Expanded Fast H_∞ Filter Toward Optimum Frequency-domain H_∞ Filter for Acoustic Echo Canceller

Kohei Hayashida* and Takanobu Nishiura*
* Ritsumeikan University, Kusatsu, Japan
E-mail: {cm012063@ed, nishiura@is}.ritsumei.ac.jp Tel/Fax: +81-77-5615075

***Abstract*—We study a system identification algorithm for an acoustic echo canceller. In research for system identification, a fast H_∞ filter (FHF) has already been developed. However, the FHF has calculation duplications and cannot finish processing in real-time with a long filter length. To overcome these problems, we propose an expanded FHF toward the optimum frequency domain for reducing the processing time. Evaluation experiments of acoustic echo canceling indicate that the expanded FHF is faster than the conventional FHF. The experimental results also indicated that the expanded FHF with restart could suppress the acoustic echo more than the conventional FHF.**

I. Introduction

The echo canceller is one of the typical applications of system identification. For example, system identification is utilized as an acoustic echo canceller for realizing comfortable communication in hands-free telephones, video conference systems, and so on. As for system identification with high-speed convergence, an H_∞ filter [1], [2] has already been developed. However, its computational complexity becomes intractable as the size of the state vector grows larger. To overcome this problem, a fast H_∞ Filter (FHF) [3], [4] for reducing the computational complexity has already been developed. FHF is conventionally utilized in the time domain. The FHF updates the estimated values at one time for one sample and has duplications that are unnecessary for an estimation of an unknown system. Therefore, real-time processing is difficult with a long filter length. To overcome these problems, we developed an expanded FHF toward the optimum frequency domain for realizing real-time processing, and we aim to reduce the number of update times based on frame processing in the frequency domain. The proposed FHF omits the duplications of conventional FHF and expands the conventional time-domain FHF to frequency-domain FHF.

II. System Identification and Acoustic Echo Canceller

In this section, a model of a signal treated in this paper is described to facilitate understanding of the following section. System identification is the decision of the relation between input and output for the unknown system. An impulse response is generally used for representing the parameter of the unknown system. The system identification is realized by estimating the coefficient of the impulse response on the basis of the adaptive algorithm. The observed echo y_k is derived from Eq. (1) by utilizing the output signal $\mathbf{h}_k = [h_k, ..., h_{k-N+1}]$ and the impulse response $\mathbf{x}_k = [x_k(1), ..., x_k(N)]^T$.

$$y_k \quad = \quad \mathbf{h}_k\mathbf{x}_k + v_k \in \Re. \tag{1}$$

The symbol k denotes the discrete time index, N denotes the length of the impulse response, and v_k denotes the observed noise. The estimated value of the observed echo $\hat{y}_k$ is derived from Eq. (2) by utilizing the estimated impulse response $\hat{\mathbf{x}}_k$.

$$\hat{y}_k \quad = \quad \mathbf{h}_k\hat{\mathbf{x}}_k \in \Re. \tag{2}$$

The acoustic echo canceller estimates the unknown impulse response $\mathbf{x}_k$ and the observed echo y_k. After the estimation, it suppresses an acoustic echo by subtracting the estimated echo $\hat{y}_k$ from the observed echo y_k.

III. Fast H_∞ Filter for Conventional Acoustic Echo Canceller

The fast H_∞ filter(FHF)[3], [4] can be recursively performed as follows.

[Step 0] Take the initial conditions for the recursions as,

$$\mathbf{K}_0 = \mathbf{0}_{N\times 2}, \mathbf{A}_{-1} = \mathbf{0}_{N\times 1}, S_{-1} = 1/\varepsilon_0,$$
$$\mathbf{D}_{-1} = \mathbf{0}_{N\times 1}, \hat{\mathbf{x}}_0 = \mathbf{0}_{N\times 1}.$$

The symbol ε_0 denotes a sufficiently large positive number, and $\mathbf{0}_{N\times M}$ denotes the $N \times M$ zero matrix.

[Step 1] Determine $\mathbf{A}_k$ and S_k recursively as,

$$\tilde{\mathbf{e}}_k \quad = \quad \mathbf{c}_k + \mathbf{C}_k\mathbf{A}_{k-1} \in \Re^{2\times 1}, \tag{3}$$
$$\mathbf{A}_k \quad = \quad \mathbf{A}_{k-1} - \mathbf{K}_k\mathbf{W}_k\tilde{\mathbf{e}}_k \in \Re^{N\times 1}, \tag{4}$$
$$\mathbf{e}_k \quad = \quad \mathbf{c}_k + \mathbf{C}_k\mathbf{A}_k \in \Re^{2\times 1}, \tag{5}$$
$$S_k \quad = \quad \rho S_{k-1} + \mathbf{e}_k^T\mathbf{W}_k\tilde{\mathbf{e}}_k \in \Re, \tag{6}$$

where

$$\mathbf{C}_k = \begin{bmatrix} \mathbf{h}_k \\ \mathbf{h}_k \end{bmatrix} \in \Re^{2\times N}, \mathbf{W}_k = \begin{bmatrix} 1 & 0 \\ 0 & -\gamma^{-2} \end{bmatrix} \in \Re^{2\times 2},$$

and $\mathbf{c}_k \in \Re^{2\times 1}$ is the first row of $\mathbf{C}_{k+1} = [\mathbf{c}_k, ..., \mathbf{c}_{k-N+1}]$, assuming that $\mathbf{c}_{k-i} = \mathbf{0}_{2\times 1}$ for $k - i < 0$. In addition, γ denotes the admissible error($\gamma > 1$) and ρ is the forgetting factor($\rho = 1 - \gamma^{-2}$).

[Step 2] Calculate $\breve{\mathbf{K}}_k$ as,

$$\breve{\mathbf{K}}_k = \begin{bmatrix} S_k^{-1}\mathbf{e}_k^T \\ \mathbf{K}_k + \mathbf{A}_k S_k^{-1}\mathbf{e}_k^T \end{bmatrix} \in \Re^{(N+1)\times 2}. \quad (7)$$

[Step 3] Partition $\breve{\mathbf{K}}_k$ as,

$$\breve{\mathbf{K}}_k = \begin{bmatrix} \mathbf{m}_k \\ \mu_k \end{bmatrix}, \mathbf{m}_k \in \Re^{N\times 2}, \mu_k \in \Re^{1\times 2}. \quad (8)$$

[Step 4] Determine $\mathbf{D}_k$ and then obtain the filter gain $\mathbf{K}_{s,k+1}$ at next time index $k+1$ through the gain matrix $\mathbf{K}_{k+1}$ as,

$$\eta_k = \mathbf{c}_{k-N} + \mathbf{C}_{k+1}\mathbf{D}_{k-1} \in \Re^{2\times 1}, \quad (9)$$

$$\mathbf{D}_k = \frac{[\mathbf{D}_{k-1} - \mathbf{m}_k\mathbf{W}_k\eta_k]}{[1-\mu_k\mathbf{W}_k\eta_k]} \in \Re^{N\times 1}, \quad (10)$$

$$\mathbf{K}_{k+1} = \mathbf{m}_k - \mathbf{D}_k\mu_k \in \Re^{N\times 2}, \quad (11)$$

$$\tilde{K}_{k+1}(i) = \rho K_{k+1}(i,1), i = 1, ..., N \quad (12)$$

$$\mathbf{K}_{s,k+1} = \frac{\tilde{\mathbf{K}}_{k+1}}{\rho + \gamma^{-2}\mathbf{h}_{k+1}\tilde{\mathbf{K}}_{k+1}}. \in \Re^{N\times 1} \quad (13)$$

[Step 5] Update the filter equation $\hat{\mathbf{x}}_{k+1}$ as,

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_k + \mathbf{K}_{s,k+1}(y_{k+1} - \mathbf{h}_{k+1}\hat{\mathbf{x}}_k) \in \Re^{N\times 1}. \quad (14)$$

[Step 6] Increase time index k ($k \to k+1$), and return to Step 1.

The FHF gradually becomes unstable by continuous updating. Therefore, the FHF with restart that regularly initializes the parameters has already been developed for realizing stable system identification[5]. The FHF can be recursively performed at the computational complexity of $\mathbf{O}(N)$ per iteration. However, the FHF has duplications that are unnecessary for a calculation of $\hat{\mathbf{x}}_k$ because $\mathbf{C}_k$ is constructed by the same rows $[\mathbf{h}_k, \mathbf{h}_k]^T$. In addition, real-time processing is difficult when filter length N is long.

IV. Expanded fast H_∞ filter toward optimum frequency domain for proposed method

The FHF is conventionally realized in the time domain. Therefore, estimated impulse response $\hat{\mathbf{x}}_k$ is updated once per sample. We expand the time-domain FHF to frequency-domain FHF for realizing real-time processing, and we aim to reduce the number of update times based on the frame processing in the frequency domain. However, frequency-domain expansion is difficult because the conventional FHF has to operate several matrices. Therefore, we represent the FHF that consists of only vector operations by omitting the duplications of the conventional FHF. After that, we expand the FHF to the frequency domain.

A. The omission of the duplication in fast H_∞ filter

The fast H_∞ filter(FHF) without duplications based on $\mathbf{C}_k$ that is constructed by same rows $[\mathbf{h}_k, \mathbf{h}_k]^T$ in Section III can be represented as follows.

[Step 0] Take the initial conditions for the recursions as,

$$\mathbf{K}_0 = \mathbf{0}_{N\times 1}, \mathbf{A}_{-1} = \mathbf{0}_{N\times 1}, S_{-1} = 1/\varepsilon_0, \quad \mathbf{D}_{-1} = \mathbf{0}_{N\times 1}, \hat{\mathbf{x}}_0 = \mathbf{0}_{N\times 1}.$$

Step 0 is exactly the same as that of the FHF in Section III.

[Step 1] Determine $\mathbf{A}_k$ and S_k recursively as,

$$\mathbf{h}_k = [h_k, ..., h_{k-N+1}] \in \Re^{1\times N}, \quad (15)$$

$$\tilde{e}_k = h_k + \mathbf{h}_{k-1}\mathbf{A}_{k-1} \in \Re, \quad (16)$$

$$\mathbf{A}_k = \mathbf{A}_{k-1} - \rho\mathbf{K}_k\tilde{e}_k \in \Re^{N\times 1}, \quad (17)$$

$$\mathbf{e}_k = h_k + \mathbf{h}_{k-1}\mathbf{A}_k \in \Re, \quad (18)$$

$$S_k = \rho(S_{k-1} + e_k\tilde{e}_k) \in \Re. \quad (19)$$

[Step 2] Calculate $\breve{\mathbf{K}}_k$ as,

$$\breve{\mathbf{K}}_k = \begin{bmatrix} S_k^{-1}e_k \\ \mathbf{K}_k + \mathbf{A}_k S_k^{-1}e_k \end{bmatrix} \in \Re^{(N+1)\times 1}. \quad (20)$$

[Step 3] Partition $\breve{\mathbf{K}}_k$ as,

$$\breve{\mathbf{K}}_k = \begin{bmatrix} \mathbf{m}_k \\ \mu_k \end{bmatrix}, \mathbf{m}_k \in \Re^{N\times 1}, \mu_k \in \Re. \quad (21)$$

[Step 4] Determine $\mathbf{D}_k$ and then obtain the filter gain $\mathbf{K}_{s,k+1}$ at next time index $k+1$ through the gain matrix $\mathbf{K}_{k+1}$ as,

$$\eta_k = h_{k-N} + \mathbf{h}_k\mathbf{D}_{k-1} \in \Re, \quad (22)$$

$$\mathbf{D}_k = \frac{\mathbf{D}_{k-1} - \rho\mathbf{m}_k\eta_k}{1-\rho\mu_k\eta_k} \in \Re^{N\times 1}, \quad (23)$$

$$\mathbf{K}_{k+1} = \mathbf{m}_k - \mathbf{D}_k\mu_k \in \Re^{N\times 1}, \quad (24)$$

$$\tilde{\mathbf{K}}_{k+1} = \rho\mathbf{K}_{k+1} \quad (25)$$

$$\mathbf{K}_{s,k+1} = \frac{\tilde{\mathbf{K}}_{k+1}}{\rho + \gamma^{-2}\mathbf{h}_{k+1}\tilde{\mathbf{K}}_{k+1}} \in \Re^{N\times 1}. \quad (26)$$

[Step 5] Update the filter equation $\hat{\mathbf{x}}_{k+1}$ as,

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_k + \mathbf{K}_{s,k+1}(y_{k+1} - \mathbf{h}_k\hat{\mathbf{x}}_k) \in \Re^{N\times 1}. \quad (27)$$

[Step 6] Increase time index k ($k \to k+1$), and return to Step 1.

B. Expanded fast H_∞ filter toward optimum frequency domain for proposed method

Expanded fast H_∞ filter toward the optimum frequency domain can be represented as follows.

[Step 0] Take the initial conditions for the recursions as,

$$\mathbf{K}_0 = \mathbf{0}_{N\times 1}, \mathbf{A}_{-1} = \mathbf{0}_{N\times 1}, \mathbf{S}_{-1} = \mathbf{1}_{N\times 1}/\varepsilon_0, \quad \mathbf{D}_{-1} = \mathbf{0}_{N\times 1}, \hat{\mathbf{x}}_0 = \mathbf{0}_{N\times 1}.$$

[Step 1] Determine $A_k(\omega)$ and $S_k(\omega)$ recursively as,

$$\mathbf{H}_k = DFT[h_k, ..., h_{k-N+1}], \quad (28)$$

$$\tilde{e}_k(\omega) = H_k(\omega) + H_{k-1}(\omega)A_{k-1}(\omega), \quad (29)$$

$$A_k(\omega) = A_{k-1}(\omega) - \rho K_k(\omega)\tilde{e}_k(\omega), \quad (30)$$

$$e_k(\omega) = H_k(\omega) + H_{k-1}(\omega)A_k(\omega), \quad (31)$$

$$S_k(\omega) = \rho(S_{k-1}(\omega) + e_k(\omega)\tilde{e}_k(\omega)). \quad (32)$$

The symbol $\mathbf{H}_k = [H_k(1), ..., H_k(N)]$ denotes the far-end signal in the frequency domain, ω denotes the discrete frequency index, and DFT denotes the discrete Fourier transform.

[Step 2] Calculate $\breve{\mathbf{K}}_k(\omega)$ as,

$$\breve{\mathbf{K}}_k(\omega) = \begin{bmatrix} S_k^{-1}(\omega)e_k(\omega) \\ \mathbf{K}_k(\omega) + \mathbf{A}_k(\omega)S_k^{-1}(\omega)e_k(\omega) \end{bmatrix}. \quad (33)$$

[Step 3] Partition $\breve{\mathbf{K}}_k(\omega)$ as,

$$\breve{\mathbf{K}}_k(\omega) = \begin{bmatrix} m_k(\omega) \\ \mu_k(\omega) \end{bmatrix}. \quad (34)$$

[Step 4] Determine $D_k(\omega)$, and then obtain the filter gain $K_{s,k+1}(\omega)$ at next time index $k+1$ through the gain matrix $K_{k+1}(\omega)$ as,

$$\eta_k(\omega) = H_{k-1}(\omega) + H_k(\omega)D_{k-1}(\omega), \quad (35)$$

$$D_k(\omega) = \frac{D_{k-1}(\omega) - \rho m_k(\omega)\eta_k(\omega)}{1 - \rho\mu_k(\omega)\eta_k(\omega)}, \quad (36)$$

$$K_{k+1}(\omega) = m_k(\omega) - D_k(\omega)\mu_k(\omega), \quad (37)$$

$$\tilde{K}_{k+1}(\omega) = \rho K_{k+1}(\omega), \quad (38)$$

$$K_{s,k+1}(\omega) = \frac{\tilde{K}_{k+1}(\omega)}{\rho + \gamma^{-2}H_{k-1}(\omega)\tilde{K}_{k+1}(\omega)}. \quad (39)$$

[Step 5] Update the filter equation $\hat{x}_{k+1}(\omega)$ as,

$$\mathbf{Y}_k = DFT[y_k, ..., y_{k-N+1}], \quad (40)$$

$$E(\omega) = Y_k(\omega) - H_k(\omega)\hat{x}(\omega)_k, \quad (41)$$

$$\hat{x}_{k+1}(\omega) = \hat{x}_k(\omega) + K_{s,k+1}(\omega)E(\omega). \quad (42)$$

$\mathbf{Y}_k = [Y_k(1), ..., Y_k(N)]$ denotes the observed signal in the frequency domain.

[Step 6] Increase time index k ($k \to k+n$), and return to Step 1. The symbol n denotes the frame shift length.

By frame processing in the frequency domain, the proposed method much reduced the execution time.

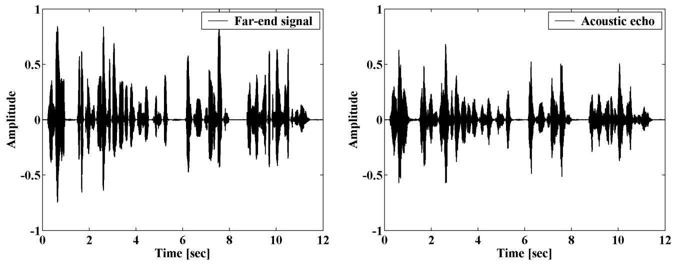
V. Experiments

A. Experimental Conditions

We carried out the suppression experiment of an acoustic echo for evaluating the performance of echo suppression and the execution time. First, an unknown impulse response was estimated based on the conventional FHF in the time domain and the proposed FHF in the frequency domain from the far-end signal and the observed signal with acoustic echo. Then, the estimated echo was designed from the far-end signal and the estimated impulse response. After that, the observed acoustic echo was suppressed by subtracting the estimated echo from the observed signal. Table I shows the experimental conditions. Figures 1(a) and (b) show the time waveforms of the far-end speech signal and its observed signal with acoustic echo. Figures 2(a) and (b) show the Fourier spectrograms of the far-end speech signal and its observed signal with acoustic echo. As for the time waveform of Fig. 1(b), the impulse response was intentionally changed six seconds later. A delay of 24 samples was inserted six seconds later. We used a laptop PC for the evaluation with Core-i5 2.67 GHz CPU and 4 Gbytes of memory. The conventional and the proposed methods were implemented with Matlab.

TABLE I
Experimental conditions

Far-end signals	Speech (male, Japanese sentence), white noise
Impulse response	Japanese-style (Tatami-floored) room ($T_{[60]} = 0.4$ [sec])
Filter length N	8192 [sample]
Frame shift n	4096 [sample]
Admissible error γ	100
Sampling frequency	24 [kHz]
Quantization	16 [bit]
Signal length	12 [sec] (288000 [sample])



(a) Far-end signal. (b) Observed signal.

Fig. 1. Time waveforms (speech signal).

We employed Echo Return Loss Enhancement(ERLE) for evaluating the performance of echo suppression. It is derived from Eq. (43).

$$ERLE_k = 10log_{10}\frac{\sum_{i=k}^{k+m} y_i^2}{\sum_{i=k}^{k+m}(y_i - \hat{y}_i)^2}. \quad (43)$$

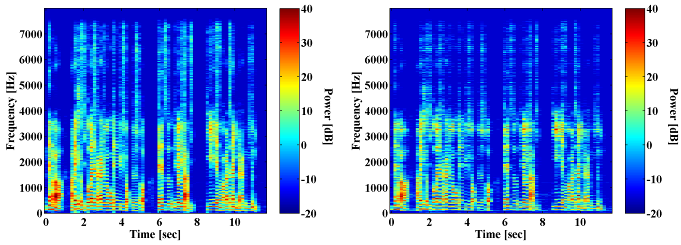
The symbol y_i denotes the observed signal, $\hat{y}_i$ denotes the estimated value of the observed signal y_i, k denotes discrete time index, and m denotes the frame length for calculating ERLE(in this paper, $m = 2400$ samples). We also used the Estimation Performance of Impulse Response(EPIR) for evaluating the estimation performance of impulse response. It is derived from Eq. (44).

$$EPIR_k = 10log_{10}\frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}. \quad (44)$$

The symbol x_i denotes the true value of the impulse response, $\hat{x}_i$ denotes the estimated value of the impulse response, and N denotes the filter length. We carried out the evaluation experiments for four methods: the conventional FHF in the time domain without restart(TD-FHF w/o restart), the FHF in the time domain with restart(TD-FHF w/ restart)[5], the expanded FHF in the frequency domain without restart (FD-FHF w/o restart), and the expanded FHF in the frequency domain with restart(FD-FHF w/ restart). The restart is executed once per 4096 samples(0.17 [sec]).

B. Experimental Results

Table II shows the execution times of the conventional FHF in the time domain and the proposed FHF in the frequency



(a) Far-end signal. (b) Observed signal.

Fig. 2. Spectrograms (speech signal).

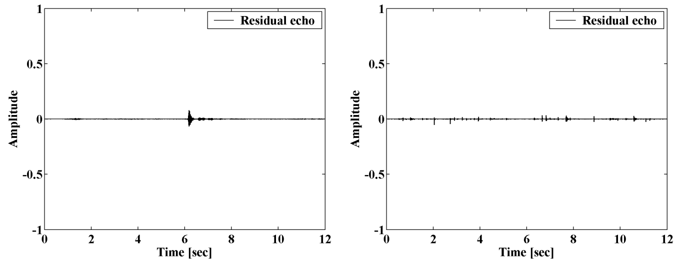
TABLE II
EXECUTION TIMES

TD-FHF (conventional method)	147.0 [sec]
FD-FHF (proposed method)	0.4 [sec]

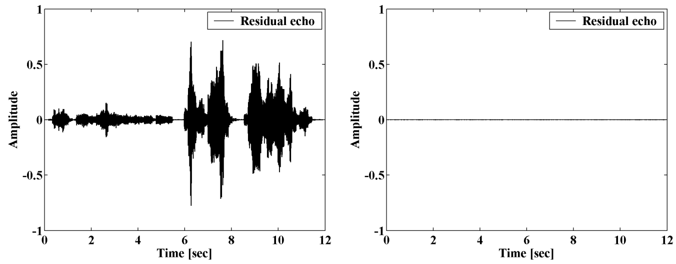
domain. The proposed method could reduce the calculation time by 99.7 % in comparison with the conventional method. Figures 3(a) - (d) show the time waveforms for residual echo with speech signal. Figures 4(a) - (d) show the Fourier spectrograms for residual echo with speech signal. Figures 5 shows the ERLE for all methods with speech signal. These results show that the suppression performance of the acoustic echo improves with the restart. In particular, the suppression performances after changing impulse response improve in the conventional and the proposed method with the restart. These results show that the proposed method with the restart could suppress the acoustic echo more than others. The results for residual echo and ERLE with the white noise show a similar tendency to the results of the speech signal. Figures 6 also shows the EPIR for all methods with speech signal. These results showed that both methods without the restart could accurately estimate the unknown impulse response compared to those methods with the restart. The results of EPIR for the white noise showed a similar tendency to that of the speech signal. All the results indicate that the proposed FHF in the frequency domain with the restart could suppress the acoustic echo in real-time much more than the conventional methods.

C. Discussion

The expanded FHF without restart improves the EPIR, but the expanded FHF with restart improves the ERLE. These facts mean that the estimation performance for the impulse response in the expanded FHF without restart is insufficient. It is caused by degrading of the correlation between the far-end signal and the observed acoustic echo based on the frame processing. Also, the expanded FHF with restart could suppress an acoustic echo by changing the impulse response. Therefore, the expanded FHF with restart converges the filter coefficients for canceling an acoustic echo in each processing frame. In other words, the expanded FHF with restart does not always estimate an unknown system.

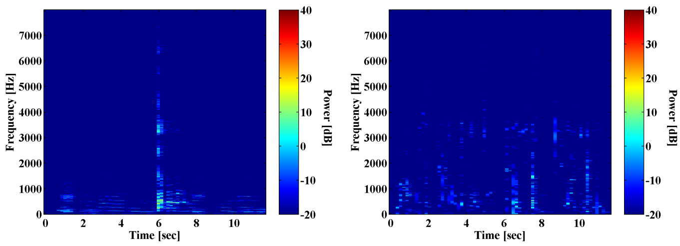


(a) TD-FHF w/o restart. (b) TD-FHF w/ restart.

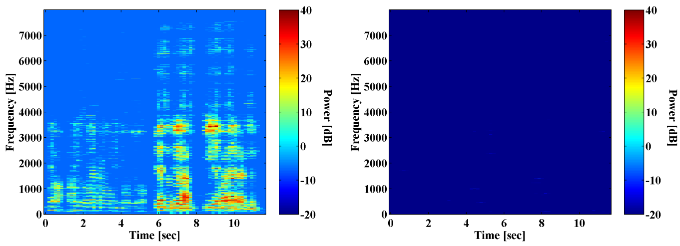


(c) FD-FHF w/o restart. (d) FD-FHF w/ restart.

Fig. 3. Time waveforms for residual echo (speech signal).



(a) TD-FHF w/o restart. (b) TD-FHF w/ restart.



(c) FD-FHF w/o restart. (d) FD-FHF w/ restart.

Fig. 4. Spectrograms for residual echo (speech signal).

VI. CONCLUSION

In this paper, we proposed the expanded fast H_∞ filter toward the optimum frequency domain for realizing an acoustic echo canceller with real-time processing. Evaluation experiments of acoustic echo canceling indicated that the expanded FHF is faster than the conventional FHF. In addition, the expanded FHF with restart could suppress the acoustic echo more than the other methods. However, the expanded FHF with restart does not always estimate an unknown system. In future work, we will try to solve the problem of degrading of the correlation between the far-end signal and the observed

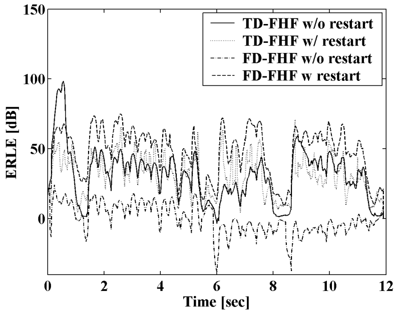


Fig. 5. ERLE (speech signal).

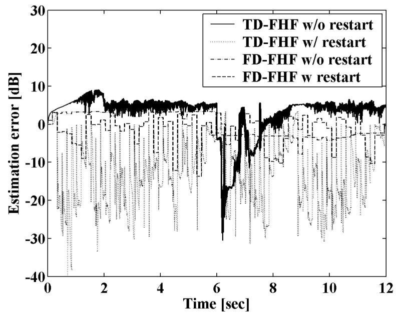


Fig. 6. EPIR (speech signal).

acoustic echo based on frame processing.

ACKNOWLEDGEMENTS

This research was partly supported by a grant-in-aid for scientific research funded by MEXT, Japan.

REFERENCES

[1] B. Hassibi *et. al.* , “Linear estimation in Krein spaces – Part I: Theory”, IEEE Trans. Automat. Contr. , vol. 41, pp. 18–33, 1996.

[2] B. Hassibi *et. al.* , “Linear estimation in Krein spaces – Part II: Applications”, IEEE Trans. Automat. Contr. , vol. 41, pp. 34–49, 1996.

[3] K. Nishiyama, “Derivation of a Fast Algorithm of Modified H_∞ Filters”, in Proc. IEEE Int. Conf. Industrial Electron. , Contr. , Instrum. , pp. 462–467, 2000.

[4] K. Nishiyama, “An H_∞ Optimization and Its Fast Algorithm for Time-Variant System Identification”, IEEE Trans. Signal Proc. , Vol. 52, No. 5, pp. 1335–1342, 2004.

[5] T. Katsumata *et. al.* , “Implementation of the Fast H_∞ Filter in Real-Time Applications”, IEEE CIMCA-IAWTIC ’06, paper ID: 103, 2006.