

## 基于 ICA 的雷达信号欠定盲分离算法

陈晓军 成昊 唐斌

(电子科技大学电子工程学院 成都 610054)

**摘要:** 该文针对源信号时域和频域不充分稀疏的情况,提出了欠定盲源分离中估计混合矩阵的一种新方法。该方法对等间隔分段的观测信号应用独立分量分析(ICA)的盲分离算法获得多个子混合矩阵,然后对其分选剔除了不属于原混合矩阵的元素,最后利用 C 均值聚类的学习算法获得对混合矩阵的精确估计,解决了源信号在时域和频域不充分稀疏的情况下准确估计混合矩阵的问题。在估计出混合矩阵的基础上,利用基于稀疏分解的统计量算法分离出源信号。由仿真结果,以及与传统的 K 均值聚类,时域检索平均算法对比的实验结果说明了该文算法的有效性和鲁棒性。

**关键词:** 信号处理; 欠定盲源分离; 独立分量分析; C 均值聚类; 稀疏分解的统计量

**中图分类号:** TN957.51

**文献标识码:** A

**文章编号:** 1009-5896(2010)04-0919-06

**DOI:** 10.3724/SP.J.1146.2009.00291

## Underdetermined Blind Radar Signal Separation Based on ICA

Chen Xiao-jun Cheng Hao Tang Bin

(College of Electronic Engineering, University of Electronic Science Technology of China, Chengdu 610054, China)

**Abstract:** A method of the mixing matrix estimation in the underdetermined source separation is proposed in which the sources are not sparse enough to estimate the mixing matrix. Getting many sub matrixes through applying Independent component analysis(ICA) for observation signals and removing the elements do not belong to the mixing matrix, the mixing matrix is estimated precisely with C-means clustering agglomeration. Then, the source signals can be recovered with the statistically sparse decomposition principle. The experiment shows that the method have better accuracy and validity than K-means and searching-and-averaging method in the time domain in estimating the mixing matrix.

**Key words:** Signal processing; Underdetermined blind signal source separation; Independent Component Analysis(ICA); C-means clustering; Statistically sparse decomposition principle

### 1 引言

近年来,盲源分离问题已成为信号处理和神经网络研究的热点领域,其特点是在混叠通道和源信号分布未知的情况下仅从观测信号中恢复源信号的波形,而且已经被广泛应用通信,语音,雷达,生物医学等领域<sup>[1-3]</sup>,但是很多盲分离算法都假设观测信号的数目不少于源信号的数目。然而在实际应用中,由于观测条件有限,有可能会发生源信号数目多于观测信号数目的欠定情况。

目前,国内外针对欠定条件下盲分离算法的研究主要是集中在稀疏性很强的信号,例如 Boffill<sup>[3]</sup>提出了利用聚类方法估计混合矩阵和最短路径法估计源信号的两阶段盲分离算法,Donoho<sup>[4]</sup>提出了用  $l^1$  范数解估计源信号, Li<sup>[5]</sup>提出了用 K 均值聚类法

估计混合矩阵, Févotte<sup>[6]</sup>利用稀疏信号的概率分布特性在贝叶斯框架下估计混合矩阵和源信号,肖明等<sup>[7]</sup>提出了利用时域检索平均法估计混合矩阵, Araki<sup>[8]</sup>进一步深化了频域欠定盲分离算法,谭北海<sup>[9]</sup>等在信源数目未知的情况下提出了基于源信号数目估计的欠定盲分离算法。这些算法在估计混合矩阵和恢复源信号时,都要求源信号满足理想的稀疏性条件。当源信号的时域上的稀疏性并不是很好时,可以通过短时傅里叶变换<sup>[3]</sup>,小波包变换<sup>[10]</sup>,时频变换<sup>[11]</sup>等方法将源信号变换到变换域上进行盲分离。但是很多信号在变换域上并不严格满足稀疏特性,此时通过上述算法并不能精确地估计出混合矩阵。因此,研究欠定盲源分离中源信号并不充分稀疏的条件下估计混合矩阵对盲分离技术的发展有重要的意义,也是目前必须予以解决的问题。

本文针对源信号在时域和频域上并不充分稀疏的情况提出了一种新的估计混合矩阵的方法。利用

独立分量分析的盲源分离算法对分段后的观测矩阵处理得到多个子混合矩阵, 并对其进行分选; 然后对分选后的子混合矩阵应用C均值聚类的学习算法获得混合矩阵的精确估计; 最后利用基于稀疏分解的统计量算法对源信号进行估计。本文算法在一定程度上克服了欠定盲分离算法对源信号稀疏性的要求, 解决了源信号在时域和频域并不充分稀疏的情况下准确估计混合矩阵的问题, 对欠定盲分离的发展有重要的意义。

## 2 算法介绍

### 2.1 盲源分离模型

为了清楚地阐述信号盲分离算法, 本文考虑一般的线性瞬时混叠模型

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

其中  $\mathbf{s}$  为  $n$  个未知的源信号,  $\mathbf{x}$  为  $m$  个已知的观测信号,  $\mathbf{e}$  为噪声,  $\mathbf{A}$  为  $m \times n$  维的混合矩阵。为了描述方便, 将混合矩阵  $\mathbf{A}$  写作  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ 。在欠定盲分离系统中, 观测信号的数目小于源信号的数目, 即  $m < n$ 。在噪声不存在或者可以忽略不计的情况下, 这时盲源分离的模型如下:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (2)$$

当源信号是稀疏信号时, 它们的混合信号具有线性聚类特性, 其在  $m$  维空间中是  $n$  条直线, 直线的斜率取决于混合矩阵的列矢量。传统的欠定盲分离算法是利用源信号时域或频域上的稀疏性通过聚类算法估计混合矩阵  $\mathbf{A}$ , 然后再去估计源信号。当源信号稀疏性不好时, 其在  $m$  维空间上不能呈现出  $n$  条清晰的直线, 因此不能用传统方法估计混合矩阵。

### 2.2 混合矩阵的估计

首先, 假设源信号  $\mathbf{s}$  在给定的时间间隔  $\Delta t$  内最多有  $m$  个源信号占主导地位。为了分析计算方便, 假设只有两个观测信号 ( $m = 2$ )。在上述假设条件下, 在  $\Delta t$  内, 欠定盲分离问题可转化成下列正定和超定盲分离问题求解

$$\left. \begin{aligned} \mathbf{x}^t &= \mathbf{A}_{j_1 j_2} [\mathbf{s}_{j_1}^t, \mathbf{s}_{j_2}^t]^T \\ \mathbf{s}_j^t &\approx 0 (j \neq j_1, j_2) \end{aligned} \right\} \quad (3)$$

式中  $\mathbf{A}_{j_1 j_2} = (\mathbf{a}_{j_1}, \mathbf{a}_{j_2})$ ,  $j_1, j_2 = 1, \dots, n$ , 如果对  $\mathbf{x}^t$  应用传统的盲源分离算法, 即可估计出混合矩阵其中的两个列向量。

为了估计出所有的列向量, 将观测矩阵  $\mathbf{x}$  按时间间隔  $\Delta t$  划分为  $N$  个子矩阵  $\mathbf{x} = [\mathbf{x}_1^t, \dots, \mathbf{x}_N^t]$ , 然后分别对  $\mathbf{x}_i^t, i = 1, \dots, N$  应用独立分量分析(ICA)算法。ICA 算法通常是假设观测信号数目大于源信号

数目, 即  $m > n$ , 然后求取一个解混矩阵  $\mathbf{B}$ , 使得观测信号  $\mathbf{x}$  通过它后所得的输出  $\mathbf{y}$  ( $\mathbf{y} = \mathbf{B}\mathbf{x}$ ) 是源信号  $\mathbf{s}$  的最优逼近。自适应的 ICA 算法是寻找一个合适的目标函数  $\varepsilon$  (通常是系统参数  $\mathbf{B}$  的函数), 通过逐步调节  $\mathbf{B}$  使  $\varepsilon$  达到最小(或最大), 实现对信号的分选。本文选取累积量迭代反演算法(Cumulant-Based Iterative Inversion, CII), 它不需要对源信号的统计分布进行假设, 而且收敛速度快, 具有很好的鲁棒性。

CII 算法是以  $\mathbf{y}$  的各个分量的互信息为目标函数  $\varepsilon$ , 同时用各个输出分量归一化的  $1 + \beta$  阶累积量来代替其边熵, 并以梯度  $\partial \varepsilon / (\partial \mathbf{B})$  调节  $\mathbf{B}$  使  $\varepsilon$  最小来达到信号的分选。其估计混合矩阵的代价函数为

$$\mathbf{B}^{(n+1)} = \mathbf{B}^{(n)} - \mu^{(n)} (\mathbf{C}_{y,y}^{1,\beta} \mathbf{S}_y^\beta - \mathbf{I}) \mathbf{B}^{(n)} \quad (4)$$

式中  $\mathbf{C}_{y,y}^{1,\beta} = \text{cum}(\mathbf{y}, \underbrace{\mathbf{y}, \mathbf{y}}_\beta, \mathbf{y})$  为互累积量矩阵,  $\mu^{(n)}$  为

自适应迭代因子,  $\mathbf{S}_y^\beta = \text{diag}(\text{sign}(\text{diag}(\mathbf{C}_{y,y}^{1,\beta})))$ 。关于如何选取  $\mu^{(n)}$ , 详见文献[12], 这里就不再赘述。当  $\beta = 3$  时,

$$\begin{aligned} \mathbf{C}_{y,y}^{1,3} &= E[\mathbf{y}(\mathbf{y} \odot \mathbf{y}^+ \odot \mathbf{y}^+)^T] - E[\mathbf{y}\mathbf{y}^T] \text{diag}(E[\mathbf{y}^+ \odot \mathbf{y}^+]) \\ &\quad - 2E[\mathbf{y}\mathbf{y}^H] \text{diag}(E[\mathbf{y} \odot \mathbf{y}^+]) \end{aligned} \quad (5)$$

将式(5)代入式(4)中, 同时对每个子观测矩阵处理后得到多个列向量归一化的子混合矩阵  $\mathbf{B} = [\mathbf{B}_1, \dots, \mathbf{B}_N]$ 。

在实际中, 在  $\Delta t$  内, 可能有多于  $m$  个源信号占主导地位, 此时, 欠定盲分离问题就不能转化成正定和超定盲分离问题求解, 对  $\Delta t$  内的观测信号应用 CII 算法迭代出的子混合矩阵并不是  $\mathbf{A}$  中的两个列向量, 因此有必要在  $\mathbf{B}$  中将这元素剔除。

为了分析方便, 定义比率向量

$$\boldsymbol{\theta} = \left[ \text{atan} \left( \frac{B_{21}}{B_{11}} \right), \dots, \text{atan} \left( \frac{B_{2K}}{B_{1K}} \right) \right] \quad (6)$$

$$\boldsymbol{\pi} = \left[ \text{atan} \left( \frac{A_{21}}{A_{11}} \right), \dots, \text{atan} \left( \frac{A_{2n}}{A_{1n}} \right) \right] \quad (7)$$

式中  $\text{atan}(\cdot)$  表示反正切函数,  $K = m \times N$  为列矢量的维数,  $B_{ij}, i = 1, 2, j = 1, \dots, K$  为混合矩阵  $\mathbf{B}$  的元素。假设  $(B_{2k}) / (B_{1k}) \in [r, R], k = 1, \dots, K$ ,  $r$  为  $\boldsymbol{\theta}$  的最小值,  $R$  为  $\boldsymbol{\theta}$  的最大值。然后把  $\boldsymbol{\theta}$  分为  $M$  ( $M$  为正整数)子区间。此时,  $\boldsymbol{\pi}_i, i = 1, \dots, n$  中所有元素都属于  $\boldsymbol{\theta}$  的某些子区间。如果某个子区间元素的个数小于  $J$  ( $J$  为正整数, 需要根据实际情况预先设定), 将子区间内所有元素删除, 这时可以近似地将  $\boldsymbol{\theta}$  中不属于  $\boldsymbol{\pi}$  的元素剔除了。记  $\boldsymbol{\theta}$  中剩余元素构成向量  $\mathbf{T}$ 。

为了准确地估计出混合矩阵, 对  $\mathbf{T}$  再使用 C 均

值聚类学习算法将其元素聚为  $n$  类。设  $\mathbf{D} = [\beta_1, \dots, \beta_n]$ , 其中  $\beta_i$  为第  $i$  个聚类中心参数。C 均值算法是最小化目标函数

$$J(\mathbf{D}, \mathbf{U}; \mathbf{T}) = \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^2 d_{ij}^2 \quad (8)$$

使得

$$\sum_{i=1}^n u_{ij} = 1, \quad j = 1, \dots, c \quad (9)$$

式中  $d_{ij}^2$  表示  $T_j$  到  $\beta_i$  距离的平方,  $\mathbf{U} = [u_{ij}]$  中的每个元素表示  $T_j$  与  $\beta_i$  的关联程度,  $c$  为  $\mathbf{T}$  的元素个数。此时可以得到估计出的混合矩阵为

$$\hat{\mathbf{A}} = [\cos(\mathbf{D}); \sin(\mathbf{D})] \quad (10)$$

### 2.3 源信号的估计

对于欠定盲源分离, 估计出混合矩阵后, 由于混合矩阵是不可逆的, 因此并不能直接估计出源信号。此时可以利用源信号在变换域上的稀疏特性, 采用基于稀疏分解的统计量 (Statistically Sparse Decomposition Principle, SSDP) 算法<sup>[13]</sup>对源信号进行估计。SSDP 算法是在固定的时间间隔  $\Delta t$  内最小化相关系数来估计源信号的, 同时是对时间间隔  $\Delta t$  寻求最短路径估计源信号, 因此误差小, 复杂度降低。

由于源信号  $\mathbf{s}$  在给定的时间间隔  $\Delta t$  内最多有  $m$  个信号占主导地位, 因此根据式(3)只需要在已估计的混合矩阵的列向量中选取两列就可以近似地估计出源信号。

首先, 定义观测信号和源信号在  $\Delta t$  内的协方差矩阵分别为

$$\mathbf{C}_x = E((\mathbf{x}^t - \mathbf{E}_x)(\mathbf{x}^t - \mathbf{E}_x)^T) \quad (11)$$

$$\mathbf{C}_s = E((\mathbf{s}^t - \mathbf{E}_s)(\mathbf{s}^t - \mathbf{E}_s)^T) \quad (12)$$

式中  $\mathbf{E}_x$  和  $\mathbf{E}_s$  分别为观测信号和源信号在  $\Delta t$  内的均值。

当  $\Delta t$  足够长时,  $\mathbf{C}_x$  可以表示为

$$\mathbf{C}_x = \frac{1}{\Delta t} \sum_{t=t_0}^{t_1} \left( \mathbf{x}^t - \frac{1}{\Delta t} \sum_{t=t_0}^{t_1} \mathbf{x}^t \right) \left( \mathbf{x}^t - \frac{1}{\Delta t} \sum_{t=t_0}^{t_1} \mathbf{x}^t \right)^T \quad (13)$$

式中  $\Delta t = t_1 - t_0 + 1$ 。当源信号  $\mathbf{s}$  在给定的时间间隔  $\Delta t$  内最多有  $m$  个源信号占主导地位时,  $\mathbf{C}_s$  可近似地表示为

$$\mathbf{C}_s \approx \begin{bmatrix} \mathbf{C}_s^{\text{sub}} & \mathbf{0}_{2 \times (n-2)} \\ \mathbf{0}_{(n-2) \times 2} & \mathbf{0}_{(n-2) \times (n-2)} \end{bmatrix} \quad (14)$$

式中  $\mathbf{C}_s^{\text{sub}} = \mathbf{A}_{j_1 j_2} \mathbf{C}_x (\mathbf{A}_{j_1 j_2}^{-1})^T$ , 此时, 可以通过式(13)和式(14)已估计的混合矩阵的两个列向量如下:

$$[j_{01}, j_{02}] = \arg \min_{\substack{j_1, j_2 \\ =1, 2, \dots, n}} \frac{|\mathbf{C}_s^{\text{sub}}(1, 2)|}{\sqrt{\mathbf{C}_s^{\text{sub}}(1, 1) \mathbf{C}_s^{\text{sub}}(2, 2)}} \quad (15)$$

最后, 在  $\Delta t$  内, 源信号可以通过下式来重构

$$\left. \begin{aligned} [\mathbf{s}_{j_{01}}^t, \mathbf{s}_{j_{02}}^t]^T &= \mathbf{A}^{-1}_{j_{01} j_{02}} [\mathbf{x}_1^t, \mathbf{x}_2^t]^T \\ \mathbf{s}_j^t &= 0, \quad j \neq j_{01}, j_{02} \end{aligned} \right\} \quad (16)$$

### 2.4 盲分离算法

本文的盲分离算法可以总结为以下步骤实现:

(1) 如果源信号不满足 2.2 节的假设条件, 对观测信号进行 FFT 变换, 记变换后的复信号的实部为  $\mathbf{X}$ , 虚部为  $\mathbf{X}'$ 。

(2) 应用 2.2 节的算法估计混合矩阵。

(3) 利用估计出的混合矩阵对  $\mathbf{X}$  和  $\mathbf{X}'$  应用基于稀疏分解的统计量算法估计出源信号 FFT 变换后的实部和虚部, 然后再进行逆 FFT 变换重构源信号。

### 3 仿真结果及其分析

**实验 1** 为了说明算法的有效性, 选取 3 个 LFM 信号, 第 1 个 LFM 信号的起始频率为 30 MHz, 带宽 450 MHz, 第 2 个 LFM 信号起始频率为 10 MHz, 带宽 220 MHz, 第 3 个 LFM 信号起始频率为 250 MHz, 带宽 150 MHz, 时宽 10  $\mu\text{s}$ , 采样率为 1 GHz, 本文选取一个时宽内的采样点作为仿真信号, 混合矩阵为

$$\mathbf{A} = \begin{bmatrix} 0.8823 & 0.9925 & 0.2713 \\ 0.4708 & 0.1224 & 0.9625 \end{bmatrix}$$

如图 1 所示, 3 个源信号在任意时间间隔  $\Delta t$  内完全重叠, 并不满足最多有 2 个源信号占主导地位。如图 2 所示, 源信号在频域上满足上述特性, 在部分频段内源信号  $s_1$  和  $s_2$  占主导地位, 在部分频段内源信号  $s_1$  和  $s_3$  占主导地位, 因此可以应用 2.4 节的算法对 3 个 LFM 信号进行盲分离。

在已知聚类的个数, 即源信号数目的情况下, 在频域中应用 K 均值聚类算法<sup>[5]</sup>估计出的混合矩阵为

$$\hat{\mathbf{A}} = \begin{bmatrix} 0.3453 & 0.9880 & 0.6539 \\ -0.9385 & -0.1543 & 0.8025 \end{bmatrix}$$

如图 2 所示, 在整个频段内几乎没有源信号  $s_3$  占主导地位, 即源信号在  $s_3$  频域上并不满足部分稀疏的

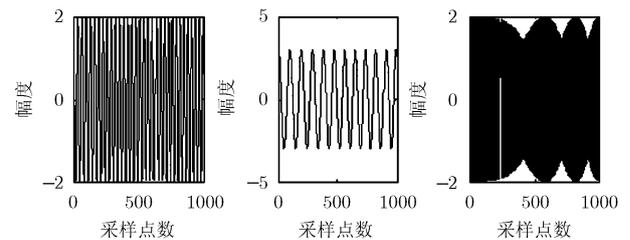


图1 源信号

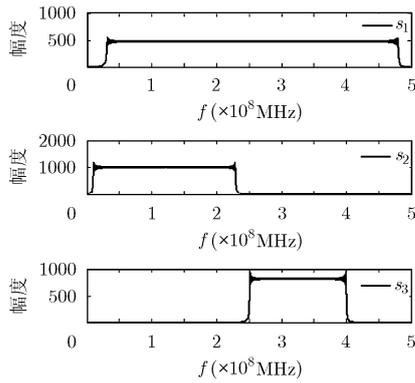


图 2 源信号频谱

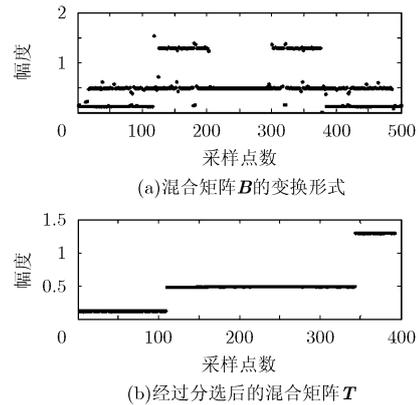


图 3 混合矩阵  $B$  的变换形式  $\theta$  和对  $\theta$  分选后得到  $T$

特性，因此利用稀疏性对观测信号直接应用  $K$  均值聚类的算法并不能准确地估计出混合矩阵。同时，由于  $K$  均值聚类法是对全部观测数据应用聚类法估计混合矩阵的，因此，即使源信号  $s_1$  和  $s_2$  在某些频段具有稀疏性，也不能估计出混合矩阵的前两列。

将文献[7]的时域检索平均法应用于频域，令检索误差  $\epsilon$  为 0.03，基矢量估计偏差  $\delta$  为 0.25，单源区间最小样本数  $\Delta$  为 3，最低样本数  $T_0$  为 20，估计出的混合矩阵为

$$\hat{A} = \begin{bmatrix} 0.8819 & 0.9921 & 0.0001 \\ 0.4699 & 0.1254 & 0 \end{bmatrix}$$

由于源信号  $s_1$  和  $s_2$  在某些频段具有单源区间，同时在算法中剔除了非单源区间，因此应用检索平均法能近似地估计出原混合矩阵的第 1 列和第 2 列，但是源信号  $s_3$  在任何频段内都不存在单源区间，因此无法估计出原混合矩阵的第 3 列，此时检索平均法也失效了。

图 3(a)表示对每个子观测矩阵应用 CII 算法得到多个估计出的混合矩阵  $B$  的变换形式  $\theta$ 。如图所示，有很多点偏离了混合矩阵列向量所在的直线，如果直接对  $\theta$  应用  $C$  均值聚类学习算法所估计出的

混合矩阵有很大的误差，因此有必要对  $\theta$  进行分选。图 3(b)表示对  $\theta$  分选后得到  $T$ ，分选后得到 3 条清晰的直线，正是混合矩阵的列向量所在的直线。

本文选取在  $\Delta t$  内有 400 个采样点，同时恰当选取  $M$  和  $J$  估计出的混合矩阵为

$$\hat{A} = \begin{bmatrix} 0.2712 & 0.9925 & 0.8825 \\ 0.9625 & 0.1224 & 0.4704 \end{bmatrix}$$

虽然估计出的混合矩阵的列向量和原混合矩阵相比有顺序上的差异性，但这并不影响源信号的估计。

在应用本文算法估计混合矩阵时，首先要注意  $\Delta t$  的选取，比较表 1 中的实验  $c$  和实验  $e$  可知，选取相同的  $M$  和  $J$ ，实验  $e$  中  $\Delta t$  内仅有 50 个采样点，估计混合矩阵的误差比实验  $c$  大很多。这是由于  $\Delta t$  太小，在 CII 算法中不能反映出源信号互累积量的特性，给分段估计子混合矩阵带来了误差。同时也要恰当选取子区间数目  $M$  和阈值  $J$ 。比较表 1 中的实验  $a$ ，实验  $b$  和实验  $c$  可知，在相同的  $\Delta t$  和  $J$  的条件下， $M$  选取越大，估计混合矩阵的误差越小。从表 1 中的实验  $c$  和实验  $d$  可知，选取相同的  $\Delta t$  和

表 1 不同参数设置对估计混合矩阵的影响

实验	$\Delta t$	$M$	$J$	$\hat{A}$	$\hat{A} - A$
$a$	400	50	20	$\begin{bmatrix} 0.8828 & 0.9922 & 0.2733 \\ 0.4698 & 0.1249 & 0.9619 \end{bmatrix}$	$\begin{bmatrix} 0.0005 & -0.0003 & 0.0020 \\ -0.0010 & 0.0025 & -0.0006 \end{bmatrix}$
$b$	400	100	20	$\begin{bmatrix} 0.8825 & 0.9925 & 0.2718 \\ 0.4704 & 0.1222 & 0.9624 \end{bmatrix}$	$\begin{bmatrix} 0.0002 & 0 & 0.0005 \\ -0.0004 & 0.0002 & 0.0001 \end{bmatrix}$
$c$	400	200	20	$\begin{bmatrix} 0.8825 & 0.9925 & 0.2712 \\ 0.4704 & 0.1224 & 0.9625 \end{bmatrix}$	$\begin{bmatrix} 0.0002 & 0 & -0.0001 \\ 0.0004 & 0 & 0 \end{bmatrix}$
$d$	400	200	5	$\begin{bmatrix} 0.8819 & 0.9922 & 0.2730 \\ 0.4714 & 0.1248 & 0.9620 \end{bmatrix}$	$\begin{bmatrix} -0.0004 & -0.0003 & 0.0017 \\ -0.0006 & 0.0024 & -0.0005 \end{bmatrix}$
$e$	50	200	20	$\begin{bmatrix} 0.8826 & 0.9924 & 0.2699 \\ 0.4701 & 0.1231 & 0.9629 \end{bmatrix}$	$\begin{bmatrix} 0.0003 & -0.0001 & -0.0014 \\ -0.0007 & 0.0007 & 0.0004 \end{bmatrix}$

$M$ , 实验  $d$  中阈值  $J$  仅为 5, 估计混合矩阵的误差比实验  $c$  大很多。这是  $J$  选取过小, 在  $B$  中不属于混合矩阵  $A$  的分量没有被剔除, 给估计混合矩阵带来了很大的误差。

最后应用基于稀疏分解的二阶统计量算法对源信号进行估计, 同样选取  $\Delta t$  内有 400 个采样点, 源信号和分离出的信号分别如图 1 和图 4 所示。

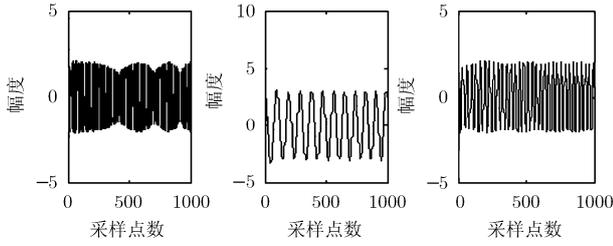


图4 分离信号

为了定量评价分离效果, 采用分离信号和源信号的相似系数  $\xi_{ij}$  作为分离的性能指标, 即

$$\xi_{ij} = \xi(y_i, s_j) = \frac{\sum_{t=1}^K y_i(t)s_j(t)}{\sqrt{\sum_{t=1}^K y_i^2(t)\sum_{t=1}^K s_j^2(t)}} \quad (17)$$

式中  $i = 1, \dots, n; j = 1, \dots, n$ 。按照式(17)计算出来的源信号和分离信号的相关系数矩阵  $\xi$  为

$$\xi = \begin{bmatrix} 0.0052 & 0.0136 & 0.9864 \\ 0.0005 & 0.9964 & 0.0014 \\ 0.9980 & 0.0036 & 0.0016 \end{bmatrix}$$

从  $\xi$  可以看出, 其每行每列都有且只有一个元素接近 1, 其它元素都接近 0, 这说明在应用本文算法估计出的混合矩阵基础上分离效果是比较理想的。

在信噪比为 20 dB 的情况下, 如表 2 所示, 本文算法估计混合矩阵的误差要比 K 均值算法要小得多。这说明本文算法在噪声条件下估计混合矩阵要比传统方法更具有优势。

**实验 2** 为了进一步验证算法的有效性, 本节选取 4 个不同的混合矩阵验证本文算法在估计混合矩

表 2 信噪比为 20 dB 时不同算法估计混合矩阵及其误差对比表

算法	估计混合矩阵	估计误差
本文算法	$\begin{bmatrix} 0.8828 & 0.9922 & 0.2733 \\ 0.4698 & 0.1249 & 0.9619 \end{bmatrix}$	$\begin{bmatrix} -0.0021 & 0.0001 & 0.0021 \\ 0.0038 & -0.0010 & -0.0006 \end{bmatrix}$
K 均值算法	$\begin{bmatrix} 0.6211 & 0.9730 & 0.3208 \\ 0.3762 & -0.1450 & -0.9419 \end{bmatrix}$	$\begin{bmatrix} -0.2612 & -0.0195 & 0.0495 \\ -0.0946 & -0.2674 & -1.9044 \end{bmatrix}$

阵时的鲁棒性, 混合矩阵具有如下的形式

$$A = \begin{bmatrix} 1 & 1 & 1 \\ \tan \omega_1 & \tan \omega_2 & \tan \omega_3 \end{bmatrix} \quad (18)$$

$\{\omega_1, \omega_2, \omega_3\}$  的设置如表 3 所见, 这些混合矩阵列向量的角度方向如图 5 所示。从图中可以看出, 5(a) 中的  $\omega_1, \omega_2, \omega_3$  的方向都非常接近; 5(b) 中的  $\omega_1, \omega_2$  的方向非常接近, 它们距离  $\omega_3$  的方向非常远; 5(c) 中的  $\omega_2, \omega_3$  的方向非常接近, 它们距离  $\omega_1$  的方向非常远; 5(d) 中的  $\omega_1, \omega_2, \omega_3$  的方向都非常远。源信号仍为实验 1 中的 3 个线性调频信号, 选取在  $\Delta t$  内有 400 个采样点, 子区间数目  $M$  为 400, 阈值  $J$  为 20。由表 3 中估计出的  $\{\omega_1, \omega_2, \omega_3\}$  可以看出, 本文算法能在不同的情况下准确地估计出原混合矩阵, 因此本文算法具有鲁棒性。

表 3 混合矩阵列向量角度的真实值及估计值

	真实值( $^\circ$ )	估计值( $^\circ$ )	
图 5(a)	$\omega_1$	55	54.9
	$\omega_2$	66	65.9
	$\omega_3$	75	75
图 5(b)	$\omega_1$	-45	-45.1
	$\omega_2$	-35	-35
	$\omega_3$	75	75.1
图 5(c)	$\omega_1$	-45	-45.1
	$\omega_2$	66	65.9
	$\omega_3$	75	75
图 5(d)	$\omega_1$	-45	-45
	$\omega_2$	10	10
	$\omega_3$	75	74.9

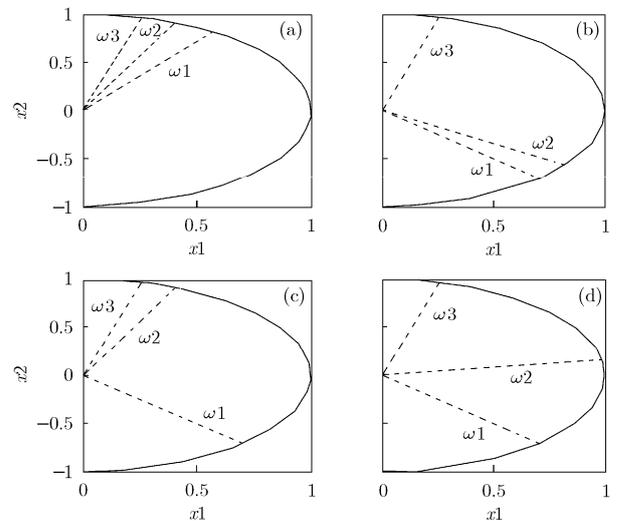


图 5 不同混合矩阵的方向图

## 4 结束语

在欠定盲源分离问题中, 准确估计混合矩阵是整个盲分离算法的关键步骤。如果混合矩阵估计不准确, 分离源信号根本无法进行。本文针对源信号在时域和频域上不充分稀疏的情况下, 提出了一种盲分离混叠矩阵的估计方法。通过精确估计混叠矩阵, 从而有效地实现盲分离。

### 参考文献

- [1] Aissa-El-Bey, Linh-Trung N, and Abed-Meraim K, *et al.* Underdetermined blind separation of nondisjoint sources in the time-frequency domain. *IEEE Transactions on Signal Processing*, 2007, 55(3): 897-907.
- [2] Delorme A, Sejnowski T, and Makeig S. Enhanced detection of artifacts in EEG data using high-order statistics and independent component analysis[J]. *Neuroimage*, 2007, 34(4): 1443-1449.
- [3] Bofill P and Zibulevsky M. Underdetermined blind source separation using sparse representations [J]. *Signal Processing*, 2001, 81(11): 2353-2362.
- [4] Donoho D L and Elad M. Optimally sparse representation in general dictionaries via l1-norm minimization[J]. *Proceedings of the National Academy of Sciences of the USA*, 2003, 100(4): 2197-2202.
- [5] Li Y Q, Cichoci A, and Amari S. Analysis of sparse representation and blind source separation[J]. *Neural Computation*, 2004, 16(6): 1193-1234.
- [6] Févotte C and Godsill S. A Bayesian approach for blind separation of sparse sources[J]. *IEEE Transactions on Audio Speech Language Process*, 2006, 16(6): 2174-2188.
- [7] 肖明, 谢胜利, 傅予力. 欠定情形下语音信号盲分离的时域检索平均法. *中国科学(E辑)*, 2007, 37(12): 1564-1575.
- Xiao M, Xie S L, and Fu Y L. Searching-and-averaging method of underdetermined blind speech signal separation in time domain[J]. *Sci China(E index)*, 2007, 37(12): 1564-1575.
- [8] Araki S, Sawada H, and Mukai R. Underdetermined blind sparse source separation for arbitrarily arranged multiple sensors[J]. *Signal Processing*, 2007, 87(8): 1833-1847.
- [9] 谭北海, 谢胜利. 基于源信号数目估计的欠定盲分离. *电子与信息学报*, 2008, 30(4): 863-867.
- Tan B H and Xie S L. Underdetermined blind separation based on source signals' number estimation[J]. *Journal of Electronics and Information Technology*, 2008, 30(4): 863-867.
- [10] Hamada T, Nakano K, and Ichijo A. Wavelet-based underdetermined blind source separation of speech mixtures. *IEEE ICCAS*, Seoul, Korea, 2007: 2790-2794.
- [11] Peng D and Xiang Y. Underdetermined blind source separation based on relaxed sparsity condition of sources[J]. *IEEE Transactions Signal Processing*, 2009, 57(2): 809-814.
- [12] Cruces S, Castedo L, and Cichocki A. Robust blind source separation algorithms using cumulants[J]. *Neural Computing*, 2002, 49(12): 87-118.
- [13] Xiao M, Xie S L, and Fu Y L. A statistically sparse decomposition principle for underdetermined blind source separation. *Intelligent Signal Processing and Communication Systems*, Guang Zhou, China, 2005: 165-168.

陈晓军: 男, 1983年生, 硕士, 研究方向为雷达信号盲分离。

成昊: 男, 1982年生, 博士生, 研究方向为信号盲处理技术。

唐斌: 男, 1964年生, 教授, 博士生导师, 研究电子对抗技术和系统、雷达信号盲分离。