CANCELLATION OF MULTIPLE ECHOES BY MULTIPLE AUTONOMIC AND DISTRIBUTED ECHO CANCELER UNITS[§]

Akihiko Sugiyama *Kenji Anzai*[†] *Hiroshi Sato*[†] *Akihiro Hirano*

Information Technology Research Laboratories
NEC Corporation
1-1 Miyazaki 4-chome, Miyamae-ku
Kawasaki-shi, KANAGAWA 216

[†] Faculty of Engineering
Nagaoka University of Technology
1603-1 Kamitomioka-cho
Nagaoka-shi, NIIGATA 940-21

ABSTRACT

This paper proposes a scalable multiecho cancellation system based on multiple autonomic and distributed echo canceler units. The proposed system does not have any common control section. Distributed control sections are equipped with in multiple echo cancelers operating autonomically. Necessary information is transferred from one unit to the next one. When the number of echoes to be canceled is changed, the necessary number of echo canceler units, each of which may be realized on a single chip, are simply plugged in or unplugged. The proposed system also provides fast convergence thanks to the novel coefficient location algorithm. The convergence time with a colored-signal input is reduced by approximately 50% over STWQ, and 30% over full-tap NLMS algorithm. With a real speech input, the proposed system cancels the echo by about 20 dB. Tap-positions have been shown to be controlled correctly.

1. INTRODUCTION

In satellite-linked communications, the delay for one relay becomes as long as 600 milliseconds [1]. In such a case, the impulse response of the echo path consists of a long and flat delay, and a dispersive region which actually models the hybrid. The dispersive region is generally 2 to 4 milliseconds, or equivalently, 16 to 32 taps at 8 kHz sampling [2],[3]. When the conversation is relayed more than once, multiple echoes are generated along with multiple flat-delay sections as illustrated in Fig. 1.

Multiecho cancellation requires an adaptive filter with a large number of coefficients when the filter is implemented by an FIR (Finite Impulse Response) filter. This is because the adaptive FIR filter must span the entire length of multiple flat-delay sections and dispersive regions. The required number of taps for the adaptive filter amounts to 4800 for a single-hop satellite link and 9600 for a double-hop link.

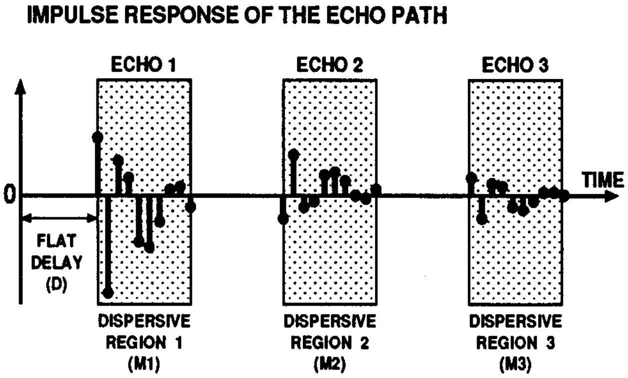


Fig. 1. Impulse Response with Multiechoes.

Such a large number of taps necessitate heavy computational load, increased excess mean squared error, and longer convergence time. However, a careful inspection leads to a good characteristics of such an impulse response of the echo path. There are quite a few coefficients that converge to zero by nature. No coefficients are actually needed at those positions where a flat delay region is located. As the positions of flat-delay regions and dispersive regions are usually unknown *a priori* and different from a transmission channel to another, a smaller number of coefficients must be located adaptively within dispersive regions.

To estimate the positions of dispersive regions, Duttweiler [3] and Yip et al. [4] introduced an auxiliary filter. After convergence of the auxiliary filter, the location information of the dispersive regions is transferred to a number of short adaptive FIR filters. Based on this information, each of the short filters is located on the tapped delay line such that a single filter covers one of the dispersive regions.

The shortcomings of these methods is "scalability." As delay estimation is performed by the auxiliary filter, it cannot be eliminated from these methods, however small the number of multiple echoes to

---

[§] This research was conducted under a student training program arranged by NEC Corporation and Nagaoka University of Technology. Hiroshi Sato is currently with Advantest.

be covered may be. The number of echoes is equal to the number of short adaptive filters and when realization as a silicon chip is considered, it must be defined before LSI design. This means that the resulting LSI cannot be used for a larger number of echoes than that considered at the time of design. Even if more than two chips are prepared for such a case, it is of no use because there is no common control section for those chips. These methods do not have "scalability."

Another series of attempts to model a sparse impulse response can be traced back to the invention of an adaptive delay filter [5]. The adaptive delay filter [6,7] is used to find the dispersive regions for multiecho cancellation [8]. However, it requires the input signal to be sufficiently white [9]. In addition, the auxiliary filter cannot be eliminated to provide overall control [8]. Therefore, multiecho cancellation based on adaptive delay filters cannot be a solution to scalability.

Scrub taps waiting in a queue STWQ algorithm [10] and one of its fast convergence modifications [11] are shown to converge with correlated signals [11]. To be applied to multiecho cancellation, a large number may be selected for the number of taps to cover all dispersive regions. Such an LSI is not economical when it is to be used for a smaller number of dispersive regions. On the other hand, a smaller number of coefficients in such an LSI prevents it from being applied to more dispersive regions. Therefore, there is a trade-off in the selection of the number of coefficients. Once it has been decided, the same LSI cannot be applied nor combined with another to a larger number of multiechoes.

This paper proposes a scalable multiecho cancellation system based on multiple autonomic and distributed echo canceler units. The proposed system does not have any common control section. Instead, multiple echo cancelers have distributed control section in themselves and each of these units operates autonomically. When the number of echoes to be canceled is changed, the necessary number of echo canceler units, each of which may be realized on a single chip, are simply plugged in or unplugged. Section 2 describes the basic algorithm of each unit to cancel a single echo. Considerations on cooperating with multiple units are given in Section 3. The distributed system based on these autonomic units is proposed in Section 4, followed by simulation results in Section 5.

2. ALGORITHM FOR EACH UNIT ECHO CANCELER

The algorithm for canceling a single-echo is based on the one presented in [11]. It consists of two stages: flat-delay estimation and constrained tap-position control. Flat-delay estimation is carried out by identifying the position of the dispersive region. This is achieved by adaptively changing positions of the available coefficients to the dispersive region. The constrained region is determined such that it is centered around a tap which has the largest absolute coefficient value. The step size μ_1 in the flat-delay estimation is set smaller than the step size μ_2 in the constrained tap-position control for steady growth of coefficients in the dispersive region.

For the constrained tap-position control, the queue is divided into two parts. The first queue K_1 is exclusively used for inactive tap indices outside of the constrained region. The second queue K_2 is assigned to those inside. The tap corresponding to the index at the top of K_2 is taken out as a new active tap and is always made active. When an active-tap index is made inactive, it is examined if it falls in the constrained region or not. If it does, it is appended to the end of queue K_2, otherwise, to the end of queue K_1.

2.1. Flat-Delay Estimation

For flat-delay estimation, only K_1, which is one of the two first-in-first-out queues in the algorithm, is used. Flat-delay estimation is performed as follows:

(I) The indices of N delay elements are divided into two groups, an active-tap group $\{J(m) | 1 \le m \le L\}$ and an inactive-tap group $\{K(n) | 1 \le n \le U, U = N - L\}$, where $\{\cdot\}$ denotes a set. A set of L equi-spaced numbers over N available coefficient positions are initially defined as J(m). Coefficients are located at active taps and are set zero as the initial value. K(n) is defined so that $\{J(m) \mid 1 \le m \le L\} \cup \{K(n) \mid 1 \le n \le U\} = \{1, 2, \cdots, N\}$ and kept in queue K_1. The order of K(n) is randomized. Counters T_0 and T_1 are cleared to zero and the tap with the index J(L) is defined as the "minimum-valued tap."

(II) Coefficients at active taps indicated by J(m) are updated for Q iterations by the NLMS algorithm [12] with a step size μ_1. At every coefficient adaptation, the ratio of the input signal power at the minimum-valued tap to the total input signal power at all active taps is compared with a threshold P_m. If the ratio is larger than P_m, a counter T_0 is incremented. Coefficient adaptation is skipped when total input signal power fed to active taps is below a threshold P_s.

(III) The active taps with the R_1 minimum absolute coefficient values are made inactive and their indices are appended to the tail of K_1. The active tap with the minimum absolute coefficient value is newly defined as the "minimum-valued tap."

(IV) The taps indicated by R_1 indices at the top of the queue are newly made active and corresponding coefficients are set zero as new values.

(V) The counter T_1 is incremented if the counter T_0 is equal to Q. The counter T_0 is then cleared to zero.

Procedures from (II) through (V) are repeated until the counter T_1 reaches $(U \times Q)/ R_1$. When this condition is satisfied, flat-delay estimation is terminated. It means that flat-delay estimation has been completed when all of the indices in queue K_1 have been tried out. The index to the maximum-valued coefficient is searched and determined as I_{max}, the center of the dispersive region.

2.2. Constrained Tap-Position Control

Tap positions are exchanged with a constraint that the newly activated tap is within a range around I_{max}, the estimated center of the dispersive region. I_{max} is updated in the constrained tap-position for tracking capability. However, when ERLE (echo return-loss enhancement) is smaller than $ERLE_{th}$, a threshold for ERLE during flat-delay estimation, I_{max} is kept unaltered to assure stable and fast convergence. The constrained tap-position control with two queues is depicted in Fig. 2. It consists of the following steps:

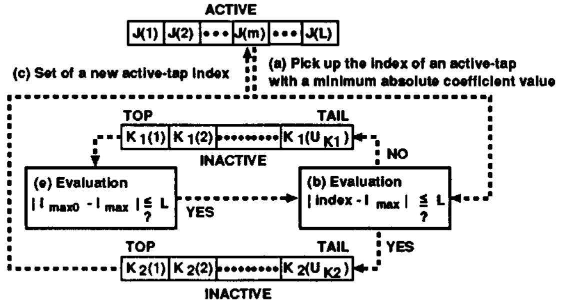
(VI) Coefficients at active taps are updated for Q iterations by the NLMS algorithm with a step size $\mu_2(>\mu_1)$.

(VII) The active taps with the R_2 minimum absolute coefficient values are made inactive, where $R_2 \le R_1$. (Fig. 2 (a))

(VIII) Their indices are evaluated one after another. (Fig. 2 (b)) The index under examination is appended to the tail of the queue K_2 if it is within $I_{max} \pm L$. Otherwise, it is appended to the tail of K_1.

(IX) The R_2 indices at the top of K_2 are taken and corresponding taps to the indices are activated with zero initial coefficient values. (Fig. 2 (c))

(X) If ERLE > $ERLE_{th}$, I_{max} is replaced with $J(m)_{max}$, the index to the tap with the maximum absolute-valued coefficient. The old I_{max} is called I_{max0} hereafter. When $|I_{max0} - I_{max}| \le L$, the W indices at the top of K_1 are taken and the conditional operation in (VIII) is repeated.



L: Number of active taps I_{max} : Tap index to the coefficient with maximum absolute value
U_{K1} : Number of inactive taps near the dispersive region
U_{K2} : Number of inactive taps away from the dispersive region $(N = L + U_{K1} + U_{K2})$

Fig. 2. Constrained Tap-Position Control (R_2=1).

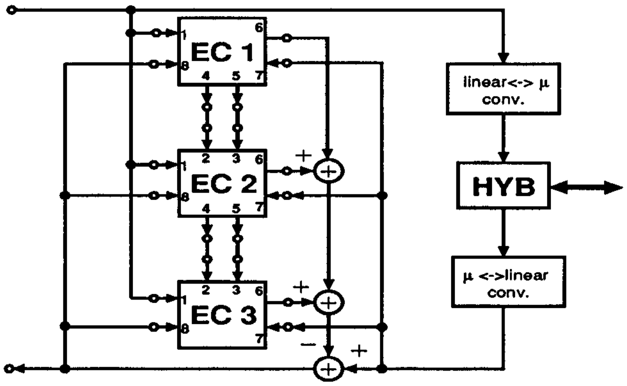


Fig. 3. Multiecho Cancellation System with Three Echo Canceler Units.

(XI) Go to (VI)."

In the flat-delay estimation, K_2 is not used. Therefore, in the beginning of the constrained tap-position control, the index to the taps to be newly activated must be supplied from K_1. In this case, as a transient operation, examination of the indices is necessary based on (VIII). Operations described in (X) are necessary for a path change. If there is no change in the echo path, (X) could be eliminated.

3. CONSIDERATIONS ON COOPERATION OF MULTIPLE ECHO CANCELER UNITS

A single echo canceler unit is assigned to cancel each echo. Thus, M units are required to cancel M echoes. Figure 1 illustrates a multiecho cancellation system for M=3 with parallel connection. Port 1 of each unit receives the reference input signal. Ports 2 is equipped with for obtaining information on the constrained region and active-tap indices outside this region, which are transmitted from Port 4 of the previous unit. Port 5 transmits the timing signal to Port 3 of the next unit when flat-delay estimation has been over. The echo replica is obtained at Port 6 of each unit. Ports 7 and 8 receive the echo and the residual echo, respectively.

Echo replicas generated by echo cancelers (ECs) are summed up and subtracted from the echo. When more than a single EC cooperates to cancel a single multiecho signal, task distribution should be considered deliberately. The overall convergence strategy with multiple units and their constrained regions are main concerns.

If unit ECs operate freely with no restrictions, more than a single unit may try to cancel the same echo, causing mutual interference. One possible procedure is that unit ECs are converged one after another. Shorter convergence could be obtained if the flat-delay estimation of an EC unit is started when that of the previous unit has been over. Therefore, constrained tap-position control of an echo canceler unit and flat-delay estimation of the subsequent EC unit are carried out simultaneously. This transfer of control is illustrated in Fig. 4 as a timing chart.

To avoid overlap of the constrained region among EC units, information on the constrained regions must be shared. These shared data are not necessarily broadcasted if *a priori* knowledge about multiechoes is utilized. In most telecommunication networks, it is true that a dispersive region with a longer flat-delay from the origin has a smaller magnitude. In addition, it is natural to find the most significant dispersive region first. Therefore, the first EC unit is assigned to the first dispersive region, the second EC unit to the second, and so on. By taking this strategy, the dispersive region which an EC unit covers must be located after the region covered by the unit which is scheduled for the previous operation. In other words, an EC unit should transfer the information on its own constrained region to the subsequent unit only. Then, each unit defines its own constrained region such that it has a longer flat delay than that of the previous unit. When flat-delay estimation has been terminated in a unit EC, the following unit is allowed to start its own flat-delay estimation.

It is likely that an EC unit locate a coefficient in a dispersive region which is to be covered by a subsequent unit. When the constrained region has been determined, active-tap indices outside the constrained region are transferred to the next EC unit. The next unit locates coefficients at these taps so that the initial convergence is speeded up. As a result, the data transferred from an EC unit to the next unit should include information on its own constrained region and the adaptation status. Because only the next unit receives information from the previous unit and no global control is necessary, each EC unit can operate autonomically.

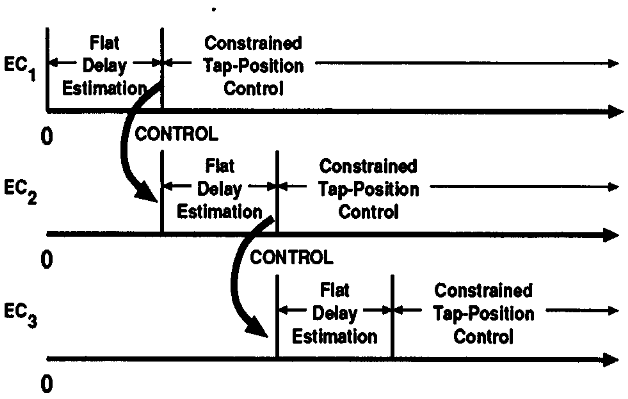


Fig. 4. Timing Chart.

4. DISTRIBUTED AND AUTONOMIC CONTROL OF MULTIPLE EC UNITS

The distributed and autonomic control of multiple echo canceler units based on the algorithm described in Section 2 is performed as follows:

(a) L_1 coefficients, where L_1 is the number of active taps for EC_1, are cleared and distributed evenly over the tapped delay line of EC_1. Inactive tap indices are kept in a queue K_1 for EC_1.

(b) EC_1 estimates the most significant dispersive region by flat-delay estimation. When the flat-delay estimation has been terminated, the coefficient location for its largest absolute coefficient is defined as I_{max1}.

(c) $I_{max1} \pm L_1$, the boundaries of the constrained region for EC_1, are transferred to EC_2.

(d) Indices to active taps are examined. Indices within $I_{max1} \pm L_1$ are kept in EC_1, while the others are transferred to EC_2. The same number of indices as those transferred to EC_2 are extracted from K_2 of EC_1 to be made active with a zero initial value.

(e) EC_2 starts its own flat-delay estimation according to (b). This estimation is carried out for tap indices larger than $I_{max1} + L_1$. I_{max2} is defined like I_{max1}. On the other hand, constrained tap-position control is started for EC_1 with a constraint on the active-tap boundaries of $I_{max1} \pm L_1$.

(f) When the flat-delay estimation for EC_2 has been terminated, the coefficient location for its largest absolute coefficient is defined as I_{max2}.

(g) $I_{max2} \pm L_2$, the boundaries of the constrained region, and active-tap indices outside these boundaries are transferred to EC_3, where L_2 is the number of active taps for EC_2. EC_3 starts its own flat-delay estimation according to (b). This estimation is carried out for tap indices larger than $I_{max2} + L_2$. Both EC_1 and EC_2 start constrained tap-position control with a constraint of $I_{max1} \pm L_1$ and $I_{max2} \pm L_2$, respectively.

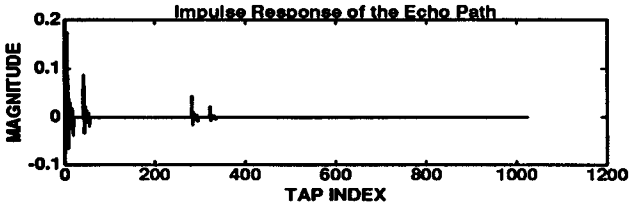


Fig. 5. Impulse Response of the Echo Path.

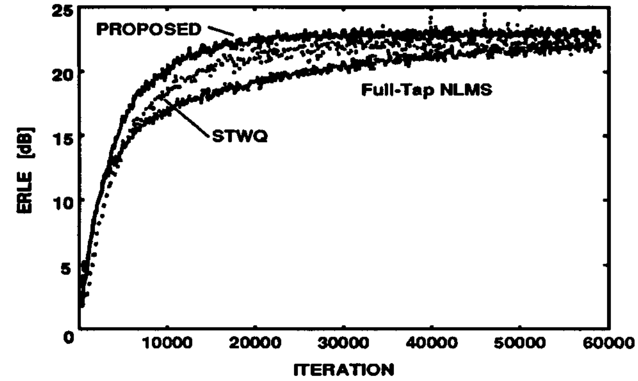


Fig. 6. Convergence Characteristics for a Colored Signal.

These procedures are repeated for each EC for M > 3.

5. SIMULATION RESULTS

Simulations were carried out to compare full-tap NLMS, STWQ [10] and the proposed algorithms. An echo-path impulse response with four echoes [4] shown in Fig. 5 was used. EC_1 cancels the first two dispersive responses and EC_2, the last two. A white *Gaussian* noise (WGN) of zero mean and unit variance is applied to drive a recursive filter with a transfer function of

$$A(Z) = \frac{0.25}{1 - 1.5Z^{-1} + Z^{-2} - 0.25Z^{-3}}. \quad (1)$$

The resulting colored signal was used as the input signal to generate multiechoes. Another WGN with an echo-to-noise ratio of 24dB in magnitude is used as an additive noise. A linear-to-*μ-law* and a *μ-law*-to-linear conversions are both applied twice in cascade to echo generation for modeling PCM (pulse code modulation) codecs. (μ_1, μ_2) are set (0.25, 0.5) for EC_1 and (0.125, 0.25) for EC_2. The step-size for all other algorithms is 0.5. $L_1=L_2=60$, N=1024 and all other parameters are same as [11].

Fig. 6 shows the ERLE defined by

$$ERLE_k = \frac{\sum_{i=k-N+1}^{k} y_i^2}{\sum_{i=k-N+1}^{k} \{y_i - \hat{y}_i\}^2}, \quad (2)$$

where y_i is the i-th sample of the echo and $\hat{y}_i$ is the corresponding echo replica obtained as a sum of replicas generated by unit echo cancelers. An ERLE curve is an average of 100 ERLE samples, each of which is an ensemble average of 50 independent runs. The proposed algorithm saves about 50% in the convergence time over STWQ, and 80% over the full-tap NLMS which requires much more number of taps.

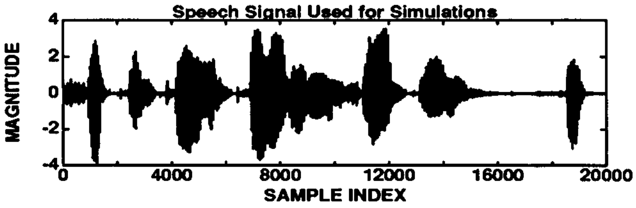


Fig. 7. Speech Signal Used for Simulations.

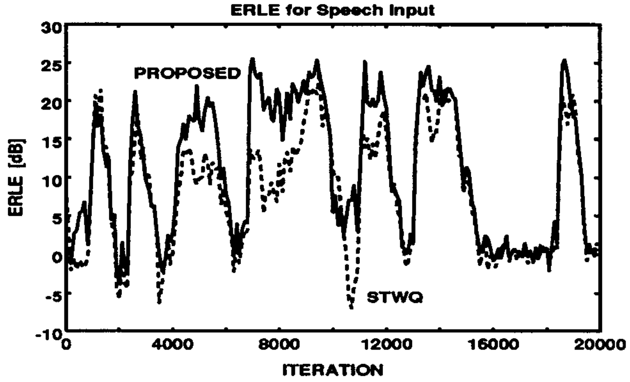


Fig. 8. Convergence Characteristics for the Above Speech Signal.

Convergence for a speech signal, which is depicted in Fig. 7, was also evaluated. Figure 8 compares the ERLE by the proposed and STWQ algorithms. Because of the superior tap-position control, the proposed algorithm outperforms STWQ algorithm. Even with a speech signal as the input, ERLE level reaches 20 dB or even more.

It is interesting to see the coefficient-error vector defined by the difference of the coefficient vector and the impulse-response vector. This is a measure of how well the coefficient vector models the impulse response. Figure 9 shows the norm trajectories of the coefficient-error vector for the proposed and STWQ algorithms. It is clear that coefficients adapted by STWQ do not model the impulse response for speech input, while those by the proposed algorithm do quite well as adaptation goes on.

Finally in Fig. 10, positions of active taps for EC_1 and EC_2 are depicted. For about 5000 iterations, EC_1 performs flat-delay estimation. During this period, active-tap positions are evenly distributed over the available positions between 0 and 1024 which is defined by the total number of taps, N.

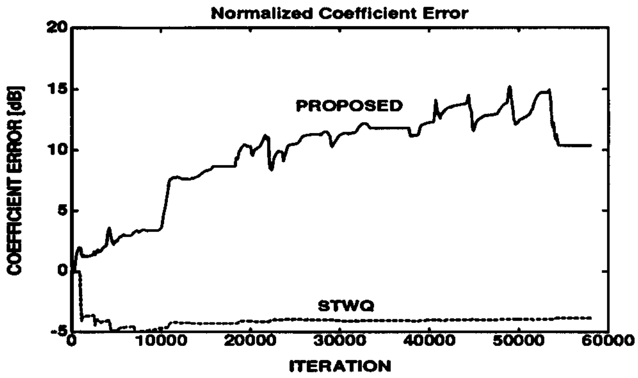


Fig. 9. Norm of the Coefficient-Error Vector.

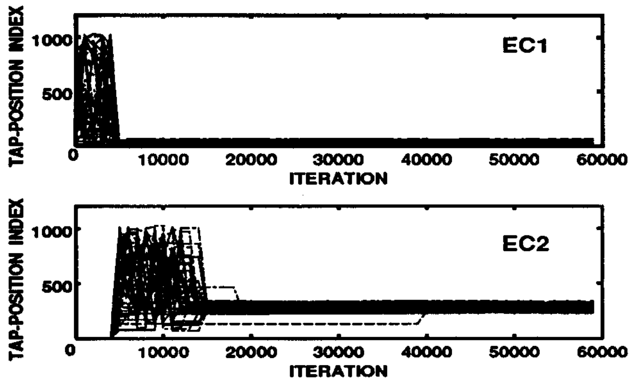


Fig. 10. Positions of Active Taps.

After flat-delay estimation has been over, the constrained region is set in the range of 0 to 60. Trajectories of active tap positions are controlled in this range. In terms of EC_2, flat-delay estimation starts at around the 5000th iteration when EC_1 terminates its own flat-delay estimation. It is over after 10000 iterations and constrained tap-position control begins with a constrained region around the 300th sample of the impulse response in Fig. 5. Because silent sections of the speech are often encountered during flat-delay estimation of EC_2, it takes about twice the number of iterations compared with those for EC_1.

6. CONCLUSION

A scalable multiecho cancellation system based on multiple autonomic and distributed echo canceler units has been proposed. A distributed control section is equipped with in each unit. To change the number of echoes to be canceled, some echo canceler units can simply be plugged in or unplugged. The convergence time with a colored-signal input is reduced by approximately 50% over STWQ, and 80% over full-tap NLMS algorithm. With a real speech input, the proposed system cancels the echo by about 20 dB. Tap-positions have been shown to be controlled correctly.

ACKNOWLEDGMENT

The authors would like to thank Shigeji Ikeda, an Assistant Manager of Radio Applications Division, for fruitful discussions on this topic. They are also indebted to Dr. Takao Nishitani, Senior Manager of Signal Processing Research Laboratory, Information Technology Research Laboratories for his encouragement and valuable comments on this paper.

REFERENCES

[1] M. M. Sondhi et al., "Silencing Echoes on the Telephone Network ," Proc. of the IEEE, Vol.68, No.8, pp.949-963, Aug. 1980.

[2] D. L. Duttweiler, "A Twelve Channel Digital Echo Canceler ," IEEE Trans. on Commun., vol.COM-26, No.5, pp.647-653, May. 1978.

[3] D. L. Duttweiler, "Subsampling to Estimate Delay with Application to Echo Canceling ," IEEE Trans. on ASSP., vol.ASSP-31, No.5, pp.1090-1099, Oct. 1983.

[4] P. C. Yip and D. M. Etter, "An Adaptive Multiple Echo Canceller for Slowly Time-Varying Echo Paths," IEEE Trans. Commun., COM38, pp.1693-1698, Oct. 1990.

[5] D. M. Etter and S. D. Stearns, "Adaptive Estimation of Time Delays in Sampled Data Systems," IEEE Trans. ASSP, vol.ASSP-29, No.3, Part II, pp.582-587, Jun. 1981.

[6] D. M. Etter and Y.-F. Cheng, "System Modeling Using an Adaptive Delay Filter ," IEEE Trans. CAS-34, pp.770-774, Jul. 1987.

[7] Y.-F. Cheng and D. M. Etter, "Analysis of an Adaptive Technique for Modeling Sparse Systems ," IEEE Trans. ASSP-37, pp.254-264, Feb. 1989.

[8] V. A. Margo, D. M. Etter, N. C. Carlson and J. H. Gross, "Multiple Short-Length Adaptive Filters for Time-Varying Echo Cancellation," ICASSP'93, pp.I-161—I-164, Apr. 1993.

[9] J. H. Gross, D. M. Etter, V. A. Margo and N. C. Carlson, "A Block Selection Adaptive Delay Filter Algorithm for Echo Cancellation," Midwest Conference on Circuits and Systems'92, pp.895-898, Aug. 1992.

[10] S. Kawamura and M. Hatori, "A Tap Selection Algorithm For Adaptive Filters," Proc. of ICASSP'86, pp.2979-2982, Apr. 1986.

[11] A. Sugiyama, H. Sato, A. Hirano, and S. Ikeda, "A Fast Adaptive FIR Filter Algorithm with Reduced Computation for Adaptive Tap-Position Control," Proc. 1993 ISPACS'93, pp.20-25, Oct. 1993.

[12] G. C. Goodwin and K. S. Sin, "Adaptive Filtering, Prediction and Control," Prentice-Hall, Info.and Sys. Sci. Ser., 1984.