

采用二阶特征窗的语音盲分离方法

徐 舜, 刘郁林, 朱行涛

(重庆通信学院DSP实验室 重庆 沙坪坝区 400035)

【摘要】利用语音信号的短时平稳特性,提出了一种二阶特征窗语音盲分离方法。该方法采用新的联合差分相关矩阵白化算法去除有色噪声影响;用长度等于语音信号基音周期的等距特征窗连续分割预白化观测数据,在每个加窗的数据帧计算不同的时滞协方差矩阵。利用联合近似对角化时滞协方差矩阵集合得到旋转参数,最终达到语音信号的盲分离。该方法消除了有色噪声的影响,只需用到二阶信息就能很好地分离语音信号。仿真实验验证了该方法的有效性。

关键词 盲信号分离; 特征窗; 相关矩阵; 白化
中图分类号 TN912 **文献标识码** A

Second-Order Characteristic Window Approach to Blind Speech Signal Separation

XU Shun, LIU Yu-lin, and ZHU Xing-tao

(DSP Lab, Chongqing Communication College Shapingba Chongqing 400035)

Abstract In this paper, a second-order characteristic window approach to blind signal separation based on short-time stationarity of speech signals is proposed. This approach uses a new whitening algorithm via joint difference correlation matrix to remove the effect of colored noise for one thing, then uses characteristic window that the length is approximately equal to the fundamental period to divide the whitening observation data into series. In every different data frame, time-lag covariance matrices are computed and jointly approximately diagonalized to estimate the rotation parameters which could make the sources obtained. This approach avoids colored noises and separates speech signals only by means of second-order information. Its effectiveness is shown via computer simulation.

Key words blind signal separation; characteristic window; correlation matrix; whitening

盲信号分离(BSS)是指在不知道源信号和缺乏混合系统参数的情况下,根据输入源信号的统计特性,仅由观测信号恢复出各个源信号的过程^[1]。盲信号分离问题成为重要的研究课题,并在许多领域得到了发展。

基于二阶的方法^[2-3]和新的联合对角化方法^[4-5]在基于信号平稳性的约束和利用高阶累积量的条件下,能够较好地实现盲信号分离。但语音是时变、非平稳的随机信号,其清音部分的分布类似于高斯白噪声^[6],四阶累积量为零。因此上述算法不适合处理语音信号的盲分离,特别是当环境中存在非i.i.d的噪声时,算法的性能会随着SNR的降低而迅速退化^[7]。本文提出了一种新的二阶特征窗(SOCW)方法,利用语音信号的短时平稳和有色噪声非消失时间相关特性之间的差异,通过特征向量的不断更新找到一组参数使联合差分相关矩阵正定,对参数化的联合差分相关矩阵进行奇异值分解(SVD),得到全

局正交化的白化矩阵,达到信号的白化;然后根据语音信号的非平稳特性和时序结构,可以将预白化的观测数据按照语音信号的基音周期分成连续的特征窗;并估计每个特征窗的时滞协方差矩阵,用联合近似对角化的方法达到分离出语音信号的目的。从实验仿真可以看到,对比二阶的SOBI算法和联合对角化JADE算法,该方法能更好地实现当存在不同分布有色噪声时的语音信号分离。

1 信号模型

本文采用典型的线性、瞬时盲信号分离问题模型,如图1所示。

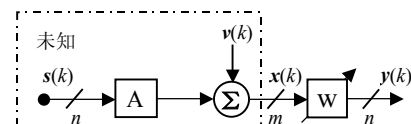


图1 盲分离模型

收稿日期: 2006-10-17; 修回日期: 2007-04-22

基金项目: 国家自然科学基金(60672157)

作者简介: 徐 舜(1980-),男,硕士,主要从事盲信号处理、语音信号处理等方面的研究。

$$\mathbf{x}(k)=\mathbf{A}\mathbf{s}(k)+\mathbf{v}(k) \quad (1)$$

$$\mathbf{y}(k)=\mathbf{W}\mathbf{x}(k) \quad (2)$$

式中 $\mathbf{A} \in R^{m \times n}$ 为未知的列满秩混合矩阵; $\mathbf{s}(k)=[s_1(k), s_2(k), \dots, s_n(k)]^T$ 为彼此独立 n 维的源信号矢量; $\mathbf{x}(k)=[x_1(k), x_2(k), \dots, x_m(k)]^T$ 为 m 维观测矢量; $\mathbf{v}(k)=[v_1(k), v_2(k), \dots, v_m(k)]^T$ 为与信号源不相关的加性噪声矢量。为了从混合信号 $\mathbf{x}(k)$ 中分离出源信号 $\mathbf{s}(k)$, 需要通过寻找一个 $n \times m$ 阶的满秩线性变换(或分离)矩阵 \mathbf{W} , 以便由 $\mathbf{y}(k)=\mathbf{W}\mathbf{x}(k)$ 定义的输出矢量 $\mathbf{y}=[y_1, y_2, \dots, y_n]^T$ 的各分量尽可能相互独立, 这样就可以获得关于随机向量 $\mathbf{x}(k)=[x_1(k), x_2(k), \dots, x_m(k)]^T$ 的盲源分离和混合矩阵 \mathbf{A} 的辨识。

2 空间预白化

根据语音信号特征, 本文利用联合差分矩阵提出了一种新的白化方法, 可以有效地去除有色噪声的影响。

2.1 定义联合差分相关矩阵

文献[6]阐述了短时基音周期的估计方法, 本文假设已估计出的语音信号基音周期为 t , 因此根据语音信号短时平稳特征, 以 8 000 Hz 为抽样频率, 取一个基音周期信号长度 $8\,000 \times t$ 为特征窗 T 。

定义差分相关矩阵^[8]为:

$$\Delta \mathbf{R}_x(T_i, T_j, p_l) = \mathbf{R}_x(T_i, p_l) - \mathbf{R}_x(T_j, p_l) \quad (3)$$

式中 T_i 和 T_j 为两个不重叠的等长特征窗; $\mathbf{R}_x(T_i, p_l)$ 为对于特征窗 T_i 的时滞 p_l 相关矩阵。

定义联合差分相关矩阵为:

$$\Delta \mathbf{R}_x(\boldsymbol{\eta}) = \sum_i \sum_j \sum_l \eta_{ijl} \Delta \mathbf{R}_x(T_i, T_j, p_l) \quad (4)$$

$$i=1,2,\dots,M; j \geq i; l=1,2,\dots,N$$

式中 $\boldsymbol{\eta}$ 为一组联合参数矢量; M 为特征窗的个数; N 为时滞 p 的个数。

2.2 白化方法的具体步骤

(1) 观测数据 $\mathbf{x}(k)$ 分成 M 个不重叠的特征窗, 并且估计了一组差分相关矩阵 $\Delta \hat{\mathbf{R}}_x(T_i, T_j, p_l)$, $i=1,2,\dots,M; j \geq i; l=1,2,\dots,N$ ($M=10, N=5$, 每个窗样本数为 160)。令 $K=MN(M-1)/2$, 可构造 $m \times mK$ 矩阵 $\mathbf{R} = [\Delta \hat{\mathbf{R}}_x(T_1, T_2, p_1), \Delta \hat{\mathbf{R}}_x(T_2, T_3, p_2), \dots, \Delta \hat{\mathbf{R}}_x(T_{M-1}, T_M, p_N)] = [\Delta \hat{\mathbf{R}}_x(\zeta_1), \Delta \hat{\mathbf{R}}_x(\zeta_2), \dots, \Delta \hat{\mathbf{R}}_x(\zeta_K)]$ 然后计算 \mathbf{R} 的奇异值分解 SVD, 即 $\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, 其中, $\mathbf{U} = [\mathbf{U}_s, \mathbf{U}_v] \in R^{m \times mK}$; $\mathbf{U}_s = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \in R^{m \times n}$; $\mathbf{V} \in R^{mK \times mK}$ 为正交矩阵; $\boldsymbol{\Sigma}$ 为一个 $m \times mK$ 矩阵, 其左 n 列包含 $\text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\}$, 其右 $(mK-n)$ 列为

零(噪声功率小于信号功率)。

(2) 当 $d=1,2,\dots,K$, 计算 $\mathbf{R}_d = \mathbf{U}_s^T \Delta \hat{\mathbf{R}}_x(\zeta_d) \mathbf{U}_s$ 。

(3) 选择任何非零的初始参数矢量 $\boldsymbol{\eta} = [\eta_1, \eta_2, \dots, \eta_K]^T$ 。

(4) 计算 $\tilde{\mathbf{R}} = \sum_{d=1}^K \eta_d \mathbf{R}_d$ 。

(5) 计算 $\tilde{\mathbf{R}}$ 的 EVD, 并检查 $\tilde{\mathbf{R}}$ 是否正定。如果正定, 转到步骤(7), 否则转到步骤(6)。

(6) 选择对应于 $\tilde{\mathbf{R}}$ 最小特征值(如果有多个, 任取一个)的特征向量 \mathbf{u} , 用 $\boldsymbol{\eta} + \boldsymbol{\delta}$ 代替 $\boldsymbol{\eta}$ 进行更新。其中, $\boldsymbol{\delta} = \frac{[\mathbf{u}^T \mathbf{R}_1 \mathbf{u} \dots \mathbf{u}^T \mathbf{R}_K \mathbf{u}]^T}{\|[\mathbf{u}^T \mathbf{R}_1 \mathbf{u} \dots \mathbf{u}^T \mathbf{R}_K \mathbf{u}]\|}$, 转步骤(4)。

(7) 计算联合差分相关矩阵 $\tilde{\mathbf{R}}_x(\boldsymbol{\eta}^*) = \sum_{d=1}^K \eta_d \times \Delta \hat{\mathbf{R}}_x(\zeta_d) = \sum_i \sum_j \sum_l \eta_{ijl} \Delta \hat{\mathbf{R}}_x(T_i, T_j, p_l)$, 对 $\tilde{\mathbf{R}}_x(\boldsymbol{\eta}^*)$ 进行奇异值分解(SVD)可得, $\tilde{\mathbf{R}}_x(\boldsymbol{\eta}^*) = [\mathbf{U}_s, \mathbf{U}_N] \times \begin{bmatrix} \boldsymbol{\Sigma}_s & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Sigma}_N \end{bmatrix} \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_N^T \end{bmatrix}$, 其中, $\boldsymbol{\eta}^*$ 是在算法收敛后, 即矩阵 $\tilde{\mathbf{R}}$ 正定后 η_d 的一组参数; \mathbf{U}_s 包含 n 个主奇值 $\boldsymbol{\Sigma}_s = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\}$ 对应的特征向量。

(8) 实现稳健的白化 $\bar{\mathbf{x}}(k) = \mathbf{Q}\mathbf{x}(k)$, 其中,

$$\mathbf{Q} = \sum_s \frac{1}{2} \mathbf{U}_s^T$$

3 信号分离

3.1 估计特征窗时滞协方差矩阵

将预白化观测数据分成连续等长的特征窗, 估计每个窗的时滞协方差矩阵 $\mathbf{R}_{\bar{\mathbf{x}}_{T_i}} = \mathbf{R}_{\bar{\mathbf{x}}}(T_i, p)$, $i=1,2,\dots,L$, 组成联合相关窗矩阵集合 $\mathfrak{R}_{\bar{\mathbf{x}}T} = \{\mathbf{R}_{\bar{\mathbf{x}}}(T_1, p), \mathbf{R}_{\bar{\mathbf{x}}}(T_2, p), \dots, \mathbf{R}_{\bar{\mathbf{x}}}(T_L, p)\}$ 。

3.2 计算 L 个相关窗矩阵的联合对角化器 \mathbf{U}

使用任意酉矩阵 \mathbf{U} , 使 L 个 $n \times n$ 矩阵 $\mathbf{R}_{\bar{\mathbf{x}}T}$ 组成的矩阵集合 $\mathfrak{R}_{\bar{\mathbf{x}}T}$ 联合对角化的代价函数定义为^[9]:

$$J(\mathfrak{R}_{\bar{\mathbf{x}}T}, \mathbf{U}) \stackrel{\text{def}}{=} \sum_{i=1}^L \text{off}(\mathbf{U}^T \mathbf{R}_{\bar{\mathbf{x}}T_i} \mathbf{U}) \quad (5)$$

一酉矩阵称为矩阵集合 $\mathfrak{R}_{\bar{\mathbf{x}}T}$ 的联合对角化器(joint diagonalizer), 若它在所有酉矩阵集合中能够使式(5)的代价函数值最小^[10]。

考虑 L 个 2×2 的子矩阵 $\mathbf{R}_i = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix}, i=1,2,\dots,L$ 的联合对角化。酉矩阵 \mathbf{U} 用实 Givens 旋转参数化,

$U = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$, 目的是寻找联合对角化器

U , 使 $R_i^* = U^T R_i U$ ($i=1, 2, \dots, L$) 可以最小化式(5)。

令 $R_i^* = \begin{bmatrix} a_i^* & b_i^* \\ c_i^* & d_i^* \end{bmatrix}$, 要使代价函数 J 最小, 根据

酉变换的二范数不变性即是求 θ , 使 $\sum_{i=1}^L (|a_i^*|^2 + |d_i^*|^2)$ 最大化。又因 $2(|a_i^*|^2 + |d_i^*|^2) = |a_i^* - d_i^*|^2 + |a_i^* + d_i^*|^2$, 而且迹 $a_i^* + d_i^*$ 在酉变换中保持不变, 因此, 代价函数 J 的最小化等价于在每个 Givens 旋转时取 S 的最大化^[10], 则有:

$$S = \sum_{i=1}^L |a_i^* - d_i^*|^2 \quad (6)$$

通过计算 $a_i^* - d_i^* = (a_i - d_i)\cos(2\theta) - (b_i + c_i)$

$\sin(2\theta)$, 定义 $\mathbf{u}^T = [a_1^* - d_1^*, a_2^* - d_2^*, \dots, a_L^* - d_L^*]$;

$\mathbf{v}^T = [\cos(2\theta), -\sin(2\theta)]$; $\mathbf{g}_i^T = [a_i - d_i, b_i + c_i]$;

$\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_L]$, 则有:

$$S = \mathbf{u}^T \mathbf{u} = \mathbf{v}^T \mathbf{G} \mathbf{G}^T \mathbf{v} \quad (7)$$

因为 $\mathbf{v}^T \mathbf{v} = 1$, 要使 S 最大, 由二次型知识可知, \mathbf{v} 必须是实对称阵 $\mathbf{G} \mathbf{G}^T$ 最大特征值对应的特征向量。这样, Givens 参数就可以从特征向量的值导出, 也可得到联合对角化器 U 。 $n \times n$ 矩阵的联合对角化器可以通过 $(C_n^2)^2$ 次 2×2 子矩阵的联合对角化更新 Givens 旋转参数后得到。

3.3 估计源信号和混合矩阵

得到了联合对角化器 U , 可以估计出分离矩阵 $W = U^T Q$ 。因此, 源信号估计为 $\hat{\mathbf{s}}(k) = W \mathbf{x}$, 混合矩阵估计为 $\hat{A} = Q^+ U$ 。

4 仿真结果

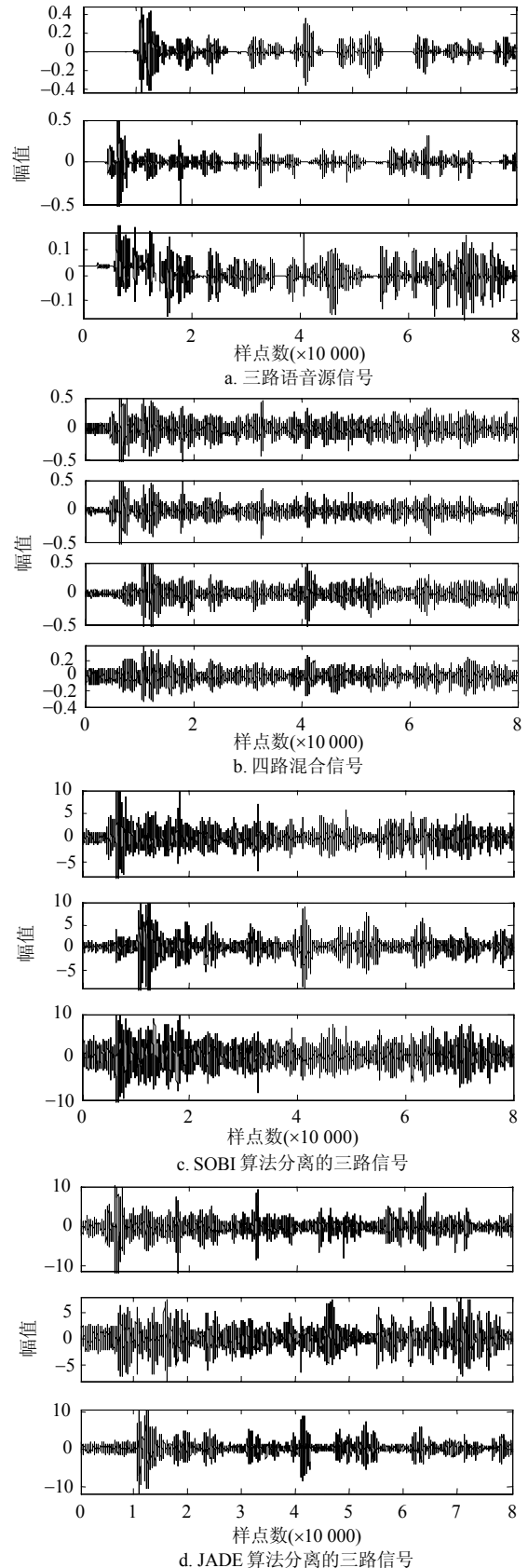
定义性能指标为:

$$PI = \frac{1}{n(n-1)} \sum_{i=1}^n \left\{ \left(\sum_{k=1}^n \frac{|g_{ik}|}{\max_i |g_{ij}|} - 1 \right) + \left(\sum_{k=1}^n \frac{|g_{ki}|}{\max_i |g_{ji}|} - 1 \right) \right\} \quad (8)$$

式中 g_{ij} 为全局系统矩阵 $G = WA$ 的第 (i, j) 个元素; $\max_i |g_{ij}|$ 为 G 的第 i 行矢量各个元素取绝对值后的最大值。同理, $\max_i |g_{ji}|$ 代表 G 的第 i 列矢量各个元素取绝对值后的最大值。分离性能越好, PI 值越接近零。

选择3个语音信号源(样点数为80 000), A 为

4×3 的混合矩阵, 输出的观测信号叠加四路不同分布的有色非语音音频噪声, 采用三种不同算法分离的波形如图2所示。



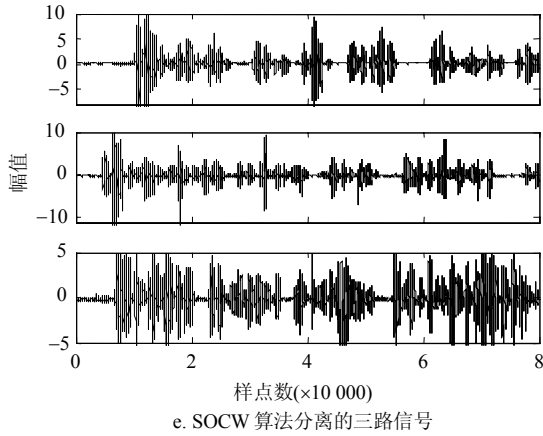


图2 三种算法分离有色噪声干扰下的语音混合信号

从图中可以看到, 对于有色噪声条件下语音信号的分离二阶的SOBI算法基本已经失效; JADE算法虽然能勉强分离出源信号的包络, 但已经受到很强的噪声干扰, 无法分辨语音; 只有SOCW算法能有效地分离出语音信号。

本文用PI值定量分析SOCW算法的性能, 如图3所示。

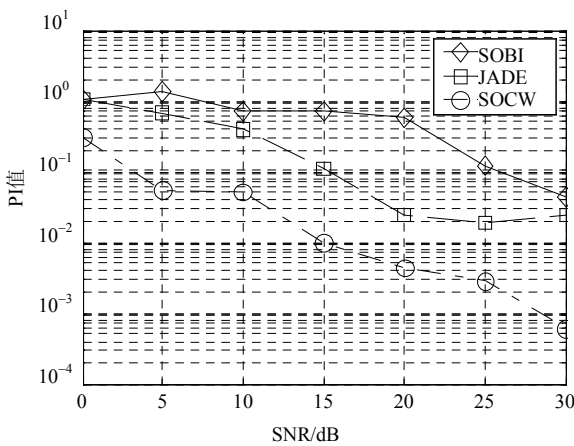


图3 三种算法性能作为SNR的函数比较

5 结束语

针对语音特性, 本文提出了一种二阶特征窗(SOCW)方法, 利用语音信号相对其他音频信号所不同的短时平稳特征, 先用联合差分相关矩阵白化算法去除分离过程中有色噪声的影响, 然后联合对角化以语音基音周期为长度的特征窗相关矩阵, 达到分离语音信号的目的。实验仿真验证了该方法的有效性。

参 考 文 献

- [1] HAYKIN S. Unsupervised adaptive filtering, volume I: blind source separation[M]. Canada: John Wiley & Sons Ltd, 2000.
- [2] BUCHNER H, AICHNER R, KELLERMANN W. A generalization of blind source separation algorithms for convolutive mixtures based on second-order statistics[J]. IEEE Trans on Speech and Audio Processing, 2005, 13(1): 120-134.
- [3] CHOI S, CICHOCKI A. Algebraic differential decorrelation for nonstationary source separation[J]. Electronics Letters, 2001, 37(23): 1414-1415.
- [4] CARDOSO J F. Multidimensional independent component analysis[C]//In Proc ICASSP. [S.l.]: [s.n.], 1998.
- [5] FUXIANG W, ZHONGKAN L, JUN Z. A new joint diagonalization algorithm with application in blind source separation[J]. IEEE Signal Processing Letters, 2006, 13(1): 41-44.
- [6] 杨行峻, 迟惠生. 语音信号数字处理[M]. 北京: 电子工业出版社, 1995.
- [7] PEDERSEN M S, WANG D. Separating underdetermined convolutive speech mixtures[C]//ICA2006. Charleston, USA: [s.n.], 2006.
- [8] CHOI S, CICHOCKI A. Algebraic differential decorrelation for nonstationary source separation[J]. Electronics Letters, 2001, 37(23): 1414-1415.
- [9] 张明键. 盲分离算法的研究[D]. 广州: 华南理工大学, 2004.
- [10] 张贤达, 保 铮. 通信信号处理[M]. 北京: 国防工业出版社, 2002.

编辑 黄 莘