

文章编号:1672-3961(2008)04-0097-05

基于最小均方误差和稀疏特征的欠定盲源分离

白树忠^{1,2}, 刘琚², 孙国霞²

(1. 山东大学电气工程学院, 山东 济南 250061;

2. 山东大学信息科学与工程学院, 山东 济南 250100)

摘要:针对欠定条件下的盲源分离问题,即观测信号个数小于信源个数的情况,提出了一种基于最小均方误差和稀疏特征的算法.首先,利用变换后信源的稀疏特征,采用一新的势函数通过聚类算法估计混叠矩阵.然后利用混叠矩阵和信源自身的相关性,通过寻找信源在聚类方向时间点上的精确值,以均方误差最小为准则寻找最佳分离矩阵实现信源的分离,克服了传统的分离算法在寻找最佳分离子矩阵方面的缺点.仿真结果显示使用该方法分离的信号具有更高的信噪比,和其他同类方法相比具有更优越的分离性能.

关键词:稀疏性;欠定分离;最小均方误差

中图分类号:TN911.7 **文献标志码:**A

An algorithm for under-determined blind source separation based on the least-mean-square error and sparse features

BAI Shu-zhong^{1,2}, LIU Ju², SUN Guo-xia²

(1. School of Electrical Engineering, Shandong University, Jinan 250061, China;

2. School of Information Science and Engineering, Shandong University, Jinan 250100, China)

Abstract: An algorithm was presented based on the least-mean-square error and sparse features for under-determined blind source separation, i. e., observed signal numbers are less than sources numbers. Based on the clustering method, the mixing matrix was first estimated by a new potential function using the sparseness of sources. By using the estimated mixing matrix and the self-correlation of sources and searching the accurate values at the source clustering directions, the optimal sub-matrix for separation was obtained according to the least-mean-square error criterion. This can overcome the disadvantages of traditional algorithm in searching the optimal sub-matrix. Simulation results show that the separated signals have higher SNR, and the proposed approach has better separation performance compared with other similar methods.

Key words: sparseness; under-determined separation; the least-mean-square error

0 引言

大多数的盲源分离是采用独立分量分析算法,利用高阶统计算法使分离出的信号尽量在统计上相互独立以实现分离,且通常假定观测信号个数大于或等于信源个数.但很多情况下这种条件很难满足,

如独立分量分析中基本的“鸡尾酒会”问题等.因此欠定条件下的盲源分离问题具有重要的实际意义.

在欠定条件下,由于系统是不可逆的,理论上分离有无穷多解,即使混叠矩阵已知,信号的分离较可逆系统要困难的多,且不同的分解方法给出不同的解.但是假若信号在时域或变换域具有某些特征,如具有超高斯分布(语音信号)或稀疏性,即大多数的

收稿日期:2008-01-18

基金项目:教育部留学回国人员科研启动基金资助项目([2005]55);教育部新世纪优秀人才支持计划基金资助项目(NECT-05-0582);高等学校博士学科点专项科研基金资助项目(20050422017);北京大学视觉与听觉信息处理国家重点实验室开放基金资助项目(0403)

作者简介:白树忠(1965-),男,山东济宁人,副教授,博士研究生,研究方向为数字信号处理,盲信号处理等.

E-mail: baishzh@sdu.edu.cn

信号值较小或接近于0,少数值取较大的值,在这种情况下可以利用信源的这种统计特征来实现信号的分离.

稀疏描述近几年受到很多学者的关注^[1-2],特别是在欠定条件下的盲源分离问题,并提出了一些算法来实现信号的分离.目前一般采用最大后验概率为准则的极大似然估计算法,Lewicki^[3]在假定信源为超高斯分布的情况下,采用贝叶斯方法来估计混叠矩阵和信源.在两步分离算法中^[1,4],首先估计混叠矩阵,常用的算法有K-均值法^[5]、中值聚类法^[6]、势函数、比幅聚类法^[7]等,在第二步的信源分离阶段,有最小化L¹范数法^[5],最短路径法^[7],统计分离法^[4]等.其中最小化L¹范数法是最基本的算法,能给出确定的解且对噪声具有较好的鲁棒性^[5].尽管如此,Takigawa等在分析了L¹范数法的性能后指出^[8],即使在信源为超高斯分布的情况下,L¹范数法也不能给出最好的信号分离.以上算法,都是在概率的意义下得到的,从统计的理论上来讲在一定的条件下可给出较好的分离效果,但在某一固定的时间点上会产生很大的误差,因此算法对信源的稀疏性要求较高.

本文采用两步的方法来实现欠定条件下的信号分离:首先,采用加权的势函数来估计信源的个数和混叠矩阵;然后在混叠矩阵已知的条件下,提出了一种新的基于最小均方误差的分离方法.最后给出仿真实验结果.

1 系统模型及问题描述

考虑如下无噪声情况下的线性混叠模型:

$$\mathbf{X} = \mathbf{A}\mathbf{S}. \quad (1)$$

为分析方便,通常将矩阵 \mathbf{A} 写成列向量 \mathbf{a}^i 的形式, i 代表列号,同时加上时间标志,如式(2)所示:

$$\mathbf{x}^t = \sum_{i=1}^N \mathbf{a}^i s_i^t, \quad t = 1, 2, \dots, T. \quad (2)$$

这里 \mathbf{X} 为 $M \times T$ 的矩阵,对应于时间 $t = 1, 2, \dots, T$,其列向量为对应于 t 时刻的观测值向量 \mathbf{x}^t ; \mathbf{S} 为 $N \times T$ 的矩阵,对应于时间 $t = 1, 2, \dots, T$,其列向量为对应于 t 时刻的信源值向量; s_i^t 为第 i 个信源在 t 时刻的值; \mathbf{A} 为 $M \times N$ 的混叠矩阵,并假设其行是满秩的; M, N 分别为观测信号和信源信号个数; T 为样本点数.欠定情况下的分离问题是在只假设信源 s_i 之间相互独立, \mathbf{A} 和 \mathbf{S} 均未知且 $M < N$ 的情况下实现信源的分离,由于式(2)不可逆,故无法采用传统的通过估计 \mathbf{A} 来实现分离的算法,为此必须充分利

用信源的先验知识在众多解中来寻找趋向信源 \mathbf{S} 的解.

2 信源个数和混叠矩阵的估计

根据公式(2),如果信源具有超高斯分布特征或稀疏性,在某一固定的时间点 t 上,只有1个信源的取较大的值,而其他信源的取值较小或趋向于0,则在观测信号 \mathbf{x}^t 的散点图上将显示出信源的聚类方向 \mathbf{a}^i ,由此可以通过聚类算法得到信源的个数和单位化的混叠矩阵,可见估计聚类方向就等价于估计混叠矩阵 \mathbf{A} .由于一般信源在时域中并不是稀疏或充分稀疏的,因此,需要对信号进行线性变换使之尽量稀疏.通过比较不同变换的性能发现,短时Fourier变换后的实部和虚部组成的观测向量具有较好的稀疏性.对 $M = 2$ 的情况,假设其列向量长度为1,则可以写成 $(\cos \alpha, \sin \alpha)^T$ 的形式,因此可以采用角度来描述聚类方向.为此在聚类方向 α 的邻域内定义基函数窗口如下:

$$\varphi(\alpha) = \exp(-|\alpha|), \quad |\alpha| < \pi/4. \quad (3)$$

则加权的势函数为

$$\Phi(\theta) = \sum_i l_i^2 \varphi(\eta(\theta_i - \theta)). \quad (4)$$

其中, θ 为待定的聚类中心的角度; θ_i 为散点图中各个样本点的角度; η 为尺度因子; l_i 为各个样本点的权重,这里取为样本点向量长度的平方,以减小多数小值对聚类方向的影响,提高估计的精度.对于 $\theta \in [0, \pi)$,通过势函数 $\Phi(\theta)$ 的极大值的个数来确定源信号的个数,并利用 $\Phi(\theta)$ 局部最大值所处的位置来估计各个向量 \mathbf{a}^i ,从而估计出混叠矩阵 \mathbf{A} .

在实际应用中,把 $\theta \in [0, \pi)$ 作 K 等分,则会得到离散的函数 $\Phi(\theta)$ 值,估计聚类方向就等价于确定 $\Phi(\theta)$ 的各个局部峰值的位置.由于预先不知道信源的个数,若尺度因子 η 大,则定位精度高,但 $\Phi(\theta)$ 不平滑,会出现多个极大值;若尺度因子 η 小,则 $\Phi(\theta)$ 平滑,极大值个数明显,可以准确地估计信源的个数,但定位精度差.为此先采用较小的 η 估计出极大值的个数和出现的区域以确定信源的个数,然后在该区域中利用较大的 η 值估计出现极大值的位置以精确确定混叠矩阵的数值.同时为了降低计算的复杂度和提高精度,对于 $l_i < C$,则置 $w_i = 0$, C 为常数.通过调整 η, C, K 等参数来获得最优估计.与传统方法相比较,本文提出的基函数 $\varphi(\alpha)$ 更具有方向性.图1显示了信号的变换及聚类过程.

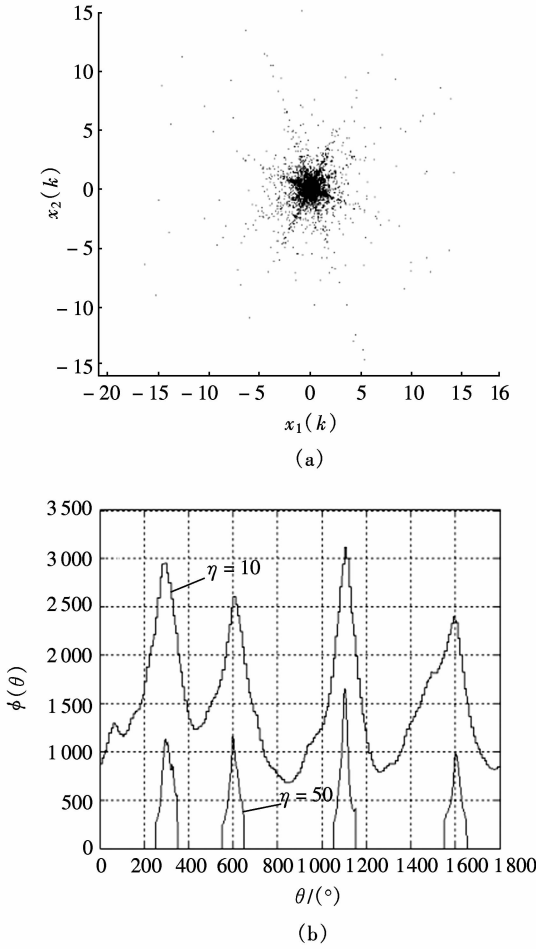


图1 (a)图为混叠信号经 Fourier 变化后实部和虚部组成的散点图; (b) 势函数对应不同尺度因子 $\eta = 10$ 和 $\eta = 50$ 的变化曲线
 Fig.1 (a) The scatter plot of the real and imaginary parts of Fourier transform of mixture, which has better sparseness; (b) The curves of the potential function respectively corresponding to $\eta = 10$ and $\eta = 50$

3 信源分离

混叠矩阵 \mathbf{A} 被估计出来之后, 下一步即估计源信号. 在信号恢复阶段, 采用最大后验概率法. 假设信号具有 Laplace 概率密度分布函数:

$$p(s_i) = \frac{1}{2\beta} \exp(-|s_i|/\beta). \quad (5)$$

其中, β 为方差参数. 由于信号间相互独立, 则信源 \mathbf{S} 的联合概率密度为

$$p(\mathbf{S}) = \prod_{i=1}^N \frac{1}{2\beta} \exp(-|s_i|/\beta) = \frac{1}{(2\beta)^N} \exp\left(-\frac{1}{\beta} \sum_{i=1}^N |s_i|\right). \quad (6)$$

由于 \mathbf{A} 已知, 在 $\mathbf{X} = \mathbf{AS}$ 的条件下对式(6)最大化得

$$\hat{\mathbf{S}} = \operatorname{argmax} \left(\frac{1}{2^N \beta^N} \exp\left(-\frac{1}{\beta} \sum_{i=1}^N |s_i|\right) \mid \mathbf{X} = \mathbf{AS} \right) = \operatorname{argmin}(\|\mathbf{S}\|_1 \mid \mathbf{X} = \mathbf{AS}). \quad (7)$$

可见, 最小化信号的 1 范数 $\|\mathbf{S}\|_1$ 可给出信源的估计. 上述问题可以转化成求解式(8)的线性规划问题^[2]:

$$\min \sum_{i=1}^T \sum_{l=1}^N |s_l^i|, \text{ 给定 } \mathbf{AS}^t = \mathbf{x}^t. \quad (8)$$

3.1 L¹ 范数法和统计稀疏分解法的缺点

L¹ 范数解在信源充分稀疏的条件下可以获得很好的分离效果, 并与 L⁰ 范数解等价^[5]. 但它对源信号的稀疏性要求很高, 本身存在不足. 文献[7]证明了最小化 1 范数法等价于最短路径法. 最短路径法通过寻找最靠近样本点的两个基向量的线性组合作为最小化 1 范数的解.

如图 2 所示为 2 个观测信号和 4 个源信号的情况, \mathbf{x}^t 为某一时间点上的观测向量, \mathbf{a}^i 为信源的聚类方向, 利用最短路径法或 L¹ 范数法只有线性组合 $(\mathbf{a}^1, \mathbf{a}^2)$ 、 $(\mathbf{a}^2, \mathbf{a}^3)$ 、 $(\mathbf{a}^3, \mathbf{a}^4)$ 中的一对有可能被选中, 而其他的组合则不可能发生, 这与实际情况不符. 由于每一个信源在某一时间点上都有可能取得较大的值概率, 且 \mathbf{x}^t 的大小与方向的随机性与 \mathbf{a}^i 无关, 信源越多, 重合的信源矢量越多, 误差就越大, 越不能反映实际的信源情况.

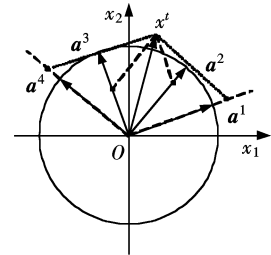


图2 观测信号 \mathbf{x}^t 的分解示意图, $M = 2, N = 4$
 Fig.2 The decomposition of \mathbf{x}^t to different column vectors combination, $M = 2, N = 4$

基于统计稀疏性的分离算法^[4], 在分离矩阵的选择上作了一定的改进, 在一定宽度的窗口上采用统计的方法, 用统计值代替窗口中心点时刻的值, 但由于统计的平滑作用, 窗口中某一两个值的改变对统计值影响很小, 造成分离矩阵的缓慢变化, 分离出的某些信号在某一时间段内出现连续的零值, 另一些信号则连续出现较大的值, 与信源的分布特性明显不同. 窗口越宽这种现象越严重, 窗口变小又不具有明显的统计特性. 因此该方法改进效果不明显.

3.2 观测信号的分布特性分析

由公式(7)可以看出, 最小化 1 范数可给出最优解, 但这是在假定信源为超高斯分布的情况下得到的. 分析观测信号不同数值范围的分布特性可知, 取值较大的点具有较好的稀疏性, 在这些点上分离效果较好, 然而对取值较小的点而言, 各个信源出现较小值的概率都较大. 从观测信号的角度来讲, 这些值的分布特性将趋向于高斯分布, 数值越小这种现象

越明显,因而其稀疏性较差或很差,即使经过变换也很难达到好的稀疏性(参见图1(a)).如果按最小化1范数进行分离,则不能保证分离的信号具有最小的1范数.

3.3 最小均方误差算法

根据以上的分析,本文提出了一种基于最小均方误差的方法,其基本思想是:利用第一步估计到的混叠矩阵,采用比幅法得到只有一个信源值不为0的点的集合,由此可得到信源在这些点上的精确值.对于2个观测信号的混叠情况,若在某一时刻 t ,只有 s_i 信源不为0,其他信源的值均为0,则有

$$k^t = \frac{x_1^t}{x_2^t} = \frac{a_{1i}s_i^t}{a_{2i}s_i^t} = \frac{a_{1i}}{a_{2i}}. \quad (9)$$

由于 a_{ij} 已知,可得到只有 s_i 信源不为0的所有点,实际应用时取一容限值,本文取 $\pm 5\%$.设其点得集合为 S_i ,则集合中 s_i 的精确值为

$$s_i^t = \frac{x_1^t}{a_{1i}} = \frac{x_2^t}{a_{2i}}. \quad (10)$$

在进行信源分离时,采用最小化1范数和最小均方误差相结合,在保证分离信源具有最小1范数的同时,利用信源相邻点之间的相关性,使分离信源在这些对应点附近与实际的信源精确值的误差最小.为此采用一窗口,以最小均方误差为准则寻找最佳分离矩阵对信源进行分离.在不同的时间点上有以下3种情况:

(1) 在精确点上,按式(10)进行分离.

(2) 若窗口内无精确值,按式(8)最小化1范数寻找最优子矩阵进行分离.

(3) 若窗口内有精确值,按最小均方误差寻找最优子矩阵进行分离,如式(11):

$$\min Err = \frac{1}{N_1} \sum_{i=1}^{N_1} \left\{ \frac{1}{N_2} \sum_{k=1}^{N_2} \left(\frac{\hat{s}_{ik} - s_{ik}}{s_{ik}} \right)^2 \right\}, S_{ik} \in S_i. \quad (11)$$

其中, N_1 为窗口内第 i 个信源具有精确值点的个数; s_{ik} 为 t 时刻的精确值; N_2 为窗口内点的个数; \hat{s}_{ik} 为按某一分离矩阵得到的第 i 信源的估计值.最小化式(11)可获得窗口中心点的最佳分离矩阵以实现分离.

按以上方法进行分离,在统计上保证了分离的信源具有最小的1范数,满足式(7)的要求,在精确值的附近可获得逼近信源解的最佳分离矩阵,同时在具体的时间点上又具有很好的跟踪信源内在变化的特性,信源的稀疏性越好或稀疏点在整个信源时段分布越均匀,则分离效果越好.

从运算量上来看,虽然较 L^1 范数法有所增大,但与其它统计算法相比,由于该方法侧重于信号动

态特性的跟踪,窗口取的较小,运算量要明显减少,仿真时也证明了这一点,因此该方法具有较好的使用价值.

4 仿真实验及分析

原始语音来自TIMIT语音数据库,采样频率为8 kHz,字长为16 bit.一共4路语音,两路男音和两路女音,每路语音取12 288个样本点.

4.1 信源个数与混叠矩阵的估计

取一 2×4 矩阵对原始语音进行混叠得到二路混叠信号,混叠矩阵取为

$$A = \begin{bmatrix} \cos 30^\circ & \cos 60^\circ & \cos 110^\circ & \cos 160^\circ \\ \sin 30^\circ & \sin 60^\circ & \sin 110^\circ & \sin 160^\circ \end{bmatrix}.$$

采用两步估计算法进行估计.由于信源在时域中不充分稀疏,采用Fourier变换,每帧加Hanning窗,窗口宽度为256点,对应时域长度32 ms.由于是线形变换,变换后其实部和虚部的系数矩阵 A 保持不变,考虑到Fourier变换的对称性,取对应于正频率的实部和虚部的组合组成变换后的观测向量,以此来估计 A .第一步估计信源个数,取小尺度因子 $\eta = 10$, $K = 180$,分辨率为 1° ,得到4个极值的平滑曲线,参见图1(b),第二步在上述4个极值点的 $\pm 5^\circ$ 的小范围内按 0.1° 的分辨率,用大尺度因子 $\eta = 50$ 进行精确定位,得到聚类方向分别为 29.6° 、 59.9° 、 110.3° 和 159.9° ,可见该方法能精确得到混叠矩阵 A ,其误差矩阵为

$$A - \hat{A} = \begin{bmatrix} -0.0035 & -0.0015 & 0.0049 & -0.0006 \\ 0.0061 & 0.0009 & 0.0018 & -0.0016 \end{bmatrix}.$$

4.2 源信号的分离

为验证该分离算法的性能,对2个观测信号和4个源信号的情况,取窗口长度 $N_2 = 8$,在稀疏性较差(情况I)和稀疏性较好(情况II)下分别进行了分离,同时与最小化1范数法进行比较.定义分离信号的信噪比如式(12),其中 \hat{S} 为分离信号, S 为原始信号,实验结果见表1.图3为一实际的分离结果.

$$SNR = 10 \log \| S \|^2 / \| \hat{S} - S \|^2. \quad (12)$$

表1 不同分离方法的信噪比性能比较表
Table 1 The SNR comparison of different methods

源信号	情况 I		情况 II	
	L^1 范数法	本文方法	L^1 范数法	本文方法
S_1	3.256	5.135	23.869	24.347
S_2	2.493	4.079	24.014	24.598
S_3	4.278	6.027	28.007	28.823
S_4	4.474	5.925	23.231	23.910

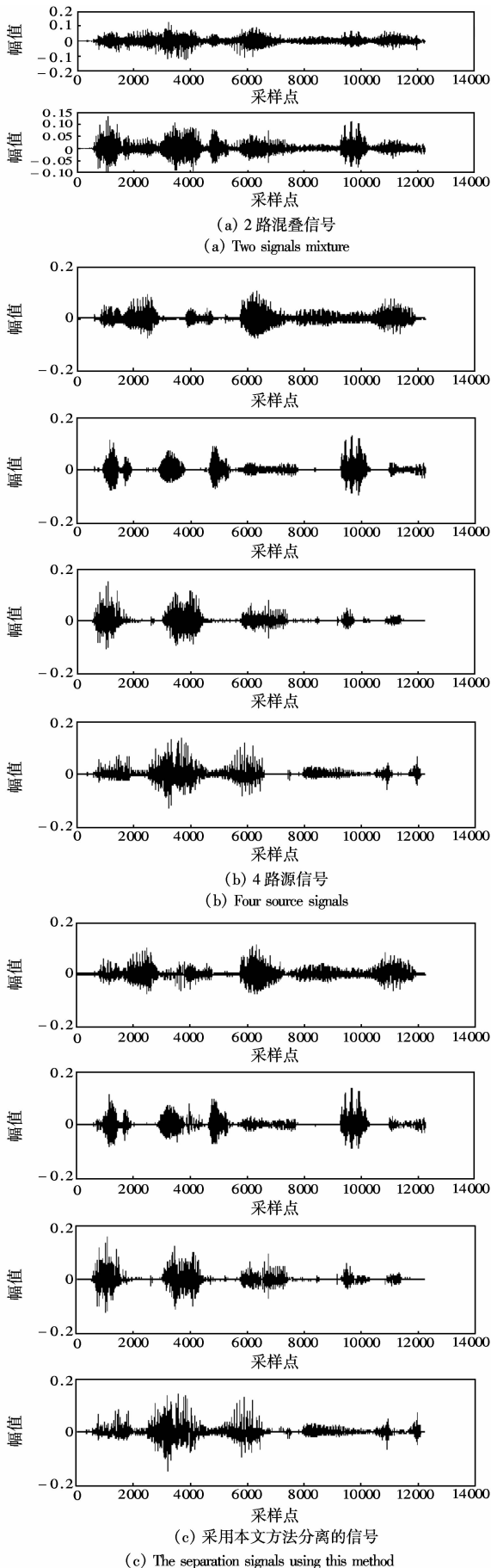


图3 语音信号的分离

Fig.3 The separation of speech signal

由表1的结果可见,在两种稀疏性的情况下,本文的方法都比最小化1范数法取得了较好的分离效果,在信源稀疏性差(情况I)的情况下效果更好.可见该方法对稀疏性的要求较小,取得了满意的分离效果.主观评价分离效果,通过试听分离后的语音信号,与最小化1范数法相比较,语音清晰度得到较好的提高,连贯性得到改善.

5 结论

欠定条件下的盲源分离问题在盲信号处理中具有重要的实际意义,在混叠矩阵估计上,采用两次聚类实现精确估计.在信源恢复上,通过分析信源在不同数值情况下的稀疏性,采用逼近信源精确值的方法以获得最佳的分离矩阵,克服了解 L^1 范数解的缺点,降低了对信源稀疏性的要求,仿真结果取得了较 L^1 范数方法更好的分离效果.

参考文献:

- [1] LI Y Q, AMARI S, CICHOCKI A, et al. Probability estimation for recoverability analysis of blind source separation based on sparse representation[J]. IEEE Transactions on Information Theory, 2006, 52(7):3139-3152.
- [2] ZIBULEVSKY M, PEARLMUTTER B A. Blind source separation by sparse decomposition in a signal dictionary[J]. Neural Computation, 2001, 13(4):863-882.
- [3] LEWICKI M, SEJNOWSKI T J. Learning overcomplete representations networks[J]. Neural Computation, 2000, 12(2):337-365.
- [4] ZHANG Wei, LIU Ju, SUN Jian-de, et al. A new two-stage approach to underdetermined blind source separation using sparse presentation[C]// ICASSP2007. USA: Hawaii, 2007: 953-956.
- [5] LI Y Q, ANDRZEJ C, AMARI S. Analysis of sparse representation and blind source separation[J]. Neural Computation, 2004, 16:1193-1234.
- [6] THEIS F J, PUNTONET C G, LANG E W. Median-based clustering for underdetermined blind signal processing [J]. IEEE Signal Processing Letter, 2006, 13(2):96-99.
- [7] THEIS F J, LANG W E, PUNTONET C G. A geometric algorithm for overcomplete linear ICA[J]. Neurocomputing, 2004, 56:381-398.
- [8] TAKIGAWA I, KUDO M, TOYAMA J. Performance analysis of minimum L^1 -norm solutions for underdetermined source separation[J]. IEEE Transactions on Signal Processing, 2004, 52(3):582-591.

(编辑:许力琴)