

行政院國家科學委員會專題研究計畫 成果報告

以極大相似性對話偵測作回音消除之研究

計畫類別：個別型計畫

計畫編號：NSC91-2219-E-009-047-

執行期間：91年08月01日至92年07月31日

執行單位：國立交通大學電信工程學系

計畫主持人：謝世福

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 92 年 10 月 14 日

# 行政院國家科學委員會專題研究計畫成果報告

## 以極大相似性對偵測作回音消除之研究

### Acoustic echo cancellation based on maximum likelihood double talk detection

計畫編號: NSC 91-2219-E-009-047

執行期限: 91 年 08 月 01 日至 92 年 07 月 31 日

主持人: 謝世福 交通大學電信系副教授

計畫參與人員: 鄭張權

#### 一、摘要

傳統的回音消除在“雙邊對話”(double-talk)時，濾波器會做錯誤的調整而失效。我們提出 IMLC AEC and Double Talk Detection (IMLC&DTD) 架構加以改進。這個架構主要的概念是觀察濾波器係數的平方值(square value) 變化的情形，我們分別假設了它們的機率分佈，進一步發展一個 Maximum Likelihood Test 偵測器並做統計分析之 Receiver Operating Characteristic (ROC) 曲線，以尋求較佳的 Neyman-Pearson decision rule 作為偵測器的法則。它能有效區別“雙邊對話”，或“脈衝響應在變化”兩種狀況。我們也從電腦模擬分析得到證明。

關鍵詞: 回音消除，雙邊對話

#### ABSTRACT

In this project, a new double-talk detection method is considered by tracking the squared coefficients errors of the AEC filter by assuming the difference of echo path change to be uniformly distributed. Based on a likelihood ratio test, the statistical detection performance is shown by a Receiver Operating Characteristic (ROC), from which the Neyman-Pearson decision rule can be determined. Therefore, cases of single-talk, double-talk and echo path changes are effectively detected.

#### 二、緣由與目的

一般的適應性濾波器(例如 NLMS, RLS 等)用來做音場回音消除，有一個嚴重的問題，就是在“雙邊對話”(double-talk)時，濾波器會做錯誤的調整而失效。近年來大量的文獻都集中在:如何有效提供一個雙邊對話偵測器，以避免濾波器做錯誤的調整[1,2,3,8,9]。但因為偵測器很難區別到底發生了“雙邊對話”，或是房間脈衝響應在變化。故仍有許多問題有待克服。本研究計畫之目的乃針對上述困難點，提出解決方法。首先我們研究近年來幾種抵抗雙邊對話適應性濾波器，其中我們較有興趣的是以 MLC (maximum length correlation) 來估計房間脈衝響應的方法[4]。MLC 方法雖然較能抵抗雙邊對話，但是它深受遠端和近端語音的影響，效能很差(縱使只有單邊遠端語音發話)以致於優點無法發揮。我們發現 MLC 方法中，遠端語音已知，故認為可以排除其影響，故提出 IMLC 的架構[6]，不過由於 IMLC 的架構中近端語音的影響還不能完全排除。我們再提出 IMLC AEC and Double Talk Detection (IMLC&DTD) 架構加以改進[7]。這個架構主要的概念是

觀察濾波器係數的平方值(square value) 變化的情形,根據分析當房間脈衝響應變化較大時,這個架構很容易區別“雙邊對話”,或脈衝響應在變化.我們分別假設了它們的機率分佈.本計畫的目的是進一步發展一個較好的 Maximum Likelihood Test 偵測器並做統計分析之 Receiver Operating Characteristic (ROC)曲線,並尋求較佳的 Neyman-Pearson decision rule 作為偵測器的法則,以便除去 IMLC&DTD 架構中需要脈衝響應變化較大的限制[10].

### 三、IMLC&DTD 回音消除模型

典型的 IMLC&DTD 的音場回音消除(Acoustic echo cancellation)方塊圖如圖一所示.回音消除是採用適應性有限脈衝響應濾波器 FIR (finite impulse response)來達成的.首先用它來模擬揚聲器與麥克風之間的房間脈衝響應 RIR (room impulse response)其次是複製回音,最後拿複製回音與其正的回音相減.由於遠端語音  $s(n)$ 已知,我們利用 IMLC 法排除其影響[6].以 IMLC 法所估計的濾波器係數可表為:

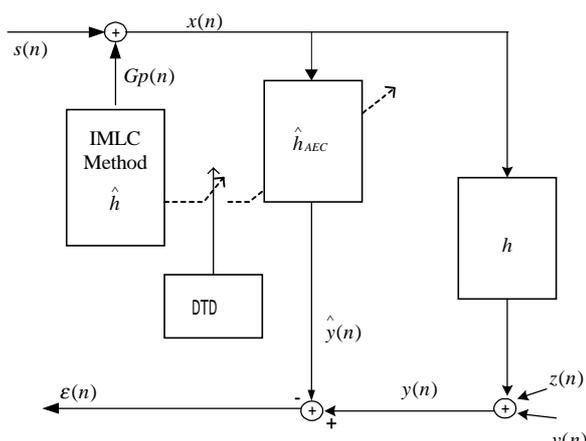
$$\hat{h}_m(n) = h(n) + I_{N,m}(n) - I_{F,m}(n) * [\hat{h}_{m-1}(n) - h(n)] \quad (1)$$

其中  $m$  是疊代的次數,而  $I_N(n)$  及  $I_F(n)$  是近端和遠端語音的干擾表為

$$I_N(n) = \frac{1}{G(L+1)} p(n) \odot u(n) - \frac{1}{(L+1)} \sum_{k=1}^M h(k) \quad (2)$$

$$I_F(n) = \frac{1}{G(L+1)} p(n) \odot s(n)$$

不過由於 IMLC 的架構中近端語音的影響還不能完全排除.我們再提出 IMLC AEC and Double Talk Detection (IMLC&DTD)架構加以改進[7].這個架構主要的概念是觀察濾波器係數的平方值  $\|e_m\|^2$  (square value) 變化的情形,根據分析當房間脈衝響應變化較大時,這個架構很容易區別“雙邊對話”,或脈衝響應在變化.



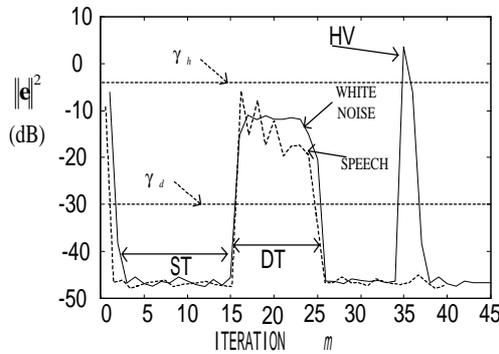
圖一 IMLC&DTD 回音消除方塊圖

器係數的平方值  $\|e\|^2$  在各種狀況(單邊發話,雙邊對話及脈衝響應變化)發生的情形如圖二所示.

$\|e_m\|^2$  的機率分佈可表為[7]:

$$\begin{aligned}
 p(\|e\|_{st}^2) &= N\left(\frac{M\sigma_v^2}{G^2(L+1)}, 2M\left[\frac{\sigma_v^2}{G^2(L+1)}\right]^2\right) \\
 p(\|e\|_{dt}^2) &= N\left(\frac{M(\sigma_z^2 + \sigma_v^2)}{G^2(L+1)}, 2M\left[\frac{\sigma_z^2 + \sigma_v^2}{G^2(L+1)}\right]^2\right) \quad (3) \\
 p(\|e\|_{hv}^2) &= N\left(\|f-h\|^2 + \frac{M(\|f-h\|^2 \sigma_s^2 + \sigma_v^2)}{G^2(L+1)}, 2M\left[\frac{\|f-h\|^2 \sigma_s^2 + \sigma_v^2}{G^2(L+1)}\right]^2\right)
 \end{aligned}$$

當房間脈衝響應變化較大時，例如  $2 < \|f-h\|^2$ ，這個架構很容易區別“雙邊對話”，或脈衝響應在變化，但是當房間脈衝響應變化不大時，例如  $0 < \|f-h\|^2 < 1$ ，雙邊對話的偵測較為困難。



圖二.  $\|e\|^2$  在各種狀況變化的情形

為了進一步發展一個較好的雙邊對話偵測器以便除去 IMLC&DTD 架構中需要脈衝響應變化較大的限制。我們的構想如下:在 Eq (3) 中為了方便分析，重新定義  $e_0(n) = e_{st}(n)$ ， $e_1(n) = e_{hv}(n)$  及  $e_2(n) = e_{dt}(n)$ 。當  $\Delta h = 0$  則  $e_1(n) = e_0(n)$ 。即  $e_0(n)$  是  $e_1(n)$  的特例。我們可以將 Eq (3) 簡化為

$$H_i: p(\mathbf{e}) = p(\mathbf{e}_i) = \frac{1}{(2\pi)^{M/2} |\mathbf{K}_i|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{e}' - \mathbf{m}_i' \mathbf{K}_i^{-1}(\mathbf{e} - \mathbf{m}_i))\right], \quad i=1, 2 \quad (4)$$

其中  $\mathbf{e} = [e(1), e(2), \dots, e(M)]^t$        $\mathbf{K}_1 = \zeta_1^2 \mathbf{I}_{M \times M} = \frac{(\Delta h)\sigma_s^2 + \sigma_v^2}{G^2 L} \mathbf{I}_{M \times M}$ ，

$\mathbf{K}_2 = \zeta_2^2 \mathbf{I}_{M \times M} = \frac{\sigma_z^2 + \sigma_v^2}{G^2 L} \mathbf{I}_{M \times M}$ ， $\mathbf{m}_1 = \Delta \mathbf{h}$  and  $\mathbf{m}_2 = \mathbf{0}$ 。依據 likelihood ratio 法

則，雙邊對話偵測可以表為： $\frac{p(\mathbf{e}_1)}{p(\mathbf{e}_2)} > \Lambda$ ，或

$$\begin{aligned}
 \frac{p(\mathbf{e}_1)}{p(\mathbf{e}_2)} > \Lambda, \quad \gamma = \frac{\zeta_1^2 \zeta_2^2}{\zeta_1^2 - \zeta_2^2} (2 \ln(\frac{\zeta_1}{\zeta_2} \Lambda) + \frac{1}{\zeta_1^2} \mathbf{m}_1' \mathbf{m}_1 - \frac{2}{\zeta_1^2} \mathbf{e}' \mathbf{m}_1) \quad \dots\dots\dots
 \end{aligned}$$

...(5)

不過 Eq.(5)中  $\mathbf{m}_1 = \mathbf{f} - \mathbf{h}$  未知, 使得  $\mathbf{e}'\mathbf{m}_1$  和  $\mathbf{m}_1'\mathbf{m}_1$  也未知, 造成雙邊對話偵測仍有困難而有待解決. 現在我們假設  $\|\mathbf{f} - \mathbf{h}\|$  為 uniformly distributed 且在  $[0, 2]$

變化. 則  $\Delta = \|\mathbf{f} - \mathbf{h}\|^2$  的機率分佈為  $f(\Delta) = 1/(4\sqrt{\Delta})$  且在  $[0, 4]$  變化

我們的策略是當  $\|\mathbf{e}\|^2 \in R_2 = [\mu_2 - x\sigma_2, \mu_2 + x\sigma_2]$  時判定為  $H_2$  其中  $x$  可以決定區

間的大小[10]. 而 probability of detection  $p_D$  表示為  $p_D = \text{prob}(\|\mathbf{e}\|^2 \in R_2 | H_2)$  另外

false alarm probability  $p_F$  表示為  $p_F = \text{prob}(\|\mathbf{e}\|^2 \in R_2 | H_1)$ . 它們可表為:

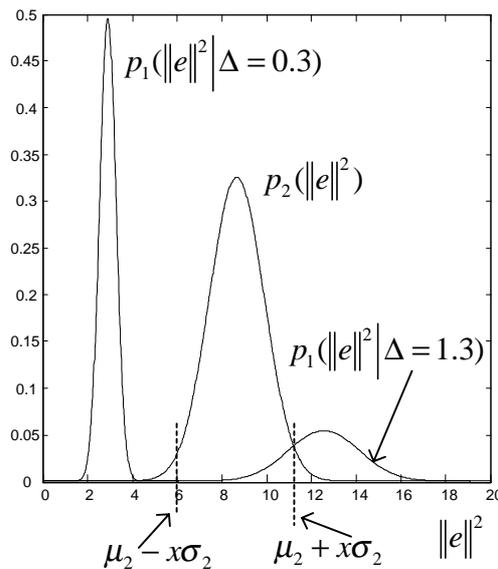
$$\begin{aligned} p_D &= \int_{\mu_2 - x\sigma_2}^{\mu_2 + x\sigma_2} p_2(\|\mathbf{e}\|^2) d\|\mathbf{e}\|^2 \\ p_F &= \int_0^4 p(\Delta h) \int_{\mu_2 - x\sigma_2}^{\mu_2 + x\sigma_2} p_1(\|\mathbf{e}\|^2) d\|\mathbf{e}\|^2 d(\Delta h) \end{aligned}, \quad (6)$$

我們藉由 ROC 曲線(即  $p_F$  對  $p_D$  做圖)並尋求較佳的 Neyman-Pearson decision rule 作為偵測器的法則, 以便除去 IMLC&DTD 架構中需要脈衝響應變化較大的限制. 它能有效區別“雙邊對話”, 或“脈衝響應在變化”兩種狀況. 我們也從電腦模擬分析得到證明.

#### 四、電腦模擬

圖三表示  $p_1(\|\mathbf{e}\|^2)$  和  $p_2(\|\mathbf{e}\|^2)$  在  $G=0.15, L=512, M=100, \sigma_z^2 = \sigma_s^2 = 1$ , 及

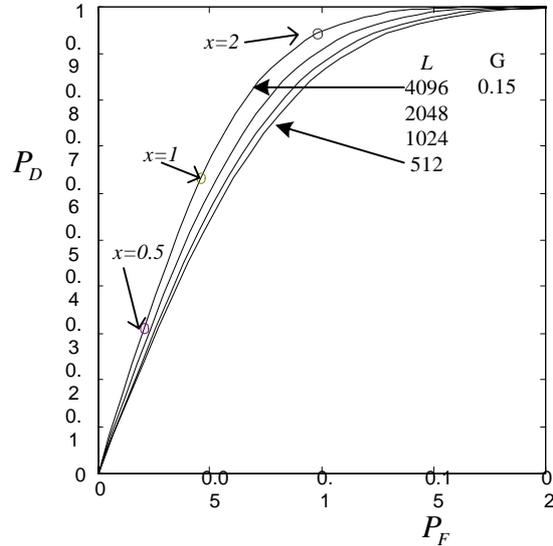
$\sigma_v^2 = 0.0001$ . 變化的情形 例如當  $\Delta = 1.3$  或  $0.3$  時  $p_1(\|\mathbf{e}\|^2 | \Delta)$  可能位於  $p_2(\|\mathbf{e}\|^2)$  的右



邊或左邊.

圖三.  $p_1(\|e\|^2)$  和  $p_2(\|e\|^2)$  隨  $\Delta$  變化情形

當  $x$  不同的時候  $p_F$  對  $p_D$  的作圖. 如圖四 所示. 訊號  $s(n)$ ,  $z(n)$  和  $v(n)$  為 white Gaussian,  $G=0.15$  (-16dB),  $L$  由 512 變化至 4096. 當區間  $[\mu_2 - x\sigma_2, \mu_2 + x\sigma_2]$  加大時,  $p_D$  和  $p_F$  也隨之加大. 由 ROC curve 可作為偵測器的法則, 例如  $L=4096$ , 且  $p_F=0.1$  被允許, 則  $x=2$ ,  $p_D$  可保證達到 0.95.



圖四. ROC curve

## 五、結論

我們提出 IMLC&DTD 架構改進發展一個較好的雙邊對話偵測器以便除去 IMLC&DTD 架構中需要脈衝響應變化較大的限制. 這個架構主要的概念是觀察濾波器係數的平方值變化的情形, 我們分別假設了它們的機率分佈, 進一步發展一個 Maximum Likelihood Test 偵測器並做統計分析之 ROC 曲線, 以尋求較佳偵測器的法則. 它能有效區別 “雙邊對話”, 或 “脈衝響應在變化” 兩種狀況. 我們也從電腦模擬分析得到證明.

## REFERENCE

- [1] E. Hansler, "The hands-free telephone problem; an annotated bibliography," *Signal Processing*, vol.27, pp. 259-271, 1992.
- [2] N. Furuya, *et al.*, "High performance custom VLSI echo canceller," in *ICC'85 Rec.*, pp. 1470-1476, 1985.
- [3] S. Minami, *et al.*, "A double-talk detection method for an echo canceller," in *ICC'85 Rec.*, pp. 1492-1497, 1985.
- [4] J. F. Doherty, *et al.*, "A robust echo canceler for acoustic environments," *IEEE Trans on circuit and systems-II.*, vol.44. no.5, pp389-396, May 1997.
- [5] M. J. Crocker, *Handbook of Acoustics*, John Wiley & Sons, Inc, 1998.
- [6] J.C. Jenq and S.F. Hsieh, "A double-talk resistant echo cancellation based on iterative maximal-length correlation," *IEEE Symposium on circuits and systems*, vol.5, pp237-241,

May 2000.

- [7] J.C. Jenq and S.F. Hsieh, " Acoustic Echo Cancellation Using Iterative Maximal Length Correlation and Double-Talk Detection, " *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 9, no. 8, pp932-942, Nov 2001
- [8] K. Ochiai *et al.*, "Echo canceller with two echo path models, " *IEEE Trans. Commun.*, vol. COM-25, pp. 589-595, June 1977.
- [9] Hua Ya and Bo-Xiu Wu, "A new double-talk detection algorithm based on the orthogonality theorem, " *IEEE Trans. Communications*, vol. 39, no. 11, pp. 1542-1545, Nov 1991.
- [10] J.C. Jenq and S.F. Hsieh, "Decision of Double-Talk and Time-Variant Echo Path for Acoustic Echo Cancellation, " *IEEE Signal Processing Letters*, to be appear at Nov. 2003.