國 立 交 通 大 學

電信工程學系

碩 士 論 文

雙邊對話偵測器在迴音消除之性能評比

Doubletalk Detector Performance Evaluation in Acoustic Echo Cancellation

研 究 生： 張 紘 睿

指導教授： 謝世福　教授

中　華　民　國　九十七 年 六 月

雙邊對話偵測器在迴音消除之性能評比

學生:張紘睿　　　　　　　　　　　指導教授:謝世福

國立交通大學電信工程研究所

中文摘要

對於迴音消除而言，當雙邊同時對話時，會使得一些採用誤差回授的適應性濾波器如 LMS 無法追蹤房間的脈衝響應。在本篇論文中，我們將採用修正型相關函數的偵測方法來解決雙邊談話的問題。除此之外，我們延伸一種技術，讓此技術可同時評比雙邊對話偵測器在發生雙邊對話和房間響應突然發生改變的情況，計算此雙邊對話偵測器發生偵測錯誤的機率。另一方面，我們推導使用相關函數的雙邊對話偵測器之偵測錯誤的理論機率，同時還簡化了理論錯誤機率。經過簡化，我們可以看到錯誤機率和假警報機率的關係。我們也提出了一個修正型相關函數可偵測雙邊對話和房間響應突然發生變化。在後面的電腦模擬將會證實我們的推導以及提出的方法。

Doubletalk Detector Performance Evaluation in Acoustic Echo Cancellation

Student: H. R. Chang　　　　**Advisor: S. F. Hsieh**

Department of Communication Engineering

National Chiao Tung University

Abstract

In the adaptive acoustic echo cancellation, double talk can make the adaptive filter diverge from the optimum. In this thesis, the cross-correlation between the microphone signal and estimate echo is used to judge whether double-talk arises. We also derive the theoretical miss probability as a function of false alarm probability. To distinguish the echo path change from double talk and we also propose a modified cross-correlation double talk detector by microphone energy. We not only develop a way to evaluate DTD algorithm whether double-talk or echo path change arises, but also calculate the miss probability. Computer simulations will validate our derivations and proposed methods.

Acknowledgement

First, I would like to express my deepest gratitude to my advisor, Dr. S. F. Hsieh. Without his assistance, enthusiastic, and patience, my thesis would not complete. I also want to think my parents and my sisters. They always encourage me when I am bad mood. Finally, thanks to my lab friends, I can always be happy to do my research

Contents

List of Figures

List of Tables

Chapter 1

Introduction

For hands-free communication systems, it is important to provide users a better quality and comfortable conversation. In these hands-free systems, acoustical echo is a major issue that leads to bad speech quality. An echo canceller removes echo due to echo path coupling between a loudspeaker and microphone. Double talk (DT) is a serious problem in the adaptive acoustic echo cancellation which can fail to trace the room impulse response especially for some error feedback adaptive filters like LMS and RLS [1].

A teleconference system with acoustics echo canceller (AEC) is shown in Fig 1.1 where a linear filter $\hat{h}$ is used to model the echo path h between the speaker and the microphone. Thus the replica of far end speaker's echo $\hat{y}(k)$ is generated, which is subtracted from the echo received by the microphone signal $d(k)$. The AEC filter is typically updated using an adaptive algorithm to account for any changes in the room impulse response.

The implementation of such a system is not as easy as it seems. Because the performance of an algorithm will be affected by long impulse response length for the linear filter, fast convergence characteristic for signal inputs such as speech and fast adaptability to variations in echo path. Among all the adaptive algorithms for AEC, the LMS algorithm and normalized LMS (NLMS) algorithm [1] are popular ones for their simplicity and predictable behavior.

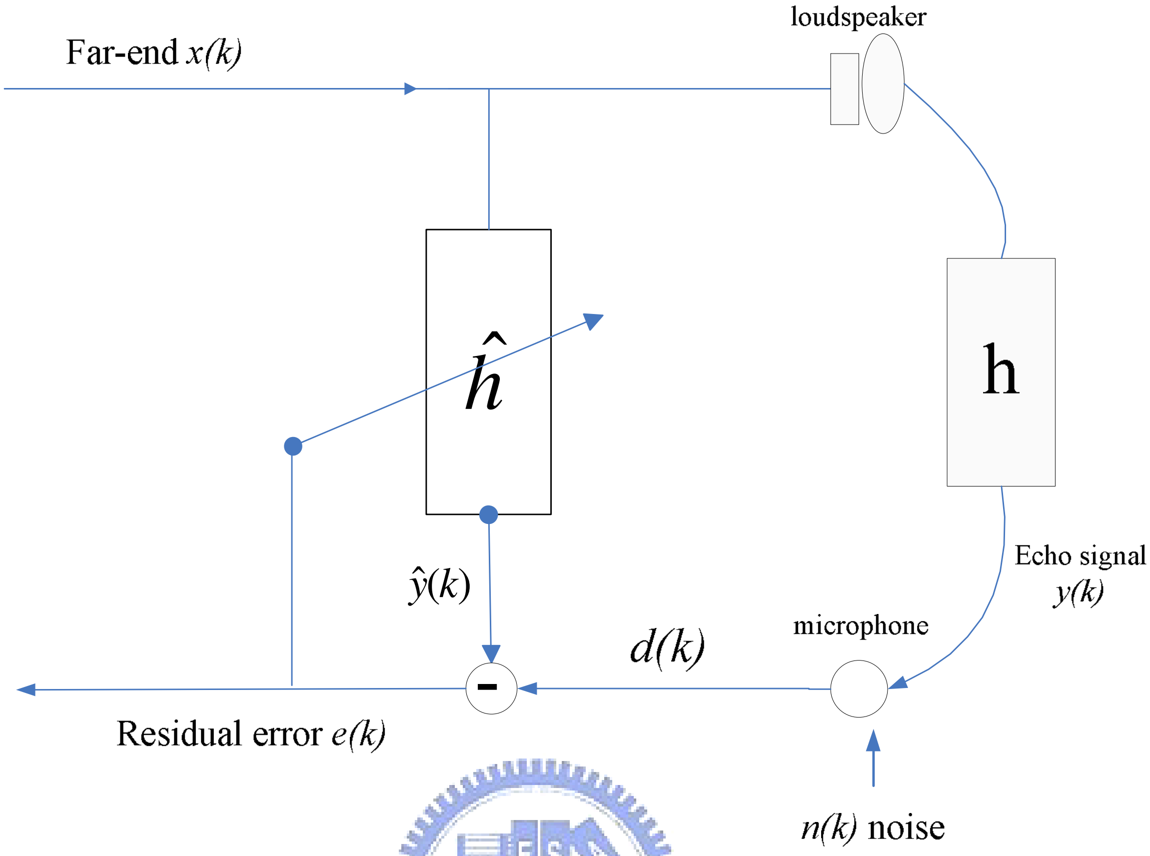


Fig 1.1 The structure of acoustics echo canceller

An adaptive echo canceller updates the tap coefficients of an adaptive filter to model echo path using an error signal $e(k)$ as shown in Fig 1.1. If the tap coefficients are updated during the double-talk situation, which means that microphone input signal includes both near-end talker signal $v(k)$ and echo signal $y(k)$, they can fluctuate greatly or diverge to misestimate the impulse response of echo path. Hence, AEC should stop the filter adaptation during the double talk period.

Several double talk detectors (DTDs) has been proposed. The conventional double talk detection algorithms are classified into several categories.

(I) Level comparison type is used to detect double talk by comparing the microphone signal level [2] or the error signal level [3] with the primary input signal level.

(II) CLMS algorithm [11] is used to distinguish DT from varying echo path and ECLMS has better performance than CLMS but they have the drawback of higher computational complexity.

(III) Cross-correlation type [7] [12] [13] [14] can detect double talk by different correlations. In this thesis, we adopt the cross correlation DTD method which is a better DTD than the other two methods because it is affected very slightly by the volume of the microphone or loudspeaker change.

(IV) Recently also some DTD algorithms have also been developed that are specifically suited for subband [15].

(V) One way to guarantee that the adaptive filter is not unnecessarily halted is to use a secondary FIR filter as in the two-path algorithm [6] [9] [16].

Several doubletalk detectors (DTDs)/step-gain controllers, which halt the adaptation during doubletalk, have been proposed. However, a badly tuned DTD induces the risk of halting the adaptive filter when it should not be halted, e.g., in an echo path change situation. A critical question is that merely measuring these signals cannot discriminate between double-talk and echo-path-change. If echo path change is mislabeled as double-talk, AEC performance degrades.

In Chapter 2, we compare several DTD's using the technique from [8]. The comparison in [8] considered only Geigel and normalized cross correlation DTD. We add different DTDs to compare. We also derive the theoretical cross correlation in double talk or echo path change. The miss probability from [8] is simulated value rather than theoretical value. We derive the theoretical miss probability in double talk period from [8]. The nonlinear loudspeaker effect is also discussed in cross correlation DTD.

In Chapter 3, we modify the cross correlation DTD. The modified cross correlation DTD by microphone energy can decide correct in any double talk or echo path change case. We also propose that the evaluating DTD technique can calculate miss probability when echo path change is present. We also use the variant threshold to improve the performance.

In chapter 4, the simulations follow to verify the results of our analysis and we will compare the simulated and analytical cross correlation. The modified cross correlation DTD will verify in different cases. We will also compare the simulated and analytical miss probability in double talk period. Finally, in chapter 5, the conclusions are given there.

Chapter 2
Double Talk Detector For Acoustic Echo Canceller

In this chapter, the serious problem, double talk, in AEC will be discussed. An adaptive echo canceller [1] updates the tap coefficient of an adaptive filter to model echo path using the error signal $e(k)$ as shown in Fig 2.1. If the tap coefficients are updated during the double talk situation, which means that microphone input signal includes both near-end talker signal $v(k)$ and echo signal $y(k)$, they can fluctuate greatly or diverge to misestimate the impulse response of echo path. Hence, AEC should freeze the filter adaptation during the double talk period.

In Section 2.1, we introduce several double talk detector algorithms. In section 2.2, we compare the Section 2.1 DTD and calculate the miss probability in DT period. In Section 2.3, we modify the DTD to more robust. In Section 2.4, we will derive the theoretical cross-correlation in double talk and echo path change. The theoretical miss probability is derived in Section 2.5. The nonlinear loudspeaker effect on the cross correlation double talk detector is discussed in Section 2.6.

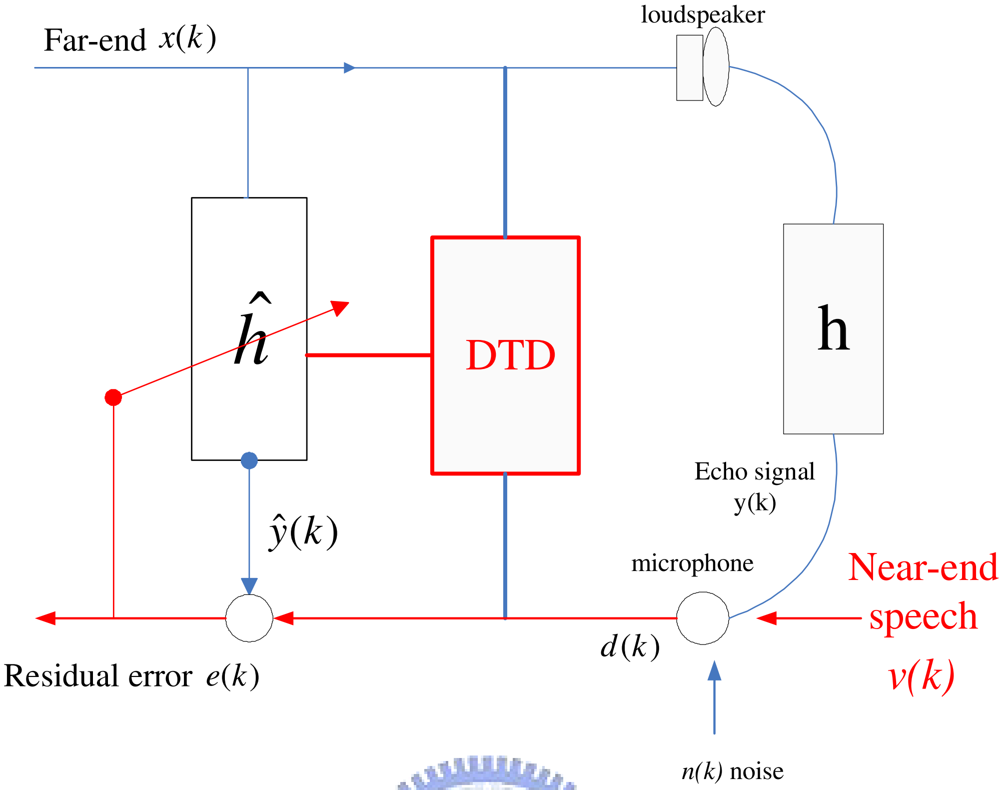


Fig 2.1 Structure of double talk detector

2.1 Double talk detectors algorithms

We have introduced several double talk detectors in introduction. Now, we discuss explicitly the Geigel, gradient vector, two echo path model, and cross correlation DTD.

2.1.1 Geigel DTD

One simple DTD algorithm due to Geigel [2]. The algorithm is given as follows.

$$\xi_{Geigel} = \frac{|d(k)|}{\max\{|x(k-1)|,......,|x(k-N)|\}} \tag{2.1.1}$$

where N is also equal the filter length.

This detection scheme is based on a waveform level comparison between the microphone signal $d(k)$ and the far-end speech $x(k)$ assuming the near-end speech $v(k)$ in the microphone signal will be typically stronger than the echo signal.

When ξ_{Geigel} is larger than the threshold T_{Geigel}, the DTD is decided that double-talk is present. Then the adaptation is halted. T_{Geigel} compensate for the energy level of the echo path response h. However, when the magnitude of $d(k)$ is -6 dB, the Geigel DTD fails to detect the double talk. For an AEC, however, it is not easy to set a universal threshold to work reliably in all the various situations because the loss through the acoustic echo path can vary greatly depending on many factors.

2.1.2 Gradient vector correlation DTD

Rohrs and Younce [4] specifically targeted the DT problem. Their algorithm considered the correlation between the instantaneous gradient estimation and the average of previous estimation. The gradient vector is defined as follows:

$$\nabla(k) \triangleq x(k) \cdot e(k)$$

$$\overline{\nabla}(k) = \overline{\nabla}(k-1) + \nabla(k) - \nabla(k-B)$$

$$S(k) = \nabla(k)\overline{\nabla}(k-1)$$

where $B > 0$ will depend on the filter length.

$\nabla(k)$ means that the instantaneous far end signal $x(k)$ multiplied by the momentary error $e(k)$. $\overline{\nabla}(k)$ is the average of previous estimation.

If the correlation $S(k)$ measured is larger than the threshold $T_{Gradient}$, the detector adjusts the weights using a fixed step size LMS update; otherwise, the coefficients are frozen. But, the algorithm is effective for a small adaptive filter length. However, their performance degraded considerably with long adaptive filter length.

Creasy and Aboulnasr [5] improved the above problem. The algorithm used a variable step size NLMS-based approach by the gradient correlation. The algorithm was given as follows:

$$\begin{aligned}
&\nabla(k) = x(k)\cdot e(k) \\
&\bar{\nabla}(k) = \bar{\nabla}(k-1) + \nabla(k) - \nabla(k-B) \\
&\bar{\bar{\nabla}}(k) = \nabla(k)\cdot \bar{\nabla}(k-1) \\
&p(k) = \beta p(k-1) + (1-\beta) sign[\bar{\bar{\nabla}}(k)] \\
&\mu(k) = \alpha\cdot\mu(k-1) + (1-\alpha) sign[p(k)] p^2(k)
\end{aligned} \qquad (2.1.2)$$

$\mu(k)$ is step size. When the step size $\mu(k)$ becomes very small, the DTD decides that double-talk is present. Even the adaptation keeps updating coefficients; the adaptive filter will not diverge. This algorithm is more robust.

2.1.3 Two Echo Path Model

Another structure of DTD, in Fig 2.2, is two echo path model [6]. This structure is a good choice to implement in real environment for its excellent stability. It is based on a structure of two path model, a background filter h_{BG} and a foreground filter h_{FG}.

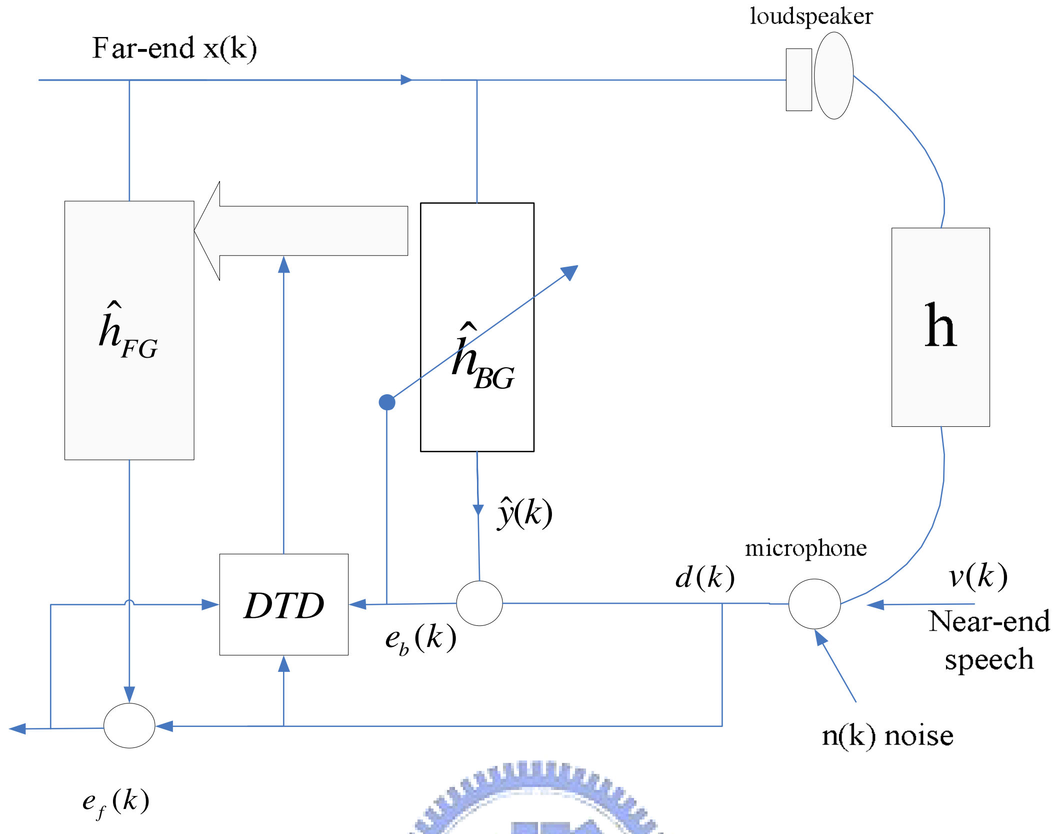


Fig. 2.2 Structure of the two echo path model

If the background filter h_{BG} is estimated to have better performance than the foreground filter h_{FG}, its filter coefficients are copied to the foreground filter. The double talk detector is controlled by comparisons between the short-term powers of the signals, $d(k)$, $x(k)$, $e_f(k)$ and $e_b(k)$.

The update conditions for the foreground filter are basically as given by

$$a = \frac{P_d(k)}{P_X(k)} \text{ and } b = \frac{P_{e_b}(k)}{P_d(k)} \text{ and } c = \frac{P_{e_b}(k)}{P_{e_f}(k)} \qquad (2.1.3)$$

where $P_X(k) = \frac{1}{M}\sum_{i=0}^{M-1} X^2(k-i)$, M is update interval.

When a, b and c is larger than the threshold T_a, T_b and T_c at the same time, the DTD decides that double-talk is present . The background filter will not be copied to foreground filter. The foreground filter retains its convergent coefficients.

2.1.4 Cross Correlation DTD

Ye and Wu [7] proposed a double-talk algorithm based on the cross-correlation between $x(k)$ and $e(k)$. However, the cross correlation DTD has a numerous correlation. We can use the different correlation based on $d(k)$, $\hat{y}(k)$, $x(k)$, and $e(k)$. Therefore, we choose one of the cross-correlations. We use the cross-correlation between $d(k)$ and $\hat{y}(k)$.The cross-correlation $\rho_{d,\hat{y}}$ is defined as:

$$\rho_{d,\hat{y}}(k) = \frac{P_{d,\hat{y}}(k)}{\sqrt{P_d(k)P_{\hat{y}}(k)}} \tag{2.1.4}$$

where

$$P_{d,\hat{y}}(k) = (1-\lambda)P_{d,\hat{y}}(k-1) + \lambda d(k)\hat{y}(k)$$

$$P_d(k) = (1-\lambda)P_d(k-1) + \lambda d^2(k)$$

$$P_{\hat{y}}(k) = (1-\lambda)P_{\hat{y}}(k-1) + \lambda \hat{y}^2(k)$$

λ is the forgetting factor , $0 < \lambda < 1$

When $\rho_{d,\hat{y}}$ is below threshold $T_{d\hat{y}}$, the DTD is decided that double-talk is present. Then the adaptation is halted.

2.2 Comparisons of Double talk detectors

There have been several algorithms to detect double talk in an acoustics echo canceller. Jun H. Cho and R. Morgan [8] proposed an objective technique to evaluate double talk detectors. The technique could calculate the double talk detector miss probability. In [8], they compare the Geigel DTD with the normalized cross correlation DTD. In this section, we extend the technique to evaluate the Gegel, gradient vector, Two echo path model, and cross correlation DTD. We compare the four kinds of the DTD in this section.

In Section 2.1.3, we introduced the two echo path model. Two echo path model can detect double talk by Eq (2.1.3). We modify the condition to decide double talk.

We only use $a = \frac{P_d(k)}{P_X(k)}$, $b = \frac{P_{e_b}(k)}{P_d(k)}$ to detect double talk. In order to simplify, we use the forgetting factor to smooth a and b.

The objective technique first step is to calculate threshold under fixed false alarm probability. The false alarm probability is measured as the proportion of the far-end speech in which doubletalk remains declared when there is no near-end speech.

The probability of false alarm at each threshold point is calculated as

$$P_f = \frac{\sum \phi \cdot \bar{x} \cdot \bar{v}}{N} \tag{2.2.1}$$

where ϕ is the DTD output, $\bar{x}$ is the activity detector output from Fig2.3. N is the length of the entire far-end speech signal x.

From Fig 2.4, the output of the far end signal speech activity detector is either one or zero, and also the near end speech. We can find that the output is zero when the far end signal is silent. From Fig 2.5, the logical AND with the activity of is necessary to disregard false alarms during innocuous periods of inactivity. Then, the threshold is determined to achieve the given false alarm probability.

Second step is to calculate miss probability. The miss probability is measured as the proportion of near-end speech duration that remains undetected at different levels of near-end to far-end speech energy ratio (NFR=$\frac{\sigma_v^2}{\sigma_x^2}$).

Once the threshold T is determined, the near-end speech is applied at different attenuation levels, and the detection procedure runs again. The miss probability is calculated as follows.

$$P_m = 1 - \frac{\sum \phi \cdot \overline{x} \cdot \overline{v}}{\sum \overline{x} \cdot \overline{v}} \qquad (2.2.2)$$

In Chap 4, we will use the technique to compare the four kind of DTD.

The complete DTD evaluation technique is summarized as follows.

1) Set near-end signal v = 0.

    (a) Select threshold T.

    (b) Compute false alarm probability P_f using (2.2.1).

    (c) Repeat steps a, b over a range of threshold values.

    (d) Select threshold value that corresponds to $P_f = 0.1$

2) Select NFR value.

    (a) Add near-end signal

    (b) Compute P_m using (2.2.2).

    (c) Repeat steps a, b, c over all conditions.

3) Repeat step 2 over a range of NFR values.

4) Plot average P_m as a function of NFR.

Table 2.1 DTD evaluation procedure in case of double talk

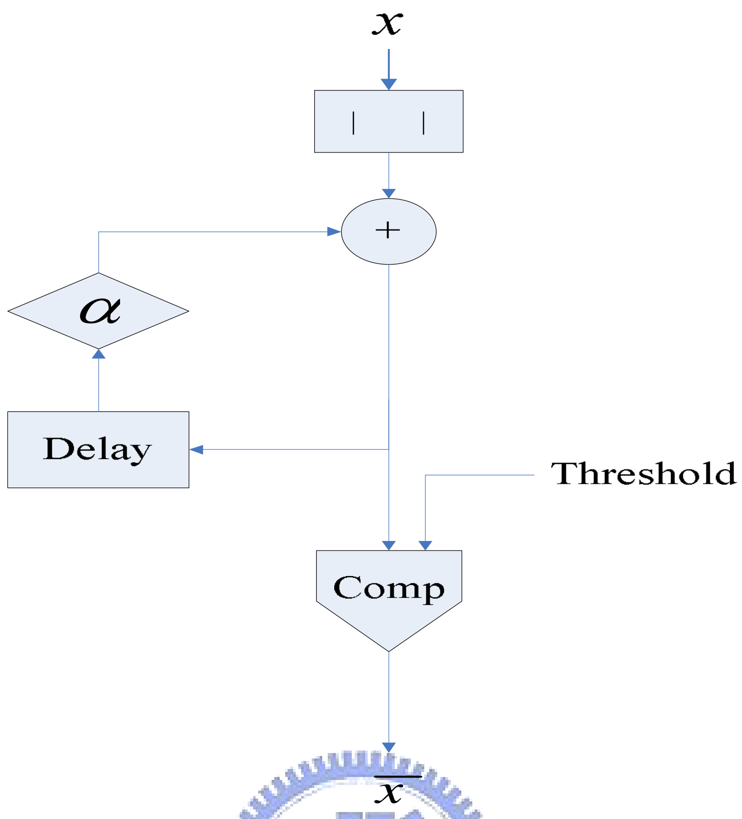


Fig 2.3 Speech activity detector

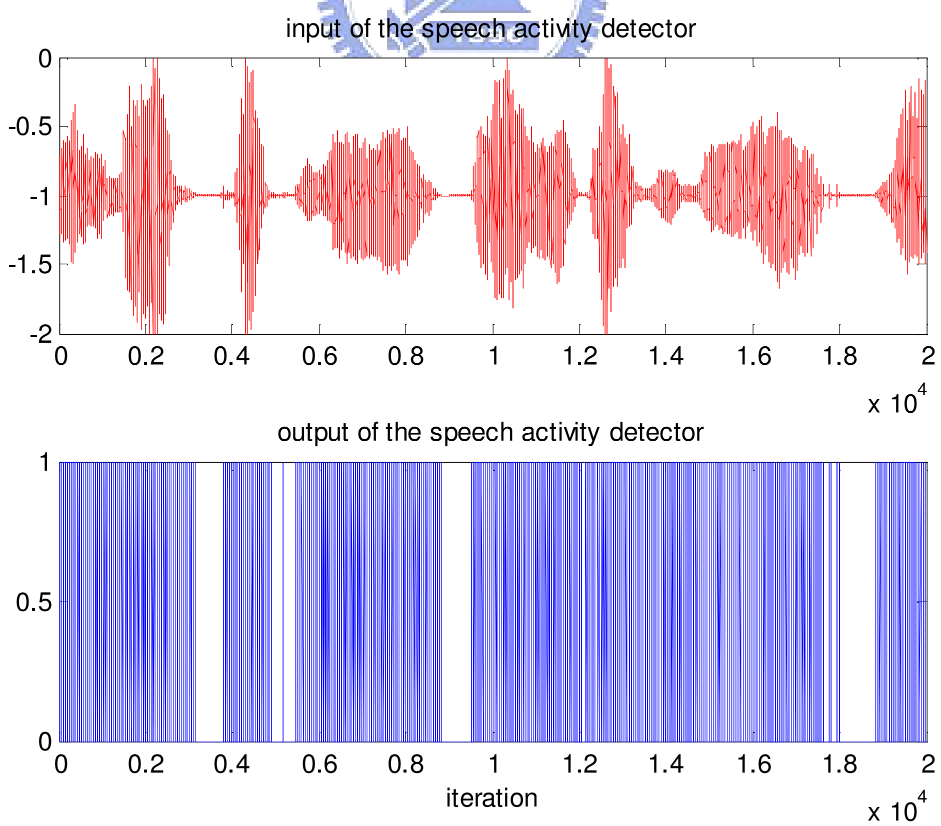


Fig 2.4 Input and output of the speech activity detector

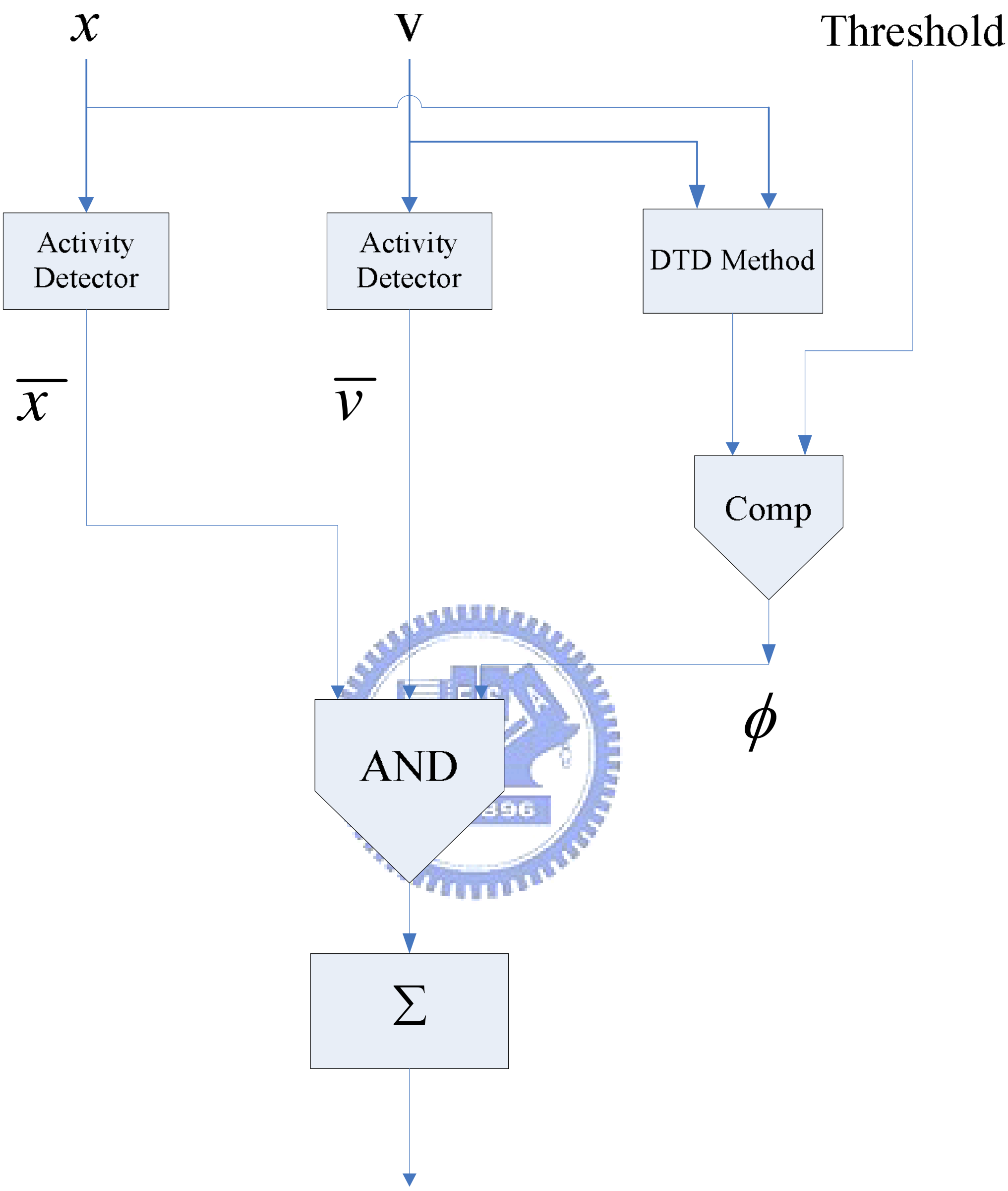


Fig 2.5 Evaluation procedure of DTD

2.3 The robust double talk detector

In section 2.1, we have introduced several double talk detectors, including Geigel DTD, two echo path model, and cross correlation DTD. These DTD decide double talk by the threshold in Fig 2.6. But, the DTD decision is dichotomous. We can modify the decision more mildly to the robust double talk detector. The robust DTD can adapt the coefficients whatever double talk or single talk is present in Fig 2.7. Therefore, the robust DTD can alleviate the miss probability, and we do not set the sensitive threshold to avoid detecting error.

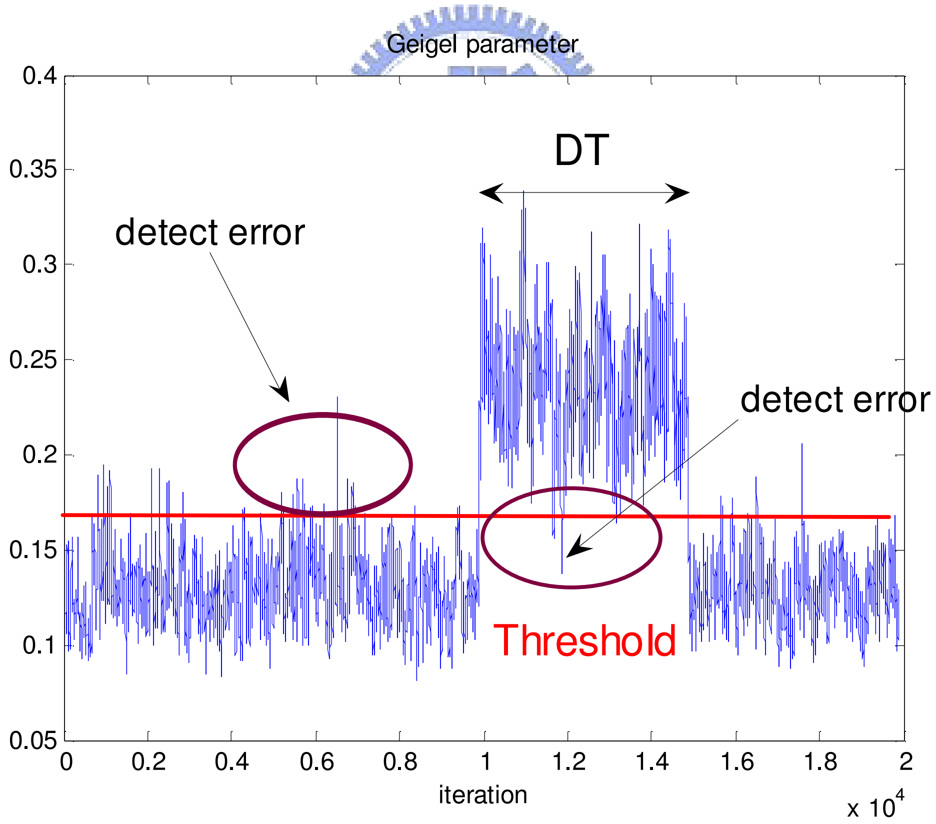


Fig 2.6 The difference ξ_{Geigel} in single talk and double talk

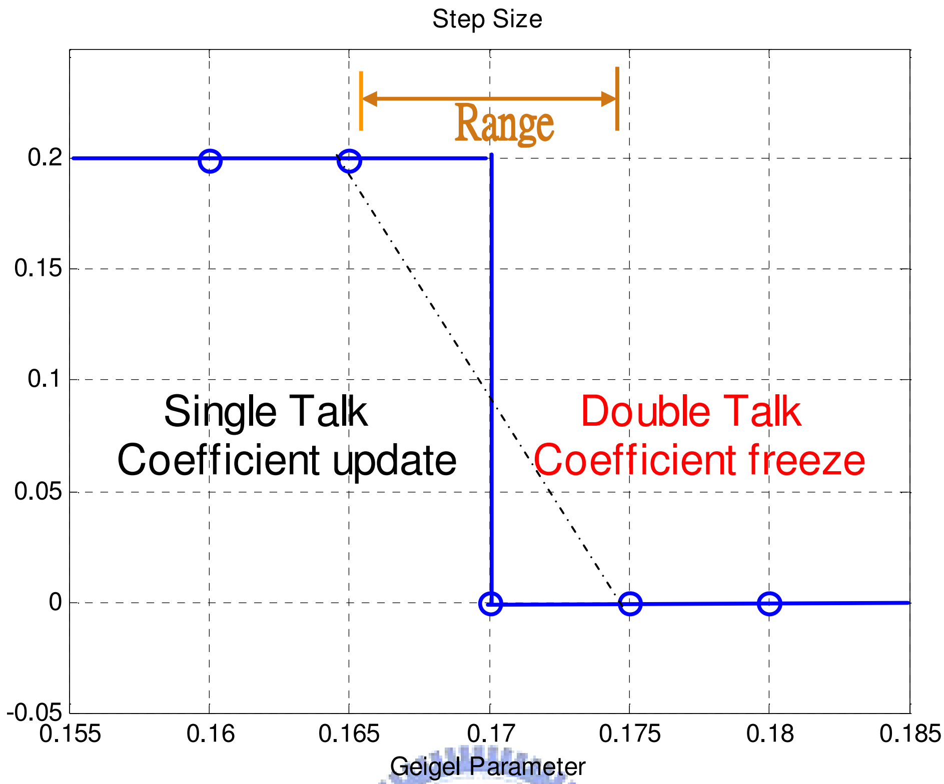


Fig 2.7 The relation of the step size and ξ_{Geigel}

2.3.1 Robust Geigel DTD

In section 2.1.1, we have introduced the Geigel DTD. The detector uses the microphone signal and far end signal energy ration to decide double talk in Eq. (2.1.1). We can find that ξ_{Geigel} is quick change in the each iteration whatever double talk or single talk and ξ_{Geigel} is increasing in DT period. If ξ_{Geigel} is larger than the T_{Geigel}, we can detect the double talk and stop adapting. On the contrary, the filter continues adapt coefficients. But, ξ_{Geigel} is the decision is dichotomous. Now, we can adapt the coefficients by ξ_{Geigel} even double talk is present. This means that the step size is very small in double talk period.

Before, we derive the ξ_{Geigel} PDF. We modify the Geigel DTD. We show that the modified Geigel DTD is given as follows.

$$P_d(k) = (1-\lambda)\cdot P_d(k-1) + \lambda \cdot abs(d(k))$$

$$P_x(k) = (1-\lambda)\cdot P_x(k-1) + \lambda \cdot \max(abs(x(k)))$$

$$\rho_{Geigel}(k) = \frac{P_d(k)}{P_x(k)} \tag{2.3.1}$$

Because, ξ_{Geigel} is fast change in every time. We smooth ξ_{Geigel} to avoid a sudden change. ρ_{Geigel} is the microphone signal amplitude and far end signal ration. However, ρ_{Geigel} is difficult to analyze. Therefore, we modify the Geigel DTD criterion to energy ratio.

The modify Geigel DTD is given as follows.

$$P_d(k) = (1-\lambda)\cdot P_d(k-1) + \lambda \cdot d^2(k)$$

$$P_x(k) = (1-\lambda)\cdot P_x(k-1) + \lambda \cdot \max(x^2(k))$$

$$P_\alpha(k) = \frac{P_d(k)}{P_x(k)} \tag{2.3.2}$$

The P_α is very like ξ_{Geigel}, and ξ_{Geigel} is amplitude ratio, and the P_α is energy ration. Now, we derive the P_α PDF [17]. From Fig 2.7, we can find that the detector make some error decision. However, we can analyze P_α PDF to alleviate the detected error.

First, we analyze the microphone signal. The microphone signal includes the echo signal, near end speech, and noise.

$$d(k) = h^T x(k) + v(k) + n(k)$$

$$d^2(k) = \|h\|^2 x^2(k) + v^2(k) + n^2(k)$$

We assume that the far end signal $x(k)$, near end speech $v(k)$, and noise $n(k)$ are normal distribution and $x^2(k)$, $v^2(k)$,and $n^2(k)$ are Chi-Square.

$$r.v\ \ x \quad f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{x^2}{2\sigma_x^2}} \qquad \text{Gaussian distribution} \qquad (2.3.3)$$

$$r.v\ \ x^2 \quad f_x(x) = \frac{1}{\sqrt{2\pi x\sigma_x^2}} e^{-\frac{x}{2\sigma_x^2}} \qquad \text{Chi-Square distribution} \qquad (2.3.4)$$

We also assume the microphone signal energy and $\max(x^2(k))$ are also Chi-Square. The P_α PDF is given as follows.

$$f_{P_\alpha}(P_\alpha) = \frac{1}{\pi\sqrt{\sigma_d^2 P_\alpha}\left(\frac{1}{\sigma_x^2} + \frac{P_\alpha}{\sigma_d^2}\right)} \qquad (2.3.5)$$

From Eq. (2.3.5), the expectancy of the random variable E_α is calculated as follows. We also assume the far end signal energy is equal one.

$$E[P_\alpha] = \mu_{P_\alpha} = \frac{(4-\pi)\sigma_d^2}{2\pi} \qquad (2.3.6)$$

$$\begin{aligned} \operatorname{var}[P_\alpha] &= \sigma_{P_\alpha}^2 = E[(P_\alpha - \mu)^2] \\ &= \frac{\sigma_d^2}{\pi}\left(-0.116 + \frac{0.02863}{\sigma_d^2} + 0.2375\sigma_d^2\right) \end{aligned} \qquad (2.3.7)$$

where the microphone signal energy $\sigma_d^2 \approx \|h\|^2 \sigma_x^2 + \sigma_v^2 + \sigma_n^2$

If $\sigma_d^2 = 1$ means only single talk, $E[P_\alpha] \approx 0.13$, and $\operatorname{var}[P_\alpha] = 0.05$. We can find the converged value is 0.13 in modified Geigel DTD. If the P_α is closed to the converged value. That means that the filter has converged. If the P_α is larger than the converged value, this means double talk is possible present.

We set the soft threshold near P_α mean. However, from (2.3.5), we can find that P_α PDF is not symmetric. The soft threshold lies between $\mu_{P_\alpha} - 2\sigma_{P_\alpha}^2 = 0.12$ and $\mu_{P_\alpha} + \sigma_{P_\alpha}^2 = 0.135$

In Fig 2.8, we discuss the relation of the step size and P_α. From Fig 2.8, Geigel DTD decides double talk by one threshold. But, this decision is too hard. However, the robust Geigel DTD set the buffer range to advance DTD performance. This means that the detector has the double threshold. Using the buffer range can robust DTD.

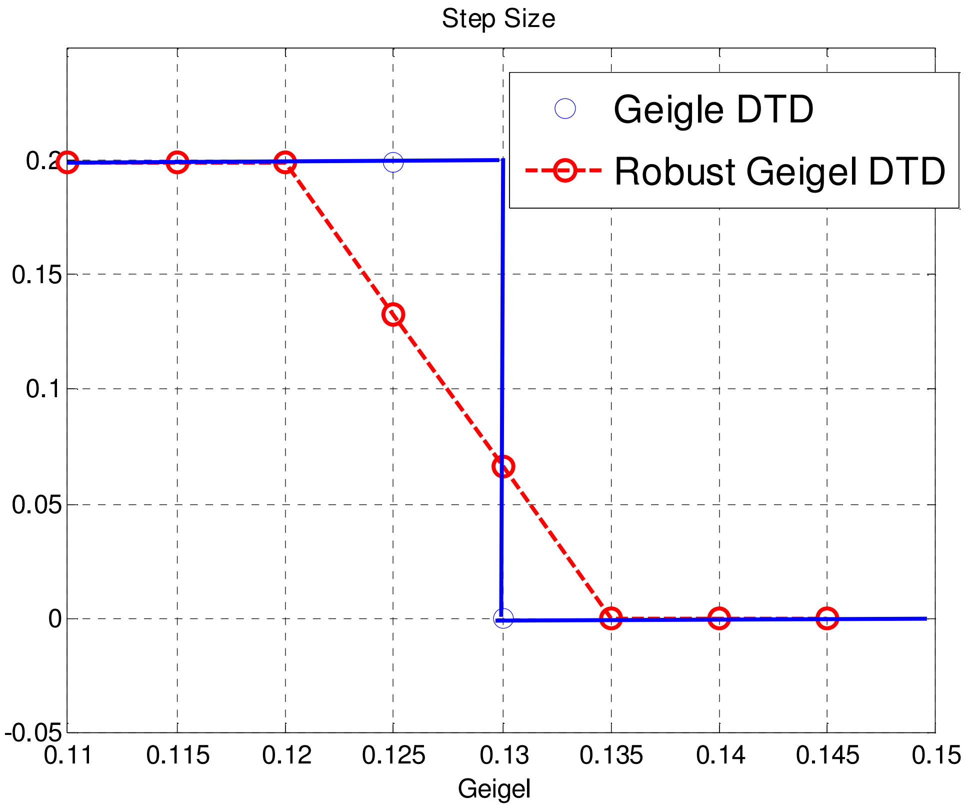


Fig 2.8 Comparison the step size of the Geigel and robust Geigel DTD

2.3.2 Robust cross-correlation DTD

Now, we discuss the cross-correlation DTD in this section. Before Section 2.3.1, we set the buffer range to robust Geigel DTD. Using the same idea, we can robust the cross-correlation DTD. We also extend two kinds of the buffer range.

In Section 2.1.4, we introduced the cross-correlation DTD. Now, we modify the decision rule. The threshold set to be 0.7. Therefore, we set the buffer range near 0.7. One set the buffer range lies between 0.6 and 0.8, and another range lies between 0.65 and 0.8.

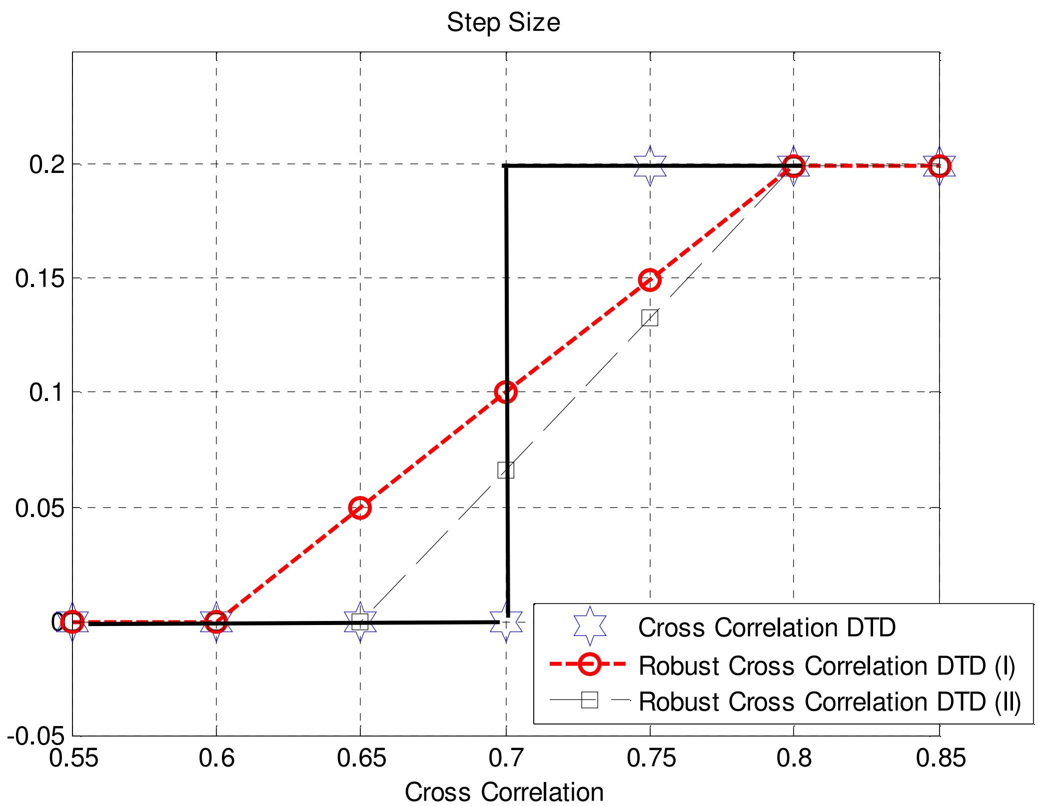


Fig 2.9 Comparison the step size of the cross and robust cross-correlation DTD

We can analyze the DTD parameter PDF to set the buffer range. If we have the buffer range, the robust DTD performance is better than the conventional DTD. The result will be simulated in Chapter 5.

2.4 Theoretical analysis of cross-correlation DTD

In this section, we analyze two kinds of cross correlation DTD. The mic/AEC correlation and mic/error correlation is discussed. When we analyze the correlation in double talk/echo path change, we assume that the adaptive filter has converged in single talk. We also assume that the far end signal $x(k)$, noise $n(k)$ and near end signal $v(k)$ are white Gaussian signals and $x(k)$, $n(k)$, $v(k)$ are mutually independent. If we know the exact cross correlation, we can set the appropriate threshold.

2.4.1 Correlation of microphone and estimated signals

Before, we discussed the cross-correlation DTD. We found the correlation values will decrease whether double-talk or echo path change arises. But we do not know the exact degraded value in any double talk or echo path change degree.

(I) Cross-correlation ρ_{DT} in double talk

Cross-correlation in Eq. (2.1.4) is rewritten here for simplicity.

$$\rho_{d,\hat{y}}(k) = \frac{P_{d,\hat{y}}(k)}{\sqrt{P_d(k)P_{\hat{y}}(k)}}$$

By assuming, the adaptive filter $\hat{h}$ is closer to echo path channel h. That means $\hat{h} \approx h$. The cross correlation DTD use the forgetting factor λ to smooth $\rho_{d,\hat{y}}(k)$ and implement online DTD. But, the forgetting factor is hard to analysis. Fortunately, the forgetting factor $\rho_{d,\hat{y}}(k)$ is closer to the expectation $\rho_{d,\hat{y}}(k)$ when the $\rho_{d,\hat{y}}(k)$ converged. It means

$$\rho_{d,\hat{y}}(k) = \frac{P_{d,\hat{y}}(k)}{\sqrt{P_d(k)P_{\hat{y}}(k)}} \approx \frac{E[d \cdot \hat{y}]}{\sqrt{E[d^2]E[\hat{y}^2]}} \triangleq \rho_{DT}$$

First, we examine $P_{d,\hat{y}}(k)$.

$$P_{d,\hat{y}} = E[d(k) * \hat{y}(k)] = E[(y(k)+v(k)+n(k)) * \hat{y}(k)]$$

$$\approx E[y(k)\hat{y}(k)]$$

And $y(k) = h^T x(k)$, $\hat{y}(k) = \hat{h}^T x(k)$ then $P_{d,\hat{y}}$ can express.

$$P_{d,\hat{y}} \approx E[h^T x(k) x^T(k) \hat{h}]$$

We assume $\hat{h} \approx h$,

$$P_{d,\hat{y}} \approx E[h^T x(k) x^T(k) h] \approx h^T h \sigma_x^2 \qquad (2.4.1)$$

Next, we proceed to find $P_d(k)$

$$P_d = E[d(k) \cdot d(k)] = E[(y(k)+v(k)+n(k)) \cdot (y(k)+v(k)+n(k))]$$

$$\approx E[h^T x(k) x^T(k) h + v^2(k) + n^2(k)]$$

$$= h^T h \sigma_x^2 + \sigma_v^2 + \sigma_n^2 \qquad (2.4.2)$$

Last,

$$P_{\hat{y}} = E[\hat{y}(k) * \hat{y}(k)^T] = E[\hat{h}^T x(k) x^T(k) \hat{h}]$$

$$= h^T h \sigma_x^2 \qquad (2.4.3)$$

Finally, we combined (2.4.1) (2.4.2) with (2.4.3)

$$\rho_{DT} = \frac{E[d \cdot \hat{y}]}{\sqrt{E[d^2]E[\hat{y}^2]}}$$

$$= \begin{cases} \dfrac{\sigma_x^2 \cdot h^T h}{\sqrt{(\sigma_x^2 \cdot h^T h + \sigma_n^2) \cdot (\sigma_x^2 \cdot h^T h)}} = \dfrac{1}{\sqrt{1 + \dfrac{\sigma_n^2}{\sigma_x^2 \cdot h^T h}}} & \text{(Single-talk)} \qquad (2.4.4) \\ \dfrac{\sigma_x^2 \cdot h^T h}{\sqrt{(\sigma_x^2 \cdot h^T h + \sigma_v^2 + \sigma_n^2) \cdot (\sigma_x^2 \cdot h^T h)}} \dfrac{1}{\sqrt{1 + \dfrac{\sigma_n^2 + \sigma_v^2}{\sigma_x^2 \cdot h^T h}}} & \text{(Doube-talk)} \qquad (2.4.5) \end{cases}$$

From (2.4.4), we find that the correlation value is much closer to 1 when single talk is present. From (2.4.5), it can find decreases in accordance with the near end signal variance in double talk period. By theoretical analysis, we can know ρ_{DT} in any DT situation.

(II) Cross-correlation ρ_{EPC} n echo path change h_c

Next, we analyze the cross correlation when the converged filter $\hat{h}$ in single talk undergoes an abrupt echo path change h_c. This means $h \approx \hat{h} \neq h_c$ where h is origin echo path.

First, we analyzed $P_{d,\hat{y}}(k)$

$$P_{d,\hat{y}} = E[d(k)\cdot \hat{y}(k)] = E[(y(k)+n(k))\cdot \hat{y}(k)]$$

$$\approx E[(y(k)\cdot \hat{y}(k)] = h_c^T \hat{h}\sigma_x^2 \tag{2.4.6}$$

Next, we proceed to find $P_d(k)$

$$P_d = E[d(k)\cdot d(k)] = E[(y(k)+n(k))\cdot (y(k)+n(k))]$$

$$\approx h_c^T h_c \sigma_x^2 + \sigma_n^2 \tag{2.4.7}$$

Last,

$$P_{\hat{y}} = E[\hat{y}(k) * \hat{y}(k)] = \hat{h}^T \hat{h}\sigma_x^2 \tag{2.4.8}$$

Finally, we combine (2.4.6) and (2.4.7) with (2.4.8)

$$\rho_{d,\hat{y}}(k) = \frac{P_{d,\hat{y}}(k)}{\sqrt{P_d(k)P_{\hat{y}}(k)}} \approx \frac{E[d\ y]}{\sqrt{d^2 y^2}} \triangleq \rho_{EPC}$$

$$= \frac{h_c^T \hat{h}\sigma_x^2}{\sqrt{(\sigma_x^2 \cdot h_c^T h_c)\cdot(\sigma_x^2 \cdot \hat{h}^T \hat{h})}} = \frac{h_c^T \hat{h}}{\sqrt{(h_c^T h_c)\cdot(\hat{h}^T \hat{h})}}$$

$$\approx \frac{h_c^T h}{\sqrt{(h_c^T h_c)\cdot(h^T h)}} = \rho_{hh_c} \tag{2.4.9}$$

From (2.4.9), we find $\rho_{d,\hat{y}}(k)$ also the decreases in accordance with origin channel h and changed channel h_c correlation. In double talk or echo path change, $\rho_{d,\hat{y}}(k)$ will decrease. The cross-correlation DTD is difficult to detect double talk and echo path change.

2.4.2 Correlation of microphone and error signals

Now, we discuss another double talk detector by the microphone signal $d(k)$ and error $e(k)$ correlation.

The micro/error correlation DTD algorithm is expressed as:

$$\rho_{d,e}(k) = \frac{P_{d,e}(k)}{\sqrt{P_d(k)P_e(k)}} \qquad (2.4.10)$$

where

$$P_{d,e}(k) = (1-\lambda)P_{d,e}(k-1) + \lambda d(k)e(k)$$

$$P_d(k) = (1-\lambda)P_d(k-1) + \lambda d^2(k)$$

$$P_e(k) = (1-\lambda)P_e(k-1) + \lambda e^2(k)$$

λ is forgetting factor , $0 < \lambda < 1$

(I) Cross-correlation in double talk

$$\rho_{d,e}(k) = \frac{P_{d,e}(k)}{\sqrt{P_d(k)P_e(k)}} \approx \frac{E[d \cdot e]}{\sqrt{E[d^2]E[e^2]}}$$

First, we examine $P_{d,e}(k)$, and define $\Delta h \triangleq h - \hat{h}$.

$$P_{d,e} = E[d(k)\cdot e(k)] = E[(h^T x(k)+n(k))\cdot(h^T x(k)-\hat{h}^T x(k)+n(k))]$$

$$= E[(h^T x(k)+n(k))\cdot(\Delta h x(k)+n(k))]$$

$$P_{d,e} \approx E[h^T x(k) x^T(k)\Delta h + n(k) n^T(k)] = h^T \Delta h \sigma_x^2 + \sigma_n^2 \qquad (2.4.11)$$

Next, we proceed to find $P_d(k)$

$$P_d = E[d(k)\cdot d(k)] = E[(h^T x(k)+n(k))\cdot(h^T x(k)+n(k))]$$

$$\approx h^T h \sigma_x^2 + \sigma_n^2 \qquad (2.4.12)$$

Last

$$P_e = E[e(k)\cdot e(k)] \approx \Delta h^T \Delta h \sigma_x^2 + \sigma_n^2 \qquad (2.4.13)$$

Finally, we combined (2.4.11) (2.4.12) and (2.4.13) to get

$$\rho_{d,e}(k) \triangleq \frac{P_{d,e}(k)}{\sqrt{P_d(k)P_e(k)}}$$

$$\approx \frac{h^T \Delta h \sigma_x^2 + \sigma_n^2}{\sqrt{(h^T h \sigma_x^2 + \sigma_n^2)(\Delta h^T \Delta h \sigma_x^2 + \sigma_n^2)}}$$

$$= \frac{h^T \Delta h \sigma_x^2 + \sigma_n^2}{\sqrt{(h^T \Delta h \sigma_x^2 + \sigma_n^2)^2 + (h-\Delta h)^2 \sigma_x^2 \sigma_n^2}}$$

Before we assumed the adaptive converged $\hat{h} = h$, $\Delta h \approx 0$.

$$\rho_{d,e}(k) = \frac{h^T \Delta h \sigma_x^2 + \sigma_n^2}{\sqrt{(h^T \Delta h \sigma_x^2 + \sigma_n^2)^2 + h^T h \sigma_x^2 \sigma_n^2}}$$

$$= \frac{1}{\sqrt{1 + \frac{h^T h \sigma_x^2 \sigma_n^2}{(h^T \Delta h \sigma_x^2 + \sigma_n^2)^2}}}$$

$$\rho_{\mathrm{d,e}}(k) \approx \begin{cases} \dfrac{1}{\sqrt{1+\dfrac{h^T h\sigma_x^2}{\sigma_n^2}}} & \text{(single talk)} \qquad (2.4.14) \\ \dfrac{1}{\sqrt{1+\dfrac{h^T h\sigma_x^2}{(\sigma_n^2+\sigma_v^2)}}} & \text{(Double talk)} \qquad (2.4.15) \end{cases}$$

From (2.4.14), the single talk is very close 0 when SNR is very large. This means that the microphone signal is significantly different from the error. So, the correlation value is small. From (2.4.15), we can find that the near end speech energy increases the noise power. Because the near end speech is view as the noise in AEC, so the noise power adds the near end speech energy in double talk period. Similarly, we can find the correlation value in (2.4.15) becomes large in double talk period.

(II) Cross-correlation in echo path change

First, we examine $P_{d,e}(k)$

$$P_{d,e} = E[d(k)\cdot e(k)] = E[(h_c^T x(k)+n(k))\cdot((h_c^T-\hat{h})x(k)+n(k))]$$

$$\approx h_c^T(h_c^T-\hat{h})\sigma_x^2+\sigma_n^2 = h_c^T\Delta h\sigma_x^2+\sigma_n^2 \qquad (2.4.16)$$

Next, we proceed to find $P_d(k)$

$$P_d = E[d(k)\cdot d(k)] \approx h_c^T h_c\sigma_x^2+\sigma_n^2 \qquad (2.4.17)$$

Last

$$P_e = E[e(k)\cdot e(k)] \approx \Delta h^T\Delta h\sigma_x^2+\sigma_n^2 \qquad (2.4.18)$$

Finally, we combined (2.4.16) (2.4.17) and (2.4.18) to get

$$\rho_{d,e}(k) \triangleq \frac{P_{d,e}(k)}{\sqrt{P_d(k)P_e(k)}}$$

$$\approx \frac{h^T \Delta h \sigma_x^2 + \sigma_n^2}{\sqrt{(h^T h \sigma_x^2 + \sigma_n^2)(\Delta h^T \Delta h \sigma_x^2 + \sigma_n^2)}}$$

$$\approx \frac{h_c^T \Delta h \sigma_x^2}{\sqrt{(h_c^T h_c^T \sigma_x^2)(\Delta h^T \Delta h \sigma_x^2)}}$$

$$\approx \rho_{h_c^T \Delta h} \qquad (2.4.19)$$

From (2.4.19), when echo path change h_c is present, $\rho_{d,e}(k)$ is equal to the correlation of the change channel h_c and the difference Δh in the new channel and origin channel h. From (2.4.15) and (2.4.19), $\rho_{d,e}(k)$ will decrease whatever double talk or echo path change. The mic/error correlation DTD is difficult to detect the two situations.

2.5 Theoretical analysis of miss probability in double talk period

In Section 2.2, we introduced the technique to evaluate double talk detectors, and calculated the miss probability. But, the miss probability is done by numerical simulation in Fig 2.10. In this section, we will derive the theoretical miss probability. We can analyze the DTD parameter probability density function (PDF). We will derive the PDF of the cross-correlation $\rho_{d,\hat{y}}(k)$, an important DTD parameter. From the PDF, the miss probability is calculated.

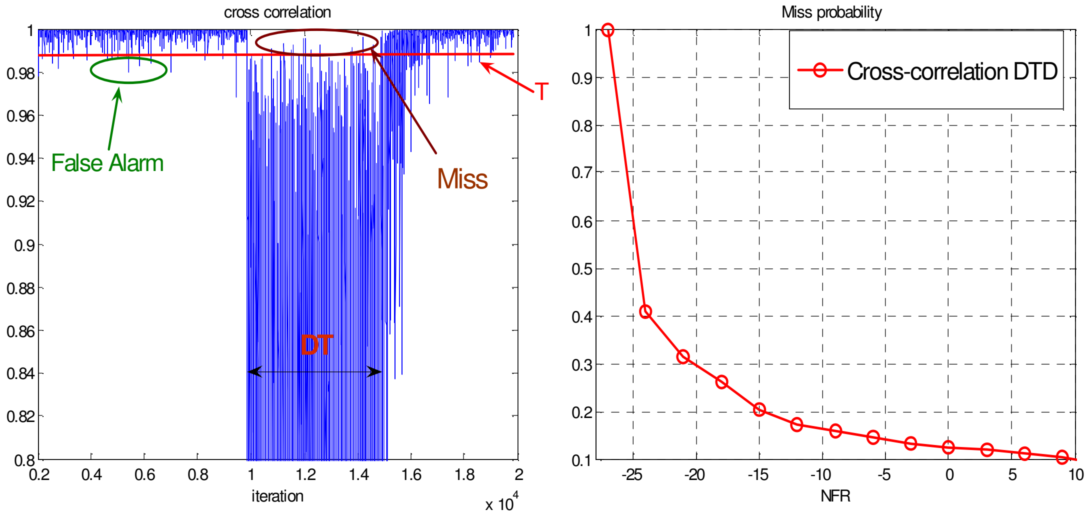


Fig 2.10 The methods of calculating miss probability

Now, we analyze the correlation $\rho_{d,\hat{y}}(k)=\dfrac{P_{d,\hat{y}}(k)}{\sqrt{P_d(k)P_{\hat{y}}(k)}}$ from (2.1.4). Because $\rho_{d,\hat{y}}(k)$ use forgetting factor to smooth. But, $\rho_{d,\hat{y}}(k)$ is a random variable. Then, by discarding the forgetting factor in (2.1.4), where the moment random variable

$$\rho_{d,\hat{y}}(k)=\frac{d(k)\cdot\hat{y}(k)}{\sqrt{d^2(k)\cdot y^2(k)}}=\frac{(h^T x(k)+v(k)+n(k))\cdot(\hat{h}^T x(k))}{\sqrt{(h^T x(k)+v(k)+n(k))^2\cdot(\hat{h}^T x(k))^2}}$$

$h\approx\hat{h}$ is assumed as the filter converges in single talk.

$$\rho_{d,\hat{y}}(k)\approx\frac{\|h\|^2 x^2(k)}{\sqrt{(\|h\|^2 x^2(k)+v^2(k)+n^2(k))(\|h\|^2 x^2(k))}}=\frac{1}{\sqrt{1+\dfrac{v^2(k)+n^2(k)}{\|h\|^2 x^2(k)}}} \qquad (2.5.1)$$

From (2.5.1), we simplify $\rho_{d,\hat{y}}(k)$ is a function of the random variables $x^2(k)$, $v^2(k)$, and $n^2(k)$. Since the random variable $x(k)$ is normally distributed, and random variable $x^2(k)$ is Chi-Square distributed. The complete DTD evaluation technique [8] calculates the miss probability. Before measuring miss probability, the DTD threshold is predetermined to meet the given false alarm probability.

First, we set that near end signal is zero ($v(k)$ =0). The correlation

$$\rho_{d,\hat{y}}(k)=\frac{h^T xxh}{\sqrt{(h^T xx^T h+n^2)\cdot(h^T xx^T h)}}=\frac{1}{\sqrt{1+\frac{n^2}{|h|^2 x^2}}} \tag{2.5.2}$$

In order to find the DTD threshold, we must derive (2.5.2) PDF.

Now, we need the $h^T hx^2$ PDF. We also define $s=\beta x^2 \triangleq h^T hx^2$, where $\beta = h^T h$, whose PDF is also Chi-Square distribution.

$$f_s(s)=\frac{1}{\sqrt{2\pi\beta s}}e^{-\frac{s}{2\beta}}$$

We define random variable $z \triangleq \frac{n^2}{s}$. By the derivation, the random variable PDF z is calculated as follows.

$$f_z(z)=\frac{1}{\pi\sqrt{\sigma_n^2\beta z}(\frac{1}{\beta}+\frac{z}{\sigma_n^2})}$$

Now, the cross correlation is simplified. It is relation with the random variable z.

$$\rho_{d,\hat{y}}(k)=\frac{1}{\sqrt{1+z}}$$

We continue to transform random variable. Defining random variable $w \triangleq \sqrt{1+z}$, the random variable w PDF is given.

$$f_w(w)=\frac{2w}{\pi}\cdot\frac{1}{\sqrt{\beta\sigma_n^2(w^2-1)}}\cdot\frac{1}{(\frac{1}{\beta}+\frac{w^2-1}{\sigma_n^2})} \quad w \ge 1$$

Then the correlation PDF is simplified. $\rho_{d,\hat{y}}(k)=\frac{1}{w}$.

Last, we define $\rho \triangleq \frac{1}{w}$. The random variable $\rho_{d,\hat{y}}$ PDF is given as follows.

$$f_{\rho_{d,y}}(\rho)=\frac{2}{\pi}\cdot\frac{1}{\sqrt{\beta\sigma_n^2(1-\rho^2)}}\cdot\frac{1}{(\frac{\rho^2}{\beta}+\frac{1-\rho^2}{\sigma_n^2})} \quad , 0<\rho<1 \tag{2.5.3}$$

We have the correlation PDF without near end signal, so we can calculate the theoretical DTD threshold under fixed false alarm probability. The false alarm probability P_f means that the detectors decide double talk when double talk is not present.

$$P_f = P(\text{DT is detected} \mid \text{DT Not happens})$$

If the correlation is below threshold, the detector will decide double talk. If we know the fixed false probability, we can calculate the theory threshold by (2.5.4).

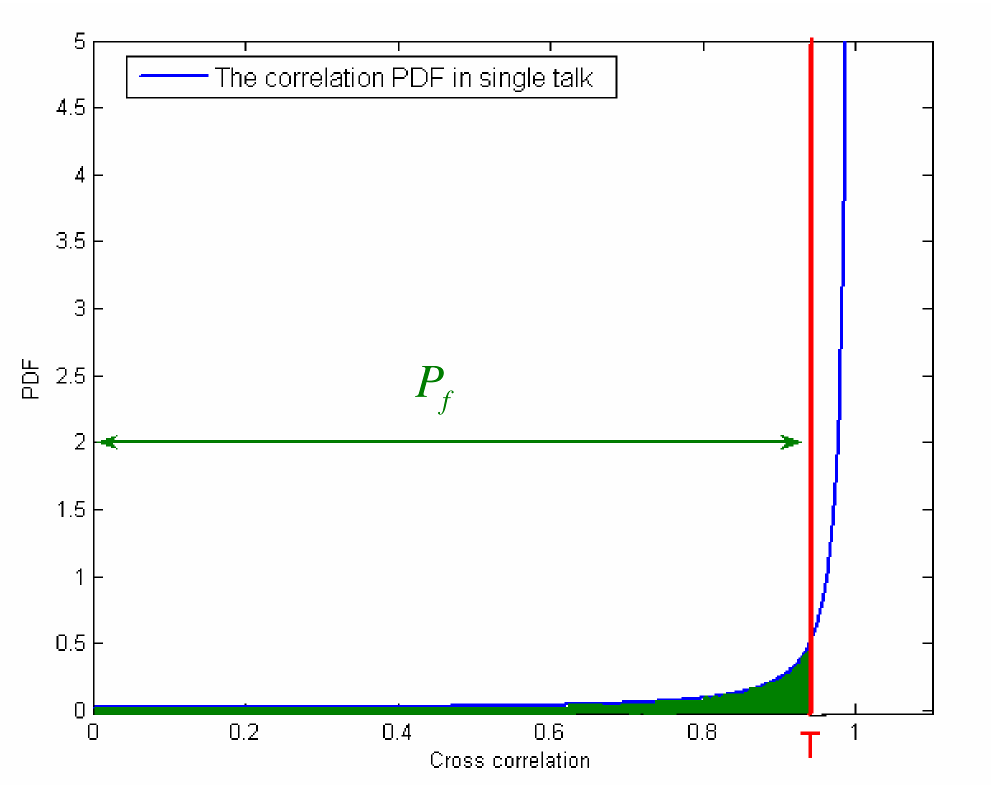


Fig 2.11 The PDF of $\rho_{d,\hat{y}}$ in single talk

In Fig 2.11, the threshold is chosen, and the false alarm probability is also determined. The dotted line area is equal the false alarm probability.

$$P_f \triangleq \int_0^T f_{\rho_{d,y}}(\rho)d\rho = \int_0^T \frac{2}{\pi} \cdot \frac{1}{\sqrt{\beta\sigma_n^2(1-\rho^2)}} \cdot \frac{1}{(\frac{\rho^2}{\beta}+\frac{1-\rho^2}{\sigma_n^2})} d\rho \qquad (2.5.4)$$

T is theoretical threshold, we assume $\beta \triangleq h \cdot h^T$

$$P_f = \frac{2}{\pi} ArcTan[\frac{T}{\sqrt{\sigma_n^{-2}\beta(1-T^2)}}] \qquad (2.5.5)$$

$$T = \sqrt{\frac{\sigma_n^{-2}\beta[\tan(\frac{P_f \pi}{2})]^2}{\{(\sigma_n^{-2}\beta[\tan(\frac{P_f \pi}{2})]^2)+1\}}}$$

$$T = \sqrt{\frac{\alpha^2}{\alpha^2+1}} \approx 1 - \frac{1}{2\alpha^2} = 1 - \frac{2\sigma_n^2}{\pi} \cdot \frac{1}{P_f^2} \qquad (2.5.6)$$

where $$\alpha \triangleq \sigma_n^{-1} \|h\| \tan(\frac{P_f \pi}{2}) \approx (\frac{P_f \pi \sigma_n^{-1} \|h\|}{2})$$

From (2.5.5), we can find the theoretical threshold as a function of the false alarm probability. In (2.5.6), we simplify the theoretic. The simplicity can more easy to see the relation between T and P_f .Now, we added near-end signal $v(k)$ in (2.5.2), and the $\rho_{d,\hat{y}}(k)$ PDF would change.

$$\rho_{d,\hat{y}}(k) = \frac{1}{\sqrt{1+\frac{n^2+v^2}{h^T h x^2}}} \qquad (2.5.7)$$

Comparing with (2.5.2), (2.5.7) has an added near-end signal random variable. We *assume the random variable* $r = n^2 + v^2$ *is still Chi-square distribution* for simplicity.

Now, the PDF of $\rho_{d,\hat{y}}(k)$ with DT can be got.

$$f_{\rho_{d,y}}(\rho) = \frac{2}{\pi} \cdot \frac{1}{\sqrt{\beta(\sigma_v^2+\sigma_n^2)(1-\rho^2)}} \cdot \frac{1}{(\frac{\rho^2}{\beta} + \frac{1-\rho^2}{(\sigma_v^2+\sigma_n^2)})}, \quad 0 < \rho < 1 \qquad (2.5.8)$$

Comparing (2.5.3) with (2.5.8), we can find that the difference of the PDF change only the noise power from the near end signal energy and noise power. Now, we can calculate the miss probability P_m from (2.5.8).

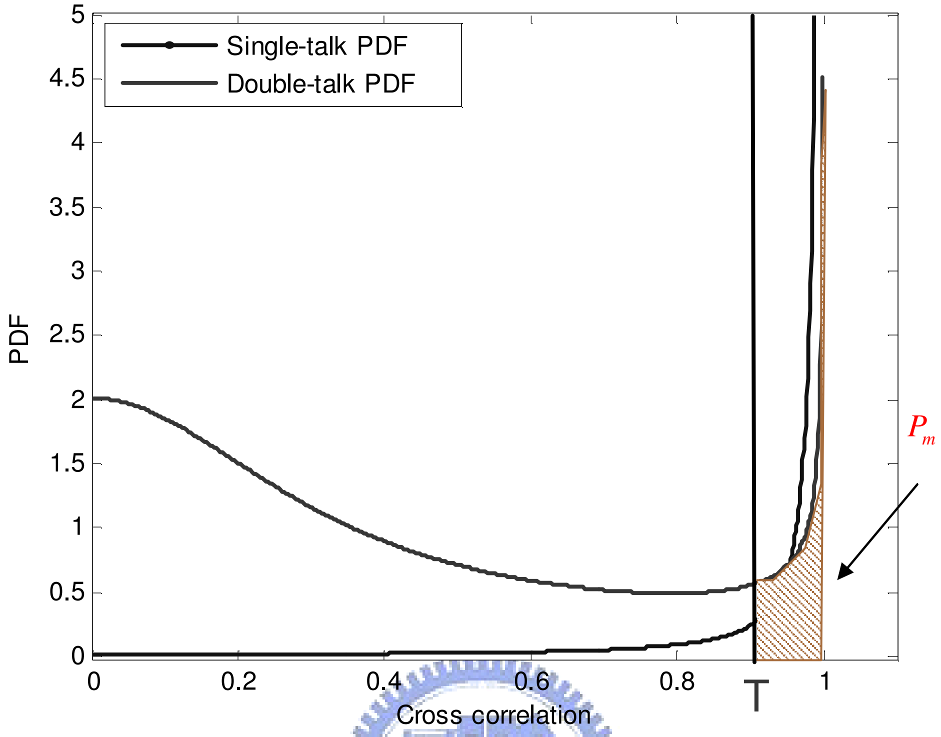


Fig 2.12 The PDF of $\rho_{d,\hat{y}}$ in DT

$P_m \triangleq P(\text{DT is Not detected} \mid \text{DT happens})$

The PDF of $\rho_{d,\hat{y}}$ in DT is in Fig 2.12, and the threshold is also the same with in Fig 2.10. When $\rho_{d,\hat{y}}$ is below threshold, the detector decides DT in double talk period. On the contrary, double talk occurs that the correlation is above threshold. The detector deices no double talk, so the detector occur error. From Fig 2.11, we define the miss probability from (2.5.9). The dotted line area is equal the miss probability.

$\int_0^T f_{\rho_{dy}}(\rho,\sigma_v^2)d\rho$ means that the detector decide correct probability. The correct probability means that the DTD detect double talk in double talk period. Therefore, the miss probability P_m = 1 – correct probability.

$$P_m = 1 - \int_0^T f_{\rho_{dy}}(\rho,\sigma_v^2)d\rho = \int_T^1 f_{\rho_{dy}}(\rho,\sigma_v^2)d\rho \qquad (2.5.9)$$

where T is theoretical threshold. Because (2.5.9) is too complicated, so it can not be written a closed form. From (2.5.9), we can find the miss probability depends on near end signal energy and threshold. In Section 2.4, we analyze the mean of $\rho_{d\hat{y}}$. In Section 2.5, we deeply analyze $\rho_{d\hat{y}}$.

Using the integration definition, we can approximate (2.5.9).

$$P_m = \int_T^1 f_{\rho_{d\hat{y}}}(\rho_{d\hat{y}})d\rho \quad , \quad T \approx 1$$

$$\approx f_{\rho_{d\hat{y}}}(\rho_{d\hat{y}} = \frac{1+T}{2}) \cdot (1-T)$$

$$\approx \kappa \cdot (1-T) \approx \kappa \frac{2\sigma_n^2}{\pi} \cdot \frac{1}{P_f^2} \approx \kappa' \frac{1}{P_f^2} \tag{2.5.10}$$

In (2.5.10), we can find a square inverse law between the false alarm and miss probability. We can use the law to simplify the computation. Without solving the DTD parameter PDF, we can find κ' to get the miss probability.

2.6 Nonlinear loudspeaker effect for DTD

Before we discussed that the acoustic echo path cancellers use linear adaptive filter structures in double talk period to model the acoustic path of the loudspeaker enclosure microphone system, such as the FIR filter $\hat{h}(k)$ described by its impulse response. However, low-cost applications employ small loudspeakers operating beyond their range of linear transduction, and mobile communication terminals may be designed to tolerate clipping of large amplitudes in the amplifier to achieve high sound levels [10].

Unlike the earlier linear loudspeaker, now, the echo component has nonlinear part and linear part. The cross-correlation double talk detector can detect double talk in linear loudspeaker. But for nonlinear loudspeaker in Fig 2.10, can the cross correlation DTD detect double talk?

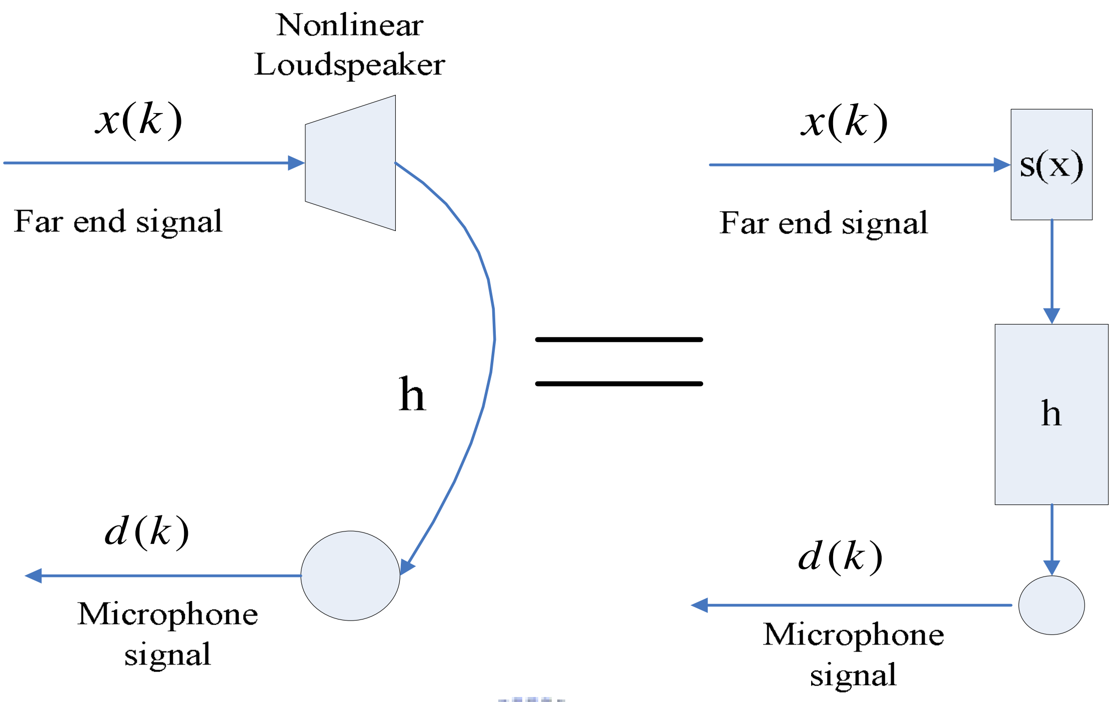


Fig 2.13 Structure of the nonlinear loudspeaker

Next, we analyze that the cross-correlation detector in case of a nonlinear loudspeaker. First, we assume that the nonlinear function in Fig 2.13 is $s(x) = a_1 x + a_3 x^3$. $s(x)$ is output signal from nonlinear loudspeaker, x is far end signal, and a_1 , a_3 is nonlinear function coefficients. We can find that nonlinear function include with linear and nonlinear part.

Now, we analyze $P_{d,\hat{y}}(k)$

$$P_{d,\hat{y}} = E[d(k) * \hat{y}(k)] = E[(y(k) + v(k) + n(k)) \cdot \hat{y}(k)]$$

$$= E[(h^T (a_1 x + a_3 x^3) + v(k) + n(k)) \cdot \hat{h}^T x(k)]$$

$$\approx h^T \hat{h} (a_3 \cdot E[x^3(k) \cdot x^T(k)] + a_1 \cdot E[x(k) x^T(k)]) \tag{2.6.1}$$

Then, we focus on $P_d(k)$

$$P_d \triangleq E[d(k) * d(k)]$$

$$= a_3^2 E[x^3(k) x^3(k)^T] + |h|^2 a_1^2 \cdot E[x(k) x^T(k)]$$

$$+2|h|^2 a_1 a_3 E[x^4(k)] + \sigma_n^2 + \sigma_v^2 \qquad (2.6.2)$$

Last

$$P_{\hat{y}} = E[\hat{y}(k) \cdot \hat{y}(k)] = E[(\hat{h}^T x(k)) \cdot (\hat{h}^T x(k))]$$

$$= E[\hat{h}^T x^2(k) \hat{h}] = \hat{h}^T \hat{h} \sigma_x^2 \qquad (2.6.3)$$

From (2.6.1) (2.6.2) and (2.6.3), $\rho_{d,\hat{y}}(k)$ becomes

$$\rho_{d,\hat{y}}(k) = \frac{P_{d,\hat{y}}(k)}{\sqrt{P_d(k) P_{\hat{y}}(k)}}$$

$$= \frac{(a_3 E[x^4] + a_1 \sigma_x^2)}{\sqrt{(\frac{\sigma_n^2 + \sigma_v^2}{h^T h} + (2a_1 a_3 E[x^4] + a_3^2 E[x^6]) + a_1^2 \sigma_x^2}}$$

$$= \frac{(1 + \frac{a_3}{a_1 \sigma_x^2} E[x^4])}{\sqrt{\frac{\sigma_n^2 + \sigma_v^2}{a_1^2 \sigma_x^2 h^T h} + 2\frac{a_3}{a_1 \sigma_x^2} E[x^4] + \frac{a_3^2}{a_1^2 \sigma_x^2} E[x^6] + 1}}$$

We assume $a_1 \gg a_3$,and $h^T h = |h|^2 = 1$. Before the far end signal $x(k)$ is white signal, $E[x^4] = 3$, and $E[x^2] = \sigma_x^2$.

$$\rho_{d,\hat{y}} \approx \frac{(1 + \frac{3a_3}{a_1 \sigma_x^2})}{\sqrt{1 + \frac{\sigma_n^2 + \sigma_v^2}{a_1^2 \sigma_x^2 h^T h} + 6\frac{a_3}{a_1 \sigma_x^2}}} \qquad (2.6.4)$$

From (2.6.4), we can find that the ratio $\frac{a_3}{a_1}$ affects the cross-correlation DTD in nonlinear loudspeaker. If the $\frac{a_3}{a_1}$ ratio is larger, $\rho_{d,\hat{y}}(k)$ is smaller in double talk period.

In Section 2.4, we also derive the theoretic cross correlation in linear loudspeaker. If a_1=1, and a_3=0, it means that the loudspeaker has only linear part.

$$\rho_{d,\hat{y}} = \frac{1}{\sqrt{1+\frac{\sigma_n^2+\sigma_v^2}{h^T h}}} \tag{2.6.5}$$

We can find that (2.6.5) is the same (2.4.5). This means that we derive the theoretic correlation in nonlinear loudspeaker is accurately. And we can find that the nonlinear loudspeaker will affect the cross-correlation value.

Chapter 3
The Modified Cross Correlation Double Talk Detector by Microphone Energy

In Chapter 2, we have derived the theoretical correlations, and miss probability. But, there is a serious problem that we discussed in section 3.1. When the echo path change happens, the correlation value will be decreased, like in double talk period. The detector detects error between double talk and echo path change. This distinction is important because the adaptive filter coefficients should be continuously updated during the echo path change but not during the double talk period. In Section 3.2, we will propose the modified cross-correlation double talk detector. The modified cross correlation DTD uses the microphone energy to distinguish the echo path change from double talk.

In Section 3.3, we will propose the technique to evaluate DTD in different echo path change. Before the Morgan's technique [8] only can evaluate the DTD in double talk period. We can combine our technique with Morgan's technique to evaluate the DTD in double talk and echo path change. In Section 3.4, the variant threshold can improve the DTD performance. Because, the cross correlation DTD is very sensitive. The cross correlation DTD is more robust by the variable threshold.

3.1 Cross-correlation DTD in echo path change and double talk

One of problems in double talk detector is that there is difficult to distinguish the echo path change [3] from DT. This distinction is important because the adaptive filter coefficients should be continuously updated during the echo path change but not during the double talk periods. On the contrary, when there is an abrupt change of the echo path change (EPC) in the near-end room, the adaptive filter with fast rate of convergence is required to track the echo path change. It is therefore necessary for a DTD to be able to distinguish between the DT situation and the echo path change in order to obtain appropriate tracking performance of the adaptive filter. For both cases of DT and echo path change, the misadjustment of the adaptive filter, and thus the error signal, is drastically increased. Thus, the error signal cannot be used as a DTD alone since it cannot distinguish between these two events.

For an example, in Fig 3.1, the cross correlation $\rho_{d\hat{y}}(k)$ is decreasing in double talk period. However, $\rho_{d\hat{y}}(k)$ is also decreasing when echo path change is present. From Fig 3.1, double talk is 1k from 1.5k, and echo path change occurs in 2.2k. So, the cross correlation double talk detector can not distinguish between DT and EPC. The variation of $\rho_{d\hat{y}}(k)$ has four cases. Therefore, the conventional cross correlation DTD can not distinguish the four cases. The conventional cross correlation DTD is not robust.

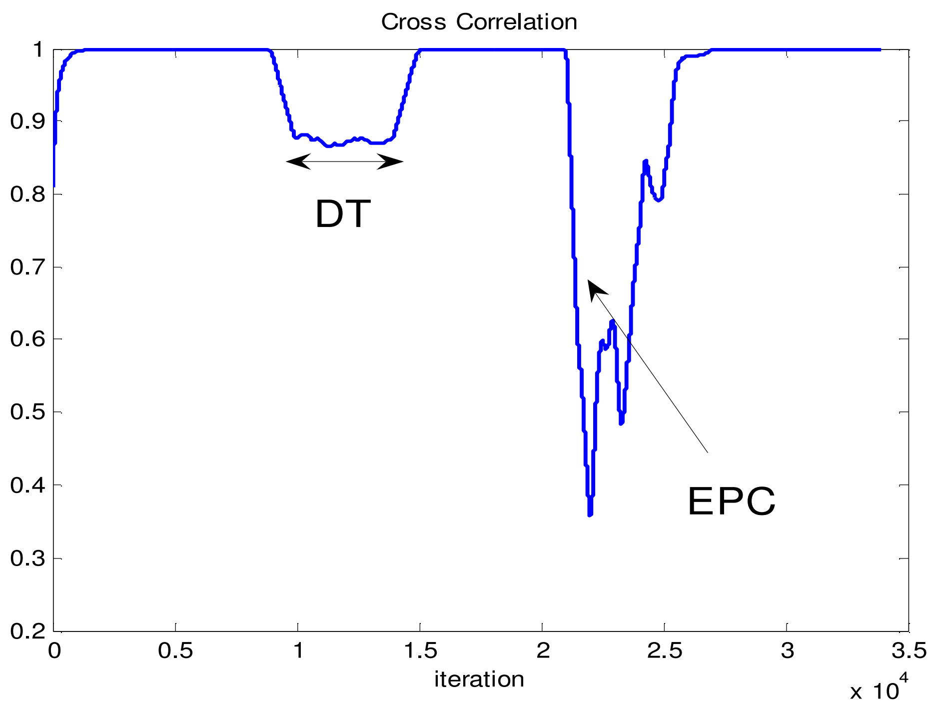


Fig 3.1 Variation of $\rho_{d,\hat{y}}$ between DT and EPC

In next section, we propose the modified cross-correlation double talk detector. The modified cross-correlation detector can distinguish the four cases. Whatever the level of double talk or echo path change is present, the modified DTD can detect correct.

3.2 The modify cross-correlation double talk detector

In Section 3.1, we discussed that the cross-correlation DTD is hard to differentiate between double talk and echo path change. In this section, we propose the modified cross -correlation DTD. The conventional cross correlation DTD although can detector one situation. But in general case, the conventional cross correlation DTD detect error between DT and EPC.

In Table 3.1, we will consider four cases of the cross-correlation, depending on the near-end signal energy and the degree of the echo path change. The small EPC means that the new change channel is close to the origin channel. The large DT means that the near end speech energy is large.

From Table 3.1, we can find four typical cases in double talk and echo path change. But the conventional cross correlation DTD works in only one case. To make the detector robust, we extend the cross-correlation DTD by incorporating microphone energy. Fig 3.2 is the structure of the modified cross-correlation DTD.

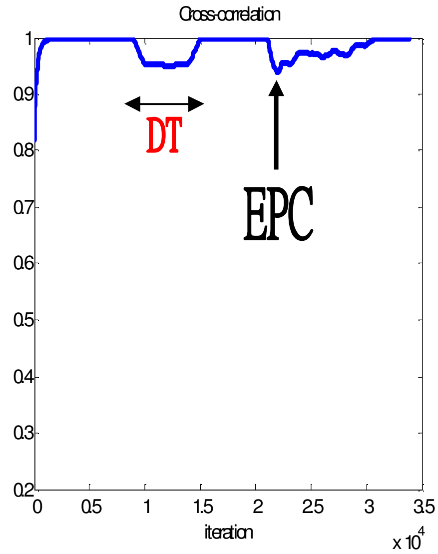
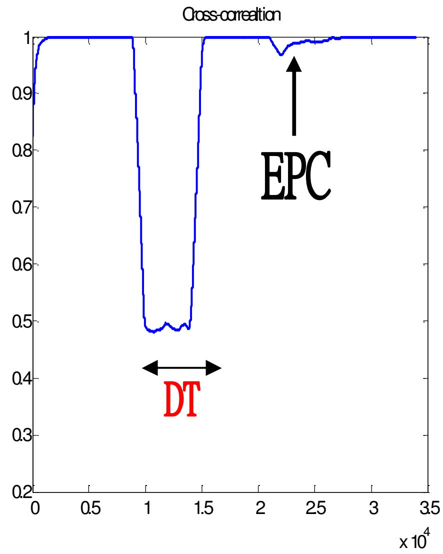
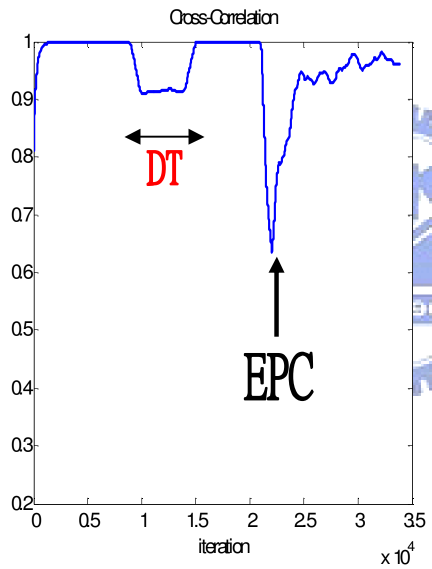
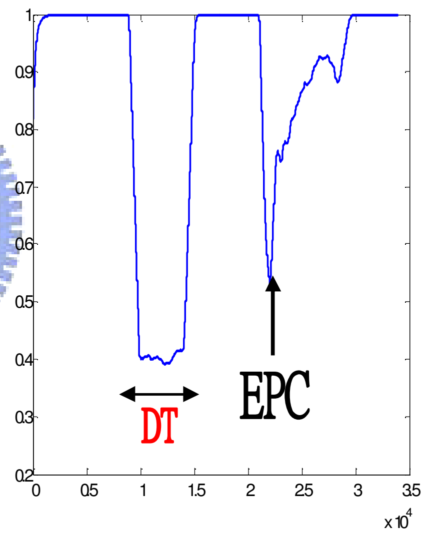
<table>
<tr><td>DT
EPC</td><td>Small</td><td>Large</td></tr>
<tr><td>Small</td><td></td><td></td></tr>
<tr><td>Large</td><td></td><td></td></tr>
</table>

Table 3.1 The cross correlation $\rho_{d\hat{y}}(k)$ in four cases

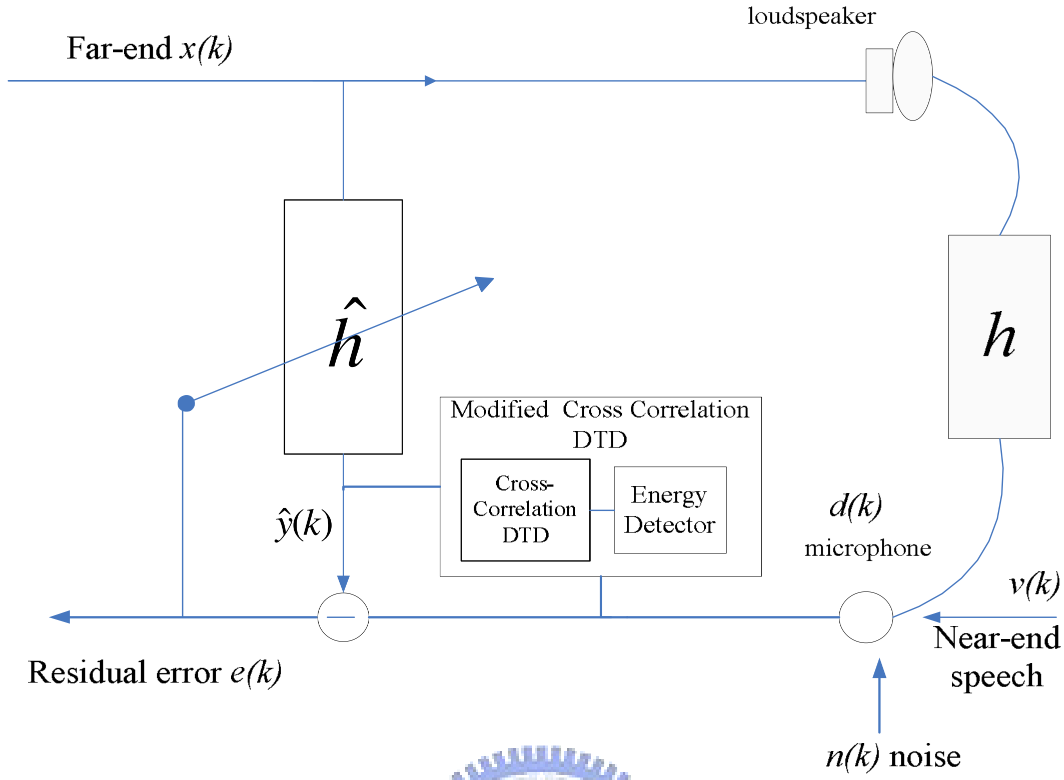


Fig 3.2 Structure of the modified cross-correlation DTD by including the microphone energy

The modified cross-correlation DTD includes the microphone energy detector. The energy detector can detect the double talk, and help cross-correlation DTD make correct decision. The energy detector is like Geigel double talk detector. With the energy detector, the modified cross-correlation DTD can detect correctly the four typical cases. Using the microphone energy, it has the drawback when the near end speech energy is small. The modified cross-correlation DTD is hard to detect double talk in near end speech small energy. The modified DTD drawback is the same with the Geigel DTD. However, the Geigel DTD is difficult to detect echo path change.

The flow chart of the modified cross-correlation DTD algorithm is given in Fig. 3.3.

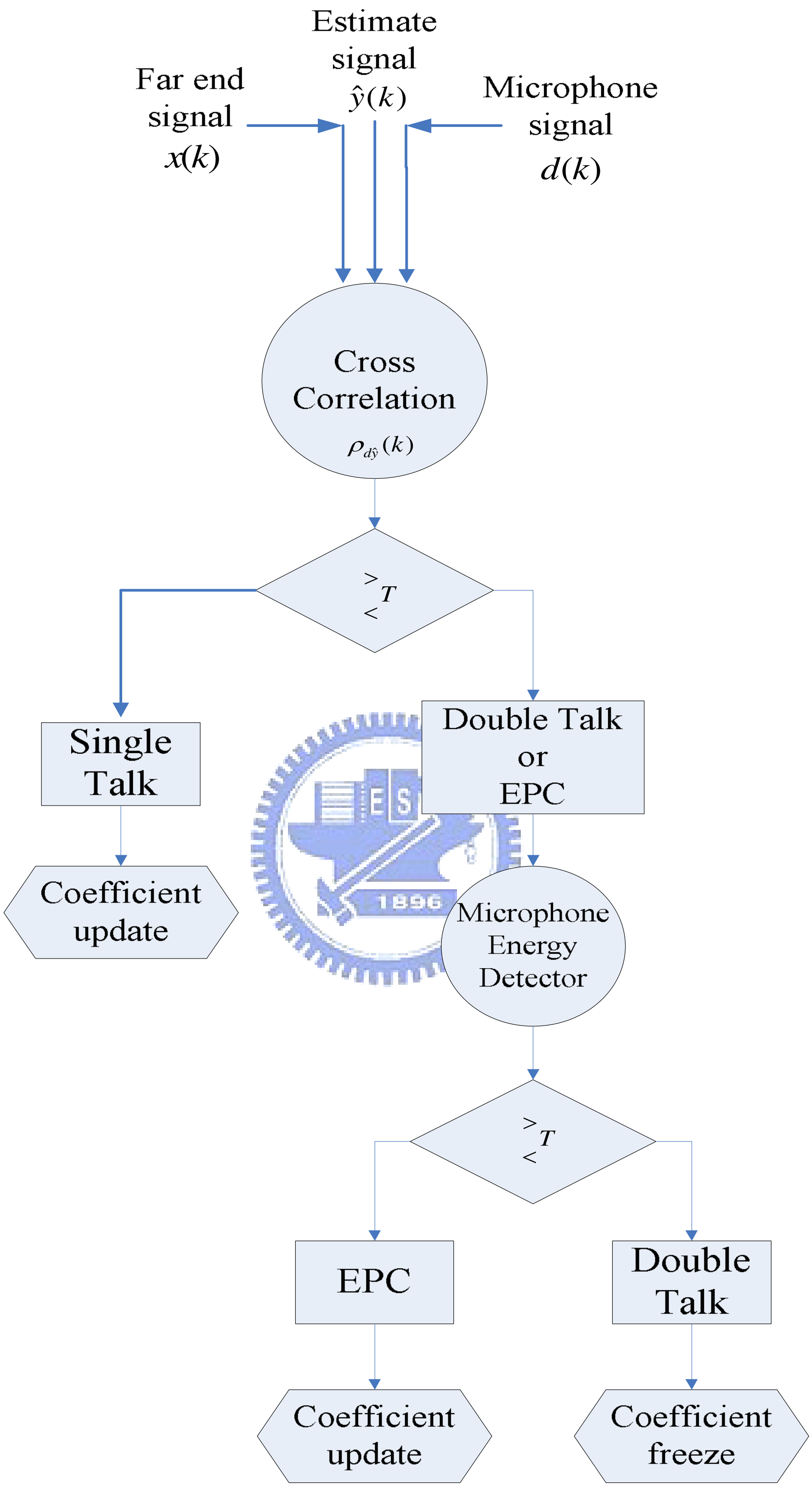


Fig 3.3 Flow chart of the modified cross correlation DTD algorithm by microphone energy

First, we use the correlation $\rho_{d\hat{y}}(k)$ in (2.1.4) to decide single talk or double talk (or echo path change). This means that double talk or echo path change is present when correlation is smaller than some threshold. Second, we use the microphone energy to detect double talk. The microphone energy detector algorithm in (2.3.2) is written for simplicity.

$$P_\alpha(k) = \frac{P_d(k)}{P_x(k)}$$

The microphone energy detector actually is a Geigel DTD with smoothed microphone and far end signal energy. If $P_\alpha(k)$ is larger than the threshold, we can decide double talk. Once doubletalk is declared, the detection is held for a minimum period of time. If $P_\alpha(k)$ is smaller than the threshold. We can decide echo path change.

Now, we use two detectors that we can detect all situations. With two detectors, we can are more confident to decide double talk or echo path change. The cross-correlation DTD can detect double talk in near end speech small energy. But, the Geigel DTD can not. We make the cross-correlation DTD to be more robust. In Chapter 4, simulations of all cases will be performed to verify the effectiveness of the modified cross-correlation DTD.

3.3 Evaluating DTD in echo path change and double talk

In Section 2.2, Morgan [8] proposed an objective technique to evaluate double talk detectors. But the technique only calculates the miss probability in double talk. The DTD should decide double talk or echo path change. Therefore, we also can calculate the miss probability when echo path change is present. This section will introduce the method to calculate the miss probability in echo path change period. We can combine [8] with our technique to calculate the miss probability in double talk and echo path change.

3.3.1 Introduction the technique evaluate DTD in echo path change

Before we introduced the technique evaluate DTD in echo path change, we discuss echo path change. From [8], we calculate the miss probability in different levels of the near end signal energy. But for the echo path change, we must calculate the miss probability in different echo paths. We define the parameter ρ_h to quantize degree how the echo path changes. The channel correlation ρ_h is defined as follows.

$$\rho_h \triangleq \frac{\hat{h} \cdot h_c^T}{\sqrt{\left|\hat{h}\right|^2 \left|h_c\right|^2}} \tag{3.3.1}$$

where h_c is the new change channel , assuming the filter has converged $\hat{h} \approx h$, h is the origin channel.

The small channel correlation means the echo path change h_c is very different to the origin channel h. The detector can easily detect that echo path change for a small ρ_h.

The channel correlation ρ_h lies between -1 and 1.

In section 2.2, the miss probability has been used. In this section, we also calculate the miss probability in echo path change. Avoiding disarraying the miss probability, we define that the miss probability in echo path change calls the change miss probability P_{cm}.
Similarly, the false alarm probability in echo path change defines the change false alarm probability P_{cf}.

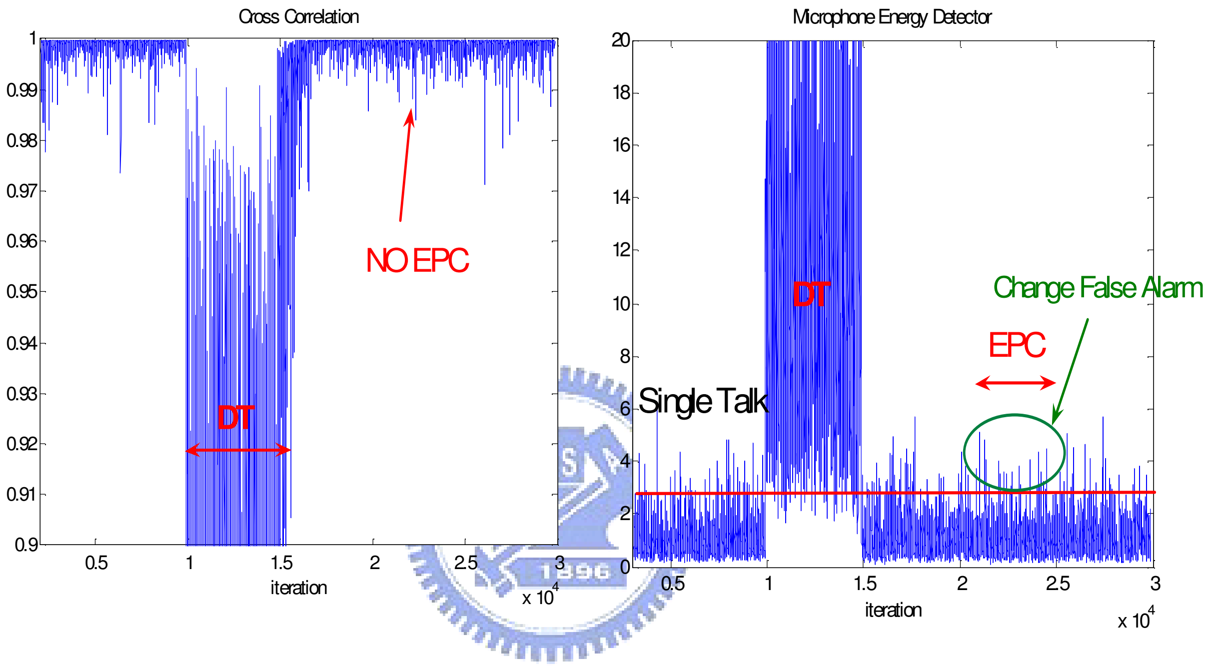


Fig 3.4 The method of the calculating false alarm probability

Now, we introduce the technique to evaluate DTD in echo path change. First, we calculate the change false alarm probability P_{cf} using the count in Fig 3.4.

$$P_{cf} \triangleq P(\text{EPC is detected} \mid \text{EPC Not happens})$$

The change false probability $$P_{cf} = \frac{\phi}{N} \qquad (3.3.2)$$

where N is trial length, ϕ is the frequency when detector makes error decision.
We find the threshold under fixed false alarm probability from (3.3.2). This step is equal to the [8].

Second step, Let EPC happen in range $\bar{C}$.

Next step, calculating the miss probability is in this step.

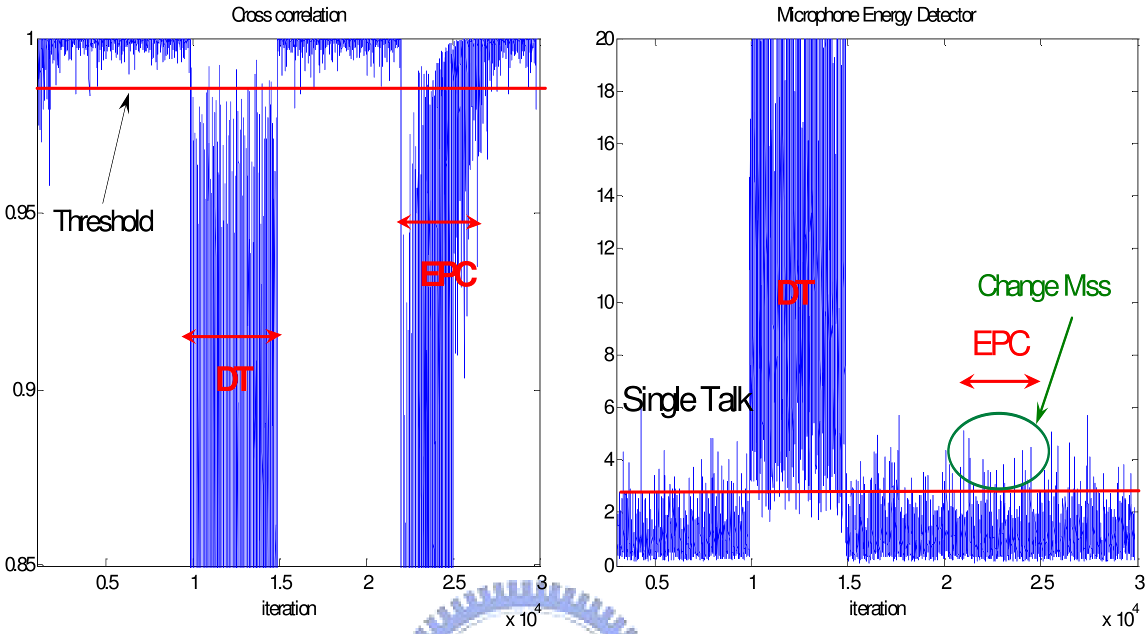


Fig 3.5 The method of the calculating miss probability

The change miss probability $P_{cm} = 1 - \frac{\phi \cdot \bar{C}}{N}$ (3.3.3)

The change miss probability is one minus correct probability. By the above procedure, we can calculate the change miss probability in EPC. We can use the technique to calculate the miss probability in different level of echo path change.

The complete DTD evaluation technique is summarized as follows.

1) a) Select threshold T. b) Compute false alarm probability P_{cf} from (3.3.2). c) Repeat steps a, b over a range of threshold values. d) Select threshold value that corresponds to $P_{cf} = 0.1$ 2) Calculate echo path change Let EPC happen in range $\bar{C}$ 3) Select different channel correlation value. a) DTD algorithm decides echo path change range ϕ. b) Compute P_m from (3.2.2) c) Repeat steps a, b, c over all conditions. 4) Repeat step 2 and step 3 over a range of channel correlation values. 5) Plot average P_m as a function of channel correlation

Table 3.2 DTD evaluation procedure in case of echo path change

3.3.2 Evaluating DTD miss probability in echo path change and double talk

In section 3.3.1, we introduced the technique which can calculate the miss probability when echo path change is present. Now, we combine Section 2.2 technique with section 3.3.1 method. The combined technique can calculate the miss probability in double talk and echo path change period.

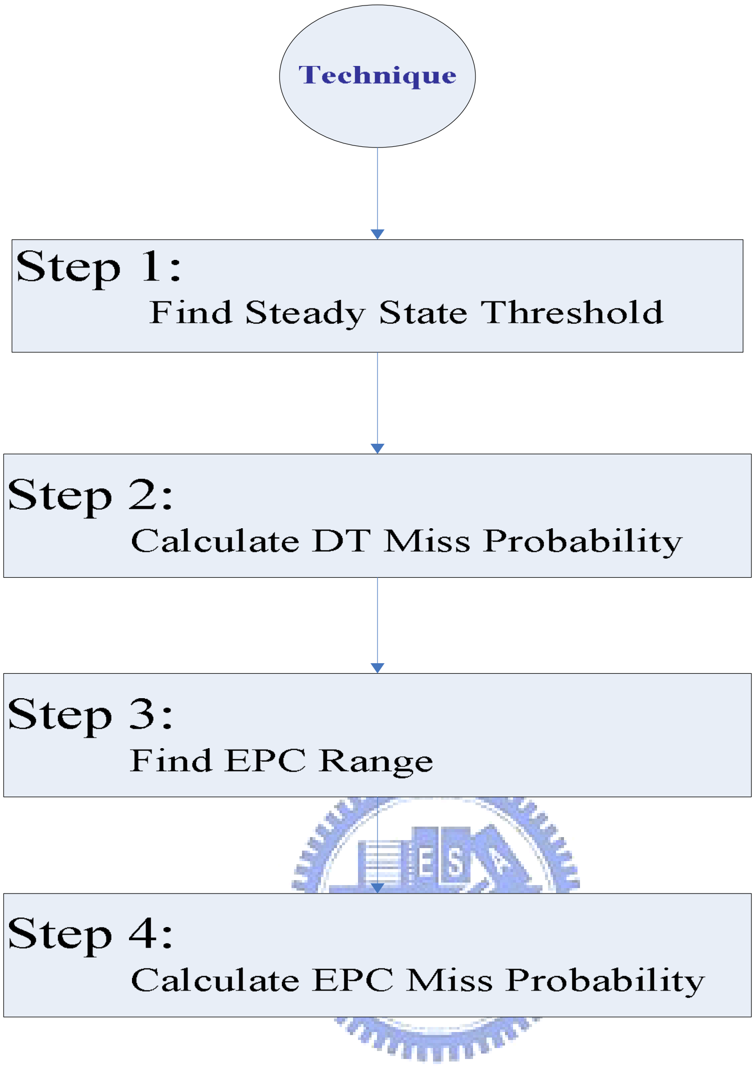


Fig 3.6 A flow chart of the calculated miss probabilities for DTD and EPC

From Fig 3.6, the step 1 and step 2 of the combined technique is the same method in [8]. The fist two steps can calculate the miss probability in double talk period. The step 3 and step 4 of the combined technique is introduced in section 3.3.1. The last two steps can calculate the miss probability when echo path change is present.

3.3.3 Weight miss probability

Before we calculated the miss probability, we have the same weight to calculate the miss probability. But, the same weight is not fair. Therefore, we want to modify the weight. The weight miss probability is more fair and robust. We use the residual error power or near end signal energy to be different weight.

In DT period, we calculate the miss probability from (2.2.2). The miss probability is given as follows.

$$P_m = 1 - \frac{\sum \phi \cdot \bar{x} \cdot \bar{v}}{\sum \bar{x} \cdot \bar{v}} = 1 - \text{correct probability}$$

We can find that the variable Φ is binary. It means that the DTD detect correct is one. On the contrary, if DTD detected wrong that the variable Φ is zero. The variable Φ is one if DTD detected correct, and we can find that the DTD had the same weight to calculate the miss probability.

Therefore, we modify the weight.

$$\Phi = \begin{cases} 0.6, & \text{near end energy} > 5\text{db} \\ 1\,, & -5\text{db} \le \text{near end energy} \le 5\text{db} \\ 1.4\,, & -5\text{db} < \text{near end energy} \end{cases} \quad (3.3.4)$$

We modify the weight in (3.3.2). When near end signal energy is large, the DTD can decide DT easier. Therefore, we change that the weight is smaller. In other words, the DTD can decide DT harder in near end signal small energy. The weight is larger than 1. Hence, we modify the weight that the miss probability is more fair and robust.

Equally, we can modify the weight when echo path change is present. In section 3.2, we proposed a technique to calculate the miss probability in echo path change. We calculate the miss probability in echo path change from (3.3.3).

$$P_m = 1 - \frac{O_d \cdot O_n}{O_n \cdot O_n}$$

The variable O_d equal one when the DTD detect correct. Similarly, the weight is the same in every point. Therefore, we modify the weight by residual error power. In section 3.3.1, we also propose the ρ_h to judge the level of echo path change. When the ρ_h is small, it means that the residual error power is also small, and the DTD detect harder correct. On the contrary, the DTD detect easier correct in large ρ_h situation. The modified weight is as follows.

$$O_d = \begin{cases} 0.7, & e^2(k) > 0\text{db} \\ 1, & -10\text{db} \le e^2(k) \le 0\text{db} \\ 1.6, & -10\text{db} < e^2(k) \end{cases} \qquad (3.3.5)$$

3.4 Variant threshold

In section 2.1, we introduced the cross correlation DTD. If $\rho_{d,\hat{y}}(k)$ is smaller than the threshold, the detector decide double talk. We only use the fixed threshold to detect double talk. But, using the fixed threshold has a drawback. From Fig 3.7, double talk is present in 1K to 1.5K. We set that the fixed threshold is equal 0.95. If $\rho_{d,\hat{y}}(k)$ is below the threshold, the detector decide that double talk is present and the correlation is decreasing to 0.5. When double talk is over, the correlation is increasing to 0.95. Therefore, $\rho_{d,\hat{y}}(k)$ is larger than the threshold and the filter continue adapting. But, the better detector can decide that the filter continue adapting when double talk is over.

In order to perform the detector, we use the variant method to prove the performance.

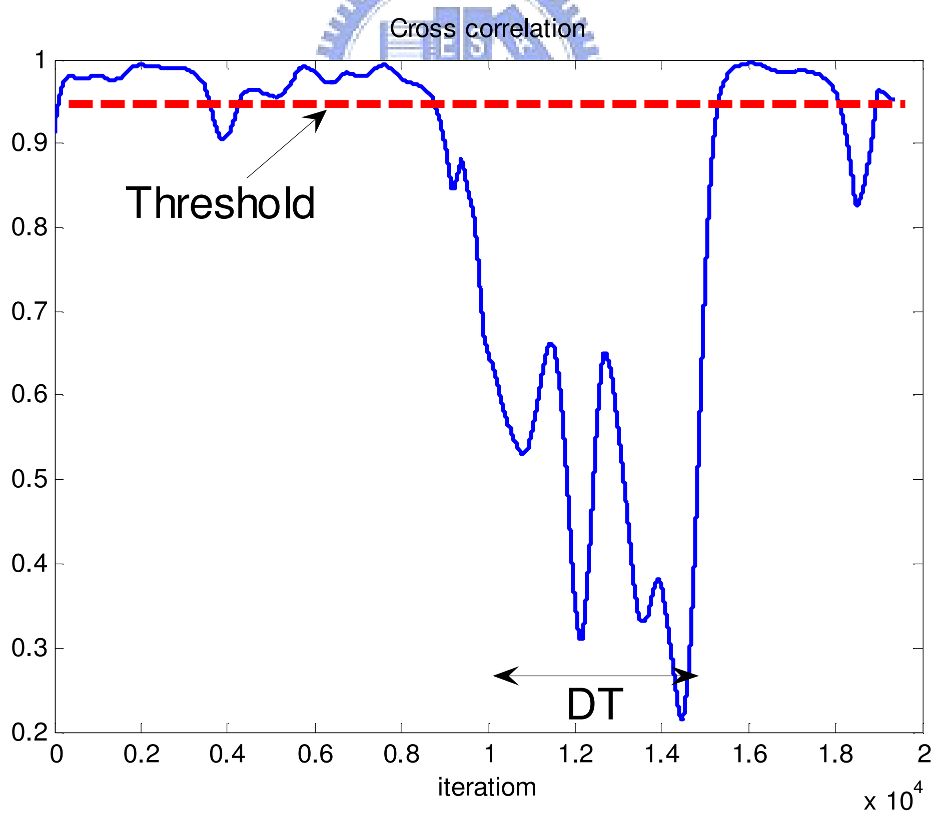


Fig 3.7 The cross correlation $\rho_{d,\hat{y}}$ under fixed threshold

Now, we use the variant threshold like Fig 3.8. If the detector decides that double talk is present, we can set that the threshold is larger than the correlation in DT period. With the new threshold, the detector can faster decide that double talk is over.

Then, the detector decides the single talk, and the threshold renews to set. This means that the double talk detector can faster decide double talk when double talk is present again.

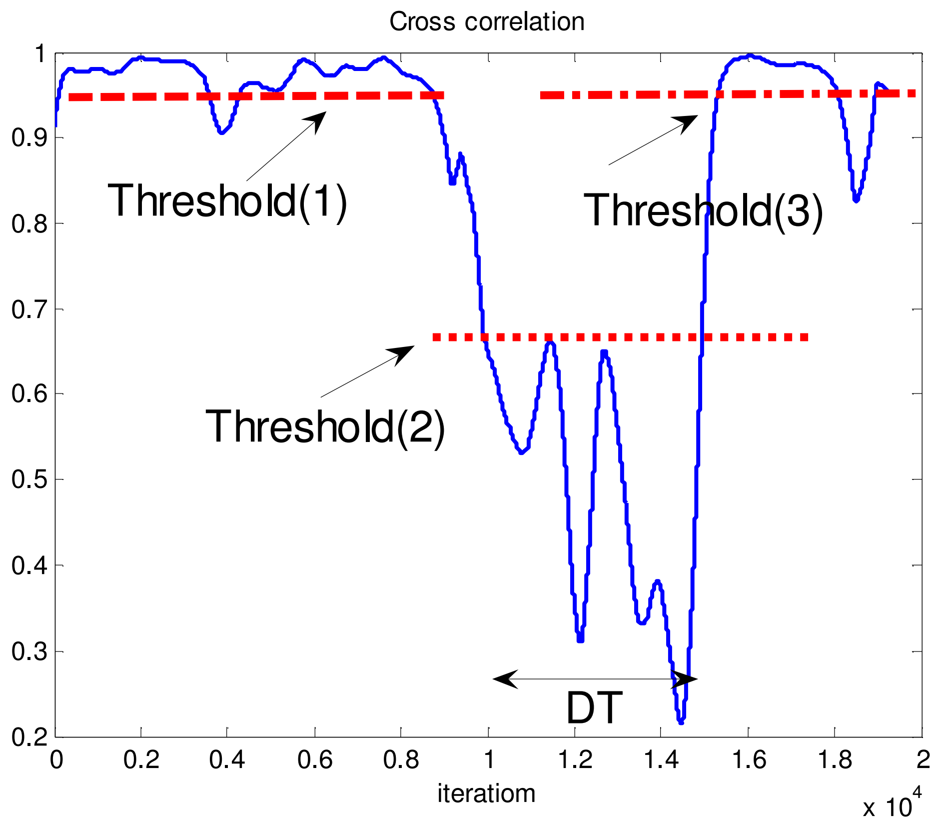


Fig 3.8 The cross correlation $\rho_{d,\hat{y}}$ under variant threshold

In section 2.4, we analyzed the theoretical cross correlation whatever double talk or echo path change. From (2.4.4) and (2.4.5), the theoretical correlation is exact value in any double talk cases. If the detector decides that double talk is present, we use the (2.4.5) to calculate the threshold again.

But, (2.4.5) is not practice, and we rewrites (2.4.5) for simplicity.

$$\rho_{d,\hat{y}}(k) = \frac{1}{\sqrt{1+\frac{\sigma_n^2+\sigma_v^2}{\sigma_x^2 \cdot h^T h}}}$$

From (2.4.4), we can find that σ_n^2 is unknown parameter in real environment. There we modify the (2.4.4).

$$\rho_{d,\hat{y}}(k) = \frac{1}{\sqrt{1+\frac{\sigma_n^2+\sigma_v^2}{\sigma_x^2 \cdot h^T h}}} = \frac{1}{\sqrt{1+\frac{\sigma_d^2-\sigma_y^2}{\sigma_x^2 \cdot h^T h}}} \approx \frac{1}{\sqrt{1+\frac{\hat{\sigma}_d^2-\hat{\sigma}_y^2}{\sigma_x^2 |w|^2}}} \qquad (3.4.1)$$

where

$$\hat{\sigma}_d^2 = \frac{1}{M}\sum_{i=0}^{M-1} d^2(k-i) \ , \ \hat{\sigma}_x^2 = \frac{1}{M}\sum_{i=0}^{M-1} x^2(k-i)$$

Now, the (3.4.1) is practice in real environment. We use (3.4.1) to calculate the threshold in double talk period. The new threshold is calculated as follows.

$$\text{Threshold} = \frac{1.2}{\sqrt{1+\frac{\hat{\sigma}_d^2-\hat{\sigma}_y^2}{\sigma_x^2 |w|^2}}} \qquad (3.4.2)$$

The (3.4.1) is accurate correlation. If the threshold is larger than correlation, the filter is safer and not diverges. The variant threshold in cross correlation DTD algorithm is from Fig 3.9.

If the filter has converged, the double talk is abrupt present. The correlation is decreasing and small the threshold. The DTD detect the double talk and freeze updating. We can calculate the threshold by (3.4.2). With the new threshold, we can faster find that the double talk is over. Therefore, the performance is better than the fixed threshold. Until the correlation is large the origin threshold, the threshold change the origin threshold.

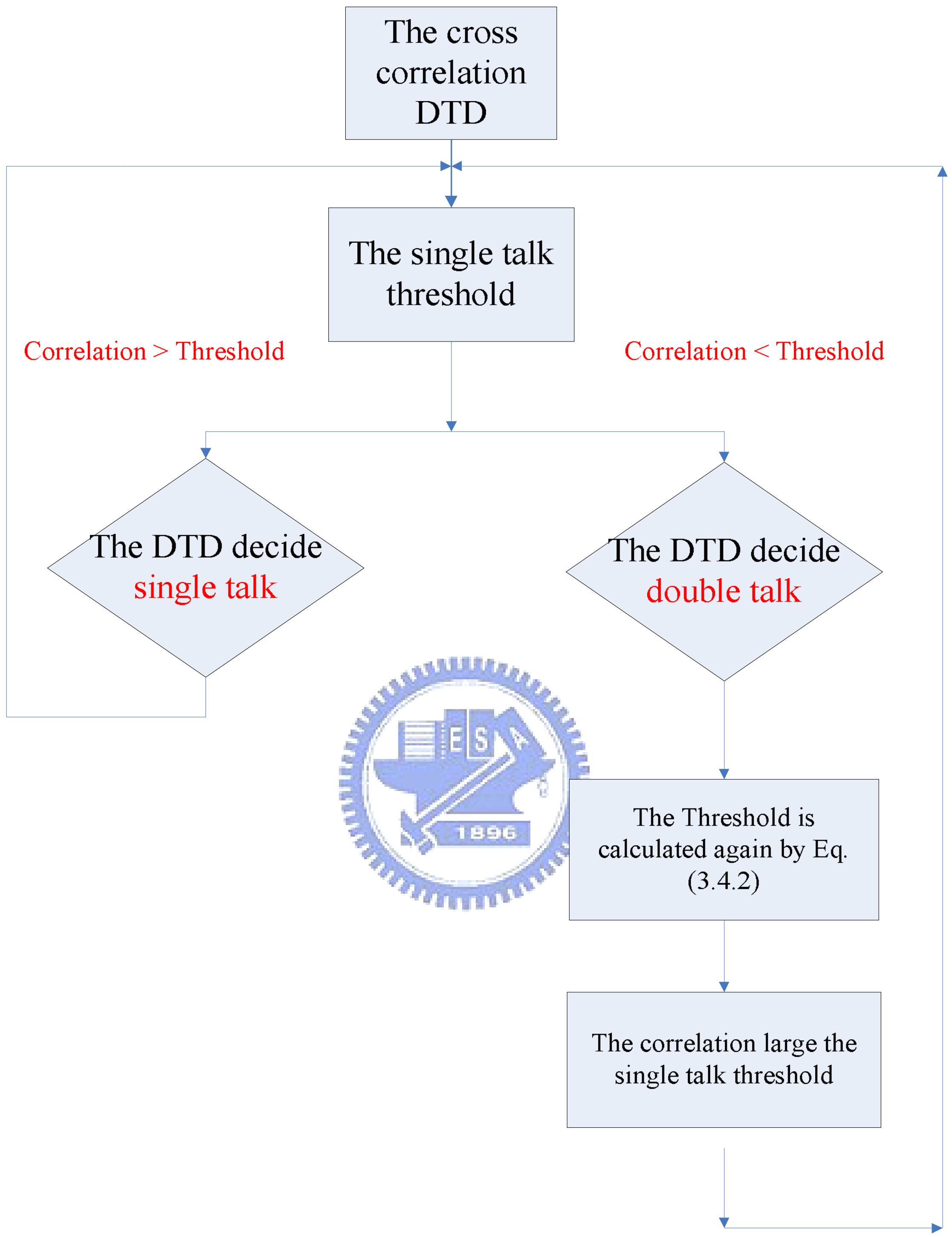


Fig. 3.9 Algorithm of the variant threshold selection in cross correlation DTD

Chapter 4

Computer Simulations

In this chapter, we will perform the computer simulations to verify the previous derived results. The difference between the simulation and theoretical results will be compared in this chapter. Both white and speech signals are also considered.

In Section 4.1, we explain the some simulation parameters, such as echo path impulse response, speech model and others parameters. In Section 4.2, we compare DTDs in section 2.1. In section 4.3, we illustrate the effectiveness of the robust DTD.

Theoretical and simulated cross-correlations DTD are compared in Section 4.4. And shown to be very close to each other. In Section 4.5, we compare the theoretical and simulated miss probabilities.

In Section 4.6, we verify that the modified cross-correlation can detect precise whatever double talk or echo path change is present. The technique for evaluating DTD in echo path change and double talk is simulated in Section 4.7. In Section 4.8, we propose the variant threshold to adapt to the real speech environment.

4.1 Simulations parameters and room impulse response

The echo impulse response $h(k)$ is shown in Fig 4.1. Fig 4.2 shows the far end signal and near end speech, where DT occurs from 10k to 15k. In following speech simulations, will use the signal in Fig 4.2

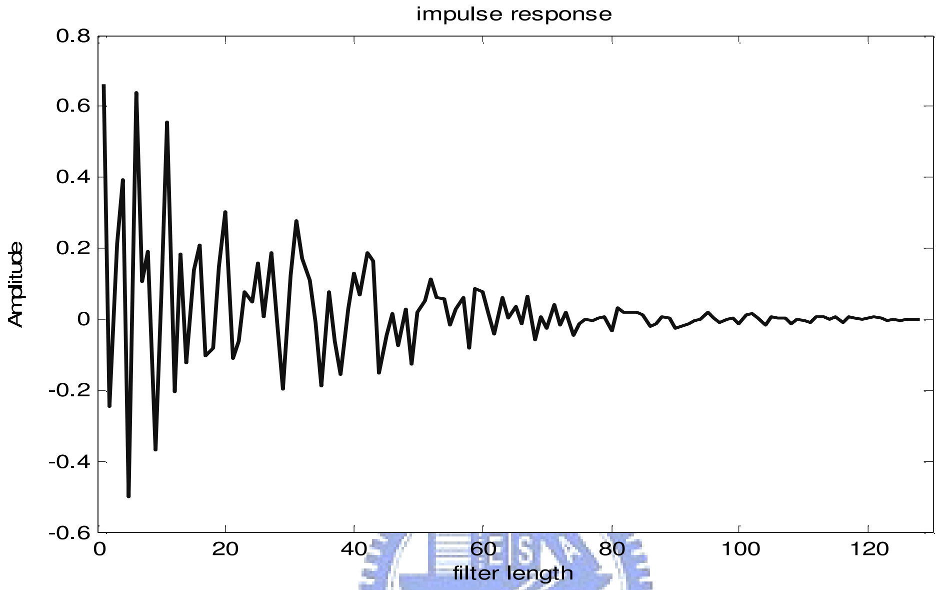


Fig 4.1 Echo path impulse response

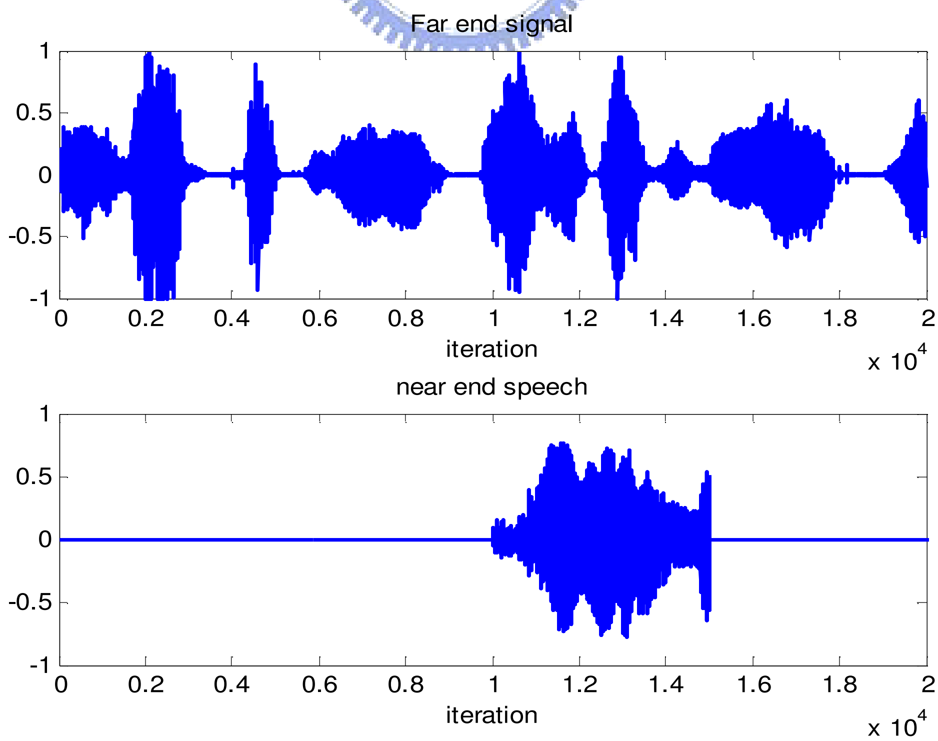


Fig 4.2 The far end signal and near end speech

4.2 Comparison of four typical double talk detectors in double talk

In Section 2.2, we have explained how the technique in [8] can calculate the miss probability in double talk period. Here we let the step size $u = 0.2$, forgetting factor $\lambda = 0.01$, and NFR=$\frac{\sigma_v^2}{\sigma_x^2}$,denotes the energy ratio of the near end signal to the echo signal.

From Fig 4.3, we found the mic/AEC correlation double DTD is better than others DTDs. Similarly, two echo path model can freeze adaptive in double talk period. But from Fig 4.3 the two echo path model performance is bad. But, two echo path model is a good choice to implement in real environment for its excellent stability. The Gradient correlation double talk detector is most effective when near end signal energy is very small. The miss probability is an inverse proportion the NFR.

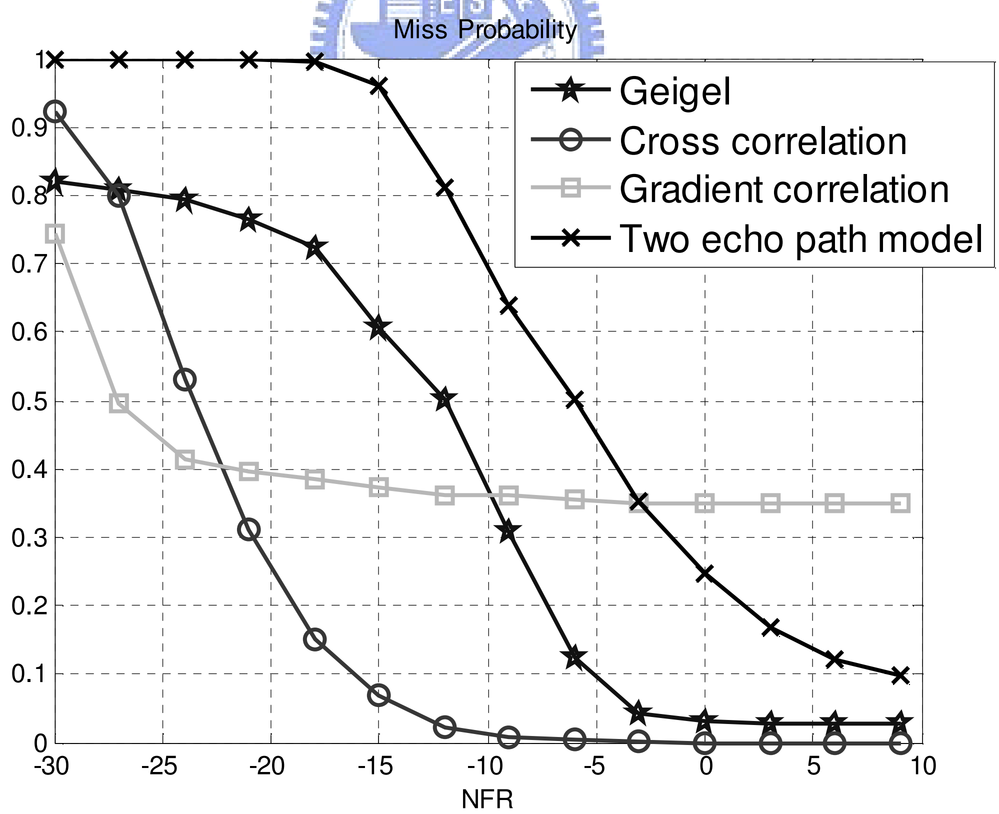


Fig 4.3 Comparison of DTDs miss probability in different near end signal energies

4.3 Robust DTD

In section 2.3.1, we introduced the robust Geigel DTD. The buffer range is set by the PDF of P_α. The buffer range is in Fig 2.8. From Fig 4.4, we can find that the robust Geigel DTD is better than the conventional Geigel DTD. In DT period, the robust ERLE is better about 3db than the convention. By the simulation, the buffer range makes the Geigel DTD more robust.

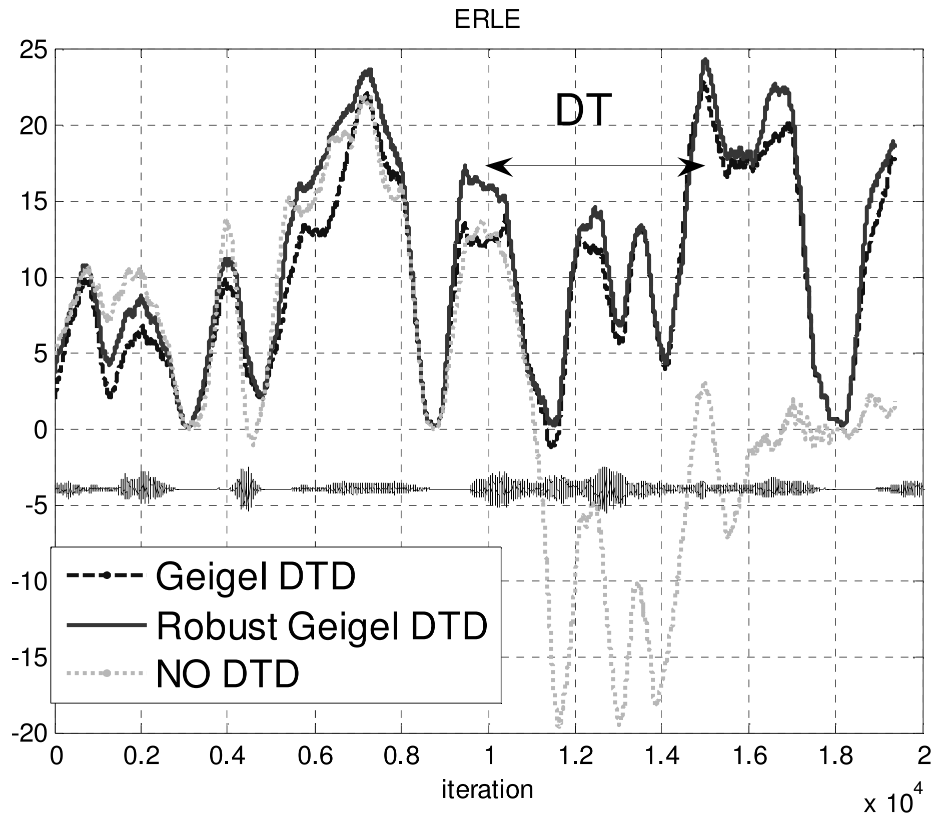


Fig 4.4 Comparison of the robust and conventional Geigel DTD

In section 2.3.2, we also introduced the robust cross-correlation DTD. We also use the different buffer range in Fig 2.9. In this section, we will simulate the different range. From Fig 4.5, we can find the wide buffer range is better than the narrow buffer range and the convention. However, the performance of the robust cross-correlation DTD is a little better than the convention.

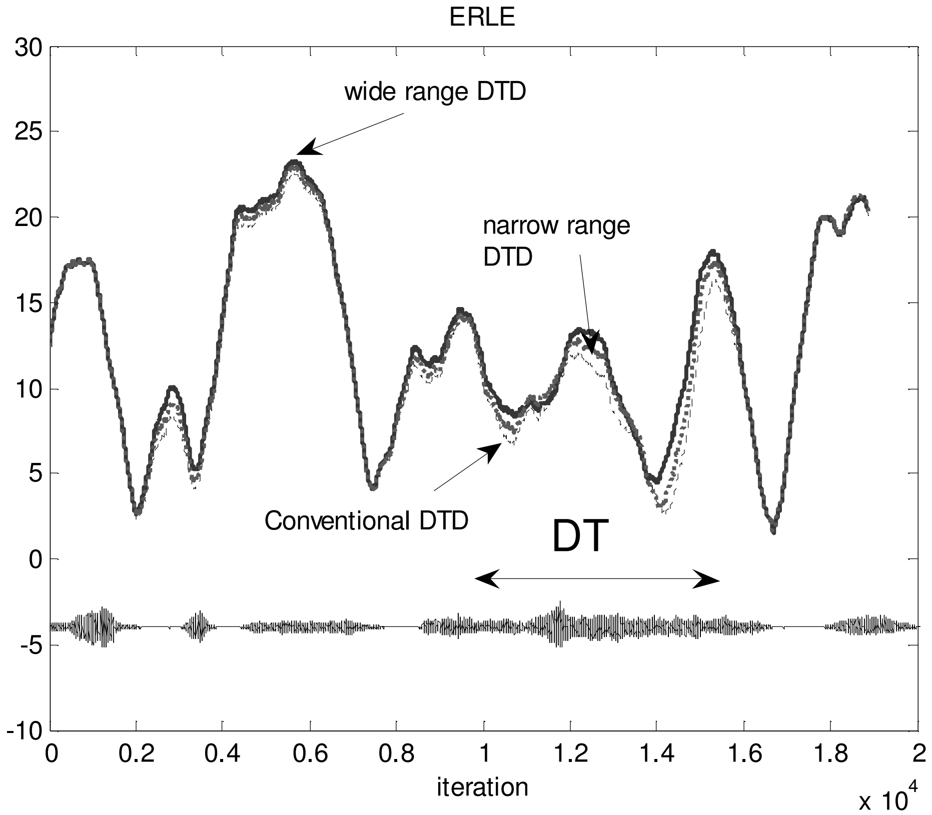


Fig 4.5 Comparison of the robust and conventional cross-correlation DTD

4.4 Analysis of the cross-correlation DTD

4.4.1 Analysis of the microphone signal and the estimated signal correlation DTD

In section 2.4, we derive the theoretical cross-correlation $\rho_{d,\hat{y}}$ of the mic/AEC correlation DTD. Now, we assume Near-end signal energy is equal to far-end signal energy. By (2.4.5), the theoretical correlation $\rho_{d,\hat{y}}$ is related with near end speech energy. The simulation is also very close to the theoretical correlation in DT period. Another case, the theoretical correlation $\rho_{d,\hat{y}}$ is equal to different channel correlation in (2.4.9) when echo path change occurs in 22k. When EPC is present, the correlation $\rho_{d,\hat{y}}$ is relation with the origin and new channel correlation. By simulation, we can find the simulation is also very close theoretical value in EPC.

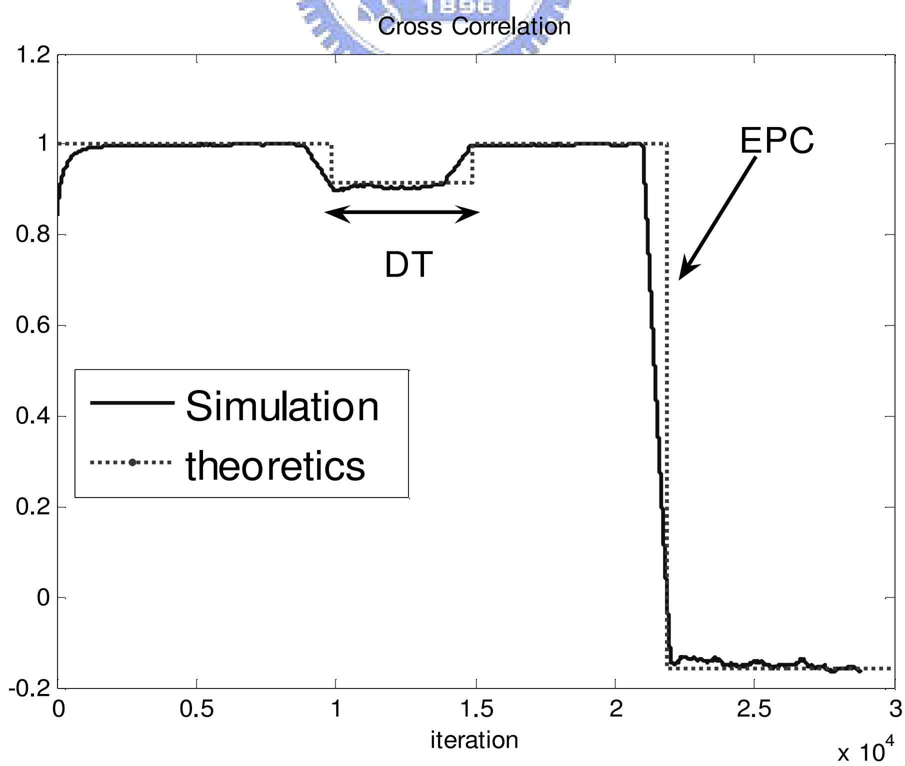


Fig 4.6 Simulated and theoretical the mic/AEC correlation $\rho_{d,\hat{y}}$ in DT and EPC

4.4.2 Analysis the microphone signal and the error signal correlation DTD

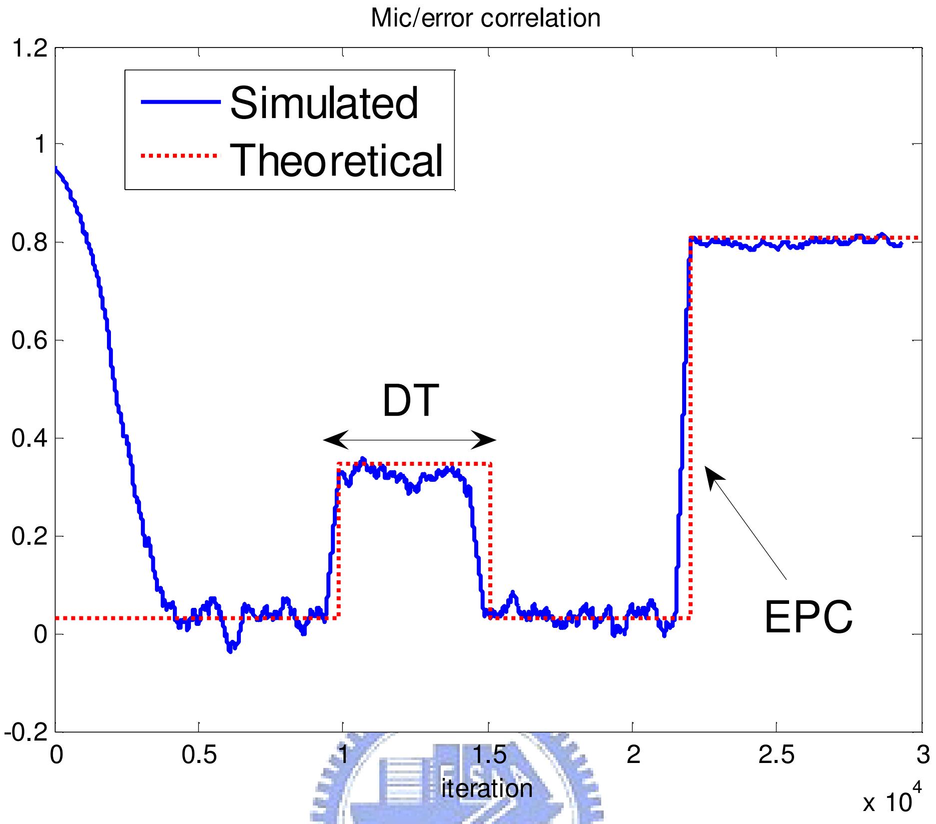


Fig 4.7 Simulated and theoretical the mic/error correlation $\rho_{d,e}$ in DT and EPC

From Fig 4.7, we can find that the simulation is also very closed the theoretic value. We add near end signal from 10k to 15k.And the near end signal energy is 1. From (2.4.15), the theoretic value is related with near end speech energy in double talk period. The theoretic correlation is relation with the change channel h_c and the difference Δh in the new channel and origin channel h in EPC from (2.4.19).

4.5 Analysis the cross-correlation DTD miss probability

In Section 2.5, we studied the miss probability of the cross-correlation DTD. First, we derived that the threshold is a function with false alarm probability in (2.5.5). In Figure 4.8, we find that the DTD threshold under fixed false probability is closed to theoretical threshold when false alarm probability P_f large 0.2. Therefore, the theoretical derivation (2.5.5) appears to be correct.

Next, we derive that the miss probability under fixed the false alarm probability. From Figure 4.8, when the false alarm probability is progressive smaller, the threshold also is smaller. The result is very intuitive. Because (2.5.3) does not include the near end signal, the correlation is very close 1. Hence, the threshold is also closed 1. But, when the threshold is getting further away from 1, the detector makes frequent errors in double talk period. Therefore, the false alarm probability is larger.

From Fig 4.9 and 4.10, we can find the miss probability P_m is decreasing when the false alarm probability P_f is increasing. In Figure 4.11, we can see that the near end signal energy is larger, and the miss probability is smaller. But the simulated line is not close to theoretical line especially when NFR is below -5db. This may be due to our assumption that the random variable $r = n^2 + v^2$ is Chi-square distribution. When the near end signal energy is small, it is added with noise. The mixed signal is not Chi-square distributed. But when the near end signal energy is large, it may be well assumed to have Chi-square distribution. When the miss probability is large 0.2, the cost of the detector is very high. Therefore, we focus on the small miss probability. We find that the theoretic is close the simulation.

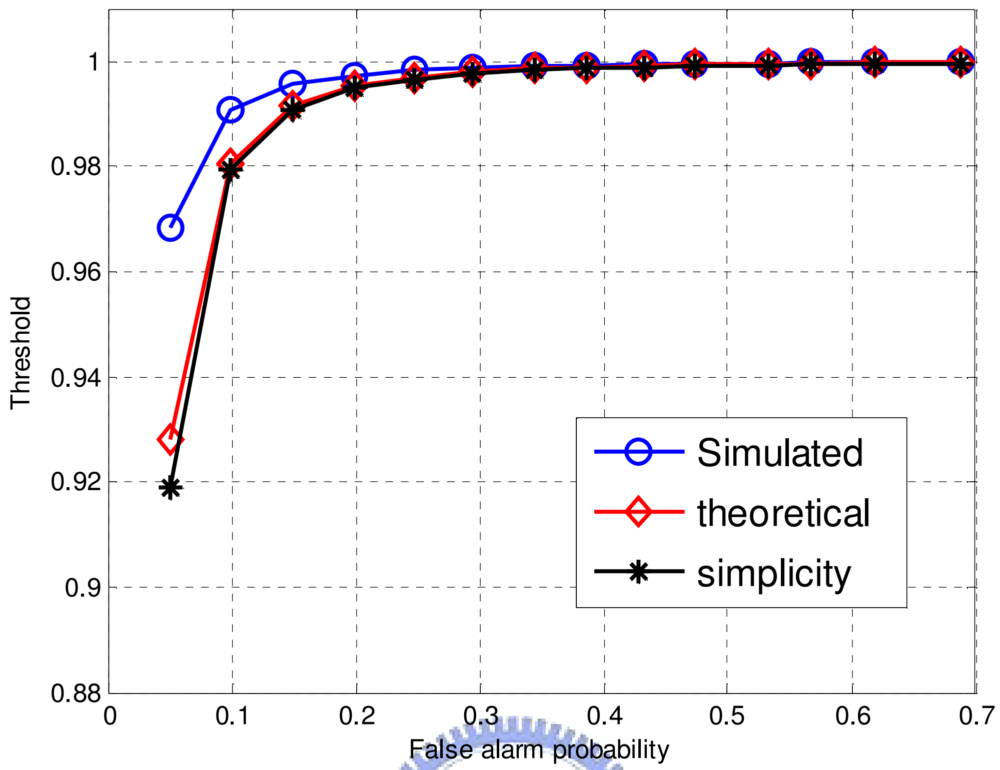


Fig 4.8 Simulated and theoretical cross correlation DTD threshold under fixed false probability

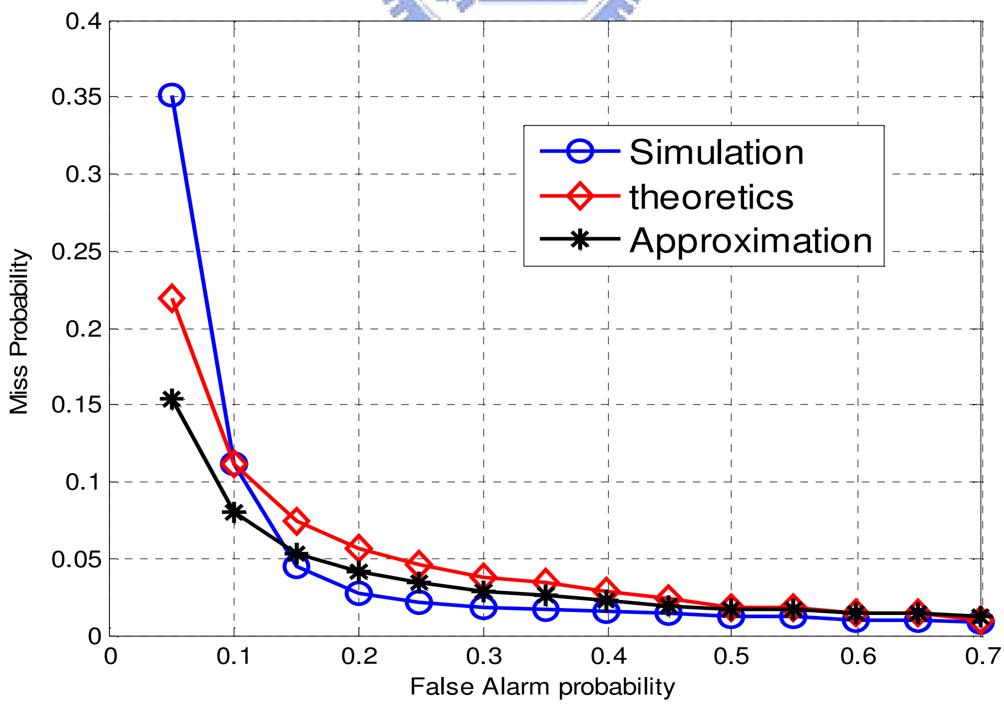


Fig 4.9 Simulated and theoretical cross correlation DTD miss probability under fixed false probability

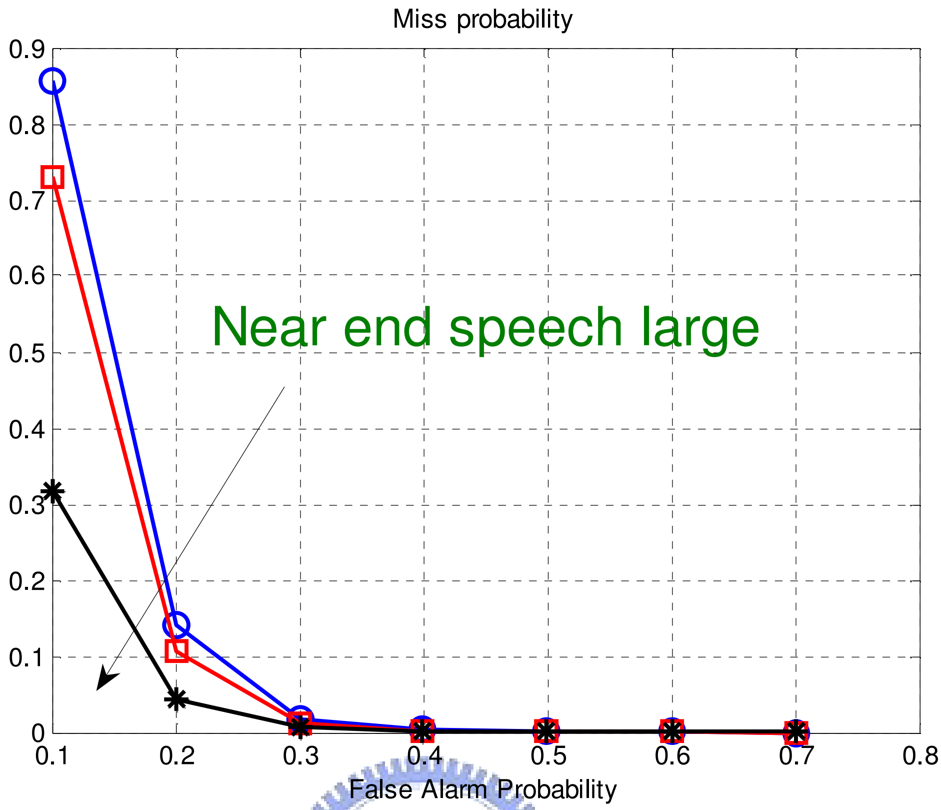


Fig 4.10 The effect of near end speech energy and the miss probability under fixed false alarm probability

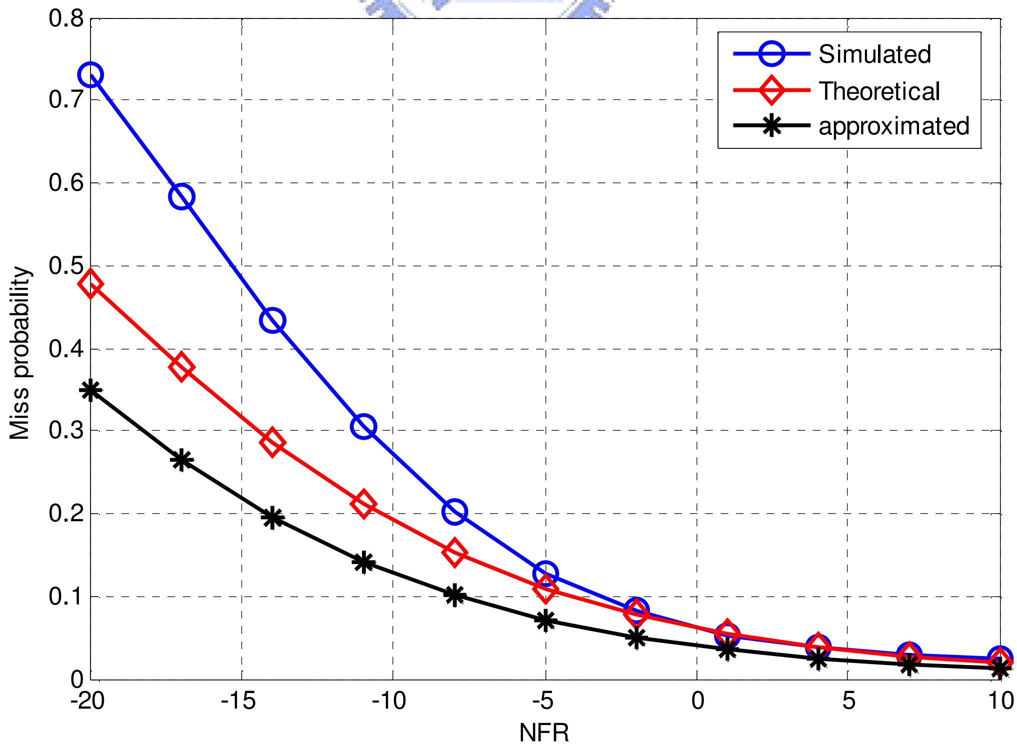


Fig 4.11 Simulated and theoretical cross correlation DTD miss probability

4.6 Nonlinear effect for cross-correlation DTD

In Section 2.6, we have derived the theoretical correlation $\rho_{d,\hat{y}}$ in nonlinear loudspeaker. From Fig 4.12, we can find that $\rho_{d,\hat{y}}$ in single talk and DT is decreased for nonlinear. The theoretical correlation $\rho_{d,\hat{y}}$ is close the simulated in single talk. However, in DT period, the theoretical correlation $\rho_{d,\hat{y}}$ is not close the simulated. Because we approximate the correlation $\rho_{d,\hat{y}}$ in (2.6.4).

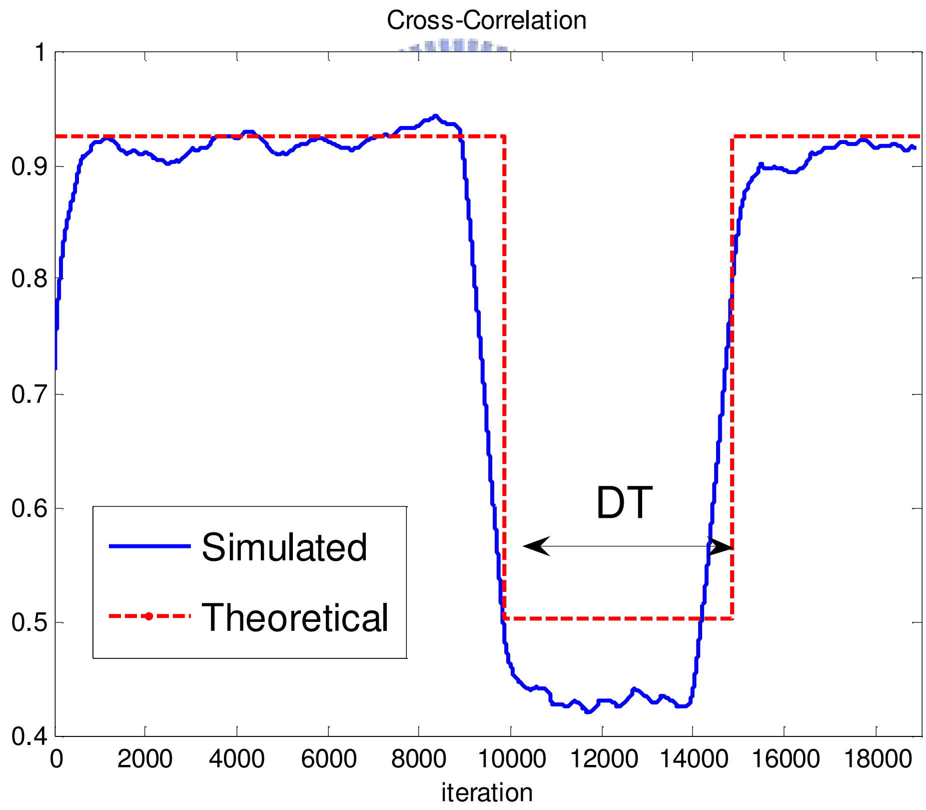


Fig 4.12 Nonlinear effect for cross-correlation $\rho_{d,\hat{y}}$

4.7 Modified cross-correlation DTD by microphone energy

In Section 3.2, we proposed the modified cross-correlation DTD. In this section, we simulate the four cases depending small/large DT or EPC. By simulation, we can find that the modified cross-correlation DTD can make corrected decisions whatever double talk or echo path change. The traditional cross-correlation DTD can not distinguish well between a double talk and echo path change. Our proposed is indeed more robust than the conventional correlation DTD.

We simulate different echo path change under fixed double talk energy. We use the cost to robust cross correlation DTD. In Fig 4.13 and 4.14, the small DT energy is equal far end signal energy, and the large or small channel correlation. We can find that the correlation $\rho_{d,\hat{y}}$ is decreasing in double talk and echo path change in Fig 4.13. We can find that the microphone energy detector is not increasing in EPC. From Fig 4.13 and F4.14, the detector can detect correct in different echo path under fixed near end speech energy in Fig 4.15. From Fig 4.15, we can find that the detector can not perfect decide in near end energy small energy.

Similarly, in Fig 4.16 and 4.17, the large DT energy larges four times to far end signal energy, and the large or small channel correlation. We can find that the correlation $\rho_{d,\hat{y}}$ is decreasing in double talk and echo path change in Fig 4.18. From Fig 4.18, we can verify that the modified cross correlation DTD can decide all case whatever double talk or echo path change.

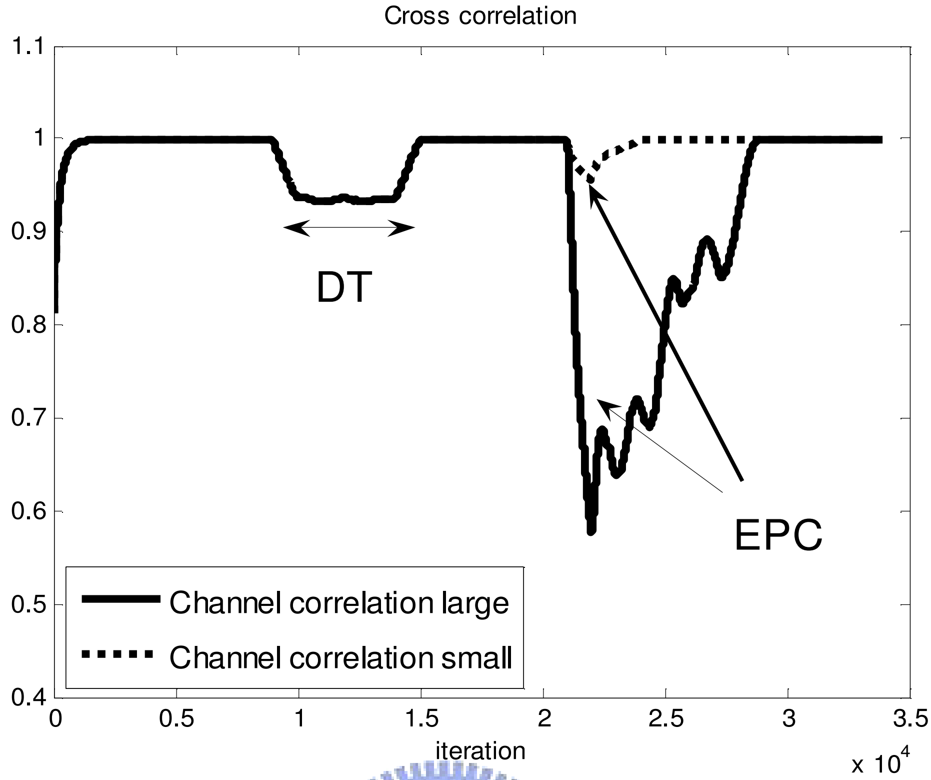


Fig 4.13 Cross correlation $\rho_{d,\hat{y}}$ of the different EPC under fixed small DT energy

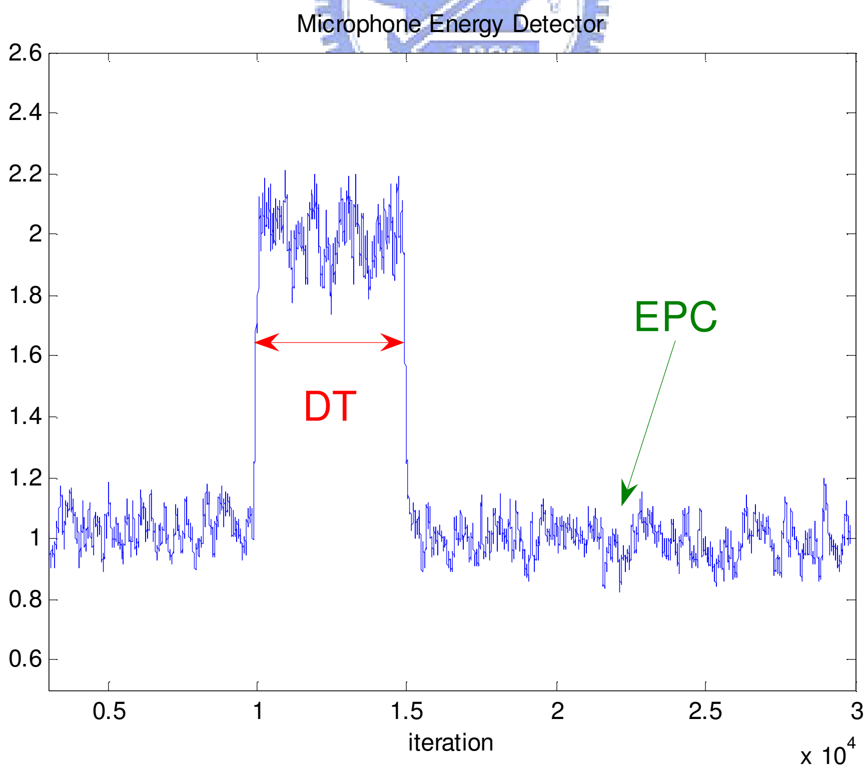


Fig 4.14 Microphone energy of the different EPC under fixed small DT energy

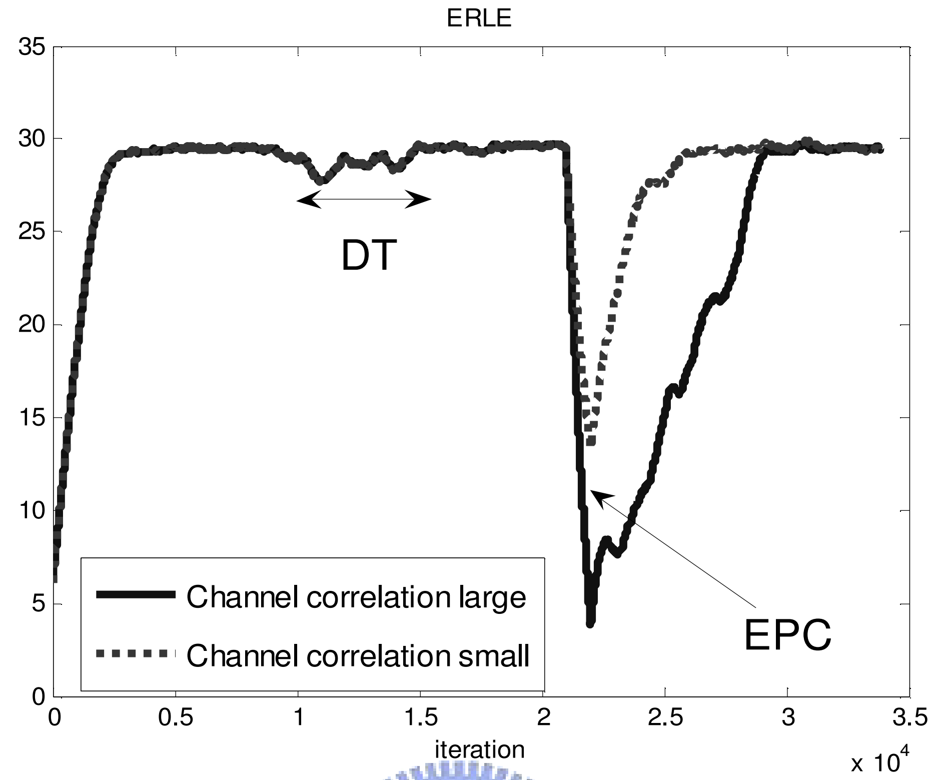


Fig 4.15 ERLE of the modified cross correlation DTD in different EPC under fixed small DT energy

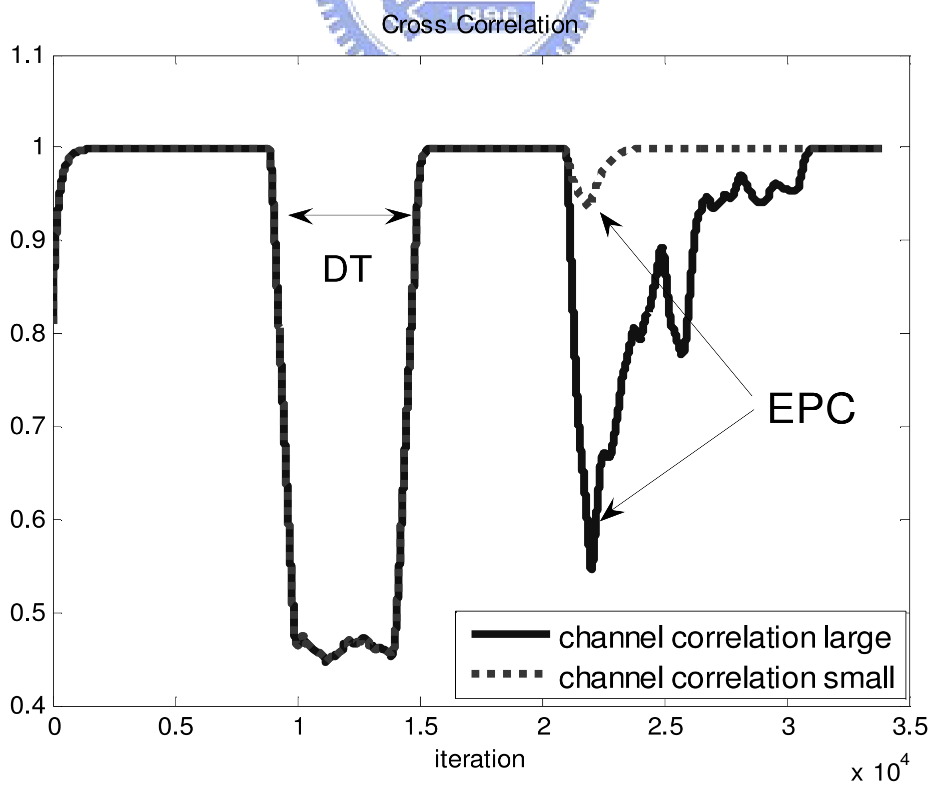


Fig 4.16 Cross correlation $\rho_{d,\hat{y}}$ of the different EPCunder fixed large DT energy

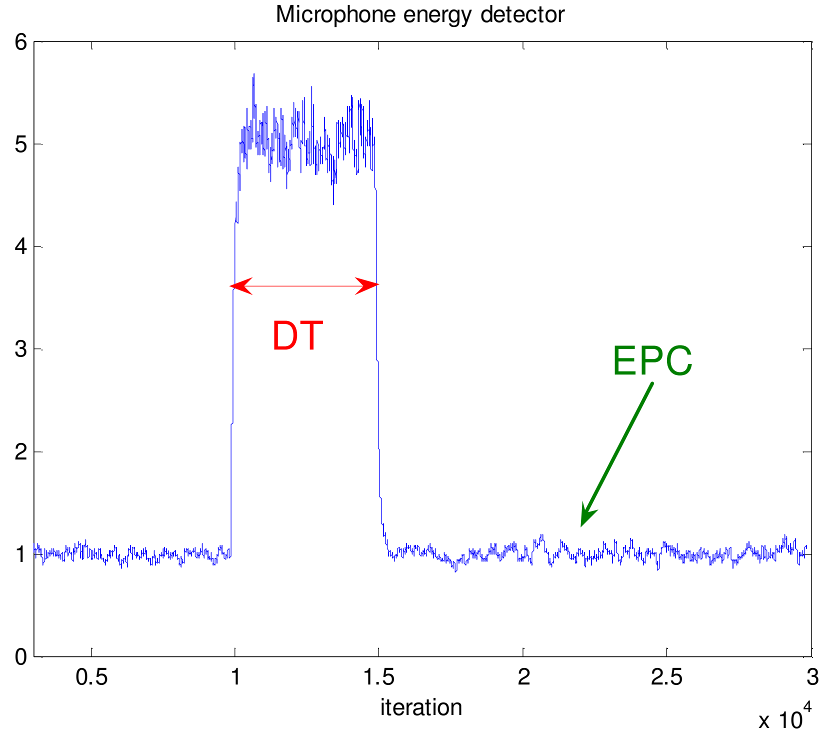


Fig 4.17 Microphone energy of the different EPC under fixed small DT energy

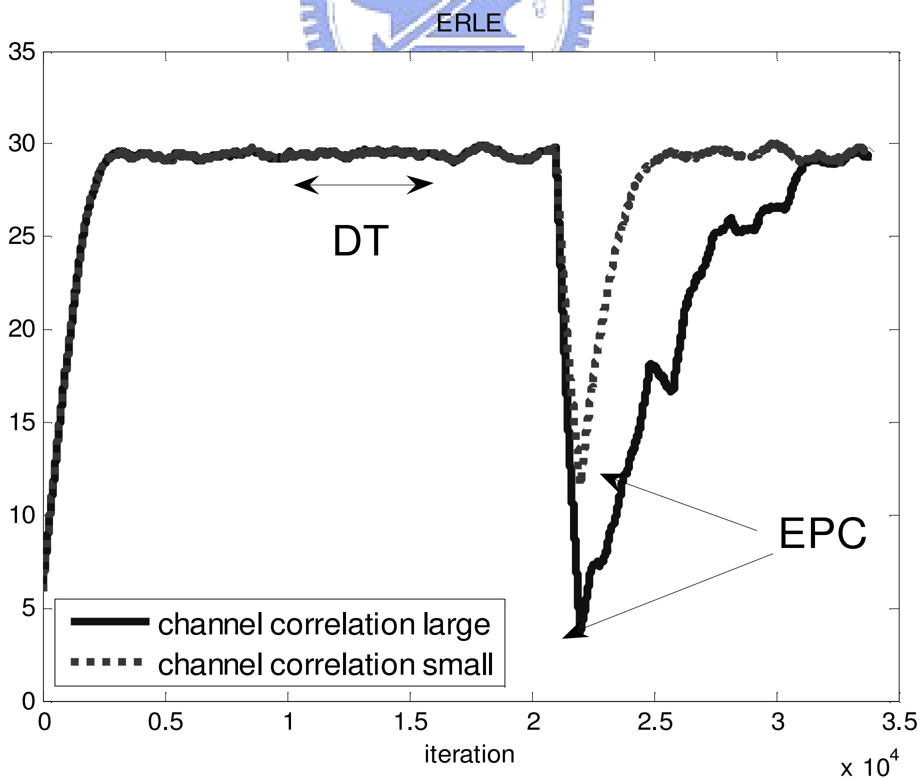


Fig 4.18 ERLE of the modified cross correlation DTD in different EPC under fixed large DT energy

4.8 An technique for evaluating DTD algorithms in double talk and echo path change

In section 3.3, we proposed that the technique could calculate the miss probability whatever double talk or echo path change. We compare the modified cross correlation DTD with the conventional cross correlation DTD. The conventional cross correlation DTD only can decide one case in table 3.1. Now, we use the technique and simulate it. From 4.19, we can find that the miss probability of the modified cross correlation DTD is an inverse proportion to near end signal energy. However, the conventional is very worse in near end speech large energy. In Fig 4.20, the channel correlation ρ_h means the change of the room impulse is correlation with the origin room impulse. When ρ_h is larger, it means the change channel room impulse is very close to the origin room impulse.

From 4.20, we can find that the channel correlation is larger, and the DTD miss probability is also larger. The result is the same to the instinct. If the change room impulse is very close to the origin room impulse, the DTD is hard to detect that echo path change is present. Therefore, the miss probability would be larger. On the contrary, the channel correlation is smaller. This means that the change channel is very different with the origin channel. Consequently; the DTD is easy to detect that echo path change is present. Even ρ_h is negative, the modified detector can detect that echo path is present. However, the conventional detect worse in large channel correlation.

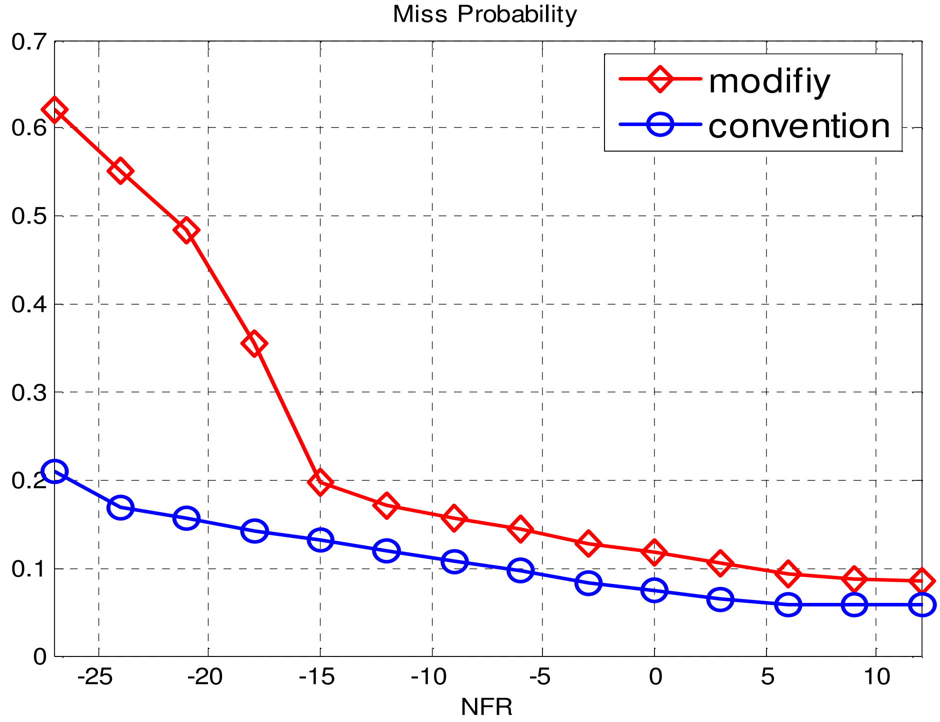


Fig 4.19 The miss probability in DT period of modified cross correlation DTD in different DT energy

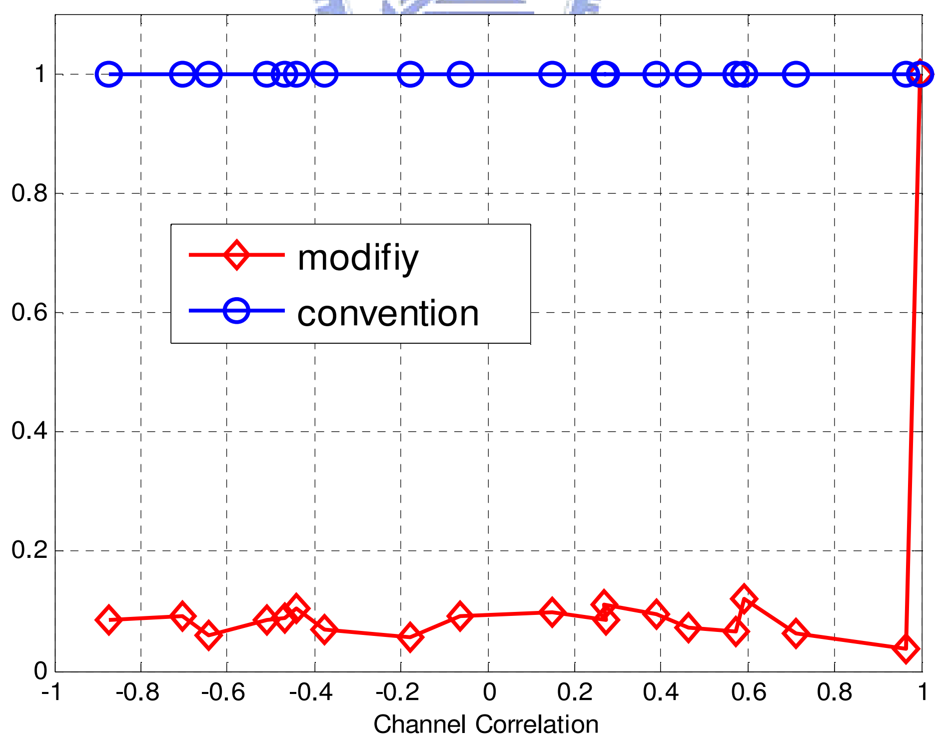


Fig 4.20 Comparison of the miss probabilities of modified cross correlation DTD and cross correlation DTD in different EPC

In section 3.3.3, we propose the weight miss probability. In Fig 4.19, we use the same weight to calculate the miss probability. In Fig 4.21, we modify the weight by near end signal energy. When the near end signal energy is larger, the weight is decreasing. This means that the detector can easy decide correct when near end signal large energy. From Fig.4.21, we can find that the miss probability in near end speech small energy is very close to zero. However, because of the weight, the weight miss probability is larger than no weight miss probability. Therefore, the detector can not detect all cases, and the weight miss probability is more penalty in near end speech larger energy.

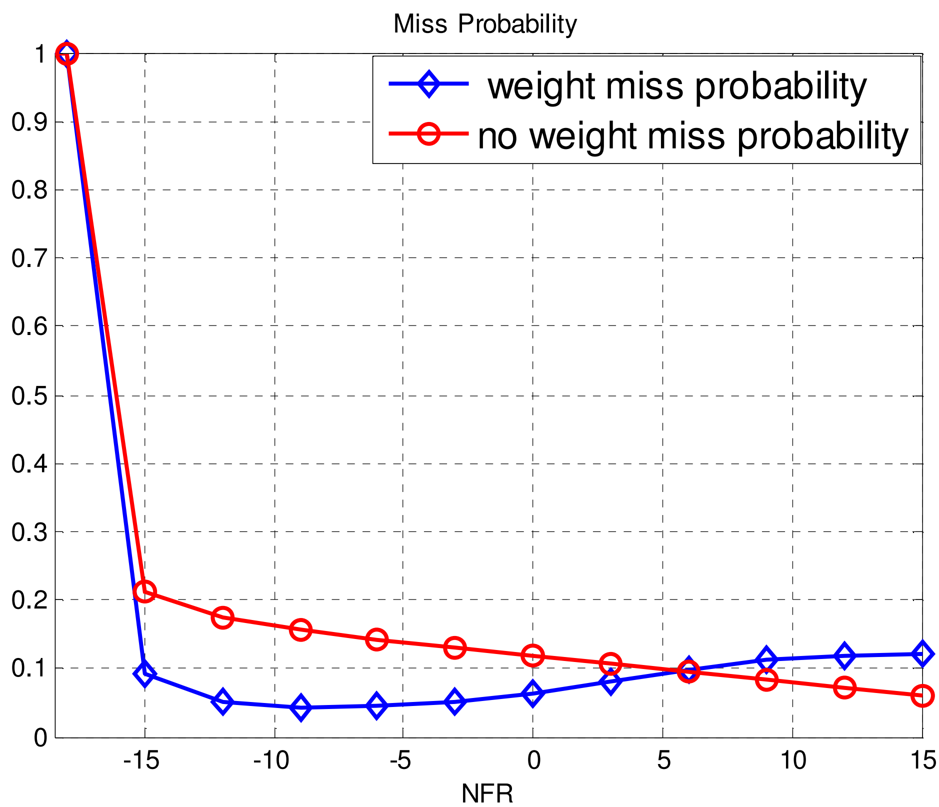


Fig 4.21 Comparison of the weighted miss probability with the miss

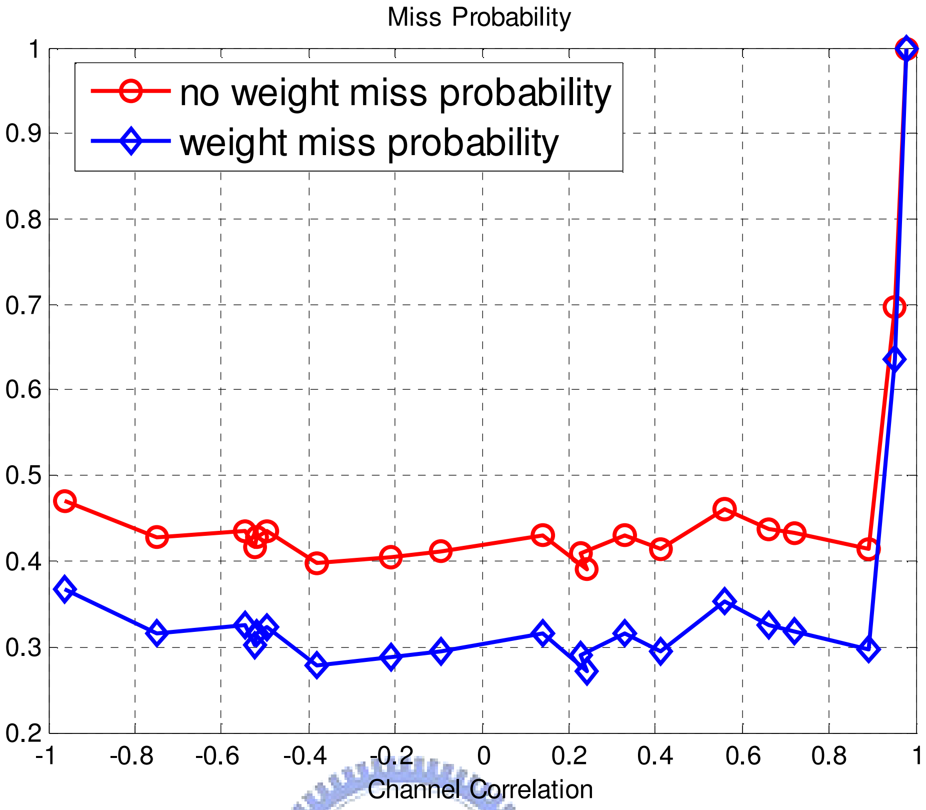


Fig 4.22 Comparison of the weighted and no weighted miss probability

From Fig 4.22, we can find that the miss probability is very large when the channel correlation is very close to one. This result is the same with the Fig 4.18. The weight miss probability is always smaller than the no weight miss probability.

4.9 Variant threshold in cross correlation DTD

In section 3.4, we discussed the selecting variant threshold can improve performance. We use the Fig 3.9 algorithm to verify our algorithm. We simulate the cross correlation DTD under the fixed threshold and the variant threshold. We also simulate the condition in real speech.

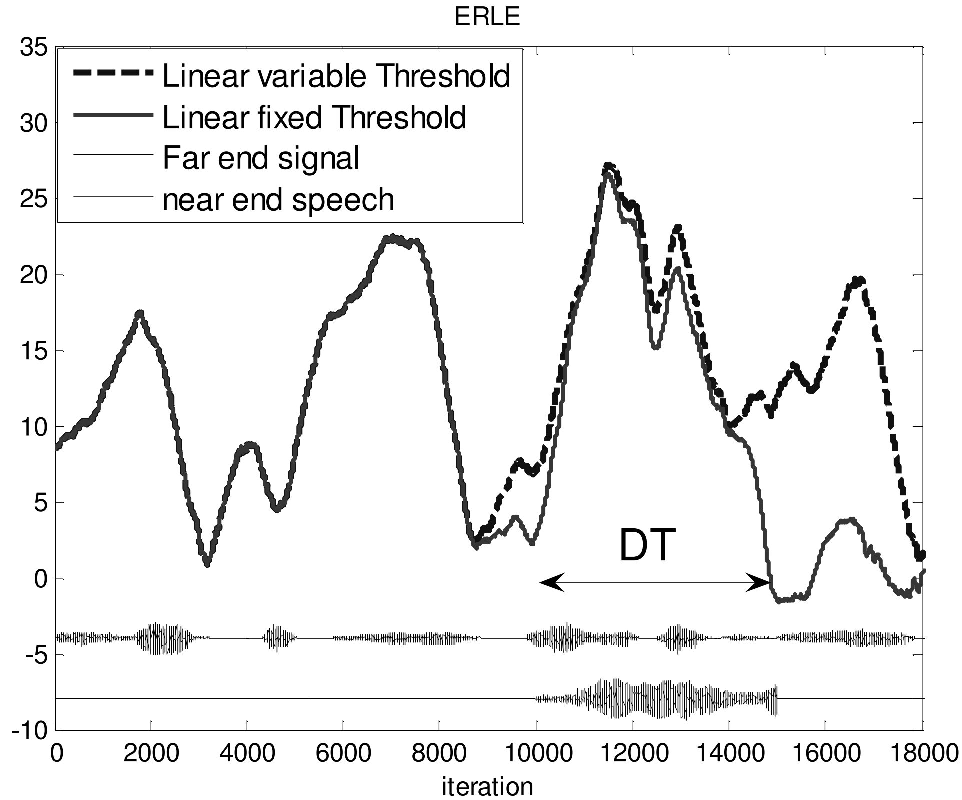


Fig 4.23 Comparison of the fixed threshold and the variant threshold in cross correlation DTD

From Fig 4.23, we can find that ERLE of the variant threshold is better than the fixed threshold when double talk is over. Because the variant threshold can faster detect that double talk is over.

4.10 Square inverse law

In section 2.5, we derive that the miss probability is function of the false alarm probability. Moreover, using some approximation, we can derive (2.5.10). From Fig 4.24, we can find that the smaller κ' is better. Therefore, the mic/AEC correlation DTD is best. The result is the same the Fig 4.3. We can use the different point to check the DTD performance in double talk period. Because we have P_m and P_f , we use the minimum least squares solution (LS) to find $\kappa' = 7 \cdot 10^{-3}$ without finding the Geigel DTD parameter PDF. Similarly, the Mic/error correlation DTD has the different $\kappa' = 4.6 \cdot 10^{-3}$. Therefore, we can extend the idea to other DTDs without solving PDF. Moreover, we can find that the smaller κ' is better. We can use κ' to decide the detector is good or bad.

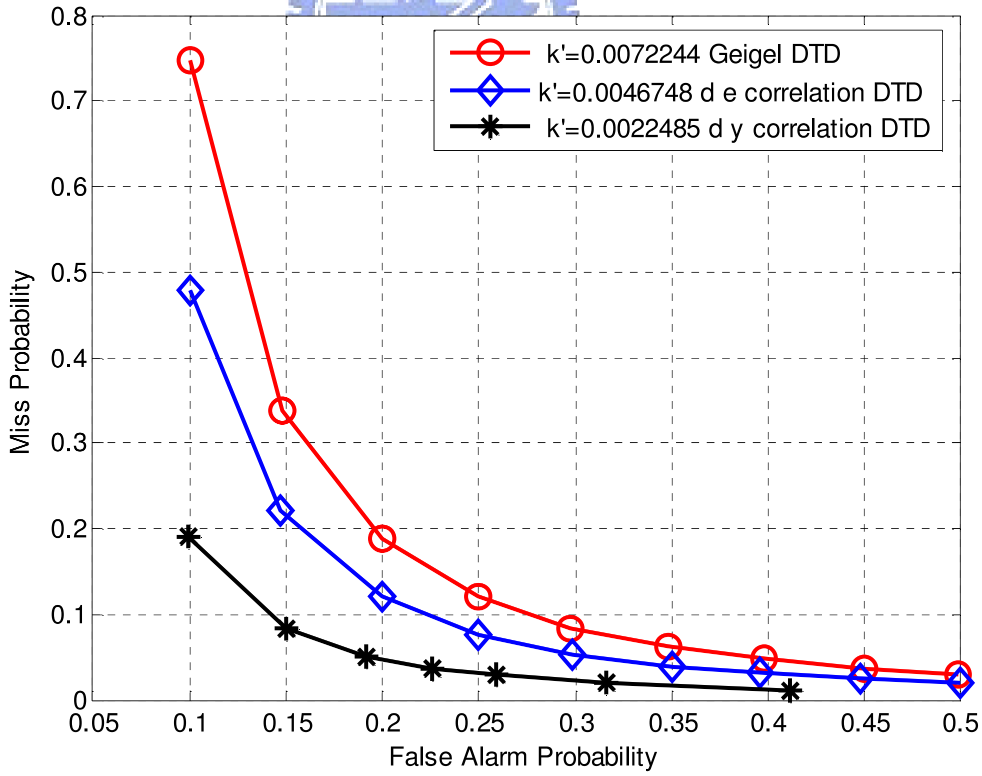


Fig 4.24 The miss probability for different κ' under fixed false alarm probability

Chapter 5

Conclusions

In the hands-free mobile radio telephone or the teleconference systems, the DT will degrade the performance of the NLMS algorithm. However, the DTD is hard to differentiate DT and EPC. We propose the modified cross-correlation DTD. The modified cross-correlation DTD can make correct decision between DT and EPC.

In Chapter 2, we derive the theoretical miss probability in DT period for mic/AEC correlation DTD. The robust DTD is used by the buffer range. The robust DTD can improve the performance. The nonlinear effect is also discussed in this section.

In Chapter 3, we propose a novel technique for evaluating DTD in EPC. The technique can calculate the miss probability in the different level of echo path change. We also propose the variant threshold instead of the fixed one.

Computer simulations is shown in Chapter 5. The result show that

(I) Robust DTD using the buffer range can improve the performance.

(II) The miss probability is a function of the false alarm probability

(III) The modified cross-correlation DTD can make correct decision between DT and EPC.

(IV) We propose a evaluating technique for DTD in EPC

Bibliography

[1] *S. Haykin, Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2002

[2] D. L. Duttweiler, “A twelve-channel digital echo canceler,” *IEEE Trans. Commu.*, vol. COM-26, pp. 647–653, May 1978

[3] P. C. Yip and D. M. Etter, “An adaptive multiple echo canceller for slowly time-varying echo paths” , *IEEE trans. Communication, vol. 38, no.10 ,pp.1693-1698,*Oct. 1990

[4] Rohrs C. and Youncer, “Double talk detector for echo canceler and method”, *U.S. patent* #4,918,727, April 17, 1990

[5] T. Creasy and T. Aboulnasr, “A projection-correlation algorithm for acoustic echo cancellation in the presence of double talk”, in Proc. *IEEE ICASSP 2000 (Istanbul, Turkey)*, vol. 1, pp. 436439, May 2000

[6] Kazuo Ochiai, Takashi Araseki and Takashi Ogihara, "Echo canceller with two echo path models", *IEEE trans*. Communication, vol.com-25,no.6, June.1977

[7] H. Ye and B.-X. Wu, “A new double-talk detection algorithm based on the orthogonality theorem,” *IEEE Trans.* Commun., vol. 39, pp.1542–1545, Nov. 1991.

[8] J. H. Cho, D. R. Morgan, and J. Benesty, “An objective technique for evaluating doubletalk detectors in acoustic echo cancelers,” *IEEE Trans. Speech Audio Processing*, vol. 7, pp. 718–724, Nov. 1999.

[9] Lindstrom, F. , Schuldt, C., and Claesson, I. "An Improvement of the Two-Path Algorithm Transfer Logic for Acoustic Echo Cancellation" *2007IEEE Trans* , vol 15, pp 1320-1326

[10] A. Stenger and W. Kellermann, “Adaptation of a memoryless preprocessor for nonlinear acoustics echo canceling” *Signal Processing* , vol 80, pp.1741-1760, Sept . 2000.

[11] Asharif, M.R. and Shimabukuro, A. and Hayashi, T. and Yamashita,K. “Expanded CLMS algorithm for double-talk echo canceling” *1999 IEEE International Conference on*, Volume: 1, On page(s): 998-1002 vol.

[12] Benesty, J. and Morgan, D.R. and Cho, J.H. "A new class of doubletalk detectors based on cross-correlation " *2000 IEEE Trans.* Volume: 8, pp168-172.

[13] T. G¨ansler and M. Hansson and C.-J. Invarsson, and G. Salomonsson, “A double-talk detector based on coherence,” *IEEE Trans*. Commun., vol.44, pp. 1421–1427, Nov. 1996.

[14] J. C. Jenq and S. F. Hsieh, “Decision of double-talk and time-variant echo path for acoustic echo cancellation,” *IEEE Signal Process*. Lett., vol. 10, no. 11, pp. 317–319, Nov. 2003.

[15] T. Jia, Y. Jia, J. Li, and Y. Hu, “Subband doubletalk detector for acoustic echo cancellation systems,” *2003 IEEE Int.* Conf.Acoustics, Speech, Signal Processing (ICASSP), Apr. 2003, vol. 5, pp.604–607.

[16] Meng Wang ,Li-rong Dai ,Yan Song ,and Ren-hua Wang “A Double-talk Robust and Fast Converging Echo Canceller” *2005 IEEE ICIICS* , pp 1316- 1319.

[17] James D. Gordy, Rafik A. Goubran “Statistical Analysis of Doubletalk Detection for Calibration and Performance Evaluation” *2006 IEEE Trans.* Volume: 15, pp 1035-1043