

# 行政院國家科學委員會專題研究計畫成果報告

## 以極長相關法作音場回音消除之研究

### Acoustic Echo Cancellation Using Iterative-Maximal-Length Correlation and Double-Talk Detection

計畫編號: NSC 89-2213-E-009-191

執行期限: 89年08月01日至90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、中文摘要

利用適應性有限脈衝響應濾波器(Adaptive FIR filter), 來處理音場回音消除, 有三個主要困難點: 第一) 房間脈衝響應非常長, 通常需要數千個濾波器係數來估計, 而計算這些係數的運算量非常大. 第二) 房間脈衝響應對室內物品人員的移動十分敏感, 適應性濾波器要迅速反應, 故運算速度要很快. 第三) 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三個困難點, 提出一序列的解決方法, 尤其希望在"雙邊對話"發生時適應性濾波器仍能有效消除回音. 在過去一年我們已經對近年來幾種抵抗雙邊對話適應性濾波器做過研究, 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]: 由遠端語音(far-end speech)加入一個低準位虛擬亂碼序列 PN (pseudo-random noise)並在麥克風輸出端得出訊號. 再與 PN 做互相關(cross correlation)運算. 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. 我們認為如果可以排除遠端語音的影響, 則 MLC 有很大的改進空間. 本計畫將集中在排除遠端和近端語音的影響, 使濾波器仍能有效消除回音, 同時運算簡單又可以平行處理, 達到即時反應的目地.

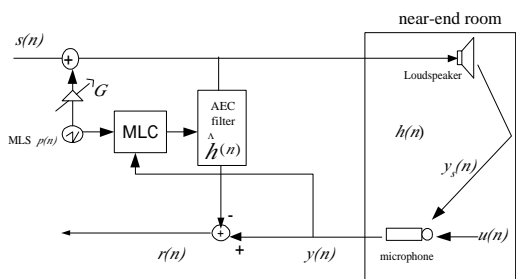
**關鍵詞:** 適應性濾波器, 音場回音消除. 雙邊對話偵測 Maximal Length Sequence, Echo cancellation, Double talk detection, Adaptive filter  
**ABSTRACT**

The conventional maximal length correlation (MLC) algorithm to estimate room impulse response for adaptive echo cancellation (AEC) is disturbed by both far-end and near-end speeches. In this paper, a new Iterative Maximal Length Correlation (IMLC) algorithm is proposed to reduce the far-end speech interference. To avoid the near-end interference, a new double-talk detection method is proposed by tracking the squared coefficients errors of the AEC filter. This DTD method has well-separated detection margins among single-talk, double-talk and echo path changes. Statistical analysis and computer simulations confirm that our proposed IMLC-DTD algorithm outperforms conventional methods.

## 二、緣由與目的

典型的音場回音消除(Acoustic echo cancellation)是採用適應性有限脈衝響應濾波器 AFIR (adaptive 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。

由於傳統適應性濾波器，會因為雙邊對話而做錯誤的調整而失效。 J.F. Doherty [4] 等人提出：由遠端語音  $s(n)$  加入一個低準位 MLS  $p(n)$  其幅度由  $G$  控制並在麥克風輸出端得出訊號，輸出訊號再與  $p(n)$  做互相關運算以估計 RIR。它的主要觀念是根據"聽覺遮沒效應 (auditory masking effect)[5]:在聽覺上,當訊的功率比大於 15dB 時,雜訊會被遮沒而無法聽見又由於語音與 MLS 不相關,故估計 RIR 時較能抵抗雙邊對話。典型的 MLC 法應用在回音消除如圖一.所示



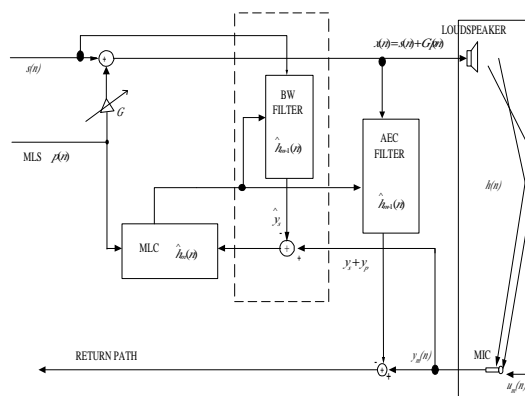
圖一. 以 MLC 法設計回音消除

不過這個方法估計的脈衝響應受遠端語音  $s(n)$ ，和近端語音  $u(n)$  的影響，雖然 PN 的長度  $L$  足夠長時，較能抵抗雙邊對話，但效能仍很差。以致於優點無法發揮。我們首先提出可以排除遠端語音影響的 IMLC 法[6]使它在單邊遠端語音發話時。效能和一般的適應性濾波器一樣好，同時運算簡單又可以平行處理。

### 三. 以 IMLC 法設計回音消除

IMLC 的架構如圖二所示。由於遠端語音  $s(n)$  已知，我們認為應該可以排除其影響。工作原理簡述如下：我們利用 BW Filter 來估計回音  $y_s(n)$ 。當以 MLC 來估計房間脈衝響應時，麥克風輸出端訊號，先減去估計之回音  $y_s(n)$  再

與 PN 做互相關，以降低受遠端語音  $s(n)$  的影響。



圖二. IMLC 的架構圖

以 IMLC 法所估計的濾波器係數可表為：

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

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

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

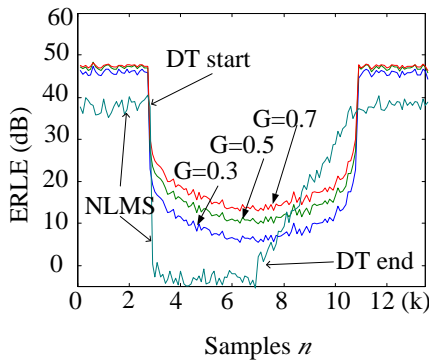
$$I_F(n) = \frac{1}{G(L+1)} p(n) \odot s(n)$$

由(1)可以發現當  $\hat{h}_{m-1}(n)$  非常接近  $h(n)$  時，遠端語音的干擾可以被排除。如果估計的濾波器係數誤差定義為  $e_m(n) \equiv \hat{h}_m(n) - h(n)$  則在雙邊對話  $dt$  和單邊發話  $st$  情形下可表為[7]:

$$H_{dt} : \begin{cases} e_{MLC,m}(n) = \frac{1}{G(L+1)} p(n) \odot [z(n) + v(n)] \\ - \frac{1}{(L+1)} \sum_{k=1}^M h(k) + \frac{1}{G(L+1)} p(n) \odot s(n) * h(n) \\ e_{IMLC,m}(n) = \frac{1}{G(L+1)} p(n) \odot [z(n) + v(n)] \\ - \frac{1}{(L+1)} \sum_{k=1}^M h(k) \end{cases}$$

$$H_{st} : \begin{cases} e_{MLC,m}(n) = \frac{1}{G(L+1)} p(n) \odot v(n) \\ -\frac{1}{(L+1)} \sum_{k=1}^M h(k) + \frac{1}{G(L+1)} p(n) \odot s(n) * h(n) \\ e_{IMLC,m}(n) = \frac{1}{G(L+1)} p(n) \odot v(n) - \frac{1}{(L+1)} \sum_{k=1}^M h(k) \end{cases}$$

這個結構經電腦模擬各種狀況(雙邊對話, 單邊發話)發生的情形, 結果如圖三所示, 可以發現, 與 NLMS 應性濾波器作比較, 較能抵抗雙邊對話, 而且在單邊遠端語音發話時, 效能比 NLMS 好。



圖三. IMLC 法在回音消除的應用

由於近端語音  $u(n)$  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[7]加以改進。

#### 四. IMLC&DTD 架構

IMLC&DTD 的方塊圖如圖四所示, 主要分為兩部份: IMLC 架構 (虛線部份) 及 Logic Decision 以偵測雙邊對話。首先我們觀察輸入訊號  $x_m(n) = x(n) \Big|_{mL+1 \leq n \leq (m+1)L}$  並定義  $ERL$

(echo return loss) 為  $ERL_{F,m} = \frac{\|\mathbf{x}_m\|^2}{\|\hat{\mathbf{a}}_F\|^2}$  和

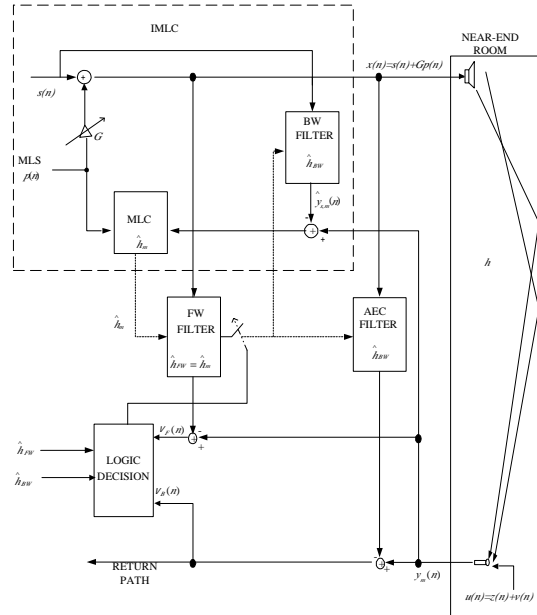
$ERL_{B,m} = \frac{\|\mathbf{x}_m\|^2}{\|\hat{\mathbf{a}}_B\|^2}$ 。其中  $\hat{\mathbf{a}}_F$  和  $\hat{\mathbf{a}}_B$  為圖三所

定義之誤差訊號。我們新提出的偵測雙邊對話的方法如圖五流程圖所示, 描述如下:

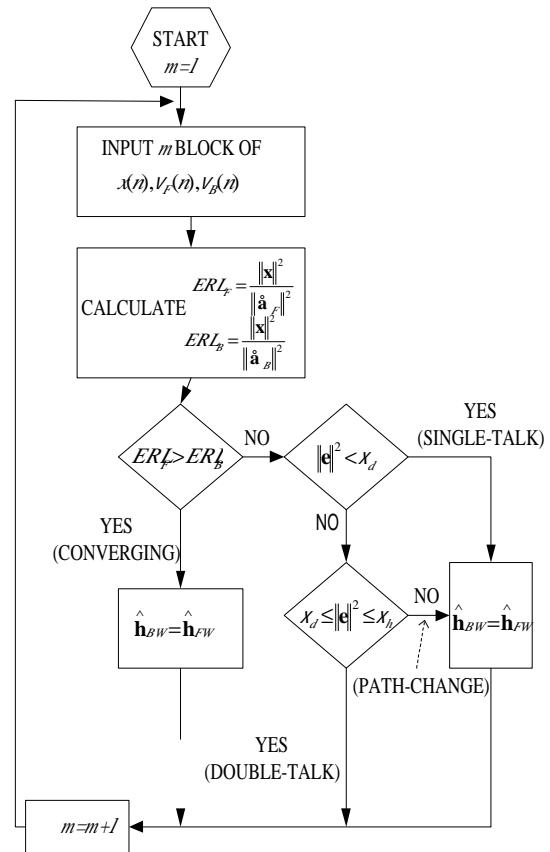
如果  $ERL_{F,m} > ERL_{B,m}$ , 則 AEC 濾波器

$\hat{h}_{BW}(n)$  的係數被  $\hat{h}_{FW}(n)$  所更新。而如果

$ERL_{F,m} \leq ERL_{B,m}$ , 則我們觀察  $\|\mathbf{e}_m\|^2$  的大小。



圖四. IMLC&DTD 的架構



圖五. 雙邊對話偵測流程圖

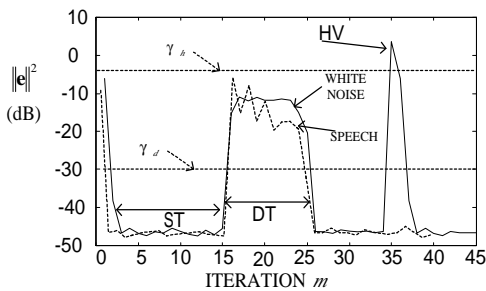
而  $\|e_m\|^2$  的機率分佈可表為[7]:

$$P(\|e_{sv}\|^2) = \mathcal{N}\left(\frac{Mf_v^2}{G^2(L+1)}, 2M\left[\frac{f_v^2}{G^2(L+1)}\right]^2\right)$$

$$P(\|e_{dv}\|^2) = \mathcal{N}\left(\frac{M(f_z^2 + f_v^2)}{G^2(L+1)}, 2M\left[\frac{f_z^2 + f_v^2}{G^2(L+1)}\right]^2\right)$$

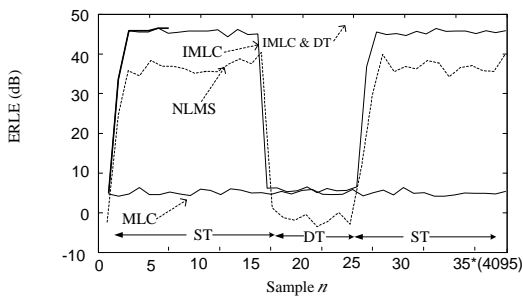
$$P(\|e_{hv}\|^2) = \mathcal{N}\left(2\|h\|^2 + \frac{2M\|h\|^2 f_s^2}{G^2(L+1)}, 2M\left[\frac{2\|h\|^2 f_s^2}{G^2(L+1)}\right]^2\right)$$

其中  $h_v$  表示房間脈衝響應在變化(麥克風與揚聲器的位置輕微變化)。 $\|e\|^2$  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, 如圖六所示



圖六.  $\|e\|^2$  在各種狀況變化的情形

IMLC&DTD 的結構經電腦模擬各種狀況發生的情形, 結果如圖七所示, 可以發現, 與 IMLC 和 NLMS 應性濾波器作比較, 效能最好。



圖七 IMLC&DTD 在各種狀況時 ERLE 的情形  
六..計畫結果與自許

在本計畫中我們研究適用於單邊發話, 雙邊對話及脈衝響應變化發生的回音消除方法。我們利用 IMLC 法排除 MLC 法中遠端語音的影響, 另外我們新提出的偵測雙邊對話的方法排除近端語音的影響, 使濾波器仍能有效消除回音,

同時運算簡單又可以平行處理, 達到即時反應的目地。

## 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. ASSP-, no.
- [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.