

# 行政院國家科學委員會專題研究計畫 成果報告

## 老年人居家使用之聲音辨識與偵測系統開發 研究成果報告(精簡版)

計畫類別：個別型  
計畫編號：NSC 100-2218-E-007-011-  
執行期間：100年06月01日至101年03月31日  
執行單位：國立清華大學電機工程學系(所)

計畫主持人：劉奕汶  
共同主持人：康仕仲、江秉穎  
計畫參與人員：碩士班研究生-兼任助理人員：羅祥友  
碩士班研究生-兼任助理人員：黃敘綱  
碩士班研究生-兼任助理人員：梁翰銘  
碩士班研究生-兼任助理人員：吳晨瑋

公開資訊：本計畫可公開查詢

中 華 民 國 101 年 06 月 28 日

中文摘要： 年長者居家，獨自處於寢室或浴室時，難免有需要協助的狀況。如屬緊急求救，傳統系統的設計，多半須要長者主動按鈴或是打電話求救。然而緊急狀況發生時，若長者的行動能力受到影響，或是由於特殊疾病別，就寢時不方便起身求助；諸如這些情況下，急救鈴或電話就可能派不上用場。本研究的目標之一，在於利用聲紋辨識的方式來監測緊急呼救聲，並即時聯絡周圍可提供幫助的人員。本計畫執行期間，團隊已蒐得長者在實際生活的場域中各種的聲音；並運用了音訊訊號處理的技術，透過傅氏變換與其他時頻分析的技巧，找到了一些合適的特徵，在電腦上初步開發了基於聲學的緊急狀況偵測演算法。

中文關鍵詞： 居家照護，聲紋辨認

英文摘要： Elderly people are likely to need assistance when being alone in their bedroom or bathroom. When the need is urgent, a conventional emergency contact system often requires the elderly to take an action, such as making a phone call or pressing an emergency button. However, the mobility of the elderly might often be compromised in an emergency. Also, due to certain illness, the elderly might not be capable of seeking for helps at night in beds. In these scenarios, an emergency bell (button) or the telephone may not be useful. A major goal of this research project is to detect emergent conditions based on acoustic cues and notify caregivers in a timely fashion. During a preceding research project funded by NSC, our team obtained vocal and non-vocal recordings from elderly people. These sound examples were put into a database so recognition methods could be developed. We utilized Fourier transform and other time-frequency analytic tools and identified a few acoustic features that are suitable for computer-recognition purposes. Emergency-detection algorithms were developed thereupon.

英文關鍵詞： home care, sound fingerprint recognition

行政院國家科學委員會專題研究計畫報告  
老年人居家使用之聲音辨識與偵測系統開發

計畫編號：NSC 100-2218-E-007-011

執行期間：100 年 6 月 1 日至 101 年 3 月 31 日

主持人：劉奕汶 清華大學電機工程系 助理教授

## 一、前言：

隨著台灣即將邁入高齡化社會，醫療照護品質與看護人力不足一直是社會議題；而高齡者的居家安全，乃是影響生活品質的重要因素。目前政府政策朝向「社區高齡者住宅」方向規劃推動，並重視數量龐大的既有住宅改建為高齡者住宅，以將高齡者社會住宅融入一般住宅社區環境——『提供優質高齡者住宅，有益高齡者的身心並減輕高齡者的照護負擔，節省照護費用與醫療費用；建構通用環境並提供多樣的高齡社會住宅，健全高齡者住宅市場之產業發展』乃政府提出的高齡化社會對策。針對於此，我們提出『VoiceCare 聲聲呼喚』產品以聲音之監測為基礎，提供高齡者安全照護服務；可針對各式各樣的住宅進行裝設，提供遠距看護服務，與住宅聲控功能。創意發想如下圖所示。

### VoiceCare 聲聲呼喚

使用情境說明、預期效益、潛在市場  
採用通用設計，高齡者或是暫時性失能者在住宅中皆能使用VoiceCare。可以減少長時間的看護人力與增進長者獨立生活的方便性與安全保障。除了高齡與無障礙考量之外，所有需要安全考量的家庭需求或是相關場所皆有潛在市場。

技術關鍵字  
語音辨識 聲紋分析 生理監測技術  
無線通訊模組 智慧住宅

搜尋關鍵字  
高齡住宅 智慧住宅 語音辨識 生理監測  
遠距看護 通用設計

### 設計主題與概念描述

隨著台灣即將邁入高齡化社會，醫療照護品質與看護人力不足一直是社會議題，高齡者在住宅當中日常生活的安全性是影響生活品質的重要因素。目前政府政策朝「社區高齡者住宅」方向規劃推動，並重視數量龐大的既有住宅改建為高齡者住宅，並將多樣的高齡者社會住宅類型融入一般住宅社區環境。

「提供優質高齡者住宅，有益高齡者的身心並減輕高齡者的照護負擔，節省照護費用與醫療費用；建構通用環境並提供多樣的高齡者社會住宅，健全高齡者住宅市場之產業發展」為此一高齡化社會對策，我們提出『VoiceCare 聲聲呼喚』產品以提供高齡者安全照護服務，可針對各式各樣的住宅進行裝設，提供遠距看護服務，與住宅聲控功能。由於高齡者所面臨到的是生理機能的全面性退化，在鼓勵子女與高齡父母同住的情形，長者在住宅中的安全是子女很急切需要關注，使用語音控制住宅機能能夠為高齡者減輕日常生活的身體負擔，並提供生活上的獨立性與成就感。

### 系統特性與功能

VoiceCare系統的特性融入通用設計的概念，結合通訊功能與智慧住宅之考量，使用VoiceCare進行之服務的優點，為適合居家使用，非接觸式、適合長期使用。

VoiceCare之設計為子母機	聲音控制住宅機能	遠距看護措施	緊急呼救	通話功能
 <p>無線控制傳輸 內嵌聲音過濾記錄裝置 功能設定 母機 子機 置於桌上或掛置牆上</p>	 <p>主機</p>	 <p>子機A</p>	 <p>子機B</p>	 <p>子機C</p>
主機體為聲音過濾紀錄儲存與功能設定 分機有防水功能與對外無線控制功能。	藉由聲紋與語音辨識功能控制整合住宅內電力設備之機能。	對日常生活進行聲音紀錄，利用過濾器過濾出需要的聲音，紀錄分析長者生活當中所呈現出的生理活動情形。當出現異常之聲音特徵頻率，即發出異常訊息通知照護人員注意。	在緊急時刻，可直接呼救即啟動警報器。選擇人名呼叫即撥出電話到特定號碼。	連結電話線路，可在日常生活當中直接通話使用。

## 摘要：

年長者居家，獨自處於寢室或浴室時，難免有需要協助的狀況。如屬緊急求救，傳統系統的設計，多半須要長者主動按鈴或是打電話求救。然而緊急狀況發生時，若長者的行動能力受到影響，或是由於特殊疾病別，就寢時不方便起身求助；諸如這些情況下，急救鈴或電話就可能派不上用場。本研究的目標之一，在於利用聲紋辨識的方式來監測緊急呼救聲，並即時聯絡周圍可提供幫助的人員。本計畫執行期間，團隊已蒐得長者在實際生活的場域中各種的聲音；並運用了音訊訊號處理的技術，透過傅氏變換與其他時頻分析的技

巧，找到了一些合適的特徵，在電腦上初步開發了基於聲學的緊急狀況偵測演算法。

### 英文摘要：

Elderly people are likely to need assistance when being alone in their bedroom or bathroom. When the need is urgent, a conventional emergency contact system often requires the elderly to take an action, such as making a phone call or pressing an emergency button. However, the mobility of the elderly might often be compromised in an emergency. Also, due to certain illness, the elderly might not be capable of seeking for helps at night in beds. In these scenarios, an emergency bell (button) or the telephone may not be useful. A major goal of this research project is to detect emergent conditions based on acoustic cues and notify caregivers in a timely fashion. During a preceding research project funded by NSC, our team obtained vocal and non-vocal recordings from elderly people. These sound examples were put into a database so recognition methods could be developed. We utilized Fourier transform and other time-frequency analytic tools and identified a few acoustic features that are suitable for computer-recognition purposes. Emergency-detection algorithms were developed thereupon.

## 二、 文獻探討

國內研究生活環境聲響辨認之先驅，乃成功大學王駿發教授團隊。在生活聲響的辨認研究目的上，主要是針對居家自動化（Home Automation）這項應用[1]，研究的聲音資料庫包含 11 類 471 筆音檔。聲音資料經由特徵萃取後，進行分類。特徵萃取的方式有二：其一乃透過對梅爾倒頻譜（Mel Frequency Cepstral Coefficients, MFCC）進行獨立成分分析（Independent Component Analysis, ICA）轉換而得；其二是由總頻譜功率、子頻段功率、質心、頻寬、音高組合成的感知特徵（Perceptual Features）。分類器的設計則是使用支持向量機（Support Vector Machine, SVM）來做聲音的辨識，此方法含有運算較為簡單的優點[2]。實驗結果可獲得相當高的辨識率(91%)。

成大團隊並曾將聲音辨識的演算法，在一個  $8 \times 11.6 \times 2.7 \text{m}^3$  的室內，於天花板架設六支麥克風進行實測；麥克風間距皆相隔 3.6m，據我們所知，此乃國內首見以多通道訊號進行生活環境聲響辨認之研究，但實驗結果準確率只有達到 74% 的效果[3]。

關於多通道聲音訊號在居家聲響分析上的應用，國內尚有台大劉佩玲教授所指導研究生，曾經進行步行聲源定位的研究 [4]。該研究係採用步行於地板所產生的訊號，包括了振動與聲音，來做分析；有別於傳統麥克風陣列演算法以波源方向角度進行定位，該篇論文中推導出了掃描最佳化法，在有限空間內可以搜尋到誤差函數的最低點，來判定波源位置。對於非腳步聲或並非透過地板傳遞的聲音定位，則未有著墨。

## 三、 研究目的

本計畫預計透過製作『聲聲呼喚』的原型，整合相關技術，以評估此方案的可行性。擬研發之功能包括：

【語音辨認】 開發特定詞彙之語音辨認，可以經由最精簡次數輸入訓練，達成 10-20 個詞彙的辨識功能。

【聲紋辨識】 透過實地錄音的方式，建立聲紋資料庫，以統計各個頻譜特徵在正常情況下的合理值與範圍；並進一步實現三至五種緊急聲響的偵測。包含下列種類：

拍打聲：以俾年長者跌倒後，可以透過拍打地面的方式求助。

腳步聲：以偵測失智年長者夜間睡眠中斷而起來走動的情形。

翻身聲：以偵測週期性肢動症患者夜間不斷翻身的情形。

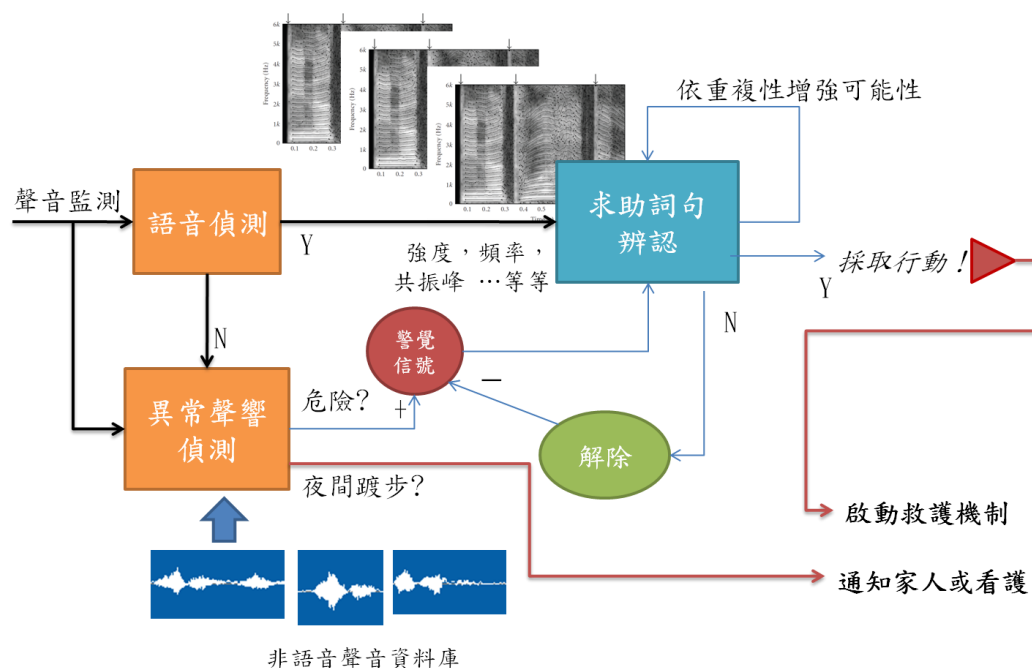
鼾聲：鼾聲與上呼吸道阻塞有直接相關性。

其他類別之緊急聲響

【聲源偵測】 透過麥克風陣列的設置，探討利用多聲道錄音偵測聲音來自何處的可行性。

【緊急狀況通報】當緊急狀況發生時，應設計確認機制，使得本系統能以語音或其他簡單的方式向求助者（高齡者）確認送出求助通知。這部分應特別考慮使用者族群之特殊需求，如聽力或視力的衰減，而作出客製化的設計。或是若長者屬於老年失智的情況，則可能預要略過此確認的機制，直接通報照護工作者。

#### 四、 研究方法



圖一：聲紋辨識系統方塊圖

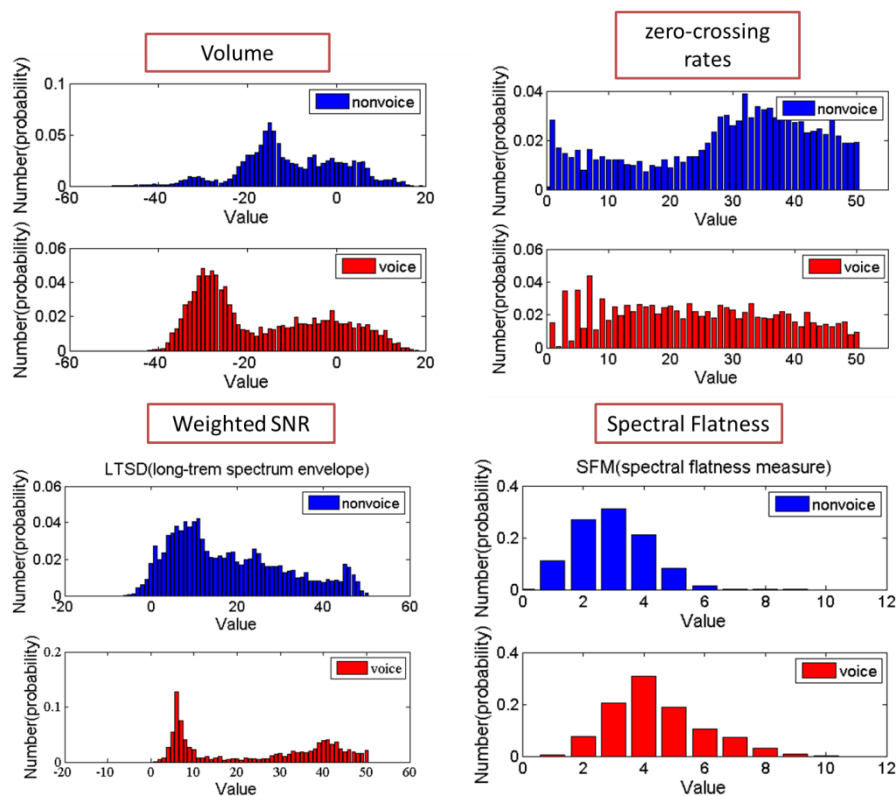
可行性評估階段，本團隊在軟體上測試了聲紋辨識系統的各個主要方塊，包含了：

- 語音偵測 (voice detection)
- 異常聲響偵測與辨認
- 求助詞句辨認

以上三者，團隊之研究助理均嘗試自行撰寫出合用的程式；而求救詞句辨認的部分，也同時試用了 microsoft Win7 作業系統所附的語音辨認軟體。我們初步認為，電腦在能夠確定輸入聲音為語音的前提之下，現有的語音辨認技術已經十分成熟；不論是採取「動態時間對整」(dynamic time warping)，高斯混和模型(Gaussian mixture models),隱藏馬可夫模型 (hidden Markov models)，或是「向量支持機」(support vector machines)等技術，文獻皆有脈絡可循[5-9]，所須考慮的僅有運算成本（含軟硬體複雜度並能量的消耗）與辨識率的權衡。值得注意的是：傳統語音辨認應用情境，多半假設使用者會確認系統開始運作後，才立刻說出指令。例如：手機的語音辨識，使用者須先按鈕；而電話的語音客服，會採目錄或問答方式明確指引並迫使 (urge) 使用者下達指令。然而，這樣的人機互動假設在本產品的『緊急求助』的應用情境上並不成立。緊急求助系統必須隨時處在待命狀態，並主動發現環境中的語音詞彙、加以判別。

所以，技術的困難點，在於看來較單純的 voice detection 與異常聲響偵測。經過了數個月來的演算法開發與程式撰寫，測試結果遇到了一些的瓶頸。大抵看來，演算法在安靜的環境中都能執行得不錯，但一遇嘈雜的環境，其表現則不免降低。以下敘述聲音偵測的演算法與其強健性的測試方式，與測試結果。





圖二：常用於語音偵測的音訊特徵值分布圖。藍色長條顯示非語音的分布，紅色長條則為語音分布。兩者分布的範圍均嚴重重疊。

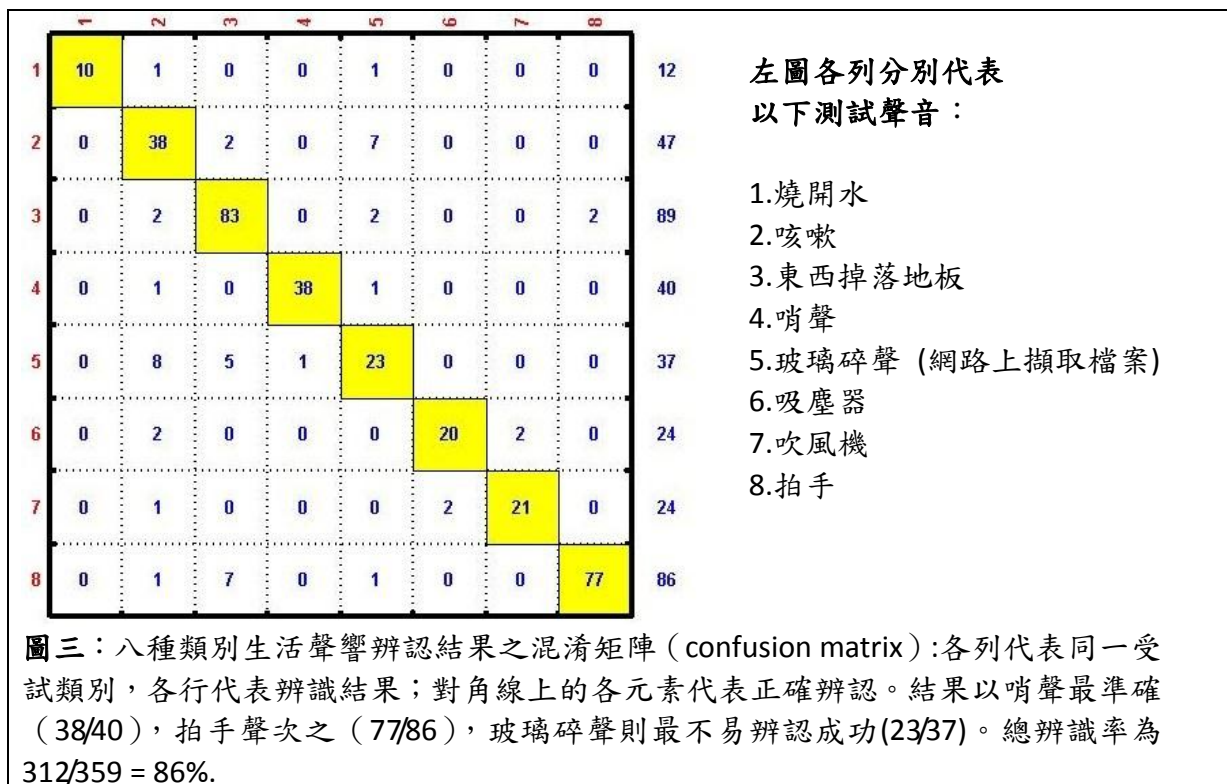
### ● 語音偵測

如圖一的左上角所示，所有的訊號進入系統時，均先辨別是否為語音。本團隊嘗試用過文獻中常用的幾個特徵，包括：音量(volume)、過零率(zero-crossing rates)、加權訊雜比(weighted SNR)、與頻譜平坦性(spectral flatness)等等。圖二顯示以預錄的聲音資料來測試這些特徵值，在語音及非語音的統計結果，發現除了音量以外，均有嚴重重疊。然而，音量在本研究之應用情境，也不適合作為語音偵測的依據；原因在於：求救者的聲音不見得比背景音的音量來得大。

故此，我們發揮了一點創意，定義出一個嶄新而簡單的方式，能藉由檢查聲音頻譜是否具有基頻(fundamental frequency)與泛頻(harmonics)的結構，來判定一段輸入的聲音是否為語音。經過反覆測試，發現這個基頻—泛頻的結構檢查，能有效區別語音與包括咳嗽、腳步聲...等八種非語音的聲響，而且能對抗白雜訊的干擾：在SNR為7dB以上時，誤判率(false positive)與錯失率(miss)均保持在10%以內。所以，這個方法目前看來十分有效，可將程式寫入第二階段之原型機中。

### ● 非語音之辨認

承1.1.1，本團隊在這段日子中，錄置了八種日常生活中聲響，共三百多個聲音的小型資料庫，以用於聲響辨識演算法之開發與測試。我們採用了梅爾導頻譜係數(Mel-frequency cepstral coefficients)與其時變率(time-derivatives)作為特徵向量，並且運用高斯混和模型嘗試分辨這八種聲響。測試結果如圖三所示。目前在這八種聲響間，發現電腦對於哨音與拍手聲辨認最正確，玻璃破碎聲則較不易辨認，而總正確率為86%。



● 「外部類別不予辨認」(outlier rejection) 機制

圖三所試測試結果，乃假設對任意聲音都強制分類。為了能夠預防其他聲響「誤觸警鈴」，本團隊開發了基於可能性比率(likelihood ratio) 的 outlier rejection 法則，目標是要對於不在這八種類別之內的聲響，都不予辨認。所以，我們又測試了水龍頭、甩垃圾袋、沖馬桶、抽油煙機、拉鍊、開包包、開關門、開關燈、敲門、撕紙、箱子摩擦地面、滾輪滑動...等等五花八門的聲音，發現技術上可以同時將誤加辨認率 (false acceptance) 與原屬正確、但被排除的遺漏率(miss) 同時降至 25% 左右，好處是：當予以辨認時，正確率就提高至 90% 以上 (230/245) 了。這個成果相信還有改良的空間。下表列出 outlier rejection 機制的測試結果：

聲響類別	予以辨認		不予辨認	總計
	正確	錯誤		
內部類別 (燒開水等八種)	230	15	114	359
外部類別 (其他聲響)	30		90	120

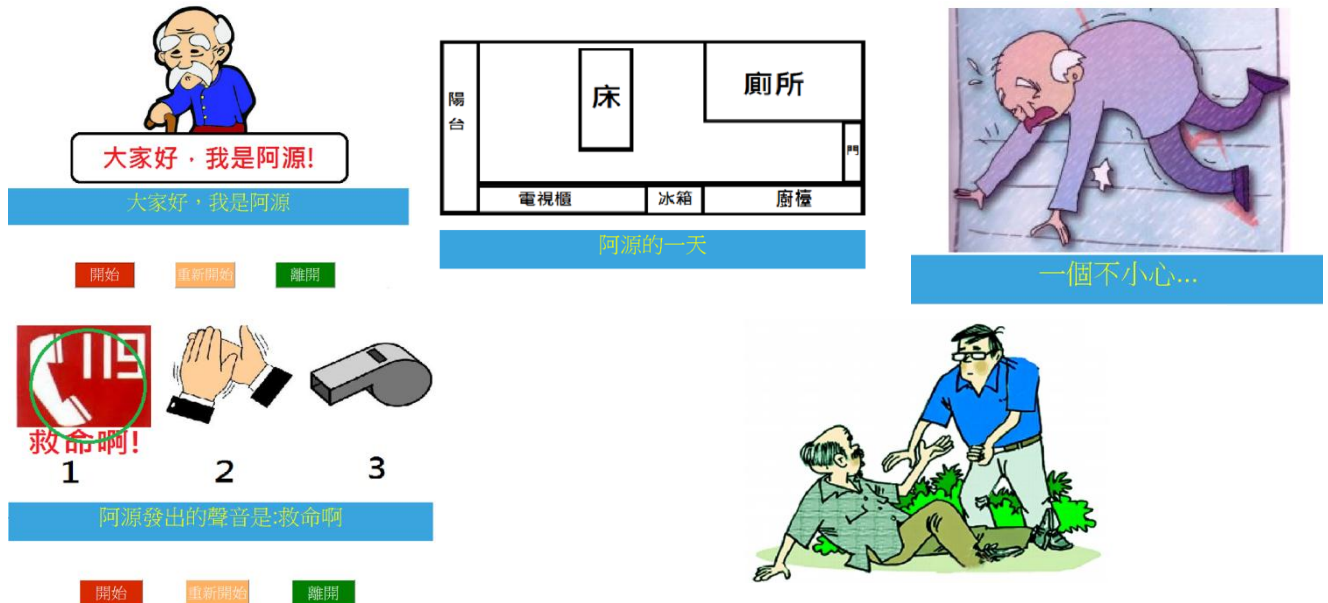
● 非語音辨認方法之強健性測試

對於非語音的辨認方法，我們另外還做了系統對抗雜訊與回音的強健性測試，由於篇幅起見，將於結案報告中陳述。大致的結論是：系統對於回音並不敏感，但雜訊會使系統辨識率下降至 60-70% 左右 ( $SNR > 5 \text{ dB}$ )。

## 五、 結果與討論

如上所述，本團隊評估與開發了語音偵測、以及環境聲響辨認方法。語音偵測部分，發現文獻中的方法有難行之處，因此找到了「基頻—泛頻結構」之特徵，能有效在嘈雜環境中辨認一段輸入的聲音是否為語音。環境聲響辨認的實測結果，亦已歸納於 1.1.2。

除此之外，在可行性評估計畫執行過程中，為配合台大智活與雙連安養中心歲末感恩聚會，本團隊花了一個多月的時間，將抽象的聲音辨識技術寫成了一個互動遊戲，描述長者「阿源」第一天到安養中心，先是遇到了小偷，後來兒孫來訪順便為孫子慶生，最後晚間如廁不慎跌倒的一個故事。邀請長者試玩，必須吹口哨、拍手、發出「救命啊」命令，才能幫助阿源順利過關。此遊戲擷取的畫面如圖四所示。



圖四：求救情境聲音互動遊戲。描述長者阿源在遷入安養院的第一天，因環境陌生而跌倒的故事。左下圖為參與者喊「救命啊！」之後，電腦辨認出來的畫面，而右下角所示乃最後「過關」的畫面。

## 六、 結論

在原計畫書中，本團隊曾預期達到下列規格

(括弧所示為技術指標：1 = 容易；5 = 困難)：

- A. 詞彙辨認：10-20 條關鍵詞句辨認，辨認準確度達統計顯著並實用的程度。(4.0)
- B. 聲紋辨認：三至五種生活中可能出現的聲響，辨認誤判率與失誤率同達最佳化。(4.5)
- C. 音源辨認：系統能辨認聲音從哪個房間傳來 (4.0)
- D. 演算即時性：所有訊號處理及辨識的演算法，所需要的延遲都不超過 10 秒 (3.5)

回顧這些技術目標，A 為可行，關鍵技術在於語音的偵測；但 10-20 條關鍵詞句乃本團隊目前尚未著力的部分，未來應可找到擁有此技術的廠商合作。B 可行，但誤判率與失誤率目前系統表現尚有改進空間。C 比當初設想的要棘手些，但可能屬於次要功能（如幫助分別長者說出的語音與電視傳出來的語音），未來可繼續開發。D 在 1.2 所述的互動遊戲中已經達成。綜言之，相較於最初的創意發想，我們有了充足的進展，使得創意開始具體落實了。

## 參考文獻

- [1] Jia-Ching Wang, Hsiao-Ping Lee, Jhing-Fa Wang & Cai-Bei Lin, "Robust environmental sound recognition for home automation," *IEEE Transactions on Automation Science and Engineering*, pp. 25-31, Jan. 2008.
- [2] 林博川博士，*private communication*.
- [3] 林財貝，「應用機率型 SVMs 與 ICA 於以內容為基礎音訊分類之研究」，國立成功



大學電機工程學系碩士論文，2006.

- [4] 陳昭廷，「室內步行聲源定位」，國立台灣大學應用力學所碩士論文，2010.
- [5] H. Sakoe, and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech and Signal Processing*, **26**(1) pp. 43–49, 1978.
- [6] Reynolds, D.A. and Rose, R.C. "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Transactions on Speech and Audio Processing* **3**(1), pp. 70-83, 1995.
- [7] Rabiner, L.R., "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, **77** (2): 257–286, Feb. 1989.
- [8] Cortes, C., and Vapnik, V. N. "Support-Vector Networks", *Machine Learning*, **20**, pp. 273-97, 1995.
- [9] Hsu, Chih-Wei, and Lin, Chih-Jen (2002). "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks* **13**(2), pp. 415-25, 2002.

# 國科會補助計畫衍生研發成果推廣資料表

日期:2012/06/26

國科會補助計畫	計畫名稱：老年人居家使用之聲音辨識與偵測系統開發	
	計畫主持人：劉奕汶	
	計畫編號：100-2218-E-007-011-	學門領域：訊號處理
無研發成果推廣資料		

100 年度專題研究計畫研究成果彙整表

計畫主持人：劉奕汶

計畫編號：100-2218-E-007-011-

計畫名稱：老年人居家使用之聲音辨識與偵測系統開發

成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	3	100%		碩士論文於今年七至九月將陸續完成三篇
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	1	1	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	4	4	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

<p>其他成果</p> <p>(無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)</p>	無
---	---

	成果項目	量化	名稱或內容性質簡述
<div>           科 教 處 計 畫 加 填 項 目         </div>	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	



# 國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

☒達成目標

☐未達成目標（請說明，以 100 字為限）

☐實驗失敗

☐因故實驗中斷

☐其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：☐已發表 ☐未發表之文稿 ☒撰寫中 ☐無

專利：☐已獲得 ☒申請中 ☐無

技轉：☐已技轉 ☐洽談中 ☒無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

人口老化係為全球各國之共同問題，其中東方人又特別不願意進住安養院接受全時照護，本項研究可提供居家之全時照護；而歐美又為非常注重個人隱私者，本專題研究還可令被關護人不會有全時遭人監視錄影的感受，使高齡者的生活，在獨立性與倚靠性中取得更美好的平衡。本專題研究目前著重於技術開發，未來將持續朝原型機製作的方向努力。