行政院國家科學委員會專題研究計畫成果報告

主動式噪音控制系統之聽覺技術研究

Computer Hearing Techniques for Active Noise Cancellation

計畫編號：NSC 89-2213-E-009-127-

執行期限：88 **年** 8 **月** 1 **日至** 89 **年** 7 **月** 31 **日**

主持人：胡竹生教授　　國立交通大學電機與控制工程系

一、中文摘要

在此計畫中我們將提出主動式噪音控制之電腦聽覺系統設計，以消除回授音場效應而準確估測噪音聲場。不同於傳統的設計，電腦聽覺系統不但解決主動式噪音控制的聲場回授問題，提供特定聲源方向之聲場估測，而且還提供具物理直覺意義的消音機制設計原理。我們的目標是利用麥克風陣列來實現聽覺技術，使其能自動偵測追蹤聲源並分離出目標聲源之信號。根據對於管路噪音控制的研究，我們首先提出兩麥克風之一維聲場聽覺系統，並實驗證明其實用性，並根據此一維聲場感測器理論經驗來發展二維及三維聲場感測器。我們在far-field聲場假設及filter-and-sum beamforming空間濾波的結構下，做聲源方向判斷及聲場的空間濾波的最佳設計。另外麥克風陣列的幾何形狀與感測器間不一致的動態特性對聽覺系統的性能分析亦於本計畫中作詳細的分析。

關鍵詞：主動式噪音控制，電腦聽覺，聲場回授，空間濾波

Abstract

In this project, we aim to design a computer audition system, which can perceive the desired sound field and eliminate the deleterious effects of acoustic feedback in the same time. In contrast to traditional design, computer audition system provides ANC a good estimation of desired noise field without running into an annoying acoustic feedback problem. It also benefits us a lot to gain a direct physical insight of the noise cancellation mechanism. We are going to design an array of microphones to intelligently track sound sources and selectively capture the desired sound. According to our study of ANC in ducts, we firstly propose a two-microphone audition system, which shows its potential to estimate sound field of a desired source. The computer audition systems of 2-D and 3-D acoustic field are more complicated compared to 1-D case. We focus on studying source localization and spatial filtering techniques based on far-field assumption and filter-and-sum beamforming structure. Under these conditions we are going to find some optimal designs. The array geometry and imperfect characteristic of microphones are another primary concern which will be considered in this project.

Keywords : Active Noise Control, Computer Audition, Acoustic Feedback, Spatial filtering, Beamforming

二、計劃緣由與目的

1933**年德國**Paul Lueg**提出主動式噪音控制**（Active Noise Control）**的構想。基本原理是利用一個與噪音**180**度反相位的聲波，使兩者產生破壞性干涉，以達到消除噪音的目的。但對於更高維的噪音控制系統，問題將變得複雜而困難。**

如何在吵雜的三度空間中製造一個寧靜區首先要感測噪音聲場的分佈與傳播方向，並預測噪音傳播至消音區的音場分佈，進而在這消音區即時產生一反相噪音場，使兩者線性重疊而相消。注意這裡所

處理的不再僅是聲音信號而是包含空間因素的聲場問題，其中面臨的是一個聲場感測與即時反相聲場重建的問題，另外要注意的是，此系統同時存在麥克風與喇叭，其中喇叭的輸出信號亦會被麥克風感測器所接收，而產生聲場迴授（Acoustic Feedback），因此有穩定性的問題產生。綜合以上分析，我們可以將三維的主動噪音控制系統設計概分二部分：（一）電腦聽覺系統。（二）反噪音聲場產生系統。其中聲場在空間與時間重現技術已有多方面的研究（Kirkeby and Nelson, 1993; Nelson,1994），而聽覺感測系統的設計在主動噪音控制領域卻未受到重視，最主要原因在主動噪音控制的發展與研究中，耳機與管路消音系統扮演非常重要的角色，而這兩種聲場模型分別侷限於點與線上，所以沒有複雜的空間與時間的關係，所以並不需要設計複雜的聽覺系統。但是在二維或三維的噪音控制中聰明的聽覺系統就變得非常的重要了，因為若系統無法準確感測其聲場，又如何能忠實的製造反相聲場，以達到局部空間消音的效果呢？而本計畫即是針對其中的電腦聽覺系統進行研究。

什麼是電腦聽覺系統？根據人類本身的經驗而言，一個聰明的聽覺系統應該包括：聲音信號的清楚接收能力、聲源方位判斷能力、聲音本質（音樂、噪音、說話）與意義的判斷能力。但對於主動噪音控制系統而言，高品質信號接收與聲源定位才是最必要的。忠實接收直接路徑傳播的噪音源信號（減少空間迴響所造成原噪音信號失真），並由噪音源的定位預測到達預定消音區的直接傳播方向，如此才能提供控制器足夠的輸入資訊以便在消音區產生傳播方向吻合的反相聲波。因此我們可以說一個好的聽覺系統必須能夠感測聲場（聲源信號、聲源分佈及傳播方向），這也是二維或三維噪音控制系統成功的先決條件。

隨著積體電路製程技術的蓬勃發展，數位訊號處理器快速的即時運算能力與價格的大眾化，使得電腦聽覺技術得以實現。而其中判斷聲源數量及位置並做選擇性的指向錄音技術，與雷達、聲納、導航與天線等領域有密切的關係。基本的原理是利用多個麥克風對聲場作空間取樣收音，以得到聲場部分點的信號變化，然後利用這些空間與時間的聲學物理特性，對這些麥克風所接收的訊號作處理，以判斷聲音信號的頻率成分與傳播方向，對不同方向聲源作信號分離，或針對特定方向聲源作加強收音並衰減其他方向聲音的接收能力，如此減少目標聲源被空間迴響或其他聲源的干擾而得到清楚的收音，此技術亦稱為空間濾波（Spatial Filtering）技術。最有名的空間濾波技術是Beamforming，見Van Veen and Buckley（1988）。而信號源的定位法則有信號延遲估測定位法(Chan, Hattin and Plant, 1978, Omologo and Svaizer, 1994; Branstein and Silverman, 1997)，Steered beamformer 定位法（Carter, 1977; Wax and Kailath, 1983）及信號空間定位法（Schmidt, 1981）。而這些技術大部分皆是由雷達、聲納及天線理論發展出來的，不見得用於主動噪音控制的聽覺系統使用，因此必須結合以往的理論並發展適合的電腦聽覺系統。

電腦聽覺系統除了在主動噪音控制領域扮演重要的角色外，也可能對其他控制應用開啟新的方向，例如電腦視覺（Computer Vision）已廣泛應用至生產自動化、機器人等的控制，提供了受控體感知環境的資訊。但以人類而言，聽覺亦是感知環境變化的一項重要工具。由於影像信號處理所花的時間較長，往往影響控制頻寬，如果輔助以聽覺，則在某些場合中，或能提高控制之效率。此外，由於聽覺系統所需的硬體成本不高且其資料量遠低於視訊。因此這方面的理論發展，極可能很快的在實際的控制問題上找到應用。

此外，本項研究在其他領域亦有多項重要的應用。隨著影像電話技術發展，以及對面對面交談的真實感要求，免持聽筒的設計越來越普遍，另外對於車內行動電

話而言，為行車安全的顧慮，免手持接聽的要求更是必要。但這種情況下聲源與接收麥克風有相當的距離，其信號接收極易受迴響及其他噪音的干擾，故傳統單麥克風的收音方法往往無法得到令人滿意的清晰聲音。所以在這種情況下需要有聰明聽覺系統，能夠自動追蹤並接收想要的聲音並且具有迴響及干擾補償的功能，使達到忠實收音的要求 (Grenier, 1992; Oh, Viswansthan and Papamichalis, 1992)。在語音辨識方面，對於免手持語音電話撥號、語音控制、語音密碼等語音辨識系統而言，清楚收音更可以增高辨識的成功率（Chen et. al., 1994）。另外在會議系統（Flanagan et. al., 1985; Khalil et. al.,1994）及助聽器（Greenberg and Zurek, 1992）方面亦有應用，可見此種電腦聽覺技術的應用潛力。

綜合以上所述，我們可以略見電腦聽覺對三度空間主動噪音控制實現的重要性，以及在日常生活中廣泛的應用性。本研究群於主動式噪音控制方面的經驗顯示（Hu,1993, 1995, 1996a, 1996b,1997a, 1997b），若缺乏對聲場量測信號的前端處理與適當描述，三度空間主動噪音控制器設計將非常困難，且系統的可靠度與性能將難以掌握，而往後的分析與改善更是困難。因此本計畫之主要目的在提出聲場感測及聲場訊號處理技術，解決聲場回授問題並分離出想要消除的噪音聲場，而實現一個聰明的電腦聽覺系統。這種技術使得主動式噪音控制設計可以分成聲場聽覺與聲場重建兩大系統，如此將使設計與理論分析大為簡化，增加實現的性能與可靠度。

三、結果與討論

Beamforming Design and Classification

Beamformer 設計可分成兩種模式:

1. Data independent :
   類似在時域性中使用 Kaiser，chebyshev 方式來設計濾波器，我們不需理會空間擷取訊號所帶來的資訊，只需在使用前，先設計好自己想要的 FILTER 規格，再計算出適當的係數。
2. Statistically optimum :
   在許多情況下，對使用的系統環境並不會很了解，這時很難先決定 SPATIAL FILTER 的形狀。

假設在一系統中，我們要收取零度角的聲音，所以我們會把 MAINLOBE 設計在零度。但是，在 IMPLEMENT 後，我們才發現空間中存在著其他的雜音，假如我們的濾波器是固定的，所以當他處在較高的 SIDELOBE 時我們就會得到較多得雜音。當然假如我們知道雜音來源的話我們還是可以使用 DATA INDEPENDENT 的方式先設計好最好的 Filter 模式。但是如果雜音是會改變的，且會由很多方向進入，DATA INDEPENDENT 的設計將不適用。此時我們需要從接收的資料中了解空間資料的訊息。之後藉由一些最佳化的方法來了解空間的雜訊訊息，進而把這些雜訊去掉。

EigenFilter Beamformer

EigenFilter 是以data independent形式來設計的。EigenFilter可以:定義預期錄取的方向及雜音滲入的方向，之後使用最佳化推導的結果來把希望錄取方向的角度響應設定成PASSBAND，而在雜音滲入方向的角度響應為STOPBAND的空間濾波器。

EigenFilter Beamformer 的計算可簡化成以下的方程式:

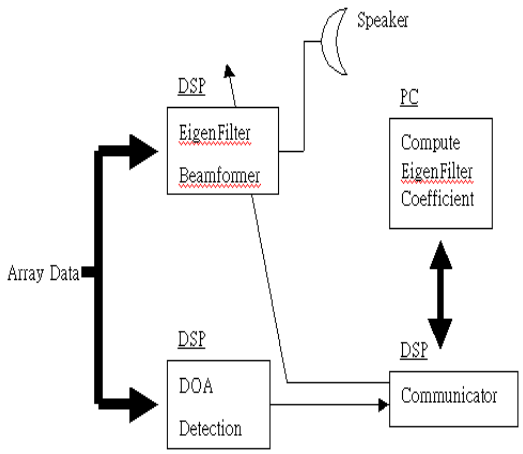
A 和 B 分別是陣列的 MANIFOLD MATRIX。最佳解 h*就可從上式中得到。h* 是 Ah=λBh 最小化 eigenvalue ª的 generalized eigenvector。也因為這個原因，我們才定義這方法為 EigenFilter。

在使用陣列系統來掃描空間時，我們可以使用DOA estimation後的結果來尋找空間中的噪音，當發現噪音的存在，就呼

$$\begin{cases} \mathbf{Ah} - {}^{a}\mathbf{Bh} = 0 \\ \mathbf{h}^{T}\mathbf{Bh} = 1 \end{cases}$$

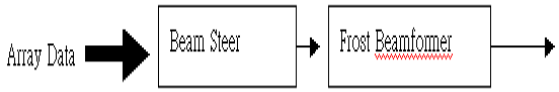
叫系統重新計算EigenFilter，計算完後再把係數載入至DSP。

整個流程如下:

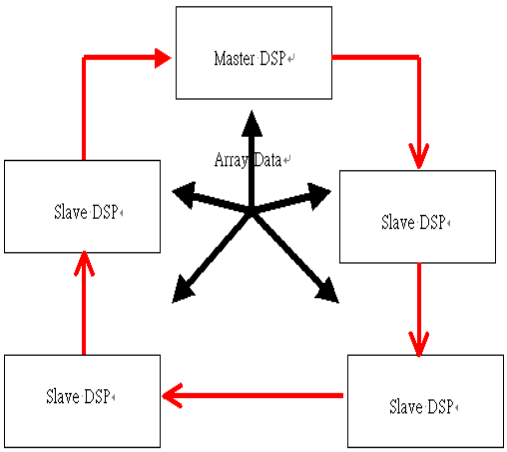


Adaptive Frost Beamformer

以上介紹的EigenFilter Beamformer，在系統只存在兩個聲源，且聲源移動不會太快時，效果會不錯。但是，當環境擁有兩個以上的聲源，在複雜的環境中快速的移動時，這種方法就不適宜了。此時就需使用adaptive signal processing一些適應特性來解決這些問題了。注意在Frost處理器之前，我們需要一組濾波器來作Beam steering的動作。訊號的處理過程如下:



陣列訊號處理系統在一些特殊情況(如精密的空間角度收音，多維空間收音，複雜 algorithm implement)下，需要龐大的數學運算。此時一顆 DSP 微處理器並不足以負擔所有的運算與控制。本計劃設計一並行處理架構如下圖。



並行處理流程說明

1. 上圖顯示的是環狀-串列埠 Multi-DSPs 架構，所有 DSP 處理器，都只能單向傳輸資料。而每顆 DSP 都可彼此間互相溝通。
2. 陣列訊號可接到任何一 DSP 處理器。
3. 需要從眾 DSP 中定義一顆 Master DSP 處理器，負責輸出最後的結果。
4. 需要傳送的 DSP 需要把欲送到的 DSPs，資料模式和資料處理方式送出。
5. 任一 DSP 都可隨時送出資料。
6. 送出的資料可同時傳給許多 DSP(Multi-cast 特性)。

特性

- DSP 與 DSP 之間可任意交換資料，因此可把工作切割得更細。譬如說:我們可以把 SLAVE DSP1 處理器處理完之後的資料再傳給 SLAVE DSP2 理器繼續處理。所以彈性更強。
- DSP 處理器可在任何時間內傳輸資料和接受資料。
- 擴充性:
  軟體 – 所有的 DSP 處理器處理都是獨立作業的，要加入一顆新的 DSP 處理器，並不需要更改其他 DSP 的程式。所以適合快速擴充的系統。
  硬體 – 新增一顆 DSP 處理器在硬體上只需 SERIAL PORT 的 RECEIVER 接到前端的 TRANSMITTER，TRANSMITTER 接到後端的 RECEIVER 即可。
- 傳輸排線為串流排，串流排的傳輸速率為 10M bits/s。

Target Hardware Design

在這個平台裡，我們可由 PC 操控所有的動作。PC 面板上可以操作以下的功能:

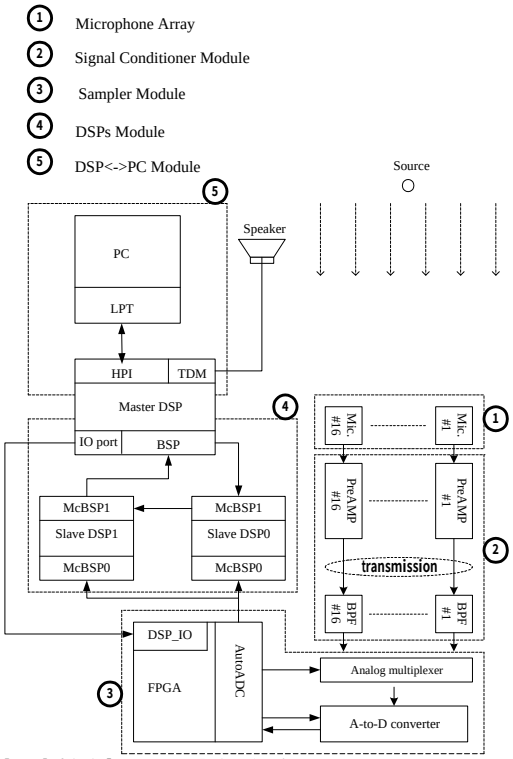
- 檢查系統是否正常
- 起始平台運作
- 停止平台運作
- 呼叫 DSP 系統自動搜尋空間聲源方向，或由面板上選擇要聽取的方向

以下列的是本平台規格:

1. 在平台上我們共使用 16 顆麥克風

2. 陣列的形狀是排列成 ULA 形式，麥克風得間隔是 7.5cm
3. 擷取的語音訊號會 band-limited 在 650~2250Hz 之間
4. Analog-to-Digital Conversion rate 為 9766Hz

為了要達到以上的效果 ,我們的平台必需要有以下的模組。



從實驗的結果，我們得知:

1. 使用 beam steer 配合 Power Estimator 可以達到非常好的聲源追蹤
2. Adaptive Frost 方法的確可以消除雜音

但是它們有著以下的 limitation:

1. Beam steering 的解析角度不夠多，我們不能定義在這九個方向間的角度
2. Frost 消除雜音的效果沒有預期的好，而且在播放時，會有電路雜音的存在。

四、計畫成果自評

項 目	完成情況
與原計畫相符程度	100%
達成預期目標	80%
研究成果學術價值	提供相關理論修正
研究成果應用價值	很有實用性

學術期刊發表合適否	投稿中
申請專利合適否	目前尚不適合
主要發現 或 其他價值	1. 並行 DSP 架構 2. Modified Frost Algorithm 3. Acoustic Beam forming 4. Noise separation

五、參考文獻

Branstein, M. S. and Silverman, H. F., 1997, “ A practical methodology for speech source localization with microphone arrays,” Computer Speech and Language, vol. 11, pp. 91-126.

Chan, Y. T., Hattin, R. V. and Plant, J. B., 1978, “The least squares estimation of time delay and its use in signal detection,” IEEE Trans. on Acoustics, Speech and Signal Processing.vol. 26, pp. 217-222.

Flanagan, J.L. and Jan, E.E., 1997, “ Sound capture with three- dimensional selectivity,”Acustica, vol. 83, pp. 644-652.

Greenberg, J. E. and Zurek, P. M., 1992, “ Evaluation of an adaptive beamforming method for hearing aids,” J. Acoust. Soc. Am., vol. 91, March, pp. 1662-1676.

Grenier, Y., 1992, “A micophone array for car environments,” In Proceedings of ICASSP-92, San Francisco, CA, pp. I-305-I-309.

Kirkeby, O. and Nelson, P.A., 1993,” Reproduction of plane wave sound fields,” J. of the Acoustical Society of America, vol. 94, pp. 2992-3000.

Knapp. C. H. and Carter, G C., 1976, “ The generalized correlation method for estimation of time delay,” IEEE Trans. on Acoustics, Speech and Signal Processing. Vol. 24, pp. 320-327.

Khalil, F., Jullien, J.P. and Gilloire, A., 1994, “ Microphone array for sound pickup in teleconference systems,” J. Audio Eng. Soc., vol. 42, No. 9, pp.691-700.

Oh, S., Viswanathan, V. and Papamichalis, P., 1992,” Hands-free voice communication in an automobile with a microphone array,” In Proceedings of ICASSP-92, San Francisco, CA, pp. I-281—I-284.