รายการอ้างอิง

1. C. G. Cho, C. Lee, D. H. Youn and S. J. Park. On integrating acoustic echo and noise cancellation systems for hands-free telephony. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. 2 (May 2001) : 961-964.
2. B. X. Wu and H. Ye. A new double-talk detection algorithm based on the orthogonality theorem. IEEE Transactions on Communications. 39, No.11 (November, 1991) : 1542-1545.
3. D. R. Morgan, J. Benesty and J. H. Cho. An objective technique for evaluating double-talk detectors in acoustic echo cancelers. IEEE Transactions on Speech and Audio Processing. 7, No.6 (November 1999) : 718-724.
4. H. K. Jung, N. S. Kim, and T. Kim. A new double-talk detector using echo path estimation. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. 2 (May 2002) :1897-1900.
5. D. R. Morgan, J. Benesty and J. H. Cho. A new class of double-talk detectors based on cross- correlation. IEEE Transactions on Speech and Audio Processing. 8, No.2 (March 2000) : 168-172.
6. J. Liu. A novel adaptation scheme in the NLMS algorithm for echo cancellation. IEEE Signal Processing Letters. 8, No.1 (January 2001) : 20-22.
7. J. Benesty, M. M. Sondhi, S. L. Gay, , and T. Gansler. Double-talk robust fast converging algorithms for network echo cancellation. . IEEE Transactions on Speech and Audio Processing. 8, No.6 (November 2000) : 656-663.
8. C.-J. Ivarsson, G. Salomonsson, M. Hansson, and T. Gansler. A double-talk detector based on coherence. IEEE Transactions on Communications. 44, No.11 (November 1996) : 1421-1427.
9. T. Creasy and T. Aboulnasr. A Projection-correlation algorithm for acoustic echo cancellation in the presence of double-talk. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. 1 (June 2000) : 436-439.

10. J. H. yoo and S. H. Cho. A new double-talk detector using the lattice predictors for an acoustic echo canceller. Proceedings of IEEE Region 10 Annual Conference on Speech and Image Technologies for Computing and Telecommunications. 2 (December 1997) : 483-486.

11. C. Koike. Adaptive threshold nonlinear algorithm for adaptive filters with robustness against impulse noise. IEEE transactions on Signal Processing. 45, No. 9 (September 1997) : 2391-2395.

12. M.J. Reed and M.O.J. Hawksford. Acoustic echo cancellation with the fast affine Projection. Proceedings of IEE Colloquium on Audio and Music Technology. No. 1998/470 (November 1998) : 16/1 - 16/8.

13. N. Tangsangiumvisai and J.A. Chambers. A fast converging and stable adaptive filtering algorithm for stereophonic acoustic echo cancellation. European Signal Processing Conference (EUSIPCO'2000, Tampere, Finland). 2 (September 2000).

14. S. Haykin. Adaptive Filter Theory. 3rd ed. : Prentice Hall, 1996.

15. S. L. Gay and J. Benesty. Acoustic signal processing for telecommunication. 1st ed. : Kluwer Academic Publisher, 2000.

16. N. Tangsangiumvisai. Introduction to adaptive filters for acoustic echo cancellation in hands-free mobile communication. Communication Engineering Laboratory Sheet. Chulalongkorn University, 2002.

17. G. Schmidt. Acoustic Echo and Noise Control for Low-Cost Processor. Internal Report . Darmstadt University of Technology.

18. S. G. Sankaran and A. A. Beex. Implementation and Evaluation of Echo Cancellation Algorithms. Meng Thesis, Bradley Department of Electrical Engineering Virginia Polytechnic Institute and State University, December 1996.

ภาคผนวก

บทความทางวิชาการที่ได้รับการเผยแพร่

1. A new double-talk detection technique for acoustic echo canellation, Proceeding of Electrical Engineering Conference (EECON-26) vol.2 pp. 1103-1107, November 2003, Pethburi, Thailand

A new double-talk detection technique for acoustic echo cancellation

N. Tangsangiumvisai *and* *N. Ithisoponkul*
Department of Electrical Engineering, Chulalongkorn University,
Bangkok, 10330, Thailand.
E-mail : Nisachon.T@chula.ac.th, N.Ithisoponkul@student.chula.ac.th

Abstract

In this paper, an improved adaptive filtering algorithm is introduced for Acoustic Echo Cancellation (AEC) in the Double-Talk (DT) situation, based on a new DT Detector (DTD). The proposed DTD aims to distinguish between the DT situation and the Abrupt Change in the Acoustic Echo Path (ACAEP) by employing the correlation behaviour of the gradient vector. In addition, another DTD is used to discriminate the DT situation from the steady-state (SS). These two detectors control the adaptation gain of the Normalised Least Mean Square (NLMS) algorithm so that it results in fast convergent rate when there is an ACAEP and slow convergent rate during the DT situation. Simulation results based on white noise sequences support these ideas.

1 Introduction

In an Acoustic Echo Cancellation (AEC), the Double-Talk (DT) situation, where the near-end and far-end speakers speak simultaneously, is a very challenging problem. The adaptation of the adaptive filter in the AEC system is severely disturbed due to the occurrence of the near-end speech signal. This results in divergence of the adaptive filter from its steady-state (SS), and thus, the AEC system fails to eliminate the echo signal. Several approaches have been proposed to tackle the DT problem by the use of DTD. Once the DT situation has been detected, these conventional DTDs, however, freeze the adaptation of the adaptive filters in order to prevent them from divergence [4, 5, 6]. This inhibits the tracking performance of the adaptive filter during the DT situation.

When there is an abrupt change of the acoustic echo path (ACAEP) in the near-end room, the fast convergent rate of the adaptive filter is required to track the ACAEP. On the other hand, a more noise robust and slow convergent rate adaptive filtering algorithm should be chosen in order to track small changes of the acoustic echo path (AEP) during the DT periods [1]. It is therefore necessary for a DT detector (DTD) to be able to distinguish between the DT situation and the ACAEP in order to obtain appropriate tracking performance of the adaptive filter. For both cases, the misadjustment of the adaptive filter, and thus the error signal, is drastically increased. Thus, the error signal cannot be used as a DTD since it cannot distinguish between these two events.

In [2], a method for DT detection is introduced, based on the *sign* of the correlation between the instantaneous gradient estimate $\nabla(n)$ and an average $\bar{\nabla}(n-1)$ of previous estimate :

$$s(n) = \mathrm{sign}[\nabla(n) \cdot \bar{\nabla}(n-1)] \tag{1}$$

The low-pass filtered version of this parameter is then compared with a threshold in order to detect the DT situation. However, the performance of this method is found to be degraded when it used in an AEC system due to long adaptive filter, i.e. slow convergence and excessive sensitivity to the DT and thus, it is unsuitable for the AEC system [3]. Hence, another DTD is proposed in [3] based on the projection-correlation algorithm. This technique utilises the distinction of the trajectories of the adaptive filter coefficients during the ACAEP and the DT situation. The average gradient vector is recursively updated, then the correlation of the new instantaneous and old average gradients are evaluated before being averaged again. An intermediate variable using the sign of the instantaneous correlation is introduced and its smoothed version is used to update the adaptation gain that will be employed in the update equation of the Affine Projection Algorithm (APA). This DT detection technique, however, requires too many stages of averaging without proper mathematical explanation.

It is therefore proposed in this paper a DTD with the ability to distinguish between the DT situation and the ACAEP. Furthermore, the adaptive filtering algorithm for the AEC system employing the proposed DTD is suggested to be able to track small changes during the DT periods, in stead of freezing the adaptation. The correlation behaviour of the gradient vector is proposed as a criterion for DT detection. First, a DTD based on the cross-correlation between the gradient vector estimates at time n and $n-1$ is proposed. This parameter is employed to indicate the *directivity* of the gradient vector and is therefore able to discriminate between the ACAEP and the DT situation. In addition, the *activity* of the gradient vector, based on the autocorrelation of the gradient vector estimates at time n, is suggested to be employed together with the directivity of the gradient vector in order to distinguish between the DT situation and the SS. Both detectors are employed as a criterion to control automatically the adaptation gain of the Normalised Least Mean Square (NLMS) adaptive filtering algorithm in order to obtain fast rate of convergence when there is an ACAEP, whereas slow convergence rate is ensured during the DT situation.

This paper is organised as follows. Section II describes the DT scenario in an AEC system. In Section III, the DTD based on the directivity and the activity of the gradient vectors is given. The improved adaptation performance of the NLMS algorithm with a control mechanism for the adaptation gain is presented in Section IV, followed by the simulation results in Section V. Finally, the conclusions are given in Section VI.

2 An AEC system with the DT situation

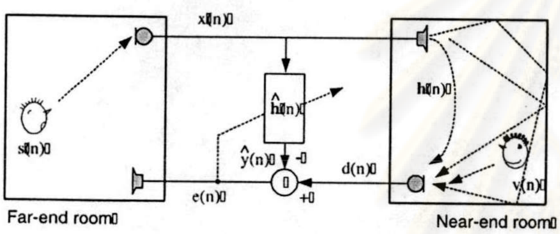


Figure 1: *An AEC diagram in the presence of DT.*

The echo signal in a full-duplex communication system can be eliminated by employing an AEC system, as illustrated in Fig. 1 where the AEC system is applied to the near-end room. Similarly, the echo signal in the far-end room can also be removed by employing the AEC system in the far-end room. By considering at the near-end room, when there exists the near-end speech $v(n)$ at the same time with the far-end speech $x(n)$, the microphone signal is given by

$$d(n) = y(n) + v(n) + w(n) \tag{2}$$

where $y(n)$ is the echo signal which is obtained from the convolution between the AEP in the near-end room, $h(n)$ and the input signal of the AEC system, $x(n)$,

$$y(n) = \mathbf{h}^T(n)\mathbf{x}(n) \tag{3}$$

the input signal vector is given by $\mathbf{x}(n) = [x(n), \quad x(n-1), \quad \ldots, \quad x(n-L+1)]^T$, and $w(n)$ is the spectrally white background noise. The unknown AEP in the near-end room can be characterised by a Finite Impulse Response (FIR) of length L, $\mathbf{h}(n) = [\hat{h}_1(n), \quad h_2(n), \quad \ldots, h_L(n)]^T$. The function of the adaptive filter with the same length L as that of the AEP, $\hat{\mathbf{h}}(n) = [\hat{h}_1(n), \quad \hat{h}_2(n), \quad \ldots, \hat{h}_L(n)]^T$, is to identify the unknown AEP $\mathbf{h}(n)$ by trying to match its output

$$\hat{y}(n) = \hat{\mathbf{h}}^T(n)\mathbf{x}(n) \tag{4}$$

to the echo signal, $y(n)$. Without the DT situation, when the adaptive filter matches the AEP in the near-end room, the error signal, which is given by

$$e(n) = d(n) - \hat{y}(n) \tag{5}$$

is approaching zero, and thus, the echo signal can be eliminated. This, however, no longer holds with the inclusion of the near-end speech signal $v(n)$. The significant energy of the near-end speech results in low Signal-to-Noise Ratio (SNR). This degrades the performance of the AEC system, i.e. the adaptive filter $\hat{\mathbf{h}}(n)$ quickly diverges from its steady-state level and the conversation is therefore disturbed. Hence, it is necessary to have an efficient DTD in order to detect the DT situation before the divergence of the AEC system. Furthermore, there should be an appropriate control mechanism for the the adaptation of the adaptive filter.

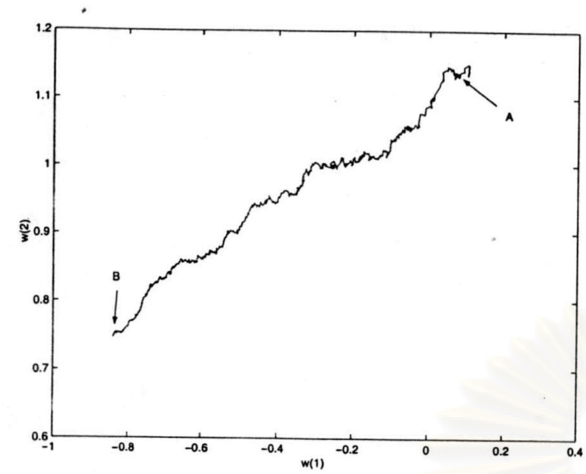
3 The proposed DTD

By considering at the trajectories of the first two coefficients of the adaptive filter with long L in an AEC system, the effect of an ACAEP on the trajectories is illustrated in Fig. 2(a). The strong direction of this trajectory of the adaptive filter is shown. At point A, the estimates of the adaptive filter for the first two coefficients of the unknown AEP in the near-end room is in the SS. After the occurrence of the ACAEP, the trajectories move consistently in one direction towards point B, which is the new optimal value of the AEP estimate. On the other hand, when there is a DT situation, the trajectories of the adaptive filter do not show any particular pattern but fluctuate around an optimal point C, as depicted in Fig. 2(b). This can be concluded that the correlation between the coefficients of the adaptive filter, thus the behavious of the gradient vector, can be utilised to distinguish between the ACAEP and the DT situation.

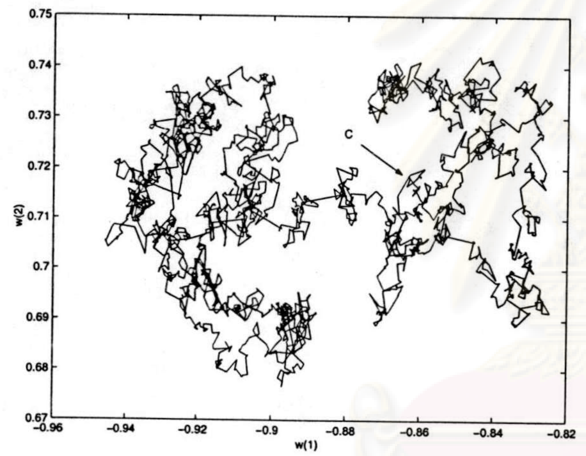
By considering at the gradient vector estimate of the adaptive filter employed in the NLMS algorithm [7], as given by

$$\nabla(n) = \mathbf{x}(n)e(n) \tag{6}$$

this parameter approaches to zero after the adaptive filter has converged. Although there is a DT situation, this relationship of the gradient vector estimate $\nabla(n)$ is still true since the near-end speech $v(n)$ that contains in the error signal is uncorrelated to the far-end speech signal $x(n)$. However, when there is an ACAEP, the above condition does not satisfied any longer due to the correlation of the echo signal $y(n)$ that contained in the error signal $e(n)$ and the far-end speech signal $x(n)$. It is therefore proposed in this paper that the **directivity** of the gradient



(a)



(b)

Figure 2: *Trajectories of the adaptive filter coefficients in (a) the ACAEP situation, (b) the DT situation.*

vector estimate should be used for DT detection. This ensures that the proposed DTD has the ability to distinguish the ACAEP from the DT situation. The directivity can be obtained from the cross-correlation between the average instantaneous gradient vector estimates at time n and $n-1$:

$$\mathbf{g}_1(n) = \bar{\nabla}(n) \cdot \bar{\nabla}(n-1) \tag{7}$$

In addition, the average version of instantaneous gradient vector estimate over a window length K_1 is taken so that the fluctuation of the gradient vector is smoothed out. When there is an ACAEP, the AEC system cannot eliminate the echo signal $y(n)$, hence, the cross-correlation between the avearge gradient vector estimate at two consecutive time, $\mathbf{g}_1(n)$, increases. On the other hand, when there is a DT situation, which is assumed to be after the AEC system has converged, the gradient vector estimate at time n and $n-1$ is no longer correlated to each other. As a result, $\mathbf{g}_1(n)$ stays as low as that in the SS.

Furthermore, the **activity** of the gradient vector estimate is proposed to be used together with the directivity of the gradient vector estimate, $\mathbf{g}_1(n)$, in order to distinguish between the DT situation and the SS condition. This second detector is based on the average of the auto-correlation of the instantaneous gradient vector estimates at time n over the window length K_2, as given by

$$\mathbf{g}_2(n) = \bar{\nabla}(n) \cdot \bar{\nabla}(n) \tag{8}$$

When the adaptive filter has converged, the auto-correlation $\mathbf{g}_2(n)$ is approaching minimum, this means that there is no activity of the gradient vector. However, this parameter increases from the steady-state level when there exists a near-end speech, i.e the activity of the gradient vector is present. Therefore, by employing the detector $\mathbf{g}_2(n)$, the DT and the SS situations can now be separated.

The control mechanisms for the adaptation gain employed in the NLMS algorithm is described as follows. In the transient period, the adaptation gain should be chosen so that fast convergent rate of the adaptive filtering algorithm is obtained. In the SS period, the adaptation gain should be decresead in order to avoid high misalignment of the adaptive filter. During the DT period, slow convergent rate of the adaptive filtering algorithm is chosen to track small changes of the AEP, instead of freezing the adaptation as in the conventional DTD methods. The time-varying adaptation gain $\mu(n)$ used in the update equantion of the NLMS algorithm

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \frac{\mu(n)}{\epsilon + ||\mathbf{x}||^2}\mathbf{x}(n)e(n) \tag{9}$$

when ϵ is a small positive constant and $||\cdot||$ is a Euclidean norm of a vector, is adjusted according to the energy of the microphone signal [8]. This time-varying adaptation gain is defined as

$$\mu'(n) = \frac{\alpha}{\beta + P_d(n)} \tag{10}$$

for small constant α, β, and the energy of the microphone signal can be found recursively by

$$P_d(n) = \lambda_d P_d(n-1) + (1-\lambda_d)d^2(n) \tag{11}$$

where $0 \ll \lambda_p < 1$ is a smoothing factor.

Different situations of the AEC system can be described as the following hypotheses, in terms of the DTD parameters $\mathbf{g}_1(n)$ and $\mathbf{g}_2(n)$.

$$H_0 \;:\; \text{ACAEP, if } \mathbf{g}_1(n) > \Theta_1 \text{ and } \mathbf{g}_2(n) > \Theta_2 \tag{12}$$
$$H_1 \;:\; \text{DT, if } \mathbf{g}_1(n) < \Theta_1 \text{ and } \mathbf{g}_2(n) > \Theta_2 \tag{13}$$
$$H_2 \;:\; \text{SS, if } \mathbf{g}_1(n) < \Theta_1 \text{ and } \mathbf{g}_2(n) < \Theta_2 \tag{14}$$

The threshold Θ_1 and Θ_2 are for the detectors $\mathbf{g}_1(n)$ and $\mathbf{g}_2(n)$, respectively. Hence, the condition for the adaptation gain of the NLMS algorithm is given by

$$\mu(n) = \begin{cases} \alpha & \text{if } \; H_0 \text{ is chosen.} \\ \mu^{'}(n) & \text{if } \; H_1 \text{ is chosen.} \\ \mu^{'} n & \text{if } \; H_2 \text{ is chosen.} \end{cases} \tag{15}$$

When there is an ACAEP, the adaptation gain should be chosen as its maximum value, α. During the DT period, the adaptation gain is variably controlled by the energy of the microphone signal from the near-end room, $\mu^{'}(n)$, as given in eq.(10). In addition, the variable adaptation gain $\mu^{'}(n)$ is employed during the SS to compromise between the convergent rate and the misadjustment level.

In terms of the computational complexity of the proposed DTD, it is insignificant although the window length K_1 and K_2 for averaging the cross-correlation and the auto-correlation can be long. This is because the moving-average characteristics of the cross-correlation and the auto-correlation of the gradient vector estimates can be found recursively, as given by

$$\begin{aligned} \mathbf{g}_1(n) = \frac{1}{K_1}(&\mathbf{g}_1(n-1)K_1 + (\nabla(n)\cdot\nabla(n-1)) \\ &- (\nabla(n-K_1)\cdot\nabla(n-K_1-1))) \end{aligned} \tag{16}$$

and

$$\begin{aligned} \mathbf{g}_2(n) = \frac{1}{K_2}(&\mathbf{g}_2(n-1)K_2 + (\nabla(n)\cdot\nabla(n)) \\ &- (\nabla(n-K_2)\cdot\nabla(n-K_2))) \end{aligned} \tag{17}$$

Note that K_1 and K_2 for each averaging can be chosen differently.

4 Experimental Results

The aim of our experiments was to observe the ability to detect the DT situation and to distinguish between the DT situation and the ACAEP of the proposed DTD. The input signals were white noise sequence of 15,000 samples and a speech signal with sampling frequency $f_s = 8kHz$. The unknown AEP in the near-end room was illustrated in Fig. 3.

The adaptive filter length was chosen to be $L = 256$. Performance was evaluated from the Weight Error Vector Norm (WEVN), which is defined as

$$\text{WEVN}(n) = 10 * \log_{10}\frac{||\mathbf{h}-\hat{\mathbf{h}}(n)||^2}{||\mathbf{h}||^2} \tag{18}$$

The background noise was added to the near-end room of $30dB$ SNR. The ACAEP was introduced at iteration $n = 4000$. Another white noise sequence of the DT-to-far-end-signal ratio (DTFR) of $6dB$ was employed as the near-end signal between $n = 8000$ to $n = 10000$. The NLMS algorithm was employed with and without the proposed DTD for comparison. In addition, the performance

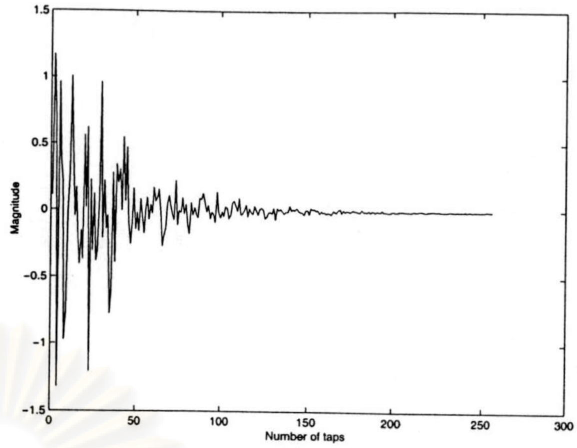


Figure 3: *The impulse response of the AEP in the near-end room.*

of a conventional DTD based on the orthogonality theorem in [4], which freezes the adaptation of the adaptive filter during the DT period, was also compared to the proposed method. This is because the DTD in [3] does not have any indicator to observe whether DTD is on or off. The parameter of the adaptation gain of the NLMS algorithm with and without the proposed DTD was chosen to be $\alpha = 0.5$. The threshold values of $\Theta_1 = 0.23$ and $\Theta_2 = 0.0006$ were set.

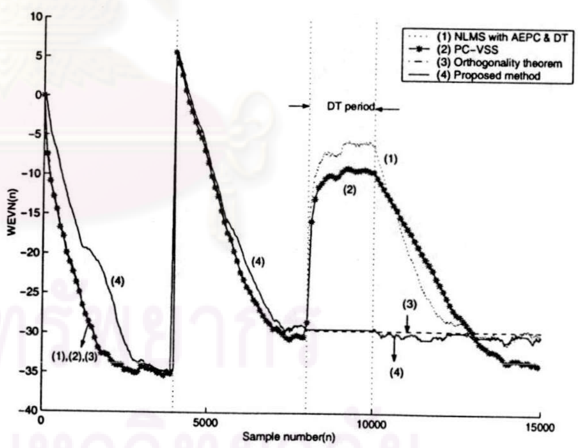


Figure 4: *WEVN performance of the NLMS algorithm with and without the proposed DTD.*

It is shown in Fig. 4 that the convergence rate during the ACAEP of the NLMS algorithm employing the proposed DTD is as fast as that of the conventional NLMS algorithm without any DTD. During the DT period, the WEVN performance of the NLMS algorithm with the proposed DTD does not diverged from its SS, as when there is no DTD. The behaviour of the gradient vector estimates,

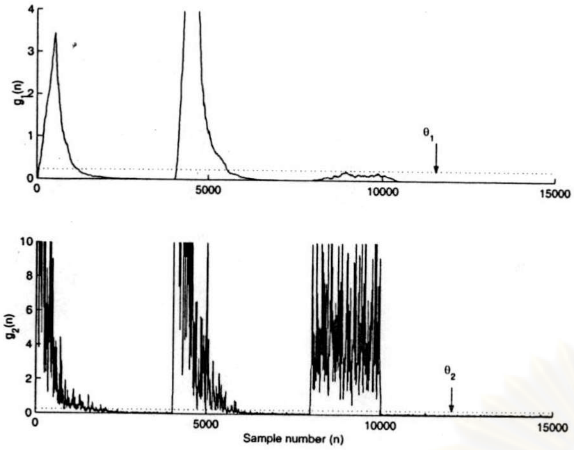


Figure 5: *The DTD* $g_1(n)$ *and* $g_2(n)$.

$g_1(n)$ and $g_2(n)$, are depicted in Fig. 5. The decision of DT detection is illustrated by the counter for the proposed DTD and the DTD in [4] in Fig. 6. The counter of the DTD in [3] cannot be shown as mentioned previously. When the DTD is on, it is represented by 1 and when it is off, 0 is plotted. The proposed method makes correct decision for the DT situation, although there are some false decision in the SS. This could be improved if appropriated values of threshold Θ_1 and Θ_2 are chosen. It is noted that the DTD in [4] still gives false alarm after the DT period.

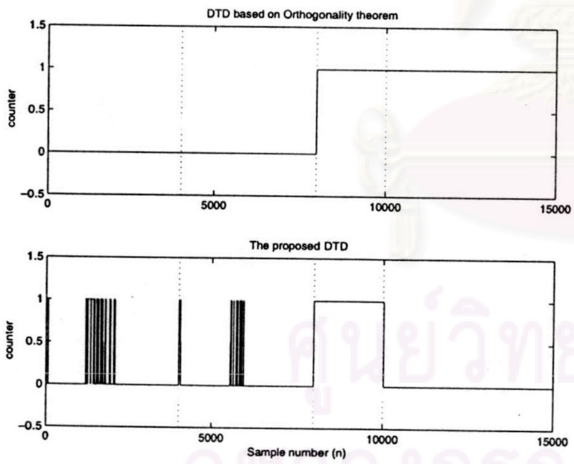


Figure 6: *The decision for DT detection.*

5 Conclusions

The proposed DTD based on the correlation behaviour of the gradient vector estimates has been shown to be capable of detecting and distinguishing between the ACAEP, the DT situation and the SS, with little extra computational complexity in evaluating the DTD. The performance of the AEC system can be improved during the DT period, i.e. no divergence. Moreover the tracking behaviour of the adaptive filter has been illustrated.

6 References

[1] N. Tangsangiumvisai and J.A. Chambers, "Double-talk detection for stereophonic acoustic echo cancellation in a time-varying environment", *EECON'2002*, DS04, pp. 5–9, November 2002.

[2] C. Rohrs and R. Younce, "Double talk detector for echo canceller and method", *U.S. patent 4,918,727*, April 1990.

[3] T. Creasy and T. Aboulnasr, "A projection-correlation algorithm for acoustic echo cancellation in the presence of double talk", *ICASSP'2000 (Istanbul, Turkey)*, vol.1, pp. 436–439, May 2000.

[4] H. Ye and B. Wu, "A New Double-Talk Detection Algorithm Based on the Orthogonality Theorem", *IEEE Transactions on Communications*, vol. 39, no. 11, pp. 1542–1545, November 1991.

[5] C. Carlemalm, F. Gustafsson and B. Wahlberg, "On the problem of detection and discrimination of double talk and change in the echo path", *ICASSP'96 (Atlanta, USA.)*, vol. 5, pp. 2742–2745, May 1996.

[6] H. K. Jung, N. S. Kim and T. Kim, "A new double-talk detector using echo path estimation", *ICASSP'2002 (Orlando, USA.)*, vol. 2, pp. 1897–1900, May 2002.

[7] S. Haykin, "Adaptive Filter Theory", *Prentice Hall*, 1996.

[8] N. Tangsangiumvisai and J.A. Chambers, "A fast converging and stable adaptive filtering algorithm for stereophonic acoustic echo cancellation", *EUSIPCO'2000 (Tampere, Finland)*, vol. 2, September 2000.

Nisachon Tangsangiumvisai received the Ph.D. degree in Electrical and Electronic Engineering from Imperial College, London, U.K., in 2001. She is currently a lecturer at Department of Electrical Engineering, Chulalongkorn University, Thailand. Her research interest includes Adaptive Signal Processing and its applications for telecommunication.

Nataporn Ithisoponkul received the B.Eng degree in Electrical and Electronic Engineering from Khonkaen University, Thailand, in 2001. He is currently persuing his M.Eng degree at the department of Electrical Engineering, Chulalongkorn University, Bangkok, Thailand.

ประวัติผู้เขียนวิทยานิพนธ์

นายณตพร อิทธิโสภณกุล เกิดวันที่ 7 พฤษภาคม พ.ศ. 2522 ที่อำเภอเมือง จังหวัดอุดรธานี เข้าศึกษาในหลักสูตรวิศวกรรมศาสตร์บัณฑิต คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่นในปีการศึกษา 2540 สำเร็จการศึกษาวิศวกรรมศาสตร์บัณฑิต สาขาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่นในปีการศึกษา 2543 เข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตร์มหาบัณฑิต คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2544