

# 基于联合对角化的声信号深度卷积混合盲分离方法

李 扬 张伟涛\* 楼顺天

(西安电子科技大学电子工程学院 西安 710071)

**摘 要:** 声信号在空间中的传播具有较强的多径效应, 在接收端往往以卷积形式相互叠加, 尤其在海洋、剧场等强混响条件下, 混合滤波器冲激响应的长度会显著增加, 现有的频域卷积盲分离算法将失效。为了消除长脉冲响应导致解混模型失效的问题, 该文对观测信号进行两次短时傅里叶变换(STFT), 第1次STFT缩短了脉冲响应长度, 第2次STFT将信号模型转化为瞬时盲分离, 最终利用联合对角化(JD)技术估计出分离矩阵。与现有方法相比, 所提方法解决了深度卷积混合下模型失效的问题, 并且当源信号数较多或存在加性噪声时, 可以得到更好的分离性能。仿真结果验证了方法的有效性和性能优势。

**关键词:** 盲源分离; 深度卷积; 联合对角化; 排序问题

中图分类号: TN911.7

文献标识码: A

文章编号: 1009-5896(2019)12-2951-06

DOI: 10.11999/JEIT190067

## Deep Convolution Blind Separation of Acoustic Signals Based on Joint Diagonalization

LI Yang ZHANG Weitao LOU Shuntian

(Institute of Electronic Engineering, Xidian University, Xi'an 710071, China)

**Abstract:** The propagation of acoustic signal in space has a strong multipath effect, and the receiver often overlaps in the form of convolution. Especially in strong reverberation conditions such as ocean and theatre, where the length of impulse response of hybrid filter increases significantly. In order to eliminate the problem that long impulse response leads to the failure of the frequency domain convolution blind separation algorithm, two Short-Time Fourier Transforms (STFT) are applied to the observed signal. The first STFT shortens the length of the hybrid filter. The second STFT converts the signal model into instantaneous blind separation. Finally, the separation matrix is estimated by Joint Diagonalization (JD) technique. Compared with the existing methods, this method solves the problem of model failure under deep convolution mixing, and can obtain better separation performance when the number of source signals is large or additive noise exists. The simulation results verify the effectiveness and performance advantages of the proposed method.

**Key words:** Blind source separation; Deep convolution; Joint Diagonalization (JD); Permutation problem

### 1 引言

盲信号分离是指在传输信道未知的情况下, 仅从观测信号估计出源信号的过程。声信号由于在传输过程中存在多径效应, 在接收端通常用卷积混合模型来描述。卷积混合可以表示为一组冲激响应滤波器, 其最大的信道延迟到达时间决定了该混合滤波器的长度。当所处环境较为空旷时, 多径效应会

造成反射信号延迟时间显著增大, 当混合滤波器的长度大于近似平稳的最大长度时, 该系统构成深度卷积混合。

盲解卷积混合方法目前主要分为时域方法<sup>[1-3]</sup>和频域方法<sup>[4,5]</sup>两种, 时域方法研究较早, 但由于计算复杂度高, 近些年来盲解卷积方法的研究主要集中在频域方法上。对于深度卷积混合盲分离方法的研究, 有通过非局部加权软约束的自然梯度迭代方法<sup>[6]</sup>, 能够实现较高强度混响盲分离, 但是受卷积深度影响, 计算量较大。另一种是通过波束形成预处理的方法<sup>[7,8]</sup>, 将非源信号方向的叠加信号抑制掉, 降低混响强度再实现卷积盲源分离, 但是这种方法对波束形成主瓣宽度和旁瓣幅度要求比较高, 否则会导致随后的盲信号分离性能变差。文

收稿日期: 2019-01-24; 改回日期: 2019-06-11; 网络出版: 2019-06-24

\*通信作者: 张伟涛 zhwt-work@foxmail.com

基金项目: 国家自然科学基金(61571339), 陕西省创新人才推进计划-青年科技新星项目(2018KJXX-019)

Foundation Items: The National Natural Science Foundation of China (61571339), The Innovative Talents Promotion Program of Shaanxi Province (2018KJXX-019)

献[9,10]提出另外一种模型,将强混响条件下的卷积混合信号经过一次时频变换后,对每个频点上的混合信号分量再做短时傅里叶变换,将频域卷积混合变换为调制谱域上的瞬时混合,再通过独立分量分析(Independent Component Analysis, ICA)的方法实现盲源分离。

本文在现有研究基础上,对深度卷积混合模型进行了推导,提出了利用联合对角化实现2次频域上瞬时混合盲分离的方法,同时分析了本模型下信号盲分离过程中排序混乱的问题,提出了频域和2次频域两次解排序混乱的方法,与现有方法相比,在分离性能上具有较大的提高。

## 2 深度卷积信号混合模型

对于一个 $M$ 个麦克风、 $N$ 个源信号构成的深度卷积混合系统( $M \geq N$ ),其中第 $m$ 个麦克风接收到的采样信号 $x_m(t)$ 可以表示为

$$x_m(t) = \sum_{n=1}^N \sum_{\tau=0}^{L-1} a_{mn}(\tau) s_n(t - \tau) + v_m(t) \quad (1)$$

其中,  $a_{mn}(\tau)$ 表示从第 $n$ 个信源到第 $m$ 个麦克风延时为 $\tau$ 时的信号幅度衰减系数,  $s_n(t)$ 表示第 $n$ 个信源,  $v_m(t)$ 为加性噪声,  $L = \text{fs} \cdot \tau_{\max}$ 为混合滤波器的最大长度,其中 $\text{fs}$ 为信号采样率,  $\tau_{\max}$ 为传输信道最大延迟时间。令  $\mathbf{a}_m(t) = [a_{m1}(t), a_{m2}(t), \dots, a_{mN}(t)]^T$ ,  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$ , 则式(1)可以表示为

$$x_m(t) = \sum_{\tau=0}^{L-1} \mathbf{a}_m^T(\tau) \mathbf{s}(t - \tau) + v_m(t) \quad (2)$$

频域盲分离方法需要对式(2)两端做窗长度为 $H$ 的短时傅里叶变换(Short-Time Fourier Transform, STFT),此时窗长度的选择要遵循一定的限制。首先,窗函数长度 $H$ 应该小于声信号最大近似平稳长度 $H_{\max}$ ,其次,窗长度 $H$ 又需要至少大于混合滤波器的长度 $L$ 。在非深度卷积条件下,  $L$ 长度小于 $H_{\max}$ ,因此窗长度可以选择 $H \in [L, H_{\max}]$ 。但是当信号处于深度卷积混合时,混合滤波器长度的显著增加将会导致 $L > H_{\max}$ ,此时窗长度的选择将陷入矛盾。

基于上述分析,在深度卷积背景下,将式(2)改写为分段形式

$$x_m(t) = \sum_{\tau=0}^{H-1} \mathbf{a}_m^T(\tau) \mathbf{s}(t - \tau) + \sum_{\tau=H}^{2H-1} \mathbf{a}_m^T(\tau) \mathbf{s}(t - \tau) + \dots + \sum_{\tau=(p-1)H}^{pH-1} \mathbf{a}_m^T(\tau) \mathbf{s}(t - \tau) + v_m(t) \quad (3)$$

其中,  $H$ 为分段长度,  $p$ 是总分段数,  $L \leq pH < L + H$ 。对式(3)两端做STFT,窗长度取 $H$ ,在频点 $\omega$ 上,可以得到混合信号频谱分量

$$X_m(\omega, \lambda) = \sum_{i=1}^p \mathbf{a}_m^T(\omega, i) \mathbf{s}(\omega, \lambda - i + 1) + V_m(\omega, \lambda) \quad (4)$$

其中,  $X_m(\omega, \lambda)$ 和 $V_m(\omega, \lambda)$ 分别为信号 $x_m(t)$ 和加性噪声 $v_m(t)$ 经过STFT在频点 $\omega$ 上的第 $\lambda$ 帧,  $\mathbf{a}_m(\omega, i) = [A_{m1}(\omega, i), A_{m2}(\omega, i), \dots, A_{mN}(\omega, i)]^T$ ,  $i = 1, 2, \dots, p$ 是 $\mathbf{a}_m(t)$ 经过STFT后的第 $i$ 帧。  $\mathbf{s}(\omega, \lambda) = [S_1(\omega, \lambda), S_2(\omega, \lambda), \dots, S_N(\omega, \lambda)]^T$ 为 $\mathbf{s}(t)$ 经过STFT后得到的第 $\lambda$ 帧。由式(4)可知,经过1次STFT后,在每个频点 $\omega$ 上,源信号分量重新构成了一个混合滤波器长度为 $p$ ( $p > 1$ )的卷积混合,这里称之为2次卷积混合。相比于深度卷积混合滤波器长度 $L$ ,  $p$ 的长度被显著缩小,2次卷积混合转变为非深度卷积混合。当 $p = 1$ 时,式(4)退化成瞬时混合,此时信号模型变为非深度卷积混合。

## 3 信号盲分离算法

### 3.1 信号模型变换

将式(4)做2次STFT,窗长度为 $h \geq p$ ,在2次频域频点 $\omega'$ 上,可以得到2次频谱分量

$$X_m(\omega, \omega', r) = \mathbf{a}_m^T(\omega, \omega') \mathbf{s}(\omega, \omega', r) + V_m(\omega, \omega', r) \quad (5)$$

其中,  $X_m(\omega, \omega', r)$ 和 $V_m(\omega, \omega', r)$ 分别为对 $X_m(\omega, \lambda)$ 和 $V_m(\omega, \lambda)$ 进行2次STFT后在2次频域频点 $\omega'$ 上的第 $r$ 帧,  $\mathbf{a}_m(\omega, \omega') = [A_{m1}(\omega, \omega'), A_{m2}(\omega, \omega'), \dots, A_{mN}(\omega, \omega')]^T$ 为 $\mathbf{a}_m(\omega, i)$ ,  $i = 1, 2, \dots, p$ 在2次频域上的STFT,  $\mathbf{s}(\omega, \omega', r)$ 为对 $\mathbf{s}(\omega, \lambda)$ 进行2次STFT后在2次频域频点 $\omega'$ 上的第 $r$ 帧。令  $\mathbf{x}(\omega, \omega', r) = [X_1(\omega, \omega', r), X_2(\omega, \omega', r), \dots, X_M(\omega, \omega', r)]^T$ ,  $\mathbf{A}(\omega, \omega') = [\mathbf{a}_1^T(\omega, \omega'), \mathbf{a}_2^T(\omega, \omega'), \dots, \mathbf{a}_M^T(\omega, \omega')]^T$ ,  $\mathbf{v}(\omega, \omega', r) = [V_1(\omega, \omega', r), V_2(\omega, \omega', r), \dots, V_M(\omega, \omega', r)]^T$ 则由式(5)可得

$$\mathbf{x}(\omega, \omega', r) = \mathbf{A}(\omega, \omega') \mathbf{s}(\omega, \omega', r) + \mathbf{v}(\omega, \omega', r) \quad (6)$$

由式(6)可知,在2次频域频点 $\omega'$ 上,混合信号2次频谱分量 $\mathbf{x}(\omega, \omega', r)$ 构成源信号2次频谱分量 $\mathbf{s}(\omega, \omega', r)$ 的瞬时混合。此时,需要找到一个分离矩阵 $\mathbf{W}(\omega, \omega')$ 使得

$$\mathbf{W}(\omega, \omega') \mathbf{A}(\omega, \omega') = \mathbf{\Pi} \mathbf{\Sigma}(\omega, \omega') \quad (7)$$

其中,  $\mathbf{\Pi}$ 是造成排序混乱的置换矩阵,  $\mathbf{\Sigma}(\omega, \omega')$ 是在频域频点 $\omega$ 和2次频域频点 $\omega'$ 上的对角矩阵。通过分离矩阵 $\mathbf{W}(\omega, \omega')$ 可以求得在对应频点上对源信号分量 $\mathbf{s}(\omega, \omega', r)$ 的估计

$$\mathbf{y}(\omega, \omega', r) = \mathbf{W}(\omega, \omega') \mathbf{x}(\omega, \omega', r) \quad (8)$$

### 3.2 2次频域分离矩阵估计

在2次频域上采用联合对角化方法<sup>[11-13]</sup>对混合信号2次谱分量的分离矩阵进行估计, 本文将文献<sup>[14]</sup>提出的非酉联合对角化方法应用在2次频域上, 首先, 对式(6)信号分量 $\mathbf{x}(\omega, \omega', r)$ 求2次频域谱密度矩阵

$$\begin{aligned} \mathbf{R}_x(\omega, \omega', q) &= \mathbf{A}(\omega, \omega') \mathbf{R}_s(\omega, \omega', q) \mathbf{A}^H(\omega, \omega') \\ &\quad + \mathbf{R}_v(\omega, \omega', q) \end{aligned} \quad (9)$$

其中,  $\mathbf{R}_x(\omega, \omega', q)$ ,  $\mathbf{R}_s(\omega, \omega', q)$ 和 $\mathbf{R}_v(\omega, \omega', q)$ 分别为混合信号、源信号和加性噪声的2次频域谱密度矩阵,  $q = 0, 1, \dots, Q$ 是帧延长度。由于加性噪声谱分量 $\mathbf{v}_1(\omega, \omega', r)$ ,  $\mathbf{v}_2(\omega, \omega', r)$ ,  $\dots$ ,  $\mathbf{v}_M(\omega, \omega', r)$ 彼此相互独立, 仅当 $q = 0$ 时,  $\mathbf{R}_v(\omega, \omega', q)$ 是一个非0对角矩阵, 当 $q \neq 0$ 时, 2次频域谱密度矩阵可表示为

$$\mathbf{R}_x(\omega, \omega', q) = \mathbf{A}(\omega, \omega') \mathbf{R}_s(\omega, \omega', q) \mathbf{A}^H(\omega, \omega'), \quad q = 1, 2, \dots, Q \quad (10)$$

非酉联合对角化的目标就是找到一个非奇异对角化器 $\mathbf{B}$ , 使得 $\mathbf{B}^H \mathbf{R}_x(\omega, \omega', q) \mathbf{B}$ ,  $q = 1, 2, \dots, Q$ 都是对角矩阵。对于频点 $\omega_j, j = 1, 2, \dots, J$ 和2次频点 $\omega'_k, k = 1, 2, \dots, K$ , 联合对角化可通过优化如式(11)的目标函数得到

$$\begin{aligned} \min \mathcal{J}(\mathbf{B}(\omega_j, \omega'_k)) &= \sum_{q=1}^Q \left\| \text{Off}(\mathbf{B}^H(\omega_j, \omega'_k) \right. \\ &\quad \left. \cdot \mathbf{R}_x(\omega_j, \omega'_k, q) \mathbf{B}(\omega_j, \omega'_k)) \right\|_F^2 \end{aligned} \quad (11)$$

其中,  $\text{Off}(\cdot)$ 为矩阵的非对角函数。如果直接对式(11)进行最小化会产生退化解, 为了避免退化解, 采用如式(12)的多约束优化模型

$$\left. \begin{aligned} \min_{\mathbf{B}(\omega_j, \omega'_k)} \mathcal{J}(\mathbf{B}(\omega_j, \omega'_k)), \quad \max_{\mathbf{B}(\omega_j, \omega'_k)} \det(\mathbf{B}^H(\omega_j, \omega'_k) \\ \mathbf{B}(\omega_j, \omega'_k)), \\ \mathbf{b}_n^H(\omega_j, \omega'_k) \mathbf{b}_n(\omega_j, \omega'_k) = 1, n = 1, 2, \dots, N \end{aligned} \right\} \quad (12)$$

其中,  $\mathbf{b}_n(\omega_j, \omega'_k)$ 是 $\mathbf{B}(\omega_j, \omega'_k)$ 的第 $n$ 列。直接对式(12)进行优化比较复杂, 本文将优化分成了 $N$ 个子优化问题, 每次优化 $\mathbf{B}(\omega_j, \omega'_k)$ 的其中一列 $\mathbf{b}_n(\omega_j, \omega'_k)$ ,  $n = 1, 2, \dots, N$ , 同时保持其他列固定。对于一个子优化问题, 将式(12)的优化模型改写为 $\mathbf{b}_n(\omega_j, \omega'_k)$ 的函数可以得到

$$\mathcal{J}(\mathbf{b}_n(\omega_j, \omega'_k)) = \text{tr}(\mathbf{b}_n^H(\omega_j, \omega'_k) \mathbf{Q}_n \mathbf{b}_n(\omega_j, \omega'_k)) \quad (13)$$

$$\begin{aligned} \det(\mathbf{B}^H(\omega_j, \omega'_k) \mathbf{B}(\omega_j, \omega'_k)) \\ = \det(\mathbf{B}_n^H(\omega_j, \omega'_k) \mathbf{B}_n(\omega_j, \omega'_k)) \mathbf{b}_n^H(\omega_j, \omega'_k) \\ \cdot \mathbf{P}_n^\perp \mathbf{b}_n(\omega_j, \omega'_k) \end{aligned} \quad (14)$$

其中,  $\text{tr}(\cdot)$ 表示矩阵的迹, 同时

$$\begin{aligned} \mathbf{Q}_n &= \sum_{q=1}^Q (\mathbf{R}_x(\omega_j, \omega'_k, q) \mathbf{B}_n(\omega_j, \omega'_k) \mathbf{B}_n^H(\omega_j, \omega'_k) \\ &\quad \cdot \mathbf{R}_x^H(\omega_j, \omega'_k, q) + \mathbf{R}_x^H(\omega_j, \omega'_k, q) \\ &\quad \cdot \mathbf{B}_n(\omega_j, \omega'_k) \mathbf{B}_n^H(\omega_j, \omega'_k) \mathbf{R}_x(\omega_j, \omega'_k, q)) \end{aligned} \quad (15)$$

$$\begin{aligned} \mathbf{P}_n^\perp &= \mathbf{I} - \mathbf{B}_n(\omega_j, \omega'_k) [\mathbf{B}_n^H(\omega_j, \omega'_k) \mathbf{B}_n(\omega_j, \omega'_k)]^{-1} \\ &\quad \cdot \mathbf{B}_n^H(\omega_j, \omega'_k) \end{aligned} \quad (16)$$

其中,  $\mathbf{B}_n(\omega_j, \omega'_k)$ 表示 $\mathbf{B}(\omega_j, \omega'_k)$ 去掉第 $n$ 列 $\mathbf{b}_n(\omega_j, \omega'_k)$ 。由此得到子优化问题的约束模型为

$$\left. \begin{aligned} \min_{\mathbf{b}_n(\omega_j, \omega'_k)} \mathbf{b}_n^H(\omega_j, \omega'_k) \mathbf{Q}_n \mathbf{b}_n(\omega_j, \omega'_k), \\ \max_{\mathbf{b}_n(\omega_j, \omega'_k)} \mathbf{b}_n^H(\omega_j, \omega'_k) \mathbf{P}_n^\perp \mathbf{b}_n(\omega_j, \omega'_k), \\ \mathbf{b}_n^H(\omega_j, \omega'_k) \mathbf{b}_n(\omega_j, \omega'_k) = 1, n = 1, 2, \dots, N \end{aligned} \right\} \quad (17)$$

对式(17)的优化过程可以分成2个阶段, 在初始阶段 $\mathbf{Q}_n$ 是可逆的,  $\mathbf{b}_n(\omega_j, \omega'_k)$ 可以取矩阵 $(\mathbf{P}_n^\perp, \mathbf{Q}_n)$ 最大特征值对应的单位范数广义特征向量。当 $\mathbf{b}_n(\omega_j, \omega'_k)$ 接近最优时,  $\mathbf{Q}_n$ 的秩变为 $N - 1$ , 令 $\mathbf{U}_0$ 为 $\mathbf{Q}_n$ 最小的 $M - N + 1$ 个特征值对应的特征向量,  $\mathbf{b}_n(\omega_j, \omega'_k)$ 可以表示为

$$\mathbf{b}_n(\omega_j, \omega'_k) = \mathbf{U}_0 \mathbf{z} \quad (18)$$

其中,  $\mathbf{z}$ 是一个权系数向量。将式(18)代入式(17)可知此时对 $\mathbf{b}_n(\omega_j, \omega'_k)$ 的优化等效于式(19)

$$\mathbf{z}_{\text{opt}} = \arg \max_{\mathbf{z}} \mathbf{w}^H \mathbf{U}_0^H \mathbf{P}_n^\perp \mathbf{U}_0 \mathbf{z} \quad (19)$$

其中, 对 $\mathbf{z}$ 的优化可以通过求解 $\mathbf{U}_0^H \mathbf{P}_n^\perp \mathbf{U}_0$ 最大特征值对应的特征向量。

通过上述优化过程可以得到在频点 $\omega_j, j = 1, 2, \dots, J$ 和2次频点 $\omega'_k, k = 1, 2, \dots, K$ 的对角化器 $\mathbf{B}(\omega_j, \omega'_k)$ , 该频点和2次频点对应的信号分离矩阵可以通过求 $\mathbf{B}(\omega_j, \omega'_k)$ 的共轭转置得到

$$\mathbf{W}(\omega_j, \omega'_k) = \mathbf{B}^H(\omega_j, \omega'_k) \quad (20)$$

### 3.3 解排序混乱

本文算法在2次频域做瞬时盲分离产生的排序不确定性会同时影响2次频域和频域上的输出结果, 因此, 在2次频域和频域需要先后解决两次排序混乱问题。假设频域共有 $J$ 个频点, 2次频域共有 $K$ 个频点, 2次频域解排序混乱采用文献<sup>[15]</sup>的思想, 将2次频点上每一次联合对角化的初始矩阵改变为

$$\begin{aligned} \mathbf{W}_0(\omega_j, \omega'_{k+1}) &= \mathbf{W}(\omega_j, \omega'_k), \\ j &= 1, 2, \dots, J, k = 1, 2, \dots, K - 1 \end{aligned} \quad (21)$$

其中,  $\mathbf{W}_0(\omega_j, \omega'_{k+1})$ 表示2次频点 $\omega'_{k+1}$ 的初始矩阵,  $\mathbf{W}(\omega_j, \omega'_k)$ 为前一个2次频点 $\omega'_k$ 的分离矩阵。

经过联合对角化后的分离矩阵由式(8)计算对应2次频点的信号分量估计  $\mathbf{y}(\omega_j, \omega'_k, r)$ ,  $k = 1, 2, \dots, K$ , 对  $\mathbf{y}(\omega_j, \omega'_k, r)$  做逆STFT, 得到频域频点  $\omega_j$  上的分离信号  $\tilde{\mathbf{Y}}(\omega_j)$ 。重复上述过程, 得到频域分离信号  $\tilde{\mathbf{Y}}(\omega_j)$ ,  $j = 1, 2, \dots, J$ 。

频域上的解排序混乱采用文献[16]基于分离信号相似性的方法, 通过相邻频点不同通道彼此之间的信号功率幅度相关性最大化原则, 调整信号排列顺序, 表示为

$$\left. \begin{aligned} \mathbf{A}_{j+1} &= \arg \max_{\mathbf{A}} \sum_{n=1}^N \rho \left[ |\mathbf{y}'_n(\omega_{j+1})|^2, |\mathbf{y}_n(\omega_j)|^2 \right] \\ \mathbf{Y}'(\omega_{j+1}) &= \mathbf{A} \tilde{\mathbf{Y}}(\omega_{j+1}), n = 1, 2, \dots, N, \\ & j = 1, 2, \dots, J-1 \end{aligned} \right\} (22)$$

其中,  $\tilde{\mathbf{Y}}(\omega_{j+1})$  是顺序调整前频域频点  $\omega_{j+1}$  上的分离信号,  $\mathbf{A}$  是所有可能的置换矩阵,  $\mathbf{y}_n(\omega_j)$  是频域频点  $\omega_j$  上的第  $n$  个分离信号分量,  $\mathbf{Y}'(\omega_{j+1})$  是  $\mathbf{Y}'(\omega_{j+1})$  的第  $n$  个分离信号分量,  $\mathbf{A}_{j+1}$  表示在频点  $\omega_{j+1}$  上的解排序混乱置换矩阵,  $\rho$  表示向量的相关系数。通过式(23)可以得到频域频点  $\omega_{j+1}$  上经过解排序混乱后的分离信号

$$\mathbf{Y}(\omega_{j+1}) = \mathbf{A}_{j+1} \tilde{\mathbf{Y}}(\omega_{j+1}) \quad (23)$$

将调整顺序后的信号  $\mathbf{Y}(\omega_j)$ ,  $j = 1, 2, \dots, J$  做第2次逆STFT, 最终得到时域上的分离信号。

#### 4 仿真实验和结果分析

仿真实验采用模拟信号和声音信号作为源信号, 实验1—实验3使用模拟源信号, 实验4和实验5使用声音信号。实验最多同时存在4个信号源, 其中, 模拟信号源表示为

$$S_i(t) = A_i \sin(\omega_i t), i = 1, 2, 3, 4 \quad (24)$$

其中,  $A_i$  为服从标准正态分布  $N(0, 1)$  的随机变量,  $\omega_i$  为角频率(实验从  $2\pi/3, \pi/2, \pi/3, \pi/4$  中选取), 模拟信号源采样率为8 kHz, 数据长度取100000。声

音信号采样率为8 kHz, 采样点数为50000, 4种声音分别为警笛声、男性话音、女性话音和男性歌声。混合滤波器为  $L$  阶FIR滤波器, 其脉冲响应服从  $U(0, 1)$  的均匀分布。声音信号的最大近似平稳时间大约为15~30 ms, 对应混合滤波器长度在120~240。可以证明模拟信号与声音信号一样, 也具有短时平稳特性, 由实验所取频率  $\omega_i$  范围可以计算出最大近似平稳时间大约为15~40 ms, 与声音信号最大近似平稳时间相近。盲分离性能采用平均信干比(SIR)来描述, 定义为

$$\text{SIR} = \frac{1}{N} \sum_{n=1}^N \frac{\max_j \tilde{g}_{nj}}{\sum_{j=1}^N \tilde{g}_{nj} - \max_j \tilde{g}_{nj}}, n = 1, 2, \dots, N \quad (25)$$

其中,  $\tilde{g}_{nj} = \sum_{f=0}^{\text{NFFT}-1} |g_{nj}(\omega_f, \omega'_f)|^2$ , NFFT是窗函数长度,  $N$  是源信号个数,  $g_{nj}(\omega_f, \omega'_f)$  是全局矩阵  $\mathbf{G}(\omega_f, \omega'_f) = \mathbf{W}(\omega_f, \omega'_f) \mathbf{A}(\omega_f, \omega'_f)$  的第  $(n, j)$  个元素。

**实验1** 比较了当源信号数  $N$  为2时, 非深度卷积频域独立分量分析(Independent Component Analysis, ICA)算法<sup>[6]</sup>、基于ICA的2次频域算法<sup>[9]</sup>和本文的联合对角化(JD)算法的SIR性能随混合滤波器长度的变化情况, 分别进行100次独立实验, 计算平均性能, 仿真结果如图1(a)所示。可以看出, 初始混合滤波器长度为150时, 只做1次STFT(1STFT)的非深度卷积频域算法已经接近失效, 随着混合滤波器长度的增大, 非深度卷积算法完全失效, 而两种深度卷积模型下的算法性能保持平稳, 相比于文献[9]方法, 本文方法在性能上始终保持一定的优势。

**实验2** 比较了当源信号数  $N$  分别为2, 3, 4时, 文献[9]方法和本文方法的SIR性能随混合滤波器长度的变化情况, 分别进行100次独立实验, 计算平均性能, 仿真结果如图1(b)所示。可以看出, 在相

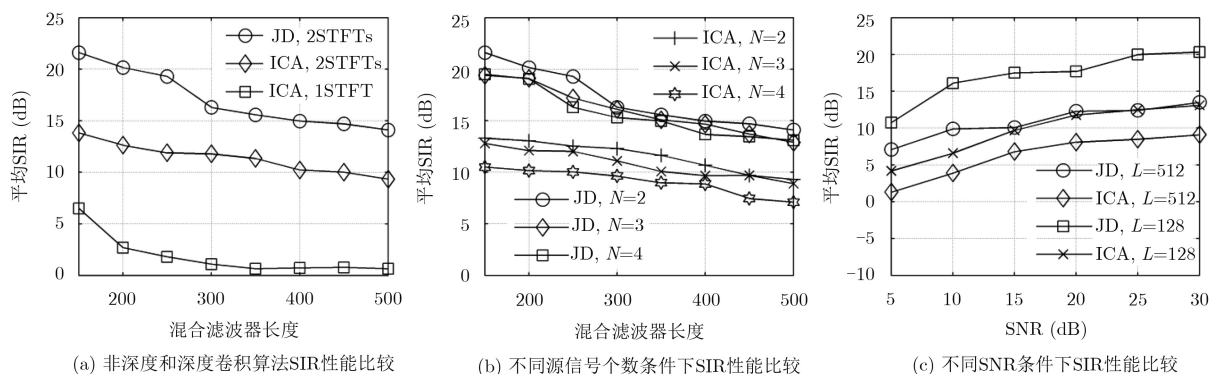


图1 算法性能比较

同条件下, 本文方法相比较文献[9]方法整体SIR有较大提高。在混合滤波器长度固定时, 文献[9]方法平均SIR随着源信号数增多会有一定程度的下降, 而本文方法则变化不明显。

**实验3** 比较了当源信号数 $N$ 为2, 混合信号上叠加不同强度的加性噪声时, 文献[9]方法和本文方法的SIR性能随信噪比(SNR)从5~30 dB的变化情况。混合滤波器长度分别取 $L = 128$ 和 $L = 512$ , 分别进行100次独立实验, 计算平均性能, 仿真结果如图1(c)所示。可以看出, 两种方法在SNR大于5 dB时, 均可以保证分离算法有效, 但是, 在相同条件下本文所用联合对角化(JD)方法的分离性能具有较大优势。

**实验4** 采用声音信号, 给出了当源信号数 $N$ 分别为2,3,4, 混合滤波器长度 $L = 512$ 时, 本文方法做排序调整前和去排序混乱后SIR性能的对

比, 分别进行100次独立试验, 计算平均性能, 结果如表1所示。从表1中可以看出, 不做排序调整, 算法的平均SIR仅为2.03 dB, 此时的分离性能接近失效, 而经过排序调整后的平均SIR约为14.9 dB, 验证了本文所述解排序混乱的必要性。

表1 不同源信号数情况下调整排序混乱前后SIR对比(dB)

分离性能	信源个数		
	2	3	4
调序前	3.4	1.5	1.2
调序后	15.4	14.9	14.3

**实验5** 采用2路声音信号, SNR = 30 dB, 混合滤波器长度为 $L = 512$ , 图2给出了源信号、混合信号和分离信号的时域波形图。可以看出, 本文方法对于深度卷积声信号具有很好的分离性能。

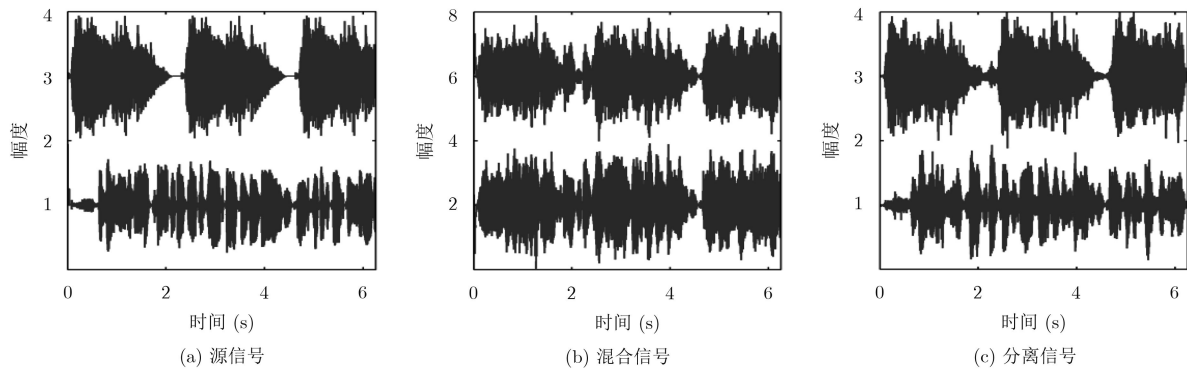


图2 声音信号时域波形

## 5 结论

多信号源且带有加性噪声的深度卷积混合信号会导致盲分离性能严重下降, 针对这一问题, 本文提出一种2次频域的联合对角化方法, 将深度卷积混合转变为2次频域的瞬时混合, 通过非酉联合对角化实现信号盲分离。仿真实验结果验证了本文方法可以对深度卷积声信号进行有效分离, 相比较现有方法, 当源信号数增多或者混合信号存在加性噪声时, 本文方法可以具有更好的分离性能。

### 参考文献

- [1] BUCHNER H, AICHNER R, and KELLERMANN W. A generalization of blind source separation algorithms for convolutive mixtures based on second-order statistics[J]. *IEEE Transactions on Speech and Audio Processing*, 2005, 13(1): 120–134. doi: [10.1109/TSA.2004.838775](https://doi.org/10.1109/TSA.2004.838775).
  - [2] 张华, 冯大政, 庞继勇. 基于二阶统计量的语音信号时域卷积盲分离算法[J]. *电子与信息学报*, 2010, 32(5): 1083–1087. doi: [10.3724/SP.J.1146.2009.00574](https://doi.org/10.3724/SP.J.1146.2009.00574).
  - [3] CHERRAK O, GHENNIoui H, THIRION-MOREAU N, et al. Preconditioned optimization algorithms solving the problem of the non unitary joint block diagonalization: Application to blind separation of convolutive mixtures[J]. *Multidimensional Systems and Signal Processing*, 2018, 29(4): 1373–1396. doi: [10.1007/s11045-017-0506-8](https://doi.org/10.1007/s11045-017-0506-8).
  - [4] RAHBAR K and REILLY J P. A frequency domain method for blind source separation of convolutive audio mixtures[J]. *IEEE Transactions on Speech and Audio Processing*, 2005, 13(5): 832–844. doi: [10.1109/TSA.2005.851925](https://doi.org/10.1109/TSA.2005.851925).
  - [5] 付卫红, 张琮. 基于步长自适应的独立向量分析卷积盲分离算法[J]. *电子与信息学报*, 2018, 40(9): 2158–2164. doi: [10.11999/JEIT171156](https://doi.org/10.11999/JEIT171156).
- FU Weihong and ZHANG Cong. Independent vector analysis convolutive blind separation algorithm based on

ZHANG Hua, FENG Dazheng, and PANG Jiyong. A time-domain blind convolutive separation algorithm for speech signals based on second-order statistic[J]. *Journal of Electronics & Information Technology*, 2010, 32(5): 1083–1087. doi: [10.3724/SP.J.1146.2009.00574](https://doi.org/10.3724/SP.J.1146.2009.00574).

- step-size adaptive[J]. *Journal of Electronics & Information Technology*, 2018, 40(9): 2158–2164. doi: [10.11999/JEIT171156](https://doi.org/10.11999/JEIT171156).
- [6] YU Meng, XIN J, QI Yingyong, *et al.* A nonlocally weighted soft-constrained natural gradient algorithm and blind separation of strongly reverberant speech mixtures[C]. The 43rd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, USA, 2009: 346–350. doi: [10.1109/ACSSC.2009.5470080](https://doi.org/10.1109/ACSSC.2009.5470080).
- [7] WANG Lin, DING Heping, and YIN Fuliang. Combining superdirective beamforming and frequency-domain blind source separation for highly reverberant signals[J]. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010, 2010(1): 797962. doi: [10.1155/2010/797962](https://doi.org/10.1155/2010/797962).
- [8] SANZ-ROBINSON J, HUANG Liechao, MOY T, *et al.* Robust blind source separation in a reverberant room based on beamforming with a large-aperture microphone array[C]. IEEE International Conference on Acoustics, Speech and Signal Processing, Shanghai, China, 2016: 440–444. doi: [10.1109/ICASSP.2016.7471713](https://doi.org/10.1109/ICASSP.2016.7471713).
- [9] HIROE A. Blind Vector Deconvolution: Convolutional Mixture Models in Short-time Fourier Transform Domain[M]. DAVIES M E, JAMES C J, ABDALLAH S A, *et al.* Independent Component Analysis and Signal Separation. Berlin, Germany, 2007: 471–479. doi: [10.1007/978-3-540-74494-8\\_59](https://doi.org/10.1007/978-3-540-74494-8_59).
- [10] 顾凡, 王惠刚, 李虎雄. 一种强混响环境下的盲语音分离算法[J]. 信号处理, 2011, 27(4): 534–540. doi: [10.3969/j.issn.1003-0530.2011.04.009](https://doi.org/10.3969/j.issn.1003-0530.2011.04.009).  
GU Fan, WANG Huigang, and LI Huxiong. A blind speech separation algorithm with strong reverberation[J]. *Signal Processing*, 2011, 27(4): 534–540. doi: [10.3969/j.issn.1003-0530.2011.04.009](https://doi.org/10.3969/j.issn.1003-0530.2011.04.009).
- [11] PHAM D T. Joint approximate diagonalization of positive definite hermitian matrices[J]. *SIAM Journal on Matrix Analysis and Applications*, 2001, 22(4): 1136–1152. doi: [10.1137/S089547980035689X](https://doi.org/10.1137/S089547980035689X).
- [12] ZENG TiaoJun, GOU Li, and WU Junhang. Non-orthogonal joint diagonalization algorithm preventable ill conditioned solutions for blind source separation[J]. *Optik*, 2017, 140: 145–150. doi: [10.1016/j.ijleo.2017.04.011](https://doi.org/10.1016/j.ijleo.2017.04.011).
- [13] ASAMIZU T, SAITO S, OISHI K, *et al.* Overdetermined blind source separation using approximate joint diagonalization[C]. The 60th IEEE International Midwest Symposium on Circuits and Systems, Boston, USA, 2017: 168–171. doi: [10.1109/MWSCAS.2017.8052887](https://doi.org/10.1109/MWSCAS.2017.8052887).
- [14] ZHANG Weitao and SUN Jinling. Nonunitary joint diagonalization for overdetermined convolutive blind signal separation[C]. The 26th European Signal Processing Conference, Rome, Italy, 2018: 1232–1236. doi: [10.23919/EUSIPCO.2018.8553132](https://doi.org/10.23919/EUSIPCO.2018.8553132).
- [15] 张天骐, 张华伟, 刘董华, 等. 基于区域增长校正的频域盲源分离排序算法[J]. 电子与信息学报, 2019, 41(3): 580–587. doi: [10.11999/JEIT180386](https://doi.org/10.11999/JEIT180386).  
ZHANG Tianqi, ZHANG Huawei, LIU Donghua, *et al.* Frequency domain blind source separation permutation algorithm based on regional growth correction[J]. *Journal of Electronics & Information Technology*, 2019, 41(3): 580–587. doi: [10.11999/JEIT180386](https://doi.org/10.11999/JEIT180386).
- [16] JAYARMAN D S, SITARAMAN G, and SESHADRI R. Blind source separation of acoustic mixtures using time-frequency domain independent component analysis[C]. The 8th International Conference on Communication Systems, Singapore, Singapore, 2002: 1016–1019. doi: [10.1109/ICCS.2002.1183286](https://doi.org/10.1109/ICCS.2002.1183286).
- 李 扬: 男, 1987年生, 博士生, 研究方向为盲信号处理。  
张伟涛: 男, 1983年生, 副教授, 硕士生导师, 研究方向为盲信号处理、语音信号处理。  
楼顺天: 男, 1962年生, 教授, 博士生导师, 研究方向为神经网络信息处理与应用、模糊信息处理与应用、盲信号处理、现代信号智能处理、智能控制技术。