科技部補助

大專學生研究計畫研究成果報告

＊ ********* ** ＊
＊ 計 畫 ＊
＊ ： 立體聲的資料隱藏 ＊
＊ 名 稱 ＊
＊ ********* ** ＊

執行計畫學生： 李東岳
學生計畫編號： MOST 103-2815-C-260-024-E
研 究 期 間 ： 103 年 07 月 01 日至 104 年 02 月 28 日止，計 8 個月
指 導 教 授 ： 吳坤熹

處理方式： 本計畫可公開查詢

執 行 單 位： 國立暨南國際大學資訊工程學系（所）

中華民國 104 年 03 月 31 日

摘要

由於資訊科技的發展以及網際網路的便利性，使得資料得以快速傳輸。為了避免機密資料輕易被竊取，資料加密及資訊隱藏的技術成為了重要的研究課題。傳統的資訊隱藏使用最低位元替換藏密法（LSB），優點是很難被人類視覺系統察覺，但是藏密量不多。像素值差異藏密法（PVD）將機密訊息透過更改兩兩相鄰之像素值差異，按規則藏入影像中，大幅增加藏密量。振幅差值藏密法(Amplitude Differencing)是將PVD應用到聲音上，不過每次藏密時需要兩個掩護音訊檔，本文將會改進需要兩個掩護檔的缺點，並且透過Minimum Audible Angle(MAA)的概念讓藏入的資訊更不易被發現。

關鍵字：PVD，MAA，Amplitude Differencing，psychoacoustical phase threshold，Interaural Phase Difference，資訊隱藏。

壹、 前言

在這資訊爆炸的時代，我們能在網路上傳遞任何媒體資料。在傳送資料前，為了不讓有心人士容易竊取我們的資料，可以將資料進行加密的動作，接收者接收時再進行解密。但若有心人士發現此資料是密碼，就會曉得這是有價值的資料，更提高他進行破解密碼的動機，這樣就會造成反效果。因此現今有了資料隱藏這項技術，它能夠將機密資訊隱藏在媒體檔案裡面，這樣一來只需要傳送一個看似正常的媒體檔案，就能夠將機密資訊傳送出去。

Bender等學者於1996年所提出的最低位元替換藏密法（Least Significant Bit, LSB）[1]，優點是藏密後很難被人類視覺系統察覺異狀，缺點是藏密量最多僅1bpp（bit per pixel）。此外，此LSB藏密技術容易被Regular-Singular（RS）[2]偵密技術所察覺。Wu與Tsai學者於2003年提出的像素值差異演算法(Pixel Value Differencing, PVD)[3]，將祕密訊息透過更改兩兩相鄰之像素值差異，按規則藏入影像中，此方法不但增加藏密量，並且可避免被RS偵密技術所察覺。在2010年Shafi等學者提出了Amplitude Differencing[4]，將PVD演算法應用在聲音上，透過兩個掩護音訊檔來進行差值藏密，缺點是每次藏密時都需要兩個掩護檔案，本文將只需要一個掩護音訊檔便能進行Amplitude Differencing，並且利用1958年Mills所提出的Minimum Audible Angle(MAA)[5]以及psychoacoustical phase threshold[6]來提升PSNR，降低對掩護資料的干擾。

貳、 研究目的

在 2010 年 Shafi 等學者提出了 Amplitude Differencing[4]，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。

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。
3. 透過 psychoacoustical phase threshold 來增加 PSNR

參、 文獻探討

1.像素值差異藏密法

在2003年Wu學者與Tsai學者所提出的像素值差異藏密法(PVD)[3]中，機密訊息是藏在影像中相鄰像素的差值裡。如果將機密訊息直接替換成差值，那麼圖像隱藏前後的差異用肉眼即可辨識。因此我們需要將像素差值做分類，讓差值只在該特定區間內做替換，如下表所示：

表一依照差值大小分為六個區間

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

表一中R_k為相鄰像素差值，u_k與l_k為該區間的最大值與最小值，$n = \log_2(u_k - l_k + 1)$ ，為可藏入的位元數，以下為PVD演算法：

步驟一：經由式子(1)計算出影像中相鄰像素P_i與P_{i+1}差值d。

$$d = | P_i - P_{i+1} | \quad (1)$$

步驟二：判斷d屬於哪個區間，向機密訊息取n個位元後，將二進位的

值轉為十進位計算出b。

步驟三：依照式子(2)計算出d’：

$$d' = \begin{cases} l_k + b & \text{for } P_i - P_{i+1} \geq 0 \\ -(l_k + b) & \text{for } P_i - P_{i+1} < 0 \end{cases} \quad (2)$$

步驟四：依照式子(3)計算出新的像素值P'_i與P'_{i+1}，$m = |d' - (P_i - P_{i+1})|$。

$$(P'_i, P'_{i+1}) = \begin{cases} \left(P_i + \left\lceil \frac{m}{2} \right\rceil, P_{i+1} - \left\lfloor \frac{m}{2} \right\rfloor\right), \text{if } P_i \geq P_{i+1} \text{ and} d' > d \\ \left(P_i - \left\lfloor \frac{m}{2} \right\rfloor, P_{i+1} + \left\lceil \frac{m}{2} \right\rceil\right), \text{if } P_i < P_{i+1} \text{ and} d' > d \\ \left(P_i - \left\lceil \frac{m}{2} \right\rceil, P_{i+1} + \left\lfloor \frac{m}{2} \right\rfloor\right), \text{if } P_i \geq P_{i+1} \text{ and} d' \leq d \\ \left(P_i + \left\lceil \frac{m}{2} \right\rceil, P_{i+1} - \left\lfloor \frac{m}{2} \right\rfloor\right), \text{if } P_i < P_{i+1} \text{ and} d' \leq d \end{cases} \quad (3)$$

要擷取訊息時，依照式子(4)計算出d’後，判斷d’屬於哪個區間，計算出$d' - l_k$便可得到訊息。

$$d' = |P'_i - P'_{i+1}| \quad (4)$$

例子：假設$P_i = 50$，$P_{i+1} = 30$，$d = 20 \in R_3$，機密訊息4個位元假設為$(1010)_2$，$b = 10$，$d' = 26$，$P'_i = 50 + 3 = 53$，$P'_{i+1} = 30 - 3 = 27$，隱藏完成。擷取訊息時，計算出$d' = 53 - 27 = 26$，$26 - 16 = (10)_{10} = (1010)_2$得到訊息。

2. Amplitude Differencing

2010年Shafi等學者[4]將PVD演算法應用在聲音上，掩護資料為兩個大小相近的音訊檔案，而機密訊息藏於兩個檔案相對應位置chunk data的差值，做法如下：

步驟一：計算出兩個檔案A、D相對應位置A_i、D_i的差值$d = |A_i - D_i|$。

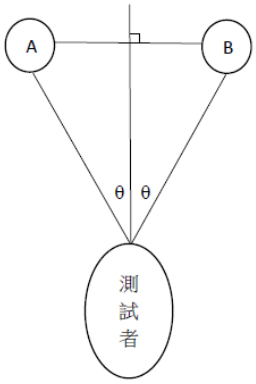
步驟二：將PVD的區間值改為固定16，向機密訊息取4個位元，把取出來二進位的值轉為十進位。

步驟三與步驟四跟PVD演算法相同，將像素值P_i、P_{i+1}改為A_i、D_i，P'_i、P'_{i+1}改為A'_i、D'_i。

例子：假設$A_i = 50$，$D_i = 30$，$d = 20$ ，為16~31區間，機密訊息4個位元假設為$(1010)_2$，$b = 10$，$d' = 26$，$A'_i = 50 + 3 = 53$，$D'_i = 30 - 3 = 27$，隱藏完成。擷取訊息時，計算出$d' = 53 - 27 = 26$，$26 - 16 = (10)_{10} = (1010)_2$得到訊息。

3.Interaural Phase Difference and Minimum Audible Angle

在1958年Mills提出了Minimum Audible Angle(MAA)[5]，關於人耳對於聲音的方位辨識。實驗如圖一，以測試者為中心，將聲源A、B分別放在測試者左右兩側，讓A、B兩點的中點與測試者所形成的直線與直線AB垂直。實驗時讓聲源AB隨機發出聲音，讓測試者分辨是在左邊或是右邊，θ為變數，測試發現θ在1˚以內時測試者無法分辨聲音來源為左邊或是右邊。



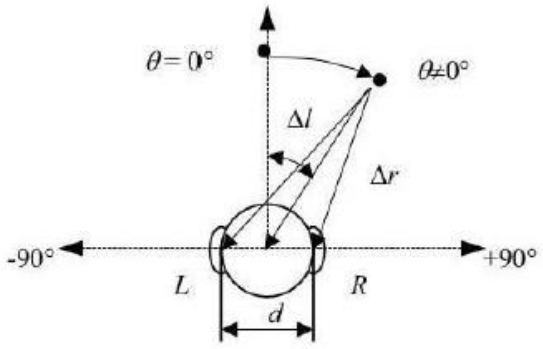
圖一　MAA實驗

利用MAA概念，可以計算出interaural phase difference(IPD)最大值，IPD的意義為聲源到左右耳的相位角差值，也就是計算差了多少個波長。首先計算出聲源到左耳與到右耳的距離差值Δd，圖二為聲源到兩耳的示意圖，Δl為聲源到左耳的距離，Δr為聲源到右耳的距離，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：

$$\Delta r = \sqrt{(r\cos\theta)^2 + (r\sin\theta - \frac{d}{2})^2} \quad (5)$$

$$\Delta l = \sqrt{(r\cos\theta)^2 + (r\sin\theta + \frac{d}{2})^2} \quad (6)$$

$$\Delta d = |\Delta r - \Delta l| \quad (7)$$



圖二azimuth plane

接下來計算IPD，IPD(Φ)可表示為頻率f的函式：

$$\Phi = \Delta d \times \frac{f}{c} \times 2\pi \ \text{or} \Delta d \times \frac{f}{c} \times 360^\circ \quad (8)$$

其中c為聲音在空氣中的傳送速率，約344m/s，由於 $v = f \times \lambda$，故$\Delta d \times \frac{f}{c}$可視為$\frac{\Delta d}{\lambda}$，也就是聲源到兩耳的波程差，最後乘上2π或360°變成相位角。

利用MAA概念，當$\theta \leq 1^\circ$時，Φ計算出的最大值Φ_{max}為$-3.104 \times 10^{-3} \times f$，也就是說，左右耳所接收到特定頻率的波的相位角差值小於Φ_{max}的話，人耳分辨不出聲音來自左方或右方，此時更改資料不易被發現，便能夠嵌入資料。透過傅立葉轉換能夠計算得到某頻率的相位角，經由psychoacoustical phase threshold式子(9)可以得知該頻率能否進行資料隱藏[6]：

$$\cos\{\text{phase}[X_R(f_i)] - \text{phase}[X_L(f_i)]\} < \cos(-3.104 \times 10^{-3} \times f_i) \quad (9)$$

$\text{phase}[X_R(f_i)]$為右聲道在頻率為f_i的相位角，傅立葉轉換會得到頻率為f_i的複數a + bi，式子(10)的φ即為此頻率的相位角：

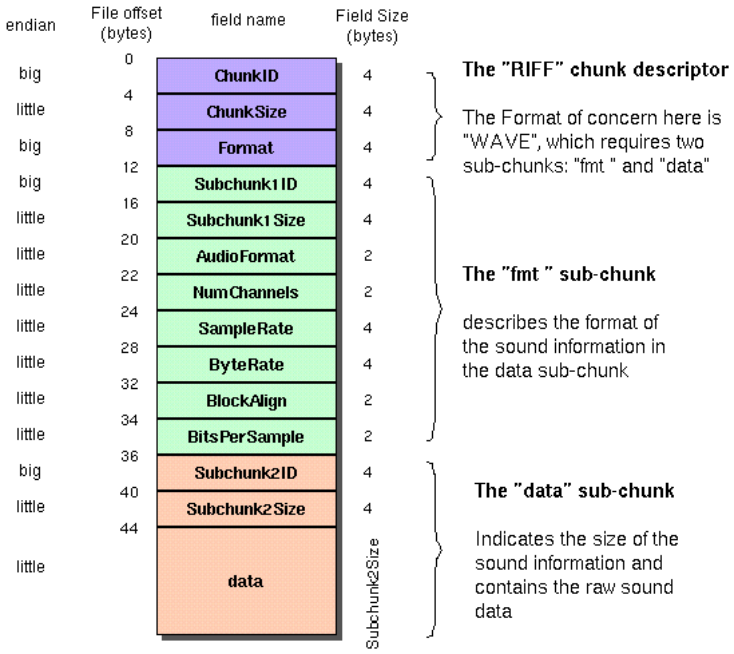
$$\varphi = \tan^{-1}(\frac{b}{a}) \quad (10)$$

肆、 研究方法

本研究實做環境如下：

- 作業系統：Ubuntu 12.04
- 語言：C++
- 編譯器：gcc 4.6.3
- 引用 library：libsndfile

要在雙聲道進行資料隱藏，首先要先了解支援雙聲道格式的檔案，本研究採用 WAV 檔，他的 chunk data 是沒有被壓縮的，處理起來較方便，圖三為 WAV 格式，雙聲道儲存格式為每兩個 byte 為一組，第一個 byte 為左聲道振幅，第二個 byte 為右聲道振幅。



圖三　WAV 格式[7]

接下來為本研究的資料隱藏方法：

步驟一：利用 libsndfile 取出 WAV 檔案的左右聲道資訊。

步驟二：將左聲道的資料每 8 個一組經過離散傅立葉轉換，右聲道的資料每 8 個一組經過傅立葉轉換，便能得到左右聲道相對應頻率的複數。

步驟三：計算出可藏入的位置。透過式子(10)計算出相位角，將頻率及其相位角代入 psychoacoustical phase threshold(式子 9)，若符合則繼續步驟四，不符合則繼續此步驟，直到隱藏資料結束。

步驟四：將符合 psychoacoustical phase threshold 的頻率取出左右聲道頻域的實部係數，進行 Amplitude Differencing 資料隱藏。

例子：在這邊以 4 組聲道資訊來做資訊隱藏。假設從 WAV 取出了音訊資料，左聲道音訊資料為 20、30、40、50，右聲道音訊資料為 90、80、70、60，需要被隱藏的資訊串流為 1100101001。

離散傅立葉公式(DFT)：

$$y_k = \sum_{j=0}^{n-1} x_j e^{-2\pi ikj/n} \qquad (11)$$

(x_j 為時域資訊，y_k 為頻域資訊)

左聲道進行傅立葉轉換：

$x_0 = 20 \qquad x_1 = 30 \qquad x_2 = 40 \qquad x_3 = 50$

$y_0 = 20 + 30 + 40 + 50 = 140$

$y_1 = 20 + 30e^{-2\pi i/4} + 40e^{-4\pi i/4} + 50e^{-6\pi i/4} = -20 + 20i$

$y_2 = 20 + 30e^{-4\pi i/4} + 40e^{-4\pi i*2/4} + 50e^{-4\pi i*3/4} = -20$

$y_3 = 20 + 30e^{-6\pi i/2} + 40e^{-6\pi i*2/4} + 50e^{-6\pi i*3/4} = -20 -20i$

右聲道進行傅立葉轉換：

$x_0 = 90 \qquad x_1 = 80 \qquad x_2 = 70 \qquad x_3 = 60$

$y_0 = 90 + 80 + 70 + 60 = 300$

$y_1 = 90 – 80i - 70 + 60i = 20 – 20i$

$y_2 = 90 – 80 + 70 – 60 = 20$

$y_3 = 90 + 80i – 70 – 60i = 20 + 20i$

計算各個頻率的相位角：

左聲道相位角：

$y_0 = 0^\circ$

$y_1 = \tan^{-1}\frac{20}{-20} = 135^\circ$

$y_2 = 0^\circ$

$y_3 = \tan^{-1}\frac{-20}{-20} = 45^\circ$

右聲道相位角：

$y_0 = 0^\circ$

$y_1 = \tan^{-1}\frac{-20}{20} = -45^\circ$

$y_2 = 0^\circ$

$y_3 = \tan^{-1}\frac{20}{20} = 45^\circ$

PsychoacousticalPhase Threshold：

f_1：$\cos[135^\circ - (-45^\circ)] < \cos\left(-3.104 \times 10^{-3} \times \frac{1}{4}\right)$

f_2：$\cos(0^\circ - 0^\circ) < \cos\left(-3.104 \times 10^{-3} \times \frac{2}{4}\right)$

f_3：$\cos(45^\circ - 45^\circ) < \cos(-3.104 \times 10^{-3} \times \frac{3}{4})$

這邊三個頻率都能藏資訊，取 f_2 來進行說明。透過 DFT 後取出實部係數，左聲道為-20 右聲道為 20，藏入資訊取 11001，十進位為 25，經過 Amplitude Differencing 後新的左聲道頻域係數為-13，右聲道頻域係數為 12，由於隱藏後相位角一樣是 0°，不影響 PsychoacousticalPhase Threshold，隱藏完成。

伍、 結果與討論

由於在實做離散傅立葉轉換時發現了幾個問題，導致藏入的資訊無法正確的取出，表二為左右聲道各取 8 份資料做離散傅立葉轉換前後數據：

表二 DFT 前後比較表

f	轉換前(左)	轉換後(左)	轉換前(右)	轉換後(右)
0	250	(766,0)	247	(797,0)
1	3	(155.138,194.655)	14	(176.957,246.392)
2	60	(149,-1.00059)	40	(125,21.9991)
3	42	(266.86,154.655)	42	(225.041,70.3914)
4	39	(92,-0.00287914)	46	(125, -0.00207511)
5	166	(266.864,-154.656)	143	(225.043, -70.3928)
6	80	(149.001,0.998238)	128	(125.001, -22.0026)
7	126	(155.145,-194.656)	137	(176.967, -246.392)

根據表二所出現的問題以下分為兩點討論：

1. 離散傅立葉轉換後有機會出現浮點數：

PVD 演算法裡掩護資料皆假設為整數，若出現浮點數，在新舊差值平均分配給兩個掩護值時則會發生問題。

例子：取頻率 3 做資料隱藏，左聲道為 266.86，右聲道為 225.041，假設藏入值為 41，則新舊差值為 |41 – (|266.86 – 225.041|)| = 0.819，根據公式新的係數會變成左聲道 265.86、右聲道 225.041，此時差值不是藏入的 41。

2. 誤差值：

離散傅立葉轉換得到的係數會有誤差值，表三為表二轉換後係數經過 IDFT 得到的數值。

表三 IDFT 值與原始資料

f	左聲道	IDFT(左)	右聲道	IDFT(右)
0	250	(250.001,0)	247	(247.001,0)
1	3	(64.5003,61.5003)	14	(75.5005,61.5004)
2	60	(69.9999,10.0001)	40	(83.9998,44)
3	42	(103.999,61.9998)	42	(92.4993,50.4996)
4	39	(38.9992, -0.000491322)	46	(45.9992, -0.00050787)
5	166	(104,-62.0002)	143	(92.4997, -50.5002)
6	80	(70.0001, -10.0001)	128	(84.0002,-44)
7	126	(64.5014, -61.4999)	137	(75.5015, -61.4997)

原本經過 IDFT 後得到的值理論上要與原始 WAV 檔案裡的值一樣，但是因為 DFT 出現了誤差值，導致再經過 IDFT 時所得到的值與預期不同，而且時域也出現了虛部，單純的經過 DFT 與 IDFT 便會發生干擾，此時若再藏入資訊則無法正確取出。

根據上述兩點問題，以下為可能的解決方法：

2001 年 Faisal Alturki 等學者所提出的隱藏演算法中，也是將資料藏在 DFT 後的頻域裡，特別的是，他在藏資料前會先經過量化的步驟，來避免 DFT 所產生的誤差值，以下為 Faisal Alturki 學者在頻域上隱藏的演算法：

步驟一：決定量化的基準Δ，之後會將掩護值設定為Δ的倍數。

步驟二：若欲藏入的值為 1，將掩護係數替換為Δ的偶數倍，並且選擇最接近原本係數的值；若欲藏入的值為 0，則將掩護係數替換為Δ的奇數倍。

例子：掩護係數為 A = 169、B = 520，Δ為 100，A 藏入 bit 1，A = 2 ×Δ = 200 ，B 藏入 bit 0，B = 5 ×Δ = 500，藏入完成。

取出資料時，假設 DFT 後的係數為 c'，擷取資訊為 d，則：

$$d = \begin{cases} 1, \text{ if } \left\lfloor \frac{c'}{\Delta} \right\rfloor \text{iseven} \\ 0, \text{ if } \left\lfloor \frac{c'}{\Delta} \right\rfloor \text{isodd} \end{cases} \qquad (12)$$

這麼一來 DFT 誤差值的影響取決於Δ，若 DFT 後的係數誤差範圍在±Δ以內，就不會影響到藏入的值。

而在時域上出現虛數則是因為 DFT 的誤差值以及藏入值所影響，DFT 之後的係數會有對稱性的規則，假設 $x_0, x_1, \ldots x_{n-1}$ 為實域上的數值，並且沒有虛部，$X_0, X_1, \ldots X_{n-1}$ 為 DFT 後頻域的複數，則 $X_{n-k} = X_k^*$，X^*標記為 X 的共軛複數，利用此一特點，將 DFT 後相對應的係數改為共軛複數，則 IDFT 回時域時便不會產生虛部。

未來發展

由於 DFT 的誤差值使得 PVD 無法正確的取出藏入資訊，若要透過量化的手法來隱藏資料，則 1 個 byte 中只能藏 1 個 bit，藏密量大幅縮減，未來可能的研究方向為：利用量化的概念實現類似 PVD 的演算法，由於 PVD 是以 bit 為單位來進行資料隱藏，若能夠讓藏入的位元在特定誤差範圍都不被影響，則有機會提升量化隱藏法的藏密量。

陸、 參考文獻

[1] W. Bender, D. Gruhl, N. Morimoto, and Aiguo Lu, "Techniques for data hiding," IBM Systems Journal,Vol.35, No. 3-4, pp. 313-336, 1996.

[2] J. Fridrich, M. Goljan, and R. Du, "Detecting LSB Steganography in Color and Gray-scale Images," IEEE Multimedia, Vol.8, No.4, pp. 22-28, OctDec 2001.D. C

[3] Wu, and W. H. Tsai, "A steganographic method for images by pixel-value differencing," Pattern Recognition Letters, Vol. 24, No. 9-10, pp. 1613-1626, 2003.

[4] K. Shafi, A. Sankaranarayanan, G. Prashanth, A. Mohan, "A Novel Audio

Steganography Scheme Using Amplitude Differencing", IEEE International Conference on Trends in Information Science and Computing, December 2010.

[5] A. W. Mills, "On the minimum audible angle". Journal of the Acoustical Society of America, pp. 2031-2041, 1958.

[6] I. Alexander, S. Michael, "AuxiliaryChannel MaskingIn An Audio Signal", Patent Application Publication, pp. 1-7, Apr. 2006.

[7] WAV format, http://knowledge-teaching.blogspot.tw/2013/09/wav.html