a_{1i}s_{iT_2}, x_{2T_2} = a_{2i}s_{iT_2}, \dots, x_{mT_2} = a_{mi}s_{iT_2} \end{aligned} \quad (6)$$

将式(6)代入式(4)的2个方程组中,可发现这2个方程组为同解方程组,即:

$$w_{i1} = r_{i1}, w_{i2} = r_{i2}, \dots, w_{im} = r_{im} \quad (7)$$

由 T_1, T_2 的任意性可知,对于第 i 个源信号取值占优的所有采样点的值都可以由同一个伪提取矢量 $w_i (i \in 1, 2, \dots, n)$ 恢复出来。进而存在 n 个伪提取矢量 $w_i (i=1, 2, \dots, n)$ 能恢复出所有 n 个源信号取

值占优的各个采样点的值。接下来只要找出所有采样点处取值占优的源信号,然后,各个观测采样点只需选取对应的伪提取矢量,就可以恢复出取值占优的源信号的采样点的值;对于采样点处非取值占优的源信号的值,近似取零即可。而 m 个伪提取矢量的获取,可以通过建立和求解 n 个类似于式(4)的线性方程组得到。

3 欠定语音信号盲分离算法步骤

对于欠定语音信号的盲分离算法,文章总结为以下步骤实现:

(1) 选取合适的处理域,以使得源信号在这一处理域中尽可能稀疏。假设观测信号在该域中的矢量形式为 $[x_1, x_2, \dots, x_n]^T$ 。

(2) 逐一判断各个采样点处 $(x_1(k), \dots, x_m(k))^T (k=1, 2, \dots, T)$ 的方向,并分别与 A 的 n 个列向量方向比较,在第 k 个采样点处,找出与 $(x_1(k), \dots, x_m(k))^T$ 的方向最接近的 A 的第 i 个列向量,则在该采样点处,第 i 个源信号即为取值占优的源信号。建立并求解类似于式(4)的 n 个 m 元线性方程组,获得 n 个 m 维伪提取矢量 $[w_{i1}, w_{i2}, \dots, w_{im}]^T, i=1, 2, \dots, n$ 。

(3) 在采样点 $k=1, 2, \dots, T$ 处,如果第 i 个源信号取值占优,则用第 i 个伪提取矢量提取并获得该源信号在该采样点处的值,其它源信号在该采样点处的值,则取零。

(4) 对稀疏域中恢复的信号进行逆 Fourier 变换或者小波变换,实现源信号的重构。

4 计算机仿真

以6个长笛声的语音信号作为源信号,语音信号来自 <http://personals.ac.upc.edu>。采样率为44 100Hz,抽取的样本采样点为58 488个。混合矩阵 A 随机产生,且它的标准化形式为:

$$A = \begin{bmatrix} -0.5361 & -0.1734 & -0.2482 & 0.9915 & -0.3148 & 0.9631 \\ -0.8441 & -0.9848 & 0.9687 & 0.1298 & 0.9492 & 0.2693 \end{bmatrix}。$$

6个长笛及其2个混合观测波形如图1所示。

图2给出了2个接收传感器观测信号的比例散点图。从图中可看出,在频域中,散点图的分布更加散落,也就是说6个长笛语音信号在频域中的稀疏性更强。因此应选择算法的处理域为频域。

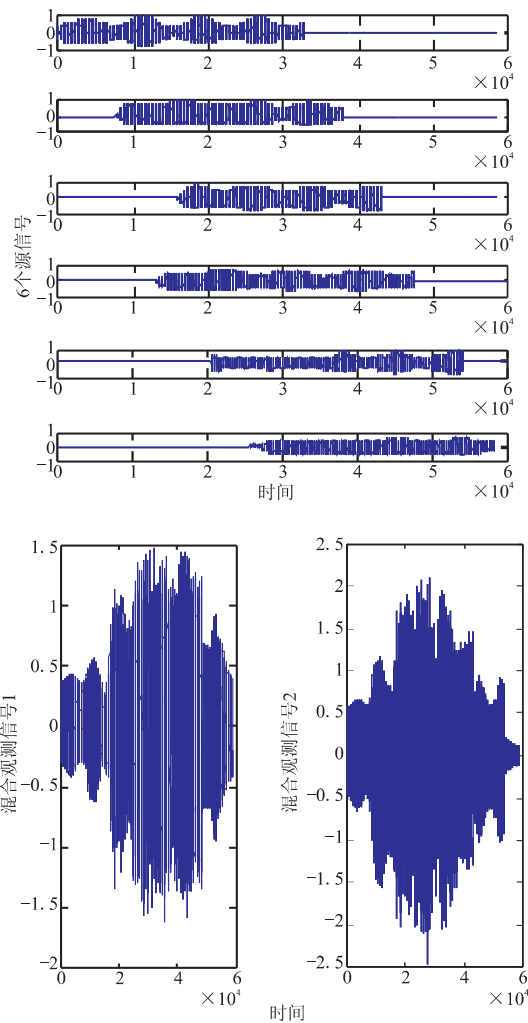


图1 6个长笛及其2个混合观测波形图

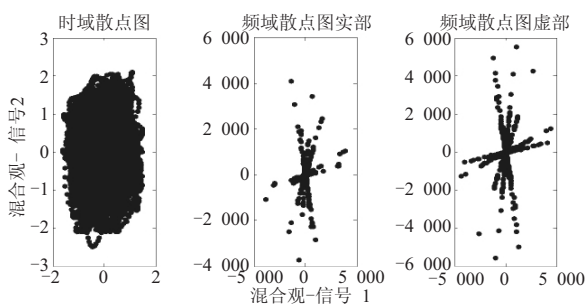


图2 2个接收传感器信号的比例散点图

图3给出了各源信号频域采样点占优的比例。

图4给出了使用文中算法分离后的6个长笛信号与6个长笛源波形对比。

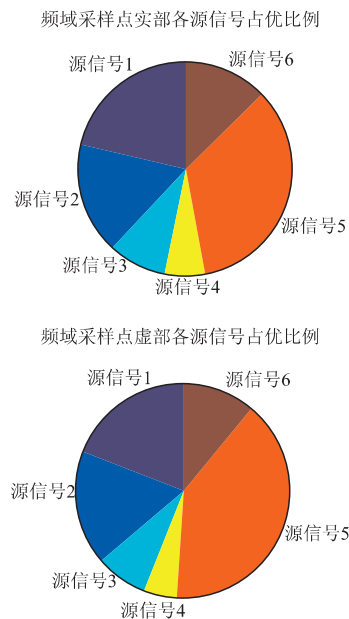


图3 各源信号频域采样点实部和虚部占优的比例

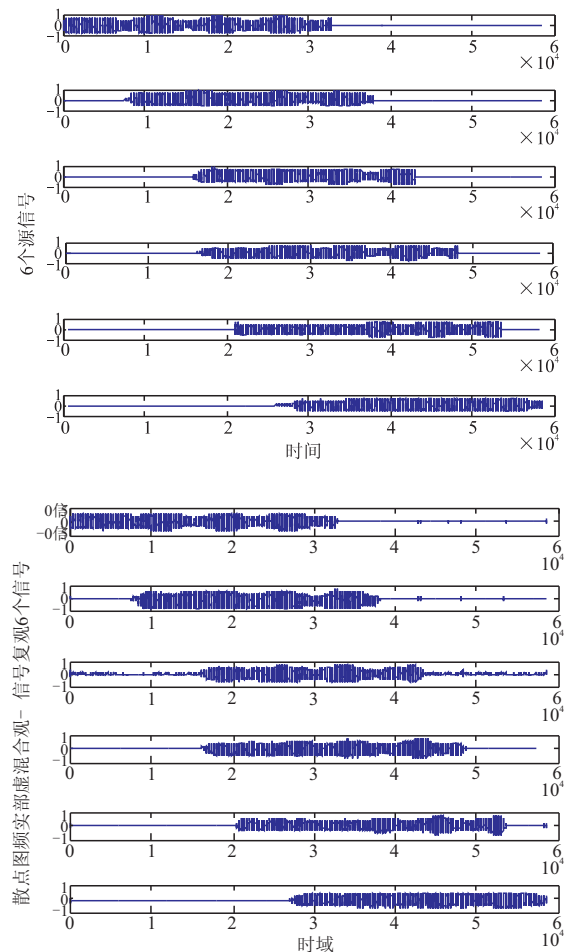


图4 使用文中算法分离后的6个长笛信号波形与6个长笛源波形对比

此外,还使用了基于线性规划的欠定盲源分离方法进行了该仿真实验,以对比说明文中提出的基于伪提取矢量的欠定盲源分离方法的性能。表 1 给出了使用两种算法的性能对比。表中的相似系数是描述所恢复信号与源信号相似性的参数,定义为:

$$\xi_{ij} = \xi(\hat{s}_i, s_j) = \frac{\left| \sum_{k=1}^T \hat{s}_i(k) s_j(k) \right|}{\sqrt{\sum_{k=1}^T \hat{s}_i^2(k) \sum_{k=1}^T s_j^2(k)}}$$

$$(i, j = 1, 2, \cdots, n) \tag{8}$$

式(8)中 $\hat{s}_i, s_j (i, j = 1, 2, \cdots, n)$ 分别表示各恢复信号与源信号。当 $i = j, \xi_{ij} = 1; i \neq j, \xi_{ij} = 0$ 时,为信号恢复的理想情况。

表 1 6 个混叠信号的分离使用两种算法的性能对比

		相似系数						运行 时间(sec)
		源语音信号						
		恢复语音信号	信号 1	信号 2	信号 3	信号 4	信号 5	
文中基于伪 提取矢量的 方法	信号 1	0.994 3	0.002 5	0.000 8	0.001 2	0.001 6	0.001 9	401.118 678 s
	信号 2	0.003 6	0.995 5	0.002 6	0.001 4	0.002 7	0.000 2	
	信号 3	0.001 5	0.000 7	0.985 7	0.000 3	0.007 0	0.000 3	
	信号 4	0.001 7	0.000 1	0.001 8	0.990 5	0.000 6	0.001 2	
	信号 5	0.000 4	0.000 5	0.015 8	0.001 1	0.986 5	0.000 1	
	信号 6	0.001 3	0.000 0	0.001 1	0.013 2	0.000 1	0.991 3	
基于线性规 划的方法	信号 1	0.995 0	0.011 5	0.000 7	0.002 8	0.002 6	0.007 0	1 322.007 85 s
	信号 2	0.018 5	0.995 2	0.014 1	0.001 2	0.002 6	0.000 1	
	信号 3	0.002 4	0.010 0	0.982 9	0.000 5	0.047 8	0.000 0	
	信号 4	0.001 7	0.000 4	0.002 1	0.991 1	0.003 6	0.020 2	
	信号 5	0.000 4	0.000 2	0.064 4	0.009 8	0.982 8	0.000 2	
	信号 6	0.009 9	0.000 4	0.000 4	0.044 0	0.001 8	0.990 7	

通过表 1 的相似系数可以看出,从分离精度上比较,两种算法都能以较高的分离精度分离出各源信号。文中的基于伪提取矢量的方法对于信号 2、信号 3、信号 5、信号 6 的分离精度略高于基于线性规划的方法;信号 1、信号 4 的分离精度稍低于基于线性规划的方法;总的来说,两种算法对信号的分离精度相差不大。从分离速度上比较,文中基于伪提取矢量的方法的分离速度是基于线性规划的方法的 3.295 8 倍。其根本原因就是,基于线性规划的方法在每个采样点处都采用了复杂的优化算法;而基于伪提取矢量的方法在源信号的整个过程没有采用任何优化算法,因此大大提高了分离速度。

5 结论

文章提出了一种新的实现欠定混合模型下的语音信号盲分离的方法,给出了该方法的理论分析以

及实现的基本算法流程,并采用该方法进行了欠定模型下多个混合语音信号盲分离的计算机仿真。此外,在源信号的分离过程中,还使用文中的算法与传统的基于线性规划的欠定盲源分离方法进行了对比。结果表明,文中所提出的方法在源信号的分离过程中,分离精度不低于基于线性规划的方法,分离速度要比基于线性规划的方法快数倍。

参考文献:

[1] Hyvarinen A, Oja E. Independent component analysis: algorithms and applications[J]. Neural Networks, 2000, 13: 411 ~ 430.
[2] 李映, 张艳宁, 许星. 基于信号稀疏表示的形态成分分析: 进展和展望[J]. 电子学报, 2009, 37(1): 146 ~ 152.

(下转第 65 页)

Design of Parameters for a Coaxial Bragg Reflectors Operating at Ka-band

WANG Ling-ling ,LI She-lei ,DING Xue-yong

(Dept. of Physics Engineering ,Hai nan University Sanya college ,Sanya 572022 ,China)

Abstract: TBased on the coupled mode theory and the CST software ,design of parameters and comparative study are carried out for the numerical simulations of a coaxial Bragg reflector operating at Ka-band corrugated with sinusoidal ripples. Results show good agreement between the multi-mode theory and the CST software. In fact ,by changing the initial phase ,sinusoidal ripple amplitude ,shape of gradient and gradient angle and windowing-function technique can improve the frequency response of a coaxial Bragg structure. These measures are favorable to improvement of its performance as a reflector or a filter in its applications.

Key words: Coaxial Bragg-reflector; Ka-Band; Reflectivity

(上接第 19 页)

- [3] Hadi Zayyani , Massoud Babaie-Zadeh , Christian Jutten , et al. An iterative bayesian algorithm for sparse component analysis in presence of noise [J]. IEEE Trans. on Signal Processing , 2009 , 57(11) : 423 ~ 437.
- [4] Bofill P , Zibulevsky M. Underdetermined blind source separation using sparse representations [J]. Signal Process 2001 81(11) : 2353 ~ 2362.
- [5] Peng De-zhong , Yong Xiang. Underdetermined blind source separation based on relaxed sparsity condition of sources [J]. IEEE Trans. on Signal Processing , 2009 , 57(2) : 809 ~ 814.
- [6] 何昭水 , 谢胜利 , 傅予力. 稀疏表示与病态混叠盲分离 [J]. 中国科学 E 辑 2006 , 36(8) : 864 ~ 879.
- [7] Li Yuan qing , Amari Shun-ichi , Cichocki Andrzej , et al.

Probability estimation for recoverability analysis of under-determined blind source separation based on sparse representation [J]. IEEE Trans. on Informations Thery , 2006 , 52: 3139 ~ 3152.

作者简介:

白 琳 1982 年生 , 中国空间技术研究院西安分院在读博士。主要研究方向为卫星通信技术以及盲信号处理。

陈 豪 1944 年生 , 中国空间技术研究院西安分院研究员 , 博士生导师 , 享受政府特殊津贴专家。主要研究方向为卫星通信技术以及通信信号处理等。

A New Method of Underdetermined Blind Separation for Voice Signals Separation for Voice Signals

BAI Lin , CHEN Hao

(China Academy of Space Technology(Xi'an) , Xi'an 710000 , China)

Abstract: A new method is put forward for accomplishing underdetermined blind signal separation (UBSS) of voice signals in the paper. A new concept of pseudo extraction vectors is introduced. The method accomplishes UBSS by choosing corresponding pseudo extraction vector to recover the signal dominating for each sampling. Some conclusions can be reached by separately simulating the method based on pseudo extraction vectors and linear programming: the method based on pseudo extraction vectors free of optimizing process increases the velocity of separating source signals. The velocity of separation of the method based on pseudo extraction vectors is several-fold of the method based on linear programming for voice signals.

Key words: Voice signals; Underdetermined BSS; Sparse representation; Pseudo extraction vectors