

行政院國家科學委員會補助專題研究計畫

期中精簡報告

免持聽筒汽車套件之迴聲與噪音控制系統

計畫類別： 個別型計畫 整合型計畫

計畫編號： NSC 97-2221-E-009-010-MY3

執行期間： 97年 8月 1日至 100年 7月 31日

計畫主持人：白明憲

共同主持人：

計畫參與人員： 林家鴻 劉冠良 劉嫻婷

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計

畫、列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權，一年二年後可公開查

詢

執行單位：國立交通大學 機械工程研究所

中文摘要

回聲消除為免持裝置中之一重要部份，負責消除由於同時包含麥克風以及揚聲器的系統所導致的回授回音。早期的回聲消除主要是透過設計一些適應性濾波器來達成，但使用這種方法需要很大的運算量。因此我們利用串接固定係數濾波器與適應性濾波器並採用依據頻率範圍分段處理來降低非必要的運算量。此方法可應用在許多方面，例如藍芽耳機、手機、視訊會議……等。在適應性濾波器的演算法中，我們採用 LMS 及 NLMS 兩種演算法，並用消相關濾波器以及次頻帶法來提高他們的運算速率和效果。在實驗的部份，我們遵照 ITU-T Recommendation G.168 上的測試方法，並將結果以統計學上的方法 MANOVA 做判定。

英文摘要

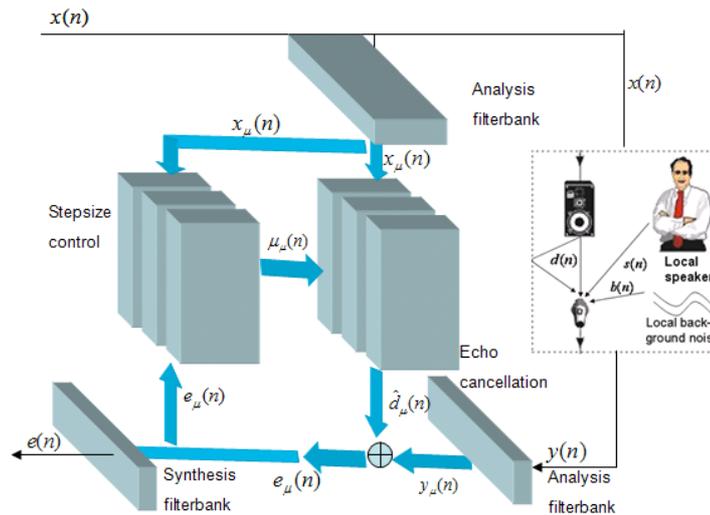
An efficient method is presented for implementing an acoustic echo canceller (AEC) that makes use of hybrid sub-band approach. The hybrid system is comprised of a fixed processor and an adaptive filter in each sub-band. The AEC aims at reducing the echo resulting from the acoustic feedback in loudspeaker-enclosure-microphone (LEM) systems such as teleconferencing and hands-free systems. In order to cancel efficiently the echo, various processing architectures including fixed filters, hybrid processors, and sub-band structure are investigated. A double-talk detector is incorporated into the proposed AEC to prevent the adaptive filter from diverging in double-talk situations. A de-correlation filter is also used alongside sub-band processing in order to enhance the performance and efficiency of AEC. All algorithms are implemented and verified on the platform of a fixed-point digital signal processor (DSP). The AECs are evaluated in terms of cancellation performance and computation complexity. In addition, listening tests are conducted to assess the subjective performance of the AECs. From the results, the proposed hybrid sub-band AEC was found to be the most effective among all methods in terms of echo reduction and timbral quality.

一、文獻探討

聲學回音產生是由於在密閉空間做通訊行為，如講手機或網路通訊 skype 對話時，遠端的訊號經過喇叭的播出經由回音路徑 (echo path) 成為回音訊號而被麥克風所接收到，其中回音路徑的定義為喇叭輸入到麥克風輸出之間的路徑。為了消除回音，吾人設計一個近似於回音路徑的數位濾波器，這種問題很像一個系統鑑別的問題 (system identification problem)，傳統 AEC 數位濾波器的設計方法有兩種，第一種是將數位濾波器用固定濾波器 (fixed filter) 來設計，其方法是事先就將動態的回音路徑給設計好，實作時訊號經過固定濾波器和經過真實的回音路徑兩個做相減，就可以把回音訊號給消除掉；第二種方法是將數位濾波器用適應性濾波器 (adaptive filter) 來設計，最常用的演算法為 least-mean-squares (LMS) 演算法，應用 LMS 演算法中輸出訊號和 wiener 濾波器作遞迴的動作，如果最終誤差訊號能收斂到很小的值，則 wiener 濾波器就能近似於真實的回音路徑，而回音訊號就可以被消掉。

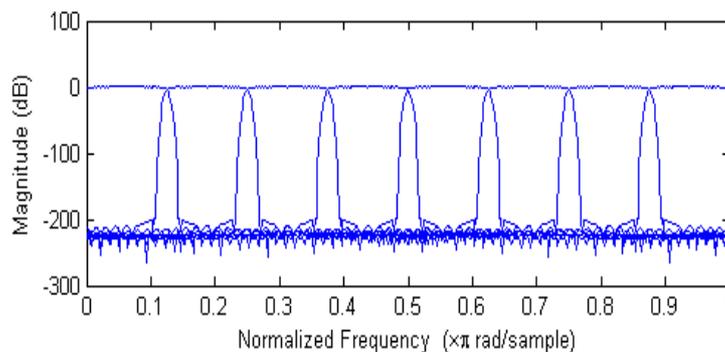
二、分頻濾波器演算法的設計與開發

為了提昇計算效率而消除迴聲的效果，因此本計畫提出的 AEC 將會使用分頻(sub-band)的概念，如圖一(a)所示。因此遠端的訊號 $x(n)$ 以及近端麥克風收到的訊號 $y(n)$ 皆被分成好幾個頻帶。



圖一(a)

圖一(b)是一組分頻率波器(8-band filter bank)將頻帶分成八分。因此我們可以在各個不同的頻帶給它不一樣的階數以及 step size，如此一來就可以針對想要處理的頻帶有更加的消除迴聲效果。

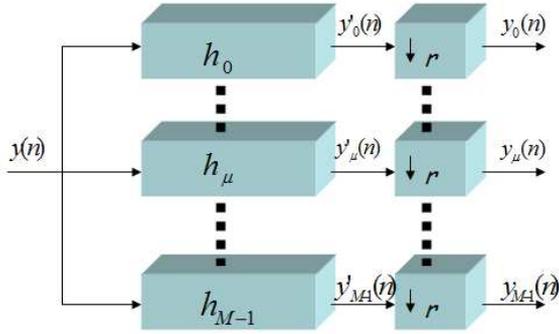


圖一(b)

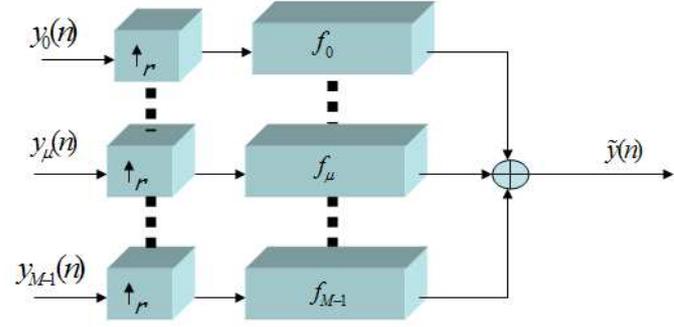
為了簡化處理程序，因此在這些 analysis 和 synthesis filter banks 中會使用 down-sampling 和 up-sampling。由於自適性濾波器在經過 down-sampling 處理

後，由於處於較低的取樣頻率下，所以自適性濾波器的階數可以被減少。而誤差訊號 $e(n)$ 則會從這些分頻的輸出訊號來重新組成並且 up-sampling。

圖一(c)與圖一(d)分別是 analysis filter-bank 和 synthesis filter-bank 的架構，圖中的 r 則是 up/down sampling 的比率。



圖一(c)



圖一(d)

本研究中是使用 cosine-modulated filter-bank (CMFB)，在 CMFB 的設計中，必須先給定一個樣本濾波器 $p_0(n)$ ，此濾波器必須是實係數而且是線性相位的有限長度的低通濾波器(FIR low-pass filter)，其截止頻率為 $\pi/2M$ 。樣本濾波器的長度 N_p

$$N_p \approx \frac{A_s - 7.95}{14.36\Delta f} \quad (1)$$

上式中 A_s 是大小衰減的程度，單位dB； Δf 則是normalized transition bandwidth (Hz)。

analysis filter-banks 的脈衝響應式子如下

$$h_k(n) = 2p_0(n) \cos \left(\frac{\pi}{M(k+0.5)(n - \frac{N}{2} + 1)} + (-1)^k \cdot \frac{\pi}{4} \right), \quad (2)$$

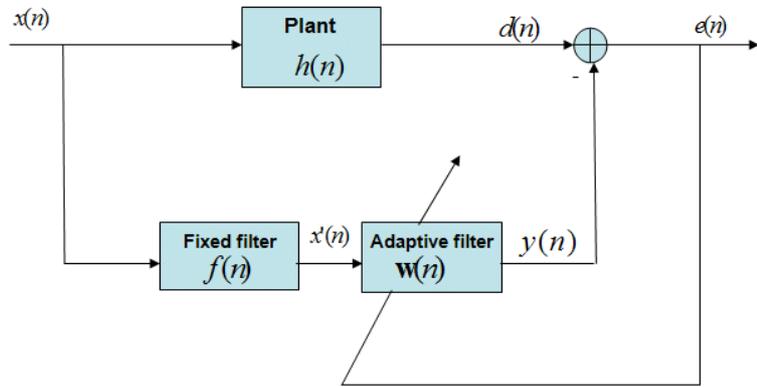
式中 n 是時域變數， $h_k(n)$ 是第 k 個 analysis filter 的係數，而 $p_0(n)$ 則是樣本濾波器的係數， N 係 analysis filter 的階數。另一方面，synthesis filter bank 則如下式

$$f_k(n) = h_k(N + 2 - n), \quad (3)$$

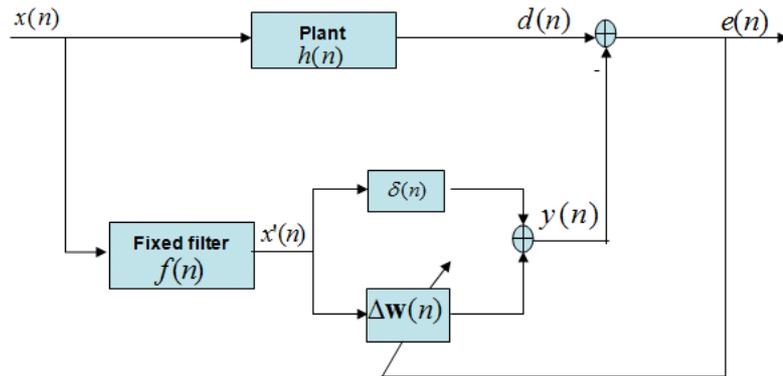
上式 $f_k(n)$ 是第 k 個 synthesis filter 的係數。

二、混合濾波器演算法的設計與開發

在許多 AEC 的問題上迴聲途徑不見得會隨時間而變動，先前描述的 LEM 系統中，麥克風與揚聲器的相對位置大多是固定的，因此我們可以結合固定式濾波器與自適性濾波器。這兩種濾波器經過串聯後，如圖二(a)與圖二(b)。



圖二(a)



圖二(b)

自適性濾波器是一個 FIR (finite impulse response) 濾波器，其中正規項(normal term) $w_0(n) = 1$ ，至於其他擾動項則是 $\Delta w(n) = [w_1(n) w_2(n) w_3(n) \cdots w_{L-1}(n)]$ ， L 是濾波器長度。所以整體可以寫成

$$\hat{h}(n) = f(n) * [\delta(n) + \Delta w(n)] \quad (4)$$

上式中 $\delta(n)$ 是單位脈衝序列，或在 z 平面則

$$\begin{aligned}\hat{H}(z) &= F(z)[1 + \Delta W(z)] \\ &= F(z)[1 + w_1(n)z^{-1} + w_2(n)z^{-2} + \cdots + w_{L-1}(n)z^{L-1}]\end{aligned}\quad (5)$$

使用如此的架構，則當迴聲途徑有些許變動將被擾動項所表示，正規項則將保證固定式濾波器能正常運作，如此混合架構就好像把固定式濾波器跟自適性濾波器的優點結合起來。因為濾波器 $\Delta \mathbf{w}(n)$ 的長度通常非常短，所以整體會有比較低的計算量而 step size 所要調動的範圍變小也將更容易收斂。至於如何更新此系統的濾波器係數，推導如下。首先從誤差訊號開始

$$e(n) = d(n) - y(n) = d(n) - \mathbf{w}^T(n)[f(n) * \mathbf{x}(n)] \quad (6)$$

上式中 $f(n)$ 是固定式濾波器的脈衝響應，單位為時間變數 n

$$\mathbf{w}(n) = [w_0(n) \ w_1(n) \ \cdots \ w_{L-1}(n)]^T \quad (7)$$

定義自適性濾波器的係數向量，單位為時間變數 n

$$\mathbf{x}(n) = [x(n) \ x(n-1) \ \cdots \ x(n-L+1)]^T \quad (8)$$

L 是自適性濾波器的階數。所以我們可以藉由最小瞬時誤差平方來獲得最佳的濾波器

$$\hat{\xi}(n) = e^2(n). \quad (9)$$

代替固定式濾波器，我們可以利用梯度搜尋找到最佳的濾波器係數，

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \frac{\mu}{2} \nabla \hat{\xi}(n) \quad (10)$$

上式中 μ 是 step size 而 $\nabla \hat{\xi}(n)$ 而定義為在時間 n 下的梯度瞬時估測

$$\nabla \hat{\xi}(n) = \nabla e^2(n) = 2[\nabla e(n)]e(n). \quad (11)$$

$$\text{和 } \nabla e(n) = \nabla \{d(n) - \mathbf{w}^T(n)[f(n) * \mathbf{x}(n)]\} = -f(n) * \mathbf{x}(n) \quad (12)$$

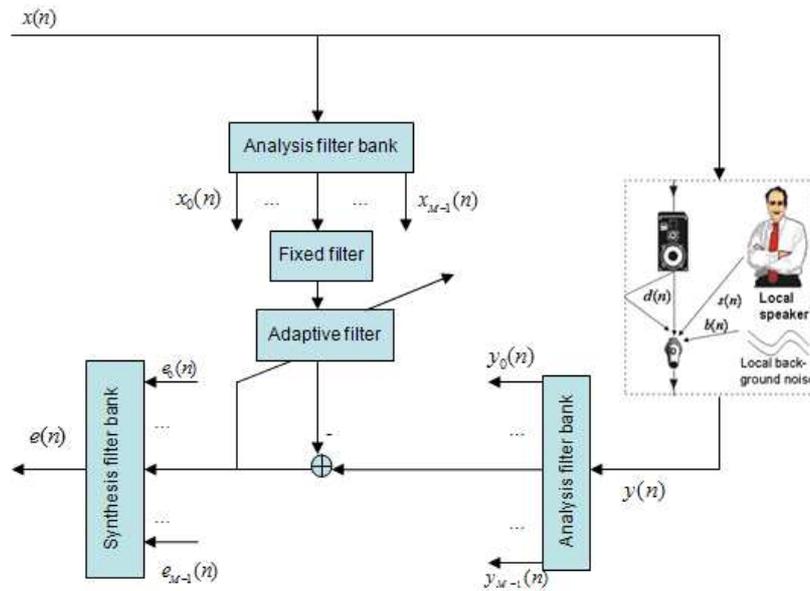
由方程式(20)-(22)可以得到更新的等式

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)[f(n) * \mathbf{x}(n)] \quad (13)$$

而 step size 收斂的範圍如下

$$0 < \mu < \frac{2}{(L + \Delta) r_{\text{rms}}^2} \quad (14)$$

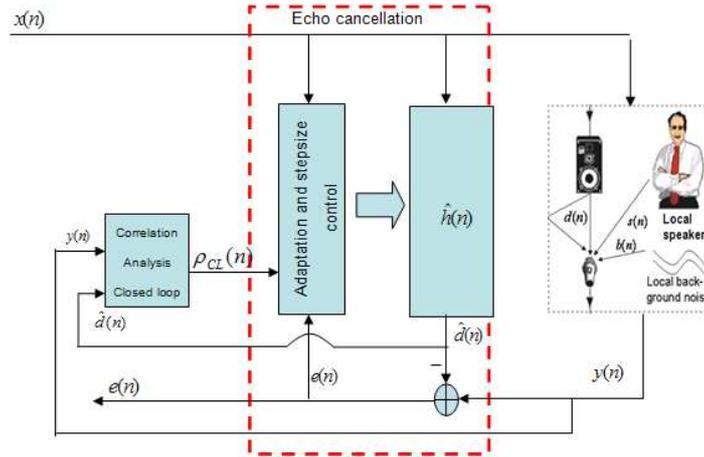
上式中 L 是濾波器長度， Δ 是受控場延遲的時間，而 r_{rms}^2 是被濾過後訊號 $\mathbf{x}'(n) = f(n) * \mathbf{x}(n)$ 的功率。因此這個混合式濾波器(Hybrid filter) 可以與先前的分頻觀念相結合，來達到最大的計算效率以及消除迴聲的效果。圖三則是分頻混合式 AEC 的方塊圖。圖中可看到 M -band analysis 和 synthesis banks 被使用。



圖三

三、相關性偵測器開發

在自適性濾波器用於LEM系統所要考慮的問題就是雙邊通話(Double talk)或單邊通話(Single talk)的情況，如果遠端與使用端同時說話，那麼對於自適性濾波器在收斂上也許會發散掉，為了避免這樣的情況，所以必須利用相關性偵測器來偵測目前是屬於雙邊通話還是單邊通話的情況，雙邊通話則讓自適性濾波器停住不更新其係數，若為單邊通話再更新。如圖十一所示：



圖四

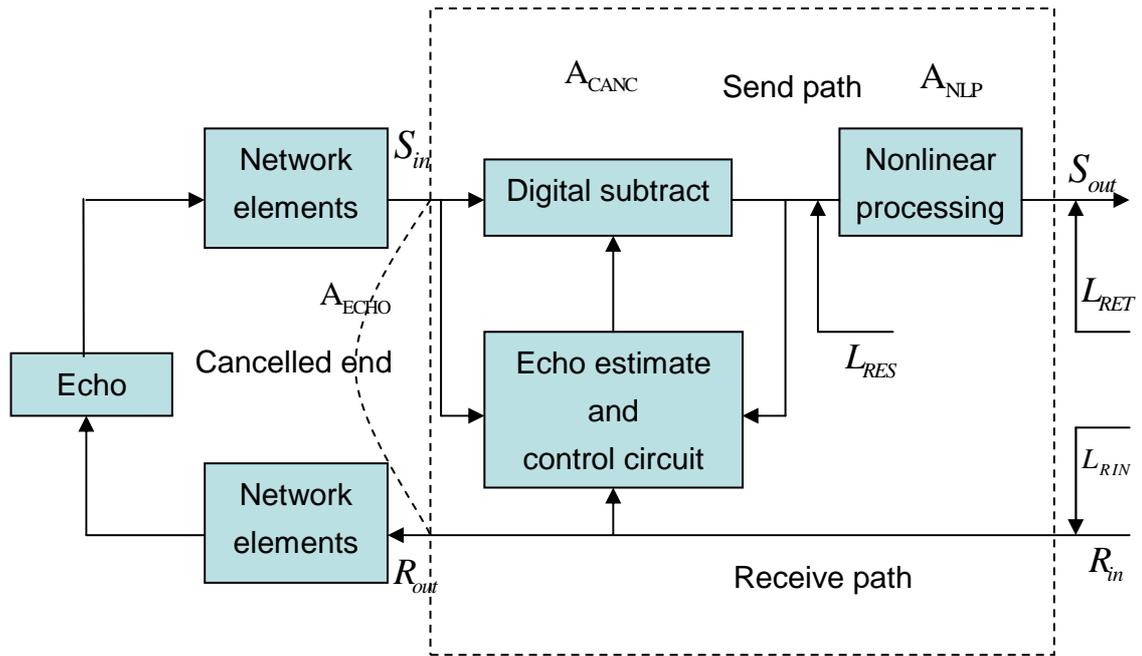
$$\rho_{CL}(n) = \frac{\left| \sum_{i=0}^{N_C-1} \hat{d}(n-i)y(n-i) \right|}{\sum_{i=0}^{N_C-1} \left| \hat{d}(n-i)y(n-i) \right|} \quad (15)$$

我們可以利用上面的公式來偵測是否為雙邊通話的情況，所用到的資訊有麥克風收到的訊號 $y(n)$ 以及自適性濾波器的輸出訊號 $\hat{d}(n)$ 。公式(15)的值介於 0 到 1 之間，當雙邊通話的情況下則此相關性係數的值會偏低，因此我們可以依照情況設定門檻來判斷現在算出來的相關性係數是雙邊通話還是單邊通話的情況。

四、改良式 AEC 的整體模擬與實驗

A. 評估的標準

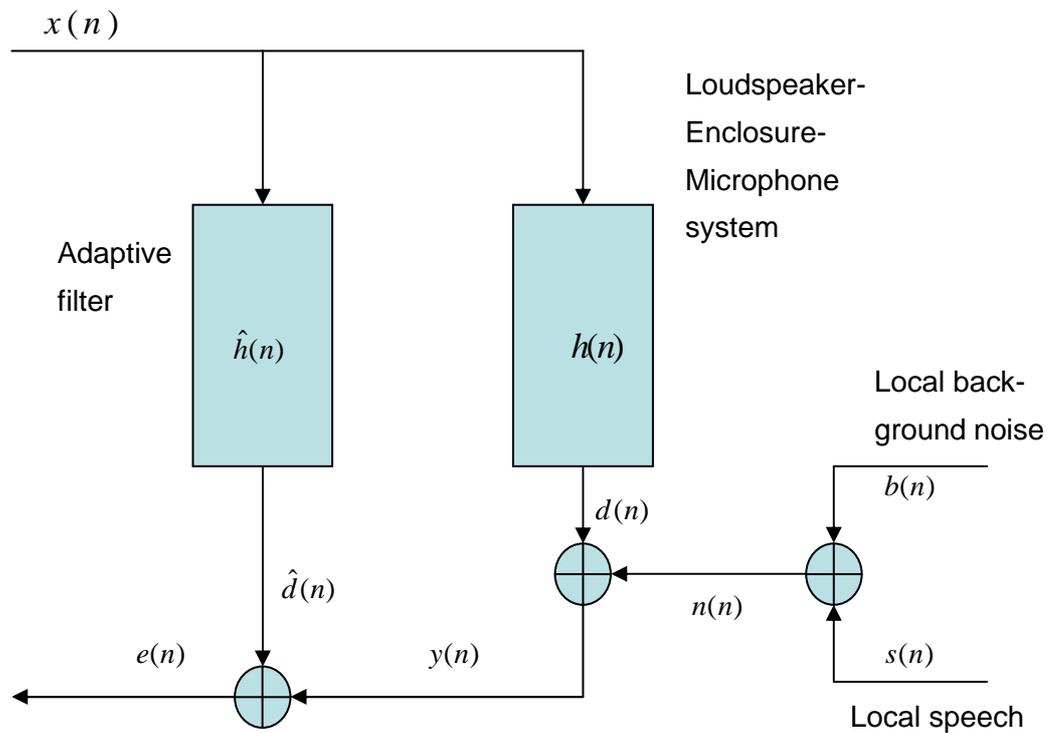
這裡運用 ITU-T G-168 的標準來評估 AEC 的表現好壞，圖五定義一些回音消除的項目。



圖五

ERL (dB)：在回音消除裝置中，從接收端(R_{out})到發送端(S_{in})訊號的衰減。

$ERLE$ (dB)：在回音消除裝置中，在接收發送過程中回音訊號的衰減。此定義不包含輸出非線性(NLP)所造成的訊號衰減。在一個 LEM 系統，回音消除濾波器 (ECF)可由圖六所示，



圖六

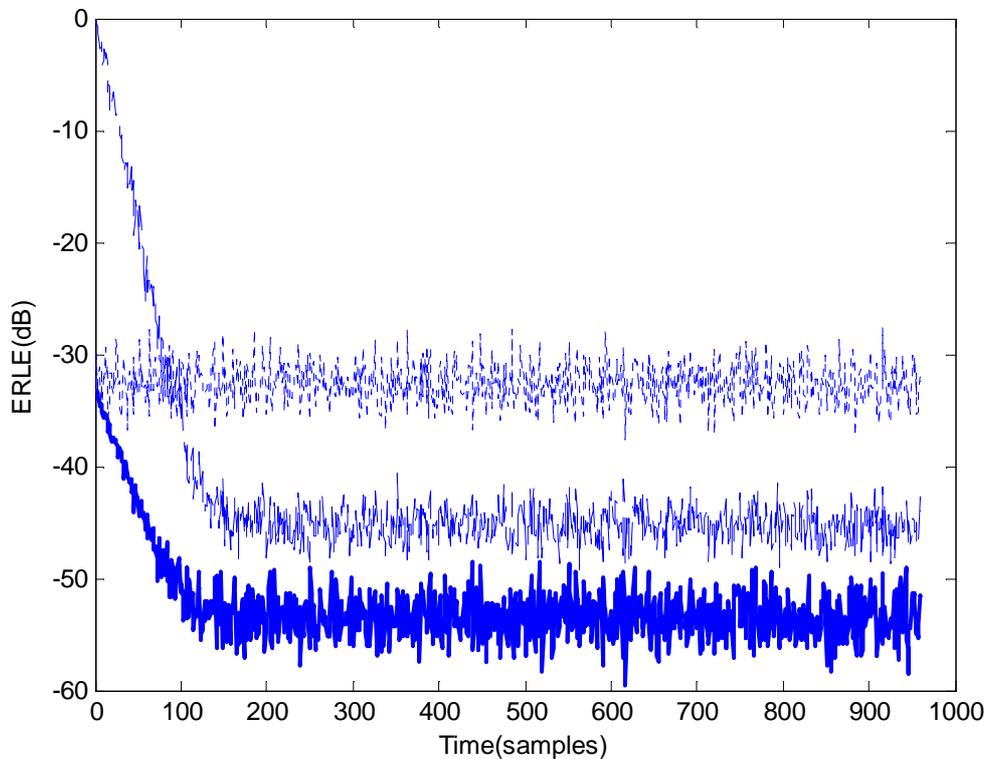
則 ERLE 可由下式所計算出

$$ERLE = 10 \cdot \log_{10} \frac{E[d^2(n)]}{E[(d(n) - \hat{d}(n))^2]} \text{ (dB)} \quad (16)$$

其中 $E[\cdot]$ 為每一百筆資料作平均的期望值。

B. 改良式 AEC 的數值模擬

數值模擬部分針對改良式 AEC 演算法來進行模擬分析。圖七為改良式 AEC 演算法模擬出來的結果，



圖七

在這個模擬中我們的主要目的是測試改良式 AEC 演算法是否更有效，其中我們運用一個 512 階的 FIR 濾波器來表現在房間裡麥克風及喇叭之間的回音路徑 (echo path)，另一個 512 階的 FIR 濾波器用來表式改良式 AEC 中固定的濾波器。對於一個白噪音的輸入，在取樣頻率 48 kHz 時 $ERLE = 53$ dB，接著我們借由濾波器 $\Delta P(z) = 1 + z^{-1} + z^{-2} + z^{-3}$ 來做個微小的回音路徑的改變，此做法是用來模

擬當回音路徑受到外界的擾動時的變化，最後我們評估當輸入白噪音時改良 AEC 演算法的表現是如何。在圖七中，虛線代表的是固定式濾波器 (fixed filter) 的 ERLE，其值大概是 32dB，由於回音路徑的擾動導致 ERLE 值明顯的降低了。另一條較細的實線代表的是運用 512 階全適應性濾波器 (fully adaptive filter) 所得到的 ERLE 值，其值大概在 45dB 時穩定下來，比固定式濾波器的表現要好一點。最後我們運用一個 3 階適應性濾波器來嘗試改良式 AEC 演算法的模擬，圖七中最粗的實線代表的是改良式 AEC 所得到的 ERLE 值，其值大概在 53dB 時穩定下來，此三種方法比較可得到改良式 AEC 的表現是最好的。

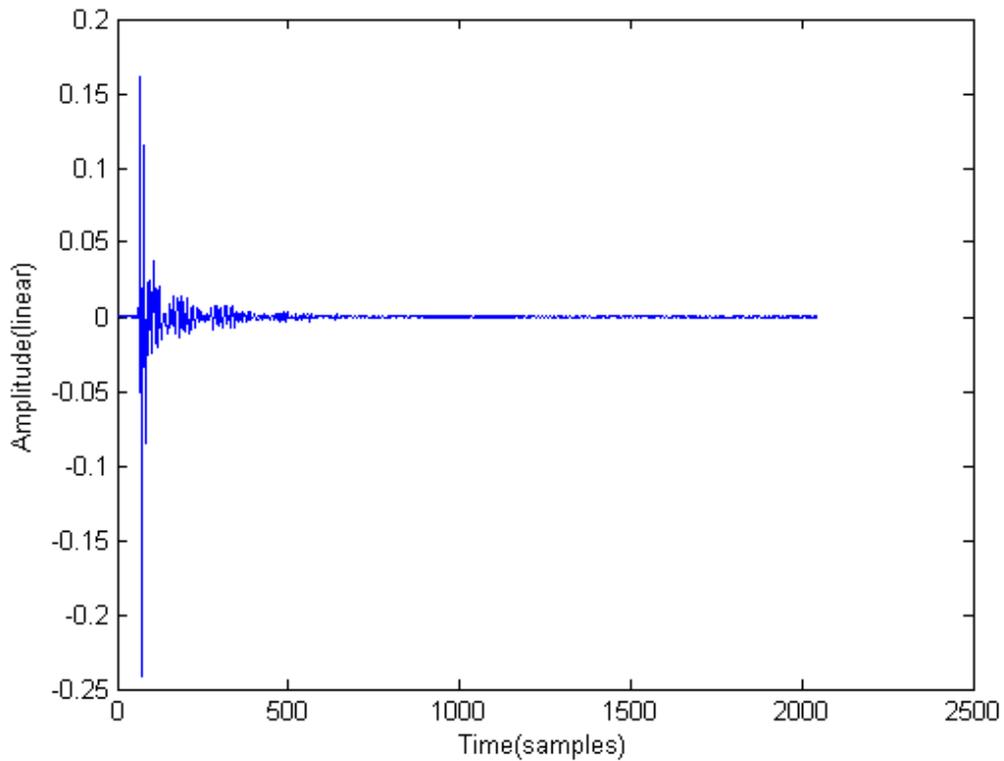
C. 改良式 AEC 的實驗

圖八為實驗的配置圖，



圖八

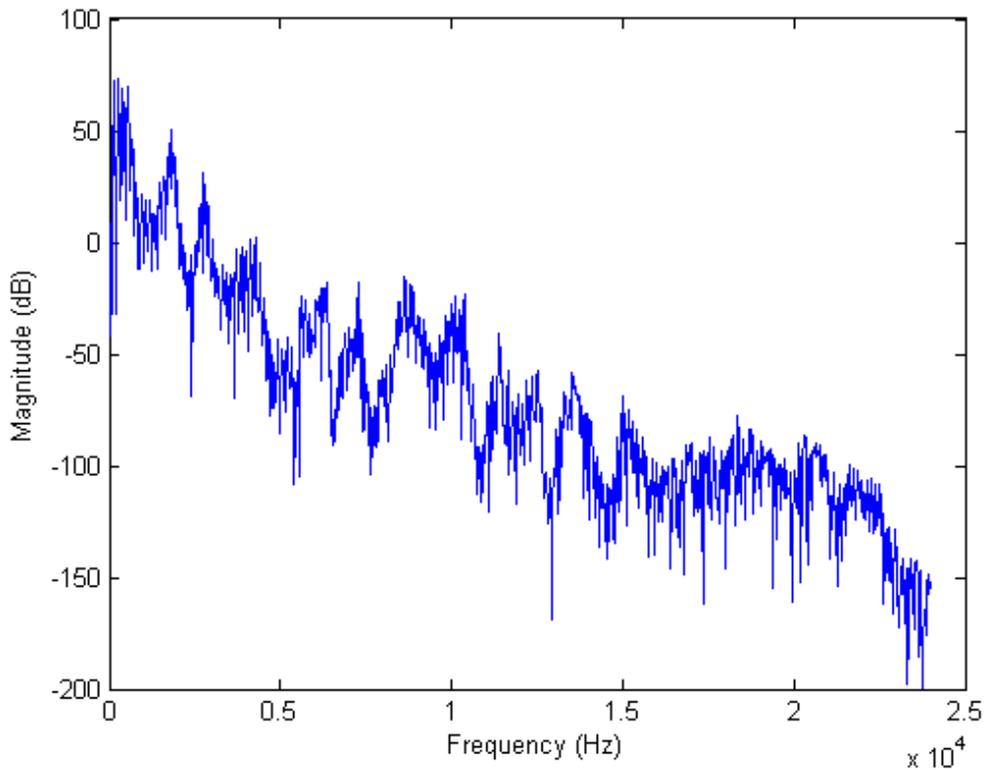
其中多媒體喇叭和麥克風的間隔為 40 公分，而圖九為回音路徑的脈衝響應，



圖九

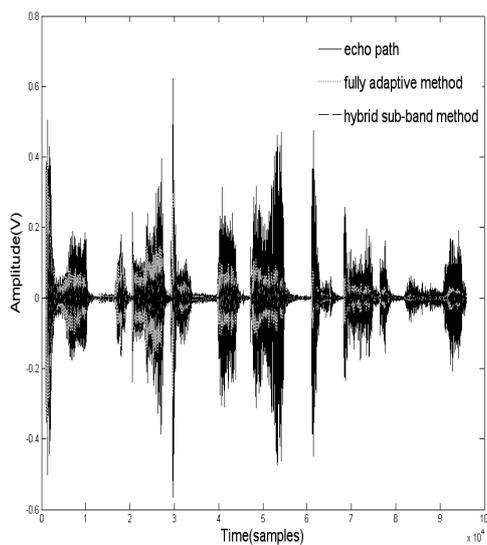
此 LEM 系統有 458ms 的迴響時間，其 $ERL = 6 \text{ dB}$ ，而所有的 AEC 演算法都實現
在一個 fixed-point DSP (ADI, BF533) 的平台上面，其取樣頻率是 48kHz。在此實
驗當中，我們運用點對點的網路通訊電話 Skype 來當作通訊的方式，在未啟動
DSP 的 AEC 演算法前，Skype 是沒有回音消除的效果的。

接著我們比較全適應性濾波的 AEC 演算法跟改良式 AEC 演算法的表現，用來
測試的訊號是一個女性的說話訊號，其頻譜圖如圖十所示，

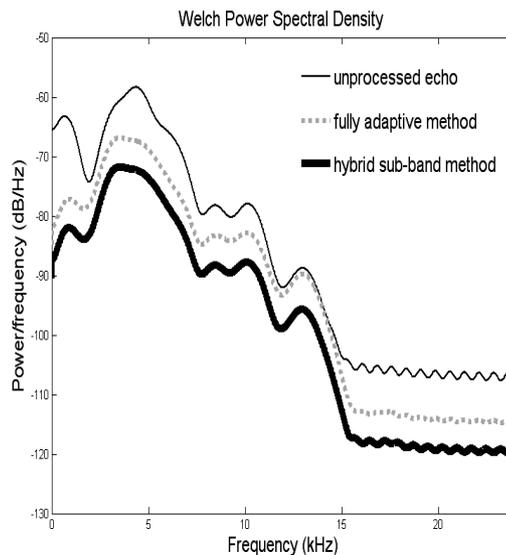


圖十

由於其訊號主要的頻帶分佈在 0-2.5k Hz 之間，所以我們降低頻率(down-sampled)到 8 kHz。接著我們將所有的頻率分成八個頻帶，適應性濾波器應用在第一到第八個分頻中，全適應性濾波器 AEC 演算法中每個階數都是 1024 階且 step size 是 0.1，而改良式 AEC 演算法中其第一到第五的分頻使用階數為 160 階的固定式濾波器，第六到第八的分頻為 75 階的固定式濾波器，使用在改良式 AEC 演算法中的適應性濾波器皆為 16 階且 step size 是 0.01。圖十一(a)是全適應性濾波 AEC(黑色的虛線)與改良式 AEC(灰色的虛線)的比較圖，圖十一(b)為其頻譜圖，由結果來看改良式 AEC(ERLE = 30.5 dB)的表現明顯優於全適應性 AEC (ERLE = 22.4 dB)。



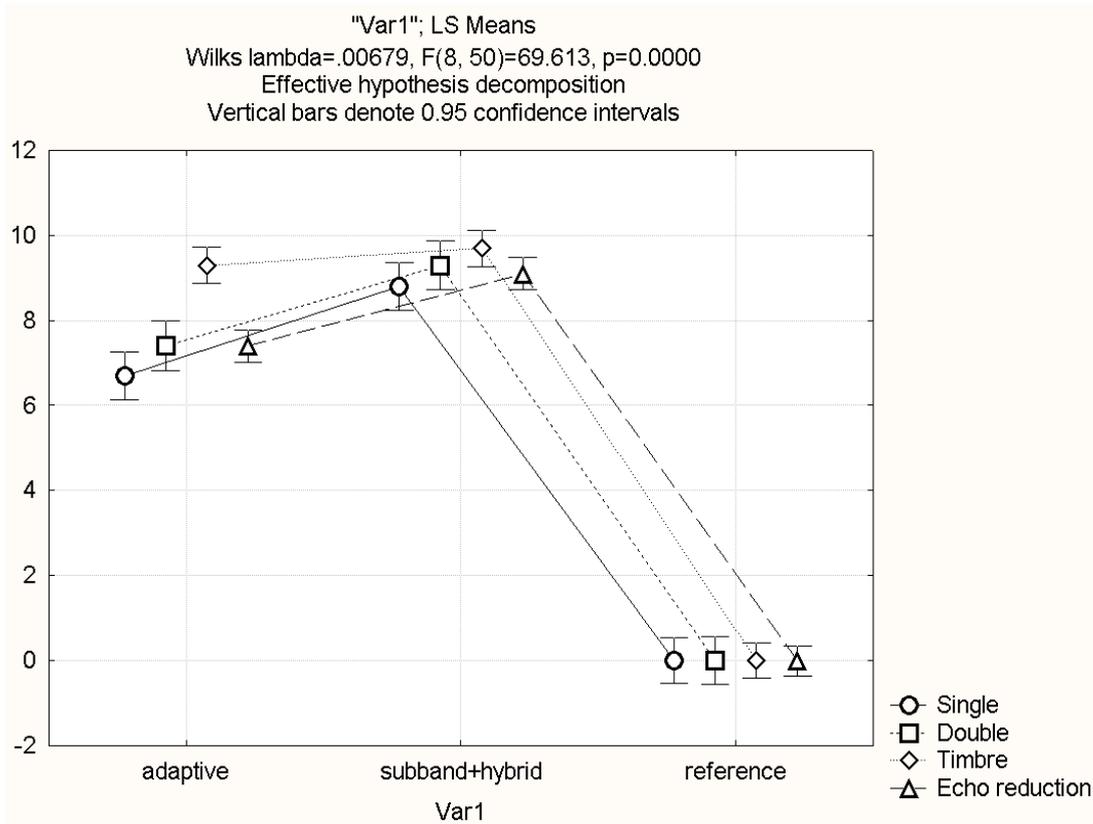
圖十一(a)



圖十一(b)

五、改良式 AEC 的聆聽測試與比較

聆聽測試的設計分為兩種，‘listening-only’測試與‘conversational’測試，一共有 20 位受測者參與本聆聽測試，在‘listening-only’測試當中，受測者將就回音減少的程度和語音失真程度來做評分，其測試訊號是在近端(near-end)已做過處理的語音檔；而在‘conversational’測試當中，評分的標準依然是回音減少的程度和語音失真程度，但測試平台是運用 Skype 與對方通話的及時語音。實驗一樣是將改良式 AEC 與全適應性 AEC 做相比較，其參數選用及測試的平台和之前的實驗一樣。在這個測試裡有 4 個主要的指標如下：1) 單邊通話型式裡的回音消除，2) 雙邊通話型式裡的回音消除，3) 近端語音(timbral quality)的品質，4) 回音消除的整體表現。受測者的評分標準可介於 1 到 10 分之間，圖十二為‘listening-only’的測試結果，其測試結果運用 MANOVA 統計分析軟體來加以統計分析其結果，在這項分析中平均 5%-95% 的受測結果將被分析，而剩下的部分因為差異太大而被刪除。

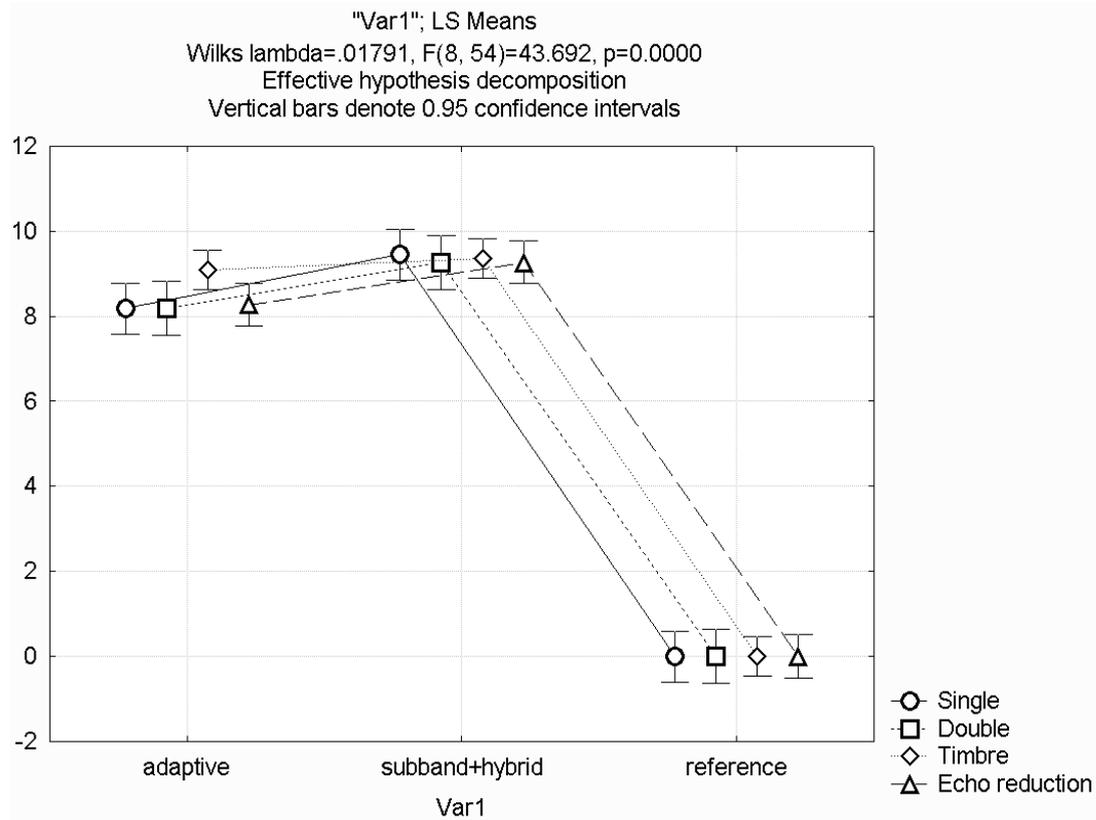


圖十二

Test module	Significance value			
	Echo reduction	Single talk	Double talk	Timbre
Listening-only test	0.000050	0.000454	0.001311	0.291625

TABLE I. The MANOVA results of the listening-only test.

由圖十二及表一可看出在 timbral quality 中改良式 AEC 和全適應性 AEC 沒有顯著的差異，但是從其他的指標來看，改良式 AEC 明顯的優於全適應性 AEC。圖十三為‘conversational’的測試結果，MANOVA 的分析結果如表二，



圖十三

Test module	Significance value			
	Echo reduction	Single talk	Double talk	Timbre
Conversational test	0.030764	0.020945	0.056050	0.501421

TABLE II. The MANOVA results of the conversational test

我們發現在 timbral quality 中改良式 AEC 和全適應性 AEC 仍然沒有顯著的差異，而從其他的指標來看，儘管在雙邊通話型式裡的回音消除的差異較小，但整體來說改良式 AEC 仍然明顯的優於全適應性 AEC。因此我們可以下一個總結，除了 timbral quality 差不多之外，改良式 AEC 在各方面都明顯優於全適應性 AEC。

參考文獻

- ¹ E. Hansler and G. Schmidt, *Acoustic echo and noise control A Practical Approach* (John-Wiley, New York, 2004).
- ² K. J. Astrom and B. Wittenmark, *Adaptive control* (Addison-wesley, New York, 1995).
- ³ S. M. Kuo and D. R. Morgan, *Active noise control system* (John-Wiley, New York, 1996).
- ⁴ M. Brandstein and D. Ward, *Microphone arrays* (Springer, New York, 2001).
- ⁵ Y. Huang and J. Benesty, *Audio signal processing. For next-generation multimedia communication systems* (Kluwer Academic, London, 2004)
- ⁶ C. Breining, P. Dreiscitel, E. Hansler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, J. Tilp, "Acoustic echo control. An application of very-high-order adaptive filters," *IEEE, Signal Processing Magazine*, **16**, 42-69 (1999).
- ⁷ B. Widrow and S. D. Stearns, *Adaptive signal processing* (Prentice-Hall PTR, Englewood Cliffs, NJ, 1985).
- ⁸ B. Farhang-Boroujeny, *Adaptive filters theory and application* (John-Wiley, New York, 2000).
- ⁹ S. Haykin, *Adaptive filter theory* (Prentice-Hall, Englewood Cliffs, NJ, 1986).
- ¹⁰ P. P. Vaidyanathan, *Multirate system and filter bank* (Prentice-Hall PTR, Englewood Cliffs, NJ, 1993).
- ¹¹ H. Yasukawa and S. Shimada, "An acoustic echo canceller using sub-band sampling and decorrelation methods," *IEEE, Trans. Signal Processing*, **4**, 926-930 (1993).
- ¹² Y. P. Lin and P. P. Vaidyanatjan, "A kaiser window approach for the design of prototype filter of cosine modulated filterbanks," *IEEE, Signal Processing Letters*, **5**, 132-134 (1998).

- ¹³ D. R. Morgan, "An analysis of multiple correlation cancellation loops with a filter in the auxiliary path," *IEEE Trans. Acoust. Speech, Signal processing.* **28**, 454-467 (1980).
- ¹⁴ L. J. Eriksson, M. C. Allie, C. D. Bremigan, and J. A. Gilbert, "Weight vector analysis of an RLMS adaptive filter with on-line auxiliary path modeling," *Proc. ICASSP 89, IEEE, Glasgow, UK, 2029-2032 (May, 1989).*
- ¹⁵ S. L. Gay and J. Benesty, *Acoustic signal processing for telecommunication* (Kluwer Academic, London, 2000).
- ¹⁶ T. Gansler, M. Hansson, C. J. Ivarsson, and G. Salomonsson, "A double talk detector based on coherence," *IEEE Trans. Commun.* **44**(11), 1421-1427, (1996).
- ¹⁷ P. Heitkamper, "An adaptation control for acoustic echo cancellers," *IEEE Signal Process.* **4**, 170-172 (1997).
- ¹⁸ H. Ye, and B. X. Wu, "A new double talk detection algorithm based on the orthogonality theorem," *IEEE Trans. Commun.* **39**(11), 1542-1545 (1991).
- ¹⁹ R. Frenzel and M. E. Heonecke, "Using prewhitening and step size control to improve the performance of the LMS algorithm for acoustic echo compensation," *IEEE International Symposium on Circuits and Systems*, **4**, 1930-1932, (San Diego, CA, 1992).
- ²⁰ S. Yamamoto, S. Kitayama, J. Tamura, and H. Ishigami, "An adaptive echo canceller with linear predictor," *Trans. IECE of Japan*, **62**, 851-857 (1979).
- ²¹ G. Keppel and S. Zedek, *Data analysis for research designs.* (W. H. Freeman and Company, New York, 1989).
- ²² ITU-T Rec. G.168, *Transmission systems and media, digital systems and networks* (International Telecommunications Union, Geneva, Switzerland, 2004).

計畫成果自評

本計畫提出的改良式分頻 AEC 能夠有效的消除在通訊時所產生的聲學回音，其中固定式濾波器用來消除大部分的回音，適應性濾波器則用來消除經過外界干擾而產生回音路徑的變化所造成的回音，運用分頻法可以更有效率的減少運算量且增加消除回音的表現，在實驗中比較改良式 AEC 與全適應性濾波 AEC 的優劣，我們可得到改良式 AEC 在計算量、收斂速度和消除回音的效果都明顯比全適應性濾波 AEC 來的要更好。

經過了一年的研究，改良式 AEC 演算法已經被詳細的模擬及實現了。本年的研究成果對於之後的研究有很大的幫助，接下來的兩年我們希望能分別就噪音消除以及麥克風陣列技術來進行研究與實驗，希望最後能整合麥克風陣列、聲學回音消除及噪音消除三大技術並實現在車內免持聽筒通訊系統當中。