自动化学报
ACTA AUTOMATICA SINICA

面向卷积混叠环境下的盲源分离新方法

解元 邹涛 孙为军 谢胜利

Novel Blind Source Separation Method for Convolutive Mixed Environment

XIE Yuan, ZOU Tao, SUN Wei-Jun, XIE Sheng-Li

在线阅读 View online: https://doi.org/10.16383/j.aas.c211207

您可能感兴趣的其他文章

基于卷积非负矩阵部分联合分解的强噪声单声道语音分离

Monaural Speech Separation by Means of Convolutive Nonnegative Matrix Partial Co-factorization in Low SNR Condition

自动化学报. 2020, 46(6): 1200-1209 https://doi.org/10.16383/j.aas.c180065

基于稀疏表示和结构自相似性的单幅图像盲解卷积算法

Single Image Blind Deconvolution Using Sparse Representation and Structural Self-similarity

自动化学报. 2017, 43(11): 1908-1919 https://doi.org/10.16383/j.aas.2017.c160357

基于多视图矩阵分解的聚类分析

Matrix Factorization for Multi-view Clustering

自动化学报. 2018, 44(12): 2160-2169 https://doi.org/10.16383/j.aas.2018.c160636

无线网络环境下数据驱动混合选别浓密过程双率控制方法

Data-driven Dual-rate Control for Mixed Separation Thickening Process in a Wireless Network Environment

自动化学报. 2019, 45(6): 1122-1135 https://doi.org/10.16383/j.aas.c180202

基于PLS交叉积矩阵非相似度分析的MPC性能监控与诊断

MPC Performance Monitoring and Diagnosis Based on Dissimilarity Analysis of PLS Cross-product Matrix

自动化学报. 2017, 43(2): 271-279 https://doi.org/10.16383/j.aas.2017.c150782

考虑多源不确定信息的路网交通拥堵状态辨识方法

Traffic Congestion Status Identification Method for Road Network with Multi-source Uncertain Information

自动化学报. 2018, 44(3): 533-544 https://doi.org/10.16383/j.aas.2018.c160373

面向卷积混叠环境下的盲源分离新方法

解 元[1, 2] 邹 涛[1] 孙为军[3, 4] 谢胜利[5, 6]

摘 要 卷积混叠环境下的盲源分离 (Blind source separation, BSS) 是一个极具挑战性和实际意义的问题. 本文在独立分量分析框架下, 建立非负矩阵分解 (Nonnegative matrix factorization, NMF) 模型, 设计新的优化目标函数, 通过严格的数学理论推导, 得到新的模型参数更新规则; 并对解混叠矩阵进行标准化处理, 避免幅度歧义性问题; 在源信号的重构阶段, 通过实时更新非负矩阵分解模型参数, 避免源信号的排序歧义性问题. 实验结果验证了所提算法在分离中英文语音混叠信号、音乐混叠信号时的有效性和优越性.

关键词 盲源分离, 卷积混叠, 独立分量分析, 非负矩阵分解

引用格式 解元, 邹涛, 孙为军, 谢胜利. 面向卷积混叠环境下的盲源分离新方法. 自动化学报, 2023, **49**(5): 1062−1072

DOI 10.16383/j.aas.c211207

Novel Blind Source Separation Method for Convolutive Mixed Environment

XIE Yuan[1, 2] ZOU Tao[1] SUN Wei-Jun[3, 4] XIE Sheng-Li[5, 6]

Abstract Blind source separation (BSS) for convolutive mixed environment is a challenging and practical topic. In this paper, a nonnegative matrix factorization (NMF) model is established based on the framework of independent component analysis, and a new optimization objective function is designed. Through strict mathematical theory derivation, new model parameters update rules are obtained, and the demixing matrix is standardized to avoid the scale ambiguity. In the stage of source reconstruction, the permutation ambiguity can be avoided by updating the parameters of the NMF model in real time. Experimental results verify the effectiveness and superiority of the proposed algorithm in separating Chinese speech mixtures, English speech mixtures, and music signal mixtures.

Key words Blind source separation (BSS), convolutive mixtures, independent component analysis, nonnegative matrix factorization (NMF)

Citation Xie Yuan, Zou Tao, Sun Wei-Jun, Xie Sheng-Li. Novel blind source separation method for convolutive mixed environment. *Acta Automatica Sinica*, 2023, **49**(5): 1062−1072

盲源分离 (Blind source separation, BSS) 是指在对混叠信道信息未知的情况下, 仅根据接收的混叠信号分离出源信号[1−2]. 盲源分离起源于经典的“鸡尾酒会问题”[3], 即在酒会派对上, 多人同时说话, 利用麦克风接收混叠的语音信号, 通过设计盲源分离方法可以把接收到的混叠信号独立地分离出来. 由于盲源分离方法只需要根据接收信号就可以分离出源信号, 因此, 具有独特的分离特性. 目前, 已经应用于多个领域, 例如: 音频信号处理[4−5]、生物信号处理[6−8]、图形信号处理[9−10]、故障盲检测[11−12] 等领域. 在音频信号处理中, 由于接收的混叠语音信号伴随着回响以及噪声等一系列不确定因素, 给源信号的分离带来了很大的困难.

盲分离是信号处理领域中一种高效的信号分离方法, 传统的盲分离方法是基于线性的混叠模型而设计的[13−15]. 然而, 线性混叠模型并不能很精确地模拟现实问题, 进而发展了卷积混叠模型, 由于卷积模型可以更好地描述实际问题, 所以目前在卷积混叠模型下研发盲源分离算法成为了主流. 主要分为三类: 超定卷积混叠、正定卷积混叠、欠定卷积混

收稿日期 2021-12-18 录用日期 2022-10-18
Manuscript received December 18, 2021; accepted October 18, 2022
国家重点研发计划 (2018YFB1802400), 国家自然科学基金 (62003095, 52171331) 资助
Supported by National Key Research and Development Project (2018YFB1802400) and National Natural Science Foundation of China (62003095, 52171331)
本文责任编委 杨健
Recommended by Associate Editor YANG Jian
1. 广州大学机械与电气工程学院 广州 510006 2. 物联网智能信息处理与系统集成教育部重点实验室 广州 510006 3. 广东省物联网信息技术重点实验室 广州 510006 4. 智能检测与制造物联教育部重点实验室 广州 510006 5. 基于物联网技术的离散制造智能化学科创新引智基地 广州 510006 6. 粤港澳离散制造智能化联合实验室 广州 510006
1. School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006 2. Key Laboratory of Intelligent Information Processing and System Integration of Internet of Things, Ministry of Education, Guangzhou 510006 3. Guangdong Provincial Key Laboratory of Information Technology of Internet of Things, Guangzhou 510006 4. Key Laboratory of Intelligent Detection and the Internet of Things in Manufacturing, Ministry of Education, Guangzhou 510006 5. Discrete Manufacturing Intelligence Discipline Innovation and Talent Introduction Base Based on Internet of Things Technology, Guangzhou 510006 6. Guangdong-Hong Kong-Macao Joint Laboratory for Smart Discrete Manufacturing, Guangzhou 510006

叠. 即, 当源信号数目小于传感器数目时是超定混叠; 当源信号数目等于传感器数目时是正定混叠; 当源信号数目大于传感器数目时是欠定混叠. 针对这三种混叠, 也有相应的论文. 例如, 基于奇异值分解的超定混叠盲源分离[16], 考虑的是不含噪声的线性混叠模型, 利用互信息与微分熵的关系, 提出基于独立分量分析的代价函数, 推导了一般梯度学习算法. 基于时频掩蔽技术的正定混叠盲源分离算法[17], 考虑的是不含噪声的卷积混叠模型, 其代价函数基于最大化对数似然估计. 基于超平面法矢量的欠定盲源分离算法[18], 以及基于平行因子分解的欠定混叠盲源分离算法[19] 等. 然而, 在盲解卷问题中, 存在两个主要问题: 一是幅度歧义性问题; 二是排序歧义性问题. 幅度歧义性指的是在每个时频点上所估计的源信号的增益大小不确定; 排序歧义性指的是在每个时频点上所估计的源信号顺序发生混乱, 难于与源信号顺序保持一致. 针对幅度歧义性问题, 盲源分离的最小失真原理[20] 是比较成熟的方法, 而且幅度歧义性问题不如排序歧义性问题那么严重. 针对排序歧义性问题, 比较流行的解决方法是基于方向到达估计排序算法和频率间相关排序算法[21−22]. 此外, 还有一种基于 Tucker 张量分解方法, 利用耦合频率消除卷积盲源分离中的排序歧义性[23]. 但是, 到目前为止, 所提出的算法只是在一定程度上缓解了排序歧义性所带来的盲源分离性能问题, 仍然没有有效手段可以彻底解决排序歧义性问题.

如何设计更好的优化算法提高盲分离性能一直是盲源分离研究的关键. 时频域方法是比较流行的方法[24−26], 它利用短时傅里叶变换把时域上的卷积混叠信号转换到频域上进行分离. 然后, 把分离后的源信号利用逆短时傅里叶变换转换到时域上, 从而实现分离. 常用的稀疏分析方法在盲源分离中发挥重要作用[27−28], 利用信号在时频域上的稀疏性, 假设在每个时频点上有且只有一个源成分是活跃的, 设计相应的稀疏盲源分离算法重构源信号. 目前, $L_{1/2}$ 范数正则化在稀疏性中扮演重要角色[29−30], 它可以平衡 L_0 范数与 L_1 范数之间的稀疏性. 同时, 为了获得更加稀疏的表示, L_p $(0 \le p < 1)$ 范数正则化通过调整参数 p 值来近似真正的稀疏补偿项[31−32], 从而提高源信号的分离性能. 然而, 参数 p 的值不是固定的, 需要根据具体数据集进行适当的调参来获得最优解.

非负矩阵分解 (Nonnegative matrix factorization, NMF) 是当前流行的一种分离方法[33−35], 文献 [36] 将独立向量分析与非负矩阵分解相结合, 设计了一种正定卷积混叠盲分离算法. 文献 [37−38]将期望最大化与非负矩阵分解相结合, 设计了欠定卷积混叠信号的盲分离. 文献 [39] 设计了一种快速多通道非负矩阵分解盲分离方法. 上述方法主要是基于低秩空间协方差模型设计的. 而基于满秩空间协方差模型的方法在解决盲分离问题中也发挥了重要作用[40−41]. 另外, 张量分解理论[42] 作为非负矩阵分解的拓展, 已应用于盲分离问题中[43−44]. 然而, 该类方法由于计算复杂度较高, 导致计算速度较慢.

为了突破传统盲分离算法的局限性, 本文设计一种新的面向卷积混叠环境下的盲源分离算法. 通过设计带权重的非负矩阵分解模型, 基于 Itakura-Saito 散度, 结合了独立分量分析和非负矩阵分解, 构建新的代价函数. 通过严格的数学理论推导和优化求解, 获得新的模型参数更新规则, 在每次迭代更新过程中对解混叠矩阵进行标准化处理, 可以避免幅度歧义性问题; 在源信号的重构阶段, 通过实时更新非负矩阵分解模型参数, 并将信号源的空间特性在所有频点上联合优化, 从而在每个时频点上更好地重构源信号, 避免了排序歧义性问题. 同时, 假设源信号在每个时频点上满足独立复高斯分布, 在独立分量分析框架下, 源信号的协方差矩阵为对角化矩阵, 对其建立带权重的非负矩阵分解新模型, 利用梯度下降法得到非负矩阵分解模型参数更新规则, 降低了计算复杂度, 提高了算法的计算速度.

本文的创新点概括如下:

1) 设计一个新的优化目标函数, 推导出新的模型参数更新规则, 并给出严格的数学理论证明, 进而提出一种高效的卷积盲源分离算法.

2) 在独立分量分析框架下, 建立非负矩阵分解新模型, 对解混叠矩阵进行实时更新学习, 避免了排序歧义性问题, 提高了盲源分离性能, 同时降低了计算复杂度.

本文实验数据来自于公开的音频信号公共数据集, 实验测试了中文语音混叠信号、英文语音混叠信号、以及音乐混叠信号在不同混响下的盲源分离性能, 通过对比几种目前比较流行的盲源分离算法, 验证了本文所提算法的有效性以及优越性.

本文的结构安排如下: 第 1 节主要介绍卷积混叠系统模型; 第 2 节是本文的核心部分, 详细介绍了所提算法的具体细节, 并给出严格的数学理论推导, 设计了新的盲源分离算法; 第 3 节给出详细的仿真实验以及对实验结果进行分析; 第 4 节对本文做总结, 并对未来工作进行展望.

1 卷积混叠系统模型

考虑多通道时域卷积混叠信号模型

$$\boldsymbol{x}(t) = A * \boldsymbol{s}(t) + \boldsymbol{v}(t) = \sum_{\tau=0}^{L-1} A(\tau)\boldsymbol{s}(t-\tau) + \boldsymbol{v}(t) \tag{1}$$

其中, $*$ 代表卷积, 时间变量 $t = 1, \cdots, T$ 是连续的, $\boldsymbol{x}(t) = [x_1(t), \cdots, x_p(t)]^{\mathrm{T}}$ 是用 p 个传感器在时刻 t 记录的混叠信号, $\boldsymbol{s}(t) = [s_1(t), \cdots, s_Q(t)]^{\mathrm{T}}$ 是 Q 个源信号, $A(\tau)$ 是在时间延迟 τ 下的混叠系统, L 是空间脉冲响应长度, $\boldsymbol{v}(t)$ 是均值为零、方差为 σ^2 的高斯白噪声. 卷积混叠信号盲源分离是指仅根据接收到的混叠信号 $\boldsymbol{x}(t)$, 分离得到源信号 $\boldsymbol{s}(t)$.

为了解决上述卷积混叠信号盲源分离问题, 使用短时傅里叶变换将时域卷积混叠信号模型 (1) 转换到频域, 得

$$\boldsymbol{x}_{fn} = A_f \boldsymbol{s}_{fn} + \boldsymbol{v}_{fn} \tag{2}$$

其中, $f = 1, \cdots, F$ 是频点指数, $n = 1, \cdots, N$ 是时间窗指数, $\boldsymbol{x}_{fn} = [x_{1,fn}, \cdots, x_{p,fn}]^{\mathrm{T}}$ 和 $\boldsymbol{s}_{fn} = [s_{1,fn}, \cdots, s_{Q,fn}]^{\mathrm{T}}$ 分别是混叠信号 $\boldsymbol{x}(t)$ 和源信号 $\boldsymbol{s}(t)$ 的短时傅里叶变换, A_f 是混叠矩阵, $\boldsymbol{v}_{fn}$ 是包括高斯白噪声在内的模型近似误差.

假设空间脉冲响应长度远远小于短时傅里叶变换窗函数长度, 则式 (2) 的线性近似模型是成立的. 同时, 假设在每个时频点上的源是相互独立的, 满足复高斯分布, 噪声是平稳的、且与源信号相互独立.

本文的目标: 在时频域上, 设计卷积盲源分离算法, 仅根据在时域上接收到的混叠信号 $\boldsymbol{x}(t)$, 利用短时傅里叶变换得到 $\boldsymbol{x}_{fn}$, 在频域上分离得到源信号 $\boldsymbol{s}_{fn}$. 然后, 利用短时傅里叶逆变换获取时域源信号 $\boldsymbol{s}(t)$, 从而实现源信号的盲分离.

2 卷积混叠信号盲源分离算法

假设传感器的数目等于源信号数目, 即 $P = Q$, 混叠矩阵 A_f 是可逆的, 定义解混叠矩阵 $W_f = A_f^{-1}$, 则分离的源信号可表示为

$$\boldsymbol{s}_{fn} \approx \boldsymbol{y}_{fn} = W_f \boldsymbol{x}_{fn} \tag{3}$$

即

$$\boldsymbol{x}_{fn} = W_f^{-1} \boldsymbol{y}_{fn} \tag{4}$$

计算下式

$$\begin{aligned} X_{fn} = \boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\mathrm{T}} = \\ (A_f\boldsymbol{s}_{fn} + \boldsymbol{v}_{fn})(A_f\boldsymbol{s}_{fn} + \boldsymbol{v}_{fn})^{\mathrm{T}} = \\ A_f\boldsymbol{s}_{fn}\boldsymbol{s}_{fn}^{\mathrm{T}}A_f^{\mathrm{T}} + A_f\boldsymbol{s}_{fn}\boldsymbol{v}_{fn}^{\mathrm{T}} + \\ \boldsymbol{v}_{fn}\boldsymbol{s}_{fn}^{\mathrm{T}}A_f^{\mathrm{T}} + \boldsymbol{v}_{fn}\boldsymbol{v}_{fn}^{\mathrm{T}} \end{aligned} \tag{5}$$

其中, $(\cdot)^{\mathrm{T}}$ 是 Hermitian 转置, 假设在每个时频点上的源成分与噪声是相互独立的, 则式 (5) 可以近似为

$$X_{fn} = \hat{X}_{fn} = A_f\boldsymbol{s}_{fn}\boldsymbol{s}_{fn}^{\mathrm{T}}A_f^{\mathrm{T}} = A_f D_{fn}^{\mathrm{T}} A_f^{\mathrm{T}} \tag{6}$$

其中, X_{fn} 是 Hermitian 正定矩阵, 它表示在时频点 (f, n) 处观测信号的瞬时协方差, X 是一个四阶张量, 假设源信号在每个时频点上满足独立复高斯分布, 在独立分量分析框架下, 源信号的协方差矩阵 $D_{fn} = \boldsymbol{s}_{fn}\boldsymbol{s}_{fn}^{\mathrm{T}}$ 是 Hermitian 正定的对角化矩阵, 可表示为

$$D_{fn} = \begin{bmatrix} d_{fn,1} & 0 & \cdots & 0 \\ 0 & d_{fn,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_{fn,Q} \end{bmatrix} \tag{7}$$

针对 D_{fn} 中的对角线的非负元素, 对其建立带权重的非负矩阵分解新模型, 可得

$$d_{fn,q} = \sum_k z_{qk} t_{fk} v_{kn} \tag{8}$$

其中, $z_{qk} \in \{0, 1\}$ 是一个潜在变量, 它表示第 k 个基是否属于第 q 个源, 如果属于, 则 $z_{qk} = 1$, 否则 $z_{qk} = 0$; t_{fk} 和 v_{kn} 分别是基矩阵 $T \in \mathbf{R}_{\geq 0}^{F\times K}$ 和激活矩阵 $V \in \mathbf{R}_{\geq 0}^{K\times N}$ 的非负元素, $k = 1, \cdots, K$ 是指数, K 被设置为比 N 和 F 小很多的值, 即 $FK + KN \ll FN$, 它的估计可利用自动相关性判定方法获得[45]. 因此, 这是一个低秩近似.

根据 X_{fn} 和 $\hat{X}_{fn}$ 的 Itakura-Saito 散度

$$d_{IS}(X_{fn}|\hat{X}_{fn}) = \frac{X_{fn}}{\hat{X}_{fn}} - \ln\frac{X_{fn}}{\hat{X}_{fn}} - 1 \tag{9}$$

把式 (5) 和式 (6) 代入式 (9) 中, 得

$$\begin{aligned} d_{IS}(X_{fn}|\hat{X}_{fn}) = \frac{\boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\mathrm{T}}}{A_f D_{fn}^{\mathrm{T}} A_f^{\mathrm{T}}} - \ln\frac{\boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\mathrm{T}}}{A_f D_{fn}^{\mathrm{T}} A_f^{\mathrm{T}}} - 1 = \\ \frac{\boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\mathrm{T}}}{A_f D_{fn}^{\mathrm{T}} A_f^{\mathrm{T}}} + \ln(\det(A_f D_{fn}^{\mathrm{T}} A_f^{\mathrm{T}})) - \\ \ln(\boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\mathrm{T}}) - 1 \end{aligned} \tag{10}$$

定义代价函数

$$\begin{aligned} F = \sum_{f,n}\left(\mathrm{tr}\left(\frac{\boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\mathrm{T}}}{A_f D_{fn}^{\mathrm{T}} A_f^{\mathrm{T}}}\right) + \right. \\ \left. \ln(\det(A_f D_{fn}^{\mathrm{T}} A_f^{\mathrm{T}}))\right) + \mathrm{const} \end{aligned} \tag{11}$$

其中, tr 代表矩阵的迹, det 代表行列式, const 是常数项. 通过最小化式 (9), 假设源信号在每个时频

点上满足独立复高斯分布, 对相互独立的源进行建模. 把式 (4) 和 $W_f = A_f^{-1}$ 代入式 (11) 中, 得

$$F = \sum_{f,n}\left(\mathrm{tr}\left(\frac{W_f^{-1}\boldsymbol{y}_{fn}\boldsymbol{y}_{fn}^{\mathrm{T}}(W_f^{\mathrm{T}})^{-1}}{W_f^{-1}D_{fn}(W_f^{\mathrm{T}})^{-1}}\right) + \ln|\det W_f^{-1}| + \ln(\det D_{fn}) + \ln|\det(W_f^{\mathrm{T}})^{-1}|\right) + \mathrm{const} = \sum_{f,n}\left(\mathrm{tr}\left(\frac{\boldsymbol{y}_{fn}\boldsymbol{y}_{fn}^{\mathrm{T}}}{D_{fn}}\right) + 2\ln|\det W_f^{-1}| + \ln(\det D_{fn})\right) + \mathrm{const} = \sum_{f,n}\left(\sum_q \frac{|y_{fn,q}|^2}{d_{fn,q}} - 2\ln|\det W_f| + \sum_q \ln d_{fn,q}\right) + \mathrm{const} \tag{12}$$

把式 (8) 代入式 (12) 中, 得

$$F = \sum_{f,n}\left(\sum_q \frac{|y_{fn,q}|^2}{\sum_k z_{qk}t_{fk}v_{kn}} - 2\ln|\det W_f| + \sum_q\left(\ln\sum_k z_{qk}t_{fk}v_{kn}\right)\right) + \mathrm{const} \tag{13}$$

通过最小化代价函数 (13) 可以推导得到 z_{qk}, t_{fk} 和 v_{kn} 的更新规则. 然而, 式 (13) 中的第一项是个凸函数, 第二项是个凹函数, 整体是一个非凸的优化问题, 求解相对困难.

因此, 针对式 (13) 中的第一项凸函数, 利用詹森 (Jensen) 不等式, 得

$$\frac{1}{\sum_k z_{qk}t_{fk}v_{kn}} \le \sum_k \frac{\alpha_{fn,k}^2}{z_{qk}t_{fk}v_{kn}} \tag{14}$$

其中, 辅助变量 $\alpha_{fn,k} \ge 0$ 且满足 $\sum_k \alpha_{fn,k} = 1$. 当且仅当 $\alpha_{fn,k} = (z_{qk}t_{fk}v_{kn})/(\sum_{k'} z_{qk'}t_{fk'}v_{k'n})$ 时, 等号成立.

针对式 (13) 中的第二项凹函数, 利用切线不等式, 得

$$\ln\frac{\sum_k z_{qk}t_{fk}v_{kn}}{\beta_{fn}} \le \frac{\sum_k z_{qk}t_{fk}v_{kn}}{\beta_{fn}} - 1 \tag{15}$$

变形为

$$\ln\left(\sum_k z_{qk}t_{fk}v_{kn}\right) - \ln\beta_{fn} \le \frac{1}{\beta_{fn}}\sum_k z_{qk}t_{fk}v_{kn} - 1 \tag{16}$$

即

$$\ln\left(\sum_k z_{qk}t_{fk}v_{kn}\right) \le \frac{1}{\beta_{fn}}\sum_k z_{qk}t_{fk}v_{kn} + \ln\beta_{fn} - 1 \tag{17}$$

其中, 辅助变量 $\beta_{fn} \ge 0$, 当且仅当 $\beta_{fn} = \sum_{k'} z_{qk'}t_{fk'}v_{k'n}$ 时, 等号成立.

然后, 定义代价函数 (13) 的辅助函数

$$F \le F^+ = \sum_{f,n}\left(\sum_{q,k}\frac{|y_{fn,q}|^2\alpha_{fn,k}^2}{\sum_k z_{qk}t_{fk}v_{kn}} - 2\ln|\det W_f| + \sum_q\left(\frac{1}{\beta_{fn}}\sum_k z_{qk}t_{fk}v_{kn} + \ln\beta_{fn} - 1\right)\right) + \mathrm{const} \tag{18}$$

因此, 非凸的优化问题转化为凸优化问题, 通过最小化式 (18), 利用梯度下降法可得到 z_{qk}, t_{fk} 和 v_{kn} 的更新规则.

首先, 令

$$\frac{\partial F^+}{\partial z_{qk}} = \sum_{f,n}\left(-\frac{|y_{fn,q}|^2\alpha_{fn,k}^2}{z_{qk}^2 t_{fk}v_{kn}} + \frac{t_{fk}v_{kn}}{\beta_{fn}}\right) = 0 \tag{19}$$

即

$$\sum_{f,n}\frac{t_{fk}v_{kn}}{\beta_{fn}} = \frac{1}{z_{qk}^2}\sum_{f,n}\frac{|y_{fn,q}|^2\alpha_{fn,k}^2}{t_{fk}v_{kn}} \tag{20}$$

变形得

$$z_{qk} = \sqrt{\frac{\sum_{f,n}\frac{|y_{fn,q}|^2\alpha_{fn,k}^2}{t_{fk}v_{kn}}}{\sum_{f,n}\frac{t_{fk}v_{kn}}{\beta_{fn}}}} \tag{21}$$

把 $\alpha_{fn,k} = (z_{qk}t_{fk}v_{kn})/(\sum_{k'} z_{qk'}t_{fk'}v_{k'n})$, $\beta_{fn} = \sum_{k'} z_{qk'}t_{fk'}v_{k'n}$ 代入式 (21), 得 z_{qk} 的更新规则如下

$$z_{qk} \leftarrow z_{qk}\sqrt{\frac{\sum_{f,n}\frac{|y_{fn,q}|^2 t_{fk}v_{kn}}{(\sum_{k'} z_{qk'}t_{fk'}v_{k'n})^2}}{\sum_{f,n}\frac{t_{fk}v_{kn}}{\sum_{k'} z_{qk'}t_{fk'}v_{k'n}}}} \tag{22}$$

其中, 计算 $z_{qk} \leftarrow z_{qk}/\sum_{q'} z_{q'k}$, 使得 $\sum_q z_{qk} = 1$.

同理, 令

$$\frac{\partial F^+}{\partial t_{fk}} = \sum_{n,q}\left(-\frac{|y_{fn,q}|^2\alpha_{fn,k}^2}{z_{qk}t_{fk}^2 v_{kn}} + \frac{z_{qk}v_{kn}}{\beta_{fn}}\right) = 0 \tag{23}$$

$$\frac{\partial F^+}{\partial v_{kn}} = \sum_{f,q}\left(-\frac{|y_{fn,q}|^2\alpha_{fn,k}^2}{z_{qk}t_{fk}v_{kn}^2} + \frac{z_{qk}t_{fk}}{\beta_{fn}}\right) = 0 \tag{24}$$

即

$$\sum_{n,\,q}\left(\frac{z_{qk}v_{kn}}{\beta_{fn}}\right)=\frac{1}{t_{fk}^2}\sum_{n,\,q}\left(\frac{|y_{fn,\,q}|^2\alpha_{fn,\,k}^2}{z_{qk}v_{kn}}\right)\tag{25}$$

$$\sum_{f,\,q}\left(\frac{z_{qk}t_{fk}}{\beta_{fn}}\right)=\frac{1}{v_{kn}^2}\sum_{f,\,q}\left(\frac{|y_{fn,\,q}|^2\alpha_{fn,\,k}^2}{z_{qk}t_{fk}}\right)\tag{26}$$

变形得

$$t_{fk}=\sqrt{\frac{\sum\limits_{n,\,q}\left(\frac{|y_{fn,\,q}|^2\alpha_{fn,\,k}^2}{z_{qk}v_{kn}}\right)}{\sum\limits_{n,\,q}\left(\frac{z_{qk}v_{kn}}{\beta_{fn}}\right)}}\tag{27}$$

$$v_{kn}=\sqrt{\frac{\sum\limits_{f,\,q}\left(\frac{|y_{fn,\,q}|^2\alpha_{fn,\,k}^2}{z_{qk}t_{fk}}\right)}{\sum\limits_{f,\,q}\left(\frac{z_{qk}t_{fk}}{\beta_{fn}}\right)}}\tag{28}$$

把 $\alpha_{fn,\,k}=(z_{qk}t_{fk}v_{kn})/(\sum_{k'}z_{qk'}t_{fk'}v_{k'n})$, $\beta_{fn}=\sum_{k'}z_{qk'}t_{fk'}v_{k'n}$ 代入式 (27) 和式 (28), 得 t_{fk} 和 v_{kn} 的更新规则分别如下

$$t_{fk}\leftarrow t_{fk}\sqrt{\frac{\sum\limits_{n,\,q}\left(\frac{|y_{fn,\,q}|^2z_{qk}v_{kn}}{(\sum\limits_{k'}z_{qk'}t_{fk'}v_{k'n})^2}\right)}{\sum\limits_{n,\,q}\left(\frac{z_{qk}v_{kn}}{\sum\limits_{k'}z_{qk'}t_{fk'}v_{k'n}}\right)}}\tag{29}$$

$$v_{kn}\leftarrow v_{kn}\sqrt{\frac{\sum\limits_{f,\,q}\left(\frac{|y_{fn,\,q}|^2z_{qk}t_{fk}}{(\sum\limits_{k'}z_{qk'}t_{fk'}v_{k'n})^2}\right)}{\sum\limits_{f,\,q}\left(\frac{z_{qk}t_{fk}}{\sum\limits_{k'}z_{qk'}t_{fk'}v_{k'n}}\right)}}\tag{30}$$

然后, 为了得到W_f的更新规则, 把 $y_{fn,\,q}=w_{f,\,q}x_{fn,\,q}$ 代入式 (18) 中, 得

$$F^+=\sum_f\left(\sum_q w_{f,\,q}^{\rm T}V_{f,\,q}w_{f,\,q}-2\ln|\det W_f|\right)+\text{const}\tag{31}$$

其中,

$$V_{f,\,q}=\frac{1}{N}\sum_n\frac{\boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\rm T}\alpha_{fn,\,k}^2}{\sum\limits_k z_{qk}t_{fk}v_{kn}}\tag{32}$$

const 是与 W_f 相互独立的项. 对于式 (31), 令

$$\frac{\partial F^+}{\partial w_{f,\,q}^{\rm T}}=V_{f,\,q}w_{f,\,q}-2\frac{\partial}{\partial w_{f,\,q}^{\rm T}}\ln|\det W_f|=0\tag{33}$$

利用矩阵对数的微分矩阵公式得

$$\frac{\partial}{\partial W_f}\ln|\det W_f|=W_f^{-1}\tag{34}$$

因此, 式 (33) 可以用向量形式表示为

$$w_{f,\,l}^{\rm T}V_{f,\,q}w_{f,\,q}=\delta_{lq}\tag{35}$$

其中,

$$\delta_{lq}=\begin{cases}1, & l=q\\ 0, & l\neq q\end{cases}\tag{36}$$

即

$$w_{f,\,q}^{\rm T}V_{f,\,q}w_{f,\,q}=1\tag{37}$$

$$w_{f,\,l}^{\rm T}V_{f,\,q}w_{f,\,q}=0\ \ (l\neq q)\tag{38}$$

等价于

$$\begin{bmatrix}w_{f,\,1}^{\rm T}\\ \vdots\\ w_{f,\,q-1}^{\rm T}\\ w_{f,\,q}^{\rm T}\\ w_{f,\,q+1}^{\rm T}\\ \vdots\\ w_{f,\,Q}^{\rm T}\end{bmatrix}V_{f,\,q}w_{f,\,q}=\begin{bmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix}\tag{39}$$

可写作

$$W_fV_{f,\,q}w_{f,\,q}=\mathbf{1}_q\tag{40}$$

其中, $\mathbf{1}_q$ 为第 q 个元素是 1 的单位列向量. 可得 $w_{f,\,q}$ 为

$$w_{f,\,q}=(W_fV_{f,\,q})^{-1}\mathbf{1}_q\tag{41}$$

为了满足式 (37), 可得 $w_{f,\,q}$ 的更新规则如下

$$w_{f,\,q}\leftarrow\frac{w_{f,\,q}}{\sqrt{w_{f,\,q}^{\rm T}V_{f,\,q}w_{f,\,q}}}\tag{42}$$

另外, 为了避免估计的源信号出现幅度歧义性问题, 可以在每次迭代过程中标准化 $w_{f,\,q}$, 即

$$w_{f,\,q}\leftarrow\frac{w_{f,\,q}}{\sqrt{\frac{1}{FN}\sum\limits_{f,\,n}|y_{fn,\,q}|}}\tag{43}$$

$$y_{fn,\,q}\leftarrow\frac{y_{fn,\,q}}{\sqrt{\frac{1}{FN}\sum\limits_{f,\,n}|y_{fn,\,q}|}}\tag{44}$$

同时, 用单位矩阵作为W_f的初始化. 从而得到频域上分离的源信号为

$$\hat{\boldsymbol{s}}_{fn}\approx\boldsymbol{y}_{fn}=W_f\boldsymbol{x}_{fn}\tag{45}$$

最后, 利用短时傅里叶变换的逆变换得到时域上的源信号 $\boldsymbol{s}(t)$. 算法 1 概述了本文所提的卷积混叠信号盲源分离算法流程.

算法 1. 卷积混叠信号盲源分离算法

输入. 卷积混叠信号 $\boldsymbol{x}(t)$.

输出. 分离的源信号 $\hat{\boldsymbol{s}}(t)$.

1) 对混叠信号 $\boldsymbol{x}(t)$ 进行短时傅里叶变换, 得到频域上的混叠信号 $\boldsymbol{x}_{fn}$;

2) 定义代价函数 (11) 和辅助函数 (18);

3) 随机初始化模型参数 z_{qk}, t_{fk}, v_{kn}, $W_f = \mathbf{1}$, 利用以下更新准则更新模型参数:

a) 利用式 (22) 更新 z_{qk};

b) 利用式 (29) 更新 t_{fk};

c) 利用式 (30) 更新 v_{kn};

d) 利用式 (42) 更新 $w_{f,q}$;

e) 利用式 (44) 更新 $y_{fn,q}$;

f) 重复步骤 a) ~ e), 直至收敛为止.

4) 利用式 (45) 分离得到频域上的源信号 $\hat{\boldsymbol{s}}_{fn}$;

5) 利用逆傅里叶变换, 把频域上的源信号 $\hat{\boldsymbol{s}}_{fn}$ 转换到时域上, 得到 $\hat{\boldsymbol{s}}(t)$.

3　仿真实验与结果分析

为了验证所提算法的有效性, 分别测试中文语音混叠信号、英文语音混叠信号、音乐混叠信号. 为了很好地模拟现实的混响环境, 创建一个可人工控制的空间脉冲响应的房间, 混响时间 RT_{60} 定义为反射声线强度比直达声线强度低 60 分贝所需的时间[46]

$$RT_{60} = \frac{24\ln(10)V}{c\sum_{i=1}^{6} S_i(1-\beta_i^2)} \tag{46}$$

其中, V 表示模拟房间的体积, c 是声音的速度 (340 m/s), β_i 和 S_i 分别表示第 i 面墙反射系数和表面积. 通过任意改变 RT_{60} 值可以控制不同的混响环境来更好地模拟现实环境, 其值越大, 说明环境越复杂, 源信号的分离就越困难.

同时, 为了评价各类算法的盲源分离性能, 选择 SDR (Source-to-distortion ratio) 和 SIR (Source-to-interference ratio) 作为源信号分离性能评价准则[47], 它们是一种已经公开建立的源信号分离质量比较评价技术, 可以很好地衡量每个算法分离源信号的能力. 其值越大, 说明分离性能越好.

另外, 为了说明所提算法的优越性, 对比几种目前比较流行的盲源分离算法: 满秩算法 (Full-Rank)[40], 体积最小化交替优化算法 (Volume minimization alternating optimization, VolMin-AO)[48], 低秩非负矩阵分解算法 (Rank1-NMF)[36], 以及校正块 Tucker-2 分解算法 (Rectified block Tucker-2 decomposition, RBTD)[23]. 其中, Full-Rank 算法是通过构建空间协方差模型, 设计模型参数估计算法. VolMin-AO 算法是在协方差域利用凸几何数学理论设计一种体积最小化交替优化算法. Rank1-NMF 算法是结合独立向量分析和非负矩阵分解理论设计的一种正定混叠信号盲源分离算法. RBTD 算法是通过构建一个张量框架消除排序歧义性问题, 进而设计一种卷积混叠信号盲源分离算法. 这几种流行的盲源分离算法是从不同的角度考虑问题, 在音频信号处理中发挥各自的优势. 通过与这些算法进行对比, 可以很好地反映所提算法的先进性和优越性.

3.1　实验一. 中文语音混叠信号盲源分离

首先, 测试中文语音信号混叠情况下的盲源分离, 数据集来自国内公共数据集[49], 选择其中两组中文语音信号进行卷积混叠, 具体信息如表 1 所示, 混响时间设置为 100 ms ~ 350 ms, 实验结果如图 1、图 2 所示, 这与我们初期预想是相符合的, 随着混响时间 RT_{60} 的增长, SDR 和 SIR 的值呈现逐渐减小的趋势. 显然, 混响时间变长给分离性能带来了很大的影响, 特别是在较高混响下, 分离性能急剧下降. 但是, 相比于其他算法, 所提算法还是具有更好的分离效果.

表 1　两组中文语音源信号

Table 1　Two groups of Chinese speech sources

中文数据	源信号	时长
语音 1	IC0936W0131	5 s
语音 2	IC0936W0134	5 s

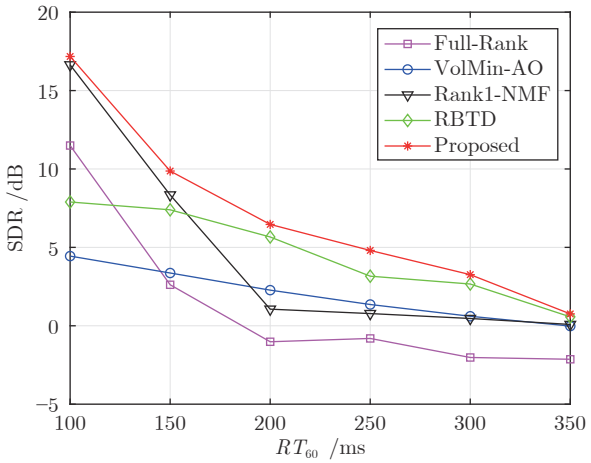

图 1　中文语音混叠信号盲源分离 SDR 性能对比

Fig. 1　SDR performance comparison for BSS of Chinese speech mixtures

3.2　实验二. 英文语音混叠信号盲源分离

其次, 测试英文语音信号混叠情况下的盲源分

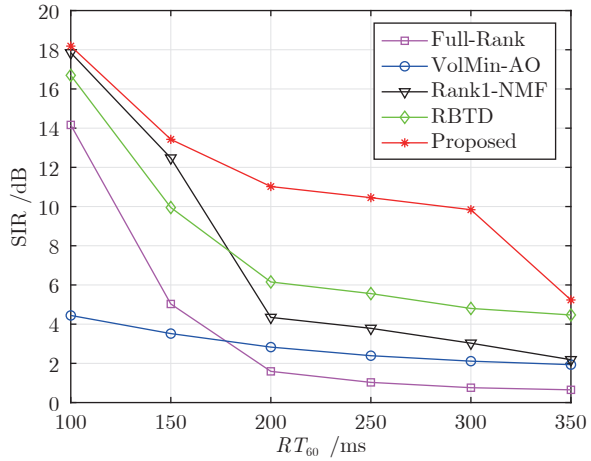

图 2 中文语音混叠信号盲源分离 SIR 性能对比

Fig. 2 SIR performance comparison for BSS of Chinese speech mixtures

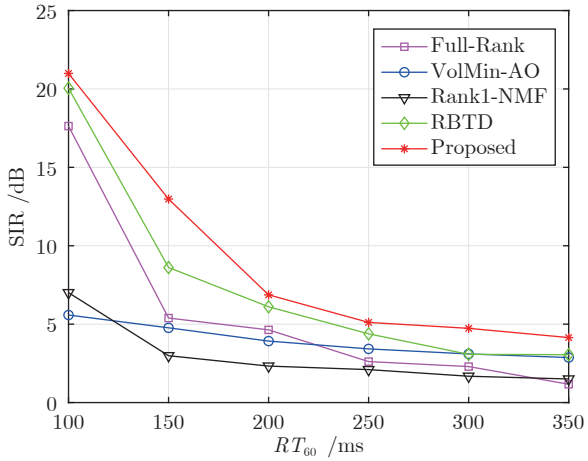

图 4 英文语音混叠信号盲源分离 SIR 性能对比

Fig. 4 SIR performance comparison for BSS of English speech mixtures

离性能, 数据集来自于国外公共数据集[50], 选择其中两组英文语音信号进行卷积混叠, 如表 2 所示, 混响时间设置为 100 ms ~ 350 ms, 实验结果如图 3、图 4 所示, 同样, 随着混响时间的增大, 盲分离性能降低, 当混响时间达到 200 ms 时, 算法的分离性能已经明显地下降, 对比于其他算法, 本文所提算法的盲源分离性能一直具有一定的优越性.

然而, 在高混响环境下, 算法的分离性能明显下降, 说明在分离英文语音卷积混叠信号时, 在低混响下效果显著, 对高混响环境比较敏感. 主要是因为混响时间的增长带来强烈的回声, 因此, 算法对高混响的鲁棒性有待进一步改进.

表 2 两组英文语音源信号

Table 2 Two groups of English speech sources

英文数据	源信号	时长
语音 1	dev1_female3_src_1	10 s
语音 2	dev1_female3_src_2	10 s

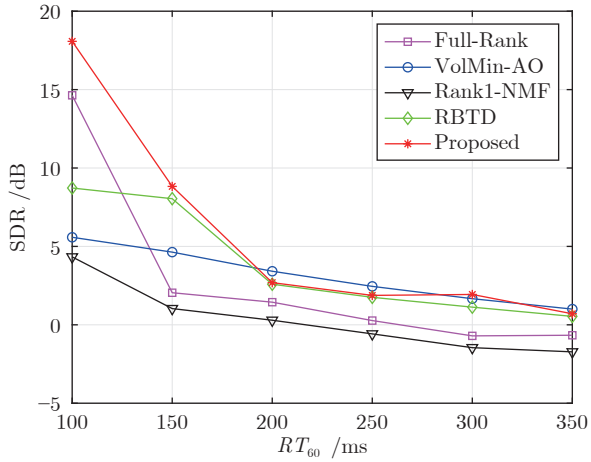

图 3 英文语音混叠信号盲源分离 SDR 性能对比

Fig. 3 SDR performance comparison for BSS of English speech mixtures

3.3 实验三. 音乐混叠信号盲源分离

为了验证所提算法在音乐混叠信号中的应用, 测试音乐信号混叠情况下的盲源分离, 数据集同样来自于国外公共数据集[50], 选择其中两组音乐信号进行卷积混叠, 如表 3 所示, 混响时间设置为 100 ms ~ 350 ms, 实验结果如图 5、图 6 所示, 可见所提算法的分离性能明显优于其他算法. 同时, 对比于实验二中的英文语音混叠信号盲源分离情况, 该算法对音乐混叠信号的盲源分离性能相对较好, 特别是在混响时间超过 200 ms 时, 盲源分离性能依然较好, 说明此算法在分离音乐混叠信号时比分离英文语音混叠信号具有更好的效果.

表 3 两组音乐源信号

Table 3 Two groups of music sources

音乐数据	源信号	时长
音乐 1	dev1_wdrums_src_1	11 s
音乐 2	dev1_wdrums_src_3	11 s

3.4 实验四. 噪声对算法盲分离性能的影响

以上实验都是在无噪声的环境下做的测试, 为了测试高斯白噪声对算法盲分离性能的影响, 对实验一中的中文语音卷积混叠信号添加一组高斯白噪声, 选择信噪比 (Source-to-noise ratio, SNR) 在 10 dB ~ 35 dB 之间进行测试. 因为算法在低混响

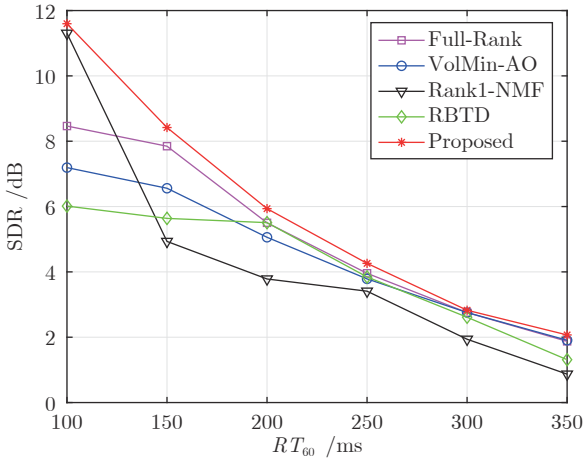

图 5　音乐混叠信号盲源分离 SDR 性能对比

Fig. 5　SDR performance comparison for BSS of music mixtures

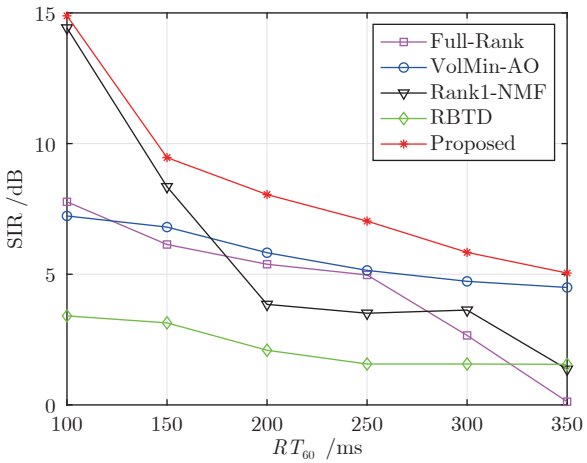

图 6　音乐混叠信号盲源分离 SIR 性能对比

Fig. 6　SIR performance comparison for BSS of music mixtures

下效果最好, 所以选择混响时间为 $RT_{60} = 100$ ms, 实验结果如图 7、图 8 所示, 对于有噪声的语音混叠信号, 所提算法的盲分离性能仍然明显超过其他的对比算法, 说明所提算法对于有噪声的语音混叠信号的盲分离有较好的鲁棒性.

同时, 为了验证噪声对音乐信号盲分离性能的影响, 对实验三中的音乐卷积混叠信号添加一组高斯白噪声, SNR 设置为 10 dB ~ 35 dB, 混响时间为 $RT_{60} = 100$ ms, 实验结果如图 9、图 10 所示, 对于有噪声的音乐混叠信号的盲分离, 所提算法仍然具有更好的分离性能, 说明所提算法对有噪声的音乐混叠信号盲分离也有较好的鲁棒性.

3.5　实验结果分析

通过以上实验结果发现, 所提算法在分离卷积

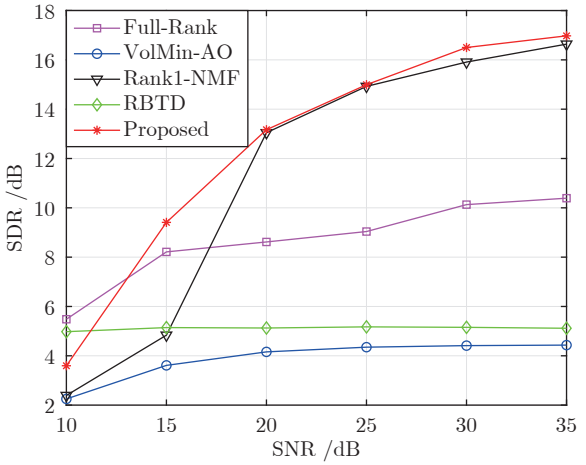

图 7　噪声对语音信号盲分离 SDR 性能的影响

Fig. 7　Effect of noise on SDR performance for BSS of Chinese speech mixtures

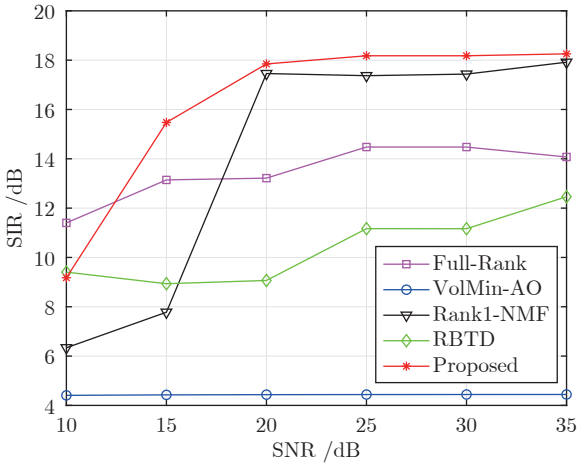

图 8　噪声对语音信号盲分离 SIR 性能的影响

Fig. 8　Effect of noise on SIR performance for BSS of Chinese speech mixtures

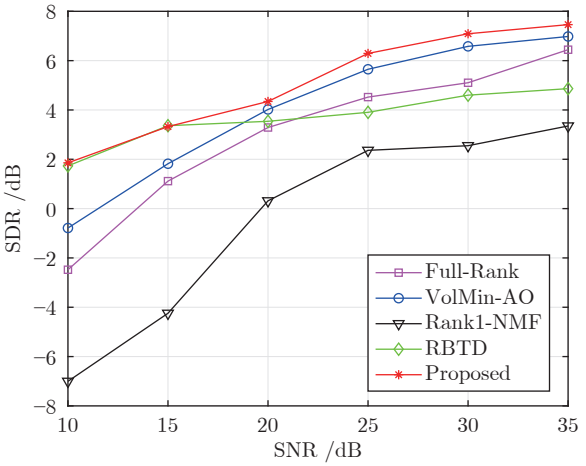

图 9　噪声对音乐信号盲分离 SDR 性能的影响

Fig. 9　Effect of noise on SDR performance for BSS of music mixtures

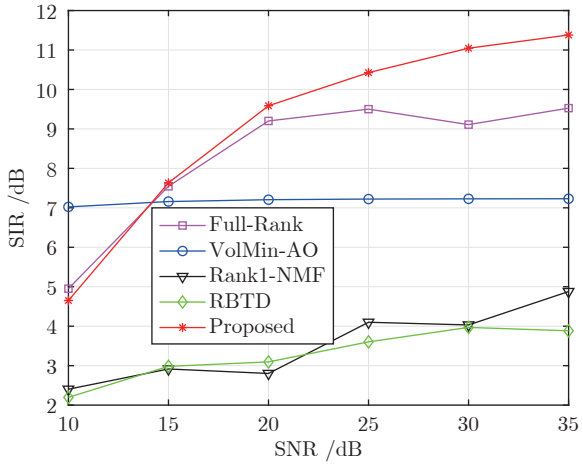

图 10 噪声对音乐信号盲分离 SIR 性能的影响

Fig. 10 Effect of noise on SIR performance for BSS of music mixtures

混叠语音信号时具有很好的效果, 在处理音乐信号时具有更好的分离性能, 主要是因为非负矩阵分解模型更适合于音乐混叠信号. 特别是在低混响环境下, 算法的盲分离性能明显优于其他对比算法; 同时在低噪声环境下, 所提算法具有较好的鲁棒性. 所提算法的计算复杂度主要集中于矩阵乘积, 本文假设源信号在每个时频点上满足独立复高斯分布, 在独立分量分析框架下, 源信号的协方差矩阵是 Hermitian 正定的对角化矩阵, 可表示为式 (7), 计算过程中非对角元素全为零, 计算复杂度主要来源于主对角元素的乘积, 由源信号数目决定, 这样减少了矩阵乘积带来的计算复杂度. 另外, 设计的优化目标函数 (11) 是基于 Itakura-Saito 散度, 通过最小化 X_{fn} 与 $\hat{X}_{fn}$ 的 Itakura-Saito 散度, 利用源信号在每个时频点上满足独立复高斯分布, 对相互独立的源进行建模. 模型参数更新方式的推导过程是利用梯度下降法进行的, 这样可以提高计算速度.

然而, 在高混响下, 盲源分离性能受到很大影响; 在高噪声环境下, 源信号淹没于噪声中, 分离将更加困难. 本文考虑音乐混叠信号的盲源分离, 选择混响时间为 $RT_{60} = 400$ ms 的高混响环境, 以及信噪比 SNR 为 5 dB 的高噪声环境, 其分离结果对比如表 4 所示, 可见在高混响、高噪声环境中, 所提算法的优势不太明显. 所以, 所提算法的优势是分离低混响、低噪声的特定环境下的卷积混叠信号, 盲源分离性能比目前比较流行的算法具有更好的效果. 针对高混响、高噪声的卷积混叠信号盲源分离问题, 由于混叠环境的复杂性和不确定性, 导致所提算法的分离性能下降, 这是该算法的局限性, 需要进一步优化.

另外, 本文提出的卷积混叠盲源分离算法不仅适用于线性混叠模型, 还适用于卷积混叠模型. 从理论上说, 本文假设混叠矩阵是可逆的, 在超定情况下仍然满足, 只需要利用主成分分析技术降维即可, 所以本文所提算法也适用于超定情况. 在实际应用中, 当传感器的数目大于或等于源信号的数目时, 该算法仍有一定的实用性. 例如, 在语音信号采集过程中使用的是全方位麦克风传感器进行收集. 当然, 当传感器的数目小于源信号的数目时, 即在欠定混叠下, 该算法受到限制. 针对欠定卷积混叠下的研究也是当下盲分离研究的难点, 作者也在积极开展这方面的研究.

表 4 高混响、高噪声环境中的实验结果

Table 4 Experimental results in high reverberation and high noise environment

	$RT_{60} = 400$ ms		SNR = 5 dB	
	SDR	SIR	SDR	SIR
Full-Rank	0.1969	4.5580	−4.2087	6.7379
VolMin-AO	1.1786	4.3729	−3.8684	6.6486
Rank1-NMF	−1.8239	0.7933	−9.8632	2.7641
RBTD	−6.7646	1.2411	−9.1111	1.8784
Proposed	1.0278	5.7190	−1.8554	4.6515

4 结论

本文主要针对卷积混叠音频信号盲源分离问题, 提出一种高效的盲源分离新算法. 通过实时更新解混叠矩阵和非负矩阵分解模型参数, 避免时频域上出现的排序歧义性问题, 提高了盲源分离性能. 大量的音频信号实验结果证明了该算法具有良好的分离性能, 优于目前流行的盲分离算法的分离性能. 特别是在低混响和低噪声环境下, 该算法具有明显的盲分离优势.

另外, 本文考虑的是时不变卷积混叠系统, 要求源信号的位置是固定不变的, 如会议室中的语音信号、录音棚中的音乐信号等应用场景. 而在实际复杂环境中, 源信号的位置常常是移动的, 同时伴随高混响、高噪声等不确定因素, 针对这种时变的卷积混叠系统, 在未来的科研工作中, 需要寻找更好的解决方案, 用新眼光新思路看问题, 设计鲁棒性、自适应性更好的盲源分离方法.

References

1 Zhang Xian-Da, Bao Zheng. Blind signal separation. *Acta Electronica Sinica*, 2001, **29**(z1): 1766−1771
(张贤达, 保铮. 盲信号分离. 电子学报, 2001, **29**(z1): 1766−1771)

2 Yilmaz O, Rickard S. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 2004, **52**(7): 1830−1847

3 Mcdermott J H. The cocktail party problem. *Neural Computation*, 2005, **17**(9): 1875−1902

4 Ozerov A, Fevotte C. Multichannel nonnegative matrix factoriz-

ation in convolutive mixtures for audio source separation. *IEEE Transactions on Audio Speech and Language Processing*, 2010, **18**(3): 550−563

5 Ito N, Ikeshita R, Sawada H, Nakatani T. A joint diagonalization based efficient approach to underdetermined blind audio source separation using the multichannel Wiener filter. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2021, **29**: 1950−1965

6 Shah G, Koch P, Papadias C B. On the blind recovery of cardiac and respiratory sounds. *IEEE Journal of Biomedical and Health Informatics*, 2015, **19**(1): 151−157

7 Xie Y, Xie K, Yang Q Y, Xie S L. Reverberant blind separation of heart and lung sounds using nonnegative matrix factorization and auxiliary function technique. *Biomedical Signal Processing and Control*, 2021, **69**(7): Article No. 102899

8 Zhang S Q, You B, Lang X, Zhou Y F, An F, Dai Y, et al. Efficient rejection of artifacts for short-term few-channel EEG based on fast adaptive multidimensional sub-bands blind source separation. *IEEE Transactions on Instrumentation and Measurement*, 2021, **70**: Article No. 4007516

9 Miettinen J, Nitzan E, Vorobyov S A, Ollila E. Graph signal processing meets blind source separation. *IEEE Transactions on Signal Processing*, 2020, **69**: 2585−2599

10 Einizade A, Sardouie S H, Shamsollahi M B. Simultaneous graph learning and blind separation of graph signal sources. *IEEE Signal Processing Letters*, 2021, **28**: 1495−1499

11 Yang Y C, Nagarajaiah S. Structural damage identification via a combination of blind feature extraction and sparse representation classification. *Mechanical Systems and Signal Processing*, 2014, **45**(1): 1−23

12 Yang Y C, Li S L, Nagarajaiah S, Li H, Zhou P. Real-time output-only identification of time-varying cable tension from accelerations via complexity pursuit. *Journal of Structural Engineering*, 2016, **142**(1): Article No. 04015083

13 Xie Sheng-Li, He Zhao-Shui, Fu Yu-Li. Underdetermined aliasing adaptive blind separation method based on sparse element analysis. *Chinese Science (Series E: Information Science)*, 2007, **37**(8): 1086−1098
(谢胜利, 何昭水, 傅予力. 基于稀疏元分析的欠定混叠自适应盲分离方法. 中国科学 (E 辑: 信息科学), 2007, **37**(8): 1086−1098)

14 Lathauwer L D, Castaing J. Blind identification of underdetermined mixtures by simultaneous matrix diagonalization. *IEEE Transactions on Signal Processing*, 2008, **56**(3): 1096−1105

15 Tang Hui, Wang Shu. Robust joint block diagonalization based convolutive blind source separation. *Acta Automatica Sinica*, 2013, **39**(9): 1502−1510
(汤辉, 王殊. 基于稳健联合分块对角化的卷积盲分离. 自动化学报, 2013, **39**(9): 1502−1510)

16 Zhu Xiao-Long, Zhang Xian-Da. Overdetermined blind signal separation based on singular value decomposition. *Journal of Electronics & Information Technology*, 2004, **26**(3): 337−343
(朱孝龙, 张贤达. 基于奇异值分解的超定盲信号分离. 电子与信息学报, 2004, **26**(3): 337−343)

17 Yatabe K, Kitamura D. Determined BSS based on time-frequency masking and its application to harmonic vector analysis. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2021, **29**: 1609−1625

18 Xiao Ming, Xie Sheng-Li, Fu Yu-Li. Underdetermined blind signal separation algorithm based on hyperplane normal vector. *Acta Automatica Sinica*, 2008, **34**(2): 142−149
(肖明, 谢胜利, 傅予力. 基于超平面法矢量的欠定盲信号分离算法. 自动化学报, 2008, **34**(2): 142−149)

19 Nion D, Mokios K N, Sidiropoulos N D, Potamianos A. Batch and adaptive PARAFAC-based blind separation of convolutive speech mixtures. *IEEE Transactions on Audio, Speech, and Language Processing*, 2010, **18**(6): 1193−1207

20 Matsuoka K. Minimal distortion principle for blind source separation. In: Procedings of the 3rd International Conference on Independent Component Analysis and Blind Signal Separation. Tobata, Japan: IEEE, 2001. 722−729

21 Sawada H, Mukai R, Araki S. A robust and precise method for solving the permutation problem of frequency-domain blind source separation. *IEEE Transactions Speech Audio Processing*, 2004, **12**(5): 530−538

22 Sawada H, Araki S, Mukai R. Underdetermined convolutive blind source separation via frequency bin-wise clustering and permutation alignment. *IEEE Transactions on Audio Speech and Language Processing*, 2011, **19**(3): 516−527

23 Xie K, Zhou G X, Yang J J, He Z S, Xie S L. Eliminating the permutation ambiguity of convolutive blind source separation by using coupled frequency bins. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, **31**(2): 589−599

24 Sawada H, Araki S, Mukai R, Makina S. Grouping separated frequency components by estimating propagation model parameters in frequency-domain blind source separation. *IEEE Transactions on Audio Speech and Language Processing*, 2007, **15**(5): 1592−1604

25 Xie S L, Yang L, Yang J M, Zhou G X, Xiang Y. Time-frequency approach to underdetermined blind source separation. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, **23**(2): 306−316

26 Liu Qiu-Hong, Xu Man-Kun, Li Tian-Jun, Lu Ming-Ming. Blind separation algorithm of APCMA signal based on complementary symmetric filter. *Acta Electronica Sinica*, 2020, **48**(12): 2394−2401
(刘秋红, 许漫坤, 李天昀, 陆明明. 基于互补对称滤波器的 APCMA 信号的盲分离算法. 电子学报, 2020, **48**(12): 2394−2401)

27 He Z S, Xie S L, Ding S X, Cichocki A. Convolutive blind source separation in the frequency domain based on sparse representation. *IEEE Transactions on Audio Speech and Language Processing*, 2007, **15**(5): 1551−1563

28 Xie Y, Xie K, Xie S L. Underdetermined blind source separation of speech mixtures unifying dictionary learning and sparse representation. *International Journal of Machine Learning and Cybernetics*, 2021, **12**(12): 3573−3583

29 Xu Z B, Zhang H, Wang Y, Chang X Y, Liang Y. $L_{1/2}$ regularization. *Science China (Information Sciences)*, 2010, **53**(6): 1159−1169

30 Xu Z B, Chang X Y, Xu F M, Zhang H. $L_{1/2}$ Regularization: A thresholding representation theory and a fast solver. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, **23**(7): 1013−1027

31 Yang J J, Guo Y, Yang Z Y, Xie S L. Underdetermined convolutive blind source separation combining density-based clustering and sparse reconstruction in time-frequency domain. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2019, **66**(8): 3015−3027

32 Xie Y, Xie K, Xie S L. Underdetermined blind separation of source using L_p-norm diversity measures. *Neurocomputing*, 2020, **411**: 259−267

33 Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization. *Nature*, 1999, **401**(6755): 788−791

34 Gillis N, Vavasis S A. Fast and robust recursive algorithms for separable nonnegative matrix factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(4): 698−714

35 Rahiche A, Cheriet M. Blind decomposition of multispectral document images using orthogonal nonnegative matrix factorization. *IEEE Transactions on Image Processing*, 2021, **30**: 5997−6012

36 Kitamura D, Ono N, Sawada H, Kameoka H, Saruwatari H. Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 2016, **24**(9): 1626−1641

37 Al-Tmeme A, Woo W L, Dlay S S, Gao B. Underdetermined convolutive source separation using GEM-MU with variational approximated optimum model order NMF2D. *IEEE/ACM*

Transactions on Audio, Speech, and Language Processing, 2017, **25**(1): 31−45

38 Xie Y, Xie K, Xie S L. Underdetermined convolutive blind separation of sources integrating tensor factorization and expectation maximization. *Digital Signal Processing*, 2019, **87**: 145−154

39 Sekiguchi K, Bando Y, Nugraha A A, Yoshii K, Kawahara T. Fast multichannel nonnegative matrix factorization with directivity-aware jointly-diagonalizable spatial covariance matrices for blind source separation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2020, **28**: 2610−2625

40 Duong N, Vincent E, Gribonval R. Under-determined reverberant audio source separation using a full-rank spatial covariance model. *IEEE Transactions on Audio, Speech, and Language Processing*, 2010, **18**(7): 1830−1840

41 Bando Y, Sekiguchi K, Masuyama Y, Nugraha A A, Fontaine M, Yoshii K. Neural full-rank spatial covariance analysis for blind source separation. *IEEE Signal Processing Letters*, 2021, **28**: 1670−1674

42 Kolda T. Tensor decompositions and applications. *Siam Review*, 2009, **51**(3): 455−500

43 Weiss A. Blind direction-of-arrival estimation in acoustic vector-sensor arrays via tensor decomposition and Kullback-Leibler divergence covariance fitting. *IEEE Transactions on Signal Processing*, 2021, **69**: 531−545

44 Mitsufuji Y, Takamune N, Koyama S, Saruwatari H. Multichannel blind source separation based on evanescent-region-aware non-negative tensor factorization in spherical harmonic domain. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2021, **29**: 607−617

45 Tan V Y F, Févotte C. Automatic relevance determination in nonnegative matrix factorization with the-divergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(7): 1592−1605

46 Audio Labs. RIR generator [Online], available: https://www.audiolabs-erlangen.de/fau/professor/habets/software/rir-generator, November 22, 2022

47 Vincent E, Gribonval R, Fevotte C. Performance measurement in blind audio source separation. *IEEE Transactions on Audio Speech and Language Processing*, 2006, **14**(4): 1462−1469

48 Fu X, Ma W K, Huang K, Sidiropoulos N. Blind separation of quasi-stationary sources: Exploiting convex geometry in covariance domain. *IEEE Transactions on Signal Processing*, 2015, **63**(9): 2306−2320

49 AISHELL-ASR0009-OS1 open source mandarin speech corpus [Online], available: http://www.aishelltech.com/kysjcp, November 22, 2022

50 SiSEC 2013. Audio source separation [Online], available: http://sisec.wiki.irisa.fr/tiki-index.php?page=Professionally+produced+music+recordings, November 22, 2022

解 元 广州大学机械与电气工程学院讲师. 主要研究方向为盲信号分离, 信号处理和机器学习.
E-mail: yuanxiemath@hotmail.com
(**XIE Yuan** Lecturer at the School of Mechanical and Electrical Engineering, Guangzhou University. His research interest covers blind signal separation, signal processing, and machine learning.)

邹 涛 广州大学机械与电气工程学院教授. 主要研究方向为工业过程建模与仿真, 模型预测控制, 先进过程控制和实时优化技术研究与应用. 本文通信作者.
E-mail: tzou@gzhu.edu.cn
(**ZOU Tao** Professor at the School of Mechanical and Electrical Engineering, Guangzhou University. His research interest covers industrial process modeling and simulation, model predictive control, advanced process control, and real-time optimization technology research and application. Corresponding author of this paper.)

孙为军 广东省物联网信息技术重点实验室、智能检测与制造物联教育部重点实验室副教授. 主要研究方向为模式识别, 机器学习.
E-mail: gdutswj@163.com
(**SUN Wei-Jun** Associate professor of Guangdong Provincial Key Laboratory of Information Technology of Internet of Things, and Key Laboratory of Intelligent Detection and the Internet of Things in Manufacturing, Ministry of Education. His research interest covers pattern recognition and machine learning.)

谢胜利 基于物联网技术的离散制造智能化学科创新引智基地、粤港澳离散制造智能化联合实验室教授. 主要研究方向为无线网络, 自动控制和盲信号处理.
E-mail: shlxie@gdut.edu.cn
(**XIE Sheng-Li** Professor of Discrete Manufacturing Intelligence Discipline Innovation and Talent Introduction Base Based on Internet of Things Technology, and Guangdong-Hong Kong-Macao Joint Laboratory for Smart Discrete Manufacturing. His research interest covers wireless networks, automatic control, and blind signal processing.)