

# 频域盲解卷积局限性分析及一种改进算法

张超, 吴小培

(安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039)

**摘要:**在频域盲解卷积问题中, 时域信号的卷积混合转化为频域信号在有限频点的瞬时混合, 使算法复杂度大大降低。但这种算法的局限是分离结果存在次序和幅度上的不确定性, 并且窗函数长度和信号非平稳性之间存在相互制约的关系。文中对语音信号频域盲解卷积算法存在的制约因素进行分析并提出一种改进的基于包络相关性的排序方法。在分裂谱法的基础上, 通过“分裂”后的多路信号求得“总包络”, 再依据“总包络”进行排序, 从而克服传统的直接依据输出信号包络相关性进行排序的不足。实验结果表明, 采用本方法可获得较高的分离质量。

**关键词:**频域; 盲解卷积; 局限性; 排序

**中图分类号:**TP301.6

**文献标识码:**A

**文章编号:**1673-629X(2008)10-0057-04

## Limitations of Frequency - Domain Blind Deconvolution and an Improved Algorithm

ZHANG Chao, WU Xiao-pei

(Ministry of Education Key Lab. of Intelligent Computing and Signal  
Processing, Anhui University, Hefei 230039, China)

**Abstract:** In frequency - domain blind deconvolution, the convolutive mixture of time - domain signals is converted to instantaneous mixtures on several frequent bins, which makes the algorithms much less computational. But this kind of approach has its basic limitations that the results of separation are indeterminate in their scaling and permutation, further more, the length of window function is added to the observation signals and the non - stationary characteristic of speech signal are restricted with each other. Analyzes the limitations of FD - blind deconvolution algorithms and proposes an improved permutation algorithm based on the correlation of envelopes on each frequent bin. On the basis of “split spectrogram” algorithm, get the “general envelop” from the splited signals and remove the permutation ambiguity by measure the “general envelop”. This overcomes the deficiency of traditional algorithms which measure the correlation of envelopes of output signals directly. The experiment results show that the algorithm proposed in this paper can get a relatively high separation quality.

**Key words:** frequency - domain; blind deconvolution; permutation; limitation

## 0 引言

盲信号分离技术是当代信号处理研究领域中的一个热点和难点问题, 其实质是在信源和信道环境等因素未知的情况下, 仅依靠观测信号估计出源信号。独立分量分析(ICA)是盲信号处理领域中较有代表性的一种方法, 近年来已取得较大的发展并已在通信信号处理、语音信号处理、生物医学信号处理等领域获得相当规模的应用<sup>[1]</sup>。

在具体的 ICA 问题中, 如果观测信号是信号源在理想情况下瞬时混合得到, 可直接利用已较为成熟的瞬时混合 ICA 算法进行分离。然而在实际情况中, 由于现场声学环境中反射和延时等效应的存在, 实际记录的混合语音信号相当于源信号进行了空间滤波, 这使得问题的解决不再仅仅是寻求一个标量分离矩阵, 而是寻求对观测信号的自适应逆滤波。

为了对真实情况中信源的卷积混合进行分离, 一种方法是直接在时域中构造反卷积滤波器, 其分离准则与瞬时混合情况基本相同。另一种方法是通过短时傅里叶变换, 将信号变换到时频域, 由于 ICA 算法存在固有的次序和尺度的不确定性, 因而还需解除尺度不确定性以及平衡加窗长度和信号非平稳特性之间的制约关系才能取得好的分离效果<sup>[2]</sup>。

文中在阐述频域盲解卷积算法及其局限性的基础

收稿日期: 2008-01-27

基金项目: 国家自然科学基金资助项目(60771033); 安徽省自然科学基金资助项目(070412038)

作者简介: 张超(1983-), 男, 安徽蚌埠人, 硕士研究生, 研究方向为盲信号处理; 吴小培, 教授, 博士生导师, 研究方向为智能信息处理。

上,对文献[3]的排序算法进行改进,并结合实际情况,探讨了窗函数长度和混合语音信号短时平稳特性之间的制约关系,在此基础上进行分离实验,取得了较好的分离效果。

## 1 语音信号频域盲解卷积

设  $x = [x_1, x_2, \dots, x_N]^T$  和  $s = [s_1, s_2, \dots, s_M]^T$  分别是观测信号和源信号,  $M$  是信源个数,  $N$  是传感器个数,为了讨论简便起见令  $M = N$  且不考虑噪声。瞬时混合情况下

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

其中  $A$  是一  $N \times N$  标量混合矩阵,分离算法的目的是寻求分离矩阵  $W$ ,使得

$$y(t) = Wx(t) \quad (2)$$

为对信源  $s$  的估计。

然而在现实情况下,由于真实语音环境中的反射和时延等效应的存在,观测信号  $x$  实际上为信源  $s$  的卷积混合,即:

$$x_i(t) = \sum_{j=1}^N \sum_{k=0}^{L-1} a_{ij}(k) s_j(t-k) \quad i=1,2,\dots,N \quad (3)$$

其中,  $L$  为混合滤波器阶数,  $a_{jk}$  为第  $k$  个源到第  $j$  个传感器的冲击响应。

通过对(3)式进行  $L$  点短时傅里叶变换,可以得到:

$$X_j(f_l, t) = \sum_{\tau=0}^{L-1} x_j(\tau+1) \text{win}(\tau) \exp(-j2\pi f_l \tau) \quad j=1,2,\dots,N \quad (4)$$

其中,  $l=1,\dots,L-1$ ,  $f_l = (l/L)f_s$  为  $L$  个频点,  $f_s$  为采样率,  $\text{win}(\tau)$  为窗函数,  $t$  为每次滑窗开始的时间位置。经过短时傅里叶变换,得到一  $L \times M$  点观测信号矩阵。图 1 所示为两观测信号经 256 点短时傅里叶变换后的时频分布图。在各个频点运用复数域情况下瞬时混合 ICA 算法,得到  $W(f)$ ,使得:

$$Y(f, t) = W(f)X(f, t) \quad (5)$$

其中,  $X(f, t) = [X_1(f, t), X_2(f, t), \dots, X_N(f, t)]^T$ ,  $Y(f, t) = [Y_1(f, t), Y_2(f, t), \dots, Y_N(f, t)]^T$  为分离结果,  $W(f)$  为各频点的分离矩阵。

文中使用的频点分离算法为基于梯度下降的最大化峭度方法<sup>[4]</sup>,其每个频点的分离矩阵优化迭代公式为:

$$\Delta W(f) = \eta \text{sign}[k(y)] \frac{\partial k(y)}{\partial W^*(f)} \quad (6)$$

$$W(f) \leftarrow W(f) / \|W(f)\| \quad (7)$$

其中,  $y = y(f, t)$  且

$$k(y) = E\{(yy^*)^2\} - 2(E\{yy^*\})^2 - E\{yy\}E\{y^*y^*\} \quad (8)$$

$$\frac{\partial k(y)}{\partial W^*(f)} = 2(E\{yy^*y^*x\} - 2E\{yy^*\}E\{y^*x\} - E\{y^*y^*\}E\{yx\}) \quad (9)$$

其中,  $x = x(f, t)$ ,  $E$  为求均值运算。算法每次迭代还需要进行正交化处理以使算法寻找到不同的极值点。

各频点分离完成后,经过排序和尺度补偿,在保证各频点分量排列正确且幅度得到恢复的情况下进行短时傅里叶逆变换,最后将得到的时域信号重新截取组合得到对源信号的估计。

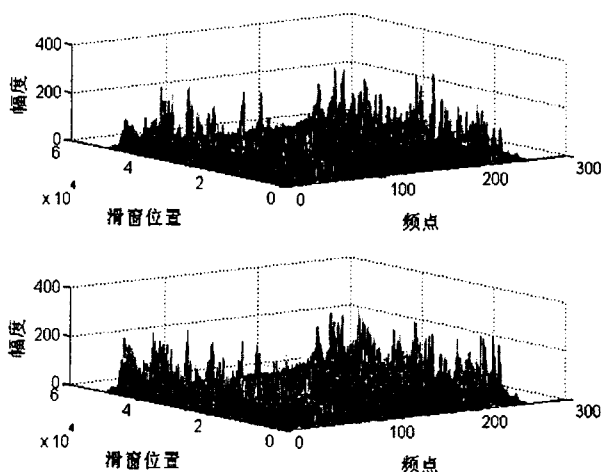


图 1 观测信号加汉明窗经 256 点短时傅里叶变换后的时-频分布图

## 2 频域盲解卷积算法基本制约因素

众所周知,排序和幅度不确定性是传统 ICA 算法的固有问题。在瞬时混合 ICA 问题中,排序不确定性最终只表现为分离信号输出通道的不确定。而在频域盲解卷积问题中,由于涉及到多个频点的分离,最后需要对各频点的分离结果进行重新组合形成恢复的时域信号,因而必须保证各频点分离信号对应同一源信号,这就造成了“排序不确定性”(Permutation Ambiguity)。同时,混合和解混过程使信号幅度获得一定的“增益”,“尺度不确定性”(Scaling Ambiguity)也因此产生。

频域算法的另一个缺陷源自语音信号的非平稳特性。语音信号频域盲解卷积算法的第一步是对观测信号进行加窗短时傅里叶变换,窗长的选取对分离效果至关重要。依据时域卷积等价频域相乘的原理,为了使用循环卷积近似线性卷积以便进行快速傅里叶变换,理论上选取窗长  $T$  为混合滤波器长度  $P$  的 2 倍。而选取窗长的另一个原则是要求窗内语音信号尽可能是平稳的,由于语音信号是短时平稳的,其只能在 10 ~ 30 毫秒左右时间内保持平稳,从而窗长  $T$  和混合滤波器长度  $P$  形成了一对制约关系。同时,对观测信号进行加窗傅里叶变换时,如果观测信号长度较短或由于

窗长较长,导致窗移动次数较少,进而导致每个频点参加分离运算的信号长度较短,就会很大程度上削弱 ICA 的源信号独立性假设,使得各频点分离效果变差。至此又形成了窗长和观测信号长度的制约关系。所以,算法实施时,应设法解决排序和尺度不确定问题,并对窗长等参数权衡选取。

### 3 一种改进的针对制约因素的频域盲解卷积算法

若  $W(f)$  为某频点分离矩阵,为了解决尺度不确定性问题,文中采用文献[5]中的方法:

$$W(f) \leftarrow \text{diag}(W^{-1}(f))W(f) \quad (10)$$

各频点输出信号经过(10)式解决尺度模糊后,每路输入转化为两路或多路输出,可视为一种“分裂”,即:

$$U_{ij}(f, t) = \text{diag}(W^{-1}(f))W(f)[0 \cdots Y_i(f, t) \cdots 0] \quad i=1, 2, \cdots, N \quad (11)$$

其中  $Y_i(f, t)$  表示解决尺度模糊前输入的已分离的第  $i$  路信号,  $U_{ij}(f, t)$  表示经过“分裂”后,第  $i$  个输入在第  $j$  个传感器的输出。

为了解决排序不确定性问题,首先获得各频点输出信号的包络,再求得总包络

$$\text{Env}(U_i(f, t)) = \frac{1}{(2d-1)n} \sum_{f'=f-d}^{f+d} \sum_{n=1}^n |U_{ik}(f', t_s)| \quad (12)$$

其中,  $\text{Env}$  代表总包络,  $d$  为滑动平均宽度,  $n$  为分裂后扩展通道数。

最后,对求得的总包络各频点利用相关性大小进行排序:

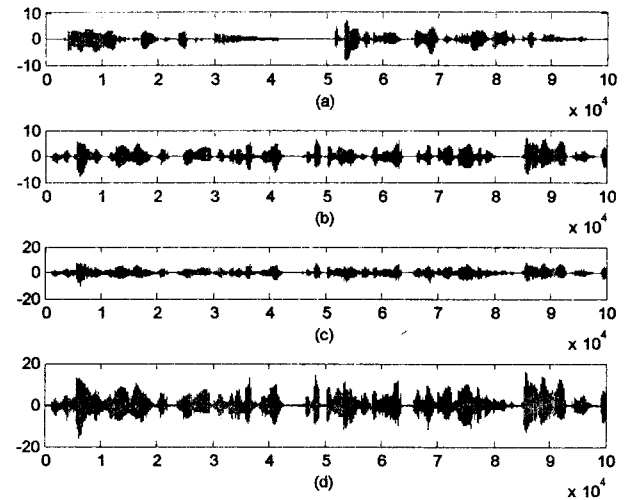
$$Y_j(f_{i+1}, t) = \{U | \max_{m=i-d}^i \sum_{n=1}^N \text{corr}(Y_j(f_m, t), U_n(f_{i+1}, t))\} \quad (13)$$

其中,  $f_m$  代表第  $m$  个频点,  $n$  表示输出通道数,  $d$  表示宽度。实际运用中,只需对(12)式求得的包络具体运用(13)式即可。

### 4 仿真实验

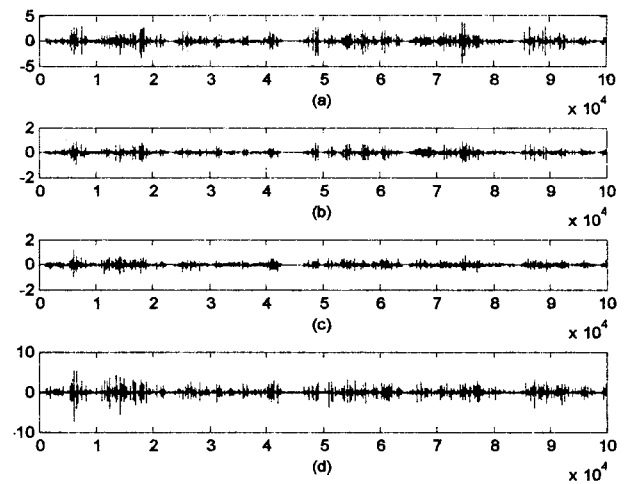
为了验证文中算法的有效性,对卷积混合语音信号进行分离实验。实验语音文件为 Stereo Audio Source Separation Evaluation Campaign 网站提供的算法开发用语音源信号,两路源信号分别为男声和女声的英语和日语语音。信号长度为 100000 点,采用一个 8 阶滤波器组进行卷积混合。源信号和混合信号的波形见图 2,对混合信号计算 STFT 时,采用汉明窗,重叠率为 0.5,先将观测信号转换到时频域如图 1,随后使用上述方法进行各频点的分离并进行排序和尺度恢

复。图 3 为最终分离结果。



(a)、(b)为源信号波形, (c)、(d)为仿真卷积混合信号波形

图 2 源信号与混合信号波形



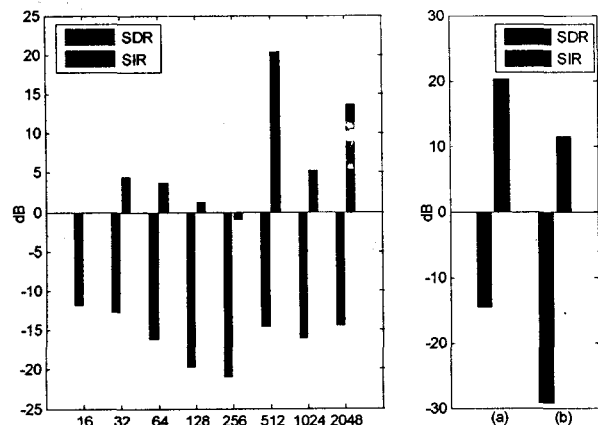
(a)(b)和(c)(d)分别为对同一源信号的分离结果

图 3 仿真实验分离结果

由第 3 部分讨论可知,理论上所加窗长应选取为混合滤波器长度的 2 倍,但实际中窗长往往和别的因素互相制约。为了寻求最佳的窗长,实验中分别使用长度不同的 8 个窗进行对比实验并引入量化指标 SDR 和 SIR<sup>[6]</sup>。SIR 表征混合信号的分离程度,SDR 可用于表征分离信号的声学质量。

不同窗长的分离效果评估见图 4 左图。通过对比实验发现,当所加窗长为理论值即为混合滤波器长度 2 倍左右时,虽然 SDR 较其他窗长时理想,但表征混合信号分离程度的 SIR 过小,说明分离结果中源信号仍呈混合状态,分离效果并不理想,其原因应是当窗长度小于 100 时,相邻频点之间“间距”过大,相关性过弱,使得排序无法有效进行。当窗长逐渐增大时,虽然 SDR 有一定的减少,但 SIR 呈上升趋势。当窗长达到 512 点时,各指标均较为理想,算法获得最佳的分离效果且从音质上较其它结果有一定改善。随后若继续增

加窗长,各指标下降分离效果变坏。此时第 3 部分中讨论的第 2 对制约关系开始显现。



左图为不同长度加窗的性能指标对比,右图(a)为使用总包络排序的分离结果,(b)为对输出信号包络直接排序的分离结果

图 4 分离效果评估

为了验证使用“总包络”进行排序的有效性,在所有实验参数相同的情况下分别使用总包络和输出信号包络进行排序,并计算 SDR 和 SIR 对比分离结果,实验结果对比分析见图 4 右图,可见使用“总包络”进行排序比直接使用输出信号包络排序具有更大的分离程度和信号质量。原因在于由一路信号分裂成多路信号时,信号的幅度和相位信息得到了较好的补偿,再由分裂后的多路信号求得“总包络”时排序信息得到充分保留。而使用分裂信号包络直接排序造成分裂前单路信号的排序信息一定程度丢失,故影响分离效果。

## 5 结束语

通过对卷积语音信号频域盲解卷积全过程的论述,较为全面地阐述了频域盲解卷积算法存在的排序、

尺度模糊和窗长选择等问题,并针对这些问题提出了解决方案。通过反复实验,证明了算法实际实施时有时会与理论情况产生一定差异,并说明了利用“总包络”进行排序的有效性。最后通过对卷积混合语音信号的分离实验验证了文中算法分离卷积混合语音信号的能力。

由于实验手段和算法自身可能存在的局限,分离音质仍需提高,从图 4 中可见即使最佳分离状态下,SDR 指标仍不理想。若分离后对分离信号增加补偿或增强环节,预计分离效果还可以进一步改善。

## 参考文献:

- [1] 吴小培,冯焕清,周荷琴. 独立分量分析及其在脑电信号预处理中的应用[J]. 北京生物医学工程,2001,20(1):35-37.
- [2] Araki S, Mukai R, Makino S, et al. Fundamental limitation of frequency-domain blind source separation for convolutive mixtures of speech[J]. IEEE Trans. on Speech and Audio Processing, 2003, 11(2):109-116.
- [3] Ikeda S, Murata N. A method of ICA in time-frequency domain[C]//Proceeding of First International Workshop on Independent Component Analysis and Blind Signal Separation. France: [s. n.], 1999:245-250.
- [4] Li H, Adali T. Gradient and fixed-point complex ICA algorithms based on kurtosis maximization [C]//Proc. IEEE Workshop. [s. l.]: [s. n.], 2006:85-90.
- [5] Sawada H, Mukai R, Araki S, et al. A robust and precise method for solving the permutation problem of frequency-domain blind source separation[J]. IEEE Trans. on Speech and Audio Processing, 2004, 12(15):530-538.
- [6] F'evotte C, Gribonval R, Vincent E. BSS EVAL toolbox user guide[R/OL]. IRISA, Rennes, France, 2005. [http://www.irisa.fr/metiss/bss\\_eval/](http://www.irisa.fr/metiss/bss_eval/).

(上接第 56 页)

## 参考文献:

- [1] Marian A, Amer-Yahia S, Koudas N. Adaptive Processing of Top-k Queries in XML[C]//Proc. of the Int'l Conf. on ICDE. Tokyo, Japan: [s. n.], 2005:162-173.
- [2] Fuhr N, Grossjohann K. XIRQL: An Extension of XQL for Information Retrieval[C]//ACM SIGIR Workshop on XML and Information Retrieval. Athens, Greece: [s. n.], 2000:132-147.
- [3] Amer-Yahia S, Lakshmanan L, Pandit S. FlexPath: Flexible Structure and Full-Text Querying for XML[C]//Proc. of the Int'l Conf. on SIGMOD. Paris, France: [s. n.], 2004:83-94.

- [4] Weigel F, Meuss H, Schulz K U, et al. Content and Structure Indexing and Ranking XML[C]//de la Chimie M. Proc. of the 7th Int'l Workshop on WebDB. [s. l.]: [s. n.], 2004:67-72.
- [5] Amer-Yahia S, Curtmola E, Deutsch A. Flexible and efficient XML search with complex full-text predicates[C]//Proc. of the Int'l Conf. on SIGMOD. Chicago: [s. n.], 2006:575-586.
- [6] 王静,孟小峰,王宇,等.以目标节点为导向的 XML 路径查询处理[J].软件学报,2005,16(5):827-837.
- [7] Polyzotis N, Garofalakis M, Ioannidis Y. Approximate XML Query Answers[C]//Proc. of the Int'l Conf. on SIGMOD. Paris, France: [s. n.], 2004:263-274.