

基於高頻音訊之行動資訊廣播系統

游家一、廖元甫、吳業寬、張量鈞
國立台北科技大學電子工程系
財團法人資訊工業策進會

摘要 一本文提出利用聲音做為傳輸媒介且利用人耳聽覺特性，將所要傳送的資訊，埋藏在高頻音訊裡，並實際應用在行動裝置上，讓店家可以使用揚聲器不停的播放聲音訊號，使消費者靠近店家即可收到該店家的消費資訊。因此，我們在傳送端，設計使用 18kHz 左右的音頻，配合 RS code、preamble encoder 和 PSK 傳遞訊號，並在接收端使用 preamble detection、PLL、RS decoder 解出訊號，讓整個系統最後備實現在 iOS 裝置上，且確實可以成功傳遞解碼商家訊息。

一、簡介

由於智慧型手機的普及，因此如何能於消費場所，感知消費者之位置是否於商家附近，進而提供商家最新消費訊息至消費者手機上之方法，為現今熱門的議題。如圖一的情境構想，消費者的行動裝置安裝由賣場提供的特定應用程式，當靠近一個商家時，行動裝置可以立刻收到該商家的消費資訊。



圖一: 情境示意圖

較為常見的推播方法可以分為兩種類型，第一是利用電磁波，第二是藉由聲音的傳播，電磁波的技术有藍芽、RFID、Wi-Fi 與 NFC。藉由聲音傳播的技术有 Audio 與 Ultrasonic，這兩者之間可用頻率來區分，20kHz 以下是 Audio，20kHz 以上是 Ultrasonic。

RFID(Radio Frequency Identification)無線射頻辨識系統與 NFC，近幾年來被應用在悠遊卡、ETC...等，其優點為可快速辨識，缺點是普遍的 RFID 與 NFC 發射器必須與接收端非常靠近才可做資料的溝通。

利用藍芽傳送，好處是裝置的擺放不必對齊，消耗的功率小，但藍芽傳送將會遇到一個問題，因為要使用藍芽功能時，兩部裝置必須配對，且會留下裝置的資訊，所以當使用者進入一個曾經配對過的裝置的範圍時，可能會被有心人士利用竊取個資。

使用 Wi-Fi 做為推播系統，好處是傳輸距離遠，資料量大，缺點與藍牙一樣需要配對，且必須另外購買無線分享器，而一部分享器可同時連線的用戶數量並不多，如果實際應用在賣場或大型商圈等場所時，會有成本問題的考量。

聲音傳遞技術裡的 Audio [1]，方法是用 Audible(可聽見的聲音)來傳遞資訊，發射端將音調編碼，或再經過調變並撥放，接收端的麥克風收到後開始解析音符的頻率，藉由不同音調的順序，可達到傳遞訊息的目的。但缺點是不適合在一般的公共場合使用，原因是一般公共場合的背景雜訊的頻帶與發射訊號的頻帶相近，會造成接收端解調的困難，且因為發射的訊號是可以被人聽到的，可能會讓其他民眾造成不適。

使用 Ultrasonic [2]，這種方法資料傳輸量夠大，可以到達 200kb/s，人類對超音波的感知並不敏感，在有超音波的環境下並不會感到不適，但因發送超音波設的備昂貴且不常見，所以並不普及。

因此，考慮人耳頻率響應與賣場背景雜訊之特性，還有一般消費者的習慣，在本論文中，我們選擇使用頻帶在 16-20kHz 之間的 Audio 訊號，來製作一個高頻音訊行動資訊廣播系統。

因大部分的人的聽覺頻帶在這之間已接近聽不到的範圍，且使用一般的揚聲設備就可以將訊號推播，所以此系統可利用一般商家已有之揚聲設備，傳送人耳不易察覺之高頻音訊訊號（亦可藏在一般音樂或廣告廣播中），並在其上承載此商家特定之消費訊息，接著消費者利用行動載具上已有之麥克風接收與解碼此高頻音訊訊號。最後於手機上即時顯示消費訊息通知，達到為此商家即時傳遞特定消費訊息的目的。

二、系統設計

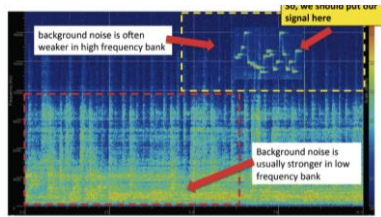
以下進一步說明完成此高頻音訊行動資訊廣播系統，所需考慮的各種系統規格。

2.1 音訊訊號頻帶選擇

為能利用人耳不易察覺的音訊訊號傳遞消費資訊，在本論文中，我們將調查人耳的聽覺特性與背景環境雜訊分佈情形。我們首先綜合各聽覺研究文獻的結論，找出人耳對各種頻率的頻率響應。基本上人耳最容易聽覺的範圍在 2~4 kHz 左右，而在 16 kHz 以上，人耳的敏感度就急劇降低。

然後我們到賣場以頻譜分析儀，實際測量環境雜訊，發現雜訊能量分佈一般如圖二所示，大部份的賣場噪音皆

分佈在低頻帶部分，尤其是與人耳最易感知的 2~4 kHz 重疊。

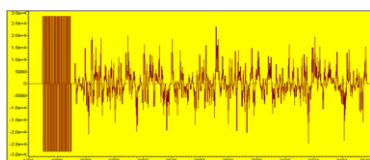


圖二 背景雜訊

因此，根據上述調查結果，我們進一步以實際音訊訊號與環境雜訊，進行不同頻帶的人耳聽覺測試。最後我們選擇以 16~20 kHz 的頻帶，作為主要音訊通訊訊號的載波頻帶。

2.2 封包方式

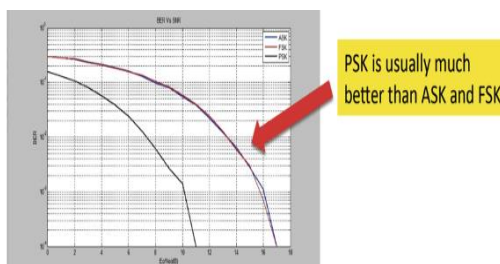
考慮連續傳送封包，我們必須在每個封包的開頭加上一段 preamble，如圖三，如此接收端可以依據此訊號作時序同步。



圖三 preamble 示意圖

2.3 調變方式選擇

調變系統我們採用 PSK 來作為基本架構，其中考慮 PSK 是因為從一般無線電與有線通訊系統文獻上可知，PSK 的效能通常比 amplitude shift keying (ASK) 和 frequency shift keying (FSK) 好。佐證資料可見圖四之 PSK, ASK 與 FSK 編碼方式與其效能比較。



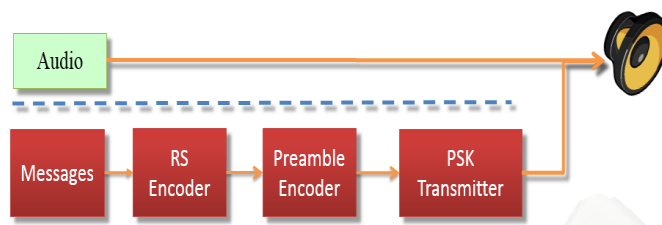
圖四 調變系統比較圖

三、系統實現方法

在本論文中，依據上述設計，提出一個基於高頻音訊通訊系統之特定消費訊息傳遞系統。圖五為這套系統的傳送端架構。以下將逐一介紹此系統重點的方塊圖的功能，原理。

3.1 傳送端

一開始我們會先把所想傳送的商家資訊(網址)，如 www.iii.org.tw，轉換成 ASCII 碼，是由二進位碼組成。



圖五 傳送端架構

3.1.1 RS code

錯誤更正碼可用來保護數位資料，主要目的就是在最小成本頻寬內降低錯誤率。在已知的錯誤更正碼中，RS(Reed-Solomon) code 對突發性的錯誤有很強的更正能力，主要功用是用來糾正隨機雜訊(random noise)所造成的錯誤，此種錯誤的發生通常是零散的。雖然 RS code 能夠糾錯 t 個未知錯誤(error)，不過需要犧牲 $2t$ 個冗餘資料(redundant data)。另外，設定每筆資料(symbol)的位元數為 m ，因此一組 code word 的長度(包括冗餘資料)最多為 $n=2^m-1$ 。

3.1.2 Preamble

是一具有與白雜訊類似的自相關性質的 0 和 1 所構成的編碼序列，此序列出現 0 與 1 的機率各為 0.5，在我們所要傳的資料開頭加上 PN code，可讓接收端知道資料的起始點，選用 PN code 作為 preamble 的原因是它與我們所傳的資料最不相像，這樣接收端在搜尋檔頭時就較不會出現錯誤。

3.1.3 PSK 調變與 Root raised cosine filter

PSK 調變我們選擇使用 QPSK，理想的 QPSK 訊號表示如(1)

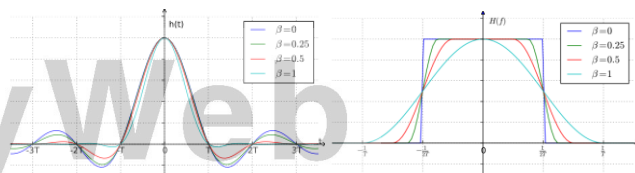
$$s(t) = A_I \cos(2\pi f_c t) - A_Q \sin(2\pi f_c t) \quad (1)$$

這裡我們定義 bit = 0, $A_I, A_Q = -1$; bit = 1, $A_I, A_Q = +1$

我們需要選擇一個波形來傳送 symbol，最理想的波形是 raised cosine [3]，第一可限制頻寬，因為我們可以利用的頻寬有限，使用 raised cosine 作為傳送波型的頻寬公式如(2)

$$BW = \frac{1}{2} R_s (\alpha + 1) \quad (2)$$

在這裡 R_s 為符號傳送速率， α 為降滾係數，介在 0 到 1 之間，所以我們可以藉由這兩個參數的調整來決定頻寬大小。圖六、圖七為不同的 α 所描繪出的 raised cosine 的時域與頻域圖。

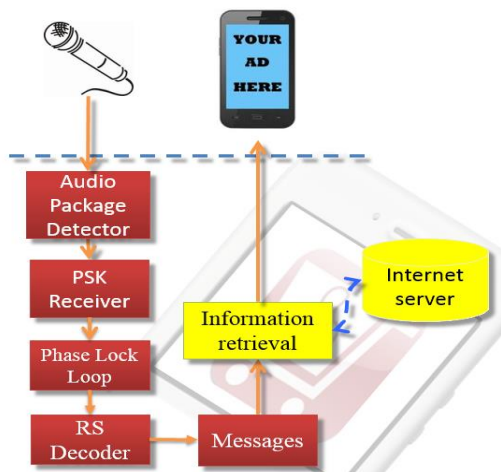


圖六 raised cosine 時域波 圖七 raised cosine 頻域波形

第二可對抗符號間干擾(ISI)，我們將符號的最高點與第二個符號零點重疊，這樣一來干擾可以減到最小。

3.2 接收端

圖八為接收端的架構圖，我們將把收到的 Audio 訊號經過 PSK 解調，preamble 偵測同步，再做 PLL 同步頻率與相位，最後經過 RS decoder 解出所傳的商家資訊，如:www.iii.org.tw。



圖八 接收端架構

以下逐一介紹各個模組功能與原理。

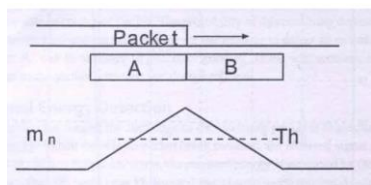
3.2.1 Preamble detector

因為我們可以事先知道 preamble 是什麼，所以我們可以事先在接收端把 preamble 波形記錄下來，偵測 preamble 的起始點有兩種方法，一種是 single sliding window，選定一個視窗的長度，對視窗內的取樣點做自相關運算，然後藉由 m_n 的大小判斷是否達到門檻值，如式(3)。

$$m_n = \sum_{k=0}^{L-1} r_{n-k} r_{n-k} \quad (3)$$

這裡 n 是資料的 bin，k=0 到 L-1 是 preamble 資料的長度。

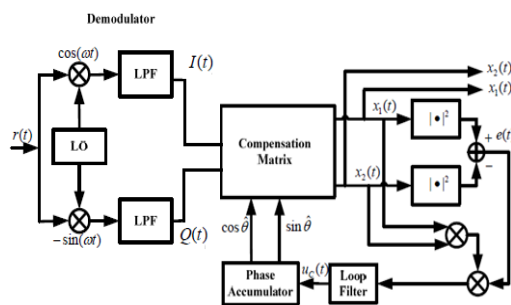
判斷 m_n 是否有達到訂定的門檻值，我們必須先知道雜訊的能量，而雜訊能量我們無法事先得知，所以我們用另一種偵測 preamble 的方法，double sliding window，示意圖如圖九，使用這個方法不再是判斷門檻值，而是直接取峰值，



圖九 double sliding window

3.2.2 PLL

訊號在接收端解調時，會碰到訊號的頻率與相位與接收端不一致的問題，這裡我們參考了 Feed-forward Compensation Algorithm [4] 的方法，這個方法好處是完全可以採用 DSP 的方式實現出來，流程如圖十



圖十 頻率與相位補償流程圖

$r(t)$ 為接收端收到的訊號，可以表示成(4)

$$r(t) = A_I \cos(\omega t + \theta) - A_Q \sin(\omega t + \theta) \quad (4)$$

θ 為相位差。在這裡我們將低通濾波器設計成 raised cosine filter，一方面是此濾波器有低通濾波器的效果，一方面是與傳送端做匹配。

訊號經過濾波器可得到 I、Q 兩個 channel

$$I(t) = \frac{1}{2} A_I \cos \theta - \frac{1}{2} A_Q \sin \theta \quad (5)$$

$$Q(t) = \frac{1}{2} A_I \sin \theta - \frac{1}{2} A_Q \cos \theta \quad (6)$$

補償矩陣運算式如下：

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} I(t) & Q(t) \\ Q(t) & -I(t) \end{pmatrix} \begin{pmatrix} \cos \hat{\theta} \\ \sin \hat{\theta} \end{pmatrix} \quad (7)$$

$\hat{\theta}$ 為補償相位，而將 $x_1(t)$ ， $x_2(t)$ 展開我們可以得到下列方程式：

$$x_1(t) = \frac{1}{2} A_I \cos(\theta - \hat{\theta}) - \frac{1}{2} A_Q \sin(\theta - \hat{\theta}) \quad (8)$$

$$x_2(t) = \frac{1}{2} A_I \sin(\theta - \hat{\theta}) - \frac{1}{2} A_Q \cos(\theta - \hat{\theta}) \quad (9)$$

接下來將 x_1 ， x_2 ，經過以下的運算，我們就可得到 error function:

$$e(t) = [x_2^2(t) - x_1^2(t)]x_1(t)x_2(t) = \frac{1}{16} \sin[4(\theta - \hat{\theta})] \quad (9)$$

到這裡就可以發現，要讓 $x_1(t)$ ， $x_2(t)$ 近似於 $I(t)$ 與 $Q(t)$ ，我們必須使 $e(t)$ 的值接近零，也就是讓補償相位接近訊號的相位差。再來利用 loop filter 來調整 $\hat{\theta}$ 的值。

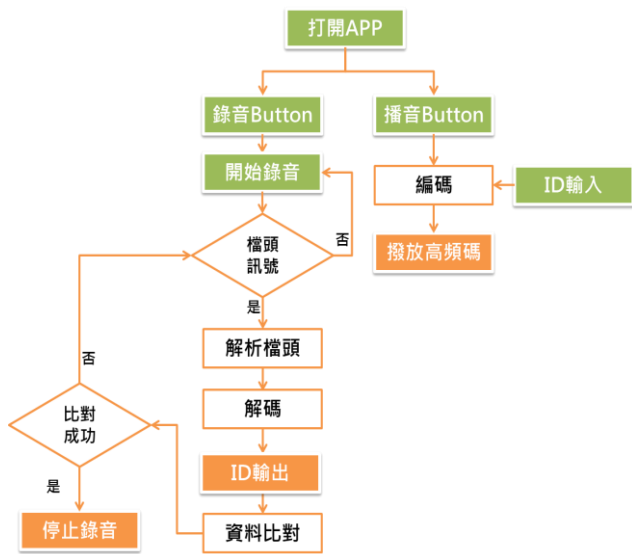
四、實驗步驟與 iOS 應用程式

4.1 iOS 程式架構及流程

APP 程式架構上我們是使用 iOS SDK 裡的 audio unit 的 framework 來實現錄放音的動作，理由是 audio unit 是 iOS 裡最底層的錄放音單元，較容易存取 raw data，這對於我們需經常使用 dsp 來說，較容易操作與實現。

雖然 iOS 的行動裝置有支援傳統的 C 語言，但如果使用傳統的 C 語言來實現 DSP 運算，將會大大增加運算時間，如 convolution 的運算，經過測試使用傳統 C 語言來撰寫，將一個 4096 長度的陣列與 257 長度陣列做摺積運算，花費時間大約要五秒，而使用 iOS 提供的 vDSP 系列的運算

函式，運算時間極快，而傳送端與接收端的程式計算時間分別只要一到兩秒。圖十一為系統實作流程圖



圖十一 系統流程

4.2 載波與頻寬

頻帶的選擇我們是依據訊號經由頻譜分析儀的能量來決定，因行動裝置的揚聲器與麥克風的取樣頻率最高可達到 44.1kHz，而根據調查結果人對 16kHz 以上的 audio 的敏感度就急遽下降，19kHz 以上的訊號因裝置的頻率響應能量有所衰減，所以我們將載波定在 18kHz。因我們所要傳送的符號為 raised cosine，訊號的頻寬我們可以依照公式(1)來決定，符號傳送速率 R_s 選擇 $\frac{32}{44100}$ 秒，也就是在取樣頻率 44.1kHz 的情況下，每 32 個 bin 傳送一次符號， α 選擇 0.5，所以訊號所占頻寬大約等於 1033Hz。

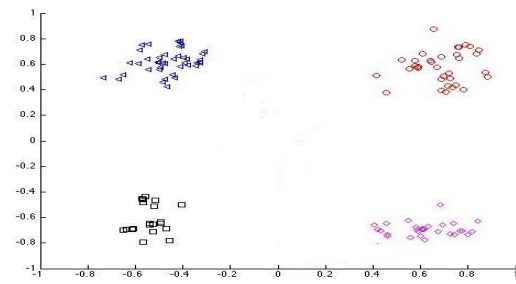
4.3 資料與傳送速率設定

我們要決定裝置錄放音的 buffer 大小，我們選擇 4096bit，如果使用 raised cosine 做為傳送波形，必須先將資料 upsample，upsample 多少取決於 R_s ，所以一個 buffer 可容納的資料為 $4096/32=128\text{bit}$ ，而我們是使用 QPSK 調變方式，可傳送的資料就可達到 256bit。

因為在實際應用上，要傳送的資訊只有一小段網址或資料，依據 ASCII 的標準，一個字元是由 8 個 bit 所編碼。preamble 的大小選擇 32bit，如果再加上錯誤更正碼、保護碼等，足以將一段完整的網址容納在一個 buffer 裡。

4.4 實現結果

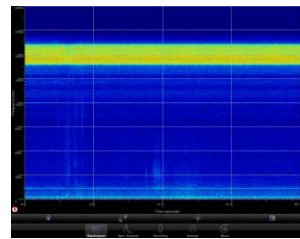
這邊我們先測試不加保護碼的情況，利用行動裝置的錄放音設備實際播放接收，並讀取行動裝置存下來的資料來驗證演算法的正確性，傳送資料為 www.iii.org.tw，圖十二為行動裝置解調完畢，經過相位補償的星座圖，依照星座圖的結果，我們可以完整解出 www.iii.org.tw 這段網址。



圖十二 經補償星座圖

4.5 實驗環境

圖十三是我們使用 iPhone 做為傳送端,iPad 做為接收端的模擬結果，在 iPhone 上我們設定的載波頻率是 18kHz，在 iPad 上可以看到資料的載波在 18KHz，資料的頻寬大約 2kHz，圖十四為接收端程式運算與解調結果，在 iPad 上可以看到我們傳送的網址及網頁。



圖十三 接收頻譜



圖十四 接收端解調結果

結論

在本論文中提出了一個用高頻音頻訊號的行動資訊廣播系統，整個系統並實際實現在 iOS 裝置上，且從實驗結果來看，此系統確實可以成功傳遞商家訊息。並讓顧客可以用手機將訊息完整解讀出來。未來本系統將再實作 RS 保護碼，使資料量可以傳更多，可傳送的距離也可以更遠，錯誤率更低。

附錄

本研究部分依經濟部補助財團法人資訊工業策進會「102 年度 資策會創新前瞻技術研究計畫」辦理。

參考文獻

- [1] Anil Madhavapeddy, David Scott, Alastair Tse, and Richard Sharp *Audio Networking: The Forgotten Wireless Technology* 2005 IEEE.
- [2] Chuan Li, David A. Hutchins, and Roger J. Green *Short-Range Ultrasonic Communications in Air Using Quadrature Modulation* IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 56, no. 10, October 2009.
- [3] http://en.wikipedia.org/wiki/Raised-cosine_filter.
- [4] Yinsheng LIU, Cheng TAO, *Feedback compensation algorithm for BPSK/QPSK carrier synchronization*, RADIOENGINEERING, VOL 19, NO 1, APRIL 2010