

卷积混迭语音信号的联合块对角化盲分离方法^{*}

张 华¹ 冯大政¹ 庞继勇²

(1 西安电子科技大学雷达信号处理国家重点实验室 西安 710071)

(2 西安电子科技大学 ISN 国家重点实验室 西安 710071)

2008 年 6 月 10 日收到

2008 年 9 月 5 日定稿

摘要 针对语音信号的卷积混迭模型, 利用不同语音信号之间的近似独立和短时平稳特性, 提出一种基于信号二阶统计量的联合块对角化方法, 解决超定卷积盲分离问题。该方法采用非对角线上各子矩阵 F -范数的平方和作为联合块对角化性能的评判准则, 将原四次代价函数转化为一组较为简单的二次子代价函数, 每一子代价函数用于估计酉混迭矩阵的一个子矩阵。依次最小化各子函数, 迭代搜索代价函数最小点, 得到混迭矩阵的估计。理论分析及实验结果表明, 所提方法不仅能够达到与类 Jacobi 经典方法同样好的分离效果, 并且具有更低的计算复杂度、更快的收敛速度和对传输信道阶数、迭代初始值不敏感的特点。

PACS 数: 43.60

Blind convolutive separation algorithm for speech signals via joint block diagonalization

ZHANG Hua¹ FENG Dazheng¹ PANG Jiyong²

(1 National Lab. of Radar Signal Process., Xidian Univ. Xi'an 710071)

(2 State Key Lab. of Integrated Service Networks, Xidian Univ. Xi'an 710071)

Received Jun. 10, 2008

Revised Sept. 5, 2008

Abstract A blind speech source separation algorithm for the overdetermined convolutive mixture model in time-domain is proposed via joint block-diagonalization based on the mutual-independence property and the short-time stationary of the speech signals. Taking the sum of the F -norms of all off-diagonal sub-matrices as a criterion, a novel joint block-diagonalization algorithm is proposed to estimate the whole mixture matrix through minimizing a sequence of quadratic subfunctions corresponding to mixture submatrices. Both theoretical analysis and simulation results show that the proposed algorithm has much lower complexity and faster convergence speed than the classical Jacobi-like method with no performance loss. In addition, there almost are no obvious impacts of the channel order and initialization values on the convergence speed.

引言

盲源分离是在源信号和传输信道参数未知的情况下, 根据输入源信号的统计特性, 仅由观测信号估计信道参数、恢复源信号波形的过程。作为信号处理领域中的研究热点之一, 近 20 年来, 盲源分离技术

持续发展、大量有效方法不断提出, 在图像处理、语音处理、生物医学和地震波检测等领域得到了广泛的应用^[1]。

盲源分离所研究的混迭模型主要分为瞬时混迭和卷积混迭两类。其中, 瞬时混迭模型, 因其较为简单, 已经得到了广泛而完善的研究。由 Cardoso 等人提出的, 基于一组四阶累积量的联合对角化方法

^{*} 国家自然科学基金 (60672128, 60702057) 和国家 863 计划课题 (2007AA01Z288)。

JADE(Joint Approximate Diagonalization of Eigenmatrices)^[2], 有效地解决了多独立分量的瞬时混迭问题。其后, 基于源信号的其它统计特性以及不同的代价函数, 人们又提出了多种有效的联合对角化方法, 如二阶盲辨识方法 SOBI(Second-Order Blind Identification)^[3]、并行因子分析法 PARAFAC (PARAllel FACtor analysis)^[4]、子空间拟合 SS-fitting (Sub-Space fitting)^[5]、二次对角化方法 QDIAG (Quadratic DIAGonalization algorithm)^[6] 等等。

然而, 实际应用场景中, 因传输信道的多径效应, 传声器接收到的信号一般是源信号与传输信道的卷积混迭信号。于是, 卷积盲分离问题逐渐受到越来越多的关注^[7-15]。不同于瞬时混迭模型, 卷积盲分离的数学模型以及求解过程则更为复杂。一些方法^[9-10]将时域卷积混迭问题转换为频域瞬时混迭问题, 采用已有的瞬时混迭盲分离方法进行求解。然而, 因为盲分离方法固有的排列模糊和尺度模糊问题, 各频点之间排列和尺度的不一致会造成合成信号频谱的重新混迭和畸变, 产生非线性失真。如果直接在时域进行^[11-15], 则排列和尺度模糊问题不会引发分离信号的非线性失真。文献 12—文献 15 利用类 Jacobi 方法解决联合块对角化问题, 实现了卷积混迭信号的时域盲分离。针对经典的类 Jacobi 方法收敛速度慢, 对传输信道阶数和初始值较为敏感的问题, 本文提出一种分步迭代的联合块对角化方

法, 分步估计每个独立信号到达传声器阵列的传输路径, 有效地克服了类 Jacobi 方法的不足。分析及实验表明, 本文方法可更加高效地分离卷积混迭的语音信号。

1 问题描述

真实环境中, 传声器所捕捉到的信号可看作是原始语音信号和信道冲激响应的卷积输出。信道冲激响应表征了声源和传声器之间的传递函数, 可用一个有限阶脉冲响应 (FIR) 滤波器来近似表示。假定有 N 个源信号 $s_j(j = 1, \dots, N)$, M 个接收传声器, 信道多径阶数为 L , 则第 n 时刻第 i 路传声器上的观测信号可表示为:

$$z_i(n) = \sum_{j=1}^N \sum_{l=0}^{L-1} h_{ij}(l) s_j(n-l), \quad i = 1, \dots, M \quad (1)$$

其中, $h_{ij}(l)$ 表示第 j 个源信号到达第 i 个传声器的第 l 径信道冲激响应。

对观测信号取长为 L' 的滑窗, 定义 ML' 维观测数据矢量 $\bar{\mathbf{z}}(n) = [\mathbf{z}_1^T(n), \dots, \mathbf{z}_M^T(n)]^T$, 其中, $\mathbf{z}_i(n) = [z_i(n), z_i(n-1), \dots, z_i(n-L'+1)]^T$, 则式 (1) 可表示为:

$$\bar{\mathbf{z}}(n) = \bar{\mathbf{H}} \bar{\mathbf{s}}(n), \quad (2)$$

其中,

$$\bar{\mathbf{s}}(n) = [\mathbf{s}_1^T(n), \dots, \mathbf{s}_N^T(n)]^T, \quad \mathbf{s}_j(n) = [s_j(n), s_j(n-1), \dots, s_j(n-L-L'+2)]^T,$$

$$\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} & \cdots & \mathbf{H}_{1N} \\ \mathbf{H}_{21} & \mathbf{H}_{22} & \cdots & \mathbf{H}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{M1} & \mathbf{H}_{M2} & \cdots & \mathbf{H}_{MN} \end{bmatrix}, \quad \mathbf{H}_{ij} = \begin{bmatrix} h_{ij}(0) & \cdots & h_{ij}(L-1) & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & h_{ij}(0) & \cdots & h_{ij}(L-1) \end{bmatrix} \in C^{L' \times (L'+L-1)}$$

此时, 式 (1) 所示的卷积混迭模型就简化为式 (2) 所示的块独立的信号 $\bar{\mathbf{s}}(n)$ 与扩展的传输信道 $\bar{\mathbf{H}}$ 的超定线性瞬时混迭模型。一般地, 有如下假设: (1) 传输矩阵 $\bar{\mathbf{H}}$ 为列满秩矩阵, 即观测滑窗长度 L' 需满足 $ML' \geq N(L'+L-1)$ 。简便起见, 记 $L_b = L'+L-1$; (2) 源信号 $s_j(n)(j = 1, \dots, N)$ 是零均值、相互独立的平稳信号。

令 $\bar{\mathbf{y}}(n) = [\mathbf{y}_1^T(n), \mathbf{y}_2^T(n), \dots, \mathbf{y}_N^T(n)]^T = \mathbf{W}_s \bar{\mathbf{z}}(n)$, 其中, $\mathbf{W}_s$ 为 $NL_b \times ML'$ 维分离矩阵, $\mathbf{y}_j(n) = [y_j(n), y_j(n-1), \dots, y_j(n-L_b+1)]^T$ 。盲分离的目的就是通过求解最优的 $\mathbf{W}_s$, 使得 $\bar{\mathbf{y}}(n)$ 的第 j 块输出数据

$\mathbf{y}_j(n)$ 仅为其中一个源信号 $\mathbf{s}_{j'}(n)$ 的滤波版本且与其它源信号独立, $j, j' \in [1, \dots, N]$ 。理想情况下, 有 $\mathbf{W}_s = \bar{\mathbf{H}}^+$, 其中“+”表示矩阵的伪逆。然而, 鉴于盲分离方法固有的排列和尺度不定性, 当分离矩阵满足 $\mathbf{W}_s \bar{\mathbf{H}} = \bar{\mathbf{E}} \bar{\mathbf{D}}$, 即分离数据 $\bar{\mathbf{y}}(n) = \bar{\mathbf{E}} \bar{\mathbf{D}} \bar{\mathbf{s}}(n)$ 时, 我们就称实现了卷积混迭盲分离。其中, $\bar{\mathbf{D}}$ 表示 NL_b 维分块对角方阵 (用来描述尺度不定性), 其位于对角线上的子矩阵由任意 N 个 L_b 维可逆方阵组成; $\bar{\mathbf{E}}$ 表示 NL_b 维的分块置换方阵 (用以描述排列不定性), 其每一行块每一列块有且仅有一个非零的 L_b 维单位子 (方) 阵。

2 分步迭代联合块对角化方法

首先, 对观测数据进行白化预处理, 然后利用白化信号的相关矩阵具有块对角结构的特点, 提出一种新颖的联合块对角化方法来估计块对角结构中的酉矩阵。最后, 结合白化矩阵来估计分离矩阵 $\mathbf{W}_s$ 。

2.1 观测信号白化处理

白化是一种有效的预处理方法, 它可以减少待估计参数的个数, 更重要的是, 对白化后的数据进行联合块对角化处理可获得更加快速和稳定的收敛性能。

与瞬时混迭盲分离中白化处理类似, $NL_b \times ML'$ 维白化矩阵为:

$$\mathbf{W} = \mathbf{D}_s^{-1/2} \mathbf{V}_s^H, \quad (3)$$

其中, $\mathbf{D}_s = \text{diag}[\sigma_1, \dots, \sigma_{NL_b}]$ 和 $\mathbf{V}_s = [\mathbf{v}_1, \dots, \mathbf{v}_{NL_b}]$ 分别由 $\mathbf{R}_{\bar{\mathbf{z}}}(0) = E\{\bar{\mathbf{z}}(n)\bar{\mathbf{z}}^H(n)\}$ 的 NL_b 个最大特征值 σ_i 以及它们所对应的特征向量 $\mathbf{v}_i$ 构成。

白化后的 NL_b 维观测数据矢量可以表示为:

$$\bar{\mathbf{x}}(n) = \mathbf{W} \bar{\mathbf{z}}(n), \quad (4)$$

其零时延自相关矩阵满足:

$$\mathbf{R}_{\bar{\mathbf{x}}}(0) = E\{\mathbf{W} \bar{\mathbf{z}}(n) \bar{\mathbf{z}}^H(n) \mathbf{W}^H\} = \mathbf{I}_{NL_b}.$$

将 $\bar{\mathbf{x}}(n)$ 的 τ_p 时延相关矩阵 $\mathbf{R}_{\bar{\mathbf{x}}}(\tau_p)$ 表示为

$$\begin{aligned} \mathbf{R}_{\bar{\mathbf{x}}}(\tau_p) &= E\{\bar{\mathbf{x}}(n) \bar{\mathbf{x}}^H(n - \tau_p)\} = \\ &= \mathbf{W} \bar{\mathbf{H}} \mathbf{R}_{\bar{\mathbf{s}}}(\tau_p) \bar{\mathbf{H}}^H \mathbf{W}^H = \\ &= \bar{\mathbf{U}} \mathbf{R}_{\bar{\mathbf{s}}}^{-1/2}(0) \mathbf{R}_{\bar{\mathbf{s}}}(\tau_p) \mathbf{R}_{\bar{\mathbf{s}}}^{-1/2}(0) \bar{\mathbf{U}}^H, \end{aligned} \quad (5)$$

其中, $\mathbf{R}_{\bar{\mathbf{s}}}(\tau_p) = E\{\bar{\mathbf{s}}(n) \bar{\mathbf{s}}^H(n - \tau_p)\}$, $\bar{\mathbf{U}} = \mathbf{W} \bar{\mathbf{H}} \mathbf{R}_{\bar{\mathbf{s}}}^{-1/2}(0)$ 为 NL_b 维酉矩阵。由 A2 可知, 块独立信号的相关矩阵 $\mathbf{R}_{\bar{\mathbf{s}}}(\tau_p)$, $\mathbf{R}_{\bar{\mathbf{s}}}(0)$ 都是 NL_b 维块对角方阵, 其子块方阵维数均为 L_b , 则它们的乘积 $\mathbf{R}_{\bar{\mathbf{s}}}^{-1/2}(0) \mathbf{R}_{\bar{\mathbf{s}}}(\tau_p) \mathbf{R}_{\bar{\mathbf{s}}}^{-1/2}(0)$ 也是 NL_b 维块对角方阵, 记作 $\bar{\mathbf{A}}(\tau_p)$ 。则式 (5) 可表示为:

$$\bar{\mathbf{U}}^H \mathbf{R}_{\bar{\mathbf{x}}}(\tau_p) \bar{\mathbf{U}} = \bar{\mathbf{A}}(\tau_p). \quad (6)$$

$$\bar{\mathbf{U}}^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \bar{\mathbf{U}} = \begin{bmatrix} \mathbf{U}_1^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_1 & \cdots & \mathbf{U}_1^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_i & \cdots & \mathbf{U}_1^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_N \\ \vdots & & \vdots & & \vdots \\ \mathbf{U}_i^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_1 & \cdots & \mathbf{U}_i^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_i & \cdots & \mathbf{U}_i^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_N \\ \vdots & & \vdots & & \vdots \\ \mathbf{U}_N^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_1 & \cdots & \mathbf{U}_N^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_i & \cdots & \mathbf{U}_N^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \mathbf{U}_N \end{bmatrix}, \quad (9)$$

式 (6) 表明, 理想情况下, 白化后的数据矢量的二阶时延相关函数具有酉正交块对角化结构。然而, 在实际应用中, 相关矩阵 $\mathbf{R}_{\bar{\mathbf{x}}}(\tau_p)$ 只能通过有限的采样数据估计得到, 且不可避免地受到各种噪声的污染, 矩阵 $\bar{\mathbf{A}}(\tau_p)$ 只能近似满足块对角化结构。一般地, 可通过联合块对角化方法得到矩阵 $\bar{\mathbf{U}}$ 的理想估计 $\bar{\mathbf{U}}_{\text{ideal}}$, 使得矩阵 $\bar{\mathbf{A}}_{\text{ideal}}(\tau_p) = \bar{\mathbf{U}}_{\text{ideal}}^H \mathbf{R}_{\bar{\mathbf{x}}}(\tau_p) \bar{\mathbf{U}}_{\text{ideal}}$ 尽可能地满足块对角矩阵。

当 $\bar{\mathbf{A}}_{\text{ideal}}(\tau_p)$ 严格满足块对角结构时, $\bar{\mathbf{U}}_{\text{ideal}}$ 可以分解为 $\bar{\mathbf{U}}_{\text{ideal}} = \bar{\mathbf{U}} \bar{\mathbf{D}}_1 \bar{\mathbf{E}}_1$ 的形式, 其中, $\bar{\mathbf{D}}_1(\bar{\mathbf{E}}_1)$ 表示与 $\bar{\mathbf{D}}(\bar{\mathbf{E}})$ 同结构的矩阵。选取分离矩阵 $\mathbf{W}_s = \bar{\mathbf{U}}_{\text{ideal}}^H \mathbf{W}$, 则有 $\mathbf{W}_s \bar{\mathbf{H}} = \bar{\mathbf{U}}_{\text{ideal}}^H \mathbf{W} \bar{\mathbf{H}} = \bar{\mathbf{E}}_1^H \bar{\mathbf{D}}_1^H \bar{\mathbf{U}}^H \mathbf{W} \bar{\mathbf{H}} = \bar{\mathbf{E}}_1^H \bar{\mathbf{D}}_1^H \mathbf{R}_{\bar{\mathbf{s}}}^{-1/2}(0)$ 。显然, $\bar{\mathbf{D}}_1^H \mathbf{R}_{\bar{\mathbf{s}}}^{-1/2}(0)$ 与 $\bar{\mathbf{D}}$ 结构相同, $\bar{\mathbf{E}}_1^H$ 与 $\bar{\mathbf{E}}$ 结构也相同。不失一般性, 记 $\bar{\mathbf{E}} = \bar{\mathbf{E}}_1^H$, $\bar{\mathbf{D}} = \bar{\mathbf{D}}_1^H \mathbf{R}_{\bar{\mathbf{s}}}^{-1/2}(0)$, 此时分离矩阵满足 $\mathbf{W}_s \bar{\mathbf{H}} = \bar{\mathbf{E}} \bar{\mathbf{D}}$, 从而可以实现卷积混迭信号的盲分离。

2.2 方法核心思想

以 P 个不同时延下的二阶相关阵 $\mathbf{R}_{\bar{\mathbf{x}}}(\tau_p)$ ($p = 1, \dots, P$), 作为目标矩阵, 简记为 $\mathbf{R}_{\bar{\mathbf{x}}}(p)$ 。令:

$$\bar{\mathbf{A}}(p) = \begin{bmatrix} \mathbf{A}_{11}(p) & \cdots & \mathbf{A}_{1N}(p) \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{N1}(p) & \cdots & \mathbf{A}_{NN}(p) \end{bmatrix} = \bar{\mathbf{U}}^H \mathbf{R}_{\bar{\mathbf{x}}}(p) \bar{\mathbf{U}}, \quad (7)$$

其中, $\mathbf{A}_{ij}(p)$ 为 L_b 维方阵。近似联合块对角化问题就是寻找矩阵 $\bar{\mathbf{U}}$ 的最优解 $\bar{\mathbf{U}}_{\text{ideal}}$, 使得 $\bar{\mathbf{A}}(p)$ 尽可能的为块对角矩阵。这可以通过最小化代价函数

$$J(\bar{\mathbf{U}}) = \sum_{p=1}^P \text{boff}\{\bar{\mathbf{A}}(p)\} \quad (8)$$

来实现。其中, $\text{boff}\{\bar{\mathbf{A}}(p)\} = \sum_{i=1}^N \sum_{j=1, j \neq i}^N \|\mathbf{A}_{ij}(p)\|_F^2$ 。

由式 (7) 和式 (8) 不难看出, 代价函数 J 为待估计矩阵 $\bar{\mathbf{U}}$ 的四次代价函数, 因此直接最小化代价函数的求解难度和计算复杂度较高。为此, 我们提出如下的分步最小化方法来简化问题的求解。

将矩阵 $\bar{\mathbf{U}}$ 表示为含有 N 个子矩阵 $\mathbf{U}_i \in C^{NL_b \times L_b}$ 的分块矩阵, 即 $\bar{\mathbf{U}} = [\mathbf{U}_1, \dots, \mathbf{U}_N]$, 从而式 (7) 可以展开成如下的形式:

式 (9) 表明, 子矩阵 $\mathbf{U}_i$ 只改变 $\overline{\mathbf{A}}(p)$ 的第 i 行块和第 i 列块, 于是, 可通过最小化 $\overline{\mathbf{A}}(p)$ 的第 i 行块和第 i 列块上非对角线上的子矩阵的 F -范数的平方和来估计子矩阵 $\mathbf{U}_i$ 。从而将式 (8) 所示的四次代价函数转换为关于子矩阵 $\mathbf{U}_i$ 的二次代价函数。

固定子矩阵 $\mathbf{U}_j (j \neq i)$, 定义:

$$J_i(\mathbf{U}_i) = \sum_{p=1}^P \sum_{j=1, j \neq i}^N \left[\|\mathbf{A}_{ij}(p)\|_F^2 + \|\mathbf{A}_{ji}(p)\|_F^2 \right] = \sum_{p=1}^P \sum_{j=1, j \neq i}^N \left[\|\mathbf{U}_i^H \mathbf{R}_{\overline{\mathbf{x}}}(p) \mathbf{U}_j\|_F^2 + \|\mathbf{U}_j^H \mathbf{R}_{\overline{\mathbf{x}}}(p) \mathbf{U}_i\|_F^2 \right]. \quad (10)$$

由式 (8) 一式 (10) 可知,

$$J(\overline{\mathbf{U}}) = \frac{1}{2} \sum_{i=1}^N J_i(\mathbf{U}_i), \quad (11)$$

于是, $\min J(\overline{\mathbf{U}})$ 的理想解 $\overline{\mathbf{U}}_{\text{ideal}}$ 可以通过依次分步执行 $\min J_i(\mathbf{U}_i)$ 操作而得到。同时, 为了避免“平凡”解, 并考虑到矩阵 $\overline{\mathbf{U}}$ 满足条件 $\overline{\mathbf{U}}^H \overline{\mathbf{U}} = \mathbf{I}_{NL_b}$, 约束任意子矩阵 $\mathbf{U}_i$ 的各列之间满足:

$$\mathbf{u}_{ic}^H \mathbf{u}_{ic'} = \delta(c - c'), \quad (12)$$

其中, $\mathbf{u}_{ic}$ 表示 $\mathbf{U}_i$ 第 c 列, $c=1, \dots, L_b$ 都是变量。

下面分析如何求解在式 (12) 约束下的式 (10) 所示的二次优化问题。引入拉格朗日乘子向量 $\boldsymbol{\beta} = [\beta_1, \dots, \beta_{L_b}]$, 构造如下的修正的无约束二次代价函数:

$$\overline{J}_i(\mathbf{U}_i) = J_i(\mathbf{U}_i) + \text{tr} \{ \overline{\boldsymbol{\beta}} [\mathbf{U}_i^H \mathbf{U}_i - \mathbf{I}_{L_b}] \}, \quad (13)$$

其中 $\boldsymbol{\beta} = \text{diag}[\boldsymbol{\beta}]$ 。基于梯度下降的方法, 代价函数 $\overline{J}_i(\mathbf{U}_i)$ 关于 $\mathbf{U}_i$ 求导并令导数为 0, 有:

$$\mathbf{C}_i \mathbf{U}_i = [\mathbf{C}_i \mathbf{u}_{i1}, \mathbf{C}_i \mathbf{u}_{i2}, \dots, \mathbf{C}_i \mathbf{u}_{iL_b}] = [\beta_1 \mathbf{u}_{i1}, \beta_2 \mathbf{u}_{i2}, \dots, \beta_{L_b} \mathbf{u}_{iL_b}], \quad (14)$$

其中:

$$\mathbf{C}_i = \sum_{p=1}^P \left[\mathbf{R}_{\overline{\mathbf{x}}}(p) \left(\sum_{j=1, j \neq i}^N \mathbf{U}_j \mathbf{U}_j^H \right) \mathbf{R}_{\overline{\mathbf{x}}}^H(p) + \mathbf{R}_{\overline{\mathbf{x}}}^H(p) \left(\sum_{j=1, j \neq i}^N \mathbf{U}_j \mathbf{U}_j^H \right) \mathbf{R}_{\overline{\mathbf{x}}}(p) \right],$$

由式 (14) 易知, $\mathbf{U}_i$ 可由 $\mathbf{C}_i$ 的 L_b 个特征向量构成。进一步展开式 (14), 可知:

$$\begin{aligned} \mathbf{C}_i \mathbf{U}_i &= \sum_{p=1}^P \left[\mathbf{R}_{\overline{\mathbf{x}}}(p) \left(\sum_{j=1, j \neq i}^N \mathbf{U}_j \mathbf{U}_j^H \right) \mathbf{R}_{\overline{\mathbf{x}}}^H(p) + \mathbf{R}_{\overline{\mathbf{x}}}^H(p) \left(\sum_{j=1, j \neq i}^N \mathbf{U}_j \mathbf{U}_j^H \right) \mathbf{R}_{\overline{\mathbf{x}}}(p) \right] \mathbf{U}_i = \\ &= \sum_{p=1}^P \sum_{j=1, j \neq i}^N [\mathbf{R}_{\overline{\mathbf{x}}}(p) \mathbf{U}_j \mathbf{U}_j^H \mathbf{R}_{\overline{\mathbf{x}}}^H(p) + \mathbf{R}_{\overline{\mathbf{x}}}^H(p) \mathbf{U}_j \mathbf{U}_j^H \mathbf{R}_{\overline{\mathbf{x}}}(p)] \mathbf{U}_i = \\ &= \sum_{p=1}^P \sum_{j=1, j \neq i}^N [\mathbf{R}_{\overline{\mathbf{x}}}(p) \mathbf{U}_j \mathbf{A}_{ij}^H(p) + \mathbf{R}_{\overline{\mathbf{x}}}^H(p) \mathbf{U}_j \mathbf{A}_{ji}(p)] \end{aligned} \quad (15)$$

理想情况下, 当 $\overline{\mathbf{A}}(p)$ 为块对角矩阵时, 对于任意 $j \neq i$, 有 $\mathbf{A}_{ij}^H(p) = \mathbf{A}_{ji}(p) = \mathbf{0}_{L_b}$ 。则由式 (15) 可知, $\mathbf{U}_i$ 应为 $\mathbf{C}_i$ 的零子空间。返回到式 (13) 和式 (14) 易知, 当 $\mathbf{U}_i$ 由 $\mathbf{C}_i$ 的 L_b 个零特征值所对应的特征向量构成时, 代价函数 $\overline{J}_i(\mathbf{U}_i) = 0$ 达到最小点。一般地, 由于不可避免地受到噪声、估计误差的影响, 当 $\mathbf{U}_i$ 由 $\mathbf{C}_i$ 的 L_b 个最小特征值所对应的特征向量构成时, 代价函数 $\overline{J}_i(\mathbf{U}_i)$ 达到最小点。

2.3 分步迭代方法实现

基于上述分析, 提出如下的分步迭代盲分离实现方法:

(1) 将 M 个传声器接收到的混迭信号重新排列成矢量 $\mathbf{z}(n)$; 并计算相关矩阵 $\mathbf{R}_{\mathbf{z}}(0)$ 。

(2) 利用式 (3) 得到白化矩阵 $\mathbf{W}$, 计算白化后的

观测数据矢量 $\mathbf{x}(n) = \mathbf{W} \mathbf{z}(n)$ 。

(3) 计算 $\mathbf{x}(n)$ 在 P 个不同时延下的二阶相关矩阵, 记为 $\mathbf{R}_{\mathbf{x}}(1), \dots, \mathbf{R}_{\mathbf{x}}(P)$ 。

(4) 选择任意 NL_b 维酉矩阵作为矩阵 $\overline{\mathbf{U}}$ 的初始值, 记为 $\overline{\mathbf{U}}(0)$, 并令 $k=1$ 。

(5) 对于每个 $i=1, \dots, N$, 计算矩阵 $\mathbf{C}_i(k)$ 的 L_b 个最小特征值所对应的特征向量 $\mathbf{v}_1(\mathbf{C}_i(k)), \dots, \mathbf{v}_{L_b}(\mathbf{C}_i(k))$, 并令 $\mathbf{U}_i(k) = [\mathbf{v}_1(\mathbf{C}_i(k)), \dots, \mathbf{v}_{L_b}(\mathbf{C}_i(k))]$ 。

(6) 如果 $\|\overline{\mathbf{U}}(k) - \overline{\mathbf{U}}(k-1)\|_F^2 > \varepsilon$, ($0 < \varepsilon \leq 1$), $k=k+1$, 返回 (5)。否则, 方法收敛, 执行 (7)。

(7) $\overline{\mathbf{U}}_{\text{ideal}} \leftarrow \overline{\mathbf{U}}(k)$, 得到酉矩阵 $\overline{\mathbf{U}}$ 的理想估计。

(8) $\mathbf{W}_s = \overline{\mathbf{U}}_{\text{ideal}}^H \mathbf{W}$, 得到分离的信号 $\mathbf{y}(n) = \mathbf{W}_s \mathbf{z}(n)$; $K=k$, 表示算法收敛所需迭代次数; 方法完毕。

表 1 子矩阵估计的计算量分析

for $i = 1, \dots, N$	NMD
for $p = 1, \dots, P$	
$M_1 = \mathbf{R}_{\mathbf{x}}(p)\mathbf{U}_i(k-1); \quad M_2 = \mathbf{R}_{\mathbf{x}}^H(p)\mathbf{U}_i(k-1);$ $C_i(k) = \overline{\mathbf{C}} - M_1 M_1^H - M_2 M_2^H;$	$4PN^2 L_b^3$
end	
$\mathbf{U}_i(k) \leftarrow C_i(k)$ 的 L_b 个最小特征值所对应的特征向量。	$3N^3 L_b^3$
for $p = 1, \dots, P$	
$M_1 = \mathbf{R}_{\mathbf{x}}(p)\mathbf{U}_i(k); \quad M_2 = \mathbf{R}_{\mathbf{x}}^H(p)\mathbf{U}_i(k); \quad \overline{\mathbf{C}} = C_i(k) + M_1 M_1^H + M_2 M_2^H; \quad 4PN^2 L_b^3$	
end	
end	

注: $k=0$ 时, 初始化 $\overline{\mathbf{C}} = \sum_{p=1}^P \left[\mathbf{R}_{\mathbf{x}}(p) \left(\sum_{j=1}^N \mathbf{U}_j \mathbf{U}_j^H \right) \mathbf{R}_{\mathbf{x}}^H(p) + \mathbf{R}_{\mathbf{x}}^H(p) \left(\sum_{j=1}^N \mathbf{U}_j \mathbf{U}_j^H \right) \mathbf{R}_{\mathbf{x}}(p) \right]$.

上文已提到, 理想情况下, 当 $\overline{\mathbf{A}}_{\text{ideal}}(p)$ 为块对角矩阵时, 分离出的信号可以表示为 $\overline{\mathbf{y}}(n) = \overline{\mathbf{E}} \overline{\mathbf{D}} \overline{\mathbf{s}}(n)$ 。则 $\overline{\mathbf{y}}(n)$ 的第 j 个子块可表示为:

$$\mathbf{y}_j(n) = \sum_{l=1}^{L_b} \mathbf{d}_{j'}(l) s_{j'}(n-l+1) \quad (16)$$

其中, $\mathbf{d}_{j'}(l)$ 表示 $\overline{\mathbf{D}}$ 对角线上的第 j' 个子矩阵 $\mathbf{d}_{j'} \in \mathbb{C}^{L_b \times L_b}$ 的第 1 行。

式 (16) 表示, $\mathbf{y}_j(n)$ 可视为源信号 $s_{j'}(n)$ 经过一个长度为 L_b 的冲击响应为 $\mathbf{d}_j(l)$ 的单输入多输出 (SIMO) 系统的输出, 即 $\mathbf{y}_j(n)$ 仅为源信号 $s_{j'}(n)$ 的滤波版本。由于块置换矩阵 $\overline{\mathbf{E}}$ 的影响, 不一定满足 $j = j'$, 即分离的块独立数据 $\overline{\mathbf{y}}(n)$ 与实际的块独立数据 $\overline{\mathbf{s}}(n)$ 可能会存在块间排列不定性。然而, 这并不影响分离效果, 因为 j 与 j' 是一一对应的, 只是改变分离信号的输出次序。

事实上, 在 NL_b 个分离的信号中, 只需要其中的 N 个信号作为真实信号的估计。由于噪声和估计误差的影响, 工程应用中很难实现精确地联合块对角化, 也就是说分离的信号只能近似满足相互独立性。因此, 可以选择其中 N 个相关系数最小的信号作为期望得到的分离信号。此外, 在高精度需求情况下, 可采用 SIMO 均衡技术修正恢复的信号 $\mathbf{y}_j(n)$, 使得系统输出与真实的源信号之间只存在相位差和幅度模糊。

2.4 方法复杂度分析

由 2.3 节中的方法实现步骤不难看出, 该方法的运算量主要取决于步骤 (4) — (7) 中求解酉矩阵 $\overline{\mathbf{U}}$ 的迭代计算。于是, 可以依据每步迭代所需的乘除运算个数, 即 NMD (Number of Multiplications and

Divisions) 来衡量方法的计算复杂度^[16]。

表 1 给出了基于式 (13) — 式 (15) 求解 $\mathbf{U}_i$ 的单次迭代伪代码, 并分析了代码实现的计算量。由表 1 可知, 求解子矩阵 $\mathbf{U}_i$ 所需的计算量为 $8PN^2 L_b^3 + 3N^3 L_b^3$, 则估计酉矩阵 $\overline{\mathbf{U}}$ 所需的总的计算量为 $N(8PN^2 L_b^3 + 3N^3 L_b^3)$ 。为了加强方法的稳健性, 常取目标矩阵的个数 P 大于源的个数 N ^[3,6]。因此, 本文提出方法的计算复杂度量级为 $O(PN^3 L_b^3)$, 这与类 Jacobi 方法实现联合块对角化的单次迭代所需的 NMD 数处于同一量级^[15]。然而, 类 Jacobi 算法的单步迭代中还需要求解 $NL(NL-1)/2$ 个雅克比旋转矩阵, 且在求解每个雅克比旋转矩阵时需求解一个一元四次 (当目标矩阵为实数时) 或一元 6 次 (当目标矩阵为复数时) 方程的根。因此本文方法的计算复杂度远小于类雅克比算法的计算复杂度。

3 仿真实验

采用两组实验通过与类 Jacobi 方法比较来验证本文方法的高效性和可行性。实验 1, 用于验证联合块对角化方法的收敛速度和稳态性能。实验 2, 与真实语音环境相结合, 验证所提方法在盲卷积分离中的实用效果。其中类 Jacobi 方法^[15]的仿真代码来自于网站 <http://mips.gsf.de/proj/cmb/researchjbd.html>。

实验 1: 评估全局拒噪水平 GRL 和方法收敛所需迭代次数两个性能指标。

构造一组具有近似块对角结构的目标矩阵 $\mathbf{R}_{\mathbf{x}}(p) = \overline{\mathbf{U}} \overline{\mathbf{A}}(p) \overline{\mathbf{U}}^H + \Delta \mathbf{R}_{\mathbf{x}}(p) (p = 1, \dots, P)$, 其中, 酉矩阵 $\overline{\mathbf{U}}$ 、块对角矩阵 $\overline{\mathbf{A}}(p)$ 以及误差矩阵 $\Delta \mathbf{R}_{\mathbf{x}}(p)$

中的每个元素随机产生 (相互独立) 且服从 $(0, 1)$ 正态分布。

定义无误差项与误差项的比 $\text{NER}(\text{Non-error to Error Ratio})$ 为:

$$\text{NER} = 10 \log_{10} \left[\frac{\|\overline{\mathbf{U}} \mathbf{A}(p) \overline{\mathbf{U}}^H\|_F^2}{\|\Delta \mathbf{R}_{\overline{\mathbf{x}}}(p)\|_F^2} \right], \quad (17)$$

定义全局拒噪水平:

$$\text{GRL} = 10 \log_{10} \left\{ \sum_{p=1}^P \left(\sum_{i=1}^N \sum_{j=1, j \neq i}^N \left\| \mathbf{A}_{ij}^{\text{ideal}}(p) \right\|_F^2 / \sum_{i=1}^N \left\| \mathbf{A}_{ii}^{\text{ideal}}(p) \right\|_F^2 \right) \right\}, \quad (18)$$

其中, $\mathbf{A}_{ij}^{\text{ideal}}(p)$ 表示近似块对角矩阵 $\overline{\mathbf{A}}_{\text{ideal}}(p) = \overline{\mathbf{U}}_{\text{ideal}}^H \mathbf{R}_{\overline{\mathbf{x}}}(p) \overline{\mathbf{U}}_{\text{ideal}}$ 的第 (i, j) 个子矩阵。由式 (18) 易知, 当 $\overline{\mathbf{A}}_{\text{ideal}}(p)$ 为块对角矩阵式, 则对于任意 $i \neq j$, $i, j \in \{1, 2, \dots, N\}$ 有 $\mathbf{A}_{ij}^{\text{ideal}}(p) = 0$, 那么 GRL 趋向于负无穷大, 估计的 $\overline{\mathbf{U}}_{\text{ideal}}$ 与真实的 $\overline{\mathbf{U}}$ 满足 $\overline{\mathbf{U}}_{\text{ideal}} = \overline{\mathbf{U}} \overline{\mathbf{D}} \overline{\mathbf{E}}$ 的关系。因此, 我们期望仅利用较少的迭代步骤就可以得到较小的 GRL 值。

图 1 所示为当信源个数 N 、目标矩阵个数 P 、以及子矩阵维数 L_b 分别为 3, 17 和 11 时, 经过 100 次独立实验, 本文提出方法 (虚线) 与类 Jacobi 联合块对角化方法 (实线) 在不同的 NER (4 dB, 8 dB, 12 dB, 16 dB) 下的平均收敛曲线 (横坐标 k 表示迭代次数)。由图 1 可知, 在所有 NER 下, 本文所提出的方法与类 Jacobi 联合块对角化方法具有相同的收敛稳态性能, 但收敛所需迭代次数远小于 Jacobi 方法。由 2.4 节已知, 本文算法单步迭代计算复杂度远小于类 Jacobi 算法, 因此图 1 表明本文提出的方法具有更快的收敛速度。

采用与图 1 相同的实验参数, 图 2 所示为, 在不同的 NER 下, 收敛所需迭代次数 K 随子矩阵维数 L_b 变化的曲线图 (当 $|\text{GRL}(k) - \text{GRL}(k-1)| < 10^{-10}$ 时, 认为方法收敛)。由图 2 易知, 随着子矩阵维数的增大, 类 Jacobi 方法收敛所需的迭代次数一直在增加, 而本文所提出的方法对子矩阵维数不敏感。也就是说, 虽然子矩阵维数随信道冲激响应阶数的增加而增大, 但本文所提方法的收敛性却不受信道阶数的限制。

此外, 需要指出的是在本实验中, 本文算法的初始值为任意酉矩阵, 而类 Jacobi 算法的初始值由 JADE 所确定 (由文献 15 知, 类 Jacobi 算法的收敛性能严重依赖于初始值的选取) 由此可知, 本文算法对初始矩阵不敏感。

实验 2: $N = 4$ 段语音信号 (如图 3(a) 所示, 其中横坐标 n 表示采样点) 通过长度为 $L = 4$ 的 FIR

滤波器矩阵, 并由 $M = 6$ 个传声器接收, 得到如图 3(b) 所示的观测信号。其中, 信道系数 $h_{ij}(l)$ ($i = 1, \dots, M; j = 1, \dots, N; l = 0, \dots, L-1$) 的相互独立且均服从 $(0, 1)$ 正态分布。为使混迭矩阵 $\overline{\mathbf{H}}$ 满足假定 $\text{A1}(ML' \geq N(L' + L - 1))$, 选择数据窗长为 $L' = 8$, 此时子矩阵维数为 $L_b = L' + L - 1 = 11$ 。图 3(c) 和 3(d) 所示为两种方法得到的分离信号波形, 可以看出二者分离效果较为接近。再引入如下两个客观评价参数来对比二者的分离性能: 算法分离性能参数 $\text{PI}^{[17]}$ 和各个原始语音与其对应的分离结果之间的巴克谱失真测度 $\text{BSD}^{[18]}$ (Bark Spectral Distortion)。在计算每个分离信号与源信号的 BSD 时, 信号采样频

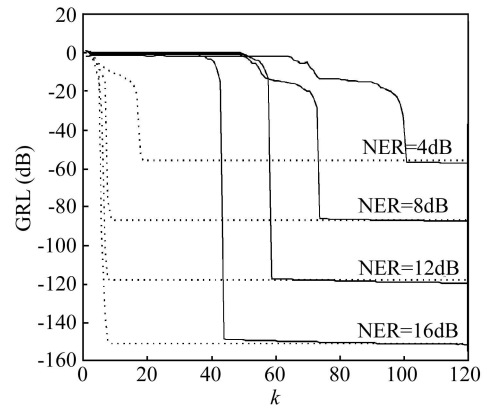


图 1 收敛曲线比较

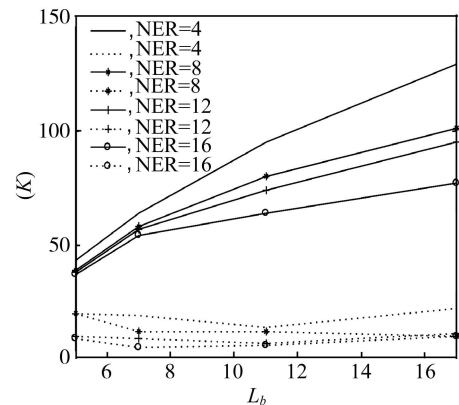


图 2 迭代次数随子矩阵维数变化曲线

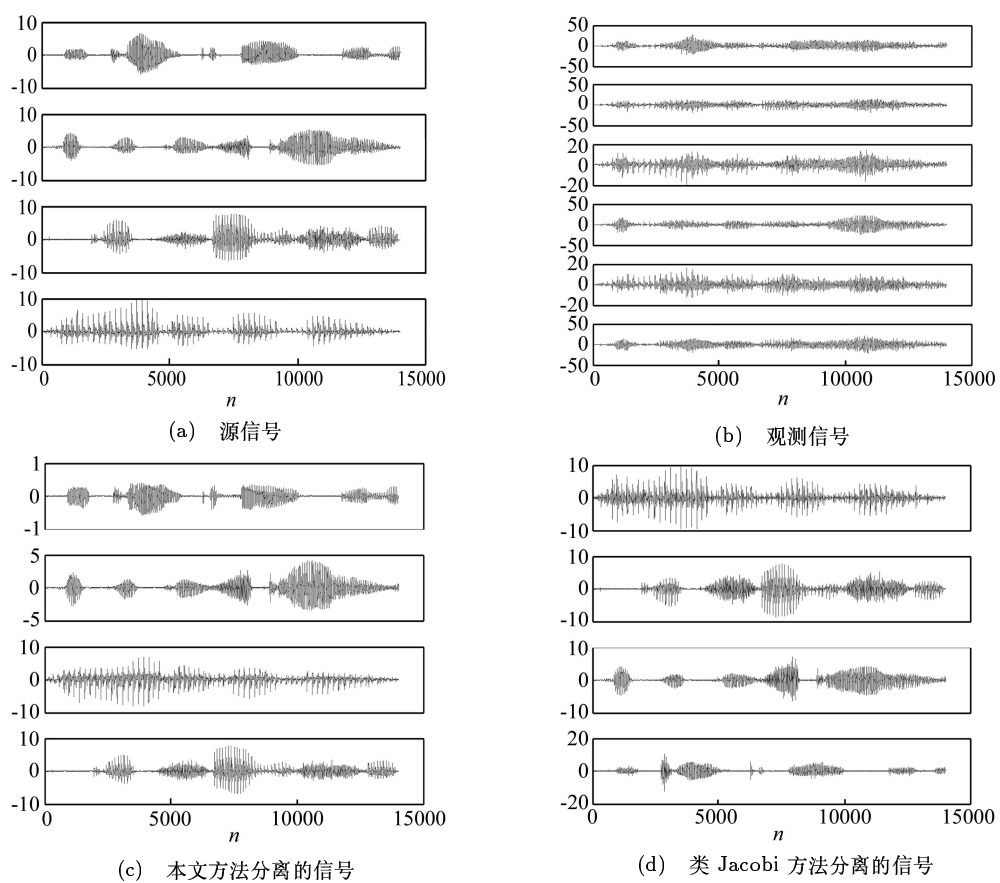


图 3 信号波形比较

表 2 分离性能测度

	本文方法				类雅克比方法			
PI	0.0016				0.0023			
BSD	0.3442	0.3053	0.0501	0.3735	0.4916	0.0195	0.3999	0.1671
BSD 均值	0.2683				0.2695			

率为 16 kHz, 共分 22 个临界带, 并采用 512 点的 FFT。表 2 所示为两种方法的分离性能测度参数(数值越小说明分离性能越好), 不难看出, 二者的 PI 值以及平均 BSD 值均较接近, 这本质上是因为者具有相同的稳态收敛性能(如实验 1 所示)。

4 结论

提出一种二次联合块对角化方法, 利用源信号间相互独立且统计短时平稳的性质, 解决超定卷积盲分离问题。该方法每步只需求解一个二次最优化问题, 迭代估计西混迭矩阵的一个子矩阵, 大大降低了求解复杂度。理论分析和实验结果表明, 该方法用于语音信号盲分离, 具有收敛速度快, 对传输信道阶数和初始值不敏感的特点, 且能够提供不亚于经典 Jacobi 方法的分离效果。此外, 该方法与 SIMO 均衡技术相结合可以实现接收信号的解卷积, 因而还可

以作为盲解卷积和语音增强的预处理方法, 扩展其应用范围。

参 考 文 献

- 1 Haykin S. Unsupervised Adaptive Filtering, vol I: Blind source separation. New York, 2000: 64
- 2 Cardoso J F, Souloumiac A. Blind beamforming for non-Gaussian signals. IEE Proc.-F, 1993; 140(6): 362—370
- 3 Belouchrani A, Abed-Meraim K, Cardoso J F, Moulines E. A blind source separation technique using second-order statistics. IEEE Trans. Signal Process., 1997; 45(2): 434—444
- 4 Sidiropoulos N D, Bro R, Giannakis G B. Parallel factor analysis in sensor array processing. IEEE Trans. Signal Process., 2000; 48(7): 2377—2388
- 5 Van der Veen A J. Joint diagonalization via subspace fitting techniques. In: Proc. ICASSP2001, Salt Lake City, UT, 2001(4): 2773—2776

- 6 Vollgraf R, Obermayer K. Quadratic optimization for simultaneous matrix diagonalization. *IEEE Trans. Signal Process.*, 2006; **54**(9): 3270—3278
- 7 陈 健, 陆信人. 噪声背景下双输入时延混合系统的盲源分离. *声学学报*, 2002; **27**(4): 477—480
- 8 徐 舜, 陈绍荣, 刘郁林. 基于非线性时频掩蔽的语音盲分离方法. *声学学报*, 2007; **32**(4): 375—381
- 9 Parra L, Spence C. Convolutional blind separation of non-stationary source. *IEEE Trans. Speech Audio Process.*, 2000; **3**(7): 320—327
- 10 Sawada H, Mukai R, Araki S, Makino S. A robust and precise method for solving the permutation problem of frequency-domain blind source separation. *IEEE Trans. Speech Audio Process.*, 2004; **12**(4): 533—538
- 11 Castella M, Rhioui S, Moreau E, Pesquet J-C. Quadratic higher order criteria for iterative blind separation of a MIMO convolutional mixture of sources. *IEEE Trans. Signal Process.*, 2007; **55**(1): 218—232
- 12 Bousbia-Salah H, Belouchrani A, Abed-Meraim K. Jacobi-like algorithm for blind signal separation of convolutional mixture. *Electron. Lett.*, 2001; **37**(16): 1049—1050
- 13 Lathauwer L D, Févotte C, Moor B D, Vandewalle J. Jacobi algorithm for joint block diagonalization in blind identification. In: Proc. 23rd Symp. Inf. Theory, Benelux, Louvain-la-Neuve, Belgium, 2002: 155—162
- 14 Theis F J. Blind signal separation into group of dependent signals using joint block diagonalization. In: Proc. ISCAS2005, Kobe, Japan, 2005: 5878—5881
- 15 Févotte C, Theis F J. Orthonormal approximate joint block-diagonalization. Technical Report GET/Télécom Pairs 2007D007, 2007
- 16 Golub G H, Van Loan C F. Matrix computations, 2nd ed. Baltimore and London, The John Hopkins University Press, 1989: 1—46
- 17 Amari S, Cichocki A, et al. A new learning algorithm for blind signal separation. *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 1996: 757—763
- 18 Wang S, Sekey A, Gersho A. An objective measure for predicting subjective quality of speech coders. *IEEE J. Sel. Areas Commun.*, 1992; **10**(4): 819—829