

DOI: 10.7652/xjtuxb201612017

利用参数结构的快速非酉联合对角化算法

刘文娟, 冯大政, 袁明冬

(西安电子科技大学雷达信号处理国家重点实验室, 710071, 西安)

摘要: 针对基于快速 Frobenius 范数对角化(FFDIAG)的盲信号分离算法不能直接处理复数数据从而导致分离性能差的问题,提出一种利用参数结构的快速非酉联合对角化(PSJD)算法。该算法首先将由观测信号的统计量得到的复目标矩阵转化为实对称矩阵;通过对代价函数的二阶近似,将解联合对角化问题转化为一系列的线性最小二乘问题,直接得到更新矩阵元素的估计。在每次迭代中,通过充分利用转化后的目标矩阵的结构信息,减少估计分离矩阵及更新目标矩阵的计算复杂度。同时,针对 FFDIAG 算法采用的固定步长难以兼顾收敛速度与更新矩阵严格对角占优性的问题,采用仅由当前更新矩阵的估计值决定的自适应学习率,提高算法的收敛性能。仿真实验表明,在一定的取值范围内,PSJD 算法的收敛速度对步长参数的变化不敏感,在步长参数同为 0.1 的情况下,PSJD 算法达到收敛所需的迭代次数比采用固定步长的算法减少了 42% 左右。

关键词: 盲信号分离;联合对角化;目标矩阵;自适应学习率

中图分类号: TN911.7 **文献标志码:** A **文章编号:** 0253-987X(2016)12-0106-08

A Fast Non-Unitary Joint Diagonalization Algorithm Based on Utilizations of Parametric Structures

LIU Wenjuan, FENG Dazheng, YUAN Mingdong

(National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China)

Abstract: A parametric structures based fast joint diagonalization (PSJD) algorithm for non-unitary diagonalization of a set of complex target matrices is presented to cope with the problem that the blind source separation by fast Frobenius diagonalization (FFDIAG) algorithm is not applicable in the complex-valued space and its separation performance is lower. The algorithm firstly transforms the complex target matrices into real-symmetric ones. Secondly, the problem of simultaneous diagonalization of matrices is transformed into a series of linear least-squares problems through second-order approximation to contract functions, and the elements of the updating matrix are directly estimated. The computational complexity for estimating the diagonalizer and for updating the target matrices is significantly reduced by making full use of the structure information of the transformed target matrices. In order to overcome the drawback of fixed step size adopted in the FFDIAG that may not strike a balance between the convergence rate and strictly diagonally dominant property of the update matrix, the proposed algorithm uses the adaptive learning rate determined from the estimation of the update matrix in each iteration to improve the convergence property. Results of numerical simulations show that the convergence

收稿日期: 2016-03-08。 作者简介: 刘文娟(1986—),女,博士生;冯大政(通信作者),男,教授,博士生导师。 基金项目: 国家自然科学基金资助项目(61271293,61373177)。

网络出版时间: 2016-10-10

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20161010.1743.002.html>

rate of PSJD algorithm is not very sensitive in a wide range of step-size values. When the step size is 0.1, the number of iterations required to reach convergence is 42% less than that of the fixed step-size method.

Keywords: blind source separation; joint diagonalization; target matrix; adaptive learning rate

盲信号分离(Blind Source Separation, BSS)由于其数学模型具有一般性,可适用于大多数观测数据,同时不需要训练序列或过多的先验信息,仅利用观测信号就可以有效恢复源信号,因此在电子侦察、无线通信、语音信号处理、图像处理及生物医学等领域具有广阔的应用前景^[1]。基于联合对角化的 BSS 算法以其分离精度高、计算复杂度低而受到广泛关注。根据表征联合对角化近似程度的代价函数的不同,现有的联合对角化算法主要分为两大类:①利用目标矩阵非对角部分的 Frobenius 范数作为代价函数^[2-5],最小化此代价函数直接得到分离矩阵的估计;②利用最小二乘代价函数^[6-8]得到混迭矩阵的估计。这两类代价函数均为待估计的矩阵的四次函数,计算复杂度高。相对于基于 Frobenius 范数代价函数的算法,基于最小二乘代价函数的算法不仅需要估计混迭矩阵,还需要估计一组对角矩阵,而这些对角矩阵的取值通常是盲信号处理不关心的,从而导致待估计的未知参数数目增加;另一方面,当估计的混迭矩阵的条件数较大时,矩阵求逆运算将会放大其估计误差,从而影响算法性能。

在现有的基于 Frobenius 范数的非正交联合对角化算法中,Ziehe 等提出的快速 Frobenius 对角化(Fast Frobenius Diagonalization, FFDIAG)算法^[3]采用乘性迭代机制更新分离矩阵,在每次迭代中,利用更新矩阵及更新后的目标矩阵的非对角线元素很小的假设,对代价函数进行合理近似,将其转化为一个最小二乘问题,同时,利用矩阵的稀疏性直接计算更新矩阵的元素,避免矩阵求逆运算,该算法因其收敛性能好、计算复杂度低而受到广泛关注,但是其只能应用于实数域,无法直接推广至复数域。徐先锋等提出的复值快速 Frobenius 对角化(Complex-Valued Fast Frobenius Diagonalization, CVFFDIAG)算法^[4]利用对未知参数的特殊表述,分开求解复未知参数的实部和虚部,将文献[3]算法推广至复数域,并利用 Hessian 矩阵的块对角结构,降低算法的计算复杂度。但是,该算法同 FFDIAG 一样,需要人为地选择一个步长因子 $\theta(0 < \theta < 1)$,使分离矩阵满足严格对角占优条件,保证分离矩阵可逆性。这一学习规则主要存在如下缺点: θ 过小,

算法收敛速度会降低; θ 过大,则可能无法保证严格对角占优条件。

本文针对复数域联合对角化问题,提出一种利用参数结构的快速非酉联合对角化(PSJD)算法。通过将复矩阵转化为实对称矩阵,该算法将 FFDIAG 推广至复数领域,利用矩阵的特殊结构,提高算法性能、降低算法的计算复杂度,实现复目标矩阵组的联合对角化,同时该算法采用自适应学习率,避免固定的步长因子存在的问题。

1 矩阵结构信息

简要介绍复矩阵转化为实对称矩阵的方法^[2],并详细分析转化后的矩阵的结构,为新算法的提出做准备。

1.1 复矩阵转化为实矩阵的方法

考虑 K 个维数为 $M \times M$ 的目标矩阵集合 $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_K\}$,其中第 k 个目标矩阵 $\mathbf{R}_k$ 满足如下可联合对角化结构

$$\mathbf{R}_k = \mathbf{A} \mathbf{D}_k \mathbf{A}^H, \quad k = 1, 2, \dots, K \quad (1)$$

式中: $\mathbf{A}$ 为可逆的混迭矩阵; $\mathbf{D}_k$ 为复对角矩阵。对于目标矩阵 $\mathbf{R}_k$,首先将其转化为 $2K$ 个 Hermitian 矩阵^[2]

$$\begin{aligned} \mathbf{R}'_{2k-1} &= (\mathbf{R}_k + \mathbf{R}_k^H)/2 = \mathbf{A} \operatorname{Re}(\mathbf{D}_k) \mathbf{A}^H \\ \mathbf{R}'_{2k} &= (\mathbf{R}_k - \mathbf{R}_k^H)/2i = \mathbf{A} \operatorname{Im}(\mathbf{D}_k) \mathbf{A}^H \end{aligned} \quad (2)$$

式中: $\operatorname{Re}(\cdot)$ 和 $\operatorname{Im}(\cdot)$ 分别表示取实部和虚部; $i = (-1)^{1/2}$ 。其次,将这 $2K$ 个维数为 $M \times M$ 的 Hermitian 矩阵转化为 $2M \times 2M$ 的实矩阵^[2]

$$\begin{aligned} \bar{\mathbf{R}}_k &= f(\mathbf{R}'_k) = \begin{bmatrix} \operatorname{Re}(\mathbf{R}'_k) & \operatorname{Im}(\mathbf{R}'_k) \\ -\operatorname{Im}(\mathbf{R}'_k) & \operatorname{Re}(\mathbf{R}'_k) \end{bmatrix} \\ k &= 1, 2, \dots, 2K \end{aligned} \quad (3)$$

$\bar{\mathbf{R}}_k$ 称之为扩展的目标矩阵。

1.2 扩展的目标矩阵的结构

从式(2)和式(3)发现,扩展的目标矩阵具有如下结构信息。

(1) $\bar{\mathbf{R}}_k$ 为具有可联合对角化结构的实对称矩阵^[2]

$$\bar{\mathbf{R}}_{2k-1} = \bar{\mathbf{A}} \begin{bmatrix} \operatorname{Re}(\mathbf{D}_k) & \mathbf{0} \\ \mathbf{0} & \operatorname{Re}(\mathbf{D}_k) \end{bmatrix} \bar{\mathbf{A}}^T \quad (4)$$

$$\bar{\mathbf{R}}_{2k} = \bar{\mathbf{A}} \begin{bmatrix} \text{Im}(\mathbf{D}_k) & \mathbf{0} \\ \mathbf{0} & \text{Im}(\mathbf{D}_k) \end{bmatrix} \bar{\mathbf{A}}^T \quad (5)$$

式中: $\bar{\mathbf{A}} = f(\mathbf{A})$ 为扩展的混迭矩阵。

(2) $\bar{\mathbf{R}}_k$ 具有块内对称结构

$$\text{Re}(\mathbf{R}'_k) = \text{Re}(\mathbf{R}'_k)^T \quad (6)$$

同时, $\bar{\mathbf{R}}_k$ 具有块内反对称结构

$$\text{Im}(\mathbf{R}'_k) = -\text{Im}(\mathbf{R}'_k)^T \quad (7)$$

(3) $\bar{\mathbf{R}}_k$ 具有块一致结构

$$R_{m,n,k} = R_{m+M,n+M,k} \quad (8)$$

$$R_{m,n+M,k} = -R_{m+M,n,k} \quad (9)$$

式中: $R_{m,n,k}$ 为 $\bar{\mathbf{R}}_k$ 的第 (m,n) 个元素, $m,n \in [1,M]$ 。

(4) $\text{Im}(\mathbf{R}'_k)$ 的对角线元素为零, 对于 $m=1, 2, \dots, M$, 有

$$R_{m,m+M,k} = R_{m+M,m,k} = 0 \quad (10)$$

2 算法描述

2.1 PSJD 算法

给定一组扩展的目标矩阵组 $\{\bar{\mathbf{R}}_1, \bar{\mathbf{R}}_2, \dots, \bar{\mathbf{R}}_{2K}\}$, 将分离矩阵分解为一系列矩阵的乘积^[3-4], 即第 t 次迭代后的分离矩阵可以表示为

$$\bar{\mathbf{V}}^{(t+1)} = (\mathbf{I} + \mathbf{W}^{(t)}) \bar{\mathbf{V}}^{(t)} \quad (11)$$

式中: $\mathbf{W}^{(t)}$ 是维数为 $2M \times 2M$ 的零对角矩阵, 上标 (t) 表示第 t 次迭代。由 Levi-Desplanques 定理^[3]可知, 给定可逆的初始值 $\bar{\mathbf{V}}^{(0)}$, 通过限制 $\mathbf{I} + \mathbf{W}^{(t)}$ 严格对角占优可确保分离矩阵 $\bar{\mathbf{V}}$ 的可逆性。令第 t 次迭代的目标矩阵更新为

$$\bar{\mathbf{R}}_k^{(t+1)} = (\mathbf{I} + \mathbf{W}^{(t)}) \bar{\mathbf{R}}_k^{(t)} (\mathbf{I} + \mathbf{W}^{(t)})^T \quad (12)$$

此时, 联合对角化问题可以描述为寻找一个更新矩阵使得矩阵 $\bar{\mathbf{R}}_k^{(t+1)}$ 尽可能地对角化。令 $\bar{\mathbf{R}}_k^{(t)} = \mathbf{E}_k^{(t)} + \mathbf{D}_k^{(t)}$, 其中 $\mathbf{D}_k^{(t)}$ 为 $\bar{\mathbf{R}}_k^{(t)}$ 对角线上的元素构成的对角矩阵, $\mathbf{E}_k^{(t)} = \bar{\mathbf{R}}_k^{(t)} - \mathbf{D}_k^{(t)}$ 。假设 $\mathbf{W}^{(t)}$ 和 $\mathbf{E}_k^{(t)}$ 的范数很小, 式(12)中的二次项可以忽略, 第 t 次迭代后的目标矩阵可近似为

$$\bar{\mathbf{R}}_k^{(t+1)} \approx \mathbf{W} \mathbf{D}_k^{(t)} + \mathbf{D}_k^{(t)} \mathbf{W}^T + \mathbf{D}_k^{(t)} + \mathbf{E}_k^{(t)} \quad (13)$$

为推导简便, 在不引起歧义的情况下, 省略表示迭代次数的上标 (t) 和 $(t+1)$, 代价函数可以表示为

$$C(\mathbf{W}) = \sum_{k=1}^{2K} \text{off}(\mathbf{W} \mathbf{D}_k + \mathbf{D}_k \mathbf{W}^T + \mathbf{E}_k) = \sum_{k=1}^{2K} \sum_{m=1}^{2M-1} \sum_{n=m+1}^{2M} |\omega_{m,n} d_{n,k} + \omega_{n,m} d_{m,k} + e_{m,n,k}|^2 \quad (14)$$

式中: $\text{off}(\cdot)$ 表示所有非对角线元素的平方和; $\omega_{m,n}$ 和 $e_{m,n,k}$ 分别表示 $\mathbf{W}$ 和 $\mathbf{E}_k$ 的第 (m,n) 个元素。定义子函数

$$J(\omega_{m,n}, \omega_{n,m}) =$$

$$\sum_{k=1}^{2K} |\omega_{m,n} d_{n,k} + \omega_{n,m} d_{m,k} + e_{m,n,k}|^2 \quad (15)$$

可以看出, 式(15)仅与元素 $\omega_{m,n}$ 和 $\omega_{n,m}$ 有关, 所以, 代价函数(14)可以分解为 $M(2M-1)$ 个相互独立的子代价函数。由于扩展的目标矩阵组为 $2K$ 个 $2M \times 2M$ 维的矩阵, 直接执行 FFDIAG 算法的运算量较大, 下面将充分利用扩展的目标矩阵的结构信息, 降低算法的计算复杂度。

由矩阵的结构信息可知, 为使扩展的目标矩阵组在迭代过程中保持块一致结构, 分离矩阵必须具有块一致结构, 因此待估计的更新矩阵参数仅为 $2M(M-1)$ 个。为充分利用上述结构, 将更新矩阵分解为 4 个 $M \times M$ 维的子块独立地进行处理, 具体步骤如下。

(1) 假设 $1 \leq m < n \leq M$, 对于该子块中的坐标为 (m,n) 元素, 子代价函数式(15)可以表示为

$$J(\omega_{m,n}) = \sum_{k=1}^{2K} |d_{1,k}^T \mathbf{w}_{m,n} + e_{m,n,k}|^2 \quad (16)$$

式中: $\mathbf{d}_{1,k} = [d_{1,k} \ d_{m,k}]^T$; $\mathbf{w}_{m,n} = [\omega_{m,n} \ \omega_{n,m}]^T$ 为待求元素的最小二乘估计

$$\mathbf{w}_{m,n} = -\mathbf{B}_{m,n}^{-1} \mathbf{e}_{m,n} \quad (17)$$

其中 $\mathbf{B}_{m,n} = \sum_{k=1}^{2K} \mathbf{d}_{1,k} \mathbf{d}_{1,k}^T$; $\mathbf{e}_{m,n} = \sum_{k=1}^{2K} e_{m,n,k} \mathbf{d}_{1,k}$ 。

(2) 对于坐标为 $(m+M, n+M)$ 的元素, 利用块一致结构式(8)可将关于该子块内元素的代价函数表示为

$$J(\omega_{m+M, n+M}) = \sum_{k=1}^{2K} |d_{1,k}^T \mathbf{w}_{m+M, n+M} + e_{m,n,k}|^2 \quad (18)$$

该函数具有与式(16)相同的形式, 该子块中元素的估计满足

$$\mathbf{w}_{m+M, n+M} = \mathbf{w}_{m,n} \quad (19)$$

(3) 对于坐标为 $(m, n+M)$ 的元素, 根据块一致结构式(8), 关于该子块元素的优化问题式(15)可以写为

$$J(\omega_{m, n+M}) = \sum_{k=1}^{2K} |d_{1,k}^T \mathbf{w}_{m, n+M} + e_{m,n+M,k}|^2 \quad (20)$$

该子块中元素的最小二乘估计为

$$\mathbf{w}_{m, n+M} = -\mathbf{B}_{m,n+M}^{-1} \mathbf{e}_{m,n+M} \quad (21)$$

式中: $\mathbf{e}_{m,n+M} = \sum_{k=1}^{2K} e_{m,n+M,k} \mathbf{d}_{1,k}$ 。

(4) 类似地, 对于坐标为 $(m+M, n)$ 的元素, 由块一致结构式(8)和块内反对称结构式(7), 该子块中元素的子代价函数为

$$J(\mathbf{w}_{n,m+M}) = \sum_{k=1}^{2K} | \mathbf{d}_{1,k}^T \mathbf{w}_{n,m+M} - e_{m,n+M,k} |^2 \quad (22)$$

对比式(20)和(22)可知

$$\mathbf{w}_{m+M,n} = -\mathbf{w}_{m,n+M} \quad (23)$$

(5)对于坐标为 $(m, m+M)$ 的元素,由块一致结构式(8)和零对角结构式(10)可得

$$J(\mathbf{w}_{m,m+M}) = \sum_{k=1}^{2K} | \mathbf{d}_{2,k}^T \mathbf{w}_{m,m+M} |^2 \quad (24)$$

式中: $\mathbf{d}_{2,k} = [d_{m,k} \quad d_{m+M,k}]^T$ 。由式(24)可以得到 $\mathbf{w}_{m,m+M} = \mathbf{0}_{2 \times 1}$, 这使得 $f^{-1}(\mathbf{W})$ 的对角元素为 0。

2.2 自适应学习率

为保证分离矩阵的可逆性, $\mathbf{I} + f^{-1}(\mathbf{W}^{(t)})$ 必须满足严格对角占优条件, 针对引言中提到的固定步长因子 θ 可能存在的问题, 本文算法采用如下自适应学习规则: 给定扩展的更新矩阵 $\mathbf{W}^{(t)}$, 若 $\mathbf{I} + f^{-1}(\mathbf{W}^{(t)})$ 不满足严格对角占优条件, 则 $\mathbf{W}^{(t)} \leftarrow \mathbf{F}(\mathbf{W}^{(t)})\mathbf{W}^{(t)}$, 其中 $\mathbf{F}(\mathbf{W}^{(t)})$ 为对角矩阵, 其第 (m, m) 个元素为 $a_m(\mathbf{w}_m^{(t)})$, 其中 $\mathbf{w}_m^{(t)}$ 为 $\mathbf{W}^{(t)}$ 的第 m 行。 $a_m(\mathbf{w}_m^{(t)})$ 由如下自适应学习规则决定: 对于 $m, n \in [1, M]$, $\eta_m^{(t)} = \sum_{n=1}^M [(\mathbf{w}_{m,n}^{(t)})^2 + (\mathbf{w}_{m,n+M}^{(t)})^2]^{1/2}$, 其中 $\mathbf{w}_{m,n}^{(t)}$ 为 $\mathbf{W}^{(t)}$ 的第 (m, n) 个元素, 若 $\eta_m^{(t)} \geq 1$, 则 $a_m(\mathbf{w}_m^{(t)}) = a_{m+M}(\mathbf{w}_{m+M}^{(t)}) = 1/(\eta_m^{(t)} + \theta)$, $\theta \in (0, 1]$; 否则, $a_m(\mathbf{w}_m^{(t)}) = a_{m+M}(\mathbf{w}_{m+M}^{(t)}) = 1$ 。

可以看出, $\eta_m^{(t)}$ 是 $f^{-1}(\mathbf{W}^{(t)})$ 的第 m 行的 1-范数。假设 $\eta_m^{(t)} \geq 1$, 由上述规则, $f^{-1}(\mathbf{F}(\mathbf{W}^{(t)})\mathbf{W}^{(t)})$ 第 m 行的 1-范数为 $\sum_{n=1}^M a_m(\mathbf{w}_m^{(t)}) [(\mathbf{w}_{m,n}^{(t)})^2 + (\mathbf{w}_{m,n+M}^{(t)})^2]^{1/2} = a_m(\mathbf{w}_m^{(t)}) \eta_m^{(t)} = \eta_m^{(t)} / [\eta_m^{(t)} + \theta]$, 可见, 参数 $\theta > 0$ 时, $\mathbf{I} + f^{-1}(\mathbf{F}(\mathbf{W}^{(t)})\mathbf{W}^{(t)})$ 满足严格对角占优条件。这里, 参数 θ 的选择主要考虑 3 个因素: 严格对角占优条件、收敛速度和稳定性。首先, θ 必须保证严格对角占优条件, 这要求 θ 必须大于 0; 由自适应信号处理^[9]的理论可知, 步长越大(对应着 θ 越小), 则算法的收敛速度越快, 稳定性越差; 步长越小(θ 越大), 则算法的失败率越低, 稳定性越好, 但收敛速度较慢。实验发现, $\theta \in (0, 1]$ 是在收敛速度与稳定性之间的良好折中, 算法的收敛速度对 θ 的变化不敏感, 且越接近于 1 算法的性能越稳健。

2.3 算法步骤

利用 PSJD 算法实现复目标矩阵联合对角化的步骤总结如下。

输入 复目标矩阵组 $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_K\}$;

初始化 根据式(2)和(3)将复目标矩阵转化为扩展

的实矩阵; 给定可逆的初始化矩阵 $\bar{\mathbf{V}}^{(0)}$ 。

重复如下步骤, 直至算法收敛:

For $t=1, 2, \dots$

步骤 1 根据如下子步骤估计更新矩阵 $\mathbf{W}^{(t)}$:

For $m=1, 2, \dots, M-1; n=m+1, m+2, \dots, M$

(1)分别根据式(17)和(21)计算 $\mathbf{w}_{m,n}$ 和 $\mathbf{w}_{m,n+M}$;

(2)分别根据式(19)和(23)计算 $\mathbf{w}_{m+M,n+M}$ 和

$\mathbf{w}_{n,m+M}$;

(3)判断 $\mathbf{I} + f^{-1}(\mathbf{W}^{(t)})$ 是否严格对角占优, 如果不是, 则 $\mathbf{W}^{(t)} \leftarrow \mathbf{F}(\mathbf{W}^{(t)})\mathbf{W}^{(t)}$;

步骤 2 根据式(11)更新联合对角化矩阵, 根据式(12)更新目标矩阵组;

输出 分离矩阵 $\mathbf{V} = f^{-1}(\bar{\mathbf{V}})$ 。

考虑式(17)和(21)有相同的形式 $\mathbf{w} = -\mathbf{B}^{-1}\mathbf{e}$, 令 $\mathbf{B}$ 的第 (i, j) 个元素为 $B_{i,j}$ ($i, j=1, 2$), 利用

$$\mathbf{B}^{-1} = \frac{1}{B_{1,1}B_{2,2} - B_{1,2}B_{2,1}} \begin{bmatrix} B_{2,2} & -B_{1,2} \\ -B_{2,1} & B_{1,1} \end{bmatrix} \quad (25)$$

可以简化计算。

3 性能分析

3.1 计算复杂度分析

由前文的分析可知, 本文算法每次迭代(执行 1 次步骤 1 和 2)需进行 $8M(M-1)K$ 次实数乘除法运算 (Number of Multiplications and Divisions, MDN), 这里忽略了阶数低于 $O(M^3K)$ 和 $O(M^4)$ 的项。CVFFDIAG 算法、基于高斯迭代的均匀加权完全对角化 (Uniformly Weighted Exhaustive Diagonalization with Gauss Iterations, UWEDGE) 算法^[10]每次更新需进行 $8M^3K$ 次实数 MDN, 双迭代算法 (bi-iterative algorithm, BIA)^[7]每次迭代需进行 $12M^3K$ 次实数 MDN。

3.2 收敛性分析

显然, 代价函数 $C(\mathbf{V})$ 的下界存在且大于等于零。这表明代价函数至少存在一个全局最小点。众所周知, 非正交联合对角化为非线性优化问题, 因此, 严格的收敛性分析往往比较困难。通常借助于动力学理论中著名的 Lyapunov 函数和 LaSalle's 不变定理^[11-12]来分析算法的收敛性能。

定义 1 (Lyapunov 函数)^[11] 令集合 $\bar{\mathbf{U}} \in \mathbf{C}^{M \times M}$, 如果关于离散序列 $\mathbf{V}^{(t)} \in \bar{\mathbf{U}}$ 的函数 $g(\mathbf{V})$ 满足:

(1)函数 $g(\mathbf{V}^{(t)})$ 连续;

(2)对于离散点 t , 有 $g(\mathbf{V}^{(t)}) \leq g(\mathbf{V}^{(t-1)})$;

(3)对于任意有限常数 γ , $\bar{\mathbf{U}}_\gamma = \{\mathbf{V}^{(t)} | g(\mathbf{V}^{(t)}) < \gamma\}$

γ 是有界集合。

那么, $g(\mathbf{V})$ 称为 Lyapunov 函数。

定义 2(不变集)^[11] 若离散序列 $\mathbf{V}^{(t)} \in \bar{U}$, 那么集合 $\bar{U}_{\text{Invar}} = \{\mathbf{V}^{(t)} | g(\mathbf{V}^{(t)}) - g(\mathbf{V}^{(t-1)}) = 0\}$ 称为 $\mathbf{V}^{(t)}$ 的不变集合。

定理 1(LaSalle's 不变定理)^[11] 若函数 $g(\mathbf{V})$ 是关于离散序列 $\mathbf{V}^{(t)} \in \bar{U}$ 的 Lyapunov 函数, 那么离散序列 $\mathbf{V}^{(t)}$ 收敛到不变集合 $\bar{U}_{\text{Invar}}$ 。

定理 2(PSJD 算法收敛性定理) 给定目标矩阵组 $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{2K}\}$, 通过 PSJD 算法估计代价函数 $C(\mathbf{V})$ 的最小值, 所得到的离散时间序列 $\mathbf{V}^{(t)}, t=1, 2, \dots$ 收敛到代价函数的不变集 $U_{\text{Invar}} = \{\mathbf{V}^{(t)} | C(\mathbf{V}^{(t)}) - C(\mathbf{V}^{(t-1)}) = 0\}$ 。对定理 2 的证明如下:

(1) 由于代价函数 $C(\mathbf{V})$ 是可微分的, 因此是连续的;

(2) PSJD 算法中关于 $\mathbf{W}^{(t)}$ 的估计为典型的最小二乘解, 因此可以得出在任意离散时间 $t, C(\mathbf{V}^{(t+1)}) \leq C(\mathbf{V}^{(t)})$ 成立^[3];

(3) 由于算法约束更新矩阵 $\mathbf{W}^{(t)}$ 满足 $\mathbf{I} + \mathbf{W}^{(t)}$ 严格对角占优, 更新矩阵的每一个行矢量需满足 $\sum_{n=1, n \neq m}^M |\mathbf{w}_{m,n}^{(t)}| < 1$, 这意味着每次迭代中, 联合对角化矩阵的估计被限制在一个等半径的超球面内。因此, 存在一个大于 0 的常数 γ , 使得集合 $U_\gamma = \{\mathbf{V} | C(\mathbf{V}) < \gamma\}$ 是有界的。

上述 3 点分析表明, 代价函数是 Lyapunov 函数。根据定理 1 可知, 离散序列 $\mathbf{V}^{(t)}$ 收敛到 $C(\mathbf{V})$ 的不变集 U_{Invar} 。

综上分析, PSJD 算法可渐进收敛到全局或至少是局部最小点。

4 仿真实验

通过与 CVFFDIAG 算法^[4]、BIA 算法^[7] 和 UWEDGE 算法^[10] 的比较, 详细评估 PSJD 算法的性能。为了便于比较, 采用盲信号分离算法中常用的性能参数——全局拒绝水平 (Global Rejection Level, GRL)^[2-7,9,12] 来衡量算法的有效性

$$L_{\text{GR}} = \sum_{m=1}^M \left(\sum_{n=1}^M |G_{m,n}|^2 / \max_l |G_{m,l}|^2 - 1 \right) + \sum_{m=1}^M \left(\sum_{n=1}^M |G_{m,n}|^2 / \max_l |G_{l,n}|^2 - 1 \right) \quad (26)$$

式中: $G_{m,n}$ 为全局传输矩阵 $\mathbf{G} = \mathbf{V}\mathbf{A}$ 的第 (m, n) 个元素。全局拒绝水平 L_{GR} 为一个非负的参数, 描述了 $\mathbf{G}$ 与广义置换矩阵 $\mathbf{E} = \mathbf{D}\mathbf{P}$ 的接近程度, 其中 $\mathbf{D}$ 为

非奇异的对角阵 (表示尺度不确定性), $\mathbf{P}$ 为置换矩阵 (表示排列不确定性)。 L_{GR} 越小, 表示 $\mathbf{G}$ 越接近于广义置换矩阵, 分离性能越好。为了便于比较, 在所有的实验中, 若满足 $\|\mathbf{V}^{(t)} - \mathbf{V}^{(t-1)}\|_F < \epsilon$ 则认为算法收敛, 停止迭代, 这里 $\mathbf{V}^{(t)}$ 表示第 t 次迭代中分离矩阵的估计值。每个算法的最大迭代 (扫描) 次数设为 500 次, 收敛门限 $\epsilon = 10^{-8}$ 。算法运行环境: MATLAB R2010a, Pentium(R) Dual-Core E5200 CPU @2.50 GHz, 2 GB 内存。

下面通过两组仿真实验比较各算法的性能。

实验 1 采用一组可近似联合对角化的矩阵来验证本文算法的性能, 包括收敛 L_{GR} 、收敛时间及收敛速度。在每次独立实验中, 构造 K 个维数为 $M \times M$ 的目标矩阵 $\mathbf{R}_k = \mathbf{A}\mathbf{D}_k\mathbf{A}^H + \sigma\Delta\mathbf{R}_k$, 其中 σ 为控制噪声扰动参数, 用无误差矩阵项和误差矩阵项 F-范数的平方比 $R_{\text{NE}} = 10\lg(\|\mathbf{A}\mathbf{D}_k\mathbf{A}^H\|_F^2 / \|\sigma\Delta\mathbf{R}_k\|_F^2)$ 来衡量噪声扰动的大小。令 $M=10, K=11$, 混迭矩阵 $\mathbf{A}$ 、对角矩阵 $\mathbf{D}_k$ 和误差矩阵 $\Delta\mathbf{R}_k$ 的实部和虚部均为随机产生, 且服从 $N(0, 1)$ 正态分布。BIA 算法的初值由 ESPRIT 算法^[7] 给定, 其余算法的初值设定为单位阵。

为了评估参数选择对算法性能的影响, 在 $R_{\text{NE}} = 0$ dB 的条件下进行了 100 次独立实验, 得到 PSJD 算法和 CVFFDIAG 算法收敛所需的平均迭代次数随参数 θ 的变化曲线, 如图 1 所示。可以看出, 当 θ 为 0.1、矩阵维数 M 分别为 6、10、12、14 时, PSJD 算法所需迭代次数比 CVFFDIAG 算法分别减少约 27%、42%、41% 和 39%, 且参数 θ 越接近于 1, 算法的性能越稳健。图 2 表示在 $R_{\text{NE}} = 0$ dB 时 PSJD 算法 ($\theta=0.1$) 以及 CVFFDIAG 算法取不同 θ 时的 GRL 随迭代次数变化的曲线。可以看出, θ 的取值较大时, 虽然收敛速度相对较快, 但收敛状态不稳定, 其原因主要是在迭代的初始阶段, 更新矩阵及目标矩阵的非对角元素不满足范数很小的假设, 较

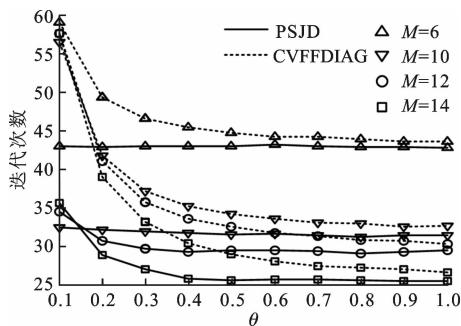


图 1 参数 θ 对收敛所需迭代次数的影响

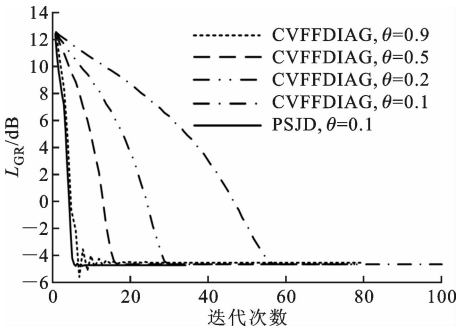
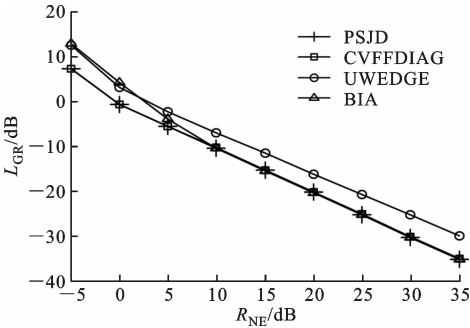


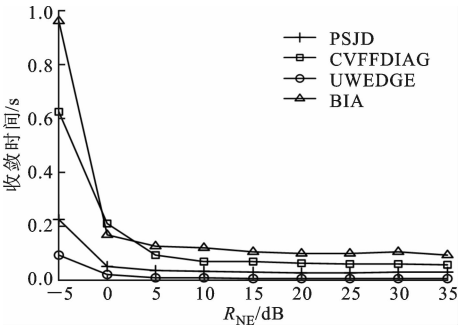
图 2 步长因子不同时 2 种算法的全局拒绝水平随迭代次数变化的曲线

大的 θ 不能保证更新矩阵满足严格对角占优条件;而自适应学习规则分开考虑分离矩阵的每一行,仅对不满足条件的行赋予较小的步长,在确保更新矩阵满足严格对角占优条件的同时,保持较快的收敛速度,因而在收敛速度与收敛状态之间取得了一个很好的平衡。

图 3 表示经过 100 次独立实验 PSJD 算法($\theta=0.17$)、CVFFDIAG 算法($\theta=0.9$)、UWEDGE 算法和 BIA 算法的收敛 L_{GR} 和收敛时间随 R_{NE} 变化的曲线。相应地,图 4 给出 $R_{NE}=0$ dB 时, L_{GR} 随迭代次数变化的曲线。可以看出,PSJD 算法具有与 CVFFDIAG 算法相似的收敛性能,且优于其他 2 个算法,PSJD 算法的收敛所需的平均时间低于 CVFFDIAG



(a) 收敛 L_{GR} 随 R_{NE} 变化的曲线



(b) 收敛所需时间随 R_{NE} 变化的曲线

图 3 R_{NE} 对 4 种算法性能的影响

算法和 BIA 算法,高于 UWEDGE 算法,但平均 L_{GR} 较 UWEDGE 算法降低了 5 dB 左右,与 CVFFDIAG 算法相比,得到相同的 L_{GR} 所需时间更短。由 4.1 节可知, CVFFDIAG 算法与 PSJD 算法每次迭代所需的 MDN 相同,图 3b 可间接说明 PSJD 算法比 CVFFDIAG 算法的平均收敛速度更快。UWEDGE 算法的参数更新是以向量或矩阵的形式进行的,避免了大量的循环迭代,因此在 MATLAB 环境下效率更高,而 BIA 算法在两组参数间交替迭代,计算复杂度最高,收敛速度最慢。

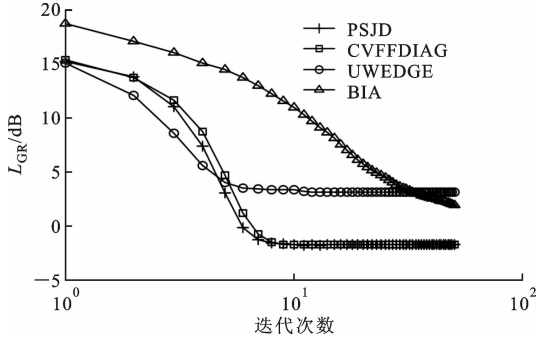


图 4 4 种算法的全局拒绝水平随迭代次数变化的曲线

实验 2 采用一组复信号测试所提算法应用于盲信号分离的性能。源信号为 4 个 0 均值的相互独立的复信号

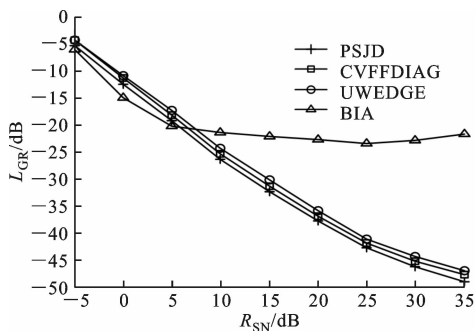
$$\left. \begin{aligned} s_1(t) &= \sin(20\pi t)\sin(600\pi t) + i\cos(20\pi t)\cos(600\pi t) \\ s_2(t) &= \text{sgn}(\sin(310\pi t)) + i\text{sgn}(\cos(100\pi t)) \\ s_3(t) &= \sin(180\pi t) + i\sin(400\pi t) \\ s_4(t) &= \sin(600\pi t + 6\cos(120\pi t)) + i\cos(900\pi t) \end{aligned} \right\} \quad (27)$$

采样点数为 1 000,接收阵列为 7 个阵元构成的均匀线阵,阵元间距半波长,即混迭矩阵 $\mathbf{A} = [\mathbf{a}(\varphi_1), \mathbf{a}(\varphi_2), \dots, \mathbf{a}(\varphi_4)]$, $\mathbf{a}(\varphi) = [1, e^{-i\pi\cos\varphi}, \dots, e^{-i6\pi\cos\varphi}]^T$, 各信号的入射角度分别为 $\varphi_1 = 20 + 10\gamma$, $\varphi_2 = 50 + 15\gamma$, $\varphi_3 = 75 + 10\gamma$, $\varphi_4 = 115 + 15\gamma$, 其中, $\gamma \in [0, 1]$ 随机产生。取时延为 $3k$ ($k=1, 2, \dots, 15$) 构造目标矩阵组,由于混迭矩阵为高矩阵,因此需要一个白化降维预操作。为衡量算法的分离性能,采用估计信号的信干噪比来表征分离信号的独立性,信干噪比为

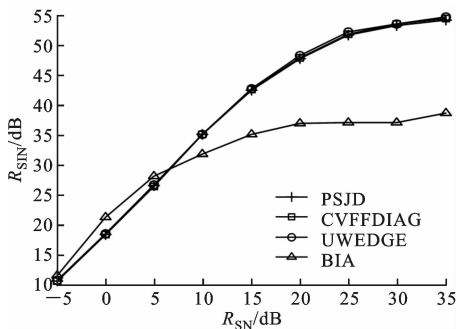
$$R_{\text{SIN}} = \frac{1}{M} \sum_{m=1}^M |\mathbf{s}_m \hat{\mathbf{s}}_m^H|^2 / \left(\sum_{n=1, n \neq m}^M \sum_{m=1}^M |\mathbf{s}_m \hat{\mathbf{s}}_n^H|^2 \right) \quad (28)$$

式中: $\mathbf{s}_m$ 和 $\hat{\mathbf{s}}_m$ 分别表示第 m 路源信号及分离信号。

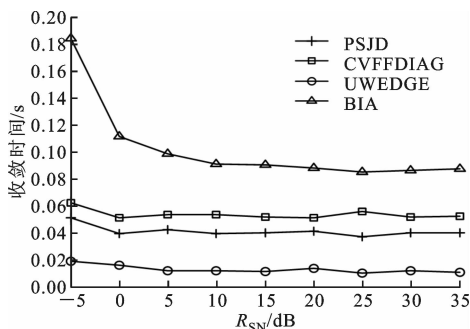
图 5 为经过 100 次独立实验,4 种算法的平均性能参数随信噪比 R_{SN} 变化的曲线。可以看出,由于预白化的影响,4 种算法的分离性能相近,但



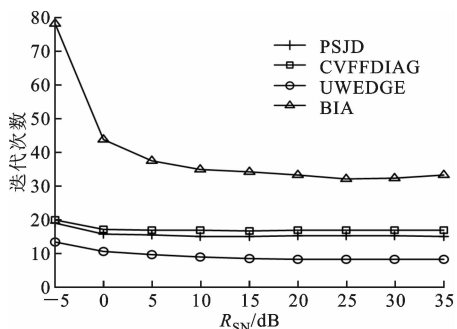
(a) 收敛 L_{GR} 随 R_{SN} 变化的曲线



(b) 收敛信干噪比随 R_{SN} 变化的曲线



(c) 收敛时间随 R_{SN} 变化的曲线



(d) 迭代次数随 R_{SN} 变化的曲线

图 5 信噪比对 4 种算法性能参数的影响

PSJD 算法 ($\theta=0.1$) 与 CVFFDIAG 算法 ($\theta=0.9$)、BIA 算法相比,得到相近的 L_{GR} 所需的迭代次数和时间较少。这是因为即使 $\mathbf{W}^{(t)}$ 的某些行不满足严格对角占优条件,CVFFDIAG 算法所用的固定步长都

对 $\mathbf{W}^{(t)}$ 所有行进行调节,这将使 $\mathbf{W}^{(t)}$ 相对于其最优值产生较大的偏离,从而影响收敛速度;而 PSJD 算法采用的自适应学习率对 $\mathbf{W}^{(t)}$ 的每行分开考虑,某行的 1-范数越小则说明经过 $t-1$ 次迭代,该行对应的分离信号与源信号越接近,因此无需对该行进行处理,从而提高了算法的收敛速度。同时,这也是 $\|\mathbf{W}^{(t)}\|_F < \epsilon$ 可以作为此算法收敛条件的原因。图 6 表示 $R_{SN}=25$ dB 时,源信号和对应的分离信号的星座图,可见所提算法能够有效地恢复源信号。

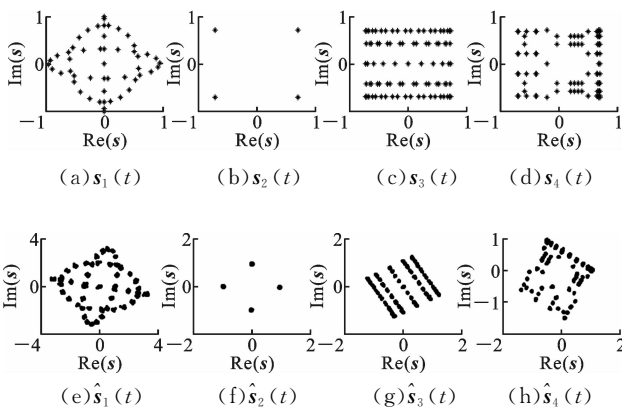


图 6 源信号、分离信号的星座图

5 结 论

本文通过将复矩阵转化为实对称矩阵,将 FF-DIAG 算法推广至复数域,利用转化后的目标矩阵的结构信息,降低算法的计算复杂度。同时,采用自适应学习规则,改善了在初始迭代状态中,更新矩阵和目标矩阵的非对角部分不满足假设条件时,采用固定步长因子无法确保严格对角占优性而导致收敛状态不稳定的情况,确保了分离矩阵的可逆性,且避免了收敛速度对参数选择的依赖,提高了算法的收敛性能。仿真实验表明,同 CVFFDIAG 算法相比,本文所提 PSJD 算法达到相同的全局拒绝水平所需的运算时间更短。

参考文献:

[1] CHABRIEL G, KLEINSTEUBER M, MOREAU E, et al. Joint matrices decompositions and blind source separation: a survey of methods, identification, and applications [J]. IEEE Signal Processing Magazine, 2014, 3(31): 34-43.

[2] MESLOUB A, ABED-MERAİM K, BELOUCHRANI A. A new algorithm for complex non orthogonal joint diagonalization based on Shear and Givens rotations [J]. IEEE Transactions on Signal Processing, 2014, 62(8): 1913-1925.

- [3] ZIEHE A, LASKOV P, NOLTE G, et al. A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation [J]. *Journal of Machine Learning Research*, 2004, 5(3): 777-800.
- [4] XU Xianfeng, FENG Dazheng, ZHENG Weixing. A fast algorithm for nonunitary joint diagonalization and its application to blind source separation [J]. *IEEE Transactions on Signal Processing*, 2011, 59(7): 3457-3463.
- [5] 丛丰裕, 雷菊阳, 许海翔, 等. 在线增强型复值混合信号盲分离算法研究 [J]. *西安交通大学学报*, 2006, 40(9): 1070-1073.
CONG Fengyu, LEI Juyang, XU Haixiang, et al. On-line blind separation of enforced mixed complex-value signal sources [J]. *Journal of Xi'an Jiaotong University*, 2006, 40(9): 1070-1073.
- [6] 聂卫科, 冯大政, 刘建强. 二维波达方向估计的非酉联合对角化方法 [J]. *西安交通大学学报*, 2008, 42(6): 747-750.
NIE Weike, FENG Dazheng, LIU Jiangiang. Non-unitary joint diagonalization method for estimating two-dimension direction of arrival [J]. *Journal of Xi'an Jiaotong University*, 2008, 42(6): 747-750.
- [7] FENG Dazheng, ZHANG Hua, ZHENG Weixing. Bi-iterative algorithm for extracting independent components from array signals [J]. *IEEE Transactions on Signal Processing*, 2011, 59(8): 3636-3646.
- [8] ZENG Tiaojun, FENG Quanyuan. Non-orthogonal joint diagonalization algorithm based on hybrid trust region method and its application to blind source separation [J]. *Neurocomputing*, 2014, 133(8): 280-294.
- [9] 陆建涛, 成玮, 訾艳阳, 等. 变步长等变自适应盲源分离算法 [J]. *西安交通大学学报*, 2015, 49(12): 83-89.
LU Jiantao, CHENG Wei, ZI Yanyang, et al. Variable step-size algorithm for equivariant adaptive separation via independent [J]. *Journal of Xi'an Jiaotong University*, 2015, 49(12): 83-89.
- [10] TICHAVSKY P, YEREDOR A. Fast approximate joint diagonalization incorporating weight matrices [J]. *IEEE Transactions on Signal Processing*, 2009, 57(3): 878-891.
- [11] LASALLE J P. The stability of dynamical system [M]. Philadelphia, PA, USA: SIAM Press, 1976: 49-50.
- [12] 张伟涛, 楼顺天, 张延良. 非对称非正交快速联合对角化算法 [J]. *自动化学报*, 2010, 36(6): 829-836.

ZHANG Weitao, LOU Shuntian, ZHANG Yanliang. Non-symmetrical non-orthogonal fast joint diagonalization algorithm [J]. *Acta Automatica Sinica*, 2010, 36(6): 829-836.

[本刊相关文献链接]

王玉玺, 黄国策, 李伟, 等. 非均匀广义对角加载稳健波束形成算法. 2016, 50(8): 64-69. [doi:10.7652/xjtuxb201608011]

刘健伶, 陈志刚, 王磊. 一种自适应广义空间调制及其低复杂度算法. 2016, 50(4): 48-53. [doi:10.7652/xjtuxb201604008]

王富平, 水鹏朗. 采用多尺度方向微分比率的角点检测算法. 2016, 50(4): 68-75. [doi:10.7652/xjtuxb201604011]

赵凯, 王闯, 李尊朝, 等. 结合平衡和滤波技术抑制 GaN 电源转换器的电磁干扰. 2016, 50(02): 38-42. [doi:10.7652/xjtuxb201602007]

孙黎, 徐洪斌. 协作式终端直通系统中星座旋转辅助的干扰避免策略. 2015, 49(12): 6-11. [doi:10.7652/xjtuxb201512002]

张东伟, 郭英, 齐子森, 等. 采用空间极化时频分布的跳频信号多参数联合估计算法. 2015, 49(8): 17-23. [doi:10.7652/xjtuxb201508004]

巴斌, 郑娜娥, 朱世磊, 等. 利用蒙特卡罗的最大似然时延估计算法. 2015, 49(8): 24-30. [doi:10.7652/xjtuxb201508005]

吴一全, 孟天亮, 吴诗姘. 人工蜂群优化的非下采样 Shearlet 域引导滤波图像增强. 2015, 49(6): 39-45. [doi:10.7652/xjtuxb201506007]

熊涛, 江桦, 崔鹏辉, 等. 应用基扩展模型的混合信号单通道盲分离算法. 2015, 49(6): 60-66. [doi:10.7652/xjtuxb201506010]

郝爱洁, 齐春. 一种鲁棒的稀疏信号重构算法. 2015, 49(4): 98-103. [doi:10.7652/xjtuxb201504016]

王云龙, 吴瑛. 联合时延与多普勒频率的直接定位改进算法. 2015, 49(4): 123-129. [doi:10.7652/xjtuxb201504020]

尚佳栋, 王祖林, 周丽娜, 等. 采用随机共振增强的混合扩频信号跳频参数估计. 2014, 48(10): 42-48. [doi:10.7652/xjtuxb201410007]

储颖, 牟轩沁, 洪伟. 采用形状一致性特征的盲图像质量评价方法. 2014, 48(8): 12-17. [doi:10.7652/xjtuxb201408003]

吴仁斌, 姚敏立, 贾维敏, 等. 采用幅度响应约束的鲁棒自适应波束形成算法. 2014, 48(4): 109-114. [doi:10.7652/xjtuxb201404019]

郝红侠, 刘芳, 焦李成, 等. 采用结构自适应窗的非局部均值图像去噪算法. 2013, 47(12): 71-76. [doi:10.7652/xjtuxb201312013]

穆为磊, 高建民, 王昭, 等. 考虑人眼视觉特性的射线检测数字图像质量评价方法. 2013, 47(7): 91-95. [doi:10.7652/xjtuxb201307017]

(编辑 刘杨)