

行政院國家科學委員會
大專學生參與專題研究計畫書

立體聲上的 PVD 資料隱藏

執行計畫學生：李東岳

指導教授：吳坤熹 教授

日期：2014 年 02 月 10 日

壹、摘要

在這個資訊爆炸的年代，資料的傳遞無所不在。在資料傳送的過程中，隨時都有可能會被有心人士竊取，資料的加密與資料隱藏越來越重要。在傳統影像上的資料隱藏，由最早的最低位元替換藏密法（LSB），固定每次藏最低位元，藏密量不高，到後來的像素值差異藏密法（PVD），改成依照兩相鄰像素差值大小決定資料的藏入量，大幅的增加藏密量。

近期開始有了聲音上的資料隱藏，相對於影像藏密量就沒那麼高了。在聲音裡只要修改超過 2 個位元便聽的出有些瑕疵，4 個以上就聽不出原本的內容了。在 2010 年有人提出了類似 PVD 的作法在聲音上，優點是會讓人類的聽覺系統察覺不出來，但是也限制住了藏密量，並且每次都需要兩個檔案才能進行資料隱藏。本計畫如果成功便能增加藏密量，並且只需要一個立體聲檔案便能完成。

貳、研究動機與研究問題

研究動機

之前閱讀了一篇 IEEE 上發表的論文，它是將資料經由 Pixel Value Differencing(PVD) 的方式隱藏在兩個 audio 檔案裡面，為了讓資料

都固定藏在最後兩個 bit，它還特地將每個區間值設為固定，這麼一來會有兩個缺點：第一，每次要藏資料必須要有兩個 audio 檔案；第二，可以藏的資料量就會被壓縮。如果將兩個單聲道的 audio 檔案改成為一個雙聲道檔案，這樣只需要一個 stereo audio 檔案就能夠進行 PVD 做資料隱藏；而區間值如果能按照原本 PVD 的精神，依照不同的差值來藏不同的資料量，capacity 就能增加。

研究問題

基於上述的研究動機，本研究問題為：

- 1. 在 stereo audio 裡進行 PVD 隱藏資料。
- 2. PVD 的區間大小修改為依據差值做調整。

參、文獻回顧

PVD 影像[1]：

在傳統圖像的 PVD 隱藏中，它是利用兩相鄰像素的差值來藏秘密訊息，根據像素差值的大小區分成若干個區間，再依照特定的區間藏入特定的訊息量。每個區間都有設定可藏入的 bit 數以及上下限值。

K	1	2	3	4	5	6
R_k	0~7	8~15	16~31	32~63	64~127	128~255
n	3	3	4	5	6	7
u_k	7	15	31	63	127	255
l_k	0	8	16	32	64	128

表一

實際例子：

假設目前兩像素值為 192、32，要藏入的資料會先轉換為二進位的字串，在這裡假設為 0100010，計算出像素差值後落在區間 6，可藏入 7 個 bit，將要藏入的訊息轉為 10 進位後得到 34，加上區間下限值 128 後平均分配到兩個像素上，便可得到新的像素值 193、31，完成 PVD 藏入演算。

取出：

將拿到的 pixel 193、31 相減得到差值 162，扣掉下限值後變可得到訊息 34。

PVD 聲音[2]：

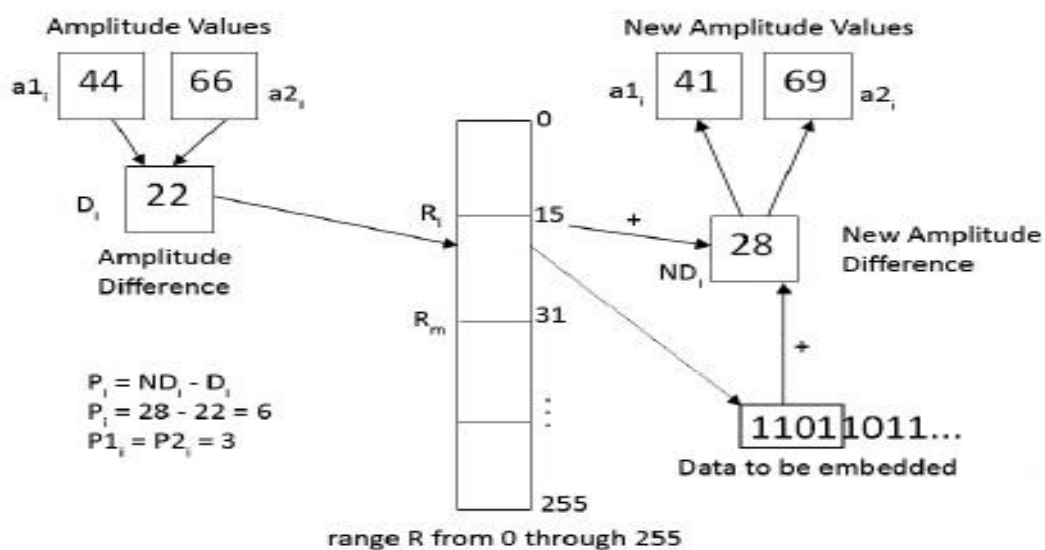
利用了兩個大小差不多的 audio 檔案，計算出這兩個檔案的相對應位置 chunk data 的差值，來進行類似 PVD 的資料隱藏。與 PVD 不同的是，由於 audio 檔案如果只修改每個 byte 的最後兩個 bit，在人的耳朵是分辨不出來的，故此篇論文將區間值設為固定 16，這樣一來就只會修改到最後兩個 bit(兩個檔案各兩個 bit 差值最大可到二的四次方)。

實際例子：

下圖為實際例子，在兩個聲道抽出相對應位置的chunk data後，計算出chunk data的差值($66-44=22$)，取message的 4 個bit(1101)換算成 10 進位後(13)加上區間下限值(15)得到新的差值(28)，最後在平均分配到兩個chunk data裡(41,69)。

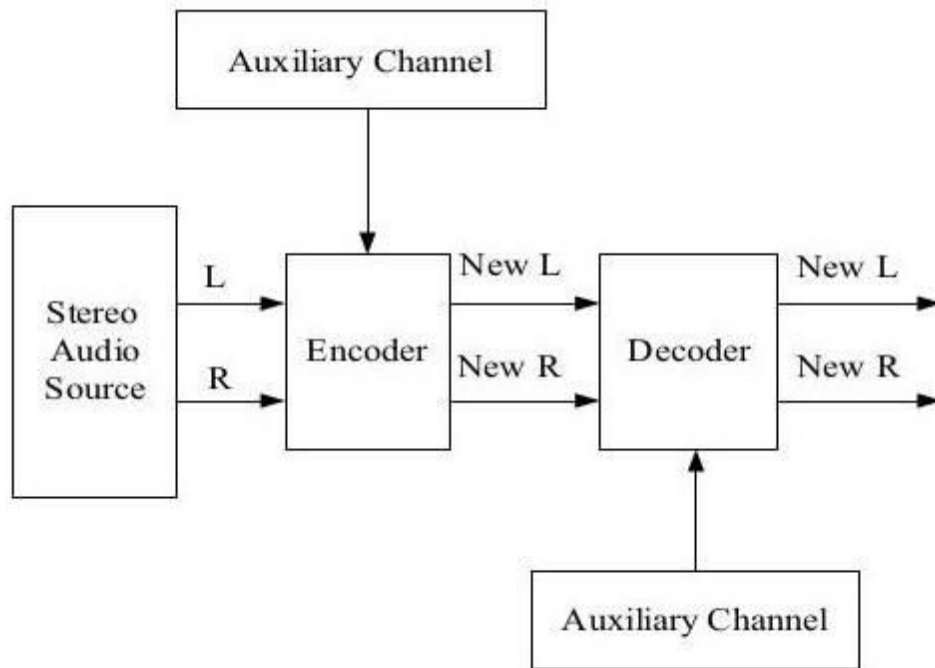
取出：

取出的部份將兩個chunk data相減得到差值 28，扣掉下限值後可得訊息 13。



圖一

在 stereo audio 上的 watermarking[3]：



圖二

圖二為 stereo audio 的 watermark 系統圖。圖中 encoder 決定了 auxiliary channel 能嵌入的資料量，最後在 decoder 會再重新製造出 auxiliary channel。

Interaural Phase Difference(IPD) Relationships：

人耳在分辨聲音的方向是透過聲音來源到兩耳距離不同來區分，如圖三所示， Δl 為聲音來源到左耳的距離， Δr 為聲音來源到右耳的距離， Δd 為兩者的距離差， d 為左耳到右耳的距離， r 為聲音來源到中心點位置。假設目前方位角為 θ ，則：

$$\Delta r^2 = (r \cdot \cos\theta)^2 + (r \cdot \sin\theta - d/2)^2$$

$$\Delta l^2 = (r \cdot \cos \theta)^2 + (r \cdot \sin \theta + d/2)^2$$

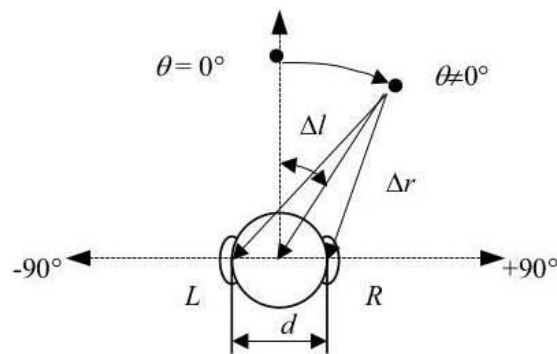
$$\Delta d = |\Delta r - \Delta l|$$

IPD(這裡假設為 Φ)值的算法為：

$$\Phi = \Delta d \cdot (f / c) \cdot 360^\circ$$

有了 Interaural Phase Difference 值，就可以進行以下的

psychoacoustical phase threshold 以及 energy threshold 計算。



圖三

psychoacoustical phase threshold：

psychoacoustical phase threshold 決定資料在什麼時候嵌入，
在頻率(Fourier analysis 提供)滿足下列式子時，可以進行
watermark。

$$\cos\{\text{phase}[X_L(f_i)] - \text{phase}[X_R(f_i)]\} < \cos(-0.003104 \cdot f_i)$$

energy threshold：

在特定頻率依照下列不同情況嵌入相對的值，嵌入時改變的是左邊

的聲道：

$$\text{phase}[X_L(f_i)] = \text{phase}[X_R(f_i)] \rightarrow \text{logical } 0$$

$$\text{phase}[X_L(f_i)] = 1/2 * \text{IPD}_{\max}(f_i) \rightarrow \text{logical } 1$$

$$\text{phase}[X_L(f_i)] \geq \text{IPD}_{\max}(f_i) \rightarrow \text{no encoding}$$

Decoder：

在取資料時，會先經由 psychoacoustical phase threshold 比較與偵測，在利用下列式子決定取出的訊息為多少：

$$| \text{phase}[X_L(f_i)] - \text{phase}[X_R(f)] | \leq r_1 \text{IPD}_{\max}(f) \rightarrow \text{logical } 0$$

$$r_1 \text{IPD}_{\max}(f) < | \text{phase}[X_L(f_i)] - \text{phase}[X_R(f)] | \leq r_2 \text{IPD}_{\max}(f) \rightarrow \text{logical } 1$$

$$| \text{phase}[X_L(f_i)] - \text{phase}[X_R(f)] | > r_2 \text{IPD}_{\max}(f) \rightarrow \text{no data}$$

$$(r_1 = 1/4, r_2 = 3/4)$$

肆、研究方法及步驟

stereo audio 格式與 psychoacoustical phase threshold：

首先要先了解 stereo audio 檔案格式。psychoacoustical phase threshold 需要計算 Interaural Phase Difference 值，所以需要在 stereo audio 格式中找到或計算出 Δd , f , c 這些值。有了 IPD 值就能夠讓 psychoacoustical phase threshold 來篩選能夠藏資訊的區

間，並且統計所有 $\cos\{\text{phase}[X_L(f_i)] - \text{phase}[X_R(f_i)]\}$ 的值，進一步研究能藏的資料量與 PVD 的區間該如何調整。

資料嵌入：

有了 psychoacoustical phase threshold 篩選出能藏資料的部份，PVD 區間也劃分好了之後，開始嵌入資料。在 watermark 裡，它以右邊的聲道當作參考，滿足了條件後修改左邊的聲道。在這邊會將篩選過後的兩邊 chunk data 相減，得出的差值依照 psychoacoustical phase threshold 分析出來的區間做分類，替換成適當藏入量的訊息後再平均分配給兩個聲道，完成藏入。

資料取出：

在資料取出部份，一樣先由 psychoacoustical phase threshold 偵測並比較兩聲道關係，之後將兩個 chunk data 相減扣掉下限值便取出訊息。

伍、預期結果

- 1.能在 stereo audio 上做 PVD 隱藏，能夠順利的藏入跟取出，並且不影響原來聲音的品質。
- 2.增加 capacity，在 A Novel Audio Steganography Scheme using Amplitude Differencing 這篇論文裡它的 capacity 就固定為 $1/4$ ，每 8 個

bit 就藏 2 個 bit，本計劃會改變區間值的設定，預期可以藏的資料量會大於 cover audio 的 $\frac{1}{4}$ 。

本計畫改善了原本需要兩個檔案的缺點，並且套用了 A high capacity watermarking technique for stereo audio 的 psychoacoustical phase threshold 來增加藏密量，期望未來可以在 MPVD、HPVD 等針對隱藏分析的隱藏演算法上發展，讓好的演算法不會只被侷限於影像上。

陸、參考文獻

1. Da-Chun Wu , Wen-Hsiang Tsai , A steganographic method for images by pixel-value differencing
<http://www.sciencedirect.com/science/article/pii/S0167865502004026>
2. Shafi K, Sankaranarayanan A, Prashanth G, Akash Mohan, A Novel Audio Steganography Scheme using Amplitude Differencing, IEEE
http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5714631&url=http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5714631
3. A high capacity watermarking technique for stereo audio
http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1327130&url=http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1327130