半盲目型浮水印植基自我參照影像

李金鳳
朝陽科技大學
資訊管理學系 副教授
lcf@cyut.edu.tw

陳星琳
朝陽科技大學
資訊科技研究所 研究生
hsingling@cyut.edu.tw

楊宗家
朝陽科技大學
資訊管理學系 研究生
s9514626@cyut.edu.tw

***摘要*—一般浮水印技術常需要原始影像、浮水印與秘密金鑰三者的配合才能進行偵測或驗證，因此需要耗費大量額外的空間來儲存或傳輸相關資訊。近年來半盲目浮水印技術逐漸被重視與發展，僅需原始浮水印與秘密金鑰便能進行浮水印的偵測與驗證為其主要特色，因此可大幅減少額外儲存空間的需求或傳輸的頻寬。本文提出一個"半盲目型浮水印植基自我參照影像"的浮水印技術，首先會利用離散小波轉換技術產生一個原始影像之參照影像，接著將原始影像與參照影像進行比對，找出原始影像中較佳的浮水印嵌入位置，再配合恰可察覺失真(Just Noticeable Distortion, JND)的技術進行浮水印嵌入。本方法僅需少量的額外資訊與浮水印便能進行浮水印偵測與驗證，此外根據實驗結果證明，相較於 Liu 等學者所提出的半盲目浮水印技術，本方法改善了其所存在安全性漏洞的問題，同時也能維持相似效能的強韌性。**

***關鍵詞*—半盲目浮水印、自我嵌入、自我參照影像、恰可察覺失真**

一、前言

由於網際網路快速的發展以及通訊技術不斷的改善，近幾年有越來越多人利用網路的便利性來呈現個人創作或進行各種宣傳與行銷。隨著網路流通的便利性及資料內容數位化，使得這些數位化資料(例如：文字、圖片、影像…等)利用網際網路進行儲存、傳送，甚至有組織性地利用網路來進行知識及經驗的分享(例如以維基百科平台)愈見普及，仍而處於在開放式的網際網路，數位資料容易在未經合法授權下遭受不法的複製、截取與竄改，進而衍生出許多智慧財產權(Intellectual Property Right, IPR)侵權的問題發生。

因此為避免非法侵權的行為，目前大多採用浮水印的技術來進行數位版權的保護，以數位影像為例，便是利用公司商標、文字或特殊圖形與符號視為浮水印來進行嵌入到原始影像中，進而產生浮水印影像，每當有侵權問題發生時，便將有爭議的浮水印影像進行浮水印的偵測與驗證，利用偵測與驗證的結果來判別其版權所屬，以遏阻非法侵權行為。

二、文獻探討

數位浮水印種類可以依視覺感官、強韌程度、嵌入方式與取出方式進行區分：

1.依視覺感官：

可分為可視浮水印(Visible watermark)與不可視浮水印(Invisible watermark)兩類，可視浮水印可被人類視覺輕易察覺它的存在，常用於遏阻的用途。例如：鈔票上的防偽線與特定視覺角度所呈現出特殊的圖案。不可視浮水印技術常用於版權宣告與授權之用，該類技術當浮水印嵌入媒體後，人類的視覺是無法輕易發現。

2.依強韌程度：

可區分為強韌型(Robust)浮水印與易碎型

(Fragile)浮水印，強韌型浮水印是指當浮水印影像遭受到攻擊後，仍能取出或辨識出所嵌入的浮水印；而易碎型浮水印技術當浮水印影像遭受到攻擊後，所嵌入的浮水印將無法被取出。

3.依嵌入方式：

可區分為空間域(Spatial domain)浮水印技術[4][6]與頻率域浮水印技術(Frequency domain) [1][8]，空間域浮水印的技術其藏量較高與演算法複雜度較低為其優點，然而強韌性不足為主要存在的問題，反觀頻率域雖藏量較低、演算法複雜度高，但高強韌性為其優點。在實務應用上為能夠在被攻擊後還能取出所嵌入的浮水印進行識別，目前大多數的浮水印技術是採用頻率域的方式進行嵌入。

4.依取出方式：

可分成非盲目(Non-blind)、半盲目(Semi-blind)與盲目(Bind) 等三類浮水印技術，當欲取出所嵌入的浮水印必需原始影像、浮水印與秘密金鑰三者配合者，此類方法稱之為非盲目型浮水印技術。然而僅需秘密金鑰便能取出所嵌入的浮水印，此類技術稱之為盲目型浮水印技術。介於兩者之間需用浮水印與秘密金鑰配合才能取出浮水印者稱之半盲目型浮水印技術。三者間以非盲目浮水印技術擁有最佳的強韌性，然而此類的方法需花費許多空間來儲存或傳輸相關資訊。而盲目型浮水印技術，僅需要秘密金鑰輔助便能取出浮水印，但此類技術強韌性較低，容易因遭受輕微的訊號干擾便會造成浮水印判讀錯誤的問題產生。近年來有許多半盲目型浮水印的研究被提出，主要的原因為該類技術僅需浮水印與秘密金鑰便能進行浮水印的偵測與驗證，因此不用浪費大量空間來儲存或傳輸原圖且能維持一定程度的強韌性。

Liu 等學者於 2005 年提出一個利用自我參照影像(Self-reference image)的強韌性浮水印的方法[5]，他們的方法是屬於半盲目型浮水印技術，因此不需原始影像協助下便能進行浮水印的偵測與驗證，經他們的實驗證明所提出的方法能有效抵擋剪裁攻擊、模糊化攻擊、亮度調整攻擊、高斯雜訊攻擊與銳利化攻擊，該方法的嵌入流程簡述如下：

首先將原始影像 I 進行一階離散小波轉換(Discrete Wavelet Transformation, DWT)，亦即將空間域的像素值轉換成頻率域的係數值，然後將高低頻(HL_1)、低高頻(LH_1)及高高頻(HH_1)的係數值設定為 0，再利用反轉函式 IDWT 的轉換將係數值轉換成像素值後，便能獲得一張與原始影像大小相同的自我參照影像 R。接著計算兩張影像像素間的差值，假設該差值符合 $s < |I(i,j) - R(i,j)| < t$，則利用一個序列 idx 進行記錄該像素位置座標，最後使用亂數的方式選取在 idx 中的位置，利用公式(1)將浮水印資料進行嵌入。

$$I'(i,j) = \begin{cases} R(idx(i,j)) + \alpha \text{ if } w = +1 \text{ and } s < I(idx(i,j)) - R(idx(i,j)) < t, \\ R(idx(i,j)) - \alpha \text{ if } w = -1 \text{ and } s < R(idx(i,j)) - I(idx(i,j)) < t. \end{cases} \quad (1)$$

其中 α =round($(s+t)/2$), 浮水印 $w \in \{+1, -1\}$。

該方法浮水印偵測與驗證流程如下簡述：首先將浮水印影像 I' 進行一階離散小波轉換，然後將將 HL_1, LH_1, HH_1 的係數值設定為 0，再利用反轉函式 IDWT 轉換獲得一張自我參照影像 R'。接著根據 idx 序列內容，找到特定位置座標，將特定位置座標的像素值與位於自我參照影像相同座標位置的像素值進行相減，利用公式(2)便能取出先前所嵌入的浮水印內容。

$$w = \begin{cases} 1, \text{ if } I'(idx(i,j)) \geq R'(idx(i,j)), \\ -1, \text{ if } I'(idx(i,j)) < R'(idx(i,j)). \end{cases} \quad (2)$$

然而 Ting 等學者[7] 提出在上述的方法中存在一個安全性的漏洞，根據 Ting 等學者指出 Liu 等學者的方法中，由於浮水印嵌入的變異性過於一致性，亦即僅$\pm\alpha$的像素值變動率，所以容易尋找出浮水印的嵌入位置，故浮水印易遭受破壞，因此強韌性便會降低，Ting 等學者所提出的攻擊方法簡述如下：

利用所獲得的浮水印影像I'，經相同的 DWT 轉換後，將HL_1, LH_1, HH_1的係數值設定為 0，再執行 IDWT，便能獲得一張自我參照影像R'，將整張浮水印影像I'進行掃描，檢查是否有$I'(i,j)=R'(i,j)+\alpha$與$I'(i,j)=R'(i,j)-\alpha$的情況存在，便能了解所嵌入的浮水印內容為何? 接下來 Ting 等學者再針對這些嵌入位置進行反相內容竄改，舉例來說，假如$I'(i,j)=R'(i,j)+\alpha$，則該像素內容將被減去2α；換言之竄改後的$I'(i,j)=R'(i,j)-\alpha$，因此取出浮水印時會將–1 取代原本浮水印內容+1 形成反相內容的浮水印結果。

有鑑於此，本文中所提出的浮水印技術便是針對改善此一安全性的問題為出發點，並利用 JND[2][3]的技術來加強浮水印影像的視覺品質。實驗結果證明我們所提出的方法，不但能維持浮水印的強韌性與 Liu 等學者的方法相同外，並將其安全性漏洞的問題加以解決。

三、本研究方法

在介紹本方法之前，本方法相關所使用到的符號及其意義如下所示：

I：8 位元灰階原始影像，大小為$M\times N$。

I'：8 位元灰階浮水印影像。

R：參照影像，大小為$M\times N$。

W：浮水印，經由隨機亂數所產生的一的組數列，$w_z\in\{-1,1\}$，$z=(1,2,\ldots,L)$且L為W的大小。

$idl_z(i,j)$：位置序列亦即浮水印嵌入的位置，視同秘密金鑰，其中$1\le i\le M$，$1\le i\le N$。

D：浮水印影像之像素值。

(一)嵌入浮水印

本研究整體浮水印嵌入流程圖如圖 1 而相關嵌入步驟如下所述：

步驟 1：將I進行一階 DWT 轉換。

步驟 2：將高低頻(HL_1)、低高頻(LH_1)及高高頻(HH_1)清除 0 後，再進行 IDWT 轉換，則得到R。

步驟 3：計算$I(i,j)$與$R(i,j)$的各像素之差值，並暫時記錄像素位置$idl_z(i,j)$，假如該差值有符合$s<|I(i,j)-R(i,j)|<t$，其中s與$t\in Z^+$。

步驟 4：利用公式(3)來計算出$I(i,j)$之 JND 值。

步驟 5：隨機選擇可嵌入位置$idl(i,j)$，並記錄該$idl_z(i,j)$值，配合步驟 4 所產生的 JND 值，將浮水印以公式(4)進行嵌入。

$$\mathrm{JND}(I(i,j))=\begin{cases}T_0\times(1-(I(i,j)/127)^{1/2})+3 & \text{if } I(i,j)\le 127,\\ \gamma\times(I(i,j)-127)+3 & \text{if } I(i,j)>127.\end{cases} \quad (3)$$

$$D(i,j)=R(idl_z(i,j))+w_z\times\mathrm{JND}(i,j) \quad (4)$$

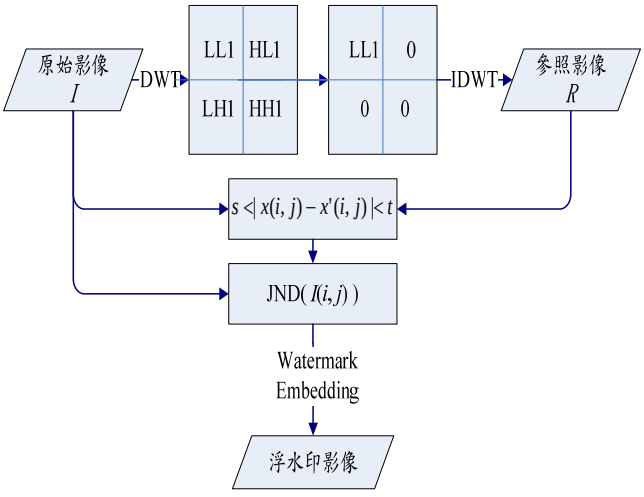


圖 1 浮水印嵌入流程圖

利用下述的例子針對本研究所提出的嵌入流程，進行更詳細的說明：

假設給於一個 8×8 的原始影像如圖 2 所示且浮水印 1 與-1 欲被嵌入其中，並設定門檻值 $s=4, t=10$, JND 參數為 $T_0=17$ 與 $\gamma=3/128$。首先針對原始影像進行一階 DWT 轉換後，再將 LH_1、HL_1 及 HH_1 等三個頻帶區域的頻率係數值全部清為 0，如圖 3 所示，接著執行 IDWT 便能獲得參照影像 R 如圖 4 所示。

100	105	120	135	140	135	135	164
125	130	124	101	85	64	90	100
130	146	158	150	140	135	105	110
125	124	101	85	130	146	158	140
100	110	120	135	140	135	164	132
110	85	140	134	95	105	126	120
130	146	158	124	101	85	90	130
135	140	140	134	130	146	158	160

圖 2 原始影像

460	480	424	489	0	0	0	0
525	494	551	513	0	0	0	0
405	529	475	542	0	0	0	0
551	556	462	538	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

圖 3 一階小波影像

接下來將 $I(i, j)$ 與 $R(i, j)$ 進行差值計算，並暫時記錄像素位置 $idl_2(i, j)$ 如圖 5 灰色區塊，假如該差值大於 4 且小於 10。目前灰色區塊代表可嵌入浮水印的候選像素位置。

115	115	120	120	106	106	122	122
115	115	120	120	106	106	122	122
131	131	123	123	138	138	128	128
131	131	123	123	138	138	128	128
101	101	132	132	119	119	128	128
101	101	132	132	119	119	128	128
138	138	139	139	115	115	135	135
138	138	139	139	115	115	135	135

圖 4 參照影像 R

115	115	120	120	106	106	122	122
115	115	120	120	106	106	122	122
131	131	123	123	138	138	128	128
131	131	123	123	138	138	128	128
101	101	132	132	119	119	128	128
101	101	132	132	119	119	128	128
138	138	139	139	115	115	135	135
138	138	139	139	115	115	135	135

圖 5 可嵌入浮水印的候選像素

100	105	120	135	140	135	135	164
125	130	123	101	85	64	90	100
130	146	158	150	140	135	105	110
125	124	101	85	135	146	158	140
100	110	120	135	140	135	164	132
110	85	140	134	95	105	126	120
130	146	158	124	101	85	90	130
135	140	140	134	130	146	158	160

圖 6 浮水印影像

然後隨機選取兩個候選像素位置，在此範例中為 $idl_1(2,3)$ 與 $idl_2(4,5)$，配合該像素的值，亦即 $JND(2,3)=3$, $JND(4,5)=3$，再利用公式(4)來進行浮水印嵌入，便可計算出偽裝像素值分別為 123 與 135，為了日後能取出所嵌入的浮水印，相關 $idl_2(i, j)$ 必需記錄起來，最終浮水印影像 I' 可被獲得如圖 6 所示。

$D(2,3) = R(idl_1(2,3)) + w_1 \times JND(2,3)$

=120+1×3=123.

$D(4,5) = R(idl_2(4,5)) + w_2 \times JND(4,5)$

=138–1×3=135.

(二)取出浮水印

在不需要原始影像的協助之下，本方法可直接將浮水印進行取出，相關取出的步驟如下：

步驟 1：將浮水印影像 I' 進行一 DWT 段的轉換。

步驟 2：將高低頻(HL_1)、低高頻(LH_1)及高高頻(HH_1)清除為 0 後，再執行 IDWT 便可獲得參照影像 R'。

步驟 3：根據先前記錄 $idl(i, j)$ 位置，利用公式(5)進行比對 I' 與 R' 後，便可取出先前所嵌入的浮水印內容。

$$w_z = \begin{cases} 1, \text{ if } I'(idl_z(i,j)) > R'(idl_z(i,j)), \\ -1, \text{ if } I'(idl_z(i,j)) < R'(idl_z(i,j)). \end{cases} \quad (5)$$

步驟 4：最後將取出的浮水印，利用公式(6)來進行類似度(Similarity)的比較，來證明浮水印的強韌性，最終類似度的值將被量化在±1 的區間內。

$$SV_z = \begin{Bmatrix} 1, & w_z = w_z' \\ -1, & w_z \neq w_z' \end{Bmatrix},$$

$$S = Sim(W, W') = \sum_{z=1}^{L} SV_z / L. \quad (6)$$

舉例來說，以圖 6 浮水印影像為例，進行一階 DWT 轉換後，將高低頻(HL_1)、低高頻(LH_1)及高高頻(HH_1)清除為 0，再執行 IDWT 轉換後，便可獲得 R' 如圖 7 所示。接著藉由先前記錄 $idl(i, j)$ 的位置亦即(2,3)與(4,5)，利用公式(5)進行浮水印的取出如下所示：

$D(2,3) > R'(2,3)$=123>119　　$w_1 = 1$

$D(4,5) < R'(4,5)$ =135<137　　$w_2 = -1$

115	115	119	119	106	106	122	122
115	115	119	119	106	106	122	122
131	131	123	123	137	137	128	128
131	131	123	123	137	137	128	128
101	101	132	132	118	118	135	135
101	101	132	132	118	118	135	135
137	137	139	139	115	115	134	134
137	137	139	139	115	115	134	134

圖 7 浮水印之參照影像

四、實驗結果

本實驗中利用一把秘密金鑰以亂數的方式產生 200 組獨立不重覆的浮水印內容，每一組浮水印由 1000 個 +1 與 –1 所組成。並將 200 組浮水印分別嵌入兩張測試影像 Lena 與 Baboon 其大小為 256×256 進行實驗測試，如圖 8 所示：

(a)Lena　　(b)Baboon

圖 8 測試影像

因此我們可各獲得 200 張浮水印影像 Lena 與 Baboon，其各浮水印影像被由 1 至 200 進行編號。相關實驗參數經反覆測試後較佳參數值設定如下：門檻值 s 與 t 分別設定為 s =4, t =6，而 JND 的參數值設定為 T_0 =17, γ =3/128。此外利用 PNSR(Peak Signal-to- Ratio)來評量原始影像與浮水印影像之間的失真程度，其計算方式如

下所示：

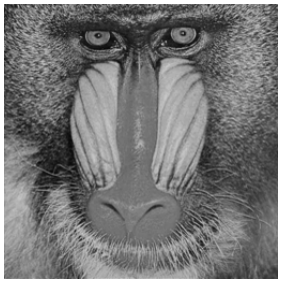
$$\mathrm{MSE}=\left(\frac{1}{M\times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\big(I(i,j)-R(i,j)\big)^2,$$

$$\mathrm{PSNR}=10\log_{10}\frac{255^2}{\mathrm{MSE}}\ \ (db).$$

其中 M 與 N 分別代表輸入影像的寬與高。一般而言 PSNR 大於或等於 30 db，以人類視覺能力是無法輕易發覺原始影像與浮水印影像之間的差異。

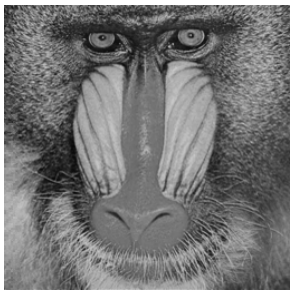
利用本研究所提出的方法，圖 9(c)與圖 9(d)為嵌入一組浮水印(編號為 100)後的影像，其 PSNR 值分別為 46.42db 與 46.85db，兩浮水印影像擁有相當高的視覺品質。

(a)Lena



(b)Baboon

(c)浮水印影像 Lena

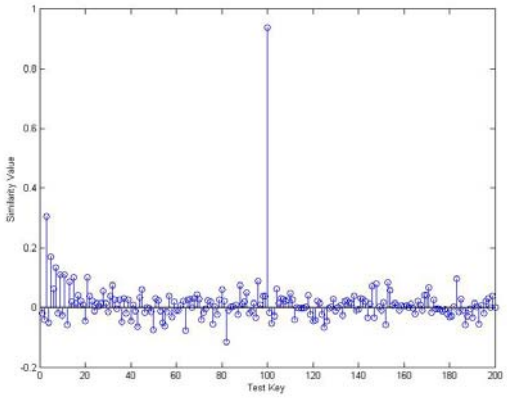


(b)浮水印影像 Baboon

圖 9 經由自我參照影像嵌入浮水印的影像(a),(b)原始影像;(c),(d)嵌入後的影像

在取出所嵌入的浮水印，本方法不需要原始影像的協助，僅需位置序列 $idl_z(i,j)$ 的指引便能將所嵌入的浮水印取出，再配合原始浮水印便可進行驗證。舉例說明，圖 10(a)為一浮水印影像，該影像內嵌入編號為 100 的浮水印資料，然後與其餘不同編號的 Lena 浮水印影像混在一起，一共是 200 張浮水印影像進行實驗偵測，作為判別本方法偵測結果的正確性。本研究利用位置序列 $idl_z(i,j)$ 與原始浮水印的協助，經偵測這 200 張浮水印影像後，正確偵測出含編號為 100 浮水印的影像，如圖 10(b)所示：

(a)

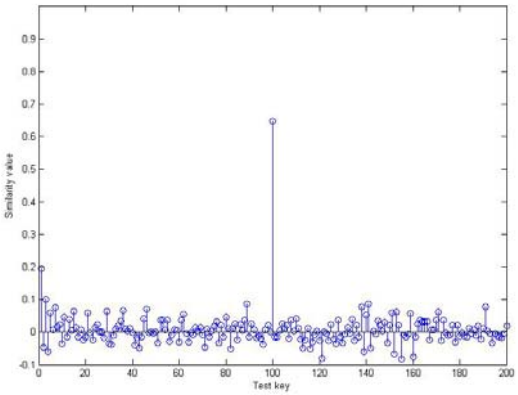


(b)

圖 10 (a)浮水印影像 PSNR 46.42 db (b) 相似值 $S=0.93$

為進一步測試本方法的強韌性，將圖 10(a)的浮水印影像進行進一步的攻擊後，再進行浮水印偵測，相關攻擊有：剪裁(50%)如圖 11(a)、模糊化如圖 12(a)、亮度調整如圖 13(a)、高斯雜訊如圖 14(a)與銳利化如圖 15(a)。

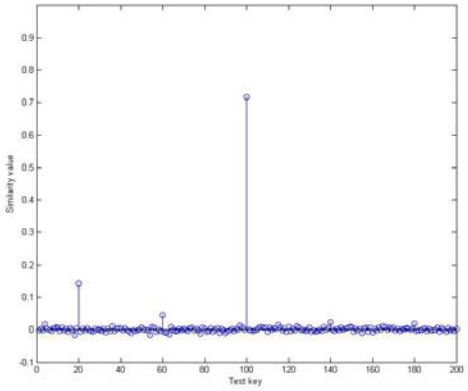
(a)



(b)

圖 11 (a)浮水印影像受剪裁(50%)攻擊 PSNR 7.08 db (b) 相似值 $S=0.64$

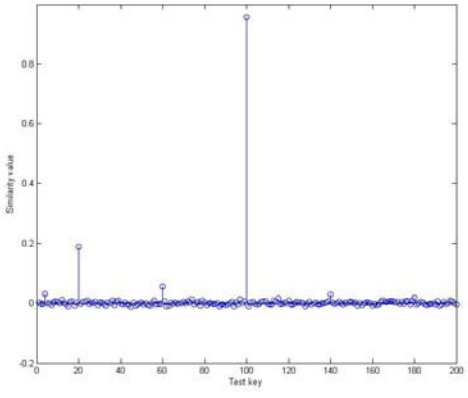
(a)



(b)

圖 12 (a)浮水印影像受模糊化攻擊 PSNR 45.51 db (b) 相似值 $S=0.93$

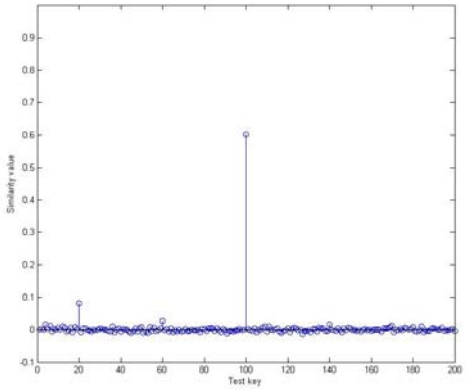
(a)



(b)

圖 13 (a)浮水印影像受亮度調整攻擊 PSNR 19.50 db (b) 相似值 $S=0.93$

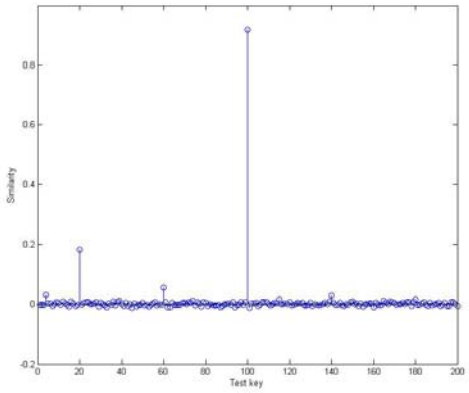
(a)高斯雜訊



(b)高斯雜訊

圖 14 (a)浮水印影像受高斯雜訊攻擊 PSNR 19.99 db (b) 相似值 $S=0.59$

(a)銳利化



(b)銳利化

圖 15 (a)浮水印影像受銳利化攻擊 PSNR 31.44 db (b) 相似值 $S=0.93$

在經過上述一系列的影像攻擊後，本方法仍能精確偵測出所嵌入的浮水印如圖 11(b)- 15(b)，足以證明本方法有足夠的強韌性。本方法也與 Liu 等學者所提出的方法進行相似度比較如表 1 所示：

表 1 相似度比對結果

影像攻擊	Liu 等學者的方法	本方法
未遭攻擊	0.97	0.93
剪裁 50%	0.70	0.64
模糊化	0.95	0.93
亮度調整	0.97	0.93
高斯雜訊	0.59	0.59
銳利化	0.91	0.93

在表 1 中呈現出本方法與 Liu 等學者的方法其強韌性的效能是差不多，然而本方法的安全性優於 Liu 等學者所提出的方法，因為浮水印嵌入的位置是根據 JND 的值來進行隨機挑選，不同於 Liu 等學者的方法是固定一個 $\pm\alpha$ 的像素值變動率，因此本方法能抵擋 Ting 等學者所提出的攻擊與竄改方法。

此外為證明本方法對於偵測浮水印的正確

率，我們利用兩個評估的指標進行實驗，分別為誤判率(False-positive errors)與漏判率(False-Negative Errors)。誤判率意指偵測影像實際上沒有嵌入浮水印，但偵測的結果卻誤判有浮水印；漏判率意味著偵測影像中藏匿有浮水印，但卻沒被偵測出來的錯誤。相關統計數據如表 2 所示：

表 2 浮水印偵測誤判率與漏判率的統計

影像攻擊	漏判率 (漏判/影像數)	誤判率 (誤判/影像數)
未遭攻擊	0/400	0/400
剪裁 50%	0/400	0/400
模糊化	0/400	0/400
亮度調整	0/400	0/400
高斯雜訊	0/400	0/400
銳利化	0/400	0/400
總　計	0/2400	0/2400

上表中所呈現的數據，證實本方法在浮水印偵測上有非常高的精確性。

五、結論

本文中我們呈了所設計的半盲目型浮水印植基自我參照影像的方法，主要的目的是改良 Liu 等學者的方法存有安全性漏洞的問題。我們的方法除能有效抵抗 Ting 等學者所提出的漏洞攻擊外，實驗結果證明我們所提出的方法能正確偵測出浮水印且誤判率與漏判率皆為 0。此外我們的方法能夠有效抵擋剪裁攻擊、模糊化攻擊、亮度調整攻擊、高斯雜訊攻擊與銳利化攻擊，然而浮水印的強韌性實驗證明能與 Liu 等學者的方法相似。

六、參考文獻

[1] S. Agreste and A. Guido, “A new approach to pre-processing digital image for wavelet-based watermark”, Journal of Computational and Applied Mathematics, vol.221, no.2, pp. 274-283, 2008.

[2] C. H. Chou and Y. C. Li, “A perceptually tuned subband image coder based on the measure of just-noticeable-distorion profile,” IEEE Trans. on Circuits and Systems for Video Technology, Vol. 5, No. 6, pp. 467-476, 1995.

[3] V. Fotopoulos and A. N. Skodras, “Improved watermark detection based on similarity diagrams”, Signal Processing: Image Communication, vol.17, no.4, pp. 337-345, 2002.

[4] C. T. Hsu and J. L. Wu “Hidden digital watermarks in images”, IEEE Transactions on Image Processing vol.8, no.1, pp. 58-68, 1999.

[5] J. L. Liu, D. C. Lou, M. C. Chang and H. K. Tso, “A robust watermarking scheme using self-reference image”, Computer Standards & Interfaces, vol.28, no.3, pp. 356-367, 2006.

[6] R. Ni, Q. Ruan and H. D. Cheng, “Secure semi-blind watermarking based on iteration mapping and image features”, Pattern Recognition, vol.38, no.3, pp. 357-368, 2005.

[7] G. C. W. Ting, B. M. Goi and S. H. Heng, “Attacks on a robust watermarking scheme based on self-reference image”, Computer Standards & Interfaces, vol.30, no.1-2, pp. 32-35, 2008.

[8] M. J. Tsai, K. Y. Yu and Y. Z. Chen, “Joint wavelet and spatial transformation for digital watermarking”, IEEE Transactions on Consumer Electronics, vol.46, no.1, pp. 237, 2000.