Low-Complexity Method of Weighted Subspace Fitting for Direction Estimation

Lei Huang, Shunjun Wu and Linrang Zhang
National Key Lab of Radar Signal Processing, Xidian University, Xi' an, Shaanxi 710071, P. R. China

Key Words: Array signal processing, direction of arrival, weighted subspace fitting, low complexity, multi-stage wiener filter.

SUMMERY

In this paper, we consider a low-complexity method of weighted subspace fitting (WSF) for direction-of-arrival (DOA) estimation. With the properties of the multi-stage wiener filter (MSWF), we derive a novel criterion function for the WSF method without the estimate of an array covariance matrix and its eigendecomposition. A new approach for noise variance estimation is also proposed. Numerical results indicate that by selecting a specific weighting matrix, the low-complexity WSF estimator can provide the comparable estimation performance with the conventional WSF method.

I. *INTRODUCTION*

Super-resolving correlated or even coherent signals is the fundamental problem in array signal processing, which is frequently encountered in many areas such as communication, radar, sonar and geophysical seismology. The subspace based methods, which resort to the decomposition of the observation space into signal subspace and noise subspace, can provide the outstanding estimation performance. In the literature, the classical subspace based methods have been investigated extensively [1]-[2]. Nevertheless, the classical subspace based methods involve the estimate of an array covariance matrix and its eigenvalue decomposition (EVD), which is rather computationally intensive for the case where the model orders in these matrices are large. To reduce the computational complexity of the classical subspace based methods, a number of low-complexity methods without eigendecomposition have been proposed in [3]-[4]. Normally, the existing linear operation based methods with low complexity [3]-[6] find the signal or noise subspace by a partition of array response matrix or exploiting the array geometry and its shift invariance property [6], and then estimate the directions of arrival (DOAs) of signals by the way similar to the classical MUSIC estimator. However, it is shown in [6] that the accuracy of the linear operation based methods [3]-[5] is generally poorer than that of the classical subspace-based methods. On the other hand, for highly correlated or even coherent incident signals, the SUMEW method presented in [6] still relies on the averaging techniques. Although the weighted subspace fitting method without eigendecomposition (WSF-E) [7] is capable of resolving the coherent signals, it still needs the estimate of the array covariance matrix and complex matrix-matrix products to derive the criterion function, thereby indicating that the WSF-E method is still computationally prohibitive

Recently, the methods termed ROCK MUSIC [8] and ROCKET algorithm [9] based on the MSWF developed by Goldstein *et al* [10] were proposed to high-resolution spectral estimation. Nevertheless, the ROCK MUSIC technique requires the forward and backward recursions of the MSWF, which increase the complexity of the algorithm. Moreover, the ROCKET algorithm still needs complex matrix-matrix products to find the reduced-rank data matrix and the reduced-rank autoregressive (AR) weight vector. This implies that additionally computational burden is included.

In this paper, we consider a low-complexity method of weighted subspace fitting (WSF) for DOA estimation. With the assumption that the training data of one desired signal are well known, the novel method is developed. Firstly, a new criterion function is derived based on the MSWF. The novel signal subspace is obtained by the forward recursion of the MSWF, which merely involves complex matrix-vector products. And then, the DOA parameters can be readily extracted by minimizing the novel criterion function. As a results, the new estimator has the attractive advantages. First, it does not require the estimate of the array covariance matrix, its eigendecomposition or the backward recursion of the MSWF. Furthermore, all operations included are merely complex matrix-vector products, thereby requiring much lower computational cost than the existing subspace based methods for direction estimation. Second, all the DOA parameters of the desired signal (with the knowledge of training data) and the "interference" signals (without the knowledge of training data) can be efficiently estimated. Simulation results imply that the proposed estimator can provide the comparable estimation accuracy with the classical WSF method.

II. *PROBLEM FORMULATION*

A. *Data Model*

Let us consider P narrow-band signals from distinct directions $\{\theta_1, \theta_2, \cdots, \theta_P\}$ impinging upon a uniform linear array (ULA) composed of M isotropic elements. The $M \times 1$ received noisy array data at the kth snapshot can be described by the following model

$$\mathbf{x}(k) = \sum_{i=1}^{P} \mathbf{a}(\theta_i)s_i(k) + \mathbf{n}(k) \quad k = 0, 1, \cdots, N-1 \tag{1}$$

where $s_i(k)$ is the scalar complex waveform referred to as the ith signal, $\mathbf{n}(k) \in \mathcal{C}^{M\times 1}$ is the additive noise vector, N denotes the number of snapshots, P represents the number of signals, and $\mathbf{a}(\theta_i)$ is the steering vector of the array toward direction θ_i that is measured relative to the normal of array, and takes the following form

$$\mathbf{a}(\theta_i) = \frac{1}{\sqrt{M}}\left[1, e^{j\varphi_i}, \cdots, e^{j(M-1)\varphi_i}\right]^T \tag{2}$$

where $\varphi_i = \frac{2\pi d}{\lambda}\sin\theta_i$ in which $\theta_i \in (-\pi/2, \pi/2)$, λ and d are the wavelength of the carry signal and the inter-element spacing measured in wave lengths, respectively, the superscript $(\cdot)^T$ denotes the transpose operator. Equation (1) can be rewritten more compactly as

$$\mathbf{x}(k) = \mathbf{A}(\theta)\mathbf{s}(k) + \mathbf{n}(k) \quad k = 0, 1, \cdots, N-1 \tag{3}$$

where

$$\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \cdots, \mathbf{a}(\theta_P)] \tag{4}$$

$$\mathbf{s}(k) = [s_1(k), s_2(k), \cdots, s_P(k)]^T \tag{5}$$

are the $M \times P$ steering matrix and the $P \times 1$ complex signal vector, respectively.

Throughout this paper we assume that $M > P$. Furthermore, the background noise uncorrelated with signals is modeled as a stationary, temporally white, zero-mean Gaussian random process, which is also spatially white and circularly symmetric with the second moments

$$E\left[\mathbf{n}(k)\mathbf{n}^H(l)\right] = \sigma_{\mathbf{n}}^2\delta_{k,l}\mathbf{I}_M \text{ and } E\left[\mathbf{n}(k)\mathbf{n}^T(l)\right] = \mathbf{0} \tag{6}$$

where $\delta_{k,l}$ is the Kronecker delta and $\mathbf{I}_M$ denotes the $M \times M$ identity matrix. We also assume that all signals are jointly stationary, temporally white, zero-mean complex Gaussian random processes. Under these assumptions, the output of the array is complex Gaussian with zero mean and the following array covariance matrix

$$\mathbf{R_x} = E\left[\mathbf{x}(k)\mathbf{x}^H(k)\right] = \mathbf{A}(\theta)\mathbf{R_s}\mathbf{A}^H(\theta) + \sigma_{\mathbf{n}}^2\mathbf{I}_M \tag{7}$$

where $\sigma_{\mathbf{n}}^2$ is the noise variance and $\mathbf{R_s} = E\left[\mathbf{s}(k)\mathbf{s}^H(k)\right]$ is the signal covariance matrix.

B. Classical Weighted Subspace Fitting

Performing the eigenvalue decomposition of the array covariance matrix $\mathbf{R_x}$ leads to

$$\mathbf{R_x} = \mathbf{V_s}\mathbf{\Lambda_s}\mathbf{V_s}^H + \mathbf{V_n}\mathbf{\Lambda_n}\mathbf{V_n}^H = \sum_{i=1}^{M}\lambda_i\mathbf{v}_i\mathbf{v}_i^H \tag{8}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{P'+1} = \cdots = \lambda_M = \sigma_{\mathbf{n}}^2$, $\mathbf{V_s} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{P'}]$ and $\mathbf{V_n} = [\mathbf{v}_{P'+1}, \mathbf{v}_{P'+2}, \cdots, \mathbf{v}_M]$, and P' denotes the number of uncorrelated signals. Accordingly, the eigendecomposition of the sample covariance matrix $\hat{\mathbf{R}}_\mathbf{x} = \frac{1}{N}\sum_{k=0}^{N-1}\mathbf{x}(k)\mathbf{x}^H(k)$ is given by

$$\hat{\mathbf{R}}_\mathbf{x} = \hat{\mathbf{V}}_\mathbf{s}\hat{\mathbf{\Lambda}}_\mathbf{s}\hat{\mathbf{V}}_\mathbf{s}^H + \hat{\mathbf{V}}_\mathbf{n}\hat{\mathbf{\Lambda}}_\mathbf{n}\hat{\mathbf{V}}_\mathbf{n}^H = \sum_{i=1}^{M}\hat{\lambda}_i\hat{\mathbf{v}}_i\hat{\mathbf{v}}_i^H \tag{9}$$

where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_{P'+1} \approx \cdots \approx \hat{\lambda}_M \approx \sigma_{\mathbf{n}}^2$, $\hat{\mathbf{V}}_\mathbf{s} = [\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \cdots, \hat{\mathbf{v}}_{P'}]$ and $\hat{\mathbf{V}}_\mathbf{n} = [\hat{\mathbf{v}}_{P'+1}, \hat{\mathbf{v}}_{P'+2}, \cdots, \hat{\mathbf{v}}_M]$. The column rank of $\mathbf{V_s}$ is in general equal to the rank $P'(P' \leq P)$ of the signal covariance matrix $\mathbf{R_s}$. Thus the columns of $\mathbf{V_s}$ span the P'-dimensional subspace of $\mathbf{A}(\theta)$. Considering (7) and (8) and performing some manipulations, we obtain

$$\mathbf{V_s} = \mathbf{A}(\theta)\mathbf{T} \tag{10}$$

where $\mathbf{T} \in C^{P\times P'}$ is the full rank matrix. Equation (10) forms a basis for the classical signal subspace fitting. θ and $\mathbf{T}$ are unknown and can be acquired by solving (10). In fact, if the theoretical $\mathbf{V_s}$ is replaced by the estimate $\hat{\mathbf{V}}_\mathbf{s}$, there will be no accurate solution to the equation above. In this case, one attempts to minimize some distance measure between $\hat{\mathbf{V}}_\mathbf{s}$ and $\mathbf{A}(\theta)\mathbf{T}$. For this purpose, the Frobenius norm is often used. Therefore, the SSF estimator is obtained by solving the following non-linear optimization problem:

$$\{\hat{\theta}, \hat{\mathbf{T}}\} = \arg\min_{\theta,\mathbf{T}} \|\hat{\mathbf{V}}_\mathbf{s} - \mathbf{A}(\theta)\mathbf{T}\|_F^2. \tag{11}$$

Since the cost function above is quadratic with respect to $\mathbf{T}$, $\hat{\mathbf{T}}$ is easily obtained. Inserting the least squares solution $\hat{\mathbf{T}} = [\mathbf{A}(\theta)^H\mathbf{A}(\theta)]^{-1}\mathbf{A}^H(\theta)\hat{\mathbf{V}}_\mathbf{s}$ into (11) yields the following equivalent optimization problem without the parameter $\mathbf{T}$:

$$\hat{\theta}_{SSF} = \arg\min_{\theta}\left\{\text{tr}\left(\mathbf{P}_\mathbf{A}^\perp\hat{\mathbf{V}}_\mathbf{s}\hat{\mathbf{V}}_\mathbf{s}^H\right)\right\} \tag{12}$$

where $\mathbf{P}_\mathbf{A}^\perp = \mathbf{I}_M - \mathbf{A}(\theta)[\mathbf{A}(\theta)^H\mathbf{A}(\theta)]^{-1}\mathbf{A}^H(\theta)$. Since the eigenvectors are estimated with a quality, commensurate with the closeness of the corresponding eigenvalues to the noise variance, it is natural to weight each eigenvectors and lead to

$$\hat{\theta}_{WSF} = \arg\min_{\theta}\left\{\text{tr}\left(\mathbf{P}_\mathbf{A}^\perp\hat{\mathbf{V}}_\mathbf{s}\mathbf{W}\hat{\mathbf{V}}_\mathbf{s}^H\right)\right\} \tag{13}$$

where $\mathbf{W}$ is the weighting matrix whose *optimal* solution [2] is given by $\mathbf{W}_{opt} = (\mathbf{\Lambda_s} - \sigma_{\mathbf{n}}^2\mathbf{I})^2\mathbf{\Lambda_s}^{-1}$.

C. Multi-Stage Wiener Filter

The multi-stage wiener filter presented by Goldstein *et al* is to find an approximate solution to the Wiener-Hopf equation which does not need the inverse of the array covariance matrix. The MSWF of rank D based on the data-level lattice structure is given by the following set of recursions:

- *Initialization*: $d_0(k) = s_1(k)$ and $\mathbf{x}_0(k) = \mathbf{x}(k)$.
- *Forward Recursion*: For $i = 1, 2, \cdots, D$:
  $\mathbf{h}_i = E[\mathbf{x}_{i-1}(k)d_{i-1}^*(k)]/\|E[\mathbf{x}_{i-1}(k)d_{i-1}^*(k)]\|_2$;
  $d_i(k) = \mathbf{h}_i^H\mathbf{x}_{i-1}(k)$;
  $\mathbf{x}_i(k) = \mathbf{x}_{i-1}(k) - \mathbf{h}_i d_i(k)$.
- *Backward Recursion*: For $i = D, D-1, \cdots, 1$ with $\varepsilon_D(k) = d_D(k)$:
  $w_i = E[d_{i-1}(k)\varepsilon_i^*(k)]/E[|\varepsilon_i(k)|^2]$;
  $\varepsilon_{i-1}(k) = d_{i-1}(k) - w_i^*\varepsilon_i(k)$.

For the MSWF of rank M, the pre-filtering matrix $\mathbf{T}_M = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_M]$ is formed by M matched filters in the forward recursion of the MSWF. To reduce the computational load, the MSWF is truncated at the Dth stage and the reduced-rank transformation matrix is $\mathbf{T}_D = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_D]$. In this paper, we assume that $D \geq P'$. Notice that the orthogonal matched filter $\mathbf{h}_i \in \mathcal{C}^M, i = 1, 2, \ldots, M$ maximizes the real part of the correlation between the new desired signal $d_i(k) = \mathbf{h}_i^H x_{i-1}(k) \in \mathcal{C}$ at the ith stage and the desired signal $d_{i-1}(k)$ at the $(i-1)th$ stage subject to $\|\mathbf{h}_i(k)\| = 1$, forcing the desired signals between successive stages to be in-phase. However, the blocking matrix $\mathbf{B}_i = \mathbf{I} - \mathbf{h}_i\mathbf{h}_i^H$ guarantees that $\mathbf{T}_M$ decorrelates all lags in the process $d_i(k)$ greater than one. It follows that the pre-filtered covariance matrix is tridiagonal:

$$\mathbf{T}_M^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_M = \begin{bmatrix} \mathbf{T}_D^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_D & \mathbf{T}_D^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_{\mathbf{n}'} \\ \mathbf{T}_{\mathbf{n}'}^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_D & \mathbf{T}_{\mathbf{n}'}^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_{\mathbf{n}'} \end{bmatrix} = \begin{bmatrix} \sigma_{d_1}^2 & \delta_2^* & & & \\ \delta_2 & \sigma_{d_2}^2 & \delta_3^* & & \\ & \delta_3 & \sigma_{d_3}^2 & \ddots & \\ & & \ddots & \ddots & \delta_M^* \\ & & & \delta_M & \sigma_{d_M}^2 \end{bmatrix} = \mathbf{D} \quad (14)$$

where $\sigma_{d_i}^2 = E\left[d_i(k)d_i^*(k)\right]$, $\delta_i = E\left[d_i(k)d_{i-1}^*(k)\right]$, $\mathbf{T}_{\mathbf{n}'} = [\mathbf{h}_{D+1}, \mathbf{h}_{D+1}, \cdots, \mathbf{h}_M]$, and $\mathbf{T}_M = [\mathbf{T}_D, \mathbf{T}_{\mathbf{n}'}]$.

III. *LOW-COMPLEXITY METHOD FOR WSF*

A. Derivation of the Novel Criterion Function for WSF

It is easy to see from (14) that $\mathbf{T}_D^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_D$ is also a tridiagonal matrix and can be expressed as

$$\mathbf{T}_D^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_D = E\left[\mathbf{d}(k)\mathbf{d}^H(k)\right] = \mathbf{R}_d \quad (15)$$

where

$$\mathbf{d}(k) = \mathbf{T}_D^H\mathbf{x}_0(k) = [d_1(k), d_2(k), \cdots, d_D(k)]^T \quad (16)$$

in which $d_i(k)$ is calculated by the MSWF algorithm above. For uncorrelated signals, performing the eigenvalue decomposition of $\mathbf{R}_d$ leads to

$$\mathbf{T}_D^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_D = \mathbf{E}_\mathbf{s}\mathbf{\Sigma}_\mathbf{s}\mathbf{E}_\mathbf{s}^H + \mathbf{E}_\mathbf{n}\mathbf{\Sigma}_\mathbf{n}\mathbf{E}_\mathbf{n}^H = \sum_{i=1}^{D}\eta_i\mathbf{e}_i\mathbf{e}_i^H \quad (17)$$

where $\{\eta_i, \mathbf{e}_i\}, i \in \{1, 2, \cdots, P'\}$ are the largest eigenpairs of $\mathbf{R}_d$. Accordingly, the eigendecomposition of the estimate $\hat{\mathbf{R}}_d$ is given by

$$\hat{\mathbf{T}}_D^H\hat{\mathbf{R}}_{\mathbf{x}_0}\hat{\mathbf{T}}_D = \hat{\mathbf{E}}_\mathbf{s}\hat{\mathbf{\Sigma}}_\mathbf{s}\hat{\mathbf{E}}_\mathbf{s}^H + \hat{\mathbf{E}}_\mathbf{n}\hat{\mathbf{\Sigma}}_\mathbf{n}\hat{\mathbf{E}}_\mathbf{n}^H = \sum_{i=1}^{D}\hat{\eta}_i\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i^H \quad (18)$$

where $\{\hat{\eta}_i, \hat{\mathbf{e}}_i\}, i \in \{1, 2, \cdots, P'\}$ are the largest eigenpairs of $\hat{\mathbf{R}}_d$. It is easy to see that the pre-filtered matrix $\mathbf{T}_M$ and its estimate $\hat{\mathbf{T}}_M$ are unitary.

Pre-multiplying and post-multiplying (17) with $\mathbf{T}_D$ and $\mathbf{E}_\mathbf{s}$, respectively, and noting the projection matrix $\mathbf{T}_D\mathbf{T}_D^H = \mathbf{I} - \mathbf{T}_{\mathbf{n}'}\mathbf{T}_{\mathbf{n}'}^H$, we obtain

$$\begin{aligned} \mathbf{T}_D\mathbf{E}_\mathbf{s}\mathbf{\Sigma}_\mathbf{s}\mathbf{E}_\mathbf{s}^H\mathbf{E}_\mathbf{s} &+ \mathbf{T}_D\mathbf{E}_\mathbf{n}\mathbf{\Sigma}_\mathbf{n}\mathbf{E}_\mathbf{n}^H\mathbf{E}_\mathbf{s} \\ &= \mathbf{T}_D\mathbf{T}_D^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_D\mathbf{E}_\mathbf{s} \\ &= \left(\mathbf{I} - \mathbf{T}_{\mathbf{n}'}\mathbf{T}_{\mathbf{n}'}^H\right)\left[\mathbf{A}(\theta)\mathbf{R}_\mathbf{s}\mathbf{A}^H(\theta) + \sigma_\mathbf{n}^2\mathbf{I}\right]\mathbf{T}_D\mathbf{E}_\mathbf{s} \\ &= \mathbf{A}(\theta)\mathbf{R}_\mathbf{s}\mathbf{A}^H(\theta)\mathbf{T}_D\mathbf{E}_\mathbf{s} + \sigma_\mathbf{n}^2\mathbf{T}_D\mathbf{E}_\mathbf{s} - \mathbf{T}_{\mathbf{n}'}\mathbf{T}_{\mathbf{n}'}^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_D\mathbf{E}_\mathbf{s}. \end{aligned} \quad (19)$$

It readily follows from (14) that

$$\mathbf{T}_{\mathbf{n}'}^H\mathbf{R}_{\mathbf{x}_0}\mathbf{T}_D = \begin{bmatrix} \mathbf{0}_{1\times(D-1)} & \delta_{D+1} \\ \mathbf{0}_{(M-D)\times(D-1)} & \mathbf{0}_{(D-1)\times 1} \end{bmatrix}. \quad (20)$$

It is shown in [10] that $\mathbf{x}_i(k)$ trends to become white as the stage of the MSWF increases. As a matter of fact, if $i \in \{P', P'+1, \cdots, M-1\}$, $\mathbf{x}_i(k)$ is a temporally white random process.

Proposition 1: Suppose that there are P' uncorrelated signals impinging on the ULA, then when the rank of the MSWF is equal to the number of signals, the process

$$\mathbf{x}_{P'}^{(P')}(k) = \left(\mathbf{I} - \mathbf{h}_{P'}\mathbf{h}_{P'}^H\right)\mathbf{x}_{P'-1}^{(P')}(k) = \mathbf{B}_{P'}\mathbf{x}_{P'-1}^{(P')}(k) \quad (21)$$

is a temporally white process, and takes the following form

$$\mathbf{x}_{P'}^{(P')}(k) = \left(\prod_{i=P'}^{1}\mathbf{B}_i\right)\mathbf{n}(k) \quad (22)$$

where the superscript $(\cdot)^{(P')}$ represents the case where the number of incident signals equals P', the subscript $(\cdot)_{P'}$ refers to the P'th stage of the MSWF.

The proof of Proposition 1 is seen in Appendix I.

Corollary 1: Suppose that there are P' uncorrelated signals received by the ULA, then when the rank of the MSWF is equal to or greater than the number of signals, the processes $\mathbf{x}_i(k) = \mathbf{B}_i\mathbf{x}_{i-1}(k), i \in \{P', P'+1, \cdots, M-1\}$ are temporally white random processes.

Proof: It is easy to see that the process after the P'th stage of the MSWF has the following form

$$\begin{aligned} \mathbf{x}_i^{(P')}(k) &= \mathbf{B}_i\mathbf{x}_{i-1}^{(P')}(k) = \mathbf{B}_i\mathbf{B}_{i-1}\mathbf{x}_{i-2}^{(P')}(k) = \cdots \\ &= \left(\prod_{j=i}^{P'+1}\mathbf{B}_j\right)\mathbf{x}_{P'}^{(P')}(k) \\ &= \left(\prod_{j=i}^{P'+1}\mathbf{B}_j\right)\left(\prod_{j=P'}^{1}\mathbf{B}_j\right)\mathbf{n}(k) \\ &= \left(\prod_{j=i}^{1}\mathbf{B}_j\right)\mathbf{n}(k) \end{aligned} \quad (23)$$

where $i = P'+1, P'+2, \cdots, M-1$. It follows from (22) and (23) that the process $\mathbf{x}_i^{(P')}(k)$, $i \in \{P', P'+1, \cdots, M-1\}$ is temporally white. Thus, we obtain the Corollary 1. For simplicity, in what follows we assume that the number of uncorrelated signals is P' and suppress the superscript $(\cdot)^{(P')}$ of the process $\mathbf{x}_i^{(P')}(k)$, namely $\mathbf{x}_i(k)$. ■

It follows from Corollary 1 that when $i = \{P', P'+1, \cdots, M-1\}$, we have

$$\begin{aligned} \delta_{i+1} &= E\left[d_{i+1}(k)d_i^*(k)\right] \\ &= E\left[\mathbf{h}_{i+1}^H\mathbf{x}_i(k)\mathbf{x}_{i-1}^H(k)\mathbf{h}_i\right] \\ &= \mathbf{h}_{i+1}^H E\left[\mathbf{x}_i(k)\mathbf{x}_{i-1}^H(k)\right]\mathbf{h}_i \\ &= \mathbf{h}_{i+1}^H E\left[\left(\prod_{t=i}^{1}\mathbf{B}_t\right)\mathbf{n}(k)\left[\left(\prod_{s=i-1}^{1}\mathbf{B}_s\right)\mathbf{A}(\theta)\mathbf{s}(k)\right.\right. \end{aligned}$$

$$+ \quad \left(\prod_{s=i-1}^{1} \mathbf{B}_s \right) \mathbf{n}(k) \Bigg]^H \Bigg] \mathbf{h}_i \tag{24}$$

$$= \mathbf{h}_{i+1}^H \left(\prod_{t=i}^{1} \mathbf{B}_t \right) E\left[\mathbf{n}(k)\mathbf{n}(k)^H\right] \left(\prod_{s=i-1}^{1} \mathbf{B}_s \right)^H \mathbf{h}_i$$

$$= \sigma_{\mathbf{n}}^2 \mathbf{h}_{i+1}^H \left(\mathbf{I} - \sum_{t=1}^{i} \mathbf{h}_t \mathbf{h}_t^H \right) \left(\mathbf{I} - \sum_{s=1}^{i-1} \mathbf{h}_s \mathbf{h}_s^H \right) \mathbf{h}_i$$

$$= 0.$$

Note that the orthogonal property of the matched filters $\mathbf{h}_i$ is used in (24). So $\mathbf{T}_{\mathbf{n}'}^H \mathbf{R}_{\mathbf{x}_0} \mathbf{T}_D = \mathbf{0}$, and (19) can be reduced to

$$\mathbf{T}_D \mathbf{E}_{\mathbf{s}} \left(\mathbf{\Sigma}_{\mathbf{s}} - \sigma_{\mathbf{n}}^2 \mathbf{I} \right) = \mathbf{A}(\theta) \mathbf{R}_{\mathbf{s}} \mathbf{A}^H(\theta) \mathbf{T}_D \mathbf{E}_{\mathbf{s}}, \tag{25}$$

namely

$$\mathbf{T}_D \mathbf{E}_{\mathbf{s}} = \mathbf{A}(\theta) \mathbf{R}_{\mathbf{s}} \mathbf{A}^H(\theta) \mathbf{T}_D \mathbf{E}_{\mathbf{s}} \left(\mathbf{\Sigma}_{\mathbf{s}} - \sigma_{\mathbf{n}}^2 \mathbf{I} \right)^{-1}. \tag{26}$$

Let

$$\begin{aligned} \mathbf{U} &= \mathbf{T}_D \mathbf{E} = \mathbf{T}_D \left[\mathbf{E}_{\mathbf{s}} \ \mathbf{E}_{\mathbf{n}} \right] \\ &= \left[\mathbf{T}_D \mathbf{E}_{\mathbf{s}} \ \mathbf{T}_D \mathbf{E}_{\mathbf{n}} \right] \\ &= \left[\mathbf{U}_{\mathbf{s}} \ \mathbf{U}_{\mathbf{n}} \right] \end{aligned} \tag{27}$$

where $\mathbf{U}_{\mathbf{s}} = \mathbf{T}_D \mathbf{E}_{\mathbf{s}}$, $\mathbf{U}_{\mathbf{n}} = \mathbf{T}_D \mathbf{E}_{\mathbf{n}}$. Therefore, (26) can be reexpressed as

$$\mathbf{U}_{\mathbf{s}} = \mathbf{A}(\theta) \mathbf{K} \tag{28}$$

where $\mathbf{K} = \mathbf{R}_{\mathbf{s}} \mathbf{A}^H(\theta) \mathbf{T}_D \mathbf{E}_{\mathbf{s}} \left(\mathbf{\Sigma}_{\mathbf{s}} - \sigma_{\mathbf{n}}^2 \mathbf{I} \right)^{-1} \in \mathcal{C}^{P \times P'}$. It is easy to see that $\mathbf{K}$ is of full rank. It follows that $\mathbf{U}_{\mathbf{s}}$ spans the signal subspace. Thereby, the relation (28) creates a novel basis for the SSF, and we have the following new criterion function:

$$\left\{ \hat{\theta}, \hat{\mathbf{K}} \right\} = \arg \min_{\theta, \mathbf{K}} \| \hat{\mathbf{U}}_{\mathbf{s}} - \mathbf{A}(\theta) \mathbf{K} \|_F^2 \tag{29}$$

where $\hat{\mathbf{U}}_{\mathbf{s}}$ is the estimate of $\mathbf{U}_{\mathbf{s}}$. Similarly to (11), (29) is also quadratic with respect to $\mathbf{K}$. Thus, the parameter $\mathbf{K}$ can be solved and replaced in the criterion function above. For the fixed unknown parameter $\mathbf{A}(\theta)$, the solution for the linear parameter $\mathbf{K}$ is

$$\hat{\mathbf{K}} = \mathbf{A}^{\dagger}(\theta) \hat{\mathbf{U}}_{\mathbf{s}} \tag{30}$$

where $\mathbf{A}^{\dagger}(\theta) = \left[\mathbf{A}^H(\theta) \mathbf{A}(\theta) \right]^{-1} \mathbf{A}^H(\theta)$. Inserting (30) into (29), we get the new criterion function for SSF without $\mathbf{K}$:

$$\hat{\theta} = \arg \min_{\theta} \| \mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{U}}_{\mathbf{s}} \|_F^2 = \arg \min_{\theta} \left\{ tr \left(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{U}}_{\mathbf{s}} \hat{\mathbf{U}}_{\mathbf{s}}^H \right) \right\}. \tag{31}$$

By introducing a weighting for each signal vector estimate $\hat{\mathbf{u}}_i$, $i = 1, 2, \cdots, P'$, we eventually obtain the following criterion function for weighted subspace fitting:

$$\begin{aligned} \hat{\theta}_{WSF} &= \arg \min_{\theta, \mathbf{K}} \| \hat{\mathbf{U}}_{\mathbf{s}} \tilde{\mathbf{W}}^{1/2} - \mathbf{A}(\theta) \mathbf{K} \|_F^2 \\ &= \arg \min_{\theta} \left\{ \mathrm{tr} \left(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{U}}_{\mathbf{s}} \tilde{\mathbf{W}} \hat{\mathbf{U}}_{\mathbf{s}}^H \right) \right\} \end{aligned} \tag{32}$$

where $\tilde{\mathbf{W}}$ is the weighting matrix whose optimal value is $\tilde{\mathbf{W}}_{opt} = \left(\mathbf{\Sigma}_{\mathbf{s}} - \sigma_{\mathbf{n}}^2 \mathbf{I} \right)^2 \mathbf{\Sigma}_{\mathbf{s}}^{-1}$ in which $\mathbf{\Sigma}_{\mathbf{s}} = diag\{\eta_1, \eta_2, \cdots, \eta_{P'}\}$.

It is not difficult to show from Lemma 1 of [11] that $\hat{\eta}_i$ and $\hat{\mathbf{u}}_i$ are the Rayleigh-Ritz (RR) values and RR vectors of *Krylov* subspace $\mathcal{K}^{(D)}(\hat{\mathbf{R}}_{\mathbf{x}_0}, \hat{\mathbf{r}}_{\mathbf{x}_0 d_0})$, which are asymptotically equivalent to the eigenvalues and eigenvectors of $\mathbf{R}_{\mathbf{x}_0}$, respectively. It follows that the noise variance can be easily estimated by

$$\begin{aligned} \hat{\sigma}_{\mathbf{n}}^2 &= \frac{1}{M - P'} \sum_{i=P'+1}^{M} \hat{\lambda}_i \\ &= \frac{1}{M - P'} \left[tr(\hat{\mathbf{R}}_{\mathbf{x}_0}) - \sum_{i=1}^{P'} \hat{\lambda}_i \right] \\ &= \frac{1}{M - P'} \left[tr(\hat{\mathbf{R}}_{\mathbf{x}_0}) - \sum_{i=1}^{P'} \hat{\eta}_i \right]. \end{aligned} \tag{33}$$

However, the array covariance matrix $\hat{\mathbf{R}}_{\mathbf{x}_0}$ is unknown yet. Computing its estimate will increase the computational cost of the proposed method. In fact, the trace of $\hat{\mathbf{R}}_{\mathbf{x}_0}$ can be computed by $tr\left(\hat{\mathbf{R}}_{\mathbf{x}_0}\right) = \frac{1}{N} \sum_{i=1}^{M} \mathbf{x}_i^H \mathbf{x}_i$, where $\mathbf{x}_i = \mathbf{X}^T \mathbf{f}_i = [x_i(0), x_i(1), \cdots, x_i(N-1)]^T$ in which $\mathbf{f}_i = [\underbrace{0, \cdots, 0, 1}_{i}, 0, \cdots, 0]^T$ and $\mathbf{X} \in \mathcal{C}^{M \times N}$. It is easy to see from (33) that the array covariance matrix can avoid to be estimated and the computational complexity of $tr(\hat{\mathbf{R}}_{\mathbf{x}_0})$ is reduced from $O(M^2 N)$ to $O(MN)$.

B. Computational Cost Requirement

It should be noted that the efficient implementation of the MSWF based on the data-lever lattice structure avoids the formation of blocking matrices, and all the operations of the MSWF only involve complex matrix-vector products, thereby requiring the computational complexity of $O(MN)$ flops for each matched filter $\mathbf{h}_i, i \in \{1, 2, \cdots, D\}$. To fulfil the design of the low-complexity WSF estimator, the eigendecomposition of the rank D tridiagonal matrix is needed, which requires $O(D^3)$ complex product operations. Thus, to estimate the signal subspace matrix $\mathbf{T}_{\mathbf{s}}$ of rank P', the computational cost of the proposed method is only $O(DMN + D^2 N + D^3)$ flops. However, the classical WSF estimator relies on the estimate of the array covariance matrix and its eigendecomposition, which require $O(M^2 N + M^3)$ flops. For the case where $D \approx P' \ll min(M, N)$, the computational complexity of the proposed method, *i.e.*, $O(DMN)$ flops, is much lower than that of the classical WSF technique.

IV. NUMERICAL RESULTS

Example 1 Suppose there are three uncorrelated signals with equal power in the far field impinging upon a ULA with 16 isotropic sensors, whose spacings equal half-wavelength. The true DOAs are $\{-5^0, 0^0, 5^0\}$. The background noise is assumed to be a stationary Gaussian white random process with zero mean. Signal-to-noise ratio (SNR) is defined by $10\log(\sigma_{\mathbf{s}}^2/\sigma_{\mathbf{n}}^2)$, where $\sigma_{\mathbf{s}}^2$ is the power of each signal in single sensor. The results shown below are all based on 500 independent trials. The root-mean-squared errors (RMSE's) of estimated DOAs versus SNR are shown in Fig. 1. We can observe from Fig. 1 that the new estimator yields the comparable estimation accuracy with the WSF method when SNR is greater than -5dB. The former slightly surpasses the latter in estimation performance when SNR varies from -10dB to -5dB. AS SNR becomes large, the two estimators achieve the CRB.

Fig. 2 shows the RMSE's of estimated DOAs versus the number of snapshots in the case where SNR equals 5dB and the rank of the MSWF is 5. From Fig. 2, it can be observed that the estimation accuracy of the proposed method nearly coincides with that of the WSF technique over the range of the number of snapshots that we simulated.

Example 2 Consider the case where there are three signals impinging upon the ULA from the same signal source. The first is a direct-path signal and the others refer to the scaled and delayed replicas of the first signal that represent the multipaths or the "smart" jammers. The propagation constants are $\{1, -0.8 + j0.6, -0.4 + j0.7\}$. The true DOAs are also assumed to be $\{-5^0, 0^0, 5^0\}$. The results shown below are all based on 500 independent trials. The RMSE's of estimated DOAs versus SNR are shown in Fig. 3. It is easy to see from Fig. 3 that the estimation performance of the proposed method is identical to that of the WSF estimator when SNR>-5dB, and the former outperforms the latter when SNR≤-5dB. As SNR becomes high, the RMSE's of the two methods approach to the CRB.

For fixed SNR equal to 5dB and the rank of the MSWF equal to 16, Fig. 4 indicates that the proposed method can provide the same estimation accuracy as the WSF estimator over the range of the number of snapshots that we simulated.

V. *CONCLUSION*

An low-complexity WSF method for DOA estimation has been discussed in this paper. By choosing a specific weighting matrix, the low-complexity WSF estimator yields the comparable estimation performance with the classical WSF technique. Unlike the classical WSF method, the proposed method finds the signal subspace merely by calculating the matched filters in the forward recursion of the MSWF, does not require the estimate of the array covariance matrix or its eigendecomposition. Thus, the proposed estimator is computationally efficient.

APPENDIX I
THE PROOF OF PROPOSITION 1

When $P' = 1$, namely the case of one signal, the observation data reads as

$$\mathbf{x}_0^{(1)}(k) = \mathbf{a}(\theta_1)s_1(k) + \mathbf{n}(k). \tag{34}$$

The matched filter $\mathbf{h}_1$ can be computed as

$$\mathbf{h}_1 = \frac{\mathbf{a}(\theta_1)\sigma_{s_1}^2}{\|\mathbf{a}(\theta_1)\sigma_{s_1}^2\|} = \frac{\mathbf{a}(\theta_1)}{\|\mathbf{a}(\theta_1)\|}. \tag{35}$$

Therefore, the new observation data at the first stage of the MSWF is given by

$$\begin{aligned}\mathbf{x}_1^{(1)}(k) &= \left(\mathbf{I} - \mathbf{h}_1\mathbf{h}_1^H\right)\mathbf{x}_0^{(1)}(k) \\ &= \left(\mathbf{I} - \frac{\mathbf{a}(\theta_1)\mathbf{a}^H(\theta_1)}{\|\mathbf{a}(\theta_1)\|^2}\right)[\mathbf{a}(\theta_1)s_1(k) + \mathbf{n}(k)] \\ &= \mathbf{B}_1\mathbf{n}(k) \qquad (36) \\ &= \left(\prod_{i=1}^{1}\mathbf{B}_i\right)\mathbf{n}(k).\end{aligned}$$

Suppose that (22) holds for $P' = K$, namely

$$\mathbf{x}_K^{(K)}(k) = \left(\prod_{i=K}^{1}\mathbf{B}_i\right)\mathbf{x}_0^{(K)}(k) = \left(\prod_{i=K}^{1}\mathbf{B}_i\right)\mathbf{n}(k). \tag{37}$$

Then, when $P' = K + 1$, we have

$$\begin{aligned}\mathbf{x}_{K+1}^{(K+1)}(k) &= \left(\prod_{i=K+1}^{1}\mathbf{B}_i\right)\mathbf{x}_0^{(K+1)}(k) \\ &= \left(\prod_{i=K+1}^{1}\mathbf{B}_i\right)\left[\mathbf{x}_0^{(K)}(k) + \mathbf{a}(\theta_{K+1})s_{K+1}(k)\right] \\ &= \mathbf{B}_{K+1}\left(\prod_{i=K}^{1}\mathbf{B}_i\right)\mathbf{x}_0^{(K)}(k) \qquad (38) \\ &+ \left(\prod_{i=K+1}^{1}\mathbf{B}_i\right)\mathbf{a}(\theta_{K+1})s_{K+1}(k).\end{aligned}$$

Considering the orthogonal property of the matched filters $\mathbf{h}_i, i \in \{1, 2, \cdots, K+1\}$ leads to

$$\begin{aligned}\left(\prod_{i=K+1}^{1}\mathbf{B}_i\right)\mathbf{a}(\theta_{K+1})s_{K+1}(k) &= \left[\prod_{i=K+1}^{1}\left(\mathbf{I} - \mathbf{h}_i\mathbf{h}_i^H\right)\right]\mathbf{a}(\theta_{K+1})s_{K+1}(k) \\ &= \left[\mathbf{I} - \sum_{i=1}^{K+1}\mathbf{h}_i\mathbf{h}_i^H\right]\mathbf{a}(\theta_{K+1})s_{K+1}(k) \qquad (39) \\ &= \left[\mathbf{I} - \mathbf{T}_{K+1}\mathbf{T}_{K+1}^H\right]\mathbf{a}(\theta_{K+1})s_{K+1}(k).\end{aligned}$$

Notice that the orthogonal matched filters $\mathbf{h}_i, i \in \{1, 2, \cdots, K+1\}$ are contained in the signal subspace, thereby $\mathbf{T}_{K+1}\mathbf{T}_{K+1}^H$ is the projection matrix of the column space of $\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \cdots, \mathbf{a}(\theta_{K+1})]$. It follows that

$$\left(\prod_{i=K+1}^{1}\mathbf{B}_i\right)\mathbf{a}(\theta_{K+1})s_{K+1}(k) = \mathbf{0}. \tag{40}$$

Thus, inserting (37) into (38) yields

$$\mathbf{x}_{K+1}^{(K+1)}(k) = \left(\prod_{i=K+1}^{1}\mathbf{B}_i\right)\mathbf{n}(k). \tag{41}$$

REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research," *IEEE Signal Processing Magazine*, pp. 67-94, July 1996.

[2] M. Viberg, B. Otterstem, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, no. 5, pp. 1110-1121, May 1991.

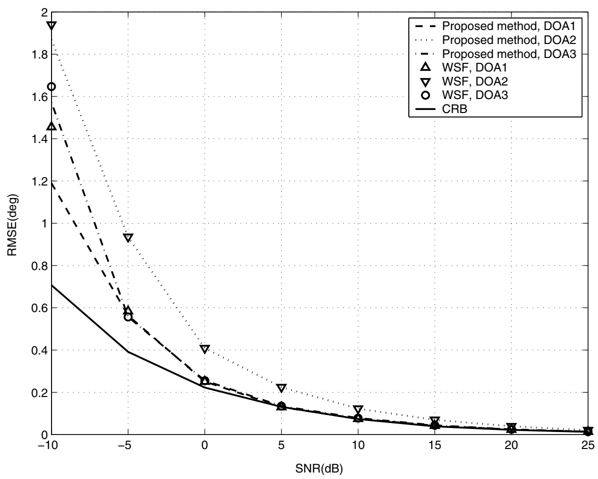

Fig. 1. RMSE's of estimated DOAs for 3 uncorrelated signals versus SNR. DOAs of signal 1 to 3 are -5^{o} 0^{o} and 5^{o}. N=64, D=5 and M=16.

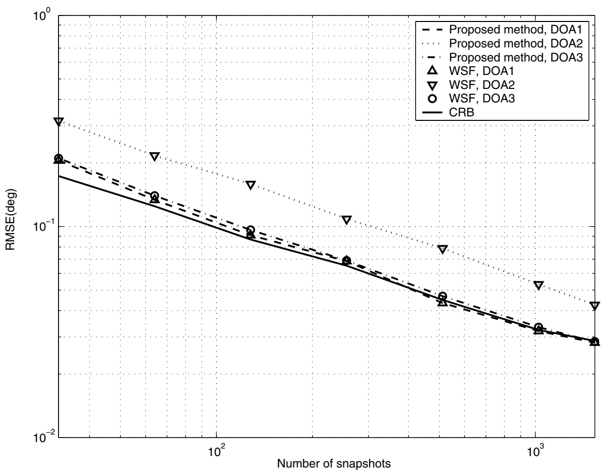

Fig. 2. RMSE's of estimated DOAs for 3 uncorrelated signals versus number of snapshots. DOAs of signal 1 to 3 are -5^{o} 0^{o} and 5^{o}. SNR=5dB, D=5 and M=16.

[3] P. Stoica and T. Söerströ, "Statistical analysis of a subspace method for bearing estimation without eigendecomposition," *Proc. Inst. Elect. Eng. F*, vol. 139, no. 4, pp. 301-305, 1992.

[4] S. Marcos, A. Marsal and M. Benidir, "The propagator method for source bearing estimation," *Signal Process.*, vol. 42, no. 2, pp. 121-138, 1995.

[5] A. Eriksson, P. Stoica and T. Söerströ, "On-line subspace algorithms for tracking moving sources," *IEEE Trans. Signal Processing*, vol. 42, pp. 2319-2330, Sept. 1994.

[6] J. Xin and A. Sano, "Computationally efficient subspace-based method for direction-of-arrival estimation without eigendecomposition," *IEEE Trans. Signal Processing*, vol. 52, no. 4, April 2004.

[7] M. Kristensson, M. Jansson and B. Ottersten, "Modified IQML and weighted subspace fitting without eigendecomposition," *Signal Processing*, vol. 79, no. 1, pp. 29-44, 1999.

[8] H. E. Witzgall and J. S. Goldstein, "ROCK MUSIC-non-unitary spectral estimation," Tech. Rep. ASE-00-05-001, SAIC, May 2000.

[9] H. E. Witzgall and J. S. Goldstein, "Detection performance of the reduced-rank linear predictor ROCKET," *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1731-1738, July 2003.

[10] J. S. Goldstein, I. S. Reed and L. L. Scharf, "A multistage representation of the wiener filter based on orthogonal projections," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2943-2959, Nov. 1998.

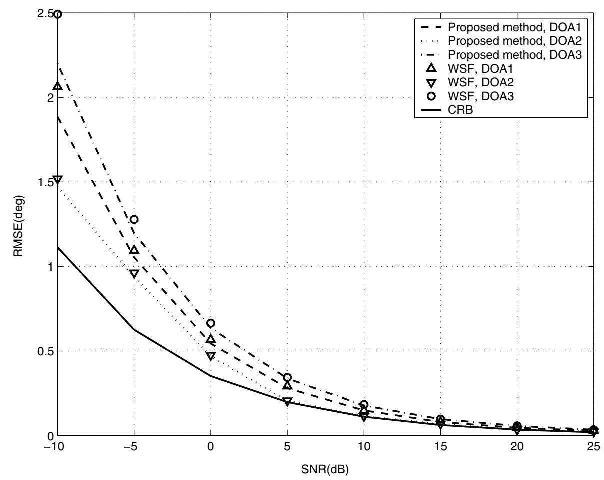

Fig. 3. RMSE's of estimated DOAs for 3 coherent signals versus SNR. DOAs of signal 1 to 3 are -5^{o} 0^{o} and 5^{o}. N=64, D=3 and M=16.

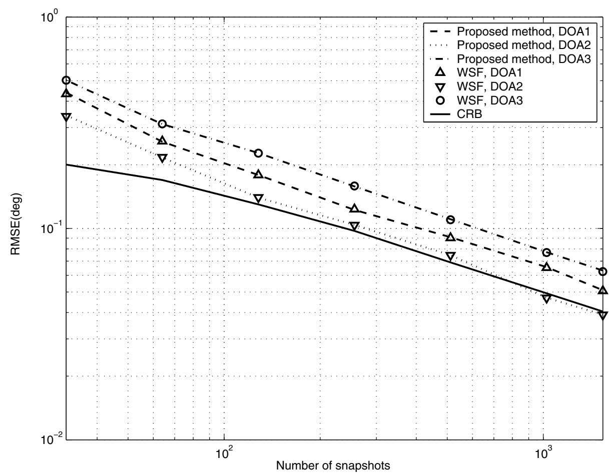

Fig. 4. RMSE's of estimated DOAs for 3 coherent signals versus number of snapshots. DOAs of signal 1 to 3 are -5^{o} 0^{o} and 5^{o}. SNR=5dB, D=3 and M=16.

[11] G. Xu and T. Kailath, "Fast subspace decomposition," *IEEE Trans. Signal Processing*, vol. 42, no. 3, pp. 539-551, March 1994.

BIOGRAPHIES

Lei Huang was born in Guangdong China. He received the B. S. and M. S. degrees in electrical engineering from Xidian University, China, in 2000 and 2003, respectively. From 2000 to 2004 he was with the National Key Lab of Radar Signal Processing, Xidian University, where he is currently pursuing the Ph. D. degree in electrical engineering.

He current research interests focus on statistical signal processing and array signal processing.

Corresponding address: National Key Lab of Radar Signal Processing, Xidian University; Xi' an, Shaanxi 710071, P. R. China.

Email: huanglei8rsp@yahoo. com. cn, sjwu@xidian.edu.cn.

0-7803-8882-8/05/$20.00 (C) 2005 IEEE