PAPER

A Simple Method to Stop an Adaptive Process for the Multistage Wiener Filter

Junichiro SUZUKI[†a)], **Yoshikazu SHOJI**[†], **Hiroyoshi YAMADA**[††], **Yoshio YAMAGUCHI**[††], *and* **Masahiro TANABE**[†], *Members*

SUMMARY The multistage Wiener filter (MWF) outperforms the full rank Wiener filter in low sample support environments. However, the MWF adaptive process should be stopped at an optimum stage to get the best performance. There are two methods to stop the MWF adaptive process. One method is to calculate until the final full-stage, and the second method is to terminate at r-stage less than full-stage. The computational load is smaller in the latter method, however, a performance degradation is caused by an additional or subtractive stage calculation. Therefore, it is very important for the r-stage calculation to stop an adaptive process at the optimum stage. In this paper, we propose a simple method based on a cross-correlation coefficient to stop the MWF adaptive process. Because its coefficient is calculated by the MWF forward recursion, the optimum stage is determined automatically and additional calculations are avoided. The performance was evaluated by simulation examples, demonstrating the superiority of the proposed method.

key words: *reduced rank, multistage Wiener filter, optimum stage, stopping criterion*

1. Introduction

This paper describes a simple method to stop an adaptive process for the multistage Wiener filter (MWF) [1]. It is often the case that the received data of the real-world are weak, complicated, colored, and non-stationary. Thus, it is effective for signal processing algorithms to adapt to such environments. The MWF is a signal-dependent reduced rank adaptive processing algorithm, and it has been developed for the space-time adaptive processing (STAP) in low sample support environments. Adaptation with a low sample support has a quick adaptability for environmental changes. Therefore, it can be said the MWF is one of the algorithms that would adapt to the real-world data. Because the MWF weight calculation is connected scalar Wiener filters to multistage (i.e., the MWF provides a stage-by-stage decomposition of the Wiener filter solution), the MWF can outperform the full rank Wiener filter when the adaptive process was stopped at the optimum rank. Namely, the optimum rank is equivalent to the number of calculation stages which show the best normalized signal-to-interference-plus-noise ratio (SINR). In addition, if the MWF adaptive process is succeeded, an error signal (i.e., a residual error at each stage) shows more and more smaller. Thus, the MWF solution behaves like removing undesired signals from the input signal at each stage. Additionally, because a performance degradation is caused by an additional or subtractive stage calculation, it is very important to select the optimum stage.

There are two methods to stop the MWF adaptive process. One method is to calculate until the final full-stage with the error loading [2] and the diagonal loading [3], [4]. Then the second method is to terminate at r-stage less than full-stage, where r is the number of stages that is not beyond the degrees-of-freedom (DOF). Because the diagonal loading and the error loading can reduce a influence of the higher stage calculations, it is not necessary to select the optimum stage. However, a large amount of the computational loads are necessary because it must calculate until the final stage. On the other hand, in the r-stage calculation, although the optimum stage must be selected in one way or another, the computational load of the MWF adaptive process is small in comparison with the full-stage calculation. Therefore, if the MWF adaptive process can be stopped at the optimum stage, we consider that the r-stage calculation is effective.

The eigenvalue decomposition (EVD) is a general method to select the optimum rank. For example, the Akaike information criterion (AIC) [5] and the minimum description length (MDL) [6] are well used as a rank selection criterion. However, an additional calculation is necessary because eigenvalues and eigenvectors must be calculated by a sample covariance matrix. On the other hand, there is a well known method which called the white noise gain constraint (WNGC) [7]. The WNGC observes that the norm of a weight vector $\boldsymbol{w}$ which grows in response to mismatch errors. Constraining $\|\boldsymbol{w}\|^2$ adds robustness. Therefore, it implies that thresholding $\|\boldsymbol{w}\|^2$ can be used as a stopping criterion. Although the computational load is smaller than the EVD, an evaluation function $\|\boldsymbol{w}\|^2$ must be calculated at each stage to select the optimum stage. In addition to the WNGC, there are conjugate gradients [7], based on the basis vectors [8], and L-curve [9]. However, an additional calculation is necessary for them because the evaluation function must be calculated.

We propose a simple method based on a cross-correlation coefficient to stop the MWF adaptive process. Because the coefficient is equivalent to a residual error between the ith and the $(i-1)$th stage (i.e., the coefficient denotes a statistical relationship between two stages), we

Manuscript received June 6, 2007.
Manuscript revised December 7, 2007.

[†]The authors are with the Radar & Sensor Systems Engineering Dept., Komukai Operations, Toshiba Corporation, Kawasaki-shi, 212-8581 Japan.
[††]The authors are with the Faculty of Engineering, Niigata University, Niigata-shi, 950-2181 Japan.
a) E-mail: junichiro.suzuki@toshiba.co.jp

DOI: 10.1093/ietcom/e91–b.5.1581

consider that the MWF adaptive process would be stopped when a value of the coefficient shows less than or equal to 1. In addition, because the coefficient is calculated by the MWF forward recursion, the proposed method has the advantage that there is no need to calculate the evaluation function. Therefore, an additional calculation is avoided and the MWF adaptive process is stopped at the optimum stage automatically.

This paper is organized as follows. In Sect. 2, we introduce the space-time signal models and the commonly used the MWF algorithm. The proposed method shows in Sect. 3. We present the performance comparison of rank estimation, the normalized SINR, and computational efficiency by simulation examples in Sect. 4. Lastly, we concluded our results in Sect. 5.

2. Space-Time Signal Models and the Multistage Wiener Filter

2.1 Space-Time Signal Models

The clutter geometry for a side-looking array is shown in Fig. 1. Corresponding to a target range, there are several clutter patches that are located at the same range along a circular ring and at various azimuths and Doppler. Therefore, we assume the received data of the clutter are sum of all clutter patches around the ring. Thus, the clutter space-time covariance matrix $\boldsymbol{R}_c$ is given by

$$\boldsymbol{R}_c = \sum_{i=1}^{N_c} \sigma_c^2(i)[\boldsymbol{v}_{f_i}\boldsymbol{v}_{f_i}^H] \otimes [\boldsymbol{v}_{\theta_i}\boldsymbol{v}_{\theta_i}^H], \tag{1}$$

where ^{H} is conjugate transpose, N_c is the number of clutter patches uniformly distributed in azimuth angle, $\sigma_c^2(i)$ is the power of the ith clutter patch, and $\otimes$ is the Kronecker product. Here, $\boldsymbol{v}_{\theta_i}$ and $\boldsymbol{v}_{f_i}$ are the spatial and temporal steering vectors are given by

$$\boldsymbol{v}_{f_i} = [1, e^{j\omega_i}, \cdots, e^{j(M-1)\omega_i}]^T, \tag{2}$$

$$\boldsymbol{v}_{\theta_i} = [1, e^{j\psi_i}, \cdots, e^{j(N-1)\psi_i}]^T, \tag{3}$$

where ^{T} is transpose, M is the number of pulses, and N is the number of elements. The normalized Doppler frequency ω_i and the normalized spatial angle ψ_i are defined as follows

$$\omega_i = 2\pi\beta\sin\theta_i, \tag{4}$$

$$\psi_i = 2\pi\frac{d}{\lambda}\sin\theta_i, \tag{5}$$

where d is the element spacing, λ is the radar wavelength, and θ_i is the azimuth angle. Here, the clutter aliasing parameter β is given by

$$\beta = \frac{2v_p}{d \cdot f_p}, \tag{6}$$

where v_p is the platform velocity, and f_p is the pulse repetition frequency (PRF).

The jammer spatial and temporal covariance matrix $\boldsymbol{R}_j$ is given by

$$\boldsymbol{R}_j = \sigma_j^2[\boldsymbol{I}_M \otimes [\boldsymbol{v}_{\theta_j}\boldsymbol{v}_{\theta_j}^H]], \tag{7}$$

where σ_j^2 is the jammer power, and $\boldsymbol{I}_M$ is the $M \times M$ identity matrix. The jth jammer spatial steering vector $\boldsymbol{v}_{\theta_j}$ is given by

$$\boldsymbol{v}_{\theta_j} = [1, e^{j\psi_j}, \cdots, e^{j(N-1)\psi_j}]^T, \tag{8}$$

where the jth normalized jammer spatial angle ψ_j is defined as follows

$$\psi_j = 2\pi\frac{d}{\lambda}\sin\theta_j, \tag{9}$$

where θ_j is the jammer direction-of-arrival (DOA).

The noise covariance matrix $\boldsymbol{R}_n$ is the identity matrix of dimension $NM \times NM$, as follows

$$\boldsymbol{R}_n = \sigma_n^2[\boldsymbol{I}_M \otimes \boldsymbol{I}_N], \tag{10}$$

where σ_n^2 is the noise power, and $\boldsymbol{I}_N$ is the $N \times N$ identity matrix.

The desired signal covariance matrix $\boldsymbol{R}_s$ is given by

$$\boldsymbol{R}_s = \sigma_s^2[\boldsymbol{s}\boldsymbol{s}^H], \tag{11}$$

where σ_s^2 is the target power, and $\boldsymbol{s}$ is the space-time steering vector, used by $\boldsymbol{s}_{\bar{f}_d}$ and $\boldsymbol{s}_{\bar{\theta}}$, as follows

$$\boldsymbol{s} = \boldsymbol{s}_{\bar{f}_d} \otimes \boldsymbol{s}_{\bar{\theta}}, \tag{12}$$

$$\boldsymbol{s}_{\bar{f}_d} = [1, e^{j\bar{f}_d}, \cdots, e^{j(M-1)\bar{f}_d}]^T, \tag{13}$$

$$\boldsymbol{s}_{\bar{\theta}} = [1, e^{j\bar{\theta}}, \cdots, e^{j(N-1)\bar{\theta}}]^T. \tag{14}$$

Here, the normalized Doppler frequency $\bar{f}_d$ and the normalized spatial angle $\bar{\theta}$ are defined as follows

$$\bar{f}_d = 2\pi\frac{f_d}{f_p}, \tag{15}$$

$$\bar{\theta} = \frac{2\pi d}{\lambda}\sin\theta, \tag{16}$$

where f_d and θ are Doppler frequency and azimuth angle

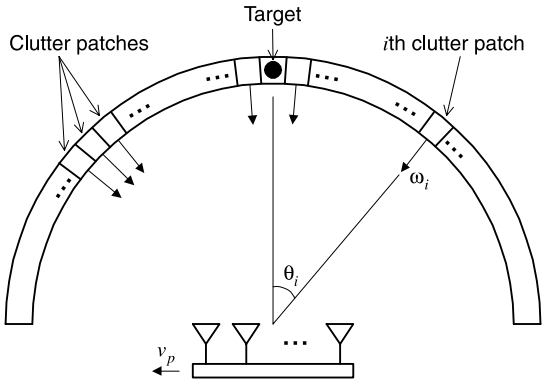


Fig. 1 The clutter geometry for a side-looking array.

of signal. Now our purpose is to maximize the signal response on simultaneously with minimizing the noise response. Therefore, the covariance matrix that we wish to estimate needs to be calculated by using only undesired signal data. Thus, we will assume that so-called "target-free" training data is available, in fact, it can be realized by using guard cells [12]. If there is the target signal in training data, the target signal is suppressed as an undesired signal. That is so-called "self-nulling," but it can be prevented by using a target-free training data. For these reasons, the total (clutter, jammer, and noise) covariance matrix $\boldsymbol{R}$ can be as the sum of the individual covariance matrices. Thus, the covariance matrix $\boldsymbol{R}$ is given by

$$\boldsymbol{R} = \boldsymbol{R}_c + \sum_{i=1}^{J} \boldsymbol{R}_j(i) + \boldsymbol{R}_n = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{x}(k)\boldsymbol{x}(k)^H, \tag{17}$$

where J is the number of jammers, $\boldsymbol{x}(k)$ is the kth training sample of the $NM \times 1$ received data vector, and K is the total number of training samples.

For example, the maximum signal-to-noise ratio (MSN) [10] weight vector $\boldsymbol{w}_{msn}$ is defined as follows

$$\boldsymbol{w}_{msn} = \frac{\boldsymbol{R}^{-1}\boldsymbol{s}}{\boldsymbol{s}^H\boldsymbol{R}^{-1}\boldsymbol{s}}. \tag{18}$$

2.2 The Multistage Wiener Filter

The MWF is a signal-dependent reduced rank adaptive algorithm. In addition, the MWF can operate fewer rank compared with the other reduced rank algorithm, e.g., the principal components (PC) and the cross spectral metric (CSM) [1]. The most commonly used the MWF filter structure is shown in Fig. 2, where the number of calculation stages is 3. The MWF analytical solution is formed in forward and backward recursion steps. The MWF recursion equations are summarized in Table 1, where $E[\cdot]$ denotes ensemble average, $\boldsymbol{r}_{x_i di}$ is the cross-correlation vectors, $d_i(k)$ is the desired signals, δ_i is the magnitude of $\boldsymbol{r}_{x_i di}$, $\boldsymbol{h}_i$ is the direction of cross-correlation vectors, w_i is the scalar weights, and $\epsilon_i(k)$ is the residual error of the ith stage, respectively. Additionally, $\boldsymbol{B}_i$ is the blocking matrices, and $null(\cdot)$ means $\boldsymbol{B}_i\boldsymbol{h}_i = \boldsymbol{0}$.

There are some methods for the blocking matrix $\boldsymbol{B}_i$ calculation, e.g., eigenvalue decomposition, effective covariance matrix, and a low computational complexity described in [1]. Above all methods, we consider that the effective covariance matrix is the most efficient because a matrix calculation is needless.

Now the blocking matrix $\boldsymbol{B}_i$ calculated by the effective covariance matrix is defined as follows

$$\boldsymbol{B}_i = \boldsymbol{I}_{N\times M} - \boldsymbol{h}_i\boldsymbol{h}_i^H. \tag{19}$$

Therefore, the ith stage output data $\boldsymbol{x}_{i+1}(k)$, i.e., next stage input data, is expanded by following equations.

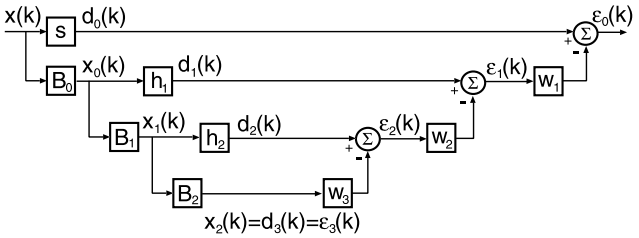


Fig. 2 The filter structure of 3 stages MWF.

Table 1 Forward and backward recursion equations.

Forward Recursion	Backward Recursion
$\boldsymbol{r}_{x_i d_i} = E\left[\boldsymbol{x}_i(k)d_i^*(k)\right]$	$\xi_N = E\left[\mid \boldsymbol{x}_{N-1}(k) \mid^2\right]$
$\delta_{i+1} = \sqrt{\boldsymbol{r}_{x_i d_i}^H \boldsymbol{r}_{x_i d_i}}$	$\delta_N(k) = \boldsymbol{r}_{x_{N-1}d_{N-1}}$
$\boldsymbol{h}_{i+1} = \boldsymbol{r}_{x_i d_i}/\delta_{i+1}$	$w_i = \delta_i/\xi_i$
$\boldsymbol{B}_{i+1} = null(\boldsymbol{h}_{i+1})$	$\sigma_{d_i}^2 = E\left[\mid d_i(k) \mid^2\right]$
$d_{i+1}(k) = \boldsymbol{h}_{i+1}^H \boldsymbol{x}_i(k)$	$\xi_i = \sigma_{d_i}^2 - \delta_{i+1}\cdot w_{i+1}$
$\boldsymbol{x}_{i+1}(k) = \boldsymbol{B}_{i+1}\boldsymbol{x}_i(k)$	$= E\left[\mid \epsilon_i \mid^2\right]$
$\boldsymbol{x}_{N-1}(k) = d_N(k) = \epsilon_N(k)$	$\epsilon_i(k) = d_i(k) - w_{i+1}\cdot\epsilon_{i+1}(k)$

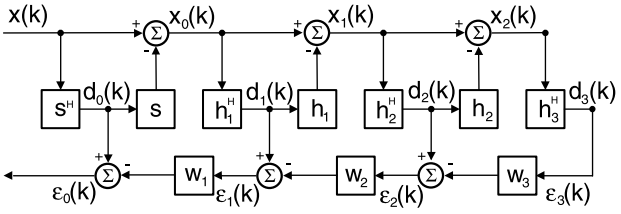


Fig. 3 The filter structure of 3 stages CSS-MWF.

$$\begin{aligned}\boldsymbol{x}_{i+1}(k) &= \boldsymbol{B}_{i+1}\boldsymbol{x}_i(k) \qquad (20)\\ &= (\boldsymbol{I}_{N\times M} - \boldsymbol{h}_{i+1}\boldsymbol{h}_{i+1}^H)\boldsymbol{x}_i(k)\\ &= \boldsymbol{x}_i(k) - \boldsymbol{h}_{i+1}\boldsymbol{h}_{i+1}^H\boldsymbol{x}_i(k)\\ &= \boldsymbol{x}_i(k) - \boldsymbol{h}_{i+1}d_{i+1}(k). \qquad (21)\end{aligned}$$

Here, this efficient MWF filter structure which called the correlation subtractive structure (CSS)-MWF [11] is shown in Fig. 3. As shown in [11], the computational cost by using (21) requires $O(NMK)$ while (20) requires $O(N^2M^2K)$. Clearly, it is more effective to use the effective covariance matrix if the covariance matrix is more and more higher rank.

Now the total output of the residual error $\epsilon_0(k)$ by the MWF filter structure can be written as follows

$$\begin{aligned}\epsilon_0(k) =& d_0(k) - w_1d_1(k) + w_1w_2d_2(k)\\ &- w_1w_2w_3d_3(k) + \cdots. \end{aligned}\tag{22}$$

Besides, the MWF weight vector $\boldsymbol{w}_{mwf}$ is given by

$$\boldsymbol{w}_{mwf} = \boldsymbol{s} - w_1\boldsymbol{h}_1 + w_1w_2\boldsymbol{h}_2 - w_1w_2w_3\boldsymbol{h}_3 + \cdots. \tag{23}$$

Thus, the rank reduction of the MWF (i.e., r-stage calculation) is provided by truncating the MWF adaptive process at the ith stage.

3. A Simple Stopping Criterion of the MWF Adaptive Process

Recall from Table 1, scalar coefficients ξ_i and w_i are written as follows

$$\xi_i = \sigma_{d_i}^2 - \delta_{i+1}^2/\xi_{i+1} = E\left[|\epsilon_i|^2\right], \tag{24}$$

$$w_i = \delta_i/\xi_i. \tag{25}$$

As mentioned previously, because the MWF stage analysis is calculated to remove undesired signal at each stage, ξ_i shows the expected value of the magnitude squared error at the ith stage. Namely, ξ_i shows almost zero when the MWF stage analysis reached the optimum rank. Therefore, because ξ_i shows almost zero, w_i grows large as shown in (25). If each coefficient can determine the optimum threshold, we would consider that these scalar coefficients ξ_i and w_i can be used as a stopping criterion. However, it is difficult to determine those optimum thresholds.

Now we introduce a scalar coefficient η_i defined by multiplying ξ_i by w_i as follows

$$\eta_i = \xi_i \cdot w_i (= \delta_i), \tag{26}$$

where η_i is equal to δ_i as shown in (25) clearly. In addition, recall from Table 1, the direction of cross-correlation vectors $\boldsymbol{h}_i$, the cross-correlation vectors $\boldsymbol{r}_{x_i d_i}$, and the desired signals d_i are written as follows

$$\boldsymbol{h}_{i+1} = \boldsymbol{r}_{x_i d_i}/\delta_{i+1}, \tag{27}$$

$$\boldsymbol{r}_{x_i d_i} = E\left[\boldsymbol{x}_i(k) d_i^*(k)\right], \tag{28}$$

$$d_{i+1}(k) = \boldsymbol{h}_{i+1}^H \boldsymbol{x}_i(k). \tag{29}$$

Thus, δ_i is represented by (26) and (27), as follows

$$\delta_i = \eta_i = \boldsymbol{h}_i^H \boldsymbol{r}_{x_{i-1} d_{i-1}}. \tag{30}$$

Substituting (28) and (29) into (30), we obtain

$$\eta_i = E\left[\boldsymbol{h}_i^H \boldsymbol{x}_{i-1}(k) d_{i-1}^*(k)\right] = E\left[d_i(k) d_{i-1}^*(k)\right]. \tag{31}$$

In general, a cross-correlation coefficient is a measure of similarity of two signals in signal processing, i.e., the coefficient shows a statistical relationship between two data. Therefore, when the value of cross-correlation coefficient shows almost 1, it can be said that the two data are numerically stable. Similarly, as shown in (31), a scalar coefficient η_i is defined as an expected value of the cross-correlation between the ith and the $(i-1)$th stage desired signal. Therefore, we would consider that the MWF stage analysis was reached the optimum stage when η_i shows almost 1. In addition, we would assume that the noise level of received data is often normalized to 0 dB in radar signal processing. Besides, the MWF adaptive process is calculated under the condition of enough training samples realizing the MWF recursion equations in Table 1. Therefore, the MWF adaptive process would be stopped at the $(i-1)$th stage forward recursion when the value of the ith stage η_i was less than or equal to 1.

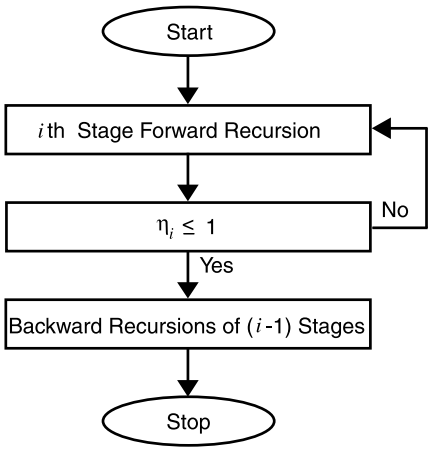


Fig. 4 The flow chart of the MWF adaptive process that used the proposed stopping criterion.

Now we propose a stopping condition of the MWF adaptive process as follows

$$\eta_i \leq 1. \tag{32}$$

Because a scalar coefficient $\eta_i (= \delta_i)$ is calculated by the MWF forward recursion (see, Table 1), an additional calculation is needless. In addition, the MWF adaptive process is stopped automatically by breaking out of the forward recursion loop when (32) was satisfied. Namely, we can determine the number of backward recursion stages automatically. Here, the flow chart of the MWF adaptive process that used the proposed stopping criterion is shown in Fig. 4.

4. Simulation Results

4.1 Performance of Rank Estimation

To demonstrate the performance of rank estimation with the proposed method, we first compared η_i to the eigenvalue of the covariance matrix $\boldsymbol{R}$. The simulation conditions are shown in Table 2, as follows: simulated side-looking array, 8 elements uniform liner array (ULA) with 0.5λ element spacing, 8 pulses, and 64 samples of target-free training data. Here, the element level signal-to-noise ratio (SNR) is 0 dB (i.e., the noise level is 0 dB), clutter-to-noise ratio (CNR) is 30 dB on each element and each pulse, and the conditions of the clutter aliasing β are 0.5, 1.0, 1.5, and 2.0.

Figure 5 shows the magnitude of eigenvalue versus the eigenvalue index. We can see that the rank of the clutter covariance matrix is 16, 15, 23, and 22, respectively. This is because that the results of the 17th, 16th, 24th, and 23rd eigenvalue index are near to the noise level and smaller than the results of the previous eigenvalue index. Next, the magnitude of η_i versus the number of stages is shown in Fig. 6. Comparing Fig. 6 with Fig. 5, we can see that the magnitude of η_i is similar to the magnitude of eigenvalue. As in the case of the eigenvalue, we can estimate that the clutter rank is 16, 15, 23, and 22 because the results of the 17th, 16th,

Table 2 Simulation conditions.

Parameter	Value
Number of Elements	8
Element Spacing	0.5λ
Number of Training Samples	64 (target-free)
Number of Pulses	8
Clutter Aliasing	0.5, 1.0, 1.5, 2.0
Signal-to-Noise Ratio (SNR)	0 dB
Clutter-to-Noise Ratio (CNR)	30 dB

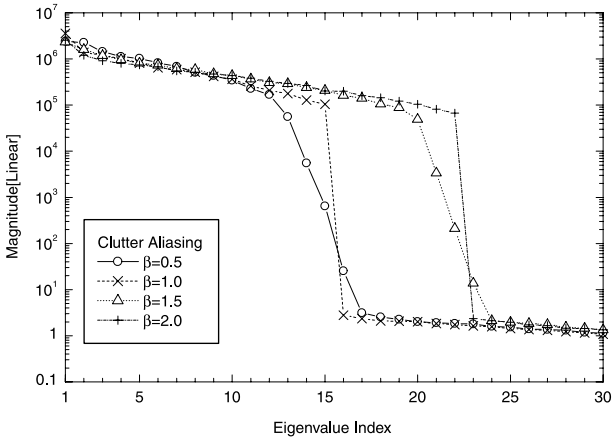


Fig. 5 Magnitude of eigenvalue versus eigenvalue index.

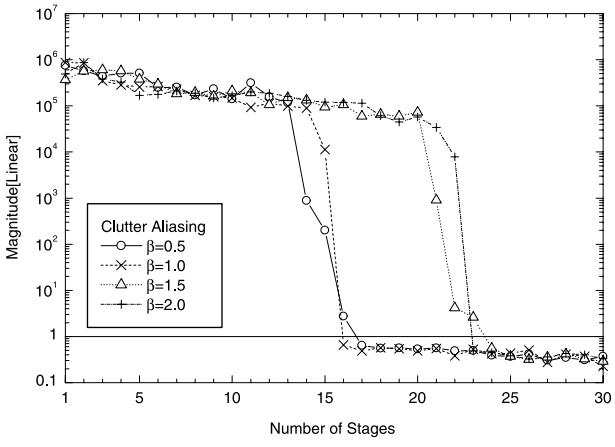


Fig. 6 Magnitude of η_i versus number of stages.

24th, and 23rd stage are less than or equal to 1 (i.e., "1" is the stopping condition of the MWF adaptive process by using the proposed method).

Next, to estimate for a change of clutter aliasing β, we run 1000 Monte Carlo trials for each β changing from 0.1 to 2.0 every 0.1 steps. Then the other simulation parameters are the same in the previous one. Here, we evaluate the parameter Δrank defined as follows

$$\Delta\text{rank} = \text{rank}\{\text{best}\} - \text{rank}\{\text{prop.}\}, \tag{33}$$

where rank{best} is the number of calculation stages that the following parameter based on the minimum variance distortionless response (MVDR) beamformer shows the smallest

Table 3 The occurrence of Δrank with the proposed method.

	Δrank								
β	−5	−4	−3	−2	−1	0	1	2	3
0.1	0	3	258	515	224	0	0	0	0
0.2	0	0	0	0	13	987	0	0	0
0.3	0	0	3	726	255	16	0	0	0
0.4	0	0	0	0	11	989	0	0	0
0.5	0	0	0	0	6	994	0	0	0
0.6	0	0	0	616	354	30	0	0	0
0.7	0	0	0	0	1	664	335	0	0
0.8	0	0	0	0	1	999	0	0	0
0.9	0	0	0	0	5	995	0	0	0
1.0	0	0	0	0	2	998	0	0	0
1.1	0	0	0	0	2	873	125	0	0
1.2	0	0	0	0	9	795	196	0	0
1.3	0	0	0	0	0	51	949	0	0
1.4	0	0	0	0	0	1000	0	0	0
1.5	0	0	0	0	33	962	5	0	0
1.6	0	0	0	0	0	235	765	0	0
1.7	0	0	0	0	0	1000	0	0	0
1.8	0	0	0	0	288	712	0	0	0
1.9	0	0	0	122	814	64	0	0	0
2.0	0	0	0	0	0	1000	0	0	0

value when it is calculated by having changed the number of stages, and rank{prop.} is the number of calculation stages determined by the proposed method.

$$\text{normalized SINR} = \frac{\boldsymbol{w}_{mvdr}^H \boldsymbol{R} \boldsymbol{w}_{mvdr}}{\boldsymbol{w}_{mwf}^H \boldsymbol{R} \boldsymbol{w}_{mwf}}, \tag{34}$$

where $\boldsymbol{w}_{mvdr}$ is weight vector derived by minimizing the interference plus noise power out of the beamformer, while retaining the desired signal without distortion.

Incidentally, we are now focused on radar applications in this paper, it is necessary to evaluate the performance by using finite sample supports. Therefore, it is desirable to estimate with the ensemble average. In general, the time average and the ensemble average may be different. However, if we run simulations for long periods of time (based on the ergodic theory), the time average is almost equal to the ensemble average. Therefore, we used the average of trials (i.e., Monte Carlo trials) to improve the simulation accuracy in the followings.

Table 3 shows the occurrence of Δrank with the proposed method. Additionally, to compare the performance of the proposed method with another one, the occurrence of the WNGC is shown in Table 4. Besides, the histogram of them is shown in Fig. 7. The threshold of the WNGC is a user-defined value, and it can be said that the typical value is 3 dB (described in [13]). However, it is necessary to apply a smaller value to improve the rank estimation accuracy. Therefore, the value was set to 1 dB in common with [14]. We can see that the proposed method is superior to the WNGC. This is because the result of the WNGC at $\Delta\text{rank} = 0$ is about 59% whereas the proposed method is about 67%. However, as shown in Table 3, we consider that the rank is overestimated at $\beta = 0.1$, 0.3, 0.6, and 1.9 by using the proposed method (Similarly, the result of the WNGC shows the same tendency as shown in Table 4). As men-

Table 4 The occurrence of Δrank with the WNGC.

	Δrank								
β	−5	−4	−3	−2	−1	0	1	2	3
0.1	0	0	0	738	262	0	0	0	0
0.2	0	0	0	0	355	645	0	0	0
0.3	0	0	0	107	856	37	0	0	0
0.4	0	0	0	0	1	996	3	0	0
0.5	0	0	0	0	68	932	0	0	0
0.6	0	0	0	0	882	118	0	0	0
0.7	0	0	0	0	0	157	842	1	0
0.8	0	0	0	0	0	999	1	0	0
0.9	0	0	0	0	2	998	0	0	0
1.0	0	0	0	0	116	884	0	0	0
1.1	0	0	0	0	0	601	399	0	0
1.2	0	0	0	0	0	598	402	0	0
1.3	0	0	0	0	0	197	803	0	0
1.4	0	0	0	0	1	999	0	0	0
1.5	0	0	0	0	0	981	19	0	0
1.6	0	0	0	0	0	159	841	0	0
1.7	0	0	0	0	0	1000	0	0	0
1.8	0	0	0	0	499	501	0	0	0
1.9	0	0	0	4	905	91	0	0	0
2.0	0	0	0	0	29	971	0	0	0

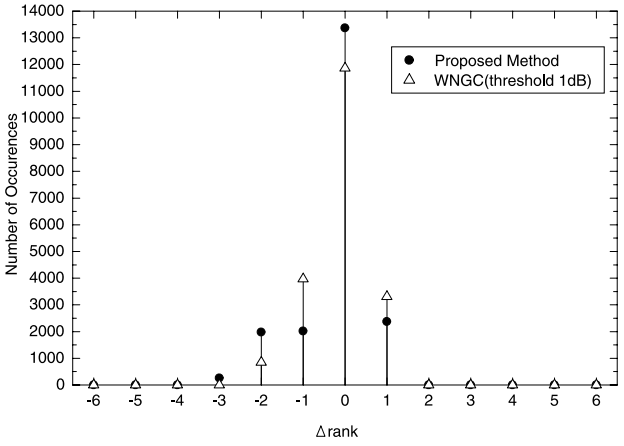


Fig. 7 The histogram of Δrank.

tioned previously, the diagonal loading and the error loading can reduce a influence of the higher stage calculations. Therefore, we consider that the performance of normalized SINR would be improved by applying them to the proposed method (we evaluate in the following paragraph).

In addition to the overestimated case, we discuss the underestimated case (e.g., β = 1.3 and 1.6). Figure 8 shows the normalized SINR versus the number of calculation stages for β = 1.3 and 1.6, and β = 1.0 is shown for comparison. As shown in this figure, we can see that the best performances of normalized SINR for β = 1.0, 1.3 and 1.6 are shown at the 15th, 22nd, and 24th stage, respectively. Additionally, the case which overestimated one stage rapidly deteriorates, however, the case which underestimated one stage hardly deteriorates. In fact, the normalized SINR performance of the underestimated case, i.e., one stage, is almost equivalent to the best performance. For this reason, the underestimated case is not a major problem in

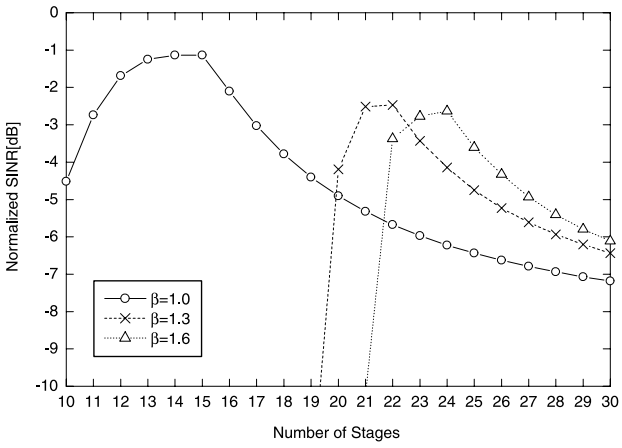


Fig. 8 Normalized SINR versus number of calculation stages.

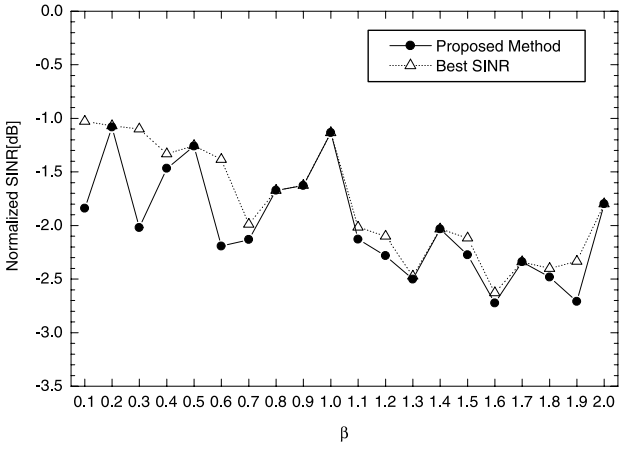


Fig. 9 Normalized SINR versus clutter aliasing parameter β.

comparison with the overestimated case.

4.2 Performance of Normalized SINR

Next, we present the performance of normalized SINR calculated by (34), where the simulation parameters are not changed. Figure 9 shows the performance of normalized SINR with the best SINR and the proposed method. We define "the best SINR" to be the smallest value of (34) for the number of given training samples. Since we consider finite sample supports in this paper, the best SINR is worse than the minimum SINR. Namely, the estimation error occurs in finite sample supports [12]. Therefore, the performance degradation occurs in the best SINR. We can see that the results of the proposed method are approximately consistent with the best SINR performance almost all β, however, we can also see that the normalized SINR deteriorates at β = 0.1, 0.3, 0.6, and 1.9. As mentioned previously, the rank is overestimated at β = 0.1, 0.3, 0.6, and 1.9. Therefore, we evaluate the performance of the proposed method with the error loading to reduce a influence of the higher stage calculations. To apply the error loading to the MWF, the following recursion equation is used instead of (25).

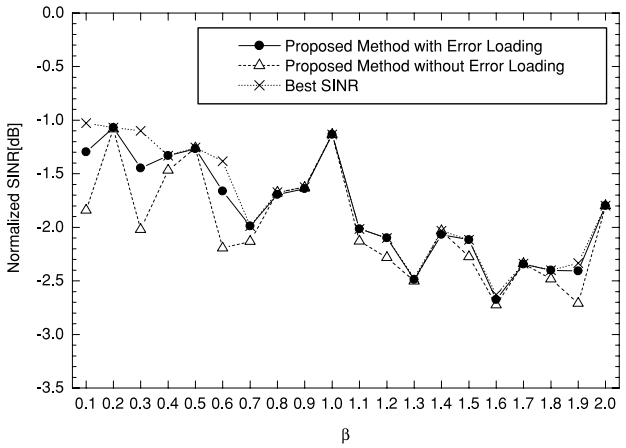


Fig. 10 The application result of error loading to the proposed method.

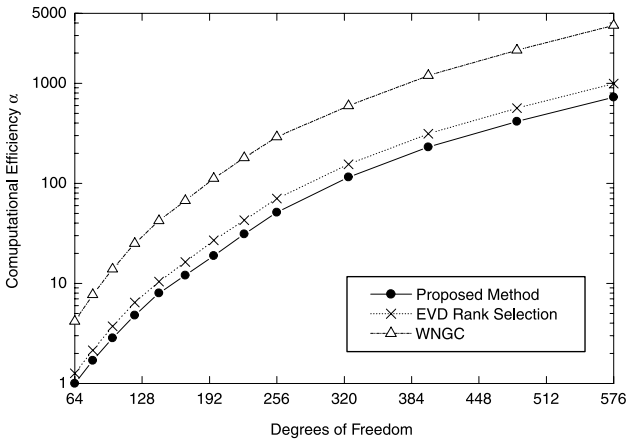


Fig. 11 Computational efficiency α versus degrees-of-freedom.

$$w_i = \delta_i / \{\xi_i + \sigma_{ld}^2\}, \tag{35}$$

where σ_{ld}^2 is the loading level.

The application result is shown in Fig. 10, where the loading level is 10 dB (loading power relative to the element noise power). Clearly, Fig. 10 shows that the performances of normalized SINR at $\beta = 0.1$, 0.3, 0.6, and 1.9 are improved by applying the error loading to the proposed method. In addition, we can see that the other results are approximately consistent with the best SINR. As a result, we consider that the proposed method is operated effectively under the condition of rank overestimation.

4.3 Performance of Computational Efficiency

Lastly, to compare the computational efficiency of some stopping criteria, we estimate the computational time using MATLAB® with *tic*/*toc* function. We used the EVD rank selection with 6 dB threshold, the WNGC, and the proposed stopping criterion, and changed the number of DOFs. In addition, we applied the error loading to all methods to evaluate under an identical computational load condition. Thereunder, we run 1000 Monte Carlo trials for each DOF using the efficient MWF filter structure (i.e., CSS-MWF), where the number of training samples K is equal to the number of DOFs, the condition of the clutter aliasing β is 1.0, and the other simulation parameters are unchanged previous simulations. Incidentally, the *tic*/*toc* functions work to measure the elapsed time. Therefore, we measured the period T_{DOF} between the completion of data input and the stopping of the MWF stage analysis for each DOF and stopping criterion. Besides, we improved the accuracy by average of 1000 trials. Now we defined the computational efficiency α as follows

$$\alpha = T_{\mathrm{DOF}} / T_{64_{\mathrm{prop.}}}, \tag{36}$$

where $T_{64_{\mathrm{prop.}}}$ is the period of 64 DOFs (i.e., $N = M = 8$) with the proposed method.

Figure 11 shows the computational efficiency α versus the DOF. We can see that the proposed method is more efficient than the other method. For example, the efficiencies of the EVD rank selection, the WNGC, and the proposed method at 64 DOFs are about 1.5, 4.2, and 1.0, respectively. In the case of 576 DOFs, the efficiencies are about 990, 3800, and 730. Here, we also see that the computational efficiency of the WNGC is the largest in comparison with the other methods. This is because the WNGC must calculate the evaluation function $\|\boldsymbol{w}\|^2$ from weight vector *at each stage* to determine whether it reached the optimum stage. For these reasons, although the computational efficiency increases depending on DOF, we consider that the proposed method is an effective stopping criterion.

5. Conclusion

In this paper, we proposed a simple stopping criterion for the MWF. Because the proposed method is based on a cross-correlation coefficient η_i calculated by the MWF forward recursion, an optimum stage is determined automatically and additional calculations are avoided. We demonstrated a performance of rank estimation, normalized SINR, and computational efficiency. The characteristic of η_i was similar to eigenvalue of covariance matrix $\boldsymbol{R}$, and the performance of rank estimation was superior to the WNGC. The normalized SINR was consistent with the best SINR except for some clutter aliasing conditions. To improve the normalized SINR of rank overestimation, we applied the error loading to the proposed method. Consequently, we demonstrated that the performance was improved by applying the error loading to the proposed method. Lastly, we confirmed the effectiveness of the proposed method from a result of computational efficiency, and we considered that the proposed method was an effective stopping criterion of the MWF adaptive process.

References

[1] J.S. Goldstein, I.S. Reed, and L.L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," IEEE Trans. Inf. Theory, vol.44, no.7, pp.2943–2959, Nov. 1998.

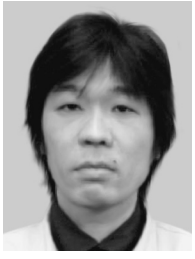
[2] J.D. Hiemstra and J.S. Goldstein, "Robust rank selection for the multistage Wiener filter," Proc. 2002 IEEE Int. Conf. Acoust., Speech, and Signal Processing, vol.III, pp.2929–2932, May 2002.
[3] J.D. Hiemstra, M.E. Weippert, H.N. Nguyen, and J.S. Goldstein, "Insertion of diagonal loading into the multistage Wiener filter," Proc. 2002 IEEE Sensor Array and Multichannel Signal Processing, pp.379–382, Aug. 2002.
[4] W.F. Gabriel, "Using spectral estimation techniques in adaptive processing antenna systems," IEEE Trans. Antennas Propag., vol.AP-34, no.3, pp.291–299, March 1986.
[5] H. Akaike, "A new look at the statistical model identification," IEEE Trans. Autom. Control, vol.AC-19, no.6, pp.716–723, Dec. 1974.
[6] J. Rissanen, "Universal coding, information, prediction, and estimation," IEEE Trans. Inf. Theory, vol.IT-30, no.4, pp.629–636, July 1984.
[7] H. Cox, R.M. Zeskind, and M.M. Owen, "Robust adaptive beamforming," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-35, no.10, pp.1365–1376, Oct. 1987.
[8] M.L. Honig and J.S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," IEEE Trans. Commun., vol.50, no.6, pp.986–994, June 2002.
[9] J.D. Hiemstra, M.E. Weippert, J.S. Goldstein, and T. Pratt, "Application of the L-curve technique to loading level determination in adaptive beamforming," Proc. 2002 IEEE the Thirty-Sixth Asilomar Conf. on Signals, Systems and Computers, vol.2, pp.1261–1266, Nov. 2002.
[10] S.P. Applebaum, "Adaptive arrays," IEEE Trans. Antennas Propag., vol.AP-24, no.5, pp.585–598, Sept. 1976.
[11] D.C. Ricks and J.S. Goldstein, "Efficient architectures for implementing adaptive algorithms," Proc. 2000 Antenna Applications Symposium, pp.29–41, Allerton Park, Monticello, Illinois, Sept. 2000.
[12] J.R. Guerci, Space-Time Adaptive Processing for Radar, Artech House, Norwood, MA, 2003.
[13] J.D. Hiemstra, Robust implementations of the multistage Wiener filter, Ph.D. dissertation, Virginia Polytechnic Institute and State University, April 2003.
[14] C.-C. Hu, H.-Y. Lin, T.-H. Liu, and Y.-S. Cheng, "Time-hopping UWB multiuser detection using adaptive multistage matrix Wiener filtering schemes," Proc. 2007 International Conf. on Commun., pp.2540–2544, Glasgow, Scotland, June 2007.

Junichiro Suzuki was born in Tokyo, Japan, on June 13, 1971. He received the B.E. and M.E. degrees in electrical engineering from Hokkaido University, Sapporo, Japan, in 1994 and 1996, respectively. In 1996, he joined the Toshiba Corporation, where he is a specialist. He is currently a student of doctoral program, Graduate School of Science and Technology, Niigata University. His research area includes adaptive signal processing, phased array antenna, and radar system design. He is a member of IEEE.

Yoshikazu Shoji was born in Niigata, Japan, on June 10, 1979. He received the B.E. and M.E. degrees in information engineering from Niigata University, Niigata, Japan, in 2005 and 2007, respectively. In 2007, he joined the Toshiba Corporation. His research area includes adaptive signal processing.

Hiroyoshi Yamada received the B.E., M.E. and Dr.Eng. degrees from Hokkaido University, Sapporo, Japan, in 1988, 1990 and 1993, respectively, all in electronic engineering. In 1993, he joined the Faculty of Engineering, Niigata University, where he is an associate professor. From 2000 to 2001, he was a Visiting Scientist at Jet Propulsion Laboratory, California Institute of Technology, Pasadena. His current interests involve in the field of array signal processing, radar polarimetry and interferometry, microwave remote sensing and imaging. Dr. Yamada is a member of IEEE.

Yoshio Yamaguchi received the B.E. degree in electronics engineering from Niigata University in 1976, and M.E. and Dr.Eng. degrees from Tokyo Institute Technology in 1978 and 1983, respectively. In 1988, he joined the Faculty of Engineering, Niigata University, where he is a professor. From 1988 to 1989, he was a Research Associate at University of Illinois at Chicago. His interests are in the field of propagation characteristics of electromagnetic waves in lossy medium, radar polarimetry, microwave remote sensing and imaging. Dr. Yamaguchi is a Fellow of IEEE.



Masahiro Tanabe was born in Kanagawa, Japan on January 24, 1964. He received the B.E. and M.E. degrees in electrical engineering from Hosei University, Tokyo, Japan, in 1986 and 1988, respectively. In 1988, he joined the Toshiba Corporation, where he is a chief specialist. His research area includes reflector antenna, phased array antenna, and antenna measurement.