

带源个数估计的 BPSK 信号盲分离算法

谭北海 赵敏 谢胜利

(华南理工大学电子与信息学院 广州 510640)

摘要: 目前盲分离研究已有算法众多,但有关数字信号或有限字符集的盲分离研究尚不多见,而带源个数估计的此类盲分离算法更鲜有涉及。针对这类问题,该文提出了一种新颖的 BPSK 数字信号的盲分离算法,首先由接收到的观测信号的特征,在无噪和有噪情况下分别给出了估计源信号的数目算法;然后再利用观测信号之间的关系估计出混叠矩阵,并在算法中给出了证明。通过估计的混叠矩阵即可恢复得到分离信号,此时得到的分离信号与源信号或者顺序发生了交换,或者产生了符号之间的差别,但并不影响盲分离。最后的仿真结果显示了该文提出的算法在估计混叠矩阵以及最后恢复源信号上都是非常成功,也表明了此算法的可行性和优异性能。

关键词: 盲分离; BPSK 信号; 源数估计; 混叠矩阵; 初等变换

中图分类号: TN911.7

文献标识码: A

文章编号: 1009-5896(2009)07-1624-06

Blind Separation Algorithms of BPSK Signals by Estimating Source Number

Tan Bei-hai Zhao Min Xie Sheng-li

(School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: At present, there exist a lot of algorithms of blind separation, among which there are few algorithms focusing on blind separation of digital signals and estimating source number. To solve the problem, this paper proposes a novel algorithm to solve the blind separation problem of BPSK signals. First, according to characteristics of observations, the algorithm to estimate the source number is given in noise circumstance and no noise circumstance. Second, mixing matrix is gotten by using the relation of observations, which is also proved feasible. Finally, the source signals are recovered by permutations and sign changes of their rows, which are allowed in blind separation. It is well shown that the algorithms are excellent and feasible to estimate the mixed matrix and recover source in the last simulation.

Key words: Blind separation; BPSK signals; Source number estimation; Mixed matrix; Primary transformation

1 引言

近年针对盲分离问题的研究已经吸引了越来越多学者^[1-6]。盲源分离(Blind Source Separation, BSS)是指在不知道源信号和传输信道的参数情况下,常根据输入源信号的统计特性,仅由根据观测信号恢复出源信号的各个独立成分的过程,因此这一过程又称作独立分量分析(Independent Component Analysis, ICA)^[7]。盲分离从开始到现在已经出现了很多研究方法,也在许多方面都得到了广泛的应用,如陈列处理、多用户通信、生物医学工程等等。

由于数字通信的发展,对数字信号盲分离的研

究也显得尤为重要。已有的关于有限字符集信号盲分离以及在数字信号盲分离算法方面,出现的方法如文献[8,9]的两步法,此算法第1步首先从 d 个天线得到的观测信号向量中进行聚类,则一共有 2^d 个不同的观测向量组成的类,第二步利用奇异值分解等方法对每一个类里面的向量找出与之对应的源信号向量,最后从所有的观测信号分别对应得到源信号,因此计算存在偏差;另外 Li Yuanqing^[10]提出了在欠定情况下的数字信号盲分离方法,但是这种方法的鲁棒性欠佳;较早出现的还有 Alle-Jan^[11]提出的基于源信号常模的特征值分解方法。除了以上的算法外,如文献[12-15]分别采取迭代算法,但是常陷入局部最优,且收敛性依赖于初始值的选择。由此可见,近年针对数字信号的盲分离研究仍不多见且以往的算法都具有一定的局限性,为了丰富数字信号盲分离的研究,本文提出了一种新的数字信号盲分离的方法,仅仅根据观测信号之间的关系来估计源

2008-08-09 收到, 2009-03-17 改回

国家自然科学基金重点项目(U0635001), 国家自然科学基金(60674033, 60774094), 广东省自然科学基金(05006508), 华南理工大学博士后创新科学基金(20080217)和中国博士后科学基金(20080430828)资助课题

信号个数以及混叠矩阵,而且估计出来的混叠矩阵只是初始混叠矩阵的一个初等列变换,在线性混叠情况下,最后得到恢复源信号只与初始源信号或者只是顺序发生了交换,或者与其相差一个符号。因此分离是成功的,而且可以完整地恢复源信号,理论和仿真证明了此方法的高效的性能。

2 数字信号的盲分离模型

在数字信号盲分离中,假设有来自于 m 个用户的 m 个BPSK信号到达 d 个接收天线,这样第 p 个观测信号如下^[8]:

$$x_p(t) = \sum_{i=1}^m q_i a_{pi} \sum_{j=1}^{\infty} b_i(j) s(t - jT - \tau_i) + w_p(t) \quad (1)$$

其中 T 为波特周期, q_i 为第 i 个源信号的放大系数, a_{pi} 为第 p 个观测信号到第 i 个源信号之间的响应系数, $b_i(\cdot) = \pm 1$ 表示第 i 个源信号上的比特流, $s(\cdot)$ 表示信号载波, τ_i 表示第 i 个源信号到天线的延迟, $w_p(\cdot)$ 表示在第 p 个天线上的高斯白噪声。

在某些假设条件下^[8],式(1)可以写成

$$x_p(n) = \sum_{i=1}^m q_i a_{pi} b_i(n) + w_p(n) \quad (2)$$

这里 $w_p(n)$ 是均值为0方差为 σ^2 的白噪声,如果用向量表示,式(2)可以写成

$$\mathbf{x}(n) = \mathbf{A} \mathbf{s}(n) + \mathbf{w}(n) \quad (3)$$

式中 $\mathbf{s}(n) = [b_1(n), \dots, b_m(n)]^T$, $\mathbf{x}(n) = [x_1(n), \dots, x_d(n)]^T$, $\mathbf{w}(n) = [w_1(n), \dots, w_d(n)]^T$, $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m] = [q_1 \tilde{\mathbf{a}}_1, \dots, q_m \tilde{\mathbf{a}}_m]$, $\tilde{\mathbf{a}}_i = [a_{1i}, \dots, a_{di}]^T$, $\mathbf{a}_i = q_i \tilde{\mathbf{a}}_i$ 。

因此,这是一个典型的数字信号盲分离模型。假设有 N 个采样,式(3)可以写成矩阵形式:

$$\mathbf{X}(n) = \mathbf{A} \mathbf{S}(n) + \mathbf{W}(n) \quad (4)$$

这里 $\mathbf{X}(n) = [\mathbf{x}(1), \dots, \mathbf{x}(N)]$, $\mathbf{S}(n) = [s(1), \dots, s(N)]$ 和 $\mathbf{W}(n) = [w(1), \dots, w(N)]$ 。

下面假设 \mathbf{A} 为列满秩混叠矩阵且 $d \geq m$,即观测信号个数大于或等于源信号数目的情况下讨论本文新的算法。在没有噪声情况下,则式(3)和式(4)分别可写为

$$\mathbf{x}(n) = \mathbf{A} \mathbf{s}(n) \quad (5)$$

$$\mathbf{X}(n) = \mathbf{A} \mathbf{S}(n) \quad (6)$$

3 源信号数目估计

在本文算法中,考虑源信号来自BPSK调制信号,且源信号满足“充分激励条件”,即当 N 充分大的时候,式(6)中构成 $\mathbf{S}(N)$ 的列向量会包含 $\mathbf{s}(n)$ 所有可能出现的情况。同时算法也假设在非欠定盲分离下进行研究,即观测器的个数大于或等于源信号的个数。

3.1 无噪声干扰情况

在式(6)中无噪声干扰,且源信号满足“充分激励条件”,假设源信号个数为 m , $M = 2^m$ 。因此每个向量 $\mathbf{s}(n)$ 可有 $M = 2^m$ 个不同取值,不妨设它们为 \mathbf{s}_i , $i = 1, 2, \dots, M$,其中 \mathbf{s}_i 形如 $[\pm 1, \pm 1, \dots, \pm 1]^T$ 。同时发现,构成 $\mathbf{S}(N)$ 的列向量包含 $\mathbf{s}(n)$ 所有可能出现的情况,让 $\mathbf{s}(n) \in \varphi = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M\}$,且这 2^m 个不同的列向量在 φ 中具有对称性,即 $\forall i, j \leq M, \exists j \neq i, \mathbf{s}_i = -\mathbf{s}_j$ 。又因为在线性瞬时混叠中, $\mathbf{x}_i = \mathbf{A} \mathbf{s}_i, i = 1, 2, \dots, M$,所以经过式(6)混叠得到的观测信号 $\mathbf{X}(N)$ 中所有的列向量必定包含 \mathbf{x}_i 所有可能出现的 M 种不同向量取值,这里称这 M 个向量为基向量, $\mathbf{x}(n) \in \phi = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$,由 φ 中元素的对称性,可知 ϕ 中元素也具有对称性,即 $\forall i, j \leq M, \exists j \neq i, \mathbf{x}_i = -\mathbf{x}_j$ 。如3个观测器3个源信号的散点图如图1,图2所示。

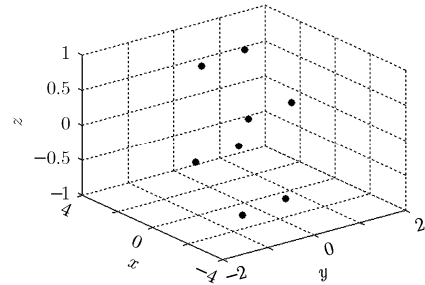


图1 $d = 3, m = 3$ 时的散点图

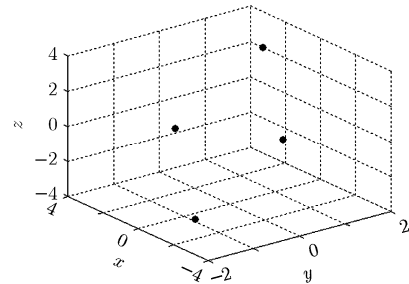


图2 $d = 3, m = 2$ 时的散点图

在无噪声干扰情况下,估计源信号的个数,首先需在观测信号 $\mathbf{X}(N)$ 中估计基向量的个数。下面用一段MATLAB程序来描述详细算法过程:

```

y(1) = x(1);
l = 1;
for j = 2 : N
    m = 0;
    for i = 1 : l
        if |x(j) - y(i)| > ε
            m = m + 1;
        end
    end
end

```

```

end;
end;
if m = l
y(l+1) = x(j);
l = l + 1;
end
end;

```

其中 ε 为一个小的正数。由上述算法，得到了 k 个不同的向量 $\mathbf{y}(i), i = 1, 2, \dots, k$ ，同时让 $\mathbf{x}_i = \mathbf{y}(i), i = 1, 2, \dots, k$ 作为基向量，则 $m = \log_2 k$ 即为源信号个数。

3.2 有噪声干扰情况

如果混叠模型中存在噪声干扰，则由式(3)可知 $\mathbf{x}(t) = \mathbf{a}_1 s_1(t) + \mathbf{a}_2 s_2(t) + \dots + \mathbf{a}_m s_m(t) + \mathbf{w}(t),$
 $t = 1, 2, \dots, N$ (7)

上式中 $\mathbf{x}(t)$ 表示 $\mathbf{X}(N)$ 的第 t 列， \mathbf{a}_i 表示混叠矩阵 \mathbf{A} 的第 i 列， $\mathbf{w}(t)$ 表示第 t 个噪声向量。

在源信号满足“充分激励条件”下，且噪声干扰不太强情况下。式(7)说明观测信号中仍然可以找到 $M = 2^m$ 个不同的基向量，散点图分别如下图 3，图 4 所示。

在噪声干扰影响不显著的情况下，观测信号的散点图呈现聚类状，因此散点图的聚类中心即为基向量，而聚类的数目即为 k ，因此在观测器个数 d 已知的条件下，源信号个数可由 $m = \log_2 k$ 得到。下面算法将从观测信号中找出所有的基向量，同时所有的聚类个数 k 。

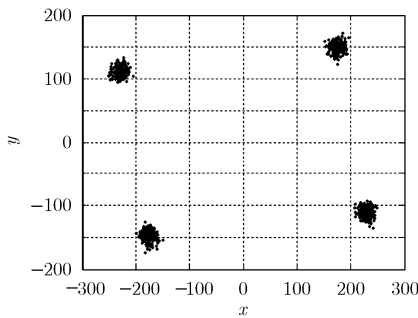


图 3 $d = 2, m = 2$ 在有噪声时的散点图

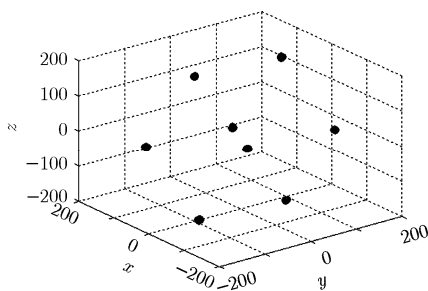


图 4 $d = 3, m = 3$ 在有噪声时的散点图

聚类的目标是使得距离相近的点聚为一类，距离比较远的点为不同的类，所以下面首先定义观测向量之间的相似关系如

$$r_{ij} = \exp\left(-\frac{\|\mathbf{x}(i) - \mathbf{x}(j)\|}{2\sigma^2}\right),$$

$$i = 1, 2, \dots, N; j = 1, 2, \dots, N$$
 (8)

上述中 $\|\mathbf{x}(i) - \mathbf{x}(j)\|$ 表示 $\mathbf{x}(i)$ 与 $\mathbf{x}(j)$ 之间的欧几里得距离， σ^2 为正态函数的方差。由式(8)定义的相似关系，可知 $0 \leq r_{ij} \leq 1$ ，且构成观测向量之间的相似矩阵 $\mathbf{R} = [r_{ij}]_{N \times N}$ 。文献[16]提出了针对解决此类问题的无监督聚类算法，本文在此算法基础上针对 BPSK 信号反映在观测信号的特征，对此算法进行了改正，以提高其性能。从上图 3，图 4 可知，观测信号聚类中心个数与观测器个数存在 2 的指数次幂的关系，并且呈现对称性，因此可在文献[16]最后得到的收敛聚类中心进行上述指导性合并。令 $\beta = \{\beta_i | \beta_i = \mathbf{a}_1 s_1 + \mathbf{a}_2 s_2 + \dots + \mathbf{a}_m s_m, s_j = \pm 1\}$ 表示在无噪声情况下所有基向量的集合，显然 β 中有 2^m 个元素 β_i ，同时让 $\alpha = \arg \min_{i \neq j, i, j \in \{1, 2, \dots, M\}} \{\|\beta_i - \beta_j\|\}, \|\cdot\|$ 表示欧几里得

距离。最后，本文提出的无监督聚类算法详细过程如下：

(1) 让 $\mathbf{y}(i) = \mathbf{x}(i), i = 1, 2, \dots, N$ ，即 $\mathbf{x}(i)$ 为 $\mathbf{y}(i)$ 的初值；

(2) 根据式(8)计算 $\mathbf{y}(i)$ 中列向量之间的相似度，并构成相似矩阵 \mathbf{R} ；

(3) 由下述方法修改相似矩阵

$$r_{ij} = \begin{cases} 0, & r_{ij} < \lambda \\ r_{ij}, & \text{其他} \end{cases}$$
 (9)

λ 为一个小的固定正常数；

(4) 通过如下方法更新 $\mathbf{y}(i), i = 1, 2, \dots, N$

$$\hat{\mathbf{y}}(i) = \frac{\sum_{j=1}^N r_{ij} \mathbf{y}(j)}{\sum_{j=1}^N r_{ij}}, i = 1, 2, \dots, N$$
 (10)

(5) 如果修改后的 $\hat{\mathbf{y}}(i)$ 与修改前的 $\mathbf{y}(i), i = 1, 2, \dots, N$ 无变化，则到第 6 步；否则令 $\mathbf{y}(i) = \hat{\mathbf{y}}(i), i = 1, \dots, N$ ，并继续第 2 步。

(6) 假设已经在收敛后得到的聚类中心为 $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_q\}, q > 2^m$ ，令 $L_k = \|\mathbf{c}_i - \mathbf{c}_j\|, i \neq j, i, j \in \{1, 2, \dots, q\}, k = 1, 2, \dots, C_q^2$ ，取 $L_0 = \arg \min_{k=1, 2, \dots, C_q^2} \{L_k\}$ ，如果

$L_0 < \varepsilon$ ，理论上 ε 需满足 $0 < \varepsilon < \alpha$ ，则将构成 L_0 的这两个聚类中心合并成一个聚类中心，得 $\mathbf{c}' = (\mathbf{c}_i + \mathbf{c}_j)/2$ ，然后在 \mathbf{C} 中去除 \mathbf{c}_i 和 \mathbf{c}_j ，加入新的 \mathbf{c}' ，继续按照上述方法合并较为相邻的类，直至在 \mathbf{C} 中所有的聚类中心不再满足上述条件。

(7)最后得到的聚类中心则为基向量 $\mathbf{x}_i, i = 1, 2, \dots, k$, k 为聚类中心的数目。因此源信号的数目 $m = \log_2 k$ 。

上述算法的修改步骤主要是提高相似度比较大的作用, 相似度比较小的降到无影响作用。第3步的 λ 通常取少于0.1的数, 或省略第3步也对算法无影响。

在上述有噪声干扰混叠和无噪声干扰混叠两种情况下, 通过本文算法都能得到源信号个数以及基向量, 又因为源信号为BPSK信号源, 且满足“充分激励条件”, 因此源信号具有对称性, 所以基向量也具有对称性, 在上述有噪声干扰混叠和无噪声干扰混叠两种情况下, 估计出的基向量, 为了方便下面更好地估计混叠矩阵, 需修正其对称性要求。

修正准则 1 在有噪和无噪声干扰情况下, 分别估计的所有基向量 $\mathbf{x}_i, i = 1, 2, \dots, M$ 中, 如果 $i \neq j$, $|\mathbf{x}_i + \mathbf{x}_j| < \varepsilon, i, j = 1, 2, \dots, M$, 这里的 ε 为一个小的正数, 则令 $\mathbf{x}_i = \frac{\text{abs}(\mathbf{x}_i) + \text{abs}(\mathbf{x}_j)}{2} \cdot \text{sgn}(\mathbf{x}_i)$, $\mathbf{x}_j = \frac{\text{abs}(\mathbf{x}_i) + \text{abs}(\mathbf{x}_j)}{2} \cdot \text{sgn}(\mathbf{x}_j)$, 这里的 abs 表示取绝对值, sgn 为符号函数, ‘ \cdot ’为矩阵的点乘。

4 算法原理及实现

在观测信号的基向量以及源信号数目得以估计并修正之后, 下面将利用它们估计混叠矩阵以及恢复源信号。

4.1 混叠矩阵的估计算法

在本文前面估计了源信号的个数, 以及得到了观测信号的基向量, 下面将通过它们来估计混叠矩阵。在本文算法中, 为了分离出数字信号, 首要任务是估计混叠矩阵 \mathbf{A} 。当 N 足够大, 则源信号满足“充分激励条件”, 也就是说源信号的每一个不同的列向量都会至少出现一次。因此(3)式中 $\mathbf{s}(k)$ 将有 $M = 2^m$ 种不同的取值, 用集合 $\varphi = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M\}$ 表示, 所以式(6)中 $\mathbf{S}(N)$ 矩阵中的每一个列向量都来自于集合 φ 的一种, 由集合 \mathbf{s} 通过(6)式可知可以得到 2^m 种不同的列观测信号, 即基向量, 用集合 $\phi = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$, 同时显然可以发现 $\mathbf{X}(N)$ 矩阵中的每一个列向量也都来自于集合 \mathbf{x} 的一种。因为

$$\mathbf{x}_i = \mathbf{A} \mathbf{s}_i \quad (i = 1, 2, \dots, M) \quad (11)$$

$\mathbf{x}_i = [x_{1i}, x_{2i}, \dots, x_{di}]^T$, $\mathbf{s}_i = [s_{1i}, s_{2i}, \dots, s_{mi}]^T$, 又因为 $s_{ji} \in \{+1, -1\}, j = 1, 2, \dots, m, i = 1, 2, \dots, M$ 得

$$\mathbf{x}_i + \mathbf{x}_k = \mathbf{A}(\mathbf{s}_i + \mathbf{s}_k), \quad i \neq k, i = 1, 2, \dots, M, k = 1, 2, \dots, M \quad (12)$$

由式(12)可知, 当 $(s_{1i} + s_{1k}) = +2$ 或者 $(s_{1i} + s_{1k}) =$

-2 , 而 $(s_{ji} + s_{jk}) = 0, j \neq 1, j \in \{1, 2, \dots, m\}$ 时, $\mathbf{x}_i + \mathbf{x}_k = (+2)\mathbf{a}_1$ 或者 $\mathbf{x}_i + \mathbf{x}_k = (-2)\mathbf{a}_1$, 即集合 \mathbf{x} 中第 i 个和第 k 个列向量的和是混叠矩阵 \mathbf{A} 的第1列 \mathbf{a}_1 的 $(+2)$ 或 (-2) 倍。同理可以知道当 $(s_{ti} + s_{tk}) = +2$ 或者 $(s_{ti} + s_{tk}) = -2$, 而 $(s_{ji} + s_{jk}) = 0, j \neq t, j \in \{1, 2, \dots, m\}$ 时, $\mathbf{x}_i + \mathbf{x}_k = (+2)\mathbf{a}_t$ 或者 $\mathbf{x}_i + \mathbf{x}_k = (-2)\mathbf{a}_t$, 即集合 \mathbf{x} 中第 i 个和第 k 个列向量的和是混叠矩阵 \mathbf{A} 的第 t 列 \mathbf{a}_t 的 $(+2)$ 或 (-2) 倍。下面为了找出混叠矩阵 \mathbf{A} 的所有列向量, 在 \mathbf{x} 中任取两个观测向量进行相加运算, 得到

$$\mathbf{y}_l = \mathbf{x}_i + \mathbf{x}_k, \quad l = 1, 2, \dots, C_M^2, i \neq k, i = 1, 2, \dots, M, \quad k = 1, 2, \dots, M \quad (13)$$

令集合 $\mathbf{Y} = \{\mathbf{y}_l, l = 1, 2, \dots, C_M^2\}$ 。

定义 1 在集合 $\mathbf{Y} = \{\mathbf{y}_l, l = 1, 2, \dots, C_M^2\}$ 中, 如果 $\mathbf{y}_i = \mathbf{y}_j$ 或 $\mathbf{y}_i = -\mathbf{y}_j, i \neq j$ 时, 将 $\mathbf{y}_i, \mathbf{y}_j$ 视为同一类 \mathbf{G}_r, r 表示这一类的集合的一个下标。

定义 2 定义一个新的集合 $\mathbf{S} = \{\mathbf{s}_l, l = 1, 2, \dots, C_M^2\}$, 其中 \mathbf{s}_l 满足以下条件

$$\mathbf{s}_l = \mathbf{s}_i + \mathbf{s}_k, \quad l = 1, 2, \dots, C_M^2, i \neq k, i = 1, 2, \dots, M, \quad k = 1, 2, \dots, M \quad (14)$$

这里 \mathbf{s}_i 和 $\mathbf{x}_i, \mathbf{s}_k$ 和 \mathbf{x}_k 以及 \mathbf{s}_l 和 \mathbf{y}_l 分别通过式(11), 式(13)和式(14)一一对应。因此也可以对集合 $\mathbf{S} = \{\mathbf{s}_l, l = 1, 2, \dots, C_M^2\}$ 按照定义1的方法进行分类, 当 $\mathbf{y}_i = \mathbf{y}_j$ 或 $\mathbf{y}_i = -\mathbf{y}_j, i \neq j$ 时, 将 $\mathbf{y}_i, \mathbf{y}_j$ 视为同一类 \mathbf{G}_r 时所对应的 $\mathbf{s}_i, \mathbf{s}_j$ 也分在同一类 \mathbf{H}_r 。

同时由定义1以及结合式(12)可知, 混叠矩阵 \mathbf{A} 的所有列向量一定来自于 m 个类 \mathbf{G}_r , 而且与这 m 个类 \mathbf{G}_r 中的相应的列向量相差 $(+2)$ 或 (-2) 倍。

定理 在按照定义1对集合 $\mathbf{Y} = \{\mathbf{y}_l, l = 1, 2, \dots, C_M^2\}$ 进行分类时, 与混叠矩阵 \mathbf{A} 的列向量有关的 m 个类和全部都是零向量的类里面包含的元素个数最多, 都是 2^{m-1} 个, 而不属于这 $m+1$ 个类的, 其里面包含的元素个数将少于 2^{m-1} 。

证明 $\mathbf{S} = \{\mathbf{s}_l, l = 1, 2, \dots, C_M^2\}$ 。其中 $\mathbf{s}_l = \mathbf{s}_i + \mathbf{s}_k, l = 1, 2, \dots, C_M^2, i \neq k, i = 1, 2, \dots, M, k = 1, 2, \dots, M$ 。当 $\mathbf{s}_l = \mathbf{0}$ 时, \mathbf{s}_i 有 2^m 种取值, 对应于每个 \mathbf{s}_i 都存在一个 \mathbf{s}_k 使得 $\mathbf{s}_i + \mathbf{s}_k = \mathbf{0}$, 但由对称性, $\mathbf{s}_i = \mathbf{0}$ 在 \mathbf{S} 中出现的次数为 2^{m-1} 次;

同理, 当 $\mathbf{s}_l = \mathbf{e}^r$ 时, 其中 \mathbf{e}^r 表示 \mathbf{s}_l 中有 r 个非零元素, 而其他元素全部为0的 $m \times 1$ 的列向量。从上述推导易知在 \mathbf{S} 中出现的次数为 2^{m-r} 次;

因此, 从 $\mathbf{s}_l \neq \mathbf{0}$ 开始, 当 \mathbf{s}_l 中非零元素个数每多一个, 属于此种情况的 \mathbf{s}_l 在 \mathbf{S} 中出现的次数就会少一半。又因为 \mathbf{s}_l 分别对应于 \mathbf{y}_l , 所以在各个不同类中, 类里面的向量含有一个非零元素或者全是零

元素组成类中元素个数最多,都为 2^{m-1} 个,其他的类里面的元素个数都少于这个数目,且随着向量中0元素的增加,类中元素个数依次减半。 证毕

4.2 源信号恢复算法

由上述定理可知,下面只要找到按照定义1中定义的类中出现次数最多的 m 个非零类,设为 $\hat{\mathbf{G}} = \{\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_m\}$,从每个类 $\mathbf{G}_i, (i = 1, 2, \dots, m)$ 中任取一个列向量除以2后作为 $\hat{\mathbf{a}}_i, (i = 1, 2, \dots, m)$,由 $\hat{\mathbf{a}}_i, (i = 1, 2, \dots, m)$ 构成新的矩阵 $\hat{\mathbf{A}}$,因此 $\hat{\mathbf{A}}$ 是将原混叠矩阵 \mathbf{A} 的一个初等列变换,其实 $\hat{\mathbf{A}}$ 只是在 \mathbf{A} 的各列乘了一个(+1)或者(-1),或者又交换了 \mathbf{A} 中一些列的位置,因此

$$\hat{\mathbf{A}} = \mathbf{A}\mathbf{P} \quad (15)$$

其中 \mathbf{P} 表示一个初等变换矩阵。

将式(15)代入式(4)和式(6)得

$$\mathbf{X}(N) = \hat{\mathbf{A}}\mathbf{P}^{-1}\mathbf{S}(N) \quad (16)$$

$$\mathbf{X}(N) = \hat{\mathbf{A}}\mathbf{P}^{-1}\mathbf{S}(N) + \mathbf{W}(N) \quad (17)$$

前者为无噪声干扰情况,后者为有噪声情况。

令 $\hat{\mathbf{S}}(N) = \mathbf{P}^{-1}\mathbf{S}(N)$,分别得

$$\mathbf{X}(N) = \hat{\mathbf{A}}\hat{\mathbf{S}}(N) \quad (18)$$

$$\mathbf{X}(N) = \hat{\mathbf{A}}\hat{\mathbf{S}}(N) + \mathbf{W}(N) \quad (19)$$

因为 \mathbf{P} 是一个初等变换矩阵,所以 \mathbf{P}^{-1} 也是一个初等矩阵,所以 $\hat{\mathbf{S}}(N)$ 为 $\mathbf{S}(N)$ 的一个初等行变换,也即 $\hat{\mathbf{S}}(N)$ 是由 $\mathbf{S}(N)$ 中的行乘以了(+1)或者(-1),或者交换行的顺序得到。

由式(18)和式(19)可知:

$$\hat{\mathbf{S}}(N) = \hat{\mathbf{A}}^{-1}\mathbf{X}(N) \quad (20)$$

$$\hat{\mathbf{S}}(N) = \hat{\mathbf{A}}^{-1}\mathbf{X}(N) - \hat{\mathbf{A}}^{-1}\mathbf{W}(N) \quad (21)$$

因此在有噪声干扰情况下,源信号恢复算法如式(20)恢复,源信号将会失真,但因为源信号来自BPSK调制信号,所以最后恢复的源信号可利用下述修正准则,以减少误码率。

修正准则 2 在 $s_{ij} \in \hat{\mathbf{S}}, i = 1, 2, \dots, m; j = 1, 2, \dots, N$,如果满足 $|s_{ij} - 1| > |s_{ij} + 1|$,则令 $s_{ij} = -1$,否则 $s_{ij} = 1$ 。

5 仿真实验

在仿真实验中,为了说明本文提出的算法在噪声干扰情况下也能顺利实现,定义源信号的误码率为:如果 \hat{s}_i 为某个源信号 s_i 的恢复,假设在某个时刻 t , $\hat{s}_i(t) \neq s_i(t)$,就说恢复信号产生了一个误码,遍历所有时刻 N ,找出总的误码为 p ,则误码率定义为

$$\text{error_rate} = (p/N) \times 100\% \quad (22)$$

下面仿真考虑在适量的高斯白噪声干扰情况下

的3个观测器,3个BPSK源信号的盲分离,取观测信号的采样数目为 $N = 100$ 进行仿真实验,混叠矩阵仍然由MATLAB随机函数命令randn(\cdot)生成,

$$\mathbf{A} = \begin{bmatrix} -1.2061 & 0.6525 & -0.1772 \\ 0.4879 & -0.3150 & -1.4956 \\ -0.2537 & 1.5930 & -0.9766 \end{bmatrix}, \text{适量高斯白噪声}$$

由MATLAB随机函数命令wgn(\cdot)生成,此时由式(4)

混叠后得到混叠信号散点图如图4所示。从观测器得到的观测信号,再由本文中以上提出的改进的无监督聚类算法得到基本向量,估计源个数估计,并在按照修正准则1修正后,利用第4节的混叠矩阵估计方法,其中得到的混叠矩阵的估计为: $\hat{\mathbf{A}} =$

$$\begin{bmatrix} -0.6421 & 1.2202 & 0.1911 \\ 0.3147 & -0.4917 & 1.4986 \\ -1.5956 & 0.2533 & 0.9694 \end{bmatrix}, \text{因此估计的混叠矩阵}$$

$\hat{\mathbf{A}}$ 也是比较准确的。最后,由第4节的式(20)源信号恢复算法恢复源信号,通过最后的源信号修正算法后,按照式(22)分别计算3个恢复信号的误码率均为0。

6 结束语

本文提出了一种新的针对BPSK数字信号盲分离的算法,此算法根据BPSK调制信号的特点,仅仅由得到的观测信号来估计源信号的个数,以及混叠矩阵,并在文中证明了估计出来的混叠矩阵的正确性,利用它即可完整地恢复源信号。仿真实验表明了此算法在估计混叠矩阵上的精确性,而且此算法实现也很简单,计算量少,从而说明了此算法在分离数字信号上的优异性能。最后不难发现,本文算法针对所有对称调制信号同样有出色的效果,而目前针对数字信号或有限字符集信号盲分离的研究仍然是一个开放性的问题。

参考文献

- [1] Naini F M, Mohimani G H, Zadeh M B, and Jutten C. Estimating the mixing matrix in sparse component analysis(SCA) based on partial k-dimensional subspace clustering[J]. *Neurocomputing*, 2008, 71(10): 2330-2343.
- [2] Tichavsky P, Koldovsky Z, Yeredor A, Gomez-herrero G, and Doron E. A hybrid technique for blind separation of non-gaussian and time-correlated sources using a multicomponent approach[J]. *IEEE Transactions on Neural Networks*, 2008, 19(3): 421-430.
- [3] Sun T Y, Liu C C, Hsieh S T, and Tsai S J. Blind separation with unknown number of sources based on auto-trimmed neural network[J]. *Neurocomputing*, 2008,

- 71(10): 2271–2280.
- [4] He Z S, Xie S L, Ding S X, and Cichocki A. Convolutional blind source separation in the frequency domain based on sparse representation[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2007, 15(5): 1551–1563.
- [5] Lin Q H, Yin F L, Mei T M, and Liang H. A blind source separation-based method for multiple images encryption[J]. *Image and Vision Computing*, 2008, 26(6): 788–798.
- [6] He Z S, Xie S L, Zhang L Q, and Cichocki A. A note on Lewicki-Sejnowski gradient for learning overcomplete representations[J]. *Neural Computation*, 2008, 20(3): 636–643.
- [7] Hyvarinen A and Oja E. Independent component analysis: algorithms and applications[J]. *Neural Networks*, 2000, 13(5): 411–430.
- [8] Anand K, Mathew G, and Reddy V U. Blind separation of multiple co-channel BPSK signals arriving at an antenna array[J]. *IEEE Signal Processing Letters*, 1995, 2(9): 176–178.
- [9] Lee C C and Lee J H. An efficient method for blind digital signal separation of array data[J]. *Signal Processing*, 1999, 77(2): 229–234.
- [10] Li Y Q, Cichocki A, and Zhang L Q. Blind separation and extraction of binary sources[J]. *Communication and Computer Sciences*, 2003, 86(3): 580–590.
- [11] Alle J and Veen V D. Analytical method for blind binary signal separation[J]. *IEEE Transactions on Signal Processing*, 1997, 45(4): 1078–1082.
- [12] Talwar S, Viberg M, and Paulraj A. Blind estimation of synchronous co-channel digital signals using an antenna array. Part I: algorithms[J]. *IEEE Transactions on Signal Processing*. 1996, 44(5): 1184–1197.
- [13] Swindlehurst A and Yang J. Using least squares to improve blind signal copy performance[J]. *IEEE Signal Processing Letters*, 1994, 1(5): 80–82.
- [14] Hansen L K and Xu G. A hyperplane-based algorithm for the digital co-channel communications problem[J]. *IEEE Transactions on Information Theory*, 1997, 43(5): 1536–1548.
- [15] Li Q Y, Bai E W W, and Ding Z. Blind source separation of signals with known alphabets using ϵ -approximation algorithms[J]. *IEEE Transactions on Signal Processing*, 2003, 51(1): 1–10.
- [16] Wong C C and Chen C C. A hybrid clustering and gradient descent approach fuzzy modeling[J]. *IEEE Transactions on System, Man, and Cybernetics-Part B*, 1999, 29(6): 686–693.
- 谭北海: 男, 1980 年生, 讲师, 博士后, 研究方向为信号盲分离及自适应信号处理.
- 赵敏: 男, 1970 年生, 工程师, 博士, 研究方向为智能信息处理、盲信号分离.
- 谢胜利: 男, 1958 年生, 教授, 博士生导师, 主要研究方向为非线性系统的稳定性分析、盲信号分离及自适应信号处理.