

改進的向量空間可適性濾波器用於聲學回聲消除

Acoustic Echo Cancellation Using an Improved Vector-Space-Based Adaptive Filtering Algorithm

李尤進*、曹昱⁺、錢膺仁[#]

Jin Li-You, Yu Tsao and Ying-Ren Chien

摘要

在回聲消除系統的應用中，濾波器係數是否能有效的快速更新是相當重要的關鍵，而收斂的效果也是影響回聲是否能消除乾淨的重要因素。因此向量空間可適性濾波器被提出，其結合機械學習向量空間的想法，引進可適性演算法中，達到有效的快速收斂的目標，但其運算複雜度也相對提高。然而為對應到現實生活中應用，運算複雜度將會是較為重要的考量。因此本篇提出改進的向量空間可適性濾波器(Improved Vector-space Adaptive Filter)與改進的向量空間仿射投影符號演算法(Improved Vector-space Affine Projection Sign Algorithm)，藉由重新設計向量空間以及濾波器係數合成的架構，將運算的矩陣維度降低，並運用組合演算法的想法，對仿射投影符號演算法與改進的向量空間仿射投影符號演算法進行組合，達到在任何環境下皆能快速且穩定收斂的目標且比起向量空間可適性濾波器有著更低的運算複雜度和更好的收斂速度與收斂效果，提升在現實生活中應用地可行性。

關鍵詞：回聲消除系統、可適性濾波器、向量空間可適性濾波器、機器學習、組合演算法、仿射投影符號演算法

* 國立臺灣大學電信工程研究所

Graduate Institute of Communication Engineering, National Taiwan University

E-mail: d05942004@ntu.edu.tw

⁺ 中央研究院資訊科技創新研究中心

Research Center for Information Technology Innovation, Academia Sinica

E-mail: yu.tsao@citi.sinica.edu.tw

[#] 國立宜蘭大學電機工程學系

Department of Electrical Engineering, National Ilan University

E-mail: yrchien@niu.edu.tw

Abstract

To eliminate acoustic echo, the convergence rate and low residual echo are very important to adaptive echo cancelers. Meanwhile, an affordable computational complexity has to be considered as well. In this paper, we proposed the improved vector space adaptive filter (IVAF) and Improved Vector-space Affine Projection Sign Algorithm (IVAPSA). The proposed can be divided into two phases: offline and online. In the offline phase, IVAF constructs a vector space to incorporate the prior knowledge of adaptive filter coefficients from a wide range of different channel characteristics. Then, in the online phase, the IVAF combines the conventional APSA and IVAPSA algorithms, where IVAPSA computes the filter coefficients based on the vector space obtained in the offline phase. By leveraging the constructed vector space, the proposed IVAF is able to fast converge and achieve a better echo return loss enhancement performance. Moreover, the computational complexity is less than a comparable work.

Keywords: Acoustic echo cancellation, Adaptive Filter, Vector-space Adaptive Filter, Machine Learning, Combined Algorithm, Affine Projection Sign Algorithm.

1. 緒論 (Introduction)

在免持通訊系統或是進行遠端會議時，通常會運用麥克風作為語音的接收，再利用揚聲器進行撥放。然而在這樣的環境下，揚聲器的聲音容易被麥克風所接收造成回聲的產生，進而嚴重的影響到通訊系統的使用品質與語音訊號的可理解性。針對上述的問題，運用可適性濾波器(Adaptive Filter, AF)的聲學回聲消除系統(Acoustic Echo Cancellation, AEC)被提了出來，並廣泛的應用在實際的通信系統中(Ahgren, 2005) (Faller & Tournery, 2006) (Hansler & Schmidt, 2006) (Wada & Juang, 2012) (Haykin, 2003) (Wada & Juang, 2009) (van Waterschoot & Moonen, 2011)。

通常，AF 的方法是利用濾波器係數來近似回聲路徑的脈衝響應(又稱未知房間系統響應)，並且一步步的從麥克風的訊號預估並去除回聲。圖 1 便是以 AF 為基礎的 AEC 系統架構，其中 $x(n)$ 、 $y(n)$ 、 $d(n)$ 和 $e(n)$ 分別是 Far-end 輸入訊號(本地端)、回聲訊號、Near-end 輸入訊號(對方端)以及估計誤差。而 \mathbf{h} 與 $\hat{\mathbf{h}}$ 分別是未知房間系統響應和濾波器係數。而 AEC 系統通過可適性濾波器產生出的估計回聲訊號 $\hat{y}(n)$ ，將 Near-end 輸入訊號與之相減去除訊號中的回聲部分 $y(n)$ ，達到消除回聲的目的。AF 的演算法有很多，最小均方演算法(Least Mean Square, LMS)與正規化最小均方演算法(Normalized Least Mean Square, NLMS)便是其中最為基本的，有著較低的運算複雜度，因此被廣泛的應用著(Widrow & Stearns, 1985) (Chien & Chu, 2014) (Soria *et al.*, 2004) (Tandon, Ahmad & Swamy, 2004) (Feuer & Weinstein, 1985) (Chien & Zeng, 2013) (Huang & Lee, 2012)。然而，上述兩個方法卻有著致命的缺點，對於彩色輸入訊號的處理能力不好，無法有好的濾波器係數估計。因此仿射投影演算法(Affine Projection Algorithm, APA) (Liao & Khong, 2010)

(Shin, Sayed & Song, 2004) (Hwang & Song, 2007) (Gil-Cacho, van Waterschoot, Moonen & Jensen, 2012)發展了出來，改善了上述問題，強化了可適性演算法對彩色輸入訊號的處理能力。但是在改善的同時，其運算複雜度上升，且對於雜訊的影響過於敏感，以至於現今較少在現實中被運用。仿射投影符號演算法(Affine Projection Sign Algorithm, APSA) (Shao, Zheng & Benesty, 2010) (Yoo, Shin & Park, 2014) (Shin, Yoo & Park, 2012)在其後被發表，比起 APA 有著更低的運算複雜度、較好的抗雜訊影響以及保持好的對彩色輸入訊號處理效果，現今也在學術討論中被廣泛的討論著。

在 AF 的應用中，演算法的收斂速度與收斂性能是相當重要的關鍵。在傳統的方法中，大多靠著對步階值的調整來影響收斂狀態，但使用這種方法無法同時顧及收斂性能與收斂速度。因此於 2015 年，向量空間可適性濾波器(Vector-space Adaptive Filter, VAF) (Tsao, Fang & Shiao, 2015)被提出，其以創新的想法將機械學習的概念加入 AF 當中，使得 AF 能擁有學習應用的能力。在離線階段可藉由 AF 的事前訓練得到先驗知識群 \mathbf{H} ，而這些先驗知識群便是一組組在不同環境設定下收斂好的濾波器係數；接著在在線階段中，更新權重向量的係數對先驗知識群的向量空間做線性組合，得到濾波器輸出。由於在先驗知識群中已有著收斂完成的各種環境係數，因此 VAF 可以比起傳統的 AF 有更快的收斂速度與收斂性能，由上述可知，先驗知識群的設計對於 VAF 來說相當重要。

然而，VAF 在線階段中的向量空間設計會造成運算複雜度大幅度的提高，無法有效的拿來作為為之，所以本篇提出了改進的向量空間可適性濾波器(Improved Vector-space Adaptive Filter, IVAF)，重新設計在線向量空間與濾波器輸出合成的結構，並搭配 APSA 提出改進的向量空間仿射投影符號演算法(Improved Vector-space Affine Projection Sign Algorithm, IVAPSA)，改善 VAF 運算複雜度過高之問題，並保持其優異之性能，詳細內容由章節四說明。

2. 演算法與系統架構 (System model and adaptive algorithm)

在這個章節將會介紹使用 AF 處理 AEC 問題的基本系統架構,以及目前較為新穎的仿射投影符號演算法,也是本篇提出方法的基石。

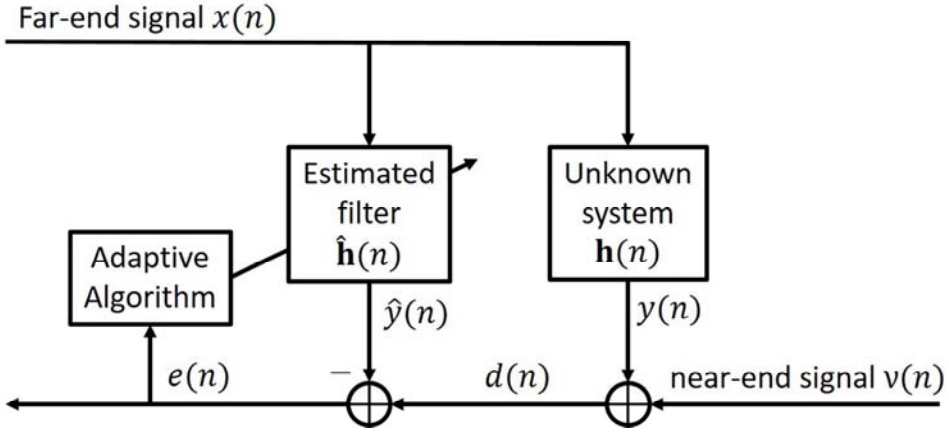


圖 1. AEC 系統架構示意圖

[Figure 1. Structure of acoustic echo cancellation using adaptive filter]

2.1 系統架構 (System model)

一個標準的 AEC 系統架構如圖 1 所示，我們定義未知房間系統響應為 $\mathbf{h} = [h_0 h_1 \cdots h_{L-1}]^T$ ，而可適性濾波器的濾波係數則是 $\hat{\mathbf{h}}(n) = [\hat{h}_0(n) \hat{h}_1(n) \cdots \hat{h}_{L-1}(n)]^T$ ，也稱之為估計系統響應，其中 L 是濾波器的長度， n 為時間指標。

我們令 Near-end 麥克風收到的信號 $d(n)$ 為：

$$d(n) = \mathbf{x}^T(n)\mathbf{h} + v(n) \quad (1)$$

其中 $v(n)$ 是由雜訊訊號與 Near-end 語音所組成。 $\mathbf{x}(n) = [x_1 x_2 \cdots x_{n-L+1}]^T$ Far-end 輸入的語音訊號，經過房間系統響應 \mathbf{h} 變成回聲訊號 $y(n)$

$$y(n) = \mathbf{x}^T(n)\mathbf{h} \quad (2)$$

而可適性濾波器的輸出信號亦是估計的回聲訊號 $\hat{y}(n)$ 為：

$$\hat{y}(n) = \mathbf{x}^T(n)\hat{\mathbf{h}}(n-1) \quad (3)$$

則 $e(n)$ 是估計的回聲訊號與 Near-end 麥克風收到的訊號 $d(n)$ 相減完所剩下的估計誤差

$$e(n) = d(n) - \hat{y}(n) \quad (4)$$

2.2 仿射投影符號演算法 (APSA)

在仿射投影符號演算法 (APSA) 中，輸入訊號 $\mathbf{X}(n)$ 是一個 $L \times p$ 的矩陣，而輸入訊號 $\mathbf{d}(n)$ 為一個 p 維的向量

$$\begin{aligned} \mathbf{X}(n) &= [\mathbf{x}(n)\mathbf{x}(n-1) \cdots \mathbf{x}(n-p+1)]^T \\ \mathbf{d}(n) &= [d(n)d(n-1) \cdots d(n-p+1)]^T \end{aligned} \quad (5)$$

其中 p 是仿射投影階數，而輸出訊號亦是一個 p 維的向量

$$\mathbf{y}(n) = [y(n)y(n-1) \cdots y(n-p+1)]^T \quad (6)$$

得到的估計誤差向量為

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1) \quad (7)$$

上式中 $\mathbf{e}(n) = [e(n)e(n-1) \cdots e(n-p+1)]^T$ 。而後驗估計誤差向量為

$$\mathbf{e}_o(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n) \quad (8)$$

由(Shao *et al.*, 2010)可以得到 APSA 的估計響應 $\hat{\mathbf{h}}(n)$ 疊代更新公式為

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu \frac{\mathbf{X}(n)\text{sgn}(e(n))}{\|\mathbf{X}(n)\text{sgn}(e(n))\|_2 + \delta} \quad (9)$$

而 $0 < \delta \leq 1$ 為一個很小的正歸化常數，以避免分母為零， μ 是更新步階值。

APSA 演算法藉由 L1-norm 的架構推導出了更新公式，解決了以往仿射投影系列的演算法運算複雜度過大之問題，在步階值 μ 較小時，可以不怕脈衝式雜訊的影響，有著相當優秀的抗雜訊能力。但是也因為 L1-norm 的架構，演算法之收斂速度大幅下降的缺點。

3. 向量空間可適性濾波器(VAF) (Tsao *et al.*, 2015)

VAF 其系統架構如圖 2 所示，比較與圖 1 的差異，可以看到 VAF 的系統架構多出了向量空間 \mathbf{H} 的結構，幫助更新可適性濾波器的係數。VAF 可以分為兩個階段，離線與在線。

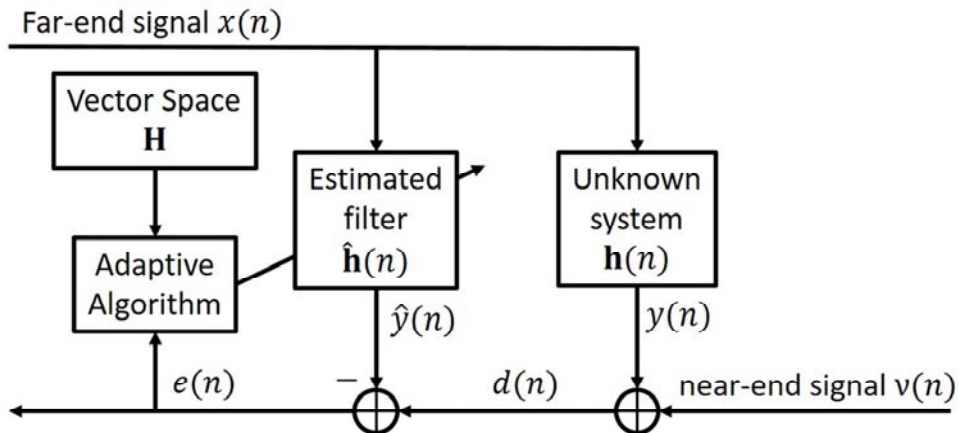


圖 2. VAF-AEC 系統架構示意圖

[Figure 2. Structure of acoustic echo cancellation using vector space based adaptive filter]

3.1 離線階段 (Offline phase)

對於離線階段的目標，是建構一個有效的擁有廣泛通道特性的濾波器輸出係數先驗知識群。該先驗知識群是由 K 組先驗知識所組成，這 K 組先驗知識便是以求得的房間響應向量，其長度為 L

$$\mathbf{H} = [\mathbf{h}^1 \mathbf{h}^2 \cdots \mathbf{h}^K] = \begin{bmatrix} h_0^1 & \cdots & h_0^K \\ \vdots & \ddots & \vdots \\ h_{L-1}^1 & \cdots & h_{L-1}^K \end{bmatrix} \quad (10)$$

其中 \mathbf{h}^K 是先驗知識，而 \mathbf{H} 便是先驗知識群。

3.2 在線階段 (Online phase)

先驗知識群矩陣 \mathbf{H} 可以看做是由不同房間響應特性所組成的向量空間，在其後方加上單位對角線矩陣 $\mathbf{I}_{L \times L}$ 重新組合成新的向量空間 \mathbf{S}

$$\mathbf{S} = [\mathbf{H} | \mathbf{I}_{L \times L}] \quad (11)$$

單位對角線矩陣的用意在於，假使 \mathbf{h} 無法靠先驗知識群合成，也能倚靠後面的單位對角線矩陣使得演算法保持持續收斂。VAF 中令 $\hat{\mathbf{h}}(n)$ 的合成為

$$\hat{\mathbf{h}}(n) = \mathbf{S}\mathbf{w}(n) \quad (12)$$

其中 $\mathbf{w}(n)$ 是 $K+L$ 維的權重向量。而 VAF 便是運用權重向量的調整，由先驗知識群合成出最恰當的濾波器輸出，而這種結合先驗知識以促進在線估計的相同技術已被證實在許多的問題中視為有效的 (Kuhn, Junqua, Nguyen & Niedzielski, 2000) (Tsao & Lee, 2009) (Belhumeur, Hespahn & Kriegman, 1997)。

但在這樣設計的向量空間下會產生矩陣過大的問題，進而造成運算複雜度大幅提高；也因為單位對角線矩陣的關係，會影響到收斂。

4. 改進的向量空間可適性濾波器 (IVAF)

由於 VAF 向量空間設計的問題，因此 IVAF 在在線階段的向量空間設計上以及濾波器係數的合成上進行了改進。

IVAF 以先驗知識群直接當作合成的向量空間，並將濾波器係數的合成改為由兩個演算法做組合，一個是傳統的 APSA，另一個則是 IVAPSA，並藉由組合係數 $\lambda(n)$ 進行組合。

$$\hat{\mathbf{h}}(n) = \lambda(n)\mathbf{H}\hat{\mathbf{a}}(n) + (1 - \lambda(n))\hat{\mathbf{b}}(n) \quad (13)$$

其中 $\hat{\mathbf{a}}(n)$ 是只有 K 維的權重向量， $\hat{\mathbf{b}}(n)$ 則是 APSA 的濾波器係數。

這樣的設計會大幅降低矩陣的大小，以解決運算複雜度過大之問題，並藉由組合 APSA 保持單位對角線矩陣的用意。而組合係數的設計則是參考 (Li-You, 2016)

$$\lambda(n) = \mathbb{E} \left[\frac{e(n)(e(n) - \rho(n))}{(e(n) - \rho(n))^2} \right] \approx \frac{\sigma_e^2(n) - R_e(n)}{\sigma_e^2(n) - 2R_e(n) + \sigma_b^2(n)} \quad (14)$$

其中 $\sigma_e^2(n)$ 、 $\sigma_\rho^2(n)$ 為平均的瞬時誤差功率， $R_e(n)$ 則是誤差訊號的交相關值

$$R_e(n) = \mathbb{E}[e(n)\rho(n)] \quad (15)$$

藉由此組合係數的組合，便能得到最適當的組合比例。當先驗知識群能提供到收斂的幫助時 $\lambda(n)$ 會趨近於 1，使得 IVAPSA 的輸出佔較大部分；反之則是使 APSA 的含量較多。然而由於 APSA 與 IVAPSA 之收斂是獨立的，因此加上一個係數繼承的條件(16)式，使得 IVAPSA 的收斂效果可以延伸到 APSA 上，保持持續收斂的效果。

$$\hat{\mathbf{b}}(n) = \alpha \hat{\mathbf{b}}(n) + (1 - \alpha) \hat{\mathbf{a}}(n), \text{ if } \lambda(n) > \beta \quad (16)$$

其中 α 為趨近於 1 的一個值， β 則是繼承門檻值。

4.1 改進的向量空間仿射投影符號演算法 (IVAPSA)

為得到 IVAPSA 權重向量 $\hat{\mathbf{a}}(n)$ 的疊代更新公式，套用 APSA 的推導算法。新的後驗估計誤差為

$$\boldsymbol{\rho}_o(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\mathbf{H}\hat{\mathbf{a}}(n) \quad (17)$$

藉由此後驗估計誤差便能得到新的成本函數

$$\mathcal{L}(\hat{\mathbf{a}}(n), \Lambda) = \|\boldsymbol{\rho}_o(n)\|_1 + \Lambda[\|\mathbf{H}\hat{\mathbf{a}}(n) - \mathbf{H}\hat{\mathbf{a}}(n-1)\|_2^2 + \epsilon^2] \quad (18)$$

其中 Λ 為拉格朗乘數，而 $0 < \epsilon \leq 1$ 。接著將 $\mathcal{L}(\hat{\mathbf{a}}(n), \Lambda)$ 對 $\hat{\mathbf{a}}(n)$ 偏微分後可以得到

$$2\Lambda\mathbf{H}^T[\mathbf{H}\hat{\mathbf{a}}(n) - \mathbf{H}\hat{\mathbf{a}}(n-1)] = \mathbf{H}^T\mathbf{X}(n)\text{sgn}(\boldsymbol{\rho}(n)) \quad (19)$$

在上式中， $\boldsymbol{\rho}(n)$ 為新的誤差錯誤

$$\boldsymbol{\rho}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\mathbf{H}\hat{\mathbf{a}}(n-1) \quad (20)$$

將(19)式重新整理過後得到

$$2\Lambda = \epsilon\|\mathbf{X}(n)\text{sgn}(\boldsymbol{\rho}(n))\|_2 \quad (21)$$

最後，便能獲得權重向量的更新公式

$$\hat{\mathbf{a}}(n) = \hat{\mathbf{a}}(n-1) + \mu'\mathbf{J}\mathbf{H}^T \frac{\mathbf{X}(n)\text{sgn}(\boldsymbol{\rho}(n))}{\|\mathbf{X}(n)\text{sgn}(\boldsymbol{\rho}(n))\|_2 + \delta} \quad (22)$$

而 $\mathbf{J} = [\mathbf{H}^T\mathbf{H}]^{-1}$ ，其值能在離線階段先求得， μ' 是新的更新步階值。有了權重向量 $\hat{\mathbf{a}}(n)$ 的更新公式後，就能藉由(13)式得到 $\hat{\mathbf{h}}(n)$ 的更新。

在運用先驗知識群作為向量空間合成的設計下，可以使可適性濾波器的估計系統響應達到更快收斂的功效。IVAF 藉由化簡向量空間的設計來克服 VAF 運算複雜度過高之問題，並運用組合係數將兩種演算法結合，解決化簡向量空間帶來的缺陷，達到改進 VAF 方法不適用於實際應用得缺點。

5. 實驗 (Experiment)

在本章中將會比較傳統 APSA 演算法與 VAPSA 演算法的回聲消除效能差異，採用回聲往返耗損增強(Echo Return Loss Enhancement, ERLE) (Rages & Ho, 2002) (Sukhumalwong & Benjangkprasert, 2006)作為評比標準。ERLE 顧名思義是用來比較回聲殘響的大小。因此，在聲學回聲消除的論文較常被運用。ERLE 越大表示著剩餘之回聲訊號越少，反之，越接近 0 代表回聲殘留的情形越嚴重，其數學式如下：

$$\text{ERLE} = 10 \log_{10} \left(\frac{\sum_{i=1}^L d^2(i)}{\sum_{i=1}^L e^2(i)} \right) \quad (23)$$

而評判運算複雜度之方法是利用 Matlab 之內建運算時間指令作為基準，該評判方法是相對比較，因為會受硬體之影響，導致不同電腦花費之時間不同，但以相對比例來說是客觀的。

5.1 實驗設置 (Simulation setup)

對於未知房間系統響應的設計，本篇運用了 RIR 工具(Habets, 2006)進行模擬。在 RIR 中有五項可以設定的參數，分別是麥克風位置、揚聲器位置、反射次數、反射係數(RC)、空間大小(RS(長,寬,高))。由於在大部分的通訊條件下，麥克風位置與揚聲器位置在本章實驗中只對 RC 與 RS 進行調整，其他參數皆是定值，麥克風位置設定(1, 0.4, 0.6)，揚聲器位置(1, 1, 1)、反射次數為 1 次。

Far-end 輸入訊號的部分有兩種，第一種採用 6000 點的彩色高斯訊號作為輸入，其數學模型如下

$$G(z) = \frac{1}{1-0.9z^{-1}} \quad (24)$$

第二種採用一段語音訊號作為輸入，該訊號為一女性英文語音，被輸入在 Aurora-4 資料庫中 (Hirsch & Pearce, 2000) (Macho *et al.*, 2002) (Parihar, Picone, Pearce & Hirsch, 2004)，檔案名"01zc020d.wv1"，其訊號如圖 3 所示。

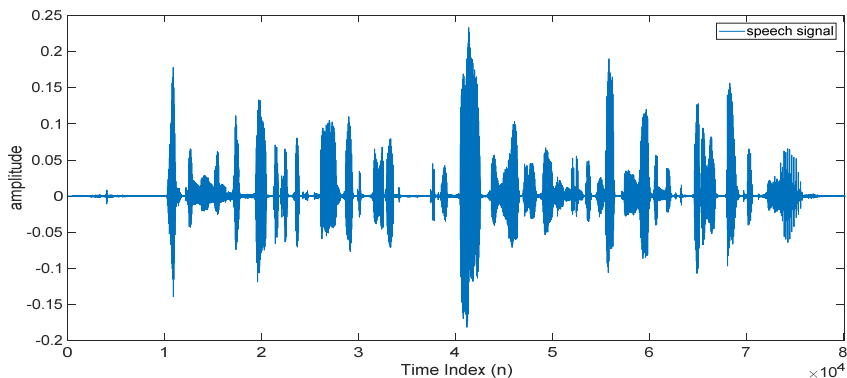


圖3. 輸入語音訊號圖
[Figure 3. The input speech signal]

為準備先驗知識群的訓練，採用了(24)式的彩色高斯訊號作為輸入，並設定背景雜訊為 SNR= 30dB 的白高斯雜訊。

本章中訓練了一組先驗知識群，並且其係數長度 L 為 100，準備了 50 種不同的 RIR 模擬結果訓練，包含了 5 種 RC 係數大小與 10 種 RS 做設定，如表 1；

表 1. 模擬訓練 RIR 50 組參數設定表
[Table 1. Configuration of training data sets]

Data Set	Room Size (RS)	Reflection Coefficient(RC)
Training Set	(1.1, 1.1, 1.1)	-0.91
	(1.2, 1.2, 1.2)	-0.82
	(1.3, 1.3, 1.3)	-0.73
	(1.4, 1.4, 1.4)	-0.64
	(1.5, 1.5, 1.5)	-0.55
	(1.6, 1.6, 1.6)	
	(1.7, 1.7, 1.7)	
	(1.8, 1.8, 1.8)	
	(1.9, 1.9, 1.9)	
	(2.0, 2.0, 2.0)	

表 2. 模擬測試 RIR 3 組參數設定表
[Table 2. Configuration of testing data sets]

Data Set	Room Size (RS)	Reflection Coefficient(RC)
Test Set A	(1.1, 1.1, 1.1)	-0.91
Test Set B	(1.1, 1.1, 1.1)	-0.75
Test Set C	(1.1, 1.2, 1.3)	-0.75

並在表 2 中設定實驗時的參數，TEST SETA 是只 RC 與 RS 都在訓練的設定中，意指未知房間系統響應包含在先驗知識中；TEST SETB 則令 RC 不再訓練的設定中；TEST SETC 最為困難，其設定完全不再訓練之中，代表先驗知識完全沒有見過該未知房間系統響應。而可適性濾波器的其他係數設定如下：

步階值 $\mu = 0.01$ ，仿射投影階數 $p = 16$ ， $\delta = 10^{-6}$ ， $\alpha = 0.99999$ ， $\beta = 0.9$ 。

5.2 實驗結果 (Simulation result)

本章中主要比較 4 種可適性濾波器之效能，分別為傳統 APSA、VAF、以及本篇提出之 IVAF 與 IVAPSA，其中 IVAPSA 為 IVAF 組合係數 $\lambda(n) = 1$ 之結果，亦代表不受 APSA 影響之 IVAF 之效能。以彩色高斯訊號作為輸入，加入 SNR= 30dB 的背景白雜訊。

由圖 4 可以看到 TEST SET A 的實驗結果，在先驗知識群含有的環境設定中，IVAF 之收斂效果比起 VAF 來的更加的快速，IVAF 在接近 2000 點時已收斂，而 VAF 只比 APSA 法好了一點，且 IVAPSA 與 IVAF 的差異也可看出先驗知識群對於收斂效果的幫助非常大，IVAPSA 在 1000 點時就已穩定收斂，但 IVAF 有 APSA 的幫助下收斂的比 IVAPSA 還要再低；在運算時間的花費上，VAF 總共運算了 1.44 秒，而 IVAF 只花費了 0.5 秒，明顯比 VAF 快了接近 3 倍的時間。

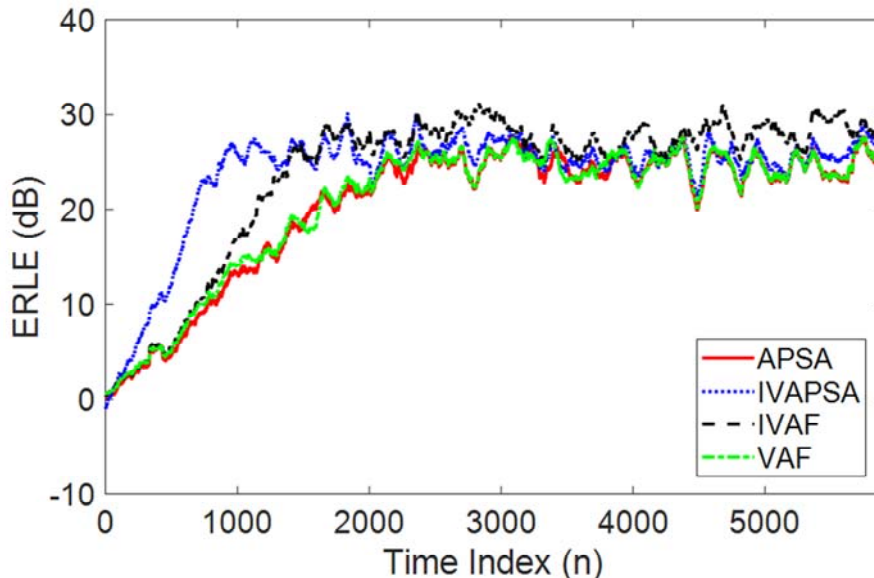


圖 4. TEST SET A 之 ERLE 比較圖

[Figure 4. ERLE of Test set A using Color Gaussian signal be the input.]

而在 TEST SET B 的環境設定下，VAF 花費了 1.48 秒，而 IVAF 則是使用了 0.52 秒的時間，收斂狀況如圖 5，IVAF 亦是能保持著相當快速的收斂狀態，IVAPSA 依舊有著最快收斂的效果，且比起 VAF 的收斂情形來的好很多。由此可以發現 RC 對於先驗知識群的可靠性影響不大，還是能使先驗知識群的向量空間有好的發揮，幫助可適性演算法進行收斂。

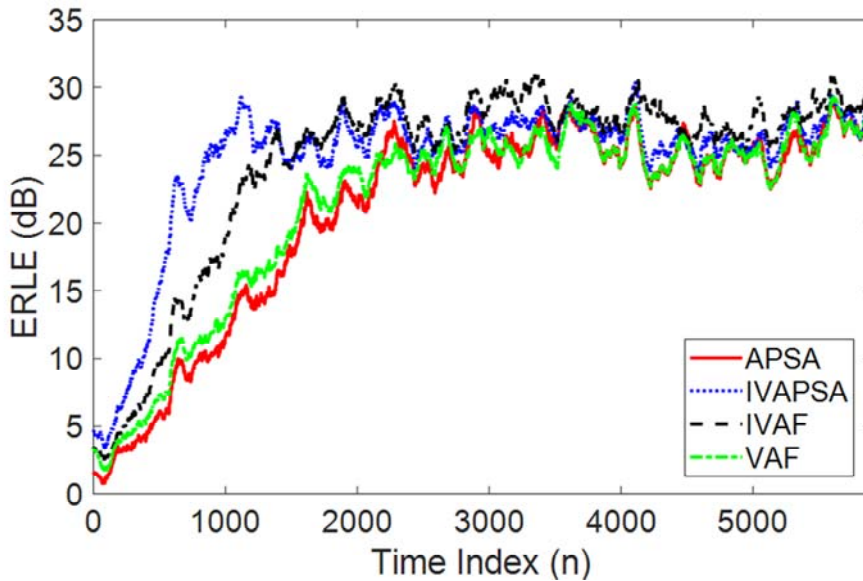


圖 5. TEST SET B 之 ERLE 比較圖

[Table 5. ERLE of Test set B using Color Gaussian signal be the input.]

則在 TEST SAETC 的實驗中，設定的 RC 與 RS 皆不再先驗知識群中，因此先驗知識群就會無法提供收斂幫助。如圖 6 所示，IVAPSA 由於完全是靠先驗知識群進行收斂，所以在這個設定中完全無法進行有效的收斂，甚至濾波器係數發散。反之，VAF 能保持著穩定的收斂，不會收到先驗知識群的影響，而這也是 VAF 加入單位矩陣的原因。再看到 IVAF 亦能持續穩定的收斂，並且收斂的效果大多優於 VAF 與 APSA。而且 VAF 在此花費了 1.51 秒的運算，IVAF 則是使用了 0.53 秒。

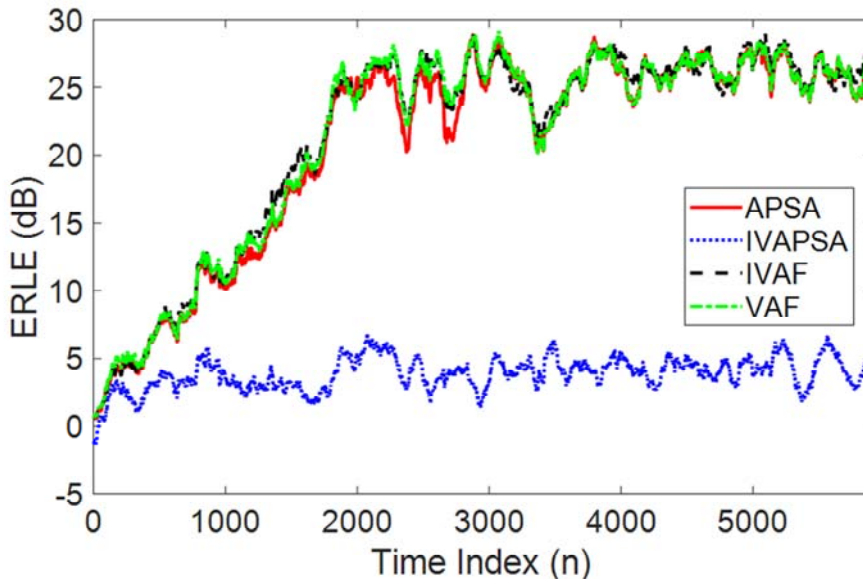


圖6. TEST SET C 之 ERLE 比較圖

[Figure 6. ERLE of Test set C using Color Gaussian signal be the input.]

在上述三個實驗結果，可以看到 IVAF 明顯改善了 VAF 的收斂情形；在先驗知識群含有的環境設定中，有著比 VAF 更加快速並且收斂更好的效能；而在先驗知識群不含的情境下，也能持續穩定的收斂，不會受到先驗知識群的影響而發散；並且 IVAF 皆能比起 VAF 減少了需多的運算時間。

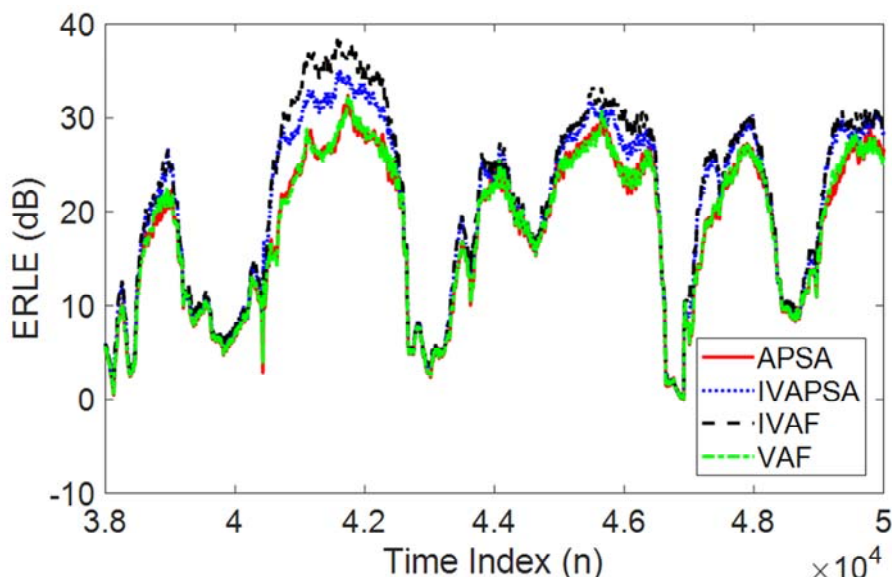


圖7. 語音輸入採用 TEST SET A 環境設定之 ERLE 比較圖

[Figure 7. ERLE of Test set A using speech signal be the input have 80202 points]

由於本篇主要是將其應用在 AEC 的系統中，因此實際測試 IVAF 模擬 AEC 的狀況，在接下來的實驗採用上述之語音訊號作為輸入，並以 TEST SETA 作為環境設定，且為能更好的觀察收斂狀況，只擷取 38000~50000 點之間做觀察。如圖 7 所示，可以觀察到就算是在真實語音資料的輸入下，IVAF 依舊比 VAF 有更好的收斂速度以及收斂效能，尤其在 40000~43000 點間可以明顯看到 IVAF 的優秀表現，與 VAF 相比 ERLE 最大差了將近 8dB；並且在這個實驗中 IVAF 總共花費了 6.8 秒的時間運算，而 VAF 花費了 19.5 秒的運算時間。

由這些實驗結果可以得知本篇提出之 IVAF 能有效的改善 VAF 的收斂問題，提升收斂速度並且保持收斂的穩定，並且比起 VAF 能花費更少量的運算時間，因此能更實用的應用在 AEC 上。

6. 結論 (Conclusion)

隨著科技的進步，免持系統與遠端通訊已成為人類相當熟悉的應用，而其中衍生出的 AEC 問題更是現在急需解決的。而評判 AEC 是否做得好的標準就是觀察 AF 的收斂速度與收斂效能，且為能應用在現實生活中，運算複雜度也不可以太大，因此本篇提出之方法針對上述三個重點進行設計。在上章實驗中可以看到本篇提出之改進的向量空間可適性濾波器，不管是在彩色高斯的輸入下，又或是真實語音輸入的設定下，皆能有效的改善了以往 VAF 收斂受到限制且運算過於複雜之問題，並且在降低運算複雜度的前提下能有效的提高收斂速度與收斂效能，達到上述三個重點的要求，更容易實現並解決 AEC 之問題。

在未來的規劃中，將會進行更多的實驗模擬，例如空間響應變換，更大量的練習資料以及使用真實聲學回聲進行實驗，並且會詳細計算演算法的運算複雜度來更好的證明 IVAF 方法的優點。

參考文獻 References

- Ahgren, P. (2005). Acoustic echo cancellation and doubletalk detection using estimated loudspeaker impulse responses. *IEEE Transactions on Speech and Audio Processing*, 13(6), 1231-1237. doi:10.1109/TSA.2005.851995
- Belhumeur, P. N., Hespanha, J. P. & Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, 19(7), 711-720. doi: 10.1109/34.598228
- Chien, Y.-R. & Chu, S.-I. (2014). A fast converging partial update LMS algorithm with random combining strategy. *Circuits, Systems, and Signal Processing*, 33(6), 1883-1898.
- Chien, Y.-R. & Zeng, W.-J. (2013). Switching-based variable step-size approach for partial update lms algorithms. *Electronics Letters*, 49(17), 1801-1803. doi: 10.1049/el.2013.1762

- Faller, C. & Tournery, C. (2006). Robust acoustic echo control using a simple echo path model. In *Proceedings of 2006 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 281-284. doi:10.1109/ICASSP.2006.1661267
- Feuer, A. & Weinstein, E. (1985). Convergence analysis of lms filters with uncorrelated gaussian data. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(1), 222-230. doi: 10.1109/TASSP.1985.1164493
- Gil-Cacho, J. M., van Waterschoot, T., Moonen, M. & Jensen, S. H. (2012). Nonlinear acoustic echo cancellation based on a parallel-cascade kernel affine projection algorithm. In *Proceedings of 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 33-36. doi: 10.1109/ICASSP.2012.6287810
- Habets, E. A. P. (2006). *Room impulse response generator*. Retrieved from Technische Universiteit Eindhoven, Tech. Rep, 2006.
- Hansler, E. & Schmidt, G. (2006). *Topics in Acoustic Echo and Noise Control*. Berlin, Germany: Springer-Verlag.
- Haykin, S. (2003). *Adaptive Filter Theory*. Upper Saddle River, NJ: Prentice-Hall.
- Hirsch, H.-G. & Pearce, D. (2000). The aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions. In *Proceedings of the ISCA workshop ASR2000*, 181-188.
- Huang, H. C. & Lee, J. (2012). A new variable step-size nlms algorithm and its performance analysis. *IEEE Transactions on Signal Processing*, 60(4), 2055-2060. doi: 10.1109/TSP.2011.2181505
- Hwang, K. Y. & Song, W. J. (2007). An affine projection adaptive filtering algorithm with selective regressors. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 54(1), 43-46. doi: 10.1109/TCSII.2006.883215
- Kuhn, R., Junqua, J.-C., Nguyen, P. & Niedzielski, N. (2000). Rapid speaker adaptation in eigenvoice space. *IEEE Transactions on Speech and Audio Processing*, 8(6), 695-707. doi: 10.1109/89.876308
- Liao, L. & Khong, A. W. (2010). Sparseness-controlled affine projection algorithm for echo cancelation. In *Proceedings of the Second APSIPA Annual Summit and Conference*, 355-361.
- Li-You, J. (2016). *A Study of Convex Combined Adaptive Filtering Algorithms* (Master's thesis). Available from <http://hdl.handle.net/11296/8tqk4w>. [In Chinese]
- Macho, D., Mauuary, L., Noé, B., Cheng, Y. M., Ealey, D., Jouvét, D., ...Saadoun, F. (2002). Evaluation of a noise-robust dsr front-end on aurora databases. In *Proceedings of INTERSPEECH 2002*, 17-20.
- Parihar, N., Picone, J., Pearce, D. & Hirsch, H. G. (2004). Performance analysis of the aurora large vocabulary baseline system. In *Proceedings of 2004 12th European Signal Processing Conference*, 553-556.

- Rages, M. & Ho, K. C. (2002). Limits on echo return loss enhancement on a voice coded speech signal. In *Proceedings of the 2002 45th Midwest Symposium on Circuits and Systems*, 2, 152-155. doi: 10.1109/MWSCAS.2002.1186820
- Shao, T., Zheng, Y. R. & Benesty, J. (2010). An affine projection sign algorithm robust against impulsive interferences. *IEEE Signal Processing Letters*, 17(4), 327-330. doi: 10.1109/LSP.2010.2040203
- Shin, H.-C., Sayed, A. H. & Song, W.-J. (2004). Variable step-size nlms and affine projection algorithms. *IEEE signal processing letters*, 11(2), 132-135. doi: 10.1109/LSP.2003.821722
- Shin, J., Yoo, J. & Park, P. (2012). Variable step-size affine projection sign algorithm. *Electronics Letters*, 48(9), 483-485. doi: 10.1049/el.2012.0751
- Soria, E., Calpe, J., Chambers, J., Martinez, M., Camps, G. & Guerrero, J. D. M. (2004). A novel approach to introducing adaptive filters based on the lms algorithm and its variants. *IEEE transactions on education*, 47(1), 127-133. doi:10.1109/TE.2003.822632
- Sukhumalwong, S. & Benjangkaprasert, C. (2006). Adaptive echo cancellation using variable step-size algorithm lattice filters. In *Proceedings of TENCON 2006 - 2006 IEEE Region 10 Conference*, 1-4. doi: 10.1109/TENCON.2006.343852
- Tandon, A., Ahmad, M. O. & Swamy, M. N. S. (2004). An efficient, low-complexity, normalized lms algorithm for echo cancellation. In *Proceedings of the 2nd Annual IEEE Northeast Workshop on Circuits and Systems, 2004. NEWCAS 2004*, 161-164. doi: 10.1109/NEWCAS.2004.1359047
- Tsao, Y. & Lee, C.-H. (2009). An ensemble speaker and speaking environment modeling approach to robust speech recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(5), 1025-1037. doi: 10.1109/TASL.2009.2016231
- Tsao, Y., Fang, S. H. & Shiao, Y. (2015). Acoustic echo cancellation using a vector-space-based adaptive filtering algorithm. *IEEE Signal Processing Letters*, 22(3), 351-355. doi: 10.1109/LSP.2014.2360099
- van Waterschoot, T. & Moonen, M. (2011). Fifty years of acoustic feedback control: State of the art and future challenges. In *Proceedings of IEEE*, 99(2), 288-327. doi:10.1109/JPROC.2010.2090998
- Wada, T. S. & Juang, B.-H. (2009). Acoustic echo cancellation based on independent component analysis and integrated residual echo enhancement. In *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2009 (WASPAA '09)*, 205-208. doi:10.1109/ASPAA.2009.5346494
- Wada, T. S. & Juang, B.-H. (2012). Enhancement of residual echo for robust acoustic echo cancellation. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(1), 175-189. doi:10.1109/TASL.2011.2159592
- Widrow, B. & Stearns, S. D. (1985). *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.

- Yoo, J., Shin, J. & Park, P. (2014). Variable step-size affine projection sign algorithm. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 61(4), 274-278. doi: 10.1109/TCSII.2014.2305013